(Knowledge for Development)

KIBABII UNIVERSITY

UNIVERSITY EXAMINATIONS

2024/2025 ACADEMIC YEAR

FIRST YEAR FIRST SEMESTER

MAIN EXAMINATION

FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS

COURSE CODE: **STA 8130E**

COURSE TITLE: **BAYESIAN INFERENCE**

DATE: 23/01/2025 **TIME:** 8:00 AM - 11:00 AM

INSTRUCTIONS TO CANDIDATES

Answer Any THREE Questions

TIME: 2 Hours

QUESTION ONE (40 MARKS)

1. (a) Bowl A contains 30 vanilla cookies and 10 chocolate cookies. Bowl B contains 20 of each. A bowl is choosen at random and a cookie picked at random, what is the probability that it come from bowl A if the cookie is vanilla? (5 mks)

   (b) State and explain any four advantages of using Bayesian Statistics. (4 mks)

   (c) Describe the procedure for finding the Bayesian estimate (5 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 (2 mks)

      ii. the Bayes estimate under quadratic loss function (3 mks)

   (f) Let $X_1, ..., X_n$ be $N \sim (\mu, \sigma^2)$ random variables with prior $\pi(\mu)$ having $N(\mu_0, \sigma_0^2)$ distribution with known σ^2. Obtain the posterior distribution of μ. (6 mks)

   (g) Let $X \sim N(\mu, \sigma^2)$, where σ^2 is known and μ is unknown. If μ behaves as a random variable whose probability distribution is $\pi(\mu)$ and also normally distributed with mean μ_p and variance σ_p^2, both assumed to be known. Find the:

      i. mean of the posterior density, $f(\mu|x)$ (5 mks)

      ii. variance of the posterior density, $f(\mu|x)$ (5 mks)

QUESTION TWO (20 MARKS)

2. (a) The cross-fertilized plant produce taller offspring than the self-fertilized plants. In order to obtain an estimate on the proportion θ of cross-fertilized that are taller, a researcher observes a random sample of 15 pairs of plants that are exactly the same age. Each pair is grown in the same conditions with some cross-fertilized and others are self-fertilized. Based on previous experience, the researcher believes that the following are possible values of θ and that the prior probability for each value of θ is $p(\theta)$.

θ	0.80	0.82	0.84	0.86	0.88	0.90
$p(\theta)$	0.13	0.15	0.22	0.25	0.15	0.10

   i. Find the Bayesian estimate for the informative prior (8 mks)
   ii. Find the Bayesian estimate for the non-informative prior (4 mks)
   iii. Comment on the Bayesian estimate in (i) and (ii) above (2 mks)

   (b) 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 (2 mks)
   ii. the Bayes estimate under quadratic loss function (4 mks)

QUESTION THREE (20 MARKS)

3. (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 FOUR (20 MARKS)

4. 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 scatter plot 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?