A Quasi Lindley Pareto Distribution

Asad Ali¹*, Qaisar Rashid¹, Muhammad Zubair², and Muhammad Tariq Jamshaid²

¹Department of Quantitative Methods, SBE, University of Management and Technology, Lahore, Pakistan
²Department of Statistics, University of Sargodha, Pakistan

Abstract: In applied sciences, the lifetime’s data models have been constructed to facilitate better modeling and significant progress towards the reliability analysis or survival analysis. For this purpose, a new distribution “Quasi Lindley Pareto Distribution (QLPD)” has been introduced. It’s Moments generating function, moments, mean, variance, coefficient of variation, skewness, kurtosis, survival function, hazard function, and distribution of extreme order statistics have been discussed. The maximum likelihood method has been discussed for estimating its parameters. The distribution has been fitted to real life data set to check its goodness of fit to which earlier the Lindley distribution and Quasi Lindley distribution have been fitted and it is found that QLPD provides closer fits than those by the Lindley distribution and Quasi Lindley distribution.

Keywords: Moments, Order statistics, Survival, Hazard, Goodness of fit

1. INTRODUCTION


Let Y be a random variable following the two parameter Quasi Lindley distribution ($\alpha, \theta$) with the following probability density function:

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

Pareto distribution was established by Pareto [11] to elaborate the unequal distribution of wealth. There are many studies on Pareto distribution such as the generalized Pareto distribution was derived by Pickands [12], the beta-Pareto distribution was discussed by Akinsete et al. [13] and the beta generalized Pareto distribution was proposed by Mahmoudi [14].
The mixed distribution is one of the most important concepts for developing a new distribution. For example, Zakerzadeh and Dolati [15] used gamma (α, θ) and gamma (α + 1, θ) to create a Generalized Lindley distribution (GLD). Nedjar and Zeghdoudi [16] determined a new distribution, based on gamma (2, θ) and one parameter Lindley distribution known as gamma Lindley distribution.

The T-X family of distributions defined by Alzaatreh et al. [17] is a new method to generate families of the continuous distributions. The cumulative distribution function (CDF) or distribution function (DF) of this family is defined as:

$$G(y) = \int_0^{\frac{F(y)}{1-F(y)}} f(t) \, dt \quad (2)$$

Using Quasi Lindley distribution, (2) becomes Quasi-Lindley X-family of distribution with CDF and Probability density function (pdf) such that:

$$G(y) = 1 - \exp(-\theta \delta) \{1 + \left(\frac{y}{\alpha+1}\right) \delta\} \quad (3)$$

Where $\delta = \frac{F(y)}{1-F(y)}$

$$g(y) = \left[\frac{\theta}{(\alpha+1)}\right] \left[\frac{F(y)}{1-F(y)}\right] \left[\frac{\theta^2}{(\alpha+1)}\right] \times e^{-\theta \left[\frac{F(y)}{1-F(y)}\right]} \quad (4)$$

We considered CDF and pdf a Pareto distribution as:

$$F(y) = 1 - \left(\frac{\beta}{y}\right)^k \quad f(y) = \frac{k \beta^k}{y^{k+1}} \quad y > \beta$$

1.1 A Quasi Lindley Pareto Distribution

By utilizing (3) and (4), we get CDF and pdf of Quasi Lindley Pareto distribution (QLPD) with scale parameter(s) $\alpha$ and $\gamma$, shape parameter(s) $\theta, \beta$ and $k$ is given by respectively:

$$Q(y) = 1 - \exp \left[\gamma - \theta \left(\frac{\gamma}{\beta}\right)^k - 1\right] \left[1 + \left(\frac{\theta}{\alpha+1}\right) \left\{\left(\frac{\gamma}{\beta}\right)^k - 1\right\}\right] \quad y > 0, k, \alpha, \gamma, \theta, \beta > 0 \quad (5)$$

and

$$q(y) = \left[\frac{\theta k e^\theta}{\beta (1+\alpha)}\right] \left[\left(\gamma - \theta \left(\frac{\gamma}{\beta}\right)^k - 1\right) + \theta \left(\frac{\gamma}{\beta}\right)^{2k-1}\right] \exp\left\{ -\theta \left(\frac{\gamma}{\beta}\right)^k \right\} \quad (6)$$

The nature of QLPD for different values of its parameters $\alpha, \gamma, \theta, \beta$ and $k$ has shown graphically in the Fig. 1 (a), (b), (c) and (d). In Fig. 1 (d) seven choices of QLPD parameters have been combined for comparisons.

2. SOME PROPERTIES

**Survival Function:** The Survival function for random variable $Y$ from QLPD($Y; k, \gamma, \alpha, \theta, \beta$) is given as:

$$S(y) = \exp \left[-\theta \left(\frac{\gamma}{\beta}\right)^k - 1\right] \left[1 + \left(\frac{\theta}{\alpha+1}\right) \left\{\left(\frac{\gamma}{\beta}\right)^k - 1\right\}\right] \quad (7)$$
A Quasi Lindley Pareto Distribution

Fig. 1 (a)

Fig. 1 (b)

Fig. 1 (c)

Fig. 1 (d)
**Hazard Function:** The failure rate function has also been called hazard function, determined, by Eq. (6) over survival function Eq. (7). So, hazard function of QLPD(Y; k, \( \gamma \), \( \alpha \), \( \theta \), \( \beta \)) is:

\[
H(y) = \frac{\theta e^{\theta (\alpha - 1) + \theta (\frac{\gamma}{\beta})^{k} \frac{y}{\beta}^{k-1}}}{\beta \left[ (1+\alpha) + \theta \left(\frac{\gamma}{\beta}\right)^{k} - 1\right]}
\]

(8)

**Moment Generating Function:** The moment generating function of QLPD(Y; k, \( \gamma \), \( \alpha \), \( \theta \), \( \beta \)) is:

\[
M_{Y}(t) = \frac{e^{\theta (\alpha - 1) + \theta (\frac{\gamma}{\beta})^{k} \frac{t}{\beta}^{k-1}}}{(1+\alpha) \sum_{i=0}^{\infty} \Gamma \left( \frac{i}{k} + 1, \theta \right) + \Gamma \left( \frac{i}{k} + 2, \theta \right)} \left(\frac{\beta}{\theta \beta^{k}}\right)^{i} \frac{t^{i}}{i!}
\]

(9)

This function has been utilized to determine \( i^{th} \) number of moment about origin.

**Moments:** The \( i^{th} \) moment about origin of random variable Y from QLPD(Y; k, \( \gamma \), \( \alpha \), \( \theta \), \( \beta \)) are given below:

\[
E(y^{i}) = \frac{e^{\theta (\alpha - 1) + \theta (\frac{\gamma}{\beta})^{k} \frac{i}{\beta}^{k-1}}}{(1+\alpha) \sum_{i=0}^{\infty} \Gamma \left( \frac{i}{k} + 1, \theta \right) + \Gamma \left( \frac{i}{k} + 2, \theta \right)} \left(\frac{\beta}{\theta \beta^{k}}\right)^{i}
\]

(10)

By putting \( i = 1, 2, 3 \) and \( 4 \) in Eq. (10), first four moments about origin have been derived to determine the mean, variance, coefficient of variation, skewness and kurtosis.

First four moments about mean are defined as:

\[
\mu_{1} = m_{1} - m_{1}
\]

(11)

\[
\mu_{2} = m_{2} - (m_{1})^{2}
\]

(12)

\[
\mu_{3} = m_{3} - 3m_{2}m_{1} + 2(m_{1})^{3}
\]

(13)

\[
\mu_{4} = m_{4} - 6m_{3}m_{1} + 3m_{2}m_{1} - 4(m_{1})^{4}
\]

(14)

From Eq. (10), we have,

\[
m_{1} = E(y^{1}) = \frac{e^{\theta (\alpha - 1) + \theta (\frac{\gamma}{\beta})^{k} \frac{1}{\beta}^{k-1}}}{(1+\alpha) \sum_{i=0}^{\infty} \Gamma \left( \frac{1}{k} + 1, \theta \right) + \Gamma \left( \frac{1}{k} + 2, \theta \right)} \left(\frac{\beta}{\theta \beta^{k}}\right)
\]

(15)

\[
m_{2} = E(y^{2}) = \frac{e^{\theta (\alpha - 1) + \theta (\frac{\gamma}{\beta})^{k} \frac{2}{\beta}^{k-1}}}{(1+\alpha) \sum_{i=0}^{\infty} \Gamma \left( \frac{2}{k} + 1, \theta \right) + \Gamma \left( \frac{2}{k} + 2, \theta \right)} \left(\frac{\beta}{\theta \beta^{k}}\right)^{2}
\]

(16)

\[
m_{3} = E(y^{3}) = \frac{e^{\theta (\alpha - 1) + \theta (\frac{\gamma}{\beta})^{k} \frac{3}{\beta}^{k-1}}}{(1+\alpha) \sum_{i=0}^{\infty} \Gamma \left( \frac{3}{k} + 1, \theta \right) + \Gamma \left( \frac{3}{k} + 2, \theta \right)} \left(\frac{\beta}{\theta \beta^{k}}\right)^{3}
\]

(17)

\[
m_{4} = E(y^{4}) = \frac{e^{\theta (\alpha - 1) + \theta (\frac{\gamma}{\beta})^{k} \frac{4}{\beta}^{k-1}}}{(1+\alpha) \sum_{i=0}^{\infty} \Gamma \left( \frac{4}{k} + 1, \theta \right) + \Gamma \left( \frac{4}{k} + 2, \theta \right)} \left(\frac{\beta}{\theta \beta^{k}}\right)^{4}
\]

(18)

From above equations, we have,

\[
\mu_{1} = 0
\]

(19)

\[
\mu_{2} = Var(y)
\]

(20)
\[\mu_3 = \frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{3}{k}, 1, \theta \right) + \Gamma_k^{3} \right] \frac{\beta^3}{\theta_k^3} - 3 \frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{2}{k}, 1, \theta \right) + \Gamma_k^{2} \right] \frac{\beta^2}{\theta_k^2} \times \]

\[\frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1} \right] \frac{\beta}{\theta_k^1} + 2 \left( \frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1} \right] \frac{\beta}{\theta_k^1} \right)^3 \]  \tag{21}

\[\mu_4 = \frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{4}{k}, 1, \theta \right) + \Gamma_k^{4} \right] \frac{\beta^4}{\theta_k^4} - 6 \frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{3}{k}, 1, \theta \right) + \Gamma_k^{3} \right] \frac{\beta^3}{\theta_k^3} \times \]

\[\frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1} \right] \frac{\beta}{\theta_k^1} + 3 \frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1} \right] \frac{\beta}{\theta_k^1} \times \]

\[\frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{2}{k}, 1, \theta \right) + \Gamma_k^{2} \right] \frac{\beta^2}{\theta_k^2} - 4 \left( \frac{e^\theta}{(1+\alpha)} \left[ (\alpha - \theta) \Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1} \right] \frac{\beta}{\theta_k^1} \right)^4 \]  \tag{22}

**Mean:** The mean of random variable Y fromQLPD(Y; k, y, \alpha, \theta, \beta) is defined as:

\[E(y) = \frac{e^\theta}{(1+\alpha)} \left( \frac{\beta}{\theta_k^1} \right) \left[ (\alpha - \theta) \Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1} \right] \]  \tag{23}

**Variance (Var):** The variance of random variable Y fromQLPD(Y; k, y, \alpha, \theta, \beta) is obtained as follow:

\[Var(y) = \frac{e^\theta}{(1+\alpha)} \left( \frac{\beta}{\theta_k^1} \right) \left[ (\alpha - \theta) \Gamma \left( \frac{2}{k}, 1, \theta \right) + \Gamma_k^{2} \right] - \left[ (\alpha - \theta) \Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1} \right]^2 \frac{e^\theta}{(1+\alpha)} \]  \tag{24}

**Coefficient of Variance (C.V):** The coefficient of variance for random variable Y fromQLPD(Y; k, y, \alpha, \theta, \beta) is obtained as follow:

\[C.V = \frac{Var(y)}{E(y)} \]  \tag{25}

\[C.V = \frac{\sqrt{\left( \frac{\alpha-\theta}{\Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1}} \right) - 1}}{\frac{\alpha-\theta}{\Gamma \left( \frac{1}{k}, 1, \theta \right) + \Gamma_k^{1}}} \]  \tag{26}

Similarly, Skewness and Kurtosis will be determined by using Eq. (21, 22 and 24):

\[Skewness = \frac{\mu_3^2}{\mu_2^3} \]  \tag{27}

\[kurtosis = \frac{\mu_4}{\mu_2^2} \]  \tag{28}

**Order Statistics (OS):** Ordering of positive continuous random variables is an important tool for judging the comparative behavior. Let, Y_1, Y_2, ..., Y_n are random samples fromQLPD(Y; k, y, \alpha, \theta, \beta) and Y_{r,n} is the r^{th} order statistics with pdf given as follow:

\[q_{r,n}(y) = \frac{n!}{(r-1)!} q(y) [Q(y)]^{r-1} [1 - Q(y)]^{n-r} \]  \tag{29}
The Quasi Lindley Pareto Distribution (QLPD) has been fitted to a data set to which earlier the Lindley distribution (LD) and Quasi Lindley Distribution (QLD) have been fitted and it was found that QLPD provides better fit than those by LD and QLD. Here the fitting of the QLPD to the data set have been presented in the Table 1. The data is regarding the survival times (in days) of 72 guinea pigs infected

\[ q_{r,n}(y) = \frac{n!}{(r-1)!} \left[ \frac{\theta k e^\theta}{\beta (1+\alpha)} \right] \left[ (y - \theta) \left( \frac{y}{\beta} \right)^{k-1} + \theta \left( \frac{y}{\beta} \right)^{2k-1} \right] \exp(-\theta \left( \frac{y}{\beta} \right)^k) \times \left[ 1 - \exp \left[ -\theta \left( \frac{y}{\beta} \right)^k \right] - 1 \left[ 1 + \left( \frac{\theta}{\alpha + 1} \right) \left( \frac{y}{\beta} \right)^k - 1 \right] \right] \times \exp(-\frac{\theta}{\alpha + 1}) \left( \frac{y}{\beta} \right)^k \right] \times \exp(-\theta \left( \frac{y}{\beta} \right)^k) \right] \times \left[ 1 + \left( \frac{\theta}{\alpha + 1} \right) \left( \frac{y}{\beta} \right)^k - 1 \right] \right] \] (30)

So, the pdf of the smallest OS at r=1, \( q_{1,n}(y) \) and the largest OS at r=n, \( q_{n,n}(y) \) are obtained by:

\[ q_{1,n}(y) = n \left[ \frac{\theta k e^\theta}{\beta (1+\alpha)} \right] \left[ (y - \theta) \left( \frac{y}{\beta} \right)^{k-1} + \theta \left( \frac{y}{\beta} \right)^{2k-1} \right] \times \left[ 1 + \left( \frac{\theta}{\alpha + 1} \right) \left( \frac{y}{\beta} \right)^k - 1 \right] \right] \times \exp(-\theta \left( \frac{y}{\beta} \right)^k) \] (31)
and

\[ q_{n,n}(y) = n \left[ \frac{\theta k e^\theta}{\beta (1+\alpha)} \right] \left[ (y - \theta) \left( \frac{y}{\beta} \right)^{k-1} + \theta \left( \frac{y}{\beta} \right)^{2k-1} \right] \exp(-\theta \left( \frac{y}{\beta} \right)^k) \times \left[ 1 - \exp \left[ -\theta \left( \frac{y}{\beta} \right)^k - 1 \right] \right] \] (32)

3. ESTIMATION OF PARAMETERS

Let, \( y_1, y_2, ..., y_n \) be a random sample of size n from the Quasi Lindley Pareto distribution and the likelihood function of QLPD is defined as:

\[ L = \left[ \frac{\theta k e^\theta}{\beta (1+\alpha)} \right] n \prod_{i=1}^{n} \left[ (y - \theta) \left( \frac{y_i}{\beta} \right)^{k-1} + \theta \left( \frac{y_i}{\beta} \right)^{2k-1} \right] \exp\left[ -\frac{\theta}{\beta (1+\alpha)} \left( \frac{y_i}{\beta} \right)^k \right] \] (33)

Further, the log likelihood function is obtained as follow:

\[ logL = n log \left[ \frac{\theta k e^\theta}{\beta (1+\alpha)} \right] + \log \left[ \frac{(y - \theta) \beta k-1}{\beta k-1} \right] \prod_{i=1}^{n} y_i \left[ \frac{e^{\frac{\theta}{\beta}} \left( \frac{y_i}{\beta} \right)^{k-1}}{\beta k-1} \right] - \sum_{i=1}^{n} \left[ -\theta \left( \frac{y_i}{\beta} \right)^k \right] \] (34)

By taking partial derivatives with respect to parameters, following estimates are obtained:

\[ \hat{\alpha}_{mle} = -1 \] (35)

\[ \hat{\beta}_{mle} = \left[ \frac{-k \sum_{i=1}^{n} y_i^{k-1}}{-n - 3k + 2} \right]^{1/k} \] (36)

\[ \hat{\theta}_{mle} = \frac{n}{\sum_{i=1}^{n} \left( \frac{y_i}{\beta} \right)^k - n} \] (37)

\[ \hat{\gamma}_{mle} = \theta \] (38)

\[ \hat{k}_{mle} = \frac{3 \ln \left[ n \beta - \sum_{i=1}^{n} \ln y_i \right] - \ln n + \ln \theta - \ln \left[ \sum_{i=1}^{n} \ln \left( \frac{y_i}{\beta} \right)^k \right] + \ln \left( \frac{y_i}{\beta} \right)}{-2 \ln} \] (39)

4. GOODNESS OF FIT

The Quasi Lindley Pareto Distribution (QLPD) has been fitted to a data set to which earlier the Lindley distribution (LD) and Quasi Lindley Distribution (QLD) have been fitted and it was found that QLPD provides better fit than those by LD and QLD. Here the fitting of the QLPD to the data set have been presented in the Table 1. The data is regarding the survival times (in days) of 72 guinea pigs infected
with virulent tubercle bacilli, observed and reported by Bjerkedal [18].

In Table 1, the expected frequencies according to the Lindley distribution and Quasi Lindley distribution have also been given for ready comparison with those obtained by the Quasi Lindley Pareto distribution. It can be seen that the QLPD gives much closer fits than the LD and QLD of Shanker and Mishra [4] and thus provides a better alternative to the LD and QLD.

### Table 1. Data of survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by Bjerkedal [18]

<table>
<thead>
<tr>
<th>Survival Time (in days)</th>
<th>Observed frequency</th>
<th>LD</th>
<th>QLD</th>
<th>QLPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-80</td>
<td>8</td>
<td>16.14</td>
<td>10.71</td>
<td>10.24</td>
</tr>
<tr>
<td>80-160</td>
<td>30</td>
<td>21.91</td>
<td>26.95</td>
<td>27.73</td>
</tr>
<tr>
<td>160-240</td>
<td>18</td>
<td>15.39</td>
<td>17.71</td>
<td>17.66</td>
</tr>
<tr>
<td>240-320</td>
<td>8</td>
<td>9.00</td>
<td>9.14</td>
<td>8.59</td>
</tr>
<tr>
<td>320-400</td>
<td>4</td>
<td>5.47</td>
<td>4.26</td>
<td>4.53</td>
</tr>
<tr>
<td>400-480</td>
<td>3</td>
<td>1.80</td>
<td>1.86</td>
<td>2.31</td>
</tr>
<tr>
<td>480-560</td>
<td>1</td>
<td>2.29</td>
<td>1.34</td>
<td>0.94</td>
</tr>
<tr>
<td>Total</td>
<td>72</td>
<td>72.00</td>
<td>72.00</td>
<td>72.00</td>
</tr>
</tbody>
</table>

5. SUMMARY

In applied sciences, the lifetime’s data models are constructed to facilitate better modeling and significant progress towards the reliability analysis or survival analysis. In this paper, we introduced five-parameter Quasi Lindley Pareto Distribution (QLPD). The probability distribution function with different values of its parameters has been shown graphically. Several properties such as probability density function, distribution function, survival function, hazard function, moment generating function, moments, mean, variance, coefficient of variance, skewness, kurtosis and distribution of extreme order statistics have been derived. The estimation of parameters by the method of maximum likelihood has been discussed. Finally, the proposed distribution has been fitted to a real life data set to check its goodness of fit to which earlier the Lindley distribution (LD) and Quasi Lindley distribution (QLD) have been fitted and it is found that the QLPD provides closer fits than those by the LD and QLD. Therefore, it is suggested to use QLPD as a lifetime’s data model for better estimation.

6. REFERENCES