رابطه‌ای برای کوواریانس شرطی و کاربردهایی از آن

حمزهٔ ترابی*

چکیده

در مباحث احتمال و فرایندهای تصادفی، بارها با رابطه‌های امید ریاضی مکرر، واریانس غیرشرطی برحسب واریانس شرطی، و حالتی ویژه از رابطهٔ کوواریانس شرطی برخوردار 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 کرانِ 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. رابطهٔ امید ریاضی مکرر عبارت است از:

$$E[Y] = E\big[E[Y|X]\big] \qquad (۱)$$

برای اثبات، می‌توان مثلاً به [۴]، صفحهٔ ۲۱۳ مراجعه کرد.

رابطهٔ دیگری که واریانس غیرشرطی را برحسب واریانس شرطی به دست می‌آورد، رابطه

$$\mathrm{Var}[Y] = \mathrm{Var}\big[E[Y|X]\big] + E\big[\mathrm{Var}[Y|X]\big] \qquad (۲)$$

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

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

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

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 رابطهٔ (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 رابطهٔ (۳) بگذاریم: $f(X) = X$، خواهیم داشت:

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

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

قضیهٔ 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

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

*) حمزهٔ ترابی، دانشکدهٔ علوم ریاضی، دانشگاه یزد

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

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

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

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

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

قضیه ۲ با داشتن شرایط قضیهٔ ۱، داریم:

$$E[X.(Y-E[Y|X])] = ۰$$

برهان. بنابر (۳′) داریم:

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

بنابراین،

$$E[X.Y]-E[X].E[Y] = E[X.E[Y|X]]+E[X].E[E[Y|X]]$$

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

$$E[X.Y] - E[X.E[Y|X]] = ۰$$

در نتیجه با بهره‌گیری ویژگی جمع‌پذیری امید ریاضی، داریم:

$$E[X.Y - X.E[Y|X]] = ۰$$

بنابراین،

$$E[X.(Y-E[Y|X])] = ۰ \qquad \square$$

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 چون رگرسیون Y روی X خطی است، داریم:

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

بنابراین،

$$\begin{aligned}\operatorname{Cov}(X,E[Y|X]) &= \operatorname{Cov}(X,\alpha+\beta X)\\ &= \operatorname{Cov}(X,\alpha)+\beta\operatorname{Cov}(X,X)\\ &= ۰ + \beta.\operatorname{Var}[X]\end{aligned}$$

ولی بنابر (۳′)، $\operatorname{Cov}(X,Y) = \operatorname{Cov}(X,E[Y|X])$. در نتیجه،

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

پس،

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

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

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

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

بنابراین،

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

از طرفی می‌دانیم برآوردی منطقی برای $E[g(X)]$، که $g(X)$ تابعی دلخواه از X است، $\overline{g(x)} = \frac{\sum_{i=۱}^{n} g(x_i)}{n}$ است که در آن $x_۱, x_۲, \ldots, x_n$ مشاهده‌های یک نمونهٔ تصادفی است. بنابراین منطقی است $E[X]$، $E[Y]$، $\operatorname{Var}[X]$، و $\operatorname{Cov}(X,Y)$ را به ترتیب با $\bar{x}$، $\bar{y}$، $\frac{۱}{n}\sum_{i=۱}^{n}(x_i-\bar{x})^۲$، و $\frac{۱}{n}\sum_{i=۱}^{n}(x_i-\bar{x})(y_i-\bar{y})$ برآورد کنیم که در نتیجه خواهیم داشت:

$$\hat{\beta} = \frac{\sum_{i=۱}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=۱}^{n}(x_i-\bar{x})^۲} \qquad \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

که همان برآوردهای روش کمترین توان دوم خطا هستند. □

کاربرد ۲ فرض کنیم می‌خواهیم پارامتر θی 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 درجهٔ دو $L(\theta,\hat{\theta}) = (\theta-\hat{\theta})^۲$ را برای این تصمیم برگزیده‌ایم. یک کران بالا برای تابع مخاطرهٔ بیز پسین، در بین ردهٔ همهٔ توزیعهای پیشین $\pi(\theta)$ با گشتاور دوم متناهی $(E[\theta^۲] < \infty)$ و توزیعهای نمونه‌ای $F(x|\theta)$ به دست آورید.

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

$$r_{\hat{\theta},\pi(\theta|X)} = E[L(\theta,\hat{\theta})] = E[(E[\theta|X]-\theta)^۲]$$

بنابراین با به کارگیری (۱) داریم:

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

اینک با بهره‌گیری تعریف واریانس شرطی

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

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

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

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

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

بنابراین

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

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 رابطهٔ خطی برقرار باشد، یعنی بین X و $E[\theta|X]$ رابطه‌ای خطی وجود داشته باشد. بنابراین،

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

اگر و تنها اگر

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

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.