CONCORDIA UNIVERSITY
Department of Economics

ECON 222/4 SECTIONS C, D and EE
STATISTICAL METHODS II
WINTER 2017 – ASSIGNMENT 2
Due: Monday, March 20 before 5:00pm

question	1	2	3	4	5	total
marks						

1. (38 **marks**) Consider the simple regression model $y_i = \beta_0 + \beta_1 x_i + e_i$, where β_0 and β_1 are unknown population parameters and $e \sim iid\, N(0, \sigma^2)$. Use the sample data given in the table below to answer the following questions.

	x_i	y_i							
	1	1							
	3	2							
	4	3							
	8	4							
	6	5							
	9	7							
	11	8							
	14	10							

a. (**2 marks**) Calculate b_0.

$$b_0 = \bar{y} - b_1 \bar{x} = 5 - \frac{92}{132}(7) \cong 0.12$$

b. (**2 marks**) Calculate b_1.

$$b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{372 - 8(7)(5)}{524 - 8(7)^2} = \frac{92}{132} \cong 0.7$$

c. (**2 marks**) ***Briefly*** interpret b_0.

$\hat{y}_i = 0.12 + 0.7X$ when $X = 0$ $y = 0.12$

d. (**2 marks**) ***Briefly*** interpret b_1.

when X changes by 1 unit, y changes by 0.7 un

	x_i	y_i					
	1	1					
	3	2					
	4	3					
	8	4					
	6	5					
	9	7					
	11	8					
	14	10					

e. **(2 marks)** Calculate SST.

$$SST = \sum (y_i - \bar{y})^2 = \sum y_i^2 - n\bar{y}^2 = 268 - 8(5)^2 = 68$$

f. **(2 marks)** Calculate SSR.

$$SSR = \sum (\hat{y}_i - \bar{y})^2 = \sum \hat{y}_i^2 - n\bar{y}^2 \cong 64$$

g. **(2 marks)** Calculate SSE.

$$SSE = \sum \hat{u}_i^2 = \sum (y_i - \hat{y}_i)^2 \cong 4$$

h. **(2 marks)** Calculate the coefficient of determination, R^2.

$$R^2 = \frac{SSR}{SST} = \frac{64}{68} \cong 0.94$$

i. **(2 marks)** ***Briefly*** interpret R^2.

Approximately 94% of the variation in y can be explained by the estimated regression

j. **(2 marks)** Calculate the correlation coefficient, r.

$$r = (\text{sign of } b_1)\sqrt{R^2} = \sqrt{0.94} \cong 0.97$$

k. **(2 marks)** ***Briefly*** interpret r.

There is a very strong, almost perfect, positive linear relationship between X and Y

l. (**2 marks**) Calculate the standard error of the estimate, $\hat{\sigma}$.

$$\hat{\sigma} = \sqrt{\frac{SSE}{n-k-1}} = \sqrt{\frac{4}{8-2}} = \sqrt{4/6} \cong 0.82$$

m. (**2 marks**) Calculate $\operatorname{var}(b_1)$.

$$var(b_1) = \frac{\hat{\sigma}^2}{\Sigma(x_i - \bar{x})^2} = \frac{4/6}{132} \cong 0.005$$

n. (**2 marks**) Calculate $\operatorname{var}(b_0)$.

$$var(b_0) = \frac{\hat{\sigma}^2 \Sigma x_i^2}{n\,\Sigma(x_i - \bar{x})^2} = \frac{4/6\,(524)}{8\,(132)} \cong 0.33$$

o. (**2 marks**) Calculate $\operatorname{cov}(b_0, b_1)$.

$$cov(b_0, b_1) = \frac{-\bar{x}\,\hat{\sigma}^2}{\Sigma(x_i - \bar{x})^2} = \frac{-7\,(4/6)}{132} \cong -0.035$$

p. (**2 marks**) Use the estimated regression line to predict y when x = 10.

$$\hat{Y}_0 = 0.12 + 0.7(10)$$
$$= 7.12$$

q. (**2 marks**) Construct a 95-percent prediction interval for y when x = 10, using $t_{0.025,6} = \pm 2.447$

$$\hat{Y}_0 = 7.12$$

$$\hat{Y}_0 \pm t_c \cdot se(f) = 10 \pm 2.447\,(0.89)$$
$$= 10 \pm 2.178$$
$$[7.822, 12.178]$$

$$se(f) = \sqrt{\hat{\sigma}^2\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}\right]}$$

$$= \sqrt{\frac{4}{6}\left[1 + \frac{1}{8} + \frac{(10-7)^2}{132}\right]} = 0.89$$

r. (**2 marks**) Construct a 95-percent confidence interval for β_1, using $t_{0.025,6} = \pm 2.447$

$$b_1 \pm t_c \, se(b_1)$$

$$= 0.7 \pm 2.447 \sqrt{0.005}$$

$$= 0.7 \pm 0.173$$

$$[0.527 , 0.7173]$$

s. (**2 marks**) Test, at the 5-pecent level of significance, whether $\beta_1 = 1$.
Clearly state the null and alternative hypotheses, the test statistic and your conclusion. Use a critical value $t_{0.025,6} = \pm 2.447$

$H_0: \beta_1 = 1$ vs $H_a: \beta_1 \neq 1$

calculated test statistic

$$t_{cal} = \frac{b_1 - \beta_1}{se(b_1)} = \frac{0.7 - 1}{\sqrt{0.005}} \cong -4.2$$

$t_{cal} = -4.2$ falls in the lower critical region

so we reject the null hypothesis

at the 5% level of significance.

2. (4 **marks**) Consider the standard linear regression model, $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ with all classical assumptions holding.

a. (**4 marks**) Derive the Ordinary Least Squares (OLS) coefficient estimators for the model above.

i. (**2 marks**) b_0, the estimator of the intercept coefficient. (Show all steps)

$\hat{Y}_i = b_0 + b_1 X_i \qquad \hat{e}_i = Y_i - \hat{Y}$

want to minimize $\Sigma \hat{e}_i^2$

$SSR = \Sigma \hat{e}_i^2 = \Sigma (y_i - \hat{y})^2$

$= \Sigma (y_i - b_0 - b_1 X_i)^2$

$\underset{b_0, b_1}{\text{minimize}} \; \Sigma (y_i - b_0 - b_1 X_i)^2$

F.O.C. conditions & set derivatives to 0

$\frac{\partial SSR}{\partial \hat{\beta}_0} = \Sigma 2(y_i - b_0 - b_1 X_i)(-1) = 0$

$\Sigma (y_i - b_0 - b_1 X_i) = 0$

$\Sigma y_i - \Sigma b_0 - \Sigma b_1 X_i = 0$

$n\bar{Y} - n b_0 - n b_1 \bar{X} = 0$

$\bar{Y} - b_0 - b_1 \bar{X} = 0$

$b_0 = \bar{Y} - b_1 \bar{X}$

$\bar{X} = \frac{\Sigma X_i}{n}$

$\Sigma X_i = n\bar{X}$

$\bar{Y} = \frac{\Sigma Y_i}{n}$

$\Sigma Y_i = n\bar{Y}$

ii. **(2 marks)** b_1, the estimator of the slope coefficient. (Show all steps)

$$\frac{\partial SSR}{\partial b_1} = \sum 2(y_i - b_0 - b_1 x_i)(-x_i) = 0$$

$$\sum (y_i - b_0 - b_1 x_i)(x_i) = 0$$

$$\sum (y_i x_i - b_0 x_i - b_1 x_i^2) = 0$$

$$\sum y_i x_i - \underbrace{b_0}_{b_0 = \bar{y} - b_1 \bar{x}} \sum x_i - b_1 \sum x_i^2 = 0$$

$$\sum y_i x_i - (\bar{y} - b_1 \bar{x}) \underbrace{\sum x_i}_{n\bar{x}} - b_1 \sum x_i^2 = 0$$

$$\sum y_i x_i - (\bar{y} - b_1 \bar{x}) n\bar{x} - b_1 \sum x_i^2 = 0$$

$$\sum y_i x_i - n\bar{y}\bar{x} + n b_1 \bar{x}^2 - b_1 \sum x_i^2 = 0$$

$$\underbrace{\sum y_i x_i - n\bar{y}\bar{x}}_{\sum (y_i - \bar{y})(x_i - \bar{x})}$$

$$\sum (y_i - \bar{y})(x_i - \bar{x}) = b_1 \sum x_i^2 - n b_1 \bar{x}^2$$

$$= b_1 \left(\sum x_i^2 - n\bar{x}^2\right)$$

$$= b_1 \sum (x_i - \bar{x})^2$$

$$b_1 = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

3. **(6 marks)** ***Briefly*** interpret b_0 and b_1 in each of the following equations.

a. **(2 marks)** $\hat{y}_i = b_0 + b_1 \ln x_i$

when $x = 1$, y should equal bo units

when x changes by 1%

y should change by b_1 units

b. **(2 marks)** $\widehat{\ln y_i} = b_0 + b_1 x_i$

when $x = 1$, ln y should equal bo units

when x changes by 1 unit,

y should change by b_1 %

c. **(2 marks)** $\widehat{\ln y_1} = b_0 + b_1 \ln x_i$

when $x = 1$, ln y should equal bo units

when x changes by 1%,

y should change by b_1 %

4. (10 **marks**) Consider the following regression models $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$
and $Y_i^* = \boldsymbol{\beta_0}^* + \beta_1^* X_i^* + \varepsilon_i^*$

Let $Y_i^* = wY_i$ where w is a constant. What is the relationship between

a. (**2 marks**) β_0 and β_0*

$$\beta_0 = \frac{\beta_0^*}{w} \quad \text{or} \quad \beta_0^* = w\beta_0$$

$$wY_i = \beta_0^* + \beta_1^* X_i^* + \varepsilon_i^*$$

$$Y_i = \frac{\beta_0^*}{w} + \frac{\beta_1^*}{w} X_i^* + \frac{\varepsilon_i^*}{w}$$

$$\beta_0 = \frac{\beta_0^*}{w} \qquad \varepsilon_i = \frac{\varepsilon_i^*}{w}$$

$$\beta_1 = \frac{\beta_1^*}{w}$$

b. (**2 marks**) β_1 and β_1*

$$\beta_1 = \frac{\beta_1^*}{w} \quad \text{or} \quad \beta_1^* = w\beta_1$$

c. (**2 marks**) var (β_0) and var (β_0*)

$$var(\beta_0) = var\left(\frac{\beta_0^*}{w}\right) = \frac{1}{w^2} var(\beta_0^*)$$

d. (**2 marks**) var (β_1) and var (β_1*)

$$var(\beta_1) = var\left(\frac{\beta_1^*}{w}\right) = \frac{1}{w^2} var(\beta_1^*)$$

e. (**2 marks**) standard deviation of (β_1) and standard deviation of (β_1*)

$$\text{standard deviation}(\beta_1) = \sqrt{var(\beta_1)} = \sqrt{\frac{1}{w^2} var(\beta_1^*)}$$

$$= \frac{1}{w} var(\beta_1^*)$$

5. (**8 marks**) **Complete textbook exercise 2.13.**

8