

Int. M.Sc. (AIML) (NEP) Sem.-3 Examination

DSC-C-AIML-231T

Statistics for Machine Learning

Time : 2-00 Hours]

November-2025

[Max. Marks : 50

Instructions: All questions are compulsory. Use of non-programmable scientific calculator is allowed.

- Q.1 (a)** A discrete random variable Y takes values 0,1,2, and 3 with the following probabilities: (05)
 $P(Y = 0) = c, P(Y = 1) = 2c, P(Y = 2) = 3c, P(Y = 3) = 4c.$
 Find the value of constant c . and also Derive the PGF.

- (b)** A petrol pump is supplied with petrol once a day. If its daily volume of sales (X) in thousands of liters is distributed by: (05)

$$f(x) = 2x \quad ; \quad 0 \leq x \leq 1,$$

What must be the capacity of its tank in order that the probability its supply will be exhausted in a given day shall be 0.05?

OR

- (a)** A discrete random variable X can take integer values from 1 to 3 with probabilities given by $P(X = k) = Cx^2$, where C is a constant. (05)
 (a) Find the value of C (b) Derive the MGF (c) Calculate first two moments.
- (b)** A jar contains 2 red marbles and 3 blue marbles. Two marbles are drawn from the jar without replacement. Let X be the random variable representing the number of blue marbles drawn. Construct the Probability Mass Function (PMF) of X , and also Calculate $P(X < 2)$ (05)

- Q.2 (a)** For two random variable X and Y , Joint probabilities are given as follows: (05)

	Y_1	Y_2
X_1	0.20	0.15
X_2	0.45	0.20

- (b)** Construct the Joint PMF table with marginal probabilities. Find $E[Y/X_1]$ & $E[Y/X_2]$ Compute $E(X)$, $V(X)$ and standard deviation of X for the following distribution. (05)

$x:$	0	1	2	3	4	5
$f(x)$	0.01	0.23	0.41	0.20	0.10	0.05

OR

- (a)** A batch of 15 items contains 4 defective and 11 non-defective items. A QC inspector selects two items, one after the other, without replacement. Let X be the number of defective items in the first draw and Y be the number of defective items found in the second draw. Find $E[Y|X=1]$. (05)

- (b)** A continuous random variable X has a p. d. f (05)

$$f(x) = 3x^2, \quad 0 < x \leq 1.$$

Find the value of 'a' such that $P(X \leq a) = P(X > a)$

E 1136.2

- Q.3**
- (a) In a medical study, 25 patients are available, and 15 of them have a specific genetic marker. A doctor randomly selects 5 patients for a trial. Assuming the Hypergeometric distribution, Find the probability that at least 4 of the selected patients have the genetic marker? (05)
 - (b) A factory produces lightbulbs, and 2% of them are defective. An inspector randomly selects lightbulbs from the assembly line until they find a defective one. What is the probability that the inspector will have to check 5 lightbulbs to find the first defective one? (05)

OR

- (a) Let X is normally distributed and the mean of X is 12 and S.D is 4: (05)
Find $P(X \geq 20)$ and $P(X < 20)$. (For $Z=2$, tabulated value is 0.4772)
- (b) In a book of 520 pages, 390 typographical error occur. Assuming Poisson distribution for the number of errors per page, find the probability that a random sample of 5 pages will contain no error. (05)

- Q.4**
- (a) Ten individuals are chosen at random from a population. The mean and standard deviation of sample heights are found to be 67.8 inches and 2.8566 inches respectively. Test the hypothesis that the mean height of the population is 66. (05)
(tabulated t with 9 d. f and at 5% level of significance is 2.26)
 - (b) The mean \bar{X} and standard deviation σ of the sample are obtained as 3.4 and 2.61 respectively. If 95% confidence limits of the population mean are 3.23 to 3.57, find the size of the sample. ($Z_t = 1.96$) (05)

OR

- (a) The 90% confidence interval for the population proportion of students is obtained as 0.4 to 0.6. The estimate of the population proportion of students is 0.5. What size of the sample must have been taken? (Consider $Z_\alpha = 1.645$) (05)
- (b) A random sample of size 100 is drawn from a population of 3000 units of a product. The mean and standard deviation of the sample are 16 and 1.2. Find the 95% confidence interval for the population mean. (Consider $Z_\alpha = 1.96$) (05)

- Q.5** Attempt any **TEN** out of **TWELVE**: (Each carries 01 mark) (10)
- (1) If third central moment (μ_3) > 0 : The distribution is _____ skewed.
 - (2) The random variable X has a mean of 5 and a variance of 2, $E(3X+10)$ is _____.
 - (3) For two independent random variables X and Y , $E[XY] =$ _____.
 - (4) If the population is perfectly normal, what is the shape of the sampling distribution of mean (\bar{X}) for a sample size of $n=9$? Answer it with reason.
 - (5) From a highly right-skewed population, A random sample of size $n=49$ is drawn. What is the approximate shape of the sampling distribution of the sample mean (\bar{X}). Answer it with reason.
 - (6) The probability level below which we reject the null hypothesis is called _____, and is denoted as ____.

- (7) Mistakes in measurement, data coding, data entries, and editing are responsible factors for which type of error?
(a) Sampling error (b) non sampling error.
- (8) The sampling error in statistics is the discrepancy between a sample statistic and the true population parameter. Is it TRUE or FALSE? Give reason.
- (9) _____ Random Sampling ensures that every possible subset (or sample) of the required size has an equal chance of selection.
- (10) In _____ Sampling, individuals are selected based on a random starting point and a fixed periodic interval (k).
- (11) The Standard error (S.E) of a mean is the standard deviation of its _____ of the mean.
- (12) An estimator is called an _____ estimator, if its expected value is equal to the True value of the parameter being estimated.
