Institute of Actuaries of India

Subject CT6 – Statistical Methods

November 2012 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.

Question 1

The cumulative numbers of claims reported in the different years are as follows.

Accident Year	Development Year		
	0	1	2
2009	105	195	240
2010	152	280	
2011	285		

The corresponding average costs (in thousands of rupees) are as follows.

Accident Year	Development Year		
	0	1	2
2009	3.047619	3.220513	3.729169
2010	3.223684	3.500000	
2011	2.631579		

The accident year 2009 is fully run off. The cumulative percentage of total claim payments for the accident year are respectively:

81.72%, 86.36%, 100.00%.

So, the ultimate average claim cost for accident year 2010 is (3.5/0.8636) = 4.052797.

With this, we can calculate the cumulative percentage figures for accident year 2010 as

79.54% and 86.36%.

For accident year 2011, we take the average of two previous year's figures for development year 0 as:

(81.72% + 79.54%) / 2 = 80.63%.

Thus, the ultimate figure for Accident year 2011 is : 2.631579 / 0.8063 = 3.263648.

In summary, we have the projected ultimate costs per claim as under.

Accident Year	Development Year			
	0	1	2	*Ult*
2009	3.04762	3.22051	3.72917	3.72917
2010	3.22368	3.5		4.05280
2011	2.63158			3.26365

Similarly, we calculate the claim number cumulative percentages and projected ultimate values as follows.

Accident Year	Development Year			
	0	1	2	*Ult*
2009	105 43.75%	195 81.25%	240 100.00%	240.0000
2010	152 44.11%	280 81.25%		344.6154
2011	285 43.93%			648.7805

Thus, the total expected ultimate claim outgo
= Rs. (240 *3.729167 + 344.6154 * 4.052797 + 648.7805 * 3.263648) * 1000
= Rs. 4,409,047.

Since the claims paid till date is Rs 2,500,000, the outstanding claim reserve is Rs 1,909,047.

[9]

Question 2

(i) The claim size can be written as $Z = 200X + 500Y$, where X is as described in the question, Y is a binary random variable assuming values 1 and 0 with probabilities 0.25 and 0.75, respectively, and X and Y are independent.

$$E(X) = \frac{2}{25}\int_0^5 (5x - x^2)dx = \left(\frac{1}{5}x^2 - \frac{2}{75}x^3\right)\Big|_0^5 = \frac{1}{5}\times 25 - \frac{2}{75}\times 125 = \frac{5}{3}.$$

$$E(X^2) = \frac{2}{25}\int_0^5 (5x^2 - x^3)dx = \left(\frac{2}{15}x^3 - \frac{1}{50}x^4\right)\Big|_0^5 = \frac{2}{15}\times 125 - \frac{1}{50}\times 625 = \frac{25}{6}.$$

Therefore,

$$V(X) = E(X^2) - [E(X)]^2 = \frac{25}{6} - \frac{25}{9} = \frac{25}{18}.$$

On the other hand,

$$E(Y) = \frac{1}{4}; \quad V(Y) = \frac{1}{4}\times\frac{3}{4} = \frac{3}{16}.$$

It follows that

$$E(Z) = 200E(X) + 500E(Y) = 200 \times \frac{5}{3} + 500 \times \frac{1}{4} = 458.33;$$

and

$$V(Z) = 200 * 200 * V(X) + 500 * 500 * V(Y) = 200 * 200 \times \frac{25}{18} + 500 * 500 \times \frac{3}{16} = 102430.66$$

(ii) The aggregate of claims S follows a compound Poisson distribution with Poisson parameter 25. Hence,

$$E(S) = 25E(Z) = 25 \times 458.33 = 11{,}458.33;$$

and

$$V(S) = 25E(Z^2) = 25 \times (102430.66 + 458.33^2) = 7812423.61.$$

[9]

Question 3

The end of the first month is the first occasion of claim payment, and hence, the first occasion of possible ruin. Let $U(t)$ be the surplus at time t and $S(t)$ denote the aggregate claims paid till time t. Then

$$U(1) = 90{,}000 + 200 \times 1{,}000 - S(1) = 290{,}000 - S(1).$$

The event of ruin corresponds to $U(1) < 0$, i.e., $S(1) > 290{,}000$.

Let $X_1, X_2, \ldots$ be the successive claims and $N(t)$ be the number of claims arising till time t. Clearly

$$S(t) = \sum_{i=1}^{N(t)} X_i.$$

Also X_i are iid, assuming values 100,000, 250,000 and 500,000 with probabilities 0.7, 0.25 and 0.05, respectively.
The following is an exhaustive list of exclusive events that permit non-ruin, i.e., $S(1) \leq 290{,}000$.
Case I: $N(1) = 0.$
Case II: $N(1) = 1;\ X_1 \neq 500{,}000.$
Case III: $N(1) = 2;\ X_1 = X_2 = 100{,}000.$

Therefore,

$$\begin{aligned} P[S(1) \leq 290{,}000] &= P[N(1) = 0] + P[N(1) = 1;\ X_1 \neq 500{,}000] \\ &\quad + P[N(1) = 2;\ X_1 = X_2 = 100{,}000] \\ &= e^{-1}\left[1 + 1 \times 0.95 + \frac{1}{2} \times 0.7 \times 0.7\right] = 2.195e^{-1} = 0.8075. \end{aligned}$$

It follows that the probability of ruin is

$$P[S(1) > 290{,}000] = 1 - 0.8075 = 0.1925.$$

[8]

Question 4

(i) The annual profit is

Number of policies sold × revenue per policy – cost – overhead.

The first term (number of policies sold × revenue per policy) for the different combinations are as follows.

Revenue Number of policies

Product	per policy	Low 1,680	Medium 2,100	High 2,520
Basic	1,500	**2,520,000**	**3,150,000**	**3,780,000**
Lean	1,000	**1,680,000**	**2,100,000**	**2,520,000**
Rich	2,000	**3,360,000**	**4,200,000**	**5,040,000**

Costs plus overhead for the different combinations are as follows.

				Number of policies		
				Low	Medium	High
Product	OH	Cost	Cost + OH	1,680	2,100	2,520
Basic	1,500,000	500,000	2,000,000	**2,000,000**	**2,000,000**	**2,000,000**
Lean	1,500,000	300,000	1,800,000	**1,800,000**	**1,800,000**	**1,800,000**
Rich	1,500,000	1,000,000	2,500,000	**2,500,000**	**2,500,000**	**2,500,000**

Finally, the annual profit for the different combinations are as follows.

	Number of policies		
	Low	Medium	High
Product	1,680	2,100	2,520
Basic	**520,000**	**1,150,000**	**1,780,000**
Lean	**–120,000**	**300,000**	**720,000**
Rich	**860,000**	**1,700,000**	**2,540,000**

(ii) The minimum profit (negative of maximum loss) for the Basic, Lean and Rich products are Rs. 520,000, Rs. –120,000 and Rs. 860,000.

This is maximized (i.e., maximum loss is minimized) when the Rich product is chosen.

(iii) The average profit for the Basic, Lean and Rich products are shown below.

	Number of policies			
	Low	Medium	High	
	1,680	2,100	2,520	
(probability)	(0.2)	(0.6)	(0.2)	Average profit
Product				
Basic	520,000	1,150,000	1,780,000	**1,150,000**
Lean	–120,000	300,000	720,000	**300,000**
Rich	860,000	1,700,000	2,540,000	**1,700,000**

The Bayes solution also turns out to be the Rich product.

[7]

Question 5

(i) The likelihood function is

$$f(X|\theta) = \frac{1}{\theta} I(X < \theta),$$

and the prior density of θ is

$$f(\theta) = \theta e^{-\theta} I(\theta > 0).$$

Therefore, the posterior density of θ is

$$f(\theta|X) = c.f(X|\theta)f(\theta) = ce^{-\theta} I(\theta > X),$$

where c is a constant. It follows that the cumulative distribution function is

$$F(\theta|X) = c\int_{-\infty}^{\theta} e^{-u} I(u > X)du = c\int_{X}^{\theta} e^{-u}du = c\left(e^{-X} - e^{-\theta}\right) I(\theta > X).$$

Since $F(\infty|X) = 1$, we have $c = e^{X}$, i.e., $F(\theta|X) = \left(1 - e^{-(\theta - X)}\right) I(\theta > X)$, and $f(\theta|X) = e^{-(\theta - X)} I(\theta > X)$.

Absolute error loss function is used. Hence, the Bayes estimator is the median of the posterior distribution. This is obtained by solving

$$\frac{1}{2} = F(\theta|X) = 1 - e^{-(\theta - X)}.$$

The solution is $\theta = X + \ln 2$, i.e., the appropriate estimator is $X + \ln 2$.

(ii) The mean absolute error is $E[|kX - \theta|]$, which simplifies to

$$\begin{aligned} E[|kX - \theta|] &= \int_0^{\theta} |kx - \theta| \frac{1}{\theta} dx = \int_0^{\theta/k} (\theta - kx)\frac{1}{\theta} dx + \int_{\theta/k}^{\theta} (kx - \theta)\frac{1}{\theta} dx \\ &= \left(x - \frac{kx^2}{2\theta}\right)\Big|_0^{\theta/k} + \left(\frac{kx^2}{2\theta} - x\right)\Big|_{\theta/k}^{\theta} = \frac{\theta}{k} - \frac{\theta}{2k} + \frac{k\theta}{2} - \frac{\theta}{2k} - \theta + \frac{\theta}{k} \\ &= \theta\left(\frac{k}{2} - 1 + \frac{1}{k}\right). \end{aligned}$$

This quantity has first derivative equal to zero when $k = \sqrt{2}$,

and the second derivative is always positive.

Evidently, $k = \sqrt{2}$ is the unique minimum.

Therefore, the requisite estimator is $\sqrt{2}X$.

(iii) The Bayes estimator is obtained by minimizing $E_{\theta}[|g(X) - \theta|]$ with respect to the function g for fixed X.

The estimator of part (ii) is obtained by minimizing $E_X[|g(X) - \theta|]$ with respect to the function g for fixed θ, subject to the constraint $g(X) = kX$.

[10]

Question 6

(i) The presumed model can be written as $\ln(\mu_i) = a + bX_i$, where the linear predictor is $a + bX_i$ and the link function is the natural log function.

The canonical link function for the exponential distribution (a special case of the gamma distribution) is the inverse function, which is different from the link function implied by the presumed model.

(ii) The likelihood is

$$\prod_{i=1}^{n} \frac{1}{\mu_i} e^{-Y_i/\mu_i} = \prod_{i=1}^{n} e^{-(a+bX_i)} e^{-Y_i e^{-(a+bX_i)}}.$$

Therefore, the log-likelihood is

$$l(a,b) = -\sum_{i=1}^{n}(a + bX_i) - \sum_{i=1}^{n} Y_i e^{-(a+bX_i)} = -na - b\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i e^{-(a+bX_i)}.$$

.

Therefore,

$$\frac{\partial l}{\partial a} = -n - \sum_{i=1}^{n} Y_i e^{-(a+bX_i)}$$

and

$$\frac{\partial l}{\partial b} = -\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} X_i Y_i e^{-(a+bX_i)}.$$

The maximum likelihood estimates of a and b are the solutions to the equations

$$n + \sum_{i=1}^{n} Y_i e^{-(a+bX_i)} = 0, \quad \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} X_i Y_i e^{-(a+bX_i)} = 0.$$

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

For the model under consideration, twice log-likelihood is

$$2l(\hat{a},\hat{b}) = -2n\hat{a} - 2\hat{b}\sum_{i=1}^{n} X_i - 2\sum_{i=1}^{n} Y_i e^{-(\hat{a}+\hat{b}X_i)},$$

where $\hat{a}$ and $\hat{b}$ are solutions to the equations obtained in part (ii).

For the saturated model, twice log-likelihood is

$$2l = -2\sum_{i=1}^{n} \ln \mu_i - 2\sum_{i=1}^{n} Y_i/\mu_i \bigg|_{\mu_i = Y_i} = -2\sum_{i=1}^{n} \ln Y_i - 2n.$$

It follows that the scaled deviance is

$$2l(\hat{a},\hat{b}) - 2l = 2\sum_{i=1}^{n} \ln Y_i + 2n - 2n\hat{a} - 2\hat{b}\sum_{i=1}^{n} X_i - 2\sum_{i=1}^{n} Y_i e^{-(\hat{a}+\hat{b}X_i)},$$

$\hat{a}$ and $\hat{b}$ being solutions to the equations obtained in part (ii).

[*Alternative derivation of expression:*

When the data have the exponential distribution, the scaled deviance is

$$2\left(-\sum_{i=1}^{n} \ln \mu_i - \sum_{i=1}^{n} Y_i/\mu_i\right)\Bigg|_{\mu_i=\hat{\mu}_i} - 2\left(-\sum_{i=1}^{n} \ln \mu_i - \sum_{i=1}^{n} Y_i/\mu_i\right)\Bigg|_{\mu_i=Y_i}$$
$$= 2\sum_{i=1}^{n} \ln Y_i + 2n - 2\sum_{i=1}^{n} \ln \hat{\mu}_i - 2\sum_{i=1}^{n} Y_i/\hat{\mu}_i,$$

where $\hat{\mu}_i$ is the estimate obtained from the fitted model.

For the model under consideration, we have $\hat{\mu}_i = e^{(\hat{a}+\hat{b}X_i)}$. Thus, the scaled deviance is

$$2\sum_{i=1}^{n} \ln Y_i + 2n - 2n\hat{a} - 2\hat{b}\sum_{i=1}^{n} X_i - 2\sum_{i=1}^{n} Y_i e^{-(\hat{a}+\hat{b}X_i)},$$

$\hat{a}$ and $\hat{b}$ being solutions to the equations obtained in part (ii). **[10]**

Question 7

(i) Two time series X and Y are cointegrated if they are $I(1)$ random processes, and there exists a non-zero vector (α,β) such that $\alpha X + \beta Y$ is stationary.

The requirement of X and Y being $I(1)$ processes means that their first order differences should be stationary.

(ii) Example 1: One process drives the other.

Example 2: Both processes are driven by a common underlying process.

(iii) A time series $X_1, X_2, X_3, \ldots$ is defined to be weakly stationary if its mean $\mu(n) = E[X_n]$ is constant (i.e., it does not depend on the time parameter n) and the covariance between two time samples of the process, $C(m,n) = E[\{X_m - \mu(m)\}\{X_n - \mu(n)\}]$ depends only on the time difference $n - m$.

(iv) It follows from the given difference equation that

$$Cov(Y_n, Z_n) = \sigma^2,$$

$$Cov(Y_n, Z_{n-1}) = 0.65Cov(Y_{n-1}, Z_{n-1}) + 0.35\sigma^2 = 0.65\sigma^2 + 0.35\sigma^2 = \sigma^2.$$

Therefore,

$$\gamma_0 = Cov(Y_n, Y_n) = Cov(Y_n, 0.65Y_{n-1} + Z_n + 0.35Z_{n-1}) = 0.65\gamma_1 + \sigma^2 + 0.35\sigma^2$$
$$= 0.65\gamma_1 + 1.35\sigma^2;$$

$$\gamma_1 = Cov(Y_n, Y_{n-1}) = Cov(0.65Y_{n-1} + Z_{n-1} + 0.35Z_{n-1}, Y_{n-1}) = 0.65\gamma_0 + 0.35\sigma^2.$$

By putting the last two equations together, we have

$$\gamma_0 = 0.65\gamma_1 + 1.35\sigma^2 = 0.65[0.65\gamma_0 + 0.35\sigma^2] + 1.35\sigma^2 = 0.4225\gamma_0 + 1.5775\sigma^2,$$
i.e., $0.5775\gamma_0 = 1.5775\sigma^2$, so that $\gamma_0 = 2.7316\sigma^2$.

Likewise,

$$\gamma_1 = 0.65\gamma_0 + 0.35\sigma^2 = 0.65 \times 2.7316\sigma^2 + 0.35\sigma^2 = 2.1255\sigma^2.$$

For $k \geq 2$,

$$\gamma_k = Cov(Y_n, Y_{n-k}) = Cov(0.65Y_{n-1} + Z_n + 0.35Z_{n-1}, Y_{n-k}) = 0.65\gamma_{k-1}.$$

It follows that $\gamma_k = 0.65^{k-1}\gamma_1 = 2.1255 \times 0.65^{k-1}\sigma^2$. Therefore,

$\rho_0 = 1$,

$\rho_1 = \dfrac{\gamma_1}{\gamma_0} = \dfrac{2.1255}{2.7316} = 0.7781$,

$\rho_k = \dfrac{\gamma_k}{\gamma_0} = 0.65^{k-1}\rho_1 = 0.7781(0.65)^{k-1}$ for $k \geq 2$.

(v) a. It follows from the given difference equation that

$$E[X_{n+1}|(X_n, X_{n-1}, X_{n-2}, X_{n-3}, \ldots)] = 1.2X_n + 0.7X_{n-1} - 0.1X_{n-2},$$

i.e., the conditional expectation of X_{n+1} depends not only on X_n, but also on X_{n-1} and X_{n-2}. Therefore, X is not a Markov process.

b. In order to construct a multivariate Markov process, define the process $\boldsymbol{Y}$ by the equation

$$\boldsymbol{Y}_n = \begin{pmatrix} X_n \\ X_{n-1} \\ X_{n-2} \end{pmatrix}.$$

c. It follows from the construction that one can write $\boldsymbol{Y}_{n+1}$ as

$$\begin{pmatrix} 1.2X_n + 0.7X_{n-1} - 0.1X_{n-2} + Z_{n+1} \\ X_n \\ X_{n-1} \end{pmatrix} = \begin{pmatrix} 1.2 & 0.7 & -0.1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \boldsymbol{Y}_n + \begin{pmatrix} Z_{n+1} \\ 0 \\ 0 \end{pmatrix}.$$

The distribution of this quantity, given $(\boldsymbol{Y}_n, \boldsymbol{Y}_{n-1}, \boldsymbol{Y}_{n-2}, \ldots)$, depends on $\boldsymbol{Y}_n$ but not on $(\boldsymbol{Y}_{n-1}, \boldsymbol{Y}_{n-2}, \ldots)$. Hence the process $\boldsymbol{Y}$ is Markov.

[15]

Question 8

(i) The overall mean, $E[m(\theta)]$, is estimated by $\bar{X} = \frac{125+85+140+175}{4} = 131.25$.

Average variance of claims within products, $E[S^2(\theta)]$, is estimated by

$$\frac{1}{4}\sum_{i=1}^{4}\left[\frac{1}{4}\sum_{j=1}^{5}\left(X_{ij} - \bar{X}_i\right)^2\right] = \frac{300+60+35+100}{4} = 123.75.$$

Variance of the mean claims across products, $V[m(\theta)]$, is estimated by

$$\frac{1}{3}\sum_{i=1}^{4}\left(\bar{X}_i - \bar{X}\right)^2 - \frac{1}{5}E\left[S^2(\theta)\right]$$
$$= \frac{(125-131.25)^2 + (85-131.25)^2 + (140-131.25)^2 + (175-131.25)^2}{3}$$
$$- \frac{123.75}{5}$$
$$= 1389.583 - 24.75 = 1364.83.$$

Therefore, the credibility factor, $\frac{5}{5+\frac{E[S^2(\theta)]}{V[m(\theta)]}}$, is estimated by

$$\frac{5}{5+\frac{123.75}{1364.83}} = 0.982189.$$

The credibility premiums for the four products are as follows.

Product 1: $0.982189 \times 125 + (1 - 0.982189) \times 131.25 = 125.11$

Product 2: $0.982189 \times \ \ 85 + (1 - 0.982189) \times 131.25 = 85.82$

Product 3: $0.982189 \times 140 + (1 - 0.982189) \times 131.25 = 139.84$

Product 4: $0.982189 \times 175 + (1 - 0.982189) \times 131.25 = 174.22$

(ii) The data shows that the variation within the products is relatively low, but there is a high variation between the average claims on products.

This means that we can put relatively little weight on the information provided by the data set as a whole, and must put more weight on the data from the individual products, leading to a relatively high credibility factor.

[9]

Question 9

In order that a density $h(x)$ can be used to generate samples from another density $f(x)$, it is necessary that there is a constant C such that $Ch(x) > f(x)$ for all x.

This condition is not fulfilled by $h_1(x)$. Specifically, the ratio

$$\frac{f(x)}{h_1(x)} = \frac{x^{1.5}}{\Gamma(2.5)}$$

is unbounded, and this makes it impossible to find an appropriate C.

Among the possible dominating densities of the form $h_k(x)$, we have considered the case $k = 1$. For $k > 1$, let

$$C_k = \max_{x>0} \frac{f(x)}{h_k(x)} = \max_{x>0} \left(\frac{x^{1.5} e^{-x}}{\Gamma(2.5)} \times \frac{k}{e^{-x/k}} \right) = \max_{x>0} \left(\frac{k x^{1.5} e^{-x(1-1/k)}}{\Gamma(2.5)} \right)$$
$$= \frac{k \exp\left[\max_{x>0}\{1.5 \ln x - x(1-1/k)\}\right]}{\Gamma(2.5)}.$$

It follows from simple calculus that the function $1.5 \ln x - x(1-1/k)$ has a unique maximum at $x = 1.5k/(k-1)$. Therefore,

$$C_k = \frac{k \exp\left[1.5 \ln\left(\frac{1.5k}{k-1}\right) - 1.5\right]}{\Gamma(2.5)}.$$

In particular, $C_2 = 2.319/\Gamma(2.5)$ and $C_3 = 2.259/\Gamma(2.5)$, i.e., $C_3 < C_2$.

Since the reciprocals of C_2 and C_3 give the fraction of accepted samples when drawn from $h_2(x)$ and $h_3(x)$, respectively, one should choose $h_3(x)$ for higher rates of acceptance (less wastage).

[6]

Question 10

The necessary/desirable conditions are as follows.

- The policyholder must have an interest in the risk being insured.
- A risk must be of a financial and reasonably quantifiable nature.
- The probability of the event should be relatively small.
- There should be scope to pool large numbers of potentially similar risks.
- Individual risk events should be independent of each other.
- There should be an ultimate limit on the liability undertaken by the insurer.
- Moral hazards should be eliminated as far as possible.

[3]

Question 11

(i) Let X be the random variable representing the total loss and M be the deductible amount. The reinsurer's requirement is that $P(X > M) = 0.5$.

This condition is equivalent to

$$\left(\frac{500}{M+500}\right)^3 = 0.5, \text{ i.e., } \frac{M}{500} + 1 = \sqrt[3]{2}$$

which leads to the solution $M = 500\left(\sqrt[3]{2} - 1\right) = 129.96$.

(ii) Let Y be the random variable representing the claim amount net of deductible. It follows that

$$P(Y > y) = P(X - M > y | X > M) = \frac{P(X > M + y)}{P(X > M)} = \frac{\left(\frac{500}{M+y+500}\right)^3}{\left(\frac{500}{M+500}\right)^3}$$
$$= \left(\frac{629.96}{y + 629.96}\right)^3.$$

The distribution of Y is evidently Pareto with parameters $\alpha = 3$ and $\lambda = 629.96$. Therefore,

$$E(Y) = \frac{\lambda}{(\alpha - 1)} = \frac{629.96}{2} = 314.98.$$

(iii) Expected claim payment made by the direct insurer is

$$MP(X > M) + E(X|X < M)P(X < M).$$

Note that

$$\begin{aligned} E(X) &= E(X|X > M)P(X > M) + E(X|X < M)P(X < M) \\ &= [E(X - M|X > M) + M]P(X > M) + E(X|X < M)P(X < M) \\ &= E(Y)P(X > M) + MP(X > M) + E(X|X < M)P(X < M). \end{aligned}$$

Therefore,

$$\begin{aligned} MP(X > M) + E(X|X < M)P(X < M) &= E(X) - E(Y)P(X > M) \\ &= \frac{500}{2} - 314.98 \times 0.5 = 250 - 157.49 = 92.51. \end{aligned}$$

(iv) The reinsurer's likelihood based on the truncated data is

$$\prod_{i=1}^{n} \left[\frac{f(x_i)}{1 - F(M)}\right],$$

where M is the size of the deductible. For the Pareto distribution, the above expression simplifies to

$$\prod_{i=1}^{n}\left[\frac{\alpha\lambda^{\alpha}(\lambda+M)^{\alpha}}{(\lambda+x_i)^{\alpha-1}}\right].$$

(v) The reinsurer's likelihood based on the censored data is

$$[1-F(M)]^{n}\prod_{i=1}^{m}f(y_i),$$

where M is the size of the deductible. For the Pareto distribution, the above expression simplifies to

$$\frac{1}{(\lambda+M)^{n\alpha}}\prod_{i=1}^{m}\frac{\alpha\lambda^{\alpha}}{(\lambda+y_i)^{\alpha-1}}.$$

[14]