

Find the following :

(i) Marginal pdf of  $x$  and  $y$

(ii) Find  $P(X > 0, Y < 1/2)$ . (7)

(b) Show that if  $Y = aX + b$  ( $a \neq 0$ ), then  $\text{Corr}(X, Y) = +1$  or  $-1$ . Under what conditions will  $\rho = +1$ ? (3)

(क) मान लीजिए कि  $X$  और  $Y$  दो संयुक्त रूप से सतत यादृच्छिक चर हैं, जिनका संयुक्त प्रायिकता घनत्व फलन निम्नलिखित है :

$$f_{xy}(x, y) = \begin{cases} \frac{1}{2}x^2 + \frac{2}{3}y & \text{for } -1 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

ज्ञात कीजिए

(i)  $X$  और  $Y$  का सीमांत प्रायिकता घनत्व फलन।

(ii)  $P(X > 0, Y < 1/2)$  का मान ज्ञात कीजिए।

(ख) यह दर्शाइए कि यदि  $Y = aX + b$  ( $a \neq 0$ ), तो  $\text{Corr}(X, Y) = +1$  या  $-1$  किन परिस्थितियों में  $\rho = +1$  होगा?

[This question paper contains 16 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2124 I

Unique Paper Code : 2272101103

Name of the Paper : Introductory Statistics for Economics

Name of the Course : B.A. (Hons) Economics

Semester : III

Duration : 3 Hours

Maximum Marks : 90

समय : 3 घण्टे

पूर्णांक : 90

### Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. All questions within each section are to be answered in a contiguous manner on the answer sheet. Start each question on a new page, and all sub-parts of a question should follow one after the other.
3. All intermediate calculations should be rounded off to 3 decimal places. The values provided in statistical tables should not be rounded off. All final calculations should be rounded off to two decimal places.

4. Simple non-programmable calculators are allowed.
5. Statistical tables are attached for your reference.
6. Answers may be written either in English or Hindi; but the same medium should be used throughout the paper.

### छात्रों के लिए निर्देश

1. इस प्रश्न-पत्र के मिलते ही ऊपर दिए गए निर्धारित स्थान पर अपना अनुक्रमांक लिखिए।
2. प्रत्येक खंड के सभी प्रश्नों के उत्तर उत्तर पुस्तिका पर एक साथ दिए जा हैं। प्रत्येक प्रश्न को एक नए पृष्ठ पर शुरू करें, और एक प्रश्न के सभी उप-भाग एक के बाद एक होने चाहिए।
3. सभी मध्यवर्ती गणनाओं को 3 दशमलव स्थानों तक पूर्णांकित किया जाना चाहिए। सांख्यिकीय तालिकाओं में दिए गए मानों को पूर्णांकित नहीं किया जाना चाहिए। सभी अंतिम गणनाओं को दो दशमलव स्थानों तक पूर्णांकित किया जाना चाहिए।
4. साधारण गैर-प्रोग्रामिय कैलकुलेटर की अनुमति है।
5. आपके संदर्भ के लिए सांख्यिकीय तालिकाएँ संलग्न हैं।
6. इस प्रश्न-पत्र का उत्तर अंग्रेजी या हिंदी किसी एक भाषा में दीजिए, लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए।

- (iv) Find the conditional distribution of Y given X=100
- (v) Compute the correlation coefficient  $\rho$  for X and Y. (2×5=10)

यादृच्छिक चर X & Y का संयुक्त प्रायिकता घनत्व फलन निम्नलिखित है :

$x \backslash y \rightarrow$	0	100	200
100	0.05	0.10	0.20
200	0.10	0.20	0.10
300	0.15	0.05	0.05

- (i) ज्ञात कीजिए  $P(Y=100|X=100)$ ।
  - (ii) क्या X और Y स्वतंत्र हैं?
  - (iii) सहसंबंध  $\text{Cov}(X, Y)$  की गणना कीजिए।
  - (iv)  $X=100$  के लिए Y का सशर्त वितरण ज्ञात कीजिए।
  - (v) X और Y के लिए सहसंबंध गुणांक  $\rho$  की गणना कीजिए।
10. (a) Let X and Y be two jointly continuous random variables with joint PDF

$$f_{xy}(x, y) = \begin{cases} \frac{1}{2}x^2 + \frac{2}{3}y & \text{for } -1 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- (ख) पौधों को पानी देने के लिए प्लास्टिक ट्यूब का अनुप्रस्थ क्षेत्र सामान्य रूप से  $\mu=12.5 \text{ mm}^2$  और  $\sigma=0.2 \text{ mm}^2$  के साथ वितरित किया जाता है। जब क्षेत्र  $12.0 \text{ mm}^2$  से कम या  $13.0 \text{ mm}^2$  से अधिक होता है, तो ट्यूब नली में ठीक से फिट नहीं होती है। यदि ट्यूबों को एक हजार के बक्से में भेजा जाता है, तो विक्रेता प्रति बॉक्स कितनी गलत आकार की ट्यूब मिलने की उम्मीद कर सकते हैं? (5)

#### SECTION IV (खण्ड IV)

All questions are compulsory.

सभी प्रश्न अनिवार्य हैं।

9. The joint PDF of random variables X&Y is given by

$\downarrow x \backslash y \rightarrow$	0	100	200
100	0.05	0.10	0.20
200	0.10	0.20	0.10
300	0.15	0.05	0.05

- Find  $P(Y=100|X=100)$ .
- Are X and Y independent?
- Compute  $\text{Cov}(X, Y)$

#### SECTION I (खण्ड I)

Do any two questions.

किन्ही दो प्रश्नों को हल कीजिए।

- Calculate 18% trimmed mean for the following data: 73, 80, 36, 75, 68, 82, 50, 48, 41, 60. Compare it with the sample mean and median. (5)
  - Identification of a spam message involves creating a list of words more likely to appear in spam than in normal messages. The list is called "spam list". Suppose that your database of 5000 messages contains 1700 that are spam. Among the spam messages, 1343 contain words from the "spam list". Of the 3300 normal messages, only 297 contain words from the "spam list". Obtain the probability that a message is spam given that the message contains words from the "spam list". (5)
- निम्नलिखित आँकड़ों के लिए 18% ट्रिम्ड माध्य की गणना कीजिए: 73, 80, 36, 75, 68, 82, 50, 48, 41, 60। इसकी तुलना नमूना माध्य और माध्यिका से कीजिए।
  - स्पैम संदेश की पहचान के लिए उन शब्दों की एक सूची बनाई जाती है जो सामान्य संदेशों की तुलना में स्पैम में अधिक बार दिखाई देते हैं। इस सूची को "स्पैम सूची" कहा जाता है। मान लीजिए कि

आपके डेटाबेस में कुल 5000 सदेश हैं, जिनमें से 1700 सदेश स्पैम हैं। स्पैम सदेशों में से 1343 सदेशों में "स्पैम सूची" के शब्द शामिल हैं। सामान्य सदेशों की संख्या 3300 है, जिनमें से केवल 297 सदेशों में "स्पैम सूची" के शब्द हैं। यह ज्ञात कीजिए कि यदि किसी सदेश में "स्पैम सूची" के शब्द पाए जाते हैं, तो उस सदेश के स्पैम होने की प्रायिकता क्या होगी।

2. (a) A class of 20 students took a statistics exam of 100 marks. Their scores were: 85, 92, 78, 89, 95, 87, 75, 90, 82, 88, 91, 84, 79, 93, 86, 80, 83, 94, 81, 77. Calculate the median, lower fourth, upper fourth, and range for this data. Does this data have any outliers? (5)

- (b) In a study, physicians were asked what the odds of breast cancer would be in a woman who was initially thought to have a 1% risk of cancer but who ended up with a positive cancer test result. This test gives a positive result in 80% of cases who have cancer and a positive result in 90% of cases for those who do not have cancer. If the test result is positive, what is the probability that the woman has breast cancer? (5)

time to board the plane. Knowing that, suppose an airline sells 178 tickets for the 168 seats. What is the approximate probability that not everyone who arrives at the gate on time can be accommodated? Under what circumstances this approximation is valid? (5)

- (b) The cross-sectional area of a plastic tube for watering plants is normally distributed with  $\mu=12.5$  mm<sup>2</sup> and  $\sigma=0.2$  mm<sup>2</sup>. When the area is less than 12.0 mm<sup>2</sup> or greater than 13.0 mm<sup>2</sup>, the tube does not fit properly in the hose. If the tubes are shipped in boxes of one thousand, how many wrong-sized tubes per box can sellers expect to find? (5)

- (क) कुछ निश्चित मार्गों पर उड़ान भरने वाले सामान्य विमानों में 168 इकॉनमी-क्लास सीटें होती हैं। अनुभव से पता चला है कि उन उड़ानों में सभी टिकट धारकों में से केवल 90% ही वास्तव में विमान में सवार होने के लिए समय पर पहुंचते हैं। यह मानते हुए कि एयरलाइन 168 सीटों के लिए 178 टिकट बेचती है, इस बात की अनुमानित प्रायिकता ज्ञात कीजिए कि जो यात्री समय पर गेट पर पहुंचते हैं, उनमें से सभी को सीट नहीं मिल पाएगी। यह निर्धारण कीजिए कि यह अनुमान किस स्थिति में मान्य होगा?



(b) Packages have a nominal net weight of 1 kg. However, their actual net weights have a uniform distribution over the interval of 980 grams to 1030 grams. Find the probability that the net weight of a package is less than 1 kg. If the net weights of packages are independent, also find the probability that, in a sample of five packages, all five net weights are less than  $w$  grams. (5)

(क) एक फैक्ट्री में प्रतिदिन 300 वस्तुएं बनती हैं। सख्त गुणवत्ता नियंत्रण के कारण, किसी भी वस्तु के दोषपूर्ण होने की प्रायिकता 0.01 है। प्रत्येक वस्तु में दोष स्वतंत्र रूप से उत्पन्न होते हैं। अधिकतम तीन वस्तुओं के दोषपूर्ण होने की प्रायिकता क्या है?

(ख) पैकेजों का अंकित शुद्ध भार 1 किलोग्राम है। हालाँकि, उनके वास्तविक शुद्ध भार का 980 ग्राम से 1030 ग्राम के अंतराल पर एक समान वितरण होता है। प्रायिकता ज्ञात कीजिए कि पैकेज का शुद्ध भार 1 किलोग्राम से कम है। यदि पैकेजों का शुद्ध भार स्वतंत्र है, तो यह भी ज्ञात कीजिए कि पाँच पैकेजों के नमूने में, सभी पाँच पैकेजों का शुद्ध भार  $w$  ग्राम से कम होने की प्रायिकता क्या है।

8. (a) Typical Plane flying on certain routes are configured to have 168 economy-class seats. Experience has shown that only 90% of all ticket holders on those flights will actually show up in

(क) 20 छात्रों की एक कक्षा ने 100 अंकों की सांख्यिकी परीक्षा दी। उनके अंक इस प्रकार हैं : 85, 92, 78, 89, 95, 87, 75, 90, 82, 88, 91, 84, 79, 93, 86, 80, 83, 94, 81, 77। इस डेटा के लिए माध्यिका, निचले चतुर्थक, ऊपरी चतुर्थक और श्रेणी की गणना कीजिए। यह भी ज्ञात कीजिए कि क्या डेटा में कोई बाह्य मान है।

(ख) एक अध्ययन में चिकित्सकों से यह पूछा गया कि उस महिला में स्तन कैंसर होने की संभावना कितनी होगी, जिसे पहले 1% कैंसर होने का जोखिम माना गया था, लेकिन उसका कैंसर परीक्षण सकारात्मक आया। यह परीक्षण कैंसर वाले 80% मामलों में सकारात्मक परिणाम देता है। यह परीक्षण उन 90% मामलों में भी सकारात्मक परिणाम देता है, जिनमें कैंसर नहीं है। यदि परीक्षण सकारात्मक है, तो महिला में स्तन कैंसर होने की प्रायिकता क्या है?

3. (a) A deadly car accident in a leading metropolitan state could have occurred as the result of potholes on the road, malfunctioning of the vehicle, rash driving, or use of mobile phones while driving. Interviews with the experts revealed that such an accident would usually occur with a probability of 0.25 as a result of potholes on the road, 0.20 as a result of malfunctioning of the vehicle, 0.40 as a result of rash driving, and 0.75 due to use of mobile phones while driving. These interviews also

yielded subjective estimates of the prior probabilities of these four causes of 0.30, 0.40, 0.15, and 0.15, respectively. What was the most likely cause of the accident? (5)

- (b) For the data given below calculate the relative frequency, cumulative frequency, cumulative relative frequency, and density.

Class	2-<4	4-<6	6-<8	8-<12	12-<20	20-<30
Frequency	9	15	5	9	8	2

(5)

- (क) किसी प्रमुख महानगरीय राज्य में एक घातक कार दुर्घटना सड़क पर गड़दों, वाहन की खराबी, लापरवाही से गाड़ी चलाने या गाड़ी चलाते समय मोबाइल फोन के इस्तेमाल के परिणामस्वरूप हो सकती है। विशेषज्ञों के साथ साक्षात्कार से पता चला कि ऐसी दुर्घटना आमतौर पर सड़क पर गड़दों के कारण 0.25, वाहन खराब होने के कारण 0.20, लापरवाही से गाड़ी चलाने के कारण 0.40 और वाहन चलाते समय मोबाइल फोन के उपयोग के कारण 0.75 होती है। इन साक्षात्कारों से इन चार कारणों की प्राथमिक प्रायिकता का व्यक्तिपरक अनुमान भी प्राप्त हुआ, जो क्रमशः 0.30, 0.40, 0.15, और 0.15 हैं। इस जानकारी के आधार पर, यह निर्धारित कीजिए कि दुर्घटना का सबसे संभावित कारण कारण क्या था?

- (b) The lifespan of a car battery averages six years. Suppose the battery lifespan follows an exponential distribution. Find the probability that a randomly selected car battery will last more than four years. Also find an expression for the 95th percentile of the battery lifespan. (5)

- (क) एक नौकरी के लिए 120 आवेदकों में से केवल 80 ही वास्तव में योग्य हैं। यदि आवेदकों में से 5 को गहन साक्षात्कार के लिए यादृच्छिक चयन किया जाता है, तो प्रायिकता ज्ञात कीजिए कि 5 में से केवल 2 ही नौकरी के लिए योग्य होंगे। नौकरी के लिए योग्य आवेदकों का अपेक्षित मान और विचरण ज्ञात कीजिए।

- (ख) कार बैटरी का औसत जीवनकाल छह वर्ष है। मान लीजिए कि बैटरी का जीवनकाल एक घातीय वितरण का अनुसरण करता है। इस बात की प्रायिकता ज्ञात कीजिए कि यादृच्छिक रूप से चयनित कार बैटरी चार वर्षों से अधिक समय तक चलेगी। बैटरी जीवनकाल के 95वें प्रतिशतक के लिए एक व्यंजक भी ज्ञात कीजिए।

7. (a) A factory produces 300 items daily. Due to strict quality control, the probability that any given item is defective is 0.01. Defects in individual items are independent of one another. What is the probability that at most three items are defective? (5)

$$f(x) = \begin{cases} \frac{3}{x^4} & \text{for } x \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

पिच व्यास के माप औसत मान के 2 मानक विचलन के भीतर होने की क्या प्रायिकता है?

(ख) किसी सुपरमार्केट में चेकआउट की अवधि  $X$  (मिनटों में मापी गई) की संचयी वितरण फलन इस प्रकार है :

$$F(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \frac{x^2}{4} & \text{for } 0 < x \leq 2 \\ 1 & \text{for } x > 2 \end{cases}$$

चेकआउट अवधि की माध्यिका और इंटरक्वार्टाइल रेंज ज्ञात कीजिए।

### SECTION III (खण्ड III)

All questions are compulsory.

सभी प्रश्न अनिवार्य हैं।

6. (a) Among the 120 applicants for a job, only 80 are actually qualified. If 5 of the applicants are randomly selected for an in-depth interview, find the probability that only 2 of the 5 will be qualified for the job. Find the expected value and variance of the applicants qualified for the job. (5)

(ख) नीचे दिए गए डेटा के लिए सापेक्ष आवृत्ति, संचयी आवृत्ति, संचयी सापेक्ष आवृत्ति और घनत्व की गणना कीजिए।

वर्ग	2-<4	4-<6	6-<8	8-<12	12-<20	20-<30
आवृत्ति	9	15	5	9	8	2

### SECTION II (खण्ड II)

All questions are compulsory.

सभी प्रश्न अनिवार्य हैं।

4. (a) A lottery ticket has a grand prize of Rs. 2 lakhs. The probability of winning the grand prize is .00023. Based on the expected value of the lottery ticket, would you pay Re 1 for a ticket? Show your calculations and reasoning clearly. (5)
- (b) The cost of a certain vehicle diagnostic test depends on the number of cylinders  $X$  in the vehicle's engine. Suppose the cost function is given by  $h(X) = 20 + 3X + 0.5X^2$ . Since  $X$  is a random variable, so is  $Y = h(X)$ . The pmf of  $X$  is as follows:

x	4	6	8
p(x)	0.5	0.3	0.2

Derive the pmf of Y and E(Y). (5)

(क) एक लॉटरी टिकट पर 2 लाख रुपये का पुरस्कार है। पुरस्कार जीतने की प्रायिकता .00023 है। लॉटरी टिकट के अपेक्षित मूल्य के आधार पर, क्या आप एक टिकट के लिए 1 रुपये का भुगतान करेंगे? अपनी गणना और तर्क स्पष्ट रूप से दिखाइए।

(ख) एक निश्चित वाहन निदान परीक्षण की लागत वाहन के इंजन में सिलेंडर X की संख्या पर निर्भर करती है। मान लीजिए कि लागत फलन  $h(X) = 20 + 3X + 0.5X^2$  द्वारा दिया गया है। चूंकि X एक यादृच्छिक चर है, इसलिए  $Y = h(X)$  का संभाव्यता द्रव्य फलन इस प्रकार है :

x	4	6	8
p(x)	0.5	0.3	0.2

Y का संभाव्यता द्रव्य फलन (pmf) और अपेक्षित मान (E(Y)) ज्ञात कीजिए।

5. (a) Certain coded measurements of the pitch diameter of threads of a fitting have the probability density function;

$$f(x) = \begin{cases} \frac{3}{x^4} & \text{for } x \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the probability that measurements of the pitch diameter of threads of a fitting are within 2 standard deviations of the mean value? (5)

- (b) The CDF function of the checkout duration, X, in a certain supermarket, measured in minutes, is

$$F(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \frac{x^2}{4} & \text{for } 0 < x \leq 2 \\ 1 & \text{for } x > 2 \end{cases}$$

Find the median and the interquartile range of the checkout duration. (5)

- (क) एक फिटिंग के थ्रेड्स के पिच व्यास के कुछ कोडित माप निम्नलिखित संभाव्यता घनत्व फलन द्वारा दर्शाए गए हैं :