Thailand Statistician
January 2023; 21(1): 110-124
http://statassoc.or.th
Contributed paper

The Poisson Inverse Pareto Distribution and Its Application

Sirithip Wasinrat [a] and Boonyarit Choopradit* [b]
[a] Faculty of Science, Chandrakasem Rajabhat University, Bangkok, Thailand.
[b] Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathumthani, Thailand.
*Corresponding author; e-mail: boonyarit@mathstat.sci.tu.ac.th

Received: 7 November 2021
Revised: 8 July 2022
Accepted: 27 July 2022

Abstract

This paper proposes a new combination of the inverse Pareto distribution as a three-parameter distribution—the Poisson inverse Pareto distribution—derived from the concept of a unified model for long-term survival analyses. This work derives the proposed distribution's probability properties, including the survival, cumulative distribution, probability density, and hazard functions. Moreover, some properties of the Poisson inverse Pareto distribution are presented, such as the value-at-risk, tail behavior, quantile function, and order statistics. The maximum likelihood estimation is then studied to obtain a parameter estimation. This study's simulation revealed that the estimated parameter's mean-squared errors decreased when the sample size increased. Finally, this work illustrates the Poisson inverse Pareto distribution's application using two real datasets to demonstrate that the proposed distribution provides a superior fit to that of other models.

Keywords: Unified model, long-term survival model, mixture distribution, value-at-risk, tail behavior.

1. Introduction

The inverse Pareto distribution (IPD) originates from an inverse transformation of the Pareto type II distribution (PD), also called the Lomax distribution. The probability density function (PDF) of a random variable X from the PD is

$$f_{PD}(x) = \frac{\eta\lambda^{\eta}}{(x+\lambda)^{\eta+1}},\ x > 0,$$

where the scale parameter $\lambda > 0$ and shape parameter $\eta > 0.$ The IPD has a decreasing hazard function and a heavy tail distribution (Dankunprasert et al. 2021), and is frequently applied to claim modeling for estimating or forecasting the behavior of claims that will occur in the future (Klugman et al., 2012). If X is a random variable from the IPD, then the PDF of the IPD with scale parameter β and shape parameter α is defined as

$$f_{IPD}(x) = \frac{\beta\alpha x^{\alpha-1}}{(x+\beta)^{\alpha+1}},\ x > 0,\ \beta > 0,\ \alpha > 0. \tag{1}$$

The corresponding cumulative distribution function (CDF) and survival function are respectively given as:

$$F_{IPD}(x) = \left(\frac{x}{x+\beta}\right)^{\alpha}, \; x > 0, \; \beta > 0, \; \alpha > 0, \tag{2}$$

$$S_{IPD}(x) = 1 - \left(\frac{x}{x+\beta}\right)^{\alpha}, \; x > 0, \; \beta > 0, \; \alpha > 0. \tag{3}$$

Recently, new models or distributions have been developed from combined models to fitting to data that cannot be fit by a commonly used distribution. However, the developed models should be sufficient for most modeling situations. The general methods for this combined distribution include the finite (McLachlan and Peel 2000, Hall and Zhou 2003, Balakrishnan et al. 2009, Erisoglu et al. 2013) and infinite mixture distributions (Bulmer 1974, Emilio et al. 2008, Withers and Nadarajah 2011). Nevertheless, to the best of our knowledge, only a few studies related to the IPD exist thus far. This paper proposes a new IPD mixture distribution through a unified approach in a long-term survival analysis: the Poisson inverse Pareto distribution (PIPD).

The remainder of this paper is organized as follows. Section 2 presents a type of mixture distribution originally derived from a unified approach in a long-term survival analysis as proposed by Rodrigues et al. (2009); this is observed as a useful way to generate a new distribution. Section 3 presents the PIPD and derives its survival function, CDF, PDF, and hazard function. Additionally, Section 4 presents some properties of the PIPD, such as its value-at-risk, tail behavior, quantile function, and order statistics. Section 5 provides a parameter estimation using the maximum likelihood method for the PIPD. Section 6 discusses the results of a Monte Carlo simulation of the maximum likelihood estimates' behavior. Section 7 illustrates the PIPD using two real datasets, and Section 8 concludes the paper.

2. Method of Mixture Distribution

Rodrigues et al. (2009) first proposed a unified long-term survival model, also known as a cured model. Their model combined the long-term survival models proposed by Berkson and Gage (1952) and Chen et al. (1999) by using the generating function of a real sequence as introduced by Feller (1968), as follows:

Let N be an unobserved random variable denoting the number of competing causes related to the occurrence of a noteworthy event with a probability distribution $p_n = P(N = n)$ of $n = 0,1,2,3,\ldots.$ Given that $N = n,$ the random variables $Z_i, i = 1,2,\ldots,n$ denote the time to the event for the i^{th} cause and are Z_i independent of $N.$ To include individuals that are not sensitive to the event of interest, the observable time to the occurrence is defined as $X = \min\{Z_1, Z_2, \ldots, Z_N\}$ if $N \geq 1,$ and $P(Z_0 = \infty) = 1$ if $N = 0.$

The long-term survival function of the random variable $X,$ denoted by $S_{LT}(x),$ is given by

$$\begin{aligned} S_{LT}(x) &= P(N = 0) + P(Z_1 > x, Z_2 > x, \ldots, Z_N > x, N \geq 1) \\ &= P(N = 0) + \sum_{n=1}^{\infty} P(N = n) P(Z_1 > x, Z_2 > x, \ldots, Z_N > x) \\ &= p_0 + \sum_{n=1}^{\infty} p_n [S(x)]^n = \sum_{n=0}^{\infty} p_n [S(x)]^n = A[S(x)], \end{aligned} \tag{4}$$

where $A[\cdot]$ is a generating function of the sequence $p_n,$ which converges if $0 \leq S(x) \leq 1,$ as defined by Feller (1968), and $S(x)$ is a survival function. Further, $S_{LT}(x)$ is an improper survival function, since $\lim_{x\to\infty} S_{LT}(x) = p_0$.

Moreover, Rodrigues et al. (2009) had demonstrated $S_{LT}(x)$ in the form of

$$S_{LT}(x) = p_0 + (1-p_0)S_M(x), \tag{5}$$

where $S_M(x) = \dfrac{\sum_{n=1}^{\infty} p_n [S(x)]^n}{1-p_0}$ is the proper survival function, and $\lim_{x\to\infty} S_M(x) = 0$. The long-term survival function in Equation (5) can be noted as the long-term survival model first proposed by Berkson and Gage (1952), also known as a mixture cure model. This paper redefines this concept of a unified model as a survival function for creating a new distribution, as in the following Theorem 1.

Theorem 1. *Let N be a random variable denoting the number of occurrences of a noteworthy event with the following probability distribution $p_n = P(N=n)$ for $n = 0,1,2,3,\ldots$. Subsequently, the mixture survival function, denoted as $S_M(x)$, is*

$$S_M(x) = \frac{A[S(x)] - p_0}{1-p_0}, \tag{6}$$

where $p_0 = P(N=0)$, $A[\cdot]$ is a probability-generating function of N, and $S(x)$ is a survival function.

Proof: From Rodrigues et al. (2009) in Equation (5), we have

$$S_M(x) = \frac{S_{LT}(x) - p_0}{1-p_0} = \frac{\sum_{n=0}^{\infty} p_n [S(x)]^n - p_0}{1-p_0} = \frac{A[S(x)] - p_0}{1-p_0}.$$

The mixture survival function in Equation (6) is perhaps called a “proper” survival function given the non-cured population in the long-term survival model. The literature mentions various mixed models, such as the geometric Birnbaum-Saunders and odd Birnbaum-Saunders geometric survival functions as presented by Cancho et al. (2012) and Ortega et al. (2017), respectively.

3. The Poisson Inverse Pareto Distribution

This section proposes a new mixture distribution of the Poisson distribution and IPD by using the mixed survival function from unifying long-term survival models as noted in Section 2.

Definition 1. Let N be a random variable of the Poisson distribution with a parameter θ. The probability-generating function (PGF) of the Poisson distribution is defined as

$$A_P(s) = \exp[-\theta(1-s)], \tag{7}$$

where $\theta > 0$, and it converges for $|s| \le 1$.

Definition 2. Let X be a random variable of the PIPD with parameters θ, β and α, denoted as $X \sim PIPD(\theta,\beta,\alpha)$ with $x > 0, \theta > 0, \beta > 0,$ and $\alpha > 0$.

Theorem 2. *The PIPD's survival function is*

$$S_{PIPD}(x) = \frac{e^{-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} - e^{-\theta}}{1-e^{-\theta}}, \tag{8}$$

where $x > 0, \theta > 0, \beta > 0$ *and* $\alpha > 0$.

Proof: From Theorem 1, let $A\left[S(x)\right]$ indicate the PGF of the Poisson distribution from Definition 1; the mixture survival function can be derived as

$$S_M(x) = \frac{\exp\left[-\theta\left(1-\mathrm{S}(x)\right)\right] - p_0}{1-p_0} = \frac{\exp\left[-\theta F(x)\right] - \exp(-\theta)}{1-\exp(-\theta)},$$

where $F(x) = 1-\mathrm{S}(x)$ and $p_0 = P(N=0) = \exp(-\theta)$. Thus, we replace $F(x)$ with $F_{IPD}(x)$ in Equation (2), and the PIPD survival function is expressed as

$$S_{PIPD}(x) = \frac{\exp\left[-\theta F_{IPD}(x)\right] - \exp(-\theta)}{1-\exp(-\theta)} = \frac{e^{-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} - e^{-\theta}}{1-e^{-\theta}}.$$

Figure 1 illustrates the survival curves for the $PIPD(\theta,\beta,\alpha)$ selected values of parameters θ, β, and α. This figure indicates that the survival curves are initially decreasing functions.

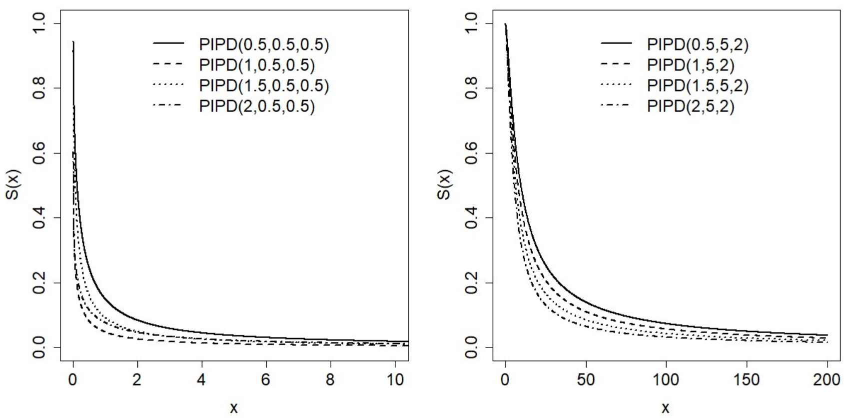

Figure 1 The PIPD's survival function with some specified parameter values

The corresponding CDF can be written as

$$F_{PIPD}(x) = 1 - S_{PIPD}(x) = \frac{1-e^{-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}}{1-e^{-\theta}}. \tag{9}$$

Figure 2 displays the CDF curves for the $PIPD(\theta,\beta,\alpha)$ selected values for the parameters θ, β, and α. This figure demonstrates that the CDF curves are initially increasing functions.

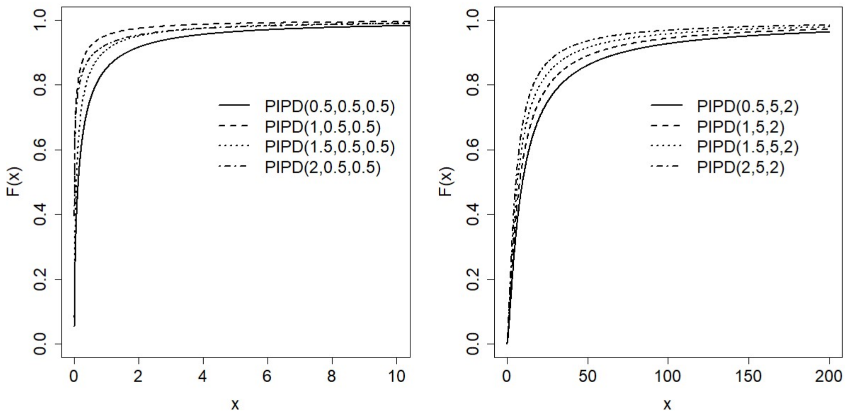

Figure 2 The PIPD's CDF with some specified parameter values

Theorem 3. *The PIPD's PDF is*

$$f_{PIPD}(x) = \frac{\theta\beta\alpha\left(\frac{x}{x+\beta}\right)^{\alpha} e^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}}{(e^{\theta}-1)x(x+\beta)}, \quad (10)$$

where $x > 0, \theta > 0, \beta > 0,$ *and* $\alpha > 0.$

Proof. As $f_{PIPD}(x) = -\frac{dS_{PIPD}(x)}{dx}$, this completes the proof of the theorem.

Figure 3 notes the PDF curves for the $PIPD(\theta,\beta,\alpha)$ selected values for the parameters θ, β, and α. The curves reveal the PIPD's positively skewed distribution. Moreover, the curves reveal that the PDF of the PIPD can have unimodal and decreasing shapes.

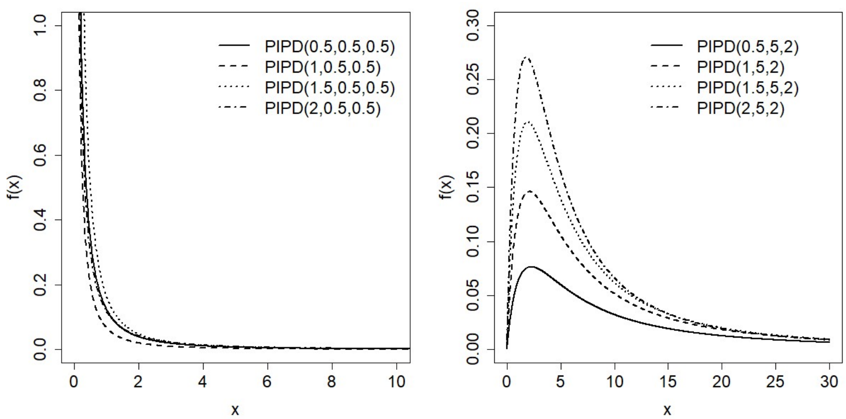

Figure 3 The PIPD's PDF with some specified parameter values

Theorem 4. *The hazard function of the PIPD is*

$$h_{PIPD}(x)=\frac{\alpha\beta\theta\left(1-\mathrm{e}^{-\theta}\right)\left(\dfrac{x}{x+\beta}\right)^{\alpha}\mathrm{e}^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}}{\left(\mathrm{e}^{\theta}-1\right)x\left(x+\beta\right)\left[\mathrm{e}^{-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}-\mathrm{e}^{-\theta}\right]}, \tag{11}$$

where $x>0,\theta>0,\beta>0,$ *and* $\alpha>0.$

Proof: As $h_{PIPD}(x)=\dfrac{f_{PIPD}(x)}{S_{PIPD}(x)}$, substituting $f_{PIPD}(x)$ with Equation (10) and $S_{PIPD}(x)$ using Equation (8) reveals the results.

Figure 4 presents the hazard curves for the $PIPD(\theta,\beta,\alpha)$ selected values of parameters θ, β, and α. The curves illustrate that the PIPD's hazard function can have unimodal and decreasing shapes. It shown that heavy-tailed distribution.

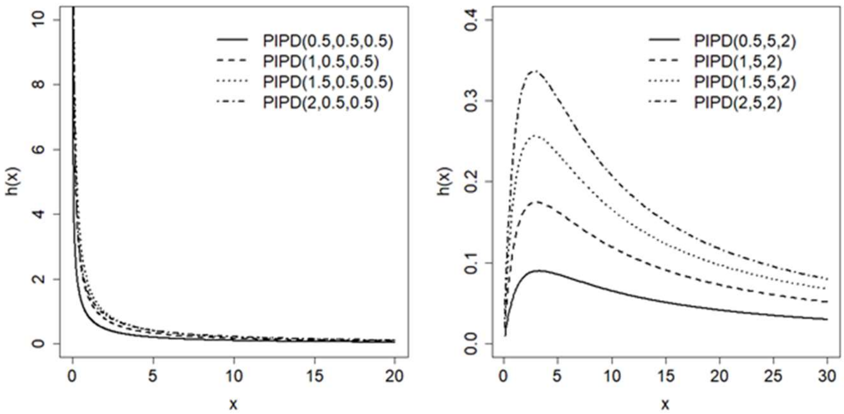

Figure 4 The PIPD's hazard function with some specified parameter values

4. Properties of the Poisson Inverse Pareto Distribution

This section derives some basic properties of the PIPD, such as its limiting behavior, value-at-risk, tail behavior, and quantile function. We also derived the density of the *k*th-order statistics.

Proposition 1. *The limit of the Poisson inverse Pareto density function* $x\to\infty$ *is 0, and the limit of the Poisson inverse Pareto hazard function* $x\to\infty$ *is 0.*

Proof: Let $f_{PIPD}\left(x\right)$ and $h_{PIPD}\left(x\right)$ as given in Equations (10) and (11), respectively. As $\lim_{x\to\infty} x/\left(x+\beta\right)=1,$ and by applying L'Hôpital's rule, the proposition is proved.

4.1. Value-at-risk

Theorem 5. *Let* X *be a random loss variable. The value-at-risk (VaR; Klugman et al. 2012) of* X *at the 100% level—denoted as* $\mathrm{VaR}_p\left(X\right)$ *or* π_p*—is the 100p percentile of the distribution of X.*

Consider a PIPD with parameters $\theta > 0, \beta > 0, \alpha > 0,$ and PDF as shown in Theorem 3. Subsequently, the VaR of the PIPD takes the form of

$$VaR_p\left(X\right) = \pi_p = \frac{\beta\left\{1-\frac{1}{\theta}\ln\left[1+\left(1-p\right)\left(e^{\theta}-1\right)\right]\right\}^{1/\alpha}}{1-\left\{1-\frac{1}{\theta}\ln\left[1+\left(1-p\right)\left(e^{\theta}-1\right)\right]\right\}^{1/\alpha}}.$$

Proof: The PDF of the PIPD is given in Equation (10). The value of π_p is expressed as follows:

$$P\left(X > \pi_p\right) = \int_{\pi_p}^{\infty} f_{PIPD}(x)dx = \frac{e^{\theta-\theta\left(\frac{\pi_p}{\pi_p+\beta}\right)^{\alpha}}-1}{e^{\theta}-1} = 1-p.$$

Hence, solving for π_p produces

$$\pi_p = \frac{\beta\left\{1-\frac{1}{\theta}\ln\left[1+\left(1-p\right)\left(e^{\theta}-1\right)\right]\right\}^{1/\alpha}}{1-\left\{1-\frac{1}{\theta}\ln\left[1+\left(1-p\right)\left(e^{\theta}-1\right)\right]\right\}^{1/\alpha}}.$$

4.2. Tail behavior

This section presents the PIPD's tail properties.

Proposition 2. *The PDF $f_{PIPD}(x)$ of $X \sim PIPD(\theta,\beta,\alpha)$ is decreasing if*

$\alpha\beta\left[1-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}\right]\Big/\left(2x+\beta\right) < 1.$

Proof: The first derivative of $f_{PIPD}(x)$ is given by

$$f'_{PIPD}(x) = -\frac{\alpha\beta\theta\left(\frac{x}{x+\beta}\right)^{\alpha}\mathrm{e}^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}\left\{\beta+2x+\alpha\beta\left[\theta\left(\frac{x}{x+\beta}\right)^{\alpha}-1\right]\right\}}{\left(\mathrm{e}^{\theta}-1\right)x^2\left(x+\beta\right)^2}.$$

If $\alpha\beta\left[1-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}\right]\Big/\left(2x+\beta\right) < 1,$ then $f'_{PIPD}(x) < 0$ for $x > 0, \theta > 0, \beta > 0,$ and $\alpha > 0,$ or specifically, the function $f_{PIPD}(x)$ is decreasing.

Proposition 3. *Let $X \sim PIPD(\theta,\beta,\alpha)$ with $x > 0, \theta > 0, \beta > 0,$ and $\alpha > 0$ with the PDF as given in Equation (10). If $A - B > 0$, then the PIPD has a heavy tail, where*

$$A = \alpha\beta\left(\frac{x}{x+\beta}\right)^{\alpha}\left[\mathrm{e}^{\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}-\mathrm{e}^{\theta}\right]$$

and

$$B = \left(\frac{x}{x+\beta}\right)^{\alpha} \left\{ (2x+\beta) \left[\mathrm{e}^{\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} - \mathrm{e}^{\theta} \right] + \alpha\beta\theta \left(\frac{x}{x+\beta}\right)^{\alpha} \mathrm{e}^{\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} \right\}.$$

Proof: The hazard function is shown in Equation (11). The first derivative of $h_{PIPD}(x)$ is then given by

$$h'_{PIPD}(x) = -\alpha\beta\theta \mathrm{e}^{\theta} (A - B) \left\{ x(x+\beta) \left[\mathrm{e}^{\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} - \mathrm{e}^{\theta} \right] \right\}^{-2},$$

where

$$A = \alpha\beta \left(\frac{x}{x+\beta}\right)^{\alpha} \left[\mathrm{e}^{\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} - \mathrm{e}^{\theta} \right],$$

$$B = \left(\frac{x}{x+\beta}\right)^{\alpha} \left\{ (2x+\beta) \left[\mathrm{e}^{\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} - \mathrm{e}^{\theta} \right] + \alpha\beta\theta \left(\frac{x}{x+\beta}\right)^{\alpha} \mathrm{e}^{\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} \right\}.$$

If $A - B > 0,$ then $h'_{PIPD}(x) < 0$ for $x > 0, \theta > 0, \beta > 0,$ and $\alpha > 0$; specifically, the hazard function for the PIPD is decreasing and the PIPD has a heavy tail. The ratio of the two survival functions can be used to indicate if one distribution has a heavier tail than another, as this ratio should diverge to infinity when $x \to \infty$ (Klugman et al., 2012).

Proposition 4. *The PIPD has a heavier tail than the exponential distribution (ED).*

Proof: As observed by applying L'Hôpital's rule,

$$\lim_{x\to\infty} \frac{S_{PIPD}(x)}{S_{ED}(x)} = \lim_{x\to\infty} \frac{S'_{PIPD}(x)}{S'_{ED}(x)} = \lim_{x\to\infty} \frac{-f_{PIPD}(x)}{-f_{ED}(x)}.$$

The PDF of the PIPD is defined by Equation (10). Subsequently, the required limit is

$$\lim_{x\to\infty} \frac{f_{PIPD}(x)}{f_{ED}(x)} = \lim_{x\to\infty} \frac{\theta\beta\alpha\left(\frac{x}{x+\beta}\right)^{\alpha} e^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} \left[(e^{\theta}-1)x(x+\beta)\right]^{-1}}{\lambda e^{-\lambda x}} = c \lim_{x\to\infty} \frac{e^{\lambda x - \theta\left(\frac{x}{x+\beta}\right)^{\alpha}}}{x(x+\beta)\left(\frac{x+\beta}{x}\right)^{\alpha}},$$

where c is constant and the exponentials progress to infinity faster than the polynomials; the limit is infinity.

Proposition 5. *The PIPD exhibits a heavier tail than the gamma distribution (GD).*

Proof: Equation (10) defines the PDF of the PIPD, while τ and δ will be used for the GD parameters instead of the typical α and β. The remaining proof is similar to that of Proposition 4; the required limit is

$$\lim_{x\to\infty} \frac{f_{PIPD}(x)}{f_{GD}(x)} = \lim_{x\to\infty} \frac{\theta\beta\alpha\left(\frac{x}{x+\beta}\right)^{\alpha} e^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}} \left[(e^{\theta}-1)x(x+\beta)\right]^{-1}}{\delta^{\tau} x^{\tau-1} e^{-\delta x} \Gamma(\tau)^{-1}} > c \lim_{x\to\infty} \frac{e^{\delta x - \theta\left(\frac{x}{x+\beta}\right)^{\alpha}}}{(x+\beta)^{\tau+1}},$$

where c is constant. Therefore, the limit is infinity.

4.3. Quantile function

Let $X \sim PIPD(\theta,\beta,\alpha)$ with $x>0, \theta>0, \beta>0,$ and $\alpha>0.$ The quantile function (QF; Gilchrist 2000) is denoted by $Q(p)$ and $Q(p)=F^{-1}(p),$ where $p\in(0,1).$

Proposition 6. *If $X \sim PIPD(\theta,\beta,\alpha),$ then the QF of X is given as*

$$Q(p)=\beta\left\{\frac{\ln\left[1-p\left(1-e^{-\theta}\right)\right]}{-\theta}\right\}^{1/\alpha}\left(1-\left\{\frac{\ln\left[1-p\left(1-e^{-\theta}\right)\right]}{-\theta}\right\}^{1/\alpha}\right)^{-1}.$$

Specifically, the median of $X \sim PIPD(\theta,\beta,\alpha)$ is given as

$$Q(0.5)=\beta\left\{\frac{\ln\left[1-0.5\left(1-e^{-\theta}\right)\right]}{-\theta}\right\}^{1/\alpha}\left(1-\left\{\frac{\ln\left[1-0.5\left(1-e^{-\theta}\right)\right]}{-\theta}\right\}^{1/\alpha}\right)^{-1}.$$

Proof. Since $Q(p)=F^{-1}(p),\ p\in(0,1).$ This implies that $F(Q(p))=p.$ As $X \sim PIPD(\theta,\beta,\alpha),$ or the CDF of X as noted in Equation (9), we can obtain

$$F(Q(p))=\frac{1-e^{-\theta\left(\frac{Q(p)}{Q(p)+\beta}\right)^{\alpha}}}{1-e^{-\theta}}=p.$$

Hence, we can solve for $Q(p)$ to obtain

$$Q(p)=\beta\left\{\frac{\ln\left[1-p\left(1-e^{-\theta}\right)\right]}{-\theta}\right\}^{1/\alpha}\left(1-\left\{\frac{\ln\left[1-p\left(1-e^{-\theta}\right)\right]}{-\theta}\right\}^{1/\alpha}\right)^{-1}.$$

4.4. Order statistics

Let $X \sim PIPD(\theta,\beta,\alpha)$ with $x>0, \theta>0, \beta>0,$ and $\alpha>0$ with PDF and CDF and as given in Equations (10) and (9), respectively. The density of the *k*th-order statistic in a random sample of size $n,$ or $X_{(1)},X_{(2)},\ldots,X_{(n)},$ is $f_{k:n}(x)=\dfrac{n!}{(k-1)!(n-k)!}f(x)\left[F(x)\right]^{k-1}\left[1-F(x)\right]^{n-k}.$

Therefore, the *k*th-order statistic of random variable $X \sim PIPD(\theta,\beta,\alpha)$ is

$$f_{k:n}(x)=\frac{n!}{(k-1)!(n-k)!}\cdot\frac{\theta\beta\alpha\left(\frac{x}{x+\beta}\right)^{\alpha}e^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}}{(e^{\theta}-1)x(x+\beta)}\left[\frac{1-e^{-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}}{1-e^{-\theta}}\right]^{k-1}\left[1-\frac{1-e^{-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}}{1-e^{-\theta}}\right]^{n-k}.$$

In simplifying this, we obtain

$$f_{k:n}(x)=\frac{n!\theta\beta\alpha e^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}x^{\alpha-1}\left[e^{\theta}-e^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}\right]^{k-1}\left[e^{\theta-\theta\left(\frac{x}{x+\beta}\right)^{\alpha}}-1\right]^{n-k}}{(k-1)!(n-k)!(e^{\theta}-1)^{n}(x+\beta)^{\alpha+1}}.$$

5. Parameter Estimation of the Poisson Inverse Pareto Distribution

A maximum likelihood estimation (MLE) was presented to calculate the PIPD's parameters, with the likelihood function as follows

$$L(\theta,\beta,\alpha)=\prod_{i=1}^{n}\frac{\theta\beta\alpha\left(\frac{x_i}{x_i+\beta}\right)^{\alpha}e^{\theta-\theta\left(\frac{x_i}{x_i+\beta}\right)^{\alpha}}}{(e^{\theta}-1)x_i(x_i+\beta)}.$$

The log-likelihood function of above expression is given by

$$LL(\theta,\beta,\alpha)=n\log\theta+n\log\beta+n\log\alpha-n\log\left(e^{\theta}-1\right)+(\alpha-1)\sum_{i=1}^{n}\log x_i$$
$$-(\alpha+1)\sum_{i=1}^{n}\log\left(x_i+\beta\right)+n\theta-\theta\sum_{i=1}^{n}\left(\frac{x_i}{x_i+\beta}\right)^{\alpha}.$$

The log-likelihood function leads to the following partial derivatives relative to $\theta,\beta,$ and α, by which the parameters' optimal values can be obtained. The score equations were derived as follows:

$$\frac{\partial LL(\theta,\beta,\alpha)}{\partial\theta}=\frac{n}{\theta}-\frac{ne^{\theta}}{e^{\theta}-1}+n-\sum_{i=1}^{n}\left(\frac{x_i}{x_i+\beta}\right)^{\alpha}$$

$$\frac{\partial LL(\theta,\beta,\alpha)}{\partial\beta}=\frac{n}{\beta}-\left(\alpha+1\right)\sum_{i=1}^{n}\left(\frac{1}{x_i+\beta}\right)+\theta\alpha\sum_{i=1}^{n}\left[\frac{x_i^{\alpha}}{\left(x_i+\beta\right)^{\alpha+1}}\right]$$

$$\frac{\partial LL(\theta,\beta,\alpha)}{\partial\alpha}=\frac{n}{\alpha}+\sum_{i=1}^{n}\log x_i-\sum_{i=1}^{n}\log\left(x_i+\beta\right)-\theta\sum_{i=1}^{n}\left\{\left(\frac{x_i}{x_i+\beta}\right)^{\alpha}\log\left(\frac{x_i}{x_i+\beta}\right)\right\}.$$

The derivatives of these equations relative to $\theta,\beta,$ and α are set equal to zero to estimate the parameters, and the following equations are obtained:

$$\frac{\partial LL(\theta,\beta,\alpha)}{\partial\theta}=0,\ \frac{\partial LL(\theta,\beta,\alpha)}{\partial\beta}=0\ \text{ and }\ \frac{\partial LL(\theta,\beta,\alpha)}{\partial\alpha}=0.$$

The MLE solutions for $\theta,\beta,$ and α can be obtained by simultaneously solving the resulting equations using a numerical procedure, such as the Newton-Raphson method. This study obtains the MLE estimates of $\hat{\theta},\hat{\beta},$ and $\hat{\alpha}$ by using the "fitdist" function in the R software suite's "fitdistrplus" package (Delignette-Muller and Dutang 2015).

6. Simulation Study

This section presents the results of a simulation study to assess the effectiveness of the MLE of the parameters $\theta,\beta,$ and α in the previous section. The estimates of $\theta,\beta,$ and α are obtained using the "fitdist" function in the R software suite's "fitdistrplus" package (Delignette-Muller and Dutang 2015).

The study was based on 2,000 simulated samples from the PIPD with different sample sizes: n = 30, 50, 100, and 200. We generate the random variables from the PIPD by using the inverse of the distribution function. Consider the identity

$$F(X)=U\Rightarrow X=F^{-1}(U),$$

where U is the standard uniform distribution, or the uniform (0,1). Let $X\sim PIPD(\theta,\beta,\alpha)$; the random variable can be generated from

$$X = \beta\left\{\frac{\ln\left[1-U\left(1-e^{-\theta}\right)\right]}{-\theta}\right\}^{1/\alpha}\left(1-\left\{\frac{\ln\left[1-U\left(1-e^{-\theta}\right)\right]}{-\theta}\right\}^{1/\alpha}\right)^{-1}.$$

Table 1 Mean estimates, standard deviation, and mean-squared errors of $\theta, \beta,$ and α

Distributions	n	Parameter	Mean Estimate	SD	MSE
PIPD(5, 5, 2)	30	θ	5.67804	1.35501	2.29488
		β	4.96627	2.34392	5.49237
		α	2.40112	0.78609	0.77854
	50	θ	5.75277	1.23825	2.09918
		β	5.31546	2.16483	4.78367
		α	2.23798	0.62319	0.44481
	100	θ	5.89390	0.99168	1.78200
		β	6.04580	1.88722	4.65354
		α	2.02301	0.38744	0.15056
	200	θ	5.92425	0.74619	1.41076
		β	6.34906	1.56491	4.26770
		α	1.92699	0.24149	0.06362
PIPD(10, 5, 2)	30	θ	11.28234	3.56331	14.33524
		β	5.31141	3.01654	9.19193
		α	2.43390	0.93418	1.06052
	50	θ	11.21392	3.36491	12.79054
		β	5.45110	2.75342	7.78101
		α	2.25388	0.69790	0.55127
	100	θ	11.08144	2.96769	9.97230
		β	5.53198	2.30517	5.59418
		α	2.10218	0.42845	0.19392
	200	θ	10.87044	2.53463	7.17879
		β	5.48472	1.94123	4.00144
		α	2.04358	0.28125	0.08096

Table 1 presents the mean of the parameter estimates as well as the standard deviation (SD) and mean-squared errors (MSEs) of the parameter estimates for different sample sizes. It is observed that the estimates of $\theta, \beta,$ and α are close to true values. Further, the MSE values for the estimates of $\theta, \beta,$ and α decrease when the sample size n increases.

7. Applications

This section evaluates the proposed model's efficiency by applying it to two real datasets. Specifically, we consider two datasets for Danish reinsurance claims and bladder cancer patients, which are applied in insurance and survival analyses, respectively. These datasets were fitted to the PIPD, PD, and IPD. The MLE was then used to estimate parameters in the PIPD and other comparative models.

The model comparison is conducted using Akaike's information criterion (AIC) and Bayesian information criterion (BIC), given by

$$AIC = -2LL(\hat{\boldsymbol{\delta}}) + 2k \text{ and } BIC = -2LL(\hat{\boldsymbol{\delta}}) + k\log(n),$$

where $LL(\hat{\boldsymbol{\delta}})$ denotes the log-likelihood function with a vector estimated parameter $\hat{\boldsymbol{\delta}}$, k is the number of estimated parameters, and n is the sample size. The model with the smallest value for these criteria was used as the preferred model to describe the dataset.

7.1. Danish reinsurance claims

The Danish reinsurance claims dataset includes 2,167 industrial fire losses gathered from the Copenhagen Reinsurance Company from 1980 to 1990. The data have been made publicly available from the "fitdistrplus" package in R programming language (Delignette-Muller and Dutang 2015), under the dataset name as *danishuni*. We list some descriptive statistics in Table 2.

Table 2 Descriptive statistics of Danish reinsurance claims

n	minimum	maximum	median	mean	SD
2,167	1	263.25	1.778	3.385	8.507452

Table 3 presents the estimated parameters from the comparative models' MLE, AIC, and BIC values, and reveals that the PIPD fits the data better than the other related models. Moreover, the PIPD provides a better fit to this dataset than the others, as illustrated in Figure 5.

Table 3 Parameter estimates (standard deviations in parentheses), AIC, and BIC from Danish reinsurance claims dataset

Distributions		Estimate		AIC	BIC
PD	$\hat{\lambda}$ = 13.84482 (1.43131)	$\hat{\eta}$ = 5.36944 (0.48191)		9,249.666	9,261.029
IPD	$\hat{\beta}$ = 0.014523 (0.00242)	$\hat{\alpha}$ = 126.31754 (21.04974)		8,548.378	8,559.741
PIPD	$\hat{\theta}$ = 5.39849 (0.15788)	$\hat{\beta}$ = 0.04431 (0.00851)	$\hat{\alpha}$ = 92.81334 (17.52509)	3,084.084	3,101.128

7.2. Bladder cancer patients

We consider an uncensored dataset corresponding to remission times (in months) of a random sample of 128 bladder cancer patients as reported by Lee and Wang (2003). We list some descriptive statistics in Table 4.

Table 4 Descriptive statistics of bladder cancer patients

n	minimum	maximum	median	mean	SD
128	0.080	79.050	6.395	9.366	10.50833

It can be observed from Table 5 that the PIPD provides the best fit for these data among all the models considered. Figure 6 presents the probability–probability plots for all the models considered for such data, indicating that the PIPD provides a better fit to this dataset than the others.

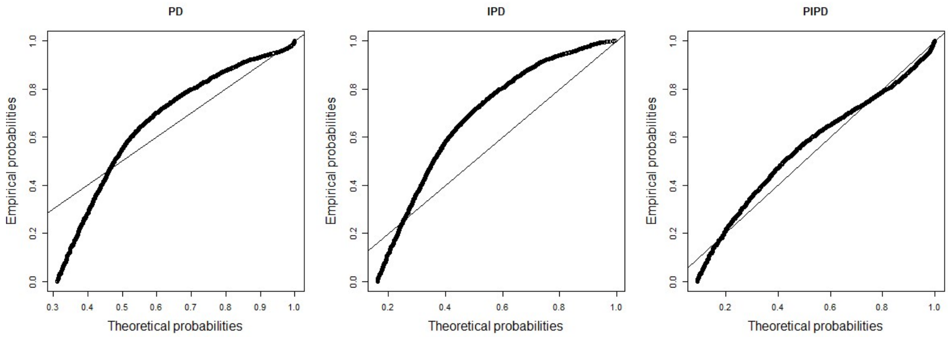

Figure 5 The probability–probability plots from the Danish reinsurance claims dataset

Additionally, Gharib et al. (2017) studied this dataset's fit to the Marshall-Olkin extended inverse Pareto distribution. Their results indicate that this distribution's AIC value is 831.5835, which is greater than that of the PIPD. Hence, the PIPD is a better fit than this distribution.

Table 5 Parameter estimates (standard deviations in parentheses), AIC, and BIC from bladder cancer patient dataset

Distributions	Estimate			AIC	BIC
PD	$\hat{\lambda}$ = 12.08017 (142.08126)	$\hat{\eta}$ = 13.9122 (15.31263)		831.6658	837.3698
IPD	$\hat{\beta}$ = 2.00351 (0.63189)	$\hat{\alpha}$ = 2.461103 (0.59331)		853.3514	859.0555
PIPD	$\hat{\theta}$ = 9.15170 (3.11380)	$\hat{\beta}$ = 3.784981 (17.28320)	$\hat{\alpha}$ = 1.43878 (0.20197)	569.8726	578.4287

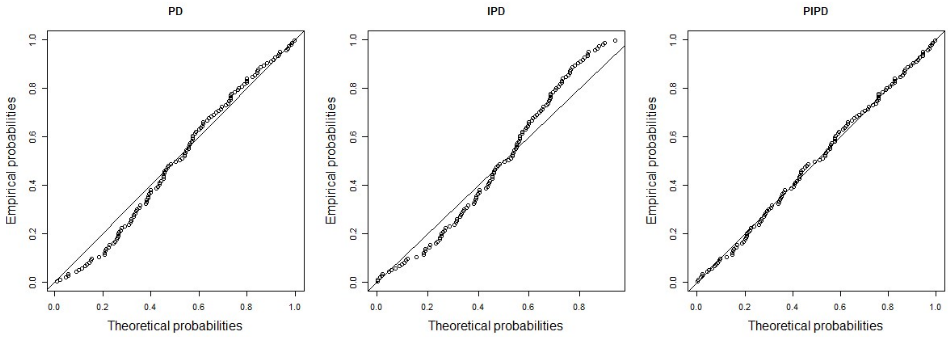

Figure 6 The probability–probability plots from the bladder cancer patient dataset

8. Conclusions

This paper provides a new combined inverse Pareto distribution called the Poisson inverse Pareto distribution. Its basic statistical properties as established in this work include the probability density, cumulative distribution, survival, and hazard functions. The results of this research indicate that the PIPD's VaR exists, while the tail behavior study revealed that the PIPD has a heavy tail and a heavier tail than the exponential and gamma distributions. Further, this work derived expressions for the QF and order statistics, with parameters estimated using the maximum likelihood method. The simulation study demonstrated that the PIPD parameters are approximate to the true parameter values; the mean-squared error values also decrease as the sample size increases. An application to the two real datasets reveals that the PIPD is more efficient than the Pareto and inverse Pareto distributions for both the Danish reinsurance claims and bladder cancer patient datasets. As it increases the parameter, it makes the distribution more flexible. The results indicate that PIPD may be used for a wider range of statistical applications. Further studies can examine a new mixture distribution given the introduced approach. Researchers can also study various other methods to estimate PIPD parameters, such as the Bayesian approach. The parameters of the proposed distribution can be estimated based on censored data. Furthermore, a study on the effect of covariates can be conducted based on the proposed distribution as the regression model.

Funding details. This work did not receive any financial support.

Disclosure statement. The authors declare no conflict of interest.

References

Balakrishnan N, Leiva V, Sanhueza A, Cabrera E. Mixture inverse Gaussian distributions and its transformations, moments and applications. Statistics. 2009; 43(1): 91-104.

Berkson J, Gage RP. Survival curve for cancer patients following treatment. J Am Stat Assoc. 1952; 47(259): 501-515.

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

Cancho VG, Louzada F, Barriga GDC. The geometric Birnbaum-Saunders regression model with cure rate. J Stat Plan Inference. 2012; 142: 993-1000.

Chen MH, Ibrahim JG, Sinha D. A new Bayesian model for survival data with a surviving fraction. J Am Stat Assoc. 1999; 94(447): 909-919.

Dankunprasert S, Jaroengeratikun U, Talangtam T. The properties of inverse Pareto distribution and its application to extreme events. Thail Stat. 2021; 19(1): 1-13.

Delignette-Muller ML, Dutang C. fitdistrplus: An R Package for Fitting Distributions. J Stat Softw. 2015; 64(4): 1-34.

Emilio GD, Sarabia JM, Enrique CO. Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications. Insur Math Econ. 2008; 42(1): 39-49.

Erisoglu M, Servi T, Erisoglu U, Calis N. Mixture gamma distribution for estimation of wind power potential. Int J Appl Math Stat. 2013; 40(10): 223-231.

Feller W. An introduction to probability theory and its applications, volume 1. New York: John Wiley & Sons; 1968.

Gharib M, Mohammed BI, Aghel WER. Marshall-Olkin extended inverse Pareto distribution and its application. Int J Stat Prob. 2017; 6(6): 71-84.

Gilchrist WG. Statistical modelling with quantile functions. Florida: Chapman & Hall/CRC; 2000.

Hall P, Zhou X. Nonparametric estimation of component distributions in a multivariate mixture. Ann Stat. 2003; 31(1): 201-224.

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

Lee ET, Wang JW. Statistical methods for survival data analysis. New York: John Wiley & Sons; 2003.

McLachlan GJ, Peel D. Finite mixture models. New York: John Wiley & Sons; 2000.

Ortega EMM, Cordeiro GM, Suzuki AK, Ramires TG. A new extended Birnbaum-Saunders model with cure fraction: classical and Bayesian approach. Commun Stat Appl Methods. 2017; 24: 397-419.

Rodrigues J, Cancho VG, de Castro M, Louzada-Neto F. On the unification of long-term survival models. Stat Probab Lett. 2009; 79(6): 753-759.

Withers CS, Nadarajah S. On the compound Poisson gamma distribution. Kybernetika. 2011; 14(1): 15-37.