# LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – STATISTICS

THIRD SEMESTER – NOVEMBER 2009

**ST 3814 - STATISTICAL COMPUTING - II**

Date & Time: 10/11/2009 / 9:00 - 12:00 Dept. No. Max. : 100 Marks

Answer ALL the Questions.

1. a). Let $\{X_n, n=0,1,2,3,4,\ldots..\}$ be a Markov chain with state space {0,1,2 } and one step matrix of transition probabilities

$$P=\begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.3 & 0.2 & 0.5 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

Find (i) $P^2$ (ii) $\lim_{n\to\infty} P^n$ (iii) $P[X_2=0]$

given $X_0$ takes the values 0,1,2 with probabilities 0.3,0.4,0.3 respectively. **(12 marks)**

(b). For a Markov chain with one step matrix of transition probabilities as

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 & 0 \end{bmatrix}$$

and with state space {0,1,2,3},clearly mention the states as transient, recurrent, positive recurrent or null recurrent. **(22 marks)**

**(OR)**

(c). An infinite Markov chain on the set of non-negative integers has the transition function as follows:

$$p_{k0} = \frac{K+1}{K+2} \quad \text{and} \quad p_{k,\,k+1} = \frac{1}{K+2}$$

i) Find whether the chain is positive recurrent, null recurrent or transient.

ii) Find the stationary distribution, in case it exists. **(20 marks)**

(d). In a genetical experiment, the following frequencies were observed:

| AB | Ab | aB | ab |
|---|---|---|---|
| 140 | 22 | 28 | 10 |

If theory predicts the probabilities to be $\frac{2+\theta}{4}$ , $\frac{1-\theta}{4}$ , $\frac{1-\theta}{4}$ , $\frac{\theta}{4}$ , obtain the maximum likelihood estimate of θ and test the goodness of fit. **(14 marks)**

2. (a). To study the effects of a drug on a particular disease 12 patients were selected in a clinical trials. The measurements on 3 variables are given below (in micrograms).

| Sl.no | $X_1$ | $X_2$ | $X_3$ |
|---|---|---|---|
| 1 | 1.40 | 0.50 | 0.71 |
| 2 | 1.18 | 0.39 | 0.69 |
| 3 | 1.23 | 0.44 | 0.70 |
| 4 | 1.19 | 0.37 | 0.72 |
| 5 | 1.38 | 0.42 | 0.71 |
| 6 | 1.17 | 0.45 | 0.70 |
| 7 | 1.31 | 0.41 | 0.70 |
| 8 | 1.30 | 0.47 | 0.67 |
| 9 | 1.22 | 0.29 | 0.68 |
| 10 | 1.00 | 0.30 | 0.70 |
| 11 | 1.12 | 0.27 | 0.72 |
| 12 | 1.09 | 0.35 | 0.73 |

(i) Estimate $\mu$, $\Sigma$ *and* the correlation matrix.

(ii) Estimate the parameters for the conditional distribution of $X_3$ given $X_1 = 1.5$, $X_2 = 0.6$ using S.

(iii) Find whether the variable $X_1$ is marginally normal.

(iv) Which of the sample correlations are significant? **(8+10+10+5)**

**(OR)**

(b). The tail length in millimeters for 15 male and female hook-billed kites are given below:

| | | |
|---|---|---|
| Male | Tail ($X_1$) | 180 186 206 184 177 177 176 200 191 193 212 181 195 187 190 |
| | Wing ($X_2$) | 278 277 308 290 273 284 267 281 287 271 302 254 297 281 284 |
| Female | Tail ($X_1$) | 191 197 208 180 180 188 210 196 191 179 208 202 200 192 199 |
| | Wing ($X_2$) | 284 285 288 273 275 280 283 288 271 257 289 285 272 282 280 |

(i) Test whether $\Sigma_1 = \Sigma_2$. **(17 marks)**

(ii) Using Behrens-Fisher method test whether the mean vectors are equal. **(16 marks)**

3. (a) The following sampling design is adopted to select a sample from a population with six units:

$$P(s) = \begin{cases} 0.2, & for \quad s = \{1,3,6\},\{2,4,5\} \\ 0.3, & for \quad s = \{1,2,5\},\{3,5,6\} \end{cases}$$

Find all the first and second order inclusion probabilities. Also, verify the result

$E[n(s)] = \sum_{i=1}^{N} \pi_i$ **(18 marks)**

(b) The following information are available from a pilot survey using a stratified random sample:

| **Stratum Size ($N_h$)** | **Sample Size($n_h$)** | **Sample std. Devn. ($s_h^2$)** | **Cost per Unit ($C_h$)** |
|---|---|---|---|
| 200 | 10 | 2.5 | 12 |
| 300 | 5 | 1.2 | 16 |
| 500 | 8 | 1.5 | 20 |
| 400 | 10 | 2.0 | 15 |
| 600 | 17 | 2.4 | 14 |

Find the optimum sample sizes to be drawn from each stratum for a full-fledged survey if the total sample size has to be 200. **(15 marks)**

**(OR)**

(c) In a survey of 100 commercial buildings in a town, it is found that 21 have not installed proper water-harvesting structures. The total number of commercial buildings in the town is known to be 1500. Compute a 99% confidence interval for the proportion of buildings without water-harvesting structures in the town. **(10 marks)**

(d) A pilot survey of 20 households in a locality gave the following information on the number of family members ($x$) and the number of mobile phones used ($y$) in each family:

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***x*** | 3 | 4 | 4 | 2 | 6 | 5 | 3 | 4 | 2 | 5 | 4 | 6 | 3 | 4 | 4 | 5 | 2 | 3 | 4 | 4 |
| ***y*** | 1 | 3 | 2 | 2 | 3 | 3 | 2 | 2 | 2 | 2 | 3 | 4 | 1 | 2 | 1 | 4 | 1 | 1 | 3 | 4 |

The number of households in the locality is known to be 700 and the number of people living in the locality is 2800. Based on the pilot survey results, would you recommend usage of 'Ratio estimate' in preference to the usual estimate $N\bar{y}$, to estimate the total number of mobile phones in the locality? Support your answer with proper theoretical justification. **(23 marks)**

************