19

**RAJARATA UNIVERSITY OF SRI LANKA**
**FACULTY OF APPLIED SCIENCES, MIHINTALE**

**B.Sc. (General) Degree**

**Third Year - Semester I Examination - February/March 2013**

**MAT 3214-APPLIED STATISTICS**

**Answer all questions.** **Time allowed: Two hours**

---

**Statistical tables and calculators will be provided.**

1.

a)

(i) As methods of data collection, what is the difference between the **census** and **sampling**?

(ii) Write short notes on two random sampling methods.

b) The life time of an electric light bulb has an exponential distribution with mean 2000 hours. If a sample of 30 such light bulbs were selected, determine the approximate probability that the average life time of the sample exceeds 1300 hours.

2.

a) The yield of rice in kg was measured from 10 experimental plots in two successive years.

| Plot number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1st year | 25 | 21 | 18 | 20 | 19 | 23 | 17 | 19 | 23 | 24 |
| 2nd year | 24 | 19 | 22 | 17 | 20 | 21 | 22 | 21 | 25 | 23 |

(i) Set up appropriate hypothesis to investigate whether there is a significant difference between the mean yield of 1st year and the mean yield of 2nd year.

(ii) Test these hypotheses using 5% level of significance.
What is your conclusion?

*[Continue…]*

b) A soft drink manufacturer claims that the average content in a bottle is $120ml$. A random sample of 12 bottles was tested, and the following results were obtained:

| Bottle number | Content(ml) |
|---|---|
| 1 | 118 |
| 2 | 115 |
| 3 | 120 |
| 4 | 122 |
| 5 | 121 |
| 6 | 125 |
| 7 | 128 |
| 8 | 113 |
| 9 | 126 |
| 10 | 115 |
| 11 | 121 |
| 12 | 118 |

(i) Set up appropriate hypothesis to investigate this claim.
(ii) Test the hypothesis using 5% level of significance, stating any assumptions you make. What is your conclusion?

3.

a) In a $2 \times 2$ contigency table, the observed frequencies are as shown below:

| | A | B | Total |
|---|---|---|---|
| Group I | $a$ | $b$ | $a+b$ |
| Group II | $c$ | $d$ | $c+d$ |
| Total | $a+c$ | $b+d$ | |

Show that in the usual notation,

$$\sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} = \frac{(a+b+c+d)(ad-bc)^2}{(a+b)(c+d)(b+d)(a+c)}$$

*[Continue...]*

b) Random samples of 100 housewives are classified according to their social status $S_1$ or $S_2$ and preference of a particular brand of a commodity $C_1$, $C_2$ and $C_3$. The data are given below:

| | $C_1$ | $C_2$ | $C_3$ | Total |
|---|---|---|---|---|
| $S_1$ | 20 | 12 | 25 | 57 |
| $S_2$ | 23 | 13 | 7 | 43 |
| Total | 43 | 25 | 32 | 100 |

(i) Find the expected frequencies.

(ii) Test at 5% level of significance, whether the preference for a brand depends upon the social status.

4.

a) Show that the chi square test statistic of goodness of fit test

$$\sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{n} \frac{O_i^2}{E_i} - N$$

where $O_i$ is the observed frequency of the i$^{th}$ category, $E_i$ is the expected frequency of the i$^{th}$ category and N is the total number of observations.

b) A packet contains ~~10~~ 07 pencils. A random sample of 400 packets were checked for defective pencils. The following table gives the results, where X is the number of defective pencils in a packet.

| x | Observed frequency |
|---|---|
| 0 | 162 |
| 1 | 88 |
| 2 | 56 |
| 3 | 35 |
| 4 | 28 |
| 5 | 19 |
| 6 | 8 |
| 7 | 4 |

(i) Suggest a suitable distribution to model the defective pieces in a packet.

(ii) Test at 5% level of significance, whether the observed results are consistent with the model specified under part (i).