

**BACHELOR'S DEGREE PROGRAMME
(BDP)**

Term-End Examination

December, 2022

AST-01 : STATISTICAL TECHNIQUES

Time : 2 hours

Maximum Marks : 50

(Weightage : 70%)

Note : *Question No. 7 is **compulsory**. Attempt any **four** questions from Questions No. 1 to 6. Use of calculators are **not** allowed. You may want to use some values given at the end.*

1. (a) Before admission in Ph.D. Programme of Statistics in IGNOU, the students have to take Basic Skills Test in Fundamentals of Statistics. In one such exam, 30 students appeared in the test. Their scores are recorded below, out of a total maximum of 30 points :

15, 12, 15, 22, 28, 30, 19, 25, 24, 28, 10, 15,
16, 20, 26, 22, 18, 20, 27, 14, 12, 19, 21, 18,
19, 30, 13, 10, 21, 24.

- (i) What class interval size would you decide and why ?
- (ii) Construct a frequency distribution for the above data with suitable class interval and number of classes.
- (iii) Also construct 'less than' and 'more than' cumulative frequency distribution. 4

- (b) With 10 observations, each on two variables X and Y, the following data were observed :
 $\bar{X} = 12$, $\sigma_X = 3$, $\bar{Y} = 15$, $\sigma_Y = 4$ and $r = 0.5$
However, on subsequent verification, it was found that one value of X (= 15) and one value of Y (= 13) were wrongly taken as 16 and 18, respectively. Find the correct value of correlation coefficient. 6

2. (a) The mean annual salary of all employees in a company is ₹ 25,000. The mean salary of male and female employees is ₹ 27,000 and ₹ 17,000, respectively. Find the percentage of males and females employed by the company. 3

- (b) For the given data in the following table :

X :	5	4	3	2	1
Y :	9	8	10	11	12

Find :

- (i) Both regression coefficients,
- (ii) Correlation coefficient,
- (iii) Equation of regression line of Y on X, and
- (iv) Estimate Y for $X = 4.5$. 7

3. (a) The following data represent the number of days absent per year in a population of five employees of a small company :

25, 20, 30, 15, 10

- (i) How many samples of size 2 are possible without replacement ? Write them.
- (ii) Construct the sampling distribution of the sample mean.
- (iii) Compute mean and standard error of the sampling distribution of the sample mean. 5

- (b) A new computer software package was developed to reduce the project completion time. To evaluate the benefit of the new software package, the following data were collected :

Time for Current Software (in hours)	Time for New Software (in hours)
300	274
280	220
344	308
385	336
372	198
360	300
288	315
321	258
376	318
290	310
301	332
283	263

It is assumed that the project completion times of both software are normally distributed with equal variances. Test whether the average project completion times of both software differ significantly at 5% level of significance.

5

[You may like to use the following values :

$$t_{22, 0.05} = 2.074, t_{20, 0.05} = 2.086,$$

$$t_{22, 0.025} = 1.717]$$

4. (a) The branch manager of a bank wanted to study the waiting times of customers for issuing the demand draft during the peak hours (11:00 am to 1:30 pm). Random samples of 4 customers were selected for 10 days and time (in minutes) was measured, when each customer earned in the queue, when she began to be served. The results are given below :

Sample No.	Time			
	X_1	X_2	X_3	X_4
1	18	16	16	15
2	18	16	20	18
3	15	18	19	12
4	18	18	20	15
5	16	16	18	16
6	20	15	16	16
7	20	25	19	20
8	16	16	17	16
9	18	18	18	17
10	15	16	17	18

Construct suitable control charts for variability as well as average time, and draw the conclusions about the process. 7

[You may use the following constants :

$$A_2 = 0.729, D_3 = 0, D_4 = 2.282]$$

- (b) The temperature of the coffee sold at 16 coffee shops was measured. If the sample mean and standard deviation of the temperature are respectively 162°F and 10°F , construct a 95% confidence interval for the population mean temperature of coffee sold. Assume that temperatures are approximately normally distributed.

3

5. (a) There are 300 misprints randomly distributed throughout a book of 500 pages. Find the probability that a given page contains :

(i) exactly 2 misprints.

(ii) 2 or more misprints.

3

- (b) A research team of a firm has randomly collected data related to the income and age of 726 employees who have quit their jobs and recorded in the following table :

Age Group	Income Category		
	I	II	III
Young	50	69	89
Middle Aged	67	78	102
Old	78	70	103

Determine whether income depends on age group of employees at 5% level of significance.

7

[You may use : $\chi_{4, 0.05}^2 = 9.49$, $\chi_{3, 0.05}^2 = 7.81$]

6. (a) In order to test whether there is any significant difference in the average number of units produced per week by each of three production methods, the following data were collected :

Method I	Method II	Method III
18	19	17
16	18	18
17	19	19
18	17	19
17		16

Use $\alpha = 0.05$ to test the difference among the units produced by the three methods. 7

[You may use, $F_{2, 11, 0.05} = 3.98$, $F_{2, 9, 0.05} = 4.24$]

- (b) The data of heights (in cm) of 8 students from each of five different classes are given as follows :

Student	Class 1	Class 2	Class 3	Class 4	Class 5
1	146	161	167	178	180
2	153	185	163	164	175
3	173	164	167	168	173
4	162	170	172	160	161
5	165	176	158	149	155
6	165	152	159	160	160
7	167	173	148	154	150
8	156	169	180	163	153

How many first stage samples of size 2 and second stage samples of size 6 can be selected ? Select one sample. 3

7. Which of the following statements are *True* and which are *False* ? Justify your answers. $5 \times 2 = 10$

- (a) If arithmetic mean of the numbers 3.2, 5.8, 7.9 and 4.5 with their corresponding frequencies Y , $(Y + 2)$, $(Y - 3)$ and $(Y + 6)$ respectively, is 4.876, then the value of Y is 5.
- (b) If the sum of the product of deviations of X and Y values from their respective means is zero, then the coefficient of correlation between X and Y will be negative.
- (c) Given that A , B and C are three mutually exclusive and exhaustive events and $\frac{1}{3} P(C) = \frac{1}{2} P(A) = P(B)$, then $P(B) = \frac{1}{6}$.
- (d) Rejecting the null hypothesis when it is not true is called type-I error.
- (e) If variance of X is 3 and $Y = 5X - 3$, then variance of Y is 5.
-

स्नातक उपाधि कार्यक्रम
(बी.डी.पी.)
सत्रांत परीक्षा
दिसम्बर, 2022

ए.एस.टी.-01 : सांख्यिकीय तकनीकें

समय : 2 घण्टे

अधिकतम अंक : 50

(कुल का : 70%)

नोट: प्रश्न सं. 7 अनिवार्य है। प्रश्न सं. 1 से 6 में से किन्हीं चार प्रश्नों के उत्तर दीजिए। कैल्कुलेटर का प्रयोग करने की अनुमति नहीं है। आप अंत में दिए गए कुछ मानों का प्रयोग कर सकते हैं।

1. (क) इन्ू के सांख्यिकी के पी एच.डी. कार्यक्रम में प्रवेश लेने से पहले विद्यार्थियों की सांख्यिकी के मूल-सिद्धांतों के आधारभूत कौशल की परीक्षा ली जाती है। ऐसी परीक्षा में, 30 विद्यार्थी बैठे। अधिकतम कुल 30 प्वाइंट्स में से प्राप्त उनके अंक नीचे दिए गए हैं :

15, 12, 15, 22, 28, 30, 19, 25, 24, 28, 10, 15,
16, 20, 26, 22, 18, 20, 27, 14, 12, 19, 21, 18,
19, 30, 13, 10, 21, 24.

- (i) आप क्या वर्ग अन्तराल आमाप निर्धारित करेंगे और क्यों ?
- (ii) उपयुक्त वर्ग अन्तराल और वर्गों की संख्या के साथ ऊपर दिए गए आँकड़ों के लिए बारंबारता बंटन बनाइए ।
- (iii) 'से कम' और 'से ज्यादा' संचयी बारंबारता बंटन भी बनाइए ।

4

(ख) दो चरों X और Y पर 10 प्रेक्षणों के साथ, निम्नलिखित आँकड़े प्रेक्षित किए गए :

$$\bar{X} = 12, \sigma_X = 3, \bar{Y} = 15, \sigma_Y = 4 \text{ और } r = 0.5$$

तथापि बाद के सत्यापन में यह पाया गया कि X (= 15) का एक मान और Y (= 13) का एक मान ग़लती से क्रमशः 16 और 18 लिए गए थे । सहसंबंध गुणांक का सही मान ज्ञात कीजिए ।

6

2. (क) एक कम्पनी के सभी कर्मचारियों का माध्य वार्षिक वेतन ₹ 25,000 है । पुरुष और महिला कर्मचारियों का माध्य वेतन क्रमशः ₹ 27,000 और ₹ 17,000 है । कम्पनी में नियुक्त पुरुषों और महिलाओं का प्रतिशत ज्ञात कीजिए ।

3

(ख) नीचे सारणी में दिए गए आँकड़ों के लिए :

X :	5	4	3	2	1
Y :	9	8	10	11	12

ज्ञात कीजिए :

- (i) दोनों समाश्रयण गुणांक
- (ii) सहसंबंध गुणांक
- (iii) X पर Y की समाश्रयण रेखा का समीकरण, और
- (iv) $X = 4.5$ के लिए Y का आकलन ।

7

3. (क) निम्नलिखित आँकड़ा पाँच कर्मचारियों वाली एक छोटी कम्पनी में कर्मचारियों के प्रति वर्ष अनुपस्थित रहने के दिनों की संख्या को निरूपित करता है :

25, 20, 30, 15, 10

- (i) बिना प्रतिस्थापन के आमाप 2 के कितने प्रतिदर्श संभव हैं ? उन्हें लिखिए ।
- (ii) प्रतिदर्श माध्य का प्रतिचयन बंटन बनाइए ।
- (iii) प्रतिदर्श माध्य के प्रतिचयन बंटन का माध्य और मानक त्रुटि परिकलित कीजिए ।

5

- (ख) परियोजना को पूरा करने के समय को कम करने के लिए एक नया कम्प्यूटर सॉफ्टवेयर पैकेज तैयार किया गया ।

नए सॉफ्टवेयर पैकेज के लाभ का मूल्यांकन करने के लिए निम्नलिखित आँकड़े एकत्रित किए गए :

मौजूदा सॉफ्टवेयर के लिए समय (घंटों में)	नए सॉफ्टवेयर के लिए समय (घंटों में)
300	274
280	220
344	308
385	336
372	198
360	300
288	315
321	258
376	318
290	310
301	332
283	263

यह मान लिया गया कि दोनों सॉफ्टवेयर का परियोजना को पूरा करने का समय बराबर प्रसरण वाला प्रसामान्यतः बंटित है। जाँच कीजिए कि क्या 5% सार्थकता-स्तर पर दोनों सॉफ्टवेयर का परियोजना को पूरा करने का औसत समय सार्थक रूप से भिन्न है।

5

[आप निम्नलिखित मानों का प्रयोग कर सकते हैं :

$$t_{22, 0.05} = 2.074, t_{20, 0.05} = 2.086,$$

$$t_{22, 0.025} = 1.717]$$

4. (क) एक बैंक का शाखा प्रबंधक सबसे अधिक व्यस्त घंटों (प्रातः 11:00 बजे से दोपहर 1:30 बजे तक) के दौरान उपभोक्ताओं को डिमांड ड्राफ्ट जारी करने के लिए प्रतीक्षा-काल का अध्ययन करना चाहता है। 10 दिनों के लिए 4 उपभोक्ताओं के यादृच्छिक प्रतिदर्श चुने गए और पंक्ति में सेवा शुरू होने से पहले प्रत्येक उपभोक्ता के लगने वाले समय (मिनटों में) को मापा गया। इसके परिणाम नीचे दिए गए हैं :

प्रतिदर्श सं.	समय			
	X ₁	X ₂	X ₃	X ₄
1	18	16	16	15
2	18	16	20	18
3	15	18	19	12
4	18	18	20	15
5	16	16	18	16
6	20	15	16	16
7	20	25	19	20
8	16	16	17	16
9	18	18	18	17
10	15	16	17	18

चर का बदलाव और औसत समय के लिए उपयुक्त नियंत्रण चार्ट बनाइए और प्रक्रिया के बारे में निष्कर्ष निकालिए।

7

[आप निम्नलिखित अचरों का प्रयोग कर सकते हैं :

$$A_2 = 0.729, D_3 = 0, D_4 = 2.282]$$

(ख) 16 कॉफी दुकानों पर बेचे जाने वाली कॉफी का तापमान मापा गया । यदि तापमान का प्रतिदर्श माध्य और मानक विचलन क्रमशः 162°F और 10°F है, तब बेची गई कॉफी के समष्टि माध्य तापमान के लिए 95% विश्वास्यता अन्तराल बनाइए । मान लीजिए कि तापमान लगभग प्रसामान्यतः बंटित हैं ।

3

5. (क) 500 पृष्ठों की एक पुस्तक में यादृच्छया बंटित 300 ग़लत छपाइयाँ हैं । इसकी प्रायिकता ज्ञात कीजिए कि किसी दिए गए पृष्ठ में

(i) यथातथ दो ग़लत छाप हैं ।

(ii) 2 या अधिक ग़लत छाप हैं ।

3

(ख) एक फर्म के शोध दल ने नौकरी छोड़ चुके 726 कर्मचारियों की आय और आयु से संबंधित यादृच्छिक रूप से आँकड़े एकत्रित किए और उन्हें निम्नलिखित सारणी में रिकॉर्ड किया :

आयु-वर्ग	आय-श्रेणी		
	I	II	III
युवा	50	69	89
मध्य वयस्क	67	78	102
बूढ़ा	78	70	103

5% सार्थकता-स्तर पर निर्धारित कीजिए कि क्या कर्मचारियों का आयु-वर्ग आय पर आधारित है या नहीं ।

7

[आप $\chi_{4, 0.05}^2 = 9.49$, $\chi_{3, 0.05}^2 = 7.81$ का प्रयोग कर सकते हैं]

6. (क) तीन उत्पादन विधियों द्वारा प्रति सप्ताह प्रत्येक विधि से उत्पादित इकाइयों की औसत संख्या में सार्थक अन्तर है या नहीं यह जाँच करने के लिए निम्नलिखित आँकड़े एकत्रित किए गए :

विधि I	विधि II	विधि III
18	19	17
16	18	18
17	19	19
18	17	19
17		16

तीनों विधियों द्वारा उत्पादित इकाइयों के बीच अंतर की जाँच करने के लिए $\alpha = 0.05$ का प्रयोग कीजिए ।

7

[आप $F_{2, 11, 0.05} = 3.98$, $F_{2, 9, 0.05} = 4.24$ का प्रयोग कर सकते हैं]

- (ख) पाँच भिन्न वर्गों के 8 – 8 विद्यार्थियों के कद (सेमी में) के आँकड़े नीचे दिए गए हैं :

विद्यार्थी	वर्ग 1	वर्ग 2	वर्ग 3	वर्ग 4	वर्ग 5
1	146	161	167	178	180
2	153	185	163	164	175
3	173	164	167	168	173
4	162	170	172	160	161
5	165	176	158	149	155
6	165	152	159	160	160
7	167	173	148	154	150
8	156	169	180	163	153

आमाप 2 की प्रथम अवस्था और आमाप 6 की द्वितीय अवस्था के कितने प्रतिदर्श चुने जा सकते हैं ? एक प्रतिदर्श चुनिए ।

3

7. निम्नलिखित में से कौन-से कथन सत्य हैं और कौन-से असत्य हैं ? अपने उत्तरों की पुष्टि कीजिए । 5×2=10

(क) यदि संख्याओं 3·2, 5·8, 7·9 और 4·5 का समांतर माध्य उनकी संगत बारंबारता क्रमशः Y , $(Y + 2)$, $(Y - 3)$ और $(Y + 6)$ के साथ 4·876 है, तब Y का मान 5 है ।

(ख) यदि X और Y मानों का उनके अपने-अपने माध्यों से विचलन के गुणनफल का योगफल शून्य है, तब X और Y के बीच सहसंबंध गुणांक ऋणात्मक होगा ।

(ग) दिया गया है कि A , B और C तीन परस्पर अपवर्जी और निश्शेष घटनाएँ हैं और $\frac{1}{3}P(C) = \frac{1}{2}P(A) = P(B)$ है, तब $P(B) = \frac{1}{6}$ होगा ।

(घ) एक शून्य परिकल्पना को अस्वीकार करना जब वह सत्य नहीं है, टाइप-I त्रुटि कहलाती है ।

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