

Indian Statistical Institute
First Semester Examination, 2025-26
Post Graduate Diploma in Business Analytics
Inference

Date: 12/12/2025

Maximum Marks: 50

Duration: 1.5 Hours

Part A

Q1. Suppose you are testing out the effectiveness of a gel to reduce lower back pain. After applying the gel and resting for 5 minutes, the participants say one of the following about the performance of the gel: 'completely relieved of the pain', 'partially relieved of the pain', or 'not relieved of the pain'. You collect data from 99 such participants. [2+2+3+(3+3)+6+(2+2)=23]

- a) What is the type of data being collected here?
- b) What would be a suitable distribution for this data?
- c) Suppose you want to check that all of the three possible responses are equally likely to occur or not. What would the concerned hypotheses look like in terms of the parameter(s) involved here?
- d) What testing procedure would you want to follow here to test the claim? What should the rejection region look like?
- e) Suppose 15 people said 'completely relieved of the pain', 45 people said 'partially relieved of the pain', and the remaining people said 'not relieved of the pain' in their responses. What would be the value of your test statistic?
- f) What would your cut-off be for a 5%-level test, and what would your verdict be, based on that cutoff and your test statistic? You may use the following: $z_{0.95} = 1.64$, $z_{0.975} = 1.96$, $\chi^2_{0.95}(2) = 5.99$, $\chi^2_{0.95}(3) = 7.81$, $\chi^2_{0.95}(4) = 9.49$, $\chi^2_{0.975}(2) = 7.38$, $\chi^2_{0.975}(3) = 9.35$, $\chi^2_{0.975}(4) = 11.14$.

Q2. Consider a dataset with 145 observations on the electricity usage (in kWh) of the residents of Baranagar. 24 of these observations are lost while reading the data from the thumb drive because of the corresponding files being corrupted. You would assume that electricity usage of consumers follows a normal distribution with mean μ and variance σ^2 . [(3+2)+(7+3)+6+6=27]

- a) What kind of missingness is there in the data? How would you deal with the missingness prior to starting the analysis?
- b) What would be the maximum likelihood estimators for μ and σ^2 , assuming both are unknown? Are the estimators based on method of moments and the estimators based on method of maximum likelihood different?
- c) What would be a 95% two-sided confidence interval for μ here? You may use the quantities provided in Question 1.f)
- d) What would be a 95% two-sided confidence interval for σ^2 here? You may use the following: $\chi^2_{0.025}(145) = 113.56$, $\chi^2_{0.025}(144) = 112.67$, $\chi^2_{0.025}(121) = 92.45$, $\chi^2_{0.025}(120) = 91.57$, $\chi^2_{0.975}(145) = 180.23$, $\chi^2_{0.975}(144) = 179.13$, $\chi^2_{0.975}(121) = 153.34$, $\chi^2_{0.975}(120) = 152.21$.