No. of Printed Pages : 21 **AST-001**

BACHELOR'S DEGREE PROGRAMME (BDP)

Term-End Examination

June, 2022

AST-001 : STATISTICAL TECHNIQUES

Time : 2 Hours *Maximum Marks : 50*

Note : *(i) Question **No. 1** is compulsory.*

*(ii) Attempt any **four** questions from question nos. **2** to **7**.*

(iii) Calculator is not allowed.

---

1. State whether the following statements are true *or* false. Give brief justification for your answers : 10

   (a) If variance of X is 3 and Y = 5X – 3, then variance of Y is 12.

   (b) If 10X – Y + 5 = 0 is the line of regression of X on Y, then for X = 3, the estimated value 35 of Y is the best estimated value.

(c) Systematic sampling is suitable for studying a population with linear trend.

(d) If the probability of being left-handed is 0.1, the probability that none of three persons selected randomly is left-handed is 0.729.

(e) The sum of deviations of a set of n values from their arithmetic mean is n.

2. (a) The following data gives the frequencies of serum cholesterol level of 1000 males aged between 25 and 35 years arrived at a particular city hospital during the last one year :

Cholesterol level	**Number of Males**
80—120	12
120—160	145
160—200	380
200—240	292
240—280	118
280—320	35
320—360	11
360—400	7

Calculate the median for the data. Is median as an average best suited for the above data ? Justify. 3

(b) To determine the yield rate of rice in a district, 5 groups were constructed of 5 clusters each. The data are given in the following table :

Cluster No.	**Group 1**	**Group 2**	**Group 3**	**Group 4**	**Group 5**
1	10	8	17	14	15
2	12	5	15	8	12
3	9	12	20	6	17
4	20	18	18	10	8
5	8	6	25	16	10

Find the sample mean along with its standard error using cluster sampling. 4

(c) The mean and standard deviation of a characteristic of 100 items were found to be 60 and 10, respectively. At the time of calculations, one item was wrongly taken as 5 instead of 20. Calculate the correct mean and standard deviation. 3

3. (a) A company is interested in forecasting the demand for one of its products. The data on demand for the last 12 months are given below :

Month	**Demand (in 100 units)**
1	15
2	14
3	16
4	17
5	15
6	18
7	20
8	22
9	23
10	21
11	24
12	26

Forecast the demand for the 13th month using exponential smoothing techniques for $w = 0.2$. Also plot the obtained values. 5

(b) Find the line of best fit y on x for the following data. Also plot scatter diagram : 5

x	y
2	3.0
3	5.0
4	5.5
5	6.0
6	8.0
7	9.5

4. (a) A bag contains 4 red and 3 blue balls. Two draws of two balls each are made. Find the chance that the first draw gives 2 red balls and the second draw 2 blue balls, if (i) the balls are returned to the bag after the first draw; (ii) the balls are not returned. 3

(b) Find the probability that the most 5 defective fuses will be found in a box of 200 fuses if experience shows that 2% of such fuses are defective. 4

(c) Define sampling and explain what advantages it has over the census approach. Explain with a situation. 3

5. (a) In the amount of cosmic radiations to which a person is exposed while flying across a specific continent is a normal random variable with mean 4.35 units and S. D. 0.59 units. Find the probabilities that the amount of exposure during such a flight is (i) between 4.00 and 5.00 units, (ii) at least 5.50 units. 4

(b) A simple random sample of size 100 has mean 15 and population variance 25. Find an interval estimate of the population mean with a confidence level of 99% and 95%. 4

(c) A hypothetical population consists of the numbers 2, 5 and 7. Write all possible simple random samples of size 2 (with replacement). Verify that the sample mean is an unbiased estimator of the population mean. 2

6. (a) The following random samples are measurements of the heat producing capacity in millions of calories per ton of specimens of coal from mines :

Mine I	**Mine II**
8260	7950
8130	7890
8350	7900
8070	8140
8340	7920
	7840

Test at 5% level of significance whether the difference between the means of two samples is significant ? (You may like to use the values given at the end of the question paper.) 4

(b) A company operates four machines on three separate shifts daily. The following table presents the data for machine break-downs resulted during a 6-month time period :

Shift	Machine				Total
	A	B	C	D	
1	10	12	6	7	35
2	10	24	9	10	53
3	13	20	7	10	50
Total	33	56	22	27	138

At 5% level of significance test the hypothesis that for an arbitrary breakdown in the machine causing the breakdown and the shifts are independent. (You may like to use the values given at the end of the question paper.) 6

7. (a) The three drying techniques for curing a glue were studied and the following times were observed :

Formula A	**Formula B**	**Formula C**
13	13	4
10	11	1
8	14	3
11	14	4
8		2
		8

At $\alpha = 0.05$, test the hypothesis that the average times for three formulae are same. (You may like to use the values given at the end of the question paper.) 6

(b) The following data represent the number of defects discovered at a factory on 20 successive batches of 10 cars each :

Batch No.	No. of Defects
1	143
2	160
3	150
4	105
5	98
6	74
7	85
8	95
9	76
10	68
11	63
12	87
13	90
14	81
15	94
16	68
17	95
18	90
19	93
20	73

Does it appear that the production process is in statistical control throughout ? Give justification. 4

Some values of use, if required :

Z-values	t-values
P $(0 < Z < 1.10) = 0.3643$	$t_{1,0.05} = 2.26$
P $(0 < Z < 0.59) = 0.2240$	$t_{10,0.05} = 1.812$
P $(0 < Z < 1.95) = 0.4744$	$t_{12,0.05} = 1.782$

F-values	χ^2-values
$F_{2,\ 12,\ 0.05} = 3.89$	$\chi^2_{0.05,6} = 12.59$
$F_{3,\ 12,\ 0.05} = 3.49$	$\chi^2_{0.05,12} = 16.919$
$F_{4,\ 12,\ 0.05} = 3.26$	$\chi^2_{0.05,9} = 21.026$

AST-001

स्नातक उपाधि कार्यक्रम (बी. डी. पी.)

सत्रांत परीक्षा

जून, 2022

ए. एस. टी.-001 : सांख्यिकीय तकनीकें

समय : 2 घण्टे *अधिकतम अंक : 50*

नोट : *(i) प्र. सं.* ***1 अनिवार्य*** *है।*

(ii) प्र. सं. ***2*** *से* ***7*** *तक किन्हीं* **चार** *प्रश्नों के उत्तर दीजिए।*

(iii) कैलकुलेटर का प्रयोग करने की अनुमति नहीं है।

1. बताइए कि निम्नलिखित कथन सत्य हैं **या** असत्य। अपने उत्तरों के पक्ष में संक्षिप्त पुष्टियाँ दीजिए : 10

   (क) यदि X का प्रसरण 3 और $Y = 5X - 3$ है, तब Y का प्रसरण 12 है।

   (ख) यदि $10X - Y + 5 = 0$, Y पर X का रेखा समाश्रयण है, तब $X = 3$ के लिए Y का अनुमानित मान 35 उत्तम अनुमानित मान है।

   (ग) रैखिक प्रवृत्ति वाली समष्टि का अध्ययन करने के लिए क्रमबद्ध प्रतिचयन उपयुक्त होता है।

(घ) यदि खब्बू (बाएँ हाथ का) होने की प्रायिकता 0.1 है, तब यादृच्छया चुने गए तीन व्यक्तियों में किसी के भी खब्बू न होने की प्रायिकता 0.729 होगी।

(ङ) n मानों के समुच्चय के विचलनों का योग उनके समांतर माध्य से n है।

2. (क) निम्नलिखित आँकड़े किसी नगर के अस्पताल में आने वाले 25 और 35 वर्ष की आयु के 1000 पुरुषों के सीरम कॉलेस्ट्रॉल की बारंबारताओं को दर्शाते हैं :

कॉलेस्ट्रॉल स्तर	**पुरुषों की संख्या**
80—120	12
120—160	145
160—200	380
200—240	292
240—280	118
280—320	35
320—360	11
360—400	7

आँकड़ों की माध्यिका परिकलित कीजिए। क्या औसत के रूप में माध्यिका उपर्युक्त आँकड़े के लिए सबसे अधिक उपयुक्त है ? पुष्टि कीजिए। 3

(ख) एक जिले में चावल की पैदावार दर निर्धारित करने के लिए 5–5 गुच्छ के पाँच समूह बनाए गए। इसके आँकड़े नीचे तालिका में दिए गए हैं :

गुच्छ संख्या	**समूह 1**	**समूह 2**	**समूह 3**	**समूह 4**	**समूह 5**
1	10	8	17	14	15
2	12	5	15	8	12
3	9	12	20	6	17
4	20	18	18	10	8
5	8	6	25	16	10

गुच्छ प्रतिचयन का प्रयोग करके प्रतिदर्श माध्य और इसकी मानक त्रुटि ज्ञात कीजिए। 4

(ग) 100 वस्तुओं की विशेषताओं का माध्य और मानक विचलन क्रमशः 60 और 10 पाए गए। गणना के दौरान एक वस्तु को गलती से 5 के स्थान पर 20 ल लिया गया। सहो माध्य और मानक विचलन परिकलित कीजिए। 3

3. (क) एक कम्पनी अपने एक उत्पाद की माँग का पूर्वानुमान लगाना चाहती है। पिछले बारह महीनों की माँग के आँकड़े नीचे दिए गए हैं :

महीना	**माँग (100 इकाइयों में)**
1	15
2	14
3	16
4	17
5	15
6	18
7	20
8	22
9	23
10	21
11	24
12	26

$w = 0.2$ के लिए चरघातांकी निष्कोणन तकनीकों का प्रयोग करके 13वें महीने के लिए माँग का पूर्वानुमान लगाइए। प्राप्त मूल्यों को भो आरेखित कीजिए। 5

(ख) निम्नलिखित आँकड़े के लिए x पर श्रेष्ठतम फिट रेखा y ज्ञात कीजिए। प्रकीर्ण चित्र भी आरेखित कीजिए : 5

x	y
2	3.0
3	5.0
4	5.5
5	6.0
6	8.0
7	9.5

4. (क) एक थैले में 4 लाल और 3 नीली गेंदें हैं। दो बार दो-दो गेंद निकाली गईं। इसकी संभावना ज्ञात कीजिए कि पहली बार में 2 लाल गेंदें और दूसरी बार में 2 नीली गेंदें निकलीं, यदि (i) पहली बार गेंदें निकालने के बाद गंदें वापस थैले में डाल दी गईं, (ii) गेंदें वापस थैले में नहीं डाली गईं। 3

(ख) इसकी प्रायिकता ज्ञात कीजिए 200 फ्यूजों के बॉक्स में अधिकांश 5 खराब फ्यूज पाए जाएँगे यदि अनुभव दर्शाते हैं कि ऐसे 2% फ्यूज खराब हैं। 4

(ग) प्रतिचयन को परिभाषित कीजिए और जनगणना उपागम में इसके लाभों का वर्णन कीजिए। एक स्थिति द्वारा उत्तर को स्पष्ट कीजिए। 3

5. (क) एक विशिष्ट महाद्वीप को पार करते हुए, एक व्यक्ति ब्रह्माण्डोय विकिरण के सपंर्क में आता है जिसकी मात्रा माध्य 4.35 इकाइयों आर मानक विचलन 0.59 इकाइयों वाले प्रसामान्य यादृच्छिक चर है। ऐसी उड़ान के लिए संपर्क में आने वाली मात्रा (i) 4.00 और 5.00 इकाइयाँ, (ii) कम से कम 5.50 इकाइयाँ होंगी, इसकी प्रायिकताएँ ज्ञात कीजिए। 4

(ख) आमाप 100 क एक सरल यादृच्छिक प्रतिदर्श का माध्य 15 और प्रतिदर्श प्रसरण 25 है। 99% और 95% विश्वास्यता अंतराल वाले समष्टि माध्य का अन्तराल आकलन ज्ञात कीजिए। 4

(ग) एक काल्पनिक समष्टि में 2, 5 और 7 संख्याएँ आती हैं। आमाप 2 के सभी संभावित सरल यादृच्छिक प्रतिदर्श (बिना प्रतिस्थापन के) लिखिए। सत्यापित कीजिए कि प्रतिदर्श माध्य समष्टि माध्य का अनभिनत आकलक है। 2

6. (क) निम्नलिखित यादृच्छिक प्रतिदर्श दो खानों के कोयले के प्रति टन कैलोरी के नमूनों क मिलियन में ऊष्मा उत्पन्न करने की क्षमता के माप हैं :

खान I	**खान II**
8260	7950
8130	7890
8350	7900
8070	8140
8340	7920
	7840

5% सार्थकता स्तर पर परीक्षण कीजिए कि क्या दो प्रतिदर्शों के माध्यों के बीच अंतर सार्थक है ? (आप प्रश्नपत्र के अंत में दिए गए मानों का प्रयोग कर सकते हैं।) 4

(ख) एक कम्पनी प्रतिदिन अलग-अलग पालियों (शिफ्टों) में चार मशीनें चलाती है। अग्रलिखित तालिका 6 माह की अवधि में मशीन के खराब होने का आँकड़ा प्रस्तुत करती है :

शिफ्ट	मशीन				कुल
	A	B	C	D	
1	10	12	6	7	35
2	10	24	9	10	53
3	13	20	7	10	50
कुल	33	56	22	27	138

5% सार्थकता स्तर पर परिकल्पना परीक्षण कोजिए कि मशीन में मनमानी खराबी के कारण मशीन का खराब होना और शिफ्टें परस्पर स्वतंत्र हैं। (आप प्रश्नपत्र के अंत में दिए गए मानों का प्रयोग कर सकते हैं।) 6

7. (क) गोंद को सुरक्षित रखने की तीन सुखाने की तकनीकों के अध्ययन किए और निम्नलिखित समय प्रेक्षित किए गए :

फार्मूला A	**फार्मूला B**	**फार्मूला C**
13	13	4
10	11	1
8	14	3
11	14	4
8		2
		8

$\alpha = 0.05$ पर परिकल्पना परीक्षण कीजिए कि तीन फॉर्मूलों का औसत समय समान है। (आप प्रश्न पत्र के अंत में दिए गए मानों का प्रयोग कर सकते हैं।) 6

(ख) निम्नलिखित आँकड़े प्रत्येक 10 कारों के 20 क्रमिक बैचों पर फैक्टरी द्वारा पाए गए दोषों की संख्या को निरूपित करत हैं :

बैच सं.	**दोषों की संख्या**
1	143

2	160
3	150
4	105
5	98
6	74
7	85
8	95
9	76
10	68
11	63
12	87
13	90
14	81
15	94
16	68
17	95
18	90
19	93
20	73

क्या ऐसा प्रतीत होता है कि पूरी प्रक्रिया के दौरान उत्पादन पक्रिया सांख्यिकीय नियंत्रण में है ? पुष्टि कीजिए। 4

यदि आवश्यक हो तो कुछ मानों का प्रयोग कर सकते हैं :

Z-मान	t-मान
P (0 < Z < 1.10) = 0.3643	$t_{1,0.05} = 2.26$
P (0 < Z < 0.59) = 0.2240	$t_{10,0.05} = 1.812$
P (0 < Z < 1.95) = 0.4744	$t_{12,0.05} = 1.782$

F-मान	χ^2-मान
$F_{2, 12, 0.05} = 3.89$	$\chi^2_{0.05,6} = 12.59$
$F_{3, 12, 0.05} = 3.49$	$\chi^2_{0.05,12} = 16.919$
$F_{4, 12, 0.05} = 3.26$	$\chi^2_{0.05,9} = 21.026$