MASENO UNIVERSITY

UNIVERSITY EXAMINATIONS 2016/2017

FOURTH YEAR FIRST SEMESTER EXAMINATIONS FOR THE DEGREE OF BACHELOR OF SCIENCE WITH INFORMATION TECHNOLOGY

MAIN CAMPUS

MIT 401: BAYESIAN MODELLING

Date: 9th December, 2016 Time: 8.30 - 11.30 am.

INSTRUCTIONS:

- Answer question ONE and any other TWO questions.
- Observe further instructions on the answer booklet.

QUESTION ONE (Compulsory) [30 Marks]

(a) Briefly describe the Bayesian method of modelling. How is it different from the frequentist of probabilistic reasoning? (6 Marks)

(b) Describe the program structure of an R code. (5 Marks)

(c) How does the following R packages work in Bayesian Modelling?

(i) **bayesSurv**

(ii) **MCMCpack**

(iii) **bayesm**

(9 Marks)

(d) In a factory, Machines M_1, M_2, M_3 are all producing springs of the same length. Machine M_1, M_2 and M_3 produce 1%, 4% and 2% defective springs, respectively. Of the total production, machine M_1 produces 30%, M_2 produces 25% and M_3 produces 45%.

(i) If a spring is selected at random from the total springs produced in a given day, determine the probability that it is defective, $P(D)$.

(ii) Given that the selected spring is defective, find the conditional probability that it was produced by machine M_2, $P(M_2|D)$. (8 Marks)

QUESTION TWO [20 Marks]

(a) Let X be a random sample of size n from a binomial distribution with parameter p. The prior density function for p is a Beta distribution.

(i) Find the Baye's estimator for p.

(ii) If $\alpha = \beta = 1$, what is the Baye's estimator for p.

(10 Marks)

(b) Three manufacturers supply clothing to a retailer. 60% of the stock comes from manufacturer 1, 30% from manufacturer 2 and 10% from manufacturer 3. 10% of the clothing from manufacturer 1 is faulty, 5% from manufacturer 2 is faulty and 15% from manufacturer 3 is faulty. What is the probability that a faulty garment comes from manufacturer 3?

(5 Marks)

(c) Under an employment promotion programme, it is proposed to allow sale of newspapers on the buses during off-peak hours. The vendor can purchase the newspapers at a special concessional rate of Shs.25 per copy against the selling price of Shs40. Any unsold copies are, however, a dead loss. A vendor has estimated the following probability distribution for the number of copies demanded.

copies	15	16	17	18	19	20
prob	0.04	0.19	0.33	0.26	0.11	0.07

How many copies should he order so that his expected profit will be maximum? (5 Marks)

ESTION THREE [20 Marks]

QUESTION THREE [20 Marks]

The following function is based on the sieve of **Eratosthenes**, the oldest known systematic method for listing prime numbers up to a given value n. The idea is as follows: begin with a vector of numbers from 2 to n.

Beginning with 2, eliminate all multiples of 2 which are larger than 2. Then move to the next number remaining in the vector, in this case, 3. Now, remove all multiples of 3 which are larger than 3. Proceed through all remaining entries of the vector in this way. The entry for 4 would have been removed in the first round, leaving 5 as the next entry to work with after 3; all multiples of 5 would be removed at the next step and so on.

```
Eratosthenes = function(n) {
  # Return all prime numbers up to n
   #(based on the sieve of Eratosthenes)
    if(n>=2){
      sieve = seq(2, n)
      primes = c()
      for (i in seq(2, n)) {
        if (any(sieve == i)) {
          primes = c(primes, i)
          sieve = c(sieve[(sieve %% i) != 0], i)
        }
      }
      return(primes)
      } else {
        stop("Input value of n should be at least 2.")
      }
 }
```

(i) Does the **Eratosthenes()** function work properly if n is not an integer? Is an error message required in this case?

(ii) Use the idea of the **Eratosthenes()** function to prove that there are infinitely many primes.

Hint: Suppose all primes were less than m, and construct a larger value n that would not be eliminated by the sieve.

QUESTION FOUR [20 Marks]

(a) We are interested in the mean, λ, of a Poisson distribution. We have a prior distribution for λ with density

$$f(\lambda) = k_0(1+\lambda)e^{-\lambda}, \quad \lambda > 0$$

(i) Find the value of k_0

(ii) Find the prior mean of λ

(iii) Find the prior standard deviation of λ

(b) We observe data $x_1, \ldots, x_n$ where, given λ, these are independent observations from the Poisson(λ) distribution.

(i) Find the likelihood.

(ii) Find the posterior density of λ.

(iii) Find the posterior mean of λ

(c) In each of the following, determine the final value of "answer"

(i)
```
answer = 0
for (j in 1:5) answer = answer + j
```

(ii)
```
answer <- NULL
for (j in 1:5) answer = c(answer, j)
```

(iii)
```
answer = 0
 for (j in 1:5) answer= c(answer, j)
```

(iv)
```
answer = 1
 for (j in 1:5) answer = answer * j
```

(v)
```
answer = 3
 for (j in 1:15) answer = c(answer, (7 * answer[j]) %% 31)
```

QUESTION FIVE [20 Marks]

The Fibonacci sequence is a famous sequence in Mathematics. The first two elements are defined as [1,1]. Subsequent elements are defined as the sum of the preceding two elements. For example, the third element is 2(=1+1), the fourth element is 3 (=1+2), the fifth element is 5 (=2+3),and so on.

To obtain the first 12 Fibonacci numbers in R, we can use

```
fibonacci=numeric(12)
fibonacci[1]=fibonacci[2]=1
for (i in 3:12)
  fibonacci[i]=fibonacci[i-2]+ fibonacci[i-1]
```

Modify the code to generate the Fibonacci sequence in the following ways:

(a) Change the first two elements to 2 and 2.

(b) Change the first two elements to 3 and 2.

(c) Change the update rule from summing successive elements to taking differences of successive elements. For example, the third element is defined as the second element minus the first element, and so on.

(d) Change the update rule so that each element is defined as the sum of three preceding elements. Set the third element as 1 in order to start the process.