

BACHELOR'S DEGREE PROGRAMME

• Term-End Examination

December, 2009

(APPLICATION ORIENTED COURSE)

AST-1 : STATISTICAL TECHNIQUES

Time : 2 hours

Maximum Marks : 50

Note : Question No. 7 is *compulsory*. Do any *four* questions from the remaining question nos. 1 to 6. No calculators are allowed.

1. (a) Two floppoies are selected at random 5
without replacement from a box containing
7 good and 3 defective floppies. Let A be
the event that the first floppy drawn is
defective, and let B be the event that the
second floppy drawn is defective.
- (i) Find the conditional probabilities
 $P(B/A)$ and $P(B/A^c)$
- (ii) Show that $P(B) = P(B/A) \cdot P(A) + P(B/A^c) \cdot P(A^c)$
- (b) A company produces one-kilogram sugar 5
packets. The specifications on the net
content are 1000 ± 5 grams. Assuming that
the net content follows normal distribution
with mean weight as 1005 grams and the

process capability equal to 30 grams, find out the proportion of packets that have weight less than lower specification limit. What should be the mean if this proportion is to be reduced to 0.01 ?

[You may like to use some of the values given at the end of the question paper]

2. Two new types of petrol, called premium and super, are introduced in the market, and their manufacturers claim that they give extra mileage. Following data were obtained on extra mileage which is defined as actual mileage minus 10. 10

Data on extra mileage				
Ordinary Petrol	1	2	2	1
Premium Petrol	2	2	1	3
Super Petrol	4	1	2	3

- (i) Using ANOVA, test whether premium or super give an extra mileage.
- (ii) What is your estimate for the error variance ?
- (iii) Assuming that the error variance is known and is equal to 1, obtain the 95% confidence interval for the mean extra mileage of super. (You may like to use some of the values given at the end of the question paper.)

3. (a) The number of total accidents during the first half of 2008 are as given in the table below. 7

Data on Fatal Accidents

Month	Jan	Feb	Mar	Apr	May	Jun
No. of fatal accidents	2	4	3	4	3	2

- (i) Draw the graph and comment
- (ii) Prepare the table of forecasts and errors by applying simple exponential smoothing for the data given in (i) (take the exponential smoothing constant as 0.1).
- (b) Among a group 50 students of class 12, five students are 16 years old, thirty students are 17 years old, five students are 18 years old and ten students are 19 years old. Let a student be drawn from this group and let X be the age of the selected students (in complete years). 3
- (i) Is the distribution of X continuous ?
- (ii) Obtain the probability density function of X .
- (iii) What is the probability that the student selected is atleast 18 years old.

4. (a) An \bar{x} -R chart is being set up for monitoring the twist angle of a certain paper board. Twenty subgroups (each of size 5) of data have been collected and their summary is furnished in the table below. 4

Sample No	1	2	3	4	5	6	7	8	9	10
\bar{x}	-1	-3	1	-3	7	0	4	-2	4	1
R	5	12	10	6	6	6	9	10	13	6
Sample No	11	12	13	14	15	16	17	18	19	20
\bar{x}	-2	4	-2	-1	-3	-1	1	-1	0	0
R	6	25	16	9	4	6	17	11	9	6

- (i) Scrutinize the data (without drawing charts) and check if there are any increasing or decreasing trends in process average.
- (ii) Assuming that there are no systematic patterns in the process, compute the control limits for \bar{x} - R charts.
- (iii) Do you suspect any assignable causes ?
If so, at what sample numbers ?

(You may like to use the values given at the end of the question paper).

- (b) Define stratified sampling and explain the basic principle that forms the strata. 3
- (c) Suppose three small towns are under study, having population $N_1 = 60000$, $N_2 = 30000$, $N_3 = 30000$, respectively. A stratified random sample is to be taken with a total sample size of $n = 500$. Using proportional allocation, determine the sample size of each stratum. 3

5. (a) The frequency distribution of the daily cost (in Rs.) of commuting back and forth to work by 100 employees of a steel plant is : 5

Cost (in Rs)	0-50	50-100	100-150	150-200
Number of employees	29	32	29	10

Find the mean daily cost, median daily cost and standard deviation of commuting.

- (b) A study was conducted to see whether the natural perception having equal number of boys and girls was followed by families or not. 160 families were considered as sample. Each family had 5 children. Frequency of families having a particular combination was recorded as follows. 5

Boy	5	4	3	2	1	0
Girl	0	1	2	3	4	5
Frequency	20	25	25	20	30	40

Test whether the standard norm is followed or not at 95% level of significance.

[You may like to use some of the values given at the end of the question paper].

6. (a) An economist wants to estimate a multiple-regression equation in which the amount saved by the i th family depends on the family's income, whether the family is graduated or non graduate , and whether 3

the family is headed by a male or a female. Explain how a regression equation of this sort can be estimated. What is the dependent variable ? What are the independent variables ? What assumptions must be made ?

- (b) A population of size 20 is sampled without replacement. The standard deviation of the population is 0.35. We require the standard error of the mean to be not more than 0.15. What is the minimum sample size ? 3
- (c) In a population of size $N=5$, the values of an attribute are 8, 3, 1, 11 and 4 corresponding to population members 1 to 5. 4
- (i) Write all the different samples of size 2 when simple random sampling (without replacement) is practised ?
- (ii) Verify that \bar{y} (sample mean) is an unbiased estimate of \bar{Y} (population mean), using the above data.
7. State whether the following statements are true or false. Give brief justification for each case. 10
- (a) Pie chart is more appropriate than bar chart for graphically presenting the profits of five plants of a manufacturing company during the financial year 2007-08.

- (b) If an event A implies another event B, then A contains B.
- (c) In a standard normal distribution, the area under the curve between $-\infty$ and 0 is 1.
- (d) A sampling plan with AQL = 0.08 and LTPD = 0.05 is a good sampling plan.
- (e) The number of possible samples in drawing 3 items from a population of 7 items without replacement is equal to 21.

Some Values for use, if required

Φ values	F-values	χ^2 -values	Statistical Constants for n = 5
$\Phi(-2) = 0.023$	$F_{2,9} = 4.26$	$\chi^2_{5,0.05} = 11.07$	$D_3 = 0$
$\Phi(-2.326) = 0.01$	$F_{2,10} = 4.10$	$\chi^2_{5,0.01} = 15.09$	$D_4 = 2.115$
	$F_{3,9} = 3.86$	$\chi^2_{6,0.05} = 12.59$	$A_2 = 0.577$
	$F_{3,10} = 3.71$	$\chi^2_{6,0.01} = 16.81$	

- o O o -

स्नातक उपाधि कार्यक्रम

सत्रांत परीक्षा

दिसम्बर, 2009

(व्यवहारमूलक पाठ्यक्रम)

ए.एस.टी.-1 : सांख्यिकीय तकनीक

समय : 2 घण्टे

अधिकतम अंक : 50

नोट : प्रश्न सं. 7 अनिवार्य हैं। प्रश्न सं. 1 से 6 तक के शेष 6 प्रश्नों में से कोई चार प्रश्न कीजिए। कैलकुलेटर के प्रयोग की अनुमति नहीं है।

1. (a) एक बॉक्स, जिसमें 7 अच्छी और 3 खराब फ्लोपियाँ हैं। 5
में से यादृच्छिक रूप से बिना प्रतिस्थापित किए दो फ्लोपियाँ निकाली जाती है। मान लीजिए की A वह घटना है जिसमें निकाली गई पहली फ्लोपी खराब होती है तथा B वह घटना है, जिसमें निकाली गई दूसरी फ्लोपी खराब होती है।
 - (i) प्रतिबंधित प्रायिकताएँ $P(B/A)$ और $P(B/A^c)$ ज्ञात कीजिए।
 - (ii) दर्शाइए कि $P(B) = P(B/A), P(A) + P(B/A^c) P(A^c) = P(A)$.
- (b) कोई कंपनी चीनी के एक-एक किलोग्राम वाले पैकट 5 बनाती है। निबल अंतर्वस्तु का विशेष विवरण 1000 ± 5 ग्राम दिया है। यह कल्पना करते हुए कि निबल अंतर्वस्तु माध्य भार 1005 ग्राम के साथ एक प्रसामान्य बंटन का

अनुसरण करती है तथा प्रक्रम क्षमता 30 ग्राम के बराबर है, उन पैकटों का समानुपात ज्ञात कीजिए जिनका भार निचली विशेष विवरण सीमा से कम है। यदि इस समानुपात को कम करके 0.01 लाना है तो माध्य क्या होना चाहिए?

[आप चाहें तो प्रश्न पत्र के अंत में दिए हुए कुछ मानों का प्रयोग कर सकते हैं।]

2. प्रामियम और सुपर कहे जाने वाले दो नए प्रकार के पेट्रोल 10 बाजार में प्रविष्ट किए जाते हैं, तथा उनके निर्माता यह दावा करते हैं की इनसे पहले की अपेक्षा अधिक दूरी तक जाया जा सकता है। इस अधिक दूरी, अर्थात् अतिरिक्त दूरी पर निम्नलिखित आँकड़े प्राप्त होते हैं, जो वास्तविक दूरी से 10 कम के रूप में परिभाषित हैं :

अतिरिक्त दूरी पर आँकड़े				
साधारण पेट्रोल	1	2	2	1
प्रीमियम पेट्रोल	2	2	1	3
सुपर पेट्रोल	4	1	2	3

- ANOVA का उपयोग करते हुए जाँच कीजिए कि क्या प्रीमियम या सुपर पेट्रोल कोई अतिरिक्त दूरी देते हैं।
- त्रुटि प्रसरण के लिए आपका क्या आकलन है?
- यह कल्पना करते हुए कि त्रुटि प्रसरण ज्ञात है और यह 1 के बराबर है, सुपर की माध्य अतिरिक्त दूरी के लिए, 95% विश्वास्थ्यता अंतराल प्राप्त कीजिए।

[आप चाहें तो इस प्रश्न पत्र के अंत में दिए कुछ मानों का प्रयोग कर सकते हैं।]

3. (a) 2008 की पहली छमाही में हुई घातक दुर्घटनाओं की संख्या नीचे दी सारणी में दी गई हैं : 7

घातक दुर्घटनाओं पर आँकड़े

महीना	जनवरी	फरवरी	मार्च	अप्रैल	मई	जून
कुल दुर्घटनाओं की संख्या	2	4	3	4	3	2

- (i) इसका आलेख खींचिए और टिप्पणी कीजिए।
- (ii) उपरोक्त आँकड़ों के लिए सरल चरघातांकी निष्कोणन का अनुप्रयोग करके पूर्वानुमानों और त्रुटियों की सारणी तैयार कीजिए (चरघातांकी निष्कोणन स्थिरांक 0.1 लीजिए)।
- (b) कक्षा 12 के 50 विद्यार्थियों के एक समूह में पाँच विद्यार्थी 16 वर्ष की आयु के हैं, तीस विद्यार्थी 17 वर्ष की आयु के हैं, पाँच विद्यार्थी 18 वर्ष की आयु के हैं तथा दस 19 वर्ष की आयु के हैं। मान लीजिए कि इस समूह में से एक विद्यार्थी को चुना जाता है तथा यह भी मान लीजिए कि इस चुने गए विद्यार्थी की आयु X (पूर्ण वर्षों में) है। 3
- (i) क्या X का बंटन संतत है?
- (ii) X का प्रायिकता घनत्व फलन प्राप्त कीजिए।
- (iii) इसकी क्या प्रायिकता है कि चुने गए विद्यार्थी की आयु कम से कम 18 वर्ष होगी?

4. (a) किसी विशेष कागज के बोर्ड के व्यावर्तन कोण के नियंत्रण के लिए, एक \bar{x} -R चार्ट स्थापित किया जा रहा है। आँकड़ों के बीस उपसमूह (प्रत्येक की माप 5) एकत्रित किए गए हैं तथा इनका सारांश नीचे दी सारणी में वर्णित है :

प्रतिदर्श सं	1	2	3	4	5	6	7	8	9	10
\bar{x}	-1	-3	1	-3	7	0	4	-2	4	1
R	5	12	10	6	6	6	9	10	13	6
प्रतिदर्श सं	11	12	13	14	15	16	17	18	19	20
\bar{x}	-2	4	-2	-1	-3	-1	1	-1	0	0
R	6	25	16	9	4	6	17	11	9	6

- (i) इन आँकड़ों का (बिना चार्ट खींचे हुए) सूक्ष्मता से निरीक्षण कीजिए तथा जाँच कीजिए कि क्या प्रक्रम औसत में कोई वर्धमान या ह्रासमान प्रवृत्तियाँ हैं।
- (ii) यह कल्पना करते हुए कि यहाँ प्रक्रम में कोई क्रमबद्ध प्रतिरूप (पैटर्न) नहीं हैं, \bar{x} -R चार्टों के लिए नियंत्रण सीमाएँ अभिकलित कीजिए।
- (iii) क्या आपको इनमें किन्हीं निर्देश्य कारणों की आशंका है? यदि हाँ, तो किन प्रतिदर्श संख्याओं पर?

[आप चाहे तो इस प्रश्न पत्र के अंत में दिए कुछ मानों का प्रयोग कर सकते हैं।]

- (b) स्तरित प्रतिचयन को परिभाषित कीजिए तथा उस मूल सिद्धांत को स्पष्ट कीजिए जिससे स्तर बनते हैं। 3
- (c) मान लीजिए कि तीन छोटे शहरों का अध्ययन किया जा रहा है, जिनकी जनसंख्या क्रमशः $N_1 = 60000$, 3

$N_2 = 30000$ तथा $N_3 = 30000$ है। कुल प्रतिदर्श साइज (माप) $n = 500$ का एक स्तरित यादृच्छिक प्रतिदर्श लिया जाना है। समानुपातिक बंटन का उपयोग करते हुए, प्रत्येक स्तर का प्रतिदर्श साइज निर्धारित कीजिए।

5. (a) किसी स्टील प्लांट के 100 कर्मचारियों के घर से आने 5
घ जाने में लगने वाली दैनिक लागत (रु.में) का बारंबारता
बंटन यह है :

लागत (रु.में)	0-50	50-100	100-150	150-200
कर्मचारियों की संख्या	29	32	29	10

इस लागत का माध्य दैनिक लागत, माध्यक दैनिक लागत
तथा मानक विचलन ज्ञात कीजिए।

- (b) परिवार का प्राकृतिक रूप से यह मानकर चलना कि 5
परिवार में लड़के लड़कियों की संख्या बराबर होती है।
इस बात की जाँच के लिए एक अध्ययन किया गया है।
प्रतिदर्श में 160 परिवारों को लिया गया। प्रत्येक परिवार
में 5 बच्चे थे। परिवारों में एक विशिष्ट संयोजन इस
प्रकार रिकार्ड किया गया :

लड़का	5	4	3	2	1	0
लड़की	0	1	2	3	4	5
बारंबारता	20	25	25	20	30	40

जाँच कीजिए कि 95% सार्थकता स्तर पर मानक नियमों
का पालन हुआ या नहीं।

[आप इस प्रश्न पत्र के अंत में दिए मानों का प्रयोग कर
सकते हैं।]

6. (a) एक अर्थशास्त्री एक बहु समाश्रयण समीकरण आकलित करता चाहता है, जिसमें i वें परिवार द्वारा की गई बचत परिवार की आय पर निर्भर करती है, चाहे उस परिवार में ग्रेजुएट हैं या नहीं तथा चाहे उस परिवार का मुखिया पुरुष है या महिला। स्पष्ट कीजिए कि इस प्रकार की समाश्रयण समीकरण किस प्रकार आकलित की जा सकती है। इसमें आश्रित चर क्या है? स्वतंत्र चर क्या-क्या हैं? कौन-सी परिकल्पनाएँ करनी चाहिए? 3
- (b) आमाप 20 की एक समष्टि का बिना प्रतिस्थापन के प्रतिचयन किया जाता है। इस समष्टि का मानक विचलन 0.35 है। हम चाहते हैं कि माध्य की मानक त्रुटि 0.15 से अधिक न रहे। न्यूनतम प्रतिदर्श आमाप क्या है? 3
- (c) आमाप $N=5$ वाली एक समष्टि में, समष्टि सदस्यों 1 से 5 के संगत किसी गुण के मान 8, 3, 1, 11 और 4 हैं। 4
- (i) आमाप 2 के सभी भिन्न-भिन्न प्रतिदर्श लिखिए, जब कि सरल यादृच्छिक प्रतिचयन (बिना प्रतिस्थापन के) का प्रयोग किया जाता है।
- (ii) उपरोक्त का प्रयोग करके, सत्यापित कीजिए कि \bar{y} (प्रतिदर्श माध्य), \bar{y} (समष्टि माध्य) का एक अनगिनत आकलन है।
7. बताइए कि निम्नलिखित कथन सत्य हैं या असत्य। प्रत्येक स्थिति के लिए एक संक्षिप्त औचित्य दीजिए। 10
- (a) एक निर्माण कंपनी के पाँच प्लांटों द्वारा वित्तीय वर्ष 2007-08 के अंतराल में अर्जित किए गए लाभों को आलेखीय रूप से प्रस्तुत करने के लिए, पाइ चार्ट दंड चार्ट की तुलना में अधिक उपयुक्त है।
- (b) यदि एक घटना A होने से एक अन्य घटना B अंतर्निहित हो, तो A, B को अंतर्विष्ट करती है।

- (c) एक मानक प्रसामान्य बंटन में, $-\infty$ और 0 के बीच वक्र के अंतर्गत क्षेत्रफल 1 होता है।
- (d) AQL=0.08 और LTPD=0.05 वाली एक प्रतिदर्श योजना एक अच्छी प्रतिदर्श योजना है।
- (e) 7 वस्तुओं वाली एक समष्टि में से, बिना प्रतिस्थापन के, 3 वस्तुएँ निकालने पर संभव प्रतिदर्शों की संख्या 21 के बराबर होती है।

यदि आवश्यक हो, तो प्रयोग करने के लिए कुछ मान :

Φ मान	F-मान	χ^2 -मान	n=5 के लिए संख्याकीय स्थिरांक
$\Phi(-2)=0.023$	$F_{2,9}=4.26$	$\chi^2_{5,0.05}=11.07$	$D_3=0$
$\Phi(-2.326)=0.01$	$F_{2,10}=4.10$	$\chi^2_{5,0.01}=15.09$	$D_4=2.115$
	$F_{3,9}=3.86$	$\chi^2_{6,0.05}=12.59$	$A_2=0.577$
	$F_{3,10}=3.71$	$\chi^2_{6,0.01}=16.81$	

- o O o -