No. of Printed Pages : 11 **MIA-002 (F2F)**

# M.Sc. ACTUARIAL SCIENCE

## Term-End Examination

## December, 2011

## MIA-002 (F2F) : PROBABILITY AND MATHEMATICAL STATISTICS

*Time : 3 hours* *Maximum Marks : 100*

**Note :** *Students are allowed to use scientific calculator and actuarial tables.*

---

### SECTION - A

Attempt *any five* questions.

**1.** (a) The following data represent the scores obtained by 27 students in a mid-term test :

79, 78, 78, 67, 76, 87, 85, 73, 66,
99, 84, 72, 66, 57, 94, 84, 72, 63,
51, 48, 50, 61, 71, 82, 93, 100, 89,

(i) Prepare stem and leaf display of the data. **3**

(ii) Calculate the range and the inter quartile range. **2**

(b) The ages in year of a group of 20 individuals are as follows : **3**

50, 56, 55, 49, 52, 57, 56, 57, 56, 59,
54, 55, 61, 60, 51, 59, 62, 52, 54, 49

If $\bar{x}$ denotes the mean and $s$ the standard deviation of the sample, find the percentage of items failing within the interval $\bar{x} \pm s$

(Given $\Sigma x = 1104$
$\Sigma x^2 = 61226$)

2. A random sample $(x_1 x_2, \ldots x_n)$ is taken from a poisson distribution, with parameter $\mu$.

(a) Show that the maximum likelihood estimator of $\mu$ is : **3**

$$\hat{\mu} = \bar{x}$$

(b) Obtain the bias and mean square error of $\hat{\mu}$. **2**

(c) Show that the variance of $\hat{\mu}$ attains the Cramer - Rao lower bound. **3**

3. (a) In an insurance company, three actuarial assistants A, B and C are assigned to process loan application forms against the policies. A processes 40%; B processes 35% ; and C processes 25% of the loan applications. The error rate while processing the loan applications by A, B and C are 0.04, 0.06 and 0.03 respectively. A loan application selected at random by the Branch manager is found to have an error. What is the probability that it was processed by A ? **4**

(b) If the probability is 0.20 that a certain bank will refuse a loan application, using normal approximation (to three decimals), find the probability that the bank will refuse at most 40 out of 225 loan applications. **4**

4. The Gamma distribution with parameters $\alpha$ and $\lambda$, has moment generating function :

$$M_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-\alpha}$$

(a) Show, using moment generating function, that the sum of two independent Gamma distributions, each with parameter $\lambda$, is also a Gamma distribution. 3

(b) A random sample $X_1, X_2 \ldots X_n$ is taken from a Gamma ( $\alpha$, $\lambda$) distribution. Derive the moment generating function of $2\lambda \Sigma x_i$ and hence show that it has a $\chi^2_{2n}$ distribution. 3

(c) Suppose that $\overline{X}$ is the mean of a random sample of size 5 taken from a Gamma (2, 0.1) distribution. Use the result from part (b) to calculate the probability that $\overline{X}$ exceeds 40. 2

5. A large office has $n$ policyholders, each with a probability of 0.01 of dying during the next year (independently of all other policyholders).

Calculate the approximate probability that these will be between 9 and 16 (both inclusive) deaths during the year, when :

(a) $n = 400$ 4

(b) $n = 3{,}000$ 4

6. The joint pmf of $x$ and $y$ is given below.

| $y$ | $x$ = 0 | $x$ = 1 | $x$ = 2 |
|---|---|---|---|
| 0 | 1/6 | 1/3 | 1/12 |
| 1 | 2/9 | 1/6 | 0 |
| 2 | 1/36 | 0 | 0 |

find :

(a) Mean values of X and Y. 2

(b) Find the covariance between X and Y. 1

(c) Conditional distribution of X given Y=1. 2

(d) Correlation between X and Y. 3

7. A survey was conducted to investigate whether people tend to marry partners of about the same age. This question was addressed to 12 married couples and their ages were given in the following table.

| Couple No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Husband's age (x) | 30 | 29 | 36 | 72 | 37 | 36 | 51 | 48 | 37 | 50 | 51 | 36 |
| Wife's age (y) | 27 | 20 | 34 | 67 | 35 | 37 | 50 | 46 | 36 | 42 | 46 | 35 |

(a) Draw the scatter plot and comment on it. 1

(b) Find the correlation coefficient and interpret it. 4

(c) If $\rho$ represent the population correlation coefficient between the ages of partners, test the significance of $\rho$ at 5% level. 3

## SECTION - B

Attempt *any four* questions.

8. (a) (i) State the central limit theorem. **2**

(ii) What is the approximate distribution of the arithmetic mean of a set of 100 independent observations from a Gamma (2.5, 0.2) distribution ? **3**

(b) House's prices in region X are normally distributed about a mean of Rs. 1,00,000 with a standard deviation of Rs. 10,000. House prices in region Y are normally distributed about a mean of Rs. 90,000 with a standard deviation of Rs. 5,000. A sample of 10 houses is taken from region X and a sample of 5 houses from region Y. Find the probability that :

(i) the region X sample mean is greater than the region Y sample mean. **5**

(ii) the difference between the sample mean is less than 5,000.

(iii) the region X sample variance is less than the region Y sample variance. **5**

(iv) the region X sample standard deviation is more than four times greater than the region Y sample standard deviation.

9. (a) Eight pairs of slow learners with similar reading capabilities are identified in a third grade class. One member of each pair is randomly assigned to the standard teaching method, while the other is assigned to a new teaching method. The scores are as given below.

| Pair | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| New method | 77 | 74 | 82 | 73 | 87 | 69 | 66 | 80 |
| Old method | 72 | 68 | 76 | 68 | 84 | 68 | 64 | 76 |

(i) Test for the difference between mean scores for the two methods. 6

(ii) Test for the equality of variances for these two methods. 4

(b) The diameter of steel rods manufactured on two different machines A and B is studied. Two random samples of sizes $n_1 = 12$ and $n_2 = 15$ are selected and the sample means and sample standard deviations respectively

are : $\bar{x}_1 = 24.6$ , $S_1 = 0.85$

$\bar{x}_2 = 22.1$ , $S_2 = 0.98$

Assuming that the diameters of the rods are as follow $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$,

(i) Test the equality of means of the diameters of the rods manufactured by the two machines assuming $\sigma_1^2 = \sigma_2^2$. **3**

(ii) Construct 95% confidence interval for $\mu_1 - \mu_2$ assuming $\sigma_1^2 = \sigma_2^2$. **2**

**10.** Two thousand individuals were choosen at random by a researcher and cross classified according to gender and colour blindness as given below :

| Description | Male | Female |
|---|---|---|
| Normal | 904 | 998 |
| Colour Blind | 91 | 7 |

(a) Apply an appropriate test to conclude that there is overwhelming evidence against the hypothesis that there is no association between gender and colour blindness. **4**

(b) A genetic model states that the human population is split in the proportion as illustrated in the following table where q $(0<q<1)$ is a parameter related to the distribution of the colour blindness.

| Description | Male | Female |
|---|---|---|
| Normal | $(1-q)/2$ | $(1-q^2)/2$ |
| Colour Blind | $q/2$ | $q^2/2$ |

Using the data in a) :

(i) Write down the likelihood function for the above model. 3

(ii) Determine the maximum likelihood estimate of q. 8

**11.** An actuary has been advised to use the following positively - skewed claim size distribution as a model for a particular type of claim, with claim sizes measured in units of Rs. 100

$$f(x\ ;\ \theta) = \frac{x^2}{2\theta^3}\ \exp\left(\frac{-x}{\theta}\right):\ 0 < x < \infty,\ \theta>0$$

with moments given by $E(x) = 3\theta$, $E(x^2) = 12\theta^2$ and $E(x^3) = 60\theta^3$.

(a) Determine the variance of this distribution and calculate the coefficient of skewness. 4

(b) Let X1, X2 ... Xn be a random sample of n claim sizes for such claims. 5

Show that the maximum likelihood estimator (MLE) of $\theta$ is given by $\hat{\theta} = \frac{\overline{x}}{3}$ and show that it is unbiased for $\theta$.

(c) A sample of $n = 50$ claim sizes yields $\Sigma x_i = 313.6$ and $\Sigma x_i^2 = 2{,}675.68$.

(i) Calculate the MLE $\hat{\theta}$. **2**

(ii) Calculate the sample variance and comment briefly on its comparison with the variance of the distribution evaluated at $\hat{\theta}$. **2**

(iii) Given that the sample coefficient of skewness is 1.149, comment briefly on its comparison with the coefficient of skewness of the distribution. **2**

**12.** A humidity influences evaporation, the solvent balance of water - reducible paints during spray-out is affected by humidity. A study is conducted to examine the relation between humidity (X) and the extent of solvent evaporation (Y). The following data summary is obtained :

$n = 25, \ \Sigma x = 1314.90, \ \ \Sigma y = 235.70$

$\Sigma x^2 = 76{,}308.53, \ \ \Sigma y^2 = 2286.07$

$\Sigma xy = 11{,}824.44$

(a) Find the estimated correlation coefficient between X and Y and test the hypothesis $H_0 : \rho = 0$, against $H_1 : \rho \neq 0$ ; $\rho$ being the population correlation coefficient between X and Y. **4**

(b) Stating the assumptions, fit a regression line of the model $\gamma_i = \beta_0 + \beta_1 x_i + ei$ for the above data. 3

(c) Obtain the unbiased estimator of $\sigma^2$. 2

(d) Test the hypothesis $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$. 3

(e) Obtain 90% confidence interval for $\beta_0$ and 99% confidence interval for $\beta_1$. 3

13. Three different drugs are being compared for their effectiveness in treating a certain illness. The mean number of days before the patient is discharged from hospital under each treatment is summarised below, together with the sample size and the sum of squares of the observations :

| Treatment | Sample size | Sample mean | sum of squares |
|---|---|---|---|
| A | 10 | 5 | 264 |
| B | 6 | 7 | 310 |
| C | 8 | 3 | 84 |

(a) For these three treatments, calculate estimates for the : 4

(i) overall mean

(ii) common underlying variance.

(b) Perform an analysis of variance to show that real differences exist among the three treatments at the 1% level. 5

(c) Show that the mean number of days before discharge under treatments A is significantly better than under treatment B. 3

(d) The cost per day for treatments A, B and C are Rs. 7.50, Rs. 5.85, Rs. 14.95 respectively. Given that it can also be shown that there are significant differences between each pair of treatments, briefly advise the hospital on which treatment it should use. 3