No. of Printed Pages : 15

AST-1

BACHELOR'S DEGREE PROGRAMME

01822

Term-End Examination

December, 2011

(APPLICATION ORIENTED COURSE)

AST-1 : STATISTICAL TECHNIQUES

Time : 2 hours *Maximum Marks : 50*

Note : *Question No. 7 is* ***compulsory****. Answer any* ***four*** *from the remaining questions* **1** *to* **6***.* ***No*** *Calculators are allowed.*

---

1. (a) Cancer is present in 22% of a population and is not present in the remaining 78%. An imperfect clinical test successfully detects the disease and with probability 0.70. Thus if a person has the disease in the serious form, the probability is 0.70 that the test will be positive and it is 0.30 if the test is negative. Moreover among the unaffected persons, the probability that the test will be positive is 0.05. **4+2=6**

(i) A person selected at random from the population is given the test and the result is positive. What is the probability that this person has cancer ?

(ii) What is the probability that the test correctly detects cancer ?

(b) Identify the population and parameter of interest in each of the following. In each case check whether the population is finite or infinite. **4**

(i) An oil company wants to know which of its service stations are working for 24 hours.

(ii) An agricultural firm wants to study the yield of mango during the year 2006-2010.

(iii) An agency wants of find out the rate of literacy in U.P. from 2007 onwards.

(iv) A firm wants to know the mean income of families in a particular locality.

2. (a) In an agricultural experiment to determine the effects of a particular insecticide, a field was planted with corn. Half of the plants were sprayed with the insecticide, and half were not sprayed with it. Several weeks later, a random sample of 200 sprayed plants was selected and a random sample of 200 unsprayed plants was selected. The number of healthy plants in each sample was as follows : **4**

	Sprayed	Unsprayed
Healthy	121	109
Unhealthy	79	91
	200	200

If the significance level is set at 0.05, does the evidence indicate that a higher proportion of sprayed than of unsprayed plants was healthy? You may like to use the values given at the end of the question paper.

(b) The total scores X obtained by 50 students in psychology test of 100 marks are given below : **1+3+1+1=6**

80	90	85	55	40	50	40	95	81	82
75	70	80	40	10	55	30	80	82	72
27	95	20	30	85	10	40	23	75	60
50	30	10	10	90	15	20	40	70	65
45	50	40	10	90	20	25	42	70	70

(i) Is the random variable X = score of a student, discrete or continuous ? What are the minimum and maximum scores ?

(ii) Draw a frequency distribution.

(iii) What percentage of students scored above the pass marks of 50 ?

(iv) How many of the students score between 50 and 80 ?

3. (a) Using a sample data on height (h) and weight (w) of 25 persons , the best linear regression fit was found to be w = 9+0.4h (height in centimetres and weight in Kgs). The sample mean and variance of heights are 165 and 200; and those of weights are 75 and 40 respectively. **6**

(i) Find the best linear fit of height based on weight for the above data.

(ii) What is your prediction for a person's height whose weight is 68 kgs ?

(b) The check up times of 200 patients at an eye clinic had a mean of 10 minutes and standard deviation of 2 minutes. Find the 95 % confidence interval for the mean check up time. What is the 95% lower confidence limit for the mean check up time ? From this information can you say that the check up time of almost every patient is less than 16 minutes ? 4

4. (a) A chemical firm wants to determine how four catalysts differ in yield. The firm runs the experiment in eight of its plants type. In each plant , the yield is measured with each catalyst. The yield (in quintals) are as follows : 7

Plant	Catalyst			
	1	2	3	4
A	2	1	2	4
B	3	2	1	3
C	1	3	3	1
D	5	4	3	2
E	1	2	4	3
F	2	1	1	2
G	4	3	4	2
H	3	1	4	3

Perform an ANOVA and comment whether yield due to a particular catalyst is significant or not at 5% level of significance. [You may like to use the values given at the end of the question paper.]

(b) An incomplete frequency distribution is given as follows : 3

C.I.	Frequency
10-20	12
20-30	30
30-40	?
40-50	65
50-60	?
60-70	25
70-80	18

Given that the median value of 200 observations is 46 , determine the missing frequencies using the median formula.

5. (a) The sales figure of a textile company is given below. Compute the moving averages for the length of 4. 5

Day	Sales
1	230
2	200
3	250
4	300
5	200
6	225
7	400
8	450
9	415
10	420
11	500
12	300
13	400
14	300
15	315

(b) A die is roled 1200 times with the following results : 5

No. that comes up	1	2	3	4	5	6
Frequency	195	289	202	242	163	109

Test if the die is unbiased at 5% level of significance. [you may like to use the values given at end.]

6. (a) The mean weekly sales of soap bars in different departmental stores was 146.3 bars per store. After an advertising campaign the mean weekly sales in 22 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was the advertising campaign successful at 5% level of significance. [you may like to use the values given at end.] **3**

(b) A Beer Company wants to determine the mean weight of a can of its beer. It takes a random sample of 80 such cans (from several thousand cans in its warehouse) and finds that the mean weight is 31.15 ounces and the standard deviation is 0.08 ounce. **2+1=3**

(i) Compute a 95% confidence interval for the mean weight of the cans in the firm warehouse.

(ii) Is your answer to (i) based on the assumption that the weights of the cans of beer in the warehouse are normally distributed ? Why or why not ?

(c) Cite two situations where stratified sampling is appropriate. Describe two advantages of stratified samplings. **4**

7. State whether the following statements are correct or not. Give brief justification. **10**

(a) Mean and standard deviation of Poisson distribution are equal.

(b) If A and B are disjoint events with positive probabilities , then they are independent.

(c) ANOVA is a tool for testing equality of variances.

(d) If LCL and UCL are the control limits on an $\bar{x}$ - chart for a process, then (UCL-LCL)/ 6 is an estimate of the process standard deviation.

(e) Population is stratified based on the different variation of the segments constituting the population.

Some values of use, if required			
χ^2 value	F - values at 5% level of significance	t - values	normal values
$\chi^2_{5, 0.05} = 11.07$	$F_{3, 7} = 4.35$	$t_{21, 0.05} = 721$	$P(Z \leq -0.02) = 0.492$
$\chi^2_{6, 0.05} = 12.59$	$F_{3, 21} = 3.07$	$t_{22, 0.05} = 1.717$	$P(Z \leq 0.02) = 0.508$
$\chi^2_{4, 0.05} = 9.48$	$F_{7, 21} = 2.49$	$t_{11, 0.05} = 1.796$	

स्नातक उपाधि कार्यक्रम

सत्रांत परीक्षा

दिसंबर, 2011

(व्यवहारमूलक पाठ्यक्रम)

ए.एस.टी.-1 : सांख्यिकीय तकनीकें

समय : 2 घण्टे *अधिकतम अंक : 50*

नोट : *प्रश्न संख्या 7 करना जरूरी है। प्रश्न संख्या 1 से 6 में से किन्हीं* ***चार*** *प्रश्न कीजिए। कैलकुलेटरों का प्रयोग करने की अनुमति* ***नहीं*** *हैं।*

---

1. (a) 22% समष्टि को कैंसर है और शेष 78% को कैंसर नहीं है। एक अपूर्ण क्लिनिकल जाँच सफलतापूर्वक रोग का पता लगाती है और इसकी प्रायिकता 0.70 है। इस प्रकार, यदि व्यक्ति के रोग का रूप गंभीर है तो टेस्ट पॉजिटिव होगा इसकी प्रायिकता 0.70 है और टेस्ट नेगेटिव है इसकी प्रायिकता 0.30 है। इसके अलावा, अप्रभावित व्यक्तियों में टेस्ट के पॉजिटिव होने की प्रायिकता 0.05 है। **4+2=6**

(i) समष्टि से एक व्यक्ति को यादृच्छया चुनकर जाँच की जाती है और परिणाम पॉजिटिव पाया जाता है। इसकी क्या प्रायिकता है कि व्यक्ति को कैंसर है ?

(ii) इसकी क्या प्रायिकता है कि टेस्ट से कैंसर का सही रूप से पता चल जाता है ?

(b) निम्नलिखित प्रत्येक में समष्टि और रूचि-प्राचल का पता लगाइए। प्रत्येक मामले में जाँच कीजिए कि समष्टि परीमित है या नहीं। 4

(i) एक तेल कम्पनी यह जानना चाहती है कि उसका कौन सा सर्विस स्टेशन 24 घंटे काम करता है।

(ii) एक कृषि-फर्म वर्ष 2006 से 2010 के दौरान आम की पैदावार का अध्ययन करना चाहती है।

(iii) एक एजेंसी 2007 के बाद से उत्तर-प्रदेश में साक्षरता दर का पता लगाना चाहती है।

(iv) एक फर्म एक विशिष्ट मुहल्ले में परिवारों की माध्य आय जानना चाहती है।

2. (a) किसी विशिष्ट कीटनाशक के प्रभावों को निर्धारित करने के लिए कृषि-संबंधी प्रयोग किया गया। इसके अंतर्गत एक खेत में मक्का लगाई गई। आधे पौधों पर कीटनाशक का छिड़काव किया गया और आधे पौधों पर छिड़काव नहीं किया गया। कुछ सप्ताह बाद, 200 छिड़काव वाले और 200 बिना छिड़काव वाले पौधे यादृच्छया चुने गए। प्रत्येक प्रतिदर्श में स्वस्थ पौधों की संख्या निम्नानुसार थी : 4

	छिड़काव किए गए पौधे	बिना छिड़काव वाले पौधे
स्वस्थ	121	109
अस्वस्थ	79	91
	200	200

यदि सार्थकता स्तर 0.05 पर निर्धारित है तो क्या प्रमाण दर्शाता है कि बिना छिड़काव वाले पौधों की तुलना में छिड़काव वाले पौधे के स्वस्थ होने का अनुपात ज्यादा है। आप प्रश्न पत्र के अन्त में दी गयी सारणी का प्रयोग कर सकते है।

(b) 50 विद्यार्थियों द्वारा मनोविज्ञान की परीक्षा में 100 में से प्राप्त कुल अंक X नीचे दिए गए हैं **1+3+1+1=6**

80	90	85	55	40	50	40	95	81	82
75	70	80	40	10	55	30	80	82	72
27	95	20	30	85	10	40	23	75	60
50	30	10	10	90	15	20	40	70	65
45	50	40	10	90	20	25	42	70	70

(i) क्या विद्यार्थी का यादृच्छिक चर X -अंक असंतत है या संतत? न्यूनतम और अधिकतम अंक क्या हैं?

(ii) बारंबारता बंटन बनाइए।

(iii) कितने प्रतिशत विद्यार्थियों ने उत्तीर्ण होने के लिए अनिवार्य 50 अंकों से ज्यादा अंक प्राप्त किए?

(iv) कितने विद्यार्थियो ने 50 और 80 के बीच अंक प्राप्त किए?

3. (a) 25 व्यक्तियों के कद (h) और वज़न (w) के प्रतिदर्श आंकड़े का प्रयोग करते हुए श्रेष्ठतम रैखिक समाश्रयण फिट w = 9+0.4h (कद से.मी. में और वज़न कि.ग्राम में) पाया गया। प्रतिदर्श माध्य और कद का प्रसरण 165 और 200 है और वज़न क्रमश: 75 और 40 है। 6

(i) ऊपर दिए गए आंकड़े के लिए वज़न पर आधारित कद का श्रेष्ठतम रैखिक फिट प्राप्त कीजिए।

(ii) 68 किग्रा वजन वाले व्यक्ति के कद के संबंध में आपका प्रागुक्ति क्या है ?

(b) एक आँख के अस्पताल में 200 रोगियों की जाँच के समय का माध्य 10 मिनट और मानक विचलन 2 मिनट है। माध्य जाँच समय का 95 % विश्वस्यता अंतराल ज्ञात कीजिए। माध्य जाँच समय के लिए 95 % निम्न विश्वस्यता सीमा क्या है ? क्या इस सूचना से हम कह सकते हैं कि प्रत्येक रोगी की जाँच का लगभग समय 16 मिनट से कम है ? 4

4. (a) एक रसायन की फर्म यह निर्धारित करना चाहती है कि किस प्रकार चार उत्प्रेरकों से फसल में अंतर आता है। फर्म अपने आठ पौधों पर इसका प्रयोग करती है। प्रत्येक प्रकार के पौधे में, प्रत्येक उत्प्रेरक वाली फसल को मापा जाता है। फसल (क्विंटल में) निम्नानुसार है। 7

पौधा	उत्प्रेरक			
	1	2	3	4
A	2	1	2	4
B	3	2	1	3
C	1	3	3	1
D	5	4	3	2
E	1	2	4	3
F	2	1	1	2
G	4	3	4	2
H	3	1	4	3

ANOVA सारणी बनाइए और टिप्पणी कीजिए कि विशिष्ट उत्प्रेरक के कारण फसल 5% सार्थकता स्तर पर सार्थक है या नहीं। (आप प्रश्न पत्र के अन्त में दी गयी तालिका का प्रयोग कर सकते हैं।)

(b) एक अधूरा (अपूर्ण) बारंबारता बंटन नीचे दिया गया हैं : 3

वर्ग-अंतराल	बारंबारता
10-20	12
20-30	30
30-40	?
40-50	65
50-60	?
60-70	25
70-80	18

दिया गया है कि 200 प्रेक्षणों का माध्यिका मान 46 है, माध्यिका फार्मूले का प्रयोग करते हुए लुप्त बारंबारता बताइए।

5. (a) एक टेक्सटाइल कम्पनी की बिक्री के आंकड़े नीचे दिए गए है। लम्बाई 4 के लिए गतिमान औसत परिकलित कीजिए : 5

दिन	बिक्री
1	230
2	200
3	250
4	300
5	200
6	225
7	400
8	450
9	415
10	420
11	500
12	300
13	400
14	300
15	315

(b) एक पाँसे को 1200 बार फेंका जाता है और निम्नलिखित परिणाम प्राप्त होते है : 5

आने वाली सं.	1	2	3	4	5	6
बारंबारता	195	289	202	242	163	109

जाँच कीजिए कि पाँसा 5% सार्थकता स्तर पर अनभिनत है या नहीं। (आप अंत में दिए गए मानों का प्रयोग कर सकते हैं।)

6. (a) विभिन्न डिपार्टमेन्टल स्टोरों में साबुन की टिकियाँ की माध्य साप्ताहिक बिक्री 146.3 टिकियाँ प्रति स्टोर है। विज्ञापन प्रचार के एक विशिष्ट सप्ताह में 22 स्टोरों मे माध्य साप्ताहिक बिक्री बढ़कर 153.7 हो गई और मानक विचलन 17.2 रहा। क्या 5% सार्थकता स्तर पर विज्ञापन प्रचार सफल रहा? (आप अंत में दिए गए मानों का प्रयोग कर सकते है) 3

(b) एक बीयर कम्पनी अपने बीयर के डिब्बों का माध्य वज़न निर्धारित करना चाहती है। वे अपने गोदाम में रखे हज़ारो डिब्बों में से 80 बीयर के डिब्बे यादृच्छया उठाती है और पाती है कि उनका माध्य वज़न 31.15 आउन्स और मानक विचलन 0.08 आउन्स हैं। 2+1=3

(i) फर्म के गोदाम के डिब्बों के माध्य वज़न के लिए 95% विश्वस्यता अंतराल परिकलित कीजिए।

(ii) क्या आपका (i) का उत्तर इस पूर्वानुमान पर आधारित है कि गोदाम के बीयर के डिब्बों का वजन प्रसामान्यत: बंटित है? क्यों और क्यों नहीं?

(c) ऐसी दो स्थितियां बताइए जिनमें स्तरित प्रतिचयन समुचित हो। स्तरित प्रतिचयन के दो लाभों का वर्णन कीजिए। 4

7. बताइए निम्नलिखित कथन सही हैं या नहीं। अपने उत्तर की संक्षिप्त पुष्टि दीजिए : 10

(a) प्वांसा बंटन का माध्य और मानक विचलन समान होते हैं।

(b) यदि A और B धनात्मक प्रायिकताओं वाली असंयुक्त घटनाएँ हैं, तब वे स्वतंत्र है।

(c) ANOVA प्रसरणों की समता का परीक्षण करने का साधन है।

(d) यदि LCL और UCL प्रक्रम के लिए $\bar{x}$ -चार्ट पर नियंत्रण सीमाएँ है, तब (UCL-LCL)/6 प्रक्रम मानक विचलन का आकलन है।

(e) समष्टि को बनाने वाले वर्गों के विभिन्न प्रसरण पर आधारित समष्टि स्तरित होती है।

कुछ मान जिनका प्रयोग आप ज़रूरत पड़ने पर कर सकते हैं।			
χ^2 मान	5% सार्थकता स्तर पर F- मान	t - मान	प्रसामान्य मान
$\chi^2_{5, 0.05} = 11.07$	$F_{3, 7} = 4.35$	$t_{21, 0.05} = 721$	$P(Z \leq -0.02) = 0.492$
$\chi^2_{6, 0.05} = 12.59$	$F_{3, 21} = 3.07$	$t_{22, 0.05} = 1.717$	$P(Z \leq 0.02) = 0.508$
$\chi^2_{4, 0.05} = 9.48$	$F_{7, 21} = 2.49$	$t_{11, 0.05} = 1.796$	