Rodrigo Pinto
Department of Economics
UCLA

Economics 103
Introduction to Econometrics
Spring, 2022

Mid-Term Practice (First Example)
Tuesday, May 16, 2023

Instructions:

- This is a 1 hour and 10 minute exam. You are allowed to use the equations sheet provided in the site of the course and handwritten notes.
- You are not allowed to use any electronic device. No connection to the internet via WiFi or any other method is allowed. It is not permitted to use any kind of mobile phone.
- When you are finished with the exam, please **turn in the exam questions**.
- Cheating of any form will result in a score of 0 (zero) for the exam, in addition to the normal university disciplinary action.
- Please sign below that you have read, understood, and fulfilled all of the above instructions and conditions.

Please fill in the following personal information:

First Name	
Last Name	
UCLA ID # **Signature**	
Exam Version A	

Please start solving the examinations only when you are instructed to do so.
Please stop immediately when instructed to do so.

Good Luck!

Part I (Questions based on Regression Output):

Questions 1–3 are based on the following regression output
Consider the following linear model based on the **Cubic** function:

$$wage = \beta_1 + \beta_2 \cdot educ3 + e, \text{ where}$$

1. *wage* means daily wage measured in dollars.
2. *educ* means years of education, measured in years of schooling.
3. $educ3$ means the cubic of education, that is $(educ)^3$.

The R Output for this linear regression is given below:

```
reg wage educ3

      Source |       SS       df       MS              Number of obs =     100
-------------+------------------------------           F(  1,  98  ) =  100.00
       Model |   10000.00     1     10000.00           Prob > F      =  0.0000
    Residual |   10000.00    98       100.00           R-squared     =  0.5000
-------------+------------------------------           Adj R-squared =  0.4950
       Total |   20000.00    99       200.00           Root MSE      =   10.00

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       educ3|   0.05         0.0050      10.00   0.000     0.0400     0.0600
       _cons |   100.00      10.0000      10.00   0.000     80.000     120.000
------------------------------------------------------------------------------
```

Answer the following questions based on the regression output:

Question 1. What is the expected wage for a person with 10 years of education?

(a) 15

(b) 50

(c) 100

(d) 150

(e) 500

Question 2. What is the marginal effect of another year of education for a person with 10 years of education?

(a) 1.5

(b) 3

(c) 5

(d) 10

(e) 15

Question 3. What is the estimated elasticity for a person with 10 years of education at his expected wage?

(a) 0.1

(b) 0.5

(c) 1

(d) 1.5

(e) 2

Questions 4–12 are based on the following regression output
Consider the quadratic model:

$$wage = \beta_1 + \beta_2 \cdot educ + \beta_3 \cdot educ2 + e.$$

1. *wage* means daily wage measured in dollars.
2. *educ* means the education variable measured in schooling years.
3. *educ*2 means the squared of education variable $(educ)^2$.

The regression output for this linear regression is given below:

```
. reg wage educ educ2

      Source |       SS           df       MS      Number of obs   =       100
-------------+----------------------------------   F(2, 97)        =     50.00
       Model |  10000.00         2     5000.00     Prob > F        =    0.0000
    Residual |  10000.00        97      100.00     R-squared       =    0.5000
-------------+----------------------------------   Adj R-squared   =    0.4900
       Total |  20000.00        99      200.00     Root MSE        =    10.000

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    -20.000    50.0000    -0.40   0.686    -120.0000    80.00000
       educ2 |      2.000     2.0000     1.00   0.448      -2.0000     6.00000
       _cons |    200.000   200.0000     1.00   0.422    -200.0000    600.0000
------------------------------------------------------------------------------
```

Question 4. What is the expected wage for a person with 10 years of education?

(a) 10

(b) 20

(c) 100

(d) 200

(e) 2000

Question 5. What is the marginal effect of another year of education for a person with 10 years of education?

(a) 200

(b) 100

(c) 2

(d) 10

(e) 20

Question 6. What is the estimated elasticity for a person with 10 years of education at his expected wage?

(a) 0.1

(b) 0.5

(c) 1

(d) 1.5

(e) 2

Question 7. Parameter λ_0 is defined as a linear combination of the model coefficients:

$$\lambda_0 = \frac{\beta_1}{200} - \frac{\beta_3}{2}.$$

What is the value of the Best Linear Unbiased Predictor (BLUP) for λ_0?

(a) 20

(b) 1

(c) 10

(d) 2

(e) 0

Question 8. The estimated covariance between estimators b_1, b_3 is $\widehat{cov}(b_1, b_3) = 200$. What is the estimated standard error for λ?

(a) -1

(b) 0

(c) 1

(d) 2

(e) 4

Question 9. What is the test statistic for the null hypothesis $H_0 : \lambda_0 = 0$ against the alternative $H_1 : \lambda_0 \neq 0$.

(a) -1

(b) 0

(c) 1

(d) 2

(e) 10

Question 10. Using the critical value of $t_c = 2$, the confidence interval for λ is given by:

(a) $[-1, 0]$

(b) $[0, 1]$

(c) $[0, 2]$

(d) $[-2, 2]$

(e) $[-1, 1]$

Question 11. A econometrician spotted a typo in the p-values of the regression output. Let $p_1 = 0.422$ be the p-value associated with intercept β_1.
Let $p_2 = 0.686$ be the p-value associated with intercept β_2 (for *educ*).
Let $p_3 = 0.448$ be the p-value associated with intercept β_3 (for *educ*2)
Given the estimated coefficients, and its respective standard errors, we should have that:

(a) We should have that $p_2 < p_3$

(b) We should have that $p_1 = p_2$

(c) We should have that $p_2 = p_3$

(d) We should have that $p_2 < p_1$

(e) We should have that $p_1 = p_3$

Part II (Multiple Choice Questions that do not use Regression Output):

Question 12. Let X_1, X_2 be two random variables and take values in $\{0, 1\}$ and whose joint distribution is given by:

Joint Distribution

	$X_1 = 0$	$X_1 = 1$	$f_{X_2}(x)$
$X_2 = 0$	0.2	0.2	0.4
$X_2 = 1$	0.3	0.3	0.6
$f_{X_1}(x)$	0.5	0.5	1

Mark the choice that is correct:

(a) X_1 and X_2 are statistically **independent**.

(b) $E(X_1) = E(X_2)$

(c) $P(X_2 = 1|X_1 = 1) = P(X_2 = 0|X_1 = 0)$

(d) $var(X_1) = var(X_2)$

(e) $Cov(X_1, X_2) \neq 0$

Question 13. Let $X_1 \sim N(1, 1)$ and $X_2 \sim N(2, 4)$ be two normally distributed random variables. Let the correlation between X_1 and X_2 be ρ. Mark the choice that is **correct**:

(a) $Var(c_1 \cdot X_1 + c_2 \cdot X_2) \neq Var(c_0 + c_1 \cdot X_1 + c_2 \cdot X_2)$ whenever $c_0 \neq 0$.

(b) $P((X_1 - 1) > 3) = P((X_2 - 2) > 3)$ *regardless* if X_1, X_2 correlate.

(c) Let $Z_1 = \left(\frac{X_1 - 1}{1}\right)$ and $Z_2 = \left(\frac{X_2 - 2}{2}\right)$, then $E(Z_1^2 + Z_2^2) = 2$ *regardless* if X_1, X_2 correlate.

(d) If X_1 and X_2 correlate, then their covariance may take any value in the real line $[-\infty, \infty]$.

(e) We have that $E\big((X_1 - 1) \cdot (X_2 - 2)\big) = 0$ *regardless* if X_1, X_2 correlate.

Question 14. Consider the following regression model:

$$Y_i = \beta_1 + \beta_2 x_i + \epsilon_i,$$

for $i = 1, ..., N$. Let $e_i \sim \left(0, \sigma^2\right)$. That is, e_i has a distribution whose mean is 0 and its variance is σ_i^2. Let

$$s_x^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2, \quad \text{where } \overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{1}$$

One ran a regression of y on x and obtained the following estimates for β_1 and β_2: $\hat{b}_1 = 4$, $\hat{b}_2 = .5$. Define $x_i^* = 10 \times x_i$. If one were to run a regression of y on x^* the estimates $\hat{b}_1^*$ and $\hat{b}_2^*$ would be

(a) $\hat{b}_1^* = 4$, $\hat{b}_2^* = 5$

(b) $\hat{b}_1^* = 4$, $\hat{b}_2^* = .5$

(c) $\hat{b}_1^* = 4$, $\hat{b}_2^* = .05$

(d) $\hat{b}_1^* = .4$, $\hat{b}_2^* = .5$

(e) $\hat{b}_1^* = .4$, $\hat{b}_2^* = 5$

Question 15. Regarding the Simple Regression Model $y = \beta_1 + \beta_2 \cdot x + e$, which of the following is **FALSE**?

(a) $\bar{y} - \hat{b}_1 - \hat{b}_2\bar{x} = 0$, where $\bar{x}, \bar{y}$ denote sample means.

(b) The LS estimates for the quadratic regression $Y = \beta_1 + \beta_2 \cdot x^2 + e$ is **not** BLUE because the relation between Y and x is not linear, so the linearity assumption is violated.

(c) In the Simple Regression Model, $y = \beta_1 + \beta_2 \cdot x + e$, $\hat{b}_2 = \frac{cov(x,y)}{var(x)}$, where $cov(x, y)$ is the sample covariance and $var(x)$ is the sample variance of x.

(d) Let $\hat{b}_1^*, \hat{b}_2^*$ be estimates for β_1, β_2 other than the least squares estimates $\hat{b}_1, \hat{b}_2$, then it must be that:

$$\sum_{i=1}^{N}(\hat{b}_1^* + \hat{b}_2^* x_i - y_i)^2 \geq \sum_{i=1}^{N}(\hat{b}_1 + \hat{b}_2 x_i - y_i)^2.$$

(e) The sign of the covariance between $\hat{b}_1, \hat{b}_1$ depends only on the sample mean of x.

Question 16. Let the simple regression model $Y = \beta_1 + \beta_2 \cdot X + e$. Consider the inference that tests the null hypothesis $H_0 : \beta_2 = 0$ against $H_1 : \beta_2 \neq 0$ at significance level α. Which of the statements is **false**?

(a) If the standard error of $\hat{b}_2$ decreases, then it is more likely to reject H_0, (everything else constant).

(b) The higher the absolute value of $\hat{b}_2$, the more likely it is to reject H_0 (everything else constant).

(c) The higher the significance level α, the more likely it is to reject H_0 (everything else constant).

(d) Hypothesis $H_0 : \beta_2 = 0$ is not rejected whenever the value 0 belongs to its confidence interval (with confidence level of $1 - \alpha$).

(e) The larger the sample size, the more likely it is to reject H_0, (everything else constant).

(f) The higher the p-value, the more likely you are to reject H_0.

Question 17. Let the simple regression model $Y = \beta_1 + \beta_2 \cdot X + e$. Consider the inference that tests the null hypothesis $H_0 : \beta_k = 0$ against $H_1 : \beta_k \neq 0$ at significance level α for $k = 1, 2$. Which of the statements is **correct**?

(a) It is less likely to reject the null hypothesis $H_0 : \beta_1 = 0$ if we do the transformation $y_{new} = y \cdot c$.

(b) It is less likely to reject the null hypothesis $H_0 : \beta_2 = 0$ if we do the transformation $x_{new} = x \cdot c$.

(c) It is less likely to reject the null hypothesis $H_0 : \beta_2 = 0$ if we do the transformation $y_{new} = y + c$.

(d) It is less likely to reject the null hypothesis $H_0 : \beta_1 = 0$ if we *standardize* both y and x.

(e) It is less likely to reject the null hypothesis $H_0 : \beta_2 = 0$ if we *standardize* both y and x.

Question 18. Consider two regressions: Y on X and Y on X. Notationally, let $Y = \beta_1^y + \beta_2^y X + \epsilon^y$, where $\hat{b}_1^y, \hat{b}_2^y$ denotes the estimates for β_1^y, β_2^y, $\hat{t}_1^y, \hat{t}_2^y$ are its t-statistics and R_y^2 denotes its goodness of fit. Similarly, let $X = \beta_1^x + \beta_2^x Y + \epsilon^x$, where $\hat{b}_1^x, \hat{b}_2^x, \hat{t}_1^x, \hat{t}_2^x, R_x^2$ denote its respective estimates. Mark the **correct** statement regarding these two regressions:

(a) It is always the case that $\hat{b}_2^y = \hat{b}_2^x$.

(b) It is always the case that $\hat{b}_2^y = 1/\hat{b}_2^x$.

(c) It is always the case that $\hat{b}_1^y = \hat{b}_1^x$.

(d) It is always the case that $\hat{t}_1^y = \hat{t}_1^x$.

(e) It is always the case that $R_y^2 = R_x^2$.

Question 19. Consider the model $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$. Mark the **correct** statement regarding the sample correlation between x_2 and x_3 :

(a) The *lower* the absolute value of the sample correlation, the *larger* the standard errors for the estimators b_2 and b_3.

(b) Any linear transformation of x_2 changes the correlation between x_2 and x_3 and thereby affects the standard error of estimator b_3.

(c) Re-scaling x_2 say $x_2^* = 5 \cdot x_2$ changes the standard error of estimator b_2 but does not change its t-statistic $\hat{t}_2 = \hat{b}_2/\widehat{se}(b_2) = \hat{b}_2^*/\widehat{se}(b_2^*) = \hat{t}_2^*$.

(d) The estimators b_2, b_3 would be most precise if the explanatory variables were equal, that is, $x_2 = x_3$.

(e) If the sample means $\bar{x}_2$ and $\bar{x}_1$ were zero, that is, $\bar{x}_2 = \bar{x}_1 = 0$, then the estimate of the intercept is also zero, $\hat{b}_1 = 0$.

Question 20. Consider the linear Regression model

$$\ln(Y_i) = \beta_1 + \beta_2 x_i + e_i$$

where x denotes annual household income (in thousands) and y denotes annual expenses on consumption goods. Let $\hat{y}$ be the estimated value of Y given the value of the explanatory variable x_0. Consider the following estimates:

$$\hat{\gamma}_x = \hat{b}_2 x_0$$
$$\hat{\gamma}_y = \hat{b}_2 \hat{y}$$
$$\hat{\gamma}_{x,y} = \hat{b}_2 \frac{x_0}{\hat{y}}$$

Which of the following statements is **TRUE**?

(a) $\hat{\gamma}_y$ estimates the percentage increase in consumption expenses associated with an additional \$1,000 in income.

(b) $\hat{\gamma}_x$ estimates the income elasticity of consumption for the average family.

(c) $\hat{\gamma}_y$ estimates the increase in consumption expenses for a 1% increase in income.

(d) $\hat{\gamma}_x$ estimates the average increase in consumption for an addition thousand dollars in income for the average household.

(e) $\hat{\gamma}_{x,y}$ estimates the average % change in food consumption for a 1% increase in income.