Institute of Actuaries of India

Subject CT6 – Statistical Methods

March 2018 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 examiner have given credit for any alternative approach or interpretation which they consider to be reasonable.

Solution 1:

i) The payoff matrix is given by,

		Player A					
		1	**2**	**3**	**4**	**5**	**6**
Player B	**1**	10	10	-	-	-	-
	2	10	10	10	-	-	-
	3	-	10	10	10	-	-
	4	-	-	10	10	10	-
	5	-	-	-	10	10	10
	6	-	-	-	-	10	10

[2]

ii) Here we can see that for Player A, number 1 and 6 are dominated by number 2 and 5 respectively. Hence the revised payoff matrix can be written as:

		Player A			
		2	**3**	**4**	**5**
Player B	**1**	10	-	-	-
	2	10	10	-	-
	3	10	10	10	-
	4	-	10	10	10
	5	-	-	10	10
	6	-	-	-	10

Now for Player B, number 1 dominates number 2 and 3 and similarly number 6 dominates number 4 and 5. Hence revised payoff matrix can be written as

		Player A			
		2	**3**	**4**	**5**
Player B	**1**	10	-	-	-
	6	-	-	-	10

Now number 3 and 4 are dominated for Player A, hence the revised payoff matrix is

		Player A	
		2	**5**
Player B	**1**	10	-
	6	-	10

[4]

iii) Player B should select either number 1 or 6 with probability 0.5.
Hence value for Player B = -10*0.5 = -5. [1]

[7 Marks]

Solution 2:

i) Under a proportional reinsurance arrangement, the direct writer (ie the original insurance company) and the reinsurer share the cost of all claims for each risk.

Under a non-proportional reinsurance arrangement, the direct writer pays a fixed premium to the reinsurer. The reinsurer will only be required to make payments where part of the claim amount falls in a particular reinsurance layer.

[2]

ii) Proportional reinsurance operates in two forms:
With quota share reinsurance, the proportions are the same for all risks
With surplus reinsurance, the proportions can vary from one risk to the next
Two forms of non-proportional reinsurance are:-
Individual excess of loss (XOL) reinsurance, the reinsurer will be required to make a payment when the claim amount for any individual claim exceeds a specified excess point or retention.

Stop loss reinsurance, the reinsurer will be required to make payments if the total claim amount for a specified group of policies exceeds a specified amount (usually expressed as a percentage of the gross premium).

[2]

iii)

a) If X represents the size of an individual claim, then we have:

$$\mathrm{E(X)} = \frac{\alpha}{\lambda}$$

and, $\mathrm{Variance(X)} = \frac{\alpha}{\lambda^2}$

Hence, E(X) = 3.33

Variance (X) = 1.11

Amount paid by Insurer is Y = 0.2X

E(Y) = 0.2 X 3.33 = 0.67

Variance(Y) = 0.2^2X1.11 = 0.04 [3]

b) Amount paid by Reinsurer Z =0.8X

E(Z) = 0.8 X 3.33 = 2.67
Variance (Z) = 0.8^2 X 1.11 = 0.71 [3]

iv) $\int_{2500}^{M}(X-2500)\lambda e^{-\lambda x}dx + \int_{M}^{\infty}(M-2500)\lambda e^{-\lambda x}dx = 5000$

Where, λ = 1/10,000

Solving the above equation using integration by parts, we get:-
$10000e^{-2500\lambda} - 10000e^{-M\lambda} = 5000$

Solving for M = 12773

[7]

[17 Marks]

Solution 3

i) The Bornhuetter-Ferguson method improves on the crude use of a loss ratio by taking account of the information provided by the latest development pattern of the claims, whilst the addition of the loss ratio to a projection method serves to add some stability against distortions in the development pattern.

The concepts behind the method are:

- That whatever claims have already developed in relation to a given origin year, the future development pattern will follow that experienced for other origin years.
- The past development for a given origin year does not necessarily provide a better clue to future claims than the more general loss ratio.

[2]

ii) The development of claims for various accident year and the expected ultimate claim amount is summarized in the table below:

Accident Year	**Development Year**			
	0	**1**	**2**	**3**
2014	40%	55%	70%	75%
2015	35%	50%	75%	
2016	60%	70%		
2017	55%			

Earned Premium	**Expected loss ratio**	**Expected loss amount**
120	80%	96
150	90%	135
140	105%	147
180	95%	171

The cumulative claim amount based on the earned premium and the development factors are given by:

Accident Year	**Development Year**			
	0	**1**	**2**	**3**
2014	48	66	84	90
2015	52.5	75	112.5	

2016	84	98		
2017	99			

Total	283.5	239	196.5	90
Total - last no	184.5	141	84	
Ratio (r)	1.295393	1.393617	1.071429	1
DF (f)	1.93423	1.493161	1.071429	1

Revised estimate of total ultimate losses, by accident year is given by

Accident Year	**2017**	**2016**	**2015**	**2014**
f	1.93423	1.493161	1.071429	1
1-1/f	0.482998	0.33028	0.066667	0
Initial UL	171	147	135	96
Emerging Liability	82.59	48.55	9.00	-
Reported Liability	99.00	98.00	112.50	90.00
Ultimate Liability	181.59	146.55	121.50	90.00

Now, there is 1 year waiting period to make payment from date of intimation of claim.

Accident Year	**2017**	**2016**	**2015**	**2014**
Claim paid	0	84	75	84
Outstanding claim reserve	181.59	62.55	46.50	6.00

Hence the outstanding total claim reserve = 296.64

[9]

[11 Marks]

Solution 4:

i) For region Metro, the log-likelihood function using poisson parameter λ1 is given by,

$LogL_1 = \log\lambda_1 \sum_{i=1}^{9} x_{1i} - 9\lambda_1 - \sum_{i=1}^{9} \log x_{1i}\,!$ ..(0.5)

Or, $LogL_1 = 37\log\lambda_1 - 9\lambda_1 - \sum_{i=1}^{9} \log x_{1i}\,!$ ..(0.5)

Taking partial derivative we get,

$\frac{\partial}{\partial\lambda_1} LogL_1 = \frac{37}{\lambda_1} - 9 = 0, \lambda_1 = 4.111$ ..(0.5)

Similarly, $LogL_2 = \log\lambda_2 \sum_{i=1}^{8} x_{2i} - 8\lambda_2 - \sum_{i=1}^{8} \log x_{2i}\,!$ ..(0.5)

Hence, $\lambda_2 = 3.625, \lambda_3 = 5.857$ [4]

ii) We need to test, $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$, where λ = 107/24 = 4.458 ..(0.5)

We can use scaled deviance to compare these models:

The difference in the scaled deviance follows chi-square with (3-1 = 2) degrees of freedom.

$$2(LogL_1 + LogL_2 + LogL_3 - LogL) = 2(37\log\lambda_1 - 9\lambda_1 + 29\log\lambda_2 - 8\lambda_2 + 41\log\lambda_3 - 7\lambda_3 - 107\log\lambda + 24\lambda)$$

(factorial terms cancel each other)
= 4.3753

The critical value at 5% level = 5.991

The calculated value is less than the critical level and hence the mean claim rate are not different. [4]

[8 Marks]

Solution 5:

Here, P(X>200) = 1/250 = 0.004

Now, under exponential distribution we get

Exp(-λ*200) = 0.004
Hence, λ = 0.027607

Now, Exp(-λ*Z)=0.0001
Hence Z = 333.62

Now, under Weibull distribution we get

Exp(-λ*200^2) = 0.004
Hence, λ = 0.000138

Now, Exp(-λ*Z^2)=0.0001
Hence Z = 258.31

Hence Z is higher under Exponential

Exponential has longer tail than Weibull and hence Z will be higher under Exponential distribution.

[6 Marks]

Solution 6:

i) The probability density function of h(x) is given by,

h(x) = exp(-x/λ)/λ, x>0

Now, $F_h(x)$ = 1 – exp(-x/λ) = U,

We get, X = - λ ln(1-U) = - λ ln U',

Step 1: Simulate a u from U(0,1)
Step 2: Return X = - λ ln(u) [1]

ii) Now f(x) follows chi-square with α degrees of freedom ie it follows Gamma(α/2,0.5)

So, f(x) = $\dfrac{0.5^{\frac{\alpha}{2}}}{\Gamma(0.5)}x^{\frac{\alpha}{2}-1}e^{-0.5x}$

Now, f(x)/ h(x)

= $\dfrac{\lambda 0.5^{\frac{\alpha}{2}}}{\Gamma(0.5)}x^{\frac{\alpha}{2}-1}e^{-0.5x+\frac{x}{\lambda}}$, x>0

Taking log both side we get,

Ln (f(x)/ h(x)) = $\ln(\dfrac{\lambda 0.5^{\frac{\alpha}{2}}}{\Gamma(0.5)})+(\dfrac{\alpha}{2}-1)\ln x-0.5x+\dfrac{x}{\lambda}$

Now taking derivative and equating to zero we get,

$x=(\dfrac{\alpha}{2}-1)/(0.5-\dfrac{1}{\lambda})$ ..(1.0)

Now the second derivate is $-(\dfrac{\alpha}{2}-1)\dfrac{1}{x^2}<0$ ..(0.5)

Hence, f/ h attains maximum value at $x_0=(\dfrac{\alpha}{2}-1)/(0.5-\dfrac{1}{\lambda})$

Hence C = max (f/h) = $\dfrac{\lambda 0.5^{\frac{\alpha}{2}}}{\Gamma(0.5)}x_0^{\frac{\alpha}{2}-1}e^{-(0.5-\frac{1}{\lambda})x_0}$

Now, the function g(x) = f(x) / [C* h(x)] = $\dfrac{x^{\frac{\alpha}{2}-1} e^{-(0.5-\frac{1}{\lambda})x}}{x_0^{\frac{\alpha}{2}-1} e^{-(0.5-\frac{1}{\lambda})x_0}}$

Step 1: Simulte U1 from U(0,1) and set Y = -ln(u1)/λ
Step 2: Simulate U2 from U(0,1) and if U2 < g(Y) set accept the value by setting X = Y otherwise move to Step 1

[6]

iii) The algorithm will be efficient by choosing value of λ such that g(x) is maximum or C is minimum.

Now C = $\dfrac{\lambda 0.5^{\frac{\alpha}{2}}}{\Gamma(0.5)} x_0^{\frac{\alpha}{2}-1} e^{-(0.5-\frac{1}{\lambda})x_0}$ = $\dfrac{0.5^{\frac{\alpha}{2}}}{\Gamma(0.5)} \lambda[(\frac{\alpha}{2}-1)/(0.5-\frac{1}{\lambda})]^{\frac{\alpha}{2}-1} e^{1-\frac{\alpha}{2}}$

Which is minimized if, $\ln[\lambda[(\frac{\alpha}{2}-1)/(0.5-\frac{1}{\lambda})]^{\frac{\alpha}{2}-1}]$ is minimized

i.e,

$\ln\lambda-(\frac{\alpha}{2}-1)\ln(0.5-\frac{1}{\lambda})$

Taking derivative and equating to zero we get,

$\frac{1}{\lambda}-(\frac{\alpha}{2}-1)\frac{1}{(0.5-\frac{1}{\lambda})}\frac{1}{\lambda^2}=0$

Or, $(0.5-\frac{1}{\lambda})\lambda=(\frac{\alpha}{2}-1)$

Or, $\lambda=\alpha$

[5]

[12 Marks]

Solution 7

i) Lundberg's inequality states that: $\psi(U) \le \exp(-RU)$

where U is the insurer's initial surplus and $\psi(U)$ is the probability of ultimate ruin. R is a parameter associated with a surplus process known as the adjustment coefficient. Its value depends upon the distribution of aggregate claims and on the rate of premium income. [2]

ii) $\exp(Rx) \le \frac{x}{M}\exp(RM) + 1 - \frac{x}{M}$ for $0 \le x \le M$

Now,

$\lambda M_x(R) = \lambda + cR$

i.e.

$\lambda + cR = \lambda\int_0^M \exp(Rx)f(x)dx$

$$\leq \lambda \int_0^M \{\frac{x}{M}\exp(RM) + 1 - \frac{x}{M}\}\ f(x)dx$$

This gives, $R > \frac{1}{M}\log\left(\frac{c}{\lambda m_1}\right)$ [4]

iii) Ruin will occur if the time of the first claim t is such that

$$U + 1.5\lambda dt < d$$

$$t < \frac{d - U}{1.5\lambda d}$$

The time until the first claim follows an exponential distribution with parameter λ, so the probability of ruin in given by:

$$\int_0^{\frac{d-U}{1.5\lambda d}} \lambda e^{-\lambda x}\, dx$$

$$= 1 - e^{\frac{-1\left(1-\frac{U}{d}\right)}{1.5}}$$ [5]

iv) $M_x(t) = E(e^{tx)} \quad = \frac{0.05^2}{(t-0.05)^2}$

The adjustment coefficient is the unique positive solution of

$M_x(R) = 1 + 1.3E(X)R$

$E(X) = M_x'(0)$

=40

Thus we need to solve

$$\frac{0.05^2}{(R - 0.05)^2} = 1 + 1.3 \times 40R$$

$52R^2 - 4.2R + 0.03 = 0$

R=0.07284 OR 0. 00791

So taking the smaller root we have R=0.00791 since that is less than 0.01 [6]

[17 Marks]

Solution 8

i)

Strategy	Minimum Profit
Freekart.com	14
Appdeal.com	9
E-mazon.com	5

Hence, the minimax strategy is to sell on Freekart.com [2]

ii)

Strategy	Maximum Profit
Freekart.com	28
Appdeal.com	30
E-mazon.com	29

Hence the maximmax strategy is to sell on Appedeal.com [1]

iii)

Strategy	Expected profit
Freekart.com	$\frac{1}{3}(28+19+14) = 20.33$
Appdeal.com	$\frac{1}{3}(9+30+15) = 18$
E-mazon.com	$\frac{1}{3}(5+16+29) = 16.67$

Hence the strategy selected by the Bayes criterion is Freekart.com [2]

iv) If probability of medium demand is p then the probabilities of High and Low demands are 0.5-0.5P each.

Hence revised expected profits from each strategy are:-

Strategy	Expected profit
Freekart.com	$28(0.5-0.5\text{P}) + 19\text{P} + 14(0.5-0.5\text{P}) = 21-2\text{P}$
Appdeal.com	$9(0.5-0.5\text{P}) + 30\text{P} + 15(0.5-0.5\text{P}) = 12+18\text{P}$
E-mazon.com	$5(0.5-0.5\text{P}) + 16\text{P} + 29(0.5-0.5\text{P}) = 17-1\text{P}$

As there are more than even chances of it being medium, so p is expected to be in between 0.5 t0 1

Hence Appdeal.com is the optimal decision. [3]

[8 Marks]

Solution 9

i) GLMs are widely used both in general and life insurance. They are used to:

- determine which rating factors to use (rating factors are measurable or categorical factors that are used as proxies for risk in setting premiums)
- Estimate an appropriate premium to charge for a particular policy given the level of risk present.

For example, in motor insurance, there are a large number of factors that may be used as proxies for the level of risk (type of car driven, age of driver, number of years past driving experience, etc). We can use a GLM both to determine which of these factors are significant to the assessment of risk (and hence which should be included) and to suggest an appropriate premium to charge for a risk that represents a particular combination of these factors. [2]

ii) The three components of a GLM are:

- A distribution for the data (Poisson, exponential, gamma, normal or binomial)
- A linear predictor (a function of the covariates that is linear in the parameters)
- A link function (that links the mean of the response variable to the linear predictor). [2]

iii)

a) In parameterized form, the linear predictors are (with i, j and k corresponding to the levels of A, B, C respectively

$\text{Model 1} = \alpha_i + \beta_j + \gamma_k$; total 4 parameters
There is one parameter to set the base level for the combination A, B and C and one additional parameter for each of the higher levels of the three factors

$\text{Model 2} = \alpha_{ij} + \gamma_k$; total 5 parameters
There are four parameters for the 2 X2 combinations of A and B and one additional parameter for the higher level of TC.

$\text{Model 3} = \alpha_{ijk}$; total 8 parameters
There are eight parameters for the 2 X 2X 2 combinations of A, B and C [5]

b)

Model	Scaled Deviance	Degrees of Freedom	Change in Scaled Deviance	Chance in Degrees of Freedom
1	40	n* (say)		
A + B +C	11	n-4	29	3
A + B + C + A.B	7	n-3	4	1
A*B*C	0	n-7	7	3

*Candidates will also get full marks if they estimate n as (2*2*2 -1 = 7)

	If Change in Deviance > 2 X change in df	Conclusion
Constant Model vs Model 1	29 > 2 X 3	Model 1 is better than constant model
Model 1 vs Model 2	4 > 2 X 1	Model 2 is better than model 1
Model 2 vs Model 3	7 > 2 X 3	Model 3 is better than Model 2

Hence Model 3 is the most appropriate model [5]

[14 Marks]

****