University of Ruhuna

B.Sc.(Special) Degree

Level I (Semester I) Examination - October 2021

Subject: Mathematics
Course Unit: MSP 3193 (Bayesian Inference and Decision Theory)
Time: Two (02) Hours

Answer <u>All</u> Questions.

---

1. a) We are interested in the parameter θ, of a *binomial*(n, θ) distribution. We have a prior distribution for θ with density

$$g(\theta) = \begin{cases} k_0\{\theta^2(1-\theta) + \theta(1-\theta)^2\} & ; 0 < \theta < 1 \\ 0 & ; \text{otherwise} \end{cases}$$

(i) Find the value of k_0.

(ii) We observe an observation x from the *binomial*(n, θ) distribution. Find the posterior density of θ. [40 marks]

b) (i) What is meant by a conjugate family distribution?

(ii) Write down an advantage of using a prior distribution from a conjugate family.

(iii) We observe independent observations $x_1, x_2, \ldots, x_n$ from the *Poisson*(λ) distribution. Find the likelihood function.

(iv) Suppose we have a *Gamma*(α, β) distribution for λ, where α and β are known. Find the posterior density of λ.

(v) Is this a conjugate family? Explain. [60 marks]

---

2. a) You are going to take a random sample of voters in a city to estimate the proportion π who support stopping the fluoridation of the municipal water supply. You need to determine your prior distribution for π. You decide that your prior mean is 0.4, and your prior standard deviation is 0.1.

(i) Determine the *beta*(a, b) prior that matches your prior belief.

(ii) What is the equivalent sample size of your prior?

(iii) Out of 100 city voters polled, $y = 21$ support the removal of fluoridation from the municipal water supply. Determine your posterior distribution. [60 marks]

b) Suppose we want to compare the proportions of a certain attribute in two populations. The true proportions in population 1 and population 2 are π_1 and π_2, respectively. We take a random sample from each of the populations and observe the number of each sample having the attribute. The distribution of $y_1|\pi_1$ is $binomial(n_1, \pi_1)$ and the distribution of $y_2|\pi_2$ is $binomial(n_2, \pi_2)$ and they are independent of each other. Let the prior for π_1 be $beta(a_1, b_1)$ and for π_2 be $beta(a_2, b_2)$.

Explain clearly how you obtain posterior distribution of $\pi_1 - \pi_2$. [40 marks]

---

3. a) Let y be a single observation normally distributed with mean μ and variance σ^2 which is assumed known. We have a prior distribution that is normal with mean m and variance s^2. Show that the posterior distribution of μ is a $normal(m', (s')^2)$, where

$$m' = \frac{1/s^2}{1/\sigma^2 + 1/s^2} \times m + \frac{1/\sigma^2}{1/\sigma^2 + 1/s^2} \times y \quad \text{and} \quad \frac{1}{(s')^2} = \frac{1}{s^2} + \frac{1}{\sigma^2}$$

[30 marks]

b) Hence write down expressions for posterior mean and posterior precision of μ for a random sample $y_1, y_2, \ldots, y_n$ taken from a normal distribution with mean μ and known variance σ^2, assuming a $normal(m, s^2)$ prior distribution for μ. [20 marks]

c) A medical researcher collected the systolic blood pressure reading for a random sample of $n = 30$ female students under age of 21 who visited the Student's Health Service. The blood pressures are:

120	122	121	108	133	119	136	108	106	105
122	139	133	115	104	94	118	93	102	114
123	125	124	108	111	134	107	112	109	125

Assume that systolic blood pressure comes from a $normal(\mu, \sigma^2)$ distribution where the standard deviation $\sigma = 12$.

(i) Use a $normal(120, 15^2)$ prior for μ. Calculate the posterior distribution of μ.

(ii) Find a 95% credible interval for μ.

(iii) Suppose we had not actually known the standard deviation σ. Instead the value $\hat{\sigma} = 12$ was calculated from the sample. Recalculate the 95% credible interval for μ. [50 marks]

---

4. a) The data consist of n ordered pairs of points (x_i, y_i), for $i = 1, 2, \ldots, n$. Let

$$y_i = \alpha_{\bar{x}} + \beta \times (x_i - \bar{x}) + e_i$$

for $i = 1, 2, \ldots, n$, where $\alpha_{\bar{x}}$ is the mean value for y given $x = \bar{x}$ and β is the slope of the linear model. Each e_i is *normally* distributed with mean 0 and known variance σ^2. The $e_i's$ are all independent of each other. If we use a $normal(m_\beta, s_\beta^2)$ prior for β, write down the posterior distribution for β. [20 marks]

b) A textile manufacturer is concerned about the strength of cotton yarn. In order to find out whether fiber length is an important factor in determining the strength of yarn, the quality control manager checked the fiber length (x) and strength (y) for a sample of 10 segments of yarn. The results are:

Fiber Length (x)	85	82	75	73	76	73	96	92	70	74
Strength (y)	99	93	103	97	91	94	135	120	88	92

(i) Suppose we know that the strength given the fiber length is $normal(\alpha_0 + \beta x, \sigma^2)$, where $\sigma^2 = 7.7^2$ is known. Use a $normal(0, 10^2)$ prior for β. What is the posterior distribution of β?

(ii) Find 95% credible interval for β.

(iii) Perform a Bayesian test for

$$H_0 : \beta \leq 0 \quad \text{versus} \quad H_1 : \beta > 0$$

at the 5% level of significance.

(iv) If we don't know σ^2 in part (i) above, how do we find an estimate for σ^2?

(v) How do we find a $(1 - \alpha) \times 100\%$ credible interval for β, using estimated σ^2? [80 marks]