CENTRAL UNIVERSITY OF HARYANA

End Semester Examinations March 2023

Programme : **M.Sc. Statistics**
Semester : **First**
Course Title : **Distribution Theory**
Course Code : **SBS ST 01 103 C 3104**

Session: 2022-2023
Max. Time : ***3 Hours***
Max. Marks : ***70***

1. Question no. 1 has seven parts and students need to answer any four. Each part carries three and half Marks.
2. Question no. 2 to 5 have three parts and student need to answer any two parts of each question. Each part carries seven marks.

Question No. 1. **(4X3.5=14)**

a) Explain compound and truncated distributions

b) Define central and non-central t and χ^2 distributions along with its underlying assumptions.

c) Explain the bivariate normal distribution.

d) Define mixture distribution.

e) In three independent throws of fair dice, let X denote the number of upper faces showing six. Then find the value of $E(3-X)^2$.

f) Let X be a random variable with the moment generating function $M_X(t)=\frac{1}{216}(5+e^t)^3$, $t \in \Re$. Then find the value of $P(X>1)$.

g) Obtain the skewness and kurtosis of the log-normal distribution.

Question No. 2. **(2X7=14)**

a) Show that a discrete random variable is geometric if and only if it satisfies memory less properties.

b) Prove the recurrence relation between the moments of Possion distribution $\mu_{r} = \lambda\left(r\mu_{r-1}+\frac{d\mu_r}{d\lambda}\right)$, where μ_r is the r^{th} moments about mean λ. Also show that $\beta_2-\beta_1-3=0$.

c) If m things are distributed among a men and b women, show that the probability that the no. of things received by men is odd is $\frac{1}{2}\left[\frac{(b+a)^m-(b-a)^n}{(b+a)^m}\right]$.

Question No. 3. **(2X7=14)**

a) For a lognormal distribution, show that

$$E(X^r)=e^{r\mu+\frac{r^2\sigma^2}{2}}.$$

b) For a gamma $G(\alpha,\beta)$ distribution prove that

$$\mu_{r+1} = \frac{r\alpha}{\beta^2}\mu_{r-1} - \frac{\partial \mu_r}{\partial \beta},$$

and hence show that

$$\beta_2 - \beta_1 - 3 = \frac{2}{\alpha},$$

where symbols have their usual meanings.

c) Define Laplace distribution. Obtain its characteristics function. Also, find its mean and variance.

Question No. 4. **(2X7=14)**

ai) For the $t-$ distribution with k degree of freedom, prove that

$$E[X^r] = \frac{k^{r/2}\Gamma\left(\frac{k-r}{2}\right)\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{k}{2}\right)},$$

and hence show that $\beta_2 - \beta_1 - 3 = \frac{6}{k-4}$, where symbols have their usual meanings.

b) Show that for t-distribution with k degrees of freedom

$$P[X \le x] = 1 - \frac{1}{2} I_{\frac{k}{k+x^2}}\left(\frac{k}{2}, \frac{1}{2}\right),$$

where

$$I_x(a,b) = \frac{1}{B(a,b)} \int_0^x y^{a-1}(1-y)^{b-1} dy$$

is the incomplete beta function?

c) Let X_i be independently distributed as $\chi^2_{n_i}$, $i = 1,2$. Obtain the probability density function of $Y = \frac{X_1}{X_2}$.

Question No. 5. **(2X7=14)**

a) A box contains N identical ball numbered $1,2,\ldots,N$ of these n balls drawn at a time.

Let $X_1, X_2, \ldots, X_n$ denote the numbers on the n ball drawn. Find $V(S_n)$, where $S_n = \sum_{i=0}^{n} X_i$

b) Obtain the moment generating function of bivariate normal distribution.

c) Show that if X_1 and X_2 are standard normal variates with correlation coefficient ρ between them, then the correlation coefficient between X_1^2 and X_2^2 is given by ρ^2.

CENTRAL UNIVERSITY OF HARYANA

Term End Examination March 2023

Programme	**: M.Sc. / M.A**	**Session: 2022-23**
Semester	**: First**	*Max. Time: 3* ***Hours***
Course Title	**: Introductory Statistics**	
Course Code	**: SBS ST 01 101 GE 3014**	*Max. Marks:* ***50***

Instructions:

1. Question no. 1 has five parts and students need to answer any four. Each part carries two and half Marks.
2. Question no. 2 to 5 have three parts and student need to answer any two parts of each question. Each part carries five marks.

Note: Scientific Non programmable Calculator is allowed.

Question No. 1. **(4x2.5=10)**

a. Distinguish Between Primary Data and Secondary data.

b. Explain the types of errors in testing of hypothesis.

c. What are the properties of a good measure of central tendency?

d. Explain with suitable example the term 'variation'. Mention some common measures of variation and describe the one which you think is widely used.

e. Explain the following with suitable examples:

(i) Random Experiment, and

(ii) Sample Space

Question No. 2 **(2x5=10)**

a. Find the mean from the following distribution:

Age in years	15-19	20-24	25-29	30-34	35-44	45-59
No. of Persons	37	81	43	24	9	6

b. Explain the term 'Skewness' and Kurtosis used in connection with the frequency distribution of a continuous variable. Explain the measures of skewness and kurtosis.

c. Compute the standard deviation form the following distribution of marks obtained by 90 students.

Marks	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
No. of Students	5	12	15	20	18	10	6	4

Question No. 3 **(2x5=10)**

a. Define probability. Explain various approaches to probability and discuss the importance of this concept in statistics.

b. A business firm receives on an average 2.5 telephone calls per dy during the time period 10.00 – 10.05 am. Find the probability that on a certain day, the firm receives

(i) No call

(ii) Exactly 4 call, during the same period.

(assuming Poisson distribution; $e^{-2.5} = 0.0821$)

c. As a result of tests on 20000 electric bulbs manufactured by company it was found that the life time of the bulb was normally distributed with an average life of 2040 hours and standard deviation of 60 hours. On the basis of the information, estimate the number of bulbs that are expected to burn for

(i) More than 2150 hours

(ii) Less than 1960 hours

Proportion of Area under the Normal Curve

Z	Area	Z	Area	Z	Area
1.23	0.3907	1.33	0.4082	1.43	0.4236
1.63	0.4484	1.73	0.4582	1.83	0.4664

Question No. 4 **(2x5=10)**

a. Define hypotheses and point out their role. Explain different types of hypotheses.

b. In a test given to two groups of students, the marks obtained are as follows:

First Group	18	20	36	50	49	36	34	49	41
Second Group	29	28	26	35	30	44	46		

The value of t at 5 % level of significance for v= 15 is 2.13

The value of t at 5 % level of significance for v= 14 is 2.14

The value of t at 5 % level of significance for v= 16 is 2.12

c. The following data present the yield in quintals of common ten subdivisions of equal area of two agricultural plots:

Plot 1	6.2	5.7	6.5	6.0	6.3	5.8	5.7	6.0	6.0	5.8
Plot 2	5.6	5.9	5.6	5.7	5.8	5.7	6.0	5.5	5.7	5.5

The value of F for v_1= 9 and v_2= 9 is 3.18

The value of F for v_1= 10 and v_2= 10 is 2.98

Question No. 5 **(2x5=10)**

a. Define Spearman's Rank correlation coefficient. In a certain examination 10 students obtained the following marks in Statistics and Physics. Find the Spearman's rank correlation coefficient.

Student Roll No.	1	2	3	4	5	6	7	8	9	10
Marks in Statistics	90	30	82	45	32	65	40	88	73	66
Marks in Physics	85	42	75	68	45	63	60	90	62	58

b. Write a brief note on Multiple Linear Regression. Also discuss various assumptions of multiple linear regression.

c. From the following data, obtain the regression equation Y on X:

X	6	2	10	4	8
Y	9	11	5	8	7

CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, March 2023

Programme: M.Sc. (Statistics) **Session: 2022-23**

Semester: First **Max. Time: 3 Hours**

Course Title: Probability Theory **Max. Marks: 70**

Course Code: SBS ST 01 102 C 3104

Instructions:

1. Question no. 1 has seven sub parts and students need to answer any four. Each sub part carries three and half Marks.

2. Question no. 2 to 5 have three sub parts and students need to answer any two sub parts of each question. Each sub part carries seven marks.

Question No. 1. (4X3.5=14)

a) What is a 'probability space'? State (i) the 'law of total probability', and (ii) Boole's inequality for events not necessarily mutually exclusive.

b) Write the definition of σ field, borel σ field and axiomatic probability.

c) What do you mean by probability mass and probability density function? Give some important result of the mathematical expectation.

d) Examine whether CLT hold for followings

$$(i) \quad P(X_k = \pm\sqrt{k}) = \frac{1}{2}$$

$$(ii) \quad P(X_k = \pm k) = \frac{1}{2}k^{-\left(\frac{1}{2}\right)}, \; P(X_k = 0) = 1 - k^{-\left(\frac{1}{2}\right)}.$$

e) Define the moment generating function and characteristic function, and give their important properties.

f) Let X and Y be independent non-degenerate variates. Prove that $Var(XY) = Var(X)Var(Y) \; iff \; E(X)=0, E(Y)=0.$

g) Explain the different types of modes of convergence. Prove that if a sequence of a random variable $\{X_n, n=0,1,2,...\}$ converges in the r^{th} mean then it also converges in probability, symbolically, $X_n \xrightarrow{r^{th}} X \Rightarrow X_n \xrightarrow{p} X$.

Question No. 2. (2X7=14)

a) For any two events A and B, show that $P(A \cap B) \le P(A) \le P(A \cup B) \le P(A) + P(B)$. If A and B are mutually exclusive events show that $P(A/\bar{B}) = P(A)/1 - P(B)$.

b) Prove that a σ field is a monotone field.

c) p is the probability that a man aged x years will die in a year. Find the probability that out of n men $A_1, A_2, ..., A_n$ each aged x, A_1 will die in a year and will be the first to die.

Question No. 3. (2X7=14)

a) Define the marginal, joint and conditional probability density functions. Let (X,Y) be jointly distributed with density $f(x,y)=\begin{cases} e^{-y}, & 0<x<y<\infty \\ 0, & otherwise \end{cases}$, then obtain the marginal and conditional densities of X and Y, respectively.

b) What are the important properties of a random variable? A symmetric die is thrown 600 times. Find the lower bound for the probability of getting 80 to 120 sixes.

c) If X is an random variable with mean 0 and standard deviation σ then show that $F(t) \le \sigma^2/\sigma^2+t^2, t<0$ and $F(t) \ge t^2/\sigma^2+t^2, t>0$, where $F(t)=P[X \le t]$. Let X and Y independent Poisson variate, show that the conditional distribution of X given $X+Y$, is binomial distribution.

Question No. 4. (2X7=14)

a) Define the probability generating function of a random variable. Let X a positive integral valued variable, such that $P(X=n)=p_n, n=0,1,2,\ldots$. Define the probability generating function $G(s)$ for X and show that $E(X)=G'(1)$, $\text{var}(X)=G''(1)+G'(1)-[G'(1)]^2$

b) Write down the necessary and sufficient conditions for the characteristic functions. Find the density function $f(x)$ corresponding to the characteristic function $\phi(t)=\begin{cases} 1-|t|, & |t| \le 1 \\ 0, & |t|>1 \end{cases}$.

c) State and prove the inversion theorem.

Question No. 5. (2X7=14)

a) Explain the Laypounov's condition for the sequence of independent random variable $\{X_i\}$ to holds CLT. Let $X_1, X_2, X_3, \ldots, X_n$ be normally independently identically distributed variates with mean 0 and variance 1 then obtain the limiting distribution of $\dfrac{\sqrt{n}(X_1+X_2+X_3+\ldots+X_n)}{X_1^2+X_2^2+X_3^2+\ldots+X_n^2}$.

b) Explain strong and weak law of large numbers. Let $\{X_n\}$ be any sequence of random variables, write $Y_n=\frac{1}{n}[S_n-E(S_n)]$ where $S_n=X_1+X_2+X_3+\ldots+X_n$. Prove that a necessary and sufficient condition for the sequence $\{X_n\}$ to satisfy the weak law of large number is that $E\left[\dfrac{Y_n^2}{1+Y_n^2}\right] \to 0$ as $n \to \infty$.

c) State and prove the Chebyshev's inequality. For geometric distribution $p(x)=2^{-x}; x=1,2,3,\ldots$, show that Chebyshev's inequality gives $P[|X-2| \le 2] > \frac{1}{2}$.

CENTRAL UNIVERSITY OF HARYANA

Term End Examination March 2023

Programme : M.Sc. Statistics
Semester : First
Course Title : Sampling Techniques
Course Code : SBS ST 01 104 C 3104

Session: 2022-23
Max. Time: 3 Hours
Max. Marks: 70

Instructions:

1. Question no. 1 has seven parts and students need to answer any four. Each part carries three and half Marks.
2. The question number two to nine carries 14 marks each and student need to attempt any four questions from eight questions.

Question No. 1. **(7x2=14)**

a. What are the principles of sampling theory?

b. Discuss the advantages of sample survey over census survey.

c. Define non-probability sampling? Also, give its disadvantages.

d. Define the following terms: population, sampling unit, and sampling frame.

e. What are the principles of stratification?

f. Discuss the advantages and disadvantages of systematic sampling.

g. Show that in simple random sampling, the bias of the ratio estimator $\hat{R}$ is given by $B(\hat{R}) = -Cov(\hat{R}, \bar{x})/\bar{X}$.

Question No. 2 **(2x7=14)**

a. Define simple random sampling. Show that sample mean is an unbiased estimator of the population mean. Also, obtain the sampling variance of the sample mean.

b. Prove that in simple random sampling, $s^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 / (n-1)$ is an unbiased estimator of $S^2 = N\sigma^2/(N-1)$.

c. In Simple random sampling without replacement for proportions, find an unbiased estimator of the population proportion. Also, obtain the variance of this estimate.

Question No. 3 **(2x7=14)**

a. If finite population correction factor is ignored, prove that $Var_{opt} \le Var_{prop} \le Var_{SRS}$.

b. If there are two strata and ϕ is the ratio of actual n_1/n_2 to optimum n_1/n_2 for fixed sample size n, show that whatever be the values of N_1, N_2, S_1 and S_2, the relative precision of the actual allocations to the optimum allocation is never less than $\frac{4\phi}{(1+\phi)^2}$.

c. What is systematic sampling? Explain how you will estimate the variance of a systematic sample with a random start.

Question No. 4 **(2x7=14)**

a. Prove that in simple random sampling without replacement, for large n, an approximation to the variance of $\hat{R}$ is given by $Var(\hat{R}) \cong \frac{(1-f)}{n\bar{X}^2} \sum_{i=1}^{N} \frac{(y_i - Rx_i)^2}{(N-1)}$.

b. Prove that in simple random sampling without replacement, for large n, the sampling variance of the regression estimator is given by $Var(\bar{y}_{lr}) = \frac{(1-f)}{n} S_y^2 (1-\rho^2)$.

c. Prove that in simple random sampling without replacement of n clusters each containing M elements from a population of N clusters, the sample mean $\bar{y}_n$ is an unbiased estimator of the population mean $\bar{Y}$ and its variance is given by

$$Var(\bar{y}_n) \cong (1-f)\frac{S_M^2}{n}\left[1+(M-1)\rho\right].$$

Question No. 5 **(2x7=14)**

a. Show that in pps sampling without replacement an unbiased estimator of the population total Y is given by $\hat{Y}_{pps} = \frac{1}{n}\sum_{i=1}^{n}(y_i/p_i)$ and its sampling variance is given by

$$Var(\hat{Y}_{pps}) = \frac{1}{n}\sum_{i=1}^{n} p_i (y_i/p_i - Y)^2.$$

b. If the n first stage units and the m second stage units from each chosen first stage unit are selected by SRSWOR, show that $\bar{y}$ is an unbiased estimator of $\bar{y}$ and its sampling variance is given by $Var(\bar{y}) = \frac{N-n}{N}\sum_{i=1}^{n}\frac{S_b^2}{n} + \frac{M-m}{M}\sum_{i=1}^{n}\frac{S_w^2}{mn}$.

c. In case of double sampling if the value of n_i do not depend on w_i, show that the estimator $\bar{y}_{std}$ is unbiased estimator of the population mean and its sampling variance is given by

$$Var(\bar{y}_{std}) = \sum_{i=1}^{k}\left[\left\{W_i^2 + \frac{g}{n'}W_i(1-W_i)\right\}(1-f_i)\frac{S_i^2}{n_i} + \frac{g}{n'}W_i(\bar{Y}_i - \bar{Y})^2\right],$$

where, $f_i = \frac{n_i}{N_i}$ and $g = \frac{(N-n')}{(N-1)}$.