सांख्यिकी / STATISTICS

प्रश्न-पत्र I / Paper I

निर्धारित समय : **तीन** घंटे	अधिकतम अंक : **250**
Time Allowed : ***Three*** *Hours*	*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 । प्राधिकृत 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.

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

Assume 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) मान लीजिए कि एक सिक्के को तीन बार उछाला जाता है । यदि X पहले उछाल पर सिरों (चितों) की संख्या दर्शाता है और Y, समस्त सिरों (चितों) की संख्या है, तो $Z = X + Y$ का प्रायिकता बंटन लिखिए ।

Suppose a coin is tossed three times. If X denotes the number of heads on first toss and Y is the total number of heads, then write down the probability distribution of $Z = X + Y$. *10*

(b) सिद्ध कीजिए कि

$X_n \xrightarrow{P} 0$ यदि और केवल यदि $X_n \xrightarrow{d} 0$.

Prove that

$X_n \xrightarrow{P} 0$ if and only if $X_n \xrightarrow{d} 0$. *10*

(c) मान लीजिए कि

$$f(x, \alpha) = \alpha\, x^{\alpha-1} e^{-x^{\alpha}},\ x > 0,\ \alpha > 0$$

के द्वारा प्रदत्त एक प्रायिकता घनत्व फलन $f(x, \alpha)$ से, X_1 और X_2 स्वतन्त्र और सर्वसम बंटित प्रेक्षण हैं ।

दर्शाइए कि $\dfrac{(\log X_1)}{(\log X_2)}$ एक सहायक प्रतिदर्शज है ।

Suppose X_1 and X_2 are independent and identically distributed (i.i.d) observations from the pdf given by

$$f(x, \alpha) = \alpha\, x^{\alpha-1} e^{-x^{\alpha}},\ x > 0,\ \alpha > 0.$$

Show that $\dfrac{(\log X_1)}{(\log X_2)}$ is an ancillary statistic. *10*

(d) एक दाग़ (धब्बा) अपसारक के निर्माता का दावा है कि उसका उत्पाद 90% दाग़ों को मिटा देता है । यदि एक यादृच्छिक प्रतिदर्श में, निर्माता का उत्पाद, 20 में से 17 दाग़ मिटाए, तो शून्य परिकल्पना $H_0 : \theta = 0{\cdot}90$ का वैकल्पिक परिकल्पना $H_1 : \theta < 0{\cdot}90$ के विरुद्ध, 0·05 सार्थकता स्तर पर परीक्षण कीजिए और टिप्पणी कीजिए ।

The manufacturer of a spot remover claims that his product removes 90% of all spots. If, in a random sample, 17 of 20 spots were removed with the manufacturer's product, test the null hypothesis $H_0 : \theta = 0{\cdot}90$ against the alternative hypothesis $H_1 : \theta < 0{\cdot}90$ at the 0·05 significance level and comment. *10*

(e) मान लीजिए कि X_1, X_2, X_3 तीन स्वतन्त्र यादृच्छिक चर हैं जिनका समान प्रायिकता घनत्व फलन (पी.डी.एफ.)

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0, & \text{अन्यथा} \end{cases} \text{ है ।}$$

मान लीजिए कि $Y = \min(X_1, X_2, X_3)$. Y का आघूर्ण उत्पादक फलन (mgf) ज्ञात कीजिए । आघूर्ण उत्पादक फलन (mgf) का उपयोग करते हुए, $E(Y)$ और $E(Y^2)$ ज्ञात कीजिए ।

Let X_1, X_2, X_3 be three independent random variables with common pdf

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Let $Y = \min(X_1, X_2, X_3)$. Find mgf of Y. Using mgf, find $E(Y)$ and $E(Y^2)$. *10*

Q2. (a) $X_1, ..., X_n$ प्रायिकता घनत्व फलन

$$f(x \mid \theta) = \theta\, x^{-2}, \;\; 0 < \theta \leq x < \infty$$

से एक यादृच्छिक प्रतिदर्श है ।

(i) θ के लिए पर्याप्त प्रतिदर्शज क्या है ?

(ii) θ का अधिकतम संभाविता आकलक (एम.एल.ई.) ज्ञात कीजिए ।

(iii) आघूर्ण विधि (एम.ओ.एम.) से θ का आकलक ज्ञात कीजिए ।

$X_1, ..., X_n$ is a random sample from probability density function

$$f(x \mid \theta) = \theta\, x^{-2}, \;\; 0 < \theta \leq x < \infty.$$

(i) What is a sufficient statistic for θ ?

(ii) Find the Maximum Likelihood Estimate (MLE) of θ.

(iii) Find the Method of Moments (MoM) estimate of θ. *20*

(b) दर्शाइए कि यादृच्छिक चरों की प्रत्येक बर्नूली श्रृंखला केन्द्रीय सीमा नियम का अनुसरण करती है ।

Show that every Bernoulli sequence of random variables obeys the central limit law. *15*

(c) मान लीजिए कि X का प्रायिकता घनत्व फलन

$$f(x, \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & 0 < x < \infty \text{ है ।} \\ 0, & \text{अन्यथा} \end{cases}$$

यह भी मान लीजिए कि $H_0 : \theta = 2$ का परीक्षण $H_1 : \theta = 4$ के विरुद्ध करने के लिए, X_1 और X_2 का 2 के आकार का एक यादृच्छिक प्रतिदर्श है । एक परीक्षण

$$C = \{(X_1, X_2);\ 9{\cdot}5 \le X_1 + X_2 < \infty\}$$

क्रान्तिक क्षेत्र के द्वारा परिभाषित है ।

परीक्षण के आकार और क्षमता का निर्धारण कीजिए ।

Let X have the pdf

$$f(x, \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & 0 < x < \infty \\ 0, & \text{otherwise.} \end{cases}$$

To test $H_0 : \theta = 2$ against $H_1 : \theta = 4$, let X_1 and X_2 be a random sample of size 2. A test is defined by taking critical region as

$$C = \{(X_1, X_2);\ 9{\cdot}5 \le X_1 + X_2 < \infty\}.$$

Determine size and power of the test. *15*

Q3. (a) मान लीजिए कि X का प्रायिकता घनत्व फलन

$$f(x \mid \theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, \quad x = -1, 0, 1, \quad 0 \le \theta \le 1$$

से एक प्रेक्षण है ।

(i) क्या X एक सम्पूर्ण पर्याप्त प्रतिदर्शज है ? यह भी जाँचिए कि क्या $|X|$ भी एक सम्पूर्ण पर्याप्त प्रतिदर्शज है ।

(ii) क्या $f(x \mid \theta)$ चरघातांकी श्रेणी से सम्बन्ध रखता है ?

(iii) θ का अधिकतम संभाविता आकलक (एम.एल.ई.) ज्ञात कीजिए ।

Let X be one observation from the pdf

$$f(x \mid \theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, \quad x = -1, 0, 1, \quad 0 \le \theta \le 1.$$

(i) Is X a complete sufficient statistic ? Check also if $|X|$ is a complete sufficient statistic.

(ii) Does $f(x \mid \theta)$ belong to exponential class ?

(iii) Find the MLE of θ. *20*

(b) मान लीजिए कि X तथा Y का संयुक्त प्रायिकता घनत्व फलन

$$f(x, y) = C \cdot \exp\{-(x^2 - 4y^2 - xy)\}$$

है, जहाँ C एक अचर है ।

ज्ञात कीजिए :

(i) E(X) और V(X)

(ii) E(Y) और V(Y)

(iii) X तथा Y के बीच सहसंबंध

X और Y के बंटनों (वितरणों) की भी पहचान कीजिए ।

Let the joint probability density function of X and Y be

$$f(x, y) = C \cdot \exp\{-(x^2 - 4y^2 - xy)\},$$

where C is a constant.

Find :

(i) E(X) and V(X)

(ii) E(Y) and V(Y)

(iii) Correlation between X and Y

Also identify the distributions of X and Y. *15*

(c) (i) एक कम्पनी में एक लेखा प्रबंधक ग्राहकों के 'साख श्रेणीयन' के आधार पर 75% ग्राहकों का वर्गीकरण 'अच्छे ऋण' की श्रेणी में करता है और बाकी ग्राहकों को 'जोखिमी ऋण' की श्रेणी में डालता है । 'जोखिमी' ऋण वाले वर्ग के ग्राहक औसतन 50% बार मियाद (अवधि) समय बीतने देते हैं जबकि 'अच्छे' ऋण वाले वर्ग के ग्राहक केवल 10% समय ही अपने ऋण की मियाद (अवधि) बीतने देते हैं । 'जोखिमी ऋण' श्रेणी में आने वाले ग्राहकों के पास अतिकालदेय लेखाओं का कितना प्रतिशत होता है ?

An accounts manager in a company classifies 75% of customers as 'good credit' and the rest as 'risky credit' depending on their 'credit rating'. Customers in the 'risky' category allow their accounts to go overdue 50% of the time on average, whereas those in the 'good' category allow their accounts to become overdue only 10% of the time. What percentage of overdue accounts are held by customers in the 'risky credit' category ?

(ii) मान लीजिए कि दो शतरंज खिलाड़ियों ने खेल खेले और यह देखा गया कि खिलाड़ी I के जीतने की प्रायिकता 0·30 है, खिलाड़ी II के जीतने की प्रायिकता 0·45 है, और खेल के बराबर होने की प्रायिकता 0·25 है । यदि इन दो शतरंज खिलाड़ियों ने दस खेल खेले हों, तो प्रायिकता ज्ञात कीजिए कि 6 खेल खिलाड़ी I जीतता है, 2 खेल खिलाड़ी II जीतता है, और शेष दो खेल बिना हार-जीत के होते हैं ।

Suppose that two chess players played games and it was seen that the probability that Player I would win is 0·30, the probability that Player II would win is 0·45, and the probability that the game would end in a draw is 0·25. If these two chess players played 10 games, find the probability that Player I wins 6 games, Player II wins 2 games, and the remaining 2 games would be drawn. *8+7=15*

Q4. (a) मान लीजिए कि $X_1, \ldots, X_n$ स्वतन्त्र और सर्वसम बंटन वाले बीटा $(\mu, 1)$ यादृच्छिक चर हैं और $Y_1, \ldots, Y_m$ स्वतन्त्र और सर्वसम बंटन वाले बीटा $(\theta, 1)$ यादृच्छिक चर हैं । X's और Y's स्वतन्त्र हैं ।

(i) $H_0 : \theta = \mu$ का परीक्षण $H_1 : \theta \neq \mu$ के विरुद्ध करने के लिए संभाविता अनुपात परीक्षण (एल.आर.टी.) ज्ञात कीजिए ।

(ii) दिखाइए कि भाग (i) में परीक्षण

$$T = \frac{\sum_{i=1}^{n} \log X_i}{\sum_{i=1}^{n} \log X_i + \sum_{j=1}^{m} \log Y_j}$$

प्रतिदर्शज पर आधारित हो सकता है ।

(iii) जब H_0 सत्य हो, तब T का बंटन ज्ञात कीजिए और फिर दिखाइए कि α आकार का परीक्षण कैसे प्राप्त किया जा सकता है ।

Suppose $X_1, \ldots, X_n$ are i.i.d beta $(\mu, 1)$ and $Y_1, \ldots, Y_m$ are i.i.d beta $(\theta, 1)$. X's are independent of Y's.

(i) Find a Likelihood Ratio Test (LRT) of $H_0 : \theta = \mu$ vs. $H_1 : \theta \neq \mu$.

(ii) Show that the test in part (i) can be based on the statistic

$$T = \frac{\sum_{i=1}^{n} \log X_i}{\sum_{i=1}^{n} \log X_i + \sum_{j=1}^{m} \log Y_j}.$$

(iii) Find the distribution of T, when H_0 is true and then show how to get a test of size α. *20*

(b) एक यादृच्छिक चर X का प्रायिकता फलन निम्नलिखित है :

X = x का मान	– 2	– 1	0	1	2	3
f(x)	0·15	k	0·25	2k	0·35	k

(i) k का मान ज्ञात कीजिए ।

(ii) बंटन फलन F(x) प्राप्त कीजिए और उसका ग्राफ़ बनाइए ।

A random variable X has the following probability function :

Value of X = x	-2	-1	0	1	2	3
f(x)	0·15	k	0·25	2k	0·35	k

(i) Find the value of k.

(ii) Obtain the distribution function F(x) and draw its graph. *5+10=15*

(c) वैडल सील मछलियाँ ऐन्टार्कटिक में रहती हैं और बर्फीले पानी में लम्बी और गहरी डुबकियों के दौरान मछलियों का भोजन करती हैं । सील मछलियाँ इन भरण डुबकियों से लाभ उठाती हैं, परन्तु इस लाभ की एक उपापचयी लागत है । अनुसंधानकर्ताओं का एक दल यह जानना चाहता था कि क्या भोजन करने में नियमित डुबकी से ऊपर ऊर्जा का इस्तेमाल होता है । उन्होंने दस भरण डुबकियों की उपापचयी लागत को मापा और प्रत्येक के लिए, बिना भरण डुबकियों के लिए भी मापा जो कि उसी जीव के लिए और समान समय तक चली । आँकड़ों का विवरण नीचे दिया गया है :

जीव	बिना भरण डुबकी के बाद ऑक्सीजन खपत ($mL\ O_2\ kg^{-1}$)	भरण डुबकी के बाद ऑक्सीजन खपत ($mL\ O_2\ kg^{-1}$)
1	42	71
2	52	77
3	60	83
4	67	96
5	82	107
6	82	113
7	81	121
8	81	80
9	96	128
10	104	143

जीव संख्याओं के बारे में अभिगृहीतों का उल्लेख कीजिए और परिकल्पना, कि भरण और बिना भरण डुबकियों की ऊर्जा लागत $\alpha = 0{\cdot}05$ स्तर पर बराबर है, का उपयुक्त अप्राचलिक परीक्षण कीजिए ।

Weddell seals live in the Antarctic and feed on fish during long, deep dives in freezing water. The seals benefit from these feeding dives, but the food they gain comes at a metabolic cost. A set of researchers wanted to know whether feeding per se was also energetically expensive, over and above the exertion of a regular dive. They measured the metabolic cost of 10 feeding dives and for each of these also measured a non-feeding dive by the same animal that lasted the same amount of time. The data are given below :

Individual	Oxygen consumption after non-feeding dive ($mL\ O_2\ kg^{-1}$)	Oxygen consumption after feeding dive ($mL\ O_2\ kg^{-1}$)
1	42	71
2	52	77
3	60	83
4	67	96
5	82	107
6	82	113
7	81	121
8	81	80
9	96	128
10	104	143

Mention assumptions about the populations and carry out an appropriate non-parametric test of hypothesis that the energetic cost is the same for feeding and non-feeding dives at level $\alpha = 0{\cdot}05$. *15*

खण्ड B

SECTION B

Q5. (a) दो प्राचलों θ और ϕ का आकलन करने के लिए, अनेक स्वतन्त्र मापन लिए जाते हैं, प्रत्येक मापन की त्रुटि का माध्य शून्य और प्रसरण σ^2 है । मान लीजिए कि

(i) n प्रेक्षणों का माध्य θ है,

(ii) m प्रेक्षणों का माध्य $(\theta - \phi)$ है, और

(iii) m प्रेक्षणों का माध्य $(\phi - \theta)$ है ।

इसे एक मानक रैखिक मॉडल के रूप में व्यवस्थित कीजिए । θ और ϕ के न्यूनतम वर्ग आकल क्या हैं ? $\text{var}(\hat{\phi})$, $\text{cov}(\hat{\theta}, \hat{\phi})$ और $\text{var}(\hat{\phi} - \hat{\theta})$ ज्ञात कीजिए ।

In order to estimate two parameters θ and ϕ, a number of independent measurements are taken, each having errors with mean zero and variance σ^2. Suppose there are

(i) n observations with mean θ,

(ii) m observations with mean $(\theta - \phi)$, and

(iii) m observations with mean $(\phi - \theta)$.

Set up this in the form of a standard linear model. What are the least square estimates of θ and ϕ ? Find $\text{var}(\hat{\phi})$, $\text{cov}(\hat{\theta}, \hat{\phi})$ and $\text{var}(\hat{\phi} - \hat{\theta})$. *10*

(b) एक मानक रैखिक मॉडल पर विचार कीजिए, जिसका अभिकल्प आव्यूह

$$X = \begin{pmatrix} -2 & -2 \\ -1 & -1 \\ 0 & 0 \\ 1 & 1 \\ 2 & 2 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}$$

है और β_1 तथा β_2 अज्ञात गुणांक हैं । कौन-से रैखिक संयोजनों $\hat{\psi} = a_1\hat{\beta}_1 + a_2\hat{\beta}_2$ का सबसे अधिक और सबसे कम प्रसरण होगा बशर्ते कि $a_1^2 + a_2^2 = 1$ है ?

Consider a standard linear model with design matrix

$$X = \begin{pmatrix} -2 & -2 \\ -1 & -1 \\ 0 & 0 \\ 1 & 1 \\ 2 & 2 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}$$

and unknown coefficients β_1 and β_2. Which linear combinations $\hat{\psi} = a_1\hat{\beta}_1 + a_2\hat{\beta}_2$ have largest and smallest variance subject to $a_1^2 + a_2^2 = 1$? *10*

(c) प्रतिस्थापन के साथ सरल यादृच्छिक प्रतिचयन (एस.आर.एस.डब्ल्यू.आर.) में (प्रचलित संकेत चिह्नों के साथ), दिखाइए कि

(i) $E(\bar{y}) = \bar{Y}$, और

(ii) $\text{Var}(\bar{y}) = \dfrac{N-1}{Nn} S_y^2$.

In Simple Random Sampling with Replacement (SRSWR) (with usual notation), show that

(i) $E(\bar{y}) = \bar{Y}$, and

(ii) $\text{Var}(\bar{y}) = \dfrac{N-1}{Nn} S_y^2$. *10*

(d) प्रचलित संकेतन के साथ, दर्शाइए कि अनुपात आकलक $\hat{R}$, R का श्रेष्ठतम रैखिक निष्पक्ष (अनभिनत) आकलक है और उल्लेख कीजिए कि उन प्रतिबंधों का जिनके अधीन यह गुणधर्म सही होता है ।

With usual notation, show that the ratio estimator $\hat{R}$ is the best linear unbiased estimator of R mentioning the conditions under which the property holds. *10*

(e) डन्कन के मल्टीपल रेंज परीक्षण की व्याख्या कीजिए ।

Explain Duncan's Multiple Range Test. *10*

Q6. (a) दर्शाइए कि (प्रचलित संकेतन के साथ), प्रतिगामी आकलक $\bar{y}_r$ का सन्निकट प्रसरण $\mathrm{Var}(\bar{y}_r) = \left(\frac{1-f}{n}\right) S_y^2 (1-\rho^2)$ है ।

Show that the approximate variance of the regression estimator $\bar{y}_r$ is $\mathrm{Var}(\bar{y}_r) = \left(\frac{1-f}{n}\right) S_y^2 (1-\rho^2)$ (with usual notation). *15*

(b) नियत प्रभाव मॉडल, यादृच्छिक प्रभाव मॉडल और मिश्रित प्रभाव मॉडल के बीच विभेदन कीजिए ।

Distinguish between fixed effect model, random effect model and mixed effect model. *15*

(c) विचार कीजिए दो द्विचर प्रसामान्य समष्टियों से आकार 10 के स्वतन्त्र और सर्वसम बंटन वाले, दो स्वतन्त्र प्रतिदर्श । परिणामों का सार नीचे दिया गया है ।

$$\bar{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \ \bar{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ S_1 = \begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix}, \ S_2 = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix},$$

जहाँ $S_j, j = 1, 2,$ प्रतिदर्श सहप्रसरण आव्यूह हैं

और $\bar{x}_j, j = 1, 2,$ प्रतिदर्श माध्य हैं ।

$\alpha = 0{\cdot}05$ स्तर पर निम्नलिखित परिकल्पनाओं के लिए उपयुक्त परीक्षण दीजिए और अपना निष्कर्ष निकालिए ।

(i) $H_0 : \boldsymbol{\mu}_1 = \begin{pmatrix} 2{\cdot}5 \\ 1 \end{pmatrix}$

(ii) $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$

(iii) $H_0 : \mu_{11} + 2\mu_{12} = 4$, जहाँ $\boldsymbol{\mu}_1 = (\mu_{11}, \mu_{12})'$

(आप निम्नलिखित मानों का उपयोग कर सकते हैं :

$F_{0{\cdot}95,\,2,\,17} = 3{\cdot}5915$, $F_{0{\cdot}95,\,2,\,18} = 3{\cdot}55$, $F_{0{\cdot}95,\,1,\,19} = 4{\cdot}38$)

Consider two independent i.i.d samples, each of size 10, from two bivariate normal populations. The results are summarised below.

$$\bar{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \bar{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, S_1 = \begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix}, S_2 = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix},$$

where S_j, j = 1, 2 are sample covariance matrices

and $\bar{x}_j$, j = 1, 2 are sample means.

Provide suitable tests for the following hypotheses at level $\alpha = 0{\cdot}05$ and draw your conclusion.

(i) $H_0 : \boldsymbol{\mu}_1 = \begin{pmatrix} 2{\cdot}5 \\ 1 \end{pmatrix}$

(ii) $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$

(iii) $H_0 : \mu_{11} + 2\mu_{12} = 4$, where $\boldsymbol{\mu}_1 = (\mu_{11}, \mu_{12})'$

(You can use the following values :

$F_{0{\cdot}95,\,2,\,17} = 3{\cdot}5915$, $F_{0{\cdot}95,\,2,\,18} = 3{\cdot}55$, $F_{0{\cdot}95,\,1,\,19} = 4{\cdot}38$) *20*

Q7. (a) मान लीजिए कि

$$\mathbf{X} \sim N_3\left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 11 & -6 & 2 \\ -6 & 10 & -4 \\ 2 & -4 & 6 \end{pmatrix}\right)$$

$Z_1 = X_2 - X_3$

$Z_2 = X_2 + X_3$

$Z_3 \mid Z_1, Z_2 \sim N(Z_1 + Z_2, 10)$

व्युत्पन्न कीजिए :

(i) Z_1, Z_2 और Z_3 का संयुक्त बंटन

(ii) X_3 का सप्रतिबन्ध माध्य जब X_1 और X_2 दिए गए हों

(iii) $\rho^2_{3{\cdot}12}$

Let

$$\mathbf{X} \sim N_3\left(\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}11 & -6 & 2\\ -6 & 10 & -4\\ 2 & -4 & 6\end{pmatrix}\right)$$

$Z_1 = X_2 - X_3$

$Z_2 = X_2 + X_3$

$Z_3 \mid Z_1, Z_2 \sim N(Z_1 + Z_2, 10)$

Derive :

(i) The joint distribution of Z_1, Z_2 and Z_3

(ii) The conditional mean of X_3 given X_1 and X_2

(iii) $\rho^2_{3\cdot 12}$ 20

(b) संकेतनों का स्पष्टतया उल्लेख करते हुए, दिखाइए कि हॉरविट्ज़-थॉम्पसन (Horvitz-Thompson) आकलक,

$$\hat{Y}_{HT} = \sum_{i=1}^{n} \frac{y_i}{\pi_i}$$

समष्टि (जनसंख्या) योग Y का एक अनभिनत आकलक है और

$$\text{Var}(\hat{Y}_{HT}) = \sum_{i=1}^{N} \left(\frac{1-\pi_i}{\pi_i}\right) y_i^2 + 2\sum_{i=1}^{N} \sum_{j>i}^{N} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j}\right) y_i y_j$$

प्रसरण है ।

Mentioning clearly about the notations, show that Horvitz-Thompson estimator of the population total

$$\hat{Y}_{HT} = \sum_{i=1}^{n} \frac{y_i}{\pi_i}$$

is an unbiased estimator of Y and variance is

$$\text{Var}(\hat{Y}_{HT}) = \sum_{i=1}^{N} \left(\frac{1-\pi_i}{\pi_i}\right) y_i^2 + 2\sum_{i=1}^{N} \sum_{j>i}^{N} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j}\right) y_i y_j.$$ 15

(c) संतुलित अपूर्ण खंडक अभिकल्पना (बी.आई.बी.डी.) की परिभाषा दीजिए । बी.आई.बी.डी. के प्राचलिक सम्बन्धों का कथन कीजिए । असमता $b \geq v + r - 1$ को भी सिद्ध कीजिए (सामान्य संकेतन के साथ) ।

Define BIBD. State the parametric relationships of the BIBD with usual notation. Also prove the inequality $b \geq v + r - 1$ (with usual notation). 15

Q8. (a) रैखिक प्रवृत्ति वाली एक समष्टि (जनसंख्या) और N = nk (N = समष्टि का आकार, n = प्रतिदर्श का आकार, k = प्रतिचयन का अंतराल) के लिए दर्शाइए कि

$$V_{st} : V_{sy} : V_{ran} :: \frac{1}{n} : 1 : \frac{nk+1}{k+1},$$

जहाँ V_{st}, V_{sy} और V_{ran} क्रमशः स्तरित प्रतिचयन, क्रमबद्ध प्रतिचयन और प्रतिस्थापन-रहित सरल यादृच्छिक प्रतिचयन के लिए, समष्टि (जनसंख्या) माध्य के प्रतिदर्श अनभिनत आकलकों के प्रतिचयन प्रसरण हैं ।

For a population with linear trend and N = nk (N = population size, n = sample size, k = sampling interval) show that

$$V_{st} : V_{sy} : V_{ran} :: \frac{1}{n} : 1 : \frac{nk+1}{k+1},$$

where V_{st}, V_{sy} and V_{ran} are respectively for sampling variances of the sample unbiased estimator of the population mean for stratified sampling, systematic sampling and SRSWOR. *15*

(b) n परिवारों में पहले पैदा हुए दो पुत्रों पर विचार कीजिए ।

मान लीजिए कि

X_1 : पहले पुत्र की सिर लम्बाई

X_2 : पहले पुत्र की सिर चौड़ाई

X_3 : दूसरे पुत्र की सिर लम्बाई

X_4 : दूसरे पुत्र की सिर चौड़ाई

मापनों द्वारा दिया गया प्रतिदर्श प्रसरण सहप्रसरण आव्यूह है

$$S = \begin{pmatrix} 91{\cdot}481 & 50{\cdot}753 & 66{\cdot}875 & 44{\cdot}267 \\ 50{\cdot}753 & 52{\cdot}186 & 49{\cdot}259 & 33{\cdot}651 \\ 66{\cdot}875 & 49{\cdot}259 & 96{\cdot}775 & 54{\cdot}278 \\ 44{\cdot}267 & 33{\cdot}651 & 54{\cdot}278 & 43{\cdot}222 \end{pmatrix}$$

जो कि सहसम्बन्ध आव्यूह

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$$ उत्पन्न करता है,

जहाँ $R_{11} = \begin{pmatrix} 1 & 0{\cdot}7346 \\ 0{\cdot}7346 & 1 \end{pmatrix}$, $R_{22} = \begin{pmatrix} 1 & 0{\cdot}8392 \\ 0{\cdot}8392 & 1 \end{pmatrix}$

$$R_{12} = R_{21}' = \begin{pmatrix} 0{\cdot}7107 & 0{\cdot}7040 \\ 0{\cdot}6931 & 0{\cdot}7085 \end{pmatrix}.$$

विहित सहसंबंधों और अतएव प्रथम विहित चरों को प्राप्त कीजिए तथा अपने परिणाम की व्याख्या कीजिए ।

Consider two sons born first in n families.

Let

X_1 : head length of first son

X_2 : head breadth of first son

X_3 : head length of second son

X_4 : head breadth of second son

The measurements provide the sample variance covariance matrix

$$S = \begin{pmatrix} 91{\cdot}481 & 50{\cdot}753 & 66{\cdot}875 & 44{\cdot}267 \\ 50{\cdot}753 & 52{\cdot}186 & 49{\cdot}259 & 33{\cdot}651 \\ 66{\cdot}875 & 49{\cdot}259 & 96{\cdot}775 & 54{\cdot}278 \\ 44{\cdot}267 & 33{\cdot}651 & 54{\cdot}278 & 43{\cdot}222 \end{pmatrix}$$

that leads to the correlation matrix

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},$$

where $R_{11} = \begin{pmatrix} 1 & 0{\cdot}7346 \\ 0{\cdot}7346 & 1 \end{pmatrix}$, $R_{22} = \begin{pmatrix} 1 & 0{\cdot}8392 \\ 0{\cdot}8392 & 1 \end{pmatrix}$

$$R_{12} = R_{21}' = \begin{pmatrix} 0{\cdot}7107 & 0{\cdot}7040 \\ 0{\cdot}6931 & 0{\cdot}7085 \end{pmatrix}.$$

Obtain the canonical correlations and hence the first canonical variables and interpret your result. *15*

(c) एक यादृच्छीकृत खंडक अभिकल्पना (RBD) में धान की नौ प्रकार की किस्मों की प्रति पौधा अनाज की फ़सल (ग्राम), नीचे प्रस्तुत की गई है :

किस्म	प्रतिकृति		
	1	2	3
1	21·0	20·0	19·5
2	19	18	18·5
3	18·5	18	18·9
4	27·5	*	27
5	31	32·5	32·6
6	31·5	30·5	32
7	25·3	25	26·6
8	39	40	38·5
9	39	38·5	40

प्रयोगात्मक आँकड़ों का विश्लेषण कीजिए ।
(तालिका मान $F(2, 25) = 3{\cdot}39$ और $F(8, 25) = 2{\cdot}34$, 5% सार्थकता स्तर पर)

Grain yields per plant (grams) of paddy of nine varieties in a Randomised Block Design are presented below :

Variety	Replication		
	1	2	3
1	21·0	20·0	19·5
2	19	18	18·5
3	18·5	18	18·9
4	27·5	*	27
5	31	32·5	32·6
6	31·5	30·5	32
7	25·3	25	26·6
8	39	40	38·5
9	39	38·5	40

Analyse the experimental data.
(Table value $F(2, 25) = 3{\cdot}39$ and $F(8, 25) = 2{\cdot}34$, at 5% level of significance) *20*