Course outline and lesson plan

1.Elementary Probability

Addition rule, Multiplicative rule, Mutually exclusive event, Independent event, conditional probability, Bayes' theorem

2.Probability distribution

Discrete random variables

- Definition, types and construction of discrete probability distribution
- Binomial distribution
- Poisson distribution

Continuous random variable

- Normal distribution
- Continuous distribution (pdf, cdf)

3.Estimation – Point and Interval estimation

4. Hypothesis Test: (using *p*-value or critical value)

Test statistics could be Normal, Binomial or Poisson

Sample notes and Specimen problem

1.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:

i. p(x) must be nonnegative for each value of the random variable, and
ii. 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. 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.

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.

3.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:

- There are two outcomes
- 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

$P(X = r) = nC_r p^r q^{n-r}$ $\quad where\ n = 0, 1,2,3, \dots\dots n$ (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 $\mu = E(x) = n \times p$ 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]

3. **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 $X \sim P_0()$, 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]

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: $Z = \dfrac{X - \mu}{\sigma}$

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]

4.

In a certain country the time taken for a common infection to clear up is normally distributed with mean μ days and standard deviation 2.6 days. 25% of these infections clear up in less than 7 days.

(i) Find the value of μ. [4]

In another country the standard deviation of the time taken for the infection to clear up is the same as in part **(i)**, but the mean is 6.5 days. The time taken is normally distributed.

(ii) Find the probability that, in a randomly chosen case from this country, the infection takes longer than 6.2 days to clear up. [3]

5. Estimation – Point and Interval estimation

Construction of Confidence Intervals:

For mean (know variance)

$$\bar{x} \pm Z \cdot \frac{\sigma}{\sqrt{n}}$$

For proportion

$$\hat{p} \pm Z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

6. Hypothesis test:

a. Hypothesis Testing Example Using Normal Distribution

Problem Statement:

A soft drink company claims that their 500ml bottles contain, on average, 500ml of soda. A consumer advocacy group suspects that the bottles contain less than 500ml on average. A random sample of 36 bottles is taken, and the mean volume is found to be 495ml, with a standard deviation of 10ml. Test the advocacy group's claim at a 5% significance level ($\alpha = 0.05$).

b. Hypothesis Testing Example Using Binomial Distribution

Problem Statement:

30% of customers in a large store present the store's loyalty card when they buy something in the store.
The store runs an advertising campaign to promote their loyalty card. After the advertising campaign has finished, a random sample of 20 sales is examined and a loyalty card was used in 10 of them.
Test at the 5% level of significance whether the campaign has been effective in increasing the use of the loyalty card.

c. Hypothesis Testing Example Using Poisson Distribution

Problem Statement:

A hospital claims that the **average number of emergency cases per hour is 4**. A researcher monitors the hospital for **6 hours** and records **36 emergency cases**. At a **5% significance level (α=0.05α = 0.05α=0.05)**, test whether the hospital's claim about the average rate of emergency cases is valid.