

B.Tech II Year II Semester (R20) Regular & Supplementary Examinations August/September 2023

STATISTICAL METHODS FOR DATA SCIENCE

CSE (Data Science)

Time: 3 hours

Max. Marks: 70

PART – A
(Compulsory Question)

1 Answer the following: (10 X 02 = 20 Marks)

- | | |
|---|----|
| (a) What is a data science? | 2M |
| (b) Define statistic. | 2M |
| (c) Write the any two importance of estimator. | 2M |
| (d) Write the any two properties of ML estimator. | 2M |
| (e) What is duality? | 2M |
| (f) Define population. | 2M |
| (g) Write equation for standard distributions. | 2M |
| (h) Why testing is important? | 2M |
| (i) Expand CRD. | 2M |
| (j) What is t-test? | 2M |

PART – B

(Answer all the questions: 05 X 10 = 50 Marks)

- | | | |
|-----------|--|-----|
| 2 | Explain in detail about Lehman scheffe's theorem with example. | 10M |
| OR | | |
| 3 | (a) List out the characteristics of a good estimator. | 5M |
| | (b) What is meant by minimal sufficiency? Explain it in detail. | 5M |
| 4 | Discuss in detail about maximum likelihood method. | 10M |
| OR | | |
| 5 | Explain in detail about large sample properties of ML estimator. | 10M |
| 6 | Discuss in detail about construction of confidence intervals for population proportion with large samples. | 10M |
| OR | | |
| 7 | (a) Write the difference between the mean and ration of two normal populations. | 5M |
| | (b) Write short notes on confidence level for mean. | 5M |
| 8 | Explain in detail about Neyman-Pearson fundamental lemma and its applications. | 10M |
| OR | | |
| 9 | Write the description and property of LR tests. | 10M |
| 10 | Discuss about equality of two population means with any one example. | 10M |
| OR | | |
| 11 | Explain the following: (i) Mann-Whitney test, (ii) Run test. | 10M |

B.Tech II Year II Semester (R20) Regular & Supplementary Examinations April/May 2024

STATISTICAL METHODS FOR DATA SCIENCE

(CSE(Data Science))

Time: 3 hours

Max. Marks: 70

PART – A

(Compulsory Question)

1 Answer the following: (10 X 02 = 20 Marks)

- | | |
|---|----|
| (a) Explain likelihood equivalence and its significance in estimation. | 2M |
| (b) Define the Uniformly Minimum Variance Unbiased Estimator (UMVUE). | 2M |
| (c) Discuss the maximum likelihood method of point estimation. | 2M |
| (d) Define the method of minimum chi-square for point estimation. | 2M |
| (e) Discuss the construction of confidence intervals for the variance of a normal population. | 2M |
| (f) Define confidence limits and confidence coefficient in the context of confidence intervals. | 2M |
| (g) Define and differentiate between Type I and Type II errors in hypothesis testing. | 2M |
| (h) Describe the Neyman-Pearson Fundamental Lemma and its applications in hypothesis testing. | 2M |
| (i) Discuss the procedure for conducting a test for the equality of two population means. | 2M |
| (j) Explain the F-test for the equality of two population variances and its assumptions. | 2M |

PART – B

(Answer all the questions: 05 X 10 = 50 Marks)

- 2 Discuss the applications of Lehmann-Scheffe's Theorem in constructing unbiased estimators with minimum variance. Explain the conditions under which Lehmann-Scheffe's Theorem can be applied and provide examples to illustrate its use in practice. 10M

OR

- 3 Define and distinguish between population, sample, parameter, and statistic in the context of statistical inference. Discuss their roles and relationships in estimating population characteristics. Provide examples to illustrate each concept. 10M

- 4 Evaluate the performance of different point estimation methods in terms of efficiency, robustness, and computational complexity. Discuss strategies for selecting the most appropriate estimation method based on the characteristics of the data and the research objectives. 10M

OR

- 5 Discuss the role of assumptions and model specifications in point estimation methods. Explain how violations of assumptions can affect the accuracy and reliability of estimators. 10M

- 6 Explain the procedure for determining confidence intervals for the difference between the means of two normal populations. Discuss the assumptions and conditions required for the validity of such intervals. 10M

OR

- 7 Discuss the method for constructing confidence intervals for the ratio of two normal populations. Highlight the differences between constructing confidence intervals for the difference between means and the ratio of means. 10M

- 8 Explain the procedure for conducting a Likelihood Ratio test for a given hypothesis testing problem. Discuss how to calculate the test statistic and determine the critical region for rejection. 10M

OR**Contd. in Page 2**

- 9 Compare and contrast the power of different statistical tests for a given hypothesis testing problem. Discuss the trade-offs between Type I and Type II errors in selecting an appropriate test. 10M
- 10 Compare and contrast the methodologies of conducting t-tests for the equality of two population means, including independent samples and paired samples. Discuss the implications of each approach and provide examples to demonstrate their application. 10M
- OR**
- 11 Explain the F-test for the equality of two population variances. Discuss the assumptions of the F-test and how it is used to assess whether the variances of two populations are equal. 10M
