(Knowledge for Development)

KIBABII UNIVERSITY

UNIVERSITY EXAMINATIONS

2016/2017 ACADEMIC YEAR

FOURTH YEAR SECOND SEMESTER

SPECIAL/ SUPPLEMENTARY EXAMINATION

FOR THE DEGREE OF BACHELOR OF SCIENCE

MATHEMATICS

COURSE CODE: **STA 446**

COURSE TITLE: **BAYESIAN STATISTICS**

DATE: 20/09/17 **TIME:** 11.30 AM -1.30 PM

INSTRUCTIONS TO CANDIDATES

Answer Question One and Any other TWO Questions

TIME: 2 Hours

QUESTION ONE (30 MARKS)

1. (a) What are the hyperparameters, Loss function and Mean square error? What do they represent? (6 mks)

   (b) Differentiate between Bayesian and Frequentist analysis (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) 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)

   (f) 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 TWO (20 MARKS)

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

   (b) What is the squared-error loss Bayes estimate for the parameter θ in a binomial pdf, where θ has a uniform distributionthat is, a noninformative prior? (Recall that a uniform prior is a beta pdf for which $r = s = 1$.) (8 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 $X_1, ..., X_n$ 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)

QUESTION FOUR (20 MARKS)

4. (a) Suppose that X is a geometric random variable, where $p_X(k|\theta) = (1-\theta)^{k-1}\theta, k = 1, 2, \ldots$. Assume that the prior distribution for θ is the beta p.d.f. with parameters r and s. Find the posterior distribution for θ. (6 mks)

(b) Suppose the binomial pdf describes the number of votes a candidate might receive in a poll conducted before the general election. Moreover, suppose a beta prior distribution has been assigned to θ, and every indicator suggests the election will be close. The pollster, then, has good reason for concentrating the bulk of the prior distribution around the value $\theta = \frac{1}{2}$. Setting the two beta parameters r and s both equal to 135 will accomplish that objective (in the event $r = s = 135$, the probability of θ being between 0.45 and 0.55 is approximately 0.90).

i. Find the corresponding posterior distribution. (4 mks)

ii. Find the squared-error loss Bayes estimate for θ and express it as a weighted average of the maximum likelihood estimate for θ and the mean of the prior pdf. (4 mks)

(c) Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a gamma pdf with parameters r and θ, where the prior distribution assigned to θ is the gamma pdf with parameters s and μ. Let $W = Y_1 + Y_2 + \ldots + Y_n$. Find the posterior pdf for θ. (6 mks)

QUESTION FIVE (20 MARKS)

5. Let $X_1, X_2, \ldots, X_n$ be $N(\mu, \sigma^2)$ random variables with prior $\pi(\mu)$ having $N(\mu_0, \sigma_0^2)$ distribution with known σ^2.

   (a) Obtain the posterior distribution of μ. (8 mks)

   (b) Suppose it is known from past experience that the weight loss for a particular combination of diet and exercise program (if followed for a month) is normally distributed with mean 10 lb and standard deviation of 2 lb. A random sample of five persons who went through this program for a month produced the following weight loss in pounds: 14 8 11 7 11. What is the point estimate of the mean, μ? Assume $\sigma^2 = 4$. (12 mks)