Department of Statistics & Operations Research
College of Science, King Saud University

STAT 335 Final Exam
Semester I, 1430 – 1431

Note: Take Σ as $\sum_{i=1}^{n}$. F and t tables are provided. Answer all questions.

Time = 3 hours. [Marks]

[60]

Q.1 The following data shows the marks of 12 students, chosen randomly, in Mathematics and Statistics courses. Consider the simple regression model $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, where ε_i's are i.i.d. $N(0, \sigma^2)$.

Student number: i	1	2	3	4	5	6	7	8	9	10	11	12
Marks in Mathematics: X	65	50	55	65	55	70	65	70	55	70	50	55
Marks in Statistics: Y	85	74	76	90	85	87	94	98	81	91	76	74

Take $\Sigma X_i = 725$, $\Sigma Y_i = 1011$, $\Sigma X_i Y_i = 61685$, $\Sigma X_i^2 = 44475$ and $\Sigma Y_i^{\,2} = 85905$.

a) Find the least square regression line of Y on X. [3]

b) Make an ANOVA table for the above data. [2]

c) Test H_0: $\beta_1 = 0$ against $\beta_1 \neq 0$. Use the level of significance $\alpha = 0.01$. [2]

d) Perform an F test to determine whether or not there is a lack of fit for the linear regression. Use $\alpha = 0.05$. Write down the hypotheses and the conclusion. [4]

e) Obtain the coefficient of determination. Give its meaning. [1]

f) Obtain a 95% confidence interval for the mean of Y when X = 52. [2]

g) Obtain a 90% prediction interval for a $Y_{(new)}$ when X = 52. [1]

Q.2 Consider the simple regression model

$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, where ε_i's are i.i.d. $N(0, \sigma^2)$.

a) Give the meanings of β_0 and β_1. [2]

b) Write down the two normal equations. Show that SST = SSR + SSE [4]

c) Show that $b_1 = \dfrac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2}$ is an unbiased estimator of β_1. [3]

d) Show that $V(b_1) = \dfrac{\sigma^2}{\sum \left(X_i - \overline{X}\right)^2}$ [3]

e) Show that $\mathrm{Cov}(\overline{Y}, b_1) = 0$ [3]

Q. 3 Consider the following data for the regression function of Y on X1 and X2.

Observation	Y	X1	X2
1	6.40	1.32	1.15
2	15.05	2.69	3.40
3	18.75	3.56	4.10
4	30.25	4.41	8.75
5	48.95	6.20	14.82
Sums	119.4	18.18	32.22
Sum of Squares		79.5402	325.8874

(a) Obtain SSTO with accuracy. [2]

(b) Find the matrix X'X and X' Y. [4]

(c) Find $b = (X' X)^{-1} X' Y$. [2]

You can take $(X'X)^{-1} = \begin{bmatrix} 3.31174 & -1.79982 & 0.53265 \\ -1.79982 & 1.17342 & -0.38280 \\ 0.53265 & -0.38280 & 0.13333 \end{bmatrix}$

(d) Write down the regression equation. [1]

(e) Let SSR = 1078.38. Obtain an ANOVA table. [2]

(f) Test $H_0 : \beta_1 = \beta_2 = 0$ at $\alpha = 0.05$ [1]

(g) Find a 95% confidence interval of $E(Y_h)$, when $X_{h1} = 6$ and $X_{h2} = 15$. [3]

Q. 4 For a regression problem, following summary data was obtained.

$n=20$, $\Sigma X_i = 100$, $\Sigma Y_i = 50$, $\Sigma X_i Y_i = 257.66$, $\Sigma X_i^2 = 509.12$ and $\Sigma Y_i^2 = 134.84$.

Consider $Y_i = \beta X_i + \varepsilon_i$, where ε_i's are i.i.d. $N(0, \sigma^2)$. Let b be an estimator of β.

(a) Write down the normal equation for the above regression line. [2]

(b) Write down the fitted regression model. [2]

(c) Find $SSE = \Sigma Y_i^2 - b^2 \Sigma X_i^2$, and the point estimate of σ^2. [4]

Q.5 For the data in Q.4, consider the regression model:

$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, where ε_i's are i.i.d. $N(0, \sigma^2)$.

(a) Write down the normal equations for the above regression line. [1]

(b) Write down the fitted regression model. [1]

(c) Find the point estimate of σ^2 [2]

(d) Find the overall 80% confidence intervals for $E(Y_h)$, when X_h are 40, and 55. [3]