Applied Mathematics & Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/amisl/060103

On Three Parameter Weighted Pareto Type II Distribution: Properties and Applications in Medical Sciences

B. A. Para and T. R. Jan*

Department of Statistics, University of Kashmir, Srinagar, India

Received: 14 Jul. 2017, Revised: 27 Sep. 2017, Accepted: 06 Oct. 2017.
Published online: 1 Jan. 2018.

Abstract: The concept of weighted distributions can be employed in the development of proper models for lifetime data in medical sciences and other fields. In this paper, we have introduced a new generalization of Pareto type II distribution using the concept of weighting. The statistical properties of this distribution are derived and the model parameters are estimated by maximum likelihood estimation along with Monte Carlo simulation procedure. Finally, an application to real data set is finally presented for illustration in medical sciences.

Keywords: Weibull Weighted Probability Models, Size biased, Pareto II distribution, Reliability Analysis, Maximum likelihood estimation, Simulation, Medical Sciences.

1 Introduction

In the last few decades, there has been a growing interest in the construction of generalized flexible parametric classes of probability models in medical sciences and other fields. Various forms of the distributions in medical sciences have appeared in the literature for data analysis and modeling. The quality of the procedures used in a statistical analysis depends heavily on the assumed probability model or distributions. Because of this, considerable effort over the years has been expended in the development of large classes of standard distributions along with relevant statistical methodologies, designed to serve as models for a wide range of real world phenomena. However, there still remain many important problems where the real data does not follow any of the classical or standard models. So to overcome such requirements, we use to develop some new models. These newly developed classes of distributions provide greater flexibility in modeling complex data from medical sciences and the results drawn from them seems quite sound and genuine. Thus our main concern becomes, to give importance especially to model specification and the data interpretation. Weighted probability models play very important role in some situations arising in various practical fields like medical sciences, engineering etc. These distributions arise in practice when observations from a sample are recorded with unequal probability and provide unifying approach for the problems when the observations fall in the non-experimental, non-replicated and non-random categories. Van Gove (2003) reviewed some of the more recent results on weighted distributions pertaining to parameter estimation in forestry. Warren (1975) was the first to apply the weighted distributions in connection with sampling wood cells. Patil and Ord (1978) introduced the concept of size-biased sampling and weighted distributions by identifying some of the situations where the underlying models retain their form. The statistical interpretation of weighted and size-biased distributions was originally identified by Cox (1964) in the context of renewal theory. Size-biased sampling situations may occur in clinical trials, reliability theory, and survival analysis and population studies, where a proper sampling frame is absent. In such situations, items are sampled at a rate proportional to their length, so that larger values of the quantity being measured are sampled with higher probabilities. Numerous works on various aspects of weighted and size-biased sampling are available in literature which include family size and sex ratio , wild life population and line transect sampling , analysis of family data, cell cycle analysis , efficacy of early screening for disease, aerial survey and visibility bias Patil and Rao (1978).

The probability density function (pdf) of the Pareto type II distribution (PTIID) is given by

*Corresponding author e-mail: parabilal@gmail.com

$$f(x;\alpha,\lambda)=\frac{\alpha}{\lambda}\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)} \qquad ; \quad x>0,\alpha,\lambda>0 \qquad (1.1)$$

The rth Moment, Mean and variance of Pareto type II distribution is given by

$$\mu_r{}' = \frac{\lambda^r \Gamma(\alpha-r)\Gamma(1+r)}{\Gamma(\alpha)} \qquad \alpha>r,\lambda>0 \qquad (1.2)$$

$$\mu_1{}' = \frac{\lambda}{(\alpha-1)} \qquad (1.3)$$

$$\mu_2 = \frac{\lambda^2\alpha}{(\alpha-1)^2\,(\alpha-2)}$$

2 Weighted Pareto Type II Distribution

Suppose X is a non negative random variable with probability density function (pdf) $f(x)$. Let $W(x)$ be the weight function which is a non negative function, then the probability density function of the weighted random variable X_w is given by:

$$f_w(x)=\frac{w(x)f(x)}{E(w(x))}\ ,\ x>0,$$

where $w(x)$ is a non-negative weight function and $E(w(x))=\int w(x)f(x)dx < \infty$.

Depending upon the choice of the weight function w(x), we have different weighted models. Clearly when $w(x)=x$, the resulting distribution is called size biased whose pdf is given by:

$$f_{SB}(x)=\frac{xf(x)}{E(x)}\ ,\ x>0$$

Weighted distributions occur frequently in research related to reliability, bio-medicine, ecology and branching process and can be seen in Patil and Rao (1986). Das and Roy (2011) discussed the size biased weighted Generalized Rayleigh distribution with its properties, also they developed the size biased weighted Weibull distribution.

In this paper, we have considered the weight function as $w(x)=x^c$ to obtain the weighted Pareto type II model. The probability density of weighted Pareto type II distribution is given as:

$$f_w(x;\alpha,\lambda,c)=\frac{x^c f(x;\ \alpha,\lambda)}{E[x^c]}, x>0,\lambda>0,c>0,\alpha>c \qquad (1.4)$$

Using (1.1) and (1.2) in (1.4), we have new generalization of Pareto type II distribution called weighted Pareto type II model (WPTII) given by

$$f_w(x;\alpha,\lambda,c)=\frac{x^c\Gamma(\alpha+1)\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}}{\lambda^{c+1}\Gamma(\alpha-c)\Gamma(1+c)}, \qquad x>0,\lambda>0,c>0,\alpha>c \qquad (1.5)$$

CDF of weighted Pareto type II distribution is given by

$$F_w(x)=\int_0^x \frac{x^c\Gamma(\alpha+1)\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}}{\lambda^{c+1}\Gamma(\alpha-c)\Gamma(1+c)}dx,\ x>0,\lambda>0,c>0,\alpha>c$$

$$= \frac{\Gamma(\alpha+1)}{\lambda^c \Gamma(\alpha-c)\Gamma(1+c)} B\left(\frac{x}{\lambda}; c+1, \alpha-c\right), \; x>0, \lambda>0, c>0, \alpha>c, \qquad (1.6)$$

Where $B(x;a,b)$ is an incomplete beta function.

Graphical overview of weighted Pareto type II distribution for different parameter combinations is given in fig. 1.1 to fig. 1.4. The newly introduced generalized distribution has right skewed nature.

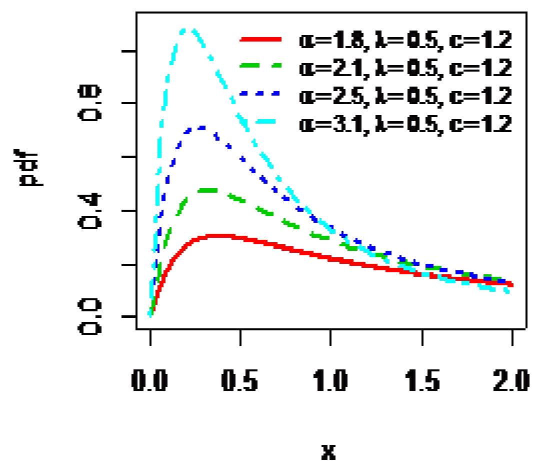

Fig 1.1 pdf plot for WPTII(α , λ ,c)

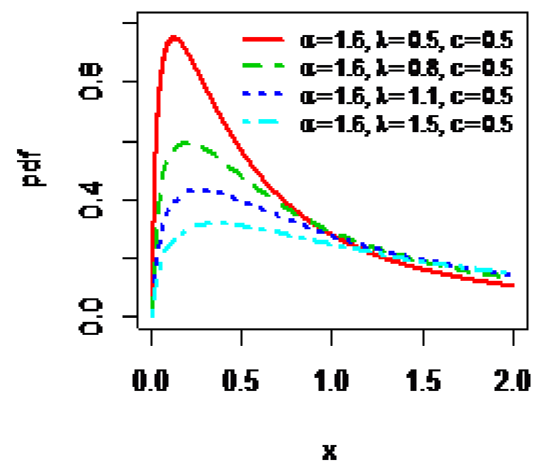

Fig 1.2 pdf plot for WPTII(α , λ ,c)

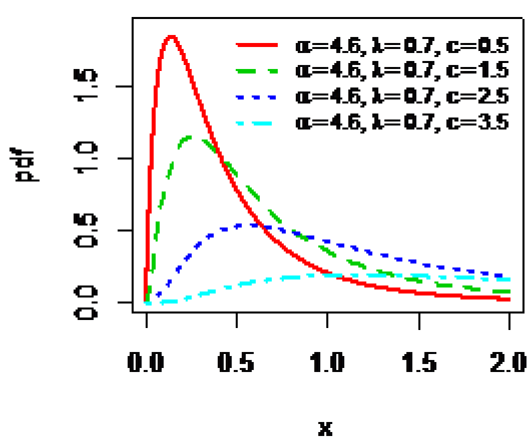

Fig 1.3 pdf plot for WPTII (α , λ ,c)

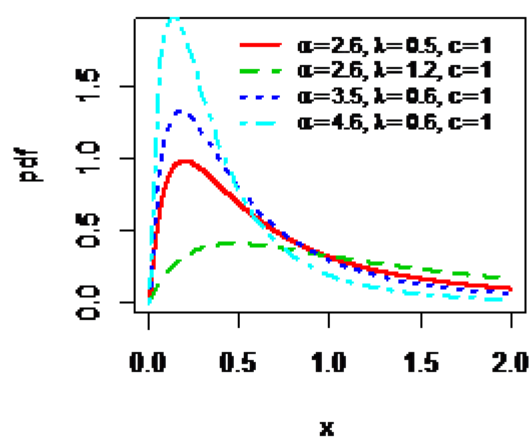

Fig 1.4 pdf plot for WPTII(α , λ ,c)

A Size biased Pareto type II distribution (SBPIID) is obtained by applying the weights x^c, where c =1 to the weighted Pareto type II distribution. We have from relation equations (1.1) and (1.3)

$$\mu_1' = \int_0^\infty x\, f(x;\alpha,\lambda)dx = \frac{\lambda}{(\alpha-1)}$$

$$\int_0^\infty \frac{\alpha(\alpha-1)}{\lambda^2} x\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)} dx = 1$$

This gives the size biased Pareto type II distribution as

$$f_{SB}(x;\alpha,\lambda) = \frac{\alpha(\alpha-1)}{\lambda^2} x\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)} \quad ; \quad x>0, \alpha>1, \lambda>0 \tag{1.7}$$

and the corresponding cdf is given by

$$F_{SB}(x;\alpha,\lambda) = 1-\left(1+\frac{x}{\lambda}\right)^{-\alpha}\left(1+\frac{\alpha x}{\lambda}\right) \tag{1.8}$$

where α and λ are shape and scale parameters, respectively. The pdf and CDF plot for different parameter combinations for size biased Pareto type II distribution is given in fig. 1.5 to fig. 1.8.

3 Reliability Analysis WPTII and SBPTII Distributions

In this sub section, we have obtained the reliability and hazard rate functions of the weighted Pareto type II model.

3.1 Reliability Function of WPTII Model.

The reliability function is defined as the probability that a system survives beyond a specified time. It is also referred to as survival or survivor function of the distribution. It can be computed as complement of the cumulative distribution function of the model. The reliability function or the survival function of weighted Pareto type II distribution is calculated as:

$$R_w(x;\alpha,\lambda,c) = 1-\frac{\Gamma(\alpha+1)}{\lambda^c\Gamma(\alpha-c)\Gamma(1+c)} B\left(\frac{x}{\lambda};c+1,\alpha-c\right),\ x>0,\ \lambda>0, c>0, \alpha>c$$

where $B(x;a,b)$ is an incomplete beta function.

3.2 Hazard Function WPTII Model

The hazard function is also known as hazard rate or instantaneous failure rate is given as:

$$h_w(x;\alpha,\lambda,c) = \frac{f_w(x;\alpha,\lambda,c)}{R_w(x;\alpha,\lambda,c)} = \frac{x^c\Gamma(\alpha+1)\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}}{\lambda^c\Gamma(\alpha-c)\Gamma(1+c)-\Gamma(\alpha+1)B\left(\frac{x}{\lambda};c+1,\alpha-c\right)}.,$$

$x>0,\ \lambda>0, c>0, \alpha>c,$

Where $B(x;a,b)$ is an incomplete beta function

3.3 Reliability Function of SBPTII Model

The reliability function of size biased Pareto type II distribution is given by

$$R_{SB}(x;\alpha,\lambda) = 1-F_{SB}(x;\alpha,\lambda)$$

$$R_{SB}(x;\alpha,\lambda) = \left(1+\frac{x}{\lambda}\right)^{-\alpha}\left(1+\frac{\alpha x}{\lambda}\right) \quad ; \quad x>0, \alpha>1, \lambda>0$$

3.4 Hazard Function of SBPTII Model

The hazard function of size biased Pareto type II distribution is given by

$$h_{SB}(x;\alpha,\lambda)=\frac{f_{SB}(x;\alpha,\lambda)}{R_{SB}(x;\alpha,\lambda)}$$

$$h_{SB}(x;\alpha,\lambda)=\frac{\alpha(\alpha-1)}{\lambda^2}x\left(1+\frac{x}{\lambda}\right)^{-1}\left(1+\frac{\alpha x}{\lambda}\right)^{-1}\quad;\quad x>0,\alpha>1,\lambda>0$$

3.5 Hazard Function SBPTII Model

The reverse hazard function for the Size biased weighted Pareto type II distribution is given as

$$h_{SB}^{r}(x)=\frac{\frac{\alpha(\alpha-1)}{\lambda^2}x\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}}{1-\left(1+\frac{x}{\lambda}\right)^{-\alpha}\left(1+\frac{\alpha x}{\lambda}\right)}\quad;\quad x>0,\alpha>1,\lambda>0$$

4 Statistical Properties of WPTII and SBPTII Distributions

In this section we shall discuss structural properties of weighted Pareto type II and size biased Pareto type II distributions. Specially moments, order statistics, maximum likelihood estimation, and moment generating function.

4.1 Moments

Suppose X denote the weighted Pareto type II distribution random variable with parameters α , λ and c then

$$E(X^r)=\mu_r'=\int_0^\infty x^r f_w(x;\alpha,\lambda,c)\,dx$$

$$=\frac{\Gamma(\alpha+1)}{\Gamma(\alpha-c)\Gamma(1+c)}\lambda^r\sum_{k=0}^{r+c}{}^{r+c}C_k(-1)^{k+1}\left(\frac{1}{r+c-k-\alpha}\right)\qquad(1.9)$$

For Size Biased Pareto type II random variable with parameters α and λ, rth moment can be directly obtained by substituting c=1 in the rth moment expression (1.9) of weighted Pareto type II distribution.

$$\mu_r'=\alpha(\alpha-1)\lambda^r\sum_{k=0}^{r+1}(-1)^{k+1}\,{}^{r+1}C_k\left(\frac{1}{r-\alpha-k+1}\right)$$

Substitute r=1, 2, 3, 4, in the rth moment expression of SBPTII distribution, we get first four moments for SBPTII distribution.

Mean = $\mu_1'=\dfrac{2\lambda}{(\alpha-2)}$

$$\mu_2'=\frac{6\lambda^2}{(\alpha-2)(\alpha-3)}$$

Variance of SBPTII distribution = $\mu_2=\dfrac{2\lambda^2\alpha}{(\alpha-2)^2(\alpha-3)}$

Standard deviation of SBPTII distribution, $\sigma=\dfrac{\lambda}{(\alpha-2)}\sqrt{\dfrac{2\alpha}{(\alpha-3)}}$

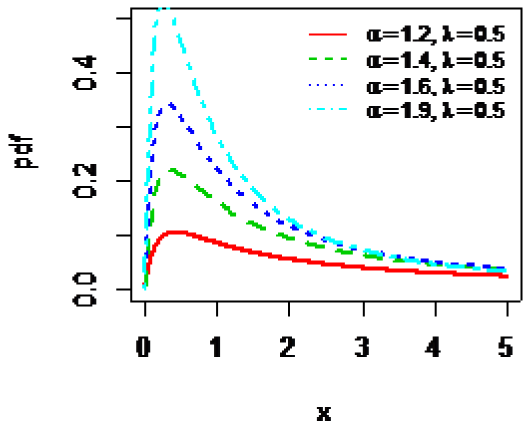

Fig 1.5 pdf plot for SBPII (α , λ)

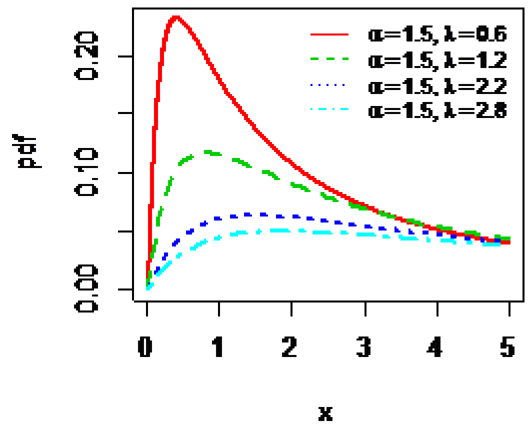

Fig 1.6 pdf plot for SBPII (α , λ)

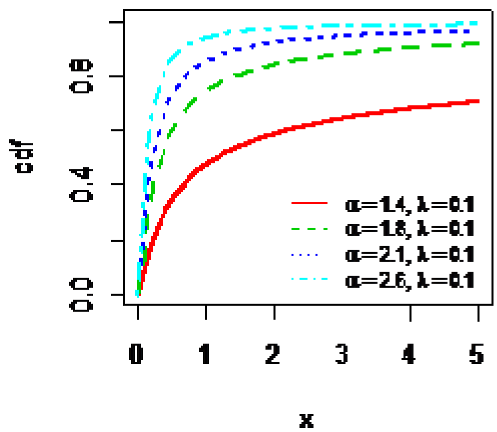

Fig 1.7 pdf plot for SBPII(α , λ)

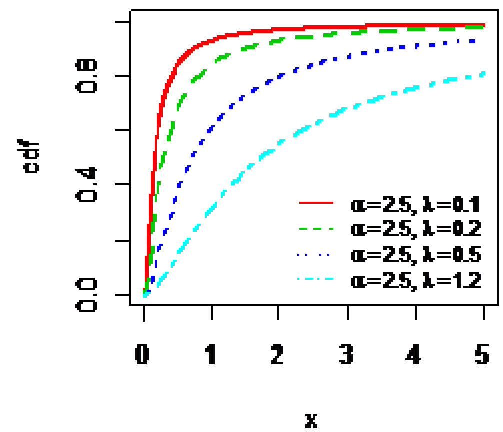

Fig 1.8 pdf plot for SBPII(α , λ)

Coefficient of variation of SBPTII distribution, $C.V = \frac{\sigma}{\mu} = \frac{1}{2}\sqrt{\frac{2\alpha}{(\alpha-3)}}$

4.2 Moment Generating Function of WPTII and SBPTII Distributions

In this sub section we derived the moment generating function of WPTII. We begin with the well known definition of the moment generating function given by

$$M_X(t) = E\left(e^{tx}\right) = \int_0^\infty e^{tx}\ f_w(x;\alpha,\lambda,c)\,dx$$

$$= \int_0^\infty \left[1 + t\,x + \frac{(t\,x)^2}{2!} + ...\right] f_w(x;\alpha,\lambda,c)\,dx$$

$$= \sum_{j=0}^{\infty} \frac{t^j}{j!}\ \mu_j{}'$$

$$\Rightarrow M_X(t) = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha-c)\Gamma(1+c)} \sum_{j=0}^{\infty} \sum_{k=0}^{j+c} (-1)^{k+1}\ \frac{t^j}{j!} \lambda^j\ {}^{j+c}C_k \left(\frac{1}{j+c-k-\alpha}\right) \qquad (1.10)$$

The moment generating function of SBPTII distribution can be directly obtained from expression (1.10) by putting c=1 as

$$M_X(t)=\alpha(\alpha-1)\sum_{j=0}^{\infty}\sum_{k=0}^{j+1}(-1)^{k+1}\frac{t^j}{j!}\lambda^j\ {}^{j+1}C_k\left(\frac{1}{j-\alpha-k+1}\right)$$

5 Order Statistics

Order statistics make their appearance in many statistical theory and practice. We know that if $X_{(1)}, X_{(2)},\ldots, X_{(n)}$ denotes the order statistics of a random sample $X_1, X_2,\ldots, X_n$ from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$, then the pdf of rth order statistics X(r) is given by

$$f^{w}_{X(r)}(x)=\frac{n!}{(r-1)!(n-r)!}f_X(x)\left(F_X(x)\right)^{r-1}\left(1-F_X(x)\right)^{n-r}$$

For r = 1, 2, . . . ,n.

We have from (1.5) and (1.6) the expression of the rth order WPTII random variable $X_{(r)}$ given by

$$g^{w}_{X(r)}(x)=\frac{n!}{(r-1)!(n-r)!}\frac{x^c\Gamma(\alpha+1)\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}}{\lambda^{c+1}\Gamma(\alpha-c)\Gamma(1+c)}\left(\frac{\Gamma(\alpha+1)}{\lambda^c\Gamma(\alpha-c)\Gamma(1+c)}B\left(\frac{x}{\lambda};c+1,\alpha-c\right)\right)^{r-1}\left(1-\frac{\Gamma(\alpha+1)}{\lambda^c\Gamma(\alpha-c)\Gamma(1+c)}B\left(\frac{x}{\lambda};c+1,\alpha-c\right)\right)^{n-r}$$

Therefore, the expression of the nth order WPTII statistic $X_{(n)}$ is given by

$$g^{w}_{X(n)}(x)=\frac{nx^c\Gamma(\alpha+1)\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}}{\lambda^{c+1}\Gamma(\alpha-c)\Gamma(1+c)}\left(\frac{\Gamma(\alpha+1)}{\lambda^c\Gamma(\alpha-c)\Gamma(1+c)}B\left(\frac{x}{\lambda};c+1,\alpha-c\right)\right)^{n-1}$$

and the expression of the first order WPTII statistic $X_{(1)}$ is given by

$$g^{w}_{X(1)}(x)=\frac{nx^c\Gamma(\alpha+1)\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}}{\lambda^{c+1}\Gamma(\alpha-c)\Gamma(1+c)}\left(1-\frac{\Gamma(\alpha+1)}{\lambda^c\Gamma(\alpha-c)\Gamma(1+c)}B\left(\frac{x}{\lambda};c+1,\alpha-c\right)\right)^{n-1}$$

Since we have cdf of size biased Pareto type II distribution in closed form, therefore we have order statistics expression for size biased Pareto type II distribution in closed form as well.

we have from (1.7) and (1.8) , the expression of the rth order statistics for SBPII random variable $X_{(r)}$given by

$$g^{sb}_{X(r)}(x)=\frac{n!}{(r-1)!(n-r)!}\frac{\alpha(\alpha-1)}{\lambda^2}x\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}\left(\left(1+\frac{x}{\lambda}\right)^{-\alpha}\left(1+\frac{\alpha x}{\lambda}\right)\right)^{r-1}\left(1-\left(1+\frac{x}{\lambda}\right)^{-\alpha}\left(1+\frac{\alpha x}{\lambda}\right)\right)^{n-r}$$

Therefore, the expression of the nth order statistic $X_{(n)}$for SBPTII distribution is given by

$$g^{sb}_{X(n)}(x)(x)=\frac{n\alpha(\alpha-1)}{\lambda^2}x\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}\left(\left(1+\frac{x}{\lambda}\right)^{-\alpha}\left(1+\frac{\alpha x}{\lambda}\right)\right)^{n-1}$$

and the expression of the first order statistic $X_{(1)}$for SBPTII distribution is given by

$$g^{sb}_{X_{(1)}}(x)=\frac{n\alpha(\alpha-1)}{\lambda^2}x\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)}\left(1-\left(1+\frac{x}{\lambda}\right)^{-\alpha}\left(1+\frac{x\alpha}{\lambda}\right)\right)^{n-1}$$

5.1 Method of Maximum Likelihood Estimation.

Maximum likelihood estimation has been the most widely used method for estimating the parameters of the probability distributions. Let $x_1, x_2, x_3.......x_n$ be a random sample from the weighted Pareto type II distribution, then the corresponding likelihood function is given as

$$L_w(x;\alpha,\lambda,c)=\frac{(\Gamma(\alpha+1))^n}{\lambda^{n(c+1)}(\Gamma(\alpha-c))^n(\Gamma(c+1))^n}\prod_{i=1}^{n}x_i^c\prod_{i=1}^{n}\left(1+\frac{x_i}{\lambda}\right)^{-(\alpha+1)}$$

The log-likelihood function is given as:

$$\log L_w(x;\alpha,\lambda)=n\log(\Gamma(\alpha+1))-n(c+1)\log(\lambda)-n\log(\Gamma(\alpha-c))-n\log(\Gamma(c+1))$$
$$+c\sum_{i=1}^{n}\log x_i-(\alpha+1)\sum_{i=1}^{n}\log\left(1+\frac{x_i}{\lambda}\right) \qquad (1.11)$$

Now, differentiating equation () with respect to α,λ and c, we obtain the normal equations

$$\frac{\partial\log L_w(x;\alpha,\lambda)}{\partial\alpha}=\frac{n}{\Gamma(\alpha+1)}\frac{\partial}{\partial\alpha}(\Gamma(\alpha+1))-\frac{n}{\Gamma(\alpha-c)}\frac{\partial}{\partial\alpha}(\Gamma(\alpha-c))-\sum_{i=1}^{n}\log\left(1+\frac{x_i}{\lambda}\right) \qquad (1.12)$$

$$\frac{\partial\log L_w(x;\alpha,\lambda)}{\partial\lambda}=-\frac{n(c+1)}{\lambda}-\frac{(\alpha+1)}{\lambda^2}\sum_{i=1}^{n}\frac{x_i}{\left(1+\frac{x_i}{\lambda}\right)} \qquad (1.13)$$

$$\frac{\partial\log L_w(x;\alpha,\lambda)}{\partial c}=n\log\lambda-\frac{n}{\Gamma(\alpha-c)}\frac{\partial}{\partial c}(\Gamma(\alpha-c))-\frac{n}{\Gamma(c+1)}\frac{\partial}{\partial c}(\Gamma(c+1))+\sum_{i=1}^{n}\log x_i \qquad (1.14)$$

The MLE $\hat{\eta}_w=(\hat{\alpha},\hat{\lambda},\hat{c})$ of $\eta_w=(\alpha,\lambda,c)$ is obtained by solving this nonlinear system of equations (1.12), (1.13) and (1.14). It is usually more convenient to use nonlinear optimization algorithms such as quasi-Newton algorithm to numerically maximize the log likelihood function given in (1.11). Applying the usual large sample approximation, the MLE $\hat{\eta}_w$ can be treated as being approximately normal with variance-covariance matrix equal to the inverse of the expected information matrix, i.e.

$$\sqrt{n}\,(\hat{\eta}_w-\eta_w)\rightarrow N\left(0,n\,I^{-1}(\eta_w)\right)$$

where $I^{-1}(\eta_w)$ is the limiting variance-covariance matrix of $\hat{\eta}_w$.

The elements of 3×3 Fisher Information matrix is given below

$$I^{-1}(\eta_w)=\begin{bmatrix}-E\left(\frac{\partial^2}{\partial\alpha^2}\log f_l(x;\alpha,\lambda)\right) & -E\left(\frac{\partial^2}{\partial\alpha\,\partial\lambda}\log f_l(x;\alpha,\lambda)\right) & -E\left(\frac{\partial^2}{\partial\alpha\,\partial c}\log f_l(x;\alpha,\lambda)\right)\\ -E\left(\frac{\partial^2}{\partial\lambda\,\partial\alpha}\log f_l(x;\alpha,\lambda)\right) & -E\left(\frac{\partial^2}{\partial\lambda^2}\log f_l(x;\alpha,\lambda)\right) & -E\left(\frac{\partial^2}{\partial\lambda\,\partial c}\log f_l(x;\alpha,\lambda)\right)\\ -E\left(\frac{\partial^2}{\partial c\,\partial\alpha}\log f_l(x;\alpha,\lambda)\right) & -E\left(\frac{\partial^2}{\partial c\,\partial\lambda}\log f_l(x;\alpha,\lambda)\right) & -E\left(\frac{\partial^2}{\partial c^2}\log f_l(x;\alpha,\lambda)\right)\end{bmatrix}$$

In case of SBPTII distribution, we have likelihood function given by

$$L_{SB}(x;\alpha,\lambda)=\frac{\alpha^n(\alpha-1)^n}{\lambda^{2n}}\prod_{i=1}^{n}x_i\prod_{i=1}^{n}\left(1+\frac{x_i}{\lambda}\right)^{-(\alpha+1)}$$

The log-likelihood function is given as:

$$\log L_{SB}(x;\alpha,\lambda)=n\log\alpha+n\log(\alpha-1)-2n\log\lambda+\sum_{i=1}^{n}\log x_i-(\alpha+1)\sum_{i=1}^{n}\log\left(1+\frac{x_i}{\lambda}\right) \tag{1.15}$$

Now, differentiating equation (1.15) with respect to parameters α and λ, we obtain the normal equations

$$\frac{n(2\alpha-1)}{\alpha(\alpha-1)}=\sum_{i=1}^{n}\log\left(1+\frac{x_i}{\lambda}\right) \tag{1.16}$$

$$\frac{2n}{\lambda}=\frac{(\alpha+1)}{\lambda^2}\sum_{i=1}^{n}\frac{x_i}{\left(1+\frac{x_i}{\lambda}\right)} \tag{1.17}$$

The MLE $\hat{\eta}_{SB}=(\hat{\alpha},\hat{\lambda})$ of $\eta_{SB}=(\alpha,\lambda)$ is obtained by solving this nonlinear system of equations. It is usually more convenient to use nonlinear optimization algorithms such as quasi-Newton algorithm to numerically maximize the log likelihood function given in (16). Applying the usual large sample approximation, the MLE $\hat{\eta}$ can be treated as being approximately bivariate normal with variance-covariance matrix equal to the inverse of the expected information matrix, i.e.

$$\sqrt{n}\,(\hat{\eta}_{SB}-\eta_{SB})\rightarrow N\left(0,nI^{-1}(\eta)\right)$$

Where $I^{-1}(\eta_{SB})$ is the limiting variance-covariance matrix of $\hat{\eta}_{SB}$.
The elements of 2×2 Fisher Information matrix is given below

$$I_{SB}(\alpha,\lambda)=\begin{bmatrix}-E\left(\frac{\partial^2}{\partial\alpha^2}\log f_l(x;\alpha,\lambda)\right) & -E\left(\frac{\partial^2}{\partial\alpha\,\partial\lambda}\log f_l(x;\alpha,\lambda)\right)\\ -E\left(\frac{\partial^2}{\partial\lambda\,\partial\alpha}\log f_l(x;\alpha,\lambda)\right) & -E\left(\frac{\partial^2}{\partial\lambda^2}\log f_l(x;\alpha,\lambda)\right)\end{bmatrix}$$

$$=\begin{bmatrix}I_{11} & I_{12}\\ I_{21} & I_{22}\end{bmatrix}$$

$$I_{11}=-E\left(\frac{\partial^2}{\partial\alpha^2}\log L(x;\alpha,\lambda)\right)=\frac{2\alpha^2-2\alpha+1}{\alpha^2(\alpha-1)^2}$$

$$I_{12}=I_{21}=-E\left(\frac{\partial^2}{\partial\alpha\,\partial\lambda}\log L(x;\alpha,\lambda)\right)=\frac{-2(\alpha-2)}{\alpha(\alpha-1)}$$

$$I_{22}=E\left(\frac{\partial^2}{\partial\lambda^2}\log L(x;\alpha,\lambda)\right)=-\frac{2}{\lambda^2}+\frac{(\alpha+1)}{\lambda^2}\frac{(4\alpha+2)}{\alpha(\alpha-1)}$$

5.2 Monte Carlo Simulation Procedure for ML Estimates of SBPII distribution

In this section, we investigate the behavior of the ML estimators for a finite sample size n. Simulation study based on different $SBPTII(x,\alpha,\lambda)$ is carried out. The random observations are generated by using the inverse cdf method presented in section 3.4 from SBPTII(α,λ). Monte Carlo simulation study was carried out for four parameter combinations as (θ=2.5, λ=0.1), (θ=2.7, λ=0.5), (θ=2.3, λ=0.2) and (θ=2.6, λ=0.3) .The process was repeated 1000 times by taking different sample sizes n = (25,50,75,100,150,200,300,500). We observe in table 1.1 that the agreement between theory and practice improves as the sample size n increases. MSE and Variance of the estimators suggest us that the estimators are consistent and the maximum likelihood method performs quite well in estimating the model parameters of the proposed distribution.

6 Applications

Example 1. n table1.2 we consider a data set reported by Efron (1988) represent the survival times of a group of patients suffering from Head and Neck cancer disease and treated using a combination of radiotherapy and chemotherapy (RT+CT).
In order to compare the two distribution models, we consider the criteria like AIC (Akaike information criterion (1974)), AICC (corrected Akaike information criterion) and BIC (Bayesian information criterion (1986)). The better distribution corresponds to lesser AIC, AICC and BIC values.

AIC = 2k-2logL, AICC = AIC+ $\frac{2k(k+1)}{n-k-1}$ and BIC = k logn-2logL

where k is the number of parameters in the statistical model, n is the sample size and –logL is the maximized value of the log-likelihood function under the considered model. From Table 1.3, it has been observed that the weighted Pareto type II and size biased Pareto type II distribution have the lesser AIC, -logL and BIC values as compared to Pareto type II Distribution. Hence we can conclude that weighted Pareto type II and size biased Pareto type II distribution leads to a better fit as compared to Pareto type II distribution.

Example 2. Here we consider an uncensored data set corresponding to the remission times (in months) of a random sample of 128 bladder cancer patients. Bladder cancer is a disease in which abnormal cells multiply without control in the bladder. The most common type of bladder cancer recapitulates the normal histology of the urothelium and is known as transitional cell carcinoma. These data were previously studied by Lemonte (2012), Zea et al. (2012) and Lee and Wang (2003). Table1.4 lists the remission times of the bladder cancer patients.

We have fitted the weighted Pareto type II and size biased Pareto type II distribution to the dataset using MLE for parameter estimations of the models and compared the proposed models with Pareto type II distribution. Estimation of parameters and other parts of analysis are carried out in R studio statistical software.

From Table 1.5, it has been observed that weighted Pareto type II and size biased Pareto type II distribution have the lesser AIC, -logL and BIC values as compared to Pareto type II distribution in case of data set corresponding to the remission times (in months) of a random sample of 128 bladder cancer patients.. Hence we can conclude that the weighted Pareto type II and size biased Pareto type II distributions leads to a better fit than the Pareto type II distribution. So weighted Pareto type II and size biased Pareto type II distributions are better models as compared to Pareto type II distribution in case of data set corresponding to the remission times (in months) of a random sample of 128 bladder cancer patients.

From table 1.7, it has been observed that weighted Pareto type II and size biased Pareto type II distribution have the lesser AIC, -logL and BIC values as compared to Pareto type II Distribution in case of data set of Survival times (in months) of patients of melanoma studied by Susarla and Vanryzin (1978). Hence we can conclude that weighted Pareto type II andsize biased Pareto type II distribution leads to a better fit than the Pareto type II distribution.

7 Conclusions

In this paper, we have introduced a new generalization of Pareto type II distribution using the concept of weighting.
The statistical properties of this distribution are derived and the model parameters are estimated by maximum likelihood estimation along with Monte Carlo simulation procedure. Finally, an application to real data set is finally presented for illustration in medical sciences.The application of the weighted Pareto type II and size biased Pareto type II distributions have also been demonstrated with real life examples from medical science. The results are compared with Pareto type II distribution, revealed that the WPTII and SBPTIImodels provides a better fit than the Pareto type II distribution.

Table 1.1 Simulated data for size biased Pareto type II distribution for ML estimates

sample size n	θ=2.5, λ=0.1			θ=2.7, λ=0.5		
	Bias	Variance	MSE	Bias	Variance	MSE
25	0.600579	0.592284	0.952979	0.344433	0.370515	0.489149
	0.035769	0.004367	0.005646	0.08103	0.03207	0.038636
50	0.282992	0.135943	0.216028	0.326387	0.28049	0.387018
	0.021599	0.001994	0.002461	0.144305	0.045631	0.066454
75	0.098617	0.164621	0.174346	0.207626	0.215917	0.259026
	0.009842	0.00175	0.001847	0.066496	0.028579	0.033001
100	0.01138	0.063056	0.063186	0.159797	0.182414	0.207949
	0.004999	0.000829	0.000854	0.079329	0.033553	0.039846
150	0.085046	0.057419	0.064651	0.123347	0.076522	0.091737
	0.006377	0.000444	0.000485	0.050436	0.011825	0.014368
200	0.047186	0.060636	0.062862	0.048678	0.069837	0.072207
	0.005217	0.000627	0.000654	0.018855	0.013152	0.013508
300	0.025181	0.026319	0.026953	0.027002	0.050838	0.051567
	-0.00059	0.000194	0.000194	0.013289	0.00809	0.008266
500	0.002418	0.011087	0.011093	0.034904	0.027641	0.028859
	-0.00048	0.000119	0.000119	0.00973	0.004338	0.004433
	θ=2.3, λ=0.2			θ=2.6, λ=0.3		
25	0.189826	0.335486	0.37152	0.271904	0.703984	0.777916
	0.050167	0.017461	0.019978	0.209723	0.324641	0.368625
50	0.142218	0.101107	0.121333	0.193477	0.240305	0.277739
	0.038076	0.006197	0.007647	0.21445	0.459922	0.505911
75	0.128496	0.145956	0.162467	0.115453	0.229064	0.242394
	0.025173	0.006789	0.007423	0.063089	0.147863	0.151843
100	0.103434	0.068826	0.079525	0.108786	0.156908	0.168742
	0.032452	0.003627	0.00468	0.102331	0.063343	0.073815
150	0.077086	0.075555	0.081497	0.097808	0.100834	0.1104
	0.025934	0.005026	0.005699	0.148577	0.065928	0.088003
200	0.060296	0.036531	0.040167	0.057534	0.045225	0.048535
	0.011341	0.002073	0.002202	0.018355	0.048957	0.049294
300	0.027856	0.017034	0.01781	0.0143	0.046247	0.046451
	0.011665	0.001351	0.001487	0.024201	0.050007	0.050593
500	0.010879	0.014777	0.014895	0.000705	0.014354	0.014354
	0.001649	0.000563	0.000565	0.017471	0.008124	0.00843

Table 1.2 Survival times of a group of patients suffering from Head and Neck cancer disease

12.2	23.6	23.7	25.9	32.0	37.0
41.4	47.4	55.5	58.4	63.5	68.5
78.3	74.5	81.4	84.0	92.0	94.0
110.0	112.0	119.0	127.0	130.0	133.0
140.0	146.0	155.0	159.0	173.0	179.0
194.0	195.0	209.0	249.0	281.0	319.0
339.0	432.0	469.0	519	633	725
817	1776				

Table 1.3 ML estimates and Criteria for Comparison for data of Survival times of a group of patients suffering from Head and Neck cancer disease.

Distribution	α	λ	c	AIC	BIC	-logL
Weighted Pareto type II	3.75	78.7	1.82	560.98	566.34	277.49
SB Pareto type II	3.319465	151.1415	-	559.44	563.01	277.72
Pareto Type II	4.409056	758.139	-	564.92	568.48	280.45

Table 1.4: Remission times (in months) of a random sample of 128 bladder cancer patients.

0.08	2.09	3.48	4.87	6.94	8.66
13.11	23.63	0.20	2.23	3.52	4.98
6.97	9.02	13.29	0.40	2.26	3.57
5.06	7.09	9.22	13.80	25.74	0.50
2.46	3.64	5.09	7.26	9.47	14.24
25.82	0.51	2.54	3.70	5.17	7.28
9.74	14.76	26.31	0.81	2.62	3.82
5.32	7.32	10.06	14.77	32.15	2.64
3.88	5.32	7.39	10.34	14.83	34.26
0.90	2.69	4.18	5.34	7.59	10.66
15.96	36.66	1.05	2.69	4.23	5.41
7.62	10.75	16.62	43.01	1.19	2.75
4.26	5.41	7.63	17.12	46.12	1.26
2.83	4.33	5.49	7.66	11.25	17.14
79.05	1.35	2.87	5.62	7.87	11.64
17.36	1.40	3.02	4.34	5.71	7.93
11.79	18.10	1.46	4.40	5.85	8.26
11.98	19.13	1.76	3.25	4.50	6.25
8.37	12.02	2.02	3.31	4.51	6.54
8.53	12.03	20.28	2.02	3.36	6.76
12.07	21.73	2.07	3.36	6.93	8.65
12.63	22.69				

Table 1.5 ML estimates and Criteria for Comparison for data of remission times (in months) of a random sample of 128 bladder cancer patients

Distribution	α	λ	c	AIC	BIC	-logL
Weighted Pareto type II	5.13 (1.80)	20.8 (14.01)	0.586 (0.28)	826.15	834.71	410.08
SB Pareto type II	4.37 (0.93)	11.20 (3.81)	-	825.74	831.44	410.87
Pareto Type II	13.94 (15.38)	121.02 (142.71)	-	831.67	837.37	413.83

Table 1.6 Survival times (in months) of patients of melanoma studied by Susarla and Vanryzin (1978)

3.25	3.50	4.75	4.75	5.00	5.25
5.75	5.75	6.25	6.50	6.50	6.75
6.75	7.78	8.00	8.50	8.50	9.25
9.50	9.50	10.00	11.50	12.50	13.25
13.50	14.25	14.50	14.75	15.00	16.25
16.25	16.50	17.50	21.75	22.50	24.50
25.50	25.75	27.50	29.50	31.00	32.50
34.00	34.50	35.25	58.50		

Table 1.7 ML estimates and Criteria for Comparison for data in table 1.6

Distribution	α	λ	c	AIC	BIC	-logL
Weighted Pareto type II	11.78	3.32	8.79	334.6	340.5	164.210
SB Pareto type II	62.92241	476.9776	-	335.6	339.3	165.836
Pareto Type II	43675.12	686106.2	-	349.1	352.8	172.546

References

[1] Buckland, W. R., & Cox, D. R. Renewal Theory. Biometrika, . doi:10.2307/2334228 **51(1/2)**, 290, 1964.

[2] Gove, J. H. Environmental and Ecological Statistics doi:10.1023/a:1026000505636. , **10(4)**, 455-467, 2003.

[3] Patil, G. P., & Rao, C. R. Weighted Distributions and Size-Biased Sampling with Applications to Wildlife Populations and Human Families. Biometric doi:10.2307/2530008s. , **34(2)**, 179, 1978.

[4] Warren, W. G. Statistical Distributions in Forestry and Forest Products Research. A Modern Course on Statistical Distributions in Scientific Work doi:10.1007/978-94-010-1845-6_27, 369-384, 1975.

[5] Das K. K. and Roy T. D.“On Some Length-Biased Weighted Weibull Distribution”, Pelagia Research Library,Advances in Applied Science Research, **2(5)**, 465-475, 2011.

[6] Lee, E.T. and Wang, J.W. “Statistical Methods for Survival Data Analysis”, 3 edn, John Willey, New York, 2003.

[7]Lemonte, A.J. The beta log-logistic distribution. To appear in Brazilian Journal of Probability and Statistics, 2012.

[8]Zea, L.M., Silva, R.B., Bourguignon, M., Santos, A.M., Cordeiro, G.M. The beta exponentiated Pareto distribution with application to bladder cancer susceptibility. International Journal of Statistics and Probability **1**, 8–19, 2012.

[9] Efron, B. Logistic regression, survival analysis and the Kaplan- Meier curve. Journal of the American Statistical Association **83**, 414-425,1988.

[10] Susarla, V., &Ryzin, J. V. Empirical Bayes Estimation of a Distribution (Survival) Function from Right Censored Observations. The Annals of Statistics. doi:10.1214/aos/1176344249, **6(4),** 740-754, 1978.

[11] Schwarz, G. Estimating the dimension of a model. Ann. Stat. **5**, 461–464, 1987.

[12] Akaike, H. A new look at the statistical model identification. IEEE Trans. Autom.Control **19**, 716–723, 1974.

T. R. Jan has done his post-graduation and M.Phil from Aligarh Muslim University and obtained Doctorial from University of Kashmir in 2004. He is currently working as Senior Assistant Professor in the Department of Statistics, University of Kashmir, Srinagar, India. He has visited some foreign universities in USA and Malaysia as well. His research interests are in the areas of Bio-statistics, Reliability theory and Generalised models in biological sciences. He has published several research articles in reputed international journals of mathematical and biosciences.

B. A. Para is a research scholar in the Department of Statistics, University of Kashmir, Srinagar, India. His area of research is Statistical models in medical sciences. He has remained as a business analyst for two years in TCS Bangalore under CSSI Nielsen Europe Project. He completed his M.phil on topic entitled "Some Burr type distributions" from University of Kashmir in 2014. He has published several research articles in reputed international journals of statistical and mathematical sciences.