

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 می‌باشد. در حقیقت برازش یک معادله رگرسیونی به دنبال تبیین معادله رفتاری میان متغیرهای مستقل و وابسته هست. معادلات رفتاری بیانگر نحوه عکس العمل متغیرهای مستقل و وابسته می‌باشد. تفاوت کلی مدل‌های ریاضی و مدل‌های رگرسیون در جمله اخلاص است. هرگاه به مدل‌های ریاضی یک جمله اخلاص - که یقیناً تصادفی است - اضافه شود به یک مدل رگرسیون تبدیل خواهد شد.

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \dots + \beta_k X_{k,t} + \epsilon_t \quad (1)$$

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

متغیر تصادفی (Random Variable)

متغیر تصادفی، متغیری است که مقادیرش در نتیجه آزمایش انتخاب می‌شود. درحقیقت متغیر تصادفی تابعی از نتایج یک آزمایش است. بنابراین هر یک از مقادیر متغیر تصادفی بیانگر وقوع یک حادثه است که هر حادثه نیز پیامد یک آزمایش است و فضای نمونه را تشکیل می‌دهد. لذا بین عناصر فضای نمونه و مقادیر متغیر تصادفی یک تناظر وجود دارد و هر یک 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 می‌باشد. بنابراین می‌توان گفت یک متغیر تصادفی متغیری است که تحت تأثیر عوامل متعددی قرار دارد و خصوصیتی دارد که توسط توزیع آن و فروض مربوط به آن بیان می‌شود (سوری، ۱۳۹۴، ۳۴).

روش حداقل مربعات معمولی ( Ordinary Least Square, OLS)

یکی از روش‌های مهم جهت تخمین پارامترهای مدل است که کاربرد زیادی در تخمین معادلات رگرسیونی دارد. طرح اولیه این روش را کارل فریدریش گوس ( Carl Friedrich Gauss) ریاضی‌دان معروف آلمانی در قرن هجدهم مطرح کرده است. با این روش در حقیقت مجموع مربعات خطا حداقل شده و با این حداقل سازی ضرایب رابطه برازش می‌شوند. تخمین‌زنده‌های حداقل مربعات معمولی دارای ویژگی‌هایی مثل بدون تورش بودن، کارایی و سازگاری هستند و اصطلاحاً آنها را تخمین‌زنده خطی ناریب یا BLUE می‌نامند. (سوری، ۱۳۹۴، ۸۴) جهت برازش یک معادله رگرسیونی با روش حداقل مربعات معمولی که تخمین‌زنده‌های خطی ناریب، کارا و سازگار نتیجه دهند، باید فروض کلاسیک برقرار باشد. این فروض، مشتمل بر خصوصیات پسماند مدل رگرسیونی است. مهم‌ترین نکته در مورد پسماند، تصادفی بودن آن است. پسماند مدل، یک متغیر تصادفی است و مثل همه متغیرهای تصادفی دارای یک تابع توزیع احتمال و در نتیجه میانگین و واریانس است. اولین فرض کلاسیک این است که میانگین یا امید ریاضی جمله اخلاص صفر است. این فرض در واقع به این معنی است که به‌ازای هر مقدار معین از متغیرهای توضیح‌دهنده، میانگین تمام مقادیر ممکن برابر  $\epsilon_t$  صفر می‌باشد. مفهوم کلی این فرض این است که مدل خطای سیستماتیک ندارد. دومین فرض کلاسیک، ثابت بودن واریانس جمله اخلاص به‌ازای مقادیر مختلف متغیرهای مستقل است

گرایش دارد. (گجراتی، ۱۳۹۱، ۱۲۰)

### کارایی (Efficiency)

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

### سازگاری (Consistency)

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

$$\lim P(|\hat{\beta} - \beta| < \varepsilon) = 1 \quad (۲)$$

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

### رگرسیون کاذب (Spurious Regression)

رگرسیون کاذب، یعنی پارامترها و آماره‌های برازش شده با روش حداقل مربعات معمولی درست و قابل اتکا نیستند. در رگرسیون کاذب یا ساختگی در عین حالی که هیچ رابطه بامفهومی بین متغیرها وجود ندارد ولی ضریب تعیین بزرگ و آمار  $t$  ضرایب برازش شده بزرگ به دست می‌آید که در نهایت استنباط‌های غلطی در مورد میزان ارتباط میان متغیرها را نشان می‌دهد. لذا 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،

$\text{Var}(\varepsilon_i) = \delta^2$ . هرگاه واریانس جمله اخلاخل ثابت باشد، مدل واریانس همسان و در غیر این صورت واریانس ناهمسان است. سومین فرض این است که  $\varepsilon_i$  و  $\varepsilon_j$  به‌ازای تمامی مقادیر  $i \neq j$  از یکدیگر مستقل باشند. یعنی کوواریانس  $\text{COV}(\varepsilon_i, \varepsilon_j) = 0$  آنها صفر است. به‌عبارت‌دیگر هرگاه دو مقدار متفاوت برای متغیرهای مستقل در نظر گرفته شوند، فرض بر این است که جمله‌های اخلاخل متناظر با آنها از یکدیگر مستقلند. در چنین حالتی بیان می‌شود که جمله‌های اخلاخل خودهمبستگی ندارند. چهارمین فرض این است که تابع توزیع جمله اخلاخل نرمال باشد؛ بنابراین با توجه به فرض‌های اول و دوم و سوم می‌توان گفت که  $\varepsilon_i$  دارای توزیع نرمال با میانگین صفر و واریانس ثابت  $\delta^2$  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 بدین معناست که  $X_{i,t}$ ‌ها از متغیر تصادفی  $\varepsilon_t$  مستقل هستند. (درخشان، ۱۳۹۵، ۸۰)

### تورش (Bias)

برای تخمین یک پارامتر، تخمین‌زننده‌های مختلفی را می‌توان به دست آورد لذا باید مناسب‌ترین تخمین‌زننده را انتخاب نمود. 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 نااریب است. لذا خاصیت نااریبی به‌صورت  $E(\hat{\beta}) = \beta$  است. حال اگر تخمین‌زننده دارای تورش باشد، در آن صورت  $E(\hat{\beta}) - \beta = \varepsilon$  است که  $\varepsilon$  میزان تورش است. در حقیقت مفهوم تورش بدان معنا است که تخمین‌زننده به‌طور متوسط به پارامتر مورد نظر

۸۵) جهت برازش با روش حداقل مربعات معمولی باید همه متغیرها پایا باشند تا به رگرسیون کاذبی منتج نشود. اگر همه متغیرها ناپایا باشند در این صورت آزمون هم‌جمعیتی انجام می‌گیرد. مفهوم هم‌جمعیتی (Cointegration) آن است که دو یا چند متغیر سری زمانی براساس مبانی نظری با یکدیگر یک رابطه تعادلی بلندمدت دارند، هرچند ممکن است خود این سری‌های زمانی دارای روند تصادفی باشند (ناپایا باشند)، اما در طول زمان یکدیگر را به خوبی دنبال می‌کنند به گونه‌ای که تقاضا بین آنها باثبات (پایا) است. بنابراین مفهوم هم‌جمعیتی تداعی‌کننده وجود یک رابطه تعادلی بلندمدت است که سیستم اقتصادی در طول زمان به سمت آن حرکت می‌کند. لذا اگر سری‌های زمانی ناپایا ولی هم‌جمع باشند، به یک رگرسیون کاذب منتج نخواهند شد.

#### کتاب‌شناسی

- درخشان، مسعود، (۱۳۹۵)، اقتصادسنجی تک معادلات با فروض کلاسیک، جلد اول، انتشارات سمت.
- سوری، علی، (۱۳۹۴)، اقتصادسنجی مقدماتی، جلد اول، نشر فرهنگ‌شناسی.
- گجراتی، دامودار، (۱۳۹۱)، مبانی اقتصادسنجی، ترجمه حمید ابریشمی، جلد اول، چاپ دوازدهم، انتشارات دانشگاه تهران.
- نوفرستی، محمد، (۱۳۹۱)، ریشه واحد و هم‌جمعیتی در اقتصادسنجی، تهران، انتشارات درسا.

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

ویدا وهرامی

هیئت علمی دانشگاه شهید بهشتی