Weighted Bivariate Pseudo-Weibull Distribution: Statistical Properties &Application

Afaq Ahmad
Department of Mathematical Sciences, Islamic University of Science & Technology Awantipoora, Kashmir, India

Fatima Bi
Department of Mathematics, Baghwant University, Ajmer, India

Abstract- Multi component systems are widely used in computer science. The reliability of these systems plays a very important role in efficient working. These systems are not always supposed to follow the standard probability distribution and so Pseudo-distributions can be thought of as suitable alternatives In this paper, a new weighted model is introduced which would be obtained by assigning weights to Bivariate Pseudo-Weibull distribution as a compound of two random variable to model the failure rate of component reliability. We have studied some properties of the proposed distribution.

Keywords – Bivariate Pseudo-Weibull distribution, weighted distribution, Maximum likelihood estimation, moments and entropy.

I. INTRODUCTION

For the system to be efficient the reliability of components in a multicomponent system serves to be very important. The component failure time generally lies on lower side which could be modelled in an efficient way by making use of the probability distributions having a longer right tail. These consequences could be modelled by making of several probability distributions, namely, the exponential, log-normal ones and Weibull. The review of literature gives the information on different versions of the bivariate exponential and bivariate Weibull distributions which could be used for the modelling of two component systems. The authors in [1-2] have modelled lifetimes of multi component systems for stochastic processes by defining a new class of probability distributions. These distributions [1] are based upon a linear combination of independent random variables. The authors in [3] proposed a bivariate Pseudo Weibull distribution as a compound distribution of two random variables X and Y each having a Weibull distribution with certain parameters. The density function of the bivariate Pseudo-Weibull distribution (BPWD) obtained in [3] has the form.

$$f(x,y)=\alpha\beta\theta\phi(x)x^{\beta-1}y^{\theta-1}\exp\left[-\left\{\alpha x^{\beta}+\phi(x)y^{\theta}\right\}\right];$$
$$\alpha>0,\beta>0,\theta>0,\phi(x)>0,x>0,y>0 \quad (1)$$

Shahbaz and Ahmad [3] studied distribution (1) with reference to concomitants of order statistics by using $\phi(x)=1$. In this work we have studied a different version of distribution (1), by $\phi(x)=x^{\beta}$ using .The distribution is introduced in the following section with some common properties.

1.1 The bivariate Pseudo-Weibull distribution

The density function of the bivariate Pseudo-Weibull distribution has given in (1).we define a new class of density functions, different from those of [3], by using $\phi(x)=x^{\beta}$ The density function in this case is given as

$$f(x,y)=\alpha\beta\theta x^{2\beta-1}y^{\theta-1}\exp\left[-x^{\beta}\left\{\alpha+y^{\theta}\right\}\right];\alpha>0,\beta>0,\theta>0,\phi(x)>0,x>0,y>0$$

For the sake of simplicity we take $\alpha=1$. The density function in this case is given as

$$f(x,y)=\beta\theta x^{2\beta-1}y^{\theta-1}\exp\left[-x^{\beta}\left\{1+y^{\theta}\right\}\right];\beta>0,\theta>0,x>0,y>0 \quad (2)$$

II. WEIGHTED BIVARIATE PSEUDO-WEIBULL DISTRIBUTION

Weighted distribution theory gives an integrated method to study with model design and data interpretation problems. Weighted distributions arise commonly in studies connected to reliability, survival analysis, analysis of family data, biomedicine, ecology and several other areas, see Stene (1981) and Oluyede and George (2002). Several authors would have been presented important consequences on weighted distributions. Rao (1965) had presented a unified model of weighted distribution and known several sampling situations which can be shown by weighted distributions. These situations occur when the recorded observations cannot be considered as a random sample from the original distributions. This imply in some cases it is not likely to work with a random sample from population. Zelen (1974) presented weighted distribution to represent what is called as a length-biased sampling. Patil and Ord (1976) studied a size biased sampling and related invariant weighted distributions. Gupta and Tripathi (1996) studied the weighted version of the bivariate logarithmic series distribution, which has applications in many fields such as: ecology, social and behavioural sciences. Ahmed et al (2016) discussed length biased weighted lomax distribution with its applications.

To existent the idea of a weighted distribution, suppose that X and Y are nonnegative random variables with its probability density function (pdf) $f(x,y)$, then the p.d.f. of the weight random variables XY_w is known by

$$f_w(x.y)=\frac{w(x.y)f(x.y)}{E(w(x.y))};x\geq 0 \qquad (3)$$

where $w(x.y)$ be a non-negative weight function. Depending upon the choice of the weight function, $w(x.y)$ we have different weighted models. The weighted bivariate Pseudo-Weibull distribution is found by taking the weights $(xy)^c$ to the bivariate Pseudo-Weibull distribution. In this paper, the weighted bivariate Pseudo-Weibull distribution is proposed with pdf

$$f(x,y)=\frac{\beta\theta x^{2\beta+c-1}y^{\theta+c-1}\exp\left[-x^{\beta}\left\{1+y^{\theta}\right\}\right]}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]},\quad \text{where} \quad \int_0^{\infty}\int_0^{\infty} f(x.y)=1 \qquad (4)$$

2.1 The Moment generating function

The moment generating function of (4) is obtained as follows:

$$M_{X,Y}(t_1,t_2)=E\{\exp(t_1X+t_2Y)\}$$

$$=\int_0^{\infty}\int_0^{\infty}\exp(t_1x+t_2y)f(x.y)dxdy$$

$$=\int_0^{\infty}\int_0^{\infty}\exp(t_1x+t_2y)\frac{\beta\theta x^{2\beta+c-1}y^{\theta+c-1}\exp\left[-x^{\beta}\left\{1+y^{\theta}\right\}\right]}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]}dxdy \qquad (5)$$

$$=\frac{\beta\theta}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]}\int_0^{\infty}e^{t_2y}y^{\theta+c-1}\int_0^{\infty}e^{t_1x}x^{2\beta+c-1}\exp\left[-x^{\beta}\left\{1+y^{\theta}\right\}\right]dxdy$$

Now expanding e^{t_1x} in a power series we have

$$M_{X,Y}(t_1,t_2)=\frac{\beta\theta}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]}\int_0^\infty e^{t_2 y}y^{\theta+c-1}\left\{\sum_{p=0}^\infty\frac{t_1^p}{p!}\int_0^\infty x^{2\beta+p+c-1}\exp\left[-x^\beta\left\{1+y^\theta\right\}\right]dx\right\}dy$$

Now simplifying the expression we get,

$$M_{X,Y}(t_1,t_2)=\frac{\theta\Gamma\left[2+\frac{c+p}{\beta}\right]}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]}\sum_{p=0}^\infty\frac{t_1^p}{p!}\int_0^\infty e^{t_2 y}\frac{y^{\theta+c-1}}{\left(1+y^\theta\right)^{\left(2+\frac{c+p}{\beta}\right)}}dy$$

Again expanding $e^{t_2 y}$ in a power series, we have

$$M_{X,Y}(t_1,t_2)=\frac{\theta\Gamma\left[2+\frac{c+p}{\beta}\right]}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]}\sum_{p=0}^\infty\frac{t_1^p}{p!}\sum_{p=0}^\infty\frac{t_2^q}{q!}\int_0^\infty\frac{y^{\theta+q+c-1}}{\left(1+y^\theta\right)^{\left(2+\frac{c+p}{\beta}\right)}}dy$$

Now simplifying the expression we get, the moment generating function of weighted bivariate Pseudo-Weibull distribution is obtained as

$$M_{X,Y}(t_1,t_2)=\frac{\Gamma\left[\frac{q+c}{\beta}\right]\Gamma\left[2+\frac{c+p}{\beta}-\frac{c+q}{\theta}\right]}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]}\sum_{p=0}^\infty\sum_{q=0}^\infty\frac{t_1^p t_2^q}{p!q!} \tag{6}$$

2.2. Conditional moments

The marginal distribution of X and Y are obtained from (4) as

$$f(x,\beta)=\frac{\beta\Gamma\left[\frac{c}{\theta}+1\right]x^{\beta+c\left(1-\frac{\beta}{\theta}\right)-1}e^{-x\beta}}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]} \tag{7}$$

And $$f(y,\theta)=\frac{\theta\Gamma\left[\frac{c}{\beta}+2\right]y^{\theta+c-1}}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]\left(1+y^\theta\right)^{\left(2+\frac{c}{\beta}\right)}} \tag{8}$$

2.3 Entropy

The entropies of the distribution is defined as

$H(x,y)=E\left[-\ln\left\{f(x,y)\right\}\right]$

So for weighted bivariate Pseudo-Weibull distribution, the entropy is

$$E\left[-\ln\{f(x,y)\}\right]=-E\ln\left[\frac{\beta\theta x^{2\beta+c-1}y^{\theta+c-1}\exp\left[-x^{\beta}\{1+y^{\theta}\}\right]}{\Gamma\left[\frac{c}{\theta}\right]\Gamma\left[2+\frac{c}{\beta}-\frac{c}{\theta}\right]}\right]$$

$$=-\log(\beta\theta)+\log\left[\Gamma\left(\frac{c}{\theta}\right)\Gamma\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)\right]-(2\beta+c-1)E[\log x]-(\theta+c-1)E[\log y]-E\left[-x^{\beta}\{1+y^{\theta}\}\right] \quad (9)$$

Now evaluate various expectations

$$E[\log x]=\int_0^{\infty}\int_0^{\infty}\log x f(x.y)dxdy\ ,\ E[\log y]=\int_0^{\infty}\int_0^{\infty}\log y f(x.y)dxdy$$

$$E\left[x^{\beta}\{1+y^{\theta}\}\right]=\int_0^{\infty}\int_0^{\infty}x^{\beta}\{1+y^{\theta}\}f(x.y)dxdy=\left(2+\frac{c}{\beta}\right) \quad (10)$$

Using the value (10) in (9) we get,

$$E\left[-\ln\{f(x,y)\}\right]=-\log(\beta\theta)+\log\left[\Gamma\left(\frac{c}{\theta}\right)\Gamma\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)\right]-(2\beta+c-1)E[\log x]-(\theta+c-1)E[\log y]-\left(2+\frac{c}{\beta}\right) \quad (11)$$

III. Estimation of parameter

In this section, we derive the estimates of parameters of distribution by using the method of maximum likelihood estimation. The method of Maximum likelihood estimation is the most popular technique used for estimating the parameters of weighted bivariate Pseudo-Weibull distribution. Let $x_1, x_2, x_3, \ldots x_n$ be a random sample from the weighted bivariate Pseudo-Weibull distribution, then the corresponding log likelihood function is given by,

$$l(\theta)=l0g(\theta|x)=\sum_{i=1}^{n}\log f(x_i|\theta)$$

$$l=n\ln\beta+n\ln\theta-\log\Gamma\left(\frac{c}{\theta}\right)-\log\Gamma\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)+(2\beta+c-1)\sum_{i=1}^{n}\log x_i+$$
$$(\theta+c-1)\sum_{i=1}^{n}\log y_i-\sum_{i=1}^{n}-x_i^{\beta}\{1+y_i^{\theta}\} \quad (12)$$

Now differentiating above with respect to the parameters, we obtain the normal equations;

$$\frac{\partial}{\partial\beta}=0\Rightarrow\frac{n}{\beta}+\frac{c}{\beta^2}\frac{\Gamma'\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)}{\Gamma\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)}+2\sum_{i=1}^{n}\log x_i-\beta\sum_{i=1}^{n}x_i^{\beta-1}\{1+y_i^{\theta}\}=0 \quad (13)$$

$$\frac{\partial}{\partial\theta}=0\Rightarrow\frac{n}{\theta}+\frac{2}{\theta^2}\frac{\Gamma'\left(\frac{c}{\theta}\right)}{\Gamma\left(\frac{c}{\theta}\right)}-\frac{2}{\theta^2}\frac{\Gamma'\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)}{\Gamma\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)}+\sum_{i=1}^{n}\log y_i-\theta\sum_{i=1}^{n}y_i^{\theta-1}x_i^{\beta}=0 \quad (14)$$

$$\frac{\partial}{\partial c}=0\Rightarrow\frac{\Gamma'\left(\frac{c}{\theta}\right)}{\Gamma\left(\frac{c}{\theta}\right)}+\left(\frac{1}{\beta}-\frac{1}{\theta}\right)\frac{\Gamma'\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)}{\Gamma\left(2+\frac{c}{\beta}-\frac{c}{\theta}\right)}+\sum_{i=1}^{n}\log x_i+\sum_{i=1}^{n}\log y_i=0 \tag{15}$$

The MLE $\hat{\eta}=(\hat{\beta},\hat{\theta},c)$ of $\eta=(\beta,\theta,c)$ is obtained by solving the above 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 (12).

IV. APPLICATIONS

DATA I: The first data set represents the survival times of a group of patients suffering from Head and Neck cancer diseases and treated using a combination of radiotherapy and chemotherapy (RT+CT). The data set has been previously used by Efron (1988) and Shanker et al., (2015). It has been successfully used to assess the superiority of the Exponential distribution over the Lindley distribution. It has forty four (44) observations and they are as follows:

12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81.43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194, 195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776.

We have fitted the weighted bivariate Pseudo-Weibull and bivariate Pseudo-Weibull distribution to this data set. The parameter estimates, log-likelihood, AIC and BIC values are reported in the Table 2. From the table it is seen that weighted bivariate Pseudo-Weibull distribution has the minimum AIC and BIC values so we can conclude that the weighted bivariate Pseudo-Weibull distribution provides better fit than bivariate Pseudo-Weibull

Table -1 Summary of above data

Min	Ist Qu.	Median	Mean	3rd Qu	Max.
12.20	**67.21**	**128.50**	**223.48**	**219.00**	**1776.00**

Table -2 ML estimates and Criteria for Comparison for data of patients suffering from head and neck cancer diseases

Distribution	Estimates	-2logL	AIC	AICC	BIC
WBPWD	$\beta=$ **0.19945** $\theta=$ **030007** **c = 0.52964**	**92.20871**	**98.20871**	**98.80871**	**103.5613**
WBPWD	$\beta=$ **0.08672** $\theta=$ **0.11339**	**150.255**	**154.255**	**154.5477**	**157.8234**

DATA II: The second data set was given by Lee and Wang (2003) and it represents the remission times (in months) of a random sample of one hundred and twenty-eight (128) bladder cancer patients. Its application in survival analysis has been identified and the data set is as follows:

0.08, 2.09, 2.73, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.22, 3.52, 4.98, 6.99, 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, 15.62, 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, 6.93, 8.65, 12.63, 22.69

We have fitted the weighted bivariate Pseudo-Weibull and bivariate Pseudo-Weibull distribution to this data set. The parameter estimates, log-likelihood, AIC and BIC values are reported in the Table 4. From the table it is seen that weighted bivariate Pseudo-Weibull distribution has the minimum AIC and BIC values so we can conclude that the weighted bivariate Pseudo-Weibull distribution provides better fit than bivariate Pseudo-Weibull.

Table -3 Summary of above data

Min	Ist Qu.	Median	Mean	3rd Qu	Max.
0.080	3.295	6.050	9.311	11.678	79.050

Table -4 ML estimates and Criteria for Comparison for data of remission times of bladder cancer patients

Distribution	Estimates	-2logL	AIC	AICC	BIC
WBPWD	$\beta =$ 0.23288 $\theta =$ 0.12483 c = 0.53820	662.3578	668.3578	668.5513	676.9139
WBPWD	$\beta =$ 0.20274 $\theta =$ 0.26185	788.687	792.687	792.783	798.3911

V.CONCLUSION

In this paper the weighted bivariate Pseudo-Weibull model has been presented along with its structural properties. Parameters have been estimated using the method of Maximum Likelihood. This distribution has also been applied to two real data sets. It has been observed that weighted bivariate Pseudo-Weibull distribution provides a better fit for data related to patients suffering from head and neck cancer diseases and of diminution times of bladder cancer patients than bivariate Pseudo-Weibull distribution.

REFERENCES

[1] J.K Filus,L.Z Filus ,On some new classes of multivariate probability distribution, Pakistan J,Statist.22(1) (2006) 21-42

[2] J.K Filus,L.Z Filus ,Construction of new continuous time stochastic processes, Pakistan J,Statist.24 (3) (2008) 227-251

[3] S.Shahbaz M.Ahmad, Concomitants of order statistics for bivariate pseudo-Weibull distribution, world Appl.Sci.J. 6(10) (2009) 1409-1412

[4] Ahmad, A., Ahmad, S.P. and Ahmed, A., (2016). "Length Biased Weighted Lomax distribution: Statistical Properties and Applications", Pakistan Journal of Statistics and Operations Research, vol. 12(2), pp. 245-255.

[5] Fisher, R.A., 1934. "The effect of methods of ascertainment upon the estimation of frequencies", Ann. Eng. 6, 13–25.

[6] Patil, G.P., Rao, C.R., (1978). Weighted distributions and size-biased sampling with applications to wildlife and human families. Biometrics 34, 179–189.

[7] Patil, G.P., Rao, C.R., Ratnaparkhi, M.V., (1986). "On discrete weighted distributions and their use in model choice for observed data", Communication in Statistics - Theory and Methods, vol. 15(3), 907–918.

[8] Rao, C.R., (1965). "On discrete distributions arising out of methods of ascertainment. Classical and Contagious Discrete Distributions", Eds. G.P. Patil Pergamon Press and Statistical Publishing Society -Calcutta.

[9] Gupta AK, Tripathi RC (1996) Weighted bivariate logarithmic series distributions. Commun Stat Theory Methods 25:1099–1117.

[10] Patil GP, Ord JK (1976) On size biased sampling and related form invariant weighted distribution. Indian J Stat 39:48–61.

[11] Stene J (1981) Probability distributions arising from the ascertainment and the analysis of data on human families and other groups. Stat Distrib Sci Work Appl Phys Soc Life Sci 6:233–244.

[12] Zelen M (1974) Problems in cell kinetics and the early detection of disease. In: Proschan F, Sering RJ (eds) Reliability and biometry. SIAM, Philadelphia.

[13] Saman S, M Qaiser Shahbaz, M Ahsanullah and M Mohsin.(2010). "On a new class probability distribution", Applied Mathematics Letters ,doi-10.1016/j.aml.2010.11.010. page no. 545-552.

[14] Gupta AK, Tripathi RC (1996) Weighted bivariate logarithmic series distributions. Commun Stat Theory Methods 25:1099–1117.

[15] Lee, E. T. and Wang, J. W. (2003). "Statistical methods for survival data analysis (3rd Edition)", John Wiley and Sons, New York, USA, 535 Pages, ISBN 0-471-36997-7.

[16] Prakash, G. and Singh, D.C. (2010). Bayesian Shrinkage estimation in a class of life testing distribution. Data Science Journal, 8(30), 243-258. Schabe, H. (1991). Bayes estimates under asymmetric loss. IEEE Trans. Reliab., 40(1), 63-67.

[17] Shanker, R., Fesshaye, H. and Selvaraj, S. (2015). "On modelling life time's data using exponential and Lindley distributions", Biometric and Biostatistics International Journal, 2(5), 1-9.

[18] Stene J (1981) Probability distributions arising from the ascertainment and the analysis of data on human families and other groups. Stat Distrib Sci Work Appl Phys Soc Life Sci 6:233–244.