1. (15 points) CONDITIONAL EXPECTATION

Consider the following three random variables: Y, D, and X, where

$$Y = \begin{cases} Y_1 & \text{if } D = 1 \\ Y_0 & \text{if } D = 0. \end{cases}$$

D is a Bernoulli random variable taking value either 1 or 0; equivalently $D = \mathbb{I}_{\{Y_1 > Y_0\}}$, where the indicator function $\mathbb{I}_{\{\cdot\}} = 1$ if $Y_1 > Y_0$, $\mathbb{I}_{\{\cdot\}} = 0$ otherwise.

(a) Write $\mathbb{E}[Y|D=1, X] - \mathbb{E}[Y|D=0, X]$ in terms of Y_1, Y_0, and X.

(b) Is $\mathbb{E}[Y|D=1, X] - \mathbb{E}[Y|D=0, X]$ equal to $\mathbb{E}[Y_1|X] - \mathbb{E}[Y_0|X]$? Why or why not?

(c) Show that

$$\mathbb{E}\left[\frac{YD}{\mathbb{E}[D=1|X]}\right] = \mathbb{E}[Y_1].$$

2. (10 points) PROBABILITY DENSITY FUNCTION

Let $X_1, \ldots, X_n$ be continuous iid random variables with probability density function $f(x)$ (cumulative distribution function $F(x)$). Find the probability density functions of

(a) $Y_1 = \min(X_1, \ldots, X_n)$

(b) $Y_n = \max(X_1, \ldots, X_n)$.

3. (15 points) JOINT CONFIDENCE REGION

Consider the following bivariate normal random variable:

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right),$$

where ρ lies between -1 and 1.

(a) Propose a test statistic to construct the joint confidence region for $H_0 : z_1 = z_2 = 0$.

(b) Given part (a) and $\rho = 0$, plot the corresponding joint confidence region.

(c) Given part (a) and $\rho = 0.6$, plot the corresponding joint confidence region.

4. (10 points) HYPOTHESIS TEST FOR DIFFERENCE OF MEANS

The STAR (Student-Teacher Achievement Ratio) project conducted a randomized controlled trial in the 1980s. In the experiment, students were randomly assigned to a small class, regular class, or regular class with an aid. This particular study lasted from 1985 to 1989 and involved 11,601 students. We are going to investigate whether the small class size improved educational performance (score) or not. We estimate the average scores for small-class and regular-sized-class groups:

見背面

- small classes: $\hat{\mu}_s = 723$ (1.9)
- regular-sized classes: $\hat{\mu}_r = 719$ (1.8)

where numbers in parentheses are standard errors.

(a) Relying on the central limit theorem, compute the 95% confidence interval for the estimated $(\mu_s - \mu_r)$.

(b) Construct the 95% joint confidence region for μ_s and μ_r.

5. (15 points)

The model is $y_i = z_i'\beta + e_i$ and $E(x_i e_i) = 0$. An economist wants to obtain the two-stage least square estimator (2SLS) estimates and standard errors for β. He uses the following steps:

(i) Regresses z_i on x_i, obtains the fitted values $\widehat{z_i}$.

(ii) Regresses y_i on $\widehat{z_i}$, obtains the coefficient estimate $\widehat{\beta}$ and standard error $s(\widehat{\beta})$ from this regression.

Is this correct? Does this produce the 2SLS estimates and standard errors? Why or why not.

6. (15 points)

Let e_t be a sequence of independent identically distributed random variables with mean zero and variance one. Define a stochastic process by $x_t = e_t - (1/3)e_{t-1} + (1/3)e_{t-2}$, $t = 1, 2, \ldots$.

(a) Find $E(x_t)$ and $Var(x_t)$.

(b) Find $Corr(x_t, x_{t+h})$ for $h = 1, 2, 3$.

(c) Is x_t a covariance stationary process? Is x_t an asymptotically uncorrelated process?

7. (20 points)

Let $E(y|x,q) = x\beta + \gamma q$, thus $y = x\beta + \gamma q + v$, where $E(v|x,q) = 0$.

(a) Suppose that we can observe y and both x and q, would OLS estimators of y on x and q provide an unbiased estimators of β? Why or why not? Would OLS estimators of y on x and q provide an best linear unbiased estimator (BLUE) of β? Why or why not?

(b) Suppose we do not observe q, under what conditions, would OLS estimators of y on x provide an unbiased estimators of β? Why ?

(c) Suppose there is a variable z where $E(q|z) = \delta z$. Suppose also that $E(y|x,q,z) = E(y|x,q)$, would OLS estimators of y on x and z provide an unbiased estimators of β? Why or why not?