REVSTAT – Statistical Journal
Volume 4, Number 2, June 2006, 131–142

A LOGNORMAL MODEL FOR INSURANCE CLAIMS DATA

Authors: Daiane Aparecida Zuanetti
– Departamento de Estatística, Universidade Federal de São Carlos, São Paulo, Brazil

Carlos A. R. Diniz
– Departamento de Estatística, Universidade Federal de São Carlos, São Paulo, Brazil
dcad@power.ufscar.br

José Galvão Leite
– Departamento de Estatística, Universidade Federal de São Carlos, São Paulo, Brazil

Received: June 2005 Revised: February 2006 Accepted: February 2006

Abstract:

- In the insurance area, especially based on observations of the number of claims, $N(w)$, corresponding to an exposure w, and on observations of the total amount of claims incurred, $Y(w)$, the risk theory arises to quantify risks and to fit models of pricing and insurance company ruin. However, the main problem is the complexity to obtain the distribution function of $Y(w)$ and, consequently, the likelihood function used to calculate the estimation of the parameters.

  This work considers the Poisson($w\lambda$), $\lambda > 0$, for $N(w)$ and lognormal(μ, σ^2), $-\infty < \mu < \infty$, and $\sigma^2 > 0$, for Z_i, the individual claims, and presents maximum-likelihood estimates for λ, μ and σ^2.

Key-Words:

- *lognormal distribution; maximum-likelihood estimation; number of claims; total amount of claims.*

AMS Subject Classification:

- 62J02, 62F03.

1. INTRODUCTION

In the insurance area, the main goals of the risk theory are to study, analyze, specify dimensions and quantify risks. The risk theory is also responsible for fitting models of pricing and insurance company ruin, especially based on observations of the random variables for the number of claims, $N(w)$, and the total amount of claims incurred, $Y(w)$, defined as

$$Y(w) = \sum_{i=1}^{N(w)} Z_i \, I_{(N(w)>0)} \tag{1.1}$$

where the Z_i's are random variables representing the individual claims, $w = vt$ corresponds to the exposure, v denotes the value insured and t is the period during which the value v is exposed to the risk of claims.

Assuming that $N(w)$, $Z_1, Z_2, ...$ are independent and the individual claims are identically distributed, Jorgensen and Souza ([4]) discussed the estimation and inference problem concerning the parameters considering the situation in which the number of claims follows a Poisson process and the individual claims follow a gamma distribution.

Using the properties of the Tweedie family for exponential dispersion models ([8]; [3]), Jorgensen and Souza ([4]) determined, using the convolution formula, that $Y(w) \,|\, N(w)$ follows an exponential dispersion model and the joint distribution of $N(w)$ and $Y(w)/w$ follows a Tweedie compound Poisson distribution. For more details about exponential dispersion models read [2] and [3].

In spite of the distribution of the individual claim values being very well represented in some situations by the gamma distribution, in other cases it could be more suitable to attribute a lognormal distribution for $Z_1, Z_2, ...$. For instance, in collision situations in car insurances and in common fires, where the individual claim values can increase almost without limits but cannot fall below zero, with most of the values near the lower limit and where the natural logarithm of the individual claim variable yields a normal distribution.

The aim of this paper is to estimate the parameters of $Y(w) = \sum_{i=1}^{N(w)} Z_i I_{(N(w)>0)}$ and $N(w)$ distributions, where $N(w)$, $Z_1, Z_2, ...$ are independent, $Z_1, Z_2, ...$ is a sequence of random variables with lognormal(μ, σ^2) distribution and $N(w)$ follows a Poisson distribution with rate λ.

Simulated examples are given to illustrate the methodology. The use of a real dataset is not possible due to the high confidentiality with which the companies deal with their database.

2. LOGNORMAL MODEL

A positive random variable Z is lognormally distributed if the logarithm of the random variable is normally distributed. Hence Z follows a lognormal(μ,σ^2) distribution if its density function is given by

$$f_Z(z;\mu,\sigma^2) = \frac{(2\pi\,\sigma^2)^{-\frac{1}{2}}}{z}\,\exp\left\{-\frac{1}{2\,\sigma^2}\big(\log(z)-\mu\big)^2\right\}, \tag{2.1}$$

for $z > 0$, $-\infty < \mu < \infty$ and $\sigma > 0$.

The moments of the lognormal distribution can be calculated from the moment generating function of the normal distribution and are defined as

$$E[Z^k] = \exp\left(k\mu + \frac{1}{2}\,k^2\sigma^2\right). \tag{2.2}$$

Thus, the mean of the lognormal distribution is given by

$$E[Z] = \exp\left(\mu + \frac{1}{2}\,\sigma^2\right) \tag{2.3}$$

and the variance is given by

$$\mathrm{Var}[Z] = \exp\big(2\mu + 2\sigma^2\big) - \exp\big(2\mu + \sigma^2\big)\ . \tag{2.4}$$

Products and quotients of lognormally distributed variables are themselves lognormally distributed, as well as Z^b and bZ, for $b \neq 0$ and Z following a lognormal(μ,σ^2) distribution ([1]). However, the distribution of the sum of independent lognormally distributed variables, that appears in many practical problems and describes the distribution of $Y(w)|N(w)$, is not lognormally distributed and does not present a recognizable probability density function ([7]).

Approximations for the distribution of the sum of lognormally distributed random variables are suggested by Levy ([5]) and Milevsky and Posner ([6]).

3. PARAMETER ESTIMATION

As described in the previous section, the distribution function for $Y(w)$, where the claims Z_i are independently and identically lognormal(μ,σ^2) distributed, is not known. Consequently, the joint distribution for $(N(w), Y(w))$ and the corresponding likelihood function for the parameters μ, σ^2 and λ cannot be exactly defined.

However, since the lognormal distribution was defined with reference to the normal distribution, estimate μ, σ^2 and λ from the likelihood function for these parameters considering the variables $N(w)$ and $Y(w)$ is equivalent to estimate μ, σ^2 and λ from the likelihood function based on the variables $N(w)$ and

$$X_+(w) \;=\; \sum_{i=1}^{N(w)} X_i \, I_{(N(w)>0)} \ ,$$

where $N(w)$ follows a Poisson($w\lambda$), $X_i = \log(Z_i)$ follows a Normal(μ,σ^2) and the Z_i's are independent identically lognormal(μ,σ^2).

Then, we have

$$X_+(w)|N(w) \;=\; n \;\sim\; \text{Normal}(n\mu, n\sigma^2), \qquad \text{for} \quad n \geq 1 \ . \tag{3.1}$$

The joint density of $X_+(w)$ and $N(w)$, for $n \geq 1$, is defined as

$$\begin{aligned} & f_{(X_+(w),N(w))}\big(x_+, n; \mu, \sigma^2, \lambda\big) \;= \\ & \qquad = \; \frac{(w\lambda)^n}{n! \, \sqrt{2\pi n \, \sigma^2}} \, \exp\!\left\{ -\frac{1}{2n\,\sigma^2}(x_+ - n\mu)^2 - w\lambda \right\} I_{(0,\infty)}(x_+) \end{aligned} \tag{3.2}$$

and

$$f_{(X_+(w),N(w))}\big(x_+, 0; \mu, \sigma^2, \lambda\big) \;=\; \exp\{-w\lambda\} \, I_{(0,\infty)}(x_+) \ . \tag{3.3}$$

In this work, without loss of generality, w is assumed to be equal to 1. Considering $(x_{+1}, n_1), (x_{+2}, n_2), ..., (x_{+m}, n_m)$ observations from the independent random vectors $(X_{+1}, N_1), (X_{+2}, N_2) ..., (X_{+m}, N_m)$, where $N_i \sim \text{Poisson}(\lambda)$, $X_{+i} \,|\, (N_i\!=\!n_i) \sim \text{Normal}(n_i\mu, n_i\sigma^2)$, $i\!=\!1, 2, ..., m$, and m is the number of groups present in the portfolio and considering $\delta_i = 0$ for $N_i = 0$ and $\delta_i = 1$ for $N_i > 0$, the log likelihood function for the parameters μ, σ^2 and λ is given by

$$\begin{aligned} & l(\mu, \sigma^2, \lambda) \;= \\ & = \sum_{i=1}^{m} \left\{ \delta_i \left(-\frac{1}{2} \log\!\big(2\pi n_i \, \sigma^2\big) + n_i \log(\lambda) - \frac{1}{2n_i\,\sigma^2}\big(x_{+i} - n_i\mu\big)^2 - \lambda \right) + (1-\delta_i)\,(-\lambda) \right\} \end{aligned} \tag{3.4}$$

If $\sigma^2\!=\!\sigma_0^2$ is known the maximum likelihood estimates of μ and λ are given by

$$\hat{\mu} \;=\; \frac{\sum_{i=1}^{m} \delta_i X_{+i}}{\sum_{j=1}^{m} \delta_j N_j} \;=\; \frac{\sum_{i=1}^{m} X_{+i}}{\sum_{j=1}^{m} N_j} \qquad \text{if} \quad \sum_{j=1}^{m} N_j > 0 \ , \tag{3.5}$$

and

$$\hat{\lambda} \;=\; \frac{\sum_{i=1}^{m} \delta_i N_i}{m} \;=\; \frac{\sum_{i=1}^{m} N_i}{m} \ . \tag{3.6}$$

Let $S=\sum_{i=1}^{m} N_i$ be the total number of claims and $U=\sum_{i=1}^{m} X_{+i}$. Hence, S follows a Poisson$(m\lambda)$ and $U\,|\,(\mathbf{N}\!=\!\mathbf{n})$ follows a Normal$\Big(\mu\sum_{i=1}^{m} n_i,\ \sigma_0^2\sum_{i=1}^{m} n_i\Big)$, where $\mathbf{N}=(N_1,N_2,...,N_m)$ and $\mathbf{n}=(n_1,n_2,...,n_m)$ is the observed vector of number of claims for m groups. Thus $U\,|\,(S\!=\!s)$ follows a Normal$(\mu s,\sigma_0^2 s)$ and the exact distribution of $\hat{\lambda}$ is given by

$$P\left(\hat{\lambda}=\frac{c}{m}\right) = P\big(S=c\big) = \frac{\exp(-m\lambda)\,(m\lambda)^c}{c!} \qquad \text{for} \quad c=0,1,2,... \tag{3.7}$$

The cumulative distribution function of $\hat{\mu}$ given $S>0$, $F_{\hat{\mu}|S>0}(v)$, for $v\in R$ is

$$\begin{aligned}
P\left(\hat{\mu}\le v|S>0\right) &= P\left[\left(\hat{\mu}\le v\right)\cap\bigcup_{j=1}^{\infty}(S=j)|S>0\right]\\
&= \frac{P\left(\hat{\mu}\le v,\ \bigcup_{j=1}^{\infty}(S=j),\ S>0\right)}{P\big(S>0\big)}\\
&= \frac{P\left(\hat{\mu}\le v,\ \bigcup_{j=1}^{\infty}(S=j)\right)}{P\big(S>0\big)}\\
&= \frac{\sum_{j=1}^{\infty} P\big(\hat{\mu}\le v,\ S=j\big)}{P\big(S>0\big)}\\
&= \frac{\sum_{j=1}^{\infty} P\big(\hat{\mu}\le v|S=j\big)\,P\big(S=j\big)}{P\big(S>0\big)}\\
&= \frac{\sum_{j=1}^{\infty} P\big(\frac{U}{S}\le v|S=j\big)\,P\big(S=j\big)}{P\big(S>0\big)}\\
&= \frac{\sum_{j=1}^{\infty} P\big(\frac{U}{j}\le v|S=j\big)\,P\big(S=j\big)}{P\big(S>0\big)}\\
&= \frac{\sum_{j=1}^{\infty} P\big(U\le jv|S=j\big)\,P\big(S=j\big)}{P\big(S>0\big)}\\
&= \sum_{j=1}^{\infty} F_U(jv)\,\frac{\exp(-m\lambda)\,(m\lambda)^j}{j!\,\big(1-\exp(-m\lambda)\big)}\ ,
\end{aligned} \tag{3.8}$$

where F_U is the cumulative distribution function of the Normal$(\mu j,\sigma_0^2 j)$ distribution, for $j=1,2,...$.

The corresponding probability density function is defined as

$$
\begin{aligned}
f_{\hat{\mu}|S>0}(v) &= \frac{dF_{\hat{\mu}|S>0}(v)}{dv} \\
&= \sum_{j=1}^{\infty} f_U(jv)\, \frac{\exp(-m\lambda)\,(m\lambda)^j}{j!\,\big(1-\exp(-m\lambda)\big)}\; j \\
&= \frac{\exp(-m\lambda)}{1-\exp(-m\lambda)} \sum_{j=1}^{\infty} f_U(jv)\, \frac{(m\lambda)^j}{(j-1)!} \\
&= \frac{\exp(-m\lambda)}{1-\exp(-m\lambda)} \sum_{r=0}^{\infty} f_U\big((r+1)v\big)\, \frac{(m\lambda)^{r+1}}{(r)!} \\
&= \frac{(m\lambda)\exp(-m\lambda)}{1-\exp(-m\lambda)} \sum_{r=0}^{\infty} f_U\big((r+1)v\big)\, \frac{(m\lambda)^{r}}{(r)!}\;,
\end{aligned}
\tag{3.9}
$$

where f_U is the probability density function of the Normal$\big(\mu(r+1), \sigma_0^2(r+1)\big)$ distribution.

Let k be the number of groups with number of claims greater that zero. If σ^2 is unknown, the maximum likelihood estimate of σ^2 is

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{\sum_{i=1}^{m} \delta_i \left(\frac{(X_{+i} - N_i\,\hat{\mu})^2}{N_i} \right)}{\sum_{i=1}^{m} \delta_i} \\
&= \frac{\sum_{j=1}^{k} \left(\frac{(X_{+j} - N_j\,\hat{\mu})^2}{N_j} \right)}{k} \qquad \text{if} \quad N_j > 0, \quad \text{for all } j = 1, 2, ..., k\;.
\end{aligned}
\tag{3.10}
$$

Using the invariant principle of maximum-likelihood estimation, the estimates of $E[Z]$, $\mathrm{Var}[Z]$, $E[N]$ and $\mathrm{Var}[N]$, where Z represents the individual claims and N the number of claims, are, respectively

$$
\begin{aligned}
\widehat{E}[Z] &= \exp\left(\hat{\mu} + \frac{1}{2}\,\hat{\sigma}^2\right), \\
\widehat{\mathrm{Var}}[Z] &= \exp\left(2\,\hat{\mu} + 2\,\hat{\sigma}^2\right) - \exp\left(2\,\hat{\mu} + \hat{\sigma}^2\right), \\
\widehat{E}[N] &= \hat{\lambda}
\end{aligned}
$$

and

$$
\widehat{\mathrm{Var}}[N] = \hat{\lambda}\;.
$$

4. THE LOCATION PARAMETER μ AS A FUNCTION OF A COVARIATE

Suppose that $(x_{+1}, n_1), (x_{+2}, n_2), ..., (x_{+m}, n_m)$ are observations of the independent random vectors $(X_{+1}, N_1), (X_{+2}, N_2) ..., (X_{+m}, N_m)$, m is the number of groups present in the insurance portfolio, $N_i \sim \text{Poisson}(\lambda)$, and $X_{+i} \,|\, (N_i = n_i) \sim \text{Normal}(\mu_i, n_i\sigma^2)$, $i = 1, 2, ..., m$, with the following regression structure for the location parameter

$$\mu_i \,=\, \alpha\, n_i + \beta \sum_{j=1}^{n_i} v_{ij} \;,$$

where v_{ij} represents the covariate of the j-th individual claims of the i-th group, for $i = 1, 2, ..., m$ and $j = 1, 2, ..., n_i$.

Defining $r_i = \sum\limits_{j=1}^{n_i} v_{ij}$, the log likelihood function for the parameters α, β, σ^2 and λ is given by

$$l\big(\alpha, \beta, \sigma^2, \lambda\big) \,=$$

$$= \sum_{i=1}^{m} \left\{ \delta_i \left(-\frac{1}{2} \log\big(2\pi n_i\, \sigma^2\big) + n_i \log(\lambda) - \frac{1}{2 n_i\, \sigma^2} \big(x_{+i} - \alpha\, n_i - \beta\, r_i\big)^2 \right) + (-\lambda) \right\} .$$

Let k, $k \le m$, be the number of groups with the number of claims greater that zero, so that $\sum\limits_{j=1}^{k} N_j > 0$. The maximum likelihood estimates of α, β, σ^2 are obtained through the data of only these k groups and are given by

$$\text{(4.1)} \qquad \overset{\wedge}{\alpha} \,=\, \frac{\sum\limits_{j=1}^{k} X_{+j} \,-\, \overset{\wedge}{\beta} \sum\limits_{j=1}^{k} r_j}{\sum\limits_{j=1}^{k} N_j} \;,$$

$$\text{(4.2)} \qquad \overset{\wedge}{\beta} \,=\, \frac{\sum\limits_{j=1}^{k} \frac{X_{+j}\, r_j}{N_j} \,-\, \frac{\sum\limits_{j=1}^{k} X_{+j} \sum\limits_{j=1}^{k} r_j}{\sum\limits_{j=1}^{k} N_j}}{\sum\limits_{j=1}^{k} \frac{r_j^2}{N_j} \,-\, \frac{\left(\sum\limits_{j=1}^{k} r_j\right)^2}{\sum\limits_{j=1}^{k} N_j}} \;,$$

$$\text{(4.3)} \qquad \overset{\wedge}{\sigma^2} \,=\, \frac{\sum\limits_{j=1}^{k} \left(\frac{\big(X_{+j} - \overset{\wedge}{\mu}_j\big)^2}{N_j} \right)}{k} \;, \qquad N_j > 0, \quad \text{for all } \; j \;,$$

where $\overset{\wedge}{\mu}_j = \overset{\wedge}{\alpha}\, n_j + \overset{\wedge}{\beta}\, r_j$.

The maximum likelihood estimates of λ is defined as (3.6).

5. APPLICATIONS

In order to illustrate the methods outlined in this article, two simulated data set, with 20 insurance groups each, are presented. For the i-th group we generated one observation of N following Poisson(λ) and n_i observations of Z following lognormal(μ, σ^2) and we considered $X_j = \log(Z_j)$ for $j=1,2,...,N_i$ and $X_{+i} = \sum_{j=1}^{N_i} X_j$ for each group. These observations, together with the values of N, are used in the estimation of the parameters. The first data set was simulated considering a small rate of occurrence of claims in each insurance group and, consequently, a large probability of groups with zero claims. The second data set was simulated considering a large rate of occurrence of claims and, consequently, a small number of groups with zero claims. In both cases the values of μ and σ^2 considered in the simulation of the data was 7.1 and 0.1, respectively. Thus

$$E[Z] \;=\; 1274.11 \qquad \text{and} \qquad \mathrm{Var}[Z] = 170728.8 \;,$$

that is, the expected individual claim value is 1274.11 MU with a variance of 170728.8 MU.

5.1. Portfolio with small rate of occurrence of claim

Considering $m = 20$, $N \sim$ Poisson(2), and the Z_i's iid lognormal$(7.1, 0.1)$, we have

$$P\big[N=0\big] \;=\; \exp(-2) \;=\; 0.135 \;,$$

that is, the probability of occurrence of no claims in each group is equal to 0.135.

The simulated individual claim values vary between 634.48 MU and 2819.6 MU and the observed values of N, X_+ and δ are presented in Table 1. Note that four of the twenty groups have no occurrence of claims.

Table 1: Observed values of N, X_+ and δ for a simulated insurance portfolio

	N	X_+	δ
1	1	6.79	1
2	3	21.12	1
3	0	0.00	0
4	3	20.98	1
5	2	13.56	1
6	0	0.00	0
7	3	21.82	1
8	2	14.63	1
9	1	6.89	1
10	1	7.29	1
11	1	6.54	1
12	3	21.97	1
13	1	7.03	1
14	4	27.69	1
15	3	21.58	1
16	0	0.00	0
17	2	13.54	1
18	2	14.48	1
19	0	0.00	0
20	2	14.24	1
Total	34	240.17	16

The estimates of λ, μ and σ^2, calculated by (3.6), (3.5) and (3.10), respectively, as well as a comparison between the true values of the parameters and their estimates are presented in Table 2.

Table 2: The parameters true values and their estimates

	True value	Estimate	Difference
λ	2	1.7	0.3
μ	7.1	7.06	0.04
σ^2	0.1	0.09	0.01

From the distribution function of $\hat{\mu}$ given $S > 0$, defined in (3.8), we can calculate $P\big(\hat{\mu} \le v|S>0\big)$ for different values of $v \in R$. Table 3 shows the probability of $\hat{\mu} \le v$ given $S > 0$, considering $\lambda = 2$, $\mu = 7.1$, $\sigma^2 = 0.1$ (used for the data simulation) and $s = 34$ (observed in this dataset).

Table 3: $P\big(\hat{\mu} \le v|S>0\big)$ for $\lambda = 2$, $\mu = 7.1$, $\sigma^2 = 0.1$ and $s = 34$

v	$P\big(\hat{\mu} \le v\|S>0\big)$	v	$P\big(\hat{\mu} \le v\|S>0\big)$	v	$P\big(\hat{\mu} \le v\|S>0\big)$
4	0.0013	7	0.8063	10	0.9960
4.25	0.0060	7.25	0.8529	10.25	0.9974
4.5	0.0193	7.5	0.8913	10.5	0.9980
4.75	0.0484	7.75	0.9208	10.75	0.9986
5	0.0980	8	0.9424	11	0.9990
5.25	0.1721	8.25	0.9585	11.25	0.9993
5.5	0.2656	8.5	0.9709	11.5	0.9995
5.75	0.3687	8.75	0.9804	11.75	0.9996
6	0.4718	9	0.9863	12	0.9997
6.25	0.5787	9.25	0.9892	12.25	0.9998
6.5	0.6635	9.5	0.9925	2.5	0.9998
6.75	0.7473	9.75	0.9952	12.75	0.9999

Note that, from the results of Table 3,

$$P\big[4.5 \le \hat{\mu} \le 9.25\big] = 0.9699 .$$

5.2. Portfolio with large rate of occurrence of claim

In the second dataset, twenty observations of N were generated from the Poisson(100) distribution and the Z_i's were generated from the lognormal$(7.1, 0.1)$ distribution. Consequently,

$$P\big[N = 0\big] = \exp\big(-100\big) \simeq 0 ,$$

that is, the probability of occurrence of no claims in each group is practically null. The observed values of N, X_+ and δ are presented in Table 4 and the simulated individual claim values vary between 440.91 MU and 3212.9 MU.

Table 4: Observed values of N, X_+ and δ for a simulated insurance portfolio

	N	X_+	δ
1	95	676.0	1
2	104	739.3	1
3	85	601.6	1
4	92	652.6	1
5	106	749.6	1
6	111	791.9	1
7	85	600.8	1
8	86	606.6	1
9	108	762.0	1
10	85	601.1	1
11	98	695.8	1
12	86	612.0	1
13	83	586.1	1
14	100	709.8	1
15	79	553.4	1
16	95	674.5	1
17	87	619.8	1
18	105	747.6	1
19	101	717.3	1
20	100	714.6	1
Total	1891	13412.5	20

The estimates of λ, μ and σ^2, calculated by (3.6), (3.5) and (3.10), respectively, as well as a comparison between the true values of the parameters and its estimates are displayed in in Table 5.

Table 5: The parameters true values and their estimates

	True value	Estimate	Difference
λ	100	94.55	5.45
μ	7.1	7.093	0.007
σ^2	0.1	0.096	0.004

Table 6 shows the probability of $\overset{\wedge}{\mu} \leq v$ given $S > 0$, considering $\lambda = 100$, $\mu = 7.1$, $\sigma^2 = 0.1$ (used for the data simulation) and $s = 1891$ (observed in this dataset).

Table 6: $P\big(\overset{\wedge}{\mu} \leq v|S > 0\big)$ for $\lambda = 100$, $\mu = 7.1$, $\sigma^2 = 0.1$ and $s = 1891$

v	$P\big(\overset{\wedge}{\mu} \leq v\vert S > 0\big)$
6.1	0.0000
6.2	0.0001
6.3	0.0019
6.4	0.0150
6.5	0.0722
6.6	0.2216
6.7	0.4639
6.8	0.7152
6.9	0.8873
7.0	0.9671
7.1	0.9929
7.2	0.9989
7.3	0.9999
7.4	1.0000
7.5	1.0000

From the results of Table 6 we have $P[6.4 \leq \overset{\wedge}{\mu} \leq 7.1] = 0.978$.

6. CONCLUDING REMARKS

The theory for exponential dispersion models cannot be applied to estimate the parameters μ, σ^2, that specify the lognormal distribution of the individual claim value (Z), and λ, the occurrence rate of claims, because the lognormal distribution and, consequently, the joint distribution of $Y(w) = \sum_{i=1}^{N(w)} Z_i \, I_{(N(w)>0)}$ and $N(w)$ does not belong to the class of the exponential dispersion model. However, from the joint distribution of $X_+(w) = \sum_{i=1}^{N(w)} \log(Z_i) \, I_{(N(w)>0)}$ and $N(w)$, maximum likelihood estimates of μ, σ^2 and λ can be defined and applied to an insurance portfolio dataset, in which $N(w)$ follows a Poisson($w\lambda$) distribution and Z is lognormally distributed.

ACKNOWLEDGMENTS

We would like to thank the associate editor and referees for their valuable suggestions and comments on an earlier version of this paper.

REFERENCES

[1] Crow, E.L. and Shimizu, K. (Eds.) (1988). *Lognormal Distributions: Theory and Application*, Dekker, New York.

[2] Jorgensen, Bent (1986). Some properties of exponential dispersion models, *Scand. J. Statist.*, **13**, 187–198.

[3] Jorgensen, Bent (1987). Exponential dispersion models (with discussion), *J. Royal Statistical Society Ser. B*, **49**, 127–162.

[4] Jorgensen, B. and Souza, M.C.P. de (1994). Fitting Tweedie's compound poisson model to insurance claims data, *Scandinavian Actuarial Journal*, 69–93.

[5] Levy, E. (1992). Pricing European average rate currency options, *Journal of International Money and Finance*, **14**, 474–491.

[6] Milevsky, M.A. and Posner, S.E. (1998). Asian Options, the sum of lognormal, and the Reciprocal Gamma distribution, *Journal of Financial and Quantitative Analysis*, **33**, 3.

[7] Slimanc, S.B. (2001). Bounds to the distribution of a sum of independent lognormal random variables, *IEEE Transations on Communications*, **20**, 6.

[8] Tweedie, M.C.K. (1947). Functions of a statistical variate with given means, with special reference to Laplacian distributions, *Proc. Camb. Phil. Soc.*, **49**, 41–49.