مدل رگرسیون کلاسیک

Classical regression Model

معادله رفتاری (Behavioral Equation)

رگرسیون به معنای تبیین رابطه بین متغیر وابسته و چند متغیر مستقل (توضیحی) است. در برازش یک رگرسیون هدف تخمین میزان تأثیر یک متغیر بر متغیر وابسته است. در حقیقت برازش یک معادله رگرسیونی به دنبال تبیین معادله رفتاری میان متغیرهای مستقل و وابسته است. معادلات رفتاری بیانگر نحوه واکنش متغیرهای مستقل و وابسته است. تفاوت کلی مدل‌های ریاضی و مدلهای رگرسیون در جمله اخلال است. هرگاه یک جمله اخلال – که یقیناً تصادفی است – به مدل‌های ریاضی اضافه شود به یک مدل رگرسیون تبدیل خواهد شد.

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \ldots\ldots\ldots + \beta_k X_{k,t} + \varepsilon_t \qquad (1)$$

به متغیر Y که در سمت چپ معادله بالا قرار دارد، متغیر وابسته و به xiها متغیرهای توضیحی یا مستقل گفته می‌شود. اصطلاحات متغیر برون‌زا و متغیر درون‌زا نیز به‌ترتیب برای xiها و Y به‌کار می‌رود، زیرا فرض بر این است که مقادیر xiها خارج از مدل مفروض تعیین‌شده و در نتیجه برون‌زا هستند، درحالی‌که مقادیر Y در داخل مدل و براساس قانونمندی تعیین می‌شود و به‌همین دلیل درون‌زا خواهد بود.

متغیر تصادفی (Random Variable)

متغیر تصادفی متغیری است که مقادیرش در نتیجهٔ آزمایش انتخاب می‌شود. درحقیقت، متغیر تصادفی تابعی از نتایج یک آزمایش است. بنابراین، هریک از مقادیر متغیر تصادفی بیانگر وقوع یک حادثه است که هر حادثه نیز پیامد یک آزمایش است و فضای نمونه را تشکیل می‌دهد. ازاین‌رو، بین عناصر فضای نمونه و مقادیر متغیر تصادفی یک تناظر وجود دارد و هریک از مقادیر متغیر تصادفی در تناظر با یک عضو در فضای نمونه است. هر متغیر تصادفی خود یک تابع است که دامنه آن فضای نمونه و حوزه آن مجموعه مقادیر متغیر تصادفی است. بنابراین، می‌توان گفت یک متغیر تصادفی متغیری است که تحت تأثیر عوامل متعددی قرار دارد و خصوصیاتی دارد که به‌واسطه توزیع آن و فروض مربوط به آن بیان می‌شود (سوری، ۱۳۹۴، ۳۴).

روش حداقل مربعات معمولی (Ordinary Least Square, OLS)

یکی از روش‌های مهم برای تخمین شاخص‌های مدل است که کاربرد زیادی در تخمین معادلات رگرسیونی دارد. طرح اولیه این روش را کارل فریدریش گوس (Carl Friedrich Gaus)، ریاضی‌دان معروف آلمانی، در قرن هجدهم میلادی مطرح کرده است. با این روش در حقیقت مجموع مربعات خطا حداقل شده و با این حداقل‌سازی ضرایب رابطه برازش می‌شوند. تخمین‌زننده‌های حداقل مربعات معمولی دارای ویژگی‌هایی مثل بدون تورش بودن، کارایی و سازگاری هستند و به‌اصطلاح آنها را تخمین‌زننده خطی ناریب یا BLUE می‌نامند (سوری، ۱۳۹۴، ۸۴). برای برازش یک معادله رگرسیونی با روش حداقل مربعات معمولی که تخمین‌زننده‌های خطی ناریب، کارا و سازگار نتیجه دهند باید فروض کلاسیک برقرار باشد. این فروض مشتمل بر خصوصیات پسماند مدل رگرسیونی است. تصادفی بودن مهم‌ترین نکته در مورد پسماند است. پسماند مدل یک متغیر تصادفی است و مثل همه متغیرهای تصادفی دارای یک تابع توزیع احتمال و درنتیجه میانگین و واریانس است. اولین فرض کلاسیک این است که میانگین یا امید ریاضی جمله اخلال صفر است. این فرض درواقع به‌این‌معنی است که میانگین تمام مقادیر ممکن به‌ازای هر مقدار معین از متغیرهای توضیح‌دهنده برابر ε_t صفر است. مفهوم کلی این فرض این است که مدل خطای سیستماتیک ندارد. ثابت بودن واریانس جمله اخلال به‌ازای مقادیر گوناگون متغیرهای مستقل دومین فرض کلاسیک است $\mathrm{Var}(\varepsilon_t) = \delta^2$. هرگاه واریانس جمله اخلال ثابت باشد

Encyclopedia of Economics
enceco.ihcs.ac.ir

مدل واریانس همسان و در غیر این صورت واریانس ناهمسان است. سومین فرض این است که ε_i و ε_j به‌ازای تمامی مقادیر $i \neq j$ از یکدیگر مستقل باشند. یعنی کوواریانس $(COV(\varepsilon_i, \varepsilon_j) = 0)$ آنها صفر است. به‌عبارت‌دیگر، هرگاه دو مقدار متفاوت برای متغیرهای مستقل در نظر گرفته شوند فرض بر این است که جمله‌های اخلال متناظر با آنها از یکدیگر مستقلند. در چنین حالتی بیان می‌شود که جمله‌های اخلال خودهمبستگی ندارند. چهارمین فرض این است که تابع توزیع جمله اخلال نرمال باشد؛ بنابراین، با توجه به فرض‌های اول و دوم و سوم می‌توان گفت که ε_t دارای توزیع نرمال با میانگین صفر و واریانس ثابت δ^2 است. پنجمین و آخرین فرض از فرض‌های کلاسیک این است که متغیرهای توضیح‌دهنده غیرتصادفی هستند. این فرض بیشتر برای سهولت در استنتاج قضایا و رسیدن به نتایج جالب‌تر در تخمین پارامترهاست. غیرتصادفی بودن متغیرهای توضیح‌دهنده بدین‌معناست که $X_{i,t}$ها از متغیر تصادفی ε_t مستقل هستند (درخشان، ۱۳۹۵، ۸۰).

تورش (Bias)

برای تخمین یک پارامتر می‌توان تخمین‌زننده‌های گوناگونی را به‌دست آورد، بنابراین، باید مناسب‌ترین تخمین‌زننده را انتخاب کرد. یکی از ویژگی‌های یک تخمین‌زننده خوب بدون تورش بودن آن است. اگر امید ریاضی یک تخمین‌زننده برابر با پارامتر موردنظر باشد بیان می‌شود که آن تخمین‌زننده بدون تورش یا ناریب است. ازاین‌رو، خاصیت ناریبی به‌صورت $E(\hat{\beta}) = \beta$ است. حال اگر تخمین‌زننده دارای تورش باشد در آن صورت $E(\hat{\beta}) - \beta = \varepsilon$ است که ε میزان تورش است. در حقیقت مفهوم تورش بدان‌معنا است که تخمین‌زننده به‌طور متوسط به پارامتر مورد نظر گرایش دارد (گجراتی، ۱۳۹۱، ۱۲۰).

کارایی (Efficiency)

کوچک بودن واریانس یکی دیگر از ویژگی‌های تخمین‌زننده‌ها است. هر تخمین‌زننده‌ای که واریانس کمتری داشته باشد کارایی بیشتری دارد (سوری، ۱۳۹۴، ۸۷).

سازگاری (Consistency)

سازگاری یکی از خواص مجانبی تخمین‌زننده‌ها است و بدین‌معنا است که اگر حجم نمونه افزایش یابد احتمال اینکه تقاضل $\hat{\beta}$ از β از مقدار ε کمتر باشد برابر یک خواهد بود، یعنی:

$$\text{LimP}(|\hat{\beta} - \beta| < \varepsilon) = 1 \qquad (۲)$$

بنابراین، وقتی حجم نمونه افزایش یابد مقدار $\hat{\beta}$ به β نزدیک می‌شود.

رگرسیون کاذب (Spurious Regression)

رگرسیون کاذب یعنی پارامترها و آماره‌های برازش‌شده با روش حداقل مربعات معمولی درست و قابل اتکا نیستند. در رگرسیون کاذب یا ساختگی درعین‌حال که هیچ رابطه بامفهومی بین متغیرها وجود ندارد ولی ضریب تعیین بزرگ و آمار t ضرایب برازش‌شده بزرگ به‌دست می‌آید که درنهایت استنباط‌های غلطی در مورد میزان ارتباط میان متغیرها را نشان می‌دهد. بنابراین، برای جلوگیری از کاذب بودن یک رگرسیون باید ابتدا در مورد پایایی (مانایی) متغیرهای مدل اطمینان حاصل کرد. به‌طورکلی، یک فرآیند تصادفی هنگامی پایا نامیده می‌شود که میانگین و واریانس آن طی زمان ثابت باشد و مقدار کوواریانس بین دو دوره زمانی تنها به‌فاصله یا وقفه بین این دو دوره بستگی داشته و ارتباطی به زمان واقعی محاسبه کوواریانس نداشته باشد (نوفرستی، ۱۳۹۱، ۸۵). برای برازش با روش حداقل مربعات

Encyclopedia of Economics
enceco.ihcs.ac.ir

معمولی باید همه متغیرها پایا باشند تا بـه رگرسـیون کـاذبی منتج نشود. اگر همه متغیرهـا ناپایـا باشـند در ایـن صـورت آزمـون هـم‌جمعـی انجـام مـی‌گیـرد. مفهـوم هـم‌جمعـی (Cointegration) آن است که دو یا چند متغیر سـری زمـانی براساس مبانی نظری با یکدیگر یک رابطه تعادلی بلندمـدت دارند، هرچند ممکن است خود این سری‌های زمـانی دارای روند تصادفی باشند (ناپایا باشند)، اما در طول زمان یکدیگر را به‌خوبی دنبال می‌کنند به‌گونه‌ای‌که تقاضل بین آنها باثبـات (پایا) است. بنابراین، مفهوم هم‌جمعـی تـداعی‌کننـده وجـود یک رابطه تعادلی بلندمدت اسـت کـه سیسـتم اقتصـادی در طـول زمـان بـه‌سـمت آن حرکـت مـی‌کنـد. ازایـن‌رو، اگـر سری‌های زمانی ناپایا ولی هم‌جمع باشند به یـک رگرسـیون کاذب منتج نخواهند شد.

کتاب‌شناسی

درخشـان، مسـعود، (۱۳۹۵)، اقتصادسـنجی تـک معـادلات بـا فـروض کلاسیک، جلد اول، انتشارات سمت.

سـوری، علـی، (۱۳۹۴)، اقتصادسـنجی مقـدماتی، جلـد اول، نشـر فرهنگ‌شناسی.

گجراتی، دامودار، (۱۳۹۱)، مبانی اقتصادسنجی، ترجمه حمیـد ابریشـمی، جلد اول، چاپ دوازدهم، انتشارات دانشگاه تهران.

نوفرستی، محمد، (۱۳۹۱)، ریشه واحـد و هـم‌جمعـی در اقتصادسـنجی، تهران، انتشارات درسا.

Wooldrige, Jeffrey, (2009), Introductory Econometrics: A Modern. Approach, Fifth Edition.

دکتر ویدا ورهرامی

هیئت علمی دانشگاه شهید بهشتی

Encyclopedia of Economics

enceco.ihcs.ac.ir