POWER WEIGHTED SUJATHA DISTRIBUTION WITH PROPERTIES AND APPLICATION TO SURVIVAL TIMES OF PATIENTS OF HEAD AND NECK CANCER

Rama Shanker
•
Department of Statistics, Assam University, Silchar, Assam, India
Shankerrama2009@gmail.com

Kamlesh Kumar Shukla
•
Department of Community Medicine, Noida International Institute
of Medical Sciences, Noida International University, Gautam Budh Nagar, India
Kkshukla22@gmail.com

Abstract

In this paper a power weighted Sujatha distribution, which includes power Sujatha distribution, weighted Sujatha distribution and Sujatha distribution as particular cases, has been proposed. Its statistical properties including behavior of density function, moments, hazard rate function, and mean residual life function have been discussed. Estimation of parameters has been discussed using the method of maximum likelihood. A simulation study has been presented to know the performance of maximum likelihood estimates of parameters. Application of the proposed distribution have been explained with a real lifetime data relating to patients suffering from head and neck cancer and goodness of fit shows quite satisfactory fit.

Keywords: *Sujatha distribution, Weighted Sujatha distribution, Power Sujatha distribution, Hazard rate function, Mean residual life function, Maximum Likelihood estimation*

1. Introduction

It has been observed that the survival times of patients suffering from head and neck cancer needs special consideration to find a suitable distribution which can be used to model the data. During recent decades several one parameter, two-parameter and three-parameter lifetime distributions have been proposed in statistics literature to model survival times of patients suffering from head and neck cancer and observed that all proposed distributions are not very much suitable due to theoretical or applied point of view. It has been observed that, in general, the survival times of patients suffering from head and neck cancer are stochastic in nature and while discussing the goodness of fit of several one parameter , two-parameter and three-parameter well-known distributions which were earlier proposed by different researchers that these distribution does not give good fit.

Shanker [1] proposed a one parameter Sujatha distribution having its probability density function (pdf) and cumulative distribution function (cdf) as

$$f_1(y;\theta) = \frac{\theta^3}{\theta^2+\theta+2}(1+y+y^2)e^{-\theta y}; y>0, \theta>0 \quad (1.1)$$

$$F_1(y;\theta) = 1-\left[1+\frac{\theta y(\theta y+\theta+2)}{\theta^2+\theta+2}\right]e^{-\theta y}; y>0, \theta>0(1.2) \quad (1.2)$$

Shanker and Shukla [2], taking a weight function $x^{\alpha-1}$in (1.1), proposed a two-parameter weighted Sujatha distribution (WSD) defined by its pdf and cdf

$$f_2(y;\theta,\alpha) = \frac{\theta^{\alpha+2}}{\theta^2+\alpha\theta+\alpha(\alpha+1)}\frac{y^{\alpha-1}}{\Gamma(\alpha)}(1+y+y^2)e^{-\theta y}; y>0, \theta>0, \alpha>0 \quad (1.3)$$

$$F_2(y;\theta,\alpha) = 1-\frac{\{\theta^2+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha,\theta y)+(\theta y)^{\alpha}(\theta y+\theta+\alpha+1)e^{-\theta y}}{\{\theta^2+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}, \quad (1.4)$$

where $\Gamma(\alpha,z)=\int_z^{\infty} e^{-y}y^{\alpha-1}dy; y\geq 0, \alpha>0$ is the upper incomplete gamma function.

Shanker and Shukla [3], taking a power transformation $x=y^{\frac{1}{\beta}}$ in (1.1), proposed a two-parameter power Sujatha distribution(PSD) defined by its pdf and cdf

$$f_3(x;\theta,\alpha) = \frac{\alpha\theta^3}{\theta^2+\theta+2}x^{\alpha-1}(1+x^{\alpha}+x^{2\alpha})e^{-\theta x^{\alpha}}; x>0, \theta>0, \alpha>0 \quad (1.5)$$

$$F_3(y;\theta,\alpha) = 1-\left[1+\frac{\theta y^{\alpha}(\theta y^{\alpha}+\theta+2)}{\theta^2+\theta+2}\right]e^{-\theta y^{\alpha}}; y>0, \theta>0, \alpha>0 \quad (1.6)$$

Ghitany *et al*. [4] proposed a two-parameter weighted Lindley distribution (WLD) having parameters θ and α defined by its probability density function (pdf) and cumulative distribution function (cdf)

$$f_4(y;\theta,\alpha) = \frac{\theta^{\alpha+1}}{\theta+\alpha}\frac{y^{\alpha-1}}{\Gamma(\alpha+1)}(1+y)e^{-\theta y}; y>0, \theta>0, \alpha>0 \quad (1.7)$$

$$F_4(y;\theta,\alpha) = 1-\frac{(\theta+\alpha)\Gamma(\alpha,\theta y)+(\theta y)^{\alpha}e^{-\theta y}}{(\theta+\alpha)\Gamma(\alpha)}; y>0, \theta>0, \alpha>0, \quad (1.8)$$

where $\Gamma(\alpha)$ and $\Gamma(\alpha,z)$ are the complete gamma function and the upper incomplete gamma function. Its structural properties including moments, hazard rate function, mean residual life function, estimation of parameters and applications for modeling survival time data has been discussed by Ghitany *et al*. [4]. Shanker *et al.* [5] discussed various moments based properties including coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion of weighted Lindley distribution and its applications to model lifetime data from biomedical sciences and engineering.

Ghitany *et al.* [6] proposed a power Lindley distribution (PLD) having parameters θ and α defined by its pdf and cdf

$$f_5(y;\theta,\alpha) = \frac{\alpha\theta^2}{(\theta+1)}y^{\alpha-1}(1+y^{\alpha})e^{-\theta y^{\alpha}}; y>0, \theta>0, \alpha>0 \quad (1.9)$$

$$F_5(x;\theta,\alpha) = 1-\left[1+\frac{\theta x^{\alpha}}{\theta+1}\right]e^{-\theta x^{\alpha}}; x>0, \theta>0, \alpha>0 \quad (1.10)$$

Note that the PLD is a convex combination of Weibull (α,θ) and a generalized gamma $(2,\alpha,\theta)$distribution with mixing proportion$\frac{\theta}{\theta+1}$. Ghitany *et al*. [6] has discussed the properties of PLD including the shapes of the density, hazard rate functions, moments, skewness and kurtosis measures, estimation of parameters using maximum likelihood estimation and application to model a real lifetime data from engineering. Recall that at $\alpha=1$ both WLD in (1.7) and PLD in (1.9) reduce to Lindley distribution introduced by Lindley [7] having pdf and cdf

$$f_6(y;\theta) = \frac{\theta^2}{\theta+1}(1+y)e^{-\theta y}; y>0, \theta>0 \quad (1.10)$$

$$F_6(y;\theta) = 1-\left[1+\frac{\theta y}{\theta+1}\right]e^{-\theta y}; y>0, \theta>0 \quad (1.11)$$

Ghitany *et al*. [8] have discussed its various statistical and mathematical properties and application. Shanker *et al*. [9] have detailed study on modeling of lifetime data using both exponential and

Lindley distributions and observed that there are many lifetime data where exponential distribution gives better fit than Lindley distribution.

Zakerzadeh and Dolati [10] have introduced a three parameter generalized Lindley distribution (GLD) having pdf and cdf given by

$$f_7(x;\theta,\alpha,\beta)=\frac{\theta^{\alpha+1}}{\theta+\beta}\frac{x^{\alpha-1}}{\Gamma(\alpha+1)}(\alpha+\beta x)e^{-\theta x};x>0,\theta>0,\alpha>0,\beta>0 \qquad (1.12)$$

$$F_7(x;\theta,\alpha,\beta)=1-\frac{\alpha(\beta+\theta)\Gamma(\alpha,\theta x)+\beta(\theta x)^{\alpha}e^{-\theta x}}{(\beta+\theta)\Gamma(\alpha+1)};x>0,\theta>0,\alpha>0,\beta>0 \quad (1.13)$$

Lindley distribution, gamma distribution and weighted Lindley distribution (WLD) are particular cases of (1.9) at $(\alpha=\beta=1)$, $(\beta=0)$ and $(\beta=\alpha)$, respectively. Shanker [11] obtained various raw moments and central moments of GLD and discussed properties based on moments including coefficient of variation, skewness, kurtosis and index of dispersion of GLD and its comparative study with generalized gamma distribution (GGD) introduced by Stacy [12] to model various lifetime data from engineering and biomedical sciences and concluded that in many cases GGD gives much better fit than GLD. Shanker and Shukla [13] have detailed comparative study on modeling of real lifetime data from engineering and biomedical sciences using GLD and generalized gamma distribution (GGD) introduced by Stacy (1962) [12] and concluded that there are several lifetime data where GGD gives much better fit than GLD.

In the present paper, a three - parameter power weighted Sujatha distribution (PWSD) which includes PSD, WSD and Sujatha distribution as particular cases, has been proposed and discussed. Its raw moments been obtained. The survival function and the hazard rate function of the distribution have been derived and their shapes have been discussed for varying values of the parameters. The estimation of its parameters has been discussed using maximum likelihood method. Finally, the goodness of fit and the application of the distribution have been explained through a real lifetime data relating to patients suffering from head and cancer and the fit has been compared with other one parameter, two-parameter and three-parameter lifetime distributions.

2. Power weighted Sujatha distribution

Assuming the power transformation $X=Y^{\frac{1}{\beta}}$ in the pdf of WSD (1.3), the pdf of the random variable Xcan be obtained as

$$f_8(x;\theta,\alpha,\beta)=\frac{\beta\theta^{\alpha+2}}{\theta+\alpha\theta+\alpha(\alpha+1)}\frac{x^{\beta\alpha-1}}{\Gamma(\alpha)}\left(1+x^{\beta}+x^{2\beta}\right)e^{-\theta x^{\beta}};x>0,\theta>0,\alpha>0,\beta>0 \qquad (2.1)$$

We would call the distribution in (2.1) as the power weighted Sujatha distribution (PWSD) and we denote it as PWSD(θ,α,β). It can be easily verified that the PWSD(θ,α,β) in (2.1) reduces to Sujatha distribution, WSD and PSD for$(\alpha=\beta=1)$,$(\beta=1)$ and$(\alpha=1)$, respectively. It can be easily verified that PWSD is a convex combination of generalized gamma distributions with different parameters, namely GGD(θ,α,β), GGD$(\theta,\alpha+1,\beta)$ and GGD$(\theta,\alpha+2,\beta)$. That is

$$f_5(x;\theta,\alpha,\beta)=p_1g_1(\theta,\alpha,\beta)+p_2g_2(\theta,\alpha+1,\beta)+(1-p_1-p_2)g_3(\theta,\alpha+2,\beta),$$

Where

$$p_1=\frac{\theta^2}{\theta^2+\alpha\theta+\alpha(\alpha+1)},\ p_2=\frac{\alpha\theta}{\theta^2+\alpha\theta+\alpha(\alpha+1)}$$

$$g_1(\theta,\alpha,\beta)=\frac{\beta\theta^{\alpha}}{\Gamma(\alpha)}x^{\beta\alpha-1}e^{-\theta x^{\beta}};x>0,\theta>0,\alpha>0,\beta>0$$

$$g_2(\theta,\alpha+1,\beta)=\frac{\beta\theta^{\alpha+1}}{\Gamma(\alpha+1)}x^{\beta(\alpha+1)-1}e^{-\theta x^{\beta}};x>0,\theta>0,\alpha>0,\beta>0$$

$$g_3(\theta,\alpha+2,\beta)=\frac{\beta\theta^{\alpha+2}}{\Gamma(\alpha+2)}x^{\beta(\alpha+2)-1}e^{-\theta x^{\beta}};x>0,\theta>0,\alpha>0,\beta>0$$

Graphs of density function of PWSD for varying values of parameters θ,αandβ have been drawn and presented in figure 1. From the figure 1, It was observed that pdf of PWSD is increasing with increased value of theta at fixed value of alpha and beta, whereas its value is decreasing with increased value of alpha at fixed valued of theta and beta.

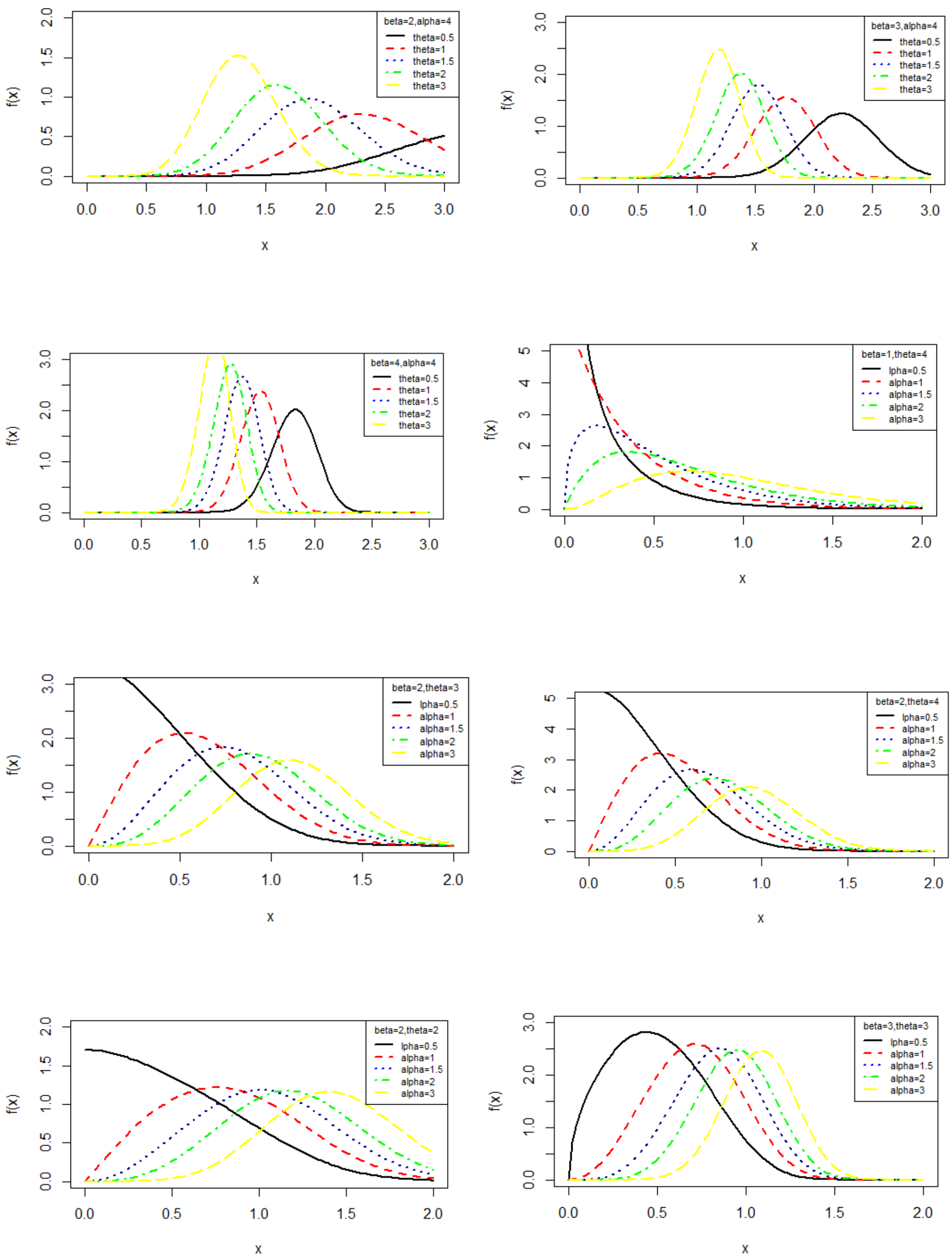

Figure 1: Graphs of the pdf of PWSD for varying values of parametersθ, αandβ

3. Moments

The rth moment about origin of PWSD (2.1) can be obtained as

$$\mu_r{}' = \frac{\beta\theta^{\alpha+2}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)} \int_0^\infty x^{\beta\alpha+r-1}\left(1+x^{\beta}+x^{2\beta}\right)e^{-\theta x^{\beta}}dx$$

$$= \frac{\beta\theta^{\alpha+2}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}\left[\int_0^\infty e^{-\theta x^{\beta}}x^{\beta\alpha+r-1}\,dx + \int_0^\infty e^{-\theta x^{\beta}}x^{\beta\alpha+\beta+r-1}\,dx + \int_0^\infty e^{-\theta x^{\beta}}x^{\beta\alpha+2\beta+r-1}\,dx\right]$$

Assuming $u=\theta x^{\beta}$, we have $du=\theta\beta x^{\beta-1}dx,\ x=\left(\frac{u}{\theta}\right)^{\frac{1}{\beta}}$ and $dx=\frac{du}{\theta\beta\left(\frac{u}{\theta}\right)^{\frac{\beta-1}{\beta}}}$

Thus, we have

$$\mu_r'=\frac{\theta^{\alpha+1}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}\begin{bmatrix}\frac{1}{\theta^{\alpha+\frac{r}{\beta}-1}}\int_0^\infty e^{-u}u^{\alpha+\frac{r}{\beta}-1}\,du+\frac{1}{\theta^{\alpha+\frac{r}{\beta}}}\int_0^\infty e^{-u}u^{\alpha+\frac{r}{\beta}+1-1}\,du\\+\frac{1}{\theta^{\alpha+\frac{r}{\beta}+1}}\int_0^\infty e^{-u}u^{\alpha+\frac{r}{\beta}+2-1}\,du\end{bmatrix}$$

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

$$=\frac{\theta^{\alpha+1}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}\left[\frac{\theta^2\Gamma\left(\alpha+\frac{r}{\beta}\right)+\theta\Gamma\left(\alpha+\frac{r}{\beta}+1\right)+\Gamma\left(\alpha+\frac{r}{\beta}+2\right)}{\theta^{\alpha+\frac{r}{\beta}+1}}\right]$$

$$=\frac{\theta^2+\theta\left(\alpha+\frac{r}{\beta}\right)+\left(\alpha+\frac{r}{\beta}\right)\left(\alpha+\frac{r}{\beta}+1\right)}{\theta^{\frac{r}{\beta}}\{\theta+\alpha\theta+\alpha(\alpha+1)\}}\frac{\Gamma\left(\alpha+\frac{r}{\beta}\right)}{\Gamma(\alpha)}$$

$$=\frac{\theta^2\beta^2+\theta\beta(\alpha\beta+r)+(\alpha\beta+r)(\alpha\beta+\beta+r)}{\beta\theta^{\frac{r}{\beta}}\{\theta+\alpha\theta+\alpha(\alpha+1)\}}\frac{\Gamma\left(\alpha+\frac{r}{\beta}\right)}{\Gamma(\alpha)};r=1,2,3,\ldots \tag{3.1}$$

Substituting $r=1,2,3$ and 4 in (3.1), the first four moments about the origin of PWSD can be obtained. Again using the relationship between moments about origin and central moments, central moments can be obtained. Since the expressions for central moments are complicated, central moments are not being given.

4. Hazard Rate Function

The survival function $S(x;\theta,\alpha,\beta)$ of PWSD can be obtained as

$$S(x;\theta,\alpha,\beta)=P(X>x)=\int_x^\infty f_5(t;\theta,\alpha,\beta)\,dt$$

$$=\frac{\beta\theta^{\alpha+2}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}\int_x^\infty t^{\beta\alpha-1}\left(1+t^\beta+t^{2\beta}\right)e^{-\theta t^\beta}dt$$

$$=\frac{\beta\theta^{\alpha+2}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}\left[\int_x^\infty t^{\beta\alpha-1}\,e^{-\theta t^\beta}dt+\int_x^\infty t^{\beta\alpha+\beta-1}\,e^{-\theta t^\beta}dt+\int_x^\infty t^{\beta\alpha+2\beta-1}\,e^{-\theta t^\beta}dt\right]$$

Taking $u=t^\beta$ and $t=(u)^{\frac{1}{\beta}}$ gives $dt=\frac{du}{\beta u^{\frac{\beta-1}{\beta}}}$ and thus we have

$$S(x;\theta,\alpha,\beta)=\frac{\theta^{\alpha+2}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}\begin{bmatrix}\int_{x^\beta}^\infty e^{-\theta u}u^{\frac{\beta\alpha-1}{\beta}}\frac{du}{u^{\frac{\beta-1}{\beta}}}+\int_{x^\beta}^\infty e^{-\theta u}u^{\frac{\beta\alpha+\beta-1}{\beta}}\frac{du}{u^{\frac{\beta-1}{\beta}}}\\+\int_{x^\beta}^\infty e^{-\theta u}u^{\frac{\beta\alpha+2\beta-1}{\beta}}\frac{du}{u^{\frac{\beta-1}{\beta}}}\end{bmatrix}$$

$$=\frac{\theta^{\alpha+2}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}\left[\int_{x^\beta}^\infty e^{-\theta u}u^{\alpha-1}du+\int_{x^\beta}^\infty e^{-\theta u}u^{\alpha}du+\int_{x^\beta}^\infty e^{-\theta u}u^{\alpha+1}du\right]$$

$$=\frac{\theta^{\alpha+2}}{\{\theta+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha)}\begin{bmatrix}\frac{\Gamma\left(\alpha,\theta x^\beta\right)}{\theta^\alpha}+\frac{e^{-\theta x^\beta}\left(\theta x^\beta\right)^\alpha+\alpha\Gamma\left(\alpha,\theta x^\beta\right)}{\theta^{\alpha+1}}\\+\frac{e^{-\theta x^\beta}\left(\theta x^\beta\right)^\alpha\left(\theta x^\beta+\alpha+1\right)+\alpha(\alpha+1)\Gamma\left(\alpha,\theta x^\beta\right)}{\theta^{\alpha+2}}\end{bmatrix}$$

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

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

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

where $\Gamma\left(\alpha,\theta x^\beta\right)$ is the upper incomplete gamma function defined as

$$\Gamma\left(\alpha,\theta x^\beta\right)=\int_{\theta x^\beta}^\infty y^{\alpha-1}e^{-y}dy;\alpha>0,\theta x^\beta>0.$$

It can be easily verified that at $(\beta=1)$,$(\alpha=1)$ and $(\alpha=\beta=1)$, the survival function of PWSD reduce to the survival function of WSD, PSD and Sujatha distribution, respectively.

Thus the cdf of PWSD can be obtained using $F_5(x;\theta,\alpha,\beta) = 1 - S(x;\theta,\alpha,\beta)$.

The hazard rate function, $h(x;\theta,\alpha,\beta)$, of PWSD can be given by

$$h(x;\theta,\alpha,\beta) = \frac{f_5(x;\theta,\alpha,\beta)}{S(x;\theta,\alpha,\beta)}$$

$$= \frac{\beta\theta^{\alpha+2}x^{\beta\alpha-1}(1+x^{\beta}+x^{2\beta})e^{-\theta x^{\beta}}}{\{\theta^2+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha,\theta x^{\beta})+e^{-\theta x^{\beta}}(\theta x^{\beta}+\theta+\alpha+1)(\theta x^{\beta})^{\alpha}}.$$

Graphs of $h(x;\theta,\alpha,\beta)$for varying values of parameters θ,αandβ are shown in figure 2.

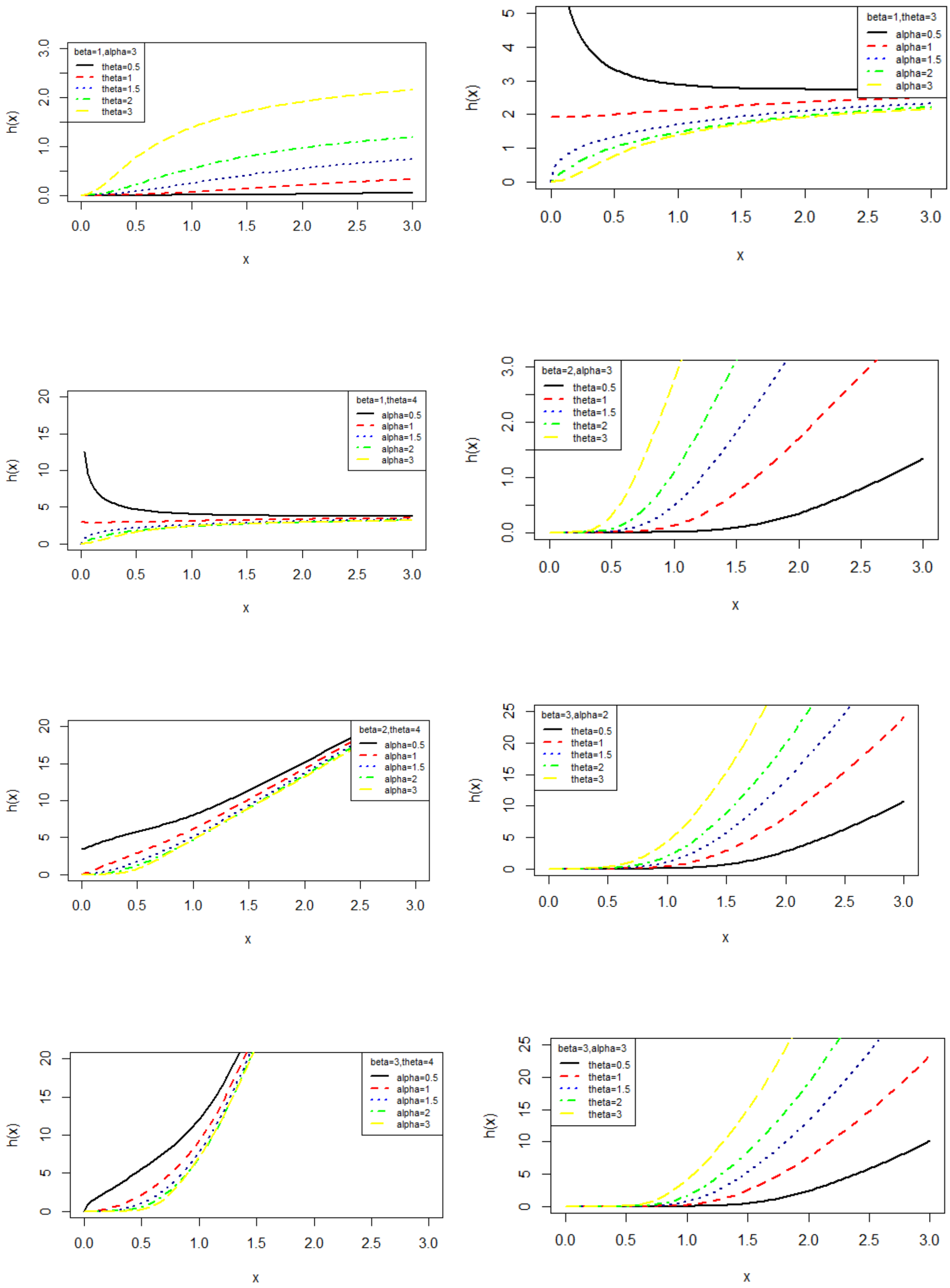

Figure 2: Graphs of hazard rate function of PWSD for varying values of parametersθ,αandβ.

5. Mean Residual life function

The mean residual life function of PWSD can be obtained as

$$m(x;\theta,\alpha,\beta)=\frac{1}{S(x;\theta,\alpha,\beta)}\int_x^{\infty} t f(t;\theta,\alpha,\beta)\,dt-x$$

$$=\frac{\beta\theta^{\alpha+2}}{\{\theta^2+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha,\theta x^{\beta})+e^{-\theta x^{\beta}}(\theta x^{\beta}+\theta+\alpha+1)(\theta x^{\beta})^{\alpha}} \times\int_x^{\infty} t^{\beta\alpha}\left(1+t^{\beta}+t^{2\beta}\right)e^{-\theta t^{\beta}}dt-x$$

$$=\frac{\beta\theta^{\alpha+2}}{\{\theta^2+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha,\theta x^{\beta})+e^{-\theta x^{\beta}}(\theta x^{\beta}+\theta+\alpha+1)(\theta x^{\beta})^{\alpha}} \times\left[\int_x^{\infty} e^{-\theta t^{\beta}}\,t^{\beta\alpha}dt+\int_x^{\infty} e^{-\theta t^{\beta}}\,t^{\beta\alpha+\beta}dt+\int_x^{\infty} e^{-\theta t^{\beta}}\,t^{\beta\alpha+2\beta}dt\right]-x$$

Taking $u=t^{\beta}$and $t=(u)^{\frac{1}{\beta}}$gives $dt=\frac{du}{\beta u^{\frac{\beta-1}{\beta}}}$ and thus we have

$$m(x;\theta,\alpha,\beta)=\frac{\theta^{\alpha+2}}{\{\theta^2+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha,\theta x^{\beta})+e^{-\theta x^{\beta}}(\theta x^{\beta}+\theta+\alpha+1)(\theta x^{\beta})^{\alpha}} \times\left[\int_{x^{\beta}}^{\infty} e^{-\theta u}\,u^{\alpha+\frac{1}{\beta}-1}du+\int_{x^{\beta}}^{\infty} e^{-\theta u}\,u^{\alpha+\frac{1}{\beta}+1-1}du+\int_{x^{\beta}}^{\infty} e^{-\theta u}\,u^{\alpha+\frac{1}{\beta}+2-1}du\right]-x$$

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

$$=\frac{\theta^2\Gamma\left(\alpha+\frac{1}{\beta},\theta x^{\beta}\right)+\theta\Gamma\left(\alpha+\frac{1}{\beta}+1,\theta x^{\beta}\right)+\Gamma\left(\alpha+\frac{1}{\beta}+2,\theta x^{\beta}\right)}{\theta^{\frac{1}{\beta}}\{\theta^2+\alpha\theta+\alpha(\alpha+1)\}\Gamma(\alpha,\theta x^{\beta})+e^{-\theta x^{\beta}}(\theta x^{\beta}+\theta+\alpha+1)(\theta x^{\beta})^{\alpha}}-x.$$

Graphs of mean residual function of PWSD for varying values of parameters θ,α andβ have been drawn and presented in figure 3. From the figure 3, It was observed that mean residual of PWSD is decreasing with increased value of theta at fixed value of alpha and beta.

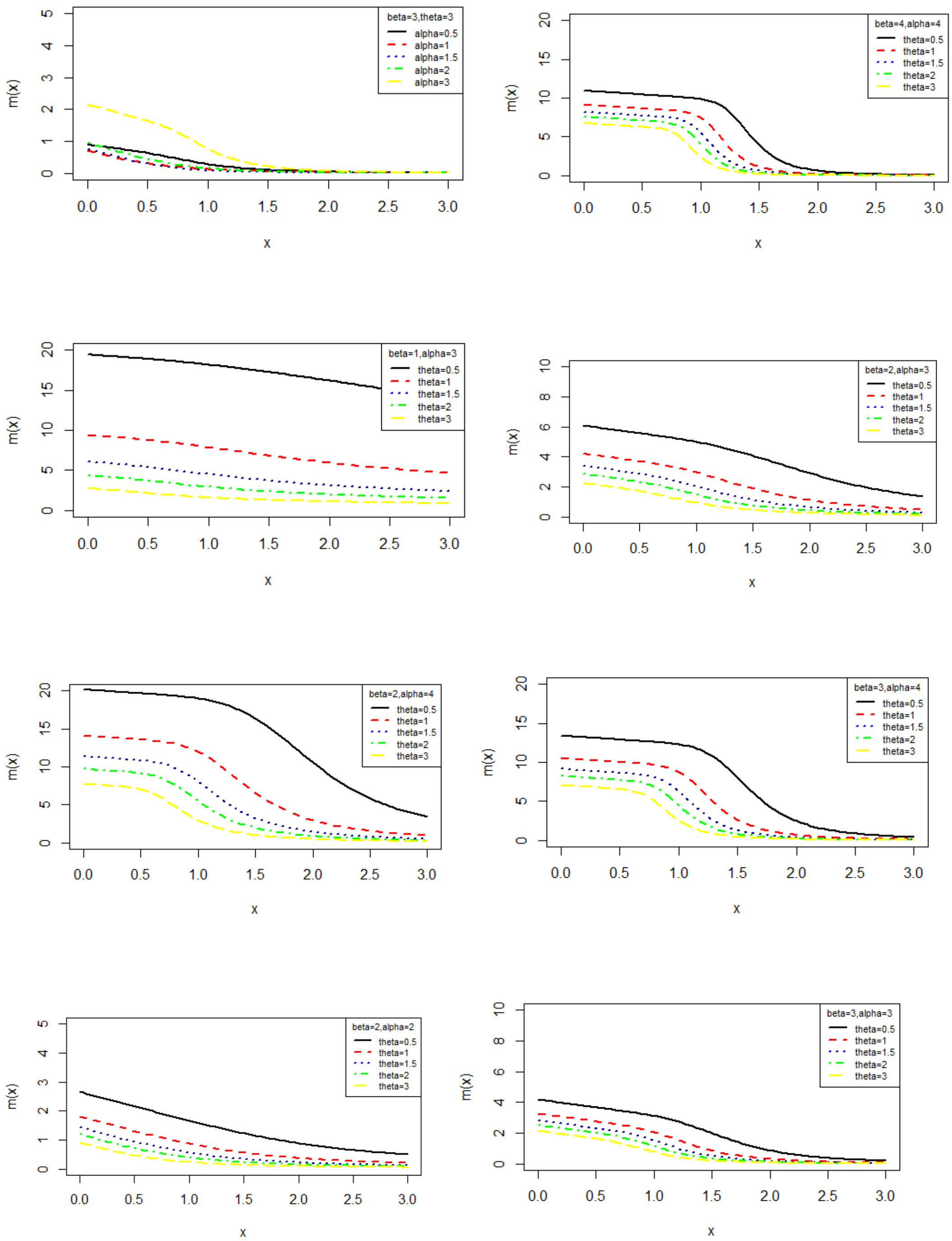

Figure 3: *Graphs of mean residual life function of PWSD for varying values of parameters* θ, α *and* β.

6. Maximum Likelihood Estimation of parameters

Suppose $(x_1, x_2, x_3, \ldots, x_n)$ be a random sample of size n from PWSD (2.1). The natural log likelihood function is thus obtained as

$$\ln L = \sum_{i=1}^{n} \ln f_8\,(x_i; \theta, \alpha, \beta)$$

$$= n[\ln \beta + (\alpha + 2) \ln \theta - \ln(\theta^2 + \alpha\theta + \alpha^2 + \alpha) - \ln \Gamma\,(\alpha)] + (\beta\alpha - 1) \sum_{i=1}^{n} \ln(x_i) + \sum_{i=1}^{n} \ln\left(1 + x_i^{\beta} + x_i^{2\beta}\right) - \theta \sum_{i=1}^{n} x_i^{\beta}.$$

The maximum likelihood estimates (MLEs) $(\hat{\theta},\hat{\alpha},\hat{\beta})$ of parameters (θ,α,β) of PWSD are the solution of the following nonlinear log likelihood equations

$$\frac{\partial\, ln\, L}{\partial\theta}=\frac{n(\alpha+2)}{\theta}-\frac{n(2\theta+\alpha)}{\theta^2+\alpha\theta+\alpha^2+\alpha}-\sum_{i=1}^{n}{x_i}^{\beta}=0$$

$$\frac{\partial\, ln\, L}{\partial\alpha}=n\, ln\,\theta-\frac{n(\theta+2\alpha+1)}{\theta^2+\alpha\theta+\alpha^2+\alpha}-n\psi(\alpha)+\beta\sum_{i=1}^{n}ln(x_i)=0$$

$$\frac{\partial\, ln\, L}{\partial\beta}=\frac{n}{\beta}+\alpha\sum_{i=1}^{n}ln\, x_i+\sum_{i=1}^{n}\frac{{x_i}^{\beta}\left(1+2{x_i}^{\beta}\right)ln(x_i)}{1+{x_i}^{\beta}+{x_i}^{2\beta}}-\theta\sum_{i=1}^{n}{x_i}^{\beta}\, ln(x_i)=0$$

where $\psi(\alpha)=\frac{d}{d\alpha}ln\,\Gamma\,(\alpha)$ is the digamma function.

These three natural log likelihood equations do not seem to be solved directly, because they cannot be expressed in closed forms. The (MLE's) $(\hat{\theta},\hat{\alpha},\hat{\beta})$ of parameters (θ,α,β) can be computed directly by solving the natural log likelihood equation using Newton-Raphson iteration method available in R-software till sufficiently close values of $\hat{\theta},\hat{\alpha}$ and $\hat{\beta}$ are obtained. The initial values of parameters θ,α and β are taken as $\theta=0.5$,$\alpha=0.5$ and $\hat{\beta}=1.5$.

7. A Simulation Study

In this section, a simulation study has been carried to check the performance of maximum likelihood estimates by taking sample sizes (n = 40,80,120,160) for values of $\theta=0.5,1.0,1.5,2.0$ and $\alpha=0.5$ $\beta=1.0$. Similarly $\alpha=0.5,2.0,4.0,5.0$&$\theta=1.5$,$\beta=1.0$, and $\beta=0.1,0.5,1.5,2.0$ & $\theta=1.5,\alpha=2.5$. Acceptance and rejection method is used to generate random number for data simulation using R-software. The process were repeated 1,000 times for the calculation of Average Bias error (ABE) and MSE (Mean square error) of parameters $\theta,\alpha$and $\beta$are presented in tables1, 2 &3 respectively.

Table1. *ABE and MSE of MLE parameters $\theta,\alpha$and β for fixed value of $\alpha=0.5$, $\beta=1.0$*

n	θ	$ABE(\hat{\theta})$	$MSE(\hat{\theta})$	$ABE(\hat{\alpha})$	$MSE(\hat{\alpha})$	$ABE(\hat{\beta})$	$MSE(\hat{\beta})$
40	0.5	0.0183	0.0134	-0.0909	0.0005	0.0249	0.0495
	1.0	0.0058	0.0014	-0.0409	0.0669	0.0198	0.0316
	1.5	-0.0066	0.0017	-0.0909	0.3305	0.0074	0.0044
	2.0	-0.0192	0.0147	-0.1159	0.5374	0.0248	0.0495
80	0.5	0.0131	0.0138	-0.0465	0.0006	0.0248	0.0495
	1.0	0.0068	0.0038	-0.0215	0.0369	0.0198	0.0316
	1.5	0.0006	0.0003	-0.0464	0.1729	0.0073	0.0043
	2.0	-0.0056	0.0025	-0.0589	0.2784	0.0011	0.0001
120	0.5	0.0132	0.0210	-0.0295	0.0002	0.0112	0.0151
	1.0	0.0091	0.0098	-0.0129	0.0199	0.0078	0.0074
	1.5	0.0049	0.0028	-0.0296	0.10491	-0.0004	0.0002
	2.0	0.0007	0.0000	-0.0379	0.1723	-0.0046	0.0025
160	0.5	0.0101	0.0164	-0.0219	0.0000	0.0080	0.0105
	1.0	0.0070	0.0078	-0.0095	0.0144	0.0056	0.0049
	1.5	0.0038	0.0024	-0.0219	0.0773	-0.0006	0.0007
	2.0	0.0007	0.0093	-0.0282	0.0127	-0.0037	0.0023

Table2. *ABE and MSE of MLE parameters $\theta, \alpha$and β for fixed value of$\theta = 1.5, \beta = 1.0$*

N	α	$ABE(\hat{\theta})$	$MSE(\hat{\theta})$	$ABE(\hat{\alpha})$	$MSE(\hat{\alpha})$	$ABE(\hat{\beta})$	$MSE(\hat{\beta})$
40	0.5	0.0157	0.0099	-0.0903	0.0003	0.0594	0.1412
	2.0	0.0033	0.0004	-0.04033	0.0650	0.0494	0.0977
	4.0	-0.0092	0.0034	-0.0903	0.3263	0.0244	0.0238
	5.0	-0.0217	0.0188	-0.1153	0.5320	0.0119	0.0056
80	0.5	0.0078	0.0049	-0.0451	0.0002	0.0297	0.0706
	2.0	0.0016	0.0002	-0.0201	0.0325	0.0247	0.0488
	4.0	-0.0046	0.0016	-0.0451	0.1632	0 .0122	0.0119
	5.0	-0.0108	0.0094	-0.0576	0.2660	0.0059	0.0028
120	0.5	0.0071	0.0060	0.0293	0.0449	0.0185	0.0413
	2.5	0.0029	0.0010	0.0168	0.0341	0.0152	0.0277
	4.0	0.0293	0.1034	0.0012	0.0002	0.0068	0.0056
	5.0	0.0377	0.1704	0.0054	0.0035	0.0027	0.0008
160	0.5	0.0056	0.0052	-0.0217	0.0000	0.0138	0.0305
	2.5	0.0025	0.0010	-0.0124	0.0245	0.0113	0.0204
	4.0	-0.0005	0.0005	-0.0217	0.0757	0.0051	0.0041
	5.0	-0.0037	0.0022	-0.0280	0.1256	0.0019	0.0005

Table 3. *ABE and MSE of MLE parameters $\theta, \alpha$and β for fixed value of $\theta = 1.5, \alpha = 2.5$*

N	β	$ABE(\hat{\theta})$	$MSE(\hat{\theta})$	$ABE(\hat{\alpha})$	$MSE(\hat{\alpha})$	$ABE(\hat{\beta})$	$MSE(\hat{\beta})$
40	0.1	0.79609	25.3479	0.7246	26.3839	0.0121	0.0059
40	0.5	0.7835519	24.5581	0.7621	23.2352	0.0021	0.0001
40	1.5	0.7710519	23.7808	0.7246	21.0050	-0.0228	0.0208
40	2.0	0.7585519	23.0160	0.6996	19.5807	-0.0353	0.0499
80	0.1	1.24654	93.2320	1.2157	97.3935	0.0051	0.0015
80	0.5	1.2382	91.9897	1.2407	92.3639	-0.0015	0.0001
80	1.5	1.2298	90.7557	1.2157	88.6793	-0.0182	0.0198
80	2.0	1.2215	89.5300	1.1990	86.2645	-0.0265	0.0422
120	0.1	0.3259	12.7464	0.3011	13.0932	0.0025	7.7334
120	0.5	0.3217	12.4226	0.3136	11.8053	-0.0007	7.5789
120	1.5	0.3175	12.1029	0.3011	10.8831	-0.0091	9.9985
120	2.0	0.3134	11.7874	0.2928	10.2891	-0.0132	2.1209
160	0.1	0.1996	6.3794	0.1856	6.8906	0.0018	5.7328
160	0.5	0.1965	6.1813	0.1950	6.0855	-0.0006	5.8974
160	1.5	0.1934	5.9863	0.1856	5.5145	-0.0068	7.5232
160	2.0	0.1903	5.7944	0.1794	05.1495	-0.0099	1.5942

8. Applications

The applications and goodness of fit of the power weighted Sujatha distribution (PWSD) has been discussed for one real dataset reported by Efron [14] relating to 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).The goodness of fit is compared with exponential distribution, Lindley distribution, Weibull distribution introduced by Weibull [15], Gamma distribution, Generalized exponential distribution (GED) introduced by Gupta and Kundu [16] , Power Lindley distribution (PLD), Weighted Lindley distribution (WLD), Power Sujatha distribution (PSD), weighted Sujatha distribution (WSD), and Generalized Lindley distribution (GLD). The dataset is as follow:

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						

In order to compare the goodness of fit of the considered distributions for the dataset, values of $-2\ln L$, AIC (Akaike information criterion), K-S Statistic (Kolmogorov-Smirnov Statistic) and p-value for two datasets have been computed The formulae for AIC and K-S Statistics are as follows: $AIC = -2\ln L + 2k$, and $K - S = \sup_x |F_n(x) - F_0(x)|$, where$k$being the number of parameters involved in the respective distributions, n is the sample size and $F_n(x)$is the empirical distribution function. The best distribution corresponds to the lower values of$-2\ln L$, AIC and K-S statistic. Note that the estimates of parameters of the considered distributions are based on maximum likelihood estimates. The initial values of the parameters for ML estimates of PWSD have been selected as$\theta = 1.5$,$\alpha = 0.5$ and$\beta = 1.5$, as the log-likelihood function is non-linear.

The ML estimate of parameters of the considered distributions for dataset is given in table 4. The goodness of fit by K-S statistics for dataset with considered distributions are presented in tables 5. From the goodness of fit given in table 5, it is crystal clear that PWSD gives much closure fit to dataset relating to survival times of patients suffering from head and neck cancer and hence it can be considered as an important distribution for modeling data relating to survival times of patients suffering from head and neck cancer. The variance-covariance matrix of the parameters (θ, α, β) of PWSD for dataset is given in table 6. The survival plots for the dataset, fitted plot of cdf for given dataset, and the fitted plot of given dataset are presented in figures 4, 5 and 6 respectively.

Table 4: *Summary of the ML estimates of parameters for dataset*

Model	ML Estimates		
	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$
PWSD	26.9742	50.3682	0.1325
GLD	0.0047	0.0524	5.0750
PSD	0.1539	………	0.5690
WSD	0.0089	0.01637	…….
WLD	0.00531	0.21191	……
PLD	0.05301	0.68893	…….
GED	0.00482	1.09367	……
Gamma	0.00489	1.08501	……
Weibull	0.00710	0.92327	…….
Lindley	0.00892	…….	……..
Exponential	0.00447	……	…….

Table 5: *Summary of Goodness of fit by K-S Statistic for dataset*

Model	$-2\ln L$	AIC	K-S	p-value
PWSD	555.41	561.41	0.104	0.572
GLD	564.09	570.09	0.150	0.248
PSD	559.45	563.45	0.135	0.512
WSD	579.96	583.96	0.350	0.000
WLD	565.91	569.91	0.161	0.181
PLD	560.78	564.78	0.118	0.529
GED	563.93	567.93	0.145	0.280
Gamma	564.10	568.10	0.149	0.249
Weibull	563.71	567.71	0.298	0.005
Lindley	579.16	581.16	0.219	0.025
Exponential	564.01	566.01	0.145	0.282

Table 6: *Variance-covariance matrix of the parameters θ, α and β of PWSD for dataset*

$$\begin{array}{c} \\ \hat{\theta} \\ \hat{\alpha} \\ \hat{\beta} \end{array} \begin{array}{c} \hat{\theta}\hat{\alpha}\hat{\beta} \\ \begin{bmatrix} 983.70956 & 1475.4914 & -2.33334 \\ 1475.4914 & 2219.87726 & -3.47289 \\ -2.33334 & -3.47289 & 0.005661 \end{bmatrix} \end{array}$$

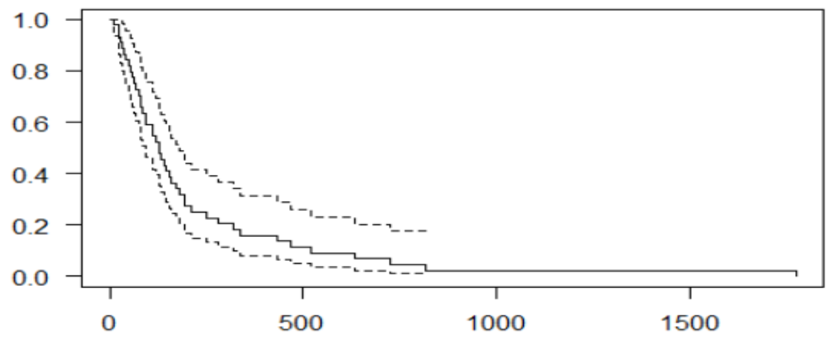

Figure 4. *Survival plots for the given dataset.*

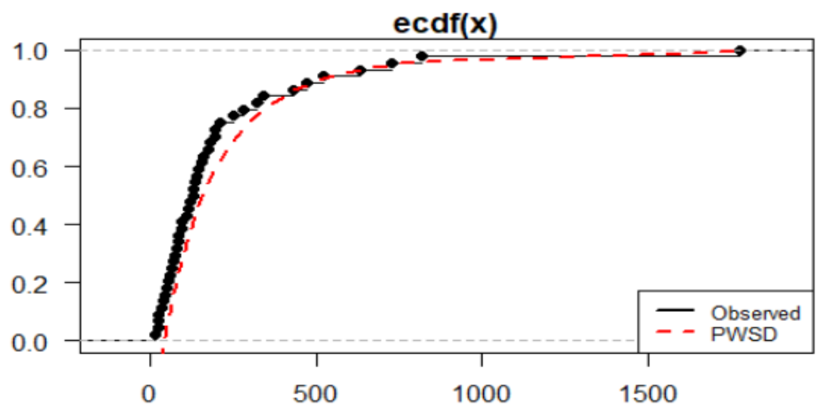

Figure 5. *Fitted plot of CDF on given dataset*

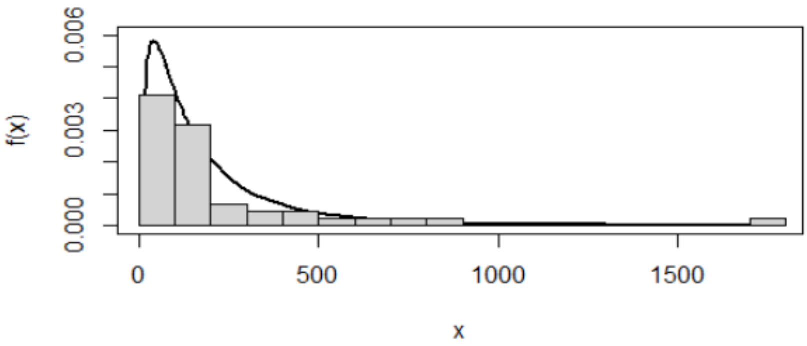

Figure 6. *Fitted plot of pdf of PWSD on dataset*

9. Concluding Remarks

In this paper a three-parameter power weighted Sujatha distribution (PWSD) which includes weighted Sujatha distribution (WSD), power Sujatha distribution (PSD) and Sujatha distribution as particular cases, has been proposed. Its statistical properties including behavior it density function, moments, hazard rate function, mean residual life function have been discussed. Estimation of parameters has been discussed using the method of maximum likelihood. A simulation study has been presented to know the performance of maximum likelihood estimates of parameters. Application of the proposed distribution have been explained with a real lifetime data which is related to patients suffering from head and neck cancer and its goodness of fit has been compared with other lifetime distributions. The proposed distribution shows better fit over the other one parameter, two-parameter and three-parameter distributions.

References

[1] Shanker, R (2016a): Sujatha distribution and its applications, *Statistics in Transition New series*, 17(3), 1-20

[2] Shanker, R. and Shukla, K.K. (2018): A two-parameter Weighted Sujatha distribution and Its Application to Model Lifetime data, International *Journal of Applied Mathematics and statistics*, 57 (3), 106 – 121

[3] Shanker, R. and Shukla, K. K. (2019): A two-parameter Power Sujatha distribution with properties and Application, *International Journal of Mathematics and Statistics*, 20(3), 11 – 22

[4] Ghitany, M.E., Alqallaf ,F., Al-Mutairi, D.K., Husain, H.A. (2011): A two-parameter weighted Lindley distribution and its applications to survival data, *Mathematics and Computers in simulation*, 81, 1190-1201.

[5] Shanker, R., Shukla, K.K. and Hagos, F. (2016): On Weighted Lindley distribution and Its Applications to Model Lifetime Data, Jacobs Journal of Biostatistics,1(1), 1- 9

[6] Ghitany, M.E., Al-Mutairi, D.K., Balakrishnan, N., and Al-Enezi, L.J. (2013): Power Lindley distribution and Associated Inference, Computational Statistics and Data Analysis 64, 20 –33.

[7] Lindley, D.V. (1958): Fiducial distributions and Bayes' Theorem, *Journal of the Royal Statistical Society, Series* B, 20, 102 – 107.

[8] Ghitany, M.E. , Atieh, B. and Nadarajah, S. (2008): Lindley distribution and its Application, *Mathematics Computing and Simulation*, 78, 493 – 506.

[9] Shanker, R., Hagos, F, and Sujatha, S. (2015): On modeling of Lifetimes data using exponential and Lindley distributions, *Biometrics & Biostatistics International Journal*, 2 (5), 1-9.

[10] Zakerzadeh, H. and Dolati, A. (2009): Generalized Lindley distribution, *Journal of Mathematical extension*, 3 (2), 13 – 25.

[11] Shanker, R. (2016b): On Generalized Lindley Distribution and its Applications to model lifetime data from B*iomedical science and engineering, Insights in Biomedicine*, 1(2), 1 – 6

[12] Stacy, E.W. (1962): A generalization of the gamma distribution, *Annals of Mathematical Statistical*, 33, 1187 – 1192

[13] Shanker, R. and Shukla, K.K. (2016): On modeling of Lifetime data using three-parameter generalized Lindley and generalized gamma distributions, *Biometrics & Biostatistics International journal*, 4(7), 283 -288

[14] Weibull, W. (1951): A statistical distribution of wide applicability, *Journal of Applied Mathematics*, 18, 293 – 297.

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

[16] Gupta, R.D. and Kundu, D. (1999): Generalized Exponential Distribution, *Austalian & New Zealand Journal of Statistics*, 41(2), 173 – 188.