**Question 1** [10 marks]

**(i)** State the axioms of probability. [3]

**(ii)** Deduce from the axioms in (i) that for events $A$, $B$ of a sample space: [3]

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

**(iii)** Give a general expression for $P\Big(A \cup B \cup C\Big)$, where $A, B, C$ are events. You are not required to prove your answer. [1]

An electronic system has three important parts (A, B and C) which are subject to failure. This system is to be used for a task, and during the task no repairs are possible. It is reckoned that the probability of failure for parts A, B and C (during the task) are 0.2, 0.4, and 0.3, respectively. The system will function if at least two parts have not failed.

**(iv)** Assume that failure of the parts occur independently of each other, find the probability that the system will function for the duration of the task. [3]

**Question 2** [10 marks]

Whenever an insurance company accepts a new customer for accident insurance (with a one-year term) it classifies the customer as low-risk or high risk. Suppose that the company reckons that the probability of an accident occurring during the year is 0.1 for a low-risk person, and 0.35 for a high-risk person. Assume that 25% of accepted customers for this accident insurance are low-risk and the rest are high-risk. Consider Mr. Z, randomly selected from accepted customers.

**(a)** Find the probability that Mr. Z has an accident during the year. [2]

**(b)** Assuming that Mr. Z has an accident, find the probability that Mr.Z is a low-risk person.[3]

**(c)** Consider two successive years, and suppose that the probabilities remain the same for these years, and also that for any particular customer, the chances of an accident in year 2 are independent of whether an accident occurred in year 1. Find the probability that Mr.Z will have an accident in both years. [3]

**(d)** With regards to (c), find the probability that Mr.Z is a high-risk person, given that he had an accident in both years. [2]

**Question 3** [10 marks]

The random variable $Y$ has an exponential distribution with probability density function (pdf) as follows:

$$\begin{aligned} f(y) &= \lambda e^{-\lambda y}, y > 0 \\ &= 0, otherwise \end{aligned}$$

**(i)** Showing your workings, find $P(Y > s|Y > t)$, for $s \geq t$. [3]

**(ii)** Derive an expression for the conditional pdf of $Y$, conditional on that $Y \leq 200$. [3]

$N(t)$ is a Poisson process with rate $\lambda$

**(iii)** Find an expression for the Cumulative Distribution Function (CDF) of the waiting time until the first event. (Hint: Consider the probability of there being 0 events in time t.) [2]

**(iv)** Explain the relationship between the mean of the Poisson distribution with rate $\lambda$ and the mean of the associated distribution for the waiting time. [2]

**Question 4** [10 marks]

In a multiple choice test, there are $n$ candidate answers to each questions. Suppose that the probability that a student knows the correct answer is $p$. When a student knows the correct answer, due to the stress of the exam situation, the correct answer is written with probability $1 - \Theta$. If the student does not know the correct answer then he selects one of the $n$ answers at random. Assuming that responses to the questions are independent of each other:

**(i)** Find an expression for the probability that the student actually knows the correct answer to a question, given that he answered correctly. [2]

**(ii)** Consider the first two questions on the test. Find an expression for the probability that the student answered exactly one of these questions correctly. State any assumptions that are needed for your derivation. (Hint: Start with the probability that the student does not know the correct answer to a question, given that they answered incorrectly.) [5]

Each member of a group of players rolls a fair, 6-sided dice. If any pair of players rolls the same number, the score of the group of players is increased by 1. $S_1$ is a random variable that measures the total score of the group of $n$ players.

**(iii)** Stating any assumptions that you make, find the probability distribution of $S_1$. Find the mean and variance of $S_1$. [3]

**Question 5** [10 marks]

The continuous random variable $X$ has finite mean and variance. The function $g(x)$ takes positive values only, and $E[g(x)]$ exists.

**(i)** Show that for any $c > 0$, [3]

$$P[g(X) \geq c] \leq \frac{1}{c}E[g(X)]$$

**(ii)** Use the result in (i) to deduce Chebyshev's Inequality:

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2},$$

where $\mu$, $\sigma^2$ are the mean and variance of $X$. [2]

**(iii)** Explain what is meant by *convergence in probability* for a sequence of random variables.[2]

**(iv)** Consider a sequence of Bernoulli trials, with probability of success denoted $p$, and let $X_n$ denote the number of successes after the first $n$ trials. Show that the sequence $\frac{X_n}{n}$ converges in probability to $p$. [3]

**Question 6** [10 marks]

Consider two continuous random variables, $X$, $Y$ each of which take values in $(0, \infty)$. We are given the following conditional and marginal probability density functions (pdfs):

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{x+y}{1+y}e^{-x} \\ f_Y(y) &= \frac{1+y}{2}e^{-y} \end{aligned}$$

Find each of the following:

**(i)** The joint pdf $f(x, y)$. [3]

**(ii)** The marginal pdf $f_X(x)$. [3]

**(iii)** The conditional pdf $f_{Y|X}(y|x)$. [1]

**(iv)** Find $E[X|y]$ [3]

**Question 7** [10 marks]

**(i)** Show that for any random variable $X$ and $Y$, $Var(Y) = E[Var(Y|X)] + Var(E[Y|X])$. [3]

**(ii)** The number $N$ of claims occurring on a collection of insurance policies has a Poisson distribution with parameter $m$. $Y_i$ is the claim amount. The random variables $\{Y_i\}$ are independent and identically distributed as $N(\mu, \sigma^2)$. [4]

Derive the moment generating function of $X_N$, the total claim amount of the aggregate claims.

**(iii)** By using (i), calculate $Var(X_N)$. [3]

**Question 8** [10 marks]

Consider a random sample of $n$ independent realizations of the random variable with distribution $N(\mu, \sigma^2)$.

**(i)** Show that $s^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$ is an unbiased estimator for $\sigma^2$. [3]

Consider the following estimator for $\sigma^2$:

$$\tilde{s}^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n+1}$$

**(ii)** Find the bias for $\tilde{s}^2$ as an estimator for $\sigma^2$. [2]

**(iii)** Find the mean square error for $\tilde{s}^2$ as an estimator of $\sigma^2$. [5]

**Question 9** [10 marks]

The random variable $Z$ has the standard normal distribution, and the random variable $U$ has a Chi-square distribution with $n$ degree of freedom. $Z$ and $U$ are independent random variables. The random variable $T$ is constructed as follows:

$$T = \frac{Z}{\sqrt{U/n}}.$$

**(i)** Write down the joint probability density function (pdf) of $Z$ and $U$. [3]

**(ii)** Find the joint pdf of $T$ and $U$. [3]

**(iii)** Use this to find the pdf of $T$. Name the distribution of $T$. [4]

**Question 10** [10 marks]

A random sample of size $n$ from a $N(\mu, \sigma^2)$ distribution is to be used to test the following null and alternative hypothesis:

$$H_0 : \mu \leq 2000; \quad H_A : \mu > 2000$$

using the usual z-test.

**(i)** The test will be constructed so that:

· the risk of making a type I error is at most 1%,

· there is a 90% chance of rejecting the null hypothesis, if the value of $\mu$ is at least 2300.

If the population standard deviation is 500, find the value of $n$ that is required. [4]

**(ii)** Compute the width of the 95% confidence interval for $\mu$ that would result from the study in (i). [1]

Consider the one-way analysis of variance (ANOVA) model with equal numbers of observations per treatment group:

$$Y_{ij} = \mu + \tau_i + e_{ij}, \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m$$

**(iii)** Give an explanation of each of the terms in the model and state the usual model assumptions. [5]