30

(University of Choice)

MASINDE MULIRO UNIVERSITY OF SCIENCE AND TECHNOLOGY

(MMUST)

MAIN CAMPUS

MAIN EXAMINATIONS

2023/2024 ACADEMIC YEAR

FOURTH YEAR FIRST SEMESTER EXAMINATION

FOR THE DEGREE OF BACHELOR OF SCIENCE IN EPIDEMIOLOGY AND BIOSTATISTICS

COURSE CODE: HEM 411

COURSE TITLE: BAYESIAN INFERENCE AND DECISION ANALYSIS

DATE: 5/12/2023 TIME: 8.00-10.00 AM

INSTRUCTIONS TO CANDIDATES: ANSWER QUESTION ONE and ANY TWO QUESTIONS

TIME: 2 Hours

Paper Consists of 2 Printed Pages. Please Turn Over

This Paper Consists of 4 Printed Pages. Please Turn Over.

QUESTION ONE (30marks)

a) Define the following terms
   i. Prior distribution
   ii. Posterior distribution
   iii. Bayes estimate. (6marks)

b) Let $x_1, x_2, \cdots, x_n$ denote a random sample from $f\left(x/_\theta\right) = \theta^x(1-\theta)^{1-x}; x = 0,1$. Suppose that the parameter space is uniformly distributed as $f(\theta) = \begin{cases} 1 & ;0 < \theta < 1 \\ 0 & ;otherwise \end{cases}$. Find the Bayes estimate of θ (9marks)

c) Define the term inferential statistics and give its types. (2marks)

d) 1% of the population has X disease. A screening test accurately detects the disease for 90% of people with it. The test also indicates the disease for 15% of people without it (false positive). Suppose a person screened for the disease tests positive. What is the probability that they have it? (5 marks)

e) Discuss four factors to consider while choosing a prior distribution. (8 marks)

QUESTION TWO (20marks)

a) Discuss the various types of priors (10marks)

b) 1% of people has the tropical disease Z. There exist a test that gives positive 80% of the times if the disease is present (true positive) but also 10% of the times when the disease is absent (false positive). If a person tests positive, what is the probability that he/she has the Z disease? (5marks)

c) Let $X_1 \ldots X_n$ denote a random sample from the Bernoulli distribution when $p(X_i = 1) = p \, and \, p(X_i = 1) = 1 - p$
   Assume that the prior distribution for p is $Beta(\alpha, \beta)$. Find the posterior distribution (5 marks)

QUESTION THREE (20marks)

a) According MMUST, the probability that a student will graduate within 4 years since he began his studies is 0.8. The score in HEM, which takes values in [0,35] of the students who graduate within 4 years follow a normal distribution with mean value of 26 and standard deviation of 2. On the contrary, the score of the students who do not graduate within 4 years follow a normal distribution with mean value of 22 and standard deviation 3. Let G denote the event that a student graduates within 4 years and let nG denote the event that a student does not graduate within 4 years. Finally, let x denote the score of a student in HEM.

i. If we know that 80% of the students that enter MMUST within 4 years, what is the probability that a student will graduate within 4 years given that his HEM score is 21.? (5marks)

ii. What is the minimum HEM score that a student must have so as to maximize the possibility of graduating within 5 years? (10marks)

b) Let $x_1, x_2, \cdots, x_n$ denote a random sample from $f\left(x/\theta\right) = \frac{e^{-\theta}\theta^x}{x!}; x \geq 0$. Assume that θ follows the distribution $\frac{1}{\Gamma\alpha\beta^\alpha}\theta^{\alpha-1}e^{\frac{-\theta}{\beta}}; \theta > 0, \alpha > 0, \beta > 0$. Find the posterior distribution (5 marks)

QUESTION FOUR (20mark)

a) Suppose that X follows binomial distribution with parameters n and p

i. If $h(p) = 1$, find the Bayes estimate of p using squared error loss function $L(p,t) = (t-p)^2$ (5marks)

ii. If $p \sim Beta(\alpha, \beta)$; $0 \leq p \leq 1$, find the Bayes estimate of p and variance of p (7marks)

b) Explain the differences between frequentist and Bayesian methods of statistical analysis. (8 marks)

QUESTION FIVE (20marks)

a) Let $x_1, x_2, \cdots, x_n$ denote the random sample from $\theta^x(1-\theta)^{1-x}; x = 0,1$. Assume that $\theta \sim Beta(\alpha, \beta)$. Find:

i. The marginal distribution (4marks)

ii. The posterior distribution function (2marks)

iii. The Bayes estimate of θ (3marks)

iv. The variance of θ (4marks)

b) Define Bayesian inference giving its importance (3 marks)

c) In hospital, patients that may have the Hepatitis C virus are submitted in tests in order to detect the virus in their blood. Let:

H: denote the event that a patient is infected with the virus

nH: denote the event that a patient is not infected with the virus

Pos: denote the event that a patient test to the virus is positive

Neg: denote the event that a patient's test to the virus is negative

Practically, the following probabilities have been determined

$$P(H) = 0.15$$

$$P(Pos \mid H) = 0.95$$

$$P(Pos \mid nH) = 0.02$$

Compute the probability that a patient is infected with the virus given that his test is positive (4 marks)