

MATH40005

Probability and Statistics

Revised Mock Examination 5

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). For $f:U \rightarrow V$ and A subset V , prove $I_A(f(u)) = I_{f^{-1}(A)}(u)$. (2 marks)

1(b). If G is a sigma-algebra on V , prove $f^{-1}(G) = \{f^{-1}(A) : A \in G\}$ is a sigma-algebra on U . (6 marks)

1(c). Construct three Bernoulli random variables that are pairwise independent but not mutually independent. (4 marks)

1(d). Let $\Omega = \{1,2,3,4\}$ and $F = \{\emptyset, \Omega, \{1,2\}, \{3,4\}\}$. Define $X(1)=X(3)=0$ and $X(2)=X(4)=1$. Is X measurable? Give a non-constant measurable random variable on the same space. (4 marks)

1(e). A fair six-sided die is rolled seven times. Give a probability space and the cardinality of its power set. (4 marks)

Total: 20 marks

Question 2

2(a). Let $X \sim N(0,1)$, $Y \sim \text{Gamma}(\alpha, \beta)$ in rate form, independently, and $Z = \sqrt{Y}X$. Find the MGF of Z and its domain. (7 marks)

2(b). Let X be continuous with density f_X and let independent Y take values $1, \dots, 10$. For $Z = XY$, derive f_Z and verify it is a density. (6 marks)

2(c). A joint density is constant c on $0 < x < y < z < 1$. Find c and $E[XYZ]$. (7 marks)

Total: 20 marks

Question 3

3(a). Ten cups contain five tea and five coffee. A person selects five cups as tea and identifies four correctly. Under random guessing, perform a one-sided exact test at level 0.05. (5 marks)

3(b). For the data {14.3, 8.3, 15.9, 12.7, 9.4, 24.0, 7.2, 6.1, 11.6, 10.5}, compute the lower and upper quartiles using the median-of-halves convention, then apply Tukey's rule for outliers. (4 marks)

3(c). Combined data favor a placebo, but within each of two risk groups a drug has a lower disease rate. Name the phenomenon and explain how it can occur. (3 marks)

3(d). Potato weights have unknown distribution and only the bound $\sigma < 10$ g is known. Using the bound 10 g to construct a distribution-free 99% Chebyshev interval, find the minimum integer n that guarantees its total width is strictly less than 4 g. (5 marks)

3(e). Define Type I and Type II errors and explain qualitatively how increasing α affects them. (3 marks)

Total: 20 marks

Question 4

- 4(a). Independent positive observations have lognormal density proportional to $\exp[-(\log x - \theta)^2 / (2\sigma^2)]/x$, where σ^2 is known. Find the MLE of θ . (6 marks)
- 4(b). Show that the fitted least-squares line in simple linear regression passes through $(\bar{x}, \bar{y})$. (4 marks)
- 4(c). Suppose all non-negative integer moments of U and V agree. State a sufficient condition for equality in distribution and justify. (4 marks)
- 4(d). Describe a percentile bootstrap 95% confidence interval for a median. (3 marks)
- 4(e). For i.i.d. observations with mean μ and variance σ^2 , construct an unbiased estimator of μ^2 using $\bar{X}$ and the unbiased sample variance S^2 . (3 marks)

Total: 20 marks