Institute of Actuaries of India

Subject CT6 – Statistical Methods

May 2015 Examinations

INDICATIVE SOLUTIONS

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 examiner have given credit for any alternative approach or interpretation which they consider to be reasonable.

Solution 1:

(i) Let Y denotes the net claim payable by the insurer.

Hence $Y = \begin{cases} X & ,X < M \\ M + 0.2(X - M) & ,X \geq M \end{cases}$

Now the net premium income by the insurer is given by,

$C_{net} = \lambda\,(1 + 0.25)\text{E(X)} - \lambda(1 + 0.6)\text{E(Z)}$

$= 1.25\lambda.\dfrac{1}{\beta} - 1.6\lambda \displaystyle\int_M^{\infty} 0.8(x - M)\beta e^{-\beta x}dx$

$= 1.25\lambda.\dfrac{1}{\beta} - 1.28\lambda[\dfrac{e^{-\beta M}}{\beta}]$

$= \frac{\lambda}{\beta}[1.25 - 1.28e^{-\beta M}]$

Now the net claim payable by the insurer is = $\lambda\, E(Y)$

$E(Y) = \int_0^M xf(x)dx + \int_M^{\infty}[M + 0.2(x - M)]f(x)dx$

$= \displaystyle\int_0^M x\beta e^{-\beta x}dx + 0.8M\int_M^{\infty} \beta e^{-\beta x}dx + 0.2\int_M^{\infty} x\beta e^{-\beta x}dx$

$= [-xe^{-\beta x}]_0^M - [\frac{1}{\beta}e^{-\beta x}]_0^M + 0.8Me^{-\beta M} + 0.2[\,[-xe^{-\beta x}]_M^{\infty} - [\frac{1}{\beta}e^{-\beta x}]_M^{\infty}]$

$= -Me^{-\beta M} - \dfrac{1}{\beta}e^{-\beta M} + \dfrac{1}{\beta} + 0.8Me^{-\beta M} + 0.2[Me^{-\beta M} + \dfrac{1}{\beta}e^{-\beta M}]$

$= \frac{1}{\beta}(1 - 0.8e^{-\beta M})$

The insurer will not make any loss if

$\lambda\frac{1}{\beta}\left[1.25 - 1.28\, e^{-\beta M}\right] \geq \lambda\frac{1}{\beta}(1 - 0.8e^{-\beta M})$

Or $0.48\, e^{-\beta M} \leq 0.25$

So, $M \geq 65.23$ **[5]**

(ii) $M_Y(R) = E(e^{RY})$

$= \int_0^M e^{Rx}\beta e^{-\beta x}dx + \int_M^{\infty} e^{R(0.8M+0.2x)}\beta e^{-\beta x}dx$

$= \frac{\beta - \beta e^{-M(\beta-R)}}{\beta-R} + \beta e^{08M} \int_M^\infty e^{-(\beta-0.2R)x} dx$

$= \frac{\beta - \beta e^{-M(\beta-R)}}{\beta - R} + \frac{\beta e^{-M(\beta-R)}}{\beta - 0.2R}$

$= \frac{\beta}{\beta-R}[1 - \frac{0.8R}{\beta-0.2R} e^{-M(\beta-R)}]$

[3]

(iii) Equation of adjustment coefficient is

$\lambda + C_{net}R = \lambda M_Y(R)$

Or $\lambda + \frac{\lambda}{\beta}[1.25 - 1.28\, e^{-\beta M}]R = \frac{\lambda\beta}{\beta-R}[1 - \frac{0.8R}{\beta-0.2R} e^{-M(\beta-R)}]$

Or $-\beta R + (\beta - R)[1.25 - 1.28\, e^{-\beta M}]R = -\frac{0.8R\beta^2}{\beta-0.2R} e^{-M(\beta-R)}]$

Or $-\beta(\beta - 0.2R) + (\beta - 0.2R)(\beta - R)[1.25 - 1.28\, e^{-\beta M}] + 0.8e^{MR}e^{-M\beta}\beta^2 = 0$

Or $0.8e^{MR}\beta^2 - \beta e^{M\beta}(\beta - 0.2R) + (\beta - 0.2R)(\beta - R)[1.25e^{M\beta} - 1.28] = 0$

[3]

(iv) Using M = 50 and β = 0.01 and $e^x = 1 + x + \frac{x^2}{2}$ we get

$1 + 50\,R + 1250\,R^2 - 206.09(0.01 - 0.2R) + (0.0001 - 1.2\beta R + 0.2R^2)[9761.27] = 0$

Solving for R and taking smallest +ve root we get R = 0.01059

[3]

(v) Here M = 0, then we get the revised equation as

$0.8\beta^2 - \beta^2 + 0.2\beta R - 0.03\,(\beta^2 - 1.2\beta R + 0.2R^2) = 0$

Solving we get, R = 0.01, taking smallest +ve root

[2]

(vi) Initial surplus U = 100

Using Lundberg's Inequality $\psi(U) \leq e^{-RU}$

Hence, maximum ruin probability under (d) = exp(-0.01059*100) = 34.68%

And maximum ruin probability under (e) = exp(-0.01 * 100) = 36.78%

We can see that if the retention level reduces then the ruin probability increases. **[2]**

[17 Marks]

Solution 2:

(i) I : Inox, P: PVR, C: Cinemax, F: Fun Cinemas, T: Watches Theatre

P(I) = 2/6, P(P) = 1/6, P(C) = 1/6, P(F) = 2/6

P(T|I) = 7/10, P(T|P) = 3/10, P(T|C) = 5/10, P(T|F) = 8/10

$$P(T) = P(I)P(T|I) + P(P)P(T|P) + P(C)P(T|C) + P(F)P(T|F)$$
$$= \frac{2}{6}\frac{7}{10} + \frac{1}{6}\frac{3}{10} + \frac{1}{6}\frac{5}{10} + \frac{2}{6}\frac{8}{10} = \frac{38}{60} = \frac{19}{30}$$

$$P(\overline{T}) = \frac{11}{30}$$

Now

$$P(I|\overline{T}) = \frac{P(I\overline{T})}{P(\overline{T})} = \frac{P(I)P(\overline{T}|I)}{P(\overline{T})} = \frac{\frac{2}{6}(1-\frac{7}{10})}{\frac{11}{30}} = \frac{6}{22}$$

Similarly

$$P(P|\overline{T}) = \frac{7}{22}$$
$$P(C|\overline{T}) = \frac{5}{22}$$
$$P(F|\overline{T}) = \frac{4}{22}$$

So it is most likely that he has been to PVR. **[4]**

(ii) The number of Sundays on which he watches theatre is a random variable X having Binomial distribution B(3, 19/30).

Hence the required probability =P(X = 2) + P(X = 3)

= $3*(19/30)^2 * (1 – 19/30) + (19/30)^3$ = 0.695 **[2]**

(iii) The probability that they meet on a given Sunday in the theatre

$$=(\frac{2}{6}\frac{7}{10})^2+(\frac{1}{6}\frac{3}{10})^2+(\frac{1}{6}\frac{5}{10})^2+(\frac{2}{6}\frac{8}{10})^2=\frac{27}{200}$$

Hence the probability that they fail to meet over two weekends = $(1 – 27/200)^2$

So, the probability that they meet at least once = $1 – (1 – 27/200)^2$ = 0.251775 **[3]**

[9 marks]

Solution 3:

(i) The lowest (worst) profit under each of the three possibilities of “masala dosa”, “noodles” and “Pizza” are 850, 800 and 500 respectively.

The highest (best) profit among them is 850.

Hence the best of the worst possible case criterion solution is to choose the strategy to sell “masala dosa”.

The maximax solution is to choose a strategy which maximizes the maximum profit under the three scenarios.

The maximum profit under each of the three possibilities of “masala dosa”, “noodles” and “Pizza” are 1200, 1500 and 1400 respectively. Hence the strategy which will maximize the maximum profit is to sell “noodles”. **[2]**

(ii) Under Bayes Criterion we will select the strategy which gives the maximum expected profit.

Probability (Low footfall) =1/4

Probability (normal footfall) =1/2

Probability (High footfall) =1/4

The expected profit under each of the three strategies is:

Expected Profit (Masala Dosa) = 850*1/4 + 950*1/2 + 1200*1/4 = 987.5

Expected Profit (Noodles) = 900*1/4 + 800*1/2 + 1500*1/4 = 1000

Expected Profit (Pizza) = 500*1/4 + 875*1/2 + 1400*1/4 = 912.5

Thus the strategy selected under Bayes Criterion is to choose to sell “noodles”. **[2]**

[4 Marks]

Solution 4:

The aggregate claim (S) from an individual health policy can take any one of the following values over the coming year:

0, 100000, 200000, 300000, 400000

We need to compute Probability (S<=s) where s can take any of the values mentioned above

Thus,

P (S<=0) = P (Number of claims = 0) = P (N=0) = 0.7

P (S<=100000) = P(S=0) + P (S=100000) = 0.7 + 0.2*0.7 = 0.84

P (S<=200000) = P(S=0) + P(S=100000) + P(S=200000) = P(S<=100000) + P(S=200000)

= 0.84 + 0.1*0.7*0.7 + 0.2*0.3

= 0.84 + 0.109 = 0.949

P (S<=300000) = P(S=0) + P(S=100000) + P(S=200000) + P(S=300000) = P(S<=200000) + P(S=300000)

= 0.949 + 2*0.1*0.7*0.3

= 0.949 + 0.042

= 0.991

Similarly,

P (S<=400000) = P(S<=300000) + P(S=400000)

= 0.991 + 0.1*0.3*0.3

= 0.991 +0.009

= 1 **[7 Marks]**

Solution 5:

(i)

(a) The sample median of our claims data is the 6^{th} observation out of the 11 observed values.

Thus the median based on the sample data is 80000.

Assume M to be the population median.

Then M will satisfy the equation

$1 - \exp(-\mu M) = ½$ ………… (i)

Substituting the value of the sample median in equation (i) we have

$1 - \exp(-80000\mu) = ½$

- ⇨ $\exp(-80000\mu) = ½$

  Taking log on both sides and solving we get

- ⇨ $-80000\mu = \ln(1/2)$
- ⇨ $\mu = - \ln(1/2)/80000$
- ⇨ $\mu = -1*-0.000008664$
- ⇨ $\mu = 0.000008664$

Hence the estimate of μ using the method of percentiles equals 0.000008664 **[3]**

(b) The likelihood of observing the 7 known claims and 4 unknown claims greater than 100000 is

$L(\mu) = f(x_1).f(x_2).f(x_3)……..f(x_7) * P(X>100000)^4$

$= \mu\exp(-\mu x_1).\ \mu\exp(-\mu x_2).\ \mu\exp(-\mu x_3)……………….\ \mu\exp(-\mu x_7).\ * (\exp(-\mu 100000)^4)$

$= \mu^7\exp(-\mu\sum x_i)\ .\ \exp(-400000\mu)$

$= \mu^7\exp(-386645\mu)\ .\ \exp(-400000\mu)$

$= \mu^7\exp(-786645\mu)$

By taking logarithm on both sides we get

$\text{Log } L(\mu) = 7 \log \mu - 786645\mu$ ………………………… (i)

To find out the mle we need to differentiate equation (i) w.r.t μ.

Thus we get

$d \text{ Log } L(\mu)/d\mu = 7/\mu - 786645$ ………………………… (ii)

To find out maxima or minima we set the derivative to zero.

Hence we get

$7/\mu = 786645$

⇨ $\mu = 7/786645 = 0.000008899$

To find out if this estimator is a maximum we differentiate equation (ii) again with respect to μ and observe that

$D^2 \text{ Log } L(\mu)/d\mu^2 = -7/\mu^2$

which is less than zero.

Thus the maximum likelihood estimate of μ is 0.000008899 **[5]**

(ii)

(a) Let X denote the random variable representing the gross claim amount which follows a pare to (α,150000) distribution.

Let Y be the random variable which represents the amount paid by the insurer on a claim.

As a policy excess of 10000 is in force the claim payments made by the insurer follow a conditional distribution.

The probability density function of Y is

$s(y) = f(x)/(P(X>10000), y>0 \ \& \ x = y + 10000$

Now Probability $(X>10000) = (150000/(150000+10000))^{\alpha}$, since X ~ pareto(α,150000)

⇨ $P(X>10000) = (150000/160000)^{\alpha}$ ……………………………….. (i)

Now the numerator is $f(x) = f(y+10000)$

$$= \alpha * 150000^{\alpha}/(150000 + y + 10000)^{\alpha+1}$$

$$= \alpha * 150000^{\alpha}/(160000 + y)^{\alpha+1}$$

Thus $s(y) = \alpha * 150000^{\alpha}/(160000 + y)^{\alpha+1} \div (150000/160000)^{\alpha}$

$$= \alpha * 160000^{\alpha}/(160000 + y)^{\alpha+1}$$

which is the pdf of the conditional claim distribution and whose form is of a pareto distribution with parameters (α, 160000) **[4]**

(b) The likelihood function $L(\alpha) = s(y_1).\ s(y_2).\ s(y_3).\ s(y_4).\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ s(y_{10})$.

Thus $L(\alpha) = \alpha * 160000^{\alpha} /(160000 + y_1)^{\alpha+1} .\ \alpha * 160000^{\alpha} /(160000 + y_2)^{\alpha+1} .\alpha * 160000^{\alpha} /(160000 + y_3)^{\alpha+1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .\ \alpha * 160000^{\alpha} /(160000 + y_{10})^{\alpha+1}$

$L(\alpha) = \prod \alpha * 160000^{\alpha}/ /(160000 + y_i)^{\alpha+1}$

$= \alpha^{10} * 160000^{10\alpha}/\prod (160000 + y_i)^{\alpha+1}$

Taking logarithm on both sides we get,

$\text{Log } L(\alpha) = \log(\alpha^{10} * 160000^{10\alpha}/\prod (160000 + y_i)^{\alpha+1})$

$= 10 \log \alpha + 10\ \alpha * \log 160000 – (\alpha+1)\sum (\log(160000+ y_i)$

To find out the mle we differentiate the above equation with respect to α

$d\log L(\alpha)/d\alpha = 10/ \alpha + 10 \log 160000 - \sum (\log(160000+ y_i)$

and set it equal to zero.

Thus,

$0 = 10/ \alpha + 10 \log 160000 - \sum (\log(160000+ y_i)$

- ⇨ $\sum (\log(160000+ y_i) = 10/ \alpha + 10 \log 160000$
- ⇨ $\sum (\log(160000+ y_i) - 10 \log 160000 = 10/ \alpha$
- ⇨ $\alpha = 10 / (\sum (\log(160000+ y_i) - 10 \log 160000) = 10 / (125.0217 – 119.8292) = 1.925$

To find out if the estimator found out is a maximum we take a second derivative of the log likelihood function with respect to α.

Thus we get

$d^2\log L(\alpha)/d\alpha^2 = -10/\alpha^2 < 0 = \max$

Hence the maximum likelihood estimate of α is 1.925 (approx 2). **[5]**

[17 Marks]

Solution 6:

(i) Adjusting past inflation for claim paid amount we get,

	Development Year		
Incidence Year	0	1	2
2012	49,612.50	33,993.75	24,375.00
2013	84,288.75	57,000.00	
2014	72,800.00		

Now Cumulative claim paid is given by,

	Development Year		
Incidence Year	0	1	2
2012	49,612.50	83,606.25	107,981.25
2013	84,288.75	141,288.75	
2014	72,800.00		

Now cumulative no of claims is given by,

	Development Year		
Incidence Year	0	1	2
2012	50	85	110
2013	95	155	
2014	80		

So the average cost per claim is given by,

	Development Year		
Incidence Year	0	1	2
2012	992.25	983.60	981.65
2013	887.25	911.54	
2014	910.00		

Now using the development factors we have completed the lower triangle for the above two table as given below:

No of claims:

	Development Year		
Incidence Year	0	1	2
2012	50.00	85.00	110.00
2013	95.00	155.00	200.59
2014	80.00	132.41	171.36

DF for Year 1 = (85+155)/(50+95) = 1.65517, for Year 2 = 110/85 = 1.294118

Average cost per claim

	Development Year		
Incidence Year	0	1	2
2012	992.25	983.60	981.65
2013	887.25	911.54	909.73
2014	910.00	917.57	915.75

Development factor for Year 1 = 1.00832, and for Year 2 = 0.99801

So total claim cost per claim can be calculated using the below table (after **adjusting future inflation**):

Average Cost per claim	Total No of claims	Total Claim
981.65	110.00	107,981.25
1,000.70	200.59	200,728.89
1,108.06	171.36	189,875.67
		498,585.80

Claims already paid = 72,800 + 1,41,288.75 + 1,07,981.25 = 3,22,070

Hence outstanding reserve requirement = 498,585.80 – 3,22,070 = 1,76,515.80 **[7]**

(ii) Assumptions:

- First year fully run-off
- The average cost per claim in each development year is a constant proportion in monetary terms of the ultimate average cost per claim for each incidence year

- The number of claims in each development year is a constant proportion of the ultimate no of claims for each incidence year **[1]**

(iii) We will project the no of claims and average cost per claim using the calculated development factors.

No of claims projected:

	Development Year		
Incidence Year	0	1	2
2012	50.00	82.76	107.10
2013	95.00	157.24	
2014	80.00		

Average cost per claim projected:

	Development Year		
Incidence Year	0	1	2
2012	992.25	1,000.51	998.52
2013	887.25	894.63	
2014	910.00		

So, total cumulative claim paid is given by,

	Development Year		
Incidence Year	0	1	2
2012	49,612.50	82,800.71	106,940.86
2013	84,288.75	140,673.59	
2014	72,800.00		

So the non-cumulative claim paid is given by,

	Development Year		
Incidence Year	0	1	2
2012	49,612.50	33,188.21	24,140.15
2013	84,288.75	56,384.84	
2014	72,800.00		

Now taking the difference between the predicted and actual claim paid we get:

	Development Year		
Incidence Year	0	1	2
2012	-	(805.54)	(234.85)
2013	-	(615.16)	
2014	-		

Expressing the above difference as % of actual claim paid we get,

	Development Year		
Incidence Year	0	1	2
2012	0.00%	-2.37%	-0.96%
2013	0.00%	-1.08%	
2014	0.00%		

We can conclude that the model is reasonably fit.

The differences are not very high with a maximum absolute difference between predicted and actual is less than 2.5%. **[4]**

(iv) Each development factor is log-normally distributed. Hence their product is also log-normally distributed.

The development factor to ultimate for Incidence Year 2014 is log-normally distributed with parameters as given below:

Average cost per claim: $\mu_1 = 0.001675 + 0.001675 = 0.00335, \sigma_1 = 0.03825 + 0.03825 = 0.0765$
No of claims: $\mu_2 = 0.190225 + 0.190225 = 0.38045, \sigma_1 = 0.454755 + 0.454755 = 0.9095$

Equation for outstanding claim for Incidence Year 2014 can be written as:

$910*1.1^2$ * exp(0.00335 + Z * 0.0765)*80*exp(0.38045 + Z * 0.9095) – 72,800

So the required probability can be found from below equation:

1,50,000 ≤ 88,088 exp(0.3838 + Z * 0.986) – 72,800 ≤ 2,00,000

Or, 0.55186 ≤ Z ≤ 0.7572

Hence, Z = 6.61% **[6]**

[18 Marks]

Solution 7:

(i)

$$X_t = A_t + B_t$$
$$A_t = 0.5A_{t-1} + 0.5B_{t-1} + e_t^{(A)}$$
$$B_t = 0.7A_{t-1} - 0.7A_{t-2} + e_t^{(B)}$$

Using matrix notation we get,

$$Y_t = \begin{pmatrix} A_t \\ B_t \end{pmatrix} = \begin{pmatrix} 0.5 0.5 \\ 0.7 0 \end{pmatrix}\begin{pmatrix} A_{t-1} \\ B_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -0.7 & 0 \end{pmatrix}\begin{pmatrix} A_{t-2} \\ B_{t-2} \end{pmatrix} + \begin{pmatrix} e_t^{(A)} \\ e_t^{(B)} \end{pmatrix}$$

Or, $Y_t = \begin{pmatrix} 0.5 0.5 \\ 0.7 0 \end{pmatrix} Y_{t-1} + \begin{pmatrix} 0 & 0 \\ -0.7 & 0 \end{pmatrix} Y_{t-2} + \begin{pmatrix} e_t^{(A)} \\ e_t^{(B)} \end{pmatrix}$

The Eigen values of first matrix are given by the following equation,

$\lambda^2 - 0.5\lambda - 0.35 = 0$

Hence, $|\lambda| < 1$

Similarly for second matrix, $|\lambda| < 1$

As all the Eigen values are less than 1, hence the process Y_t is stationary. **[3]**

(ii)

a) $X_t = (\alpha + 1)X_{t-1} - (\alpha + 0.25\alpha^2)X_{t-2} + 0.25\alpha^2 X_{t-3} + e_t$

Or, $X_t - X_{t-1} = \alpha\,(X_{t-1} - X_{t-2}) - 0.25\alpha^2(X_{t-2} - X_{t-3}) + e_t$

Or, $Y_t = \alpha\, Y_{t-1} - 0.25\alpha^2 Y_{t-2} + e_t,\ \ assuming\ Y_t = X_t - X_{t-1}$

So X_t is ARIMA(2,1,0) process if it is I(1)

Now, $(1 - \alpha B + 0.25\alpha^2 B^2)\ Y_t = e_t$

The characteristic equation is, $1 - \alpha\lambda + 0.25\alpha^2\lambda^2 \quad = 0,\ or\ \lambda = \frac{2}{\alpha}$

To meet stationary condition, $|\lambda| > 1, or\ |\alpha| < 2$

Hence X_t is ARIMA(2,1,0) process with $|\alpha| < 2$ **[2]**

b) Now Cov(Y_t , e_t) = Cov(e_t , e_t) = σ^2

$\gamma_0 = Cov(Y_t, Y_t) = \alpha\gamma_1 - 0.25\,\alpha^2\gamma_2 + \sigma^2$ ……….(1)

Taking co-variances with Y_{t-1} , Y_{t-2} and Y_{t-k} we ge

$\gamma_1 = \alpha\gamma_0 - 0.25\,\alpha^2\gamma_1$ ……………(2)

$\gamma_2 = \alpha\gamma_1 - 0.25\,\alpha^2\gamma_0$ ……………..(3)

$\gamma_k = \alpha\gamma_{k-1} - 0.25\,\alpha^2\gamma_{k-2}$

From (2), $\gamma_1 = \frac{\alpha\gamma_0}{(1+0.25\,\alpha^2)}$ ……………(4)

Substituting (3) in (1) we get,

$\gamma_0 = \alpha\gamma_1 - 0.25\,\alpha^2(\alpha\gamma_1 - 0.25\,\alpha^2\gamma_0) + \sigma^2$

Or, $(1 - (0.25\,\alpha^2)^2\,)\gamma_0 = \frac{\alpha(1-0.25\,\alpha^2)\alpha}{1+0.25\,\alpha^2}\gamma_0 + \sigma^2$

Or, $\gamma_0 = \frac{(1+0.25\,\alpha^2)}{(1-0.25\,\alpha^2)^3}\sigma^2$

Hence, $\gamma_1 = \frac{\alpha}{(1+0.25\,\alpha^2)}\frac{(1+0.25\,\alpha^2)}{(1-0.25\,\alpha^2)^3}\sigma^2 = \frac{\alpha}{(1-0.25\,\alpha^2)^3}\sigma^2$

For k ≥ 2, $\gamma_k = \alpha\gamma_{k-1} - 0.25\,\alpha^2\gamma_{k-2}$ ……….(5)

Now γ_k follows the below equation

$\lambda_1 + k\lambda_2 = (0.5\alpha)^{-k}\ \gamma_k$ ………………… (6)

Substituting k by k-1 and k-2 we get

$\lambda_1 + (k-1)\lambda_2 = (0.5\alpha)^{-(k-1)}\ \gamma_{k-1}$

$\lambda_1 + (k-2)\lambda_2 = (0.5\alpha)^{-(k-2)}\ \gamma_{k-2}$

Substituting the above two equation in eqn (5) we get,

$\gamma_k = \alpha(0.5\alpha)^{(k-1)}[\lambda_1 + (k-1)\lambda_2] - 0.25\,\alpha^2(0.5\alpha)^{(k-2)}[\lambda_1 + (k-2)\lambda_2]$

Or, $\gamma_k = \lambda_1\alpha^k\,(0.5)^{(k-1)}\left[1 - \frac{0.25}{0.5}\right] + \lambda_2\alpha^k\,(0.5)^{(k-1)}[(k-1) - \frac{1}{2}(k-2)]$

Or, $\gamma_k = \lambda_1\alpha^k\,(0.5)^k + k\lambda_2\alpha^k\,(0.5)^k$

Or, $\lambda_1 + k\lambda_2 = (0.5\alpha)^{-k}\ \gamma_k$

Which is the original equation. Hence γ_k follows equation (6)

Now putting k = 0 we get,

$$\lambda_1 = \gamma_0 = \frac{(1+0.25\,\alpha^2)}{(1-0.25\,\alpha^2)^3}\sigma^2$$

Putting k = 1 we get,

$$\lambda_2 = \frac{1}{0.5\alpha}\gamma_1 - \lambda_1 = \frac{\sigma^2}{(1-0.25\,\alpha^2)^2}$$ **[11]**

c) α = 0.04, hence $Y_t = 0.04\,Y_{t-1} - 0.0004Y_{t-2} + e_t$

$Y_t = X_t - X_{t-1}$ *Or* $X_t = Y_t + X_{t-1}$

Since $x_1, x_2, \ldots\ldots, x_{50}$ are observed values

$X_{51} = Y_{51} + X_{50}$

$X_{52} = Y_{52} + X_{51}$

So the forecasted values are

$x_{51} = y_{51} + x_{50}$ and $x_{52} = y_{52} + x_{51}$

Where

$y_{51} = 0.04\,(x_{50} - x_{49}) - 0.0004(x_{49} - x_{48})$

And $y_{52} = 0.04\,y_{51} - 0.0004(x_{50} - x_{49})$ **[2]**

[18 Marks]

Solution 8:

(i)

The Likelihood function based on previous year's claims data is

L(q) = $\binom{3500}{p}$. q^p. $(1-q)^{3500-p}$, as the policies are independent and maximum permissible claim under a single policy is 1

Taking logarithm on both sides we get ,

Log L(q) = constant + p log q + (3500-p). log(1-q)

Differentiating with respect to q we get,

$d \text{ Log } L(q)/dq = p/q - (3500-p)/(1-q)$

Equating the above equation to zero and rearranging we have

$p/q = (3500-p)/(1-q)$

- ⇨ $p(1-q) = q.(3500-p)$
- ⇨ $p - pq = 3500q - pq$
- ⇨ $p = 3500q$
- ⇨ $q = p/3500$

which gives the estimate of q using the method of mle.

Posterior distribution of q

We know that the posterior distribution is proportional to prior*likelihood ……(i)

The prior distribution of q is Beta (α,β)

Thus we have $f(q) = \Gamma(\alpha+\beta)^* q^{\alpha-1}.(1-q)^{\beta-1} / (\Gamma(\alpha)\,\Gamma(\beta))$

Also Likelihood function is represented by

$L(q) = \binom{3500}{p} . q^p. (1-q)^{3500-p}$

Thus equation (i) becomes

Posterior distribution is

$$f(q|x) \propto \binom{3500}{p} . q^p. (1-q)^{3500-p} . \Gamma(\alpha+\beta)^* q^{\alpha-1}.(1-q)^{\beta-1} / (\Gamma(\alpha)\,\Gamma(\beta))$$

$$\propto \Gamma(\alpha+\beta)^* q^{\alpha+p-1}.(1-q)^{\beta+3500-p-1} / (\Gamma(\alpha)\,\Gamma(\beta))$$

This is Beta distribution with parameters (p+α, β –p +3500) **[5]**

(ii) Bayesian Estimate under quadratic loss function is the mean of the posterior distribution

Since the given posterior distribution is Beta with parameters (p+α, β –p +3500) the Bayesian Estimate is equal to

$$(p+\alpha) / (\alpha + \beta + 3500)$$

For p = 500, α = 1, β = 4 , the Bayesian Estimate under quadratic loss function is

$$= (500+1)/(1+4+3500) = 501/3505 = .1429$$ **[2]**

(iii) Let us find out if $p+\alpha / (\alpha + \beta + 3500)$ can be written in the form of a credibility estimate

Thus $(p+\alpha) / (\alpha + \beta + 3500) = p / (\alpha + \beta + 3500) + \alpha / (\alpha + \beta + 3500)$

⇨ $(p+\alpha) / (\alpha + \beta + 3500) = 3500/(\alpha + \beta + 3500)*p/(3500) + (\alpha + \beta)/ (\alpha + \beta + 3500)* \alpha/(\alpha + \beta)$

which is in the form of a credibility estimate, $Z*x + (1-Z) * \mu$

where $x = p/(3500)$ is the sample mean and $\mu = \alpha/(\alpha + \beta)$is the population mean(from the prior)

and $Z = 3500/(\alpha + \beta + 3500)$

Thus it can be represented in the form of a credibility estimate.

And Z(credibility factor) = 3500/3505 = 0.998573 for $p = 500$, $\alpha = 1$, $\beta = 4$ **[3]**

[10 Marks]

****