King Fahd University of Petroleum and Minerals
College of Computing and Mathematics
Department of Mathematics

STAT 510 – Term 261 – Midterm Exam

Time allowed: 90 minutes

Name: ..... **ID number:** .....

- Answer all five questions on separate paper. Write neatly and legibly.
- Textbooks, notes, and mobile phones are not allowed. A calculator is allowed; no software is needed.
- Show all your working. Answers without justification receive no credit.
- Use $\alpha = 0.05$ for all tests and report numerical answers to three decimal places.
- Formulas and critical values are provided on the final page. For derivation questions, show the mathematical steps even if the result appears on the formula page.

Question	Marks	Max Marks
1		6
2		8
3		9
4		6
5		6
Total		**35**

Question 1 [6 marks]

Consider $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, with fixed x_i, $\mathrm{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, and uncorrelated errors. Assume $S_{xx} > 0$ and use the definitions of S_{xx}, k_i, and $\widehat{\beta}_1$ on the formula page.

(a) Show that $\sum_i k_i = 0$ and $\sum_i k_i x_i = 1$, and hence prove that $\widehat{\beta}_1$ is unbiased for β_1. [3]

(b) Derive $\mathrm{Var}(\widehat{\beta}_1)$. [2]

(c) Is normality of the errors needed for either result above? Explain briefly. [1]

Question 2 [8 marks]

A study relates a response Y to a predictor x. For $n = 12$ observations,

$$\bar{x} = 5, \qquad \bar{y} = 20, \qquad S_{xx} = 50, \qquad S_{xy} = 100, \qquad \text{SSE} = 90.$$

Assume independent $N(0, \sigma^2)$ errors.

(a) Find the fitted regression line and the unbiased estimate of σ^2. [2]

(b) Test whether the slope differs from zero. [3]

(c) At $x_0 = 5$, construct a 95% confidence interval for the mean response and a 95% prediction interval for an independent new observation. Explain why one interval is wider. [3]

Question 3 [9 marks]

The model

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

is fitted to $n = 15$ observations with independent $N(0, \sigma^2)$ errors. The columns of X correspond, in order, to the intercept, x_1, and x_2. You are given

$$(X'X)^{-1} = \begin{bmatrix} 1/15 & 0 & 0 \\ 0 & 0.04 & -0.02 \\ 0 & -0.02 & 0.05 \end{bmatrix}, \qquad X'\mathbf{y} = \begin{bmatrix} 150 \\ 80 \\ 20 \end{bmatrix},$$

$$\text{SSE} = 48, \qquad \text{SST} = 260.$$

(a) Compute $\widehat{\boldsymbol{\beta}}$, write the fitted equation, and interpret the coefficient of x_1. [2]

(b) Compute MSE and explain the residual degrees of freedom. [2]

(c) Test whether x_2 contributes to the model after accounting for x_1. [2]

(d) Calculate R^2 and adjusted R^2. [2]

Question 4 [6 marks]

Using the *same 15 observations*, the model in Question 3 is extended by adding x_3 and x_4. The extended model has an intercept and four regressors, and its SSE is 24. Assume this model has independent $N(0, \sigma^2)$ errors.

(a) Test whether x_3 and x_4 jointly contribute after accounting for x_1 and x_2. Identify the reduced and full models and use the appropriate error mean square. [3]

(b) In a separate auxiliary regression, x_3 is regressed on x_1, x_2, x_4 and an intercept, giving $R_3^2 = 0.96$. Calculate VIF_3 and explain what it indicates about the precision of $\widehat{\beta}_3$. [3]

Question 5 [6 marks]

Let $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ for a model with an intercept. Show that

(a) $\mathbf{e} = (\mathbf{I}_n - \mathbf{H})\mathbf{y}$ and that $\mathbf{H}$ and $\mathbf{I}_n - \mathbf{H}$ are symmetric and idempotent.

(b) $\mathbf{e} = (\mathbf{I}_n - \mathbf{H})\boldsymbol{\varepsilon}$.

(c) $\mathbf{X}'\mathbf{e} = \mathbf{0}$ and deduce that $\sum_i e_i = 0$ and $\mathbf{H1} = \mathbf{1}$.

End of examination questions