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*

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

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

$$\mathrm{E}(aX + bY) = a\mathrm{E}X + b\mathrm{E}Y.$$

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

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

which is $a\mathrm{E}X + b\mathrm{E}Y$. □

2. Covariance and correlation

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

$$|\mathrm{E}(XY)| \le \sqrt{\mathrm{E}(X^2)\ \mathrm{E}(Y^2)}.$$

Proof. Note that

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

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

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

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

$$\mathrm{E}(X^2)\mathrm{E}(Y^2) \geq \left(\mathrm{E}(XY)\right)^2.$$

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

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

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

Because $(X - \mathrm{E}X)(Y - \mathrm{E}Y) = XY - X\mathrm{E}Y - Y\mathrm{E}X + \mathrm{E}X\mathrm{E}Y$, we obtain the following, which is the computationally useful formula for covariance:

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

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

Theorem 15.4. *Suppose* $\mathrm{E}(X^2)$ *and* $\mathrm{E}(Y^2)$ *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 $\mu = \mathrm{E}X$ and $\nu = \mathrm{E}Y$ for brevity. We then have

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

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. □