Handout 2: Covariance & Correlation

EC 282: Introduction to Econometrics

Spring 2026

1 Setup

A career center surveys recent graduates and classifies them by two random variables:

- X = number of internships completed during college (1, 2, or 3)
- Y = starting salary in thousands of dollars (60, 80, or 100)

The **joint distribution** of X and Y is given below.

	$Y = 60$	$Y = 80$	$Y = 100$	Marginal of X
$X = 1$	0.15	0.10	0.05	
$X = 2$	0.05	0.25	0.10	
$X = 3$	0.05	0.05	0.20	
Marginal of Y				

2 Marginal Distributions

Question 2.1: Compute the marginal distributions of X and Y by filling in the row and column totals in the table above.

3 Expected Values

Question 2.2: Compute $E[X]$ and $E[Y]$ using the marginal distributions:

$$E[X] = \mu_X = \sum_i x_i \cdot \Pr(X = x_i) \qquad E[Y] = \mu_Y = \sum_j y_j \cdot \Pr(Y = y_j)$$

Question 2.3: Interpret $E[X]$ and $E[Y]$ in plain language. What do these numbers tell us about the typical graduate?

4 Variance and Standard Deviation

Question 2.4: Compute the variance of X using the definition:

$$\sigma_X^2 = E\big[(X - \mu_X)^2\big] = \sum_i (x_i - \mu_X)^2 \cdot \Pr(X = x_i)$$

Question 2.5: Similarly, compute σ_Y^2, and then the standard deviations σ_X and σ_Y.

5 Covariance

Question 2.6: Compute the covariance using the definition:

$$\sigma_{XY} = E\big[(X - \mu_X)(Y - \mu_Y)\big] = \sum_i \sum_j (x_i - \mu_X)(y_j - \mu_Y) \cdot \Pr(X = x_i,\ Y = y_j)$$

Hint: You need to evaluate the expression for all 9 cells in the joint probability table.

Question 2.7: Interpret the sign of the covariance. Does the result make intuitive sense? What does it tell you about the relationship between internship experience and starting salary?

Question 2.8: Suppose salary were measured in **dollars** instead of thousands (i.e., multiply each Y value by 1,000). How would σ_{XY} change? What does this tell you about using covariance to compare the strength of relationships?

6 Correlation

Question 2.9: Compute the correlation coefficient:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Question 2.10: Interpret ρ_{XY}. What does its value (between -1 and $+1$) tell you about the *strength* and *direction* of the linear relationship?

Question 2.11: Would ρ_{XY} change if salary were measured in dollars instead of thousands? Why is this property important?

7 Excel Verification

Question 2.12: Verify your calculations by entering the joint probability table in Excel and using `SUMPRODUCT()` to compute each quantity. Do your manual calculations match?

8 Thinking Deeper

Question 2.13: We found a positive correlation between internships and starting salary. Can we conclude that doing more internships *causes* a higher starting salary? What other factors might explain this relationship?

Question 2.14: A news article reports: "People who eat breakfast every day earn 20% more than people who skip breakfast." What is wrong with concluding that eating breakfast causes higher earnings?