CS (MAIN) EXAM:2018

EGT-P-STSC

सांख्यिकी / STATISTICS

प्रश्न-पत्र I / Paper I

निर्धारित समय : तीन घंटे Time Allowed : **Three** Hours	अधिकतम अंक : **250** Maximum Marks : **250**

प्रश्न-पत्र सम्बन्धी विशिष्ट अनुदेश

कृपया प्रश्नों के उत्तर देने से पूर्व निम्नलिखित प्रत्येक अनुदेश को ध्यानपूर्वक पढ़ें :

*इसमें **आठ** प्रश्न हैं जो **दो खण्डों** में विभाजित हैं तथा **हिन्दी** और **अंग्रेज़ी** दोनों में छपे हैं ।*

*परीक्षार्थी को कुल **पाँच** प्रश्नों के उत्तर देने हैं ।*

*प्रश्न संख्या **1** और **5** अनिवार्य हैं तथा बाकी प्रश्नों में से प्रत्येक खण्ड से कम-से-कम **एक** प्रश्न चुनकर किन्हीं **तीन** प्रश्नों के उत्तर दीजिए ।*

प्रत्येक प्रश्न/भाग के अंक उसके सामने दिए गए हैं ।

प्रश्नों के उत्तर उसी प्राधिकृत माध्यम में लिखे जाने चाहिए जिसका उल्लेख आपके प्रवेश-पत्र में किया गया है, और इस माध्यम का स्पष्ट उल्लेख प्रश्न-सह-उत्तर (क्यू.सी.ए.) पुस्तिका के मुख-पृष्ठ पर निर्दिष्ट स्थान पर किया जाना चाहिए । उल्लिखित माध्यम के अतिरिक्त अन्य किसी माध्यम में लिखे गए उत्तर पर कोई अंक नहीं मिलेंगे ।

यदि आवश्यक हो, तो उपयुक्त आँकड़ों का चयन कीजिए तथा उनको निर्दिष्ट कीजिए ।

जब तक उल्लिखित न हो, संकेत तथा शब्दावली प्रचलित मानक अर्थों में प्रयुक्त हैं ।

प्रश्नों के उत्तरों की गणना क्रमानुसार की जाएगी । यदि काटा नहीं हो, तो प्रश्न के उत्तर की गणना की जाएगी चाहे वह उत्तर अंशतः दिया गया हो । प्रश्न-सह-उत्तर पुस्तिका में खाली छोड़ा हुआ पृष्ठ या उसके अंश को स्पष्ट रूप से काटा जाना चाहिए ।

Question Paper Specific Instructions

Please read each of the following instructions carefully before attempting questions :

*There are **EIGHT** questions divided in **TWO SECTIONS** and printed both in **HINDI** and in **ENGLISH**.*

*Candidate has to attempt **FIVE** questions in all.*

*Questions no. **1** and **5** are compulsory and out of the remaining, any **THREE** are to be attempted choosing at least **ONE** question from each section.*

The number of marks carried by a question / part is indicated against it.

An... be written in the medium authorized in the Admission Certificate which must be ...rly on the cover of this Question-cum-Answer (QCA) Booklet in the space provided. No ...will be given for answers written in a medium other than the authorized one.

...ssume suitable data, if considered necessary, and indicate the same clearly.

Unless and otherwise indicated, symbols and notations carry their usual standard meanings.

Attempts of questions shall be counted in sequential order. Unless struck off, attempt of a question shall be counted even if attempted partly. Any page or portion of the page left blank in the Question-cum-Answer Booklet must be clearly struck off.

खण्ड A

SECTION A

Q1. (a) हमारे पास एक थैले में 3 सिक्के हैं । उनमें से एक सिक्का अनभिनत और बाकी चालाकी वाले अभिनत सिक्के हैं । जब तीनों सिक्के उछाले जाते हैं, तो उन पर 'चित' आने की प्रायिकता क्रमश: 0·5, 0·6, 0·1 है । यदि इन सिक्कों में से एक सिक्का यादृच्छिक तरीके से चुना जाता है और तीन बार उछाला जाता है, तो

(i) P(HTT) क्या है ?

(ii) यह मानते हुए कि {HTT} आता है, तो अनभिनत सिक्का चुनने की प्रायिकता क्या है ?

यहाँ पर {H, T} क्रमश: चित तथा पट निर्दिष्ट करते हैं ।

We have a bag with 3 coins in it. One of them is a fair coin, but the others are biased trick coins. When flipped the three coins come up 'heads' with probability 0·5, 0·6, 0·1 respectively. If one of these coins is picked at random and flipped three times,

(i) What is P(HTT) ?

(ii) Assuming that {HTT} occurred, what is the probability of having chosen a fair coin ?

Here {H, T} denote Heads and Tails respectively. *10*

(b) मान लीजिए कि $X_1, X_2, \ldots, X_n$ उस बंटन से एक प्रतिदर्श है जिसका घनत्व फलन

$$f(x \mid \theta) = \theta(\theta + 1)\, x^{\theta - 1} (1 - x), \;\; 0 < x < 1, \;\; \theta > 0 \text{ है ।}$$

θ का आघूर्ण विधि से आकलन प्राप्त कीजिए ।

Let $X_1, X_2, \ldots, X_n$ be a sample from a distribution with density function

$$f(x \mid \theta) = \theta(\theta + 1)\, x^{\theta - 1} (1 - x), \;\; 0 < x < 1, \;\; \theta > 0.$$

Obtain an estimate of θ by method of moments. *10*

(c) मान लीजिए कि दो यादृच्छिक चर एक ही जैसे हैं और स्वतंत्र रूप से बंटित हैं जिनका बंटन एकसमान U(0, 1) है । $U = \frac{X}{Y}$ का बंटन व्युत्पन्न कीजिए । E(U) पर टिप्पणी कीजिए ।

Consider two independent and identically distributed uniform U(0, 1) random variables. Derive the distribution of $U = \frac{X}{Y}$. Comment on E(U). *10*

(d) निम्नलिखित आँकड़े दो विभिन्न प्रकार की बैटरियों की जीवन अवधि (घण्टों में) प्रकट करते हैं :

प्रकार A	40	30	40	45	55	30
प्रकार B	50	50	45	55	60	40

दोनों प्रतिदर्श एक ही समष्टि से आए हैं, के परीक्षण के लिए माध्यिका परीक्षण का प्रयोग कीजिए ।

The following data represents lifetime (hours) of batteries for two different brands :

Brand A	40	30	40	45	55	30
Brand B	50	50	45	55	60	40

Use the Median test to conclude if the two samples come from the same population. *10*

(e) केन्द्रीय सीमा प्रमेय का प्रयोग करते हुए दर्शाइए कि

$$\lim_{n \to \infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

Using central limit theorem, show that

$$\lim_{n \to \infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$ *10*

Q2. (a) X, Y का संयुक्त घनत्व है

$$f(x, y) = e^{-y}, \quad 0 \le x \le y < \infty.$$

X, Y के उपान्त बंटनों तथा X के दिए होने पर Y के सप्रतिबन्ध बंटन को ज्ञात कीजिए ।

The joint density of X, Y is

$$f(x, y) = e^{-y}, \quad 0 \le x \le y < \infty.$$

Find the marginal distributions of X, Y and the conditional distribution of Y given X. *20*

(b) मान लीजिए कि $X_1, X_2, ..., X_n$ स्वतंत्र व समरूप बंटित द्विपद (1, p) यादृच्छिक चर हैं । $p(1 - p)$ का एक प्रतिदर्श के आधार पर UMVUE ज्ञात कीजिए ।

Let $X_1, X_2, ..., X_n$ be iid binomial (1, p) random variables. Find the UMVUE for $p(1 - p)$, based on the sample. *15*

(c) मान लीजिए कि $x_1, x_2, ..., x_n$, $N(0, \theta)$ से एक यादृच्छिक प्रतिदर्श है । प्रसरण θ अज्ञात है । $H_0 : \theta = \theta_0$ का $H_1 : \theta > \theta_0$ के विरुद्ध परीक्षण के लिए सर्वाधिक शक्तिमान परीक्षण व्युत्पन्न कीजिए, जबकि θ_0 एक निश्चित धनात्मक संख्या है । क्या यह UMP परीक्षण है ?

Let $x_1, x_2, ..., x_n$ be a random sample from $N(0, \theta)$. The variance θ is unknown. Derive the most powerful test for $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$, where θ_0 is a fixed positive number. Is it a UMP test ? 15

Q3. (a) एक परिक्षेत्र में बने लक्ष्य को भेदने के लिए निशाने लगाए जाते हैं जो केन्द्र से R की दूरी पर परिक्षेत्र पर पहुँचते हैं, जहाँ कि R का प्रायिकता घनत्व फलन है

$$f_R(r; \sigma^2) = \frac{1}{\sigma^2}\, r\, e^{-r^2/2\sigma^2};\ r \geq 0.$$

यदि $R \leq a$ तो निशाने को एक अंक मिलता (सफल माना जाता) है । यदि n निशाने लगाए जाते हैं और उनमें से m वास्तव में सफल होते हैं तथा भेद बिन्दुओं की केन्द्र से दूरियाँ $r_1, r_2, ..., r_m$ हैं, तो सिद्ध कीजिए कि σ^2 का अधिकतम संभाविता आकलक

$$\hat{\sigma}^2 = \frac{1}{2m} \sum_{i=1}^{m} r_i^2 + \frac{n-m}{2m} a^2 \text{ है ।}$$

Shots aimed at a target reach the plane of the target at a distance R from the centre, where R has the probability density function

$$f_R(r; \sigma^2) = \frac{1}{\sigma^2}\, r\, e^{-r^2/2\sigma^2};\ r \geq 0.$$

A hit is scored if $R \leq a$. If n shots are fired and there are exactly m hits, these being at distances $r_1, r_2, ..., r_m$ from the centre, prove that the maximum likelihood estimator of σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{2m} \sum_{i=1}^{m} r_i^2 + \frac{n-m}{2m} a^2.$$ 20

(b) मान लीजिए कि $\{X_n\}$ यादृच्छिक चरों का एक क्रम है तथा $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

सिद्ध कीजिए कि क्रम $\{X_n\}$ के बृहत् संख्याओं के दुर्बल नियम को संतुष्ट करने के लिए आवश्यक तथा पर्याप्त प्रतिबन्ध

$$E\left[\frac{Y_n^2}{1+Y_n^2}\right] \to 0 \text{ जैसे } n \to \infty \text{ है ।}$$

Let $\{X_n\}$ be a sequence of random variables and $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Prove that a necessary and sufficient condition for the sequence $\{X_n\}$ to satisfy the weak law of large numbers is that

$$E\left[\frac{Y_n^2}{1+Y_n^2}\right] \to 0 \text{ as } n \to \infty.$$ *15*

(c) (i) मान लीजिए कि S_n एक सुडौल (सममित) पाँसे को n बार उछालने से प्राप्त अंकों का योग है ।

$$P\left[\left|\frac{S_n}{n} - 3{\cdot}5\right| > \varepsilon\right], \text{ जहाँ } \varepsilon > 0$$

की उपरि सीमा का आकलन कीजिए ।

Let S_n be equal to the total obtained in n tosses of a symmetric die. Estimate an upper bound for

$$P\left[\left|\frac{S_n}{n} - 3{\cdot}5\right| > \varepsilon\right], \text{ for } \varepsilon > 0.$$

(ii) मान लीजिए कि $n = 1, 2, \ldots$ के लिए X_n का घनत्व फलन

$P[X_n = 1] = \frac{1}{n}$ और $P[X_n = 0] = 1 - \frac{1}{n}$ है ।

जाँच कीजिए कि क्या X_n प्रायिकता में शून्य को अभिसरित होता है ।

Let X_n have the density function, for $n = 1, 2, \ldots,$

$P[X_n = 1] = \frac{1}{n}$ and $P[X_n = 0] = 1 - \frac{1}{n}$.

Check whether X_n converges to zero in probability. *15*

Q4. (a) यदि $X_n \xrightarrow{L} X, Y_n \xrightarrow{P} C$, तो दर्शाइए कि

(i) $X_n + Y_n \xrightarrow{L} X + C$

(ii) $X_n \cdot Y_n \xrightarrow{L} CX$

(iii) $X_n / Y_n \xrightarrow{L} X/C$, बशर्ते $C \neq 0$.

If $X_n \xrightarrow{L} X,\ Y_n \xrightarrow{P} C$, then show that

(i) $X_n + Y_n \xrightarrow{L} X + C$

(ii) $X_n \cdot Y_n \xrightarrow{L} CX$

(iii) $X_n / Y_n \xrightarrow{L} X/C$, provided $C \neq 0$. *20*

(b) मान लीजिए कि $X \sim f_j$, $j = 0, 1$, जहाँ

x :	1	2	3	4	5
f_0 :	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
f_1 :	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$

$\frac{2}{5}$ के आकार के सर्वोत्तम क्रान्तिक क्षेत्र को ज्ञात कीजिए । इसकी क्षमता क्या है ?

Let $X \sim f_j$, $j = 0, 1$, where

x :	1	2	3	4	5
f_0 :	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
f_1 :	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$

Find the best critical region of size $\frac{2}{5}$. What is its power ? *15*

(c) (i) एक साधारण परिकल्पना $H_0 : \theta = \theta_0$ जहाँ कि वैकल्पिक परिकल्पना $H_1 : \theta = \theta_1$ है, के परीक्षण के लिए सिद्ध कीजिए कि अनुक्रमिक प्रायिकता अनुपात परीक्षण (SPRT) प्रायिकता एक के साथ समाप्त होता है ।

For testing a simple hypothesis $H_0 : \theta = \theta_0$ against simple alternative $H_1 : \theta = \theta_1$, prove that the Sequential Probability Ratio Test (SPRT) terminates with probability one.

(ii) मान लीजिए कि यादृच्छिक प्रेक्षण $X_1, X_2, \ldots$ बर्नूली बंटन $P(X = 1) = p$ तथा $P(X = 0) = 1 - p$ से हैं । $H_0 : p = p_0$ विकल्प (विरुद्ध) $H_1 : p = p_1$ के लिए अनुक्रमिक प्रायिकता अनुपात परीक्षण प्राप्त कीजिए ।

Let $X_1, X_2, \ldots$ be random observations from the Bernoulli distribution with $P(X = 1) = p$ and $P(X = 0) = 1 - p$. Obtain a sequential probability ratio test for $H_0 : p = p_0$ against $H_1 : p = p_1$. *15*

खण्ड B

SECTION B

Q5. (a) एक दोलक के दोलन का अवधिकाल t, $2\pi\sqrt{l/g}$ है जहाँ कि l दोलक की लम्बाई तथा g गुरुत्वीय अचर है । एक प्रयोग में दोलक की लम्बाइयों l_i (i = 1, ..., k) के लिए प्राप्त किए गए अवधिकाल t_{ij} (j = 1, 2, ..., n_i) हैं । प्रेक्षणों की त्रुटियों को असहसंबंधित, शून्य माध्य और σ^2 प्रसरण वाली मानते हुए दर्शाइए कि $\beta = 2\pi/\sqrt{g}$ का सर्वोत्तम रैखिक आकलक

$$\hat{\beta} = \frac{\sum (l_i)^{\frac{1}{2}} \sum_j t_{ij}}{\sum_i n_i l_i} \text{ है और } V(\hat{\beta}) = \frac{\sigma^2}{\sum_i n_i l_i}$$

The period of oscillation, t of a pendulum is $2\pi\sqrt{l/g}$ where l is the length of the pendulum and g is the gravitational constant. The periods observed are t_{ij} (j = 1, 2, ..., n_i) and lengths l_i (i = 1, ..., k) of the pendulum, in an experiment. Assuming the errors of observations to be uncorrelated with zero means and variances σ^2, show that the best linear estimator of $\beta = 2\pi/\sqrt{g}$ is

$$\hat{\beta} = \frac{\sum (l_i)^{\frac{1}{2}} \sum_j t_{ij}}{\sum_i n_i l_i} \text{ and } V(\hat{\beta}) = \frac{\sigma^2}{\sum_i n_i l_i}$$ *10*

(b) प्रधान घटकों की परिभाषा दीजिए । क्या प्रधान घटक मानकीकरण से प्रभावित होता है ? सहप्रसरण आव्यूह

$$\sum = \begin{bmatrix} 1 & 4 \\ 4 & 100 \end{bmatrix}$$

के उदाहरण के साथ अपने तर्क की व्याख्या कीजिए ।

Define principal components. Is the principal component affected by standardization ? Establish your claim by illustrating the same with the covariance matrix

$$\sum = \begin{bmatrix} 1 & 4 \\ 4 & 100 \end{bmatrix}.$$ *10*

(c) (i) एक 2^k बहु-उपादानी प्रयोग में पूर्ण तथा आंशिक संकरण में अंतर बताइए ।

Distinguish between total and partial confounding in a 2^k factorial experiment.

(ii) यदि 4 के आकार के खंडकों में चार पुनरावृत्तियों के साथ एक 2^3 बहु-उपादानी प्रयोग लगाया जाता है तथा उच्चतम कोटि के अन्योन्यक्रिया विन्यास पूर्ण-संकरणित हैं, तो परिणामों का विश्लेषण करने हेतु, स्वतंत्रता की कोटियों का नियतन कीजिए ।

If a 2^3 factorial experiment is laid in a block of size 4 with four replicates, write down the allocations of degrees of freedom for analyzing the results of such a design when the highest order interaction is totally confounded. *10*

(d) एक समष्टि में सात इकाइयाँ हैं जिनमें 30, 10, 25, 5, 20, 6 और 4 क्षेत्र हैं । यादृच्छिक संख्याओं 72, 46 और 94 का प्रयोग करते हुए 3 इकाइयों के आकार का एक PPS प्रतिदर्श चुनिए ।

A population has seven units consisting of 30, 10, 25, 5, 20, 6 and 4 fields. Select a PPS sample of units of size 3, using random numbers 72, 46 and 94. *10*

(e) यादृच्छिकीकृत खण्डक अभिकल्पना, जिसमें t उपचार तथा r पुनरावृत्तियाँ हैं, में एक प्रेक्षण खोया पाया जाता है । खोए हुए प्रेक्षण के आकलन की विधि की व्याख्या कीजिए तथा प्रसरण-विश्लेषण को सम्पन्न कीजिए ।

In a randomized block design with t treatments and r replicates, one observation was found missing. Explain the method of estimating the missing value and perform the analysis of variance. *10*

Q6. (a) संतुलित अभिकल्पना से आप क्या समझते हैं ? मान लीजिए कि चार उपचार (1, 2, 3, 4) पाँच खण्डकों (I, II, III, IV, V) में निम्न तरह से लगाए जाते हैं :

I	1	1	2
II	1	2	2
III	3	4	
IV	3	4	
V	3	4	

स्थापित कीजिए कि उपर्युक्त अभिकल्पना असंतुलित है ।

What do you understand by Balanced Design ? Suppose four treatments (1, 2, 3, 4) are laid out in five blocks (I, II, III, IV, V) as given below :

I	1	1	2
II	1	2	2
III	3	4	
IV	3	4	
V	3	4	

Establish that the above design is unbalanced. *20*

(b) यदि यादृच्छिक चरों वाले p-घटक सदिश X का बंटन $N_p(0, \Sigma)$ है, तो दर्शाइए कि $X' \Sigma^{-1} X$, p स्वतंत्रता कोटि वाले काई-वर्ग (chi-square) की भाँति बंटित है ।

If X is a p-component vector of random variables distributed as $N_p(0, \Sigma)$, then show that $X' \Sigma^{-1} X$ is distributed as chi-square with p degrees of freedom. 15

(c) आकलन की अनुपात एवं समाश्रयण विधियों की तुलना कीजिए । दर्शाइए कि अनुपात विधि समष्टि के योग का अभिनत आकलन देती है । अभिनति के सन्निकट व्यंजक को व्युत्पन्न कीजिए ।

Compare ratio and regression methods of estimation. Show that ratio method gives a biased estimation of population total. Derive an approximate expression of the bias. 15

Q7. (a) $N_4(\mu, \Sigma)$ से n आकार के यादृच्छिक प्रतिदर्श से हम निम्न प्राप्त करते हैं :

$\hat{\mu} = (18, 15, 18, 14)$ और

$$\hat{\Sigma} = \begin{bmatrix} 9{\cdot}5 & & & \\ 5{\cdot}2 & 5{\cdot}4 & & \\ 6{\cdot}9 & 5{\cdot}1 & 10{\cdot}0 & \\ 4{\cdot}6 & 3{\cdot}5 & 5{\cdot}6 & 4{\cdot}5 \end{bmatrix}.$$

For a random sample of size n from $N_4(\mu, \Sigma)$, we obtain

$\hat{\mu} = (18, 15, 18, 14)$ and

$$\hat{\Sigma} = \begin{bmatrix} 9{\cdot}5 & & & \\ 5{\cdot}2 & 5{\cdot}4 & & \\ 6{\cdot}9 & 5{\cdot}1 & 10{\cdot}0 & \\ 4{\cdot}6 & 3{\cdot}5 & 5{\cdot}6 & 4{\cdot}5 \end{bmatrix}.$$

(i) (X_1, X_2) के दिए होने पर (X_3, X_4) के सप्रतिबन्ध बंटन के प्राचलों के आकलकों को ज्ञात कीजिए ।

Find the estimates of the parameters of the conditional distribution of (X_3, X_4) given (X_1, X_2).

(ii) आंशिक सहसंबंध $r_{34 \cdot 12}$ को ज्ञात कीजिए ।

Find the partial correlation $r_{34 \cdot 12}$. 20

(b) A, B, C, D और E उपादानों वाला एक 2^5 प्रयोग 8 के आकार वाले खण्डकों में लगाया जाता है । 1 और 2 पुनरावृत्तियों में कुंजिका खंडक क्रमश:

1, bd, abde, cde, acd, bce, abc, ae और

1, ac, abd, ade, bcd, cde, be, abce हैं ।

प्रत्येक पुनरावृत्ति में संकरणित प्रभावों की पहचान कीजिए ।

A 2^5 experiment involving factors A, B, C, D and E is conducted in the blocks of size 8. The key blocks in the replicates 1 and 2 are respectively given by

1, bd, abde, cde, acd, bce, abc, ae and

1, ac, abd, ade, bcd, cde, be, abce.

Identify the confounded effects in each replication. 15

(c) दो दाँत के डॉक्टर, A और B, एक गाँव के 200 बच्चों के दाँतों की स्थिति का सर्वेक्षण करते हैं । डॉक्टर A, 20 बच्चों का एक साधारण यादृच्छिक प्रतिदर्श चुनता है और प्रत्येक बच्चे के क्षयित दाँतों को गिन कर निम्न परिणाम प्राप्त करता है :

क्षयित दाँतों की संख्या	0	1	2	3	4	5	6	7	8	9	10
बच्चों की संख्या	8	4	2	2	1	1	0	0	0	1	1

डॉक्टर B दाँतों की उसी विधि का प्रयोग करते हुए सभी 200 बच्चों की जाँच करता है और केवल उन बच्चों को रिकॉर्ड करता है जिनके कोई दाँत क्षयित नहीं हैं । वह 60 बच्चों में कोई दाँत क्षयित नहीं पाता ।

क्षयित दाँतों की कुल संख्या का आकलन कीजिए :

(i) केवल डॉक्टर A के परिणामों द्वारा

(ii) दोनों डॉक्टर A और डॉक्टर B के परिणामों द्वारा

किस आकलक को आप वरीयता देंगे और क्यों ?

Two dentists A and B make a survey of the state of the teeth of 200 children in a village. Dr. A selects a simple random sample of 20 children and counts the number of decayed teeth for each child with the following results :

No. of decayed teeth	0	1	2	3	4	5	6	7	8	9	10
No. of children	8	4	2	2	1	1	0	0	0	1	1

Dr. B using the same dental technique, examines all 200 children, recording merely those who have no decayed teeth. He finds 60 children with no decayed teeth.

Estimate the total number of decayed teeth using

(i) Dr. A's results only

(ii) Both Dr. A's and Dr. B's results

Which estimate do you prefer and why ? 15

Q8. (a) स्तरित यादृच्छिक प्रतिचयन की व्याख्या कीजिए और इसकी साधारण यादृच्छिक प्रतिचयन से वरीयता के कारण बताइए ।

एक प्रतिदर्श लेने वाला नेमेन नियतन उपयोग करने के बजाय दो स्तरों में प्रशासनिक सुविधा के लिए $n_1 = n_2$ लेना पसन्द करता है । यदि $V(\bar{y}_{st})$ तथा $V_{opt}(\bar{y}_{st})$ क्रमशः $n_1 = n_2$ एवं नेमेन नियतन के प्रसरणों को प्रकट करते हैं, तो दर्शाइए कि प्रसरण में भिन्नात्मक वृद्धि है

$$\frac{V(\bar{y}_{st}) - V_{opt}(\bar{y}_{st})}{V_{opt}(\bar{y}_{st})} = \left(\frac{r-1}{r+1}\right)^2$$

जहाँ कि $r = n_1/n_2$ जो कि नेमेन नियतन से मिलता है ।

Explain stratified random sampling and give reasons for its preference to simple random sampling.

With two strata, a sampler would like to have $n_1 = n_2$ for administrative convenience instead of using the values given by the Neyman allocation. If $V(\bar{y}_{st})$ and $V_{opt}(\bar{y}_{st})$ denote the variances given by the $n_1 = n_2$ and Neyman allocations respectively, show that the fractional increase in the variance is

$$\frac{V(\bar{y}_{st}) - V_{opt}(\bar{y}_{st})}{V_{opt}(\bar{y}_{st})} = \left(\frac{r-1}{r+1}\right)^2$$

where $r = n_1/n_2$ as given by Neyman allocation. *20*

(b) दर्शाइए कि हॉटेलिंग (Hotelling) का T^2, स्टूडेन्ट (Student) के t का व्यापकीकरण है । चरों के व्युत्क्रमणीय रैखिक रूपांतरण के लिए T^2 की निश्चरता के गुण को सिद्ध कीजिए । T^2 के विभिन्न उपयोगों का संक्षेप में उल्लेख कीजिए ।

Show that Hotelling's T^2 is a generalization of the Student's t. Prove the invariance property of T^2 under the non-singular linear transformations of the variables. Briefly mention different uses of T^2. *15*

(c) एक निदर्श

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \; i = 1, 2, 3$$

के लिए जहाँ कि $X_1 = -1, \; X_2 = 0, \; X_3 = 1, \; \beta_1$ और β_0 के BLUE प्राप्त कीजिए । मान लीजिए कि यह निदर्श सही नहीं है और सही निदर्श

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i \text{ है ।}$$

प्राप्त किए हुए BLUE's की अभिनति को ज्ञात कीजिए ।

For a model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \; i = 1, 2, 3$$

where $X_1 = -1, \; X_2 = 0, \; X_3 = 1$, find BLUE's of β_1, β_0. Suppose the model is not correct and the true model is

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i.$$

Find the bias of the BLUE's obtained. *15*