

Seat No.	
----------	--

Set **P**

M.Sc. (Semester - III) (CBCS) Examination Oct/Nov-2019
Statistics
REGRESSION ANALYSIS

Day & Date: Saturday, 09-11-2019
 Time: 03:00 PM To 05:30 PM

Max. Marks: 70

Instructions: 1) All questions are compulsory.
 2) Figures to the right indicate full marks.

Q.1 Fill in the blanks by choosing correct alternatives given below. 14

- 1) The LSE in general linear model is unique if _____.
 a) coefficient matrix is full rank
 b) coefficient matrix is non-full rank
 c) if generalized inverse of coefficient matrix exist
 d) none of these
- 2) Any vector in estimation space is _____ to any vector in error space.
 a) linear
 b) orthogonal
 c) projected
 d) normalized
- 3) The model $Y = \beta_0 e^{\beta_1 X} \epsilon$ can be linearized by using _____ transformation.
 a) square root
 b) reciprocal
 c) logarithmic
 d) none of these
- 4) Forward selection procedure begins with the assumption that there are _____.
 a) no regressors in the model
 b) all regressors in the model
 c) some regressors in the model
 d) none of these
- 5) In simple linear regression model $Y = \beta_0 + \beta_1 X + \epsilon$, β_0 and β_1 are _____.
 a) slope and intercept
 b) intercept and slope
 c) error and slope
 d) intercept and error
- 6) In multiple linear regressions, variance of LSE of β is _____.
 a) $(X'X)\sigma^2$
 b) $(X'X)^{-1}\sigma^2$
 c) $X(X'X)^{-1}X'\sigma^2$
 d) σ^2
- 7) The condition number of $(X'X)$ matrix is given as _____.
 a) $\lambda_{max} - \lambda_{min}$
 b) $\lambda_{max} + \lambda_{min}$
 c) $\frac{\lambda_{max}}{\lambda_{min}}$
 d) $\frac{\lambda_{min}}{\lambda_{max}}$
- 8) If we use unit length scaling for the regressor variables then $X'X$ matrix of scaling regressors will be in the form of _____.
 a) covariance matrix
 b) correlation matrix
 c) identity matrix
 d) none of these
- 9) If the coefficient of determination (R^2) is near to 1 then it leads to the conclusion that _____.
 a) a good linear relation exists
 b) there is a lack of linear relationship
 c) there is a curvilinear relation
 d) none of these

- 10) The hat matrix $H = X(X'X)^{-1}X'$ is _____.
 - a) symmetric and orthogonal
 - b) symmetric and idempotent
 - c) skew symmetric matrix
 - d) identity matrix
- 11) The multicollinearity in linear regression concerns with _____.
 - a) The error terms
 - b) The regressors
 - c) The response variable values
 - d) The coefficient
- 12) The LSE of β for the model $Y = X\beta + \epsilon$ can be written as _____.
 - a) $\beta + (X'X)^{-1}\epsilon$
 - b) $\beta + (X'X)\epsilon$
 - c) $\beta + X'\epsilon$
 - d) $\beta + (X'X)^{-1}X'\epsilon$
- 13) The regression model $Y = \beta_0 + \beta_1X + \beta_2X^2$ is called _____ model.
 - a) linear
 - b) non-linear
 - c) polynomial
 - d) none of these
- 14) In usual notations, $var(\hat{Y}) =$ _____.
 - a) $H\sigma^2$
 - b) σ^2
 - c) $(I - H)\sigma^2$
 - d) $H(I - H)\sigma^2$

Q.2 A) Answer the following questions. (Any Four) 08

- 1) Define the coefficient of determination R^2 and adj. R^2 . Derive the relation between them.
- 2) Define K^{th} order polynomial regression model in one variable.
- 3) Define condition number and condition indices of $X'X$ matrix.
- 4) Explain the procedure of computing λ , the parameter of power transformation.
- 5) Define intrinsically model. Give an example.

B) Write short notes. (Any Two) 06

- 1) Variance stabilizing transformation
- 2) Prediction interval for the model $Y = X\beta + \epsilon$
- 3) Cubic spline and cubic-B spline

Q.3 A) Answer the following questions. (Any Two) 08

- 1) Define residual. Obtain its mean and variance.
- 2) With usual notations, prove that R^2 is the square of correlation between Y and its predicted value \hat{Y} .
- 3) Show that any solution to normal equations minimizes the residual sum of squares.

B) Answer the following questions. (Any One) 06

- 1) Describe cochrane-orkut method for parameter estimation in the presence of autocorrelation.
- 2) Propose an unbiased estimator of error variance σ^2 in the regression model and prove your claim.

Q.4 A) Answer the following questions. (Any Two) 10

- 1) Describe polynomial models in one variable and two variables.
- 2) Define mallow's c_p statistic and explain how it is used for variable selection in regression.
- 3) Describe detection of multicollinearity using variance inflation factor.

B) Answer the following questions. (Any One) 04

- 1) Define ridge estimator of regression coefficients. Obtain the mean square error of the ridge estimator.

2) Justify whether the following are linear models or not.

i) $Y = \alpha + \beta X$

ii) $Y = \alpha \beta \epsilon$

iii) $Y = \beta_0 + \beta_1 X + \epsilon$

iv) $Y = \alpha + \frac{\beta}{X} + \epsilon$

Where $\epsilon \sim iid N(0, \sigma^2)$

Q.5 Answer the following questions. (Any Two)

14

- 1) State and prove Gauss-Mark off theorem.
- 2) Describe multiple linear regression model stating the assumptions, obtain mean and variance of LSE $\hat{\beta}$ of β .
- 3) Define non-linear regression model. Discuss least squares method for parameter estimation in non-linear regression.



Seat No.	
-------------	--

M.Sc. (Part – II) (Semester – III) Examination, 2015
STATISTICS (Paper – XV) (Elective – II)
Regression Analysis (New CGPA)

Day and Date : Thursday, 26-11-2015

Max. Marks : 70

Time : 2.30 p.m. to 5.00 p.m.

- Instructions:** 1) Attempt **five** questions.
2) Q. No. (1) and Q. No. (2) are **compulsory**.
3) Attempt **any three** from Q. No. 3 to Q. No. 7.
4) Figures to the **right** indicate **full** marks.

1. A) Select the correct alternative :

5

- 1) The model $y = \theta_1 e^{\theta_2 x} + \epsilon$ is
- linear regression model
 - non-linear regression model
 - polynomial regression model
 - none of these
- 2) The sum of the residuals in any regression model with intercept β_0 is always
- positive
 - zero
 - non-zero
 - one
- 3) The variance of i^{th} press residual is
- $\frac{\sigma^2}{1 - h_{ii}}$
 - σ^2
 - $\sigma^2 (1 - h_{ii})$
 - $\frac{1 - h_{ii}}{\sigma^2}$
- 4) The multicollinearity problem in a multiple linear regression is concern with
- the error terms
 - response variable
 - the regressors
 - none of these

P.T.O.



5) Coefficient of determination R^2 is defined as

a) $\frac{SS_{\text{Regression}}}{SS_{\text{Total}}}$

b) $\frac{SS_{\text{residual}}}{SS_{\text{Total}}}$

c) $1 - \frac{SS_{\text{Regression}}}{SS_{\text{Total}}}$

d) None of these

B) Fill in the blanks :

5

- 1) Any model that is linear in the unknown parameters is called _____ regression model.
- 2) The hat matrix $H = x(x^1x^{-1})x^1$ is symmetric and _____ matrix.
- 3) _____ test is used to test the significance individual regression coefficient in linear regression model.
- 4) The regression model $y = \beta_0 + \beta_1x + \beta_2x^2$ is called polynomial regression model with _____ variable(s).
- 5) $\text{Cor} \left(\hat{\beta} \right) = \text{_____}$, $\hat{\beta}$ is OLS estimator of β .

C) State whether the following statements are **True** or **False** :

4

- 1) OLS estimator of regression coefficient is BLUE.
- 2) Condition indices of matrix $x'x$ is defined as $\frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} + \lambda_j$.
- 3) Auto correlation is concern with predictor variables.
- 4) Residuals are useful for detecting outlier observation in x-space.

2. a) Explain the terms :

- 1) Variance Inflation Factor (VIF)
- 2) Standardized and studentized residual.

b) Write short notes on the following :

- 1) Variable selection problem.
- 2) Box-cox power transformation.

(6+8)



3. a) Describe multiple linear regression model with K predictor variables. Write model in matrix form and state the basic assumptions. Derive the least square estimator of regression coefficients.
- b) In usual notations, outline the procedure of testing a general linear hypothesis $T\beta = 0$. **(7+7)**
4. a) Describe the problem of multicollinearity with suitable example. What are the effects of the same on least squares estimates of the regression coefficients.
- b) Define Mallows' C_p -Statistic and derive the same. **(7+7)**
5. a) Explain the following terms :
- 1) Influential observation
 - 2) Mallows' class of estimators
 - 3) Breakdown point.
- b) Define M-estimator and writedown the computational procedure of M-estimator. **(6+8)**
6. a) Describe the least square method for parameter estimation in non-linear regression. Discuss the same for $y = \theta_1 e^{\theta_2 x} + \epsilon$.
- b) Describe Cochran-Orkut method for parameter estimation in the presence of autocorrelation. **(7+7)**
7. a) Explain :
- i) Kernal regression
 - ii) Locally weighted regression.
- b) Discuss Durbin-Watson test for deteching auto correlation. **(7+7)**
-

Seat No.	
----------	--

M.Sc. (Semester - III) (CBCS) Examination March/April-2019
Statistics
REGRESSION ANALYSIS

Day & Date: Thursday, 02-05-2019
 Time: 03:30 PM To 06:00 PM

Max. Marks: 70

Instructions: 1) All questions are compulsory.
 2) Figures to the right indicate full marks.

Q.1 Choose the correct alternative from the given alternatives:- **14**

- 1) In classical linear regression, the distribution of response variable Y is _____.
 - a) Poisson
 - b) Binomial
 - c) Normal
 - d) Uniform
- 2) The hat matrix $X(X'X)^{-1}X'$ is _____.
 - a) idempotent
 - b) symmetric
 - c) both a and b
 - d) neither (a) nor (b)
- 3) The difference between observed value Y_i and corresponding fitted value \hat{Y}_i is called _____.
 - a) slope
 - b) intercept
 - c) error
 - d) residual
- 4) If distribution of Y is Poisson then variance stabilizing transformation used is _____.
 - a) $Y' = Y$
 - b) $Y' = \sqrt{Y}$
 - c) $Y' = \frac{1}{Y}$
 - d) $Y' = 109 Y$
- 5) The sum of residuals weighted by corresponding fitted values is always _____.
 - a) 0
 - b) Non-zero
 - c) One
 - d) None of these
- 6) If eigen values of matrix $(X'X)$ are 4.2, 0.3, 1.0 and 0.03. Then condition number is _____.
 - a) 14
 - b) 140
 - c) 1.4
 - d) None of these
- 7) The model $Y = \theta_1 e^{\theta_2 X} + \epsilon$ is _____ regression model.
 - a) Polynomial
 - b) Linear
 - c) Non-linear
 - d) None of these
- 8) In usual notation, $E[Cp/Bias = 0] =$ _____.
 - a) p
 - b) $p - 1$
 - c) $\frac{p(p+1)}{2}$
 - d) None of these
- 9) In a Linear model $Y = X\beta + \epsilon$, with usual assumptions on ϵ , the estimate $\hat{\beta}$ is _____.
 - a) uniquely obtained
 - b) not uniquely obtained
 - c) may or may not be uniquely obtained
 - d) none of these

B) Attempt any one of the following:-**04**

- 1) Explain Kernel regression.
- 2) Describe the Box-Cox method for transforming the response variable.

Q.5 Attempt any two of the following:-**14**

- a) Define multiple linear regression model. Obtain least squares estimator of β . Show that it is unbiased estimator β .
- b) In usual notations, outline the procedure of testing general linear hypothesis $T\beta = 0$
- c) Describe the problem of multicollinearity with suitable example. What are the effects of the same on LSE of regression coefficients?



Seat No.	
-------------	--

M.Sc. (Part – II) (Semester – III) Examination, 2014
STATISTICS (Paper – XV) (Elective – II)
Regression Analysis

Day and Date: Wednesday, 30-4-2014

Total Marks : 70

Time: 3.00 p.m. to 6.00 p.m.

- Instructions :** 1) Attempt **five** questions.
2) Q. No. (1) and Q. No. (2) are **compulsory**.
3) Attempt **any three** from Q. No. (3) to Q. No. (7).
4) Figures to the **right** indicate **full** marks.

1. A) Select correct alternative :

- 1) Least squares estimator in the linear regression model is
 - a) unbiased but not BLUE
 - b) unbiased and BLUE
 - c) biased
 - d) none of these
- 2) Autocorrelation is concerned with
 - a) correlation among regressor variables
 - b) correlation among error terms
 - c) correlation among response variable and regressor variables
 - d) none of these
- 3) If ρ is the simple correlation coefficient, the quantity ρ^2 is known as
 - a) coefficient of determination
 - b) coefficient of non-determination
 - c) coefficient of alienation
 - d) none of these
- 4) The least squares estimator for the model $y = x\beta + \epsilon$ can be written as
 - a) $\beta + (x'x)^{-1}\epsilon$
 - b) $\beta + (x'x)^{-1}x'\epsilon$
 - c) $\beta - (x'x)^{-1}x'\epsilon$
 - d) $\beta (x'x)^{-1}x'\epsilon$
- 5) The Hat matrix $x(x'x)^{-1}x'$ is
 - a) symmetric
 - b) idempotent
 - c) both (a) and (b)
 - d) neither (a) or (b)



B) Fill in the blanks :

- 1) The model $y = \beta_0 x^{\beta_1} \epsilon$ can be linearized by using _____ transformation.
- 2) Multiple correlation is a measure of _____ association of a variable with other variables.
- 3) Significance of a individual regression coefficient can be tested by _____ test.
- 4) The regression equation having two or more independent variables is called _____
- 5) $E(Cp/Bias = 0) = \text{_____}$, Cp is Mallows's Cp – statistic.

C) State whether following statements are **True** or **False** :

- 1) Variance of least squares estimator of β in linear regression model is $(x'x)^{-1} \sigma^2$.
- 2) Any model not linear in the unknown parameters is called simple linear regression model.
- 3) The sum of the residuals is any regression model with intercept (β_0) is always non-zero.
- 4) Residuals are useful in detecting outliers in Y-shape. **(5+5+4)**

2. a) Answer the following :

- i) Derive the relation between R^2 and adj R^2 .
- ii) Explain press residual. **(3+3)**

b) Answer the following :

- i) Describe a linear regression model with P-regressors. State all the assumptions.
- ii) Discuss the prediction interval for future observation in the context of multiple linear regression. **(4+4)**

3. a) Define problem of variable selection in linear regression. Describe forward selection method for the same.

b) Derive the null distribution of sample correlation coefficient. **(7+7)**



4. a) Define autocorrelation. Derive the Durbin-Watson test to determine the autocorrelation in errors.
- b) Briefly explain the Box-Cox transformation method. Also discuss the computational process of λ . **(7+7)**
5. a) What are the uses of residual plots ? Write a note on normal probability plot and plot of residuals against fitted values.
- b) What is multicollinearity ? Describe different sources of multicollinearity. **(7+7)**
6. a) Define residual and residual sum of squares. Propose an unbiased estimator of error variance σ^2 in the multiple linear regression and prove your claim.
- b) Describe the test procedure for testing $H_0 : T\beta = 0$ in the context of multiple regression. **(7+7)**
7. a) Discuss the least squares method for estimation of regression coefficients of linear and non-linear regression model.
- b) Write notes on the following :
- i) Variance inflation factor (VIF)
 - ii) Malow's Cp-statistic. **(7+7)**
-