STAT 234 Lecture 10B
Expected Values, Covariance,and Correlation
Section 5.2

Yibi Huang
Department of Statistics
University of Chicago

Expected Values of Functions of X & Y

For two random variable X, Y with

- a joint pmf $p(x, y)$, or
- a joint cdf $f(x, y)$,

the expected value of a function $g(X, Y)$ of X and Y is defined as

$$\mathrm{E}[g(X, Y)] = \begin{cases} \sum_{xy} g(x, y)p(x, y) & \text{for discrete case,} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f(x, y)\, dx\, dy & \text{in continuous case.} \end{cases}$$

Example (Gas Station)

Recall the joint pmf for the Gas Station Example in L09 is the table on the right. Suppose we are interested in

$p(x,y)$		Y (full-service) 0	1	2
X	0	0.10	0.04	0.02
self-	1	0.08	0.20	0.06
service	2	0.06	0.14	0.30

$$g(X,Y) = |X - Y|$$
$$= \text{the absolute diff in the \# of hoses in use}$$
$$\text{of the self-service and full-service islands.}$$

The expected value is

$$\begin{aligned} \mathrm{E}\,|X - Y| &= \sum\nolimits_{xy} |x - y| p(x,y) \\ &= |0-0| \cdot 0.10 + |0-1| \cdot 0.04 + |0-2| \cdot 0.02 \\ &\quad + |1-0| \cdot 0.08 + |1-1| \cdot 0.20 + |1-2| \cdot 0.06 \\ &\quad + |2-0| \cdot 0.06 + |2-1| \cdot 0.14 + |2-2| \cdot 0.30 \\ &= 0.48 \end{aligned}$$

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

If $g(X, Y) = aX + bY$ for two random variables X and Y and two constants a and b, we have

$$\mathrm{E}[g(X, Y)] = \mathrm{E}(aX + bY) = a\,\mathrm{E}(X) + b\,\mathrm{E}(Y)$$

no matter X and Y are both discrete, both continuous, or one discrete and one continuous.

Proof. We will prove it for the case when X and Y are continuous with joint pdf $f(x, y)$. The proof for the discrete case is similar. By definition, the expected value of the function $g(X, Y) = aX + bY$ of X and Y is

$$\begin{aligned} \mathrm{E}(aX + bY) &= \iint (ax + by) f(x, y)\mathrm{d}x\mathrm{d}y \\ &= \underbrace{\iint ax f(x, y)\mathrm{d}x\mathrm{d}y}_{\text{Part I}} + \underbrace{\iint by f(x, y)\mathrm{d}x\mathrm{d}y}_{\text{Part II}} \end{aligned}$$

For Part I, we first integrate over y, and then over x.

$$\text{Part I} = \iint axf(x,y)\mathrm{d}x\mathrm{d}y = a\int\left(\int xf(x,y)dy\right)\mathrm{d}x$$

$$= a\int x\underbrace{\int f(x,y)dy}_{f_X(x)}\mathrm{d}x = a\underbrace{\int xf_X(x)\mathrm{d}x}_{\mathrm{E}(X)} = a\,\mathrm{E}(X)$$

For Part II, we first integrate over x, and then over y.

$$\text{Part II} = \iint byf(x,y)\mathrm{d}x\mathrm{d}y = b\int\left(\int yf(x,y)dx\right)\mathrm{d}y$$

$$= b\int y\underbrace{\int f(x,y)dx}_{f_Y(y)}\mathrm{d}y = b\underbrace{\int yf_Y(y)\mathrm{d}y}_{\mathrm{E}(Y)} = b\,\mathrm{E}(Y)$$

Putting Parts I & II together, we get

$$\mathrm{E}(aX + bY) = \mathrm{E}(aX) + \mathrm{E}(bY).$$

Expected Value for Linear Combination of Random Variables

The result $\mathrm{E}(aX + bY) = a\,\mathrm{E}(X) + b\,\mathrm{E}(Y)$ can be generalized to linear combinations of several random variables

$$\mathrm{E}(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1\,\mathrm{E}(X_1) + a_2\,\mathrm{E}(X_2) + \cdots + a_n\,\mathrm{E}(X_n),$$

no matter the rv's are discrete or continuous, independent or not.

$\mathrm{E}[g(X)h(Y)] = \mathrm{E}[g(X)]\,\mathrm{E}[h(Y)]$ if X & Y are independent

When X and Y are **independent**, for any functions g and h,

$$\mathrm{E}[g(X)h(Y)] = \mathrm{E}[g(X)]\,\mathrm{E}[h(Y)].$$

In particular, $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$.

Proof. We prove the discrete case. The continuous case is similar. Using that $p(x,y) = p_X(x)p_Y(y)$ when X, Y are indep, one has

$$\begin{aligned}
\mathrm{E}[g(X)h(Y)] &= \sum\nolimits_{xy} g(x)h(y)p(x,y) \\
&= \sum\nolimits_x \sum\nolimits_y g(x)h(y)p_X(x)p_Y(y) \quad (p(x,y) = p_X(x)p_Y(y) \text{ by indep.}) \\
&= \underbrace{\sum\nolimits_x g(x)p_X(x)}_{\mathrm{E}[g(X)]} \underbrace{\sum\nolimits_y h(y)p_Y(y)}_{\mathrm{E}[h(Y)]} = \mathrm{E}[g(X)]\,\mathrm{E}[h(Y)]
\end{aligned}$$

Covariance

The **covariance** of X and Y, denoted as $\text{Cov}(X, Y)$ or σ_{XY}, is defined as

$$\text{Cov}(X, Y) = \sigma_{XY} = \text{E}[(X - \mu_X)(Y - \mu_Y)],$$

in which $\mu_X = \text{E}(X)$, $\mu_Y = \text{E}(Y)$

- Covariance is a generalization of variance:

$$\text{Var}(X) = \text{Cov}(X, X) = \text{E}[(X - \mu_X)^2]$$

- Covariance can be positive or negative:
  - $\text{Cov}(X, Y) > 0$ means positive association between X, Y
  - $\text{Cov}(X, Y) < 0$ means negative association between X, Y

Shortcut Formula for Covariance

$$\boxed{\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)}$$

- Like the Shortcut Formula for Variance $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$.
- If X & Y are indep., then $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$, which implies $\mathrm{Cov}(X,Y) = 0$.
- However $\mathrm{Cov}(X,Y) = 0$ does **not** imply the independence of X and Y. In this case, we say X and Y are **uncorrelated.**
- Proof of the shortcut formula:

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