

# Quasi Lindley Geometric Distribution

L.S. Diab

College of Science for (girls)  
 Dept. of Mathematics, Al-Azhar University  
 Nasr City, 11884, Egypt

Hiba Z. Muhammed

Institute of Statistical Studies and Research  
 Department of Mathematical Statistics  
 Cairo University, Cairo, Egypt

## ABSTRACT

In this paper, we introduce a new class of lifetime distributions which is called the Quasi Lindley Geometric (QLG) distribution. This distribution obtained by compounding the Quasi Lindley and geometric distributions. Some structural properties of the proposed new distribution are discussed, including probability density function and explicit algebraic formulas for its survival and hazard functions, moment, moment generating function and mean deviations. We propose the method of maximum likelihood for estimating the model parameters and obtain the observed information matrix. A real data set is used to illustrate the importance and flexibility of the new distribution.

## Keywords

Quasi Lindley distribution, Geometric distribution, Moments, Maximum likelihood

## 1. INTRODUCTION

The Lindley distribution was introduced by Lindley (1958) as a new distribution useful to analyze lifetime data especially in applications modeling stress-strength reliability. In a recent paper Ghitany et al. (2008) studied the properties of the Lindley distribution under a carefully mathematical treatment. They also showed in a numerical example that the Lindley distribution gives better modeling for waiting times and survival times data than the exponential distribution. The use of the Lindley distribution could be a good alternative to analyse lifetime data within the competing risks approach as compared with the use of standard Exponential or even the Weibull distribution commonly used in this area. The Exponential distribution assumes constant hazard function, usually not an appropriated assumption for many competing risks data. Mazucheli and Achcar (2011) studied the applications of Lindley distribution to competing risks lifetime data. Also, the Lindley distribution has some nice properties to be used in lifetime data analysis as closed forms for the survival and hazard functions and good flexibility of fit.

Lindley (1958), introduced a one- parameter distribution, known as Lindley distribution, given by its probability density function

$$f(x, \theta) = g(x, \theta) = \frac{\theta^2}{\theta + 1} (1 + x)e^{-\theta x}; x > 0, \theta > 0,$$

the cumulative distribution function (cdf) of Lindley distribution is obtained as

$$F(x, \theta) = 1 - e^{-\theta x} \left[ 1 + \frac{\theta x}{\alpha + 1} \right], x > 0, \theta > 0.$$

Rama and Mishra (2013) introduced a new two-parameter Quasi Lindley distribution (QLD), of which the Lindley distribution (LD) is a particular case. They studied several properties of the QLD, and shown that the QLD is more flexible than Lindley and exponential distributions. Quasi Lindley distribution with parameters  $\alpha$  and  $\theta$  is defined by its probability density function (p.d.f)

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

It can easily be seen that at  $\alpha = \theta$ , the Equation (1.3) reduces to the Lindley distribution. And at  $\alpha = 0$ , it reduces to the gamma distribution with parameters  $(2, \theta)$ . The p.d.f. Equation (1.3) can be shown as a mixture of exponential ( $\theta$ ) and gamma(2,  $\theta$ ), distributions as follows

$$f(x, \theta, \alpha) = pg_1(x) + (1 - p)g_2(x)$$

where

$$p = \frac{1}{\alpha + 1}, g_1(x) = \theta e^{-\theta x} \text{ and } g_2(x) = \theta^2 e^{-\theta x}.$$

The cumulative distribution function (cdf) of QLD is obtained as

$$F(x, \theta, \alpha) = 1 - e^{-\theta x} \left[ 1 + \frac{\theta x}{\alpha + 1} \right], x > 0, \theta > 0, \alpha > -1.$$

where  $\theta$  is scale parameter.

Adamidis and Loukas (1998) introduced a two-parameter lifetime distribution with decreasing failure rate by compounding exponential and geometric distributions, which was named exponential geometric (EG) distribution. In the same way, the exponential Poisson (EP) and exponential logarithmic (EL) distributions were introduced and studied by Kus (2007) and Tahmasbi and Rezaei (2008), respectively. Recently, Chahkandi and Ganjali (2009) proposed a class of distributions, named exponential power series (EPS) distributions, by compounding exponential and power series distributions, where compounding procedure follows the same way that was previously carried out by Adamidis and Loukas (1998). In the same way, Barreto-Souza et al. (2010) and Lu and Shi (2011) introduced the Weibull-geometric (WG) and Weibull-Poisson (WP) distributions which naturally extend the EG and EP distributions, respectively. Barreto et al. (2009) presented a generalization of the exponential-Poisson distribution. Morais and Barreto-Souza (2011) defined the Weibull power series (WPS) class of distributions which contains the (EPS) distributions as sub-models. Adamidis et al. (2005) proposed the extended exponential-geometric (EEG) distribution which generalizes the EG distribution and discussed several of its statistical properties along with its reliability features.

In this paper we introduce a Quasi Lindley-geometric (QLG) distribution which generalizes the Geometric and Quasi Lindley distributions and study some of its properties. The paper is organized as follows. In Section 2, we define the QLG distribution, density, hazard rate function and survival function. In Section 3, we give some statistical properties of the new distribution. The order statistics and its moment are given in Section 4. Residual life and reversed residual life functions of QLG distribution are discussed in Section 5. Mean deviations from the mean and median are derived in Section 6. The least squares and weighted least squares estimators are introduced in Section 7. In Section 8, we demonstrate the maximum likelihood estimates of the unknown parameters.

Finally, Section 8 gives applications involving a real data set.

## 2. QUASI LINDLEY-GEOMETRIC DISTRIBUTION

Our class can be derived as follows. Suppose that a company has  $N$  systems functioning independently and producing a certain product at a given time, where  $N$  is a random variable, which is often determined by economy, customers demand, etc. The reason for considering  $N$  as a random variable comes from a practical viewpoint in which failure (of a device for example) often occurs due to the present of an unknown number of initial defects in the system. In this paper, we focus on the case in which  $N$  is discrete random variable following a geometric distribution (truncated at zero) with the probability mass function given by

$$P(n; p) = P(N = n) = (1-p)p^{n-1}, \text{ for } n \in \mathbb{N} \text{ and } p \in (0,1).$$

Note that  $N$  can also be taken to follow other discrete distributions, such as binomial, Poisson, logarithmic, etc, whereas they need to be truncated zero because one must have  $N \geq 1$ . Now, let  $X_1, X_2, \dots, X_N$  be  $N$  independent and identically distributed (iid) random variables following the Quasi Lindley distribution cumulative distribution function (cdf) (1.4). Let

$$X_{(1)} = \min\{X_i\}_{i=1}^N,$$

The conditional cumulative distribution of  $X_{(1)} | N = n$  is given by

$$\begin{aligned} G_{X_{(1)} | N=n}(x) &= 1 - [1 - F(x)]^n \\ &= 1 - \left[ e^{-\theta x} \left[ 1 + \frac{\theta x}{\alpha + 1} \right] \right]^n, \end{aligned}$$

and the unconditional cumulative distribution function of  $X_{(1)}$  is

$$\begin{aligned} F(x; \alpha, \theta, p) &= (1-p) \sum_{n=1}^{\infty} p^{n-1} \left\{ 1 - \left[ e^{-\theta x} \left[ 1 + \frac{\theta x}{\alpha + 1} \right] \right]^n \right\} \\ &= \frac{1 - e^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right)}{1 - p e^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right)}, x > 0. \end{aligned}$$

We shall refer to the distribution given by Equation (2.2) as the Quasi Lindley-geometric (QLG) distribution. If a random variable  $X$  has the QLG distribution, then we write  $X \sim QLG(\alpha, \theta, p)$ . The corresponding probability density function (pdf) of  $X_{(1)}$  is given by

$$\begin{aligned} f(x; \alpha, \theta, p) &= (1-p) \frac{\theta}{\alpha + 1} (\alpha + \theta x) e^{-\theta x} \times \\ &\quad \left[ 1 - p e^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right) \right]^{-2}, x > 0. \end{aligned}$$

It is evident that (2.3) is much more flexible than the Quasi Lindley distribution. It is clear that the QLD distribution contains as limiting cases: (i) the Lindley geometric distribution for  $\alpha = \theta$  (see, Zakerzadeh and Mahmoudi (2012)), (ii) the gamma geometric distribution for  $\alpha = 0$ ; and (iii) the Quasi Lindley distribution for  $p \downarrow 0$ . Figure 1 and Figure 2 illustrate some of the possible shapes of the pdf and cdf of QLG distribution for selected values of the parameters  $\theta, p$  and  $\lambda$ , respectively.

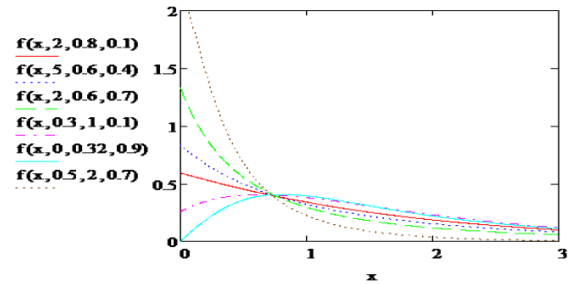


Figure 1: pdf of Quasi Lindley geometric distribution

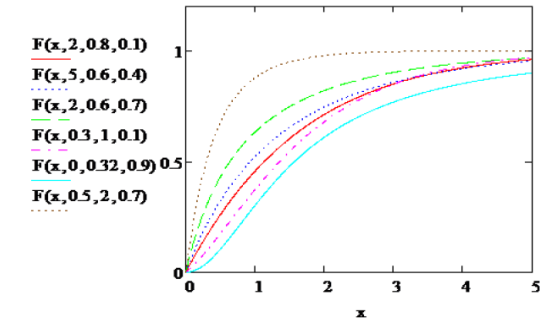


Figure 2: cdf of Quasi Lindley geometric distribution

The expressions given by Equations (2.2) and (2.3) are clearly complicated. For mathematical tractability, using the series expansion

$$(1-z)^{-k} = \sum_{j=0}^{\infty} \frac{\Gamma(k+j)}{\Gamma(k)j!} z^j,$$

Where  $|z| < 1$  and  $k > 0$  One can re-express Equations (2.2) and (2.3) as

$$F(x; \phi) = \sum_{j=0}^{\infty} p^j e^{-\theta j x} \left( 1 + \frac{\theta x}{\alpha + 1} \right)^j \left[ 1 - e^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right) \right],$$

and

$$f(x; \phi) = (1-p) \sum_{j=0}^{\infty} (j+1) p^j (\alpha + \theta x) \left( 1 + \frac{\theta x}{\alpha + 1} \right)^j e^{-\theta(j+1)x},$$

Again using the binomial expansion of  $\left( 1 + \frac{\theta x}{\alpha + 1} \right)^j$  given by

$$\left( 1 + \frac{\theta x}{\alpha + 1} \right)^j = \sum_{k=0}^j \binom{j}{k} \left( \frac{\theta}{\alpha + 1} \right)^k x^k,$$

we can write (2.5) and (2.6) as

$$F(x; \phi) = \sum_{j=0}^{\infty} \sum_{k=0}^j \binom{j}{k} \left( \frac{\theta}{\alpha + 1} \right)^k p^j e^{-\theta j x} x^k \left[ 1 - e^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right) \right],$$

and

$$\begin{aligned} f(x; \phi) &= (1-p) \sum_{j=0}^{\infty} (j+1) p^j (\alpha x^k + \theta x^{k+1}) \times \\ &\quad \sum_{k=0}^j \binom{j}{k} \left( \frac{\theta}{\alpha + 1} \right)^k e^{-\theta(j+1)x}, \end{aligned}$$

respectively. Note that the expressions involving  $X$  on the right-hand sides of Equations (2.8) and (2.9) are mathematically more tractable. So, Equation (2.9) can be used, for example, to derive some mathematical properties of the QLG distribution.

The hazard function (also known as the failure rate, hazard rate, or force of mortality), reverse hazard functions and survival function are given by

$$h(x) = \frac{f(x)}{1-F(x)} = \frac{\theta(\alpha + \theta x)}{1 + \alpha + \theta x} \left[ 1 - pe^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right) \right]^{-1},$$

$$\tau(x) = \frac{f(x)}{F(x)} = \frac{\theta(1-p)(\alpha + \theta x)e^{-\theta x}}{(1+\alpha) \left[ 1 - e^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right) \right] \left[ 1 - pe^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right) \right]},$$

and

$$R(x) = \frac{(1-p)e^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right)}{1 - pe^{-\theta x} \left( 1 + \frac{\theta x}{\alpha + 1} \right)},$$

respectively. Figure 3, 4 illustrate the graphical behavior of the hazard rate and survival function of QLG distribution for selected values of the parameters.

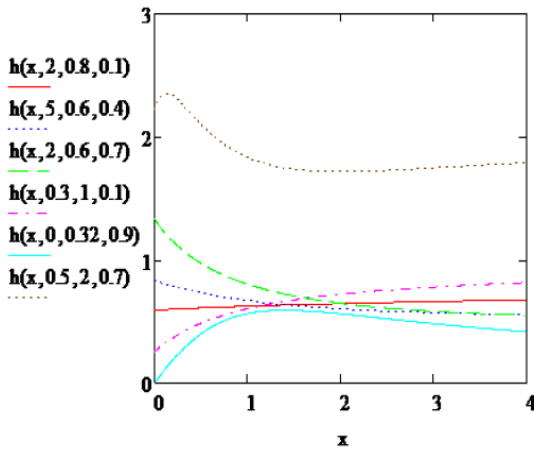


Figure 3: the hazard rate function of Quasi Lindley geometric distribution

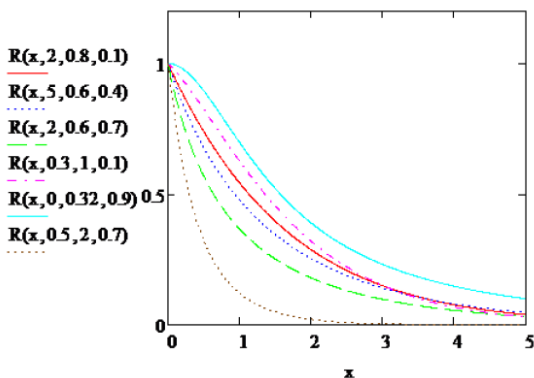


Figure 4: the hazard rate and survival function of Quasi Lindley geometric distribution

### 3. STATISTICAL PROPERTIES

In this section we discuss the statistical properties of the Quasi Lindley geometric distribution, in particular, quantile function, moment and moment generating function.

#### 3.1 Quantile function

The quantile  $X_q$  of the  $QLG(\alpha, \theta, p)$  is the real solution of the following equation:

$$\theta x_q - \log\left(1 + \frac{\theta x_q}{1 + \alpha}\right) + \log\left(\frac{1-p}{1-pq}\right) = 0$$

we present the shortcomings of the classical kurtosis measure are well-known. There are many heavy tailed distributions for which this measure is infinite. So, it becomes uninformative precisely when it needs to be. The Bowley's skewness (see Kenney and Keeping, 1962) is based on quartiles

$$S_K = \frac{Q_{0.75} - 2Q_{0.5} + Q_{0.25}}{Q_{0.75} - Q_{0.25}},$$

And the Moors' kurtosis (1998) is based on octiles

$$K_u = \frac{Q_{0.875} - Q_{0.625} - Q_{0.375} + Q_{0.125}}{Q_{0.75} - Q_{0.25}}.$$

Where  $Q(\cdot)$  represents the quantile function.

#### 3.2 Moments and associated measures

Moments are necessary and important in any statistical analysis, especially in applications. It can be used to study the most important features and characteristics of a distribution (e.g., tendency, dispersion, skewness and kurtosis).

##### Theorem (3.1).

If  $X$  has  $QLG(x; \phi)$ ,  $\phi = (\alpha, \theta, p)$  then the  $r_{th}$  raw moment of  $X$  is given by the following

$$\mu_r(x) = (1-p) \sum_{j=0}^{\infty} \sum_{k=0}^j \binom{j}{k} \left( \frac{\theta}{\alpha + 1} \right)^k (j+1)p^j \times \left\{ \frac{\Gamma(r+k+1)}{[\theta(j+1)]^{r+k+1}} \left( \alpha + \frac{\theta(r+k+1)}{\theta(j+1)} \right) \right\}.$$

##### Proof:

Let  $X$  be a random variable with density function (2.9). The  $r_{th}$  ordinary moment of the  $QLG$  distribution is given by

$$\mu_r(x) = E(X^r) = \int_0^{\infty} x^r f(x, \phi) dx$$

$$\begin{aligned} &= (1-p) \sum_{j=0}^{\infty} \sum_{k=0}^j \binom{j}{k} \left( \frac{\theta}{\alpha + 1} \right)^k (j+1)p^j \int_0^{\infty} x^{r+k} (\alpha + \theta x) e^{-\theta(j+1)x} dx \\ &= w_{j,k} \left\{ \alpha \int_0^{\infty} x^{r+k} e^{-\theta(j+1)x} dx + \theta \int_0^{\infty} x^{r+k+1} e^{-\theta(j+1)x} dx \right\} \\ &= w_{j,k} \left\{ \frac{\Gamma(r+k+1)}{[\theta(j+1)]^{r+k+1}} \left( \alpha + \frac{\theta(r+k+1)}{\theta(j+1)} \right) \right\} \end{aligned}$$

where

$$w_{j,k} = (1-p) \sum_{j=0}^{\infty} \sum_{k=0}^j \binom{j}{k} \left( \frac{\theta}{\alpha + 1} \right)^k (j+1)p^j.$$

Which completes the proof.

Based on the first four moments of the  $QLG$  distribution, the measures of skewness  $A(\Phi)$  and kurtosis  $K(\Phi)$  of the  $QLG$  distribution can be obtained as

$$A(\Phi) = \frac{\mu_3(\theta) - 3\mu_1(\theta)\mu_2(\theta) + 2\mu_1^3(\theta)}{[\mu_2(\theta) - \mu_1^2(\theta)]^{\frac{3}{2}}},$$

and

$$k(\Phi) = \frac{\mu_4(\theta) - 4\mu_1(\theta)\mu_3(\theta) + 6\mu_1^2(\theta)\mu_2(\theta) - 3\mu_1^4(\theta)}{[\mu_2(\theta) - \mu_1^2(\theta)]^2}$$

**Theorem (3.2):**

If  $X$  has the  $QLG(x;\phi)$ , then the moment generating function (mgf) of  $X$  is given as follows

$$M_X(t) = w_{j,k} \left\{ \frac{\Gamma(k+1)}{[\theta(j+1)-t]^{k+1}} \left( \alpha + \frac{\theta(k+1)}{\theta(j+1)-t} \right) \right\}$$

**Proof:**

Starting with

$$M_X(t) = \int_0^\infty e^{tx} f_{QLG}(\phi, x) dx$$

$$= w_{j,k} \left\{ \alpha \int_0^\infty x^k e^{-(\theta(j+1)-t)x} dx + \theta \int_0^\infty x^{k+1} e^{-(\theta(j+1)-t)x} dx \right\}$$

$$w_{j,k} \left\{ \frac{\Gamma(k+1)}{[\theta(j+1)-t]^{k+1}} \left( \alpha + \frac{\theta(k+1)}{\theta(j+1)-t} \right) \right\}$$

which completes the proof.

**4. ORDER STATISTICS AND MOMENT OF ORDER STATISTICS**

Order statistics are among the most fundamental tools in non-parametric statistics and inference. They enter in the problems of estimation and hypothesis tests in a variety of ways. Let  $X_1, X_2, \dots, X_n$  be a simple random sample from QLG distribution with cdf and pdf given by (2.2) and (2.3), respectively. Let  $X_{(1:n)} \leq X_{(2:n)} \leq \dots \leq X_{(n:n)}$  denote the order statistics obtained from this sample. The probability density function of  $i$ th order statistics, say  $X_{i:n}$ , and the moments of  $X_{i:n}; i = 1, 2, \dots, n$ . Therefore, the measures of skewness and kurtosis of the distribution of the  $X_{i:n}$  are presented. The pdf of  $X_{i:n}$  is given by

$$f_{i:n}(x, \phi) = \frac{f(x)}{B(i, n-i+1)} [F(x, \phi)]^{i-1} [1-F(x, \phi)]^{n-i}$$

Where  $F(x, \phi)$  and  $f(x, \phi)$  are the cdf and pdf of the QLG distribution given by (2.2), (2.3), respectively, and  $B(.,.)$  is the beta function, since  $0 < F(x, \phi) < 1$ , for  $X > 0$ , by using the binomial series expansion of

$$[1 - F(x, \phi)]^{n-i} \text{ given by}$$

$$[1 - F(x, \phi)]^{n-i} = \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} [F(x, \phi)]^j$$

We have

$$f_{i:n}(x, \phi) = \frac{f(x, \phi)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} [F(x, \phi)]^{i+j-1}$$

Substituting from (2.2) and (2.3) into (4.3), we get

$$f_{i:n}(x, \phi) = \frac{\sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \left[ \frac{1 - e^{-\theta x} \left(1 + \frac{\theta x}{\alpha+1}\right)}{1 - p e^{-\theta x} \left(1 + \frac{\theta x}{\alpha+1}\right)} \right]^{i+j-1}}{B(i, n-i+1)} \times (1-p) \frac{\theta}{\alpha+1} (\alpha + \theta x) e^{-\theta x} \left[ 1 - p e^{-\theta x} \left(1 + \frac{\theta x}{\alpha+1}\right) \right]^{-2}$$

we can express the  $k$ th ordinary moment of the  $i$ th

order statistics  $X_{i:n}$  say  $E(X_{i:n}^k)$  as a linear combination of the  $k$ th moments of the QLG distribution with different shape parameters. Therefore, the measures of skewness and kurtosis of the distribution of  $X_{i:n}$  can be calculated.

**5. RESIDUAL LIFE AND REVERSED FAILURE RATE FUNCTION**

Given that a component survives up to time  $t \geq 0$ , the residual life is the period beyond  $t$  until the time of failure and defined by the conditional random variable  $X - t | X > t$ . In reliability, it is well known that the mean residual life function and ratio of two consecutive moments of residual life determine the distribution uniquely (Gupta and Gupta, 1983).

Therefore, we obtain the  $r$ th-order moment of the residual life via the general formula

$$\mu_r(t) = E((X - t)^r | X > t)$$

$$= \frac{1}{F(t)} \int_t^\infty (x - t)^r f(x, \phi) dx, r \geq 1.$$

Applying the binomial expansion of  $(X - t)^r$  and substituting  $f(x, \phi)$  given by (2.9) into the above formula gives

$$\mu_r(t) = \frac{(1-p)}{F(t)} \sum_{i=0}^r \sum_{j=0}^i (-t)^j (j+1) p^j \binom{j}{k} \left( \frac{\theta}{\alpha+1} \right)^k$$

$$\times \int_t^\infty (\alpha x^{r+k-i} + \theta x^{r+k-i+1}) e^{-\theta(j+1)x} dx$$

$$= \frac{(1-p)}{F(t)} \sum_{i=0}^r \sum_{j=0}^i (-t)^j (j+1) p^j \binom{j}{k} \left( \frac{\theta}{\alpha+1} \right)^k$$

$$\times \frac{1}{[\theta(j+1)]^{r+k-i+2}} \{ \alpha \theta (j+1) \Gamma(r+k-i+1, \theta t(j+1))$$

$$+ \theta \Gamma(r+k-i+2, \theta t(j+1)) \}$$

where  $\Gamma(s, t) = \int_t^\infty x^{s-1} e^{-x} dx$  is the upper incomplete gamma function.

Also the mean residual life of the QLG distribution is given by

$$\mu(t) = E((X - t) | X > t) = \frac{1}{F(t)} \int_t^\infty x f(x, \phi) dx - t$$

$$= \frac{(1-p)}{F(t)} \sum_{j=0}^\infty \sum_{k=0}^j (j+1) p^j \binom{j}{k} \left( \frac{\theta}{\alpha+1} \right)^k$$

$$\times \frac{1}{[\theta(j+1)]^{k+3}} \{ \alpha \theta (j+1) \Gamma(k+2, \theta t(j+1))$$

$$+ \theta \Gamma(k+3, \theta t(j+1)) \} - t.$$

On the other hand, we analogously discuss the reversed residual life and some of its properties. The reversed residual life can be defined as the conditional random variable  $t - X | X \leq t$  which denotes the time elapsed from the failure of a component given that its life is less than or equal to  $t$ . This random variable may also be called the inactivity time (or time since failure); for more details you may see (Kundu

and Nanda, 2010; Nanda, Singh, Misra, and Paul, 2003). Also, in reliability, the mean reversed residual life and ratio of two consecutive moments of reversed residual life characterize the distribution uniquely. The reversed failure (or reversed hazard)

rate function is given by Equation (2.5). The  $r$ th-order moment of the reversed residual life can be obtained by the well-known formula

$$m_r(t) = E((t - X)^r | X \leq t) \\ = \frac{1}{F(t)} \int_0^t (t - x)^r f(x, \phi) dx, r \geq 1$$

Applying the binomial expansion of  $(t - X)^r$  and substituting  $f(x, \phi)$  given by (2.9) into the above formula gives

$$m_r(t) = \frac{(1-p)}{F(t)} \sum_{i=0}^r \sum_{j=0}^{\infty} (-1)^{r+i} t^i (j+1) p^j \binom{j}{k} \left(\frac{\theta}{\alpha+1}\right)^k \\ \times \frac{1}{[\theta(j+1)]^{k+i+2}} \{ \alpha\theta(j+1)\gamma(r+k-i+1, \theta t(j+1)) \\ + \theta\gamma(r+k-i+2, \theta t(j+1)) \}$$

where  $\gamma(s, t) = \int_0^t x^{s-1} e^{-x} dx$  is the lower incomplete gamma function. Thus the mean of the reversed residual life of the QLG distribution is given by

$$m_1(t) = m(t) = t - \frac{(1-p)}{F(t)} \sum_{j=0}^{\infty} \sum_{k=0}^j (j+1) p^j \binom{j}{k} \left(\frac{\theta}{\alpha+1}\right)^k \\ \times \frac{1}{[\theta(j+1)]^{k+3}} \{ \alpha\theta(j+1)\gamma(k+2, \theta t(j+1)) \\ + \theta\gamma(k+3, \theta t(j+1)) \}.$$

Using  $m(t)$  and  $m_2(t)$  we obtain the variance of the reversed residual life of the QLG distribution, and hence the coefficient of variation of the reversed residual life of the QLG distribution can be easily obtained.

## 6. MEAN DEVIATIONS

The amount of scatter in a population can be measured by the totality of deviations from the mean and median. For a random variable  $X$  with pdf,  $f(x)$ , distribution function  $F(x)$ , mean  $\mu = E(X)$  and  $M = Median(X)$ , the mean deviation about the mean and the mean deviation about the median, respectively, are defined by

$$\eta_1(x) = \int_0^{\infty} |x - \mu| f(x) dx = 2\mu F(\mu) - 2I(\mu),$$

and

$$\eta_2(x) = \int_0^{\infty} |x - M| f(x) dx = \mu - 2I(M),$$

where

$$I(q) = \int_0^q x f(x) dx$$

$$= \frac{(1-p)}{F(t)} \sum_{j=0}^{\infty} \sum_{k=0}^j (j+1) p^j \binom{j}{k} \left(\frac{\theta}{\alpha+1}\right)^k \\ \times \frac{1}{[\theta(j+1)]^{k+3}} \{ \alpha\theta(j+1)\gamma(k+2, \theta q(j+1)) \\ + \theta\gamma(k+3, \theta q(j+1)) \},$$

Thus the mean deviation about the mean and the mean deviation about the median are given by

$$\eta_1(x) = 2\mu \left[ \frac{1 - e^{-\theta\mu} \left(1 + \frac{\theta\mu}{\alpha+1}\right)}{1 - p e^{-\theta\mu} \left(1 + \frac{\theta\mu}{\alpha+1}\right)} \right] - 2I(\mu)$$

and

$$\eta_2(x) = \mu - 2I(M)$$

respectively, where  $\mu$  is given by Equation (3.1) at  $r = 1$ ,  $I(\mu)$  and  $I(M)$  are obtained by substituting  $\mu$  and  $M$  in Equation (6.3).

## 7. LEAST SQUARES AND WEIGHTED LEAST SQUARES ESTIMATORS

In this section we provide the regression based method estimators of the unknown parameters of the modified Weibull geometric distribution which was originally suggested by Swain, Venkatraman and Wilson (1988) to estimate the parameters of beta distributions. It can be used some other cases also. Suppose  $Y_1, \dots, Y_n$  is a random sample of size  $n$  from a distribution function  $G(\cdot)$  and suppose  $Y_{(j)}$ ;  $j = 1, 2, \dots, r$  denotes the ordered sample. The proposed method uses the distribution of  $(Y_{(j)})$ . For a sample of size  $n$ , we have

$$E(G(Y_{(j)})) = \frac{j}{n+1}, V(G(Y_{(j)})) = \frac{j(n-j+1)}{(n+1)^2(n+2)}$$

$$\text{and } Cov(G(Y_{(j)}), G(Y_{(k)})) = \frac{j(n-k+1)}{(n+1)^2(n+2)}; \text{ for } j < k,$$

see Johnson, Kotz and Balakrishnan (1995). Using the expectations and the variances, two variants of the least squares methods can be used.

**Method 1 (Least Squares Estimators)**. Obtain the estimators by minimizing

$$\sum_{j=1}^n \left( G(Y_{(j)}) - \frac{j}{n+1} \right)^2,$$

with respect to the unknown parameters. Therefore in case of QLG distribution the least squares estimators of  $\alpha, \theta$  and  $p$ , say  $\widehat{\alpha}_{LSE}$ ,  $\widehat{\theta}_{LSE}$  and  $\widehat{p}_{LSE}$  respectively, can be obtained by minimizing

$$\sum_{j=1}^n \left\{ \frac{1 - e^{-\theta x_j} \left(1 + \frac{\theta x_j}{\alpha+1}\right)}{1 - p e^{-\theta x_j} \left(1 + \frac{\theta x_j}{\alpha+1}\right)} - \frac{j}{n+1} \right\}^2$$

with respect to  $\alpha, \theta$  and  $p$ .

**Method 2 (Weighted Least Squares Estimators).** The weighted least squares estimators can be obtained by minimizing

$$\sum_{j=1}^n w_j \left( G(Y_{(j)}) - \frac{j}{n+1} \right)^2,$$

with respect to the unknown parameters, where

$$w_j = \frac{1}{V(G(Y_{(j)}))} = \frac{(n+1)^2(n+2)}{j(n-j+1)}.$$

Therefore, in case of QLG distribution the weighted least squares estimators of  $\alpha, \theta$  and  $p$ , say  $\widehat{\alpha}_{WLSSE}, \widehat{\theta}_{WLSSE}$  and  $\widehat{p}_{WLSSE}$  respectively, can be obtained by minimizing

$$\sum_{j=1}^n w_j \left\{ \frac{1 - e^{-\theta x_j} \left( 1 + \frac{\theta x_j}{\alpha+1} \right)}{1 - p e^{-\theta x_j} \left( 1 + \frac{\theta x_j}{\alpha+1} \right)} - \frac{j}{n+1} \right\}^2$$

with respect to the unknown parameters only.

## 8. MAXIMUM LIKELIHOOD ESTIMATORS

In this Section we consider the maximum likelihood estimators (MLE's) of QLG  $(\varphi, X)$ . Let  $X_1, \dots, X_n$  be a random sample of size  $n$  from QLG  $(\varphi, X)$ , let  $\varphi = (\alpha, \theta, p)^T$  be the parameter vector. The log likelihood function can be written as

$$L(\varphi) = n \log(1-p) + n \log \theta - n \log(1+\alpha) + \sum_{i=1}^n \log(\alpha + \theta x_i)$$

$$- \theta \sum_{i=1}^n x_i - 2 \sum_{i=1}^n \log \left[ 1 - p e^{-\theta x_i} \left( 1 + \frac{\theta x_i}{\alpha+1} \right) \right]$$

The associated score function is given by

$$U_n(\varphi) = \left[ \frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \theta}, \frac{\partial L}{\partial p} \right]^T$$

where

$$\frac{\partial L}{\partial \alpha} = \frac{-n}{(1+\alpha)} + \sum_{i=1}^n \frac{1}{(\alpha + \theta x_i)} - 2 \sum_{i=1}^n \frac{p \theta x_i e^{-\theta x_i}}{[1+\alpha]^2 \left[ 1 - p e^{-\theta x_i} \left( 1 + \frac{\theta x_i}{\alpha+1} \right) \right]},$$

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

$$- 2p \sum_{i=1}^n \frac{x_i e^{-\theta x_i} \left[ \left( 1 + \frac{\theta x_i}{\alpha+1} \right) - \frac{1}{(1+\theta)^2} \right]}{\left[ 1 - p e^{-\theta x_i} \left( 1 + \frac{\theta x_i}{\alpha+1} \right) \right]},$$

and

$$\frac{\partial L}{\partial p} = \frac{-n}{1-p} + 2 \sum_{i=1}^n \frac{e^{-\theta x_i} \left( 1 + \frac{\theta x_i}{\alpha+1} \right)}{\left[ 1 - p e^{-\theta x_i} \left( 1 + \frac{\theta x_i}{\alpha+1} \right) \right]}.$$

The maximum likelihood estimation (MLE) of  $\varphi$ , say  $\widehat{\varphi}$ , is obtained by solving the nonlinear system  $U_n(\varphi) = 0$ . These equations cannot be solved analytically, and statistical software can be used to solve them numerically via iterative methods. We can use iterative techniques such as a Newton–Raphson type algorithm to obtain the estimate  $\widehat{\varphi}$ . The Broyden–Fletcher–Goldfarb–Shanno method (see e.g. [1999, 2007]) with analytical derivatives has been used for maximizing the log-likelihood function  $L(\varphi)$ . For interval estimation and hypothesis tests on the model parameters, we

require the information matrix. The  $3 \times 3$  observed information matrix is given by

$$I_n(\varphi) = - \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\theta} & I_{\alpha p} \\ I_{\theta\alpha} & I_{\theta\theta} & I_{\theta p} \\ I_{p\alpha} & I_{p\theta} & I_{pp} \end{bmatrix}$$

whose elements are given in Appendix. Applying the usual large sample approximation, MLE of  $\varphi$ , i.e.  $\widehat{\varphi}$  can be treated as being approximately  $N_3(\varphi, J_n(\varphi)^{-1})$ , where  $J_n(\varphi) = E[I_n(\varphi)]$ . Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of  $\sqrt{n}(\widehat{\varphi} - \varphi)$  is  $N_3(0, J(\varphi)^{-1})$ , where  $J(\varphi) = \lim_{n \rightarrow \infty} n^{-1} I_n(\varphi)$  is the unit information matrix. This asymptotic behavior remains valid if  $J(\varphi)$  is replaced by the average sample information matrix evaluated at  $\widehat{\varphi}$ , say  $n^{-1} I_n(\widehat{\varphi})$ . The estimated asymptotic multivariate normal  $N_3(\varphi, I_n(\widehat{\varphi})^{-1})$  distribution of  $\widehat{\varphi}$  can be used to construct approximate confidence intervals for the parameters and for the hazard rate and survival functions. An  $100(1 - \gamma)$  asymptotic confidence interval for each parameter  $\varphi_r$  is given by

$$ACI_r = \left( \widehat{\varphi}_r - z_{\frac{\gamma}{2}} \sqrt{I_{rr}}, \widehat{\varphi}_r + z_{\frac{\gamma}{2}} \sqrt{I_{rr}} \right)$$

where  $\widehat{I}_{rr}$  is the  $(r, r)$  diagonal element of  $I_n(\widehat{\varphi})^{-1}$  for  $r = 1, 2, 3$ , and  $Z_{\frac{\gamma}{2}}$  is the quantile  $1 - \frac{\gamma}{2}$  of the standard normal distribution.

The following table represents the mean square error (MSEs) of the MLEs.

**Table 1 The mean square errors of the MLEs**

QLG( $\alpha, \theta, P$ )	SampleSize(n)	MSE( $\alpha$ )	MSE( $\widehat{\theta}$ )	MSE( $P$ )
	15	0.4206	0.2343	0.0726
	25	0.3008	0.2168	0.0269
	35	0.1800	0.0961	0.0195
QLG(0.5,0.25,0.1)	45	0.1190	0.0460	0.0138
	55	0.1128	0.0041	0.0122
	65	0.1086	0.0031	0.0101
	75	0.0176	0.0027	0.0054
	15	0.4825	0.0141	0.201
	25	0.4632	0.0130	0.1668
	35	0.2701	0.0114	0.0712
QLG(1.5,0.75,0.3)	45	0.2673	0.0059	0.0624
	55	0.2239	0.0042	0.0678
	65	0.2043	0.0012	0.0048
	75	0.0896	0.0011	0.0022
	15	0.0129	0.0118	0.1013
	25	0.0006	0.0748	0.0454
	35	0.0005	0.0612	0.022
QLG(0.05,1.0,0.5)	45	0.0004	0.0509	0.0148
	55	0.0003	0.0386	0.0123
	65	0.0002	0.0116	0.0104
	75	0.0001	0.0015	0.0018

We noticed from the above Table 1 that all MSEs decrease as the sample size increases, while they increase with increasing of the true parameter.

### 9. APPLICATIONS

In this Section, we use two real data set to show how the QLG distribution can be applied in practice and the Quasi Lindley geometric distribution can be a better model than one based on the Lindley geometric and Lindley distribution. For each data set, we estimate the unknown parameters of each distribution by the maximum-likelihood method, and with these obtained estimates, we obtain the values of the Kolmogorov-Smirnov (K-S) statistic (the distance between the empirical CDFs and the fitted CDFs), Akaike information criterion (AIC), Bayesian information criterion (BIC), and the consistent Akaike information criterion (CAIC) are used to compare the candidate distributions. The AIC and BIC values are given by  $-2 \log \hat{L} + 2k$  and  $-2 \log \hat{L} + k \log n$ , respectively, where  $\hat{L}$  is the value of the log likelihood function for obtained estimates of the unknown parameters,  $k$  is the number of the estimated parameters and  $n$  is the sample size. The better distribution corresponds to smaller  $KS, -2 \log \hat{L}, AIC$  and  $CAIC$  values.

The first data set represents remission times (in months) of a random sample of 128 bladder cancer patients reported in Lee and Wang (2003). The data are as follows:  
0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

**Table 2. Maximum-likelihood estimates, AIC, BIC and CAIC values, and Kolmogorov--Smirnov statistics for the data.**

The model	MLEs				Measurements			
	$\alpha$	$\hat{\theta}$	$p$	$KS$	$-2\log L$	$AIC$	$BIC$	$CAIC$
QLG	0.922	0.033	0.035	0.583	480.618	486.618	492.355	487.14
LG		0.043	0.005	0.606	829.978	833.978	837.802	834.234
L		0.043		0.624	502.861	504.861	506.773	504.944

From table2, we observe that the QLG distribution is a competitive distribution compared with other distributions. In fact, based on the values of the AIC and BIC criteria as well as the value of the KS -statistic, we observe that the QLG distribution provides the best fit for these data among all the models considered.

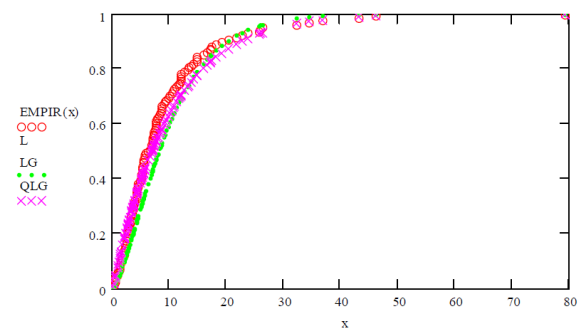
The second data set have been obtained from Aarset (1987) and it is provided below. It represents the lifetimes of 50 devices.  
0.1, 0.2, 1, 1, 1, 1, 1, 2, 3, 6, 7, 1, 1, 12, 18, 18, 18, 18, 18, 21, 32, 36, 40, 45, 46, 47, 50, 55, 60, 63, 63, 67, 67, 67, 67, 72, 75, 79, 82, 82, 83, 84, 84, 84, 85, 85, 85, 85, 85, 86, 86.

**Table 3. Maximum-likelihood estimates, AIC, BIC and CAIC values, and Kolmogorov--Smirnov statistics for the data**

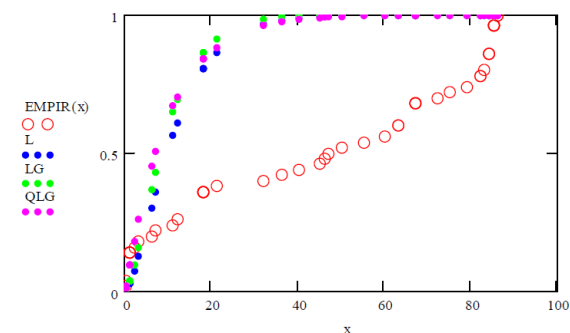
The model	MLEs				Measurements			
	$\alpha$	$\hat{\theta}$	$p$	$KS$	$-2\log L$	$AIC$	$BIC$	$CAIC$
QLG	6.639	0.12	0.009	0.082	829.008	835.008	843.564	835.201
LG		0.192	0.026	0.123	1350.098	2866.984	2872.689	5867.08
L		0.196		0.117	839.06	841.06	843.912	841.092

Again, the values in Table 3 indicate that the QLG distribution is a strong competitor to other distributions commonly used in literature for fitting lifetime data. From the above results, it is evident that the QLG distribution is the best distribution for fitting these data sets compared to other distributions considered here.

A cdf plot compares the fitted cdf of the models with the empirical curve of the observed data (Fig. 5) and (Fig.6) The fitted cdf for the QLG model is closer to the empirical graph than the fits of the LG and L models.



**Figure5: Empirical, fitted QLG, LG and L cdf of the bladder cancer patients data**



**Figure6: Empirical, fitted QLG, LG and L cdf of the lifetimes of 50 devices**

### 10. REFERENCES

- [1] Adamidis K., Dimitrakopoulou,T., Loukas,S. (2005). On ageneralization of the exponential- geometric distribution. Statistics & Probability Letters, 73,259-269.
- [2] Adamidis K.,Loukas,S.(1998).A lifetime distribution with decreasing failure rate. Statistics & Probability Letters, 39, 35-42.
- [3] Aarset M. V. (1987). How to identify bathtub hazard rate, IEEEET ransactions Reliability, 36,106-108.
- [4] Barreto-Souza,W, Silva,R. Band Cordeiro,G.M (2010) A new distribution with decreasing, increasing and upside-

- down bathtub failure rate, *Comput. Statist. Data Anal.* 54pp.935-944.
- [5] Barreto-Souza, W., Cribari-Neto, F. (2009). A generalization of the exponential- Poisson distribution. *Statistics & Probability Letters*, 79, 2493-2500.
- [6] Chahkandi, M., Ganjali, M. (2009). On some life time distributions with decreasing failure rate. *Computational Statistics & Data Analysis*, 53, 4433-4440.
- [7] Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78, 493-506.
- [8] Gupta, P.L. and Gupta, R.C. (1983). On the moments of residual life in reliability and some characterization results. *Comm. Statist. Theory Methods* 12(4), 449-461.
- [9] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*. volume 1. John Wiley & Sons, New York.
- [10] Kenney, J. F. and Keeping, E.S. (1962). *Mathematics of statistics*, pp. 101-102, Part 1, 3rd ed. Princeton, NJ.
- [11] Kundu, C. and Nanda, A.K. (2010). Some reliability properties of the inactivity time. *Communications in Statistics-Theory and Methods* 39(899-911).
- [12] Kus, C. (2007). A new lifetime distribution. *Computational Statistics & Data Analysis*, 51, 4497-4509.
- [13] Lee, E.T. and Wang, J.W. *Statistical Methods for Survival Data Analysis*, 3rd ed., Wiley, New York, 2003.
- [14] Lindley, D.V. (1958). Fiducial distributions and Bayes theorem, *J. Royal Stat. Soc. Series B*, 20(102-107).
- [15] Lu, W., Shi, D. (2011). A new compounding life distribution: the Weibull- Poisson distribution. *Journal of Applied Statistics*, DOI:10.1080/02664763.2011.575126.
- [16] Mazucheli, J. and Achcar, J.A. (2011). The Lindley distribution applied to competing risks lifetime data. *Comput. Methods Programs Biomed.* 104 (2):188-192.
- [17] Morais, A.L., Barreto-Souza, W. (2011). A Compound Class of Weibull and Power Series Distributions. *Computational Statistics & Data Analysis*, 55, 1410-1425.
- [18] Moors, J. J. A. (1998). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society Ser. D., The Statistician*, 37, 25-32.
- [19] Nanda, A.K., Singh, H. Misra, N. Paul, P. (2003). Reliability properties of reversed residual lifetime. *Communications in Statistics- Theory and Methods* 32(2031-2042).
- [20] Rama, S. and Mishra, A. (2013). A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research*. 6(4), 64-71.
- [21] Swain, J. S., Venkatraman, S., and Wilson, J. (1988). Least squares estimation of distribution function in Johnson translation system. *Journal of Statistical Computation and Simulation*, 29, 271-297.
- [22] Tahmasbi, R., Rezaei, S. (2008). A two-parameter lifetime distribution with decreasing failure rate. *Computational Statistics & Data Analysis*, 52, 3889-3901.
- [23] Zakerzadeh, H. and Mahmoudi, E. (2012). A new two parameter lifetime distribution: model and properties. arXiv:1204.4248[stat.CO].