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

STAT 310: Linear Regression
Semester 161
First Major Exam
Wednesday October 26, 2016
7:30 − 9:30 pm

Name: ID #:

Question No	Full Marks	Marks Obtained
1	**09**	
2	**06**	
3	**08**	
4	**09**	
5	**10**	
6	**03**	
Total	**45**	

Q.No.1:- (4+5 = 9 points) Consider a simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$ where the intercept β_0 is known.

(a) Find the least square 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).

Q.No.2:- (2+4 = 6 points) Consider the following data and assume that steam usage and average temperature are jointly normally distributed.

Temp (X)	21	24	32	47	50	59	68	74	62	50	41	30
Usage (Y)	185.79	214.47	288.03	424.84	454.68	539.03	621.55	675.06	562.03	452.93	369.95	273.98

$\sum x = 558$ $\sum y = 5062.34$ $\sum x^2 = 29256$ $\sum y^2 = 2416234.61$ $\sum xy = 265869.63$

(a) Find the correlation between steam usage and monthly average ambient temperature.

(b) Test the hypothesis that the correlation coefficient is at least $\frac{1}{2}$. Assume $\alpha = 0.05$ and write down all the testing steps.

Q.No.3:- (8 points) Suppose that the actual regression model was $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ but you ignored the second variable (x_2) and fit a simple linear regression model $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$. The least square estimators for β_0 and β_1 are as follows:

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$ and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.

Is the least square estimator of the slope in the original simple linear regression model unbiased? If no, then find the bias of $\hat{\beta}_1$.

Hint: $\text{Bias}(\hat{\beta}_1) = \text{E}(\hat{\beta}_1) - \beta_1$

Q.No.4:- (3+3+3 = 9 points) Suppose we have a simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$ where ε follows the following distribution:

$f(\varepsilon) = \frac{1}{\sqrt{\theta\pi}} e^{-\frac{\varepsilon^2}{\theta}}, \quad -\infty < \varepsilon < \infty$

with mean $\mathrm{E}(\varepsilon) = 0$ and variance $\mathrm{V}(\varepsilon) = \frac{\theta}{2}$.

If we assume that the x's are fixed, then the distribution of y becomes

$f(y) = \frac{1}{\sqrt{\theta\pi}} e^{-\frac{(y-\beta_0-\beta_1 x)^2}{\theta}}, \quad -\infty < y < \infty.$

Using this density function, make a likelihood function and then finally derive the maximum likelihood estimators (MLEs) for all the 3 parameters i.e. β_0, β_1 and θ.

Q.No.4:- continues...

Q.No.5:- (3+4+3 = 10 points) Consider the following fuel consumption data. The automobile engineer believes that the fuel consumption decreases for each one C^0 increase in initial boiling point.

	Fuel Consumption (F)	Initial Boiling Point (P)	F^2	P^2	F*P
1	343	220	117649	48400	75460
2	356	220	126736	48400	78320
3	344	223	118336	49729	76712
4	356	223	126736	49729	79388
5	352	221	123904	48841	77792
6	361	221	130321	48841	79781
7	372	190	138384	36100	70680
8	355	190	126025	36100	67450
9	375	180	140625	32400	67500
10	359	180	128881	32400	64620
11	364	180	132496	32400	65520
12	357	180	127449	32400	64260
13	368	176	135424	30976	64768
14	360	176	129600	30976	63360
15	372	175	138384	30625	65100
16	352	175	123904	30625	61600
Sum	**5746**	**3130**	**2064854**	**618942**	**1122311**

(a) Fit s simple linear regression model and interpret both the estimated coefficients.

(b) Do the data support the engineer's belief? Justify your answer by performing a relevant hypothesis test. Assume $\alpha = 0.05$ and write down all the testing steps.

(c) Predict the value of fuel consumption when the initial boiling point is at 170 C^0. Also, make 95% prediction interval and interpret it.

Q.No.6:- (3 points) If $\mathrm{Var}(X+Y)=\mathrm{Var}(X-Y)$, then find the correlation coefficient between the two variables X and Y.

With the Best Wishes