(Knowledge for Development)

KIBABII UNIVERSITY

UNIVERSITY EXAMINATIONS

2020/2021 ACADEMIC YEAR

FOURTH YEAR SECOND SEMESTER

SPECIAL/SUPPLIMENTARY EXAMINATION

FOR THE DEGREE OF BACHELOR OF EDUCATION AND BACHELOR OF SCIENCE

(MATHEMATICS)

COURSE CODE: **STA 446**

COURSE TITLE: **BAYESIAN STATISTICS**

DATE: 18/01/2021 **TIME:** 8 AM -10 AM

INSTRUCTIONS TO CANDIDATES

Answer Question One and Any other TWO Questions

TIME: 2 Hours

QUESTION ONE (30 MARKS)

1. (a) State and explain any three advantages of Bayesian estimation in statistical modelling and data analysis (6 mks)

   (b) Differentiate between Informative and Non-informative priors (4 mks)

   (c) In order to determine how effective a magazine is reaching its target audince, a market research company selects a random sample of people from the target audience and interviews them. Out of 150 people in the sample, 29 had seen the latest issue.

      i. What is the distribution of y, the number who have seen the latest issue? (2 mks)

      ii. Use the uniform prior for π, the proportion of the target audience that has seen the latest issue. What is the posterior distribution of π? (3 mks)

   (d) Let $X \sim B(n, p)$. Assume the prior distribution of p is uniform on [1,0]. Show that the posterior is essentially the likelihood function. (5 mks)

   (e) Suppose $X_1, ..., X_n$ is a sample from geometric distribution with parameter p, $0 \leq p \leq 1$. Assume that the prior distribution of p is beta with $a = 4$ and $b = 4$. Find

      i. the posterior distribution of p (3 mks)

      ii. the Bayes estimate under quadratic loss function (2 mks)

   (f) An urn containing a total of 5 balls, some of which are red and the rest of which are green. Let the random variable X be the number of red balls in the urn. Find the Bayesian estimate of X (5 mks)

QUESTION TWO (20 MARKS)

2. (a) Define a Bayesian Credible interval (2 mks)

   (b) Suppose $P(B) = 0.02$, $P(T|B) = 0.99$ and $P(T|B^c) = 0.05$. Calculate:

      i. $P(T)$ (3 mks)

      ii. $P(B|T)$ (2 mks)

      iii. $P(B|T^c)$ (3 mks)

   (c) Let $Y|\pi$ be binomial $(n = 4, \pi)$. Suppose we consider that there are only three possible values for π, 0.4, 0.5 and 0.6 and that $Y = 3$. Find the posterior probabilty distribution of π (10 mks)

QUESTION THREE (20 MARKS)

3. (a) In a research program of human health from recreational contactwith contaminated with pathogenic microbiological material, the National Environmental Management Authority (NEMA) instituted a study to determine the quality of Kenya stream water in a variety of catchment types in Nakuru County. This study showed that where $n = 116$ one-litre water samples from sites identified as having a heavy environmental impact from birds (flamingo) and waterfowl. Out of these samples, $y = 17$ samples contained *Giardia cysts.*

      i. What is the distribution of y, the number of samples containing *Giardia cysts* (2 mks)

      ii. Let π be the true probability that a one-litre water sample from this type of site contains *Giardia cysts.* Use a beta(1,4) prior for π. Find the posterior distribution of π given y. (2 mks)

      iii. Summarize the posterior distribution by its first two moments (4 mks)

      iv. Find the normal approximation to the posterior distribution $g(\pi|y)$ (2 mks)

      v. Compute a 95 percent credible interval for π using the normal approximation found in part (iv) (5 mks)

(b) Suppose X is a normal random variable with μ and variance σ^2, where σ^2 is known and μ is unknown. Suppose μ behaves as a r.v whose probability distribution is $\pi(\mu)$ and is normally distributed with mean μ_p and variance σ_p^2 both assumed to be estimated. Find the mean and variance of the posterior pdf $f(\mu|X)$ if $\mu_p = 50$, $\sigma_p = 6$ and $\sigma = 5$, $x = 52$. (5 mks)

QUESTION FOUR (20 MARKS)

4. (a) In the past million days, the sun has been predicted to rise the next morning or not. Each evening, the sun is said to rise the next morning ($\widehat{R}$) and found right (R) all these days. Suppose on the 10^6 evenings it was predicted that the sun will rise on the next day. What is the probability that the sun will rise the next day? (12 mks)

   (b) Suppose $x_1, ..., x_n$ is a random sample from $N(\mu, \sigma^2)$ with $\sigma^2 = 4$. If the prior pdf of μ is $N(0, 1)$ that is, $\pi(\mu) \sim N(0, 1)$. Find 95 percent credible interval of μ. (8 mks)

QUESTION FIVE (20 MARKS)

5. A researcher measured heart rate (z) and oxygen uptake (y) for one person under varying exercise conditions. He wishes to determine if heart rate, which is easier to measure, can be used to predict oxygen uptake. If so, then the estimated oxygen uptake based on the measured heart rate can be used in place of the measured oxygen uptake for later experiments on the individual:

Heart rate, x	94	96	94	95	104	106	108	113	115	121	131
Oxygen Uptake, y	0.47	0.75	0.83	0.98	1.18	1.29	1.40	1.60	1.75	1.90	2.23

(a) Plot a scatterplot of oxygen uptake y versus heart rate x.

(b) Calculate the parameters of the least squares line.

(c) Graph the least squares line on your scatterplot.

(d) Calculate the estimated variance about the least squares line.

(e) Suppose that we know that oxygen uptake given the heart rate is $(\alpha_0 + \beta \times x\sigma^2)$, where $\sigma^2 = 0.13^2$is known. Use a normal $(0, l^2)$ prior for β. What is the posterior distribution of β?