Pricing catastrophe reinsurance risk premium using Peaks Over Threshold (POT) model

A S Jaladri, S Nurrohmah and I Fithriani
Department of Mathematics, Faculty of Mathematics and Natural Sciences (FMIPA), Universitas Indonesia, Depok 16424, Indonesia
Corresponding author’s email: snurrohmah@sci.ui.ac.id

Abstract. An incident that cost more lives than a certain threshold is called a catastrophic event. In the world of insurance, the incidence of catastrophe causes many policies to make a claim together, so it can cause big losses for insurance companies. The risk of catastrophic events can be minimized by purchasing a catastrophic reinsurance contract. The reinsurance company then calculates the amount of risk premium for catastrophic reinsurance to be paid by the insurance company. The model used to calculate the risk premium is the peaks over threshold (POT) model. The POT model is generally used to model extreme events. In this model, the number of catastrophic events is modeled using the Poisson distribution. Then, the number of casualties is modeled using the discrete generalized Pareto distribution (DGPD). To estimate the parameters of the model, the maximum likelihood method is used and the data on the number of fatalities resulting from a particular event in Indonesia is collected and compiled. Furthermore, the number of claims is modeled with the Beta-Binomial distribution and the size of claims resulting from a catastrophic event is modeled by Exponential distribution. Numerical simulations are then performed to obtain the total size of claims resulting from all catastrophic events within 1 year of catastrophic reinsurance contracts. Ultimately the risk premium can be calculated using the standard deviation premium principle by utilizing the expected value and the variance of the total amount of the claim.

Keywords: Beta-binomial, catastrophic event, discrete generalized pareto, POT, risk premium

1. Introduction
One of the risk concerning life insurance companies is the risk of catastrophic events. Indonesia territory is located in tropical climate area with the characteristics of climate change, temperature and wind direction that is quite critical. This condition can cause some natural disasters such as floods, landslides, forest fires, and drought. Another natural disasters that occur in Indonesia are earthquake, volcano eruption, and tsunami. Given the enormous risk of catastrophic threats in Indonesia, it is important for insurance companies to control the risk of catastrophic events. Based on [1], insurance company has a limited capital to cover all losses caused by catastrophic events, therefore they need to transfer the risk to the reinsurer, that is to purchase a catastrophic reinsurance contract. The catastrophic reinsurance contract will protect the insurance company from the immense risk of claims derived from a group of policyholders that collectively claim their loss caused by a particular event.

In the catastrophic reinsurance scheme the ceding company must determine the deductible retention or the part paid by ceding company. If a catastrophic event occurs with a minimum number of victims
as agreed, the reinsurer will help to pay the loss of the insurance company. The reinsurer pays the excess of loss that exceed retention but the maximum payable by reinsurer (aggregate limit) also needs to be limited in accordance with the agreement. The catastrophic reinsurance contract is unique.

The model that has been developed by Strickler is usually used to calculate the risk premium of catastrophic reinsurance contract. Strickler uses data from The Statistical Bulletin of the Metropolitan Life Insurance Company in New York that provides a summary of the events that led to the deaths of more than five individuals in the United States in the period 1946–1950. According to Erland et al. [2], the model has some drawbacks which are there is no statistical method to update the model according to new data. Besides that, the model is also limited to the occurrence of catastrophic event that cause a maximum of 1500 deaths. Therefore, to overcome the drawbacks in the model, this paper suggests peaks over threshold model.

2. Peaks over threshold model

The extreme value theory (EVT) method is one of the most preferred methods in analyzing extreme events. The EVT method provides a strong theoretical basis for constructing statistical models that can describe extreme events [3]. This method construct a statistical model which focuses on extreme events that has extraordinarily large values. This method can provide a better estimate in the tail of a distribution rather than the overall fitting method.

There are two different models in EVT which are peaks over thresholds (POT) and block maxima. These two models can be differentiated by how each model classifies extreme observations in the data and use it in the process of data analysis. In block maxima model, for each period, e.g. weekly, monthly, or yearly, samples of maximum or minimum observations are taken and considered as extreme observations. In peaks over threshold model, a fixed threshold \( u \) are needed rather than the period of observation. An observation will be considered as an extreme event if it exceeds the value of the threshold \( u \).

The use of block maxima model is usually found in analyzing seasonal data, e.g. weather and hydrological data. The challenge of using block maxima model is to determine the size of the blocks of each period. Block maxima model is usually used to determine the distribution of the maximum value of the existing data. While the POT model is often used in analyzing financial and insurance data that tend to be independent of time. The advantages of the POT model compared to the maxima block model are the POT model using the data more efficiently as it maximizes all data information by not grouping the data and taking into account all the extreme values in the data.

2.1. The number of catastrophic events

Suppose that the number of catastrophic events can be expressed by a random variable \( K_m(T) \), where \( T \) denotes the duration of the contract and usually the contract period is one year. \( K_m(T = 1) = K_m \) will satisfy the definition of the Poisson process. Hence the first assumption is,

**Assumption 1.** \( K_m \sim \text{Poisson} \left( \lambda_m \right) \)

where the parameter \( \lambda_m \) is the rate of the Poisson distribution. \( \lambda_m \) can be estimated by using maximum likelihood method.

2.2. The number of deaths

Suppose a random variable \( X'' \) represents the number of deaths resulting from a particular event. Then suppose a random variable \( X' = X'' \mid X'' \geq m \) which states the number of deaths from a catastrophic event that causes the victim at least \( m \) people. Based on the POT model, it is known that \( X' \) will have generalized Pareto distribution (GPD), so it is assumed,
Assumption 2. $X' \sim GPD \left( m - \frac{1}{2}, \sigma, \xi \right)$, $X'$ has a distribution function,

$$F \left( (m, \sigma, \xi) \right) (x') = 1 - \left( 1 + \xi \frac{(x' - m + \frac{1}{2})}{\sigma} \right)^{-\frac{1}{\xi}},$$

where $m \in \mathbb{R}^+, x' \geq m - \frac{1}{2}$, and $\sigma > 0$ with $m - \frac{1}{2}$ are location parameters for GPD.

Generalized Pareto distributions are distributions of continuous random variables, whereas $X'$ which states the number of fatalities of a catastrophic event is unlikely to be a continuous random variable, but a discrete random variable. So a technique is required to convert a continuous $X'$ random variable into a discrete random variable namely $X$ commonly referred to as a discretization method. This method will dissipate a continuous $X'$ random variable to a discrete $X$ random variable, where the discrete distribution is a discrete generalized Pareto distribution (DGPD). The density function form of DGPD is as follows,

$$f_X(x) = \left( 1 + \xi \frac{(x - m)}{\sigma} \right)^{-\frac{1}{\xi}} \left( 1 + \frac{x - m + 1}{\sigma} \right)^{-\frac{1}{\xi}},$$

where $x \in \{m, m + 1, m + 2, \ldots \}, \sigma > 0$, and $\xi \in \mathbb{R}$.

Parameters $\sigma$ and $\xi$ are estimated by using the maximum likelihood method, so the required assumption is

Assumption 3. $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables to the distribution of DGPD $(m, \sigma, \xi)$.

2.3. The number of claims
Suppose a random variable $Y'$ represents the number of claims to an insurance company resulting from the occurrence of a catastrophic incident that claimed the lives of as many as $x$ people. Then defined the market penetration of insurance companies $q$ as follows,

$$q = \frac{\text{Number of Policies sold}}{\text{Total Population}}$$

Then assumed,

Assumption 4. $E \left[ Y' \right] = xq$.

On this assumption the number of claims of a catastrophic event will be proportional to the market penetration of the insurer. The distribution assumption used for the number of claims is,

Assumption 5. $Y' \mid p \sim Bin \left( x, p \right)$.

The probability of a victim make a claim may vary depending on the scale of the catastrophic event. In small-scale events, casualties are mutually dependent and on large-scale events, casualties are mutually independent. Therefore, $P$ follows a certain distribution. Since the value of $P$ lies between 0 and 1, then a suitable distribution is a Beta distribution with parameters $\alpha$ and $\beta$. However, the parameters of the Beta distribution, $\alpha$ and $\beta$, are not sufficient to provide information about the mean and variation of the Beta distribution. Thus, re-parameterization of the parameters $\alpha, \beta$ is required [4].
Note that $E[P] = \frac{\alpha}{\alpha + \beta} = \mu$, so $E[Y'] = E[E[Y'|P]] = E[xP] = xE[P] = x\mu$. In accordance with Assumption 4, that $E[Y'] = xq$, it is necessary to re-parameterize the parameters of the Beta-distributed random variable to $q = \frac{\alpha}{\alpha + \beta}$, that describes the mean of the Beta distribution.

The variance of $P, Var(P) = \frac{\alpha \beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}$, can be written as,

$$Var(P) = \frac{\alpha}{1 + \frac{\beta}{\alpha}} \frac{1}{(\alpha + \beta + 1)(\alpha + \beta)} \left(1 - \frac{\alpha}{\alpha + \beta}\right)$$

$$= \frac{\mu(1-\mu)}{\alpha + \beta + 1}.$$

The variance of $P$ depends on the value of $\alpha + \beta$, so that $\alpha + \beta$ can be viewed as a variation of $P$. Therefore $\alpha + \beta$ will be used as the new parameter that describes the variation of the Beta distribution.

Based on [4] and [5], we make re-parameterization of Beta distribution parameters by using formula $q = \frac{\alpha}{\alpha + \beta}$ and $d(x) = \alpha + \beta$. Parameters $q$ and $d(x)$ will be more meaningful since $q$ is the mean of the Beta distribution, and $d(x)$ measures the variation of the Beta distribution. As previously known, $q$ is also a notation for market penetration of insurance companies. While $d(x)$ is a function that maps $x$ deaths to $d(x)$, where $0 \leq d(x) < \infty$. By solving the equation $q = \frac{\alpha}{\alpha + \beta}$ and $d(x) = \alpha + \beta$, the new parameter of the Beta distribution is $\alpha = d(x)q$ and $\beta = d(x)(1-q)$. The assumption used for the distribution of $P$ is

**Assumption 6.** $P \sim \text{Beta}(d(x)q, d(x)(1-q)), 0 < d(x) < \infty$

The concept of mixing is used to obtain the density function of $Y'$ that is as follows,

$$f_{Y'}(y') = \binom{x}{y'} \frac{B(y' + d(x)q, x - y' + d(x)(1-q))}{B(d(x)q, d(x)(1-q))}, y' = 0, 1, 2, ..., x.$$

$Y'$ is also called Beta-Binomial distribution with parameters $x$, $d(x)$ and $q$. The function $d(x)$ denoting the variation of deaths should be chosen so that when $d(x) \to \infty$, $x$ is large and when $d(x) \to 0$, $x$ is small, or in other words $d(x)$ is an increasing function. So, the assumption used for the selection of function $d(x)$ is as follows:

**Assumption 7.** $d(x) = \theta \log(x), \theta \in \mathbb{R}^+.$

In the catastrophic reinsurance contract, it is stated that at least $m$ insured person die on a catastrophic event in order for the reinsurer to help paying the losses received by the ceding company. In other words, the catastrophic reinsurance contract is active, if at least there are $m$ claim resulting from a catastrophic event. Suppose that a catastrophic event yields claims that comes from $x$ deaths, then

$$y = \begin{cases} y', & \text{if } y' \geq m \\ 0, & \text{if } y' < m \end{cases}$$

states the number of claims covered in the catastrophic reinsurance contract.
2.4. The amount of claims

Suppose the amount of the \(i\)-th individual claim of the dead insured is denoted by the random variable \(Z_i\). The total amount of claim from a catastrophic event coming from the \(y\) insured dead is \(Z' = \sum_{i=1}^{y} Z_i\). The standard value of \(E[Z_i] = 1\) is then used, which states the expected amount of each person’s claim is 1, where this value may vary depending on the actual situation.

The distribution of \(Z_i\) can be obtained from the insurance company’s portfolio. In cases where the catastrophe reinsurance contract protects the policy in group insurance, the sum insured of each insured may be the same, so \(Z_i = 1\) and \(Z'\) will be equal to the number of claims \(y\). But the catastrophe reinsurance contract can not only protect the policy in this group of insurance, but other types of insurance products can also be protected. In such a case, the assumption that is used to model the magnitude of a claim is that the amount of claim is exponentially distributed with mean 1, or \(Pr(Z \leq z' | y = 1) = 1 - e^{-z'} [4]\).

In the catastrophic reinsurance contract, the ceding company and reinsurer have determined the retention and aggregate limit of the contract. The retention of the catastrophic reinsurance contract is denoted by \(S\) and the aggregate limit is denoted by \(L\). From a catastrophic event, the amount of claim paid by the reinsurer for the occurrence of that catastrophe is

\[
Z = \begin{cases} 
0, & \text{if } z' < S \\
 z' - S, & \text{if } S \leq z' < S + L \\
 L, & \text{if } S + L \geq z' 
\end{cases}
\]

2.5. Pricing catastrophe reinsurance risk premium

Suppose the amount of claims due to the \(k\) catastrophic event is \(z_k\). Then the total magnitude of claims resulting from all catastrophic events occurring within 1 year of the catastrophic reinsurance contract is defined as follows,

\[
C = \sum_{k=1}^{k_m} z_k
\]

where \(k_m\) is the number of catastrophic events occurring within 1 year of catastrophic reinsurance contracts.

The value of \(C\) depends on the parameters of the catastrophic reinsurance coverage contract that is the minimum of deaths \(m\), the overall retention of \(S\) and the aggregate limit \(L\) and the parameters in the model are \(\lambda_m, \sigma, \xi, q, \theta\). After performing the simulation, the expected value, variance, and distribution of \(C\) can be obtained and can be used to calculate the risk premium of catastrophic reinsurance.

The risk premium calculation of catastrophe reinsurance can use standard deviation premium principle, as follows:

\[
P = E[C] + \alpha \times SD(C)
\]

where \(\alpha > 0\) is a safety loading.

3. Numerical example

The data is the number of deaths from disasters in Indonesia taken in the period 1917–2017. The data source is the website of the National Disaster Management Agency (BNPB). The types of disaster coverage include floods and landslides, earthquakes, earthquakes and tsunami, industrial accidents, transportation accidents, volcanic eruptions, landslides, and tsunami.

An insurance company \(A\) buys the reinsurance contract of catastrophe to a reinsurance company \(B\). The catastrophic reinsurance contract is purchased to protect the insurance company \(A\) from claims made
by the policyholders jointly resulting from a 1-year catastrophic event. The number of policies entered into the contract is as many as 190,000 policies, and it is known that the insurance company A market is 10,250,000 people. The average sum insured of each policy is Rp 75,000,000.

In this contract, the ceding company and reinsurer decided that the minimum claim requirement to activate the catastrophic reinsurance contract is 10. It is also determined that the retention and aggregate limit of the contract is $S = Rp\, 500,000,000$ and $L = Rp\, 30,000,000,000$. This means that if a catastrophe reinsurance contract is active and will result in a loss to the ceding company, the reinsurer will help cover the loss with a minimum coverage of 500 million, and the maximum by the reinsurer is 30 billion.

Based on information above, the actuary of the reinsurance company has data on the number of casualties due to a catastrophic event and additional information about the insurance company. But in the data from BNPB only consist the number of deaths. For calculating the amount of risk premium it requires the amount of claim from one catastrophic event. Hence it is necessary to do simulation to get the total amount of claims due to the entire catastrophic events in 1 year catastrophic reinsurance contract.

After doing 100,000 simulations, which from Law of Large Number we know it’s already sufficient, and then we get 100,000 values from $C$ that are $C_1, C_2, \ldots, C_{100,000}$. Then we get $E[C] = 320,501,500$ and $SD(C) = 2,501,729,122$ The amount of risk premium with some sample of the value of safety loading are:

a. $\alpha = 0.1, \quad P = 320,501,500 + 0.1 \times 2,501,729,122 = 570,674,412$.

b. $\alpha = 0.2, \quad P = 320,501,500 + 0.2 \times 2,501,729,122 = 820,847,324$.

c. $\alpha = 0.5, \quad P = 320,501,500 + 0.5 \times 2,501,729,122 = 1,571,366,061$.

d. $\alpha = 0.8, \quad P = 320,501,500 + 0.8 \times 2,501,729,122 = 2,321,884,798$.

There are several risk premium options from the catastrophic reinsurance contract that will be charged to insurance company A by the reinsurance company, but not all of these options can be taken. For the price of risk premium with $\alpha = 0.1$, the price is too low, and if the amount of the claims is very large, it would be very risky for the reinsurance company to incur losses. For the price of risk premium with $\alpha = 0.5$ and $\alpha = 0.8$, the price is too high, so it is likely that the insurance company A won’t continue to buy the contract. For that of the above options, the risk premium price when $\alpha = 0.2$ is Rp 820,847,324 can be an option, because the price offered is not too high, and also large enough to cope with the risk of catastrophic events.

4. Conclusion

There are many steps for modelling risk premium calculation on catastrophic reinsurance contract using POT, those are modelling the number of catastrophic events with Poisson distribution, modelling the number of deaths resulting from catastrophic event with discrete Generalized Pareto distribution, modelling the number of claims resulting from the one catastrophic event with Beta-Binomial distribution and modelling the total amount of a claim resulting from a catastrophic event with Gamma distribution. With these steps of modelling, the model depends on the data to produce its parameter, so the data has to be updated to get an update result. Furthermore, this model allows us to use data without any threshold, unlike the previous model gives us data limit of 1500 casualties.

Risk premium of catastrophic reinsurance is calculating by generated total claim data as a result from all catastrophic events within 1 year of catastrophic reinsurance contract with simulation. Based on numerical example, the amount of risk premium of catastrophic reinsurance contracts sold by reinsurance company B to insurance company A in 2017 is Rp 667,095,448. Risk premiums charged to each insurance company may vary depending on the type of disaster that are covered, and the ability of the insurance company.
References
[1] Ajay S and Jinjing W 2018 Risk *Insurance: Mathematics and Economics* **82** 55-72
[2] Erland E and Ola H 2014 *Scandinavian Actuarial Journal* **4** 352-67
[3] Manfred G and Evis K 2006 *Computational Economics* **27** 207-28
[4] Daniel M E and Jian Bi 1997 *Journal of Sensory Studies* **13** 389-412
[5] Edson Z M, Jorge A A and Davi C A 2015 *Ciência e Natura* **37** 12-9