Handout for Lecture 11
Independence and Correlation

ECON 340: Economic Research Methods — Instructor: Div Bhagia

Consider two random variables X and Y.

- The *joint probability* $f(x, y) = Pr(X = x, Y = y)$ represents the likelihood that X equals x and Y equals y.
- The marginal probability of $Y = y$, denoted $f(y)$, is obtained by summing the joint probability $f(x, y) = Pr(X = x, Y = y)$ over all possible values of x.
- The *conditional probability* $f(y|x) = Pr(Y = y|X = x)$ represents the likelihood that Y is equal to y, given that X is equal to x.

$$f(y|x) = Pr(Y = y|X = x) = \frac{Pr(X = x, Y = y)}{Pr(X = x)} = \frac{f(x, y)}{f(x)}$$

- The *conditional expectation* $E(Y|x)$ is the expected value of Y given that $X = x$.

$$E(Y|X = x) = \sum_y yPr(Y = y|X = x) = \sum_y yf(y|x)$$

<u>Uncorrelated vs Independent</u>

- Covariance and correlation:

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

$$\rho_{XY} = corr(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \quad \text{where } -1 \leq \rho \leq 1$$

- Two random variables are *uncorrelated* if $corr(X, Y) = 0$.
- Two random variables are *independent* if $f(y|x) = f(y)$ for all x and y or equivalently $E(Y|X) = E(Y)$.
- If X and Y are independent, then they are also uncorrelated. (The converse is not necessarily true.)

1. The following table gives the joint-probability distribution of rain (X) *and* commute time in minutes (Y).

	Rain $(X = 1)$	No Rain $(X = 0)$	Total
60-min commute ($Y = 60$)	0.3	0.2	**0.5**
30-min commute ($Y = 30$)	0.1	0.4	**0.5**
Total	**0.4**	**0.6**	**1**

   (a) Fill the marginal probabilities in the column and row labeled *Total*. For example, the first row in column *Total* should contain the marginal probability of having a 60-minute commute ($Pr(Y = 60)$).

   (b) Find the following conditional probabilities:

   - Probability of having a 60-min commute conditional on *raining*

$$Pr(Y = 60|X = 1) = \frac{Pr(Y = 60, X = 1)}{Pr(X = 1)} = \frac{0.3}{0.4} = \frac{3}{4}$$

   - Probability of having a 60-min commute conditional on *not raining*

$$Pr(Y = 60|X = 0) = \frac{Pr(Y = 60, X = 0)}{Pr(X = 0)} = \frac{0.2}{0.6} = \frac{1}{3}$$

   (c) Calculate $E(Y|X = 1)$ and $E(Y|X = 0)$.

$$\begin{aligned} E(Y|X = 1) &= \sum_y y Pr(Y = y|X = 1) \\ &= 60 \cdot Pr(Y = 60|X = 1) + 30 \cdot Pr(Y = 30|X = 1) \\ &= 60 \cdot \frac{3}{4} + 30 \cdot \frac{1}{4} = 52.5 \end{aligned}$$

$$
\begin{aligned}
E(Y|X=0) &= \sum_y yPr(Y=y|X=0) \\
&= 60 \cdot Pr(Y=60|X=0) + 30 \cdot Pr(Y=30|X=0) \\
&= 60 \cdot \frac{1}{3} + 30 \cdot \frac{2}{3} = 40
\end{aligned}
$$

(d) How does rain impact the expected commute time in this example?

Rain increases the expected commute time by 12.5 minutes.

2. You flipped a coin six times and got tails each time. The likelihood of getting a head in your seventh flip is

   ☒ 1/2

   ☐ More than 1/2

   ☐ Less than 1/2

3. Note that for sums of two random variables X and Y:

$$E(aX + bY) = aE(X) + bE(Y)$$

$$Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$$

In the above expressions, a and b are constants.

You are contemplating investing in two stocks with the same average return and spread.

$$\mu_X = \mu_Y = \mu \qquad \sigma_X^2 = \sigma_Y^2 = \sigma^2$$

Should you pick any one stock at random or invest equally in both?

Consider a new random variable W that represents the return from investing equally in both stocks:

$$W = 0.5X + 0.5Y.$$

Using the formula above

$$E(W) = 0.5E(X) + 0.5E(Y) = 0.5\mu + 0.5\mu = \mu.$$

The expected return from investing equally in both stocks is the same as the return from investing in individual stocks.

Using the formula for the variance of the sum of two random variables, we have

$$\begin{aligned}
\text{Var}(W) &= 0.5^2\text{Var}(X) + 0.5^2\text{Var}(Y) + 2 \times 0.5 \times 0.5 \times \text{Cov}(X, Y) \\
&= 0.25\sigma^2 + 0.25\sigma^2 + 0.5\text{Cov}(X, Y) \\
&= 0.5\sigma^2 + 0.5\text{Cov}(X, Y).
\end{aligned}$$

If X and Y are uncorrelated, then $\text{Cov}(X, Y) = 0$, and $\text{Var}(W) = 0.5\sigma^2$, which is lower than the variance of individual stocks. In this case, it would be better to invest equally in both stocks to minimize risk.

If X and Y are negatively correlated, i.e., $\text{Cov}(X, Y) < 0$, then $\text{Var}(W)$ would be even lower than $0.5\sigma^2$, making the portfolio considerably less risky. Therefore, if the stocks are negatively correlated, diversifying by investing equally in both would be an even more advantageous strategy to minimize risk.

The only case where investing equally in both stocks would not offer a risk-reduction advantage is when the stocks are perfectly positively correlated. Specifically, if $\text{Corr}(X, Y) = 1$, then $\text{Cov}(X, Y) = \sigma^2$ and the variance of W would equal σ^2, the same as the variance for individual stocks. In such a scenario, diversifying by investing in both stocks would offer no risk-minimizing benefits, making us indifferent between the two investment options.