Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 154 (2021) 1-33 EISSN 2392-2192

Some Inferences on Dagum (4P) Distribution: Statistical Properties, Characterizations and Applications

M. Shakil[1,a], **B. M. G. Kibria**[2,b] and **M. Ahsanullah**[3,c]

[1]Department of Liberal Arts and Sciences, Faculty of Mathematics, Miami Dade College, Hialeah, FL, USA

[2]Department of Mathematics and Statistics, Faculty of Statistics, Florida International University, FIU, Miami, USA

[3]Department of Management Sciences, Rider University, Lawrenceville, NJ, USA

[a-c]E-mail address: mshakil@mdc.edu , kibriag@fiu.edu , ahsan@rider.edu

ABSTRACT

The Dagum (4P) distribution is one of the most popular statistical models in the fields of actuarial science, biological sciences, engineering, finance, hydrology, medical sciences, reliability, transportation, etc. The objective of this paper is to discuss the Dagum (4P) distribution, its various statistical properties, characterizations and applications to some real life data sets, and draw some inferences on it.

***Keywords*:** Characterizations, Dagum (4P) distribution, Goodness of fit tests, Reliability

1. INTRODUCTION

In order to deal with the random phenomena and data occurring in many applied problems in the fields of actuarial science, biological sciences, engineering, finance, hydrology, medical sciences, reliability, transportation, etc., probability distributions can be applied to make

(Received 10 January 2021; Accepted 05 February 2021; Date of Publication 07 February 2021)

predictions and informed decisions under uncertainty. Therefore, the statistical treatment of such data is an important aspect of their analysis and interpretation. Different probability models have been applied to characterize such data. Thus, a better selection of the best fitting probability distribution may be helpful in extrapolating the observed values to those which are more significant from the point of view of quality and reliability engineering standards. Motivated by the importance of such studies, in this paper, we have considered the fitting of the Dagum (4 P) distribution function to some real life-time data, namely, the tensile fatigue characteristics of the yarn data as reported by Quesenberry and Kent [1]. The goodness of fit test of the Dagum (4 P) distribution to these data was carried out by the Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared distribution tests, and compared with those of Burr (3P), Birnbaum-Saunders (3P) and Dagum (3P) distributions.

The organization of the paper is as follows: In Section 2, we give a description the Dagum (4P) distribution, including several new statistical properties, namely, the moments, Shannon entropy, reliability analysis and computations of percentage points. The characterizations of the Dagum (4P) distribution are provided in Section 3. The estimation of the parameters and the applications of the Dagum (4P) distribution to some real life-time data based on some goodness of fit tests are given in Section 4. We have provided some concluding remarks in Section 5.

2. DAGUM (4P) DISTRIBUTION

The Dagum (4P) distribution was proposed by Camilo Dagum in the 1970s in the study of the income distribution when he investigated the size distribution of personal income. It is a continuous probability distribution defined over positive real numbers. For details, see, for example, Dagum [2, 3], Johnson et al. [4], Kleiber and Kotz [5], Kleiber [6], Dey et al. [7] and Domma et al. [8], among others.

Dagum (4P) Distribution: A continuous non-negative random variable, X , is said to have a Dagum (4P) distribution if its probability density function (pdf) and cumulative distribution function (cdf) are respectively given by

$$f\left(x\right)=\frac{\alpha k\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}}{\beta\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}, \tag{1}$$

and

$$F\left(x\right)=\left(1+\left(\frac{x-\gamma}{\beta}\right)^{-\alpha}\right)^{-k}, \tag{2}$$

where $k\left(>0\right)$: shape parameter; $\alpha\left(>0\right)$: shape parameter; $\beta\left(>0\right)$: scale parameter; $-\infty<\gamma<+\infty$: location parameter; and domain: $\gamma\leq x<\infty$. Note that Dagum (4P) distribution is also known as the inverse Burr (4P) distribution. When $\gamma=0$, Dagum (4P) reduces to the

Dagum (3P) distribution. The possible shapes of the pdf (1) and cdf (2) of Dagum (4P) distribution are given for some selected values of the parameters in **Figures 1** & **2**, respectively.

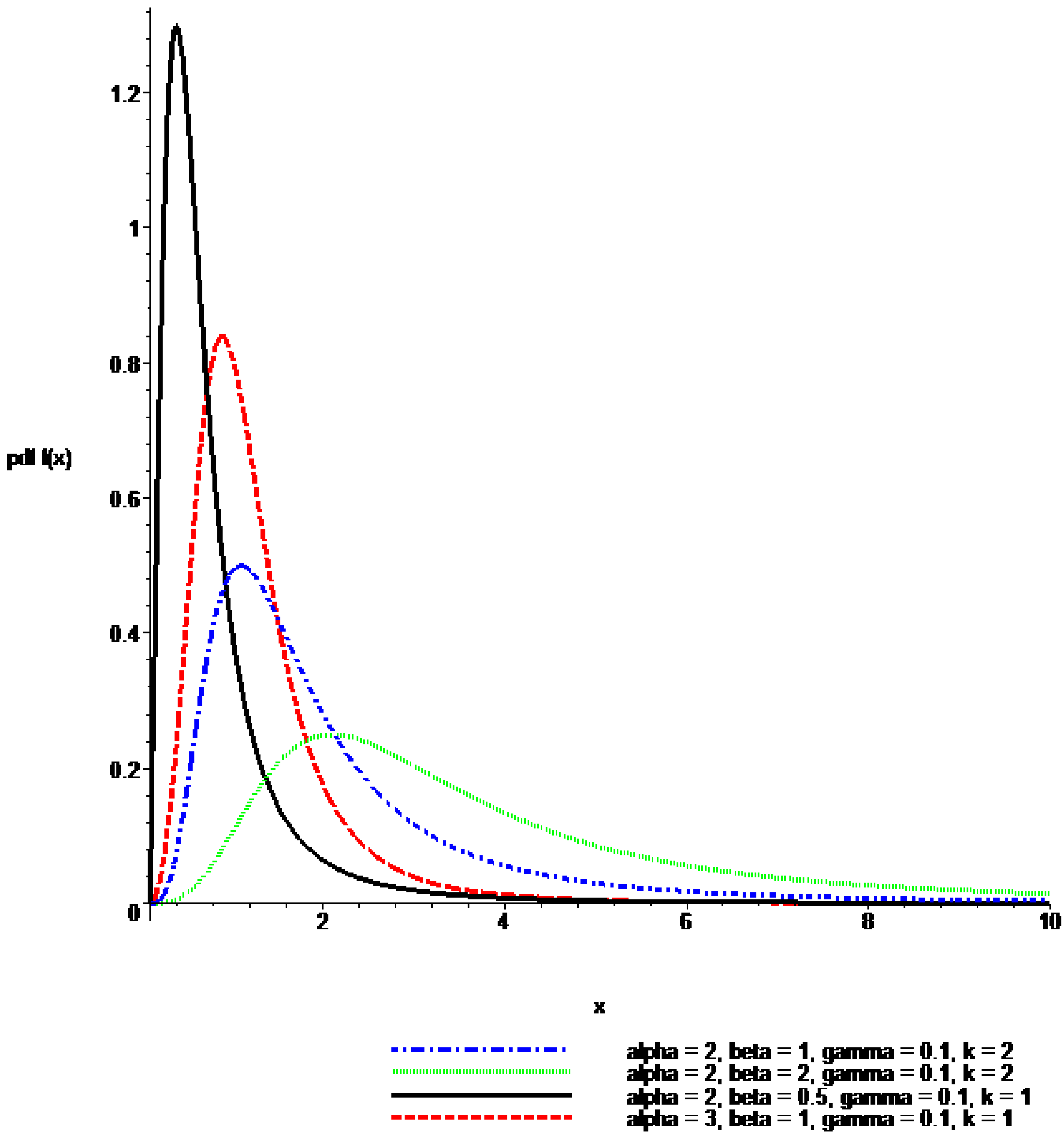


Figure 1. Plots of the Dagum (4P) pdf (1).

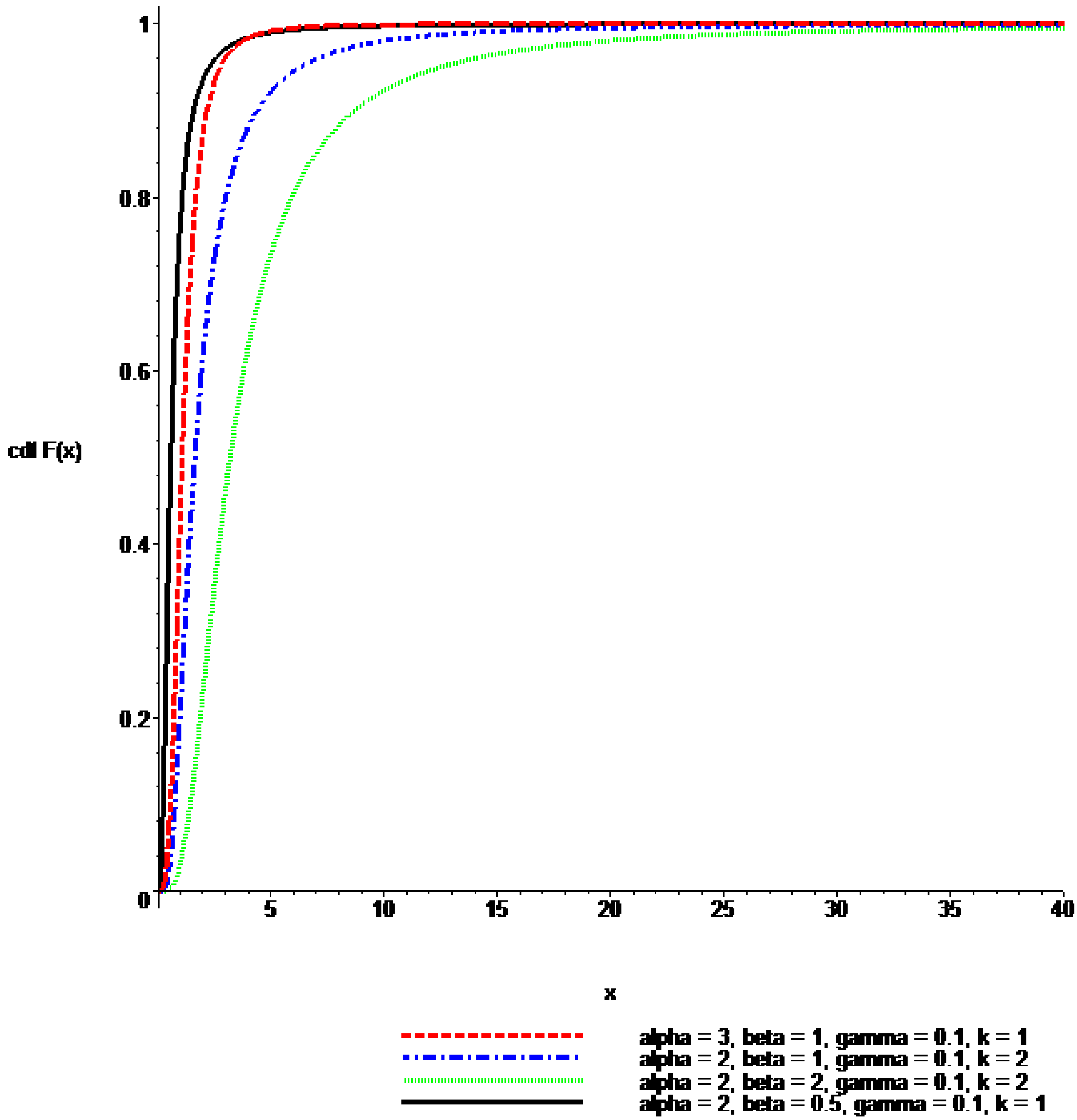


Figure 2. Plots of the Dagum (4P) cdf (2).

The effects of the parameters can easily be seen from these graphs. For example, it is clear from these plots that the Dagum (4P) distribution is positively right skewed with longer and heavier right tails for the selected values of the parameters.

2. 1. *jth* Moment of the Dagum (4P) Distribution

For a positive integer j, the jth moment, α_j, of the Dagum (4P) distribution is given by

$$\alpha_j = E\left(X^j\right) = \int_\gamma^\infty x^j \frac{\alpha k \left(\frac{x-\gamma}{\beta}\right)^{\alpha k - 1}}{\beta \left(1 + \left(\frac{x-\gamma}{\beta}\right)^\alpha\right)^{k+1}} dx. \tag{3}$$

Letting $\frac{x-\gamma}{\beta} = u$ in equation (3), we have

$$\alpha_j = E\left(X^j\right) = \left(\alpha k \beta^j\right) \int_0^\infty u^{\alpha k - 1} \left(u + \frac{\gamma}{\beta}\right)^j \left(1 + u^\alpha\right)^{-(k+1)} du. \tag{4}$$

Now, using the binomial expansion for $\left(u + \frac{\gamma}{\beta}\right)^j$ in equation (4) and simplifying, we obtain

$$\alpha_j = E\left(X^j\right) = \left(\alpha k \beta^j\right) \sum_{m=0}^{j} \binom{j}{m} \left(\frac{\gamma}{\beta}\right)^m \int_0^\infty u^{\alpha k + j - m - 1} \left(1 + u^\alpha\right)^{-(k+1)} du. \tag{5}$$

Thus, using the Eq. 3.251.11, page 295, of Gradshteyn and Ryzhik [9], the jth moment of the Dagum (4P) distribution is easily given by

$$\alpha_j = E\left(X^j\right) = \left(k \beta^j\right) \sum_{m=0}^{j} \binom{j}{m} \left(\frac{\gamma}{\beta}\right)^m B\left(k + \frac{j-m}{\alpha}, 1 - \frac{j-m}{\alpha}\right), \tag{6}$$

where $0 \le m \le j < \alpha$, $j > 0$ (a positive integer), $k > 0$, $\alpha > 0$, $\beta > 0$, $-\infty < \gamma < +\infty$, and $B(\ )$ denotes the complete beta function . Taking $j = 1$ in (6) and simplifying, the first moment (or the mean), α_1, of the Dagum (4P) distribution is easily given by

$$\alpha_1 = E\left(X\right) = \gamma + \left(k \beta\right) B\left(k + \frac{1}{\alpha}, 1 - \frac{1}{\alpha}\right), 0 < \frac{1}{\alpha} < 1, \tag{7}$$

where $k > 0$, $\alpha > 0$, $\beta > 0$, $-\infty < \gamma < +\infty$.

2. 2. *jth* (Central) Moment

The jth (central) moment of the Dagum (4P) distribution can easily be derived as follows:

$$\beta_j = E\left[\ X - E(\ X\ )\right]^j = \int_\gamma^\infty \left[\ x - E(\ X\ )\right]^j f_X(x)\, dx$$

$$= \sum_{m=0}^{j} (-1)^m \binom{j}{m} \left(E(X)\right)^m E\left(X^{j-m}\right), \tag{8}$$

where $E\left(X^{j-m}\right)$ and $\left(E(X)\right)^{m}$ can be obtained from the equations (6) and (7) respectively. From the equation (8), one can easily obtain the second, third, and higher central moments.

2. 3. Mean, Variance, Coefficients of Skewness and Kurtosis:

Mean: As stated above, **t**aking $j=1$ in the equation (6), the mean, $\alpha_1 = E(X)$, of the random variable X can easily be obtained.

Variance: Taking $j=2$ in equation (8), the variance (or the second central moment), β_2, is given by

$$\beta_2 = E\left[X - E(X)\right]^2 = \int_{\gamma}^{\infty}\left[x - E(X)\right]^2 f_X(x)\,dx = E\left[X^2\right] - \left(E[X]\right)^2. \tag{9}$$

Coefficients of Skewness and Kurtosis: By taking $j=3$ and $j=4$ in the equation (8), the third and fourth central moments are respectively given by

$$\beta_3 = E\left[X - E(X)\right]^3 = \sum_{m=0}^{3}(-1)^m\binom{3}{m}\left(E(X)\right)^m E\left(X^{3-m}\right), \tag{10}$$

and

$$\beta_4 = E\left[X - E(X)\right]^4 = \sum_{m=0}^{4}(-1)^m\binom{4}{m}\left(E(X)\right)^m E\left(X^{4-m}\right). \tag{11}$$

Thus, using equations (11) and (12), the measure of skewness, γ_1, and kurtosis, γ_2, are respectively given by

$$\gamma_1 = \frac{\sum_{m=0}^{3}(-1)^m\binom{3}{m}\left(E(X)\right)^m E\left(X^{3-m}\right)}{\left(E\left[X^2\right] - E[X]^2\right)^{\frac{3}{2}}} = \frac{\beta_3}{(\beta_2)^{3/2}}, \tag{12}$$

and

$$\gamma_2 = \frac{\sum_{m=0}^{4}(-1)^m\binom{4}{m}\left(E(X)\right)^m E\left(X^{4-m}\right)}{\left(E\left[X^2\right] - E[X]^2\right)^{2}} = \frac{\beta_4}{(\beta_2)^{2}}, \tag{13}$$

where $E\left(X^{j-m}\right)$ and $\left(E(X)\right)^{m}$ can be obtained from the equations (6) and (7) respectively.

2. 4. Empirical Analysis of Mean, Variance, Skewness and Kurtosis

In what follows, we will conduct some empirical analysis of the mean, variance, skewness and kurtosis of the Dagum (4P) distribution by considering some actual data. This is due to the fact that the difference between the theoretical and empirical probabilities is that the theoretical probability assumes that certain outcomes are equally likely, while empirical probability relies on the actual data to determine the likelihood of outcomes.

2. 4. 1. Example

For the empirical analysis of the mean, variance, skewness and kurtosis of the Dagum (4P) distribution based on some actual data, we consider the following example of a random sample of the tensile strength of 100 carbon fibers data as reported by Nichols and Padgett [10], which are provided in **Table 1**.

Table 1. The Tensile Strength of 100 Carbon Fibers Data (Sample Size n = 100).

3.70	2.74	2.73	2.50	3.60	3.11	3.27	2.87	1.47	3.11
4.42	2.41	3.19	3.22	1.69	3.28	3.09	1.87	3.15	4.90
3.75	2.43	2.95	2.97	3.39	2.96	2.53	2.67	2.93	3.22
3.39	2.81	4.20	3.33	2.55	3.31	3.31	2.85	2.56	3.56
3.15	2.35	2.55	2.59	2.38	2.81	2.77	2.17	2.83	1.92
1.41	3.68	2.97	1.36	0.98	2.76	4.91	3.68	1.84	1.59
3.19	1.57	0.81	5.56	1.73	1.59	2.00	1.22	1.12	1.71
2.17	1.17	5.08	2.48	1.18	3.51	2.17	1.69	1.25	4.38
1.84	0.39	3.68	2.48	0.85	1.61	2.79	4.70	2.03	1.80
1.57	1.08	2.03	1.61	2.12	1.89	2.88	2.82	2.05	3.65

The descriptive statistics of the above-mentioned carbon fibers data are computed in **Table 2** below.

Table 2. Descriptive Statistics of the Tensile Strength of 100 Carbon Fibers Data

Statistic	Value	Statistic	Value	Percentile	Value
Sample Size	100	Kurtosis	3.26597	Min	0.39
Range	5.17	Mode	2.17, 3.68	5%	1.082
Mean	2.6214	Midrange	2.975	10%	1.261
Variance	1.028			25% (Q1)	1.84

Std. Deviation	1.0139			50% (Median)	2.7
Coef. of Variation	0.38677			75% (Q3)	3.22
Std. Error	0.10139			90%	3.698
Skewness	0.37378			95%	4.686
Excess Kurtosis	0.17287			Max	5.56

2. 4. 2. Histogram, pdf, Empirical CDF and P-P plots of Dagum (4P) Distribution fitted to Carbon Fibers Data

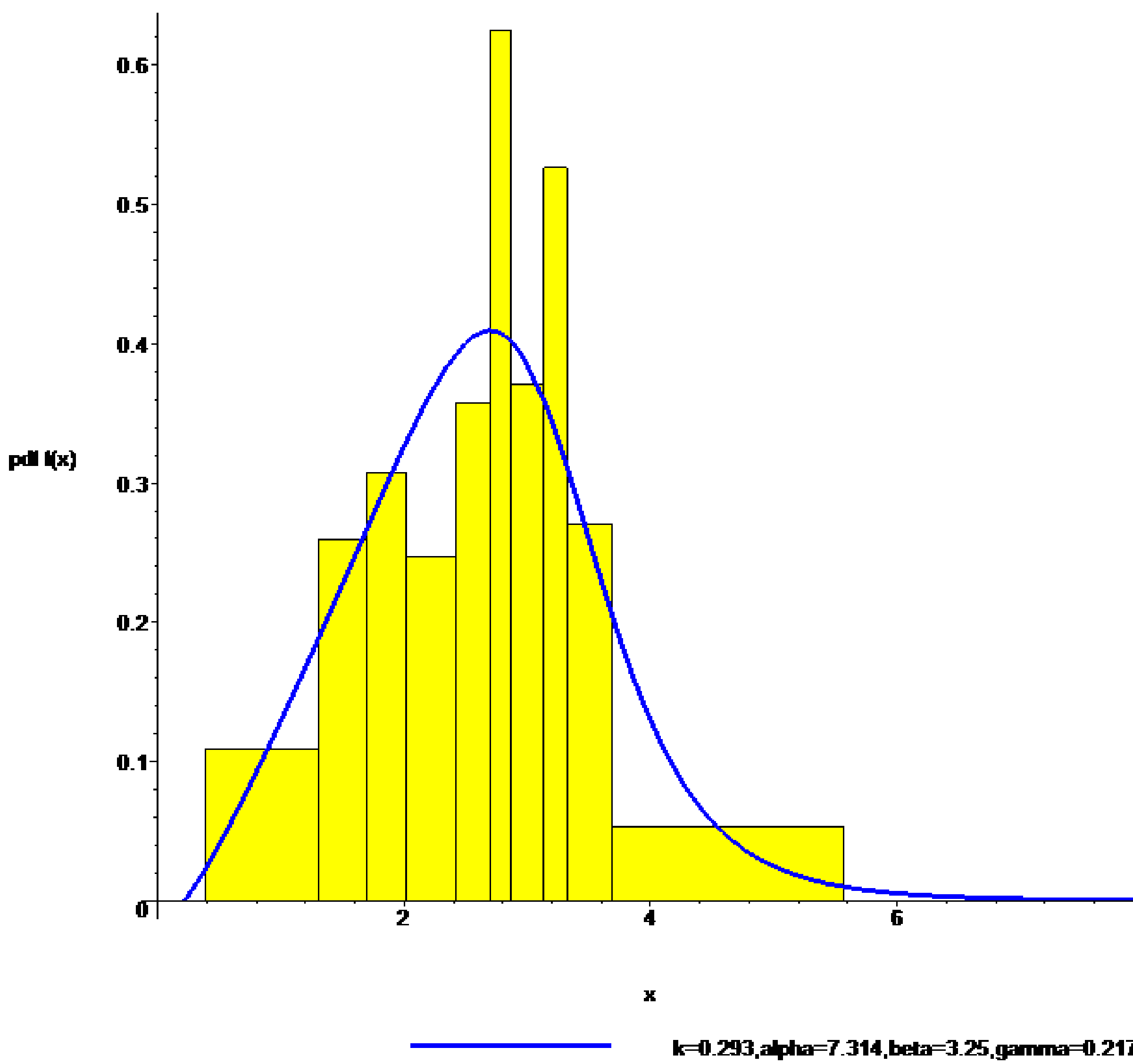
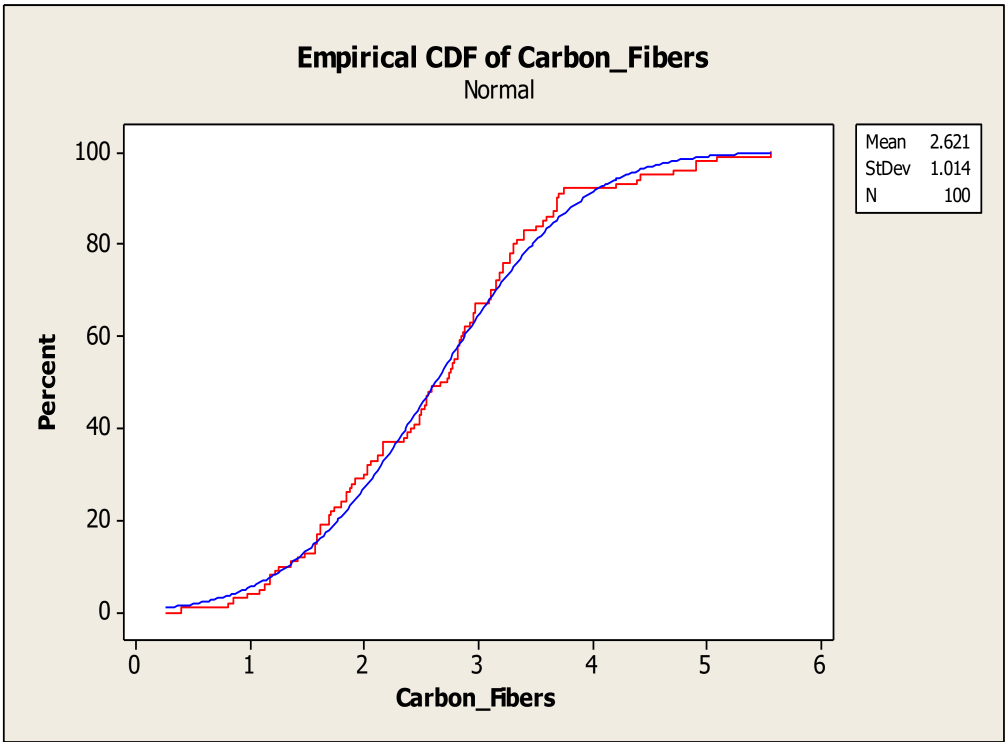
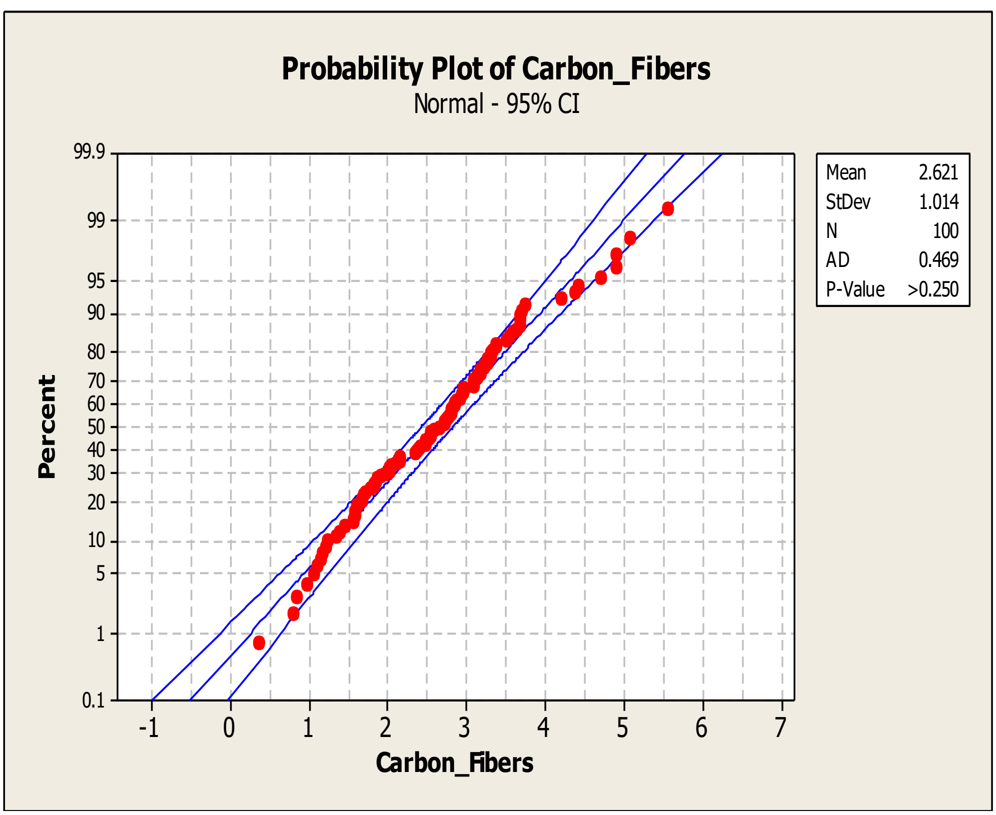


Figure 3 (a). Fitting of the pdfs of the Dagum (4 P) Distributions to the Histogram of the Tensile Strength of 100 Carbon Fibers Data.



(A)



(B)

Figure 4. Empirical CDF Plot **(A)** and P-P Plot **(B)** of the Tensile Strength of 100 Carbon Fibers Data

Since fitting of a probability distribution to the tensile strength of 100 carbon fibers data may be helpful in predicting the probability or forecasting the frequency of occurrence of the carbon fibers data, this suggests that ‘ y ’, the carbon fibers data, could possibly be modeled by some skewed distributions. As such, we have tested the fitting of the Dagum (4P) distribution to the tensile strength of 100 carbon fibers data. The probability density functions (pdf’s) of the Dagum (4P) distribution have been superimposed on the histogram the tensile strength of 100 carbon fibers data, as provided in **Figure 3**.

Furthermore, it should be noted that, in order to determine how well a specific distribution fits to the observed data, we usually draw the probability-probability, that is, the P-P plot, which should be approximately linear if the specified theoretical distribution is the correct model. It is a graph of the empirical cumulative distribution function (CDF) values which are plotted against the fitted, that is, the theoretical CDF values. Thus, for the Dagum (4P) distribution, we have also plotted the empirical CDF and P-P plots for the tensile strength of 100 carbon fibers data in **Figure 4**.

Thus, based on our plotted graphs in **Figures 3** and **4** respectively, it is observed that the Dagum (4P) distribution models for the tensile strength of 100 carbon fibers data reasonably well.

2. 4. 3. Confidence Interval Estimates of the Mean and Variance

From **Figures 3** (fitting of the pdf) and 4 (empirical CDF and P-P plots), it is obvious that the shape of the tensile strength of 100 carbon fibers data is positively skewed. This is also confirmed from the skewness (0.37378) and kurtosis (3.26597) of the data as computed in Table 2 above as well as the data points on the P-P plot do not adhere well to a straight line which also suggests that the tensile strength of 100 carbon fibers data are not normally distributed. Since we have a large sample with n = 100 and the data are skewed, therefore, the 95 % confidence interval (CI) estimates of mean, standard deviation and variance of the above-mentioned random sample of the tensile strength of 100 carbon fibers data are respectively computed as follows:

(i) 95% Confidence Interval for the Mean:

Margin of error, E = 0.2011796

95% Confident the population mean is within the range:

2.42022 < mean <2.82258

(ii) 95% Confidence Interval for the St. Dev.:

95% Confident the population S.D. is within the range:

0.8902116 < SD < 1.177822

(iii) 95% Confidence Interval for the variance:

95% Confident the population variance is within the range:

0.7924768 < VAR < 1.387265

2. 4. 4. Empirical Analysis of Skewness

For this, we have first estimated the parameters of the Dagum (4P) distribution based on the tensile strength of 100 carbon fibers data, which are provided in the **Table 3**.

Table 3. Fitting Results (Estimation of the Parameters)

#	Distribution	Parameter Estimates
1	Dagum (4P)	$n = 100$, $k = 0.293$, $\alpha = 7.314$, $\beta = 3.25$, $\gamma = 0.217$

The mean, α_1, is given by the equation (7), and the variance, β_2, and the third and fourth central moments, β_3 and β_4, are obtained from the equation (8) by taking $j = 2$, $j = 3$ and $j = 4$ respectively. Using a Maple 11 program, numerical values of skewness, γ_1, and kurtosis, γ_2, for the above-mentioned estimated values of the parameters of the Dagum (4P) distribution based on the tensile strength of 100 carbon fibers data, are provided in the **Tables 4** below:

Table 4. Numerical Values of skewness, γ_1, and Kurtosis, γ_2

n	α_1	β_2	β_3	β_4	γ_1	γ_2
100	**2.621**	**1.030**	**0.581**	**5.115**	**0.55552**	**4.82202**

It is evident from these computations that the skewness, γ_1, is positive which implies that distribution of the random variable X is positively skewed. Moreover, it is observed from these tables that, for all estimated values of the parameters, the kurtosis, $\gamma_2 > 3$, implying that the distribution is heavier tailed.

Following Cramer [11] and Joanes and Gill [12], we compute the sample skewness of the tensile strength of 100 carbon fibers data by the following formula

$$G_1 = \left(\frac{\sqrt{n(n-1)}}{n-2} \right) \gamma_1,$$

from which we obtain

$$G_1 = \left(\frac{\sqrt{100 \times (100-1)}}{(100-2)} \right) \times (0.55552) \approx 0.56402.$$

The standard error of skewness (SES) and the test statistic, which measures how many standard errors separate the sample skewness from zero, are computed as follows:

$$SES = \sqrt{\frac{6\,n(n-1)}{(n-2)(n+1)(n+3)}} = \sqrt{\frac{6\times100\times(100-1)}{(100-2)\times(100+1)\times(100+3)}} \approx 0.241380,$$

from which we get the test statistic as

$$Z_{\gamma_1} = \frac{G_1}{SES} = \frac{2.331456}{0.241380} \approx 2.3366.$$

The critical value of Z_{γ_1} is approximately 2. (This is a two-tailed test of skewness $\neq 0$ at roughly the 0.05 significance level.)

Since $Z_{\gamma_1} > 2$, it is inferred that the population is very likely skewed positively.

Confidence Interval for Skewness: 95 % confidence interval of the skewness is given by

$$G_1 \pm 2\,SES,$$

from which we obtain the 95 % confidence interval of the skewness as $(0.08126, 1.04677)$.

2. 4. 5. Empirical Analysis of Kurtosis

Following Cramer [11] and Joanes and Gill [12], we first compute the excess kurtosis by the following formula

$$g_2 = \gamma_2 - 3,$$

from which we obtain

$$g_2 = 4.82202 - 3 = 1.82202.$$

The sample excess kurtosis of the tensile strength of 100 carbon fibers data is computed by the following formula

$$G_2 = \left(\frac{n-1}{(n-2)(n-3)}\right)\left[(n+1)g_2 + 6\right],$$

from which we obtain

$$G_2 = \left(\frac{100-1}{(100-2)\times(100-3)}\right)\left[(100+1)\times1.82202 + 6\right] \approx 1.9790.$$

The standard error of kurtosis (SEK) and the test statistic, which measures how many standard errors the sample excess kurtosis is from zero, are computed as follows:

$$SEK = 2 \times (SES) \times \sqrt{\frac{(n^2 - 1)}{(n-3)(n+5)}}$$

$$= 2 \times (0.241380) \times \sqrt{\frac{(100^2 - 1)}{(100-3)(100+5)}} \approx 0.47833$$

where SES denotes the standard error of skewness, from which we get the test statistic as

$$Z_{\gamma_2} = \frac{G_2}{SEK} = \frac{1.9790}{0.47833} \approx 4.1373$$

The critical value of Z_{γ_2} is approximately 2. (This is a two-tailed test of excess kurtosis ≠ 0 at approximately the 0.05 significance level.)

If Z_{γ_2} > +2, it is inferred that the population very likely has positive excess kurtosis.

Confidence Interval for Kurtosis: 95 % confidence interval of the kurtosis is given by

$$G_2 \pm 2\,SEK,$$

from which we obtain the 95 % confidence interval of the kurtosis as $(1.02234, 2.93566)$.

2. 5. Moment Generating Function, Characteristic Function and *rth* Cumulant

It is easy to see that, for the Dagum (4P) distribution, the moment generating and characteristic functions of the random variable X are respectively given by

$$M_X(t) = E\left(e^{tX}\right) = \sum_{j=0}^{\infty} \frac{(t)^j}{j!} E\left(X^j\right), \qquad (14)$$

and

$$\phi_X(t) = M_X(it) = E\left(e^{itX}\right) = \sum_{j=0}^{\infty} \frac{(it)^j}{j!} E\left(X^j\right), \qquad (15)$$

where $i = \sqrt{-1}$ is the imaginary number, $i^2 = -1$, and $E\left(X^j\right)$ denotes the jth moment about the origin of the random variable X which can easily obtained from the equation (6). By taking the natural log of the characteristic function ($\phi_X(t)$ in (15), the rth cumulant, κ_r, of the random variable X is then given by

$$\ln\left(\phi_X(t)\right)=\sum_{r=1}^{\infty}\kappa_r\frac{(it)^r}{r!},$$

see Eq. 26.1.12, Page. 928, of Abramowitz and Stegun [13] or Stuart and Ord [14], from which, taking the Maclaurin series of the left-hand side of the equation and equating the coefficients of various terms on both sides, gives the following expression for the required *rth* cumulant κ_r:

$$\kappa_r=\frac{1}{i^r}\left[\frac{d^r\left(\ln\left(\phi_X(t)\right)\right)}{dt^r}\right]_{t=0}, \quad r=1,2,\ldots, \tag{16}$$

from which, by successive differentiation, it can be easily seen that

$$\kappa_1=E(X)=\alpha_1,\ \kappa_2=Var(X)=\beta_2,\ \kappa_3=E\left[X-E(X)\right]^3=\beta_3, \text{ etc.,}$$

which can easily obtained by using the equations (7), (9) and (10), respectively.

2. 6. *jth* Incomplete Moment of the Dagum (4P) Distribution

For a positive integer j, the *jth* incomplete moment of the Dagum (4P) distribution is given by

$$I_j(x)=\int_{\gamma}^{x}u^j\,f(u)\,du=\int_{\gamma}^{x}u^j\frac{\alpha k\left(\frac{u-\gamma}{\beta}\right)^{\alpha k-1}}{\beta\left(1+\left(\frac{u-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}du. \tag{17}$$

Letting $\left(\frac{u-\gamma}{\beta}\right)^{\alpha}=t$ in equation (17) and using the binomial expansion for $\left(t^{\frac{1}{\alpha}}+\frac{\gamma}{\beta}\right)^{j}$, we have

$$\begin{aligned}I_j(x)&=\int_0^{\left(\frac{x-\gamma}{\beta}\right)^{\alpha}}k\beta^j\left(t^{\frac{1}{\alpha}}+\frac{\gamma}{\beta}\right)^{j}\frac{(t)^{k-1}}{(1+t)^{k+1}}dt\\&=\left(k\beta^j\right)\sum_{m=0}^{j}\binom{j}{m}\left(\frac{\gamma}{\beta}\right)^{m}\int_0^{\left(\frac{x-\gamma}{\beta}\right)^{\alpha}}\frac{(t)^{\frac{j-m}{\alpha}+k-1}}{(1+t)^{k+1}}dt\end{aligned}$$

from which, using the Eq. 3.194.1, page 284, of Gradshteyn and Ryzhik [9], the *jth* incomplete moment of the Dagum (4P) distribution is easily given by

$$I_x = \left(\alpha k \beta^j\right) \sum_{m=0}^{j} \left[\binom{j}{m} \left(\frac{\gamma}{\beta}\right)^m \frac{1}{j-m+\alpha k} \left(\frac{x-\gamma}{\beta}\right)^{j-m+\alpha k} \times {}_2F_1\left(k+1, \frac{j-m+\alpha k}{\alpha}; 1+\frac{j-m+\alpha k}{\alpha}; -\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)\right]$$

$$= P_k(x), \text{ say,} \tag{18}$$

where $0 \le m \le j < \alpha$, $j > 0$ (a positive integer), $\frac{j-m+\alpha k}{\alpha} > 0$, $k > 0$, $\alpha > 0$, $\beta > 0$, $-\infty < \gamma < +\infty$, $\gamma \le x < \infty$ and ${}_2F_1(\ )$ denotes the hypergeometric function; for details, see Gradshteyn and Ryzhik [9], Eqs. 9.10-9.11, Pages 1039 - 1040. Taking $j = 1$ in (18) and simplifying, the first incomplete moment of the Dagum (4P) distribution is easily given by

$$P_1(x) = \int_\gamma^x u\, f(u)\, du$$

$$= \left(\alpha \beta k\right) \sum_{m=0}^{1} \left[\binom{1}{m} \left(\frac{\gamma}{\beta}\right)^m \frac{1}{1-m+\alpha k} \left(\frac{x-\gamma}{\beta}\right)^{1-m+\alpha k} \times {}_2F_1\left(k+1, \frac{1-m+\alpha k}{\alpha}; 1+\frac{1-m+\alpha k}{\alpha}; -\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)\right] \tag{19}$$

2. 7. Shannon Entropy

According to Shannon [15], the entropy measure of a continuous real random variable X is defined as

$$H_X\left[f_X(X)\right] = E[-\ln(f_X(X)] = -\int_{-\infty}^{\infty} f_X(x) \ln\left[f_X(x)\right] dx .$$

Thus, the Shannon entropy of the Dagum (4P) distribution is given by

$$H_X\left[f_X(X)\right] = -\int_\gamma^\infty \left[\frac{\alpha k \left(\frac{x-\gamma}{\beta}\right)^{\alpha k - 1}}{\beta \left(1+\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)^{k+1}} \right] \left[\ln \frac{\alpha k \left(\frac{x-\gamma}{\beta}\right)^{\alpha k - 1}}{\beta \left(1+\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)^{k+1}} \right] dx$$

$$= (k+1) E\left(\ln\left(1+\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)\right) - (\alpha k - 1) E\left(\ln\left(\frac{x-\gamma}{\beta}\right)\right) - \ln\left(\frac{\alpha k}{\beta}\right). \tag{20}$$

Now, in equation (20), taking $\left(\frac{x-\gamma}{\beta}\right)^{\alpha}=u$, and using the Eq. 4.293.14, page 558, of Gradshteyn and Ryzhik [9], for $E\left(\ln\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)\right)$, and, also, at the same time, taking $1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}=t$, and using the Eq. 4.293.9, page 558, of Gradshteyn and Ryzhik [9], for $E\left(\ln\left(\frac{x-\gamma}{\beta}\right)\right)$, we easily obtain

$$E\left(\ln\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)\right)=\psi(k+1)-\psi(1), \tag{21}$$

and

$$E\left(\ln\left(\frac{x-\gamma}{\beta}\right)\right)=\left(\frac{k}{\alpha}\right)\sum_{m=0}^{\infty}\frac{(-1)^{m+k-1}(k-1)(k-2)\ldots(k-m)}{m!} \times\left[\left(\frac{1}{m-k}\right)\left\{\pi\, ctg((m-k)\pi)+\psi(m-k+1)-\psi(1)\right\}\right], \tag{22}$$

where $m-k<0$, $k>0$, $\alpha>0$, $\psi(\ )$ denotes the digamma (or psi) function, and $-\psi(1)\approx 0.577216$ is known as the Euler's constant; for details, see Gradshteyn and Ryzhik [9]. Using (21) and (22) respectively in equation (20), we obtain the required Shannon entropy of the Dagum (4P) distribution.

2. 8. Reliability Analysis

The hazard function (hf) of the Dagum (4P) distribution is given by

$$h(x)=\frac{f_X(x)}{1-F_X(x)}=\frac{\alpha k\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}}{\beta\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}\left(1-\left(1+\left(\frac{x-\gamma}{\beta}\right)^{-\alpha}\right)^{-k}\right)}, \tag{23}$$

where $k(>0)$: shape parameter; $\alpha(>0)$: shape parameter; $\beta(>0)$: scale parameter; $-\infty<\gamma<+\infty$: location parameter; and domain: $\gamma\le x<\infty$. The possible shapes of the hf (23) of Dagum (4P) distribution are given for some selected values of the parameters in **Figure 5** below.

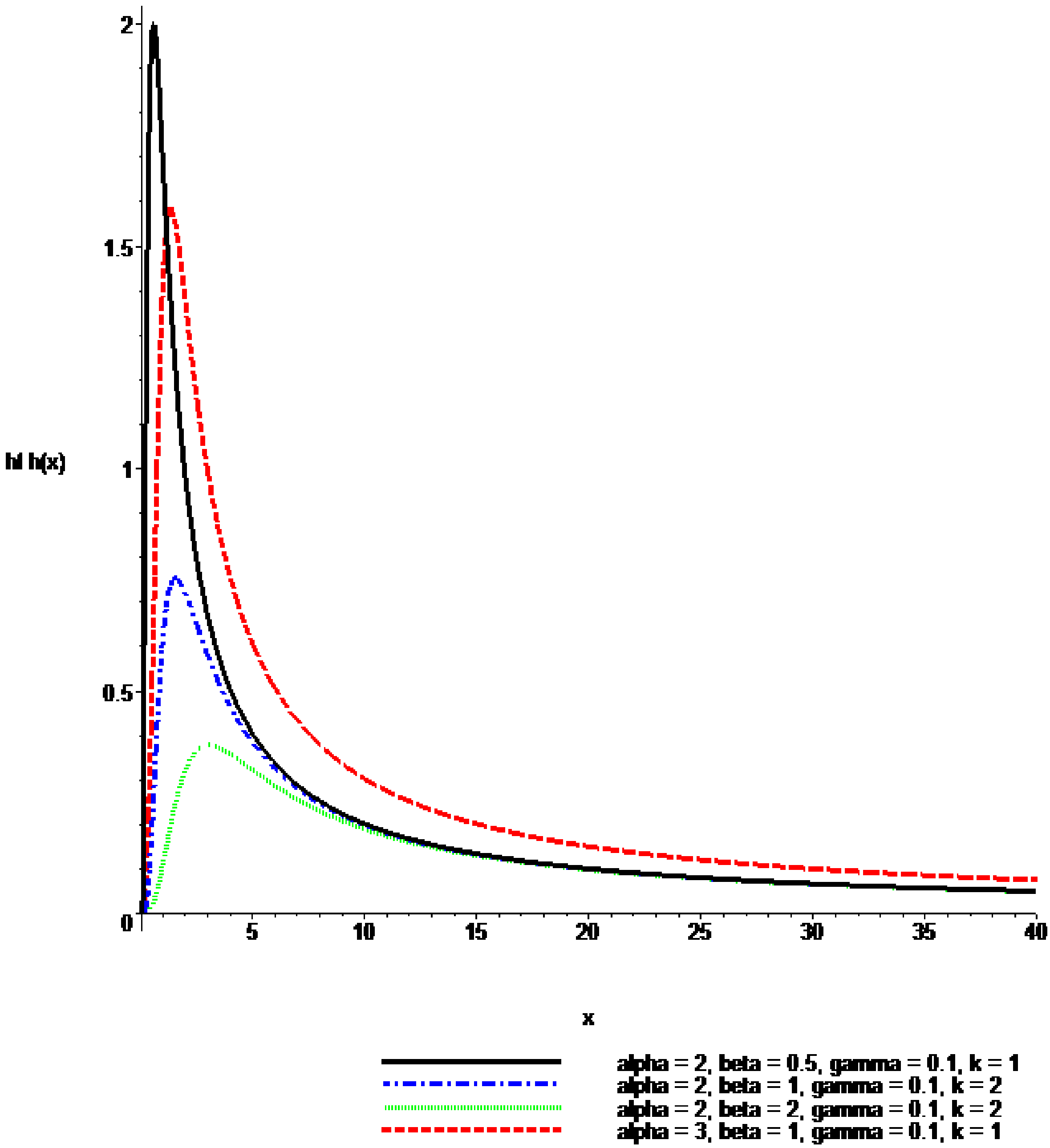


Figure 5. Plots of the Dagum (4P) hazard function (hf) (23)

The effects of the parameters are obvious from these figures. The increasing, then decreasing and upside-down bathtub shape behaviors of the failure rate function, $h(x)$, are also evident from the **Figure 3.** Also, differentiating equation (23) with respect to x, we have

$$h^{/}(x)=\frac{f^{/}(x)}{f(x)}h(x)+\left[h(x)\right]^{2}, \tag{24}$$

for $x > 0$, where $f(x)$ and $h(x)$ are given by equations (1) and (23) respectively, and $f^{/}(x)$ is obtained by differentiating equation (1) with respect to x, that is,

$$f^{/}(x)=\frac{\alpha k\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}\left[\alpha k-(\alpha+1)\left(\frac{x-\gamma}{\beta}\right)^{\alpha}-1\right]}{\beta(x-\gamma)\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+2}}. \tag{25}$$

In order to discuss the behavior of the failure rate function, $h(x)$, let $h^{/}(x)=0$. We observe that the nonlinear equation $h^{/}(x)=0$ does not have a closed form solution, but could be solved numerically by using some computer software. It is obvious from the equation (24) that $h^{/}(x)$ is positive irrespective of the values of the parameters. This shows that the Dagum (4P) distribution has the increasing failure rate (IFR) property.

2. 9. Percentile Points

The percentage points x_p of the Dagum (4P) distribution are computed by numerically solving the equation $F(x_p)=\int_{\gamma}^{x_p} f_X(u)\,du = p$ (say), for any $0 < p < 1$, for different sets of values of the parameters as provided in **Table 5**.

Table 5. Percentile Points of the Dagum (4P) Distribution.

Percentiles p / Parameters		0.75	0.80	0.85	0.90	0.95	0.99
$k=1, \gamma=0,$ $\alpha=2, \beta=0.5$	x_p	**0.8660**	**1.0000**	**1.1902**	**1.5000**	**2.1794**	**4.9749**
$k=2, \gamma=0.1,$ $\alpha=2, \beta=1$	x_p	**2.6425**	**3.0107**	**3.5370**	**4.3996**	**6.3043**	**14.1890**
$k=2, \gamma=0.1,$ $\alpha=2, \beta=2$	x_p	**5.1849**	**5.9214**	**6.9740**	**8.6993**	**12.5086**	**28.2779**
$k=1, \gamma=0.1,$ $\alpha=2, \beta=0.5$	x_p	**0.9660**	**1.1000**	**1.2902**	**1.6000**	**2.2794**	**5.0749**
$k=1, \gamma=0.1,$ $\alpha=3, \beta=1$	x_p	**1.5422**	**1.6874**	**1.8828**	**2.1801**	**2.7684**	**4.7261**

$k=3, \gamma=0.5,$ $\alpha=3, \beta=2$	x_p	**4.7997**	**5.1967**	**5.7380**	**6.5715**	**8.2412**	**13.8588**

3. CHARACTERIZATIONS

Since a characterization of a particular probability distribution states that it is the only distribution that satisfies some specified conditions, therefore in what follows, we will provide characterization results of the Dagum (4P) distribution by truncated moments. For various methods of characterizations of probability distributions, we refer to Ahsanullah (2017) and references therein.

3. 1. Characterization by Truncated Moment

We provide two new characterization results of the Dagum (4P) distribution by truncated moments. For this, we will also need the following assumption and lemmas.

Assumption 3.1. Suppose the random variable X is absolutely continuous with the cumulative distribution function $F(x)$ and the probability density function $f(x)$. We assume that $\omega = \inf\{x \mid F(x) > 0\}$, and $\delta = \sup\{x \mid F(x) < 1\}$. We also assume that $f(x)$ is a differentiable for all x, and $E(X)$ exists.

Lemma 3.1. If the random variable X satisfies the Assumption 3.1 with $\omega = \gamma$ and $\delta = \infty$, where $-\infty < \gamma < +\infty$, and if $E(X \mid X \le x) = g(x)\tau(x)$, where $\tau(x) = \frac{f(x)}{F(x)}$ and $g(x)$ is a continuous differentiable function of x with the condition that $\int_\gamma^x \frac{u - g'(u)}{g(u)} du$ is finite for $\gamma \le x$, $-\infty < \gamma < +\infty$, then $f(x) = c\, e^{\int_\gamma^x \frac{u - g'(u)}{g(u)} du}$, where c is a constant determined by the condition $\int_\gamma^\infty f(x)dx = 1$.

Proof. For proof, see Shakil et al. [16].

Lemma 3.2. If the random variable X satisfies the Assumption 3.1 with $\omega = \gamma$ and $\delta = \infty$, and if $E(X \mid X \ge x) = \tilde{g}(x) r(x)$, where $r(x) = \frac{f(x)}{1 - F(x)}$ and $\tilde{g}(x)$ is a continuous differentiable function of x with the condition that

$$\int_x^\infty \frac{u + \left[\tilde{g}(u)\right]'}{\tilde{g}(u)} du$$

is finite for for $\gamma \le x$, $-\infty < \gamma < +\infty$, then

$$f(x) = c\, e^{-\int_x^{\infty} \frac{u + \left[\tilde{g}(u)\right]'}{\tilde{g}(u)} du},$$

where c is a constant determined by the condition $\int_\gamma^\infty f(x)dx = 1$.

Proof. For proof, see Shakil et al. [16].

Theorem 3.1. If the random variable X satisfies the Assumption 3.1 with $\omega = \gamma$ and $\delta = \infty$, then $E(X|X \le x) = g(x)\, \frac{f(x)}{F(x)}$, where

$$g(x) = \frac{\left[\beta\left(1 + \left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}\right] P_1(x)}{\alpha k \left(\frac{x-\gamma}{\beta}\right)^{\alpha k - 1}}, \tag{26}$$

where $P_1(x)$ is given by (19), if and only if X has the pdf

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

Proof. Suppose that $E(X|X \le x) = g(x)\, \frac{f(x)}{F(x)}$. Then, since $E(X|X \le x) = \frac{\int_\gamma^x u\, f(u)\, du}{F(x)}$, we have $g(x) = \frac{\int_\gamma^x u\, f(u)\, du}{f(x)}$. Now, if the random variable X satisfies the Assumption 3.1 and has the distribution with the pdf (1), then we have

$$g(x)=\frac{\int_{\gamma}^{x} u\, f(u)\,du}{f(x)}=\frac{P_1(x)}{f(x)}=\frac{\left[\beta\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}\right]P_1(x)}{\alpha k\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}},$$

where $P_1(x)$ is given by (19). Consequently, the proof of "if" part of the Theorem 3.1 follows from Lemma 3.1.

Conversely, suppose that

$$g(x)=\frac{\left[\beta\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}\right]P_1(x)}{\alpha k\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}},$$

where $P_1(x)$ is given by (19). Now, from Lemma 3.1, we have

$$g(x)=\frac{\int_{\gamma}^{x} u\, f(u)\,du}{f(x)},$$

or

$$\int_{\gamma}^{x} u\, f(u)\,du = f(x)g(x).$$

Differentiating the above equation with respect to respect to x, we obtain

$$x\, f(x)= f^{/}(x)g(x)+ f(x)g^{/}(x),$$

from which, using the definition of the pdf (1) and $f^{/}(x)$ being given by (25), we easily obtain

$$g^{/}(x)=x-g(x)\frac{\left[\alpha k-(\alpha+1)\left(\frac{x-\gamma}{\beta}\right)^{\alpha}-1\right]}{(x-\gamma)\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)},$$

or,

$$\frac{x-g^{/}(x)}{g(x)}=\frac{\left[\alpha k-(\alpha+1)\left(\frac{x-\gamma}{\beta}\right)^{\alpha}-1\right]}{(x-\gamma)\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)}. \tag{27}$$

Since, by Lemma 3.1, we have

$$\frac{x-g^{/}(x)}{g(x)}=\frac{f^{/}(x)}{f(x)}\text{, (see Shakil, et al. [16]),} \tag{28}$$

therefore, from (27) and (28), it follows that

$$\frac{f^{/}(x)}{f(x)}=\frac{\left[\alpha k-(\alpha+1)\left(\frac{x-\gamma}{\beta}\right)^{\alpha}-1\right]}{(x-\gamma)\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)}. \tag{29}$$

Now, integrating Eq. (29) with respect to x and simplifying, we easily have

$$\ln(f(x))=\ln\left(c\left(\frac{\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}\right)\right),$$

or

$$f(x)=c\left(\frac{\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}\right), \tag{30}$$

where c is the normalizing constant to be determined. Thus, on integrating the above equation (30) with respect to x from $x=\gamma$ to $x=\infty$, and using the condition $\int_{\gamma}^{\infty} f(x)\,dx=1$, we obtain

$$\frac{1}{c}=\int_{\gamma}^{\infty}\left(\frac{\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}\right)dx\,. \tag{31}$$

Now, letting $\frac{x-\gamma}{\beta}=u$ in equation (31), and noting that $\int_{0}^{\infty}t^{\mu-1}\left(1+\theta t^{p}\right)^{-\nu}dt=\frac{1}{p}\theta^{-\frac{\mu}{p}}B\left(\frac{\mu}{p},\nu-\frac{\mu}{p}\right)$, where $B(\ )$ denotes the complete beta function, see Gradshteyn and Ryzhik [9], Eq. 3.251.11, Page 295, we easily obtain $c=\frac{\alpha k}{\beta}$. This completes the proof of Theorem 3.1.

Theorem 3.2. If the random variable X satisfies the Assumption 3.1 with $\omega=\gamma$ and $\delta=\infty$, then $E\left(X|X\geq x\right)=\tilde{g}(x)\frac{f(x)}{1-F(x)}$, where $\tilde{g}(x)=\frac{\left(E(X)-g(x)f(x)\right)}{f(x)}$, $g(x)$ being given by Eq. (26) and $E(X)$ being given by Eq. (7), if and only if X has the pdf

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

Proof. Suppose that $E\left(X|X\geq x\right)=\tilde{g}(x)\frac{f(x)}{1-F(x)}$. Then, since $E\left(X|X\geq x\right)=\frac{\int_{x}^{\infty}u\,f(u)du}{1-F(x)}$, we have $\tilde{g}(x)=\frac{\int_{x}^{\infty}u\,f(u)du}{f(x)}$. Now, if the random variable X satisfies the Assumptions 3.1 and has the distribution with the pdf (1), then, using the Theorem 3.1, we have

$$\tilde{g}(x)=\frac{\int_{x}^{\infty}u\,f(u)\,du}{f(x)}=\frac{\int_{\gamma}^{\infty}u\,f(u)\,du-\int_{\gamma}^{x}u\,f(u)\,du}{f(x)}$$

$$=\frac{\left(E(X)-g(x)\,f(x)\right)}{f(x)},$$

where $f(x)$ denotes the pdf of Dagum (4P) distribution given by Eq. (1), $g(x)$ being given by Eq. (26) and $E(X)$ being given by Eq. (7). Consequently, the proof of "if" part of the Theorem 3.2 follows from Lemma 3.2.

Conversely, suppose that $\tilde{g}(x)=\frac{\left(E(X)-g(x)\,f(x)\right)}{f(x)}$. Now, from Lemma 3.2, we have

$$\tilde{g}(x)=\frac{\int_x^\infty u\, f(u)du}{f(x)},$$

or

$$\int_x^\infty u\, f(u)\,du = f(x).\tilde{g}(x).$$

Differentiating the above equation with respect to respect to x, we obtain

$$-x\, f(x) = f'(x).\tilde{g}(x) + f(x).\left(\tilde{g}(x)\right)'.$$

Thus, proceeding in the same way as in Theorem 3.1, we easily obtain

$$f(x)=c\left(\frac{\left(\frac{x-\gamma}{\beta}\right)^{\alpha k-1}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}\right),$$

from which we easily obtain $c=\frac{\alpha k}{\beta}$. This completes the proof of Theorem 3.2.

3. 2. Characterizations by Order Statistics

If $X_1, X_2, \dots, X_n$ be the n independent copies of the random variable X with absolutely continuous distribution function $F(x)$ and pdf $f(x)$, and if $X_{1,n} \le X_{2,n} \le \dots \le X_{n,n}$ be the corresponding order statistics, then it is known from Ahsanullah et al. [17], chapter 5, or Arnold et al. [18], chapter 2, that $X_{j,n} \mid X_{k,n} = x$, for $1 \le k < j \le n$, is distributed as the $(j-k)th$ order statistics from $(n-k)$ independent observations from the random variable V having the pdf $f_V(v \mid x)$ where $f_V(v \mid x)=\frac{f(v)}{1-F(x)}$, $0 \le v < x$, and $X_{i,n} \mid X_{k,n} = x, 1 \le i < k \le n$, is distributed as ith order statistics from k independent observations from the random variable W having the pdf $f_W(w|x)$ where $f_W(w|x)=\dfrac{f(w)}{F(x)}$, $w<x$. Let $S_{k-1}=\frac{1}{k-1}\left(X_{1,n}+X_{2,n}+\ldots+X_{k-1,n}\right)$, and $T_{k,n}=\frac{1}{n-k}\left(X_{k+1,n}+X_{k+2,n}+\ldots+X_{n.n}\right)$.

Theorem 3.3: Suppose the random variable X satisfies the Assumption 3.1 with $\omega=\gamma$ and $\delta=\infty$, then $E(S_{k-1}|X_{k,n}=x)=g(x)\tau(x)$, where $\tau(x)=\dfrac{f(x)}{F(x)}$ and $g(x)$ being given by Eq. (26), if and only if X has the pdf

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

Proof: It is known that $E(S_{k-1}|X_{k,n}=x)=E(X|X\le x)$; see Ahsanullah et al. [17], and David and Nagaraja [19]. Hence, by Theorem 3.1, the result follows.

Theorem 3.4: Suppose the random variable X satisfies the Assumption 3.1 with $\omega=\gamma$ and $\delta=\infty$, then $E(T_{k,n}|X_{k,n}=x)=\tilde{g}(x)\dfrac{f(x)}{1-F(x)}$, where

$$\tilde{g}(x)=\frac{\left(E(X)-g(x)f(x)\right)}{f(x)},$$

$g(x)$ being given by Eq. (26) and $E(X)$ being given by Eq. (7), if and only if X has the pdf

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

Proof: Since $E(T_{k,n}|X_{k,n}=x)=E(X|X\ge x)$, see Ahsanullah et al. [17], and David and Nagaraja [19], the result follows from Theorem 3.2.

3. 3. Characterization by Upper Record Values

For details on record values, see Ahsanullah [20]. Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed absolutely continuous random variables with distribution function $F(x)$ and pdf $f(x)$. If $Y_n=\max(X_1, X_2, \ldots, X_n)$ for $n\ge 1$ and

$Y_j > Y_{j-1}$, $j > 1$, then X_j is called an upper record value of $\{X_n,\ n \geq 1\}$. The indices at which the upper records occur are given by the record times $\{U(n) > \min(j \mid j > U(n+1), X_j > X_{U(n-1)}, n > 1)\}$ and $U(1) = 1$. Let the *nth* upper record value be denoted by $X(n) = X_{U(n)}$.

Theorem 3.5: Suppose the random variable X satisfies the Assumption 3.1 with $\omega = 0$ and $\delta = \infty$, then $E(X(n+1) \mid X(n) = x) = \tilde{g}(x) \dfrac{f(x)}{1 - F(x)}$, where

$$\tilde{g}(x) = \frac{\left(E(X) - g(x) f(x)\right)}{f(x)},$$

$g(x)$ being given by Eq. (26) and $E(X)$ being given by Eq. (7), if and only if X has the pdf

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

Proof: It is known from Ahsanullah et al. [17], and Nevzorov [21] that $E(X(n+1) \mid X(n) = x) = E(X \mid X \geq x)$. Then, the result follows from Theorem 3.2.

4. ESTIMATION OF PARAMETERS AND APPLICATIONS

The estimation of the parameters of Dagum (4P) distribution was carried out by using the method of maximum likelihood (MLE). In Sub-Section 4.1, the maximum likelihood system of equations of Dagum (4 P) distribution is derived. In Sub-Section 4.2, using the EasyFit software, parameters of Dagum (4P) distribution are estimated using MLEs by considering a real-world data set example. The results are presented in Table 4, including those of the Birnbaum-Saunders (3P), Burr (3P) and Dagum (3P) distributions for testing the goodness of fit of Dagum (4P) distribution.

4. 1. The Method of Maximum Likelihood

Given a sample $\{x_i\}$, $i = 1, 2, 3, \ldots, n$, the likelihood function of the Dagum (4P) distribution pdf (1) is given by $L = \prod_{i=1}^{n} f(x_i)$, that is,

$$L=\prod_{i=1}^{n}\frac{\alpha k\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha k-1}}{\beta\left(1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}.$$

The objective of the likelihood function approach is to determine those values of the parameters that maximize the function L. Suppose its log-likelihood function is given by

$$\begin{aligned}R=\ln(L)= \ln(\alpha)-\ln(\beta)+\ln(k)+(\alpha k-1)\sum_{i=1}^{n}\ln\left(\frac{x_i-\gamma}{\beta}\right)\\ -(k+1)\sum_{i=1}^{n}\ln\left(1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\right)\end{aligned}. \tag{32}$$

Differentiating Eq. (32) partially with respect to the respective parameters, the maximum likelihood system of equations is given by

$$\left.\begin{aligned}
&\frac{\partial R}{\partial k}=\frac{1}{k}+(\alpha)\sum_{i=1}^{n}\ln\left(\frac{x_i-\gamma}{\beta}\right)-\sum_{i=1}^{n}\ln\left(1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\right)=0,\\
&\frac{\partial R}{\partial \alpha}=\frac{1}{\alpha}+(k)\sum_{i=1}^{n}\ln\left(\frac{x_i-\gamma}{\beta}\right)-(k+1)\sum_{i=1}^{n}\frac{\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\ln\left(\frac{x_i-\gamma}{\beta}\right)}{1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}}=0,\\
&\frac{\partial R}{\partial \beta}=\frac{1}{\beta}+\frac{(\alpha k-1)}{\beta}-\frac{\alpha(k+1)}{\beta^{\alpha+1}}\sum_{i=1}^{n}\frac{(x_i-\gamma)^{\alpha}}{1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}}=0,\\
&\frac{\partial R}{\partial \gamma}=(\alpha k-1)\sum_{i=1}^{n}\frac{1}{(x_i-\gamma)}-\frac{\alpha(k+1)}{\beta^{\alpha}}\sum_{i=1}^{n}\frac{(x_i-\gamma)^{\alpha-1}}{1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}}=0
\end{aligned}\right\}. \tag{33}$$

The maximum likelihood estimates (MLE) of the parameters $\{k,\alpha,\beta,\gamma\}$ can be obtained by solving the maximum likelihood system of equations (33) numerically by Newton-Raphson's iteration method using some computer software.

4. 2. Applications

In this sub-section, the goodness of fit tests of the Dagum (4 P) distribution vis-à-vis Birnbaum-Saunders (3P), Burr (3P) and Dagum (3P) distributions will be provided by considering a real-world data set example.

4. 2. 1. Example

This example considers a random sample of the tensile fatigue characteristics of 100 yarn data as reported by Quesenberry and Kent [1], which are provided in **Table 6**. Based on this example, we tested the chi-squared goodness-of-fit of Dagum (4 P) distribution to this data and compared it with some competitor distributions: Birnbaum-Saunders (3P), Burr (3P) and Dagum (3P) distributions.

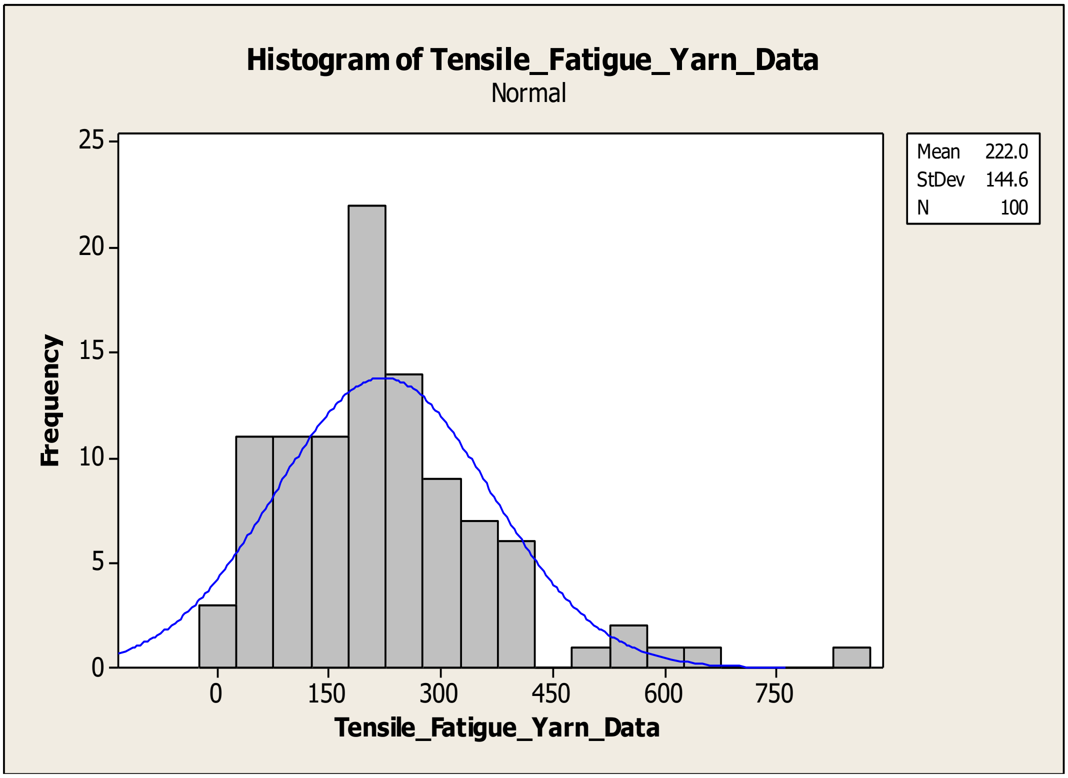
Table 6. Failure time data on 100 cms. Yarn at 2.3 % strain level (sample size n = 100).

86	**175**	**157**	**282**	**38**	**211**	**497**	**246**	**393**	**198**
146	**176**	**220**	**224**	**337**	**180**	**182**	**185**	**396**	**264**
251	**76**	**42**	**149**	**65**	**93**	**423**	**188**	**203**	**105**
653	**264**	**321**	**180**	**151**	**315**	**185**	**568**	**829**	**203**
98	**15**	**180**	**325**	**341**	**353**	**229**	**55**	**239**	**124**
249	**364**	**198**	**250**	**40**	**571**	**400**	**55**	**236**	**137**
400	**195**	**38**	**196**	**40**	**124**	**338**	**61**	**286**	**135**
292	**262**	**20**	**90**	**135**	**279**	**290**	**244**	**194**	**350**
131	**88**	**61**	**229**	**597**	**81**	**398**	**20**	**277**	**193**
169	**264**	**121**	**166**	**246**	**186**	**71**	**284**	**143**	**188**

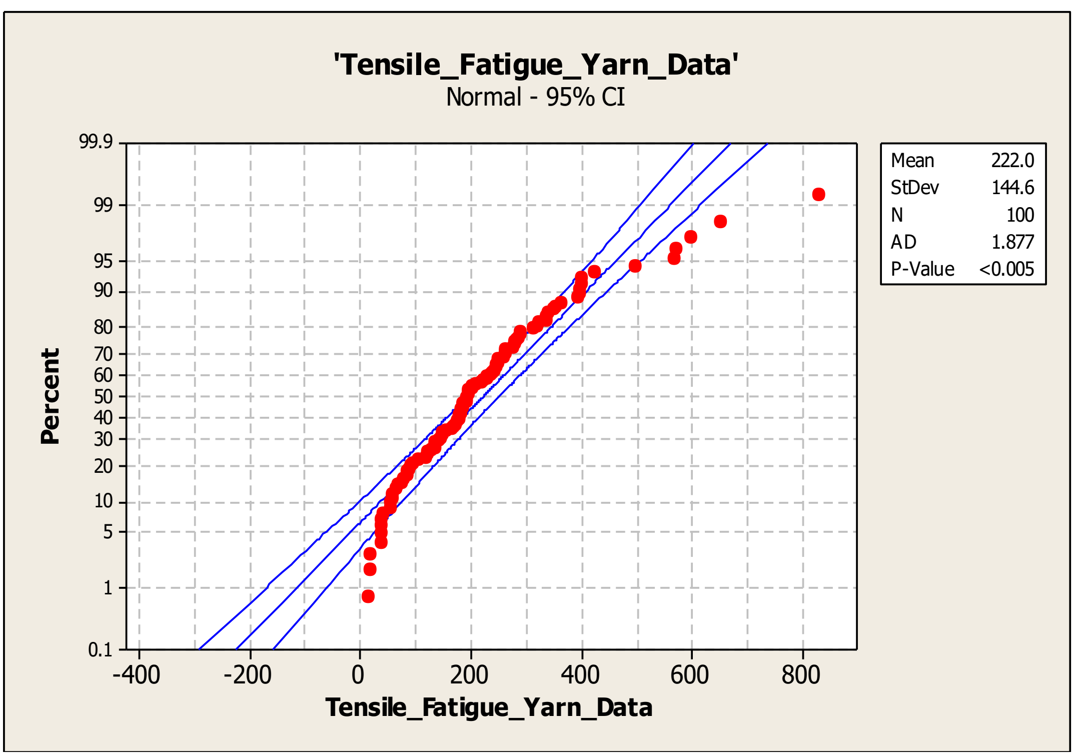
The descriptive statistics of the above-mentioned tensile data are computed in **Table 7**. Furthermore, we have drawn the histogram and normal quantile plot of the data, which are given in **Figure 6**.

Table 7. Descriptive Statistics of the Tensile Fatigue Characteristics of 100 Yarn Data.

Statistic	Value	Statistic	Value	Percentile	Value
Sample Size	100	Kurtosis	6.16403	Min	15
Range	814	Mode	180, 264	5%	38.1
Mean	221.98	Midrange	422	10%	55.6
Variance	20914.0			25% (Q1)	125.75
Std. Deviation	144.62			50% (Median)	195.5
Coef. of Variation	0.65149			75% (Q3)	283.5
Std. Error	14.462			90%	397.8
Skewness	1.3808			95%	564.45
Excess Kurtosis	3.0709			Max	829



(A)



(B)

Figure 6. Histogram **(A)** and Normality Assessment **(B)** of the Tensile Fatigue Characteristics of Yarn Data

From the **Figure 6** (of histogram and normal quantile plot), it is obvious that the shape of tensile fatigue characteristics of 100 yarn data is positively skewed. This is also confirmed from the skewness (1.3808) and kurtosis (6.16403) of the data as computed in **Table 5**. Since fitting of a probability distribution to the tensile fatigue characteristics of 100 yarn data may be helpful in predicting the probability or forecasting the frequency of occurrence of the tensile fatigue characteristics of 100 yarn data, this suggests that ‘ y ’, the tensile fatigue data, could possibly be modeled by some skewed distributions. As such we have tested the fitting of the Dagum (4P), Birnbaum-Saunders (3P), Burr (3P) and Dagum (3P) distributions to the tensile fatigue characteristics of 100 yarn data. For this, we have estimated the parameters of these distributions, which are provided in the **Table 8**. Using Chi-Squared Goodness-of-Fit Test, the goodness of fit of the Dagum (4P), Birnbaum-Saunders (3P), Burr (3P) and Dagum (3P) distributions to the tensile fatigue characteristics of 100 yarn data is provided in the **Table 9**.

Table 8. Fitting Results (Estimation of the Parameters)

#	Distributions	Parameter Estimates
1	Dagum (4P)	k = 0.27144, α = 4.3517, β = 321.03, γ = 13.869
2	Dagum (3P)	k = 0.3546, α = 4.1986, β = 305.54
3	Burr (3P)	k = 5.7813, α = 1.7986, β = 610.81
4	Birnbaum-Saunders (3P) (or Fatigue-Life (3P)).	α = 0.44646, β = 289.33, γ = - 96.201

Table 9. Comparison Criteria/Ranking of Fitted Distributions (P-Values and Test Statistics Analysis, Based on the Chi-Squared Goodness-of-Fit at the Level of Significance = 0.05)

	Dagum (4P) (Rank 1)	**Dagum (3P) (Rank 2)**	**Burr (3P) (Rank 3)**	**Birnbaum-Saunders (3P) (or Fatigue-Life (3P)) (Rank 4)**
Test Statistic	4.0796	4.1655	6.1234	6.5999
Critical Value	12.592	12.592	12.592	12.592
P-Value	0.66591	0.65429	0.40951	0.35944

From **Table 9**, we observed that the Dagum (4P), Dagum (3P), Burr (3P) and Birnbaum-Saunders (3P) distributions fitted reasonably well to the tensile fatigue data. However, the Dagum (4P) distribution model produces the highest p-value and the smallest test statistic value, and therefore fitted better than the Dagum (3P), Burr (3P) and Birnbaum-Saunders (3P) distributions. For the parameters estimated in Table 8, the probability density functions (pdf s) of the Dagum (4 P), Dagum (3P), Burr (3P) and Birnbaum-Saunders (3P) distributions have been superimposed on the histogram the tensile fatigue characteristics of 100 yarn data, as provided in Figure 7, from which we also observed that the Dagum (4P) distribution models the tensile fatigue characteristics of 100 yarn data reasonably well.

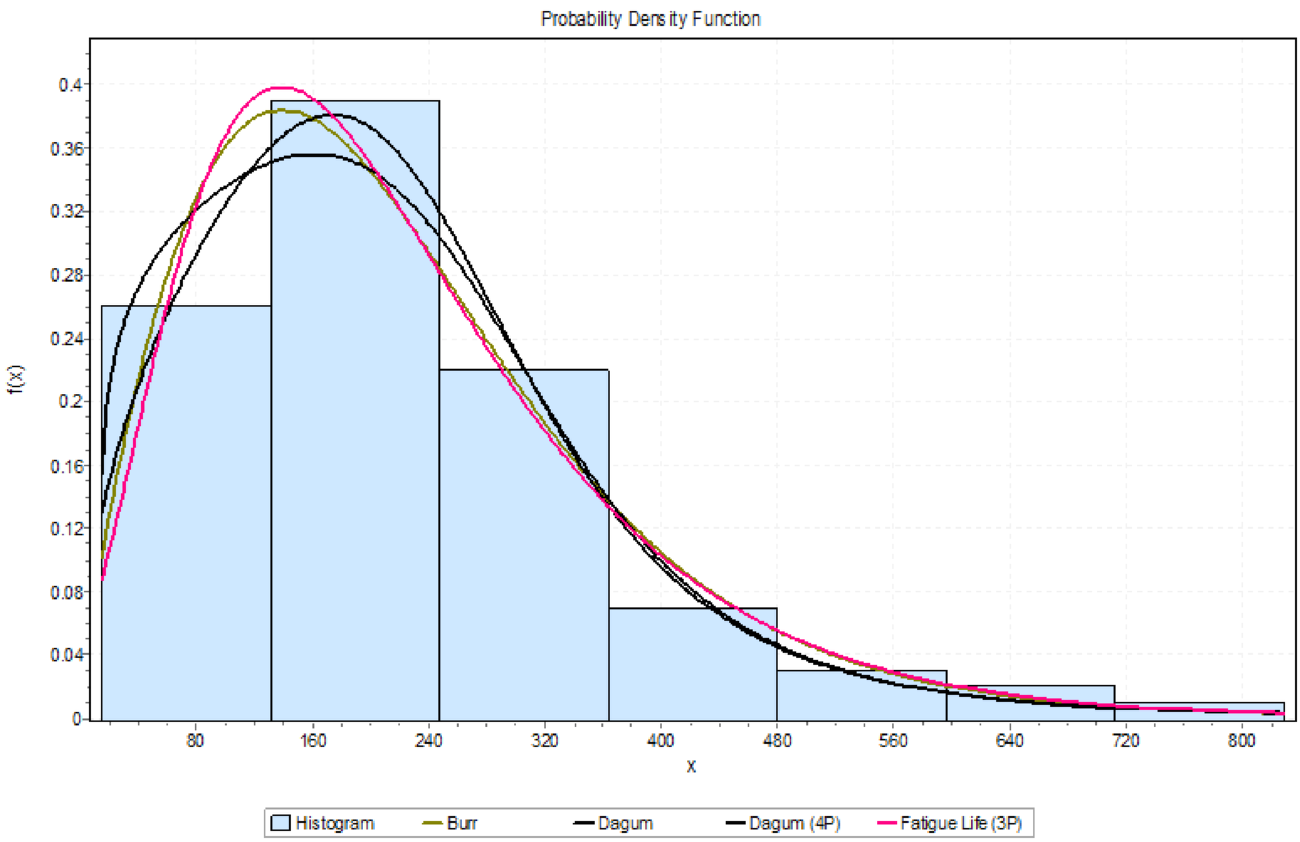


Figure 7. Fitting of the pdfs of the Dagum (4 P), Dagum (3P), Burr (3P) and Birnbaum-Saunders (3P) Distributions to the Tensile Fatigue Characteristics of Yarn Data

5. SOME CONCLUDING REMARKS

In this paper, we have considered the Dagum (4P) distribution introduced by Dagum [2, 3]. We have reviewed the Dagum (4P) distribution first, and then established its several new statistical properties, including the reliability analysis, the estimation of the parameters, computations of percentage points and characterizations. We have used a real life-time data sets to show the applications of the Dagum (4P) distribution. It is hoped that the findings of this paper will be quite useful to the researchers and practitioners in various fields of theoretical and applied sciences. One can also consider of developing bootstrap control charts for the percentiles of the Dagum (4P) distribution, which is an important area of studies in quality and reliability engineering.

ACKNOWLEDGEMENT

The authors are thankful to the comments and suggestions of the Editor-in-Chief and Reviewers which considerably improved the presentation of the paper.

References

[1] Quesenberry, C. P., and Kent, J., Selecting among probability distributions used in reliability. *Technometrics,* 24(1) (1982) 59-65. https://doi.org/10.2307/1267579

[2] Dagum, C., A model of income distribution and the conditions of existence of moments of finite order. *Bulletin of the International Statistical Institute,* 46 (1975) 199-205

[3] Dagum, C., A new model of personal income distribution: Specification and estimation. *Economie Appliquée,* 30 (1977) 413-437. https://doi.org/10.1007/978-0-387-72796-7_1

[4] Johnson, N. L., Kotz, S., and Balakrishnan, N., Continuous Univariate Distributions, Volumes 1 – 2 (1994), (6th Edition), John Wiley & Sons, New York, USA.

[5] Kleiber, C., and Kotz, S. (2003). Statistical Size Distributions in Economics and Actuarial Sciences (2003), Wiley, New York, USA.

[6] Kleiber, C., A guide to the Dagum distributions. In Modeling income distributions and Lorenz curves (2008) (pp. 97-117), Springer, New York, USA. https://doi.org/10.1007/978-0-387-72796-7_6

[7] Dey, S., Al-Zahrani, B., and Basloom, S., Dagum Distribution: Properties and Different Methods of Estimation. *International Journal of Statistics and Probability*, 6(2) (2017) 74-92. DOI:10.5539/ijsp.v6n2p74

[8] Domma, F., Condino, F., and Giordano, S., A new formulation of the Dagum distribution in terms of income inequality and poverty measures. *Physica A: Statistical Mechanics and its Applications*, 511 (2018) 104-126. https://doi.org/10.1016/j.physa.2018.07.027

[9] Gradshteyn, I. S., and Ryzhik, I. M., Table of integrals, series, and products (2000), Academic Press, Inc., San Diego, California, USA.

[10] Nichols, M.D., and Padgett, W.J., A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, 22 (2006) 141-151. https://doi.org/10.1002/qre.691

[11] Cramer, D., Basic Statistics for Social Research (1997), Routledge, USA.

[12] Joanes, D. N., and Gill, C. A., Comparing Measures of Sample Skewness and Kurtosis, *The Statistician* 47(1) (1998) 183-189. https://doi.org/10.1111/1467-9884.00122

[13] Abramowitz, M., and Stegun, I. A., Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables (1972), Dover, New York, NY, USA.

[14] Stuart, A., and Ord, J. K., Kendall's Advanced Theory of Statistics, Volume 1, Distribution Theory (2010), 6th Edition, Wiley, New York, USA.

[15] Shannon, C. E., A mathematical theory of communication. *Bell System Tech. J.* Vol. 27 (1948) 379-423; 623-656. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[16] Shakil, M., Ahsanullah, M., and Kibria, B. G., On the Characterizations of Chen's Two-Parameter Exponential Power Life-Testing Distribution. *Journal of Statistical Theory and Applications*, 17 (3) (2018) 393-407. https://doi.org/10.2991/jsta.2018.17.3.1

[17] Ahsanullah, M., Nevzorov, V. B., and Shakil, M., An Introduction to Order Statistics (2013), Atlantis Press, Paris, France.

[18] Arnold, B.C., Balakrishnan, and Nagaraja, H. N., First Course in Order Statistics (2005), Wiley, New York, USA.

[19] David, H. A., and Nagaraja, H. N., Order Statistics (2003), Third Edition, Wiley, New York, USA.

[20] Ahsanullah, M., Introduction to Record Statistics (1995), Nova Science Publishers, Inc., New York, USA.

[21] Nevzorov, V. B., Records: Mathematical Theory (2001), Translation of Mathematical Monograph, American Mathematical Society, Rhode Island, USA.