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

STAT 335 Final Exam
Semester I, 1431 – 1432

Note: Take Σ as $\sum_{i=1}^{n}$ .F and t tables are provided. Answer questions: (1,2,3,4) or (1,3,4,5) or (1,2,4,5); that is you can only omit Q.2 or Q.3 or Q.5.

Time = 3 hours. [Marks]

Q.1 The following data shows the weights (in Kg) of 10 wives and their husbands, chosen randomly, from a population. Consider the simple regression model $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, where ε_i's are i.i.d. $N(0, \sigma^2)$.

Observation i	1	2	3	4	5	6	7	8	9	10
Weight of wife: X	65	50	55	65	55	70	65	70	55	70
Weight of husband: Y	85	74	76	90	85	87	94	98	81	91

Take $\Sigma X_i = 620$, $\Sigma Y_i = 861$, $\Sigma X_i Y_i = 53815$, $\Sigma X_i^2 = 38950$, and $\Sigma Y_i^2 = 74653$.

a) Find the least square regression line of Y on X. [2]

b) Find the point estimate of σ^2 [2]

c) Obtain a 90% prediction interval for a $Y_{(new)}$ when $X = 63$. [2]

d) Find an estimate of the mean of Y when X is increased by 1 Kg. [2]

e) Find the coefficient of determination. Give its meaning. [2]

f) Write down the expression for b_1. Show that it is an unbiased estimator of β_1. [2.5]

Q.2 Consider the data and the simple regression model given in Q. 1.

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

b) Test H_0: $\beta_1 = 0$ against $\beta_1 \neq 0$. Use the level of significance $\alpha = 0.05$. [1.5]

c) Perform an F test to determine whether or not there is a lack of fit for the linear regression. Use $\alpha = 0.10$. Write down the hypotheses and draw the conclusion. [4]

d) Find the overall 80% confidence intervals for $E(Y_h)$, when X_h are 50, and 55. [5]

Q.3 For the data given in Q. 1,

a) Give the meaning of β_0. [1]

b) Define SST, SSR and SSE. Show that SST = SSR + SSE [3.5]

c) Show that $Cov(\overline{Y}, b_1) = 0$ [5]

d) Show that $V(b_0) = \dfrac{\sigma^2 \sum (X_i)^2}{\sum (X_i - \overline{X})^2}$ [3]

Q. 4 Consider the following data for the regression function of Y on X1 and X2.

Observation	Y	X1	X2
1	6.3	1.3	1.15
2	15.4	2.7	3.40
3	17.8	3.6	4.10
4	30.9	4.4	8.75
5	52.9	6.2	14.82
6	71.2	18.2	14.82
Sums	194.5	36.4	47.04
Sum of Squares	$\Sigma Y_i^2 = 9416.35$	$\Sigma X1_i^2 = 410.98$	$\Sigma X2_i^2 = 545.52$

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

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

You can take $(X'X)^{-1} = \begin{bmatrix} 0.514758 & 0.0019179 & -0.0458835 \\ 0.001918 & 0.0126610 & -0.0100418 \\ -0.045884 & -0.0100418 & 0.0136230 \end{bmatrix}$

$\begin{bmatrix} 194.5 \\ 1873.63 \\ 2242.122 \end{bmatrix}$

$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$

$SSE = Y'Y - b'X'Y$

$b' = [b_0 \ \ b_1]$

$* SSTO = \Sigma Y_i^2 - \frac{(\Sigma Y_i)^2}{n}$

(c) Find b' X' Y and SSR [2.5]

$* SSE = \Sigma Y_i^2 - b'X'Y$

(d) Make an ANOVA table [2]

(e) Test $H_0 : \beta_1 = \beta_2 = 0$ at $\alpha = 0.05$ [2]

$SSR = SSTo - SSE$

Q. 5 Consider the data for the regression function of Y on X1 and X2, given in Q. 4.
You can use MSE obtained from ANOVA table in Q. 4 or take MSE = 4.0

(a) Find the variance covariance matrix of vector **b**. [2.5]

Compute V(Y_h), when X_{h1} = 6 and X_{h2} = 14.8. [4]

(b) Find a 95% confidence interval of E(Y_h), when X_{h1} = 6 and X_{h2} = 14.8. [4]

(c) Find a 95% prediction interval for $Y_{h(new)}$, when X_{h1} = 6 and X_{h2} = 14.8. [2]

variance - covariance

$s^2(b) = MSE\ (X'X)^{-1}$

mult k1 m9 m6 # s2(b)

$s^2(\hat{Y}_h) = X_h'\ s^2(b)\ X_h$

$= \boxed{}$

read 3 1 m7 # X_h

1

6

14.8

trans m7 m8 # X_h'

multiply m8 m6 m9

multiply m9 m7 m10 # $s^2(\hat{Y}_h)$

$s(\hat{Y}_h)$

$\hat{Y}_h = X_h'\ b$ $\hat{Y}_h = \boxed{}$