Journal of Population Therapeutics & Clinical Pharmacology

RESEARCH ARTICLE
DOI: 10.53555/jptcp.v31i1.4013

MODELING BREAST CANCER DATA USING A NOVEL THREE PARAMETER GUL ALPHA POWER EXPONENTIAL MODEL

Abdur Rehman[1], Alamgir[2*], Sajjad Ahmad Khan[3], Gul Karam Khan[4], Muhammad Ilyas[5], Sundas Hussain[6]

[1]Department of Statistics, University of Peshawar, abdurrehmanshad@uop.edu.pk
[2*]Department of Statistics, University of Peshawar, alamgir_khalil@uop.edu.pk
[3]Department of Statistics, Islamia College Peshawar, sajjadkhan@icp.edu.pk
[4]Gajju Khan Medical College Swabi, gulkaramkhann@gmail.com
[5]Department of Statistics, University of Malakand, KP, ilyas.farhan@uom.edu.pk
[6]Department of Statistics, Shaheed Benazir Bhutto Women University Peshawar, sundashussain@sbbwu.edu.pk

[*]**Corresponding Author:** Alamgir
alamgir_khalil@uop.edu.pk

ABSTRACT

Breast Cancer is a common type of cancer found in women. In this paper, we model breast cancer data using a new three parameter family of distribution, called the Three Parameter Gull Alpha Power Exponential (TP-GAPE) Distribution. The new distribution is more useful because it can correspond to various hazard rate functions, which are widely used in reliability investigations. With the help of the proposed model, data with rising, uni-modal, and modified uni-modal hazard rate functions can be analyzed. We determine the essential statistical and reliability properties of the suggested model. The goodness-of-fit criteria have been studied using two real-life data sets. The proposed model is compared to other current modifications of the exponential distribution with the aim of evaluating the model's efficacy using a variety of goodness of fit measures, including the Akaike Information Criterion, Bayesian Information Criterion, etc. These results suggest that the proposed model fits the cancer data as well as some other scientific data more precisely than any recently developed extensions of exponential distribution.

Keywords: Akaike Information Criterion, Bayesian Information Criterion, Anderson-Darling estimation; Cramér-von Mises estimation; data analysis; exponential distribution; mean residual life.

INTRODUCTION

Applications for lifetime data analysis can be extended to a number of areas, including business, engineering, finance, and health [1, 2]. The key objective of such analyses is often to simulate the probability distribution of the time to an event and/or the factors influencing the time to an event of interest. Various probability models, such as log-logistic, beta, gamma, Weibull, exponential, and many more, are available for modeling lifetime data and in several cases.These traditional approaches [3, 4] are inadequate for modeling lifetime data, prompting the use of updated iterations of current distributions [5].

The modeling of data, including survival analysis, is frequently performed through statistical distributions. The literature focuses a lot on the exponential, Weibull, Rayleigh, gamma, and lognormal distributions because they are some of the most flexible distributions implemented in survival analysis. These distributions may not be enough for modeling, however, considering the many different ways that data can be generated [6].

The development of new probability model in statistical distribution theory frequently involves the addition of a new parameter to an existing family of distribution functions. A class of distribution functions become more flexible by adding an additional parameter, which can be very valuable whenever performing data analysis. For instance, in order to provide the normal distribution greater flexibility, Azzalini [7] devised the skew-normal distribution by adding one additional parameter. A two-parameter the exponentiated Weibull model was suggested by Mudholkar and Srivastava [8] consisting of two shape parameters and one scale parameter. The new exponentiated Weibull model is able to adapt better than the two-parameter Weibull model since it has an additional shape parameter.Marshall and Olkin [9] and Eugene et al. [10] proposed methods to add additional parameters to any distribution function.Later, several other exponentiated distributions have been introduced by several authors. As a result, the exponential distribution has been frequently employed to model data sets from survival analysis. The exponential distribution's problem is that it can only accurately convey data with a constant hazard function.

In order to increase flexibility of exponential family, the purpose of this research is to add an additional parameter to a family of distribution functions. The new model is termed as "Three Parameter Gull alpha Power Exponential (TP-GAPE) distribution ". As a result of the suggested TP-GAPE method's elegance, it can be utilized for analyzing many data sets very effectively. The Weibull, Gamma, or GE distributions can have forms that are comparable to the PDF and hazard functions of the TP-GAPE distribution. The well-known Weibull, Gamma, or GE distributions can thus be replaced with it. Since the CDF of the TP-GAPE distribution can be written clearly, it may be used relatively easily to the analysis of censored data as well. Moreover, we deal with the maximum likelihood estimation method for the unknown parameters and introduce the three-parameter TP-GAPE distribution, which is mostly used for data analysis. For illustration purpose, an analysis of an actual dataset has been implemented.

The Proposed TP-GAP Method

Ijaz [11] introduced a Gull Power Alpha family of distribution. For some other families, we refer to [12-24]. In this article, a new family (TP-GAP) is presented by introducing a new shape parameter "γ" which is given by

$$F(y,\alpha,\gamma)=\frac{\alpha^{\gamma}[F(y)]^{\gamma}}{\alpha^{\gamma F(y)}} \text{ If } \alpha > 1 \qquad (1)$$

Where F(y) represents the CDF of the baseline distribution and probability density function of *f*(*y*) is

$$f(\mathcal{Y},\alpha,\gamma)=\gamma\alpha^{\gamma[1-\mathcal{F}(\mathcal{Y})]}[\mathcal{F}(\mathcal{Y}]^{\gamma}\,\mathcal{F}(\mathcal{Y})\left(\frac{1}{\mathcal{F}(\mathcal{Y})}-ln\alpha\right) \qquad (2)$$

TP-GAPE distribution and its statistical properties

This section uses the CDF of the exponential distribution known as "a three parameter Gull alpha power exponential distribution" (TP-GAPE) to illustrate the unique form of TP-GAPE. The cumulative distribution function (CDF) of the exponential distribution is presented by

$$F(Y)=1-e^{-\lambda y} \qquad y>0 \qquad (3)$$

Where λ in the above equation represents scale parameter.

By replacing (3) in (1), the CDF and PDF of TP-GAPE are, respectively, given by

$$F(\mathcal{Y})=\alpha^{\gamma e^{-\lambda \mathcal{Y}}}\left[1-e^{-\lambda \mathcal{Y}}\right]^{\gamma}, y>0 \;\&\alpha;\ \beta;\ \gamma>0 \qquad (4)$$

$$f(y,\alpha,\lambda,\gamma)=\lambda\gamma\alpha^{\gamma e^{-\lambda \mathcal{Y}}}\left[1-e^{-\lambda y}\right]^{\gamma}e^{-\lambda \mathcal{Y}}\left[\frac{1}{1-e^{-\lambda y}}-ln\alpha\right] \quad , y>0 \qquad (5)$$

Figure 1 and 2 presents numerous shapes of the CDF and PDF with different set of parameter values.

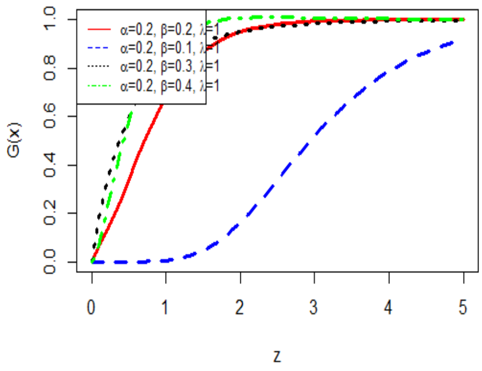


Figure 1: Cumulative distribution function of TP-GAPE distribution

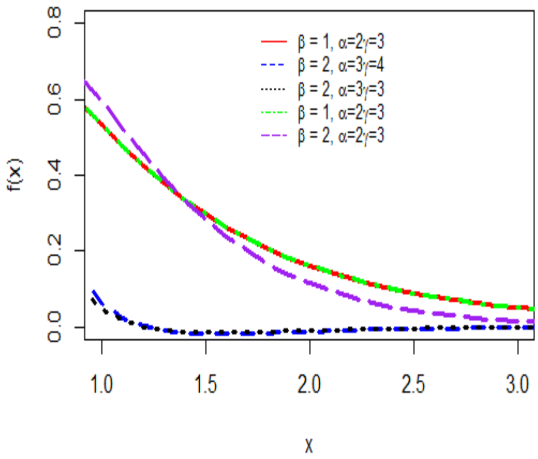


Figure 2: Probability density functions of TP-GAPE distribution.

The Survival and Hazard Rate Function.
The survival and hazard rate function of TP-GAPE is defined by

$\mathcal{S}(y) = 1 - F(\mathcal{Y})$And using equation (4) we get

$$\mathcal{S}(y) = 1 - \left(\alpha^{\gamma e^{-\lambda y}}\left[1 - e^{-\lambda y}\right]^{\gamma}\right) \qquad (6)$$

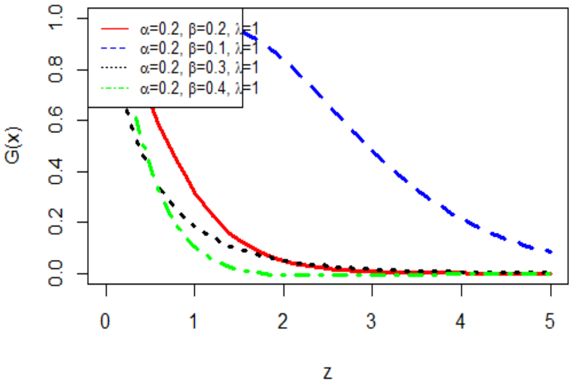


Figure 3: Survival functions of TP-GAPE

The hazard rate function of TP-GAPE is provided by

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

Using equation (5) & equation (6), we get

$$\mathcal{H}(y) = \frac{\lambda\gamma\alpha^{\gamma e^{-\lambda y}}\left[1-e^{-\lambda y}\right]^{\gamma} e^{-\lambda y}\left[\frac{1}{1-e^{-\lambda y}} - ln\alpha\right]}{1-\left(\alpha^{\gamma e^{-\lambda y}}\left[1-e^{-\lambda y}\right]^{\gamma}\right)} \quad (7)$$

Figure 4 demonstrates the nature of the hazard rate function for different values of parameter.

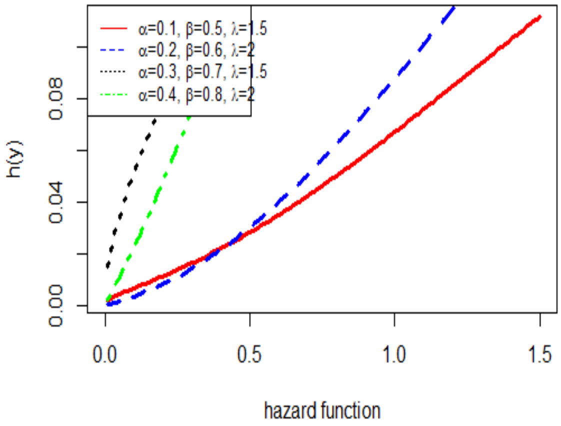


Figure 4: Hazard function of TP-GAPE distribution

The Quantile Function and Median of TP-GAPE distribution

The quantile function, an inverse CDF function, is used to calculate the median, octal, decile, percentiles, and quantile, among various other measurements. The TP-GAPE distribution's quantile function is obtained as

$$F(y) = P(Y \leq y) = u$$

Where "u"is a standard uniform random variable, Substituting (4), we get the result

$$\gamma\alpha^{\gamma e^{-\lambda y}}\left[1 - e^{-\lambda y}\right]^{\gamma} = u \quad (8)$$

The solution of equation (8) for $y$will give the following result:

$$Y = \frac{\log\left[\frac{\log(\beta)}{w\left(-\alpha\sqrt{\frac{\beta-\alpha^{u}}{\alpha}}log(\beta)\right)+lob(\beta)}\right]}{\gamma} \quad (9)$$

For median, consider u=0.5 in equation (9).

Order Statistics

Let $y_1, y_2, y_{3,\ldots} y_n$be ordered random variables from NFW, then the PDF of the **HQIC** order statistic is given by

$$f_{i;n}(y) = \frac{n!}{(i-1)!(n-i)!} f(y)(F(y))^{-(i-1)}[1 - F(y)]^{-(n-1)} \quad (10)$$

Using equation (4) and equation (5), of TP-GAPE, we get

$$f_{i;n}(y) = \frac{n!}{(i-1)!(n-i)!}\left[\lambda\gamma\alpha^{\gamma e^{-\lambda y}}\left[1 - e^{-\lambda y}\right]^{\gamma} e^{-\lambda y}\left[\frac{1}{1-e^{-\lambda y}} - ln\alpha\right]\right]\left[\alpha^{\gamma e^{-\lambda y}}\left[1 - e^{-\lambda y}\right]^{\gamma}\right]^{i-1}\left[1 - \alpha^{\gamma e^{-\lambda y}}\left[1 - e^{-\lambda y}\right]^{\gamma}\right]^{n-i} \quad (11)$$

Using equation (11), the smallest and largest order statistic of TP-GAPE is defined by

$$\mathrm{f}(1;\mathcal{Y}) = \frac{\mathrm{n}!}{(1-1)!(\mathrm{n}-\mathrm{i})!}\left[\lambda\gamma\alpha^{\gamma \mathrm{e}^{-\lambda y}}\left[1 - \mathrm{e}^{-\lambda y}\right]^{\gamma} \mathrm{e}^{-\lambda y}\left[\frac{1}{1-\mathrm{e}^{-\lambda y}} - \ln\alpha\right]\right]\left[\alpha^{\gamma \mathrm{e}^{-\lambda y}}\left[1 - \mathrm{e}^{-\lambda y}\right]^{\gamma}\right]^{1-1}\left[1 - \alpha^{\gamma \mathrm{e}^{-\lambda y}}\left[1 - \mathrm{e}^{-\lambda y}\right]^{\gamma}\right]^{n-1} \quad (12)$$

and the largest order statistic is;

$$\boldsymbol{f}(n;n)(\mathcal{Y}) = n\left\{\lambda\gamma\alpha^{\gamma e^{-\lambda y}}\left[1-e^{-\lambda y}\right]^{\gamma} e^{-\lambda y}\left[\frac{1}{1-e^{-\lambda y}} - ln\alpha\right]\right\}\left[\mathbf{1} - \alpha^{\gamma e^{-\lambda y}}\left[1-e^{-\lambda y}\right]^{\gamma}\right]^{\boldsymbol{n-1}} \quad (13)$$

PDF of Median, smallest and largest order statistics

Let us consider a random sample of size k from aTP-GAPE distribution having Parameters α, β and γ. Then the smallest order statistics is given by:

$$g_{1-k}(y) = kg(y)[1-G(y)]^{k-1} \quad (14)$$

By putting the equation (4) and equation (5) of TP-GAPE distribution, we have

$$g_{1-k}(y) = k\left(\lambda\gamma\alpha^{\gamma e^{-\lambda y}}\left(1-e^{-\lambda y}\right)^{\gamma} e^{-\lambda y}\left(\frac{1}{1-e^{-\lambda y}} - ln\alpha\right)\right)\left(1 - \alpha^{\gamma e^{-\lambda y}}\left(1-e^{-\lambda y}\right)^{\gamma}\right)^{k-1} \quad (15)$$

Order statistics for median

We know that

$$g_{(m+k)}(\tilde{y}) = \frac{(2m+1)!}{n!m!}\left(g(\tilde{y})\{G(\tilde{y})\}^{n}\right)\{1-G(\tilde{y})\}^{m} \quad (16)$$

$$g_{(m+k)}(\tilde{y}) = \frac{(2m+1)!}{n!m!}\left[\lambda\gamma\alpha^{\gamma e^{-\lambda y}}\left(1-e^{-\lambda y}\right)^{\gamma} e^{-\lambda y}(-\ln\alpha)\right]\left(\alpha^{\gamma e^{-\lambda y}}\left(1 - e^{-\lambda y}\right)^{\gamma}\right)^{m}\left(\begin{matrix}1-\\ \alpha^{\gamma e^{-\lambda y}}\left(1-e^{-\lambda y}\right)^{\gamma}\end{matrix}\right)^{m} \quad (17)$$

Density of the Maximum order statistics:

$$g_{1-k}(y) = kg[G(y)]^{k-1} \quad (18)$$

By putting the equation (4) and equation (5) of TP-GAPE distribution, we get

$$g_{k-k}(y) = k\left(\lambda\gamma\alpha^{\gamma e^{-\lambda y}}\left(1-e^{-\lambda y}\right)^{\gamma} e^{-\lambda y}\left(\frac{1}{1-e^{-\lambda y}} - \ln\alpha\right)\right)\left(\alpha^{\gamma e^{-\lambda y}}\left(1-e^{-\lambda y}\right)^{\gamma}\right)^{k-1} \quad (19)$$

Joint PDF of smallest and largest order statistics

We can derive the joint density for the smallest and largest order statistics of TP-GAPE distribution. The expression for joint density is as follows:

$$g_{i.j.k}\left(x_i, x_j\right) = \frac{k_i}{(i-1)!(j-i-1)!(k-j)!} \quad (20)$$

Putting CDF and PDF of TP-GAPE distribution in equation (20), we have the resultant form,

Let $w=\frac{k_i}{(i-1)!(j-i-1)!(k-j)!}$ (21)

$$g_{i.j.k}\left(x_i, x_j\right) = w\left[\alpha^{\gamma e^{-\lambda y_i}}\left(1-e^{-\lambda y_i}\right)^{\gamma}\right]^{i-j}\left[\left(\alpha^{\gamma e^{-\lambda y_i}}\left(1-e^{-\lambda y_i}\right)^{\gamma}\right) - \left(\alpha^{\gamma e^{-\lambda y_i}}\left(1 - e^{-\lambda y_i}\right)^{\gamma}\right)\right]^{j-i-1}\left[-\alpha^{\gamma e^{-\lambda y_i}}\left(1-e^{-\lambda y_i}\right)^{\gamma}\right]^{k-j}\left[\lambda\gamma\alpha^{\gamma e^{-\lambda y_i}}\left(1-e^{-\lambda y_i}\right)^{\gamma} e^{-\lambda y_i}\left(\frac{1}{1-e^{-\lambda y_i}} - \ln\alpha\right)\right] \quad (22)$$

For special case if we put k=j and i=1 we will have the following form of minimum and maximum order statistics.

$$g_{1:j:j}(x_1, x_k) = \frac{j!}{(1-1)!(j-1-1)!(j-j)!}[G(x_1)]^{1-1}.\left[G(x_j) - G(x_1)\right]^{j-1-1}.\left[1-G\left(x_j\right)\right]^{j-j} g(x_1)g(j) \quad (23)$$

$$g_{1:j:j}(x_1, x_k) = j(j-1)\left[G(x_j) - G(x_i)\right]^{j-2} g(x_1)g(x_j) \quad (24)$$

$$g_{1:j:j}(x_1, x_k) = j(j-1)\left[\left(\alpha^{\gamma e^{-\lambda y_i}}\left(1-e^{-\lambda y_i}\right)^{\gamma}\right).g(x_1)g(x_j)\right] \quad (25)$$

Skewness and Kurtosis of TP-GAPE

The Bowley's Skewness and Moors kurtosis are

$$SK = \frac{\mu 3 - 3\mu 2'\mu 1' + 2\mu 1'^3}{\left(\mu 2' - \mu 1'^2\right)^{\frac{3}{2}}} \ \textbf{OR} \qquad SK = \frac{Q\left(\frac{3}{4}\right) + Q\left(\frac{1}{4}\right) - 2Q\left(\frac{1}{2}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)} \quad (26)$$

$$KU = \frac{\mu 4' - 4\mu 3'\mu 1' + 6\mu 2'\mu 1'^2 - 3\mu 1'^4}{\left(\mu 2' - \mu 1'^2\right)^2} \quad \textbf{OR} \quad KU = \frac{Q\left(\frac{7}{8}\right) + Q\left(\frac{3}{8}\right) - Q\left(\frac{5}{8}\right) - Q\left(\frac{1}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{2}{8}\right)} \tag{27}$$

APPLICATIONS

Two real data sets with extreme values are studied, one with non-monotonic and the other with monotonic hazard rate shapes, in order to evaluate the performance of the suggested model. To assess the effectiveness of the suggested model, several goodness of fit statistics were taken into assessment, including the Akaike information criteria (AIC), Hannan and Quinn information criteria (HQIC), Anderson darling (A), Cramer-von Mises (W), consistent Akaike information criteria (CAIC), and Bayesian information criteria (BIC). The aforementioned circumstances are described mathematically as follows:

Where "xi" is the random sample, $L = L(\Psi, X_i)$ is the maximized likelihood function is the MLE, and "p" is the number of parameters in the model. Typically, the probability model that satisfies a smaller number of these various prerequisites has been considered to be the best fitted one.

Data Set 1: Kevlar 49/epoxy strands under 90 percent pressure.

The data set comprises 101 observations corresponding to the failure time in hours of Kevlar 49/epoxy strands with pressure at 90%. The data set displayed in Table 1can be found in [25, 26].

Table1. Kevlar 49/epoxy strands under 90 percent pressure
0.01 ,0.01, 0.02, 0.02, 0.02, 0.03, 0.03, 0.04, 0.05, 0.06, 0.07, 0.07, 0.08, 0.09, 0.09, 0.10, 0.10, 0.11, 0.11, 0.12, 0.13, 0.18, 0.19, 0.20, 0.23, 0.24, 0.24, 0.29, 0.34, 0.35, 0.36, 0.38, 0.40, 0.42, 0.43, 0.52, 0.54, 0.56, 0.60, 0.60, 0.63, 0.65, 0.67, 0.68, 0.72, 0.72, 0.72, 0.73, 0.79, 0.79, 0.80, 0.80, 0.83, 0.85, 0.90, 0.92, 0.95, 0.99, 1, 1.01, 1.02, 1.03, 1.05, 1.10, 1.10, 1.11, 1.15, 1.18, 1.20, 1.29, 1.31, 1.33, 1.34, 1.40, 1.43, 1.45, 1.50, 1.51, 1.52, 1.53, 1.54, 1.54, 1.55, 1.58, 1.60, 1.63, 1.64, 1.80, 1.80, 1.81, 2.02, 2.05, 2.14, 2.17, 2.33, 3.03, 3.03, 3.34, 4.20, 4.69, 7.89

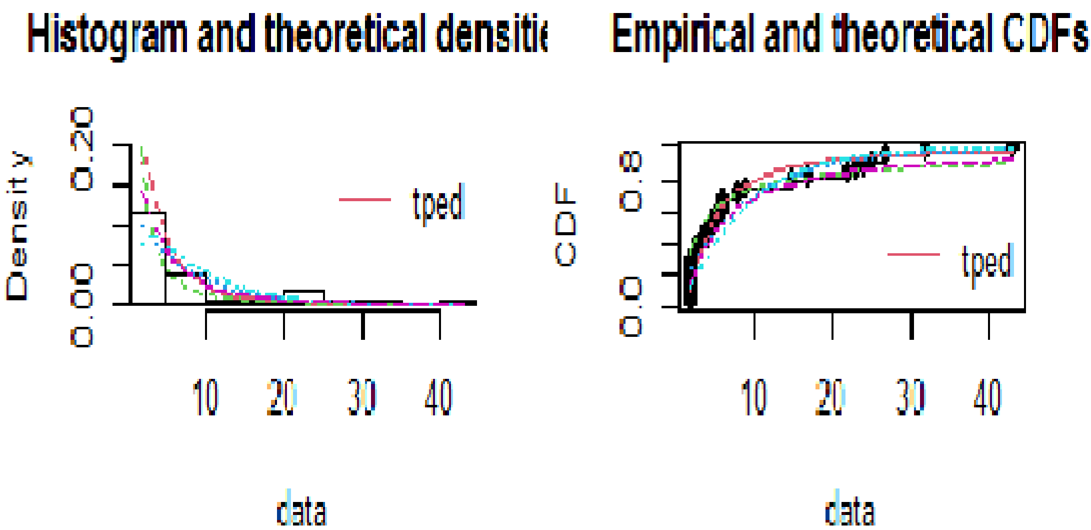


Fig. 5 Empirical and fitted density and CDF plots of Kevlar data

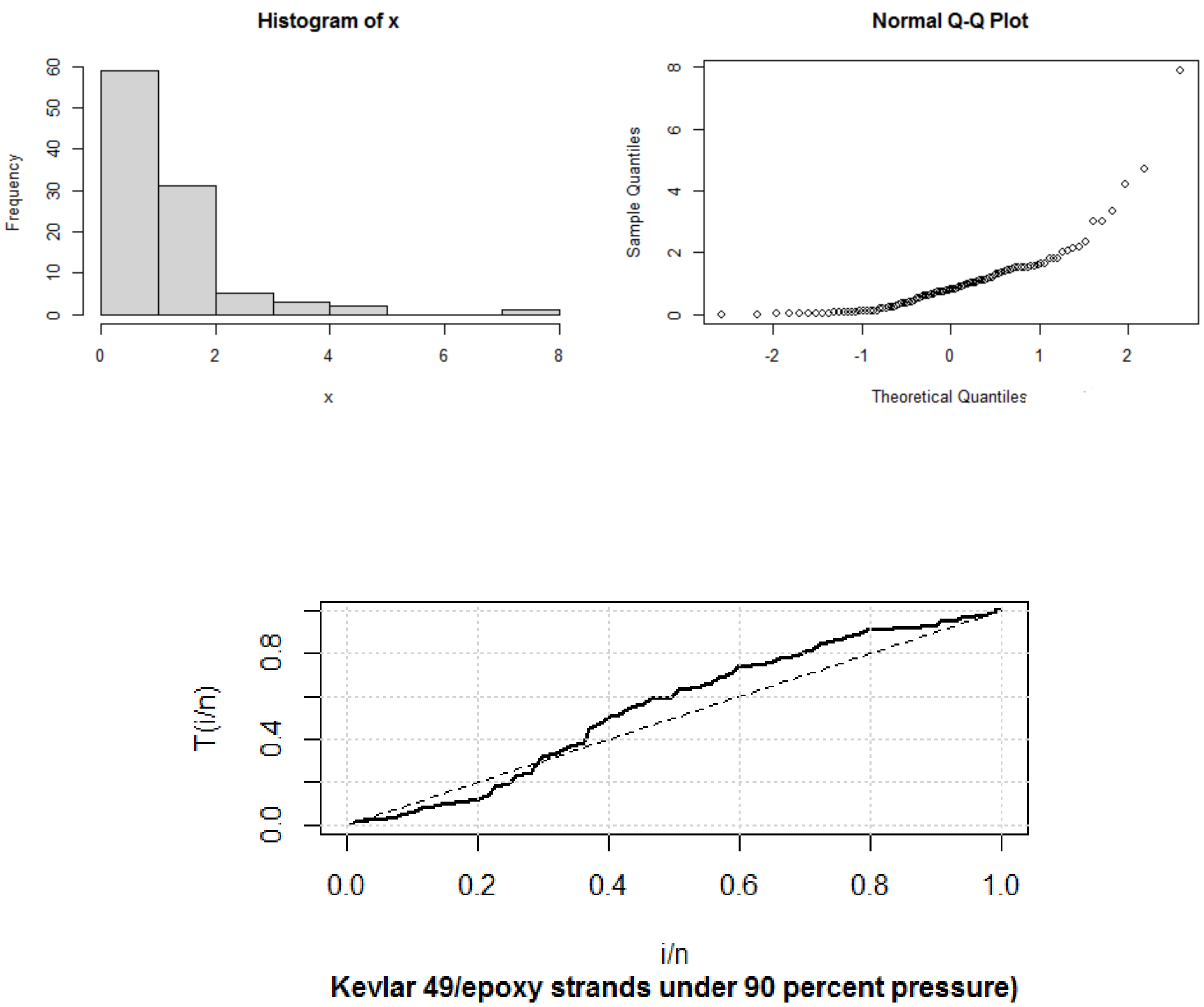


Figure.6 Q-Q, TTTand Histogram plot for Kevlar 49/epoxy strands under 90 percent pressure

Table2. MLE and standard error of each distribution parameters of a data failure times of Kevlar 49/epoxy strands under than 90 percent pressure.

MODELS	*TPGAPE*	*EGPIE*	*EGBIE*	*EGGIE*	*EGLIE*
MLE	0.4744 1.0609 0.7170	26.062 7.320 0.175 0.002	11.644 8.862 0.313 0.003	0.664 20.525 0.498 0.002	0.018 19.277 0.616 0.002
STANDARD ERROR	*0.3560* *0.1672* *0.1669*	0.009 1.770 0.019 0.002	(1.925×10−5) (1.136×10−4) (2.141×10−2) (8.755×10−5)	(2.412×10−1) (4.798×10−3) (1.360×10−1) (9.954×10−4)	(5.004×10−1) (3.845×10−3) (6.431×10−2) (6.301×10−4)

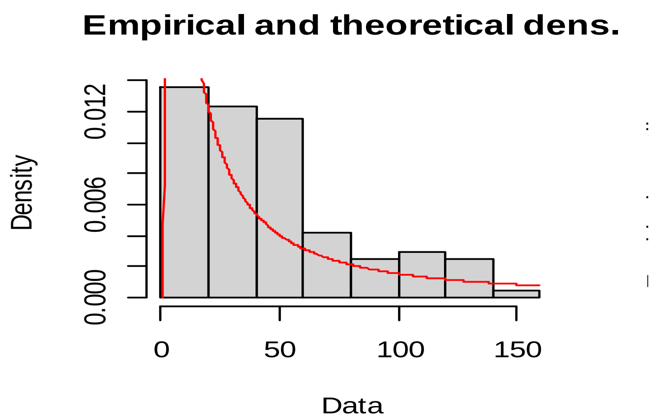
Table3. Goodness of fit criteria for failure times of Kevlar 49/epoxy strands under 90 percent pressure.

Model	**Log-likelihood**	**AIC**	**AICc**	**BIC**	**A**	**W**	**K-S**	**Ranks**
TP-GAPE	**-**	**210.5921**	**210.8395**	**218.4375**	**0.7101897**	**0.10929**	**NA**	**1**
EGPIE	−116.660	241.314	241.946	251.774	NA	0.738	0.182	2
EGBIE	−122.930	253.868	254.500	264.328	NA	0.926	0.195	3
EGGIE	−140.090	288.170	288.802	298.631	NA	1.386	0.237	4
EGLIE	−134.010	276.025	276.657	286.486	NA	1.211	0.203	5

The TP-GAPE distribution utilizing the Kevlar data is shown in Figure 5 along with the empirically determined and fitted density and CDF charts. The results for Kevlar are plotted in Figure 6 using a Q-Q plot, TTT plot and Histogram. Table 2 shows the estimates with the highest probability and their standard errors. The goodness of fit measurements for strands made of Kevlar 49 and epoxy at 90 percent pressure are shown in Table 3. In comparison to the other models, the TP-GAPE distribution provides a better fit to the data set. The TP-GAPE distribution, when compared to the other fitted models, has the highest log-likelihood and the lowest values for W, AIC, AICc, and BIC, as shown in Table 3.

Data Set 2: Breast cancer patients Data acquired from a large hospital between 1929 and 1938. The data set comprises 121 breast cancer patients acquired from a large hospital between 1929 and 1938. The data set displayed in Table 4 can be found in [27].

Table 4. 121 breast cancer patients acquired from a large hospital between 1929 and 1938.
0.3,0.3,4.0,5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3,13.5,14.4,14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8,20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0,31.0, 31.0, 32.0,35.0, 35.0,37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0,42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0,54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0,69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0,115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0, 154.0



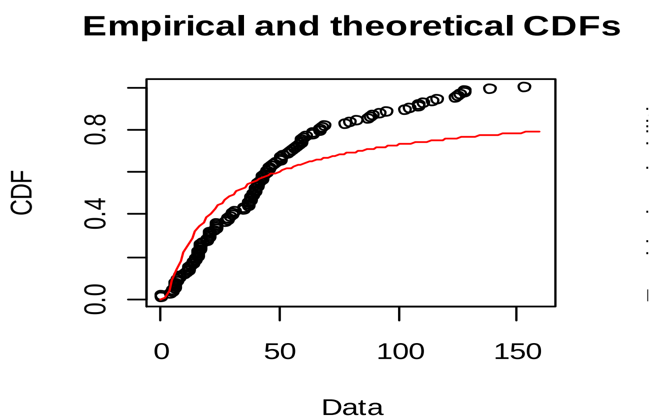


Figure 7: Empirical and theoretical densities and Empirical and theoretical CDFs

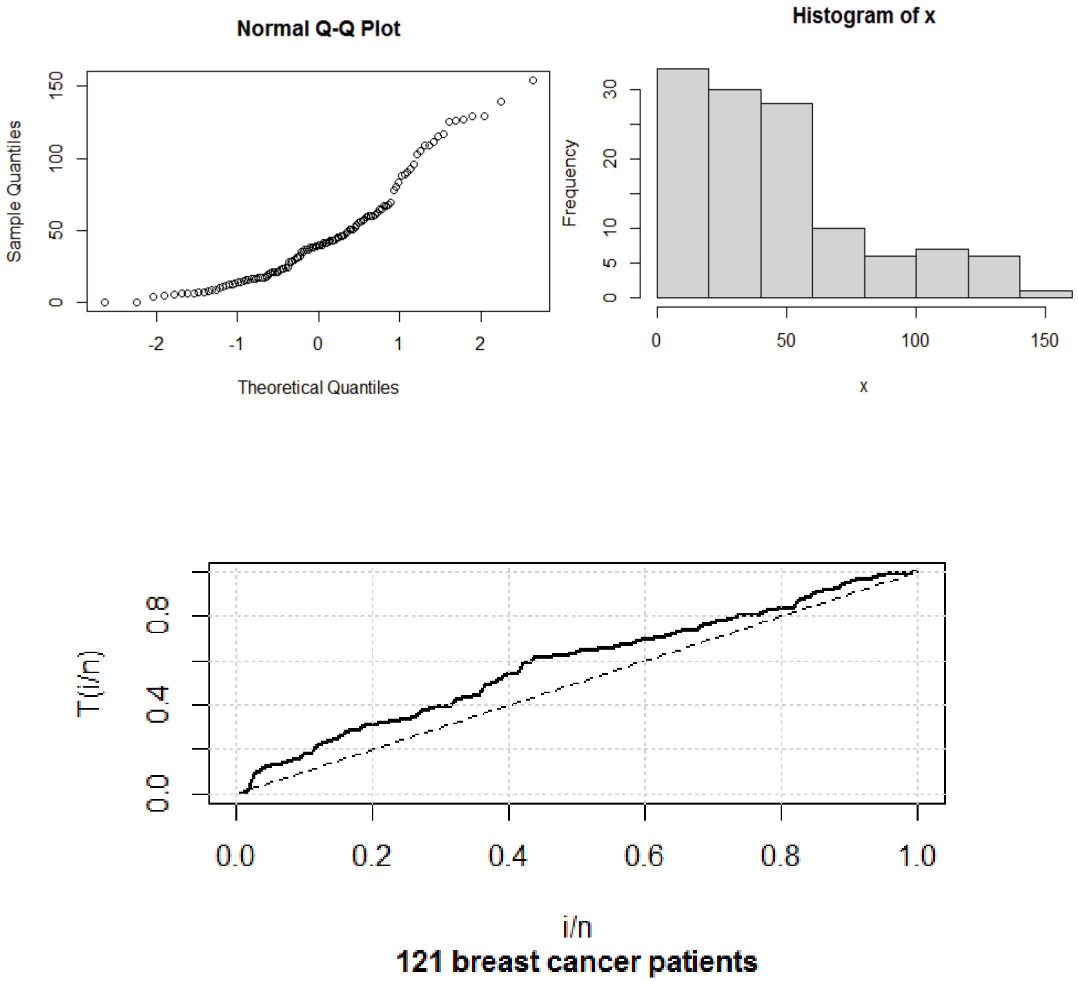


Figure.8 Q-Q, TTT and Histogram plot for 121 breast cancer patients acquired from a large hospital between 1929 and 1938.

Table5. Descriptive Statistics of 121 breast cancer patients acquired from a large hospital between 1929 and 1938.

No of samples	maximum	minimum	mean	Median	variance	Kurtosis	Skewness
121	154	0.30	46.33	40	1244.464	3.402139	1.04318

Table6. Goodness of fit criteria 121 breast cancer patients acquired from a large hospital between 1929 and 1938.

Distribution	Parameters	Log-likehood	AIC	Rank
TP-GAPE	**0.36227649, 0.3190185, 1.0506364**	**NA**	**1164.204**	**1**
EIE	*0.350733* 0.007599	-584.9013	*1173.803*	*2*
IE	*10.3215*	*-677.2791*	*1,356.558*	*3*

Figure 5 shows the Empirical and theoretical densities of PDF and Empirical and theoretical CDFs Charts of the TP-GAPE distribution using the Breast cancer patient's data. Figure 6 shows the Q-Q plot, TTT plot and Histogram plot for 121 breast cancer patients acquired from a large hospital between (1929 and 1938) data. In Table 5, the breast cancer data set is positively skewed, with a coefficient of skew-ness of 1.04318 and a variance of 1,244.464. Table 6 lists the goodness of fit measure for Kevlar 49/epoxy strands fewer than 90 percent pressure. Compared to the other models, the TP-GAPE distribution offers a superior match to the data set. Table 6 shows that when compared to the other fitted models, the TP-GAPE distribution has the lowest W, AIC, AICc, and BIC values.

CONCLUSION

Probability distributions have significance for modeling various sets of data in a number of fields, including engineering, business, reliability analysis, insurance, and biostatistics. The present study offers a new method to obtain a new probability model to adequately model various data sets including breast cancer data. The TP-GAPE (α, γ, λ) distribution, a unique probability distribution with three parameters, has been proposed.

Several attractive properties have been discussed for the new model including the cumulative distribution and probability density function, the survival function, hazard function, quantile function, median, order statistics, etc. The parameters were estimated using the greatest likelihood method. Plots of the PDF, CDF, and hazard rate function are made to show how the suggested distribution behaves. Furthermore, the performance of the TP-GAPE (α, γ, λ) distribution is compared to other distributions using various model selection criteria. The results show that TP-GAPE (α, γ, λ) gives results that are better than the other distributions.

REFERENCES

1. Z. Ahmad, E. Mahmoudi, G. G. Hamedani, and O. Kharazmi (2020). "New methods to define heavy-tailed distributions with applications to insurance data," *Journal of Taibah University forScience*, vol. 14, no. 1, pp. 359–382.
2. S. Nasiru, P. N. Mwita, and O. Ngesa,(2019) "Exponentiated generalized power series family of distributions," *Annals of DataScience*, vol. 6, no. 3, pp. 463–489.
3. S. Nasiru, P. N. Mwita, and O. Ngesa (2019). "Exponentiated generalized exponential Dagum distribution," *Journal of KingSaud University Science*, vol. 31, no. 3, pp. 362–371.
4. A. H. Muse, S. Mwalili, O. Ngesa, S. J. Almalki, and G. A. Abd- Elmougod (2021). "Bayesian and classical inference for the generalized log-logistic distribution with applications to survival data," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5820435, 24 pages.
5. Y. L. Tung, Z. Ahmad, O. Kharazmi, C. B. Ampadu, E. H. Hafez, and S. A. Mubarak (2021). "On a new modification of the Weibull model with classical and Bayesian analysis," *Complexity*, vol. 2021, Article ID 5574112, 19 pages.
6. O. S. Balogun, M. Z. Iqbal, M. Z. Arshad, A. Z. Afify, and P. E. Oguntunde (2021). "A new generalization of Lehmann type-II distribution: theory, simulation, and applications to survival and failure rate data," *Scientific African*, vol. 12, e00790.
7. Azzalini, A. (1985).A class of distributions which includes the normal ones. *Scand. J. Stat.* 12(2):171–178.
8. Mudholkar, G.S., Srivastava, D.K. (1993). ExponentiatedWeibull family for analyzing bathtub failureratedata.*IEEE Trans. Reliab*. 42(2):299–302.
9. Marshall, A.W., Olkin, I. (1997).A newmethod for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*84(3):641 652.
10. Eugene, N., et al. (2002).Beta-normal distribution and its applications.*Communications in Statistics- Theory and Methods*. 31(4):497–512.
11. Ijaz, M., Asim, S. M., Alamgir, Farooq, M., Khan, S. A., & Manzoor, S. (2020). A Gull Alpha Power Weibull distribution with applications to real and simulated data. *Plos one*, *15*(6), e0233080.
12. M. Amini, S. M. T. K. MirMostafaee, and J. Ahmadi (2014). "Loggamma- generated families of distributions," *Statistics*, vol. 48, no. 4, pp. 913–932.
13. A. Alzaatreh, C. Lee, and F. Famoye (2013)."A new method for generating families of continuous distributions," *Metron*, vol. 71, no. 1, pp. 63–79.
14. G. M. Cordeiro, M. Alizadeh, and P. R. DinizMarinho (2016). "The type I half-logistic family of distributions," *Journal of StatisticalComputation and Simulation*, vol. 86, no. 4, pp. 707– 728.
15. M. Aldeni, C. Lee, and F. Famoye (2017). "Families of distributions arising from the quantile of generalized lambda distribution," *Journal of Statistical Distributions and Applications*, vol. 4, no. 1, pp. 1–18.

16. A. Alzaghal, F. Famoye, and C. Lee (2013). “Exponentiated TX family of distributions with some applications,” *International Journalof Statistics and Probability*, vol. 2, no. 3, p. 31.
17. M. Ijaz, S. M. Asim, and Alamgir (2019). “Lomax exponential distribution with an application to real-life data,” *PLoS One*, vol. 14, no. 12, Article ID e0225827.
18. R. Al-Aqtash, F. Famoye, and C. Lee (2015). “On generating a new family of distributions using the logit function,” *Journal ofProbability and Statistical Science*, vol. 13, no. 1, pp. 135–152.
19. G. M. Cordeiro, M. Alizadeh, T. G. Ramires, and E. M. M. Ortega (2017). “*e generalized odd half-Cauchy family of distributions: properties and applications,” *Communications in Statistics - theory and Methods*, vol. 46, no. 11, pp. 5685– 5705.
20. M. A. Nasir, M. Aljarrah, F. Jamal, and M. H. Tahir (2017). “A new generalized Burr family of distributions based on quantile function,” *Journal of Statistics Applications and Probability*, vol. 6, no. 3, pp. 1–14.
21. M. E. Ghitany, F. A. Al-Awadhi, and L. A. Alkhalfan(2007). “Marshall-olkin extended lomax distribution and its application to censored data,” *Communications in Statistics-_eoryand Methods*, vol. 36, no. 10, pp. 1855–1866.
22. M. Ijaz, M. Asim, and A. Khalil (2019). “Flexible lomax distribution,” *Songklanakarin Journal of Science and Technology*, vol. 42, no. 5, pp. 1125–1134.
23. M. Ali, A. Khalil, M. Ijaz, and N. Saeed (2021). “Alpha-Power Exponentiated Inverse Rayleigh distribution and its applications to real and simulated data,” *PLoS One*, vol. 16, no. 1, Article ID e0245253.
24. A. EL-Baset, A. Ahmad, and M. G. M. Ghazal(2020). “Exponentiated additive Weibull distribution,” *Reliability Engineering andSystem Safety*, vol. 193, Article ID 106663
25. Andrews DF, Herzberg AM (2012) Data: a collection of problems from many fields for the studentand research worker. Springer, New York
26. Barlow R, Toland R, Freeman T (1984) A Bayesian analysis of stress-rupture life of Kevlar 49/epoxy spherical pressure vessels. In: Proceedings of conference on applications of statistics. Marcel Dekker, New York.
27. Ramos, M. A., Cordeiro, G. M., Marinho, P. D., Dias, C. B. and Hamadani, G. G. (2013). The zografos-balakrishman log-logistic distribution: properties and applications, Journal of Statistical Theory and Applications, 12(3): 225-244.