اقتصادسنجی

جلد اول

تک معادلات با فروض کلاسیک

جزء اول

دکتر مسعود درخشان

اقتصادسنجی

جلد اول

تک‌معادلات با فروض کلاسیک

جزء اول

دکتر مسعود درخشان

تهران
۱۳۷٤

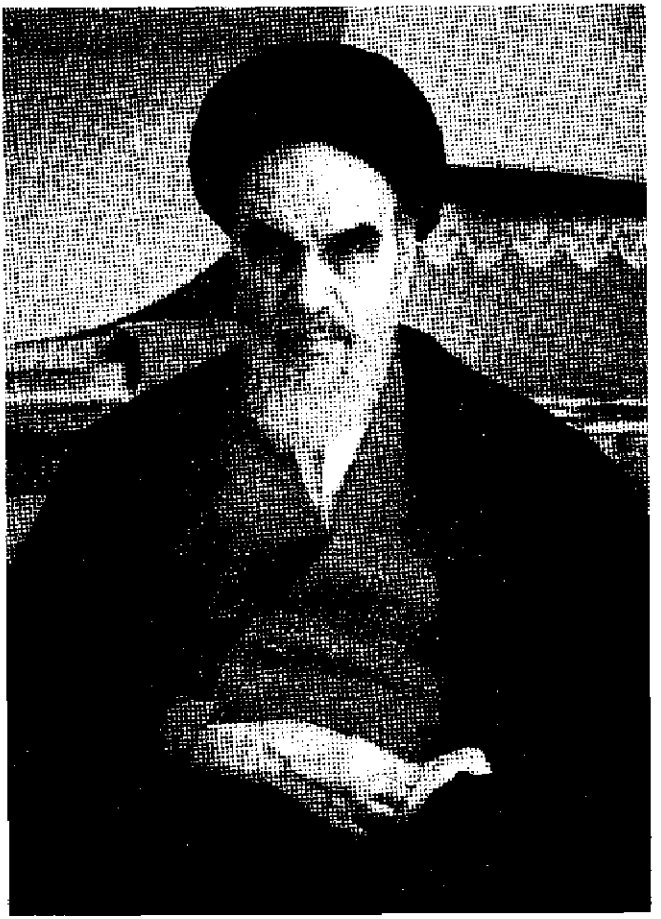
سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)

اقتصاد ۹

درخشان، مسعود
اقتصادسنجی / مسعود درخشان. — تهران: سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)، ۱۳۷٤.
ج.: جدول، نمودار. —(سازمان مطالعه و ... ؛ ۱٤۲ و ۱٤۳ : اقتصاد؛ ۹ و ۱۰)
عنوان پشت جلد به انگلیسی: Econometrics; vol. 1, part 1: Single Equations with Classical Assumptions.
جلد اول شامل ۲ جزء: تک معادلات با فروض کلاسیک.
واژه‌نامه.
کتابنامه.
۱. اقتصادسنجی. ۲. اقتصاد ریاضی. الف. فروست. ب. عنوان.
الف ٤۳۷ د ۱۵۱۹۵/۰ ۳۳۰ ۷ الف ٤ د ۱۳۹ HB

اقتصادسنجی: جلد اول؛ تک‌معادلات با فروض کلاسیک (جزء اول)
دکتر مسعود درخشان
سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)
چاپ اول: پاییز ۱۳۷٤
تعداد: ۵۰۰۰
حروفچینی، صفحه‌آرایی و لیتوگرافی: سمت
چاپ: مهر (قم)

بسم الله الرحمن الرحیم



بدون شک جهانخواران به همان میزان که از شهادت‌طلبی و سایر ارزشهای ایثارگرانه ملت ما واهمه دارند، از گرایش و رواج اقتصاد اسلام به طرف حمایت از پابرهنگان در هراسند و مسلم هر قدر کشور ما به طرف فقرزدایی و دفاع از محرومان حرکت کند، امید جهانخواران از ما منقطع و گرایش ملتهای جهان به اسلام زیادتر می‌شود.

صحیفهٔ نور؛ جلد ۲۰، صفحه ۱۲۹

سخن «سمت»

یکی از اهداف مهم انقلاب فرهنگی، ایجاد دگرگونی اساسی در دروس علوم انسانی دانشگاهها بوده است و این امر، مستلزم بازنگری منابع درسی موجود و تدوین منابع مبنایی و علمی معتبر و مستند با درنظرگرفتن دیدگاه اسلامی در مبانی و مسائل این علوم است.

ستاد انقلاب فرهنگی در این زمینه گامهایی برداشته بود، اما اهمیّت موضوع اقتضا می‌کرد که سازمانی مخصوص این کار تأسیس شود و شورای عالی انقلاب فرهنگی در تاریخ ۶۳/۱۲/۷ تأسیس «سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها» را که به اختصار «سمت» نامیده می‌شود، تصویب کرد.

بنابراین، هدف سازمان این است که با استمداد از عنایت خداوند و همّت و همکاری دانشمندان و استادان متعهد و دلسوز، به مطالعات و تحقیقات لازم بپردازد و در هرکدام از رشته‌های علوم انسانی به تألیف و ترجمه منابع درسی اصلی، فرعی و جنبی اقدام کند.

دشواری چنین کاری بر دانشمندان و صاحبنظران پوشیده نیست و به‌همین جهت مرحلهٔ کمال مطلوب آن، باید بتدریج و پس از انتقادها و یادآوریهای پیاپی ارباب‌نظر به‌دست آید و انتظار دارد که این بزرگواران از این همکاری دریغ نورزند.

کتاب حاضر که دربرگیرنده مفاهیم اقتصادسنجی است برای دانشجویان رشته‌های اقتصاد در مقطع کارشناسی و کارشناسی ارشد به عنوان منبع اصلی درس «اقتصادسنجی» به ارزش ۷ واحد تدوین شده است و برای رشته‌های مدیریت، آمار و مهندسی صنایع نیز قابل استفاده است. امید است علاوه بر جامعه دانشگاهی سایر علاقه‌مندان نیز بتوانند از آن استفاده کنند.

از استادان و صاحبنظران ارجمند تقاضا می‌شود با همکاری، راهنمایی و پیشنهادهای اصلاحی خود، این سازمان را در جهت اصلاح کتاب حاضر و تدوین دیگر آثار مورد نیاز جامعهٔ دانشگاهی جمهوری اسلامی ایران یاری دهند.

فهرست مطالب

تقدیم به:

شهدای دانشکده‌های اقتصاد

و آنان که

«محرومیت‌زدایی» را «عقیده و راه و رسم زندگی خود می‌دانند».

پیشگفتار

امیدوارم مطالب این کتاب برای دانشجویان عزیز و سایر پژوهشگران اقتصادی سودمند بوده، بتواند زمینهٔ مطالعات دقیقتر و جامعتری را در مباحث اقتصادی فراهم آورد. کتاب حاضر اولین جلد از یک مجموعه است که برای سطوح کارشناسی و بالاتر در رشتهٔ اقتصادسنجی تدوین شده است. در این پیشگفتار به چند نکته اشاره می‌کنیم.

۱. ضرورت

اولین سؤال این است که آیا ضرورتی برای انتشار این کتاب وجود دارد؟ این سؤال بویژه با توجه به کتابهای متعددی که در سالهای اخیر در رشته اقتصادسنجی به زبان فارسی به چاپ رسیده است[۱] اهمیت فراوان دارد. اکثر کتابهای موجود، ترجمه‌های بسیار خوب از منابع ارزشمندی است که اغلب از متون درسی دانشگاههای معتبر جهان به‌شمار می‌آیند. تدریس یکی از این کتابها که متناسب با دورهٔ کارشناسی است، و کتابی دیگر که هماهنگی بیشتری با دورهٔ کارشناسی ارشد دارد، می‌تواند برای دانشجویان عزیز و اساتید بزرگوار تضمینی باشد که معیار آموزشی، حداقل در سطح دانشگاههای معتبر جهان حفظ شده است. این تضمین، دارای اهمیت فراوان بوده و آثار مثبت زیادی دارد. آنچه در حال حاضر مورد نیاز نظام آموزش عالی دانشگاههای کشور است اعتماد به نفس و تکیه بر تواناییهای علمی داخل برای استفاده از امکانات خارجی است؛ زیرا ظرفیت علمی دانشجویان عزیز و استادان محترم قطعاً از موارد مشابه در

---

۱. برای ملاحظه فهرستی از این کتابها، به پیوست «ه» مراجعه کنید.

دانشگاههای خارج کمتر نیست.

روش فوق می‌تواند نقطه ضعفی نیز داشته باشد، که دقیقاً ناشی از ماهیت آن است. آن دسته از تألیفات جدید اقتصادسنجی به زبان خارجی، که از کتابهای درسی دانشگاههای معتبر جهان در دوره‌های کارشناسی و کارشناسی ارشد به شمار می‌آیند، اولاً از سطح ریاضیات دانشجویان و ثانیاً از کیفیت تقاضا در بازار اشتغال برای فارغ‌التحصیلان متأثر است. متأسفانه سطح ریاضیات دانشجویان اقتصاد در کشورهای خارج کاهش یافته و بنابراین اکثر نویسندگان خارجی برای اینکه بتوانند کتابهای «موفق‌تری» عرضه کنند از عمق و گستره ریاضیات در تألیفات خود کاسته‌اند. همچنین نیاز مبرم اقتصاد کشورهای صنعتی به فارغ‌التحصیلان آشنا به استفاده از برنامه‌های کامپیوتری در حل مسائل اقتصادسنجی، موجب شده است که تحلیلهای نظری و دقتهای مبنایی در مطالعات اقتصادسنجی به نحو قابل ملاحظه‌ای محدودتر شود. در مواردی نیز که از ریاضیات در سطوح خوب و عالی برای ارائه مباحث نظری اقتصادسنجی استفاده شده است مطالب بقدری تخصصی ارائه گردیده که فقط قابل استفاده برای متخصصان این رشته است و دانشجویان دوره‌های کارشناسی و کارشناسی ارشد اقتصاد، از آن چندان بهره‌ای نمی‌برند. با وجود این، امتیاز بعضی از کتابهای جدید، یعنی آشنا کردن دانشجویان با مثالهای کاربردی و برنامه‌های کامپیوتری نباید فراموش شود؛ هرچند که به قیمت کاهش دقتهای نظری صورت گرفته است.

کتاب حاضر که با توجه به سطح خوب و ظرفیت عالی ریاضیات دانشجویان عزیز کشور و علاقه آنها به دقت در مباحث نظری تدوین شده است می‌تواند به عنوان مکملی برای اکثر کتابهای موجود محسوب شود. معمولاً نمی‌توان یک کتاب خارجی در سطوح کارشناسی یا کارشناسی ارشد یافت که از هر جهت نسبت به سایر کتابهای موجود برتری داشته باشد؛ اما مباحث یا مثالهایی در هر کتاب وجود دارد که جامعتر و عمیق‌تر از موارد مشابه در سایر کتابها است. برای تدوین هر مبحث دراین کتاب، اکثر منابع اصلی و معتبر ملاحظه شده و سعی بر این بوده است که از نکات مثبت موجود در هر یک از آنها استفاده شود. بنابراین شاید بتوان گفت که اگر ترجمه یکی از منابع خوب اقتصادسنجی مبنا

قرار گیرد کتاب حاضر می‌تواند مکملی برای آن باشد. بر فرض درستی این نکته، ضرورت انتشار این کتاب اجمالاً به ثبوت می‌رسد.

۲. هدف

در تمام فصلهای این کتاب سعی بر این بوده است که مباحث اقتصادسنجی، در سطوح پایه و متوسطه، در کمال سادگی بیان شود به گونه‌ای که بتواند به قول دانشجویان عزیز یک «خودآموز اقتصادسنجی» باشد. مباحث پیشرفته اقتصادسنجی کتاب مستقلی است که ان‌شاءالله در آینده نزدیک منتشر خواهد شد.

بیان مباحث به زبانی بسیار ساده، هدفی است که آگاهانه در این کتاب تعقیب شده است. هرچند که ساده‌نویسی می‌تواند در ابعادی با اهداف پژوهشی و نوآوریهای علمی منافات داشته باشد اما به نظر می‌رسد که در مقطع فعلی، نمی‌توان آموزش و پژوهش برای نوآوری در اقتصادسنجی را به عنوان یک هدف اصلی در صدر برنامه‌های اقتصاد دانشگاهها قرار داد. به دانشجویان توصیه می‌شود که با توجه به شرایط و مقتضیات اقتصادی کشور، مطالعات اقتصادسنجی را تنها به عنوان «وسیله‌ای» بسیار خوب و قوی برای درک و تحلیل سیاستها و مدلهای اقتصادی در کشورهای در حال توسعه، و نیز برای شناخت مکانیسم تحولات و سیاستگذاریهای اقتصادی کشورهای صنعتی، با دقت کامل فراگیرند.

اگر هدف، مطلق کردن مطالعات اقتصادسنجی و رسیدن به مرحلهٔ نوآوری در این قلمرو می‌بود، آنگاه به طور قطع نحوهٔ نگارش اینجانب در این کتاب تغییر اساسی می‌کرد و از ساده‌نویسی و تبیین همه نکات و جوانب قضایا و حل مسائل، به شدت اجتناب می‌شد. بنابراین تأکید براین است که دانشجویان در حداقل زمان و با فراگیری حداکثر مطالب، هر چه سریعتر با ابعاد مختلف روشهای اقتصادسنجی آشنا شوند تا به مرحلهٔ اصلی فراگیری و تحقیقات اقتصادی خود برسند که همانا تجزیه و تحلیل استراتژیهای رشد و توسعه و بررسی ساختار سیاستهای مختلف اقتصادی در کشورهای در حال توسعه و آثار بلندمدت آن در اقتصاد جهانی است.

۳. الگوی تدوین

کیفیت طبقه‌بندی و ارائه مباحث تابعی از ضرورت و هدف تدوین این کتاب بوده است. در هر فصل، مطالب در دو سطح کارشناسی و کارشناسی ارشد ارائه می‌شود. آن قسمت از مباحثی که به دوره‌های بالاتر از کارشناسی مربوط می‌شود با علامت (*) مشخص شده است. همین امر زمینهٔ مناسبی را ایجاد می‌کند که دانشجویان دورهٔ کارشناسی که علاقه یا فرصت بیشتری برای مطالعه دارند بتوانند قلمرو مطالعات خود را در هر مبحث افزایش دهند.

مطالب این کتاب به گونه‌ای تنظیم شده است که سطح فعلی ریاضیات و آمار عمومی که در حال حاضر در رشته‌های اقتصاد و بازرگانی در دانشگاهها تدریس می‌شود به عنوان پیش‌نیاز کفایت می‌کند. با وجود این، پیوست «الف» نکات ریاضی که در انتهای کتاب قرار دارد باید دقیقاً مطالعه شود. برای آن دسته از دانشجویان دورهٔ کارشناسی اقتصاد که به هر دلیل ترجیح می‌دهند اقتصاد سنجی را بدون استفاده از جبر ماتریسی فراگیرند مشکل خاصی وجود نخواهد داشت، زیرا تنظیم مطالب به نحوی است که می‌توان مباحث مطرح شده به زبان ماتریسی را از سایر عناوین جدا کرد به گونه‌ای که بدون آشنایی با جبر ماتریسی بتوان مفاهیم و روشهای اصلی اقتصادسنجی را فراگرفت. با وجود این، تحلیل مباحث اقتصادسنجی به کمک جبر ماتریسی، برای دانشجویان دوره‌های کارشناسی ارشد ضروری است.

با توجه به اینکه کتاب حاضر در واقع مکملی برای سایر کتابهاست، نه تنها سعی شده است که مطالب در چهارچوب مثالهای ساده مطرح شود، بلکه جمعاً ۸۳ مثال جامع در خلال ده فصل، نقش عمده‌ای در تبیین موضوعات داشته باشد. همچنین جمعاً ۱۱۳ مسأله در پایان فصلها آمده که همگی با روشنی و سادگی تمام حل شده است. از نظر کثرت مثال و مسأله و نبودن ابهام در راه‌حلها و تبیین موضوعات تا سطوح پیشرفته در هر مبحث، شاید کتاب حاضر را بتوان به عنوان اولین کوشش در این زمینه معرفی کرد. با مطالعه این کتاب، دانشجویان عزیز می‌توانند از طریق تسلط بیشتر در درک و حل مسائل، تمرینات موجود در کتابهای دیگر را که استادان بزرگوار توصیه می‌نمایند

بهتر حل و بحث کنند.

چند نکته تاریخی را در بخش پیوستها مطرح کرده‌ایم. مباحث تاریخی همواره شیرین و بسیار آموزنده است. در خلال بررسیهای تاریخی در زمینهٔ روش حداقل مربعات، موفق شدم که در کتابخانهٔ قدیمی دانشگاه آکسفورد، اولین کتابنامه در روش حداقل مربعات را به دست آورم که در سال ۱۸۷۷ یعنی همزمان با دوران سلطنت ناصرالدین شاه قاجار منتشر شده است. در این کتابنامه، جمعاً ۴۰۸ مقاله و کتاب گردآوری شده که دربارهٔ این روش از سال ۱۷۲۲ (مصادف با شکست سلطان حسین صفوی از محمود افغانی) تا سال ۱۸۷۶ (سه سال بعد از اولین سفر ناصرالدین شاه به فرنگستان) به چاپ رسیده است. همچنین موفق شدم اولین کتاب معروف درسی در روش حداقل مربعات را که در سال ۱۸۷۷ به چاپ رسیده و در سال ۱۸۸۵ با تجدید نظر کلی مجدداً منتشر شده است ملاحظه نموده و صفحاتی از آن را به همراه دو جدول آماری، در کتاب حاضر منعکس کنم تا بدین ترتیب دانشجویان عزیز با سطح مباحث در آن زمان بیشتر آشنا شوند.

در پیوست نکات تاریخی، سعی کرده‌ام که قدر و منزلت «گاست»[1] آماردان معروف قرن بیستم و صاحب توزیع t را برای دانشجویان رشته اقتصاد، بیشتر روشن کنم. متأسفانه تنها نکته‌ای که در بعضی از تألیفات آماری دربارهٔ این آماردان بزرگ مطرح می‌شود اشتغال نامبرده در یک کارخانه آبجوسازی است. همین امر موجب شده است که این شخصیت بزرگ علمی، که توانست برای اولین بار توزیع و آزمون t را در سال ۱۹۰۸، یعنی همزمان با انقلاب مشروطیت و حکومت محمدعلی شاه قاجار، کشف کند مورد کم‌مهری و بی‌توجهی در بعضی از محافل علمی ایران قرار گیرد.

تدوین این کتاب، مبتنی بر متون اصلی اقتصادسنجی دانشگاههای معروف دنیاست. از تمام کتابها و مقالاتی که در قسمت منابع و مآخذ ذکر گردیده مستقیماً بهره‌برداری شده است. در مواردی نیز اثبات بعضی از قضایا به روشهای ساده‌تر، ارائه

1. William Sealy Gosset

نظریات یا اشکالاتی در تجزیه و تحلیل مباحث، و یا طراحی و حل بعضی از مثالها و مسائل آخر هر فصل از اینجانب است. با وجود این، منابع زیر بیشترین سهم را در تدوین این کتاب داشته است: جانستون[1]، کمنتا[2]، مادالا[3]، استیوارت و والیس[4]، کندی[5]، بیکن[6]، هاروی[7]، و گرین[8].

یافتن واژه‌های فارسی برای اصطلاحات علمی یکی از سخت‌ترین مراحل تدوین کتاب و مقاله علمی به زبان فارسی است. متأسفانه کوشش جدی و مستمری از طرف جامعه متخصصان زبان فارسی در حل این مشکل ملاحظه نمی‌شود. رشد سریع علوم و تکنولوژی و تداوم ظهور اصطلاحات علمی جدید، و همچنین علاقه فراوان دانشجویان و استادان ارجمند به ترجمهٔ آخرین دستاوردهای علمی جهان، زبان فارسی را با کمبود شدید واژه‌ها و اصطلاحات علمی مواجه کرده است. این امر دلالت بر فوریت اتخاذ تصمیماتی دارد که برای حفظ زبان فارسی و رشد و توسعه آن ضروری است.

در مقطع کنونی، که رابطهٔ نزدیکی بین جامعه اقتصاددانان کشور و متخصصان ادبیات فارسی وجود ندارد، شاید روش بهینه، همفکری و تبادل نظر استادان و پژوهشگران اقتصاد با یکدیگر و کوشش در استفاده از ترجمه‌های یکسان برای اصطلاحات اقتصادی است. به همین دلیل سعی شده است که در این کتاب، تا حد امکان، از واژه‌هایی استفاده شود که استادان محترم در تألیفات و ترجمه‌های خود برگزیده‌اند. بنابراین در کتاب حاضر بسیاری از اصطلاحاتی را که در خلال چندین سال تدریس اقتصادسنجی از آنها استفاده کرده‌ام کنار گذاشته و به جای آنها، اصطلاحات موجود در کتابهای منتشر شده را قرار داده‌ام. با وجود این، نظر به ناهماهنگیهای رایج در ترجمه‌ها، طبعاً فقط به فصل مشترکات اکتفا شده است. در مواردی نیز انتقال مفاهیم در قالب جمله‌ها، به ساختن و پرداختن کلمات یا اصطلاحات جدید ترجیح داده شده تا بدین وسیله زمینه برای طراحی واژه‌های زیبا و صحیح فارسی توسط اهل فن از بین نرود.

---

1. J. Johnston (1984)
2. Jan Kementa (1986)
3. G. S. Maddala (1992)
4. M. B. Stewart and K. F. Wallis (1981)
5. P. Kennedy (1992)
6. R. Bacon (1988)
7. A. C. Harvey (1990)
8. W. H. Greene (1993)

٤. تاریخچهٔ تدوین

ضرورت تدوین یک کتاب اقتصادسنجی به زبان فارسی، در سالهای ۱۳٤٦ ـ ۱۳۵۰، که اینجانب دانشجوی دانشکدهٔ اقتصاد دانشگاه تهران بودم، کاملاً احساس می‌شد، زیرا هیچ کتاب درسی در این زمینه وجود نداشت. تدوین مقدمات کتاب حاضر عملاً در سالهای ۱۳۵۱ و ۱۳۵۲ و همزمان با تحصیلات دورهٔ فوق لیسانس نگارنده در مدرسه اقتصاد لندن (LSE) آغاز شد. توجه خاص آن مدرسه به آموزش و پژوهش در اقتصادسنجی، انگیزه بسیار مناسبی در تدوین این کتاب بود. در دوره تحصیلات دکتری و تهیهٔ رساله در آکسفورد به نظر نگارنده رسید که تألیف کتابی در زمینه کاربرد روشهای سیستم و کنترل بهینه در اقتصادسنجی سریعتر به نتیجه خواهد رسید. بدین‌ترتیب اتمام کتاب اقتصادسنجی به تعویق افتاد.

حضور در ایران در سال ۱۳۵۵ به مدت یک ترم تحصیلی برای تکمیل یک مدل اقتصادسنجی کلان، فرصت مناسبی را فراهم آورد تا جزوه‌های تدوین شده در اقتصادسنجی و در کاربرد نظریه سیستم و کنترل در مدلهای اقتصادی را در دانشکده اقتصاد دانشگاه تهران تدریس کنم. علاقه و استقبال بسیار زیاد دانشجویان موجب شد که بعد از بازگشت به آکسفورد، کوشش برای تکمیل و انتشار هر چه سریعتر کتاب اقتصادسنجی آغاز شود. خوشبختانه وقوع انقلاب اسلامی در سال ۱۳۵۷ باعث شد که از «قیل و قال» اقتصادسنجی و نظریه‌های کنترل رها شده و به «شور و حال» اقتصاد اسلامی، تاریخ تحولات اندیشه‌ها و نظامهای اقتصادی، که سالها مورد علاقه‌ام بود، بازگردم.

بعد از انقلاب فرهنگی و بازگشایی دانشگاهها، افتخار همکاری در تدریس اقتصادسنجی در دوره‌های کارشناسی، کارشناسی ارشد، و دکتری دانشگاههای تهران، تربیت مدرس، و امام صادق(ع) بر عهده اینجانب قرار گرفت، و در نتیجه جزوه‌های درسی سابق به مراتب کاملتر شد. امّا در خلال این دوره، همکاران دانشمند و متخصصان بزرگوار اقتصادسنجی در دانشگاههای تهران، شهید بهشتی، اصفهان، مازندران، شیراز، و بعضی دیگر از مراکز مطالعات اقتصادی، کتابهای بسیار خوب و معتبری را در این زمینه

ترجمه یا تألیف کردند، در نتیجه دیگر هیچ دلیلی بر ضرورت انتشار آن جزوه‌ها، به عنوان کتابی مستقل وجود نداشت. به هر حال با توجه به شرحی که در بند اول آمد و با اصرار دانشجویان عزیزم، سرانجام انتشار این کتاب، بعد از تجدیدنظر اساسی و بهره‌برداری از جدیدترین متون، به تحقق پیوست، امیدوارم که مفید باشد.

۵. تشکر و قدردانی

از دانشجویان عزیزم که در خلال سالها تدریس اقتصادسنجی همواره در تدوین این کتاب مشوقم بوده‌اند صمیمانه تشکر می‌کنم. خاطرات شیرین کلاسهای درس، صفای قلب، اراده راسخ، و آمادگی کامل این جوانان عزیز برای حرکت در جهت رشد و توسعه اقتصادی کشور، انگیزه اصلی در تداوم خدمات ناچیز دانشگاهی نگارنده بوده است. از درگاه خداوند متعال موفقیت هرچه بیشتر برای این جوانان عزیز مسئلت می‌نمایم؛ جوانانی که با داشتن حداقل امکانات مادی می‌کوشند تا با اعتلای سطح علمی خود، دانشگاههای کشور را به مرحله رقابت با بهترین دانشگاههای معتبر جهان برسانند.

لازم است از حمایتهای جناب آقای دکتر احمدی، ریاست محترم سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)، در انتشار کتاب حاضر، صمیمانه سپاسگزاری کنم. همه می‌دانیم که سازماندهی و مدیریت نهادهای علمی و پژوهشی در سالهای اخیر، با توجه به کمبود منابع انسانی و مالی، تا چه حد طاقت‌فرسا است. تأسیس و توسعه نهاد علمی «سمت» مدیون فداکاریهای ایشان است؛ و اینجانب به سهم خود از این همه تلاش و کوشش صمیمانه تشکر می‌کنم. از جناب آقای عبدالرضا حسنی که با دقت، حوصله و تسلط فراوان متن کتاب را ویراستاری نمودند و از سرکار خانم دهقان‌نیری، دبیر محترم گروه اقتصاد، صمیمانه سپاسگزاری می‌کنم. از زحمات اداره تولید «سمت» که مسئولیت حروف‌چینی، نمونه‌خوانی، لیتوگرافی و چاپ کتاب را بر عهده داشته است قدردانی می‌شود.

مسعود درخشان

مرداد ۱۳۷۳

فصل اوّل

مفاهیم و تخمین مدل رگرسیون خطی ساده

۱ـ۱ مقدمه

اقتصادسنجی[1] معمولاً با بحث مدلهای رگرسیون خطی ساده آغاز می‌شود. در اینجا تعریف واژه‌های اقتصادسنجی، مدل رگرسیون، و مدل رگرسیون خطی ساده مفید است. بدیهی است این تعاریف در حد اجمال بوده و صرفاً برای ایجاد زمینهٔ مناسب برای تبیین مفاهیم کلی اقتصادسنجی به کار می‌رود.

۱. اقتصادسنجی

در کتابهای اقتصادسنجی معمولاً فصل اول به تعریفی مبسوط از اقتصادسنجی و ماهیت آن اختصاص یافته است. با توجه به تخصصی بودن این تعاریف، به نظر می‌رسد در این کتاب بهتر است تعریف و ماهیت اقتصادسنجی بعد از مطالعه دقیق روشهای اقتصادسنجی (فصل آخر جلد دوم) ارائه شود.[2] در این قسمت به تعریفی اجمالی از اقتصادسنجی اکتفا می‌کنیم.

اقتصادسنجی در لغت به معنای علم اندازه‌گیری در اقتصاد است. قلمرو این تعریف بسیار وسیع است؛ برای مثال، موضوع حسابداری ملی نیز به طور عمده اندازه‌گیری کمیّات اقتصادی مانند تولید ناخالص ملی، شاخص قیمتها، یا سرمایه‌گذاری ملی است. تعریف دقیقتر این است که «اقتصادسنجی علم تحلیلهای آماری از مدلهای

---

1. Econometrics

۲. برای مطالعه شکل‌گیری و پیدایش اقتصادسنجی، و بررسی توسعه‌های اولیه و پیشرفتهای جدید، به مقاله محمدهاشم پسران در *The New Palgrave.A Dictionary of Economics*, 1987 مراجعه کنید.

اقتصادی است»). منظور از مدلهای اقتصادی، صورت منظم و ریاضی توابع یا روابط [۱] اقتصادی است که در ذیل تعریف می‌شود.

۲. مدل

هرگاه متغیرهای اقتصادی مانند تولید، مصرف، سرمایه‌گذاری و قیمت در روابط معینی تعریف شوند زمینهٔ مناسبی برای شکل‌گیری یک مدل اقتصادی فراهم می‌شود. در حالی که نظریه‌های اقتصادی سعی دارند تا این روابط و تحولات آنها را تفسیر کنند، اقتصادسنجی می‌خواهد این روابط را به زبان آماری بیان کند. یک مدل اقتصادی در واقع بیان دیگری از رابطه یا روابط اقتصادی است که معمولاً به زبان ریاضی بیان می‌شود.

یک متغیر اقتصادی مانند مصرف را در نظر می‌گیریم؛ آنچه در خارج وجود دارد نوسانات یا تغییرات مصرف است. مشاهدهٔ تغییر، انگیزهٔ تفسیر تغییر را به وجود می‌آورد که خود زمینه‌ساز مطالعات و تحلیلهای اقتصادی است. اولین قدم در تفسیر تغییر، ساختن مدل تغییر است؛ یعنی مدلی که در آن، تغییر می‌تواند تفسیر شود. چگونگی ساختن مدلها به روش مدل‌سازی و مبانی نظری مدل بر می‌گردد که از موضوع این کتاب خارج است. یکی از ساده‌ترین و در عین حال رایجترین روشهای مدل‌سازی، طراحی یک یا چند رابطهٔ ریاضی است که بتواند چند متغیر را با یکدیگر مرتبط سازد. برای مثال، در تفسیر تغییرات مصرف، یک متغیر دیگر مانند درآمد را در نظر گرفته، فرض می‌کنیم که مصرف تابعی از درآمد است. اگر مصرف و درآمد در زمان را به ترتیب با C_t و Y_t نشان دهیم خواهیم داشت

$$C_t = f(Y_t) . \qquad (۱-۱)$$

به عبارت فوق که در واقع یک رابطهٔ ریاضی بین مصرف و درآمد است، معادله یا مدل مصرف گفته می‌شود. برای تفسیر تغییرات مصرف می‌توان علاوه بر Y_t متغیرهای دیگری را نیز در نظر گرفت. برای مثال، معادلهٔ

---

۱. در اینجا تفاوت مختصر بین تابع و رابطه را در نظر نگرفته‌ایم.

$$C_t = f(Y_t , C_{t-1}) , \qquad (۱-۲)$$

مدل دیگری برای مصرف است که در آن C_{t-1} مقدار مصرف در سال گذشته است.

در گذشته واژهٔ مدل فقط برای سیستم معادله‌ها به معنای چند معادلهٔ مرتبط با یکدیگر به کار می‌رفت. برای مثال، اگر فرض کنیم مصرف تابعی از درآمد است در حالی که درآمد خود به صورت مجموع مصرف و سرمایه گذاری تعریف شود، آنگاه خواهیم داشت

$$\begin{aligned} C_t &= f(Y_t) , \\ Y_t &= C_t + I_t , \end{aligned} \qquad (۱-۳)$$

که در آن I_t سرمایه گذاری است. به معادله‌های ۱-۳ ، اصطلاحاً یک «مدل» می‌گوییم. با وجود این، واژهٔ مدل را می‌توان در حال حاضر علاوه بر سیستم معادله‌ها، برای تک‌معادله نیز به کار برد.

تک‌معادله‌های ۱-۱ و ۱-۲ و سیستم معادله‌های ۱-۳ بسیار عمومی است؛ پس نمی‌تواند کاربرد چندانی داشته باشد؛ بنابراین برای اولین قدم، باید شکل ریاضی این معادله‌ها معیّن شود. فرض کنیم که این معادله‌ها همگی خطی هستند، آنگاه تک‌معادله‌های ۱-۱ و ۱-۲ و سیستم معادله‌های ۱-۳ را می‌توان به ترتیب چنین نوشت،

$$C_t = \alpha + \beta Y_t , \qquad (۱-۴)$$

$$C_t = \alpha + \beta Y_t + \gamma C_{t-1} , \qquad (۱-۵)$$

$$\begin{aligned} C_t &= \alpha + \beta Y_t , \\ Y_t &= C_t + I_t , \end{aligned} \qquad (۱-۶)$$

که در آن α و β را اصطلاحاً «پارامترهای»[1] مدل می‌گویند؛ بنابراین این مدلها هم

1. Parameters

برحسب پارامترها و هم برحسب متغیرها، خطی است.

فرض خطی‌بودن این مدلها در سهولت ریاضیات و ابعاد محاسباتی آن بسیار مفید است؛ سؤالی که مطرح می‌شود این است که چگونه می‌توان فرض خطی بودن مدل مصرف را توجیه کرد؟ در پاسخ باید گفت که یکی از اهداف اصلی مدل مصرف این است که بتواند تغییرات مصرف را که در عینیت مشاهده شده است به نحو رضایت‌بخشی تفسیر کند. اگر با فرض خطی بودن مدل مصرف بتوان به این هدف رسید، این فرض بر هر فرض دیگری ترجیح دارد؛ زیرا از هر فرض دیگری ساده‌تر است. البته معیار و نحوهٔ محاسبهٔ درجهٔ توانایی مدل در تفسیر عینیت، مسأله‌ای است که در قسمت ٤ـ١ بررسی خواهد شد. در پایان مفید است به این نکته نیز توجه شود که به مدلهای ٤ـ١ و ٥ـ١ و ٦ـ١ در اصطلاح «مدلهای ریاضی» مصرف می‌گویند؛ زیرا در این مدلها فقط به ابعاد ریاضی رابطهٔ مصرف با درآمد توجه شده است.

۳. رگرسیون

رگرسیون[1] در لغت به معنای «بازگشت به مراحل قبلی در یک مسیر تحول و توسعه» است. نظریه رگرسیون تقریباً در قلمرو همین معنی در روانشناسی و جامعه‌شناسی به کار می‌رود. در آمار، اصطلاح رگرسیون را اولین بار گالتن[2] در مطالعات جمعیت‌شناختی به کار بُرد. گالتن در نمونه‌های مختلف آماری سعی کرد رابطه‌ای بین قد فرزندان با قد والدین برقرار کند. با اینکه معمولاً والدین قد بلند فرزندان قد بلندی دارند و برعکس، گالتن به این نتیجه رسید که در جامعه گرایشی وجود دارد که قد فرزندان به سمت متوسط قد جامعه میل می‌کند.[3] این میل کردن را گالتن «رگرسیون» نامید.

در متون اقتصادسنجی، رگرسیون هیچگاه در این معنی به کار نمی‌رود. در اقتصادسنجی بیشتر اصطلاحات «تحلیل رگرسیون» و «مدل رگرسیون» رایج است.

---

1. Regression 2. Sir Francis Galton (1822 - 1911).

۳. گالتن این حقیقت را به صورت regression towards mediocrity بیان می‌کند. برای توضیح بیشتر به مقاله گالتن در سال ۱۸۸٦ مراجعه شود.

تحلیل رگرسیون در واقع بدنهٔ اصلی مطالعات اقتصادسنجی را تشکیل می‌دهد و به طور کلی دربارهٔ مدلهای رگرسیون و نحوهٔ تخمین آنها بحث می‌کند؛ بنابراین باید به تعریف مدلهای رگرسیون بپردازیم تا بدین ترتیب مفهوم رگرسیون در اقتصادسنجی روشن شود.

٤. مدل رگرسیون خطی ساده

فرض می‌کنیم در یک جامعهٔ مفروض تغییرات مصرف را در ٥ سال مشاهده کرده و داده‌های $C_t = ۲, ۳, ٥, ٦, ۹$ را برحسب میلیارد تومان به دست آورده‌ایم. می‌خواهیم چگونگی این تغییرات را تفسیر کنیم. برای این منظور باید متغیر یا متغیرهایی را در نظر بگیریم که بتواند این تغییرات را توضیح دهد. فرض کنید مدل ٤ـ۱ مدل مطلوب مصرف است؛ یعنی این مدل می‌تواند تغییرات مصرف را به صورت رضایت‌بخشی توضیح دهد، مگر آنکه خلاف آن ثابت شود. این مدل خطی مصرف را، که در واقع یک مدل ریاضی است، یک بار دیگر می‌نویسیم:

$$C_t = \alpha + \beta Y_t \cdot$$

بدیهی است باید آمار و مشاهدات Y_t را نیز داشته باشیم؛ زیرا بر اساس مدل مفروض، تغییرات مصرف بر اساس تغییرات درآمد بیان می‌شود. فرض کنید تغییرات Y_t را نیز برای همان ٥ سال به ترتیب $Y_t = ۳, ٥, ۹, ۱۰, ۱۳$ مشاهده کرده‌ایم. مفید است در اینجا به این نکته اشاره کنیم که داده‌ها و مشاهدات متغیرهای موجود در یک مدل معمولاً در سه نوع مختلف می‌تواند وجود داشته باشد: داده‌های سری زمانی[1]، داده‌های مقطع زمانی[2] و داده‌های تلفیقی یا «پانل»[3].

داده‌های سری زمانی، مقادیر یک متغیر را در نقاط متوالی در زمان، اندازه گیری می‌کند. این توالی می‌تواند سالانه، فصلی، ماهانه، هفتگی یا حتی به صورت پیوسته باشد. مثالی که تا به حال موضوع بحث ما بوده است، یک سری زمانی است که

1. Time Series Data 2. Cross - Section Data 3. Panel Data

مشاهدات را به صورت سالانه ارائه می‌کند. معمولاً از اندیس t برای سری زمانی استفاده می‌کنند. برای مثال، اگر متغیر مورد نظر، X ناپیوسته باشد سری زمانی آن را با X_t و در حالتی که X پیوسته باشد با X(t) نشان می‌دهند؛ مثلاً داده مصرف در زمان، یعنی C_t، یک متغیر ناپیوسته است، چون می‌توان فقط در فواصل معینی از زمان مقادیر آن را ملاحظه کرد. متغیرهای پیوسته در اقتصاد، کاربرد چندانی ندارد.

داده‌های مقطع زمانی، مقادیر یک متغیر را در زمان معین و روی واحدهای متعدد اندازه‌گیری می‌کند. این واحدها می‌تواند افراد، خانوارها، واحدهای تولیدی، صنایع، نواحی مختلف و حتی کشورهای مختلف باشد. می‌توان داده‌های درآمد و مصرف خانوارهای مختلف را در سال معینی جمع‌آوری کرد. معمولاً از اندیس i برای داده‌های مقطعی استفاده می‌کنند؛ بنابراین C_i و Y_i نشان‌دهندهٔ مصرف و درآمد خانوارهای مختلف در زمان معین است.

داده‌های تلفیقی در واقع بیان‌کنندهٔ داده‌های مقطعی در طی زمان است. بنابراین حجم مشاهدات در داده‌های تلفیقی نسبتاً زیاد است. در سالهای اخیر، کاربرد داده‌های تلفیقی در اقتصادسنجی افزایش بسیاری یافته است. معمولاً داده‌های تلفیقی و داده‌های مقطعی در «اقتصادسنجی خرد[1]» به کار می‌رود که موضوع آن بررسی روشهای اقتصادسنجی در اقتصاد خرد است، در حالی که داده‌های سری زمانی به طور کلی موضوع کار «اقتصادسنجی کلان»[2] است که روشهای اقتصادسنجی را در سطح اقتصاد کلان مطالعه می‌کند.

حال که با انواع مختلف آمار و مشاهدات در اقتصادسنجی آشنا شدیم یادآوری می‌کنیم که به مدل ۱ـ۴ یک مدل ریاضی خطی ساده گفته می‌شود. واژه‌های ریاضی، خطی و ساده را به ترتیب تعریف می‌کنیم. می‌دانیم مدل فوق یک مدل ریاضی است؛ زیرا فقط رابطهٔ ریاضی بین مصرف و درآمد را منعکس کرده است. همچنین یک مدل خطی است، زیرا پارامترها و متغیرهای این مدل در یک رابطهٔ خطی تعریف شده است؛

1. Microeconometrics 2. Macroeconometrics

و سرانجام یک مدل ساده است؛ زیرا در سمت راست فقط یک متغیر وجود دارد. اگر بیش از یک متغیر در سمت راست باشد، یک مدل خطی «چند متغیره» خواهیم داشت.

برای سهولت ارائه مباحث، مشاهدات مصرف و درآمد را در جدول ۱ـ۱ تنظیم می‌کنیم.

جدول ۱ـ۱

t	مصرف C_t	درآمد Y_t
۱	۲	۳
۲	۳	۵
۳	۵	۹
٤	٦	۱۰
۵	۹	۱۳

سؤال این است که آیا تغییرات درآمد، در چهارچوب شکل ریاضی مدل ٤ـ۱ می‌تواند تغییرات مصرف را به نحو رضایت‌بخشی توضیح دهد؟ پاسخ منفی است:

اولاً، علاوه بر درآمد، به منزلهٔ عامل اصلی تغییرات مصرف، قطعاً عوامل دیگری نیز وجود دارد که می‌تواند تغییرات مصرف را توضیح دهد. بسیاری از این عوامل، مانند مصرف سال قبل، را می‌توان شناخت که معمولاً آنها را وارد مدل می‌کنیم؛ و در نتیجه مدل رگرسیون ساده به رگرسیون چند متغیره تبدیل می‌شود. اما عوامل دیگری مانند چگونگی انتظارات مصرف‌کننده نسبت به تغییر در پارامترهای مختلف اقتصادی و درجهٔ عدم اطمینان نسبت به تحولات آیندهٔ اقتصادی وجود دارد که بیان کمی آنها معمولاً بسیار مشکل است.

ثانیاً، هیچ دلیلی وجود ندارد که رابطهٔ بین مصرف و متغیر یا متغیرهای سمت راست، خطی باشد.

ثالثاً، مصرف، خصوصیتی از رفتار انسانی است و می‌دانیم در رفتار انسان همواره عناصر تصادفی غیرقابل پیش‌بینی وجود دارد که اساساً نمی‌توان آنها را با مدلهای معین ریاضی از نوع مدل ٤ـ۱ بیان کرد.

البته می‌توان دلایل دیگری مبنی بر عدم قابلیت مدل ٤ـ۱ در بیان دقیق تغییرات مصرف بیان کرد، مانند خطای اندازه‌گیری[1] متغیرمصرف. این دلایل برای قبول این

1. Measurement Error

حقیقت که مدل ۱ـ۴ دقیق نیست و خطا دارد کافی است. به این خطا اصطلاحاً «جمله اختلال[1]» می‌گویند؛ زیرا تعادل ریاضی مدل مفروض را مختل می‌کند. اگر تأثیر تمام عوامل فوق را بر تغییرات مصرف با U_t نشان‌دهیم، آنگاه مدل ۱ـ۴ را می‌توان چنین نوشت،

$$C_t = \alpha + \beta Y_t + U_t \cdot \qquad (۱ـ۷)$$

که در آن U_t جملهٔ اختلال بوده و منعکس‌کنندهٔ خطای ما در بیان مدل مصرف است. بدیهی است U_t یک متغیر تصادفی است؛ زیرا عناصری که مقادیر آن را تعیین می‌کند، تصادفی است. به معادلهٔ ۱ـ۷ مدل رگرسیون مصرف می‌گویند؛ بنابراین تفاوت کلی مدلهای ریاضی و مدلهای رگرسیون در جملهٔ اختلال است. هرگاه به مدلهای ریاضی یک جملهٔ اختلال ـ که یقیناً تصادفی است ـ اضافه کنیم، مدل ریاضی به مدل رگرسیون تبدیل خواهد شد. با توجه به تعاریف قبلی می‌توان گفت که معادلهٔ ۱ـ۷ یک «مدل رگرسیون خطی ساده»[2] است. همچنین به معادلهٔ ۱ـ۷ یک مدل ساختاری[3] مصرف نیز می‌گویند؛ زیرا ساختار مصرف را مبتنی بر یک تئوری مصرف بیان می‌کند. به همین دلیل به پارامترهای α و β پارامترهای ساختاری[4] گفته می‌شود.

۲ـ۱ خصوصیات آماری مدل رگرسیون خطی ساده

مدل رگرسیون خطی زیر را در نظر می‌گیریم،

$$Y_t = \alpha + \beta X_t + U_t \cdot \qquad (۱ـ۸)$$

در این مدل می‌خواهیم تغییرات Y_t را به کمک تغییرات X_t توضیح دهیم. به همین دلیل به X_t متغیر توضیحی[5] و به Y_t متغیر توضیح داده شده[6] می‌گویند. اصطلاحات دیگری نیز برای X_t و Y_t به کار می‌رود؛ برای مثال، به X_t متغیر مستقل[7] و به Y_t متغیر وابسته[8]

---

1. Disturbance Term
2. Simple Linear Regression Model
3. Structural Model
4. Structural Parameters
5. Explanatory Variable
6. Explained Variable
7. Independent Variable
8. Dependent Variable

می‌گویند؛ زیرا مقادیری که X_t می‌گیرد به طور مستقل تعیین می‌شود؛ یعنی تابعی از مدل ۱ـ۸ نیست. به عبارت دیگر مقادیر X_t داده شده است، در حالی که مقادیر Y_t دقیقاً تابعی از مقادیر X_t و کیفیت مدل رگرسیون مفروض است و به همین دلیل Y_t یک متغیر تابع یا وابسته محسوب می‌شود. اصطلاحات متغیر برون‌زا[1] و متغیر درون‌زا[2] به ترتیب برای X_t و Y_t نیز به کار می‌رود؛ زیرا فرض بر این است که مقادیر X_t خارج از مدل رگرسیون مفروض تعیین شده و در نتیجه برون‌زا است، در حالی که مقادیر Y_t در داخل و براساس قانونمندی مدل رگرسیون تعیین می‌شود و به همین دلیل درون‌زا خواهد بود. همچنین متغیرهای X_t و Y_t را می‌توان متغیرهای «غیرتصادفی»[3] و «تصادفی»[4] نامید؛ زیرا مقادیری که X_t می‌گیرد به طور مستقل تعیین شده و تابعی از جملهٔ اختلال مدل (U_t) نیست که خود یک متغیر تصادفی است، در حالی که Y_t دقیقاً تابعی از U_t بوده و یک متغیر تصادفی خواهد بود. در این کتاب بیشتر از واژهٔ متغیرهای درون‌زا و برون‌زا استفاده می‌شود، هر چند در بعضی موارد و برحسب ضرورت از اصطلاحات دیگر نیز استفاده خواهد شد.

در مدل رگرسیون فوق، α و β پارامترهای مدل هستند. مقادیر واقعی این پارامترها را نمی‌دانیم، بنابراین سعی می‌کنیم مقادیر آنها را تخمین[5] بزنیم. یکی از وظایف اصلی اقتصادسنجی تخمین پارامترهای مدل رگرسیون است. با اینکه مقادیر واقعی پارامترهای α و β را نمی‌دانیم اما مطمئن هستیم این پارامترها غیرتصادفی هستند؛ زیرا فرض بر این است که تابعی از جملهٔ خطای مدل نیستند. با توجه به اینکه X_t نیز غیرتصادفی است به $(\alpha + \beta X_t)$ در اصطلاح قسمت معیّن[6] مدل رگرسیون می‌گویند. همان گونه که دیدیم U_t نیز قسمت نامعیّن یا تصادفی مدل است. بنابراین نتیجه می‌گیریم که ساختار ریاضی مدل رگرسیون مفروض به گونه‌ای است که برای تفسیر تغییرات متغیر درون‌زای Y_t، تغییرات آن را به دو قسمت تقسیم می‌کند: تغییرات معیّن،

---

1. Exogenous
2. Endogenous
3. Non-stochastic
4. Stochastic
5. Estimation
6. Deterministic Part

یعنی $(\alpha + \beta X_t)$ که در واقع همان مدل ریاضی است و تغییرات تصادفی، یعنی U_t که جملهٔ اختلال مدل محسوب می‌شود. به مجموعهٔ این دو تغییرات مدل رگرسیون گفته می‌شود.

تحلیلهای مدل رگرسیون می‌تواند در حوزه‌های مختلف صورت پذیرد. در این قسمت فقط به سه قلمرو اشاره می‌کنیم:

۱. آثار مختلف سیاستگذاریهای مرتبط با متغیرهای برون‌زا را می‌توان بر تغییرات متغیر درون‌زا تحلیل کرد. که در اصطلاح به آن تحلیل «سیاستگذاری[1]» می‌گویند؛ برای مثال، در بحث مدل مصرف ۷ـ۱ می‌توان آثار تغییر در سطوح مختلف درآمد را بر کیفیت ایجاد تغییر در سطوح مختلف مصرف تجزیه و تحلیل کرد.

۲. به ازای مقادیر معینی از متغیر یا متغیرهای برون‌زا می‌توان مقادیر متغیر درون‌زا را پیش‌بینی[2] کرد. برای مثال، می‌توان به ازای مقدار معینی از سطح درآمد در آینده مقدار مصرف را پیش‌بینی کرد.

۳. به کمک مدلهای رگرسیون می‌توان تحلیلهای آماری، تأثیر و یا بهتر بگوییم معنی‌دار بودن[3] تأثیر متغیرهای برون‌زا را بر متغیر درون‌زا مطالعه کرد. برای مثال، می‌توان به این سؤال پاسخ داد که آیا از نظر آماری تأثیر تغییرات درآمد بر تغییرات مصرف معتبر است؟ به عبارت دیگر آیا می‌توان تغییرات مصرف را به کمک تغییرات درآمد توضیح داد؟

با بررسی مدل رگرسیون ۸ـ۱ به سهولت ملاحظه می‌شود که هر گونه پیشرفت در تحلیلهای رگرسیونی متوقف به شناخت بیشتر از جملهٔ اختلال مدل (U_t) است. در واقع در یک مدل رگرسیون، U_t با اینکه نقش مهمی ایفا می‌کند اما بنا بر تعریف ناشناخته است. هر گاه کوشش کنیم اجزایی از U_t را بشناسیم و آنها را اندازه گیری کنیم، این اجزای شناخته شده، در قسمت معیّن مدل قرار می‌گیرد و مجموعهٔ عوامل مجهولی که باقی می‌ماند U_t را شکل می‌دهد؛ بنابراین U_t هیچگاه قابل مشاهده و اندازه گیری نیست.

---

1. Policy Analysis
2. Forecast or Prediction
3. Statistical Significance

در نتیجه راه خروج از این تنگنای نظری این است که یک سری فرضهای منطقی در مورد U_i مطرح کنیم تا بر آن اساس بتوان به تحلیلهای رگرسیونی ادامه داد. این فرضها با یک فرض در مورد متغیر برون‌زا، در ذیل با عنوان فرضهای کلاسیک مدلهای رگرسیون مطرح می‌شود.

۱. فرضهای کلاسیک مدلهای رگرسیون خطی ساده

در قسمت قبل، تعریف U_i و عوامل ایجادکنندهٔ آن را به طور مختصر بررسی کردیم. دیدیم که U_i را در اصطلاح جملهٔ اختلال می‌گویند؛ زیرا وجود U_i موجب می‌شود که قسمت معیّن مدل که در واقع همان مدل ریاضی ($Y_i = \alpha + \beta X_i$) است مختل شود. اصطلاحات دیگری نیز در متون اقتصادسنجی برای U_i به کار می‌رود؛ برای مثال، U_i را خطای معادله یا خطای مدل نیز می‌گویند؛ زیرا وجود U_i مبین خطای موجود در بیان دقیق عوامل و متغیرهایی است که می‌تواند تغییرات متغیر درون‌زا یعنی Y_i را در مدل ۸ـ۱ توضیح دهد. اصطلاح دیگری که برای U_i به کار می‌رود «جملهٔ تصادفی»[1] است؛ زیرا تأثیری که U_i بر تغییرات Y_i می‌گذارد، کاملاً تصادفی است. در گذشته بیشتر اصطلاح جملهٔ اختلال در کتابهای اقتصادسنجی رایج بود، اما به نظر می‌رسد امروز گرایش به کاربرد جملهٔ خطای مدل بیشتر شده است. با توجه به اینکه هنوز در کتابهای مرجع اقتصادسنجی، از واژهٔ جملهٔ اختلال استفاده می‌شود، ما نیز در این کتاب همین اصطلاح را برگزیده‌ایم.

مهمترین نکته در مورد U_i، خصوصیت تصادفی بودن آن است. با توجه به تعریفی که از U_i ارائه شد، بدیهی است این فرض قابل قبول است و خلاف آن را نمی‌توان تصور نمود. U_i یک متغیر تصادفی است و مانند همهٔ متغیرهای تصادفی دارای یک تابع توزیع احتمال است. می‌دانیم هر تابع توزیع احتمال دارای میانگین و واریانس است. سؤال مهمی که می‌توان مطرح کرد این است که خصوصیات آماری و شکل تابع

1. Random or Stochastic Term

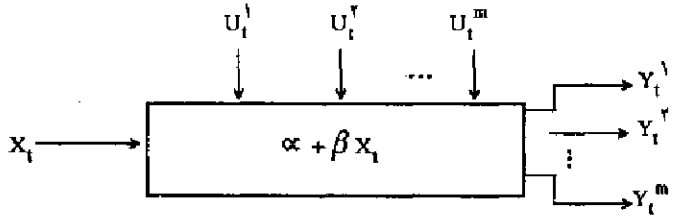
توزیع احتمال متغیر تصادفی U_t چیست؟ پاسخ به این سؤال با عنوان فرضهای کلاسیک چنین است:

اولین فرضی که در مورد U_t مطرح می‌شود این است که میانگین یا امید ریاضی آن صفر است. می‌دانیم امید ریاضی را با عملگر[1] E نشان می‌دهند؛ بنابراین در یک نمونه n تایی داریم،

$$E(U_t) = ۰ , t = ۱ , ۲ , ۳ , \dots , n , \qquad (۱-۹)$$

مفهوم این فرض اهمیت بسیار دارد. این فرض در واقع بدین معنی است که به ازای هر مقدار معین از X_t، میانگین تمام مقادیر ممکن U_t برابر صفر است. ابتدا باید به این سؤال پاسخ داد که به ازای یک مقدار ثابت و معیّن X_t چگونه U_t می‌تواند مقادیر مختلفی داشته باشد؟ بدیهی است دقت در میانگین مقادیر مختلف U_t که آیا صفر است یا خیر بعد از پاسخ به این سؤال مطرح می‌شود.

ظهور مقادیر مختلف U_t به اعتبار فرض آزمایشهای فرضی تکراری[2] به ازای مقدار معیّن و ثابت X_t است. برای تبیین این مطلب به نمودار ۱-۱ توجه می‌کنیم.



نمودار ۱-۱ مفهوم امید ریاضی U_t

مقدار X_t را ثابت فرض کرده، با توجه به مقادیر ثابت α و β نتیجه می‌گیریم که $(\alpha + \beta X_t)$ نیز مقدار معیّنی خواهد داشت. مقدار Y_t چیست؟ فرض می‌کنیم در دورهٔ زمانی t هستیم. در طی این دوره کلیه عواملی که در تعیّن U_t مؤثر هستند فعال می‌شود و مقدار U_t را شکل می‌دهند. می‌دانیم U_t یک متغیر تصادفی است، یعنی هر مقداری که

1. Operator
2. Repeated Hypothetical Experiment

U_i بدین ترتیب اتخاذ کند در واقع می‌توانست مقدار دیگری باشد. اگر بخواهیم این نکته را به زبانی دیگر بیان کنیم، کافی است بگوییم که در آزمایش فرضی اول، U_i مقدار U_i^1 را می‌گیرد. این مقدار با $(\alpha + \beta X_i)$ جمع می‌شود و Y_i^1 را نتیجه می‌دهد. حال می‌گوییم اگر این آزمایش را دوباره انجام دهیم، با توجه به تصادفی بودن U_i، هیچ دلیلی وجود ندارد که U_i درست همان مقدار قبلی را بگیرد؛ بنابراین باید گفت U_i در آزمایش فرضی دوم، مقدار U_i^2 را خواهد داشت و وقتی با مقدار ثابت $(\alpha + \beta X_i)$ جمع شود، Y_i^2 را تشکیل می‌دهد. به همین ترتیب می‌توان m آزمایش فرضی را تصور کرد که مقادیر U_i^m، ...، U_i^2 و U_i^1، با فرض ثباتِ $(\alpha + \beta X_i)$ به ترتیب مقادیر Y_i^m، ...، Y_i^2 و Y_i^1 را نتیجه می‌دهند. انتظار این است که به ازای مقادیر بزرگ m، میانگین یا امید ریاضی U_i صفر باشد؛ زیرا مقادیر مختلف U_i^m، ...، U_i^2 و U_i^1 در حول محور ثابت $(\alpha + \beta X_i)$ نوسان دارند؛ به عبارت دیگر بعضی از مقادیر فوق U_i در بالای این خط ثابت قرار می‌گیرد و بعضی دیگر در پایین و در مجموع مقادیر مثبت و منفی، همدیگر را خنثی می‌کند؛ به گونه‌ای که در نهایت امید ریاضی یا میانگین U_i برابر صفر خواهد شد. با وجود این، باید در نظر داشت که هیچیک از مقادیر U_i در آزمایشهای تکراری قابل مشاهده نیست و آنچه مطرح شد در واقع دقت در خصوصیات آماری این متغیر تصادفی است.

دومین فرضی که برای U_i مطرح می‌کنیم، ثابت بودن واریانس آن به ازای مقادیر مختلف X_i است،

$$\mathrm{Var}(U_i) = \mathrm{E}(U_i)^2 = \sigma^2 \quad \text{و} \quad i = 1, 2, \ldots, n \ . \qquad (۱.۱۰)$$

اگر مقدار X_i در دورهٔ i، یعنی X_i را ملاحظه کنیم و نمودار ۱-۱ را برای آن بسازیم، m مقدار مختلف برای U_i به دست خواهیم آورد. واریانس این مقادیر مختلف U_i را با σ_i^2 نشان می‌دهیم؛ یعنی $\mathrm{Var}(U_i) = \sigma_i^2$. به همین ترتیب اگر مقدار X_i را در دورهٔ j ملاحظه کرده، واریانس U_i را به ترتیبی مشابه تعریف کنیم، خواهیم داشت: $\mathrm{Var}(U_j) = \sigma_j^2$. فرض ثابت بودن واریانس U_i این است که برای تمام مقادیر i و j این واریانسها با یکدیگر برابر باشد،

$$\mathrm{Var}(U_i) = \mathrm{Var}(U_j) = \sigma_i^2 = \sigma_j^2 \quad , \quad i, j = ۱, ۲, \ldots, n .$$

با توجه به تعریف واریانس،

$$\mathrm{Var}(U_i) = E\left[U_i - E(U_i)\right]^2 ,$$

و با در نظر گرفتن فرض اول، یعنی $E(U_i) = ۰$، خواهیم داشت

$$\mathrm{Var}(U_i) = E(U_i)^2 = \sigma^2 .$$

هر گاه واریانس جملهٔ اختلال ثابت باشد، می‌گوییم U_i واریانس همسانی[1] دارد؛ در غیر این صورت U_i واریانس ناهمسان[2] است.

سومین فرض U_i این است که U_i و U_j به ازای تمامی مقادیر $i \neq j$ از یکدیگر مستقلند؛ یعنی کوواریانس آنها صفر است.

$$\mathrm{cov}(U_i, U_j) = E(U_i U_j) = ۰ \quad , \quad i, j = ۱, ۲, \ldots, n , \quad i \neq j \qquad (۱\text{-}۱۱)$$

به عبارت دیگر، هر گاه دو مقدار X_i، برای مثال، X_i و X_j را در نظر بگیریم، فرض بر این است که جمله‌های اختلال متناظر با آنها، یعنی U_i و U_j از یکدیگر مستقلند. همچنین با توجه به تعریف کوواریانس داریم

$$\mathrm{cov}(U_i, U_j) = E\left[U_i - E(U_i)\right]\left[U_j - E(U_j)\right] = ۰ ,$$

و با در نظر گرفتن فرض اول، یعنی $E(U_i) = ۰$، رابطهٔ مذکور را می‌توان به صورت زیر نوشت،

$$\mathrm{cov}(U_i, U_j) = E(U_i U_j) = ۰ .$$

در چنین حالتی می‌گوییم که جمله‌های اختلال خود همبستگی[3] ندارد و در غیر این صورت (هنگامی که $E(U_i, U_j) \neq ۰$) جمله‌های اختلال خودهمبستگی خواهد داشت. اگر جمله‌های اختلال، واریانس همسان و خودهمبستگی نداشته باشند، می‌گوییم U_i

---

1. Homoskedastic 2. Heteroskedastic 3. Autocorrelation

«کروی»[1] است.

چهارمین فرض U_t این است که تابع توزیع احتمال آن را نرمال بدانیم. بنابراین با توجه به فرضهای اول، دوم و سوم می‌توان گفت که U_t دارای توزیع مستقل نرمال با میانگین صفر و واریانس ثابت σ^2 است. اگر علامت $\sim$ را برای توزیع و IN را برای توزیع مستقل نرمال[2] به کار بریم، آنگاه خواهیم داشت

$$U_t \sim IN(0, \sigma^2). \qquad (۱-۱۲)$$

در بعضی از منابع، به جای علامت IN از NID[3] استفاده می‌کنند،

$$U_t \sim NID(0, \sigma^2).$$

پنجمین و آخرین فرض از فرضهای کلاسیک این است که X_t یک متغیر غیرتصادفی است. این فرض بیشتر برای سهولت در استنتاج قضایا و نیز رسیدن به نتایج جالبتر در تخمین پارامترهاست. بدیهی است که می‌توان این فرض را نقض کرد و X_t را به صورت یک متغیر تصادفی در نظر گرفت. به هر حال فرض غیرتصادفی بودن X_t بدین معنی است که X_t از متغیر تصادفی U_t مستقل است (در این جلد تمام نتایج به صورت شرطی و با فرض غیرتصادفی بودن X_t بیان شده است).

در بعضی از کتابهای اقتصادسنجی، تصادفی بودن U_t و متساوی نبودن مقادیر مختلف X_t در یک نمونه را نیز فرضهای کلاسیک می‌نامند و در نتیجه هفت فرض را با این عنوان مطرح می‌کنند. در معادلهٔ ۲۵-۱ خواهیم دید که در صورت تساوی مقادیر X_t با یکدیگر، تخمین پارامترها ممکن نخواهد بود.

۲. خصوصیات آماری متغیر درون‌زا

یک بار دیگر مدل رگرسیون ۸-۱ را در نظر می‌گیریم،

---

1. Spherical Disturbances
2. Independent Normal
3. Normally and Independently Distributed

$$Y_t = \alpha + \beta X_t + U_t .$$

ملاحظه می‌شود که چون Y_t تابعی از جملهٔ اختلال (U_t) است یک متغیر تصادفی خواهد بود؛ زیرا U_t یک متغیر تصادفی است. Y_t مانند هر متغیر تصادفی دیگر دارای تابع توزیع احتمال است و یک میانگین و یک واریانس خواهد داشت.

با توجه به اینکه در مدل رگرسیون خطی ۸ـ۱ متغیر درون‌زای Y_t تابعی از U_t است و علاوه بر این چون فرض کرده‌ایم U_t دارای توزیع نرمال است بنابراین متغیر تصادفی Y_t نیز دارای توزیع احتمال نرمال خواهد بود. برای یافتن میانگین آن کافی است از دو طرف رابطهٔ ۸ـ۱ امید ریاضی بگیریم خواهیم داشت

$$E(Y_t) = E(\alpha) + E(\beta X_t) + E(U_t) .$$

با توجه به اینکه α ثابت است پس $E(\alpha) = \alpha$. همچنین فرض کرده‌ایم X_t یک متغیر غیرتصادفی بوده و در آزمایشهای تکراری ثابت است؛ بنابراین با توجه به ثابت بودن β داریم

$$E(\beta X_t) = \beta E(X_t) = \beta X_t .$$

اگر اولین فرض مربوط به U_t، یعنی $E(U_t) = 0$ را نیز درنظربگیریم، آنگاه خواهیم داشت

$$E(Y_t) = \alpha + \beta X_t . \qquad (۱ـ۱۳)$$

به رابطهٔ ۱۳ـ۱ معمولاً تابع رگرسیون جامعه[1] می‌گویند. همان گونه که بعداً خواهیم دید اگر به جای α و β تخمین آنها، یعنی $\hat{\alpha}$ و $\hat{\beta}$ را بگذاریم در آن صورت تابع رگرسیون نمونه[2] حاصل خواهد شد[3]. تابع خطی ۱۳ـ۱ در حقیقت مکان هندسی نقاطی است که برای آنها $E(U_t) = 0$ است.

برای محاسبهٔ واریانس Y_t کافی است به تعریف واریانس مراجعه کنیم،

$$Var(Y_t) = E[Y_t - E(Y_t)]^2 .$$

---

1. Population Regression Function
2. Sample Regression Function
۳. معمولاً تخمین پارامترها را با علامت ^ نشان می‌دهیم و آن را کلاه (Hat) می‌خوانیم.

با جایگزینی مقادیر Y_t و $E(Y_t)$ به ترتیب از رابطه‌های ۱ـ۸ و ۱ـ۱۳، مقدار $Var(Y_t)$ به ترتیب زیر محاسبه می‌شود،

$$Var(Y_t) = E[\alpha + \beta X_t + U_t - (\alpha + \beta X_t)]^2$$
$$= E(U_t)^2 = \sigma^2 . \qquad (۱ـ۱۴)$$

واریانس متغیر درون‌زای Y_t دقیقاً برابر واریانس جملهٔ اختلال مدل رگرسیون است.

نتیجه می‌گیریم که Y_t یک متغیر درون‌زای نرمال با میانگین $(\alpha + \beta X_t)$ و واریانس σ^2 است. چون فرض کرده بودیم مقادیر مختلف U_t دارای توابع توزیع نرمال مستقل از یکدیگر هستند بنابراین Y_t نیز دقیقاً چنین خصوصیتی خواهد داشت. این خصوصیات آماری Y_t را می‌توان به صورت زیر نشان داد،

$$Y_t \sim IN[(\alpha + \beta X_t) \ , \ \sigma^2] . \qquad (۱ـ۱۵)$$

۳. بیان هندسی خصوصیات آماری مدلهای رگرسیون خطی ساده

آنچه را که تا به حال در مورد خصوصیات آماری یک مدل رگرسیون خطی گفته‌ایم، در یک نمودار نشان خواهیم داد. در یک مدل رگرسیون خطی ساده سه متغیر اصلی داریم: X_t، Y_t و U_t. برای سهولت کافی است در یک نمودار سه بعدی مقادیر X_t، Y_t و $P(U_t)$ را منعکس سازیم.

ابتدا باید نمودارهای تابع رگرسیون جامعه و تابع رگرسیون نمونه را رسم کرد. تابع رگرسیون جامعه، یعنی مدل ۱ـ۱۳ را یک بار دیگر همراه با مدل رگرسیون خطی ۱ـ۸ می‌نویسیم،

$$Y_t = \alpha + \beta X_t + U_t ,$$

$$E(Y_t) = \alpha + \beta X_t .$$

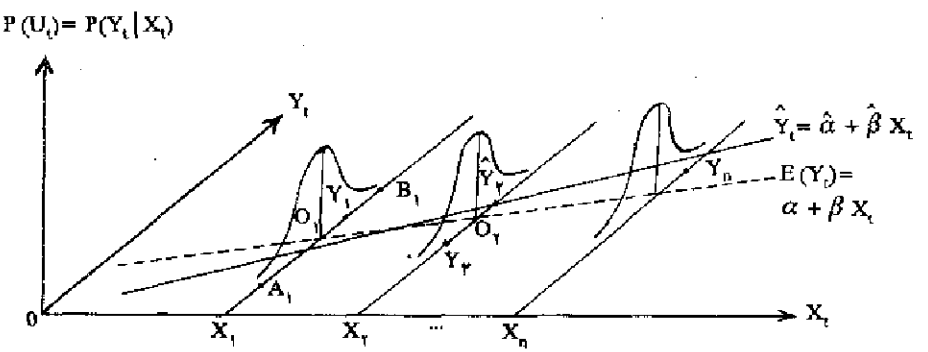
یکی از مباحث اصلی تحلیلهای رگرسیونی تخمین مدلهای رگرسیون است. منظور از تخمین یک مدل رگرسیونی در واقع تخمین پارامترهای آن است؛ بنابراین اگر تخمینهای

α، β و Y_i را به ترتیب با $\hat{\alpha}$، $\hat{\beta}$ و $\hat{Y}_i$ نشان دهیم، تابع یا مدل رگرسیون نمونه ـ که در واقع تخمین مدل رگرسیون ۸ـ۱ است ـ عبارت خواهد بود از

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i . \qquad (۱\text{ـ}۱۶)$$

رسیدن به معادلهٔ ۱ـ۱۶ یکی از اهداف مهم تخمین در تحلیلهای رگرسیونی است. در مباحث آینده خواهیم دید که $\hat{\alpha}$ و $\hat{\beta}$ قابل اندازه‌گیری است و در نتیجه می‌توان مدل رگرسیون نمونهٔ ۱ـ۱۶ را رسم کرد. اما پارامترهای واقعی جامعه (α و β) هیچگاه قابل مشاهده و اندازه‌گیری نیستند؛ زیرا U_i اساساً قابل مشاهده نیست؛ بنابراین نمودار هندسی مدل رگرسیون جامعهٔ ۱ـ۱۳ در عمل امکان‌پذیر نیست.

در نمودار ۱ـ۲ ابتدا منحنی تغییرات مدل رگرسیون جامعه را به صورت یک خط مستقیم و فرضی با نقطه‌چین رسم می‌کنیم؛ آنگاه منحنی تغییرات مدل رگرسیون نمونه را به صورت یک خط مستقیم نشان می‌دهیم. از این به بعد به مدل رگرسیون نمونه «تخمین مدل رگرسیون جامعه» یا به طور خلاصه «تخمین مدل رگرسیون» می‌گوییم.



نمودار ۱ـ۲ خصوصیات آماری مدل رگرسیون خطی

X_i در آزمایشهای فرضی تکراری ثابت فرض می‌شود. برای n دوره مشاهده، یعنی $t = ۱, ۲, ۳, \ldots, n$ فرض کنیم $X_۱$ (مقدار X) را در اولین دوره مشاهده ثابت گرفته‌ایم، این مقدار ثابت در نمودار ۱ـ۲ با $OX_۱$ مشخص شده است. می‌دانیم به ازای $X_۱$ معین، $Y_۱$ می‌تواند مقادیر متفاوتی داشته باشد. اما این مقادیر دقیقاً تابعی از این نکته

است که U_1 چه مقادیری کسب می‌کند. فرض بر این است که تابع توزیع احتمال U_1 نرمال است؛ بنابراین قلمرو تغییرات U_1 مثلاً محدود به $A_1 B_1$ است. حال با فرض ثبات X_1 می‌گوییم که در آزمایش اول فرض می‌کنیم Y مقدار $X_1 Y_1$ را گرفته است.

اگر مدل رگرسیون ۸ـ۱ را برای دورهٔ اول، یعنی $t = 1$ بنویسیم، خواهیم داشت

$$Y_1 = \alpha + \beta X_1 + U_1 .$$

این واقعیت که در آزمایش اول، Y در سال اول مقدار Y_1 را گرفته است، این نکته را بیان می‌کند که جملهٔ اختلال (U) در این آزمایش، مقدار مثبت $O_1 Y_1$ را داشته است؛ زیرا می‌دانیم $X_1 O_1 = E(Y_1)$ در واقع همان $(\alpha + \beta X_1)$ است. البته در آزمایشهای فرضی دیگر، U می‌تواند هر مقدار دیگری را در قلمرو $A_1 B_1$ بگیرد و بدین ترتیب مقادیر مثبت یا منفی دیگری به مقدار ثابت $\alpha + \beta X_1$ که در واقع همان $E(Y_1)$ است اضافه کند. نتیجه می‌گیریم که به ازای X_1 ثابت و بعد از ظهور مقدار Y_1 در عینیت، می‌توان مقدار U_1 را به صورت زیر نوشت،

$$\begin{aligned} U_1 &= X_1 Y_1 - (\alpha + \beta X_1) , \\ &= Y_1 - E(Y_1) . \end{aligned}$$

بنابراین با توجه به اینکه مقدار مشاهده شدهٔ Y_1 در عینیت بالای خط فرضی رگرسیون جامعه $(E(Y_1) = \alpha + \beta X_1)$ یا پایین آن قرار می‌گیرد، می‌توان نتیجه گرفت که مقدار U_1 مثبت یا منفی خواهد بود.

همین بحث را دقیقاً برای دوره‌های $t = 2, 3, \ldots, n$ تکرار می‌کنیم و به ازای X_2، X_3، تا X_n و مشاهده مقادیر Y_2، Y_3 تا Y_n مقادیر U را استنتاج می‌کنیم. در حالت کلی می‌توان چنین نوشت

$$\begin{aligned} U_t &= X_t Y_t - (\alpha + \beta X_t) , \\ &= Y_t - E(Y_t) . \end{aligned} \qquad (۱ـ۱۷)$$

در نمودار ۲ـ۱ فرض $E(U_t) = 0$ را می‌توان بدین صورت منعکس کرد که به ازای

هرمقدار X_t ثابت، منحنی تابع توزیع‌احتمال U_t در حول محور $E(Y_t)=\alpha+\beta X_t$ متمرکز بوده و دارای میانگین صفر است. به همین دلیل است که منحنی $E(Y_t)=\alpha+\beta X_t$ را می‌توان مکان هندسی نقاطی دانست که برای آنها $E(U_t)$ برابر صفر است فرض ثابت بودن واریانس U_t به ازای مقادیر مختلف t (فرض واریانس همسانی) چیزی نیست جز اینکه بگوییم منحنی نمایش هندسی توابع توزیع احتمال U_t به ازای مقادیر ثابت X_1 ، X_2 تا X_n دارای پراکندگی یکسانی است. فرض عدم وجود خودهمبستگی (صفر بودن کوواریانس U_i و U_j) به این معنی است که مقدار U در دورهٔ i مستقل از مقداری است که U در دورهٔ j گرفته است یا خواهد گرفت و برعکس. سرانجام تعبیر هندسی فرض غیرتصادفی بودن X_t بدین گونه است که مقدار X_t مثلاً در دورهٔ دوم، t=۲ ، به طور مطلق تابعی از مقادیر جملهٔ اختلال، (U_t) در این دوره، دورهٔ قبل یا دورهٔ بعد نیست؛ به عبارت دیگر، مقدار X در دورهٔ i اساساً تابعی از این نیست که جملهٔ اختلال در دورهٔ i + ۱ چه مقداری خواهد داشت یا در دورهٔ i − ۱ چه مقداری داشته است.

۱ـ۳ تخمین مدل رگرسیون خطی: روش حداقل مربعات معمولی (OLS)[1]

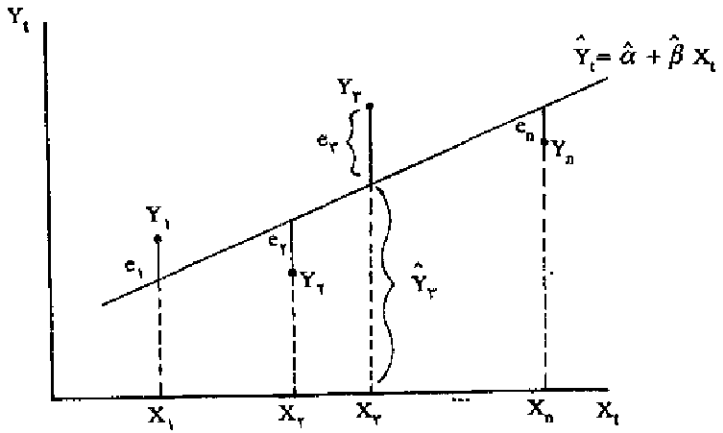
مدل رگرسیون خطی ساده ۱ـ۸ را یک بار دیگر می‌نویسیم،

$$Y_t=\alpha+\beta X_t+U_t.$$

می‌خواهیم پارامترهای این مدل $(\alpha$ و $\beta)$ را تخمین بزنیم. برای این منظور مشاهدات متغیرهای برون‌زا و درون‌زا $(X_t$ و $Y_t)$ را برای n دوره، $t=1,2,3,\ldots,n$ ، جمع‌آوری کرده‌ایم. در نمودار ۱ـ۲ این مشاهدات را با $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ نشان داده‌ایم. برای سهولت بحث، صفحهٔ مختصات XOY از نمودار سه بعدی ۱ـ۲ را در نمودار ۱ـ۳ نشان می‌دهیم. برای تخمین پارامترهای α و β ضرورتاً نیاز به یک روش تخمین داریم. در اقتصادسنجی روشهای مختلفی برای تخمین پارامترها وجود دارد. ساده‌ترین و بهترین روش تخمین، روش حداقل مربعات معمولی است. طرح اولیهٔ این

1. Ordinary Least Squares

روش را که معمولاً با OLS نشان داده می‌شود گاس[۱] ریاضیدان معروف آلمانی در قرن هیجدهم مطرح کرده است.



نمودار ۱.۳ روش تخمین OLS

مدل رگرسیون ۱ـ۸ در واقع یک مدل خطی است که می‌خواهد تغییرات متغیر درون‌زای Y_i را در خلال مشاهدات n دوره توضیح دهد. زیربنای فکری روش حداقل مربعات معمولی (OLS) این است که بتوانیم این مدل خطی را با یک مدل خطی دیگر از نوع

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i ,$$

چنان تخمین بزنیم که در بالاترین سطح ممکن، بتواند n مشاهدهٔ ما را توضیح دهد. مدل فوق که همان مدل ۱ـ۱۶ است تخمین مدل رگرسیون ۱ـ۸ است. در حقیقت باید $\hat{\alpha}$ و $\hat{\beta}$ چنان مقادیری داشته باشند که مدل ۱ـ۱۶ نسبت به هر مدل خطی دیگر، بیشترین نزدیکی را به مشاهدات Y_n , ... , Y_2 , Y_1 داشته باشد؛ به عبارت دیگر کمترین انحراف را از مشاهدات فوق نشان دهد.

اگر معادلهٔ خط تخمین مدل رگرسیون (۱ـ۱۶) را در نمودار ۱ـ۳ رسم کنیم

---

۱. Carl Friedrich Gauss 1777-1855 ؛ با اینکه روش حداقل مربعات ابتدا در کتاب لژاندر (A. M. Legendre) ریاضیدان معروف فرانسوی در سال ۱۸۰۵ به چاپ رسید، اما گاس از سال ۱۷۹۵ ازاین روش در حل بسیاری از مسائل استفاده کرده است. برای توضیحات بیشتر به پیوست «د» مراجعه کنید.

بنا بر اقتضای بحث این خط باید کمترین فاصله را با مشاهدات ما داشته باشد. به ازای مقادیر $X_1, X_2, \ldots, X_n$، تخمین مدل رگرسیون، یعنی معادله ۱-۱۶، مقادیر متناظر Y_t را به ترتیب به صورت $\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n$ تخمین می‌زند. بدیهی است تخمین ما از مدل رگرسیون نمی‌تواند مقداری برای $\hat{Y}_1$ بدهد که دقیقاً برابر Y_1 باشد. به همین ترتیب در حالت کلی نمی‌توان انتظار داشت $\hat{Y}_t$ برابر Y_t شود. اختلاف بین مشاهده (Y_t) و تخمین ($\hat{Y}_t$) را اصطلاحاً «پسماند»[1] یا «انحراف»[2] گفته و با e_t نشان می‌دهیم:

$$e_t = Y_t - \hat{Y}_t . \qquad (۱-۱۸)$$

اصطلاح «خطای تخمین»[3] نیز در مواردی برای این مفهوم به کار رفته است. ناگفته نماند که e_t را با $\hat{U}_t$ نیز می‌توان نشان داد.

معیار روش حداقل مربعات معمولی این است که α و β را باید چنان تخمین زد که مجموع مربعات پسماند به حداقل برسد؛ به عبارت دیگر باید $\hat{\alpha}$ و $\hat{\beta}$ را چنان محاسبه کرد که Σe_t^2 به حداقل برسد. علت اینکه به جای مجموع سادهٔ پسماند، مجموع مربع آنها را در نظر می‌گیریم، این است که می‌خواهیم نه تنها انحرافات مثبت و منفی یکدیگر را خنثی نکنند، بلکه انحرافهای بزرگ نسبت به انحرافهای کوچک از اهمیت بیشتری برخوردار شود. می‌دانیم، $\sum_1^n e_t^2$ برابر است با

$$\sum_1^n e_t^2 = \sum_1^n (Y_t - \hat{Y}_t)^2 ,$$

$$= \Sigma (Y_t - \hat{\alpha} - \hat{\beta} X_t)^2 . \qquad (۱-۱۹)$$

بنابراین مسأله تخمین α و β به مسأله حداقل‌سازی رابطهٔ ۱-۱۹ نسبت به $\hat{\alpha}$ و $\hat{\beta}$ تبدیل می‌شود. تخمین پارامترها با روش حداقل مربعات معمولی (OLS) همواره می‌تواند به حداقل‌سازی یک تابع تبدیل شود. بدین ترتیب باید Σe_t^2 را در رابطهٔ ۱-۱۹

1. Residual 2. Deviation 3. Estimation Error

نسبت به $\hat{\alpha}$ و $\hat{\beta}$ حداقل کنیم. کافی است از $\sum e_t^2$ نسبت به $\hat{\alpha}$ و $\hat{\beta}$ مشتق بگیریم و آنها را مساوی صفر قرار داده و $\hat{\alpha}$ و $\hat{\beta}$ مطلوب را محاسبه کنیم.

$$\frac{\partial \sum e_t^2}{\partial \hat{\alpha}} = \frac{\partial \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t)^2}{\partial \hat{\alpha}} = 0 ,$$

$$= 2 \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t) (-1) = 0 ,$$

$$= \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t) = 0 .$$

$$\frac{\partial \sum e_t^2}{\partial \hat{\beta}} = \frac{\partial \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t)^2}{\partial \hat{\beta}} = 0 ,$$

$$= 2 \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t) (-X_t) = 0 ,$$

$$= \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t) (X_t) = 0 .$$

نتایج این مشتق‌گیریهای جزئی را «معادله‌های اول و دوم نرمال»[1] می‌گویند. این دو معادله را به صورت زیر می‌نویسیم،

$$\begin{cases} \sum Y_t = \sum \hat{\alpha} + \sum \hat{\beta} X_t , \\ \sum X_t Y_t = \sum \hat{\alpha} X_t + \sum \hat{\beta} X_t^2 , \end{cases}$$

با توجه به اینکه n مشاهده داریم و مقادیر $\hat{\alpha}$ و $\hat{\beta}$ نیز ثابت است، خواهیم داشت

$$\begin{cases} \sum Y_t = n \hat{\alpha} + \hat{\beta} \sum X_t , \\ \sum X_t Y_t = \hat{\alpha} \sum X_t + \hat{\beta} \sum X_t^2 . \end{cases} \qquad (۱-۲۰)$$

از حل دستگاه دو معادلهٔ دو مجهولی ۱-۲۰ می‌توان $\hat{\alpha}$ و $\hat{\beta}$ را به دست آورد،

1. Normal Equations

$$\hat{\alpha}_{OLS} = \frac{\sum X_t^2 \sum Y_t - \sum X_t \sum X_t Y_t}{n \sum X_t^2 - (\sum X_t)^2}, \qquad (۱\text{-}۲۱)$$

$$\hat{\beta}_{OLS} = \frac{n \sum X_t Y_t - \sum X_t \sum Y_t}{n \sum X_t^2 - (\sum X_t)^2}. \qquad (۱\text{-}۲۲)$$

منظور از $\hat{\alpha}_{OLS}$ و $\hat{\beta}_{OLS}$ در فرمولهای فوق این است که تخمینهای α و β با استفاده از روش حداقل مربعات معمولی به دست آمده است. به فرمولهای ۱ـ۲۱ و ۱ـ۲۲ که در واقع فرمولهای تخمین هستند، «تخمین زننده»[1] می‌گویند.

می‌توان به فرمولهای محاسباتی ساده‌تری برای $\hat{\alpha}_{OLS}$ و $\hat{\beta}_{OLS}$ رسید. اما قبل از آن دو نتیجهٔ بسیار مهم را که از معادله‌های نرمال حاصل می‌شوند بررسی می‌کنیم. معادلهٔ اول نرمال را یک بار دیگر می‌نویسیم،

$$\frac{\partial \sum e_t^2}{\partial \hat{\alpha}} = \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t) = 0.$$

با توجه به معادلهٔ ۱ـ۱۸ می‌توان از معادله فوق نتیجه گرفت که مجموع پسماندها و میانگین آنها صفر است،

$$\bar{e} = 0 \quad \text{یا} \quad \sum e_t = 0. \qquad (۱\text{-}۲۳)$$

معادلهٔ دوم نرمال را نیز یک بار دیگر می‌نویسیم،

$$\frac{\partial \sum e_t^2}{\partial \hat{\beta}} = \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t)(X_t) = 0,$$

با توجه به معادلهٔ ۱ـ۱۸ می‌توان گفت که مجموع حاصلضرب پسماندها در متغیر توضیحی برابر صفر است،

$$\sum e_t X_t = 0.$$

همچنین می‌توان نشان داد که $\Sigma e_t X_t = \Sigma e_t x_t = 0$؛ زیرا با توجه به معادلهٔ ۱ـ۲۳

1. Estimator

داریم:

$$\sum e_t X_t = \sum e_t (x_t + \bar{X}) = \sum e_t x_t - \bar{X} \sum e_t = \sum e_t x_t = ۰ .$$

بنابراین جملهٔ پسماند و متغیر توضیحی از یکدیگر مستقل هستند؛ زیرا

$$\sum (e_t - \bar{e}) (X_t - \bar{X}) = \sum e_t x_t = ۰ . \qquad (۱\text{-}۲۴)$$

با استفاده از تعاریف $y_t = Y_t - \bar{Y}$ و $x_t = X_t - \bar{X}$ که در آن $\bar{X}$ و $\bar{Y}$ به ترتیب میانگین X_t و Y_t هستند، می‌توان به فرمولهای ساده‌تری برای $\hat{\alpha}$ و $\hat{\beta}$ نیز رسید. می‌دانیم،

$$\sum x_t y_t = \frac{n \sum X_t Y_t - \sum X_t \sum Y_t}{n} \quad \text{و} \quad \sum x_t^۲ = \frac{n \sum x_t^۲ - (\sum X_t^۲)}{n} .$$

اگر صورت و مخرج تخمین‌زنندهٔ $\hat{\beta}_{OLS}$ در رابطهٔ ۱ـ۲۲ را بر n تقسیم کرده، از رابطه‌های فوق استفاده کنیم، خواهیم داشت

$$\hat{\beta}_{OLS} = \frac{\sum x_t y_t}{\sum x_t^۲} . \qquad (۱\text{-}۲۵)$$

دو طرف معادلهٔ اول از سیستم معادله‌های ۱ـ۲۰ را بر n تقسیم می‌کنیم،

$$\frac{\sum Y_t}{n} = \hat{\alpha} + \hat{\beta} \frac{\sum X_t}{n} , \qquad \text{یا}$$

$$\bar{Y} = \hat{\alpha} + \hat{\beta} \bar{X} . \qquad (۱\text{-}۲۶)$$

رابطه فوق بر این دلالت می‌کند که مختصات $\bar{Y}$ و $\bar{X}$ در تخمین مدل رگرسیون (معادلهٔ ۱ـ۱۶)، صدق می‌کند؛ به عبارت دیگر می‌توان گفت که مدل $\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t$ از نقطهٔ میانگین X_t و میانگین Y_t می‌گذرد. به سهولت ملاحظه می‌شود که

$$\hat{\alpha}_{OLS} = \bar{Y} - \hat{\beta} \bar{X} .$$

برای تبیین کیفیت استفاده از تخمین‌زننده‌های $\hat{\beta}_{OLS}$ و $\hat{\alpha}_{OLS}$ در حل مسائل،

به ذکر چند مثال می‌پردازیم. (در ادامه بحث هر گاه تخمین‌زننده‌ای بدون اندیس نوشته شود، منظور تخمین زنندهٔ حداقل مربعات معمولی (OLS) است).

مثال ۱ـ۱ مدل رگرسیون مصرف زیر مفروض است،

$$C_t = \alpha + \beta Y_t + U_t .$$

با استفاده از آمار موجود در جدول ۱ـ۱ پارامترهای α و β را با استفاده از روش حداقل مربعات معمولی تخمین بزنید. مقادیر پسماند را نیز حساب کنید و مجموع مربع آنها را به دست آورید. اگر در دورهٔ ٦ = t مقدار درآمد ۱۲ میلیارد تومان باشد، سطح مصرف را پیش‌بینی کنید.

بهتر است نام متغیرهای موجود در مسأله را به X_t و Y_t تغییر دهیم تا بهتر بتوان از فرمولها یا تخمین‌زننده‌های ۱ـ۲۵ و ۱ـ۲٦ استفاده کرد.

$$Y_t = \text{مصرف} \quad \text{و} \quad X_t = \text{درآمد}$$

محاسبات لازم را در جدولی به شرح زیر تنظیم می‌کنیم.

جدول ۱ـ۲

t	Y_t	X_t	$y_t = Y_t - \bar{Y}$	$x_t = X_t - \bar{X}$	$x_t y_t$	x_t^2	$\hat{Y}_t$	$e_t = Y_t - \hat{Y}_t$	e_t^2
۱	۲	۳	۲ - ۵ = -۳	۳ - ۸ = -۵	۱۵	۲۵	۱/٦۵	۲ - ۱/٦۵ = ۰/۳۵	۰/۱۲۲۵
۲	۳	۵	۳ - ۵ = -۲	۵ - ۸ = -۳	٦	۹	۲/۹۹	۳ - ۲/۹۹ = ۰/۰۱	۰/۰۰۰۱
۳	۵	۹	۵ - ۵ = ۰	۹ - ۸ = ۱	۰	۱	۵/٦۷	۵ - ۵/٦۷= -۰/٦۷	۰/٤٤۸۹
٤	٦	۱۰	٦ - ۵ = ۱	۱۰ - ۸ = ۲	۲	٤	٦/۳٤	٦ - ٦/۳٤ = -۰/۳٤	۰/۱۱۵٦
۵	۹	۱۳	۹ - ۵ = ٤	۱۳ - ۸ = ۵	۲۰	۲۵	۸/۳۵	۹ - ۸/۳۵ = ۰/٦۵	۰/٤۲۲۵
Σ	۲۵	٤۰	۰	۰	٤۰	٦٤	۲۵	۰	۱/۱۰۹٦

$$\hat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} \quad , \quad \bar{Y} = \frac{\Sigma Y_t}{n} = \frac{۲۵}{۵} = ۵ \quad , \quad \bar{X} = \frac{\Sigma X_t}{n} = \frac{٤۰}{۵} = ۸ ,$$

$$= \frac{٤۳}{٦٤} = ۰/٦۷ ,$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{X}\,,$$

$$= ۵ - ۰/۶۷(۸) = -۰/۳۶\,,$$

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta}\,X_i\,,$$

$$\hat{Y}_i = -۰/۳۶ + ۰/۶۷\,X_i\,.$$

بر اساس نامگذاریهای موجود در مسأله داریم

$$\hat{C}_i = -۰/۳۶ + ۰/۶۷\,Y_i\,.$$

میل نهایی به مصرف[1]، ۰/۶۷ = mpc است. تخمینهایی که از پارامترها زده‌ایم، بر این دلالت می‌کند که اگر درآمد صفر شود، مصرف برابر ۰/۳۶ میلیارد تومان است. بدیهی است چنین نتیجه‌ای مطلقاً بی‌معنی است. بنابراین، محاسبات متغیر درون‌زا (متغیر موجود در سمت چپ مدل رگرسیون) فقط باید به ازای قلمرو محدودی از تغییرات متغیر برون‌زا صورت گیرد و در خارج از این قلمرو معمولاً نتایج، بی‌معنی خواهد بود.

با توجه به $\hat{Y}_i = \hat{\alpha} + \hat{\beta}\,X_i$، مقادیر $\hat{Y}_i$ را به ترتیب زیر محاسبه می‌کنیم.

$$\hat{Y}_۱ = -۰/۳۶ + ۰/۶۷\,(۳) = ۱/۶۵\,,$$

$$\hat{Y}_۲ = -۰/۳۶ + ۰/۶۷\,(۵) = ۲/۹۹\,,$$

$$\hat{Y}_۳ = -۰/۳۶ + ۰/۶۷\,(۹) = ۵/۶۷\,,$$

$$\hat{Y}_۴ = -۰/۳۶ + ۰/۶۷\,(۱۰) = ۶/۳۴\,,$$

$$\hat{Y}_۵ = -۰/۳۶ + ۰/۶۷\,(۱۳) = ۸/۳۵\,.$$

با استفاده از رابطهٔ $e_i = Y_i - \hat{Y}_i$، به راحتی می‌توان پسماندها را برای هر یک از

1. Marginal Propensity to Consume

دوره‌های مورد مطالعه محاسبه کرد. این مقادیر در ستون ۹ جدول منعکس شده است. مجموع مربعات پسماند ($\Sigma e_i^2 = ۱/۱۰۹۶$) نیز به راحتی قابل محاسبه است. مقدار مجموع مربعات پسماند به این معنی است که با روش حداقل مربعات معمولی مقادیر $\hat{\alpha}$ و $\hat{\beta}$ چنان تعیین شده است که مجموع مربعات پسماند کمترین مقدار ممکن را دارد و این کمترین مقدار برابر ۱/۱۰۹۶ میلیارد تومان شده است. برای پیش‌بینی مقدار مصرف در $t=۶$ کافی است که در مدل

$$\hat{C}_t = -۰/۳۶ + ۰/۶۷ Y_t ,$$

به جای Y_t، مقدار $Y_t=۱۲$ را قرار داده، در آن صورت $\hat{C}_t$ برابر ۷/۶۸ محاسبه می‌شود. همچنین می‌توان نشان داد که $\Sigma e_t X_t$ یا $\Sigma e_t x_t$ تقریباً برابر صفر است.

مثال ۲-۱ برای بررسی رابطهٔ تولید با ساعتهای نیروی کار از مدل رگرسیون زیر استفاده کرده‌ایم،

$$Q_i = \alpha + \beta L_i + U_i ,$$

که در آن Q_i تولید و L_i ساعت کار است.

برای تخمین α و β با استفاده از روش حداقل مربعات معمولی این مشاهدات را داریم

جدول ۳-۱

Q_i	۱۱	۱۰	۱۲	۶	۱۰	۷	۹	۱۰	۱۱	۱۰
L_i	۱۰	۷	۱۰	۵	۸	۸	۶	۷	۹	۱۰

$\hat{\alpha}_{OLS}$ و $\hat{\beta}_{OLS}$ را بدون استفاده از محاسبات انحراف از میانگین به دست آورید.

درمسألهٔ قبل با استفاده از انحراف ازمیانگین، یعنی $x_i = X_i - \bar{X}$ و $y_i = Y_i - \bar{Y}$، پارامترها را تخمین زدیم. می‌توان به طور مستقیم و به کمک مشاهدات اصلی نیز پارامترها را تخمین زد. می‌دانیم،

$$\sum x_i y_i = \sum (X_i - \bar{X})(Y_i - \bar{Y}) = \sum X_i Y_i - n \bar{X} \bar{Y} ,$$

$$\sum x_i^2 = \sum (X_i - \bar{X})^2 = \sum X_i^2 - n\,\bar{X}^2 .$$

بنابراین:

$$\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2},$$

$$= \frac{\sum X_i Y_i - n\,\bar{X}\,\bar{Y}}{\sum X_i^2 - n\,\bar{X}^2} .$$

با تغییر در نامگذاری متغیرها خواهیم داشت:

Y_i = تولید و X_i = ساعت کار

بنابراین مدل رگرسیون مفروض را می‌توان به صورت زیر نوشت،

$$Y_i = \alpha + \beta X_i + U_i .$$

با استفاده از محاسبات جدول ٤ـ١ کمیتهای لازم برای $\hat{\beta}_{OLS}$ را به دست می‌آوریم،

$$\sum x_i y_i = \sum X_i Y_i - n\,\bar{X}\,\bar{Y},$$

$$= ۷۸۹ - ۱۰(۸)(۹/۶) = ۷۸۹ - ۷۶۸ = ۲۱ ,$$

$$\sum x_i^2 = \sum X_i^2 - n\,\bar{X}^2 ,$$

$$= ۶۶۸ - ۱۰(۸)^2 = ۶۶۸ - ۶۴۰ = ۲۸ .$$

بنابراین

$$\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{۲۱}{۲۸} = ۰/۷۵ .$$

برای تخمین $\hat{\alpha}_{OLS}$ کافی است از رابطهٔ زیر استفاده کنیم،

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{X},$$

$$= ۹/۶ - ۰/۷۵(۸) = ۳/۶ .$$

جدول ٤ـ١

مشاهدات	Y_i	X_i	$X_i Y_i$	X_i
١	١١	١٠	١١٠	١٠٠
٢	١٠	٧	٧٠	٤٩
٣	١٢	١٠	١٢٠	١٠٠
٤	٦	٥	٣٠	٢٥
٥	١٠	٨	٨٠	٦٤
٦	٧	٨	٥٦	٦٤
٧	٩	٦	٥٤	٣٦
٨	١٠	٧	٧٠	٤٩
٩	١١	٩	٩٩	٨١
١٠	١٠	١٠	١٠٠	١٠٠
Σ	٩٦	٨٠	٧٨٩	٦٦٨

$$\bar{X} = \frac{\Sigma X_i}{n} , = \frac{٨٠}{١٠} = ٨ ,$$

$$\bar{Y} = \frac{\Sigma Y_i}{n} , = \frac{٩٦}{١٠} = ٩/٦ .$$

بدین‌ترتیب تخمین مدل رگرسیون مفروض عبارت است از

$$\hat{Q}_i = ٣/٦ + ٠/٧٥ L_i .$$

ملاحظه می‌شود که بهره‌وری نهایی نیروی کار[1] برابر ٠/٧٥ است. تخمین مدل رگرسیون بر این دلالت می‌کند که اگر ساعت ـ کار، صفر باشد؛ یعنی اساساً هیچ کاری انجام نشود، تولید ٣/٦ واحد خواهد بود. علت رسیدن به چنین نتیجهٔ بی‌معنا این است که مقدار متغیر درون‌زا را به ازای آن مقدار از متغیر برون‌زا حساب کرده‌ایم که فاصلهٔ بسیاری از قلمرو تغییرات متغیر برون‌زا دارد. در جدول ٤ـ١ ملاحظه می‌کنیم که قلمرو تغییرات متغیر برون‌زا بین ٥ تا ١٠ است. هر مقدار متغیر برون‌زا از این فاصله دورتر شود، محاسبه و پیش‌بینی مقدار متناظر متغیر برون‌زا معنای خود را از دست خواهد داد (اثبات ریاضی این نکته در مبحث پیش‌بینیها در ٢ـ٣ آمده است).

1. Marginal Productivity of Labour

۱-٤ ضریب تعیین

مدل رگرسیون ۱-۸ را یک بار دیگر می‌نویسیم،

$$Y_t = \alpha + \beta X_t + U_t .$$

می‌دانیم تخمین این مدل با استفاده از روش حداقل مربعات معمولی عبارت است از

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t .$$

سؤال این است که چگونه $\hat{Y}_t$ توانسته است تغییرات Y_t را توضیح دهد؟ در واقع می‌دانیم یکی از اهداف اساسی در تحلیلهای رگرسیونی این است که بتوان تغییرات متغیر درون‌زا را توضیح داد. متغیر درون‌زای ما در اینجا Y_t است. به کمک X_t و با استفاده از مدل ۱-۸ کوشش کرده‌ایم تغییرات Y_t را توضیح دهیم و سرانجام بعد از به دست آوردن $\hat{\alpha}$ و $\hat{\beta}$، مقادیر $\hat{Y}_t$ به دست آمده است. بنابراین می‌توان در این قسمت این سؤال را مطرح کرد که مسیری از تغییرات Y_t ـ که ما تخمین زده‌ایم ـ $(\hat{Y}_t)$، چقدر با مسیر اصلی تغییرات Y_t تناسب دارد. برای این منظور کمیتی را محاسبه می‌کنیم که بتواند درجهٔ برازندگی یا تناسب $\hat{Y}_t$ را نسبت به Y_t تعیین کند. به این کمیت «ضریب تعیین درجهٔ تناسب»[1] تخمین مدل رگرسیون گفته می‌شود. از این به بعد، برای سهولت، این کمیت را «ضریب تعیین» می‌خوانیم. مسلماً هر چه تخمین ما از تغییرات Y_t $(\hat{Y}_t)$ به Y_t نزدیکتر باشد، نشان‌دهندهٔ بهتر بودن تناسب تخمین با مقادیر واقعی است.

توجه به این نکته اهمیت دارد که وقتی از تغییرات یک متغیر صحبت می‌شود، باید بتوانیم کل تغییرات ملاحظه شده در یک دورهٔ زمانی را با یک کمیت واحد نشان دهیم؛ برای مثال، وقتی می‌گوییم تغییرات Y_t یا $\hat{Y}_t$ در فاصلهٔ زمانی $t = ۱, ۲, \ldots, n$ است، به دو کمیت واحد نیاز داریم تا از طریق مقایسه آنها با یکدیگر بتوان نسبت به تناسب دو مسیر از تغییرات Y_t و $\hat{Y}_t$ اظهارنظر کرد. راههای مختلفی برای بیان میزان تغییرات یک متغیر وجود دارد؛ یکی از بهترین روشها این است که ابتدا میانگین تغییرات

---

1. Coefficient of Determination of Goodness of Fit

را مبنای سنجش قرار دهیم و انحراف مشاهدات را از میانگین به دست آوریم، سپس مربع انحرافات را حساب کرده، مجموع آنها را به عنوان شاخص کل تغییرات معرفی کنیم. بدین ترتیب شاخص تغییرات Y_t و $\hat{Y}_t$ به ترتیب عبارتند از

$$\sum (Y_t - \bar{Y})^2 = \sum y_t^2, \qquad (۱\text{-}۲۷)$$

$$\sum (\hat{Y}_t - \bar{\hat{Y}})^2 = \sum \hat{y}_t^2. \qquad (۱\text{-}۲۸)$$

$\sum y_t^2$ در حقیقت کل تغییرات مشاهده شدهٔ Y_t است که معمولاً آن را با TSS یا TSSQ[1] نشان می‌دهند. $\sum \hat{y}_t^2$ نیز تغییرات توضیح داده شده توسط تخمین مدل رگرسیون است که می‌توان آن را با ESS یا ESSQ[2] نشان داد. نزدیکی هر چه بیشتر ESS به TSS مبین این واقعیت است که تناسب تخمین مدل رگرسیون با مدل واقعی رگرسیون بیشتر است. بنابراین، بهتر است کسری را در نظر بگیریم که صورت آن تغییرات توضیح داده شده و مخرج آن کل تغییرات باشد. این کسر همان ضریب تعیین خواهد بود،

$$\text{ضریب تعیین} = \frac{\text{تغییرات توضیح داده شده}}{\text{کل تغییرات}},$$

$$= \frac{\text{ESS}}{\text{TSS}} = \frac{\sum (\hat{Y}_t - \bar{\hat{Y}})^2}{\sum (Y_t - \bar{Y})^2} = \frac{\sum \hat{y}_t^2}{\sum y_t^2}. \qquad (۱\text{-}۲۹)$$

برای اینکه بتوان به فرمولهای ساده‌تری رسید، ابتدا دو نکته را مطرح می‌کنیم:

نکتهٔ اول: ثابت می‌کنیم که میانگین Y_t با میانگین $\hat{Y}_t$ برابر است یعنی $\bar{Y} = \bar{\hat{Y}}$.

برای اثبات کافی است معادلهٔ اول نرمال ۱-۲۰ را یک بار دیگر بنویسیم،

$$\sum Y_t = n\hat{\alpha} + \hat{\beta} \sum X_t.$$

دو طرف را بر n تقسیم می‌کنیم،

1. Total Sum of Squares
2. Explained Sum of Squares

$$\bar{Y} = \hat{\alpha} + \hat{\beta}\,\bar{X}\,.$$

سپس تخمین مدل رگرسیون، یعنی معادلهٔ ۱ـ۱۶ را نیز دوباره می‌نویسیم،

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta}\,X_t\,.$$

دو طرف رابطهٔ فوق را برای $t = ۱, ۲, \ldots, n$ جمع می‌کنیم،

$$\sum \hat{Y}_t = n\,\hat{\alpha} + \hat{\beta} \sum X_t\,,$$

و یا تقسیم دو طرف بر n خواهیم داشت

$$\bar{\hat{Y}} = \hat{\alpha} + \hat{\beta}\,\bar{X}\,,$$

پس نتیجه می‌گیریم که

$$\bar{Y} = \bar{\hat{Y}}\,. \qquad (۱\text{ـ}۳۰)$$

نکتهٔ دوم: نشان می‌دهیم که $\hat{y}_t = \hat{\beta}\,\hat{x}$ است. اثبات این نکته بسیار ساده است. مدل تخمین رگرسیون (معادلهٔ ۱ـ۱۶) را یک بار دیگر می‌نویسیم،

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta}\,X_t\,.$$

در نکتهٔ اول دیدیم که

$$\bar{Y} = \hat{\alpha} + \hat{\beta}\,\bar{X}\,.$$

کافی است دو رابطهٔ فوق را از یکدیگر کم کنیم،

$$\hat{Y}_t - \bar{Y} = \hat{\beta}\,(X_t - \bar{X})\,.$$

با استفاده از رابطهٔ ۱ـ۳۰ ، معادلهٔ فوق را می‌توان چنین نوشت،

$$\hat{Y}_t - \bar{\hat{Y}} = \hat{\beta}\,(X_t - \bar{X})\,,$$

یا

$$\hat{y}_t = \hat{\beta}\,x_t\,. \qquad (۱\text{ـ}۳۱)$$

بعد ازبیان این دو نکته به بحث ضریب تعیین برمی‌گردیم. رابطهٔ ۱ـ۲۹ را یک بار

دیگر می‌نویسیم،

$$\text{ضریب تعیین} = \frac{\Sigma\, \hat{y}_t^2}{\Sigma\, y_t^2}.$$

کافی است به جای $\hat{y}_t$ مقدار آن را از رابطهٔ ۱ـ۳۱ قرار دهیم،

$$\text{ضریب تعیین} = \frac{\Sigma\, (\hat{\beta}\, x_t)^2}{\Sigma\, y_t^2} = \frac{\hat{\beta}^2\, \Sigma\, x_t^2}{\Sigma\, y_t^2}.$$

با توجه به رابطهٔ ۱ـ۲۵ می‌دانیم

$$\hat{\beta} = \frac{\Sigma\, x_t\, y_t}{\Sigma\, x_t^2},$$

در نتیجه خواهیم داشت

$$\text{ضریب تعیین} = \frac{(\Sigma\, x_t\, y_t)^2 \,.\, \Sigma\, x_t^2}{(\Sigma\, x_t^2)^2 \,.\, \Sigma\, y_t^2},$$

$$= \frac{(\Sigma\, x_t\, y_t)^2}{\Sigma\, x_t^2 \Sigma\, y_t^2}.$$

اما در آمار دیده‌ایم که ضریب همبستگی[1] بین دو متغیر X_t و Y_t از رابطهٔ زیر به دست می‌آید،

$$r_{xy} = \frac{\Sigma\, x_t\, y_t}{\sqrt{\Sigma\, x_t^2 \,.\, \Sigma\, y_t^2}}. \qquad (۱ـ۳۲)$$

بنابراین، خواهیم داشت

$$\text{ضریب تعیین} = \frac{(\Sigma\, x_t\, y_t)^2}{\Sigma\, x_t^2 \,.\, \Sigma\, y_t^2} = r^2.$$

در واقع نشان دادیم که ضریب تعیین برابر مجذور ضریب همبستگی است.

---

1. Coefficient of Correlation

به عبارت دیگر، r^2 که معمولاً در کتابهای اقتصادسنجی برای نشان دادن ضریب تعیین به کار می‌رود یک علامتگذاری یا قرارداد نیست، بلکه ثابت می‌شود که ضریب تعیین برابر r^2 است، بنابراین

$$r^2 = \frac{\Sigma \hat{y}_t^2}{\Sigma y_t^2} \quad (۱\text{-}۳۳)$$

$$r^2 = \frac{(\Sigma x_t y_t)^2}{\Sigma x_t^2 . \Sigma y_t^2} . \quad (۱\text{-}۳۴)$$

می‌توان فرمول سومی نیز برای r^2 استخراج کرد که از نظر محاسباتی بسیار ساده‌تر است. فرمول ۱-۳۴ را یک بار دیگر می‌نویسیم و آن را تفکیک می‌کنیم،

$$r^2 = \frac{(\Sigma x_t y_t)^2}{\Sigma x_t^2 . \Sigma y_t^2} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} \cdot \frac{\Sigma x_t y_t}{\Sigma y_t^2} .$$

اما با توجه به رابطهٔ ۱-۲۵ می‌دانیم $\hat{\beta} = \frac{\Sigma (x_t y_t)}{\Sigma x_t^2}$ در نتیجه خواهیم داشت

$$r^2 = \hat{\beta}^2 \frac{\Sigma x_t^2}{\Sigma y_t^2} \quad \text{یا} \quad r^2 = \hat{\beta} \frac{\Sigma x_t y_t}{\Sigma y_t^2} . \quad (۱\text{-}۳۵)$$

برای رسیدن به فرمول دیگری برای r^2 ـ که جنبهٔ مفهومی بسیار خوبی دارد ـ ابتدا دو نکته را مطرح کرده، سپس یک قضیه را ثابت می‌کنیم.

نکتهٔ اول: با استفاده از نمودار ۱-۳ ملاحظه می‌شود که

$$Y_t = \hat{Y}_t + e_t , \quad (۱\text{-}۳۶)$$

یعنی مقدار مشاهده شده متغیر درون‌زا برابر است با مقدار تخمین زده شده به علاوهٔ مقدار پسماند.

این نکته را به بیان ساده‌تر می‌توان چنین نوشت.

پسماند + تخمین = مشاهده

از دو طرف رابطهٔ ۱ـ۳۶ مقدار $\bar{Y}$ را کم می‌کنیم،

$$(Y_t - \bar{Y}) = (\hat{Y}_t - \bar{Y}) + e_t .$$

با استفاده از رابطهٔ ۱ـ۳۰ و جایگزینی $\bar{\hat{Y}} = \bar{Y}$ در طرف راست آن خواهیم داشت

$$(Y_t - \bar{Y}) = (\hat{Y}_t - \bar{\hat{Y}}) + e_t ,$$

یا:

$$y_t = \hat{y}_t + e_t . \qquad (۱ـ۳۷)$$

در نتیجه نشان دادیم، رابطهٔ ۱ـ۳۶ که برای مشاهدات اصلی صادق است برای حالت تفاوت از میانگین نیز برقرار است.

نکتهٔ دوم: در معادلهٔ ۱ـ۲۳ دیدیم که می‌توان به کمک معادلات نرمال نشان داد که $\bar{e} = ۰$ است. در اینجا می‌خواهیم با روش دیگر به همین نتیجه برسیم. از تعریف میانگین e_t شروع می‌کنیم،

$$\bar{e} = \frac{\Sigma e_t}{n} .$$

با جایگزینی رابطهٔ ۱ـ۱۸ ، یعنی $e_t = Y_t - \hat{Y}_t$ خواهیم داشت

$$\bar{e} = \frac{\Sigma (Y_t - \hat{Y}_t)}{n} = \frac{\Sigma Y_t}{n} - \frac{\Sigma \hat{Y}_t}{n} ,$$

$$= \bar{Y} - \bar{\hat{Y}} .$$

با توجه به رابطهٔ ۱ـ۳۰ ، یعنی $\bar{Y} = \bar{\hat{Y}}$ به سهولت ملاحظه می‌شود که

$$\bar{e} = ۰ . \qquad (۱ـ۳۸)$$

قضیه در مدل رگرسیون $Y_t = \alpha + \beta X_t + U_t$ ، اگر پارامترها را با روش حداقل مربعات معمولی تخمین بزنیم، رابطهٔ زیر برقرار است.

تغییرات توضیح داده نشده + تغییرات توضیح داده شده = کل تغییرات متغیر درون‌زا

با توجه به اینکه مقادیر توضیح داده نشده، همان پسماندهاست؛ تغییرات توضیح داده

نشده نیز برابر تغییرات پسماندها خواهد بود. اگر تغییرات پسماندها را نیز به صورت مجموع مربعات انحراف از میانگین پسماندها تعریف کنیم، خواهیم داشت

$$\text{تغییرات پسماندها} = \Sigma e_t^2 = \Sigma (e_t - \bar{e})^2 .$$

رابطهٔ ۱ـ۳۸ را در رابطهٔ فوق جایگزین می‌کنیم،

$$\text{تغییرات پسماندها} = \Sigma e_t^2 . \qquad (۱ـ۳۹)$$

با توجه به تعریف $e_t = Y_t - \hat{Y}_t$، می‌توان Σe_t^2 را مجموع مربعات پسماند[1] (RSS) نامید. با توجه به رابطه‌های ۱ـ۲۷ ، ۱ـ۲۸ و ۱ـ۳۹ ،قضیهٔ فوق را می‌توان چنین نوشت،

$$\sum y_t^2 = \sum \hat{y}_t^2 + \sum e_t^2 , \qquad (۱ـ۴۰)$$

یا

$$TSS = ESS + RSS .$$

اثبات این قضیه بسیار ساده است. رابطهٔ ۱ـ۳۷ را دوباره می‌نویسیم،

$$y_t = \hat{y}_t + e_t .$$

دو طرف آن را مجذور کرده و برای تمام مشاهدات $t = ۱ , ۲ , \dots , n$ جمع می‌کنیم،

$$y_t^2 = (\hat{y}_t + e_t)^2 ,$$

$$\sum y_t^2 = \sum (\hat{y}_t^2 + e_t)^2 .$$

طرف راست را بسط می‌دهیم،

$$\sum y_t^2 = \sum \hat{y}_t^2 + \sum e_t^2 + ۲ \sum e_t \hat{y}_t .$$

برای اثبات رابطهٔ ۱ـ۴۰ کافی است ثابت کنیم $\Sigma e_t \hat{y}_t$ برابر صفر است. ابتدا معادلهٔ

---

1. Residual Sum of Squares

۱ـ۳۱ را دوباره می‌نویسیم،

$$\hat{y}_t = \hat{\beta}\, x_t .$$

دو طرف را در e_t ضرب کرده و نسبت به تمام مشاهدات جمع می‌کنیم،

$$e_t\, \hat{y}_t = \hat{\beta}\, x_t\, e_t ,$$

$$\sum e_t\, \hat{y}_t = \hat{\beta} \sum x_t\, e_t . \qquad (۱ـ۴۱)$$

با استفاده از رابطه‌های ۱ـ۳۷ و ۱ـ۳۱ داریم

$$e_t = y_t - \hat{y}_t ,$$

$$= y_t - \hat{\beta}\, x_t ,$$

و با جایگزینی در طرف راست رابطهٔ ۱ـ۴۱ خواهیم داشت

$$\sum e_t\, \hat{y}_t = \hat{\beta} \sum x_t\, (y_t - \hat{\beta}\, x_t) ,$$

$$= \hat{\beta}\, [\sum x_t\, y_t - \hat{\beta} \sum x_t^2] .$$

مقدار $\hat{\beta}$ از رابطهٔ ۱ـ۲۵ را در این رابطه جایگزین می‌کنیم،

$$\sum e_t\, \hat{y}_t = \hat{\beta}\, [\sum x_t\, y_t - \frac{\sum x_t\, y_t}{\sum x_t^2} \cdot \sum x_t^2] ,$$

در نتیجه:

$$\sum e_t\, \hat{y}_t = 0 . \qquad (۱ـ۴۲)$$

با استفاده از معادلهٔ ۱ـ۲۴ نیز به طور مستقیم می‌توان صفر بودن $\Sigma\, e_t\, \hat{y}_t$ را از معادلهٔ ۱ـ۴۱ نتیجه گرفت. بدیهی است با اثبات صفر شدن $\Sigma\, e_t\, \hat{y}_t$ ، رابطهٔ ۱ـ۴۰ ثابت می‌شود.[۱]

---

۱. در فصل دوم در مبحث آنالیز واریانس، به کمک معادلات ۲ـ۴۵ و ۲ـ۴۶ اثبات دیگری برای این قضیه ارائه شده است.

حال که این قضیهٔ بسیار مهم به اثبات رسید به استخراج یک فرمول دیگر برای r^2 می‌پردازیم. رابطهٔ ۱ـ۴۰ را به صورت زیر می‌نویسیم،

$$\sum \hat{y}_t^2 = \sum y_t^2 - \sum e_t^2 ,$$

و دو طرف آن را بر $\sum y_t^2$ تقسیم می‌کنیم،

$$\frac{\sum \hat{y}_t^2}{\sum y_t^2} = \frac{\sum y_t^2}{\sum y_t^2} - \frac{\sum e_t^2}{\sum y_t^2} .$$

با استفاده از رابطهٔ ۱ـ۳۳ خواهیم داشت

$$r^2 = 1 - \frac{\sum e_t^2}{\sum y_t^2} = 1 - \frac{RSS}{TSS} . \qquad (۱ـ۴۳)$$

فرمول ۱ـ۴۳ به دو دلیل بسیار مهم است:

اولاً، این فرمول می‌تواند قلمرو تغییرات r^2 را مشخص کند. در واقع حد بالای r^2 موقعی است که مطلقاً پسماندی در تخمین نداشته باشیم و آنچه تخمین زده‌ایم دقیقاً برابر مشاهدات باشد. در این حالت $\sum e_t^2$ برابر صفر و r^2 مساوی یک می‌شود. حد پایین r^2 هنگامی‌است‌که تخمین مدل رگرسیون ما مطلقاً قدرت توضیحی ندارد؛ بنابراین پسماندها به قدری بزرگند که برابر مشاهدات می‌باشند. در چنین حالتی $\sum e_t^2$ برابر $\sum y_t^2$ و مقدار r^2 برابر صفر خواهد بود. بنابراین قلمرو تغییرات r^2 را می‌توان به صورت زیر خلاصه کرد،

$$0 \le r^2 \le 1 . \qquad (۱ـ۴۴)$$

ثانیاً، با استفاده از این فرمول به راحتی می‌توان مجموع مربعات پسماند را محاسبه کرد. یادآوری می‌کنیم که فرمول ۱ـ۴۳ فرمول مناسبی برای محاسبه r^2 نیست؛ زیرا مستلزم محاسبهٔ جمله‌های پسماند است. از بین فرمولهایی که تا اینجا برای محاسبه r^2 مطرح کردیم، به نظر می‌رسد فرمول ۱ـ۳۵ مناسبتر باشد،

$$r^2 = \hat{\beta} \frac{\sum x_t y_t}{\sum y_t^2} .$$

دلیل این امر کاملاً روشن است. می‌دانیم r^2 را معمولاً بعد از تخمین پارامترهای یک مدل محاسبه می‌کنند؛ بنابراین در موقع محاسبهٔ r^2 مقدار $\hat{\beta}$ مشخص است. برای تخمین $\hat{\beta}$ نیز مقدار $\Sigma x_t y_t$ باید ضرورتاً محاسبه شود. بنابراین وقتی شروع به محاسبهٔ r^2 می‌کنیم، $\hat{\beta}$ و $\Sigma x_t y_t$ از قبل محاسبه شده است. فقط کافی است Σy_t^2 محاسبه شود تا r^2 به دست آید. حال اهمیت فرمول ۱ـ۴۳،

$$r^2 = 1 - \frac{\Sigma e_t^2}{\Sigma y_t^2},$$

در این است که اگر r^2 را از فرمول ۱ـ۳۵ حساب کنیم و در آن قرار دهیم، Σe_t^2 به راحتی محاسبه می‌شود. محاسبهٔ Σe_t^2 با روشهای دیگر، ممکن است مستلزم تقریب و محاسبات طولانی در مورد هر یک از پسماندها باشد؛ در حالی که با استفاده از فرمول فوق مقدار Σe_t^2 به طور مستقیم قابل محاسبه خواهد بود. در فصل دوم نه تنها فرمولهای محاسباتی دقیقتری مانند ۲ـ۴۶ را برای Σe_t^2 به دست خواهیم آورد؛ بلکه نشان خواهیم داد که محاسبه Σe_t^2 قدم اساسی در تخمین واریانس پارامترهاست.

برای تبیین نحوهٔ محاسبهٔ ضریب تعیین چند مثال می‌آوریم:

مثال ۳ـ۱ مدل مصرف. موضوع مثال ۱ـ۱ را در نظر بگیرید. با استفاده از جدول ۲ـ۱ ضریب تعیین r^2 را برای این مدل محاسبه کنید و مجموع مربعات پسماند، یعنی Σe_t^2 را به دست آورید.

با استفاده از فرمول ۱ـ۳۵ داریم

$$r^2 = \hat{\beta} \frac{\Sigma x_t y_t}{\Sigma y_t^2}.$$

در مثال ۱ـ۱ دیدیم که $\hat{\beta} = ۰/۶۷$ و $\Sigma x_t y_t = ۴۳$. بنابراین باید فقط Σy_t^2 محاسبه شود.

با استفاده از جدول ۲ـ۱ خواهیم داشت

$$\Sigma y_t^2 = ۹ + ۴ + ۱ + ۱۶ = ۳۰,$$

به این ترتیب

$$r^2 = \frac{۰/۶۷\,(۴۳)}{۳۰} = \frac{۲۸/۸۱}{۳۰} = ۰/۹۶\,,$$

یعنی متغیر درآمد، با توجه به مدل رگرسیون خطی مصرف، ۹۶ درصد تغییرات مصرف را توضیح داده است.

برای محاسبهٔ مجموع مربعات پسماند، می‌توان از فرمول ۱ـ۴۳ استفاده کرد،

$$r^2 = ۱ - \frac{\Sigma e_t^2}{\Sigma y_t^2}\,,$$

$$\frac{۰/۶۷\,(۴۳)}{۳۰} = ۱ - \frac{\Sigma e_t^2}{۳۰}, \quad ۲۸/۸۱ - ۳۰ = -\Sigma e_t^2\,,$$

$$\Sigma e_t^2 = ۱/۱۹\,.$$

مثال ۱ـ۴ مدل رابطهٔ تولید با ساعتهای نیروی کار، موضوع مثال ۱ـ۲ را ملاحظه کرده، با استفاده از جدول ۱ـ۴ ضریب تعیین r^2 را برای این مدل محاسبه کنید و مجموع مربعات پسماند Σe_i^2 را به دست آورید.

با توجه به معادلهٔ ۱ـ۳۵ می‌دانیم

$$r^2 = \hat{\beta}\,\frac{\Sigma x_i y_i}{\Sigma y_i^2}\,.$$

در مثال ۱ـ۲ دیدیم که $\hat{\beta} = ۰/۷۵$ و $\Sigma x_i y_i = ۲۱$؛ بنابراین باید فقط Σy_i^2 محاسبه شود. با توجه به اینکه محاسبات در جدول ۱ـ۴ بر اساس مشاهدات واقعی و نه انحراف از میانگین ارائه شده است، ابتدا ΣY_i^2 را محاسبه می‌کنیم،

$$\Sigma Y_i^2 = (۱۱)^2 + (۱۰)^2 + (۱۲)^2 + (۶)^2 + (۱۰)^2 + (۷)^2 + (۹)^2 + (۱۰)^2 + (۱۱)^2 + (۱۰)^2 = ۹۵۲\,.$$

Σy_i^2 به راحتی و با استفاده از ΣY_i^2 به دست می‌آید. می‌دانیم

$$\Sigma y_i^2 = \Sigma Y_i^2 - n\,(\bar{Y})^2\,,$$

$$= ۹۵۲ - ۱۰\,(۹/۶)^2 = ۹۵۲ - ۹۲۱/۶ = ۳۰/۴\,,$$

بدین ترتیب r^2 محاسبه می‌شود،

$$r^2 = \hat{\beta} \frac{\Sigma x_i y_i}{\Sigma y_i^2},$$

$$= \frac{۰/۷۵(۲۱)}{۳۰/٤} \simeq ۰/۵۲ ,$$

یعنی متغیر برون‌زا (ساعتهای کار)، فقط ۵۲ درصد از تغییرات متغیر درون‌زا (تولید) را توضیح داده است.

به عنوان تمرین، r^2 را از فرمول ضریب همبستگی نیز به دست می‌آوریم. با توجه به معادلهٔ ۱ـ۳۲ داریم

$$r_{x,y} = \frac{\Sigma x_i y_i}{\sqrt{\Sigma x_i^2 . \Sigma y_i^2}} ,$$

$$= \frac{۲۱}{\sqrt{۲۸ (۳۰/٤)}} = \frac{۲۱}{۲۹} = ۰/۷۲٤ ,$$

یا

$$r^2 = ۰/۵۲ .$$

برای محاسبهٔ Σe_i^2، از فرمول ۱ـ٤۳ استفاده می‌کنیم.

$$r^2 = ۱ - \frac{\Sigma e_i^2}{\Sigma y_i^2} ,$$

$$\frac{۰/۷۵ (۲۱)}{۳۰/٤} = ۱ - \frac{\Sigma e_i^2}{۳۰/٤} ,$$

$$\Sigma e_i^2 = ۱٤/٦۵ .$$

۱ـ۵ خطای معیار تخمین

در این قسمت به بررسی روش دیگری برای ارزیابی تخمین یک مدل رگرسیون

می‌پردازیم. ـ همان‌گونه که اشاره خواهد شد ـ با اینکه این روش جامعیت معیار ضریب تعیین را ندارد، در مواردی می‌تواند بسیار مفید باشد.

سؤالی را که در قسمت ۱.٤ مطرح کردیم، یک بار دیگر مطرح می‌کنیم. مدل رگرسیون مفروض، عبارت است از

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

که با روش حداقل مربعات معمولی آن را به صورت زیر تخمین زده‌ایم،

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t \ .$$

می‌دانیم مبنای روش حداقل مربعات معمولی این است که مجموع مربعات پسماند باید حداقل شود؛ پس این سؤال که آیا $\hat{Y}_t$ تخمین مناسبی از Y_t است یا خیر، می‌تواند به این سؤال تبدیل شود که آیا مقدار عددی مجموع مربعات پسماند رضایت‌بخش است یا خیر؟ به عبارت دیگر Σe_t^2 می‌تواند بالقوه اساس ارزیابی ما در تخمین یک مدل رگرسیون باشد. حال به بررسی جوانب این مسأله می‌پردازیم.

مهمترین اشکالی که به Σe_t^2 به منزلهٔ شاخص ارزیابی به تخمین مدل رگرسیون وارد می‌شود، این است که مقدار آن به طور مستقیم، تابعی از حجم مشاهدات است. با افزایش مشاهدات، Σe_t^2 به طور مرتب زیاد می‌شود و اساساً حدی ندارد. این اشکال را می‌توان بدین گونه رفع کرد که مجموع مربعات پسماند را بر تعداد مشاهدات تقسیم کرده و نتیجه را به عنوان شاخصی در ارزیابی خوبی تخمین مدل رگرسیون معرفی کنیم،

$$\text{معیار ارزیابی تخمین مدل رگرسیون} = \frac{\Sigma e_t^2}{n} = \frac{RSS}{n} \ ,$$

که در آن n حجم مشاهدات است.

با توجه به مباحث آماری، می‌توان گفت که اگر Σe_t^2 را به جای اینکه بر n تقسیم کنیم، بر درجات آزادی[1] آن تقسیم نماییم، معیار بهتری خواهیم داشت. این نکته را در

1. Degrees of Freedom

فصل دوم دربحث واریانس جملهٔ اختلال و درمعادلهٔ ۲۸ـ۲ ثابت خواهیم کرد. به هر حال می‌دانیم که در محاسبهٔ تغییرات توضیح داده نشده (RSS) یا Σe_t^2 دو درجهٔ آزادی را از دست می‌دهیم. یادآوری می‌شود که تعداد درجات آزادی که از دست می‌دهیم برابر با تعداد پارامترهایی است که باید برای محاسبهٔ مجموع مربعات پسماند تخمین بزنیم. واضح است که برای محاسبهٔ تغییرات توضیح داده نشده باید α و β تخمین زده شود؛ بنابراین می‌گوییم که در محاسبهٔ تغییرات توضیح داده نشده دو درجهٔ آزادی را از دست می‌دهیم. اگر درجات آزادی را ـ که از دست داده‌ایم ـ از تعداد کل مشاهدات کم کنیم، آنگاه به درجات آزادی باقیمانده یا به طور خلاصه درجات آزادی می‌رسیم. بنابراین معیار بهتر، عبارت خواهد بود از

$$\text{معیار ارزیابی تخمین مدل رگرسیون} = \frac{\Sigma e_t^2}{n-2} = \frac{RSS}{n-2} \quad .$$

با وجود این، هنوز می‌توان اشکال دیگری به این معیار وارد کرد. با این معیار می‌خواهیم توانایی $\hat{Y}_t$ در تفسیر تغییرات Y_t را اندازه‌گیری کنیم، در حالی که بُعد[1] این معیار با بُعد Y_t متفاوت است؛ به عبارت دیگر، $\frac{\Sigma e_t^2}{n-2}$ از رتبهٔ دو است در حالی که Y_t از رتبهٔ یک است. بدیهی است این اشکال را نیز می‌توان به سهولت رفع کرد. کافی است از $\frac{\Sigma e_t^2}{n-2}$ جذر بگیریم تا همرتبهٔ Y_t یا $\hat{Y}_t$ شود. به معیاری که به این ترتیب به دست می‌آید «خطای معیار تخمین»[2] می‌گویند و آن را «SEE» نشان می‌دهند،

$$SEE = \sqrt{\frac{\Sigma e_t^2}{n-2}} \quad . \qquad (۱ـ۴۵)$$

می‌توان برای خطای معیار تخمین (SEE) فرمول ساده‌تری نیز به دست آورد،

---

1. Dimension
2. Standard Error of Estimation

برای توضیح بیشتر این فرمول می‌توان به بحث واریانس جملهٔ اختلال در قسمت ۲ـ۲ و فرمول ۲۸ـ۲ مراجعه کرد.

به گونه‌ای که با داشتن مشاهدات اصلی قابل محاسبه باشد. با استفاده از معادله‌های ۱ـ۳۷ و ۱ـ۳۱ و جایگزینی آنها در فرمول ۱ـ۴۵ داریم

$$SEE = \sqrt{\frac{\Sigma (y_i - \hat{y}_i)^2}{n-2}} ,$$

$$= \sqrt{\frac{\Sigma (y_i - \hat{\beta} x_i)^2}{n-2}} .$$

می‌دانیم

$$\Sigma (y_i - \hat{\beta} x_i)^2 = \Sigma y_i^2 - \frac{(\Sigma x_i y_i)^2}{\Sigma x_i^2} ,$$

زیرا کافی است طرف چپ رابطهٔ فوق را بسط داده، به جای $\hat{\beta}$ مقدار آن را از فرمول ۱ـ۲۵ قرار دهیم،

$$\Sigma (y_i - \hat{\beta} x_i)^2 = \Sigma y_i^2 + \hat{\beta} \Sigma x_i^2 - 2\hat{\beta} \Sigma x_i y_i ,$$

$$= \Sigma y_i^2 + \frac{(\Sigma x_i y_i)^2}{(\Sigma x_i^2)^2} \cdot \Sigma x_i^2 - 2 \frac{(\Sigma x_i y_i)^2}{\Sigma x_i^2} ,$$

$$= \Sigma y_i^2 - \frac{(\Sigma x_i y_i)^2}{\Sigma x_i^2} .$$

بدین ترتیب خطای معیار تخمین (SEE) عبارت خواهد بود از

$$SEE = \sqrt{\left[\Sigma y_i^2 - \frac{(\Sigma x_i y_i)^2}{\Sigma x_i^2}\right] / (n-2)} . \qquad (۱ـ۴۶)$$

قبل از اینکه به نکته‌های مثبت و منفی خطای معیار تخمین (SEE) بپردازیم، برای تبیین بیشتر فرمول ۱ـ۴۶ مثالی می‌زنیم.

مثال ۱ـ۵ مدل رابطهٔ تولید با ساعتهای نیروی کار. موضوع مثال ۱ـ۲ را در نظر بگیرید. با استفاده از جدول ۱ـ۴ خطای معیار تخمین را حساب کنید.

می‌دانیم $\Sigma y_i^2 = ۳۰/٤$ ، $\Sigma x_i^2 = ۲۸$ و $\Sigma x_i y_i = ۲۱$ است، بنابراین با توجه به $n = ۱۰$ و با استفاده از فرمول ٤٦ـ۱ خواهیم داشت

$$SEE = \sqrt{\left[۳۰/٤ - \frac{(۲۱)^2}{۲۸}\right] / (۱۰ - ۲)} ,$$

$$= \sqrt{۱٤/٦٥/۸} = ۱/۳٥۳ .$$

با استفاده از فرمول ٤٥ـ۱ نیز می‌توان خطای معیار تخمین (SEE) را حساب کرد. در مثال ٤ـ۱ دیدیم که $\Sigma e_i^2 = ۱٤/٦٥$؛ بنابراین مقدار خطای معیار تخمین به صورت زیر محاسبه می‌شود،

$$SEE = \sqrt{\frac{۱٤/٦٥}{۱۰ - ۲}} = ۱/۳٥۳ .$$

حال که با کیفیت محاسبهٔ خطای معیار تخمین آشنا شدیم، به موارد مثبت و منفی آن اشاره می‌کنیم. برای مواردی که می‌خواهیم حالتهای مختلف یک مدل رگرسیون مفروض را مقایسه کنیم، خطای معیار تخمین می‌تواند بسیار مفید باشد؛ برای مثال، اگر بخواهیم به این سؤال پاسخ دهیم که آیا در مدل تولید و نیروی کار مذکور می‌توان به جای ۱۰ مشاهده فقط از ۸ مشاهده استفاده کرد یا خیر، کافی است خطای معیار تخمین را یک بار برای ۱۰ مشاهده و بار دیگر برای ۸ مشاهده محاسبه می‌کنیم. اگر نتایج به دست آمده تفاوت چندانی با یکدیگر نداشته باشد، می‌توان نتیجه گرفت که ضرورتی برای ۱۰ مشاهده نیست و ۸ مشاهده کافی است. بدیهی است با معیار r^2 نیز می‌توان به این سؤال پاسخ داد.

اما اشکال اساسی روش خطای معیار تخمین این است که نمی‌تواند معیار مناسبی برای مقایسهٔ تخمینهای دو یا چند مدل رگرسیون باشد؛ زیرا خطای معیار تخمین تابعی از مقیاس اندازه‌گیری متغیر درون‌زاست. به عبارت دیگر، اگر بخواهیم برای چند متغیر درون‌زای متفاوت، مدلهای رگرسیون بسازیم، برای مثال، برای تابع مصرف و شاخص قیمتها، بر همین اساس، جمله‌های پسماند، مقیاسهای متفاوتی خواهد داشت و در نتیجه نمی‌توان به طور مستقیم آنها را با یکدیگر مقایسه کرد. در نتیجه به منظور رسیدن به

معیاری برای ارزیابی تخمین یک مدل رگرسیون باید سعی کنیم معیار ما تابعی از مقیاس اندازه گیری متغیرها نباشد و دیدیم که معیار ضریب تعیین (r^2) این خصوصیت مطلوب را دارد. بنابراین r^2 نسبت به خطای معیار تخمین مرجح است، هر چند در مواردی نیز می‌توان از خطای معیار تخمین استفاده کرد. کاربرد اساسی خطای معیار تخمین در محاسبهٔ واریانس تخمین پارامترها و آزمون فرضیه‌ها موضوع فصل آینده است.

مسائل فصل اول

۱ـ۱ مدل رگرسیون خطی

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

مفروض است. U_t جملهٔ اختلال مدل می‌باشد.

۱. معمولاً در اقتصادسنجی چه دلایلی برای ضرورت وجود U_t در یک مدل رگرسیون مطرح می‌شود؟

۲. اگر بخواهیم α و β را با روش حداقل مربعات معمولی تخمین بزنیم، چه فرضهایی برای U_t لازم است؟

۳. فرض کنید X_t برای تمام مشاهدات $t = ۱, ۲, ۳, \ldots, n$، مقدار ثابتی دارد. تأثیر این فرض را بر تخمین پارامترهای α و β ارزیابی کنید.

٤. فرض $E(U_t) = ۰$ را با استفاده از نمودار، دقیقاً توضیح دهید.

٥. نشان دهید که واریانس متغیر درون‌زا برابر واریانس جملهٔ اختلال مدل است.

٦. امید ریاضی متغیر درون‌زا را به دست آورید.

۷. با استفاده از نمودار، مفاهیم جملهٔ اختلال مدل و جملهٔ پسماند را دقیقاً توضیح دهید و با یکدیگر مقایسه کنید. می‌دانیم

$$U_t = Y_t - E(Y_t) \quad , \quad e_t = Y_t - \hat{Y}_t \ ,$$

آیا می‌توان e_t را به صورت تخمینی از U_t به کار برد. چرا؟

۸. با استفاده از نمودار، فرض واریانس همسانی U_t را توضیح دهید.

۹. با استفاده از نمودار، فرض استقلال X_t از U_t را توضیح دهید.

۱۰. برای هر یک از X_t و Y_t، ٥ اصطلاح رایج در اقتصادسنجی را نام ببرید.

۱ـ۲ مدل رگرسیون خطی زیر مفروض است،

$$Y_t = \alpha + \beta X_t + U_t \ .$$

۱. نشان دهید $\bar{Y} = \hat{\bar{Y}}$.

۲. نشان دهید $\hat{y}_t = \hat{\beta}\, x_t$، که y و x به ترتیب عبارتند از $y_t = Y_t - \bar{Y}$ و $x_t = X_t - \bar{X}$.

۳. نشان دهید $\bar{e} = 0$.

۴. ثابت کنید $\Sigma y_t^2 = \Sigma \hat{y}_t^2 + \Sigma e_t^2$ است.

۵. ثابت کنید $\Sigma e_t \hat{y}_t = 0$.

که درآن، $e_t = y_t - \hat{y}_t$.

۶. ثابت کنید $\Sigma e_t x_t = 0$.

۷. چرا ضریب تعیین r^2 بین صفر و یک است؟

۸. چرا r^2 به خطای معیار تخمین (SEE) برتری دارد؟

۱ـ۳ مدل رگرسیون

$$y_t = \beta x_t + u_t \ ,$$

مفروض است. x_t و y_t به ترتیب انحراف X_t و Y_t از میانگین آنهاست. با استفاده از روش حداقل مربعات معمولی (OLS) پارامتر β را به طور مستقیم از این مدل تخمین بزنید. آیا $\hat{\beta}$ به دست آمده از این مدل با $\hat{\beta}$ به دست آمده از مدل

$$Y_t = \alpha + \beta X_t + U_t$$

متفاوت است؟ چرا؟

۱ـ۴ مدل رگرسیون

$$CF_i = \alpha + \beta DI_i + U_i \ ,$$

مفروض است. CF_i و DI_i به ترتیب هزینهٔ مصرف مواد غذایی و درآمد قابل تصرف خانوار i است. $i = ۱, ۲, \ldots, ۱۰$. مشاهدات زیر را داریم

DI_i	۲۰	۳۰	۳۳	۴۰	۱۵	۱۳	۲۶	۳۸	۳۵	۴۳
CF_i	۷	۹	۸	۱۱	۵	۴	۸	۱۰	۹	۱۰

۱. با روش حداقل مربعات معمولی پارامترهای α و β را تخمین بزنید.

۲. $e_i = CF_i - \hat{CF}_i$ را برای تمام مقادیر i حساب کنید.

۳. Σe_i^2 را به طور مستقیم محاسبه کنید.

۴. ضریب تعیین r^2 را حساب کنید.

۵. با استفاده از r^2 به دست آمده در بند ۴، مقدار Σe_i^2 را حساب کنید.

۶. خطای معیار تخمین را به دست آورید.

۷. نشان دهید که $\Sigma cf_i^2 = \Sigma \hat{cf}_i^2 + \Sigma e_i^2$

۸. اگر خانوار یازدهم، درآمد قابل تصرفی برابر ۵۰ هزار تومان داشته باشد، مصرف مواد غذایی آن خانوار را تخمین بزنید.

۹. اگر DI_i را برابر صفر فرض کنیم، مقدار مصرف چقدر خواهد بود؟ این نتیجه را چگونه توجیه می‌کنید؟

۵-۱ مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t \quad .$$

مفروض است. برای تخمین پارامترهای این مدل، ۱۶ مشاهده روی X و Y داشته و این کمیتها را محاسبه کرده‌ایم،

$$\Sigma Y_t^2 = ۵۲۶ \quad , \qquad \Sigma X_t^2 = ۶۵۷ \quad , \qquad \Sigma X_t Y_t = ۴۹۲ \quad .$$

$$\Sigma Y_t = ۶۴ \quad , \qquad \Sigma X_t = ۹۶ \quad ,$$

۱. با روش حداقل مربعات معمولی پارامتر β را تخمین بزنید.

۲. r^2 را محاسبه کنید.

۳. خطای معیار تخمین را به دست آورید.

۱-۶ برای تخمین یک مدل رگرسیون خطی بین دو متغیر X و Y ـ که Y درون‌زاست ـ تعداد ۲۷ مشاهده روی X و Y انجام داده‌ایم و این نتایج را به دست آورده‌ایم،

$$\bar{X} = ۱۰۰ , \quad \sum_{i=۱}^{۲۷} (X_i - \bar{X})^۲ = ۱۰۰ , \quad \sum_{i=۱}^{۲۷} (X_i - \bar{X})(Y_i - \bar{Y}) = ۲۰۰ .$$

$$\bar{Y} = ۱۵۰ , \quad \sum_{i=۱}^{۲۷} (Y_i - \bar{Y}) = ۵۰۰ ,$$

۱. با روش حداقل مربعات معمولی پارامترهای α و β را تخمین بزنید.

۲. $r^۲$ را محاسبه کنید.

۳. $\Sigma e_i^۲$ را به دست آورید.

٤. خطای معیار تخمین را حساب کنید.

۱-۷ مدل زیر مفروض است،

$$Y_t = a_۱ + b_۱ X_t + U_t . \qquad (۱)$$

۱. فرض می‌شود جملهٔ اختلال U_t، تابعی از متغیر توضیحی X_t است، یعنی

$$U_t = a_۲ + b_۲ X_t + \varepsilon_t , \qquad (۲)$$

که ε_t جملهٔ اختلال مدل (۲) بوده است و تمام فرضهای کلاسیک در مورد آن صادق است. همچنین فرض می‌کنیم که $b_۲ < ۰$ است. نشان دهید که پارامتر $b_۱$ در مدل (۱) تأثیر X_t بر Y_t را کمتر از مقدار واقعی آن منعکس می‌کند.

۲. اگر معادلهٔ (۲) را به صورت زیر بنویسیم،

$$U_t = a_۲ + b_۲ X_t^۲ + \varepsilon_t , \qquad (۳)$$

در این صورت آیا همهٔ فرضهای کلاسیک جملهٔ اختلال در مدل (۱) به قوّت خود باقی می‌ماند؟

۱-۸ مدل رگرسیون

$$Y_t = a + b X_t + U_t . \qquad (۱)$$

را در نظر می‌گیریم. فرض کنید در اندازه گیری متغیر توضیحی X_t خطای اندازه گیری داریم؛ یعنی به جای اینکه به طور مستقیم X_t را اندازه گیری کنیم، مقدار X_t^* را محاسبه کرده‌ایم، به گونه‌ای که

$$X_t^* = X_t + \varepsilon_t \ , \qquad (۲)$$

ε_t جملهٔ اختلال مدل (۲) و مستقل از X_t است و تمام فرضهای کلاسیک نیز در مورد آن صادق است. اگر فرض کنیم ε_t و U_t از یکدیگر مستقل باشند:

۱. نشان دهید که X_t^* از U_t مستقل است.

۲. یک مدل رگرسیون بسازید که Y_t را با X_t^* مرتبط سازد.

۳. در این مدل رگرسیون که ساخته‌اید آیا همهٔ فرضهای کلاسیک در مورد جملهٔ اختلال آن صادق است؟

۱-۹ مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t \ .$$

مفروض است. نشان دهید که هیچگونه تضمینی وجود ندارد که تخمینهای به دست آمده با روش حداقل مربعات ($\hat{\alpha}$ و $\hat{\beta}$)، دقیقاً با مقادیر واقعی پارامترها (α و β) برابر باشد.

۱-۱۰ درمدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

۱. نشان دهید که اگر تمام مقادیر X_t با یکدیگر برابر باشند تخمین $\hat{\beta}$ ممکن نیست (بند ۳ مسأله ۱-۱).

۲. آیا می‌توان یک مدل رگرسیون داشت که علی‌رغم ثابت بودن مقادیر متغیر توضیحی، تخمین β ممکن باشد؟

۱-۱۱ برای مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

۱. نشان دهید که مجموع پسماندها، برابر صفر است؛ یعنی $\Sigma e_t = 0$.

۲. آیا به نظر شما می‌توان گفت $\Sigma U_t = 0$؟

۱-۱۲ مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t \ .$$

مفروض است. فرض کنید مشاهدات مربوط به Y_t را به دو قسمت تقسیم کرده‌ایم: Y_{1t} و Y_{2t} و بدیهی است که $Y_t = Y_{1t} + Y_{2t}$. به کمک Y_{1t} و Y_{2t} دو مدل رگرسیون به شرح زیر می‌سازیم،

$$Y_{1t} = \alpha_1 + \beta_1 X_t + U_{1t} \ ,$$

$$Y_{2t} = \alpha_2 + \beta_2 X_t + U_{2t} \ .$$

نشان دهید که $\hat{\alpha} = \hat{\alpha}_1 + \hat{\alpha}_2$ و $\hat{\beta} = \hat{\beta}_1 + \hat{\beta}_2$.

۱-۱۳ متغیر تصادفی X_t مفروض است. این متغیر را به k قسمت تقسیم می‌کنیم، به گونه‌ای که برای تمام مقادیر t داشته باشیم

$$\sum X_{it} = X_t \quad , \quad i = 1, 2, 3, \ldots, k \ .$$

X_{it} را تابعی از X_t گرفته و مدلهای زیر را می‌سازیم.

$$X_{1t} = \alpha_1 + \beta_1 X_t + U_{1t} \ ,$$

$$X_{2t} = \alpha_2 + \beta_2 X_t + U_{2t} \ ,$$

$$\vdots$$

$$X_{kt} = \alpha_k + \beta_k X_{kt} + U_{kt} \ ,$$

که در آن $t = 1, 2, \ldots, n$. نشان دهید $\sum_{i=1}^{k} \hat{\alpha}_i = 0$ و $\sum_{i=1}^{k} \hat{\beta}_i = 1$.

۱-۱۴ مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t \ .$$

مفروض است. این مدل را تخمین زده و $\hat{Y}_t$ را به دست می‌آوریم. نشان دهید که اگر مدل رگرسیون Y_t روی $\hat{Y}_t$ را بسازیم،

$$Y_t = a + b\,\hat{Y}_t + \varepsilon_t \ ,$$

آنگاه $\hat{a} = ۰$، $\hat{b} = ۱$.

۱ـ۱۵ مدل رگرسیون

$$Y_t = \beta\, X_t + U_t \ ,$$

مفروض است. آیا تخمین این مدل با روش حداقل مربعات معمولی ($\hat{Y} = \hat{\beta}\, X_t$) از نقطهٔ ($\bar{Y}$ و $\bar{X}$) می‌گذرد؟

۱ـ۱۶ می‌دانیم مدل واقعی برای تخمین پارامتر β عبارت است از:

$$Y_t = \beta\, X_t + U_t \ . \qquad (۱)$$

حال فرض کنید β را از این مدل تخمین می‌زنیم،

$$Y_t = \alpha + \beta\, X_t + U_t \ .$$

نشان دهید هنگامی تخمین β از مدل (۲) دقیقاً به همان نتیجهٔ تخمین β از مدل (۱) می‌رسد که مدل (۲) را مقید تخمین بزنیم؛ یعنی شرط کرده باشیم که تخمین ضریب ثابت در مدل دوم صفر است.

۱ـ۱۷ مدل رگرسیون

$$Y_t = \beta\, X_t + U_t \ .$$

مفروض است. بدون استفاده از مشتق‌گیری نشان دهید که

$$\hat{\beta} = \frac{\Sigma\, X_t\, Y_t}{\Sigma\, X_t^{۲}} \ .$$

۱ـ۱۸ مدل رگرسیون

$$Y_t = \alpha + \beta\, X_t + U_t \ .$$

مفروض است. مشاهدات X_t و Y_t را برحسب انحراف از مقادیر معلوم و اختیاری X^* و Y^* می‌نویسم،

$$Y_t^* = Y_t - Y^* \quad , \quad X_t^* = X_t - X^*$$

حال مدل

$$Y_t^* = a + b\, X_t^* + U_t \qquad (۲)$$

را تخمین می‌زنیم. به نظر شما چه رابطه‌ای بین تخمین پارامترها در مدل اولیه و مدل جدید وجود دارد؟

۱-۱۹ در مدل رگرسیون

$$Y_t = \alpha + \beta\, X_t + U_t \ ,$$

نشان دهید که ضریب تعیین r^2 را، می‌توان به کمک هر یک از فرمولهای زیر محاسبه کرد.

۱) $r^2 = \dfrac{\Sigma\,(x_t\, y_t)^2}{\Sigma\, x_t^2\, \Sigma\, y_t^2}$,

۲) $r^2 = \hat{\beta}\, \dfrac{\Sigma\, x_t\, y_t}{\Sigma\, y_t^2}$,

۳) $r^2 = \dfrac{(\Sigma\, y_t\, \hat{y}_t)^2}{\Sigma\, y_t^2\, \Sigma\, \hat{y}_t^2}$,

۴) $r^2 = 1 - \dfrac{\Sigma\, e_t^2}{\Sigma\, y_t^2}$.

۱-۲۰ در مدل رگرسیون

$$Y_t = \beta\, X_t + U_t \ ,$$

نشان دهید که ضریب تعیین r^2 را می‌توان به کمک هر یک از فرمولهای زیر محاسبه کرد:

۱) $r^2 = \dfrac{(\Sigma\, X_t\, Y_t)^2}{\Sigma\, X_t^2\, \Sigma\, Y_t^2}$,

۲) $r^2 = \dfrac{\Sigma\,(Y_t\, \hat{Y}_t)^2}{\Sigma\, Y_t^2 \Sigma\, Y_t^2}$,

۳) $r^2 = 1 - \dfrac{\Sigma\, e_t^2}{\Sigma\, Y_t^2}$.

۱-۲۱ در مدل رگرسیون

$$Y_t = \beta\, X_t + U_t \ ,$$

نشان دهید که اگر بخواهیم r^2 را با استفاده از فرمول

$$r^2 = 1 - \frac{RSS}{TSS} ,$$

محاسبه کنیم که در آن $RSS = \Sigma e_t^2$ و $TSS = \Sigma y_t^2$، چه بسا r^2 ممکن است منفی شود.

۲۲ـ۱ مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t . \qquad (1)$$

مفروض است. از دو طرف مدل (۱) مقدار X_t را کم می‌کنیم:

$$Y_t - X_t = \alpha + (\beta - 1) X_t + U_t . \qquad (2)$$

۱. نشان دهید که تخمین هر یک از معادلات (۱) و (۲) برای $\hat{\alpha}$ و $\hat{\beta}$ به نتایج یکسانی می‌رسد.

۲. آیا تغییرات توضیح داده نشده (RSS) و r^2 هر یک از دو مدل فوق نیز با یکدیگر برابر است؟

حل مسائل فصل اول

مسائل ۱ـ۱ تا ۱ـ۳ در واقع سؤالاتی است که از فصل اول کتاب، انتخاب شده و پاسخ آنها در متن کتاب آمده است.

مسائل ۱ـ٤ تا ۱ـ٦ دقیقاً مشابه مثالهایی است که در متن فصل اول کتاب حل شده است و اساساً شامل نکته‌های جدید نیست.

۱ـ۷ برای پاسخ به قسمت اول این مسأله کافی است معادلهٔ (۲) را در مدل رگرسیون مفروض قرار دهیم. خواهیم داشت

$$Y_t = (a_1 + a_2) + (b_1 + b_2) X_t + \varepsilon_t \ .$$

ملاحظه می‌شود که تأثیر X_t بر Y_t در مدل فوق ـ که در واقع صورت صحیح مدل (۱) است ـ با $(b_1 + b_2)$ برابر است. چون فرض بر مثبت بودن b_2 است؛ بنابراین

$$b_1 < b_1 + b_2 \ ,$$

و در نتیجه b_1 تأثیر X_t بر Y_t در مدل (۱) را کمتر از میزان واقعی آن، یعنی $(b_1 + b_2)$ منعکس می‌سازد.

پاسخ قسمت دوم این مسأله منفی است؛ یعنی با فرض معادلهٔ (۳) ، بیشتر فرضهای کلاسیک جملهٔ اختلال در معادلهٔ (۱) برقرار نخواهد بود. برای تبیین این نکته معادلهٔ (۳) را در معادلهٔ (۱) قرار می‌دهیم،

$$Y_t = (a_1 + a_2) + b_1 X_t + (b_2 X_t^2 + \varepsilon_t) \ .$$

جملهٔ اختلال در مدل فوق که صورت واقعی مدل رگرسیون (۱) است برابر با

$$w_t = b_2 X_t^2 + \varepsilon_t \ .$$

است. در اینجا فرضهای کلاسیک در مورد w_t را بررسی می‌کنیم. ملاحظه می‌شود که

امید ریاضی w_t برابر صفر نیست؛ زیرا

$$E(w_t) = b_2 E(X_t^2) + E(\varepsilon_t) ,$$

است. با فرض $E(\varepsilon_t) = 0$، خواهیم داشت

$$E(w_t) = b_2 X_t^2 \neq 0 .$$

چون مقادیر مختلف X_t در زمان از یکدیگر مستقل نیست؛ بنابراین مقادیر w_t نیز در دوره‌های زمانی مختلف از هم مستقل نبوده و در نتیجه فرض عدم خودهمبستگی برقرار نیست

$$\text{Cov}(w_t, w_s) \neq 0 .$$

با توجه به اینکه X_t و X_t^2 از یکدیگر مستقل نیستند، w_t که تابعی از X_t^2 است تابعی از X_t نیز خواهد بود و می‌دانیم X_t متغیر توضیحی در مدل رگرسیون است؛ بنابراین فرض استقلال متغیر توضیحی از جملهٔ اختلال نیز باطل می‌شود،

$$E(w_t X_t) \neq X_t E(w_t) \neq 0 .$$

همچنین چون w_t تابعی از X_t است و X_t در زمان تغییر می‌کند؛ در نتیجه واریانس w_t نیز در طی زمان ثابت نخواهد بود؛ در نتیجه به جای فرض کلاسیک واریانس همسانی، باید فرض واریانس ناهمسانی را بپذیریم. تنها جزئی که از فرضهای کلاسیک تغییر نخواهد کرد فرض توزیع نرمال جملهٔ اختلال است.

۸ـ۱ ۱. چون X_t^* تابعی از ε_t و ε_t بنا بر فرض از U_t مستقل است؛ بنابراین X_t^* از U_t مستقل خواهد بود.

۲. با استفاده از معادلهٔ (۲) داریم

$$X_t = X_t^* - \varepsilon_t ,$$

که با جایگزینی در معادلهٔ (۱) خواهیم داشت

$$Y_t = a + b X_t^* + (U_t - b \varepsilon_t) ,$$

یا

$$Y_t = a + b X_t^* + U_t^* .$$

۳. خیر؛ زیرا در این مدل جدید، متغیر توضیحی از جملهٔ خطای مدل مستقل نیست. صحت این امر با ملاحظهٔ معادلهٔ (۲) واضح است. X_t^* تابعی از ε_t است و چون ε_t قسمتی از جملهٔ اختلال مدل جدید را تشکیل می‌دهد؛ یعنی $U_t^* = U_t - b\ \varepsilon_t$، پس X_t^* از U_t^* مستقل نخواهد بود. نتیجه می‌گیریم که فرض غیر تصادفی بودن متغیر توضیحی دیگر برقرار نیست.

۱-۹ ابتدا دربارهٔ $\hat{\beta}$ و β بحث می‌کنیم. می‌دانیم $\hat{\beta}_{OLS}$ با تخمین‌زننده زیر به دست می‌آید،

$$\hat{\beta}_{OLS} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} .$$

با جایگزینی $y_t = \beta x_t + U_t$، خواهیم داشت

$$\hat{\beta}_{OLS} = \frac{\Sigma x_t (\beta x_t + U_t)}{\Sigma x_t^2} ,$$

$$= \beta + \frac{\Sigma x_t U_t}{\Sigma x_t^2} .$$

برای اینکه $\hat{\beta} = \beta$ باشد، باید $\frac{\Sigma x_t U_t}{\Sigma x_t^2}$ ضرورتاً برابر صفر شود. اما می‌دانیم Σx_t^2 همواره مثبت است، بجز موردی که تمام مقادیر X_t با یکدیگر برابر باشند. بدیهی است در چنین صورتی محاسبهٔ $\hat{\beta}$ ممکن نیست؛ چون مخرج کسر $\hat{\beta}_{OLS}$ برابر صفر است؛ بنابراین برای $\hat{\beta} = \beta$ کافی است $\Sigma x_t U_t$ برابر صفر شود. با توجه به اینکه مقادیر U_t، هیچگاه شناخته نخواهند شد، نمی‌توان صحت رابطهٔ $\Sigma x_t U_t = 0$ را ارزیابی کرد؛ بدین ترتیب، هیچ حالتی وجود نخواهد داشت که مطمئن باشیم $\hat{\beta} = \beta$.

در مورد امکان $\hat{\alpha} = \alpha$، ابتدا می‌گوییم

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X} .$$

با توجه به رابطهٔ

$$\bar{Y} = \alpha + \beta \bar{X} + \bar{U} ,$$

خواهیم داشت

$$\hat{\alpha} = \alpha + \beta \bar{X} + \bar{U} - \hat{\beta} \bar{X} \ ,$$

$$= \alpha + \beta \bar{X} + \bar{U} - \left(\beta + \frac{\Sigma x_t U_t}{\Sigma x_t^2}\right) \bar{X} \ ,$$

$$= \alpha + \bar{U} - \frac{\Sigma x_t U_t}{\Sigma x_t^2} \bar{X} \ .$$

برای اینکه $\hat{\alpha}$ برابر α شود، نه تنها باید شرط $\Sigma x_t U_t = 0$ برقرار باشد، بلکه $\bar{U}$ نیز باید برابر صفر شود. دوباره این استدلال تکرار می‌شود که چون مقادیر U_t بنا بر تعریف، نامعلوم است، هیچ راهی وجود ندارد که بتوان، برای مجموعهٔ معینی از مشاهدات داده‌ای، نشان داد که آیا شرایط $\Sigma x_t U_t = 0$ یا $\bar{U} = 0$ برقرار است یا خیر.

۱۰ـ۱ ۱. این سؤال دقیقاً موضوع بند ۳ مسألهٔ (۱ـ۱) است. می‌دانیم

$$\hat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} \ ,$$

بنابراین، اگر مقادیر X_t ثابت باشد، آنگاه $\bar{X} = X_t$ و در نتیجه $x_t = X_t - \bar{X} = 0$ یا $\Sigma x_t^2 = 0$. در اینصورت محاسبهٔ $\hat{\beta}$ ممکن نخواهد بود.

۲. در مدل $Y_t = \beta X_t + U_t$ داریم

$$\hat{\beta} = \frac{\Sigma X_t Y_t}{\Sigma X_t^2} \ ,$$

و اگر فرض کنیم مقادیر X_t ثابت باشد، خواهیم داشت

$$\hat{\beta} = \frac{\Sigma X_t Y_t}{\Sigma X_t^2} = \frac{X_t \Sigma Y_t}{n X_t^2} = \frac{\bar{Y}}{X} \ .$$

نتیجهٔ فوق را می‌توان تعمیم داد. مدلهای رگرسیون را که فاقد متغیر توضیحی بوده اما ضریب ثابت دارند، می‌توان به صورت زیر نوشت،

$$Y_t = \alpha + U_t \ .$$

اگر α را با روش حداقل مربعات معمولی تخمین بزنیم، خواهیم داشت

$$\hat{\alpha}_{OLS} = \bar{Y} \ .$$

اثبات این نکته بسیار ساده است. مدل فوق را به صورت زیر می‌نویسیم،

$$Y_t = \alpha X_t + U_t \ ,$$

که در آن برای تمام مقادیر t داریم $X_t = 1$. می‌دانیم

$$\hat{\alpha}_{OLS} = \frac{\Sigma X_t Y_t}{\Sigma X_t^2} \ ,$$

با توجه به $X_t = 1$، خواهیم داشت $\hat{\alpha}_{OLS} = \bar{Y}$.

۱۱-۱ ۱. می‌دانیم $e_t = Y_t - \hat{Y}_t$ است؛ بنابر

$$e_t = Y_t - \hat{\alpha} - \hat{\beta} X_t \ .$$

می‌توان نوشت،

$$\Sigma e_t = \Sigma Y_t - n\hat{\alpha} - \hat{\beta} \Sigma X_t \ ,$$

یا

$$\Sigma e_t = \Sigma Y_t - n(\bar{Y} - \hat{\beta}\bar{X}) - \hat{\beta} \Sigma X_t \ ,$$

$$= \hat{\beta} \Sigma X_t - \hat{\beta} \Sigma X_t \ ,$$

$$= 0 \ .$$

۲. با توجه به مفهوم U_t، ملاحظه می‌شود که اگر مجموع پسماندها برابر صفر باشد، هیچ دلیلی وجود ندارد که مجموع جملات اختلال نیز صفر شود. U_t و e_t مفاهیم کاملاً متفاوتی هستند و چون مقادیر U_t مطلقاً ناشناخته است نمی‌توان در مورد مجموع مقادیر آنها هیچگونه اظهارنظری کرد.

۱۲-۱ از مدل اولیه داریم

$$\hat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} \quad , \quad \hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \ .$$

از دو مدل دوم داریم

$$\hat{\beta}_i = \frac{\Sigma x_t y_{it}}{\Sigma x_t^2} \quad , \quad i = ۱ , ۲$$

$$\hat{\alpha}_i = \bar{Y}_i - \hat{\beta}_i \bar{X} \quad , \quad i = ۱ , ۲$$

می‌توان نوشت،

$$\Sigma \hat{\beta}_i = \hat{\beta}_۱ + \hat{\beta}_۲ ,$$

$$= \frac{\Sigma (y_{۱t} + y_{۲t}) x_t}{\Sigma x_t^2} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} = \hat{\beta} .$$

حال برای نشان دادن صحت $\hat{\alpha} = \hat{\alpha}_۱ + \hat{\alpha}_۲$ چنین می‌نویسیم،

$$\Sigma \hat{\alpha}_i = \hat{\alpha}_۱ + \hat{\alpha}_۲ = \Sigma \bar{Y}_i - \Sigma \hat{\beta}_i \bar{X} ,$$

$$= \bar{Y}_۱ + \bar{Y}_۲ - (\hat{\beta}_۱ + \hat{\beta}_۲) \bar{X} ,$$

$$= \bar{Y} - \hat{\beta} \bar{X} ,$$

$$= \hat{\alpha} .$$

البته باید توجه کرد که این قاعده در مورد متغیر توضیحی صادق نیست؛ یعنی اگر X_t را به k قسمت تقسیم کنیم و به کمک Y_t برای هر سری از مشاهدات X_i ، $i = ۱ , ۲ , \dots , k$ ، یک مدل رگرسیون بسازیم، مشابه نتیجه فوق حاصل نخواهد شد.

۱-۱۳ می‌دانیم

$$\hat{\beta}_i = \frac{\sum_t x_{it} x_t}{\sum_t x_t^2} \quad , \quad \hat{\alpha}_i = \bar{X}_i - \hat{\beta}_i \bar{X} ,$$

اگر برای تمام مقادیر i جمع کنیم، خواهیم داشت

$$\sum_{i=۱}^{k} \hat{\beta}_i = \frac{\sum_i \sum_t x_{it} x_t}{\sum_t x_t^2} = \frac{\sum x_t^2}{\sum x_t^2} = ۱ ,$$

$$\sum_i \hat{\alpha}_i = \sum_i \bar{X}_i - \sum \hat{\beta}_i \bar{X} \ .$$

می‌دانیم $\sum_i \bar{X}_i = \bar{X}$ و همچنین $\sum \hat{\beta}_i \bar{X} = \bar{X}$؛ زیرا

$$\sum \hat{\beta}_i \bar{X} = \bar{X} \sum \hat{\beta}_i = \bar{X} \ (1) = \bar{X} \ ,$$

بنابراین خواهیم داشت

$$\sum_i \hat{\alpha}_i = \bar{X} - \bar{X} = 0 \ .$$

۱-۱۴ می‌دانیم

$$\hat{b} = \frac{\sum y_t \hat{y}_t}{\sum \hat{y}_t^2} \ .$$

با توجه به $\hat{y}_t = \hat{\beta} x_t$ داریم

$$\hat{b} = \frac{\sum y_t (\hat{\beta} x_t)}{\sum (\hat{\beta} x_t)^2} = \frac{\hat{\beta} \sum x_t y_t}{\hat{\beta}^2 \sum x_t^2} = 1 \ .$$

از طرف دیگر می‌دانیم

$$\hat{a} = \bar{Y} - \hat{b} \bar{\hat{Y}} \ . \qquad (1)$$

اما:

$$\bar{\hat{Y}} = \frac{\sum (\hat{\alpha} + \hat{\beta} X_i)}{n} = \hat{\alpha} + \hat{\beta} \bar{X} \ .$$

با جایگزینی $\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}$ در رابطهٔ فوق داریم.

$$\bar{\hat{Y}} = \bar{Y} \ .$$

اگر در نتیجهٔ فوق یعنی $\hat{b} = 1$ و $\bar{\hat{Y}} = \bar{Y}$ را در فرمول (۱) قرار دهیم خواهیم داشت

$$\hat{a} = 0 \ .$$

۱-۱۵ اگر $\hat{Y} = \hat{\beta} X_t$ بخواهد از نقطهٔ $(\bar{X} , \bar{Y})$ بگذرد، باشد داشته باشیم

$$\bar{Y} = \hat{\beta}\,\bar{X} \ .$$

و چون می‌دانیم

$$\hat{\beta} = \frac{\sum X_t Y_t}{\sum X_t^2} \ ,$$

بنابراین رابطهٔ زیر باید همواره برقرار باشد،

$$\bar{Y} = \frac{\bar{X} \sum X_t Y_t}{\sum X_t^2} \ .$$

ملاحظه می‌شود که هیچ دلیلی بر صحت رابطهٔ فوق به ازای جمیع مقادیر t وجود ندارد. با اینکه مدلهای رگرسیون با ضریب ثابت، همواره از نقطهٔ $(\bar{X}\ ,\ \bar{Y})$ می‌گذرند، این خصوصیت برای مدلهای رگرسیون فاقد ضریب ثابت معمولاً برقرار نیست.

۱-۱۶ هر چند در فصل هشتم، این گونه تخمینهای مقید به طور گسترده مطرح خواهد شد، در این فصل به طور خلاصه به آن اشاره می‌شود. همان گونه که مسألهٔ تخمین در اقتصادسنجی به مسأله حداقل‌سازی توابع تبدیل می‌شود، تخمینهای مقید نیز به حداقل‌سازی مقید تبدیل می‌گردد. در چنین مواردی برای اینکه مجموع مربعات پسماندها را با توجه به شرط یا قید مفروض حداقل کنیم، دو راه وجود دارد: یا باید قید را در معادله‌ای که قرار است حداقل شود جایگزین کنیم یا با روش ضریب لاگرانژ عمل کنیم. راه حل دوم بیشتر معمول است. اگر λ را ضریب لاگرانژ بنامیم، با توجه به اینکه قید ما صفر بودن $\hat{\alpha}$ است، باید مجموع مربعات پسماند را با توجه به این قید حداقل کنیم. در واقع باید عبارت زیر حداقل شود،

$$S = \sum e_t + \lambda\,\hat{\alpha} \ .$$

با توجه به اینکه $e_t = Y_t - \hat{\alpha} - \hat{\beta}\,X_t$ یا $e_t = Y_t - \hat{Y}_t$؛ بنابراین

$$S = \sum (Y_t - \hat{\alpha} - \hat{\beta}\,X_t)^2 + \lambda\,\hat{\alpha} \ .$$

از S نسبت به $\hat{\alpha}$ ، $\hat{\beta}$ و λ مشتق می‌گیریم،

$$\frac{\partial s}{\partial \hat{\alpha}} = -۲ \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t) + \lambda = ۰ ,$$

$$\frac{\partial s}{\partial \hat{\beta}} = -۲ \sum X_t (Y_t - \hat{\alpha} - \hat{\beta} X_t) = ۰ ,$$

$$\frac{\partial s}{\partial \lambda} = \hat{\alpha} = ۰ .$$

با جایگزینی معادلهٔ سوم در معادلهٔ دوم، به معادلهٔ نرمال برای مدل رگرسیون بدون ضریب ثابت می‌رسیم. اگر آن را برای $\hat{\beta}$ حل کنیم، خواهیم داشت

$$\hat{\beta} = \frac{\sum X_t Y_t}{\sum X_t^۲} . \qquad (۱)$$

همچنین از معادلهٔ اول نیز داریم

$$\lambda = ۲ \sum (Y_t - \hat{\beta} X_t) .$$

اگر λ صفر شود، در آن صورت قید مفروض بی‌تأثیر است؛ یعنی اساساً فرقی نمی‌کند که تخمین با توجه به وجود قید صورت پذیرد یا اینکه یک تخمین آزاد باشد. حال باید دید که در چه صورتی λ برابر صفر است. بدیهی است موقعی λ برابر صفر است که

$$\sum (Y_t - \hat{\beta} X_t) = ۰ ,$$

یعنی $\bar{Y} = \hat{\beta} \bar{X}$ رابطهٔ اخیر موقعی صحیح است که تخمین مدل رگرسیون $(\hat{Y}_t = \hat{\beta} X_t)$ از نقطهٔ $(\bar{X} , \bar{Y})$ بگذرد.

با جایگزینی $\hat{\beta}$ در معادلهٔ $\bar{Y} = \hat{\beta} \bar{X}$ خواهیم داشت

$$\bar{Y} = \frac{\bar{X} \sum X_t Y_t}{\sum X_t^۲} . \qquad (۲)$$

قبلاً در مسألهٔ ۱.۱۵ نیز دقیقاً به معادلهٔ (۲) رسیده بودیم. از طرف دیگر می‌دانیم که تخمین β از مدل $Y_t = \alpha + \beta X_t + U_t$ برابر است با

$$\hat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2} \qquad (۳)$$

با یک جایگزینی ساده می‌توان نشان داد که موقعی تخمین β از دو مدل فوق به یک جواب می‌رسد ـ یعنی (۱) و (۳) برابر است ـ که رابطه (۲) برقرار باشد.

۱ـ۱۷ این مسأله نکتهٔ خاصی در اقتصادسنجی ندارد و بیشتر کاربرد روش حداقل‌سازی معادله‌های، بدون استفاده از مشتق‌گیری است. می‌دانیم

$$\sum e_t^2 = \sum (Y_t - \hat{Y}_t)^2 = \sum (Y_t - \hat{\beta} X_t)^2 ,$$

$$= \sum Y_t^2 + \hat{\beta}^2 \sum X_t^2 - 2\hat{\beta} \sum X_t Y_t .$$

متغیر ما، $\hat{\beta}$ است. معادلهٔ فوق را برحسب توانهای $\hat{\beta}$ مرتب می‌کنیم،

$$\sum e_t^2 = A_0 \hat{\beta}^2 + A_1 \hat{\beta} + A_2 , \qquad (۱)$$

که در آن $A_0 = \sum X_t^2$ و $A_1 = -2 \sum X_t Y_t$ و $A_2 = \sum Y_t^2$. سعی می‌کنیم یک مربع کامل برحسب $\hat{\beta}$ از رابطهٔ (۱) به دست آوریم،

$$\sum e_t^2 = A_0 \left(\hat{\beta}^2 + \frac{A_1}{A_0}\hat{\beta} + \frac{A_1^2}{4A_0^2}\right) - \left(\frac{A_1^2}{4A_0} - A_2\right) ,$$

$$= A_0 \left(\hat{\beta} + \frac{A_1}{2A_0}\right)^2 - \left(\frac{A_1^2}{4A_0} - A_2\right) .$$

فقط جملهٔ اول، تابعی از $\hat{\beta}$ است؛ بنابراین $\sum e_t^2$ هنگامی حداقل است که این جمله حداقل باشد. با توجه به اینکه این جمله مربع کامل است، حداقل آن برابر صفر است.

$$\hat{\beta} + \frac{A_1}{2A_0} = 0 ,$$

$$\hat{\beta} = -\frac{A_1}{2A_0} = -\frac{-2\sum X_t Y_t}{2\sum X_t^2} = \frac{\sum X_t Y_t}{\sum X_t^2} .$$

۱-۱۸ هر یک از دو مدل را یک بار دیگر می‌نویسیم،

$$Y_t = \alpha + \beta X_t + U_t \ , \qquad (۱)$$

$$Y_t^* = \alpha + \beta X_t^* + U_t \ . \qquad (۲)$$

توجه داریم که U_t تغییر نمی‌کند؛ زیرا فقط مقیاس اندازه‌گیری هر یک از دو متغیر را تغییر داده‌ایم که قاعدتاً در ساختار دینامیکی مدل تأثیری ندارد. بدیهی است که

$$\hat{b} = \frac{\sum x_t^* y_t^*}{\sum x_t^{*2}} \ . \qquad (۳)$$

حال می‌گوییم $y_t^* = y_t$؛ برای اثبات کافی است در دو طرف این معادله، مقادیر هر یک را جایگزین کنیم خواهیم داشت

$$Y_t^* - \bar{Y}^* = Y_t - \bar{Y} \ .$$

با جایگزینی $Y_t^* = Y_t - Y^*$، داریم:

$$Y_t - Y^* - \frac{\sum (Y_t - Y^*)}{n} = Y_t - \bar{Y} \ ,$$

$$Y_t - Y^* - \bar{Y} + \frac{n Y^*}{n} = Y_t - \bar{Y} \ .$$

ملاحظه می‌شود که رابطهٔ فوق همواره برقرار است. به همین ترتیب می‌توان نشان داد که $x_t^* = x_t$. با جایگزینی مقادیر $y_t^* = y_t$ و $x_t^* = x_t$ در معادلهٔ (۳) خواهیم داشت

$$\hat{b} = \frac{\sum x_t y_t}{\sum x_t^2} = \hat{\beta} \ .$$

نتیجه می‌گیریم که تخمین شیب معادلهٔ رگرسیونی در هر دو مدل دقیقاً با یکدیگر برابر است. اما در مورد ضریب ثابت به ترتیب زیر عمل می‌کنیم. با استفاده از مدل (۲) داریم

$$\hat{a} = \bar{Y}^* - \hat{b} \bar{X}^* \ ,$$

$$= \frac{\sum (Y_t - Y^*)}{n} - \hat{b}\,\frac{\sum (X_t - X^*)}{n} ,$$

$$= \bar{Y} - Y^* - \hat{b}\,\bar{X} + \hat{b}\,X^* .$$

با توجه به معادلهٔ (۱) می‌دانیم

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{X} ,$$

در نتیجه، خواهیم داشت

$$\hat{a} = \hat{\alpha} - Y^* + \hat{\beta}\,X^* .$$

۱.۱۹ تساوی معادله‌های (۱)، (۲) و (٤) را در متن دیدیم. در اینجا تنها به اثبات تساوی دو معادلهٔ (۱) و (۳) می‌پردازیم. می‌دانیم

$$\hat{y}_t = \hat{\beta}\,x_t .$$

با جایگزینی در معادلهٔ (۳) خواهیم داشت

$$r^2 = \frac{\left[\sum y_t\,(\hat{\beta}\,x_t)\right]^2}{\sum y_t^2 \sum (\hat{\beta}\,x_t)^2} ,$$

$$= \frac{\hat{\beta}^2\,(\sum x_t\,y_t)^2}{\hat{\beta}^2 \sum x_t^2 \sum y_t^2} = \frac{(\sum x_t\,y_t)^2}{\sum x_t^2 \sum y_t^2} ,$$

که در واقع همان معادلهٔ (۱) است.

۱.۲۰ می‌دانیم، $\hat{Y}_t = \hat{\beta}\,X_t$. با جایگزینی در معادلهٔ (۲) داریم

$$r^2 = \frac{(\sum Y_t \hat{\beta}\,X_t)^2}{\sum Y_t^2 \sum (\hat{\beta}\,X_t)^2} = \frac{\hat{\beta}^2\,(\sum X_t\,Y_t)^2}{\hat{\beta}^2 \sum X_t^2 \sum Y_t^2} ,$$

$$= \frac{\sum (X_t\,Y_t)^2}{\sum X_t^2 \sum Y_t^2} ,$$

که در واقع همان معادلهٔ (۱) است.

برای نشان‌دادن تساوی معادله‌های (۳) و (۱) کافی است معادلهٔ (۳) را به صورت زیر بنویسیم،

$$r^2 = 1 - \frac{\sum (Y_t - \hat{\beta} X_t)^2}{\sum Y_t^2} .$$

با جایگزینی $\hat{\beta} = \frac{\Sigma X_t Y_t}{\Sigma X_t^2}$ خواهیم داشت

$$r^2 = 1 - \frac{\left[\sum Y_t^2 - \frac{\Sigma (X_t Y_t)^2}{\Sigma X_t^2}\right]}{\sum Y_t^2} ,$$

$$= \frac{(\sum X_t Y_t)^2}{\sum X_t^2 \sum Y_t^2} ,$$

که دقیقاً همان معادلهٔ (۱) است. و معادله (۱) نیز تعریف r^2 در مدلهایی است که جمله ثابت ندارند.

۲۱ـ۱ راه‌حل اول: برای اثبات، ابتدا به ذکر یک مقدمهٔ ریاضی می‌پردازیم. نامساوی

$$\sum a_t^2 \sum b_t^2 \geqq (\sum a_t b_t)^2 ,$$

به نامساوی کوچی ـ شوارتز[1] معروف است. با توجه به اینکه اثبات دقیق این نامساوی مستلزم دقتهای ریاضی است ـ که ضرورتاً کاربردی در تحلیلهای اقتصادسنجی ندارد ـ در اینجا به طرح روش بسیار ساده اثبات این نامساوی می‌پردازیم. ابتدا نامساوی را برای حالتی می‌نویسیم که ۲ و ۱ = t. خواهیم داشت

$$\sum_{t=1}^{2} a_t^2 \sum_{t=1}^{2} b_t^2 \geqq (\sum_{t=1}^{2} a_t b_t)^2 ,$$

$$(a_1^2 + a_2^2)(b_1^2 + b_2^2) \geqq (a_1 b_1 + a_2 b_2)^2 .$$

1. Cauchy - Schwartz

دو طرف نامساوی را بسط می‌دهیم و ساده می‌کنیم،

$$a_1^2 b_2^2 + a_2^2 b_1^2 - 2 a_1 a_2 b_1 b_2 \geq 0 \ ,$$

یا

$$(a_1 b_2 - a_2 b_1)^2 \geq 0 \ ,$$

که همواره مثبت است، مگر در حالتی که $\frac{a_1}{b_1} = \frac{a_2}{b_2}$ که در آن صورت برابر صفر می‌شود.

در حالت کلی، اگر جمله‌های $a_t^2 b_t^2$ را برای تمام مقادیر t از دو طرف نامساوی حذف کنیم، خواهیم داشت

$$\sum_{s=1}^{n} \sum_{v=1}^{n} (a_s b_v - a_v b_s)^2 \geq 0 \ , \qquad s \neq v$$

که همواره مثبت است، مگر در حالتی که $\frac{a_s}{b_s} = \frac{a_v}{b_v}$ که در آن صورت برابر صفر می‌شود.

بعد از یادآوری این نکته ریاضی به بحث خود برمی‌گردیم. می‌دانیم

$$r^2 = 1 - \frac{RSS}{TSS} \ .$$

در مدل $Y_t = \beta X_t + U_t$، داریم

$$RSS = \sum e_t^2 = \sum (Y_t - \hat{Y}_t)^2 \ ,$$

$$= \sum (Y_t - \hat{\beta} X_t)^2 \ ,$$

$$= \sum Y_t^2 + \hat{\beta}^2 \sum X_t^2 - 2 \hat{\beta} \sum X_t Y_t \ ,$$

که با جایگزینی $\hat{\beta} = \frac{\sum X_t Y_t}{\sum X_t^2}$، خواهیم داشت

$$RSS = \sum Y_t^2 - \frac{(\sum X_t Y_t)^2}{\sum X_t^2} \ .$$

همچنین می‌دانیم

$$TSS = \sum y_t^2 = \sum (Y_t - \bar{Y})^2 \ .$$

بنابراین

$$r^2 = 1 - \frac{RSS}{TSS} ,$$

$$= 1 - \frac{\sum Y_t^2 - \frac{(\Sigma X_t Y_t)^2}{\Sigma X_t^2}}{\sum (Y_t - \bar{Y})^2} ,$$

$$= 1 - \left[\frac{\sum Y_t^2}{\sum (Y_t - \bar{Y})^2} - \frac{(\sum X_t Y_t)^2}{\sum (Y_t - \bar{Y})^2 \sum X_t^2} \right] ,$$

$$= 1 - \left[\frac{\sum Y_t^2}{\sum (Y_t - \bar{Y})^2} - \frac{(\sum X_t Y_t)^2}{\sum X_t^2 \sum Y_t^2} \cdot \frac{\sum Y_t^2}{\sum (Y_t - \bar{Y})^2} \right] .$$

اگر فرض کنیم

$$q = \frac{\sum Y_t^2}{\sum (Y_t - \bar{Y})^2} \quad , \quad r^{*2} = \frac{(\sum X_t Y_t)^2}{\sum X_t^2 \sum Y_t^2} ,$$

آنگاه خواهیم داشت

$$r^2 = 1 - [q - r^{*2} q] ,$$

$$= 1 - (1 - r^{*2}) q . \qquad (1)$$

در اینجا به بررسی q و r^{*2} می‌پردازیم. می‌دانیم

$$q = \frac{\sum_{t=1}^{n} Y_t^2}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2} = \frac{Y_1^2 + Y_2^2 + \cdots + Y_n^2}{(Y_1 - \bar{Y})^2 + (Y_2 - \bar{Y})^2 + \cdots + (Y_n - \bar{Y})^2} .$$

صورت و مخرج کسر q را بر $\bar{Y}^2$، تقسیم می‌کنیم،

$$q = \frac{\frac{Y_1^2}{\bar{Y}^2} + \frac{Y_2^2}{\bar{Y}^2} + \dots + \frac{Y_n^2}{\bar{Y}^2}}{\frac{Y_1^2 + \bar{Y}^2 - 2Y_1\bar{Y}}{\bar{Y}^2} + \frac{Y_2^2 + \bar{Y}^2 - 2Y_2\bar{Y}}{\bar{Y}^2} + \dots + \frac{Y_n^2 + \bar{Y}^2 - 2Y_n\bar{Y}}{\bar{Y}^2}},$$

$$= \frac{\sum \left(\frac{Y_t}{\bar{Y}}\right)^2}{\sum \left(\frac{Y_t}{\bar{Y}}\right)^2 + n - \frac{2\bar{Y}\sum Y_t}{\bar{Y}^2}} = \frac{\sum \left(\frac{Y_t}{\bar{Y}}\right)^2}{\sum \left(\frac{Y_t}{\bar{Y}}\right)^2 - n}.$$

هنگامی که تمام مقادیر Y_t با یکدیگر برابر شوند حد بالای q برابر $+\infty$ است و وقتی که تغییرات Y_t بسیار زیاد باشد، حد پایین آن برابر یک است. اما با توجه به اینکه r^{*2} بنا بر قضیهٔ کوچی ـ شوارتز بین صفر و یک نوسان می‌کند، می‌توان اینگونه نتیجه گرفت که به ازای مقادیر کوچک r^{*2} و مقادیر بزرگ q، چه بسا مقدار r^2 بتواند منفی شود. البته توجه داریم که مقادیر r^{*2} و q مستقل از یکدیگر تعیین می‌شوند.

راه‌حل دوم: راه‌حل ساده‌تری نیز برای این مسأله وجود دارد؛ با وجود این، به این نکته اشاره می‌کنیم که مزیت راه‌حل فوق این است که می‌تواند رابطهٔ بین ضریب تعیین معمولی r^2 و ضریب تعیین برای مدلهای بدون جملهٔ ثابت r^{*2} را نشان دهد. به هر حال راه‌حل دوم بسیار خلاصه و به شرح زیر است.

در مدل $Y_t = \beta X_t + U_t$، می‌دانیم

$$Y_t = \hat{Y}_t + e_t .$$

با توجه به $\hat{Y}_t = \hat{\beta} X_t$، خواهیم داشت

$$\hat{Y}_t = \hat{\beta} X_t + e_t .$$

دو طرف رابطه فوق را مجذور کرده و برای تمام مقادیر t جمع می‌کنیم،

$$\sum Y_t^2 = \hat{\beta}^2 \sum X_t^2 + \sum e_t^2 + 2\hat{\beta} \sum X_t e_t .$$

می‌دانیم $\Sigma X_t e_t$، بنا بر معادلهٔ نرمال صفر است؛ زیرا

$$\frac{d\sum e_t^2}{d\hat{\beta}} = \frac{d\sum (Y_t - \hat{Y}_t)^2}{d\hat{\beta}} = \frac{d\sum (Y_t - \hat{\beta} X_t)^2}{d\hat{\beta}} .$$

$$= -2\sum X_t (Y_t - \hat{\beta} X_t) = -2\sum X_t e_t = 0 .$$

بنابراین خواهیم داشت:

$$\sum Y_t^2 = \sum \hat{Y}_t^2 + \sum e_t^2 .$$

اگر r^2 را برای اینگونه مدلها که فاقد جملهٔ ثابت هستند به صورت

$$r^2 = 1 - \frac{\sum e_t^2}{\sum Y_t^2}$$

تعریف کنیم، آنگاه $0 < r^2 < 1$. در غیر این صورت اگر r^2 را به روال معمول تعریف نماییم، خواهیم داشت

$$r^2 = 1 - \frac{\sum e_t^2}{\sum y_t^2} = 1 - \frac{\sum e_t^2}{\sum Y_t^2 - n\bar{Y}^2} = \frac{\sum Y_t^2 - n\bar{Y}^2 - (\sum Y_t^2 - \hat{\beta}^2 \sum X_t^2)}{\sum Y_t^2 - n\bar{Y}^2}$$

در نتیجه داریم

$$r^2 = \frac{\hat{\beta}^2 \sum X_t^2 - n\bar{Y}^2}{\sum y_t^2} .$$

مخرج کسر همواره مثبت است، اما در مواردی که $\hat{\beta}^2 \Sigma X_t^2 < n\bar{Y}$، مقدار r^2 منفی خواهد شد. یادآوری می‌کنیم در مدلهایی که ضریب ثابت ندارند $\bar{\hat{Y}} \neq \bar{Y}$؛ بنابراین $\hat{\beta}^2 \Sigma X_t^2 = \Sigma \hat{Y}_t^2 \neq \Sigma \hat{y}_t^2 + n\bar{Y}^2$.

۱-۲۲ ۱. مدل (۲) را به صورت زیر می‌نویسیم،

$$Y_t^* = a + b X_t + U_t \qquad (۳)$$

که در آن، $Y_t^* = Y_t - X_t$ و $a = \alpha$ و $b = (\beta - ۱)$. به سهولت ملاحظه می‌شود که با تخمین b از مدل (۲) یا (۳) می‌توان β را به دست آورد؛ زیرا

$$\hat{b} = \frac{\sum x_t y_t}{\sum x_t^۲} ,$$

$$= \frac{\sum x_t (y_t - x_t)}{\sum x_t^۲} = \hat{\beta} - ۱ .$$

همچنین ملاحظه می‌شود که $\hat{a} = \hat{\alpha}$؛ به عبارت دقیقتر از معادلهٔ (۳) داریم

$$\hat{a} = \bar{Y}^* - \hat{b}\,\bar{X} .$$

با توجه به $\hat{b} = \hat{\beta} - ۱$ داریم

$$\hat{a} = \bar{Y} - \bar{X} - (\hat{\beta} - ۱)\,\bar{X} ,$$

$$= \bar{Y} - \bar{X} - \hat{\beta}\,\bar{X} + \bar{X} ,$$

$$= \bar{Y} - \hat{\beta}\,\bar{X} ,$$

$$= \hat{\alpha} .$$

۲. با استفاده از فرمول $r^۲ = ۱ - \frac{RSS}{\sum y_t^۲}$، برای مدل (۱) داریم

$$\frac{(\sum x_t y_t)^۲}{\sum x_t^۲ \sum y_t^۲} = ۱ - \frac{RSS}{\sum y_t^۲} ,$$

یا:

$$RSS = \sum y_t^۲ - \frac{(\sum x_t y_t)^۲}{\sum x_t^۲} .$$

اگر RSS (تغییرات توضیح داده نشده) را برای مدل (۲) به صورت RSS^* تعریف کنیم، خواهیم داشت

$$RSS^* = \sum y^{*۲} - \frac{(\sum x_t y_t^*)^۲}{\sum x_t^۲} .$$

می‌دانیم $y_t^* = y_t - x_t$، در نتیجه داریم

$$RSS^* = \sum (y_t - x_t)^2 - \frac{[\sum x_t (y_t - x_t)]^2}{\sum x_t^2} ,$$

که بعد از ساده کردن، عبارت است از

$$RSS^* = \sum y_t - \frac{(\sum x_t y_t)^2}{\sum x_t^2} ,$$

$$= RSS .$$

نتیجه می‌گیریم که مجموع مربعات پسماند در هر دو مدل با یکدیگر برابر است.

در اینجا به بررسی r^2 در هر یک از دو مدل (۱) و (۲) می‌پردازیم. اگر r^2 را برای مدل دوم به صورت r^{*2} تعریف کنیم، خواهیم داشت

$$r^{*2} = \frac{(\sum x_t y_t^*)^2}{\sum x_t^2 \sum y_t^{*2}} \qquad (۴)$$

رابطهٔ $y_t^* = y_t - x_t$ را جایگزین می‌کنیم،

$$r^{*2} = \frac{(\sum x_t y_t)^2 + M}{\sum x_t^2 \sum y_t^2 + M} , \qquad (۵)$$

که در آن:

$$M = (\sum x_t^2)^2 - 2 (\sum x_t y_t) (\sum x_t^2) .$$

با ملاحظهٔ معادلهٔ (۴) و بنا بر نامساوی کوچی ـ شوارتز، می‌دانیم

$$\sum x_t^2 \sum y_t^{*2} \geq (\sum x_t y_t)^2 ,$$

در معادلهٔ (۴) مخرج کسر معمولاً از صورت آن بزرگتر است. می‌دانیم اگر به صورت و مخرج کسر $\frac{A}{B}$، که در آن $A<B$، یک عدد مثبت اضافه کنیم، مقدار کسر بزرگتر می‌شود

و در صورتی که آن عدد منفی باشد، مقدار کسر کوچکتر خواهد شد،

$$\frac{A}{B} < \frac{A + M}{B + M} .$$

برای اثبات کافی است که مخرجها را به یک مخرج مشترک تبدیل کنیم. با توجه به اینکه در معادلهٔ (۵)، M نمی‌تواند مقداری باشد که r^{*2} منفی شود، داریم

$$M > ۰ \qquad \text{هرگاه} \qquad r^{*} > r^{۲} ,$$

$$M < ۰ \qquad \text{هرگاه} \qquad r^{*} < r^{۲} .$$

اما $M > ۰$ بدین معنی است که

$$(\sum x_t^{۲})^{۲} > ۲ (\sum x_t y_t) (\sum x_t^{۲}) ,$$

یا

$$\frac{\sum x_t y_t}{\sum x_t^{۲}} < \frac{۱}{۲} ,$$

یعنی

$$\hat{\beta} < \frac{۱}{۲} .$$

نتیجه می‌گیریم که اگر در مدل اولیه $\hat{\beta} < \frac{۱}{۲}$، می‌توان مدل را چنان تبدیل کرد که ضریب تعیین جدید آن بزرگتر از حالت قبلی باشد. باید توجه داشت که این قاعده عمومی است و برای حالت زیر نیز صادق خواهد بود،

$$Y_t - k X_t = \alpha + (\beta - k) X_t + U_t .$$

فصل دوم

آزمونهای آماری و خصوصیات مطلوب تخمین‌زننده‌ها در مدل رگرسیون خطی ساده

۲ـ۱ مقدمه

اگر مدل رگرسیون و تخمین آن به ترتیب عبارت باشند از:

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t \ ,$$

آنگاه مقادیر $\hat{\alpha}_{OLS}$ و $\hat{\beta}_{OLS}$ را می‌توان از طریق فرمولهای زیر به دست آورد،

$$\hat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2} \quad , \quad \hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X} \ .$$

$\hat{\beta}$ یک متغیر تصادفی است؛ زیرا تابعی از y_t می‌باشد که خود تابعی از U_t است. همچنین $\hat{\alpha}$ که تابعی از $\hat{\beta}$ است تصادفی خواهد شد. $\hat{\alpha}$ و $\hat{\beta}$ مانند هر متغیر تصادفی دارای تابع توزیع احتمال با میانگین و واریانس معیّن هستند.

یکی از هدفهای اصلی در محاسبهٔ $\hat{\alpha}$ و $\hat{\beta}$ به دست آوردن اطلاعات دربارهٔ α و β واقعی است؛ به عبارت دیگر، با داشتن $\hat{\alpha}$ و $\hat{\beta}$ باید بتوان فرضیه‌های مختلفی را در مورد α و β آزمون کرد. در قسمت ۲ـ۳ آزمون فرضیه‌های مربوط به هر یک از پارامترها را مطالعه خواهیم کرد. در این آزمونها از آماره‌های Z و t استفاده خواهد شد. تعیین فاصله‌های اطمینان برای پارامترهای α و β و رابطهٔ بین فاصله اطمینان و دقت تخمین پارامترها از موضوعات دیگر این قسمت است. آزمون واریانس جملهٔ اختلال، که به کمک آمارهٔ χ^2 انجام می‌شود، نیز در همین قسمت

بررسی خواهد شد.

کاربرد مسألهٔ آنالیز واریانس در آزمون معنی‌دار بودن کل مدل رگرسیون موضوع قسمت ٤ـ۲ است. بعد از معرفی اجمالی آنالیز واریانس و ارائهٔ تعاریف توزیع و آزمون F، به چگونگی اجرای آزمون معنی‌دار بودن مدل رگرسیون می‌پردازیم. رابطهٔ بین آزمونهای F و t و نیز رابطهٔ بین آزمونهای معنی‌دار بودن ضریب تعیین و معنی‌دار بودن کل رگرسیون از اهمیت خاصی برخوردار است که در همین قسمت به طور خلاصه بررسی می‌شود.

خصوصیات مطلوب تخمین‌زننده‌ها در حالت کلی، موضوع قسمت ۲ـ۵ را تشکیل می‌دهد. این سؤال از لحاظ نظری اهمیت بسیاری دارد که اگر $\hat{\theta}$ تخمینی از پارامتر θ باشد، آنگاه خصوصیات مطلوب $\hat{\theta}$ چیست؟ خصوصیات مطلوب $\hat{\theta}$ را در دو بخش بررسی خواهیم کرد. ابتدا این خصوصیات را در نمونه‌هایی مطالعه می‌کنیم که حجم آنها محدود است، سپس همین مسأله، با عنوان خصوصیات حدی تخمین‌زننده‌ها، برای نمونه‌های بسیار بزرگ بررسی خواهد شد. بدیهی است، مطالعهٔ این خصوصیات، مستقل از روشی است که به کمک آن پارامتر θ را تخمین زده‌ایم. در قسمت ۲ـ٦ به تحلیل این مسأله می‌پردازیم که اگر پارامتر مفروض را با روش حداقل مربعات معمولی (OLS) تخمین بزنیم، آیا خصوصیات مطلوب تخمین‌زننده‌ها را خواهد داشت؟ خواهیم دید که بنا بر قضیهٔ گاس ـ مارکف پاسخ این سؤال مثبت است.

۲ـ۲ خصوصیات آماری $\hat{\alpha}$ و $\hat{\beta}$

مدل رگرسیون ۱ـ۸ و تخمین آن، معادلهٔ ۱ـ۱٦ را یک بار دیگر می‌نویسیم،

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t \ .$$

بحث را روی خصوصیات آماری $\hat{\beta}$ متمرکز می‌کنیم. برای $\hat{\alpha}$ می‌توان دقیقاً استدلال مشابهی عرضه کرد.

در معادلهٔ ۲۵ـ۱ دیدیم که[۱]

$$\hat{\beta}_{OLS} = \frac{\sum x_t y_t}{\sum x_t^2} .$$

می‌دانیم

$$y_t = \beta x_t + u_t , \qquad (۱ـ۲)$$

زیرا کافی است از دو طرف مدل ۸ـ۱ میانگین گرفته و معادلهٔ به دست آمده را از ۸ـ۱ کم کنیم، خواهیم داشت

$$(Y_t - \bar{Y}) = (\alpha - \alpha) + \beta (X_t - \bar{X}) + (U_t - \bar{U})$$

یا:

$$y_t = \beta x_t + u_t .$$

معادلهٔ ۱ـ۲ را در ۲۵ـ۱ جایگزین می‌کنیم

$$\hat{\beta} = \frac{\sum x_t (\beta x_t + u_t)}{\sum x_t^2} ,$$

$$\hat{\beta} = \beta + \frac{\sum x_t u_t}{\sum x_t^2} . \qquad (۲ـ۲)$$

$\hat{\beta}$ تابعی از u_t است[۲]، در نتیجه یک متغیر تصادفی است.

---

۱. در فرمول $\hat{\beta}$ می‌توان در صورت کسر به جای $\Sigma x_t y_t$ مقدار $\Sigma x_t Y_t$ را نوشت؛ زیرا

$$\Sigma x_t y_t = \Sigma x_t (Y_t - \bar{Y}) = \Sigma x_t Y_t - \bar{Y} \Sigma x_t = \Sigma x_t Y_t .$$

۲. در معادلهٔ ۲ـ۲ می‌توان در صورت کسر به جای $\Sigma x_t u_t$، مقدار $\Sigma x_t U_t$ را نوشت؛ زیرا

$$\Sigma x_t u_t = \Sigma x_t (U_t - \bar{U}) = \Sigma x_t U_t - \bar{U} \Sigma x_t = \Sigma x_t U_t .$$

توجه داریم که $\Sigma x_t = 0$؛ زیرا

$$\Sigma x_t = \Sigma (X_t - \bar{X}) = \Sigma X_t - n \bar{X} = 0 .$$

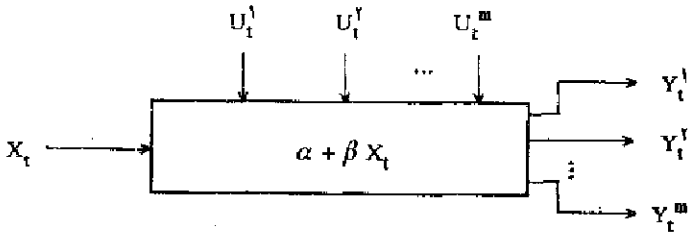
قبلاً فرض کرده‌ایم U_t دارای توزیع نرمال است؛ بنابراین با توجه به اینکه $\hat{\beta}$ یک تابع خطی از U_t است، $\hat{\beta}$ نیز دارای توزیع نرمال خواهد بود. سؤال این است که میانگین و واریانس $\hat{\beta}$ چیست؟

۱. میانگین $\hat{\beta}$

ابتدا سعی می‌کنیم مفهوم میانگین یا امید ریاضی $\hat{\beta}$ روشن شود. مدل رگرسیون ۸ـ۱ را دوباره می‌نویسیم:

$$Y_t = \alpha + \beta X_t + U_t \quad .$$

می‌دانیم X_t در آزمایشهای فرضی تکراری ثابت فرض می‌شود. برای تمام مقادیر t، $t = ۱, ۲, ۳, \dots, n$، مقدار X_t را ثابت نگاه داشته، m آزمایش انجام می‌دهیم. در هر آزمایش جملهٔ اختلال، یک مقداری به خود می‌گیرد و با افزوده شدن به مقدار معیّن $(\alpha + \beta X_t)$، مقدار متغیر درون‌زا را مشخص می‌کند. این نکته‌ها را می‌توان در نمودار ۱ـ۲ ملاحظه کرد.



نمودار ۱ـ۲ مفهوم $E(\hat{\beta})$

به ازای X_t معیّن و در آزمایش اول، جملهٔ اختلال مدل، مقدار U_t^1 را می‌گیرد. این مقدار به کمیت ثابت $(\alpha + \beta X_t)$ اضافه می‌شود و Y_t^1 را نتیجه می‌دهد. در آزمایش دوم، U_t مقدار U_t^2 را گرفته که بعد از ورود به سیستم، مقدار Y_t^2 به دست می‌آید. به همین ترتیب با تعیین U_t^m مقدار Y_t^m حاصل خواهد شد.

در اینجا به ازای سری مشاهدات X_t، $t = ۱, ۲, \dots, n$ و سری مشاهدات Y_t^1 برای تمام $t = ۱, ۲, \dots, n$، می‌توان یک رگرسیون ساخت و پارامتر β را تخمین زد. این

تخمین را $\hat{\beta}^{1}$ می‌نامیم. به همین ترتیب با استفاده از n مشاهدهٔ X_t و مشاهدات n تایی Y_t در آزمایش دوم (Y_t^{2}) می‌توان مدل رگرسیون را دوباره تخمین زد و این بار $\hat{\beta}^{2}$ را به دست آورد. این استدلال به همین روال ادامه پیدا می‌کند تا $\hat{\beta}^{m}$ حاصل شود. امید ریاضی $\hat{\beta}$ در واقع میانگین این m مقدار است که به طور فرضی برای $\hat{\beta}$ به دست آمده است. با افزایش m می‌توان امید ریاضی $\hat{\beta}$ را به صورت $E(\hat{\beta})$ تعریف کرد.

بعد از آشنایی با مفهوم $E(\hat{\beta})$ به بررسی مقدار آن می‌پردازیم. معادلهٔ ۲ـ۲ را یک بار دیگر می‌نویسیم،

$$\hat{\beta} = \beta + \frac{\sum x_t U_t}{\sum x_t^{2}} .$$

می‌دانیم X_t در آزمایشهای تکراری، ثابت است؛ بنابراین x_t نیز همین خصوصیت را خواهد داشت. نتیجه می‌گیریم که $\sum x_t^{2}$ نیز برای تمام مقادیر t ($t = 1, 2, \dots, n$) در مجموع کمیت ثابتی است. این مقدار را با A نشان می‌دهیم؛ یعنی $\sum x_t^{2} = A$. خواهیم داشت

$$\hat{\beta} = \beta + \frac{x_1 U_1 + x_2 U_2 + \cdots + x_n U_n}{A} ,$$

$$= \beta + \frac{x_1}{A} U_1 + \frac{x_2}{A} U_2 + \cdots + \frac{x_n}{A} U_n .$$

با توجه به اینکه مقادیر x_t برای $t = 1, 2, \dots, n$ ثابت است؛ بنابراین

$$\frac{x_t}{A} = \frac{x_t}{\sum x_t^{2}} = w_t = \text{ ثابت} . \qquad (۲ـ۳)$$

در نتیجه خواهیم داشت

$$\hat{\beta} = \beta + w_1 U_1 + w_2 U_2 + \cdots + w_n U_n , \qquad (۲ـ۴)$$

یعنی $\hat{\beta}$ یک تابع خطی از جمله‌های اختلال است.

از دو طرف معادلهٔ ۲ـ۴ امید ریاضی می‌گیریم. با توجه به ثابت بودن w_t داریم

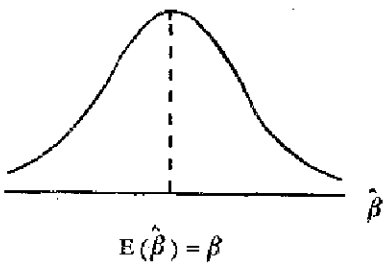
$$E(\hat{\beta}) = E(\beta) + w_1 E(U_1) + w_2 E(U_2) + \cdots + w_n E(U_n) .$$

چون β ثابت است بنابراین $E(\hat{\beta}) = \beta$، با در نظر گرفتن $E(U_i) = 0$ برای تمام مقادیر $t = 1, 2, \ldots, n$ خواهیم داشت

$$E(\hat{\beta}_{OLS}) = \beta \quad . \qquad (۲\text{-}۵)$$

نتیجه می‌گیریم که میانگین یا امید ریاضی مقادیر مختلف $\hat{\beta}$ که در آزمایش‌های فرضی تکراری و با ثابت بودن X_i به دست می‌آید، برابر مقدار واقعی پارامتر جامعه، یعنی β است. یادآوری میشود که β واقعی قابل مشاهده نیست، همانگونه که انجام چنین آزمایشهای فرضی نیز در عمل ممکن نخواهد بود. معادلهٔ ۵ـ۲ در واقع منعکس‌کنندهٔ یک خصوصیت آماری است و اساساً قابل محاسبه نیست.

در تمام مواردی که امید ریاضی یک تخمین‌زننده، برابر مقدار واقعی پارامتر باشد، می‌گویند آن تخمین‌زننده ناااریب[1] است. بنابراین $\hat{\beta}_{OLS}$ نیز یک تخمین‌زنندهٔ ناااریب از β است. برای درک مفهوم خصوصیت ناااریبی، کافی است به نمودار ۲ـ۲ توجه کنیم. می‌دانیم تابع توزیع احتمال $\hat{\beta}$ نرمال و میانگین این توزیع برابر $E(\hat{\beta})$ است. $\hat{\beta}$ موقعی ناااریب است که این میانگین، یعنی $E(\hat{\beta})$ دقیقاً بر مقدار واقعی پارامتر، یعنی β منطبق باشد.



نمودار ۲ـ۲ خصوصیت ناااریبی $\hat{\beta}$

۲. میانگین $\hat{\alpha}$

در این قسمت، ثابت می‌کنیم که میانگین $\hat{\alpha}$ نیز برابر α است، یعنی $E(\hat{\alpha}) = \alpha$، که در واقع بیان دیگری از این واقعیت است که $\hat{\alpha}$ یک تخمین‌زنندهٔ ناااریب از α است.

با توجه به معادلهٔ ۲۶ـ۱ می‌دانیم $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$ و با توجه به مدل رگرسیون ۸ـ۱، می‌توان نوشت

$$\bar{Y} = \alpha + \beta \bar{X} + \bar{U} \quad . \qquad (۲\text{-}۶)$$

1. Unbiased

معادلهٔ ۶ـ۲ را در ۲۶ـ۱ جایگزین می‌کنیم،

$$\hat{\alpha} = \alpha + \beta\,\bar{X} + \bar{U} - \hat{\beta}\,\bar{X}\ ,$$

$$\hat{\alpha} = \alpha + (\beta - \hat{\beta})\,\bar{X} + \bar{U}\ . \qquad (۲ـ۷)$$

از دو طرف رابطهٔ فوق امید ریاضی می‌گیریم. با توجه به غیر تصادفی بودن X_i خواهیم داشت

$$E(\hat{\alpha}) = \alpha + \bar{X}\,E(\beta - \hat{\beta}) + E(\bar{U})\ .$$

فرض صفر بودن امید ریاضی $E(U_i)$ ضرورتاً ایجاب می‌کند که $E(\bar{U})$ نیز صفر شود،

$$E(\hat{\alpha}) = \alpha + \bar{X}\,[E(\beta) - E(\hat{\beta})] + ۰\ ,$$

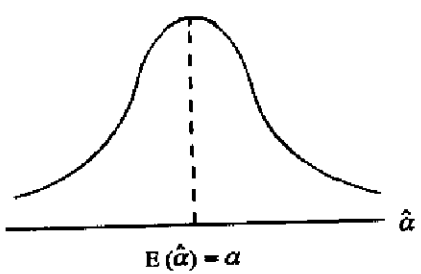
که با توجه به $E(\hat{\beta}) = \beta$، به صورت زیر نوشته می‌شود،

$$E(\hat{\alpha}) = \alpha + \bar{X}\,[\beta - \beta] + ۰\ ,$$

یا

$$E(\hat{\alpha}_{OLS}) = \alpha\ . \qquad (۲ـ۸)$$

نمودار ۳ـ۲ نشان می‌دهد که تخمین‌زنندهٔ $\hat{\alpha}$ ناریب است. ملاحظه می‌شود که مانند $\hat{\beta}$، تابع توزیع $\hat{\alpha}$ نیز نرمال است؛ زیرا با توجه به معادلهٔ ۲۶ـ۱ $\hat{\alpha}$ یک تابع خطی از $\hat{\beta}$ است، در نتیجه از توزیع $\hat{\beta}$ تبعیت می‌کند.



نمودار ۳ـ۲ خصوصیت ناریبی $\hat{\alpha}$

بعد از آشنایی با میانگین تخمین زننده‌های $\hat{\alpha}$ و $\hat{\beta}$، به بررسی واریانس آنها می‌پردازیم. مطابق معمول بحث را با واریانس $\hat{\beta}$ آغاز می‌کنیم.

۳. واریانس $\hat{\beta}$

با استفاده از تعریف واریانس، می‌دانیم

$$\mathrm{Var}\,(\hat{\beta}) = \mathrm{E}\,[\hat{\beta} - \mathrm{E}\,(\hat{\beta})]^2 \ ,$$

در معادلهٔ ۵ـ۲ دیدیم که، $\mathrm{E}\,(\hat{\beta}) = \beta$ است. با جایگزینی ۵ـ۲ در معادلهٔ فوق خواهیم داشت

$$\mathrm{Var}\,(\hat{\beta}) = \mathrm{E}\,(\hat{\beta} - \beta)^2 \ . \qquad (۲ـ۹)$$

در اینجا باید مقدار $(\hat{\beta} - \beta)$ را به دست آوریم. در معادلهٔ ٤ـ۲ دیدیم که $\hat{\beta}$ یک تابع خطی از جمله‌های اختلال است. این معادله را یک بار دیگر می‌نویسیم،

$$\hat{\beta} = \beta + w_1 U_1 + w_2 U_2 + \cdots + w_n U_n \ ,$$

یا

$$(\hat{\beta} - \beta) = w_1 U_1 + w_2 U_2 + \cdots + w_n U_n \ . \qquad (۲ـ۱۰)$$

با جایگزینی ۱۰ـ۲ در ۹ـ۲ خواهیم داشت

$$\mathrm{Var}\,(\hat{\beta}) = \mathrm{E}\,(w_1 U_1 + w_2 U_2 + \cdots + w_n U_n)^2 \ .$$

طرف راست را بسط می‌دهیم،

$$\mathrm{Var}\,(\hat{\beta}) = \mathrm{E}(w_1^2 U_1^2 + w_2^2 U_2^2 + \cdots + w_n^2 U_n^2 + 2\, w_1 w_2 U_1 U_2 + 2\, w_1 w_3 U_1 U_3 + \cdots + 2\, w_1 w_n U_1 U_n + \cdots)$$

در معادلهٔ ۳ـ۲ دیدیم که w_i مقدار ثابتی است، بنابراین

$$\mathrm{Var}\,(\hat{\beta}) = w_1^2 \mathrm{E}\,(U_1)^2 + w_2^2 \mathrm{E}\,(U_2)^2 + \cdots + w_n^2 \mathrm{E}\,(U_n)^2 + 2\, w_1 w_2 \mathrm{E}\,(U_1 U_2) + 2\, w_1 w_3 \mathrm{E}\,(U_1 U_3) + \cdots + 2\, w_1 w_n \mathrm{E}\,(U_1 U_n) + \cdots \ . \qquad (۲ـ۱۱)$$

به تعاریف واریانس و کوواریانس جملهٔ اختلال برمی‌گردیم،

$$\mathrm{Var}\,(U_t) = \mathrm{E}\left[U_t - \mathrm{E}\,(U_t)\right]^2 \ .$$

با توجه به $\mathrm{E}\,(U_t) = 0$، نتیجه می‌گیریم که

$$\mathrm{Var}\,(U_t) = \mathrm{E}\,(U_t)^2 \ . \qquad (۲\text{-}۱۲)$$

همچنین بنا بر تعریف می‌دانیم

$$\mathrm{Cov}\,(U_t\ ,\ U_s) = \mathrm{E}\left[U_t - \mathrm{E}\,(U_t)\right]\left[U_s - \mathrm{E}\,(U_s)\right] \ ,$$

و با توجه به $\mathrm{E}\,(U_t) = 0$، برای تمام مقادیر t، خواهیم داشت

$$\mathrm{Cov}\,(U_t\ ,\ U_s) = \mathrm{E}\,(U_t\,U_s) \ . \qquad (۲\text{-}۱۳)$$

معادله‌های ۲ـ۱۲ و ۲ـ۱۳ را در معادلهٔ ۲ـ۱۱ جایگزین می‌کنیم،

$$\mathrm{Var}\,(\hat{\beta}) = w_1^2\ \mathrm{Var}\,(U_1) + w_1^2\ \mathrm{Var}\,(U_2) + \cdots + w_n^2\ \mathrm{Var}\,(U_n) +$$

$$2\ w_1\ w_2\ \mathrm{Cov}\,(U_1\ ,\ U_2) + 2\ w_1\ w_3\ \mathrm{Cov}\,(U_1\ ,\ U_3) +$$

$$\cdots + 2\ w_1\ w_n\ \mathrm{Cov}\,(U_1\ ,\ U_n) + \cdots \ .$$

فرض بر این است که U_t کلاسیک است، یعنی از خصوصیات واریانس همسانی و عدم خودهمبستگی برخوردار است. به عبارت دیگر

$$\mathrm{Var}\,(U_t) = \sigma^2 \quad , \quad t = 1\ ,\ 2\ ,\ \ldots\ ,\ n \ ,$$

$$\mathrm{Cov}\,(U_t\ ,\ U_s) = 0 \ , \quad t,s = 1\ ,\ 2\ ,\ \ldots\ ,\ n \ , \quad t \neq s \ .$$

بنابراین خواهیم داشت

$$\mathrm{Var}\,(\hat{\beta}) = w_1^2\ \sigma^2 + w_2^2\ \sigma^2 + \cdots + w_n^2\ \sigma^2 + 2\ w_1\ w_2\ (0) +$$

$$۲ w_۱ w_۲ (۰) + ... + ۲ w_۱ w_n (۰) + ۰ + ۰ + ... + ۰ .$$

$$= \sigma^۲ \sum w_t^۲ . \qquad (۲\text{-}۱٤)$$

از معادلهٔ ۳ـ۲ داریم

$$w_t^۲ = \frac{x_t^۲}{A^۲} ,$$

یا

$$\sum w_t^۲ = \sum \frac{۱}{A^۲} x_t^۲ .$$

با توجه به تعریف $A = \Sigma x_t^۲$ و نظر به اینکه A مقدار ثابتی است، خواهیم داشت

$$\sum w_t^۲ = \frac{۱}{(\sum x_t^۲)^۲} \sum x_t^۲ = \frac{۱}{\sum x_t^۲} . \qquad (۲\text{-}۱٥)$$

معادلهٔ ۱۵ـ۲ را در ۱٤ـ۲ جایگزین کرده، داریم

$$\text{Var}\,(\hat{\beta}_{OLS}) = \frac{\sigma^۲}{\sum x_t^۲} . \qquad (۲\text{-}۱٦)$$

٤. واریانس $\hat{\alpha}$

از تعریف واریانس $\hat{\alpha}$ شروع می‌کنیم،

$$\text{Var}\,(\hat{\alpha}) = E\,[\hat{\alpha} - E\,(\hat{\alpha})]^۲ ,$$

که با توجه به معادلهٔ ۸ـ۲ خواهیم داشت

$$\text{Var}\,(\hat{\alpha}) = E\,(\hat{\alpha} - \alpha)^۲ .$$

باید به بررسی مقدار $(\hat{\alpha} - \alpha)$ بپردازیم. از معادلهٔ ۷ـ۲ داریم

$$(\hat{\alpha} - \alpha) = (\beta - \hat{\beta})\,\bar{X} + \bar{U} ,$$

که در فرمول $\mathrm{Var}(\hat{\alpha})$ به صورت زیر جایگزین می‌شود،

$$\mathrm{Var}(\hat{\alpha}) = \mathrm{E}[(\beta - \hat{\beta})\,\bar{X} + \bar{U}]^2 ,$$

$$= \mathrm{E}(\beta - \hat{\beta})^2\,\bar{X}^2 + 2\,\mathrm{E}[(\beta - \hat{\beta})\,\bar{X}\,\bar{U}] + \mathrm{E}(\bar{U})^2 ,$$

$$= \bar{X}^2\,\mathrm{E}(\beta - \hat{\beta})^2 + 2\,\bar{X}\,\mathrm{E}[(\beta - \hat{\beta})\,\bar{U}] + \mathrm{E}(\bar{U})^2 .$$

می‌دانیم بنا بر تعریف $\mathrm{Var}(\hat{\beta}) = \mathrm{E}(\hat{\beta} - \beta)^2$، بنابراین

$$\mathrm{Var}(\hat{\alpha}) = \bar{X}^2\,\mathrm{Var}(\hat{\beta}) + 2\,\bar{X}\,\mathrm{E}[(\beta - \hat{\beta})\,\bar{U}] + \mathrm{E}(\bar{U})^2 . \quad (۲\text{-}۱۷)$$

به بررسی مقدار $\mathrm{E}[(\beta - \hat{\beta})\,\bar{U}]$ می‌پردازیم، از معادلهٔ ۲ـ۴ داریم

$$(\hat{\beta} - \beta) = \sum w_i U_i ,$$

که در آن $w_i = \frac{x_i}{\sum x_i^2}$. یادآوری می‌کنیم که چون ثابت $= A = \sum x_i^2$؛ بنابراین

$$\sum w_i = \frac{\sum x_i}{A} = \frac{\sum x_i}{\sum x_i^2} ,$$

و چون $\sum x_i = 0$، بنابراین

$$\sum w_i = 0 . \quad (۲\text{-}۱۸)$$

در اینجا به خاصیت دیگری از w_i اشاره می‌شود. می‌توان نشان داد که $\sum w_i X_i = 1$. کافی است دو طرف معادلهٔ ۲ـ۳ را در X_i ضرب کرده،

$$w_i X_i = \frac{x_i X_i}{A} ,$$

یا

$$\sum w_i X_i = \frac{\sum x_i X_i}{A} = \frac{\sum x_i^2}{A} = \frac{\sum x_i^2}{\sum x_i^2} ,$$

$$\sum w_t X_t = ۱ \ .$$

بعد از ذکر این مقدمه به بررسی مقدار $\bar{U}\,(\beta - \hat{\beta})$ می‌رسیم، می‌دانیم

$$(\beta - \hat{\beta})\,\bar{U} = (\sum_{t=۱}^{n} w_t U_t)\,\frac{۱}{n}\sum_{t=۱}^{n} U_t = \frac{۱}{n}\sum w_t U_t \sum U_t \ ,$$

$$= \frac{۱}{n}\left[(w_۱ U_۱ + w_۲ U_۲ + \cdots + w_n U_n)\,(U_۱ + U_۲ + \cdots + U_n)\right] \ ,$$

$$= \frac{۱}{n}\left[U_۱ (w_۱ U_۱ + w_۲ U_۲ + \cdots + w_n U_n) + U_۲ (w_۱ U_۱ + \cdots + w_n U_n) + \cdots\right] .$$

در نتیجه با گرفتن امید ریاضی خواهیم داشت

$$E(\beta - \hat{\beta})\,\bar{U} = \frac{۱}{n}\left[w_۱ E(U_۱)^۲ + w_۲ E(U_۱ U_۲) + \cdots + w_n E(U_۱ U_n) + \right.$$

$$\left. w_۱ E(U_۱ U_۲) + w_۲ E(U_۲)^۲ + \cdots + w_n E(U_۲ U_n) + \cdots\right] .$$

با فرض واریانس همسانی $E(U_t)^۲ = \sigma^۲$ و عدم خودهمبستگی $E(U_t U_s)$، داریم

$$E(\beta - \hat{\beta})\,\bar{U} = \frac{۱}{n}\left[w_۱ \sigma^۲ + ۰ + \cdots + ۰ + ۰ + W_۲ \sigma^۲ + ۰ + \cdots + ۰ + \cdots \right.$$

$$\left. + w_n \sigma^۲\right] \ ,$$

$$= \frac{۱}{n}\left[\sum w_t \sigma^۲\right] .$$

با توجه به ثابت بودن $\sigma^۲$ داریم

$$E(\beta - \hat{\beta})\,\bar{U} = \frac{۱}{n}\sigma^۲ \sum w_t \ ,$$

و با استفاده از معادلهٔ ۱۸ـ۲ داریم،

$$E(\beta - \hat{\beta})\,\bar{U} = ۰ \ . \qquad (۲\text{ـ}۱۹)$$

مقدار $E(\bar{U}_t)^2$ را حساب می‌کنیم،

$$E(\bar{U}_t)^2 = E\left(\frac{\sum U_t}{n}\right)^2 = \frac{1}{n^2} E(\sum U_t)^2 ,$$

$$= \frac{1}{n^2} E(U_1 + U_2 + \cdots + U_n)^2 ,$$

$$= \frac{1}{n^2} E(U_1^2 + U_2^2 + \cdots + U_n^2 + 2U_1U_2 + 2U_1U_3 + \cdots) .$$

طبق فرض واریانس همسانی و عدم خودهمبستگی، داریم

$$= \frac{1}{n^2} (\sigma^2 + \sigma^2 + \cdots + \sigma^2 + 0 + 0 + \cdots + 0) .$$

در نتیجه

$$E(\bar{U}_t)^2 = \frac{\sigma^2}{n} . \qquad (۲\text{-}۲۰)$$

معادله‌های ۱۹-۲ و ۲۰-۲ را در ۱۷-۲ جایگزین می‌کنیم،

$$\text{Var}(\hat{\alpha}) = \bar{X}^2 \text{Var}(\hat{\beta}) + 0 + \frac{\sigma^2}{n} .$$

معادلهٔ ۱۶-۲ را در معادلهٔ فوق جایگزین می‌کنیم خواهیم داشت

$$\text{Var}(\hat{\alpha}) = \bar{X}^2 \frac{\sigma^2}{\sum x_t^2} + \frac{\sigma^2}{n} ,$$

در نتیجه

$$\text{Var}(\hat{\alpha}) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum x_t^2}\right] . \qquad (۲\text{-}۲۱)$$

در بعضی موارد، واریانس $\hat{\alpha}$ را به صورت دیگری نیز ارائه می‌کنند. از معادلهٔ ۲۱-۲ داریم

$$\text{Var}(\hat{\alpha}) = \frac{n\bar{X}^2\sigma^2 + \sigma^2\sum x_i^2}{n\sum x_i^2} = \frac{\sigma^2}{n\sum x_i^2}\left[n\bar{X}^2 + \sum x_i^2\right] .$$

نشان می‌دهیم که $n\bar{X}^2 + \sum x_i^2 = \sum X_i^2$. برای این منظور $\sum x_i^2$ را چنین می‌نویسیم،

$$\sum x_i^2 = \sum (X_i - \bar{X})^2 = \sum X_i^2 + n\bar{X}^2 - 2\bar{X}\sum X_i ,$$

$$= \sum X_i^2 - n\bar{X}^2 .$$

بدین ترتیب

$$n\bar{X}^2 + \sum x_i^2 = n\bar{X}^2 + \sum X_i^2 - n\bar{X}^2 ,$$

$$= \sum X_i^2 .$$

رابطهٔ فوق را در فرمول $\text{Var}(\hat{\alpha})$ جایگزین می‌کنیم،

$$\text{Var}(\hat{\alpha}) = \frac{\sigma^2}{n\sum x_i^2}\left[\sum X_i^2\right] ,$$

یا

$$\text{Var}(\hat{\alpha}) = \sigma^2\left(\frac{\sum X_i^2}{n\sum x_i^2}\right) . \qquad (a\ 2\text{-}21)$$

۵. روش دیگری برای استخراج واریانسهای $\hat{\beta}$ و $\hat{\alpha}$

با اینکه روش فوق در به دست آوردن واریانس $\hat{\beta}$ و $\hat{\alpha}$، یعنی فرمولهای ۱۶-۲ و ۲۱-۲ تا حدی طولانی است، از نظر مفهومی بسیار جالب و هماهنگ با روش استخراج واریانس پارامترها در مدل رگرسیون چند متغیره است. البته روش ساده‌تری نیز وجود دارد که در زیر به آن اشاره می‌کنیم. این روش به این بستگی دارد که ابتدا ثابت کنیم $\hat{\beta}$ و $\hat{\alpha}$، یک ترکیب خطی از مشاهدات Y_i است. این قضیه را ابتدا برای $\hat{\beta}$ و سپس برای $\hat{\alpha}$ ثابت می‌کنیم.

قضیهٔ ۱. ثابت کنید $\hat{\beta}_{OLS}$ یک ترکیب خطی از Y_t است.

معادلهٔ ۲۵ـ۱ را می‌نویسیم،

$$\hat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2} .$$

رابطهٔ $y_t = Y_t - \bar{Y}$ را جایگزین می‌کنیم،

$$\hat{\beta} = \frac{\sum x_t (Y_t - \bar{Y})}{\sum x_t^2} = \frac{\sum x_t Y_t - \bar{Y} \sum x_t}{\sum x_t^2} .$$

با توجه به $\sum x_t = 0$، خواهیم داشت

$$\hat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2} . \qquad (۲\text{ـ}۲۲)$$

طرف راست فرمول ۲۲ـ۲ را بسط داده، با فرض $A = \sum x_t^2$ داریم

$$\hat{\beta} = \frac{x_1}{A} Y_1 + \frac{x_2}{A} Y_2 + \cdots + \frac{x_n}{A} Y_n .$$

از معادلهٔ ۳ـ۲ استفاده می‌کنیم،

$$\hat{\beta}_{OLS} = w_1 Y_1 + w_2 Y_2 + \cdots + w_n Y_n = \sum w_t Y_t . \qquad (۲\text{ـ}۲۳)$$

با توجه به اینکه ضرایب w_t ثابت است، معادلهٔ ۲۳ـ۲ نشان می‌دهد که $\hat{\beta}_{OLS}$ یک ترکیب خطی از مشاهدات Y_t است.

قضیهٔ ۲. ثابت کنید $\hat{\alpha}_{OLS}$ یک ترکیب خطی از Y_t است.

معادلهٔ ۲۶ـ۱ را دوباره می‌نویسیم،

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X} .$$

طبق معادلهٔ ۲۳ـ۲ می‌دانیم، $\hat{\beta} = \sum w_t Y_t$ که در آن $w_t = \frac{x_t}{\sum x_t^2}$.

بنابراین خواهیم داشت

$$\hat{\alpha} = \bar{Y} - \bar{X} \sum w_t Y_t \ ,$$

$$= \frac{\sum Y_t}{n} - \bar{X} \sum w_t Y_t \ ,$$

و در نتیجه

$$\hat{\alpha}_{OLS} = \sum \left[\frac{1}{n} - \bar{X} w_t\right] Y_t \ . \qquad (۲\text{-}۲۴)$$

معادلهٔ ۲۴ـ۲ نشان می‌دهد که $\hat{\alpha}$ یک ترکیب خطی از Y_t است که $(\frac{1}{n} - \bar{X} w_t)$ در واقع ضرایب ثابت این ترکیب خطی است. در زیر به بحث واریانس $\hat{\beta}$ و $\hat{\alpha}$ می‌پردازیم.

روش دیگری برای واریانس $\hat{\beta}$

طبق معادلهٔ ۲۳ـ۲ می‌دانیم $\hat{\beta}$ یک تابع خطی از Y_t است،

$$\hat{\beta} = \sum w_t Y_t \ ,$$

که در آن w_t ضرایب ثابت بوده و از Y_t مستقل هستند. از دو طرف معادلهٔ ۲۳ـ۲ واریانس می‌گیریم،

$$\text{Var}(\hat{\beta}) = \text{Var}(\hat{\beta} \sum w_t Y_t) = \sum w_t^2 \text{Var}(Y_t) \ .$$

مطابق معادلهٔ ۱۴ـ۱ می‌دانیم، $\text{Var}(Y_t) = \sigma^2$. در نتیجه

$$\text{Var}(\hat{\beta}) = \sum w_t^2 \sigma^2 \ ,$$

$$= \sigma^2 \sum w_t^2 \ .$$

با استفاده از معادلهٔ ۱۵ـ۲ خواهیم داشت

$$\text{Var}(\hat{\beta}_{OLS}) = \frac{\sigma^2}{\sum x_t^2} \ ,$$

که دقیقاً همان معادلهٔ ۱۶ـ۲ است.

روش دیگری برای واریانس $\hat{\alpha}$

طبق معادلهٔ ۲۴ـ۲ می‌دانیم $\hat{\alpha}$ یک تابع خطی از Y_t است،

$$\hat{\alpha} = \sum [\frac{1}{n} - \bar{X}\, w_t]\, Y_t \ .$$

از دو طرف این معادله واریانس می‌گیریم،

$$\mathrm{Var}\,(\hat{\alpha}) = \mathrm{Var}\,[\sum (\frac{1}{n} - \bar{X}\, w_t)\, Y_t] \ ,$$

$$= \sum (\frac{1}{n} - \bar{X}\, w_t)^2\, \mathrm{Var}\, Y_t \ .$$

از معادلهٔ ۱۴ـ۱ می‌دانیم $\mathrm{Var}\,(Y_t) = \sigma^2$. در نتیجه

$$\mathrm{Var}\,(\hat{\alpha}) = \sigma^2 \sum (\frac{1}{n^2} - \frac{2\bar{X}\, w_t}{n} + \bar{X}^2\, w_t^2) \ .$$

با توجه به معادلهٔ ۱۸ـ۲ ($\Sigma\, w_t = 0$) و معادلهٔ ۱۵ـ۲ ($\Sigma\, w_t^2 = \frac{1}{\sum x_t^2}$) و با در نظرگرفتن $\Sigma \frac{1}{n^2} = \frac{1}{n}$ خواهیم داشت

$$\mathrm{Var}\,(\hat{\alpha}) = \sigma^2\, [\frac{1}{n} - \frac{2\bar{X}}{n} \sum w_t + \bar{X}^2 \sum w_t^2] \ ,$$

$$= \sigma^2\, [\frac{1}{n} + \frac{\bar{X}^2}{\sum x_t^2}] \ ,$$

که دقیقاً همان معادلهٔ ۲۱ـ۲ است. مانند گذشته می‌توان به فرمول دیگری برای $\mathrm{Var}\,(\hat{\alpha})$ نیز رسید. می‌دانیم

$$\mathrm{Var}\,(\hat{\alpha}) = \sigma^2\, [\frac{\sum x_t^2 + n\, \bar{X}^2}{n \sum x_t^2}] \ .$$

با توجه به

$$\sum x_t^2 = \sum (X_t - \bar{X})^2 = \sum X_t^2 - n\, \bar{X}^2 \ ,$$

خواهیم داشت

$$\text{Var}\,(\hat{\alpha}_{\text{OLS}}) = \sigma^2\,(\frac{\sum X_i^2}{n\sum x_i^2})\ .$$

با مفاهیم و فرمولهای میانگین و واریانس $\hat{\alpha}$ و $\hat{\beta}$ آشنا شدیم. قبل از پایان مبحث خصوصیات آماری تخمین‌زننده‌های حداقل مربعات معمولی، کوواریانس $\hat{\alpha}$ و $\hat{\beta}$ را مطرح نموده و به یک نکته در مورد نرمال بودن تابع توزیع احتمال $\hat{\alpha}$ و $\hat{\beta}$ اشاره می‌کنیم.

۶. کواریانس $\hat{\alpha}$ و $\hat{\beta}$

از تعریف کواریانس $\hat{\alpha}$ و $\hat{\beta}$ آغاز می‌کنیم،

$$\text{Cov}\,(\hat{\alpha}\,,\hat{\beta}) = \text{E}\,[\hat{\alpha} - \text{E}\,(\hat{\alpha})]\,[\hat{\beta} - \text{E}\,(\hat{\beta})]\ .$$

با استفاده از معادله‌های ۲ـ۵ و ۲ـ۸ داریم

$$\text{Cov}\,(\hat{\alpha}\,,\hat{\beta}) = \text{E}\,(\hat{\alpha} - \alpha)\,(\hat{\beta} - \beta)\ .$$

از معادلهٔ ۲ـ۷ داریم

$$(\hat{\alpha} - \alpha) = (\hat{\beta} - \beta)\,\bar{X} + \bar{U}\ ,$$

در نتیجه

$$\text{Cov}\,(\hat{\alpha}\,,\hat{\beta}) = \text{E}\,[(\beta - \hat{\beta})\,\bar{X} + \bar{U}]\,(\hat{\beta} - \beta)\ ,$$

$$= \text{E}\,[(\beta - \hat{\beta})\,(\hat{\beta} - \beta)\,\bar{X} + \bar{U}\,(\hat{\beta} - \beta)]\ ,$$

$$= \bar{X}\,\text{E}\,[(\beta - \hat{\beta})\,(\hat{\beta} - \beta)] + \text{E}\,[\bar{U}\,(\hat{\beta} - \beta)]\ ,$$

$$= -\,\bar{X}\,\text{E}\,(\hat{\beta} - \beta)^2 + \text{E}\,[\bar{U}\,(\hat{\beta} - \beta)]\ .$$

از معادلهٔ ۲ـ۱۹ می‌دانیم

$$\text{E}\,[\bar{U}\,(\hat{\beta} - \beta)] = 0\ ,$$

بنابراین

$$\mathrm{Cov}(\hat{\alpha}, \hat{\beta}) = -\bar{X}\, E(\hat{\beta} - \beta)^2 + 0 ,$$

و با توجه به $\mathrm{Var}(\hat{\beta}) = E(\hat{\beta} - \beta)^2$ و با استفاده از معادلهٔ ۲ـ۱۶ خواهیم داشت

$$\mathrm{Cov}(\hat{\alpha}, \hat{\beta}) = -\bar{X}\, \mathrm{Var}(\hat{\beta}) ,$$

یا

$$\mathrm{Cov}(\hat{\alpha}, \hat{\beta}) = \frac{-\bar{X}\sigma^2}{\sum x_i^2} . \qquad (۲ـ۲۵)$$

نتیجه می‌گیریم که $\hat{\alpha}$ و $\hat{\beta}$ دارای توزیع نرمال با خصوصیات زیر هستند،

$$\hat{\alpha}_{OLS} \sim N\left(\alpha, \frac{\sigma^2 \sum X_i^2}{n \sum x_i^2}\right) , \qquad (۲ـ۲۶)$$

$$\hat{\beta}_{OLS} \sim N\left(\beta, \frac{\sigma^2}{\sum x_i^2}\right) , \qquad (۲ـ۲۷)$$

یک بار دیگر، به بحث نرمال بودن تابع توزیع احتمال $\hat{\beta}$ و $\hat{\alpha}$ برمی‌گردیم. چون $\hat{\beta}$ در معادلهٔ ۲ـ۲۲ یک ترکیب خطی از Y_i است و با توجه به اینکه Y_i در معادلهٔ ۲ـ۸ تابع خطی از U_i است؛ اگر U_i دارای توزیع نرمال باشد، Y_i نیز نرمال بوده و در نتیجه توزیع $\hat{\beta}$ نرمال خواهد بود. $\hat{\alpha}$ نیز طبق معادلهٔ ۲ـ۲۶ تابع خطی از $\hat{\beta}$ است و بنابراین توزیع نرمال دارد. حال اگر توزیع U_i نرمال نباشد، Y_i توزیع نرمال نخواهد داشت؛ با وجود این، با استفاده از قضیهٔ حد مرکزی[1] می‌توان گفت که $\hat{\beta}$ برای نمونه‌های بزرگ توزیع نرمال دارد. قضیهٔ حد مرکزی به طور بسیار ساده، یعنی اگر حجم یک نمونه به طور مرتب زیاد شود، توزیع میانگین این نمونه به سمت نرمال میل می‌کند. $\hat{\beta}$ و $\hat{\alpha}$ نیز طبق معادله‌های ۲ـ۲۳ و ۲ـ۲۴ یک نوع میانگین وزنی از Y_i هستند؛ زیرا در واقع یک ترکیب خطی از Y_i هستند. در مورد $\hat{\beta}$ و $\hat{\alpha}$ این وزنه‌ها به ترتیب w_i و $(\frac{1}{n} - \bar{X} w_i)$ هستند. بنابراین ملاحظه

1. Central Limit Theorem

می‌شود که روش حداقل مربعات معمولی خصوصیت بسیار برجسته‌ای دارد، بدین معنی که حتی اگر جمله‌های اختلال دارای توزیع نرمال نباشد، تخمین‌زننده‌های $\hat{\alpha}$ و $\hat{\beta}$ برای نمونه‌های بزرگ توزیع نرمال دارند.

۷. تخمین واریانس جملهٔ اختلال

واریانس $\hat{\beta}$ (معادلهٔ ۱۶ـ۲)، واریانس $\hat{\alpha}$ (معادلهٔ ۲۱ـ۲) و کوواریانس $\hat{\alpha}$ و $\hat{\beta}$ (معادلهٔ ۲۵ـ۲)، همگی تابعی از واریانس جملهٔ اختلال (U_i) هستند. به طور قطع ما مقدار واقعی واریانس U_i را نمی‌دانیم، چون مقادیر U_i قابل مشاهده نیستند؛ بنابراین باید واریانس U_i (σ_U^2) را تخمین بزنیم. یادآوری می‌کنیم که σ_U^2 اگر هم بدون اندیس نوشته شود به معنای واریانس U_i است. بدیهی است اگر σ^2 اندیس دیگری به غیر از U داشته باشد بر واریانس همان اندیس دلالت می‌کند.

برای محاسبهٔ واریانس U_i، ضرورتاً باید U_i را داشته باشیم. اگر مقادیر U_i موجود نباشد، باید متغیری که بیشترین نزدیکی را با U_i دارد، در نظر گرفته و واریانس آن را به عنوان تقریبی از واریانس U_i به حساب آوریم. در اقتصادسنجی معمول است که e_i را به عنوان چنین متغیری معرفی می‌کنند؛ زیرا از بین متغیرهای موجود در یک سیستم رگرسیون خطی، هیچ متغیر دیگری نزدیکتر به U_i یافت نمی‌شود.

بدین ترتیب، کافی است بعد از تخمین پارامترهای مدل رگرسیون، جمله‌های پسماند را به دست آورده، واریانس آن را محاسبه کنیم تا تخمینی از واریانس U_i به دست آید. بنا بر تعریف، واریانس e_i عبارت است از

$$\text{Var}(e_i) = S^2 = \frac{\sum (e_i - \bar{e})^2}{n}$$

با توجه به معادلهٔ ۳۸ـ۱ می‌دانیم $\bar{e} = 0$، پس

$$\text{Var}(e_i) = S^2 = \frac{\sum e_i^2}{n}$$

بنابراین، باید واریانس e_i را که به ترتیب فوق به دست آمده است به عنوان واریانس U_i

بپذیریم. اما در قسمت ۵ـ۱ و بدون اثبات، دیدیم که برای رسیدن به تخمین ناااریب از واریانس U_t، باید مجموع مربعات پسماند را بر درجات آزادی آن تقسیم کنیم،

$$\hat{\text{Var}}\,(U_t) = \hat{\sigma}^2 = \frac{\sum e_t^2}{n-2} \ .$$

یادآوری می‌کنیم که برای محاسبهٔ $\sum e_t^2$ باید ابتدا $\hat{\alpha}$ و $\hat{\beta}$ به دست آید و این امر به آن معنی است که در محاسبهٔ $\sum e_t^2$ دو درجهٔ آزادی از دست داده‌ایم. به همین دلیل است که برای رسیدن به تخمین ناااریب از σ^2، مجموع مربعات پسماند را به جای n، بر $(n-2)$ تقسیم کرده‌ایم. در اینجا باید ثابت کنیم که $\frac{\sum e_t^2}{n-2}$ تخمین ناااریبی از σ^2 است.

برای اثبات، به ترتیب زیر عمل می‌کنیم. ابتدا در تعریف e_t، به جای Y_t و $\hat{Y}_t$ مقادیر آن را قرار می‌دهیم،

$$\begin{aligned} e_t &= Y_t - \hat{Y}_t \ , \\ &= \alpha + \beta X_t + U_t - \hat{\alpha} + \hat{\beta} X_t \ , \\ &= U_t - (\alpha - \hat{\alpha}) - (\hat{\beta} - \beta) X_t \ . \end{aligned}$$

مقدار $(\alpha - \hat{\alpha})$ را به ترتیب زیر به دست می‌آوریم. می‌دانیم

$$\begin{aligned} \hat{\alpha} &= \bar{Y} - \hat{\beta} \bar{X} \ , \\ &= (\alpha + \beta \bar{X} + \bar{U}) - \hat{\beta} \bar{X} \ , \\ &= \alpha - (\hat{\beta} - \beta) \bar{X} + \bar{U} \ , \end{aligned}$$

در نتیجه

$$(\hat{\alpha} - \alpha) = \bar{U} - (\hat{\beta} - \beta) \bar{X} \ .$$

بدین ترتیب e_t را می‌توان به صورت زیر نوشت،

$$e_t = (U_t - \bar{U}) - (\hat{\beta} - \beta) x_t \ .$$

با اینکه $E(U_t) = 0$، اما توجه داریم که $\bar{U}$ صفر نیست؛ زیرا میانگین نمونه است.

دو طرف رابطهٔ فوق را مجذور کرده و جمع می‌کنیم،

$$\sum e_t^2 = \sum (U_t - \bar{U})^2 + (\hat{\beta} - \beta) \sum x_t^2 - 2 (\hat{\beta} - \beta) \sum (U_t - \bar{U}) x_t \ ,$$

$$= \sum U_t^2 - n \bar{U}^2 + (\hat{\beta} - \beta)^2 \sum x_t^2 - 2 (\hat{\beta} - \beta) \sum U_t x_t \ .$$

از دو طرف رابطهٔ فوق امید ریاضی می‌گیریم،

$$E(e_t^2) = E(\sum U_t^2) - n E(\bar{U}^2) + E(\hat{\beta} - \beta)^2 \sum x_t^2$$

$$- 2 E[(\hat{\beta} - \beta) \sum U_t x_t] \ .$$

اما می‌دانیم که

$$E(\sum U_t^2) = \sum E(U_t^2) = \sum E[U_t - E(U_t)]^2 = n \sigma_U^2 \ ,$$

$$E(\bar{U})^2 = Var(\bar{U}) = \frac{\sigma_U^2}{n} \ ,$$

$$E(\hat{\beta} - \beta)^2 = E[\hat{\beta} - E(\hat{\beta})]^2 = Var(\hat{\beta}) = \frac{\sigma_U^2}{\sum x_t^2} \ .$$

برای محاسبهٔ جملهٔ آخر، از معادلهٔ ۱۰ـ۲ استفاده می‌کنیم،

$$E[(\hat{\beta} - \beta) \sum U_t x_t] = E[(\sum w_t U_t) (\sum U_t x_t)] \ ,$$

$$= \sigma_U^2 \sum w_t x_t \ .$$

با توجه به معادله‌های ۳ـ۲ و ۱۹ـ۲ می‌دانیم $\sum w_t x_t = 1$؛ بنابراین

$$E[(\hat{\beta} - \beta) \sum U_t x_t] = \sigma_U^2 \ .$$

بدین ترتیب $E(e_t^2)$ به صورت زیر نوشته می‌شود،

$$E(\sum e_t^2) = n \sigma_U^2 - \sigma_U^2 + \sigma_U^2 - 2 \sigma_U^2 = (n - 2) \sigma^2 \ ,$$

پس

$$E\left(\frac{\sum e_t^2}{n-2}\right) = \sigma_U^2 \ .$$

نتیجه می‌گیریم که برای تخمین واریانس جملهٔ اختلال (σ_U^2) باید واریانس پسماندها، یعنی

$$\sigma_e^2 = Var\ (e_i) = \frac{\sum e_i^2}{n}$$

را به دست آورد. اما برای اینکه واریانس جمله‌های پسماند بتواند تخمین نااریبی از واریانس جملهٔ اختلال را نتیجه دهد، باید مجموع مربعات پسماند (Σe_i^2) را بر $n-2$ که در واقع همان درجات آزادی است تقسیم کنیم. بنابراین

$$S^2 = Var\ (e_i) = \sigma_e^2 = \hat{\sigma}_U^2 = \hat{Var}\ (U_i) = \frac{\Sigma e_i^2}{n-2} \ . \qquad (۲\text{ـ}۲۸)$$

در معادلهٔ ۱ـ۴۵ نیز جذر رابطهٔ فوق را به دست آورده، آن را خطای معیار تخمین یا SEE نامیدیم. با توجه به معادلهٔ ۲ـ۲۸ جذر واریانس U_i را می‌توان خطای معیار رگرسیون[1] نیز نامید و آن را با SER (خطای معیار رگرسیون) نشان داد

$$SER = \sqrt{Var\ (U_i)} = \hat{\sigma}_U \ .$$

مثال ۱ـ۲ مدل رابطهٔ تولید با ساعتهای نیروی کار. موضوع مثال ۲ـ۱ را در نظر می‌گیریم،

$$Q_i = \alpha + \beta\ L_i + U_i \ .$$

الف) واریانس U_i را به دست آورید.

ب) واریانس $\hat{\alpha}$ و واریانس $\hat{\beta}$ را تخمین بزنید.

ج) خطای معیار $\hat{\alpha}$ و خطای معیار $\hat{\beta}$ را حساب کنید.

د) کوواریانس $\hat{\alpha}$ و $\hat{\beta}$ را محاسبه کنید.

---

1. Standard Error of Regression

در مثال ۲ـ۱ ، تخمین مدل فوق را به صورت زیر به دست آوردیم،

$$\hat{Q}_i = ۳/۶ + ۰/۷۵\, L_i \ ,$$

و با فرض $Y_i = Q_i$ و $X_i = L_i$ داریم

$$\hat{Y}_i = ۳/۶ + ۰/۷۵\, X_i \ .$$

همچنین نتایج زیر را نیز به دست آوردیم،

$$\sum x_i^۲ = ۲۸ \ , \qquad \bar{X} = ۸ \ ,$$

$$\sum x_i\, y_i = ۲۱ \ , \qquad \bar{Y} = ۹/۶ \ .$$

در مثال ۱ـ٤ نیز نتایج زیر را داشتیم،

$$\sum y_i^۲ = ۳۰/٤ \qquad , \qquad \sum e_i^۲ = ۱٤/٦۵ \ .$$

الف) برای تخمین واریانس جملهٔ اختلال، از معادلهٔ ۲ـ۲۸ استفاده می‌کنیم،

$$\hat{Var}\,(U_i) = \hat{\sigma}_U^۲ = \frac{\Sigma\, e_i^۲}{n - ۲} \ ,$$

$$\hat{\sigma}_U^۲ = \frac{۱٤/٦۵}{۱۰ - ۲} = ۱/۸۳ \ .$$

ب) فرمول ۲ـ۲۱ را می‌نویسیم،

$$Var\,(\hat{\alpha}) = \sigma^۲ \left[\frac{۱}{n} + \frac{\bar{X}^۲}{\Sigma\, x_i^۲}\right] \ .$$

چون مقدار $\sigma^۲$ را نمی‌دانیم، باید از مقدار تخمین آن، یعنی $\hat{\sigma}^۲$ استفاده کنیم. در آن صورت تخمین واریانس $\hat{\alpha}$ عبارت خواهد بود از

$$\hat{Var}\,(\hat{\alpha}) = \hat{\sigma}^۲ \left[\frac{۱}{n} + \frac{\bar{X}^۲}{\Sigma\, x_i^۲}\right] \ ,$$

یا

$$\hat{Var}\,(\hat{\alpha}) = ۱/۸۳ \left[\frac{۱}{۱۰} + \frac{٦٤}{۲۸}\right] = ٤/۳٦۵ \ .$$

برای تخمین واریانس $\hat{\beta}$ نیز به همین ترتیب عمل می‌کنیم،

$$\hat{\text{Var}}(\hat{\beta}) = \frac{\hat{\sigma}^2}{\Sigma x_i^2} ,$$

$$= \frac{۱/۸۳}{۲۸} = ۰/۰۶۵ .$$

ج) برای محاسبهٔ خطای معیار $\hat{\alpha}$ و $\hat{\beta}$ کافی است از واریانس آنها جذر بگیریم،

$$\text{SE}(\hat{\alpha}) = \sqrt{\text{Var}(\hat{\alpha})} ,$$

$$= \sqrt{۴/۳۶۵} = ۲/۰۹ .$$

$$\text{SE}(\hat{\beta}) = \sqrt{\text{Var}(\hat{\beta})} ,$$

$$= \sqrt{۰/۰۶۵} = ۰/۲۵۶ .$$

معمولاً مقادیر خطای معیار را در پرانتز و زیر تخمین پارامتر مربوط به هر یک می‌نویسند؛ بنابراین

$$\hat{Q}_i = ۳/۶ + ۰/۷۵ L_i .$$

$$(۲/۰۹) \quad (۰/۲۵۶)$$

د) فرمول کواریانس $\hat{\alpha}$ و $\hat{\beta}$ را از معادلهٔ ۲۵ـ۲ می‌نویسیم،

$$\text{Cov}(\hat{\alpha}, \hat{\beta}) = \frac{-\bar{X}\hat{\sigma}^2}{\Sigma x_i^2} ,$$

$$= \frac{-۸(۱/۸۳)}{۲۸} = -۰/۵۲۳ .$$

در پایان، دو نکته را مطرح می‌کنیم: اول اینکه، خطای معیار تخمین پارامترها $\text{SE}(\hat{\alpha})$ یا $\text{SE}(\hat{\beta})$ را نباید با خطای معیار تخمین رگرسیون (SER)، که به آن خطای معیار تخمین (SEE) نیز گفته می‌شود، اشتباه کرد. مفهوم اول، پراکندگی $\hat{\alpha}$ و $\hat{\beta}$ را در حول α و β اندازه گیری می‌کند و مفهوم دوم، معیاری برای اندازه گیری پراکندگی جمله‌های پسماند

است. دوم اینکه، وقتی $\text{Cov}(\hat{\alpha}$ و $\hat{\beta})$ منفی شود، می‌توان نتیجه گرفت که اگر $\hat{\alpha}$ بیشتر از اندازه تخمین زده شود، $\hat{\beta}$ کمتر از میزان تخمین می‌خورد و برعکس. این نتیجه در تمام مواردی که $\bar{X}$ مثبت باشد صادق است؛ زیرا $\hat{\sigma}^2$ و Σx_i^2 همواره مثبت‌هستند.

۲.۳ آزمون فرضیه برای هر یک از پارامترها

می‌دانیم $\hat{\alpha}$ و $\hat{\beta}$ دقیقاً برابر α و β واقعی نیست. سؤال این است که با داشتن $\hat{\alpha}$ و $\hat{\beta}$ چگونه می‌توان نسبت به مقادیر α و β واقعی استنتاج آماری کرد. باید پاسخ را در مباحث آزمون فرضیه[1] یافت. با استفاده از توابع توزیع احتمال $\hat{\alpha}$ و $\hat{\beta}$ و با دانستن مقادیر میانگین و واریانس آنها به راحتی می‌توان فرضیه‌های مختلف دربارهٔ α و β واقعی را آزمون کرد. برای سهولت، بحث را در مورد β متمرکز می‌کنیم. به همین صورت نیز می‌توان برای α استنتاج کرد. در این قسمت به ترتیب آزمونهای Z و t برای α و β و آزمون χ^2 برای σ^2 مطرح می‌شود. فرض بر این است که خواننده با مفاهیم مقدماتی آزمونهای آماری به طور خلاصه آشنایی دارد.

۱. آزمون Z برای پارامترها

هرگاه تابع توزیع احتمال $\hat{\beta}$ نرمال باشد، می‌توان از آزمون Z استفاده کرد. برای استفاده از آزمون Z لازم است که واریانس جملهٔ اختلال، یعنی $\text{Var}(U_i) = \sigma_U^2$ را بدانیم. معمولاً مقدار واقعی واریانس U_i، یعنی σ_U^2، مشخص نیست؛ بلکه تخمینی از آن در دست است. همان گونه که در بحث از آزمون t نشان خواهیم داد، موقعی می‌توان از تخمین واریانس U_i، یعنی $\hat{\sigma}_U^2$ در آزمون Z استفاده کرد که حجم نمونه بزرگ باشد و معمولاً $n > ۲۵$ را به عنوان نمونه‌های بزرگ می‌شناسیم. در ادامهٔ بحث آزمون Z، فرض بر این است که مقدار واقعی واریانس U_i را می‌دانیم.

در آزمون فرضیه‌های مختلف در مورد پارامترهای α یا β معمولاً با استفاده از

---

1. Test of Hypothesis

یک نمونهٔ n تایی، β را تخمین زده و برای مثال، داریم: $\hat{\beta} = ۲/۵$. می‌دانیم مقدار واقعی β اصولاً باید با این مقدار متفاوت باشد. سؤال می‌کنیم که با $\hat{\beta} = ۲/۵$، آیا می‌توان این فرضیه را قبول کرد که β واقعی، برای مثال، برابر ۲/۲۵ باشد؟ معمولاً فرضیه‌ای را که می‌خواهیم آزمون کنیم، فرضیهٔ «صفر، پوچ، تهی یا عدم»[1] نامیده و آن را با $H_۰$ نشان می‌دهیم. در مقابل، خلاف آن فرضیه را فرضیهٔ رقیب[2] می‌خوانیم و آن را با $H_۱$ مشخص می‌سازیم. در مورد این مثال داریم

$$H_۰ : \beta = ۲/۲۵ \quad , \qquad H_۱ : \beta \neq ۲/۲۵ \quad .$$

در حقیقت اگر این فرضیه را قبول کنیم یا به طور دقیقتر، نتوانیم رد کنیم که مقدار واقعی β بیشتر یا کمتر از ۲/۲۵ است، فرضیهٔ $H_۰$ قابل قبول نخواهد بود. با توجه به اینکه در این مثال مقادیر بیشتر یا کمتر از ۲/۲۵ ، مبنای استنتاج در آزمون $H_۰$ است، می‌گوییم که آزمون دو طرفه است. در موارد دیگر که فقط یک دامنه مورد توجه است، آزمون، یکطرفه خواهد بود.

خصوصیت تصادفی بودن متغیرها ایجاب می‌کند که در این گونه مسائل باید آزمون را در سطح احتمال معینی انجام داد؛ به عبارت دیگر، باید گفت که اگر مثلاً $\hat{\beta}$ برابر ۲/۵ است با ۹۵ درصد احتمال این فرضیه را آزمون کنید که مقدار واقعی β برابر ۲/۲۵ باشد.

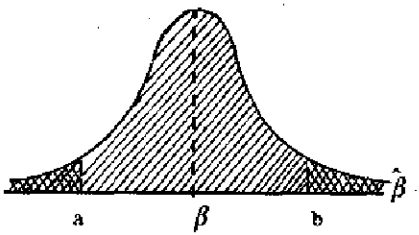
ماهیت اساسی این گونه آزمونها چنین است. این سؤال که اگر $\hat{\beta}$ برابر ۲/۵ باشد، آیا مقدار واقعی β برابر ۲/۲۵ است یا خیر؛ به این بر می‌گردد که بگوییم یک جامعه با پارامتری به نام β داریم که مقدار آن ۲/۲۵ است. یک نمونه با آماره‌ای[3] به نام $\hat{\beta}$ داریم که مقدار آن ۲/۵ است. با احتمال ۹۵ درصد این فرضیه را آزمون کنید، که این نمونه متعلق به آن جامعه باشد.

---

۱. Null Hypothesis البته این اصطلاح، به جز در مورد $\beta = ۰$ چندان روشن‌کننده نیست. متأسفانه فرضیه مثلاً $H_۰ : \beta = ۲۰$ نیز فرضیهٔ صفر نامیده می‌شود. این اصطلاح را نیمن و پیرسون ـ که از آماردانان معروفند ـ در دههٔ ۱۹۳۰ معرفی کرده‌اند.

2. Alternative Hypothesis

3. Statistic



نمودار ٤ـ۲ مفهوم آزمون فرضیه

می‌دانیم چون $\hat{\beta}$ از یک نمونه تصادفی به دست آمده، پس متغیری تصادفی با توزیع احتمال نرمال است که میانگین آن دقیقاً همان میانگین جامعه ($\beta = ۲/۲۵$) خواهد بود. این توزیع، در نمودار ٤ـ۲ رسم شده است. با توجه به اینکه این آزمون دو طرفه است، احتمال ۹۵ درصد به این معنی است که باید ۹۵ درصد سطح زیر منحنی توزیع احتمال را در حول محور مرکزی ملاحظه کرد. سطح هاشور خورده که منعکس‌کنندهٔ این ۹۵ درصد است، نقاط a و b را مشخص می‌کند. اگر $\hat{\beta}$ به‌دست آمده از نمونه ($\hat{\beta} = ۲/۵$) در فاصلهٔ b و a قرار گیرد، نتیجه می‌گیریم که این نمونه درواقع متعلق به جامعه‌ای است که میانگین آن ۲/۲۵ است؛ بنابراین، $H_0 : \beta = ۲/۲۵$ نمی‌تواند رد شود. در غیر این صورت اگر $\hat{\beta}$ در خارج از این فاصله افتاد، H_0 رد خواهد شد. به همین دلیل است که فاصلهٔ b و a را فاصلهٔ اعتماد یا فاصلهٔ اطمینان[1] می‌گویند. نواحی خارج از این فاصله که با هاشور به صورت ضربدر مشخص شده است ناحیه بحرانی[2] نامیده می‌شود؛ زیرا اگر $\hat{\beta}$ در آنجا واقع شود، H_0 قبول نخواهد شد.

بنابراین، مهمترین مسأله در آزمون فرضیه، پیدا کردن نقاط a و b است. برای یافتن این نقاط، کافی است از تابع توزیع احتمال $\hat{\beta}$ انتگرال گرفته و دو نقطه در روی محور $\hat{\beta}$ چنان تعیین کنیم که سطح زیر منحنی در حول محور مرکزی برابر ۰/۹۵ شود. با تعیین این نقاط، نواحی بحرانی نیز مشخص می‌شود. یک روش ساده آزمون فرضیه، این است که اگر $\hat{\beta}$ در ناحیهٔ بحرانی قرار گرفت، نتیجه می‌گیریم که نمی‌توانیم فرضیه H_0 را قبول کنیم.

بدیهی است این روش، مستلزم انتگرال‌گیری از تابع توزیع احتمال نرمال در هر

1. Confidence Interval 2. Critical Region

مسأله‌ای است که می‌خواهیم آن را حل کنیم. برای رفع این مشکل، کافی است تابع توزیع احتمال نرمال را استاندارد کنیم و برای یک بار و همیشه سطوح زیر منحنی آن را به ازای مقادیر مختلف، با روش انتگرال‌گیری محاسبه کنیم. این محاسبات که در حال حاضر با دقت بسیار انجام شده است در جدول Z و به صورت طبقه‌بندی شده، آماده است.[1] در ذیل این نکته را بیشتر توضیح می‌دهیم.

می‌دانیم اگر یک متغیر تصادفی دارای توزیع نرمال باشد، می‌توان آن را به راحتی استاندارد کرد. برای مثال، در حالت کلی، اگر متغیر تصادفی X_t دارای توزیع احتمال نرمال و میانگین آن $E(X_t)$ و واریانس آن σ_x^2 باشد؛

$$X_t \sim N\left[E(X_t), \sigma_x^2\right] ,$$

در آن صورت می‌توان X_t را به متغیر استاندارد Z_t تبدیل کرد، به گونه‌ای که Z_t دارای توزیع احتمال نرمال با میانگین صفر و واریانس یک شود:

$$Z_t = \sim N(0, 1) .$$

برای این منظور کافی است متغیر تصادفی را از میانگین آن کم کرده، بر جذر واریانس (خطای معیار)[2] تقسیم کنیم،

$$Z_t = \frac{X_t - E(X_t)}{\sigma_x} \sim N(0, 1) .$$

یادآوری می‌شود که X_t در این مثال تنها برای تبیین مفهوم استاندارد کردن یک متغیر تصادفی در حالت کلی به کار رفته است و هیچ رابطه‌ای با X_t به عنوان متغیر برون‌زا یا توضیحی مدل رگرسیون ندارد.

می‌دانیم $\hat{\beta}_{OLS}$ نیز یک متغیر تصادفی است که تابع توزیع احتمال دارد؛ بنابراین، اگر $\hat{\beta}$ را از میانگین آن کم و بر جذر واریانس (خطای معیار) تقسیم کنیم، کسر به دست

---

۱. به پیوست «ب» مراجعه شود.

2. Standard Error

آمده، دارای توزیع Z خواهد بود،

$$Z = \frac{\hat{\beta} - E(\hat{\beta})}{\sqrt{Var(\hat{\beta})}} = \frac{\hat{\beta} - E(\hat{\beta})}{SE(\hat{\beta})} \quad , \qquad (۲\_۲۹)$$

که در آن $SE(\hat{\beta})$ همان خطای خعیار $\hat{\beta}$ است. از معادله‌های ۵ـ۲ و ۱٦ـ۲ می‌دانیم

$$E(\hat{\beta}) = \beta \quad ، \quad Var(\hat{\beta}) = \frac{\sigma^2}{\Sigma x_i^2} \quad .$$

با جایگزینی این مقادیر در معادلهٔ ۲۹ـ۲ خواهیم داشت

$$Z = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} = \frac{\hat{\beta} - \beta}{\sqrt{\sigma^2 / \Sigma x_i^2}} \quad . \qquad (۲\_۳۰)$$

ملاحظه می‌شود Z در معادلهٔ ۳۰ـ۲ دارای توزیعی نرمال با میانگین صفر و واریانس یک است. برای نشان دادن این خصوصیت، باید E (Z) و Var (Z) را حساب کنیم،

$$E(Z) = \frac{1}{\sqrt{\sigma^2 / \Sigma x_i^2}} \; E(\hat{\beta} - \beta) \quad ,$$

و چون

$$E(\hat{\beta} - \beta) = E(\hat{\beta}) - E(\beta) = \beta - \beta = 0 \quad ,$$

بنابراین $E(Z) = 0$ است.

با گرفتن واریانس و توجه به این نکته که Σx_i^2 و σ^2 هر یک کمیتهای ثابتی هستند خواهیم داشت

$$Var(Z) = \frac{1}{\sigma^2 / \Sigma x_i^2} \; Var(\hat{\beta} - \beta) \quad ,$$

و چون

$$Var(\hat{\beta} - \beta) = Var(\hat{\beta}) + Var(\beta) \quad ,$$

و با توجه به ثابت بودن مقدار β، خواهیم داشت؛ $Var(\beta) = 0$، در نتیجه

$$\text{Var}(\hat{\beta} - \beta) = \text{Var}(\hat{\beta}) = \frac{\sigma^2}{\Sigma x_t^2} .$$

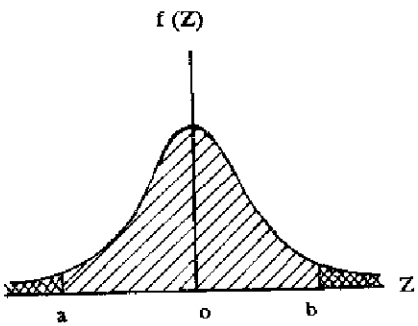
واریانس Z به صورت زیر خواهد بود

$$\text{Var}(Z) = \frac{1}{\sigma^2 / \Sigma x_t^2} \sigma^2 / \Sigma x_t^2 = 1$$

نتیجهٔ کلی این است که $\hat{\beta}$ استاندارد دارای توزیع Z با میانگین صفر و واریانس یک است،

$$\frac{\hat{\beta} - \beta}{\text{SE}(\hat{\beta})} = Z \sim N(0, 1) . \qquad (۲\text{-}۳۱)$$

نمودار ۲-۵ توزیع احتمال Z را نشان می‌دهد. با توجه به اینکه این آزمون دو طرفه است، احتمال ۹۵ درصد بدین معنی خواهد بود که باید ۹۵ درصد سطح زیر منحنی توزیع احتمال را در حول محور مرکزی ملاحظه کرد. سطح هاشور خورده که منعکس‌کنندهٔ این ۹۵ درصد است، نقاط a و b را مشخص می‌کند. در اینجا باید دید، آیا $\hat{\beta}$ به دست آمده از نمونه در فاصلهٔ a و b قرار می‌گیرد یا خیر؟ اما نکتهٔ بسیار مهم این است که ابتدا باید $\hat{\beta}$ را استاندارد کرد تا بتواند با نقاط موجود دیگر روی محور Z قابل مقایسه شود. اگر $\hat{\beta}$ استاندارد شده درفاصلهٔ b و a قرار گیرد، نتیجه می‌گیریم که این نمونه در واقع متعلق به جامعه‌ای است که میانگین آن صفر است؛ بنابراین فرضیهٔ $H_0 : \beta = ۲/۲۵$ نمی‌تواند رد شود. البته صفر بودن میانگین جامعه در یک مقیاس استاندارد به معنی ۲/۲۵ بودن این میانگین در مقیاس عادی است. اگر $\hat{\beta}$ استاندارد شده در خارج از فاصلهٔ b و a قرار بگیرد، فرضیهٔ H_0 قبول نخواهد شد. به همین دلیل است که فاصلهٔ b و a را فاصلهٔ اطمینان می‌گویند؛ مانند بحث قبل، نواحی خارج از این فاصله ـ که



نمودار ۲-۵ توزیع متغیر استاندارد Z

با هاشور ضربدری مشخص شده است ـ نواحی بحرانی نامیده می‌شود؛ زیرا اگر $\hat{\beta}$ استاندارد شده در آنجا واقع شود، H_0 قبول نخواهد شد.

به این ترتیب مهمترین مسأله در آزمون فرضیه، یافتن نقاط b و a است و این امر با انتگرال‌گیری از تابع توزیع نرمال استاندارد، به سهولت به دست می‌آید. برخلاف گذشته لازم نیست برای هر مسأله، جداگانه انتگرال‌گیری شود. چون فقط و فقط یک تابع برای منحنی نرمال استاندارد وجود دارد و با توجه به اینکه میانگین و واریانس این تابع همواره صفر و یک است بنابراین کافی است به ازای مقادیر مختلف موجود در محور Z ، انتگرال‌گیری شده و سطوح مختلف زیر منحنی نرمال استاندارد محاسبه شود. این محاسبات در جدول Z طبقه‌بندی شده است.

بدین ترتیب در سطح احتمال ۹۵ درصد، نقاط a و b که قرینه نیز هستند از جدول Z به دست می‌آید. این مقادیر را «Z جدول» می‌نامیم. اکنون باید $\hat{\beta}$ را استاندارد کرد؛ یعنی آن را به متغیر Z تبدیل کرد. کافی است $\hat{\beta}$ را از میانگین آن کم کرده و بر خطای معیارش تقسیم کنیم؛ این مقدار را «Z محاسباتی» یا «آمارهٔ آزمون»[1] می‌گوییم،

$$\text{آمارهٔ آزمون } Z = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} \quad (۲\text{ـ}۳۲)$$

اگر آمارهٔ آزمون در فاصلهٔ مقادیر Z جدول قرار گرفت، نتیجه می‌گیریم که $\hat{\beta}$ به دست آمده از نمونه در واقع متعلق به جامعه‌ای است که میانگین آن در H_0 تحت آزمون قرار گرفته است. در چنین حالتی، فرضیهٔ H_0 رد نمی‌شود. اگر آمارهٔ آزمون در ناحیهٔ بحرانی قرار گرفت، H_0 رد خواهد شد، به عبارت دقیقتر قبول نمی‌شود. قبل از پایان مباحث این قسمت، ضروری است که به یک اصطلاح و یک نکته اشاره کنیم.

معمولاً برای هر آزمون ابتدا یک «اندازه یا سطح آزمون»[2] تعیین می‌شود که آن را با α نشان می‌دهند. اندازه یا سطح آزمون ۱ درصد یا ۵ درصد بسیار معمول است.[3] البته

---

1. Test Statistic
2. Size or Level of Test

۳. آزمون ۱ درصد و ۵ درصد بیشتر توسط فیشر (۱۹۶۲ ـ ۱۸۹۰ R. A. Fisher) از بنیانگذاران آمار جدید رایج شده است.

به جای لفظ اندازه یا سطح آزمون به طور خلاصه لفظ آزمون یا سطح نیز به کار می‌رود؛ یعنی گفته می‌شود آزمون ۱ درصد و آزمون ۵ درصد یا سطح ۱ درصد و سطح ۵ درصد. α در حقیقت سطح معنی‌داری آماری[1] آزمون است و نشان می‌دهد چقدر احتمال دارد فرضیهٔ H_0 را که در واقع صحیح است، رد کنیم؛ یعنی اشتباه نوع اول را مرتکب شویم. اصطلاح سطح معنی‌داری آماری را به طور خلاصه «سطح معنی‌دار» می‌گوییم. بنابراین اصطلاحات سطح معنی‌داری ۵ درصد، سطح ۵ درصد و آزمون ۵ درصد نه تنها همگی یک حقیقت را بیان می‌کنند بلکه با اصطلاحات سطح اطمینان ۹۵ درصد، فاصلهٔ اطمینان ۹۵ درصد و احتمال ۹۵ درصد ـ در واقع همان $(\alpha - 1)$ ـ نیز معادل است. ناگفته نماند که بعضی ازمتخصصان اقتصادسنجی به α، اندازه یا سطح آزمون، ولی به $(\alpha - 1)$ سطح معنی‌داری آزمون می‌گویند؛ بنابراین در متون مختلف باید به تفاوت این اصطلاحات توجه کرد. البته همهٔ متون، اصطلاح «فاصلهٔ اطمینان» را به کار می‌برند؛ مثلاً فاصلهٔ اطمینان ۹۵ درصد منعکس‌کنندهٔ این واقعیت است که مجموع سطوح نواحی بحرانی برابر ۵ درصد است.

با توجه به قرینگی منحنی تغییرات z و t، محاسبات انتگرال، فقط برای نیمی از منحنی تغییرات انجام شده است؛ بنابراین در آزمونهای دو طرفه باید به مقدار $\frac{\alpha}{2}$ توجه کرد؛ مثلاً در مثال فوق و با توجه به نمودار ۵ـ۲ می‌توان گفت که چون مساحت سطح زیر منحنی برابر یک است؛ بنابراین a نقطه‌ای است که سطح زیرمنحنی از $\alpha -$ تا آن نقطه برابر ۰/۰۲۵ است. به همین ترتیب b نقطه‌ای است که سطح زیرمنحنی از آن نقطه تا $\alpha +$، برابر ۰/۰۲۵ خواهد بود. مجموع این دو سطح، برابر $\alpha =$ ۰/۰۵ می‌شود؛ به عبارت دیگر نقاط a و b متناظر با کمیت $Z_{\frac{\alpha}{2}}$ در جدول Z است؛ یعنی نقاطی است که سطح زیر منحنی از آن به بعد (یا قبل از آن) برابر $\frac{۰/۰۵}{۲}$ است.

مثال ۲ـ۲ فرض کنید که در تخمین یک مدل رگرسیون خطی، $\hat{\beta}$ برابر ۲۹/۴۸ شده است. همچنین فرض می‌کنیم که واریانس $\hat{\beta}$ را می‌دانیم، به گونه‌ای که جذر آن ـ که در واقع

1. Statistical Level of Significance

همان خطای معیار است ـ برابر ۳۶ شده است. می‌خواهیم در سطح معنی‌دار ۵ درصد این فرضیه را آزمون کنیم که β واقعی برابر با ۲۵ است.

حلّ داریم:

$$\hat{\beta} = ۲۹/۴۸ \qquad H_۰ : \beta = ۲۵$$

$$SE(\hat{\beta}) = ۳۶ \qquad H_۱ : \beta \neq ۲۵$$

ملاحظه می‌شود که آزمون ما دو طرفه است. ابتدا آمارهٔ Z را می‌سازیم؛ به عبارت دیگر $\hat{\beta}$ را استاندارد می‌کنیم که بتوان آن را روی محور Z مشخص کرد. کافی است $\hat{\beta}$ را از میانگین آن کم کرده و بر خطای معیارش تقسیم کنیم،

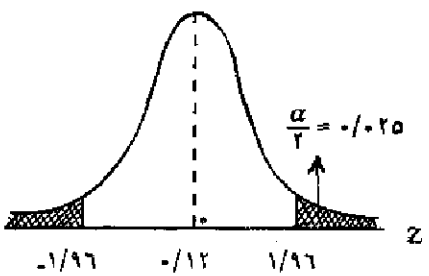
$$Z = \frac{\hat{\beta} - E(\hat{\beta})}{SE(\hat{\beta})}$$

می‌دانیم $E(\hat{\beta}) = \beta$ است و چون بنا بر فرضیهٔ $H_۰$ داریم $\beta = ۲۵$، بنابراین

$$Z = \frac{۲۹/۴۸ - ۲۵}{۳۶} = ۰/۱۲$$

در اینجا باید این آمارهٔ آزمون را با مقدار به دست آمده از جدول Z مقایسه کنیم اگر در ناحیهٔ بحرانی قرار گرفت، فرضیهٔ $H_۰$ قبول می‌شود. چون می‌خواهیم در سطح معنی‌دار ۵ درصد، یعنی به ازای $\alpha = ٪۵$ این فرضیه را آزمون کنیم و با توجه به اینکه آزمون ما دو طرفه است، باید در جدول Z، مقدار Z متناظر با $\frac{\alpha}{۲} = \frac{۰/۰۵}{۲} = ۰/۰۲۵$ را پیدا کنیم. ملاحظه می‌شود که

$$Z_{۰/۰۲۵} = \pm ۱/۹۶$$



آزمون Z: مقایسهٔ آمارهٔ آزمون و Z جدول

مفید است مقدار آمارهٔ آزمون را در نمودار مشخص کنیم. آمارهٔ آزمون برابر

۰/۱۲ است؛ بنابراین در ناحیهٔ بحرانی قرار نمی‌گیرد. نتیجه می‌گیریم که آزمون H_0 یعنی $\beta = ۲۵$ نمی‌تواند رد شود.

فرضیهٔ $\beta = ۰$ در آزمون Z

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_i = \alpha + \beta X_i + U_i$$

اگر بخواهیم به این سؤال پاسخ دهیم که آیا X_i تأثیری بر Y_i دارد یا خیر، باید این فرضیه را آزمون کنیم که آیا $\beta = ۰$ قابل قبول است؟ بنابراین فرضیهٔ H_0 و فرضیه H_1 در این مورد چنین خواهد بود،

$$H_0 : \beta = ۰ \qquad \text{و} \qquad H_1 : \beta \neq ۰$$

در این حالت، مقدار $\hat{\beta}$ استاندارد شده یعنی آمارهٔ آزمون عبارت خواهد بود از

$$Z = \frac{\hat{\beta} - E(\hat{\beta})}{SE(\hat{\beta})} = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})}$$

و چون بنا بر فرضیهٔ H_0 داریم، $\beta = ۰$، آنگاه

$$Z = \frac{\hat{\beta}}{SE(\hat{\beta})} \qquad (۲\text{-}۳۳)$$

نتیجه می‌گیریم که نحوهٔ عمل در آزمون فرضیهٔ $\beta = ۰$ بسیار ساده است. باید مقدار $\hat{\beta}$ را بر خطای معیار آن تقسیم کنیم تا آمارهٔ آزمون به دست آید. اگر آمارهٔ آزمون در ناحیهٔ بحرانی قرار گرفت، معنی‌دار است و فرضیهٔ H_0 رد می‌شود، در غیر این صورت نمی‌توان آن را رد کرد.

معمولاً، فرضیهٔ $\beta = ۰$ را می‌توان به روش ساده‌تر ـ که البته تقریبی است ـ نیز آزمود. می‌دانیم در سطح احتمال ۹۵ درصد، مقدار Z به دست آمده از جدول Z برابر با ۱/۹۶ ± است. اگر این مقدار را برابر ۲ ± بگیریم (که قاعدتاً متناظر با احتمال بیشتر از

۹۵ درصد نیز هست)، آنگاه هنگامی H_0 رد می‌شودکه قدر مطلق آمارهٔ آزمون بیشتر از $|\pm ۲|$ بشود،

$$\left| \frac{\hat{\beta}}{SE(\hat{\beta})} \right| > |\pm ۲|$$

به عبارت دیگر، بدون توجه به علامت، مقدار $\hat{\beta}$ باید از دو برابر خطای معیار خود بزرگتر باشد،

$$|\hat{\beta}| > ۲\, SE(\hat{\beta}) \qquad (۲\text{ـ}۳۴)$$

در صورت درست بودن نامساوی (۲ـ۳۴)، فرضیهٔ H_0 رد می‌شود، یعنی نتیجه می‌گیریم که متغیر توضیحی X_t در واقع تأثیر قابل ملاحظه‌ای بر تغییرات Y_t دارد.

مثال ۳ـ۲ فرض کنید با استفاده از یک نمونهٔ ۵۰۰ تایی مشاهدات درآمد و مصرف، تابع مصرف زیر را تخمین زده‌ایم،

$$\hat{C}_t = ۲۵۰/۷ + ۰/۸۴\, Y_t$$

می‌دانیم انحراف معیار میل نهایی به مصرف برابر ۰/۱۴ است $(SE(\hat{\beta}) = ۰/۱۴)$ که β میل نهایی به مصرف فرض شده است. می‌خواهیم در سطح معنی‌دار ۵ درصد این فرضیه را آزمون کنیم که آیا درآمد، بر تغییرات مصرف تأثیری دارد یا خیر؟

به دو دلیل می‌توان از آزمون Z استفاده کرد: اولاً حجم مشاهدات بسیار است؛ ثانیاً، فرض شده است که مقدار واقعی واریانس تخمین پارامتر را می‌دانیم. البته وجود یکی از این موارد، کافی است که به ما اجازه دهد از آزمون Z استفاده کنیم. فرضیه‌ای که می‌خواهیم آزمون کنیم چنین است.

$$H_0 : \beta = ۰$$

$$H_1 : \beta \neq ۰$$

ابتدا باید آمارهٔ آزمون را تشکیل دهیم. با استفاده از معادلهٔ ۲ـ۳۳ داریم

$$Z = \frac{\hat{\beta}}{SE(\hat{\beta})} = \frac{۰/۸٤}{۰/۱٤} = ٦$$

چون آزمون دو طرفه است، در سطح معنی‌دار ۵ درصد، یعنی به ازای $\frac{\alpha}{۲} = \frac{۰/۰۵}{۲} = ۰/۰۲۵$ داریم

$$Z_{۰/۰۲۵} = \pm ۱/۹٦$$

بدیهی است که آمارهٔ آزمون به مراتب از این مقدار بیشتر می‌شود و معنی‌دار است. در نتیجه فرضیهٔ $H_۰$ رد می‌شود؛ یعنی فرض تأثیر درآمد بر مصرف را نمی‌توان رد کرد.

راه حل بسیار ساده‌تر این است که چون مقدار تخمین $\hat{\beta}$ یعنی ۰/۸٤ از دو برابر انحراف معیار آن بزرگتر است،

$$۰/۸٤ > ۲\,(۰/۱٤)$$

در نتیجه بنا بر رابطهٔ ۲.۳٤ فرضیهٔ $H_۰$ رد خواهد شد.

فاصلهٔ اطمینان برای پارامترها در آزمون Z

مدل رگرسیون زیر را در نظر می‌گیریم،

$$Y_t = \alpha + \beta X_t + U_t \quad .$$

می‌دانیم $\hat{\beta}_{OLS}$ دقیقاً برابر مقدار واقعی β نخواهد بود. سؤال این است که آیا می‌توان با دانستن $\hat{\beta}$، دربارهٔ مقدار واقعی β اظهارنظر کرد؟ بدیهی است چون مقدار واقعی β را نمی‌دانیم، با آگاهی از $\hat{\beta}$، فقط می‌توان فاصله‌ای را تعیین کرد که با احتمال معینی انتظار می‌رود β واقعی در آن محدوده قرار گیرد. این محدوده را فاصلهٔ اعتماد یا فاصلهٔ اطمینان می‌گویند.

برای به دست آوردن این فاصله، کافی است $\hat{\beta}$ را به روش سابق استاندارد کنیم. $\hat{\beta}$ را از میانگین آن کم کرده و حاصل را بر خطای معیار $\hat{\beta}$ تقسیم کنیم. مقدار $\hat{\beta}$ استاندارد شده عبارت خواهد بود از

$$Z = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} .$$

فاصلهٔ اطمینان را باید با یک احتمال معینی تخمین زد که به آن سطح معنی‌داری می‌گویند. فرض کنیم آزمون ما مثلاً α درصد است؛ پس سطح اطمینان آزمون $(1 - \alpha)$ خواهد بود. مطابق نمودار ۲ـ۵ ، هدف ما در واقع این است که حدودی مانند a و b را تعیین کنیم که آمارهٔ آزمون در آن فاصله قرار گیرد. نقطهٔ a متناظر با $Z_{\frac{\alpha}{2}}$ و نقطهٔ b متناظر با $-Z_{\frac{\alpha}{2}}$ است که ازجدول Z به دست می‌آید؛ بنابراین آمارهٔ آزمون باید درفاصلهٔ بین این دو قرار گیرد، یعنی

$$-Z_{\frac{\alpha}{2}} \text{ جدول} < \text{آمارهٔ آزمون} < Z_{\frac{\alpha}{2}} \text{ جدول} ,$$

یا

$$-Z_{\frac{\alpha}{2}} < \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} < Z_{\frac{\alpha}{2}} .$$

دو طرف نامساوی را در $SE(\hat{\beta})$ ضرب می‌کنیم،

$$-Z_{\frac{\alpha}{2}} SE(\hat{\beta}) < \hat{\beta} - \beta < Z_{\frac{\alpha}{2}} SE(\hat{\beta}) .$$

از دو طرف $\hat{\beta}$ را کم کرده، بعد از ضرب در (۱-) ، خواهیم داشت

$$\hat{\beta} - Z_{\frac{\alpha}{2}} SE(\hat{\beta}) < \beta < \hat{\beta} + Z_{\frac{\alpha}{2}} SE(\hat{\beta}) . \qquad (۲\text{ـ}۳۵)$$

نامساوی ۲ـ۳۵ فاصلهٔ اطمینان β واقعی را تعیین می‌کند. با داشتن فاصلهٔ اطمینان، می‌توان روش دیگری نیز برای آزمون فرضیه‌ها پیشنهاد کرد. کافی است بعد از ساختن فاصلهٔ اطمینان، ببینیم آیا فرضیهٔ $H_0 : \beta = k$ در این فاصله قرار می‌گیرد یا خیر. اگر $\beta = k$ در فاصلهٔ اطمینان واقع شد، در آن صورت می‌گوییم فرضیهٔ H_0 رد نمی‌شود، در غیر این صورت نمی‌توان H_0 را قبول کرد.

مثال ٤ـ۲ در مثال ۲ـ۲، اولاً، برای β یک فاصلهٔ اطمینان ۹۵ درصد بسازید. ثانیاً، این

فرضیه را آزمون کنید که آیا β برابر ۲۵ است؟

اولاً، با استفاده از معادلهٔ ۳۵ـ۲ داریم:

$$۲۹/٤۸ - ۱/۹٦(۳٦) < \beta < ۲۹/٤۸ + ۱/۹٦(۳٦) ,$$

$$-٤۱/۰۸ < \beta < ۱۰۰/۰٤ .$$

ثانیاً، چون مقدار $\beta = ۲۵$ در این فاصلهٔ اطمینان قرار می‌گیرد، بنابراین نمی‌تواند رد شود.

مثال ۵ـ۲ در مورد مثال ۳ـ۲، اولاً، برای میل نهایی به مصرف، یک فاصلهٔ اطمینان ۹۵ درصد بسازید، ثانیاً، فرضیه $H_0 : \beta = ۰$ را آزمون کنید.

اولاً، با استفاده از رابطهٔ ۳۵ـ۲ داریم

$$۰/۸٤ - ۱/۹٦(۰/۱٤) < \beta < ۰/۸٤ + ۱/۹٦(۰/۱٤) ,$$

$$۰/٥٦٥٦ < \beta < ۱/۱۱٤٤ .$$

ثانیاً، چون مقدار $\beta = ۰$ در این فاصلهٔ اطمینان قرار نمی‌گیرد، بنابراین فرضیهٔ H_0 نمی‌تواند قبول شود.

مثال ٦ـ۲ در تخمین یک مدل رگرسیون، $\hat{\beta} = ۸/٤$ و $SE(\hat{\beta}) = ۲/۲$ به دست آمده است. برای مقدار واقعی β یک فاصله اطمینان ۹۵ درصد بسازید.

می‌دانیم

$$\hat{\beta} - Z_{\frac{\alpha}{۲}} SE(\hat{\beta}) < \beta < \hat{\beta} + Z_{\frac{\alpha}{۲}} SE(\hat{\beta}) ,$$

با توجه به $Z_{\frac{\alpha}{۲}} = \pm ۱/۹٦$، خواهیم داشت

$$-۸/٤ + ۱/۹٦(۲/۲) < \beta < ۸/٤ + ۱/۹٦(۲/۲) ,$$

یا

$$٤/۱ < \beta < ۱۲/۷ .$$

یعنی $\hat{\beta} = ۸/٤$ که حاصل یک نمونهٔ آماری است از جامعه‌ای گرفته شده است که مقدار واقعی β در آن با ۹۵ درصد احتمال بین ٤/۱ و ۱۲/۷ قرار دارد.

۲. آزمون t برای پارامترها

می‌دانیم آزمون هر فرضیه‌ای در مورد $\hat{\beta}$ مستلزم استاندارد کردن $\hat{\beta}$ است. $\hat{\beta}$ استاندارد شده (Z) را که از فرمول ۳۰ـ۲ به دست می‌آید، یک بار دیگر می‌نویسیم،

$$Z = \frac{\hat{\beta} - \beta}{\sqrt{\sigma^2 / \Sigma x_i^2}} .$$

محاسبهٔ Z مستلزم داشتن σ^2 است و معمولاً σ^2 را نمی‌دانیم؛ زیرا مقدار واقعی واریانس U_i یا واریانس Y_i مجهول است؛ بنابراین باید σ^2 را تخمین زد. در معادلهٔ ۲۸ـ۲ و در بحث مربوط به U_i و e_i، تخمین واریانس U_i به صورت زیر به دست آمد،

$$\hat{\sigma}_U^2 = \frac{\Sigma e_i^2}{n - ۲} .$$

اگر در فرمول ۳۰ـ۲ به جای σ^2، مقدار تخمین آن ($\hat{\sigma}^2$) را قرار دهیم، در آن صورت آمارهٔ به دست آمده توزیع t با (n − ۲) درجهٔ آزادی دارد،

$$\frac{\hat{\beta} - \beta}{\sqrt{\hat{\sigma}_U^2 / \Sigma x_i^2}} \sim t\,(n - ۲) ,$$

با فرض $s^2 = \hat{\sigma}_U^2$، خواهیم داشت

$$\frac{\hat{\beta} - \beta}{\sqrt{s^2 / \Sigma x_i^2}} \sim t\,(n - ۲) . \qquad (۲\text{ـ}۳٦)$$

برای تبیین فرمول ۳٦ـ۲ به طور خلاصه به بعضی مفاهیم آماری اشاره می‌کنیم. برای مباحث مقدماتی اقتصادسنجی همین تعاریف کافی است، ولی برای تحلیل دقیق قاعدتاً باید به کتابهای آمار مراجعه کرد. می‌دانیم اگر n متغیر تصادفی داشته باشیم که هر یک دارای توزیع نرمال مستقل با میانگین صفر و واریانس یک باشد آنگاه مجموع

مربعات آنها، توزیع «مربع کای»[1] با n درجه آزادی خواهد داشت. معمولاً توزیع مربع کای با با علامت χ^2 نشان می‌دهند؛[2] بنابراین اگر $Z \sim N(0, 1)$ یک متغیر نرمال استاندارد باشد و n مقدار تصادفی $z_1, z_2, \ldots, z_n$ را از این توزیع به دست آوریم و هر یک را مجذور و نتایج را با هم جمع کنیم، آنگاه آمارهٔ به دست آمده توزیع χ^2 با n درجهٔ آزادی خواهد داشت،[3]

$$(z_1^2 + z_2^2 + \ldots + z_n^2) \sim \chi^2(n) \quad . \qquad (۲\text{-}۳۷)$$

وقتی درجات آزادی به سمت بی‌نهایت میل کند، توزیع χ^2 به سمت توزیع نرمال میل می‌کند. می‌توان نشان داد که $\frac{\sum_{i=1}^{n} e_i^2}{\sigma_U^2}$، دارای توزیع χ^2 با $(n-2)$ درجه آزادی است،

$$\frac{\sum_{i=1}^{n} e_i^2}{\sigma_U^2} \sim \chi^2(n-2) \quad . \qquad (۲\text{-}۳۸)$$

در اینجا فرض می‌کنیم که دو متغیر تصادفی با توزیعهای نرمال استاندارد و χ^2 به شرح زیر داریم

$$Z \sim N(0, 1) \quad , \quad V \sim \chi^2(\nu) \quad ,$$

که Z و V از یکدیگر مستقل و ν درجات آزادی توزیع V است. اگر متغیر نرمال استاندارد را بر جذر متغیری که توزیع χ^2 داشته باشد و بر درجات آزادیش نیز تقسیم شده باشد، تقسیم کنیم آمارهٔ به دست آمده، توزیع t با ν درجهٔ آزادی خواهد داشت؛ بنابراین

$$\frac{Z}{\sqrt{V/\nu}} \sim t(\nu) \quad .$$

با توجه به معادلهٔ ۲ـ۳۰ می‌دانیم که مقدار $\hat{\beta}$ استاندارد، توزیع نرمال استاندارد دارد،

---

1. Chi - square Distribution

۲. χ بیست و دومین حرف از حروف الفبای یونانی است که صورت بزرگ آن X است. به پیوست «ج» مراجعه شود.

۳. به پیوستهای «۲ـ ب» و «۳ـ ب» مراجعه شود.

$$Z = \frac{\hat{\beta} - \beta}{\sqrt{\sigma_U^2 / \Sigma x_t^2}} \sim N(0, 1) \ .$$

از طرف دیگر با توجه به معادلهٔ ۲ـ۳۸ نیز می‌دانیم که $\frac{\Sigma e_t^2}{\sigma_U^2}$ دارای توزیع χ^2 با $(n - 2)$ درجهٔ آزادی است؛ اگر $\frac{\Sigma e_t^2}{\sigma_U^2}$ را بر درجات آزادی آن تقسیم کنیم، خواهیم داشت

$$\frac{\Sigma e_t^2}{(n - 2)\sigma_U^2} \ ,$$

و اگر Z به دست آمده از معادلهٔ ۲ـ۳۰ را بر آن تقسیم کنیم، آمارهٔ به دست آمده توزیع t به $(n - 2)$ درجه آزادی خواهد داشت،

$$\left[\frac{(\hat{\beta} - \beta)\sqrt{\Sigma x_t^2}}{\sqrt{\sigma_U^2}} \div \frac{\sqrt{\Sigma e_t^2}}{\sqrt{n - 2}\sqrt{\sigma_U^2}}\right] \sim t(n - 2) \ ,$$

یا

$$\frac{(\hat{\beta} - \beta)}{\sqrt{\frac{\Sigma e_t^2}{n - 2}} / \sqrt{\Sigma x_t^2}} \sim t(n - 2) \ .$$

با جایگزینی معادلهٔ ۲ـ۲۸ خواهیم داشت

$$\frac{(\hat{\beta} - \beta)}{\sqrt{\hat{\sigma}_U^2 / \Sigma x_t^2}} \sim t(n - 2) \ ,$$

و با فرض $s^2 = \hat{\sigma}_U^2$، داریم

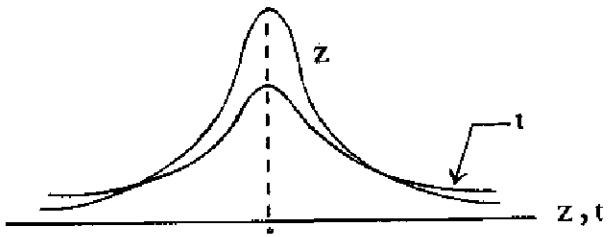
$$\frac{(\hat{\beta} - \beta)}{\sqrt{s^2 / \Sigma x_t^2}} \sim t(n - 2) \ ,$$

که دقیقاً همان معادلهٔ ۲ـ۳۶ است. بنابراین نتیجه می‌گیریم که اگر در فرمول Z به جای σ^2، مقدار تخمین آن یعنی $\hat{\sigma}^2$ را قرار دهیم، توزیع t به دست می‌آید.[۱]

---

۱. توزیع t را ویلیام گاست در سال ۱۹۰۸ در مجله بیومتریکا *(Biometrika)* منتشر کرد. گاست مقالات خود را با نام مستعار استیودنت (Student) به چاپ می‌رساند. به پیوست «د» مراجعه شود.

مانند توزیع نرمال، توزیع t نیز در حول میانگین صفر، قرینگی دارد. در مقایسه منحنی تغییرات t با Z می‌توان گفت که دنباله‌های منحنی t پهن‌تر است و این امر بویژه در حجم کم مشاهدات بسیار قابل ملاحظه است. به موازات افزایش حجم نمونه، منحنی تغییرات t به سمت Z میل می‌کند. معمولاً به ازای $n > ۳۰$ تفاوت مقادیر t و Z بسیار کم و قابل اغماض است. به همین دلیل در مواردی که حجم مشاهدات بیشتر از ۳۰ است، علی‌رغم نداشتن مقدار واقعی واریانس U_1، می‌توان از جدول Z استفاده کرد. منحنی Z و منحنی t در نمودار ۶ـ۲ نشان داده شده است.



نمودار ۶ـ۲ نمایش منحنیهای t , Z

آزمون فرضیه‌های مختلف دربارهٔ $\hat{\beta}$ در تابع توزیع احتمال t دقیقاً مشابه مباحثی است که برای Z مطرح کردیم. ابتدا باید فرضیه‌های H_0 و H_1 را تشکیل دهیم، آنگاه سطح معنی‌دار بودن آزمون را تعیین کنیم. $\hat{\beta}$ را استاندارد کرده، آمارهٔ آزمون را به دست می‌آوریم. برای به دست آوردن t از جدول، علاوه بر سطح معنی‌دار بودن آزمون، باید درجات آزادی را نیز در نظر بگیریم. در واقع عناصر موجود در جدول t برحسب سطوح معنی‌داری و درجات آزادی، مدون شده است.[۱] همچنین با توجه به قرینگی منحنی تغییرات t، فقط نیمی از محاسبات انتگرال انجام شده است. اگر قدر مطلق آماره آزمون از $|t|$ در جدول بزرگتر باشد در آن صورت آماره آزمون معنی‌دار است و فرضیهٔ H_0 رد می‌شود؛ به عبارت دیگر هرگاه

$$\left| \frac{\hat{\beta} - \beta}{\sqrt{s^2 / \Sigma x_i^2}} \right| > |\pm t| , \qquad (۲\text{ـ}۳۹)$$

۱. به پیوست «۴ ـ ب» مراجعه شود.

باشد آنگاه H_0 رد خواهد شد، در غیر این صورت نمی‌توان H_0 را رد کرد.

مثال ۷ـ۲ مدل رابطهٔ تولید با ساعتهای نیروی کار، موضوع مثال ۱ـ۲ را دوباره در نظر می‌گیریم؛

$$Q_i = \alpha + \beta L_i + U_i \quad .$$

اولاً، فرضیهٔ $H_0 : \beta = 1$ را در مقابل فرضیهٔ $H_1 : \beta \neq 1$ در سطح معنی دار ۵ درصد آزمون کنید.

ثانیاً، در همین سطح معنی‌داری، فرضیهٔ $H_0 : \beta = 1$ را در مقابل فرضیهٔ $H_1 : \beta > 1$ آزمون کنید.

از مثال ۱ـ۲ این اطلاعات را داریم

$$\hat{\beta} = ۰/۷۵ \quad , \quad SE(\hat{\beta}) = ۰/۲۵۶ \ ,$$

$$\hat{\sigma}^2 = ۱/۸۳ \quad .$$

با توجه به اینکه به جای مقدار واقعی واریانس U_i، فقط تخمین آن را داریم؛ بنابراین $\hat{\beta}$ استاندارد توزیع t خواهد داشت،

$$t = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} \ ,$$

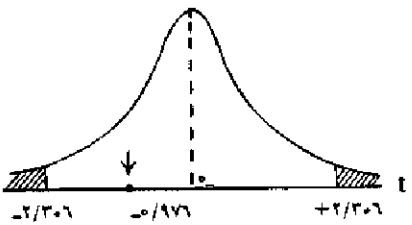
$$= \frac{۰/۷۵ - ۱}{۰/۲۵۶} = -۰/۹۷۶ \quad .$$

اولاً، با حجم مشاهدات $n = ۱۰$ و درجات آزادی $۸ = ۲ - ۱۰ = \nu$، مقدار t را از جدول t و در سطح معنی‌دار ۵ درصد به دست می‌آوریم. چون آزمون دو طرفه است، باید t متناظر با سطح بحرانی $\frac{۰/۰۵}{۲} = ۰/۰۲۵$ را ملاحظه کنیم، داریم

$$t_{۰/۰۲۵} = \pm ۲/۳۰۶ \quad .$$

باید قدر مطلق آمارهٔ آزمون را با قدر مطلق مقدار جدول t مقایسه کنیم. اگر آمارهٔ آزمون از مقدار جدول t بیشتر شد، معنی‌دار است و فرضیه H_0 رد می‌شود. ملاحظه می‌شود که

۲/۳۰۶ > ۰/۹۷۶ – است؛ بنابراین فرضیهٔ H_0 رد نخواهد شد.



نمودار ۲.۷ نواحی بحرانی و آزمون t

با استفاده از نمودار نیز می‌توان آزمون را انجام داد. اگر آمارهٔ آزمون در ناحیهٔ بحرانی قرار بگیرد، فرضیهٔ H_0 رد می‌شود. نواحی بحرانی با استفاده از جدول t به‌دست می‌آید و به ترتیب با سطوح از ∞ – تا ۲/۳۰۶ – و از ۲/۳۰۶ تا ∞ + مشخص می‌شود. چون آمارهٔ آزمون، یعنی ۰/۹۷۶ – در هیچکدام از نواحی بحرانی واقع نشده است؛ پس فرضیهٔ H_0 رد نمی‌شود.

ثانیاً، اگر بخواهیم فرضیهٔ $H_0 : \beta = 1$ را در مقابل فرضیهٔ $H_1 : \beta > 1$ آزمون کنیم، در آن صورت آزمون ما یکطرفه است. دیگر لازم نیست سطح معنی‌داری را نصف کنیم، بلکه با درجات آزادی $\nu = 8$ و $\alpha = 0/05$، مقدار t را از جدول به دست می‌آوریم؛ یعنی نقطه‌ای را در روی محور t مشخص می‌کنیم که سطح زیر منحنی از آن نقطه به بعد برابر ۵ درصد باشد و این دقیقاً همان ناحیهٔ بحرانی است، خواهیم داشت $t = 1/860$. بنابراین چون آمارهٔ آزمون در ناحیهٔ بحرانی قرار نمی‌گیرد، فرضیهٔ H_0 رد نمی‌شود.

فرضیهٔ $\beta = 0$ در آزمون t

در اینجا دقیقاً مانند مباحثی که در آزمون Z مطرح شد، عمل می‌کنیم. فرضیهٔ H_0 و H_1 به صورت زیر خواهد بود:

$$H_0 : \beta = 0 \quad , \quad H_1 : \beta \neq 0 \ .$$

$\hat{\beta}$ استاندارد شده (آمارهٔ آزمون) را به دست می‌آوریم،

$$t = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} \ .$$

چون بنا بر فرضیهٔ H_0 داریم، $\beta = 0$، پس

$$t = \frac{\hat{\beta}}{SE(\hat{\beta})} \quad . \qquad (۲\text{-}۴۰)$$

بنابراین باید $\hat{\beta}$ را بر انحراف معیار آن تقسیم کنیم تا آمارهٔ آزمون به دست آید. سپس مقدار به دست آمده را با جدول t مقایسه می‌کنیم، اگر در ناحیهٔ بحرانی قرار گرفت، معنی‌دار است و فرضیهٔ H_0 رد می‌شود؛ در غیر این صورت نمی‌توان آن را رد کرد.

مثال ۸-۲ در مثال ۷-۲ فرضیهٔ $H_0 : \beta = 0$ را در مقابل فرضیهٔ $H_1 : \beta \neq 0$ در سطح معنی‌دار ۵ درصد آزمون کنید.

حلّ: با اینکه از لحاظ نظریه اقتصادی، این فرضیه مفهومی ندارد، برای بررسی چگونگی استفاده از فرمول ۴۰-۲ آمارهٔ آزمون را محاسبه می‌کنیم،

$$t = \frac{۰/۷۵}{۰/۲۵۶} = ۲/۹۲۹ \quad .$$

چون آزمون دو طرفه است، بنابراین $t = \pm ۲/۳۰۶$. ملاحظه می‌شود که آمارهٔ آزمون در ناحیهٔ بحرانی قرار می‌گیرد و فرضیهٔ H_0 رد می‌شود؛ یعنی متغیر توضیحی از نظر آماری بر متغیر درون‌زا تأثیر می‌گذارد.

فاصلهٔ اطمینان برای پارامترها در آزمون t

همچنان که در مباحث آزمون Z گفتیم، با داشتن $\hat{\beta}$ می‌توان برای مقدار واقعی β یک فاصلهٔ اطمینان ساخت. اگر بخواهیم در سطح معنی‌دار α این فاصله را بسازیم کافی است مقدار t را از جدول برای سطح معنی‌دار $\frac{\alpha}{۲}$ به دست آورده و آمارهٔ آزمون را در فاصلهٔ $\pm t$ جدول قرار دهیم،

$$-t_{\frac{\alpha}{۲}} \text{ جدول} < \text{آمارهٔ آزمون} < t_{\frac{\alpha}{۲}} \text{ جدول} \quad ,$$

یعنی

$$-t_{\frac{\alpha}{۲}} < \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} < t_{\frac{\alpha}{۲}} \quad .$$

دو طرف نامساوی را در $\mathrm{SE}(\hat{\beta})$ ضرب کرده و از دو طرف $\hat{\beta}$ را کم می‌کنیم. اگر نتیجه را در (۱ −) ضرب کنیم خواهیم داشت

$$\hat{\beta} - t_{\frac{\alpha}{2}}\ \mathrm{SE}(\hat{\beta}) < \beta < \hat{\beta} + t_{\frac{\alpha}{2}}\ \mathrm{SE}(\hat{\beta}) \quad . \qquad (۲\text{-}۴۱)$$

نامساوی ۲-۴۱ فاصلهٔ اطمینان α درصد برای β واقعی است. با داشتن فاصلهٔ اطمینان، می‌توان روش دیگری نیز برای آزمون فرضیه‌ها پیشنهاد کرد. اگر $\beta = k$ در فاصلهٔ اطمینان واقع شد، می‌گوییم فرضیهٔ H_0 رد نمی‌شود، در غیر این صورت H_0 را رد می‌کنیم.

مثال ۲-۹ در مثال ۲-۷ اولاً، برای α و β فاصلهٔ اطمینان ۹۵ درصد بسازید. ثانیاً، برای β فاصله اطمینان یکطرفه ۹۵ درصدی بسازید، ثالثاً، برای حالت یک، این فرضیه را آزمون کنید که آیا β برابر ۱/۱۲ است.

اولاً، با استفاده از رابطهٔ ۲-۴۱ داریم

$$۰/۷۵ - ۲/۳۰۶(۰/۲۵۶) < \beta < ۰/۷۵ + ۲/۳۰۶(۰/۲۵۶) ,$$

$$۰/۱۵۹ < \beta < ۱/۳۴۰ .$$

می‌توان راه حل فوق را به بیان دیگری نیز مطرح کرد. فاصلهٔ اطمینان ۹۵ درصد، یعنی فاصله‌ای که با احتمال ۹۵ درصد، مقدار آمارهٔ آزمون در نواحی بحرانی قرار نگیرد؛ به عبارت دیگر، آمارهٔ آزمون در فاصلهٔ مقادیر $t \pm$ به دست آمده از جدول t واقع شود؛ بنابراین

$$\Pr\left[-۲/۳۰۶ < \frac{\hat{\beta} - \beta}{\mathrm{SE}(\hat{\beta})} < ۲/۳۰۶\right] = ۰/۹۵ ,$$

که در آن Pr به معنای احتمال است. به ازای $\mathrm{SE}(\hat{\beta}) = ۰/۲۵۶$، داریم

$$۰/۱۵۹ < \beta < ۱/۳۴۰ .$$

به همین ترتیب برای α عمل می‌کنیم.

$$\Pr\left[-۲/۳۰۶ < \frac{\hat{\alpha} - \alpha}{\mathrm{SE}(\hat{\alpha})} < ۲/۳۰۶\right] = ۰/۹۵ ,$$

که به ازای $\mathrm{SE}(\hat{\alpha}) = ۲/۰۹$، داریم:

$$-۱/۲۲ < \alpha < ۸/۴۲ .$$

ثانیاً، فاصله‌های اطمینان یکطرفه، به حد بالا یا حد پایین پارامترها مربوط است. باید نقاطی را در روی محور t پیدا کنیم که سطح زیر منحنی از آن به بعد یا قبل از آن ۹۵ درصد باشد. با مراجعه به جدول t داریم

$$\Pr(t < ۱/۸۶) = ۰/۹۵ , \quad \Pr(t > -۱/۸۶) = ۰/۹۵ .$$

بنابراین برای یک فاصلهٔ اطمینان یکطرفه، حد بالای β عبارت است از

$$\hat{\beta} + ۱/۸۶\,\mathrm{SE}(\hat{\alpha}) = ۰/۷۵ + ۱/۸۶(۰/۲۵۶) = ۰/۷۵ + ۰/۴۷ = ۱/۲۲$$

و در نتیجه فاصلهٔ اطمینان ۹۵ درصدی یکطرفه برابر است با $(۱/۲۲ , -\infty)$. به همین ترتیب حد پایین β را به دست می‌آوریم،

$$\hat{\beta} - ۱/۸۶\,\mathrm{SE}(\hat{\alpha}) = ۰/۷۵ - ۱/۸۶(۰/۲۵۶) = ۰/۷۵ - ۰/۴۷ = -۰/۲۷ ,$$

ملاحظه می‌شود که فاصلهٔ اطمینان ۹۵ درصد یکطرفه، برابر است با $(+\infty , -۰/۲۷)$.

ثالثاً، چون $\beta = ۱/۱۲$ در فاصله اطمینان β برای حالت اول یعنی $۰/۱۶۴ < \beta < ۱/۳۳۵$ قرار می‌گیرد؛ بنابراین H_0 را نمی‌توان رد کرد.

مثال ۱۰-۲ تخمین تابع مصرف به شرح زیر مفروض است،

$$\underset{(۵/۶)}{C_t} = \underset{(۰/۰۱)}{۱۳ + ۰/۸۹\, Y_t^d} ,$$

که در آن C_t مصرف و Y_t^d درآمد قابل تصرف است. اعداد داخل پرانتز مقادیر انحراف معیار تخمین پارامترها هستند. می‌دانیم $n = ۱۰$.

اولاً، اگر ضریب ثابت را α بگیریم، فرضیهٔ $H_0 : \alpha = 0$ را در مقابل فرضیهٔ $H_1 : \alpha \neq 0$ در سطوح معنی‌دار ۱ درصد و ۵ درصد آزمون کنید.

ثانیاً، اگر میل نهایی به مصرف را β بگیریم، برای β یک بار فاصلهٔ اطمینان ۹۵ درصد و بار دیگر فاصلهٔ اطمینان ۹۹ درصد بسازید.

اولاً، چون آزمون، دو طرفه است باید مقدار t را از جدول به ازای $\alpha = \frac{0/01}{2} = 0/005$ و درجهٔ آزادی $\nu = 10 - 2 = 8$ به دست آوریم. خواهیم داشت $t = \pm 3/355$، حال باید آمارهٔ آزمون را به دست آوریم،

$$t = \frac{\hat{\alpha} - 0}{SE(\hat{\alpha})} = \frac{13}{5/6} = 2/32 .$$

با توجه به اینکه این آماره در ناحیهٔ بحرانی قرار نمی‌گیرد، فرضیهٔ H_0 رد نمی‌شود. اما در سطح معنی‌دار ۵ درصد مقدار به دست آمده از جدول t برابر است با $t_{0/025} = 2/306$ و ملاحظه می‌شود که آمارهٔ آزمون در ناحیهٔ بحرانی قرار می‌گیرد؛ در نتیجه فرضیهٔ H_0 رد می‌شود. این مثال نشان می‌دهد که در یک سطح معیّنی از معنی‌داری، می‌توان فرضیه‌ای را قبول کرد، در حالی که در سطح دیگر رد خواهد شد.

ثانیاً، برای به دست آوردن فاصلهٔ اطمینان ۹۵ درصد برای β، باید مقدار t را از جدول با ۸ درجه آزادی و به ازای سطح زیر منحنی ۰/۰۲۵ به دست آورد. خواهیم داشت $t = \pm 2/306$؛ بنابراین

$$0/89 - 2/306(0/01) < \beta < 0/89 + 2/306(0/01) ,$$

یا

$$0/866 < \beta < 0/913 .$$

اما برای فاصلهٔ اطمینان ۹۹ درصد کافی است مقدار t را به ازای سطح زیر منحنی ۰/۰۰۵ حساب کنیم؛ خواهیم داشت: $t = \pm 3/355$. در نتیجه فاصلهٔ اطمینان برای β عبارت است از:

$$0/89 - 3/355(0/01) < \beta < 0/89 + 3/355(0/01) ,$$

یا

$$۰/۸۵٦ < \beta < ۰/۹۲۳ \quad .$$

ملاحظه می‌شود که به ازای سطوح بالاتری از احتمال، فاصلهٔ اطمینان بیشتر می‌شود.

فاصله اطمینان و دقت تخمین پارامترها

در مثال ۱۰ـ۲ دیدیم که هر چه سطح احتمال بالاتر باشد، فاصلهٔ اطمینان بیشتر می‌شود. بدیهی است این نتیجه، چندان رضایت‌بخش نیست. حالت مطلوب این است که در سطوح بالای احتمال، فواصل کوچک اطمینان داشته باشیم. سؤال این است که چگونه ممکن است به این هدف نزدیک شویم.

یک بار دیگر رابطهٔ ٤۱ـ۲ را می‌نویسیم.

$$\hat{\beta} - t\,SE(\hat{\beta}) < \beta < \hat{\beta} + t\,SE(\hat{\beta}) \quad .$$

چون مقادیر $\hat{\beta}$ و t خارج از کنترل ما هستند بنابراین برای اینکه فاصلهٔ اطمینان β کمتر شود باید $SE(\hat{\beta})$ کوچکتر شود. اما می‌دانیم $SE(\hat{\beta})$ در واقع جذر واریانس $\hat{\beta}$ است؛ بنابراین باید واریانس $\hat{\beta}$ به کمترین مقدار خود برسد تا فاصلهٔ اطمینان برای β حداقل شود. حال ببینیم چگونه می‌توان واریانس $\hat{\beta}$ را کوچک کرد.

می‌دانیم

$$Var(\hat{\beta}) = \frac{\hat{\sigma}^2}{\Sigma x_t^2} \quad ,$$

که در آن $\hat{\sigma}^2 = s^2 = \frac{\Sigma e_t^2}{n-2}$. برای اینکه واریانس $\hat{\beta}$ کوچک شود، باید صورت کسر؛ یعنی $\hat{\sigma}^2$ کم شود و مخرج کسر، یعنی Σx_t^2 افزایش یابد. به نظر می‌رسد که ممکن است افزایش حجم مشاهدات صورت کسر $Var(\hat{\beta})$ را از طریق افزایش مخرج کسر $\frac{\Sigma e_t^2}{n-2}$ کوچکتر کند. افزایش حجم مشاهدات می‌تواند مخرج کسر $Var(\hat{\beta})$ را نیز از طریق افزایش $\sum_{t=1}^{n} x_t^2$ بزرگتر کند و در نتیجه واریانس $\hat{\beta}$ را کاهش دهد. چون افزایش حجم مشاهدات باعث بزرگتر شدن Σe_t^2 خواهد شد ـ که خود $\hat{\sigma}^2$ را بیشتر می‌کند ـ بنابراین

تأثیر مستقیمی در افزایش واریانس $\hat{\beta}$ نیز خواهد داشت. البته انتظار این است که خالص تأثیر افزایش حجم مشاهدات باعث کمتر شدن واریانس $\hat{\beta}$ شود؛ می‌دانیم هرچه واریانس یک تخمین زننده کمتر باشد، دقت آن تخمین‌زننده بیشتر خواهد بود.

اما مهمترین عاملی که می‌تواند واریانس $\hat{\beta}$ را کاهش دهد، این است که از طریق پراکندگی بیشتر مشاهدات، مقدار Σx_i^2 بیشتر شود. می‌دانیم

$$\sum x_i^2 = \sum (X_i - \bar{X})^2 \quad , \qquad (۲\text{-}۴۲)$$

در نتیجه هر چقدر مقادیر X_i پراکندگی بیشتری داشته باشد، تفاوت هر یک از آنها با میانگین $\bar{X}$ کمیت بزرگتری خواهد بود. وقتی این تفاوتها، یعنی $(X_i - \bar{X})$، مجذور شود، به مراتب بزرگتر شده و در نتیجه حاصل جمع بزرگتری خواهد داد. بدین ترتیب، می‌توان گفت که بهترین روش در کاهش واریانس $\hat{\beta}$، که خود موجب افزایش دقت تخمین شده و در نهایت فاصله اطمینان β را محدودتر می‌کند، این است که پراکندگی مشاهدات متغیر برون‌زا یا توضیحی بیشتر باشد.

این نکته در مباحث اقتصادسنجی کاربردی، اهمیت فراوان دارد؛ برای مثال، وقتی می‌خواهیم تابع مصرف را تخمین بزنیم که در آن درآمد به منزلهٔ متغیر توضیحی در نظر گرفته شده است، نباید نمونه‌ای که می‌گیریم به گروه خاصی از افرد در طبقهٔ معینی از درآمد منحصر باشد؛ زیرا در این صورت درآمد افراد موجود در نمونه، اختلاف چندانی با هم ندارد، بنابراین تفاوت بسیاری با میانگین نخواهد داشت. در نتیجه $\Sigma (X_i - \bar{X})^2$ بسیار کوچک خواهد بود. برعکس وقتی مشاهدات ما به تمام گروههای درآمدی مربوط باشد، اختلاف عناصر نمونه از میانگین کل بزرگ شده و Σx_i^2 افزایش خواهد یافت. به همین دلیل نمونه گیریها نباید به مناطق جغرافیایی خاص یا به دوره‌هایی محدود شود که سیاستهای شدید کنترل درآمد بر روند تغییرات اقتصادی حاکم است. برای مثال، اگر می‌خواهیم تابع سرمایه گذاری در بخش خصوصی را تخمین بزنیم که قیمت به عنوان یک متغیر توضیحی فرض شده است در آن صورت نباید دامنهٔ تغییرات قیمت را به دوره‌هایی محدود کرد که سیاست شدید کنترل قیمتها اعمال می‌شود.

آزمونهایی که تا اینجا مطرح کردیم، به فرضیه‌هایی مربوط می‌شود که می‌تواند در مورد هر یک از پارامترها مطرح شود. از مرحلهٔ بعد باید از آزمونهایی صحبت کنیم که دربارهٔ تخمین کل مدل نظر می‌دهد؛ یعنی به اصطلاح معنی‌دار بودن مدل راآزمون می‌کند. اما قبل از طرح آن، به آزمون واریانس جملهٔ اختلال می‌پردازیم.

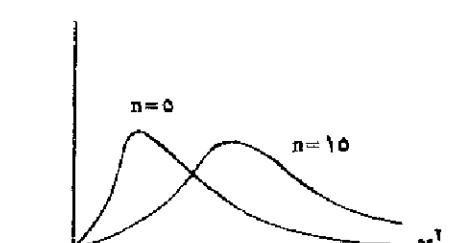
۳. آزمون واریانس جملهٔ اختلال

در مباحث آزمون t و در معادلهٔ ۳۸ـ۲ دیدیم که کمیت $\frac{\sum_{t=1}^{n} e_t^2}{\sigma^2}$، دارای توزیع χ^2 با (n − ۲) درجهٔ آزادی است،

$$\frac{\sum_{t=1}^{n} e_t^2}{\sigma^2} \sim \chi^2 (n-2) \quad .$$

بنابراین با استفاده از جدول χ^2 می‌توان در سطح معنی‌دار α، برای $\frac{\sum e_t^2}{\sigma^2}$ یک فاصلهٔ اطمینان به دست آورد.

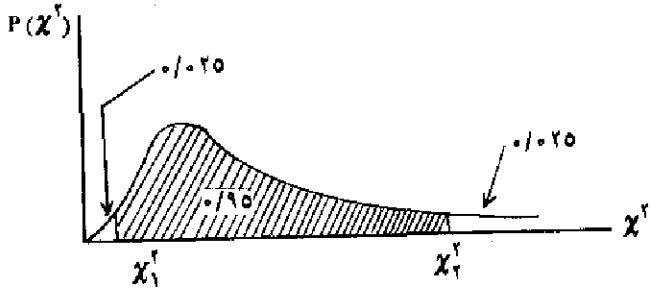
توزیع χ^2 همان‌گونه که در نمودار ۸ـ۲ ملاحظه می‌شود، معمولاً از مرکز مختصات شروع شده و به سمت راست تا بی‌نهایت ادامه‌دارد. شکل دقیق منحنی تغییرات χ^2 تابعی از درجات آزادی است. به ازای افزایش درجات آزادی، قرینگی منحنی توزیع χ^2 به طور مرتب بیشتر می‌شود. وقتی مقدار درجات آزادی خیلی بزرگ شود، توزیع χ^2 به‌سمت توزیع نرمال‌میل خواهد کرد. مقادیر مختلف χ^2 را می‌توان به ازای سطوح مختلف معنی‌داری و با درجات آزادی ازجدول χ^2 به‌دست آورد. چون توزیع χ^2 قرینگی‌ندارد، هر مقدار χ^2 که از جدول به دست آید، در واقع نقطه‌ای است که مساحت سطح زیر منحنی از آن نقطه به سمت راست برابر سطح معنی‌داری مورد نظر است؛ برای مثال، با درجه آزادی ۱۰، نقطه‌ای روی محور χ^2 که مساحت سطح زیر منحنی از آن نقطه به سمت راست ۵ درصد است



نمودار ۲.۸ توزیع χ^2

برابر ۱۸/۳۰۷ خواهد بود.[۱] در جدولهایی که برحسب توزیع تراکمی χ^2 تنظیم شده‌اند[۲] معمولاً مساحت سطح زیر منحنی از مبدأ مختصات تا نقطهٔ χ^2، در ردیف اول از بالای جدول قرار دارد. مقدار χ^2 که از متن جدول به دست می‌آید در واقع نقطه‌ای است که مساحت سطح زیر منحنی از مبدأ مختصات تا آن نقطه برابر سطح مندرج در ردیف اول است، برای‌مثال، به ازای درجهٔ آزادی ۱۰ مقدار $\chi^2 = ۱۸/۳$ نقطه‌ای روی محور χ^2 است که مساحت سطح زیر منحنی از مبدأ مختصات تا آن نقطه برابر ۹۵ درصد است.

در اینجا فرض کنید می‌خواهیم یک فاصلهٔ اطمینان ۹۵ درصدی برای $\frac{\sum_{i=1}^{n} e_i^2}{\sigma^2}$ به دست آوریم؛ یعنی سطحی از منحنی تغییرات χ^2 را مشخص کنیم که ۹۵ درصد در وسط و ۰/۰۲۵ در هر طرف باشد. مقدار هاشور خورده نمودار ۹ـ۲ جواب مسأله است؛ بنابراین باید نقاط χ_1^2 و χ_2^2 را با استفاده از جدول به دست آوریم.



نمودار ۲.۹ فاصلهٔ اطمینان ۹۵ درصد برای χ^2

می‌توان از جدول توزیع تراکمی χ^2 استفاده کرد. χ_1^2 نقطه‌ای است که سطح زیر منحنی از مبدأ مختصات تا آن برابر ۰/۰۲۵ است؛ بنابراین آن را با $\chi^2_{۰/۰۲۵}$ نشان می‌دهیم. به همین ترتیب χ_2^2 نقطه‌ای است که انتگرال $P(\chi^2)$ از ۰ تا χ_2^2 برابر ۰/۹۷۵ است. بنابراین می‌توان χ_2^2 را با $\chi^2_{۰/۹۷۵}$ نشان داد. در نتیجه، فاصلهٔ اطمینان ۹۵ درصد برای $\frac{\sum e_i^2}{\sigma^2}$ برابر است با

---

۱. به پیوست «۲ ـ ب» رجوع شود.

۲. به پیوست «۳ ـ ب» رجوع شود.

$$\chi^2_{۰/۰۲۵} < \frac{\sum e_t^2}{\sigma^2} < \chi^2_{۰/۹۷۵} .$$

به عبارت دقیقتر، احتمال اینکه $\frac{\sum e_t^2}{\sigma^2}$ بین $\chi^2_{۰/۹۷۵}$ و $\chi^2_{۰/۰۲۵}$ قرار گیرد، ۹۵ درصد است، یعنی

$$\Pr\left(\chi^2_{۰/۰۲۵} < \frac{\sum e_t^2}{\sigma^2} < \chi^2_{۰/۹۷۵}\right) = ۰/۹۵ .$$

اما هدف اصلی ما پیدا کردن فاصلهٔ اطمینان برای σ^2 است. به عبارت دیگر، وقتی در یک تخمین $\hat{\sigma}^2_u$ را به دست می‌آوریم، به طور طبیعی این سؤال مطرح می‌شود که آیا می‌توان با داشتن $\hat{\sigma}^2$ به حدود یا فاصلهٔ اطمینانی برای مقدار واقعی σ^2 رسید؟ برای پاسخ به این سؤال کافی است از معادلهٔ ۲ـ۲۸ استفاده کنیم، می‌دانیم

$$\hat{\sigma}^2_u = \frac{\sum e_t^2}{n-۲} ,$$

در نتیجه

$$\sum e_t^2 = (n-۲)\,\hat{\sigma}^2 . \qquad (۲ـ۴۳)$$

با جایگزینی ۲ـ۴۳ در فاصلهٔ اطمینان ۹۵ درصد برای $\frac{\sum e_t^2}{\sigma^2}$، خواهیم داشت

$$\chi^2_{۰/۰۲۵} < \frac{(n-۲)\,\hat{\sigma}^2}{\sigma^2} < \chi^2_{۰/۹۷۵} .$$

دو طرف نامساوی را معکوس کرده، در $(n-۲)\,\hat{\sigma}^2$ ضرب می‌کنیم. خواهیم داشت

$$\frac{(n-۲)\,\hat{\sigma}^2}{\chi^2_{۰/۹۷۵}} < \sigma^2 < \frac{(n-۲)\,\hat{\sigma}^2}{\chi^2_{۰/۰۲۵}} . \qquad (۲ـ۴۴)$$

نامساوی ۲ـ۴۴ فاصله اطمینان ۹۵ درصد را برای σ^2 مشخص می‌کند.

در پایان به این نکته اشاره می‌کنیم که فاصله‌های اطمینان $\hat{\alpha}$ و $\hat{\beta}$ـ که قبلاً به دست آوردیم ـ حول مقادیر α و β قرینگی دارند؛ در حالی که فاصلهٔ اطمینان σ^2 حول مقدار $\hat{\sigma}^2$

قرینه نیست. دلیل این امر عدم وجود قرینگی در توزیع χ^2 است؛ در حالی که توزیع Z و t هر کدام قرینگی کامل دارند.

مثال ۱۱ـ۲ با توجه به مدل تولید و نیروی کار، موضوع مثال ۷ـ۲، برای σ^2 فاصلهٔ اطمینان ۹۵ درصد بسازید.

نامساوی ٤٤ـ۲ را می‌نویسیم،

$$\frac{(n-2)\,\hat{\sigma}^2}{\chi^2_{/۹۷۵}} < \sigma^2 < \frac{(n-2)\,\hat{\sigma}^2}{\chi^2_{/۰۲۵}} .$$

با استفاده از مثال ۷ـ۲ می‌دانیم $n = ۱۰$ و $\hat{\sigma}^2 = ۱/۸۳$ است، در اینجا باید $\chi^2_{/۰۲۵}$ و $\chi^2_{/۹۷۵}$ را با استفاده از جدول χ^2 تراکمی و با $۸ = ۲ - ۱۰$ درجهٔ آزادی به دست آوریم. خواهیم داشت

$$\chi^2_{/۰۲۵}(۸) = ۲/۱۸ \quad , \quad \chi^2_{/۹۷۵}(۸) = ۱۷/۵ ,$$

بنابراین:

$$\Pr\left(\frac{۸\,\hat{\sigma}^2}{۱۷/۵} < \sigma^2 < \frac{۸\,\hat{\sigma}^2}{۲/۱۸}\right) = ۰/۹۵ .$$

به عبارت دیگر در سطح احتمال ۹۵ درصد، فاصلهٔ اطمینان برای σ^2 عبارت است از

$$\frac{۸\,(۱/۸۳)}{۱۷/۵} < \sigma^2 < \frac{۸\,(۱/۸۳)}{۲/۱۸} ,$$

یا

$$۰/۸۳ < \sigma^2 < ٦/۷۱ .$$

۲ـ٤ آنالیز واریانس و آزمون معنی‌دار بودن مدل رگرسیون

می‌دانیم آنالیز واریانس[1] یکی از مفاهیم بسیار مهم آمار است. در این قسمت به کاربرد آن در رگرسیون خطی ساده اشاره می‌کنیم. در فصل اول، قسمت ۱ـ٤ وقتی که بحث ما

1. Analysis of Variance

دربارهٔ ضریب تعیین بود، قضیهٔ بسیار مهم ۴۰ـ۱ را ثابت کردیم، که

$$\sum y_t^2 = \sum \hat{y}_t^2 + \sum e_t^2 \ ,$$

یا

$$\text{TSS} = \text{ESS} + \text{RSS} \ .$$

این قضیه نشان می‌دهد که تغییرات متغیر درون‌زا (TSS) را می‌توان به دو قسمت تقسیم کرد: تغییرات توضیح داده شده (ESS)، و تغییرات توضیح داده نشده (RSS).

یکی از هدفهای آنالیز واریانس این است که بتواند معنی‌دار بودن تغییرات توضیح داده شده را آزمون کند. البته در مورد مدل رگرسیون ساده، مفهوم این آزمون چیزی نیست جز اینکه $\beta = 0$ را آزمون کنیم.

معمولاً تقسیم‌بندی تغییرات متغیر درون‌زا به تغییرات توضیح داده شده و توضیح داده نشده را همراه با درجات آزادی و میانگین هر کدام در یک جدول به شرح ذیل تنظیم می‌کنند و آن را جدول تجزیهٔ واریانس می‌نامند.

جدول ۱ـ۲ آنالیز واریانس برای مدل رگرسیون ساده

منبع تغییرات	مجموع مربعات	درجات آزادی	میانگین مربعات یا واریانس
تخمین رگرسیون	$\text{ESS} = \sum \hat{y}_t^2 = \hat{\beta} \sum x_t y_t$	۱	$\text{ESS} / ۱$
پسماند	$\text{RSS} = \sum e_t^2 = \sum y_t^2 - \hat{\beta} \sum x_t y_t$	$(n-۲)$	$\text{RSS} / (n-۲)$
کل تغییرات	$\text{TSS} = \sum y_t^2$	$(n-۱)$	$F = \dfrac{\text{ESS} / ۱}{\text{RSS} / (n-۲)}$

برای تبیین بیشتر این جدول، به نکات زیر اشاره می‌کنیم. اثبات رابطهٔ

$$\text{ESS} = \hat{\beta} \sum x_t y_t \ ,$$

بسیار ساده است. با استفاده از معادلهٔ ۳۱ـ۱، داریم

$$\text{ESS} = \sum \hat{y}_t^2 = \sum \hat{\beta}^2 x_t^2 \ ,$$

$$= \hat{\beta}^2 \sum x_t^2 = \hat{\beta} \frac{\sum x_t y_t}{\sum x_t^2} \cdot \sum x_t^2 \ ,$$

یا

$$ESS = \hat{\beta} \sum x_t y_t \ . \qquad (۲\text{-}۴۵)$$

برای اثبات

$$RSS = \sum y_t^2 - \hat{\beta} \sum x_t y_t \ ,$$

می‌گوییم بنابر تعریف

$$RSS = \sum (Y_t - \hat{\alpha} - \hat{\beta} X_t)^2 \ ,$$

$$= \sum [(Y_t - \bar{Y}) - \hat{\beta} (X_t - \bar{X})]^2 \ ,$$

$$= \sum y_t^2 + \hat{\beta}^2 \sum x_t^2 - 2 \hat{\beta} \sum x_t y_t \ ,$$

$$= \sum y_t^2 - \frac{(\sum x_t y_t)^2}{\sum x_t^2} \ ,$$

در نتیجه

$$\sum e_t^2 = RSS = \sum y_t^2 - \hat{\beta} \sum x_t y_t \ , \qquad (۲\text{-}۴۶)$$

و با جایگزینی $\hat{\beta}$ خواهیم داشت[1]

$$\sum e_t^2 = RSS = \sum y_t^2 - \frac{(\sum x_t y_t)^2}{\sum x_t^2} \ . \qquad (۲\text{-}۴۷)$$

---

۱. فرمول ۲ـ۴۷ را می‌توان به صورت زیر نیز نوشت. از $\sum y_t$ فاکتور می‌گیریم.

$$\sum e_t^2 = \sum y_t \left[1 - \frac{(\sum x_t y_t)^2}{\sum x_t^2 \sum y_t^2} \right] \ ,$$

در نتیجه

$$\sum e_t^2 = (1 - r^2) \sum y_t^2$$

از این فرمول در فصل چهارم و در قسمت ۴ـ۳ استفاده خواهیم کرد.

ملاحظه می‌شود که اگر تغییرات توضیح داده شده (ESS) را با تغییرات توضیح داده شده (RSS) جمع کنیم حاصل، Σy_i^2 کل تغییرات خواهد بود. بحث فوق در واقع روش دیگری برای اثبات قضیهٔ بسیار مهم ۴۰-۱ است.

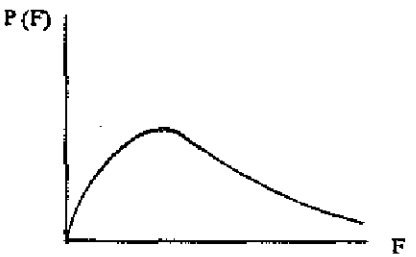
بحث بیشتر در مورد جدول ۱-۲ مستلزم بررسی توزیع احتمال F است؛ بنابراین در همینجا به مفاهیم اساسی این توزیع اشاره کرده، سپس به آزمون معنی‌دار بودن مدل رگرسیون بر می‌گردیم.

۱. توزیع و آزمون F

توزیع F برحسب دو متغیر مستقل از هم که هر یک توزیع χ^2 دارند تعریف می‌شود. متغیرهای تصادفی U و V با توزیعهای χ^2 و مستقل از یکدیگر مفروض است. همچنین فرض بر این است که درجات آزادی U و V به ترتیب ν_1 و ν_2 است. در این صورت آماره زیر، توزیع F با درجات آزادی ν_1 و ν_2 خواهد داشت،

$$F = \frac{\frac{U}{\nu_1}}{\frac{V}{\nu_2}} .$$

جدول F برحسب درجات آزادی ν_1 و ν_2 مدون شده است. ν_1 و ν_2 به ترتیب درجات آزادی صورت و مخرج کسر F است. توزیع F مانند توزیع χ^2 از مرکز مختصات شروع می‌شود و به طرف راست تا بی‌نهایت ادامه دارد[1]. نمودار ۱۰-۲ منحنی تغییرات F را نشان می‌دهد. توزیع F معمولاً برای مواردی به کار می‌رود که می‌خواهیم چند فرضیه را به طور همزمان در مورد



نمودار ۱۰-۲ تابع توزیع احتمال F

۱. به بخش ضمائم، جدول ۵ ـ ب رجوع کنید.

پارامترها آزمون کنیم؛ بنابراین آزمونF بیشتر برای آن دسته از مدلهای رگرسیون مفید است که بیش از یک متغیر توضیحی دارند (این نکته را در فصل ششم و در قسمت ٤ـ٦، در مباحث آنالیز واریانس در رگرسیون چند متغیره توضیح خواهیم داد). همچنین ـ همان گونه که در آمار دیده‌ایم ـ می‌توان از توزیع F برای آزمون تساوی دو واریانس نیز استفاده کرد.

در اینجا پس از یادآوری تعریف توزیع F به بحث آنالیز واریانس در مدلهای رگرسیون ساده می‌پردازیم. با توجه به معادلهٔ ۳۰ـ۲ داریم

$$\frac{\hat{\beta} - \beta}{\sqrt{\sigma^2 / \Sigma x_t^2}} \sim N(0, 1) .$$

می‌دانیم اگر توزیع نرمال استاندارد را مجذور کنیم، توزیع χ^2 خواهیم داشت؛ بنابراین با توجه به رابطهٔ ۳۷ـ۲ اگر کسر فوق را مجذور کنیم، توزیع χ^2 با یک درجهٔ آزادی به دست می‌آید،

$$\frac{(\hat{\beta} - \beta)^2}{\sigma^2 / \Sigma x_t^2} \sim \chi^2(1) . \qquad (۲ـ٤۸)$$

از معادلهٔ ۳۸ـ۲ نیز می‌دانیم که

$$\frac{\Sigma e_t^2}{\sigma^2} \sim \chi^2(n-2) .$$

با توجه به تعریف توزیع F، اگر معادلهٔ ٤۸ـ۲ و ۳۸ـ۲ را به ترتیب بر درجات آزادی آنها تقسیم کرده و نتایج را نیز بر هم تقسیم کنیم و کسر حاصل را F بنامیم، آنگاه F دارای توزیع F با درجات آزادی ۱ و (n − ۲) خواهد بود؛ به گونه‌ای که درجه آزادی ۱ مربوط به صورت و درجه آزادی (n − ۲) متعلق به مخرج است؛

$$F = \frac{(\hat{\beta} - \beta)\Sigma x_t^2 / 1}{\Sigma e_t^2 / (n-2)} \sim F(1, n-2) . \qquad (۲ـ٤۹)$$

گفتیم که هدف آنالیز واریانس آزمون معنی‌دار بودن تغییرات توضیح داده شده است. به عبارت دیگر، اگر بخواهیم آنچه از تغییرات متغیر درون‌زا را ـ که توضیح داده شده است ـ آزمون کنیم، در واقع مانند این است که فرضیهٔ $H_0 : \beta = 0$، یعنی کل مدل را آزمون کرده باشیم. حال اگر فرض کنیم که فرضیهٔ H_0 صحیح است؛ یعنی $\beta = 0$، آنگاه آمارهٔ F در معادلهٔ ۴۹ـ۲ را می‌توان چنین نوشت،

$$F = \frac{\hat{\beta}^2 \Sigma x_t^2 / 1}{\Sigma e_t^2 / (n-2)} \sim F(1, n-2) \quad . \qquad (۲\text{ـ}۵۰)$$

مقدار F در معادلهٔ ۵۰ـ۲ را آمارهٔ آزمون می‌نامیم و می‌گوییم اگر فرضیهٔ $\beta = 0$ صحیح باشد، آنگاه مقدار آمارهٔ آزمون در معادلهٔ ۵۰ـ۲ توزیع F خواهد داشت. بدیهی است برای هر سطح معنی‌داری که برای آزمون در نظر می‌گیریم، می‌توان F متناظر با نقطهٔ بحرانی را از جدول F به دست آورد. اگر آمارهٔ آزمون در ناحیهٔ بحرانی قرار بگیرد، فرضیهٔ H_0 رد می‌شود، در غیر این صورت H_0 رد نمی‌شود؛ پس نتیجه می‌گیریم که مدل، معنی‌دار نیست؛ یعنی X_t نمی‌تواند تغییرات Y_t را توضیح دهد.

رابطهٔ ۵۰ـ۲ را به صورت ساده‌تری می‌نویسیم. با استفاده از معادلهٔ ۳۱ـ۱ داریم

$$\hat{\beta}^2 \sum x_t^2 = \sum \hat{y}_t^2 = \text{ESS} \quad ,$$

در نتیجه خواهیم داشت

$$F = \frac{\text{ESS} / 1}{\text{RSS} / (n-2)} \sim F(1, n-2) \quad . \qquad (۲\text{ـ}۵۱)$$

به فرمول ۵۱ـ۲ که بر اساس فرضیهٔ $H_0 : \beta = 0$ استوار است، آزمون مدل رگرسیون[1] یا آزمون معنی‌دار بودن مدل رگرسیون[2] می‌گویند. ملاحظه می‌شود که رابطهٔ ۵۱ـ۲ چیزی نیست جز نسبت دو مقداری که در ستون آخر جدول تجزیه واریانس (جدول ۱ـ۲)

1. Testing the Regression Equation
2. Testing the Significance of Regression Equation

به دست آوردیم.

مثال ۱۲ـ۲ مدل رابطهٔ تولید با ساعتهای نیروی کار، موضوع مثالهای ۲ـ۱ و ۱ـ۲ را در نظر بگیرید. برای این مدل، جدول آنالیز واریانس را تشکیل دهید و آزمون F را برای $\beta = 0$ در سطح معنی‌دار ۵ درصد انجام دهید.

با استفاده از محاسبه‌های قبلی می‌دانیم

$$\sum x_i y_i = ۲۱ \ , \qquad \hat{\beta} = ۰/۷۵ \ ,$$

$$\sum x_i^2 = ۲۸ \ , \qquad \sum y_i^2 = ۳۰/۴ \ ,$$

و با استفاده از فرمول ۴۵ـ۲ داریم

$$ESS = ۰/۷۵ \ (۲۱) = ۱۵/۷۵ \ ,$$

و با استفاده از فرمول ۴۶ـ۲ خواهیم داشت

$$RSS = ۳۰/۴ - ۰/۷۵ \ (۲۱) = ۱۴/۶۵ \ .$$

به این ترتیب جدول آنالیز واریانس برای این مثال به صورت جدول ۲ـ۲ به دست می‌آید.

جدول ۲ـ۲ مقادیر عددی جدول آنالیز واریانس برای اعداد جدول ۱ـ۴

منبع تغییرات	مجموع مربعات	درجات آزادی	میانگین مربعات یا واریانس
تغییرات X_i	۱۵/۷۵	۱	۱۵/۷۵
تغییرات e_i	۱۴/۶۵	۸	۱/۸۳
کل تغییرات	۳۰/۴۰	۹	$F = \frac{۱۵/۷۵}{۱/۸۳} = ۸/۶۰۶$

آمارهٔ آزمون از فرمول ۵۱ـ۲ به سهولت قابل محاسبه است.

$$F = \frac{۱۵/۷۵}{۱/۸۳} = ۸/۶۰۶ \ .$$

مقدار به دست آمده از جدول F با درجات آزادی صورت و مخرج به ترتیب برابر ۱ و ۸ و در سطح معنی‌دار ۵ درصد، برابر با ۵/۳۲ است. بدین ترتیب در سطح معنی دار ۵ درصد، فرضیهٔ H_0 رد می‌شود.

۲. رابطهٔ بین F، t، r^2

در مثال فوق دیدیم که آمارهٔ آزمون F برابر ۸/٦۰٦ به دست آمد. قبلاً در مثال ۸ـ۲ ملاحظه شد که برای همین مورد، مقدار t برابر با ۲/۹۳ است؛ بنابراین مقدار F برابر مجذور مقدار t است. می‌توان این نکته را در حالت کلی ثابت کرد. از معادلهٔ ۳٦ـ۲ می‌دانیم که

$$t = \frac{\hat{\beta} - \beta}{\sqrt{s^2 / \Sigma x_t^2}} ,$$

و بنابر فرضیهٔ $\beta = 0 : H_0$، خواهیم داشت

$$t = \frac{\hat{\beta}}{\sqrt{s^2 / \Sigma x_t^2}} ,$$

مقدار s^2 را از معادلهٔ ۲۸ـ۲ در رابطهٔ فوق جایگزین کرده، t^2 را محاسبه می‌کنیم،

$$t^2 = \frac{\hat{\beta}^2 \Sigma x_t^2}{\Sigma e_t^2 / (n-2)} = \frac{\hat{\beta}^2 \Sigma x_t^2 / 1}{\Sigma e_t^2 / (n-2)} ,$$

که دقیقاً همان فرمول ۵۰ـ۲ است. نتیجه می‌گیریم که برای آزمون $\beta = 0 : H_0$، خواهیم داشت

$$F = t^2 . \qquad (۲ـ۵۱)$$

در اینجا برای یافتن رابطهٔ بین F و r^2، معادلهٔ ۲۹ـ۱ را دوباره می‌نویسیم،

$$r^2 = \frac{ESS}{TSS} ,$$

در نتیجه

$$ESS = r^2 TSS . \qquad (۲ـ۵۲)$$

همچنین از معادلهٔ ۴۳ـ۱ داریم

$$RSS = (1 - r^2)\, TSS \quad . \qquad (۲ـ۵۳)$$

با جایگزینی معادله‌های ۵۲ـ۲ و ۵۳ـ۲ در ۵۱ـ۲ خواهیم داشت

$$F = \frac{r^2/1}{(1 - r^2)/(n - 2)} \quad ,$$

یا

$$F = \frac{(n - 2)\, r^2}{1 - r^2} \quad . \qquad (۲ـ۵۴)$$

برای یافتن رابطهٔ بین r و t کافی است معادلهٔ ۵۱ـ۲ را در معادلهٔ ۵۴ـ۲ قرار دهیم،

$$t^2 = \frac{(n - 2)\, r^2}{1 - r^2} \quad , \qquad (۲ـ۵۵)$$

به همین ترتیب، خواهیم داشت

$$r^2 = \frac{t^2}{t^2 + (n - 2)} \quad . \qquad (۲ـ۵۶)$$

فرمول ۵۶ـ۲، رابطه بین مقدار آمارهٔ t برای فرضیهٔ $\beta = 0$ و ضریب تعیین r^2 را مشخص می‌کند.

در پایان مباحث آزمونها، به این نکته اشاره می‌کنیم که سه آزمون فوق در حقیقت صورتهای مختلف یک واقعیت است؛ به عبارت دیگر، اولاً، آزمونی که شامل r^2 (فرمول ۵۴ـ۲)، آزمون معنی‌دار بودن ضریب تعیین است؛ ثانیاً، آزمونی که بر اساس $\hat{\beta}$ (فرمول ۳۶ـ۲) استوار می‌شود، در واقع آزمون معنی‌دار بودن شیب مدل رگرسیون خواهد بود؛ ثالثاً، از طریق آنالیز واریانس کوشش می‌شود که معنی‌دار بودن تغییرات توضیح داده شده آزمون شود.

مثال ۱۳ـ۲ با استفاده از t یا F در مثال ۱۲ـ۲ مقدار r^2 را به دست آورید.

می‌دانیم t^2 یا F در این مثال برابر ۸/۶ است. با استفاده از فرمول ۵۶ـ۲، r^2

به سهولت به دست می‌آید،

$$r^2 = \frac{۸/۶}{۸/۶ + ۸} = \frac{۸/۶}{۱۶/۶} = ۰/۵۲ \quad .$$

۵ـ۲ خصوصیات مطلوب تخمین‌زننده‌ها

در این قسمت، ابتدا مفهوم خصوصیات مطلوب یک تخمین‌زننده را در حالت کلی بررسی کرده، سپس انطباق این خصوصیات مطلوب را با تخمین‌زننده‌های حداقل مربعات معمولی (OLS) در یک مدل رگرسیون ساده مطالعه خواهیم کرد. بنابراین ضرورتاً از مباحث اقتصادسنجی خارج شده، مطالب صرفاً آماری را ارائه می‌دهیم.

متغیر تصادفی X_i مفروض است. فرض کنید θ یکی از پارامترهایی است که تابع توزیع احتمال این متغیر تصادفی را مشخص می‌کند. هدف ما تخمین θ است جامعهٔ مشاهدات X_i ، شامل تمام مقادیر ممکن این متغیر بوده و θ نیز یکی از مشخصه‌های پارامتریکی[1] این جامعه است. X_i می‌تواند یک متغیر پیوسته یا ناپیوسته فرض شود؛ برای مثال، می‌توان X_i را درآمد خانوار فرض کرد؛ بنابراین میانگین درآمد، یعنی θ ، یکی از مشخصه‌های پارامتر یکی این جامعه محسوب خواهد شد. می‌خواهیم میانگین این جامعه را تخمین بزنیم. بدیهی است این مثال ساده‌ترین حالت در مسأله تخمین است. در موارد پیشرفته، هدف این است که بتوان چند پارامتر در مورد چند متغیر را همزمان تخمین زد؛ با وجود این، برای بررسی مشخصه‌های مطلوب تخمین‌زننده‌ها، بهتر است از ساده‌ترین حالت شروع کنیم؛ زیرا نتایج حاصل برای مدلهای پیچیده‌تر نیز دقیقاً صادق خواهد بود.

برای تخمین یک پارامتر، معمولاً دو سری اطلاعات در دسترس است؛ اطلاعات اولیه[2] و اطلاعات حاصل از نمونه. معلوماتی که در مورد جامعه مشاهدات X_i داریم، اطلاعات اولیه نامیده می‌شود؛ مانند فرضهای صورت تابع توزیع احتمال، یا مقادیر بعضی دیگر از پارامترهای جامعه غیر از θ یا حتی فرضهایی که در مورد قلمرو تغییرات

1. Parametric Characteristics
2. Prior Information

θ ارائه شده است. اطلاعات حاصل از نمونه، شامل کلیه اطلاعاتی است که از نمونهٔ n تایی $X_1, X_2, \ldots, X_n$ به دست می‌آید. سؤال این است که چگونه می‌توان از این دو سری اطلاعات برای تخمین پارامتر θ استفاده کرد. برای پاسخ به این سؤال باید به فرمولی مراجعه کرد که برای تخمین پارامتر به دست می‌آوریم، به این فرمول «تخمین‌زننده» می‌گویند. معمولاً همواره می‌توان به بیش از یک فرمول برای تخمین θ رسید؛ به عبارت دیگر، برای تخمین θ، همواره تخمین‌زننده‌های متعددی وجود دارد. سؤال این است که چگونه می‌توان از بین تخمین‌زننده‌های مختلف، مناسبترین را انتخاب کرد. بدیهی است هر تخمین‌زننده، خصوصیات خاص خود را دارد. بنابراین برای اینکه بتوان مناسبترین تخمین‌زننده را برگزید؛ ابتدا باید خصوصیات مطلوب تخمین‌زننده‌ها را بررسی کرد. بهترین تخمین‌زننده آن است که بیشترین خصوصیات مطلوب را دارا باشد.

در قسمت ۲ـ۱ دیدیم که به یک تخمین‌زننده از θ اصطلاحاً $\hat{\theta}$ می‌گوییم. می‌دانیم $\hat{\theta}$ نیز یک متغیر تصادفی است، بنابراین دارای میانگین و واریانس است. اگر میانگین $\hat{\theta}$ را با $E(\hat{\theta})$ نشان دهیم، آنگاه بنا بر تعریفْ واریانس $\hat{\theta}$ برابر است با

$$\mathrm{Var}(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2 .$$

می‌دانیم انحراف معیار $\hat{\theta}$ برابر است با جذر واریانس $\hat{\theta}$،

$$SE(\hat{\theta}) = \sqrt{\mathrm{Var}(\hat{\theta})} .$$

حال مفاهیم زیر را تعریف می‌کنیم.

$$\text{خطای نمونه‌گیری} = \hat{\theta} - \theta ,$$

$$\text{اریب} = E(\hat{\theta}) - \theta ,$$

$$\text{میانگین مربع خطا} = E(\hat{\theta} - \theta)^2 .$$

«خطای نمونه گیری»[1]، در واقع تفاوت مقدار تخمین و مقدار واقعی پارامتری است که باید تخمین زده شود. «اریب»[2] به تفاوت بین میانگین توزیع تخمین‌زننده $E(\hat{\theta})$ و مقدار واقعی پارامتر θ گفته می‌شود. مقدار اریب برای هر تخمین‌زننده معمولاً مقدار معینی است که می‌تواند صفر یا غیر صفر باشد. سرانجام میانگین مربع خطا (MSE)[3] مفهومی است که با پراکندگی توزیع یک تخمین‌زننده مرتبط بوده بنابراین بسیار نزدیک به مفهوم واریانس است. تفاوت بین واریانس و میانگین مربع خطای یک تخمین‌زننده این است که واریانس در واقع پراکندگی توزیع تخمین زننده را در حول مقدار میانگین آن اندازه گیری می‌کند؛ در حالی که میانگین مربع خطا، پراکندگی توزیع تخمین‌زننده را در اطراف مقدار واقعی پارامتر مشخص می‌سازد. اگر میانگین توزیع $\hat{\theta}$، یعنی $E(\hat{\theta})$ بر مقدار واقعی پارامتر یعنی θ منطبق شود، $\text{Var}(\hat{\theta})$ و میانگین مربع خطای $\hat{\theta}$، یعنی $\text{MSE}(\hat{\theta})$، با هم برابر می‌شود، در غیر این صورت قطعاً با یکدیگر متفاوتند.

رابطهٔ بین میانگین مربع خطای $\hat{\theta}$ و واریانس $\hat{\theta}$ را می‌توان به صورت زیر به دست آورد. می‌دانیم

$$\text{MSE}(\hat{\theta}) = E(\hat{\theta} - \theta)^2 .$$

به طرف راست و در داخل پرانتز $\pm E(\hat{\theta})$ را اضافه می‌کنیم

$$\begin{aligned}
\text{MSE}(\hat{\theta}) &= E[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2 , \\
&= E\left\{[\hat{\theta} - E(\hat{\theta})] + [E(\hat{\theta}) - \theta]\right\}^2 , \\
&= E[\hat{\theta} - E(\hat{\theta})]^2 + [E(\hat{\theta}) - \theta]^2 + \\
&\quad 2E\left\{[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta]\right\} .
\end{aligned}$$

جملهٔ آخر برابر صفر است؛ زیرا بنا بر قواعد امید ریاضی، می‌دانیم $E(a\,x) = a\,E(x)$ و

1. Sampling Error 2. Bias 3. Mean Square Error

$E(a) = a$، بنابراین

$$E\left\{[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta]\right\} = [E(\hat{\theta}) - \theta]\, E[\hat{\theta} - E(\hat{\theta})]\ ,$$

$$= [E(\hat{\theta}) - \theta][E(\hat{\theta}) - E(\hat{\theta})]\ ,$$

$$= 0\ .$$

بدین ترتیب $MSE(\hat{\theta})$ عبارت است از

$$MSE(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2 + [E(\hat{\theta}) - \theta]^2\ ,$$

یا

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + (\text{اریب})^2\ . \qquad (۲\text{-}۵۷)$$

یعنی میانگین مربع خطای $\hat{\theta}$ برابر است با واریانس $\hat{\theta}$ به علاوه مربع مقدار اریب. در نتیجه مقدار میانگین مربع خطای $\hat{\theta}$، هیچگاه نمی‌تواند از مقدار واریانس $\hat{\theta}$ کمتر باشد:

$$MSE(\hat{\theta}) \geq Var(\hat{\theta})\ . \qquad (۲\text{-}۵۸)$$

تفاوت $MSE(\hat{\theta})$ و $Var(\hat{\theta})$ دقیقاً برابر مربع مقدار اریب است.

بعد از ذکر این مقدمه، به بررسی خصوصیاتی از $\hat{\theta}$ می‌پردازیم که معمولاً به عنوان خصوصیات مطلوب تخمین‌زننده‌ها شناخته می‌شود. بحث را در دو قسمت خصوصیات مطلوب در نمونه‌های محدود یا نمونه‌های کوچک[1] و خصوصیات مطلوب حدی در نمونه‌های بزرگ[2] ارائه می‌دهیم. در حالت اول خصوصیاتی از $\hat{\theta}$ را بررسی می‌کنیم که حاصل نمونه‌هایی با حجم محدود و معین است و در حالت دوم خصوصیات $\hat{\theta}$ متعلق به نمونه‌هایی را به دست می‌آوریم که حجم آنها به سمت بی‌نهایت میل می‌کند.

---

1. Finite or Small Sample Properties
2. Asymptotic or Large Sample Properties

۱. خصوصیات مطلوب تخمین‌زننده‌ها در نمونه‌های کوچک

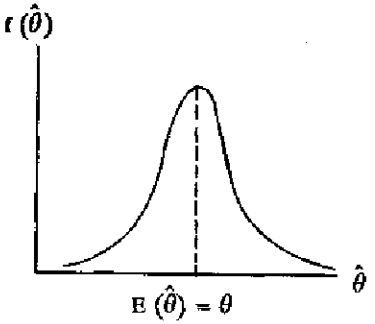
در این قسمت به بررسی خصوصیات مطلوب $\hat{\theta}$، هنگامی که حاصل نمونه‌ای با حجم معین و محدود است، می‌پردازیم:

نااریبی

اولین و شاید مهمترین خصوصیات $\hat{\theta}$ نااریبی آن است. بنا بر تعریف $\hat{\theta}$ یک تخمین‌زنندهٔ نااریب از θ است.

هرگاه

$$E(\hat{\theta}) = \theta \quad . \qquad (۲\_۵۹)$$



نمودار ۱۱ـ۲ ناریبی $\hat{\theta}$

نمودار ۱۱ـ۲ خصوصیت ناریبی $\hat{\theta}$ را نشان می‌دهد. ملاحظه می‌شود که چون توزیع $\hat{\theta}$ قرینگی دارد؛ بنابراین میانگین این توزیع یعنی $E(\hat{\theta})$ دقیقاً در محور مرکزی توزیع، قرار می‌گیرد. اگر مقدار واقعی پارامتر، یعنی θ، نیز دقیقاً در همین نقطه واقع شود، می‌گوییم $\hat{\theta}$ یک تخمین‌زنندهٔ ناریب از θ است.

مثالی که در مورد خصوصیت ناریبی یک تخمین‌زننده می‌توان ذکر کرد، این است که بگوییم میانگین نمونه، یعنی $\bar{X}$، یک تخمین ناریب از میانگین جامعه، یعنی μ، است؛ زیرا

$$E(\bar{X}) = E\left[\frac{1}{n}\sum_{i}^{n} X_i\right] = \frac{1}{n}\sum E(X_i) = \mu \quad .$$

باید توجه کرد که ناریب بودن ـ به منزلهٔ تنها معیار مطلوب ـ نمی‌تواند چندان رضایت‌بخش باشد، زیرا هیچ نکته‌ای دربارهٔ پراکندگی توزیع تخمین‌زننده بیان نمی‌کند. یک تخمین‌زننده ناریب که واریانس بسیار بزرگی داشته باشد می‌تواند در بسیاری موارد به نتایجی برسد که ممکن است قابل قبول نباشد؛ زیرا می‌دانیم واریانس بزرگ به معنای دقت کم در تخمین پارامترهاست؛ از طرف دیگر، تخمین‌زننده‌ای که

واریانس بسیار کم دارد، اما مقدار اریب آن مشخص نیست با همین استدلال، نمی‌تواند چندان مفید باشد.

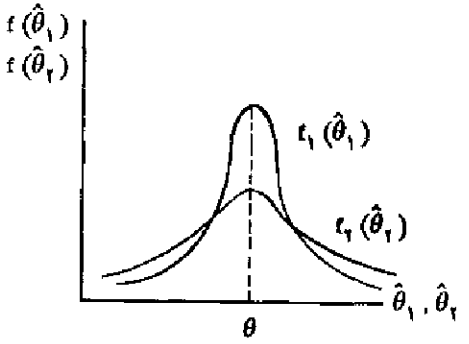
از بحث فوق می‌توان نتیجه گرفت که تخمین‌زننده‌ای مطلوب است که میانگین مربع خطای $\hat{\theta}$ را حداقل کند؛ زیرا میانگین مربع خطا ($\hat{\theta}$) عبارت است از مجموع واریانس و مربع اریب. اما متأسفانه هنوز نمی‌توان تخمین‌زننده‌ای یافت که بتواند حداقل بودن میانگین مربع خطا را تضمین کند؛ زیرا هر تابعی از میانگین مربع خطا شامل مقدار واقعی θ است، در حالی که ما می‌خواهیم خود θ را تخمین بزنیم. بدین ترتیب معیار حداقل نمودن میانگین مربع خطا یک بحث نظری بیش نیست و در عمل ممکن است هیچگاه تحقق نیابد.

کارآیی[1]

مباحثی که تا به حال مطرح کردیم، زمینهٔ مناسبی برای طرح خصوصیت کارآیی تخمین‌زننده‌هاست که در این قسمت مطرح خواهد شد. متأسفانه در آمار هنوز تعریف واحدی برای کارآیی یک تخمین‌زننده ارائه نشده است. بعضی از نویسندگان، اصطلاح کارآیی را در مواردی به کار می‌برند که یک تخمین‌زننده بتواند میانگین مربع خطا را حداقل کند. همان گونه که دیدیم حداقل نمودن میانگین مربع خطا تنها یک مفهوم نظری است و در عمل نمی‌توان $\hat{\theta}$ را چنان تعیین کرد که میانگین مربع خطا حداقل شود؛ بنابراین منظور، یک کارآیی نسبی است؛ یعنی تخمین‌زننده‌ای از کارآیی نسبی بیشتری برخوردار است که در مقایسه با تخمین‌زننده‌های دیگر، واریانس و اریب کمتری داشته باشد. عده‌ای دیگر نیز اصطلاح کارآیی را فقط برای خصوصیات مطلوب حدی تخمین‌زننده‌ها به کار می‌برند و در نمونه‌های محدود و کوچک از این مفهوم استفاده نمی‌کنند. گروهی نیز معتقدند، تخمین‌زننده‌ای کارآست که اولاً نااریب باشد، ثانیاً حداقل واریانس را داشته باشد. به نظر می‌رسد که تعریف آخر بیشتر معمول است

1. Efficiency

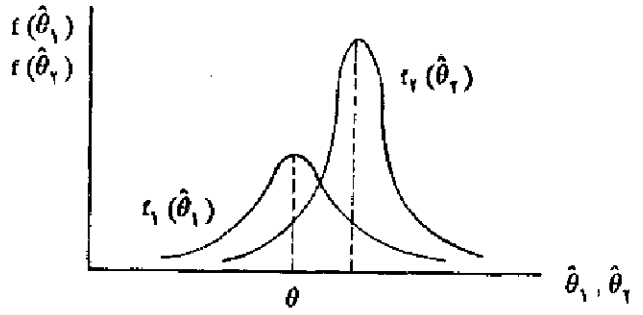
و ما نیز در این کتاب از این تعریف استفاده خواهیم کرد.



نمودار ۲.۱۲ اریب و واریانس

برای تبیین بیشتر مفهوم کارآیی به دو نمودار ۱۲ـ۲ و ۱۳ـ۲ توجه می‌کنیم. در نمودار ۱۲ـ۲ دو تخمین‌زنندهٔ $\hat{\theta}_1$, $\hat{\theta}_2$ با توابع توزیع احتمال $f_1(\hat{\theta}_1)$ و $f_2(\hat{\theta}_2)$ مفروض هستند. هر دو می‌خواهند θ را تخمین بزنند. با اینکه هر دو تخمین‌زننده نااریب هستند، $\hat{\theta}_2$ پراکندگی بیشتری نسبت به $\hat{\theta}_1$ دارد. در این حالت می‌توان گفت که $\hat{\theta}_2$ از کارآیی کمتری نسبت به $\hat{\theta}_1$ برخوردار است. اگر نتوانیم هیچ تخمین‌زنندهٔ نااریب دیگری پیدا کنیم که نسبت به $\hat{\theta}_1$ واریانس کمتری داشته باشد، می‌گوییم $\hat{\theta}_1$ در میان تمام تخمین‌زننده‌های نااریب، حداقل واریانس را دارد؛ بنابراین یک تخمین‌زنندهٔ کارآست.

در نمودار ۱۳ـ۲ ملاحظه می‌کنیم که $\hat{\theta}_1$ نااریب است، اما واریانس بیشتری نسبت به $\hat{\theta}_2$ دارد ـ که خود دارای اریب است. نمی‌توان گفت کدام تخمین‌زننده بر دیگری برتری دارد، مگر اینکه وزنه‌ای ـ که بیان‌کنندهٔ درجهٔ اهمیت است ـ به معیارهای نااریبی و واریانس کمتر منسوب کنیم. بدیهی است تعیین این ضرایب یا وزنه‌ها تابعی از



نمودار ۲.۱۳ اریب و واریانس

مشخصات مسأله‌ای است که در عمل با آن مواجهیم؛ به عبارت دیگر، در اقتصادسنجی،

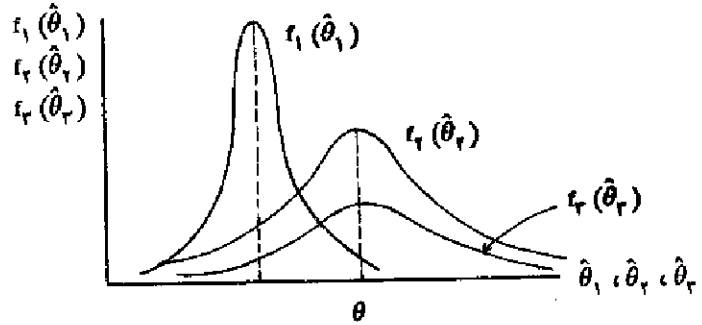
محقق با توجه به خصوصیات مسأله‌ای که در مقابل اوست، تصمیم می‌گیرد که آیا نااریبی نسبت به واریانس کمتر مرجح است یا برعکس. بعد از ذکر این مقدمه، می‌توان تعریف کارآیی را به صورت دقیقتری بیان کرد.

$\hat{\theta}$ یک تخمین‌زنندهٔ کارآ از θ است، اگر شرایط زیر برقرار باشد. اولاً، $\hat{\theta}$ نااریب باشد؛ $E(\hat{\theta}) = \theta$، ثانیاً، واریانس $\hat{\theta}$ از واریانس هر تخمین‌زنندهٔ نااریب دیگری از θ کمتر باشد،

$$\mathrm{Var}(\hat{\theta}) \leq \mathrm{Var}(\tilde{\theta}) \ , \qquad (۲.۶۰)$$

که در آن $\tilde{\theta}$ هر تخمین‌زنندهٔ نااریب دیگری از θ است.

در بعضی موارد به تخمین‌زننده‌های کارآ تخمین‌زننده نااریب حداقل واریانس (MVUE)[1] نیز می‌گویند. اصطلاح دیگری که برای تخمین‌زننده‌های کارآ به کار می‌رود، بهترین تخمین‌زننده نااریب[2] است. باید به این نکته توجه داشت که بنا بر تعریف فوق، یک تخمین‌زننده با حداقل واریانس ـ که کوچکترین مقدار اریب را داشته باشد ـ دیگر یک تخمین‌زنندهٔ کارآ نخواهد بود. این نکته در نمودار ۱۴ـ۲ نشان داده شده است. در این نمودار سه تخمین‌زنندهٔ $\hat{\theta}_1$ ، $\hat{\theta}_2$ و $\hat{\theta}_3$ می‌خواهند θ را تخمین بزنند.



نمودار ۱۴ـ۲ اریب و واریانس

ملاحظه می‌شود که $\hat{\theta}_1$ حداقل واریانس را دارد، اما کارآ نیست، زیرا اریب دارد. از طرف

1. Minimum Variance Unbiased Estimator
2. Best Unbiased Estimator

دیگر، $\hat{\theta}_1$ و $\hat{\theta}_2$ هر دو نااریب هستند اما $\hat{\theta}_2$ واریانس بیشتری نسبت به $\hat{\theta}_1$ دارد، بنابراین $\hat{\theta}_2$ کارآ نیست. در مجموع می‌توان گفت هرگاه هیچ تخمین‌زنندهٔ نااریب دیگری نتوان یافت که واریانس کمتری از $\hat{\theta}_1$ داشته باشد، $\hat{\theta}_1$ یک تخمین‌زنندهٔ کارآست.

تعیین کارآیی در عمل بسیار مشکلتر از این است که مشخص کنیم آیا یک تخمین‌زننده اریب دارد یا خیر؛ زیرا برای یافتن اریب کافی است از $\hat{\theta}$ امید ریاضی بگیریم و اگر برابر θ شد، در آن صورت نااریبی $\hat{\theta}$ را نتیجه می‌گیریم. اما برای کارآیی باید معیار کمترین واریانس در بین همه تخمین‌زننده‌ها را دقیقاً بررسی کنیم. ممکن است دهها تخمین‌زننده وجود داشته باشد؛ بنابراین محاسبهٔ واریانس همه آنها برای یافتن کمترین واریانس، قاعدتاً بسیار مشکل است. به همین دلیل مفید است که از مفهوم کارآیی نسبی[1] نیز استفاده شود. در این صورت می‌گوییم مثلاً $\hat{\theta}_1$ از کارآیی نسبی بیشتری نسبت به $\hat{\theta}_2$ در نمودار ۱۴-۲ برخوردار است[2]. راه دیگر این است که قلمرو تخمین‌زننده‌ها را محدود کنیم تا محاسبهٔ واریانسها در آن قلمرو محدود، راحت‌تر انجام شود. برای ین منظور ابتدا باید خطی بودن را خصوصیت مطلوب تخمین‌زننده‌ها معرفی کنیم.

تخمین‌زننده‌های خطی

یکی دیگر از خصوصیات مطلوب تخمین‌زننده‌ها این است که $\hat{\theta}$ یک تخمین‌زنندهٔ خطی باشد. خطی بودن در اینجا بدین معنی است که $\hat{\theta}$ یک تابع خطی از مشاهدات موجود در نمونه است؛ به عبارت دیگر، برای نمونه‌ای شامل $X_1, X_2, \ldots, X_n$، تخمین‌زنندهٔ $\hat{\theta}$ موقعی خطی است که داشته باشیم

$$\hat{\theta} = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n \quad . \qquad (۲\text{-}۶۱)$$

---

1. Relative Efficiency

۲. البته برای تعیین درجه کارآیی یک تخمین‌زننده، وقتی تابع توزیع احتمال آن کاملاً مشخص باشد، می‌توان از نامساوی کرامر - رائو (Cramer - Rao Inequality) استفاده کرد که در کتابهای آمار به تفصیل بحث شده است.

با اضافه کردن خصوصیت خطی بودن تخمین‌زننده‌ها به خصوصیت کارآیی، می‌توان مفهوم دیگری را در این زمینه مطرح کرد که در ذیل به آن اشاره می‌کنیم.

خصوصیت بهترین تخمین‌زنندهٔ نااریب خطی

اصطلاح BLUE به معنای «بهترین تخمین‌زنندهٔ نااریب خطی»[1] است. در اقتصادسنجی اصطلاح «بهترین» معمولاً مترادف «حداقل واریانس» به کار می‌رود؛ یعنی وقتی می‌گوییم $\hat{\theta}$ بهترین است، بدین معنی نیست که $\hat{\theta}$ از هر نظر بهترین تخمین‌زننده است، بلکه منظور این است که $\hat{\theta}$ حداقل واریانس را دارد.

$\hat{\theta}$ بهترین تخمین‌زنندهٔ خطی (BLUE) از θ است اگر سه شرط زیر را داشته باشد:

۱. $\hat{\theta}$ یک تابع خطی از مشاهدات موجود در نمونه، یعنی $\hat{\theta} = \Sigma\, a_i X_i$ باشد؛

۲. $\hat{\theta}$ نااریب، یعنی $E(\hat{\theta}) = \theta$ باشد؛

۳. واریانس $\hat{\theta}$ کمتر از واریانس هر تخمین‌زنندهٔ نااریب دیگر برای θ، یعنی $\mathrm{Var}(\hat{\theta}) \leq \mathrm{Var}(\tilde{\theta})$ باشد ـ که $\tilde{\theta}$ هر تخمین‌زنندهٔ نااریب دیگری از θ است.

برای مقایسهٔ خصوصیت بهترین تخمین‌زننده خطی و کارآیی، مفید است سه حالت زیر را ملاحظه کنیم.

حالت اول: $\hat{\theta}$ یک تخمین‌زنندهٔ کارآ و نیز یک تابع خطی از مشاهدات X_i است. دراین حالت بهترین تخمین‌زننده‌های خطی و کارآ در واقع یکی هستند. برای مثال، چون می‌توان نشان داد که $\bar{X}$ یک تخمین‌زننده کارآ از میانگین یک توزیع احتمال نرمال و $\bar{X}$ یک تابع خطی از X_i است؛ بنابراین $\bar{X}$ بهترین تخمین‌زنندهٔ خطی از μ نیز خواهد بود.

حالت دوم: $\hat{\theta}$ یک تخمین‌زننده کارآ، اما تقریباً یک تابع خطی از مشاهدات X_i است. در این حالت بهترین تخمین‌زننده‌های خطی دیگر تخمین‌زننده‌های کارآ نیستند. البته در این حالت می‌توان نشان داد که واریانس بهترین تخمین‌زنندهٔ خطی بسیار نزدیک به واریانس تخمین‌زنندهٔ کارآست.

---

1. Best Linear Unbiased Estimator

حالت سوم: $\hat{\theta}$ یک تخمین‌زنندهٔ کارآ، اما تابع غیرخطی از مشاهدات X_i است. در این حالت بهترین تخمین‌زنندهٔ خطی دیگر کارآ نیست، اما می‌توان نشان داد که واریانس آن به مراتب بیشتر از واریانس تخمین‌زنندهٔ کارآ خواهد بود.

آخرین نکته این است که خصوصیت خطی بودن در واقع یک خصوصیت بسیار مطلوب در تخمین پارامترهایی چون میانگین است ـ که در واقع گشتاورهای مرتبهٔ اول است. برای تخمین پارامترهایی که صورت گشتاورهای مرتبهٔ دوم را دارد؛ مانند واریانس، خطی بودن تخمین‌زننده، یک خصوصیت بسیار نامطلوب است. در تخمین این گونه پارامترها، خصوصیت مطلوب در واقع خصوصیت بهترین تخمین‌زننده نااریب درجه دو[1] خواهد بود، که آن را با BQUE نشان می‌دهیم.

جامعیت

آخرین خصوصیت مطلوب تخمین‌زننده‌ها در نمونه‌های محدود، شرط «جامعیت»[2] است. بنا بر تعریف، تخمین‌زننده‌ای جامع است که بتواند در تخمین پارامتر جامعه، از کلیهٔ اطلاعات موجود در نمونه استفاده کند. با توجه به اینکه هر مشاهده‌ای در مجموعهٔ مشاهدات موجود در نمونه باید مشخص کننده ساختار جامعه باشد، در محاسبهٔ $\hat{\theta}$ انتظار این است که تمام اطلاعات موجود در نمونه مشارکت داشته باشند. برای مثال، میانه نمی‌تواند تخمینی جامع از میانگین جامعه μباشد؛ زیرا برای محاسبه میانه به طور کلی از اولویت و طبقه‌بندی مشاهدات موجود در نمونه استفاده می‌شود و مقادیر تمام مشاهدات نقش چندانی ایفا نمی‌کند.

البته ممکن است استدلال شود که معیار استفاده از تمام اطلاعات موجود در نمونه نباید امتیاز برجسته‌ای در خصوصیات مطلوب یک تخمین‌زننده به حساب آید. در اقتصادسنجی مهم این است که از پارامترهای موجود در جامعه تخمینهای خوب و رضایت‌بخشی به دست آوریم، بنابراین مهم نیست که آیا از تمام اطلاعات موجود در

---

1. Best Quadratic Unbiased Estimator
2. Sufficiency

نمونه استفاده کرده‌ایم یا خیر. با وجود این، به نظر می‌رسد، این استدلال چندان معتبر نباشد؛ زیرا می‌توان ثابت کرد که معیار جامعیت شرط لازم برای کارآیی است[۱]. به عبارت دیگر، یک تخمین‌زنندهٔ کارآ نیست، مگر اینکه از تمام اطلاعات موجود در نمونه استفاده کرده باشد. همچنین ثابت شده است که مجموع خصوصیات ناارییی و جامعیت، دقیقاً موجب کارآیی می‌شود ـ که در آمار به نام قضیهٔ بلک ول ـ راثو[۲] معروف است.

در نتیجه سه شرط ناارییی، کارآیی و بهترین تخمین‌زنندهٔ نااریب خطی، در واقع، مهمترین خصوصیات مطلوب تخمین‌زننده‌ها هستند. این خصوصیات مطلوب ـ همان گونه که دیدیم ـ برحسب میانگین و واریانس تخمین‌زننده‌ها تعریف می‌شوند، بنابراین برای استنتاج در مورد هر یک از آنها باید میانگین و واریانس توزیع احتمال تخمین‌زننده‌ها را دقیقاً شناخت.

۲. خصوصیات حدی تخمین‌زننده‌ها

قبلاً گفتیم که خصوصیات حدی[۳] به نمونه‌هایی مربوط می‌شود که حجم آنها بسیار بزرگ است و به بی‌نهایت میل می‌کند. ابتدا مفهوم توزیع حدی[۴] را تعریف می‌کنیم. در بسیاری موارد، شکل توزیع یک متغیر تصادفی، تابعی از حجم نمونه است؛ برای مثال، بنا بر قضیهٔ حد مرکزی، می‌دانیم که به ازای افزایش حجم نمونه، توزیع میانگین نمونه، یعنی $\bar{X}$، به سمت یک توزیع نرمال میل می‌کند. بنابراین، می‌گوییم توزیع نرمال، یک توزیع حدی برای میانگین نمونه است. به طور کلی، می‌توان گفت هر گاه توزیع یک تخمین‌زننده به ازای افزایش حجم نمونه به سمت خاصی میل کند، آن شکل خاص، توزیع حدی تخمین‌زننده نامیده می‌شود.

واژهٔ حدی در اینجا مطلقاً بدین معنی نیست که توزیع حدی برای یک

---

۱. برای توضیحات بیشتر در این زمینه به کتاب B. W. Lindgren(۱۹۷۶) صفحات ۲٦٤ به بعد مراجعه شود.

2. Blackwell - Rao Theorem

3. Asymptotic Properties

4. Asymptotic Distribution

تخمین‌زننده، در واقع صورت نهایی توزیع $\hat{\theta}$ است. وقتی n به سمت بی‌نهایت میل کند، آنچه در عمل اتفاق می‌افتد، این است که به ازای افزایش حجم نمونه، توزیع احتمال تخمین‌زننده در یک نقطهٔ معینی به مرحلهٔ فروپاشی[1] می‌رسد، و چه بسا ممکن است این نقطه دقیقاً منطبق با مقدار واقعی پارامتر شود. برای مثال، دیدیم که میانگین توزیع $\bar{X}$ برابر μ و واریانس آن $\frac{\sigma^2}{n}$ است که σ^2 واریانس جامعه است. بدیهی است وقتی n به سمت بی‌نهایت میل کند، $\frac{\sigma^2}{n}$ به سمت صفر میل می‌کند و توزیع $\bar{X}$ در نقطهٔ میانگین جامعه (μ) فرو می‌پاشد. از نظر هندسی، توزیع در این نقطه با یک خط عمودی مشخص می‌شود که قاعدتاً منحنی نرمال نخواهد بود؛ بنابراین ملاحظه می‌شود که منظور از توزیع حدی یک تخمین‌زننده، شکل نهایی توزیع آن نیست، بلکه در واقع صورتی از توزیع است که تخمین‌زننده قبل از مرحلهٔ فروپاشی به آن رسیده است. در مثال توزیع میانگین نمونه $\bar{X}$، دیدیم که به ازای افزایش حجم نمونه، واریانس این توزیع به طور مرتب کمتر می‌شود و در عین حال شکل توزیع نیز بیشتر به نرمال میل می‌کند. قبل از مرحلهٔ فروپاشی، توزیع به صورت نرمال و با واریانس بسیار کوچکی خواهد بود.

حال که با مفهوم توزیع حدی آشنا شدیم، این سؤال مطرح می‌شود که چگونه می‌توان توزیع حدی تخمین‌زننده‌ها را به دست آورد. برای بسیاری از تخمین‌زننده‌ها پاسخ به این سؤال بسیار ساده است؛ زیرا شکل توزیع آنها تابعی از حجم نمونه نیست. توزیع میانگین نمونه برای یک جامعهٔ نرمال، مثال خوبی در این مورد است؛ زیرا این توزیع، به رغم حجم نمونه، همواره نرمال و میانگین آن μ و واریانس آن $\frac{\sigma^2}{n}$ است. در بسیاری موارد، شکل توزیع تخمین‌زننده در نمونه‌های کوچک چندان روشن نیست؛ زیرا معمولاً شکل توزیع به موازات افزایش حجم نمونه و میل آن به سمت بی‌نهایت مشخص می‌شود. برای مثال، می‌توان گفت توزیع میانگین نمونه، به عنوان تخمینی از میانگین جامعه، دارای چنین خصوصیتی برای جامعه‌های غیر نرمال است؛ زیرا در حقیقت بر اساس قضیهٔ حد مرکزی می‌دانیم که به ازای افزایش حجم نمونه و میل آن

1. Degenerate

به سمت بی‌نهایت، $\bar{X}$ توزیع نرمال خواهد داشت.

توزیع حدی نیز با میانگین و واریانس آن مشخص می‌شود. میانگین این توزیعها را در اصطلاح «میانگین حدی»[1] و واریانس آن را «واریانس حدی»[2] می‌گویند. برای یافتن میانگین حدی، کافی است از امید ریاضی آن حد بگیریم. برای مثال، می‌دانیم میانگین $\hat{\theta}$ برابر با $E(\hat{\theta})$ است، اما میانگین حدی آن برابر با

$$\text{میانگین حدی } \hat{\theta} = \lim_{n\to\infty} E(\hat{\theta}) \quad . \qquad (۲\text{-}۶۲)$$

خواهد بود. به همین ترتیب واریانس حدی $\hat{\theta}$ را تعریف می‌کنیم،

$$\lim_{n\to\infty} \text{Var}(\hat{\theta}) = \lim_{n\to\infty} E[\hat{\theta} - E(\hat{\theta})]^2 \quad . \qquad (۲\text{-}۶۳)$$

در ذیل به سه خصوصیت حدی برای تخمین‌زننده‌ها اشاره می‌شود.

نااریب حدی

یک تخمین‌زننده، موقعی نااریب حدی[3] است که به ازای افزایش حجم نمونه، به طور مرتب از مقدار اریب آن کم شود و وقتی حجم نمونه به سمت بی‌نهایت میل می‌کند، مقدار اریب صفر شود،

$$\lim_{n\to\infty} E(\hat{\theta}) = 0 \quad . \qquad (۲\text{-}۶۴)$$

به عبارت دیگر، یک تخمین‌زننده، موقعی نااریب حدی است که در حد نااریب شود. باید در نظر داشت که اگر یک تخمین‌زننده نااریب باشد آنگاه نااریب حدی نیز هست، اما نه برعکس. دلیل این امر روشن است: وقتی تخمین‌زننده‌ای نااریب باشد، امید ریاضی آن به ازای هر حجمی از نمونه برابر مقدار واقعی پارامتر است؛ به بیان دیگر، اگر برای تخمین پارامتر θ، نمونه‌ای با حجم n داشته باشیم و به کمک این نمونه مقدار $\hat{\theta}$ را به دست آوریم، آنگاه با اینکه ممکن است برای مقادیر کوچک n مقدار $E(\hat{\theta})$ برابر θ

1. Asymptotic Mean
2. Asymptotic Variance
3. Asymptotically Unbiased

نشود، اما اگر برای مقادیر بزرگ و بزرگتر n، مقدار $E(\hat{\theta})$ به سمت θ میل کند، به گونه‌ای که در حد مقدار $E(\hat{\theta})$ دقیقاً برابر θ شود، در آن صورت می‌گوییم $\hat{\theta}$ یک تخمین نااریب حدی از θ است.

مثال ۲ـ۱۴ نشان دهید که واریانس نمونهٔ s^2، یک تخمین نااریب حدی از واریانس جامعهٔ σ^2 است. اگر واریانس نمونه را به صورت زیر تعریف کنیم،

$$s^2 = \frac{\Sigma (X_i - \bar{X})^2}{n} \quad , \qquad (۲ـ۶۵)$$

با اضافه کردن $\mu \pm$ به داخل پرانتز صورت کسر خواهیم داشت

$$s^2 = \frac{1}{n} \left\{ \sum [(X_i - \mu) - (\bar{X} - \mu)]^2 \right\} \quad ,$$

$$= \frac{1}{n} [\sum (X_i - \mu)^2 + n (\bar{X} - \mu)^2 - 2 (\bar{X} - \mu) \sum (X_i - \mu)] \quad .$$

اما می‌دانیم $\sum (X_i - \mu) = n (\bar{X} - \mu)$ بنابراین

$$s^2 = \frac{1}{n} [\sum (X_i - \mu)^2 - n (\bar{X} - \mu)] \quad ,$$

$$= \frac{1}{n} \sum (X_i - \mu)^2 - (\bar{X} - \mu)^2 \quad .$$

حال اگر امید ریاضی بگیریم، با توجه به $E(\bar{X}) = \mu$ خواهیم داشت

$$E(s^2) = \frac{1}{n} \sum E (X_i - \mu)^2 - E (\bar{X} - \mu)^2 \quad ,$$

$$= \frac{1}{n} \sum E [X_i - E (X_i)]^2 - E [\bar{X} - E (\bar{X})]^2 \quad ,$$

$$= \frac{1}{n} \sum \text{Var} (X_i) - \text{Var} (\bar{X}) \quad .$$

با توجه به اینکه $\text{Var}(X_i) = \sigma^2$ و $\text{Var}(\bar{X}) = \frac{1}{n}\sigma^2$ و $\Sigma \text{Var}(X_i) = n\sigma^2$، در نتیجه

$$E(s^2) = \sigma^2 - \frac{1}{n}\sigma^2 \quad ,$$

یا

$$E(s^2) = \left(1 - \frac{1}{n}\right)\sigma^2 \quad . \qquad (۲\text{-}۶۶)$$

ملاحظه می‌شود که $E(s^2) \neq \sigma^2$ و s^2 تخمین نااریبی از σ^2 است. اما نکتهٔ مهم این است که اگر n به سمت بی‌نهایت میل کند؛ یعنی حجم نمونه به طور مرتب بزرگ و بزرگتر شود، آنگاه در حد $E(s^2)$ برابر σ^2 خواهد شد؛ زیرا با افزایش n مقدار $\frac{1}{n}$ به سمت صفر میل می‌کند،

$$\lim_{n\to\infty} E(s^2) = \lim_{n\to\infty}\left(1 - \frac{1}{n}\right)\sigma^2 = \sigma^2 \quad , \qquad (۲\text{-}۶۷)$$

یعنی s^2 یک تخمین‌زنندهٔ نااریب حدی از σ^2 است. به همین دلیل است که s^2 را در مباحث آماری به صورت زیر تعریف می‌کنند که تخمین نااریبی از σ^2 باشد،

$$s^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1} \quad . \qquad (۲\text{-}۶۸)$$

زیرا از معادلهٔ ۲ـ۶۶ داریم

$$E(s^2) = \frac{n-1}{n}\sigma^2 \quad ,$$

یا

$$\frac{n}{n-1}E(s^2) = \sigma^2 \quad ,$$

$$E\left(\frac{n s^2}{n-1}\right) = \sigma^2 \quad . \qquad (۲\text{-}۶۹)$$

با توجه به معادلهٔ ۲ـ۶۵ می‌دانیم

$$n s^2 = \sum (X_i - \bar{X}) \quad ,$$

که با جایگزینی در ۲ـ۶۹ خواهیم داشت

$$E\left[\frac{\sum (X_i - \bar{X})^2}{n-1}\right] = \sigma^2 \quad .$$

بنابراین اگر واریانس نمونه را به صورت معادلهٔ ۲ـ٦٨ تعریف کنیم، آنگاه $E(s^2) = \sigma^2$ است، در نتیجه واریانس نمونه تخمین نااریبی از واریانس جامعه خواهد بود.

سازگاری

در مقدمه بحث خصوصیات حدی دیدیم که به ازای افزایش حجم نمونه و میل آن به سمت بی‌نهایت، معمولاً شکل توزیع تخمین‌زننده به مرحله فروپاشی می‌رسد. تخمین‌زننده $\hat{\theta}$ را در نظر می‌گیریم. نقطه‌ای را که در آن توزیع $\hat{\theta}$ فرو می‌پاشد، نقطهٔ احتمال حدی[1] $\hat{\theta}$ می‌گوییم و آن را با علامت $\text{Plim}\ \hat{\theta}$ نشان می‌دهیم. فرض کنید θ^* نقطهٔ احتمال حدی $\hat{\theta}$ است، البته نمی‌دانیم که آیا θ^* با مقدار واقعی پارامتر θ برابر است یا خیر. معادلهٔ

$$\text{Plim}\ \hat{\theta} = \theta^* \ ,$$

در واقع بدین معنی است که احتمال برابری $\hat{\theta}$ با θ^* در حد برابر یک است؛ یعنی توزیع در نقطهٔ θ^* فرو می‌پاشد؛ بنابراین می‌توان نوشت

$$\text{Lim Pr}\ (\theta^* - \varepsilon \leq \hat{\theta} \leq \theta^* + \varepsilon) = ۱ \ ,$$

که در آن ε عدد مثبت بسیار کوچکی است.

در اینجا می‌گوییم تخمین‌زننده‌ای سازگار است که توزیع آن در نقطه‌ای فرو می‌ریزد که دقیقاً برابر مقدار واقعی پارامتر جامعه است؛ به عبارت دیگر $\hat{\theta}$ یک تخمین‌زنندهٔ سازگار از θ است هرگاه

$$\text{Plim}\ \hat{\theta} = \theta \ . \qquad (۲ـ۷۰)$$

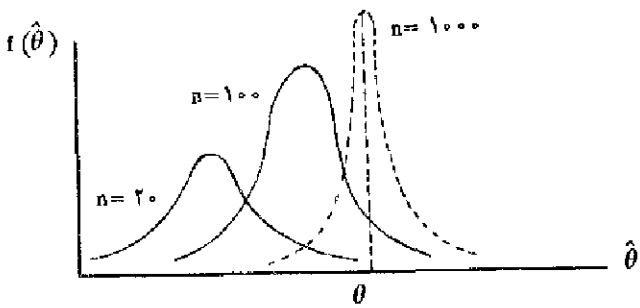
برای اینکه بدانیم یک تخمین‌زننده سازگار است یا خیر، باید به جهت حرکت مقدار اریب و واریانس آن به ازای افزایش حجم نمونه و میل آن به سمت بی‌نهایت توجه کنیم. اگر به موازات افزایش حجم نمونه، مقدار اریب و مقدار واریانس هر دو کاهش

1. Probability Limit

پیدا کند و این سیر تا مرحله‌ای ادامه یابد که به ازای $n \to \infty$، مقادیر اریب و واریانس به سمت صفر میل کنند، تخمین‌زننده سازگار خواهد بود. بنابراین احتمال حدی $\hat{\theta} = \theta$ است، هرگاه دو شرط زیر صادق باشد.

$$\lim_{n\to\infty} E(\hat{\theta}) = \theta \ ,$$
$$\lim_{n\to\infty} \text{Var}(\hat{\theta}) = 0 \ . \qquad (۲\text{-}۷۱)$$

نمودار ۱۵-۲ یک تخمین‌زنندهٔ سازگار را نشان می‌دهد.



نمودار ۱۵-۲ سازگاری تخمین‌زننده

بنا بر معادلهٔ ۵۷-۲ میانگین مربع خطای $\hat{\theta}$ برابر با واریانس به علاوهٔ مربع اریب است، بنابراین وقتی در سازگاری، اریب و واریانس هر دو به سمت صفر میل کنند، $\text{MSE}(\hat{\theta})$ نیز به سمت صفر میل خواهد کرد؛ در نتیجه می‌توان سازگاری تخمین‌زننده را برحسب میانگین مربع خطا به صورت زیر تعریف کرد.

$\hat{\theta}$ یک تخمین‌زننده سازگار از θ است، هرگاه در حد میانگین مربع خطا برابر صفر شود،

$$\text{Plim}_{n\to\infty} \hat{\theta} = \theta \quad \text{اگر} \quad \lim_{n\to\infty} \text{MSE}(\hat{\theta}) = 0 \ . \qquad (۲\text{-}۷۲)$$

می‌توان ثابت کرد که شرط فوق درواقع شرط لازم است ولی کافی نیست؛ به عبارت دیگر می‌توان یک تخمین‌زنندهٔ سازگار یافت که میانگین مربع خطای آن به‌ازای $n \to \infty$ به سمت صفر میل نکند (در مورد این موضوع در کتابهای آمار پیشرفته بحث بیشتری شده است).

خصوصیت بسیار مهم تخمین‌زننده‌های سازگار این است که هر تابع پیوسته‌ای از یک تخمین‌زنندهٔ سازگار، خود سازگار است. این قضیه را در ذیل توضیح می‌دهیم.

قضیهٔ اسلاتسکی

بنا بر قضیه اسلاتسکی[1]، اگر احتمال حدی $\theta = \hat{\theta}$ و $g(\hat{\theta})$ یک تابع پیوسته از $\hat{\theta}$ باشد، آنگاه

$$\text{Plim } g(\hat{\theta}) = g(\theta) \ . \qquad (۲\text{-}۷۳)$$

در اینجا تنها یادآوری می‌کنیم که این قضیه، کاربردهای بسیاری در روشهای تخمین دارد، که در فصلهای آینده به آن اشاره خواهد شد. بر طبق این قضیه می‌توان گفت که اگر $\hat{\theta}$ یک تخمین‌زنندهٔ سازگار از θ باشد، $\frac{۱}{\hat{\theta}}$ نیز تخمین‌زنندهٔ سازگاری از $\frac{۱}{\theta}$ و $\log \hat{\theta}$ نیز تخمین‌زنندهٔ سازگاری از $\log \theta$ خواهد بود. این قاعده، در حالت کلی برای خصوصیت ناااریبی صادق نیست؛ یعنی اگر $\hat{\theta}$ یک تخمین‌زنندهٔ ناااریب از θ باشد، بدین‌عنی نخواهد بود که $\frac{۱}{\hat{\theta}}$ یا $\log \hat{\theta}$ نیز تخمینهای ناااریبی از $\frac{۱}{\theta}$ یا $\log \theta$ هستند.

یکی از امتیازهای احتمال حدی سادگی عملیات آن است. بدون ورود به اثبات قضایا، بعضی از خصوصیات عملیاتی احتمال حدی را مطرح می‌کنیم. احتمال حدی یک عدد ثابت، عدد ثابت است، یعنی $\text{Plim } a = a$. همچنین داریم

$$\text{Plim } g(\hat{\theta})^۲ = (\text{Plim } \hat{\theta})^۲ \ ,$$

$$\text{Plim } g(\hat{\theta})^{-۱} = (\text{Plim } \hat{\theta})^{-۱}$$

اگر $\hat{\theta}_۱$ و $\hat{\theta}_۲$ به ترتیب تخمینهای سازگاری از $\theta_۱$ و $\theta_۲$ باشند، $f_۱(\hat{\theta}_۱)\ f_۲(\hat{\theta}_۲)$ نیز تخمین سازگاری از $f_۱(\theta_۱)\ f_۲(\theta_۲)$ خواهد بود که در آن $f_۱$ و $f_۲$ هر تابع پیوسته‌ای می‌تواند فرض شود. بنابراین برای مثال داریم

1. Slutsky Theorem

$$\text{Plim}\,(\hat{\theta}_1^{\,2}\log\hat{\theta}_2) = \theta_1^{\,2}\log\theta_2 \ ,$$

$$\text{Plim}\,(\hat{\theta}_1\,\hat{\theta}_2) = \text{Plim}\,\hat{\theta}_1\ \text{Plim}\,\hat{\theta}_2 = \theta_1\,\theta_2 \ ,$$

$$\text{Plim}\,\frac{\hat{\theta}_1}{\hat{\theta}_2} = \frac{\text{Plim}\,\hat{\theta}_1}{\text{Plim}\,\hat{\theta}_2} = \frac{\theta_1}{\theta_2} \ .$$

همچنین داریم

$$\text{Plim}\,(\hat{\theta}_1 \pm \hat{\theta}_2) = \text{Plim}\,\hat{\theta}_1 \pm \text{Plim}\,\hat{\theta}_2 = \theta_1 \pm \theta_2 \ .$$

امتیاز عملیاتی احتمال حدی بر امید ریاضی به وضوح روشن است. برای مثال، می‌دانیم $\mathrm{E}\,(\frac{\hat{\theta}_1}{\hat{\theta}_2}) \neq \frac{\mathrm{E}\,(\hat{\theta}_1)}{\mathrm{E}\,(\hat{\theta}_2)}$ یا $\mathrm{E}\,(\hat{\theta}_1\,\hat{\theta}_2)$ فقط موقعی برابر $\mathrm{E}\,(\hat{\theta}_1)\,\mathrm{E}\,(\hat{\theta}_2)$ است که $\hat{\theta}_1$ و $\hat{\theta}_2$ از یکدیگر مستقل باشد، در حالی که این شرط برای احتمال حدی ضروری نیست.

کارآیی حدی

آخرین خصوصیت حدی که در اینجا مطرح می‌کنیم، خصوصیت کارآیی حدی[1] است که به پراکندگی توزیع حدی یک تخمین‌زننده مربوط می‌شود. این خصوصیت معمولاً برای تخمین‌زننده‌هایی تعریف می‌شود که میانگین حدی و واریانس حدی محدود و معینی دارند. کارآیی حدی در واقع معیار خوبی است که می‌توان با آن از بین گروهی از تخمین‌زننده‌های نااریب حدی، مناسبترین را انتخاب کرد. می‌دانیم توزیع تخمین‌زننده‌های سازگار به ازای $n\to\infty$ در نقطه‌ای فرو می‌پاشد که برابر مقدار واقعی پارامتر جامعه است؛ بنابراین اگر تعداد بسیاری از تخمین‌زننده‌ها این خصوصیت را داشته باشند آنگاه تخمین‌زننده‌ای بر بقیه برتری دارد که سریعتر به این نقطه فروپاشی می‌رسد. واضح است که توزیع حدی این تخمین‌زننده کمترین واریانس را خواهد داشت؛ زیرا می‌دانیم، توزیع حدی در واقع صورتی از توزیع است که $\hat{\theta}$ قبل از نیل به

1. Asymptotic Efficiency

مرحلهٔ فروپاشی به آن رسیده است؛ و تخمین‌زننده‌ای که حداقل واریانس را دارد به طور طبیعی زودتر از تخمین‌زننده‌های سازگار دیگر به این مرحله خواهد رسید. با توجه به تعاریف فوق، می‌توان تعریف دقیقتری از کارآیی حدی ارائه کرد.

$\hat{\theta}$ یک تخمین‌زننده کارای حدی از θ است اگر تمام شرایط زیر صادق باشد.

۱. $\hat{\theta}$ یک توزیع حدی با میانگین و واریانس حدی محدود و معین باشد؛

۲. $\hat{\theta}$ سازگار باشد؛

۳. هیچ تخمین‌زنندهٔ سازگار دیگری از θ را نتوان یافت که واریانس حدی کوچکتری از $\hat{\theta}$ داشته باشد.

مثال ۱۵ـ۲ فرض کنید متغیر تصادفی X_i دارای توزیع نرمال با میانگین μ و واریانس σ^2 است. می‌خواهیم μ را از یک نمونهٔ تصادفی با مشاهدات $X_1, X_2, \ldots, X_n$ تخمین بزنیم. سه تخمین‌زننده به شرح زیر پیشنهاد شده است،

$$\bar{X} = \frac{1}{n}\sum_{1}^{n} X_i \ ,$$

$$\hat{\mu} = \frac{1}{n+1}\sum_{1}^{n} X_i \ ,$$

$$\tilde{\mu} = \frac{1}{2}X_1 + \frac{1}{2n}\sum_{2}^{n} X_i \ .$$

خصوصیات مطلوب تخمین‌زننده‌ها را در مورد هر یک از این سه تخمین‌زننده بررسی کنید.

ابتدا خصوصیات ناااریبی، کارآیی و بهترین تخمین‌زننده ناااریب خطی را در حالت نمونه‌های محدود مطالعه می‌کنیم، آنگاه فرض می‌شود که n به سمت بی‌نهایت میل می‌کند. در این حالت خصوصیات ناااریبی حدی، سازگاری و سرانجام کارآیی حدی در مورد هر یک از سه تخمین‌زننده فوق به طور خلاصه بررسی خواهد شد.

ناااریبی

از $\bar{X}$ ، $\hat{\mu}$ و $\tilde{\mu}$ به ترتیب امید ریاضی می‌گیریم،

$$E(\bar{X}) = E\left[\frac{1}{n}\sum_{1}^{n} X_i\right] = \frac{1}{n}\sum E(X_i) \ ,$$

$$= \frac{1}{n}\sum \mu = \frac{1}{n} \cdot n\mu = \mu \ ,$$

یعنی $\bar{X}$ یک تخمین‌زنندهٔ نااریب از μ است.

$$E(\bar{\mu}) = E\left[\frac{1}{n+1}\sum_{1}^{n} X_i\right] = \frac{1}{n+1}\sum E(X_i) \ ,$$

$$= \frac{n}{n+1}\ \mu \ ,$$

یعنی $\hat{\mu}$ تخمینی از μ و دارای اریب است.

$$E(\tilde{\mu}) = E\left[\frac{1}{2} X_1 + \frac{1}{2n}\sum_{2}^{n} X_i\right] \ ,$$

$$= \frac{1}{2} E(X_1) + \frac{1}{2n}\sum_{2}^{n} E(X_i) \ ,$$

$$= \frac{1}{2}\mu + \left(\frac{n-1}{2n}\right)\mu = \frac{2n-1}{2n}\mu \ ,$$

یعنی $\tilde{\mu}$ اریب دارد.

کارآیی

می‌دانیم کارایی فقط برای تخمین‌زننده‌های نااریب تعریف می‌شود. با توجه به اینکه $\hat{\mu}$ و $\tilde{\mu}$ هر دو اریب دارند فقط $\bar{X}$ می‌تواند به عنوان مورد مناسبی برای کارآیی مطرح شود. واریانس $\bar{X}$ برابر با $\frac{\sigma^2}{n}$ است. همچنین می‌دانیم که نمونهٔ مفروض از یک جامعهٔ نرمال گرفته شده است. بدون اثبات می‌گوییم که با استفاده از نامساوی کرامر ـ راثو می‌توان نشان داد که $\bar{X}$ یک تخمین‌زننده کارآ از μ است.

بهترین تخمین‌زنندهٔ نااریب خطی

ملاحظه می‌شود که فقط $\bar{X}$ می‌تواند مورد مناسب این خصوصیت باشد؛ چون دو تخمین‌زنندهٔ دیگر اریب دارند. $\bar{X}$ خصوصیت خطی بودن را نیز دارا است؛ زیرا

$$\bar{X} = \frac{1}{n}\sum_{1}^{n} X_i = \frac{1}{n} X_1 + \frac{1}{n} X_2 + \cdots + \frac{1}{n} X_n \ .$$

در قسمت قبل دیدیم که $\bar{X}$ کارآست، یعنی حداقل واریانس را در میان تمام تخمین‌زننده‌های ناااریب دیگر از جمله خطی دارد. بنابراین $\bar{X}$ بهترین تخمین‌زنندهٔ ناااریب خطی است. در اینجا به بررسی خصوصیات حدی در مثال مذکور می‌پردازیم.

ناااریب حدی

کافی است از امید ریاضی تخمین‌زننده‌ها حد بگیریم.

$$\lim_{n\to\infty} E(\bar{X}) = \lim_{n\to\infty} \mu = \mu \ ,$$

یعنی $\bar{X}$ ناااریب حدی است.

$$\lim_{n\to\infty} E(\hat{\mu}) = \lim_{n\to\infty} \frac{n}{n+1}\ \mu = \mu \ ,$$

یعنی $\hat{\mu}$ نیز ناااریب حدی است.

$$\lim_{n\to\infty} E(\tilde{\mu}) = \lim_{n\to\infty} \left(\frac{2n-1}{2n}\right)\mu = \mu \ ,$$

و بدین ترتیب $\tilde{\mu}$ نیز ناااریب حدی خواهد بود.

سازگاری

چون $\bar{X}$ ناااریب است، پس

$$\text{MSE}(\bar{X}) = \text{Var}(\bar{X}) = \frac{\sigma^2}{n} \ .$$

اگر از میانگین مربع خطای $(\bar{X})$ حد بگیریم، داریم

$$\lim_{n\to\infty} \text{MSE}(\bar{X}) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0 \ ,$$

یعنی $\bar{X}$ سازگار است.

برای بررسی سازگاری $\hat{\mu}$ ابتدا $\text{MSE}(\hat{\mu})$ را به دست می‌آوریم. با توجه به معادلهٔ ۲ـ۵۷

داریم

$$MSE(\hat{\mu}) = Var(\hat{\mu}) + (\hat{\mu} \text{ اریب})^2 .$$

بنابراین

$$MSE(\hat{\mu}) = Var\left[\frac{1}{n+1}\sum_{1}^{n} X_i\right] + \left[\left(\frac{n}{n+1}\right)\mu - \mu\right]^2 ,$$

$$= \left(\frac{1}{n+1}\right)^2 \sum_{1}^{n} Var(X_i) + \left(\frac{-1}{n+1}\right)^2 \mu^2 ,$$

$$= \frac{n\sigma^2}{(n+1)^2} + \frac{1}{(n+1)^2}\mu^2 = \frac{n\sigma^2 + \mu^2}{(n+1)^2} .$$

از $MSE(\hat{\mu})$ حد گرفته، داریم

$$\lim_{n\to\infty} MSE(\hat{\mu}) = \lim_{n\to\infty} \frac{n\sigma^2 + \mu^2}{(n+1)^2} = 0 ,$$

یعنی $\hat{\mu}$ سازگار است. به همین ترتیب برای $\tilde{\mu}$ عمل می‌کنیم،

$$MSE(\tilde{\mu}) = Var(\tilde{\mu}) + (\tilde{\mu} \text{ اریب})^2 .$$

بنابراین

$$MSE(\tilde{\mu}) = Var\left[\frac{1}{2}X_1 + \frac{1}{2n}\sum_{2}^{n} X_i\right] + \left[\left(\frac{2n-1}{2n}\right)\mu - \mu\right]^2 ,$$

$$= \frac{1}{4} Var(X_1) + \left(\frac{1}{2n}\right)^2 \sum_{2}^{n} Var(X_i) + \left(\frac{-1}{2n}\right)^2 \mu^2 ,$$

$$= \frac{n^2\sigma^2 + (n-1)\sigma^2 + \mu^2}{4n^2} = \frac{(n^2+n-1)\sigma^2 + \mu^2}{4n^2} .$$

از $MSE(\tilde{\mu})$ حد می‌گیریم، داریم

$$\lim_{n\to\infty} MSE(\tilde{\mu}) = \frac{\sigma^2}{4} \neq 0 .$$

چون $\text{Lim MSE}\ (\tilde{\mu})$ برابر صفر نیست، پس $\tilde{\mu}$ یک تخمین‌زنندهٔ سازگار نخواهد بود.

کارآیی حدی

فقط $\bar{X}$ و $\hat{\mu}$ موردهای مناسبی برای کارآیی حدی هستند؛ زیرا شرط سازگاری را دارند. چون $\bar{X}$ برای هر حجمی از نمونه کارآست؛ برای حالتی که حجم نمونه به سمت بی‌نهایت میل کند نیز کارآ خواهد بود؛ بنابراین $\bar{X}$ کارآیی حدی دارد.

در مورد $\hat{\mu}$، می‌دانیم

$$\text{Var}\ (\hat{\mu}) = \frac{1}{(n+1)^2} \sum_{1}^{n} \text{Var}\ (X_i)\ ,$$

$$= \frac{1}{(n+1)^2} \cdot n\sigma^2 = \frac{n^2}{(n+1)^2} \cdot \frac{\sigma^2}{n}\ .$$

از $\text{Var}\ (\hat{\mu})$ حد می‌گیریم،

$$\lim_{n\to\infty} \text{Var}\ (\hat{\mu}) = \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^2 \cdot \frac{\sigma^2}{n}\ ,$$

$$= \text{Lim}\ \left(\frac{n}{n+1}\right)^2 \text{Var}\ (\bar{X})\ ,$$

چون به ازای $n\to\infty$، مقدار $\frac{n}{n+1}$ به سمت یک میل می‌کند

$$\lim_{n\to\infty} \text{Var}\ (\hat{\mu}) = \text{Var}\ \bar{X}\ .$$

با توجه به اینکه $\bar{X}$ کارآیی حدی دارد و واریانس $\hat{\mu}$ در حد برابر واریانس $\bar{X}$ شده است نتیجه می‌گیریم که $\hat{\mu}$ نیز کارآیی حدی دارد.

نتایج ششگانه فوق را می‌توان در جدول ۳ـ۲ خلاصه کرد.

جدول ۲.۳ خصوصیات مطلوب برای سه تخمین‌زننده

	خصوصیات مطلوب	$\bar{X}$	$\hat{\mu}$	$\tilde{\mu}$
نمونه‌های محدود	ناااریبی	+	-	-
	کارآیی	+	-	-
	BLUE	+	-	-
نمونه‌های نامحدود	ناااریبی‌حدی	+	+	+
	سازگاری	+	+	-
	کارآیی حدی	+	+	-

نتیجهٔ کلی این که $\bar{X}$ مناسبترین تخمین‌زننده برای نمونه‌های محدود است، اما برای نمونه‌های بزرگ، $\bar{X}$ و $\hat{\mu}$ به یک اندازه مناسبند. تخمین‌زنندهٔ سوم، یعنی $\tilde{\mu}$، هیچ خصوصیت مطلوبی بجز ناااریبی حدی ندارد.

۲.۶ خصوصیات مطلوب تخمین‌زننده‌های حداقل مربعات معمولی: قضیهٔ گاس ـ مارکف*[1]

در قسمت ۲.۵ به بررسی خصوصیات یک تخمین‌زننده در حالت کلی پرداختیم و به این نتیجه رسیدیم که خصوصیات مطلوب یک تخمین‌زننده را می‌توان در حالت کلی به صورت بهترین تخمین‌زنندهٔ ناااریب خطی جمعبندی کرد. در واقع یک تخمین‌زننده، موقعی مطلوب است که اولاً تابع خطی از مشاهدات متغیر درون‌زا باشد، ثانیاً ناااریب باشد و بالاخره در مقایسه با تخمین‌زننده‌های خطی ناااریب دیگر، از حداقل واریانس برخوردار باشد. در مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

دیدیم که تخمین‌زننده‌های $\hat{\alpha}$ و $\hat{\beta}$، با توجه به معادله‌های ۱.۲۵ و ۱.۲۶ عبارتند از

۱. علامت * بدین معنی است که سطح موضوع از دوره کارشناسی بالاتر است.

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{X} \quad , \qquad \hat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2} \ .$$

بنا بر قضیهٔ گاس ـ مارکف[1] می‌توان نشان داد که $\hat{\alpha}$ و $\hat{\beta}$ خصوصیات مطلوب بهترین تخمین‌زنندهٔ نااریب خطی را دارند. اثبات این قضیه را در ذیل با دو روش مطرح می‌کنیم.

۱. روش مستقیم

در این روش باید نشان دهیم که تخمین‌زننده‌های $\hat{\alpha}$ و $\hat{\beta}$ در معادله‌های ۲۵ـ۱ و ۲۶ـ۱ خصوصیات بهترین تخمین‌زنندهٔ نااریب خطی را دارا هستند. بحث را روی $\hat{\beta}$ متمرکز می‌کنیم؛ زیرا برای $\hat{\alpha}$ می‌توان از استدلالی مشابه، استفاده کرد.

اولاً، برای اثبات اینکه $\hat{\beta}$ یک تابع خطی از مشاهدات Y_t است، کافی است معادلهٔ ۲۳ـ۲ را یک بار دیگر بنویسیم،

$$\hat{\beta}_{OLS} = w_1 Y_1 + w_2 Y_2 + \cdots + w_n Y_n = \sum w_t Y_t \ ,$$

که در آن w_t با توجه به معادلهٔ ۳ـ۲ عبارت است از

$$w_t = \frac{x_t}{\sum x_t^2} \ .$$

بنابراین، با توجه به ثابت بودن مقادیر w_t، می‌توان نتیجه گرفت که $\hat{\beta}_{OLS}$ یک تابع خطی برحسب Y_t است.

ثانیاً، برای اثبات نااریبی $\hat{\beta}_{OLS}$ به معادلهٔ ۵ـ۲ رجوع می‌کنیم،

$$E(\hat{\beta}_{OLS}) = \beta \ .$$

ثالثاً، باید نشان دهیم که اگر هر تخمین‌زنندهٔ خطی نااریب دیگری را در نظر بگیریم، واریانس آن بیشتر از واریانسی است که در معادلهٔ ۱۶ـ۲ یعنی

$$\mathrm{Var}(\hat{\beta}_{OLS}) = \frac{\sigma^2}{\sum x_t^2} \ ,$$

---

۱. شکل اولیهٔ این قضیه را ابتدا گاس در سال ۱۸۲۱ اثبات کرد. به پیوست «د» مراجعه شود.

برای $\hat{\beta}_{OLS}$ به دست آورده‌ایم. برای اثبات، یک تخمین‌زنندهٔ خطی ناااریب دیگر از β مثلاً $\hat{\beta}_*$ را در نظر می‌گیریم. چون $\hat{\beta}_*$ خطی است پس

$$\hat{\beta}_* = \sum d_t Y_t \ , \qquad (۲\text{-}۷۴)$$

و با توجه به اینکه $\hat{\beta}_*$، ناااریب است

$$E(\hat{\beta}_*) = \beta \ . \qquad (۲\text{-}۷۵)$$

مدل $Y_t = \alpha + \beta X_t + U_t$ را در معادلهٔ ۲ـ۷۴ قرار داده خواهیم داشت

$$\hat{\beta}_* = \alpha \sum d_t + \beta \sum d_t X_t + \sum d_t U_t \ .$$

از دو طرف رابطهٔ فوق امید ریاضی می‌گیریم،

$$E(\hat{\beta}_*) = \alpha \sum d_t + \beta \sum d_t X_t \ .$$

برای اینکه $\hat{\beta}_*$ یک تخمین‌زنندهٔ ناااریب از β شود، ضروری است رابطه‌های زیر برقرار باشد،

$$\sum d_t = ۰ \quad , \quad \sum d_t X_t = ۱ \ . \qquad (۲\text{-}۷۶)$$

بر اساس فرض استقلال مقادیر Y_t از یکدیگر و نیز فرض واریانس همسانی، اگر از دو طرف معادلهٔ ۲ـ۷۴ واریانس بگیریم، با توجه به $Var(Y_t) = \sigma^۲$، خواهیم داشت

$$Var(\hat{\beta}_*) = \sum d_t^۲ \sigma^۲ = \sigma^۲ \sum d_t^۲ \ . \qquad (۲\text{-}۷۷)$$

باید نشان دهیم که واریانس $\hat{\beta}_*$ که از معادلهٔ ۲ـ۷۷ به دست می‌آید از مقدار واریانس $\hat{\beta}_{OLS}$ کمتر است (قبلاً آن را از معادلهٔ ۲ـ۱۶ به دست آوردیم). در معادلهٔ ۲ـ۲۳ دیدیم که $\hat{\beta}_{OLS} = \sum w_t Y_t$ است. حال رابطهٔ زیر را تعریف می‌کنیم،

$$d_t = w_t + (d_t - w_t) \ .$$

دو طرف رابطهٔ فوق را مجذور کرده و برای تمام مقادیر t جمع می‌کنیم،

$$\sum d_t^2 = \sum w_t^2 + \sum (d_t - w_t)^2 + 2 \sum w_t (d_t - w_t) \quad . \qquad (۲\text{-}۷۸)$$

نشان می‌دهیم که

$$\sum w_t d_t = \frac{1}{\sum x_t^2} \quad . \qquad (۲\text{-}۷۹)$$

برای اثبات، اگر مقدار w_t از معادلهٔ ۳-۲، یعنی $w_t = \frac{x_t}{\sum x_t^2}$، را در طرف چپ رابطهٔ ۲-۷۹ قرار دهیم، خواهیم داشت

$$\sum w_t d_t = \sum \frac{d_t x_t}{\sum x_t^2} = \frac{\sum d_t x_t}{\sum x_t^2} \quad . \qquad (۲\text{-}۸۰)$$

با استفاده از رابطهٔ ۲-۷۶ داریم

$$\sum d_t X_t = \sum d_t (x_t + \bar{X}) \quad ,$$

$$= \sum d_t x_t + \bar{X} \sum d_t = \sum d_t x_t = 1 \quad ,$$

که با جایگزینی نتیجهٔ فوق در معادلهٔ ۲-۸۰ خواهیم داشت

$$\sum w_t d_t = \frac{1}{\sum x_t^2} \quad ,$$

که دقیقاً همان رابطهٔ ۲-۷۹ است. از طرف دیگر، طبق معادلهٔ ۲-۳ داریم

$$w_t^2 = \frac{x_t^2}{(\sum x_t^2)^2} \quad ,$$

یا

$$\sum w_t^2 = \frac{\sum x_t^2}{(\sum x_t^2)^2} = \frac{1}{\sum x_t^2} \quad . \qquad (۲\text{-}۸۱)$$

بنابراین

$$\sum w_t (d_t - w_t) = \sum w_t d_t - \sum w_t^2 \quad ,$$

که با استفاده از معادله‌های ۲ـ۷۹ و ۲ـ۸۱ خواهیم داشت

$$\sum w_t (d_t - w_t) = \frac{1}{\sum x_t^2} - \frac{1}{\sum x_t^2} = 0 \quad .$$

نتیجهٔ فوق را در معادلهٔ ۲ـ۷۸ قرار می‌دهیم،

$$\sum d_t^2 = \sum w_t^2 + \sum (d_t - w_t)^2 \quad .$$

دو طرف رابطهٔ فوق را در σ_0^2 ضرب می‌کنیم،

$$\sigma^2 \sum d_t^2 = \sigma^2 \sum w_t^2 + \sigma^2 \sum (d_t - w_t)^2 \quad . \qquad (۲ـ۸۲)$$

در معادلهٔ ۲ـ۱۴ دیدیم که

$$\mathrm{Var}\,(\hat{\beta}_{\mathrm{OLS}}) = \sigma^2 \sum w_t^2 \quad .$$

با استفاده از رابطهٔ فوق و معادلهٔ ۲ـ۷۷ معادلهٔ ۲ـ۸۲ را به صورت زیر می‌نویسیم،

$$\mathrm{Var}\,(\hat{\beta}_*) = \mathrm{Var}\,(\hat{\beta}_{\mathrm{OLS}}) + \sigma^2 \sum (d_t - w_t)^2 \quad .$$

جملهٔ $\sigma^2 \Sigma (d_t - w_t)^2$ همواره مثبت و بزرگتر از صفر است، بنابراین

$$\mathrm{Var}\,(\hat{\beta}_{\mathrm{OLS}}) < \mathrm{Var}\,(\hat{\beta}_*) \quad .$$

به ازای $d_t = w_t$، خواهیم داشت

$$\mathrm{Var}\,(\hat{\beta}_*) = \mathrm{Var}\,(\hat{\beta}_{\mathrm{OLS}}) \quad .$$

اما واضح است که در این حالت تخمین‌زنندهٔ $\hat{\beta}_*$ دقیقاً همان $\hat{\beta}_{\mathrm{OLS}}$ خواهد بود.

۲. روش غیرمستقیم

روش دیگر، اثبات خصوصیت بهترین تخمین‌زنندهٔ نااریب خطی برای تخمین‌زننده‌های حداقل مربعات معمولی این است که از ابتدا $\hat{\alpha}$ و $\hat{\beta}$ را چنان تخمین بزنیم که نه تنها

مجموع مربعات پسماند، یعنی Σe_t^2 ، را حداقل کند؛ بلکه خصوصیت بهترین تخمین‌زنندهٔ نااریب خطی را نیز دارا باشد. این روش، بسیار عمومی است و می‌تواند کاربردهای بسیاری در مسائل مختلف داشته باشد. مانند گذشته بحث را روی $\hat{\beta}$ متمرکز می‌کنیم؛ زیرا به راحتی می‌توان نتایج مشابهی را برای $\hat{\alpha}$ استنتاج کرد.

مدل رگرسیون $Y_t = \alpha + \beta X_t + U_t$ را در نظر می‌گیریم. با توجه به معادلهٔ ۲۵ـ۱ می‌دانیم

$$\hat{\beta}_{OLS} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} .$$

مقدار $\hat{\beta}_{OLS}$ که از رابطهٔ فوق به دست می‌آید یقیناً مقدار Σe_t^2 را حداقل می‌کند. در اینجا یک تخمین‌زنندهٔ غیر حداقل مربعات معمولی از β، مانند $\hat{\beta}_*$ را چنان تعریف می‌کنیم که بهترین تخمین‌زنندهٔ نااریب خطی باشد. می‌خواهیم ثابت کنیم که $\hat{\beta}_*$ دقیقاً برابر $\hat{\beta}_{OLS}$ خواهد بود.

با توجه به معادله‌های ۷۶ـ۲ و ۷۷ـ۲ می‌دانیم لازمهٔ اینکه $\hat{\beta}_*$ بهترین تخمین‌زنندهٔ نااریب خطی باشد این است که:

$$\sum d_t = 0 \quad , \quad \sum d_t X_t = 1 \quad , \quad \text{Var}(\hat{\beta}_*) = \sigma^2 \sum d_t^2 .$$

در اینجا اول اینکه، باید ضرایب d_t را چنان تعیین کنیم که سه رابطهٔ فوق به طور همزمان برقرار باشد و دوم اینکه، باید نشان دهیم که در چنین حالتی $\hat{\beta}_*$ دقیقاً همان $\hat{\beta}_{OLS}$ خواهد بود. کافی است d_t را چنان تعیین کنیم که معادلهٔ

$$\text{Var}(\hat{\beta}_*) = \sigma^2 \sum d_t^2 ,$$

به شرط اینکه رابطه‌های زیر حداقل شود،

$$\sum d_i = 0 \quad , \quad \sum d_i X_i = 1 .$$

ملاحظه می‌شود که این مسأله در واقع یک حداقل‌سازی مقید است که به راحتی می‌توان آن را به استفاده از ضرایب لاگرانژ حل کرد. تابع زیر را تعریف می‌کنیم،

$$\varphi = \sigma^2 \sum d_i^2 - \lambda_1 \left(\sum d_i\right) - \lambda_2 \left(\sum d_i X_i - 1\right) .$$

باید از φ نسبت به d_i و λ_1 و λ_2 مشتق جزئی گرفته و آنها را مساوی صفر قرار دهیم؛

$$\frac{\partial \varphi}{\partial d_1} = 0 \ , \ \frac{\partial \varphi}{\partial d_2} = 0 \ , \ldots , \frac{\partial \varphi}{\partial d_n} = 0 \text{ و } \frac{\partial \varphi}{\partial \lambda_1} = 0 \ , \ \frac{\partial \varphi}{\partial \lambda_2} = 0 .$$

بنابراین خواهیم داشت،

$$\begin{aligned}
&2 d_1 \sigma^2 - \lambda_1 - \lambda_2 X_1 = 0 \ , \\
&2 d_2 \sigma^2 - \lambda_1 - \lambda_2 X_2 = 0 \ , \\
&\vdots \\
&2 d_n \sigma^2 - \lambda_1 - \lambda_2 X_n = 0 \ , \qquad (۲\text{-}۸۳) \\
&- \sum d_i = 0 \ , \\
&- \left(\sum d_i X_i\right) + 1 = 0 .
\end{aligned}$$

ملاحظه می‌شود که $(n + 2)$ معادله داریم و مجهولات ما عبارتند از: n مقدار d_i و نیز λ_1 و λ_2. می‌توان n معادلهٔ اول را به صورت زیر نوشت،

$$\begin{aligned}
d_1 &= \frac{1}{2\sigma^2} \left(\lambda_1 + \lambda_2 X_1\right) \ , \\
d_2 &= \frac{1}{2\sigma^2} \left(\lambda_1 + \lambda_2 X_2\right) \ , \qquad (۲\text{-}۸۴) \\
&\vdots \\
d_n &= \frac{1}{2\sigma^2} \left(\lambda_1 + \lambda_2 X_n\right) .
\end{aligned}$$

معادله‌های فوق را با هم جمع می‌کنیم، خواهیم داشت

$$\sum_1^n d_i = \frac{1}{2\sigma^2} \left(\lambda_1 n + \lambda_2 \sum X_i\right) . \qquad (۲\text{-}۸۵)$$

همچنین اگر معادلهٔ اول از معادله‌های ۲ـ۸۴ را در X_1 و معادلهٔ دوم رادر X_2 و معادلهٔ سوم را در X_3 و ... ضرب کنیم و نتایج را با یکدیگر جمع کنیم، خواهیم داشت.

$$\sum d_t X_t = \frac{۱}{۲\sigma^۲} [\lambda_۱ \sum X_t + \lambda_۲ \sum X_t^۲] \quad . \qquad (۲\text{-}۸۶)$$

معادله‌های ۲ـ۸۵ و ۲ـ۸۶ را در دو معادلهٔ آخر سیستم معادله‌های ۲ـ۸۳ جایگزین کرده داریم

$$-\frac{۱}{۲\sigma^۲} (\lambda_۱ n + \lambda_۲ \sum X_t) = ۰ \quad ,$$

$$-\frac{۱}{۲\sigma^۲} [\lambda_۱ \sum X_t + \lambda_۲ \sum X_t^۲] = -۱ \quad .$$

دو معادلهٔ فوق را حل کرده و $\lambda_۱$ و $\lambda_۲$ را به صورت زیر به دست می‌آوریم،

$$\lambda_۱ = \frac{-۲\sigma^۲ \sum X_t}{n \sum X_t^۲ - (\sum X_t)^۲} \quad , \quad \lambda_۲ = \frac{۲ n \sigma^۲}{n \sum X_t^۲ - (\sum X_t)^۲} \quad .$$

مقادیر $\lambda_۱$ و $\lambda_۲$ از معادله‌های فوق را در سیستم معادله‌های ۲ـ۸۴ قرار می‌دهیم و معادله‌های حاصل را برای $d_۱$ ، $d_۲$ تا d_n حل می‌کنیم. خواهیم داشت

$$d_t = \frac{-(\sum X_t) + n X_t}{n (\sum X_t^۲) - (\sum X_t)^۲} = \frac{(X_t - \bar{X})}{\sum (X_t - \bar{X})^۲} \quad ,$$

یا

$$d_t = \frac{x_t}{\sum x_t^۲} \quad , \quad t = ۱, ۲, \ldots, n \quad . \qquad (۲\text{-}۸۷)$$

مقادیر d_t که از رابطهٔ فوق به دست می‌آید، $\hat{\beta}_*$ را نااریب و خطی و واریانس آن را حداقل می‌کند. مقدار d_t از معادلهٔ فوق را در معادلهٔ ۲ـ۷۴ قرار می‌دهیم،

$$\hat{\beta}_* = \sum d_t Y_t = \sum \frac{x_t Y_t}{\sum x_t^۲} = \frac{\sum x_t Y_t}{\sum x_t^۲} \quad ,$$

یا

$$\hat{\beta}_* = \frac{\sum x_t Y_t}{\sum x_t^۲} \quad ,$$

که دقیقاً برابر تخمین‌زنندهٔ $\hat{\beta}$ است و با استفاده از روش حداقل مربعات معمولی به دست

می‌آید؛ یعنی اگر قرار است $\hat{\beta}_*$ خصوصیت بهترین تخمین‌زنندهٔ نااریب خطی را داشته باشد، باید برابر $\hat{\beta}_{OLS}$ باشد.

روش فوق بسیار عمومی است، زیرا: اولاً، خصوصیت بهترین تخمین‌زنندهٔ نااریب خطی را تضمین می‌کند؛ ثانیاً، مقدار $\hat{\beta}$ را نیز معین می‌نماید؛ و ثالثاً می‌تواند واریانس $\hat{\beta}$ را نیز محاسبه کند. برای به دست آوردن واریانس $\hat{\beta}_*$، ابتدا معادلهٔ ۲ـ۷۷ را می‌نویسیم،

$$\text{Var}\,(\hat{\beta}_*) = \sigma^2 \sum d_t^2 \ .$$

برای محاسبهٔ $\sum d_t^2$ از معادلهٔ ۲ـ۸۷ استفاده می‌شود. دو طرف این معادله را در d_t ضرب کرده و برای تمام مشاهدات جمع می‌کنیم. خواهیم داشت

$$\sum d_t^2 = \frac{-(\sum X_t)\sum d_t + n\,(\sum d_t X_t)}{n\sum X_t^2 - (\sum X_t)^2} \ .$$

از دو معادلهٔ آخر سیستم معادله‌های ۲ـ۸۳ داریم

$$\sum d_t = 0 \quad , \quad \sum d_t X_t = 1 \ ,$$

بنابراین $\sum d_t^2$ را می‌توان به صورت زیر نوشت،

$$\sum d_t^2 = \frac{n}{n\sum X_t^2 - (\sum X_t)^2} = \frac{1}{\sum x_t^2} \ .$$

بدین ترتیب واریانس $\hat{\beta}_*$ عبارت است از

$$\text{Var}\,(\hat{\beta}_*) = \frac{\sigma^2}{\sum x_t^2} \ ,$$

که دقیقاً برابر $\text{Var}\,(\hat{\beta}_{OLS})$ است که در معادلهٔ ۲ـ۱۶ به دست آوردیم.

مسائل فصل دوم

۲-۱ برای مدل رگرسیون

$$Y_i = \alpha + \beta X_i + U_i \ .$$

مشاهدات زیر مفروض است.

X_i:	۲	۳	۱	۵	۹
Y_i:	٤	۷	۳	۹	۱۷

۱. با استفاده از روش حداقل مربعات معمولی، پارامترهای α و β را تخمین بزنید.

۲. مقادیر پسماند را حساب کنید و نشان دهید $\Sigma e_i = ۰$.

۳. تغییرات توضیح داده شده (ESS) و تغییرات توضیح داده نشده (RSS) را حساب کنید.

٤. نشان دهید که تغییرات توضیح داده نشده + تغییرات توضیح داده شده = کل تغییرات (TSS = ESS + RSS).

۵. r^2 را از چهار روش مختلف به دست آورید.

٦. با استفاده از r^2، مقدار تغییرات توضیح داده نشده را حساب کنید.

۷. واریانس U_i را تخمین بزنید (محاسبهٔ s^2).

۸. واریانس $\hat{\alpha}$ و واریانس $\hat{\beta}$ را تخمین بزنید.

۹. انحراف معیار $\hat{\alpha}$ و انحراف معیار $\hat{\beta}$ را تخمین بزنید.

۱۰. فاصلهٔ اطمینان ۹۵ درصد را برای α حساب کنید.

۱۱. فاصله اطمینان ۹۵ درصد را برای β حساب کنید.

۱۲. فرضیهٔ $H_0: \alpha = ۰$ را در مقابل فرضیهٔ $H_1: \alpha \neq ۰$ در سطح معنی‌دار ۵ درصد آزمون کنید.

۱۳. فرضیهٔ $H_0: \beta = 0$ را در مقابل فرضیهٔ $H_1: \beta \neq 0$ در سطح معنی‌دار ۵ درصد آزمون کنید.

۱۴. برای σ_U^2 فاصلهٔ اطمینان ۹۵ درصد بسازید.

۱۵. برای این مسأله جدول آنالیز واریانس را تشکیل دهید.

۱۶. فرضیهٔ $H_0: \beta = 0$ را در مقابل فرضیهٔ $H_1: \beta \neq 0$ با استفاده از توزیع F آزمون کنید.

۱۷. معنی‌دار بودن مدل رگرسیون را در سطح معنی‌داری ۵ درصد آزمون کنید (آزمون r^2).

۱۸. آمارهٔ t را از سه راه مختلف زیر برای فرضیهٔ $H_0: \beta = 0$ محاسبه کرده و تساوی نتایج را تفسیر کنید.

اولاً، فقط با استفاده از r^2؛

ثانیاً، با استفاده از استاندارد کردن $\hat{\beta}$؛

ثالثاً، فقط با استفاده از F.

۱۹. با داشتن آمارهٔ آزمون t، برای فرضیهٔ $H_0: \beta = 0$، مقدار ضریب تعیین r^2 را حساب کنید.

۲-۲ در مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

نشان دهید که اگر $\bar{X} = 0$ باشد، کوواریانس بین $\hat{\alpha}$ و $\hat{\beta}$ برابر صفر است. آیا می‌توانید این نتیجه را توجیه کنید.

۲-۳ مدلهای رگرسیون زیر را ملاحظه کنید،

$$Y_t = \alpha_1 + \beta_1 C_t + U_{1t} \ ,$$

$$C_t = \alpha_2 + \beta_2 Y_t + U_{2t} \ ,$$

که در آن Y_t و C_t به ترتیب درآمد کل و مصرف کل است. این دو مدل را به صورت زیر

تخمین زده‌ایم.

$$\hat{y}_t = ۱/۲\ c_t \ ,$$

$$\hat{c}_t = ۰/٦\ y_t \ ,$$

که در آن $\hat{y}_t$ و c_t و y_t برحسب انحراف از میانگین هستند. اگر داشته باشیم

$$Y_t \equiv C_t + S_t \ ,$$

که در آن S_t پس‌انداز است، آنگاه

۱. نسبت انحراف معیار S_t به انحراف معیار Y_t را حساب کنید.

۲. ضریب همبستگی بین Y_t و S_t را به دست آورید.

۳. ضریب همبستگی بین C_t و S_t را محاسبه کنید.

۲ـ٤ مشاهدات زیر در مورد X_t و Y_t مفروض است. X_t نرخ بیکاری و Y_t درصد کارگرانی است که از صنایع تولیدی خارج می‌شوند.

t	Y_t	X_t	t	Y_t	X_t
۱	۱/۳	٦/۲	۸	۲/۳	۳/٦
۲	۱/۲	۷/۸	۹	۲/٥	۳/۳
۳	۱/٤	٥/۸	۱۰	۲/۷	۳/۳
٤	۱/٤	٥/۷	۱۱	۲/۱	٥/٦
٥	۱/٥	٥	۱۲	۱/۸	٦/۸
٦	۱/۹	٤	۱۳	۲/۲	٥/٦
۷	۲/٦	۳/۲			

۱. مدل زیر را تخمین بزنید،

$$Y_t = \alpha + \beta\ X_t + U_t \ .$$

۲. فاصلهٔ اطمینان ۹٥ درصد برای β بسازید.

۳. فرضیهٔ $H_0: \beta = 0$ را در مقابل فرضیهٔ $H_1: \beta \neq 0$ در سطح معنی‌دار ۵ درصد آزمون کنید.

۴. فاصلهٔ اطمینان ۹۰ درصد برای σ^2 بسازید.

۲.۵ مدل رگرسیون زیر مفروض است،

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

که در آن U_t جملهٔ اختلال با میانگین صفر و واریانس ثابت است. برای تخمین α و β دو پژوهشگر، مستقل از یکدیگر دو نمونهٔ ۸ تایی به طور تصادفی انتخاب کرده‌اند و با روش حداقل مربعات معمولی پارامترها را تخمین زده‌اند. نتایج تخمین با مشاهدات هر یک از نمونه‌ها در جدول زیر ارائه شده است.

پژوهشگر اول		پژوهشگر دوم	
Y_t	X_t	Y_t	X_t
۴	۳	۲	۱
۴/۵	۳	۲/۵	۱
۴/۵	۳	۲/۵	۱
۳/۵	۳	۱/۵	۱
۴/۵	۴	۱۱/۵	۱۰
۴/۵	۴	۱۰/۵	۱۰
۵/۵	۴	۱۰/۵	۱۰
۵	۴	۱۱	۱۰
$\hat{Y}_t = ۱/۸۷۵ + ۰/۷۵۰ X_t$		$\hat{Y}_t = ۱/۵ + ۰/۹۷۰ X_t$	
(۱/۲۰) (۰/۳۳۹)		(۰/۲۷) (۰/۰۳۸)	
$r^2 = ۰/۴۵$		$r^2 = ۰/۹۹$	
$\sigma^2 = ۰/۴۸$		$\sigma^2 = ۰/۴۸$	

این نکته را توضیح دهید که چرا انحراف معیار $\hat{\beta}$ در مدلی که پژوهشگر اول تخمین زده است از مقدار انحراف معیار $\hat{\beta}$ در مدل دوم بیشتر است.

۲ـ۶ با توجه به اینکه واریانس $\hat{\beta}$ تابع معکوسی از واریانس X_t است، بهتر است مشاهدات مرکزی X_t را در نمونه حذف کنیم تا به واریانس بیشتری برای X_t و در نتیجه به واریانس کمتری برای $\hat{\beta}$ برسیم. بدین ترتیب به تخمینهای دقیقتری از β خواهیم رسید. آیا موافق چنین نظری هستید؟ چرا.

۲ـ۷ مدل رگرسیون زیر مفروض است،

$$Y_t = \beta X_t + U_t \ ,$$

که در آن U_t همهٔ فرضهای کلاسیک را دربر دارد. جمله‌های پسماند را به صورت زیر تعریف می‌کنیم،

$$e_t = Y_t - \hat{\beta} X_t \ .$$

نشان دهید که

$$۱. \quad E(e_t^2) = \sigma^2 \left(1 - \frac{X_t^2}{\Sigma X_t^2}\right) \ ,$$

که در آن σ^2 واریانس جملهٔ اختلال است.

$$۲. \quad E(e_t e_s) = \frac{-\sigma^2 X_t X_s}{\Sigma X_t^2} \ , \quad s \neq t \ .$$

۳. در حالت کلی $\Sigma e_t \neq 0$. این نتیجه را با مدلی که ضریب ثابت دارد، مقایسه کنید. آیا $E(\Sigma e_t)$ برابر صفر است؟

۴. نشان دهید که

$$E\left(\frac{\Sigma e_t^2}{n-1}\right) = \sigma^2 \ .$$

۲ـ۸ مدل رگرسیون زیر مفروض است،

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

که در آن U_t شامل تمام فرضهای کلاسیک است.

۱. نشان دهید که شکل عمومی تخمین‌زننده‌های خطی نااریب به صورت زیر است،

$$\tilde{\beta} = \frac{\Sigma z_t y_t}{\Sigma z_t x_t} ,$$

z_t در آن یک متغیر برون‌زاست که در نمونه‌گیریهای تکراری ثابت فرض می‌شود. همچنین می‌دانیم $z_t = Z_t - \bar{Z}$.

۲. واریانس صورت عمومی تخمین‌زننده‌های خطی نااریب را به دست آورید و با واریانس تخمین‌زننده‌های حداقل مربعات معمولی مقایسه کنید.

۳. نشان دهید کارآیی نسبی $\tilde{\beta}$ (در مقایسه با $\hat{\beta}_{OLS}$) تابعی از همبستگی بین x_t و z_t است، به گونه‌ای که وقتی این همبستگی کامل باشد، واریانس $\tilde{\beta}$ با واریانس $\hat{\beta}$ برابر می‌شود.

حل مسائل فصل دوم

۲-۱ برای پاسخ به این سؤال داریم:

۱. تخمینهای α و β را یک بار با استفاده از مشاهدات اصلی و بار دیگر با روش انحراف از میانگین به دست می‌آوریم. ابتدا جدول زیر را تشکیل می‌دهیم.

جدول ۱

	X_i	Y_i	$X_i Y_i$	X_i^2	$\hat{Y}_i$	$e_i = Y_i - \hat{Y}_i$
	۲	٤	۸	٤	٤/٥٠	-۰/٥٠
	۳	۷	۲۱	۹	٦/۲٥	۰/۷٥
	۱	۳	۳	۱	۲/۷٥	۰/۲٥
	٥	۹	٤٥	۲٥	۹/۷٥	-۰/۷٥
	۹	۱۷	۱٥۳	۸۱	۱٦/۷٥	۰/۲٥
Σ	۲۰	٤۰	۲۳۰	۱۲۰	٤۰	۰

$$\hat{\beta} = \frac{n \Sigma X_i Y_i - \Sigma X_i \Sigma Y_i}{n \Sigma X_i^2 - (\Sigma X_i)^2} ,$$

$$= \frac{٥ (۲۳۰) - ۲۰ (٤۰)}{٥ (۱۲۰) - (۲۰)^2} = \frac{۱۱٥۰ - ۸۰۰}{٦۰۰ - ٤۰۰} = ۱/۷٥ .$$

$$\hat{\alpha} = \frac{\Sigma X_i^2 \Sigma Y_i - \Sigma X_i \Sigma X_i Y_i}{n \Sigma X_i^2 - (\Sigma X_i)^2} ,$$

$$= \frac{۱۲۰ (٤۰) - ۲۰ (۲۳۰)}{٥ (۱۲۰) - (۲۰)^2} = \frac{٤۸۰۰ - ٤٦۰۰}{٦۰۰ - ٤۰۰} = ۱ .$$

در اینجا $\hat{\alpha}$ و $\hat{\beta}$ را با استفاده از مشاهداتی به دست می‌آوریم که برحسب انحراف از

میانگین محاسبه شده است. ابتدا جدول زیر را تشکیل می‌دهیم.

جدول ۲

	x_i	y_i	$x_i y_i$	x_i^2	y_i^2
	-۲	-۴	۸	۴	۱۶
	-۱	-۱	۱	۱	۱
	-۳	-۵	۱۵	۹	۲۵
	۱	۱	۱	۱	۱
Σ	۵ / ۰	۹ / ۰	۴۵ / ۷۰	۲۵ / ۴۰	۸۱ / ۱۲۴

$$\hat{\beta} = \frac{\Sigma x_i y_i}{\Sigma x_i^2} = \frac{۷۰}{۴۰} = ۱/۷۵ \ .$$

$$\bar{X} = \frac{۲۰}{۵} = ۴ \quad , \quad \bar{Y} = \frac{۴۰}{۵} = ۸ \ ,$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{X} = ۸ - ۱/۷۵\,(۴) = ۱ \ ,$$

بدین ترتیب تخمین مدل رگرسیون به شرح زیر خواهد بود،

$$\hat{Y}_i = ۱ + ۱/۷۵\,X_i \ .$$

۲. با استفاده از جدول ۱ ملاحظه می‌شود که $\Sigma e_i = ۰$.

۳. با استفاده از معادلهٔ ۴۵ـ۲ داریم

$$ESS = \sum \hat{y}_i^2 = \hat{\beta} \sum x_i y_i \ ,$$

$$= ۱/۷۵\,(۷۰) = ۱۲۲/۵ \ .$$

مجموع مربعات پسماند (RSS) را از معادلهٔ ۴۶ـ۲ به دست می‌آوریم،

$$RSS = \sum e_i^2 = \sum y_i^2 - \hat{\beta} \sum x_i y_i \ ,$$

$$= ۱۲۴ - ۱/۷۵\,(۷۰) = ۱۲۴ - ۱۲۲/۵ = ۱/۵ \ .$$

٤. برای ثبات: تغییرات توضیح‌داده‌نشده + تغییرات توضیح‌داده‌شده = کل تغییرات کافی است مقادیر مربوط را جایگزین کنیم،

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 \ ,$$

$$۱۲٤ = ۱۲۲/٥ + ۱/٥ \ .$$

٥. ابتدا از فرمول ۳۲ـ۱ استفاده می‌کنیم. می‌دانیم ضریب تعیین برابر مجذور ضریب همبستگی است،

$$r_{x,y} = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \ . \ \sum y_i^2}} \ ,$$

$$= \frac{۷۰}{\sqrt{٤۰ (۱۲٤)}} = ۰/۹۹۳۹۳۳۲۰۹ \ ,$$

در نتیجه:

$$r^2 \cong ۰/۹۸۸ \ .$$

با استفاده از فرمول ۳۳ـ۱ داریم

$$r^2 = \frac{ESS}{TSS} = \frac{\sum \hat{y}_i^2}{\sum y_i^2} \ ,$$

$$= \frac{۱۲۲/٥}{۱۲٤} = ۰/۹۸۸ \ .$$

روش سوم این است که از فرمول ۳۵ـ۱ استفاده کنیم،

$$r^2 = \hat{\beta} \frac{\sum x_i y_i}{\sum y_i^2} \ ,$$

$$= ۱/۷٥ \frac{۷۰}{۱۲٤} = \frac{۱۲۲/٥}{۱۲٤} = ۰/۹۸۸ \ .$$

روش چهارم این است که مبتنی بر فرمول ٤۳ـ۱ عمل کنیم،

$$r^2 = ۱ - \frac{RSS}{TSS} \ ,$$

$$= ۱ - \frac{۱/۵}{۱۲٤} = \frac{۱۲۲/۵}{۱۲٤} = ۰/۹۸۸ \ .$$

۶. با استفاده از فرمول ۴۳ـ۱ داریم

$$r^۲ = ۱ - \frac{RSS}{TSS} \ ,$$

$$۰/۹۸۸ = ۱ - \frac{RSS}{۱۲٤} \ ,$$

$$\sum e_i^۲ = RSS = ۰/۰۱۲ \ (۱۲٤) = ۱/٤۸۸ \ .$$

ملاحظه می‌شود مقدار دقیق تغییرات توضیح داده نشده (RSS) برابر ۱/۵ است؛ در حالی که این روش به مقدار ۱/٤۸۸ می‌رسیم. دلیل این امر وجود تقریبی است که معمولاً در $r^۲$ وجود دارد.

۷. می‌دانیم

$$s^۲ = \hat{\sigma}^۲ = \frac{\Sigma e_i^۲}{n - ۲} \ ,$$

$$= \frac{۱/۵}{۵ - ۲} = ۰/۵ \ .$$

۸. ابتدا واریانس $\hat{\alpha}$ را تخمین می‌زنیم. از فرمول ۲۱ـ۲ داریم

$$Var \ (\hat{\alpha}) = \hat{\sigma}^۲ \left[\frac{۱}{n} + \frac{\bar{X}^۲}{\Sigma x_i^۲}\right] \ ,$$

$$= ۰/۵ \left[\frac{۱}{۵} + \frac{۱۶}{٤۰}\right] = \frac{۱۲}{٤۰} = ۰/۳ \ .$$

برای تخمین واریانس $\hat{\beta}$ کافی است از معادلهٔ ۱۶ـ۲ استفاده کنیم،

$$Var \ (\hat{\beta}) = \frac{\hat{\sigma}^۲}{\Sigma x_i^۲} \ ,$$

$$= \frac{۰/۵}{٤۰} = ۰/۰۱۲۵ \ .$$

۹.

$$SE(\hat{\alpha}) = \sqrt{Var(\hat{\alpha})} = \sqrt{۰/۳} = ۰/۵۴۷۷ ,$$

$$SE(\hat{\beta}) = \sqrt{Var(\hat{\beta})} = \sqrt{۰/۰۱۲۵} = ۰/۱۱۱۸ .$$

۱۰. برای فاصلهٔ اطمینان ۹۵ درصد باید از سطح معنی‌دار ۵ درصد استفاده کنیم؛ بنابراین مساحت ناحیهٔ بحرانی در هر طرف برابر ۰/۰۲۵ خواهد بود. مقدار به دست آمده از جدول t در سطح معنی‌دار ۲/۵ درصد و با ۳ درجه آزادی برابر است با ۳/۱۸۲؛ بنابراین خواهیم داشت

$$\hat{\alpha} - t_{\alpha/۲} SE(\hat{\alpha}) < \alpha < \hat{\alpha} + t_{\alpha/۲} SE(\hat{\alpha}) ,$$

$$۱ - ۳/۱۸۲ (۰/۵۴۷۷) < \alpha < ۱ + ۳/۱۸۲ (۰/۵۴۷۷) ,$$

$$- ۰/۷۴ < \alpha < ۲/۷۴ .$$

۱۱. فاصلهٔ اطمینان ۹۵ درصد برای β عبارت است از

$$\hat{\beta} - t_{\alpha/۲} SE(\hat{\beta}) < \beta < \hat{\beta} + t_{\alpha/۲} SE(\hat{\beta}) ,$$

$$۱/۷۵ - ۳/۱۸۲ (۰/۱۱۱۸) < \beta < ۱/۷۵ + ۳/۱۸۲ (۰/۱۱۱۸) ,$$

$$۱/۳۹ < \beta < ۲/۱۱ .$$

۱۲. برای آزمون فرضیهٔ $H_۰: \alpha = ۰$، در مقابل $H_۱: \alpha \neq ۰$ مقدار آمارهٔ آزمون را به دست می‌آوریم.

$$t = \frac{\hat{\alpha} - ۰}{SE(\hat{\alpha})} = \frac{۱ - ۰}{۰/۵۴۷۷} = ۱/۸۲۶ .$$

ملاحظه می‌شود که چون آمارهٔ آزمون از مقدار جدول t کمتر است، این آماره معنی‌دار نبوده، در نتیجه فرضیهٔ $H_۰$ رد نمی‌شود.

مقدار به دست آمده از جدول t > آمارهٔ آزمون

$$۱/۸۲۶ < ۳/۱۸۲ .$$

بنابراین ضریب ثابت در مدل رگرسیون مفروض در سطح معنی‌دار ۵ درصد معتبر نیست. البته به محاسبهٔ آمارهٔ آزمون نیازی نبود؛ زیرا قبلاً فاصلهٔ اطمینان را ساخته‌ایم. در واقع چون $\alpha = ۰$ در فاصلهٔ اطمینان ۹۵ درصد برای α قرار دارد، فرضیهٔ $H_۰: \alpha = ۰$ را نمی‌توان رد کرد.

۱۳. ابتدا آمارهٔ آزمون را به دست می‌آوریم،

$$t = \frac{\hat{\beta} - ۰}{SE(\hat{\beta})} = \frac{۱/۷۵ - ۰}{۰/۱۱۱۸} = ۱۵/۶۵۳ \ ,$$

و چون این آماره از مقدار جدول t (۳/۱۸۲) بیشتر شده است، بنابراین در ناحیهٔ بحرانی قرار گرفته و فرضیهٔ $H_۰$ رد می‌شود. مانند بحث بند ۱۲، می‌توان گفت به محاسبهٔ t نیازی نیست؛ زیرا قبلاً فاصلهٔ اطمینان β را ساخته‌ایم. چون $\beta = ۰$ در فاصله اطمینان ۹۵ درصد برای β قرار نمی‌گیرد؛ پس فرضیهٔ $\beta = ۰$ قابل قبول نیست.

۱۴. برای ساختن فاصلهٔ اطمینان ۹۵ درصد برای واریانس جامعه $\sigma^۲$ ابتدا باید دو نقطه در روی محور $\chi^۲$ تعیین کنیم که مساحت سطح زیر منحنی از مبدأ مختصات تا آن نقاط به ترتیب برابر ۰/۰۲۵ و ۰/۹۷۵ باشد. این نقاط را می‌توان با ۳ درجه آزادی از جدول $\chi^۲$ به دست آورد. خواهیم داشت

$$\chi^۲_{/۰۲۵}(۳) = ۰/۲۱۶ \quad , \quad \chi^۲_{/۹۷۵}(۳) = ۹/۳۵ \ .$$

با توجه به معادلهٔ ۲ـ۴۴ داریم

$$\frac{(n-۲)\hat{\sigma}^۲}{\chi^۲_{/۹۷۵}} < \sigma^۲ < \frac{(n-۲)\hat{\sigma}^۲}{\chi^۲_{/۰۲۵}} \ .$$

می‌دانیم

$$\hat{\sigma}^۲ = \frac{\sum e_i^۲}{n-۲} \ ,$$

در نتیجه

$$(n-۲)\hat{\sigma}^۲ = \sum e_i^۲ = ۱/۵ \ ,$$

بنابراین

$$\frac{۱/۵}{۹/۳۵} < \sigma^۲ < \frac{۱/۵}{۰/۲۱۶} \ ,$$

$$۰/۱۶ < \sigma^۲ < ۶/۹٤ .$$

۱۵. برای تشکیل جدول آنالیز واریانس به کمیتهای زیر نیاز داریم،

$$ESS \quad ، \quad RSS \quad ، \quad TSS \quad ، \quad ESS/۱ \quad ، \quad RSS/(n - ۲)$$

ملاحظه می‌شود تمام این کمیتها را در بندهای قبلی محاسبه کرده‌ایم؛ بنابراین جدول تجزیه واریانس را به شرح زیر می‌نویسیم

منبع تغییرات	مجموع مربعات	درجات آزادی	میانگین مربعات یا واریانس
X_i	ESS = ۱۲۲/۵	۱	ESS/۱ = ۱۲۲/۵
e_i	RSS = ۱/۵	۳	RSS/۳ = ۰/۵
کل	TSS = ۱۲٤	٤	

۱۶. از معادلهٔ ۱۵ـ۲ مقدار آمارهٔ آزمون F را به دست می‌آوریم، داریم

$$F = \frac{ESS/۱}{RSS/(n - ۲)} ,$$

$$= \frac{۱۲۲/۵}{۰/۵} = ۲٤۵ .$$

در اینجا باید مقدار F از جدول پیوست F را با درجات آزادی ۱ (متعلق به صورت) و ۳ (متعلق به مخرج) و در سطح معنی‌دار ۵ درصد به دست آورد. خواهیم داشت

$$F_{۰/۹۵} (۱ و ۳) = ۱۰/۱ .$$

ملاحظه می‌شود آمارهٔ F در ناحیهٔ بحرانی قرار می‌گیرد و فرضیهٔ $H_۰ : \beta = ۰$ رد می‌شود.

۱۷. با استفاده از معادلهٔ ۵٤ـ۲ داریم

$$F = \frac{(n - ۲) r^۲}{۱ - r^۲} ,$$

$$= \frac{(۳)\ ۰/۹۸۸}{۱ - ۰/۹۸۸} = ۲۴۷ \ .$$

مقدار F که با روش $r^۲$ به دست آمده است کمی با مقدار دقیق F که در بند ۱۶ محاسبه شده متفاوت است. علت این امر، وجود تقریب در محاسبهٔ $r^۲$ است. کافی است که آمارهٔ آزمون F را با مقدار جدول F مقایسه کنیم. می‌دانیم

$$F_{۰/۹۵}\ (۱ \text{ و } ۳) = ۱۰/۱ \ ,$$

نتیجه می‌گیریم که آمارهٔ آزمون در ناحیهٔ بحرانی قرار می‌گیرد و فرضیهٔ $H_۰$ رد می‌شود؛ یعنی تخمین مدل رگرسیون معنی‌دار است.

۱۸. **اولاً**، برای محاسبهٔ t با روش $r^۲$ باید از معادلهٔ ۵۵ـ۲ استفاده کنیم،

$$t^۲ = \frac{(n - ۲)\ r^۲}{۱ - r^۲} \ ,$$

$$= \frac{۳\ (۰/۹۸۸)}{۱ - ۰/۹۸۸} = \frac{۲/۹۶۴}{۰/۰۱۲} = ۲۴۷ \ ,$$

$$t = \pm \sqrt{۲۴۷} = \pm\ ۱۵/۷۱ \ .$$

ثانیاً، محاسبهٔ t با روش استاندارد کردن $\hat{\beta}$، در بند ۱۳ انجام شده است.

$$t = \frac{\hat{\beta}}{SE\ (\hat{\beta})} = \frac{۱/۷۵}{۰/۱۱۱۸} = ۱۵/۶۵ \ .$$

ثالثاً، می‌دانیم

$$t = \pm \sqrt{F} \ .$$

در بند ۱۶ مقدار F را برای فرضیهٔ $\beta = ۰$ برابر $F = ۲۴۵$ به دست آوردیم؛ بنابراین

$$t = \sqrt{۲۴۵} = \pm\ ۱۵/۶۵ \ .$$

اگر تقریبهای موجود در محاسبهٔ $r^۲$ را نادیده بگیریم، مقادیری که برای t با سه

روش فوق به دست آمده است با یکدیگر برابر است تفسیر این نتایج بدین ترتیب است که سه راه حل فوق در واقع صورتهای مختلف بیان یک واقعیت است؛ به عبارت دیگر، راه حل اول ـ که از r^2 استفاده می‌شود ـ بر آزمون ضریب تعیین مبتنی است. در یک مدل رگرسیون ساده، آزمون فرضیهٔ معنی‌دار بودن مدل رگرسیون چیزی نیست جز آزمون $\beta = 0$؛ زیرا تمام ساختار مدل رگرسیون ساده به β بستگی دارد. اگر $H_0: \beta = 0$ رد نشود بدین معنی است که کل مدل قابل قبول نیست؛ راه حل دوم نیز که دقیقاً همان فرضیهٔ $H_0: \beta = 0$ است؛ و راه حل سوم که بر آزمون F مبتنی است، در حقیقت معنی‌دار بودن تغییرات توضیح داده شده را آزمون می‌کند. می‌دانیم در یک مدل رگرسیون ساده، معنی‌دار بودن تغییرات توضیح داده شده دقیقاً بر فرض معنی‌داری شیب مدل متکی است؛ به همین دلیل آزمون دوم و سوم نیز از نظر ماهیت یکی هستند.

۱۹. با استفاده از معادلهٔ ۵۶ـ۲ داریم

$$r^2 = \frac{t^2}{t^2 + (n-2)} ,$$

$$= \frac{(۱۵/۶۵۳)^2}{(۱۵/۶۵۳)^2 + ۳} = \frac{۲۴۵/۰۱۶}{۲۴۸/۰۱۶} ,$$

$$\cong ۰/۹۸۸ .$$

۲ـ۲ با توجه به معادلهٔ ۲۵ـ۲ می‌دانیم

$$\text{Cov}(\hat{\alpha}, \hat{\beta}) = \frac{-\bar{X}\sigma^2}{\Sigma x_i^2} .$$

بنابراین وقتی $\bar{X} = 0$، آنگاه $\text{Cov}(\hat{\alpha}, \hat{\beta}) = 0$ است. این نتیجه را می‌توان چنین توجیه کرد که می‌دانیم

$$\hat{\alpha} = \bar{Y} - \beta \bar{X} ,$$

وقتی $\bar{X} = 0$ در نتیجه $\hat{\alpha} = \bar{Y}$. حال می‌گوییم که اگر به Y_i مقدار ثابت k را اضافه کنیم

یا از آن کم کنیم، مقدار $\hat{\beta}$ تغییری نمی‌کند؛ زیرا

$$\hat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} = \frac{\Sigma x_t Y_t}{\Sigma x_t^2} \quad .$$

اگر به Y_t مقدار k را اضافه کنیم، خواهیم داشت

$$\hat{\beta} = \frac{\Sigma x_t (Y_t + k)}{\Sigma x_t^2} = \frac{\Sigma x_t Y_t + k \Sigma x_t}{\Sigma x_t^2} \quad ,$$

$$= \frac{\Sigma x_t y_t}{\Sigma x_t^2} \quad .$$

بنابراین می‌توان به راحتی نتیجه گرفت که $\hat{\beta}$، تابعی از $\bar{Y}$ نیست. در نتیجه اگر $\bar{X} = 0$ باشد، از یک طرف $\hat{\alpha}$ دقیقاً برابر $\bar{Y}$ می‌شود و از طرف دیگر $\hat{\beta}$ از $\bar{Y}$ مستقل خواهد بود؛ بنابراین باید انتظار داشت که کوواریانس $\hat{\alpha}$ و $\hat{\beta}$ برابر صفر شود.

۲.۳ الف) از رابطهٔ $Y_t \equiv C_t + S_t$، داریم $y_t \equiv c_t + s_t$ واریانس S_t را حساب می‌کنیم.

$$\text{Var}(S_t) = \frac{\Sigma (S_t - \bar{s})^2}{n} \quad ,$$

$$= \frac{\Sigma (s_t)^2}{n} = \frac{\Sigma (y_t - c_t)}{n} = \frac{\Sigma y_t^2}{n} + \frac{\Sigma c_t^2}{n} - \frac{2 \Sigma y_t c_t}{n} \quad . \quad (1)$$

دو طرف رابطهٔ فوق را بر واریانس Y_t تقسیم می‌کنیم،

$$\frac{\text{Var}(S_t)}{\text{Var}(Y_t)} = 1 + \frac{\text{Var}(C_t)}{\text{Var}(Y_t)} - \frac{2 \Sigma y_t c_t}{\Sigma y_t^2} \quad ,$$

$$= 1 + \frac{\text{Var}(C_t)}{\text{Var}(Y_t)} - 2 \hat{\beta}_2 \quad . \quad (2)$$

در یک مدل رگرسیون $Y_t = \alpha + \beta X_t + U_t$، ضریب همبستگی بین X_t و Y_t برابر است با

$$r = \frac{\Sigma x_t y_t}{\sqrt{\Sigma x_t^2 \Sigma y_t^2}} .$$

عبارت فوق را می‌توان به صورت زیر نوشت،

$$r = \frac{\Sigma x_t y_t}{\Sigma x_t^2} \cdot \frac{\sqrt{\Sigma x_t^2}}{\sqrt{\Sigma y_t^2}} .$$

کسر اول برابر $\hat{\beta}$ است و اگر صورت و مخرج کسر دوم را بر $\sqrt{n}$ تقسیم کنیم، به انحراف معیار X و انحراف معیار Y خواهیم رسید؛ بنابراین

$$r = \hat{\beta} \frac{SE(X_i)}{SE(Y_i)} ,$$

یا:

$$\frac{SE(Y_i)}{SE(X_i)} = \frac{\hat{\beta}}{r} \qquad (۳)$$

فرمول کلی فوق را با مدل رگرسیون $C_t = \alpha_۲ + \beta_۲ Y_t + U_{۲t}$ تطبیق می‌دهیم،

$$\frac{SE(C_i)}{SE(Y_i)} = \frac{\hat{\beta}_۲}{r} = \frac{۰/۶}{r} . \qquad (٤)$$

به همین ترتیب، با توجه به معادلهٔ (۳)، برای مدل رگرسیون $Y_t = \alpha_۱ + \beta_۱ C_t + U_{۱t}$ خواهیم داشت

$$\frac{SE(Y_i)}{SE(C_i)} = \frac{\hat{\beta}_۱}{r} = \frac{۱/۲}{r} . \qquad (۵)$$

دو طرف معادلهٔ (٤) را بر (۵) تقسیم می‌کنیم،

$$\frac{[SE(C_i)]^۲}{[SE(Y_i)]^۲} = \frac{\hat{\beta}_۲}{\hat{\beta}_۱} = \frac{۰/۶}{۱/۲} = ۰/۵ ,$$

یا

$$\frac{Var(C_i)}{Var(Y_i)} = ۰/۵ .$$

اگر این رابطه را در معادلهٔ (۲) قرار دهیم، خواهیم داشت

$$\frac{\text{Var }(S_t)}{\text{Var }(Y_t)} = ۱ + ۰/۵ - ۲\,(۰/۶) = ۰/۳ \ ,$$

و در نتیجه

$$\frac{\text{SE }(S_t)}{\text{SE }(Y_t)} = \sqrt{۰/۳} \ . \qquad (۶)$$

ب) برای به دست آوردن ضریب همبستگی بین Y_t و S_t و نیز بین C_t و S_t به ترتیب زیر عمل می‌کنیم. ابتدا نسبت $\frac{\text{SE }(S_t)}{\text{SE }(C_t)}$ را به دست می‌آوریم. دو طرف معادلهٔ (۱) را بر Var (C_t) تقسیم می‌کنیم،

$$\frac{\text{Var }(S_t)}{\text{Var }(C_t)} = \frac{\text{Var }(Y_t)}{\text{Var }(C_t)} + ۱ - ۲\,\hat{\beta}_۱ \ ,$$

$$= ۲ + ۱ - ۲\,(۱/۲) = ۳ - ۲/۴ = ۰/۶ \ ,$$

در نتیجه

$$\frac{\text{SE }(S_t)}{\text{SE }(C_t)} = \sqrt{۰/۶} \ . \qquad (۷)$$

حال به محاسبهٔ $r_{s,y}$ و $r_{s,c}$ می‌پردازیم. مطابق فرمول (۳) داریم

$$r_{s,y} = \hat{\beta}_{s,y} \frac{\text{SE }(Y_t)}{\text{SE }(S_t)} \ ,$$

$$r_{s,c} = \hat{\beta}_{s,c} \frac{\text{SE }(C_t)}{\text{SE }(S_t)} \ .$$

قبلاً در معادله‌های (۶) و (۷) نسبت انحراف معیارها را حساب کرده‌ایم. فقط باید $\hat{\beta}_{s,y}$ و $\hat{\beta}_{s,c}$ را به دست آورد،

$$\hat{\beta}_{s,y} = \frac{\Sigma\, s_t\, y_t}{\Sigma\, y_t^۲} = \frac{\Sigma\,(y_t - c_t)\, y_t}{\Sigma\, y_t^۲} \ ,$$

$$= ۱ - \frac{\Sigma c_t y_t}{\Sigma y_t^۲} = ۱ - \hat{\beta}_{c,y} ,$$

$$= ۱ - \hat{\beta}_۲ = ۱ - ۰/۶ = ۰/۴ .$$

به همین ترتیب

$$\hat{\beta}_{s,c} = \frac{\Sigma s_t c_t}{\Sigma c_t^۲} = \frac{\Sigma (y_t - c_t) c_t}{\Sigma c_t^۲} ,$$

$$= \frac{\Sigma y_t c_t}{\Sigma c_t^۲} - ۱ = \hat{\beta}_{y,c} - ۱ ,$$

$$= \hat{\beta}_۱ - ۱ = ۱/۲ - ۱ = ۰/۲ .$$

بدین ترتیب، خواهیم داشت

$$r_{s,y} = ۰/۴ \left(\frac{۱}{\sqrt{۰/۳}}\right) = ۰/۷۳ ,$$

$$r_{s,c} = ۰/۲ \left(\frac{۱}{\sqrt{۰/۶}}\right) = ۰/۲۶ ,$$

۴ـ۲ الف) تخمین مدل به صورت زیر خواهد بود،

$$\hat{Y}_t = \underset{(۰/۳۳۱)}{۳/۳۶۶} - \underset{(۰/۰۶۳)}{۰/۲۸۶}\ X_t ,$$

$$r^۲ = ۰/۶۵۳ ,$$

$$RSS = \sum e_t^۲ = ۱/۱۴۳۵ .$$

ب) فاصلهٔ اطمینان ۹۵ درصد برای β، با توجه به مقدار t با ۱۱ درجهٔ آزادی، یعنی $t = \pm ۲/۲۰۱$، عبارت است از

$$-۰/۲۸۶ - ۲/۲۰۱ (۰/۰۶۳) < \beta < -۰/۲۸۶ + ۲/۲۰۱ (۰/۰۶۳) ,$$

$$-۰/۴۲۵ < \beta < -۰/۱۴۷ .$$

ج) با توجه به اینکه $\beta = ۰$ در فاصلهٔ اطمینان ۹۵ درصد برای β واقع نمی‌شود، بنابراین فرضیهٔ $H_۰$ را در سطح معنی‌دار ۵ درصد رد می‌کنیم.

د) برای ساختن یک فاصلهٔ اطمینان ۹۰ درصدی برای $\sigma^۲$، به این نکته توجه می‌کنیم که $\frac{\Sigma e_i^۲}{\sigma^۲}$ دارای یک توزیع $\chi^۲$ با ۱۱ درجه آزادی است.
با استفاده از معادلهٔ ۲.۴۴ داریم

$$\frac{(n-۲)\,\hat{\sigma}^۲}{\chi^۲_{۰/۹۵}} < \sigma^۲ < \frac{(n-۲)\,\hat{\sigma}^۲}{\chi^۲_{۰/۰۵}} ,$$

$$\frac{۱/۱۴۳۵}{۱۹/۷} < \sigma^۲ < \frac{۱/۱۴۳۵}{۴/۵۷} ,$$

$$۰/۰۵۸ < \sigma^۲ < ۰/۲۵۰ .$$

۲.۵ می‌دانیم واریانس $\hat{\beta}$ از فرمول زیر به دست می‌آید،

$$\text{Var}\,(\hat{\beta}) = \frac{\hat{\sigma}^۲}{\Sigma x_i^۲} = \frac{\hat{\sigma}^۲}{\Sigma\,(X_i - \bar{X})^۲} .$$

ملاحظه می‌شود واریانس $\hat{\beta}$ با $\Sigma\,(X_i - \bar{X})^۲$ نسبت معکوس دارد. به موازات افزایش واریانس یا پراکندگی متغیر برون‌زای X_i، مقدار $\Sigma\,(X_i - \bar{X})^۲$ زیاد می‌شود و واریانس $\hat{\beta}$ کمتر می‌شود. در نمونهٔ ۸ تایی پژوهشگر دوم پراکندگی مشاهدات X_i بسیار زیاد است(۱ و ۱۰)؛ در حالی که این پراکندگی برای نمونهٔ اول بسیار کم است (۳ و ۴). بنابراین باید انتظار داشت که واریانس متغیر X_i در نمونهٔ دوم بسیار کمتر از مقدار مشابه برای نمونهٔ اول باشد. در نتیجه واریانس $\hat{\beta}$ و نیز انحراف معیار آن برای نمونهٔ دوم بسیار کمتر از انحراف معیار $\hat{\beta}$ برای نمونه اول خواهد بود.

۲.۶ این مسأله شبیه نکته‌ای است که در مسألهٔ ۲.۵ مطرح شده است. می‌دانیم

$$\text{Var}\,(\hat{\beta}) = \frac{\hat{\sigma}^2}{\Sigma\, x_t^2} = \frac{\hat{\sigma}^2}{\Sigma\,(X_t - \bar{X})^2} \ .$$

همچنین

$$\text{Var}\,(X_t) = \frac{\Sigma\,(X_t - \bar{X})^2}{n} \ ,$$

یا

$$\Sigma\,(X_t - \bar{X})^2 = n\ \text{Var}\,(X_t) \ .$$

با جایگزینی در فرمول $\text{Var}\,(\hat{\beta})$ خواهیم داشت

$$\text{Var}\,(\hat{\beta}) = \frac{\hat{\sigma}^2}{n\ \text{Var}\,(X_t)} \ .$$

با حذف مشاهدات مرکزی X_t، با اینکه واریانس X_t زیاد و $\text{Var}\,(\hat{\beta})$ کم می‌شود، اما

اولاً، با کم شدن حجم نمونه و کاهش n، مقدار $\text{Var}\,(\hat{\beta})$ زیاد می‌شود ـ که با هدف اصلی مغایر است ـ

ثانیاً، با کم شدن n، ممکن است مقدار $\hat{\sigma}^2$ نیز افزایش یابد؛ زیرا

$$\hat{\sigma}^2 = \frac{\Sigma\, e_t^2}{n-2} \ ,$$

در نتیجه صورت کسر $\text{Var}\,(\hat{\beta})$ زیاد شده و مقدار $\text{Var}\,(\hat{\beta})$ افزایش نشان خواهد داد که این نیز خلاف هدف اولیه است ـ البته با کاهش n، باید به جهت حرکت $\Sigma\, e_t^2$ نیز توجه داشت؛ اما معمولاً به ازای کاهش n انتظار می‌رود که واریانس U_t زیاد شود.

ثالثاً، با حذف مشاهدات مرکزی X_t، نمونهٔ ما، دیگر یک نمونهٔ تصادفی نخواهد بود و این امر زیربنای تمام محاسبات ما را متزلزل می‌کند؛ زیرا می‌دانیم از فرضهای اولیهٔ تخمین‌زننده‌های روش حداقل مربعات معمولی تصادفی بودن نمونهٔ مشاهدات X_t و Y_t است.

۷ـ۲ الف) می‌دانیم $e_t = Y_t - \hat{\beta}\, X_t$. به جای Y_t و $\hat{\beta}$ مقادیر آن را قرار می‌دهیم.

$$e_t = \beta X_t + U_t - \frac{X_t \Sigma X_t Y_t}{\Sigma X_t^2} ,$$

$$= U_t + \beta X_t - \frac{X_t \Sigma (\beta X_t + U_t) X_t}{\Sigma X_t^2} ,$$

$$= U_t - \frac{X_t \Sigma U_t X_t}{\Sigma X_t^2} . \qquad (۱)$$

$$E(e_t^2) = E\left[U_t^2 + \frac{X_t^2 (\Sigma U_t X_t)^2}{(\Sigma X_t^2)^2} - \frac{2 X_t U_t \Sigma U_t X_t}{\Sigma X_t^2}\right] ,$$

$$= \sigma^2 - \frac{\sigma^2 X_t^2}{\Sigma X_t^2} = \sigma^2 \left(1 - \frac{X_t^2}{\Sigma X_t^2}\right) . \qquad (۲)$$

ب) می‌دانیم، $e_t = U_t - \frac{X_t \Sigma U_t X_t}{\Sigma X_t^2}$. به همین ترتیب خواهیم داشت

$$e_s = U_s - \frac{X_s \Sigma U_t X_t}{\Sigma X_t^2} .$$

در نتیجه داریم

$$E(e_t e_s) = E\left[U_t U_s - \frac{X_t U_t \Sigma X_t U_t}{\Sigma X_t^2} - \frac{X_s U_t \Sigma X_t U_t}{\Sigma X_t^2} + \frac{X_t X_s (\Sigma U_t X_t)^2}{(\Sigma X_t^2)^2}\right] .$$

می‌دانیم بنا بر فرضهای کلاسیک، $E(U_t U_s) = 0$. امید ریاضی سه جملهٔ دیگر را باید حساب کنیم؛ برای مثال، امید ریاضی یکی از آنها را نشان می‌دهیم،

$$E\left[-\frac{X_t U_s \Sigma X_t U_t}{\Sigma X_t^2}\right] = -\frac{X_t E[U_s \Sigma X_t U_t]}{\Sigma X_t^2} ,$$

$$= \frac{-X_t E[U_s (X_1 U_1 + X_2 U_2 + \cdots + X_s U_s + \cdots + X_n U_n)]}{\Sigma X_t^2} ,$$

$$= \frac{-Xt\left[X_1 E(U_1 U_s) + X_2 E(U_2 U_s) + \cdots + X_s E(U_s)^2 + \cdots + X_n E(U_n U_s)\right]}{\Sigma X_t^2},$$

$$= \frac{-X_t\left[0 + 0 + \cdots + X_s \sigma^2 + 0 + \cdots + 0\right]}{\Sigma X_t^2},$$

$$= \frac{-X_t X_s \sigma^2}{\Sigma X_t^2}.$$

به همین ترتیب برای دو جملهٔ دیگر عمل می‌کنیم. خواهیم داشت

$$E(e_t e_s) = 0 - \frac{X_t X_s \sigma^2}{\Sigma X_t^2} - \frac{X_t X_s \sigma^2}{\Sigma X_t^2} + \frac{X_t X_s \sigma^2}{\Sigma X_t^2}.$$

در نتیجه

$$E(e_t e_s) = -\frac{\sigma^2 X_t X_s}{\Sigma X_t^2}, \qquad t \neq s.$$

ج) برای نشان دادن $\Sigma e_t \neq 0$ می‌نویسیم

$$\Sigma e_t = \Sigma (Y_t - \hat{\beta} X_t),$$

$$= \Sigma \left[Y_t - \frac{X_t \Sigma X_t Y_t}{\Sigma X_t^2}\right].$$

هیچ دلیلی وجود ندارد که رابطهٔ فوق در حالت کلی برابر صفر باشد. اما در مدلهایی که ضریب ثابت دارند مقدار Σe_t برابر است با

$$\Sigma e_t = \Sigma (Y_t - \hat{\alpha} - \hat{\beta} X_t).$$

طرف راست رابطهٔ فوق در واقع معادلهٔ اول نرمال است که حاصل مشتق‌گیری مجموع مربعات پسماند نسبت به $\hat{\alpha}$ است و می‌دانیم که باید صفر باشد تا $\hat{\alpha}$ بتواند مجموع مربعات پسماند را حداقل کند. بنابراین، چون چنین شرطی در مدلهای رگرسیون بدون ضریب ثابت موجود نیست، در حالت کلی هیچ دلیلی بر صفر بودن Σe_t در این گونه

مدلها وجود ندارد.

البته می‌توان نشان داد که $E(\Sigma e_t)$ در مدلهایی که ضریب ثابت ندارند همواره برابر صفر است. از دو طرف رابطهٔ (۱)، Σ می‌گیریم،

$$\sum e_t = \sum \left[U_t - \frac{X_t \Sigma X_t U_t}{\Sigma X_t^2} \right] ,$$

و با گرفتن امید ریاضی خواهیم داشت

$$E[\sum e_t] = \sum \left[E(U_t) - \frac{X_t \Sigma E(X_t U_t)}{\Sigma X_t^2} \right] ,$$

$$= 0 .$$

د) رابطهٔ (۲) را یک بار دیگر می‌نویسیم،

$$E(e_t^2) = \sigma^2 - \frac{\sigma^2 X_t^2}{\Sigma X_t^2} .$$

بنابراین

$$E\left(\frac{\Sigma e_t^2}{n-1}\right) = \sum \left(\sigma^2 - \frac{\sigma^2 X_t^2}{\Sigma X_t^2}\right) / (n-1) ,$$

$$= \frac{n\sigma^2 - \sigma^2}{n-1} = \sigma^2 .$$

۲-۸ الف) چون $\tilde{\beta}$ یک تخمین‌زنندهٔ خطی است، داریم

$$\tilde{\beta} = w_1 Y_1 + w_2 Y_2 + \cdots + w_n Y_n ,$$

$$\tilde{\beta} = \sum w_t Y_t . \qquad (1)$$

در اینجا باید شرایطی را پیدا کنیم که $\tilde{\beta}$ ناااریب شود. از رابطهٔ (۱) امید ریاضی می‌گیریم.

$$E(\tilde{\beta}) = E[\sum w_t Y_t] ,$$

$$= E[\sum w_t (\alpha + \beta X_t + U_t)] ,$$

$$= \alpha \sum w_t + \beta \sum w_t X_t + \circ \ .$$

برای $E(\tilde{\beta}) = \beta$، ضروری است شرایط زیر صادق باشد،

$$\sum w_t = \circ \ ,$$
$$\sum w_t X_t = 1 \ . \qquad (۲)$$

اگر متغیری مانند Z_t داشته باشیم، به گونه‌ای که $z_t = Z_t - \bar{Z}$، سپس w_t را به صورت زیر تعریف کنیم

$$w_t = \frac{z_t}{\sum z_t x_t} \ ,$$

آنگاه خواهیم داشت

$$w_t = \frac{z_t}{\sum z_t X_t} \ . \qquad (۳)$$

واضح است که $\sum w_t = \circ$ و $\sum w_t X_t = 1$. بنابراین w_t که در رابطهٔ (۳) تعریف شده است شرایط (۲) را تأمین می‌کند. اگر w_t را در رابطهٔ (۱) قرار دهیم، تخمین‌زننده‌های خطی نااریب به دست می‌آیند،

$$\tilde{\beta} = \frac{\sum z_t Y_t}{\sum z_t X_t} \ , \qquad (۴)$$

یا

$$\tilde{\beta} = \frac{\sum z_t y_t}{\sum z_t x_t} \ . \qquad (۵)$$

ب) از دو طرف رابطهٔ (۱) واریانس می‌گیریم،

$$\mathrm{Var}(\tilde{\beta}) = \mathrm{Var}(w_1 Y_1 + w_2 Y_2 + \cdots + w_n Y_n) \ .$$

با توجه به اینکه $\mathrm{Var}(U_t) = \sigma^2$ و بنا بر فرض $\mathrm{Var}(Y_t) = \mathrm{Var}(U_t)$ ، $\mathrm{Cov}(U_t, U_s) = \circ$ ،

بنابراین:

$$\text{Var}(\tilde{\beta}) = w_1^2\sigma^2 + w_2^2\sigma^2 + \cdots + w_n^2\sigma^2 ,$$

$$= \sigma^2 \sum w_i^2 .$$

با جایگزینی (۳) در رابطهٔ فوق خواهیم داشت

$$\text{Var}(\tilde{\beta}) = \frac{\sigma^2 \Sigma z_i^2}{(\Sigma z_i x_i)^2} . \qquad (۶)$$

می‌دانیم واریانس $\hat{\beta}_{OLS}$ برابر است با

$$\text{Var}(\hat{\beta}) = \frac{\sigma^2}{\Sigma x_i^2} . \qquad (۷)$$

ج) برای به دست آوردن کارآیی نسبی، کافی است رابطهٔ (۷) را بر (۶) تقسیم کنیم،

$$\frac{\text{Var}(\hat{\beta})}{\text{Var}(\tilde{\beta})} = \frac{(\Sigma z_i x_i)^2}{\Sigma x_i^2 . \Sigma z_i^2} ,$$

که با توجه به نامساوی، کوچی ـ شوارتز بین صفر و یک است. ملاحظه می‌شود به ازای $z_i = k x_i$، داریم

$$\frac{\text{Var}(\hat{\beta})}{\text{Var}(\tilde{\beta})} = \frac{(\Sigma k x_i^2)^2}{\Sigma x_i^2 . k^2 \Sigma x_i^2} = \frac{k^2 (\Sigma x_i^2)^2}{k^2 (\Sigma x_i^2)^2} = 1 .$$

فصل سوم

پیش‌بینی و مباحث تکمیلی در مدل رگرسیون خطی ساده

۳ـ۱ مقدمه

در فصلهای اول و دوم دیدیم که با تخمین یک مدل رگرسیون می‌توان به کمک تغییرات متغیر برون‌زا، تغییرات متغیر درون‌زا را توضیح داد. یکی از مهمترین موارد کاربرد تخمین مدلهای رگرسیون این است که بتوانیم مقادیر آیندهٔ متغیر درون‌زا را پیش‌بینی کنیم. همان‌گونه که خواهیم دید، مقدار پیش‌بینی شدهٔ متغیر درون‌زا، یک متغیر تصادفی است و باید تابع توزیع احتمال با میانگین و واریانس آن دقیقاً مطالعه شود. با داشتن تابع توزیع احتمال و خصوصیات آن می‌توان فرضیه‌های مختلف را در مورد مقادیر پیش‌بینی شده متغیر درون‌زا به راحتی آزمون کرد. این نکات موضوع قسمت ۳ـ۲ خواهد بود.

هدف ما، در فصلهای اول و دوم، ارائه مفاهیم کلیدی در تخمین یک مدل رگرسیون خطی ساده و آزمون فرضیه‌های مختلف در مورد پارامترهای آن بوده است، بنابراین بسیاری از مباحث جانبی، ناگفته ماند. طرح این نکته‌های حاشیه‌ای در متن مباحث اولیه، ممکن بود توجه خواننده را از مسیر اصلی ارائه مطالب منحرف سازد. به نظر رسید که بهتر است این موضوعها در فصلی مجزا مطرح شود. در این فصل بعد از پایان مبحث پیش‌بینی، به ذکر چند عنوان در این مورد می‌پردازیم.

توجه به مسأله مقیاس اندازه‌گیری متغیرهای برون‌زا و درون‌زا اهمیت فراوان دارد. سؤال این است: اگر مقیاس اندازه‌گیری یک متغیر را تغییر دهیم، آیا تخمین پارامترهای یک مدل رگرسیون نیز تغییر خواهد کرد؟ این نکته‌ای است که در قسمت ۳ـ۳ به آن می‌پردازیم.

مطالعات اجمالی در رگرسیون معکوس، موضوع قسمت ۳ـ٤ خواهد بود. تا اینجا

ما، در مدلهای رگرسیون، Y_t را تابعی از تغییرات X_t فرض می‌کردیم. آیا می‌توان جهت تأثیرپذیری متغیرها را عوض کرده، این بار X_t را تابعی از تغییرات Y_t بگیریم؟ اگر جواب، مثبت باشد، سؤال این است که چه رابطه‌ای بین تخمین پارامترهای دو مدل وجود دارد.

در مباحث دو فصل گذشته، بویژه در بعضی از مسائل پایانی هر فصل، احتمال دارد خواننده با این نکته مواجه شده باشد که بعضی از مقادیر «دور افتاده» در متغیرهای برون‌زا و درون‌زا میتواند بر تخمین پارامترهای یک مدل رگرسیون تأثیر بسیار قابل ملاحظه‌ای داشته باشد. آیا می‌توان این متغیرهای دور افتاده را حذف کرد؟ تأثیر این متغیرها دقیقاً به چه صورت است؟ این نکته‌ها را به طور مختصر در قسمت ۳ـ۵ بررسی خواهیم کرد.

مدل رگرسیون ساده خطی که در فصلهای گذشته مطالعه شد در واقع مدلی است که هم برحسب متغیرها و هم برحسب پارامترها خطی است. در بعضی از مدلهای رگرسیون، متغیرها به صورت غیرخطی وارد می‌شوند؛ و در بعضی دیگر، پارامترها صورت غیرخطی دارند. آیا می‌توان با تبدیلهای مناسب، مدلهای غیرخطی برحسب متغیرها و مدلهای غیرخطی برحسب پارامترها را به مدلهای خطی تبدیل کرد؟ این سؤالی است که در قسمت ۳ـ۶ مطرح خواهد شد.

۳ـ۲ پیش‌بینی در مدلهای رگرسیون ساده

یکی از هدفهای اساسی در تخمین یک مدل رگرسیون، این است که بتوان تغییرات متغیر درون‌زا را به ازای مقدار معینی از متغیر برون‌زا پیش‌بینی[1] کرد. مدل رگرسیون ساده خطی زیر را ملاحظه کنید.

$$Y_t = \alpha + \beta X_t + U_t \quad . \qquad (۳ـ۱)$$

فرض کنید می‌خواهیم برای سال f در آینده پیش‌بینی کنیم. ابتدا باید مقدار متغیر برون‌زا

1. Prediction or Forecasting

در سال f را بدانیم. مقداری را که متغیر X در سال f خواهد داشت، با X_f نشان می‌دهیم. پیش‌بینی مقدار متغیر درون‌زا را به ازای X_f به دو صورت می‌تواند انجام شود: پیش‌بینی نقطه‌ای[1] و پیش‌بینی فاصله‌ای[2].

پیش‌بینی نقطه‌ای در واقع تخمین یک نقطه از فضای Y در آینده است که به ازای مقدار معین X_f صورت می‌پذیرد. اگر دورهٔ زمانی مشاهدات را سالانه فرض کنیم، مدل رگرسیون مفروض برای سال f عبارت است از

$$Y_f = \alpha + \beta X_f + U_f \quad . \qquad (۳\text{-}۲)$$

ملاحظه می‌شود فرض بر عدم تغییر ساختاری در مدل رگرسیون است؛ به عبارت دیگر، پارامترها و شکل ریاضی مدل رگرسیون مفروض ـ در آینده‌ای که می‌خواهیم برای آن پیش‌بینی کنیم ـ ثابت مانده است. منظور از پیش‌بینی نقطه‌ای محاسبه مقدار Y_f به ازای X_f است.

تخمین رگرسیون (۳ـ۱) عبارت است از

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t \quad . \qquad (۳\text{-}۳)$$

مدل ۳ـ۳ را برای سال f در آینده می‌نویسیم،

$$\hat{Y}_f = \hat{\alpha} + \hat{\beta} X_f \quad . \qquad (۳\text{-}٤)$$

به ازای X_f، مقدار $\hat{Y}_f$ به منزلهٔ پیش‌بینی نقطه‌ای از متغیر درون‌زا به راحتی از معادلهٔ ۳ـ٤ به دست می‌آید. این پیش‌بینی قاعدتاً یک پیش‌بینی شرطی است، بدین معنی که تحقق $\hat{Y}_f$ اولاً، مشروط به وقوع X_f است، و ثانیاً، فرض بر این است که ترکیب و ساختار رابطهٔ بین X و Y در آینده، (تا دوره‌ای که می‌خواهیم پیش‌بینی کنیم) کوچکترین تغییری نکرده است و در واقع مقادیر α و β و شکل ریاضی مدل مفروض، ثابت خواهد بود.

---

1. Point Prediction 2. Interval Prediction

۱. پیش‌بینی فاصله‌ای برای Y_f

فرض کنید پیش‌بینی نقطه‌ای ($\hat{Y}_f$) را برای یک دورهٔ زمانی در آینده به دست آورده‌ایم. بدیهی است در سال f، متغیر درون‌زا مقدار معینی خواهد داشت که آن را با Y_f نشان می‌دهیم. اینکه Y_f چه مقداری دارد، تابعی است از X_f، مقادیر α و β و مهمتر از همه مقداری که جملهٔ اختلال در سال f، یعنی U_f، خواهد داشت. بجز X_f، از عوامل دیگر هیچ اطلاعی نداریم. یعنی α و β و U_f برای ما مجهول هستند. تنها چیزی که از Y_f می‌دانیم پیش‌بینی نقطه‌ای آن، یعنی $\hat{Y}_f$، است. سؤال این است که آیا می‌توان با داشتن $\hat{Y}_f$ برای مقدار واقعی Y_f یک فاصلهٔ اطمینان ساخت؟ به عبارت دیگر، آیا می‌توان به کمک $\hat{Y}_f$ مثلاً در سطح احتمال ۹۵ درصد گفت که Y_f بین دو مقدار معلوم a و b قرار خواهد گرفت؟ پیش‌بینی فاصله‌ای دقیقاً همین مسأله است؛ یعنی ساختن فاصله‌ای که با احتمال معیّنی بتواند مقدار متغیر درون‌زا را در آینده مورد نظر شامل شود.

برای ساختن این فاصله اطمینان، ابتدا می‌گوییم $\hat{Y}_f$ یک متغیر تصادفی است؛ زیرا تابعی از $\hat{\alpha}$ و $\hat{\beta}$ است که هر دو تصادفی هستند. در سال f، مقداری که Y_f می‌گیرد نیز یک متغیر تصادفی خواهد بود؛ زیرا تابعی از U_f تصادفی است. اما قاعدتاً انتظار این است که مقدار پیش‌بینی ما از متغیردرون‌زا در آینده ($\hat{Y}_f$) دقیقاً با مقداری که واقعاً این متغیر در آینده خواهد داشت (Y_f) برابر نباشد. تفاوت یا انحراف پیش‌بینی نقطه‌ای $\hat{Y}_f$ از مقدار واقعی متغیردرون‌زا (Y_f) را اصطلاحاً خطای پیش‌بینی[1] می‌نامیم و آن را با e_f نشان می‌دهیم،

$$e_f = Y_f - \hat{Y}_f \quad . \qquad (۳\text{-}۵)$$

مقادیر Y_f و $\hat{Y}_f$ را از معادله‌های ۳-۲ و ۳-۳ در معادلهٔ ۳-۵ جایگزین می‌کنیم،

$$e_f = (\alpha + \beta X_f + U_f) - (\hat{\alpha} + \hat{\beta}) X_f \quad ,$$

یا

$$e_f = U_f - (\hat{\alpha} - \alpha) - (\hat{\beta} - \beta) X_f \quad . \qquad (۳\text{-}۶)$$

---

1. Prediction Error

در سمت راست معادلهٔ ۳ـ۶ متغیرهای $\hat{\alpha}$، $\hat{\beta}$ و U_i متغیرهای تصادفی هستند، بنابراین e_i نیز یک متغیر تصادفی خواهد بود. می‌دانیم توابع توزیع احتمال $\hat{\alpha}$، $\hat{\beta}$ و U_i نرمال هستند؛ در نتیجه e_i نیز ـ که یک تابع خطی از آنهاست ـ توزیع احتمال نرمال خواهد داشت. در اینجا باید میانگین و واریانس خطای پیش‌بینی را به دست آوریم.

از معادلهٔ ۳ـ۶ امید ریاضی می‌گیریم،

$$E(e_i) = E(U_i) - E(\hat{\alpha} - \alpha) - X_i E(\hat{\beta} - \beta) \quad .$$

با توجه به ناااریب بودن $\hat{\alpha}$، $\hat{\beta}$ و فرض صفر بودن امید ریاضی U_i خواهیم داشت

$$E(e_i) = 0 \quad . \qquad (۳ـ۷)$$

بنابراین نخستین نتیجه‌ای که به دست آوردیم این است که میانگین خطای پیش‌بینی، صفر است. قبل از وارد شدن به بحث واریانس e_i، به یک اصطلاح، اشاره می‌کنیم. می‌دانیم $\hat{\alpha}$ یا $\hat{\beta}$ را تخمین‌زننده‌های روش حداقل مربعات معمولی از α، β می‌گوییم. به همین ترتیب می‌توان $\hat{Y}_i$ را پیش‌بینی کنندهٔ[1] روش حداقل مربعات معمولی از Y_i نامید.

برای به دست آوردن واریانس e_i، از تعریف آن شروع می‌کنیم،

$$\text{Var}(e_i) = E[e_i - E(e_i)]^2 \quad .$$

با استفاده از معادلهٔ ۳ـ۷ داریم

$$\text{Var}(e_i) = E(e_i)^2 \quad .$$

دو طرف معادلهٔ ۳ـ۶ را مجذور کرده خواهیم داشت

$$e_i^2 = U_i^2 + (\hat{\alpha} - \alpha)^2 + (\hat{\beta} - \beta)^2 X_i^2 - 2 U_i (\hat{\alpha} - \alpha)$$
$$- 2 U_i (\hat{\beta} - \beta) X_i + 2 (\hat{\alpha} - \alpha)(\hat{\beta} - \beta) X_i \quad .$$

1. Predictor

از معادله فوق امید ریاضی می‌گیریم،

$$\begin{aligned}\mathrm{Var}\,(e_f) = {} & \mathrm{Var}\,(U_f) + \mathrm{Var}\,(\hat{\alpha}) + X_f^2\,\mathrm{Var}\,(\hat{\beta}) \\ & - 2\,\mathrm{Cov}\,[U_f, (\hat{\alpha} - \alpha)] - 2\,\mathrm{Cov}\,[U_f, (\hat{\beta} - \beta)\,X_f] \\ & + 2\,X_f\,\mathrm{Cov}\,(\hat{\alpha}, \hat{\beta})\ .\end{aligned}$$

بنا بر فرض واریانس همسانی، می‌دانیم: $\mathrm{Var}\,(U_f) = \mathrm{Var}\,(U_t) = \sigma^2$. می‌توان مقادیر واریانس $\hat{\alpha}$ و واریانس $\hat{\beta}$ و کوواریانس $\hat{\alpha}$ و $\hat{\beta}$ را به ترتیب از فرمولهای ۲ـ۱۶ ، ۲ـ۲۱ و ۲ـ۲۵ به دست آورد. می‌توان دربارهٔ $\mathrm{Cov}\,[U_f\ ,\ (\hat{\beta} - \beta)\ X_f]$ ، توجه را فقط به کوواریانس U_t و $\hat{\beta}$ محدود کرد، زیرا β و X_t مقادیر ثابتی هستند. از معادلهٔ ۲ـ۱۰ می‌دانیم که $\hat{\beta}$ ، یک تابع خطی از مقادیر $U_1, U_2, \ldots, U_n$ است و چون بنا بر فرض عدم خودهمبستگی، این مقادیر از U_f مستقل هستند؛ بنابراین $\hat{\beta}$ نیز از U_f مستقل بوده و کوواریانس آنها صفر می‌شود:

$$\mathrm{Cov}\,[U_f, (\hat{\beta} - \beta)\,X_f] = 0\ .$$

به همین ترتیب می‌توان نشان داد که

$$\mathrm{Cov}\,[U_f, (\hat{\alpha} - \alpha)] = 0\ .$$

بنابراین خواهیم داشت

$$\begin{aligned}\mathrm{Var}\,(e_f) &= \sigma^2 + \sigma^2\left[\frac{1}{n} + \frac{\bar{X}^2}{\Sigma x_t^2}\right] + \frac{X_f^2\sigma^2}{\Sigma x_t^2} - \frac{2\,X_f\,\bar{X}\,\sigma^2}{\Sigma x_t^2}\ , \\ &= \sigma^2\left[1 + \frac{1}{n} + \frac{\bar{X}^2}{\Sigma x_t^2} + \frac{X_f^2}{\Sigma x_t^2} - \frac{2\,X_f\,\bar{X}}{\Sigma x_t^2}\right]\ ,\end{aligned}$$

در نتیجه

$$\mathrm{Var}\,(e_f) = \sigma^2\left[1 + \frac{1}{n} + \frac{(X_f - \bar{X})^2}{\Sigma x_t^2}\right]\ . \qquad (۳ـ۸)$$

با ملاحظهٔ معادلهٔ ۳ـ۶ دیدیم که e_t توزیع نرمال دارد. معادله‌های ۳ـ۷ و ۳ـ۸ به ترتیب میانگین و واریانس خطای پیش‌بینی را مشخص می‌کند. بدین ترتیب به راحتی می‌توان e_t را استاندارد کرد. کافی است e_t را از میانگین آن کم کرده و بر انحراف معیارش تقسیم کنیم، آنگاه توزیع Z با میانگین صفر و واریانس یک خواهد داشت،

$$\frac{e_f - E(e_f)}{\sqrt{Var(e_f)}} \sim Z(0,1) \ ,$$

$$\frac{e_f}{\sigma_U \sqrt{1 + \frac{1}{n} + \frac{(X_f - \bar{X})^2}{\Sigma x_t^2}}} \sim Z(0,1) \ .$$

اگر به جای σ_U مقدار تخمین آن را قرار دهیم، توزیع کسر فوق از Z به t تبدیل می‌شود:

$$\frac{e_f}{\hat{\sigma}_U \sqrt{1 + \frac{1}{n} + \frac{(X_f - \bar{X})^2}{\Sigma x_t^2}}} \sim t(0,1) \ , \qquad (۳ـ۹)$$

که در آن $\hat{\sigma}_U = \sqrt{\frac{\Sigma e_t^2}{(n-2)}}$.

سؤالی که در مقدمه مطرح کردیم، این بود که آیا می‌توان با داشتن پیش‌بینی نقطه‌ای ($\hat{Y}_f$) به فاصلهٔ اطمینانی برای مقدار واقعی Y_f رسید؟ تاکنون بحث ما بیشتر روی خطای پیش‌بینی و محاسبهٔ میانگین و واریانس آن متمرکز بوده است. با جایگزینی معادلهٔ ۳ـ۵ در معادلهٔ ۳ـ۹ به راحتی می‌توان به چنین فاصلهٔ اطمینانی دست یافت،

$$\frac{Y_f - \hat{Y}_f}{\hat{\sigma}_U \sqrt{1 + \frac{1}{n} + \frac{(X_f - \bar{X})^2}{\Sigma x_t^2}}} \sim t \ .$$

اگر در سطح معنی‌دار α، مقدار t را با $(n-2)$ درجه آزادی از جدول t به دست آوریم $(\pm t_{\alpha/2})$، دقیقاً می‌توان به روش گذشته فاصلهٔ اطمینان برای Y_f را به دست آورد. یادآوری می‌کنیم که در رابطهٔ فوق، بجز Y_f، تمام مقادیر دیگر را می‌دانیم. بدین ترتیب

فاصلهٔ اطمینان مطلوب عبارت است از

$$\hat{Y}_f - t_{\alpha/2}\hat{\sigma}_U\sqrt{1+\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\sum x_i^2}} < Y_f < \hat{Y}_f + t_{\alpha/2}\hat{\sigma}_U\sqrt{1+\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\sum x_i^2}}, \qquad (۳\text{-}۱۰)$$

یا به صورت ساده‌تر

$$\hat{Y}_f - t_{\alpha/2}\,SE(e_f) < Y_f < \hat{Y}_f + t_{\alpha/2}\,SE(e_f) \quad . \qquad (۳\text{-}۱۱)$$

مثال ۳-۱ با استفاده ازیک نمونه دوازده‌تایی، تابع‌مصرف را به‌صورت‌زیر تخمین زده‌ایم،

$$\hat{C}_t = ۱۰ + ۰/۹۰\,Y_t \ ,$$

که در آن C_t و Y_t به ترتیب مصرف و درآمد است (محاسبات کاملاً فرضی است). همچنین این کمیتها را داریم،

$$\hat{\sigma}^2 = ۰/۰۱ \quad , \quad \bar{Y} = ۲۰۰ \quad , \quad \sum (Y_t - \bar{Y}) = ۴۰۰۰ \ .$$

سطح درآمد در پنج سال دیگر $X_f = ۲۵۰$ فرض می‌شود.

اولاً، پیش‌بینی نقطه‌ای از سطح مصرف در آن سال ($\hat{C}_f$) را به دست آورید.

ثانیاً، یک فاصلهٔ اطمینان ۹۵ درصد برای مقدار واقعی مصرف در سال f پیش‌بینی کنید.

بهتر است علامت‌گذاریهای استفاده شده در مثال را با علامتهای به کار رفته در متن کتاب، هماهنگ کنیم،

$$\hat{Y}_t = ۱۰ + ۰/۹۰\,X_t \ ,$$

$$\hat{\sigma}^2 = ۰/۰۱ \quad , \quad \bar{X} = ۲۰۰ \quad , \quad \sum (X_t - \bar{X}) = ۴۰۰۰ \quad , \quad Y_f = ۲۵۰ \ .$$

اولاً،

$$\hat{Y}_f = \hat{\alpha} + \hat{\beta}\,X_f \ ,$$
$$= ۱۰ + ۰/۹۰\,(۲۵۰) = ۲۳۵ \ ,$$

یعنی مصرف در پنج سال دیگر به ازای $Y = ۲۵۰$ ، برابر ۲۳۵ خواهد بود.

ثانیاً، برای پیش‌بینی فاصلهٔ اطمینان ۹۵ درصد برای مصرف در سال f ، ابتدا واریانس خطای پیش‌بینی را به دست می‌آوریم. از معادلهٔ ۸ـ۳ داریم

$$\text{Var}(e_f) = \hat{\sigma}^2 \left[۱ + \frac{۱}{n} + \frac{(X_f - \bar{X})^2}{\sum x_i^2} \right] ,$$

$$= ۰/۰۱ \left[۱ + \frac{۱}{۱۲} + \frac{(۲۵۰ - ۲۰۰)^2}{۴۰۰۰} \right] = ۰/۰۱۷ .$$

برای رسیدن به انحراف معیار خطای پیش‌بینی از واریانس e_f جذر می‌گیریم،

$$\text{SE}(e_f) = \sqrt{۰/۰۱۷} = ۰/۱۳۱ .$$

برای محاسبهٔ فاصله اطمینان ۹۵ درصد باید مقدار t را با $۱۰ = ۲ - ۱۲$ درجهٔ آزادی و $\frac{\alpha}{۲} = ۰/۰۲۵$ از جدول t به دست آورد. خواهیم داشت $t = \pm ۲/۲۲۸$. با استفاده از نامساوی ۱۱ـ۳ فاصله اطمینان مورد نظر براحتی قابل محاسبه است،

$$۲۳۵ - ۲/۲۲۸ (۰/۱۳۱) < Y_f < ۲۳۵ + ۲/۲۲۸ (۰/۱۳۱) ,$$

$$۲۳۵ - ۰/۲۹ < Y_f < ۲۳۵ + ۰/۲۹ ,$$

یا

$$۲۳۴/۷۱ < Y_f < ۲۵۳/۲۹ .$$

در پایان این مثال به دو اصطلاح اشاره می‌کنیم. نمونهٔ مورد استفاده در این مثال شامل دوازده مشاهده است؛ بنابراین محدودهٔ تغییرات متغیر برون‌زا (X_i) کاملاً مشخص است. اگر X_f، یعنی مقداری از X ـ که می‌خواهیم به ازای آن پیش‌بینی کنیم ـ در محدودهٔ تغییرات X_i در نمونه قرار گیرد، آنگاه می‌گوییم «پیش‌بینی درون ـ نمونه‌ای»[1] داریم. در غیر این صورت پیش‌بینی ما «برون ـ نمونه‌ای»[2] خواهد بود؛ بنابراین اصطلاح پیش‌بینی در اقتصادسنجی، همواره ناظر به مقادیر آینده متغیرها نیست.

---

1. Within-Sample Prediction
2. Out-of-Sample Prediction

۲. پیش‌بینی فاصله‌ای برای $E(Y_f)$

در قسمت قبل، برای Y_f (یک نقطه در آینده)، فاصله اطمینان پیش‌بینی کردیم. در بسیاری از مواقع، آنچه مورد توجه ماست، میانگین یا امید ریاضی Y_f است؛ بنابراین می‌خواهیم برای میانگین یا حد متوسط Y_f یک فاصله اطمینان بسازیم. از تعریف میانگین یا امید ریاضی Y_f آغاز می‌کنیم. از معادلهٔ ۲ـ۳ امید ریاضی می‌گیریم،

$$E(Y_f) = \alpha + \beta X_f \quad . \qquad (۳ـ۱۲)$$

برای تخمین مقدار $E(Y_f)$ به ازای X_f، کافی است به جای α و β مقادیر $\hat{\alpha}$ و $\hat{\beta}$ را قرار دهیم،

$$\hat{E}(Y_f) = \hat{\alpha} + \hat{\beta} X_f \quad .$$

اما در معادلهٔ ٤ـ۳ ، دیدیم $\hat{Y}_f = \hat{\alpha} + \hat{\beta} X_f$ در نتیجه

$$\hat{E}(Y_f) = \hat{Y}_f \quad . \qquad (۳ـ۱۳)$$

بنابراین فرقی نمی‌کند که Y_f یا $E(Y_f)$ را پیش‌بینی کنیم، زیرا در هر دو حالت به یک جواب می‌رسیم. با اینکه پیش‌بینیهای نقطه‌ای برای Y_f و میانگین Y_f یکی هستند، فاصله‌های اطمینان این دو کاملاً از یکدیگر متفاوت است؛ زیرا همانگونه که خواهیم دید خطای پیش‌بینی و واریانس خطای پیش‌بینی آنها متساوی نیست.

اگر خطای پیش‌بینی را در این مورد با e_f^m نشان دهیم ـ که منعکس‌کنندهٔ خطای پیش‌بینی برای میانگین Y_f است ـ خواهیم داشت

$$e_f^m = E(Y_f) - \hat{E}(Y_f) \quad ,$$

یا

$$e_f^m = E(Y_f) - \hat{Y}_f \quad . \qquad (۳ـ۱٤)$$

با جایگزینی معادله‌های ۱۲ـ۳ و ٤ـ۳ در معادله (۱٤ـ۳) داریم

$$e_f^m = \alpha + \beta X_f - (\hat{\alpha} + \hat{\beta} X_f) \quad ,$$

در نتیجه

$$e_f^m = -(\hat{\alpha} - \alpha) - (\hat{\beta} - \beta) X_f \quad . \qquad (۳\text{-}۱۵)$$

با مقایسهٔ خطای پیش‌بینی در این حالت، یعنی معادلهٔ ۳ـ۱۵، با خطای پیش‌بینی در حالت اول، یعنی معادلهٔ ۳ـ۶، نتیجه می‌گیریم که خطاهای پیش‌بینی از یکدیگر متفاوت است.

در اینجا باید میانگین و واریانس خطای پیش‌بینی e_f^m را به دست آورد. از معادلهٔ ۳ـ۱۵ امید ریاضی می‌گیریم. خواهیم داشت

$$E(e_f^m) = 0 \quad . \qquad (۳\text{-}۱۶)$$

برای محاسبه واریانس e_f^m می‌گوییم که

$$\text{Var}(e_f^m) = E\left[e_f^m - E(e_f^m)\right]^2 = E(e_f^m)^2 \quad .$$

باید دو طرف معادلهٔ ۳ـ۱۵ را مجذور کنیم و امید ریاضی بگیریم،

$$\text{Var}(e_f^m) = E(e_f^2) = E\left[-(\hat{\alpha} - \alpha) - (\hat{\beta} - \beta) X_f\right]^2 \quad .$$

اگر به ترتیبی کاملاً مشابه با آنچه در مورد واریانس e_f انجام دادیم، عمل کنیم، خواهیم داشت

$$\text{Var}(e_f^m) = \sigma^2 \left[\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\Sigma x_i^2}\right] \quad . \qquad (۳\text{-}۱۷)$$

می‌توان نشان داد که e_f^m همانند e_f توزیعی نرمال با میانگین صفر و واریانس معلوم دارد که می‌توان آن را از معادلهٔ ۳ـ۱۷ به دست آورد. اگر به جای σ^2 مقدار تخمین آن ($\hat{\sigma}^2$) را قرار دهیم و e_f^m را استاندارد کنیم، توزیع t با $(n - 2)$ درجه آزادی خواهد داشت،

$$\frac{e_f^m - 0}{\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\Sigma x_i^2}}} \sim t(n - 2) \quad .$$

با استفاده از معادلهٔ ۱۴ـ۳ داریم

$$\frac{E(Y_f) - \hat{Y}_f}{\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum x_i^2}}} \sim t(n-2) \ .$$

با به دست آوردن مقدار t از جدول در سطح معنی‌دار α درصد، فاصلهٔ اطمینان $(1-\alpha)$ درصد برای $E(Y_f)$ به دست می‌آید،

$$\hat{Y}_f - t_{\alpha/2}\hat{\sigma}_U\sqrt{\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum x_i^2}} < E(Y_f) < \hat{Y}_f + t_{\alpha/2}\hat{\sigma}_U\sqrt{\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum x_i^2}} \ , \qquad (۳\text{ـ}۱۸)$$

یا به صورت ساده‌تر

$$\hat{Y}_f - t_{\alpha/2}\,SE(e_f) < E(Y_f) < \hat{Y}_f + t_{\alpha/2}\,SE(e_f) \qquad (۳\text{ـ}۱۹)$$

مثال ۲ـ۳ مدل رگرسیون زیر برای پیش‌بینی سطح فروش یک شرکت تجاری مفروض است،

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

که در آن Y_t درآمد حاصل از فروش و X_t هزینهٔ تبلیغات است. جدول ۱ـ۳ هزینهٔ تبلیغات و درآمد حاصل از فروش ماهانهٔ ۵ ماه را نشان می‌دهد.

جدول ۱ـ۳

t	درآمد حاصل از فروش Y_t (به صد هزار تومان)	هزینهٔ تبلیغات X_t (به ده‌هزار تومان)
۱	۳	۱
۲	٤	۲
۳	۲	۳
٤	٦	٤
٥	۸	٥

الف) مدل رگرسیون را تخمین بزنید.

ب) اگر در آینده، هزینهٔ تبلیغات ماهی ۶۰ هزار تومان باشد، میزان فروش چقدر خواهد بود.

ج) برای سطح فروش به دست آمده در بند «ب»، یک فاصلهٔ اطمینان ۹۰ درصدی بسازید.

د) اگر هزینهٔ تبلیغات ماهی ۶۰ هزار تومان را برای ۱۰ ماه ادامه دهیم، متوسط میزان فروش ماهانه چقدر خواهد بود.

ه) برای متوسط فروش ماهانه که در بند «د» به دست آورده‌اید فاصلهٔ اطمینان ۹۰ درصدی بسازید.

الف) بر اساس کمیتهای مندرج در جدول ۱-۳ می‌توان مدل رگرسیون مفروض را به صورت زیر تخمین زد،

$$\hat{Y}_i = ۱ + ۱/۲\, X_i$$

همچنین می‌دانیم

$$\bar{X} = ۳ \quad ، \quad RSS = \sum e_i^۲ = ۸/۸ \quad , \quad \sum x_i^۲ = ۱۰$$

ب) پیش‌بینی نقطه‌ای به ازای $X_f = ۶$ برابر است با

$$\hat{Y}_f = ۱ + ۱/۲\, X_f \ ,$$

$$= ۱ + ۱/۲\, (۶) = ۸/۲ \ ,$$

بنابراین به ازای ماهی ۶۰ هزار تومان هزینهٔ تبلیغات، پیش‌بینی می‌شود، فروش ماهانهٔ این شرکت برابر ۸۲۰ هزار تومان باشد.

ج) برای به دست آوردن فاصلهٔ اطمینان ۹۰ درصد برای Y_f، ابتدا انحراف معیار خطای پیش‌بینی را حساب می‌کنیم. براساس معادلهٔ ۳-۸ داریم:

$$\text{Var}\,(e_f) = \hat{\sigma}^۲ \left[۱ + \frac{۱}{n} + \frac{(X_f - \bar{X})^۲}{\sum x_i^۲} \right] \ ,$$

با توجه به

$$\hat{\sigma}^2 = \frac{\Sigma e_t^2}{n - 2} = \frac{۸/۸}{۵ - ۳} = ۲/۹۳ \ ,$$

خواهیم داشت

$$\text{Var}(e_f) = ۲/۹۳ \left[۱ + \frac{۱}{۵} + \frac{(۶ - ۳)^2}{۱۰} \right] = ۲/۹۳ \ (۲/۱) = ۶/۱۵۳$$

بنابراین انحراف معیار خطای پیش‌بینی عبارت است از

$$\text{SE}(e_f) = \sqrt{\text{Var}(e_f)} = \sqrt{۶/۱۵۳} = ۲/۴۸ \ .$$

با استفاده از نامساوی ۱۱-۳ و با توجه به اینکه مقدار t با ۳ درجه آزادی و سطح معنی‌دار ۱۰ درصد برابر با ۲/۳۵۳ است، فاصله اطمینان ۹۰ درصدی برای Y_f به صورت زیر خواهد بود،

$$\hat{Y}_f - t_{\alpha/۲} \text{SE}(e_f) < Y_f < \hat{Y}_f + t_{\alpha/۲} \text{SE}(e_f) \ ,$$

$$۸/۲ - ۲/۳۵۳ \ (۲/۴۸) < Y_f < ۸/۲ + ۲/۳۵۳ \ (۲/۴۸) \ ,$$

یا

$$۲/۳۶ < Y_f < ۱۴/۰۴ \ ,$$

یعنی به ازای ماهانه ۶۰ هزار تومان هزینهٔ تبلیغات، پیش‌بینی می‌شود که با ۹۰ درصد احتمال سطح فروش بین ۲۳۶ هزار تومان تا ۱ میلیون و ۴۰۴ هزار تومان نوسان داشته باشد.

د) در این حالت به جای Y_f باید $E(Y_f)$ را پیش‌بینی کرد. با توجه به معادلهٔ ۱۳-۳ می‌دانیم تخمین نقطه‌ای Y_f یا $E(Y_f)$ متساویند:

$$\hat{E}(Y_f) = \hat{\alpha} + \hat{\beta} X_f$$

$$= ۱ + ۱/۲ \ (۶) = ۸/۲$$

ه) همان گونه که در معادلهٔ ۱۷-۳ ملاحظه می‌شود واریانس خطای پیش‌بینی در

این حالت با حالت قبلی تفاوت دارد،

$$\text{Var}\,(e_f^m) = \hat{\sigma}^2 \left[\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\Sigma x_i^2}\right] .$$

با استفاده از کمیتهای داده شده خواهیم داشت

$$\text{Var}\,(e_f^m) = ۲/۹۳ \left[\frac{1}{۵} + \frac{(۶ - ۳)^2}{۱۰}\right] = ۲/۹۳\,(۱/۱) = ۳/۲۲۳ ,$$

بنابراین انحراف معیار خطای پیش‌بینی عبارت است از

$$\text{SE}\,(e_f^m) = \sqrt{\text{Var}\,(e_f)} = \sqrt{۳/۲۲۳} = ۱/۷۹۵ .$$

با استفاده از نامساوی ۳.۱۹ و با توجه به اینکه مقدار t با ۳ = ۲ − ۵ درجهٔ آزادی و سطح معنی‌دار ۱۰ درصد برابر با ۲/۳۵۳ است؛ بنابراین فاصلهٔ اطمینان ۹۰ درصدی برای $E\,(Y_f)$ به صورت زیر خواهد بود،

$$\hat{Y}_f - t_{\alpha/۲}\,\text{SE}\,(e_f^m) < E\,(Y_f) < \hat{Y}_f + t_{\alpha/۲}\,\text{SE}\,(e_f^m) ,$$

$$۸/۲ - ۲/۳۵۳\,(۱/۷۹۵) < E\,(Y_f) < ۸/۲ + ۲/۳۵۳\,(۱/۷۹۵) ,$$

$$۳/۹۸ < E\,(Y_f) < ۱۲/۴۲ .$$

یعنی به ازای ماهانه ۶۰ هزار تومان هزینهٔ تبلیغات و با این فرض که این هزینه تا ۱۰ ماه ادامه یابد، پیش‌بینی می‌شود که با ۹۰ درصد احتمال متوسط سطح فروش در این دوره ۱۰ ماهه بین ۳۹۸ هزار تومان تا ۱ میلیون و ۲۴۲ هزار تومان در نوسان باشد. ملاحظه می‌شود که فاصلهٔ اطمینان در این حالت از حالت قبلی کمتر است. دلیل این امر کوچکتر بودن واریانس e_f^m نسبت به e_f می‌باشد که به معنای دقیقتر بودن تخمین و درنتیجه کوچکتر شدن فاصله اطمینان است.

۳. دقت پیش‌بینی

در قسمتهای قبل دیدیم که می‌توان فاصله‌های اطمینان را برای Y_f و $E\,(Y_f)$ به ترتیب از

نامساویهای ۱۰ـ۳ و ۱۸ـ۳ به دست آورد. مطلوب این است که دقت پیش‌بینی حداکثر باشد. فاصله‌های اطمینان بزرگ نشان‌دهندهٔ دقتهای کم در پیش‌بینی است، برعکس هر چه دقت در پیش‌بینی بیشتر باشد، فاصله‌های اطمینان کوچکتر می‌شود. سؤال این است که در چه شرایطی می‌توان به حداکثر دقت در پیش‌بینی رسید؟

نامساویهای ۱۰ـ۳ و ۱۸ـ۳ را یک بار دیگر می‌نویسیم،

$$\hat{Y}_f - t_{\alpha/2}\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\sum x_i^2}} < Y_f < \hat{Y}_f + t_{\alpha/2}\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\sum x_i^2}} ,$$

یا به طور خلاصه

$$\hat{Y}_f - t\,SE(e_f) < Y_f < \hat{Y}_f + \hat{Y}_f - t\,SE(e_f) .$$

و برای $E(Y_f)$ داریم

$$\hat{Y}_f - t_{\alpha/2}\hat{\sigma}\sqrt{\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\sum x_i^2}} < E(Y_f) < \hat{Y}_f + t_{\alpha/2}\hat{\sigma}\sqrt{\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\sum x_i^2}} ,$$

یا به طور خلاصه

$$\hat{Y}_f - t\,SE(e_f^m) < E(Y_f) < \hat{Y}_f + t\,SE(e_f^m) .$$

ملاحظه می‌شود مهمترین عاملی که می‌تواند فاصله‌های اطمینان برای Y_f و $E(Y_f)$ را حداقل کند، این است که انحراف معیار e_f یا e_f^m حداقل شود. سؤال این است که چگونه می‌توان این انحراف معیارها را حداقل کرد.

بدیهی است با افزایش حجم نمونه n، مقدار $\frac{1}{n}$ کاهش می‌یابد و در نتیجه $SE(e_f)$ و $SE(e_f^m)$ نیز کمتر می‌شود. با وجود این افزایش حجم نمونه، حدی دارد و معمولاً محدودیتهای آماری به ما اجازه نمی‌دهد که n را به دلخواه و در ابعاد بزرگ افزایش دهیم؛ بنابراین باید توجه خود را به کسر $\frac{(X_f-\bar{X})^2}{\sum x_i^2}$ معطوف کنیم. برای رسیدن به مقدار حداقل این کسر، صورت و مخرج آن باید به ترتیب حداکثر و حداقل شود. مخرج کسر در واقع پراکندگی مشاهدات X_i را اندازه گیری می‌کند. قاعدتاً نمی‌توان

به طور دقیق شرایطی را تعریف کرد که این پراکندگی حداکثر شود. تنها کاری که می‌توان انجام داد این است که با نمونه گیریهای صحیح و تصادفی به سمت پراکندگیهای بیشتر در X_i میل کنیم. می‌دانیم افزایش پراکندگی X_i به همراه افزایش حجم نمونه، شرایط عمومی افزایش دقت در تخمین پارامترهاست که قبلاً در معادلهٔ ۲ـ۴۲ به طور خلاصه مطرح شد. با مدلی که دقیقتر تخمین زده شده است می‌توان بهتر و دقیقتر پیش‌بینی کرد.

حداقل کردن صورت کسر اهمیت فراوان دارد با توجه به این شرط میتوان گفت که به ازای مقادیری از X_i که به $\bar{X}$ نزدیکتر باشد، فاصلهٔ اطمینان از Y_i یا $E(Y_i)$ کوچکتر خواهد بود. می‌دانیم پراکندگی X_i با توجه به مقدار میانگین آن، یعنی $\bar{X}$، مشخص می‌شود. بنابراین اگر پیش‌بینی به ازای مقادیری از متغیر برون‌زا صورت گیرد که در محدودهٔ تغییرات X_i قرار دارد، مکانیسم پیش‌بینی در محدوده‌ای انجام خواهد شد که دقت تخمینها بسیار بالاست، و در نتیجه پیش‌بینی ما نیز از دقت خوبی برخوردار خواهد بود، یعنی واریانس کمتری خواهد داشت. هر چه X_i به سمت $\bar{X}$ میل کند، بیشتر در محورهای مرکزی تغییرات X_i قرار می‌گیرد و پیش‌بینی ما دارای واریانس کمتر و دقت بیشتر خواهد بود. آوردن مثالی در این مورد مفید است.

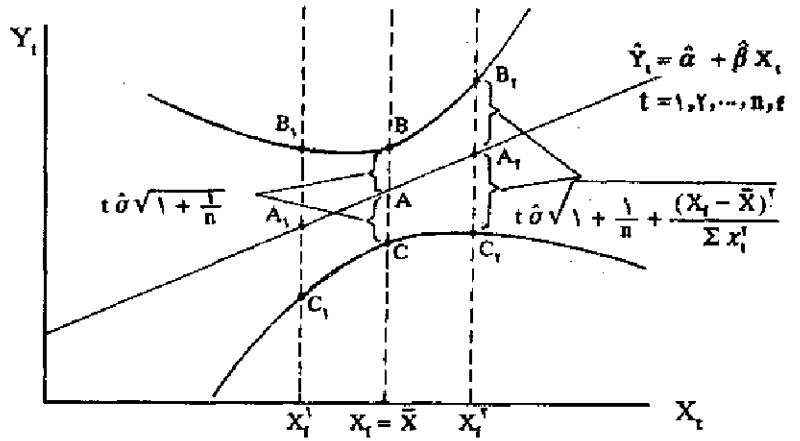
فرض کنید حجم واردات را تابعی از سطح صادرات مواد نفتی گرفته و می‌خواهیم یک مدل رگرسیون بر این مبنا تخمین بزنیم. یک نمونه ۱۰ تایی از متغیر برون‌زا داریم که صادرات نفت را طی ۱۰ سال گذشته نشان می‌دهد. یادآوری می‌کنیم که این ۱۰ عدد و درجهٔ پراکندگی آنها (واریانس X_i)، از مهمترین عوامل تخمین پارامترهای مدل است؛ زیرا به کمک تغییرات X_i است که می‌خواهیم تغییرات Y_i را توضیح دهیم. فرض کنید میانگین صادرات نفت $\bar{X}$ برابر ۲۰ میلیارد دلار است، سؤال این است که اگر در صادرات آیندهٔ نفت، ـ مثلاً در پنجمین سال ـ X_i برابر ۲۰۰ میلیارد دلار باشد، فاصلهٔ اطمینان ۹۵ درصد برای حجم واردات در آن سال چقدر خواهد بود؟

بدیهی است با چنین مقدار از X_i ـ که بسیار دور از $\bar{X}$ است ـ مقدار خطای پیش‌بینی دارای واریانسی بسیار بزرگ بوده و پیش‌بینیهای ما از دقت بسیار کمی

برخوردار خواهد بود. وقتی می‌گوییم X_t بسیار دور از $\bar{X}$ است، در واقع می‌خواهیم بگوییم که X_t از واقعیت بسیار دور است. توجه داریم که «دور از واقعیت» بدین معنی نیست که بعد از ۵ سال صادرات نفت مطلقاً نمی‌تواند به مرز ۲۰۰ میلیارد دلار برسد. ممکن است این طور باشد یا نباشد. به هر حال مدل رگرسیون مفروض برای پیش‌بینی صادرات مواد نفتی طراحی نشده است. منظور ما از واقعیتها، در واقع اوضاع حاکم بر ساختار مدلی است که بر اساس مشاهدات گذشته، ساختار مدلی را تخمین می‌زند که می‌خواهد برای آینده پیش‌بینی کند. بنابراین هر چه مقدار X_t به $\bar{X}$ نزدیکتر شود، پیش‌بینی مدل با شرایط حاکم بر ساختار مدل هماهنگی بیشتری خواهد داشت. به ازای $X_t = \bar{X}$، مقدار پیش‌بینی دقیقاً در محور تغییرات X_t صورت می‌گیرد و بهترین مقادیر پیش‌بینی را نتیجه خواهد داد.

از نظر ریاضی، وقتی X_t با $\bar{X}$ برابر شود، صورت کسر صفر شده و مقدار حداقل کسر حاصل می‌شود. با حذف این کسر، مقدار $SE(e_t)$ یا $SE(e_t^m)$ حداقل می‌شود و فاصله‌های اطمینان برای Y_t و $E(Y_t)$ نیز حداقل خواهد شد. البته در عمل هیچگاه کوشش ما این نیست که این فاصله‌ها حداقل شود؛ زیرا ما نمی‌توانیم X_t را به دلخواه انتخاب کنیم. مقدار X_t به عنوان متغیر برون‌زا در خارج از سیستم تعیین می‌شود و ما فقط می‌توانیم مقادیر Y_t یا $E(Y_t)$ را به ازای آن پیش‌بینی کنیم. معیار $X_t = \bar{X}$ تنها شاخصی است که با توجه به آن می‌توان، قبل از محاسبه فاصله‌های اطمینان، گفت که آیا پیش‌بینیهای خوبی خواهیم داشت یا خیر. هر چقدر X_t به $\bar{X}$ نزدیکتر باشد، فاصله‌های اطمینان برای Y_t و $E(Y_t)$ کوچکتر و دقت پیش‌بینی بیشتر خواهد شد و برعکس نمودار (۱ـ۳).

تخمین مدل رگرسیون $(\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t)$ را در روی صفحهٔ مختصات رسم می‌کنیم. t می‌تواند مقدار f را نیز بگیرد. ابتدا نقطهٔ $\bar{X}$ را در روی محور X_t مشخص نموده و می‌گوییم اگر بخواهیم به ازای $X_t = \bar{X}$ برای Y_t پیش‌بینی کنیم، پیش‌بینی نقطه‌ای ما از Y_t، یعنی $\hat{Y}_t$، برابر مقدار $A\bar{X}$ خواهد بود. برای رسیدن به پیش‌بینی فاصله‌ای، باید مقدار $t\,SE(e_t)$ را یک بار به $A\bar{X}$ اضافه و بار دیگر از آن کم کنیم تا به ترتیب نقاط B و



نمودار ۳.۱ فاصله‌های اطمینان برای Y_t

C به دست آید. بنابراین پیش‌بینی فاصله‌ای برای Y_t به ازای $X_t = \bar{X}$ برابر است با

$$A\bar{X} - AC < Y_t < A\bar{X} + AB \quad ,$$

که در آن $AB = AC$. چون در نقطهٔ $X_t = \bar{X}$ مقدار $\frac{(X_t - \bar{X})^۲}{\Sigma x_t^۲}$ صفر می‌شود، در این نقطه مقدار $t\hat{\sigma}\sqrt{۱ + \frac{۱}{n} + \frac{(X_t - \bar{X})^۲}{\Sigma x_t^۲}}$ برابر $t\hat{\sigma}\sqrt{۱ + \frac{۱}{n}}$ خواهد شد که در واقع کمترین مقدار ممکن آن است. در نتیجه AB یا AC کمترین مقدار خود را در این نقطه خواهند داشت. به ازای مقادیری از X_t که از $\bar{X}$ بزرگتر یا کوچکتر باشد، مقادیر AB یا AC بزرگتر خواهد شد. برای مثال، به ازای $X_t^۱$، مقدار $A_۱B_۱ = A_۱C_۱$ به دست می‌آید، که قطعاً $A_۱B_۱ > AB$ و $A_۱C_۱ > AC$. با توجه به اینکه انحراف X_t از $\bar{X}$ در $SE(e_t)$ به صورت مربع ظاهر می‌شود، قرینگی دارد؛ یعنی فرقی نمی‌کند که X_t از $\bar{X}$ بزرگتر باشد یا کوچکتر. بنابراین اگر $X_t^۱ - \bar{X}$ برابر $X_t^۲ - \bar{X}$ باشد، خواهیم داشت $B_۱C_۱ = B_۲C_۲$. با حرکت X_t بر روی محور X، مجموعه نقاط $B_۱, B_۲, \ldots$ و نیز $C_۱, C_۲, \ldots$ به دست می‌آید که دو منحنی را در صفحه مختصات نشان می‌دهد. فضای بین این دو منحنی منعکس‌کنندهٔ تغییرات فاصله‌های اطمینان Y_t به ازای تغییر مکان X_t در روی محور X است.

حال که با مسألهٔ پیش‌بینی در مدلهای رگرسیون خطی آشنا شدیم، به ذکر مباحث تکمیلی می‌پردازیم. این مباحث در واقع مجموعه‌ای از نکات مهم در تخمین مدلهای خطی است که فرصت مناسب برای طرح آنها در فصلهای قبل نبوده است. تغییر مقیاس در اندازه گیری متغیرها، رگرسیون معکوس، تأثیر مقادیر دور افتادهٔ متغیرهای درون‌زا یا برون‌زا و تبدیل مدلهای رگرسیونی از جمله عناوین مهمی است که در ادامهٔ این فصل به آنها اشاره خواهد شد.

۳ـ۳ تغییر مقیاس در متغیرها

در مباحثی که تا به حال داشتیم، هیچگونه اشاره‌ای به مقیاس متغیرهای موجود در یک مدل رگرسیون نشده است. در عمل، می‌توان متغیرهای برون‌زا و درون‌زا را با مقیاسهای مختلف وارد مدل رگرسیون کرد. برای مثال، وقتی متغیرهای ما تعداد بسیاری صفر در سمت راست دارند معمولاً بهتر است صفرها را تا حد ممکن حذف کرد و در عوض مقیاس متغیرها را بالا برد و آنها را با مقیاس ۱۰ هزار یا میلیون و مانند آن بیان کرد. در این قسمت پاسخ به این سؤال را مطرح می‌کنیم که اگر مقیاس اندازه گیری متغیرهای برون‌زا و درون‌زا را تغییر دهیم، آیا تخمین پارامترهای مدل، تغییر می‌کند؟ بحث را به صورت یک مثال مطرح کرده و سه حالت را در نظر می‌گیریم، تغییر مقیاس در متغیر برون‌زا، تغییر مقیاس در متغیر درون‌زا و تغییر مقیاس در هر دو متغیر.

۱. تغییر مقیاس در متغیر برون‌زا

تابع تقاضا برای نان را در نظر می‌گیریم که در آن Y_t، مقدار تقاضا شده برحسب کیلوگرم در هفته و X_t، قیمت نان برحسب ریال برای هر کیلوگرم است

$$Y_t = \alpha + \beta X_t + U_t \quad . \qquad (۳ـ۲۰)$$

حال مقیاس قیمت نان (X_t) را تغییر داده و بجای ریال آن را با تومان اندازه گیری می‌کنیم. اگر X_t^* قیمت نان برحسب تومان باشد، مدل جدید به صورت زیر خواهد بود،

$$Y_t = \alpha^* + \beta^* X_t^* + U_t \ . \qquad (۳\text{-}۲۱)$$

توجه داریم که پیش فرض مدل ۳ـ۲۱ این است که ورود X_t^* به مدل، هم تخمین α و هم تخمین β را تغییر می‌دهد، در حالی که ممکن است اینچنین نباشد؛ با وجود این، برای اینکه در این مقطع از بحث، حالت عمومی را حفظ کرده باشیم، فرض را بر تغییر هر دو می‌گذاریم. تخمین β^* و α^* با روش حداقل مربعات معمولی عبارت است از

$$\hat{\beta}^* = \frac{\Sigma x_t^* y_t}{\Sigma x_t^{*2}} \ , \qquad (۳\text{-}۲۲)$$

$$\hat{\alpha}^* = \bar{Y} - \beta^* \bar{X}^* \ . \qquad (۳\text{-}۲۳)$$

با توجه به تغییر مقیاس در متغیر X_t، می‌توان نوشت

$$X_t^* = \lambda X_t \ , \qquad (۳\text{-}۲۴)$$

که بر فرض تغییر مقیاس قیمت نان از ریال به تومان، مقدار λ برابر ۰/۱ خواهد بود. می‌توان معادلهٔ ۳ـ۲۴ را برحسب انحراف از میانگین نیز نوشت،

$$x_t^* = \lambda x_t \ . \qquad (۳\text{-}۲۵)$$

با جایگزینی در معادلهٔ ۳ـ۲۲ خواهیم داشت

$$\hat{\beta}^* = \frac{\Sigma (\lambda x_t y_t)}{\Sigma (\lambda x_t)^2} = \frac{\lambda \Sigma x_t y_t}{\lambda^2 \Sigma x_t^2} = \frac{1}{\lambda} \cdot \frac{\Sigma x_t y_t}{\Sigma x_t^2} \ ,$$

در نتیجه داریم

$$\hat{\beta}^* = \frac{1}{\lambda} \hat{\beta} \ . \qquad (۳\text{-}۲۶)$$

می‌توان از معادلهٔ ۳ـ۲۴ نتیجه گرفت که $\bar{X}^* = \lambda \bar{X}$ با جایگزینی آن و همراه با جایگزینی ۳ـ۲۶ در معادلهٔ ۳ـ۲۳ خواهیم داشت

$$\hat{\alpha}^* = \bar{Y} - \frac{1}{\lambda}\hat{\beta}\,(\lambda\,\bar{X}) \ ,$$

$$= \bar{Y} - \hat{\beta}\,\bar{X} \ .$$

بنابراین نتیجه می‌گیریم که

$$\hat{\alpha}^* = \hat{\alpha} \ , \qquad (۳\_۲۷)$$

یعنی تغییر مقیاس متغیر برون‌زا به اندازهٔ λ باعث می‌شود که تخمین ضریب آن متغیر به اندازهٔ $\frac{1}{\lambda}$ تغییر کند؛ در حالی که تخمین ضریب ثابت مدل رگرسیون بدون تغییر باقی می‌ماند.

نتیجه دیگری که از بحث فوق می‌توان به دست آورد این است که اگر یک مدل رگرسیون را تخمین زده و $\hat{\alpha}$ و $\hat{\beta}$ را محاسبه کرده باشیم و سپس تصمیم به تغییر مقیاس متغیر برون‌زا بگیریم به تخمین مجدد پارامترها نیازی نیست. کافی است $\hat{\beta}$ را در معکوس مقیاس تغییر ضرب کنیم تا تخمین جدیدی از شیب مدل داشته باشیم. در $\hat{\alpha}$ نیز هیچگونه تغییری ایجاد نمی‌شود.

می‌توان تأثیر تغییر مقیاس در متغیر برون‌زا بر r^2 و بر مجموع مربعات پسماند را به صورت یک بحث تکمیلی بررسی کرد. با توجه به معادلهٔ ۳٤_۱ ضریب تعیین برای مدل اولیه عبارت است از

$$r^2 = \frac{(\Sigma\, x_t\ y_t)^2}{\Sigma\, x_t^2\ \Sigma\, y_t^2} \ ,$$

که برای مدل جدید به صورت زیر خواهد بود،

$$r^{*2} = \frac{(\Sigma\, x_t^*\ y_t)^2}{\Sigma\, x_t^{*2}\ \Sigma\, y_t^2} \ .$$

معادلهٔ ۲۵_۳ را در فرمول فوق جایگزین می‌کنیم،

$$r^{*2} = \frac{\Sigma (\lambda x_t y_t)^2}{\Sigma (\lambda x_t)^2 \Sigma y_t^2} = \frac{\lambda^2 (\Sigma x_t y_t)^2}{\lambda^2 \Sigma x_t^2 \Sigma y_t^2} ,$$

در نتیجه می‌توان نوشت

$$r^{*2} = r^2 , \qquad (۳\text{ـ}۲۸)$$

یعنی تغییر مقیاس متغیر برون‌زا بر ضریب تعیین تأثیری ندارد.

برای ارزیابی تأثیر متغیر مقیاس متغیر برون‌زا بر مجموع مربعات پسماند (RSS)، ابتدا با توجه به معادلهٔ ۲ـ۴۷ مقدار مجموع مربعات پسماند را برای مدل اولیه می‌نویسیم،

$$RSS = \sum e_t^2 = \sum y_t^2 - \frac{(\Sigma x_t y_t)^2}{\Sigma x_t^2} ,$$

که برای مدل جدید به صورت زیر خواهد بود

$$RSS^* = \sum y_t^{*2} - \frac{(\Sigma x_t^* y_t)^2}{\Sigma x_t^{*2}} .$$

معادلهٔ ۳ـ۲۵ را در فرمول فوق قرار می‌دهیم.

$$RSS^* = \sum y_t^2 - \frac{(\Sigma \lambda x_t y_t)^2}{\Sigma (\lambda x_t)^2} = \sum y_t^2 - \frac{\lambda^2 (\Sigma x_t y_t)^2}{\lambda^2 \Sigma x_t^2}$$

$$= \sum y_t^2 - \frac{(\Sigma x_t y_t)^2}{\Sigma x_t^2} ,$$

در نتیجه

$$RSS^* = RSS , \qquad (۳\text{ـ}۲۹)$$

یعنی تغییر مقیاس متغیر برون‌زا، مجموع مربعات پسماند را تغییر نمی‌دهد. بدیهی است تخمین واریانس U_t نیز تغییر نمی‌کند و در نتیجه آزمون فرضیه‌های مختلف در مورد پارامترهای مدل نیز متأثر نخواهد شد.

۲. تغییر مقیاس در متغیر درون‌زا

مانند گذشته مدل رگرسیون زیر را در نظر گرفته

$$Y_t = \alpha + \beta X_t + U_t \ ,$$

و این بار مقیاس متغیر درون‌زا (Y_t) را به اندازهٔ μ تغییر می‌دهیم، به گونه‌ای که

$$Y_t^* = \mu Y_t \ . \qquad (۳\_۳۰)$$

سؤال این است که آیا با تغییر Y_t به μY_t، تخمین پارامترهای α و β تغییر خواهد کرد؟ همچنین آیا این مسأله بر ضریب تعیین r^2 و نیز بر مجموع مربعات پسماند $\sum e_t^2 = RSS$ تأثیر می‌گذارد؟

برای پاسخ به این سؤال دقیقاً مطابق حالت اول عمل می‌کنیم. ابتدا دو طرف مدل اولیه را در μ ضرب می‌کنیم،

$$\mu Y_t = \mu \alpha + \mu \beta X_t + \mu U_t \ ,$$

یا

$$Y_t^* = \alpha^* + \beta^* X_t + U_t^* \ ,$$

β^* را تخمین می‌زنیم،

$$\hat{\beta}^* = \frac{\sum x_t y_t^*}{\sum x_t^2} \ .$$

از معادلهٔ ۳۰ـ۳ می‌دانیم که $y_t^* = \mu y_t$. با جایگزینی در رابطهٔ فوق خواهیم داشت

$$\hat{\beta}^* = \frac{\sum x_t \mu y_t}{\sum x_t^2} = \mu \frac{\sum x_t y_t}{\sum x_t^2} \ .$$

نتیجه می‌گیریم که

$$\hat{\beta}^* = \mu \hat{\beta} \ , \qquad (۳\_۳۱)$$

یعنی تغییر مقیاس متغیر درون‌زا به اندازهٔ μ، باعث می‌شود که تخمین جدید از β

به اندازهٔ μ برابر تخمین قبلی باشد.

برای تخمین α^* رابطهٔ زیر را در نظر می‌گیریم

$$\hat{\alpha}^* = \bar{Y}^* - \hat{\beta}^* \bar{X} \ ,$$

که با استفاده از معادلهٔ ۳ـ۳۱ خواهیم داشت

$$\hat{\alpha}^* = \mu \bar{Y} - \mu \hat{\beta} \bar{X} \ ,$$

$$= \mu \ (\bar{Y} - \hat{\beta} \bar{X}) \ .$$

نتیجه می‌گیریم که

$$\hat{\alpha}^* = \mu \hat{\alpha} \ , \qquad (۳ـ۳۲)$$

یعنی تغییر مقیاس متغیر درون‌زا به اندازهٔ μ، باعث می‌شود که تخمین جدید از α به اندازهٔ μ برابر تخمین قبلی باشد.

با توجه به معادلهٔ ۱ـ۳٤ می‌دانیم ضریب تعیین برای مدل اولیه عبارت است از

$$r^2 = \frac{(\Sigma x_t y_t)^2}{\Sigma x_t^2 \Sigma y_t^2} \ ,$$

که برای مدل جدید به صورت زیر خواهد بود

$$r^{*2} = \frac{(\Sigma x_t y_t)^2}{\Sigma x_t^2 \Sigma y_t^{*2}} \ .$$

$$= \frac{[\Sigma x_t (\mu y_t)]^2}{\Sigma x_t^2 \ \Sigma (\mu y_t)^2} = \frac{\mu^2 (\Sigma x_t y_t)^2}{\mu^2 \Sigma x_t^2 \Sigma y_t^2} \ ,$$

در نتیجه خواهیم داشت

$$r^{*2} = r^2 \ , \qquad (۳ـ۳۳)$$

یعنی تغییر مقیاس متغیر درون‌زا بر ضریب تعیین تأثیری ندارد.

برای ارزیابی تأثیر تغییر مقیاس متغیر درون‌زا بر مجموع مربعات پسماند (RSS) ابتدا با توجه به معادلهٔ ۲ـ۴۷، مقدار مجموع مربعات پسماند را برای مدل اولیه می‌نویسیم،

$$RSS = \sum e_t^2 = \sum y_t^2 - \frac{(\Sigma x_t\ y_t)^2}{\Sigma x_t^2} \ ,$$

که برای مدل جدید به صورت زیر خواهد بود

$$RSS^* = \sum y_t^{*2} - \frac{(\Sigma x_t y_t^*)^2}{\Sigma x_t^2} \ ,$$

$$= \sum (\mu\, y_t)^2 - \frac{(\Sigma \mu\, x_t y_t)^2}{\Sigma x_t^2} \ ,$$

$$= \mu^2 \sum y_t^2 - \mu^2 \frac{(\Sigma x_t y_t)^2}{\Sigma x_t^2} \ ,$$

در نتیجه

$$= RSS^* = \mu^2\, RSS \ , \qquad (۳ـ۳۴)$$

یعنی تغییر مقیاس متغیر درون‌زا، مجموع مربعات پسماند را تغییر می‌دهد و بر تخمین واریانس U_t و نیز آزمون فرضیه‌های مختلف تأثیر خواهد گذاشت؛ به عبارت دیگر

$$\hat{\sigma}_U^{*2} = \frac{\Sigma e_t^{*2}}{n-2} = \frac{RSS^*}{n-2} \ ,$$

$$= \frac{\mu^2\, RSS}{n-2} = \mu^2\, \hat{\sigma}^2 \ ,$$

در نتیجه خطای معیار U_t یا خطای معیار تخمین (SEE) در مدل جدید دقیقاً μ برابر خطای معیار تخمین مدل اولیه است،

$$SEE^* = \sqrt{\hat{\sigma}_U^{*2}} = \mu\, SEE \ .$$

۳. تغییر مقیاس در متغیرهای برون‌زا و درون‌زا

مدل رگرسیون اولیه را می‌نویسیم،

$$Y_t = \alpha + \beta X_t + U_t \ .$$

مقیاس متغیر برون‌زا را به اندازهٔ λ و مقیاس متغیر درون‌زا را به اندازهٔ μ تغییر می‌دهیم،

$$X_t^* = \lambda X_t \ ,$$

$$Y_t^* = \mu Y_t \ .$$

از رابطهٔ اول داریم، $X_t = \frac{1}{\lambda} X_t^*$، که اگر آن را با معادلهٔ دوم در معادلهٔ

$$\mu Y_t = \mu \alpha + \mu \beta X_t + \mu U_t$$

قرار دهیم، خواهیم داشت

$$Y_t^* = \mu \alpha + \mu \beta \left(\frac{1}{\lambda} X_t^*\right) + \mu U_t \ ,$$

یا

$$Y_t^* = \alpha^* + \beta^* X_t^* + U_t^* \ , \qquad (۳\text{-}۳۵)$$

که در آن $\alpha^* = \mu \alpha$ و $\beta^* = \frac{\mu}{\lambda} \beta$. سؤال این است که آیا تخمین پارامترها در مدل ۳ـ۳۵ با مقادیر مشابه در مدل اولیه تفاوت دارد؟

برای پاسخ، کافی است به ترتیب $\hat{\beta}^*$ و $\hat{\alpha}^*$ را به دست آورده با $\hat{\beta}$ و $\hat{\alpha}$ مقایسه کنیم.

می‌دانیم

$$\hat{\beta}^* = \frac{\Sigma x_t^* y_t^*}{\Sigma x_t^{*2}} \ ,$$

$$= \frac{\Sigma (\lambda x_t)(\mu y_t)}{\Sigma (\lambda x_t)^2} = \frac{\mu \lambda \Sigma x_t y_t}{\lambda^2 \Sigma x_t^2} \ ,$$

در نتیجه خواهیم داشت

$$\hat{\beta}^* = \frac{\mu}{\lambda}\hat{\beta} \ . \qquad (۳\text{ـ}۳۶)$$

یعنی برای رسیدن به تخمین جدید از β، باید تخمین قبلی را در نسبت دو مقیاس $(\frac{\mu}{\lambda})$ ضرب کنیم. برای تخمین $\hat{\alpha}^*$ رابطهٔ زیر را در نظر می‌گیریم

$$\hat{\alpha}^* = \bar{Y}^* - \hat{\beta}^* \bar{X}^* \ ,$$

که با استفاده از معادلهٔ ۳ـ۳۶ خواهیم داشت

$$\hat{\alpha}^* = \frac{\Sigma \mu Y_i}{n} - \frac{\mu}{\lambda}\hat{\beta}\frac{\Sigma \lambda X_i}{n} \ ,$$

$$= \mu \bar{Y} - \frac{\mu}{\lambda}\hat{\beta}\lambda \bar{X} \ ,$$

$$= \mu \bar{Y} - \mu \hat{\beta} \bar{X} = \mu (\bar{Y} - \hat{\beta} \bar{X}) \ ,$$

بنابراین

$$\hat{\alpha}^* = \mu \hat{\alpha} \ . \qquad (۳\text{ـ}۳۷)$$

نتیجهٔ کلی این است که تغییر مقیاس، تخمین پارامترها را دقیقاً به همان صورت تغییر می‌دهد که بر رابطهٔ بین پارامترهای واقعی تأثیر دارد؛ به عبارت دیگر، در معادلهٔ ۳ـ۳۵ دیدیم که

$$\beta^* = \frac{\mu}{\lambda}\beta \quad , \quad \alpha^* = \mu \alpha \ .$$

معادله‌های فوق تأثیر مستقیم تغییر مقیاس بر رابطهٔ بین پارامترهای واقعی را نشان میدهد. در معادله‌های ۳ـ۳۶ و ۳ـ۳۷ نیز دیدیم که

$$\hat{\beta}^* = \frac{\mu}{\lambda}\beta \quad , \quad \hat{\alpha}^* = \mu \hat{\alpha}$$

دقیقاً همان ساختار روابط قبلی را دارد.

برای بررسی آثار تغییر مقیاس بر r^2 و بر مجموع مربعات پسماند و خطای معیار

تخمین کافی است که فرمول هر یک از معیارهای فوق را بعد از تغییر مقیاس بررسی کنیم. برای مدل جدید، یعنی معادلهٔ ۳۵ـ۳ داریم

$$r^{*2} = \frac{(\Sigma x_t^* y_t^*)^2}{\Sigma x_t^{*2} \Sigma y_t^{*2}} .$$

با جایگزینی $x_t^* = \lambda x_t$ و $y_t^* = \mu y_t$ در معادلهٔ فوق خواهیم داشت

$$r^{*2} = \frac{(\Sigma \lambda x_t \mu y_t)^2}{\Sigma (\lambda x_t)^2 . \Sigma (\mu y_t)^2} ,$$

$$= \frac{\lambda^2 \mu^2 (\Sigma x_t y_t)^2}{\lambda^2 \Sigma x_t^2 . \mu^2 \Sigma y_t^2} ,$$

در نتیجه

$$r^{*2} = r^2 \qquad (۳\text{ـ}۳۸)$$

یعنی تغییر همزمان مقیاس متغیرهای درون‌زا و برون‌زا، بر ضریب تعیین تأثیری ندارد.

برای بررسی خطای معیار تخمین کافی است معادلهٔ ۴۷ـ۲ را برای مدل ۳۵ـ۳ بنویسیم،

$$RSS^* = \sum y_t^{*2} - \frac{(\Sigma x_t^* y_t^*)^2}{\Sigma x_t^{*2}} ,$$

$$= \sum (\mu y_t)^2 - \frac{(\Sigma \lambda x_t \mu y_t)^2}{\Sigma (\lambda x_t)^2} ,$$

$$= \mu^2 \sum y_t^2 - \frac{\lambda^2 \mu^2 (\Sigma x_t y_t)^2}{\lambda^2 \Sigma x_t^2} ,$$

$$= \mu^2 \left[\sum y_t^2 - \frac{(\Sigma x_t y_t)^2}{\Sigma x_t^2} \right] ,$$

در نتیجه

$$RSS^* = \mu^2 RSS \ , \qquad (۳\text{-}۳۹)$$

یعنی مجموع مربعات پسماند در مدل جدید، فقط می‌تواند به اندازهٔ μ^2، مقدار مشابه از مدل اولیه را تغییر دهد؛ در حالی که λ مطلقاً در مجموع مربعات پسماند جدید تأثیری ندارد. به همین ترتیب داریم

$$\hat{\sigma}_U^{*2} = \frac{RSS^*}{n-2} = \frac{\mu^2 RSS}{n-2} = \mu^2 \hat{\sigma}_U^2 \ ,$$

و همچنین خواهیم داشت

$$SEE^* = \sqrt{\hat{\sigma}_U^{*2}} = \mu \, SEE \ , \qquad (۳\text{-}۴۰)$$

یعنی تغییر همزمان مقیاس متغیرهای برون‌زا و درون‌زا را به اندازه‌های λ و μ، واریانس U_t را به اندازهٔ μ^2 و خطای معیار تخمین را به اندازهٔ μ تغییر می‌دهد.

۳-۴ رگرسیون معکوس

مدلی را که تا اینجا موضوع مطالعهٔ ما بوده است یک بار دیگر ملاحظه کنید،

$$Y_t = \alpha + \beta X_t + U_t \ .$$

در این مدل اصطلاحاً می‌گوییم Y_t را روی X_t «رگرس»[1] کرده‌ایم؛ یعنی Y_t را متغیر وابسته و X_t را متغیر توضیحی گرفته‌ایم. به مدل فوق، یک مدل «رگرسیون مستقیم»[2] نیز می‌گویند. در مواردی ضروری است که رگرسیون X_t روی Y_t را بررسی کنیم؛ یعنی برعکس مدل فوق عمل کرده، در واقع X_t را متغیر وابسته و Y_t را متغیر توضیحی در نظر بگیریم. خواهیم داشت

$$X_t = \alpha' + \beta' Y_t + V_t \ .$$

مدل فوق را یک مدل «رگرسیون معکوس»[3] می‌نامیم.

---

1. Regress
2. Direct Regression
3. Reverse Regression

برای مثال، اگر بخواهیم تأثیر مسألهٔ تبعیض جنسیت یا تبعیض نژادی را بر سطح دستمزدها بررسی کنیم، می‌توان از مدلهای رگرسیون معکوس استفاده کرد. فرض کنید مشاهدات ما به صورت مقطعی است. Y_i سطح دستمزدها و X_i شاخص صلاحیت و تخصص نیروی کار است. اگر مسألهٔ مورد نظر ما تأثیر تبعیض جنسیت بر دستمزدها باشد آنگاه می‌توان دو سؤال زیر را مطرح کرد.

الف) آیا مردان و زنانی که از یک سطح تخصصی معیّن برخوردارند (مقادیر X_i آنها یکسان است) دستمزدهای یکسان دارند؟ می‌توان این سؤال را با یک مدل رگرسیون مستقیم (مدل رگرسیون Y_i روی X_i) پاسخ داد.

ب) آیا مردان و زنانی که دستمزدهای یکسان دارند (مقادیر Y_i آنها مساوی است) از تخصصهای یکسان بهره‌مند هستند؟ می‌توان پاسخ به این سؤال را با یک مدل رگرسیون معکوس (مدلی که X_i را برحسب Y_i بیان می‌کند) به دست آورد.

در مورد این مثال خاص، مدلهای رگرسیون مستقیم و معکوس به ترتیب عبارتند از

$$Y_i = \alpha + \beta X_i + U_i \ ,$$

$$X_i = \alpha' + \beta' Y_i + V_i \ , \qquad (۳\text{-}۴۱)$$

که در آن V_i جملهٔ اختلال در رگرسیون معکوس بوده و شامل تمام فرضهای کلاسیک است.

با روشی همانند تخمین پارامترها در مدل رگرسیون مستقیم، می‌توان $\hat{\alpha}'$ و $\hat{\beta}'$ را به دست آورد. خواهیم داشت

$$\hat{\beta}' = \frac{\Sigma x_i y_i}{\Sigma y_i^2} \ , \qquad (۳\text{-}۴۲)$$

$$\hat{\alpha}' = \bar{X} - \hat{\beta}' \bar{Y} \ . \qquad (۳\text{-}۴۳)$$

همانند فرمول ۲-۴۷ ـ که برای به دست آوردن مجموع مربعات پسماند (RSS) در مدل رگرسیون مستقیم داشتیم ـ می‌توان مجموع مربعات پسماند را در مدل رگرسیون معکوس

(RSS') از فرمول زیر به دست آورد،

$$\sum e_i'^2 = RSS' = \sum x_i^2 - \frac{(\sum x_i\ y_i)^2}{\sum y_i^2} \quad . \qquad (۳\text{-}۴۴)$$

اگر $\hat{\beta}$ مانند قبل، تخمین β در مدل رگرسیون مستقیم باشد، آنگاه

$$\hat{\beta} = \frac{\sum x_i\, y_i}{\sum x_i^2} \quad ,$$

در نتیجه به سهولت ملاحظه می‌شود که

$$\hat{\beta}\,\hat{\beta}' = \frac{\sum x_i\, y_i}{\sum x_i^2} \cdot \frac{\sum x_i\, y_i}{\sum y_i^2} \quad ,$$

و خواهیم داشت

$$\hat{\beta}\,\hat{\beta}' = r_{x,y}^2 \quad . \qquad (۳\text{-}۴۵)$$

یعنی حاصلضرب تخمین شیبهای دو خط رگرسیون مستقیم و معکوس برابر ضریب تعیین است؛ بنابراین شیب رگرسیون معکوس، هنگامی معکوس شیب رگرسیون مستقیم است که ضریب تعیین مدل برابر یک باشد.

مثال ۳ـ۳ مدل رابطه بین تولید و نیروی کار، موضوع مثالهای ۲ـ۱ و ۱ـ۲ و جدول ۳ـ۱ را یک بار دیگر ملاحظه می‌کنیم.

Q_i	۱۱	۱۰	۱۲	۶	۱۰	۷	۹	۱۰	۱۱	۱۰
L_i	۱۰	۷	۱۰	۵	۸	۸	۶	۷	۹	۱۰

با فرض $Q_i = Y_i$ و $L_i = X_i$، نتایج محاسبات قبلی را دوباره می‌نویسیم،

$$\bar{X} = ۸ \ , \qquad \bar{Y} = ۹/۶ \ ,$$

$$\sum x_i^2 = ۲۸ \ , \qquad \sum y_i^2 = ۳۰/۴ \ ,$$

$$\sum x_i y_i = ۲۱ \ , \qquad r^۲ = ۰/۵۲ \ ,$$

$$\hat{\beta} = ۰/۷۵ \ , \qquad \hat{\alpha} = ۳/٦ \ ,$$

بنابراین تخمین مدل رگرسیون مستقیم، عبارت است از

$$\hat{Y}_i = ۳/٦ + ۰/۷۵ \, X_i \ .$$

برای تخمین مدل رگرسیون معکوس، معادلهٔ ۳ـ٤۲ را می‌نویسیم،

$$\hat{\beta}' = \frac{\Sigma x_i y_i}{\Sigma y_i^۲} \ ,$$

$$= \frac{۲۱}{۳۰/٤} = ۰/٦۹ \ .$$

برای تخمین $\hat{\alpha}'$ از معادلهٔ ۳ـ٤۳ استفاده می‌کنیم،

$$\hat{\alpha}' = \bar{X} - \hat{\beta}' \, \bar{Y} \ ,$$

$$= ۸ - ۰/٦۹ \, (۹/٦) = ۱/۳۷ \ .$$

در نتیجه مدل رگرسیون معکوس عبارت خواهد بود از

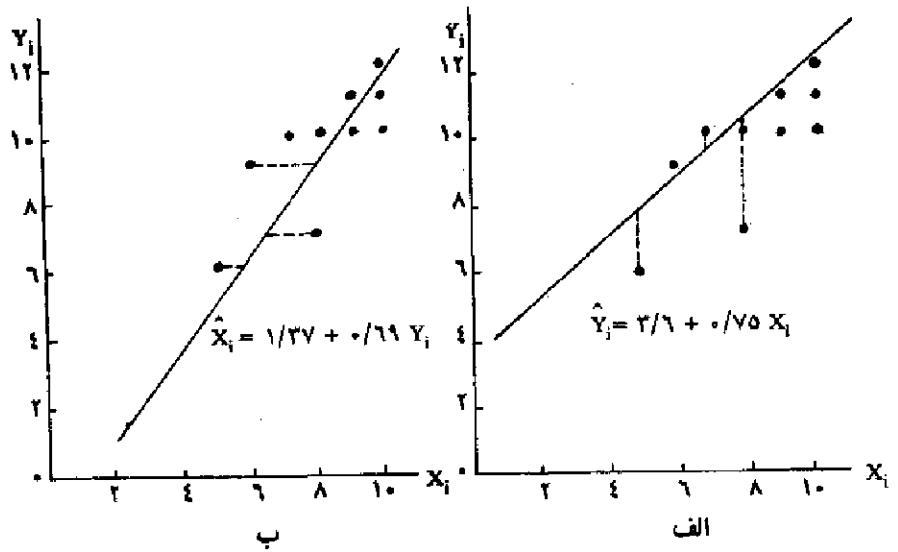
$$\hat{X}_i = ۱/۳۷ + ۰/٦۹ \, Y_i \ .$$

در محاسبات قبلی دیدیم که $r^۲ = ۰/۵۲$. این نتیجه را می‌توان با توجه به معادلهٔ ۳ـ٤۵ نیز به دست آورد. کافی است تخمین شیبهای دو خط رگرسیون مستقیم و معکوس را در هم ضرب کنیم،

$$\hat{\beta}\,\hat{\beta}' = r^۲ \ ,$$

$$۰/۷۵ \, (۰/٦۹) = ۰/۵۲ \ .$$

تبیین بیشتر مدلهای رگرسیون مستقیم و معکوس، مستلزم نمایش هندسی آنهاست. در نمودار ۲ـ۳ الف تخمین مدل رگرسیون مستقیم و در نمودار ۲ـ۳ ب مدل رگرسیون معکوس نشان داده شده است. هر دو نمودار در صفحهٔ مختصات XOY رسم شده است.



نمودار ۲-۳ نمودار هندسی رگرسیونهای مستقیم و معکوس

ملاحظه می‌شود که در نمودار ۲-۳ الف، وقتی می‌خواهیم مجموع مربعات پسماند، یعنی $\Sigma (Y_i - \hat{\alpha} - \hat{\beta} X_i)^2$ برای رگرسیون مستقیم را حداقل کنیم، در واقع مجموع مربع خطوط قائمی حداقل می‌شود که فاصلهٔ نقاط مشاهده شده تا خط تخمین رگرسیون را نشان می‌دهد. برای مثال، سه خط از این فاصله‌ها به صورت نقطه‌چین مشخص شده است. اما مجموع مربعات پسماند برای مدل رگرسیون معکوس عبارت است از $\Sigma (X_i - \hat{\alpha}' - \hat{\beta}' Y_i)$، که حداقل کردن آن، به معنای به حداقل رساندن مجموع مربع خطوط افقی است که فاصلهٔ نقاط مشاهده شده را تا خط تخمین رگرسیون نشان می‌دهد. در نمودار ۲-۳ ب سه خط از این فاصله‌ها به صورت نقطه‌چین مشخص شده است. البته می‌توان فاصلهٔ نقاط مشاهده شده تا خط تخمین رگرسیون را، به جای خطوط قائم یا افقی، با خطوط عمود بر خط تخمین رگرسیون تعریف کرد، در این صورت می‌گوییم مدل ما یک «رگرسیون متعامد»[1] است ـ که بحث آن از موضوع این کتاب خارج است. ناگفته نماند که اگر بخواهیم تخمین یک مدل رگرسیون معکوس را در صحفهٔ مختصات

1. Orthogonal Regression

نشان دهیم، مطابق قاعده باید X_i را روی محور عمودی و Y_i را روی محور افقی اندازه گیری کنیم. اما در نمودار ۲ـ۳ب، Y_i را روی محور عمودی برده‌ایم. علت تنها این بوده است که می‌خواستیم تفاوت دو مفهوم «پسماند در جهت Y_i» و «پسماند در جهت X_i» را در یک نمودار نشان دهیم. در واقع اگر e_i را به صورت زیر تعریف کنیم:

$$\text{پسماند در جهت } Y_i = e_i = Y_i - \hat{Y}_i \ ,$$

یک رگرسیون مستقیم خواهیم داشت که نمایش هندسی آن نمودار ۲ـ۳ الف و مقادیر e_i به صورت خطوط عمودی است. اما اگر داشته باشیم

$$\text{پسماند در جهت } X_i = e_i = X_i - \hat{X}_i \ ,$$

در آن صورت مدل رگرسیون ما معکوس بوده و همان گونه که در نمودار ۲ـ۳ب ملاحظه می‌شود، مقادیر e_i به صورت خطوط افقی خواهد بود.

آخرین نکته این است که آیا معیاری وجود دارد که بتوان گفت در چه مواردی باید از رگرسیون معکوس استفاده کرد؟ پاسخ منفی است؛ اما می‌توان ملاحظات کلی زیر را در انتخاب نوع مدل رگرسیون در نظر گرفت.

الف) اگر نظریه‌های اقتصادی، جهت رابطهٔ علت و معلولی را بین دو متغیر X و Y، مشخص کرده باشد، به راحتی می‌توان نسبت به نوع مدل رگرسیون تصمیم گرفت؛ یعنی متغیری که علت تغییرات است را به عنوان متغیر برون‌زا در سمت راست مدل قرار می‌دهیم و معلول را در سمت چپ می‌بریم. برای مثال، فرض کنید براساس نظریه‌های اقتصادی و ملاحظات دیگر، مطمئن هستیم که در واحد تولیدی i، حجم نیروی کار در زمان t بر سطح تولیدات در زمان t تأثیر می‌گذارد، نه بر عکس. در این صورت یک مدل رگرسیون خواهیم داشت که در آن حجم نیروی کار، متغیر توضیحی یا برون‌زا، و سطح تولیدات، یک متغیر وابسته یا درون‌زاست، خواهیم داشت

$$Q_i = \alpha + \beta L_i + U_i \ .$$

در چنین شرایطی، ساختن یک مدل رگرسیون معکوس، مفهوم چندان روشنی ندارد و

توصیه نمی‌شود.

بنابراین نتیجه می‌گیریم که در تمام مواردی که علت تغییرات را می‌دانیم، ضروری است در مدل رگرسیون، علت به عنوان متغیر توضیحی و معلول به صورت متغیر وابسته در نظر گرفته شود. یادآوری می‌کنیم که در چهارچوب این مثال مدل رگرسیون Q_i بر L_i می‌تواند به دو سؤال زیر پاسخ دهد.

۱. به ازای مقدار معیّنی از حجم نیروی کار در آینده، چگونه می‌توان سطح تولیدات را پیش‌بینی کرد؟

۲. برای رسیدن به سطح معیّنی از تولیدات در آینده، چگونه می‌توان حجم نیروی کار لازم را تخمین زد؟

به عبارت دیگر نیازی نیست که برای پاسخ به سؤال دوم، یک مدل رگرسیون معکوس بسازیم؛ بلکه رگرسیون مستقیم کافی است.

ب) در مواردی که جهت رابطهٔ علت و معلولی بین دو متغیر X و Y چندان روشن نیست، معمولاً هر دو مدل رگرسیون Y روی X و X روی Y می‌تواند مفید باشد. در این موارد بهتر است ابتدا هر دو مدل رگرسیون را تخمین زده، سپس با توجه به نتایج تخمین و شرایط مسأله مفروض، نسبت به انتخاب نهایی یکی از آن دو تصمیم بگیریم.

ج) مواردی نیز وجود دارد که از مفروضات خود مسأله می‌توان دریافت که آیا ساختن یک مدل رگرسیون معکوس ضروری است یا خیر؛ برای مثال، در بحث مسأله تبعیض جنسیت و سطح دستمزدها در معادلهٔ ۳.۴۱ دیدیم که طراحی مدل رگرسیون معکوس می‌تواند روشنگر باشد.

د) و سرانجام این مسأله که آیا رگرسیون مستقیم بر رگرسیون معکوس برتری دارد یا برعکس، در بسیاری موارد به چگونگی تولید آمار و کیفیت عرضهٔ آن بر می‌گردد؛ برای مثال، جدول مثال ۳.۳ را دوباره ملاحظه کنید. L_i یا X_i ساعتهای کار انجام شده و Q_i یا Y_i حجم تولیدات بوده است. تعداد ۱۰ مشاهدهٔ مندرج در این جدول مربوط به ۱۰ کارگر است. اینکه آیا Y_i را تابعی از X_i بگیریم یا برعکس، تا حد بسیاری به این نکته بر می‌گردد که آمار مندرج در این جدول چگونه ایجاد شده است. این دو حالت را

در نظر می‌گیریم:

۱. اگر ساعتهای معین و متفاوتی از زمان کار را به کارگران مختلف بدهیم (X_i)، سپس حجم تولیدات حاصل از این ساعتهای کار را مشاهده و ثبت کنیم (Y_i)، بهتر است که از مدل رگرسیون Y_i بر X_i استفاده کنیم.

۲. اگر به هر یک از کارگران مقادیر مختلفی از تولیدات (Y_i) را سفارش بدهیم و سپس ساعتهای کار انجام شده برای تولید مقادیر مفروض (X_i) را مشاهده و ثبت کنیم آنگاه مفید است که مدل رگرسیون ما X_i را برحسب Y_i بیان کند.

یکی از معیارهایی که می‌تواند در انتخاب بین مدلهای رگرسیون مستقیم و معکوس استفاده شود این است که متغیر تحت کنترل ما کدام است. اگر توانستیم چنین متغیری را شناسایی کنیم، آن را به عنوان متغیر توضیحی یا برون‌زا در سمت راست معادله قرار می‌دهیم؛ برای مثال، در حالت اول، متغیر تحت کنترل ما ساعتهای کار انجام شده است که آن را به کارگران ابلاغ کرده‌ایم، بنابراین به عنوان متغیر توضیحی یا برون‌زا وارد مدل رگرسیون می‌شود. اما در حالت دوم، با تعیین حجم تولیدات و درخواست از کارگران برای تولید آنها، در واقع حجم تولیدات را کنترل می‌کنیم و ساعتهای کار را فقط کارگران تعیین می‌کنند. به همین دلیل است که در حالت دوم (Y_i) به عنوان متغیر توضیحی وارد مدل رگرسیون می‌شود؛ در حالی که X_i متغیر درون‌زاست.

۵ـ۳ مشاهدات دورافتاده

در بعضی از مدلهای رگرسیون ممکن است چند مشاهده «غیرعادی» یا «دورافتاده»[1] بتواند بر تخمین پارامترهای مدل تأثیر بسیاری بگذارد، به گونه‌ای که از دقت آن کم کند و نتایج را غیرقابل اعتماد نماید. دقت در مشاهدات X و Y برای تفسیر نتایج حاصل از تخمین یک مدل رگرسیون همواره ضروری است. باید همیشه به دنبال این سؤال باشیم که آیا مجموعه مشاهدات ما شامل موارد دور افتاده هست یا خیر. تشخیص اینکه آیا

1. Outliers

یک یا چند مشاهده دورافتاده است یا بررسی جمله‌های پسماند صورت می‌پذیرد. اگر X_t شامل موارد دورافتاده باشد، اثر آن با معادلهٔ $\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t$ در $\hat{Y}_t$ ظاهر می‌شود؛ در نتیجه با توجه به $e_t = Y_t - \hat{Y}_t$، در e_t منعکس خواهد شد. اگر مشاهدات Y_t دارای موارد دورافتاده باشد، با رابطهٔ $e_t = Y_t - \hat{Y}_t$، به طور قطع این موارد بر e_t تأثیر می‌گذارد؛ بنابراین بررسی سیر تغییرات e_t برای کشف موارد دورافتاده در مشاهدات X و Y ضروری است.

معمولاً بعد از تخمین یک مدل رگرسیون ساده، فقط به $\hat{\alpha}$، $\hat{\beta}$، r^2 و خطای معیار تخمین زننده‌های $\hat{\alpha}$ و $\hat{\beta}$، یعنی $SE(\hat{\alpha})$ و $SE(\hat{\beta})$ توجه می‌شود؛ در حالی که به نظر می‌رسد مفید است که مقادیر مختلف پسماندها را نیز محاسبه و به روند تغییرات آن توجه کامل کنیم. بنابراین مسأله «آنالیز پسماندها»[1] یکی از مباحث مهم اقتصادسنجی است که باید به طور مستقل بررسی شود. در این قسمت فقط به طرح مطلب اکتفا می‌شود.

یک مشاهده دورافتاده، در واقع مشاهده‌ای است که با مشاهدات دیگر هماهنگی لازم را ندارد. دلایل مختلفی برای به وجود آمدن این مشاهدات وجود دارد که موضوع تحلیلها و نظریه‌های اقتصادی است. اما مسأله مهم این است که روش حداقل مربعات معمولی در تخمین پارامترها حتی از یک مشاهده دورافتاده نیز می‌تواند بسیار متأثر باشد. در مورد رگرسیون ساده، به راحتی می‌توان با استفاده از نمودار هندسی مشاهدات در صفحهٔ مختصات، وجود موارد دور افتاده را کشف کرد؛ اما وقتی رگرسیون ما چند متغیره باشد، یعنی چند متغیر توضیحی را شامل شود، نمودار هندسی مشاهدات، معمولاً غیرممکن است و حتماً باید از آنالیز جملهٔ اختلال استفاده کرد.

مثال ٤ـ٣ این مثال نشان می‌دهد چگونه فقط تخمین پارامترها همراه با r^2 و انحراف معیار تخمین پارامترها، نمی‌تواند تمام خصوصیات موجود در تخمین یک مدل رگرسیون را بیان کند[2]. جدول ۲ـ۳ چهار سری مشاهدات X_t و Y_t را نشان می‌دهد.

---

1. Analysis of Residuals

۲. به مقاله F. J. Anscombe (۱۹۷۳) مراجعه شود.

جدول ۲ـ۳

متغیر		X_i	Y_i	Y_i	Y_i	X_i	Y_i
سری مشاهدات		(۱ـ۳)	۱	۲	۳	٤	٤
مشاهدات	۱	۱۰	۸/۰٤	۹/۱٤	۷/٤٦	۸	٦/۵۸
	۲	۸	٦/۹۵	۸/۱٤	٦/۷۷	۸	۵/۷٦
	۳	۱۳	۷/۵۸	۸/۷٤	۱۲/۷٤	۸	۷/۷۱
	٤	۹	۸/۸۱	۸/۷۷	۷/۱۱	۸	۸
	۵	۱۱	۸/۳۳	۹/۲٦	۷/۸۱	۸	۸/٤۷
	٦	۱٤	۹/۹٦	۸/۱۰	۸/۸٤	۸	۷/۰٤
	۷	٦	۷/۲٤	٦/۱۳	٦/۰۸	۸	۵/۲۵
	۸	٤	٤/۲٦	۳/۱۰	۵/۳۹	۱۹	۱۲/۵
	۹	۱۲	۱۰/۸٤	۹/۱۳	۸/۱۵	۸	۵/۵٦
	۱۰	۷	٤/۸۲	۷/۲٦	٦/٤۲	۸	۷/۹۱
	۱۱	۵	۵/٦۸	٤/۷٤	۵/۷۳	۸	٦/۸۹

برای سه سری مشاهدات اول و دوم و سوم مربوط به Y_i، مشاهدات X_i مندرج در ستون دوم ثابت فرض‌شده است. دو ستون آخرجدول منعکس‌کنندهٔسری چهارم مشاهدات X_i و Y_i است. نکته جالب این است که هر چهار سری این مشاهدات در مورد X_i و Y_i به یک تخمین واحد از مدل رگرسیون $Y_i = \alpha + \beta X_i + U_i$ می‌رسد.

از هر چهار سری مشاهدات فوق کمیتهای زیر را داریم

$$n = ۱۱ \ , \qquad \bar{X} = ۹ \ , \qquad \bar{Y} = ۷/۵ \ ,$$

$$\sum x_i^۲ = ۱۱۰ \ , \qquad \sum y_i^۲ = ٤۱/۲۵ \ , \qquad \sum x_i y_i = ۵۵ \ ,$$

در نتیجه تخمین مدل رگرسیون به صورت زیر خواهد بود

$$\hat{Y}_i = ۳ + ۰/۵ X_i \ .$$
$$(۰/۱۱۸)$$

ملاحظه می‌شود که $SE(\hat{\beta}) = ۰/۱۱۸$. همچنین برای مدل فوق داریم:

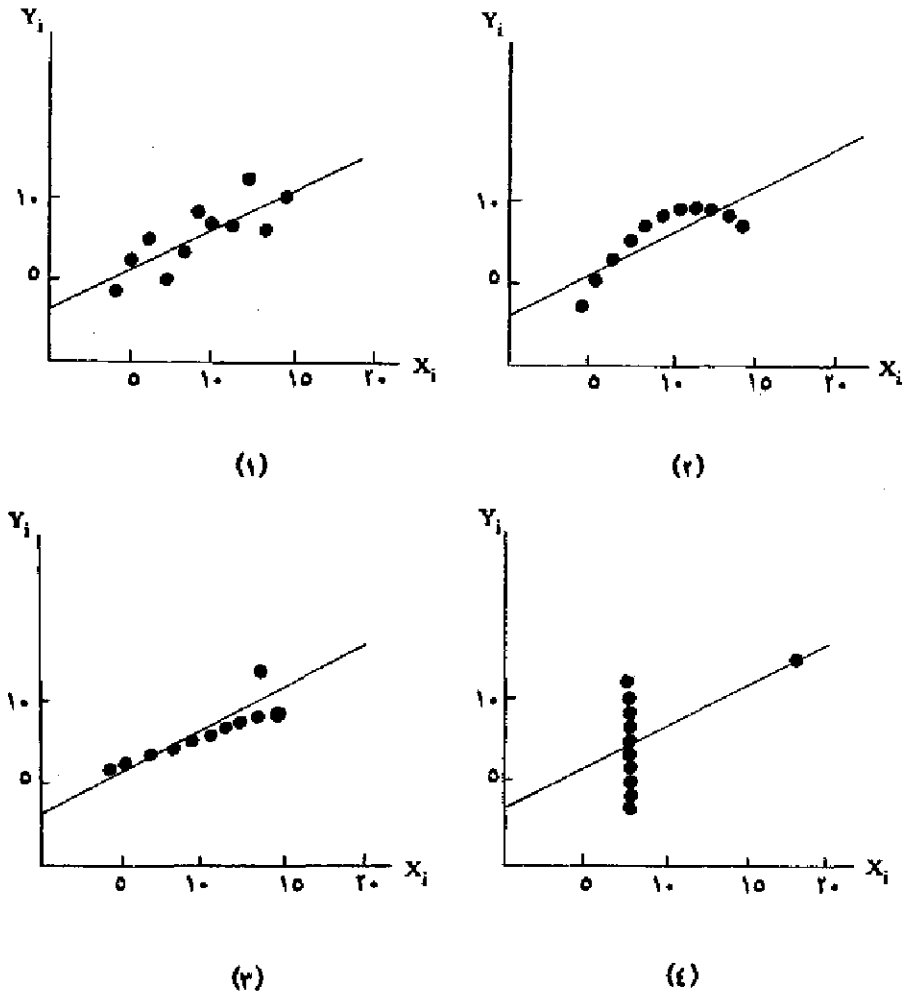
$$r^۲ = ۰/٦٦۷ \ ,$$

$$\sum \hat{y}_i^2 = ESS = ۲۷/۵ \ ,$$

$$\sum e_i^2 = RSS = ۱۳/۷۵ \ .$$

با اینکه چهار سری مشاهدات فوق به تخمین واحدی از مدل رگرسیون مفروض منتهی می‌شود، اما به وضوح ملاحظه می‌گردد که این چهار سری مشاهده، خصوصیات کاملاً متفاوتی با یکدیگر دارند. برای تبیین این نکته کافی است که آنها را در چهار نمودار ۳-۳ منعکس کنیم. نمودار سری اول از مشاهدات مسألهٔ خاصی ندارد؛ اما نمودار هندسی سری دوم از مشاهدات، نشان می‌دهد که یک مدل رگرسیون خطی برای تبیین تغییرات Y_i صلاحیت کافی را ندارد زیرا همان گونه که در نمودار ملاحظه می‌کنیم تغییرات Y_i غیرخطی است. در نمودار سری سوم از مشاهدات ملاحظه می‌کنیم که چگونه یک مشاهدهٔ دورافتاده ($X = ۱۳ \ , Y = ۱۲/۷۴$) شیب تخمین مدل رگرسیون را نسبتاً زیاد تغییر داده و آن را به اصطلاح تندتر کرده است. نکتهٔ مهم این است که اگر مقادیر مختلف پسماندها را حساب کنیم، مقدار پسماند مربوط به این مشاهده ـ که در واقع مشاهدهٔ سوم است ـ یعنی $e_۳$، نسبت به مقادیر دیگر e_i، مقدار کاملاً متفاوت و بیشتری را نشان می‌دهد. ممکن است استدلال شود که باید مشاهدهٔ سوم را حذف کرد تا دیگر مشاهدات که هماهنگی بیشتری با یکدیگر دارند تخمینهای بهتری را نتیجه دهد این نکته را در ادامهٔ بحث دوباره مطرح خواهیم کرد. در نمودار سری چهارم از مشاهدات، تأثیر یک مشاهدهٔ دورافتاده در نمونه به وضوح منعکس شده است. اگر هشتم ($X = ۱۹ \ ، Y = ۱۲/۵$) را از نمونه حذف کنیم، تخمین مدل رگرسیون مفروض به صورت یک خط عمودی ظاهر خواهد شد. بنابراین معلوم می‌شود که چگونه تأثیر یک مشاهدهٔ دورافتاده توانسته است وضعیت تخمین را کاملاً تغییر دهد.

سؤال این است که آیا می‌توان مشاهدات دورافتاده را از نمونه حذف کرد؟ به این سؤال نمی‌توان پاسخ مثبت یا منفی داد؛ شرایط هر مسأله در مجموع می‌تواند رهگشا باشد، با وجود این، می‌توان این ملاحظات کلی را مطرح کرد: در مواردی که شرایط حاکم بر تعیین مقادیر مشاهدات دورافتاده کاملاً با شرایط مشاهدات دیگر متفاوت است، می‌توان تصمیم بر حذف آن مشاهدات گرفت. برای تبیین این نکته دو مثال ذکر می‌کنیم.



نمودار ۳-۳ نمودار هندسی چهارسری مشاهدات

فرض کنید مشاهدات ما به صورت مقطعی و مربوط به تابع تولید در n واحد صنعتی است. همچنین فرض کنید که m مشاهدهٔ دورافتاده در نمونه ملاحظه شده است. اگر این m مشاهدهٔ دورافتاده به واحدهای تولیدی دولتی و در بخش صنایع نظامی محدود باشد، قطعاً شرایط حاکم بر نظام تولیدی این واحدها، مانند نیروی کار تخصصی، بازار فروش، استفاده از ارز دولتی با واحدهای صنعتی دیگر، متفاوت است. در چنین شرایطی

می‌توان این m مشاهده را حذف کرد و برای (n – m) مشاهدهٔ دیگر، یک مدل رگرسیون به صورت مستقل تخمین زد.

مثالی دیگر؛ فرض کنید مشاهدات ما به صورت سری زمانی از تابع تولید در یک بخش صنعتی برای n سال مختلف است. همچنین فرض می‌شود که m مشاهده دورافتاده در این نمونه ملاحظه شده است اگر فرض کنیم که این m مشاهده متعلق به m سالی است که شرایط استثنایی، مانند جنگ، بر نظام تولیدی این بخش صنعتی حاکم بوده است، چه بسا بتوان این چند نمونه دورافتاده را حذف کرد و برای سایر مشاهدات که هماهنگی بهتری با یکدیگر دارند تخمین جداگانه‌ای به دست آورد. البته در فصلهای آینده خواهیم دید که مشاهدات دورافتاده را در اکثر موارد می‌توان با استفاده از «متغیرهای مجازی»[1] در مدل حفظ کرد، به نحوی که در آنالیز پسماند مشاهدات دورافتاده دیگر دیده نشود.

در بسیاری موارد، ظهور مشاهدات دورافتاده بویژه در آنالیز پسماند، نه به علت شرایط استثنایی حاکم بر دوره‌های خاصی از مشاهدات است و نه اینکه عدم هماهنگی (مثلاً) واحدهای تولیدی زمینه‌ساز ایجاد آنها بوده است، بلکه علت اصلی، سادگی مبانی نظری مدل رگرسیون خطی است که نتوانسته است قانونمندی موجود در مشاهدات عینی را به خوبی منعکس کند. در این موارد تخمین مدل رگرسیون در واقع تقریب بسیار ضعیفی از عینیت اقتصادی بوده، در نتیجه مقادیر پسماند در مواردی کاملاً غیرعادی است. راه حل این مسأله قطعاً نمی‌تواند حذف مشاهداتی باشد که عامل ظهور چنین رفتار غیرعادی در پسماند است؛ بلکه برعکس باید این گونه مشاهدات را حفظ کرد و در عوض کوشید تا مدل رگرسیون مفروض را کاملتر کرده و قدرت توضیحی آن را افزایش داد. این کار ـ همان‌گونه که بعد خواهیم دید ـ با اضافه کردن متغیرهای توضیحی جدید، تبدیل مدل رگرسیون خطی به غیرخطی، تخمین سیستم معادلات به جای یک تک‌معادله، یا حتی استفاده از روشهای دیگر تخمین، می‌تواند عملی شود. بنابراین نتیجهٔ کلی این است که حذف مشاهدات دورافتاده همیشه باید نامناسبترین و در عین حال

---

1. Dummy Variables

آخرین راه‌حل تلقی شود.

۶ـ۳ رابطه‌های غیرخطی و تبدیل متغیرها

تاکنون بحث ما دربارهٔ مدلهای رگرسیون خطی ساده بود. در قسمت ۱ـ۱ به طور خلاصه به این نکته اشاره کردیم که منظور از خطی بودن رگرسیون این است که متغیرها (Y, X) و پارامترها $(\alpha$ و $\beta)$ به صورت خطی وارد مدل می‌شوند. یک معادله برحسب مثلاً متغیر X خطی است اگر X فقط با توان یک در آن معادله ظاهر شده و در متغیر دیگری ضرب یا بر متغیر دیگری تقسیم نشده باشد؛ بنابراین معادله‌های

$$Y_t = \alpha + \beta X_t^1 + U_t \ ,$$

$$Y_t = \alpha + \beta \sqrt{x_t} + U_t \ ,$$

$$Y_t = \alpha + \beta X_t Z_t + U_t \ ,$$

که در آن Z_t یک متغیر است، دیگر برحسب X_t خطی نیستند. همچنین یک معادله برحسب مثلاً پارامتر β خطی است، اگر β فقط با توان یک در آن معادله ظاهر شده و در پارامتر دیگری ضرب یا بر پارامتر دیگری تقسیم نشده باشد. بنابراین معادله‌های

$$Y_t = \alpha + \sqrt{\beta} X_t + U_t \ ,$$

$$Y_t = \alpha + \ln \beta X_t + U_t \ ,$$

$$Y_t = \frac{\alpha}{\beta} \ln X_t + U_t \ ,$$

که در آن ln لگاریتم بر مبنای e یا لگاریتم طبیعی است، دیگر برحسب β خطی نیستند. ناگفته نماند که معادلهٔ آخر، هم برحسب β و هم برحسب X_t غیرخطی است.

از دو حالت غیرخطی برحسب پارامترها و غیرخطی برحسب متغیرها، حالت اول اهمیت فراوان دارد. اگر در یک مدل رگرسیون، پارامترها غیرخطی باشد و هیچ «تبدیلی» برای خطی کردن آنها نتوان یافت، روش حداقل مربعات معمولی دیگر

کاربردی نخواهد داشت. اما معمولاً می‌توان تبدیلهایی یافت که متغیرهای غیرخطی را خطی کند. به همین دلیل می‌توان از واژهٔ «خطی» در عبارت «مدل رگرسیون خطی ساده» ، خطی بودن پارامترها را نتیجه گرفت. در این قسمت به بررسی تبدیلهایی می‌پردازیم که می‌تواند یک مدل غیرخطی برحسب متغیرها را به یک مدل خطی تبدیل کند.

۱. مدلهای خطی ـ لگاریتمی

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \alpha X_t^{\beta} e^{U_t} , \qquad (۳\text{ـ}۴۶)$$

که برحسب X_t غیرخطی است. اگر از دو طرف این معادله لگاریتم بگیریم خواهیم داشت

$$\ln Y_t = \ln \alpha + \beta \ln X_t + U_t ,$$

و با فرض $\alpha_0 = \ln \alpha$ داریم

$$\ln Y_t = \alpha_0 + \beta \ln X_t + U_t . \qquad (۳\text{ـ}۴۷)$$

مدل ۳ـ۴۷ برحسب پارامترها و همچنین برحسب لگاریتم متغیرهای X_t و Y_t خطی است. به همین دلیل به آن مدل «خطی ـ لگاریتمی»[1] می‌گویند. با توجه به اینکه در مدل ۳ـ۴۷ متغیرهای سمت راست و سمت چپ هر دو برحسب لگاریتم است، به آن مدل "Log - log" یا مدل «لگاریتم مضاعف»[2] نیز گفته میشود.

برای تخمین پارامترهای α و β این روابط را تعریف می‌کنیم،

$$Y_t^* = \ln Y_t \quad , \quad X_t^* = \ln X_t ,$$

که با جایگزینی در معادله ۳ـ۴۷ خواهیم داشت

1. Log-linear Model 2. Double-log Model

$$Y_t^* = \alpha + \beta X_t^* + U_t \ .$$

اگر U_t تمام فرضهای کلاسیک را داشته‌باشد، به راحتی می‌توان α_* و β را با روش حداقل مربعات معمولی تخمین زد. باید توجه داشت که در تخمین β مسأله‌ای نخواهیم داشت زیرا از تخمین مدل فوق مستقیماً $\hat{\beta}$ به دست می‌آید. اما با توجه به معادلهٔ ۳-۴۶ پارامتر دیگر مورد نظر ما α است؛ در حالی که از تخمین مدل فوق به $\hat{\alpha}_*$ می‌رسیم نه $\hat{\alpha}$؛ بنابراین با توجه به فرض $\alpha_* = \ln \alpha$، باید از $\hat{\alpha}_*$ آنتی‌لگاریتم بگیریم تا $\hat{\alpha}$ به دست آید. می‌توان نشان داد که تخمین $\hat{\alpha}$ اریب دارد؛ در حالی که $\hat{\beta}$ تخمین نااریبی از β است. با اینکه در این روش $\hat{\alpha}$ اریب دارد، اما چندان باعث تضعیف تبدیل «خطی ـ لگاریتمی» نیست؛ زیرا مهم این است که توانسته‌ایم به یک تخمین نااریب از شیب معادله رگرسیون برسیم.

می‌دانیم بین مفاهیم کشش و لگاریتم رابطهٔ بسیاری نزدیکی وجود دارد. در واقع در معادله $Y_t = f(X_t)$، کشش نقطه‌ای Y_t نسبت به X_t (η) برابر است با[1]

---

۱. برای اثبات، به تعریف کشش مراجعه می‌کنیم. در معادلهٔ $Y_t = f(X_t)$، اگر ΔX بتواند ΔY را ایجاد کند آنگاه رابطهٔ

$$\frac{\Delta Y_t}{Y_t} \div \frac{\Delta X_t}{X_t} = \frac{\Delta Y_t}{\Delta X_t} \cdot \frac{X_t}{Y_t}$$

می‌تواند تغییر نسبی در Y_t را به ازای یک واحد تغییر نسبی از X_t اندازه گیری نماید. کشش Y_t نسبت به X_t در واقع مقدار حدی عبارت فوق است وقتی ΔX_t به سمت صفر میل کند؛

$$\eta = \text{کشش نقطه‌ای } Y_t \text{ نسبت به } X_t = \frac{dY_t}{dX_t} \cdot \frac{X_t}{Y_t} \ .$$

حال ثابت می‌کنیم $\eta = \dfrac{d(\ln Y_t)}{d(\ln X_t)}$. ابتدا تعریف می‌کنیم $z_t = \ln X_t$ و نیز $X_t = e^{w_t}$ و در نتیجه $w_t = \ln X_t$. بنابراین

$$\frac{d(\ln Y_t)}{d(\ln X_t)} = \frac{dZ_t}{dX_t} = \frac{dZ_t}{dY_t} \cdot \frac{dY_t}{dX_t} \cdot \frac{dX_t}{dw_t} \ .$$

می‌دانیم، $\dfrac{dZ_t}{dY_t} = \dfrac{1}{Y_t}$ ، $\dfrac{dX_t}{dw_t} = \dfrac{1}{dw_t/dX_t} = \dfrac{1}{1/X_t} = X_t$ و در نتیجه با جایگزینی در عبارت فوق خواهیم داشت

$$\frac{d(\ln Y_t)}{d(\ln X_t)} = \frac{dY_t}{dX_t} \cdot \frac{X_t}{Y_t} = \eta \ .$$

$$\eta = \frac{d\ (\ln Y_t)}{d\ (\ln X_t)} \ .$$

با ملاحظهٔ معادلهٔ ۳ـ۴۷ می‌توان نتیجه گرفت که مشتق $(\ln Y_t)$ نسبت به $(\ln X_t)$ برابر کشش Y_t نسبت به X_t است که دقیقاً برابر β است:

$$\text{کشش نقطه‌ای } Y_t \text{ نسبت به } X_t = \eta = \frac{d\ (\ln Y_t)}{d\ (\ln X_t)} = \beta \ .$$

نتیجه می‌گیریم که اگر مدل ۳ـ۴۶ را به یک مدل خطی ـ لگاریتمی تبدیل کنیم، نه تنها پارامترهای α و β به راحتی با روش حداقل مربعات معمولی تخمین زده می‌شوند بلکه $\hat{\beta}$ تخمینی از کشش متغیر درون‌زا نسبت به متغیر برون‌زا نیز هست. با توجه به اینکه این کشش ثابت است، به مدل ۳ـ۴۷ «مدل کشش ثابت»[1] نیز می‌گویند.

به نمودار هندسی تغییرات مدل ۳ـ۴۶ اشاره‌ای می‌کنیم. جملهٔ اختلال (e^{U_t}) را در نظر نگرفته و فقط به رابطهٔ زیر توجه می‌کنیم،

$$Y_t = \alpha\ X_t^{\beta} \ . \qquad (۳ـ۴۸)$$

اولاً، باید فقط مقادیر مثبت X_t و Y_t را در نظر گرفت؛ چون لگاریتم برای کمیتهای منفی تعریف نمی‌شود. ثانیاً، بهتر است نمودار تغییرات تابع فوق را یک بار برای $\beta > 0$ و بار دیگر برای $\beta < 0$ بررسی کنیم. ابتدا از معادلهٔ ۳ـ۴۸ مشتق می‌گیریم تا شیب تابع به دست آید،

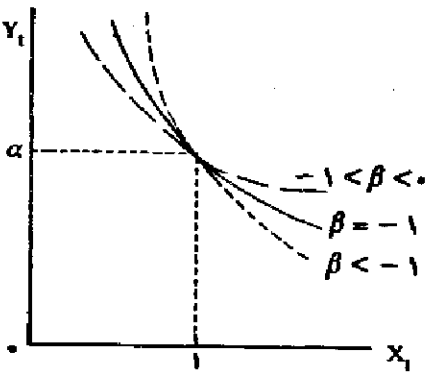
$$\frac{d\ Y_t}{d\ X_t} = \alpha\ \beta\ X_t^{\beta-1} \ . \qquad (۳ـ۴۹)$$

اگر β مثبت باشد، با دیدن ۳ـ۴۹ می‌توان گفت که شیب $\left(\frac{d\ Y_t}{d\ X_t}\right)$ همواره مثبت است و به ازای افزایش X_t و میل آن به سمت بی‌نهایت، Y_t نیز زیاد شده و به سمت بی‌نهایت میل می‌کند.

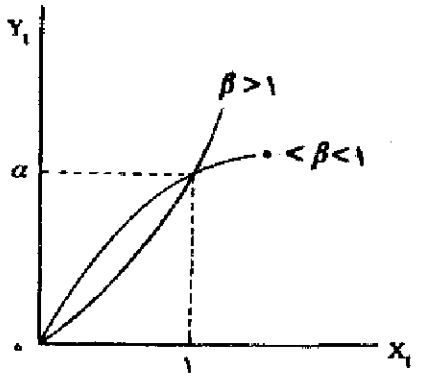
---

1. Constant-elasticity Model

دو حالت را می‌توان از یکدیگر تمیز داد. اگر $0 < \beta < 1$، آنگاه شیب منحنی، به ازای افزایش X_t، به طور مرتب کمتر می‌شود، هر چند همواره مثبت باقی می‌ماند. اما اگر $\beta > 1$، آنگاه به ازای افزایش X_t، شیب منحنی به طور مرتب زیاد خواهد شد. این دو حالت را می‌توان برای $\beta > 0$ در نمودار ۳ـ٤ مشاهده کرد توجه داریم که در معادلهٔ ۳ـ٤۸ مقدار Y_t به ازای $X_t = 1$ برابر α می‌شود.



نمودار ۳ـ۵ تابع خطی ـ لگاریتمی و $\beta < 0$



نمودار ۳ـ٤ تابع خطی ـ لگاریتمی و $\beta > 0$

منحنی تغییرات تابع ۳ـ٤۸ را در حالت $\beta < 0$ بررسی می‌کنیم. با توجه به معادلهٔ ۳ـ٤۹ می‌توان گفت که اگر β منفی باشد، شیب همواره منفی خواهد بود. البته می‌دانیم که α همواره مثبت است؛ زیرا $\ln \alpha$ برای $\alpha < 0$ نمی‌تواند وجود داشته باشد. مانند نمودار ۳ـ٤، می‌توان حالتهای مختلفی را برای β در نظر گرفت و نمودار تغییرات تابع ۳ـ٤۸ را برای آنها ترسیم کرد. در نمودار ۳ـ۵ سه حالت $-1 < \beta < 0$، $\beta = -1$، و $\beta < -1$ بررسی شده است.

۲. مدلهای نیمه لگاریتمی

در مدل رگرسیون زیر

$$Y_t = e^{(\alpha + \beta X_t + U_t)}, \qquad (۳ـ۵۰)$$

که برحسب X_t غیرخطی است، اگر از دو طرف این مدل لگاریتم بگیریم خواهیم داشت:

$$\ln Y_t = \alpha + \beta X_t + U_t \quad . \qquad (۳\text{-}۵۱)$$

با فرض $Y_t^* = \ln Y_t$، معادلهٔ ۳ـ۵۱ را می‌توان به صورت زیر نوشت،

$$Y_t^* = \alpha + \beta X_t + U_t \quad ,$$

که در واقع یک مدل رگرسیون خطی است که می‌توان پارامترهای آن را به راحتی با روش حداقل مربعات معمولی تخمین زد. با توجه به اینکه در مدل ۳ـ۵۱ فقط متغیر درون‌زا برحسب لگاریتم است، به آن «مدل نیمه‌لگاریتمی»[1] می‌گویند.

مدلهای رگرسیون از نوع

$$Y_t = \alpha\, e^{(\beta X_t + U_t)} \qquad (۳\text{-}۵۲)$$

نیز وقتی به مدلهای نیمه‌لگاریتمی تبدیل شوند برحسب X_t خطی خواهند بود. اگر از دو طرف معادلهٔ ۳ـ۵۲ لگاریتم بگیریم خواهیم داشت

$$\ln Y_t = \ln \alpha + \beta X_t + U_t \quad ,$$

که با تعریف $\alpha_* = \ln \alpha$ به صورت زیر خواهد بود،

$$\ln Y_t = \alpha_* + \beta X_t + U_t \quad ,$$

که دقیقاً با معادلهٔ ۳ـ۵۱ ساختاری مشابه دارد. با فرض $Y_t^* = \ln Y_t$، می‌توان مدل فوق را به صورت زیر نوشت،

$$Y_t^* = \alpha_* + \beta X_t + U_t \quad .$$

ناگفته نماند که مدلهایی از نوع ۳ـ۵۱ در تخمین توابع سرمایه‌های انسانی کاربردهای فراوانی دارند، که معمولاً X_t، سالهای اشتغال به تحصیل و Y_t، درآمد فرد است.

برای آشنایی بیشتر با توابع نیمه‌لگاریتمی باید در نظر داشت که اگر Y_t سریعتر از

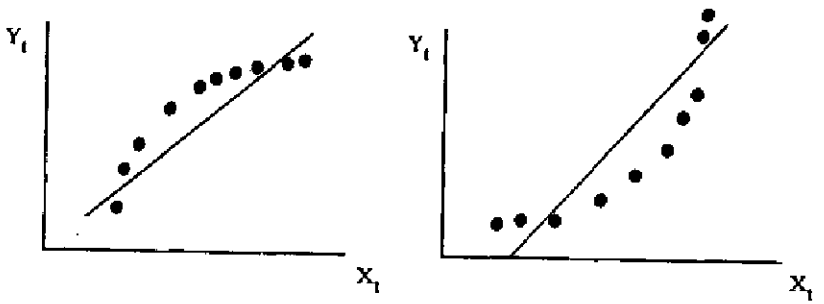
---

1. Semilog Model

X_t افزایش یابد، از مدل ۳ـ۵۱ و در حالتی که Y_t آهسته‌تر از X_t رشد کند، می‌توان از مدل زیر استفاده کرد

$$Y_t = \alpha' + \beta' \ln X_t + U_t \quad . \qquad (۳ـ۵۳)$$

به مدل ۳ـ۵۳ نیز یک مدل نیمه‌لگاریتمی می‌گویند؛ زیرا فقط X_t برحسب لگاریتم بیان شده است. در نمودار ۳ـ۶ الف حالتی نشان داده شده است که مدل ۳ـ۵۱ کاربرد دارد و در نمودار ۳ـ۶ ب کاربرد مدل ۳ـ۵۳ منعکس است.



الف) Y_t سریعتر از X_t رشد می‌کند
$\ln Y_t = \alpha + \beta X_t$

ب) Y_t آهسته‌تر از X_t رشد می‌کند
$Y_t = \alpha' + \beta' \ln X_t$

نمودار ۳ـ۶ مدلهای نیمه‌لگاریتمی

مدل نیمه‌لگاریتمی ۳ـ۵۱ را در نظر می‌گیریم. برای اینکه بتوانیم تفسیر مناسبی از β ارائه کنیم، از $\ln Y_t$ نسبت به X_t مشتق می‌گیریم،

$$\frac{d \ln Y_t}{d X_t} = \beta \quad .$$

رابطه فوق را می‌توان به صورت زیر نوشت،

$$\beta = \frac{1}{Y_t} \cdot \frac{d Y_t}{d X_t} = \frac{d Y_t}{Y_t} \cdot \frac{1}{d X_t}$$

از طرف دیگر می‌دانیم $\frac{d Y_t}{Y_t}$ با تغییر نسبی در Y_t برابر است و $d X_t$ نیز چیزی جز تغییر مطلق در X_t نیست به این ترتیب

$$\beta = \frac{\text{تغییر نسبی در } Y_t}{\text{تغییر مطلق در } X_t} .$$

در مواردی که به ازای یک مقدار معین از تغییر مطلق در X_t، متغیر Y_t به صورت یک درصد ثابتی تغییر می‌کند، می‌توان نتیجه گرفت که بهترین مدل رگرسیون برای آنها، مدل نیمهٔ لگاریتمی از این نوع است

$$\ln Y_t = \alpha + \beta X_t + U_t .$$

این گونه مدلها را «مدل رشد ثابت»[1] نیز می‌گویند. می‌دانیم مدلهای رشد ثابت کاربرد فراوانی در اندازه‌گیری نرخ رشد روند تغییرات متغیرهایی چون قیمت، بیکاری، صادرات، واردات و مانند آن دارد.

با ترتیبی کاملاً مشابه می‌توان β' را در مدل ۳ـ۵۳ به صورت زیر تفسیر کرد،

$$\beta' = \frac{\text{تغییر مطلق در } Y_t}{\text{تغییر نسبی در } X_t} .$$

بنابراین در مواردی که مشاهده میشود به ازای یک مقدار معیّن از تغییر نسبی در X_t، متغیر Y_t به صورت مطلق تغییر می‌کند، نتیجه می‌گیریم که مدل نیمه‌لگاریتمی به صورت زیر

$$Y_t = \alpha' + \beta' \ln X_t + U_t ,$$

بهترین مدل برای تخمین رابطهٔ بین آنها خواهد بود.

برای بررسی نمودار هندسی مدلهای نیمه‌لگاریتمی، مدل ۳ـ۵۱ را یک بار دیگر ملاحظه کنید،

$$\ln Y_t = \alpha + \beta X_t + U_t .$$

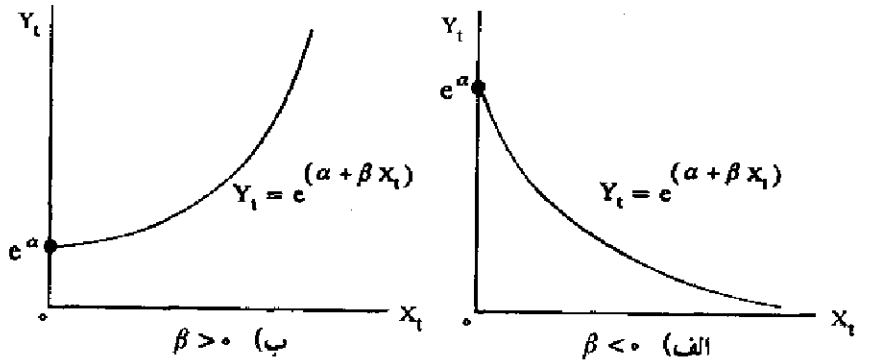
می‌دانیم این تابع فقط برای مقادیر مثبت Y_t تعریف می‌شود، زیرا کمیتهای منفی، لگاریتم ندارند. اگر جملهٔ اختلال را نادیده بگیریم، می‌توان مدل فوق را به صورت زیر نوشت،

---

1. Constant Growth Model

$$Y_t = e^{(\alpha + \beta X_t)}$$

که در واقع همان معادله ۵۰ـ۳ و بدون جملهٔ اختلال است. این معادله را می‌توان برای دو حالت $\beta < 0$ و $\beta > 0$ در نمودار ۷ـ۳ ملاحظه کرد.



نمودار ۷ـ۳ مدلهای نیمه‌لگاریتمی برای $\beta < 0$ و $\beta > 0$

به ازای $X = 0$ مقدار Y_t برابر e^{α} خواهد بود؛ بنابراین عرض از مبدأ در هر دو نمودار برابر e^{α} است. مشاهده می‌شود که علامت شیب تابع، از علامت β تبعیت می‌کند. اگر β منفی باشد شیب منفی و تابع نزولی خواهد بود، که در نمودار ۷ـ۳ الف نشان داده شده است. برای مواردی که β مثبت است، همچنانکه در نمودار ۷ـ۳ ب ملاحظه می‌کنیم منحنی تغییرات تابع ۵۰ـ۳ صعودی است.

می‌توان حالت خاصی را از معادلهٔ ۵۰ـ۳ در نظر گرفت که X_t متغیر زمان باشد. در این صورت خواهیم داشت

$$Y_t = e^{\alpha + \beta t} \qquad (۵۴ـ۳)$$

معادلهٔ ۵۴ـ۳ نشان می‌دهد که Y_t به ازای $\beta > 0$ ، یک نرخ نسبی رشد ثابت داشته و به ازای $\beta < 0$ از یک نرخ رکود نسبی و ثابت برخوردار است.

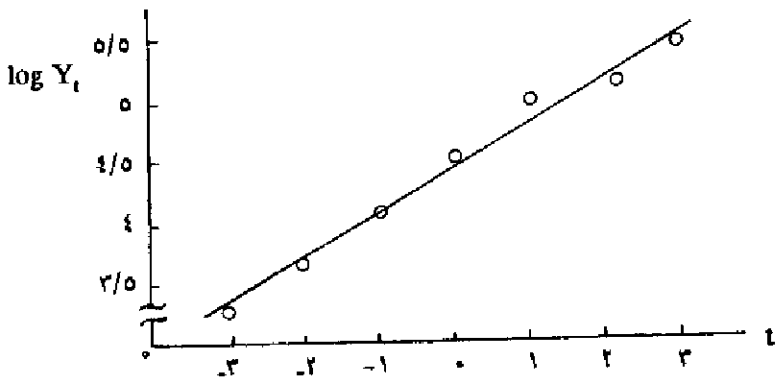
مثال ۵ـ۳ آمار تولیدات یک بخش صنعتی در ۷ دههٔ گذشته در جدول زیر ارائه شده

است. به نظر شما برای این مشاهدات چه نوع مدل رگرسیون می‌توان پیشنهاد کرد. آیا می‌توان گفت که این صفت از نرخ رشد ثابتی برخوردار بوده است.

جدول ۳.۳

دهه	Y_t: متوسط تولید	$\log Y_t$	$X = t$
۱	۱۸۳۷	۳/۲۶۴۱	-۳
۲	۴۸۶۸	۳/۶۸۷۳	-۲
۳	۱۲۴۱۱	۴/۰۹۳۷	-۱
۴	۳۲۶۱۷	۴/۵۱۳۵	۰
۵	۸۲۷۷۰	۴/۹۱۷۹	۱
۶	۱۴۸۴۵۷	۵/۱۷۱۸	۲
۷	۳۲۲۹۵۸	۵/۵۰۹۲	۳

لگاریتم متوسط تولید را در دهه‌های مختلف حساب کرده و آن را در یک نمودار و در مقابل t نشان دهیم. به سهولت ملاحظه می‌شود که مشاهدات موجود در نمودار ۳.۸



نمودار ۳.۸

یک رابطهٔ تقریباً خطی را نشان می‌دهد؛ بنابراین مدلی که می‌توان پیشنهاد کرد عبارت است از

$$\ln Y_t = \alpha + \beta t + U_t$$

که دقیقاً همان مدل ۳ـ۵٤ است. با استفاده از روش حداقل مربعات معمولی می‌توان به تخمین پارامترهای α و β رسید. خواهیم داشت

$$\widehat{\ln Y_t} = ٤/٤٥١٠ + ٠/٣٧٦٠\, t \quad .$$

توجه داریم که در اندازه گیری متغیر زمان، دورهٔ دهه چهارم را مبنا گرفته و مقدار آن را صفر قرار داده‌ایم.

۳. مدلهای وارون

به مدل رگرسیون زیر

$$Y_t = \alpha + \beta \left(\frac{1}{X_t}\right) + U_t \quad , \qquad (۳ـ۵۵)$$

که در آن Y_t برحسب معکوس X_t بیان شده است یک «مدل وارون»[1] می‌گویند.
برای بررسی منحنی تغییرات این تابع، ابتدا جملهٔ اختلال مدل (U_t) را در معادلهٔ ۳ـ۵۵ در نظر نگرفته، از آن مشتق می‌گیریم:

$$\frac{d Y_t}{d X_t} = -\frac{\beta}{X_t^2} \quad .$$

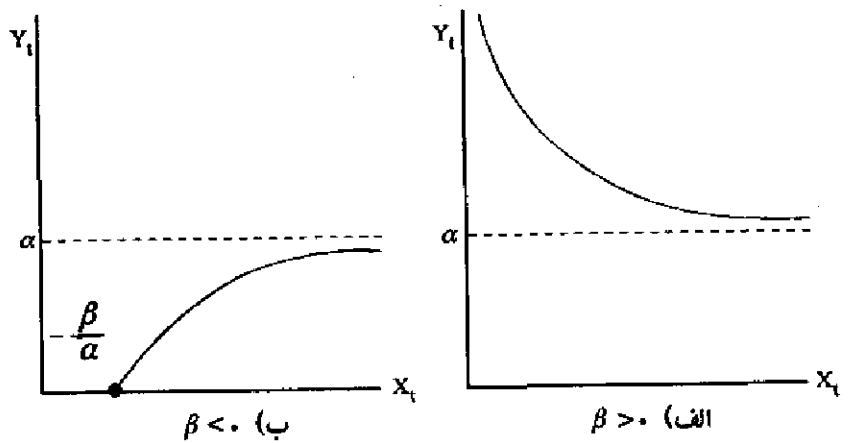
چون علامت X_t^2 همواره مثبت است؛ بنابراین علامت مشتق دقیقاً برعکس علامت β است. حال منحنی تغییرات تابع Y_t را در دو حالت بررسی خواهیم کرد:

حالت اول $\beta > ٠$

هرگاه β مثبت باشد، علامت مشتق منفی و منحنی تغییرات Y_t نزولی است. اگر X_t به سمت بی‌نهایت میل کند، مقدار $\frac{1}{X_t}$ به صفر میل کرده، در نتیجه مقدار Y_t در معادلهٔ ۳ـ۵۵ به سمت α میل خواهد کرد؛ بنابراین خط $Y_t = \alpha$ در واقع مجانب منحنی تغییرات

1. Reciprocal Model

Y_t است. نمودار ۹ـ۳ الف، منحنی تغییرات تابع را به ازای $\beta > 0$ نشان می‌دهد.



نمودار ۹ـ۳ مدلهای معکوس

حالت دوم $\beta < 0$

به ازای مقادیر منفی β، مقدار $\frac{d\,Y_t}{d\,X_t}$ مثبت شده و منحنی تغییرات Y_t یک سیر صعودی را نشان می‌دهد. در این حالت نیز اگر $X_t \to \infty$، $Y_t = \alpha$ مجانب منحنی تغییرات Y_t خواهد بود. به ازای $Y_t = 0$ خواهیم داشت

$$\alpha + \frac{1}{\beta} X_t = 0 \ ,$$

در نتیجه

$$X_t = \frac{-\beta}{\alpha} \ ,$$

یعنی منحنی تغییرات Y_t محور X را در نقطهٔ $\frac{-\beta}{\alpha}$ قطع می‌کند.

مدلهای رگرسیون وارون کاربردهای بسیاری بویژه در اقتصادسنجی خرد دارد؛ برای مثال، می‌توان گفت که نمودار ۹ـ۳ الف در واقع مبیّن شکل عمومی یک «منحنی فیلیپس»[1] است. می‌دانیم در بسیاری موارد می‌توان نرخ تغییر سطح دستمزدها

1. Phillips Curve

را با رگرسیون W_t بر U_t به صورت زیر تخمین زد،

$$W_t = \alpha + \beta \frac{1}{U_t} + \varepsilon_t ,$$

که در آن W_t و U_t به ترتیب نرخ تغییر سطح دستمزدها و نرخ بیکاری است. مشاهده می‌شود که با تخمین α، می‌توان حد پایین تغییر در نرخ دستمزدها را به دست آورد، به گونه‌ای که با افزایش بیشتر نرخ بیکاری، دستمزدها کاهش نمی‌یابد.

مدلهای تخمین تابع مخارج خانوار برای کالاها و خدمات خاص، می‌تواند در بسیاری موارد به صورت یک مدل وارون طراحی شود که در آن β منفی است. فرض کنید با استفاده از داده‌های مقطعی روی یک نمونه n تایی از خانوارها می‌خواهیم تابع مخارج خانوار را برای کالا یا خدمت معینی تخمین بزنیم،

$$Y_i = \alpha + \beta \frac{1}{X_i} + U_i ,$$

که در آن Y_i مخارج خانوار i برای کالای مفروض و X_i درآمد کل یا مخارج کل این خانوار است. به نمودار ۹ـ۳ب مراجعه می‌کنیم. قبل از اینکه درآمد یا مخارج کل خانوار به نقطهٔ $X_i = -\frac{\beta}{\alpha}$ برسد، هیچ تقاضایی برای این کالا یا خدمت مفروض وجود ندارد. ولی از نقطهٔ $-\frac{\beta}{\alpha}$ به بعد، به ازای افزایش درآمد، مخارج خانوار بشدت افزایش می‌یابد، اما بزودی به نقطهٔ تعادل $Y_i = \alpha$ می‌رسد، به گونه‌ای که با افزایش بیشتر درآمد، خانوار حاضر نیست برای آن کالا یا خدمت هزینهٔ بیشتر بپردازد. بسیاری از مواد غذایی لوکس یا پوشاک وارداتی و مانند آن خصوصیت فوق را دارد. لازم است یادآوری کنیم که هنگامی می‌توان از مدل وارون برای تخمین این گونه توابع مصرف خانوار استفاده کرد که $\hat{\beta}$ منفی و $\hat{\alpha}$ مثبت باشد.

حال که با مفهوم توابع وارون و کاربرد آن در تخمین مدلهای رگرسیون آشنا شدیم به چگونگی تخمین این گونه مدلها می‌پردازیم. می‌دانیم مدل ۵۵ـ۳ برحسب X_i غیرخطی است. برای اینکه بتوان α و β را با روش حداقل مربعات معمولی تخمین زد، باید متغیرها و پارامترها همه خطی باشند. با توجه به اینکه در این مدل α و β به صورت

خطی وارد شده‌اند فقط X_t را خطی می‌کنیم. با استفاده از تعریف

$$X_t^* = \frac{1}{X_t} \ ,$$

و جایگزینی آن در مدل ۳ـ۵۵ خواهیم داشت

$$Y_t = \alpha + \beta X_t^* + U_t \ ,$$

که یک مدل خطی برحسب متغیرها و پارامترهاست و می‌توان براحتی آن را با روش حداقل مربعات معمولی تخمین زد.

٤. مدلهای وارون لگاریتمی

به مدل رگرسیون زیر

$$\ln Y_t = \alpha - \beta \left(\frac{1}{X_t}\right) + U_t \qquad (۳ـ۵۶)$$

که در آن Y_t به صورت لگاریتمی و برحسب وارون X_t بیان شده است یک «مدل وارون لگاریتمی»[1] می‌گویند. ساختار ریاضی این مدل با مدل ۳ـ۵۵ تشابه دارد و تفاوت آنها فقط این است که در ۳ـ۵۶ متغیر Y_t به صورت لگاریتمی وارد شده است.

می‌توان گفت که مدل ۳ـ۵۶ در واقع عبارت است از مدل

$$Y_t = e^{(\alpha - \beta \frac{1}{X_t} + U_t)} \ , \qquad (۳ـ۵۷)$$

زیرا کافی است از معادله ۳ـ۵۷ لگاریتم گرفته تا معادلهٔ ۳ـ۵۶ به دست آید. برای بررسی منحنی تغییرات Y_t ابتدا U_t را نادیده گرفته خواهیم داشت

$$Y_t = e^{\left(\alpha - \frac{\beta}{X_t}\right)} \ . \qquad (۳ـ۵۸)$$

مقدار Y_t به ازای $X_t = ۰$ تعریف نمی‌شود زیرا Y_t در معادلهٔ ۳ـ۵۸ تابعی از $\frac{1}{X_t}$ است. اما

1. Logarithmic Reciprocal Model

اگر X_t به سمت صفر میل کند، مقدار Y_t نیز به سمت صفر میل خواهد کرد؛ بنابراین با صرف‌نظر کردن از دقتهای ریاضی، می‌توان به طور خلاصه گفت که $Y_0 = 0$. در نتیجه با این قرارداد، می‌توانیم منحنی تغییرات Y_t را از مبدأ مختصات شروع کنیم؛ چون Y_t همواره مثبت است؛ منحنی تغییرات آن از مبدأ مختصات و به طور پیوسته به سمت راست ادامه خواهد داشت.

برای بررسی شیب تغییرات این منحنی، باید از معادلهٔ ۳ـ۵۸ مشتق بگیریم، خواهیم داشت

$$\frac{d Y_t}{d X_t} = \left(\frac{\beta}{X_t^2}\right) e^{\left(\alpha - \frac{\beta}{X_t}\right)} \quad , \qquad (۳ـ۵۹)$$

یعنی علامت شیب دقیقاً از علامت β تبعیت می‌کند. اگر β مثبت باشد، شیب تابع مثبت بوده و منحنی تغییرات Y_t صعودی است و به ازای مقادیر منفی β، علامت $\frac{d Y_t}{d X_t}$ منفی خواهد شد که در آن صورت، منحنی تغییرات Y_t نزولی است.

حال به بررسی مشتق دوم می‌پردازیم. خواهیم داشت

$$\frac{d^2 Y_t}{d X_t^2} = -\frac{2\beta X_t}{X_t^4} e^{\left(\alpha - \frac{\beta}{X_t}\right)} + \frac{\beta}{X_t^2}\left[\frac{\beta}{X_t^2} e^{\left(\alpha - \frac{\beta}{X_t}\right)}\right] ,$$

یا

$$\frac{d^2 Y_t}{d X_t^2} = \left(\frac{\beta^2}{X_t^4} - \frac{2\beta}{X_t^3}\right) e^{\left(\alpha - \frac{\beta}{X_t}\right)} .$$

اگر مشتق دوم را مساوی صفر قرار دهیم، نقطهٔ عطف منحنی تغییرات Y_t در نقطه‌ای به طول $X_t = \frac{\beta}{2}$ به دست می‌آید. در سمت چپ این نقطه، یعنی $X_t < \frac{\beta}{2}$، شیب منحنی به ازای افزایش X_t، زیاد می‌شود و در سمت راست آن، به ازای افزایش X_t، شیب منحنی کاهش خواهد یافت. برای به دست آوردن عرض نقطهٔ عطف کافی است $X_t = \frac{\beta}{2}$ را در معادلهٔ ۳ـ۵۸ قرار دهیم،

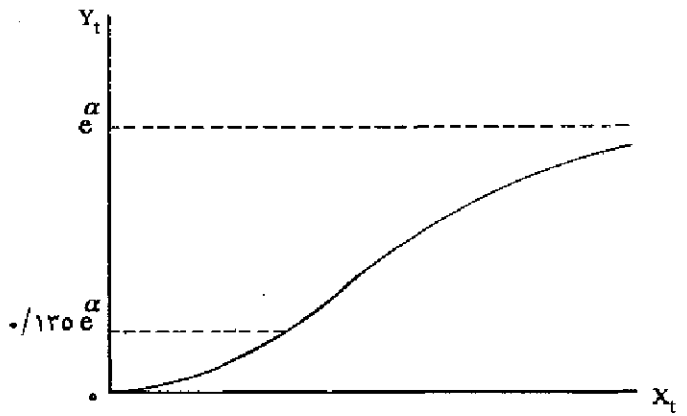
$$Y_t = e^{\left(\alpha - \frac{\beta}{\beta/2}\right)}$$

$$= e^{\alpha - ۲} = e^{\alpha} \cdot e^{-۲} ,$$

که با توجه به $e = ۲/۷۱۸$ خواهیم داشت

$$Y_t = ۰/۱۳۵\, e^{\alpha}.$$

از طرف دیگر می‌دانیم به ازای افزایش X_t و میل آن به سمت بی‌نهایت، مقدار Y_t به سمت e^{α} میل می‌کند؛ بنابراین خط $Y_t = e^{\alpha}$ در واقع مجانب منحنی تغییرات Y_t است. مشاهده می‌شود که عرض نقطهٔ عطف، یعنی $۰/۱۳۵\, e^{\alpha}$، حدود ۱۳ درصد مقدار حدی Y_t، یعنی e^{α} است. با این اطلاعات، می‌توانیم نمودار تغییرات منحنی Y_t را در معادلهٔ ۳ـ۵۸ در نمودار ۱۰ـ۳ نشان دهیم.



نمودار ۳.۱۰ مدل وارون لگاریتمی

مشاهده می‌کنیم که نرخ رشد Y_t به موازات افزایش X_t بشدت کاهش می‌یابد این خصوصیت مدلهای وارون لگاریتمی را باید در کاربرد آنها همواره در نظر داشت. از نظر ریاضی نیز می‌توان این خصوصیت را بسادگی نشان داد. معادلهٔ ۳ـ۹۵ را یک بار دیگر می‌نویسیم،

$$\frac{d\,Y_t}{d\,X_t} = \left(\frac{\beta}{X_t^{۲}}\right) e^{\left(\alpha - \frac{\beta}{X_t}\right)} .$$

دو طرف این معادله را بر معادلهٔ ۳ـ۵۸ تقسیم می‌کنیم. جملهٔ $e^{(\alpha - \frac{\beta}{X_t})}$ حذف شده خواهیم داشت

$$\frac{1}{Y_t} \cdot \frac{dY_t}{dX_t} = \frac{\beta}{X_t^2} \ ,$$

یعنی نرخ رشد Y_t تابعی از معکوس X_t^2 است و با افزایش X_t بشدت کم خواهد شد.

بعد از آشنایی با خصوصیات ریاضی و نمودار تغییرات منحنی مدلهای وارون لگاریتمی به بررسی چگونگی تخمین آنها می‌پردازیم. معادلهٔ ۳ـ۵۶ را یک بار دیگر می‌نویسیم،

$$\ln Y_t = \alpha - \beta \left(\frac{1}{X_t}\right) + U_t \ .$$

با تبدیل متغیرها به

$$\ln Y_t = Y_t^* \ , \qquad \frac{1}{X_t} = X_t^* \ ,$$

خواهیم داشت

$$Y_t^* = \alpha - \beta X_t^* + U_t \ ,$$

که برحسب متغیرها و پارامترها خطی است؛ بنابراین با روش حداقل مربعات معمولی براحتی می‌توان پارامترهای آن را تخمین زد.

در پایان، یک بار دیگر نمودار ۳ـ۱۰ را ملاحظه کنیم. منحنیهای دیگری هم وجود دارند که شبیه نمودار تغییرات منحنی مدلهای وارون لگاریتمی هستند؛ اما:

اولاً، علاوه بر یک مجانب در بالا، یک مجانب دیگر نیز دارند که همان محور X_t است؛ به عبارت دیگر، معادله‌های $Y_t = k$ و $Y_t = 0$ مجانبهای آن است که k عدد ثابت و معینی است.

ثانیاً، منحنی تغییرات آنها در بین دو مجانب تقریباً قرینگی دارد.

به معادله‌های این گونه منحنیها «مدلهای لجیستیک»[1] می‌گویند. این مدلها در تخمین

1. Logistic Models

مدلهای رگرسیون مربوط به روند رشد متغیرها کاربرد فراوانی دارند و در جلد دوم این کتاب با نام «مدلهای عکس‌العمل کیفی»[1] بررسی خواهند شد.

۵. مدلهای غیرخطی برحسب پارامترها

تا به حال بحث ما عمدتاً دربارهٔ آن دسته از مدلهای رگرسیون بود که متغیرهای آن به صورت غیرخطی وارد مدل شده‌اند. در این قسمت به بررسی حالت خاصی می‌پردازیم که پارامترها غیرخطی هستند. مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \alpha + \frac{\beta}{X_t + \gamma} + U_t \quad . \qquad (۳\text{-}۶۰)$$

بدیهی است به علت غیرخطی بودن γ، نمی‌توان پارامترهای مدل را با روش حداقل مربعات معمولی تخمین زد. در اینجا به روشی اشاره می‌کنیم که حالت عمومی دارد و می‌تواند برای همهٔ مدلهایی به کار رود که با ثابت نگهداشتن یک پارامتر، مدل غیرخطی به یک مدل خطی برحسب پارامترها تبدیل می‌شود؛ برای مثال، اگر در مدل ۶۰-۳ مقدار γ را ثابت بگیریم، مدل مزبور برحسب پارامترها خطی شده و به راحتی می‌توان آن را، با فرض $X_t^* = \frac{1}{X_t}$، به کمک روش حداقل مربعات معمولی تخمین زد.

می‌دانیم مبنای روش حداقل مربعات معمولی در تخمین مدل ۶۰-۳ این است که مجموع مربعات پسماند را حداقل کند؛ بنابراین باید $\hat{\alpha}$، $\hat{\beta}$ و $\hat{\gamma}$ را چنان به دست آوریم که عبارت زیر حداقل شود،

$$\sum e_t^2 = \sum \left(Y_t - \hat{\alpha} - \frac{\hat{\beta}}{X_t + \hat{\gamma}} \right)^2 \quad .$$

برای هر مقدار ممکن از $\hat{\gamma}$، متغیر Z_t را به صورت زیر تعریف می‌کنیم،

$$Z_t = \frac{1}{X_t + \hat{\gamma}} \quad . \qquad (۳\text{-}۶۱)$$

---

1. Qualitative Response Models

با جایگزینی معادلهٔ ۳ـ۶۱ در مجموع مربعات پسماند خواهیم داشت

$$\sum e_t^2 = \sum (Y_t - \hat{\alpha} - \hat{\beta} Z_t)^2 \quad . \qquad (۳ـ۶۲)$$

از معادلهٔ ۳ـ۶۲ برحسب $\hat{\alpha}$ و $\hat{\beta}$ مشتق گرفته و مقادیری از $\hat{\alpha}$ و $\hat{\beta}$ که مشتق را صفر می‌کند، تخمینهای روش حداقل مربعات معمولی از α و β خواهد بود.

نکتهٔ مهم این است که مقدار Z_t در معادلهٔ ۳ـ۶۱ باید به ازای تمام مقادیر ممکن $\hat{\gamma}$ تعریف شود. اگر m مقدار برای $\hat{\gamma}$ فرض شود، در آن صورت m مقدار مختلف برای Z_t خواهیم داشت. مشاهده می‌شود که به ازای m مقدار Z_t می‌توان m معادلهٔ از ۳ـ۶۲ به دست آورد؛ بنابراین m مقدار مختلف $\sum e_t^2$ خواهیم داشت. از بین m مقدار $\sum e_t^2$، کوچکترین آن را انتخاب کرده و Z_t آن را مشخص می‌سازیم. آن مقدار از $\hat{\gamma}$ ـ که Z_t را تعیین کرده است ـ تخمین‌زنندهٔ روش حداقل مربعات معمولی از γ خواهد بود. مقادیر $\hat{\alpha}$ و $\hat{\beta}$ متناظر با کمترین مقدار $\sum e_t^2$ نیز تخمین زننده‌های روش حداقل مربعات معمولی α و β است.

بدیهی است لازمهٔ اصلی کاربرد روش فوق این است که با استفاده از نظریه‌ها و مشاهدات اقتصادی، محدودهٔ تغییرات γ را از قبل بدانیم تا بتوانیم مقادیر Z_t و $\sum e_t^2$ را در قلمروی محدود محاسبه کنیم. روشهای دیگری نیز برای خطی کردن پارامترهای غیرخطی وجود دارد که از برنامه این کتاب خارج است.

مسائل فصل سوم

۳ـ۱ مسأله ۱ـ۲ را یک بار دیگر ملاحظه کنید. مدل و مشاهدات زیر مفروض است،

$$Y_i = \alpha + \beta X_i + U_i \quad .$$

X_i :	۲	۳	۱	۵	۹
Y_i :	٤	۷	۳	۹	۱۷

می‌دانیم نتایج زیر به دست آمده است،

$$\hat{Y}_i = ۱ + ۱/۷۵ X_i \quad ,$$

$$\bar{X} = ٤ \quad ,$$

$$\hat{\sigma}_U^۲ = s^۲ = \frac{\Sigma e_i^۲}{n - ۲} = \frac{۱/۵}{۳} = ۰/۵ \quad ,$$

$$\sum x_i^۲ = ٤۰ \quad .$$

به ازای $X_f = ۱۰$، مطلوب است

۱. پیش‌بینی نقطه‌ای برای Y_f، یعنی $\hat{Y}_f$؛

۲. فاصلهٔ اطمینان ۹۵ درصد برای Y_f؛

۳. فاصلهٔ اطمینان ۹۵ درصد برای میانگین Y_f، یعنی $E(Y_f)$.

۳ـ۲ مدل رگرسیون زیر مفروض است

$$Y_t = \alpha + \beta X_t + \varepsilon_t \quad ,$$

که در آن ε_t جملهٔ اختلال مدل بوده و توزیع نرمال با میانگین صفر و واریانس $\sigma^۲$ دارد. این کمیتها را از نمونه‌ای با ۲۰ مشاهده محاسبه کرده‌ایم،

$$\sum Y_t = ۲۱/۹ \quad ، \quad \sum (Y_t - \bar{Y})^۲ = ۸۶/۹ \quad , \quad \sum (X_t) = ۱۸۶/۲ \quad ,$$

$$\sum (X_t - \bar{X})(Y_t - \bar{Y}) = ۱۰۶/٤ \quad ، \quad \sum (X_t - \bar{X})^۲ = ۲۱۵/٤ \quad .$$

الف) پارامترهای α و β را تخمین بزنید.

ب) واریانس $\hat{\alpha}$ و $\hat{\beta}$ را تخمین بزنید.

ج) مقدار میانگین شرطی Y_t را به شرط $X_t = ۱۰$ تخمین بزنید،

$$\hat{E}\,(Y_t \mid X_t = ۱۰)$$

د) برای میانگین شرطی Y_t ـ که در بند (ج) تعریف شده است ـ یک فاصلهٔ اطمینان ۹۵ درصد بسازید.

۳ـ۳ مدل رگرسیون زیر را که جملهٔ ثابت ندارد، مشاهده کنید،

$$Y_t = \beta X_t + U_t \ ,$$

که در آن $t = ۱, ۲, \ldots, n$ و U_t تمام فرضهای کلاسیک را شامل است. به ازای مقدار X_t، مقدار Y_t تعریف می‌شود. می‌خواهیم $Y_t^۲$ را پیش‌بینی کنیم. پیش‌بینی کنندهٔ $\hat{Y}_t^۲$ را چنان تعیین کنید که

$$E\,(\hat{Y}_t^۲) = E\,(Y_t^۲) \ .$$

۴ـ۳ مدل رگرسیون زیر مفروض است،

$$Y_t = \beta X_t + U_t \ ,$$

که در آن $t = ۱, ۲, ۳, \ldots, T, T+۱, T+۲, \ldots, T+F$. تمام فرضهای کلاسیک در مورد U_t صادق است. برای هر دو متغیر X_t و Y_t مشاهدات $t = ۱, ۲, \ldots, T$ موجود است، اما مشاهدات $t = T+۱, T+۲, \ldots, T+F$ را فقط برای X_t داریم. پیش‌بینی کنندهٔ ما از Y_{T+m} عبارت است از

$$\hat{Y}_{T+m} = \hat{\beta}\, X_{T+m} \ ,$$

که در آن $\hat{\beta}$ تخمین روش حداقل مربعات معمولی از β است که با استفاده از مشاهدات $t = ۱, ۲, \ldots, T$ به دست آمده است. نشان دهید که واریانس پیش‌بینی کننده در حول مقدار واقعی آن برابر است با واریانس پیش‌بینی کننده در حول مقدار واقعی آن برابر

است با واریانس پیش‌بینی کننده در حول میانگین آن، بعلاوه واریانس جملهٔ اختلال:

$$E\left(\hat{Y}_{T+F} - Y_{T+F}\right)^2 = E\left[\hat{Y}_{T+F} - E\left(Y_{T+F}\right)\right]^2 + \sigma_U^2 \ .$$

۳ـ۵ نشان دهید که اگر در مدل رگرسیون ساده، یک تبدیل خطی روی هر دو متغیر برون‌زا و درون‌زا انجام شود؛

$$Y_t^* = p_1 + q_1 Y_t \ ,$$

$$X_t^* = p_2 + q_2 X_t \ ,$$

آنگاه ضریب تعیین، (r^2) بدون تغییر باقی می‌ماند.

۳ـ۶ نشان دهید که فقط هنگامی تخمین شیب رگرسیون Y_t بر X_t برابر معکوس تخمین شیب رگرسیون X_t بر Y_t است که داشته باشیم $r^2 = 1$.

۳ـ۷ در مدل رگرسیون $Y_i = \alpha + \beta X_i + U_i$ ، با استفاده از یک نمونهٔ ۲۰۰ تایی مشاهده در مورد X_i و Y_i ، کمیتهای زیر را محاسبه کرده‌ایم،

$$\sum X_i = ۱۱/۳۴ \ , \quad , \quad \sum Y_i = ۲۰/۷۲ \ ,$$

$$\sum X_i^2 = ۱۲/۱۶ \ , \quad , \quad \sum Y_i^2 = ۸۴/۹۶ \ , \quad , \quad \sum X_i Y_i = ۲۲/۱۳ \ .$$

α و β را تخمین زده و آنها را با $\hat{\alpha}'$ و $\hat{\beta}'$ در مدل رگرسیون زیر مقایسه کنید،

$$X_i = \alpha' + \beta' Y_i + V_i \ .$$

۳ـ۸ مدل رگرسیون $Y_t = \beta X_t + U_t$ مفروض است

$\hat{\beta}$، ضریب تعیین (r^2)، و تخمین واریانس جملهٔ اختلال ($\hat{\sigma}_U^2$) را در نظر می‌گیریم. مدل رگرسیون معکوس را بصورت $X_t = \beta' Y_t + V_t$ می‌نویسیم. $\hat{\beta}'$ ، r^2 و $\hat{\sigma}_U^2$ را برای این مدل به دست آورده، با مقادیر مشابه در رگرسیون مستقیم مقایسه کنید.

۳ـ۹ آیا مدلهای رگرسیون زیر را می‌توان به مدلهایی تبدیل کرد که برحسب پارامترها خطی باشد؟

الف) $Y_t = \alpha e^{\beta X_t} U_t$. ج) $Y_t = \alpha e^{(\alpha + \beta X_t + U_t)}$.

ب) $Y_t = \alpha e^{-\beta X_t} + U_t$. د) $Y_t = \frac{\alpha}{(\beta - X_t)} + U_t$.

۳-۱۰ فرض کنید آمار X_t و Y_t را داریم. توضیح دهید که چگونه می‌توان پارامتر مدلهای زیر را با روش حداقل مربعات معمولی تخمین زد:

الف) $Y_t = \alpha X_t^{\beta}$. هـ) $Y_t = \frac{e^{\alpha + \beta X_t}}{1 + e^{\alpha + \beta X_t}}$.

ب) $Y_t = \alpha e^{\beta X_t}$. و) $Y_t = \alpha + \beta \sqrt{X_t}$.

ج) $Y_t = \alpha + \beta \ln X_t$ ز) $Y_t = \alpha + e^{\beta X_t}$.

د) $Y_t = \frac{X_t}{\alpha X_t - \beta}$. ح) $Y_t = \alpha + \frac{\beta}{2 X_t^2 - \gamma}$.

درباره چگونگی ورود جملهٔ اختلال یعنی U_t، در هر یک از مدلهای فوق بحث کنید.

حل مسائل فصل سوم

۱ـ۳الف) با استفاده از تخمین مدل مفروض داریم،

$$\hat{Y}_t = ۱ + ۱/۷۵ X_t \ .$$

به ازای $X_t = ۱۰$، پیش‌بینی نقطه‌ای عبارت است از

$$\hat{Y}_t = ۱ + ۱/۷۵ (۱۰) = ۱۸/۵ \ .$$

ب) نامساوی ۱۰ـ۳ را می‌نویسیم،

$$\hat{Y}_t - t_{\alpha/۲}\, \hat{\sigma}_U \sqrt{۱ + \frac{۱}{n} + \frac{(X_t - \bar{X})^۲}{\Sigma x_t^۲}} < Y_t < \hat{Y}_t + t_{\alpha/۲}\, \hat{\sigma}_U \sqrt{۱ + \frac{۱}{n} + \frac{(X_t - \bar{X})^۲}{\Sigma x_t^۲}}$$

مقدار به دست آمده از جدول t با $n - ۲ = ۵ - ۲ = ۳$ درجهٔ آزادی و در سطح معنی‌دار $\alpha/۲ = ۰/۰۲۵$ برابر است با $\pm ۳/۱۸۲$. با استفاده از کمیتهای داده شده، خواهیم داشت

$$۱۸/۵ - ۳/۱۸۲ \sqrt{۰/۵} \sqrt{۱ + \frac{۱}{۵} + \frac{(۱۰ - ٤)^۲}{٤۰}} < Y_t < ۱۸/۵$$
$$+ ۳/۱۸۲ \sqrt{۰/۵} \sqrt{۱ + \frac{۱}{۵} + \frac{(۱۰ - ٤)^۲}{٤۰}}$$

در نتیجه فاصلهٔ اطمینان ۹۵ درصد برای Y_t عبارت خواهد بود از

$$۱۸/۵ - ۳/۲۵ < Y_t < ۱۸/۵ + ۳/۲٦ \ ,$$

یا:

$$۱۵/۲٤ < Y_t < ۲۱/۷٦ \ .$$

ج) برای به دست آوردن فاصلهٔ اطمینان ۹۵ درصد برای میانگین Y_t، از نامساوی

۱۸ـ۳ استفاده می‌کنیم،

$$\hat{Y}_f - t_{\alpha/2}\hat{\sigma}_U\sqrt{\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\Sigma x_i^2}} < E(Y_f) < \hat{Y}_f + t_{\alpha/2}\hat{\sigma}_U\sqrt{\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\Sigma x_i^2}} .$$

در نتیجه خواهیم داشت

$$۱۸/۵ + ۳/۱۸۲\sqrt{۰/۵}\sqrt{۱+\frac{۱}{۵}+\frac{(۱۰-٤)^2}{٤۰}} < E(Y_f) < ۱۸/۵ + ۳/۱۸۲\sqrt{۰/۵}\sqrt{۱+\frac{۱}{۵}+\frac{(۱۰-٤)^2}{٤۰}} ,$$

$$۱۸/۵ - ۲/۳٦ < E(Y_f) < ۱۸/۵ + ۲/۳٦ ,$$

یا

$$۱٦/۱٤ < E(Y_f) < ۲۰/۸٦ .$$

ملاحظه می‌شود فاصلهٔ اطمینان برای $E(Y_f)$ از فاصلهٔ اطمینان برای Y_f کمتر است.

۲ـ۳الف) $\hat{\alpha}$ و $\hat{\beta}$ را می‌توان به ترتیب زیر محاسبه کرد،

$$\hat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} = \frac{۱۰٦/٤}{۲۱۵/٤} = ۰/٤۹٤ ,$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} = ۱/۰۹۵ - ۰/٤۹٤\,(۹/۳۱) ,$$

ب) ابتدا σ_U^2 را تخمین می‌زنیم،

$$\hat{\sigma}_U^2 = \frac{\Sigma e_t^2}{n-۲} = \frac{\Sigma y_t^2 - \hat{\beta}\Sigma x_t y_t}{n-۲} ,$$

$$= \frac{۸٦/۹ - ۰/٤۹٤\,(۱۰٦/٤)}{۲۰-۲} ,$$

$$= ۱/۹۰۷۷ .$$

با داشتن واریانس U_t، تخمین واریانسهای $\hat{\alpha}$ و $\hat{\beta}$ به راحتی به دست می‌آید،

$$\hat{\text{Var}}(\hat{\alpha}) = \hat{\sigma}_u^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{\Sigma x_t^2}\right) ,$$

$$= ۱/۹۰۷۷ \left(\frac{۱}{۲۰} + \frac{۹/۳۱}{۲۱۵/٤}\right) ,$$

$$= ۰/۸٦۳۰ .$$

$$\text{Var}(\hat{\beta}) = \frac{\hat{\sigma}^2}{\Sigma x_t^2} ,$$

$$= \frac{۱/۹۰۷۷}{۲۱۵/٤} = ۰/۰۰۸۹ .$$

ج) می‌دانیم منظور از میانگین شرطی Y_t به شرط $X_t = ۱۰$، دقیقاً همان مفهوم پیش‌بینی است؛ به عبارت دیگر می‌خواهیم بگوییم اگر X مقدار ۱۰ بگیرد، میانگین Y چقدر خواهد بود؟ در معادلهٔ ۱۳ـ۳ دیدیم که

$$\hat{E}(Y_t) = \hat{Y}_t ,$$

تخمین یا پیش‌بینی میانگین Y_t، یعنی $\hat{E}(Y_t)$، دقیقاً برابر تخمین یا پیش‌بینی Y_t است. در بند (ج) می‌خواهیم مقدار $\hat{E}(Y_t)$ به ازای $X_t = ۱۰$ را به دست آوریم؛ بنابراین کافی است $\hat{Y}_t$ به ازای $X_t = ۱۰$ را حساب کنیم. از بند (الف) می‌دانیم که

$$\hat{Y}_t = -۳/۵۰٤ + ۰/٤۹٤ X_t ,$$

یا

$$\hat{Y}_t = -۳/۵۰٤ + ۰/٤۹٤ X_t .$$

بنابراین به ازای $X_t = ۱۰$ داریم

$$\hat{Y}_t = -۳/۵۰٤ + ۰/٤۹٤ (۱۰) = ۱/٤۳٦ .$$

د) ابتدا مقدار جدول t را با $n - ۲ = ۲۰ - ۲ = ۱۸$ درجهٔ آزادی و در سطح معنی‌دار $\alpha = ۰/۰۲۵$ به‌دست می‌آوریم، خواهیم داشت $t = \pm ۲/۱۰۱$. فاصلهٔ اطمینان اطمینان برای $E(Y_t)$ عبارت است از:

$$\hat{Y}_f - t_{\alpha/2}\hat{\sigma}\sqrt{\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\Sigma x_i^2}} < E(Y_f) < \hat{Y}_f + t_{\alpha/2}\hat{\sigma}\sqrt{\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\Sigma x_i^2}} ,$$

$$۱/۴۳۶ - ۲/۱۰۱\sqrt{۱/۹۰۷۷}\sqrt{\frac{۱}{۲۰}+\frac{(۱۰-۹/۳۱)^2}{۲۱۵/۴}} < E(Y_f) < ۱/۴۳۶$$

$$+ ۲/۱۰۱\sqrt{۱/۹۰۷۷}\sqrt{\frac{۱}{۲۰}+\frac{(۱۰-۹/۳۱)^2}{۲۱۵/۴}} ,$$

$$۱/۴۳۶ - ۰/۶۶۳ < E(Y_f) < ۱/۴۳۶ + ۰/۶۶۳ ,$$

در نتیجه

$$۰/۷۷۳ < E(Y_f) < ۲/۰۹۹ .$$

۳ـ۳ می‌دانیم که تخمین β در یک مدل رگرسیون ـ که جملهٔ ثابت ندارد ـ عبارت است از

$$\hat{\beta} = \frac{\Sigma X_i Y_i}{\Sigma X_i^2} .$$

می‌خواهیم $\hat{Y}_f^2$ را چنان به دست آوریم که

$$E(\hat{Y}_f^2) = E(Y_f^2 \mid X_f) . \qquad (۱)$$

ابتدا $E(Y_f^2 \mid X_f)$ را بررسی می‌کنیم. می‌دانیم

$$Y_f \mid X_f = \beta X_f + U_f ,$$

در نتیجه خواهیم داشت

$$E(Y_f^2 \mid X_f) = E(\beta X_f + U_f)^2 = E(\beta^2 X_f^2 + U_f^2 + 2\beta X_f U_f) .$$

با توجه به استقلال X_f از U_f و نیز $E(U_f) = ۰$ و همچنین بنابر فرض واریانس همسانی داریم

$$E(Y_t^2 | X_t) = \beta^2 X_t^2 + \sigma_U^2 \qquad (۲)$$

با توجه به مدل رگرسیون مفروض، ملاحظه می‌شود که یک مورد مناسب برای $\hat{Y}_t^2$ مقدار $(\hat{\beta} X_t)^2$ است، اما در زیر نشان می‌دهیم که این تخمین‌زننده اریب دارد.

در یک مدل رگرسیون بدون جمله ثابت، می‌دانیم

$$\hat{\beta} = \beta + \frac{\Sigma X_t U_t}{\Sigma X_t^2} .$$

بنابراین اگر از $(\hat{\beta} X_t)^2$ امید ریاضی بگیریم، خواهیم داشت

$$E(\hat{\beta} X_t)^2 = X_t^2 E(\hat{\beta})^2 = X_t^2 E\left(\beta + \frac{\Sigma X_t U_t}{\Sigma X_t^2}\right)^2 ,$$

$$= X_t^2 E\left[\beta^2 + \frac{(\Sigma X_t U_t)^2}{(\Sigma X_t^2)^2} + 2\beta \frac{\Sigma X_t U_t}{\Sigma X_t^2}\right] .$$

با بسط طرف راست و گرفتن امید ریاضی و با توجه به $E(U_i U_j) = 0$، نتیجه می‌گیریم که

$$E(\hat{\beta} X_t)^2 = X_t^2\left(\beta^2 + \frac{\sigma_U^2}{\Sigma X_t^2}\right) . \qquad (۳)$$

خلاصهٔ راه حل تا اینجا این است که می‌خواستیم Y_t^2 را چنان تخمین بزنیم که معادلهٔ (۱) صادق باشد. دیدیم طرف راست معادلهٔ (۱) با معادلهٔ (۲) برابر است. بنابراین باید $\hat{Y}_t^2$ را چنان بدست آورد که اگر از آن امید ریاضی بگیریم با طرف راست معادله (۲) برابر شود. گفتیم $(\hat{\beta} X_t)^2$ جواب مسأله است. اما دیدیم که این تخمین‌زننده، اریب دارد؛ زیرا امید ریاضی آن، بر طبق معادله (۳)، با سمت راست معادلهٔ (۲) برابر نیست. البته به تعریف «اریب» در عبارتهای فوق توجه داریم ـ که تنها به معنای عدم صدق در رابطهٔ (۱) به کار می‌رود ـ در اینجا باید کوشش کنیم مورد مناسب دیگری به عنوان یک پیش‌بینی‌کننده از Y_t^2 پیدا کنیم که این اریب را نداشته باشد.

با مقایسهٔ معادلات (۲) و (۳) می‌گوییم که پیش‌بینی کننده زیر مطلوب ماست،

$$\hat{Y}_f^2 = E(\hat{\beta} X_f)^2 - \frac{\sigma_U^2 X_f^2}{\Sigma X_i^2} + \sigma_U^2 \quad . \qquad (٤)$$

برای اثبات کافی است به جای $E(\hat{\beta} X_f)^2$ مقدار آن را از معادلهٔ (۳) قرار دهیم؛

$$\hat{Y}_f^2 = \left(\beta^2 X_f^2 + \frac{\sigma_U^2 X_f^2}{\Sigma X_i^2}\right) - \frac{\sigma_U^2 X_f^2}{\Sigma X_i^2} + \sigma_U^2 \quad ,$$

$$= \beta^2 X_f^2 + \sigma_U^2 \quad .$$

مشاهده می‌شود که سمت راست معادله فوق دقیقاً با سمت راست معادله (۲) برابر است. بنابراین می‌توان نوشت:

$$\hat{Y}_f = E(Y_f^2 \mid X_f) \quad ,$$

با توجه به معادلهٔ (۲) و ثابت بودن مقادیر β ، X_f و σ^2، می‌توان نوشت

$$E(\hat{Y}_f^2) = E(Y_f^2 \mid X_f) \quad ,$$

که همان معادلهٔ (۱) است.

بدین ترتیب نشان دادیم که پیش‌بینی کننده‌ای که در معادلهٔ (٤) پیشنهاد شد، جواب مسأله است. در عمل باید به جای σ_U^2 در معادلهٔ (٤)، تخمین آن ($\hat{\sigma}_U^2$) را قرار داد. می‌دانیم تخمین زنندهٔ نااریب از σ^2 عبارت است از:

$$\hat{\sigma}^2 = s^2 = \frac{\Sigma e_i^2}{n-2} \quad .$$

اما یادآوری می‌کنیم که در مدلهای رگرسیون بدون جملهٔ ثابت، باید در مخرج کسر به جای $(n-2)$ مقدار $(n-1)$ را قرار دهیم؛ زیرا برای به دست آوردن Σe_i^2 فقط به $\hat{\beta}$ نیاز داریم و در نتیجه فقط یک درجهٔ آزادی را از دست می‌دهیم.

۳.٤ به داخل پرانتز سمت چپ $\pm E(Y_{T+F})$ را اضافه می‌کنیم. خواهیم داشت

$$E(\hat{Y}_{T+F} - Y_{T+F})^2 = E\left\{[\hat{Y}_{T+F} - E(Y_{T+F})] + [E(Y_{T+F}) - (Y_{T+F})]\right\}^2 ,$$

$$= E\left\{[\hat{Y}_{T+F} - E(Y_{T+F})]^2 + [E(Y_{T+F}) - Y_{T+F}]^2\right.$$

$$\left. + 2[\hat{Y}_{T+F} - E(Y_{T+F})][E(Y_{T+F}) - Y_{T+F}]\right\} .$$

می‌دانیم که

$$E[\hat{Y}_{T+F} - E(Y_{T+F})]^2 = Var^*(Y_{T+F}) ,$$

و منظور از $Var^*(Y_{T+F})$ این است که واریانس در حول مقدار میانگین Y_{T+F} تعریف شده است. همچنین از رابطهٔ

$$Y_{T+F} = \alpha + \beta X_{T+F} + U_{T+F} ,$$

می‌دانیم که

$$E(Y_{T+F}) = \alpha + \beta X_{T+F} ,$$

بنابراین با کم کردن دو رابطهٔ فوق از یکدیگر داریم

$$E(Y_{T+F}) - Y_{T+F} = -U_{T+F} .$$

بدین ترتیب خواهیم داشت

$$E(\hat{Y}_{T+F} - Y_{T+F})^2 = Var^*(Y_{T+F}) + \sigma_U^2$$

$$+ 2E\left\{[\hat{Y}_{T+F} - E(Y_{T+F})][E(Y_{T+F}) - (Y_{T+F})]\right\} .$$

نشان می‌دهیم که در معادلهٔ فوق، جملهٔ سوم سمت راست، برابر صفر است،

$$E\left\{[\hat{Y}_{T+F} - E(Y_{T+F})][E(Y_{T+F}) - (Y_{T+F})]\right\} =$$

$$E[\hat{Y}_{T+F}E(Y_{T+F}) - \hat{Y}_{T+F}Y_{T+F} - E(Y_{T+F})E(Y_{T+F}) + E(Y_{T+F})Y_{T+F}] . (۱)$$

در معادلهٔ ۱۲ـ۳ دیدیم که

$$E(Y_f) = \alpha + \beta X_f \ ,$$

و اگر از معادلهٔ ٤ـ۳ نیز امید ریاضی بگیریم، خواهیم داشت

$$E(\hat{Y}_f) = \alpha + \beta X_f \ .$$

بنابراین داریم

$$E(Y_f) = E(\hat{Y}_f) \ .$$

به همین ترتیب خواهیم داشت

$$E(Y_{T+F}) = E(\hat{Y}_{T+F}) \ , \qquad (۲)$$

و در نتیجه طرف راست معادلهٔ (۱) برابر صفر می‌شود؛ بنابراین می‌توان گفت

$$\mathrm{Var}\,(\hat{Y}_{T+F}) = \mathrm{Var}^*\,(\hat{Y}_{T+F}) + \sigma^2 \ .$$

۳ـ۵ از رابطهٔ $Y_t^* = p_1 + q_1 Y_t$ داریم $\bar{Y}^* = p_1 + q_1 \bar{Y}$، و در نتیجه خواهیم داشت

$$y_t^* = q_1 y_t \ . \qquad (۱)$$

به همین ترتیب می‌توان نشان داد که

$$x_t^* = q_2 x_t \ . \qquad (۲)$$

می‌دانیم در مدل رگرسیون

$$Y_t^* = a + b X_t^* + U_t^* \ ,$$

ضریب تعیین (r^{*2}) عبارت است از

$$r^{*2} = \frac{(\Sigma x_t^* y_t^*)^2}{\Sigma x_t^{*2} \Sigma y_t^{*2}} \ .$$

با جایگزینی رابطه (۱) و (۲) در معادله فوق خواهیم داشت

$$r^{*2} = \frac{(\Sigma\, q_2 x_t q_1 y_t)^2}{\Sigma\,(q_2 x_t)^2 \Sigma\,(q_1 y_t)^2} ,$$

$$= \frac{q_1^2 q_2^2 (\Sigma\, x_t y_t)^2}{q_1^2 q_2^2 \Sigma\, x_t^2 \Sigma\, y_t^2} = r^2 .$$

۳-۶ پاسخ به این سؤال در متن کتاب و در بحث از معادلهٔ ۳-۴۵ ارائه شده است. با وجود این، در اینجا به راه‌حل دیگری اشاره می‌کنیم. می‌دانیم در مدل رگرسیون

$$Y_t = \alpha + \beta X_t + U_t , \qquad (۱)$$

و با توجه به معادلهٔ ۳-۱۵ داریم

$$r^2 = \frac{\hat{\beta} \Sigma\, x_t y_t}{\Sigma\, y_t^2} = \frac{\hat{\beta}^2 \Sigma\, x_t^2}{\Sigma\, y_t^2} .$$

هرگاه $r^2 = 1$، آنگاه خواهیم داشت

$$\hat{\beta}^2 = \frac{\Sigma\, y_t^2}{\Sigma\, x_t^2} ,$$

یا

$$\Sigma\, x_t^2 = \frac{\Sigma\, y_t^2}{\hat{\beta}^2} . \qquad (۲)$$

اما برای مدل (۱) داریم

$$\hat{\beta} = \frac{\Sigma\, x_t y_t}{\Sigma\, x_t^2} ,$$

که با جایگزینی معادلهٔ (۲) خواهیم داشت

$$1 = \hat{\beta} \frac{\Sigma\, x_t y_t}{\Sigma\, x_t^2} ,$$

یا

$$\hat{\beta} = \frac{\Sigma\, y_t^2}{\Sigma\, x_t y_t} . \qquad (۳)$$

اما برای مدل رگرسیون معکوس $X_t = \alpha' + \beta' Y_t + V_t$ داریم

$$\hat{\beta}' = \frac{\Sigma x_t y_t}{\Sigma x_t^2} ,$$

یا

$$\frac{1}{\hat{\beta}'} = \frac{\Sigma y_t^2}{\Sigma x_t y_t} . \qquad (٤)$$

با مقایسهٔ معادله‌های (۳) و (٤) نتیجه می‌گیریم که

$$\hat{\beta} = \frac{1}{\hat{\beta}'} . \qquad ۳.۷$$

$$\Sigma x_i^2 = \Sigma X_i^2 - \frac{1}{n} (\Sigma X_i)^2 ,$$

$$= ۱۲/۱۶ - \frac{1}{۲۰۰} (۱۱/۳٤)^2 = ۱۱/۵۱۷ ,$$

$$\Sigma y_i^2 = \Sigma Y_i^2 - \frac{1}{n} (\Sigma Y_i)^2 ,$$

$$= ۸٤/۹٦ - \frac{1}{۲۰۰} (۲۰/۷۲)^2 = ۸۲/۸۱۳ ,$$

$$\Sigma x_i y_i = \Sigma X_i Y_i - \frac{1}{n} (\Sigma X_i)(\Sigma Y_i) ,$$

$$= ۲۲/۱۳ - \frac{1}{۲۰۰} (۱۱/۳٤)(۱۲/۱٦) = ۲۰/۹۵۵ .$$

با محاسبات فوق به راحتی می‌توان β و α را تخمین زد،

$$\hat{\beta} = \frac{\Sigma x_i y_i}{\Sigma x_i^2} = ۱/۸۱۹ \quad , \quad \hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X} = ۰/۰۰۰٤ .$$

پارامترهای رگرسیون معکوس را به ترتیب زیر تخمین می‌زنیم،

$$\hat{\beta}' = \frac{\Sigma x_i y_i}{\Sigma y_i^2} = ۰/۲۵۳۰ \quad , \quad \hat{\alpha}' = \bar{X} - \hat{\beta}' \bar{Y} = ۰/۰۳۰ .$$

۳.۸ می‌دانیم برای مدل رگرسیون مستقیم داریم

$$\hat{\beta} = \frac{\Sigma X_t Y_t}{\Sigma X_t^2} .$$

برای تخمین β' در مدل رگرسیون معکوس $X_t = \beta' Y_t + V_t$ خواهیم داشت

$$\hat{\beta}' = \frac{\Sigma X_t Y_t}{\Sigma Y_t^2} .$$

فرق $\hat{\beta}$ و $\hat{\beta}'$ فقط در مخرج کسر است. اگر $\Sigma X_t^2 > \Sigma Y_t^2$، آنگاه $\hat{\beta} < \hat{\beta}'$. می‌دانیم r^2 در مدل رگرسیون مستقیم برابر است با

$$r^2 = \frac{(\Sigma X_t Y_t)^2}{\Sigma X_t^2 \Sigma Y_t^2} ,$$

و برای رگرسیون معکوس نیز دقیقاً همین مقدار خواهد بود؛ بنابراین مقادیر r^2 با یکدیگر برابر است. برای تخمین واریانس جملهٔ اختلال یا $\hat{\sigma}^2$، باید ابتدا مجموع مربعات پسماند را در هر یک از دو مدل به دست آورد. برای رگرسیون مستقیم داریم،

$$r^2 = 1 - \frac{\Sigma e_t^2}{\Sigma Y_t^2} ,$$

$$\frac{(\Sigma X_t Y_t)^2}{\Sigma X_t^2 \Sigma Y_t^2} = 1 - \frac{\Sigma e_t^2}{\Sigma Y_t^2} ,$$

در نتیجه:

$$\Sigma e_t^2 = \frac{\Sigma X_t^2 \Sigma Y_t^2 - (\Sigma X_t Y_t)^2}{\Sigma X_t^2} .$$

برای تخمین σ^2 کافی است Σe_t^2 را بر درجات آزادی، یعنی $(n-1)$ تقسیم کنیم،

$$\hat{\sigma}^2 = \frac{\Sigma X_t^2 \Sigma Y_t^2 - (\Sigma X_t Y_t)^2}{(n-1) \Sigma X_t^2} .$$

به روشی کاملاً مشابه می‌توان $\hat{\sigma}^2$ را برای مدل رگرسیون معکوس به دست آورد.

خواهیم داشت

$$\hat{\sigma}^{-2} = \frac{\Sigma X_t^2 \Sigma Y_t^2 - (\Sigma X_t Y_t)^2}{(n-1)\Sigma X_t^2} .$$

ملاحظه می‌شود تفاوت $\hat{\sigma}^2$ در مدلهای مستقیم و معکوس فقط در مخرج کسر است. اگر داشته باشیم $\Sigma X_t^2 > \Sigma Y_t^2$ آنگاه $\hat{\sigma}^{-2} < \hat{\sigma}^2$ ، یعنی تخمین واریانس جملهٔ اختلال در مدل رگرسیون مستقیم، از مقدار مشابه در مدل معکوس کمتر خواهد بود.

۹ـ۳ الف) از دو طرف لگاریتم می‌گیریم،

$$\ln Y_t = \ln \alpha + \beta X_t + \ln U_t ,$$

و با تعریف جدیدی از متغیرها خواهیم داشت

$$Y_t^* = \alpha^* + \beta X_t + U_t^* ,$$

که برحسب پارامتر β و $\ln \alpha$ خطی است. به هر حال هر دو پارامتر را می‌توان با روش حداقل مربعات معمولی تخمین زد و برای رسیدن به تخمین α کافی است از $\hat{\alpha}^*$، آنتی‌لگاریتم بگیریم.

ب) این مدل را نمی‌توان خطی کرد زیرا U_t در مدل، خصوصیت جمع‌پذیری دارد.

ج) از دو طرف لگاریتم می‌گیریم،

$$\ln Y_t = \alpha + \beta X_t + U_t ,$$

که یک مدل خطی برحسب α و β است.

د) با اینکه در ظاهر به نظر می‌رسد که این مدل را نمی‌توان خطی کرد، چون U_t جمع‌پذیر است، امّا با استفاده از روش «ثابت نگه داشتن یک پارامتر» ـ که در بند ۵ قسمت ۳ـ۶ ارائه شد ـ بسادگی می‌توان مدل مزبور را خطی کرد. ابتدا مجموع مربعات پسماند را تشکیل می‌دهیم،

$$\Sigma e_t^2 = \left[Y_t - \frac{\hat{\alpha}}{(\hat{\beta} - X_t)} \right]^2 . \qquad (۱)$$

با ثابت نگه داشتن $\hat{\beta}$، متغیر Z_t را به صورت زیر تعریف می‌کنیم،

$$Z_t = \frac{1}{\hat{\beta} - X_t} \quad . \qquad (۲)$$

اگر در محدودهٔ تغییرات $\hat{\beta}$، یک مقدار اختیاری انتخاب کرده و در معادلهٔ (۲) قرار دهیم و Z_t به دست آمده را در معادلهٔ (۱) بگذاریم، خواهیم داشت

$$\sum e_t^2 = (Y_t - \hat{\alpha}\, Z_t)^2 \quad .$$

به راحتی می‌توان مقداری از $\hat{\alpha}$ که $\sum e_t^2$ را حداقل می‌کند به دست آورد. با تغییر مقدار $\hat{\beta}$ در معادلهٔ (۲) می‌توان به مقادیر جدیدی از Z_t و $\sum e_t^2$ رسید. اگر تمام مقادیر ممکن $\hat{\beta}$ را بدین ترتیب آزمایش کنیم، به تمام مقادیر ممکن $\sum e_t^2$ دست یافته و سپس کمترین آن را مشخص می‌کنیم. $\hat{\alpha}$ و $\hat{\beta}$ متناظر با این مقدار از $\sum e_t^2$، تخمین‌زننده‌های روش حداقل مربعات معمولی از پارامترهای مدل مفروض خواهد بود. بدیهی است راه‌حل فوق، مستلزم شناخت دقیقی از محدودهٔ تغییرات β است.

۱۰-۳ الف) از دو طرف معادله لگاریتم می‌گیریم،

$$\ln Y_t = \ln \alpha + \beta \ln X_t \quad .$$

متغیرهای زیر را تعریف می‌کنیم،

$$Y_t^* = \ln Y_t \quad , \quad \alpha^* = \ln \alpha \quad , \quad X_t^* = \ln X_t \quad ,$$

بنابراین خواهیم داشت

$$Y_t^* = \alpha^* + \beta X_t^* \quad .$$

اگر جملهٔ اختلال، جمع‌پذیر باشد، روش فوق موفقیت‌آمیز نبوده و باید از روشهای تخمین غیرخطی استفاده کرد.

ب) از دو طرف معادله، لگاریتم می‌گیریم،

$$\ln Y_t = \ln \alpha + \beta X_t \quad .$$

متغیرهای زیر را تعریف می‌کنیم،

$$Y_t^* = \ln Y_t \quad , \quad \alpha^* = \ln \alpha \ ,$$

بنابراین خواهیم داشت

$$Y_t^* = \alpha^* + \beta X_t^* \ .$$

اگر U_t جمع‌پذیر باشد، باید از روشهای تخمین غیرخطی استفاده کرد.

ج) تعریف می‌کنیم $X_t^* = \ln X_t$. با جایگزینی در مدل اولیه داریم،

$$Y_t = \alpha^* + \beta X_t^* \ .$$

U_t می‌تواند جمع‌پذیر باشد.

د) دو طرف مدل مفروض را معکوس می‌کنیم،

$$\frac{1}{Y_t} = \alpha - \frac{\beta}{X_t} \ .$$

این متغیرها را تعریف می‌کنیم،

$$Y_t^* = \frac{1}{Y_t} \quad , \quad X_t^* = \frac{1}{X_t} \ ,$$

در نتیجه خواهیم داشت

$$Y_t^* = \alpha + \beta X_t^* \ .$$

اگر U_t جمع‌پذیر باشد آنگاه روش فوق قابل اجرا نبوده و باید از روشهای تخمین غیرخطی استفاده کرد.

ه) ابتدا $\frac{Y_t}{1 - Y_t}$ را تشکیل می‌دهیم،

$$\frac{Y_t}{1 - Y_t} = \frac{e^{(\alpha + \beta X_t)} \Big/ [1 + e^{(\alpha + \beta X_t)}]}{1 - e^{(\alpha + \beta X_t)} \Big/ [1 + e^{(\alpha + \beta X_t)}]} \ ,$$

$$= e^{(\alpha + \beta X_t)} \ .$$

از دو طرف معادله لگاریتم می‌گیریم،

$$\ln\left(\frac{Y_t}{1-Y_t}\right) = \alpha + \beta X_t \ .$$

متغیر زیر را تعریف می‌کنیم.

$$Y_t^* = \frac{Y_t}{1-Y_t} \ ,$$

در نتیجه خواهیم داشت

$$Y_t^* = \alpha + \beta X_t \ .$$

اگر U_t جمع‌پذیر باشد، باید از روشهای تخمین غیرخطی استفاده کرد.

و) تعریف می‌کنیم $X_t^* = \sqrt{X_t}$، در این صورت خواهیم داشت

$$Y_t = \alpha + \beta X_t^* \ .$$

جملهٔ اختلال (U_t) می‌تواند جمع‌پذیر باشد.

ز) در تخمین این مدل باید از روشهای تخمین غیرخطی استفاده کرد، مگر آنکه محدودهٔ تغییرات β معلوم باشد. در این صورت، مشابه بند (د) مسألهٔ ۹ـ۳ عمل می‌کنیم. برای هر مقدار معلوم β، به راحتی می‌توان این کمیت را محاسبه کرد،

$$X_t^* = e^{\beta X_t} \ .$$

فرض می‌کنیم U_t جمع‌پذیر باشد. با جایگذاری معادلهٔ فوق در مدل مفروض داریم،

$$Y_t = \alpha + X_t^* + U_t \ ,$$

$$(Y_t - X_t^*) = \alpha + U_t \ ,$$

در نتیجه

$$Y_t^* = \alpha + U_t \ , \qquad (۱)$$

که در آن $Y_t^* = Y_t - X_t^*$ است. ملاحظه می‌شود که در مدل (۱)، Y_t^* فقط تابعی از α و U_t است. در مسألهٔ ۱۰ـ۱ دیدیم که در این گونه مدلهای رگرسیون تخمین α با میانگین متغیر درون‌زا برابر است،

$$\hat{\alpha} = \bar{Y}^* .$$

خلاصهٔ بحث این است که به ازای مقدار معیّنی از β، مدل مفروض را به یک مدل خطی تبدیل کرده و $\hat{\alpha}$ را تخمین زده‌ایم. حال باید مقدار $\sum e_t^2$ را به دست آوریم،

$$\sum e_t^2 = \sum (Y_t^* - \hat{\alpha})^2 ,$$

$$= \sum (Y_t^* - \bar{Y}^*)^2 = \sum y_t^{*2} .$$

به همین ترتیب از مقادیر ممکن دیگر β نیز استفاده کرده، با به دست آوردن تخمینهایی از $\hat{\alpha}$ به مقادیر ممکن $\sum e_t^2$ می‌رسیم. کمترین مقدار $\sum e_t^2$ رامشخص کرده، مقادیر $\hat{\alpha}$ و $\hat{\beta}$ متناظر با آن، جواب مسأله خواهد بود.

ک) این مسأله شبیه مثالی است که در بند ۵ قسمت ۶ـ۳ و با عنوان «مدلهای غیرخطی برحسب پارامترها» مطرح شده است.

فصل چهارم

مدل رگرسیون خطی با دو متغیر توضیحی

۴-۱ مقدمه

تاکنون مدلی را مطالعه کرده‌ایم که فقط یک متغیر توضیحی داشته است. در مطالعات کاربردی معمولاً باید بیش از یک متغیر توضیحی را در نظر گرفت؛ زیرا بیان تغییرات یک متغیر درون‌زا به کمک چند متغیر توضیحی به مراتب دقیقتر خواهد بود. ورود بیش از یک متغیر توضیحی در مدلهای رگرسیون، مسائل مختلفی را مطرح می‌کند. برای اینکه بتوان این مسائل را به نحو منظمی بررسی کرد، ابتدا مدلی را در نظر می‌گیریم که فقط دو متغیر توضیحی دارد و سپس در فصل پنجم این مدل را به k متغیر توضیحی تعمیم می‌دهیم. معمولاً به مدلی که بیش از یک متغیر توضیحی دارد، مدل «رگرسیون چند متغیره»[1] می‌گوییم. بنابراین مدلهای رگرسیون با دو متغیر توضیحی در واقع ساده‌ترین مدل رگرسیون چند متغیره است. این مدل را می‌توان چنین نوشت،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \quad .$$

تعمیم این مدل به حالت عمومی مدلهای رگرسیون چند متغیره که این است که به جای دو متغیر از k متغیر توضیحی استفاده کنیم،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + U_t \quad .$$

تخمین یک مدل رگرسیون چند متغیره با k متغیر توضیحی، به کمک جبر ماتریسی بسیار ساده‌تر انجام می‌شود، بنابراین بررسی حالت عمومی مدلهای رگرسیون

---

1. Multiple Regression

با استفاده از ماتریسها[۱]، موضوع فصل پنجم خواهد بود.

تخمین پارامترهای α، β_1 و β_2 در مدل رگرسیون با دو متغیر توضیحی موضوع قسمت ۲ـ٤ است. در این قسمت مسائل ضریب تعیین نیز بررسی خواهد شد. تحلیلهای آماری تخمین پارامترها شامل واریانس، کوواریانس و آزمون فرضیه‌های مختلف، موضوع قسمت ۳ـ٤ است. با اینکه جدول تجزیه واریانس و آزمون F، در تحلیلهای رگرسیون از اهمیت ویژه‌ای برخوردار است، اما به نظر می‌رسد بهتر است مباحث تجزیه واریانس در مدلهای رگرسیون با دو متغیر توضیحی را به صورت مورد خاص از تجزیه واریانس در حالت کلی در نظر بگیریم؛ بنابراین در فصل ششم و در مطالعهٔ تجزیه واریانس در رگرسیون چند متغیره، ابعاد این مسأله در مورد رگرسیونهایی با دو متغیر توضیحی نیز روشن خواهد شد.

مسألهٔ پیش‌بینی، همواره یکی از مهمترین مباحث تخمین در مدلهای رگرسیون بوده است. روال کلی بحث در یک مدل رگرسیون با دو متغیر توضیحی دقیقاً شبیه رگرسیون ساده است؛ با این تفاوت که در محاسبهٔ خطای پیش‌بینی، باید از واریانس و کوواریانس تمام پارامترها استفاده شود، در قسمت ٤ـ٤ به طور خلاصه مسأله پیش‌بینی را بررسی خواهیم کرد.

هنگامی که یک مدل رگرسیون ساده را تعمیم می‌دهیم و به جای یک متغیر توضیحی، از دو یا چند متغیر توضیحی استفاده می‌کنیم، مسألهٔ جدیدی مطرح می‌شود و آن ضرایب همبستگی جزئی و ضرایب تعیین جزئی است که در تحلیل و تفسیر نتایج حاصل از تخمین پارامترهای مدل، اهمیت فراوان دارد. آیا نمی‌توان رگرسیون چند متغیره را به چند رگرسیون ساده تجزیه کرد؟ به عبارت دیگر آیا نمی‌توان یک بار به کمک Y_i و X_{1i} و بار دیگر با Y_i و X_{2i} دو رگرسیون جداگانه ساخت و پارامترهای هر یک را تخمین زد؟ اساساً ورود همزمان X_{1i} و X_{2i} به عنوان متغیرهای توضیحی، چه نکته‌هایی را می‌تواند روشن کند که دو مدل رگرسیون ساده مجزا از هم نمی‌تواند آن را

---

۱. پیوست «۵ ـ الف»

تبیین کند؟ اینها نکته‌هایی است که در قسمت ٤ـ٥ و با عنوان ضرایب همبستگی و ضرایب تعیین بررسی خواهد شد. در این قسمت نشان خواهیم داد که چگونه می‌توان ضریب تعیین کلی یک مدل رگرسیون را بر حسب ضرایب همبستگی جزئی نوشت. همچنین به این نکته نیز توجه خواهیم کرد که چگونه تخمین پارامترهای یک مدل رگرسیون چند متغیره را می‌توان بر حسب تخمین پارامترهای رگرسیون ساده یا برحسب ضرایب همبستگی ساده بیان کرد.

٤ـ٢ تخمین مدل

مدل رگرسیون زیر مفروض است،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \quad , \qquad (٤\text{ـ}١)$$

که در آن $t = ١ , ٢ , \ldots , n$ و Y_t متغیر درون‌زا و X_{1t} و X_{2t} متغیرهای برون‌زا است. مانند گذشته فرض بر این است که U_t یک متغیر تصادفی است. و تمام فرضهای کلاسیک را به این شرح داراست.

۱. میانگین آن صفر است؛ یعنی $E(U_t) = ٠$.

۲. واریانس همسانی دارد؛ یعنی واریانس آن به ازای تمام مقادیر t ثابت است، $Var(U_t) = \sigma^2$.

۳. خودهمبستگی ندارد؛ یعنی U_t و U_s به ازای تمام مقادیر $t \neq s$ از یکدیگر مستقل است؛ یعنی $Cov(U_t , U_s) = ٠$.

٤. به ازای تمام مقادیر t، تابع توزیع احتمال U_t نرمال است.

فرضهای فوق در مورد جملهٔ اختلال است. با ارائه فرضهای پنجم و ششم در مورد متغیرهای توضیحی، فرضهای کلاسیک مدلهای رگرسیون با بیش از یک متغیر توضیحی کامل می‌شود.

۵. متغیر توضیحی X_t غیرتصادفی است؛ یعنی X_t از جملهٔ اختلال مستقل است.

با استفاده از ٥ فرض فوق و دقیقاً مانند مباحث فصل دوم می‌توان نشان داد که $\hat{\alpha}$، $\hat{\beta}_1$ و

$\hat{\beta}_۲$ نااریب است و حداقل واریانس را در گروه تخمین زننده‌های خطی دارد. فرض ششم که در ذیل مطرح خواهد شد، فقط به مدلهای رگرسیون چند متغیره اختصاص دارد.

۶. متغیرهای توضیحی «همخطی»[1] ندارند؛ یعنی هیچ نوع همبستگی کامل خطی بین $X_۱$ و $X_۲$ وجود ندارد؛ به عبارت دیگر نمی‌توان هیچگونه رابطهٔ خطی بین متغیرهای برون‌زا برقرار کرد. مسألهٔ همخطی از مباحث بسیار مهم اقتصادسنجی است و در جلد دوم این کتاب بررسی خواهد شد. اما برای تبیین بیشتر به ذکر مثالی اکتفا می‌کنیم.

در مدل رگرسیون ۱ـ٤، این رابطهٔ خطی بین متغیرهای توضیحی مفروض است:

$$۳ X_{۱t} + X_{۲t} = ۵ \quad .$$

می‌توان $X_{۲t}$ را بر حسب $X_{۱t}$ نوشت،

$$X_{۲t} = ۵ - ۳ X_{۱t} \quad .$$

با جایگزینی معادلهٔ فوق در مدل رگرسیون ۱ـ٤ خواهیم داشت

$$Y_t = \alpha + \beta_۱ X_{۱t} + \beta_۲ (۵ - ۳ X_{۱t}) + U_t \quad ,$$

یا

$$Y_t = (\alpha + ۵ \beta_۲) + (\beta_۱ - ۳ \beta_۲) X_{۱t} + U_t \quad .$$

با تخمین مدل فوق آنچه به دست خواهیم آورد، تخمینهای $(\alpha + ۵ \beta_۲)$ و $(\beta_۱ - ۳ \beta_۲)$ است؛ در حالی که هدف ما این بود که α، $\beta_۱$ و $\beta_۲$ را تخمین بزنیم. مشاهده می‌شود که در موارد «همخطی کامل»[2]، یعنی وجود یک رابطهٔ معین و مشخص خطی بین متغیرها، تخمین هر یک از پارامترهای مدل ممکن نیست، مگر اینکه بتوان به کمک نظریه‌های اقتصادی، رابطهٔ دیگری نیز بین پارامترها برقرار کرد. در مثال فوق فقط دو تخمین $(\hat{\alpha} + ۵ \hat{\beta}_۲)$ و $(\hat{\beta}_۱ - ۳ \hat{\beta}_۲)$ را به دست می‌آوریم؛ در حالی که سه مجهول $\hat{\alpha}$، $\hat{\beta}_۱$ و $\hat{\beta}_۲$ داریم. باید یک رابطهٔ خطی دیگر به کمک نظریه‌های اقتصادی بین پارامترها به دست آورد تا تخمینهای جداگانه از آنها ممکن شود. بدیهی است در بسیاری موارد چنین

1. Multicolinearity 2. Perfect Multicolinearity

روابطی را نمی‌توان از نظریه‌های اقتصادی استنتاج کرد؛ بنابراین مسألهٔ همخطی از مباحث قابل توجه در اقتصادسنجی است که ابعاد مختلف آن در جلد دوم این کتاب بررسی خواهد شد.

در مواردی (مانند مثال فوق) که همبستگی خطی بین متغیرهای توضیحی کامل باشد؛ یعنی «همخطی کامل»، داشته باشیم، ضریب همبستگی بین X_{1t} و X_{2t} برابر ± 1 است. در این فصل فرض بر این است که بین متغیرهای توضیحی همبستگی وجود دارد، اما کامل نیست. حال که با فرض ششم (مختص رگرسیونهای چند متغیره) آشنا شدیم به تخمین مدل ۱ـ٤ می‌پردازیم. فرض کنید مدل ۱ـ٤ را به صورت زیر تخمین زده‌ایم،

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta}_1 X_{1t} + \hat{\beta}_2 X_{2t} \quad . \qquad (٤ـ٢)$$

پسماند را به روال گذشته تعریف می‌کنیم،

$$e_t = Y_t - \hat{Y}_t \quad ,$$

و مجموع مربعات پسماند را می‌نویسیم،

$$\sum_{t=1}^{n} e_t^2 = \sum_{t=1}^{n} (Y_t - \hat{\alpha} - \hat{\beta} X_{1t} - \hat{\beta}_2 X_{2t})^2 \qquad (٤ـ٣)$$

معیار روش حداقل مربعات معمولی این است که $\hat{\alpha}$ ، $\hat{\beta}_1$ و $\hat{\beta}_2$ را چنان تعیین کنیم که Σe_t^2 در معادلهٔ ٤ـ٣ حداقل شود؛ بنابراین باید مشتق Σe_t^2 را نسبت به متغیرهای $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ به دست آوریم و مساوی صفر قرار دهیم،

$$\frac{\partial \Sigma e_t^2}{\partial \hat{\alpha}} = \sum 2 (Y_t - \hat{\alpha} - \hat{\beta} X_{1t} - \hat{\beta}_2 X_{2t}) (-1) = 0 \quad ,$$

$$\frac{\partial \Sigma e_t^2}{\partial \hat{\beta}_1} = \sum 2 (Y_t - \hat{\alpha} - \hat{\beta}_1 X_{1t} - \hat{\beta}_2 X_{2t}) (-X_{1t}) = 0 \quad , \qquad (٤ـ٤)$$

$$\frac{\partial \Sigma e_t^2}{\partial \hat{\beta}_2} = \sum 2 (Y_t - \hat{\alpha} - \hat{\beta}_1 X_{1t} - \hat{\beta}_2 X_{2t}) (-X_{2t}) = 0 \quad .$$

به معادله‌های ٤ـ٤ سیستم معادلات نرمال می‌گویند. این معادله‌ها را می‌توان به شرح زیر ساده نمود،

$$\sum Y_i = n\hat{\alpha} + \hat{\beta}_1 \sum X_{1i} + \hat{\beta}_2 \sum X_{2i} ,$$

$$\sum X_{1i} Y_i = \hat{\alpha} \sum X_{1i} + \hat{\beta}_1 \sum X_{1i}^2 + \hat{\beta}_2 \sum X_{1i} X_{2i} , \qquad (٤ـ٥)$$

$$\sum X_{2i} Y_i = \hat{\alpha} \sum X_{2i} + \hat{\beta}_1 \sum X_{1i} X_{2i} + \hat{\beta}_2 \sum X_{2i}^2 .$$

سیستم سه معادلهٔ نرمال ٥ـ٤ را ـ که بر حسب سه مجهول $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ است ـ می‌توان به سهولت حل کرد و تخمینهای حداقل مربعات معمولی از پارامترهای α، β_1 و β_2 را به دست آورد.

یک روش ساده وجود دارد که می‌توان سیستم معادله‌های نرمال ٥ـ٤ را بدون مشتق‌گیری از Σe_i^2 نوشت. معادلهٔ ١ـ٤ را یک بار دیگر می‌نویسیم،

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i .$$

از دو طرف معادله فوق Σ گرفته و جملهٔ آخر، یعنی ΣU_i را مساوی صفر فرض می‌کنیم[1]. بدین ترتیب معادلهٔ اول نرمال در سیستم معادله‌های ٥ـ٤ به دست می‌آید، با این تفاوت که به جای $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ مقادیر α، β_1 و β_2 را خواهیم داشت. در ادامهٔ بحث، اشاره خواهیم کرد که این امر مسألهٔ خاصی را به وجود نمی‌آورد. یک بار دیگر دو طرف معادلهٔ ١ـ٤ را در X_{1i} ضرب کرده و Σ می‌گیریم و با حذف جملهٔ آخر، یعنی $\Sigma X_{1i} U_i$، به دومین معادلهٔ نرمال سیستم ٥ـ٤ می‌رسیم. سرانجام دو طرف معادلهٔ ١ـ٤ را در X_{2i} ضرب کرده و با گرفتن Σ و حذف جملهٔ آخر به سومین معادله نرمال خواهیم رسید. اگر به جای دو متغیر توضیحی، k متغیر توضیحی نیز داشتیم، به همین ترتیب، متغیرهای توضیحی دیگر را نیز یک به یک در دو طرف معادلهٔ ١ـ٤ ضرب کرده و به ترتیب فوق معادله‌های نرمال دیگر حاصل می‌شود. سیستم معادلاتی را ـ که به این ترتیب و بر حسب α، β_1 و

---

١. در مسألهٔ ١١ـ١ دیدیم که ضروری نیست ΣU_i برابر صفر باشد، هر چند $E(U_i)$ صفر است.

$\beta_۲$ به دست می‌آید ـ حل می‌کنیم. جوابهای حاصل $\hat{\alpha}$، $\hat{\beta}_۱$ و $\hat{\beta}_۲$ خواهد بود. این راه‌حل بیشتر شبیه روشی برای از بر کردن معادله‌های نرمال است و مبنای نظری ندارد؛ با وجود این، کاربردهای دیگر این روش را در جلد دوم یا عنوان روش تخمین «متغیرهای ابزاری»[1] بررسی خواهیم کرد.

سیستم معادله‌های ٤ـ٥ را می‌توان به صورت ساده‌تری نیز نوشت. با تقسیم دو طرف معادلهٔ اول نرمال بر n، خواهیم داشت:

$$\bar{Y} = \hat{\alpha} + \hat{\beta}_۱ \bar{X}_۱ + \hat{\beta}_۲ \bar{X}_۲ \ . \qquad (٤ـ٦)$$

با جایگزینی مقدار $\hat{\alpha}$ از معادلهٔ ٤ـ٦ در معادله دوم نرمال در سیستم معادله‌های ٤ـ٥ خواهیم داشت:

$$\sum X_{۱t} Y_t = n \bar{X}_۱ (\bar{Y} - \hat{\beta}_۱ \bar{X}_۱ - \hat{\beta}_۲ \bar{X}_۲) + \hat{\beta}_۱ \sum X_{۱t}^۲ + \hat{\beta}_۲ \sum X_{۱t} X_{۲t} \ . \qquad (٤ـ٧)$$

برای ساده‌تر کردن معادلهٔ ٤ـ٧، متغیرهای زیر را تعریف می‌کنیم،

$$\sum x_{۱t}^۲ = \sum (X_{۱t} - \bar{X}_۱)^۲ = \sum X_{۱t}^۲ - n \bar{X}_۱^۲ \ ,$$

$$\sum x_{۱t} x_{۲t} = \sum (X_{۱t} - \bar{X}_۱)(X_{۲t} - \bar{X}_۲) = \sum X_{۱t} X_{۲t} - n \bar{X}_۱ \bar{X}_۲ \ ,$$

$$\sum x_{۲t}^۲ = \sum (X_{۲t} - \bar{X}_۲)^۲ = \sum X_{۲t}^۲ - n \bar{X}_۲^۲ \ ,$$

$$\sum x_{۱t} y_t = \sum (X_{۱t} - \bar{X}_۱)(Y_t - \bar{Y}) = \sum X_{۱t} Y_t - n \bar{X}_۱ \bar{Y} \ ,$$

$$\sum x_{۲t} y_t = \sum (X_{۲t} - \bar{X}_۲)(Y_t - \bar{Y}) = \sum X_{۲t} Y_t - n \bar{X}_۲ \bar{Y} \ ,$$

$$\sum y_t^۲ = \sum (Y_t - \bar{Y})^۲ = \sum Y_t^۲ - n \bar{Y}^۲ \ .$$

با استفاده از تعاریف فوق، معادلهٔ ٤ـ٧ را می‌توان به صورت زیر نوشت،

---

1. Instrumental Variables

$$\sum x_{1t} y_t = \hat{\beta}_1 \sum x_{1t}^2 + \hat{\beta}_2 \sum x_{1t} x_{2t} \quad . \qquad (٤\text{ـ}٨)$$

با روشی کاملاً مشابه، معادلهٔ سوم نرمال را نیز به صورت زیر می‌نویسیم،

$$\sum x_{2t} y_t = \hat{\beta}_1 \sum x_{1t} x_{2t} + \hat{\beta}_2 \sum x_{1t}^2 \qquad (٤\text{ـ}٩)$$

معادله‌های ٤ـ٨ و ٤ـ٩ را می‌توان برای $\hat{\beta}_1$ و $\hat{\beta}_2$ حل کرد،

$$\hat{\beta}_1 = \frac{\sum x_{2t}^2 \sum x_{1t} y_t - \sum x_{1t} x_{2t} \sum x_{2t} y_t}{\sum x_{1t}^2 \sum x_{2t}^2 - (\sum x_{1t} x_{2t})^2}$$

$$\hat{\beta}_2 = \frac{\sum x_{1t}^2 \sum x_{2t} y_t - \sum x_{1t} x_{2t} \sum x_{1t} y_t}{\sum x_{1t}^2 \sum x_{2t}^2 - (\sum x_{1t} x_{2t})^2} \qquad (٤\text{ـ}١٠)$$

با به دست آوردن $\hat{\beta}_1$ و $\hat{\beta}_2$ می‌توان $\hat{\alpha}$ را از معادلهٔ ٤ـ٦ به دست آورد.

به طور خلاصه برای تخمین پارامترهای مدل ٤ـ١ به این ترتیب عمل می‌کنیم.

۱. مقادیر میانگینها را حساب می‌کنیم: $\bar{X}_1$ ، $\bar{X}_2$ ، $\bar{Y}$.

۲. مجموع مربعات و مجموع حاصلضرب تمام متغیرها را به دست می‌آوریم:

$\sum X_{1t}^2$ ، $\sum X_{1t} X_{2t}$ و ... .

۳. مقادیر $\sum x_{1t}^2$ ، $\sum x_{1t} x_{2t}$ ، $\sum x_{2t}^2$ ، $\sum x_{1t} y_t$ ، $\sum x_{2t} y_t$ و $\sum y_t^2$ را محاسبه می‌کنیم.

٤. با استفاده از معادله‌های ٤ـ١٠ مقادیر $\hat{\beta}_1$ و $\hat{\beta}_2$ را به دست می‌آوریم.

٥. از معادلهٔ ٤ـ٦ مقدار $\hat{\alpha}$ را حساب می‌کنیم.

بعد از آشنایی با نحوهٔ تخمین پارامترهای مدل ٤ـ١ به چند نکته اشاره می‌شود.

الف) با توجه به معادلهٔ ٤ـ٦ می‌توان گفت که صفحهٔ تخمین رگرسیون از نقطهٔ میانگینها می‌گذرد.

ب) اگر معادلهٔ ٤ـ٢ را برای تمام مقادیرنمونه جمع و بر n تقسیم‌کنیم، خواهیم داشت:

$$\bar{Y} = \hat{\alpha} + \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 \quad .$$

با مقایسهٔ معادله فوق با معادلهٔ ٤ـ٦ می‌توان نتیجه گرفت که

$$\hat{Y} = \bar{Y} \quad . \qquad (٤\text{ـ}١١)$$

ج) می‌دانیم $Y_t = \hat{Y}_t + e_t$. دو طرف معادله را برای تمام مقادیر t جمع کرده و بر n تقسیم می‌کنیم،

$$\bar{Y} = \bar{\hat{Y}} + \bar{e} \quad ,$$

که با توجه به معادلهٔ ٤ـ١١ نتیجه می‌گیریم که

$$\Sigma e_t = ٠ \quad \text{یا} \quad \bar{e} = ٠ \quad . \qquad (٤\text{ـ}١٢)$$

معادله‌های ٤ـ١١ و ٤ـ١٢ بر این دلالت می‌کند که میانگین مقادیر تخمین زده شدهٔ Y_t با میانگین مقادیر Y_t برابر است؛ یعنی مجموع پسماندها برابر صفر است.

نتیجه حاصل از معادلهٔ ٤ـ١٢ را می‌توان به روش دیگری نیز به دست آورد. از معادلهٔ اول در سیستم معادله‌های نرمال ٤ـ٤ داریم

$$\sum (Y_t - \hat{Y}_t) = ٠ \quad ,$$

یا

$$\sum_{t=١}^{n} e_t = ٠ \quad \text{و یا} \quad \bar{e} = ٠ \quad .$$

د) معادله دوم نرمال در سیستم معادله‌های ٤ـ٤ را میتوان چنین نوشت،

$$\frac{\partial \Sigma e_t^٢}{\partial \hat{\beta}_١} = -٢ \sum (Y_t - \hat{Y}_t) X_{t١} = ٠ \quad ,$$

و در نتیجه

$$\sum X_{t١} e_t = ٠ \quad . \qquad (٤\text{ـ}١٣)$$

به همین ترتیب از معادلهٔ سوم نرمال در سیستم معادله‌های ٤ـ٤ نتیجه می‌گیریم که

$$\sum X_{t٢} e_t = ٠ \quad . \qquad (٤\text{ـ}١٤)$$

اگر دو طرف معادلهٔ ۲ـ٤ را در e_t ضرب کنیم، خواهیم داشت

$$e_t \hat{Y}_t = e_t \hat{\alpha} + e_t \hat{\beta}_1 X_{1t} + e_t \hat{\beta} X_{2t} ,$$

و برای مجموع مشاهدات موجود در نمونه، داریم

$$\sum e_t \hat{Y}_t = \hat{\alpha} \sum e_t + \hat{\beta}_1 \sum X_{1t} e_t + \hat{\beta}_2 \sum X_{2t} e_t .$$

با جایگزینی معادله‌های ۱۲ـ٤، ۱۳ـ٤ و ۱٤ـ٤ در معادلهٔ فوق نتیجه می‌گیریم که

$$\sum e_t \hat{Y}_t = 0 . \qquad (٤\text{ـ}۱۵)$$

از معادله‌های ۱۳ـ٤ ، ۱٤ـ٤ و ۱۵ـ٤ می‌توان این سه نتیجهٔ بسیار مهم را گرفت که در روش حداقل مربعات معمولی، پسماندها، از متغیرهای توضیحی X_{1t} و X_{2t} و تخمین متغیر درون‌زا $\hat{Y}_t$ ، مستقل است.

٤ـ۳ ضریب تعیین

برای به دست آوردن فرمول ضریب تعیین در مدل رگرسیون با دو متغیر توضیحی به ترتیب زیر عمل می‌کنیم. می‌دانیم

$$Y_t = \hat{Y}_t + e_t .$$

دو طرف معادلهٔ فوق را برای تمام مقادیر t جمع کرده، بر n تقسیم می‌کنیم،

$$\bar{Y} = \bar{\hat{Y}} + \bar{e}_t .$$

دو طرف این دو معادله را از هم کم کرده خواهیم داشت

$$y_t = \hat{y}_t + e_t .$$

دو طرف معادله فوق را به توان ۲ می‌رسانیم و برای تمام مقادیر n جمع می‌کنیم،

$$\sum y_t^2 = \sum \hat{y}_t^2 + \sum e_t^2 + 2 \sum \hat{y}_t e_t .$$

جمله آخر صفر است؛ زیرا با توجه به معادلهٔ ٤ـ۱۵ و نیز معادلهٔ ٤ـ۱۲ داریم

$$\sum e_t \hat{Y}_t = \sum e_t (\hat{y}_t + \hat{\bar{Y}}) = \sum e_t \hat{y}_t + \hat{\bar{Y}} \sum e_t = 0 \ .$$

بنابراین می‌توان نوشت

$$\sum y_t^2 = \sum \hat{y}_t^2 + \sum e_t^2 \ ,$$

یا:

$$TSS = ESS + RSS \ .$$

ملاحظه می‌شود که معادلهٔ فوق دقیقاً همان قضیهٔ بسیار مهمی است که به صورت معادلهٔ ٤۰ـ۱ در رگرسیون ساده مطرح کردیم.

با استفاده از تعریف ضریب تعیین (معادلهٔ ۲۹ـ۱) می‌توان نوشت

$$R^2 = \frac{ESS}{TSS} = \frac{\sum \hat{y}_t^2}{\sum y_t^2} = 1 - \frac{\sum e_t^2}{\sum y_t^2} \ . \qquad (٤ـ۱٦)$$

ساختار فرمول ضریب تعیین در رگرسیون چند متغیره دقیقاً همان فرمولی است که در رگرسیون ساده و به صورت معادله‌های ۳۳ـ۱ و ٤۳ـ۱ به دست آمد. البته ضریب تعیین را در رگرسیون چند متغیره با R^2 نشان می‌دهیم که از مقدار مشابه در رگرسیون ساده متمایز شود.

از نظر محاسباتی، با توجه به معادله‌های ٤۵ـ۲ و ٤٦ـ۲ می‌دانیم در یک رگرسیون ساده

$$ESS = \sum \hat{y}_t^2 = \hat{\beta} \sum x_t y_t \ ,$$

$$RSS = \sum e_t^2 = \sum y_t^2 - \hat{\beta} \sum x_t y_t \ .$$

اگر دقیقاً به همان روش به کار رفته در استخراج معادله‌های ٤۵ـ۲ و ٤٦ـ۲ عمل می‌کنیم، می‌توان تغییرات توضیح داده شده (ESS) و مجموع مربعات پسماند (RSS) را برای رگرسیون ٤ـ۱ به شرح زیر به دست آورد،

$$ESS = \sum \hat{y}_t^2 = \hat{\beta}_1 \sum x_{1t} y_t + \hat{\beta}_2 \sum x_{2t} y_t \ , \qquad (٤ـ۱۷)$$

$$RSS = \sum e_t^2 = \sum y_t^2 - \hat{\beta}_1 \sum x_{1t} y_t - \hat{\beta}_2 \sum x_{2t} y_t \ . \qquad (٤ـ۱۸)$$

در نتیجه ، R^2 با استفاده از تعریف ۲۹ـ۱ به صورت زیر محاسبه می‌شود،

$$R^2 = \frac{\hat{\beta}_1 \sum x_{1t} y_t + \hat{\beta}_2 \sum x_{2t} y_t}{\sum y_t^2} \quad (۴ـ۱۹)$$

معمولاً از علامت‌گذاریهای مختلفی برای R^2 استفاده می‌کنند. در بعضی موارد، R^2 را با ۳ اندیس معرف می‌کنند. اندیس اول منعکس‌کنندهٔ متغیر درون‌زاست. بعد از این اندیس، یک ممیز گذاشته می‌شود و ۲ اندیس بعدی برای متغیرهای توضیحی به کار می‌رود. برای مثال، $R^2_{y/x_1, x_2}$ نشان می‌دهد که ضریب تعیین برای مدلی محاسبه می‌شود که Y متغیر درون‌زا و X_1 و X_2 متغیرهای برون‌زا هستند. برای سهولت، $R^2_{y/x_1, x_2}$ را به صورت $R^2_{y/12}$ یا حتی $R^2_{1/23}$ نیز نشان می‌دهند. منظور از «۱۲» متغیرهای اول و دوم توضیحی است. می‌توان $R^2_{y/x_1, x_2}$ را به صورت $R^2_{1/23}$ نیز نوشت. در اینجا «۱/۲۳» بدین صورت تفسیر می‌شود که «۱» منعکس‌کنندهٔ متغیر درون‌زا است و «۲۳» به معنای دومین و سومین متغیر موجود در مدل است که در واقع همان X_1 و X_2 هستند؛ بنابراین $R^2_{1/23}$ یعنی ضریب تعیین یک مدل رگرسیون که اولین متغیر آن (Y_t) تابعی از دومین و سومین متغیر آن (X_2 , X_1) فرض شده است

در پایان این بحث، به این نکته اشاره می‌شود که R^2 برابر مجذور ضریب همبستگی بین Y_t و $\hat{Y}_t$ است،

$$R^2_{y/12} = r^2_{y, \hat{y}} = \frac{(\sum y_t \hat{y}_t)^2}{\sum y_t^2 \sum \hat{y}_t^2} \quad (۴ـ۲۰)$$

در مسألهٔ ۱۹ـ۱ بند (۳)، این فرمول را برای رگرسیون ساده ثابت کردیم. برای رگرسیون چند متغیره نیز به ترتیبی مشابه عمل می‌کنیم. با استفاده از رابطهٔ $y_t = \hat{y}_t + e_t$ داریم

$$\sum y_t \hat{y}_t = \sum (\hat{y}_t + e_t) \hat{y}_t ,$$

$$= \sum \hat{y}_t^2 + \sum e_t \hat{y}_t .$$

با توجه به معادلهٔ ۱۵ـ۴ به سهولت مشاهده می‌شود که جملهٔ آخر ($\sum e_t \hat{y}_t$) برابر صفر

می‌شود، و در نتیجه

$$\sum y_t \hat{y}_t = \sum \hat{y}_t^2 \quad . \qquad (۴\text{-}۲۱)$$

معادلهٔ ۴ـ۲۱ را در معادلهٔ ۴ـ۲۰ جایگزین می‌کنیم،

$$r^2_{y,\hat{y}} = \frac{(\sum \hat{y}_t^2)^2}{\sum y_t^2 \sum \hat{y}_t^2} = \frac{\sum \hat{y}_t^2}{\sum y_t^2} \quad .$$

بر اساس معادلهٔ ۲۹ـ۱ می‌دانیم

$$R^2 = \frac{\sum \hat{y}_t^2}{\sum y_t^2} \quad ,$$

در نتیجه $r^2_{y,\hat{y}}$ با R^2 برابر خواهد شد. لازم است توضیح دهیم که هر جا R^2 را بدون اندیس بنویسیم، منظور همان $R^2_{y/x_1 . x_2}$ یا صورتهای مشابه آن است.

مثال ۴ـ۱ یک نمونهٔ ۵ تایی از کارمندان یک شرکت تولیدی را به طور تصادفی انتخاب کرده‌ایم و داده‌های حقوق ماهانه (Y_i)، تعداد سالهای تحصیلی بعد از اخذ دیپلم (X_{1i}) و سالهای خدمت (X_{2i})، را در جدول ۴ـ۱ نشان داده‌ایم.

جدول ۴ـ۱ داده‌های حقوق، سالهای تحصیلی و خدمت

Y_i	X_{1i}	X_{2i}	$Y_i - \bar{Y}$	$X_{1i} - \bar{X}_1$	$X_{2i} - \bar{X}_2$
۳۰	۴	۱۰	۰	-۱	۰
۲۰	۳	۸	-۱۰	-۲	-۲
۳۶	۶	۱۱	۶	۱	۱
۲۴	۴	۹	-۶	-۱	-۱
۴۰	۸	۱۲	۱۰	۳	۲

۱. این مدل رگرسیون را با استفاده از آمار جدول ۴ـ۱ تخمین بزنید،

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i \quad .$$

۲. R^2 را به دست آورید.
۳. دربارهٔ تخمین پارامترهایی که به دست آورده‌اید، بحث کنید.

۱. برای پاسخ به قسمت یک، ابتدا میانگینها را حساب می‌کنیم.

$$\bar{Y} = ۳۰ \quad , \quad \bar{X}_۱ = ۵ \quad , \quad \bar{X}_۲ = ۱۰ \quad .$$

با توجه به اینکه مشاهدات، بسیار کم است نوشتن تمام محاسبات در جدول ضروری نیست. ابتدا مقادیر هر یک از متغیرها را بر حسب انحراف از میانگین، در ستونهای چهارم، پنجم و ششم جدول می‌نویسیم و سپس مقادیر لازم را محاسبه می‌کنیم.

$$\sum x_{۱i}^۲ = ۱۶ \quad . \qquad \sum x_{۱i} y_i = ۶۲ \quad ,$$

$$\sum x_{۲i}^۲ = ۱۰ \quad , \qquad \sum x_{۲i} y_i = ۵۲ \quad ,$$

$$\sum y_i^۲ = ۲۷۲ \quad , \qquad \sum x_{۱i} x_{۲i} = ۱۲ \quad .$$

با استفاده از معادلات ۱۰ـ۴ مقادیر $\hat{\beta}_۱$ و $\hat{\beta}_۲$ به ترتیب زیر محاسبه می‌شود،

$$\hat{\beta}_۱ = \frac{\sum x_{۲i}^۲ \sum x_{۱i} y_i - \sum x_{۱i} x_{۲i} \sum x_{۲i} y_i}{\sum x_{۱i}^۲ \sum x_{۲i}^۲ - (\sum x_{۱i} x_{۲i})^۲} \quad ,$$

$$\hat{\beta}_۱ = \frac{۱۰ (۶۲) - ۱۲ (۵۲)}{۱۶ (۱۰) - (۱۲)^۲} = \frac{-۴}{۱۶} = -۰/۲۵ \quad .$$

و برای $\hat{\beta}_۲$ داریم

$$\hat{\beta}_۲ = \frac{\sum x_{۱i}^۲ \sum x_{۲i} y_i - \sum x_{۱i} x_{۲i} \sum x_{۱i} y_i}{\sum x_{۱i}^۲ \sum x_{۲i}^۲ - (\sum x_{۱i} x_{۲i})^۲} \quad ,$$

$$\hat{\beta}_۲ = \frac{۱۶ (۵۲) - ۱۲ (۶۲)}{۱۶ (۱۰) - (۱۲)^۲} = \frac{۸۸}{۱۶} = ۵/۵ \quad .$$

برای محاسبهٔ $\hat{\alpha}$، معادلهٔ ٤ـ٦ را می‌نویسیم،

$$\bar{Y} = \hat{\alpha} + \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 \ ,$$

در نتیجه داریم:

$$۳۰ = \hat{\alpha} + (-۰/۲۵)(۵) + (۵/۵)(۱۰) \ ,$$

$$\hat{\alpha} = -۲۳/۷۵ \ .$$

بدین ترتیب تخمین مدل رگرسیون مفروض عبارت است از

$$\hat{Y}_i = -۲۳/۷۵ - ۰/۲۵ X_{1i} + ۵/۵ X_{2i}$$

۲. برای محاسبه R^2 ابتدا معادلهٔ ۱۹ـ٤ را می‌نویسیم،

$$R^2_{y/12} = R^2 = \frac{\hat{\beta}_1 \Sigma x_{1i} y_i + \hat{\beta}_2 \Sigma x_{2i} y_i}{\Sigma y_i^2}$$

که با استفاده از کمیتهای محاسبه شده، خواهیم داشت

$$R^2_{y.12} = R^2 = \frac{-۰/۲۵ (۶۲) + ۵/۵ (۵۲)}{۲۷۲} = \frac{۲۷۰/۵}{۲۷۲} = ۰/۹۹٤$$

۳. با مشاهدهٔ تخمین مدل رگرسیون می‌توان دو نکته را استنتاج کرد. اول اینکه ضریب متغیر سابقهٔ تحصیلی منفی است؛ یعنی افزایش سالهای تحصیلی، بر سطح حقوق تأثیر منفی دارد. نکته دوم این است که ضریب متعلق به سالهای خدمت از ضریب مشابه برای سابقهٔ تحصیلی بسیار بزرگتر است. به عبارت دیگر، با استفاده از تخمین پارامترها می‌توان گفت که اگر سابقهٔ تحصیلی را ثابت بگیریم، به ازای هر سال سابقهٔ خدمت بیشتر، مبلغ a ۵/۵ به سطح حقوق اضافه خواهد شد که در آن a واحدی است که بر اساس آن حقوق ماهانه را اندازه گیری کرده‌ایم. از طرف دیگر، اگر دو نفر را با سالهای خدمت یکسان در نظر بگیریم، پیش‌بینی می‌شود فردی که یک سال بیشتر درس خوانده باشد، a ۰/۲۵ کمتر از دیگری حقوق خواهد گرفت. به این

ترتیب به نظر می‌رسد که نتایج تخمین پارامترهای متغیرهای برون‌زا اساساً رضایت‌بخش نیست.

تخمین جملهٔ ثابت مدل ($\hat{\alpha}$ = -۲۳/۷۵) را نیز باید ارزیابی کرد. $\hat{\alpha}$ در مدل اصلی، نشان‌دهندهٔ حقوق ماهانه فردی دیپلمه است که هیچ سابقهٔ خدمت نیز ندارد. بدیهی است حقوق نمی‌تواند منفی شود، اما در تخمین مدل، مقدار $\hat{\alpha}$ منفی شده است؛ یعنی بر اساس پیش‌بینی مدل، حقوق یک کارمند جدید منفی است؛ بنابراین تخمین جملهٔ ثابت مدل نیز چندان جالب نیست.

نتیجه می‌گیریم که نمونهٔ ٥ تایی ما، نمونهٔ خوبی از جامعه کارمندان شرکت مفروض نیست. به نظر می‌رسد که این نمونه از گروه خاصی از کارمندان گرفته شده باشد؛ زیرا ملاحظه می‌شود که سالهای خدمت اعضای متعلق به این نمونه، بین ۸ تا ۱۲ سال است. در چنین حالتی قاعدتاً نمی‌توان خارج از این محدوده را برای متغیر درون‌زا پیش‌بینی کرد و تخمین مدل رگرسیون مفروض نمی‌تواند به این سؤال پاسخ دهد که حقوق یک کارمند تازه‌وارد چه مقدار خواهد بود.

برای تحلیل بیشتر در نتایج تخمین، دو رگرسیون ساده نیز برای این رگرسیون چندمتغیره می‌سازیم. ابتدا برای حقوق ماهانه (Y_i) و سالهای تحصیلی (X_{1i}) و سپس برای حقوق ماهانه و سالهای خدمت (X_{2i})، مدلهای جداگانه‌ای تخمین می‌زنیم. با استفاده از داده‌های موجود در جدول ٤ـ١ تخمین این مدلها به شرح زیر خواهد بود،

$$\hat{Y}_i = ۱۰/۶۲۵ + ۳/۸۷۵\ X_{1i}\ , \qquad R^۲ = ۰/۸۸۳$$

$$\hat{Y}_i = -۲۲ + ۵/۲\ X_{2i}\ . \qquad R^۲ = ۰/۹۹٤$$

تخمین مدل رگرسیون سادهٔ اول نشان می‌دهد که به ازای هر سال تحصیل بیشتر، مبلغ ۳/۸۷۵ a به حقوق ماهانه اضافه خواهد شد، که در آن a واحد اندازه گیری حقوق است. اما در مدل رگرسیون چندمتغیره دیدیم که اضافه شدن سالهای تحصیل هیچگونه کمکی به افزایش سطح حقوق ندارد. این نتیجه می‌تواند بر این حقیقت دلالت کند که عدم حضور متغیر توضیحی سالهای خدمت در مدل رگرسیون ساده اول، باعث شده است که

ما به نتایج کاملاً غلطی از تأثیر متغیر «سوابق تحصیلی» بر سطح حقوق ماهانه برسیم.

قدم بعدی در تفسیر تخمین پارامترهای مدل مفروض این است که واریانس تخمینها را محاسبه و فرضیه‌های مختلف را نسبت به آنها آزمون کنیم. اما قبلاً باید به بحث آزمون فرضیه در یک مدل رگرسیون با دو متغیر توضیحی بپردازیم.

٤ـ٤ آزمون فرضیه

مدل رگرسیون ٤ـ١ و تخمین آن (٤ـ٢) را یک بار دیگر می‌نویسیم،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \ ,$$

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_{1t} + \hat{\beta} X_{2t} \ .$$

می‌دانیم $\hat{\alpha}$ ، $\hat{\beta}_1$ و $\hat{\beta}_2$ متغیرهای تصادفی هستند. دقیقاً مانند بحث قسمت ٢ـ٢ (خصوصیات آماری تخمین زننده‌های حداقل مربعات معمولی) می‌توان نشان داد که $\hat{\alpha}$ ، $\hat{\beta}_1$ و $\hat{\beta}_2$ توابع خطی از U_t هستند. اگر فرض نرمال بودن تابع توزیع احتمال U_t را بپذیریم، نتیجه می‌گیریم که تخمین پارامترهای مدل ٤ـ١ نیز توزیع نرمال دارد،

$$\hat{\alpha} \sim N\left[E(\hat{\alpha}) \ , \ Var(\hat{\alpha})\right] \ ,$$

$$\hat{\beta}_1 \sim N\left[E(\hat{\beta}_1) \ , \ Var(\hat{\beta}_1)\right] \ ,$$

$$\hat{\beta}_2 \sim N\left[E(\hat{\beta}_2) \ , \ Var(\hat{\beta}_2)\right] \ .$$

با توجه به اینکه اولاً، روش محاسبه میانگین و واریانس تخمین زننده‌های $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ دقیقاً مشابه مدل رگرسیون ساده است که در قسمت ٢ـ٢ ارائه شد و ثانیاً اثبات نحوهٔ استخراج فرمولهای میانگین و واریانس تخمین پارامترها در حالت کلی با k پارامتر در فصل ششم مطرح خواهد شد؛ بنابراین کافی است در این قسمت فقط به ذکر نتایج مدل رگرسیون ٤ـ١ می‌پردازیم.

می‌توان نشان داد که تخمین‌زننده‌های $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ نااریب هستند؛

$$E(\hat{\alpha}) = \alpha \ ,$$

$$E(\hat{\beta}_1) = \beta_1 \ , \qquad (4\text{-}22)$$

$$E(\hat{\beta}_2) = \beta_2 \ .$$

همچنین اگر r_{12} به معنای ضریب همبستگی بین X_{1i} و X_{2i} باشد، از معادلهٔ ۱-۳۲ می‌دانیم که

$$r_{12}^2 = \frac{(\Sigma x_{1i} x_{2i})^2}{\Sigma x_{1i}^2 \Sigma x_{2i}^2} \ ,$$

می‌توان نشان داد که واریانس و کواریانس $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ به شرح زیر خواهد بود،

$$\text{Var}(\hat{\alpha}) = \frac{\sigma^2}{n} + \bar{X}_1^2 \text{Var}(\hat{\beta}_1) + 2\bar{X}_1 \bar{X}_2 \text{Cov}(\hat{\beta}_1, \hat{\beta}_2) + \bar{X}_2^2 \text{Var}(\hat{\beta}_2) \qquad (4\text{-}23)$$

$$\text{Var}(\hat{\beta}_1) = \frac{\sigma^2}{(1 - r_{12}^2) \Sigma x_{1i}^2} \ , \qquad (4\text{-}24)$$

$$\text{Var}(\hat{\beta}_2) = \frac{\sigma^2}{(1 - r_{12}^2) \Sigma x_{2i}^2} \ , \qquad (4\text{-}25)$$

$$\text{Cov}(\hat{\beta}_1, \hat{\beta}_2) = \frac{-\sigma^2 r_{12}^2}{(1 - r_{12}^2) \Sigma x_{1i} \Sigma x_{2i}} \ , \qquad (4\text{-}26)$$

$$\text{Cov}(\hat{\alpha}, \hat{\beta}_1) = -[\bar{X}_1 \text{Var}(\hat{\beta}_1) + \bar{X}_2 \text{Cov}(\hat{\beta}_1, \hat{\beta}_2)] \qquad (4\text{-}27)$$

$$\text{Cov}(\hat{\alpha}, \hat{\beta}_2) = -[\bar{X}_1 \text{Cov}(\hat{\beta}_1, \hat{\beta}_2) + \bar{X}_2 \text{Var}(\hat{\beta}_2)] \qquad (4\text{-}28)$$

روش آزمون فرضیه و ساختن فاصله‌های اطمینان برای تخمین پارامترهای مدل ۱-۴ دقیقاً مانند رگرسیون ساده است که در قسمت ۲-۲ بررسی شد. در اینجا تنها به ذکر چند نکته اکتفا می‌کنیم،

۱. از معادله‌های ۴ـ۲۳ ، ۴ـ۲۴ و ۴ـ۲۵ مشاهده می‌شود که به ازای افزایش ضریب همبستگی بین X_{i1} و X_{i2} ، مقادیر واریانس زیاد می‌شود و از دقت تخمینها کاسته خواهد شد. اگر همبستگی بین X_{i1} و X_{i2} خیلی زیاد شود؛ یعنی r^2 به سمت یک میل کند، عملاً واریانس تخمینها به قدری زیاد می‌شود که دیگر تخمینها قابل اطمینان نخواهند بود. هنگامی که همبستگی بین X_{i1} و X_{i2} کامل شود؛ یعنی $r^2 = 1$ باشد، آنگاه تخمین پارامترها ممکن نخواهد بود. به همین ترتیب اگر بین X_{i1} و X_{i2} مطلقاً رابطه‌ای نباشد؛ یعنی $r^2 = 0$ ، در آن صورت مقادیر واریانسها دقیقاً برابر مقادیری است که از رگرسیون سادهٔ Y_i بر X_{i1} و نیز Y_i بر X_{i2} به دست می‌آید. اینها مسائلی است که به طور خلاصه در مقدمه قسمت ۴ـ۲ و در چهارچوب یک مثال مطرح شد و دوباره در فصل ششم مطرح خواهد شد.

۲. اگر مجموع مربعات پسماند برای مدل ۴ـ۱ را با RSS نشان دهیم، در آن صورت مطابق آنچه در معادلهٔ ۲ـ۳۸ گفتیم، کمیت $\frac{RSS}{\sigma^2}$ ، دارای توزیع χ^2 با $(n-3)$ درجهٔ آزادی است. توجه داریم که در مدل ۴ـ۱ سه پارامتر، تخمین زده می‌شود و در محاسبهٔ RSS سه درجهٔ آزادی از دست می‌دهیم؛ بنابراین

$$\frac{RSS}{\sigma^2} \sim \chi^2 (n-3) \ .$$

دقیقاً مشابه نامساوی ۲ـ۴۴ ، می‌توان برای σ^2 نیز فاصلهٔ اطمینان ساخت؛ برای مثال، یک فاصلهٔ اطمینان ۹۵ درصد برای σ^2 عبارت خواهد بود از

$$\frac{(n-3)\,\hat{\sigma}^2}{\chi^2_{.975}} < \sigma^2 < \frac{(n-3)\,\hat{\sigma}^2}{\chi^2_{.025}} \ . \qquad (۴ـ۲۹)$$

۳. برای به دست آوردن تخمین واریانس U_i ، مقدار مجموع مربعات پسماند را بر درجات آزادی آن تقسیم می‌کنیم. مشابه استدلالی که برای معادلهٔ ۲ـ۲۸ ارائه شده می‌توان ثابت کرد که بدین ترتیب تخمین نااریبی از واریانس U_i خواهیم داشت؛

$$\hat{\sigma}_u^2 = \frac{RSS}{(n-3)} \ . \qquad (۴ـ۳۰)$$

۴. با جایگزینی $\hat{\sigma}^2$ از معادلهٔ ۴ـ۳۰ در معادله‌های ۴ـ۲۳ و ۴ـ۲۶ به تخمین واریانس پارامترها می‌رسیم. اگر جذر واریانسها را حساب کنیم، انحراف معیار تخمین پارامترها را به دست خواهیم آورد. با داشتن میانگین و انحراف معیار پارامترها، به راحتی می‌توان آنها را استاندارد کرد. پارامترهای استاندارد زیر هر یک توزیع t با (n-۳) درجه آزادی دارند:

$$t = \frac{\hat{\alpha} - \alpha}{SE\ (\hat{\alpha})}\ , \qquad (۴\text{ـ}۳۱)$$

$$t = \frac{\hat{\beta}_1 - \beta_1}{SE\ (\hat{\beta}_1)}\ , \qquad (۴\text{ـ}۳۲)$$

$$t = \frac{\hat{\beta}_2 - \beta_2}{SE\ (\hat{\beta}_2)}\ , \qquad (۴\text{ـ}۳۳)$$

مثال ۴ـ۲ تابع تولید زیر را در نظر می‌گیریم:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i\ ,$$

که در آن $U_i \sim IN\ (۰\ ,\ \sigma^2)$ و

$Y_i =$ لگاریتم تولید ,

$X_{1i} =$ لگاریتم کار ,

$X_{2i} =$ لگاریتم سرمایه ,

X_{1i} و X_{2i} غیرتصادفی است. این محاسبات بر اساس یک نمونه شامل ۲۳ شرکت مختلف تولیدی به دست آمده است.

$$\bar{X}_1 = ۱۰\ ,\quad \sum x_{1i}^2 = ۱۲\ ,\quad \sum x_{1i} y_i = ۱۰\ ,$$

$$\bar{X}_2 = ۵\ ,\quad \sum x_{1i}\ x_{2i} = ۸\ ,\quad \sum x_{2i} y_i = ۸\ ,$$

$$\bar{Y} = ۱۲\ ,\quad \sum x_{2i}^2 = ۱۲\ ,\quad \sum y_i^2 = ۱۰\ .$$

الف) $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ و مقادیر انحراف معیار آنها را محاسبه کنید و تخمین مدل را بنویسید.

ب) فاصلهٔ اطمینان ۹۵ درصد برای α، β_1 و β_2 و σ^2 را به دست آورید و فرضیه‌های $\beta_1=1$ و $\beta_2=0$ را به صورت جداگانه در سطح معنی‌دار ۵ درصد آزمون کنید.

الف) از معادله‌های ۱۰ـ۴ استفاده می‌کنیم و مقادیر $\hat{\beta}_1$ و $\hat{\beta}_2$ را به دست می‌آوریم،

$$\hat{\beta}_1 = \frac{\Sigma x_{2i}^2 \Sigma x_{1i} y_i - \Sigma x_{1i} x_{2i} \Sigma x_{2i} y_i}{\Sigma x_{1i}^2 \Sigma x_{2i}^2 - (\Sigma x_{1i} x_{2i})^2} ,$$

$$= \frac{12 (10) - 8 (8)}{12 (12) - (8)^2} = \frac{56}{80} = 0/7 .$$

$$\hat{\beta}_2 = \frac{\Sigma x_{1i}^2 \Sigma x_{2i} y_i - \Sigma x_{1i} x_{2i} \Sigma x_{1i} y_i}{\Sigma x_{1i}^2 \Sigma x_{2i}^2 - (\Sigma x_{1i} x_{2i})^2} ,$$

$$= \frac{12 (8) - 8 (10)}{12 (12) - (8)^2} = \frac{16}{80} = 0/2 .$$

برای محاسبهٔ $\hat{\alpha}$، معادلهٔ ۶ـ۴ را می‌نویسیم،

$$\bar{Y} = \hat{\alpha} + \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 ,$$

در نتیجه داریم:

$$12 = \hat{\alpha} + 0/7 (10) + 0/2 (5) ,$$

و بدین ترتیب

$$\hat{\alpha} = 4 .$$

برای تخمین مقادیر انحراف معیار، ابتدا σ^2 را تخمین می‌زنیم. برای این کار ضروری است که ابتدا R^2 را به دست آوریم و سپس به کمک آن RSS (مجموع مربعات پسماند) را محاسبه کنیم تا $\hat{\sigma}^2$ به دست آید. از فرمول ۱۹ـ٤ استفاده می‌کنیم،

$$R^2_{y/12} = \frac{\hat{\beta}_1 \Sigma x_{1i} y_i + \hat{\beta}_2 \Sigma x_{2i} y_i}{\Sigma y_i^2} ,$$

$$= \frac{۰/۷ (۱۰) + ۰/۲ (۸)}{۱۰} = ۰/۸٦ .$$

برای محاسبهٔ مجموع مربعات پسماند از معادلهٔ ۱۸ـ٤ استفاده می‌کنیم،

$$RSS = \sum e_i^2 = \sum y_i^2 - \hat{\beta} \sum x_{1i} y_i - \hat{\beta}_2 \sum x_{2i} y_i ,$$

$$= ۱۰ - ۰/۷ (۱۰) - ۰/۲ (۸) = ۱/٤ .$$

با استفاده از معادلهٔ ۳۰ـ٤ داریم

$$\hat{\sigma}_u^2 = \frac{RSS}{(n-۳)} = \frac{۱/٤}{۲۰ - ۳} = ۰/۰۷ .$$

با به دست آوردن تخمین واریانس U_i ابتدا باید از معادله‌های ۲٤ـ٤ و ۲۵ـ٤ استفاده کرد تا واریانس $\hat{\beta}_1$ و $\hat{\beta}_2$ به دست آید. اما لازمهٔ این کار محاسبهٔ مجذور ضریب همبستگی بین X_{1i} و X_{2i} یعنی r_{12}^2 است. از معادلهٔ ۳۲ـ۱ می‌دانیم که

$$r_{12}^2 = \frac{(\Sigma x_{1i} x_{2i})^2}{\Sigma x_{1i}^2 \Sigma x_{2i}^2} ,$$

$$= \frac{(۸)^2}{۱۲ (۱۲)} = \frac{٦٤}{۱٤٤} ,$$

بنابراین

$$Var (\hat{\beta}_1) = \frac{\hat{\sigma}^2}{\Sigma x_{1i}^2 (۱ - r_{12}^2)} ,$$

$$= \frac{۰/۰۷}{۱۲\left(۱ - \frac{٦٤}{١٤٤}\right)} = \frac{۰/۰۷}{\frac{۸۰}{۱۲}} = \frac{۷}{۱۰۰}\left(\frac{۳}{۲۰}\right) = \frac{۲۱}{۲۰۰۰} .$$

$$\text{Var}\,(\hat{\beta}_۲) = \frac{\hat{\sigma}^۲}{\Sigma\, x_{۲i}^۲\,(۱ - r_{۱۲}^۲)} ,$$

$$= \frac{۰/۰۷}{۱۲\left(۱ - \frac{٦٤}{۱۰۰}\right)} = \frac{۲۱}{۲۰۰۰} .$$

برای تخمین واریانس $\hat{\alpha}$ به تخمین $\text{Cov}\,(\hat{\beta}_۱ , \hat{\beta}_۲)$ نیاز داریم. از معادلهٔ ٤ـ۲٦ استفاده می‌کنیم،

$$\text{Cov}\,(\hat{\beta}_۱ , \hat{\beta}_۲) = \frac{-\hat{\sigma}^۲ r_{۱۲}^۲}{\Sigma\, x_{۱i}\, x_{۲i}\,(۱ - r_{۱۲}^۲)} ,$$

$$= \frac{-۰/۰۷\left(\frac{٦٤}{١٤٤}\right)}{۸\left(\frac{۸۰}{١٤٤}\right)} = \frac{-۰/۰۷\,(٦٤)}{٦٤۰} = -۰/۰۰۷ ,$$

و با توجه به $\bar{X}_۱ = ۱۰$ و $\bar{X}_۲ = ۵$، واریانس $\hat{\alpha}$ به صورت زیر محاسبه می‌شود،

$$\text{Var}\,(\hat{\alpha}) = \frac{\sigma^۲}{n} + \bar{X}_۱^۲\,\text{Var}\,(\hat{\beta}_۱) + ۲\,\bar{X}_۱\bar{X}_۲\,\text{Cov}\,(\hat{\beta}_۱ , \hat{\beta}_۲) + \bar{X}_۲^۲\,\text{Var}\,(\hat{\beta}_۲) ,$$

$$= \frac{۰/۰۷}{۲۳} + ۱۰۰\,(۰/۷)\left(\frac{۳}{۲۰}\right) + ۲\,(۵)\,(۱۰)\,(۰/۰۷)\left(\frac{۱}{۱۰}\right)$$

$$+ ۲۵\,(۰/۰۷)\left(\frac{۳}{۲۰}\right) ,$$

$$= ۰/۰۷\left[\frac{۱}{۲۳} + \frac{۳۰۰}{۲۰} + ۲۵ + \frac{۷۵}{۲۰}\right] = ۰/۰۷\,(۸/۷۹۳۵) ,$$

$$= ۰/٦۱۵٥٤٥$$

برای محاسبهٔ مقادیر انحراف معیار $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ از واریانس آنها جذر می‌گیریم،

$$SE(\hat{\alpha}) = \sqrt{۰/۶۱۵۵۴۵} = ۰/۷۸ ,$$

$$SE(\hat{\beta}_1) = \sqrt{\frac{۲۱}{۲۰۰۰}} = ۰/۱۰۲ ,$$

$$SE(\hat{\beta}_2) = ۰/۱۰۲ .$$

تخمین مدل رگرسیون مفروض به شرح زیر خواهد بود،

$$\hat{Y}_i = \underset{(۰/۷۸)}{۴} + \underset{(۰/۱۰۲)}{۰/۷ X_{1i}} + \underset{(۰/۱۰۲)}{۰/۲ X_{2i}} ,$$

$$R^2 = ۰/۸۶ .$$

ب) مقدار t از جدول t با ۲۰=۳ - ۲۳ درجهٔ آزادی و در سطح معنی‌دار ۵ درصد برابر است با $t = \pm ۲/۰۸۶$. فاصله‌های اطمینان ۹۵ درصد برای α، β_1 و β_2 به شرح زیر است،

$$\hat{\alpha} - t_{\alpha/2} SE(\hat{\alpha}) < \alpha < \hat{\alpha} + t_{\alpha/2} SE(\hat{\alpha}) ,$$

$$۴ - ۲/۰۸۶ (۰/۷۸) < \alpha < ۴ + ۲/۰۸۶ (۰/۷۸) ,$$

$$۲/۳۷ < \alpha < ۵/۶۳ .$$

برای β_1 داریم:

$$\hat{\beta}_1 - t_{\alpha/2} SE(\hat{\beta}_1) < \beta_1 < \hat{\beta}_1 + t_{\alpha/2} SE(\hat{\beta}_1) ,$$

$$۰/۷ - ۲/۰۸۶ (۰/۱۰۲) < \beta_1 < ۰/۷ + ۲/۰۸۶ (۰/۱۰۲) ,$$

$$۰/۴۹ < \beta < ۰/۹۱ .$$

برای β_2 خواهیم داشت

$$\hat{\beta}_2 - t_{\alpha/2} SE(\hat{\beta}_2) < \beta_2 < \hat{\beta}_2 + t_{\alpha/2} SE(\hat{\beta}_2) ,$$

$$۰/۲ - ۲/۰۸۶\,(۰/۱۰۲) < \beta_۲ < ۰/۲ + ۲/۰۸۶\,(۰/۱۰۲)\ ,$$

$$-۰/۰۱ < \beta_۲ < ۰/۴۱\ .$$

فاصلهٔ اطمینان برای $\sigma^۲$ از توزیع $\chi^۲$ به دست می‌آید. از معادلهٔ ۴ـ۲۹ داریم:

$$\frac{(n-۳)\,\hat{\sigma}^۲}{\chi^۲_{۰/۹۷۵}} < \sigma^۲ < \frac{(n-۳)\,\hat{\sigma}^۲}{\chi^۲_{۰/۰۲۵}}\ .$$

مقادیر به دست آمده از جدول $\chi^۲$ با ۲۰=۳ - ۲۳=(n-۳) درجهٔ آزادی عبارت است از

$$\chi^۲_{۰/۹۷۵} = ۳۴/۲\ ,\qquad ،\qquad \chi^۲_{۰/۰۲۵} = ۹/۵۹\ ,$$

در نتیجه داریم:

$$\Pr\left[\frac{۲۰\,(۰/۰۷)}{۳۴/۲} < \sigma^۲ < \frac{۲۰\,(۰/۰۷)}{۹/۵۹}\right] = ۰/۹۵\ .$$

به عبارت دیگر با ۹۵ درصد احتمال، $\sigma^۲$ بین دو مقدار زیر قرار می‌گیرد،

$$\frac{۲۰\,(۰/۰۷)}{۳۴/۲} < \sigma^۲ < \frac{۲۰\,(۰/۰۷)}{۹/۵۹}\ ,$$

یا

$$۰/۰۴۱ < \sigma^۲ < ۰/۱۴۵\ .$$

فرضیهٔ $\beta_۱=۱$ در سطح معنی‌دار ۵ درصد رد می‌شود؛ زیرا $\beta_۱$ در فاصلهٔ اطمینان ۹۵ درصد قرار نمی‌گیرد. اما فرضیهٔ $\beta_۲=۰$ نقطه‌ای است که در فاصلهٔ اطمینان ۹۵ درصد برای $\beta_۲$ قرار دارد.

۴ـ۵ پیش‌بینی

مسألهٔ پیش‌بینی در مدلهای رگرسیون با دو متغیر توضیحی شبیه پیش‌بینی در مدلهای رگرسیون ساده است؛ با این تفاوت که برای محاسبهٔ خطای پیش‌بینی

باید از واریانس و کوواریانس تمام پارامترها استفاده کرد.

مدل رگرسیون ٤ـ۱ و تخمین آن ٤ـ۲ را یک بار دیگر می‌نویسیم،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \ ,$$

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_{1t} + \hat{\beta}_2 X_{2t} \ .$$

فرض کنید می‌خواهیم به ازای مقادیر معینی از متغیرهای برون‌زا، مقدار متغیر درون‌زا را پیش‌بینی کنیم. دوره‌ای را که می‌خواهیم پیش‌بینی کنیم، t=f می‌نامیم. سؤال این است که به ازای $X_{1f}=X_{1t}$ و $X_{2f}=X_{2t}$ اولاً مقدار $\hat{Y}_f$ چقدر است و ثانیاً چگونه می‌توان یک فاصلهٔ اطمینان ۹۵ درصد برای آن ساخت؟

مدل ٤ـ۱ برای دورهٔ f عبارت است از

$$Y_f = \alpha + \beta_1 X_{1f} + \beta_2 X_{2f} + U_f \ ,$$

که تخمین آن را می‌توان چنین نوشت،

$$\hat{Y}_f = \hat{\alpha} + \hat{\beta} X_{1f} + \hat{\beta}_2 X_{2f} \ . \qquad (٤ـ۳٤)$$

مانند معادلهٔ ٥ـ۳ خطای تخمین را برای معادلهٔ فوق تعریف می‌کنیم،

$$e_f = Y_f - \hat{Y}_f \ , \qquad (٤ـ۳٥)$$

خواهیم داشت

$$e_f = U_f - (\hat{\alpha} - \alpha) - (\hat{\beta}_1 - \beta_1) X_{1f} - (\hat{\beta}_2 - \beta_2) X_{2f} \ . \qquad (٤ـ۳٦)$$

می‌دانیم توابع توزیع احتمال $\hat{\alpha}$، $\hat{\beta}_1$ و $\hat{\beta}_2$ و U_f نرمال هستند و چون e_f یک تابع خطی از آنهاست توزیع احتمال نرمال خواهد داشت. برای پیدا کردن میانگین e_f باید از معادلهٔ ٤ـ۳٦ امید ریاضی بگیریم،

$$E(e_f) = E(U_f) - E(\hat{\alpha} - \alpha) - X_{1f} E(\hat{\beta}_1 - \beta_1) - X_{2f} E(\hat{\beta}_2 - \beta_2) \ .$$

با توجه به اینکه $\hat{\alpha}$ ، $\hat{\beta}_1$ و $\hat{\beta}_2$ تخمینهای ناااریب هستند و نیز با توجه به $E(U_i)=0$ خواهیم داشت

$$E(e_i) = 0 \quad . \qquad (4\text{-}37)$$

با گرفتن امید ریاضی از معادلهٔ ۴-۳۵ و با توجه به معادلهٔ ۴-۳۷ خواهیم داشت

$$E(\hat{Y}_i) = E(Y_i) \quad . \qquad (4\text{-}38)$$

بر اساس معادلهٔ ۴-۳۸ می‌توان گفت که $\hat{Y}_i$ یک پیش‌بینی کنندهٔ ناااریب است. باید دقت کنیم که در اینجا ناااریبی را به صورت $E(\hat{Y}_i)$ برابر با امید ریاضی Y_i گرفته‌ایم، Y_i یک متغیر تصادفی است؛ در حالی که امید ریاضی Y_i دیگر تصادفی نیست، بلکه مقدار معنی دارد.

برای محاسبهٔ واریانس e_i، به ترتیبی کاملاً مشابه روش استخراج معادلهٔ ۳-۸ عمل می‌کنیم. خواهیم داشت

$$\mathrm{Var}(e_i) = \sigma_u^2 \left(1 + \frac{1}{n}\right) + (X_{1i} - \bar{X}_1)^2 \mathrm{Var}(\hat{\beta}_1) + 2(X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)$$
$$\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_2) + (X_{2i} - \bar{X}_2)^2 \mathrm{Var}(\hat{\beta}_2) \quad . \qquad (4\text{-}39)$$

با داشتن میانگین و واریانس e_i، به راحتی می‌توان آن را استاندارد کرد. اگر به جای σ^2 در معادلهٔ ۴-۳۹ مقدار تخمین آن را از معادلهٔ ۴-۳۰ قرار دهیم،

$$\hat{\sigma}_u^2 = \frac{RSS}{(n-3)} \quad ,$$

به تخمین واریانس e_i می‌رسیم. بدین ترتیب مقدار استاندارد شدهٔ e_i دارای توزیع t با $(n-3)$ درجهٔ آزادی خواهد بود،

$$\frac{e_i - E(e_i)}{SE(e_i)} \sim t(n-3) \quad .$$

با توجه به معادله‌های ۴-۳۵ و ۴-۳۷ خواهیم داشت

$$\frac{Y_i - \hat{Y}_i}{SE(e_i)} \sim t(n-3) \quad ,$$

در نتیجه فاصلهٔ اطمینان ۹۵ درصد برای Y_f عبارت است از

$$\hat{Y}_f - t_{\alpha/2}\, SE\,(e_f) < Y_f < \hat{Y}_f + t_{\alpha/2}\, SE\,(e_f) \quad , \qquad (۴\text{-}۴۰)$$

که از نظر ساختار ریاضی دقیقاً مانند نامساوی ۱۱-۳ در مدلهای رگرسیون ساده است. مشاهده می‌شود که تفاوت معادله‌های ۴۰-۴ و ۱۱-۳ فقط در فرمول $SE\,(e_f)$ است.

به ترتیبی کاملاً مشابه با رگرسیون ساده و معادلهٔ ۱۸-۳، می‌توان برای $E\,(Y_f)$ نیز در مدل رگرسیون با دو متغیر توضیحی، فاصلهٔ اطمینان ساخت. با توجه به اینکه حالت کلی این مسأله را در فصل نهم مطرح خواهیم کرد، به نظر می‌رسد، تبیین بیشتر آن در این قسمت ضروری نیست.

مثال ۳-۴ مدل رگرسیون، موضوع مثال ۲-۴ را یک بار دیگر در نظر می‌گیریم:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i \quad .$$

می‌دانیم تخمین این مدل به صورت زیر به دست آمده است:

$$\hat{Y}_i = ۴ + ۰/۷\, X_{1i} + ۰/۲\, X_{2i} \quad .$$

همچنین نتایج محاسباتی زیر را از مثال ۲-۴ داریم،

$$\bar{X}_1 = ۱۰ \quad , \qquad Var\,(\hat{\beta}_1) = \frac{۳}{۲۰}\,(۰/۰۷) \quad ,$$

$$\bar{X}_2 = ۵ \quad , \qquad Var\,(\hat{\beta}_2) = \frac{۳}{۲۰}\,(۰/۰۷) \quad ,$$

$$\hat{\sigma}_u^2 = ۰/۰۷ \quad , \qquad Cov\,(\hat{\beta}_1 , \hat{\beta}_2) = -\frac{۱}{۱۰}\,(۰/۰۷) \quad .$$

اولاً، مقدار Y_f را به ازای $X_{1f}=۱۲$ و $X_{2f}=۷$، پیش‌بینی کنید.

ثانیاً، برای مقدار پیش‌بینی شدهٔ Y_f یک فاصلهٔ اطمینان ۹۵ درصد بسازید.

برای قسمت اول تخمین مدل را برای دورهٔ f می‌نویسیم:

$$\hat{Y}_f = ۴ + ۰/۷\, X_{1f} + ۰/۲\, X_{2f} \quad ,$$

که به ازای $X_{1f}=۱۲$ و $X_{2f}=۷$ خواهیم داشت

$$\hat{Y}_f = ٤ + ۰/۷ (۱۲) + ۰/۲ (۷) = ۱۳/۸ .$$

برای قسمت دوم، نامساوی ٤۰ـ٤ را برای فاصلهٔ اطمینان Y_f می‌نویسیم،

$$\hat{Y}_f - t_{۲/۲} SE (e_f) < Y_f < \hat{Y}_f + t_{۲/۲} SE (e_f) .$$

بنابراین باید ابتدا $Var (e_f)$ را حساب کنیم. با توجه به

$$X_{1f} - \bar{X} = ۱۲ - ۱۰ = ۲ ,$$

$$X_{2f} - \bar{X} = ۷ - ۵ = ۲ ,$$

و با استفاده از معادلهٔ ٤۰ـ۳۹ خواهیم داشت

$$Var(e_f) = ۰/۰۷\left(۱+\frac{۱}{۲۳}\right) + (٤)\left(\frac{۳}{۲۰}\right)۰/۰۷ - ۲(۲)(۲)\left(\frac{۱}{۱۰}\right)۰/۰۷ + ٤\left(\frac{۳}{۲۰}\right)۰/۰۷,$$

$$= ۰/۰۷\left(۱+\frac{۱}{۲۳}\right) + ٤ (۰/۰۷)\left(\frac{۳}{۲۰} - \frac{۲}{۱۰} + \frac{۳}{۲۰}\right) = ۰/۱۰۱ .$$

با گرفتن جذر از $Var (e_f)$، به انحراف معیار e_f خواهیم رسید،

$$SE (e_f) = \sqrt{۰/۱۰۱} = ۰/۳۱۸ .$$

بدین ترتیب نامساوی ٤۰ـ٤ را می‌توان با توجه به $t = \pm ۲/۰۸٦$ چنین نوشت،

$$۱۳/۸ - ۲/۰۸٦ (۰/۳۱۸) < Y_f < ۱۳/۸ + ۲/۰۸٦ (۰/۳۱۸) ,$$

$$۱۳/۸ - ۰/٦٦ < Y_f < ۱۳/۸ + ۰/٦٦ ,$$

یا

$$۱۳/۱٤ < Y_f < ۱٤/٤٦ .$$

دقت پیش‌بینی

فاصلهٔ اطمینان برای Y_f در مدل رگرسیون‌ساده، یعنی رابطه ۳ـ۱۰ رایک‌بار دیگر می‌نویسیم:

$$\hat{Y}_f - t_{\alpha/2}\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\sum x_i^2}} < Y_f < \hat{Y}_f + t_{\alpha/2}\hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(X_f-\bar{X})^2}{\sum x_i^2}}.$$

در بحث مربوط به نمودار ۱ـ۳ دیدیم که به ازای $X_f = \bar{X}$، فاصلهٔ اطمینان برای Y_f در کمترین حد ممکن است؛ زیرا به ازای این مقدار، جملهٔ $\frac{(X_f-\bar{X})^2}{\sum x_i^2}$ حذف خواهد شد؛ به عبارت دیگر، هر چه X_f از $\bar{X}$ فاصله بگیرد، واریانس و انحراف معیار e_f بیشتر می‌شود و از دقت پیش‌بینی فاصلهٔ اطمینان برای Y_f کاسته خواهد شد.

نتیجهٔ فوق در مورد مدلهایی که بیش از یک متغیر توضیحی دارد، معمولاً صادق نیست. علت آن، وجود جملهٔ حاوی کوواریانس $\hat{\beta}_1$ و $\hat{\beta}_2$ در فرمول محاسبهٔ واریانس e_f است. در معادلهٔ ٤ـ۳۹ می‌بینیم که واریانس e_f نه فقط تابعی از واریانس U_f، واریانس $\hat{\beta}_1$ و واریانس $\hat{\beta}_2$ است، بلکه از کوواریانس $\hat{\beta}_1$ و $\hat{\beta}_2$ نیز تبعیت می‌کند. معمولاً علامت کوواریانس، که با توجه به فرمول ٤ـ۲٦ منفی است، و ضریب این کوواریانس، که حاصل ضرب $(X_{1f}-\bar{X}_1)$ و $(X_{2f}-\bar{X}_2)$ است، عوامل مهمی در نادرستی نتیجهٔ حاصل از رگرسیون ساده به رگرسیون چند متغیره است.

برای تبیین بیشتر این نکته، به معادلهٔ ٤ـ۳۹ مراجعه می‌کنیم. اگر X_{1f} از $\bar{X}_1$ فاصله بگیرد، چون اختلاف آن از $\bar{X}_1$ به توان ۲ می‌رسد، در نهایت، نتیجه یک عدد مثبت خواهد بود. بنابراین فرقی نمی‌کند که جهت انحراف X_{1f} از $\bar{X}_1$ چه باشد. برای هر گونه انحراف X_{2f} از $\bar{X}_2$ نیز این بحث دقیقاً معتبر خواهد بود. اما برای ضریب کوواریانس $\hat{\beta}_1$ و $\hat{\beta}_2$ این گونه نیست. فرض می‌کنیم $\text{Cov}(\hat{\beta}_1, \hat{\beta}_2)$ یک عدد منفی است. اگر X_{1f} به $\bar{X}_1$ چنان نزدیک شود که نتیجهٔ $(X_{1f}-\bar{X}_1)$ یک عدد منفی شود، حاصل ضرب $(X_{1f}-\bar{X}_1)$ و $\text{Cov}(\hat{\beta}_1, \hat{\beta}_2)$ یک عدد مثبت شده و واریانس e_f بشدت افزایش خواهد یافت؛ در حالی که با نزدیک شدن X_f به میانگین $\bar{X}_1$ انتظار ما این بود که واریانس e_f کاهش یابد. بدین ترتیب نتیجه می‌گیریم که در مدلهای رگرسیون چند متغیره نمی‌توان

به همان نتیجهٔ ساده و قابل توجهی برسیم که در مبحث دقت پیش‌بینی در رگرسیونهای ساده حاصل شد. در این مورد مثالی می‌آوریم.

مثال ٤ـ٤ در مثال ٤ـ٣ دیدیم که $\bar{X}_2$ برابر ٥ بود و به ازای $X_{2f}=7$ مقدار واریانس e_f را برابر ۰/۱۰۱ تخمین زدیم. با توجه به معادلهٔ ٤ـ۳۹ مشاهده می‌شود که $(X_{2f}-\bar{X}_2)$ در مثال ٤ـ٣ با ۲ = (٥ـ٧) برابر است و هنگامی که به صورت ضریب واریانس $\hat{\beta}_1$ به توان ۲ برسد، برابر ٤ خواهد شد.

حال فرض کنید که می‌خواهیم به ازای $X_{2f}=3$ پیش‌بینی کنیم که مقدار $(X_{2f}-\bar{X}_2)$ در معادلهٔ ٤ـ۳۹ برابر ۲- = (٥ـ٣) می‌شود که مربع آن یا ضریب واریانس $\hat{\beta}_1$ همان مقدار ٤ را خواهد داشت. اما تأثیر مقدار جدید X_{2f} در ضریب کوواریانس، کاملاً با مثال قبلی متفاوت است. مقدار این ضریب در مثال ٤ـ٣ عبارت بود از

$$2\,(X_{1f}-\bar{X}_1)(X_{2f}-\bar{X}_2)=2\,(12-10)(7-5)=2\,(4)\ ,$$

در حالی که در این مثال برابر است با

$$2\,(X_{1f}-\bar{X}_1)(X_{2f}-\bar{X}_2)=2\,(12-10)(3-5)=-2\,(4)\ .$$

با توجه به مقدار کوواریانس $\hat{\beta}_1$ و $\hat{\beta}_2$ خواهیم داشت

در مثال ٤ـ٣: $2\,(X_{1f}-\bar{X}_1)(X_{2f}-\bar{X}_2)\,\mathrm{Cov}\,(\hat{\beta}_1,\hat{\beta}_2)=8\,(-0/007)=-0/056$,

در مثال ٤ـ٤: $2\,(X_{1f}-\bar{X}_1)(X_{2f}-\bar{X}_2)\,\mathrm{Cov}\,(\hat{\beta}_1,\hat{\beta}_2)=-8\,(-0/007)=0/056$,

و بدیهی است که این دو مقدار، کاملاً با یکدیگر تفاوت دارند و تأثیر چشمگیری بر $\mathrm{Var}\,(e_f)$ خواهند داشت؛ به عبارت دقیقتر، برای این مثال داریم

$$\mathrm{Var}\,(e_f)=0/07\left(1+\frac{1}{23}\right)+4\,(0/07)\left(\frac{3}{20}\right)+0/056+4\left(\frac{3}{20}\right)0/07=0/213,$$

که از مقدار قبلی، یعنی ۰/۱۰۱، بزرگتر است؛ زیرا به جای یک جملهٔ منفی شامل یک جمله مثبت است.

٤ـ٦ ضرایب همبستگی و ضرایب تعیین*

مدل رگرسیون ٤ـ١ را یک بار دیگر مشاهده کنید،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \quad .$$

ضریب همبستگی Y_t و X_{1t} را با r_{y1}، ضریب همبستگی Y_t و X_{2t} را با r_{y2} و ضریب همبستگی X_{1t} و X_{2t} را با r_{12} نشان می‌دهیم. به r_{y1}، r_{y2} و r_{12} «ضرایب همبستگی ساده»[1] می‌گوییم. همچنین سه تخمین زیر را تعریف می‌کنیم،

$\hat{\beta}_{y1}$ = تخمین شیب رگرسیون Y_t بر X_{1t} ،

$\hat{\beta}_{y2}$ = تخمین شیب رگرسیون Y_t بر X_{2t} ،

$\hat{\beta}_{12}$ = تخمین شیب رگرسیون X_{1t} بر X_{2t} .

به همین ترتیب می‌توان $\hat{\beta}_{1y}$، $\hat{\beta}_{2y}$ و $\hat{\beta}_{21}$ را تعریف کرد. برای مثال

$$\hat{\beta}_{12} = \frac{\Sigma x_{1t} x_{2t}}{\Sigma x_{2t}^2} , \qquad \hat{\beta}_{y1} = \frac{\Sigma x_{1t} y_t}{\Sigma x_{1t}^2} , \qquad \hat{\beta}_{1y} = \frac{\Sigma x_{1t} y_t}{\Sigma y_t^2} .$$

به شش پارامتر فوق «پارامترهای مدل جزئی» می‌گوییم. مدل جزئی، هر زیر مدلی است که بتوان از مدل رگرسیون مرکب ٤ـ١ به دست آورد.

ابتدا این سؤال را مطرح می‌کنیم که آیا می‌توان رابطهٔ بین تخمین پارامترهای مدل و پارامترهای مدل جزئی را به دست آورد؟ فرمول $\hat{\beta}_1$ را از معادله‌های ٤ـ١٠ می‌نویسیم،

$$\hat{\beta}_1 = \frac{\Sigma x_{2t}^2 \Sigma x_{1t} y_t - \Sigma x_{1t} x_{2t} \Sigma x_{2t} y_t}{\Sigma x_{1t}^2 \Sigma x_{2t}^2 - (\Sigma x_{1t} x_{2t})^2} .$$

صورت و مخرج کسر فوق را بر $\Sigma x_{1t}^2 \Sigma x_{2t}^2$ تقسیم می‌کنیم،

$$\hat{\beta}_1 = \frac{\hat{\beta}_{y1} - \hat{\beta}_{21} \hat{\beta}_{y2}}{1 - \hat{\beta}_{12} \hat{\beta}_{21}} . \qquad (٤ـ٤١)$$

1. Simple Correlation Coefficients

با استفاده از فرمول $\hat{\beta}_{۲}$ و به صورتی کاملاً مشابه خواهیم داشت

$$\hat{\beta}_{۲} = \frac{\hat{\beta}_{y۲} - \hat{\beta}_{۱۲}\ \hat{\beta}_{y۱}}{۱ - \hat{\beta}_{۱۲}\ \hat{\beta}_{۲۱}} \quad . \qquad (٤\text{-}٤۲)$$

مفید است به طور خلاصه تحلیلی از معادله‌های ٤ـ٤١ و ٤ـ٤٢ ارائه دهیم. معادلهٔ ٤ـ٤١ را به صورت زیر می‌نویسیم،

$$\hat{\beta}_{۱} = \hat{\beta}_{y۱} - \frac{\hat{\beta}_{۲۱}\ \hat{\beta}_{y۲}}{۱ - \hat{\beta}_{۱۲}\ \hat{\beta}_{۲۱}} \quad ,$$

یا:

$$\hat{\beta}_{۱} = \hat{\beta}_{y۱} - P \quad ,$$

که P را می‌توان «جملهٔ تعدیل» نامید؛ به عبارت دیگر، رابطهٔ فوق نشان می‌دهد که اگر $\hat{\beta}_{y۱}$ را تعدیل کنیم، $\hat{\beta}_{۱}$ به دست می‌آید. $\hat{\beta}_{۱}$ ضریب $X_{۱۱}$ در مدل رگرسیون دو متغیره ١ـ٤ و $\hat{\beta}_{y۱}$ ضریب $X_{۱۱}$ در مدل جزئی زیر است

$$Y_{i} = \alpha + \beta_{۱} X_{۱۱} + V_{i} \quad .$$

اگر $X_{۱۱}$ و $X_{۲۱}$ در نمونه با یکدیگر همبستگی نداشته باشند، یعنی $\Sigma x_{۱۱} \Sigma x_{۲۱} = ۰$، آنگاه $\hat{\beta}_{۱۲} = \hat{\beta}_{۲۱} = ۰$؛ زیرا

$$\hat{\beta}_{۲۱} = \frac{\Sigma x_{۲۱} x_{۱۱}}{\Sigma x_{۱۱}^{۲}} = ۰ \quad ,$$

و در نتیجه جملهٔ تعدیل در معادله‌های ٤ـ٤١ و ٤ـ٤٢ برابر صفر می‌شود؛ بنابراین

$$\hat{\beta}_{۱} = \hat{\beta}_{y۱} \quad , \qquad \hat{\beta}_{۲} = \hat{\beta}_{y۲} \quad .$$

در چنین حالتی می‌گوییم که متغیرهای توضیحی، «متعامد»[1] هستند و بدین ترتیب تخمین پارامترها در رگرسیون ١ـ٤ دقیقاً برابر تخمین پارامترها در مدلهای رگرسیون

1. Orthogonal

ساده جزئی خواهد بود. این مسأله را در جلد دوم این کتاب با عنوان «مسألهٔ همخطی» مطالعه خواهیم کرد.

سؤال دوم این است که آیا می‌توان $\hat{\beta}_1$ و $\hat{\beta}_2$ را برحسب ضرایب همبستگی ساده نوشت؟ پاسخ مثبت است. ابتدا نشان می‌دهیم که چگونه می‌توان ضریب همبستگی ساده بین X_t و Y_t را برحسب β نوشت.

مدل رگرسیون ساده زیر را مشاهده کنید،

$$Y_t = \alpha + \beta X_t + U_t \ .$$

در معادلهٔ ۳۲ـ۱ دیدیم که ضریب همبستگی ساده بین دو متغیر X_t و Y_t عبارت است از

$$r_{yx} = \frac{\Sigma x_t y_t}{\sqrt{\Sigma x_t^2 \Sigma y_t^2}} \ ,$$

$$= \frac{\Sigma x_t y_t}{\Sigma x_t^2} \cdot \frac{\sqrt{\Sigma x_t^2}}{\sqrt{\Sigma y_t^2}} \ .$$

واریانس X_t در نمونه برابر است با

$$\text{Var}\,(X_t) = S^2 = \frac{\Sigma (X_t - \bar{X})^2}{n} = \frac{\Sigma x_t^2}{n} \ ,$$

و انحراف معیار X_t در نمونه عبارت است از

$$S = \sqrt{\frac{\Sigma x_t^2}{n}} \ . \qquad (٤\text{ـ}٤٣)$$

بنابراین با استفاده از معادلهٔ ٤ـ٤٣ خواهیم داشت

$$r_{yx} = \hat{\beta}_{yx} \frac{S_x}{S_y} \ , \qquad (٤\text{ـ}٤٤)$$

یا

$$\hat{\beta}_{yx} = r_{yx} \frac{S_y}{S_x} \ . \qquad (٤\text{ـ}٤٥)$$

با استفاده از معادلهٔ ٤ـ٤٥ می‌توان معادلهٔ ٤ـ٤١ را به صورت زیر نوشت،

$$\hat{\beta}_1 = \frac{r_{y1}\frac{S_y}{S_1} - r_{21}\frac{S_2}{S_1}r_{y2}\frac{S_y}{S_2}}{1 - r_{12}\frac{S_1}{S_2}r_{21}\frac{S_2}{S_1}} = \frac{r_{y1} - r_{21}r_{y2}}{1 - r_{12}r_{21}} \cdot \frac{S_y}{S_1} ,$$

یا:

$$\hat{\beta}_1 = \frac{r_{y1} - r_{21}r_{y2}}{1 - r_{12}^2} \cdot \frac{S_y}{S_1} . \qquad (٤ـ٤٦)$$

به ترتیبی کاملاً مشابه خواهیم داشت

$$\hat{\beta}_2 = \frac{r_{y2} - r_{12}r_{y1}}{1 - r_{12}^2} \cdot \frac{S_y}{S_2} . \qquad (٤ـ٤٧)$$

به تعریف دو اصطلاح می‌پردازیم: به تخمین پارامترهایی، مانند $\hat{\beta}_{21}$ یا $\hat{\beta}_{y2}$ «ضرایب درجهٔ صفر»[1] می‌گویند. به تخمین پارامترهایی مانند $\hat{\beta}_{y1/2}$ «ضرایب درجه اول»[2] گفته می‌شود. منظور از $\hat{\beta}_{y1/2}$، تخمین ضریب اولین متغیر توضیحی در مدلی است که متغیر درون‌زای آن Y_i است. عدد ۲ بعد از ممیز نشان می‌دهد که این مدل، یک متغیر توضیحی دیگر نیز دارد؛ بنابراین، معادله‌های ٤ـ٤١ و ٤ـ٤٢ ، ضرایب درجهٔ اول را برحسب ضرایب درجهٔ صفر در یک مدل رگرسیون چند متغیره مشخص می‌سازد. سؤالی که مطرح می‌شود، این است که آیا در ضرایب همبستگی نیز چنین مراتبی وجود دارد؟

۱. ضرایب همبستگی درجهٔ صفر و یک

ابتدا می‌گوییم ضرایب همبستگی ساده (در مقدمه قسمت ٤ـ٥) را می‌توان «ضرایب همبستگی درجهٔ صفر»[3] نامید. مدل رگرسیون ٤ـ١ را یک بار دیگر می‌نویسیم،

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i .$$

---

1. Zero-order Coefficients
2. First-order Coefficients
3. Zero-order Correlation Coefficients

می‌دانیم X_{1t} بر متغیر Y_t تأثیر می‌گذارد. از مدل فوق ملاحظه می‌شود که متغیر دیگری نیز به نام X_{2t} وجود دارد که Y_t از آن متأثر است. ضریب همبستگی ساده بین Y_t و X_{1t} عبارت است از

$$r_{y1} = \frac{\Sigma y_t x_{1t}}{\sqrt{\Sigma y_t^2 \Sigma x_{1t}^2}} .$$

برای محاسبهٔ r_{y1}، تغییرات ملاحظه شده در Y_t را با تغییراتی که در X_{1t} وجود داشته است مرتبط کرده‌ایم. اما باید توجه داشت که تمام تغییرات Y_t به علت تغییر در X_{1t} نبوده، بلکه قسمتی از آن بر اثر تغییرات X_{2t} حاصل شده است؛ بنابراین r_{y1} واقعاً ضریب همبستگی تغییرات X_{1t} و تغییرات Y_t حاصل از آن را اندازه گیری نمی‌کند، زیرا Y_t از X_{2t} نیز متأثر است.

اگر مدل رگرسیون Y_t روی X_{2t} را بسازیم، خواهیم داشت

$$Y_t = \alpha + \beta_{y2} X_{2t} + V_t .$$

مقادیر پسماند این مدل عبارت است از

$$e_t = Y_t - \hat{\beta}_{y2} X_{2t} . \qquad (٤\text{-}٤٨)$$

مقادیر پسماند را می‌توان این گونه تفسیر کرد که اگر تأثیر خطی تغییرات X_{2t} بر Y_t را کنار بگذاریم، آنچه باقی می‌ماند، تغییرات پسماند است، یعنی آن قسمتی از تغییرات Y_t که به تغییرات X_{2t} مربوط نیست؛ بنابراین باید ضریب همبستگی این قسمت از تغییرات Y_t را با X_{1t} حساب کنیم؛ به عبارت دیگر، باید ضریب همبستگی X_{1t} و $(Y_t - \hat{\beta}_{y2} X_{2t})$ را به دست آوریم. این ضریب همبستگی را با r^*_{y1} نشان می‌دهیم.

$$r^*_{y1} = \frac{\Sigma (y_t - \hat{\beta}_{y2} x_{2t}) x_{1t}}{\sqrt{\Sigma (y_t - \hat{\beta}_{y2} x_{2t})^2 \Sigma x_{1t}^2}} . \qquad (٤\text{-}٤٩)$$

نکتهٔ مهمی که باید در نظر داشت این است که X_{1t} و X_{2t} نیز در مدل (٤-١) از یکدیگر مستقل نیستند. همین امر باعث می‌شود که نتوانیم ضریب همبستگی X_{1t} و Y_t را دقیقاً حساب کنیم؛ زیرا وقتی تغییرات X_{1t} را با Y_t یا با $(Y_t - \hat{\beta}_{y2} X_{2t})$، مرتبط

می‌کنیم، این «تغییرات خالص» X_{1t} نیست که با Y_t مرتبط است، بلکه قسمتی از تغییرات X_{1t} در واقع ناشی از تغییرات X_{2t} است. اکنون باید به همین ترتیب بتوان تغییرات X_{1t} را که به علت X_{2t} بوده است حذف کرد. کافی است یک مدل رگرسیون بسازیم که X_{1t} را تابعی از X_{2t} فرض کند. داریم

$$X_{1t} = \alpha + \beta_{12} X_{2t} + \varepsilon_t \ .$$

مقادیر پسماند عبارت است از

$$e_t = X_{1t} - \hat{\beta}_{12} X_{2t} \ . \qquad (۴\text{-}۵۰)$$

یعنی اگر تأثیر خطی تغییرات X_{2t} بر X_{1t} را کنار بگذاریم، آنچه باقی می‌ماند، تغییرات پسماند است، یعنی آن قسمتی از تغییرات X_{1t} که به تغییرات X_{2t} مربوط نیست؛ بنابراین باید در معادلهٔ ۴۹ـ۴ به جای X_{1t}، مقدار پسماند ۵۰ـ۴ را قرار داد. خواهیم داشت

$$r_{y1/2} = \frac{\Sigma (y_t - \hat{\beta}_{y2}\ x_{2t})(x_{1t} - \hat{\beta}_{12}\ x_{2t})}{\sqrt{\Sigma (y_t - \hat{\beta}_{y2}\ x_{2t})^2 \Sigma (x_{1t} - \hat{\beta}_{12}\ x_{2t})^2}} \ . \qquad (۴\text{-}۵۱)$$

معادلهٔ ۵۱ـ۴ فرمول «ضریب همبستگی درجه اول»[1] است. $r_{y1/2}$ نشان می‌دهد که می‌خواهیم ضریب همبستگی بین Y_t و اولین متغیر توضیحی (X_{1t}) را در مدلی حساب کنیم که یک متغیر توضیحی دیگر نیز به نام X_{2t} دارد؛ بنابراین باید توجه کرد که عدد ۲ بعد از ممیز بدین معنی نیست که متغیر توضیحی دیگر «ثابت» فرض شده است، بلکه می‌گوییم X_{2t} به منزلهٔ یک متغیر توضیحی وجود دارد، اما سعی می‌کنیم تأثیر تغییرات آن را در Y_t و نیز در X_{1t} به نحوی حذف کنیم که بتوانیم خالص تغییرات X_{1t} و Y_t را با یکدیگر مرتبط ساخته، ضریب همبستگی آنها را محاسبه کنیم. به $r_{y1/2}$، یعنی ضریب همبستگی درجه اول، معمولاً «ضریب همبستگی جزئی»[2] می‌گویند.

باید توجه داشت که اگر بین X_{1t} و X_{2t} در نمونه، همبستگی وجود نداشته باشد،

---

1. First-order Correlation Cofficient
2. Partial Correlation Coefficient

در معادلهٔ ٤ـ٤٩ جملهٔ $\Sigma x_{1t}x_{2t}$ برابر صفر شده و ضریب همبستگی جزئی عبارت خواهد بود از

$$r^{*}_{y1/2} = \frac{\Sigma (y_t\ x_{1t})}{\sqrt{\Sigma (y_t - \hat{\beta}_{y2}\ x_{2t})^2\ \Sigma (x^2_{1t})}} \ .$$

۲. ضریب همبستگی جزئی برحسب ضرایب همبستگی ساده

در معادلهٔ ٤ـ٥١ توانستیم ضریب همبستگی جزئی بین Y_t و X_{1t} را در مدلی به دست آوریم که شامل X_{2t} است. در این قسمت می‌خواهیم این ضریب همبستگی جزئی را برحسب ضرایب همبستگی ساده بنویسیم. ابتدا صورت کسر ٤ـ٥١ را به شرح زیر می‌نویسیم:

$$r_{y1/2} = \Sigma y_t\ x_{1t} - \hat{\beta}_{12} \Sigma y_t\ x_{2t} - \hat{\beta}_{y2} \Sigma x_{1t}\ x_{2t} + \hat{\beta}_{y2} \hat{\beta}_{12} \Sigma x^2_{2t} \ .$$

صورت کسر جملهٔ اول ($\Sigma y_t\ x_{1t}$) را ملاحظه می‌کنیم. از معادلهٔ ۱ـ۳۲ می‌دانیم

$$r_{y1} = \frac{\Sigma y_t\ x_{1t}}{\sqrt{\Sigma y^2_t}\ \sqrt{\Sigma x^2_{1t}}} \ ,$$

در نتیجه

$$\Sigma y_t\ x_{1t} = n\, r_{y1} \sqrt{\frac{\Sigma y^2_t}{n}}\ \sqrt{\frac{\Sigma x^2_{1t}}{n}} \ .$$

همچنین از معادلهٔ ٤ـ٤٣ می‌دانیم

$$S_y = \sqrt{\frac{\Sigma y^2_t}{n}} \ , \qquad S_1 = \sqrt{\frac{\Sigma x^2_{1t}}{n}} \ ,$$

بنابراین جملهٔ اول صورت کسر را می‌توان به صورت زیر نوشت،

$$\Sigma y_t\ x_{1t} = n\, r_{y1}\, S_y\, S_1 \ . \qquad (٤ـ٥٢)$$

حال به جملهٔ دوم توجه می‌کنیم. از معادلهٔ ٤ـ٤٥ می‌دانیم

$$\hat{\beta}_{12} = r_{12} \frac{S_1}{S_2} \ .$$

دقیقاً مشابه معادلهٔ ٤ـ۵۲ خواهیم داشت

$$\sum y_t \; x_{2t} = n \, r_{y2} \, S_y \, S_2 \;\; ,$$

در نتیجه جملهٔ دوم رابه صورت زیر می‌نویسیم،

$$\hat{\beta}_{12} = \sum y_t \; x_{2t} = r_{12} \frac{S_1}{S_2} \; n \, r_{y2} \, S_y \, S_2 \;\; .$$

به همین ترتیب خواهیم داشت

$$\hat{\beta}_{y2} \sum x_{1t} \; x_{2t} = r_{y2} \frac{S_y}{S_2} \; n \, r_{12} \, S_1 \, S_2 \;\; ,$$

$$\hat{\beta}_{y2} \hat{\beta}_{12} \sum x_{2t}^2 = r_{y2} \, r_{12} \frac{S_y \, S_1}{S_2^2} \; n \, S_2^2 \;\; .$$

عبارتهای فوق را در صورت کسر $r_{y1/2}$ جایگزین می‌کنیم،

$$\text{صورت کسر} = n \, r_{y1} \, S_y \, S_1 - r_{12} \frac{S_1}{S_2} \; n \, r_{y2} \, S_y \, S_2 - r_{y2} \frac{S_y}{S_2} \; n \, r_{12} \, S_1 \, S_2$$

$$+ r_{y2} \, r_{12} \frac{S_y \, S_1}{S_2^2} \; n \, S_2^2 \;\; .$$

و بعد از ساده کردن داریم

$$\text{صورت کسر} = n \, S_y \, S_1 \, (r_{y1} - r_{y2} \, r_{12}) \;\; .$$

مخرج کسر ٤ـ۵۱ را نیز می‌توان به راحتی ساده کرد. یک بار دیگر مخرج کسر را می‌نویسیم،

$$\text{مخرج کسر} = \sqrt{\sum (y_t - \hat{\beta}_{y2} \; x_{2t})^2 \sum (x_{1t} - \hat{\beta}_{12} \, x_{2t})^2} \;\; .$$

جملهٔ $\Sigma \, (y_t - \hat{\beta}_{y2} x_{2t})^2$ ، در واقع مجموع مربعات پسماند برای مدل رگرسیون زیر است،

$$Y_t = \alpha + \beta_{y2} \, X_{2t} + V_t \;\; .$$

با استفاده از معادلهٔ ٤ـ۴۳ـ۱ ، برای معادلهٔ فوق داریم

$$r_{y2}^2 = 1 - \frac{RSS}{\sum y_t^2} ,$$

یا

$$RSS = \sum y_t^2 (1 - r_{y2}^2) .$$

با توجه به معادلهٔ ٤٣ـ٤ می‌دانیم

$$\sum y_t^2 = \frac{n \sum y_t^2}{n} = n S_y^2 ,$$

بنابراین

$$RSS = n S_y^2 (1 - r_{y2}^2) .$$

جملهٔ $\sum (y_t - \hat{\beta}_{y2} x_{2t})^2$ (مجموع مربعات پسماند برای مدل جزئی فوق) را می‌توان به‌صورت زیر نوشت،

$$\sum (y_t - \hat{\beta}_{y2}\ x_{2t})^2 = n S_y^2 (1 - r_{y2}^2) .$$

به همین ترتیب می‌توان گفت که جملهٔ $\sum (x_{1t} - \hat{\beta}_{12} x_{2t})^2$ یعنی مجموع مربعات پسماند برای مدل

$$X_{1t} = \alpha + \beta_{12} x_{2t} + \varepsilon_t ,$$

عبارت خواهد بود از

$$\sum (x_{1t} - \hat{\beta}_{12}\ x_{2t})^2 = n S_1^2 (1 - r_{12}^2) .$$

اکنون مخرج کسر ٥١ـ٤ را به صورت زیر می‌نویسیم،

$$\text{مخرج کسر} = n S_y S_1 \sqrt{(1 - r_{y2}^2)(1 - r_{12}^2)} .$$

با توجه به نتایج فوق برای صورت و مخرج، می‌توان کسر ٥١ـ٤ را به شرح زیر نوشت،

$$r_{y1/2} = \frac{r_{y1} - r_{y2} r_{12}}{\sqrt{(1 - r_{y2}^2)(1 - r_{12}^2)}} . \qquad (٤ـ٥٣)$$

معادلهٔ ۵۳ـ۴ مقدار ضریب همبستگی جزئی را بین متغیر درون‌زای Y_t و متغیر برون‌زای X_{1t} مشخص می‌کند. به همین ترتیب می‌توان ضریب همبستگی جزئی را بین متغیر درون‌زای Y_t و متغیر برون‌زای X_{2t} تعریف کرد. به همین ترتیب و با توجه به آنچه که تاکنون داشتیم، می‌توان گفت که

$$r_{y2/1} = \frac{r_{y2} - r_{y1}\, r_{12}}{\sqrt{(1 - r_{y1}^2)\,(1 - r_{12}^2)}} \quad . \qquad (۵۴ـ۴)$$

یادآوری این نکته مفید است که $r_{y2/1}$، یعنی ضریب همبستگی جزئی بین Y_t و X_{2t}، در حقیقت یک ضریب همبستگی ساده بین دو سری پسماند به شرح زیر است.

۱. $(y_t - \hat{\beta}_{y1} x_{1t})$ که مقدار e_t برای مدل زیر است،

$$Y_t = \alpha + \beta_{y1} X_{1t} + V_t \quad .$$

۲. $(x_{2t} - \hat{\beta}_{21} x_{1t})$ که مقدار e_t برای مدل زیر خواهد بود،

$$X_{2t} = \alpha + \beta_{21} x_{1t} + \varepsilon_t \quad .$$

نتیجهٔ کلی بحث را یک بار دیگر ذکر می‌کنیم: ضریب همبستگی جزئی بین هر دو متغیر در حقیقت همبستگی بین پسماندهای هر یک از این متغیرهاست وقتی تأثیرات خطی تمام متغیرهای توضیحی دیگر بر روی آنها را حذف کرده باشیم.

در پایان این قسمت به معرفی یک اصطلاح می‌پردازیم. به مجذور ضریب همبستگی جزئی اصطلاحاً «ضریب تعیین جزئی»[1] می‌گویند.

۳. ضرایب تعیین و پسماندها

در این قسمت می‌خواهیم اهمیت پسماندها و ارتباط آن را با ضرایب تعیین بیشتر تبیین کنیم. اولین نکته این است که می‌توان به کمک پسماندهای به دست آمده از مدلهای جزئی رگرسیون، به تخمین پارامترهای یک رگرسیون چند متغیره رسید.

---

1. Partial Coefficient of Determination

مدل رگرسیون ٤ـ١ را یک بار دیگر می‌نویسیم،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \ .$$

مدل رگرسیون جزئی که Y_t را برحسب X_{2t} بیان می‌کند، در نظر می‌گیریم.

$$Y_t = \alpha + \beta_{y2} X_{2t} + V_t \ .$$

پسماند این مدل جزئی عبارت است از

$$(y_t - \hat{\beta}_{y2} x_{2t}) \ . \qquad (٤ـ٥٥)$$

به همین ترتیب یک مدل رگرسیون جزئی دیگر می‌سازیم که در آن X_{1t} برحسب X_{2t} بیان شده باشد.

$$X_{1t} = \alpha + \beta_{12} X_{2t} + \varepsilon_t \ . \qquad (٤ـ٥٦)$$

پسماند این مدل به شرح زیر است،

$$(x_{1t} - \hat{\beta}_{12}\ x_{2t}) \ . \qquad (٤ـ٥٧)$$

یک مدل جزئی می‌سازیم که متغیرهای ٤ـ٥٦ و ٤ـ٥٧ به ترتیب متغیرهای درون‌زا و برون‌زای آن باشد. تخمین شیب این مدل خطی ساده را می‌نویسیم،

$$\text{شیب مدل جزئی} = \frac{\Sigma\,(y_t - \hat{\beta}_{y2}\ x_{2t})\,(x_{1t} - \hat{\beta}_{12}\ x_{2t})}{\Sigma\,(x_{1t} - \hat{\beta}_{12}\ x_{2t})^2} \ ,$$

$$= \frac{\Sigma\, y_t\,(x_{1t} - \hat{\beta}_{12}\ x_{2t}) - \hat{\beta}_{y2}\,\Sigma\, x_{2t}\,(x_{1t} - \hat{\beta}_{12}\ x_{2t})}{\Sigma\,(x_{1t} - \hat{\beta}_{12}\ x_{2t})^2} \ .$$

در مدل رگرسیون ٤ـ٥٦ می‌دانیم X_{2t} متغیر توضیحی و $(x_{1t} - \hat{\beta}_{12}\ x_{2t})$ همان پسماند یا e_t است. از طرف دیگر از معادلهٔ ١ـ٢٤ می‌دانیم که در تخمینهای حداقل مربعات معمولی، متغیر توضیحی از جملهٔ پسماند مستقل است،

$$\Sigma\, x_{2t}\,(x_{1t} - \hat{\beta}_{12}\ x_{2t}) = 0 \ ,$$

بدین ترتیب شیب مدل جزئی برابر است با

$$\text{شیب مدل جزئی} = \frac{\Sigma y_t\ x_{1t} - \hat{\beta}_{12}\Sigma y_t\ x_{2t}}{\Sigma (x_{1t} - \hat{\beta}_{12}\ x_{2t})^2} .$$

با جایگزینی $\hat{\beta}_{12} = \frac{\Sigma x_{1t} x_{2t}}{\Sigma x_{2t}^2}$ در رابطهٔ مذکور، خواهیم داشت

$$\text{شیب مدل جزئی} = \frac{\Sigma x_{2t}^2 \Sigma x_{1t} y_t - \Sigma x_{1t} x_{2t} \Sigma x_{2t} y_t}{\Sigma x_{1t}^2 \Sigma x_{2t}^2 - (x_{1t}\ x_{2t})^2} = \beta_1 .$$

با توجه به معادلهٔ ۱۰ـ٤ معلوم می‌شود که طرف راست معادلهٔ فوق دقیقاً همان $\hat{\beta}_1$ است. بنابراین توانستیم به کمک پسماندهای حاصل از مدلهای جزئی رگرسیون، به تخمین پارامتر β_1 در مدل رگرسیون چند متغیره برسیم.

به روشی کاملاً مشابه، می‌توان به کمک پسماندهای حاصل از مدل رگرسیون جزئی

$$Y_t = \alpha + \beta_{y1} X_{1t} + V_t ,$$

یعنی $(y_t - \hat{\beta}_{y1}\ x_{1t})$ و نیز پسماندهای حاصل از مدل رگرسیون جزئی

$$X_{2t} = \alpha + \beta_{21} X_{1t} + V_t ,$$

یعنی $(x_{2t} - \hat{\beta}_{21}\ x_{1t})$، یک مدل رگرسیون ساخت و نشان داد که تخمین شیب این مدل دقیقاً برابر $\hat{\beta}_2$ است.

٤. ضریب تعیین برحسب ضرایب همبستگی جزئی

فرض کنید ضرایب همبستگی جزئی در یک مدل رگرسیون چند متغیره را می‌دانیم. آیا می‌توان ضریب تعیین این مدل را برحسب ضرایب همبستگی جزئی نوشت؟ معادلهٔ ۱۹ـ٤ را یک بار دیگر می‌نویسیم،

$$R^2 = R^2_{y/12} = \frac{\hat{\beta}_1 \Sigma x_{1t} y_t + \hat{\beta}_2 \Sigma x_{2t} y_t}{\Sigma y_t^2} .$$

قبلاً در معادله‌های ٤٦ـ٤ و ٤٧ـ٤ توانستیم تخمین پارامترهای مدل رگرسیون چند متغیره، یعنی $\hat{\beta}_1$ ، $\hat{\beta}_1$، را بر حسب ضرایب همبستگی ساده بنویسیم. این مقادیر را در معادلهٔ فوق قرار می‌دهیم. در صورت کسر، دو جمله وجود دارد. ابتدا هر یک از این دو جمله را جداگانه محاسبه می‌کنیم،

$$\hat{\beta}_1 \sum x_{1t} y_t = \frac{(r_{y1} - r_{21}\, r_{y2})\, S_y}{(1 - r_{12}^2)\quad S_1} \cdot R \frac{\sum x_{11}\, y_t}{1} ,$$

$$= \frac{(r_{y1} - r_{21}\, r_{y2})}{(1 - r_{12}^2)} \cdot \frac{\sqrt{\sum y_t^2}}{\sqrt{\sum x_{11}^2}} \cdot \frac{\sqrt{\sum x_{11}\, y_t}}{1} \cdot \frac{\sqrt{\sum x_{11}\, y_t}}{\sqrt{\sum y_t^2}} \cdot \frac{\sqrt{\sum y_t^2}}{1} ,$$

$$= \frac{(r_{y1} - r_{21}\, r_{y2})}{(1 - r_{12}^2)} \cdot \sqrt{\hat{\beta}_{y1}} \cdot \hat{\beta}_{1y} \cdot \sum y_t^2 .$$

با توجه به معادلهٔ ٤٥ـ٤ می‌توان چنین نوشت،

$$\hat{\beta}_{y1} = r_{y1} \frac{S_y}{S_1} \quad , \quad \hat{\beta}_{1y} = r_{1y} \frac{S_1}{S_y} .$$

در نتیجه خواهیم داشت

$$\hat{\beta}_1 \sum x_{1t}\, y_t = \frac{(r_{y1} - r_{21}\, r_{y2})}{(1 - r_{12}^2)} \cdot \sqrt{r_{y1} \frac{S_y}{S_1} r_{1y} \frac{S_1}{S_y}} \cdot \sum y_t^2 .$$

با توجه به اینکه $r_{y1} = r_{1y}$، نتیجه می‌شود که

$$\hat{\beta}_1 \sum x_{1t}\, y_t = \frac{(r_{y1}^2 - r_{21}\, r_{y2}\, r_{y1})}{(1 - r_{12}^2)} \cdot \sum y_t^2 . \qquad (٤\text{ـ}٥٨)$$

برای جملهٔ دوم در صورت کسر ١٩ـ٤ داریم

$$\hat{\beta}_2 \sum x_{2t}\, y_t = \frac{(r_{y2}^2 - r_{21}\, r_{y2}\, r_{y1})}{(1 - r_{12}^2)} \cdot \sum y_t^2 . \qquad (٤\text{ـ}٥٩)$$

معادله‌های ٥٨ـ٤ و ٥٩ـ٤ را در معادلهٔ ١٩ـ٤ جایگزین می‌کنیم و بعد از ساده کردن داریم

$$R_{y/12}^2 = \frac{r_{y1}^2 + r_{y2}^2 - 2\, r_{y1}\, r_{y2}\, r_{21}}{(1 - r_{12}^2)} , \qquad (٤\text{ـ}٦٠)$$

و بـه ایـن تـرتیب بـا داشـتن ضـرایب هـمبستگی سـاده مـی‌توان بـه ضـریب تـعیین رسید.

۵. تجزیهٔ تغییرات توضیح داده شده به کمک ضرایب تعیین

در یک مدل رگرسیون می‌توان تغییرات توضیح داده شده (ESS) را به کمک ضرایب تعیین ساده و جزئی تجزیه کرد و مکانیسم شکل‌گیری آنها را بررسی نمود. بدیهی است این مسأله در تحلیل آنالیز واریانس برای رگرسیون چند متغیره که در فصل ششم بررسی خواهد شد، اهمیت فراوان دارد.

فرض کنید یک متغیر درون‌زا (Y_t) و دو متغیر توضیحی (X_{1t} , X_{2t}) داریم. ابتدا مدل رگرسیون Y_t بر روی X_{1t} را می‌سازیم،

$$Y_t = \alpha + \beta_{y1} X_{1t} + V_t \ .$$

با استفاده از تعریف ضریب تعیین، معادلهٔ ۲۹ـ۱ داریم

$$r_{y1}^2 = \frac{ESS}{\sum y_t^2} \ ,$$

یا

$$ESS = r_{y1}^2 \sum y_t^2 \ . \qquad (۴ـ۶۱)$$

با توجه به معادلهٔ ۴۳ـ۱ می‌توان مجموع مربعات پسماند را به صورت زیر نوشت،

$$RSS = (1 - r_y^2 1) \sum y_t^2 \ . \qquad (۴ـ۶۲)$$

بدیهی است که این پسماندها، یعنی $(y_t - \hat{\beta}_{y1} x_{1t})$ باید توضیح داده شوند.

X_{2t} را وارد مدل می‌کنیم. انتظار این است که بتوانیم پسماندهای Y_t را با X_{2t} توضیح دهیم. اما X_{2t} از X_{1t} مستقل نیست؛ بنابراین قسمتی از تغییرات X_{2t} با X_{1t} و با استفاده از مدل زیر توضیح داده می‌شود.

$$X_{2t} = \alpha + \beta_{21} X_{1t} + \varepsilon_t \ ,$$

در نتیجه پسماندهای X_{2t} عبارت است از $(x_{2t} - \hat{\beta}_{21}\, x_{1t})$. باید بتوانیم به کمک پسماندهای X_{2t}، پسماندهای Y_t را توضیح دهیم. یک رگرسیون می‌سازیم که پسماندهای Y_t را بر حسب پسماندهای X_{2t} بیان کند.

ضریب تعیین این رگرسیون در واقع یک ضریب تعیین جزئی است که می‌توان آن را به صورت $r^2_{y2/1}$ نوشت، که دلالت بر مدلی می‌کند که می‌خواهد تأثیر X_{2t} را بر Y_t اندازه‌گیری کند، وقتی X_{1t} تأثیر خود را گذاشته است. با توجه به معادلهٔ ۱ـ۲۹ می‌توان تغییرات توضیح داده شده را در این مدل حساب کرد. می‌دانیم برای به دست آوردن تغییرات توضیح داده شده (ESS) باید ضریب تعیین را در کل تغییرات (TSS) ضرب کنیم. اما کل تغییراتی که در این مدل باید توضیح داده شود چیزی نیست جز همان مجموع مربعات پسماند Y_t، یعنی معادله ٤ـ٦٢، که از مرحلهٔ قبل باقی مانده است. بنابراین برای به دست آوردن تغییرات توضیح داده شده کافی است که $r^2_{y2/1}$ را در معادلهٔ ٤ـ٦١ ضرب کنیم،

$$\text{ESS} = r^2_{y2/1}\left[(1 - r^2_{y1})\sum y_t^2\right] . \quad (٤\text{ـ}٦٣)$$

در مدل رگرسیون پسماندهای Y_t روی پسماندهای X_{2t}

به سهولت می‌توان مجموع مربعات پسماند در مدل فوق را نیز به دست آورد. با توجه به معادلهٔ ۱ـ٤۳ داریم،

$$\text{RSS} = (1 - r^2_{y2/1})\left[(1 - r^2_{y1})\sum y_t^2\right] . \quad (٤\text{ـ}٦٤)$$

در مدل رگرسیون پسماندهای Y_t روی پسماندهای X_{2t}

بدین ترتیب نتیجه می‌گیریم که تغییرات توضیح داده شدهٔ Y_t، در دو مرحله ساخته شده است. مرحلهٔ اول که در واقع معادلهٔ ٤ـ٦١ است، و آنچه که در مرحلهٔ دوم به دست آمده و در معادلهٔ ٤ـ٦۳ منعکس است. این دو را با هم جمع می‌کنیم تا کل تغییرات توضیح داده شدهٔ Y_t در مدل رگرسیون مفروض به دست آید،

$$\text{کل ESS} = \sum y_t^2\left[r^2_{y1} + r^2_{y2/1}\,(1 - r^2_{y1})\right] . \quad (٤\text{ـ}٦٥)$$

به عبارت دیگر کل تغییرات توضیح داده شدهٔ Y_t شامل دو قسمت است: تغییراتی که از رگرسیون سادهٔ Y_t روی X_{1t} به دست می‌آید و تغییراتی که به رگرسیون ساده پسماندهای Y_t روی پسماندهای X_{2t} مربوط است؛ بنابراین با ورود X_{2t} به مدل، تغییرات توضیح داده شده به اندازهٔ $r^2_{y2/1}\,[(1 - r^2_{y1})\,\Sigma\, y_t^2]$ اضافه می‌شود، که در آن $(1 - r^2_{y1})\,\Sigma\, y_t^2$ مجموع مربعات پسماندهایی است که حاصل رگرسیون Y_t روی X_{1t} است.

یک بار دیگر به تعریف $r^2_{y2/1}$ اشاره می‌کنیم. در یک رگرسیون سادهٔ Y_t روی X_{1t}، تغییراتی که توضیح داده‌نشده است، (پسماندها) را در نظر می‌گیریم. $r^2_{y2/1}$ در واقع درصدی از این تغییرات توضیح داده نشده را حساب می‌کند که با ورود X_{2t} به مدل می‌تواند توضیح داده شود.

معادلهٔ ٤ـ٦٥ را دوباره ملاحظه کنید. وقتی متغیر X_{2t} به X_{1t} اضافه شود، ESS کل تغییرات توضیح داده نشده در Y_t است؛ به عبارت دیگر، کل تغییرات توضیح داده شده چیزی نیست جز تغییرات توضیح داده شده در مدل زیر

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \ .$$

اما از معادلهٔ ٤ـ١٦ می‌دانیم که

$$R^2_{y/12} = \frac{ESS}{\Sigma\, y_t^2} \ ,$$

یا

$$ESS = R^2_{y/12}\,\Sigma\, y_t^2 \ . \qquad (٤ـ٦٦)$$

با جایگزینی معادلهٔ ٤ـ٦٦ در ٤ـ٦٥ خواهیم داشت

$$R^2_{y/12} = r^2_{y1} + r^2_{y2/1}\,(1 - r^2_{y1}) \ . \qquad (٤ـ٦٧)$$

معادلهٔ فوق رابطهٔ بین ضریب تعیین کل، ضریب تعیین جزئی و ضریب تعیین ساده در یک مدل رگرسیون، شامل دو متغیر توضیحی را نشان می‌دهد.

مثال ٤ـ٥ برای تخمین تغییرات دستمزدها مدل رگرسیون زیر مفروض است،

$$W_t = \alpha + \beta_1 \frac{1}{U_t} + \beta_2 P_{t-1} + \varepsilon_t \ ,$$

که در آن W_t درصد تغییر دستمزد و U_t نرخ بیکاری به صورت درصدی از نیروی کار و P_t درصد تغییر در شاخص قیمت کالاهای مصرفی است. آمار این متغیرها به صورت فصلی و برای ۲۶ دوره جمع‌آوری و محاسبات زیر انجام شده است،

$$r_{w1} = ۰/۶۵۰۸ \quad ، \quad r_{w2} = ۰/۳۵۶۷ \quad ، \quad r_{12} = ۰/۰۷۲۶ \ .$$

الف) ضرایب تعیین سادهٔ r^2_{w1} و r^2_{w2} را حساب کنید.

ب) ضرایب همبستگی جزئی و ضرایب تعیین جزئی بین W_t و $\frac{1}{U_t}$ و نیز بین W_t و P_{t-1} را به دست آورید.

ج) ضریب تعیین مدل رگرسیون مفروض، یعنی $R^2_{w/12}$، را محاسبه کنید.

د) نتایج به دست آمده را توضیح دهید.

الف) باید ضرایب همبستگی ساده را مجذور کنیم،

$$r^2_{w1} = (r_{w1})^2 = (۰/۶۵۰۸)^2 = ۰/۴۲۳۵ \ ,$$

$$r^2_{w2} = (r_{w2})^2 = (۰/۳۵۶۷)^2 = ۰/۱۲۷۲ \ .$$

ب) برای محاسبهٔ ضریب همبستگی جزئی بین W_t و $\frac{1}{U_t}$ از معادلهٔ ۵۳ـ۴ استفاده می‌کنیم،

$$r_{w1/2} = \frac{r_{w1} - r_{w2}\, r_{12}}{\sqrt{(1 - r^2_{w2})\,(1 - r^2_{12})}} \ ,$$

$$= \frac{۰/۶۵۰۸ - ۰/۳۵۶۷\,(۰/۰۷۲۶)}{\sqrt{(1 - ۰/۱۲۷۲)\,(1 - ۰/۰۷۲۶^2)}} = ۰/۶۷۰۷ \ .$$

برای به دست آوردن ضریب همبستگی جزئی بین W_t و P_{t-1} کافی است از معادلهٔ ۵۴ـ۴ استفاده شود،

$$r_{w2/1} = \frac{r_{w2} - r_{w1}\, r_{12}}{\sqrt{(1 - r^2_{w1})\,(1 - r^2_{12})}} \ ,$$

$$= \frac{۰/۳۵۶۷ - ۰/۶۵۰۸\,(۰/۰۷۲۶)}{\sqrt{(1 - ۰/۴۲۳۵)\,(1 - ۰/۰۷۲۶^2)}} = ۰/۴۰۸۷ \ .$$

برای به دست آوردن ضریب تعیین جزئی $r^2_{w1/2}$، باید $r_{w1/2}$ را مجذور کنیم،

$$r^2_{w1/2} = (r_{w1/2})^2 = (۰/۶۷۰۷)^2 = ۰/۴۴۹۸ ,$$

و به همین ترتیب

$$r^2_{w2/1} = (r_{w2/1})^2 = (۰/۴۰۸۷)^2 = ۰/۱۶۷۰ .$$

ملاحظه می‌شود که در این مثال چون همبستگی بین متغیرهای توضیحی بسیار ضعیف، یعنی $r_{12} = ۰/۰۷۲۶$؛ تفاوت بسیار مختصری بین ضرایب همبستگی مرتبهٔ صفر و یک وجود دارد.

ج) از معادلهٔ ۶۰-۴ می‌توان برای به دست آوردن $R^2_{w/12}$ استفاده کرد؛

$$R^2_{w/12} = \frac{r^2_{w1} + r^2_{w2} - ۲\, r_{w1}\, r_{w2}\, r_{12}}{(۱ - r^2_{12})} .$$

بنابراین خواهیم داشت

$$R^2_{w/12} = \frac{۰/۴۲۳۵ + ۰/۱۲۷۲ - ۲\,(۰/۶۵۰۸)\,(۰/۳۵۶۷)\,(۰/۰۷۲۶)}{۱ - (۰/۰۷۲۶)^2} ,$$

$$= \frac{۰/۵۱۷}{۰/۹۹۴۸} = ۰/۵۱۹۷ .$$

د) با ملاحظهٔ $r^2_{w1} = ۰/۴۲۳۵$ می‌توان گفت که متغیر $\frac{۱}{U_t}$ به تنهایی می‌تواند بیشتر از ۴۲ درصد تغییرات دستمزدها را توضیح دهد و با توجه به $r^2_{w2} = ۰/۱۲۷۲$ به این نتیجه می‌رسیم که حدود ۱۳ درصد از تغییر در دستمزدها می‌تواند به تنهایی به تغییر در قیمتها نسبت داده شود. اگر هر دو متغیر توضیحی، یعنی بیکاری و قیمت، با هم عمل کنند حدود ۵۲ درصد تغییرات دستمزدها توضیح داده خواهد شد و این حقیقتی است که با ملاحظهٔ $R^2_{w/12} = ۰/۵۱۹۷$ به دست می‌آید. وقتی قیمت تنها متغیر توضیحی باشد، نمی‌تواند همه تغییرات دستمزدها را توضیح دهد و قسمتی از این تغییرات به صورت پسماند باقی می‌ماند. با توجه به ضریب تعیین جزئی $r^2_{w1/2} = ۰/۴۴۹۸$، می‌توان گفت

که بیکاری می‌تواند حدود ٤٥ درصد از این تغییرات توضیح داده نشده را توضیح دهد. همچنین وقتی بیکاری به عنوان تنها متغیر توضیحی عمل کند، قطعاً پسماندهایی در تغییرات دستمزد خواهد بود که متغیر قیمت، با توجه به ضریب تعیین جزئی $r^2_{w1/1} = ۰/۱۶۷۰$، می‌تواند حدود ۱۷ درصد تغییرات آن را تبیین کند.

مسائل فصل چهارم

۴-۱ تخمین معادلات رگرسیون زیر مفروض است،

$$C_t = \hat{\alpha}_1 + ۰/۹۲\ Y_t + e_{1t}\ ,$$

$$C_t = \hat{\alpha}_2 + ۰/۸۴\ C_{t-1} + e_{2t}\ ,$$

$$C_{t-1} = \hat{\alpha}_3 + ۰/۷۸\ Y_t + e_{3t}\ ,$$

$$Y_t = \hat{\alpha}_4 + ۰/۵۵\ C_{t-1} + e_{4t}\ ,$$

که در آن $\hat{\alpha}_i$ تخمین جملهٔ ثابت و e_{it} مقادیر پسماند است. با استفاده از اطلاعات مذکور، پارامترهای β_1 و β_2 در مدل رگرسیون زیر را تخمین بزنید،

$$C_t = \alpha + \beta_1 Y_t + \beta_2 C_{t-1} + U_t\ .$$

۴-۲ فرض کنید دو نمونه با مشخصات زیر موجود است.

نمونهٔ اول	نمونهٔ دوم
$n_1 = ۲۰$	$n_2 = ۲۵$
$\bar{X}_1 = ۲۰$	$\bar{X}_2 = ۲۳$
$\bar{Y}_1 = ۲۵$	$\bar{Y}_2 = ۲۸$
$\sum x_{1t}^2 = ۸۰$	$\sum x_{2t}^2 = ۱۰۰$
$\sum x_{1t} y_{1t} = ۱۲۰$	$\sum x_{2t} y_{2t} = ۱۵۰$
$\sum y_{1t}^2 = ۲۰۰$	$\sum y_{2t}^2 = ۲۵۰$

الف) برای هر نمونه یک معادلهٔ رگرسیون خطی تخمین بزنید.

ب) دو نمونه را با یکدیگر تلفیق کرده، یک معادلهٔ رگرسیون خطی تخمین بزنید.

ج) برای اینکه تخمین پارامترها در «رگرسیون تلفیقی»[1] معتبر باشد، چه فرضهایی را باید پذیرفت.

٤ـ۳ مدل رگرسیون زیر مفروض است،

$$Y_t = \alpha + \beta X_t + U_t \ .$$

این مدل را تخمین زده، سپس متغیر توضیحی دیگری، مانند Z_t به آن اضافه می‌کنیم:

$$Y_t = \alpha' + \beta' X_t + \gamma' Z_t + U'$$

در چه شرایطی موارد زیر صحیح است.

الف) $\hat{\beta} = \hat{\beta}'$.

ب) $\sum e_t^2 \geq \sum e_t'^2$.

ج) در سطح معنی‌دار ۵ درصد، $\hat{\beta}$ معنی‌دار است، اما $\hat{\beta}'$ معنی‌دار نیست.

د) در سطح معنی‌دار ۵ درصد، $\hat{\beta}'$ معنی‌دار است، اما β' معنی‌دار نیست.

٤ـ٤ ابتدا مدل شماره (۱) را تخمین می‌زنیم،

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i \ , \qquad (۱)$$

آنگاه مدل زیر را تخمین می‌زنیم

$$X_{1i} = \gamma + \lambda X_{2i} + V_i \ , \qquad (۲)$$

و پسماندهای رگرسیون مدل آخر، یعنی $\hat{V}_i$ را حساب می‌کنیم. سپس مدل شمارهٔ (۳) را تخمین می‌زنیم،

$$Y_i = \alpha' + \beta_1' \hat{V}_i + \beta_2' X_{2i} + \varepsilon_i \qquad (۳)$$

---

1. Pooled Regression

نشان دهید که، $\hat{\beta}'_1 = \hat{\beta}_1$. این نتیجه را چگونه توضیح می‌دهید؟

۴ـ۵ در مدل رگرسیون

$$y_t = \beta x_{1t} + \gamma x_{2t} + U_t \ ,$$

که $t = 1, 2, \ldots, n$، تمام متغیرها بر حسب انحراف از میانگین نوشته شده است. دو روش مختلف زیر را برای تخمین β در نظر می‌گیریم:

الف) مدل فوق را به همین صورت موجود تخمین می‌زنیم و $\hat{\beta}$ و $\hat{\gamma}$ را حساب می‌کنیم.

ب) ابتدا رگرسیون y_t بر روی x_{2t} را تخمین زده، پسماندهای رگرسیون، یعنی e_t^y را به‌دست می‌آوریم. آنگاه رگرسیون x_{1t} روی x_{2t} را تخمین می‌زنیم و پسماندهای این رگرسیون، یعنی e_{1t}^x را محاسبه می‌کنیم. سپس رگرسیون e_t^y بر روی e_{1t}^x را تخمین می‌زنیم و شیب این تخمین را $\hat{\rho}$ می‌نامیم.

اولاً، نشان دهید $\hat{\rho} = \hat{\beta}$.

ثانیاً، نشان دهید پسماندهای رگرسیون در دو روش فوق با یکدیگر برابر است.

۴ـ۶ به جای اینکه پارامترهای β_1 و β_2 را از مدل رگرسیون زیر تخمین بزنیم،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \ ,$$

از این مدل استفاده می‌شود:

$$Y_t = \alpha' + \beta_1 X_{1t}^* + \beta_2 X_{2t} + V_t \ ,$$

که در آن X_{1t}^*، پسماند رگرسیون X_{1t} روی X_{2t}، و V_t جملهٔ اختلال است.

الف) نشان دهید تخمین β_2 از مدل دوم دقیقاً برابر است با تخمین شیب مدلی که از رگرسیون Y_t روی X_{2t} به دست می‌آید.

ب) مقدار اریب این تخمین را حساب کنید.

۴ـ۷ یک پژوهشگر اقتصادسنجی می‌خواهد مدل مصرف زیر را تخمین بزند،

$$C_t = \alpha + \beta Y_t + \gamma S_t + U_t \ ,$$

که در آن C_t مصرف، Y_t درآمد و S_t پس‌انداز در سطح کلان است. می‌دانیم

$$Y_t = C_t + S_t \ .$$

به نظر شما چه مسائلی در تخمین این مدل مصرف وجود دارد؟

۴ـ۸ نتایج تخمین دو مدل مختلف رگرسیون را ارائه می‌کنیم. چگونه می‌توان گفت که در کدام موارد حتماً اشتباهاتی صورت گرفته است.

الف) $R^2_{y/123} = ۰/۸۹$ ، $R^2_{y/1234} = ۰/۸۶$

ب) $r^2_{y1/2} = ۰/۲۳$ ، $r^2_{y1/3} = ۰/۱۳$ ، $R^2_{y/123} = ۰/۷۰$

۴ـ۹ نتایج تخمین دو مدل مختلف رگرسیون را بیان می‌کنیم. به نظر شما در چه مواردی قطعاً اشتباهاتی در محاسبات صورت گرفته است.

الف) $r^2_{y1} = ۰/۲۲۷$ ، $r^2_{y2} = ۰/۱۲۶$ ، $R^2_{y/12} = ۰/۷۰۱$

ب) $(\sum x_t^2)(\sum y_t^2) - (\sum x_t y_t)^2 = -۱۷۳۲/۸۶$

۴ـ۱۰ مدل زیر را در نظر می‌گیریم که در آن U_t، تمام فرضهای کلاسیک را شامل است،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \ .$$

نشان دهید که

الف) $$R^2 = \frac{\hat{\beta}_1^2 \Sigma x_{1t}^2 + \hat{\beta}_2^2 \Sigma x_{2t}^2 + 2\hat{\beta}_1 \hat{\beta}_2 \Sigma x_{1t} x_{2t}}{\Sigma y_t^2} \ .$$

ب) $$R^2 = \frac{\hat{\beta}_1 \Sigma x_{1t} y_t + \hat{\beta}_2 \Sigma x_{2t} y_t}{\Sigma y_t^2} \ .$$

ج) $$R^2 = \frac{r_{y1}^2 + r_{y2}^2 - 2\, r_{y1}\, r_{y2}\, r_{12}}{1 - r_{12}^2}$$

د) $$r_{y1/2}^2 = \frac{(r_{y1} - r_{y2}\, r_{12})^2}{(1 - r_{12}^2)\,(1 - r_{y2}^2)} \quad .$$

۴-۱۱ مدل رگرسیون زیر را در نظر می‌گیریم،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \quad .$$

این مدل را با روش حداقل مربعات معمولی تخمین می‌زنیم. اگر ضریب همبستگی بین $\hat{\beta}_1$ و $\hat{\beta}_2$ را با $r_{\hat{\beta}_1 \hat{\beta}_2}$ تعریف کنیم، نشان دهید که مجذور ضریب همبستگی بین $\hat{\beta}_1$ و $\hat{\beta}_2$ برابر با مجذور ضریب همبستگی بین X_{1t} و X_{2t} است،

$$r_{\hat{\beta}_1 \hat{\beta}_2} = r_{12} \quad .$$

۴-۱۲ در مدل رگرسیون

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + U_t \quad ,$$

داده‌های متغیرهای Y_t، X_{1t} و X_{2t} را تغییر مقیاس داده‌ایم، به طوری که:

$$Y_t^* = \lambda\, Y_t \quad ,$$

$$X_{1t}^* = \mu_1\, X_{1t} \quad ,$$

$$X_{2t}^* = \mu_2\, X_{2t} \quad .$$

الف) به نظر شما چه رابطه‌ای بین تخمین پارامترها در مدل اولیه، یعنی $\hat{\beta}_2$ و $\hat{\beta}_1$ با تخمین) پارامترها در مدل جدید، یعنی $\hat{\beta}_2^*$ و $\hat{\beta}_1^*$ وجود دارد؟

ب) تغییر مقیاس در اندازه گیری متغیرها، چه تغییری در R^2 و SEE (خطای معیار تخمین) ایجاد می‌کند؟

حل مسائل فصل چهارم

۴ـ۱ با توجه به‌معادلهٔ ۴ـ۴۴ می‌دانیم در مدل رگرسیون $Y_t=\alpha+\beta_1X_{1t}+\beta_2X_{2t}+U_t$، داریم

$$\hat{\beta}_1=\frac{\hat{\beta}_{y1}-\hat{\beta}_{21}\hat{\beta}_{y2}}{1-\hat{\beta}_{12}\hat{\beta}_{21}}\ .$$

با توجه به تخمین معادلات مفروض داریم

$$\hat{\beta}_{y1}=۰/۹۲\ ,\qquad \hat{\beta}_{y2}=۰/۸۴\ ,$$

$$\hat{\beta}_{21}=۰/۷۸\ ,\qquad \hat{\beta}_{12}=۰/۵۵\ ,$$

در نتیجه خواهیم داشت

$$\hat{\beta}_1=\frac{۰/۹۲-۰/۷۸\,(۰/۸۴)}{۱-۰/۵۵\,(۰/۷۸)}=\frac{۰/۲۶۴۸}{۰/۵۷۱}=۰/۴۶۴\ .$$

برای $\hat{\beta}_2$ نیز از معادلهٔ ۴ـ۴۲ داریم

$$\hat{\beta}_2=\frac{\hat{\beta}_{y2}-\hat{\beta}_{12}\hat{\beta}_{y1}}{1-\hat{\beta}_{12}\hat{\beta}_{21}}\ ,$$

$$=\frac{۰/۸۴-۰/۵۵\,(۰/۹۲)}{۱-۰/۵۵\,(۰/۷۸)}=\frac{۰/۳۳۴}{۰/۵۷۱}=۰/۵۸۵$$

۴ـ۲ برای نمونه‌های اول و دوم، به ترتیب اندیسهای یک و دو می‌گذاریم. در محاسبات زیر اگر متغیری بدون اندیس باشد، بدین معنی است که برای نمونهٔ تلفیقی محاسبه شده است.

$$n = n_1 + n_2 \ ,$$

$$\bar{X} = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2}{n_1 + n_2} \ , \qquad \bar{Y} = \frac{n_1 \bar{Y}_1 + n_2 \bar{Y}_2}{n_1 + n_2} \ .$$

برای محاسبهٔ Σx_i^2، ابتدا به این نکته توجه می‌کنیم که

$$\sum_{1}^{n_1+n_2} x_i^2 = \sum_{1}^{n_1+n_2} X_i^2 - n \bar{X}^2 \ . \qquad (1)$$

اما می‌دانیم که

$$\sum_{1}^{n_1+n_2} X_i^2 = \sum_{1}^{n_1} X_{1i}^2 + \sum_{1}^{n_2} X_{2i}^2 \ .$$

باید هر یک از دو جملهٔ سمت راست را محاسبه کنیم،

$$\sum_{1}^{n_1} X_{1i}^2 = \sum_{1}^{n_1} x_{1i}^2 + n_1 \bar{X}_1^2 \ , \quad \sum_{1}^{n_2} X_{2i}^2 = \sum_{1}^{n_2} x_{2i}^2 + n_2 \bar{X}_2^2 \ ,$$

بنابراین

$$\Sigma X_i^2 = \left(\sum_{1}^{n_1} x_{1i}^2 + \sum_{1}^{n_2} x_{2i}^2\right) + (n_1 \bar{X}_1^2 + n_2 \bar{X}_2^2) \ .$$

رابطه فوق را در معادله (۱) قرار می‌دهیم،

$$\sum_{1}^{n_1+n_2} x_i^2 = \left(\Sigma x_{1i}^2 + \Sigma x_{2i}^2\right) + (n_1 \bar{X}_1^2 + n_2 \bar{X}_2^2 - n \bar{X}^2) \ . \quad (2)$$

جملهٔ دوم سمت راست عبارت فوق را چنین می‌نویسیم،

$$n_1 \bar{X}_1 + n_2 \bar{X}_2^2 - n \bar{X}^2 = n_1 \bar{X}_1^2 + n_2 \bar{X}_2^2 - (n_1 + n_2) \left[\frac{(n_1 \bar{X}_1 + n_2 \bar{X}_2)^2}{(n_1 + n_2)^2}\right] \ ,$$

$$= n_1 \bar{X}_1^2 + n_2 \bar{X}_2^2 - \frac{(n_1 \bar{X}_1 + n_2 \bar{X}_2)^2}{(n_1 + n_2)} \ ,$$

$$= \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1^2 - \bar{X}_2^2) \ .$$

رابطهٔ فوق را در معادلهٔ (۲) جایگزین می‌کنیم،

$$\sum_{1}^{n_1+n_2} x_t^2 = \sum x_{1t}^2 + \sum x_{2t}^2 + \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1^2 - \bar{X}_2^2) \ .$$

بدین ترتیب می‌توان نوشت،

$$\sum_{1}^{n_1+n_2} x_t y_t = \sum x_{1t}\ y_{1t} + \sum x_{2t}\ y_{2t} + \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2)(\bar{Y}_1 - \bar{Y}_2) \ ,$$

$$\sum_{1}^{n_1+n_2} y_t^2 = \sum y_{1t}^2 + \sum y_{2t}^2 + \frac{n_1 n_2}{n_1 + n_2} (\bar{Y}_1^2 - \bar{Y}_2^2) \ .$$

با اطلاعات فوق به محاسبات زیر می‌پردازیم.

الف) تخمین معادلهٔ رگرسیون برای نمونه اول به صورت زیر محاسبه می‌شود،

$$\hat{Y}_{1t} = \underset{(۲/۳۷)}{-۵} + \underset{(۰/۱۲)}{۱/۵}\ X_{1t} \ ,$$

$$\sum e_{1t}^2 = RSS_1 = ۲۰ \ , \quad r_1^2 = ۰/۹ \ , \quad d.f. = ۱۸ \ .$$

برای مدل دوم داریم

$$\hat{Y}_{2t} = \underset{(۲/۹۲)}{-۶/۵} + \underset{(۰/۱۲)}{۱/۵}\ X_{2t} \ ,$$

$$\sum e_{2t}^2 = RSS_2 = ۲۵ \ , \quad r_2^2 = ۰/۹ \ , \quad d.f. = ۲۳ \ .$$

ب) برای نمونهٔ تلفیقی داریم:

$$n = ۴۵ \ , \quad \bar{X} = \frac{۶۵}{۳} \ , \quad \bar{Y} = \frac{۸۰}{۳} \ ,$$

$$\sum x_t^2 = ۲۸۰ \ , \quad \sum x_t y_t = ۳۷۰ \ , \quad \sum y_t^2 = ۵۵۰ \ ,$$

تخمین مدل رگرسیون تلفیقی، عبارت است از

$$\hat{Y}_t = -۱/۹۶ + ۱/۳۲\ X_t \ , \ RSS = ۶۱/۱ \ , \ r^2 = ۰/۸۹ \ .$$

ج) برای اینکه تخمین پارامترها در مدل رگرسیون تلفیقی معتبر باشد، باید فرض کنیم که مدلهای رگرسیون نمونه‌های اول و دوم در واقع یکی است،

$$\alpha_1 = \alpha_2 \ , \ \beta_1 = \beta_2 \ , \ \sigma_1^2 = \sigma_2^2 \ ,$$

در فصل هشتم، قسمت ۵ـ۸ با عنوان آزمون تغییر ساختاری، خواهیم دید که چگونه این فرضها را می‌توان آزمون کرد. بعلاوه در مورد این مسألهٔ خاص می‌توان نشان داد که فرضیهٔ $\sigma_1^2 = \sigma_2^2$ را نمی‌شود رد کرد، اما فرضیهٔ $\beta_1 = \beta_2$ و $\alpha_1 = \alpha_2$ در سطح معنی‌دار یک درصد، رد می‌شود، چنین نتیجه‌ای بدین معنی است که نباید این دو نمونه را تلفیق کرد. لازم است یادآوری شود که این نتیجه از اهمیت خاصی برخوردار است؛ زیرا با ملاحظه تخمین مدلهای رگرسیون در هر یک از دو نمونه ملاحظه می‌شود که $\hat{\beta}_1 = ۱/۵$ و $\hat{\beta}_2 = ۱/۵$ و تخمینهای $\hat{\alpha}_1$ و $\hat{\alpha}_2$ نیز تفاوت بسیار مختصری با هم دارند. بنابراین اگر بخواهیم تنها با توجه به صورت تخمین معادله‌ها دربارهٔ مجوز تلفیق نظر بدهیم، قطعاً جواب مثبت است؛ زیرا این دو تخمین بسیار نزدیک به هم بوده و بر این دلالت می‌کند که از یک جامعهٔ آماری اخذ شده است. اما در آزمون فرضیه مشاهده می‌شود که با وجود $\hat{\beta}_1 = \hat{\beta}_2 = ۱/۵$، حتی نمی‌توان فرضیهٔ $\beta_1 = \beta_2$ را قبول کرد؛ بنابراین نتایج حاصل از تخمینها به تنهایی نمی‌تواند ما را به استنتاج آماری برساند و باید حتماً فرضیه‌های مختلف را آزمون کرد. ادامهٔ مباحث مربوط به این مسأله را می‌توان در مسألهٔ ۸ـ۸ ملاحظه کرد.

۴ـ۳ الف) هرگاه Z_t از X_t مستقل باشد، اضافه کردن Z_t به مدل رگرسیون اول، در تخمین β تغییری نمی‌دهد؛ بنابراین موقعی $\hat{\beta}' = \hat{\beta}$ که Z_t و X_t همبستگی نداشته باشد.

ب) می‌دانیم اگر تعداد متغیرهای توضیحی اضافه شود، مجموع مربعات پسماند کاهش می‌یابد یا حداقل تغییری نمی‌کند؛ بنابراین رابطهٔ $\Sigma e_t^2 \geq \Sigma e_t'^2$ همواره صادق است.

ج) وقتی متغیر جدیدی که اضافه می‌کنیم (مثلاً Z_t) همبستگی خیلی زیاد با X_t

داشته باشد، ممکن است معنی‌دار بودن تخمین ضریب X_t در مدل جدید، یعنی $\hat{\beta}'$، را تغییر دهد.

د) این مسأله، حالتی است که یک متغیر جدید، می‌تواند ضریبی را که معنی‌دار نبوده است، معنی‌دار کند؛ یعنی می‌تواند ضریبی را که مثلاً در سطح معنی‌دار ۵ درصد رد شده است قابل قبول کند. این حالت در شرایطی ایجاد می‌شود که متغیر جدید، یعنی Z_t، اولاً با متغیر قدیمی یعنی X_t همبستگی ضعیفی داشته باشد و ثانیاً از قدرت توضیحی بسیار قوی برخوردار باشد. برای تبیین این نکته، واریانس $\hat{\beta}$ و $\hat{\beta}'$ را می‌نویسیم.

$$\text{Var}(\hat{\beta}) = \frac{\Sigma e_t^2}{(n-2)\Sigma x_t^2} \quad , \quad \text{Var}(\hat{\beta}') = \frac{\Sigma e_t'^2}{(n-3)\Sigma x_t^2 (1 - r_{x,z}^2)} \; .$$

فرض می‌کنیم $\text{Var}(\hat{\beta})$ بقدری زیاد است که در آزمونهای آماری، فرضیهٔ $\beta = 0$ رد نشده است؛ یعنی $\hat{\beta}$ معنی‌دار نیست. اما چون ورود متغیر Z_t ـ که قدرت توضیحی بسیاری هم دارد ـ باعث می‌شود که مجموع مربعات پسماند در مدل جدید ($\Sigma e_t'^2$) کمتر از Σe_t^2 بشود؛ در نتیجه قطعاً $\text{Var}(\hat{\beta}')$ از $\text{Var}(\hat{\beta})$ کمتر خواهد شد. بنابراین چه بسا در آزمونهای آماری $\hat{\beta}'$ معنی‌دار شود در حالی که قبلاً معنی‌دار نبود. البته این امر به شرطی واقع می‌شود که ضریب همبستگی بین X_t و Z_t بسیار کم باشد، این $r_{x,z}^2$ بسیار کوچک باشد؛ زیرا در غیر این صورت، با اینکه صورت کسر $\text{Var}(\hat{\beta}')$ کم می‌شود، اما مخرج کسر نیز کم شده، هیچ دلیلی در دست نخواهد بود که $\text{Var}(\hat{\beta}')$ کاهش یابد.

۴ـ۴ با توجه به معادلهٔ ۱۰ـ۴ می‌دانیم اگر β_1 را از مدل اول تخمین بزنیم، خواهیم داشت

$$\hat{\beta}_1 = \frac{\Sigma x_{2i}^2 \; \Sigma x_{1i} y_i - \Sigma x_{1i} x_{2i} \Sigma x_{2i} y_i}{\Sigma x_{1i}^2 \; \Sigma x_{2i}^2 - (\Sigma x_{1i} x_{2i})^2} \; .$$

نشان می‌دهیم که نتیجهٔ تخمین β_1' از مدل سوم نیز دقیقاً به همین نتیجه خواهد رسید. برای تخمین β_1' از مدل سوم، می‌توان از مدل زیر استفاده کرد،

$$Y_i = \alpha'' + \beta_1' \hat{V}_i + \varepsilon_i' \; ,$$

زیرا می‌دانیم $\hat{V}_i$ پسماند رگرسیون X_{2i} بر X_{2i} است و با توجه به معادله‌های نرمال می‌توان گفت که $\hat{V}_i$ از X_{2i} مستقل است؛ بنابراین وقتی در معادلهٔ سوم، متغیر X_{2i} از $\hat{V}_i$ مستقل باشد، می‌توان به جای تخمین یک مدل رگرسیون دو متغیره، β_1' را با استفاده از مدل رگرسیون ساده فوق تخمین زد:

$$\hat{\beta}_1' = \frac{\Sigma y_i \hat{v}_i}{\Sigma \hat{v}_i^2}$$

می‌دانیم $\hat{v}_i = x_{1i} - \hat{\lambda}\, x_{2i}$. با جایگزینی در معادلهٔ فوق داریم

$$\hat{\beta}_1' = \frac{\Sigma y_i (x_{1i} - \hat{\lambda}\, x_{2i})}{\Sigma (x_{1i} - \hat{\lambda}\, x_{2i})^2} .$$

مقدار $\hat{\lambda} = \dfrac{\Sigma x_{1i}\, x_{2i}}{\Sigma x_{2i}^2}$ را در معادلهٔ فوق قرار می‌دهیم،

$$\hat{\beta}_1' = \frac{\Sigma x_{2i}^2 \Sigma y_i x_{1i} - \Sigma x_{1i} x_{2i} \Sigma x_{2i} y_i}{\Sigma x_{1i}^2 \Sigma x_{2i}^2 - (\Sigma x_{1i} x_{2i})^2} ,$$

که دقیقاً همان فرمول $\hat{\beta}_1$ است؛ بنابراین $\hat{\beta}_1' = \beta_1$. توضیح این نتیجه در متن کتاب آمده است.

۴.۵ الف) تخمین β از مدل اولیه به صورت زیر است،

$$\hat{\beta} = \frac{\Sigma x_{2t}^2 \Sigma x_{1t}\, y_t - \Sigma x_{1t} x_{2t} \Sigma x_{2t} y_t}{\Sigma x_{1t}^2 \Sigma x_{2t}^2 - (\Sigma x_{1t} x_{2t})^2} \qquad (۱)$$

رگرسیون y_t روی x_{2t} عبارت است از

$$y_t = \alpha\, x_{2t} + \varepsilon_t ,$$

و پسماندهای آن برابر است با

$$y_t^e = y_t - \hat{\alpha}\, x_{2t} .$$

رگرسیون x_{1t} روی x_{2t} عبارت است از

$$x_{1t} = \lambda\, x_{2t} + V_t , \qquad (۲)$$

و پسماندهای آن برابر است با

$$x_{1t}^{e} = x_{1t} - \hat{\lambda}\, x_{2t} \ .$$

رگرسیون y_t^e روی x_{1t}^e عبارت است از

$$y_t^e = \rho\, x_{1t}^e + w_t \ , \qquad (۳)$$

و پسماندهای آن برابر است با

$$y_t^{ee} = y_t^e - \hat{\rho}\, x_{1t}^e \ . \qquad (۴)$$

باید ثابت کنیم که $\hat{\rho} = \hat{\beta}$. برای این منظور، به جای y_t^e و x_{1t}^e در معادلهٔ (۳) مقادیرشان را قرار می‌دهیم،

$$(y_t - \hat{\alpha}\, x_{2t}) = \rho\, (x_{1t} - \hat{\lambda}\, x_{2t}) + w_t \ .$$

$\hat{\rho}$ از این معادله برابر است با

$$\hat{\rho} = \frac{\Sigma\, (y_t - \hat{\alpha}\, x_{2t})\, (x_{1t} - \hat{\lambda}\, x_{2t})}{\Sigma\, (x_{1t} - \hat{\lambda}\, x_{2t})^2} \ . \qquad (۵)$$

می‌دانیم $(x_{1t} - \hat{\lambda}\, x_{2t})$ در واقع پسماند رگرسیون در معادلهٔ (۲) است که بنا بر معادلات نرمال از متغیر توضیحی یعنی x_{2t} مستقل است؛ بنابراین

$$\hat{\alpha}\, \Sigma\, x_{2t}\, (x_{1t} - \hat{\lambda}\, x_{2t}) = 0 \ .$$

با جایگزینی رابطهٔ فوق در معادلهٔ (۵)، خواهیم داشت

$$\hat{\rho} = \frac{\Sigma\, y_t\, (x_{1t} - \hat{\lambda}\, x_{2t})}{\Sigma\, (x_{1t} - \hat{\lambda}\, x_{2t})^2} \ .$$

اما از معادلهٔ (۲) می‌دانیم که

$$\hat{\lambda} = \frac{\Sigma\, x_{1t}\, x_{2t}}{\Sigma\, x_{2t}^2} \ ,$$

در نتیجه خواهیم داشت

$$\hat{\rho} = \frac{\Sigma x_{۲t}^{۲}\, x_{۱t}\, y_t - \Sigma x_{۱t}\, x_{۲t}\, \Sigma x_{۲t}\, y_t}{\Sigma x_{۱t}^{۲}\, \Sigma x_{۲t}^{۲} - (x_{۱t}\, x_{۲t})^{۲}} ,$$

که دقیقاً همان معادلهٔ (۱) است؛ یعنی $\hat{\rho} = \hat{\beta}$.

ب) پسماندهای مدل مفروض عبارت است از $y_t - \hat{\beta}\, x_{۱t} - \hat{\gamma}\, x_{۲t}$. با توجه به معادله (٤) ، باید ثابت کنیم که

$$y_t^e - \hat{\rho}\, x_{۱t}^e = y_t - \hat{\beta}\, x_{۱t} - \hat{\gamma}\, x_{۲t} . \qquad (٦)$$

برای این منظور ابتدا $(y_t^e - \rho\, x_{۱t}^e)$ را بسط می‌دهیم،

$$y_t^e - \hat{\rho}\, x_{۱t}^e = (y_t - \hat{\alpha}\, x_{۲t}) - \hat{\rho}\, (x_{۱t} - \hat{\lambda}\, x_{۲t}) .$$

در قسمت اول ثابت کردیم که $\hat{\beta} = \hat{\rho}$. در نتیجه داریم

$$y_t^e - \hat{\rho}\, x_{۱t}^e = y_t - \hat{\beta}\, x_{۱t} - (\hat{\alpha} - \hat{\beta}\, \hat{\lambda})\, x_{۲t} .$$

بنابراین برای اثبات معادلهٔ (٦) باید ثابت کنیم که

$$(\hat{\alpha} - \hat{\beta}\, \hat{\lambda}) = \hat{\gamma} . \qquad (۷)$$

برای اثبات رابطهٔ (۷) کافی است از معادله‌های ٤۱ـ٤ و ٤۲ـ٤ استفاده کنیم. بر اساس این معادله‌ها خواهیم داشت

$$\hat{\beta} = \frac{\hat{\beta}_{yx_۱} - \hat{\beta}_{x_۲x_۱}\, \hat{\alpha}}{۱ - \hat{\lambda}\, \hat{\beta}_{x_۲x_۱}} \quad ، \quad \hat{\gamma} = \frac{\hat{\alpha} - \hat{\lambda}\, \hat{\beta}_{yx_۱}}{۱ - \hat{\lambda}\, \hat{\beta}_{x_۲x_۱}} . \qquad (۸)$$

بنابراین تعداد $(\hat{\alpha} - \hat{\beta}\, \hat{\lambda})$ برابر است با

$$(\hat{\alpha} - \hat{\beta}\, \hat{\lambda}) = \hat{\alpha} - \frac{(\hat{\lambda}\, \hat{\beta}_{yx_۱} - \hat{\beta}_{x_۲x_۱}\, \hat{\alpha})}{۱ - \hat{\lambda}\, \hat{\beta}_{x_۲x_۱}} ,$$

$$= \frac{\hat{\alpha} - \hat{\alpha}\hat{\lambda}\hat{\beta}_{x_2x_1} - \hat{\lambda}\hat{\beta}_{yx_1} + \hat{\alpha}\hat{\lambda}\hat{\beta}_{x_2x_1}}{1 - \hat{\lambda}\hat{\beta}_{x_2x_1}} ,$$

$$= \frac{\hat{\alpha} - \hat{\lambda}\hat{\beta}_{yx_1}}{1 - \hat{\lambda}\hat{\beta}_{x_2x_1}} ,$$

با توجه به معادله (۸)، خواهیم داشت

$$(\hat{\alpha} - \hat{\beta}\hat{\lambda}) = \hat{\lambda} ,$$

که دقیقاً اثبات رابطهٔ (۷) است.

۴ـ۶ الف) می‌دانیم اگر یک مدل رگرسیون را با روش حداقل مربعات معمولی تخمین بزنیم، مقادیر پسماند از متغیرهای توضیحی مستقل است؛ بنابراین X_{1t}^* ـ که پسماند مدل رگرسیون مفروض است ـ از متغیر توضیحی X_{2t} مستقل خواهد بود. در چنین حالتی، رگرسیون Y_t بر روی X_{1t}^* و X_{2t} دقیقاً به همان تخمینی از β_2 می‌رسد که رگرسیون Y_t روی X_{2t} نتیجه خواهد داد.

ب) برای تعیین مقدار اریب $\hat{\beta}_2$، ابتدا فرمول آن را می‌نویسیم،

$$\hat{\beta}_2 = \frac{\Sigma x_{2t} y_t}{\Sigma x_{2t}^2} .$$

مقدار y_t را از مدل اولیه (مدل صحیح) در فرمول فوق قرار می‌دهیم،

$$\hat{\beta}_2 = \frac{\Sigma x_{2t}(\beta_1 x_{1t} + \beta_2 x_{2t} + U_t)}{\Sigma x_{2t}^2} ,$$

$$= \beta_2 + \beta_1 \frac{\Sigma x_{1t} x_{2t}}{\Sigma x_{2t}^2} + \frac{\Sigma x_{2t} U_t}{\Sigma x_{2t}^2} .$$

از دو طرف رابطهٔ بالا امید ریاضی می‌گیریم. با توجه به $E(\Sigma x_{2t} U_t) = 0$، خواهیم داشت

$$E(\hat{\beta}_2) = \beta_2 + \beta_3 \frac{\Sigma x_{3i} x_{2i}}{\Sigma x_{2i}^2} .$$

اگر یک مدل رگرسیون بسازیم که در آن X_{3i} بر حسب X_{2i} بیان شده باشد،

$$X_{3i} = \gamma + \beta_{32} X_{2i} + \varepsilon_i ,$$

آنگاه

$$\hat{\beta}_{32} = \frac{\Sigma x_{3i} x_{2i}}{\Sigma x_{2i}^2} .$$

نتیجه می‌گیریم که

$$\text{مقدار اریب} = E(\hat{\beta}_2) - \beta_2 = \beta_3 \hat{\beta}_{32} .$$

۴ـ۷ با توجه به $C_i = Y_i - S_i$، به وضوح ملاحظه می‌شود که در تخمین تابع مصرف خواهیم داشت: $\hat{\alpha} = 0$ ، $\hat{\beta} = 1$ و $\hat{\gamma} = -1$. ضریب تعیین (R^2) نیز برابر یک خواهد بود. همچنین مجموع پسماندها نیز صفر می‌شود؛ بنابراین نتیجه می‌گیریم که مدل مفروض در واقع تابع تغییرات ساختاری مصرف را بیان نمی‌کند، بلکه مبتنی بر یک رابطهٔ حسابداری است. در این گونه موارد باید در مدل اصلی تجدیدنظر کرد به گونه‌ای که یکی از متغیرهای توضیحی حذف شود. استفاده از نظریه‌های اقتصادی، باید محور اصلی این تجدیدنظرها باشد.

۴ـ۸ الف) $R^2_{y/1234} = 0/86$ بر این دلالت می‌کند که ضریب تعیین در یک مدل رگرسیون که Y_i را بر حسب چهار متغیر توضیحی X_{1i}، X_{2i}، X_{3i} و X_{4i} بیان می‌کند، برابر است با ۰/۸۶. اما $R^2_{y/123} = 0/89$ نشان می‌دهد که Y_i بر حسب سه متغیر توضیحی X_{1i}، X_{2i} و X_{3i} بیان شده است؛ با وجود این، ضریب تعیین آن ۰/۸۹ است. می‌دانیم که اگر تعداد متغیرهای توضیحی اضافه شود، قدرت توضیحی مدل بیشتر شده یا حداقل ثابت باقی می‌ماند؛ بنابراین $R^2_{y/1234}$ نمی‌تواند کمتر از $R^2_{y/123}$ شود. نتیجه می‌گیریم که در مورد (الف) قطعاً اشتباهی واقع شده است.

ب) $R^2_{y/123}$ نشان می‌دهد که در مدل رگرسیون مفروض سه‌متغیر توضیحی وجود

دارد. اطلاعات موجود فقط در مورد دو ضریب تعیین جزئی ارائه شده است؛ بنابراین هیچ راهی وجود ندارد که بتوان صحت رابطهٔ $R^2_{y/12} = 0/70$ را با توجه به دو اطلاع داده شده ارزیابی کرد.

۴ـ۹ الف) با توجه به معادلهٔ ۴ـ۶۰ داریم

$$R^2_{y/12} = \frac{r^2_{y1} + r^2_{y2} - 2\, r_{y1}\, r_{y2}\, r_{21}}{(1 - r^2_{12})} .$$

بنابراین چهار متغیر در ارتباط با یکدیگر قرار دارند. در قسمت (الف) این مسأله سه کمیت در مورد سه متغیر ارائه شده است: r^2_{y1} ، r^2_{y2} و $R^2_{y/12}$. اما مقدار r_{12} ـ که در واقع همان r_{21} است ـ معین نشده است و می‌تواند مقادیر متعددی داشته باشد که با توجه به معادلهٔ فوق قابل محاسبه است. کافی است این مقادیر را در معادلهٔ فوق قرار دهیم. خواهیم داشت

$$0/701 = \frac{0/227 + 0/126 - 2\left(\sqrt{0/227\,(0/126)}\right) r_{21}}{(1 - r^2_{12})} ,$$

بعد از ساده کردن داریم

$$0/701\ r^2_{12} - 0/33382\ r_{12} - 0/348 = 0 .$$

ریشه‌های این معادله برابر است با

$$r_{12} = 0/987 \quad , \quad r_{12} = -0/504 .$$

مشاهده می‌شود که هر دو مقدار r_{12} ، مقادیر قابل قبولی از ضریب همبستگی است؛ بنابراین نتیجه می‌گیریم که داده‌های مسأله می‌تواند صحیح باشد.

ب) داده‌های مسأله در این قسمت غلط است؛ زیرا می‌دانیم

$$r^2_{x,y} = \frac{(\Sigma\, x_t\, y_t)^2}{\Sigma\, x_t^2\ \Sigma y_t^2} .$$

از طرف دیگر می‌دانیم

$$0 \leq r^2_{x,y} \leq 1 \ ,$$

باید داشته باشیم:

$$\sum x_t^2 \ \sum y_t^2 \geq (\sum x_t \ y_t)^2 \ ,$$

یا

$$\sum x_t^2 \ \sum y_t^2 - (\sum x_t \ y_t)^2 \geq 0 \ .$$

اما مقدار داده شده برای عبارت فوق برابر ۱۷۳۲/۸۶- است؛ پس نتیجه می‌گیریم که قطعاً اشتباهاتی در محاسبات واقع شده است.

۴ـ۱۰ الف) از تعریف R^2 شروع می‌کنیم،

$$R^2 = \frac{ESS}{TSS} = \frac{\Sigma \hat{y}_t^2}{\Sigma y_t^2} \ .$$

به جای $\hat{y}_t$ مقدار آن را از رابطه زیر قرار می‌دهیم،

$$\hat{y}_t = \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t} \ .$$

نتیجه می‌گیریم که

$$R^2 = \frac{\hat{\beta}_1^2 \Sigma x_{1t}^2 + \hat{\beta}_2^2 \Sigma x_{2t}^2 + 2 \hat{\beta}_1 \hat{\beta}_2 \Sigma x_{1t} \ x_{2t}}{\Sigma y_t^2} \ .$$

ب) دقیقاً همان فرمول ۱۹ـ۴ در متن کتاب است.

ج) دقیقاً همان فرمول ۶۰ـ۴ در متن کتاب است.

د) اگر فرمول ۵۳ـ۴ را مجذور کنیم، این رابطه به اثبات می‌رسد.

۴ـ۱۱ با استفاده از تعریف ضریب همبستگی بین دو متغیر، می‌توان چنین نوشت

$$r_{\hat{\beta}_1 \hat{\beta}_2} = \frac{E\left[(\hat{\beta}_1 - \beta_1)(\hat{\beta}_2 - \beta_2)\right]}{\sqrt{E(\hat{\beta}_1 - \beta_1)^2 \, E(\hat{\beta}_2 - \beta_2)^2}} \ ,$$

یا

$$r^{۲}_{\hat{\beta}_۱\hat{\beta}_۲} = \frac{\left\{E\left[(\hat{\beta}_۱ - \beta_۱)(\hat{\beta}_۲ - \beta_۲)\right]\right\}^{۲}}{E(\hat{\beta}_۱ - \beta_۱)^{۲} E(\hat{\beta}_۲ - \beta_۲)^{۲}} ,$$

$$= \frac{\left[\text{Cov}(\hat{\beta}_۱, \hat{\beta}_۲)\right]^{۲}}{\text{Var}(\hat{\beta}_۱)\,\text{Var}(\hat{\beta}_۲)}$$

از طرف دیگر، با توجه به معادلهٔ ٤-٢٦ می‌دانیم

$$\text{Cov}(\hat{\beta}_۱, \hat{\beta}_۲) = \frac{-\sigma^{۲} r^{۲}_{۱۲}}{\Sigma x_{۱t} x_{۲t}(۱ - r^{۲}_{۱۲})} = \frac{-\sigma^{۲}(\Sigma x_{۱t} x_{۲t})^{۲}}{\Sigma x^{۲}_{۱t} \Sigma x^{۲}_{۲t} \Sigma x_{۱t} x_{۲t}(۱ - r^{۲}_{۱۲})} ,$$

$$= \frac{-\sigma^{۲} \Sigma x_{۱t} x_{۲t}}{\Sigma x_{۱t} \Sigma x_{۲t}(۱ - r_{۱۲})} .$$

با استفاده از فرمول فوق و نیز فرمولهای ٤-٢٤ و ٤-٢٥ و جایگزینی در فرمول $r^{۲}_{\hat{\beta}_۱\hat{\beta}_۲}$ خواهیم داشت

$$r^{۲}_{\hat{\beta}_۱\hat{\beta}_۲} = \frac{\sigma^{٤}(\Sigma x_{۱t} x_{۲t})^{۲}(\Sigma x^{۲}_{۱t})(۱ - r^{۲}_{۱۲})(\Sigma x^{۲}_{۲t})(۱ - r^{۲}_{۱۲})}{(\Sigma x^{۲}_{۱t})^{۲}(\Sigma x^{۲}_{۲t})^{۲}(۱ - r^{۲}_{۱۲})^{۲}\sigma^{۲}\sigma^{۲}} ,$$

$$= \frac{(\Sigma x_{۱t} x_{۲t})^{۲}}{\Sigma x^{۲}_{۱t} \Sigma x^{۲}_{۲t}} = r^{۲}_{۱۲} .$$

٤-١٢ الف) مدل جدید را که در آن مقیاس اندازه گیری متغیرها تغییر یافته است، به صورت زیر می‌نویسیم،

$$\frac{Y^{*}_t}{\lambda} = \frac{\beta_۱ X^{*}_{۱t}}{\mu_۱} + \frac{\beta_۲ X^{*}_{۲t}}{\mu_۲} + U_t ,$$

یا

$$Y_t^* = \frac{\lambda \beta_1 X_{1t}^*}{\mu_1} + \frac{\lambda \beta_2 X_{2t}^*}{\mu_2} + \lambda U_t \ ,$$

$$= \beta_1^* X_{1t}^* + \beta_2^* X_{2t}^* + U_t^* \ . \qquad (1)$$

تخمین β_1^* از مدل جدید عبارت است از

$$\hat{\beta}_1^* = \left(\frac{\hat{\lambda} \beta_1}{\mu_1}\right) = \frac{\lambda}{\mu_1} \hat{\beta}_1 \ ,$$

در نتیجه

$$\hat{\beta}_1^* = \frac{\lambda}{\mu_1} \hat{\beta}_1 \ .$$

به همین ترتیب می‌توان نشان داد که

$$\hat{\beta}_2^* = \frac{\lambda}{\mu_2} \hat{\beta}_2 \ .$$

ب) اگر ضریب تعیین در مدل جدید را با R^{*2} نشان دهیم، با توجه به معادلهٔ ۴-۲۰ خواهیم داشت

$$R^{*2} = \frac{\Sigma (Y_t^* \hat{Y}_t^*)^2}{\Sigma Y_t^{*2} \ \Sigma \hat{Y}_t^{*2}} \ .$$

عبارت $Y_t^* = \lambda Y_t$ و $\hat{Y}_t^* = \lambda \hat{Y}_t$ را در فرمول فوق جایگزین می‌کنیم،

$$R^{*2} = \frac{\Sigma (\lambda Y_t \ \lambda \hat{Y}_t)^2}{\Sigma \lambda^2 Y_t^2 \ \Sigma \lambda^2 \hat{Y}_t^2} = \frac{\Sigma (Y_t \hat{Y}_t)^2}{\Sigma Y_t^2 \ \Sigma \hat{Y}_t^2} = R^2 \ .$$

نتیجه می‌گیریم که تغییر در مقیاس اندازه‌گیری متغیرها در ضریب تعیین تأثیری ندارد. همچنین اگر انحراف معیار رگرسیون در مدل جدید را با SEE^* نشان دهیم،

با توجه به معادلهٔ ۱.۴۵ داریم

$$SEE^* = \sqrt{\frac{\Sigma e_t^{*2}}{n-2}} = \sqrt{\frac{\Sigma \hat{U}_t^{*2}}{n-2}} \ .$$

با توجه به معادلهٔ (۱)، می‌دانیم: $\hat{U}_t^* = \lambda \hat{U}_t^*$

$$SEE^* = \sqrt{\frac{\Sigma (\lambda \hat{U}_t^{*2})}{(n-2)}} = \lambda \sqrt{\frac{\Sigma \hat{U}_t^2}{n-2}} = \lambda \sqrt{\frac{\Sigma e_t^2}{n-2}} \ ,$$

در نتیجه خواهیم داشت

$$SEE^* = \lambda \, SEE \ ,$$

یعنی تغییر مقیاس در متغیر درون‌زا، انحراف معیار رگرسیون را به اندازهٔ λ تغییر می‌دهد.

فصل پنجم

مفاهیم و تخمین مدل رگرسیون خطی چند متغیره

۵ـ۱ مقدمه

مدلهایی که تا به حال مطالعه کرده‌ایم، شامل یک متغیر توضیحی و به صورت

$$Y_t = \alpha + \beta X_t + U_t ,$$

یا دو متغیر توضیحی و از نوع:

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t ,$$

بوده‌اند. اما می‌توان فرض کرد که در یک مدل رگرسیون، k متغیر توضیحی وجود داشته باشد،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt}$$

برای سهولت در محاسبات، تخمین این گونه مدلها با جبر ماتریسی انجام می‌شود؛ بنابراین باید ابتدا مدل رگرسیون چند متغیره را به زبان ماتریسی نوشت، سپس آن را با روش حداقل مربعات معمولی تخمین زد. یک پیوست ماتریسی نیز در بخش پیوستها در نظر گرفته شده است[1] تا بتوان تعاریف و عملیات اصلی ماتریسها را در حد ضرورت مرور کرد. بنابراین آشنایی با جبر ماتریسی پیش‌نیاز این فصل محسوب می‌شود. با وجود این، سعی شده است بسیاری از مطالب اصلی به گونه‌ای ارائه شود که آشنایی با مفاهیم و عملیات مقدماتی ماتریسها کافی باشد. در قسمت ۵ـ۲ یک مدل رگرسیون چند متغیره و نیز فرضهای کلاسیک آن به زبان ماتریسی ارائه شده است. رسیدن به فرمولهای تخمین

---

۱. پیوست «۵ ـ الف»

پارامترها در رگرسیون چند متغیره، با استفاده از مشتق‌گیری ماتریسی بسیار ساده است. باید این نکته را نیز در نظر گرفت که محاسبات تخمین پارامترها می‌تواند برحسب مشاهدات اصلی یا انحراف از میانگین باشد. توجه به خصوصیات ماتریسها در هریک از این دو حالت اهمیت فراوان دارد. این نکات موضوع قسمت ۳ـ۵ است.

ضریب تعیین در مدلهای رگرسیون چند متغیره، موضوع قسمت ٤ـ۵ خواهد بود. در این قسمت فرمولهای مختلفی برای محاسبهٔ R^2 به زبان ماتریسی ارائه شده است. در فصل اول دیدیم که اگر فرمول عمومی ضریب تعیین را برای مدلی به کار بریم که فاقد جملهٔ ثابت است، چه بسا ممکن است، مقدار r^2 منفی شود. در این قسمت، این نکته را برای مدلهای رگرسیون چندمتغیره تعمیم داده‌ایم و همچنین شرایطی را که R^2 می‌تواند منفی شود، ارزیابی کرده‌ایم.

یکی از مسائل مهمی که در رگرسیونهای چندمتغیره مطرح می‌شود، افزایش یا حداقل عدم کاهش R^2 به ازای افزایش متغیرهای توضیحی است. این مسأله را ـ که ریشه در تعریف R^2 دارد ـ می‌توان با وارد کردن درجات آزادی در تعریف R^2 تا حدی تعدیل کرد. به همین دلیل در مواردی باید ضرورتاً از معیار ضریب تعیین تعدیل شده یا $\bar{R}^2$ استفاده کرد. مباحث $\bar{R}^2$ و رابطهٔ آن با R^2 در قسمت ۵ـ۵ مطرح شده است. همچنین در این قسمت نشان داده‌ایم که آن دسته از متغیرهای توضیحی که واریانس جملهٔ اختلال را حداقل می‌کنند، ضریب تعیین تعدیل شده را حداکثر خواهند کرد. بدیهی است این امر می‌تواند به منزلهٔ یکی از معیارهای مفید در حفظ یا حذف متغیرهای توضیحی مورد استفاده قرار گیرد.

در قسمت ٦ـ۵ نکاتی پیرامون R^2 و $\bar{R}^2$ مطرح شده است. معمول این است که در تخمین یک مدل رگرسیون، هرگاه مقدار R^2 زیاد باشد، این امر را نشانه‌ای از موفقیت یا خوبی تخمین می‌دانند. در این قسمت سعی بر این است تا به این سؤال پاسخ داده شود که آیا چنین استنتاجی صحیح است؟ مسأله حداکثرسازی R^2 یا استفاده از R^2 در مقایسه تخمینهای مختلف از یک مدل یا مقایسه تخمین مدلهای مختلف، یکی از مباحث بحث‌انگیز اقتصادسنجی است، بنابراین سعی شده است که در این قسمت ابعاد مختلف

آن روشنتر شود. سرانجام در قسمت ۵ـ۷ مسائل بیان ماتریسی R^2 برحسب مقادیر اصلی متغیرها به طور خلاصه بررسی خواهد شد.

۵ـ۲ بیان ماتریسی رگرسیون چندمتغیره

مدل رگرسیون چندمتغیره زیر را ملاحظه کنید،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t$$

که در آن $t = 1, 2, \ldots, n$. اگر بخواهیم این مدل را به زبان ماتریسی بیان کنیم، پارامترهای مدل باید با یک بردار مشخص شود. انتظار این است که این بردار را $\boldsymbol{\beta}$ بنامیم. در این صورت عناصر بردار $\boldsymbol{\beta}$ عبارت است از:

$$\boldsymbol{\beta}' = [\alpha, \beta_1, \beta_2, \ldots, \beta_k] \cdot$$

بیان فوق از بردار $\boldsymbol{\beta}$ چندان مرسوم نیست؛ زیرا اولین عنصر آن α است؛ بنابراین باید نوع نوشتن مدل رگرسیون چندمتغیره را تغییر دهیم. یک راه حل این است که مدل را به صورت زیر بنویسیم،

$$Y_t = \beta_1 + \beta_2 X_{1t} + \beta_3 X_{2t} + \cdots + \beta_{k+1} X_{kt} + U_t \cdot$$

در این حالت، بردار $\boldsymbol{\beta}$ عبارت خواهد بود از:

$$\boldsymbol{\beta}' = [\beta_1, \beta_2, \beta_3, \ldots, \beta_k] \cdot$$

ملاحظه می‌شود که مشکل فوق برای $\boldsymbol{\beta}$ رفع شده است؛ زیرا دیگر پارامتر α در بین عناصر بردار $\boldsymbol{\beta}$ وجود ندارد. اما مسألهٔ دیگری ایجاد شده و آن عدم تطابق اندیس پارامترهای $\boldsymbol{\beta}$ با اندیس متغیرهای توضیحی است؛ به عبارت دیگر، معمول این است که مثلاً β_2 ضریب X_{2t} باشد، اما در اینجا β_2 ضریب X_{1t} فرض شده است. برای رفع این مشکل بهترین راه حل این است که مدل رگرسیون را به صورت زیر بنویسیم،

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t ,$$

سپس تعریف کنیم که $X_{i1} \equiv 1$.

بدین ترتیب، خواهیم داشت:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + U_i \ . \qquad (5\text{-}1)$$

بنابراین مدل رگرسیون مفروض شامل k متغیر توضیحی است که مقدار متغیر توضیحی اول آن همواره یک فرض شده است. همچنین مدل رگرسیون مفروض، شامل k پارامتر βاست که باید تخمین زده شود. سؤال این است که چگونه می‌توان معادلهٔ ۵ـ۱ را به زبان ماتریسی نوشت؟

اگر نمونهٔ ما، شامل n مشاهده باشد، در آن صورت n معادله از نوع ۵ـ۱ خواهیم داشت. چون فرض بر این است که آمار برحسب سری زمانی است؛ بنابراین برای n دوره، n معادله به شرح زیر حاصل می‌شود،

$$\begin{aligned}
Y_1 &= \beta_1 + \beta_2 X_{21} + \beta_3 X_{31} + \cdots + \beta_k X_{k1} + U_1 \ , \\
Y_2 &= \beta_1 + \beta_2 X_{22} + \beta_3 X_{32} + \cdots + \beta_k X_{k2} + U_2 \ , \\
&\vdots \\
Y_n &= \beta_1 + \beta_2 X_{2n} + \beta_3 X_{3n} + \cdots + \beta_k X_{kn} + U_n \ .
\end{aligned} \qquad (5\text{-}2)$$

سیستم n معادله فوق را می‌توان به صورت زیر نوشت،

$$\underbrace{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} 1 & X_{21} & X_{31} \cdots & X_{k1} \\ 1 & X_{21} & X_{32} \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{2n} & X_{3n} \cdots & X_{kn} \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix}}_{\mathbf{u}} , \qquad (5\text{-}3)$$

یا

$$\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \mathbf{u} \qquad (5\text{-}4)$$

که در آن $\mathbf{y}$ بردار مقادیر متغیر درون‌زا و $n\times 1$ است $(\mathbf{y} \to n\times 1)$، و $\mathbf{X}$ ماتریس مقادیر متغیرهای برون‌زا یا توضیحی و $n\times k$ است، $(\mathbf{X} \to n\times k)$. پارامترهای مدل با بردار $\boldsymbol{\beta}$ مشخص شده است؛ بنابراین $k\times 1$ خواهد بود $(\boldsymbol{\beta} \to k\times 1)$، سرانجام بردار $\mathbf{u}$ شامل مقادیر جمله‌های اختلال بوده، $n\times 1$ می‌باشد $(\mathbf{u} \to n\times 1)$. لازم است یادآوری شود که عناصر ماتریس $\mathbf{X}$ برخلاف روال معمول در ماتریسها نوشته شده است. می‌دانیم در ماتریسها، عنصر a_{ij} بیان کنندهٔ مقداری است که در ردیف i و ستون j قرار دارد؛ مثلاً $X_{۳۲}$ باید عنصری باشد که در ردیف سوم و ستون دوم قرار می‌گیرد. اما در معادلهٔ ۵ـ۴ ملاحظه می‌شود که مثلاً عنصر $X_{۳۲}$ در ماتریس $\mathbf{X}$، در ردیف دوم و ستون سوم واقع شده است. این جابجایی به طور اساسی مسألهٔ خاصی را ایجاد نمی‌کند و بهتر است به همین شکل حفظ شود؛ زیرا در غیر این صورت علامتگذاری را در معادلات پیچیده‌تر خواهد کرد.

معادلهٔ ۳ـ۵ را می‌توان به صورت زیر نیز نوشت،

$$\mathbf{y}= \beta_۱ + \beta_۲\mathbf{X}_۲ + \beta_۳\mathbf{X}_۳ + \cdots + \beta_k\mathbf{X}_k + \mathbf{u} \qquad (۵\_۵)$$

که در آن $\mathbf{X}_k \to n\times 1$. معادلهٔ ۵ـ۵ بر این دلالت می‌کند که بردار مشاهدات متغیر درون‌زا برابر با مجموع بردار جمله‌های اختلال $\mathbf{u}$ و یک ترکیب خطی از ستونهای ماتریس $\mathbf{X}$ است.

بعد از آشنایی با بیان ماتریسی یک مدل رگرسیون چندمتغیره، به بررسی فرضهای کلاسیک در مدلهای رگرسیون چندمتغیره می‌پردازیم.

۱. فرضهای کلاسیک مدلهای رگرسیون چندمتغیره

در این قسمت فرضهای کلاسیک جمله‌های اختلال در بردار $\mathbf{u}$ و نیز فرضهای کلاسیک متغیرهای توضیحی را بررسی می‌کنیم. در فصل چهارم دیدیم که باید فرضهای میانگین، واریانس، کوواریانس و سرانجام تابع توزیع احتمال U_i را بررسی کنیم. در مورد متغیرهای توضیحی نیز دو فرض را مطرح کردیم. روال بحث دقیقاً مانند مطالب مطرح

شده در قسمت ۲ـ٤ است، با این تفاوت که در این قسمت از بیان ماتریسی استفاده خواهد شد.

ابتدا می‌گوییم که در معادلهٔ ۵ـ۵، مقادیر $\beta_1, \beta_2, \ldots, \beta_k$ به عنوان پارامترهای جامعه، برای ما مجهول است و باید آنها را تخمین بزنیم. حتی اگر این پارامترها معلوم باشد، ترکیب خطی $(\beta_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_k X_k)$ نمی‌تواند بردار $\mathbf{y}$ را به طور دقیق برای ما مشخص کند؛ زیرا قطعاً متغیرهای دیگری هم وجود دارد که می‌تواند تغییرات متغیر درون‌زا را توضیح دهد، اما در مدل ما وارد نشده است. معادله‌ای که می‌خواهد رفتار یک متغیر اقتصادی را توضیح دهد، معمولاً دقیق و معین نیست بلکه تصادفی است. نتیجه می‌گیریم که بردار $\mathbf{u}$ در مدل ۵ـ۵ شامل متغیرهای تصادفی است و در نتیجه یک بردار تصادفی خواهد بود. سؤال این است که میانگین، واریانس، کوواریانس و تابع توزیع احتمال آن چیست؟ لازم است توضیح داده شود که اطلاعات ما از خصوصیات $\mathbf{u}$ در نهایت روشن کنندهٔ این نکته خواهد بود که بردار $\mathbf{y}$ چگونه ایجاد شده است.

اولین فرض، به میانگین یا امید ریاضی $\mathbf{u}$ مربوط است. مانند آنچه در فصل اول و در نمودار ۱ـ۱ و معادلهٔ ۹ـ۱ مطرح شد، می‌توان گفت که در معادله‌های ۲ـ۵، برای هریک از مقادیر U_t، $t = 1, 2, \ldots, n$، فرض بر این است که $E(U_t) = 0$؛ بنابراین

$$\begin{bmatrix} E(U_1) \\ E(U_2) \\ \vdots \\ E(U_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

یا

$$E(\mathbf{u}) = \mathbf{0} \qquad (۵\text{ـ}۶)$$

که در آن $\mathbf{0}$ ماتریس $n \times 1$ است. از دو طرف معادلهٔ ۴ـ۵، امید ریاضی می‌گیریم. داریم

$$E(\mathbf{y}) = E(\mathbf{X}\boldsymbol{\beta}) + E(\mathbf{u}) \quad .$$

با توجه به معادلهٔ ۵ـ۶ و نیز در نظر گرفتن این نکته که X در آزمایشهای فرضی تکراری ثابت نگه‌داشته می‌شود، خواهیم داشت:

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} . \qquad (۵\text{ـ}۷)$$

معادلهٔ ۵ـ۷ نتیجهٔ مستقیم اولین فرض u است.

دومین و سومین فرض u به واریانس و کوواریانس آن مربوط است. ابتدا ماتریسی می‌سازیم که عناصر قطری و غیرقطری آن به‌ترتیب واریانس و کوواریانس عناصر بردار u باشد. این ماتریس را معمولاً با $\boldsymbol{\Sigma}_u$ نشان می‌دهند. برای ساختن این ماتریس باید ابتدا ترانهاد بردار u را حساب کرده، سپس ماتریس $\mathbf{uu}'$ را تشکیل دهیم. چون $\mathbf{u} \to n\times ۱$، بنابراین $\mathbf{u}' \to ۱\times n$، و لذا $\mathbf{uu}' \to n\times n$. اگر از ماتریس $\mathbf{uu}'$ امید ریاضی بگیریم، ماتریس واریانس ـ کوواریانس مقادیر مختلف جمله‌های اختلال به ترتیب زیر به دست می‌آید،

$$E(\mathbf{uu}') = E\begin{bmatrix} U_۱ \\ U_۲ \\ \vdots \\ U_n \end{bmatrix}[U_۱ U_۲ \cdots U_n] = E\begin{bmatrix} U_۱^۲ & U_۱U_۲ & \cdots & U_۱U_n \\ U_۲U_۱ & U_۲^۲ & \cdots & U_۲U_n \\ \vdots & & & \\ U_nU_۱ & U_nU_۲ & \cdots & U_n^۲ \end{bmatrix},$$

یا

$$E(\mathbf{uu}') = \begin{bmatrix} E(U_۱^۲) & E(U_۱U_۲) & \cdots & E(U_۱U_n) \\ E(U_۲U_۱) & E(U_۲)^۲ & \cdots & E(U_۲U_n) \\ \vdots & & & \\ E(U_nU_۱) & E(U_nU_۲) & \cdots & E(U_n)^۲ \end{bmatrix}.$$

می‌دانیم واریانس U_i برابر است با

$$\mathrm{Var}(U_i) = E[U_i - E(U_i)]^۲ = E(U_i - ۰)^۲,$$

$$= E(U_i)^۲ .$$

عناصر قطری ماتریس $E(uu')$، درواقع واریانس مقادیر $U_1, U_2, \ldots, U_n$ است. همچنین می‌دانیم کوواریانس U_i و U_j برابر است با:

$$\begin{aligned} \text{Cov}(U_i, U_j) &= E[U_i - E(U_i)][U_j - E(U_j)], \\ &= E(U_i - 0)(U_j - 0), \\ &= E(U_i U_j). \end{aligned}$$

بنابراین کوواریانس U_i و U_j به ازای تمام مقادیر i و j به استثنای مقادیر $i=j$، برابر عناصر غیرقطری ماتریس $E(uu')$ است. بدین ترتیب می‌توان نوشت،

$$\Sigma_u = E(uu') = \begin{bmatrix} \text{Var}(U_1) & \text{Cov}(U_1, U_2) & \cdots & \text{Cov}(U_1, U_n) \\ \text{Cov}(U_2, U_1) & \text{Var}(U_1) & \cdots & \text{Cov}(U_2, U_n) \\ \vdots & & & \\ \text{Cov}(U_n, U_1) & \text{Cov}(U_n, U_2) & \cdots & \text{Var}(U_n) \end{bmatrix}, \tag{۵-۸}$$

می‌دانیم یکی از فرضهای کلاسیک، U_i، فرض واریانس همسانی است. براساس این فرض، معادلهٔ ۵-۸ را می‌توان چنین نوشت،

$$\Sigma_u = E(uu') = \begin{bmatrix} \sigma^2 & \text{Cov}(U_1, U_2) & \cdots & \text{Cov}(U_1, U_n) \\ \text{Cov}(U_2, U_1) & \sigma^2 & \cdots & \text{Cov}(U_2, U_n) \\ \vdots & & & \\ \text{Cov}(U_n, U_1) & \text{Cov}(U_n, U_2) & \cdots & \sigma^2 \end{bmatrix}, \tag{۵-۹}$$

که در آن σ^2 برابر واریانس جملهٔ اختلال است.

عدم خودهمبستگی، یعنی صفر بودن کوواریانس هر دو مقدار U_i و U_j، یکی دیگر از فرضهای کلاسیک جمله‌های اختلال است؛ یعنی $\text{Cov}(U_i, U_j) = 0$. براساس

این فرض، تمام جمله‌های غیرقطری ماتریس Σ_u برابر صفر می‌شود؛ یعنی Σ_u یک ماتریس قطری خواهد بود. با استفاده از معادلهٔ ۸ـ۵ داریم:

$$\Sigma_u = E(uu') = \begin{bmatrix} \text{Var}(U_1) & 0 & \cdots & 0 \\ 0 & \text{Var}(U_2) & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \text{Var}(U_n) \end{bmatrix} \quad (۵ـ۱۰)$$

اگر هر دو فرض واریانس همسانی و عدم خود همبستگی را با هم در نظر بگیریم، در معادلهٔ ۸ـ۵، تمام عناصر قطری برابر σ^2 و عناصر غیرقطری برابر صفر می‌شود و ماتریس واریانس ـ کوواریانس Σ_u را می‌توان به صورت زیر نوشت،

$$\Sigma_u = E(uu') = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}.$$

از σ^2 فاکتور گرفته خواهیم داشت:

$$\Sigma_u = E(uu') = \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

می‌دانیم ماتریسی که هریک از عناصر قطری آن یک و همهٔ عناصر غیرقطری آن صفر باشد، یک ماتریس «یکه»[1] است که معمولاً آن را با $\mathbf{I}$ نشان می‌دهند. بدین ترتیب، ماتریس واریانس ـ کوواریانس جمله‌های اختلال مدل رگرسیون، با فرضهای واریانس

1. Identity Matrix

ناهمسانی و عدم خودهمبستگی برابر است با

$$\boldsymbol{\Sigma}_{\mathbf{u}} = \mathrm{E}(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{I}. \qquad (۵\text{-}۱۱)$$

چهارمین فرض $\mathbf{u}$ به تابع توزیع احتمال آن مربوط است. فرض می‌شود که تابع توزیع احتمال $\mathbf{u}$ نرمال است. با توجه به میانگین $\mathbf{0}$ و واریانس $\sigma^2\mathbf{I}$، می‌توان چنین نوشت،

$$\mathbf{u} \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I}) \qquad (۵\text{-}۱۲)$$

بعد از بیان فرضهای $\mathbf{u}$، باید به بررسی فرضهای ماتریس متغیرهای توضیحی، یعنی $\mathbf{X}$ بپردازیم. اولین فرض این است که $\mathbf{X}$ یک ماتریس غیرتصادفی است. این فرض بدین معنی است که اگر یک نمونهٔ $\mathbf{n}$ تایی دیگر از مشاهدات داشته باشیم، ماتریس متغیرهای توضیحی، یعنی $\mathbf{X}$ بدون تغییر باقی می‌ماند. در ابتدا به نظر می‌رسد، قبول چنین فرضی بسیار دشوار باشد؛ زیرا انتظار این است که مقادیر متغیرهای توضیحی از یک نمونه به نمونهٔ دیگر فرق کند. اما ـ همان گونه که در فصل اول اشاره کردیم ـ منظور از فرض اخیر این است که متغیرهای توضیحی را در آزمایشهای تکراری ثابت نگه‌می‌داریم تا به این ترتیب بتوانیم تأثیر تغییرات U_i بر Y_i را مشاهده کنیم. ممکن است این نکته مطرح شود که در علوم اجتماعی، ثابت نگه‌داشتن متغیرهای توضیحی و تکرار آزمایش معمولاً امکانپذیر نیست. دقیقاً به همین علت است که باید از آزمایشهای فرضی تکراری صحبت کنیم. بنابراین فرض غیرتصادفی بودن $\mathbf{X}$، معمولاً در اقتصاد نمی‌تواند مابه ازای عینی داشته باشد و تنها به این سبب مطرح می‌شود که بتواند خصوصیات تأثیرگذاری متغیر تصادفی U_i را بر تغییرات Y_i بیان کند. درواقع درک استنتاجات آماری به شرط ثبات $\mathbf{X}$، می‌تواند بسیار روشن کننده باشد. همان گونه که در جلد دوم این کتاب خواهیم دید، می‌توان فرض غیرتصادفی بودن $\mathbf{X}$ را کنار گذاشت و تأثیر آن را در استنتاجات قبلی مشاهده کرد.

دومین فرض مربوط به ماتریس $\mathbf{X}$ این است که بگوییم متغیرهای توضیحی، همخطی کامل ندارد؛ یعنی در یک رابطهٔ خطی کامل با یکدیگر قرار نمی‌گیرد. ـ همان‌گونه که در قسمت ۲ـ۴ بیان شد ـ اگر بین متغیرهای توضیحی همبستگی کامل

خطی برقرار باشد، تخمین مستقل هریک از پارامترها، عملاً غیرممکن خواهد بود. بررسی مسألهٔ همخطی، موضوع جلد دوم این کتاب خواهد بود. با استفاده از تعریف «رتبهٔ ماتریس»[1] می‌توان این فرض را به صورت زیر بیان کرد،

$$\rho(\mathbf{X}) = \mathbf{k}$$

که در آن $\rho(\mathbf{X})$، درواقع مرتبهٔ ماتریس $\mathbf{X}$ و $\mathbf{k}$ تعداد متغیرهای توضیحی است[2].

اکنون که با فرضهای کلاسیک مدلهای رگرسیون چندمتغیره آشنا شدیم، می‌توان به تخمین مدل به زبان ماتریسی پرداخت. اما بهتر است قبل از آن، به بیان ماتریسی جمله‌های پسماند نیز اشاره کنیم تا مبحث بیان ماتریسی رگرسیون چندمتغیره کاملتر شود.

۲. بیان ماتریسی جمله‌های پسماند

مدل رگرسیون چندمتغیره (۱ـ۵) را یک بار دیگر می‌نویسیم،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t .$$

اگر به جای پارامترهای مدل، یعنی $\beta_k, \ldots, \beta_2, \beta_1$، مقادیر تخمین آنها را قرار دهیم، تخمین مدل رگرسیون به دست می‌آید؛ بنابراین:

$$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_{2t} + \cdots + \hat{\beta}_k X_{kt} . \qquad (۵ـ۱۳)$$

تفاوت بین مقدار $\hat{Y}_t$ ـ که از معادلهٔ ۱۳ـ۵ حاصل می‌شود ـ با مقدار مشاهده شدهٔ Y_t برابر با مقدار پسماند است؛ بنابراین با توجه به تعریف

پسماند = تخمین ـ مشاهده

---

1. Rank of a Matrix

به پیوست «۵ ـ الف» مراجعه شود.

۲. می‌توان فرض سومی نیز برای متغیرهای توضیحی مطرح کرد: تعداد مشاهدات بیشتر از تعداد پارامترهایی است که باید تخمین زده شود؛ یعنی $n > k$.

خواهیم داشت:

$$e_t = Y_t - (\hat{\beta}_1 + \hat{\beta}_2 X_{t1} + \cdots + \hat{\beta}_k X_{kt}) \qquad (۵\_۱۴)$$

که در آن e_t پسماند و $t = ۱, ۲, \ldots, n$.

برای بیان ماتریسی پسماندها، باید به جای مقادیر β در سیستم معادله‌های ۳ـ۵، مقادیر $\hat{\beta}$ را قرار دهیم. اگر بردار $\hat{\beta}$ را به صورت زیر تعریف کنیم،

$$\hat{\beta} = [\hat{\beta}_1\ \hat{\beta}_2 \cdots \hat{\beta}_k]\,,$$

با توجه به معادلهٔ ۴ـ۵ خواهیم داشت:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}\,, \qquad (۵\_۱۵)$$

که در آن $\hat{\mathbf{y}} \rightarrow n\times ۱$. با استفاده از معادلهٔ ۱۴ـ۵، می‌توان بردار پسماند، یعنی $\mathbf{e}$ را چنین نوشت،

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}\,, \qquad (۵\_۱۶)$$

که در آن $\hat{\mathbf{e}} \rightarrow n\times ۱$. معادلهٔ ۱۶ـ۵ را می‌توان به صورت زیر نوشت،

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\beta}\,. \qquad (۵\_۱۷)$$

۵ـ۳ تخمین مدل رگرسیون خطی چندمتغیره: روش حداقل مربعات معمولی

می‌دانیم معیار روش حداقل مربعات معمولی این است که باید پارامترها را چنان تخمین زد که مجموع مربعات پسماند حداقل شود؛ بنابراین ابتدا Σe_t^2 را به‌زبان ماتریسی می‌نویسیم و سپس از آن نسبت به پارامترها مشتق می‌گیریم. مقادیری از پارامترها که این مشتق را صفر کند، جواب مسأله است. ابتدا نشان می‌دهیم که $\mathbf{e}'\mathbf{e}$، بیان ماتریسی مجموع مربعات پسماند، یعنی Σe_t^2 است. برای این منظور چنین می‌نویسیم،

$$e'e = [e_1\ e_2 \cdots e_n] \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} ,$$

$$= e_1^2 + e_2^2 + \cdots + e_n^2 = \sum_{i=1}^{n} e_i^2 .$$

با جایگزینی معادلهٔ ۵ـ۱۷ در $e'e$ خواهیم داشت:

$$e'e = (y - X\hat{\beta})' (y - X\hat{\beta}) ,$$

$$= (y' - \hat{\beta}' X') (y - X\hat{\beta}) ,$$

$$= y'y - y'X\hat{\beta} - \hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta} . \qquad (۵ـ۱۸)$$

به ابعاد $y'X\hat{\beta}$ و $\hat{\beta}'X'y$ توجه می‌کنیم.

$$y'X\hat{\beta} \rightarrow (1\times n)\ (n\times k)\ (k\times 1) \rightarrow (1\times 1) ,$$

$$\hat{\beta}'X'y \rightarrow (1\times k)\ (k\times n)\ (n\times 1) \rightarrow (1\times 1) .$$

بنابراین $y'X\hat{\beta}$ و $\hat{\beta}'X'y$، هر دو اسکالر هستند و چون یکی ترانهاد دیگری است، پس با یکدیگر برابرند؛ یعنی:

$$y'X\hat{\beta} = \hat{\beta}'X'y .$$

با جایگزینی رابطه فوق در معادلهٔ ۵ـ۱۸، خواهیم داشت:

$$e'e = y'y - 2y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} . \qquad (۵ـ۱۹)$$

باید از $e'e$ در معادلهٔ فوق نسبت به $\hat{\beta}$ مشتق بگیریم. در پیوست (۵ـالف) دیدیم که در

معادلهٔ $y= ax$، داریم:

$$\frac{\partial y}{\partial x} = a' .$$

همچنین می‌دانیم که اگر داشته باشیم: $y= x' Ax$ آنگاه

$$\frac{\partial y}{\partial x} = ۲Ax .$$

در معادلهٔ ۵ـ۱۹ مجهول ما بردار $\hat{\beta}$ است. اگر $X' y$ را a و $X' X$ را برابر A بگیریم، خواهیم داشت:

$$\frac{\partial (e'e)}{\partial \hat{\beta}} = -۲X' y + ۲X' X\hat{\beta} .$$

با توجه به اینکه k پارامتر داریم؛ k معادله از این رابطه به دست می‌آید. مقادیری از این k پارامتر که معادلهٔ فوق را صفر می‌کند، تخمینهای ما خواهد بود. بدین ترتیب

$$-۲X' y + ۲X' X\hat{\beta} = ۰ ,$$

$$(X' X)\hat{\beta} = X' y \quad . \qquad (۵-۲۰)$$

معادله‌های فوق درواقع سیستم معادله‌های نرمال برای یک مدل رگرسیون با k متغیر توضیحی است. دو طرف معادلهٔ ۵ـ۲۰ را در $(X' X)^{-1}$ ضرب می‌کنیم، خواهیم داشت:

$$\hat{\beta} = (X' X)^{-1} X' y . \qquad (۵-۲۱)$$

بدیهی است برای اینکه معادلهٔ ۵ـ۲۱ برقرار باشد، باید $(X' X)$ معکوس داشته باشد. دومین شرط مربوط به متغیرهای توضیحی در مدلهای رگرسیون چندمتغیره، یعنی معادلهٔ ۵ـ۱۲ دلالت بر استقلال کامل ستونهای ماتریس X از یکدیگر دارد؛ بنابراین با توجه به معادلهٔ ۵ـ۱۲ داریم:

$$\rho (X' X) = \rho (X) = k$$

و همین امر تضمینی است که دترمینال $(X' X)$ صفر نباشد و معکوس داشته باشد.

اکنون که فرمول $\hat{\beta}$ را در مدلهای رگرسیون چندمتغیره به دست آوردیم، بهتر است به‌ساختار ماتریسهای موجود در آن توجه بیشتری‌داشته باشیم تا چگونگی استفاده از این فرمول در حل مسائل روشن شود. ابتدا به مطالعهٔ ماتریس $\mathbf{X}'\mathbf{X}$ می‌پردازیم. ماتریس $\mathbf{X}$ را در معادلهٔ ۳ـ۵ می‌نویسیم و آن را از سمت راست در ترانهاد خود ضرب می‌کنیم،

$$\mathbf{X}'\mathbf{X}=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{21} & X_{22} & & X_{2n} \\ X_{31} & X_{32} & & X_{3n} \\ \vdots & \vdots & & \vdots \\ X_{k1} & X_{k2} & & X_{kn} \end{bmatrix}\begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \cdots & X_{k2} \\ \vdots & & & & \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix},$$

و در نتیجه بعد از ضرب کردن این دو ماتریس داریم

$$\mathbf{X}'\mathbf{X}=\begin{bmatrix} n & \Sigma X_{2i} & \Sigma X_{3i} & \cdots & \Sigma X_{ki} \\ \Sigma X_{2i} & \Sigma X_{2i}^2 & \Sigma X_{2i}X_{3i} & \cdots & \Sigma X_{2i}X_{ki} \\ \Sigma X_{3i} & \Sigma X_{3i}X_{2i} & \Sigma X_{3i}^2 & \cdots & \Sigma X_{3i}X_{ki} \\ \vdots & & & & \\ \Sigma X_{ki} & \Sigma X_{ki}X_{2i} & \Sigma X_{ki}X_{3i} & \cdots & \Sigma X_{ki}^2 \end{bmatrix}. \qquad (۵ـ۲۲)$$

ملاحظه می‌شود که با داشتن مقادیر X_i به راحتی می‌توان کمیتهای لازم برای عناصر مختلف ماتریس $\mathbf{X}'\mathbf{X}$ را محاسبه و سپس این ماتریس را معکوس کرد. یادآوری می‌کنیم که $(\mathbf{X}'\mathbf{X})^{-1}$ یک ماتریس $(k\times k)$ است.

به بررسی بردار $\mathbf{X}'\mathbf{y}$ می‌پردازیم. ابتدا می‌گوییم که

$$\mathbf{X}'\mathbf{y}\rightarrow (k\times n)\,(n\times 1)\rightarrow (k\times 1)\ .$$

برای مطالعهٔ عناصر $\mathbf{X}'\mathbf{y}$، باید ماتریس $\mathbf{X}'$ را در بردار $\mathbf{y}$ ضرب کنیم،

$$\mathbf{X}'\mathbf{y}=\begin{bmatrix} ۱ & ۱ & \cdots & ۱ \\ X_{۲۱} & X_{۲۲} & & X_{۲n} \\ X_{۳۱} & X_{۳۲} & & X_{۳n} \\ \vdots & \vdots & & \vdots \\ X_{k۱} & X_{k۲} & & X_{kn} \end{bmatrix}\begin{bmatrix} Y_۱ \\ Y_۲ \\ Y_۳ \\ \vdots \\ Y_n \end{bmatrix},$$

بعد از ضرب کردن داریم،

$$\mathbf{X}'\mathbf{y}=\begin{bmatrix} \Sigma Y_t \\ \Sigma X_{۲t} Y_t \\ \Sigma X_{۳t} Y_t \\ \vdots \\ \Sigma X_{kt} Y_t \end{bmatrix}. \qquad (۵\text{-}۲۳)$$

با داشتن مقادیر $X_{۲}, X_{۳}$، محاسبهٔ عناصر بردار $\mathbf{X}'\mathbf{y}$ بسیار ساده است. با استفاده از ماتریس $(\mathbf{X}'\mathbf{X})^{-۱}$ و بردار $\mathbf{X}'\mathbf{y}$، محاسبهٔ بردار $\hat{\boldsymbol{\beta}}$، به راحتی با استفاده از فرمول ۲۱ـ۵ انجام خواهد شد.

دقیقاً مانند آنچه در رگرسیون ساده و در معادله‌های ۲۳ـ۱ و ۲٤ـ۱ دیدیم، می‌توان نشان داد که اولاً میانگین جمله‌های پسماند صفر است، و ثانیاً جمله‌های پسماند از متغیرهای توضیحی مستقل است. برای این منظور، ابتدا بردار $\mathbf{y}$، از معادلهٔ ۱۷ـ۵ را به صورت زیر می‌نویسیم:

$$\mathbf{y}=\mathbf{e}+\mathbf{X}\hat{\boldsymbol{\beta}},$$

و سپس آن را در معادلهٔ ۲۰ـ۵ قرار می‌دهیم

$$(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}}=\mathbf{X}'\mathbf{e}+(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}},$$

در نتیجه داریم:

$$\mathbf{X}'\mathbf{e}=\mathbf{o}, \qquad (۵\text{-}۲٤)$$

که در آن $n\times 1 \to 0$ و شامل عناصرصفر است. معادلهٔ ۵ـ۲۴ را به‌صورت زیربسط می‌دهیم.

$$\mathbf{X'e}=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{12} & & X_{1n} \\ X_{21} & X_{22} & & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{k1} & X_{k2} & & X_{kn} \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$=\begin{bmatrix} \Sigma e_i \\ \Sigma X_{2i}e_i \\ \Sigma X_{3i}e_i \\ \\ \Sigma X_{ki}e_i \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

بدین ترتیب ملاحظه‌می‌شود که $\Sigma e_i=0$ یا $\bar{e}=0$ و نیز $\Sigma X_{ji}e_i=0$، $i=1,2,\ldots,k$ است. یادآوری می‌کنیم که نتیجهٔ تخمینهای حداقل مربعات معمولی همواره پسماندهایی است که میانگین آنها صفر است به شرط اینکه مدل رگرسیون دارای جملهٔ ثابت باشد.

قبل از پایان این مبحث دوباره به‌ساختار معادله‌های نرمال در رگرسیون چندمتغیره توجه می‌کنیم. معادلهٔ ۵ـ۲۰ را می‌نویسیم:

$$(\mathbf{X'X})\hat{\boldsymbol{\beta}}=\mathbf{X'y}.$$

مقادیر $(\mathbf{X'X})$ و $\mathbf{X'y}$ را به ترتیب از معادله‌های ۵ـ۲۲ و ۵ـ۲۳ در معادله فوق قرار می‌دهیم. بعد از ضرب کردن داریم:

$$n\hat{\beta}_1+\hat{\beta}_2\Sigma X_{2i}+\hat{\beta}_3\Sigma X_{3i}+\cdots+\hat{\beta}_k\Sigma X_{ki}=\Sigma Y_i,$$

$$\hat{\beta}_1\Sigma X_{2i}+\hat{\beta}_2\Sigma X_{2i}^2+\hat{\beta}_3\Sigma X_{2i}X_{3i}+\cdots+\hat{\beta}_k\Sigma X_{2i}X_{ki}=\Sigma X_{2i}Y_i,$$

$$\hat{\beta}_1 \Sigma X_{1t} + \hat{\beta}_2 \Sigma X_{2t} X_{1t} + \hat{\beta}_3 \Sigma X_{2t}^2 + \cdots + \hat{\beta}_k \Sigma X_{kt} X_{1t} = \Sigma X_{1t} Y_t ,$$

$$\vdots$$

$$\hat{\beta}_1 \Sigma X_{kt} + \hat{\beta}_2 \Sigma X_{kt} X_{1t} + \hat{\beta}_3 \Sigma X_{kt} X_{2t} + \cdots + \hat{\beta}_k \Sigma X_{kt}^2 = \Sigma X_{kt} Y_t .$$

بدیهی است از حل دستگاه **k** معادلهٔ **k** مجهول فوق مقادیر مجهولات $(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k)$ به دست می‌آید. اگر معادلهٔ اول نرمال را در نظر گرفته و دو طرف آن را بر **n** تقسیم کنیم، خواهیم داشت:

$$\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{X}_2 + \hat{\beta}_3 \bar{X}_3 + \cdots + \hat{\beta}_k \bar{X}_k , \qquad (۵\text{-}۲۵)$$

که دقیقاً تعمیم نتیجه‌ای است که در معادلهٔ ۲۶ـ۱ برای رگرسیون ساده به دست آوردیم.

مثال ۵ـ۱ مدل رگرسیون زیر مفروض است:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + U_t .$$

برای تخمین پارامترهای β_1، β_2 و β_3، مشاهدات Y_t، X_{2t} و X_{3t} را درجدول ۵ـ۱ نوشته‌ایم.

جدول ۵ـ۱

Y_t	۳	۱	۸	۳	۵
X_{2t}	۳	۱	۵	۲	۴
X_{3t}	۵	۴	۶	۴	۶

به کمک روش حداقل مربعات معمولی مقادیر $\hat{\beta}_1$، $\hat{\beta}_2$ و $\hat{\beta}_3$ را به دست آورید.

با استفاده از فرمول ۵ـ۲۱ می‌دانیم:

$$\underset{\hat{\beta}}{\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}} = \underset{(X'X)^{-1}}{\begin{bmatrix} n & \Sigma X_{2t} & \Sigma X_{3t} \\ \Sigma X_{2t} & \Sigma X_{2t}^2 & \Sigma X_{2t} X_{3t} \\ \Sigma X_{3t} & \Sigma X_{3t} X_{2t} & \Sigma X_{3t}^2 \end{bmatrix}^{-1}} \underset{X'y}{\begin{bmatrix} \Sigma Y_t \\ \Sigma X_{2t} Y_t \\ \Sigma X_{3t} Y_t \end{bmatrix}} ,$$

بنابراین ابتدا $X'X$ و سپس $X'y$ را محاسبه می‌کنیم،

$$X'X=\underbrace{\begin{bmatrix}۱ & ۱ & ۱ & ۱ & ۱\\ ۳ & ۱ & ۵ & ۲ & ۴\\ ۵ & ۴ & ۶ & ۴ & ۶\end{bmatrix}}_{X'}\underbrace{\begin{bmatrix}۱ & ۳ & ۵\\ ۱ & ۱ & ۴\\ ۱ & ۵ & ۶\\ ۱ & ۲ & ۴\\ ۱ & ۴ & ۶\end{bmatrix}}_{X}=\underbrace{\begin{bmatrix}۵ & ۱۵ & ۲۵\\ ۱۵ & ۵۵ & ۸۱\\ ۲۵ & ۸۱ & ۱۲۹\end{bmatrix}}_{(X'X)},$$

$$X'y=\begin{bmatrix}۱ & ۱ & ۱ & ۱ & ۱\\ ۳ & ۱ & ۵ & ۲ & ۴\\ ۵ & ۴ & ۶ & ۴ & ۶\end{bmatrix}\begin{bmatrix}۳\\ ۱\\ ۸\\ ۳\\ ۵\end{bmatrix}=\begin{bmatrix}۲۰\\ ۷۶\\ ۱۰۹\end{bmatrix}.$$

X' $\qquad$ y $\qquad$ $(X'y)$

مقادیر فوق را در فرمول $\hat{\beta}$ جایگزین می‌کنیم

$$\begin{bmatrix}\hat{\beta}_۱\\ \hat{\beta}_۲\\ \hat{\beta}_۳\end{bmatrix}=\begin{bmatrix}۵ & ۱۵ & ۲۵\\ ۱۵ & ۵۵ & ۸۱\\ ۲۵ & ۸۱ & ۱۲۹\end{bmatrix}\quad\begin{bmatrix}۲۰\\ ۷۶\\ ۱۰۹\end{bmatrix}.$$

ملاحظه می‌شود تخمین مدلهای رگرسیون چندمتغیره مستلزم معکوس کردن ماتریسهایی با ابعاد ۳ و بالاتر است. در پروژه‌های تحقیقاتی و با استفاده از کامپیوتر به راحتی می‌توان هر ماتریسی را معکوس کرد، البته به شرطی که معکوس آن وجود داشته باشد. در حل آن دسته از مسائل اقتصادسنجی که بیشتر جنبهٔ آموزشی دارد، معمول این است که یا معکوس ماتریس را می‌دهند، یا $(X'X)$ به صورتی است که محاسبهٔ معکوس آن بسیار ساده است؛ مثلاً می‌دانیم اگر $(X'X)$ یک ماتریس قطری باشد، معکوس آن برابر عکس عناصر قطری است. البته اگر $(X'X)$ یک ماتریس (۳×۳) باشد، معکوس کردن آن

چندان دشوار نیست و محاسبهٔ آن درواقع یادآوری جبر ماتریسی نیز هست. در مورد این‌گونه ماتریسهای (۳×۳) اگر فرمول معکوس کردن آنها نیز فراموش شود، همواره می‌توان از راه معادله‌های نرمال استفاده کرد و با روش حذفی به راحتی یک دستگاه سه معادلهٔ سه مجهولی را حل کرد. این روشی است که در این قسمت از آن استفاده خواهیم کرد. اگر دو طرف رابطهٔ فوق را در مقدار عددی $(\mathbf{X}'\mathbf{X})$ ضرب کنیم خواهیم داشت

$$\begin{bmatrix} ۵ & ۱۵ & ۲۵ \\ ۱۵ & ۵۵ & ۸۱ \\ ۲۵ & ۸۱ & ۱۲۹ \end{bmatrix} \begin{bmatrix} \hat{\beta}_۱ \\ \hat{\beta}_۲ \\ \hat{\beta}_۳ \end{bmatrix} = \begin{bmatrix} ۲۰ \\ ۷۶ \\ ۱۰۹ \end{bmatrix} ,$$

که در واقع همان معادله‌های نرمال ۲۰-۵ برای این مثال است.

کافی است سه برابر سطر اول ماتریس $(\mathbf{X}'\mathbf{X})$ را از سطر دوم و پنج برابر سطر اول را از سطر سوم کم کنیم. سیستم جدید معادله‌های نرمال به صورت زیر خواهد بود،

$$\begin{bmatrix} ۵ & ۱۵ & ۲۵ \\ ۰ & ۱۰ & ۶ \\ ۰ & ۶ & ۴ \end{bmatrix} \begin{bmatrix} \hat{\beta}_۱ \\ \hat{\beta}_۲ \\ \hat{\beta}_۳ \end{bmatrix} = \begin{bmatrix} ۲۰ \\ ۱۶ \\ ۹ \end{bmatrix} .$$

اکنون باید ۰/۶ سطر دوم را از سطر سوم کم کنیم. در نتیجه

$$\begin{bmatrix} ۵ & ۱۵ & ۲۵ \\ ۰ & ۱۰ & ۶ \\ ۰ & ۰ & ۰/۴ \end{bmatrix} \begin{bmatrix} \hat{\beta}_۱ \\ \hat{\beta}_۲ \\ \hat{\beta}_۳ \end{bmatrix} = \begin{bmatrix} ۲۰ \\ ۱۶ \\ -۰/۶ \end{bmatrix} .$$

معادلهٔ سوم سیستم معادله‌های نرمال فوق عبارت است از

$$۰/۴\ \hat{\beta}_۳ = -۰/۶ ,$$

یا

$$\hat{\beta}_۳ = -۱/۵ .$$

معادلهٔ دوم در سیستم معادله‌های نرمال فوق را می‌نویسیم،

$$۱۰\hat{\beta}_۲ + \hat{\beta}_۳ = ۱۶ .$$

در نتیجه:

$$۱۰\hat{\beta}_۲ + ۶(-۱/۵) = ۱۶ ,$$

$$\hat{\beta}_۲ = ۲/۵ .$$

از معادلهٔ اول نرمال هم داریم:

$$۵\hat{\beta}_۱ + ۱۵\hat{\beta}_۲ + ۲۵\hat{\beta}_۳ = ۲۰ ,$$

$$۵\hat{\beta}_۱ + ۱۵(۲/۵) + ۲۵(-۱/۵) = ۲۰ ,$$

$$\hat{\beta}_۱ = ۴ .$$

بدین ترتیب تخمین مدل رگرسیون مفروض عبارت است از

$$\hat{Y}_i = ۴ + ۲/۵X_{۲i} - X_{۳i} .$$

در پایان این مثال، به یک قرارداد در نمایش محاسبات مدلهای رگرسیون چندمتغیره اشاره می‌کنیم. در مثال فوق دیدیم که برای محاسبهٔ $\hat{\beta}_۱$، $\hat{\beta}_۲$ و $\hat{\beta}_۳$ کمیتهایی که باید محاسبه شود عبارت است از:

الف) محاسباتی که باید برای $(X'X)$ انجام شود

$$\Sigma X_{۲i} = ۱۵ , \quad \Sigma X_{۳i} = ۲۵ , \quad \Sigma X_{۲i}X_{۳i} = ۸۱ ,$$

$$\Sigma X_{۲i}^۲ = ۵۵ , \quad \Sigma X_{۳i}^۲ = ۱۲۹ , \quad \Sigma X_{۳i}X_{۲i} = ۸۱ .$$

یادآوری می‌کنیم که ماتریس $(X'X)$ یک ماتریس قرینه است و مقدار $\Sigma X_{۲i}X_{۳i}$ برابر با $\Sigma X_{۳i}X_{۲i}$ است.

ب) محاسباتی که برای $(X'y)$ لازم است عبارتند از

$$\Sigma Y_i = ۲۰ , \quad \Sigma X_{۲i}Y_i = ۷۶ , \quad \Sigma X_{۳i}Y_{۲i} = ۱۰۹ .$$

معمولاً در مسائل آموزشی، به جای مشاهدات خام متغیرهای Y_t ، X_{tt} و X_{rt} ، مقادیر محاسبه شده فوق را می‌آورند تا به طور مستقیم ماتریسهای $(X'X)$ و $X'y$ محاسبه شود. روش زیر برای نمایش مقادیر مجموع مربعات ΣX_{it}^2 و مجموع حاصلضربهای $\Sigma X_{it}X_{jt}$، به صورت قراردادی پذیرفته شده است.

	Y_t	X_{tt}	X_{rt}
Y_t	□	۷۶	۱۰۹
X_{rt}	۷۶	۵۵	۸۱
X_{rt}	۱۰۹	۸۱	۱۲۹

یادآوری می‌کنیم که در محاسبات این مثال، مقدار ΣY_t^2 محاسبه نشده و مقدار آن در جدول فوق خالی مانده و با علامت □ مشخص شده است. قرینگی ماتریس $X'X$ نیز از جدول فوق به سهولت ملاحظه می‌شود. همچنین یادآوری می‌شود که جدول فوق فقط برای نشان دادن مقادیر مجموع مربعات و مجموع حاصلضربهاست؛ بنابراین ΣX_{it} یا ΣY_t در آن منعکس نیست. روش فوق برای حالتی که k متغیر توضیحی داریم نیز به همین صورت به کار می‌رود.

تخمین مدل رگرسیون خطی چندمتغیره برحسب انحراف از میانگین

در بسیاری موارد، مشاهدات متغیرهای برون‌زا و درون‌زا را برحسب انحراف از میانگین می‌نویسیم تا به این ترتیب اولاً محاسبات $(X'X)^{-1}$ و $X'y$ ساده‌تر انجام شود و ثانیاًهمان‌گونه که بزودی خواهیم دید از ابعاد ماتریس $(X'X)$ یکی کم شده، معکوس‌کردن آن بسیار ساده‌تر خواهد شد؛ بنابراین ضروری است که تخمین مدلهای رگرسیون چندمتغیره را در این حالت نیز به طور خلاصه مطالعه کنیم.

مدل ۱ـ۵ را یک بار دیگر می‌نویسیم،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t .$$

دو طرف را برای تمام مقادیر $t = 1, 2, \ldots, n$، جمع کرده و بر n تقسیم می‌کنیم.

$$\bar{Y}= \beta_1 + \beta_2 \bar{X}_2 +\cdots+ \beta_k \bar{X}_k + \bar{U} .$$

اگر دو طرف این دو معادله را از هم کم کنیم، خواهیم داشت:

$$y_t= \beta_2 x_{2t} + \beta_3 x_{3t} +\cdots+ \beta_k x_{kt} + u_t , \qquad (۵\_۲۶)$$

که در آن، $x_{it}= X_{it} - \bar{X}_i$ و $y_t= Y_t - \bar{Y}$ و $u_t= U_t - \bar{U}$. معادلهٔ ۵ـ۲۶ را برای n دوره و به ترتیبی مشابه با معادلهٔ ۵ـ۲ می‌نویسیم،

$$\begin{aligned}
y_1 &= \beta_2 x_{21} + \beta_3 x_{31} +\cdots+ \beta_k x_{k1} + u_1 ,\\
y_2 &= \beta_2 x_{22} + \beta_3 x_{32} +\cdots+ \beta_k x_{k2} + u_2 ,\\
&\vdots\\
y_n &= \beta_2 x_{2n} + \beta_3 x_{3n} +\cdots+ \beta_k x_{kn} + u_n .
\end{aligned} \qquad (۵_۲۷)$$

سیستم n معاله فوق را می‌توان به صورت زیر نوشت

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} x_{21} & x_{31} & \cdots & x_{k1} \\ x_{22} & x_{32} & \cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{2n} & x_{3n} & \cdots & x_{kn} \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}}_{\mathbf{u}} , \qquad (۵\_۲۸)$$

یا

$$\mathbf{y}= \mathbf{X}\boldsymbol{\beta} + \mathbf{u} , \qquad (۵\_۲۹)$$

که در آن $\mathbf{y} \to (n \times 1)$ و $\mathbf{X} \to n \times (k-1)$ و $\boldsymbol{\beta} \to (k-1) \times 1$ و $\mathbf{u} \to n \times 1$. ملاحظه می‌شود که در این حالت از تعداد ستونهای ماتریس $\mathbf{X}$ و سطرهای بردار $\boldsymbol{\beta}$ یکی کم شده است.

با توجه به‌اینکه بردار u از نظر تعداد سطرها تغییری نمی‌کند واریانس ـ کوواریانس

جمله‌های اختلال، یعنی Σ_uـ که در معادله‌های ۵ـ۹ ، ۵ـ۱۰ و ۵ـ۱۱ تعریف شده است ـ به همان ترتیب برای این حالت نیز تعریف خواهدشد. $e'e$ نیز به‌همان ترتیب معادلهٔ ۵ـ۱۶ است؛ با وجود این، ماتریس $\mathbf{X'X}$ و $\mathbf{X'y}$ تغییراتی خواهد کرد که بررسی خواهد شد.

با استفاده از معادلهٔ ۵ـ۲۸، ماتریس $\mathbf{X'X}$ را می‌نویسیم،

$$\mathbf{X'X}= \underbrace{\begin{bmatrix} x_{21} & x_{22} & \dots & x_{2n} \\ x_{31} & x_{32} & & x_{3n} \\ \vdots & \vdots & & \vdots \\ x_{k1} & x_{k2} & \dots & x_{kn} \end{bmatrix}}_{\mathbf{X'}} \underbrace{\begin{bmatrix} x_{21} & x_{31} & \dots & x_1 \\ x_{22} & x_{31} & \dots & x_2 \\ \vdots & \vdots & & \vdots \\ x_{2n} & x_{3n} & & x_{kn} \end{bmatrix}}_{\mathbf{X}} ,$$

و بعد از ضرب کردن این دو ماتریس، داریم:

$$\mathbf{X'X}= \begin{bmatrix} \Sigma x_{2t}^2 & \Sigma x_{2t}x_{3t} & \dots & \Sigma x_{2t}x_{kt} \\ \Sigma x_{3t}x_{2t} & \Sigma x_{3t}^2 & \dots & \Sigma x_{3t}x_{kt} \\ \vdots & & & \\ \Sigma x_{kt}x_{2t} & \Sigma x_{kt}x_{3t} & \dots & \Sigma x_{kt}^2 \end{bmatrix} , \qquad (۵\text{ـ}۳۰)$$

بنابراین ماتریس $\mathbf{X'X}$ ، در حالتی که مشاهدات برحسب انحراف از میانگین باشد، بسیار ساده است. عناصر قطری آن مجموع مربعات x_{it} و عناصر غیرقطری آن مجموع حاصلضربها، یعنی $\Sigma x_{it}x_{jt}$ ، است. ملاحظه می‌شود که $\mathbf{X'X}$ یک ماتریس مربع و $(k-1)\times(k-1)$ است. در حالی که در حالت قبلی، یعنی وقتی مشاهدات برحسب انحراف از میانگین نبود، این ماتریس $(k\times k)$ است. نتیجه می‌گیریم که اگر مشاهدات را برحسب انحراف از میانگین بنویسیم از ابعاد ماتریس $\mathbf{X'X}$ یکی کم می‌شود و معکوس کردن آن ساده‌ترخواهد بود. البته در عمل، از این امتیاز فقط در حالتی می‌توان استفاده کرد که بخواهیم سه پارامتر یا چهار پارامتر تخمین بزنیم؛ زیرا با این تبدیل، ابعاد ماتریس $\mathbf{X'X}$به جای اینکه (3×3) یا (4×4) بشود، (2×2) یا (3×3) خواهد بود که

معکوس کردن آن بسیار ساده‌تر است. اما برای مدلهای رگرسیون که بیش از چهار پارامتر دارد، باید عملاً از کامپیوتر استفاده شود؛ بنابراین فرق چندانی نمی‌کند که مشاهدات را برحسب انحراف از میانگین بنویسیم.

برای بررسی ماتریس $\mathbf{X}'\mathbf{y}$ به ترتیب زیر عمل می‌کنیم،

$$\mathbf{X}'\mathbf{y}=\underbrace{\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{k1} & x_{k2} & & x_{kn} \end{bmatrix}}_{\mathbf{X}'} \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{y}} = \begin{bmatrix} \Sigma x_{1i}\, y_i \\ \Sigma x_{2i}\, y_i \\ \vdots \\ \Sigma x_{ki}\, y_i \end{bmatrix} \quad . \qquad (۵\_۳۱)$$

ماتریس $\mathbf{X}'\mathbf{y}$ در این حالت، $k \times 1$ و عناصر آن $\Sigma x_{ji}\, y_i$ است.

بردار جمله‌های پسماند نیز همانند معادلهٔ ۵ـ۱۶ تعریف می‌شود

$$\mathbf{e}=\mathbf{y}-\hat{\mathbf{y}},$$

یا

$$\mathbf{e}=\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}, \qquad (۵\_۳۲)$$

که در آن، $\mathbf{e} \to n \times 1$ و $\mathbf{X} \to n \times k$ و $\hat{\boldsymbol{\beta}} \to k \times 1$. مجموع مربعات پسماند نیز دقیقاً همان رابطهٔ ۵ـ۱۹ است، با این تفاوت که از تعداد ستونها و ردیفهای ماتریس $\mathbf{X}$ و بردار $\hat{\boldsymbol{\beta}}$ به ترتیب یکی کم شده است. با روشی کاملاً مشابه با حالت قبلی می‌توان از $\mathbf{e}'\mathbf{e}$ نسبت به $\hat{\boldsymbol{\beta}}$ مشتق گرفت و به معادله‌های نرمال به شرح زیر رسید،

$$(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}}=\mathbf{X}'\mathbf{y}, \qquad (۵\_۳۳)$$

که در آن $\mathbf{X}'\mathbf{X} \to k \times k$ و $\mathbf{X}'\mathbf{y} \to k \times 1$ و $\hat{\boldsymbol{\beta}} \to k \times 1$. معادلهٔ ۵ـ۷۸ بیان دقیقتری از این معادله است. فرمول $\hat{\boldsymbol{\beta}}$ نیز عبارت خواهد بود از

$$\hat{\boldsymbol{\beta}}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \qquad (۵\_۳۴)$$

که در آن $\hat{\boldsymbol{\beta}} \to (k-1)\times 1$، $(\mathbf{X}'\mathbf{X})^{-1} \to (k-1)\times(k-1)$ و $\mathbf{X}'\mathbf{y} \to (k-1)\times 1$. معادلهٔ ۵ـ۷۹ صورت دقیقتری از این فرمول است. بدین ترتیب تخمین مدل رگرسیون ۵ـ۲۹ عبارت است از

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}\,, \qquad (۵\text{ـ}۳۵)$$

مثال ۵ـ۲ مدل رگرسیون زیر مفروض است،

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t\,.$$

با استفاده از مشاهدات مندرج در جدول ۵ـ۲، پارامترهای α، β و γ را با حداقل مربعات معمولی و با روش انحراف از میانگین تخمین بزنید.

جدول ۵ـ۲

Y_t	۳	۱	۸	۳	۵
X_t	۳	۱	۵	۲	٤
Z_t	۵	٤	٦	٤	٦

این مثال دقیقاً همان مثال ۵ـ۱ است، با این تفاوت که اولاً حروف‌گذاری متغیرهای توضیحی و نیز پارامترها را تغییر داده‌ایم و ثانیاً خواسته شده است که ابتدا مشاهدات را برحسب انحراف از میانگین بنویسیم تا کاربرد فرمول ۵ـ۳٤ روشن شود. معمولاً راه‌حل ساده این‌گونه مسائل این است که حروف متغیرها و پارامترها را چنان تغییر دهیم که با فرمولهای متعارف منطبق‌شود. اما در این مثال و به عنوان یک تمرین، با همین حروف جدید کارمی‌کنیم تا مفهوم ماتریس $\mathbf{X}'\mathbf{X}$ و بردارهای $\hat{\boldsymbol{\beta}}$ و $\mathbf{X}'\mathbf{y}$ موجود در فرمول ۵ـ۳٤ بیشتر روشن شود.

ابتدا مدل را برحسب انحراف از میانگین می‌نویسیم:

$$y_t = \beta x_t + \gamma z_t + u_t\,.$$

با استفاده از معادلهٔ ۵ـ۳۰ ماتریس $\mathbf{X}'\mathbf{X}$ را برای این مدل می‌نویسیم،

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} \Sigma x_t^2 & \Sigma x_t z_t \\ \Sigma z_t x_t & \Sigma z_t^2 \end{bmatrix}\,.$$

به کمک معادلهٔ ۳۱ـ۵ می‌توان بردار $\mathbf{X}'\mathbf{y}$ را برای این مدل به صورت زیر نوشت:

$$\mathbf{X}'\mathbf{y}=\begin{bmatrix}\Sigma x_t y_t\\ \Sigma z_t y_t\end{bmatrix},$$

و سرانجام بردار $\hat{\beta}$ برای این مثال، به این ترتیب است:

$$\hat{\beta}=\begin{bmatrix}\hat{\beta}\\ \hat{\gamma}\end{bmatrix}.$$

برای محاسبهٔ کمیتهای لازم برای تشکیل ماتریس $\mathbf{X}'\mathbf{X}$ و بردار $\mathbf{X}'\mathbf{y}$، مشاهدات را برحسب انحراف از میانگین می‌نویسیم. کمیتهای لازم در جدول ۳ـ۵ محاسبه شده است.

جدول ۳ـ۵

Y_t	X_t	z_t	y_t	x_t	z_t	x_t^2	z_t^2	$x_t y_t$	$z_t y_t$	$x_t z_t$
۳	۳	۵	-۱	۰	۰	۰	۰	۰	۰	۰
۱	۱	۴	-۳	-۲	-۱	۴	۱	۶	۳	۲
۸	۵	۶	۴	۲	۱	۴	۱	۸	۴	۲
۳	۲	۴	-۱	-۱	-۱	۱	۱	۱	۱	۱
۵	۴	۶	۱	۱	۱	۱	۱	۱	۱	۱
Σ			۰	۰	۰	۱۰	۴	۱۶	۹	۶

$$\bar{Y}=۴\,،\qquad \bar{X}=۳\,،\qquad \bar{Z}=۵$$

ماتریس $\mathbf{X}'\mathbf{X}$ و $\mathbf{X}'\mathbf{y}$ را تشکیل می‌دهیم،

$$\mathbf{X}'\mathbf{X}=\begin{bmatrix}\Sigma x_t^2 & \Sigma x_t z_t\\ \Sigma z_t x_t & \Sigma z_t^2\end{bmatrix}=\begin{bmatrix}۱۰ & ۶\\ ۶ & ۴\end{bmatrix},$$

$$\mathbf{X}'\mathbf{y}=\begin{bmatrix}\Sigma x_t y_t\\ \Sigma z_t y_t\end{bmatrix}=\begin{bmatrix}۱۶\\ ۹\end{bmatrix}.$$

فرمول ۳٤ـ۵ را برای این مثال می‌نویسیم،

$$\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \Sigma x_t^2 & \Sigma x_t z_t \\ \Sigma z_t x_t & \Sigma z_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma x_t y_t \\ \Sigma x_t z_t \end{bmatrix},$$

$$= \begin{bmatrix} ۱۰ & ٦ \\ ٦ & ٤ \end{bmatrix}^{-1} \begin{bmatrix} ۱٦ \\ ۹ \end{bmatrix} = \begin{bmatrix} ۱ & -\frac{۳}{۲} \\ -\frac{۳}{۲} & \frac{۵}{۲} \end{bmatrix}.$$

و بعد از ضرب کردن، خواهیم داشت:

$$\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} ۲/۵ \\ -۱/۵ \end{bmatrix},$$

یعنی $\hat{\beta}= ۲/۵$ و $\hat{\gamma}= -۱/۵$. برای محاسبهٔ $\hat{\alpha}$ باید از معادلهٔ ۲۵ـ۵ استفاده کنیم. ابتدا این معادله را برای حالت خاص این مثال می‌نویسیم،

$$\bar{Y}= \hat{\alpha} + \hat{\beta}\bar{X} + \hat{\gamma}\bar{Z},$$

و در نتیجه خواهیم داشت:

$$٤= \hat{\alpha} + ۲/۵(۳) - ۱/۵(۵),$$

$$\hat{\alpha}= ٤$$

بدین ترتیب، تخمین مدل مفروض به صورت زیر است:

$$\hat{Y}_t= ٤ + ۲/۵\, X_t - ۱/۵\, Z_t,$$

در پایان این قسمت به یک نکته اشاره می‌کنیم. فرمولهای ۲۱ـ۵ و ۳٤ـ۵ را یک‌بار دیگر ملاحظه کنید. هر دو فرمول به صورت زیر است،

$$\hat{\beta}= (X'X)^{-1} X'y$$

در حالی که فرمول ۳٤ـ۵ برای حالتی است که مشاهدات برحسب انحراف از میانگین

است. در متون اقتصادسنجی کوشش بر این است که تنوع استفاده از حروف یا علامتگذاری به حداقل برسد. با اینکه به راحتی می‌توان از حروف دیگر برای فرمول ۵ـ۳٤ استفاده کرد، اما معمولاً آن را به همین صورت می‌نویسند؛ با وجود این، به این نکته توجه می‌کنند که ابعاد ماتریس $(X'X)$ و بردارهای $X'y$ و $\hat{\beta}$ با فرمول ۵ـ۲۱ متفاوت است و عناصر ماتریس و بردارهای فوق برحسب انحراف از میانگین محاسبه شده است.

۵ـ٤ ضریب تعیین

قبلاً از معادله ۱ـ۲۹ دیدیم که ضریب تعیین در رگرسیون ساده بنا به تعریف برابر است با نسبت تغییرات توضیح داده شده به کل تغییرات Y_t،

$$r^2 = \frac{\Sigma \hat{y}_t^2}{\Sigma y_t^2} \quad \text{(ضریب تعیین)}.$$

در مبحث رگرسیون چندمتغیره و در معادلهٔ ٤ـ۱٦ نیز دیدیم که همین تعریف صادق است. با وجود این، آن را با R^2 نشان می‌دهیم تا از ضریب تعیین در رگرسیون ساده، یعنی r^2، قابل تمیز باشد. در این قسمت ابتدا به بررسی ضریب تعیین به زبان ماتریسی می‌پردازیم و سپس مفهوم «ضریب تعیین تعدیل شده»[1] را بررسی می‌کنیم.

بیان ماتریسی R^2 مستلزم تعریف مجموع تغییرات توضیح داده شده، یعنی $\Sigma \hat{y}_t^2$ و مجموع کل تغییرات Y_t، یعنی Σy_t^2 به زبان ماتریسی است. می‌دانیم

$$\Sigma \hat{y}_t^2 = \Sigma (\hat{Y}_t - \bar{\hat{Y}})^2$$

که با توجه به معادلهٔ ۱ـ۳۰، یعنی $\bar{\hat{Y}} = \bar{Y}$، خواهیم داشت

$$\Sigma \hat{y}_t^2 = \Sigma (\hat{Y}_t - \bar{Y})^2 .$$

اگر $t = 1, 2, \ldots, n$، می‌توان نشان داد که

$$\Sigma \hat{y}_t^2 = \hat{y}' \hat{y}, \qquad (۵ـ۳٦)$$

1. Adjusted Coefficient of Determination

که در آن $\hat{\mathbf{y}} \to n\times 1$. برای اثبات، دو بردار $\hat{\mathbf{y}}'$ و $\hat{\mathbf{y}}$ را در هم ضرب کرده و نشان می‌دهیم که برابر $\sum \hat{y}_t^2$ است. ملاحظه می‌شود که

$$\hat{\mathbf{y}}'\hat{\mathbf{y}} = [\hat{y}_1\ \hat{y}_2\ \cdots\ \hat{y}_n]\begin{bmatrix}\hat{y}_1\\ \hat{y}_2\\ \vdots\\ \hat{y}_n\end{bmatrix} = \hat{y}_1^2+\hat{y}_2^2+\cdots+\hat{y}_n^2 = \sum_1^n \hat{y}_t^2.$$

به همین ترتیب نشان می‌دهیم که مجموع کل تغییرات Y_t برابر است با

$$\sum y_t^2 = \mathbf{y}'\mathbf{y}, \qquad (۵\text{-}۳۷)$$

که در آن، $\mathbf{y} \to n\times 1$. برای اثبات کافی است که در بردار $\mathbf{y}'$ و $\mathbf{y}$ را در هم ضرب کنیم،

$$\mathbf{y}'\mathbf{y} = [y_1\ y_2\ \cdots\ y_n]\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix} = y_1^2+y_2^2+\cdots+y_n^2 = \sum_1^n y_t^2.$$

بدین ترتیب طبق معادلهٔ ۲۹ـ۱ می‌توان ضریب تعیین برای رگرسیون چندمتغیره را به صورت زیر تعریف کرد،

$$R^2 = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}}{\mathbf{y}'\mathbf{y}}. \qquad (۵\text{-}۳۸)$$

معادلهٔ ۳۸ـ۵ از نظر محاسباتی تا حدی مشکل است زیرا مستلزم محاسبهٔ مقادیر $\hat{y}_t$ برای تمام مقادیر t است. برای به دست آوردن یک فرمول ساده‌تر برای R^2، ابتدا معادلهٔ ۳۵ـ۵، یعنی $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ را می‌نویسیم و سپس آن را در ترانهاد خود از سمت راست ضرب می‌کنیم، خواهیم داشت:

$$\hat{\mathbf{y}}'\hat{\mathbf{y}} = (\hat{\boldsymbol{\beta}}'\mathbf{X}')(\mathbf{X}\hat{\boldsymbol{\beta}}).$$

مقدار $\hat{\boldsymbol{\beta}}$ از معادلهٔ ۳۴ـ۵ را در معادلهٔ فوق قرار می‌دهیم،

$$\hat{y}'\hat{y}=\hat{\beta}'X'X\,[(X'X)^{-1}X'y]\,.$$

مقدار $(X'X)^{-1}$ داخل کروشه با مقدار $(X'X)$ خارج از کروشه حذف شده است و در نتیجه

$$\hat{y}'\hat{y}=\hat{\beta}'X'y\,. \qquad (۵\text{-}۳۹)$$

معادلهٔ ۳۹ـ۵ را در معادلهٔ ۳۸ـ۵ جایگزین می‌کنیم. داریم

$$R^2=\frac{\hat{\beta}'X'y}{y'y} \qquad (۵\text{-}۴۰)$$

معادلهٔ فوق، مشابه معادلهٔ ۳۵ـ۱ برای رگرسیون ساده و بیان ماتریسی معادله ۱۹ـ٤ است. از این معادله می‌توان در حل مسائل استفاده فراوانی کرد؛ زیرا وقتی یک مدل رگرسیون را تخمین زده‌ایم، مقادیر $\hat{\beta}$ و $X'y$ را داریم، بنابراین $y'y$ را که همان Σy_i^2 است محاسبه می‌کنیم تا R^2 به سهولت و از طریق معادلهٔ ۴۰ـ۵ به دست آید.

محاسبهٔ R^2 با فرمول فوق مستلزم داشتن مقادیر تخمین پارامترها است. می‌توان معادلهٔ ۴۰ـ۵ را تغییر داد به نحوی که R^2 فقط برحسب مشاهدات متغیرهای درون‌زا و برون‌زا نوشته شود. برای این منظور، باید مقدار $\hat{\beta}$ را در معادلهٔ ۴۰ـ۵ قرار دهیم. ابتدا معادلهٔ ۳٤ـ۵ را یک بار دیگر می‌نویسیم:

$$\hat{\beta}=(X'X)^{-1}X'y\,.$$

معادله فوق را ترانهاد می‌کنیم،

$$\hat{\beta}'=y'X(X'X)^{-1}\,. \qquad (۵\text{-}۴۱)$$

یادآوری می‌شود که چون $(X'X)$، متقارن است، با ترانهاد خود برابر است؛ یعنی

$$(X'X)'=(X'X)\,.$$

معادلهٔ ۴۱ـ۵ را در ۴۰ـ۵ قرار می‌دهیم،

$$R^2=\frac{y'X(X'X)^{-1}X'y}{y'y}\,. \qquad (۵\text{-}۴۲)$$

معادلهٔ ۴۲ـ۵ فرمول مستقیم محاسبه R^2 است.

در مدل رگرسیون ساده، قضیهٔ مهمی را در قالب معادلهٔ ۴۰ـ۱ ثابت کردیم. اثبات این قضیه به زبان ماتریسی به قرار زیر است. می‌خواهیم ثابت کنیم که در یک مدل رگرسیون چندمتغیره، کل تغییرات متغیر درون‌زا برابراست با مجموع تغییرات توضیح داده شده، بعلاوهٔ تغییرات توضیح داده نشده؛ به عبارت‌دیگر، می‌دانیم در رگرسیون ساده داریم

$$\Sigma y_i^2 = \Sigma \hat{y}_i^2 + \Sigma e_i^2 .$$

باید ابتدا مشابه معادله فوق را برای مدلهای رگرسیون چندمتغیره بنویسیم. با استفاده از معادله‌های ۳۶ـ۵ و ۳۷ـ۵ می‌دانیم که

$$\Sigma y_i^2 = \mathbf{y}'\mathbf{y} \qquad \text{و} \qquad \Sigma \hat{y}_i^2 = \hat{\mathbf{y}}'\hat{\mathbf{y}}$$

همچنین می‌توان نشان داد که در یک مدل رگرسیون که متغیرها برحسب انحراف از میانگین اندازه‌گیری شده است، یعنی مدل ۲۶ـ۵، تغییرات توضیح داده نشده، یعنی مجموع مربعات پسماند، برابر است با

$$\Sigma e_i^2 = \mathbf{e}'\mathbf{e} . \qquad (۵\text{ـ}۴۳)$$

برای اثبات کافی است به این نکته توجه کنیم که $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$؛ پس نبودن جملهٔ ثابت، یعنی β_1، در مدل رگرسیون ۲۶ـ۵ در ابعاد بردار $\mathbf{y}$ و $\hat{\mathbf{y}}$ تأثیری ندارد، بلکه فقط مقدار $\hat{\mathbf{y}}$ را متأثر می‌کند؛ زیرا $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$. بنابراین:

$$\mathbf{e}'\mathbf{e} = [e_1 \; e_2 \cdots e_n] \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = e_1^2 + e_2^2 + \cdots + e_n^2 = \Sigma e_i^2 . \qquad (۵\text{ـ}۴۴)$$

در نتیجه بیان ماتریسی قضیه‌ای که می‌خواهیم ثابت کنیم به شرح زیر خواهد بود،

$$\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \mathbf{e}'\mathbf{e} \qquad (۵\text{ـ}۴۵)$$

یادآوری می‌شود که هریک از جمله‌های معادلهٔ ۴۵ـ۵ یک اسکالر است.

برای اثبات، ابتدا تغییرات توضیح داده نشده، یعنی مجموع مربعات پسماند را می‌نویسیم،

$$e'e = (y-\hat{y})' (y-\hat{y}) .$$

با جایگزینی معادلهٔ ۵ـ۳۵، یعنی $\hat{y} = X\hat{\beta}$، در معادله فوق داریم

$$e'e = (y-X\hat{\beta})' (y-X\hat{\beta}) ,$$

$$= (y' - \hat{\beta}'X') (y-X\hat{\beta}) .$$

مقدار $\hat{\beta}$ از معادلهٔ ۵ـ۳٤ را در این معادله قرار می‌دهیم:

$$e'e = [y' - y'X(X'X)^{-1}X'] [y - X(X'X)^{-1}X'y] .$$

و بعد از ضرب کردن داریم:

$$e'e = y'y - y'X(X'X)^{-1}X'y - y'X(X'X)^{-1}X'y + y'X(X'X)^{-1}X'y .$$

معادله فوق را ساده می‌کنیم،

$$e'e = y'y - y'X(X'X)^{-1}X'y . \qquad (۵ـ٤٦)$$

حال نشان می‌دهیم که جملهٔ آخر با $\hat{y}'\hat{y}$ برابر است. برای این منظور با استفاده از معادلهٔ ۵ـ۳۵، یعنی $\hat{y} = X\hat{\beta}$، چنین می‌نویسیم،

$$\hat{y}'\hat{y} = \hat{\beta}'X'X\hat{\beta} = [y'X(X'X)^{-1}] (X'X) [(X'X)^{-1}X'y] ,$$

در نتیجه

$$\hat{y}'\hat{y} = y'X(X'X)^{-1}X'y . \qquad (۵ـ٤٧)$$

البته قبلاً نیز به نتیجهٔ فوق رسیده بودیم؛ زیرا معادلهٔ فوق درواقع صورت کسر ۵ـ٤۲ است. به هرحال با جایگزینی معادلهٔ ۵ـ٤۷ در ۵ـ٤٦ خواهیم داشت

$$y'y = \hat{y}'\hat{y} + e'e ,$$

که اثبات معادلهٔ ۵ـ۴۵ است.

با استفاده از معادله فوق می‌توان سه نکتهٔ دیگر را نیز بررسی کرد. اولاً، قلمرو تغییرات R^2، ثانیاً، فرمول دیگری برای R^2، و ثالثاً به دست آوردن مجموع مربعات پسماند $(e'e)$. برای به دست آوردن قلمرو تغییرات R^2 باید دو طرف معادلهٔ ۵ـ۴۵ را بر کل تغییرات متغیر درون‌زا، یعنی $y'y$ تقسیم کنیم:

$$\frac{\hat{y}'\hat{y}}{y'y} = \frac{y'y}{y'y} - \frac{e'e}{y'y}.$$

باتوجه به تعریف R^2 در رابطهٔ ۵ـ۳۸ معادله فوق را به صورت زیر می‌نویسیم؛

$$R^2 = 1 - \frac{e'e}{y'y}. \qquad (۵ـ۴۸)$$

این معادله، نه تنها فرمول دیگری برای R^2 محسوب می‌شود، بلکه می‌تواند قلمرو تغییرات R^2 را نیز تعیین کند. بهترین حالت موقعی است که پسماندی نداشته باشیم؛ یعنی تخمین ما توانسته باشد تمام تغییرات y_i در مدل ۵ـ۲۶ را توضیح دهد. در چنین حالتی مجموع مربعات پسماند برابر صفر می‌شود و $e'e = 0$، در نتیجه $R^2 = 1$. بدترین حالت موقعی است که تخمین ما هیچ قدرت توضیحی نداشته باشد، یعنی $\hat{y}'\hat{y} = 0$، و در نتیجه $e'e = y'y$ باشد، در این صورت $R^2 = 0$ است؛ بنابراین

$$0 \leqslant R^2 \leqslant 1. \qquad (۵ـ۴۹)$$

استفاده دیگری که معمولاً از معادلهٔ ۵ـ۴۸ می‌شود، این است که با داشتن R^2، مثلاً از معادلهٔ ۵ـ۴۲ و نیز با محاسبهٔ $y'y$، می‌توان $e'e$ را محاسبه کرد:

$$e'e = (1 - R^2)y'y. \qquad (۵ـ۵۰)$$

همان گونه که خواهیم دید، محاسبهٔ مجموع مربعات پسماند، یعنی $e'e$، شرط لازم برای تخمین واریانس جملهٔ اختلال است.

قبلاً در فصل اول و در مسألهٔ ۱ـ۱۹ دیدیم که می‌توان r^2 را از فرمول زیر به دست آورد،

$$r^2 = \frac{\Sigma(y_i \hat{y}_i)^2}{\Sigma y_i^2 \Sigma \hat{y}_i^2},$$

یعنی r^2 با مجذور ضریب همبستگی بین Y_i و $\hat{Y}_i$ برابر است. در اینجا می‌خواهیم ثابت کنیم که برای یک رگرسیون چندمتغیره نیز این خصوصیت صادق است؛ یعنی R^2 برابر است با مجذور ضریب همبستگی بین $\mathbf{y}$ و $\hat{\mathbf{y}}$.

برای اثبات، از تعریف ضریب همبستگی بین $\mathbf{y}$ و $\hat{\mathbf{y}}$ شروع کرده، نشان خواهیم داد که این ضریب برابر R^2 است. می‌دانیم

$$\text{ضریب همبستگی بین } \mathbf{y}, \hat{\mathbf{y}} = \frac{\mathbf{y}'\hat{\mathbf{y}}}{\sqrt{\mathbf{y}'\mathbf{y}}\sqrt{\hat{\mathbf{y}}'\hat{\mathbf{y}}}}$$

دو طرف را مجذور می‌کنیم،

$$\text{مجذور ضریب همبستگی بین } \mathbf{y}, \hat{\mathbf{y}} = \frac{(\mathbf{y}'\hat{\mathbf{y}})^2}{(\mathbf{y}'\mathbf{y})(\hat{\mathbf{y}}'\hat{\mathbf{y}})}. \qquad (۵\text{-}۵۱)$$

با توجه به $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e}$، می‌توان صورت کسر را به صورت زیر نوشت

$$\mathbf{y}'\hat{\mathbf{y}} = (\hat{\mathbf{y}}' + \mathbf{e}')\hat{\mathbf{y}} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \mathbf{e}'\hat{\mathbf{y}}.$$

با توجه به $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ داریم

$$\mathbf{y}'\hat{\mathbf{y}} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \mathbf{e}'\mathbf{X}\hat{\boldsymbol{\beta}}.$$

با توجه به معادلهٔ ۵ـ۲۴ می‌دانیم $\mathbf{e}'\mathbf{X} = \mathbf{0}$، بنابراین با جایگزینی در معادلهٔ فوق خواهیم داشت

$$\mathbf{y}'\hat{\mathbf{y}} = \hat{\mathbf{y}}'\hat{\mathbf{y}}. \qquad (۵\text{-}۵۲)$$

معادلهٔ ۵ـ۵۲ را در معادلهٔ ۵ـ۵۱ قرار می‌دهیم،

$$\text{مجذور ضریب همبستگی بین } \mathbf{y}, \hat{\mathbf{y}} = \frac{(\hat{\mathbf{y}}'\hat{\mathbf{y}})^2}{(\mathbf{y}'\mathbf{y})(\hat{\mathbf{y}}'\hat{\mathbf{y}})} = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}}{\mathbf{y}'\mathbf{y}}.$$

اما با توجه به معادلهٔ ۵ـ۳۸ می‌دانیم که

$$\frac{\hat{y}'\hat{y}}{y'y} = R^2 ,$$

بنابراین خواهیم داشت

$$R^2 = \text{مجذور ضریب همبستگی بین } y \text{ و } \hat{y} . \qquad (۵ـ۵۳)$$

R^2 در مدلهای رگرسیون بدون جملهٔ ثابت

در فصل اول و مسألهٔ ۱ـ۲۱ دیدیم که اگر بخواهیم یک رگرسیون ساده و بدون جملهٔ ثابت، یعنی

$$Y_t = \beta X_t + U_t$$

را با فرمول معمولی ضریب تعیین، یعنی

$$r^2 = 1 - \frac{\Sigma e_t^2}{\Sigma y_t^2} ,$$

تخمین بزنیم، ممکن است r^2 منفی شود. در این قسمت می‌خواهیم این نکته را برای حالت عمومی مدلهای رگرسیون که فاقد جملهٔ ثابت هستند ثابت کنیم. برای این منظور کافی است نشان دهیم که قضیهٔ ۵ـ۴۵ برای مدلهای فاقد جملهٔ ثابت نیز صادق است.

مدل رگرسیون ۵ـ۱ را بدون جملهٔ ثابت β_1 می‌نویسیم،

$$Y_t = \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + \beta_k X_{kt} + U_t . \qquad (۵ـ۵۴)$$

اگر معادلهٔ ۵ـ۵۴ را دقیقاً مانند سیستم معادله‌های ۵ـ۳ نوشته و سپس آن را به صورت ماتریسی نمایش دهیم، خواهیم داشت:

$$y = X\beta + u , \qquad (۵ـ۵۵)$$

که در آن $y \to n\times 1$ و $X \to n\times(k-1)$ و $\beta \to (k-1)\times 1$ و $u \to n\times 1$. توجه داریم که

تفاوت معادلهٔ ۵۵ـ۵ با معادلهٔ ٤ـ۵ این است که در معادلهٔ ۵۵ـ۵ اولین ستون ماتریس $\mathbf{X}$ ـ که در معادلهٔ ٤ـ۵ شامل عناصر یک است ـ حذف شده است و بردار $\boldsymbol{\beta}$ نیز فاقد اولین عنصر، یعنی β_1، است. دقیقاً مشابه معادلهٔ ۱۹ـ۵ مقدار $\mathbf{e}'\mathbf{e}$ را نوشته و از آن نسبت به $\hat{\boldsymbol{\beta}}$ مشتق می‌گیریم. در نتیجه

$$\frac{\partial \mathbf{e}'\mathbf{e}}{\partial \hat{\boldsymbol{\beta}}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} .$$

مشتق را مساوی صفر قرار داده و برای $\hat{\boldsymbol{\beta}}$ حل می‌کنیم،

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} . \qquad (۵\text{ـ}۵۶)$$

فرمول ۵۶ـ۵ که برای مدلهای رگرسیون فاقد جملهٔ ثابت است دقیقاً مشابه فرمول ۲۱ـ۵ است، با این تفاوت که $(\mathbf{X}'\mathbf{X})$ یک ماتریس $(k-1)\times(k-1)$ و $\mathbf{X}'\mathbf{y}$ نیز یک بردار $(k-1)\times 1$ است. عناصر این ماتریس و بردار برحسب مقادیر اصلی است نه انحراف ازمیانگین. می‌دانیم $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e}$ و با مراجعه به $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$، داریم

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} .$$

از ضرب معادله فوق در ترانهاد خود، خواهیم داشت

$$\mathbf{y}'\mathbf{y} = (\hat{\boldsymbol{\beta}}'\mathbf{X}' + \mathbf{e}')(\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}) ,$$

$$= \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{e} + \mathbf{e}'\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}'\mathbf{e} .$$

دقیقاً مشابه معادلهٔ ۲٤ـ۵، می‌توان نشان داد که از معادله‌های نرمال برای مدل ٥٤ـ۵ نتایج زیر به دست می‌آید،

$$\mathbf{X}'\mathbf{e} = \mathbf{o} \qquad \text{یا} \qquad \mathbf{e}'\mathbf{X} = \mathbf{o} ,$$

در نتیجه با توجه به $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ داریم

$$\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \mathbf{e}'\mathbf{e} .$$

ملاحظه می‌شود قضیهٔ ٤٥ـ۵ در مورد مدلهای رگرسیون فاقد جملهٔ ثابت نیز صدق

می‌کند؛ بنابراین معادلهٔ فوق دقیقاً مشابه این قضیه برای حالت عمومی است. با این تفاوت که اولاً، با یک روش تقریباً متفاوتی به دست آمده، ثانیاً تمام عناصر بردارهای آن برحسب مقادیر اصلی است نه انحراف از میانگین. با توجه به اینکه در ادامه بحث، باید مرتباً به عناصر ماتریسها توجه کنیم که آیا برحسب انحراف از میانگین است یا مشاهدات اصلی؛ بنابراین تا جایی که امکان دارد، عبارتها را به زبان غیرماتریسی می‌نویسیم. ابتدا معادلهٔ فوق را به صورت غیرماتریسی می‌نویسیم،

$$\Sigma Y_t^2 = \Sigma \hat{Y}_t^2 + \Sigma e_t^2. \qquad (۵\text{-}۵۷)$$

اگر R^2 را به صورت زیر تعریف کنیم،

$$R^2 = 1 - \frac{\Sigma e_t^2}{\Sigma Y_t^2} = \frac{\Sigma \hat{Y}_t^2}{\Sigma Y_t^2},$$

آنگاه با توجه به معادلهٔ ۵ـ۵۷ خواهیم داشت $0 < R^2 < 1$. اما اگر بخواهیم R^2 را به روال معمول، یعنی مطابق معادلهٔ ۵ـ۴۸ تعریف کنیم:

$$R^2 = 1 - \frac{\Sigma e_t^2}{\Sigma y_t^2},$$

آنگاه مقدار R^2 می‌تواند منفی شود. با توجه به اینکه مقدار Σe_t^2 برای مشاهدات اصلی و انحراف از میانگین یکی است، ابتدا R^2 را به صورت زیر می‌نویسیم،

$$R^2 = 1 - \frac{\Sigma Y_t^2 - \Sigma \hat{Y}_t^2}{\Sigma y_t^2}.$$

می‌دانیم

$$\Sigma y_t^2 = \Sigma (Y_t - \bar{Y})^2 = \Sigma Y_t^2 - n\bar{Y}^2$$

بنابراین

$$R^2 = \frac{\Sigma \hat{Y}_t^2 - n\bar{Y}^2}{\Sigma Y_t^2 - n\bar{Y}^2}. \qquad (۵\text{-}۵۸)$$

مخرج کسر همواره مثبت است، زیرا برابر با Σy_i^2 است. اما در مواردی که ΣY_i^2 از $n\bar{Y}^2$ کمتر بشود، مقدار R^2 منفی خواهد بود. یادآوری می‌کنیم که:

$$\Sigma \hat{Y}_i^2 - n\bar{Y}^2 \neq [\Sigma \hat{y}_i^2 = \Sigma (\hat{Y}_i - \bar{\hat{Y}})^2],$$

زیرا در مدلهایی که فاقد جملهٔ ثابت هستند اولاً، $\bar{\hat{Y}} \neq \bar{Y}$ و ثانیاً، $\bar{e} \neq 0$ است. درواقع این دو شرط، همان‌گونه که در معادله‌های ۱ـ۳۰ و ۱ـ۳۸ دیدیم، حاصل معادلهٔ اول نرمال است که خود از مشتق‌گیری Σe_i^2 نسبت به جملهٔ ثابت به دست آمده است. بدیهی است در مدلهای فاقد جملهٔ ثابت، این شرایط نیز محقق نخواهد بود. در پایان، به این نکته نیز توجه داریم که بحث از R^2 در مدلهایی است که جملهٔ ثابت ندارد. نکته‌های مربوط به بیان ماتریسی R^2، برحسب مشاهدات اصلی در قسمت ۵ـ۷ مطرح خواهد شد.

۵ـ۵ ضریب تعیین تعدیل شده یا $\bar{R}^2$

می‌دانیم با افزایش متغیرهای توضیحی در یک مدل رگرسیون، مقدار R^2 معمولاً زیاد شده، یا حداقل کاهش نمی‌یابد؛ برای مثال، یک مدل رگرسیون شامل یک متغیر توضیحی X_{i2} را در نظر گرفته و مقادیر Σe_i^2 و Σy_i^2 آن را حساب می‌کنیم. به این مدل یک متغیر توضیحی دیگر، مانند X_{i3} اضافه می‌کنیم. اقتضای روش حداقل مربعات معمولی این است که مجموع مربعات پسماند را حداقل کند. اگر مجموع مربعات پسماند را برای مدل جدید Σe_i^{*2} بنامیم، می‌توان چنین نوشت

$$RSS^* = \Sigma e_i^{*2} \leq \Sigma e_i^2 \qquad (۵ـ۵۹)$$

زیرا مکانیسم حداقل‌سازی مجموع مربعات پسماند در مدل دوم به گونه‌ای است که اگر حذف متغیر جدید، یعنی X_{i3} در به حداقل رساندن Σe_i^{*2} کمکی باشد، قطعاً در مشتق‌گیری، ضریب متغیر X_{i3}، صفر خواهد شد. دلیل این امر شمولیت مدل دوم نسبت به مدل اول است؛ یعنی مدل دوم شامل مدل اول نیز بوده، بنابراین

مکانیسم حداقل‌سازی Σe_i^2 در مدل اول را نیز به طور ضمنی دربردارد؛ بنابراین نتیجه می‌گیریم که به موازات افزایش متغیرهای توضیحی، مقدار Σe_i^2 هیچگاه افزایش نخواهد یافت.

با توجه به معادلهٔ ۱۶ـ۴، فرمول ضریب تعیین برای مدل جدید را به‌صورت زیر می‌نویسیم،

$$R^2 = 1 - \frac{\Sigma e_i^{*2}}{\Sigma y_i^2}.$$

ضریب تعیین برای مدل رگرسیون اول که فقط شامل یک متغیر توضیحی X_{11} است از معادلهٔ ۴۳ـ۱ به دست می‌آید،

$$r^2 = 1 - \frac{\Sigma e_i^2}{\Sigma y_i^2}.$$

از طرف دیگر می‌دانیم که مقدار کل تغییرات (Σy_i^2)، برای هر دو مدل یکسان است؛ بنابراین با توجه به رابطهٔ ۵۹ـ۵ خواهیم داشت:

$$R^2 \geq r^2.$$

به همین ترتیب می‌توان نشان داد که اگر به مدل دوم، متغیرهای توضیحی جدید اضافه کنیم، مقادیر ضریب تعیین آنها مرتباً زیاد شده، یا حداقل کاهش نمی‌یابد.

از نظر محاسباتی، می‌توان مسأله را از زاویهٔ دیگری نیز حداقل ملاحظه کرد. یک متغیر توضیحی جدید، فقط صورت کسر معادلهٔ ۱۶ـ۴ را تغییر می‌دهد و مخرج آن را مطلقاً متأثر نمی‌کند. علت این امر، تعریف خاصی است که از ضریب تعیین شده است. در واقع این ضریب را به صورت نسبت دو مجموعه از تغییرات معرفی کرده‌ایم، به گونه‌ای که تغییرات منعکس شده در مخرج، تابعی از تعداد متغیرهای توضیحی نیست. حال اگر به جای تغییر، از مفهوم واریانس استفاده کنیم، دیگر مشکل فوق را نداریم و یک متغیر توضیحی جدید، می‌تواند صورت و مخرج کسر R^2 را همزمان تغییر دهد.

از تعریف ضریب تعیین ۱۶ـ۴ شروع می‌کنیم،

$$R^2 = \frac{ESS}{TSS} = \frac{\Sigma \hat{y}_t^2}{\Sigma y_t^2}.$$

صورت و مخرج آن را بر n تقسیم کرده، خواهیم داشت:

$$R^2 = \frac{\Sigma \hat{y}_t^2/n}{\Sigma y_t^2/n} = 1 - \frac{\Sigma e_t^2/n}{\Sigma y_t^2/n}. \qquad (۵\text{-}۶۰)$$

در معادلهٔ ۲۸ـ۲ ثابت کردیم که اگر بخواهیم واریانس پسماندها یک تخمین نااریب از جملهٔ اختلال باشد، باید Σe_t^2 را بر درجات آزادی آن تقسیم کنیم. مدل جدید را که شامل $X_{۲t}$ است می‌نویسیم،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + U_t.$$

می‌دانیم برای محاسبهٔ Σe_t^2 در این مدل، سه درجه آزادی از دست می‌دهیم؛ زیرا محاسبهٔ Σe_t^2، مستلزم تخمینهای β_1، β_2 و β_3 است؛ بنابراین درجات آزادی Σe_t^2 در معادلهٔ ۶۰ـ۵ و برای این مدل برابر (n - ۳) خواهد بود.

البته بیان دقیقتر مفهوم درجات آزادی برای Σe_t^2 در مدل فوق‌الذکر به این صورت است که اگر بخواهیم سه پارامتر β_1، β_2 و β_3 را تخمین بزنیم، آنگاه با توجه به معادله‌های نرمال، یعنی معادله‌های ۱۲ـ۴، ۱۳ـ۴ و ۱۴ـ۴، خواهیم داشت

$$\Sigma e_t = 0 \;،\; \Sigma e_t X_{1t} = 0 \;،\; \Sigma e_t X_{2t} = 0$$

بنابراین مقادیر پسماندها به گونه‌ای تعیین می‌شود که سه شرط فوق در آنها صدق کند. درواقع فقط (n - ۳) مقدار e_t می‌تواند به «آزادی» مقدار بگیرد، بنابراین می‌گوییم درجات آزادی Σe_t^2 برابر (n - ۳) است؛ به عبارت دیگر، به ازای هر مجموعه از مقادیری که (n - ۳) پسماند اختیار می‌کند، مقادیر سه پسماند باقیمانده، فقط با حل یک دستگاه معادله، شامل سه شرط فوق امکانپذیر است؛ یعنی درواقع برای محاسبهٔ Σe_t^2، سه درجهٔ آزادی از دست داده‌ایم. نتیجه می‌گیریم که باید Σe_t^2 در معادلهٔ ۶۰ـ۵ را به جای n بر (n - ۳) تقسیم کنیم. همچنین برای اینکه بتوان به تخمین نااریبی از واریانس

متغیر درون‌زا رسید، باید Σy_i^2 در معادلهٔ ۵ـ۶۰ را بر $(n-1)$ تقسیم کنیم؛ زیرا محاسبهٔ Σy_i^2 مستلزم محاسبهٔ $\bar{Y}$ است؛ بنابراین یک درجهٔ آزادی از دست خواهم داد. اصطلاحاً می‌گوییم R^2 را نسبت به درجات آزادی تعدیل کرده‌ایم. به ضریب تعیینی که بدین ترتیب به دست می‌آید، «ضریب تعیین تعدیل شده» گفته و آن را با $\bar{R}^2$ نشان می‌دهند.

$$\bar{R}^2 = 1 - \frac{\Sigma e_i^2/(n-3)}{\Sigma y_i^2/(n-1)},$$

یا

$$\bar{R}^2 = 1 - \left(\frac{n-1}{n-3}\right)\frac{\Sigma e_i^2}{\Sigma y_i^2}. \qquad (۵\text{ـ}۶۱)$$

با استفاده از تعریف ضریب تعیین، یعنی معادلهٔ ۱۶ـ۴، می‌دانیم که:

$$\frac{\Sigma e_i^2}{\Sigma y_i^2} = 1 - R^2.$$

با جایگزینی معادلهٔ فوق در معادلهٔ ۶۱ـ۵، خواهیم داشت

$$\bar{R}^2 = 1 - \left(\frac{n-1}{n-3}\right)(1-R^2), \qquad (۵\text{ـ}۶۲)$$

که در آن R^2، ضریب تعیین مدلی است که دارای دو متغیر توضیحی است. معادلهٔ ۶۲ـ۵ رابطهٔ بین ضریب تعیین معمولی و ضریب تعیین تعدیل شده را نشان می‌دهد.

بحث فوق را برای یک مدل رگرسیون با k متغیر توضیحی تعمیم می‌دهیم. مدل ۱ـ۵ را یک بار دیگر می‌نویسیم،

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + U_i.$$

با استفاده از تعریف ضریب تعیین در رابطهٔ ۱۶ـ۴ می‌توان نوشت،

$$R^2 = \frac{ESS}{TSS} = \frac{\Sigma \hat{y}_i^2}{\Sigma y_i^2} = 1 - \frac{\Sigma e_i^2}{\Sigma y_i^2}.$$

می‌دانیم برای محاسبهٔ مجموع مربعات پسماند، یعنی Σe_i^2، در مدل فوق باید $\hat{\beta}_1$، $\hat{\beta}_2$ تا $\hat{\beta}_k$ را داشته باشیم؛ بنابراین k درجه آزادی از دست داده، در نتیجه درجات آزادی برای Σe_i^2 برابر $(n-k)$ است؛ با وجود این، درجات آزادی برای Σy_i^2 همان $(n-1)$ خواهد بود. با تقسیم Σe_i^2 و Σy_i^2 بر درجات آزادی آنها به $\bar{R}^2$ می‌رسیم.

$$\bar{R}^2 = 1 - \frac{\Sigma e_i^2/(n-k)}{\Sigma y_i^2/(n-1)},$$

یا

$$\bar{R}^2 = 1 - \left(\frac{n-1}{n-k}\right)\frac{\Sigma e_i^2}{\Sigma y_i^2}, \qquad (۵\_۶۳)$$

$$= 1 - \left(\frac{n-1}{n-k}\right)\frac{e'e}{y'y}.$$

با توجه به معادلهٔ ۴ـ۱۶ داریم:

$$\frac{\Sigma e_i^2}{\Sigma y_i^2} = \frac{e'e}{y'y} = 1 - R^2,$$

با جایگزینی در معادلهٔ ۵ـ۶۳، خواهیم داشت:

$$\bar{R}^2 = 1 - \left(\frac{n-1}{n-k}\right)(1-R^2), \qquad (۵\_۶۴)$$

که در آن n، تعداد مشاهدات و k، تعداد پارامترهایی است که در مدل رگرسیون باید تخمین زد. یادآوری می‌شود که در این مدل $(k-1)$ متغیر توضیحی داریم. با ملاحظهٔ معادلهٔ ۵ـ۶۴ به سهولت می‌توان گفت که

$$\bar{R}^2 \leq R^2. \qquad (۵\_۶۵)$$

$\bar{R}^2$، درواقع آماره‌ای است که خودمان ساخته‌ایم و همبستگی هیچیک از دو متغیر را نشان نمی‌دهد. به همین دلیل، برای مقادیر کوچک R^2، می‌توان نشان داد که $\bar{R}^2$ ممکن است منفی شود. برای این منظور، ابتدا معادلهٔ ۵ـ۶۴ را به صورت زیر می‌نویسیم،

$$1 - \bar{R}^2 = \left(\frac{n-1}{n-k}\right)(1-R^2). \qquad (۵\_۶۶)$$

به ازای $R^2 < \frac{k-1}{n-1}$، خواهیم داشت:

$$(1-R^2) > 1-\frac{k-1}{n-1},$$

یا

$$(1-R^2) > \frac{n-k}{n-1}.$$

با جایگزینی رابطهٔ فوق در معادلهٔ ۵ـ۶۶، داریم:

$$1-\bar{R}^2 > 1,$$

یا

$$\bar{R}^2 < 0.$$

(۵ـ۶۷)

برای مثال، فرض کنید، می‌خواهیم مدلی را تخمین بزنیم که ۴ پارامتر دارد. با ۲۱ مشاهده روی متغیرها داریم: $k=4$ و $n=21$. اگر R^2 برابر ۰/۱ باشد، چون

$$R^2 < \frac{k-1}{n-1},$$

$$0/1 < \frac{4-1}{21-1},$$

بنابراین $\bar{R}^2$ منفی خواهد بود. برای ارزیابی صحت این نتیجه می‌توان مستقیماً از معادلهٔ ۵ـ۶۴ نیز استفاده کرد

$$\bar{R}^2 = 1-\left(\frac{21-1}{21-4}\right)(1-0/1),$$

$$= 1-\frac{20}{17}(0/9) = 1-\frac{18}{17} = -0/058.$$

تأثیر متغیرهای توضیحی در $\bar{R}^2$ و $\hat{\sigma}_u^2$

می‌توان ثابت کرد که آن دسته از متغیرهای توضیحی که واریانس جملهٔ اختلال، یعنی $\hat{\sigma}_u^2$ را حداقل کند، ضریب تعیین تعدیل شده، یعنی $\bar{R}^2$، را حداکثر خواهد کرد. برای اثبات کافی است دو طرف معادلهٔ ۵ـ۶۶ را بر $(n-1)$ تقسیم کنیم،

$$\frac{(1-\bar{R}^2)}{(n-1)} = \frac{(1-R^2)}{(n-k)}.$$

دو طرف معادلهٔ فوق را در Σy_i^2 ضرب می‌کنیم،

$$\frac{(1-\bar{R}^2)\,\Sigma y_i^2}{(n-1)} = \frac{(1-R^2)\,\Sigma y_i^2}{(n-k)} \quad . \qquad (۵\_۶۸)$$

با توجه به معادلهٔ ۴ـ۱۶، می‌دانیم $(1-R^2)\,\Sigma y_i^2 = \Sigma e_i^2$ است. همچنین در معادله‌های ۲ـ۲۸ و ۵ـ۶۱ دیدیم که اگر مجموع مربعات پسماند، یعنی Σe_i^2 را بر درجات آزادی آن، یعنی $(n-k)$ تقسیم کنیم، یک تخمین نااریب از واریانس جملهٔ اختلال خواهیم داشت؛ بنابراین

$$\frac{(1-R^2)\,\Sigma y_i^2}{(n-k)} = \frac{\Sigma e_i^2}{(n-k)} = \hat{\sigma}_u^2 . \qquad (۵\_۶۹)$$

معادلهٔ ۵ـ۶۹ را در معادلهٔ ۵ـ۶۸ جایگزین می‌کنیم،

$$\frac{(1-\bar{R}^2)\,\Sigma y_i^2}{(n-1)} = \frac{\Sigma e_i^2}{(n-k)} = \hat{\sigma}_u^2 . \qquad (۵\_۷۰)$$

مشاهده می‌شود که به موازات افزایش تعداد متغیرهای توضیحی، یعنی با افزایش k، مقدار $(n-k)$ کمتر شده؛ بنابراین مخرج کسر $\frac{\Sigma e_i^2}{n-k}$ کاهش می‌یابد. می‌دانیم با افزایش متغیرهای توضیحی، مقدار Σe_i^2 کاهش یافته یا حداقل ثابت می‌ماند؛ بنابراین در بهترین حالت، یک متغیر توضیحی جدید می‌تواند، صورت و مخرج کسر $\hat{\sigma}_u^2$ را همزمان کاهش دهد. اما اینکه آیا $\hat{\sigma}_u^2$ کم‌شده یا افزایش می‌یابد، در فصل هفتم و در قسمت ۷ـ۳ بررسی خواهد شد. در این قسمت فقط می‌خواهیم رابطهٔ بین $\hat{\sigma}_u^2$ و $\bar{R}^2$ را بررسی کنیم. با فرض ثبات n و Σy_i^2 و با توجه به معادلهٔ ۵ـ۷۰، می‌توان نتیجه گرفت که جهت تغییرات $\hat{\sigma}_u^2$ و $(1-\bar{R}^2)$ به ازای افزایش یا کاهش متغیرهای توضیحی یکسان است. این نتیجه را می‌توان به این صورت نیز بیان کرد که به ازای افزایش متغیرهای توضیحی، تغییرات $\hat{\sigma}_u^2$ در جهت مخالف تغییرات $\bar{R}^2$ خواهد بود؛ یعنی آن دسته از متغیرهای توضیحی که $\hat{\sigma}_u^2$ را کاهش می‌دهند $\bar{R}^2$ را افزایش خواهند داد.

۵ـ۶ نکته‌هایی دربارهٔ R^2 و $\bar{R}^2$

به نظر می‌رسد قبل از پایان دادن به مباحث ضریب تعیین لازم باشد بعضی مسائل را فهرست‌وار مطرح کنیم.

۱. با توجه به معادلهٔ ۱۶ـ۴ داریم:

$$R^2 = 1 - \frac{\Sigma e_i^2}{\Sigma y_i^2}.$$

می‌دانیم تخمین زننده حداقل مربعات معمولی، مجموع مربعات پسماند، یعنی Σe_i^2، را حداقل می‌کند. با توجه به فرمول فوق، ملاحظه می‌شود که وقتی Σe_i^2 حداقل باشد، R^2 حداکثر خواهد بود. بدین ترتیب می‌توان گفت که معیار روش حداقل مربعات معمولی در تخمین پارامترها حداکثر نمودن R^2 است. نتیجه می‌گیریم که معیارهای حداقل کردن مجموع مربعات پسماند و حداکثر نمودن R^2 درواقع یکی هستند.

مسألهٔ حداکثرسازی R^2 یک مبحث بسیار مهم در اقتصادسنجی است. وقتی در تخمین یک مدل رگرسیون به مقداری از ضریب تعیین می‌رسیم که نزدیک به یک است، معمولاً این امر بدین صورت تفسیر می‌شود که پارامترهای مدل بسیار خوب تخمین زده شده است. در بین نکاتی که در این قسمت مطرح خواهیم کرد، سعی می‌کنیم به این سؤال پاسخ دهیم که آیا واقعاً چنین استنتاجی صحیح است؟ ابتدا باید دو حالت زیر را از یکدیگر تفکیک کرد.

در حالت اول، فرض بر این است که مدل رگرسیون از هر نظر دقیقاً تعریف شده است به گونه‌ای که ما حق هیچگونه تصرفی را در آن نداریم؛ به عبارت دقیقتر، اولاً شکل ریاضی تابع رگرسیون کاملاً مشخص است که مثلاً خطی است یا غیرخطی، ثانیاً تعداد و نوع متغیرهای توضیحی کاملاً شناخته شده است به نحوی که نه می‌توان متغیری به مدل اضافه کرد و نه متغیر یا متغیرهایی از آن را حذف کرد، و ثالثاً فرضهای جملهٔ اختلال را نیز به طور دقیق و غیرقابل تغییری قبول کرده‌ایم. در چنین شرایطی از ما خواسته شده است که پارامترهای این مدل را تخمین بزنیم. نکتهٔ مهم این است که فقط در این حالت

می‌توان گفت که معیار روش حداقل مربعات معمولی همان حداکثرسازی R^2 است؛ یعنی باید پارامترهای مدل را چنان تخمین زد که $\hat{Y}_i$ بتواند در بالاترین سطح ممکن، تغییرات Y_i را توضیح دهد. بدیهی است این معیار چیزی نیست جز حداقل‌سازی مجموع تغییرات توضیح داده نشده یا پسماندها، که همان معیار متعارف حداقل مربعات معمولی است.

در حالت دوم، فرض می‌کنیم که می‌خواهیم پارامترهای یک مدل رگرسیون را چنان تخمین بزنیم که R^2 حداکثر شود. اما برای رسیدن به این هدف نه تنها اجازه داریم شکل ریاضی تابع رگرسیون را به هر نحوی که مقتضی باشد تغییر دهیم، بلکه می‌توانیم هر متغیر توضیحی جدیدی را که بخواهیم به مدل اضافه کنیم یا هر متغیر دیگری را نیز هنگامی که لازم است، حذف کنیم. همچنین می‌توانیم هر فرض جدیدی را نیز که موجب افزایش R^2 شود مطرح کنیم، یا در فرضهایی که قبلاً ارائه شده است به سهولت تصرف کنیم. به طور خلاصه، هر اقدامی که بتواند به افزایش R^2 منجر شود، مجاز خواهد بود. همهٔ بحث ما این است که چنین تفسیری از حداکثر نمودن R^2 کاملاً غلط است و هیچگونه ارزشی از نظر کاربردی ندارد؛ زیرا ماهیت این نحوه برخورد با R^2 این است که سعی کنیم مدلی بسازیم که بتواند از مجموعه مشاهدات موجود در نمونهٔ مفروض یک R^2 بسیار بالا نتیجه دهد[1]. در صورت موفقیت در ساختن چنین مدل مصنوعی، بدیهی است اگر نمونه را تغییر داده و یا مجموعهٔ دیگری از مشاهدات کار کنیم، چه بسا مدلی که با اینهمه زحمت ساخته‌ایم، R^2 بسیار پایینی را نشان دهد.

نتیجه‌ای که می‌گیریم، این نیست که اولاً نباید به R^2 توجه چندانی کرد و ثانیاً نباید در طراحی مدل هیچگونه تغییری داد. بلکه برعکس، باید با توجه به مقادیر جملات پسماند، یعنی تغییرات توضیح داده نشده، سعی نمود مدل را به نحوی تغییر داد که قدرت توضیحی بیشتری داشته باشد ـ که درواقع دقیقاً به معنای افزایش R^2 است. اما نکتهٔ مهم این است که این تغییراتی که در مدل می‌دهیم، باید مجوز نظری داشته باشد. نارسایی عملکرد تخمین ما در تفسیر تغییرات مشاهده شده در متغیر درون‌زا، به طور

---

۱. در اقتصادسنجی اصطلاحاً به این روش، Data Mining می‌گویند.

کلی روشن کنندهٔ ضعف نظری مدل رگرسیون است. همهٔ کوشش باید این باشد که بتوان در ساختار ریاضی مدل و فرضهای آن، یا فرضهای جملهٔ اختلال چنان تصرف کرد که این ضعف نظری کاهش یابد. بدیهی است هر اقدام مؤثری در این زمینه ممکن است به افزایش R^2 منجر شود، اما چه بسا ضرورتاً R^2 بسیار بالایی نتیجه ندهد. بنابراین، اگر دو مدل در تفسیر تغییرات یک متغیر درون‌زا داشته باشیم که اولی R^2 بسیار بالا مثلاً ۰/۹۹ و دومی R^2 مثلاً ۰/۵۲ داشته باشد، اما مدل اول چنان «ساخته» شده باشد که R^2 بالا نتیجه دهد، در حالی که مدل دوم از مبنای نظری قوی‌تری برخوردار باشد، آنگاه قطعاً مدل دوم مرجح است. در بین نکاتی که مطرح شد دوباره به این مبحث اشاره خواهد شد. اما به این نکته اشاره می‌کنیم که یکی از راههای تشخیص اینکه آیا یک مدل، تنها برای رسیدن به R^2 بالا طراحی شده یا خیر، این است که ابتدا به موازین نظری آن توجه شود و سپس از آزمونهای مختلف آماری برای معنی‌دار بودن پارامترهای آن استفاده کرده و سرانجام با ایجاد تغییر مختصر در نمونه، تغییرات R^2 را ارزیابی کنیم.

۲. معمولاً این سؤال مطرح می‌شود که «(R^2 بالا)» دقیقاً به چه معنی است؟ بدیهی است نمی‌توان به این سؤال به طور دقیق پاسخ داد، اما توجه به نکات کلی زیر مفید است. در مواردی که مشاهدات ما در مورد متغیرهای درون‌زا و برون‌زای یک مدل به صورت سریهای زمانی است، R^2 مقدار بسیاری نشان می‌دهد؛ زیرا تغییر متغیرها در زمان، روند تقریباً مشابهی دارند. می‌توان گفت[1] که اگر یک متغیر اقتصادی را به طور تصادفی انتخاب کنیم، مثلاً Y_t و مقدار این متغیر را در دورهٔ قبل در نظر بگیریم، Y_{t-1}، و سپس مدل رگرسیون Y_t بر Y_{t-1} را بسازیم، به طور متوسط R^2 حدود ۰/۷ خواهد بود. همچنین اگر یک متغیر اقتصادی را به طور تصادفی انتخاب کرده و رگرسیون آن را با دو تا شش متغیر اقتصادی دیگر ـ که آنها نیز به طور تصادفی انتخاب می‌شود ـ در نظر بگیریم، مقدار R^2 به طور متوسط از ۰/۵ بیشتر خواهد شد به شرط اینکه مشاهدات مربوط به متغیرها برحسب سری زمانی باشد. با وجود این، باید توجه کرد که وقتی در یک مدل

---

۱. به مقالهٔ E. Ames and S. Reiter (۱۹۶۱) مراجعه شود.

رگرسیون، مشاهدات به صورت مقطعی باشد، مقادیر R^2 در مقایسه با حالت سری زمانی بسیار کمتر خواهد بود.

برای تبیین این مسأله مثالی می‌زنیم. فرض کنید مقادیر یک متغیر اقتصادی به‌طور مداوم در زمان، یک سیر نزولی را نشان می‌دهد، به گونه‌ای که می‌توان مسیر تغییرات زمانی آن را با معادلهٔ $Y_t = ۵۰ - ۰/۲\,t$ نشان داد که در آن t زمان است. همچنین فرض می‌کنیم یک متغیر دیگری نیز، مانند X_t، داریم که اساساً هیچگونه ارتباطی با Y_t ندارد، اما ملاحظه می‌کنیم که مقادیر آن در زمان ـ به طور مرتب زیاد می‌شود. اگر مسیر تغییرات زمانی X_t را از نظر ریاضی با معادلهٔ $X_t = ۵ + ۰/۴\,t$، تقریب کنیم، به راحتی می‌توان از این معادله، t را برحسب X_t به دست آورده و در معادلهٔ Y_t قرار داد تا بدین ترتیب معادلهٔ Y_t برحسب X_t به دست آید،

$$t = ۲/۵\,X_t - ۱۲/۵\,,$$

$$Y_t = ۵۰ - ۰/۲\,(۲/۵\,X_t - ۱۲/۵)\,,$$

یا:

$$Y_t = ۵۲/۵ - ۰/۵\,X_t\,.$$

بدیهی است با چنین ساختار ریاضی بین متغیرها، اگر با استفاده از سری زمانی تغییرات X_t و Y_t، مدل رگرسیون

$$Y_t = \alpha + \beta\,X_t + U_t\,,$$

را تخمین بزنیم، قطعاً R^2 بسیار بالایی را به دست خواهیم آورد، و حال آنکه می‌دانیم با توجه به موازین نظری، اساساً هیچگونه رابطه‌ای بین X_t و Y_t موجود نیست.

نتیجهٔ دیگری که از مثال فوق می‌توان گرفت این است که در یک مدل رگرسیون Y_t روی X_t، اگر مقدار عددی R^2 حتی نزدیک به یک باشد، مطلقاً دلالت نمی‌کند که ضرورتاً یک رابطهٔ علت و معلولی قوی بین X_t و Y_t وجود دارد. اثبات وجود رابطه علت و معلولی، در حوزهٔ تحلیلهای نظری صورت می‌پذیرد. در مثال قبل دیدیم که اساساً هیچگونه رابطه‌ای بین X_t و Y_t وجود نداشت؛ در چنین مواردی، مقدار عددی R^2، تنها

منعکس کنندهٔ روند موجود در متغیرهای X_t و Y_t است. راه حلی که معمولاً برای خنثی کردن تأثیر روند بر مقدار R^2 می‌توان پیشنهاد کرد، این است که به جای استفاده از مقادیر مطلق متغیرهای درون‌زا و برون‌زا از نسبت یا نرخ تغییرات آنها استفاده شود، زیرا وقتی با نرخ تغییرات یک متغیر کار می‌کنیم، تأثیر روند را تا حد بسیاری از بین خواهیم برد. برای مثال، در اکثر کشورهای صنعتی ملاحظه می‌شود که بعد از جنگ جهانی دوم، تغییرات تولید و تورم، یک روند صعودی را نشان داده است، در حالی که این روند در نرخ رشد تولید یا نرخ رشد تورم ملاحظه نمی‌شود. با وجود این، کار کردن با نرخ تغییر متغیرها نیز مسائل خاصل خود را دارد که بحث بیشتر دربارهٔ آنها از موضوع این قسمت خارج است و باید در اقتصادسنجی کاربردی بررسی شود.

۳. می‌دانیم که یکی از اهداف مهم روشهای تخمین در اقتصادسنجی این است که بتوان به تخمینهای «خوبی» از پارامترهای مدل رگرسیون رسید. با وجود این، نباید «خوبی» تخمین پارامترها را با توجه به معیار «مقدار R^2» تفسیر کرد؛ بدین معنی که اگر R^2 زیاد باشد نتیجه بگیریم که تخمین پارامترها خوب است و برعکس البته تخمینهای خوب از پارامترهای یک مدل ممکن است موجب افزایش مقدار R^2 شود، اما عکس آن صحیح نیست. بنابراین، می‌توان نتیجه گرفت که معیار R^2 نقش چندانی در استنتاج نهایی دربارهٔ خوبی تخمین مدل رگرسیون ندارد.

با اینکه در اقتصادسنجی نسبت به عدم اهمیت زیاد مقدار R^2، اتفاق‌نظر وجود دارد، اما در تمام محاسبات کامپیوتری در تخمین پارامترهای یک مدل رگرسیون، حتماً مقدار R^2 منعکس می‌شود و همواره مقادیر بالای R^2 به عنوان معیاری در موفقیت تخمین یک مدل معرفی می‌شود.[1]

۴. با توجه به معادلهٔ ۴۸ـ۵ می‌دانیم R^2 تابعی از Σe_t^2 و Σy_t^2 و در نتیجه تابعی از e_t و y_t است؛ بنابراین در نمونه گیریهای تکراری، R^2 از یک نمونه به نمونهٔ دیگر تغییر می‌کند. درواقع، چون e_t یک متغیر تصادفی است، R^2 نیز یک متغیر تصادفی خواهد بود

---

۱. برای توضیحات بیشتر به مقالهٔ J. S. Cramer (۱۹۸۷) مراجعه شود.

که دارای توزیع احتمال است و می‌توان فرضیه‌های مختلف در مورد آن را آزمود. در معادلهٔ ۵۴ـ۲ نیز دیدیم که می‌توان آمارهٔ F را برحسب r^2 نوشت. در فصل آینده و در قسمت ۶ـ۴ نیز دوباره به این مسأله مراجعه نموده و آن را برای حالت رگرسیون چندمتغیره مطرح خواهیم کرد. در این قسمت به مفهوم آزمون فرضیهٔ $H_0: R^2 = 0$ می‌پردازیم:

در مدل رگرسیون

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t ,$$

فرض بر این است که تمام پارامترهای β غیر از β_1 صفر است؛ یعنی متغیر درون‌زای Y_t مطلقاً تحت تأثیر متغیرهای توضیحی $X_2 , X_3 , \ldots , X_k$ نیست. چنین فرضی دلالت بر این می‌کند که در جامعهٔ آماری، مقدار مجموع تغییرات توضیح داده شده، یعنی $\sum \hat{y}_t^2$، برابر صفر می‌شود؛ بنابراین مقدار R^2 جامعه نیز صفر خواهد شد. اما در عمل ما فقط یک نمونه‌داریم و با اینکه درواقع، تمام پارامترهای مدل غیر از β_1 صفراست درتخمین حاصل از نمونهٔ مفروض، برای اکثر پارامترها، تخمینهایی غیرصفر خواهیم داشت؛ بنابراین مقداری نیز برای R^2 به دست خواهد آمد که ضرورتاً غیر صفر است. سؤال این است که چگونه به کمک R^2 به دست آمده از نمونه می‌توان دربارهٔ این فرضیه ـ که R^2 جامعه صفر است ـ قضاوت آماری کرد؟ در فصل آینده و درمعادلهٔ ۶ـ۴۱ خواهیم دید که کمیت

$$\left(\frac{n-k}{k-1}\right) \left(\frac{R^2}{1-R^2}\right) ,$$

دارای توزیع F با $(k-1 , n-k)$ درجهٔ آزادی است.

۵. در ارائه بحث فوق می‌توان به یک مورد دیگر نیز اشاره کرد. فرض کنید در یک مدل رگرسیون چند متغیره یک یا چند متغیر توضیحی دیگر وارد مدل کنیم. سؤال این است که با چه معیاری می‌توان ضرورت یا اهمیت متغیرهای جدید در مدل رگرسیون مفروض را ثابت کرد؟ با توجه به مباحثی که در ابتدای قسمت ۵ـ۵ و در تعریف $\bar{R}^2$ گفتیم، به نظر می‌رسد که چون ورود هر متغیر توضیحی جدید باعث می‌شود که R^2

افزایش یافته و یا حداقل کاهش نیابد بنابراین با معیار R^2، مدلی که شامل تعداد بیشتری از متغیرهای توضیحی شده است بر مدل قبلی مرجح خواهد بود. همچنین در مباحث قبلی دیدیم که ورود متغیر یا متغیرهای توضیحی، درجات آزادی Σe_t^2 و Σy_t^2 را تغییر می‌دهد؛ پس معیار $\bar{R}^2$ نسبت به R^2 از برتری نسبی برخوردار می‌شود. درواقع، می‌توان حالتی را در نظر گرفت که اضافه کردن یک متغیر توضیحی جدید، موجب افزایش R^2 و کاهش $\bar{R}^2$ می‌شود.

برای توضیح بیشتر این نکته، معادلهٔ ۵ـ۶۳ را یک بار دیگر می‌نویسیم،

$$\bar{R}^2 = ۱ - \left(\frac{n-۱}{n-k}\right)\frac{\Sigma e_t^2}{\Sigma y_t^2}.$$

فرض کنید $n=۲۱$، $k=۶$، $\Sigma e_t^2=۱۵$ و $\Sigma y_t^2=۱۰۰$ است؛ بنابراین:

$$\bar{R}^2 = ۱ - \left(\frac{۲۱-۱}{۲۱-۶}\right)\frac{۱۵}{۱۰۰} = ۱ - ۰/۲۰ = ۰/۸۰$$

برای همین مثال، مقدار R^2 را نیز محاسبه می‌کنیم:

$$R^2 = ۱ - \frac{\Sigma e_t^2}{\Sigma y_t^2} = ۱ - \frac{۱۵}{۱۰۰} = ۰/۸۵ .$$

یک متغیر توضیحی دیگر وارد مدل می‌کنیم. فرض بر این است که این متغیر جدید، قدرت توضیحی بسیاری ندارد؛ با وجود این، توانسته است مجموع مربعات پسماند را تا حدی کاهش دهد، به گونه‌ای که داریم، $\Sigma e_t^2=۱۴/۵$. مقدار $\bar{R}^2$ در این حالت به قرار زیر خواهد بود،

$$\bar{R}^2 = ۱ - \left(\frac{۲۱-۱}{۲۱-۷}\right)\frac{۱۴/۵}{۱۰۰} = ۱ - ۰/۲۰۷ = ۰/۷۹۳,$$

یعنی مقدار $\bar{R}^2$ کاهش یافته است. مفید است مقدار R^2 را نیز حساب کنیم،

$$R^2 = ۱ - \frac{۱۴/۵}{۱۰۰} = ۰/۸۵۵ ,$$

ملاحظه می‌شود که R^2 افزایش نشان می‌دهد.

نتیجه می‌گیریم که معیار $\bar{R}^2$ از R^2 دقیقتر است. با وجود این، برای گرفتن تصمیم نهایی در مورد حفظ یا حذف متغیرهای توضیحی جدید، باید علاوه بر دقت در موازین نظری، از آزمونهای آماری نیز استفاده کرد که موضوع بحث فصل آینده خواهد بود.

۶. در فصل نهم و در قسمت ۳ـ۹ به این نکته اشاره خواهیم کرد که اگر یک متغیر توضیحی جدید به مدل رگرسیون اضافه کنیم، موقعی $\bar{R}^2$ زیاد می‌شود که مقدار t مربوط به تخمین پارامتر این متغیر توضیحی جدید بیشتر از یک باشد. بنابراین، ملاحظه می‌شود که مسألهٔ حداکثرسازی $\bar{R}^2$ کاملاً با قاعدهٔ حفظ متغیرهای توضیحی در یک مدل رگرسیون متفاوت است؛ زیرا می‌دانیم موقعی می‌توان یک متغیر توضیحی مانند X_{ik} را مفید و ضروری تشخیص داد که فرضیهٔ $H_0: \beta_i = 0$ در سطح معنی‌دار، مثلاً ۵ درصد رد شود. اگر این فرضیه با $t < 1$ رد شود، نتیجه می‌گیریم که متغیر توضیحی مورد نظر باید در مدل حفظ شود، در حالی که می‌دانیم ورود این متغیر، مقدار $\bar{R}^2$ را افزایش نداده است؛ زیرا مقدار t متعلق به تخمین پارامتر آن از یک کمتر است. نتیجه می‌گیریم که در مورد حفظ یا حذف یک متغیر توضیحی، دو معیار آزمونهای آماری و $\bar{R}^2$ ممکن است در مواردی به یک استنتاج منتهی نشود.

مثال ۳ـ۵ مدل رگرسیون را ـ که در مثال ۲ـ۵ مطرح کردیم ـ یک بار دیگر در نظر بگیرید:

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t .$$

با استفاده از جدول ۱ـ۵،

اولاً، R^2 را یک بار با فرمول ۴۰ـ۵ و بار دیگر با استفاده از فرمول ۴۸ـ۵ به دست آورید.

ثانیاً، مقدار $\bar{R}^2$ را محاسبه کنید.

۱. ابتدا فرمول ۴۰ـ۵ را می‌نویسیم،

$$R^2 = \frac{\hat{\beta}' X' y}{y' y} .$$

می‌دانیم عناصر بردارهای موجود در فرمول فوق باید برحسب انحراف از میانگین باشد. بنابراین، ضرورتاً از محاسبات مربوط به مثال ۲ـ۵ استفاده می‌کنیم. با توجه به جدول ۳ـ۵ می‌دانیم

$$\Sigma = x_t y_t = ۱۶ \quad , \quad \Sigma z_t y_t = ۹ .$$

همچنین در محاسبات مثال ۲ـ۵ دیدیم

$$\hat{\beta} = ۲/۵ \quad , \quad \hat{\gamma} = -۱/۵ .$$

برای محاسبهٔ $R^۲$ به $\Sigma y_t^۲$ نیز نیاز داریم که باید محاسبه کنیم،

$$\Sigma y_t^۲ = \mathbf{y}'\mathbf{y} = ۱ + ۹ + ۱۶ + ۱ + ۱ = ۲۸ .$$

صورت کسر $R^۲$، یعنی $\hat{\beta}'\mathbf{X}'\mathbf{y}$ عبارت است از

$$\hat{\beta}'\mathbf{X}'\mathbf{y} = [\hat{\beta} \quad \hat{\gamma}] \begin{bmatrix} \Sigma x_t y_t \\ \Sigma z_t y_t \end{bmatrix} = \hat{\beta} \Sigma x_t y_t + \hat{\gamma} \Sigma z_t y_t ,$$

$$= ۲/۵ (۱۶) - ۱/۵ (۹) = ۴۰ - ۱۳/۵ = ۲۶/۵ .$$

بنابراین

$$R^۲ = \frac{۲۶/۵}{۲۸} = ۰/۹۴۶۴ ,$$

یعنی تخمین ما از مدل توانسته است حدود ۹۵ درصد تغییرات Y_t را توضیح دهد.

اگر بخواهیم $R^۲$ را با استفاده از فرمول ۴۸ـ۵ به دست آوریم، باید چنین نوشت،

$$R^۲ = ۱ - \frac{\mathbf{e}'\mathbf{e}}{\mathbf{y}'\mathbf{y}} .$$

برای محاسبهٔ $\mathbf{e}'\mathbf{e}$ ابتدا معادلهٔ ۴۵ـ۵ را دوباره می‌نویسیم،

$$\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \mathbf{e}'\mathbf{e} ,$$

$$۲۸ = ۲۶/۵ + \mathbf{e}'\mathbf{e} ,$$

در نتیجه

$$e'e = ۱/۵ .$$

بنابراین مجموع مربعات پسماند یا تغییرات توضیح داده نشده برابر ۱/۵ است؛ در نتیجه

$$R^۲ = ۱ - \frac{۱/۵}{۲۸} = \frac{۲۸ - ۱/۵}{۲۸} = ۰/۹۴۶۴ .$$

۲. طبق معادلهٔ ۶۴-۵ می‌دانیم که

$$\bar{R}^۲ = ۱ - \left(\frac{n-۱}{n-k}\right)(۱ - R^۲) ,$$

در نتیجه خواهیم داشت

$$\bar{R}^۲ = ۱ - \left(\frac{۵-۱}{۵-۳}\right)(۱ - ۰/۹۴۶۴) ,$$

$$= ۱ - ۲\,(۰/۰۵۳۶) = ۰/۸۹۲۸ .$$

۵-۷ بیان ماتریسی $R^۲$ برحسب مشاهدات اصلی*

تاکنون در تمام فرمولهایی که برای رگرسیون چند متغیره استخراج کرده‌ایم هیچگاه علامت خاصی را به کار نبرده‌ایم که بیان کنندهٔ این واقعیت باشد که محاسبات برحسب انحراف از میانگین انجام شده است. دلیل این امر، پرهیز از پیچیده‌تر کردن علائم در بیان فرمولهاست. تا زمانی که می‌توان از متن بحث به این نکته پی برد که آیا محاسبات با استفاده از مشاهدات اصلی است یا انحراف از میانگین، ضرورتی برای معرفی علائم جدید نیست. اما مطرح کردن این سؤال مفید است که آیا روشی وجود دارد که بتوان به آن وسیله ماتریسهایی را که برحسب مشاهدات اصلی محاسبه شده است، به مقادیر انحراف از میانگین تبدیل کرد؟ در صورتی که پاسخ مثبت باشد به راحتی می‌توان در مواردی که به مقادیر انحراف از میانگین نیاز است، از این تبدیل استفاده کرده و آن را مستقیماً روی مشاهدات اصلی اجرا کرد. این مسأله، بویژه در محاسبهٔ $R^۲$ بسیار مهم است؛ زیرا می‌دانیم در فرمول

$$R^۲ = \frac{\hat{y}'\hat{y}}{y'y} ,$$

مقادیر صورت و مخرج باید برحسب انحراف از میانگین محاسبه شود. سؤال این است که چگونه می‌توان در محاسبهٔ R^2 مستقیماً از ماتریسهایی استفاده کرد که برحسب مشاهدات اصلی تنظیم شده است.

برای پاسخ به این سؤال، ابتدا ماتریس $\mathbf{A}$ را معرفی می‌کنیم و آن را ماتریس تبدیل می‌نامیم،

$$\mathbf{A} = \mathbf{I} - \frac{1}{n}\mathbf{i}\mathbf{i}' , \qquad (۵\text{-}۷۱)$$

که در آن $\mathbf{A} \to n \times n$. $\mathbf{i}$ یک بردار شامل n عدد یک و $\mathbf{I}$ یک ماتریس $n \times n$ و تمام عناصر قطری آن یک و عناصر غیرقطری آن صفر است (ماتریس یکه)،

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & 1 \end{bmatrix} - \frac{1}{n} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \\ 1 & 1 & \cdots & 1 \end{bmatrix} . \qquad (۵\text{-}۷۲)$$

نشان می‌دهیم که اگر ماتریس تبدیل ($\mathbf{A}$) را در هر برداری از سمت چپ ضرب کنیم، بردار به دست آمده برحسب انحراف از میانگین خواهد بود. برای این منظور، بردار ستونی $\mathbf{y}$، شامل n مقدار را در نظر می‌گیریم. می‌دانیم $\mathbf{y}$ برحسب مشاهدات اصلی است، یعنی

$$\mathbf{y}' = [Y_1 \quad Y_2 \quad , \ldots , \quad Y_n] .$$

ماتریس تبدیل ($\mathbf{A}$) را از سمت چپ در $\mathbf{y}$ ضرب می‌کنیم.

$$\mathbf{A}\mathbf{y} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} - \frac{1}{n} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} ,$$

$$= \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} - \begin{bmatrix} \bar{Y} \\ \bar{Y} \\ \cdot \\ \cdot \\ \bar{Y} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{bmatrix} .$$

ملاحظه می‌شود $\mathbf{\Lambda y}$، بردار $\mathbf{y}$ را به برداری تبدیل کرد که عناصر آن انحراف از میانگین مقادیر Y است.

در اینجا اشاره به دو مورد خاص ضروری است. نکتهٔ اول این است که اگر $\mathbf{\Lambda}$ را در هر برداری ضرب کنیم که عناصر آن با یکدیگر مساوی است، یک بردار صفر خواهیم داشت؛ مثلاً

$$\mathbf{\Lambda i} = \mathbf{o} . \qquad (۵\text{-}۷۳)$$

نکتهٔ دوم این است که چون میانگین مقادیر پسماند صفر است، بنابراین

$$\mathbf{\Lambda e} = \mathbf{e} , \qquad (۵\text{-}۷۴)$$

یعنی بردار $\mathbf{e}$ همواره برحسب انحراف از میانگین است. همچنین می‌توان نشان داد که $\mathbf{\Lambda}$ یک ماتریس متقارن و «خودتوان[1]» است؛ یعنی $\mathbf{\Lambda} = \mathbf{\Lambda}'$ و $\mathbf{\Lambda\Lambda} = \mathbf{\Lambda}^2 = \mathbf{\Lambda}$.

اکنون به بررسی مدلهای رگرسیون می‌پردازیم. مدل رگرسیون ۵-۱ یا ۵-۴ مفروض است

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{u} ,$$

که در آن، $\mathbf{y} \to n \times 1$، $\mathbf{X} \to (n \times k)$، $\mathbf{\beta} \to k \times 1$ و $\mathbf{u} \to n \times 1$. فرض بر این است که مقادیر تمام متغیرهای این مدل برحسب مشاهدات اصلی بیان شده است. با توجه به معادلهٔ ۵-۱۷ می‌دانیم

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} . \qquad (۵\text{-}۷۵)$$

---

1. Idempotent Matrix

به پیوست «۵ ـ الف» مراجعه شود.

ماتریس $\mathbf{X}$ را به صورت زیر افراز می‌کنیم،

$$\mathbf{X}= [\boldsymbol{x}_1 \quad \mathbf{X}_2] ,$$

که در آن، $\boldsymbol{x}_1 = \mathbf{i}$، یک بردار ستونی شامل اعداد یک است که منعکس کنندهٔ جملهٔ ثابت در مدل رگرسیون ۵ـ۱ است. $\mathbf{X}_2$ نیز یک ماتریس $n \times (k - 1)$ از مشاهدات متغیرهای توضیحی X_2، X_3 تا X_k است. با جایگزینی معادلهٔ فوق در معادلهٔ ۵ـ۷۵ داریم

$$\mathbf{y}= \mathbf{i}\hat{\beta}_1 + \mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{e} , \qquad (۵ـ۷۶)$$

که در آن $\hat{\boldsymbol{\beta}}' = [\hat{\beta}_1 \quad \hat{\boldsymbol{\beta}}_2]$ بدین معنی است که بردار $\hat{\boldsymbol{\beta}}$ به دو قسمت افراز شده است که قسمت اول شامل یک ضریب $\hat{\beta}_1$ و قسمت دوم شامل بردار $\hat{\boldsymbol{\beta}}_2$ با $(k - 1)$ عضو است. دو طرف معادلهٔ ۵ـ۷۶ را از سمت چپ در $\mathbf{A}$ ضرب می‌کنیم،

$$\mathbf{A}\,\mathbf{y}= \mathbf{A}\,\mathbf{i}\hat{\beta}_1 + \mathbf{A}\,\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{A}\,\mathbf{e} .$$

با توجه به معادله‌های ۵ـ۷۳ و ۵ـ۷۴، معادلهٔ فوق را می‌توان چنین نوشت،

$$\mathbf{A}\,\mathbf{y}= \mathbf{A}\,\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{e} . \qquad (۵ـ۷۷)$$

دو طرف معادله فوق را در $\mathbf{X}_2'$ ضرب می‌کنیم،

$$\mathbf{X}_2'\,\mathbf{A}\,\mathbf{y}= \mathbf{X}_2'\,\mathbf{A}\,\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{X}_2'\,\mathbf{e} .$$

از معادلهٔ ۵ـ۲۴ می‌دانیم که جمله‌های پسماند از متغیرهای توضیحی مستقل هستند؛ یعنی $\mathbf{X}_2'\,\mathbf{e}= \mathbf{0}$. بنابراین:

$$\mathbf{X}_2'\,\mathbf{A}\,\mathbf{y}= \mathbf{X}_2'\,\mathbf{A}\,\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 . \qquad (۵ـ۷۸)$$

با توجه به تقارن و خودتوانی ماتریس $\mathbf{A}$، معادلهٔ فوق را می‌توان چنین نوشت،

$$(\mathbf{A}\,\mathbf{X}_2)'\,(\mathbf{A}\,\mathbf{y})= (\mathbf{A}\,\mathbf{X}_2)'\,(\mathbf{A}\,\mathbf{X}_2)\hat{\boldsymbol{\beta}}_2 . \qquad (۵ـ۷۹)$$

معادلهٔ ۵ـ۷۹ درواقع همان معادله‌های نرمال برای مدل رگرسیون چند متغیره ۵ـ۴ است و

صورت دیگری از معادلهٔ ۳۳ـ۵ می‌باشد. برای توضیح بیشتر معادلهٔ ۳۳ـ۵ را دوباره می‌نویسیم،

$$(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} \ .$$

فرق بین معادلهٔ فوق با معادلهٔ ۷۹ـ۵ این است که در معادلهٔ ۷۹ـ۵ ماتریس $\mathbf{\Lambda}$ توانسته است $\mathbf{X}_2$ و $\mathbf{y}$ را برحسب انحراف از میانگین بنویسد. بدیهی است معادلهٔ ۷۹ـ۵ را می‌توان برای $\hat{\boldsymbol{\beta}}_2$ به این صورت حل کرد،

$$\hat{\boldsymbol{\beta}}_2 = [(\mathbf{\Lambda}\mathbf{X}_2)'(\mathbf{\Lambda}\mathbf{X}_2)]^{-1}(\mathbf{\Lambda}\mathbf{X})'(\mathbf{\Lambda}\mathbf{y}) \ . \qquad (۵\text{ـ}۸۰)$$

معادلهٔ فوق مقادیر $\hat{\beta}_k$ ، ... ، $\hat{\beta}_3$ ، $\hat{\beta}_2$ را در یک مدل رگرسیون چند متغیره محاسبه می‌کند. یادآوری می‌کنیم که در این معادله، $\mathbf{X}_2 \to n\times(k-1)$ و $\mathbf{y} \to n\times 1$ و $\hat{\boldsymbol{\beta}}_2 \to (k-1)\times 1$. این معادله دقیقاً صورت دیگری از معادلهٔ ۳۴ـ۵ است.

حال به بررسی R^2 می‌پردازیم. دیدیم که $\mathbf{\Lambda}\mathbf{y}$ بردار مقادیر متغیر درون‌زا و $\mathbf{\Lambda}\mathbf{X}_2$ نیز ماتریس مقادیر $(k-1)$ متغیر برون‌زا برحسب انحراف از میانگین هستند. ابتدا مجموع تغییرات متغیر درون‌زا $(\mathbf{y}'\mathbf{y})$ را برحسب انحراف از میانگین به دست می‌آوریم. کافی است $\mathbf{\Lambda}\mathbf{y}$ را ترانهاد نموده و از سمت چپ در $\mathbf{\Lambda}\mathbf{y}$ ضرب کنیم، خواهیم داشت

$$\text{TSS} = \mathbf{y}'\mathbf{\Lambda}'\mathbf{\Lambda}\mathbf{y} \ , \quad \text{برحسب انحراف از میانگین}$$

که با توجه به تقارن و خودتوانی، ماتریس $\mathbf{\Lambda}$ عبارت خواهد بود از

$$\text{TSS} = \mathbf{y}'\mathbf{\Lambda}\mathbf{y}$$

در معادلهٔ ۴۵ـ۵ دیدیم

$$\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \mathbf{e}'\mathbf{e} \ .$$

باید رابطهٔ فوق را برای این حالت نیز ثابت کنیم. کافی است معادلهٔ ۷۷ـ۵ را ترانهاد نموده،

$$\mathbf{y}'\mathbf{\Lambda}' = \hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{\Lambda}' + \mathbf{e}' \ ,$$

و سپس در معادلهٔ ۵ـ۷۷ ضرب کنیم؛ در نتیجه

$$\mathbf{y}'\mathbf{\Lambda}'\mathbf{\Lambda}\mathbf{y}=\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{\Lambda}'(\mathbf{\Lambda}\mathbf{X}_1\hat{\boldsymbol{\beta}}_2)+\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{\Lambda}'\mathbf{e}+\mathbf{e}'\mathbf{\Lambda}\mathbf{X}_1\hat{\boldsymbol{\beta}}_2'+\mathbf{e}'\mathbf{e}.$$

با مراجعه به ۵ـ۷۴ و استفاده از تقارن $\mathbf{\Lambda}'$، داریم

$$\mathbf{\Lambda}'\mathbf{e}=\mathbf{e} \quad , \quad \mathbf{e}'\mathbf{\Lambda}=\mathbf{e},$$

و همچنین با توجه به معادلهٔ ۵ـ۲۴ می‌دانیم که

$$\mathbf{X}_2'\mathbf{e}=\mathbf{0} \quad , \quad \mathbf{e}'\mathbf{X}=\mathbf{0}.$$

بنابراین جمله‌های دوم و سوم از سمت راست معادله فوق حذف شده، خواهیم داشت

$$\mathbf{y}'\mathbf{\Lambda}\mathbf{\Lambda}\mathbf{y}=\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{\Lambda}'(\mathbf{\Lambda}\mathbf{X}_1\hat{\boldsymbol{\beta}}_1)+\mathbf{e}'\mathbf{e},$$

که با استفاده از تقارن و خودتوانی $\mathbf{\Lambda}$ به صورت زیر نوشته می‌شود،

$$\mathbf{y}'\mathbf{\Lambda}\mathbf{y}=\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{\Lambda}\mathbf{X}_1\hat{\boldsymbol{\beta}}_2'+\mathbf{e}'\mathbf{e}. \qquad (۵ـ۸۱)$$

مفهوم معادله فوق عبارت است از

$$\mathrm{TSS}=\mathrm{ESS}+\mathrm{RSS}.$$

می‌دانیم R^2 بنابر تعریف برابر است با $\frac{\text{تغییرات توضیح داده شده (ESS)}}{\text{کل تغییرات (TSS)}}$ و

$$R^2=\frac{\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{\Lambda}\mathbf{X}_1\hat{\boldsymbol{\beta}}_2'}{\mathbf{y}'\mathbf{\Lambda}\mathbf{y}}, \qquad (۵ـ۸۲)$$

معادلهٔ ۵ـ۷۸ را در معادلهٔ فوق جایگزین می‌کنیم،

$$R^2=\frac{\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{\Lambda}\mathbf{y}}{\mathbf{y}'\mathbf{\Lambda}\mathbf{y}}. \qquad (۵ـ۸۳)$$

فرمول فوق، مقدار R^2 را برای حالتی محاسبه می‌کند که $\mathbf{y}$ و $\mathbf{X}_2$ برحسب مشاهدات اصلی ـ نه انحراف از میانگین ـ محاسبه شده است.

روش فوق در استخراج R^2 اهمیت نظری بسیاری دارد؛ زیرا یک روش عمومی در حل این گونه مسائل را نشان می‌دهد. یکی از موارد کاربرد این روش، در فصل ششم و در مسأله ۶ـ۱۰ مطرح شده است؛ با وجود این، برای حل مسائل R^2 ـ وقتی محاسبات برحسب مقادیر اصلی انجام شده است ـ می‌توان از روش بسیار ساده زیر استفاده کرد:

فرض می‌کنیم محاسبات برحسب مقادیر اصلی است،

$$\mathbf{y}'\mathbf{y}=\sum_{i}^{n} Y_i^2.$$

می‌دانیم

$$\mathbf{y}'\mathbf{A}\mathbf{y}=\Sigma (Y_i-\bar{Y})^2,$$

$$=\Sigma Y_i^2-n\bar{Y}^2,$$

$$=\mathbf{y}'\mathbf{y}-n\bar{Y}^2. \qquad (۵ـ۸۴)$$

کافی است از $\mathbf{y}'\mathbf{y}$ که برحسب مقادیر اصلی محاسبه شده است مقدار $n\bar{Y}^2$ را کم کنیم تا مجموع کل تغییرات (TSS) به دست آید. به همین ترتیب اگر $\hat{\mathbf{y}}'\hat{\mathbf{y}}$ را برحسب مشاهدات اصلی حساب کرده باشیم، می‌توان، $n\bar{Y}^2$ را از آن کم کرد تا تغییرات توضیح داده شده (ESS)، یعنی $\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{A}\mathbf{X}_2\hat{\boldsymbol{\beta}}_2$، نتیجه شود؛ زیرا

$$\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{A}\mathbf{X}_2\hat{\boldsymbol{\beta}}_2=\Sigma (\hat{Y}_i-\bar{\hat{Y}})^2.$$

از معادلهٔ ۳۰ـ۱، می‌دانیم که $\bar{\hat{Y}}=\bar{Y}$، بنابراین

$$\hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{A}\mathbf{X}_2\hat{\boldsymbol{\beta}}_2=\Sigma \hat{Y}_i^2+n\bar{Y}^2-2\bar{Y}\Sigma \hat{Y}_i,$$

$$=\Sigma \hat{Y}_i^2+n\bar{Y}^2,$$

$$=\hat{\mathbf{y}}'\mathbf{y}-n\bar{Y}^2. \qquad (۵ـ۸۵)$$

با استفاده از معادله‌های ۵ـ۸۴ و ۵ـ۸۵ مقدار R^2 به سهولت محاسبه می‌شود:

$$R^2=\frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}-n\bar{Y}^2}{\mathbf{y}'\mathbf{y}-n\bar{Y}^2}. \qquad (۵ـ۸۶)$$

فرمولهای دیگری نیز می‌توان برای R^2 به دست آورد. می‌دانیم $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$، در نتیجه

$$R^2 = \frac{\hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{X}\hat{\boldsymbol{\beta}} - n\bar{Y}^2}{\mathbf{y}'\mathbf{y} - n\bar{Y}^2}. \qquad (۵\text{-}۸۷)$$

با قراردادن $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ در فرمول ۵ـ۸۷ خواهیم داشت

$$R^2 = \frac{\hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{y} - n\bar{Y}^2}{\mathbf{y}'\mathbf{y} - n\bar{Y}^2}. \qquad (۵\text{-}۸۸)$$

یادآوری می‌شود که در فرمولهای ۵ـ۸۶، ۵ـ۸۷ و ۵ـ۸۸ مقادیر $\mathbf{y}$ و $\mathbf{X}$ برحسب مشاهدات اصلی است. همچنین باید توجه داشت که اگر بخواهیم R^2 را از فرمولی به دست آوریم که شامل $\mathbf{e}'\mathbf{e}$ است، دیگر ضرورتی ندارد که $n\bar{Y}^2$ را از $\mathbf{e}'\mathbf{e}$ کسر نماییم؛ زیرا $\mathbf{e}'\mathbf{e}$ همواره به صورت انحراف از میانگین است؛ چون میانگین مقادیر پسماندها صفر است.

مثال ۴ـ۵ مدلی را که در مثال ۲ـ۵ مطرح شد، یک بار دیگر ملاحظه می‌کنیم،

$$Y_t = \alpha_t + \beta X_t + \gamma Z_t + U_t$$

می‌دانیم بردار $\mathbf{y}$ و ماتریس $\mathbf{X}$ برابر است با:

$$\mathbf{y} = \begin{bmatrix} ۳ \\ ۱ \\ ۸ \\ ۳ \\ ۵ \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} ۱ & ۳ & ۵ \\ ۱ & ۱ & ۴ \\ ۱ & ۵ & ۶ \\ ۱ & ۲ & ۴ \\ ۱ & ۴ & ۶ \end{bmatrix}.$$

اولاً، با استفاده از ماتریس تبدیل $\mathbf{A}$، مشاهدات را برحسب انحراف از میانگین نوشته، سپس با به دست آوردن معادله‌های نرمال، پارامترهای β و γ را تخمین بزنید.

ثانیاً، R^2 را با استفاده از معادلهٔ ۵ـ۸۳ محاسبه کنید.

ثالثاً، R^2 را به روش ساده و با استفاده از مقادیر اصلی و به کمک یکی از فرمولهای ۵ـ۸۶ یا ۵ـ۸۷ یا ۵ـ۸۸ به دست آورید.

۱. ماتریس $\mathbf{A}$، تعریف شده در معادلهٔ ۵ـ۷۲ را از سمت چپ در $\mathbf{y}$ و $\mathbf{X}$ ضرب

می‌کنیم تا مشاهدات متغیر درون‌زا و متغیرهای برون‌زا برحسب انحراف از میانگین، به دست آید،

$$\Lambda y=\begin{bmatrix}-1\\-3\\4\\-1\\1\end{bmatrix}, \qquad \Lambda X_2=\begin{bmatrix}0 & 0\\-2 & -1\\2 & 1\\-1 & -1\\1 & 1\end{bmatrix}.$$

با توجه به معادلهٔ ۷۹ـ۵، می‌دانیم معادله‌های نرمال عبارت است از

$$(\Lambda X_2)'(\Lambda y)=(\Lambda X_2)'(\Lambda X_2)\hat{\beta}_2,$$

در نتیجه داریم

$$\begin{bmatrix}0 & -2 & 2 & -1 & 1\\0 & -1 & 1 & -1 & 1\end{bmatrix}\begin{bmatrix}-1\\-3\\4\\-1\\1\end{bmatrix}=\begin{bmatrix}0 & -2 & 2 & -1 & 1\\0 & -1 & 1 & -1 & 1\end{bmatrix}\begin{bmatrix}0 & 0\\-2 & -1\\2 & 1\\-1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}\hat{\beta}\\\hat{\gamma}\end{bmatrix},$$

$$\begin{bmatrix}16\\9\end{bmatrix}=\begin{bmatrix}10 & 6\\6 & 4\end{bmatrix}\begin{bmatrix}\hat{\beta}\\\hat{\gamma}\end{bmatrix}.$$

از حل این دستگاه معادلات نرمال، مقادیر $\hat{\beta}=2/5$ و $\hat{\gamma}=-1/5$ به دست می‌آید.

۲. برای محاسبه R^2 از فرمول ۸۳ـ۵، ابتدا فرمول مزبور را می‌نویسیم،

$$R^2=\frac{\hat{\beta}_2'X_2'\Lambda y}{y'\Lambda y}.$$

بردار Λy را در ترانهاد خود ضرب کرده و با توجه به تقارن و خودتوانی Λ داریم

$$y'\Lambda'\Lambda y=y'\Lambda y=TSS,$$

در نتیجه

$$TSS=\begin{bmatrix}-۱ & -۳ & ٤ & -۱ & ۱\end{bmatrix}\begin{bmatrix}-۱\\-۳\\٤\\-۱\\۱\end{bmatrix}=۲۸ .$$

برای به دست آوردن صورت کسر R^2، ابتدا $\mathbf{X}_1'\mathbf{\Lambda}\mathbf{y}$ را محاسبه می‌کنیم،

$$\mathbf{X}_1'\mathbf{\Lambda}\mathbf{y}=\mathbf{X}_1'\mathbf{\Lambda}'(\mathbf{\Lambda}\mathbf{y}),$$

$$=\begin{bmatrix}۰ & -۲ & ۲ & -۱ & ۱\\ ۰ & -۱ & ۱ & -۱ & ۱\end{bmatrix}\begin{bmatrix}-۱\\-۳\\٤\\-۱\\۱\end{bmatrix}=\begin{bmatrix}۱٦\\۹\end{bmatrix} .$$

صورت کسر R^2، یعنی تغییرات توضیح داده شده (ESS)، برابر است با

$$ESS=\hat{\beta}_2'(\mathbf{X}_1'\mathbf{\Lambda}\mathbf{y}),$$

$$=\begin{bmatrix}۲/۵ & -۱/۵\end{bmatrix}\begin{bmatrix}۱٦\\۹\end{bmatrix}=۲٦/۵ ,$$

در نتیجه

$$RSS=TSS-ESS=۲۸-۲٦/۵=۱/۵ .$$

بدین ترتیب:

$$R^2=\frac{۲٦/۵}{۲}=۰/۹٤٦٤ .$$

۳. برای محاسبه R^2، به کمک مقادیر اصلی مشاهدات، از معادلهٔ ۵-۸۸ استفاده می‌کنیم،

$$R^2=\frac{\hat{\beta}'\mathbf{X}'\mathbf{y}-n\bar{Y}^2}{\mathbf{y}'\mathbf{y}-n\bar{Y}^2} .$$

یادآوری می‌کنیم که در فرمول فوق $\mathbf{y}$ و $\mathbf{X}$ برحسب مشاهدات اصلی است. در مثال ۱ـ۵ دیدیم که

$$\bar{Y}= ٤ \quad و \quad \hat{\beta}' = [٤ \quad ٢/٥ \quad -١/٥] \quad و \quad \mathbf{X}'\mathbf{y}= \begin{bmatrix} ٢٠ \\ ٧٦ \\ ١٠٩ \end{bmatrix} .$$

همچنین در همین مثال می‌توان نشان داد که $\mathbf{y}'\mathbf{y}= ١٠٨$.

$$\hat{\beta}'\mathbf{X}'\mathbf{y}= [٤ \quad ٢/٥ \quad -١/٥] \begin{bmatrix} ٢٠ \\ ٧٦ \\ ١٠٩ \end{bmatrix} = ١٠٦/٥ ,$$

در نتیجه خواهیم داشت

$$R^٢= \frac{١٠٦/٥ - ٥(٤)^٢}{١٠٨ - ٥(٤)^٢} = \frac{٢٦/٥}{٢٨} = ٠/٩٤٦٤ ,$$

که دقیقاً برابر همان مقدار $R^٢$ است که با استفاده از مقادیر انحراف از میانگین محاسبه شده است.

مسائل فصل پنجم

۵ـ۱ مسأله شمارهٔ ۱ـ٤ را یک بار دیگر مطرح می‌کنیم،

$$C_t = \hat{\alpha}_1 + ۰/۹۲\, Y_t + e_{1t},$$

$$C_t = \hat{\alpha}_2 + ۰/۸٤\, C_{t-1} + e_{2t},$$

$$C_{t-1} = \hat{\alpha}_3 + ۰/۷۸\, Y_t + e_{3t},$$

$$C_t = \hat{\alpha}_4 + ۰/۵۵\, C_{t-1} + e_{4t},$$

که در آن $\hat{\alpha}_i$، تخمین جملهٔ ثابت و e_{it} مقادیر پسماند است. با استفاده از اطلاعات فوق، پارامترهای β_1 و β_2 را در مدل رگرسیون زیر فقط از راه ماتریسی تخمین بزنید،

$$C_t = \alpha + \beta_1 Y_t + \beta_2 C_{t-1} + U_t.$$

۵ـ۲ مدل رگرسیون

$$Y_i = \alpha + \beta X_i + \gamma Z_i + U_i,$$

مفروض است. با استفاده از یک نمونه، شامل ۲۳ مشاهده، محاسبات زیر را برحسب انحراف از میانگین انجام داده‌ایم،

$$\Sigma x_i^2 = ۱۲, \qquad \Sigma x_i z_i = ۸,$$

$$\Sigma z_i^2 = ۱۲, \qquad \Sigma y_i x_i = ۱۰,$$

$$\Sigma y_i^2 = ۱۰, \qquad \Sigma y_i z_i = ۸.$$

اولاً، پارامترهای β و γ را تخمین بزنید.
ثانیاً، R^2 و $\bar{R}^2$ را محاسبه کنید.

۵ـ۳ مدل زیرا را ملاحظه کنید،

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t .$$

با استفاده از یک نمونه شامل ۱۰ مشاهده روی Y_t، X_{1t} و X_{2t}، محاسبات زیر را انجام داده‌ایم،

$$\Sigma Y_t = ۲۰ , \quad \Sigma Y_t^2 = ۸۸/۲ , \quad \Sigma X_{1t} Y_t = ۵۹ ,$$

$$\Sigma X_{1t} = ۳۰ , \quad \Sigma X_{1t}^2 = ۹۲ , \quad \Sigma X_{2t} Y_t = ۸۸ ,$$

$$\Sigma X_{2t} = ۴۰ , \quad \Sigma X_{2t}^2 = ۱۶۳ , \quad \Sigma X_{1t} X_{2t} = ۱۱۹ .$$

اولاً، پارامترهای α، β_1 و β_2 را تخمین بزنید.
ثانیاً، R^2 و $\bar{R}^2$ را به دست آورید.
ثالثاً، مجموع مربعات پسماند را حساب کنید.

۵ـ۴ مدل

$$\ln Y_i = \beta_1 + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + U_i ,$$

مفروض است. نشان دهید که تخمین پارامترهای این مدل $(\hat{\beta}_2, \hat{\beta}_3)$ به ترتیب برابر با کشش Y_i نسبت به X_{2i} و X_{3i} است. آیا این کشش‌ها ثابت هستند؟

۵ـ۵ مدل رگرسیون زیر را مشاهده کنید،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 (X_{2t} - X_{3t}) + \beta_5 X_{5t} + U_t .$$

به نظر شما کدامیک از پارامترهای مدل فوق را می‌توان تخمین زد؟ توضیح دهید.

۵ـ۶ فرض کنید تقاضا برای کالای a، تابعی از قیمت آن کالا (P_{at})، قیمت k کالاهای دیگر $(P_{1t}, P_{2t}, \ldots, P_{kt})$، سطح عمومی قیمتها $(\bar{P}_t)$ و درآمد (Y_t) است؛ بنابراین

$$D_{at} = \alpha_0 + \beta_0 P_{at} + \beta_1 P_{1t} + \beta_2 P_{2t} + \cdots + \beta_k P_{kt} + \gamma \bar{P}_t + \lambda Y_t + U_t ,$$

که در آن D_{at}، مقدار تقاضا شده برای کالای a در دورهٔ t است. فرض کنید سطح

عمومی قیمتها به صورت بسیار ابتدایی چنین تعریف شده است،

$$\bar{P}_t = \frac{\sum_{i=1}^{m} P_{it}}{M},$$

که در آن $i = a, 1, 2, \ldots, k, k+1, k+2, \ldots, m$. به نظر شما چه مشکل اساسی در تخمین این مدل وجود دارد؟ چرا نمی‌توان تمام پارامترهای آن را تخمین زد؟

۵ـ۷ به نظر شما در مدل رگرسیون

$$Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + \alpha_3 (X_{1t} - X_{2t}) + \alpha_4 X_{1t} X_{2t} + U_t,$$

کدامیک از پارامترهای مدل را نمی‌توان تخمین زد؟

۵ـ۸ مدل رگرسیون زیر مفروض است:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + U_t.$$

اولاً، آیا تمام پارامترهای این مدل را می‌توان با روش حداقل مربعات معمولی تخمین زد؟

ثانیاً، با استفاده از مشاهدات زیر

Y_t	-۱	-۱	۲	۴	۵
X_t	۰	۱	۲	۵	۶

معادله‌های نرمال را بنویسید.

۵ـ۹ مدل رگرسیون

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + U_t.$$

مفروض است. این مدل راتخمین زده‌ایم و $\bar{R}^2$ برابر ۰/۸۶ به دست آمده است. متغیر توضیحی X_{4t} را به مدل اضافه کرده و دوباره آن را تخمین می‌زنیم. این باره $\bar{R}^2$ برابر ۰/۸۲ محاسبه شده است. نشان دهید که حتماً اشتباهی در محاسبات صورت گرفته است، مگر اینکه حجم نمونه ۸ یا کمتر از آن باشد.

۵.۱۰ می‌خواهیم مدل رگرسیون زیر را تخمین بزنیم،

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t , \qquad (1)$$

که در آن جملهٔ ثابت وجود ندارد. فرض کنید کامپیوتری که در اختیار داریم، چنان برنامه‌ریزی شده است که هر مجموعه مشاهداتی که به آن بدهیم، یک جملهٔ ثابت را نیز در نظر می‌گیرد؛ یعنی کامپیوتر همواره مدل زیر را تخمین می‌زند،

$$Y_t = \beta_1 + \beta_2 X_{1t} + \beta_3 X_{2t} + \cdots + \beta_k X_{kt} + U_t . \qquad (2)$$

نشان دهید که اگر به ازای هر زوج مشاهده که به کامپیوتر می‌دهیم، یعنی (Y_t, X_{it})، $t = 1, 2, \ldots, n$ و $i = 2, 3, \ldots, k$، یک زوج مشاهده دیگر با علامت مخالف، یعنی $(-Y_t, -X_{it})$، نیز وارد کامپیوتر کنیم، آنگاه

اولاً، تخمینهای به دست آمده از معادلهٔ (۲) دقیقاً برابر تخمینهای مطلوب ما برای معادلهٔ (۱) خواهد بود.

ثانیاً، مجموع مربعات پسماند برای مدل (۲) دقیقاً دو برابر مجموع مربعات پسماند برای مدل (۱) است.

ثالثاً، نسبت انحراف معیار تخمین (SEE) در مدل (۲) نسبت به مدل (۱) برابر است با:

$$\frac{\text{SEE}\ (2)}{\text{SEE}\ (1)} = \frac{\sqrt{2\,(n-k+1)}}{\sqrt{(2n-k)}} .$$

حل مسائل فصل پنجم

۵ـ۱ معادلهٔ ۵ـ۲۰ یعنی $(\mathbf{X}'\mathbf{X})\beta = \mathbf{X}'\mathbf{y}$ را ـ که همان سیستم معادله‌های نرمال است ـ برای این مسأله می‌نویسیم،

$$\begin{bmatrix} \Sigma y_t^2 & \Sigma y_t c_{t-1} \\ \Sigma c_{t-1} y_t & \Sigma c_{t-1}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \Sigma c_t y_t \\ \Sigma c_{t-1} c_t \end{bmatrix}.$$

با استفاده از اطلاعات موجود می‌دانیم

$$\frac{\Sigma c_t y_t}{\Sigma y_t^2} = ۰/۹۲ \quad \rightarrow \quad \Sigma c_t y_t = ۰/۹۲ \Sigma y_t^2, \qquad (۱)$$

$$\frac{\Sigma c_t c_{t-1}}{\Sigma c_{t-1}^2} = ۰/۸۴ \quad \rightarrow \quad \Sigma c_t c_{t-1} = ۰/۸۴ \Sigma c_{t-1}^2, \qquad (۲)$$

$$\frac{\Sigma c_{t-1} y_t}{\Sigma y_t^2} = ۰/۷۸ \quad \rightarrow \quad \Sigma c_{t-1} y_t = ۰/۷۸ \Sigma y_t^2, \qquad (۳)$$

$$\frac{\Sigma y_t c_{t-1}}{\Sigma c_{t-1}^2} = ۰/۵۵ \quad \rightarrow \quad \Sigma y_t c_{t-1} = ۰/۵۵ \Sigma c_{t-1}^2. \qquad (۴)$$

از معادلهٔ (۴) داریم

$$\Sigma c_{t-1}^2 = \frac{\Sigma y_t c_{t-1}}{۰/۵۵}.$$

معادلهٔ (۳) را در معادلهٔ فوق قرار می‌دهیم،

$$\Sigma c_{t-1}^2 = \frac{۰/۷۸ \Sigma y_t^2}{۰/۵۵}. \qquad (۵)$$

بدین ترتیب، تمام عناصر ماتریس $\mathbf{X}'\mathbf{X}$ را برحسب Σy_t^2 به‌دست آورده‌ایم. با معادله‌های

(۳) و (۵) این ماتریس را می‌سازیم،

$$\begin{bmatrix} \Sigma y_t^2 & \Sigma y_t c_{t-1} \\ \Sigma c_{t-1} y_t & \Sigma c_{t-1}^2 \end{bmatrix} = \begin{bmatrix} \Sigma y_t^2 & ۰/۷۸ \Sigma y_t^2 \\ ۰/۷۸ \Sigma y_t^2 & \frac{۰/۷۸ \Sigma y_t^2}{۰/۵۵} \end{bmatrix}.$$

به ترتیبی مشابه، بردار $\mathbf{X}'\mathbf{y}$ را برحسب Σy_t^2 می‌نویسیم. در معادلهٔ (۱) عنصر اول این بردار را برحسب Σy_t^2 به دست آورده‌ایم. باید عنصر دوم آن، یعنی $\Sigma c_{t-1} c_t$، را برحسب Σy_t^2 بنویسیم. کافی است معادلهٔ (۵) را در (۲) قرار دهیم،

$$\Sigma c_t c_{t-1} = ۰/۸۴ \left(\frac{۰/۷۸ \Sigma y_t^2}{۰/۵۵}\right) = ۱/۱۹۱ \Sigma y_t^2.$$

بنابراین ماتریس $\mathbf{X}'\mathbf{y}$ برحسب Σy_t^2 به ترتیب زیر خواهد بود،

$$\begin{bmatrix} \Sigma c_t y_t \\ \Sigma c_{t-1} c_t \end{bmatrix} = \begin{bmatrix} ۰/۹۲ \Sigma y_t^2 \\ ۱/۱۹۱ \Sigma y_t^2 \end{bmatrix},$$

معادله‌های نرمال ۲۰ـ۵ را نوشته، بعد از حذف Σy_t^2 خواهیم داشت

$$\begin{bmatrix} ۱ & ۰/۷۸ \\ ۰/۷۸ & \frac{۰/۷۸}{۰/۵۵} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} ۰/۹۲ \\ ۱/۱۹۱ \end{bmatrix},$$

که با حل این دستگاه در معادلهٔ دو مجهولی داریم $\hat{\beta}_1 = ۰/۴۶۴$ و $\hat{\beta}_2 = ۰/۵۸۵$.

۲ـ۵ الف) ابتدا ماتریس $(\mathbf{X}'\mathbf{X})$ را تشکیل می‌دهیم،

$$(\mathbf{X}'\mathbf{X}) = \begin{bmatrix} \Sigma x_i^2 & \Sigma x_i z_i \\ \Sigma z_i x_i & \Sigma z_i^2 \end{bmatrix} = \begin{bmatrix} ۱۲ & ۸ \\ ۸ & ۱۲ \end{bmatrix}.$$

$(\mathbf{X}'\mathbf{X})$ را معکوس می‌کنیم. داریم

$$(\mathbf{X}'\mathbf{X})^{-1}=\begin{bmatrix} ۰/۱۵ & -۰/۱ \\ -۰/۱ & ۰/۱۵ \end{bmatrix} .$$

بردار $\mathbf{X}'\mathbf{y}$ را برای این مسأله محاسبه می‌کنیم،

$$\mathbf{X}'\mathbf{y}=\begin{bmatrix} \Sigma x_i y_i \\ \Sigma z_i y_i \end{bmatrix} = \begin{bmatrix} ۱۰ \\ ۸ \end{bmatrix} .$$

بردار $\hat{\boldsymbol{\beta}}$ به راحتی قابل محاسبه است،

$$\hat{\boldsymbol{\beta}}=\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \Sigma x_i^۲ & \Sigma x_i z_i \\ \Sigma z_i x_i & \Sigma z_i^۲ \end{bmatrix} \begin{bmatrix} \Sigma x_i y_i \\ \Sigma z_i y_i \end{bmatrix} ,$$

$$= \begin{bmatrix} ۰/۱۵ & -۰/۱ \\ -۰/۱ & ۰/۱۵ \end{bmatrix} \begin{bmatrix} ۱۰ \\ ۸ \end{bmatrix} = \begin{bmatrix} ۰/۷ \\ ۰/۲ \end{bmatrix} .$$

بنابراین $\hat{\beta}= ۰/۷$ و $\hat{\gamma}= ۰/۲$.

ب) می‌دانیم

$$R^۲= \frac{\hat{\beta}'\mathbf{X}'\mathbf{y}}{\mathbf{y}'\mathbf{y}} .$$

صورت کسر $R^۲$ برای این مسأله عبارت است از

$$\hat{\beta}'\mathbf{X}'\mathbf{y}= [\hat{\beta} \quad \hat{\gamma}] \begin{bmatrix} \Sigma x_i y_i \\ \Sigma z_i y_i \end{bmatrix} = \hat{\beta} \Sigma x_i y_i + \hat{\gamma} \Sigma z_i y_i ,$$

$$= ۰/۷ (۱۰) + ۰/۲ (۸)= ۸/۶ .$$

مخرج کسر $R^۲$ برای این مسأله برابر است با

$$\mathbf{y}'\mathbf{y}= \Sigma y_i^۲= ۱۰ ,$$

در نتیجه R^2 به راحتی محاسبه می‌شود،

$$R^2 = \frac{۸/۶}{۱۰} = ۰/۸۶$$

برای محاسبهٔ $\bar{R}^2$ از فرمول زیر استفاده می‌کنیم،

$$\bar{R}^2 = ۱ - (\frac{n-۱}{n-k})\ (۱-R^2)\ ,$$

$$= ۱ - (\frac{۲۳-۱}{۲۳-۳})\ (۱ - ۰/۸۶) = ۰/۸۴۶\ .$$

۵.۳ الف) با توجه به اینکه کمیتهای محاسبه شده برحسب مشاهدات اصلی است، باید آنها را به انحراف از میانگین تبدیل نماییم. کمیتهای لازم را به شرح زیر محاسبه می‌کنیم.

$$\bar{Y} = ۲ \quad , \quad \bar{X}_۱ = ۳ \quad , \quad \bar{X}_۲ = ۴\ .$$

$$\Sigma x_{۱t}^2 = \Sigma X_{۱t}^2 - n\bar{X}_۱^2 = ۹۲ - ۱۰\ (۳)^2 = ۲\ ,$$

$$\Sigma x_{۲t}^2 = \Sigma X_{۲t}^2 - n\bar{X}_۲^2 = ۱۶۳ - ۱۰\ (۴)^2 = ۳\ ,$$

$$\Sigma x_{۱t} x_{۲t} = \Sigma X_{۱t} x_{۲t} - n\bar{X}_۱ \bar{X}_۲ = ۱۱۹ - ۱۰\ (۳)(۴) = -۱\ ,$$

$$\Sigma x_{۱t} y_t = \Sigma X_{۱t} Y_t - n\bar{X}_۱ \bar{Y} = ۵۹ - ۱۰\ (۳)(۲) = -۱\ ,$$

$$\Sigma x_{۲t} y_t = \Sigma X_{۲t} Y_t - n\bar{X}_۲ \bar{Y} = ۸۸ - ۱۰\ (۴)(۲) = ۸\ ,$$

$$\Sigma y_t^2 = \Sigma Y_t^2 - n\bar{Y}^2 = ۸۸/۲ - ۱۰\ (۲)^2 = ۴۸/۲\ .$$

ماتریس $(\mathbf{X}'\mathbf{X})$ و بردار $\mathbf{X}'\mathbf{y}$ را محاسبه می‌کنیم. خواهیم داشت

$$(\mathbf{X}'\mathbf{X}) = \begin{bmatrix} \Sigma x_{۱t}^2 & \Sigma x_{۱t} x_{۲t} \\ \Sigma x_{۲t} x_{۱t} & \Sigma x_{۲t}^2 \end{bmatrix} = \begin{bmatrix} ۲ & -۱ \\ -۱ & ۳ \end{bmatrix}\ ,$$

$$(\mathbf{X}'\mathbf{y}) = \begin{bmatrix} \Sigma x_{1t} y_t \\ \Sigma x_{2t} y_t \end{bmatrix} = \begin{bmatrix} -1 \\ 8 \end{bmatrix} .$$

$\hat{\beta}$ به صورت زیر محاسبه می‌شود،

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} ,$$

یعنی $\hat{\beta}_1 = 1$ و $\hat{\beta}_2 = 2$. برای تخمین α به صورت زیر عمل می‌کنیم،

$$\bar{Y} = \hat{\alpha} + \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 ,$$

$$2 = \hat{\alpha} + (1)(3) + (3)(4) ,$$

$$\hat{\alpha} = -13$$

بدین ترتیب تخمین مدل عبارت است از

$$\hat{Y}_t = -13 + X_{1t} + 3X_{2t} .$$

پ) برای محاسبهٔ R^2، ابتدا فرمول ۴۰ـ۵ را می‌نویسیم،

$$R^2 = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}}{\mathbf{y}'\mathbf{y}} = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{\mathbf{y}'\mathbf{y}} .$$

صورت کسر R^2 برای این مسأله عبارت است از

$$\hat{\boldsymbol{\beta}}\mathbf{X}'\mathbf{y} = [\hat{\beta}_1 \quad \hat{\beta}_2] \begin{bmatrix} \Sigma x_{1t} y_t \\ \Sigma x_{2t} y_t \end{bmatrix} ,$$

$$= [1 \quad 3] \begin{bmatrix} -1 \\ 8 \end{bmatrix} = 23 .$$

مخرج کسر R^2 برابر است با

$$y' y = \Sigma y_i^2 = ۴۸/۲ ,$$

بدین ترتیب داریم

$$R^2 = \frac{۲۳}{۴۸/۲} = ۰/۴۷۷۱ .$$

می‌دانیم

$$\bar{R}^2 = ۱ - (\frac{n-۱}{n-k}) (۱-R^2) ,$$

$$= ۱ - (\frac{۱۰-۱}{۱۰-۳}) (۱ - ۰/۴۷۷۱) ,$$

$$= ۱ - ۰/۶۷۲۳ = ۰/۳۲۷۷۷ .$$

ج) برای به دست آوردن مجموع مربعات پسماند (Σe_i^2) از فرمول ۴۸ـ۵ استفاده می‌کنیم.

$$R^2 = ۱ - \frac{e' e}{y' y} ,$$

یا

$$e' e = (۱-R^2) y' y ,$$

$$= (۱ - ۰/۴۷۷۱) (۴۸/۲) = ۲۵/۲۰ .$$

۵ـ۴ می‌دانیم:

$$\beta_2 = \frac{\partial \ln Y_i}{\partial \ln X_{2i}} = \frac{(\frac{۱}{Y_i}) d Y_i}{(\frac{۱}{X_{2i}}) d X_{2i}} = \frac{d Y_i}{d X_{2i}} \cdot \frac{X_{2i}}{Y_i} .$$

$$\beta_3 = \frac{\partial \ln Y_i}{\partial \ln X_{3i}} = \frac{(\frac{۱}{Y_i}) d Y_i}{(\frac{۱}{X_{3i}}) d X_{3i}} = \frac{d Y_i}{d X_{3i}} \cdot \frac{X_{3i}}{Y_i} .$$

برای اثبات رابطه‌های فوق به توضیحات موجود در پاورقی مبحث مدلهای

خطی ـ لگاریتمی، مندرج در فصل سوم قسمت ۶ـ۳ مراجعه کنید. بنابراین، $\hat{\beta}_2$ و $\hat{\beta}_3$ به ترتیب تخمین کشش Y_i نسبت به X_{2i} و X_{3i} است. همچنین این کششها ثابت نیز هستند زیرا می‌دانیم تخمین پارامترهای یک مدل رگرسیون با روش حداقل مربعات معمولی مقادیر ثابتی را نتیجه می‌دهد.

۵ـ۵ مدل را به صورت زیر می‌نویسیم،

$$Y_t = \beta_1 + (\beta_2 + \beta_4) X_{2t} + (\beta_3 - \beta_4) X_{3t} + \beta_5 X_{5t} + U_t .$$

با داشتن مشاهدات Y_t ، X_{2t} ، X_{3t} و X_{5t} می‌توان مدل را به صورت زیر تخمین زد،

$$\hat{Y}_t = \hat{\beta}_1 + \hat{\gamma}\, X_{2t} + \hat{\lambda}\, X_{3t} + \hat{\beta}_5\, X_{5t} .$$

ملاحظه می‌شود برای پارامترهای β_1 و β_5 می‌توان به تخمینهای مستقلی رسید، اما نمی‌توان $\hat{\beta}_2$، $\hat{\beta}_4$ و $\hat{\beta}_3$ را به دست آورد. آنچه می‌توان تخمین زد، عبارت است از

$$\hat{\gamma} = \hat{\beta}_2 + \hat{\beta}_4 \qquad , \qquad \hat{\lambda} = \hat{\beta}_3 - \hat{\beta}_4 .$$

بدیهی است که دو معادلهٔ سه مجهولی داریم. اگر بتوان یک رابطهٔ دیگر بین $\hat{\beta}_2$، $\hat{\beta}_3$ و $\hat{\beta}_4$ به دست آورد، تخمینهای مستقل از این پارامترها ممکن خواهد بود. در این گونه موارد معمولاً باید از نظریه‌های اقتصادی کمک گرفت تا بتوان به چنین رابطه‌هایی رسید.

۶ـ۵ با توجه به صورت مسأله می‌دانیم

$$\bar{P}_t = \frac{1}{k} [\, P_{at} + P_{1t} + P_{2t} + \cdots + P_{kt} + P_{(k-1)t} + \cdots + P_{mt}] .$$

بنابراین $\bar{P}_t$ با متغیرهای $P_{at}, P_{1t}, \ldots, P_{kt}$، همخطی کامل خواهد داشت و به همین دلیل نمی‌توان پارامترهای متغیرهایی را که همخطی دارند تخمین زد. برای تبیین بیشتر این نکته می‌توانیم $\bar{P}_t$ را در مدل مفروض جایگزین کنیم. بعد از ساده کردن خواهیم داشت

$$D_{at} = \alpha_0 + (\beta_0 + \frac{\gamma}{M}) P_{at} + (\beta_1 + \frac{\gamma}{M}) P_{1t} + \cdots + (\beta_k + \frac{\gamma}{M} P_{kt}) + \gamma^* \bar{P}^*_t + \lambda\, Y_t + U_t ,$$

که در آن

$$\bar{P}_t^* = \frac{1}{k}\left[P_{(k+1)t} + P_{(k+2)t} + \cdots + P_{mt}\right].$$

ملاحظه می‌شود که فقط می‌توان α_0، λ و γ^* را تخمین زد. البته γ^* مطلوب ما نیست؛ بنابراین از $k+4$ پارامتری که در مدل اولیه وجود دارد، فقط میتوان به تخمین دو پارامتر رسید.

۵ـ۷ پارامترهایی را که نمی‌توان با روش حداقل مربعات معمولی تخمین زد، عبارت است از α_1، α_2 و α_3، زیرا سومین متغیر توضیحی مدل مفروض، یعنی $(X_{1t}-X_{2t})$ یک ترکیب خطی از دو متغیر اول و دوم، یعنی X_{1t} و X_{2t} است؛ به عبارت دیگر، بین این سه متغیر توضیحی همخطی کامل برقرار است. مدل مفروض را می‌توان به‌صورت زیر نوشت

$$Y_t = \alpha_0 + (\alpha_1 + \alpha_2) X_{1t} + (\alpha_2 - \alpha_3) X_{2t} + \alpha_4 X_{1t} X_{2t} + U_t.$$

بنابراین فقط می‌توان به تخمینهای مستقلی از α_0 و α_4 رسید. توجه داریم که چهارمین متغیر توضیحی مدل مفروض، یعنی $(X_{1t} X_{2t})$، یک ترکیب «غیرخطی» از متغیرهای اول و دوم است و به همین دلیل مشکلی را در امر تخمین ایجاد نکرده است و α_4 را می‌توان به راحتی تخمین زد.

۵ـ۸ الف) هیچ مشکل نظری در تخمین پارامترهای این مدل وجود ندارد؛ زیرا X_t و X_t^2 یک رابطهٔ غیرخطی با یکدیگر دارند.

ب) می‌دانیم برای نوشتن معادله‌های نرمال، باید از Σe_t^2 نسبت به β_0، β_1 و β_2 مشتق گرفته و آنها را مساوی صفر قرار دهیم. اما قبلاً در فصل چهارم و در ادامهٔ بحث معادله‌های نرمال ۴ـ۵ دیدیم که یک روش ساده برای نوشتن معادله‌های نرمال وجود دارد. با استفاده از این روش داریم

$$\Sigma Y_t = n\hat{\beta}_0 + \hat{\beta}_1 \Sigma X_t + \hat{\beta}_2 \Sigma X_t^2,$$

$$\Sigma X_t Y_t = \hat{\beta}_0 \Sigma X_t + \hat{\beta}_1 \Sigma X_t^2 + \hat{\beta}_2 \Sigma X_t^3,$$

$$\Sigma X_t^2 Y_t = \hat{\beta}_0 \Sigma X_t^2 + \hat{\beta}_1 \Sigma X_t^3 + \hat{\beta}_2 \Sigma X_t^4,$$

با استفاده از مشاهدات داده‌شده، می‌توان معادله‌های نرمال فوق را به صورت زیر نوشت،

$$۹= ۵\hat{\beta}_۰ + ۱۴\hat{\beta}_۱ + ۶۶\hat{\beta}_۲ ,$$

$$۵۳= ۱۴\hat{\beta}_۰ + ۶۶\hat{\beta}_۱ + ۳۵۰\hat{\beta}_۲ ,$$

$$۲۸۷= ۶۶\hat{\beta}_۰ + ۳۵۰\hat{\beta}_۱ + ۱۹۳۸\hat{\beta}_۲ .$$

۵ـ۹ می‌دانیم اگر آمارهٔ t، متعلق به تخمین یک پارامتر از یک کمتر باشد، ورود آن متغیر توضیحی در مدل می‌تواند $\bar{R}^۲$ را کاهش دهد. در این مسأله ملاحظه می‌شود که بعد از اضافه کردن $X_{۴i}$ مقدار $\bar{R}^۲$ کاهش یافته است؛ بنابراین اگر t مربوط به $X_{۴i}$ کمتر از یک باشد، این نتیجه قابل قبول است. اما با اطلاعات بسیار محدودی که در مسأله داده شده است، نمی‌توان این نکته را ارزیابی کرد که واقعاً t از یک کمتر است. باید به دنبال معیار دیگری برویم. مسلماً $\bar{R}^۲$ در مدل جدید کاهش یافته، بنابراین سؤال این است که چه شرطی لازمهٔ تحقق چنین نتیجه‌ای است؟ با استفاده از معادلهٔ ۶۳ـ۵ یعنی

$$\bar{R}^۲= ۱-(\frac{n-۱}{n-k})(۱-R^۲) ,$$

یا مستقیماً از معادلهٔ ۶۶ـ۵ می‌دانیم

$$\frac{۱-\bar{R}^۲}{n-۱} = \frac{۱-R^۲}{n-k} ,$$

یا

$$(۱-R^۲)= \frac{(n-k)}{(n-۱)}(۱-\bar{R}^۲) . \qquad (۱)$$

واضح است که با ورود یک متغیر توضیحی جدید به مدل رگرسیون، مقدار $R^۲$ اضافه شده یا حداقل ثابت می‌ماند؛ بنابراین باید انتظار داشته باشیم که در مدل رگرسیون جدید، یعنی بعد از اضافه کردن $X_{۴i}$، مقدار $(۱-R^۲)$ از حالت قبل کمتر بشود. در نتیجه، همواره باید رابطهٔ زیر برقرار باشد،

$$(۱-R^۲) \text{ در مدل جدید} \leq (۱-R^۲) \text{ در مدل قدیم} . \qquad (۲)$$

رابطهٔ (۱) را در (۲) قرار می‌دهیم،

$$\frac{(n-k)}{(n-1)}(1-\bar{R}^2) \text{ در مدل قدیم} \leq \frac{(n-k)}{(n-1)}(1-\bar{R}^2) \text{ در مدل جدید},$$

و با حذف $(n-1)$ خواهیم داشت

$$(n-k)(1-\bar{R}^2) \text{ در مدل قدیم} \leq (n-k)(1-\bar{R}^2) \text{ در مدل جدید}. \quad (۳)$$

می‌دانیم تعداد پارامترهایی که تخمین زده می‌شود، یعنی k، در مدلهای جدید و قدیم به ترتیب برابر با ۵ و ۴ است. برای اینکه محاسبات ارائه شده در مسأله صحیح باشد، باید در رابطهٔ (۳) صدق کند؛ بنابراین باید داشته باشیم

$$(n-۵)(۱-۰/۸۶) \leq (n-۴)(۱-۰/۸۲),$$

$$۰/۱۸(n-۵) \leq ۰/۱۴(n-۴),$$

$$n \leq ۸/۵.$$

یعنی شرط صحت محاسبات ارائه شده این است که حجم نمونه ۸ یا کمتر از آن باشد.

۵-۱۰ الف) وقتی به ازای هر زوج‌مشاهده (Y_i, X_{ij}) یک زوج‌مشاهده با علامت مخالف، یعنی $(-Y_i, -X_{ij})$ وارد کامپیوتر کنیم، ماتریس $\mathbf{X}$ و بردار $\mathbf{y}$ به صورت زیر خواهد بود،

$$\mathbf{y}^* = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \\ -Y_1 \\ -Y_2 \\ \vdots \\ -Y_n \end{bmatrix}, \quad \mathbf{X}^* = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & & & & \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \\ 1 & -X_{11} & -X_{21} & \cdots & -X_{kn} \\ 1 & -X_{12} & -X_{22} & \cdots & -X_{kn} \\ \vdots & & & & \\ 1 & -X_{1n} & -X_{2n} & \cdots & -X_{kn} \end{bmatrix},$$

که $\mathbf{y}^* \to ۲n\times ۱$ و $\mathbf{X}^* \to ۲n\times k$. درواقع مدل (۲) با $\mathbf{y}^*$ و $\mathbf{X}^*$ تخمین زده می‌شود. باید نشان دهیم مقادیر $\hat{\beta}_۱, \hat{\beta}_۲, \ldots, \hat{\beta}_k$ که با استفاده از $\mathbf{y}^*$ و $\mathbf{X}^*$ به دست می‌آید دقیقاً برابر تخمینهایی است که از مدل (۱) و به کمک $\mathbf{y}$ و $\mathbf{X}$ حاصل می‌شود. برای این منظور $\mathbf{X}^{*\prime}\mathbf{X}^*$ و $\mathbf{X}^{*\prime}\mathbf{y}^*$ را تشکیل می‌دهیم. ابتدا $\mathbf{y}^*$ و $\mathbf{X}^*$ را به صورت زیر تعریف می‌کنیم،

$$\mathbf{y}^* = \begin{bmatrix} \mathbf{y} \\ -\mathbf{y} \end{bmatrix} \quad , \quad \mathbf{X}^* = \begin{bmatrix} \mathbf{i} & \mathbf{X} \\ \mathbf{i} & -\mathbf{X} \end{bmatrix} \quad ,$$

که در آن $\mathbf{i}$ یک بردار ستونی $(n\times ۱)$ و شامل عدد یک است و $\mathbf{X} \to n\times(k-۱)$؛ بنابراین

$$\mathbf{X}^{*\prime}\mathbf{X}^* = \begin{bmatrix} \mathbf{i}' & \mathbf{i} \\ \mathbf{X}' & -\mathbf{X}' \end{bmatrix} \begin{bmatrix} \mathbf{i} & \mathbf{X} \\ \mathbf{i} & -\mathbf{X} \end{bmatrix} = \begin{bmatrix} ۲n & ۰ \\ ۰ & ۲\mathbf{X}'\mathbf{X} \end{bmatrix} ,$$

که در آن n، تعداد مشاهدات اولیه در مدل (۱) است. به همین ترتیب می‌توان نوشت

$$\mathbf{X}^{*\prime}\mathbf{y}^* = \begin{bmatrix} \mathbf{i}' & \mathbf{i} \\ \mathbf{X}' & -\mathbf{X}' \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ -\mathbf{y} \end{bmatrix} = \begin{bmatrix} ۰ \\ ۲\mathbf{X}'\mathbf{y} \end{bmatrix} .$$

اگر تخمین $\boldsymbol{\beta}$ را در مدل (۲)، $\hat{\boldsymbol{\beta}}^*$ بنامیم، خواهیم داشت

$$\hat{\boldsymbol{\beta}}^* = (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-۱}(\mathbf{X}^{*\prime}\mathbf{y}) ,$$

$$= \begin{bmatrix} ۲n & ۰ \\ ۰ & ۲\mathbf{X}'\mathbf{X} \end{bmatrix}^{-۱} \begin{bmatrix} ۰ \\ ۲\mathbf{X}'\mathbf{y} \end{bmatrix} = \frac{۱}{۲} \begin{bmatrix} \frac{۱}{n} & ۰ \\ ۰ & (\mathbf{X}'\mathbf{X})^{-۱} \end{bmatrix} \begin{bmatrix} ۰ \\ ۲\mathbf{X}'\mathbf{y} \end{bmatrix} .$$

بنابراین داریم:

$$\hat{\boldsymbol{\beta}}^* = \begin{bmatrix} ۰ \\ (\mathbf{X}'\mathbf{X})^{-۱}\mathbf{X}'\mathbf{y} \end{bmatrix} . \qquad (۳)$$

اما اگر بخواهیم مدل (۱) را بدون جملهٔ ثابت تخمین بزنیم، باید از فرمول زیر استفاده کنیم،

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-۱}\mathbf{X}'\mathbf{y} \qquad (۴)$$

ملاحظه می‌شود که $\hat{\beta}^*$ در معادلهٔ (۳) برای جملهٔ ثابت مدل (۲) تخمینی برابر صفر و برای پارامترهای β_1، β_2 تا β_k تخمینی برابر $(X'X)^{-1}X'y$ نتیجه می‌دهد که دقیقاً برابر $\hat{\beta}$ در معادلهٔ (٤) است. بدین ترتیب توانستیم از مدلی که جملهٔ ثابت دارد، تخمین پارامترهای مدل فاقد جملهٔ ثابت را، با استفاده از مشاهدات اصلی، به دست آوریم.

$$\hat{\beta}^* = \begin{bmatrix} 0 \\ \hat{\beta} \end{bmatrix} . \qquad (۵)$$

ب) بردار جمله‌های پسماند برای مدل (۲) عبارت است از

$$e^* = y^* - \hat{y}^* = y^* - X^*\hat{\beta}^* .$$

با استفاده از معادلهٔ (۵)، داریم:

$$e^* = \begin{bmatrix} y \\ -y \end{bmatrix} - \begin{bmatrix} i & X \\ i & -X \end{bmatrix} \begin{bmatrix} 0 \\ \hat{\beta} \end{bmatrix} ,$$

$$= \begin{bmatrix} (y - X\hat{\beta}) \\ -(y - X\hat{\beta}) \end{bmatrix} = \begin{bmatrix} e \\ -e \end{bmatrix} .$$

با داشتن e^* می‌توان مجموع مربعات پسماند برای مدل (۲) را حساب کرد:

$$e^{*\prime}e^* = [e' \; -e'] \begin{bmatrix} e \\ -e \end{bmatrix} = e'e + e'e = 2e'e , \qquad (٦)$$

که در آن $e'e$، مجموع مربعات پسماند برای مدل (۱) است.

ج) با توجه به معادلهٔ ۱ـ۴۵ می‌دانیم که انحراف معیار تخمین در یک مدل رگرسیون ساده شامل دو پارامتر، عبارت است از

$$SEE = \sqrt{\frac{\Sigma e_i^2}{n-2}} .$$

درجات آزادی برای Σe_t^2 در حالتی که k پارامتر داریم، برابر است با $(n-k)$، بنابراین خطای معیار تخمین برای یک رگرسیون چند متغیره با k پارامتر عبارت است از

$$SEE = \sqrt{\frac{e'e}{n-k}}.$$

در مدل (۱) مقدار پارامترها $(k-1)$ است؛ بنابراین

$$SEE\ (1) = \sqrt{\frac{e'e}{n-(k-1)}}, \qquad (۷)$$

و برای مدل (۲) داریم

$$SEE\ (2) = \sqrt{\frac{e^{*\prime}e^{*}}{2n-k}}.$$

با استفاده از معادلهٔ (۶)، می‌توان خطای معیار تخمین (۲) (SEE) را به صورت زیر نوشت،

$$SEE\ (2) = \sqrt{\frac{2e'e}{2n-k}}, \qquad (۸)$$

در نتیجه، خواهیم داشت

$$\frac{SEE\ (2)}{SEE\ (1)} = \sqrt{\frac{2e'e\ (n-k+1)}{(2n-k)\ e'e}} = \frac{\sqrt{2\ (n-k+1)}}{\sqrt{(2n-k)}}.$$

فصل ششم

آزمون پارامترهای مدل رگرسیون خطی چندمتغیره

۶ـ۱ مقدمه

بعد از تخمین پارامترهای یک مدل رگرسیون چندمتغیره در فصل پنجم، در این فصل به آزمون این پارامترها می‌پردازیم. بردار پارامترهای β شامل k پارامتر را در نظر می‌گیریم. $\hat{\beta}$ را با روش حداقل مربعات معمولی و به عنوان تخمین بردار β به دست آورده‌ایم. سؤال این است که چگونه به کمک $\hat{\beta}$ می‌توان دربارهٔ β استنباط آماری داشت؟

با توجه به اینکه $\hat{\beta}$ یک بردار تصادفی است باید ابتدا تابع توزیع احتمال و میانگین و واریانس آن را مطالعه کنیم. این نکات در قسمت ۶ـ۲ با نام خصوصیات آماری $\hat{\beta}$ بررسی شده است. در این قسمت قضیهٔ گاس ـ مارکف به طور خلاصه و نیز خصوصیت «بهترین تخمین زنندهٔ نااریب خطی» برای $\hat{\beta}$ مطرح می‌شود. همچنین نشان داده‌ایم که اگر مجموع مربعات پسماند را بر درجات آزادی آن تقسیم کنیم، یک تخمین نااریب از واریانس جملهٔ اختلال خواهیم داشت. آزمون واریانس جملهٔ اختلال نیز در همین قسمت مطرح شده است.

با شناخت خصوصیات آماری $\hat{\beta}$ به بررسی آزمون β می‌پردازیم. قسمت ۶ـ۳ آزمون هریک از پارامترهاست. آزمونهای $\beta_i = a$ که a عدد ثابت و معلومی است و نیز آزمون $\beta_i = 0$ در این قسمت ارزیابی خواهد شد. فرضیهٔ صفر بودن همزمان زیرمجموعه‌ای از پارامترهای یک مدل رگرسیون چندمتغیره اهمیت فراوان دارد و کاربردهای متعددی در اقتصادسنجی کاربردی دارد. نکات آزمون این فرضیه که $\beta_i = \beta_{i+1} = \cdots = \beta_j = 0$، در همین قسمت مطرح می‌شود. آزمون معنی‌دار بودن کل مدل

رگرسیون که درواقع آزمون اعتبار مدل است در قسمت ۶.۴ بررسی می‌شود. در این قسمت فرضیه $\beta_2 = \beta_3 = \cdots = \beta_k = 0$ را به کمک آنالیز واریانس، آزمون خواهیم کرد. ملاحظه می‌شود که در صورت صحت فرضیهٔ H_0، مدل رگرسیون چندمتغیره مفروض، به مدل $Y_i = \beta + U_i$ تبدیل خواهد شد و بدین معنی است که تغییرات Y_i کاملاً تصادفی بوده و تنها تابعی از U_i است. این سؤال مهم معمولاً در بحث آزمونها مطرح می‌شود: آیا بین آزمونهای معنی‌دار بودن هریک از پارامترها و آزمون همزمان آنها تفاوت وجود دارد؟ به عبارت دیگر، اگر برای هریک از مقادیر $i = 2, 3, \ldots, k$ فرضیه‌های $\beta_i = 0$ را به طور جداگانه انجام دهیم و به این نتیجه برسیم که تمام آنها قبول می‌شود، آیا ضرورتاً فرضیهٔ $\beta_2 = \beta_3 = \cdots = \beta_k = 0$ نیز قبول خواهد شد. در همین قسمت نشان خواهیم داد که ضرورتاً این گونه نیست.

آزمون یک ترکیب خطی از پارامترها موضوع قسمت ۶.۵ است. در این قسمت خواهیم دید که یک ترکیب خطی را می‌توان با آمارهٔ t آزمون کرد، ولی برای بیش از یک ترکیب خطی باید ضرورتاً از آزمون F استفاده کنیم که در قسمت ۶.۸ مطرح شده است. سؤال بسیار مهم دیگری مطرح می‌شود که اگر برای دو یا چند پارامتر به طور جداگانه فواصل اطمینان بسازیم، آیا دقیقاً به همان جوابی می‌رسیم که از ساختن همزمان ناحیهٔ اطمینان برای همان پارامترها به دست می‌آید؟ در قسمت ۶.۶ خواهیم دید که پاسخ به این سؤال همواره مثبت نیست و به موازات افزایش همبستگی بین متغیرهای توضیحی، این تفاوت به مراتب بیشتر می‌شود.

در بسیاری از تحلیلهای اقتصادسنجی، موضوع مورد علاقه پژوهشگر می‌تواند این باشد که آیا پارامترهای چندمتغیر توضیحی در یک مدل رگرسیون چندمتغیره با یکدیگر برابر است؟ این سؤال با نام آزمون فرضیهٔ $\beta_i = \beta_{i+1} = \cdots = \beta_j$ در قسمت ۶.۷ مطرح شده است. سرانجام در قسمت ۶.۸ آزمون همزمان چند ترکیب خطی از پارامترها را به زبان ماتریسی بررسی کرده‌ایم. در این قسمت، عمومی‌ترین آمارهٔ آزمون ارائه شده است و نشان داده‌ایم که بسیاری از آزمونهایی که در قسمتهای قبل مطرح گردیده است، می‌تواند حالتهای خاصی از این آماره باشد. با اینکه مباحث این فصل مدلهای رگرسیون

با چند متغیر توضیحی است، اما به علت مشکلات محاسباتی، سعی شده است در مثالهای متعددی که آورده‌ایم، از مدلهای بسیار کوچک استفاده شود.

۶ـ۲ خصوصیات آماری $\hat{\beta}$

مدل رگرسیون چندمتغیره ۴ـ۵ را یک بار دیگر می‌نویسیم:

$$\mathbf{y}= \mathbf{X}\boldsymbol{\beta} + \mathbf{u} ,$$

که در آن، $\mathbf{y} \to n\times 1$، $\mathbf{X} \to n\times k$، $\boldsymbol{\beta} \to k\times 1$ و $\mathbf{u} \to n\times 1$ است. می‌دانیم تخمین پارامترهای این مدل، بر طبق معادلهٔ ۲۱ـ۵، عبارت است از

$$\hat{\boldsymbol{\beta}}= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} .$$

با جایگزینی معادلهٔ ۴ـ۵ در ۲۱ـ۵ داریم

$$\hat{\boldsymbol{\beta}}= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' (\mathbf{X}\boldsymbol{\beta} + \mathbf{u}) ,$$

یا

$$\hat{\boldsymbol{\beta}}= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u} . \qquad (۶ـ۱)$$

ملاحظه می‌شود $\hat{\boldsymbol{\beta}}$ در معادلهٔ ۱ـ۶ تابعی از جملهٔ اختلال ($\mathbf{u}$) است؛ بنابراین یک بردار تصادفی است. سؤال این است که تابع توزیع، میانگین، واریانس و کوواریانس $\hat{\boldsymbol{\beta}}$ چیست؟ با توجه به اینکه $\hat{\boldsymbol{\beta}}$ یک تابع خطی از $\mathbf{u}$ است، اگر $\mathbf{u}$ توزیع نرمال داشته باشد، $\hat{\boldsymbol{\beta}}$ نیز توزیع نرمال خواهد داشت. در رابطهٔ ۱۱ـ۵ دیدیم که فرض چهارم از فرضهای کلاسیک جملهٔ اختلال، درواقع نرمال بودن توزیع آن با میانگین $\mathbf{0}$ و واریانس $\sigma^2\mathbf{I}$ است؛ بنابراین نتیجه می‌گیریم که $\hat{\boldsymbol{\beta}}$ نیز دارای توزیع نرمال است. حال به بررسی میانگین و واریانس آن می‌پردازیم.

۱. میانگین $\hat{\boldsymbol{\beta}}$

معادلهٔ ۱ـ۶ را ملاحظه کنید. می‌دانیم که اولین فرض از فرضهای کلاسیک متغیرهای

توضیحی این است که هریک از متغیرهای توضیحی غیرتصادفی بوده و در آزمایشهای فرضی تکراری ثابت حفظ شده‌اند. با توجه به اینکه، از دو طرف معادلهٔ ۶ـ۱ امید ریاضی می‌گیریم، خواهیم داشت

$$E(\hat{\boldsymbol{\beta}}) = E(\boldsymbol{\beta}) + E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}],$$

$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u}).$$

بنابر فرض اول از فرضهای کلاسیک، یعنی معادلهٔ ۶ـ۵، می‌دانیم $E(\mathbf{u}) = \mathbf{0}$؛ بنابراین

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}. \qquad (۶ـ۲)$$

معادلهٔ ۶ـ۲ دلالت بر این می‌کند که $\hat{\boldsymbol{\beta}}$ یک تخمین ناااریب از $\boldsymbol{\beta}$ است.

۲. واریانس و کوواریانس $\hat{\boldsymbol{\beta}}$

چون $\hat{\boldsymbol{\beta}}$ یک بردار $(k \times ۱)$ است، باید یک ماتریس $(k \times k)$ تعریف کنیم که عناصر قطری آن واریانس مقادیر $\hat{\beta}_k, \ldots, \hat{\beta}_۲, \hat{\beta}_۱$ و عناصر غیرقطری آن کوواریانس β_i و β_j به ازای تمام مقادیر i و j و $i \neq j$ باشد. برای این منظور می‌گوییم ماتریس $E[(\hat{\beta}-\beta)(\hat{\beta}-\beta)']$ مطلوب است، زیرا $(\hat{\beta}-\beta)$ یک بردار $(k \times ۱)$ بوده. و بنابراین ترانهاد آن، یعنی $(\hat{\beta}-\beta)'$ یک بردار $(۱ \times k)$ می‌شود. بدین ترتیب حاصلضرب $(\hat{\beta}-\beta)'(\hat{\beta}-\beta)$ یک ماتریس $(k \times k)$ خواهد بود. اگر از این ماتریس امید ریاضی بگیریم، ماتریس واریانس - کوواریانس $\hat{\boldsymbol{\beta}}$ به دست می‌آید،

$$E[(\hat{\beta}-\beta)(\hat{\beta}-\beta)'] = E\begin{bmatrix}(\hat{\beta}_۱-\beta_۱)\\(\hat{\beta}_۲-\beta_۲)\\\vdots\\(\hat{\beta}_k-\beta_k)\end{bmatrix}[(\hat{\beta}_۱-\beta_۱)(\hat{\beta}_۲-\beta_۲)\ldots(\hat{\beta}_k-\beta_k)].$$

این دو بردار را در هم ضرب کرده، یک ماتریس $(k \times k)$ به دست می‌آوریم. اگر عملگر E را دخالت دهیم، خواهیم داشت

$$E(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})' = \begin{bmatrix} E(\hat{\beta}_1-\beta_1)^2 & E(\hat{\beta}_1-\beta_1)(\hat{\beta}_2-\beta_2) & \dots & E(\hat{\beta}_1-\beta_1)(\hat{\beta}_k-\beta_k) \\ E(\hat{\beta}_2-\beta_2)(\hat{\beta}_1-\beta_1) & E(\hat{\beta}_2-\beta_2)^2 & \dots & E(\hat{\beta}_2-\beta_2)(\hat{\beta}_k-\beta_k) \\ \vdots & & & \\ E(\hat{\beta}_k-\beta_k)(\hat{\beta}_1-\beta_1) & E(\hat{\beta}_k-\beta_k)(\hat{\beta}_2-\beta_2) & \dots & E(\hat{\beta}_k-\beta_k)^2 \end{bmatrix}. \quad (۶\text{-}۳)$$

بنابر تعریف می‌دانیم

$$\mathrm{Var}(\hat{\beta}_i) = E[\hat{\beta}_i - E(\hat{\beta}_i)] = E(\hat{\beta}_i-\beta_i)^2,$$

$$\mathrm{Cov}(\hat{\beta}_i,\hat{\beta}_j) = E[\hat{\beta}_i - E(\hat{\beta}_i)][\hat{\beta}_j - E(\hat{\beta}_j)] = E(\hat{\beta}_i-\beta_i)(\hat{\beta}_j-\beta_j).$$

با جایگزینی تعاریف فوق در معادلهٔ ۶ـ۳ داریم

$$E(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})' = \begin{bmatrix} \mathrm{Var}(\hat{\beta}_1) & \mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_2) & \dots & \mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_k) \\ \mathrm{Cov}(\hat{\beta}_2,\hat{\beta}_1) & \mathrm{Var}(\hat{\beta}_2) & \dots & \mathrm{Cov}(\hat{\beta}_2,\hat{\beta}_k) \\ \vdots & & & \\ \mathrm{Cov}(\hat{\beta}_k,\hat{\beta}_1) & \mathrm{Cov}(\hat{\beta}_k,\hat{\beta}_2) & \dots & \mathrm{Var}(\hat{\beta}_k) \end{bmatrix}. \quad (۶\text{-}۴)$$

بدین ترتیب نشان دادیم که $E(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'$، یک ماتریس $(k\times k)$ و قرینه است که عناصر قطری آن واریانسها و عناصر غیرقطری آن کوواریانسهاست. با این حال هنوز نشان نداده‌ایم که فرمول به دست آوردن مقادیر عددی واریانسها یا کوواریانسها چیست؟ برای به دست آوردن این فرمول، به ترتیب زیر عمل می‌کنیم.

معادلهٔ ۶ـ۱ را به صورت زیر می‌نویسیم،

$$(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}. \quad (۶\text{-}۵)$$

دو طرف معادلهٔ فوق را ترانهاد می‌کنیم. می‌دانیم ترانهاد معکوس برابر معکوس ترانهاد

است. بدین ترتیب، ترانهاد $(\mathbf{X}'\mathbf{X})^{-1}$ برابر معکوس ترانهاد $(\mathbf{X}'\mathbf{X})$ است. چون $(\mathbf{X}'\mathbf{X})$ یک ماتریس متقارن است؛ بنابراین با ترانهاد خود برابر خواهد بود. نتیجه می‌گیریم که ترانهاد $(\mathbf{X}'\mathbf{X})^{-1}$ برابر $(\mathbf{X}'\mathbf{X})^{-1}$ است؛ بنابراین

$$(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})' = \mathbf{u}'\mathbf{X}\,(\mathbf{X}'\mathbf{X})^{-1}\,.$$

دو طرف معادله فوق را در معادلهٔ ۶ـ۵ ضرب می‌کنیم،

$$(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\,(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})' = (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{u}\,[\mathbf{u}'\mathbf{X}\,(\mathbf{X}'\mathbf{X})^{-1}]\,.$$

از دو طرف معادلهٔ فوق امیدریاضی می‌گیریم. می‌دانیم $\mathbf{X}$ ثابت فرض‌شده است؛ بنابراین

$$\mathrm{E}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})' = (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\,\mathrm{E}(\mathbf{u}\mathbf{u}')\,\mathbf{X}\,(\mathbf{X}'\mathbf{X})^{-1}\,. \qquad (۶\text{ـ}۶)$$

سمت راست معادلهٔ ۶ـ۶ شامل ۵ ماتریس به شرح زیر است،

$$(\mathbf{X}'\mathbf{X})^{-1} \to (k\times k)\,, \qquad \mathrm{E}(\mathbf{u}\mathbf{u}') \to (n\times n)\,,$$
$$\mathbf{X}' \to (k\times n)\,, \qquad \mathbf{X} \to (n\times k)\,.$$

حاصلضرب این ۵ ماتریس، یک ماتریس مربع و متقارن $(k\times k)$ خواهد بود. این ماتریس درواقع همان ماتریس واریانس ـ کوواریانس تخمینهای $\hat{\boldsymbol{\beta}}$ است که در معادلهٔ ۶ـ۴ تعریف شده است. اگر سمت راست معادلهٔ ۶ـ۴ را که یک ماتریس $(k\times k)$ است، مساوی سمت راست معادلهٔ ۶ـ۶ ـ که آن هم یک ماتریس $(k\times k)$ است ـ بگیریم، ضرورتاً عناصر متناظر این دو ماتریس یک به یک با هم برابرند. می‌دانیم عناصر قطری ماتریس سمت راست معادلهٔ (۶ـ۴) واریانس تخمینهای $\hat{\boldsymbol{\beta}}$ است؛ بنابراین عناصر قطری ماتریس سمت راست معادلهٔ ۶ـ۶ نیز دقیقاً مقادیر همین واریانسها خواهد بود. به همین ترتیب می‌توان برای کوواریانس استدلال کرد. پس

$$\mathrm{Var}\,(\hat{\beta}_i) = \{[(\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\,\mathrm{E}(\mathbf{u}\mathbf{u}')\,\mathbf{X}\,(\mathbf{X}'\mathbf{X})^{-1}] \text{ امین عنصر قطری ماتریس } i\}\,,$$

$$\mathrm{Cov}(\hat{\beta}_i\,,\hat{\beta}_j) = \mathrm{Cov}(\hat{\beta}_j\,,\hat{\beta}_i)$$
$$= \{[(\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\,\mathrm{E}(\mathbf{u}\mathbf{u}')\,\mathbf{X}\,(\mathbf{X}'\mathbf{X})^{-1}] \text{ از ماتریس } a_{ij} \text{ یا } a_{ji} \text{ عنصر}\}\,.$$

برای مثال، در مدل رگرسیون زیر،

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k x_{ki} + U_i ,$$

اگر بخواهیم واریانس β_3 را به دست آوریم، کافی است که سومین عنصر قطری ماتریس

$$(\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\,E(\mathbf{u}\mathbf{u}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

را به دست آوریم. به همین ترتیب کوواریانس $\hat{\beta}_2$ و $\hat{\beta}_3$ برابر با عنصر ردیف ۲ و ستون ۶ یا عنصر ردیف ۶ و ستون ۲ ماتریس فوق است.

محاسبهٔ واریانس و کوواریانس تخمینهای $\hat{\beta}$، برای حالتی که جملهٔ اختلال دارای دو خصوصیت واریانس همسانی و عدم خودهمبستگی باشد، بسیار ساده است. درواقع، در چنین حالتی $E(\mathbf{u}\mathbf{u}')$ برابر $\sigma^2\mathbf{I}$ می‌شود و این نکته را در معادلهٔ ۱۱ـ۵ نشان داده‌ایم. معادلهٔ

$$E(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{I} ,$$

را که در آن σ^2 واریانس جملهٔ اختلال و $\mathbf{I}$ ماتریس $(n \times n)$ شامل عناصر قطری یک و غیرقطری صفر است، در معادلهٔ ۶ـ۶ قرار می‌دهیم،

$$E(\hat{\beta}-\beta)(\hat{\beta}-\beta)' = (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\,(\sigma^2\mathbf{I})\,\mathbf{X}\,(\mathbf{X}'\mathbf{X})^{-1} ,$$

$$= \sigma^2\,(\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{X}\,(\mathbf{X}'\mathbf{X})^{-1} ,$$

در نتیجه خواهیم داشت

$$E(\hat{\beta}-\beta)(\hat{\beta}-\beta)' = \sigma^2\,(\mathbf{X}'\mathbf{X})^{-1} . \qquad (۶ـ۷)$$

ملاحظه می‌شود در این حالت، محاسبهٔ واریانسها و کوواریانسها بسیار ساده است، مثلاً برای به دست آوردن $\mathrm{Var}\,(\hat{\beta}_2)$، سومین عنصر قطری $(\mathbf{X}'\mathbf{X})^{-1}$ را گرفته، در σ^2 ضرب می‌کنیم. به همین ترتیب $\mathrm{Cov}(\hat{\beta}_2 , \hat{\beta}_3)$ برابر با حاصلضرب σ^2 در عنصر a_{23} یا a_{32} از $(\mathbf{X}'\mathbf{X})^{-1}$ خواهد بود. در نتیجه می‌توان چنین نوشت،

$$\mathrm{Var}\,(\hat{\beta}_i) = \sigma^2\,[\text{امین عنصر قطری ماتریس } (\mathbf{X}'\mathbf{X})^{-1}] , \qquad (۶ـ۸)$$

$$\text{Cov}(\hat{\beta}_i, \hat{\beta}_j) = \sigma^2 [\text{عنصر } a_{ij} \text{ از ماتریس } (\mathbf{X}'\mathbf{X})^{-1}], \qquad (۶.۹)$$

در محاسبهٔ واریانس و کوواریانس تخمینهای $\hat{\beta}$ باید به این نکته توجه کنیم که اگر ماتریس $\mathbf{X}'\mathbf{X}$ برحسب مشاهدات اصلی و نه انحراف از میانگین محاسبه شده و مدل رگرسیون مفروض نیز دارای جملهٔ ثابت باشد، می‌توان از معادله‌های ۶.۸ و ۶.۹ استفاده کرد. اما معمولاً ماتریس $(\mathbf{X}'\mathbf{X})$ را به کمک مقادیر انحراف از میانگین متغیرهای توضیحی حساب می‌کنند. در این حالت اگر مدل رگرسیون مفروض جملهٔ ثابت داشته باشد، دیگر نمی‌توان از معادله‌های فوق استفاده کرد؛ زیرا در مدل k پارامتر وجود دارد، در حالی که ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$، یک ماتریس مربع $(k-1) \times (k-1)$ است. بدیهی است باید از فرمولهای زیر استفاده شود،

$$\text{Var}(\hat{\beta}_i) = \sigma^2 [\text{امین عنصر قطری ماتریس } (\mathbf{X}'\mathbf{X})^{-1}], \qquad (۶.۱۰)$$

$$\text{Cov}(\hat{\beta}_i, \hat{\beta}_j) = \sigma^2 [\text{عنصر } a_{(i-1)(j-1)} \text{ از ماتریس } (\mathbf{X}'\mathbf{X})^{-1}], \qquad (۶.۱۱)$$

۳. قضیهٔ گاس ـ مارکف *

می‌دانیم تخمین β در مدل $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$، طبق معادله‌های ۵.۲۱ و ۶.۱ برابر است با

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y},$$

$$\hat{\beta} = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}.$$

اگر $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{A}$ باشد، معادله‌های ۵.۲۱ و ۶.۱ را می‌توان به ترتیب به صورت زیر نوشت،

$$\hat{\beta} = \mathbf{A}\mathbf{y}, \quad \hat{\beta} = \mathbf{A}\mathbf{u},$$

با توجه به ثابت بودن $\mathbf{X}$ در آزمایشهای تکراری، نتیجه می‌گیریم که $\mathbf{A}$ ثابت است. بنابراین می‌توان گفت که $\hat{\beta}$ یک ترکیب خطی از $\mathbf{u}$ یا $\mathbf{y}$ است. ناریب بودن $\hat{\beta}$، یعنی $E(\hat{\beta}) = \beta$ را نیز در معادلهٔ ۶.۲ ملاحظه کردیم. بنابراین، تا این قسمت از بحث به این

نتیجه رسیدیم که $\hat{\beta}_{OLS}$ یک تخمین زننده نااریب و خطی از β است. برای اینکه $\hat{\beta}_{OLS}$ بتواند بهترین تخمین زنندهٔ نااریب خطی باشد، باید ثابت کنیم که کمترین واریانس را نیز دارد؛ یعنی هیچ تخمین زنندهٔ دیگری را نمی‌توان یافت که در قلمرو تخمین زننده‌های خطی نااریب، واریانس کمتری از واریانس $\hat{\beta}_{OLS}$ داشته باشد.

برای اثبات، فرض می‌کنیم که یک تخمین زنندهٔ خطی نااریب غیر OLS مانند $\hat{\beta}_*$ وجود دارد. با توجه به خطی بودن این تخمین زننده، خواهیم داشت

$$\hat{\beta}_* = \mathbf{D}\mathbf{y} .$$

اگر $\mathbf{D}$ برابر $\mathbf{A}$ باشد، آنگاه $\hat{\beta}_* = \hat{\beta}_{OLS}$. فرض می‌کنیم که $\mathbf{D} = \mathbf{A} + \mathbf{C}$؛ بنابراین

$$\hat{\beta}_* = (\mathbf{A} + \mathbf{C})\mathbf{y} .$$

با جایگزینی معادلهٔ ٤ـ٥ در معادله فوق داریم:

$$\hat{\beta}_* = (\mathbf{A} + \mathbf{C})(\mathbf{X}\beta + \mathbf{u}) ,$$

$$= (\mathbf{A} + \mathbf{C})\mathbf{X}\beta + (\mathbf{A} + \mathbf{C})\mathbf{u} .$$

با توجه به مقدار ماتریس $\mathbf{A}$، ملاحظه می‌شود که

$$\mathbf{AX} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{I} ,$$

در نتیجه

$$\hat{\beta}_* = \beta + \mathbf{CX}\beta + (\mathbf{A} + \mathbf{C})\mathbf{u} ,$$

$$= (\mathbf{I} + \mathbf{CX}) + (\mathbf{A} + \mathbf{C})\mathbf{u} .$$

از دو طرف معادلهٔ فوق امید ریاضی می‌گیریم،

$$E(\hat{\beta}_*) = (\mathbf{I} + \mathbf{CX})\beta .$$

اگر قرار است $\hat{\beta}_*$ نااریب باشد، باید ضرورتاً

$$\mathbf{CX} = \mathbf{0}$$

بنابراین بر فرض ناااریب بودن $\hat{\beta}_*$ خواهیم داشت

$$\hat{\beta}_* = \beta + (A + C)\,u\ .$$

با توجه به تعریف واریانس، می‌توان چنین نوشت

$$\text{Var}\,(\hat{\beta}_*) = E[(\hat{\beta}_* - \beta)\,(\hat{\beta}_* - \beta)'\,]\ ,$$

$$= E\{[(A + C)\,u]\,[(A + C)\,u]'\}\ ,$$

$$= E[(A + C)\,uu'\,(A + C)'\,]\ ,$$

$$= (A + C)\,E(uu')\,(A + C)'\ .$$

براساس فرضهای واریانس همسانی و عدم خودهمبستگی، داریم $E(uu') = \sigma^2 I$؛ بنابراین

$$\text{Var}\,(\hat{\beta}_*) = \sigma^2\,(A + C)\,(A + C)'\ .$$

اما می‌دانیم که

$$(A + C)\,(A + C)' = A'A + CA' + AC' + CC'\ .$$

مقدار A را در رابطهٔ فوق قرار می‌دهیم، خواهیم داشت

$$(A + C)\,(A + C)' = (X'X)^{-1}\,X'X\,(X'X)^{-1} + CX(X'X)^{-1}$$

$$+ (X'X)^{-1}\,X'C' + CC'\ .$$

با توجه به اینکه $CX = 0$ بنابراین $X'C = 0$، در نتیجه

$$(A + C)\,(A + C)' = (X'X)^{-1} + CC'\ .$$

با جایگزینی معادله فوق در واریانس $\hat{\beta}_*$، داریم

$$\text{Var}\,(\hat{\beta}_*) = \sigma^2\,[(X'X)^{-1} + CC'] = \sigma^2\,(X'X)^{-1} + \sigma^2\,CC'\ .$$

اما می‌دانیم $\sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$، درواقع همان واریانس $\hat{\beta}_{OLS}$است؛ بنابراین

$$\text{Var}(\hat{\beta}_*) = \text{Var}(\hat{\beta}_{OLS}) + \sigma^2 \mathbf{CC}' .$$

همچنین می‌دانیم $\mathbf{CC}'$ یک «ماتریس نیمه معین مثبت»[1] است. برای اینکه یک فرم درجهٔ دوم این ماتریس صفر باشد، باید $\mathbf{C}=0$، یعنی تمام عناصر $\mathbf{C}$ باید صفر شود. حال می‌گوییم $\text{Var}(\hat{\beta}_*)$ برابر مجموع دو مقدار غیرمنفی شده است. واریانس $\hat{\beta}_{OLS}$نمی‌تواند صفر باشد؛ بنابراین برای رسیدن به مقدار حداقل برای واریانس $\hat{\beta}_*$، ضرورتاً جملهٔ دوم، یعنی $\sigma^2\mathbf{CC}'$، باید مقداری برابر صفر اختیار کند. می‌دانیم لازمهٔ این کار صفر شدن عناصر $\mathbf{C}$ است. اما اگر $\mathbf{C}=0$، آنگاه

$$\mathbf{D} = \mathbf{A} + \mathbf{C} = \mathbf{A} ,$$

یعنی $\hat{\beta}_* = \mathbf{Dy}$ با $\hat{\beta}_{OLS} = \mathbf{Ay}$ دقیقاً متساوی خواهد بود. به عبارت دیگر، در قلمرو تخمین زننده‌های خطی نااریب، تخمین زننده‌ای که حداقل واریانس را دارد، ضرورتاً حداقل مربعات معمولی خواهد بود. بدین ترتیب ثابت شد که $\hat{\beta}_{OLS}$بهترین تخمین زنندهٔ نااریب خطی است، که درواقع اثبات قضیهٔ گاس ـ مارکف است.

دیدیم که $\hat{\beta}$توزیع نرمال دارد و میانگین و ماتریس واریانس ـ کوواریانس آن را نیز به دست آوردیم. این اطلاعات را با توجه به فرضهای کلاسیک می‌توان به صورت زیر نوشت،

$$\hat{\beta} \sim \mathrm{N}[\beta , \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}] . \qquad (۶.۱۲)$$

علاوه بر اینکه $\hat{\beta}$ بهترین تخمین زنندهٔ نااریب خطی است، می‌توان نشان داد که سازگار نیز هست. می‌دانیم عناصر ماتریس $(\mathbf{X}'\mathbf{X})$ مقادیر ثابتی است؛ بنابراین $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ را می‌توان به صورت $\frac{\sigma^2}{n}(\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}$ نوشت. برای بررسی خصوصیت سازگاری، باید حد واریانس $\hat{\beta}$را به دست آوریم، وقتی n به سمت بی‌نهایت میل

1. Positive Semidefinite Matrix

به پیوست «۵ ـ الف» مراجعه شود.

می‌کند. خواهیم داشت

$$\lim_{n \to \infty} \sigma^2 (X'X)^{-1} = \lim_{n \to \infty} \frac{\sigma^2}{n} \left(\frac{X'X}{n}\right)^{-1} = 0 ,$$

زیرا اولاً مقدار $\frac{\sigma^2}{n}$ در حد وقتی n به سمت بی‌نهایت میل کند برابر صفر می‌شود؛ یعنی

$$\lim_{n \to \infty} \frac{\sigma^2}{n} = 0 ,$$

و ثانیاً با توجه به فرضهای کلاسیک در مورد متغیرهای توضیحی، می‌دانیم که ماتریس X تصادفی نیست و عناصر آن در آزمایشهای تکراری ثابت فرض شده است. همچنین به علت عدم‌خودهمبستگی کامل، ماتریس $\frac{1}{n}(X'X)$ دترمینان غیرصفر دارد و وارون‌پذیر است و عناصر آن به ازای $n \to \infty$، مقادیر نامعین ندارد؛ یعنی

$$\lim_{n \to \infty} \left(\frac{X'X}{n}\right)^{-1} = \text{مقدار معین} . \qquad (۶-۱۳)$$

می‌توان نشان داد که اگر فرض نرمال بودن تابع توزیع احتمال جملهٔ اختلال را کنار بگذاریم، تخمین زنندهٔ $\hat{\beta}_{OLS}$ هنوز خصوصیت بهترین تخمین زنندهٔ نااریب خطی را حفظ می‌کند و مقدار واریانس آن نیز تغییری نخواهد کرد، اما دیگر یک تخمین زننده «کارآ» نخواهد بود.[۱]

۴. تخمین و آزمون واریانس جملهٔ اختلال *

در فرمولهای ۶-۸ و ۶-۹ و همچنین ۶-۱۰ و ۶-۱۱ دیدیم که محاسبهٔ واریانس و کوواریانس $\hat{\beta}$ مستلزم داشتن مقدار واریانس جملهٔ اختلال، یعنی σ^2_u است. می‌دانیم واریانس جملهٔ اختلال به علت تصادفی بودن u، باید تخمین زده شود. قبلاً در فصل دوم و معادلهٔ ۲-۲۸ فرمول تخمین جملهٔ اختلال را برای یک مدل رگرسیون ساده با دو

۱. به کتاب *اقتصادسنجی* تألیف P. Schmidt (۱۹۷۶) مراجعه شود.

پارامتر ملاحظه کردیم،

$$\hat{\sigma}^2 = \frac{\Sigma e_i^2}{n-2}.$$

در مواردی نیز این نتیجه را به حالت کلی با k پارامتر تعمیم داده‌ایم و از فرمول زیر استفاده کرده‌ایم،

$$\hat{\sigma}^2 = \frac{\Sigma e_i^2}{n-2} = \frac{e'e}{n-k}.$$

در این قسمت می‌خواهیم نه تنها صحت فرمول فوق را ثابت کنیم، بلکه نشان دهیم که اگر از این فرمول استفاده شود یک تخمین نااریب از σ^2 خواهیم داشت.

باید از مجموع مربعات پسماند، یعنی $e'e$ امید ریاضی بگیریم. برای این منظور ابتدا از معادلهٔ ۱۷ـ۵ شروع می‌کنیم،

$$e = y - X\hat{\beta}.$$

مقدار $\hat{\beta}$ از معادلهٔ ۲۱ـ۵ را در معادلهٔ فوق قرار می‌دهیم،

$$e = y - X(X'X)^{-1}X'y,$$

یا

$$e = [I - X(X'X)^{-1}X']y. \qquad (۶\text{ـ}۱۴)$$

ماتریس $(n \times n)$ زیر را تعریف می‌کنیم،

$$M = I - X(X'X)^{-1}X'. \qquad (۶\text{ـ}۱۵)$$

ماتریس M بسیار مهم است؛ زیرا، اولاً یک ماتریس متقارن $(n \times n)$ است، ثانیاً $M.M = M^2 = M$، یعنی M یک ماتریس «خودتوان» است، یعنی وقتی M را در خودش ضرب کنیم، حاصل با خودش برابر است و ثالثاً، وقتی M را از سمت چپ در X ضرب کنیم یک ماتریس صفر به دست می‌آوریم، یعنی

$$MX = 0 \qquad (۶\text{ـ}۱۶)$$

زیرا $$\mathbf{MX}= [\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{X},$$

$$= \mathbf{X}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})= \mathbf{0}.$$

با این مقدمات، معادلهٔ ٦ـ١٥ را در ٦ـ١٤ قرار می‌دهیم، خواهیم داشت

$$\mathbf{e}= \mathbf{My},$$

$$= \mathbf{M}(\mathbf{X}\beta + \mathbf{u}),$$

$$= \mathbf{MX}\beta + \mathbf{Mu},$$

با توجه به معادلهٔ ٦ـ١٦ داریم

$$\mathbf{e}= \mathbf{Mu}. \qquad (٦ـ١٧)$$

مجموع مربعات پسماند، یعنی Σe_t^2، را می‌نویسیم و معادلهٔ فوق را در آن جایگزین می‌کنیم،

$$\Sigma e_t^2 = \mathbf{e}'\mathbf{e},$$

$$= \mathbf{u}'\mathbf{M}'\mathbf{Mu}.$$

می‌دانیم چون $\mathbf{M}$ متقارن است؛ پس $\mathbf{M}' = \mathbf{M}$ بنابراین

$$\mathbf{M}'\mathbf{M}= \mathbf{MM}= \mathbf{M}.$$

در نتیجه

$$\mathbf{e}'\mathbf{e}= \mathbf{u}'\mathbf{Mu}.$$

همچنین ملاحظه می‌شود که $\mathbf{u}'\mathbf{Mu}$ یک شکل درجه دوم است که مقدار آن برابر است با

$$\mathbf{e}'\mathbf{e}= [U_1 \ U_2 \ \ldots U_n]\begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ & & & \\ M_{n1} & M_{1n2} & \cdots & M_{nn} \end{bmatrix}\begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix},$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} U_i U_j M_{ij}.$$

از رابطهٔ فوق امید ریاضی می‌گیریم،

$$E(e'e)= E\Big[\sum_{i=1}^{n}\sum_{j=1}^{n} U_i U_j M_{ij}\Big].$$

با توجه به فرضهای واریانس همسانی و عدم خودهمبستگی، می‌دانیم،

$$Var(U_i)= E(U_i)^2= \sigma^2,$$

$$Cov(U_i, U_j)= E(U_i U_j)= 0,$$

در نتیجه خواهیم داشت

$$E(e'e)= \sigma^2 \sum_i M_{ii}. \qquad (۶\text{-}۱۸)$$

اما می‌دانیم که $\sum_i M_{ii}$ با مجموع عناصر قطری ماتریس $\mathbf{M}$ که «اثر ماتریس»[1] نام دارد و با $tr(\mathbf{M})$ نشان داده می‌شود برابر است. در پیوست ماتریسها (پیوست ۵ ـ الف) دو خاصیت مهم tr عبارت است از

$$tr(\mathbf{A}+\mathbf{B})= tr(\mathbf{A})+tr(\mathbf{B}),$$

که $\mathbf{A}$ و $\mathbf{B}$ ماتریسهای مربع بوده و همرتبه هستند و همچنین

$$tr(\mathbf{AB})= tr(\mathbf{BA}),$$

که $\mathbf{A}$ و $\mathbf{B}$ به ترتیب $(m\times n)$ و $(n\times m)$ هستند. با استفاده از این خصوصیات، معادلهٔ ۶ـ۱۸ را چنین می‌نویسیم،

$$\begin{aligned} E(e'e)&= \sigma^2\, tr(\mathbf{M}),\\ &= \sigma^2\, tr[\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'],\\ &= \sigma^2\,\{tr(\mathbf{I})-tr[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\}. \end{aligned} \qquad (۶\text{-}۱۹)$$

1. Trace of a Matrix

اما می‌دانیم $tr(\mathbf{I}) = n$، زیرا $\mathbf{I}$ یک ماتریس $(n \times n)$ است که عناصر قطری آن عدد یک است. همچنین می‌دانیم

$$tr[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] = tr\{[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{X}'\},$$

$$= tr\mathbf{X}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}],$$

$$= tr(\mathbf{I})$$

$$= k,$$

زیرا $(\mathbf{X}'\mathbf{X})^{-1}$ یک ماتریس $(k \times k)$ است. بنابراین معادلهٔ ٦ـ١٩ را می‌توان چنین نوشت،

$$E(\mathbf{e}'\mathbf{e}) = \sigma^2(n-k),$$

با تقسیم دو طرف معادلهٔ فوق بر $(n-k)$ و توجه به این نکته که $(n-k)$ عدد ثابتی است، داریم

$$E\left(\frac{\mathbf{e}'\mathbf{e}}{n-k}\right) = \sigma^2. \qquad (٦ـ٢٠)$$

به عبارت دیگر، اگر $\frac{\mathbf{e}'\mathbf{e}}{n-k}$ را تخمینی از σ^2 بدانیم، آنگاه امید ریاضی این تخمین برابر مقدار واقعی پارامتر است؛ یعنی $\frac{\mathbf{e}'\mathbf{e}}{n-k}$ یک تخمین ناارریب از σ^2 است. بنابراین

$$\hat{\sigma}_U^2 = \frac{\mathbf{e}'\mathbf{e}}{n-k}. \qquad (٦ـ٢١)$$

اگر تخمین واریانس جملهٔ اختلال را در فرمولهای ٦ـ٨ و ٦ـ٩ یا ٦ـ١٠ و ٦ـ١١ قرار دهیم، فرمولهای محاسباتی واریانس و کوواریانس $\hat{\beta}$ به دست می‌آید. برای مدلهای رگرسیون با ضریب ثابت که $(\mathbf{X}'\mathbf{X})$ برحسب مشاهدات اصلی محاسبه شده و جمله اختلال کلاسیک است، داریم

$$Var(\hat{\beta}_i) = \hat{\sigma}_U^2[(\mathbf{X}'\mathbf{X})^{-1} \text{ امین عنصر قطری ماتریس } i], \qquad (٦ـ٢٢)$$

$$Cov(\hat{\beta}_i, \hat{\beta}_j) = \hat{\sigma}_U^2[(\mathbf{X}'\mathbf{X})^{-1} \text{ از ماتریس } a_{ij} \text{ عنصر}]. \qquad (٦ـ٢٣)$$

برای مدلهایی که $(\mathbf{X}'\mathbf{X})$ برحسب انحراف از میانگین محاسبه شده و فرضهای کلاسیک را نیز شامل است خواهیم داشت

$$\text{Var}\,(\hat{\beta}_i) = \hat{\sigma}_U^2[(\mathbf{X}'\mathbf{X})^{-1}\text{امین عنصر قطری ماتریس } (i-1)] \,, \qquad (۶\text{-}۲۴)$$

$$\text{Cov}\,(\hat{\beta}_i, \hat{\beta}_j) = \hat{\sigma}_U^2[(\mathbf{X}'\mathbf{X})^{-1}\text{از ماتریس } a_{(i-1)\,(j-1)}\text{عنصر}] \,. \qquad (۶\text{-}۲۵)$$

آزمون واریانس جمله اختلال

بحث آزمون σ^2 در رگرسیون چندمتغیره دقیقاً مشابه مطالبی است که در فصل دوم با همین عنوان برای رگرسیون ساده مطرح کردیم. بنابراین می‌توان در این قسمت، مطالب را فهرست‌وار مطرح کرد. براساس معادلهٔ ۲ـ۳۸ می‌توان گفت که

$$\frac{\mathbf{e}'\mathbf{e}}{\sigma^2} \sim \chi^2\,(n-k) \,,$$

که در آن $\mathbf{e}'\mathbf{e}$ مجموع مربعات پسماند در مدلی است که k پارامتر دارد. ملاحظه می‌شود که $\frac{\mathbf{e}'\mathbf{e}}{\sigma^2}$ دارای توزیع χ^2 با $(n-k)$ درجه آزادی است. اگر بخواهیم یک فاصلهٔ اطمینان ۹۵ درصد برای $\frac{\mathbf{e}'\mathbf{e}}{\sigma^2}$ به دست آوریم، باید سطحی از منحنی تغییرات χ^2 را تعیین کنیم که ۹۵ درصد سطح زیر منحنی در ناحیهٔ مرکزی و ۰/۰۲۵ در هریک از دو طرف باشد. اگر نقاط متناظر با ابتدا و انتهای این سطح در روی محور χ^2 را $\chi^2_{۰/۹۷۵}$ و $\chi^2_{۰/۰۲۵}$ بنامیم، خواهیم داشت

$$\Pr\left(\chi^2_{۰/۰۲۵} < \frac{\mathbf{e}'\mathbf{e}}{\sigma^2} < \chi^2_{۰/۹۷۵}\right) = ۰/۹۵ \,,$$

یا به عبارت دیگر، در سطح احتمال ۹۵ درصد داریم

$$\chi^2_{۰/۰۲۵} < \frac{\mathbf{e}'\mathbf{e}}{\sigma^2} < \chi^2_{۰/۹۷۵} \,.$$

اما در آزمون واریانس جملهٔ اختلال، هدف ما این است که یک فاصلهٔ اطمینان برای σ^2 به دست آوریم. با مراجعه به معادلهٔ ۶ـ۲۱ داریم

$$\hat{\sigma}_U^2 = \frac{\mathbf{e}'\mathbf{e}}{n-k} \,,$$

یا

$$e'e = (n-k)\,\hat{\sigma}_U^2 .$$

معادلهٔ فوق را در فاصلهٔ اطمینان برای $\frac{e'e}{\sigma^2}$ جایگزین می‌کنیم،

$$\chi^2_{۰/۰۲۵} < \frac{(n-k)\,\hat{\sigma}^2}{\sigma^2} < \chi^2_{۰/۹۷۵} .$$

دو طرف رابطه فوق را معکوس کرده، سپس در $(n-k)\,\hat{\sigma}^2$ ضرب می‌کنیم. خواهیم داشت

$$\frac{(n-k)\,\hat{\sigma}^2}{\chi^2_{۰/۹۷۵}} < \sigma^2 < \frac{(n-k)\,\hat{\sigma}^2}{\chi^2_{۰/۰۲۵}} .$$

رابطهٔ فوق فاصلهٔ اطمینان ۹۵ درصد برای σ^2 را تعیین می‌کند. فرضیه‌های مختلف برای σ^2 را می‌توان با داشتن این فاصلهٔ اطمینان آزمون کرد. اگر H_0 در قلمرو این فاصلهٔ اطمینان قرار گیرد، قبول شده، در غیر این صورت رد خواهد شد.

مثال ۱ـ۶ مدل رگرسیون را که در مثال ۲ـ۵ مطرح شده است، ملاحظه کنید. می‌خواهیم برای واریانس جملهٔ اختلال این مدل، یک فاصلهٔ اطمینان ۹۵ درصد بسازیم.

براساس محاسبات انجام شده در مثال ۲ـ۵ می‌دانیم $\hat{\sigma}^2_U = ۰/۷۵$. همچنین می‌دانیم $n = ۵$ و $k = ۳$. مقادیر $\chi^2_{۰/۰۲۵}$ را از جدول χ^2 و با درجات آزادی $n - k = ۲$ به دست می‌آوریم. خواهیم داشت

$$\chi^2_{۰/۹۷۵} = ۷/۳۸ \quad , \quad \chi^2_{۰/۰۲۵} = ۰/۰۵۰۶ .$$

بنابراین

$$\Pr\left(\frac{(۵-۳)\,\hat{\sigma}^2}{\chi^2_{۰/۹۷۵}} < \sigma^2 < \frac{(۵-۳)\,\hat{\sigma}^2}{\chi^2_{۰/۰۲۵}}\right) = ۰/۹۵ ,$$

یا

$$\frac{(۵-۳)\,۰/۷۵}{۷/۳۸} < \sigma^2 < \frac{(۵-۳)\,(۰/۷۵)}{۰/۰۵۰۶} ,$$

$$۰/۲۰۳ < \sigma^2 < ۲۹/۶۴ .$$

ملاحظه می‌شود که فاصلهٔ اطمینان بسیار بزرگ است که به معنی پایین بودن دقت در تخمین واریانس جملهٔ اختلال است. ضعف تخمینها در تمام مواردی که در مورد مثال ۲ـ۵ مطرح کردیم یا خواهیم کرد، ملاحظه می‌شود. این امر به دلیل تعداد بسیار کم مشاهدات (n=۵) در مقایسه با تعداد پارامترها (k=۳) است.

۳ـ۶ آزمون هریک از پارامترها

در این قسمت، ابتدا آزمون $\beta_i = a$، که a عدد ثابت و معلومی است، و نیز آزمون $\beta_i = ۰$ بررسی می‌شود. در ادامهٔ این قسمت فرضیهٔ معنی‌دار بودن زیر مجموعه‌ای از پارامترهای یک مدل رگرسیون چند متغیره، یعنی فرضیهٔ $\beta_j = \dots = \beta_{i+1} = \beta_i = ۰$ مطرح می‌شود.

۱. آزمون فرضیهٔ $\beta_i = a$

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \beta_۱ + \beta_۲ X_{۲t} + \dots + \beta_k X_{kt} + U_t .$$

روش آزمون فرضیه در مورد هریک از پارامترهای این مدل دقیقاً مانند آزمون فرضیه در رگرسیون ساده است. پارامتر β_i را در نظر می‌گیریم. فرض کنید می‌خواهیم فرضیهٔ $H_۰: \beta_i = a$ را در مقابل فرضیهٔ $H_۱: \beta_i \neq a$ در سطح معنی‌دار ۵ درصد آزمون کنیم. ابتدا مدل را تخمین می‌زنیم و $\hat{\beta}_i$ را به دست آورده و آن را استاندارد می‌کنیم. می‌دانیم برای استاندارد کردن $\hat{\beta}_i$ باید آن را از میانگین خود کم کرده و بر انحراف معیارش تقسیم کرد. با توجه به اینکه انحراف معیار همان جذر واریانس است، ابتدا باید واریانس $\hat{\beta}_i$ را تخمین زد. فرمول ۲۲ـ۶ یا ۲۴ـ۶ تخمین واریانس $\hat{\beta}_i$ را تعیین می‌کند. با داشتن تخمین واریانس $\hat{\beta}_i$، مقدار استاندارد شده $\hat{\beta}$ توزیع t با (n - k) درجهٔ آزادی خواهد داشت؛ بنابراین

$$t = \frac{\hat{\beta}_i - E(\hat{\beta}_i)}{\sqrt{Var(\hat{\beta}_i)}} \sim t(n-k) .$$

با توجه به $E(\hat{\beta}_i)=\beta_i$ و $SE(\hat{\beta}_i)=\sqrt{Var(\hat{\beta}_i)}$ خواهیم داشت

$$t=\frac{\hat{\beta}_i-\beta_i}{SE(\hat{\beta}_i)} \sim t(n-k) .$$

اگر فرضیهٔ H_0 صحیح باشد، داریم

$$t=\frac{\hat{\beta}_i-a}{SE(\hat{\beta}_i)} \sim t(n-k) . \qquad (۶-۲۶)$$

مقدار t که از معادلهٔ فوق به دست می‌آید، یعنی آمارهٔ آزمون، را با مقدار جدول t مقایسه می‌کنیم؛ اگر آمارهٔ آزمون در ناحیهٔ بحرانی قرار گرفت فرضیه H_0 رد شده و در غیر این صورت رد نخواهد شد. اگر آزمون در سطح معنی دار α درصد باشد، مقدار t را از جدول، باید در سطح معنی‌دار $\alpha/2$ و با $(n-k)$ درجه آزادی به دست آورد؛ زیرا آزمون ما دو طرفه است.

براساس رابطهٔ ۶-۲۶ می‌توان برای β_i فاصلهٔ اطمینان نیز ساخت. فرض کنید می‌خواهیم یک فاصلهٔ اطمینان $(1-\alpha)$ درصدی برای β_i بسازیم. با توجه به معادلهٔ ۴۱-۲ داریم

$$-t_{\alpha/2} < \frac{\hat{\beta}_i-\beta_i}{SE(\hat{\beta}_i)} < t_{\alpha/2} ,$$

یا

$$\hat{\beta}_i - t_{\alpha/2} SE(\hat{\beta}_i) < \beta_i < \beta_i + t_{\alpha/2} SE(\hat{\beta}_i) . \qquad (۶-۲۷)$$

بدیهی است با داشتن فاصلهٔ اطمینان برای β_i، می‌توان فرضیهٔ $H_0: \beta_i = a$ را نیز آزمون کرد. اگر $\beta_i = a$ در فاصلهٔ اطمینان واقع شود، H_0 را قبول می‌کنیم و در غیر این صورت H_0 رد خواهد شد.

آزمون معنی‌دار بودن β_i

آزمون فرضیهٔ $H_0: \beta_i = a$ را آزمون فرضیه معنی‌دار بودن β_i می‌گویند. اگر H_0 را نتوان

رد کرد، X_{it} در توضیح تغییرات Y_t تأثیری نخواهد داشت. کافی است در رابطهٔ ۶ـ۲۶ به جای a مقدار صفر قرار دهیم. خواهیم داشت

$$t = \frac{\hat{\beta}_i - 0}{SE(\hat{\beta}_i)} = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} . \qquad (۶ـ۲۸)$$

مقدار آمارهٔ آزمون را با مقدار جدول t مقایسه می‌کنیم. اگر آمارهٔ آزمون در ناحیهٔ بحرانی افتاد، H_0 رد می‌شود، در غیر این صورت H_0 را قبول می‌کنیم. یادآوری می‌شود که اگر $H_1: \beta_i \neq 0$، آزمون ما دوطرفه خواهد بود. اگر آزمون در سطح معنی‌دار α درصد اجرا شود، باید مقدار t را با $(n-k)$ درجه آزادی و در سطح معنی‌دار $\alpha/2$ درصد از جدول t به دست آوریم.

مثال ۶ـ۲ مدلی را که در مثال ۵ـ۲ مطرح کردیم، یک بار دیگر در نظر می‌گیریم،

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t .$$

با استفاده از مشاهدات مندرج در جدول ۵ـ۲ این مدل را به صورت زیر تخمین زده‌ایم،

$$\hat{Y}_t = ۴ + ۲/۵ X_t - ۱/۵ Z_t .$$

همچنین در مثال ۵ـ۲ دیدیم که:

$$(X'X)^{-1} = \begin{bmatrix} ۱ & -۱/۵ \\ -۱/۵ & ۲/۵ \end{bmatrix} , \quad ESS = ۲۶/۵ , \quad TSS = ۲۸ ,$$
$$RSS = ۱/۵ , \quad R^2 = ۰/۹۴۶۴ .$$

با اطلاعات فوق

الف) واریانس U_t را تخمین بزنید.

ب) واریانس $\hat{\beta}$ و واریانس $\hat{\gamma}$ را تخمین بزنید.

ج) فرضیهٔ $H_0: \beta_i = ۲$ را در مقابل فرضیهٔ $H_1: \beta_i \neq ۲$ در سطح معنی‌دار ۵ درصد آزمون کنید.

د) فرضیهٔ $H_0: \gamma = 0$ را در مقابل فرضیهٔ $H_1: \gamma \neq 0$ در سطح معنی‌دار ۵ درصد آزمون کنید.

هـ) برای β یک فاصلهٔ اطمینان ۹۵ درصد بسازید.

برای قسمت الف، یعنی تخمین واریانس U_i از فرمول ۶ـ۲۱ استفاده می‌کنیم،

$$\hat{\sigma}_U^2 = \frac{\mathbf{e}'\mathbf{e}}{n-k}.$$

می‌دانیم $\mathbf{e}'\mathbf{e} = ۱/۵$، $n = ۵$ و $k = ۳$. بنابراین

$$\hat{\sigma}_U^2 = \frac{۱/۵}{۵-۳} = ۰/۷۵.$$

در قسمت ب، تخمین واریانس $\hat{\beta}$ و تخمین واریانس $\hat{\gamma}$ را باید از فرمول ۶ـ۲۴ به دست آورد؛ زیرا اولاً مدل رگرسیون مفروض شامل جملهٔ ثابت است. ثانیاً ماتریس $(\mathbf{X}'\mathbf{X})$ برحسب مقادیر انحراف از میانگین محاسبه شده است. β_1 درواقع دومین پارامتر در مدل رگرسیون است؛ بنابراین، باید $\hat{\sigma}_U^2$ را در اولین عنصر قطری $(\mathbf{X}'\mathbf{X})^{-1}$ ضرب کنیم. به‌همین ترتیب برای محاسبهٔ واریانس $\hat{\gamma}$ ـ که سومین پارامتر است ـ باید دومین عنصر قطری ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ در $\hat{\sigma}_U^2$ ضرب شود. خواهیم داشت

$$\mathrm{Var}(\hat{\beta}) = ۰/۷۵\,(۱) = ۰/۷۵,$$

$$\mathrm{Var}(\hat{\gamma}) = ۰/۷۵\,(۲/۵) = ۱/۸۷۵.$$

در قسمت ج، رابطهٔ ۶ـ۲۶ را می‌نویسیم،

$$t = \frac{\hat{\beta}_i - a}{\mathrm{SE}(\hat{\beta}_i)},$$

که با توجه به $H_0: \beta = ۲$ مقدار a برابر ۲ است. مقدار آمارهٔ آزمون برابر است با

$$t = \frac{۲/۵ - ۲}{\sqrt{۰/۷۵}} = \frac{۰/۵}{۰/۸۶۶} = ۰/۵۷۷.$$

می‌دانیم با دو درجهٔ آزادی و سطح معنی‌دار ۰/۰۲۵ داریم $t_{۰/۰۲۵}(۲) = ۴/۳۰۳$. بنابراین آماره آزمون درناحیهٔ بحرانی قرارنمی‌گیرد، پس معنی‌دارنیست و فرضیهٔ H_0 رد نمی‌شود.

در قسمت د، برای آزمون معنی‌دار بودن فرضیهٔ $H_0: \gamma = 0$ از رابطهٔ ۶ـ۲۸ استفاده می‌کنیم. آماره آزمون عبارت است از

$$t = \frac{\hat{\gamma}}{SE(\hat{\gamma})} = \frac{-1/5}{\sqrt{1/875}} = \frac{-1/5}{1/369} = -1/1 .$$

ملاحظه می‌شود که آماره آزمون معنی‌دار نیست، زیرا در ناحیهٔ بحرانی قرار نمی‌گیرد. بنابراین $H_0: \gamma = 0$ رد نخواهد شد.

در قسمت هـ، برای به دست آوردن فاصلهٔ اطمینان ۹۵ درصد برای β، می‌توان از رابطهٔ ۶ـ۲۷ استفاده کرد،

$$\hat{\beta} - t_{\alpha/2}\, SE(\hat{\beta}) < \beta < \hat{\beta} + t_{\alpha/2}\, SE(\hat{\beta}) ,$$

یا

$$2/5 - 4/303\,(0/866) < \beta < 2/5 + 4/303\,(0/866) ,$$

$$2/5 - 3/7 < \beta < 2/5 + 3/7 ,$$

$$-1/2 < \beta < 6/2 .$$

۲. آزمون معنی‌دار بودن زیرمجموعه‌ای از پارامترها: «آزمون والد»[1]

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_m X_{mt} + V_t . \qquad (6\text{ـ}29)$$

فرض می‌کنیم براساس یک نظریه، نه تنها تمام متغیرهای توضیحی فوق ضروری است، بلکه باید r متغیر توضیحی دیگر نیز به مدل اضافه کرد، یعنی

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_m X_{mt} + \beta_{m+1} X_{(m+1)t} + \cdots + \beta_n X_{kt} + U_t ,$$

(۶ـ۳۰)

---

۱. Wald Test، برای مطالعه بیشتر به مقالهٔ A. Wald (۱۹۴۳) مراجعه شود.

که در آن $r = k - m$. برای آزمون صحت نظریه مفروض باید فرضیه زیر را آزمود.

$$H_0: \beta_{m+1} = \beta_{m+2} = \cdots = \beta_k = 0 \qquad (٦-٣١)$$

بنابراین موضوع بحث ما در این قسمت تخمین مدل ٦-٣٠ و آزمون فرضیهٔ ٦-٣١ است.

می‌توان رابطهٔ ٦-٣١ را مجموعه‌ای از محدودیتهای صفری برای پارامترهای مدل ٦-٣٠ تعبیر کرد. بنابراین، اگر این محدودیتها را در مدل ٦-٣٠ قرار دهیم، یک مدل رگرسیون مقید[1] خواهیم داشت که درواقع همان مدل ٦-٢٩ است. در ادامهٔ بحث، هرگاه برای یک متغیر از اندیس r استفاده کنیم، بدین معنی است که آن متغیر متعلق به مدل رگرسیون مقید ٦-٢٩ است. متغیرهای بدون اندیس متعلق به مدل غیرمقید ٦-٣٠ خواهد بود.

در قسمت ٥-٥ و در رابطهٔ ٥-٥٩ دیدیم که اگر در یک رگرسیون چند متغیره، تعداد متغیرهای توضیحی را افزایش دهیم، مجموع مربعات پسماند کاهش یافته، یا حداقل ثابت می‌ماند. این امر بدین معنی است که R^2 افزایش یافته و یا حداقل ثابت خواهد ماند. بنابراین، انتظار ما این است که مجموع مربعات پسماند در مدل ٦-٣٠، یعنی $RSS = \Sigma e_i^2$، از مقدار مشابه آن در مدل مقید ٦-٢٩، یعنی $RSS_r = \Sigma e_{ir}^2$، کمتر بوده یا حداقل مساوی آن باشد. در اینجا کمیت زیر را تشکیل می‌دهیم،

$$\frac{(RSS_r - RSS)}{RSS}.$$

اگر فرضیهٔ H_0 صحیح باشد؛ یعنی متغیرهای توضیحی $X_{(m+1)i}$ تا X_{ki} نتوانند تغییرات Y_i را توضیح دهند، دو مدل مقید ٦-٢٩ و غیرمقید ٦-٣٠ درواقع برای جامعه مشاهدات متغیرها، با یکدیگر تفاوت ندارند و یکی هستند. در چنین حالتی خواهیم داشت $RSS_r = RSS$؛ بنابراین کسر فوق برابر صفر می‌شود. البته اگر در عمل و با استفاده از نمونه‌گیری، این دو مدل را تخمین بزنیم، قطعاً نتایج تخمین پارامترها با یکدیگر متفاوت خواهد بود، که می‌توان گفت این تفاوت ناشی از خطای نمونه‌گیری است.

1. Restricted Regression Model

نتیجه می‌گیریم که مقدار کسر ۶.۳۱ یا صفر است یا عددی مثبت و هیچگاه منفی نخواهد بود. اگر صفر شد یا تفاوت مختصری با صفر داشت، فرضیهٔ H_0 را می‌توان قبول کرد. بدیهی است برای پاسخ دادن به این سؤال که این «تفاوت مختصر» در عمل دقیقاً چقدر است، باید از آزمونهای آماری استفاده کرد؛ بنابراین باید یک تفسیر آماری از کسر فوق ارائه دهیم. به ترتیب زیر عمل می‌کنیم.

الف) مدل رگرسیون ۶.۲۹ را تخمین زده و مجموع مربعات پسماند آن، یعنی RSS_r را به دست می‌آوریم.

ب) به همین ترتیب مقدار RSS را از مدل ۶.۳۰ حساب می‌کنیم.

ج) تفاوت $(RSS_r - RSS)$، درواقع برابرکاهشی است که درمجموع مربعات پسماند مدل ۶.۲۹ حاصل می‌شود و این کاهش به علت افزایش تعداد r متغیر توضیحی است.

د) براساس معادلهٔ ۲.۳۸ می‌دانیم که در یک مدل رگرسیون با k پارامتر، آمارهٔ $\frac{\sum e_i^2}{\sigma^2}$ دارای توزیع χ^2 با $(n-k)$ درجهٔ آزادی است؛ یعنی

$$\frac{RSS}{\sigma^2} \sim \chi^2 (n-k) . \qquad (۶.۳۲)$$

می‌گوییم که اگر فرضیهٔ H_0 صحیح باشد، یعنی درحقیقت مدلهای ۶.۲۹ و ۶.۳۰ با یکدیگر تفاوتی نداشته باشد، مقادیر واقعی واریانس جملهٔ اختلال در هریک از این دو معادله با یکدیگر برابر است. اگر صورت و مخرج کسر ۶.۳۱ را بر واریانس جملهٔ اختلال تقسیم کنیم، هرکدام دارای توزیع χ^2 خواهند بود. درجات آزادی هریک به صورت زیر است. درجهٔ آزادی صورت r است؛ زیرا درجات آزادی RSS_r و RSS به ترتیب برابر است با $(n-m)$ و $(n-k)$؛ بنابراین درجات آزادی تفاوت آنها عبارت خواهد بود از

$$(n-m)-(n-k)= k-m= r .$$

درجات آزادی مخرج، یعنی تغییرات توضیح داده نشده نیز برابر $(n-k)$ است.

ه) می‌دانیم نسبت دو توزیع χ^2 ـ که هریک بر درجات آزادی خود تقسیم شده باشد ـ توزیع F دارد؛ بنابراین اگر صورت و مخرج کسر به دست آمده در مرحلهٔ «د» را به ترتیب بر r و $(n-k)$ تقسیم کنیم، کسر به دست آمده توزیع F خواهد داشت. با توجه به

اینکه واریانس جملهٔ اختلال، یعنی σ^2 از صورت و مخرج حذف می‌شود، در نتیجه

$$\frac{(RSS_r - RSS)/r}{RSS/(n-k)} \sim F(r, n-k) . \qquad (۶.۳۳)$$

یادآوری می‌کنیم که رابطهٔ فوق به شرطی برقرار است که H_0 صحیح باشد.

و) با تعیین سطح معنی‌دار بودن آزمون، به راحتی می‌توان فرضیهٔ H_0 را آزمون کرد. مقدار F با درجات آزادی r و (n-k) را در سطح معنی‌دار α از جدول F به دست می‌آوریم و با آمارهٔ آزمون که در رابطهٔ ۶.۳۳ محاسبه شده است مقایسه می‌کنیم. اگر آمارهٔ آزمون از مقدار جدول F بیشتر بود معنی‌دار است و فرضیه H_0 رد می‌شود.

مثال ۶.۳ برای تخمین تابع مصرف، دو مدل به شرح زیر پیشنهاد شده است،

$$C_t = \beta_1 + \beta_2 Y_t + U_t ,$$

$$C_t = \beta_1 + \beta_3 Y_t + \beta_4 C_{t-1} + \beta_5 L_{t-1} + U_t ,$$

که در آن C_t مصرف، Y_t درآمد و L_t داراییهای نقدی است. برای بررسی تفاوت این دو مدل، می‌خواهیم فرضیه زیر را

$$H_0 : \beta_3 = \beta_5 = 0$$

در مقابل فرضیهٔ مخالف صفر و در سطح معنی‌دار یک درصد آزمون کنیم. همچنین می‌دانیم

$$n = ۳۱ , \quad RSS_r = ۱۳/۸۲۴ , \quad RSS = ۳/۵۸۴ .$$

ابتدا آمارهٔ آزمون را با استفاده از رابطهٔ ۶.۳۳ می‌نویسیم،

$$F = \frac{(RSS_r - RSS)/r}{RSS/(n-k)} .$$

در مثال فوق $r = ۲ , k = ۴$. بنابراین مقدار عددی آماره آزمون عبارت است از

$$F = \frac{(۱۳/۸۲۴ - ۳/۵۸۴)/۲}{۳/۵۸۴/(۳۱ - ۴)} = ۳۵/۵۷$$

مقدار F، با درجات آزادی (۲ , ۲۷) و در سطح معنی‌دار ۰/۰۱ برابر است با

$$F_{0/01}(2, 27) = 5/49 .$$

ملاحظه می‌شود آمارهٔ آزمون معنی‌دار است؛ زیرا در ناحیهٔ بحرانی قرار می‌گیرد. در نتیجه فرضیهٔ H_0 رد شده و ضرورت وجود C_{t-1} و L_{t-1} به عنوان متغیرهای توضیحی در تابع مصرف اثبات می‌شود.

۳. آمارهٔ آزمون F برحسب R^2

می‌توان رابطهٔ ۶ـ۳۳ را فقط برحسب ضرایب تعیین در مدل رگرسیون غیرمقید و مدل رگرسیون مقید نوشت. با توجه به معادلهٔ ۵ـ۴۸ می‌دانیم

$$R^2 = 1 - \frac{e'e}{y'y},$$

یا

$$RSS = (1 - R^2)\, TSS .$$

معادلهٔ فوق می‌تواند برای مدل غیرمقید مورد استفاده قرار گیرد. برای مدل مقید خواهیم داشت

$$RSS_r = (1 - R_r^2)\, TSS .$$

یادآوری می‌شود که مقدار $TSS = \Sigma y_i^2$ در هر دو مدل یکسان است. با جایگزینی معادله‌های فوق در رابطهٔ ۶ـ۳۳ داریم

$$\frac{[(1 - R_r^2)\, TSS - (1 - R^2)\, TSS] / r}{[(1 - R_r)\, TSS] / (n - k)} \sim F(r, n - k) ,$$

با حذف کل تغییرات (TSS) از صورت و مخرج خواهیم داشت

$$\frac{(n - k)}{r} \cdot \frac{R^2 - R_r^2}{1 - R^2} \sim F(r, n - k) . \qquad (۶ـ۳۴)$$

مثال ۶ـ۴ دو مدل مصرف (مثال ۶ـ۳) را یک بار دیگر ملاحظه کنید،

$$C_t = \beta_1 + \beta_2 Y_t + U_t ,$$

$$C_t = \beta_1 + \beta_2 Y_t + \beta_3 C_{t-1} + \beta_4 L_{t-1} + U_t .$$

فرض کنید این دو مدل را تخمین زده و مقادیر ضریب تعیین تعدیل شده زیر را به دست آورده‌ایم،

$$\bar{R}^2 = ۰/۹۸۴ \qquad \bar{R}_r^2 = ۰/۹۴۴ .$$

با توجه به $n = ۳۱$، فرضیهٔ $H_0: \beta_3 = \beta_4 = ۰$ را در سطح معنی‌دار ۰/۰۱ و در مقابل فرضیهٔ مخالف صفر آزمون کنید.

ابتدا باید R^2 را حساب کنیم. معادلهٔ ۶۴ـ۵ را نوشته، خواهیم داشت

$$\bar{R}^2 = ۱ - \left(\frac{n-۱}{n-k}\right)(۱-R^2) ,$$

$$۰/۹۸۴ = ۱ - \left(\frac{۳۰-۱}{۳۰-۴}\right)(۱-R^2) ,$$

$$R^2 = ۰/۹۸۶ .$$

برای مدل مقید داریم

$$\bar{R}_r^2 = ۱ - \left(\frac{n-۱}{n-k}\right)(۱-R_r^2) ,$$

$$۰/۹۴۴ = ۱ - \left(\frac{۳۱-۱}{۳۱-۲}\right)(۱-R_r^2) ,$$

$$R_r^2 = ۰/۹۴۶$$

در معادلهٔ ۶۴ـ۳۴ آمارهٔ آزمون را برحسب R^2 و R_r^2 به دست آوردیم، یعنی

$$\frac{(n-k)}{r} \cdot \frac{R^2 - R_r^2}{۱-R^2} \sim F(r, n-k) ,$$

بنابراین

$$F = \frac{(۳۱-۴)}{۲} \cdot \frac{۰/۹۸۶ - ۰/۹۴۶}{۱-۰/۹۸۶} = \frac{۲۷}{۲}\left(\frac{۰/۰۴}{۰/۰۱۴}\right) = ۳۸/۵۷ .$$

با توجه به اینکه ۵/۴۹ =$F_{۰/۰۱}$ (۲ , ۲۷)؛ بنابراین آمارهٔ آزمون معنی‌دار بوده و فرضیهٔ $H_۰$ رد می‌شود.

۴. آزمون معنی‌دار بودن β_i با آمارهٔ F

قبلاً در معادلهٔ ۶-۲۸ دیدیم که چگونه می‌توان آزمون‌معنی‌دار بودن β_iرا به کمک آمارهٔ t انجام داد. در این قسمت نشان می‌دهیم که این آزمون، حالت خاصی از آزمون معنی‌دار بودن زیرمجموعه‌ای از پارامترهاست. مدل زیر را در نظر می‌گیریم،

$$Y_t = \beta_۱ + \beta_۲ X_{۲t} + \cdots + \beta_k X_{kt} + U_t .$$

می‌خواهیم فرضیهٔ $H_۰: \beta_i = ۰$ را در مقابل فرضیهٔ $H_۱: \beta_i \neq ۰$ در سطح معنی‌دار α درصد آزمون کنیم.

مدل فوق را مدل غیرمقید می‌نامیم. فرضیهٔ $\beta_i = ۰$را در این مدل قرار می‌دهیم تا مدل مقید به دست آید. هر دو مدل را تخمین می‌زنیم و به کمک RSS و RSS_r، آمارهٔ F را ـ که در ۶-۳۳ تعریف شده است ـ می‌سازیم. با مقایسهٔ آمارهٔ آزمون، با مقدار موجود در جدول F نسبت به فرضیهٔ $H_۰$، استنباط آماری خواهیم کرد. بنابراین، مراحل آزمون به صورت زیر است.

الف) متغیر توضیحی X_{it} را از مدل حذف کرده، مدل مقید به دست آمده را تخمین می‌زنیم تا RSS_r به دست آید.

ب) مدل مفروض را با وجود X_{it} تخمین می‌زنیم.

ج) مقدار کاهش در مجموع مربعات پسماند را که از مرحله (الف) به (ب) ملاحظه می‌شود حساب می‌کنیم؛ یعنی $(RSS_r - RSS)$.

د) آمارهٔ آزمون را به شرح زیر می‌سازیم،

$$\frac{(RSS_r - RSS) / ۱}{RSS / (n-k)} \sim F(۱ , n-k) ,$$

ه) از مقایسهٔ آماره آزمون با مقدار جدول F، می‌توان در سطح معنی‌دار

α نسبت به فرضیهٔ مورد نظر استنباط آماری کرد.

مثال ۶ـ۵ مدل رگرسیون زیر را که موضوع مثال ۶ـ۲ بوده است یک بار دیگر ملاحظه کنید،

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t .$$

در بند «د» مثال مزبور، فرضیهٔ $\gamma = 0$: H_0 را در مقابل فرضیهٔ $\gamma \neq 0$: H_1 در سطح معنی‌دار ۵ درصد آزمون کرده‌ایم. همین آزمون را با استفاده از آمارهٔ F انجام دهید.

برای این منظور، ابتدا متغیر توضیحی Z_t را از مدل حذف کرده و مدل مقید به دست آمده را تخمین می‌زنیم. با استفاده از مثال ۲ـ۵ و محاسبات جدول ۳ـ۵ می‌دانیم که $\Sigma x_t y_t = ۱۶$ و $\Sigma x_t^2 = ۱۰$.

خواهیم داشت

$$\hat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} = \frac{۱۶}{۱۰} = ۱/۶$$

باید مقدار RSS_r را برای این مدل مقید حساب کنیم. با توجه به معادلهٔ ۱ـ۳۵ داریم

$$R_r^2 = \frac{\hat{\beta}\,\Sigma x_t y_t}{\Sigma y_t^2} = \frac{۱/۶(۱۶)}{۲۸} = ۰/۹۱۴۳ .$$

طبق معادلهٔ ۱ـ۴۳ می‌دانیم

$$\Sigma e_t^2 = (۱ - r^2)\,\Sigma y_t^2 ,$$

در نتیجه برای این مثال خواهیم داشت

$$RSS_r = (۱ - R_r^2)\,TSS ,$$

$$= (۱ - ۰/۹۱۴۳)\,۲۸ = ۲/۴ .$$

حال مدل مفروض را با وجود X_t تخمین زده و تغییرات توضیح داده نشده را به‌دست می‌آوریم. البته این محاسبات در مثال ۶ـ۲ انجام شده است، داریم $RSS = ۱/۵$.

مقدار کاهش در مجموع مربعات پسماند برابر است با ۰/۹ = ۱/۵ - ۲/۴ . علت این بهبود وجود متغیر توضیحی X_t در مدل است که توانسته قدرت توضیحی مدل را افزایش دهد. بدین ترتیب آمارهٔ F را به شرح زیر محاسبه می‌کنیم،

$$F = \frac{(RSS_r - RSS) / 1}{RSS / (n - k)},$$

$$= \frac{(2/4 - 1/5) / 1}{1/5 / (5 - 3)} = \frac{0/9}{0/75} = 1/2 .$$

مقدار F در سطح معنی‌دار ۵ درصد برابر است با $F_{0/05}(1 , 2) = 18/51$؛ بنابراین آمارهٔ آزمون معنی‌دار نیست؛ زیرا در ناحیهٔ بحرانی قرار نگرفته و در نتیجه فرضیه H_0 رد نمی‌شود. قبلاً در مثال ۲-۶ که همین مسأله را با آمارهٔ t حل کردیم - مقدار t برابر ۱/۱- به‌دست آمد، درحالی‌که در اینجا مقدار F برابر ۱/۲ شده است. بدیهی است چنین نتیجه‌ای نشان دهندهٔ درستی محاسبات است؛ زیرا با توجه به معادلهٔ ۲-۵۱ می‌دانیم

$$F = t^2 ,$$

بنابراین، $\sqrt{1/2} = 1/1$.

۶-۴ آنالیز واریانس و آزمون معنی‌دار بودن مدل رگرسیون

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t .$$

در این قسمت می‌خواهیم این فرضیه را آزمون کنیم که هیچکدام از متغیرهای توضیحی بر تغییرات Y_t تأثیری ندارد؛ یعنی

$$H_0 : \beta_2 = \beta_3 = \cdots = \beta_k = 0 .$$

فرضیهٔ مخالف، یعنی H_1 این است که $H_0 = 0$ صحیح نیست؛ یعنی حداقل یک متغیر

توضیحی وجود دارد که در تغییرات Y_i مؤثر واقع می‌شود؛ به عبارت دیگر، فرضیهٔ H_1 بر این دلالت دارد که حداقل یک β وجود دارد که صفر نیست. اگر فرضیهٔ H_0 صحیح باشد، تغییرات Y_i صرفاً تصادفی بوده، فقط تابعی از U_i خواهد بود، زیرا در چنین حالتی خواهیم داشت

$$Y_i = \beta_1 + U_i .$$

توجه به این نکته اهمیت فراوان دارد که آزمون «تمام پارامترها» برابر صفر غیر از آزمونهای جداگانه «هریک از پارامترها» برابر صفر است. این نکته را در ادامهٔ مباحث این قسمت مطرح خواهیم کرد.

همان‌گونه که در فصل دوم قسمت ۳ـ۲ دیدیم، آزمون معنی‌دار بودن کل رگرسیون را به کمک تجزیهٔ مجموع تغییرات Y_i، یعنی Σy_i^2، به تغییرات توضیح‌داده‌شده $(\Sigma \hat{y}_i^2)$ و تغییرات توضیح داده نشده (Σe_i^2) انجام می‌دهند. با توجه به معادلهٔ ۴۵ـ۵ می‌دانیم

$$\mathbf{y'y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \mathbf{e'e} .$$

یکی از اهداف مسألهٔ تجزیه واریانس آزمون معنی‌دار بودن تغییرات توضیح داده شده (قدرت توضیحی مدل) است. در رگرسیون ساده، این آزمون دقیقاً همان $H_0: \beta = 0$ ست؛ زیرا در این گونه مدلها فقط یک متغیر توضیحی وجود دارد، اما در رگرسیون چندمتغیره، به علت تعدد متغیرهای توضیحی، این گونه نیست و در ادامهٔ بحث خواهیم دید که باید از آزمون F استفاده شود؛ بنابراین ملاحظه می‌شود که جدول تجزیهٔ واریانس اطلاعاتی دارد که مستقیماً در آزمون معنی‌دار بودن رگرسیون مورد نیاز است.

مانند جدول ۱ـ۲ که تجزیهٔ واریانس را برای رگرسیون ساده نشان می‌دهد کل تغییرات متغیر درون‌زا $(\mathbf{y'y})$، تغییرات توضیح داده شده $(\hat{\mathbf{y}}'\hat{\mathbf{y}})$ و تغییرات توضیح داده نشده یا مجموع مربعات پسماند $(\mathbf{e'e})$ را همراه با درجات آزادی و میانگین هرکدام در جدول زیر (۶ـ۱) تنظیم کرده، آن را جدول تجزیهٔ واریانس برای رگرسیون چندمتغیره می‌نامیم.

جدول ۶ـ۱ آنالیز واریانس برای مدل رگرسیون چندمتغیره

منبع تغییرات	مجموع مربعات	درجات آزادی	میانگین مربعات یا واریانس
تخمین رگرسیون	$ESS=\hat{y}'\hat{y}=\hat{\beta}'X'y$	$(k-1)$	$ESS/(k-1)$
پسماند	$RSS=e'e=y'y-\hat{\beta}'X'y$	$(n-k)$	$RSS/(n-k)$
کل تغییرات	$TSS=y'y$	$(n-1)$	$F=\frac{ESS/(k-1)}{RSS/(n-k)}$

برای تبیین بیشتر جدول تجزیهٔ واریانس به توضیحات زیر اشاره می‌کنیم:

الف) معادلهٔ $ESS=\hat{y}'\hat{y}=\hat{\beta}'X'y$ درواقع همان معادلهٔ ۵ـ۳۹ است که قبلاً ثابت شده است.

ب) برای نشان دادن درستی رابطهٔ $RSS=e'e=y'y-\hat{\beta}'X'y$، ابتدا معادلهٔ ۵ـ۴۶ را می‌نویسیم،

$$e'e=y'y-y'X(X'X)^{-1}X'y.$$

با توجه به معادلهٔ ۵ـ۳۴ می‌دانیم

$$y'X(X'X)^{-1}=\hat{\beta}'$$

در نتیجه خواهیم داشت

$$e'e=y'y-\hat{\beta}'X'y \qquad (۶ـ۳۵)$$

ج) برای تبیین این نکته که نسبت $\frac{ESS/(k-1)}{RSS/(n-k)}$ دارای توزیع F با $(k-1)$ و $(n-k)$ درجه آزادی است، ابتدا معادلهٔ ۵ـ۵۱ را یک بار دیگر می‌نویسیم،

$$F=\frac{ESS/1}{RSS/(n-2)}\sim F(1,n-2).$$

می‌دانیم معادلهٔ فوق برای مدل رگرسیونی است که فقط دو پارامتر دارد. می‌توان مشابه این فرمول را برای یک مدل رگرسیون با k پارامتر نوشت. خواهیم داشت

$$F=\frac{\hat{y}\,\hat{y}'/(k-1)}{e'e/(n-k)}=\frac{ESS/(k-1)}{RSS/(n-k)}\sim F(k-1,n-k). \qquad (۶ـ۳۶)$$

با وجود این، برای اینکه اشاره‌ای به نحوهٔ استخراج این فرمول داشته باشیم، می‌گوییم آمارهٔ F نسبت دو توزیع مستقل χ^2 است که هریک بر درجات آزادی خود تقسیم شده‌اند. با توجه به تعریف توزیع χ^2 در معادلهٔ ۲ـ۳۷ می‌توان گفت که

$$\frac{(\hat{\beta}-\beta)^2}{\mathrm{Var}(\hat{\beta})} \sim \chi^2(k-1). \qquad (۶ـ۳۷)$$

یادآوری می‌کنیم که فرض ما، در این قسمت این است که تمام محاسبات برحسب انحراف از میانگین انجام شده است؛ بنابراین بردار β در رابطهٔ (۶ـ۳۶) شامل $(k-1)$ پارامتر است. همچنین مشابه معادلهٔ ۲ـ۳۸، می‌توان نشان داد که

$$\frac{e'e}{\sigma_u^2} \sim \chi^2(n-k) \qquad (۶ـ۳۸)$$

اگر آماره‌های موجود در رابطه‌های ۶ـ۳۷ و ۶ـ۳۸ را بر درجات آزادی خود تقسیم کنیم، نسبت آنها به یکدیگر توزیع F خواهدداشت. با توجه به $\mathrm{Var}(\hat{\beta}) = \sigma^2(X'X)^{-1}$، داریم

$$\frac{[(\hat{\beta}-\beta)^2/\sigma^2(X'X)^{-1}]/(k-1)}{[e'e/\sigma^2]/(n-k)} \sim F(k-1, n-k).$$

اگر فرضیهٔ H_0 صحیح باشد، یعنی $\beta_2 = \beta_3 = \cdots = \beta_k = 0$؛ آنگاه $\beta = 0$، در نتیجه خواهیم داشت

$$\frac{\hat{\beta}^2(X'X)/(k-1)}{e'e/(n-k)} \sim F(k-1, n-k),$$

$$\frac{[\hat{\beta}'(X'X)\hat{\beta}]/(k-1)}{e'e/(n-k)} \sim F(k-1, n-k),$$

با استفاده از معادلهٔ ۵ـ۳۵، یعنی $\hat{y} = X\hat{\beta}$، داریم

$$\hat{y}'\hat{y} = \hat{\beta}'X'X\hat{\beta},$$

. و با جایگزینی معادلهٔ فوق در آمارهٔ F خواهیم داشت

$$F=\frac{\hat{\mathbf{y}}\hat{\mathbf{y}}'(k-1)}{\mathbf{e}'\mathbf{e}(n-k)}=\frac{ESS/(k-1)}{RSS/(n-k)}\sim F(k-1,n-k),$$

که دقیقاً همان رابطهٔ ۶ـ۳۶ است. یادآوری می‌شود که این رابطه، براساس $\beta=0$ به دست آمده و بنابراین موقعی برقرار است که H_0 صحیح باشد. برای آزمون درستی H_0، کافی است آمارهٔ آزمون را که از رابطهٔ ۶ـ۳۶ به دست می‌آید با مقدار جدول F و در سطح معنی‌دار α درصد مقایسه کنیم. اگر آمارهٔ آزمون معنی‌دار بود؛ یعنی در ناحیهٔ بحرانی قرار گرفت، H_0 رد می‌شود.

مثال ۶ـ۶ مدل رگرسیون زیر را که موضوع مثال ۶ـ۲ بوده است دوباره ملاحظه کنید،

$$Y_t=\alpha+\beta X_t+\gamma Z_t+U_t.$$

می‌خواهیم فرضیهٔ $H_0:\beta=\gamma=0$ را در سطح معنی‌دار ۵ درصد آزمون کنیم.

آمارهٔ آزمون را با استفاده از رابطهٔ ۶ـ۳۶ می‌نویسیم،

$$\frac{ESS/(k-1)}{RSS/(n-k)}.$$

با استفاده از اطلاعات موجود در مثال ۶ـ۲ می‌دانیم ۲۶/۵ =تغییرات توضیح داده شده و ۱/۵ =تغییرات توضیح داده نشده، $n=5$، $k=3$، در نتیجه

$$F=\frac{۲۶/۵/(۳-۱)}{۱/۵/(۵-۳)}=۱۷/۶۷$$

مقدار F را با درجات آزادی ۲ و ۲ در سطح معنی‌دار ۵ درصد از جدول به دست می‌آوریم،

$$F_{۵٪}(۲,۲)=۱۹,$$

در نتیجه آمارهٔ آزمون کمتر از مقدار F در جدول است و معنی‌دار نیست؛ بنابراین فرضیهٔ H_0 رد نمی‌شود.

۱. آمارهٔ آزمون F برحسب R^2

می‌توان رابطهٔ ۶ـ۳۶ را فقط برحسب R^2 نوشت. در چنین حالتی محاسبات آزمون معنی‌دار بودن مدل رگرسیون بسیار ساده‌تر خواهد بود. با توجه به معادلهٔ ۴۸ـ۵ می‌توان چنین نوشت

$$RSS = e'e = (1 - R^2)\, y'y, \qquad (۶ـ۳۹)$$

$$ESS = y'y = R^2 y'y \qquad (۶ـ۴۰)$$

با جایگزینی معادله‌های ۶ـ۳۹ و ۶ـ۴۰ در معادلهٔ ۶ـ۳۶ داریم

$$F = \frac{[R^2 y'y]/(k-1)}{[(1-R^2)\, y'y]/(n-k)} \sim F(k-1, n-k),$$

بعد از ساده کردن خواهیم داشت

$$F = \left(\frac{n-k}{k-1}\right) \frac{R^2}{(1-R^2)} \sim F(k-1, n-k). \qquad (۶ـ۴۱)$$

ملاحظه می‌شود که برای آزمون معنی‌دار بودن یک مدل رگرسیون، فقط کافی است مقدار R^2 را بدانیم.

فرمول ۶ـ۴۱ را می‌توان از راه دیگری نیز به دست آورد. آزمون صفر بودن زیرمجموعه‌ای از پارامترهای یک مدل رگرسیون را که در معادلهٔ ۶ـ۳۴ مطرح کردیم یک بار دیگر می‌نویسیم،

$$F = \frac{(n-k)}{r} \cdot \frac{R^2 - R_r^2}{1 - R^2} \sim F(r, n-k).$$

بحث ما در معادلهٔ فوق این بود که می‌خواستیم فرضیهٔ صفر بودن r پارامتر را آزمون کنیم. اگر بخواهیم فرضیهٔ صفر بودن تمام پارامترها، غیر از جملهٔ ثابت را آزمون کنیم، یعنی

$$H_0: \beta_2 = \beta_3 = \cdots = \beta_k = 0,$$

r برابر $(k-1)$ خواهد بود. از طرف دیگر، بر فرض درستی فرضیهٔ H_0، می‌توان مدل

رگرسیون مقید را به صورت زیر نوشت

$$Y_t = \beta_1 + U_t .$$

چون هیچ متغیر برون‌زا برای توضیح دادن تغییرات Y_t وجود ندارد؛ بنابراین R_r^2 برابر صفر می‌شود. مقادیر $r = (k-1)$ و $R_r^2 = 0$ را در معادلهٔ ۶.۳۴ قرار می‌دهیم،

$$F = \left(\frac{n-k}{k-1}\right) . \frac{R^2}{(1-R^2)} \sim F(k-1, n-k) ,$$

که دقیقاً همان رابطهٔ ۶.۴۱ است.

مثال ۶.۷ مدل رگرسیون زیر را که موضوع مثال ۶.۶ بوده است ملاحظه کنید،

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t .$$

فرضیهٔ $H_0 : \beta = \gamma = 0$ را در سطح معنی‌دار ۵ درصد آزمون کنید. می‌دانیم $R^2 = 0/9464$

با استفاده از مقدار R^2، آمارهٔ آزمون را از رابطهٔ ۶.۴۱ محاسبه کرده و با مقدار به دست آمده از جدول F مقایسه می‌کنیم. خواهیم داشت

$$F = \left(\frac{n-k}{k-1}\right) \frac{R^2}{1-R^2} = \frac{5-3}{3-1} \left(\frac{0/9464}{1-0/9464}\right) = 17/6 ,$$

که با توجه به $F_{0/05}(2,2) = 19$ معنی‌دار نیست و فرضیهٔ H_0 رد نمی‌شود.

۲. آزمون معنی‌دار بودن کل مدل رگرسیون برای مدلهای فاقد جمله ثابت

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t .$$

می‌خواهیم فرضیه زیر را آزمون کنیم،

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0 .$$

فرضیهٔ مخالف این است که حداقل یکی از پارامترها صفر نیست. ملاحظه می‌شود که

مدل رگرسیون فاقد جمله ثابت است. از آزمون والد استفاده می‌کنیم. معادلهٔ ۶۳ـ۶ را دوباره می‌نویسیم،

$$F=\frac{(RSS_r - RSS_U)/r}{RSS_U/(n-k)} \sim F(r, n-k)$$

با جایگزینی H_0 در مدل رگرسیون مفروض، مدل مقید به شرح زیر به دست می‌آید،

$$Y_t = U_t .$$

با توجه به اینکه در مدل مقید هیچ متغیر توضیحی وجود ندارد، مدل اساساً فاقد قدرت توضیحی بوده در نتیجه مجموع مربعات پسماند برابر مجموع تغییرات متغیر درون‌زا می‌شود؛ یعنی $RSS_r = \Sigma Y_t^2$ یا $\Sigma e_t^2 = \Sigma Y_t^2$. از طرف دیگر، با توجه به معادلهٔ ۵۷ـ۵ داریم

$$\Sigma Y_t^2 = \Sigma \hat{Y}_t^2 + \Sigma e_t^2 .$$

بنابراین برای مدل غیرمقید می‌توان چنین نوشت

$$\Sigma e_t^2 = RSS_U = \Sigma Y_t^2 - \Sigma \hat{Y}_t^2 ,$$

با استفاده از تساوی $RSS_r = \Sigma Y_t^2$ خواهیم داشت

$$RSS_U - RSS_r = \Sigma \hat{Y}_t^2 .$$

با جایگزینی رابطهٔ فوق در صورت کسر آمارهٔ F داریم

$$F=\frac{\Sigma \hat{Y}_t^2/(k-1)}{RSS_U/(n-k+1)} = \frac{\Sigma \hat{Y}_t^2/(k-1)}{\Sigma e_t^2/(n-k+1)} \sim F(k-1, n-k+1) ,$$

اگر آمارهٔ مزبور معنی‌دار باشد، فرضیهٔ H_0 رد می‌شود.

۳. مقایسهٔ آزمون معنی‌دار بودن کل رگرسیون و معنی‌دار بودن هریک از پارامترها*

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + \beta_k X_{kt} + U_t .$$

دو فرضیهٔ زیر را به ترتیب با آمارهٔ F و آمارهٔ t آزمون می‌کنیم،

$$H_0 : \beta_2 = \beta_3 = \cdots = \beta_k = 0$$

$$H_0 : \beta_i = 0 , \qquad i = 2, 3, \ldots, k .$$

آیا اگر یکی از این دو فرضیه قبول یا رد شود، فرضیهٔ دیگر نیز ضرورتاً قبول یا رد خواهد شد؟

ابتدا پاسخ را مطرح کرده، سپس به اثبات آن می‌پردازیم. در جواب باید گفت که ضرورتاً این گونه نیست؛ زیرا آزمونهای F و t، از نظر استنباط آماری، قابل مقایسه با یکدیگر نیستند، چون بر فرضهای متفاوتی مبتنی هستند. برای مثال، وقتی بخواهیم $\beta_i = 0$ را با آمارهٔ t آزمون کنیم، وجود هیچ فرضی برای β_j ضروری نیست. درواقع با آماره t، آزمون هریک از پارامترها مستقل از دیگری انجام می‌شود، در حالی که آزمون F مستلزم ملاحظهٔ همزمان تمام پارامترهاست. برای اثبات این قضیه کافی است آمارهٔ F را برحسب آمارهٔ t بنویسیم. بحث را در قالب یک مدل رگرسیون با دو متغیر توضیحی ادامه می‌دهیم.

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + U_t .$$

سؤال این است که آیا فرضیهٔ $H_0 : \beta_2 = \beta_3 = 0$ و فرضیهٔ $H_0 : \beta_i = 0$ و $i = 2, 3$، که اولی با آمارهٔ F و دومی با آمارهٔ t آزمون می‌شود، ضرورتاً به یک نتیجه منتهی می‌گردد؟ برای این منظور ابتدا مقادیر تغییرات توضیح داده شده و تغییرات توضیح داده نشده را محاسبه کرده و سپس آنها را در معادلهٔ ۶-۳۶ قرار می‌دهیم تا رابطهٔ F با t به دست آید. با توجه به

تعریف تغییرات توضیح داده شده و نیز معادله‌های ۳۰ـ۱ و ۶ـ۴ داریم

$$\begin{aligned} ESS &= \Sigma (\hat{Y}_t - \bar{\hat{Y}})^2 = \Sigma (\hat{Y}_t - \bar{Y})^2 , \\ &= \Sigma (\hat{\beta}_1 + \hat{\beta}_2 X_{2t} + \hat{\beta}_3 X_{3t} - \bar{Y})^2 , \\ &= \Sigma [(\bar{Y} - \hat{\beta}_2 \bar{X}_2 - \hat{\beta}_3 \bar{X}_3) + \hat{\beta}_2 X_{2t} + \hat{\beta}_3 X_{3t} - \bar{Y}]^2 , \\ &= \Sigma (\hat{\beta}_2 x_{2t} + \hat{\beta}_3 x_{3t})^2 . \end{aligned}$$

بدین ترتیب خواهیم داشت

$$ESS = \hat{\beta}_2^2 \Sigma x_{2t}^2 + \hat{\beta}_3^2 \Sigma x_{3t}^2 + 2\hat{\beta}_2 \hat{\beta}_3 \Sigma x_{2t} x_{3t} . \qquad (6\text{ـ}42)$$

از معادلهٔ ۲۱ـ۶ استفاده کرده واریانس U_t را برای این حالت می‌نویسیم،

$$\hat{\sigma}_U^2 = \frac{e'e}{n-3} = \frac{RSS}{n-3} . \qquad (6\text{ـ}43)$$

معادله‌های ۴۲ـ۶ و ۴۳ـ۶ را در آمارهٔ F در معادلهٔ ۳۶ـ۶ قرار می‌دهیم. خواهیم داشت

$$F = \frac{\hat{\beta}_2^2 \Sigma x_{2t}^2 + \hat{\beta}_3^2 \Sigma x_{3t}^2 + 2\hat{\beta}_2 \hat{\beta}_3 \Sigma x_{2t} x_{3t}}{2\hat{\sigma}^2} . \qquad (6\text{ـ}44)$$

حال به بررسی آزمون t می‌پردازیم. آمارهٔ t برای آزمون $\beta_2 = 0$ و آمارهٔ t برای آزمون $\beta_3 = 0$ به ترتیب برابر است با

$$t_2 = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)} , \qquad t_3 = \frac{\hat{\beta}_3}{SE(\hat{\beta}_3)} .$$

مقادیر فوق را مجذور کرده داریم

$$t_2^2 = \frac{\hat{\beta}_2^2}{Var(\hat{\beta}_2)} , \qquad t_3^2 = \frac{\hat{\beta}_3^2}{Var(\hat{\beta}_3)} . \qquad (6\text{ـ}45)$$

با توجه به معادله‌های ٤ـ٢٤ و ٤ـ٢٥ می‌دانیم که

$$\mathrm{Var}\,(\hat{\beta}_2)=\frac{\hat{\sigma}^2}{(1-r_{23}^2)\,\Sigma x_{21}^2}, \qquad \mathrm{Var}\,(\hat{\beta}_3)=\frac{\hat{\sigma}^2}{(1-r_{23}^2)\,\Sigma x_{31}^2}.$$

معادله‌های فوق را در معادلهٔ ٦ـ٤٥ قرار می‌دهیم، خواهیم داشت

$$t_2^2=\frac{\hat{\beta}_2^2\,(1-r_{23}^2)\,\Sigma x_{21}^2}{\hat{\sigma}^2}, \qquad t_3^2=\frac{\hat{\beta}_3^2\,(1-r_{23}^2)\,\Sigma x_{31}^2}{\hat{\sigma}^2}.$$

مقادیر $\hat{\beta}_2^2$ و $\hat{\beta}_3^2$ را از معادله‌های فوق به دست می‌آوریم،

$$\hat{\beta}_2^2=\frac{t_2^2\,\hat{\sigma}^2}{(1-r_{23}^2)\,\Sigma x_{21}^2}, \qquad \hat{\beta}_3^2=\frac{t_3^2\,\hat{\sigma}^2}{(1-r_{23}^2)\,\Sigma x_{31}^2}.$$

با استفاده از دو معادلهٔ فوق می‌توان $\hat{\beta}_2\ \hat{\beta}_3$ را به دست آورد،

$$\hat{\beta}_2\,\hat{\beta}_3=\frac{t_2\,t_3\,\hat{\sigma}_2}{(1-r_{23}^2)\sqrt{\Sigma x_{21}^2\,\Sigma x_{31}^2}}.$$

مقادیر $\hat{\beta}_2$، $\hat{\beta}_3$ و $\hat{\beta}_2\ \hat{\beta}_3$ را در معادلهٔ ٦ـ٤٤ جایگزین می‌کنیم. در نتیجه

$$F=\frac{t_2^2+t_3^2+2\,(t_2\,t_3\,\Sigma x_{21}\,x_{31})\,/\sqrt{\Sigma x_{21}^2\,\Sigma x_{31}^2}}{2\,(1-r_{23}^2)},$$

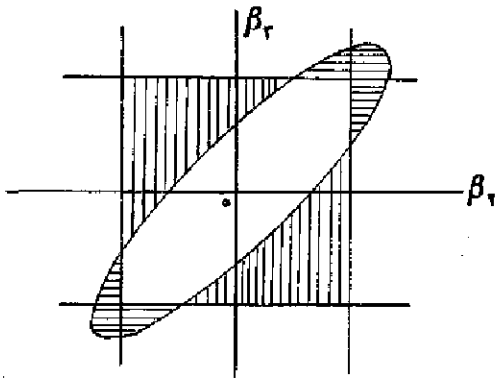
و با توجه به معادلهٔ ١ـ٣٢ داریم

$$F=\frac{t_2^2+t_3^2+2t_2\,t_3\,r_{23}}{2\,(1-r_{23}^2)}\sim F(2,n-3) \qquad (٦ـ٤٦)$$

فرض می‌کنیم که مقادیر t_2 و t_3 کوچک هستند، به گونه‌ای که از مقادیر t در جدول کمتر شده، معنی‌دار نمی‌شوند. در نتیجه فرضیه‌های $H_0: \beta_2=0$ و $H_0: \beta_3=0$ هر دو قبول خواهد شد. اما اگر متغیرهای توضیحی مدل، یعنی X_{21} و X_{31}، از شدت همبستگی تقریباً زیادی برخوردار باشد، یعنی مقدار r_{23}^2 به یک نزدیک شود، آنگاه $(1-r_{23}^2)$ عدد بسیار

کوچکی شده، بنابراین کسر F بسیار بزرگ خواهد شد، به گونه‌ای که به راحتی از مقدار جدول F تجاوز کرده، فرضیهٔ $H_0: \beta_2 = \beta_3 = 0$ قبول می‌شود. به همین ترتیب می‌توان گفت که در مواردی ممکن است مقادیر $\hat{\beta}_2$ یا $\hat{\beta}_3$ یا هر دو به طور معنی‌داری با صفر متفاوت باشد، در حالی که آمارهٔ F، معنی‌دار نیست. علت این امر تفاوت ماهوی آزمونهای F و t است. وقتی بخواهیم فرضیهٔ $\beta_2 = 0$ را با آمارهٔ t آزمون کنیم، هیچ ضرورتی ندارد که فرضهایی دربارهٔ β_3 مطرح کنیم. اساساً آزمونهای t برای پارامترهای مختلف مستقل از یکدیگر انجام می‌شود، در حالی که آزمون F براساس فرضیه‌ای شکل می‌گیرد که شامل تمام پارامترهای متغیرهای توضیحی است.

تفاوت در استنباط آماری حاصل از آزمونهای F و t را می‌توان در یک نمودار نیز نشان داد. وقتی بخواهیم فرضیهٔ $H_0: \beta_2 = \beta_3 = 0$ را آزمون کنیم درواقع به جای یک فاصلهٔ اطمینان، باید از یک «ناحیهٔ اطمینان»[1] صحبت کرد؛ یعنی ناحیه‌ای که اگر آمارهٔ آزمون F در آن واقع شود، معنی‌دار نیست و H_0 قبول می‌شود. در آزمون t نیز دو فاصلهٔ اطمینان برای β_2 و β_3 به دست می‌آوریم که در نتیجه یک ناحیهٔ اطمینان به وجود می‌آورند. همان گونه که در نمودار ۶.۱ ملاحظه می‌شود قسمتهایی وجود دارد که این دو ناحیهٔ اطمینان بر هم منطبق نیست. در بند (۶) همین قسمت نشان خواهیم داد که اگر



نمودار ۶.۱ نواحی اطمینان برای آزمونهای F و t

1. Confidence Region

بخواهیم با آمارهٔ F ، دو فرضیه را همزمان آزمون کنیم، ناحیهٔ اطمینان، یک بیضی خواهد بود. همچنین می‌توان نشان داد که هرقدر همبستگی دو متغیر توضیحی در مدل رگرسیون بیشتر باشد، این بیضی باریکتر خواهد بود و برعکس. در نمودار فوق، قسمتهایی که با هاشور عمودی مشخص شده است، ناحیهٔ اطمینان برای $\beta_۲= ۰$ و $\beta_۳= ۰$ است که از آمارهٔ t به دست می‌آید. ملاحظه می‌شود که فقط ناحیهٔ هاشور نخوردهٔ داخل بیضی، فصل مشترک نواحی اطمینان آماره‌های F و t است. اگر آمارهٔ F در قسمتهایی قرار گیرد که با هاشورهای افقی مشخص شده است، فرضیهٔ $H_.$ قبول می‌شود، در حالی که دقیقاً در همین نواحی فرضیهٔ $H_.$ با آزمون t رد خواهد شد. به همین ترتیب در نواحی مشخص شده با هاشورهای عمودی، فرضیهٔ $H_.$ با آزمون t قبول شده و با آزمون F رد می‌شود؛ بنابراین، با افزایش همبستگی بین متغیرهای توضیحی، نواحی هاشور زده شده، بزرگتر می‌شود در نتیجه تفاوت نتایج حاصل از آزمونهای F و t بزرگتر خواهد شد.

مثال ۶.۸ مدل رگرسیون زیر مفروض است،

$$Y_t= \beta_۱ + \beta_۲ X_{۲t} + \beta_۳ X_{۳t} + U_t .$$

با استفاده از ۲۰ مشاهده، کمیتهای زیر را برحسب انحراف از میانگین محاسبه کرده‌ایم.

$$\Sigma x_{۲t}^۲= ۱۰۰ , \qquad \Sigma y_t^۲= ۱۰۰ ,$$

$$\Sigma x_{۳t}^۲= ۱۰۰ , \qquad \Sigma y_t x_{۲t}= ۹۰ ,$$

$$\Sigma x_{۲t} x_{۳t}= ۹۵ , \qquad \Sigma y_t x_{۳t}= ۹۰ ,$$

اولاً، پارامترهای $\beta_۲$ و $\beta_۳$ را تخمین بزنید.

ثانیاً، با استفاده از آزمون t نشان دهید که هریک از دو فرضیهٔ $\beta_۲= ۰$ و $\beta_۳= ۰$ رد نمی‌شود.

ثالثاً، با استفاده از آزمون F نشان دهید دو فرضیهٔ $\beta_۲= ۰$ و $\beta_۳= ۰$ به طور همزمان رد می‌شود؛ یعنی: $H_.: \beta_۲= \beta_۳= ۰$.

۱. از معادله‌های ۱۰ـ۴ داریم

$$\hat{\beta}_2 = \frac{\Sigma x_{3t}^2 \Sigma x_{1t} y_t - \Sigma x_{1t} x_{3t} \Sigma x_{3t} y_t}{\Sigma x_{2t}^2 \Sigma x_{3t}^2 - (\Sigma x_{2t} \Sigma x_{3t})^2},$$

$$\hat{\beta}_3 = \frac{\Sigma x_{1t}^2 \Sigma x_{3t} y_t - \Sigma x_{2t} x_{3t} \Sigma x_{1t} y_t}{\Sigma x_{2t}^2 \Sigma x_{3t}^2 - (\Sigma x_{2t} \Sigma x_{3t})^2}.$$

با استفاده از محاسبات انجام شده، خواهیم داشت

$$\hat{\beta}_2 = \frac{100(90) - 95(90)}{100(100) - (95)^2} = \frac{450}{975} = 0/461,$$

$$\hat{\beta}_3 = \frac{100(90) - 95(90)}{100(100) - (95)^2} = \frac{450}{975} = 0/461.$$

۲. برای آزمون t باید واریانس $\hat{\beta}_2$ و $\hat{\beta}_3$ را به دست آوریم. ابتدا باید تغییرات توضیح داده نشده را به دست آورد، که خود مستلزم محاسبهٔ تغییرات توضیح داده شده است. از معادلهٔ ۳۹ـ۵ داریم

$$ESS = \hat{\mathbf{y}}'\hat{\mathbf{y}} = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y},$$

$$= [\hat{\beta}_2 \quad \hat{\beta}_3] \begin{bmatrix} \Sigma x_{2t} y_t \\ \Sigma x_{3t} y_t \end{bmatrix} = \hat{\beta}_2 \Sigma x_{2t} y_t + \hat{\beta}_3 \Sigma x_{3t} y_t,$$

در نتیجه

$$ESS = 0/461(90) + 0/461(90) = 83$$

از صورت مسأله می‌دانیم که: $\Sigma y_t^2 = \mathbf{y}'\mathbf{y} = TSS = 100$.

بنابراین

$$RSS = TSS - ESS = 100 - 83 = 17.$$

از معادلهٔ ۶ـ۲۱ مقدار واریانس U_t را حساب می‌کنیم، خواهیم داشت

$$\hat{\sigma}_U^2 = \frac{\mathbf{e}'\mathbf{e}}{n-k} = \frac{17}{20-3} = 1.$$

از معادله‌های ۴ـ۲۴ و ۴ـ۲۵ داریم

$$\mathrm{Var}(\hat{\beta}_2)=\frac{\hat{\sigma}_U^2}{(1-r_{23}^2)\sum x_{2i}^2}, \qquad \mathrm{Var}(\hat{\beta}_3)=\frac{\hat{\sigma}_U^2}{(1-r_{23}^2)\sum x_{3i}^2}.$$

ابتدا باید r_{23}^2 را محاسبه کرد

$$r_{23}^2=\frac{(\sum x_{2i}x_{3i})^2}{\sum x_{2i}^2\sum x_{3i}^2}=\frac{(۹۵)^2}{۱۰۰(۱۰۰)}=۰/۹۰۲۵.$$

بنابراین

$$\mathrm{Var}(\hat{\beta}_2)=\frac{۱(۱۰۰)^2}{(۱۰۰^2-۹۵^2)۱۰۰}=\frac{۱۰۰}{۹۷۵}=۰/۱۰۲۵۶۴,$$

$$\mathrm{Var}(\hat{\beta}_3)=\frac{۱(۱۰۰)^2}{(۱۰۰^2-۹۵^2)۱۰۰}=\frac{۱۰۰}{۹۷۵}=۰/۱۰۲۵۶۴.$$

آمارهٔ آزمون t عبارت است از

$$t_2=\frac{\hat{\beta}_2}{\mathrm{SE}(\hat{\beta}_2)}=\frac{۰/۴۶۱}{\sqrt{۰/۱۰۲۵۶۴}}=۱/۴۴۴,$$

$$t_3=\frac{\hat{\beta}_3}{\mathrm{SE}(\hat{\beta}_3)}=\frac{۰/۴۶۱}{\sqrt{۰/۱۰۲۵۶۴}}=۱/۴۴۴.$$

مقدار به دست آمده از جدول t با ۱۷ درجه آزادی و در سطح معنی‌دار ۵ درصد برابر است با ۲/۱۱۰ ، در نتیجه آمارهٔ آزمون معنی‌دار نیست و فرضیه‌های $\beta_2=۰$ و $\beta_3=۰$ رد نمی‌شود.

۳. برای آزمون F از آمارهٔ ۶ـ۳۶ استفاده می‌کنیم،

$$F=\frac{\mathrm{ESS}/(k-۱)}{\mathrm{RSS}/(n-k)},$$

$$F=\frac{۸۳/(۳-۱)}{۱۷(۲۰-۳)}=۴۱/۵$$

مقدار به دست آمده از جدول F با درجات آزادی ۲ و ۱۷ و در سطح معنی‌دار ۵ درصد برابر است با

$$F_{\%5}(2, 17) = 3/59 ,$$

بنابراین آمارهٔ آزمون از مقدار F در جدول به مراتب بیشتر شده، در ناحیهٔ بحرانی قرار می‌گیرد، نتیجه می‌گیریم که فرضیهٔ $H_0: \beta_2 = \beta_3 = 0$ رد می‌شود.

بنابراین ملاحظه می‌شود که با آزمون t هر دو فرضیه قبول می‌شود. در حالی که با آزمون F هر دو فرضیه رد خواهد شد. علت این امر را می‌توان این گونه توضیح داد که تأثیر مستقل X_{2i} و X_{3i} روی Y_i بسیار ضعیف است، در حالی که تأثیر مشترک آنها بسیار قوی است. البته دلیل اشتراک قدرت توضیحی بسیار X_{2i} و X_{3i} را می‌توان در ضریب همبستگی بسیار قوی آنها جستجو کرد. می‌دانیم $r_{23}^2 = 0/9025$ در نتیجه $r_{23} = 0/95$. این نکته را در بحث از فرمول ۶.۴۶ بیان کردیم. برای تبیین بیشتر این نکته تأثیر مستقل X_{2i} و X_{3i} و تأثیر مشترک آنها را در تغییرات توضیح داده شدهٔ Y_i محاسبه می‌کنیم. معادلهٔ ۶.۴۲ جواب این سؤال است. این معادله را یک بار دیگر می‌نویسیم،

$$ESS = \hat{\beta}_2^2 \Sigma x_{2i}^2 + \hat{\beta}_3^2 \Sigma x_{3i}^2 + 2\hat{\beta}_2 \hat{\beta}_3 \Sigma x_{2i} x_{3i} .$$

دو جمله اول سمت راست، تأثیرات مستقیم و مستقل X_{2i} و X_{3i} روی Y_i را اندازه گیری می‌کند و جملهٔ سوم تأثیر مشترک را منعکس می‌سازد. با اطلاعاتی که داریم، هریک را محاسبه می‌کنیم،

$$\hat{\beta}_2^2 \Sigma x_{2i}^2 = (0/461)^2 (100) = 21/3 ,$$

$$\hat{\beta}_3^2 \Sigma x_{3i}^2 = (0/461)^2 (100) = 21/3 ,$$

$$\frac{2\hat{\beta}_2 \hat{\beta}_3 \Sigma x_{2i} x_{3i}}{\Sigma \hat{y}_i^2} = \frac{2 (0/461) (0/461) (95) = 40/4}{83} ,$$

ملاحظه می‌شود که تأثیر مشترک X_{2i} و X_{3i} روی Y_i تقریباً دو برابر تأثیر مستقل هریک

از آنهاست. همان گونه که قبلاً اشاره شد این امر می تواند به علت همبستگی بسیار شدید $X_{۲t}$ و $X_{۳t}$ در مدل رگرسیون مفروض باشد.

۶.۵ آزمون یک ترکیب خطی از پارامترها

آزمون محدودیتهای خطی بین پارامترها، کاربرد وسیعی در تحلیلهای اقتصادسنجی دارد. فرض کنید در مدل

$$Y_t = \beta_۱ + \beta_۲ X_{۲t} + \cdots + \beta_k X_{kt} + U_t ,$$

می‌خواهیم ترکیب خطی زیر را بین پارامترها آزمون کنیم،

$$H_۰ : p\beta_۲ + q\beta_۳ = d .$$

روش‌کار دقیقاً مانند آزمون $\beta_۱ = a$ است. همان گونه که $\hat{\beta}_۱$ را استاندارد کرده، آمارهٔ آزمون را محاسبه و یا مقدار t از جدول مقایسه می‌کردیم، در اینجا نیز باید $(p\hat{\beta}_۲ + q\hat{\beta}_۳)$ را استاندارد کرده، آمارهٔ آزمون را به دست آوریم. کافی است $(p\hat{\beta}_۲ + q\hat{\beta}_۳)$ را از میانگین خود کم کرده، بر انحراف معیارش تقسیم کنیم، خواهیم داشت

$$t = \frac{(p\hat{\beta}_۲ + q\hat{\beta}_۳) - E(p\hat{\beta}_۲ + q\hat{\beta}_۳)}{\sqrt{Var(p\hat{\beta}_۲ + q\hat{\beta}_۳)}} \sim t(n-k) ,$$

رابطهٔ فوق را می‌توان چنین نوشت،

$$t = \frac{(p\hat{\beta}_۲ + q\hat{\beta}_۳) - (p\beta_۲ + q\beta_۳)}{p^۲ Var\,\hat{\beta}_۲ + q^۲ Var\,\hat{\beta}_۳ + ۲ p q Cov(\hat{\beta}_۲, \hat{\beta}_۳)} \sim t(n-k) .$$

با داشتن تخمینهای $\hat{\beta}_۲$ و $\hat{\beta}_۳$ و با توجه به اینکه $(p\beta_۲ + q\beta_۳) = d$ و نیز محاسبهٔ واریانس $\hat{\beta}_۲$ و واریانس $\hat{\beta}_۳$ و کوواریانس $\hat{\beta}_۲$ و $\hat{\beta}_۳$، آمارهٔ آزمون به سهولت محاسبه می‌شود. اگر آمارهٔ آزمون معنی‌دار باشد، یعنی از مقدار جدول t بزرگتر شود و در ناحیهٔ بحرانی قرار بگیرد، فرضیهٔ $H_۰$ رد می‌شود، در غیر این صورت قبول خواهد شد.

مثال ۹ـ۶ مدل رگرسیون زیر را که در مثال ۲ـ۶ مطرح کردیم، یک بار دیگر ملاحظه کنید،

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t .$$

می‌دانیم

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} ۱ & -۱/۵ \\ -۱/۵ & ۲/۵ \end{bmatrix} ,$$

و نیز $\hat{\sigma}^2 = ۰/۷۵$، $\mathrm{Var}(\hat{\beta}) = ۰/۷۵$ و $\mathrm{Var}(\hat{\gamma}) = ۱/۸۷۵$، $n = ۵$. همچنین می‌دانیم

$$\hat{Y}_t = ۴ + ۲/۵ X_t - ۱/۵ Z_t .$$

می‌خواهیم فرضیهٔ $H_۰ : ۲\beta - ۳\gamma = ۱۰$ را در مقابل فرضیهٔ $H_۰ \neq ۰$ در سطح معنی‌دار ۵ درصد آزمون کنیم.

برای این منظور آمارهٔ آزمون را محاسبه می‌کنیم. می‌دانیم

$$\mathrm{Var}(۲\hat{\beta} - ۳\hat{\gamma}) = ۴\,\mathrm{Var}(\hat{\beta}) + ۹\,\mathrm{Var}(\hat{\gamma}) - ۱۲\,\mathrm{Cov}(\hat{\beta}, \hat{\gamma}) .$$

با توجه به $\mathrm{Cov}(\hat{\beta}, \hat{\gamma}) = ۰/۷۵(-۱/۵) = -۱/۱۲۵$، خواهیم داشت

$$\mathrm{Var}(۲\hat{\beta} - ۳\hat{\gamma}) = ۴(۰/۷۵) + ۹(-۱/۸۷۵) + ۱۲(۱/۱۲۵) = ۳۳/۳۷۵ ,$$

در نتیجه

$$t = \frac{(۲\hat{\beta} - ۳\hat{\gamma}) - (۲\beta - ۳\gamma)}{\mathrm{SE}(۲\hat{\beta} - ۳\hat{\gamma})} = \frac{۲(۲/۵) - ۳(-۱/۵) - ۱۰}{\sqrt{۳۳/۳۷۵}} = \frac{-۰/۵}{۵/۷۷۱} = -۰/۰۸۶ .$$

با توجه به مقدار جدول t در سطح معنی‌دار ۰/۰۲۵، یعنی $t = ۴/۳۰۳$، ملاحظه می‌شود که آمارهٔ آزمون معنی‌دار نیست و $H_۰$ رد نمی‌شود.

مثال ۱۰ـ۶ برای مثال ۹ـ۶ فرضیه زیر را آزمون کنید،

$$H_۰ : \beta + \gamma = ۱$$

یادآوری می‌کنیم که اگر مدل مفروض در مثال ۶ـ۹ بیان لگاریتمی از یک تابع تولید گاب ـ داگلاس به صورت $Q_t = AK^{\alpha} L^{\beta} \epsilon_t$ باشد، فرضیهٔ H_0 همان بازده ثابت نسبت به مقیاس است. آمارهٔ آزمون در این حالت برابر است با

$$t = \frac{(\hat{\beta} + \hat{\gamma}) - (\beta + \gamma)}{\sqrt{\text{Var}(\hat{\beta}) + \text{Var}(\hat{\gamma}) + 2\,\text{Cov}(\hat{\beta}, \hat{\gamma})}},$$

$$t = \frac{2/5 + (-1/5) - 1}{\sqrt{0/75 + 1/875 - 2(0/75)(-1/5)}} = 0,$$

که با توجه به مقدار جدول t، یعنی ۴/۳۰۳ معنی‌دار نیست و H_0 قبول می‌شود.

۱. آزمون $\mathbf{C}'\boldsymbol{\beta} = \mathbf{r}$

حالت عمومی آزمون یک محدودیت خطی در مدل رگرسیون $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ را می‌توان به صورت زیر نشان داد،

$$H_0: \mathbf{C}'\boldsymbol{\beta} = \mathbf{r} \qquad (۶ـ۴۷)$$

ابتدا بردارهای $\mathbf{C} \to (k \times 1)$ و $\mathbf{r} \to (1 \times 1)$ را تعریف می‌کنیم. برای این منظور، $\mathbf{C}'\boldsymbol{\beta} = \mathbf{r}$ را برای مثالهای ۶ـ۹ و ۶ـ۱۰ می‌نویسیم. فرضیهٔ $H_0: 2\beta - 3\gamma = 10$ در مثال ۶ـ۱۰ به زبان معادلهٔ ۶ـ۴۷ عبارت است از

$$\begin{bmatrix} 2 & -3 \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 10 \end{bmatrix}, \qquad (۶ـ۴۸)$$

بنابراین بردار $\mathbf{C}'$ برابر است با $[2 \quad -3]$ و $\mathbf{r}$ نیز یک اسکالر یک در یک بوده، برابر $[10]$ است. برای مثال ۶ـ۱۰، فرضیهٔ $H_0: \beta + \gamma = 1$ به زبان ماتریسی برابر است با

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix},$$

در نتیجه $\mathbf{C}' = [1 \quad 1]$ و $\mathbf{r} = [1]$.

در نوشتن بردار $\mathbf{C}'$ باید دقت شود که علاوه‌بر ضرایب پارامترهایی که در H_0 وجود دارد، باید برای سایر پارامترهای مدل نیز ضریب صفر را منظور کرد تا بدین ترتیب $\mathbf{C}'$ همواره یک بردار $(1 \times k)$ شود. برای مثال، اگر بخواهیم در مدل رگرسیون زیر

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + U_t$$

فرضیهٔ $H_0: 2\beta_2 - 3\beta_3 = 10$ را آزمون کنیم، معادلهٔ $\mathbf{C}'\beta = r$ عبارت است از

$$\begin{bmatrix} 2 & -3 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = [10] \ .$$

یادآوری می‌کنیم که پارامتر β_1 در $\mathbf{C}'$ وارد نشده؛ زیرا فرض بر این است که $(\mathbf{X}'\mathbf{X})$ براساس مقادیر انحراف از میانگین محاسبه شده است. اما اگر در محاسبهٔ $(\mathbf{X}'\mathbf{X})$ از مشاهدات اصلی استفاده کنیم، ضرورتاً پارامتر β_1 نیز در $\mathbf{C}'$ ظاهر خواهد شد. در حالت کلی، اگر ماتریس $(\mathbf{X}'\mathbf{X})$ براساس مقادیر اصلی محاسبه‌شده باشد، $\mathbf{C}'$ یک بردار $(1 \times k)$ است زیرا $(\mathbf{X}'\mathbf{X}) \to (n \times k)$؛ در غیر این صورت هرگاه $(\mathbf{X}'\mathbf{X}) \to (k-1) \times (k-1)$، آنگاه $\mathbf{C}' \to (1 \times k-1)$، یعنی $\mathbf{C}'$ فاقد جملهٔ ثابت مدل رگرسیون خواهد بود.

بعد از آشنایی با $\mathbf{C}'\beta = r$، به این نکته اشاره می‌کنیم که برای آزمون H_0 در معادلهٔ ۶ـ۴۷ باید $\mathbf{C}'\beta = r$ را استاندارد کرده، آمارهٔ آزمون به دست آمده را با مقدار جدول t مقایسه کنیم؛ بنابراین ابتدا میانگین و واریانس $\mathbf{C}'\hat{\beta}$ باید محاسبه شود،

$$E(\mathbf{C}'\hat{\beta}) = \mathbf{C}'E(\hat{\beta}) = \mathbf{C}'\beta \ ,$$

$$\text{Var}(\mathbf{C}'\hat{\beta}) = E[\mathbf{C}'(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\mathbf{C}] \ ,$$

$$= \mathbf{C}'E[(\hat{\beta}-\beta)(\hat{\beta}-\beta)']\mathbf{C} \ .$$

با توجه به معادلهٔ ۶ـ۷ می‌توان نوشت،

$$\mathrm{Var}(\mathbf{C}'\hat{\boldsymbol{\beta}}) = \sigma^2 \mathbf{C}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{C} \qquad (۶ـ۴۹)$$

بدین ترتیب آمارهٔ آزمون عبارت خواهد بود از

$$t = \frac{\mathbf{C}'\hat{\boldsymbol{\beta}} - \mathbf{C}'\boldsymbol{\beta}}{\sqrt{\sigma^2 \mathbf{C}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{C}}} \sim t(n-k),$$

$$t = \frac{\mathbf{C}'\hat{\boldsymbol{\beta}} - \mathbf{r}}{\sigma\sqrt{\mathbf{C}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{C}}} \sim t(n-k). \qquad (۶ـ۵۰)$$

اگر آمارهٔ آزمون فوق معنی‌دار بود؛ یعنی از مقدار موجود در جدول t بزرگتر شد، فرضیهٔ H_0 رد می‌شود.

مثال ۶ـ۱۱ فرضیهٔ $H_0: 2\beta - 3\beta = ۱۰$ را که موضوع آزمون مثال ۶ـ۹ بود به زبان ماتریسی و با استفاده از رابطهٔ ۶ـ۵۰ آزمون کنید.

فرضیهٔ H_0 را قبلاً در معادلهٔ ۶ـ۴۸ به زبان ماتریسی و به صورت $\mathbf{C}'\boldsymbol{\beta}$ نوشته‌ایم. $\mathbf{C}'\hat{\boldsymbol{\beta}}$ نیز به همین ترتیب به دست می‌آید،

$$\mathbf{C}'\hat{\boldsymbol{\beta}} = [۲ \quad -۳] \begin{bmatrix} ۲/۵ \\ -۱/۵ \end{bmatrix} = ۹/۵ .$$

برای محاسبهٔ واریانس $\mathbf{C}'\hat{\boldsymbol{\beta}}$ از معادلهٔ ۶ـ۴۹ استفاده می‌کنیم،

$$\mathrm{Var}(\mathbf{C}'\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2 \mathbf{C}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{C},$$

$$= ۰/۷۵ [۲ \quad -۳] \begin{bmatrix} ۱ & -۱/۵ \\ -۱/۵ & ۲/۵ \end{bmatrix} \begin{bmatrix} ۲ \\ -۳ \end{bmatrix} = ۳۳/۳۷۵ .$$

آمارهٔ آزمون از معادلهٔ ۶ـ۵۰ به دست می‌آید،

$$t = \frac{۹/۵ - ۱۰}{\sqrt{۳۳/۳۷۵}} = \frac{-۰/۵}{۵/۷۷۱} = -۰/۰۸۶ .$$

با توجه به مقدار جدول t در سطح معنی‌دار ۰/۰۲۵، یعنی ۴/۳۰۳ =t، ملاحظه می‌شود که آمارهٔ آزمون معنی‌دار نیست و فرضیهٔ H_0 رد نمی‌شود.

سؤال این است که اگر در یک مدل رگرسیون چندمتغیره، دو یا چند پارامتر را در نظر بگیریم و نسبت به هر پارامتر فرضیه‌ای را مطرح نماییم و سپس بخواهیم این فرضیه‌ها را همزمان آزمون کنیم از چه روشی می‌توان استفاده کرد. با توجه به مطالبی که مطرح شده است به نظر می‌رسد که نمی‌توان از آزمون t استفاده کرد. در ادامهٔ بحث خواهیم دید که آزمون F جواب این مسأله است.

۶.۶ آزمون همزمان پارامترها و تعیین نواحی اطمینان *

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t .$$

فرض کنید می‌خواهیم فرضیه‌های زیر را به طور همزمان آزمون کنیم،

$$H_0: \beta_2 = a_1 ,$$

$$H_0: \beta_3 = a_2 ,$$

$$\vdots$$

$$H_0: \beta_i = a_i .$$

برای این منظور ابتدا فرض می‌کنیم که فرضیه‌های H_0 همه صحیح است. مقادیر β در فرضیه‌های H_0 را در مدل رگرسیون مفروض جایگزین می‌کنیم تا مدل رگرسیون مقید به‌دست آید. بدیهی است تعداد پارامترهایی که باید در مدل رگرسیون مقید تخمین زده شود، از تعداد پارامترهای مدل اصلی کمتر است. تفاوت این دو دقیقاً به اندازهٔ تعداد فرضیه‌هایی است که می‌خواهیم آزمون کنیم. اگر مجموع مربعات پسماند را در مدل اصلی با RSS و در مدل مقید با RSS_r مشخص کنیم، مسألهٔ مفروض دقیقاً «آزمون معنی‌دار بودن زیرمجموعه‌ای از پارامترها» می‌شود که در قسمت ۶.۳ مطرح شد. اگر تعداد فرضیه‌ها را r بنامیم، آمارهٔ آزمون عبارت خواهد بود از

$$\frac{(RSS_r - RSS)/r}{RSS/(n-k)} \sim F(r, n-k) \qquad (۶.۵۱)$$

که همان فرمول ۶.۳۳ است.

مثال ۶.۱۲ تابع تولید زیر را در نظر می‌گیریم،

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i ,$$

که در آن Y_i، X_{1i} و X_{2i} به ترتیب لگاریتم تولید، لگاریتم کار و لگاریتم سرمایه است. براساس یک نمونه شامل ۲۳ شرکت مختلف تولیدی و محاسبات انجام شده در مثال ۴.۲ تخمین زیر را به دست آورده‌ایم،

$$\hat{Y}_i = ۴ + ۰/۷ X_{1i} + ۰/۲ X_{2i} .$$

همچنین ازمثال ۴.۲ می‌دانیم‌که $\Sigma y_i^2 = ۱۰$، $\Sigma x_{1i}^2 = ۱۲$ و $\Sigma x_{1i} y_i = ۱۰$، $RSS = ۱/۴$.

می‌خواهیم دو فرضیه زیر را به طور همزمان در سطح معنی‌دار ۵ درصد آزمون کنیم،

$$H_0: \beta_1 = ۱ , \qquad H_0: \beta_2 = ۰ ,$$
$$H_1: \beta_1 \neq ۱ , \qquad H_1: \beta_2 \neq ۰ ,$$

برای اینکه بتوان آمارهٔ آزمون را از معادلهٔ ۶.۳۳ به دست آورد کافی است مجموع مربعات پسماند را برای مدل مقید، یعنی RSS_r محاسبه کنیم. برای این منظور می‌گوییم که اگر فرضیه‌های فوق صحیح باشد، مدل رگرسیون مقید به صورت زیر خواهد بود،

$$Y_i = \alpha + (۱) X_{1i} + ۰ (X_{2i}) + U_i ,$$
$$Y_i = \alpha + X_{1i} + U_i .$$

برای به دست آوردن $\hat{\alpha}$ باید Σe_i^2 را حداقل کنیم،

$$\frac{d \Sigma e_i^2}{d \hat{\alpha}} = \frac{d \Sigma (Y_i - \hat{\alpha} - X_{1i})^2}{d \hat{\alpha}} ,$$
$$= ۲ (-۱) \Sigma (Y_i - \hat{\alpha} - X_{1i}) = ۰ ,$$

در نتیجه خواهیم داشت

$$\hat{\alpha} = \bar{Y} - \bar{X}_1 .$$

با استفاده از تعریف RSS_r داریم

$$RSS_r = \Sigma (Y_i - \hat{Y}_i)^2 ,$$

$$= \Sigma [(Y_i - (\hat{\alpha} + X_{1i})]^2 ,$$

$$= \Sigma [(Y_i - (\bar{Y} - \bar{X}_1 + X_{1i})] = \Sigma [(Y_i - \bar{Y}) - (X_{1i} - \bar{X}_1)]^2 .$$

بدین ترتیب نتیجه می‌گیریم که

$$RSS_r = \Sigma y_i^2 + \Sigma x_{1i}^2 - 2 \Sigma x_{1i} y_i ,$$

$$= ۱۰ + ۱۲ - ۲۰ = ۲ .$$

معادلهٔ ۶۵۱ را می‌نویسیم،

$$F = \frac{(RSS_r - RSS) / r}{RSS / (n - k)} .$$

می‌دانیم تعداد محدودیتها که درواقع همان تعداد فرضیه‌هاست برابر است با $r = ۲$. در نتیجه:

$$F = \frac{(۲ - ۱/۴)/۲}{۱/۴/(۲۳ - ۳)} = \frac{۰/۳}{۰/۰۷} = ۴/۳ .$$

آمارهٔ آزمون فوق را باید با مقدار جدول F با درجات آزادی ۲۰ و ۲ و سطح معنی‌دار ۵ درصد آزمون کرد. می‌دانیم

$$F_{٪۵} (۲ , ۲۰) = ۳/۴۹ ,$$

در نتیجه آمارهٔ آزمون در ناحیهٔ بحرانی قرار گرفته و معنی‌دار است؛ بنابراین، فرضیه‌های H_0 به طور همزمان رد می‌شوند.

ناحیهٔ اطمینان برای چند پارامتر

همان گونه که در قسمت ۶ـ۴ و در توضیحات نمودار ۶ـ۱ اشاره کردیم، در رگرسیون چندمتغیره می توان از نواحی اطمینان نیز صحبت کرد. حالت عمومی در به دست آوردن این نواحی اطمینان در قسمت ۶ـ۸ بررسی می‌شود. در اینجا برای تبیین این مسأله، قلمرو بحث را به یک مدل رگرسیون با دو متغیر توضیحی محدود می‌کنیم. مدل مثال ۶ـ۱۲ را یک بار دیگر می‌نویسیم،

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i .$$

در مثال ۴ـ۲ برای β_1 و β_2 فواصل اطمینان را به طور جداگانه محاسبه کردیم. در این قسمت می‌خواهیم ناحیهٔ اطمینان برای هر دو پارامتر را در سطح معنی‌دار ۵ درصد به دست آوریم.

بدون اثبات، می‌گوییم که در رگرسیونهایی که دو متغیر توضیحی به شرح فوق دارند، آماره زیر دارای توزیع F با درجات آزادی $(۳-n, ۲)$ است.

$$F = \frac{۱}{۲\hat{\sigma}_U^۲}\left[\Sigma x_{1i}^۲ (\hat{\beta}_1-\beta_1)^۲ + ۲\Sigma x_{1i}x_{2i}(\hat{\beta}_1-\beta_1)(\hat{\beta}_2-\beta_2) + \Sigma x_{2i}^۲ (\hat{\beta}_2-\beta_2)^۲\right] . \quad (۶ـ۵۲)$$

به کمک معادلهٔ ۶ـ۵۲ می توان اولاً، نواحی اطمینان برای پارامترهای β_1 و β_2 را به دست آورد؛ ثانیاً، فرضیه‌های مختلف برای β_1 و β_2 را به طور همزمان آزمون کرد. البته در همین قسمت دیدیم که به کمک معادلهٔ ۶ـ۵۱ می توان آزمونهای همزمان پارامترها را نیز انجام داد. در این قسمت نشان می‌دهیم که استفاده از معادله‌های ۶ـ۵۱ و ۶ـ۵۲ به یک جواب منتهی می‌شوند. بنابراین می توان گفت که اهمیت معادلهٔ ۶ـ۵۲ بیشتر در به دست آوردن نواحی اطمینان است.

مثال ۶ـ۱۳ برای به دست آوردن ناحیهٔ اطمینان ۹۵ درصدی برای β_1 و β_2 در مثال ۶ـ۱۲، معادلهٔ ۶ـ۵۲ را به صورت زیر می‌نویسیم. توجه داریم که $F_{۰.۰۵}(۲,۲۰) = ۳/۴۹$. همچنین می‌دانیم در ناحیهٔ اطمینان، آمارهٔ آزمون باید کمتر یا مساوی مقدار F از جدول باشد.

$$\Sigma x_{1i}^{2}(\hat{\beta}_1-\beta_1)^2+2\Sigma x_{1i}x_{2i}(\hat{\beta}_1-\beta_1)(\hat{\beta}_2-\beta_2)+\Sigma x_{2i}^{2}(\hat{\beta}_2-\beta_2)^2\le 3/49\,(2\hat{\sigma}^2)\,.$$

(۶ـ۵۳)

با استفاده از محاسبات مثال ۴ـ۲ می‌دانیم

$$\Sigma y_i^2=10\,,\quad \Sigma x_{1i}^2=12\,,\quad \Sigma x_{2i}^2=12\,,\quad \Sigma x_{1i}x_{2i}=8\,,\quad \hat{\sigma}^2=0/07$$

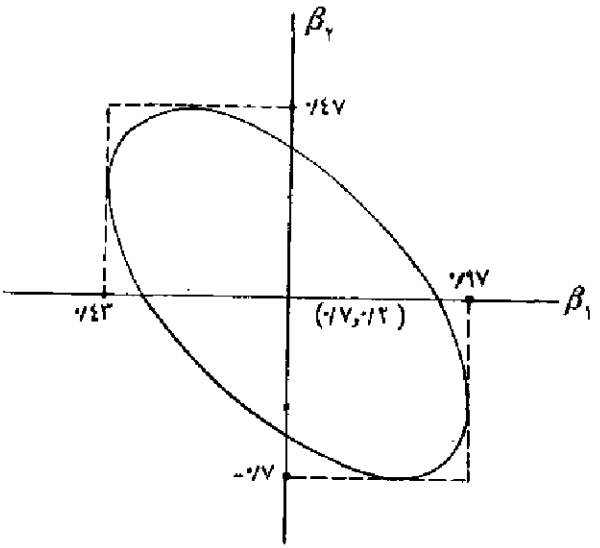
بنابراین معادلهٔ ۶ـ۵۳ به صورت زیر نوشته می‌شود،

$$12\,(0/7-\beta_1)^2+2\,(8)\,(0/7-\beta_1)(2-\beta_2)+12\,(0/2-\beta_2)^2\le 3/49\,(2)\,(0/07)\,.$$

دو طرف نامساوی فوق را بر ۱۲ تقسیم کرده، با تبدیل $(\hat{\beta}_1-\beta_1)$ به $(\beta_1-\hat{\beta}_1)$ و نیز $(\hat{\beta}_2-\beta_2)$ به $(\beta_2-\hat{\beta}_2)$ خواهیم داشت

$$(\beta_1-0/7)^2+\tfrac{4}{3}(\beta_1-0/7)\,(\beta_2-0/2)+(\beta_2-0/2)^2\le 0/41\,. \quad (۶ـ۵۴)$$

ملاحظه می‌شود که منحنی نمودار تغییرات رابطهٔ ۶ـ۵۴ یک بیضی است که مرکز آن نقطهٔ (۰/۷, ۰/۲) است. نمایش هندسی این بیضی در نمودار ۶ـ۲ منعکس است. ذکر دو نکته



نمودار ۶ـ۲ ناحیهٔ اطمینان برای آزمونهای همزمان

در مورد این نمودار ضروری است. نکتهٔ اول این است که اگر $\text{Cov}(\hat{\beta}_1, \hat{\beta}_2) < 0$، آنگاه این بیضی به سمت چپ متمایل است. در مثال ۲ـ۴ نیز دیدیم که $\text{Cov}(\hat{\beta}_1, \hat{\beta}_2) = -0/007$؛ بنابراین در نمودار ۲ـ۶ جهت بیضی به سمت چپ است. اگر $\text{Cov}(\hat{\beta}_1, \hat{\beta}_2) > 0$، بیضی به سمت راست متمایل خواهد بود. با توجه به معادلهٔ ۲۶ـ۴ می‌توان گفت که علامت کوواریانس $\hat{\beta}_1, \hat{\beta}_2$ تابع معکوسی از علامت $\Sigma x_{1i} x_{2i}$ است. نتیجه می‌گیریم که در مدلهای رگرسیون با دو متغیر توضیحی، هرگاه $\Sigma x_{1i} x_{2i}$ مثبت باشد، ناحیهٔ اطمینان برای آزمون همزمان دو پارامتر به سمت چپ متمایل است و برعکس. دومین نکته در مورد این نمودار این است که ناحیهٔ اطمینان ۹۵ درصد برای β_1, β_2 درواقع سطح محصور در بیضی است. این سطح با فواصل اطمینان ۹۵ درصد که با آزمون t به طور جداگانه برای β_1, β_2 به دست می‌آید متفاوت است.

برای تبیین این نکته مفید است که فاصلهٔ اطمینان ۹۵ درصد برای β_1, β_2 را که در مثال ۲ـ۴ به دست آورده‌ایم یک بار دیگر می‌نویسیم،

$$\text{فاصله اطمینان برای } \beta_1 : (0/49, 0/91) ,$$

$$\text{فاصله اطمینان برای } \beta_2 : (-0/01, 0/41) ,$$

این فاصله‌های اطمینان را می‌توان با قلمرو تغییرات β_1 و β_2 که در آزمون همزمان تعیین شده است و به راحتی از طریق مماسهای بر بیضی به دست می‌آید، مقایسه نمود،

$$\text{فاصله اطمینان برای } \beta_1 : (0/43, 0/97) ,$$

$$\text{فاصله اطمینان برای } \beta_2 : (-0/07, 0/47) ,$$

استفادهٔ دیگری که می‌توان از معادلهٔ ۵۲ـ۶ کرد، آزمون همزمان فرضیه‌های مختلف در مورد پارامترهای β_1 و β_2 است. بدیهی است با داشتن نواحی اطمینان، به راحتی میتوان فرضیه‌های مختلف را به طور همزمان آزمون کرد. اگر فرضیه‌های مورد نظر در داخل ناحیهٔ اطمینان قرار بگیرد، قبول و در غیر این صورت رد می‌شود. اما معمولاً تعیین نواحی اطمینان تا حدی مشکل است و به همین دلیل در این موارد

مستقیماً و به کمک معادلهٔ ۶ـ۵۲ از آزمون F استفاده می‌شود. می‌دانیم آمارهٔ آزمون در رابطهٔ ۶ـ۵۲ دارای توزیع F با درجات آزادی (۳ - n , ۲) است. مقدار عددی آمارهٔ آزمون را حساب می‌کنیم. مقادیر $H_0: \beta_1 = 1$ و $H_0: \beta_2 = 0$ را در ۶ـ۵۲ قرار داده خواهیم داشت

$$F = \frac{1}{2(0/07)}[12(0/7-1)^2 + 2(8)(0/7-1)(0/2-0) + 12(0/2-0)^2] = 4/3.$$

مقدار موجود در جدول F با درجات آزادی ۲ , ۲۰ و در سطح معنی‌دار ۵ درصد برابر است با

$$F_{5\%}(2, 20) = 3/49.$$

با توجه به اینکه آمارهٔ آزمون در ناحیهٔ بحرانی قرار می‌گیرد؛ معنی‌دار است و فرضیه‌های H_0 هر دو همزمان رد می‌شوند. یادآوری می‌شود که مقدار آمارهٔ آزمون به دست آمده از ۶ـ۵۲ دقیقاً برابر مقداری است که از معادلهٔ ۶ـ۵۱ به دست آوردیم. با توجه به سهولت معادلهٔ ۶ـ۵۱ ملاحظه می‌شود که آزمون همزمان پارامترها بهتر است به کمک معادلهٔ ۶ـ۵۱ صورت پذیرد. تعمیم نتایج به حالت کلی نیز با استفاده از این معادله بهتر انجام می‌شود. این نکته را در قسمت ۶ـ۸ نشان خواهیم داد.

۶ـ۷ آزمون تساوی چند پارامتر با یکدیگر

آزمون دیگری که می‌توان در مدلهای رگرسیون چند متغیره مطرح کرد، آزمون تساوی دو یا بیشتر از دو پارامتر با یکدیگر است. برای مثال، مدل مصرف زیر را در نظر می‌گیریم،

$$C_t = \beta_1 + \beta_2 W_t + \beta_3 NW_t + \beta_4 C_{t-1} + U_t,$$

که در آن C_t مصرف، W_t درآمدهای حاصل از دستمزد و NW_t درآمدهای حاصل از منابع غیر دستمزد است. می‌توان این فرضیه را مطرح کرد که آیا میل نهایی به مصرف در افرادی که حقوق یا دستمزد می‌گیرند با میل نهایی به مصرف در صاحبان مشاغل آزاد که درآمدشان از طریق حقوق و دستمزد نیست، یکسان است. در

این صورت باید فرضیه زیر را آزمون کرد،

$$H_0: \beta_2 = \beta_4 ,$$

$$H_1: \beta_2 \neq \beta_4 .$$

در حالت کلی می‌توان یک مدل رگرسیون با k پارامتر را در نظر گرفت و فرضیه زیر را مطرح کرد،

$$H_0: \beta_i = \beta_j ,$$

$$H_1: \beta_i \neq \beta_j .$$

دو راه حل می‌توان در نظر گرفت. در راه حل اول کافی است که توزیع تفاوت $\hat{\beta}_i$ و $\hat{\beta}_j$ را ملاحظه کرده، فرضیهٔ $H_0: (\hat{\beta}_i - \hat{\beta}_j) = 0$ را در مقابل فرضیهٔ مخالف صفر آزمون کرد. بنابراین، باید ابتدا $\hat{\beta}_i - \hat{\beta}_j$ را استاندارد کرد تا آمارهٔ آزمون به دست آید. می‌دانیم

$$E(\hat{\beta}_i - \hat{\beta}_j) = \hat{\beta}_i - \hat{\beta}_j ,$$

$$Var(\hat{\beta}_i - \hat{\beta}_j) = Var(\hat{\beta}_i) + Var(\hat{\beta}_j) - 2\,Cov(\hat{\beta}_i, \hat{\beta}_j) .$$

با تخمین واریانس جملهٔ اختلال از معادلهٔ ۶ـ۲۱ و تخمین واریانسهای $\hat{\beta}_i$ و $\hat{\beta}_j$ از معادلهٔ ۶ـ۲۴ و نیز کوواریانس $\hat{\beta}_i$، $\hat{\beta}_j$ از معادلهٔ ۶ـ۲۵، مقدار تخمین $Var(\hat{\beta}_i - \hat{\beta}_j)$ به دست می‌آید. بدین ترتیب آمارهٔ آزمون از رابطهٔ زیر نتیجه خواهد شد،

$$t = \frac{(\hat{\beta}_i - \hat{\beta}_j) - E(\hat{\beta}_i - \hat{\beta}_j)}{\sqrt{Var(\hat{\beta}_i - \hat{\beta}_j)}} ,$$

$$= \frac{\hat{\beta}_i - \hat{\beta}_j - (\beta_i - \beta_j)}{SE(\hat{\beta}_i - \hat{\beta}_j)} .$$

اگر فرضیهٔ H_0 صحیح باشد، یعنی اگر $(\beta_i - \beta_j) = 0$، آنگاه آمارهٔ آزمون زیر دارای

توزیع t با (n - k) درجهٔ آزادی خواهد داشت،

$$t = \frac{\hat{\beta}_i - \hat{\beta}_j}{SE(\hat{\beta}_i - \hat{\beta}_j)} \sim t(n-k) . \qquad (۶.۵۵)$$

مقدار جدول t را با (n - k) درجه آزادی و در سطح معنی‌دار $\alpha/۲$ درصد به دست می‌آوریم. اگر آمارهٔ آزمون (۶.۵۵) از این مقدار بیشتر شد، در آن صورت معنی‌دار بوده، فرضیهٔ H_0 رد می‌شود، در غیر این صورت رد نخواهد شد.

مثال ۱۴.۶ مدل رگرسیون زیر، موضوع مثال ۲.۶ را دوباره می‌نویسیم،

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t .$$

با استفاده از محاسبات انجام شده در مثال ۱.۶ فرضیه زیر را در سطح معنی‌دار ۵ درصد آزمون کنید،

$$H_0: \beta = \gamma ,$$

$$H_1: \beta \neq \gamma .$$

می‌دانیم $\hat{\beta} = ۲/۵$ ، $\hat{\gamma} = -۱/۵$ ، $Var(\hat{\beta}) = ۰/۷۵$ ، $Var(\hat{\gamma}) = ۱/۸۷۵$ و $Cov(\hat{\beta},\hat{\gamma}) = ۰/۷۵(-۱/۵) = -۱/۱۲۵$. با این اطلاعات، آمارهٔ آزمون قابل محاسبه است.

$$(\hat{\beta} - \hat{\gamma}) = ۲/۵ - (-۱/۵) = ۴ ,$$

$$Var(\hat{\beta} - \hat{\gamma}) = Var(\hat{\beta}) + Var(\hat{\gamma}) - ۲\,Cov(\hat{\beta},\hat{\gamma}) ,$$

$$= ۰/۷۵ + ۱/۸۷۵ - ۲(-۱/۱۲۵) = ۴/۸۷۵ .$$

از معادلهٔ ۶.۵۵ داریم

$$t = \frac{\hat{\beta} - \hat{\gamma}}{\sqrt{Var(\hat{\beta} - \hat{\gamma})}} \sim t(n-۳) ,$$

$$t = \frac{۴}{\sqrt{۴/۸۷۵}} = \frac{۴}{۲/۲۰۸} = ۱/۸۱۱ .$$

با توجه به اینکه آزمون دو طرفه است، مقدار جدول t با ۲ =(n - ۳) درجهٔ آزادی و در سطح معنی‌دار ۲۵ درصد برابر است با ۴/۳۰۳؛ بنابراین، آمارهٔ آزمون در ناحیهٔ بحرانی قرار نگرفته، معنی‌دار نیست. بدین ترتیب فرضیهٔ H_0 رد نمی‌شود.

آزمون تساوی بیش از دو پارامتر با یکدیگر

مدل رگرسیون چندمتغیره زیر مفروض است،

$$Y_t = \beta_1 + \beta_{2t} + \cdots + \beta_k X_{kt} + U_t .$$

می‌خواهیم فرضیهٔ زیر را آزمون کنیم،

$$H_0: \beta_2 = \beta_3 = \beta_4 .$$

فرضیهٔ H_1 این است که حداقل یکی از تساویهای فوق برقرار نباشد. همان‌گونه که می‌دانیم در روال عمومی آزمون F، فرض می‌شود، H_0 صحیح است؛ سپس مدل رگرسیون را مشروط به درستی H_0 می‌نویسیم و آن را مدل رگرسیون مقید می‌نامیم. بنابراین

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \cdots + \beta_k X_{kt} + U_t .$$

از β_2 فاکتور می‌گیریم،

$$Y_t = \beta_1 + \beta_2 (X_{2t} + X_{3t} + X_{4t}) + \cdots + \beta_k X_{kt} + U_t ,$$

یا

$$Y_t = \beta_1 + \beta_2 X^*_{2t} + \beta_5 X_{5t} + \cdots + \beta_k X_{kt} + U_t ,$$

که در آن

$$X^*_{2t} = (X_{2t} + X_{3t} + X_{4t}) .$$

مجموع مربعات پسماند در مدل آخر را RSS_r می‌نامیم و با توجه به معادله‌های ۶.۳۳ یا ۶.۵۱ خواهیم داشت

$$F = \frac{(RSS_r - RSS)/r}{RSS/(n-k)} \sim F(r, n-k) ,$$

که در آن، RSS مجموع مربعات پسماند در مدل اولیه و r تعداد محدودیتهاست. باید توجه کنیم که در این مثال، r =۲ است، نه ۳؛ زیرا اگر دو محدودیت برقرار باشد، مثلاً $\beta_2=\beta_3$ و $\beta_2=\beta_4$، محدودیت سوم، یعنی $\beta_3=\beta_4$ نیز برقرار خواهد بود. به عبارت دیگر فقط می‌توان دو محدودیت مستقل در H_0 تشخیص داد. بنابراین، r تعداد محدودیتهای مشاهده شده نبوده، بلکه تعداد محدودیتهای مستقل است. اگر آمارهٔ آزمون فوق معنی‌دار باشد، فرضیهٔ H_0 رد می‌شود، در غیر این صورت قبول خواهد شد.

مثال ۱۵ـ۶ در مثال ۱۴ـ۶ فرضیهٔ $H_0: \beta=\gamma$ را در مقابل فرضیهٔ $H_1: \beta\neq\gamma$ در سطح معنی‌دار ۵ درصد آزمون کنید.

گفتیم که فرضیه‌های تساوی چند پارامتر را باید با آمارهٔ F آزمون کرد؛ زیرا فرضیه‌های تساوی دو پارامتر را همواره می‌توان با آمارهٔ t آزمون کرد؛ با وجود این، با توجه به اینکه محاسبات مثال ۱۴ـ۶ کامل است، بنابراین، خیلی ساده‌تر خواهد بود اگر آزمون F را در مورد همین مثال بررسی کنیم. روش کار مهم است وگرنه تساوی دو یا بیشتر از دو پارامتر، از نظر روش آزمون تفاوتی ندارد. مدل رگرسیون را یک بار دیگر می‌نویسیم،

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t ,$$

در مثال ۲ـ۶ دیدیم که مجموع مربعات پسماند برای این مدل برابر است با RSS= ۱/۵. حال باید مدل مقید را بسازیم. با جایگزینی فرضیهٔ H_0 در این مدل خواهیم داشت

$$Y_t = \alpha + \beta X_t + \beta Z_t + U_t$$

$$= \alpha + \beta (X_t + Z_t) + U_t ,$$

یا با تعریف $X_t^* = (X_t + Z_t)$، داریم

$$Y_t = \alpha + \beta X_t^* + U_t .$$

باید این مدل مقید را تخمین زده، مجموع مربعات پسماند را به دست آوریم. با استفاده از

مثال ۲ـ۵ و جدول ۳ـ۵ محاسبات زیر را انجام می‌دهیم.

جدول ۶ـ۱

$X_t^* = X_t + Z_t$	x_t^*	$x_t^* y_t$	x_t^{*2}	y_t^2
۸	۰	۰	۰	۱
۵	-۳	۹	۹	۹
۱۱	۳	۱۲	۹	۱۶
۶	-۲	۲	۴	۱
۱۰	۲	۲	۴	۱
Σ ۴۰	۰	۲۵	۲۶	۲۸

با مراجعه به معادلهٔ ۴۷ـ۲ می‌دانیم

$$RSS_r = \Sigma e_{t(r)}^2 = \Sigma y_t^2 - \frac{(\Sigma x_t^* y_t)^2}{\Sigma x_t^{*2}},$$

$$= ۲۸ - \frac{(۲۵)^2}{۲۶} = \frac{۱۰۳}{۲۶} = ۳/۹۶ .$$

معادلهٔ ۵۱ـ۶ را می‌نویسیم،

$$F = \frac{(RSS_r - RSS)/r}{RSS/(n-k)} \sim F(r, n-k),$$

$$= \frac{(۳/۹۶ - ۱/۰۵)/۱}{(۱/۰۵)(۵-۳)} \sim F(۱, ۲) .$$

بنابراین مقدار آمارهٔ آزمون برابر است با $F = ۳/۲۸$. مقدار جدول F با درجات آزادی (۱ , ۲) و در سطح معنی‌دار ۵ درصد برابر است با ۱۸/۵۱؛ بنابراین، آمارهٔ آزمون معنی‌دار نبوده و H_0 رد نمی‌شود. همان گونه که انتظار می‌رفت دو آزمون F و t در این مورد خاص، که تساوی دو پارامتر را می‌خواهیم آزمون کنیم، دقیقاً به یک جواب می‌رسند؛ زیرا بنابر معادلهٔ ۵۱ـ۲ می‌دانیم $F = t^2$ و با توجه به $F = ۳/۲۸$ در مثال ۱۴ـ۶ و $t = ۱/۸۱۱$

در مثال ۱۳ـ٦ نتیجه می‌گیریم

$$\sqrt{۳/۲۸} = ۱/۸۱۱$$

٦ـ۸ آزمون همزمان چند ترکیب خطی از پارامترها: آماره عمومی آزمون و کاربرد آن*

عمومی‌ترین حالت در آزمون فرضیه این است که مجموعه‌ای از ترکیبهای خطی بین پارامترها را آزمون کنیم. ابتدا این آزمون را به زبان ماتریسی ارائه داده، سپس نشان می‌دهیم که چگونه بسیاری از آزمونهایی که تا به حال مطرح کرده‌ایم، می‌تواند حالت خاصی از این آزمون باشد.

مدل رگرسیون $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ مفروض است که در آن $\mathbf{y} \to n \times ۱$، $\mathbf{X} \to n \times k$ و $\boldsymbol{\beta} \to k \times ۱$ و $\mathbf{u} \to n \times ۱$. می‌خواهیم مجموعهٔ r فرضیه‌زیر را به‌طور همزمان آزمون کنیم،

$$H_۰: \mathbf{R}\boldsymbol{\beta} = \mathbf{r}, \qquad (٦ـ۵٦)$$

که در آن $\mathbf{R} \to r \times k$ و $r \leq k$ و $\mathbf{r} \to r \times ۱$. عناصر ماتریس $\mathbf{R}$ و بردار $\mathbf{r}$ مقادیر معلومی است. تعداد ردیفهای ماتریس $\mathbf{R}$، یعنی r درواقع تعداد محدودیتهای خطی است که می‌خواهیم آزمون کنیم؛ بنابراین r فرضیه برای آزمون داریم و بدیهی است که باید r از k کوچکتر یا مساوی آن باشد. توجه به این نکته نیز اهمیت دارد که ماتریس $\mathbf{R}$ باید دارای «رتبهٔ کامل سطری»[1] باشد؛ یعنی سطرهای آن از یکدیگر استقلال کامل داشته باشد که درواقع به معنی عدم همبستگی خطی بین فرضیه‌هاست. معادلهٔ ٦ـ۵٦ دارای اهمیت خاصی است، زیرا می‌توان فرضیه‌های مختلفی را در قالب آن بیان نمود. برای مثال، فرض کنید می‌خواهیم دو فرضیه زیر را به طور همزمان آزمون کنیم،

---

1. Full Row Rank

* برای توضیح بیشتر به پیوست «۵ ـ الف» مراجعه شود.

$$\begin{cases} H_0: 2\beta_2 - 3\beta_3 = 10, \\ H_1: 2\beta_1 - 3\beta_3 \neq 10, \end{cases} \quad \text{و} \quad \begin{cases} H_0: 2\beta_4 - 4\beta_5 = 2, \\ H_1: 2\beta_4 - 4\beta_5 \neq 2. \end{cases}$$

معادلهٔ ۶ـ۵۶ به صورت زیر خواهد بود،

$$\underbrace{\begin{bmatrix} 0 & 2 & -3 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & -4 & \dots & 0 \end{bmatrix}}_{\mathbf{R}} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \vdots \\ \beta_k \end{bmatrix}}_{\beta} = \underbrace{\begin{bmatrix} 10 \\ 2 \end{bmatrix}}_{\mathbf{r}}.$$

ملاحظه می‌شود که $\mathbf{R} \to 2 \times k$ و $\boldsymbol{\beta} \to k \times 1$ و $\mathbf{r} \to 2 \times 1$؛ بنابراین، فرض بر این است که مدل برحسب مشاهدات اصلی محاسبه شده است.

برای آزمون فرضیهٔ $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$، ابتدا بردار $\hat{\boldsymbol{\beta}}$ را به جای $\boldsymbol{\beta}$ قرار می‌دهیم. هرقدر اختلاف $\mathbf{R}\hat{\boldsymbol{\beta}}$ با $\mathbf{r}$ بیشتر باشد، احتمال اینکه فرضیهٔ H_0 رد شود، بیشتر خواهد بود. بنابراین، مسأله به بررسی توزیع $\mathbf{R}\hat{\boldsymbol{\beta}}$ و یافتن روشی که بتوان آمارهٔ آزمون را ساخت، تبدیل می‌شود. اگر آمارهٔ آزمون معنی‌دار باشد، فرضیهٔ H_0 رد می‌شود و برعکس. ابتدا میانگین $\mathbf{R}\hat{\boldsymbol{\beta}}$ را به دست می‌آوریم:

$$E(\mathbf{R}\hat{\boldsymbol{\beta}}) = \mathbf{R}E(\hat{\boldsymbol{\beta}}) = \mathbf{R}\boldsymbol{\beta}. \qquad (۶\text{ـ}۵۷)$$

واریانس $\mathbf{R}\hat{\boldsymbol{\beta}}$ برابر است با

$$\mathrm{Var}(\mathbf{R}\hat{\boldsymbol{\beta}}) = E[\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\mathbf{R}'],$$

$$= \mathbf{R}E[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \mathbf{r})']\mathbf{R}'.$$

با مراجعه به معادلهٔ ٦ـ٧ خواهیم داشت

$$\mathrm{Var}(\mathbf{R}\hat{\boldsymbol{\beta}}) = \sigma^2 \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' . \qquad (٦\text{ـ}٥٨)$$

با توجه به معادلهٔ ٦ـ١٢ و با توجه به اینکه $\hat{\boldsymbol{\beta}}$ دارای توزیع نرمال است، با استفاده از معادله‌های ٦ـ٥٧ و ٦ـ٥٨ داریم

$$\mathbf{R}\hat{\boldsymbol{\beta}} \sim \mathrm{N}[\mathbf{R}\boldsymbol{\beta}, \sigma^2 \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'] . \qquad (٦\text{ـ}٥٩)$$

حال به بررسی توزیع $\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ توجه می‌کنیم،

$$\mathrm{E}[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})] = \mathbf{R}\mathrm{E}(\hat{\boldsymbol{\beta}}) - \mathbf{R}\mathrm{E}(\hat{\boldsymbol{\beta}}) ,$$

$$= \mathbf{R}\boldsymbol{\beta} - \mathbf{R}\boldsymbol{\beta} = \mathbf{o} ,$$

همچنین

$$\mathrm{Var}[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})] = \mathrm{Var}(\mathbf{R}\hat{\boldsymbol{\beta}}) ,$$

زیرا $\boldsymbol{\beta}$ و $\mathbf{R}$ مقادیر ثابتی هستند. با جایگزینی معادلهٔ ٦ـ٥٨ در معادلهٔ فوق داریم

$$\mathrm{Var}[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})] = \sigma^2 \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' .$$

بدین ترتیب، با توجه به معادلهٔ ٦ـ٥٩ خواهیم داشت

$$\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) \sim \mathrm{N}[\mathbf{o}, \sigma^2 \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'] . \qquad (٦\text{ـ}٦٠)$$

اگر فرضیهٔ $H_0: \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ صحیح باشد، با جایگزینی $\mathbf{r}$ به جای $\mathbf{R}\boldsymbol{\beta}$ در رابطهٔ ٦ـ٦٠ داریم

$$(\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r}) \sim \mathrm{N}[\mathbf{o}, \sigma^2 \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'] . \qquad (٦\text{ـ}٦١)$$

از جبر ماتریسی می‌دانیم[1] که اگر $\boldsymbol{x} \sim \mathrm{N}(\mathbf{o}, \boldsymbol{\Sigma})$، آنگاه

$$\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x} \sim \chi^2(n) , \qquad (٦\text{ـ}٦٢)$$

---

۱. برای اثبات به پیوست «۵ـالف» مراجعه شود.

که در آن $(n \times 1) \to x$. رابطهٔ ٦ـ٦١ را با استفاده از معادلهٔ ٦ـ٦٢ به‌صورت زیر می‌نویسیم،

$$(R\hat{\beta} - r)' \left[\sigma^2 R (X'X)^{-1} R'\right]^{-1} (R\hat{\beta} - r) \sim \chi^2 (r) , \qquad (٦ـ٦٣)$$

که در آن، r درجات آزادی توزیع χ^2 است و برابر با مقدار ردیفهای بردار $R\hat{\beta}$، یعنی مقدار فرضیه‌هایی است که می‌خواهیم آزمون کنیم. همچنین می‌توان ثابت کرد که معکوس ماتریس $[R (X'X)^{-1} R']$ وجود دارد.

تمام اجزای رابطهٔ ٦ـ٦٣ بغیر از σ^2، مقادیر عددی معلوم و مشخصی است. از معادلهٔ ٦ـ٣٨ داریم

$$\frac{e'e}{\sigma_0^2} \sim \chi^2 (n - k) .$$

می‌دانیم، اگر دو توزیع مستقل χ^2 را بر درجات آزادی خود تقسیم کرده، سپس نسبت آنها را تشکیل دهیم، یک توزیع F خواهیم داشت. با تقسیم رابطهٔ ٦ـ٦٣ بر r و رابطهٔ ٦ـ٣٨ بر $(n - k)$ و تشکیل نسبت آنها اولاً σ^2 حذف می‌شود ثانیاً یک توزیع F با درجات آزادی $[r, (n - k)]$ خواهیم داشت. نتیجهٔ کلی اینکه اگر $H_0 : R\beta = r$ صحیح باشد آنگاه

$$\frac{\{(R\hat{\beta} - r)' \left[R (X'X)^{-1} R'\right]^{-1} (R\hat{\beta} - r)\} / r}{e'e / (n - k)} \sim F (r, n - k) . \qquad (٦ـ٦٤)$$

رابطهٔ فوق عمومی‌ترین آمارهٔ آزمون در مدلهای رگرسیون چندمتغیره است. اگر مقدار عددی آمارهٔ آزمون از مقدار جدول F بیشتر شود، یعنی آمارهٔ آزمون در ناحیهٔ بحرانی قرار بگیرد، فرضیهٔ H_0 رد می‌شود.

در اینجا باید نشان داد که چگونه در محاسبات عددی می‌توان از آمارهٔ عمومی آزمون استفاده کرد. در اینجا به چند مورد اشاره می‌کنیم تا قدرت و ظرفیت بسیار فرمول فوق در اجرای آزمون فرضیه‌های مختلف روشن شود.

۱. آزمون هریک از پارامترها

فرض کنید می‌خواهیم در مدل رگرسیون

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + U_t ,$$

فرضیهٔ $H_0: \beta_i = a$ را در مقابل $H_1: \beta_i \neq a$ آزمون کنیم. برای استفاده از آمارهٔ آزمون ٤ـ٦٤ باید $\mathbf{R}$ و $\mathbf{r}$ را تعریف کنیم:

$$\mathbf{R} = [0 \quad 0 \quad \cdots \quad 0 \quad \underset{\downarrow}{1} \quad \cdots \quad 0], \qquad \mathbf{r} = a .$$

i امین عنصر

بنابراین

$$(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}) = \beta_i - a .$$

باید $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ را برای حالت $\beta_i = a$ به دست آوریم. با توجه به ساختار بردار $\mathbf{R}$ ملاحظه می‌شود که $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$، یک اسکالر بوده، مقدار آن برابر با i امین عنصر قطری ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ است، یعنی

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' = a_{ii} ,$$

که a_{ii} درواقع i امین عنصر قطری ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ است. بدین ترتیب آمارهٔ عمومی آزمون، یعنی رابطهٔ ٦ـ٦٤ و با توجه به $r = 1$ به قرار زیر خواهد بود،

$$F = \frac{[(\beta_i - a)'(a_{ii})^{-1}(\beta_i - a)]/1}{\Sigma e_t^2/(n-k)} \sim F(1, n-k) .$$

معادلهٔ فوق را می‌توان به صورت زیر نوشت:

$$F = \frac{(\beta_i - a)^2}{a_{ii}\,\hat{\sigma}_u^2} \sim F(1, n-k) ,$$

که با توجه به معادلهٔ ٦ـ٢٤ داریم

$$F = \frac{(\beta_i - a)^2}{\mathrm{Var}(\hat{\beta}_i)} \sim F(1, n-k) . \qquad (٦ـ٦٥)$$

ملاحظه می‌شود که آمارهٔ آزمون در رابطهٔ ٦ـ٦٥ دقیقاً مجذور آمارهٔ t است که در رابطهٔ

۶ـ۲۶ به دست آوردیم. انتظار چنین نتیجه‌ای را نیز داشتیم زیرا طبق معادلهٔ ۲ـ۵۱ می‌دانیم $F= t^2$.

آزمون معنی‌دار بودن β_i

اگر بخواهیم فرضیهٔ $H_0: \beta_i = 0$ را در مقابل $H_1: \beta_i \neq 0$ برای یک مدل رگرسیون چندمتغیره آزمون کنیم، بردارهای $\mathbf{R}$ و $\mathbf{r}$ به صورت زیر تعریف می‌شود،

$$[0 \quad 0 \quad \dots \quad \underset{\downarrow}{1} \quad 0 \quad \dots \quad 0], \qquad r= 0.$$

i امین عنصر

بنابراین

$$\mathbf{R}\hat{\beta} - r = \beta_i .$$

اسکالر $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$، در این حالت برابر است با i امین عنصر قطری ماتریس $(\mathbf{X}'\mathbf{X})$، یعنی

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' = a_{ii} .$$

آمارهٔ عمومی آزمون برای این حالت و از رابطهٔ ۶ـ۶۴ و با توجه به $r=1$ عبارت است از

$$F= \frac{[(\beta_i)'(a_{ii})^{-1}(\beta_i)]/1}{\Sigma e_t^2/(n-k)} \sim F(1, n-k) .$$

در نتیجه خواهیم داشت

$$F= \frac{\beta_i^2}{\mathrm{Var}(\hat{\beta}_i)} \sim F(1, n-k) , \qquad (۶ـ۶۶)$$

که دقیقاً مجذور آمارهٔ t است که در رابطهٔ ۶ـ۲۸ به دست آوردیم.

۲. آزمون معنی‌دار بودن زیرمجموعه‌ای از پارامترها

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t= \beta_1 + \beta_2 X_{t1} + \dots + \beta_k X_{tk} + U_t .$$

می‌خواهیم، فرضیه زیر را آزمون کنیم،

$$H_0: \beta_{k-r+1} = \beta_{k-r+2} = \cdots = \beta_k = 0 .$$

فرضیه فوق دلالت بر این می‌کند که باید r پارامتر آخر مدل رگرسیون آزمون شود. ممکن است در بعضی موارد تمام پارامترهایی که می‌خواهیم تساوی آنها را با صفر آزمون کنیم، به طور منظم پارامترهای مدل را تشکیل ندهند. در این صورت می‌توان با یک جابجایی ساده در ترتیب نوشتن متغیرهای توضیحی، پارامترهای مورد نظر را در انتهای مدل قرار داد.

برای آزمون تساوی r پارامتر با صفر، مطابق معمول باید $\mathbf{R}$ و $\mathbf{r}$ را برای این حالت تعریف کنیم،

$$\mathbf{R} = [\mathbf{0} \quad \mathbf{I}_r], \qquad \mathbf{r} = \mathbf{0} .$$

توجه داریم که در اینجا $\mathbf{R}$ را به دو ماتریس $\mathbf{0}$ و $\mathbf{I}_r$ افراز کرده‌ایم. همچنین می‌دانیم $\mathbf{R} \to r \times k$، $\mathbf{0} \to k \times (k-r)$، $\mathbf{I} \to k \times r$ و $\mathbf{r} \to r \times 1$. ماتریس $\mathbf{X}$ و بردار $\hat{\beta}$ را مدل رگرسیون $\mathbf{y} = \mathbf{X}\hat{\beta} + \mathbf{e}$ را نیز به ترتیبی مشابه افراز می‌کنیم،

$$\mathbf{y} = [\mathbf{X}_s \quad \mathbf{X}_r] \begin{bmatrix} \hat{\beta}_s \\ \hat{\beta}_r \end{bmatrix} + \mathbf{e} = \mathbf{X}_s\hat{\beta}_s + \mathbf{X}_r\hat{\beta}_r + \mathbf{e} , \qquad (۶.۶۷)$$

که در آن $\mathbf{X}_s \to n \times (k-r)$ و $\mathbf{X}_r \to n \times r$. همچنین $\mathbf{X}_s$ و $\mathbf{X}_r$ به ترتیب $(k-r)$ ستون اول و r ستون آخر ماتریس $\mathbf{X}$ است.

برای محاسبهٔ آمارهٔ آزمون، یعنی رابطهٔ ۶.۶۴، مطابق معمول باید $(\mathbf{R}\hat{\beta} - \mathbf{r})$ و $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ را به دست آوریم. با توجه به تعاریف $\mathbf{r}$ و $\mathbf{R}$ داریم

$$\mathbf{R}\hat{\beta} - \mathbf{r} = \hat{\beta}_r , \qquad (۶.۶۸)$$

و نیز با دقت در ماتریس $\mathbf{X}$ و بردار $\mathbf{R}$، به سهولت می‌توان گفت که $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ برابر است با ماتریس مربع مرتبهٔ r که در سمت راست پایین ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ وجود دارد.

اگر این ماتریس را با $\mathbf{C}_{۲۲}$ نشان دهیم، آنگاه با توجه به ساختار ماتریس $(\mathbf{X}'\mathbf{X})$، یعنی

$$\mathbf{X}'\mathbf{X}=\begin{bmatrix}\mathbf{X}_۱'\mathbf{X}_۱ & \mathbf{X}_۱'\mathbf{X}_۲\\ \mathbf{X}_۲'\mathbf{X}_۱ & \mathbf{X}_۲'\mathbf{X}_۲\end{bmatrix},$$

می‌توان نشان داد که[۱]

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-۱}\mathbf{R}'=\mathbf{C}_{۲۲}=(\mathbf{X}_۲'\mathbf{M}_۱\mathbf{X}_۲)^{-۱}, \qquad (۶.۶۹)$$

که در آن:

$$\mathbf{M}_۱=\mathbf{I}-\mathbf{X}_۱(\mathbf{X}_۱'\mathbf{X}_۱)^{-۱}\mathbf{X}_۱'.$$

با مراجعه به معادله‌های ۶.۶۸ و ۶.۶۹ می‌توان آمارهٔ عمومی آزمون، یعنی معادلهٔ ۶.۶۴ را برای این حالت به قرار زیر نوشت،

$$F=\frac{[\hat{\boldsymbol{\beta}}_۲'(\mathbf{X}_۲'\mathbf{M}_۱\mathbf{X}_۲\hat{\boldsymbol{\beta}}_۲)]/r}{\mathbf{e}'\mathbf{e}/(n-k)} \sim F(r,n-k). \qquad (۶.۷۰)$$

با مقایسهٔ آمارهٔ آزمون با مقدار جدول F، در صورتی فرضیهٔ $H_۰$ رد می‌شود که آمارهٔ آزمون در ناحیهٔ بحرانی واقع شود. میتوان نشان داد که معادلهٔ ۶.۳۳، یعنی $F=\frac{(RSS_r-RSS)/r}{RSS/(n-k)}$، حاصل معادلهٔ ۶.۷۰ است.[۲]

۳. آزمون معنی‌دار بودن مدل رگرسیون

مدل رگرسیون زیر را ملاحظه کنید،

$$Y_t=\beta_۱+\beta_۲X_{۲t}+\cdots+\beta_kX_{kt}+U_t.$$

می‌خواهیم فرضیهٔ

$$H_۰:\beta_۲=\beta_۳=\cdots=\beta_k=۰$$

---

۱. به قاعدهٔ معکوس کردن ماتریسهای افراز شده در پیوست «۵ ـ الف» مراجعه شود.
۲. برای اثبات به مسألهٔ ۶.۹ مراجعه کنید.

را آزمون کنیم. ابتدا $\mathbf{R}$ و $\mathbf{r}$ را در رابطهٔ ٤ـ٦٤ برای این حالت تعریف می‌کنیم،

$$\mathbf{R}=\begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ & & & & & \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}, \qquad \mathbf{r}=\mathbf{0},$$

که در آن $\mathbf{R} \to (k-1)\times k$ و $\mathbf{r} \to (k-1)\times 1$. برای ساختن آمارهٔ عمومی آزمون، یعنی معادلهٔ ٦ـ٦٤، باید مقادیر $(\mathbf{R}\hat{\beta}-\mathbf{r})$ و $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ را برای این حالت به دست آورد. خواهیم داشت

$$(\mathbf{R}\hat{\beta}_r-\mathbf{r})=\begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}=\hat{\beta}_r. \qquad (٦ـ٧١)$$

همچنین با توجه به ماتریس $\mathbf{R}$، می‌توان بسادگی نشان داد که $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ (که یک ماتریس $(k-1)\times(k-1)$ است) با آخرین $(k-1)$ سطر و ستون ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ برابر است. یادآوری می‌کنیم که بردار $\hat{\beta}_r$ در معادلهٔ ٦ـ٧١ دارای $(k-1)$ عنصر بوده، فاقد $\hat{\beta}_1$ است. بدین ترتیب آمارهٔ عمومی آزمون ٦ـ٦٤ برای فرضیهٔ H_0 عبارت است از

$$F=\frac{\{\hat{\beta}_r'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\hat{\beta}_r\}/(k-1)}{\mathbf{e}'\mathbf{e}/(n-k)} \sim F(k-1,n-k). \qquad (٦ـ٧٢)$$

اگر آمارهٔ آزمون فوق معنی‌دار باشد، فرضیهٔ H_0 رد می‌شود.[1]

٤. نواحی اطمینان در رگرسیون چند متغیره

فاصلهٔ اطمینان برای یک پارامتر، به راحتی از طریق آمارهٔ t به دست می‌آید. در قسمت

---

۱. می‌توان نشان داد که معادلهٔ ٦ـ٧٢ دقیقاً همان معادلهٔ ٦ـ٣٦ یا ٦ـ٤١ است. برای اثبات به مسألهٔ ٦ـ١٠ مراجعه شود.

۶-۶ به بررسی ناحیهٔ اطمینان برای دو پارامتر پرداختیم. در این قسمت این مسأله را در حالت کلی بررسی خواهیم کرد.

رابطهٔ ۶-۶۰ را یک بار دیگر می‌نویسیم،

$$\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) \sim N[o, \sigma^2 \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'] .$$

با توجه به معادلهٔ ۶-۶۲ خواهیم داشت:

$$[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]'[\sigma^2 \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})] \sim \chi^2(r) ,$$

که در آن r درجات آزادی توزیع χ^2 بوده، با تعداد ردیفهای ماتریس $\mathbf{R}$ برابر است. از معادلهٔ ۶-۳۸ نیز می‌دانیم که

$$\frac{e'e}{\sigma_u^2} \sim \chi^2(n-k) .$$

اگر دو توزیع مستقل χ^2 را بر درجات آزادی خود تقسیم نموده، سپس نسبت آنها را تشکیل دهیم نتیجه، یک توزیع F خواهد شد. اگر دقیقاً مشابه روش به کار رفته در معادلهٔ ۶-۶۴ عمل کنیم، خواهیم داشت

$$F=\frac{[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]/r}{e'e/(n-k)} \sim F(r,n-k) . \quad (۶-۷۳)$$

به کمک معادلهٔ ۶-۷۳ می‌توان برای هر گروهی از پارامترهای موجود در یک مدل رگرسیون چندمتغیره، یک ناحیهٔ اطمینان ساخت. بنابراین از نظر محاسباتی، باید اولاً، پارامترهایی را که می‌خواهیم برای آنها ناحیهٔ اطمینان مشترک بسازیم، تعیین کنیم و ثانیاً، ماتریس $\mathbf{R}$ را به تناسب پارامترهایی که مورد نظر است کامل کنیم و مقدار موجود در جدول F را به دست آوریم. ناحیهٔ اطمینان، فضایی است که آمارهٔ آزمون، یعنی مقدار محاسبهٔ شدهٔ F، در معادلهٔ ۶-۷۳ از مقدار F در جدول F کمتر باشد؛ بنابراین اگر نامساوی زیر برقرار باشد، H_0 صحیح است،

$$\frac{[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}[\mathbf{R}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]/r}{e'e/(n-k)} < F(r,n-k) . \quad (۶-۷۴)$$

مثال ۶.۱۶ مدلی که در مثال ۶.۲ مطرح کردیم را یک بار دیگر در نظر می‌گیریم،

$$Y_t = \alpha + \beta X_t + \gamma Z_t + U_t .$$

با استفاده از محاسبات مثال ۵.۲، این مدل را به صورت زیر تخمین زده‌ایم،

$$\hat{Y}_t = ۴ + ۲/۵ X_t - ۱/۵ Z_t .$$

اولاً، برای β و γ ناحیهٔ اطمینان مشترک در سطح احتمال ۹۵ درصد بسازید.
ثانیاً، با استفاده از آمارهٔ عمومی آزمون، یعنی معادلهٔ ۶.۶۴، فرضیهٔ

$$H_0 : \beta + \gamma = ۰$$

را در مقابل فرضیهٔ $H_1 : \beta + \gamma \neq ۰$ در سطح معنی‌دار ۵ درصد آزمون کنید.

۱. برای ساختن نواحی اطمینان باید از معادلهٔ ۶.۷۳ استفاده کنیم. ابتدا، ماتریس $\mathbf{R}$ را تشکیل می‌دهیم. با توجه به اینکه می‌خواهیم برای دو پارامتر β و γ ناحیهٔ اطمینان بسازیم، خواهیم داشت

$$\mathbf{R} = \begin{bmatrix} ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix} .$$

باید $\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ و $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ را محاسبه کنیم. داریم

$$\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \begin{bmatrix} ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix} \begin{bmatrix} (\hat{\alpha} - \alpha) \\ (\hat{\beta} - \beta) \\ (\hat{\gamma} - \gamma) \end{bmatrix} = \begin{bmatrix} (\hat{\beta} - \beta) \\ (\hat{\gamma} - \gamma) \end{bmatrix} ,$$

در نتیجه با استفاده از $\hat{\beta} = ۲/۵$ و $\hat{\gamma} = ۱/۵$ نتیجه می‌شود که

$$\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \begin{bmatrix} (۲/۵ - \beta) \\ (-۱/۵ - \gamma) \end{bmatrix} .$$

مفید است به این نکته توجه کنیم که چون نمی‌خواهیم برای α ناحیهٔ اطمینان بسازیم؛ بنابراین می‌توان ستون اول ماتریس $\mathbf{R}$ و سطر اول بردار $(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ را حذف کرد.

برای محاسبهٔ $\mathbf{R}(\mathbf{X}'\mathbf{X})\mathbf{R}'$، با توجه به اینکه ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ در مثال ۶ـ۱ برحسب مقادیر انحراف از میانگین ارائه شده است، باید ستون اول ماتریس $\mathbf{R}$ و سطر اول بردار $(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ را حذف کنیم. به این ترتیب

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' = \begin{bmatrix} ۰ & ۱ \\ ۰ & ۱ \end{bmatrix} \begin{bmatrix} ۱ & -۱/۵ \\ -۱/۵ & ۲/۵ \end{bmatrix} \begin{bmatrix} ۰ & ۱ \\ ۰ & ۱ \end{bmatrix},$$

$$= \begin{bmatrix} ۱ & -۱/۵ \\ -۱/۵ & ۲/۵ \end{bmatrix}.$$

باید $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ را معکوس کرد

$$[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} = \begin{bmatrix} ۱ & -۱/۵ \\ -۱/۵ & ۲/۵ \end{bmatrix}^{-1} = \begin{bmatrix} ۱۰ & ۶ \\ ۶ & ۴ \end{bmatrix}.$$

محاسبات فوق را در معادلهٔ ۶ـ۷۳ قرار داده، آمارهٔ آزمون به شرح زیر محاسبه می‌شود:

$$F = \left\{ [(۲/۵-\beta) \quad (-۱/۵-\gamma)] \begin{bmatrix} ۱۰ & ۶ \\ ۶ & ۴ \end{bmatrix} \begin{bmatrix} (۲/۵-\beta) \\ (-۱/۵-\beta) \end{bmatrix} \right\} / ۲ \div \frac{۱/۵}{۵-۳},$$

$$= \frac{۲۶/۵ - ۳۲\beta - ۱۸\gamma + ۱۲\beta\gamma + ۱۰\beta^۲ + ۴\gamma^۲}{۱/۵}.$$

مقدار F در سطح معنی‌دار ۵ درصد و با درجات آزادی ۲ و ۲ برابر است با

$$F_{٪۵}(۲,۲) = ۱۹.$$

بنابراین براساس معادلهٔ ۶ـ۷۴ می‌توان چنین نوشت،

$$\frac{۱۰\beta^۲ + ۱۲\beta\gamma + ۴\gamma^۲ - ۳۲\beta - ۱۸\gamma + ۲۶/۵}{۱/۵} - ۱۹ = ۰,$$

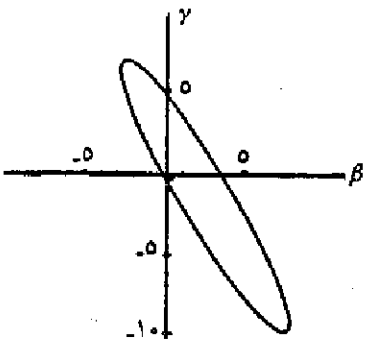
یا بعد از ساده کردن داریم

$$۱۰\beta^۲ + ۱۲\beta\gamma + ۴\gamma^۲ - ۳۲\beta - ۱۸\gamma - ۲ = ۰.$$

معادلهٔ فوق یک بیضی است که در نمودار ۶ـ۳ ترسیم شده است. مرکز این بیضی در نقاط تخمین ($\hat{\beta}=2/5, \hat{\gamma}=-1/5$) قرار دارد. در مثال ۶ـ۱۳ و در بحث نمودار ۶ـ۲ دیدیم که اگر کوواریانس بین پارامترها مثبت باشد، بیضی به سمت راست متمایل است، در غیر این صورت تمایل به چپ دارد. با توجه به محاسبات موجود در مثال ۶ـ۱ ملاحظه می‌شود که کوواریانس $\hat{\beta}, \hat{\gamma}$ منفی است، زیرا

$$\mathrm{Cov}(\hat{\beta},\hat{\gamma}) = \hat{\sigma}^2 (a_{12}) = 0/75\,(-1/5) = -1/125\,,$$

که در آن a_{12} عنصر ردیف اول و ستون دوم ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ است؛ بنابراین بیضی ناحیهٔ اطمینان برای β و γ به سمت چپ متمایل است.



نمودار ۶ـ۳ ناحیهٔ اطمینان برای β و γ

به وسیله ناحیهٔ اطمینان مشترک می‌توان فرضیه‌های مشترک در مورد پارامترها را آزمون کرد. مثلاً در مثال ۶ـ۴ فرضیهٔ $H_0: \beta=\gamma=0$ را آزمون کردیم و ملاحظه شد که H_0 رد نمی‌شود. در نمودار ۶ـ۳ نیز می‌توان گفت که چون $\beta=\gamma=0$ درون بیضی قرار دارد؛ بنابراین در ناحیهٔ اطمینان قرار گرفته است و در نتیجه H_0 رد نخواهد شد. بدیهی است در مواردی که می‌خواهیم ناحیهٔ اطمینان را برای بیش از دو پارامتر تعیین کنیم، ترسیم هندسی مشکل یا غیرممکن است. در این موارد کافی است بعد از به دست آوردن رابطهٔ ناحیهٔ اطمینان، یعنی نامساوی ۶ـ۷۴، فرضیهٔ مشترک پارامترها را در آن، قرار دهیم. اگر H_0 در این نامساوی صدق کند صحیح است و برعکس.

۲. برای آزمون فرضیهٔ $H_0: \beta + \gamma = 0$ به کمک آمارهٔ عمومی آزمون، یعنی معادلهٔ ٦٤-٦، باید $(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})$ و $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ را محاسبه کنیم. با توجه به اینکه $(\mathbf{X}'\mathbf{X})^{-1}$ در مثال ٦-١ برحسب انحراف از میانگین گزارش شده است، به نوشتن عنصر اول در بردار $\mathbf{R}$ نیازی نیست. خواهیم داشت

$$\mathbf{R} = [1 \quad 1] \qquad \text{و} \qquad \mathbf{r} = 0 .$$

بنابراین $(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})$ برابر است با

$$(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}) = \hat{\beta} + \hat{\gamma} .$$

با توجه به ساختار بردار $\mathbf{R}$ ملاحظه می‌شود که $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ یک ماتریس ۲ × ۲ بوده و برابر است با

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' = [1 \quad 1] \begin{bmatrix} 1 & -1/5 \\ -1/5 & 2/5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0/5 .$$

بدین ترتیب آمارهٔ آزمون عبارت است از

$$F = \frac{\{(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})' [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})\} / r}{\mathbf{e}'\mathbf{e} / (n - k)} ,$$

$$= \frac{(\hat{\beta} + \hat{\gamma})' [0/5]^{-1} (\hat{\beta} + \hat{\gamma}) / 1}{15 / (5 - 3)} = \frac{(\hat{\beta} + \hat{\gamma})^2}{0/375} = 2/66 .$$

مقدار جدول F در سطح معنی‌دار ۵ درصد و با درجات آزادی (۲ , ۱) برابر است با ۱۸/۵۱. نتیجه می‌گیریم که H_0 رد نمی‌شود.

مسائل فصل ششم

۶-۱ با روش حداقل مربعات معمولی و به کمک ۸۰ مشاهده فصلی ($n=۸۰$)، یک مدل رگرسیون را به صورت زیر تخمین زده‌ایم،

$$\hat{Y}_t = \underset{(۳/٤)}{۲/۲۰} + \underset{(۰/۰۰۵)}{۰/۱۰٤\,X_{۱t}} + \underset{(۲/۲)}{۳/٤۸\,X_{۲t}} + \underset{(۰/۱۵)}{۰/۳٤\,X_{۳t}}\ .$$

اعداد داخل پرانتز مقادیر انحراف از معیار تخمین پارامترهاست. مجموع مربع مقادیر توضیح داده شده و مجموع مربعات پسماند به ترتیب برابر است با تغییرات توضیح داده شده $=۱۱۲/۵$ و تغییرات توضیح داده نشده $=۱۹/۵$.

الف) در سطح معنی‌دار ۵ درصد، کدام یک از ضرایب متغیرهای توضیحی به‌طور معنی‌دار با صفر متفاوتند؟

ب) مقدار $R^۲$ را برای این مدل حساب کنید.

ج) مقدار $\bar{R}^۲$ را نیز به دست آورید.

۶-۲ مدل رگرسیون زیر مفروض است،

$$Y_t = \beta_۰ + \beta_۱ X_{۱t} + \beta_۲ X_{۲t} + \beta_۳ X_{۳t} + U_t\ .$$

با استفاده از ۲۶ مشاهده، این مدل را تخمین زده و نتایج زیر را به دست آورده‌ایم،

$$\hat{Y}_t = \underset{(۱/۹)}{۲} + \underset{(۲/۲)}{۳/۵\,X_{۱t}} - \underset{(۱/۵)}{۰/۷X_{۲t}} + ۲\,X_{۳t}\ .$$

اعداد داخل پرانتز، مقادیر آمارهٔ آزمون t برای هریک از پارامترهاست. همچنین مقدار $R^۲$ برابر است با ۰/۹۸۲ . یک بار دیگر همین مدل را با فرض $\beta_۱=\beta_۲$ تخمین زده و نتیجه زیر را ملاحظه کرده‌ایم،

$$\hat{Y}_t = \underset{(۲/۷)}{۱/۵} + \underset{(۲/٤)}{۳(X_{۱t} + X_{۲t}) - ۰/٦X_{۳t}}\ .$$

همچنین مقدار R^2 برای این مدل برابر ۰/۸۷۶ است.

الف) فرضیهٔ $H_0: \beta_1 = \beta_2$ را آزمون کنید.

ب) اگر متغیر توضیحی X_{t1} را از این معادله حذف کنیم، آیا به نظر شما $\bar{R}^2$ زیاد می‌شود یا کم؟

ج) آیا با حذف X_{t1} مقدار R^2 کم می‌شود؟

۶-۳ مدل تغییرات نرخ دستمزد به صورت زیر مفروض است،

$$Y_t = \alpha + \beta + X_t + \gamma X_t^2 + U_t ,$$

که در آن Y_t نرخ تغییرات دستمزد و X_t نرخ بیکاری است. با استفاده از ۱۷ مشاهده، تخمین زیر را به دست آورده‌ایم،

$$\hat{Y}_t = \underset{(۹/۵۲)}{۲۳/۵۱۱۲} - \underset{(۱۲/۶۸)}{۲۱/۶۷۱۵ X_t} + \underset{(۴/۱۱)}{۵/۸۲۵۷ X_t^2} .$$

همچنین می‌دانیم $R^2 = ۰/۴۰۶۲$.

الف) دو فرضیهٔ $H_0: \beta = ۰$ و $H_0: \gamma = ۰$ را در مقابل فرضیهٔ مخالف صفر در سطح معنی‌دار ۵ درصد به طور جداگانه آزمون کنید.

ب) فرضیهٔ $H_0: \gamma = \beta = ۰$، یعنی معنی‌دار بودن کل مدل را در سطح ۵ درصد آزمون کنید.

ج) به نظر شما تفاوت بین دو استنتاج آماری در بندهای «الف» و «ب» را چگونه می‌توان توضیح داد؟

۶-۴ مدل زیر مفروض است،

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i .$$

محاسبات زیر برحسب انحراف از میانگین و با استفاده از ۲۳ مشاهده انجام شده است،

$$\Sigma x_{1i}^2 = ۱۲ \quad , \quad \Sigma x_{1i} x_{2i} = ۸ \quad , \quad \Sigma x_{2i}^2 = ۱۲ ,$$

$$\Sigma x_{1i} y_i = ۱۰ \quad , \quad \Sigma x_{2i} y_i = ۸ \quad , \quad \Sigma y_i^2 = ۱۰ .$$

با استفاده از مشاهدات فوق، مدل را به صورت زیر تخمین زده‌ایم،

$$\hat{Y}_i = ۴ + ۰/۷X_{1i} + ۰/۲X_{2i} \quad , \quad R^2 = ۰/۸۶ .$$

فرضیه زیر را در سطح معنی‌دار ۵ درصد با آمارهٔ F و آمارهٔ t آزمون کنید.

$$H_0: \frac{\beta_1}{\beta_2} = \frac{۵}{۳} \quad , \qquad H_1: \frac{\beta_1}{\beta_2} \neq \frac{۵}{۳} .$$

۶ـ۵ مدل رگرسیون زیر را که برحسب انحراف از میانگین است ملاحظه نمایید،

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + U_t .$$

اطلاعات زیر موجود است،

$$n = ۱۰۰ \quad , \quad \Sigma y_t^2 = \frac{۴۹۳}{۳} \quad , \quad \Sigma x_{1t}^2 = ۳۰ \quad , \quad \Sigma x_{2t}^2 = ۳ ,$$

$$\Sigma x_{1t} y_t = ۳۰ \quad , \quad \Sigma x_{2t} y_t = ۲۰ \quad , \quad \Sigma x_{1t} x_{2t} = ۰ .$$

الف) پارامترهای β_1 و β_2 را تخمین بزنید. R^2 را نیز به دست آورید.

ب) فرضیهٔ $H_0: \beta_2 = ۷$ را در مقابل $H_1: \beta_2 \neq ۷$ آزمون کنید.

ج) فرضیهٔ $H_0: \beta_1 = \beta_2 = ۰$ را در مقابل $\beta_2 \neq ۰$ یا $H_1: \beta_1 \neq ۰$ آزمون کنید.

د) فرضیهٔ $H_0: \beta_2 = ۷\beta_1$ را در مقابل $H_1: \beta_2 \neq ۷\beta_1$ آزمون کنید.

می‌توان سطح معنی‌دار بودن هریک از آزمونها را ۵ درصد گرفت. تمام محاسبات را به زبان ماتریسی انجام دهید.

۶ـ۶ مدل رگرسیون زیر به عنوان یک تابع تولید تخمین زده شده است،

$$\ln Q_i = \underset{}{۱/۳۷} + \underset{(۰/۲۵۷)}{۰/۶۳۲} \ln K_i + \underset{(۰/۲۱۹)}{۰/۴۵۲} \ln L_i ,$$

$$R^2 = ۰/۹۸ \quad , \qquad \mathrm{Cov}(\beta_K , \beta_L) = ۰/۰۵۵ .$$

اعداد داخل پرانتز، مقادیر انحراف از معیار تخمین پارامترها هستند.

فرضیه‌های زیر را در سطح معنی‌دار ۵ درصد آزمون کنید.

الف) کشش تولید نسبت به سرمایهٔ K و نسبت به نیروی کار L با یکدیگر برابر است.

ب) بازده ثابت نسبت به مقیاس وجود دارد.

همان‌گونه که ملاحظه می‌شود، در این مسأله تعداد مشاهدات داده نشده است. آیا این امر در استنتاجات شما تأثیری دارد؟

۶ـ۷ مدل رگرسیون زیر مفروض است،

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + U_t .$$

محاسبات زیر را براساس ۲۴ مشاهده و برحسب انحراف از میانگین داریم،

$$\Sigma y_t^2 = ۶۰ \quad , \quad \Sigma x_{1t} y_t = ۷ \quad , \quad \Sigma x_{1t} x_{2t} = ۱۰ ,$$

$$\Sigma x_{1t}^2 = ۱۰ \quad , \quad \Sigma x_{2t} y_t = -۷ \quad , \quad \Sigma x_{1t} x_{3t} = ۵ ,$$

$$\Sigma x_{2t}^2 = ۳۰ \quad , \quad \Sigma x_{3t} y_t = -۲۶ \quad , \quad \Sigma x_{2t} x_{3t} = ۱۵ .$$

$$\Sigma x_{3t}^2 = ۲۰ \quad ,$$

الف) فرضیه‌های $\beta_1 = ۱$ ، $\beta_2 = ۱$ و $\beta_3 = -۲$ را آزمون کنید.

ب) فرضیهٔ $\beta_1 + \beta_2 + \beta_3 = ۰$ را آزمون کنید.

ج) سه فرضیهٔ بند «الف» را به طور همزمان آزمون کنید؛ یعنی

$$H_0 : [\beta_1 \quad \beta_2 \quad \beta_3] = [۱ \quad ۱ \quad -۲] .$$

فرضیه‌ها را در سطح معنی‌دار ۵ درصد آزمون نمایید.

۶ـ۸ از یک مجموعه اعداد، نتایج تخمینهای زیر برای دو مدل رگرسیون به‌دست آمده است،

$$Y_t = \underset{(۲)}{۱} + \underset{(۴)}{۲} X_t \quad , \qquad R^2 = ۰/۲ ,$$

$$Y_t = \underset{(۱/۸)}{۱/۵} + \underset{(۲)}{۳} X_t + \underset{(۳)}{۴} W_t \quad , \qquad R^2 = ۰/۶ .$$

اعداد داخل پرانتز مقادیر آمارهٔ t مربوط به پارامترهاست. تعداد مشاهدات را به دست آورید.

۶ـ۹ در مدل رگرسیون $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ شامل k پارامتر، می‌خواهیم فرضیهٔ صفر بودن r پارامتر را آزمون کنیم، یعنی

$$H_0: \beta_{k-r+1} = \beta_{k-r+2} = \cdots = \beta_k = 0 .$$

در معادلهٔ ۶ـ۷۰ نشان دادیم که برای این آزمون می‌توان از آمارهٔ زیر استفاده کرد،

$$F = \frac{[\hat{\boldsymbol{\beta}}_r' (\mathbf{X}_r' \mathbf{M}_s \mathbf{X}_r) \hat{\boldsymbol{\beta}}_r] / r}{\mathbf{e}'\mathbf{e} / (n-k)} \sim F(r, n-k) .$$

نشان دهید که از رابطهٔ فوق، می‌توان به آمارهٔ زیر رسید،

$$F = \frac{(RSS_r - RSS) / r}{RSS / (n-k)} \sim F(r, n-k) .$$

۶ـ۱۰ در مدل رگرسیون $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ شامل k پارامتر، می‌خواهیم فرضیهٔ معنی‌دار بودن کل مدل را آزمون کنیم، یعنی

$$H_0: \beta_2 = \beta_3 = \cdots = \beta_k = 0 .$$

در معادلهٔ ۶ـ۷۲ نشان دادیم که برای این آزمون میتوان از آمارهٔ زیر استفاده کرد،

$$F = \frac{\{\hat{\boldsymbol{\beta}}_r' [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} \hat{\boldsymbol{\beta}}_r\} / (k-1)}{\mathbf{e}'\mathbf{e} / (n-k)} \sim F(k-1, n-k) .$$

نشان دهید که از رابطهٔ فوق می‌توان به آمارهٔ زیر رسید،

$$F = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}} / (k-1)}{\mathbf{e}'\mathbf{e} / (n-k)} = \frac{ESS / (k-1)}{RSS / (n-k)} \sim F(k-1, n-k) .$$

حل مسائل فصل ششم

۶-۱ الف) درجات آزادی برابر است با ۷۶ = ٤ - ۸۰ = d.F. . مقدار t از جدولی در سطح معنی‌دار ۲۵ درصد برابر است با ۱/۹۸ . هرگاه آمارهٔ آزمون از ۱/۹۸ بیشتر باشد، H_0 رد می‌شود؛ یعنی ضریب مورد نظر به طور معنی‌داری متفاوت با صفر است. اگر مدل مفروض را به صورت زیر بنویسیم،

$$Y_t = \beta_1 + \beta_2 X_{t1} + \beta_3 X_{t2} + \beta_4 X_{t3} + U_t$$

با فرض $H_0: \beta_i = 0$ و برای مقادیر $i = 2, 3, 4$، داریم

$$t_2 = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)} = \frac{0/104}{0/005} = 20/8 ,$$

بنابراین t_2 معنی‌دار بوده، H_0 رد می‌شود، یعنی β_2 معنی‌دار است.

$$t_3 = \frac{\hat{\beta}_3}{SE(\hat{\beta}_3)} = \frac{3/48}{2/2} = 1/581 ,$$

پس آمارهٔ آزمون معنی‌دار نیست و H_0 رد نمی‌شود.

$$t_4 = \frac{\hat{\beta}_4}{SE(\hat{\beta}_4)} = \frac{0/34}{0/15} = 2/26 ,$$

در نتیجه H_0 رد خواهد شد.

ب) می‌دانیم تغییرات توضیح داده نشده + تغییرات توضیح داده شده = کل تغییرات
بنابراین ۱۳۲ = ۱۹/۵ + ۱۱۲/۵ = کل تغییرات. در نتیجه

$$R^2 = \frac{ESS}{TSS} = \frac{112/5}{132} = 0/85 .$$

ج) $\bar{R}^2 = 1 - \left(\frac{n-1}{n-k}\right)(1-R^2) = 1 - \frac{80-1}{80-4}(1-0/85) = 0/84$.

۶ـ۲ الف) مدل دوم را ـ که براساس درستی H_0 به دست آمده است ـ مدل مقید می‌نامیم. آمارهٔ آزمون F را با توجه به مبحث ۶ـ۷ و نیز معادله‌های ۶ـ۳۳ یا ۶ـ۵۱ محاسبه می‌کنیم. می‌دانیم

$$F = \frac{(RSS_r - RSS)/r}{RSS/(n-k)} \sim F(r, n-k),$$

که در آن RSS_r و RSS به ترتیب مجموع مربعات پسماند در مدل مقید و مدل اولیه است. کل تغییرات Σy_t^2 برای هر دو مدل یکی است. برای مدل اول و مدل مقید داریم

$$RSS_r = (1 - R_r^2)\Sigma y_t^2 = (1 - ۰/۸۷۶)\Sigma y_t^2 = ۰/۱۲۴ \Sigma y_t^2,$$

$$RSS = (1 - R^2)\Sigma y_t^2 = (1 - ۰/۹۸۲)\Sigma y_t^2 = ۰/۰۱۸ \Sigma y_t^2.$$

بنابراین با جایگزینی معادله‌های فوق در آمارهٔ آزمون F خواهیم داشت

$$F = \frac{(۰/۱۲۴ \Sigma y_t^2 - ۰/۰۱۸ \Sigma y_t^2)/۱}{(۰/۰۱۸ \Sigma y_t^2)/(۲۶ - ۴)} = ۱۲۹/۶.$$

با درجات آزادی ۱ و ۲۲، در سطح معنی‌دار ۵ درصد، از جدول F داریم

$$F_{۵٪}(۱, ۲۲) = ۴/۳۰,$$

نتیجه می‌گیریم که آمارهٔ آزمون معنی‌دار است؛ بنابراین فرضیهٔ $H_0: \beta_1 = \beta_2$ رد می‌شود.

ب) در بند ۶ قسمت ۶ـ۵ به طور خلاصه به این نکته اشاره کردیم که هنگامی یک متغیر توضیحی می‌تواند $\bar{R}^2$ را زیاد کند که مقدار عددی آمارهٔ آزمون t پارامتر آن متغیر از یک بیشتر باشد و برعکس. در این مسأله مقدار t مربوط به $X_{۳i}$ برابر ۰/۷ـ است. بنابراین با حذف $X_{۳i}$ از مدل مقدار $\bar{R}^2$ افزایش خواهد یافت.

ج) با حذف یک متغیر توضیحی مقدار R^2 معمولاً کاهش می‌یابد؛ بنابراین پاسخ این قسمت می‌تواند مثبت باشد.

۶ـ۳ الف) مقدار جدول t با $n - k = ۱۷ - ۳ = ۱۴$ درجهٔ آزادی و در سطح معنی‌دار ۲۵ درصد برابر است با ۲/۱۴۵ ± . مقادیر آمارهٔ t برای β و γ به ترتیب عبارت است از

(۱/۱۷۱-) و (۱/۴۲)؛ بنابراین از مقدار جدول t کمتر است و معنی‌دار نیست. پس فرضیه‌های $\beta = ۰$ و $\gamma = ۰$ رد نخواهد شد.

ب و ج) آمارهٔ F را به کمک $R^۲$ تشکیل می‌دهیم. می‌دانیم اگر $H_۰$ صحیح باشد، آمارهٔ زیر دارای توزیع F با درجات آزادی $(k-۱)$ و $(n-k)$ خواهد بود،

$$F = \frac{R^۲/(k-۱)}{(۱-R^۲)/(n-k)} \sim F(k-۱, n-k),$$

$$F = \frac{۰/۴۰۶۲}{(۱-۰/۴۰۶۲)} \cdot \frac{(۱۷-۳)}{(۳-۱)} = ۴/۷۸۷۹.$$

اما مقدار جدول F برابر است با

$$F_{۵\%}(۲, ۱۴) = ۳/۷۴,$$

بنابراین آمارهٔ آزمون از این مقدار بیشتر می‌شود و معنی‌دار است. بنابراین فرضیهٔ $H_۰: \beta = \gamma = ۰$ رد می‌شود.

ملاحظه می‌شود که هریک از پارامترها در آزمونهای جداگانه معنی‌دار نیست، در حالی که در آزمون همزمان معنی‌دار است. علت این امر می‌تواند همبستگی بین متغیرهای توضیحی باشد.

۶-۴ فرضیهٔ $H_۰$ را می‌توان به صورت زیر نوشت،

$$H_۰: \beta_۲ = ۰/۶\beta_۱ \quad , \quad H_۱: \beta_۲ \neq ۰/۶\beta_۱ .$$

فرض می‌کنیم $H_۰$ صحیح است. با جایگزینی $H_۰$ در مدل مفروض، مدل مقید را به دست می‌آوریم،

$$Y_i = \alpha + \beta_۱ X_{۱i} + ۰/۶\beta_۱ X_{۲i} + U_i ,$$

$$= \alpha + \beta_۱ (X_{۱i} + ۰/۶ X_{۲i}) + U_i .$$

باید مجموع مربعات پسماند برای این مدل مقید را حساب کنیم. ابتدا مدل مقید را تخمین می‌زنیم،

$$\hat{\beta}_۱^* = \frac{\Sigma x_i^* y_i}{\Sigma x_i^{*۲}},$$

که در آن علامت * بیان کنندهٔ این نکته است که محاسبات برای مدل مقید انجام شده است. باید صورت و مخرج کسر فوق را محاسبه کنیم. می‌دانیم

$$x_i^* = x_{1i} + ۰/۶\, x_{2i} .$$

دو طرف رابطهٔ فوق را مجذور کرده و برای تمام مشاهدات i جمع می‌کنیم،

$$\Sigma x_i^{*2} = \Sigma x_{1i}^2 + (۰/۶)^2 \Sigma x_{2i}^2 + ۲\,(۰/۶) \Sigma x_{1i} x_{2i} ,$$
$$= ۱۲ + ۰/۳۶\,(۱۲) + ۱/۲\,(۸) = ۲۵/۹۲ .$$

به همین ترتیب خواهیم داشت

$$\Sigma x_i^* y_i = \Sigma x_{1i}\, y_i + ۰/۶ \Sigma x_{2i}\, y_i ,$$
$$= ۱۰ + ۰/۶\,(۸) = ۱۴/۸ .$$

در نتیجه، β_1^* به دست می‌آید،

$$\beta_1^* = \frac{۱۴/۸}{۲۵/۹۲} = ۰/۵۷ .$$

مجموع مربعات پسماند برای مدل مقید، یعنی RSS_r را حساب می‌کنیم،

$$RSS_r = \Sigma y_i^2 - \hat{\beta}_1^* \Sigma x_i^*\, y_i ,$$
$$= ۱۰ - ۰/۵۷\,(۱۴/۸) = ۱/۵۶۴ .$$

حال باید مجموع مربعات پسماند را برای مدل اصلی، یعنی RSS، نیز به دست آوریم. می‌دانیم

$$RSS = \Sigma y_i^2 - \hat{\beta}_1 \Sigma x_{1i}\, y_i - \hat{\beta}_2 \Sigma x_{2i}\, y_i ,$$
$$= ۱۰ - ۰/۷\,(۱۰) - ۰/۲\,(۸) = ۱/۴ .$$

اگر H_0 صحیح باشد، آمارهٔ زیر توزیع F خواهد داشت

$$F = \frac{(RSS_r - RSS)/۱}{RSS/(۲۳-۳)} \sim F(۱, ۲۰) ,$$

$$= \frac{(۱/۵۶۴ - ۱/۴)/۱}{۱/۴/۲۰} = \frac{۰/۱۶۴}{۰/۰۷} = ۲/۳۴۲ .$$

مقدار جدول F در سطح معنی‌دار ۵ درصد برابر است با

$$F_{\%5}(1,20) = 4/35 ,$$

در نتیجه آمارهٔ آزمون معنی‌دار نبوده و فرضیهٔ H_0، رد نمی‌شود.

راه‌حل دیگر، استفاده از آزمون t است. فرضیهٔ H_0 را به صورت زیر می‌نویسیم،

$$H_0: \beta_2 - 0/6\,\beta_1 = 0 .$$

باید $\hat{\beta}_2 - 0/6\,\hat{\beta}_1$ را استاندارد کنیم. ابتدا محاسبات زیر را انجام می‌دهیم،

$$E(\hat{\beta}_2 - 0/6\,\hat{\beta}_1) = \beta_2 - 0/6\,\beta_1 ,$$

$$Var(\hat{\beta}_2 - 0/6\,\hat{\beta}_1) = Var(\hat{\beta}_2) + (0/6)^2\,Var(\hat{\beta}_1) - 2(0/6)\,Cov(\hat{\beta}_1, \hat{\beta}_2) ,$$

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{(1 - r_{12}^2)\sum x_{1i}^2} \quad , \qquad r_{12}^2 = \frac{(\sum x_{1i} \sum x_{2i})^2}{\sum x_{1i}^2 \sum x_{2i}^2} ,$$

$$r_{12}^2 = \frac{(8)^2}{12(12)} = \frac{64}{144} = \frac{4}{9} ,$$

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{(1 - \frac{4}{9})\,12} = \frac{3}{20}\sigma^2 \quad , \qquad Var(\hat{\beta}_1) = \frac{3}{20}\sigma^2 ,$$

$$Cov(\hat{\beta}_1, \hat{\beta}_2) = \frac{-\sigma^2 r_{12}^2}{(1 - r_{12}^2)\sum x_{1i} x_{2i}} ,$$

$$= \frac{-\sigma^2(4)(9)}{9(5)8} = -\frac{1}{10}\sigma^2 .$$

بدین ترتیب خواهیم داشت

$$Var(\hat{\beta}_2 - 0/6\,\hat{\beta}_1) = \frac{3}{20}\sigma^2 + (0/6)^2\frac{3}{20}\sigma^2 + 2(0/6)\frac{1}{10}\sigma^2 = 0/324\,\sigma^2 ,$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{(n-k)} = \frac{1/4}{20} = 0/07 .$$

با اطلاعات فوق می‌توان آمارهٔ t را به شرح زیر محاسبه کرد،

$$t=\frac{(\hat{\beta}_2-۰/۶\hat{\beta}_1)-E(\hat{\beta}_2-۰/۶\hat{\beta}_1)}{\sqrt{\text{Var}(\hat{\beta}_2-۰/۶\hat{\beta}_1)}}=\frac{[۰/۲-۰/۶(۰/۷)]-۰}{\sqrt{۰/۳۲۴(۰/۰۷)}},$$

$$=\frac{-۰/۲۲}{۰/۱۵۰۵}=-۱/۴۶۱.$$

مقدار موجود در جدول t با ۲۰ =۳- n درجهٔ آزادی و در سطح معنی‌دار ۲/۵ درصد برابر است با ۲/۰۸۶ ± و در نتیجه آمارهٔ t معنی‌دار نبوده و فرضیهٔ H_0 رد نمی‌شود. انتظار چنین نتیجه‌ای را نیز داشتیم زیرا

$$t^2=\frac{(-۰/۲۲)^2}{۰/۳۲۴(۰/۰۷)}=۲/۱۳۴,$$

که تقریباً برابر آمارهٔ آزمون F است. اختلاف ناشی از تقریب در محاسبات است.

۶.۵ الف) ماتریس $(\mathbf{X}'\mathbf{X})$ و بردار $\mathbf{X}'\mathbf{y}$ را تشکیل می‌دهیم،

$$(\mathbf{X}'\mathbf{X})=\begin{bmatrix}۳۰ & ۰\\ ۰ & ۳\end{bmatrix},\qquad \mathbf{X}'\mathbf{y}=\begin{bmatrix}۳۰\\ ۲۰\end{bmatrix}.$$

درنتیجه خواهیم داشت

$$\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{bmatrix}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}=\begin{bmatrix}\frac{۱}{۳۰} & ۰\\ ۰ & \frac{۱}{۳}\end{bmatrix}\begin{bmatrix}۳۰\\ ۲۰\end{bmatrix}=\begin{bmatrix}۱\\ \frac{۲۰}{۳}\end{bmatrix},$$

بنابراین ۱ =$\hat{\beta}_1$ و $\hat{\beta}_2=\frac{۲۰}{۳}$. همچنین داریم

$$R^2=\frac{\hat{\beta}'\mathbf{X}'\mathbf{y}}{\mathbf{y}'\mathbf{y}}=\frac{\hat{\beta}_1\Sigma x_{1t}y_t+\hat{\beta}_2\Sigma x_{1t}y_t}{\Sigma y_t^2},$$

$$=\frac{۱(۳۰)+\frac{۲۰}{۳}(۲۰)}{\frac{۴۹۳}{۳}}=\frac{۴۹۰}{۴۹۳}=۰/۹۹.$$

ب) ابتدا واریانس U_i را به شرح زیر حساب می‌کنیم،

$$RSS = (۱ - R^۲)\, TSS\ ,$$

$$= (۱ - \frac{۴۹۰}{۴۹۳})\ (\frac{۴۹۳}{۳}) = ۱\ ,$$

$$\hat{\sigma}_U^۲ = \frac{RSS}{n-k} = \frac{۱}{۱۰۰ - ۳} = \frac{۱}{۹۷}\ .$$

حال آمارهٔ t را به دست می‌آوریم،

$$t = \frac{\hat{\beta}_۲ - \beta_۲}{\hat{\sigma}\sqrt{a_{۲۲}}} = \frac{\frac{۲۰}{۳} - ۷}{\sqrt{\frac{۱}{۹۷}}\sqrt{\frac{۱}{۳}}} = -۵/۶۹\ .$$

یادآوری می‌شود که $a_{۲۲}$ به معنی دومین عنصر قطری ماتریس $(X'X)^{-۱}$ است. مقدار t از جدول t برابر است با ۱/۹۸ ± ؛ بنابراین آمارهٔ آزمون در ناحیهٔ بحرانی قرار گرفته و معنی‌دار است. درنتیجه فرضیهٔ $H_۰$ رد می‌شود.

ج) این فرضیه به آزمون معنی‌دار بودن کل رگرسیون مربوط است. از فرمول F برحسب $R^۲$ در معادلهٔ ۶.۴۱ استفاده می‌کنیم،

$$F = (\frac{n-k}{k-۱})\ \frac{R^۲}{۱ - R^۲} \sim F(k-۱, n-k)\ ,$$

$$F = (\frac{۱۰۰ - ۳}{۲})\ \frac{۴۹۰}{۴۹۳(۱ - ۴۹۰/۴۹۳)} = \frac{۹۷}{۲} \cdot \frac{۴۹۰\,(۴۹۳)}{۴۹۳\,(۳)} = ۷۹۲۱/۶\ .$$

ملاحظه می‌شود که آمارهٔ آزمون معنی‌دار است و فرضیهٔ $H_۰$ رد می‌شود.

د) آزمون فرضیهٔ $H_۰: \beta_۲ = ۷\beta_۱$ از راه ماتریسی می‌تواند به دو صورت انجام شود. استفاده از فرمول ۶.۵۰ با استاندارد کردن $\mathbf{C}'\hat{\boldsymbol{\beta}}$ یا نوشتن $H_۰$ به صورت $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ و استفاده از آمارهٔ عمومی آزمون، یعنی فرمول ۶.۶۴ . از هر دو راه این مسأله را حل می‌کنیم تا رابطهٔ بین این دو فرمول نیز روشن شود.

راه حل اول: از معادلهٔ ۶.۵۰ می‌دانیم که

$$t=\frac{\mathbf{c}'\hat{\boldsymbol{\beta}}-\mathbf{r}}{\sigma\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}}\sim t(n-k).$$

H_0 را به صورت $\beta_2-7\beta_1=0$ نوشته، داریم

$$[-7 \quad 1]\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}=0,$$

در نتیجه $\mathbf{c}'\hat{\boldsymbol{\beta}}$ برابر است با

$$\mathbf{c}'\hat{\boldsymbol{\beta}}=[-7 \quad 1]\begin{bmatrix}1\\ \frac{20}{3}\end{bmatrix}=-\frac{1}{3}.$$

صورت کسر t برابر خواهد بود با

$$\mathbf{c}'\hat{\boldsymbol{\beta}}-\mathbf{r}=-\frac{1}{3}-0=-\frac{1}{3}.$$

برای سهولت محاسبات، ابتدا مجذور مخرج کسر t را حساب می‌کنیم،

$$\sigma^2\,\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}=\frac{1}{97}[-7 \quad 1]\begin{bmatrix}\frac{1}{30} & 0\\ 0 & \frac{1}{3}\end{bmatrix}\begin{bmatrix}-7\\ 1\end{bmatrix}=\frac{1}{97}\left(\frac{59}{30}\right)=0/020274.$$

آمارهٔ t به شرح زیر محاسبه می‌شود،

$$t=\frac{-1}{3\sqrt{0/020274}}=-2/34091,$$

که معنی‌دار بوده و فرضیهٔ H_0 رد می‌شود.

راه حل دوم: از معادلهٔ ۶.۶۴ می‌دانیم که

$$F=\frac{\{(\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r})\}/r}{\mathbf{e}'\mathbf{e}/(n-k)}\sim F(r,n-k),$$

فرمول فوق می‌تواند به طور همزمان چند ترکیب خطی را با هم آزمون کند. با توجه

به اینکه در این مسأله، فقط یک ترکیب خطی را می‌خواهیم آزمون کنیم، بنابراین ماتریس $\mathbf{R}$ به همان بردار $\mathbf{C}'$ تبدیل خواهد شد و در نتیجه، $(\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r})$ دقیقاً همان $(\mathbf{C}'\hat{\boldsymbol{\beta}}-\mathbf{r})$ می‌شود. یعنی

$$\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r}=-\frac{۱}{۳}.$$

به ترتیبی مشابه خواهیم داشت

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-۱}\mathbf{R}'=\mathbf{C}'(\mathbf{X}'\mathbf{X})^{-۱}\mathbf{C}=\frac{۵۹}{۳۰}.$$

بنابراین، صورت کسر F برابر است با

$$[(-\frac{۱}{۳})(\frac{۵۹}{۳۰})^{-۱}(-\frac{۱}{۳})]/۱,$$

و مخرج کسر F نیز عبارت خواهد بود از

$$\frac{\mathbf{e}'\mathbf{e}}{\mathbf{n}-\mathbf{k}}=\hat{\sigma}_{ü}^{۲}=\frac{۱}{۹۷}.$$

در نتیجه آمارهٔ F برابر خواهد بود با

$$F=\frac{(\frac{۱}{۳})^{۲}/۱}{\frac{۵۹}{۳۰}(\frac{۱}{۹۷})}=۵/۴۸.$$

ملاحظه می‌شود $\sqrt{۵/۴۸}=۲/۳۴۰۹$، یعنی $\sqrt{F}=t$.

۶-۶ این مسأله اهمیت فراوان دارد؛ زیرا تعداد مشاهدات و واریانس جملهٔ اختلال داده نشده است. آمارهٔ عمومی آزمون، یعنی معادلهٔ ۶-۶۴، را می‌نویسیم. می‌دانیم مخرج کسر، یعنی $\mathbf{e}'\mathbf{e}/(\mathbf{n}-\mathbf{k})$، درواقع همان تخمین واریانس جملهٔ اختلال، یعنی $\hat{\sigma}_{ü}^{۲}$، است.

$$F=\frac{\{(\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-۱}\mathbf{R}']^{-۱}(\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r})'\}/\mathbf{r}}{\hat{\sigma}^{۲}}.$$

استفاده از فرمول فوق مستلزم داشتن مقادیر جداگانه برای $\hat{\sigma}^{۲}$ و $[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-۱}\mathbf{R}']^{-۱}$

است که متأسفانه در این مسأله موجود نیست؛ با وجود این، اگر آمارهٔ F را به صورت زیر بنویسیم،

$$F = \{(R\hat{\beta} - r)' [R\hat{\sigma}^2 (X'X)^{-1} R']^{-1} (R\hat{\beta} - r)'\} / r , \quad (۱)$$

می‌توان، $[R\hat{\sigma}^2 (X'X)^{-1} R']^{-1}$ را به راحتی محاسبه کرد. برای اثبات، ابتدا براساس معادلهٔ ۶.۷، معادلهٔ واریانس ـ کوواریانس $\hat{\beta}$ را می‌نویسیم،

$$\Sigma_{\hat{\beta}} = \hat{\sigma}^2 (X'X)^{-1} .$$

ماتریس فوق را از سمت چپ و راست به ترتیب در R و R' ضرب می‌کنیم،

$$R\hat{\sigma}^2 (X'X)^{-1} R' . \quad (۲)$$

با معکوس کردن ماتریس (۲)، جملهٔ مجهول آمارهٔ F در (۱) به دست می‌آید. بدیهی است مقدار $(R\hat{\beta} - r)$ کاملاً معلوم و قابل محاسبه است. محاسبهٔ آمارهٔ F، مستلزم داشتن $\hat{\sigma}^2$ و $(X'X)^{-1}$ به طور جداگانه نیست. نتیجهٔ کلی این است که اگر ماتریس واریانس ـ کوواریانس $\hat{\beta}$ را داشته باشیم کافی است با استفاده از R، ماتریس (۲) را ساخته و آن را معکوس کنیم تا آمارهٔ F در رابطهٔ (۱) قابل محاسبه شود.

با استفاده از صورت مسأله می‌توان ماتریس واریانس ـ کوواریانس $\hat{\beta}$ را به دست آورد.

$$\Sigma_{\hat{\beta}} = \begin{bmatrix} \text{Var}(\hat{\beta}_K) & \text{Cov}(\hat{\beta}_K, \hat{\beta}_L) \\ \text{Cov}(\hat{\beta}_L, \hat{\beta}_K) & \text{Var}(\hat{\beta}_L) \end{bmatrix} ,$$

$$= \hat{\sigma}^2 (X'X)^{-1} = \begin{bmatrix} (۰/۲۵۷)^2 & ۰/۰۵۵ \\ ۰/۰۵۵ & (۰/۲۱۹)^2 \end{bmatrix} = \begin{bmatrix} ۰/۰۶۶ & ۰/۰۵۵ \\ ۰/۰۵۵ & ۰/۰۴۸ \end{bmatrix} ,$$

الف) با توجه به اینکه در مدل مفروض، متغیرها برحسب لگاریتم نوشته شده است، کشش تولید نسبت به K و L به ترتیب برابر است با β_K و β_L. بنابراین فرضیه

تساوی دو کشش عبارت است از فرضیهٔ $H_0: \beta_K = \beta_L$. از آمارهٔ عمومی آزمون استفاده کرده، $R\beta = r$ را می‌سازیم،

$$R = [1 \quad -1] \ , \quad r = 0 \ ,$$

درنتیجه خواهیم داشت

$$R\hat{\beta} - r = [1 \quad -1] \begin{bmatrix} ۰/۶۳۲ \\ ۰/۴۵۲ \end{bmatrix} - 0 = ۰/۱۸ \ .$$

ماتریس (۲) را تشکیل می‌دهیم،

$$R\,\hat{\sigma}^2\,(X'X)^{-1}\,R' = [1 \quad -1] \begin{bmatrix} ۰/۰۶۶ & ۰/۰۵۵ \\ ۰/۰۵۵ & ۰/۰۴۸ \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = ۰/۰۰۴ \ .$$

آمارهٔ F از رابطهٔ (۱) به سهولت قابل محاسبه است،

$$F = \{(۰/۱۸)[۰/۰۰۴]^{-1}(۰/۱۸)\} / 1 = \frac{(۰/۱۸)^2}{۰/۰۰۴} = ۸/۱ \ .$$

در این قسمت باید آمارهٔ آزمون را با مقدار F از جدول مقایسه کنیم که خود مستلزم داشتن تعداد مشاهدات است که در این مسأله موجود نیست؛ البته بدون داشتن تعداد مشاهدات نیز می‌توان فرضیهٔ مورد نظر را در این مسألهٔ خاص آزمون کرد. می‌دانیم مقدار $R^2 = ۰/۹۸$ و نیز مقادیر نسبتاً کوچک واریانس تخمین پارامترها می‌توانند درواقع دلالت بر وجود حجم قابل قبولی از مشاهدات داشته باشد. با توجه به اینکه در این مدل سه پارامتر تخمین زده می‌شود انتظار این است که حجم مشاهدات از ۱۰ بیشتر باشد. حتی اگر حجم مشاهدات را ۷ بگیریم، درجات آزادی مخرج کسر F در معادلهٔ ۶۴ـ۶ برابر است با $n - k = ۷ - ۳ = ۴$. مقدار F به دست آمده از جدول را برای کمترین و بیشترین مقادیر ممکن درجات آزادی مخرج به دست می‌آوریم،

$$F_{۵٪}(1, 4) = ۷/۷۱ \quad , \quad F_{۵٪}(1, \infty) = ۳/۸۴ \ .$$

در هر دو حالت مقدار آمارهٔ آزمون بیشتر از مقدار جدول F می‌شود و فرضیهٔ H_0، یعنی برابری کششها رد می‌شود.

ب) فرضیهٔ بازده ثابت نسبت به مقیاس در این مسأله را می‌توان به صورت زیر نوشت،

$$\beta_K + \beta_L = ۱ .$$

دوباره از آمارهٔ عمومی آزمون، یعنی معادلهٔ (۱) استفاده می‌کنیم،

$$R = [۱ \quad ۱] , \quad r = ۱ ,$$

$$(R\hat{\beta} - r) = ۰/۰۸٤ ,$$

$$R\,\hat{\sigma}^{۲}\,(X'X)^{-۱}\,R' = [۱ \quad ۱] \begin{bmatrix} ۰/۰٦٦ & ۰/۰۵۵ \\ ۰/۰۵۵ & ۰/۰٤۸ \end{bmatrix} \begin{bmatrix} ۱ \\ ۱ \end{bmatrix} = ۰/۲۲٤ .$$

بدین ترتیب داریم

$$F = \{(۰/۰۸٤)[۰/۲۲٤]^{-۱}(۰/۰۸٤)\} / ۱ = \frac{(۰/۰۸٤)^{۲}}{۰/۲۲٤} = ۰/۰۳۱۵ .$$

برای هر تعدادی از مشاهدات که بتوان تصور کرد، مقدار آمارهٔ آزمون معنی‌دار نیست؛ بنابراین فرضیهٔ بازده ثابت به مقیاس را نمی‌توان رد کرد.

۶.۷ ابتدا پارامترهای مدل را تخمین می‌زنیم. برای این منظور باید محاسبات زیر را انجام داد،

$$(X'X) = \begin{bmatrix} ۱۰ & ۱۰ & ۵ \\ ۱۰ & ۳۰ & ۱۵ \\ ۵ & ۱۵ & ۲۰ \end{bmatrix} \qquad X'y = \begin{bmatrix} ۷ \\ -۷ \\ -۲۶ \end{bmatrix} ,$$

$$(X'X)^{-۱} = \begin{bmatrix} ۰/۱۵ & -۰/۰۵ & ۰ \\ -۰/۰۵ & ۰/۰۷ & -۰/۰٤ \\ ۰ & -۰/۰٤ & ۰/۰۸ \end{bmatrix} .$$

در نتیجه بردار $\hat{\beta}$ عبارت است از

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} ۰/۱۵ & -۰/۰۵ & ۰ \\ -۰/۰۵ & ۰/۰۷ & -۰/۰۴ \\ ۰ & -۰/۰۴ & ۰/۰۸ \end{bmatrix} \begin{bmatrix} ۷ \\ -۷ \\ -۲۶ \end{bmatrix} = \begin{bmatrix} ۱/۴ \\ ۰/۲ \\ -۱/۸ \end{bmatrix}$$

برای تخمین واریانس جملهٔ اختلال، ابتدا ESS را حساب می‌کنیم،

$$ESS = \hat{\beta}' \mathbf{X}'\mathbf{y} = [۱/۴ \quad ۰/۲ \quad -۱/۸] \begin{bmatrix} ۷ \\ -۷ \\ -۲۶ \end{bmatrix} = ۵۵/۲ ,$$

$$RSS = \mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \hat{\beta}'\mathbf{X}'\mathbf{y} ,$$

$$= ۶۰ - ۵۵/۲ = ۴/۸ ,$$

و بدین ترتیب $\hat{\sigma}^2$ عبارت است از

$$\hat{\sigma}^2 = \frac{\mathbf{e}'\mathbf{e}}{n-k} = \frac{۴/۸}{۲۴-۴} = ۰/۲۴ \quad , \quad \hat{\sigma}^2 = ۰/۴۹۰ .$$

الف) از آزمون t استفاده می‌کنیم،

$$t = \frac{\hat{\beta}_i - a}{SE(\hat{\beta}_i)} .$$

برای آزمون $\beta_1 = ۱$، داریم

$$t = \frac{۱/۴ - ۱}{۰/۴۹۰\sqrt{۰/۱۵}} = ۲/۱۱ .$$

با توجه به اینکه مقدار t از جدول با ۲۰ درجهٔ آزادی برابر است با ۲/۰۸۶ ؛ بنابراین آمارهٔ آزمون معنی‌دار است و فرضیهٔ $\beta_1 = ۱$ رد می‌شود.

برای آزمون $\beta_2 = ۱$، داریم

$$t = \frac{۰/۲ - ۱}{۰/۴۹۰\sqrt{۰/۰۷}} = -۶/۱۷ ,$$

بنابراین آمارهٔ آزمون معنی‌دار است و فرضیهٔ $\beta_۲ = ۱$ رد می‌شود.

برای آزمون $\beta_۳ = -۲$، داریم

$$t = \frac{-۱/۸ - (-۲)}{۰/۴۹۰\sqrt{۰/۰۸}} = ۱/۴۴ ,$$

در نتیجه آمارهٔ آزمون معنی‌دار نبوده و فرضیهٔ $\beta_۳ = -۲$ رد نمی‌شود.

ب) برای آزمون $H_۰: \beta_۱ + \beta_۲ + \beta_۳ = ۰$، از آمارهٔ عمومی آزمون، یعنی معادلهٔ ۶-۶۴ استفاده می‌کنیم،

$$F = \frac{\{(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})' [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-۱}\mathbf{R}']^{-۱}(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})\} / r}{\mathbf{e}'\mathbf{e} / (n-k)} .$$

می‌دانیم

$$\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r} = \begin{bmatrix} ۱ & ۱ & ۱ \end{bmatrix} \begin{bmatrix} ۱/۴ \\ ۰/۲ \\ -۱/۸ \end{bmatrix} = -۰/۲ ,$$

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-۱}\mathbf{R}' = \begin{bmatrix} ۱ & ۱ & ۱ \end{bmatrix} \begin{bmatrix} ۰/۱۵ & -۰/۰۵ & ۰ \\ -۰/۰۵ & ۰/۰۷ & -۰/۰۴ \\ ۰ & -۰/۰۴ & ۰/۰۸ \end{bmatrix} \begin{bmatrix} ۱ \\ ۱ \\ ۱ \end{bmatrix} = ۰/۱۲ ,$$

بنابراین آمارهٔ آزمون به شرح زیر محاسبه می‌شود،

$$F = \frac{\{(-۰/۲)\,[۰/۱۲]^{-۱}(-۰/۲)\} / ۱}{۴/۸ / (۲۴ - ۴)} = \frac{۰/۰۴}{۰/۰۲۸۸} = ۱/۳۸۹ .$$

مقدار F از جدول برابر است با

$$F_{۰/۰۵}(۱, ۲۰) = ۴/۳۵ ,$$

بنابراین آمارهٔ آزمون معنی‌دار نبوده و فرضیهٔ $H_۰$، یعنی $\beta_۱ + \beta_۲ + \beta_۳ = ۰$ رد نمی‌شود.

ج) برای آزمون همزمان

$$H_0: [\beta_1 \quad \beta_2 \quad \beta_3] \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix},$$

از آمارهٔ عمومی آزمون استفاده می‌کنیم. داریم

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{r} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}.$$

بنابراین، خواهیم داشت

$$\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/4 \\ 0/2 \\ -1/8 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0/4 \\ -0/8 \\ 0/2 \end{bmatrix}.$$

همچنین می‌دانیم که

$$[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} = \mathbf{X}'\mathbf{X},$$

زیرا $\mathbf{R}$ یک ماتریس واحد است. بدین ترتیب، آمارهٔ عمومی آزمون از معادلهٔ ۶.۶۴ به شرح زیر محاسبه می‌شود،

$$F = \frac{\{(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})'[\mathbf{X}'\mathbf{X}]^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})'\} / (k-1)}{\mathbf{e}'\mathbf{e} / (n-k)}.$$

ابتدا صورت کسر F را حساب می‌کنیم.

$$\left\{ (0/4 \quad 0/8 \quad 0/2) \begin{bmatrix} 0/15 & -0/05 & 0 \\ -0/05 & 0/07 & -0/04 \\ 0 & -0/04 & 0/08 \end{bmatrix} \begin{bmatrix} 0/4 \\ -0/8 \\ 0/2 \end{bmatrix} \right\} / 3 = \frac{11/2}{3},$$

مقدار آمارهٔ عمومی آزمون عبارت خواهد بود از

$$F = \frac{11/2/3}{4/8/20} = 15/56.$$

با توجه به مقدار $F_{\%۵}(۳,۲۰)=۳/۱۰$، نتیجه می‌گیریم که آمارهٔ آزمون معنی‌دار است و فرضیهٔ $H_۰$، یعنی $[\beta_۱\ \beta_۲\ \beta_۳]=[۱\ \ ۱\ \ -۲]$، رد می‌شود.

۶ـ۸ مدل اول حالت مقید مدل دوم است؛ به عبارت دیگر اگر در مدل دوم بخواهیم فرضیهٔ تساوی ضریب متغیر W_t برابر صفر را آزمون کنیم و سپس این قید را در مدل دوم جایگزین کنیم، مدل اول به‌دست می‌آید. در مدل دوم، مقدار آمارهٔ t متعلق به W_t برابر ۳ محاسبه شده است. با توجه به اینکه $F=t^۲$، بنابراین مقدار F متعلق به همین آزمون برابر ۹ خواهد بود. اما می‌دانیم آمارهٔ F برابر است با

$$F=\frac{(RSS_r-RSS)/r}{RSS/(n-k)}.$$

از طرف دیگر، $r=۱$ و $(n-k)=n-۳$؛ بنابراین

$$F=\frac{(RSS_r-RSS)}{RSS/(n-۳)}=(n-۳)\left(\frac{RSS_r}{RSS}-۱\right).$$

مقادیر کل تغییرات برای هر دو مدل متساوی است. می‌دانیم

$$RSS_r=(۱-R_r^۲)\,TSS=(۱-۰/۲)\,TSS=۰/۸\,TSS,$$

$$RSS=(۱-R^۲)\,TSS=(۱-۰/۶)\,TSS=۰/۴\,TSS,$$

بنابراین

$$F=(n-۳)\left(\frac{۰/۸\,TSS}{۰/۴\,TSS}-۱\right).$$

همچنین با توجه به اینکه $F=۹$، پس

$$۹=(n-۳)(۲-۱)\quad,\quad n=۱۲.$$

۶ـ۹ برای اثبات، باید نشان دهیم که

$$\hat{\beta}_r'(X_r'M_s X_r)\hat{\beta}_r=RSS_r-RSS=e_s'e_s-e'e.$$

معادلهٔ ۶.۶۷ را در نظر گرفته و دو طرف آن را از سمت چپ در $\mathbf{M}_2$ ضرب می‌کنیم،

$$\mathbf{M}_2 \mathbf{y} = \mathbf{M}_2 \mathbf{X}_1 \hat{\boldsymbol{\beta}}_1 + \mathbf{M}_2 \mathbf{X}_2 \hat{\boldsymbol{\beta}}_2 + \mathbf{M}_2 \mathbf{e} . \qquad (۱)$$

با توجه به تعریف ماتریس $\mathbf{M}_2$، یعنی

$$\mathbf{M}_2 = \mathbf{I} - \mathbf{X}_2 (\mathbf{X}_2' \mathbf{X}_2)^{-1} \mathbf{X}_2' , \qquad (۲)$$

می‌توان نشان داد که $\mathbf{M}_2 \mathbf{X}_2 = \mathbf{0}$؛ زیرا

$$\mathbf{M}_2 \mathbf{X}_2 = \mathbf{I} \mathbf{X}_2 - \mathbf{X}_2 (\mathbf{X}_2' \mathbf{X}_2)^{-1} (\mathbf{X}_2' \mathbf{X}_2) = \mathbf{0} \qquad (۳)$$

همچنین $\mathbf{M}_2 \mathbf{e} = \mathbf{e}$؛ زیرا

$$\mathbf{M}_2 \mathbf{e} = \mathbf{I} \mathbf{e} - \mathbf{X}_2 (\mathbf{X}_2' \mathbf{X}_2)^{-1} \mathbf{X}_2' \mathbf{e} . \qquad (۴)$$

اما با توجه به تعریف ماتریس $\mathbf{X}$ در معادلهٔ ۶.۶۷، داریم

$$\mathbf{X}' \mathbf{e} = [\mathbf{X}_1' \quad \mathbf{X}_2'] \mathbf{e} = [\mathbf{X}_1' \mathbf{e} \quad \mathbf{X}_2' \mathbf{e}] = [\mathbf{0} \quad \mathbf{0}] ,$$

زیرا می‌دانیم جمله‌های پسماند از متغیرهای توضیحی مستقل است. با جایگزینی معادله‌های (۳) و (۴) در معادلهٔ (۱)، خواهیم داشت

$$\mathbf{M}_2 \mathbf{y} = \mathbf{M}_2 \mathbf{X}_2 \hat{\boldsymbol{\beta}}_2 + \mathbf{e} .$$

معادلهٔ فوق را ترانهاد کرده، مجذور می‌کنیم، در نتیجه

$$\mathbf{y}' \mathbf{M}_2 \mathbf{y} = \hat{\boldsymbol{\beta}}_2' \mathbf{X}_2' \mathbf{M}_2 \mathbf{X}_2 \hat{\boldsymbol{\beta}}_2 + \mathbf{e}' \mathbf{e} . \qquad (۵)$$

نشان می‌دهیم که $\mathbf{y}' \mathbf{M}_2 \mathbf{y}$ برابر با $\mathbf{e}' \mathbf{e}$ است. برای این منظور دو طرف معادلهٔ (۲) را از سمت راست در $\mathbf{y}$ و از سمت چپ در $\mathbf{y}'$ ضرب می‌کنیم،

$$\begin{aligned} \mathbf{y}' \mathbf{M}_2 \mathbf{y} &= \mathbf{y}' \mathbf{y} - \mathbf{y}' \mathbf{X}_2 (\mathbf{X}_2' \mathbf{X}_2)^{-1} \mathbf{X}_2' \mathbf{y} . \\ &= \mathbf{y}' \mathbf{y} - \mathbf{y}' \mathbf{X}_2 \hat{\boldsymbol{\beta}}_2 , \\ &= \mathbf{y}' \mathbf{y} - \mathbf{y}' \hat{\mathbf{y}}_2 . \qquad (۶) \end{aligned}$$

اما با توجه به معادلهٔ ۵ـ۵۲ می‌دانیم که

$$= \mathbf{y}'\hat{\mathbf{y}}_s = \hat{\mathbf{y}}_s'\hat{\mathbf{y}}_s ,$$

و با جایگزینی در معادلهٔ (۶)، داریم

$$\mathbf{y}'\mathbf{M}_s\mathbf{y} = \mathbf{y}'\mathbf{y} - \hat{\mathbf{y}}_s'\hat{\mathbf{y}}_s = \mathbf{e}_s'\mathbf{e}_s .$$

با جایگزینی معادلهٔ فوق در معادلهٔ (۵)، خواهیم داشت

$$\hat{\boldsymbol{\beta}}_r'\mathbf{X}_r'\mathbf{M}_s\mathbf{X}_r\hat{\boldsymbol{\beta}}_r = \mathbf{e}_s'\mathbf{e}_s - \mathbf{e}'\mathbf{e} ,$$

$$= RSS_r - RSS .$$

رابطهٔ فوق را در آمارهٔ F جایگزین می‌کنیم،

$$F = \frac{(RSS_r - RSS)/r}{RSS/(n-k)} \sim F(r, n-k) .$$

۶ـ۱۰ می‌دانیم در معادلهٔ ۶ـ۷۲، ماتریس $\mathbf{R}$ به صورت زیر تعریف شده است،

$$\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} ,$$

که در آن $\mathbf{R} \to (k-1) \times k$. به کمک این ماتریس باید نشان دهیم که

$$\hat{\boldsymbol{\beta}}_r'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\hat{\boldsymbol{\beta}}_r = \hat{\mathbf{y}}'\hat{\mathbf{y}} = ESS .$$

ملاحظه می‌شود که $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$، برابر با آخرین $(k-1)$ سطر و ستون ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ است. برای اینکه ساختار این ماتریس را بررسی کنیم ابتدا ماتریس $\mathbf{X}$ به صورت زیر افراز می‌شود،

$$\mathbf{X} = [\mathbf{i} \quad \mathbf{X}_r] ,$$

که در آن $\mathbf{i}$ یک بردار ستونی شامل n عدد یک است و $\mathbf{X}_2$ یک ماتریس $n \times (k-1)$ از مشاهدات تمام متغیرهای توضیحی است. ماتریس $(\mathbf{X}'\mathbf{X})$ را تشکیل می‌دهیم،

$$(\mathbf{X}'\mathbf{X}) = \begin{bmatrix} n & \mathbf{i}'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{i} & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix} .$$

با استفاده از قاعدهٔ معکوس کردن ماتریسهای افراز شده[1]، می‌دانیم که اگر $(\mathbf{X}'\mathbf{X})$ را معکوس کنیم، زیر ماتریس تشکیل شده از آخرین $(k-1)$ سطر و ستون آن عبارت خواهد بود از

$$(\mathbf{X}_2'\mathbf{X}_2 - \mathbf{X}_2'\mathbf{i}\frac{1}{n}\mathbf{i}'\mathbf{X}_2)^{-1} .$$

پیش از این نیز گفتیم که ماتریس $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ برابر با آخرین $(k-1)$ سطر و ستون ماتریس $(\mathbf{X}'\mathbf{X})^{-1}$ است؛ بنابراین

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' = (\mathbf{X}_2'\mathbf{X}_2 - \mathbf{X}_2'\mathbf{i}\frac{1}{n}\mathbf{i}'\mathbf{X}_2)^{-1} .$$

با توجه به ساختار بردار $\mathbf{i}$ و با مراجعه به تعریف ماتریس تبدیل $\mathbf{A}$ که در معادلهٔ ۵ـ۷۱ تعریف شده است، داریم

$$(\mathbf{X}_2'\mathbf{X}_2 - \mathbf{X}_2'\mathbf{i}\frac{1}{n}\mathbf{i}'\mathbf{X}_2)^{-1} = (\mathbf{X}_2'\mathbf{A}\mathbf{X}_2)^{-1} ,$$

در نتیجه

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' = (\mathbf{X}_2'\mathbf{A}\mathbf{X}_2)^{-1} , \qquad (۱)$$

که در آن عناصر ماتریس $\mathbf{X}_2$ برحسب مقادیر اصلی است. معادلهٔ (۱) را در آمارهٔ عمومی آزمون، یعنی معادلهٔ ۶ـ۷۲ جایگزین می‌کنیم،

$$F = \frac{\{\hat{\boldsymbol{\beta}}_2'(\mathbf{X}_2'\mathbf{A}\mathbf{X}_2)\hat{\boldsymbol{\beta}}_2\}/(k-1)}{\mathbf{e}'\mathbf{e}/(n-k)} \sim F(k-1, n-k) . \qquad (۲)$$

---

۱. به پیوست «۵ ـ الف» مراجعه شود.

با مراجعه به معادلهٔ ۵.۸۱ می‌دانیم

$$\hat{\beta}_r' X_r' \Lambda X_r \hat{\beta}_r = \hat{y}' \hat{y} = ESS ,$$

که در آن $\hat{y}$ برحسب انحراف از میانگین است؛ بنابراین با جایگزینی در معادلهٔ (۲) داریم

$$F = \frac{ESS / (k-1)}{RSS / (n-k)} \sim F(k-1, n-k) .$$