LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

B.Sc. DEGREE EXAMINATION – **STATISTICS**

FIFTH SEMESTER – NOVEMBER 2011

ST 5507 – COMPUTATIONAL STATISTICS

Date : 08-11-2011 Dept. No. Max. : 100 Marks

Time : 9:00 - 12:00

Answer any THREE of the following questions: **Max.Mark:100**

(1) (a) A firm that runs a string of retail outlets across a city receives complaints from its clients regarding quality and other aspects and maintains a register of complaints. The following are data on the number of complaints received on 100 randomly chosen days:

No. of Complaints	0	1	2	3	4	5	6	7
No. of days	10	31	26	18	7	4	3	1

Test at 5% level of significance whether the number of complaints per day follows Poisson distribution.

(b) The following table gives the distances that a particular brand of battery-operated vehicle ran before developing technical troubles. Data on 600 trial vehicles are available:

Distance in kms	150-250	250-350	350-450	450-550	550-650	650-750	750-850
No. of vehicles	2	4	14	50	65	105	127

Distance in kms	850-950	950-1050	1050-1150	1150-1250	1250-1350	1350-1450
No. of vehicles	87	61	53	22	8	2

Fit a normal distribution to the data and test for goodness of fit at 5 % level of significance. Estimate the probability for a randomly chosen vehicle to develop troubles before completing 150 kms.

(13 +20)

2) (a) A population consists of 6 units with 'Y' values 3, 5, 8, 11, 12, 15. By choosing simple random samples (WOR) of size 2, verify the results $E(\bar{y}) = \bar{Y}$ and $E(s^2) = S^2$.

(b) A population with 15000 units is stratified into 4 strata. The stratum variance S_h^2 for each stratum was estimated from a pilot survey and the estimates are reported below along with the sizes of the strata and the per-unit costs of conducting the study:

Stratum No.	Stratum Size	Stratum variance (S_h^2)	Per-unit cost (in Rs)
1	2000	85.54	50
2	3000	104.25	75
3	4500	96.82	90
4	5500	72.93	40

The investigators wish to choose a sample of 300 units using stratified random sampling. Compute the sample sizes to be drawn from the different strata according to

(i) Equal Allocation

(ii) Proportional Allocation

(iii) Optimum Allocation (18 + 15)

3) (a) Compute index number for the given data using the following methods (i) Laspeyre's method, (ii) Paasche's method and (iii) Fisher's ideal formula (8)

Item (Rs.)	Base year		Current year	
	Price (in Rs)	Expenditure	Price (in Rs)	Expenditure
A	6	300	10	560
B	2	200	2	240
C	4	240	6	360
D	10	300	12	288
E	8	320	12	432

(b) Construct Index number by chain base method from the following data of wholesale prices of a certain commodity: (5)

Year	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Price	75	50	65	60	72	70	69	75	84	80

(c) Given the following information, calculate the seasonal indices using the method of ratios to moving averages. (Multiplicative model) (20)

	Quarter			
Year	I	II	III	IV
2000	106	124	104	90
2001	84	114	107	88
2002	90	112	101	85
2003	76	94	91	76
2004	80	104	95	83
2005	104	112	102	84

4) (a) In a large city , 20 % of a random sample of 900 school children had defective eye sight. In other large city B, 15% of random sample of 1600 children had the same defect. Is this difference between the two proportions significant ? Test at 1% level of significance.

(8marks)

(b)Ten specimens of copper wires drawn from a large lot have the following breaking strength (in kg weight) : 578 572 568 572 571 570 572 596 548 570 . Test whether the mean breaking strength of the lot may be taken to be 578 kg wt . Use 5% significance level.

(9marks)

(c) Seven coins were tossed and the number of heads noted. The experiment was repeated 128 times and the following distribution was obtained.

No. of heads :	0	1	2	3	4	5	6	7
Frequency :	7	6	19	35	30	23	7	1

Fit a binomial distribution to the given data and test the goodness of fit at 1% level of significance. (16marks)

5) (a) Let X denote the length of time in seconds between two calls entering a college switchboard. Let m be the unique median of this continuous-type distribution. Test the null hypothesis $H_0 : m = 6.2$ against the alternative hypothesis $H_1: m < 6.2$ using a random sample of size 20 given below:
6.8, 5.7, 6.9, 5.3, 4.1, 9.8, 1.7, 7.0, 2.1, 19.0, 18.9, 16.9, 10.4, 44.1, 2.9, 2.4, 4.8, 18.9, 4.8, 7.9.

Find the **significance level α** if the critical region **C = {y | y ≥ 14},** where **'y'** is the number of lengths of time in a random sample of size 20 that are less than 6.2. Find also the **p – value** of this sign test. (13)

(b) A vendor of milk products produces and sells low-fat dry milk to a company that uses it to produce baby formula. In order to determine the fat content of the milk, both the company and the vendor take a sample from each lot and test it for fat content in percent. Ten sets of paired test results are

Lot Number	Company Test Results (X)	Vendor Test Results (Y)
1	0.50	0.79
2	0.58	0.71
3	0.90	0.82
4	1.17	0.82
5	1.14	0.73
6	1.25	0.77
7	0.75	0.72
8	1.22	0.79
9	0.74	0.72
10	0.80	0.91

Test the hypothesis $H_0 : p = P[X > Y] = \frac{1}{2}$ against the one – sided alternative $H_1 : p > \frac{1}{2}$ using the critical region **C = { w | w ≥ 8 },** where **'w'** is the number of pairs for which $X_i - Y_i > 0$. Find the **significance level α** and **p – value** of this test. (20)

****