Roll No. .....

Question Booklet Number

O. M. R. Serial No.

M. Sc. (Fourth Semester)

(NEP) EXAMINATION, 2025-26

STATISTICS

(Applied Regression Analysis)

Paper Code							
B	0	6	1	0	0	1	T

Questions Booklet Series
D

Time : 1:30 Hours] [*Maximum Marks : 75*

Instructions to the Examinee :

1. Do not open the booklet unless you are asked to do so.

2. The booklet contains 100 questions. Examinee is required to answer 75 questions in the OMR Answer-Sheet provided and not in the question booklet. All questions carry equal marks.

3. Examine the Booklet and the OMR Answer-Sheet very carefully before you proceed. Faulty question booklet due to missing or duplicate pages/questions or having any other discrepancy should be got immediately replaced.

परीक्षार्थियों के लिए निर्देश :

1. प्रश्न-पुस्तिका को तब तक न खोलें जब तक आपसे कहा न जाए।

2. प्रश्न-पुस्तिका में 100 प्रश्न हैं। परीक्षार्थी को 75 प्रश्नों को केवल दी गई OMR आन्सर-शीट पर ही हल करना है, प्रश्न-पुस्तिका पर नहीं। सभी प्रश्नों के अंक समान हैं।

3. प्रश्नों के उत्तर अंकित करने से पूर्व प्रश्न-पुस्तिका तथा OMR आन्सर-शीट को सावधानीपूर्वक देख लें। दोषपूर्ण प्रश्न-पुस्तिका जिसमें कुछ भाग छपने से छूट गए हों या प्रश्न एक से अधिक बार छप गए हों या उसमें किसी अन्य प्रकार की कमी हो, तो उसे तुरन्त बदल लें।

(Remaining instructions on the last page) **(शेष निर्देश अन्तिम पृष्ठ पर)**

(Only for Rough Work)

1. Nonlinear regression models are widely used in :

(A) Growth models

(B) Pharmacokinetics

(C) Population studies

(D) All of the above

2. The asymptotic variance of MLE is approximately equal to :

(A) Inverse of Fisher information

(B) Fisher information

(C) Square of Fisher information

(D) Log likelihood

3. An estimator that achieves the Cramér-Rao lower bound is called :

(A) Consistent estimator

(B) Efficient estimator

(C) Biased estimator

(D) Robust estimator

4. The Gauss-Newton algorithm is mainly used for :

(A) Linear regression

(B) Nonlinear regression estimation

(C) Time series analysis

(D) Sampling design

5. Nonlinear least squares estimators are obtained by minimizing :

(A) Sum of squared residuals

(B) Sum of absolute residuals

(C) Variance of residuals

(D) Mean of residuals

6. Mean square error (MSE) of an estimator is :

(A) Variance + Bias^2

(B) Variance – Bias^2

(C) Variance only

(D) Bias only

7. Which of the following properties ensures estimators become normally distributed for large samples ?

(A) Central Limit Theorem

(B) Law of Large Numbers

(C) Markov inequality

(D) Chebyshev inequality

8. Asymptotic efficiency refers to :

(A) Efficiency for small samples

(B) Efficiency when sample size approaches infinity

(C) Efficiency for biased estimators

(D) Efficiency of predictors

9. The Fisher Information measures :

(A) Information in predictors

(B) Amount of information data provides about parameters

(C) Variance of predictors

(D) Bias of estimator

10. The variance of an estimator can be approximated asymptotically using :

(A) Information matrix

(B) Covariance matrix

(C) Variance matrix

(D) Design matrix

11. Under regularity conditions, MLE is :

(A) Consistent

(B) Asymptotically normal

(C) Efficient

(D) All of the above

12. The asymptotic distribution of MLE is generally :

(A) Binomial

(B) Poisson

(C) Normal

(D) Gamma

13. Efficiency of an estimator refers to :

(A) Minimum variance among unbiased estimators

(B) Maximum bias

(C) Maximum likelihood

(D) Minimum sample size

14. An estimator is unbiased if :

(A) Its variance is zero

(B) Its expected value equals the true parameter

(C) It has minimum mean square error

(D) Sample size is large

15. Consistency of an estimator means :

(A) Estimator equals true value always

(B) Estimator converges to true value as sample size increases

(C) Estimator has minimum variance

(D) Estimator is unbiased

16. Which method is commonly used to obtain MLEs numerically ?

(A) Newton-Raphson method

(B) Fisher scoring

(C) EM algorithm

(D) All of the above

17. The likelihood function is based on :

(A) Distribution of explanatory variables

(B) Joint probability of observed data

(C) Mean of observations

(D) Variance of errors

18. When errors are normally distributed, the MLE of parameters coincides with :

(A) Method of moments

(B) Least squares estimator

(C) Bayesian estimator

(D) Ridge estimator

19. Maximum Likelihood Estimation (MLE) chooses parameter values that :

(A) Maximize the likelihood function

(B) Minimize variance

(C) Maximize mean

(D) Minimize bias

20. In nonlinear regression, least squares estimation often requires :

(A) Closed-form solution

(B) Iterative computation

(C) No assumptions

(D) Random sampling only

21. The Least Squares method estimates parameters by minimizing :

(A) Absolute errors

(B) Sum of squared residuals

(C) Sum of residuals

(D) Variance of predictors

22. Which of the following is a commonly used iterative method in nonlinear regression ?

(A) Newton-Raphson method

(B) Gauss-Newton method

(C) Gradient descent

(D) All of the above

23. Non-linear regression models are generally estimated using :

(A) Direct algebraic solution

(B) Iterative methods

(C) Graphical methods only

(D) Simple averaging

24. Which of the following is an example of a non-linear regression model ?

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

(B) $Y = \beta_0 + \beta_1 X + \beta_2 X^2$

(C) $Y = \beta_0 e^{\beta_1 X}$

(D) $Y = \beta_0 + \beta_1 X + \in$

25. A regression model is called non-linear when :

(A) The predictors are squared

(B) The parameters appear non-linearly in the model

(C) The response variable is discrete

(D) The errors are correlated

26. Which statistical software is commonly used for fitting GLM ?

(A) R

(B) SAS

(C) SPSS

(D) All of the above

27. A saturated model in log-linear analysis :

(A) Fits data perfectly

(B) Has no parameters

(C) Has minimum parameters

(D) Cannot be estimated

28. The deviance statistic in GLM is used for :

(A) Testing goodness of fit

(B) Estimating parameters

(C) Finding variance

(D) Measuring bias

29. Which method is commonly used for estimating parameters in log-linear models ?

(A) Least squares

(B) Maximum likelihood

(C) Bayesian estimation

(D) Ridge regression

30. In a two-way contingency table, independence implies :

(A) No interaction effect

(B) Perfect correlation

(C) Linear regression

(D) Equal means

31. The main purpose of log-linear models is :

(A) Predict continuous variables

(B) Analyze interactions in contingency tables

(C) Estimate means

(D) Reduce variance

32. Log-linear models analyze relationships between :

(A) Continuous variables

(B) Categorical variables

(C) Time variables

(D) Dependent variables only

33. Log-linear models assume that cell counts follow :

(A) Binomial distribution

(B) Poisson distribution

(C) Normal distribution

(D) Gamma distribution

34. Log-linear models are mainly used for :

(A) Continuous data

(B) Categorical data in contingency tables

(C) Time series

(D) Regression analysis

35. In logistic regression the relationship between predictors and probability is :

(A) Linear

(B) Non-linear

(C) Quadratic

(D) Cubic

36. Logistic regression is mainly used when the dependent variable is :

(A) Nominal

(B) Binary

(C) Continuous

(D) Ordinal

37. If the probability of success is 0.75, the odds are :

(A) 3

(B) 0.75

(C) 4

(D) 1.5

38. The odds of success are defined as :

(A) p

(B) $1 - p$

(C) $p/(1 - p)$

(D) $1/p$

39. For binary response data, the distribution assumed is :

(A) Normal

(B) Binomial

(C) Poisson

(D) Gamma

40. Grouped binary data refers to :

(A) Continuous observations

(B) Data grouped by categories with counts of successes and failures

(C) Time series observations

(D) Multivariate observations

41. In logistic regression the parameters are usually estimated by :

(A) Least squares

(B) Maximum likelihood

(C) Method of moments

(D) Bayesian estimation

42. The logit transformation is defined as :

(A) $\log(p)$

(B) $\log\left(\frac{p}{1-p}\right)$

(C) $\log(1-p)$

(D) $p/(1-p)$

43. The logistic function is given by :

(A) $p = e^x$

(B) $p = \frac{1}{1+e^{-x}}$

(C) $p = x^2$

(D) $p = \log(x)$

44. In logistic regression the response variable is :

(A) Continuous

(B) Count

(C) Binary

(D) Multivariate

45. The canonical link function for the Poisson distribution is :

(A) Identity

(B) Log

(C) Logit

(D) Probit

46. Which of the following is a common link function for binomial data ?

(A) Identity link

(B) Logit link

(C) Log link

(D) Square root link

47. The link function in GLM connects :

(A) Response variable and predictors

(B) Mean of response and linear predictor

(C) Error and predictors

(D) Variance and predictors

48. In GLM, the distribution of the response variable belongs to :

(A) Normal distribution family

(B) Exponential family

(C) Binomial family only

(D) Poisson family only

49. A Generalised Linear Model consists of :

(A) Random component

(B) Systematic component

(C) Link function

(D) All of the above

50. Generalised Linear Models (GLM) were introduced by :

(A) Fisher

(B) Nelder and Wedderburn

(C) Cochran

(D) Montgomery

51. Robust regression methods provide estimates that are :

(A) Highly sensitive to outliers

(B) Resistant to outliers

(C) Biased always

(D) Independent of data

52. Which method helps assess stability of a model ?

(A) Bootstrap

(B) Cross-validation

(C) Robust regression

(D) All of the above

53. The bootstrap distribution is obtained from :

(A) Original sample

(B) Repeated resampling of data

(C) Population data

(D) Theoretical distribution

54. Prediction error measures :

(A) Model complexity

(B) Model accuracy on new data

(C) Sample variance

(D) Mean of predictors

55. Which method is commonly used for model validation ?

(A) Cross-validation

(B) Regression

(C) Sampling

(D) Transformation

56. Robust regression is preferred when :

(A) Data contains outliers

(B) Data is perfectly normal

(C) Predictors are independent

(D) Sample size is small

57. The bootstrap estimate of standard error is obtained from :

(A) Variance of bootstrap estimates

(B) Mean of bootstrap estimates

(C) Median of bootstrap estimates

(D) Minimum of bootstrap estimates

58. Bootstrap methods are useful when :

(A) Sample size is large

(B) Analytical distribution is difficult to derive

(C) Data is perfectly normal

(D) Parameters are known

59. In bootstrap, each resample usually has size :

(A) Smaller than original sample

(B) Equal to original sample

(C) Twice the original sample

(D) Half the original sample

60. Bootstrap sampling involves :

(A) Sampling without replacement

(B) Sampling with replacement

(C) Random sampling only

(D) Stratified sampling

61. Bootstrap is mainly used for :

(A) Increasing sample size

(B) Estimating sampling distribution

(C) Eliminating predictors

(D) Removing errors

62. The bootstrap method was introduced by :

(A) Fisher

(B) Efron

(C) Cochran

(D) Neyman

63. Prediction error in cross-validation is calculated using :

(A) Training data only

(B) Test data only

(C) Both training and test data

(D) Predictor variables only

64. Leave-one-out cross-validation (LOOCV) is a special case of :

(A) 2-fold cross-validation

(B) k-fold cross-validation where $k = n$

(C) Bootstrap method

(D) Random sampling

65. In k-fold cross-validation, the model is trained :

(A) Once

(B) k times

(C) Twice

(D) n times

66. In k-fold cross-validation, the data is divided into :

(A) k equal parts

(B) k random variables

(C) k dependent variables

(D) k predictors

67. The main goal of cross-validation is to :

(A) Increase predictors

(B) Estimate prediction error

(C) Estimate parameters

(D) Reduce sample size

68. Huber's estimator is an example of :

(A) M-estimator

(B) Least squares estimator

(C) Maximum likelihood estimator

(D) Bayesian estimator

69. A common robust regression method is :

(A) Ridge regression

(B) Lasso regression

(C) M-estimation

(D) Principal component regression

70. Least squares estimation minimizes :

(A) L-1 norm

(B) L-2 norm

(C) L-∞ norm

(D) Log likelihood

71. The L-2 norm corresponds to :

(A) Sum of absolute residuals

(B) Sum of squared residuals

(C) Maximum likelihood function

(D) Absolute deviation

72. Minimization of L-1 norm leads to :

(A) Least squares estimator

(B) Least absolute deviations estimator

(C) Maximum likelihood estimator

(D) Ridge estimator

73. The L-1 norm is defined as :

(A) Sum of squared residuals

(B) Sum of absolute residuals

(C) Sum of residuals

(D) Product of residuals

74. Robust regression techniques are designed to be less sensitive to :

(A) Small samples

(B) Multicollinearity

(C) Outliers

(D) Large samples

75. Robust regression methods are mainly used to :

(A) Increase sample size

(B) Reduce effect of outliers

(C) Increase correlation

(D) Remove predictors

76. Model diagnostics in regression analysis involve :

(A) Checking residuals and influential points

(B) Increasing predictors

(C) Removing dependent variable

(D) Increasing variance

77. Multicollinearity refers to :

(A) Correlation among independent variables

(B) Correlation among dependent variables

(C) Correlation between residuals

(D) Correlation between errors

78. A good regression model should have residuals that are :

(A) Randomly distributed

(B) Increasing with predictors

(C) Decreasing with predictors

(D) Perfectly correlated

79. If $\lambda = 0$ in Box-Cox transformation, the transformation becomes :

(A) Square root transformation

(B) Logarithmic transformation

(C) Reciprocal transformation

(D) Exponential transformation

80. The Box-Cox transformation parameter is commonly denoted by :

(A) α

(B) β

(C) λ

(D) θ

81. Power transformations help to stabilize :

(A) Mean

(B) Variance

(C) Sample size

(D) Parameters

82. Square root transformation is often applied to :

(A) Count data

(B) Continuous data only

(C) Time series

(D) Multivariate data

83. A logarithmic transformation is commonly used when :

(A) Data grows exponentially

(B) Data decreases linearly

(C) Data is symmetric

(D) Data has constant variance

84. The Box-Cox transformation is used for :

(A) Transforming dependent variable

(B) Transforming independent variable only

(C) Removing outliers

(D) Increasing correlation

85. Power transformations are mainly used to :

(A) Increase observations

(B) Improve model assumptions

(C) Reduce sample size

(D) Eliminate predictors

86. Outliers in regression analysis are observations with :

(A) Large residuals

(B) Small residuals

(C) Zero residuals

(D) Equal residuals

87. An observation with high leverage and large residual is likely to be :

(A) Normal observation

(B) Influential observation

(C) Independent observation

(D) Random error

88. Leverage measures :

(A) Influence of response variable

(B) Influence of predictor values on fitted values

(C) Variance of residuals

(D) Mean of predictors

89. Cook's distance is used to detect :

(A) Multicollinearity

(B) Influential observations

(C) Correlation

(D) Mean values

90. An observation that strongly affects regression coefficients is called :

(A) Residual observation

(B) Influential observation

(C) Random observation

(D) Independent observation

91. Standardized residuals are used to :

(A) Detect outliers

(B) Estimate regression coefficients

(C) Calculate mean

(D) Measure correlation

92. If variance of errors changes with predictor values, the condition is called :

(A) Homoscedasticity

(B) Heteroscedasticity

(C) Multicollinearity

(D) Normality

93. Homoscedasticity means :

(A) Constant variance of errors

(B) Increasing variance

(C) Decreasing variance

(D) Zero variance

94. If residuals display a clear pattern in a residual plot, it indicates

(A) Good model fit

(B) Model inadequacy

(C) Large sample size

(D) No correlation

95. A residual plot is mainly used to detect :

(A) Linearity and homoscedasticity

(B) Sample size

(C) Mean of residuals

(D) Regression coefficients

96. Residuals in regression analysis are used to :

(A) Estimate mean

(B) Check model adequacy

(C) Increase sample size

(D) Estimate predictors

97. The difference between observed and predicted value is called :

(A) Deviation

(B) Residual

(C) Error variance

(D) Bias

98. In a simple linear regression model $Y = \beta_0 + \beta_1 X + \in$, the $\in$ represents :

(A) Predictor variable

(B) Random error term

(C) Parameter

(D) Constant

99. A regression model describes the relationship between :

(A) Two categorical variables

(B) Dependent and independent variables

(C) Only dependent variables

(D) Only independent variables

100. A statistical model is primarily used to :

(A) Increase sample size

(B) Represent relationships between variables

(C) Eliminate errors

(D) Reduce observations

(Only for Rough Work)

4. Four alternative answers are mentioned for each question as—A, B, C & D in the booklet. The candidate has to choose the correct answer and mark the same in the OMR Answer-Sheet as per the direction :

Example :

Question :

Q. 1 (A) ● (C) (D)
Q. 2 (A) (B) ● (D)
Q. 3 (A) ● (C) (D)

Illegible answers with cutting and over-writing or half filled circle will be cancelled.

5. Each question carries equal marks. Marks will be awarded according to the number of correct answers you have.

6. All answers are to be given on OMR Answer Sheet only. Answers given anywhere other than the place specified in the answer sheet will not be considered valid.

7. Before writing anything on the OMR Answer Sheet, all the instructions given in it should be read carefully.

8. After the completion of the examination candidates should leave the examination hall only after providing their OMR Answer Sheet to the invigilator. Candidate can carry their Question Booklet.

9. There will be no negative marking.

10. Rough work, if any, should be done on the blank pages provided for the purpose in the booklet.

11. To bring and use of log-book, calculator, pager and cellular phone in examination hall is prohibited.

12. In case of any difference found in English and Hindi version of the question, the English version of the question will be held authentic.

Impt. : On opening the question booklet, first check that all the pages of the question booklet are printed properly. If there is any discrepancy in the question Booklet, then after showing it to the invigilator, get another question Booklet of the same series.

4. प्रश्न-पुस्तिका में प्रत्येक प्रश्न के चार सम्भावित उत्तर—A, B, C एवं D हैं। परीक्षार्थी को उन चारों विकल्पों में से सही उत्तर छाँटना है। उत्तर को OMR आन्सर-शीट में सम्बन्धित प्रश्न संख्या में निम्न प्रकार भरना है :

उदाहरण :

प्रश्न :

प्रश्न 1 (A) ● (C) (D)
प्रश्न 2 (A) (B) ● (D)
प्रश्न 3 (A) ● (C) (D)

अपठनीय उत्तर या ऐसे उत्तर जिन्हें काटा या बदला गया है, या गोले में आधा भरकर दिया गया, उन्हें निरस्त कर दिया जाएगा।

5. प्रत्येक प्रश्न के अंक समान हैं। आपके जितने उत्तर सही होंगे, उन्हीं के अनुसार अंक प्रदान किये जायेंगे।

6. सभी उत्तर केवल ओ. एम. आर. उत्तर-पत्रक (OMR Answer Sheet) पर ही दिये जाने हैं। उत्तर-पत्रक में निर्धारित स्थान के अलावा अन्यत्र कहीं पर दिया गया उत्तर मान्य नहीं होगा।

7. ओ. एम. आर. उत्तर-पत्रक (OMR Answer Sheet) पर कुछ भी लिखने से पूर्व उसमें दिये गये सभी अनुदेशों को सावधानीपूर्वक पढ़ लिया जाये।

8. परीक्षा समाप्ति के उपरान्त परीक्षार्थी कक्ष निरीक्षक को अपनी OMR Answer Sheet उपलब्ध कराने के बाद ही परीक्षा कक्ष से प्रस्थान करें। परीक्षार्थी अपने साथ प्रश्न-पुस्तिका ले जा सकते हैं।

9. निगेटिव मार्किंग नहीं है।

10. कोई भी रफ कार्य, प्रश्न-पुस्तिका के अन्त में, रफ-कार्य के लिए दिए खाली पेज पर ही किया जाना चाहिए।

11. परीक्षा-कक्ष में लॉग-बुक, कैलकुलेटर, पेजर तथा सेल्युलर फोन ले जाना तथा उसका उपयोग करना वर्जित है।

12. प्रश्न के हिन्दी एवं अंग्रेजी रूपान्तरण में भिन्नता होने की दशा में प्रश्न का अंग्रेजी रूपान्तरण ही मान्य होगा।

महत्वपूर्ण : प्रश्नपुस्तिका खोलने पर प्रथमतः जाँच कर देख लें कि प्रश्न-पुस्तिका के सभी पृष्ठ भलीभाँति छपे हुए हैं। यदि प्रश्नपुस्तिका में कोई कमी हो, तो कक्षनिरीक्षक को दिखाकर उसी सिरीज की दूसरी प्रश्न-पुस्तिका प्राप्त कर लें।