

[This question paper contains 12 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 3016

C

Unique Paper Code : 62273506

Name of the Paper : Data Analysis

Name of the Course : BA Prog. (CBCS)

Semester : V

Duration : 3 Hours

Maximum Marks : 65

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. All questions carry equal marks (13 marks each).
3. Attempt any Five questions.
4. Use of simple calculator is allowed.
5. Required Tables will be provided by the colleges.
6. Answers may be written either in English or Hindi; but the same medium should be used throughout the paper.

P.T.O.

3016

2

छात्रों के लिए निर्देश

1. इस प्रश्न-पत्र के मिलते ही ऊपर दिए गए निर्धारित स्थान पर अपना अनुक्रमांक लिखिए।
2. सभी प्रश्नों के अंक समान हैं (प्रत्येक के 13 अंक)।
3. किन्हीं पाँच प्रश्नों के उत्तर दीजिए।
4. साधारण कैलकुलेटर के उपयोग की अनुमति है।
5. आवश्यक तालिकाएं महाविद्यालयों द्वारा उपलब्ध कराई जाएंगी।
6. इस प्रश्न-पत्र का उत्तर अंग्रेजी या हिंदी किसी एक भाषा में दीजिए, लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए।

1. (i) How much time do college students spend using their Mobile phones? A 2014 Ahmadabad University study showed college students spend an average of 9 hours a day using their mobile phones for the following activities :

(6,2)

Mobile phone activity	Percentage
Banking	2%
Checking Data & Time	8%

3016

3

Listening to Music	12%
Playing Games	4%
Reading	3%
Sending Emails	9%
Social Media	18%
Surfing the Internet	7%
Taking Photos	3%
Talking	6%
Texting	18%
YouTube	2%
Others	8%

- (a) How can we construct a bar chart, a pie chart, and a Pareto chart of the above data by using Excel/R? Write syntax also.
- (b) Which graphical method do you think is best for portraying these data?
- (ii) What are the various methods of sampling? Explain any two. (5)
- (i) कॉलेज के छात्र अपने मोबाइल फोन का उपयोग करने में कितना समय व्यतीत करते हैं? 2014 के अहमदाबाद विश्वविद्यालय

P.T.O.

के एक अध्ययन से पता चला है कि कॉलेज के छात्र निम्नलिखित गतिविधियों के लिए अपने मोबाइल फोन का उपयोग करके दिन में औसतन 9 घंटे खर्च करते हैं :

Mobile phone activity	Percentage
Banking	2%
Checking Data & Time	8%
Listening to Music	12%
Playing Games	4%
Reading	3%
Sending Emails	9%
Social Media	18%
Surfing the Internet	7%
Taking Photos	3%
Talking	6%
Texting	18%
YouTube	2%
Others	8%

(अ) हम Excel/R का उपयोग करके उपरोक्त डेटा का एक बार चार्ट, एक पाई चार्ट और एक परेडो चार्ट कैसे बना सकते हैं? वाक्य रचना भी लिखें।

(ब) आपके विचार में इन आँकड़ों को चित्रित करने के लिए कौन-सी चित्रण विधि सर्वोत्तम है?

(ii) प्रतिचयन की विभिन्न विधियाँ क्या हैं? किन्हीं दो को स्पष्ट कीजिए।

2. (i) The following data represents the per litre consumption of two wheelers : (6,2)

26	38	26	30	24	26	28	28	24	26	24	39
32	23	22	24	28	37	31	40	25	25	33	30

(a) How variance, standard deviation, range can be calculated by using either excel/R of the above data? Write syntax only.

(b) Graphically represent the skewness of the above data.

(ii) Highlight the differences among the mean, median, and mode. (5)

(i) निम्नलिखित डेटा दोपहिया वाहनों की प्रति लीटर खपत को दर्शाता है :

26	38	26	30	24	26	28	28	24	26	24	39
32	23	22	24	28	37	31	40	25	25	33	30

3016

6

(अ) उपरोक्त डेटा के एक्सेल/आर का उपयोग करके भिन्नता, मानक विचलन, सीमा की गणना कैसे की जा सकती है? केवल सिटैक्स लिखें।

(ब) उपरोक्त आँकड़ों की विषमता को आलेखीय रूप से निरूपित करें।

(ii) माध्य, माध्यिका और बहुलक के बीच के अंतरों पर प्रकाश डालिए।

3. (i) The following is a set of data from a sample of $n = 7$: (6.2)

12	7	4	9	0	7	3
----	---	---	---	---	---	---

(a) Compute the first quartile (Q_1), Second Quartile (Q_2) and third quartile (Q_3) for the above data

(b) Compute the Z-score of the above data

(ii) Explain Chebyshev theorem. (5)

(i) निम्नलिखित $n = 7$ के नमूने से डेटा का एक सेट है :

12	7	4	9	0	7	3
----	---	---	---	---	---	---

3016

7

(अ) उपरोक्त आँकड़ों के लिए प्रथम चतुर्थक (Q_1), द्वितीय चतुर्थक (Q_2) और तृतीय चतुर्थक (Q_3) की गणना करें।

(ब) उपरोक्त डेटा के Z-स्कोर की गणना करें।

(ii) चेबीशेव प्रमेय की व्याख्या करें।

4. (i) The Cereals lists the calories and sugar, in grams, in one serving of seven breakfast cereals :

(6.2)

Cereals	Calories	Suger
Kellogg's All Bran	80	6
Kellogg's Corn Flakes	100	2
Wheaties	100	4
Nature's Path Organic Multigrain Flakes	110	4
Kellogg's Rice Krispies	130	4
Post Shredded Wheat	190	11
Vanilla Almond		
Kellogg's Mini Wheats	200	10

(a) Compute the covariance.

(b) Compute the coefficient of correlation.

P.T.O.

- (ii) Why is the sample mean an unbiased estimator of the population mean? Prove it with the help of given data set (5)

Administrative Assistant	Number of Errors
A	$X_1 = 3$
B	$X_2 = 2$
C	$X_3 = 1$
D	$X_4 = 4$

- (i) अनाज कैलोरी और चीनी को ग्राम में, सात में से एक सर्विंग में सूचीबद्ध करता है नाश्ता अनाज :

Cereals	Calories	Suger
Kellogg's All Bran	80	6
Kellogg's Corn Flakes	100	2
Wheaties	100	4
Nature's Path Organic Multigrain Flakes	110	4
Kellogg's Rice Krispies	130	4
Post Shredded Wheat	190	11
Vanilla Almond		
Kellogg's Mini Wheats	200	10

- (अ) सहप्रसरण की गणना कीजिए।
(ब) सहसंबंध के गुणांक की गणना करें।

- (ii) नमूना माध्य जनसंख्या का निष्पक्ष अनुमानक माध्य क्यों है? दिए गए डेटा सेट की मदद से इसे साबित करें।

प्रशासनिक सहायक	त्रुटियों की संख्या
A	$X_1 = 3$
B	$X_2 = 2$
C	$X_3 = 1$
D	$X_4 = 4$

5. (i) What are the risks in decision making using hypothesis testing? (5)
(ii) The following is a series of annual sales (in \$millions) over an 11-year period (2004 to 2014): (6,2)

Year:	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Sales:	13.0	17.0	19.0	20.0	20.5	20.5	20.5	20.0	19.0	17.0	13.0

- (a) Construct a time-series plot.
(b) Does there appear to be any change in annual sales over time? Explain.
(i) परिकल्पना परीक्षण का उपयोग करके निर्णय लेने में जोखिम क्या है?

- (ii) निम्नलिखित 11 साल की अवधि (2004 से 2014) में वार्षिक बिक्री (\$ मिलियन में) की एक श्रृंखला है :

Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Sales	13.0	17.0	19.0	20.0	20.5	20.5	20.5	20.0	19.0	17.0	13.0

- (अ) एक समय-श्रृंखला प्लॉट का निर्माण करें।
 (ब) क्या समय के साथ वार्षिक बिक्री में कोई बदलाव दिखाई देता है? समझाना।
6. (i) Determining the interval that includes a fixed proportion of the sample means by using symmetrically distributed around the population mean that will include 95% of the sample means, based on the sample of 25 boxes. (6)
- (ii) Explain type I type II Error (4)
- (iii) Explain graphically the empirical rule (3)
- (i) अंतराल का निर्धारण जिसमें नमूना का एक निश्चित अनुपात शामिल है, जनसंख्या के चारों ओर सममित रूप से वितरित का उपयोग करके मतलब है कि 25 बक्से के नमूने के आधार पर नमूना साधनों का 95% शामिल होगा।

- (ii) टाइप I टाइप II एरर की व्याख्या करें।
 (iii) अनुभवजन्य नियम को आलेखीय रूप से समझाइए।

7. (i) A company that manufactures chocolate bars is particularly concerned that the mean weight of a chocolate bar is not greater than 6.03 ounces. A sample of 50 chocolate bars is selected; the sample mean is 6.034 ounces, and the sample standard deviation is 0.02 ounce. Using the $\alpha = 0.01$ level of significance, is there evidence that the population mean weight of the chocolate bars is greater than 6.03 ounces? Apply One-Tail Test for the mean on the given statement. (8)
- (ii) Explain p - value approach. (5)
- (i) चॉकलेट बार बनाने वाली कंपनी विशेष रूप से चिंतित है कि चॉकलेट बार का औसत वजन 6.03 औंस से अधिक नहीं है। 50 चॉकलेट बार का एक नमूना चुना गया है नमूना माध्य 6.034 औंस है, और नमूना मानक विचलन 0.02 औंस है। $\alpha = 0.01$ स्तर के महत्व का उपयोग करते हुए, क्या इस बात

का प्रमाण है कि जनसंख्या का मतलब चॉकलेट बार का वजन 6.03 औंस से अधिक है? दिए गए कथन पर माध्य के लिए वन-टेल टेस्ट लागू करें।

(ii) पी-मूल्य दृष्टिकोण की व्याख्या करें।

8. Write any **two** of the followings : (6.5,6.5)

(i) A Statistic and A parameter?

(ii) Six steps of Critical Value Approach to Hypothesis Testing

(iii) Box plot and Scatter plot

निम्नलिखित में से कोई दो लिखिए :

(i) एक आँकड़ा और एक पैरामीटर?

(ii) परिकल्पना परीक्षण के लिए महत्वपूर्ण मूल्य दृष्टिकोण के छह चरण

(iii) बॉक्स प्लॉट और स्कैटर प्लॉट

(10,000)