Lecture 15

1. Expectations

Theorem 15.1. *Let g be a real-valued function of two variables, and (X, Y) have joint mass function f. If the sum converges then*

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)f(x,y).$$

Corollary 15.2. *For all a, b real,*

$$E(aX + bY) = aEX + bEY.$$

Proof. Setting $g(x,y) = ax + by$ yields

$$\begin{aligned} E(aX + bY) &= \sum_x \sum_y (ax + by)f(x,y) \\ &= \sum_x ax \sum_y f(x,y) + \sum_x \sum_y byf(x,y) \\ &= a \sum_x xf_X(x) + b \sum_y y \sum_x f(x,y) \\ &= aEX + b \sum_y f_Y(y), \end{aligned}$$

which is $aEX + bEY$. □

2. Covariance and correlation

Theorem 15.3 (Cauchy–Schwarz inequality)**.** *If* $E(X^2)$ *and* $E(Y^2)$ *are finite, then*

$$|E(XY)| \leqslant \sqrt{E(X^2)\, E(Y^2)}.$$

Proof. Note that

$$\begin{aligned}&\left(XE(Y^2) - YE(XY)\right)^2\\ &\qquad = X^2 \left(E(Y^2)\right)^2 + Y^2 \left(E(XY)\right)^2 - 2XYE(Y^2)E(XY).\end{aligned}$$

Therefore, we can take expectations of both side to find that

$$\begin{aligned}&E\left[\left(XE(Y^2) - YE(XY)\right)^2\right]\\ &\qquad = E(X^2)\left(E(Y^2)\right)^2 + E(Y^2)\left(E(XY)\right)^2 - 2E(Y^2)\left(E(XY)\right)^2\\ &\qquad = E(X^2)\left(E(Y^2)\right)^2 - E(Y^2)\left(E(XY)\right)^2.\end{aligned}$$

The left-hand side is $\geqslant 0$. Therefore, so is the right-hand side. Solve to find that

$$E(X^2)E(Y^2) \geqslant \left(E(XY)\right)^2.$$

[If $E(Y^2) > 0$, then this is OK. Else, $E(Y^2) = 0$, which means that $P\{Y = 0\} = 1$. In that case the result is true, but tautologically.] □

Thus, if $E(X^2)$ and $E(Y^2)$ are finite, then $E(XY)$ is finite as well. In that case we can define the *covariance* between X and Y to be

$$\mathrm{Cov}(X, Y) = E\left[(X - EX)(Y - EY)\right]. \tag{12}$$

Because $(X - EX)(Y - EY) = XY - XEY - YEX + EXEY$, we obtain the following, which is the computationally useful formula for covariance:

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y). \tag{13}$$

Note, in particular, that $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

Theorem 15.4. *Suppose* $E(X^2)$ *and* $E(Y^2)$ *are finite. Then, for all nonrandom* a, b, c, d*:*

(1) $\mathrm{Cov}(aX + b\,, cY + d) = ac\mathrm{Cov}(X, Y)$;
(2) $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$.

Proof. Let $\mu = EX$ and $\nu = EY$ for brevity. We then have

$$\begin{aligned}\mathrm{Cov}(aX + b\,, cY + d) &= E\left[(aX + b - (a\mu + b))(cY + d - (c\nu + d))\right]\\ &= E\left[(a(X - \mu))(c(Y - \nu))\right]\\ &= ac\mathrm{Cov}(X, Y).\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathrm{E}\left[(X+Y-(\mu-\nu))^2\right] \\ &= \mathrm{E}\left[(X-\mu)^2\right] + \mathrm{E}\left[(Y-\nu)^2\right] + 2\mathrm{E}\left[(X-\mu)(Y-\nu)\right].\end{aligned}$$

Now identify the terms. $\square$