

MATH40005

Probability and Statistics

Revised Mock Examination 2

Time allowed	2 hours
Questions	4
Total marks	80
Format	Closed book; justify all answers

- Answer all four questions.
- Start each question on a new page in an actual timed attempt.
- State assumptions and named results used.
- Statistical tables or stated critical values may be used where supplied.

Question 1

- 1(a). Passengers use routes A,B,C with probabilities 0.5,0.3,0.2. Their late probabilities are 0.04,0.10,0.05 respectively. Find the probability that a randomly selected passenger is late. (5 marks)
- 1(b). Given that the passenger is late, find the probability that route A was used. (3 marks)
- 1(c). A signal failure now makes every train on routes B and C late, while route A keeps late probability 0.04. Find $P(A|L)$ under the new conditions. (4 marks)
- 1(d). A 10-character password contains 5 letters (52 choices, repetition allowed), 3 digits (10 choices, repetition allowed), and 2 distinct symbols from 5 symbols. Count all passwords. (5 marks)
- 1(e). State the most suitable plot for each purpose: (i) the relationship between house price and number of bedrooms; (ii) median, spread and outliers of prices in one area; (iii) daily stock prices over three years. (3 marks)

Total: 20 marks

Question 2

2(a). Let $U \sim \text{Uniform}(0,1)$ and $\lambda > 0$. Identify the distribution of $T = -\log(U)/\lambda$ by deriving its CDF. (4 marks)

2(b). For independent $X_1, \dots, X_n \sim \text{Uniform}(0,1)$, find the CDF and density of $M = \max(X_1, \dots, X_n)$. (4 marks)

2(c). Let X be standard normal. Derive its moment generating function. (4 marks)

2(d). If $X_1, \dots, X_n$ are independent standard normal variables, use MGFs to identify $S = \sum X_i$. (4 marks)

2(e). Suppose $E[U^k] = E[V^k]$ for every non-negative integer k . State a sufficient additional condition under which U and V must have the same distribution, and justify. (4 marks)

Total: 20 marks

Question 3

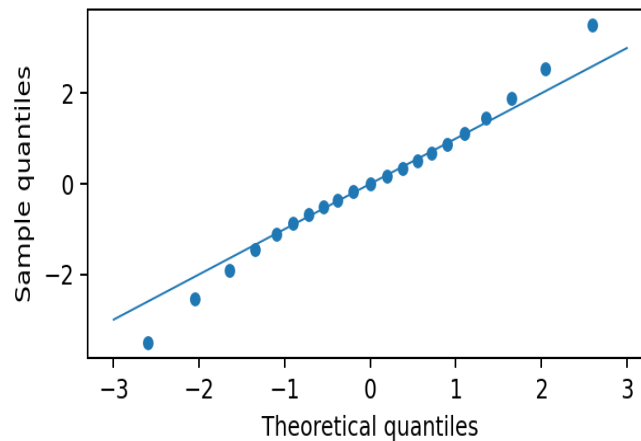
3(a). Independent $X_i \sim N(\mu_i, \sigma^2)$ may have different means. For $\bar{X} = n^{-1} \sum X_i$, prove $\text{Var}(\bar{X}) = \sigma^2/n$. (2 marks)

3(b). With $S^2 = (n-1)^{-1} \sum (X_i - \bar{X})^2$ and $\bar{\mu} = n^{-1} \sum \mu_i$, derive $E[S^2]$. (6 marks)

3(c). For $n=12$ normal observations with $\bar{x}=181$ and $s^2=3$, compute a 90% confidence interval for μ . Use $t_{(11,0.95)}=1.796$. (4 marks)

3(d). Using the same sample, compute a 95% confidence interval for σ^2 . Use $\chi^2_{(11,0.025)}=3.816$ and $\chi^2_{(11,0.975)}=21.920$. (4 marks)

3(e). Interpret the normal Q-Q plot shown. Does it support a normal model? (4 marks)



Total: 20 marks

Question 4

4(a). Independent observations are modeled as $\text{Uniform}[\theta_1, \theta_2]$, where $0 < \theta_1 < \theta_2$ are unknown and the observed sample is not constant. Find the MLEs. (5 marks)

4(b). In simple linear regression with fixed x_i and independent errors of variance σ^2 , prove $S_{xy} = \sum (x_i - \bar{x})y_i$ and derive $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\sigma^2 \bar{x} / S_{xx}$. (4 marks)

4(c). Describe how to construct a percentile bootstrap 95% confidence interval for a population median from an observed sample of size n . (4 marks)

4(d). Rugby and basketball heights are independent normal samples. The population variances are known: $\sigma_1^2 = 10$ and $\sigma_2^2 = 4$. Data give $\bar{x} = 182, n = 15$ and $\bar{y} = 185, m = 12$. Test $H_0: \mu_1 = \mu_2$ against a two-sided alternative at $\alpha = 0.01$, compute the p-value, and conclude. (7 marks)

Total: 20 marks