INSTITUTE OF ACTUARIES OF INDIA

EXAMINATIONS

27th October 2015

Subject CT3 – Probability & Mathematical Statistics

Time allowed: Three Hours (10.30 – 13.30 Hrs.)

Total Marks: 100

INSTRUCTIONS TO THE CANDIDATES

1. ***Please read the instructions on the front page of answer booklet and instructions to examinees sent along with hall ticket carefully and follow without exception***
2. ***Mark allocations are shown in brackets.***
3. ***Attempt all questions, beginning your answer to each question on a separate sheet. However, answers to objective type questions could be written on the same sheet.***
4. ***Please check if you have received complete Question Paper and no page is missing. If so, kindly get new set of Question Paper from the Invigilator.***

AT THE END OF THE EXAMINATION

Please return your answer book and this question paper to the supervisor separately.

Q. 1) The following is the frequency distribution of 125 claims made from a certain portfolio.

Number of Claims	0	1	2	3	4	5	6
Frequency	10	?	35	?	15	7	8

The mean number of claims of this distribution is 2.504. However, some frequencies are missing.

i) Calculate the missing frequencies. (3)

ii) Compute median and comment on symmetry of the distribution. (2)
[5]

Q. 2) i) Show that the probability generating function of binomial distribution with mean 12 and n = 20 is $G_x(t) = (0.4 + 0.6\text{t})^{20}$.

Deduce the moment generating function of the above distribution. (4)

ii) Independent random samples are taken from two normal populations A and B. The population variances are σ_A^2 = 10 and σ_B^2 = 8. The sample sizes are n_A = 13 and n_B = 15. The sample variances are S_A^2 and S_B^2.

Determine P $\left(\frac{S_A^2}{S_B^2} < 4.75\right)$ (3)

iii) The number of complaints, X, handled by a call centre officer in a day is modelled as Poisson with mean 5. The time (in minutes) the call centre officer takes, Y, to process x complaints is modelled as having a distribution with a conditional mean and variance given by:

E (Y|X=x) = 2x + 3 and V (Y|X=x) = x + 1

Calculate the unconditional variance of the time the call centre officer takes to process complaints in a day. (3)
[10]

Q. 3) The probability function of the discrete random variable N is

P (N = n) = $\text{n}/(\text{n}+1)!$; n=0, 1, 2…
0 otherwise

i) Show that P (N = n) = $\frac{1}{\text{n}!} - \frac{1}{(\text{n}+1)!}$

Hence, or otherwise, show that the distribution function of N is

F (n) = 0 for $n < 0$

= 1- $\frac{1}{([\text{n}]+1)!}$ for $0 \leq n < \infty$

where [n] is the integral part of n (4)

ii) Show that the probability generating function of N is

$$G_N(t) = \frac{1+(t-1)e^t}{t}$$ (4)

iii) Hence or otherwise find the mean of N. (3)

[11]

Q. 4) **i)** At a hospital patients were categorised as 'critical', 'serious' and 'stable on arrival'.

It is seen from past records of the hospital that 20% patients were critical, 35% were serious and rest were stable on arrival. The hospital claims that the probabilities of deaths for critical, serious and stable patients respectively are 0.50, 0.20 and 0.05.

Given that a patient survived, what is the probability that patient was critical upon arrival. (4)

ii) The pregnancy time for human beings follows a normal distribution with a mean period of 268 days and standard deviation of 16 days. A random sample of 100 women was selected. Find the probability that the average pregnancy time period of the sample will be more than 265 days. (3)

[7]

Q. 5) **i)** In an opinion poll, each individual in a random sample of 500 individuals from a large population is asked whether he/she supports the economic reforms of the country. If 40% of the population supports the economic reforms, calculate the probability that at least 220 of the sample support them. (3)

ii) Given that the probability of an Indian male is diagnosed of testicular cancer is 0.000125. In a random sample of 50,000 males, find the probability of observing 2 or fewer males with diagnosis of testicular cancer. (3)

[6]

Q. 6) An automated machine fills a drug on an average of 250 mg per capsule. It has been observed that the amount of drug filled by the machine follows normal distribution with standard deviation of 1 mg. On a given day a random sample of size 4 capsules is selected and the amount of drug filled in each capsule is measured.

i) Calculate probability of the sample mean lies in the interval (249.8 mg, 250.2 mg). (3)

ii) Determine the minimum sample required for the sample mean to lie in the interval (249.6 mg, 250.4 mg) with at least 99% probability. (3)

[6]

Q. 7) The claim amount X for a policy follows log-normal distribution with parameters μ and σ^2.

i) Show that the MLE of μ and σ^2 based on a random sample $X_1, X_2, \dots, X_n$ are given by

$$\hat{\mu} = \frac{1}{n}\sum_1^n lnX_i \ ; \qquad \hat{\sigma}^2 = \frac{1}{n}\langle \sum_1^n (lnX_i)^2 - 2\hat{\mu}\sum_1^n lnX_i + n\,\hat{\mu}^2 \rangle$$

(It is not required to check for maximum) (4)

ii) Derive an expression in terms of CRLB, for the asymptotic 95% confidence interval of μ. (3)

iii) Show that $\hat{\mu}$ is unbiased and consistent. (4)

iv) Examine if the variance of $\hat{\mu}$ attains CRLB (1)

[12]

Q. 8) A diamond merchant desires to carry out a study to ascertain the proportion of population P who are aware of the existence of his shop.

i) Determine the minimum sample size needed to ensure that the widest 90% confidence interval for p lies in the interval $\hat{p} \pm 0.1$ where $\hat{p}$ is an estimator of p. (4)

ii) The shop in-charge is recording the inter-arrival time between consecutive consumers. She believes that the inter-arrival time X has an exponential distribution with parameter λ. A random sample of size 9 gives the values of inter arrival times as:

4 14 7 5 6 7 8 1 15

Obtain an exact 95% confidence interval for λ. (4)

[8]

Q. 9) A marketing manager of a large insurance company assumes that the number of contacts by its agents follow normal distribution with mean no more than 60. In order to verify his assumption a random sample of 36 agents have been selected. The sample mean and sample variance of the number of contacts are 68 and 144 respectively.

i) Perform a suitable test to verify his assumption about the mean at 10% level (4)

ii) Calculate the power of the above test at mean 64 and at mean 66. (4)

iii) Comment on the powers obtained in (ii) above. (1)

[9]

Q. 10) An actuarial student working in a life insurance company is analysing policy withdrawals on a sample of 2000 policies. She classified 500 policies with annual premium greater than or equal to INR 50,000 as large policies and remaining policies with annual premium less than INR 50,000 as small policies. She said, “Within a policy year, 20% of large policies have been withdrawn by policyholders whereas 30% of the small policies have been withdrawn by policyholders. Our aim in the next business plan is to reduce the small policies in order to reduce the policy withdrawals”.

Carry out a test to assess whether there is association between policy size (small / large) and policy withdrawals by policyholders at 1% level of significance. **[5]**

Q. 11) The following table shows the data on height (cm) and the weight (kg) of a group of nine people

Height X	142	155	147	188	165	128	166	161
Weight Y	52	60	49	69	60	44	55	56

The sample statistics are: $\overline{X}$= 159; $\overline{Y}$ = 55; S_{XX}= 2740; S_{YY} = 438 and S_{XY}= 782

i) Fit a Linear regression model: $Y_i = \alpha + \beta X_i + e_i$ for the above data, where $e_i's$ are i.i.d ~ N (0, σ^2) (2)

ii) Obtain 90% confidence limits for ρ, the underlying correlation coefficient. (3)

iii) Assuming the full normal model, obtain a 90% confidence interval for σ^2. (3)

iv) Mention four ways to check the fit of linear regression models. (4)

[12]

Q. 12) With the ongoing energy crisis, researchers have developed three methods for extracting oil from shale. Twelve bits of shale (of the same size) were randomly subjected to these three methods and the amount of oil (in litres per cubic meter) extracted from these methods are given below.

Method 1	**Method 2**	**Method 3**
3	4	5
2	2	2
1	3	4
2	2	5

i) Test whether the three methods differ significantly in the average amount of oil at 5% level of significance. State the assumptions made. (6)

ii) It is claimed by the researchers that Method 3 is superior to Method 1. Perform an appropriate test to verify the claim at 5% level of significance. (3)

[9]

****