Research Article

On the New Weight Parameter of the Mixture Pareto Distribution and Its Application to Real Data

Chanaphun Chananet and Wikanda Phaphan*
Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

* Corresponding author. E-mail: wikanda.p@sci.kmutnb.ac.th DOI: 10.14416/j.asep.2020.10.002
Received: 22 April 2020; Revised: 29 July 2020; Accepted: 5 August 2020; Published online: 5 October 2020

Abstract
This research aims to investigate the theoretical properties of new distribution which are based on the preceding well-known distributions by adding the new weight parameter. The new distribution is called the new weight parameter of the mixture of Pareto distribution. The advantage of the new distribution is the right-skewed distribution that is depending on the shape parameter. This property made the new mixture of Pareto distribution have high benefits of reliability engineering and survival analysis. Besides, we applied the purposed distribution to three sets of actual data: cement industry data, hard drive failure data, and heart attack data to show that new mixture Pareto distribution is fit to real data for generating data models.

Keywords: Length-biased Pareto distribution, Weight parameter, Cement industry, Failure, Heart attack

1 Introduction

Reliability engineering is used for analyzing periods of time deterioration of non-living things. The values of the measurements might express in the frequencies of a piece of equipment or machine that fails. These values of the variables are always non-negative real values which are called lifetime data. In statistics, the values of the measurements were random variables that resulted from random experimentation. When assuming the processes of random experimentation occurred continuously and endlessly. Therefore, the conclusions of lifetime data are necessary to use probability distribution to control random variables.

Reliability engineering, survival analysis, event history analysis in sociology, and duration analysis in economics mostly deal with non-negative random variables T. The distribution of the random variable T which has non-negative real values will have a well-known specific name in the above-fields. It has a well-known specific name which is lifetime distribution or survival distribution.

The lifetime distribution contains the following distributions: exponential distribution, gamma distribution, Pareto distribution, Weibull distribution, log-normal distribution, extreme value distribution, Gompertz distribution, generalized F distribution, and log-logistic distribution, and so forth. For this research, the authors are interested in studying a Pareto distribution.

The Pareto distribution arises as a tractable model in reliability engineering and survival analysis. It is used in the frequency modeling of data with non-negative data and skewed curves, for example, Hager and Bain [1], Boyd [2], and Brazauskas and Serfling [3]. A random variable X_1 is said to have the Pareto distribution (Doostparast and Balakrishnan, [4]), denoted by $X_1 \sim$ *Pareto* (α, β), its probability density function (pdf) is:

$$g_p(x_1;\alpha,\beta)=\frac{\alpha}{\beta}\left(\frac{x_1}{\beta}\right)^{-(\alpha+1)} I_{\{x_1\geq\beta\}}, x_1\geq\beta, \alpha>0, \beta>0 \quad (1)$$

Please cite this article as: C. Chananet and W. Phaphan, "On the new weight parameter of the mixture Pareto distribution and its application to real data," *Applied Science and Engineering Progress*, vol. 14, no. 3, pp. 460–467, Jul.–Sep. 2021, doi: 10.14416/j.asep.2020.10.002

Patil and Rao [5] presented a length biased Pareto (LBP) distribution by the concept of a weighted distribution. If X_1 is Pareto random variable with pdf Equation (1), then the pdf for the length biased distribution of the random variable X_2 is [Equation (2)]:

$$g_{LBP}(x_2;\alpha,\beta)=\frac{(\alpha-1)}{\beta}\left(\frac{x_2}{\beta}\right)^{-\alpha} I_{\{x_2\geq\beta\}}, x_2\geq\beta,\alpha>1,\beta>0. \quad (2)$$

It may be observed that can be written in the following form [Equation (3)]:

$$g_{LBP}(x_2;\alpha,\beta)=\frac{x_2(\alpha-1)}{\alpha\beta}g_p(x_2;\alpha,\beta)I_{\{x\geq\beta\}},$$
$$x_2\geq\beta,\alpha>1,\beta>0. \quad (3)$$

The mixture of the Pareto distribution with its complementary reciprocal is introduced by Nanuwong *et al*. [6]. It contains three parameters: scale parameter, shape parameter, and weight parameter. In modeling, the distribution with numerous parameters will be more complicated in applications. Therefore, the main aim of this paper is to introduce a new alternative distribution more suitable for the right tail data by adjusting new weight parameters. Causing the number of a mixture of Pareto distribution parameters will be reducing one parameter, and this makes it easier to apply. Hence, the proposed distribution will give great flexibility for modeling in reliability engineering, actuarial science, economics, finance, and survival analysis.

This paper is organized as follows. Section 2, briefly describes the mixture of the Pareto distribution. Section 3, gives the new weight parameter of the mixture Pareto distribution. Section 4, gives an application of the presented distribution with actual data. Finally, Section 5 contains conclusions.

2 A Brief Review of the Mixture Pareto Distribution

Nanuwong *et al*. [6] introduced the mixture Pareto distribution (MP) generated from the concept of the weighted two-component distribution which is mixed between a Pareto (P) and a length biased Pareto distributions (LBP). Special sub-models include the Pareto, exponential, chi-square, and logistic distributions. For this mixture Pareto distribution, it is formed by considering the random variable X such that [Equation (4)],

$$X=\begin{cases} X_1 & with \quad probability \quad (1-\varpi) \\ X_2 & with \quad probability \quad \varpi \end{cases} \quad (4)$$

where $x\geq\beta$, $\alpha>1$, $\beta>0$, and $0\leq\varpi\leq1$. Obviously, X is a mixture of X_1 and X_2. The pdf of X is given by the following formula:

$$f_{MP}(x)=(1-\varpi)g_p(x)+\varpi g_{LBP}(x) \quad (5)$$

From Equation (5), the pdf, can be expressed as:

$$f_{MP}(x;\alpha,\beta)=\frac{1}{\beta}\left(\frac{x}{\beta}\right)^{-(\alpha+1)}\left[(1-\varpi)\alpha+\frac{\varpi(\alpha-1)x}{\beta}\right],$$
$$x\geq\beta,\ \alpha>1,\ \beta>0,\ 0\leq\varpi\leq1.$$

3 The New Weight Parameter of the Mixture Pareto Distribution

The new distribution, which was introduced in this paper, was a combination of two distributions. They are Pareto distribution and length-biased Pareto distribution in the form of adjusted weight parameters.

3.1 *Probability density function*

Definition 3.1 Define X_1 and X_2 as an independent random variable which X_1 ~ *Pareto* (α, β) and X_2 ~ *LBP* (α, β). Then X is a new random variable that follows a new mixture Pareto distribution X ~ *NMP* (α, β) which there is the pdf as this formula [Equation (7)] and (Figure 1):

$$f_{NMP}(x)=\left(\frac{\alpha}{\alpha+1}\right)g_p(x)+\left(\frac{1}{\alpha+1}\right)g_{LBP}(x), \quad (7)$$

where $x\geq\beta$, $\alpha>1$, $\beta>0$.

Theorem 3.1 If X is a random variable of the new mixture Pareto distribution X ~ *NMP* (α, β), then the pdf of X is given by [Equation (8)] and (Figure 2):

$$f_{NMP}(x)=\frac{1}{(\alpha+1)\beta}\left(\frac{x}{\beta}\right)^{-\alpha}\left[\alpha^2\left(\frac{x}{\beta}\right)^{-1}+(\alpha-1)\right], \quad (8)$$

where $x\geq\beta$, $\alpha>1$, $\beta>0$.

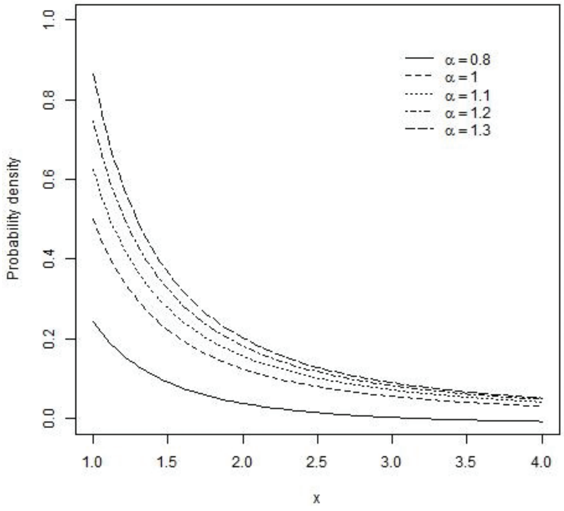

Figure 1: Density function of the new weight parameter of the mixture Pareto distribution for different values of α, where $\beta = 2$.

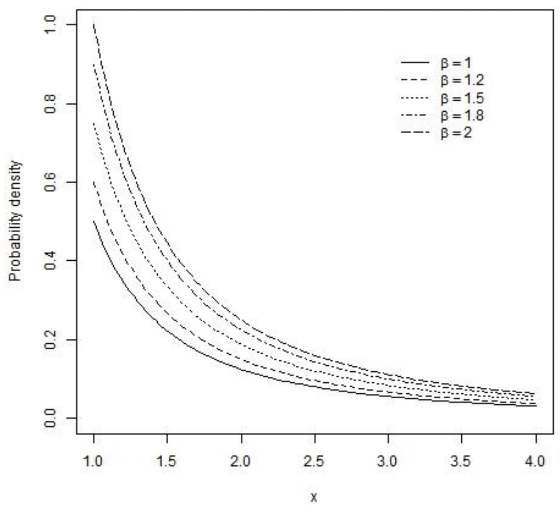

Figure 2: Density function of the new weight parameter of the mixture Pareto distribution for different values of β, where $\alpha = 2$.

Proof:

$$f_{NMP}(x)$$

$$=\left(\frac{\alpha}{\alpha+1}\right)g_p(x)+\left(\frac{1}{\alpha+1}\right)g_{LBP}(x)$$

$$=\left(\frac{\alpha}{\alpha+1}\right)\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{-(\alpha+1)}+\left(\frac{1}{\alpha+1}\right)\frac{(\alpha-1)}{\beta}\left(\frac{x}{\beta}\right)^{-\alpha}$$

$$=\frac{\alpha^2}{(\alpha+1)\beta}\left(\frac{x}{\beta}\right)^{-(\alpha+1)}+\frac{(\alpha-1)}{(\alpha+1)\beta}\left(\frac{x}{\beta}\right)^{-\alpha}$$

$$=\frac{1}{(\alpha+1)\beta}\left[\alpha^2\left(\frac{x}{\beta}\right)^{-(\alpha+1)}+(\alpha-1)\left(\frac{x}{\beta}\right)^{-\alpha}\right]$$

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

3.2 *Cumulative distribution function*

Theorem 3.2 If X is a random variable which $X \sim NMP(\alpha, \beta)$, then the cumulative distribution function of X can be written as [Equation (9)]:

$$F(x)=1-\left(\frac{\alpha}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-\alpha}-\left(\frac{1}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-(\alpha-1)}. \tag{9}$$

Proof: Let X be a continuous non-negative random variable, then the cumulative distribution function of X is expressed as:

$$F(x)$$

$$=\left(\frac{\alpha}{\alpha+1}\right)G_p(x)+\left(\frac{1}{\alpha+1}\right)G_{LBP}(x)$$

$$=\left(\frac{\alpha}{\alpha+1}\right)\left[1-\left(\frac{x}{\beta}\right)^{-\alpha}\right]+\left(\frac{1}{\alpha+1}\right)\left[1-\left(\frac{x}{\beta}\right)^{-(\alpha-1)}\right]$$

$$=\left(\frac{\alpha}{\alpha+1}\right)-\left(\frac{\alpha}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-\alpha}+\left(\frac{1}{\alpha+1}\right)-\left(\frac{1}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-(\alpha-1)}$$

$$=\left(\frac{\alpha+1}{\alpha+1}\right)-\left(\frac{\alpha}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-\alpha}-\left(\frac{1}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-(\alpha-1)}$$

$$=1-\left(\frac{\alpha}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-\alpha}-\left(\frac{1}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-(\alpha-1)}.$$

3.3 *Survival function*

Theorem 3.3 If X is a random variable which $X \sim NMP(\alpha, \beta)$, then the survival function of X can be written as [Equation (10)]:

$$S(x)=\frac{1}{(\alpha+1)}\left[\alpha\left(\frac{x}{\beta}\right)^{-\alpha}+\left(\frac{x}{\beta}\right)^{-(\alpha-1)}\right]. \quad (10)$$

Proof: Let X be a continuous random variable with cumulative distribution function $F(x)$ in the period of $[0, \infty)$ then survival function of X is:

$$S(x)=\int_x^\infty f(t)dt=1-F(x)$$

$$=1-\left[1-\left(\frac{\alpha}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-\alpha}-\left(\frac{1}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-(\alpha-1)}\right]$$

$$=\left(\frac{\alpha}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-\alpha}+\left(\frac{1}{\alpha+1}\right)\left(\frac{x}{\beta}\right)^{-(\alpha-1)}$$

$$=\frac{1}{(\alpha+1)}\left[\alpha\left(\frac{x}{\beta}\right)^{-\alpha}+\left(\frac{x}{\beta}\right)^{-(\alpha-1)}\right].$$

3.4 *Hazard rate function*

Theorem 3.4 If X is a random variable and $X \sim NMP$ (α, β), then the hazard rate function of X will be [Equation (11)] and (Figure 3):

$$h(x)=\frac{\frac{1}{\beta}\left[\left(\frac{x}{\beta}\right)^{-1}\alpha^2+(\alpha-1)\right]}{\left[\alpha+\frac{x}{\beta}\right]}. \quad (11)$$

Proof: Let X be a continuous non-negative random variable with a probability density function $f(x)$ and a survival function $S(x)$. Then, the hazard rate of X can be written as:

$$h(x)=\frac{f(x)}{S(x)}$$

$$=\frac{\frac{1}{(\alpha+1)\beta}\left(\frac{x}{\beta}\right)^{-\alpha}\left[\alpha^2\left(\frac{x}{\beta}\right)^{-1}+(\alpha-1)\right]}{\frac{1}{(\alpha+1)}\left[\alpha\left(\frac{x}{\beta}\right)^{-\alpha}+\left(\frac{x}{\beta}\right)^{-(\alpha-1)}\right]}$$

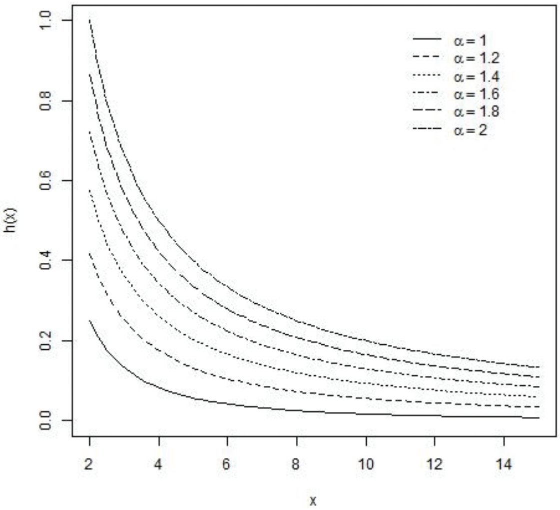

Figure 3: Hazard rate function.

$$=\frac{\frac{1}{\beta}\left(\frac{x}{\beta}\right)^{-\alpha}\left[\left(\frac{x}{\beta}\right)^{-1}\alpha^2+(\alpha-1)\right]}{\alpha\left(\frac{x}{\beta}\right)^{-\alpha}+\left(\frac{x}{\beta}\right)^{-(\alpha-1)}}$$

$$=\frac{\frac{1}{\beta}\left(\frac{x}{\beta}\right)^{-\alpha}\left[\left(\frac{x}{\beta}\right)^{-1}\alpha^2+(\alpha-1)\right]}{\left(\frac{x}{\beta}\right)^{-\alpha}\left[\alpha+\frac{x}{\beta}\right]}$$

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

3.5 *The r^{th} moments*

Theorem 3.5 Let $X \sim NMP$ (α, β), the r^{th} moment will be written in this form [Equation (12)]:

$$E(X^r)=\left(\frac{1}{\alpha+1}\right)\left[\frac{\alpha^2\beta^r}{\alpha-r}+\frac{(\alpha-1)\beta^r}{\alpha-(r+1)}\right]. \quad (12)$$

Proof: The r^{th} moment of a random variable X with $X \sim NMP$ (α, β) is:

$$E(X^r)=\left(\frac{\alpha}{\alpha+1}\right)E_p(X^r)+\left(\frac{1}{\alpha+1}\right)E_{LBP}(X^r)$$
$$=\left(\frac{\alpha}{\alpha+1}\right)\frac{\alpha\beta^r}{\alpha-r}+\left(\frac{1}{\alpha+1}\right)\left[\frac{(\alpha-1)\beta^r}{\alpha-(r+1)}\right]$$
$$=\left(\frac{1}{\alpha+1}\right)\left[\frac{\alpha^2\beta^r}{\alpha-r}+\frac{(\alpha-1)\beta^r}{\alpha-(r+1)}\right].$$

From the r^{th} moment, we can calculate the value of the mean of a new mixture Pareto distribution [Equation (13)]:

$$E(X)=\left(\frac{\alpha}{\alpha+1}\right)\frac{\alpha\beta}{\alpha-1}+\left(\frac{1}{\alpha+1}\right)\left[\frac{(\alpha-1)\beta}{\alpha-2}\right]$$
$$=\frac{1}{\alpha+1}\left[\frac{\alpha^2\beta}{\alpha-1}+\frac{(\alpha-1)\beta}{\alpha-2}\right]. \quad (13)$$

3.6 *Parameter estimation*

Parameters estimation of the mixture Pareto distribution with the new weighted-parameter will be done via the maximum likelihood estimation (MLE) procedure. The likelihood function of this distribution with parameters α and β is given by [Equation (14)]:

$$L(\alpha,\beta)=\prod_{i=1}^{n}\left\{\frac{1}{(\alpha+1)\beta}\left(\frac{x}{\beta}\right)^{-\alpha}\left[\alpha^2\left(\frac{x}{\beta}\right)^{-1}+(\alpha-1)\right]\right\}. \quad (14)$$

Its associated log-likelihood function can be written as:

$$l(\alpha,\beta)$$
$$=\log L(\alpha,\beta)$$
$$=-n\log(\alpha+1)-n\log\beta-\alpha\sum_{i=1}^{n}\log x_i+\alpha n\log\beta$$
$$+2n\log\alpha-\sum_{i=1}^{n}\log x_i+n\log\beta+n\log(\alpha-1)$$
$$=-n\log(\alpha+1)-(\alpha+1)\sum_{i=1}^{n}\log x_i+\alpha n\log\beta$$
$$+2n\log\alpha+n\log(\alpha-1). \quad (15)$$

Since $x \geq \beta$ then the MLE of β is min x_1. Since the Equation (15) cannot be solved analytically for α, the iterative method must be used. In this paper, we applied the Nelder and Mead [7] method by using the default function of the R program (R Core Team [8]), called the "optim" function to obtain the MLEs of α via the Nelder and Mead method.

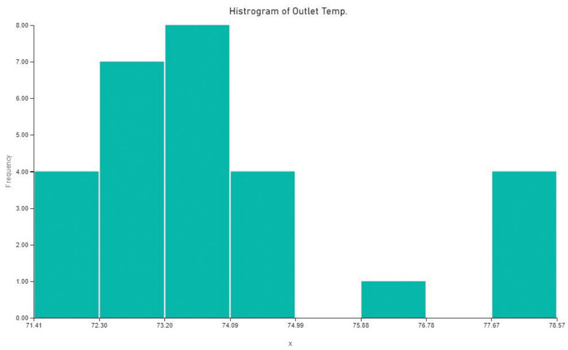

Figure 4: Histogram of the outlet hot air temperature of VRM (°C).

4 Application to Real Data Set

4.1 *Cement industry data*

In this subsection has presented an analysis of a comparison of the mixture Pareto distribution (MP) with the new parameter distribution (NMP) and the Pareto distribution with the two data of cement industry collected from the Siam Cement (Ta Luang) company limited, Khao Wong plant, which are: 1) the outlet hot air temperature of the vertical roller mill (°C) in 30 min (Table 1), and 2) the electrical current of kiln drive (ampere) in 24 h (Table 2). Criterions used for these distributions are the Akaike information criterion (AIC).

Table 1: The outlet hot air temperature of the vertical roller mill (VRM) in 30 min (°C)

73.11	73.11	73.94	73.34	73.69
73.83	74.16	74.53	74.19	73.48
73.46	73.33	72.79	72.16	72.41
72.98	73.94	73.06	71.7	71.41
71.68	72.37	74.22	78.56	78.57
78.08	77.82	76.67	74.07	73.02

Table 1 shows the hot temperature that leaves the vertical roller mill (VRM), degree Celsius, measured every one minute each half-hour (°C/min) (Figure 4).

In the first step, we fit the NMP distribution to the outlet hot air temperature of the vertical roller mill (VMR) data and compare the fitness with the MP distribution and the Pareto distribution. In order to compare distribution, we estimate parameters of those distributions by using the "optim" function (Nelder and Mead method) and consider the AIC statistics for the outlet hot air temperature of the VMR dataset. The MLE estimates of the parameters, the AIC measure for the fitted models are shown in Table 2.

Table 2: The MLE of the model parameters for the outlet hot air temperature of VRM, AIC measure

Fitting Dist.	Estimate Parameters			AIC
	α	β	ϖ	
Pareto	28.45	71.41	-	121.3157
MP	28.4531	71.41	0.000059	123.3157
NMP	28.4875	71.41	-	121.3162

The results indicate that the AIC statistics of the NMP distribution and the Pareto distribution are smaller than the MP distribution. Therefore, the NMP distribution and the Pareto distribution are better fitted for the outlet hot air temperature of VMR data than MP distribution.

Table 3: The electrical current of kiln drive (ampere) in 24 h

1425.28	1359.15	1410.95	1430.5
1410.01	1427.4	1514.7	1541.85
1605.44	1653.2	1686.01	1647.7
1561.11	1537.12	1461.71	1386.61
1421.21	1454.45	1504.6	1520.14
1500.23	1487.95	1453.38	1408.39

Table 3 shown the current used to drive the kiln, measure every 1 h (Amp/h) (Figure 5).

In the second step, we fit those three distributions to the electrical current of the kiln drive (ampere) in 24 h of data. The MLE estimates of the parameters, the AIC measure for the electrical current of the kiln drive dataset are shown in Table 3.

Regarding the results in Table 4, it was observed that the NMP distribution and the Pareto distribution fits the data better than the MP distribution. Hence, we have positive evidence that the NMP distribution is superior to the MP distribution for two example data in the cement industry.

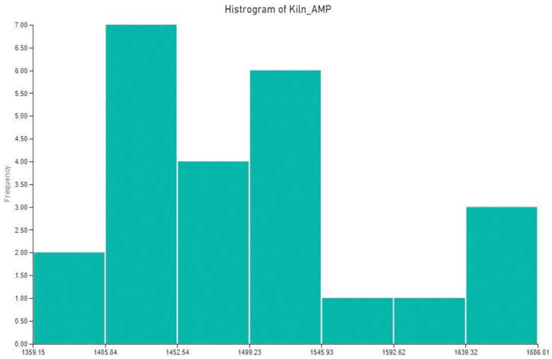

Figure 5: Histogram of the electrical current of kiln drive (ampere).

Table 4: The MLE of the model parameters for the electrical current of kiln drive, AIC measure

Fitting Dist.	Estimate Parameters			AIC
	α	β	ϖ	
Pareto	10.9125	1359.15	-	287.9769
MP	10.9158	1359.15	0.000019	289.9769
NMP	11.0125	1359.15	-	287.9870

4.2 *Hard drive failure data*

This subsection is an example of a lifetime dataset. The dataset collected from the BackBlaze dataset [9], contains only model ST8000DM002 in December 2017 (Table 5) and (Figure 6).

Table 5: The lifetime and serial number of the collected hard disk (days)

Serial_Number	Lifetime	Serial_Number	Lifetime
ZA11RN01	490	ZA13RJA4	319
ZA11X9QN	497	ZA1261SY	431
ZA11VZVN	521	ZA12ME03	484
ZA13Z5S6	394	ZA10NDSM	547
ZA11M7BK	489	ZA136W64	383
ZA13Z8H4	323	ZA123ZWK	534
ZA13R2P0	376	ZA13R3T3	316

In order to fit the hard drive failure data to an NMP distribution, we estimated the unknown parameters with the maximum likelihood estimation by using the "optim" function (Nelder and Mead method) in the R program. The MLE estimate of the parameter β is min x_i = 316 and the MLE estimate of the parameter α is 3.575.

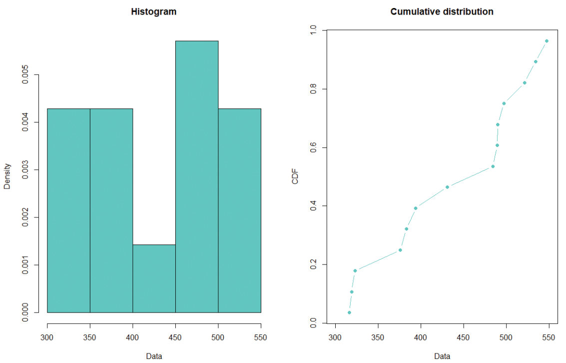

Figure 6: Histogram and cumulative distribution of the lifetime of the collected hard disk (days).

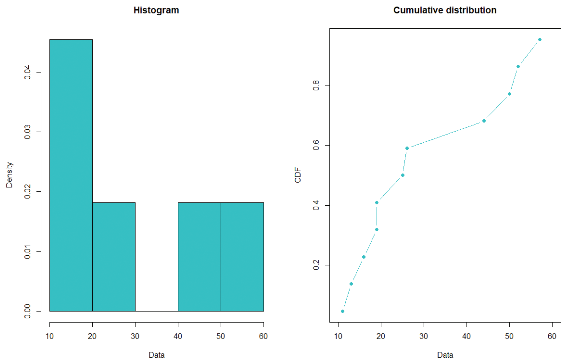

Figure 7: Histogram and cumulative distribution of the number of months patient survived.

Next, we fit the hard drive failure data to an NMP distribution by considering the Kolmogorov-Smirnov test (KS test). The value of the KS test is equal to 0.1854609 and the critical value is equal to 0.3629675. So the hard drive failure data follows the NMP distribution. Then the average of a lifetime of the hard disk model ST8000DM002 is

$$E(X)=\left(\frac{1}{3.575+1}\right)\times\left[\frac{3.575^2(316)}{3.575-1}+\frac{(3.575-1)(316)}{3.575-2}\right]$$

$= 455.7494$ days.

And the probability of the lifetime of the hard disk model ST8000DM002 will more than one year is

$$S(365)=\left(\frac{1}{3.575+1}\right)\times\left[3.575\left(\frac{365}{316}\right)^{-3.575}+\left(\frac{365}{316}\right)^{-(3.575-1)}\right]$$

$= 0.617533.$

4.3 *Heart attack data*

The heart attack data is an example of a survival time dataset, collected from www.kaggle.com [10]. In this research, we interested in group one and pericardial effusion is no fluid around the heart. All the patients suffered heart attacks at some point in the past and they are dead at the end of the survival period.

In order to fit the NMP distribution to the number of months the patient survived, we eliminate missing values and consider the KS test. Also, we estimated parameters by using the maximum likelihood estimation via the Nelder and Mead method. The MLE estimate of the parameter β is min $x_i = 11$ and the MLE estimate of the parameter α is 1.744141. So we get the value of the KS test is 0.1487127 and the critical value is 0.4094826. Hence it indicates that the NMP distribution fits this data (Table 6) and (Figure 7).

Table 6: The number of months patient survived

11	19	16	57	26	13	50	19
25	52	44					

Therefore the probability of the patient who in group one and pericardial effusion is no fluid around the heart will survive more than one year is

$$S(12)=\left(\frac{1}{1.744141+1}\right)\times\left[1.744141\left(\frac{12}{11}\right)^{-1.744141}+\left(\frac{12}{11}\right)^{-(1.744141-1)}\right]$$

$= 0.8876582.$

5 Conclusions

This research presented the new distribution by adjusting the new weighted parameter to the mixture Pareto distribution based on a mixture of Pareto distribution according to Nanuwong *et al*. [6]. Furthermore, we have studied various properties of the proposed distribution such as the probability density function, the cumulative function, the survival function, the hazard rate function, and the r^{th} moment. We have also estimated the estimators with the maximum likelihood estimation of the proposed distribution, as well as presenting the analysis of the comparison of the proposed distribution with the mixture of Pareto distribution and the Pareto distribution with the real data. The criteria used for these distributions are the Akaike information criterion (AIC). Also, we give an example of a descriptive data analysis based on NMP distribution. The results revealed that the proposed distribution is suitable for real data which has a small sample size and a small range of data.

Acknowledgments

This research was funded by King Mongkut's University of Technology North Bangkok. Contract no. KMUTNB-62-DRIVE-04. Also, we would like to express my sincere thanks to the Siam Cement (Ta Luang) Co.,Ltd. Khao Wong Plant for support data.

References

[1] H. W. Hager and L. J. Bain, "Inferential procedures for the generalized gamma distribution," *Journal of the American Statistical Association*, vol. 65, no. 332, pp. 1601–1609, 1970.

[2] A. V. Boyd, "Fitting the truncated Pareto distribution to loss distributions," *Journal of the Staple Inn Actuarial Society*, vol. 31, pp. 151–158, 1988.

[3] V. Brazauskas and S. Robert, "Favorable estimator for fitting Pareto models: A study using goodness-of-fit measures with actual data," *ASTIN Bulletin*, vol. 33, no. 2, pp. 365–381, 2003.

[4] M. Doostparast and N. Balakrishnan, "Pareto analysis based on records, *Statistics*, vol. 47, no. 5, pp. 1075–1089, 2013.

[5] G. P. Patil and C. R. Rao, "Weighted distributions and size-biased sampling with applications to wildlife populations and human families," *Biometrics*, vol. 34, no. 2, pp. 179–189, 1978.

[6] N. Nanuwong, W. Bodhisuwan, and C. Pudprommarat, "A new mixture Pareto distribution and its application," *Thailand Statistician*, vol. 13, no. 2, pp. 191–207, 2015.

[7] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.

[8] The R Development Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018.

[9] Backblaze, "The Raw Hard Drive Test Data," 2002. [Online]. Available: https://www.backblaze.com/b2/hard-drive-test-data.html

[10] R. Bhaskaran, "Echocardiogram-UCI," 2018. [Online]. Available: https://www.kaggle.com/loganalive/echocardiogram-uci