MIT 14.32/14.320 **Spring 2023**
J. Angrist

Pop Quiz #1

Note: Please type answers directly in GradeScope. Use _ for subscripts, as in X_i for X_i. Use ^ for superscripts as in s^2 for s^2.

1. Multiple Choice

   (a) Suppose that X_i is drawn from a distribution with mean μ_X and variance σ^2_X. Let $\bar{X} = \frac{\sum X_i}{n}$ denote the sample mean and let s^2_X denote the sample variance. We reject the null hypothesis that $\mu_X = 0$ at the 5% significance level if:

   i. $\frac{\bar{X}}{s^2_X}$ is less than about 2 in absolute value
   ii. $\frac{\bar{X}}{\sigma_X}$ exceeds about 2 in absolute value
   iii. $\frac{\sqrt{n}\bar{X}}{s_X}$ exceeds about 2 in absolute value
   iv. $\frac{n\bar{X}}{s_X}$ is less than about 2 in absolute value

   $t = \frac{\bar{X} - \mu_X}{SE}$

   SE

   $SE = \frac{S_x}{\sqrt{n}}$

   (b) The Finkelstein, et al (2012) study on our reading list reports that randomly-assigned Medicaid eligibility (mark all that apply)

   i. increases health insurance coverage ✓
   ii. improves self-reported health ✓
   iii. increases emergency department visits
   iv. reduces medical debt

   (c) The Aron-Dine, Einav, and Finkelstein (2013) retrospective article on the RAND health insurance experiment focuses on how health insurance offered through the RAND experiment affects experimental subjects' health.

   i. True
   ii. False

2. What are two ways to write $C(X_i, Y_i)$, the covariance between random variables X_i and Y_i

$$E[(X_i - \mu_x)(Y_i - \mu_y)] = E[X_i Y_i] - \mu_x \mu_y = E[X_i(Y_i - \mu_y)] = E[Y_i(X_i - \mu_x)]$$

3. Suppose there are 600 Course 6 majors at MIT, 100 of whom do 6-14 and the rest do something else. Data from an MIT career services survey done one year after graduation show that 80 percent of 6-14 majors are happy with their jobs, while only 40% of other Course 6 majors are happy with their jobs. The same survey shows that half of Course 14 majors are happy with their jobs. Use the law of iterated expectations (showing your work) to determine for which major, Course 6 or Course 14, the prospects of happiness are higher (note: for purposes of this problem, 6-14 majors are considered part of Course 6, not Course 14).

$$E(h_6) = E(h_{614}) \cdot \frac{100}{600} + E(h_{6,\text{not } 614}) \cdot \frac{500}{600}$$

$$= 0.8 \cdot \frac{100}{600} + 0.4 \cdot \frac{500}{600} = \frac{80}{600} + \frac{200}{600} = \frac{280}{600} < \frac{1}{2} = E(h_{14})$$

Bivariate regression

① assume a linear CEF $E[Y_i|X_i] = \alpha + \beta X_i$

② define CEF residual $\varepsilon_i = Y_i - E[Y_i|X_i]$

$\Rightarrow$ $E[\varepsilon_i|X_i] = 0$ $\qquad \varepsilon_i = Y_i - \alpha - \beta X_i$

$$E[Y_i - E[Y_i|X_i]\,|X_i] = E[Y_i|X_i] - E[Y_i|X_i] = 0$$

(1) $\Rightarrow$ $\boxed{E[\varepsilon_i] = 0}$ $\qquad E[\varepsilon_i] = E[\underbrace{E(\varepsilon_i|X_i)}_{=0}] = 0$

(2) $\Rightarrow$ $\boxed{E[X_i\varepsilon_i] = 0}$ $\qquad E[X_i\varepsilon_i] = E[E[X_i\varepsilon_i|X_i]] =$

$$= E[X_i \underbrace{E[\varepsilon_i|X_i]}_{=0}] = 0$$

Use (1) & (2) + Linearity CEF

$E(\varepsilon_i) = E[Y_i - E[Y_i|X_i]]$ by def

$= E[Y_i - \alpha - \beta X_i]$ by linearity of CEF

$$E[Y_i - \alpha - \beta X_i] = 0$$

$$\boxed{E[Y_i] - \beta E[X_i] = \alpha}$$

$$E[X_i\varepsilon_i] = 0$$

$$E[X_i(Y_i - \alpha - \beta X_i)] = 0$$

$$E[X_iY_i] - \alpha E[X_i] - \beta E[X_i^2] = 0$$

$$E[X_iY_i] - (E[Y_i] - \beta E[X_i])E[X_i] - \beta E[X_i^2] = 0$$

$$\underbrace{E[X_iY_i] - E[X_i]E[Y_i]}_{cov(X_iY_i)} = \beta\underbrace{(E[X_i^2] - (E[X_i])^2)}_{var(x_i)}$$

$$\boxed{\beta = \frac{cov(X_i, Y_i)}{var(X_i)}}$$