

13 MAY 2022

13/05/22

[This question paper contains 12 printed pages.]

Your Roll No.....



Sr. No. of Question Paper : 3729

Unique Paper Code : 62273601

Name of the Paper : Basic Computational Techniques
for Data Analysis

Name of the Course : B. A.(Prog.) SEC

Semester : VI

Duration : 3 Hours

Maximum Marks : 50

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. You are expected to answer any **four (4)** out of the **eight (8)** questions given below.
3. Each question is of **twelve and a half** marks (12½). Each question has sub parts.
4. Use of simple calculator is allowed.
5. Answers may be written either in English or Hindi; but the same medium should be used throughout the paper.

P.T.O.

छात्रों के लिए निर्देश

1. इस प्रश्न-पत्र के मिलते ही ऊपर दिए गए निर्धारित स्थान पर अपना अनुक्रमांक लिखिए।
2. आपसे नीचे दिए गए आठ (8) प्रश्नों में से किन्हीं चार (4) का उत्तर देने की अपेक्षा की जाती है।
3. प्रत्येक प्रश्न साढ़े बारह अंक ($12\frac{1}{2}$) का है। प्रत्येक प्रश्न के उप भाग होते हैं।
4. साधारण कैलकुलेटर के उपयोग की अनुमति है।
5. इस प्रश्न-पत्र का उत्तर अंग्रेजी या हिंदी किसी एक भाषा में दीजिए, लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए।

1. (a) How would you interpret the sign and magnitude of a calculated 'r'? Consider in particular the values $r = 0$, $r = -1$ and $r = +1$. (4)

(b) "Compound interest is an interest on interest". Explain this statement by using a suitable numerical example. (4)

(c) The linear correlation coefficient for the following data is 0.985658. Interpret the relationship

between these two variables X and Y on the basis of the following data. (4.5)

	A	B	C
1	X	Y	
2	5	4	
3	8	7	
4	9	10	
5	12	14	
6	14	15	
7	16	17	
8			Ctrl

(अ) आप परिकलित 'r' के चिह्न और परिमाण की व्याख्या कैसे करेंगे? विशेष रूप से मानों $r = 0$, $r = -1$ और $r = +1$ पर विचार करें।

(ब) "चक्रवृद्धि ब्याज ब्याज पर ब्याज है"। एक उपयुक्त संख्यात्मक उदाहरण की सहायता से इस कथन की व्याख्या कीजिए।

(स) निम्नलिखित डेटा के लिए रैखिक सहसंबंध गुणांक 0.985658 है। निम्नलिखित आँकड़ों के आधार पर इन दो चरों X और Y के बीच संबंध की व्याख्या कीजिए।

	A	B	C
1	X	Y	
2	5	4	
3	8	7	
4	9	10	
5	12	14	
6	14	15	
7	16	17	
8			

2. (a) In addition to RBI data base, which other agencies bring out the data on macroeconomic aggregates? Briefly discuss the main features and applications of such data. (4)

- (b) Write the syntax for calculation of skewness in excel from the following data set and also interpret the result of skewness found if skewness is -0.20435 . (4)

Class	A	B	C	D	E	F	G	H	I
Marks	78	58	98	87	67	54	34	58	80

- (c) (i) Find out the confidence interval and interpret it using the following data.

Data	
Sample Standard deviation	25.28
Sample mean	43.89
Sample Size	27
Confidence level	95%

- (ii) What is the level of significance from the above data?

- (iii) Define Type I and Type II error. (4.5)

- (अ) आरबीआई डेटा बेस के अलावा, कौन सी अन्य एजेंसियां मैक्रो-इकोनॉमिक एग्रीगेट्स पर डेटा लाती हैं? ऐसे डेटा की मुख्य विशेषताओं और अनुप्रयोगों पर संक्षेप में चर्चा करें।

- (ब) निम्नलिखित डेटा सेट से एक्सेल में स्क्यूनेस की गणना के लिए सिंटैक्स लिखें और यदि स्क्यूनेस -0.20435 है तो स्क्यूनेस के परिणाम की व्याख्या भी करें।

Class	A	B	C	D	E	F	G	H	I
Marks	78	58	98	87	67	54	34	58	80

- (स) (i) विश्वास अंतराल का पता लगाएं और निम्नलिखित डेटा का उपयोग करके इसकी व्याख्या करें।

Data	
Sample Standard deviation	25.28
Sample mean	43.89
Sample Size	27
Confidence level	95%

- (ii) उपरोक्त आंकड़ों से महत्व का स्तर क्या है ?
- (iii) टाइप I और टाइप II त्रुटि को परिभाषित करें।

3. (a) Explain the nature and types of data used in social science research. Discuss the various methods of analysis of quantitative data. (4)

(b) The following table indicates the percentage of residential electricity consumption in the United States, in a recent year organized by type of use. (4)

Type of Use	Percentage %
Cooking	2%
Cooling	15%
Electronics	9%
Heating	15%
Lighting	13%

Refrigeration	10%
Water Heating	10%
Wet Cleaning	3%
Other	23%

- (i) Construct a bar chart, and a pie chart.
- (ii) Which graphical method do you think is best for portraying above data?
- (c) Can we compare the standard deviation for both the following data sets? State how can you determine the variability of these data sets?

(4.5)

	A	B	C	D
1	Student	Height (C)	Weight (KG)	
2	1	178	65	
3	2	180	70	
4	3	175	71	
5	4	185	72	
6	5	177	73	
7	6	190	79	

(अ) सामाजिक विज्ञान अनुसंधान में प्रयुक्त डेटा की प्रकृति और प्रकारों की व्याख्या करें। मात्रात्मक डेटा के विश्लेषण के विभिन्न तरीकों पर चर्चा करें।

- (ब) निम्न तालिका संयुक्त राज्य अमेरिका में हाल के वर्ष में उपयोग के प्रकार द्वारा आयोजित आवासीय बिजली खपत के प्रतिशत को इंगित करती है।

Type of Use	Percentage %
Cooking	2%
Cooling	15%
Electronics	9%
Heating	15%
Lighting	13%
Refrigeration	10%
Water Heating	10%
Wet Cleaning	3%
Other	23%

- (i) एक बार चार्ट और एक पाई चार्ट बनाइए।
- (ii) उपरोक्त डेटा को चित्रित करने के लिए आपको कौन सी चित्रमय विधि सबसे अच्छी लगती।
- (स) क्या हम निम्नलिखित दोनों डेटा सेट के लिए मानक विचलन की तुलना कर सकते हैं? बताएं कि आप इन डेटा सेटों की परिवर्तनशीलता कैसे निर्धारित कर सकते हैं?

	A	B	C	D
1	Student	Height (C)	Weight (KG)	
2	1	178	65	
3	2	180	70	
4	3	175	71	
5	4	185	72	
6	5	177	73	
7	6	190	79	

4. Write short notes on any **two** of the following :
- (a) Methods of data collection
- (b) R Studio
- (c) Covariance (12.5)
- निम्नलिखित में से किन्हीं दो पर संक्षिप्त टिप्पणी लिखिए :
- (अ) डेटा संग्रह के तरीके
- (ब) R स्टूडियो
- (स) सहप्रसरण
5. (a) "A coefficient estimate is statistically significant at the 10% significance level". Explain. (4)
- (b) What is NPV? What function would be required to calculate NPV in Excel. (4.5)

(c) Write significance of Boxplot in Excel. (4)

(अ) "एक गुणांक अनुमान 10% महत्व स्तर पर सांख्यिकीय रूप से महत्वपूर्ण है"। समझाइये।

(ब) NPV क्या है ? एक्सेल में NPV की गणना करने के लिए किस फंक्शन की आवश्यकता होगी।

(स) एक्सेल में बॉक्सप्लॉट का महत्व लिखिए।

6. (a) Write the syntax in MS Excel for finding Kurtosis and also determine the type of Kurtosis of the following data set. (4)

X	22	26	62	54	34	40	28
---	----	----	----	----	----	----	----

(b) Explain the importance of Line diagram and Histogram. What are the advantages and disadvantages of using them in data analysis? (4)

(c) The following table exhibits the relation between change in temperature and sale of ice-cream. Build a linear regression equation from given ANOVA table below. (4.5)

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	8.050741	4.691719	1.715947	0.161314	-4.97556	21.07704	-4.97556	21.0770
sale of ice-cream	0.000408	8.6E-05	4.749017	0.008977	0.00017	0.000647	0.00017	0.00064

(अ) Kurtosis खोजने के लिए एम एस एक्सेल में सिटैक्स लिखें और निम्नलिखित डेटा सेट के Kurtosis के प्रकार को भी निर्धारित करें।

X	22	26	62	54	34	40	28
---	----	----	----	----	----	----	----

(ब) रेखा आरेख और हिस्टोग्राम के महत्व की व्याख्या करें। डेटा विश्लेषण में उनका उपयोग करने के क्या फायदे और नुकसान हैं?

(स) निम्न तालिका तापमान में परिवर्तन और आइसक्रीम की बिक्री के बीच संबंध दर्शाती है। नीचे दी गई एनोवा तालिका से एक रैखिक समाश्रयण समीकरण बनाइए।

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	8.050741	4.691719	1.715947	0.161314	-4.97556	21.07704	-4.97556	21.0770
sale of ice-cream	0.000408	8.6E-05	4.749017	0.008977	0.00017	0.000647	0.00017	0.00064

7. (a) The file Sedans contains the overall miles per gallon (MPG) of 2021 midsize sedans :

32	46	16	30	34	56	48	28	44	36	54	20
69	22	33	22	14	38	37	32	30	45	33	25

(i) Compute mean and median.

(ii) Compute the variance, standard deviation, range, coefficient of variation. (6.5)

- (b) What do you mean by hypothesis? Is it necessary to formulate hypothesis in every research study in social sciences? (6)

- (अ) फाइल सेडान में 2021 मध्यम आकार की सेडान की कुल मील प्रति गैलन (एकीपीजी) शामिल है :

32	46	16	30	34	56	48	28	44	36	54	20
69	22	33	22	14	38	37	32	30	45	33	25

- (i) माध्य और माध्यिका की गणना करें।
(ii) विचरण, मानक विचलन, श्रेणी, भिन्नता के गुणांक की गणना करें।
- (ब) परिकल्पना से आप क्या समझते हैं ? क्या सामाजिक विज्ञान के प्रत्येक शोध अध्ययन में परिकल्पना तैयार करना आवश्यक है ?

8. Write short notes on any **two** of the following : (12.5)

- (i) Goodness of Fit R^2
(ii) Equated Monthly Instalment (EMI)
(iii) OLS Regression Method

निम्नलिखित में से किन्हीं दो पर संक्षिप्त टिप्पणी लिखिए :

- (i) फिट R^2 की अच्छाई
(ii) समान मासिक किस्त (ईएमआई)
(iii) OLS रीग्रेशन विधि