Institute of Actuaries of India

Subject CT6 – Statistical Methods

May 2011 Examination

INDICATIVE SOLUTION

Introduction

The indicative solution has been written by the Examiners with the aim of helping candidates. The solutions given are only indicative. It is realized that there could be other points as valid answers and examiners have given credit for any alternative approach or interpretation which they consider to be reasonable.

Question 1

For a risk to be insurable:

- the policyholder must have an interest in the risk being insured, to distinguish between insurance and a wager,
- a risk must be of a financial and reasonably quantifiable nature.

Other desirable criteria:

- Individual risk events should be independent of each other.
- The probability of the event should be relatively small.
- There should be scope of pooling a large numbers of potentially similar risks in order to reduce the variance and hence achieve more certainty.
- There should be an ultimate limit on the liability undertaken by the insurer.
- It should be possible to eliminate moral hazards.

Total – 4 Marks

Question 2

Let the ACFs be ρ_1 and ρ_2, and the PACFs be ϕ_1, ϕ_2, ϕ_3,… . We have
$\phi_1 = \rho_1 = 0.8,$

$$\phi_2 = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} = \frac{0.46 - 0.8^2}{1-0.8^2} = \frac{0.46-0.64}{1-0.64} = -\frac{0.18}{0.36} = -0.5,$$

Since the process is AR(2), $\phi_i = 0$ for $i \geq 3$.

Total – 3 Marks

Question 3

(i) $f(x) = \frac{d}{dx} P(X \leq x) = \frac{d}{dx}\left[1 - \left(1 + \frac{\alpha}{\beta}x\right)^{-1/\alpha}\right] \Rightarrow f(x) = \frac{1}{\beta}\left(1+\frac{\alpha}{\beta}x\right)^{-1/\alpha - 1}, x > 0$

(ii) $E(X) = \int_0^\infty x f(x)dx = \frac{1}{\beta}\int_0^\infty x\left(1+\frac{\alpha}{\beta}x\right)^{-1/\alpha-1} dx$

Use the substitution $1 + \frac{\alpha}{\beta}x = t^\alpha$.

$\Rightarrow \frac{\alpha}{\beta}dx = \alpha t^{\alpha-1}dt$; $x = 0$ corresponds to $t = 1$ and $x = \infty$ corresponds to $t = \infty$. Also, the stated restriction on the value of α implies that $t^{\alpha-1}$ tends to 0 as $t \to \infty$.

$$E(X) = \int_1^\infty \frac{\beta}{\alpha}\left(t^\alpha - 1\right)t^{-2}dt = \frac{\beta}{\alpha}\left[\left.\frac{t^{\alpha-1}}{\alpha-1}\right|_1^\infty + \left.t^{-1}\right|_1^\infty\right] = \frac{\beta}{1-\alpha}.$$

$$E(X^2) = \int_0^\infty x^2 f(x)dx = \frac{1}{\beta}\int_0^\infty x^2\left(1+\frac{\alpha}{\beta}x\right)^{-1/\alpha-1}dx$$

Once again, we substitute t^{α} for $1+\frac{\alpha}{\beta}x$, and note that the stated restriction on the value of α implies that $t^{2\alpha-1}$ tends to 0 as $t \to \infty$.

$$E(X^2) = \int_1^{\infty} \frac{\beta^2}{\alpha^2}(t^{\alpha}-1)^2 t^{-2}dt = \int_1^{\infty} \frac{\beta^2}{\alpha^2}(t^{2\alpha-2}+t^{-2}-2t^{\alpha-2})dt$$

$$= \frac{\beta^2}{\alpha^2}\left[\left.\frac{t^{2\alpha-1}}{2\alpha-1}\right|_1^{\infty} - \left.t^{-1}\right|_1^{\infty} - \left.\frac{2t^{\alpha-1}}{\alpha-1}\right|_1^{\infty}\right] = \frac{\beta^2}{\alpha^2}\left(\frac{-1}{2\alpha-1}+1+\frac{2}{\alpha-1}\right) = \frac{2\beta^2}{(1-2\alpha)(1-\alpha)}.$$

Equating first and second order moments with those of the sample observations

$$E(X) = \frac{\beta}{(1-\alpha)} = 0.6. \Rightarrow \beta = 0.6(1-\alpha).$$

$$E(X^2) - E^2(X) = Var(X) = 2. \Rightarrow E(X^2) = 2 + 0.6^2 = 2.36.$$

$$\Rightarrow \frac{2\beta^2}{(1-2\alpha)(1-\alpha)} = 2.36. \Rightarrow \frac{2\times 0.6^2}{(1-2\alpha)(1-\alpha)} = 2.36. \Rightarrow \alpha = 0.41.$$

$$\Rightarrow \beta = 0.6(1-\alpha) = 0.354 .$$

Total – 8 Marks

Question 4

The farmer's expected profit in the three scenarios from each fruit (in 1000 rupees) is

	Warm and wet (Θ_1)	Cool (Θ_2)	Warm and dry (Θ_3)
Apples	100	70	90
Pears	60	90	105
Oranges	75	100	50

(i) Maximum profit on growing apples = 100
Maximum profit on growing pears = 105
Maximum profit on growing oranges = 100
Hence, maximum possible profit = 105, which is on growing pears.

(ii) Probability of occurrence of $\Theta_3 = p$.
Let probability of occurrence of Θ_1 and Θ_2 be x.
Then, $x + x + p = 1 \Rightarrow x = (1 - p)/2$.

Expected profit on growing apples = $100\frac{(1-p)}{2}+70\frac{(1-p)}{2}+90p = 85+5p$.

Expected profit on growing pears = $60\frac{(1-p)}{2}+90\frac{(1-p)}{2}+105p = 75+30p$.

Expected profit on growing oranges = $75\frac{(1-p)}{2}+100\frac{(1-p)}{2}+50p = 87.5-37.5p$.

Extreme points of the three linear functions (expected profit) at $p = 0$ and $p = 1$

	Apples	Pears	Oranges
$p = 0$	85	75	87.5
$p = 1$	90	105	50

Apples are more profitable than pears when $85+5p > 75+30p$, i.e., when $p < 0.4$.

Apples are more profitable than oranges when $85+5p > 87.5-37.5p$, i.e., when $p > 0.059$.

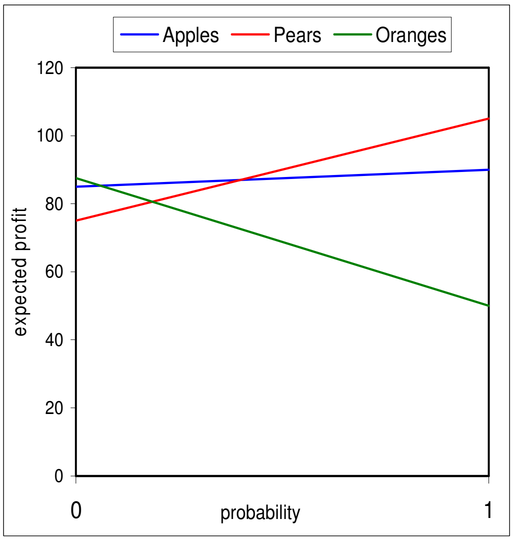
At $p = 0.059$ oranges and apples are equally profitable and;

At $p = 0.4$ pears and apples are equally profitable.

It follows that in order to maximize profit, the farmer should grow

Oranges if $0 \le p \le 0.059$,
Apples if $0.059 \le p \le 0.4$,
Pears if $0.4 \le p \le 1$.

This fact also becomes clear from the plot of the profits from each crop as a function of *p*, shown below.



(iii) Each scenario is equally likely when p = ⅓ = 0.333. In this case, the solution to part suggests the farmer grows apples.

Total Marks - 7

Question 5

Let λ be the rate of claim (Poisson process). The rate of premium is $c = \lambda(1+\theta)m_1$.
The upper bound on the adjustment coefficient follows from the observation

$$\lambda + cR = \lambda M_X(R) = \lambda\int_0^\infty e^{Rx} f(x)dx > \lambda\int_0^\infty \left(1+Rx+\frac{R^2x^2}{2}\right) f(x)dx = \lambda\left[1+Rm_1+\frac{R^2m_2}{2}\right]$$

$$\Rightarrow (c-\lambda m_1)R > \frac{1}{2}R^2 m_2 \lambda \quad \Rightarrow R < \frac{2(c-\lambda m_1)}{\lambda m_2}$$

$$\therefore R < \frac{2\theta m_1}{m_2}.$$

On the other hand,

$$e^{Rx} = 1 + Rx + \frac{R^2x^2}{2!} + \frac{R^3x^3}{3!} + \cdots$$

$$\le 1 + Rx + \frac{R^2Mx}{2!} + \frac{R^3M^2x}{3!} + \cdots = \frac{x}{M}\left(1 + RM + \frac{R^2M^2}{2!} + \frac{R^3M^3}{3!} + \cdots\right) + 1 - \frac{x}{M} = \frac{x}{M}e^{RM} + 1 - \frac{x}{M}.$$

Now,

$$\lambda + cR = \lambda M_X(R) = \lambda\int_0^M e^{Rx} f(x)dx \le \lambda\int_0^M \left(\frac{x}{M}e^{RM} + 1 - \frac{x}{M}\right) f(x)dx = \frac{\lambda}{M}e^{RM}m_1 + \lambda - \frac{\lambda}{M}m_1$$

$$\Rightarrow cR \le \frac{\lambda m_1}{M}\left(e^{RM} - 1\right)$$

$$\Rightarrow \frac{c}{\lambda m_1} \le \frac{1}{RM}\left(1 + RM + \frac{R^2M^2}{2!} + \cdots - 1\right) = 1 + \frac{RM}{2!} + \frac{R^2M^2}{3!} + \cdots < 1 + RM + \frac{R^2M^2}{2!} + \cdots = e^{RM}$$

$$\Rightarrow R > \frac{1}{M}\log\left(\frac{c}{\lambda m_1}\right) = \frac{1}{M}\log\left(\frac{\lambda m_1(1+\theta)}{\lambda m_1}\right) = \frac{1}{M}\log(1+\theta).$$

Total Marks - 6

Question 6

(i) The outstanding claims according to the chain-ladder method are given in the last column of the bottom table.

Cumulative claim amounts reported

	Development Year >>>				Earned	Paid
AY	0	1	2	3	Premiums	to-date
2008	500	1500	1700	1800	2000	1690
2009	700	1900	2000		2500	1700
2010	600	1400			3000	800
2011	1200				3300	100
Age-to Age	2.6667	1.0882	1.0588	1.0000		
CLDF	3.0727	1.1522	1.0588	1.0000		

AY	Reported Loss	CLDF	Reported %	CL Ult Loss	Paid to-date	CL O/S Payments
2008	1800	1.0000	100.0%	1800.00	1690	110.00
2009	2000	1.0588	94.4%	2117.65	1700	417.65
2010	1400	1.1522	86.8%	1613.15	800	813.15
2011	1200	3.0727	32.5%	3687.20	100	3587.20
	6400			9217.99	4290	4927.99

(ii) The outstanding claims according to the B-F method are given in the last column of the bottom table.

Years more than 90% developed: 2008 and 2009	
CL Ultimate Loss for 2008 + 2009 =	3917.65
Earned Premiums for 2008 + 2009 =	4500.00
IELR for BF	87.06%

	Earned Premiums	Initial exptd LR	Initial exptd Ult	BF Ult Loss	BF O/S Payments
AY	2000	87.06%	1741.18	1800.00	110.00
2008	2500	87.06%	2176.47	2120.92	420.92
2009	3000	87.06%	2611.76	1745.10	945.10
2010	3300	87.06%	2872.94	3137.94	3037.94
2011	10800		9402.35	8803.95	4513.95

(iii) The BF and the CL estimates are similar for older years that are sufficiently developed. AY 2008 is fully developed and there is no difference between the BF and the CL estimate. As we move towards more recent years, the reported percentage declines and the BF estimate moves away from the CL estimate.

This is as expected because the BF estimate is based on a credibility approach that weighs the chain-ladder estimate using the reported percentage for each year and assigns the rest of the weight to the initial expected loss. Hence, the more the year is developed, the greater is the weight given by the BF approach to the CL estimate.

However, the CL estimate for 2011 is much higher than the BF estimate for the same year. This is because this year has suffered a major loss early on and the chain-ladder method has projected this poor experience into the future. This is contrary to the BF approach which assumes the future is not impacted by the losses observed so far.

Total Marks - 9

Question 7

(i) $\overline{X}_{Fire} = (20+12+25+36)/4 = 23.25$. Likewise, $\overline{X}_{Motor} = 181.25$, $\overline{X}_{Marine} = 61.25$.

It follows that $E[m(\theta)] = \overline{X} = \frac{1}{3}\sum_{i=1}^{3}\overline{X}_i = 88.583$.

On the other hand, for Fire ($i = 1$),

$$\frac{1}{n-1}\sum_{j=1}^{n}(X_{ij} - \overline{X}_i)^2 = \frac{1}{3}\left[\left(20-23.25^2\right)+\left(12-23.25^2\right)+\left(25-23.25^2\right)+\left(36-23.25^2\right)\right]$$
$$= 100.917.$$

The corresponding sample variance for Motor (i = 2) and Marine (i = 3) are 3906.25 and 68.917, respectively. It follows that

$$E\left[s^2(\theta)\right] = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{n-1}\sum_{j=1}^{n}(X_{ij} - \overline{X}_i)^2\right\} = \frac{1}{3}\left(100.917 + 3906.25 + 68.917\right) = 1358.694.$$

$$Var[m(\theta)] = \frac{1}{N-1}\sum_{i=1}^{N}(\overline{X}_i - \overline{X})^2 - \frac{1}{Nn}\sum_{i=1}^{N}\left\{\frac{1}{n-1}\sum_{j=1}^{n}\left(X_{ij} - \overline{X}_i\right)^2\right\}$$

$$= \frac{1}{2}\left[(23.25 - 88.583)^2 + (181.25 - 88.583)^2 + (61.25 - 88.583)^2\right] - \frac{1}{4}\times 1358.694$$

$$= 6801.333 - 339.674 = 6461.66.$$

So the credibility factor is given by: $z = \dfrac{n}{\left(n + \dfrac{E(s^2(\theta))}{Var(m(\theta))}\right)} = \dfrac{4}{4 + \dfrac{1358.694}{6461.66}} = 0.95006$

$\therefore$ EBCT model 1 expected claim payment (in Rs.'000s) for

Fire $= (23.25 \times 0.95006) + (1 - 0.95006) \times 88.583 \approx 26.513$;
Motor $= (181.25 \times 0.95006) + (1 - 0.95006) \times 88.583 \approx 176.622$;
Marine $= (61.25 \times 0.95006) + (1 - 0.95006) \times 88.583 \approx 62.615$.

(ii) Assumptions for each of the three risks indexed as i = 1, 2, 3:

- The distribution of each X_{ij}, j = 1, 2, 3, 4, depends on the value of a parameter, say θ_i, whose value is fixed (and the same for each value of j) but is unknown.
- Given θ_i, the Xij 's, j = 1, 2, 3, 4, are independent and identically distributed.

Assumption about the relationship between different risks:

- For $i \neq k$, the pairs (θ_i, X_{ij}) and (θ_k, X_{km}) are independent and identically distributed.

Total Marks - 7

Question 8

Let X be a typical claim amount and the retention limit be k. Clearly, $k > 5$.
The expected part of X ceded to the reinsurer is

$$E[(X-k)I(X>k)] = \int_k^\infty (x-k)\frac{1}{5}\exp[-(x-5)/5]dx$$

$$= \exp[-(k-5)/5]\int_k^\infty (x-k)\frac{1}{5}\exp[-(x-k)/5]dx = \exp[-(k-5)/5]\int_0^\infty y\frac{1}{5}\exp[-y/5]dy$$

$$= 5\exp[-(k-5)/5].$$

Also, $E(X) = \int_5^\infty x \times \frac{1}{5}\exp[-(x-5)/5]dx = \int_0^\infty (5y+5)\exp[-y]dy = 5 + 5 = 10.$

Expected part of X retained by insurer is

$$E(X) - E[(X-k)I(X>k)] = 10 - 5\exp[-(k-5)/5].$$

Expected annual profit (P)
= Expected annual premium income, net of premium ceded to reinsurer
– Cost of cover, net of cover received from reinsurer

$= 1.15\times10\times1000-1.30\times5\exp[-(k-5)/5]\times1000-10\times1000+5\exp[-(k-5)/5]\times1000$

$= 0.15\times10\times1000-0.30\times5\exp[-(k-5)/5]\times1000$

$= 1500\times[1-\exp\{-(k-5)/5\}]$.

Conversely, $k = 5\times\left[1-\ln\left(1-\frac{P}{1500}\right)\right]$.

Using these relations, the table can be completed as under.

Retention limit	Expected annual profit
8	676.78
20	1425.32
∞	1500

Total Marks - 6

Question 9

(i) $f(\theta|x)$.is proportional to the product of the prior and the likelihood, i.e.,

$$f(\theta|x)\propto f(x|\theta)\times g(\theta)=\frac{1}{\theta}I(x<\theta)\times\theta\exp(-\theta)=\exp(-\theta)I(\theta>x).$$

In order that the integral of the density over θ is 1, the above expression has to be divided by the constant $\exp(-x)$.

Thus, the posterior density is $f(\theta|x)=\exp[-(\theta-x)]I(\theta>x)$.

The Bayes estimator for θ with respect to the absolute error loss function, based on a single observation X, is the median of this density, which is obtained by solving for m from the equation

$$\frac{1}{2}=\int_x^m\exp[-(\theta-X)]d\theta=1-\exp[-(m-X)].$$

The solution is $m = X + \ln(2)$.

(ii) The overall expected loss is

$$E\left|\theta-kX\right|^2=E_\theta\left[E_{X|\theta}\left|\theta-kX\right|\right]=E_\theta\left[\int_0^{\theta/k}\frac{1}{\theta}(\theta-kx)dx+\int_{\theta/k}^{\theta}\frac{1}{\theta}(kx-\theta)dx\right]$$

$$=E_\theta\left(\frac{\theta}{k}-\frac{\theta}{2k}+\frac{k\theta}{2}-\frac{\theta}{2k}-\theta+\frac{\theta}{k}\right)=E_\theta(\theta)\times\left(\frac{k}{2}+\frac{1}{k}-1\right).$$

The second factor depends only on k. Its derivative is $\frac{1}{2}-\frac{1}{k^2}$, which is negative for $0<k<\sqrt{2}$ and positive for $k>\sqrt{2}$. Thus, the overall expected loss is minimized by choosing $k=\sqrt{2}$, irrespective of the distribution of θ.

(iii) The prior mean is $\int_0^\infty \theta \exp(-\theta) d\theta = 1$.

An estimate based solely on the single data from the risk itself is $\sqrt{2}X$.
However, there is no value of a credibility factor *Z*, such that the estimator *X* + ln(2) can be expressed as $Z \times (\sqrt{2}X) + (1-Z) \times 1$. Therefore, the Bayes estimate given is part (i) is NOT a credibility estimate.

Total Marks – 10

Question 10

20th July will fall in the 3rd week if it is a Saturday or Sunday, otherwise it will fall in the 4th week. Therefore, *all* the claim sizes arising on this day are 3 lacs with probability 2/7, and *all* the claim sizes are 4 lacs with probability 5/7. No other combination of values is possible.

When all the claim sizes are 3 lacs (which happens with probability 2/7), the total claim amount is less than or equal to 7 lacs if and only if the number of claims is less than or equal to 2.

When all the claim sizes are 4 lacs (which happens with probability 5/7), the total claim amount is less than or equal to 7 lacs if and only if the number of claims is less than or equal to 1.

Therefore,
P(total claim amount is less than or equal to 7)
= P(total claim amount is less than or equal to 7|20th is in 3rd week) x P(20th is in 3rd week)
+ P(total claim amount is less than or equal to 7|20th is in 4th week) x P(20th is in 4th week)

$$= P(N \le 2) \times \frac{2}{7} + P(N \le 1) \times \frac{5}{7} = P(N \le 1) + P(N = 2) \times \frac{2}{7}$$

$$= \exp(-5) \times \left(\frac{5^0}{1} + \frac{5^1}{1} \right) + \exp(-5) \times \frac{5^2}{2} \times \frac{2}{7} = \exp(-5) \times \left(1 + 5 + \frac{25}{7} \right) = \exp(-5) \times \frac{67}{7}$$

= 0.0645.

Total Marks - 6

Question 11

(i) Let *X* be the amount of a claim.

Then $X \sim LN(\mu, \sigma^2)$ such that $E(X) = e^{\left(\mu + \frac{1}{2}\sigma^2\right)} = 250$ and $V(X) = e^{(2\mu + \sigma^2)}\left(e^{\sigma^2} - 1\right) = 375^2$.
Solving for the parameters μ and σ^2, we have

$$e^{\sigma^2} - 1 = \frac{375^2}{250^2} \qquad \Rightarrow \sigma^2 = \ln\left(\frac{375^2}{250^2} + 1 \right) = 1.1786;$$

$$\mu + \frac{1}{2}\sigma^2 = \ln(250) \quad \Rightarrow \mu = 4.9321.$$

Let Y_1 and Y_2 denote the amount of the claim paid by the insurer under proportional cover and XOL cover respectively. Let Z_1 and Z_2 denote those paid by the reinsurer.

$$Z_2 = \begin{cases} 0 & , X < 300 \\ X - 300 & , X \ge 300 \end{cases}$$

$$\Rightarrow E(Z_2) = \int_{300}^{\infty} (x - 300) f(x) dx = \int_{300}^{\infty} x f(x) dx - 300 \int_{300}^{\infty} f(x) dx = I_1 - 300 I_2$$

$$I_1 = \int_{300}^{\infty} x.\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2} dx$$. Use the substitution $u = \left(\frac{\ln x - \mu}{\sigma}\right) - \sigma$.

$\Rightarrow du = \frac{1}{x\sigma} dx$ and $x = \exp(\sigma u + \sigma^2 + \mu)$.

$$I_1 = \int_{\frac{\ln 300-\mu}{\sigma}-\sigma}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(u+\sigma)^2} .\sigma.e^{(\sigma u+\sigma^2+\mu)} du = \int_{\frac{\ln 300-\mu}{\sigma}-\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \times e^{(\mu+\frac{1}{2}\sigma^2)}$$

$$= \left[1 - \Phi\left(\frac{\ln 300 - \mu}{\sigma} - \sigma\right)\right] \times e^{(\mu+\frac{1}{2}\sigma^2)}$$

$$I_2 = P(X > 300) = P(\ln X > \ln 300) = P\left(N\left(\mu, \sigma^2\right) > \ln 300\right)$$

$$= P\left[N(0,1) > \frac{\ln 300 - \mu}{\sigma}\right] = 1 - \Phi\left(\frac{\ln 300 - \mu}{\sigma}\right)$$

Therefore,

$$E(Z_2) = I_1 - 300 I_2 = 250\left[1 - \Phi\left(\frac{\ln 300 - \mu}{\sigma} - \sigma\right)\right] - 300\left[1 - \Phi\left(\frac{\ln 300 - \mu}{\sigma}\right)\right].$$

$$= 250[1 - \Phi(-0.375)] - 300[1 - \Phi(0.711)] = 90.3525$$

.

100α% ceded under the proportional cover equate $E(Y_1)$ and $E(Y_2)$.
Now, $E(X) = E(Y_1) + E(Z_1)$ and $E(X) = E(Y_2) + E(Z_2)$
=> $E(Y_1) = E(Y_2)$ if and only if $E(Z_1) = E(Z_2)$.
The requisite condition is $250\alpha = 90.3525$, which corresponds to $100\alpha\% = 36.141\%$.

(ii) Under XOL reinsurance, profit from portfolio to Alpha
= Premium income net of reinsurance premium ceded – cost of cover net of reinsurance
= $(1000\times 100 - 35{,}000) - 0.2\times 1000\times(250 - 90.3525) = 65{,}000 - 31{,}929.50 = 33{,}070.50$.
Under proportional reinsurance, profit from portfolio to Alpha
= Premium income net of reinsurance premium ceded – cost of cover net of reinsurance
= $1000\times 100\times(1 - 0.36141) - 0.2\times 1000\times 250\times(1 - 0.36141) = 63{,}859 - 31{,}929.50$.
= $31{,}929.50$.
Profit under XOL is greater.

Total Marks - 8

Question 12

(i) The ratio is

$$\frac{f(x)}{\phi(x)} = \begin{cases} \dfrac{\sqrt{3}\times\sqrt{2\pi}}{4\left(1 + x^2/2\right)^{3/2}\exp\left(-\frac{1}{2}x^2\right)} & for\ |x| \le 2, \\ 0 & for\ |x| > 2. \end{cases}$$

The function is strictly positive in the range $|x| \le 2$. Therefore, we can maximize the log of the ratio, which is an easier problem.

$$\frac{\partial}{\partial x}\ln\left(\frac{f(x)}{\phi(x)}\right) = -\frac{3}{2}\times\frac{x}{(1+x^2/2)} + x = x\left(1-\frac{3}{2+x^2}\right) = \frac{x(x-1)(x+1)}{2+x^2}.$$

This function turns
from negative to positive at $x = -1$,
from positive to negative at $x = 0$,
from negative to positive at $x = 1$.

Thus, the original ratio $\frac{f(x)}{\phi(x)}$ has
a minimum at $x = -1$,
a maximum at $x = 0$,
a minimum at $x = 1$.

Thus, we need to evaluate the ratio at $x = 0$ and at the end-points of the interval $[-2,2]$, and pick the maximum.

It is easy to see that

$$\frac{f(2)}{\phi(2)} = \frac{f(-2)}{\phi(-2)} = \frac{\sqrt{6\pi}e^2}{4(3)^{3/2}} = \frac{e^2\sqrt{2\pi}}{12};$$

$$\frac{f(0)}{\phi(0)} = \frac{\sqrt{6\pi}}{4} < \frac{e^2\sqrt{2\pi}}{12}.$$

The last inequality is verified numerically. Thus, the maximum value of the ratio is as stated in the question.

(ii) The following steps may be used.

1. Generate U_1 from uniform (0,1).
2. Compute $Z = \Phi^{-1}(U_1)$.
3. If $|Z| > 2$, go to Step 1.
4. Generate U_2 from uniform (0,1).
5. Compute $C = \frac{e^2\sqrt{2\pi}}{12}$, $Z_2 = \frac{\sqrt{3}}{4(1+Z^2/2)^{3/2}}$, $Z_3 = \frac{e^{-Z^2/2}}{\sqrt{2\pi}}$.
6. If $U_2 > \frac{Z_2}{CZ_3}$, go to Step 1; else return Z.

(iii) The proportion of accepted samples
= probability of acceptance in Step 3
x conditional probability of further acceptance in Step 6
= $(\Phi(2)-\Phi(-2))\times\frac{1}{C} = 0.9545\times 0.6479 = 0.6184$.

Total Marks - 10

Question 13

(i) The model is
$(X_n - X_{n-1}) = \alpha(X_{n-1} - X_{n-2}) + \varepsilon_n$, ε_n's are uncorrelated, with mean 0 and variance σ^2.
Since the mean of the process is known to be 0, the usual estimator of the parameter α is

$$\hat{\alpha} = \hat{\rho}_1 = \frac{\sum_{i=3}^{200}(X_i - X_{i-1})(X_{i-1} - X_{i-2})}{\sum_{i=2}^{200}(X_i - X_{i-1})^2} = \frac{587.83}{936.49} = 0.6277.$$

The estimator of autocovariance at lag 0 is

$$\hat{\gamma}_0 = \frac{1}{200}\sum_{i=2}^{200}(X_i - X_{i-1})^2 = \frac{936.49}{200} = 4.6825.$$

Using the relation $\gamma_0 = \frac{\sigma^2}{1-\rho_1^2} = \frac{\sigma^2}{1-\alpha^2}$, we estimate

$$\hat{\sigma}^2 = (1-\hat{\alpha}^2)\hat{\gamma}_0 = (1-0.6277^2)\times 4.6825 = 2.8376 .$$

(ii) The forecast of x_{201} is obtained from
$(\hat{x}_{200}(1) - x_{200}) = \hat{\alpha}(x_{200} - x_{199}) = 0.6277\times(1.93-0.82) = 0.6967$.
Thus, $\hat{x}_{200}(1) = x_{200} + 0.6967 = 1.93 + 0.6967 = 2.6267$.

Total Marks – 7

Question 14

(i) The exponential density $f(y) = \frac{1}{\mu}e^{-\frac{y}{\mu}}$ for $y > 0$ can be written in the general form of the exponential family as $\exp\left[\frac{y\theta - b(\theta)}{a(\phi)} - c(y,\phi)\right]$, with $\theta = 1/\mu$, $b(\theta) = \ln(\theta)$, $\phi = 1$, $a(\phi) = -1$ and $c(y,\phi) = 0$.

Since $\theta = 1/\mu$, the canonical link function is the reciprocal function.

Therefore, for $i = 1,2,\ldots,n,$ we have

$$\frac{1}{\mu_i} = \frac{1}{E(Y_i|X_i)} = \alpha + \beta X_i .$$

The likelihood is $\prod_{i=1}^{n}\left(\frac{1}{\mu_i}\exp\left(-\frac{Y_i}{\mu_i}\right)\right) = \prod_{i=1}^{n}\left[(\alpha+\beta X_i)\exp\{-Y_i(\alpha+\beta X_i)\}\right]$.

The log-likelihood for the GLM parameters is

$$L(\alpha,\beta) = \sum_{i=1}^{n}\ln(\alpha+\beta X_i) - \sum_{i=1}^{n}Y_i(\alpha+\beta X_i).$$

(ii) Setting the derivatives of $L(\alpha,\beta)$ with respect to α and β equal to 0, we have

$$\sum_{i=1}^{n} \frac{1}{\alpha+\beta X_i} = \sum_{i=1}^{n} Y ,$$

$$\sum_{i=1}^{n} \frac{X_i}{\alpha+\beta X_i} = \sum_{i=1}^{n} Y_i X_i .$$

(iii) The equations reduce to

$$\frac{1}{\alpha+40\beta}+\frac{1}{\alpha+50\beta}+\frac{1}{\alpha+60\beta} = 250+1000+250 ,$$

$$\frac{40}{\alpha+40\beta}+\frac{50}{\alpha+50\beta}+\frac{60}{\alpha+60\beta} = 250\times 40+1000\times 50+250\times 60 .$$

We eliminate the middle term on the left hand side by multiplying the first equation by 50 and subtracting the second from it, to get

$$\frac{10}{\alpha+40\beta}-\frac{10}{\alpha+60\beta} = 250\times 10-250\times 10 = 0 .$$

It follows that $\alpha+40\beta = \alpha+60\beta$, i.e., $\hat{\beta}=0$.

The likelihood maximized with respect to α and β is the same with and without the constraint β = 0.

Therefore, the scaled deviance is 0, and the hypothesis β = 0 cannot be rejected. There is no evidence that age has an effect on claim size.

Total Marks – 9

[Total Marks 100]

****