3 /51

1311E1137

Candidate's Seat No :_____

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

DSC-C-DTSC-231T

Statistics for Data Science

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

$$f(x) = 3x \quad ; 0 \leq x \leq 1,$$

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?

(b) 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.

OR

(a) A jar contains 3 red marbles and 2 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)**

(b) 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.

Q.2 **(a)** A batch of 10 items contains 4 defective and 6 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)**

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

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

OR

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

	Y_1	Y_2
X_1	0.20	0.25
X_2	0.35	0.20

Construct the Joint PMF table with marginal probabilities. Find E [Y/X_1] & E [Y/X_2]

(b) 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	0.26	0.35	0.20	0.13	0.05

P.T.O.

E1137-2

Q.3 (a) Let X is normally distributed and the mean of X is 12 and S.D is 4: **(05)**
Find P(X≥20) and P(X<20). (For Z=2, tabulated value is 0.4772)

(b) In a book of 400 pages, 350 typographical error occur. Assuming Poission distribution **(05)**
for the number of errors per page, find the probability that a random sample of 5 pages will contain no error.

OR

(a) In a medical study, 20 patients are available, and 12 of them have a specific genetic **(05)**
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?

(b) A factory produces lightbulbs, and 4% of them are defective. An inspector randomly **(05)**
selects lightbulbs from the assembly line until they find a defective one. What is the probability that the inspector will have to check 3 lightbulbs to find the first defective one?

Q.4 (a) Ten individuals are chosen at random from a population. The mean and standard **(05)**
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.
(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 **(05)**
respectively. If 95% confidence limits of the population mean are 3.23 to 3.57, find the size of the sample. (Zt=1.96)

OR

(a) The 90% confidence interval for the population proportion of students is obtained as 0.4 **(05)**
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)

(b) A random sample of size 100 is drawn from a population of 2000 units of a product. The **(05)**
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)

Q.5 Attempt any **TEN** out of **TWELVE:** (Each carries 01 mark) **(10)**

(1) For two independent random variables X and Y, E[XY] =_____.

(2) 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.

(3) 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.

(4) The random variable X has a mean of 5 and a variance of 2, E(2X+5) is_____.

(5) If third central moment (μ_3) >0: The distribution is _____skewed.

(6) The probability level below which we reject the null hypothesis is called_____, and is denoted as ___.

(7) The sampling error in statistics is the discrepancy between a sample statistic and the true population parameter. Is it TRUE or FALSE? Give reason.

E 1137. 3

(8) Mistakes in measurement, data coding, data entries, and editing are responsible factors for which type of error?
(a) Sampling error (b) non sampling error.

(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) An estimator is called an _____estimator, if its expected value is equal to the True value of the parameter being estimated.

(12) The Standard error (S.E) of a mean is the standard deviation of its_____of the mean.

_____ **** _____ **** _____