

Econ 3040 Midterm, Oct. 29, 2020

Total marks: 103

Time: 70 minutes

Upload your answers to the UM Learn dropbox marked “Midterm”

Part A – Short Answer [7 marks each]

1.) Let the probability function for a random variable, Y , be:

$$Y = 1 \text{ with probability } 0.8; Y = 6 \text{ with probability } 0.2$$

- Find the expected value of Y .
- Find the variance of Y .
- Let $Z = 4 \times Y$. What is the expected value and the variance of Z ?
- What is the correlation between Y and Z ?

2.) Explain why the sample average, \bar{y} , and the OLS estimators, b_0 and b_1 , are considered random variables. Why might they follow a Normal distribution?

3.) Derive the variance of the sample average, \bar{y} . Explain how this variance is used to determine that \bar{y} is efficient and consistent. Explain how this variance is used in hypothesis testing.

4.) Explain the difference between the “z test” and the “t test”.

5.) Explain why we should minimize the sum of the squared “vertical distances” (why we should minimize the sum-of-squared residuals), instead of minimizing the sum of absolute vertical distance. That is, why should we use the OLS estimators, b_0 and b_1 in order to estimate the unknown β_0 and β_1 , instead of some other estimator?

6.) Describe a situation where $R^2 = 0$, and where $R^2 = 1$.

7.) Use the following data: $Y = \{0, 2, 5\}$; $X = \{1, -1, -4\}$

- Calculate the OLS estimators b_1 and b_0 .
- Calculate the OLS predicted value and residual, for *only* the 2nd observation in the sample.

8.) Briefly explain why the estimator:

$$s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

is better than the estimator:

$$\hat{\sigma}_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$

9.) Consider a different estimator for the population mean of y (μ_y):

$$\tilde{\mu}_y = \frac{y_1 + y_n}{2}$$

That is, only the first and last value in the sample is used. Assume that y_i has population mean μ_y and population variance σ_y^2 , and that y_i is i.i.d.

- Is $\tilde{\mu}_y$ unbiased? Explain/prove.
- What is the variance of $\tilde{\mu}_y$? Is $\tilde{\mu}_y$ efficient? Explain/prove.
- Is $\tilde{\mu}_y$ consistent? Explain/prove.

Part B – Long Answer [Each part worth 5 marks]

10.) The following question uses the data on the quantity consumed of spirits (hard liquor), and the price of spirits. The estimated demand equation is:

$$\hat{Q} = 12.34 - 0.74 \times P, \quad R^2 = 0.93$$

(0.212) (0.024)

- If price is increased by 1, what is the estimated effect on quantity?
- What does the estimated model predict for quantity consumed, when price equals 5?
- How good is price at explaining variation in the quantity of liquor consumed? What are some other factors that might determine quantity consumed?
- Construct a 95% confidence interval for b_1 . Using only this interval, test the null hypothesis that *price* has no effect on *quantity* consumed.

11.) The following question refers to a regression of *wage* on *female*, where *wage* is measured in dollars per hour, and where *female* is a dummy variable that takes on the value 1 if the worker is female, and 0 if the worker is male. The estimated results are:

$$\widehat{wage} = 11.24 - 0.85 \times female, \quad R^2 = 0.068$$

(1.23) (0.56)

- What is the estimated difference between the wages of men and women?
- What is the sample mean wage for males and the sample mean wage for females?
- Test the null hypothesis that there is no difference in the wages of men and women. Use a p -value for the conclusion of your test.
- Another researcher uses the same data set, but instead estimates the model:

$$wage = \beta_0 + \beta_1 \times male + \epsilon,$$

where *male* is a dummy variable defined in the *opposite* way as *female* ($male = 1$ if the worker is male, and $male = 0$ otherwise). What will be the estimated values b_0 and b_1 for this new model?

END

