Thailand Statistician
January 2021; 19(1): 1-13
http://statassoc.or.th
Contributed paper

The Properties of Inverse Pareto Distribution and its Application to Extreme Events

Sukanda Dankunprasert [a]*, Uraiwan Jaroengeratikun [b] and Tosaporn Talangtam [a]
[a] Faculty of Science, Khon Kaen University, Khon Kaen, Thailand.
[b] Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.
*Corresponding author; e-mail: Sukanda_d@kkumail.com

Received: 7 May 2019
Revised: 15 August 2019
Accepted: 1 November 2019

Abstract

In this paper we present some properties of the inverse Pareto distribution (IPD). We compare IPD to gamma distribution, exponential distribution and generalized Pareto distribution (GPD). The data sets over threshold u are analyzed and obtained by the Monte Carlo simulation and the use of Danish fire data. The maximum likelihood estimation (MLE) is the parameter estimation. The various measurements of model fitting are the Kolmogorov-Smirnov test (KS test), the Anderson-Darling test (AD test), Akaike information criterion (AIC) and Bayesian information criterion (BIC). We found that the IPD is a good competitor with GPD for the modeling of extreme events.

Keywords: Danish fire data, generalized Pareto distribution, heavy tailed distribution, infinite mixture distribution, peak over threshold.

1. Introduction

Losses of non-life insurance are referred to as claims or liabilities of insurance company which are responsible for the insured or policyholders. Claim modeling is an important task for actuaries who need to be able to estimate or forecast the behavior of claims which will occur in the future. The building of claim models has been developed by many authors relative to their different types of data. New models or distributions are built from mixture models for fitting to the data which cannot be fitted by a commonly used distribution. Some loss distribution and their modeling are described by Hogg et al. (2005), Catherine et al. (2011) and Klugman et al. (2012). The mixture models are known in terms of urn-to-urn models in risk theory that is combined with continuous and discrete distributions, while in terms of loss models, they are finite mixture distributions and infinite mixture distributions. Many authors presented the modeling of finite mixture models. For example, Mohamed et al. (2010) proposed a compound Pareto distribution for claim severity modeling. Sattayatham and Talangtam (2012) presented finite mixture lognormal distributions and applied the models to motor insurance claims data. Mauro et al. (2012) proposed finite mixture skew normal distributions and applied them to the insurance claim data set of Danish fire losses. Moreover, Erisoglu et al. (2013) used two mixture gamma distributions for the estimation of heterogeneous wind data sets. Matthayomnan (2016) constructed new models for claim amount by finite mixture lognormal and Frechet distributions. There are a few loss models for the infinite mixture distribution for claim severity in textbooks and papers. Therefore, we are interested in the modeling of infinite mixture distributions for claim severity because most papers of infinite mixture distributions are published regarding claim frequency. The popular posterior distribution of Poisson is presented for modeling by Rai (1971), Bulmer (1974), Sichel (1975), Irwin (1975), Kempton (1975), Albrecht (1982) and others. Emilio et al. (2006) proposed a negative binomial inverse gaussian distribution (NBIG) which has been applied for the pricing of automobile premiums. Hossein Zamani and Noriszura Ismail (2010) introduced the negative binomial-Lindley distribution which provides a better fit than the Poisson and the negative binomial for two claim frequency samples of insurance data. Some papers describe and present infinite mixture models for both claim frequency and claim severity. For example, Pacáková and Zapletal (2013) explain a model of claim frequency and claim severity by using Poisson-gamma distribution and exponential-gamma (Pareto) distribution, respectively. Anantasopon, Sattayatham and Talangtam (2015) propose an infinite mixture of inverse exponential-gamma for fitting to motor insurance claims data. Dankunprasert (2017) presented gamma-exponential distribution (GED) and exponential-exponential distribution (EED) constructed from an infinite mixture model for severity claims of motor insurance.

Reinsurance in excess of loss treaty is considered to be the threshold u, whereby the reinsurer takes the risks or losses over u. This means that the reinsurer's liability is based on the remaining data from the right censored data of the ceding company. Therefore, we focus on the modeling depending on u of claims by the reinsurer that the extreme value theory (EVT) is appropriately applied and compared for this work. EVT is divided into two types depending on the characterization of censored data. These types are block maximum and peak over threshold (POT) that lead to the use of generalized extreme value distribution (GEVD) and generalized Pareto distribution (GPD), respectively. GPD is frequently used in many fields and it is a good tool for modeling data by the POT method. The EVT is described in a book of Fisher and Tippett (1928), Coles (2001), Kluppelberg (2004) and Beirlant et al. (2004). The paper of Berning (2010) explained how EVT is able to be applied to various fields, such as those of financial and insurance risks. Vladimir et al. (2012) presented the extreme events which were relative to the global crisis of 2008-2009 and they employed the POT and GPD for the modeling of financial returns based on the daily losses of the Russian stocks index (RTSI) in 1995-2009. Yang (2013) presented GPD for insurance modeling by using a software package R for simulated data and then applied these to fire losses in Denmark. Recently, Puangkaew (2018) applied the GPD to Danish fire claims severity and compared the estimated parameters using various methods.

The organization of this paper is composed by 7 sections as follows. In Section 2, the derivative of IPD and GED are explained. In Section 3, the tail properties of IPD is presented by using a hazard function. In Section 4, the materials and methods for model fitting are employed to some loss distributions, IPD and GPD. In Section 5, the results of models are compared based on both simulated data and actual claim data. The conclusions and discussions are specified in Sections 6 and 7, respectively.

2. Models

The IPD and GED are formed by the same distribution, but they are derived from different constructions. The IPD comes from the inverse transformation of the Pareto distribution, type II Lomax distribution whereas the GED comes from the infinite mixture distribution of Gamma and Exponential distributions. The construction of IPD or GED for inverse transformation and infinite mixture distribution are explained in Sections 2.1 and 2.2, respectively. Wordpress.com (2017) presented the relative between distribution as follows.

2.1. Inverse transformation

The type II Lomax distribution with shape parameter α and scale parameter θ for cumulative distribution function (CDF) and probability density function (PDF) of random variable X are as follows.

Cumulative distribution function is

$$F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha}, \alpha > 0, \theta > 0, x > 0.$$

Probability density function is

$$f(x) = \frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}}.$$

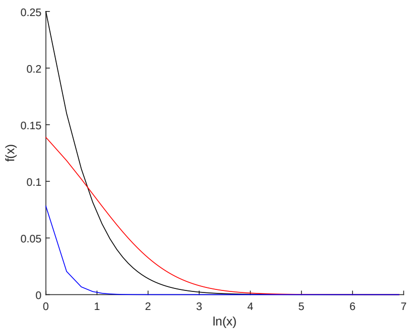

Figure 1 PDF of Pareto

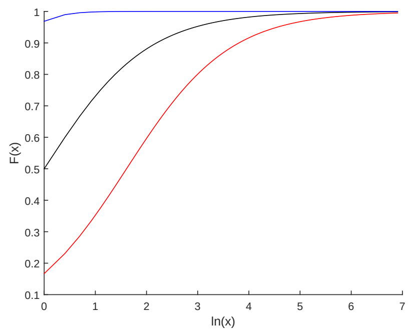

Figure 2 CDF of Pareto

— $\alpha = 1$ and $\theta = 1$ — $\alpha = 5$ and $\theta = 1$ — $\alpha = 5$ and $\theta = 5$

We take the approach of raising a base Pareto distribution with shape parameter α and scale parameter θ^{-1}. Both approaches lead to the same CDF:

$$F(y) = 1 - F_X(y^{-1}) = \left(\frac{\theta}{\frac{1}{y}+\theta}\right)^{\alpha} = \left(\frac{y}{y+\lambda}\right)^{\alpha}, \text{where } \lambda = \frac{1}{\theta}.$$

The PDF is given by

$$f(y) = \frac{\lambda\alpha y^{\alpha-1}}{(y+\lambda)^{\alpha+1}}; \alpha > 0, \lambda > 0, y > 0.$$

2.2. Infinite mixture distribution

Dankunprasert (2017) presented the IPD by using infinite mixture model. The PDF and CDF are described as below.

Suppose a random variable X follows a gamma distribution with parameters c and α. Denote its PDF by $f(x; c, \alpha)$ as

$$f(x; c, \alpha) = \frac{\alpha(x\alpha)^{c-1}}{\Gamma(c)} \exp(-\alpha x); c > 0, \alpha > 0, x > 0,$$

where $\Gamma(c)$ is the gamma function, $\Gamma(c) = \int_0^{\infty} x^{c-1} \exp(-x)\, dx$ for $c > 0$.

The exponential distribution of random variable α will be used as the mixing distribution. The PDF of α is

$$g(\alpha) = \lambda \exp(-\alpha\lambda); \lambda > 0, \alpha > 0.$$

Then the unconditional PDF of X is

$$\begin{aligned} f(x) &= \int_0^\infty \frac{\alpha(\alpha x)^{c-1}\exp(-\alpha x)}{\Gamma(c)}\lambda\exp(-\alpha\lambda)d\alpha \\ &= \frac{\lambda c x^{c-1}}{(x+\lambda)^{c+1}}. \end{aligned}$$

Hence,

$$f(x) = \frac{\lambda c x^{c-1}}{(x+\lambda)^{c+1}}; c > 0, \lambda > 0, x > 0. \tag{1}$$

The CDF is given by

$$F(x) = \left(\frac{x}{x+\lambda}\right)^c.$$

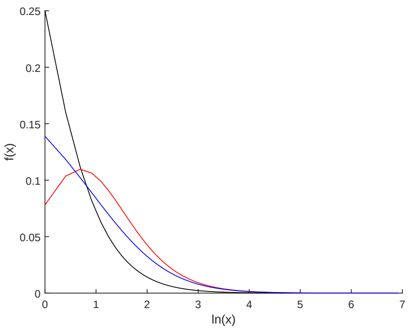

Figure 3 PDF of Inverse Pareto

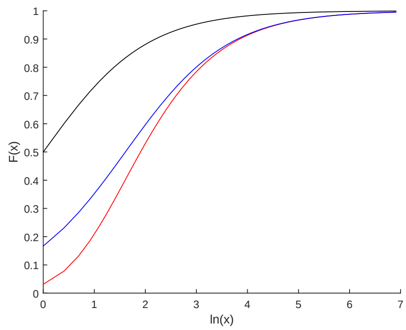

Figure 4 CDF of Inverse Pareto

—— $\lambda = 1$ and $c = 1$ —— $\lambda = 5$ and $c = 1$ —— $\lambda = 5$ and $c = 5$

3. Tail Behavior

This section presents the tail properties of the IPD in Theorems 1 and 2. First of all, we will show some properties of IPD as survival, hazard functions and Value-at-risk.

Survival function is given by

$$S(x) = 1 - F(x) = 1 - \left(\frac{x}{(x+\lambda)}\right)^c.$$

Hazard function is

$$h(x) = \frac{f(x)}{S(x)} = \frac{\frac{\lambda c x^{c-1}}{(x+\lambda)^{c+1}}}{(\frac{x}{x+\lambda})^c} = \frac{\lambda c x^{c-1}}{(x+\lambda)[(x+\lambda)^c - x^c]}. \tag{2}$$

The Value-at-risk (VaR) (Klugman 2012) of a loss random variable X at the $100p\%$ level, denoted by $VaR_p(X)$ or π_p is the $100p$ percentile (or quantile) of the distribution of X. The VaR of IPD is in the form of

$$\pi_p = \frac{\lambda}{p^{-1/c} - 1}.$$

Theorem 1 *Let X be a random variable. The probability distribution function is an inverse Pareto distribution (IPD) in the form of*

$$f(x) = \frac{\lambda c x^{c-1}}{(x+\lambda)^{c+1}}; c > 0, \lambda > 0, x > 0.$$

Then IPD has a heavy tail.

Proof: The hazard function for the IPD is Equation (2).
Thus we obtain

$$\begin{aligned} h^{'}(x) &= \frac{\lambda c x^{c-2}(A-B)}{\left[(x+\lambda)[(x+\lambda)^c - x^c]\right]^2} < 0, \\ \text{where} \quad A &= (c-1)(x+\lambda)[(x+\lambda)^c - x^c], \\ B &= x[(c+1)(x+\lambda)^c - x^{c-1}(cx + c\lambda + x)]. \end{aligned}$$

Since IPD has a decreasing hazard function, the IPD has a heavy tail.

Theorem 2 *Let X be a random variable. A heavy tailed distribution has a tail that's heavier than an Exponential distribution Then IPD has a heavy tail.*

Proof: Consider the survival function,

$$\lim_{x\to\infty} \frac{S_{IPD}(x)}{S_{Exp}(x)}.$$

Then, the limit of the ration will be the same, as can be seen by an application of L'Hpital's rule:

$$\begin{aligned} \lim_{x\to\infty} \frac{S_{IPD}(x)}{S_{Exp}(x)} &= \lim_{x\to\infty} \frac{S'_{IPD}(x)}{S'_{Exp}(x)} = \lim_{x\to\infty} \frac{-f_{IPD}(x)}{-f_{Exp}(x)} = \lim_{x\to\infty} \frac{f_{IPD}(x)}{f_{Exp}(x)} \\ &= \lim_{x\to\infty} \frac{\frac{\lambda c x^{c-1}}{(x+\lambda)^{c+1}}}{a\exp(-ax)} = \lim_{x\to\infty} \frac{\lambda c x^{c-1}\exp(ax)}{a(x+\lambda)^{c+1}}. \end{aligned}$$

Since exponential goes to infinity faster than polynomials, the limit is infinity. So, the IPD has a heavier tail than exponential. Therefore, the IPD is a heavy tailed distribution.

4. Materials and Methods

Embrechts et al. (2011) presented the distribution of EVT which is used for modeling exceedance over a threshold. Modeling with a GPD model and statistical analysis of real data are as follow: Let basic losses data $X_1, X_2, \ldots, X_n$ be random variables with independent identically distributed (iid) functions F. The ordered data are denoted by $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$. The right endpoint, x_F, is defined as $x_F = \sup\{x; F(x) < 1\}$. For all $u < x_F$, the function

$$F_u(x) = P\{X - u \leq x \mid X > u\}, x \geq u$$

is called the distribution function of exceedances above threshold u.

Let $Y = X - u$ for $X > u$ and for n observed variables $X_1, X_2, \ldots, X_n$ we can write $Y_j = X_i - u$ such that i is the index of the j^{th} exceedance, $j = 1, 2, \ldots, n_u$. The distribution of the exceedances $(Y_1, \ldots, Y_{n_u})$ can be fitted by the models.

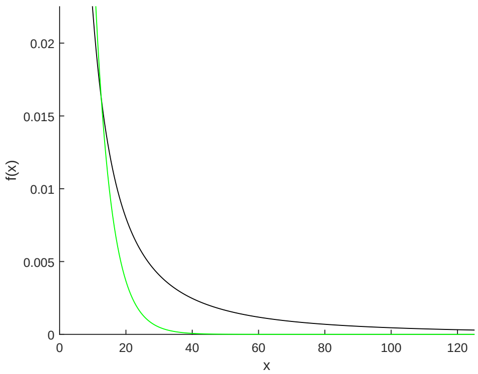

Figure 5 The curve of the PDF

— is PDF of IPD with $\lambda = 5$ and $c = 1$ — is PDF of Exponential with $rate = 0.2$.

4.1. The models

The loss distributions and extreme event models are selected for our models that are composed of IPD , gamma, exponential and GPD. We make a comparison of IPD modeling among the selected distributions. The GPD is a popular model from EVT which is

$$G_{\xi,\beta}(y) = \begin{cases} 1-(1+\xi\frac{y}{\beta})^{-\frac{1}{\xi}}, & \text{if } \xi \neq 0 \\ 1-\exp(-\frac{y}{\beta}), & \text{if } \xi = 0, \end{cases}$$

where $\xi \in \mathbb{R}$ and $\beta > 0$. We require $0 \leq y < \infty$ if $\xi \geq 0$, and $0 \leq y \leq -\frac{\beta}{\xi}$ if $\xi < 0$. The parameter $\beta = \beta(u)$ is the scale parameter and its depends on the threshold u.

The PDF of GPD is given by

$$g_{\xi,\beta}(y) = \begin{cases} \frac{1}{\beta}(1+\xi\frac{y}{\beta})^{-\frac{1}{\xi}-1}, & \text{if } \xi \neq 0 \\ \frac{1}{\beta}\exp(-\frac{y}{\beta}), & \text{if } \xi = 0. \end{cases}$$

4.2. Parameters estimation

The method of maximum likelihood estimation (MLE) provides estimators which are usually quite satisfactory and frequently used in actuarial mathematics, see in Klugman (2012).

Consider the amount x_i paid for the i^{th} contract, where $i = 1, 2, ..., n$. Dankunprasert (2017) fit the IPD in Equation (1) to the data set by MLE. The likelihood function can be written as follows.

$$L = \prod_{i=1}^{n} \frac{\lambda c x_i^{c-1}}{(x_i+\lambda)^{c+1}},$$

and the log-likelihood function is in the form.

$$\ln L = \sum_{i=1}^{n} \ln\left\{\frac{\lambda c x_i^{c-1}}{(x_i+\lambda)^{c+1}}\right\}.$$

Taking the partial derivatives of the log-likelihood function with respect to parameters are as follows:

$$\frac{\partial}{\partial c}\ln L = \frac{n}{c} + \sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\ln(x_i+\lambda)$$
$$\frac{\partial}{\partial \lambda}\ln L = \frac{n}{\lambda} - (c+1)\sum_{i=1}^{n}\frac{1}{x_i+\lambda}.$$

We estimate $\widehat{c}$ and $\widehat{\lambda}$ for c and λ, respectively by $\frac{\partial}{\partial c}\ln L = 0$ and $\frac{\partial}{\partial \lambda}\ln L = 0$.
Thus,

$$\frac{n}{c} + \sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\ln(x_i+\lambda) = 0,$$
$$\frac{n}{\lambda} - (c+1)\sum_{i=1}^{n}\frac{1}{x_i+\lambda} = 0.$$

Solve the equations numerically for estimated parameters by fixed point iteration method.

4.3. Goodness of fit test and model selection

There are two measurements for model fitting and two paradox in model selections are as follows:

1) The Kolmogorov-Smirnov test

The K-S test statistic is defined by

$$D = \sup_x |F_n(x) - F_X^*(x)|,$$

where $F_n(x)$ is empirical cumulative distribution function of X with n data and $F_X^*(x)$ is the theoretical cumulative distribution of the distribution being tested.

2) The Anderson-Darling Test

The AD test statistic is defined by

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[\ln F(x_i) + \ln\{1 - F(x_{n-i+1})\}],$$

where F is the theoretical cumulative distribution of the distribution being tested.

3) The Akaike information criterion (AIC)

$$AIC = 2k - 2\ln(L(\theta)),$$

where k is the number of parameters estimated and $\ln(L(\theta))$ is the log-likelihood function.

4) The Bayesian information criterion (BIC)

$$BIC = -2\ln(L(\theta)) + k\ln(n),$$

where $\ln(L(\theta))$ is the log-likelihood function and n is the number of observations.

4.4. Data

We apply the models to the simulated data and the Danish fire data sets for model fitting. Some descriptions are as follows:

1) The simulation data

The Monte Carlo simulation of data are generated by three different distributions such as the distributions of Log-logistic, BurrXII and Log-normal with sample size $n = 200$, and 3000 replicates.

2) The actual data

The Danish fire data consist of 2167 losses which were over one million Danish Krone (DKK) for the years $1980 - 1990$. The loss includes damage of buildings, personal property and loss of profits.

5. Results

The thresholds u are varied for our model fitting in tail distributions. We found that the IPD is mostly suitable for the data sets on u of 84^{th} to 97.5^{th} percentile for some simulated data and u of 6, 12 and 15 million Krone for the Danish fire data. The results are presented as the following items.

5.1. Simulation data

Tables 1 and 2 show the estimated parameters and values of model fitting with censored data at 84^{th} and 94^{th} percentile, respectively, for the generated data by log-logistic distribution. The number of exceedance are 480 for 84^{th} percentile (around 4 million Krone) and 180 for 94^{th} percentile (around 9 million Krone).

At the significant level $\alpha = 0.05$, IPD and GPD are fitted to the data whereas the gamma and exponential distributions cannot be fitted to any data sets. By AIC and BIC, the IPD is the most suitable for the data following by GPD, gamma and exponential distributions, respectively.

Table 1 Threshold u at 84^{th} percentile (≈ 4 million Krone)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 2.6951$ $c = 1.1286$	$D = 0.0292$ p-value $= 0.8093$	$AD = 0.3795$ p-value $= 0.8687$	3000.942	3009.290
Gamma	Rate $= 5.5164 \times 10^{-3}$ Shape $= 0.2624$	D $= 0.2893$ p-value $< 2.2 \times 10^{-16}$	AD $> 10^6$ p-value $< 10^{-5}$	3585.763	3594.110
Exponential	Rate $= 0.0210$	D $= 0.6011$ p-value $= 1.06 \times 10^{-9}$	AD $> 10^6$ p-value $< 10^{-5}$	4669.567	4673.741
GPD	$\beta = 3.3380$ $\xi = 0.9174$	D $= 0.0338$ p-value $= 0.642$	AD $= 0.5064$ p-value $= 0.7402$	3001.911	3010.258

Table 2 Threshold u at 94^{th} percentile (≈ 9 million Krone)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 9.7906$ $c = 0.8375$	$D = 0.0494$ p-value $= 0.7712$	$AD = 0.3444$ p-value $= 0.9014$	1459.492	1465.877
Gamma	Rate $= 2.1380 \times 10^{-3}$ Shape $= 0.2541$	D $= 0.2493$ p-value $< 3.8110 \times 10^{-10}$	AD $> 10^6$ p-value $< 10^{-15}$	1649.987	1656.373
Exponential	Rate $= 8.4189 \times 10^{-3}$	D $= 0.5727$ p-value $< 2.2 \times 10^{-16}$	AD $> 10^6$ p-value $< 10^{-15}$	2081.821	2085.014
GPD	$\beta = 7.2519$ $\xi = 1.0637$	D $= 0.0446$ p-value $= 0.8667$	AD $= 0.3774$ p-value $= 0.8706$	1460.173	1466.559

Tables 3 and 4 show the estimated parameters and values of model fitting with censored data at 92^{th} and 96^{th} percentile, respectively, for the generated data by BurrXII distribution. The number of exceedance are 240 for 92^{th} percentile (around 6 million Krone) and 120 for 96^{th} percentile (around 10 million Krone).

At the significant level $\alpha = 0.05$, IPD and GPD are fitted to the data whereas the gamma and exponential distributions cannot be fitted to any data sets. By AIC and BIC, the GPD is the most suitable for the data following by IPD, gamma and exponential distributions, respectively.

Table 3 Threshold u at 92^{th} percentile ($\approx$ 6 million Krone)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 2.4943$ $c = 1.3706$	$D = 0.0489$ p-value $= 0.6153$	$AD = 0.8364$ p-value $= 0.4554$	1522.827	1529.788
Gamma	Rate $= 0.0593$ Shape $= 0.6151$	D $= 0.1484$ p-value $= 5.13 \times 10^{-5}$	AD $= 6.7125$ p-value $= 4.5350 \times 10^{-4}$	1560.770	1567.731
Exponential	Rate $= 0.0964$	D $= 0.2292$ p-value $= 2.23 \times 10^{-11}$	AD $= 23.37$ p-value $= 2.5 \times 10^{-6}$	1604.774	1608.254
GPD	$\beta = 4.5877$ $\xi = 0.6346$	D $= 0.0547$ p-value $= 0.4693$	AD $= 0.9701$ p-value $= 0.3733$	1519.946	1526.907

Table 4 Threshold u at 96^{th} percentile ($\approx$ 10 million Krone)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 5.0759$ $c = 1.0833$	$D = 0.0620$ p-value $= 0.7460$	$AD = 0.6793$ p-value $= 0.5761$	878.4522	884.0271
Gamma	Rate $= 0.0406$ Shape $= 0.6266$	D $= 0.1048$ p-value $= 0.1432$	AD $= 1.8404$ p-value $= 0.1127$	880.0029	885.5779
Exponential	Rate $= 0.0647$	D $= 0.2129$ p-value $= 3.7860 \times 10^{-5}$	AD $= 10.1260$ p-value $= 1.0360 \times 10^{-5}$	899.0689	901.8564
GPD	$\beta = 7.2555$ $\xi = 0.6521$	D $= 0.079245$ p-value $= 0.4383$	AD $= 0.9436$ p-value $= 0.3881$	876.0974	881.6723

Tables 5 and 6 show the estimated parameters and values of model fitting with censored data at 95^{th} and 97.5^{th} percentile, respectively, for the generated data by log-normal distribution. The number of exceedance are 120 for 96^{th} percentile (around 10 million Krone) and 75 for 97.5^{th} percentile (around 13 million Krone).

At the significant level $\alpha = 0.05$, for 95^{th} percentile, IPD and GPD are fitted to the data whereas the gamma and exponential distributions cannot be fitted to any data sets. For 97.5^{th} percentile, the data sets are fitted by all models. By AIC and BIC, for both 95^{th} and 97.5^{th} percentile, the GPD is the most suitable for the data following by IPD, gamma and exponential distributions, respectively.

Table 5 Threshold u at 96^{th} percentile (≈ 10 million Krone)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 4.3443$ $c = 0.8985$	$D = 0.0703$ p-value $= 0.4482$	$AD = 1.1134$ p-value $= 0.3028$	939.9067	945.9280
Gamma	Rate $= 0.0448$ Shape $= 0.6070$	D $= 0.1249$ p-value $= 0.0186$	AD $= 2.8339$ p-value $= 0.0333$	1055.4330	1061.4550
Exponential	Rate $= 0.0738$	D $= 0.2357$ p-value $= 1.1540 \times 10^{-7}$	AD $= 14.2080$ p-value $= 4 \times 10^{-6}$	1084.0190	1087.0290
GPD	$\beta = 5.1616$ $\xi = 0.4463$	D $= 0.0699$ p-value $= 0.4558$	AD $= 1.0221$ p-value $= 0.3457$	930.3706	936.3920

Table 6 Threshold u at 97.5^{th} percentile (≈ 13 million Krone)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 4.9149$ $c = 1.3063$	$D = 0.1137$ p-value $= 0.2658$	$AD = 1.8030$ p-value $= 0.1183$	540.0031	544.6381
Gamma	Rate $= 0.0365$ Shape $= 0.7632$	D $= 0.0595$ p-value $= 0.9390$	AD $= 0.3558$ p-value $= 0.8908$	606.0451	610.6801
Exponential	Rate $= 0.0478$	D $= 0.0844$ p-value $= 0.6287$	AD $= 0.4622$ p-value $= 0.7851$	608.0573	610.3748
GPD	$\beta = 10.4731$ $\xi = 0.1100$	D $= 0.0674$ p-value $= 0.8623$	AD $= 0.3266$ p-value $= 0.9166$	524.3504	528.9854

5.2. Actual data

Tables 7 to 9 show the estimated parameters and values of model fitting with censored data at 6, 12 and 15 million Krone, respectively, for the Danish fire data in Danish Krone (DKK). The number of exceedance are 273 for threshold u at 6 million Krone, 87 for threshold u at 12 million Krone and 65 for threshold u at 15 million Krone.

At the significant level $\alpha = 0.05$, the Danish fire data is fitted by all models, except for the exponential distribution. By AIC and BIC with u at 6 and 12 million Krone, GPD is the most suitable for the data following by IPD , gamma and exponential distributions, respectively. For the threshold of 15 million Krone the IPD is the most suitable for the data following by GPD, gamma and exponential distributions, respectively.

Table 7 Threshold $u = 6$ million Krone ($\approx 87.5^{\text{th}}$ percentile)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 3.8477$ $c = 1.1118$	$D = 0.0771$ p-value $= 0.2183$	$AD = 1.5299$ p-value $= 0.1695$	1221.153	1227.605
Gamma	Rate $= 0.0548$ Shape $= 0.6138$	D $= 0.0894$ p-value $= 0.1025$	AD $= 2.3554$ p-value $= 0.05912$	1238.782	1245.234
Exponential	Rate $= 0.0893$	D $= 0.1447$ p-value $= 8.275 \times 10^{-4}$	AD $= 10.8430$ p-value $= 4.28 \times 10^{-6}$	1272.741	1275.966
GPD	$\beta = 5.8444$ $\xi = 0.4704$	D $= 0.0356$ p-value $= 0.9725$	AD $= 0.3285$ p-value $= 0.9152$	1207.654	1214.106

Table 8 Threshold $u = 12$ million Krone ($\approx 95^{\text{th}}$ percentile)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 4.6439$ $c = 1.2121$	$D = 0.0784$ p-value $= 0.6737$	$AD = 0.8680$ p-value $= 0.4834$	611.1520	616.0370
Gamma	Rate $= 0.0382$ Shape $= 0.6033$	D $= 0.1354$ p-value $= 0.0887$	AD $= 2.0649$ p-value $= 0.0848$	625.5176	630.4029
Exponential	Rate $= 0.0633$	D $= 0.2023$ p-value $= 1.9010 \times 10^{-3}$	AD $= 6.7169$ p-value $= 4.5750 \times 10^{-4}$	641.3073	643.7500
GPD	$\beta = 7.5512$ $\xi = 0.5213$	D $= 0.0553$ p-value $= 0.9572$	AD $= 0.2476$ p-value $= 0.9716$	606.2441	611.1294

Table 9 Threshold $u = 15$ million Krone ($\approx 97^{\text{th}}$ percentile)

Distributions	Estimated Parameters	Measurements of model fitting			
		K-S test	AD test	AIC	BIC
IPD	$\lambda = 2.1163$ $c = 2.5724$	$D = 0.0647$ p-value $= 0.9489$	$AD = 0.2651$ p-value $= 0.9617$	446.1715	450.3602
Gamma	Rate $= 0.0345$ Shape $= 0.6496$	D $= 0.1709$ p-value $= 0.0530$	AD $= 2.4551$ p-value $= 0.0525$	467.4470	471.6357
Exponential	Rate $= 0.0531$	D $= 0.2246$ p-value $= 3.8600 \times 10^{-3}$	AD $= 5.6394$ p-value $= 1.4460 \times 10^{-3}$	474.2738	476.3681
GPD	$\beta = 8.7134$ $\xi = 0.5430$	D $= 0.0762$ p-value $= 0.8508$	AD $= 0.4967$ p-value $= 0.7495$	448.9685	453.1571

6. Conclusions

IPD and GPD are mostly suitable for the simulation data and actual data with percentile of 84^{th} to 97.5^{th} following by gamma and exponential distributions.

By K-S test and AD test, the IPD is the best fit to censored data with threshold u at 84^{th} percentile that are simulated from log-logistic and BurrXII for 92^{th} to 96^{th}. Whereas the IPD is suitable for threshold u at 84^{th} to 94^{th} percentile that are simulated from log-logistic by AIC and BIC. In actual data, the IPD is the best fit to censored data with threshold u of 15 million Krone for all tests.

7. Discussion

The IPD can be fitted to some tails of distribution depending on threshold u. Although it cannot be fitted to all censored data, it provides a better fit than the traditional distributions such as the gamma and the exponential distributions. The IPD can be a good model for extreme events modeling in the same way that GPD is. In future research, we should consider a model with different parameters estimation for analysis estimated parameters according to point estimation and interval estimation. The aggregate claim and collective claim models are interested for insurance pricing that we take them into the account.

Acknowledgements

We would like to thank the referees for their comments and suggestions on the manuscript. This research is supported by a Science Achievement Scholarship of Thailand.

References

Albrecht P. On some statistical methods connected with the mixed Poisson process. Scand Actuar J, 1982; 9(1): 1-14.

Anantasopon S, Sattayatham P, Talangtam T. The modelling of motor insurance claims with infinite mixture distribution. Int J Appl Math Stat, 2015; 53(4): 40-49.

Beirlant J, Goegebeur Y, Teugels J, Segers J. Statistics of extremes: theory and applications. Hoboken: John Wiley & Sons Inc., 2004.

Bernardi M, Maruotti A, Lea P. Skew mixture models for loss distributions: a Bayesian approach. Munich Personal RePEc Archive; 2012 [cited 2017 Nov 6]. Available from: http://mpra.ub.uni-muenchen.de/40883/.

Berning L. Improved estimation procedures for a positive extreme value index. PhD [dissertation]. Matieland: Stellenbosch University; 2010.

Bulmer MG. On Fitting The Poisson lognormal distribution to species abundance data. Biometrics, 1974; 30(1): 101-110.

Catherine F, Merran E, Nicholas H, Brian P. Statistical distribution. Hoboken: John Wiley & Sons Inc., 2011.

Coles S. An Introduction to statistical modeling of extreme values. Springer Series in Statistics. London: Springer Verlag, 2001.

Dankunprasert S. The modeling of claim severity with infinite mixture models for individual data of motor insurance. MSc. [thesis]. Khon Kaen: Khon Kaen University; 2017.

Embrechts P, Klüppelberg C, Mikosch T. Modelling extremal events for insurance and finance. New York: Springer, 2011.

Erisoglu M, Servi T, Erisoglu U, Calis N. Mixture gamma distribution for estimation of wind power potential. Int J Appl Math Stat, 2013; 40: 223-231.

Fisher RA, Tippett LHC. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In Mathematical Proceedings of the Cambridge Philosophical Society. Cambridge University Press, 1928; 24, pp. 180-190.

Hogg RV, Craig A, McKean JW. Introduction to mathematical statistical. New Jersey: Prentice Hall, 2005.

Hossein Z, Noriszura I. Negative binomial-Lindley distribution and its application. J Math Stat, 2010; 6 (1): 4-9.

Irwin J. The generalized waring distribution Parts I, II, III. J R Stat Soc Ser A Stat Soc, 1975; 18-31(Part I), 204-227(Part II), 374-384(Part III).

Kempton RA. A generalized form of Fisher's logarithmic series. Biometrics, 1975; 62(1): 29-38.

Klugman SA, Panjer HH, Willmot GE. Loss models: from data to decisions. New York: John Wiley & Sons Inc., 2012.

Kluppelberg C. Risk management with extreme value theory. In: Finkenstaedt B. and Rootzen H. Extreme values in finance, telecommunication and the environment, 2004; pp. 101-168.

Matthayomnan W. Finite mixture models of non-life insurance and application to motor insurance claims. MSc. [thesis]. Khon Kaen: Khon Kaen University; 2016.

Mohamed MA, Ahmad MR, Noriszura I. Approximation of aggregate losses using simulation. J Math Stat, 2010; 6(3): 233-239.

Pacáková V, Zapletal D. Mixture distributions in mModelling of insurance losses. AMCME2013: Proceedings of the 2013 International Conference on Applied Mathematics and Computational Methods in Engineering; 2013. pp. 122-125.

Puangkaew S. Modelling extreme events for insurance claims using peak over threshold Model. MSc. [thesis]. Khon Kaen: Khon Kaen University; 2018.

Rai G. A Mathematical model for accident proneness. Trabajos Estadistica, 1971; 22: 207-212.

Sattayatham P, Talangtam T. Fitting of finite mixture distributions to motor insurance claims. J Math Stat, 2012; 8(1): 49-56.

Sichel HS. On a distribution law for word frequencies. J Am Stat Assoc, 1975; 70(351): 542-547.

Vladimir OA, Sergey ET, Oksana PO, Gennady PP. Extreme value theory and peak over threshold model in the Russian stock market, J Siberian Fed Uni Eng & Tech . 2012; 1(5): 111-121.

Wordpress [Internet]. n.p.: Topics in Actuarial Modeling, A catalog of parametric severity models; 2017 [cited 2018 May 6]. Available from: https:actuarialmodelingtopics.wordpress.com /tag/inverse-pareto-distribution/.

Yang H. Extreme value mixture modelling with simulation study and applications in finance and insurance. MSc. [thesis]. New Zealand: Canterbury University; 2013.