

رابطه‌ای برای کوواریانس شرطی و کاربردهایی از آن

حمزه ترابی*

چکیده

رابطه‌ی سومی که گشتاورهای غیرشرطی را به گشتاورهای شرطی مربوط می‌کند، رابطه‌ی کوواریانس شرطی است و آن عبارت است از

$$\text{Cov}(f(X), Y) = \text{Cov}(f(X), E[Y|X]) \quad (۳)$$

که در آن، $f(X)$ تابعی از X و خود یک متغیر تصادفی است. اثبات این رابطه، در [۳] صفحه‌های ۵۱ تا ۵۳ آورده شده است. نکته شگفت‌آور رابطه (۳) این است که طرف راست آن کوواریانس دو تابع از X است، در حالی که طرف چپ آن کوواریانس تابعی از X و Y است.

در حالت ویژه اگر در رابطه (۳) بگذاریم: $f(X) = X$ ، خواهیم داشت:

$$\text{Cov}(X, Y) = \text{Cov}(X, E[Y|X]) \quad (۳')$$

اینک به بیان و اثبات قضیه‌ای درباره‌ی کوواریانس شرطی می‌پردازیم که در آن کوواریانس دو تابع از X ، برابر کوواریانس دو تابع از Y است.

قضیه ۱. فرض کنیم X و Y دو متغیر تصادفی با گشتاورهای اول و دوم متناهی و توزیع احتمال توأم دلخواهی باشند. در این صورت اگر $f(X)$ و $g(Y)$ به ترتیب دو تابع از X و Y باشند به گونه‌ای که هر دو متغیر تصادفی باشند، داریم

$$\text{Cov}(f(x), g(Y)) = \text{Cov}(E[f(X)|Y], E[g(Y)|X])$$

در مباحث احتمال و فرایندهای تصادفی، بارها با رابطه‌های امید ریاضی مکرر، واریانس غیرشرطی برحسب واریانس شرطی، و حالتی ویژه از رابطه کوواریانس شرطی برخورد کرده‌ایم و کم و بیش با کاربردهایی از آنها آشنا شده‌ایم.

در این مقاله، پس از بیان این سه رابطه، دو قضیه و سپس دو کاربرد را که اولی در رابطه با برآوردهای پارامترهای مدل رگرسیون خطی، و دومی در رابطه با به دست آوردن کران بالایی برای تابع مخاطره بیز است، بررسی و اثبات می‌کنیم.

۱ پیشگفتار

پنداریم X و Y دو متغیر تصادفی با گشتاورهای اول و دوم متناهی با توزیع احتمال توأم دلخواهی باشند. رابطه امید ریاضی مکرر عبارت است از:

$$E[Y] = E[E[Y|X]] \quad (۱)$$

برای اثبات، می‌توان مثلاً به [۴]، صفحه ۲۱۳ مراجعه کرد. رابطه دیگری که واریانس غیرشرطی را برحسب واریانس شرطی به دست می‌آورد، رابطه

$$\text{Var}[Y] = \text{Var}[E[Y|X]] + E[\text{Var}[Y|X]] \quad (۲)$$

است که اثبات آن را می‌توان، مثلاً در [۲]، صفحه ۱۵۹ ملاحظه کرد.

* حمزه ترابی، دانشکده علوم ریاضی، دانشگاه یزد

برهان. بنا به فرض $g(Y)$ یک متغیر تصادفی است. بنابراین با گذاشتن $g(Y)$ به جای Y در (۳) خواهیم داشت:

$$\begin{aligned} \text{Cov}(X, E[Y|X]) &= \text{Cov}(X, \alpha + \beta X) \\ &= \text{Cov}(X, \alpha) + \beta \text{Cov}(X, X) \\ &= 0 + \beta \cdot \text{Var}[X] \end{aligned}$$

$$\text{Cov}(f(X), g(Y)) = \text{Cov}(f(X), E[g(Y)|X]) \quad (۴)$$

ولی بنا بر (۳)، $\text{Cov}(X, Y) = \text{Cov}(X, E[Y|X])$. در نتیجه،

ولی $\text{Cov}(f(X), g(Y)) = \text{Cov}(g(Y), f(X))$. بنابراین با تعریف متغیرهای تصادفی جدید و به کارگیری دوباره (۳)، داریم:

$$\text{Cov}(X, Y) = \beta \text{Var}[X]$$

$$\text{Cov}(f(X), g(Y)) = \text{Cov}(g(Y), E[f(X)|Y]) \quad (۵)$$

پس،

$$\beta = \frac{\text{Cov}(X, Y)}{\text{Var}[X]}$$

با مقایسه (۴) و (۵)، حکم ثابت می‌شود. □

اینک با امیدگیری از دو طرف (۶) و با به کارگیری رابطه (۱)، خواهیم داشت:

قضیه ۲ با داشتن شرایط قضیه ۱، داریم:

$$E[Y] = \alpha + \beta E[X]$$

$$E[X \cdot (Y - E[Y|X])] = 0$$

ولی

$$\beta = \frac{\text{Cov}(X, Y)}{\text{Var}[X]}$$

برهان. بنا بر (۳) داریم:

$$\alpha = E[Y] - \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \cdot E(X)$$

بنابراین،

$$\text{Cov}(X, Y) = \text{Cov}(X, E[Y|X])$$

بنابراین،

از طرفی می‌دانیم برآوردی منطقی برای $E[g(X)]$ که $g(X)$ تابعی دلخواه از X است، $\bar{g} = \frac{\sum_{i=1}^n g(x_i)}{n}$ است که در آن x_1, x_2, \dots, x_n مشاهده‌های یک نمونه تصادفی است. بنابراین منطقی است $E[Y], E[X]$ و $\text{Cov}(X, Y)$ را به ترتیب با \bar{y}, \bar{x} و $\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ برآورد کنیم که در نتیجه خواهیم داشت:

$$E[X \cdot Y] - E[X] \cdot E[Y] = E[X \cdot E[Y|X]] + E[X] \cdot E[E[Y|X]]$$

اینک با به کارگیری رابطه (۱)، خواهیم داشت:

$$E[X \cdot Y] - E[X \cdot E[Y|X]] = 0$$

در نتیجه با بهره‌گیری ویژگی جمعپذیری امید ریاضی، داریم:

$$E[X \cdot Y - X \cdot E[Y|X]] = 0$$

که همان برآوردهای روش کمترین توان دوم خطا هستند. □

بنابراین،

کاربرد ۲ فرض کنیم می‌خواهیم پارامتر θ یک متغیر تصادفی با تابع توزیع $F(x|\theta)$ را برآورد کنیم و تصمیم ما برگزیدن $\hat{\theta}$ به عنوان برآورد θ باشد. همچنین فرض کنیم تابع زیان درجه دو $L(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$ را برای این تصمیم برگزیده‌ایم. یک کران بالا برای تابع مخاطره بیز پسین، در بین رده همه توزیعهای پیشین $\pi(\theta)$ با گشتاور دوم متناهی ($E[\theta^2] < \infty$) و توزیعهای نمونه‌ای $F(x|\theta)$ به دست آورید.

$$\square \quad E[X \cdot (Y - E[Y|X])] = 0$$

اینک به بیان و بررسی دو کاربرد مهم از رابطه‌های بالا می‌پردازیم:

کاربرد ۱. گیریم متغیرهای تصادفی X و Y دارای توزیع احتمال دلخواهی باشند و همچنین رگرسیون Y روی X خطی باشد. برآورد پارامترهای مدل رگرسیون خطی بالا را به دست آورید.

حل می‌دانیم که برآورد کننده بیز $\hat{\theta} = E[\theta|X]$ است. بنابراین اگر تابع مخاطره پسین برآورد کننده $\hat{\theta}$ و چگالی پسین، $\pi(\theta|X)$ را با $r_{\hat{\theta}, \pi(\theta|X)}$ نشان دهیم، بنا بر تعریف خواهیم داشت:

حل چون رگرسیون Y روی X خطی است، داریم:

$$E[Y|X] = \alpha + \beta X \quad (۶)$$

$$r_{\hat{\theta}, \pi(\theta|X)} = E[L(\theta, \hat{\theta})] = E[(E[\theta|X] - \theta)^2]$$

بنابراین،

بنابراین با به کارگیری (۱) داریم:

بنابراین

$$r_{\hat{\theta}, \pi(\theta|X)} = E[E\{\{E[\theta|X] - \theta\}^2|X\}]$$

اینک با بهره‌گیری تعریف واریانس شرطی

$$r_{\hat{\theta}, \pi(\theta|X)} \leq \text{Var}[\theta] \cdot (1 - \text{Corr}^2(X, \theta))$$

توجه می‌کنیم که برابری برقرار است اگر و تنها اگر رگرسیون θ روی X خطی باشد؛ چرا که در نابرابری کوشی شوارتس، برابری برقرار است اگر و تنها اگر بین دو متغیر یک رابطه خطی برقرار باشد، یعنی بین X و $E[\theta|X]$ رابطه‌ای خطی وجود داشته باشد. بنابراین،

$$r_{\hat{\theta}, \pi(\theta|X)} = \text{Var}[\theta] \cdot (1 - \text{Corr}^2(X, \theta))$$

اگر و تنها اگر

$$\square \quad E[\theta|X] = \alpha + \beta X$$

$$\text{Var}[Y|X] = E[(Y - E[Y|X])^2|X]$$

مشروط بر اینکه $E[\theta^2] < \infty$ و با توجه به اینکه $E[\theta^2] < \infty$ و رابطه (۲)، خواهیم داشت

$$r_{\hat{\theta}, \pi(\theta|X)} = E[\text{Var}[\theta|X]] = \text{Var}[\theta] - \text{Var}[E[\theta|X]]$$

اکنون نابرابری کوشی - شوارتس را برای دو متغیر تصادفی X و $E[\theta|X]$ به کار می‌بندیم. خواهیم داشت:

$$\begin{aligned} r_{\hat{\theta}, \pi(\theta|X)} &\leq \text{Var}[\theta] - \frac{\text{Cov}^2(X, E[\theta|X])}{\text{Var}[X]} \\ &= \text{Var}[\theta] \cdot (1 - \text{Corr}^2(X, \theta)) \end{aligned}$$

مراجع

- [1] Ash, R. (1970), *Basic Probability Theory*, Wiley, New York.
- [2] Mood, A. M., Graybill, F. A., and Boes, D. C. (1974), *Introduction to the Theory of Statistics*, 3rd ed., McGraw-Hill Book Company, New, York.
- [3] Parzen, E. (1962), *Stochastic Processes*, Holden Dag.
- [4] Ross, S. M. (1976), *A First Course in Probability*, Thrd. Edition, MacMillan Publishing Co., New York.
- [5] Wallenius, K. T. (1971), *A Conditional Covariance Formula with Applications*, The American Statistician, Vol. 25, No. 3.