دانشکده‌ی علوم ریاضی

احتمال و کاربرد آن — ۱۶ اردیبهشت ۹۳

جلسه‌ی ۲۰: کوواریانس، واریانس مجموع، همبستگی

مدرّس: دکتر شهرام خزائی — *نگارنده: سجاد گودرزی*

۱ کوواریانس

تعریف ۱ *اگر X و Y دو متغیر تصادفی دلخواه باشند، کوواریانس[۱] آن‌ها با نماد $\mathrm{Cov}(X,Y)$ نشان داده شده و بدین صورت تعریف می‌شود:*

$$\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])]$$

با بسط دادن رابطه‌ی بالا خواهیم داشت:

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])]\\ &= \mathrm{E}[XY - X\mathrm{E}[Y] - Y\mathrm{E}[X] + \mathrm{E}[X]\mathrm{E}[Y]]\\ &= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]\end{aligned}$$

گزاره ۱ *با استفاده از تعریف کوواریانس، می‌توان اثبات کرد:*

$$\begin{aligned}&\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)\\ &\mathrm{Cov}(X,X) = \sigma_X^2\\ &\mathrm{Cov}(aX+b, cY+d) = ac\,\mathrm{Cov}(X,Y)\\ &\mathrm{Cov}(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m}\mathrm{Cov}(X_i,Y_j)\end{aligned}$$

نکته ۱

$$\mathrm{Var}(\sum_{i=1}^{N} X_i) = \sum_{i=1}^{N}\mathrm{Var}(X_i) + 2\sum\sum_{i<j}\mathrm{Cov}(X_i,X_j)$$

تعریف ۲ *اگر X و Y دو متغیر تصادفی باشند که $\infty > \sigma_X^2 > 0$ و $\infty > \sigma_Y^2 > 0$، آنگاه کوواریانس بین متغیرهای استاندارد شده‌ی X و Y، ضریب همبستگی[۲] بین X و Y نامیده شده و به صورت زیر تعریف می‌گردد:*

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

[۱] Covariance
[۲] Correlation coefficient

گزاره ۲ *برای هر دو متغیر تصادفی X و Y داریم:*

$$-1 \leq \rho(X,Y) \leq 1$$

برهان. *با توجه به این انتقال متغیرهای تصادفی X و Y کواریانس و ضریب همبستگی آنها را تغییر نمی‌دهد، برای سادگی فرض می‌کنیم که X و Y دارای متوسط صفر هستند. حال پارامتر حقیقی t و واریانس متغیر تصادفی $X - tY$ را که همواره کمیتی نامنفی است، در نظر بگیرید. داریم:*

$$\begin{aligned} 0 &\leq \mathrm{Var}(tX-Y) \\ &= \mathrm{E}[(tX-Y)^2] \\ &= t^2\mathrm{E}[X^2] - 2t\mathrm{E}[XY] + \mathrm{E}[Y^2] \\ &= t^2\mathrm{Var}[X] - 2t\mathrm{Cov}(X,Y) + \mathrm{Var}[Y] \\ &= t^2\sigma_X^2 - 2t\mathrm{Cov}(X,Y) + \sigma_Y^2 \end{aligned}$$

برای اینکه عبارت درجه دوم فوق همواره نامنفی باشد، باید داشته باشیم:

$$\mathrm{Cov}^2(X,Y) - \sigma_X^2\sigma_Y^2 \leq 0$$

یا به طور معادل

$$|\rho(X,Y)| = |\frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}| \leq 1$$

■

نکته ۲ *اگر X,Y مستقل باشند، آنگاه داریم:*

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])] \\ &= \mathrm{E}[(X-\mathrm{E}[X])]\mathrm{E}[(Y-\mathrm{E}[Y])] \\ &= 0 \end{aligned}$$

نتیجه ۱ *اگر X_i ها مستقل باشند:*

$$\mathrm{Var}(\sum_{i=1}^{N} X_i) = \sum_{i=1}^{N} \mathrm{Var}(X_i)$$

نکته ۳ *دو متغیر تصادفی با کوواریانس صفر، ناهمبسته نامیده می‌شوند. متغیرهای تصادفی ناهمبسته، لزوماً مستقل نیستند. ناهمبستگی دو متغیر تصادفی را می‌توان به عدم وجود ربط سیستماتیک بین آنها تفسیر کرد.*

۲ متغیرهای همبسته

تعریف ۳ *اگر دو متغیر تصادفی ناهمبسته نباشند، همبسته نامیده می‌شوند. همبستگی می‌تواند مثبت ($\rho_{X,Y} > 0$) یا منفی ($\rho_{X,Y} < 0$) باشد.*

مثال ۱ *قد و وزن افراد همبستگی مثبت دارند؛ یعنی اگر یکی بیشتر از متوسط باشد، به طور معمول دیگری نیز، بیشتر از متوسط است.*

مثال ۲ *میزان مصرف سیگار و طول عمر افراد دو متغیر همبسته‌ی منفی هستند.*

۳ متوسط و واریانس شرطی

تعریف ۴ *برای متغیرهای تصادفی X و Y، تابع جرم احتمال X به شرط $Y = y$ بدین صورت است:*

$$p_{X|Y}(x, y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

دقت کنید که به ازای هر y که $p_Y(y) > 0$، $p_{X|Y}(x, y)$ خود یک تابع چگالی احتمال است. متوسط و واریانس شرطی برای این تابع چگالی احتمال به صورت زیر تعریف می‌شود:

$$\mathrm{E}[X|y] = \sum_x x p_{X|Y}(x, y)$$
$$\mathrm{Var}(X|y) = \mathrm{E}[(X - \mathrm{E}[X|y])^۲|y]$$

دقت کنید که متوسط و واریانس‌های شرطی بالا، توابعی حقیقی از y هستند. با استفاده از این توابع، می‌توان متغیرهای تصادفی $\mathrm{E}[X|Y]$ و $\mathrm{Var}(X|Y)$ را تعریف نمود، که توابعی از متغیر تصادفی Y هستند.
با توجه به نکات گفته شده، قضیه‌ی مهم زیر را بیان می‌کنیم.

قضیه ۱ (متوسط شرطی)
$$\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X|Y]]$$

برهان. *با استفاده از قضیه متوسط کل داریم:*

$$\mathrm{E}[X] = \sum_y \mathrm{E}[X|Y = y] \Pr\{Y = y\}$$

زیرا طبق تعریف $\mathrm{E}[X]$ داریم:

$$\begin{aligned}
\mathrm{E}[X] &= \sum_x x\, p_X(x) = \sum_x x \sum_y p_{X,Y}(x, y) \\
&= \sum_y \sum_x x\, p(x, y) = \sum_y \sum_x x\, p_{X|Y}(x|y)\, p_Y(y) \\
&= \sum_y \Big(\sum_x x\, p_{X|Y}(x|y)\Big) p_Y(y) \\
&= \sum_y \mathrm{E}[X|Y = y]\ \Pr\{Y = y\}
\end{aligned}$$

از طرف دیگر می‌دانیم برای هر تابع حقیقی مقدار مانند h داریم

$$\mathrm{E}[h(Y)] = \sum_y h(y)\ \Pr\{Y = y\}\,.$$

پس با جای‌گذاری $h(y) = \mathrm{E}[X|y]$ می‌توان نوشت:

$$\mathrm{E}[\mathrm{E}[X|Y]] = \sum_y \mathrm{E}[X|Y = y]\ \Pr\{Y = y\}$$

از تساوی‌های فوق می‌توان حکم را نتیجه گرفت. ■

به طور مشابه، قضیه‌ی زیر را برای واریانس شرطی داریم.

قضیه ۲ (واریانس شرطی)
$$\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\mathrm{E}[X|Y])$$

۴ متغیرهای تصادفی مستقل با توزیع یکسان

تعریف ۵ *دنباله‌ای از متغیرهای تصادفی را مستقل با توزیع یکسان $(i.i.d)$*[۳] *می‌نامیم، اگر همه‌ی آنها دارای توزیع یکسان بوده و هر زیر مجموعه‌ی دلخواهی از آنها، متغیرهای تصادفی مستقل باشند.*

مثال ۳ فرض کنید N یک متغیر تصادفی با مقادیر طبیعی باشد. متغیر تصادفی Y را به صورت زیر در نظر بگیرید

$$Y = X_1 + X_2 + X_3 + \cdots + X_N$$

که در آن X_i ها $i.i.d$ و همگی مستقل از N هستند. فرض کنید:

$$\begin{aligned} \mathrm{E}[X_i] &= \mu_X \\ \mathrm{Var}(X_i) &= \sigma_X \\ \mathrm{E}[N] &= \mu_N \\ \mathrm{Var}(N) &= \sigma_N \end{aligned}$$

برای محاسبه‌ی متوسط و واریانس Y ابتدا متوسط شرطی $\mathrm{E}[Y|N]$ را محاسبه می‌کنیم:

$$\begin{aligned} \mathrm{E}[Y|N] &= \mathrm{E}[X_1 + X_2 + X_3 + \cdots + X_N|N] \\ &= \mathrm{E}[\mathrm{E}[X_1|N] + \mathrm{E}[X_2|N] + \mathrm{E}[X_3|N] + \cdots + \mathrm{E}[X_N|N]] \\ &= \mathrm{E}[\mathrm{E}[X_1] + \mathrm{E}[X_2] + \mathrm{E}[X_3] + \cdots + \mathrm{E}[X_N]] \\ &= N\mu_X \end{aligned}$$

با جایگذاری در قضیه متوسط شرطی داریم:

$$\begin{aligned} \mathrm{E}[Y] &= \mathrm{E}[\mathrm{E}[Y|N]] \\ &= \mathrm{E}[N\mu_X] \\ &= \mu_X \mathrm{E}[N] \\ &= \mu_N \mu_X \end{aligned}$$

برای محاسبه‌ی واریانس Y به $\mathrm{Var}(Y|N)$ نیز نیاز داریم:

$$\begin{aligned} \mathrm{Var}(Y|N) &= \mathrm{Var}(X_1 + X_2 + X_3 + \cdots + X_N|N) \\ &= \mathrm{Var}(X_1|N) + \mathrm{Var}(X_2|N) + \mathrm{Var}(X_3|N) + \cdots + \mathrm{Var}(X_N|N) \\ &= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3) + \cdots + \mathrm{Var}(X_N) \\ &= N\sigma_X^2 \end{aligned}$$

با جای‌گذاری در قضیه متوسط شرطی خواهیم داشت:

$$\begin{aligned} \mathrm{Var}(Y) &= \mathrm{E}[\mathrm{Var}(Y|N)] + \mathrm{Var}(\mathrm{E}[Y|N]) \\ &= \mathrm{E}[N\sigma_X^2] + \mathrm{Var}(N\mu_X) \\ &= \sigma_X^2 \mu_N + \mu_X^2 \sigma_N^2 \end{aligned}$$

[۳] Independent and identically distributed