FUTURE INSTITUTE OF ENGINEERING AND MANAGEMENT

SUGGESTION

PAPER CODE: PCCDS501

Chapter 1:

Marks: 1

1. The number of students in a class is an example of
   a) Continuous variable b) Discrete variable c) Definite variable d) None of these
2. Suppose 4 coins are tossed, the value of a random variable for number of heads is
   a) 1,2,3,4 b) 0,1,2,3,4 c) 0,1,2,3 d) 0,1
3. A random variable is also called
   a) Constant b) Variable c) Attribute d) Chance variable
4. A quantity which can vary from one individual to another is called
   a) Constant b) Variable c) Data d) None of these
5. A variable which can assume each and every value within a given range is called
   a) Discrete variable b) Random variable c) Qualitative variable d) Continuous variable
6. Random number can be generated mechanically by
   a) Use of random numbers table
   b) Use of digital computers
   c) Ordinary calculators
   d) None of these
7. A random variable assuming only a finite number of variables is called
   a) Discrete random variable
   b) Continuous random variable
   c) Random variable
   d) None of these
8. A random variable assuming an infinite number of values called
   a) Absolute variable
   b) Discrete random variable
   c) Continuous random variable
   d) None of these
9. A variable whose values is determined by the outcome of a random experiments is called
   a) Random
   b) Random Variable
   c) Constant
   d) None of these
10. If x is a discrete random variable, the function f(x) is
    a) Distribution function
    b) Probability function

c) Density function
d) None of these

11. From a group of 7 men and 6 women, five persons are to be selected to form a committee so that at least 3 men are there on the committee. In how many ways can it be done?

Answer: 756

12. In how many different ways can the letters of the word 'LEADING' be arranged in such a way that the vowels always come together?

Answer: 720 (M)

13. What is the probability that two cards drawn at random from a deck of playing cards will both be aces? Answer: 1/221 (m)

14. A die is cast twice and a coin is tossed twice. What is the probability that the die will turn a 6 each time and the coin will turn a tail every time? Answer: = 1/144 (m)

15. A die is cast 6 times. What is the probability that each throw will return a prime number?
Answer: 1/64

Marks: 5

1. Find the range for each of the following random variables.
   1. I toss a coin 100 times. Let X be the number of heads I observes.
   2. I toss a coin until the first heads appears. Let Y be the total number of coin tosses.
   3. The random variable T is define as the times (in hours) from now until the next earthquake occurs in a certain city.
   Answer:
1. The random variable X can take any integer from 0 to 100 so $R_X = \{0,1,2,3,\ldots..100\}$.
2. The random variable Y can take any positive integer, so $R_Y = \{1,2,3\ldots.\} = N$.
3. The random Variable T can take any nonnegative real number so, $R_T = [0,\alpha)$.

2. The probability that a management trainee will remain with a company is 0.6. The probability that an employee earns more than Rs. 10,000 per month is 0.50. The probability that an employee is a management trainee who remained with the company or who earns more than Rs. 10,000 per month is 0.70. What is the probability that an employee earns more than Rs. 10,000 per month, given that he is a management trainee who stayed with the company?
   Answer: = 2/3

3. There are 12 balls in a bag, 8 red and 4 green. Three balls are drawn successively without replacement. What is the probability that there are alternately of the same colour?
   Answer: 8/33
4. A box contains 3 red and 7 white balls. One ball is drawn at random and in its place a ball of the other colour is put in the box. Now one ball is drawn at random from the box. Find the probability that it is red.
   Answer: 0.34

5. A bag contains 30 balls numbered 1 through 30. Suppose drawing an even numbered ball is considered a success . Two balls are drawn from the bag with replacement. Find the probability of getting:

(i) Two successes

(ii) Exactly one success

(iii) At least one success

(iv) No successes

Marks: 15

1. A committee of 4 persons is to be appointed from 3 officers of the production department, 4 officers of the purchase department, 2 officers of the sales department and 1 chartered accountant. Find the probability of forming the committee in the following manners: (H)
   a) There must be one from each category
   b) It should have at least one from the purchase department
   c) The chartered accountant must be in the committee.
   1) 0.1143
   2) 0.9286
   3) 0.4
2. Prove that For 3 events A, B,C the probability occurrence of any one of the mutually exclusive events, is equal to the sum of their individual probabilities given by,
   1. $P(A \cup B \cup C) = P(A) + P(B) + P(C)$
   2. Prove that $P(A \cap B) = P(A) * P(A|B)$ where A and B are the two events simultaneously happening.
   3. Prove that P(A’) = 1 – P(A) for an event A

3. If P(A) = 0.4 , P(B) = 0.7 and $P(at\ least\ one\ of\ A\ and\ B) = 0.8$ find P(only one of A and B)

Answer: 0.3

4. In a village of 21 inhabitants, a person tells a rumor to a second person, who in turn repeats it to a third person etc. At each step the recipient of the rumor is chosen at random from the 20 people available. Find the probability that the rumor will be told 10 times without
   a) Returning to the originator
   b) Being repeated to any person
   Answer: $(19/20)^9$ $(20*19*18\ldots11)/20^{10}$

5. A) A five figure number is formed by the digits 0,1,2,3,4 (without repetitions). Find the probability that the number formed is divisible by 4.
   Answer: 5/16
   B) There are four hotels in a certain town. If 3 men check into hotels in a day, what is the probability that each check into a different hotels?
   Answer: 0.375

6. a) State and prove the addition theorem of probability for any two events A and B. Rewrite the theorem when A and B are mutually exclusive.

   b) State the axioms of probability

   c) Explain the meaning of conditional probability.

   d) Define independent events. Obtain the necessary and sufficient condition for the independence of two events A and B. (5 +3+3+4)

7. a) Explain with the example the rules of addition and multiplication in the theory of Probability.

   b) What is the difference between permutation and combination. Write the formula for both.

CHAPTER – 2

MARKS: 1

1. Consider a dice with the property that that probability of a face with n dots showing up is proportional to n. The probability of face showing 4 dots is? 4/21
2. Runs scored by batsman in 5 one day matches are 50, 70, 82, 93, and 20. The standard deviation is : The mean of 5 innings is
   (50+70+82+93+20)÷5 = 63
   S.D = $[\frac{1}{n}(x(n)\text{-mean})^2]^{0.5}$
   S.D = 25.79.
3. Find median and mode of the messages received on 9 consecutive days 15, 11, 9, 5, 18, 4, 15, 13, 17. Arranging the terms in ascending order 4, 5, 9, 11, 13, 14, 15, 18, 18.
   Median is (n+1)/2 term as n = 9 (odd) = 13.
   Mode = 18 which is repeated twice.
4. If E denotes the expectation the variance of a random variable X is denoted as? By property of Expectation
   $V(X) = E(X^2)-(E(X))^2$.
5. What is the area under a conditional Cumulative density function?

6. X is a variate between 0 and 3. The value of $E(X^2)$ is _______
   Explain: Integrating $f(x) = x^2$ from 0 to 3 we get $E(X^2)$ = 32 = 9.
7. The random variables X and Y have variances 0.2 and 0.5 respectively. Let Z= 5X-2Y. The variance of Z is?
   Var(X) = 0.2, Var(Y) = 0.5
   Z = 5X – 2Y
   Var(Z) = Var(5X-2Y)
   = Var(5X) + Var(2Y)
   = 25Var(X) + 4Var(Y)
   Var(Z) = 7.

8. If E(x) = 2 and E(z) = 4, then E(z – x) =? = 2

9. Let X be a random variable with probability distribution function f (x)=0.2 for |x|<1
   = 0.1 for 1 < |x| < 4
   = 0 otherwise
   The probability P (0.5 < x < 5) is ______
   Explanation: P (0.5 < x < 5) = Integrating f (x) from
   0.5 to 5 by splitting in 3 parts that is from 0.5 to 1
   and from 1 to 4 and 4 to 5 we get
   P (0.5 < x < 5) = 0.1 + 0.3 + 0
   P (0.5 < x < 5) = 0.4.

10. If f(x) is a probability density function of a continuous random variable, then $\int_{-\infty}^{\infty} f(x)$ = ? Answer: 1

MARKS:5

1. (i) If the probability that a bomb dropped from a place will strike the target is 60% and if 10 bombs are dropped, find mean and variance?
   (ii) If P(1) = P(3) in Poisson's distribution, what is the mean?
   (iii) Find the mean of tossing 8 coins.
   (i) Explanation: Here, p = 60% = 0.6 and q = 1-p = 40% = 0.4 and n = 10
   Therefore, mean = np = 6
   Variance = npq = (10)(0.6)(0.4)
   = 2.4.
   (ii)$P(x)=(e^{-\lambda}\lambda^{x})/x!$
   Therefore, $P(3)=(e^{-\lambda}\lambda^{3})/3!$
   and $P(1)=(e^{-\lambda}\lambda^{1})/1!$
   P(1) = P(2)
   Therefore, $\lambda=\sqrt{6}$.
   (iii) Explanation: p = ½
   n = 8
   q = ½
   Therefore, mean = np = 8 * ½ = 4.
2. (i) If 40% of boys opted for maths and 60% of girls opted for maths, then what is the probability that maths is chosen if half of the class's population is girls?
   Let E be the event of electing boy or a girl and A be the event of selecting a maths student.
   $P(A) = P(E_1)\ P(A|E_1) + P(E_2)\ P(A|E_2)$
   =(12)(40100)+(12)(60100)
   = 0.5.
   (ii) Company A produces 10% defective products, Company B produces 20% defective products and C produces 5% defective products. If choosing a company is an equally likely event, then find the probability that the product chosen is defective. Answer: 0.12
3. (i)Suppose 5 men out of 100 men and 10 women out of 250 women are colour blind, then find the total probability of colour blind people. (Assume that both men and women are in equal numbers.)= 0.045
   (ii) A problem is given to 5 students P, Q, R, S, T. If the probability of solving the problem individually is 1/2, 1/3, 2/3, 1/5, 1/6 respectively, then find the probability that the problem is solved. = 0.37
4. Let X and Y be jointly continuous random variable with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} cx+1 & x, y \geq 0, x+y < 1 \\ 0 & otherwsie \end{cases}$$

   (i) Show the range of (X,Y), R_{XY}, in x-y plane.
   (ii) Find the constant c
   (iii) Find the marginal pdfs $f_X(x)$ and $f_Y(y)$
   (iv) Find $P(Y<2x^2)$

5. Let X and Y be jointly continuous random variable with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 6e^{-(2x+3y)} & x,y \geq 0, \\ 0 & otherwsie \end{cases}$$

(i)Are X and Y independent?

(ii) Find E[Y|X>2].

(iii) Find P(X>Y)

Marks: 15

1. (i) What do you mean by joint probability distribution? What is cumulative distribution? (Explain with example)
   (ii) What do you mean by expectation, variance and standard deviation of a random variable?
   (iii) Write the significance of expectation value, and variance in data science.
2. (i) What do you mean by moment? What is the significance of it in data science?
   (ii) Define various types of quartiles with example.
3. (i) Show that mean of a random variable X = Expected value of the random variable X
   (ii) Prove that Var(X +c) = var(X)
   (iii) Prove that Var(aX+b) = a^2Var(X)
   (iv) Prove that Cov(aX,Y) = aCov(X,Y)
   (v) How standard deviation related to variance?
4. (i) Define conditional probability with example.
   (ii) Define density function with example.
   (iv) Define cumulative distribution with example.
   (iii) A box contains two coins: a regular coin and one fake two headed coin (P(H) = 1). I choose a coin at random and toss it twice. Define the following events.
   (a) A = first coin toss results in an H
   (b) B = Second coin toss results in an H
   (c) Coin 1 (regular) has been selected.
   (d) Find P(A|C), P(B|C), P(A∩B|C), P(A), P(B). Note that A and B are not independent but they are conditionally independent given C.

CHAPTER 3

Marks: 1

1. Six men and five women apply for an executive position in a small company. Two of the applicants are selected for an interview. Let X denote the number of women in the interview pool. We have found the probability mass function of X. (easy)

$X = x$	0	1	2
$P(x)$	$\frac{2}{11}$	$\frac{5}{11}$	$\frac{4}{11}$

How many women do you expect in the interview pool?

Answer: Expected number of women in the interview pool is

$$E(X) = \sum_{x} x P_X(x)$$

$$= \left[\left(0 \times \frac{2}{11}\right) + \left(1 \times \frac{5}{11}\right) + \left(2 \times \frac{4}{11}\right)\right]$$

$$= \frac{13}{11}$$

2.An urn contains four balls of red, black, green and blue colours. There is an equal probability of getting any coloured ball. What is the expected value of getting a blue ball out of 30 experiments with replacement? (easy)

Answer: Probability of getting a blue ball = (p) = 1/4 = 0.25

Total experiments (N) = 30

Expected value = Number of experiments × Probability

= N × p

= 30 × 0.25

= 7.50

Therefore, the expected value of getting blue ball is approximately 8

3. The following information is the probability distribution of successes.

No. of Successes	0	1	2
Probability	$\frac{6}{11}$	$\frac{9}{22}$	$\frac{1}{22}$

Determine the expected number of success. (easy)

Answer: Expected number of success is

$$E(X) = \sum_x x P_X(x)$$
$$= \left(0 \times \frac{6}{11}\right) + \left(1 \times \frac{9}{22}\right) + \left(2 \times \frac{1}{22}\right)$$
$$= \frac{11}{22}$$
$$= 0.5$$

Therefore, the expected number of success is 0.5. Approximately one success

4. Write the expression of Poisson's Distribution. (Easy)
5. Write the expression of mean and variance of Binomial distribution. (Easy)
6. Define correlation. (med)
7. What do you mean by covariance? (Med)
8. A coin is tossed12 times. What is the probability of getting exactly 7 heads?(Hard)

Solution:

Given that a coin is tossed 12 times. (i.e) n= 12

Thus, a probability pf getting head in single toss = ½. (i.e) p = ½.
So, 1-p = 1-½ = ½.
We know that the binomial probability distribution is $P(r) = {}^{n}C^{r} \cdot p^{r} (1 - p)^{n-r}$.
Now, we have to find the probability of getting exactly 7 heads.(i.e) r = 7.
Substituting the values in the binomial distribution formula, we get
$P(7) = {}^{12}C_7 \cdot (½)^7 (½)^{12-7}$
$P(7) = 792 \cdot (½)^7 (½)^5$
$P(7) = 792.(½)^{12}$

P(7) = 792 (1/4096)

P(7) = 0.193
Therefore, the probability of getting exactly 7 heads is 0.193.

9. The probability that a person can achieve a target is 3/4. The count of tries is 5. What is the probability that he will attain the target at least thrice? (Med)
   Solution:
   Given that, p = ¾, q = ¼, n = 5.

Using binomial distribution formula, we get $P(X) = {}^nC_x \cdot p^x (1 - p)^{n-x}$
Thus, the required probability is: $P(X = 3) + P(X=4) + P(X=5)$
$= {}^5C_3 \cdot (¾)^3 (¼)^2 + {}^5C_4 \cdot (¾)^4 (¼)^1 + {}^5C_5 \cdot (¾)^5$
$= 459/512$.
Therefore, the probability that the person will attain the target atleast thrice is 459/512.

10. A coin that is fair in nature is tossed n number of times. The probability of the occurrence of a head six times is the same as the probability that a head comes 8 times, then find the value of n. (Hard)

Solution:

The probability that head occurs 6 times = ${}^nC_6 (½)^6 (½)^{n-6}$

Similarly, the probability that head occurs 8 times = ${}^nC_8 (½)^8 (½)^{n-8}$

Given that, the probability of the occurrence of a head six times is the same as the probability that a head comes 8 times,

(i.e) ${}^nC_6 (½)^6 (½)^{n-6} = {}^nC_8 (½)^8 (½)^{n-8}$

$\Rightarrow {}^nC_6 (½)^n = {}^nC_8 (½)^n$

$\Rightarrow {}^nC_6 = {}^nC_8$

$\Rightarrow 6 = n-8$

$\Rightarrow n= 14$.

Therefore, the value of n is 14.

Marks: 5

1. Determine the mean and variance of the random variable *X* having the following probability distribution. (Easy)

$X = x$	1	2	3	4	5	6	7	8	9	10
$P(x)$	0.15	0.10	0.10	0.01	0.08	0.01	0.05	0.02	0.28	0.20

Answer:

Mean of the random variable $X = E(X) = \sum_x x\, P_X(x)$

$= (1 \times 0.15) + (2 \times 0.10) + (3 \times 0.10) + (4 \times 0.01) + (5 \times 0.08) + (6 \times 0.01) +$

$(7 \times 0.05) + (8 \times 0.02) + (9 \times 0.28) + (10 \times 0.20)$

$E(X) = 6.56$

$$E\left(X^2\right) = \sum_x x^2\, P_X(x)$$

$= (1^2 \times 0.15) + (2^2 \times 0.10) + (3^2 \times 0.10) + (4^2 \times 0.01) +$

$(5^2 \times 0.08) + (6^2 \times 0.01) + (7^2 \times 0.05) + (8^2 \times 0.02) +$

$(9^2 \times 0.28) + (10^2 \times 0.20).$

$= 50.38$

Variance of the Random Variagble $X = V(X) = E\left(X^2\right) - [E(X)]^2$

$= 50.38 - (6.56)^2$

$= 7.35$

2. Define Normal distribution. What is SNV? Explain the advantages of normal distribution. (Easy)
3. Consider a random variable X with probability density function

$$f(x) = \begin{cases} 4x^3, if\ 0 < x < 1 \\ 0,\ otherwise \end{cases}$$

Find E(X) and V(X). (Me)

4. If $f(x)$ is defined by $(x) = ke^{-2x}$, $0 \le x < \infty$ is a density function. Determine the constant k and also find mean. (Hard)

Solution

We know that

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \text{, since } f(x) \text{ is a density function.}$$

$$\int_0^{\infty} k e^{-2x}\,dx = 1$$

$$k\int_0^{\infty} e^{-2x}\,dx = 1$$

$$k\left[\frac{e^{-2x}}{-2}\right]_0^{\infty} = 1$$

$$k = 2$$

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\,dx$$

$$= \int_0^{\infty} x k e^{-2x}\,dx$$

$$= 2\int_0^{\infty} x e^{-2x}\,dx$$

$$= 2\left\{\left[\frac{xe^{-2x}}{-2}\right]_0^{\infty} - \int_0^{\infty} \frac{e^{-2x}}{-2}\,dx\right\} \quad \left(\because \int u\,dv = uv - \int v\,du\right)$$

$$= \int_0^{\infty} e^{-2x}\,dx$$

$$= \frac{1}{2}$$

Marks: 15

1. What do you mean by Monte Carlo estimation? Write its characterises. Write 3 advantages and 3 disadvantages of this method. Draw the flow diagram of this method. (2+2+3+3+5) (easy)
2. What do you mean by conditional expectation? What do you mean by convergence of probability? What is the application of convergence of probability in data science? What do you mean by generating function? Define Binomial distribution and its application with example. (2+2+2+2+7) (easy)

3. Determine the mean and variance of a discrete random variable, given its distribution as follows: (M)

$X = x$	1	2	3	4	5	6
$F_x(x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	1

Answer:

From the given data, you first calculate the probability distribution of the random variable. Then using it you calculate mean and variance.

X	$p(x)$
1	$F(1) = \frac{1}{6}$
2	$F(2) - F(1) = \frac{2}{6} - \frac{1}{6} = \frac{1}{6}$
3	$F(3) - F(2) = \frac{3}{6} - \frac{2}{6} = \frac{1}{6}$
4	$F(4) - F(3) = \frac{4}{6} - \frac{3}{6} = \frac{1}{6}$
5	$F(5) - F(4) = \frac{5}{6} - \frac{4}{6} = \frac{1}{6}$
6	$F(6) - F(5) = 1 - \frac{5}{6} = \frac{1}{6}$

The probability mass function is

$X = x$	1	2	3	4	5	6
$P(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Mean of the random variable $X = E(X) = \sum_x x P_X(x)$

$$= \left(1\times\frac{1}{6}\right)+\left(2\times\frac{1}{6}\right)+\left(3\times\frac{1}{6}\right)+\left(4\times\frac{1}{6}\right)+\left(5\times\frac{1}{6}\right)+\left(6\times\frac{1}{6}\right)$$

$$= \frac{1}{6}(1+2+3+4+5+6)$$

$$= \frac{7}{2}$$

$$E(X^2) = \sum_x x^2 P_X(x)$$

$$= \left(1^2\times\frac{1}{6}\right)+\left(2^2\times\frac{1}{6}\right)+\left(3^2\times\frac{1}{6}\right)+\left(4^2\times\frac{1}{6}\right)+\left(5^2\times\frac{1}{6}\right)+\left(6^2\times\frac{1}{6}\right)$$

$$= \frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2)$$

$$= \frac{91}{6}$$

Variance of the Random Variable $X = V(X) = E(X^2) - [E(X)]^2$

$$= \frac{91}{6} - \left(\frac{7}{2}\right)^2$$

$$= \frac{35}{12}$$

4. (a)A radar unit is used to measure speeds of cars on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. What is the probability that a car picked at random is travelling at more than 100 km/hr?

Answer: Let x be the random variable that represents the speed of cars. x has μ = 90 and σ = 10. We have to find the probability that x is higher than 100 or P(x > 100)
For x = 100 , z = (100 - 90) / 10 = 1

$P(x > 90) = P(z > 1)$ = [total area] - [area to the left of $z = 1$]
$= 1 - 0.8413 = 0.1587$
The probability that a car selected at a random has a speed greater than 100 km/hr is equal to 0.1587

(b) For a certain type of computers, the length of time between charges of the battery is normally distributed with a mean of 50 hours and a standard deviation of 15 hours. John owns one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours.
Answer: Let x be the random variable that represents the length of time. It has a mean of 50 and a standard deviation of 15. We have to find the probability that x is between 50 and 70 or $P(50< x < 70)$
For $x = 50$, $z = (50 - 50) / 15 = 0$
For $x = 70$, $z = (70 - 50) / 15 = 1.33$ (rounded to 2 decimal places)
$P(50< x < 70) = P(0< z < 1.33)$ = [area to the left of $z = 1.33$] - [area to the left of $z = 0$]
$= 0.9082 - 0.5 = 0.4082$
The probability that John's computer has a length of time between 50 and 70 hours is equal to 0.4082.

(c) Entry to a certain University is determined by a national test. The scores on this test are normally distributed with a mean of 500 and a standard deviation of 100. Tom wants to be admitted to this university and he knows that he must score better than at least 70% of the students who took the test. Tom takes the test and scores 585. Will he be admitted to this university?

Answer: Let x be the random variable that represents the scores. x is normally ditsributed with a mean of 500 and a standard deviation of 100. The total area under the normal curve represents the total number of students who took the test. If we multiply the values of the areas under the curve by 100, we obtain percentages.
For $x = 585$, $z = (585 - 500) / 100 = 0.85$
The proportion P of students who scored below 585 is given by
P = [area to the left of $z = 0.85$] = 0.8023 = 80.23%
Tom scored better than 80.23% of the students who took the test and he will be admitted to this University.

1. X is a normally normally distributed variable with mean $\mu = 30$ and standard deviation $\sigma = 4$. Find
   a) $P(x < 40)$
   b) $P(x > 21)$
   c) $P(30 < x < 35)$
2. The length of similar components produced by a company are approximated by a normal distribution model with a mean of 5 cm and a standard deviation

of 0.02 cm. If a component is chosen at random

a) what is the probability that the length of this component is between 4.98 and 5.02 cm?

b) what is the probability that the length of this component is between 4.96 and 5.04 cm?

3. The length of life of an instrument produced by a machine has a normal distribution with a mean of 12 months and standard deviation of 2 months. Find the probability that an instrument produced by this machine will last

   a) less than 7 months.

   b) between 7 and 12 months.

4. The time taken to assemble a car in a certain plant is a random variable having a normal distribution of 20 hours and a standard deviation of 2 hours. What is the probability that a car can be assembled at this plant in a period of time

   a) less than 19.5 hours?

   b) between 20 and 22 hours?

5. A large group of students took a test in Physics and the final grades have a mean of 70 and a standard deviation of 10. If we can approximate the distribution of these grades by a normal distribution, what percent of the students

   a) scored higher than 80?

   b) should pass the test (grades≥60)?

   c) should fail the test (grades<60)?

6. The annual salaries of employees in a large company are approximately normally distributed with a mean of $50,000 and a standard deviation of $20,000.

   a) What percent of people earn less than $40,000?

   b) What percent of people earn between $45,000 and $65,000?

   c) What percent of people earn more than $70,000?

   Answer (1):

1. Note: What is meant here by area is the area under the standard normal curve.

   a) For x = 40, the z-value z = (40 - 30) / 4 = 2.5

   Hence P(x < 40) = P(z < 2.5) = [area to the left of 2.5] = 0.9938

   b) For x = 21, z = (21 - 30) / 4 = -2.25

   Hence P(x > 21) = P(z > -2.25) = [total area] - [area to the left of -2.25]

   = 1 - 0.0122 = 0.9878

   c) For x = 30 , z = (30 - 30) / 4 = 0 and for x = 35, z = (35 - 30) / 4 = 1.25

   Hence P(30 < x < 35) = P(0 < z < 1.25) = [area to the left of z = 1.25] - [area to the left of 0]

   = 0.8944 - 0.5 = 0.3944

2. A) P(4.98 < x < 5.02) = P(-1 < z < 1)

   = 0.6826

   b) P(4.96 < x < 5.04) = P(-2 < z < 2)

   = 0.9544

3. a) $P(x < 7) = P(z < -2.5)$
   $= 0.0062$
   b) $P(7 < x < 12) = P(-2.5 < z < 0)$
   $= 0.4938$

4. a) $P(x < 19.5) = P(z < -0.25)$
   $= 0.4013$
   b) $P(20 < x < 22) = P(0 < z < 1)$
   $= 0.3413$

5. a) For $x = 80$, $z = 1$
   Area to the right (higher than) $z = 1$ is equal to 0.1586 = 15.87% scored more that 80.
   b) For $x = 60$, $z = -1$
   Area to the right of $z = -1$ is equal to 0.8413 = 84.13% should pass the test.
   c)100% - 84.13% = 15.87% should fail the test.

6. a) For $x = 40000$, $z = -0.5$
   Area to the left (less than) of $z = -0.5$ is equal to 0.3085 = 30.85% earn less than $40,000.
   b) For $x = 45000$, $z = -0.25$ and for $x = 65000$, $z = 0.75$
   Area between $z = -0.25$ and $z = 0.75$ is equal to 0.3720 = 37.20 earn between $45,000 and $65,000.
   c)For $x = 70000$, $z = 1$
   Area to the right (higher) of $z = 1$ is equal to 0.1586 = 15.86% earn more than $70,000.

CHAPTER 4

Marks: 1

1. What is the formula of calculating the confidence interval in confidence interval estimation?
2. What does range or set of values having chances to contain value of population parameter with particular confidence level considered as?
3. If sample size is greater than or equal to 30 then sample standard deviation can be approximated to population standard deviation for the?
4. Considering sample satistic if mean of sampling distribution is equal to population mean then what does sample satistics is classified as?
5. What is the value of any sample statistics which is used to estimate parameters of population classifies as?
6. What does the method in which sample statistics is used to estimate value of parameters of population classified as?
7. What is the confidence interval if point estimate is 8 and margin of error is 5 ?
8. What does the distance between true value of population parameter and estimated value of population parameter called?

9. Which hypothesis test should be used to as certain improvement of the worker's performance before and after training?
10. Write one advantage of simple random sampling.
11. What do you mean by likelihood inference
12. What do you mean by finite population?
13. What do you mean by poster distribution?

Marks 5

1. Write the steps of simple random sampling.
2. What do you mean by simple random sampling?
3. What do you mean by Maximum likelihood estimation?
4. What is the requirement of MLE in data science?
5. Suppose that you would like to estimate the portion of voters in your town that plan to vote for Party A in an upcoming election. To do so, you take a random sample of size n from the likely voters in the town. Since you have a limited amount of time and resources, your sample is relatively small. Specifically, suppose that n=20. After doing your sampling, you find out that 6 people in your sample say they will vote for Party A.

6. 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

a) What is the posterior distribution of p?
b) What is the posterior probability that p exceeds 50%?

Setting up the Bayes rule computations, we get

p	Pr(p)	Pr(X=28 | p)	Pr(X=28, p)	Pr(p | X=28)
0	1/6	0	0	0
0.10	1/6	.0000312	5.21 (10^(-6))	.00037
0.20	1/6	.0724	.012	.8656
0.30	1/6	.0112	.00186	.1339
0.40	1/6	.000008	1.38 (10^(-6))	.0001
0.50	1/6	6.2 (10^(-11))	1.04 (10^(-11))	7.4 (10^(-10))

Pr(X=28) = .013933

Each entry in the third column is obtained by using the binomial formula for the corresponding value of p. For example,
Pr(X=28 | p=0.20) = (127!)/(28! 99!) $.2^{28}$ $.8^{99}$ = .0724

Each entry in the fourth column is obtained from the multiplication rule:
Pr(X=28, p) = Pr(p) Pr(X=28 | p)

Pr(X=28) is obtained by summing Pr(X=28, p) for all values of p.

Pr(p | X=28) is obtained from the definition of conditional probability:
Pr(p | X=28) = Pr(X=28, p) / Pr(X=28)

b) Pr(p < .30) = .00037 + .8656

Note that this prior distribution is very strong, in that it forces p to equal only one of 6 values. A more realistic prior distribution would allow p to range from 0 to 1. But, that's more complicated computationally than we need to show the general idea of Bayesian statistics.

7. Write the differences between Bayesian and classical inference.

Marks: 15
A machine is built to make mass-produced items. Each item made by the machine has a probability p of being defective. Given the value of p, the items are independent of each other. Because of the way in which the machines are made, p could take one of several values. In fact p = X/100 where X has a discrete uniform distribution on the interval [0, 5]. The machine is tested by counting the number of items made before a defective is produced. Find the conditional probability distribution of X given that the first defective item is the thirteenth to be made.

2. What do you mean by hypothetical population? Give example. Define the term population and sample. Write the differences between sample and population. What do you mean by population parameter and sample statistics?
3. What do you mean by optimal interference? Define inference based on MLE. What is the Procedure for Statistical Inference? What are the three components of statistical inference?

Chapter 5

Marks 1

1. What type of data do you need for a chi-square test?
   Answer: The data used in calculating a chi-square statistic must be **random, raw, mutually exclusive, drawn from independent variables, and drawn from a large enough sample**. For example, the results of tossing a fair coin meet these criteria. Chi-square tests are often used to test hypotheses.
2. What is stochastic process?
3. What is markov chain?
4. Define cluster sample.
5. What are all the four types of stochastic process?
   Some basic types of stochastic processes include **Markov processes, Poisson processes (such as radioactive decay), and time series**, with the index variable referring to time. This indexing can be either discrete or continuous, the interest being in the nature of changes of the variables with respect to time.
6. What is the need of stochastic process?

Marks 5

1. What is Markov process? Define Markov chain. Give an example.
2. Define Poisson process. What is the need of it?
3. What do you mean by time serise data? Why this kind of data is requered?
4. Under what conditions a stochastic process becomes counting process?
   A counting process is a stochastic process $\{N(t), t \geq 0\}$ with values that are non-negative, integer, and non-decreasing: $\mathbf{N(t) \geq 0}$. N(t) is an integer. If $s \leq t$ then $N(s) \leq N(t)$.
5. What is state space in stochastic process?
   The range (possible values) of the random variables in a stochastic process is called the state space of the process.
6. Why do we need stochastic process?

Since stochastic processes **provides a method of quantitative study through the mathematical model**, it plays an important role in the modern discipline or operations research.

7. Is every stochastic process a time series?

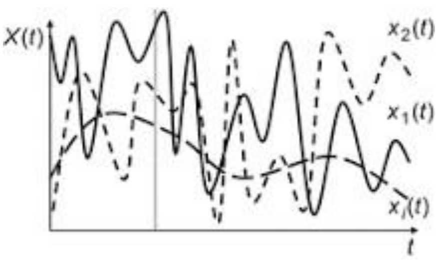
The stochastic process is a model for the analysis of time series. The stochastic process is considered to generate the infinite collection (called the ensemble) of all possible time series that might have been observed. Every member of the ensemble is a possible realization of the stochastic process.

Marks: 15

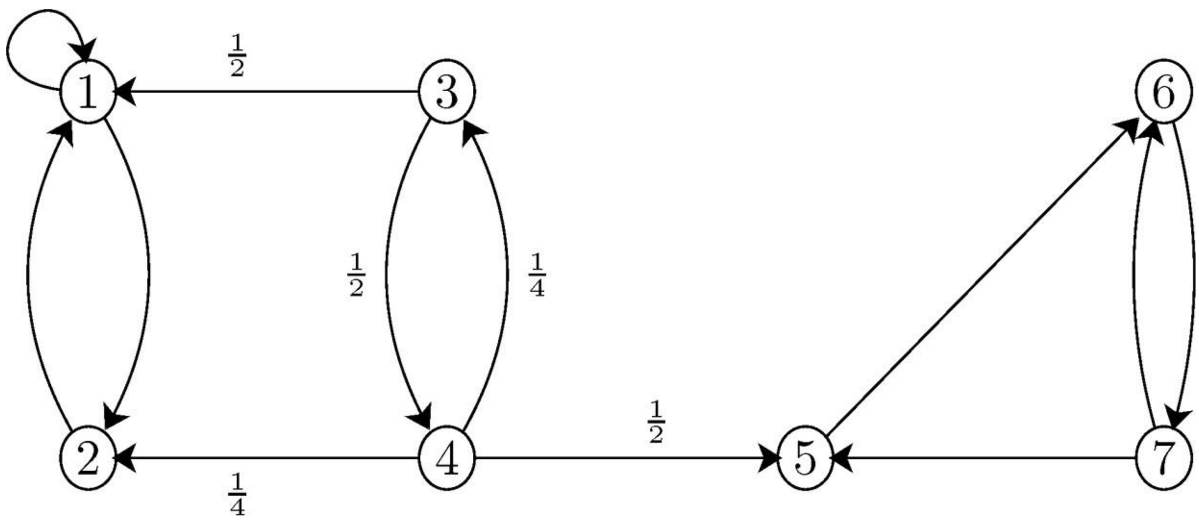
1. What is the difference between random and stochastic?

Stochastic means nondeterministic or unpredictable. Random generally means unrecognizable, not adhering to a pattern. A random variable is also called a stochastic variable.

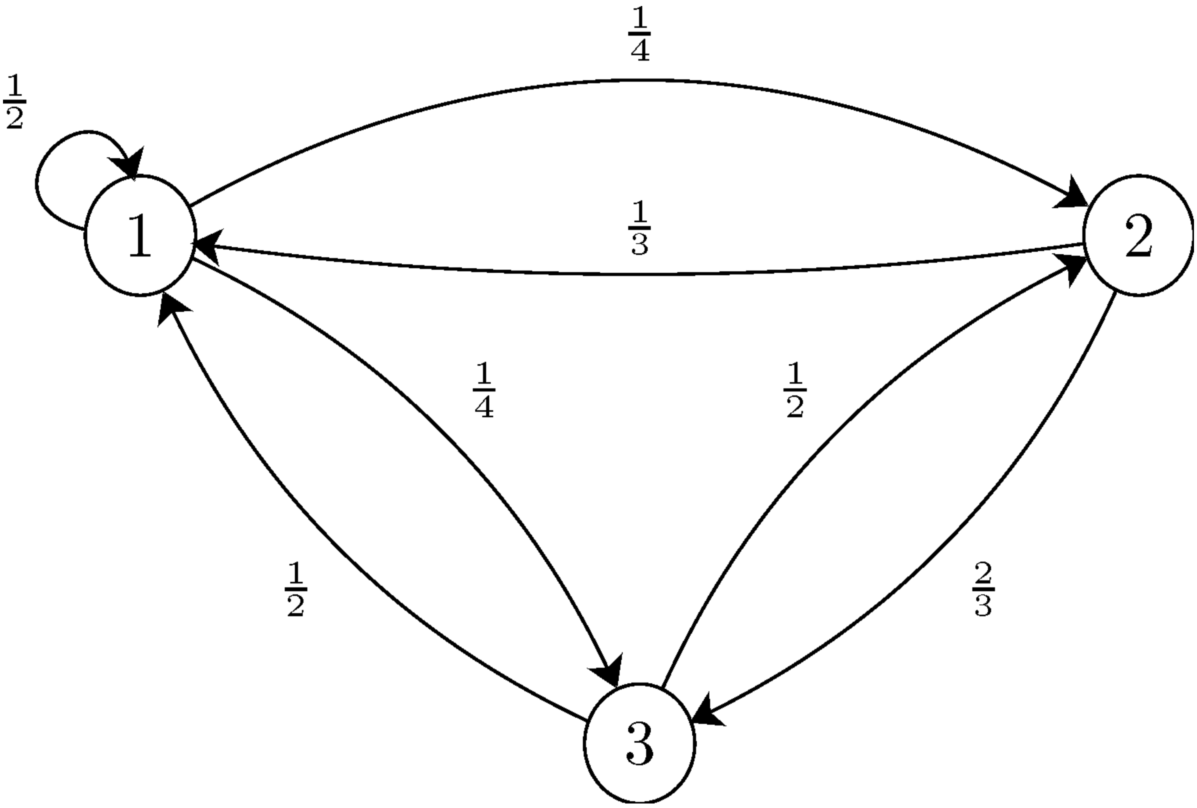
2. Does stochastic mean random?
   In general, **stochastic is a synonym for random**. For example, a stochastic variable is a random variable. A stochastic process is a random process. Typically, random is used to refer to a lack of dependence between observations in a sequence.
3. What is stochastic function?



   A stochastic (random) function X(t) is **a many-valued numerical function of an independent argument t, whose value for any fixed value t ∈ T (where T is the domain of the argument) is a random variable, called a cut set**
4. What are the two primary components that make up a stochastic process?
   Stochastic Process Meaning is one that has a system for which there are **observations at certain times, and that the outcome, that is, the observed value at each time is a random variable**.
5. Consider the Markov chain in the following figure. There are two recurrent classes, R1={1,2} and R2={5,6,7}. Assuming X0=3, find the probability that the chain gets absorbed in R1.



6. Consider the Markov chain shown in Figure.



a. Is this chain irreducible?

b. Is this chain aperiodic?

c. Find the stationary distribution for this chain.

d. Is the stationary distribution a limiting distribution for the chain?