Pricing Catastrophe Put Option and Related Issues

X. Sheldon Lin
Department of Statistics
University of Toronto
Toronto, Canada
sheldon@utstat.utoronto.ca

This presentation is based on joint works with Tao Wang and Simon Lee

IFID/MITACS Conference on Financial Engineering for Actuarial Mathematics

November 9, 2008

Outline of the Presentation

- Motivation
- Description of the Option and the Model
- Useful Results from Ruin Theory
- Derivation of Pricing Formulae
- Modelling Catastrophe Loss Data
- Risk Measures

Motivation

- In 1996, Centre Re issued a catastrophe equity put option contract to RLI Corporation. The contract gave RLI Corporation the right to issue up to $50 million cumulative convertible preferred shares in a catastrophic event. The contract was underwritten by AON.
- In 1997, Horace Mann Educators Corporation entered into a multi-year equity put program that allowed it to receive up to $100 million from Centre Re in exchange for an equivalent value of its convertible preferred shares.
- The catastrophic event was defined as that when the PCS catastrophic losses exceed a pre-specified level (the trigger).
- Catastrophe equity put options thus provide insurers with access to additional equity in the wake of huge catastrophic losses and hence diversify catastrophe risks from insurers.
- They allow the participation of investors in the insurance market, which improves the capital capacity of the reinsurance market.

Option Payoff

Let $S(t)$ be the stock price of an insurer and $L_R(t)$ the aggregate losses of the insured over the time period $(0, t]$.

Let $\mathcal{L}_R$ be the trigger level and K the strike price of a put.
If the option is exercised at time t, its payoff can be expressed as

$$\max\{K - S(t),\ 0\}\ \mathbb{I}_{\{L_R(t) > \mathcal{L}_R\}}, \tag{1}$$

where $\mathbb{I}_A$ is the indicator function.

Model for Stock Price

The aggregate losses $L_R(t)$ is modelled by a compound Poisson process:

$$L_R(t) = \sum_{i=1}^{N(t)} Y_i,$$

where Y_i's are individual catastrophic losses.

The underlying stock price $S(t)$ under a risk-neutral probability measure is of the form

$$S(t) \;=\; s\exp\{ct - \alpha L_R(t) + \sigma W(t)\} = s\exp\{ct - L(t) + \sigma W(t)\}, \tag{2}$$

where $0 < \alpha < 1$ is the impact parameter, $W(t)$ is a standard Brownian motion stochastically independent of $L_R(t)$.

For notational simplicity, let

$$L(t) = \alpha L_R(t) = \sum_{i=1}^{N(t)} X_i \ , \tag{3}$$

where X_i's have the common distribution function $P(x)$ and the scaled loss trigger $\mathcal{L} = \alpha\mathcal{L}_R$. Throughout, we assume a mixture of Erlangs for $P(x)$:

$$p(x) = \sum_{j=1}^{n} a_j \frac{x^{j-1}e^{-x/\theta}}{\theta^j (j-1)!} \ , \quad x > 0 \ , \tag{4}$$

where $a_j \geq 0, \ j = 1, 2, \cdots, n$, and $\sum_{j=1}^{n} a_j = 1$.

The risk-neutral assumption implies:

$$c = \delta - \frac{\sigma^2}{2} + k,$$

where k is the loss compensation rate and given by

$$k = \lambda[1 - \tilde{p}(1)] \ ,$$

where $\tilde{p}(\xi) = \int_0^\infty e^{-\xi x} dP(x)$ is the Laplace transform of the loss distribution function $P(x)$.

Both the stock price process and the aggregate losses process are stationary and of independent increments.

As a result, the optimal exercise strategy of a perpetual American put with the payoff (1) is of the form

$$T_{\varphi,\mathcal{L}} = \inf \{t;\ S(t) < \varphi \text{ and } L(t) > \mathcal{L}\}\,, \tag{5}$$

where φ is the level optimal exercise boundary that is to be determined.

That is, the option is exercised as soon as the stock price falls below the level φ and the aggregate losses exceed the trigger $\mathcal{L}$.

Price of the Perpetual American Option

Let $C_{\mathcal{L}}(s,\varphi)$ denote the price of the perpetual CatEPut with such a strategy $T_{\varphi,\mathcal{L}}$.

$$C_{\mathcal{L}}(s,\varphi) \;=\; \mathbb{E}\,[\,e^{-\delta T_{\varphi,\mathcal{L}}}\Pi(S(T_{\varphi,\mathcal{L}}))\mid S(0)=s\,], \tag{6}$$

where

$$\Pi(S) = \max\{K - S,\ 0\},$$

and δ is the force of interest or annual rate of interest compounded continuously.

$$
\begin{aligned}
C_{\mathcal{L}}(s,\varphi) &= \mathbb{E}\,[\, e^{-\delta T_{\mathcal{L}}} C(S(T_{\mathcal{L}}),\varphi)\, \mathbb{I}_{\{S(T_{\mathcal{L}})\geq\varphi\}}\; |S(0)=s\,] \\
&+ \mathbb{E}\,[\, e^{-\delta T_{\mathcal{L}}} \Pi(S(T_{\mathcal{L}}))\, \mathbb{I}_{\{S(T_{\mathcal{L}})<\varphi\}}\; |S(0)=s\,] \qquad (7)
\end{aligned}
$$

where $T_{\mathcal{L}} = \inf\{t;\ L(t) > \mathcal{L}\}$.

The first term represents the case where the stock price is still above the boundary when the aggregate losses exceed the trigger. At time $T_{\mathcal{L}}$ the holder holds the perpetual put as a regular perpetual American put.

The second term is for the case where the stock price is below the boundary when the aggregate losses exceed the trigger. the holder will thus exercise the option immediately.

The optimal exercise boundary for this perpetual put is the same as the optimal exercise boundary for the regular perpetual American put (with no trigger).

Let $U(t) = \ln \frac{S(t)}{\varphi}$.

$$U(t) = u + ct - L(t) + \sigma W(t), \quad t \geq 0, \tag{8}$$

where $u = \ln(s/\varphi)$.

The log-price process $U(t)$ is interpreted in ruin theory as an insurer's surplus, where u is the initial surplus, c is the premium rate, $L(t)$ is the aggregate losses up to time t and the diffusion term $\sigma W(t)$ represents the fluctuation of the surplus.

The exercise time $T_\varphi = \inf\{t|S(t) < \varphi\}$ is the same as the time of ruin $T = \inf\{t|U(t) < 0\}$.

The expected discounted penalty function

$$\phi(u) = \mathbb{E}[e^{-\delta T} w(U(T))\ \mathbb{I}_{\{T<\infty\}} \mid U(0) = u]\ , \tag{9}$$

where the penalty function

$$w(x_2) = \Pi(\varphi e^{-x_2}) \tag{10}$$

represents a penalty when the deficit at ruin is $|x_2|$.

The parameter δ is the force of interest.

Gerber and Landry's Results

$C(s,\varphi) = \phi(u)$ with $\ln(s/\varphi) = u$ satisfies the defective renewal equation

$$\begin{aligned} C(s,\varphi) &= \int_0^{\ln(s/\varphi)} C(se^{-y},\varphi) g(y) dy + \left(\frac{\varphi}{s}\right)^{\beta} \Pi(\varphi) + \int_{\ln(s/\varphi)}^{\infty} \Pi(se^{-y}) g(y) dy \\ &- \left(\frac{\varphi}{s}\right)^{\beta} \int_0^{\infty} \Pi(\varphi e^{-y}) g(y) dy \ . \end{aligned}$$

where

$$D = \frac{\sigma^2}{2}, \quad \beta = c/D + 1 \ ,$$

$c = \delta - D + \lambda[1 - \tilde{p}(1)]$, and

$$g(y) = \frac{\lambda}{D} \int_0^y e^{-\beta(y-z)} \int_z^{\infty} e^{-(x-z)} dP(x) dz.$$

The optimal exercise boundary:

$$\varphi^* = K \frac{\delta}{c + 2D - \lambda \int_0^{\infty} x e^{-x} p(x) dx} \ .$$

Lin and Willmot's Results

$$C(s,\varphi) = \frac{1+\gamma}{\gamma} \int_0^{\ln(s/\varphi)} \hat{H}(se^{-x},\varphi) dK(x) + \hat{H}(s,\varphi) , \quad s \geq \varphi , \tag{11}$$

where

$$\gamma = \frac{\delta}{k} ,$$

$$K(x) = \sum_{n=0}^{\infty} \frac{\gamma}{1+\gamma} \left(\frac{1}{1+\gamma} \right)^n G^{*n}(x) , \quad u \geq 0 ,$$

and

$$\hat{H}(s,\varphi) = \left(\frac{\varphi}{s}\right)^{\beta} \Pi(\varphi) + \int_{\ln(s/\varphi)}^{\infty} \Pi(se^{-y}) g(y) dy - \left(\frac{\varphi}{s}\right)^{\beta} \int_0^{\infty} \Pi(\varphi e^{-y}) g(y) dy .$$

The Joint Distribution of $T_{\mathcal{L}}$ and $L(T_{\mathcal{L}}) - \mathcal{L}$

Introduce the surplus process $\widetilde{U}(t) = \mathcal{L} - L(t)$.

This surplus process has the initial surplus $\mathcal{L}$ and no premium.

The time of ruin coincides with the triggering time $T_{\mathcal{L}}$.

The deficit at ruin $|\widetilde{U}(T_{\mathcal{L}})|$ is the same as the excess loss $L(T_{\mathcal{L}}) - \mathcal{L}$.

Define the penalty $w(x_2) = e^{zx_2}$.

The expected discounted penalty function is of the expression

$$\psi(\mathcal{L}, \delta, z) = \mathbb{E}\left[\, e^{-\delta T_{\mathcal{L}} - z|\widetilde{U}(T_{\mathcal{L}})|} \mid \widetilde{U}(0) = \mathcal{L}\right],$$

which is the Laplace transform of the joint distribution of $T_{\mathcal{L}}$ and $|\widetilde{U}(T_{\mathcal{L}})| = L(T_{\mathcal{L}}) - \mathcal{L}$.

Associated Defective Renewal Equation

The Laplace transform satisfies the defective renewal equation

$$\psi(u,\delta,z)=\frac{\lambda}{\lambda+\delta}\int_0^u \psi(u-x,\delta,z)p(x)dx+\frac{\lambda}{\lambda+\delta}\int_u^\infty w(x-u)p(x)dx,\ u\geq 0.$$

The same approach as Lin and Willmot's leads to an explicit formula for the joint distribution:

$$f(t,x) \ = \ \sum_{j=1}^{n} a_j\theta^{-j}\sum_{k=1}^{j}\frac{x^{j-k}e^{-x/\theta}}{(j-k)!}\sum_{i=0}^{\infty} h_i(t)\frac{\mathcal{L}^{i+k-1}}{\theta^i(i+k-1)!}e^{-\mathcal{L}/\theta}, \quad x>0, t>0\ , \quad (12)$$

where

$$h_i(t)=\sum_{j=0}^{i} b_{ij}\frac{\lambda^{j+1}t^j}{j!}e^{-\lambda t}\ ,$$

and b_{ij} is computed recursively by

$$b_{ij}=\sum_{k=j-1}^{i-1} a_{i-k}b_{k,j-1},$$

with $b_{00}=1$ and $b_{k0}=0, k=1,2,\cdots$.

$$C_{\mathcal{L}}(s,\varphi^*) = I(s,\varphi^*) + II(s,\varphi^*),$$

where

$$\begin{aligned} &I(s,\varphi^*) \\ &= \int_0^\infty \int_0^\infty \int_{\frac{\ln(\varphi^*/s)-ct+\mathcal{L}+x}{\sigma\sqrt{t}}}^\infty e^{-\delta t} C\left(s\exp\left\{\sigma\sqrt{t}y + ct - \mathcal{L} - x\right\}, \varphi^*\right) n(y) f(t,x) dy dt dx, \end{aligned}$$

where $n(y)$ is the density of the standard normal distribution, $f(t,x)$ is given in (12), and $C(s,\varphi^*)$ is given in (11), and

$$\begin{aligned} &II(s,\varphi^*) \\ &= \int_0^\infty \int_0^\infty \left[K e^{-\delta t} N(-d_2) - s e^{kt-\mathcal{L}-x} N(-d_1)\right] f(t,x) dt dx \ , \end{aligned}$$

where

$$d_{1,2} = \frac{\ln(s/\varphi^*) + (\delta+k)t - \mathcal{L} - x}{\sigma\sqrt{t}} \pm \frac{\sigma\sqrt{t}}{2} \ ,$$

and $N(y)$ is the cdf of the standard normal distribution.

Remarks

The approaches may be used for pring other types of options.

Let $w(x_2) = e^{zx_2}$ in (9). One obtains the joint distribution of the barrier hitting time of the stock price and the stock price at the hitting time.

This distribution may be used to price digital options and barrier options.

Fitting the Mixture of Erlangs to Catastrophe Loss Data

The proposed distribution for $P(x)$:

$$p(x|\theta, \{a_{n_i}\}) = \sum_{i=1}^{M} a_{n_i} \frac{x^{n_i-1} e^{-x/\theta}}{\theta^{n_i}(n_i - 1)!} \ , \quad x > 0 \ , \tag{13}$$

where θ and a_{n_i}'s are to be estimated.

Theoretical justification (Tijms Approximation):

Let

$$\hat{p}(x|\theta) = \sum_{i=1}^{\infty} (P(i\theta) - P((i-1)\theta)) \frac{x^{i-1} e^{-x/\theta}}{\theta^{i}(i-1)!}$$

Then,

$$\lim_{\theta \to 0} \hat{P}(x|\theta) = P(x),$$

for all continuous x.

An EM Algorithm

A MLE based expectation-maximization algorithm for incomplete data.

Let $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ be an incomplete sample generated from a pair of random variables (X, Y) with joint density $p(x, y|\Phi)$,

where Y is an unobservable random variable and Φ is the set of parameters to be estimated.

The corresponding complete data is $\{(x_1, Y_1), (x_2, Y_2), \cdots, (x_n, Y_n)\}$,

The complete-data log-likelihood is given by

$$l(\Phi|\boldsymbol{x}, \boldsymbol{Y}) = \sum_{i=1}^{n} \ln p(x_i, Y_i|\Phi)$$

Given the sample $\boldsymbol{x}$ and the current estimate of the parameters $\Phi^{(k-1)}$, the posterior distribution of Y_i is given by

$$q(y_i|x_i, \Phi^{(k-1)}) = \frac{p(x_i, y_i|\Phi^{(k-1)})}{p(x_i|\Phi^{(k-1)})},$$

where $p(x|\Phi^{(k-1)})$ is the marginal density.
The predictive log-likelihood is given by

$$Q(\Phi|\Phi^{(k-1)}) = \sum_{i=1}^{n} E\{\ln p(x_i, Y_i|\Phi)\} = \sum_{i=1}^{n} \int [\ln p(x_i, y_i|\Phi)] q(y_i|x_i, \Phi^{(k-1)}) dy_i$$

The next step is to maximize the log-likelihood:

$$\Phi^{(k)} = \max_{\Phi} Q(\Phi|\Phi^{(k-1)})$$

The maximization procedure is often difficult.

An EM Algorithm for the Mixture of Erlangs

The joint distribution is given by

$$p(x, y|\Phi) = a_y \frac{x^{y-1} e^{-x/\theta}}{\theta^y (y-1)!}$$

The posterior diatribution is given by

$$q(y_i|x_i, \Phi) = a_{y_i} \frac{x_i^{y_i-1} e^{-x_i/\theta}/[\theta^{y_i}(y_i-1)!]}{\sum_{y=1}^{M} a_y x_i^{y-1} e^{-x_i/\theta}/[\theta^y (y-1)!]}$$

and the corresponding log-likelihood

$$Q(\Phi|\Phi^{(k-1)}) \propto \sum_{i=1}^{n} \sum_{y=1}^{M} \left(\ln a_y - \frac{x_i}{\theta} - y \ln \theta \right) q(y|x_i, \Phi^{(k-1)})$$

The usual optimization method gives

$$a_y = \frac{1}{n}\sum_{i=1}^{n} q(y|x_i, \Phi^{(k-1)}),\ \ y = 1, 2, \cdots, M,$$

and

$$\theta = \frac{\sum_{y=1}^{M}\sum_{i=1}^{n} x_i q(y|x_i, \Phi^{(k-1)})}{\sum_{y=1}^{M} y \sum_{i=1}^{n} q(y|x_i, \Phi^{(k-1)})}$$

This algorithm is purely an interative algorithm and no maximization is involved!

As a result, efficiency is guaranteed.

Testing the Algorithm Using Common Distributions

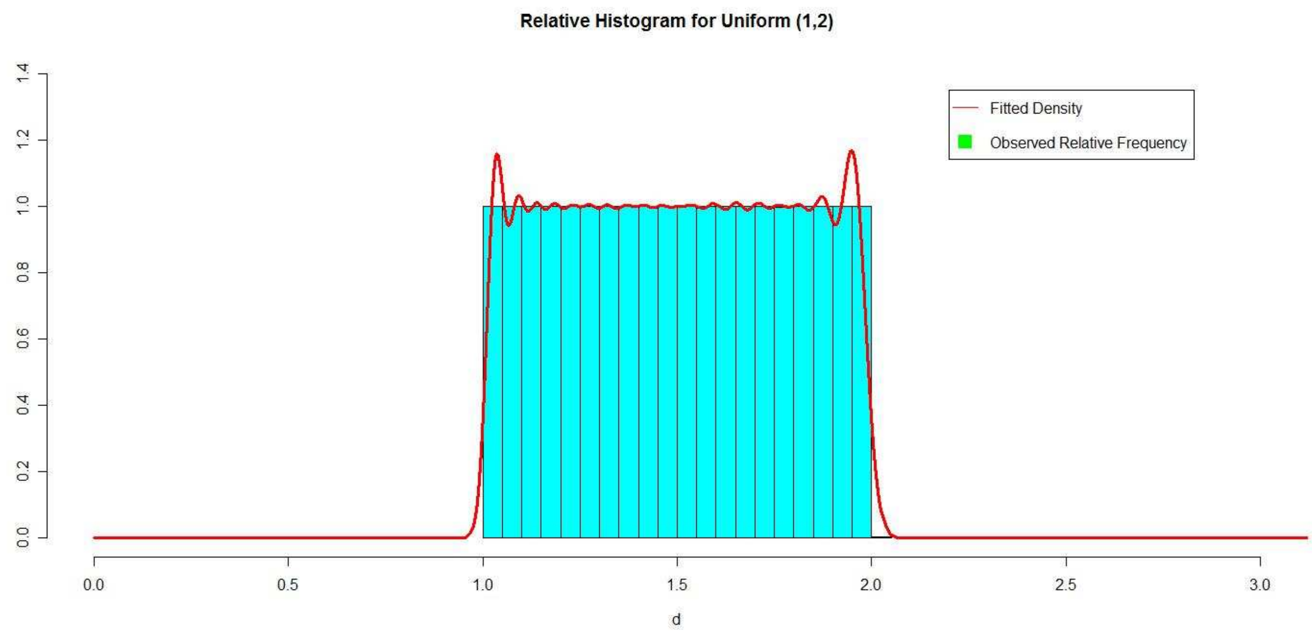

Figure 1: Histogram of uniform distribution and line for the fitted distribution

generated uniformly for study. In this study, we will use Dufresne's result as a comparison.

The EM algorithm is applied to the data. A mixture of 19 Erlangs with a fixed θ of 0.000498 fits the data well with each Erlang having significant weight. The parameters that maximize the log-likelihood function can be found in the appendix. The resulting log-likelihood value is -83.18.

5.1 Graphical comparison of the fitted distribution and underlying distribution

The fitted distribution is graphed with the empirical histogram overlay in the Figure 1. It can be observed that the fitted curve tightly envelops the relative histogram with slight overshoot near $x = 1^-$ and $x = 2^+$. In Figure 2, the QQ plot provides another evidence of overshooting at the extremes. PP plot suggests that the proposed model has a good fitness.

5.2 Statistical tests

The following table summarizes the results of the statistical tests suggested in the previous section.

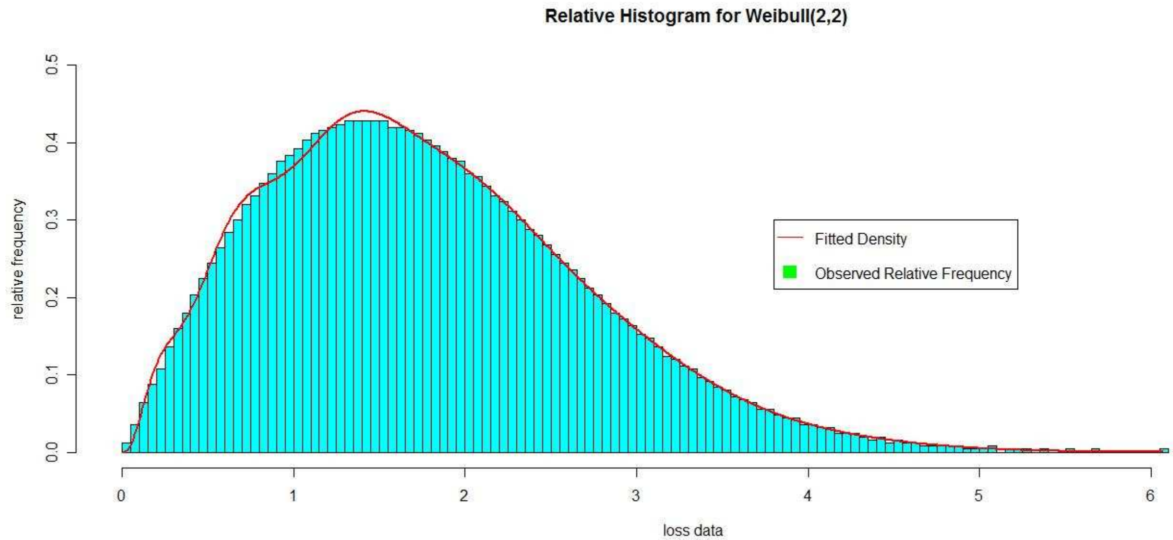

Figure 3: Histogram for Weibull(2,2) and the fitted density using a mixture of 6 Erlangs

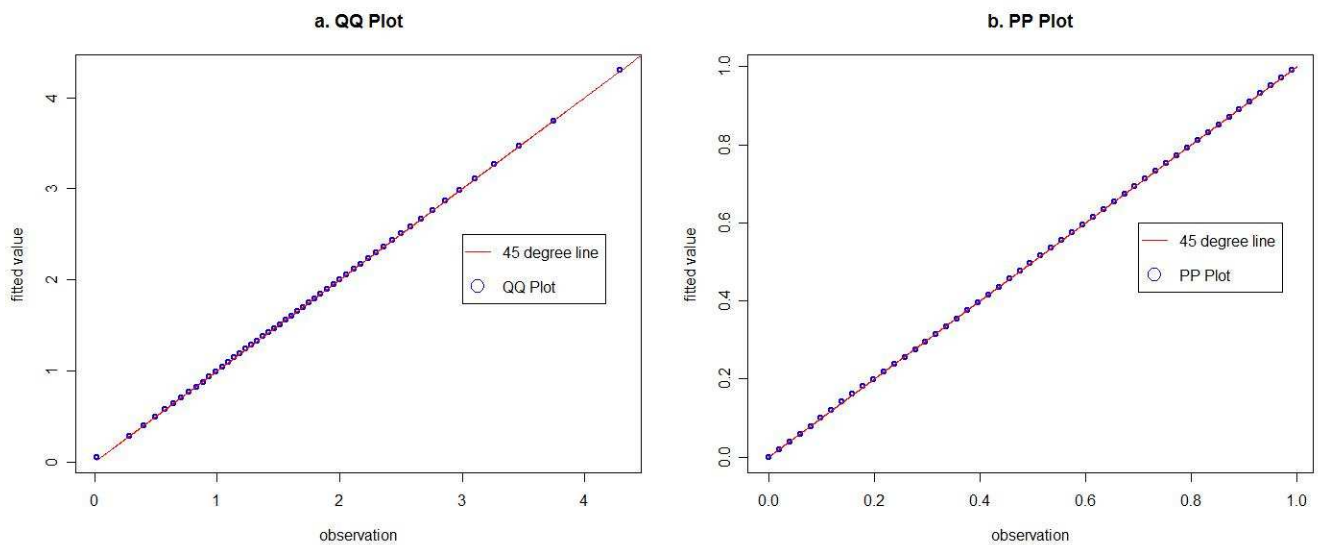

Figure 4: PP and QQ plots for Weibull(2,2) and the fitted distribution

The plots in Figures 3 and 4 show our proposed model has an almost complete fit for the light tailed Weibull distribution.

Example 2: Heavy tailed Weibull distribution

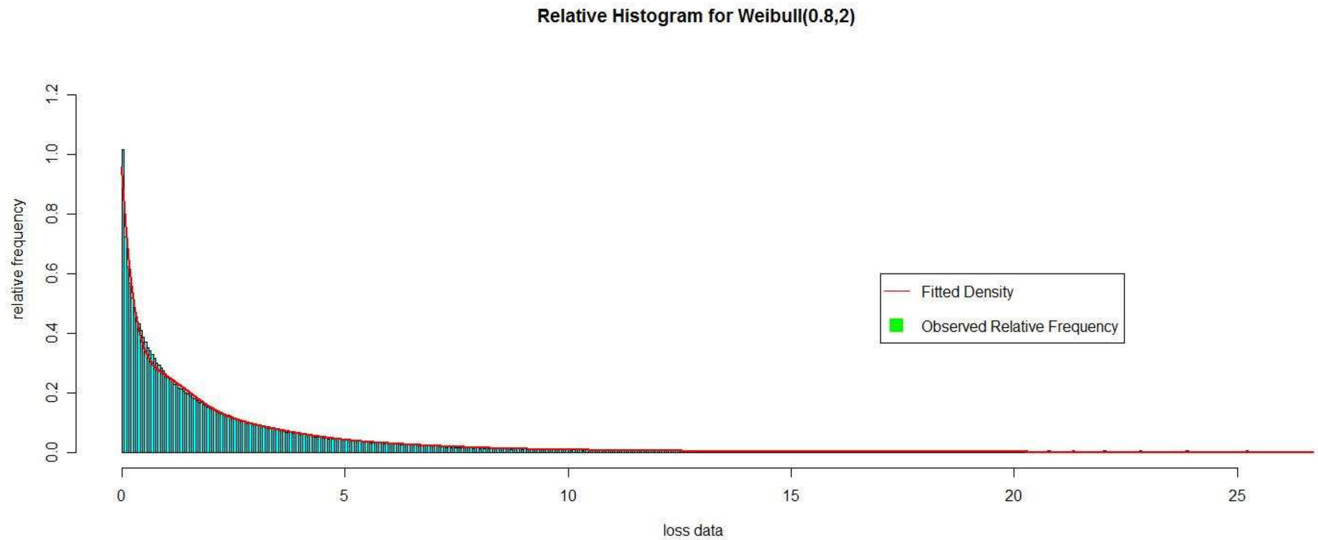

Figure 5: Histogram for Weibull(0.8, 2) and the fitted density using a mixture of 10 Erlangs

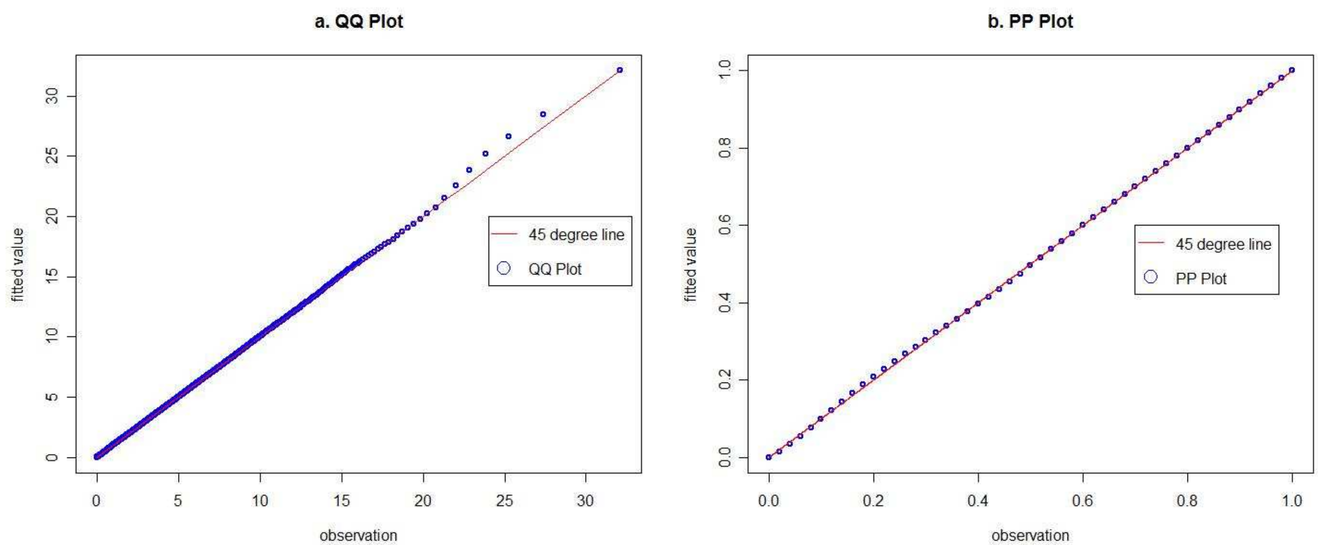

Figure 6: PP and QQ plots for Weibull(0.8, 2)

The plots in Figures 5 and 6 show our proposed model again performs well in fitting for the heavy tailed Weibull distribution. The QQ plot suggests that the fitness of the tail is not perfect. However, the problem can be solved by using more Erlangs. The effect of increasing the number of Erlangs will be shown in the example for the lognormal distribution.

Example 3: Pareto distribution

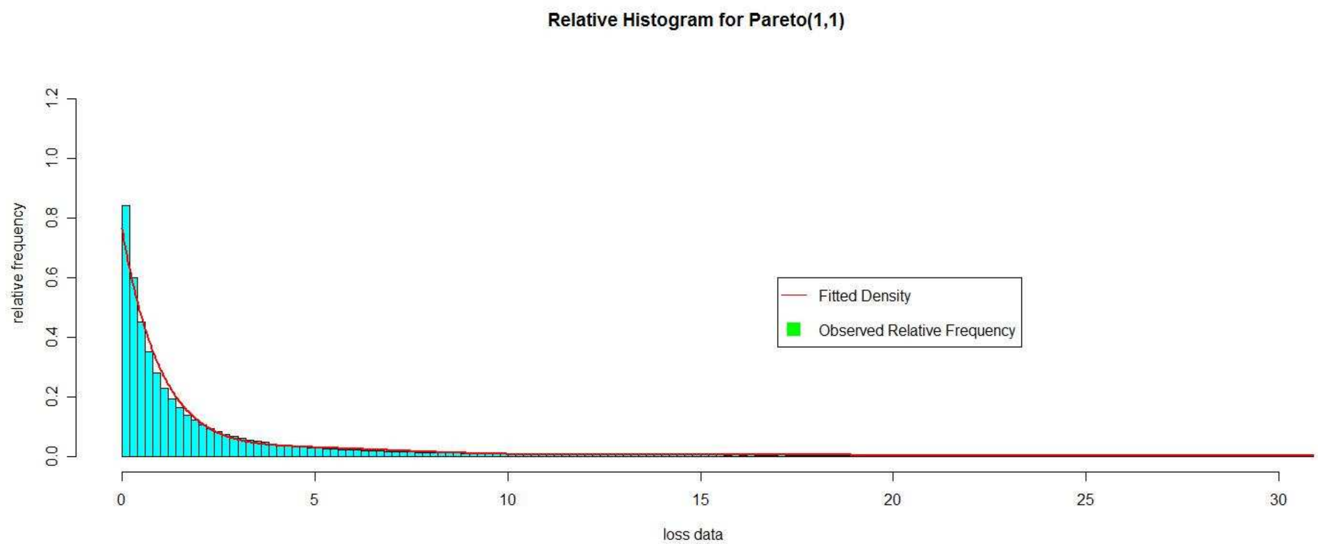

Figure 7: Histogram for Pareto(1, 1) and the fitted density using a mixture of 5 Erlangs

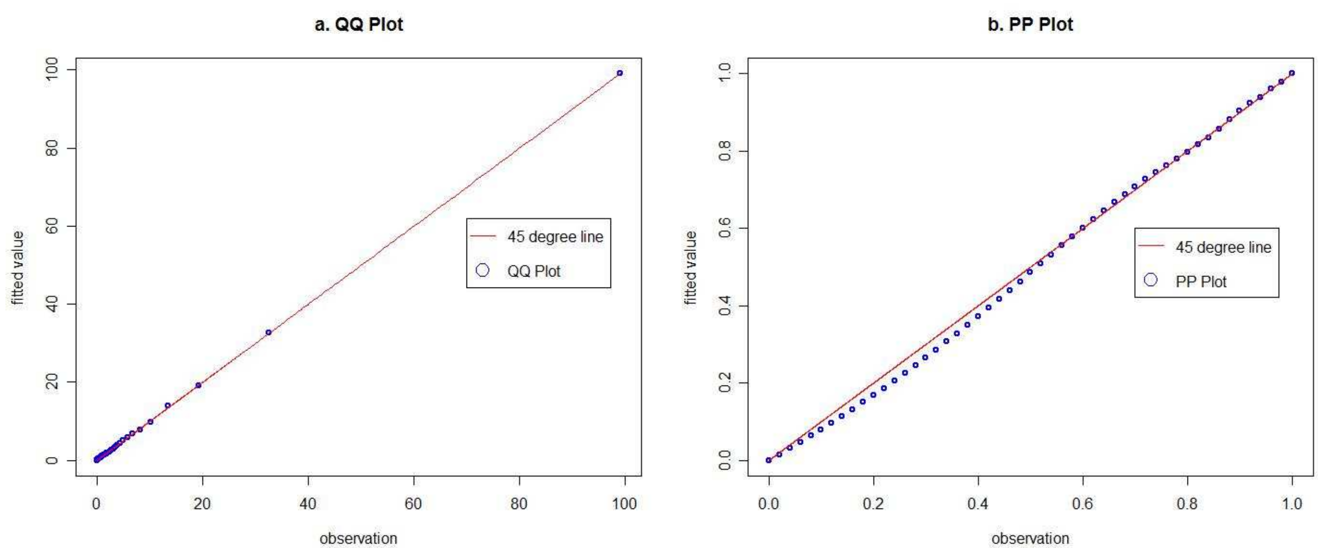

Figure 8: PP and QQ plots for Pareto(1, 1)

As shown in Figures 7 and 8, the plots again show our model has a good fitness for Pareto distribution.

Example 4: Lognormal distribution

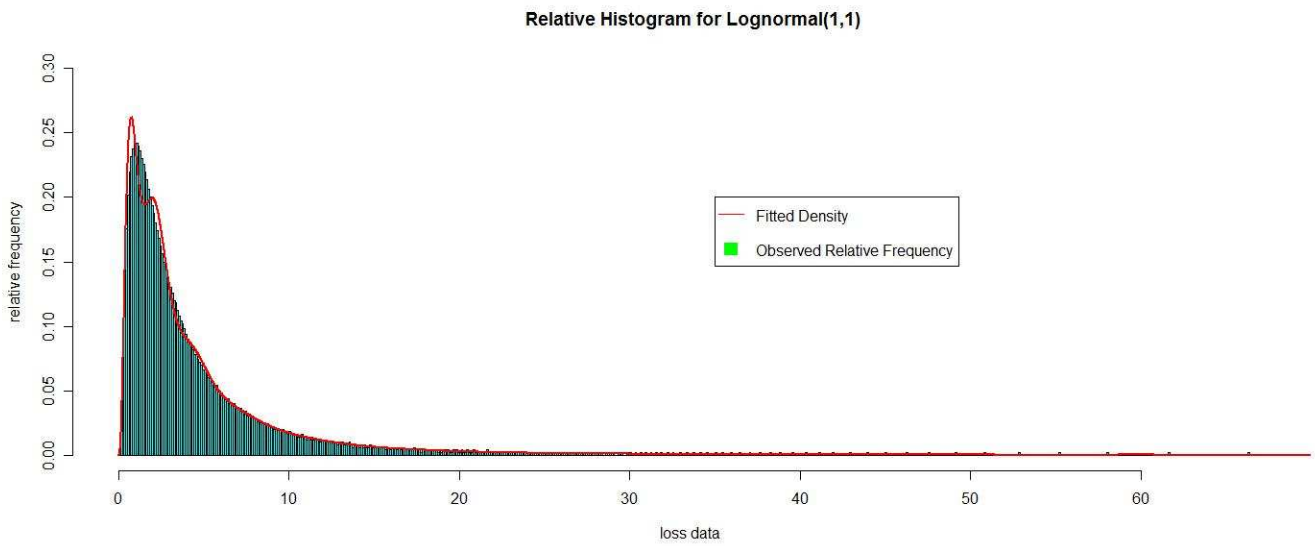

Figure 9: Histogram for lognormal(1, 1) and the fitted density using a mixture of 15 Erlangs

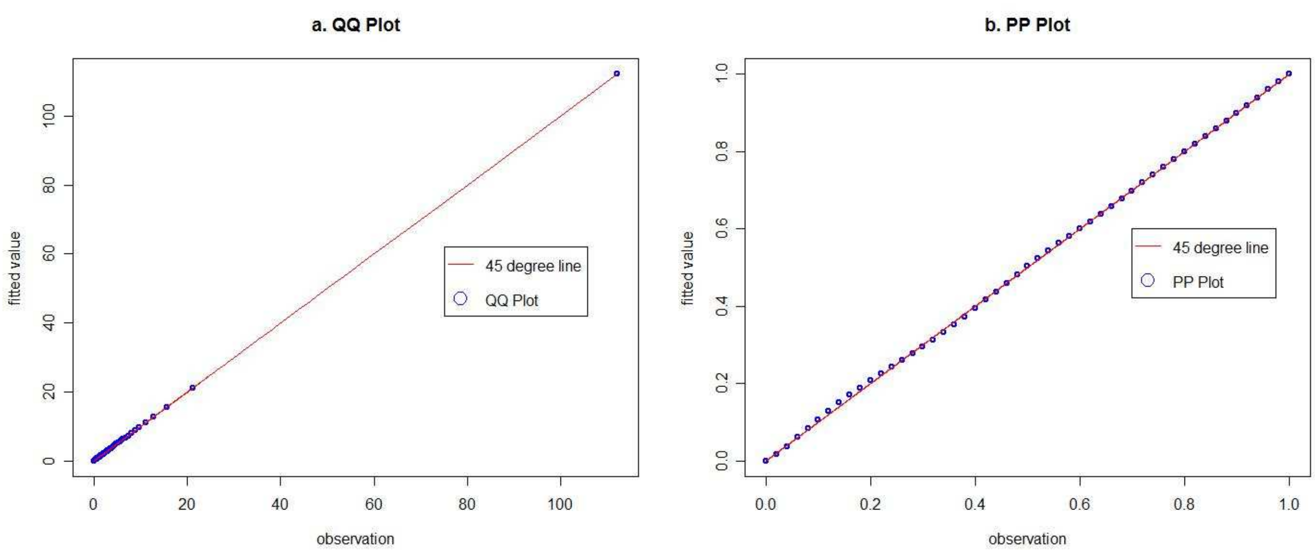

Figure 10: PP and QQ plots for lognormal(1, 1)

By increasing the number of Erlangs used for fitting to 15, the fitness for heavy tailed distribution improves significantly. As shown in Figures 9 and 10, the plots again show our model has an almost perfect

Fitting PCS Catastrophe Data

Data: 1271 catastrophe losses in US from 1997 to 2005.

Some stylish facts:

1. The data is multi-modal.
2. The maximum value of the data is 247 times of the mean.
3. There are 9.13% of the observations categorized as outliers.
4. The skewness and kurtosis for the data are 23.04 and 619.63.
5. 56% of the data is smaller than 0.1% of the maximum value while 96.6% of the data is smaller than 1% of the maximum value.

All of the above point to that the data is irregular and heavy tailed.

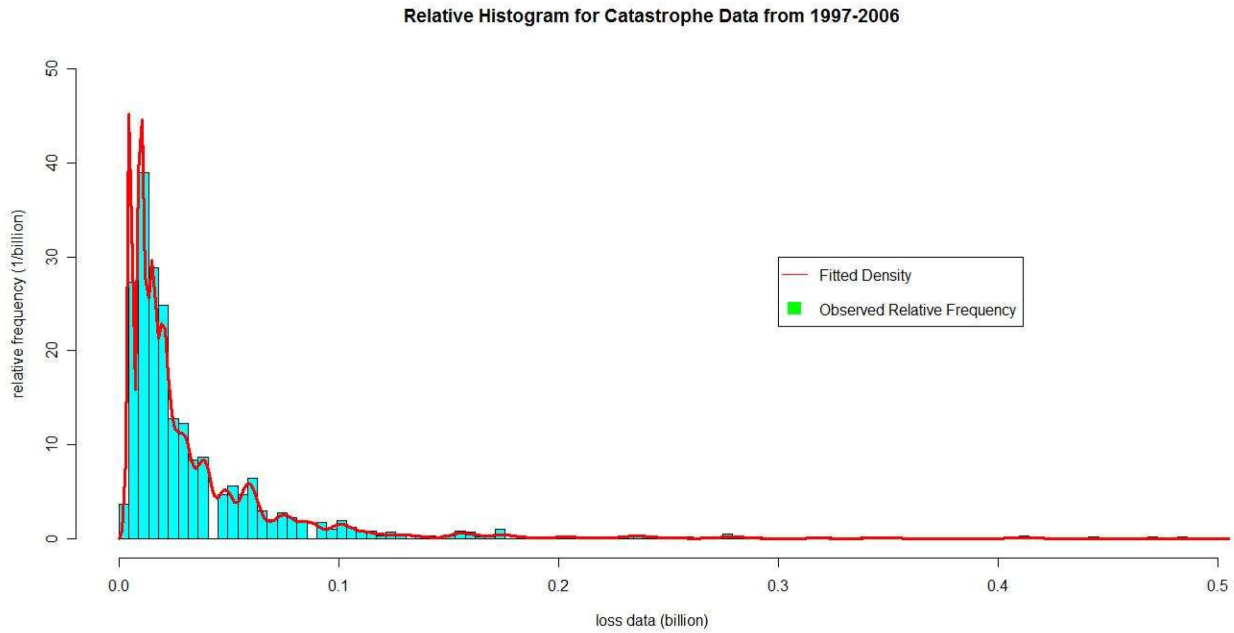

Figure 11: Histogram of observed loss and line for the fitted distribution

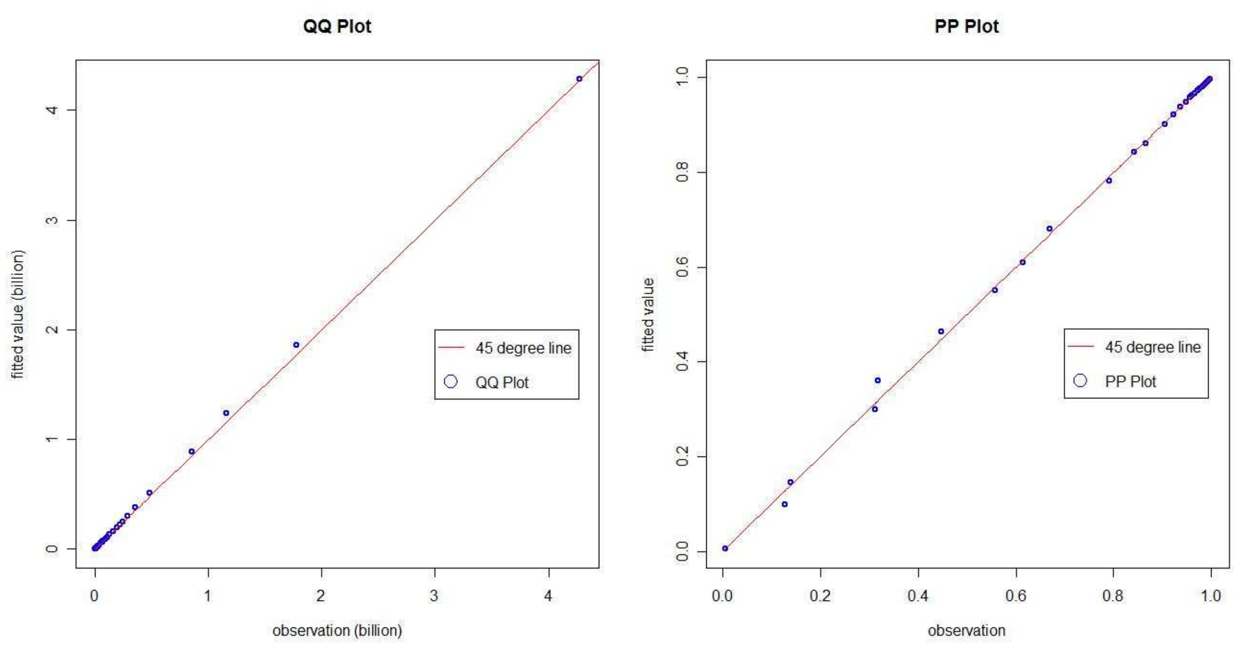

Figure 12: Histogram of observed loss and line for the fitted distribution

Goodness of Fit

nth moment	Empirical	Fitted	Fitted/Empirical	Percentage Difference (%)
1	$9.833 * 10^7$	$9.833 * 10^7$	1.0000	0.00%
2	$6.917 * 10^{17}$	$6.908 * 10^{17}$	0.9987	-0.13%
3	$1.317 * 10^{28}$	$1.315 * 10^{28}$	0.9983	-0.17%
4	$2.932 * 10^{38}$	$2.926 * 10^{38}$	0.9979	-0.21%
5	$6.857 * 10^{48}$	$6.840 * 10^{48}$	0.9975	-0.25%

Table 1: Catastrophe data–raw moments

Quantities	Empirical	Fitted	Fitted/Empirical	Percentage Difference (%)
Mean	98.33 million	98.33 million	1.0000	0.00%
Standard Deviation	825.85 million	825.31 million	0.9993	-0.07%
Skewness	23.03	23.04	1.0003	0.03%
Kurtosis	619.28	619.63	1.0006	0.06%

Table 2: Catastrophe data–central moments

Related Issues
Calculation of VaR and CVaR (TVaR, CTE, Expected Shortfall)

VaR V_p at security level p:

V_p is the solution of

$$e^{-V_p/\theta} \sum_{i=0}^{M} Q_i \frac{V_p^i}{\theta^i i!} = 1 - p$$

where $Q_i = \sum_{j=i+1}^{M} a_j$.

Conditional VaR at security level p:

$$CVaR_p = \frac{\theta e^{-V_p/\theta}}{1-p} \sum_{i=0}^{M} Q_i^* \frac{V_p^i}{\theta^i i!} + V_p,$$

where $Q_i^* = \sum_{j=i}^{M} Q_j$.

Distribution and Risk Measures of Aggregate Losses

Collective Risk Model:

$$S = \sum_{n=1}^{N} X_n$$

Where N is the number of losses and the iid sequence $X_n : n = 1, 2, \cdots$, are successive loss amounts following a mixture of Erlangs.

Its distribution is a zero-modified mixture of Erlangs with positive density

$$f_S(x) = \sum_{k=1}^{\infty} \eta_k \frac{x^{i-1} e^{-x/\theta}}{\theta^i (i-1)!}.$$

Here $\eta_k, \ k = 0, 1, \cdots$, are the coefficients of the power series $P_N(P_a(z))$, where $P_N(z)$ is the probability generating function of N and $P_a(z) = \sum_{i=1}^{M} a_i z^i$.

Its Var and CVaR are obtainable.

Individual Risk Model:

$$S = S_1 + S_2 + \cdots + S_n.$$

S_j's are mixtures of Erlangs, independent but not identical.

Each S_j may represent a individual risk or a collective risk.

Its distribution is again a zero-modified mixture of Erlangs with positive density

$$f_S(x) = \sum_{k=1}^{\infty} \eta_k \frac{x^{i-1} e^{-x/\theta}}{\theta^i (i-1)!},$$

Where $\eta_k,\ k = 0, 1, \cdots$, are the coefficients of the power series $\prod_{j=1}^{n} P_{a_j}(z)$.