

MA329 Statistical linear models

Assignment 6 (Due date: Dec 29, 11:00pm. For late submission, each day costs 10 percent)

1. A research was conducted to study how communities respond to a violent crime. Monthly violent crime statistics were collected in two neighboring cities, A and B, before and after the murder of a citizen in city A. The following regression model was adopted.

$$\text{MODEL G: } E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$$

where

y is the violent crime rate (number of crimes per 1000 population)

$x_1 = \{1 \text{ if city A, } 0 \text{ if city B}\}$

$x_2 = \{1 \text{ if after the murder, } 0 \text{ if before the murder}\}$

- a. Express, in terms of the β s, the expected violent crime rate for months after the murder in city A.
 - b. Express, in terms of the β s, the expected violent crime rate for months before the murder in city B.
 - c. With $\alpha = 0.05$, state clearly an α -level test (including the full model, the reduced model, the hypotheses, the test statistics and the rejection rule) for testing the usefulness of the independent variable x_2 in the model.
2. The administration of an airline company conducted a salary equity study to help establish benchmarks for employee salaries. The administration utilized the following regression model for annual salary,

$$E(Y) = \beta_0 + \beta_1 X$$

where $X = 0$ if pilot, 1 if mechanic, 2 if flight attendant. The administration wanted to use the model to compare the mean salaries of flight crew members in different job positions.

- a. Explain the flaw (problem) in the model.
 - b. Propose a better model that will achieve the administration's objective.
3. An economist wants to compare the mean dollar amounts owed by delinquent credit card customers in the three different socioeconomic classes: (1) lower class, (2) middle class, and (3) upper class. Let

μ_1 = mean amount owed by lower class delinquent credit card customers

μ_2 = mean amount owed by middle class delinquent credit card customers

μ_3 = mean amount owed by upper class delinquent credit card customers

A sample of 10 customers with delinquent accounts is selected from each group, and the amount owed by each is recorded, as shown in the following table.

Dollars Owed		
<u>Group 1 (lower class)</u>	<u>Group 2 (middle class)</u>	<u>Group 3 (upper class)</u>
148	513	335
76	264	643
393	433	216
520	94	536
236	535	128
134	327	723
55	214	258
166	135	380
415	280	594
153	304	465

A regression model is used to predict the amount owed (y) using socioeconomic class as an independent variable. Use indicator variables

$$X_1 = \begin{cases} 1 & \text{if group 2} \\ 0 & \text{if not} \end{cases} \quad X_2 = \begin{cases} 1 & \text{if group 3} \\ 0 & \text{if not} \end{cases}$$

The regression model is

$$E(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Computer output

ANOVA

<u>Model</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>
Regression	2	198772	99386
Error	27	770671	28543
Total	29	969443	

Coefficients

	Coefficients	Std. Error	<i>t</i>
Constant	229.6	53.43	4.3
X_1	80.3	75.56	1.06
X_2	198.20	75.56	2.62

- Express the parameters β_0 , β_1 and β_2 in terms of μ_1 , μ_2 and μ_3 .
- Compute the estimate of the difference between the population mean delinquent amounts for the upper- and lower-class groups.
- Compute the estimate of the difference between the population mean delinquent amounts for the upper- and middle-class groups.
- Test the null hypothesis that the population mean delinquent amounts for the three groups are equal ($\alpha = 0.05$). State clearly the null and alternative hypotheses, the test statistics, the accept/reject decision and the conclusion.