

[This question paper contains 20 printed pages.]

9902

Your Roll No. ....

M.Com. / Sem. I

B

Paper-6102 : STATISTICAL ANALYSIS

(Admissions of 2011 and onwards)

Time : 3 Hours

Maximum Marks : 100

*(Write your Roll No. on the top immediately  
on receipt of this question paper.)*

*Note :- Answers may be written either in English or in  
Hindi; but the same medium should be used  
throughout the paper.*

*Attempt All the questions. All the parts of a  
question must be answered together and in the  
given order. Use of simple calculator is allowed.*

1. (a) Which probability approach would you use to calculate the following probabilities? Give reason also.
  - (i) Getting an ace in the next draw of cards.
  - (ii) A nuclear accident at the site of a nuclear power plant.
  - (iii) A student of M.Com. scoring a first division in Statistics.

P.T.O.

- (iv) The share price of Coal India Ltd. increasing next week. (6)
- (b) A telemarketer sells magazine subscription over the telephone. The probability of a busy signal or no answer is 65%. If the telemarketer does make contact, the probability of 0, 1, 2 or 3 magazine subscriptions is .5, .2, .20, and .05, respectively. Find the probability that in one call she sells no magazines. (7)
- (c) Bad gums may mean a bad heart. Researchers at a medical college discovered that 85% of people who have suffered a heart attack had periodontal disease, an inflammation of the gums. Only 29% of healthy people (who have not suffered a heart attack) have this disease. Suppose that in a certain community, heart attacks are occurring with 40% probability.
- (i) If someone has periodontal disease, what is the probability that he or she will have a heart attack ?
- (ii) What is the probability of having a periodontal disease ? (7)

OR

- (d) Explain Bayes' theorem with an example. (6)

- (e) Researchers at a Medical College have determined that children under 2 years old who sleep with the lights on have a 42% chance of becoming myopic before they are 16. Children who sleep in darkness have a 21% probability of becoming myopic. A survey indicates that 28% of children under 2 sleep with some light on. Find the probability that a child under 16 is myopic. (7)
- (f) Three airlines serve a small city. Airline A has 50% of all the scheduled flights, airline B has 30% and airline C has the remaining 20%. Their on-time rates are 80%, 65% and 40% respectively. A plane has just left on time.
- (i) What is the probability that it was Airline A ?
- (ii) What is the probability that it was Airline C ? (7)
- (क) निम्नलिखित प्रायिकताओं का परिकलन करने के लिए आप किस प्रायिकता उपागम का प्रयोग करेंगे ? कारण बताइए ।
- (i) पत्तों की अगली खींच में इक्का प्राप्त करना ।
- (ii) न्यूक्लियर शक्ति संयंत्र के स्थल पर न्यूक्लीयर दुर्घटना का होना ।
- (iii) M.Com का एक छात्र जो सांख्यिकी में प्रथम डिवीज़न के अंक प्राप्त करता है ।

(iv) कोल इन्डिया लिमिटेड के शेयर की कीमत अगले सप्ताह बढ़ रही है ।

(ख) एक टेलीमार्केटर टेलीफोन पर मैगजीन का चंदा बेचती है । सिग्नल के व्यस्त होने या कोई उत्तर न मिलने की प्रायिकता 65% है । यदि टेलीमार्केटर संपर्क करता है तो 0, 1, 2 या 3 मैगजीन चंदों की प्रायिकता क्रमशः .5, .2, .20, और .05 है । इसकी प्रायिकता ज्ञात कीजिए कि एक कॉल में वह कोई भी मैगजीन नहीं बेचती है ।

(ग) खराब मसूड़ों का अर्थ खराब हृदय हो सकता है । मैडिकल कॉलेज के अनुसंधाताओं ने पता लगाया है कि 85% लोगों को, जिन्हें हृदय रोध हुआ है, पेरियोडोन्टल अर्थात् मसूड़ों की सूजन का रोग था केवल 29% स्वस्थ लोगों को यह रोग है । जिन्हें हृदय रोध नहीं हुआ है) । मान लीजिए कि एक खास समुदाय में हृदय-रोध की प्रायिकता 40% है ।

(i) यदि किसी को पेरियोडोन्टल रोग है तो इसकी प्रायिकता क्या है कि उसे हृदय रोध होगा ?

(ii) पेरियोडोन्टल रोग होने की प्रायिकता क्या है ?

अथवा

(घ) बेयस प्रमेय की उदाहरण देकर ब्याख्या कीजिए ।

(ङ) मैडिकल कॉलेज के अनुसंधाताओं ने निर्धारित किया है कि 2 वर्ष से कम आयु के बच्चों में, जो रोशनी में सोते हैं, 42% बच्चों में 16 वर्ष की आयु से पहले निकटदृष्टि रोग हो जाता है । जो

बच्चे अँधेरे में सोते हैं उनमें निकटदृष्टि रोग की प्रायिकता 21% है। सर्वेक्षण से पता चला है कि 2 वर्ष से कम आयु वाले 28% बच्चे कुछ रोशनी में सोते हैं। इसकी प्रायिकता ज्ञात कीजिए कि 16 वर्ष से कम आयु के बच्चों में निकटदृष्टि रोग है।

(च) तीन ऐयरलाइन एक छोटे शहर में आती-जाती है। A ऐयरलाइन की 50% उड़ानें अनुसूचित उड़ाने हैं, B ऐयरलाइन की 30% और C ऐयरलाइन की शेष 20% उड़ाने हैं। उनकी यथासमय दरें क्रमशः 80%, 65% और 40% हैं। एक जहाज हाल ही में समय पर उड़ा है।

(i) इसकी प्रायिकता क्या है कि यह A ऐयरलाइन थी।

(ii) इसकी प्रायिकता क्या है कि यह C ऐयरलाइन थी।

2. (a) In the following cases, state whether the random variable is discrete or continuous, giving reason thereof.

(i) Gender of the customer while selling an automobile.

(ii) The distance a car travels on a litre of petrol.

(iii) The time between consecutive incoming calls in a bank.

(iv) Number of returns containing error in an audit of 100 tax returns. (6)

(b) An automatic filling machine fills cola into 750-milliliter bottles. At the start of shift an inspector will take a random sample of 50 bottles and measure the volume of cola in them accurately. After computing the sample mean, she will approve the machine operation only if the sample mean falls between 749.5 and 750.5 milliliter. Assume that during a particular shift the population mean was 750.4 milliliters and the standard deviation was 1.2 milliliters. What is the probability that the inspector will approve the machine operation ? (7)

(c) An economist wants to estimate the average amount in savings account at banks in a given region. A random sample of 100 accounts gives a mean of Rs. 357.60 and standard deviation of Rs. 140.

(i) Give a 95% confidence interval for the average amount in any savings account at a bank in the given region.

(ii) 90% confidence interval for the average amount in any savings account at a bank in the given region.

(iii) On the basis of the above results, what would you conclude about the relationship between confidence level and confidence interval ?

(7)

OR

- (d) Explain the concept of standard error of mean. (6)
- (e) The number of accidents that occur at a busy intersection is Poisson distributed with a mean of 3.5 per week. Find the probability of the following events.
- (i) Five or more accidents in one week.
  - (ii) One accident today. (7)
- (f) Audience profile data collected at a sports web site showed that 26% of the users were women. Assume that this percentage was based on a sample of 400 users.
- (i) At 95% confidence, what is the margin of error associated with the estimated proportion of users who are women?
  - (ii) How large a sample should be taken if the desired margin of error is 0.03 ?
  - (iii) On the basis of the above results, what would you conclude about the relationship between margin of error and required sample size ? (7)

(क) निम्नलिखित केशों में उल्लेख कीजिए कि यादृच्छिक परिवर्ती विविक्त हैं या सतत और इसका कारण बताइए :

- (i) एक ऑटोमोबाइल को बेचते समय ग्राहक का लिंग ज्ञात करना
- (ii) एक लीटर पेट्रोल में कार द्वारा तय की गई दूरी
- (iii) बैंक में आने वाली क्रमिक कॉलों के बीच की अवधि
- (iv) 100 कर रिटर्नों के लेखा परीक्षण में त्रुटि वाली रिटर्नों की संख्या

(ख) एक स्वचालित फिलिंग मशीन 750 मिलीलीटर बोतलों में कोला भरती है। पारी के शुरू होने पर निरीक्षक 50 बोतलों को यादृच्छिक सैम्पल लेकर उनमें कोला का सही माप करता है। सैम्पल के माध्य को कम्प्यूट करने पर वह मशीन के प्रचालन की इजाजत तभी देता है जब सैम्पल माध्य 749.5 और 750.5 मिलीलीटर के बीच होता है। मान लीजिए कि किसी खास पारी में जनसंख्या माध्य 750.4 मिलीलीटर था और मानक विचलन 1.2 मिलीलीटर था। इसकी प्रायिकता क्या है कि निरीक्षक मशीन के प्रचालन की स्वीकृति देगा ?

(ग) एक अर्थशास्त्री बताए गए क्षेत्र में एक बैंक में बचत खाते में औसत राशि का आकलन करना चाहता है। 100 खातों का एक यादृच्छिक सैम्पल 357.60 रु. का माध्य और 140 रु. का मानक विचलन बताता है।



- (i) बताए गए क्षेत्र के बैंक में बचत खाते में औसत राशि के लिए 95% विश्वास्थ्यता अंतराल दीजिए ।
- (ii) बताए गए क्षेत्र के बैंक में बचत खाते में औसत राशि के लिए 90% विश्वास्थ्यता अंतराल दीजिए ।
- (iii) उपर्युक्त परिणामों के आधार पर आप विश्वास्थ्यता स्तर और विश्वास्थ्यता अंतराल में संबंध के बारे में क्या निष्कर्ष निकालेंगे ?

#### अथवा

- (घ) माध्य की मानक त्रुटि की संकल्पना की व्याख्या कीजिए ।
- (ङ) किसी व्यस्त प्रतिच्छेदन पर होने वाली दुर्घटनाओं की संख्या 3.5 प्रति सप्ताह के माध्य के साथ प्वासों बंटन है । निम्नलिखित घटनाओं की प्रायिकता ज्ञात कीजिए :
- (i) पाँच या अधिक दुर्घटनाएँ प्रति सप्ताह
- (ii) आज एक दुर्घटना
- (च) एक स्पोर्ट्स वेबसाइट पर संकलित दर्शक स्थिति संबंधी आँकड़े, प्रदर्शित करते हैं कि 20% उपयोक्ता महिलाएँ थीं । मान लीजिए कि यह प्रतिशतता 400 उपयोक्ताओं के सैम्पल पर आधारित थी ।
- (i) 95% विश्वास्थ्यता पर महिला उपयोक्ताओं के अनुमानित अनुपात से संबंधित त्रुटि की सीमा क्या है ?

- (ii) यदि वांछित त्रुटि सीमा 0.03 है तो सैम्पल कितना बड़ा लिया जाना चाहिए ?
- (iii) उपर्युक्त परिणामों के आधार पर त्रुटि की सीमा और आवश्यक सैम्पल के आकार के बीच संबंध के बारे में आप क्या निष्कर्ष निकालेंगे ?
3. (a) Suppose a new production method will be implemented if it reduces the mean operating cost per hour. The mean cost for the current production method is Rs. 220 per hour.
- (i) Develop the appropriate null and alternative hypotheses for the above situation.
- (ii) What is the Type I error in this situation? Explain.
- (ii) What is the Type II error in this situation? Explain. (2+2+2)
- (b) X Ltd.'s advertising agency has developed two new packaging designs for a soap. To determine which design is better, the marketing manager selected two supermarkets. The first packaging design A (more expensive) was used in one supermarket and the second packaging design B was used in the other supermarket. A random sample of soap buyers was taken in each of the supermarkets and the proportion of people buying

X Ltd.'s brand of soap was determined. The following data was obtained :

	Sample size	People buying X Ltd.'s soap
Supermarket 1	904	180
Supermarket 2	1038	155

As design A is more expensive, the company will switch to design A only when it outsells design B by more than 3%. Should the management switch to design A ? ( $\alpha = 0.05$ ). (14)

OR

- (c) State the characteristics of the t-distribution. (6)
- (d) A company wants to determine whether individuals who are expected at the time of employment to be better salesperson actually turnout to have better sales records. To investigate this, the HR manager ranked 10 salespersons in terms of their potential for success, basing the assessment solely on the basis of information available at the time of employment. On the basis of actual sales performance, a second ranking of these 10 salespersons was obtained. The rankings were as follows :

Salesperson	Ranking of potential	Ranking according to sales
A	2	1
B	4	3
C	7	5
D	1	6
E	6	7
F	3	4
G	10	10
H	9	8
I	8	9
J	5	2

Do these data indicate that the expected potential at the time of employment and sales record are related to each other? ( $\alpha = 0.01$ ). (14)

(क) माम लीजिए कि एक नई उत्पादन विधि निष्पादित हो जाएगी यदि यह माध्य प्रचालन लागत प्रति घंटा कम कर देती है। चालू उत्पादन विधि के लिए माध्य लागत 220 रु. प्रति घंटा है।

(i) उपर्युक्त स्थिति के लिए उपयुक्त वातिल और वैकल्पिक परिकल्पना को विकसित कीजिए।

(ii) इस स्थिति में टाइप I त्रुटि क्या है? व्याख्या कीजिए।

(iii) इस स्थिति में टाइप II त्रुटि क्या है? व्याख्या कीजिए।

(ख) X लिमिटेड नामक विज्ञापन एजेन्सी ने साबुन के लिए दो नए पैकेजिंग डिजाइन तैयार किए हैं। इनमें से कौनसा डिजाइन बहतर है यह निर्धारित करने के लिए विपणन प्रबंधक ने दो सुपरबाजारों को चुना है। प्रथम पैकेजिंग डिजाइन A (अधिक कीमती) को एक सुपरबाजार में और दूसरे पैकेजिंग डिजाइन B को दूसरे सुपरबाजार में प्रयोग में लिया गया। प्रत्येक सुपरबाजार से साबुन खरीदने वालों का एक यादृच्छिक सैम्पल लिया गया और X लिमिटेड के ब्रांड वाले साबुन को खरीदने वाले लोगों के अनुपात का निर्धारण किया गया।

निम्नलिखित आँकड़े प्राप्त हुए :

	सैम्पल का आकार	X लिमिटेड के साबुन को खरीदने वाले लोगों की संख्या
सुपरबाजार 1	904	180
सुपरबाजार 2	1038	155

चूँकि डिजाइन A अधिक कीमती है इसलिए कंपनी डिजाइन A को इस्तेमाल तभी करेगी जब डिजाइन B 3% से भी अधिक बिक जाता है। बताइए कि क्या प्रबंध को डिजाइन A का इस्तेमाल करना चाहिए ? ( $\alpha = 0.05$ )

अथवा

(ग) t-बंटन की विशेषताओं का उल्लेख कीजिए।

(घ) एक कंपनी यह निर्धारित करना चाहती है, कि रोजगार के समय जिन व्यक्तियों से बहतर विक्रयकर्ता होने की आशा थी वे वास्तव में वे व्यक्ति थे जिनका विक्रय रिकार्ड बहतर था। इसका निरीक्षण करने के लिए मानव संसाधन प्रबंधक ने 10 विक्रयकर्ताओं को उनकी सफलता के लिए उनकी शक्ति के पदों में रैंकिंग दिया और यह निर्धारण केवल रोजगार देते समय उपलब्ध सूचना के आधार पर किया गया। वास्तविक विक्रय निष्पादन के आधार पर इन 10 विक्रयकर्ताओं को दूसरा रैंकिंग दिया गया। ये रैंकिंग इस प्रकार थे:

विक्रयकर्ता	शक्ति का रैंकिंग	विक्रयानुसार रैंकिंग
A	2	1
B	4	3
C	7	5
D	1	6
E	6	7
F	3	4
G	10	10
H	9	8
I	8	9
J	5	2

क्या ये आँकड़े यह संकेत देते हैं कि रोजगार के समय प्रत्याशित शक्ति और विक्रय रिकार्ड परस्पर संबंधित हैं? ( $\alpha = 0.01$ )

4. (a) Distinguish between parametric and non-parametric tests. Compare the conditions under which we should use parametric and non-parametric tests. (6)
- (b) To test whether the mean time needed to mix a batch of materials is the same for machines produced by 3 manufacturers, a chemical company obtained the following data on the time (in minutes) needed to mix the material. Use these data to determine whether the population mean times for mixing a batch of material differ for the 3 manufacturers? (Use  $\alpha = 0.05$ )

**Manufacturer**

1	2	3
20	28	20
26	26	19
24	31	23
22	27	22

(14)

**OR**

- (c) Which control chart is used for quality control related to process variability? Explain its working. (6)

P.T.O.

(d) For the given set of data.

(i) Calculate the multiple regression equation.

(ii) Predict Y when  $X_1 = 3.0$  and  $X_2 = 2.7$ .

Y	$X_1$	$X_2$
25	3.5	5.0
30	6.7	4.2
11	1.5	8.5
22	0.3	1.4
27	4.6	3.6
19	2.0	1.3

(14)

(क) प्राचलिक और अप्राचलिक परीक्षणों में अंतर बताइए। उन दशाओं की तुलना कीजिए जिनमें हमें प्राचलिक और अप्राचलिक परीक्षणों का प्रयोग करना चाहिए।

(ख) यह जाँच करने के लिए कि पदार्थों के एक बैच को मिश्रित करने के लिए आवश्यक माध्य समय 3 मशीन निर्माताओं द्वारा तैयार की गई मशीनों के लिए समान होता है, एक रसायन कंपनी को पदार्थ को मिश्रित करने के लिए आवश्यक समय (मिनटों में) के बारे में निम्नलिखित आँकड़े प्राप्त हुए। इन आँकड़ों का प्रयोग यह निर्धारित करने के लिए कीजिए कि क्या तीन मशीन निर्माताओं के लिए पदार्थों के एक बैच को मिश्रित करने के लिए जनसंख्या माध्य समय भिन्न होते हैं? (प्रयोग कीजिए  $\alpha = 0.05$ )



## मशीन निर्माता

1	2	3
20	28	20
26	26	19
24	31	23
22	27	22

## अथवा

(ग) परिवर्तिता की प्रक्रिया से संबंधित गुणवत्ता नियंत्रण के लिए किस नियंत्रण चार्ट को प्रयोग में लिया जाता है ? इसकी क्रियाविधि की व्याख्या कीजिए ।

(घ) आँकड़ों के दिए गए समुच्चय के लिए :

(i) बहुसमाश्रयण समीकरण का परिकलन कीजिए ।

(ii)  $y$  का पूर्वानुमान कीजिए जब  $X_1 = 3.0$  और  $X_2 = 2.7$

Y	$X_1$	$X_2$
25	3.5	5.0
30	6.7	4.2
11	1.5	8.5
22	0.3	1.4
27	4.6	3.6
19	2.0	1.3

5. (a) Distinguish between simple random sampling and systematic sampling. Give examples. (6)
- (b) A tyre manufacturer randomly samples 40 tyres at the end of each shift to test for tyres that are defective. The number of defectives in 12 shifts is as follows : 4, 2, 0, 5, 2, 3, 14, 2, 3, 4, 12, 3. Construct a p-chart for this process. Is the production process under control ? (14)

**OR**

- (c) Explain the least squares method of finding the regression line. (6)
- (d) A local pizza restaurant and a local branch of a national pizza chain are located across the street from a college campus. The local pizza restaurant advertises that it delivers to the hostels faster than the national chain. In order to determine whether this advertisement is valid, two groups of hostel students decide to order pizzas from the local pizza restaurant as well as from the national chain. Each time a group orders a pizza from the local pizza restaurant, the other group orders a pizza from the national pizza chain. For each of the ten times pizzas were ordered, the following measurements (in minutes) are available.

Time	Local pizza restaurant	National chain
1	16.8	22.0
2	11.7	15.2
3	15.6	18.7
4	16.7	15.6
5	17.5	20.8
6	18.1	19.5
7	14.1	17.0
8	21.8	19.5
9	13.9	16.5
10	20.8	24.0

At a 5% level of significance, is there evidence to conclude that the mean delivery time for the local pizza restaurant is less than the mean delivery time for the national pizza chain? (14)

(क) सरल यादृच्छिक सैम्पलिंग और व्यवस्थित सैम्पलिंग में अंतर बताइए। उदाहरण दीजिए।

(ख) एक टायर निर्माता प्रत्येक पारी के अंत में 40 टायरों का यादृच्छिक सैम्पल दोषपूर्ण टायरों की जाँच करने के लिए लेता है। दोषपूर्ण टायरों की संख्या 12 पारियों में इस प्रकार है : 4, 2, 0, 5, 2, 3, 14, 2, 3, 4, 12, 3. इस प्रक्रिया के लिए p-चार्ट की रचना कीजिए। क्या उत्पादन प्रक्रिया नियंत्रण में है ?

अथवा

(ग) समाश्रयण रेखा ज्ञात करने की न्यूनतम वर्ग विधि की व्याख्या कीजिए।

P.T.O.

(घ) एक स्थानीय पिज़ा रेस्टोरेन्ट और राष्ट्रीय पिज़ा शृंखला की एक स्थानीय शाखा कॉलेज परिसर की सड़क के पार स्थित हैं। स्थानीय पिज़ा रेस्टोरेन्ट विज्ञापन करता है कि वह राष्ट्रीय शृंखला से अधिक तीव्रता से छात्रावासों को पिज़ा पहुँचाता है। यह निर्धारित करने के लिए कि क्या यह विज्ञापन वैध है छात्रावास के छात्रों के दो समूह स्थानीय पिज़ा रेस्टोरेन्ट और साथ ही राष्ट्रीय शृंखला को पिज़ा का आदेश देने का निश्चय करते हैं। हर बार जब एक समूह स्थानीय पिज़ा रेस्टोरेन्ट को पिज़ा का आदेश देता है, दूसरा समूह राष्ट्रीय पिज़ा शृंखला को पिज़ा का आदेश देता है। दस बार दिए गए पिज़ा के आदेशों में से प्रत्येक के लिए निम्नलिखित माप (मिनटों में) उपलब्ध हैं :

समय	स्थानीय पिज़ा रेस्टोरेन्ट	राष्ट्रीय पिज़ा शृंखला
1	16.8	22.0
2	11.7	15.2
3	15.6	18.7
4	16.7	15.6
5	17.5	20.8
6	18.1	19.5
7	14.1	17.0
8	21.8	19.5
9	13.9	16.5
10	20.8	24.0

5% सार्थकता स्तर पर क्या कोई प्रमाण यह निष्कर्ष निकालने के लिए है कि स्थानीय पिज़ा रेस्टोरेन्ट का माध्य पिज़ा पहुँचाने का समय राष्ट्रीय पिज़ा शृंखला के माध्य पिज़ा पहुँचाने के समय से कम है ?