Week 2 - AYU - Individual

Type 1: Parameter Estimators

Problem 1

You are given the following data

y	x_1	x_2
2	1	1
3	1	1
4	2	2
6	3	2
10	3	4

You are using the following model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i, \quad i = 1, 2, ..., 6$$

You have determine

$$(X'X)^{-1} = \begin{bmatrix} 1.2 & -0.5 & 0 \\ -0.5 & 0.75 & -0.5 \\ 0 & -0.5 & 0.5 \end{bmatrix}$$

Determine $\hat{\beta_1}$.

(A) -2.00
(B) -0.75
(C) -0.5
(D) 0.5
(E) 0.75

Problem 2

You are given the following data

y	x_1	x_2
2	1	1
3	1	0
4	2	0
6	0	2
10	0	0

You are using the following model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i, \quad i = 1, 2, ..., 6$$

You have determine

```
     [,1]       [,2] [,3]
```

[1,] 0.8285714 -0.4857143 -0.4 [2,] -0.4857143 0.4571429 0.2 [3,] -0.4000000 0.2000000 0.4

$$(X'X)^{-1} = \begin{bmatrix} 0.8285714 & -0.4857143 & -0.4 \\ -0.4857143 & 0.4571429 & 0.2 \\ -0.4000000 & 0.2000000 & 0.4 \end{bmatrix}$$

Determine $\hat{\beta}_0$.

(A) -15.00
(B) -10.75
(C) 8.8
(D) 10.5
(E) 9.75

Type 2: F-test and Coefficient of Determination

Problem 3 (SRM - Sample Question 24)

Sarah performs a regression of the return on a mutual fund (y) on four predictors plus an intercept. She uses monthly returns over 105 months. Her software calculates the F statistic for the regression as $F = 20.0$, but then it quits working before it calculates the value of R^2. While she waits on hold with the help desk, she tries to calculate R^2 from the F-statistic.

Determine which of the following statements about the attempted calculation is true.

(A) There is insufficient information, but it could be calculated if she had the value of the residual sum of squares (RSS).
(B) There is insufficient information, but it could be calculated if she had the value of the total sum of squares (TSS) and RSS.
(C) $R^2 = 0.44$
(D) $R^2 = 0.56$
(E) $R^2 = 0.80$

Problem 4 You fit a multiple linear regression to a data of 20 observation and 4 predictors. You have determined that the coefficient of determination of the model is 0.8. Calculate the F-statistics to test the significant of the model.

(A) 4
(B) 8
(C) 12
(D) 15
(E) 18

Generalized F-test Problem 5 (CAS Exam ST Spring 2016 Question 22)

The following two models were fit to 18 observations:

Model 1: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$
Model 2: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2 + \epsilon$

The result of the regression are:

Model Number	Error Sum of Squares	Regression Sum of Squares
1	102	23
2	78	47

Calculate the value of the F-statistics used to test the hypothesis that $\beta_3 = \beta_4 = \beta_5 = 0$

(A) Less than 1.30
(B) At least 1.30, but less than 1.40
(C) At least 1.40, but less than 1.50
(D) At least 1.50, but less than 1.60
(E) At least 1.60

Problem 6 (SOA Course)

A professor ran an experiment in three sections of a psychology courses to show that the more digits in a number, the more difficult it is to remember. The following variables were used in a multiple regression:

\$x_1 = \$ number of digits in the number $x_2 = 1$ if student was in section 1, 0 otherwise $x_3 = 1$ if student was in section 2, 0 otherwise \$y = \$ percentage of students correctly remembering the number

You are given

- A total of 42 students participated in the study
- The regression equation
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_2 + \beta_4 x_3 + \epsilon$$
was to fit the data and resulted in $R^2 = 0.940$
- A second regression equation $y = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_1^2 + \epsilon$ was to fit to the data and resulted $R^2 = 0.915$

Determine the F statistic used to test whether class section is significant variable.

(A) 5.4
(B) 7.3
(C) 7.7
(D) 7.9
(E) 8.3

Problem 7 (SOA Course)

You are determining the relationship of salary (y) to experience (x_1) for both men ($x_2 = 1$) and woman ($x2 = 0$). You fit the model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon$$

You are given

Sources	Sum of Squares	Degree of Freedom
Regression	330.0117	3
Error	12.8156	7

You also fit the below model to the same observations.

$$y = \gamma_0 + \gamma_1 x_1 + \epsilon$$

You are given

Sources	Sum of Squares	Degree of Freedom
Regression	315.0992	1
Error	27.7281	9

Determine the F-ratio to test whether the linear relationship between salary and experience is identical for men and women.

(A) 0.6
(B) 2.0
(C) 3.5
(D) 4.1
(E) 6.2

Problem 8 (SRM - Sample Question 27)

Trevor is modeling monthly incurred dental claims. Trevor has 48 monthly claims observations and three potential predictors: - Number of weekdays in the month - Number of weekend days in the month - Average number of insured members during the month

Trevor obtained the following results from a linear regression:

	Coefficient	Standard Error	Stat	p-value
Intercept	–45,765,767.76	20,441,816.55	–2.24	0.0303
Number of weekdays	513,280.76	233,143.23	2.20	0.0330
Number of weekend days	280,148.46	483,001.55	0.58	0.5649
Average number of members	38.64	6.42	6.01	0.0000

Determine which of the following variables should be dropped, using a 5% significance level.

I. Intercept
II. Number of weekdays
III. Number of weekend days
IV. Number of members.

(A) I only
(B) II only
(C) III only
(D) IV only
(E) None should be dropped from the model

Type 3: Application of MLM Models

Problem 9 (SRM - Sample Question 49)

The actuarial student committee of a large firm has collected data on exam scores. A generalized linear model where the target is the exam score on a 0-10 scale is constructed using a log link, resulting in the following estimated coefficients

Predictor Variables	Coefficient
Intercept	-0.1
Study Time (in units of 100 hours)	0.5
Attempt (1 for first attempt, else 0)	0.5
Master's degree (1 for Yes, 0 for No)	–0.1
Interaction of Attempt and Master's degree	0.2

The company is about to offer a job to an applicant who has a Master's degree and for whom the exam would be a first attempt. It would like to offer half of the study time that will result in an expected exam score of 6.0. Calculate the amount of study time that the company should offer.

(A) 123 hours
(B) 126 hours
(C) 129 hours
(D) 132 hours
(E) 135 hours