CHRIST (DEEMED TO BE UNIVERSITY),BENGALURU - 560029

End Semester Examination October/November - 2018
Bachelor of Science V SEMESTER

Code: STA531   Max.Marks: 100
Course: LINEAR REGRESSION MODELS   Duration: 3Hrs

SECTION A

Answer any TEN questions   10X3=30

1   Distinguish the relationship between deterministic and probabilistic models.

2   What are the assumptions involved in simple linear regression?

3   Explain multiple linear regression model with k-explanatory variables.

4   Define Residual term. Explain the role of residual term in MLR Models.

5   Explain the need of mean square error criteria in model selection.

6   Explain forward and backward methods for variable selection in regression modeling.

7   Define the term residual and give some of the assumption of residuals in regression modeling.

8   Explain various types of residual plots in regression.

9   Explain the role of residuals in regression model building.

10 Explain the role of outliers in regression model building.

11 If the assumptions are not satisfied in regression modeling, what are the possible consequences? Explain in brief.

12 Explain the term "variance inflation factor".

SECTION B

Answer any FIVE questions   5X6=30

13 A regression between foot length (dependent variable in cm) and height (independent variable in inches) for 33 students resulted in the following regression equation: $\hat{y} = 10.9 + 0.23 x$

    (i) One student in the sample was 73 inches tall with a foot length of 29 cm. What is the predicted foot length for this student?

    (A) 17.57 cm (B) 27.69 cm (C) 29 cm (D) 33 cm

    (ii) One student in the sample was 75 inches tall with a foot length of 30 cm. What is the residual for this student?

    (A) 29 cm (B) 1.31 cm (C) 0.00 cm (D) -1.31 cm

    (iii) What is the estimated average foot length for students who are 70 inches tall?

    (A) 27 cm (B) 28 cm (C) 29 cm (D) 30 cm

14 Explain the method of least square estimation to estimate the regression coefficients in $Y_i = \beta_0 + \beta_1 X_i + \in_i, (i = 1, 2, \ldots, n)$.

15 Explain five important assumptions of multiple linear regression analysis in detail.

16 Explain ordinary least squares estimation method for a multiple linear regression model $Y = X \beta + \ominus$, where Y is (n x1), X is (n x k), β is (nx1) and $\ominus$ is an (nx1) order matrices.

17 Explain the need of transformation of variables in regression analysis. Explain Box Cox transformation.

18 (a) Describe the consequences of model misspecification in multiple linear regression model.

    (b) Calculate the amount of bias occured and the estimate of variance for the misspecified model.

SECTION C

Answer any FOUR questions.   4X10=40

19 Explain ANOVA table for simple linear regression model and derive the test for

significance of regression parameters.

20 (a) Explain the procedure to construct 95% and 99% confidence intervals for the true regression coefficient β_1 in a simple linear regression model $Y_i = \beta_0 + \beta_1 X_i + \in_i, (i = 1, 2, \ldots, n)$.

(b) The effect of the temperature of the deodorizing process on the color of the finished product was determined experimentally. A sample of 15 observations collected from the process and the regression analysis output using SPSS is as given below:

Coefficients[a]		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta		
1	(Constant)	2.536	?		4.802	.000
	Temperature	?	.001	-.726	-3.809	.002
a. Dependent Variable: Color						

(i) Estimate the missing values from the above table.

(ii) Form a simple linear regression model

(iii) Is the regression coefficients are significant? (Use $\alpha = 0.05$)

(iv) Interpret your result.

21 Let $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i, (i = 1, 2, 3)$, where $Y_i = \{5, 3, 4\}, X_{1i} = \{1, 4, 2\}$ and $X_{2i} = \{2, 1, 0\}$ respectively. Find the least square estimates of β_0 , β_1 and β_2.

22 (a) Explain the test procedure for testing the significance of regression coefficients in MLR model.

(b) A study was conducted to predict the BMI of persons based on several factors like Education, Calories, income and Expenditure of persons. Data were collected and the output of SPSS is given below:

Coefficients[a]		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta		
1	(Constant)	20.693	.208		?	.000
	calorie	.002	.000	.753	38.969	.000
	exercise	?	.003	-.163	-8.434	.000
	income	8.819E-005	.000	.035	1.837	.067
	education	-.001	?	-.002	-.086	.932
a. Dependent Variable: BMI						

(i) Fill the missing values (?).

(ii) From the above results form a fitted multiple linear regression model.

(iii) Identify the variables which are having significant influence on BMI.

(iv) Give your conclusion based on Beta values.

(v) Give your conclusion from this output.

23 (a) Explain the usage of leverages in regression analysis.

(b) Explain any two methods to identify the influencial data points.

24 (a) Define autocorrelation.

(b) Explain Durbin-Watson test. A study was conducted to investigate the relationship

between monthly income and expenditure of 30 households and the following result obtained:

Model Summary[b]					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.931[a]	.866	.861	5.19203	2.060

Using above table give some useful conclusions about the model assumptions.