Final Syllabus and Chapter wise practicing problem

1. Simple Probability and Bays theorem
2. Probability distribution
3. Binomial, Poisson, and Normal distribution
4. Continuous random variable
5. Hypothesis test and Estimation

Simple Probability and Bays theorem

1.

Find the probability that the number rolled with an ordinary fair die is:

a a prime number or a 4 **b** a square number or a multiple of 3 **c** more than 3 or a factor of 8.

2.

Given that $P(A) = 0.4$, $P(B) = 0.7$ and that $P(A \cup B) = 0.8$, find

a $P(A \cup B')$ **b** $P(A' \cap B)$

3.

Abha passes through three independent sets of traffic lights when she drives to work. The probability that she has to stop at any particular set of lights is 0.2. Find the probability that Abha:

a first has to stop at the second set of lights

b has to stop at exactly one set of lights

c has to stop at any set of lights.

4.

1% of a population have a certain disease and the remaining 99% are free from this disease. A test is used to detect this disease. This test is positive in 95% of the people with the disease and is also (falsely) positive in 2% of the people free from the disease. If a person, selected at random from this population, has tested positive, what is the probability that she/he has the disease?

5. Jamia has 75% chance to attend a training session before a football match. If he attends, he is certain to be chosen for the team which plays in the match. If he does not attend, there is a probability of 0.6 that he is certain for the team

(i) Find the probability that Jamia is chosen for the team

(ii) Find the probability that Jamia attended the training session, given that he was chosen for the team.

Hypothesis test and Estimation

1.

In the past, the mean height of plants of a particular species has been 2.3 m. A random sample of 60 plants of this species was treated with fertiliser and the mean height of these 60 plants was found to be 2.4 m. Assume that the standard deviation of the heights of plants treated with fertiliser is 0.4 m.

Carry out a test at the 2.5% significance level of whether the mean height of plants treated with fertiliser is greater than 2.3 m. [5]

2.

In the past the time, in minutes, taken by students to complete a certain challenge had mean 25.5 and standard deviation 5.2. A new challenge is devised and it is expected that students will take, on average, less than 25.5 minutes to complete this challenge. A random sample of 40 students is chosen and their mean time for the new challenge is found to be 23.7 minutes.

(a) Assuming that the standard deviation of the time for the new challenge is 5.2 minutes, test at the 1% significance level whether the population mean time for the new challenge is less than 25.5 minutes. [5]

3.

Harry has a five-sided spinner with sectors coloured blue, green, red, yellow and black. Harry thinks the spinner may be biased. He plans to carry out a hypothesis test with the following hypotheses.

$$H_0\text{: P(the spinner lands on blue)} = \tfrac{1}{5}$$
$$H_1\text{: P(the spinner lands on blue)} \neq \tfrac{1}{5}$$

Harry spins the spinner 300 times. It lands on blue on 45 spins.

Use a suitable approximation to carry out Harry's test at the 5% significance level. [5]

4.

The number of cars arriving at a certain road junction on a weekday morning has a Poisson distribution with mean 4.6 per minute. Traffic lights are installed at the junction and a council officer wishes to test at the 2% significance level whether there are now fewer cars arriving. He notes the number of cars arriving during a randomly chosen 2-minute period.

(a) State suitable null and alternative hypotheses for the test. [1]

(b) Find the critical region for the test.

5.

In the past, the time, in hours, for a particular train journey has had mean 1.40 and standard deviation 0.12. Following the introduction of some new signals, it is required to test whether the mean journey time has decreased.

(a) State what is meant by a Type II error in this context. [1]

(b) The mean time for a random sample of 50 journeys is found to be 1.36 hours.

Assuming that the standard deviation of journey times is still 0.12 hours, test at the 2.5% significance level whether the population mean journey time has decreased. [5]

6.

The local council claims that the average number of accidents per year on a particular road is 0.8. Jane claims that the true average is greater than 0.8. She looks at the records for a random sample of 3 recent years and finds that the total number of accidents during those 3 years was 5.

(a) Assume that the number of accidents per year follows a Poisson distribution.

(i) State null and alternative hypotheses for a test of Jane's claim. [1]

(ii) Test at the 5% significance level whether Jane's claim is justified.

Estimation

1.

A construction company notes the time, t days, that it takes to build each house of a certain design. The results for a random sample of 60 such houses are summarised as follows.

$$\Sigma t = 4820 \qquad \Sigma t^2 = 392\,050$$

Calculate a 98% confidence interval for the population mean time.

2.

Lengths of a certain species of lizard are known to be normally distributed with standard deviation 3.2 cm. A naturalist measures the lengths of a random sample of 100 lizards of this species and obtains an α% confidence interval for the population mean. He finds that the total width of this interval is 1.25 cm.

Find α. [5]

Continuous Random variable

1.

The length, X centimetres, of worms of a certain type is modelled by the probability density functic

$$f(x) = \begin{cases} \frac{6}{125}(10-x)(x-5) & 5 \leqslant x \leqslant 10, \\ 0 & \text{otherwise.} \end{cases}$$

(a) State the value of $E(X)$. [

(b) Find $\mathrm{Var}(X)$.

2.

A random variable X has probability density function given by

$$f(x) = \begin{cases} \dfrac{k}{x^2} & 1 \leqslant x \leqslant a, \\ 0 & \text{otherwise,} \end{cases}$$

where k and a are positive constants.

(a) Show that $k = \dfrac{a}{a-1}$.

(b) Find $E(X)$ in terms of a.

3.

The length of time, T minutes, that a passenger has to wait for a bus at a certain bus stop is modelled by the probability density function given by

$$f(t) = \begin{cases} \frac{3}{4000}(20t - t^2) & 0 \leqslant t \leqslant 20, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Sketch the graph of $y = f(t)$. [1]

(b) Hence explain, without calculation, why $E(T) = 10$.

(c) Find $\mathrm{Var}(T)$.

The probability distribution

The probability distribution for a random variable describes how the probabilities are distributed over the values of the random variable. In the development of the probability function for a discrete random variable, two conditions must be satisfied: (1) p(x) must be nonnegative for each value of the random variable, and (2) the sum of the probabilities for each value of the random variable must equal one.

X	x_1	x_2	x_3	- - -	x_n
$P(X = x)$	p_1	p_2	p_3	- - -	p_n

Mean = Expected value $\mu = E(x) = \sum x\, p(x) = x_1p_1 + x_2p_2 - - - + x_np_n$

Var(x) = $\sum x^2\, p(x) - (E(x))^2$

Exercise:

1. **1.** A fair 4 sided die, numbered 1, 2,3, and 5 is rolled twice. The random variable X is the sum of the two outcomes on which the die comes to rest.

   (i) Show that P(x = 8) = $\frac{1}{8}$

(ii) Draw up the probability distribution table for X, and find p(x >6)

2.

A bag contains 7 orange balls and 3 blue balls. 4 balls are selected at random from the bag, without replacement. Let X denote the number of blue balls selected.

(i) Show that $P(X = 0) = \frac{1}{6}$ and $P(X = 1) = \frac{1}{2}$.

(ii) Construct a table to show the probability distribution of X.

(iii) Find the mean and variance of X.

3.

A fair cubical die with faces numbered 1, 1, 1, 2, 3, 4 is thrown and the score noted. The area A of a square of side equal to the score is calculated, so, for example, when the score on the die is 3, the value of A is 9.

(i) Draw up a table to show the probability distribution of A.

(ii) Find $E(A)$ and $Var(A)$.

4.

The discrete random variable X has the following probability distribution.

x	1	3	5	7
$P(X = x)$	0.3	a	b	0.25

(i) Write down an equation satisfied by a and b.

(ii) Given that $E(X) = 4$, find a and b.

5.

Two fair dice are thrown. Let the random variable X be the smaller of the two scores if the scores are different, or the score on one of the dice if the scores are the same.

(i) Copy and complete the following table to show the probability distribution of X.

x	1	2	3	4	5	6
$P(X = x)$						

(ii) Find $E(X)$.

Binomial /Bernoulli Distribution:

A binomial distribution is one kind of discrete probability distribution that has **two possible outcomes** (Success or failure / Pass or Fail)

Properties/Criteria

Binomial distributions must also meet the following three criteria:

- The number of observations or trials is fixed
- Trails are independent
- The probability of success or pass) is exactly the same from one trial to another.

When the random variable X, satisfies these conditions we denote it by

$$X \sim \beta(n, p)$$

The random variable X, which represents the number of successes in the ***n*** trials of this experiment, has a probability distribution given by

$P(X = r) = nC_r p^r q^{n-r}$ $\quad$ $where\ n = 0, 1,2,3, \ldots\ldots n$ (numbers of trails)

$p = probability\ of\ success$

$q = 1 - p$ (probability of failure)

Mean (Expected Value) and Variance of Binomial Distribution

If $X \sim \beta(n, p)$, then mean $\mu = E(x) = n \times p$ $\quad$ Variance $\sigma^2 = npq$

Exercise:

1. A driving test is passed by 70% of people at their attempt. Find the probability that
   (i) exactly 5 people out of 10 people will passed the driving test.
   (ii) More than 1 people out of 8 people will passed the driving test.

2.

65% of all watches sold by a shop have a digital display and 35% have an analog display.

(i) Find the probability that, out of the next 12 customers who buy a watch, fewer than 10 choose one with a digital display. [4]

3.

(i) A garden shop sells polyanthus plants in boxes, each box containing the same number of plants. The number of plants per box which produce yellow flowers has a binomial distribution with mean 11 and variance 4.95.

(a) Find the number of plants per box. [4]

(b) Find the probability that a box contains exactly 12 plants which produce yellow flowers. [2]

Poisson Distribution

In probability theory and statistics, the Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event.

Conditions /Properties

- Occurs singly
- The average rate at which events occur is always the same
- Events are independent

The random variable X, satisfying Poisson distribution $X \sim P_0(\mu)$, then probability distribution given by

$$P(X = r) = \frac{e^{-\mu} \times \mu^r}{r!}$$

Mean (Expected Value) and Variance of Poisson Distribution

$E(x) =$ Variance $\sigma^2 = \mu$

Exercise:

1.

Computer breakdowns occur randomly on average once every 48 hours of use.

(i) Calculate the probability that there will be fewer than 4 breakdowns in 60 hours of use.

(ii) there will be no breakdowns in 24 hours of use.

2.

Between 7 p.m. and 11 p.m., arrivals of patients at the casualty department of a hospital occur at random at an average rate of 6 per hour.

(i) Find the probability that, during any period of one hour between 7 p.m. and 11 p.m., exactly 5 people will arrive. [2]

(ii) A patient arrives at exactly 10.15 p.m. Find the probability that at least one more patient arrives before 10.35 p.m. [3]

3.

The number of radioactive particles emitted per second by a certain metal is random and has mean 1.7. The radioactive metal is placed next to an object which independently emits particles at random such that the mean number of particles emitted per second is 0.6. Find the probability that the total number of particles emitted in the next 3 seconds is 6, 7 or 8. [4]

Normal Distribution

The normal distribution is a continuous probability distribution that is symmetrical around its mean, most of the observations cluster around the central peak, and the probabilities for values further away from the mean taper off equally in both directions.

Properties of a normal distribution

- The mean, mode and median are all equal.
- The curve is symmetric at the center (i.e. around the mean, μ).
- Exactly half of the values are to the left of center and exactly half the values are to the right.
- The total area under the curve is 1.

To standardize your data, you need to convert the raw measurements into Z-scores.

To calculate the standard score for an observation, take the raw measurement, subtract the mean, and divide by the standard deviation. Mathematically, the formula for that process is the following:

The z-score formula that we have been using is: $Z = \frac{X - \mu}{\sigma}$

The probability distribution function for Normal Distribution is defined as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$$

$\mu =$ Mean
$\sigma =$ Standard Deviation
$\pi \approx 3.14159\cdots$
$e \approx 2.71828\cdots$

$$P(a \le x \le b) = \int_a^b f(x)dx \ge 0$$

Exercise:

1.

The distance in metres that a ball can be thrown by pupils at a particular school follows a normal distribution with mean 35.0 m and standard deviation 11.6 m.

(i) Find the probability that a randomly chosen pupil can throw a ball between 30 and 40 m. [3]

(ii) The school gives a certificate to the 10% of pupils who throw further than a certain distance. Find the least distance that must be thrown to qualify for a certificate. [3]

2.

(i) In a normal distribution with mean μ and standard deviation σ, $P(X > 3.6) = 0.5$ and $P(X > 2.8) = 0.6554$. Write down the value of μ, and calculate the value of σ. [4]

(ii) If four observations are taken at random from this distribution, find the probability that at least two observations are greater than 2.8. [4]

3.

The waiting time in a doctor's surgery is normally distributed with mean 15 minutes and standard deviation 4.2 minutes.

(i) Find the probability that a patient has to wait less than 10 minutes to see the doctor. [3]

(ii) 10% of people wait longer than T minutes. Find T. [3]

(iii) In a given week, 200 people attend the surgery. Estimate the number of these who wait more than 20 minutes. [3]