Assignment 1

Samuel Fraley
Dmitrii Kuptsov
Felipe Manzi

September 30, 2025

Contents

Question 1

(a) Describe, in one precise sentence, what term (1) captures.

Term 1, $E(y_i \mid d_i = 1)$, captures the expected value of of y_i given $d_i = 1$, specifically the expected monthly pay of a person given they have a college degree ($d_i = 1$)

(b) What sign would you expect for (1) – (2)? Briefly justify.

Term 1 is the expected monthly pay of individuals with a college degree, while Term 2 is the expected monthly pay of those without a college degree. Generally we would expect those with a college degree to, on average, have a higher monthly pay, so Term 1 would be larger than Term 2. As a result, we would expect (1) - (2) to be positive.

(c) Describe what does term (3) capture. Be specific.

Term 3 describes the expected difference in monthly pay between going and not going to college given a person did go to college. y_{i1} - y_{i0} calculates the difference in monthly wage, while $d_i = 1$ is the condition that they did go to college. This is a casual parameter that we defined as the Average Treatment Effect for the Treated (ATT).

(d) Consider we are given a sample of n pairs of observations (yi, di) (observational data). Explain how the sample can be used to provide information for term (1) - (2) but not for term (3) nor (4). Justify.

Term 3 and Term 4 both focus on counterfactuals that we cannot observe. We cannot observe y_{i0} for those given $d_i = 1$, or y_{i1} for those given $d_i = 0$. We know that Term 3 requires the wages given attended college (y_{i0}) and not attended college (y_{i1}) for the treatment group ($d_i = 1$), and Term 4 requires y_{i0} for both those given thehy did ($d_i = 1$) and did not ($d_i = 0$) attend college. In short, we cannot observe the counterfactual, what the wage would have been if those who attended college did not as well as if those who did not attend college actually did.

(e) What would it mean for term (4) to be positive?

Term 4 is the difference in the counterfactual wage (no college) between those that did and did not attend college. If it was positive, that would mean $E(y_{i0} \mid d_i = 1)$ is greater than $E(y_{i0} \mid d_i = 1)$, so the expected monthly wage without college is greater for those that did attend college compared to those that did not. It would mean that those that did attend college would have had a higher wage even without college compared to those that did not attend college. This captures systemic differences between the two populations that would impact monthly wage, such as ability.

(f) What are the implications of term (4) being positive for measuring the effect of a degree on earnings?

If Term (4) is positive, it means that individuals who attended college would still earn more, on average, than those who did not, even in the absence of a degree. As a result, the observed wage gap between graduates and non-graduates would reflect not only the effect of college itself, but also pre-existing differences between the two groups, making the observed effect larger than the true causal effect of a degree.

Question 2

Discrete random variables X and Y can take 10 equally likely pairs:

$$(1,2),(1,4),(1,6),(2,1),(2,3),(2,5),(3,2),(3,4),(3,6),(3,8).$$

Verify the Law of Total Expectations: $E[E(Y|X)] = E(Y)$. **Answer:** ...

Question 3

Variable $X \sim N(0,1)$ and $Y = X^2 - 1$. Show how this example illustrates that uncorrelated $\not\Rightarrow$ independent. **Answer:** ...

Question 4

Question 4

Joint pdf: $f(x,y) = \dfrac{3(x^2+y)}{11}$ for $0 \le x \le 2$, $0 \le y \le 1$. (For consistency with Appendix 2 notation, set $x_2 \equiv x$.)

We want the best linear approximation to the conditional expectation function (CEF),

$$\ell(x) = \beta_1 + \beta_2 x.$$

The population least squares problem

$$\min_{\beta} \; E\big[(y - x'\beta)^2\big],$$

where

$$x = \begin{bmatrix} 1 \\ x_2 \end{bmatrix}, \qquad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}.$$

Normal equations:

$$E(xx')\,\beta = E(xy).$$

Rewrite:

$$\beta \equiv [E(xx')]^{-1} E(xy).$$

Working the algebra:

$$\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} = [E(xx')]^{-1}E(xy) = \left(E\Big(\begin{bmatrix}1\\ x_2\end{bmatrix}\begin{bmatrix}1 & x_2\end{bmatrix}\Big)\right)^{-1} E\begin{bmatrix}1\\ x_2\end{bmatrix} y = \begin{bmatrix}1 & E(x_2)\\ E(x_2) & E(x_2^2)\end{bmatrix}^{-1}\begin{bmatrix}E(y)\\ E(x_2y)\end{bmatrix}.$$

Using the 2×2 inverse formula,

$$\begin{bmatrix}1 & E(x_2)\\ E(x_2) & E(x_2^2)\end{bmatrix}^{-1} = \frac{1}{E(x_2^2)-[E(x_2)]^2}\begin{bmatrix}E(x_2^2) & -E(x_2)\\ -E(x_2) & 1\end{bmatrix} = \frac{1}{\mathrm{var}(x_2)}\begin{bmatrix}E(x_2^2) & -E(x_2)\\ -E(x_2) & 1\end{bmatrix}.$$

Hence,

$$\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} = \frac{1}{\mathrm{var}(x_2)}\begin{bmatrix}E(x_2^2) & -E(x_2)\\ -E(x_2) & 1\end{bmatrix}\begin{bmatrix}E(y)\\ E(x_2y)\end{bmatrix} = \begin{bmatrix}\dfrac{E(x_2^2)E(y)-E(x_2)E(x_2y)}{\mathrm{var}(x_2)}\\ \dfrac{E(x_2y)-E(x_2)E(y)}{\mathrm{var}(x_2)}\end{bmatrix}.$$

Identify the slope:

$$\beta_2 = \frac{E(x_2y)-E(x_2)E(y)}{\mathrm{var}(x_2)} = \frac{\mathrm{cov}(x_2,y)}{\mathrm{var}(x_2)}.$$

Regarding the intercept (first element), notice

$$\frac{E(x_2^2)E(y)-E(x_2)E(x_2y)}{E(x_2^2)-[E(x_2)]^2} = E(y) - \frac{E(x_2y)-E(x_2)E(y)}{E(x_2^2)-[E(x_2)]^2}E(x_2) = E(y)-\beta_2E(x_2).$$

Therefore,

$$\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} = \begin{bmatrix}E(y)-\beta_2E(x_2)\\ \dfrac{\mathrm{cov}(x_2,y)}{\mathrm{var}(x_2)}\end{bmatrix}.$$

Plugging in the given moments (with $x_2 \equiv x$):

$$E(x_2)=\tfrac{15}{11},\quad \mathrm{var}(x_2)=\tfrac{151}{605},\quad E(y)=\tfrac{6}{11},\quad \mathrm{cov}(x_2,y)=-\tfrac{2}{121}.$$

For the slope:

$$\beta_2 = \frac{\mathrm{cov}(x_2,y)}{\mathrm{var}(x_2)} = \frac{-\frac{2}{121}}{\frac{151}{605}} = \frac{-2}{121}\cdot\frac{605}{151} = \frac{-1210}{18271} = -\tfrac{10}{151}.$$

For the intercept:

$$\beta_1 = E(y)-\beta_2E(x_2) = \tfrac{6}{11} - \Big(-\tfrac{10}{151}\Big)\tfrac{15}{11}.$$

Compute the second term:

$$-\beta_2E(x_2) = \frac{10}{151}\cdot\frac{15}{11} = \frac{150}{1661}.$$

So,

$$\beta_1 = \frac{6}{11}+\frac{150}{1661}.$$

Writing with common denominator 1661:

$$\frac{6}{11} = \frac{906}{1661},$$

hence

$$\beta_1 = \frac{906}{1661} + \frac{150}{1661} = \frac{1056}{1661} = \tfrac{96}{151}.$$

Therefore,

$$\beta_2 = -\tfrac{10}{151}, \qquad \beta_1 = \tfrac{96}{151}.$$

Final answer:

$$\ell(x) = \beta_1 + \beta_2 x = \frac{96}{151} - \frac{10}{151}\, x.$$

Question 5

Simple regression model: $y = \beta_1 + \beta_2 x^2 + \epsilon$.

(a) Prove β_1, β_2 solve the least squares problem.

(b) Reverse regression $x^2 = \alpha_1 + \alpha_2 y + v$: write expressions for α_1, α_2.

(c) When does $\alpha_2 = 1/\beta_2$? Justify.