CHAHAL ACADEMY
LEARN TODAY! LEAD TOMORROW!
96259-93995

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 । 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 (क्यू.सी.ए.) 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 (QCA) Booklet must be clearly struck off.

CS (Main) Exam, 2021

खण्ड A

SECTION A

Q1. (a) सर्जिकल मास्क बनाने वाली एक उत्पादन इकाई की मास्क की गुणवत्ता जाँचने में रुचि है । दोषपूर्ण मास्क बनाने की प्रायिकता, 'p', के आकलन के लिए n मास्क के एक यादृच्छिक प्रतिदर्श का निरीक्षण किया गया । प्रतिदर्श कितना बड़ा होना चाहिए ताकि 0·95 प्रायिकता के साथ p के आकलक का परिसर $p \pm 0{\cdot}1$ हो ?

A production unit manufacturing surgical masks is concerned about the quality of their masks. A random sample of n masks are inspected to estimate 'p', the probability of manufacturing a defective mask. How large a sample is required so that the estimate of p lies in the range $p \pm 0{\cdot}1$ with probability 0·95 ? *10*

(b) एक बीमा कम्पनी ने 150 पॉलिसी-धारकों के प्रतिदर्श का अध्ययन किया । पॉलिसी के तीन वर्ग हैं : वाहन, गृह और चिकित्सा । पॉलिसी-धारकों द्वारा गृहित पॉलिसियों के संबंध में निम्न परिणाम प्राप्त हुए :

(i) 30 के पास केवल गृह बीमा है

(ii) 10 के पास केवल चिकित्सा बीमा है

(iii) 98 के पास वाहन बीमा है, लेकिन सभी तीन प्रकार के बीमा नहीं हैं

(iv) 27 के पास चिकित्सा बीमा है, लेकिन सभी तीन प्रकार के बीमा नहीं हैं

(v) 13 के पास वाहन और चिकित्सा बीमा है

यदि यह दिया हुआ है कि पॉलिसी-धारक के पास चिकित्सा बीमा है, तो उसके पास गृह बीमा होने की प्रायिकता परिकलित कीजिए ।

An insurance company studies a sample of 150 policy-holders. There are three categories of policies : auto, home and medical. The following results are obtained about the policies held by the policy-holders :

(i) 30 have only home insurance

(ii) 10 have only medical insurance

(iii) 98 have auto insurance, but not all three types of insurance

(iv) 27 have medical insurance, but not all three types of insurance

(v) 13 have auto and medical insurance

Given that a policy-holder has medical insurance, calculate the probability that he has home insurance. *10*

(c) माना X और Y स्वतंत्र और सर्वसम बंटित चरघातांकी यादृच्छिक चर हैं जिनका माध्य $\lambda > 0$ है ।

परिभाषित है

$$Z = \begin{cases} 1, & \text{यदि} \quad X < Y \\ 0, & \text{यदि} \quad X \geq Y \end{cases}$$

ज्ञात कीजिए :

$$E[X \mid Z = 1] + E[X \mid Z = 0]$$

Let X and Y be independent and identically distributed exponential random variables with mean $\lambda > 0$.

Define

$$Z = \begin{cases} 1, & \text{if} \quad X < Y \\ 0, & \text{if} \quad X \geq Y \end{cases}$$

Find $E[X \mid Z = 1] + E[X \mid Z = 0]$. 10

(d) माना $X_1, X_2, \ldots, X_n$

$$f(x, \theta) = \frac{\log(\theta)}{\theta - 1} \theta^x; \; 0 < x < 1, \; \theta > 1$$

से लिया गया कोई यादृच्छिक प्रतिदर्श है । क्या θ का एक फलन, माना $g(\theta)$, के लिए कोई अनभिनत आकलक है जिसका प्रसरण सी.-आर. (C-R) निम्न परिबंध प्राप्त करता हो ? यदि हाँ, तो ज्ञात कीजिए । यदि नहीं, तो दर्शाइए क्यों नहीं ।

Let $X_1, X_2, \ldots, X_n$ be a random sample from

$$f(x, \theta) = \frac{\log(\theta)}{\theta - 1} \theta^x; \; 0 < x < 1, \; \theta > 1$$

Is there a function of θ, say $g(\theta)$, for which there exists an unbiased estimator whose variance attains the C-R lower bound ? If yes, find it. If not, show why not. 10

(e) माना

$$f(x, \theta) = \frac{\theta}{\pi} \frac{1}{\theta^2 + x^2}; \; -\infty < x < \infty, \; \theta > 0$$

कौशी का प्रायिकता घनत्व फलन है ।

(i) दर्शाइए कि इस बंटन संवर्ग के लिए एकदिष्ट संभाविता अनुपात (एम.एल.आर.) नहीं है ।

(ii) यदि X, $f(x, \theta)$ से लिया गया एक प्रेक्षण है, तो दर्शाइए कि $|X|$, θ का पर्याप्त प्रतिदर्शज है और इसलिए $|X|$ के बंटन के लिए एम.एल.आर. है ।

Let f(x, θ) be the Cauchy pdf

$$f(x, \theta) = \frac{\theta}{\pi} \frac{1}{\theta^2 + x^2}; \; -\infty < x < \infty, \; \theta > 0$$

(i) Show that this family does not have Monotone Likelihood Ratio (MLR).

(ii) If X is one observation from f(x, θ), show that |X| is sufficient for θ and hence the distribution of |X| does have an MLR. *5+5*

Q2. (a) माना $Y_1, Y_2, Y_3, \ldots$ स्वतंत्र और सर्वसम प्वासों यादृच्छिक चर हैं जिनका प्राचल 1 है । केन्द्रीय सीमा प्रमेय का उपयोग करते हुए स्थापित कीजिए

$$n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n,$$

जबकि धनात्मक पूर्णांक n बृहत् है ।

Let $Y_1, Y_2, Y_3, \ldots$ be independent and identical Poisson random variables with parameter 1. Use central limit theorem to establish

$$n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

for large value of positive integer n. *20*

(b) माना $X_1, X_2, \ldots, X_n$ ऐसा यादृच्छिक प्रतिदर्श है जिसका बंटन $\log X_i \sim N(\theta, \theta), \theta > 0$ अज्ञात है । दर्शाइए कि संभाविता समीकरण का एक हल θ का एकमात्र MLE है । θ के MLE के लिए उपगामी बंटन प्राप्त कीजिए ।

Let $X_1, X_2, \ldots, X_n$ be a random sample such that $\log X_i \sim N(\theta, \theta)$ distribution with θ > 0 unknown. Show that one of the solutions of the likelihood equation is the unique MLE of θ. Obtain asymptotic distribution of MLE of θ. *15*

(c) (i) फलन $\phi(t)$ के अभिलक्षण फलन होने के लिए पर्याप्त प्रतिबंधों को लिखिए ।

(ii) जाँच कीजिए कि क्या निम्न फलन अभिलक्षण फलन हैं :

1. e^{-t^4}
2. $[1 + |t|]^{-1}$

अपने उत्तर को तर्कसंगत सिद्ध कीजिए ।

(i) State the sufficient conditions for a function $\phi(t)$ to be a characteristic function.

(ii) Investigate if the following functions are characteristic functions :

1. e^{-t^4}
2. $[1 + |t|]^{-1}$

Justify your answer. *5+10*

Q3. (a) माना X और Y चरघातांकी बंटन से लिए गए दो स्वतंत्र यादृच्छिक चर हैं जिनका माध्य क्रमशः $\frac{1}{\lambda}$ और $\frac{1}{\mu}$, $\lambda > 0, \mu > 0$ है । माना $(X_1, X_2, ..., X_n)$ और $(Y_1, Y_2, ..., Y_n)$ क्रमशः X और Y से लिए गए प्रेक्षणों के अनुक्रम हैं ।

एक यादृच्छिक चर U_i इस प्रकार से परिभाषित है

$$U_i = \begin{cases} 1, & \text{यदि} \quad X_i \geq Y_i, \ i = 1, 2, ..., n \\ 0, & \text{अन्यथा} \end{cases}$$

U_i पर आधारित $H : \lambda = \mu$ विरुद्ध $K : \lambda = 2\mu$ के परीक्षण के लिए वॉल्ड SPRT विधि की रचना कीजिए जिसकी शक्ति (α, β) है ।

Let X and Y be two independent random variables following exponential distribution with mean $\frac{1}{\lambda}$ and $\frac{1}{\mu}$ respectively, $\lambda > 0, \mu > 0$. Suppose that $(X_1, X_2, ..., X_n)$ and $(Y_1, Y_2, ..., Y_n)$ are sequences of observations on X and Y respectively.

A random variable U_i is defined as

$$U_i = \begin{cases} 1, & \text{if} \quad X_i \geq Y_i, \ i = 1, 2, ..., n \\ 0, & \text{otherwise} \end{cases}$$

Construct Wald's SPRT procedure based on U_i's for testing $H : \lambda = \mu$ versus $K : \lambda = 2\mu$ with strength (α, β). *20*

(b) माना Y_i, $i \geq 1$, स्वतंत्र और सर्वसम $U(-1, 1)$ यादृच्छिक चर हैं । ज्ञात कीजिए कि क्या निम्न अनुक्रम प्रायिकता में अभिसरित हैं :

(i) $\left\{\frac{Y_i}{i}\right\}$

(ii) $\left\{(Y_i)^i\right\}$

Let Y_i, $i \geq 1$ be independent and identical $U(-1, 1)$ random variables. Determine if the following sequences converge in probability : *5+10*

(i) $\left\{\frac{Y_i}{i}\right\}$

(ii) $\left\{(Y_i)^i\right\}$

(c) माना $X_1, X_2, ..., X_n$ एकसमान बंटन $U(-\theta, \theta), \theta > 0$ से लिया गया एक यादृच्छिक प्रतिदर्श है । θ का पूर्ण पर्याप्त प्रतिदर्शज ज्ञात कीजिए । इससे θ का सर्वोत्तम अनभिनत आकलक प्राप्त कीजिए ।

Let $X_1, X_2, ..., X_n$ be a random sample from uniform distribution $U(-\theta, \theta), \theta > 0$. Find the complete sufficient statistic for θ. Hence, obtain the best unbiased estimator of θ. *15*

Q4. (a) माना $X_1, X_2, ..., X_n$ प्वासों बंटन, जिसका माध्य $\lambda > 0$, से लिया गया एक यादृच्छिक प्रतिदर्श है । एक प्रतिदर्शज परिभाषित है

$$W = \left(1 - \frac{1}{n}\right)^T, \quad T = \sum_{i=1}^{n} X_i$$

(i) दर्शाइए कि T पूर्ण पर्याप्त प्रतिदर्शज है ।

(ii) दर्शाइए कि T, $e^{-\lambda}$ के लिए अनभिनत है ।

(iii) भले ही T, UMVUE (यू.एम.वी.यू.ई.) है, दर्शाइए कि यह $g(\lambda) = e^{-\lambda}$ के लिए CRLB (सी.आर.एल.बी.) प्राप्त नहीं करता है ।

Let $X_1, X_2, ..., X_n$ be a random sample from Poisson distribution with mean $\lambda > 0$. Define a statistic

$$W = \left(1 - \frac{1}{n}\right)^T, \quad T = \sum_{i=1}^{n} X_i$$

(i) Show that T is complete sufficient statistic.

(ii) Show that T is unbiased for $e^{-\lambda}$.

(iii) Show that even though T is UMVUE, it does not attain the CRLB for $g(\lambda) = e^{-\lambda}$. *20*

(b) माना

$$f(x, y) = \frac{e^{\frac{-yx^2}{2}} y^{3/2} e^{-y}}{\sqrt{2\pi}}, -\infty < x < \infty, y > 0.$$

(i) Y का उपान्त बंटन और Y के दिए होने पर X का सप्रतिबंध बंटन प्राप्त कीजिए ।

(ii) $E(Y), V(Y), E(X|Y), V(X|Y)$ ज्ञात कीजिए ।

(iii) (ii) का उपयोग करते हुए $E(X), V(X)$ ज्ञात कीजिए ।

Let $f(x, y) = \dfrac{e^{\frac{-yx^2}{2}} \; y^{3/2} \; e^{-y}}{\sqrt{2\pi}}, \; -\infty < x < \infty, \; y > 0.$

(i) Obtain the marginal distribution of Y and conditional distribution of X given Y.

(ii) Find $E(Y)$, $V(Y)$, $E(X \mid Y)$, $V(X \mid Y)$.

(iii) Use (ii) to find $E(X)$, $V(X)$. *5+5+5*

(c) तीन विभिन्न तरीकों से एक निश्चित औद्योगिक निरीक्षण प्रक्रिया द्वारा एक कम्पनी के प्रशिक्षार्थियों को यादृच्छया समूहों में नियत किया गया । प्रशिक्षण अवधि की समाप्ति पर निरीक्षण प्रदर्शन गुणवत्ता के लिए उनका परीक्षण किया गया । उनके स्कोर निम्न हैं :

रीति A :	80	83	79	85	90	68	
रीति B :	82	84	60	72	86	67	91
रीति C :	93	65	77	78	88		

उपयुक्त अप्राचलिक परीक्षण का उपयोग करते हुए, 0·05 सार्थकता स्तर पर निर्धारित कीजिए कि क्या तीनों रीतियाँ समान रूप से प्रभावी हैं ।

A company's trainees are randomly assigned to groups which are through a certain industrial inspection procedure by three different methods. At the end of the instructing period they are tested for inspection performance quality. The following are their scores :

Method A :	80	83	79	85	90	68	
Method B :	82	84	60	72	86	67	91
Method C :	93	65	77	78	88		

Using the appropriate non-parametric test, determine at 0·05 level of significance whether the three methods are equally effective. *15*

Chi-Square (χ^2) Distribution

Area to the Right of Critical Value

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	–	–	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997

खण्ड B
SECTION B

Q5. (a) एक साधारण रैखिक समाश्रयण निदर्श

$Y = \beta_0 + \beta_1 X_i + \varepsilon_i , i = 1, ..., n$ के लिए

(i) माने गए प्रतिबंधों को स्पष्टत: लिखते हुए, β_0 और β_1 के न्यूनतम वर्ग आकलकों को व्युत्पन्न कीजिए ।

(ii) $e_i = Y_i - \hat{Y}_i$ जहाँ $\hat{Y}_i$ आसंजित मान है, के लिए दर्शाइए कि

1. $$\sum_{i=1}^{n} e_i = 0$$

2. $$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \hat{Y}_i$$

3. $$\sum_{i=1}^{n} X_i e_i = 0$$

4. $$\sum_{i=1}^{n} \hat{Y}_i e_i = 0$$

5. समाश्रयण रेखा $(\overline{X}, \overline{Y})$ से गुज़रती है ।

For a simple linear regression model

$Y = \beta_0 + \beta_1 X_i + \varepsilon_i , i = 1, ..., n$

(i) Derive the least square estimators of β_0 and β_1, clearly stating the conditions assumed.

(ii) For $e_i = Y_i - \hat{Y}_i$ where $\hat{Y}_i$ is the fitted value, show that

1. $$\sum_{i=1}^{n} e_i = 0$$

2. $$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \hat{Y}_i$$

3. $$\sum_{i=1}^{n} X_i e_i = 0$$

4. $$\sum_{i=1}^{n} \hat{Y}_i e_i = 0$$

5. The regression line passes through $(\overline{X}, \overline{Y})$. 5+5

(b) प्रचलित संकेतनों में, यदि v, b, r, k और λ किसी संतुलित अपूर्ण खंडक अभिकल्पना के प्राचल हैं, तो दर्शाइए कि :

(i) $b \geq r + 1 \geq \lambda + 2$

(ii) $v \leq b \leq \dfrac{r^2 - 1}{\lambda}$

In usual notations, if v, b, r, k and λ are the parameters of a Balanced Incomplete Block Design, then show that : 10

(i) $b \geq r + 1 \geq \lambda + 2$

(ii) $v \leq b \leq \dfrac{r^2 - 1}{\lambda}$

(c) एक बहु रैखिक समाश्रयण निदर्श जिसमें X_1 और X_2 दो प्राग्वक्ता चर हैं, के लिए दर्शाइए कि जब भी X_1 और X_2 असहसंबंधित होंगे, समाश्रयण निदर्श में X_2 को जोड़ने पर X_1 के समाश्रयण गुणांक का आकलक अपरिवर्तित रहेगा ।

For the multiple linear regression model with two predictor variables X_1 and X_2, show that the estimate of regression coefficient of X_1 is unchanged when X_2 is added to the regression model, whenever X_1 and X_2 are uncorrelated. 10

(d) प्रतिस्थापन रहित सरल यादृच्छिक प्रतिचयन द्वारा समष्टि की N इकाइयों से n आकार का एक प्रतिदर्श चुना गया । प्रतिस्थापन रहित सरल यादृच्छिक प्रतिचयन द्वारा n इकाइयों से n_1 इकाई का एक उप-प्रतिदर्श चुना गया । माना कि n_1 इकाइयों पर आधारित माध्य को $\bar{y}_1$ और $n_2 = n - n_1$ इकाइयों पर आधारित माध्य को $\bar{y}_2$ से व्यक्त किया गया । समष्टि माध्य $\bar{Y}_N$ का आकलक दिया गया है :

$$\hat{\bar{Y}}_N = w\,\bar{y}_1 + (1 - w)\,\bar{y}_2 ; 0 < w < 1$$

दर्शाइए कि $E(\hat{\bar{Y}}_N) = \bar{Y}_N$, और इसका प्रसरण प्राप्त कीजिए ।

A sample of size n is drawn from a population having N units by simple random sampling without replacement. A sub-sample of n_1 units is drawn from the n units by simple random sampling without replacement. Let $\bar{y}_1$ denote the mean based on n_1 units and $\bar{y}_2$, the mean based on $n_2 = n - n_1$ units. Consider the estimator of the population mean $\bar{Y}_N$ given by :

$$\hat{\bar{Y}}_N = w\,\bar{y}_1 + (1 - w)\,\bar{y}_2 ; 0 < w < 1$$

Show that $E(\hat{\bar{Y}}_N) = \bar{Y}_N$, and obtain its variance. 10

(e) किसी अभिकल्पना की दक्षता कैसे मापी जाती है ? पूर्णत: यादृच्छिकीकृत अभिकल्पना पर यादृच्छिकीकृत खंडक अभिकल्पना की दक्षता को मापने का व्यंजक व्युत्पन्न कीजिए ।

How is the efficiency of a design measured ? Derive the expression to measure the efficiency of a Randomised Block Design over a Completely Randomised Design. *10*

Q6. (a) किसी बहु रैखिक समाश्रयण निदर्श जिसमें तीन सह-विचर X_1, X_2 और X_3 हैं, के लिए, माना r_{ij}, X_i और X_j में सहसंबंध गुणांक दर्शाता है । किन्हीं आँकड़ों के लिए, देखा गया कि $r_{12} = 0{\cdot}77$, $r_{23} = 0{\cdot}52$, $r_{13} = 0{\cdot}72$ है ।

(i) उपर्युक्त आँकड़ों की संगतता जाँचिए ।

(ii) यदि r_{13} अज्ञात हो, तो ऊपर दिए गए r_{12} और r_{23} के मानों से r_{13} की सीमाएँ प्राप्त कीजिए ।

For a multiple linear regression model with three covariates X_1, X_2 and X_3, let r_{ij} denote the correlation coefficient between X_i and X_j. For a data, it was found $r_{12} = 0{\cdot}77$, $r_{23} = 0{\cdot}52$, $r_{13} = 0{\cdot}72$.

(i) Check the consistency of the above data.

(ii) If r_{13} is unknown, obtain the limits within which r_{13} lies given the above values for r_{12} and r_{23}. *20*

(b) समान आकार वाले गुच्छों के गुच्छ प्रतिचयन में, समष्टि माध्य का अनभिनत आकलक प्राप्त कीजिए । इसका प्रतिचयन प्रसरण भी निम्न रूप में ज्ञात कीजिए :

$$V(\bar{\bar{y}}) = (1 - f)\,(NM - 1)\,S^2\,\frac{\{1 + (M - 1)\,\rho_{cl}\}}{M^2(N - 1)\,n}$$

जहाँ संकेतों के अपने सामान्य अर्थ हैं ।

In cluster sampling with equal size clusters, obtain the unbiased estimate of population mean. Also obtain its sampling variance as

$$V(\bar{\bar{y}}) = (1 - f)\,(NM - 1)\,S^2\,\frac{\{1 + (M - 1)\,\rho_{cl}\}}{M^2(N - 1)\,n},$$

where notations have their usual meanings. *15*

(c) माना $Z_{3\times1} = \begin{pmatrix} X_{1\times1} \\ Y_{2\times1} \end{pmatrix} \sim N_3\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 2 \end{pmatrix}\right).$

दर्शाइए कि $X_{1\times1}$ के प्रतिबंध पर, $Y_{2\times1}$ के दो घटक स्वतंत्र हैं लेकिन उपांतीय वे स्वतंत्र नहीं हैं ।

Let $Z_{3\times1} = \begin{pmatrix} X_{1\times1} \\ Y_{2\times1} \end{pmatrix} \sim N_3\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 2 \end{pmatrix}\right).$

Show that conditional on $X_{1\times1}$, the two components of $Y_{2\times1}$ are independent but marginally they are not. *15*

Q7. (a) (i) बहु-उपादानी प्रयोगों में संकरण क्या है ?

(ii) एक 2^6 बहु-उपादानी प्रयोग 2^3 आकार के खंडकों में संचालित किया गया । संकरित प्रभावों को लिखिए जिसमें कोई भी मुख्य उपादान या दो घटक अन्योन्यक्रिया संकरित न हों । संकरित होने वाले स्वतंत्र व व्यापकीकृत अन्योन्यक्रियाओं की सूची लिखिए, साथ ही केवल प्रमुख खंडक के अवयव लिखिए ।

(iii) 2^k खंडकों में 2^n बहु-उपादानी प्रयोग के लिए स्वातंत्र्य कोटियों का विभाजन दीजिए ।

(i) What is confounding in factorial experiments ?

(ii) A 2^6 factorial experiment is conducted in blocks of size 2^3. Write the confounded effects such that no main effect or two factor interaction are confounded. Give the list of independent and generalised interactions confounded along with the elements of key block only.

(iii) Give the break-up of degrees of freedom for a 2^n factorial experiment in 2^k blocks. *20*

(b) मुख्य घटक क्या हैं ? सदिश $X_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ और $X_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ के मुख्य घटकों के परिकलन का विवरण दीजिए । X_1 और X_2 को मुख्य घटकों के रूप में लिखिए ।

What are principal components ? Describe how to compute the principal components of the vectors $X_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $X_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. Give X_1 and X_2 in terms of the principal components. *15*

(c) समाश्रयण आकलक परिभाषित कीजिए । दर्शाइए अभिनति = – सहप्रसरण $(\overline{x}, b)$ । किन प्रतिबन्धों के अन्तर्गत अभिनति नगण्य होती है ? प्रथम घात के सन्निकट आकलक की त्रुटि वर्ग माध्य ज्ञात कीजिए । समाश्रयण आकलक की अनुपात आकलक के साथ तुलना कीजिए ।

Define Regression estimator. Show bias = – Cov $(\overline{x}, b)$. Under what conditions is bias negligible ? Find the mean square error of the estimator to first degree of approximation. Give comparison of Regression estimator with Ratio estimator. *15*

Q8. (a) (i) स्तरित प्रतिचयन में अनुकूलतम नियतन के अंतर्गत यदि एक या अधिक $n_i, N_i \ (i \geq 2)$ से ज्यादा बड़े हैं, तो आप विभिन्न स्तरों से इकाइयों का चयन किस प्रकार करेंगे ?

(ii) हिमाचल प्रदेश के किसी जिले में एक प्रतिदर्श सर्वेक्षण किया गया । राजस्व अभिलेखों द्वारा प्राप्त फलदार पेड़ों के क्षेत्रफल के आधार पर गाँवों के चार स्तर A, B, C और D बनाए गए । प्रत्येक स्तर से गाँवों का एक यादृच्छिक प्रतिदर्श चुना गया और प्रत्येक चुने गए गाँव से सेब के बगीचों की संख्या लिखी गई । आँकड़े नीचे दर्शाए गए हैं :

स्तर	गाँवों की कुल संख्या (N_i)	प्रतिदर्श में गाँवों की संख्या (n_i)	चुने गए गाँवों में बगीचों की संख्या
A (0 – 3 एकड़)	275	15	2, 5, 1, 9, 6, 7, 0, 4, 7, 0, 5, 0, 0, 3, 0
B (3 – 6 एकड़)	146	10	21, 11, 7, 5, 6, 19, 5, 24, 30, 24
C (6 – 15 एकड़)	93	12	3, 10, 4, 11, 38, 11, 4, 46, 4, 18, 1, 39
D (15 एकड़ और अधिक)	62	11	30, 42, 20, 38, 29, 22, 31, 28, 66, 14, 15

जिले में बगीचों की संख्या का आकलन कीजिए ।

(i) In stratified sampling under optimum allocation, how will you proceed to select units from different strata, if one or more n_i's happens to be greater than N_i ($i \geq 2$) ? 10

(ii) A sample survey was conducted in a certain district of Himachal Pradesh. Four strata A, B, C and D of villages were formed according to the acreage of fruit trees as obtained from revenue records. A random sample of villages was selected from each stratum and the number of apple orchards in each selected village was noted. The data are shown below :

Stratum	Total number of villages (N_i)	Number of villages in sample (n_i)	Number of orchards in the selected villages
A (0 – 3 acres)	275	15	2, 5, 1, 9, 6, 7, 0, 4, 7, 0, 5, 0, 0, 3, 0
B (3 – 6 acres)	146	10	21, 11, 7, 5, 6, 19, 5, 24, 30, 24
C (6 – 15 acres)	93	12	3, 10, 4, 11, 38, 11, 4, 46, 4, 18, 1, 39
D (15 acres and above)	62	11	30, 42, 20, 38, 29, 22, 31, 28, 66, 14, 15

Estimate the number of orchards in the district. 10

(b) (i) द्विघातीय बहुपद निदर्श जिसमें एक प्राग्वक्ता चर है, के लिए माने गए प्रतिबन्धों को स्पष्टत: लिखते हुए, न्यूनतम वर्ग प्रसामान्य समीकरण व्युत्पन्न कीजिए । आप इस निदर्श में प्राचलों की व्याख्या कैसे करेंगे ?

(ii) वर्णन कीजिए कि क्यों माध्य के परित: केन्द्रित प्राग्वक्ता चरों को संस्तुत किया जाता है । इस विषय में अनुक्रिया चर के आसंजित मानों पर टिप्पणी लिखिए । अपने दावे को सिद्ध कीजिए ।

(i) For a second order polynomial model with one predictor variable, derive the least squares normal equations clearly stating the conditions assumed. How will you interpret the parameters in this model ?

(ii) Describe why it is recommended to work with predictor variables centred around the mean. Comment on fitted values of the response variable in this case. Prove your claim. 7+8

(c) विभक्त-क्षेत्र अभिकल्पनाएँ क्या हैं ? आप इन अभिकल्पनाओं के उपयोग को कब संस्तुत करेंगे ? यदि e_1 और e_2 क्रमशः मुख्य क्षेत्र और उप-क्षेत्र त्रुटियाँ हैं, दोनों ही एकल उप-क्षेत्र इकाइयों में आकलित हैं, तो स्पष्ट कीजिए कि क्यों e_1, e_2 से अधिक बड़ा अनुमानित होता है ।

What are split-plot designs ? When do you recommend the use of such designs ? If e_1 and e_2 are the main plot and sub-plot errors respectively, both estimated in units of a single sub-plot, explain why e_1 is expected to be larger than e_2. *15*