End-Term Exam (1/2)

Basic Econometrics
Fall 2018
Tadao Hoshino

Name. ______________________

ID. ______________________

Section $\boxed{1}$

1. Let X be a continuous uniform random variable on $[-2, 4]$.

1. Calculate the probability $\Pr(1 \le X \le 2)$.
2. Calculate the probability $\Pr(-4 \le X \le 1)$.
3. Calculate the expectation $E(X)$.
4. Calculate the variance $V(X)$.

2. Let S be a dummy random variable with $\Pr(S = 1) = 0.7$ and $\Pr(S = 0) = 0.3$, and let X and Y be normally distributed as $N(0, 3)$ and $N(0, 9)$, respectively. Further, let $Z = SX + (1-S)Y$. X, Y, and S are assumed to be independent.

1. Calculate the expected value of $(XY)^2$, $E((XY)^2)$.
2. Calculate the expected value of Z, $E(Z)$.
3. Calculate the variance of Z, $V(Z)$. (*Hint*: As S is a dummy variable, $S^2 = S$.)

Section $\boxed{2}$

Decide whether the following statements are True or False.

1. For random events A and B, if they are independent, the joint probability $\Pr(A, B)$ is equal to the marginal probability $\Pr(A)$.

   ☐ True ☐ False

2. For random variables X and Y, if they are uncorrelated, $E(Y|X = x_1) = E(Y|X = x_2)$ holds for any (x_1, x_2) in the support of X.

   ☐ True ☐ False

3. Consistency implies unbiasedness.

   ☐ True ☐ False

4. Suppose that there are two consistent estimators $\hat{\theta}_1$ and $\hat{\theta}_2$. If the variance of $\hat{\theta}_1$ is larger than that of $\hat{\theta}_2$, $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$.

   ☐ True ☐ False

5. Let Z be distributed as the standard normal $N(0, 1)$. Then, $\Pr(|Z| \le 2) > 0.95$.

   ☐ True ☐ False

6. Reducing the probability of Type I error increases the probability of Type II error.

   ☐ True ☐ False

7. Consider the following simple regression model:

   $$Y = \beta_0 + X\beta_1 + \epsilon,$$

   where Y is a dependent variable, X is an explanatory variable, and ϵ is an error term with $E(\epsilon) = 0$. Assume that ϵ is independent of X.

   - X is an exogenous explanatory variable.

     ☐ True ☐ False

   - If $\beta_1 = 0$, X does not affect Y.

     ☐ True ☐ False

8. Let Y be a dependent variable, and X be an endogenous explanatory variable. When a random variable Z is correlated with Y, Z is not a valid instrument for X.

   ☐ True ☐ False

9. Let Y be an outcome variable, T be a binary treatment variable, and X be a set of covariates. Further, let Y_1 and Y_0 be the potential value of Y when $T = 1$ and when $T = 0$, respectively. When the data are collected in a randomized experiment, it holds that

   $$(Y_1, Y_0) \perp\!\!\!\perp T \mid X,$$

   i.e., (Y_1, Y_0) and T are independent conditional on X.

   ☐ True ☐ False

End-Term Exam (2/2)

Basic Econometrics
Fall 2018
Tadao Hoshino

Name. ____________________

ID. ____________________

Section $\boxed{3}$

1. Consider the following simple regression model:

$$Y_i = \beta_0 + X_i\beta_1 + \epsilon_i \;\; i = 1, \ldots, n.$$

Derive the OLS estimator for β_1.

2. The following is the regression result on a sample of 1000 apartment houses in a Japanese city (standard errors in parentheses):

$$\widehat{rprice} = \underset{(9.9)}{71.3} + \underset{(0.3)}{1.4} \cdot area - \underset{(0.5)}{1.9} \cdot age + \underset{(0.4)}{0.9} \cdot renov - \underset{(0.1)}{0.3} \cdot AirPollut$$

where $rprice$ is the monthly rental price (1,000JPY), $area$ is the area size (m^2) of the house, age is the age of the apartment (years), $renov$ is a dummy variable $renov = 1$ if the apartment was renovated, and $AirPollut$ is the amount of air pollutant emission from nearby manufacturing plants (measured in some unit, e.g., ppm).

1. Suppose that an apartment owner renovates the apartment. What is the predicted increase in the rental price of a house in the apartment?
2. Suppose that instead of measuring $rprice$ and $area$ in 1,000JPY and m^2, these variables are measured in 100JPY and $10m^2$. What is the regression coefficient estimate of $area$ from this new regression?
3. Is the hypothesis H_0: *renovated and un-renovated apartments have the same economic value* rejected at the 5% significance level? If yes (no), what about the significance level at 1% (10%)?
4. Do you think the variable $AirPollut$ should be treated as endogenous? Why?

Section $\boxed{4}$

Describe (1) the problem of selection bias in the estimation of treatment effects, and (2) how randomized experiments can eliminate the bias. (You can answer either mathematically or in words.)