JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF AGRICULTURAL AND FOOD SCIENCES

UNIVERSITY EXAMINATION FOR DEGREE OF BACHELOR OF SCIENCE

ACTUARIAL SCIENCE

4^{TH} YEAR 2^{ND} SEMESTER 2021/2022 ACADEMIC YEAR

MAIN REGULAR

COURSE CODE: WAB 2410

COURSE TITLE: BAYESIAN INFERENCE AND DECISION THEORY

EXAM VENUE: **STREAM:**

DATE: EXAM SESSION:

TIME: 2.00 HOURS

<u>Instructions:</u>

1. **Answer question 1 (Compulsory) and ANY other 2 questions in SECTION B**
2. **Candidates are advised not to write on the question paper.**
3. **Candidates must hand in their answer booklets to the invigilator while in the examination room.**

QUESTION ONE COMPULSORY (30 MARKS)

a) Briefly explain the difference between Bayesian statistics and Frequentist statistics(4mks)

b) When IRS receives tax forms, it puts them through a computer to flag forms that need to be investigated further. The computer looks for mistakes in the forms, for example addition mistakes or incorrect deduction amounts. Suppose the computer correctly flags 80% of all returns that have mistakes, and it incorrectly flags 5% of error-free returns. Further, suppose that 15% of all tax returns have errors.
A tax return is flagged by the computer. What is the chance that it actually contains mistakes, given that the computer flagged it? (5mks)

c) Let X be a continuous random variable with the following PDF

$$f_x(x) = \begin{cases} 3x^2 \; if \; o \leq x \leq 1 \\ 0 \qquad\qquad otherwise \end{cases}$$

Suppose that we know

Y/X = x ~Geometric (x)

Find the MAP estimate of X given Y = 5 (5mks)

d) It's a typically hot morning in June in Durham. You look outside and see some dark clouds rolling in. Is it going to rain?

Historically, there is a 30% chance of rain on any given day in June. Furthermore, on days when it does in fact rain, 95% of the time there are dark clouds that roll in during the morning. But, on days when it does not rain, 25% of the time there are dark clouds that roll in during the morning.
Given that there are dark clouds rolling in, what is the chance that it will rain? (5mks)

e) At a certain assembly plant, three machines make 30%, 45%, and 25%, respectively, of the products. It is known from the past experience that 2%, 3% and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected.

i) What is the probability that it is defective? (3mks)
ii) and found to be defective, what is the probability that it was made by machine 3 (4mks)

f) Let X be a continuous random variable with the following PDF

$$f(x) = \begin{cases} 6x(1-x) \quad if \; 0 \leq x \leq 1 \\ \qquad 0, \quad otherwise \end{cases}$$

Suppose that we know Y/X = x ~Geometric(x), find the posterior density of X given Y = 2, (4mks)

QUESTION TWO (20 MARKS)

a) A particular species of fish makes an annual migration up a river. On a particular day there is a probability of 0.4 that the migration will start. If it does then an observer will have to wait T minutes before seeing a fish, where T has an exponential distribution with mean 20 (i.e. an exponential(0.05) distribution). If the migration has not started then no fish will be seen.

(i) Find the conditional probability that the migration has not started given that no fish has been seen after one hour. (5mks)

(ii) How long does the observer have to wait without seeing a fish to be 90% sure that the migration has not started? (5mks)

b) Angioplasty is a medical procedure in which clogged heart arteries are widened by inserting and partially filling a balloon in the arteries. Some people have serious reactions to angioplasty, such as severe chest pains, heart attacks, or sudden death. In a recent study published in *Science*, researchers reported that 28 out of 127 adults (under age 70) who had undergone angioplasty had severe reactions.

For simplicity, suppose your prior beliefs on the population percentage of adults (under age 70) who have severe reactions to angioplasty has the following distribution:

p	Pr(p)
0	1/11
0.10	1/11
0.20	1/11
0.30	1/11
0.40	1/11
0.50	1/11
0.60	1/11
0.70	1/11
0.80	1/11
0.90	1/11
1.00	1/11

i) What is the posterior distribution of p?
ii) What is the posterior probability that p exceeds 50%? (6mks)

QUESTION THREE (20 MARKS)

a) In a forest area of Northern Europe there may be wild lynx. At a particular time the number X of lynx can be between 0 and 5 with

$$\Pr(X = x) = \binom{5}{x} 0.6^{x} 0.4^{5-x} \qquad (x = 0, \ldots, 5).$$

A survey is made but the lynx is difficult to spot and, given that the number present is x, the number Y observed has a probability distribution with

$$\boldsymbol{Pr\left(Y = \frac{y}{x} = x\right) = \begin{cases} \binom{x}{y} 0.3^{y} 0.7^{x-y} & 0 \le y \le x \\ 0 \qquad x < y \end{cases}}$$

Find the conditional probability distribution of X given that Y = 2.

(That is, find $\Pr(X = 0 \mid Y = 2), \ldots, \Pr(X = 5 \mid Y = 2)$)

(10mks)

b) Let X~Uniform (0,1). Suppose that we know Y|X=x~Geometric(x).

Find the posterior density of X given Y=2, $f_{X|Y}(x|2)$. (10mks)

QUESTION FOUR (20 MARKS)

a) Legal cases of disputed paternity in many countries are resolved using blood tests. Laboratories make genetic determinations concerning the mother, child, and alleged father.
You are on a jury considering a paternity suit. The mother has blood type O, and the alleged father has blood type AB.

A blood test shows that the child has blood type B. What is the chance that the alleged father is in fact the real father, given that the child has blood type B?
Here's some information we need to solve the problem. According to genetics, there is a 50% chance that this child will have blood type B if this alleged father is the real father. Furthermore, based on incidence rates of B genes in the population, there is a 9% chance that this child would have blood type B if this alleged father is not the real father. Based on other evidence (e.g., testimonials, physical evidence, records) presented before the DNA test, you believe there is a 75% chance that the alleged father is the real father. This assessment is your prior belief. Now, we need to use Bayes Rule to update it for the results of the child's blood test (10mks)

b) Assume our data $Y=(y_1,y_2,\ldots,y_n)T$ given X is independently identically distributed, $Y|X=x\sim$i.i.d. Exponential(λ=x) and we chose the prior to be $X\sim$Gamma(α,β). (10mks)

i. Find the likelihood of the function, $L(Y;X)= f_{Y1,Y2,\ldots,Yn}|X(y_1,y_2,\ldots,y_{n|x})$

ii. Using the likelihood function of the data, show that the posterior distribution is Gamma(α+n, $\beta+\sum_{i=1}^{n} y_i$)

iii. Write out the PDF for the posterior distribution, $f_{X|Y}(x|y)$.

iv. Find mean and variance of the posterior distribution, E[X|Y]and Var(X|Y).

QUESTION FIVE (20 MARKS)

Assume our data Y given X is distributed Y|X=x$\sim$Binomial(n, p=x) and we chose the prior to be X$\sim$Beta(α, β). Then the PMF for our data is

$$P_{Y/X}(y/x) = \binom{n}{y} x^y (1-x)^{n-y}, for\ x \in [0,1], y \in \{0,1,\ldots,n\}$$

And the PDF of the prior is given by

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} x^{\alpha-1}(1-x)^{\beta-1}, for\ 0 \le x \le 1, \alpha > 0, \beta > 0$$

(i) Show that the posterior distribution is Beta(α+y, β+n−y). (10mks)

(ii) Write out the PDF for the posterior distribution, $f_{X|Y}(x|y)$. (5mks)

(iii) Find mean and variance of the posterior distribution, E[X|Y] and Var(X|Y). (5mks)