KING FAHD UNIVERSITY OF PETROLEUM & MINERALS, DHAHRAN, SAUDI ARABIA
DEPARTMENT OF MATHEMATICS

STAT 310: Regression Analysis

Term 221, First Major Exam
Saturday October 15, 2022, 06:00 PM

Name: ______________________________ ID #: ___________________

Question No	Full Marks	Marks Obtained
1	**08**	
2	**08**	
3	**06**	
4	**20**	
Total	**42**	

Instructions:

1. Formula sheet will be provided to you in exam. You are not allowed to bring, with you, formula sheet or any other printed/written paper.
2. Mobiles are not allowed in exam. If you have your **mobile** with you, **turn it off** and put it **under your seat** so that it is visible to proctor.
3. Show all the calculation steps. There are points for the steps so if your miss them, you lose points.
4. Derive every result that you use in your solution, unless mentioned otherwise.
5. Anything bold in a question indicates that it is a vector or matrix.

Q1: (4+4 = 8 pts.) Consider a simple linear regression model $y_i = \beta_1 X_i + \epsilon_i$ where β_1 is the unknown parameter.

(a) Find the ordinary least square (OLS) estimator of β_1 for this model.

(b) Derive the variance of the slope ($\hat{\beta}_1$) for the least square estimator found in part (a).

Q2: (4+4 = 8 pts.) For a multiple linear regression model $\boldsymbol{y}_{n\times 1} = \boldsymbol{X}_{n\times(k+1)}\boldsymbol{\beta}_{(K+1)\times 1} + \boldsymbol{\epsilon}_{n\times 1}$, the OLS estimates of $\boldsymbol{\beta}$ vector are given as $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$, the fitted values as $\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}$ and the residuals as $\boldsymbol{e} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$. Define the sum of squares of error to be $SSE = \boldsymbol{e}'\boldsymbol{e}$.

(a) Mathematically, show that $SSE = \boldsymbol{y}'\boldsymbol{y} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}$

(b) Mathematically, show that $SSE = \boldsymbol{y}'\boldsymbol{e}$

Q3: (3+3 = 6 pts.) Consider a simple linear regression model $y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ where β_0 and β_1 are the unknown parameter and $\epsilon_i \sim N(0, \sigma^2)$. The estimates of β_0 and β_1 are given as $\hat{\beta}_1 = \frac{S_{XY}}{S_{XX}}$, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{X}$.

(a) Show that $\hat{\beta}_0$ follow a normal distribution.

(b) Show that $\hat{\beta}_0$ is unbiased estimator of β_0 i.e. $E(\hat{\beta}_0) = \beta_0$.

[Blank page 1]

[Blank page 2]

Name: ________________________________ ID #: ____________________

Q4: Code ______

Download the dataset from Blackboard and write down the code number in above blank.

(6+4+4+6 = 20 pts.) A commercial real estate company evaluates vacancy rates, square footage, rental rates, and operating expenses for commercial properties in a large metropolitan area in order to provide clients with quantitative information upon which to make rental decisions. The data below are taken from 50 suburban commercial properties that are the newest, best located, most attractive, and expensive for five specific geographic areas. Shown in the data file are the age (X_1), operating expenses and taxes (X_2), vacancy rates (X_3), total square footage (X_4), and rental rates (Y).

(a) Test the hypothesis that there is significant positive correlation between rental rates (Y) and total square footage (X_4) i.e. $\rho_{y,X_4} > 0$.

H_0: ____________________________ H_1: ________________________________

Test Statistic = ____________________

P-value = __________________

Conclusion:

(b) Fit a simple linear regression model $y_i = a + bX_{4i} + \delta_i$ and construct a 90% confidence interval for the unknown parameter γ_4.

$\hat{b} =$ ________________ Critical Value = __________________

Lower Confidence Limit = _______________ Upper Confidence Limit = _______________

(c) Fit a multiple linear regression model $y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \epsilon_i$. Find the following:

$\widehat{\text{Var}}(\hat{\beta}_2)$ = ____________________ $\widehat{\text{Var}}(\hat{\beta}_4)$ = ____________________

$\widehat{\text{Cov}}(\hat{\beta}_2, \hat{\beta}_4)$ = ____________________ $r_{\hat{\beta}_2, \hat{\beta}_4}$ = ________________________

(d) Using the model in part (c) test that operating expenses and taxes (X_2) and vacancy rates (X_3) are not significantly affecting the rental rates (Y) i.e. $\beta_2 = \beta_3 = 0$.

H_0: ________________________________ H_1: _____________________________________

Test Statistic = ________________________

P-value = _____________________

Conclusion:

Good Luck