BINDURA UNIVERSITY OF SCIENCE EDUCATION SFM421

BACHELOR OF STATISTICS AND FINANCIAL MATHEMATICS

GENERAL LINEAR MODELS

MAR 2023

Time : 3 hours

Candidates should attempt ALL questions in section A and at most TWO questions in section B.

SECTION A (40 marks)

Candidates may attempt ALL questions being careful to number them A1 to A3

A1. (a) Explain using examples the meaning of the following pre- test statistics in regression modeling

(i) Multicollinearity. [3]

(ii) Unit root testing. [3]

(iii) Normality. [3]

(iv) Differencing. [3]

(b) State and explain three main components of general linear models. [6]

A2. (a) Let y be a $k \times 1$ random vector with $E(y) = \mu$ and $Var(y) = V$. Let **A** be a $k \times k$ matrix of real numbers. Prove that $E(y'Ay) = tr(AV) + \mu' A\mu$. [7]

(b) Let $y = X\beta + e$ where **X** is an $n \times p$ matrix of full rank, β is a $p \times 1$ vector of unknown parameters, and **e** is an $n \times 1$ normally distributed random vector with mean **0** and variance $\sigma^2 I$. Derive the maximum likelihood estimator of β and $\sigma^2 I$. [8]

A3. (a) What is data standardisation. [2]

(b) Explain what is meant by training and testing data. Give one possible training-testing ratio and its two possible advantages over others [5]

SECTION B (60 marks)

Candidates may attempt TWO questions being careful to number them B4 to B6.

B4. Let $y = X\beta + e$ where **X** is an $n \times p$ matrix of full rank, β is a $p \times 1$ vector of unknown parameters, and **e** is an $n \times 1$ normally distributed random vector with mean **0** and variance $\sigma^2 I$. Assume that $t'\beta$ is estimable where **t** is a $1 \times p$ non zero vector of real numbers and that $\hat{\beta}$ denotes any solution to the normal equations.

(a) Show that $E(t'\hat{\beta}) = t'\beta$. [3]

(b) Prove that $Var(t'\hat{\beta}) = t'(X'X)^{-1}t\sigma^2$. [4]

(c) Deduce that the random variable $\frac{t'\hat{\beta}-t'\beta}{\sigma\sqrt{t'(X'X)^{-1}t}}$ follows a standard normal distribution. [10]

(d) Deduce that the random variable $\frac{t'\hat{\beta}-t'\beta}{s\sqrt{t'(X'X)^{-1}t}}$ follows a student t distribution with $n-p$ degrees of freedom. [9]

(e) Derive the $100(1-\alpha\%)$ confidence bounds on the estimable function $t'\beta$ [4]

B5. An experiment was conducted to estimate the demand for cars (y) based on cost (x_1) and current employment rate (x_2). The proposed linear model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$$

These data are obtained,

x_1	x_2	y
2.1	4.4	1
2.2	2.7	2
2.9	5.2	3
2.3	3.3	4
3.5	2.4	5
3.1	3.7	6
3.0	5.1	7
3.2	4.9	8

(a) Find $\hat{\beta}$. [7]

(b) Construct the ANOVA table using the uncorrected sum total. [8]

(c) Test if the model is adequate. [5]

(d) Calculate $R(\beta_0)$ and $R(\beta_1/\beta_0)$. [10]

B6. (a) Show that $\frac{(n-p)s^2}{\sigma^2} = \frac{SS_{Res}}{\sigma^2}$ follows a chi square distribution with $n-p$ degrees of freedom. [6]

(b) Let $y = X\beta + e$ where $\mathbf{X}$ is an $n \times p$ matrix of full rank, β is a $p \times 1$ vector of unknown parameters, and $\mathbf{e}$ is an $n \times 1$ normally distributed random vector with mean $\mathbf{0}$ and variance $\sigma^2 I$. Show that $t'\hat{\beta}$is BLUE for $t'\beta$. [10]

(c) It is known that humidity influences evaporation in water reducible paints. These data are obtained on x, the relative humidity, and y, the extent of solvent evaporation in the paint during sprayout.

x	31	24	30	54	60	70	71	76	67	55
y	11	12	13	10	9	9.9	8.8	11	14	13

(i) Find 90 percent joint confidence interval for β assuming that both β_0 and β_1 are significant. [10]

(ii) Find a 90 percent confidence interval for the mean extent of solvent evaporation with the relative humidity of 40. [4]

END OF QUESTION PAPER