LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

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

FIFTHSEMESTER – APRIL 2017

ST 5503- COMPUTATIONAL STATISTICS

Date: 26-04-2017 Dept. No. Max. : 100 Marks
01:00-04:00

Answer any three questions. Each question carries 34 marks.

1. (a) In a population of size N = 6 values of the population characteristic are 1,2,4,6,8 and 10. A sample of size n=2 is drawn without replacement. Verify that $\bar{y}$ is an unbiased estimate of $\bar{Y}$ and that Var($\bar{y}$) is(S^2/n). (N-n) /N . **(9marks)**

   (b) A population of size 120 is divided into four strata whose sizes are N_1 = 50 , N_2 = 10 ,N_3 = 20 and N_4 = 40 and variances (S_i^2) are 3,4,5,6 respectively. How will you draw samples of size 20 according to the two allocations(proportional and optimum) ? **(10 marks)**

   (c) 2020 cultivator's holdings in U.P. are stratified according to size. The number of holdings (N_h), mean area under wheat per holding ($\bar{Y}_h$) and standard deviation of area under wheat per holding (S_h) are given below for each stratum.

Stratum no.	Holding size (acres)	No. of Holdings (N_h)	$\bar{Y}_h$	S_h
1	0 - 40	395	5.5	8.4
2	41 - 80	464	16.4	13.4
3	81 -120	393	24.4	15.2
4	121-160	335	34.6	19.9
5	161-200	171	42.2	24.6
6	200andabove	262	58.0	31.3

   A sample of 110 holdings is taken to estimate the mean area under wheat per holding by

   (i) Simple random sampling (ii) stratified random sampling with proportional allocations

   (iii) Stratified random sampling with optimum allocations.

   Compare the standard errors of the estimates in the three cases.**(15 marks)**

2. (a) Find the maximum likelihood estimates for $\theta_1 = \mu$ and $\theta_2 = \sigma^2$ if a random sample of size 15 from N(μ , σ^2) yielded the following values: 32.7 37.1 34.0 31.2 34.1 36.4 30.8 35.5 31.7 35.4 32.8 37.9 37.0 35.7 33.9. **(6 marks)**

   (b) As a clue to the amount of organic waste in Lake Macatawa, a count was made of the number of bacteria colonies in 100 milliliters of water. The number of colonies, in hundreds, for n = 30 samples of water from the east basin yielded the following observations:

96	14	31	11	23	61	23	39	73	94
64	10	10	32	21	10	11	13	26	17
97	60	12	69	57	31	79	61	12	27

   Find an approximate 95% confidence interval for the mean number of colonies in 100 milliliters of water in the east basin , μ_E. **(8 marks)**

   (c) A botanist measured the growths of pea stem segments, in millimeters , for n = 12 observations of and m=14 observations of

   X: 0.9 1.9 1.1 0.2 1.0 1.8 1.1 1.5 1.0 1.3 0.6

   Y: 1.2 1.0 1.8 2.8 1.5 1.3 2.6 2.0 2.7 1.6 2.1 2.2 1.4

   Test the hypothesis $H_0 : \sigma^2_X = \sigma^2_Y$ against $H_1 : \sigma^2_X \neq \sigma^2_Y$ at $\alpha = 0.01$. **(12 marks)**

(d) The intelligent quotient of 12 students are as follows: 115 125 150 165 133 131 143 160 134 145 . Test $H_0 : \mu = 150$ against $H_1 : \mu \neq 150$ at $\alpha = 0.05$. **(8 marks)**

3. (a) The number of mistakes per page in a book of 265 pages are given below:

No.of mistakes per page (x) :	0	1	2	3	4	5
No. of pages(f) :	159	61	23	13	6	3

Fit a Poisson distribution to the given data and test the goodness of fit at 1% level of significance. **(15 marks)**

(b) The random samples from two different populations are given below:

X: 64	75	68	80	53	66	86	96	60	80
Y: 70	80	60	65	95	45	70	82	59	64

Test $H_0 : \mu_X = \mu_Y$ against $H_1 : \mu_X \neq \mu_Y$ at $\alpha = 0.01$. **(9 marks)**

(c) Let X and Y be the percentages of body fat for women and men , respectively with distribution functions F(x) and G(y). Using runs ,test the hypothesis H_0: F(z) = G(z) against H_1: F(z) < G(z) at $\alpha = 0.01$. Twelve observations of both X and Y that have been ordered are

X: 16.8	16.9	18.7	19.4	21.7	22.6	22.8	23.4	24.4	26.5	28.2
Y: 9.2	9.9	11.5	12.0	13.5	15.8	16.3	16.7	18.4	21.9	22.4

(10 marks)

4. (a) For the following data compute Fisher , Dorbish-Bowley , Marshall-Egeworth and Walsh price index numbers:

Commodity	Base year price	Current year price	Base year quantity	Current year quantity
A	45	55	11	6
B	25	35	9	5
C	35	45	9	7
D	15	25	10	10

Also verify the time and factor reversal tests. **(16 marks)**

(b) In the construction of certain cost of living index number , the following group index numbers were found. Calculate the cost of living index number using (i) weighted arithmetic mean (ii) weighted geometric mean.

Group:	A	B	C	D	E
Index number:	352	200	230	160	190
Weight :	50	10	10	15	15

(8 marks)

(c) Given below are two sets of indices:

Year :	1939	1940	1945	1947	1949	1950	1951	1952
Index(old) A :	100	120	125	135	150	…	...	...
Index (new) B :	…	…	…	…	100	115	130	140

(i) Splice new series to old series.

(ii) Splice old series to new series. **(6 marks)**

(d)The annual wages of a worker in rupees along with price index numbers are given below:

Year :	1971	1972	1973	1974	1975
Wages :	225	250	280	288	325
Index number :	100	120	135	145	155

Prepare index numbers for real wages of workers. **(4 marks)**

5.(a) Fit a straight line trend by least squares to the following data and calculate trend values:

Year	: 1985	1986	1987	1988	1989	1990
Production ('000 tons)	:85	93	119	139	144	158

(14 marks)

(b) Find out the seasonal indices by the method of moving averages for the following data:

Quarter / Year	I	II	III	IV
2002	35	45	40	39
2003	39	57	55	49
2004	45	63	59	52
2005	59	81	63	72
2006	85	97	85	87

(20 marks)

$$$$$$$$