Institute of Actuaries of India

Subject CT6 – Statistical Methods

May 2013 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 :-

$E(X) = b+1$ & $var(X) = b$.

So, for a given x, R is a uniform discrete variable over (0, x – 1). So, R can takes values 0, 1, 2,....x-1 with equal probability.

So, $E(R/x) = (x-1) / 2$ & $var(R/x) = (x – 1) (x+1) / 12$.

So, $E(R) = E(E(R/x)) = E((x- 1) / 2) = b/2$.

$Var(R) = Var(E(R/x)) + E(var(R/x)) = var((x-1)/2) + E((x^2 – 1) / 12) = var(X) / 4 + E((X^2) – 1) / 12$.

$E(X^2) – 1 = var(X) +(E(X))^2 – 1 = 3b + b^2$.

So, $Var(R) = b/4 + (3b + b^2) / 12$. **(9 Marks)**

Solution 2 :-

a. The first step is to accumulate the claims data must form the table below:

Development Year				
Accident Year	0	1	2	3
Y1	2,000	4,400	5,000	5,250
Y2	2,500	5,000	5,810	
Y3	3,000	5,600		
Y4	3,500			

The next step is to calculate the development factors:

- Development factor for Development Year 1 (DY1) = $\frac{4400 + 5000 + 5600}{2000 + 2500 + 3000} = 2$
- Development factor for Development Year 2 (DY2) = $\frac{5000 + 5810}{4400 + 5600} = 1.15$
- Development factor for Development Year 3 (DY3) = $\frac{5250}{5000} = 1.05$

The lower half of the run-off triangle can then be completed as follows:

Development Year				
Accident Year	0	1	2	3
Y1	2,000	4,400	5,000	5,250
Y2	2,500	5,000	5,810	**5810 * 1.05 = 6,100.5**
Y3	3,000	5,600	**5600 * 1.15 = 6440**	**6440 * 1.05 = 6762.0**
Y4	3,500	**3500 * 2 = 7000**	**7000 * 1.15 = 8050**	**8050 * 1.05 = 8452.5**

The estimated claims can then be estimated as:
(6100.5 – 5810) + (6762 – 5600) + (8452.5 – 3500) = 6,405

(6)

b. Assumptions underlying the inflation adjusted Chain Ladder method are as follows:

- Payments from each origin year will develop in the same way in real terms.
- Rates of past and future claims inflation are appropriate.
- The first year is fully run-off. **(3)**

c. First step is to adjust the incremental claim data for past inflation (i.e. change the figures to present day Y4 values):

Development Year				
Accident Year	0	1	2	3
Y1	2000 * 1.02 * 1.025 * 1.0275 = **2,149**	2400 * 1.025 * 1.0275 = **2,528**	600 * 1.0275 = **617**	**250**
Y2	2500 * 1.025 * 1.0275 = **2,633**	2500 * 1.0275 = **2,569**	**810**	
Y3	3000 * 1.0275 = **3,083**	**2,600**		
Y4	**3,500**			

The next step is to accumulate the inflation adjusted figures, as shown in table below:

Development Year				
Accident Year	0	1	2	3
Y1	2,149	4,676	5,293	5,543
Y2	2,633	5,202	6,012	
Y3	3,083	5,683		
Y4	3,500			

The next step is to calculate the development factors:

- Development factor for Development Year 1 (DY1) = $\frac{4676 + 5202 + 5683}{2149 + 2633 + 3083}$ = 1.979
- Development factor for Development Year 2 (DY2) = $\frac{5293 + 6012}{4676 + 5202}$ = 1.144
- Development factor for Development Year 3 (DY3) = $\frac{5543}{5293}$ = 1.047

The lower half of the run-off triangle can then be completed as follows:

Development Year				
Accident Year	0	1	2	3
Y1	2,149	4,676	5,293	5,543
Y2	2,633	5,202	6,012	**6012 * 1.047 = 6,296**
Y3	3,083	5,683	**5683 * 1.144 = 6,503**	**6503 * 1.047 = 6,810**
Y4	3,500	**3500 * 1.979 = 6,925**	**6,925 * 1.144 = 7,926**	**7926 * 1.047 = 8300**

Finally, we need incremental data again, so we can adjust for future inflation (i.e. calculate the actual money to be paid):

Development Year				
Accident Year	0	1	2	3
Y1				
Y2				284 * 1.025 = 291
Y3			821 * 1.025 = 841	307 * 1.025^2 = 323
Y4		3,426 * 1.025 = 3511	1,000 * 1.025^2 = 1051	374 * 1.025^3 = 403

The estimated claims for year Y3 & Y4 can then be estimated as equal to 6132.

(12)

[21 Marks]

<u>Solution 3</u> :-

(a) <u>EBCT Model 1</u>

Using the data given and formulas from the tables:

N=2, n=3

Let Xij be the claims paid.

$\overline{X_A}$ = 1/3 * (3,112 + 2,124 + 1,106) = 2,114

$\overline{X_B}$ = 1/3 * (5,129 + 4,116 + 3,154) = 4,133

$\overline{X}$ = ½ * ($\overline{X_A}$ + $\overline{X_B}$) = 3,123.5

E [m (Θ)] = $\overline{X}$ = 3,123.5

E [s² (Θ)] = $\frac{1}{2}\sum\left[\frac{1}{2}\sum_{j=1}^{3}(X_{ij} - \overline{X_i})^2\right]$

= ½ * 19,81,457

= 9,90,728.5

V [m (Θ)] = $\sum_{i=1}^{2}\sum(\overline{X_i} - \overline{X})^2 - \frac{1}{6}\sum\left[\frac{1}{2}\sum_{j=1}^{3}(X_{ij} - \overline{X_i})^2\right]$

= $\sum_{i=1}^{2}\sum(\overline{X_i} - \overline{X})^2$ - 1/3 E [s² (Θ)]

= 20,38,630.5 – 3,30,242.83

= 17,08,387.67

Credibility factor, z, would be:

$$z = \frac{n}{n + \frac{E[s2(\Theta)]}{V[m(\Theta)]}} = \frac{3}{3 + \frac{9,90,728.5}{17,08,387.67}}$$

= 0.838

Thus the EBCT premium for the two states for the coming year will be:

State A: $\overline{P_A} = \overline{X_A}$ * z + (1 - z) * E [m (Θ)]

= 2,114 * 0.838 + 3,138.5 * 0.162

= 2,277.54

State B: $\overline{P_B} = \overline{X_B}$ * z + (1 - z) * E [m (Θ)]

= 4,133 * 0.838 + 3,138.5 * 0.162

= 3,969.46

(6)

(b) <u>EBCT Model 2</u>

Let Y_{ij} be the claims amounts.

Let P_{ij} be the policy volume.

Then $X_{ij} = \frac{Y_{ij}}{P_{ij}}$; N = 2, n = 3

Using the formulae from the tables:

$\overline{P_A} = \sum_j \overline{P_{Aj}}$ = 332 + 242 + 125 = 699

$\overline{P_B} = \sum_j \overline{P_{Bj}}$ = 427 + 326 + 198 = 951

$\overline{P} = \overline{P_A} + \overline{P_B}$ = 699 + 951 = 1,650

$P* = \frac{1}{5}\left[699\left(1-\frac{699}{1,650}\right)+951\left(1-\frac{951}{1,650}\right)\right]$

$= \frac{1}{5}(805.75)$ = 161.15

Table for claims per unit volume, X_{ij} :

	Yr 1	Yr 2	Yr 3
State A	9.3735	8.7769	8.848
State B	12.012	12.626	15.929

Using formulae from tables:

$\overline{X_A}$ = 9.0734

$\overline{X_B}$ = 13.0382

$\overline{X} = \frac{1}{2\left(\overline{X_A} + \overline{X_B}\right)}$ = 11.3582

E [m (Θ)] = $\overline{X}$ = 11.3582

E [s^2 (Θ)] = $\frac{1}{2}\sum\left[\frac{1}{2}\sum_{j=1}^{3} P_{ij}\,[(X]_{ij} - \overline{X_i})^2\right]$

= ½ [1/2 * 2,217.875]

= 554.4687

V[m[Θ]] =*1/P** $\left[\frac{1}{Nn-1}\sum_{j=1}^{3} P_{ij}\,[(X]_{ij} - \overline{X_i})^2 - \frac{1}{2}\sum\left[\frac{1}{2}\sum_{j=1}^{3} P_{ij}\,[(X]_{ij} - \overline{X_i})^2\right]\right]$

= *1/P** $\left[\frac{1}{Nn-1}\sum_{j=1}^{3} P_{ij}\,[(X]_{ij} - \overline{X_i})^2 - \mathrm{E}\,[s2\,(\Theta)]\right]$

=1/161.15 [1/5 * 8,551.247 – 554.469]

= 7.1721

Substituting the values to get the credibility factors;

$$Z_A = \frac{699}{699 + \frac{554.469}{7.1721}} = 0.9004$$

$$Z_B = \frac{951}{951 + \frac{554.469}{7.1721}} = 0.9248$$

Credibility premium per unit of risk volume would be:

PA = ZA * $\overline{X_A}$ + (1- ZA) * E [m (Θ)]

= 0.9004 * 9.073 + 0.0996 * 11.3582

= 9.3

PB = ZB * $\overline{X_B}$ + (1- ZB) * E [m (Θ)]

= 0.9248 * 13.038 + 0.0752 * 11.3582

= 12.9116

Thus, EBCT premium for coming year for two states will be:

State A = 9.3 * 100 = 930 lakhs

State B = 12.9116 * 200 = 2,582.3 lakhs **(8)**

[14 Marks]

Solution 4 :-

(a) Prior distribution: λ ~ Gamma (α', λ')

Likelihood: X ~ Gamma (α,λ)

Posterior distribution α Prior distribution * Likelihood

Prior: f(λ) = $\frac{\lambda'^{\alpha'}}{\Gamma\alpha'}\lambda^{\alpha'-1}\, e^{-\lambda\lambda'}$

Likelihood: Π f(xi) = $\prod_{i=1}^{n} \frac{\lambda^{\alpha}}{\Gamma\alpha}\, x_i^{\alpha-1}\, e^{-\lambda x_i}$

$= \frac{\lambda^{n\alpha}}{(\Gamma\alpha)^n}\, e^{-\lambda\Sigma x_i}\, \Pi x_i^{\alpha-1}$

$\alpha \quad \frac{1}{const} * \lambda^{n\alpha}\, e^{-\lambda\Sigma x_i} * Const$

⇨ Posterior $\alpha \quad \frac{\lambda'^{\alpha'}}{\Gamma\alpha'}\lambda^{\alpha'-1}\, e^{-\lambda\lambda'} * \lambda^{n\alpha}\, e^{-\lambda\Sigma x_i}$

$\alpha \quad \frac{1}{K} * \lambda^{(n\alpha+\alpha'-1)}\, e^{-\lambda(\Sigma x_i+\lambda')}$

» Gamma (nα + α', $\Sigma\, x_i + \lambda'$)

The above expression shows that the posterior distribution is Gamma (nα + α', $\Sigma\, x_i + \lambda'$) ; as desired.

(5)

(b)

Under quadratic loss, Bayesian estimate that minimises the expected loss is the mean of the posterior distribution.

Bayesian estimate = (ά+ nα) / (λ' + Σxi) **(2)**

[7 Marks]

<u>Solution 5</u> :-

X ~ Pareto (α, λ)

E(X) = 600

S.D (X) = 1,200

$$XR = \begin{cases} 0 & ; \quad 0 < X \leq 1{,}600 \\ X - 1{,}600 & ; \quad 1{,}600 < X \leq 2{,}800 \\ 1{,}200 & ; \quad X > 2{,}800 \end{cases}$$

Notations used: d = 1,600 ; retention

L = 1,200 ; maximum amount reinsurer would pay on any claim

$$\frac{\lambda}{\alpha - 1} = 600$$

$$\left(\frac{\lambda}{\alpha - 1}\right)^2 \frac{\alpha}{\alpha - 2} = 1{,}200^2$$

$$\Rightarrow \frac{\alpha}{\alpha - 2} = \frac{1{,}200^2}{6{,}00^2} = 4$$

$$\Rightarrow \alpha = \frac{8}{3}$$

$$\Rightarrow \lambda = 600 * \frac{5}{3} = 1{,}000$$

Now we need to calculate E[XR|X > d] = $\frac{E[X_R]}{P\,(X > d)}$

E[XR] = 0. P(X < 1600) + $\int_d^{d+L} (x-d) f(x) dx + \int_{d+L}^{\infty} L\, f(x) dx$

$$= \int_d^{d+L} x\, f(x) dx \;-\; d \int_d^{d+L} f(x) dx + \; L \int_{d+L}^{\infty} f(x) dx$$

$$= \int_d^{d+L} x\, f(x) dx - \; d\; P(d < X < d+L) + \; L\; P(X > d+L)$$

A B C

Denoting the three terms in the above equation as A, B & C for ease of showing calculations.

Evaluating A:

$$\int_d^{d+L} x\, f(x) dx = \int_d^{d+L} \frac{x\, \alpha\, \lambda^{\alpha}}{(\lambda + x)^{\alpha+1}} dx$$

Integrating by parts

$$= \alpha\lambda^{\alpha} \left\{ \left[\frac{-x}{\alpha\, (\lambda + x)^{\alpha}} \right]_d^{d+L} + \int_d^{d+L} \frac{1}{\alpha(\lambda + x)^{\alpha}} dx \right\}$$

$$= \lambda^{\alpha} \left\{ \left[\frac{-x}{(\lambda + x)^{\alpha}} \right]_d^{d+L} - \left[\frac{1}{(\alpha - 1)(\lambda + x)^{\alpha-1}} \right]_d^{d+L} \right\}$$

$$= \lambda^{\alpha} \left\{ \left[-\frac{x}{(\lambda + x)^{\alpha}} - \frac{1}{(\alpha - 1)(\lambda + x)^{\alpha-1}} \right]_d^{d+L} \right\}$$

$$= \lambda^{\alpha} \left[\frac{-x(\alpha - 1) - (\lambda + x)}{(\alpha - 1)(\lambda + x)^{\alpha}} \right]_d^{d+L}$$

= $\lambda^{\uparrow}\alpha/((\alpha - 1))$ [− +"

Substituting limits

$$= \frac{\lambda^{\alpha}}{(\alpha - 1)} \left[\frac{\alpha d + \lambda}{(\lambda + d)^{\alpha}} - \frac{\alpha(L + d) + \lambda}{(\lambda + L + d)^{\alpha}} \right]$$

Substituting the values of d, L, α & λ

A = 102.76

For B & C

F(x) = 1 - $\left(\frac{\lambda}{\lambda + x}\right)^{\alpha}$

B: d * P (d < X< d + L) = d * [F(d+L) – F(d)]

$= d\left[\left(\frac{\lambda}{\lambda + d}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + d + L}\right)^{\alpha}\right]$

Substituting values;

= 1,600 * 0.0498

= 79.68

C: L * P (X > d + L) = L * [1 – F (d + L)]

= L * $\left(\frac{\lambda}{\lambda + d + L}\right)^{\alpha}$

Substituting values;

= 1,200 * 0.0284

= 34.08

Thus E[XR] = 102.76 – 79.68 + 34.08

= 57.16

Now P(X>d) is the probability of claims hitting the reinsurer;

P(X>d) = 1 – F (d)

$= \left(\frac{\lambda}{\lambda + d}\right)^{\alpha}$

Substituting values;

= 0.0782

⇨ E [XR| X>d] = 57.16/ 0.0782 = 730.62

[9 Marks]

Solution 6 :-

a. Assumptions underlying Compound Poisson Process

The following three important assumptions are made:

- the random variables of claim amounts, $\{Xi\}_{i=1}^{\infty}$, are independent and identically distributed
- the random variables $\{Xi\}_{i=1}^{\infty}$ are independent of the number of claims, N(t), for all $t \geq 0$
- The stochastic process $\{N(t)\}_{t>0}$ is a Poisson process. **(2)**

b. Assuming reinsurer applies 50% premium loading

The net premium is given by:

C_{net} $= (1+\theta)\ E(S) - (1+\varepsilon)\ E\ (S_R)$
$= (1+\theta)\ \lambda\ E(X) - (1+\varepsilon)\ \lambda\ (1-\beta)\ E\ (Z)$
$= 200 * 1.4 * \lambda - 200 * 1.5 * \lambda\ (1-\beta)$
$= 300\lambda\beta - 20\ \lambda$

Claims (net are given by):

$E\ (S_i)$ $= \lambda\ E(Y)$
$= \lambda\beta\ E(X)$
$= 200\lambda\beta$

Hence, for net income to be greater than net claims, we must have:

$300\lambda\beta - 20\ \lambda > 200\lambda\beta$

Or $\beta > 1/5$

Assuming reinsurer applies 20% premium loading

The net premium is given by:

C_{net} $= (1+\theta)\ E(S) - (1+\varepsilon)\ E\ (S_R)$
$= (1+\theta)\ \lambda\ E(X) - (1+\varepsilon)\ \lambda\ (1-\beta)\ E\ (Z)$
$= 200 * 1.4 * \lambda - 200 * 1.2 * \lambda\ (1-\beta)$
$= 240\lambda\beta + 40\ \lambda$

Claims (net are given by) remains the same as above.

Hence, for net income to be greater than net claims, we must have:

$240\lambda\beta + 40\ \lambda > 200\lambda\beta$

Or $\beta > -1$

The value is coming negative as the reinsurers' profit loading is less than the insurers' profit loading. **(5)**

c. Direct Insurer's adjustment coefficient

The adjustment coefficient is the unique positive root, R, of the equation:

$\lambda + c_{net}\ R = \lambda\ M_Y(R)$ equation (1)

We know c_{net} from part b above.

Since the Gamma distribution has parameter $\alpha = 1$, we can say that individual claims follow exponential distribution.

Then:

$M_Y(R) = E(e^{RY}) = E(e^{R\beta X}) = M_X(\beta R) = (1 - 200\beta R)^{-1}$
Putting into equation (1), we get:
$\lambda + (300\lambda\beta - 20\lambda) R = \lambda (1 - 200\beta R)^{-1}$
or $1 + (300\beta - 20) R = (1 - 200\beta R)^{-1}$
or $1 - 200\beta R + (300\beta - 20) * R * (1 - 200\beta R) = 1$
or $- 200\beta R + 300\beta R - 60{,}000\, \beta^2 R^2 - 20R + 4000\, \beta R^2 = 0$
or $R^2 (200\beta - 3000\, \beta^2) + R (5\beta - 1) = 0$
or $R = 0$ or $(1 - 5\beta) / (200\beta - 3000\, \beta^2)$
Since R is the unique positive root, we have $R = (1 - 5\beta) / (200\beta - 3000\, \beta^2)$

(4)

d. ***Value of β***

In order to maximize R we need to differentiate the above equation with respect to β. Using the quotient rule, we get:

$$\frac{dR}{d\beta} = \frac{(200\,\beta - 3000\,\beta^2) * (-5) | (1 - 5\beta) * (200 - 6000\beta)}{(200\,\beta - 3000\,\beta^2)}$$

$$\frac{dR}{d\beta} = \frac{6000\,\beta - 15000\,\beta^2 - 200}{(200\,\beta - 3000\,\beta^2)}$$

Dividing by 200, we get:

$$\frac{dR}{d\beta} = \frac{30\,\beta - 75\,\beta^2 - 1}{(\beta - 15\,\beta^2)}$$

Setting this equal to zero, we get:

$$-30\,\beta + 75\,\beta^2 + 1 = 0$$

Therefore,

$$\beta = \frac{30 \pm \sqrt{900 - 300}}{150} = 0.3633 \text{ or } 0.0367$$

But from b) we know that $\beta > 1/5$, therefore $\beta = 0.3633$.
Substituting this value back into the equation in part c), we get $R = 0.002526$

(4)

[15 Marks]

<u>Solution 7</u> :-

a. Three important components which need to be specified in order to define a Generalized Linear Model (GLM) are as follows:

1. The first item that needs to be specified is the ***distribution of the response variable***, which should belong to exponential family of distributions.

2. ***Linear Predictor***

   In context of a GLM, the covariates (or explanatory variable) enter the model through the linear predictor. Thus, a linear predictor is a function of the covariates that allows relating them to the response variable.

3. ***Link Function***

In specifying a GLM, it is necessary to connect the mean response to the linear predictor. In general, this is done by taking some function of the mean response and this function is called the link function. **(3)**

b.

1. ***Only age is the explanatory variable***

   The linear predictor in this case can take the form $\eta = \beta_0 + \beta_1 x$, where:

   - x is the age of the policyholder; and
   - β_0 and β_1 are parameters to be estimated. **(1)**

2. ***Age and gender are the explanatory variables (but independent)***

   The linear predictor in this case can take the form $\eta = \alpha_i + \beta x$, where:

   - x is the age of the policyholder; and
   - i = 1 for a male and i = 2 for a female. Thus, there are three parameters (α_1, α_2 and β) that need to be estimated. **(1)**

3. ***Age and gender are the explanatory variables and the effect of the age is different for males and females.***

   The linear predictor in this case can take the form $\eta = \alpha_i + \beta_i x$, where:

   - x is the age of the policyholder; and
   - i = 1 for a male and i = 2 for a female. Thus, there are four parameters (α_1, α_2, β_1 and β_2) that need to be estimated. **(1)**

c. The probability distribution function of the Poisson distribution can be written as follows:

f (y) = $\dfrac{\lambda^y e^{-y}}{y!} = \exp(y \, Log\, \lambda - \lambda - Log\, y!)$

The log likelyhood function can thus be written as:

Log L = $\sum y_i \, Log\, \lambda_i - \sum \lambda_i - \sum Log\, y_i!$ ……. Eqn

The canonical link function us $\eta_i = Log\, \lambda_i = \alpha + \beta\, x_i$

Or $\lambda_i = \exp(\alpha + \beta\, x_i)$

Putting this in equation (1), we get:

Log L = $\sum y_i \, (\alpha + \beta\, \mathbf{x_i}) - \sum \exp(\boldsymbol{\alpha} + \boldsymbol{\beta}\, \mathbf{x_i}) - \sum Log\, y_i!$

Differentiating above with respect to α and β, we get:

$$\frac{\partial}{\partial \alpha} Log\, L = \sum y_i - \sum \mathbf{exp}(\alpha + \beta\, \mathbf{x_i})$$

$$\frac{\partial}{\partial\beta} Log\ L = \sum y_i x_i - \sum x_i \exp(\alpha + \beta\, \mathbf{x_i})$$

So the equations satisfied by the Maximum Likelyhood Estimators of α and β are:

$$\sum y_i - \sum \exp(\alpha + \beta\, \mathbf{x_i}) = 0$$

$$\sum y_i x_i - \sum x_i \exp(\alpha + \beta\, \mathbf{x_i}) = 0$$

(4)

d. The **scaled deviance** is defined as twice the difference between the log-likelihood of the model under consideration (known as the current model) and the saturated model.

In the context of GLM, the decision on which model to use is usually based on consideration of the deviances for a range of models. The smaller the deviance, the better the model from the point of view of model fit.

The concept of deviance is thus used to assess the adequacy of a model for describing a set of data. **(3)**

[13 Marks]

<u>Solution 8</u> :-

(a)

Using compound distribution formulae:

E[N] = E[E(N|p)]

= E[1,000 p]

= 1,000 * (0.5 *0.03 + 0.5 * 0.01)

= 20

V[N] = V[E(N|p)] + E[V(N|p)]

= V[1,000 p] + E[1,000 p (1-p)]

E[p] = 0.02

$E[p^2]$ = $0.5 * 0.03^2 + 0.5 * 0.01^2 = 0.000455$

V[p] = $E[p^2] – (E[p])^2$

= $0.000455 – 0.02^2$

$= 0.000055$

$E[p(1-p)] = 0.5 * 0.01 * 0.99 + 0.5 * 0.03 * 0.97 \quad = 0.0195$

$V[N] \quad = 1{,}000^2 \, V[p] + 1000 \, E[p(1-p)]$

$= 1{,}000^2 * 0.000055 + 1000 * 0.0195$

$=74.5$

(4)

(b) Using binomial approximation: $N \sim Bin(1{,}000, 0.02)$

$E[N] = np = 1{,}000 * 0.02 = 20$

$V[N] = np(1-p) = 1{,}000 * 0.02 * 0.98 = 19.6$ **(2)**

(c)

$S_1 = 4{,}000$; poisson rate = 1/5; n = 400

$S_2 = \begin{cases} 8{,}000, & p = 0.6 \\ 10{,}000, & p = 0.4 \end{cases}$; Poisson rate = 1/10 ; n = 400

$S = S_1 + S_2$

For compound poisson process:

$E[S] = \lambda \, E[X] \qquad V[S] = \lambda \, E[X^2] \qquad$ at t = 1

$E[S] \quad = 4{,}000 * 1/5 * 400 + 8{,}000 * 1/10 * 0.6 * 400 + 10{,}000 * .4 * 400 * .2$

$= 6{,}72{,}000$

$V[S] \quad = 4{,}000^2 * 1/5 * 400 + 8{,}000^2 * 1/10 * 0.6 * 400 + 10{,}000^2 * 1/10 * 0.4 * 400$

$= 4{,}41{,}60{,}00{,}000$

$\approx 66453 \text{ ^ } 2$ **(3)**

(d)

S ~ N(6,72,,000, $66{,}453^2$)

P(S>Y) = P[((Y – 672000) / 66453) < ((S – 672000) / 66453)] = 0.2

Solving the equation with the help of the table, it comes as Y ≈ 727926.

(3)

[12 Marks]

****