INSTITUTE OF ACTUARIES OF INDIA

CT6 – STATISTICAL METHODS

OCTOBER 2009 EXAMINATION

INDICATIVE SOLUTION

Q 1) (i) Any four of the following (or other reasonable answer) should suffice.

(a) The company may be exposed to legal liability for the death of or bodily injury to a third party that could result from a faulty product.

A product liability cover can indemnify the insured against this risk.

(b) The company may be exposed to loss of or damage to its material property, including building, material assets and so on, from such perils as explosion, lightning, theft, storm, fire, flood, etc.

A property insurance can indemnify the company against this risk.

(c) The company is likely to be exposed to financial losses caused by dishonest actions by its employees (e.g., fraud, embezzlement, etc.). These would include loss of money or goods, either owned by the company or for which the company is responsible.

A fidelity guarantee insurance would cover the company for such losses.

(d) The pharmaceutical company may be exposed to losses made as a result of not being able to conduct business for various reasons.

A business interruption cover is appropriate for this risk.

(e) The company may be legally liable for the death of or bodily injury to a third party or for damage to property belonging to a third party, other than those liabilities covered by other liability insurance.

A public liability cover would indemnify the company against this risk.

(ii) The conditions are as follows.

(a) The policyholder must have an interest in the risk being insured.

(b) The risk must be of financial nature and should be reasonably quantifiable.

(c) Individual risks should be independent of one another.

(d) The probability that the insured event will occur should be small.

(e) The insurer's liability should be limited.

(f) There should be scope of pooling a large number of risks in order to reduce variability.

[7]

Q 2) We work with the unit of Rs. 1000's.

(i) Let X represent the original claim amount. The probability that a claim would involve the excess of loss reinsurer is

$$P(X > 80) = \int_{80}^{\infty} \frac{3(20)^3}{(x+20)^4}\,dx = \left(-\frac{(20)^3}{(x+20)^3}\right)\Bigg|_{80}^{\infty} = \left(\frac{20}{80+20}\right)^3 = \frac{1}{125} = 0.008.$$

(ii) The direct insurer's share of claim, net of excess of loss reinsurer's share, is

$$Y = \begin{cases} X & \text{if} \quad X \le 80, \\ 80 & \text{if} \quad X > 80. \end{cases}$$

The direct insurer's share of claim, net of excess of loss reinsurer and quota share reinsurer's shares, is $Z = 0.75Y = 0.75\min\{X, 80\}$.

$$E(Z)=\int_0^{\infty}0.75\min\{x,80\}\frac{3(20)^3}{(x+20)^4}dx=0.75\int_0^{80}x\frac{3(20)^3}{(x+20)^4}dx+0.75\int_{80}^{\infty}80\frac{3(20)^3}{(x+20)^4}dx$$

$$=0.75\int_0^{80}(x+20)\frac{3(20)^3}{(x+20)^4}dx-15\int_0^{80}\frac{3(20)^3}{(x+20)^4}dx+60\int_{80}^{\infty}\frac{3(20)^3}{(x+20)^4}dx$$

$$=22.5\int_0^{80}\frac{2(20)^2}{(x+20)^3}dx-15\left(-\frac{(20)^3}{(x+20)^3}\right)\Bigg|_0^{80}+60\left(-\frac{(20)^3}{(x+20)^3}\right)\Bigg|_{80}^{\infty}$$

$$=22.5\left(-\frac{(20)^2}{(x+20)^2}\right)\Bigg|_0^{80}-15\left(-\frac{(20)^3}{(x+20)^3}\right)\Bigg|_0^{80}+60\left(-\frac{(20)^3}{(x+20)^3}\right)\Bigg|_{80}^{\infty}$$

$$=22.5\left(1-\frac{(20)^2}{(80+20)^2}\right)-15\left(1-\frac{(20)^3}{(80+20)^3}\right)+60\left(\frac{(20)^3}{(80+20)^3}\right)=\frac{108}{5}-\frac{372}{25}+\frac{12}{25}=7.2.$$

Thus, the mean of direct insurer's share is Rs. 7,200.

[6]

Q 3) Expected aggregate loss over n days is $\sum_{i=1}^{n}p_i\mu_i=1000000\sum_{i=1}^{n}e^{-10-i}=1000000\times e^{-11}\frac{1-e^{-n}}{1-e^{-1}}$. As n goes to infinity, this reduces to $1000000e^{-11}/(1-e^{-1})$, which is only about Rs. 26.42.

[3]

Q 4) (i) Assigning letters to the three unknown figures for 1999, the incremental (non cumulative) claims are:

Claim payments (Rs. '000s)		Development year 0	1	2	3
Accident year	1997	32.7	19.7	x	4.7
	1998	41.6	y	12.9	*
	1999	z	37	*	6.7
	2000	40.2	32.1	11.1	6.7

So the cumulative amounts are:

Claim payments (Rs. '000s)		Development year 0	1	2	3
Accident year	1997	32.7	52.4	52.4+x	57.1+x
	1998	41.6	$41.6+y$	54.5+y	*
	1999	z	$37+z$	*	*
	2000	40.2	72.3	83.4	90.1

The factor relating to development years 2 to 3 is:

$$\frac{57.1+x}{52.4+x}=\frac{90.1}{83.4}.$$

$x=6.1.$

The factor relating to development years 1 to 2 is:

$$\frac{52.4 + x + 54.5 + y}{52.4 + 41.6 + y} = \frac{83.4}{72.3}.$$

$y = 29.8$.

The factor relating to development years 0 to 1 is:

$$\frac{52.4 + 41.6 + y + 37 + z}{32.7 + 41.6 + z} = \frac{72.3}{40.2}.$$

z = 34.

So the total amount of claims paid during the 1999 calendar year is

$x + y + z$ = 69.9, i.e., Rs. 69,900.

(ii) The cumulative claim incurred amount is as per the table below:

Claims incurred (Rs. '000s)		Earned premium (Rs. '000s)	Development year			
			0	1	2	3
Accident year	1996	1050	650	796	868	891
	1997	1135	685	837	921	
	1998	1180	704	848		
	1999	1260	722			

The development factors are as follows:

$f_{2:3}$ = 1.02650.

$f_{1:2}$ = 1.09553.

$f_{0:1}$ = 1.21677.

Assuming that AY 1996 is fully run off, the estimated loss ratio is

891/1050 = 0.84857.

We assume that the ultimate loss for AY 1996 is 891 and estimate the ultimate losses for the other accident years as follows:

<u>1997</u>

An initial estimate of the ultimate loss is 1135*0.84857 = 963.128.

Given this estimate and the above development factors, the IBNR is

963.128*(1 – 1/1.02650) = 24.8639.

The claims incurred to date is 921. Hence the revised estimate of the ultimate loss in respect of AY 1997 is 921+24.8639 = 945.86.

<u>1998</u>

An initial estimate of the ultimate loss is 1180*0.84857 = 1001.314.

Given this estimate and the above development factors, the IBNR is

1001.314*[1 – 1/(1.09553*1.02650)] = 110.91.

The claims incurred to date is 848. Hence the revised estimate of the ultimate loss in respect of AY 1998 is 848+110.91 = 958.91.

<u>1999</u>

An initial estimate of the ultimate loss is 1260*0.84857 = 1069.1999.

Given this estimate and the above development factors, the IBNR is

1069.1999*(1-1/(1.09553*1.02650*1.21677)) = 287.81.

The claims incurred to date is 722. Hence the revised estimate of the ultimate loss in respect of AY 1998 is 722+287.81 = 1009.81.

Summing the figures over all accident years gives an estimate of the total claims incurred of 3805.6, i.e., Rs. 3.8056m. Claims of Rs. 2.5m have been paid to date, so the estimated outstanding claims reserve is Rs. 1.3056m.

[14]

Q 5) (i) The likelihood is

$$p(X_1, X_2, \cdots, X_t|\theta) \propto \prod_{i=1}^{t} \exp\left[-\frac{(X_i-\theta)^2}{2\sigma^2}\right]$$

$$= \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{t}(X_i-\theta)^2\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{t}\left(\theta^2-2\theta X_i\right)\right]$$

$$= \exp\left[-\frac{1}{2\sigma^2}\left(t\theta^2-2\theta\sum_{i=1}^{t}X_i\right)\right].$$

The prior density of θ is $p(\theta) \propto \exp\left[-\frac{(\theta-\mu)^2}{2\tau^2}\right] \propto \exp\left[-\frac{1}{2\tau^2}\left(\theta^2-2\theta\mu\right)\right].$

Therefore, the posterior density is

$$p(\theta|X_1, X_2, \cdots, X_t) \propto p(X_1, X_2, \cdots, X_t|\theta)p(\theta)$$

$$\propto \exp\left[-\frac{1}{2\sigma^2}\left(t\theta^2-2\theta\sum_{i=1}^{t}X_i\right)-\frac{1}{2\tau^2}\left(\theta^2-2\theta\mu\right)\right]$$

$$= \exp\left[-\left(\frac{t}{2\sigma^2}+\frac{1}{2\tau^2}\right)\theta^2+2\left(\frac{\sum_{i=1}^{t}X_i}{2\sigma^2}+\frac{\mu}{2\tau^2}\right)\theta\right]$$

$$= \exp\left[-\frac{\theta^2-2\left(\frac{t}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}\left(\frac{\sum_{i=1}^{t}X_i}{\sigma^2}+\frac{\mu}{\tau^2}\right)\theta}{2\left(\frac{t}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}}\right]$$

$$\propto \exp\left[-\frac{\left\{\theta-\left(\frac{t}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}\left(\frac{\sum_{i=1}^{t}X_i}{\sigma^2}+\frac{\mu}{\tau^2}\right)\right\}^2}{2\left(\frac{t}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}}\right].$$

Thus, the posterior distribution is normal with mean $\left(\frac{t}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}\left(\frac{\sum_{i=1}^{t} X_i}{\sigma^2}+\frac{\mu}{\tau^2}\right)$ and variance $\left(\frac{t}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}$.

(ii) The Bayes estimate of θ, under the squared error loss function, is the mean of the posterior distribution, which is $\left(\frac{t}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}\left(\frac{\sum_{i=1}^{t} X_i}{\sigma^2}+\frac{\mu}{\tau^2}\right)$.

(iii) The Bayes estimate of θ can be rewritten as

$$\frac{\frac{t}{\sigma^2}\overline{X}+\frac{1}{\tau^2}\mu}{\frac{t}{\sigma^2}+\frac{1}{\tau^2}}, \text{ where } \overline{X}=\sum_{i=1}^{t} X_i.$$

This is of the form of a credibility estimate $Z\overline{X}+(1-Z)\mu$, where the credibility factor is

$$Z=\frac{\frac{t}{\sigma^2}}{\frac{t}{\sigma^2}+\frac{1}{\tau^2}}.$$

(iv) The Bayes estimate of θ, under the all-or-nothing loss function, is the mode of the posterior distribution. This is the same as the mean, $Z\overline{X}+(1-Z)\mu$, which has the form of a credibility estimate.

[9]

Q 6) (i)

Discount level	n-year premium if claim is made (A)	n-year premium if claim is not made (B)	Minimum loss for making a claim (A – B)	
0%	300 + 210 + 150(n – 2)	210 + 150(n – 1)	Rs. 150	[1]
30%	300 + 210 + 150(n – 2)	150n	Rs. 210	[1]
50%	210 + 150(n – 1)	150n	Rs. 60	[1]

Clearly, the threshold values of loss for making a claim for the 0%, 30% and 50% discount categories are Rs. 150, Rs. 210 and Rs. 60, respectively.

(ii) Let X denote the claim size. We know that a policyholder at 0% discount will make a claim if $X > 150$.

$P(X > 150) = P(\log X > 5.011)$

Since log X has Normal distribution with parameters $\mu = 6.012$ and $\sigma^2 = 1.792$,

$P(X > 150) = P(Z > -0.7477)$.

$= 0.773$.

Likewise, the probabilities in respect of policyholders in the other two discount classes are 0.690 and 0.924.

(iii) A policyholder paying full premium (0% discount level) has a loss incidence rate of 0.1 but he will only claim in respect of a fraction 0.773 of claims. Therefore, his claim occurrences follow another Poisson process with rate 0.0773 per year.

If N denotes the number of claims in a year for such a policyholder, then

$P(N=0)=\exp(-0.0773)=0.9256.$

$P(N\geq 0)=1-0.9256=0.0744.$

For a policyholder in the 30% discount category the corresponding probabilities are 0.9332 and 0.0667.

For a policyholder in the 50% discount category the corresponding probabilities are 0.9118 and 0.0883.

The required transition matrix is therefore

$$\begin{bmatrix} 0.0744 & 0.9256 & 0 \\ 0.0667 & 0 & 0.9338 \\ 0 & 0.0883 & 0.9117 \end{bmatrix}.$$

Let us denote the proportion of policyholders in NCD categories 0%, 30% and 50% in the stationary state by π_0, π_1 and π_2.

Then

$$\begin{pmatrix} \pi_0 & \pi_1 & \pi_2 \end{pmatrix}\begin{bmatrix} 0.0744 & 0.9256 & 0 \\ 0.0667 & 0 & 0.9333 \\ 0 & 0.0883 & 0.9117 \end{bmatrix}=\begin{pmatrix} \pi_0 & \pi_1 & \pi_2 \end{pmatrix}.$$

We get the following equations.

$\pi_0+\pi_1+\pi_2=1,$

$0.0744\pi_0+0.0667\pi_1=\pi_0,$

$0.9256\pi_0+0.0883\pi_2=\pi_1,$

$0.9333\pi_1+0.9117\pi_2=\pi_2.$

Solving these simultaneous equations, we obtain the required proportions:

$\pi_0=0.0062, \quad \pi_1=0.0858 \quad \pi_2=0.9080.$

[12]

Q 7) (i) A linear predictor is a linear function of covariates or independent variables (or transformed versions of these). Covariates may be numerical, e.g., age of driver or categorical, e.g., gender of driver. In a generalised linear model, the mean response is assumed to depend on the covariates only through this linear function.

(ii) (a) The linear predictor, when the area belongs to the i^{th} category and the class belongs to the j^{th} category, is

$\eta=\alpha_i+\beta_j,$

where α_i and β_j are unspecified parameters to be estimated from the data (i = 1,2,…,10; j = 1,2,…,6).

There are 15 parameters, excluding redundant parameters.

(b) The linear predictor, when the area belongs to the i^{th} category, the class belongs to the j^{th} category and the age is x, is

$\eta = \alpha_i + \beta_j + \gamma x + \delta_i x$,

where α_i, β_j, γ and δ_i are unspecified parameters to be estimated from the data (i = 1,2,…,10; j = 1,2,…,6).

The model can equivalently be written as

$\eta = \alpha_i + \beta_j + \gamma_i x$,

where α_i, β_j and γ_i are unspecified parameters to be estimated from the data (i = 1,2,…,10; j = 1,2,…,6).

[Either answer will do.]

There are 25 parameters, excluding redundant parameters.

[5]

Q 8) (i)

$$\gamma_k = Cov(X_t, X_{t-k})$$

$$= \frac{1}{(m+1)^2} Cov\left(\sum_{r=0}^{m} e_{t-r}, \sum_{s=0}^{m} e_{t-k-s}\right)$$

$$= \frac{1}{(m+1)^2} \sum_{r=0}^{m}\sum_{s=0}^{m} Cov(e_{t-k-r}, e_{t-s})$$

$$= \frac{1}{(m+1)^2} \sum_{r=0}^{m}\sum_{s=0}^{m} \sigma_e^2 I(r = k+s).$$

Clearly, if $|k| > m$, all terms are zero, so that $\gamma_k = 0$

For $0 \le |k| \le m$, there are exactly $(m - |k| + 1)$ non-zero terms out of the total of $(m+1)^2$ terms.

Thus,

$$\gamma_k = \frac{m - |k| + 1}{(m+1)^2}\sigma_e^2, \qquad \text{if } |k| = 0,1,2,\ldots,m;$$

$$= 0, \qquad \text{if } |k| > m.$$

The autocorrelation function is:

$$\rho_k = \frac{m - |k| + 1}{m+1}, \quad \text{if } |k| = 0,1,2,\ldots,m;$$

$$= 0, \qquad \text{if } |k| > m.$$

(ii) For the process to be invertible we require that the roots of the characteristic equation should be greater than 1 in absolute value.

We can rewrite the MA model with the aid of the backward shift operator B as follows:

$$X_t - \mu = \frac{1 + B + B^2}{3} e_t.$$

The roots of the characteristic equation $1 + B + B^2 = 0$ are

$$\frac{-1 + i\sqrt{3}}{2} \text{ and } \frac{-1 - i\sqrt{3}}{2}.$$

In both cases, $|B| = 1$. Thus, the process is not invertible.

[6]

Q 9) (i) The model can be written as:

$(1 - 1.7B + 0.4B^2 + 0.3B^3)\, X_t = (1 - 0.7B + 0.12B^2)\, e_t$.

The term in the bracket on the LHS is divisible by $1 - B$. We have

$(1 - 1.7B + 0.4B^2 + 0.3B^3) = (1 - B)(1 - 0.7B - 0.3B^2)$.

The last term is also divisible by $(1 - B)$, giving

$(1 - 1.7B + 0.4B^2 + 0.3B^3) = (1 - B)(1 - B)(1 + 0.3B) = (1 - B)^2 (1 + 0.3B)$.

Thus, the model has the representation:

$(1 + 0.3B)\, (1 - B)^2 X_t = (1 - 0.7B + 0.12B^2)\, e_t$.

(ii) The model can be identified as ARIMA(1,2,2).

(iii) The process $\{X_t : \ t = 0,1,2,\ldots\}$ is clearly a non-stationary process, as any ARIMA (p,d,q) process with $d > 0$ is non-stationary.

(iv) $(1 + 0.3B)\, W_t = (1 - 0.7B + 0.12B^2)\, e_t$.

$W_t = (1 + 0.3B)^{-1}(1 - 0.7B + 0.12B^2)\, e_t$.

By expanding the first term we get,

$W_t = (1 - 0.3B + 0.09B^2 + \ldots)(1 - 0.7B + 0.12B^2)\, e_t$.

$W_t = (1 - B + 0.42B^2 + \ldots)\, e_t$.

Therefore, $\psi_0 = 1$, $\psi_1 = -1$ and $\psi_2 = 0.42$.

(v) Since $d = 2$, the twice differenced series $W_t = (1 - B)^2 X_t = X_t - 2X_{t-1} + X_{t-2}$ is stationary (assuming the model has been identified correctly). The values W_{1001} and W_{1002} are forecast, in that order, from the fitted model using the equation

$\hat{W}_t = -0.3W_{t-1} + \hat{e}_t - 0.7\hat{e}_{t-1} + 0.12\hat{e}_{t-2}$, $\quad t = 1001, 1002,$

where W_{1000} is the observed value of $X_{1000} - 2X_{999} + X_{998}$; $\hat{e}_{1002} = \hat{e}_{1001} = 0$; $\hat{e}_{1000}$ and $\hat{e}_{999}$ are residuals obtained from fitting the series $W_3, W_4, \ldots, W_{1000}$; and W_{1001} is replaced by $\hat{W}_{1001}$.

Subsequently, the values X_{1001} and X_{1002} are forecast, in that order, from the fitted model using the equation

$\hat{X}_t = \hat{W}_t + 2X_{t-1} - X_{t-2}$, $\quad t = 1001, 1002,$

where X_{1001} is replaced by $\hat{X}_{1001}$.

[Full credit may be given in case one uses the model equation for X_t for forecasting, provided the substitutions are correctly specified, and the residuals are identified as those obtained from the fitting of the W_t series.]

[9]

Q 10) (i) The three integers to be chosen are:

- The multiplier – denoted here by a.
- The increment – denoted here by c. (It is often set to 0 to speed up the generation process.)

- The modulus – denoted here by m, where $m > a$ and $m > c$; the generator will produce a series of pseudo random integers in the range 0 to $m - 1$, so the modulus is usually set as a large number.

The recursive relationship is $x_{n+1} = (ax_n + c) \bmod m$, for $n = 0, 1, 2, \ldots$.

We then set $X_n = (2x_n/m - 1)$ for $n = 1, 2, \ldots$.

The desired sequence is $X_1, X_2, \ldots$.

(ii) Let $X = -(\log U)/5$.

Then, for $x > 0$, the cumulative distribution function is

$$F(x) = P[X \le x] = P[\log U \ge -5x] = P\left[U \ge e^{-5x}\right] = 1 - P\left[U < e^{-5x}\right] = 1 - e^{-5x}.$$

By differentiating this expression, we get $f(x) = 5\exp(-5x)$ as required.

Alternatively, use the inverse distribution function method:

$$F(x) = \int_0^x 5e^{-5y}\,dy = 1 - e^{-5x}.$$

We set $F(X) = U$, and solve for X. This gives $X = -(\log U)/5$.

(Half credit should be given for any valid justification, and half credit for correct formula.)

(iii) Use $f(x) = 5\exp(-5x)$ as the base density function. We need to find a constant c such that

$$\frac{ke^{-5x}}{1+x} \le c \cdot 5e^{-5x}.$$

This is accomplished when $c = k/5$.

The procedure is:

a. Generate a sample Y from the density function $f(x) = 5\exp(-5x)$.

b. Generate another pseudo random sample U from the uniform distribution over the interval [0,1].

c. If $U < g(Y)/(cf(Y))$, i.e., if $U < 1/(1 + Y)$, then we accept the value Y; otherwise we reject it and return to step a.

As the above procedure demonstrates, it is not necessary to compute k.

[7]

Q 11) (i) Revenue under different combinations of strategies and scenarios are as follows.

	Protection	Saving	Investment
Cautious	50 x 10,000 = **500,000**	100 x 10,000 = **1,000,000**	150 x 10,000 = **1,500,000**
Best estimate	50 x 20,000 = **1,000,000**	100 x 20,000 = **2,000,000**	150 x 20,000 = **3,000,000**
Optimistic	50 x 30,000 = **1,500,000**	100 x 30,000 = **3,000,000**	150 x 30,000 = **4,500,000**

Costs under different combinations of strategies and scenarios are as follows.

	Protection	Saving	Investment
Cautious	2,000,000 + 300,000	2,000,000 + 600,000	2,000,000 + 900,000

	= **2,300,000**	= **2,600,000**	= **2,900,000**
Best estimate	2,000,000 + 300,000 = **2,300,000**	2,000,000 + 600,000 = **2,600,000**	2,000,000 + 900,000 = **2,900,000**
Optimistic	2,000,000 + 300,000 = **2,300,000**	2,000,000 + 600,000 = **2,600,000**	2,000,000 + 900,000 = **2,900,000**

Thus, loss under different combinations of strategies and scenarios are as follows.

	Protection	Saving	Investment
Cautious	2,300,000 – 500,000 = **1,800,000**	2,600,000 – 1,000,000 = **1,600,000**	2,900,000 – 1,500,000 = **1,400,000**
Best estimate	2,300,000 – 1,000,000 = **1,300,000**	2,600,000 – 2,000,000 = **600,000**	2,900,000 – 3,000,000 = **– 100,000**
Optimistic	2,300,000 – 1,500,000 = **800,000**	2,600,000 – 3,000,000 = **– 400,000**	2,900,000 – 4,500,000 = **– 1,600,000**

(ii) Maximum loss occurs when the Cautious estimate is realized. The maximum loss is minimized if the company chooses the Investment sector.

(iii) Expected loss (risk) under the three schemes:

Protection: 0.15 x 1,800,000 + 0.75 x 1,300,000 + 0.10 x 800,000
= 1,325,000.

Saving: 0.15 x 1,600,000 + 0.75 x 600,000 – 0.10 x 400,000
= 650,000.

Investment: 0.15 x 1,400,000 – 0.75 x 100,000 – 0.10 x 1,600,000
= – 25,000.

The risk is also minimized if the company chooses the Investment sector. This is the Bayes strategy.

[Full credit is given if any candidate does not make any additional computation for part (iii), but simply says that the strategy of choosing the Investment sector dominates the other two strategies, and hence is optimal under any criterion.]

[7]

Q 12) (i) The moment generating function is $M_S(t) = \exp[\lambda(M_X(t)-1)]$. The skewness is the third moment of the cumulant generating function $\log M_S(t)$ [page 222 of Core Reading], evaluated at 0.

$$E\left[(S-E(S))^3\right] = \frac{\partial^3}{\partial t^3}\log M_S(t)\bigg|_{t=0} = \frac{\partial^3}{\partial t^3}\lambda(M_X(t)-1)\bigg|_{t=0} = \lambda\frac{\partial^3}{\partial t^3}M_X(t)\bigg|_{t=0} = \lambda E(X^3).$$

Since X is a non-negative random variable, its third moment is positive, and hence the skewness of S is positive.

Alternatively,

$$E\left[(S-E(S))^3\right]=E\left[S^3-3S^2E(S)+3S\{E(S)\}^2-\{E(S)\}^3\right].$$

The moments of S are computed as follows.

$$E(S)=\frac{\partial}{\partial t}M_S(t)\Big|_{t=0}=\frac{\partial}{\partial t}\exp[\lambda(M_X(t)-1)]\Big|_{t=0}=\exp[\lambda(M_X(t)-1)]\lambda M'_X(t)\big|_{t=0}=\lambda m_1\,.$$

$$E(S^2)=\frac{\partial^2}{\partial t^2}M_S(t)\Big|_{t=0}=\frac{\partial}{\partial t}\{\exp[\lambda(M_X(t)-1)]\lambda M'_X(t)\}\Big|_{t=0}$$

$$=\exp[\lambda(M_X(t)-1)]\left[\lambda^2\{M'_X(t)\}^2+\lambda M''_X(t)\right]\Big|_{t=0}=\lambda^2m_1^2+\lambda m_2\,.$$

$$E(S^3)=\frac{\partial^3}{\partial t^3}M_S(t)\Big|_{t=0}=\frac{\partial}{\partial t}\left\{\exp[\lambda(M_X(t)-1)]\left[\lambda^2\{M'_X(t)\}^2+\lambda M''_X(t)\right]\right\}\Big|_{t=0}$$

$$=\exp[\lambda(M_X(t)-1)]\left[\lambda^3\{M'_X(t)\}^3+\lambda^2M''_X(t)M'_X(t)+2\lambda^2M'_X(t)M''_X(t)+\lambda M'''_X(t)\right]\Big|_{t=0}$$

$$=\lambda^3m_1^3+3\lambda^2m_1m_2+\lambda m_3\,. \qquad [1]$$

After substitution of these moments, the expression for $E\left[(S-E(S))^3\right]$ simplifies as follows.

$$E\left[(S-E(S))^3\right]=\left(\lambda^3m_1^3+3\lambda^2m_1m_2+\lambda m_3\right)-3\left(\lambda^2m_1^2+\lambda m_2\right)\lambda m_1+3\lambda^3m_1^3-\lambda^3m_1^3=\lambda m_3\,.$$

Since X is a non-negative random variable, its third moment is positive, and hence the skewness of S is positive.

(ii) The expression for the skewness of S is

$$\frac{E\left[(S-E(S))^3\right]}{(Var(S))^{3/2}}=\frac{\lambda m_3}{(\lambda m_2)^{3/2}}=\frac{m_3}{\left(\sqrt{\lambda m_2^3}\right)}\,.$$

Clearly, $\lim_{\lambda\to\infty}E\left[(Z-E(Z))^3\right]=\lim_{\lambda\to\infty}\frac{m_3}{\left(\sqrt{\lambda m_2^3}\right)}=0\,.$

Thus, even though the distribution of Z is positively skewed for finite λ, the skewness vanishes as $\lambda\to\infty$.

Alternation solution:

For every finite λ, the distribution of the normalized statistic $Z=[S-E(S)]/\sqrt{Var(S)}$ is positively skewed because the distribution of S is positively skewed.

The limiting distribution of Z (as $\lambda\to\infty$) is symmetric, and so the skewness of that distribution is zero. Since S and Z have the same coefficient of skewness, it follows that the limiting value of the skewness of the distribution of S is also zero.

[In part (ii) of the Question 12, S was misprinted as Z. Full credit should be given to anyone who gives a reasonable answer assuming (a) that Z is S, or (b) that Z is the normalized version of S, or (c) that the question as stated cannot be answered.]

[5]

Q 13) (i) Let insurer's surplus at the end of day t be U_t. Clearly,

$$U_t = U_0 + ct - \sum_{i=1}^{N(t)} X_i,$$

where $N(t)$ is the number of claim till day t, and X_1, X_2, … are the successive claim amounts.

Let the day of the first claim be T. This has a geometric distribution. On the other hand, since T is the smallest value of t such that $N(t) = 1$, we have $U_T = U_0 + cT - X_1$. The probability of ruin at first claim is

$$P(U_T < 0) = P(U_0 + cT - X_1 < 0) = P(X_1 > U_0 + cT)$$

$$= \sum_{t=1}^{\infty} P(X_1 > U_0 + cT | T = t) P(T = t) = \sum_{t=1}^{\infty} P(X_1 > U_0 + ct) P(T = t)$$

$$= \sum_{t=1}^{\infty} e^{-\lambda(U_0 + ct)} (1-p)^{t-1} p \qquad = \sum_{s=0}^{\infty} e^{-\lambda(U_0 + c + cs)} (1-p)^s p$$

$$= pe^{-\lambda(U_0 + c)} \sum_{s=0}^{\infty} \left\{(1-p)e^{-\lambda c}\right\}^s$$

$$= \frac{pe^{-\lambda(U_0 + c)}}{1 - (1-p)e^{-\lambda c}}.$$

(ii) By setting $\dfrac{0.01e^{-0.0125(U_0+10)}}{1 - 0.99e^{-0.125}} \le 0.01$, we get $U_0 \ge \dfrac{\log\left(1 - 0.99e^{-0.125}\right)}{0.0125} - 10 = 155.51$, i.e., the minimum required initial surplus is Rs. 155.51.

(iii) Note that $N(t)$ has a binomial distribution, and $X_1 + X_2$ has a gamma distribution. The required probability is

$$P(U_2 < 0) = P\left(U_0 + 2c - \sum_{i=1}^{N(2)} X_i < 0\right) = \sum_{n=0}^{2} P\left(U_0 + 2c - \sum_{i=1}^{N(2)} X_i < 0 \middle| N(t) = n\right) P(N(t) = n)$$

$$= P(U_0 + 2c < 0) P(N(t) = 0) + P(U_0 + 2c - X_1 < 0) P(N(t) = 1)$$
$$+ P(U_0 + 2c - (X_1 + X_2) < 0) P(N(t) = 2)$$

$$= P(X_1 > U_0 + 2c) P(N(t) = 1) + P(X_1 + X_2 > U_0 + 2c) P(N(t) = 2)$$

$$= e^{-\lambda(U_0 + 2c)} 2p(1-p) + \left(\int_{U_0 + 2c}^{\infty} \lambda^2 x e^{-\lambda x} dx\right) p^2$$

$$= e^{-\lambda(U_0 + 2c)} 2p(1-p) + \left(\int_{(U_0 + 2c)\lambda}^{\infty} y e^{-y} dy\right) p^2$$

$$= e^{-\lambda(U_0 + 2c)} 2p(1-p) + ((U_0 + 2c)\lambda + 1) e^{-(U_0 + 2c)\lambda} p^2$$

$= 0.0022.$

[10]

[Total 100 Marks]

****