Harris Generalized Linear Exponential Distribution and Its Applications

Albin Paul[a], K.K. Jose[b]

•

[a]St.Thomas College Palai, Kerala, India
albinpaul0333@gmail.com
[b]School of Mathematics and Statistics,
Mahatma Gandhi University,
Kottayam, Kerala, India.
kkj.smsda.mgu@gmail.com

Abstract

In this paper we consider an extension of the linear exponential distribution based on the Harris generalization method, which includes some of the life time distribution as sub models. Along with the four parameter generalization namely Harris generalized linear exponential distribution, we derive some of its properties such as moments, quantiles and moment generating function. A compound form expression of the density function is given. For the given model the R´enyi entropy, mean residual life and the distribution of order statistics is derived. The problem of estimation of parameters is considered and validated with respect to two real data sets.

Keywords: Harris generalization, Linear exponential distribution, Order Statistics, Moment generating function, Entropy, Mean residual life, Data analysis

1 Introduction

Introduction of new parameters to the well established classical distributions may result in more flexible new families of distributions. Several extended forms of distributions have been studied by many researchers like Azzalini (1985), Marshall and olkin (1997), Ferreira and Steel (2007), Jose et al (2010), Krishna et al (2013), Jose and Sivdas (2015) etc. Recently Aly and Benkherouf (2011) introduced a method for developing new classes of distributions by adding two new parameters to an existing distribution, which includes the baseline distribution as a special case and gives more flexible models for various types of data. This method is based on probability generating function introduced by Harris (1948). Hence the resulting family of distributions is known as Harris Generalized (HG) family of distributions. This family of distributions can be considered as a generalization of Marshall Olkin family of distributions introduced by Marshall and olkin (1997). Some properties and applications of HG family of distributions are studied by Aly and Benkherouf (2011), Batsidis and lemonte (2014) and Cordeiro et al (2015) etc. Aly and Benkherouf (2011) derived the general structure of HG family of distributions as follows.

The survival function of HG family of distributions is given by

$$\bar{G}(x;\ \theta, r) = \left(\frac{\theta(\bar{F}(x))^r}{1-\bar{\theta}(\bar{F}(x))^r}\right)^{\frac{1}{r}},\ 0 < \theta < \infty,\ r > 0,\ \bar{\theta} = 1 - \theta \tag{1.1}$$

where $\bar{F}(.)$ is the survival function of the baseline distribution. The corresponding density function is

$$g(x) = \frac{\theta^{\frac{1}{r}} f(x)}{\left(1-\bar{\theta}(\bar{F}(x))^r\right)^{\left(\frac{r+1}{r}\right)}} \quad , \quad x > 0 \tag{1.2}$$

If r=1 in equations (1.1) and (1.2), reduces to the corresponding survival function and density function of Marshall-Olkin family of distributions. The density function of HG family of distributions can be expressed as a linear combination of exponentiated family of distributions as given in Batsidis and lemonte (2014) and Barreto-Souza et al (2013) as follows.

For $\theta \in (0, 1)$

$$g(x) = f(x) \sum_{i=0}^{\infty} w_i \left(\bar{F}(x)\right)^{r_i} \tag{1.3}$$

where $w_i = w_i(\theta, r) = \theta^{\frac{1}{r}} \bar{\theta}^i \frac{\Gamma(r^{-1}+i+1)}{\Gamma(r^{-1}+1)i!}$

For $\theta > 1$

$$g(x) = f(x) \sum_{i=0}^{\infty} v_i \left(\bar{F}(x)\right)^{r_i} \tag{1.4}$$

where $v_i = v_i(\theta, r) = (-1)^i \theta^{-1} \sum_{j=1}^{\infty} \binom{j}{i} \left(\frac{\theta-1}{\theta}\right)^j \frac{\Gamma(r^{-1}+j+1)}{\Gamma(r^{-1}+1)j!}$

From (1.3) and (1.4) it is clear that the HG family of distributions can be expressed as the baseline distribution $f(x)$, multiplied by an infinite power series which differ only for the coefficients.

The aim of this paper is to study a new univariate family of distribution based on the Harris generalization method. The contents are organized as follows. In Section 2 we discuss the Linear exponential distribution. In Section 3 we introduce the Harris generalized linear exponential distribution. In Section 4 we represent the Harris generalized linear exponential distribution as a compound distribution with exponential density. In Section 5 and 6 we evaluate the Entropy and Mean residual life. Section 7 gives the distribution of order statistics. Section 8 and 9 discuss the maximum likelihood estimation of the parameters and an application to a real data set.

2 Linear Exponential Distribution

The Linear Exponential (LE) distribution is an important distribution that has rich variety of applications for modeling life time data. The LE distribution is also known as Linear Failure rate distribution. The LE distribution contains Rayleigh and Exponential distribution as special cases and they are well known in the literature for a variety of applications.

From Lai et al (2006) and Zang et al (2005), the LE distribution models phenomenon with increasing failure rate, the accuracy of the statistical procedures depends on the probability model or distributions which are considered for the analysis. Recently there has been a renewed interest in the study of extended versions of conventional classical distributions. Since the real data is affected by various factors, the statistical models derived using the extended form of distribution shows more significant applications in reliability, medical science, finance, economics etc. The LE distribution has many applications in applied statistics, reliability analysis, medical studies (See Carbone et al (1967), Broadbent (1958)). Recently many studies have been done on LE distribution and its generalizations by introducing additional parameters (See Mahmoud and Alam (2010), Cordeiro et al (2015), Nadarajah et al (2014) etc).

The LE distribution with the parameters β_1 and β_2, (LE(β_1, β_2)) has the following cumulative distribution function

$$F(x;\ \beta_1, \beta_2) = 1 - \exp\left(-\beta_1 x - \frac{\beta_2}{2} x^2\right) \tag{2.1}$$

If we put β_2=0 in (2.1) we can obtain the exponential distribution with parameter β_1 and if we put β_1=0 in (2.1) we get the Rayleigh distribution with parameter β_2.

The probability density function (pdf) of LE(β_1, β_2) distribution is given by

$$f(x;\beta_1,\beta_2) = (\beta_1 + \beta_2 x)\exp\left(-\beta_1 x - \frac{\beta_2}{2}x^2\right), \ \ x > 0, \quad \beta_1 \ ,\beta_2 > 0.$$

3 Harris Generalized Linear Exponential Distribution

By applying the method given in Aly and Benkherouf (2011) Harris Extended Linear exponential distribution is introduced by Batsidis and lemonte (2014). Here we call this family of distributions as Harris Generalized distribution. By taking LE distribution as the baseline distribution, we get the Harris generalized Linear exponential (HGLE) distribution. The application of this family of distribution in the context of reliability test plan were studied by JoseandPaul(2018).

The substitution of the survival function of LE(β_1, β_2) distribution in (1.1) gives the HGLE distribution with parameters θ, r, β_1, β_2 and is denoted as HGLE(θ, r, β_1, β_2),

The survival function of the HGLE distribution is obtained as

$$\bar{G}(x) = \left(\frac{\theta\exp\left(-r\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)}{1-\bar{\theta}\exp\left(-r\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)}\right)^{\frac{1}{r}} \tag{3.1}$$

The pdf of HGLE distribution is given by

$$g(x) = \frac{\theta^{\frac{1}{r}}(\beta_1 + \beta_2 x)\exp\left(-\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)}{\left(1-\bar{\theta}\exp\left(-r\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)\right)^{1+\frac{1}{r}}} \tag{3.2}$$

where $x > 0, \theta > 0, r > 0, \beta_1 > 0, \beta_2 > 0, \bar{\theta} = 1 - \theta$.

The corresponding hazard rate function is given by

$$h(x) = \frac{(\beta_1 + \beta_2 x)}{1-\bar{\theta}\exp\left(-r\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)} \tag{3.3}$$

By taking r=1 in (3.2) we get the Marshall-Olkin Linear Exponential distribution (MOLE), and when r=1 and $\theta = 1$, (3.2) reduces to the LE distribution. When β_2=0, (3.2) gives the Harris Extended Exponential distribution as given in Pinho et al (2015). In the same way it will reduce to Marshall Olkin Exponential distribution and Exponential distribution when r=1 and r=1, $\theta = 1$ respectively. If $\beta_1 = 0$ this distribution reduces to the Harris extended form of Rayleigh distribution and it includes the Marshall-Olkin generalization of Rayleigh distribution (if r=1), Rayleigh distribution (if $\theta = 1$) as its special cases.

The quantile function of HGLE distribution can be obtained by inverting the cdf given in (2.1), and is obtained as

$$x_p = \frac{\beta_1}{\beta_2}\left[\sqrt{1 + \frac{2\beta_2}{r\beta_1^2}\log(\bar{\theta} + \theta(1-p)^{-r})} - 1\right], \tag{3.4}$$

where $0 < p < 1$.

Some possible shapes of the pdf of HGLE distribution for different values of parameters are given in Fig. 1.

Some possible shapes of the hazard rate function of the distribution for different values of parameters are given in Fig. 2

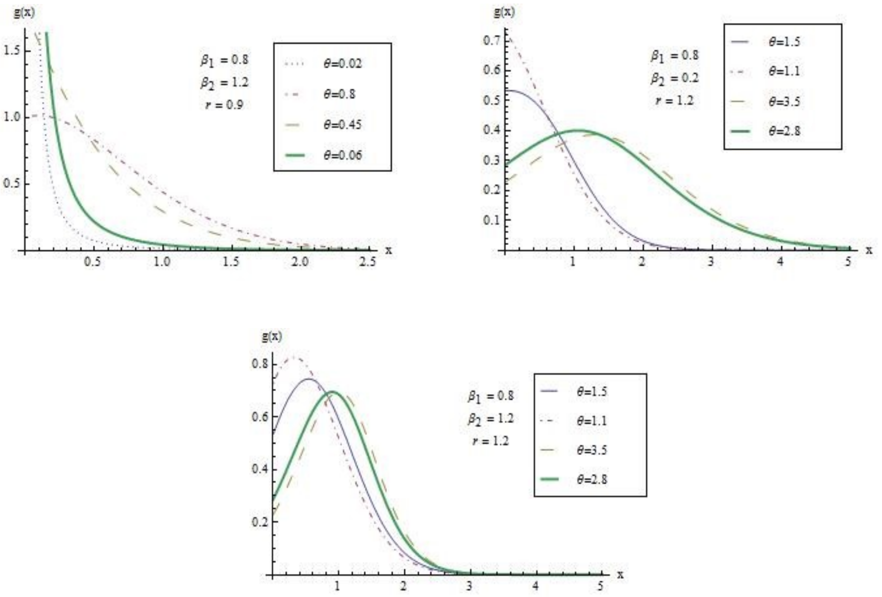

Figure 1: pdf of HGLE Distribution for different values of parameters.

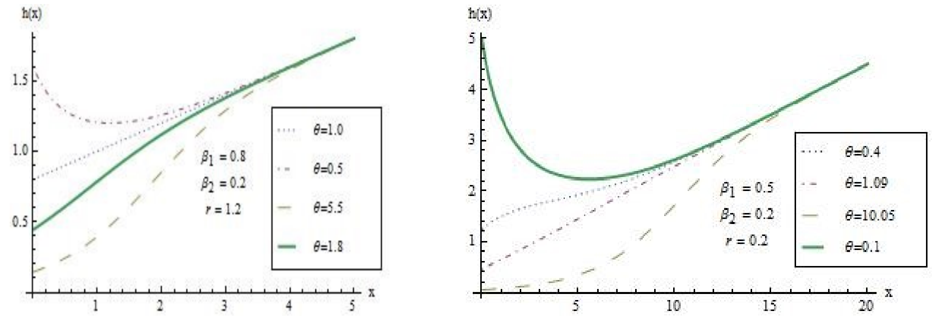

Figure 2: Hazard rate function of HGLE Distribution for different values of parameters.

Theorem: 1 *The pdf of HGLE distribution can be represented as a linear combination of Linear exponential density function as*

$$g(x) = \sum_{i=0}^{\infty} \eta_i f_{(\beta_1^*,\beta_2^*)}(x)$$

where $f_{(\beta_1^*,\beta_2^*)}(x)$ follows LE distribution with parameters $\beta_1^* = (ri+1)\beta_1$ and $\beta_2^* = (ri+1)\beta_2$.

Proof: Consider the pdf of HGLE distribution given in (3.2),

For $\theta < 1$,

the expression given by (3.2) can be expanded using the negative binomial power series as follows,

$$g(x) = \theta^{\frac{1}{r}}(\beta_1 + \beta_2 x)\exp\left(-\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)\sum_{i=0}^{\infty}\binom{i+r^{-1}}{i}\bar{\theta}^i \exp\left(-\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)ri\right)$$

$$= \sum_{i=0}^{\infty} \frac{\bar{\theta}^i \theta^{\frac{1}{r}}}{(ri+1)}\binom{i+r^{-1}}{i}\left((ri+1)(\beta_1+\beta_2 x)\exp\left(-(ri+1)\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)\right)$$

$$= \sum_{i=0}^{\infty} \eta_i f_{(\beta_1^*,\beta_2^*)}(x)$$

Here for $\theta < 1$, $\eta_i = \frac{\bar{\theta}^i \theta^{\frac{1}{r}}}{(ri+1)} \binom{i + r^{-1}}{i}$ (3.5)

and $f_{(\beta_1^*,\beta_2^*)}(x) = \left((ri+1)(\beta_1 + \beta_2 x)\exp\left(-(ri+1)\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)\right)$

where $f_{(\beta_1^*,\beta_2^*)}(x)$ follows LE distribution with parameters $\beta_1^* = (ri+1)\beta_1$ and $\beta_2^* = (ri+1)\beta_2$.

For $\theta > 1$, let us take $\theta = \theta_1^{-1}$ so that $0 < \theta_1 < 1$.

On simplification (3.2) reduces to

$$g(x) = \frac{\theta_1(\beta_1 + \beta_2 x)y}{\left(1 - \bar{\theta}_1(1 - y^r)\right)^{1+\frac{1}{r}}}$$

where $y = \exp\left(-\left(\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)\right)$ and $\bar{\theta}_1 = 1 - \theta_1$. Since the denominator lies in the interval (0, 1), we can use the negative binomial power series expansion, then we have

$$g(x) = \theta_1(\beta_1 + \beta_2 x)y \sum_{j=0}^{\infty} \bar{\theta}_1{}^j (1 - y^r)^j \binom{j + r^{-1}}{j}$$

$$= \theta_1(\beta_1 + \beta_2 x)y \sum_{j=0}^{\infty} \sum_{i=0}^{j} (-1)^i (\bar{\theta}_1)^j y^{ri} \binom{j + r^{-1}}{j} \binom{j}{i}$$

Interchanging the order of summation and substituting y we get

$$g(x) = \sum_{i=0}^{\infty} \frac{\theta_1(-1)^i}{(ri+1)} \left(\sum_{j=i}^{\infty} (\bar{\theta}_1)^j \binom{j + r^{-1}}{j} \binom{j}{i} \right)$$
$$\left\{(ri+1)(\beta_1 + \beta_2 x)\exp\left(-(ri+1)\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)\right\}$$

which leads to

$g(x) = \sum_{i=0}^{\infty} \eta_i f_{(\beta_1^*,\beta_2^*)}(x)$ where

$$\eta_i = \frac{\theta_1(-1)^i}{(ri+1)} \left(\sum_{j=i}^{\infty} (\bar{\theta}_1)^j \binom{j + r^{-1}}{j} \binom{j}{i} \right) \quad (3.6)$$

and

$$f_{(\beta_1^*,\beta_2^*)}(x) = \left((ri+1)(\beta_1 + \beta_2 x)\exp\left(-(ri+1)\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)\right)$$

which is the pdf of LE distribution with parameters $\beta_1^* = (ri+1)\beta_1$ and $\beta_2^* = (ri+1)\beta_2$.

Theorem: 2 *The moment generating function of HGLE distribution, denoted as M(t) is given by* $M(t) = \sum_{i=0}^{\infty} \eta_i \omega_{s,ri}$ *and* $\omega_{s,ri}$ *can be expressed as*

$$\omega_{s,ri} = \int_0^{\infty} \exp(tx)(ri+1)(\beta_1 + \beta_2 x)\exp\left(-(ri+1)\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)dx$$

For $\theta < 1$, $\eta_i = \frac{\bar{\theta}^i \theta^{\frac{1}{r}}}{(ri+1)} \binom{i + r^{-1}}{i}$

For $\theta \geq 1$, $\eta_i = \frac{\theta_1(-1)^i}{(ri+1)} \sum_{j=0}^{\infty} \bar{\theta}_1^j \binom{j + r^{-1}}{j} \binom{j}{i}$, *where* $\theta_1 = \frac{1}{\theta}, 0 \leq \theta_1 \leq 1$.

Proof:

We have $M(t) = \int_0^{\infty} \exp(tx) f(x; \theta, r, \beta_1, \beta_2)dx$

$$= \int_0^{\infty} \exp(tx) \sum_{i=0}^{\infty} \eta_i (ri+1)(\beta_1 + \beta_2 x)\exp\left(-(ri+1)\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right)dx$$

(by using theorem 1)

$$= \sum_{i=0}^{\infty} \eta_i \int_0^{\infty} (ri+1)(\beta_1 + \beta_2 x)\exp\left(-\left((ri+1)\beta_1 - t\right)x\right)\exp\left(-(ri+1)\left(\frac{\beta_2}{2}x^2\right)\right)dx$$

consider

$$\int_0^\infty (ri+1)\beta_1 x^{2s}\exp\left(-\left((ri+1)\beta_1-t\right)x\right)dx+\int_0^\infty (ri+1)\beta_2 x^{2s+1}\exp\left(-\left((ri+1)\beta_1-t\right)x\right)dx$$

$$=\frac{(ri+1)\beta_1\Gamma(2s+1)}{\left((ri+1)\beta_1-t\right)^{2s+1}}+\frac{(ri+1)\beta_2\Gamma(2s+2)}{\left((ri+1)\beta_1-t\right)^{2s+2}};\quad t\le\beta_1$$

Then

$$M(t)=\sum_{i=0}^{\infty}\eta_i\sum_{s=0}^{\infty}\frac{\left((ri+1)\frac{\beta_2}{2}\right)^s}{s!}(-1)^s\left(\frac{(ri+1)\beta_1\Gamma(2s+1)}{\left((ri+1)\beta_1-t\right)^{2s+1}}+\frac{(ri+1)\beta_2\Gamma(2s+2)}{\left((ri+1)\beta_1-t\right)^{2s+2}}\right);\ t\le\beta_1$$

where η_i can be chosen suitably for $\theta\le 1$ and $\theta>1$ as given in equation (3.5) and (3.6).

Theorem: 3

If X has $HGLE(r,\beta_1,\beta_2,\theta)$ distribution. The the k^{th} moment of X denoted by μ_k is given by,

$$\mu_k=\sum_{i=0}^{\infty}\sum_{s=0}^{\infty}\theta^{\frac{1}{r}}\bar{\theta}^i\binom{1+r^{-1}}{i}\frac{\left(-(ri+1)\frac{\beta_2}{2}\right)^s}{s!}\left(\frac{\beta_1\Gamma(k+2s+1)}{\left((ri+1)\beta_1\right)^{(k+2s+1)}}+\frac{\beta_2\Gamma(k+2s+2)}{\left((ri+1)\beta_1\right)^{(k+2s+2)}}\right)$$

Proof:

We have,

$$\mu_k=\int_0^\infty x^k\frac{\theta^{\frac{1}{r}}(\beta_1+\beta_2x)\exp\left(-\left(\beta_1+\frac{\beta_2}{2}x^2\right)\right)}{\left(1-\bar{\theta}\exp\left(-r\left(\beta_1+\frac{\beta_2}{2}x^2\right)\right)\right)^{1+\frac{1}{r}}}dx\qquad \int_0^\infty x^k g(x)dx$$

$$=\int_0^\infty x^k\theta^{\frac{1}{r}}(\beta_1+\beta_2x)\sum_{i=0}^{\infty}\binom{1+r^{-1}}{i}\bar{\theta}^i\exp\left(-(ri+1)\left(\beta_1x+\frac{\beta_2}{2}x^2\right)\right)dx$$

Interchanging the integration and summation we get

$$=\sum_{i=0}^{\infty}\binom{1+r^{-1}}{i}\theta^{\frac{1}{r}}\bar{\theta}^i\int_0^\infty x^k(\beta_1+\beta_2x)\exp\left(-(ri+1)\left(\beta_1x+\frac{\beta_2}{2}x^2\right)\right)dx$$

$$=\sum_{i=0}^{\infty}\binom{1+r^{-1}}{i}\theta^{\frac{1}{r}}\bar{\theta}^i\sum_{s=0}^{\infty}\frac{\left(-(ri+1)\frac{\beta_2}{2}\right)^s}{s!}\left(\int_0^\infty \beta_1x^{k+2s}\exp\left(-(ri+1)(\beta_1x)\right)dx+\right.$$
$$\left.\int_0^\infty \beta_2x^{k+2s+1}\exp\left(-(ri+1)(\beta_1x)\right)dx\right)$$

$$=\sum_{i=0}^{\infty}\sum_{s=0}^{\infty}\theta^{\frac{1}{r}}\bar{\theta}^i\binom{1+r^{-1}}{i}\frac{\left(-(ri+1)\frac{\beta_2}{2}\right)^s}{s!}\left(\frac{\beta_1\Gamma(k+2s+1)}{\left((ri+1)\beta_1\right)^{(k+2s+1)}}+\frac{\beta_2\Gamma(k+2s+2)}{\left((ri+1)\beta_1\right)^{(k+2s+2)}}\right)$$

This completes the proof of the theorem.

4 Compounding

Ghitany et al (2005), Ghitany and Kotz (2007) and Krishna et al (2013) expressed Marshall Olkin extended forms of distributions Marshall and olkin (1997) as compound distributions with exponential as mixing density. This gives a new parametric family of distributions in terms of existing ones.

Let $\bar{F}(x/\alpha), x\in\Re, \alpha\in\Re$, be the conditional survival function of a continuous random variable X given . Let follows a distribution with probability density function $m(\alpha)$. A distribution with survival function

$$\bar{F}(x)=\int_{-\infty}^{\infty}\bar{F}(x/\alpha)m(\alpha)\,d\alpha,\ x\in\Re$$

is called a compound distribution with mixing density $m(\alpha)$.

The following theorem shows that under suitable conditions the HGLE distribution can be obtained as a compound distribution.

Theorem: 4 *Let X be a continuous random variable with conditional pdf given by* $\bar{F}(x/\alpha)=$

$$exp\left(\left(\left(exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)\right)^{-1}-1\right)\alpha, \qquad x,r,\beta_1,\beta_2,\alpha\ >0.$$

Let α follows an exponential distribution with with pdf given by $m(\alpha)=\theta e^{-\theta\alpha}$, $\theta,\ \alpha>0$. Then the proportional failure rate model of the compound distribution of X becomes the HGLE (r, θ, β_1, β_2) distribution.

Proof: For all $x>0$, β_1, β_2, $\theta>0$, the unconditional survival function of x is given by

$$\bar{F}(x)=\int_0^\infty \bar{F}(x/\alpha)m(\alpha)d\alpha$$

$$=\alpha\int_0^\infty \exp\left(\left(\left(\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)\right)^{-1}-1\right)\alpha\ \exp(-\alpha\theta)\ d\alpha$$

$$\bar{F}(x)=\frac{\alpha}{\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)^{-1}-\bar{\alpha}}$$

$$\bar{F}(x)=\frac{\alpha\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)}{1-\bar{\alpha}\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)}$$

Let us take the proportional failure rate model of $\bar{F}(x)$, Then

$$\bar{G}(x)=\bar{F}(x)^{\frac{1}{r}}$$

$$\bar{G}(x)=\left(\frac{\alpha\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)}{1-\bar{\alpha}\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)}\right)^{\frac{1}{r}}$$

which is the survival function of a random variable with HGLE distribution with parameter $(r,\theta,\beta_1,\beta_2)$.

5 Renyi Entropy

Numerous measures of entropy are discussed and studied by many researchers and these Entropy measures has been used in various situations of science and technology especially in communications engineering and information technology. The entropy of a random variable X with density function $g(x)$ is a measure of uncertainty. The Rényi entropy is defined by

$$I_g(\delta)=(1-\delta)^{-1}\log\int_{-\infty}^{\infty} g^\delta(x)dx$$

where $\delta>0$ and $\delta\neq 1$

$$I_g(\delta)=(1-\delta)^{-1}\log\left(\begin{array}{c}\sum_{k=0}^{\infty}\sum_{s=0}^{\infty}\sum_{i=0}^{\delta}\theta^{\frac{\delta}{r}}(1-\theta)^k\frac{\Gamma\delta\left(1+\frac{1}{r}\right)+k}{\Gamma\delta\left(1+\frac{1}{r}\right)k!}\binom{\delta}{i}\beta_1^i\beta_2^{\delta-i}\\ (-1)^s\frac{\left((\delta+rk)\frac{\beta_2}{2}\right)^s}{s!}\frac{\Gamma(\delta+2s-i+1)}{((\delta+rk)\beta_1)^{(\delta+2s-i+1)}}\end{array}\right)$$

Table 1 displays the Renyi entropy for HGLE distribution at $\delta=2.0, 2.5, 3.0$, $\beta_1=0.5$, $\beta_2=0.9$, $r=1.2$ and for different choices of $\theta>1$.

θ	$I_g(\delta)$ $\delta=2.0$	$I_g(\delta)$ $\delta=2.5$	$I_g(\delta)$ $\delta=3.0$
1.5	0.8290	0.7549	0.7336
2.0	0.8530	0.8003	0.7783
2.5	0.8661	0.8228	0.7995
3.0	0.8731	0.8342	0.8094
3.5	0.8765	0.8396	0.8134
4.0	0.8777	0.8415	0.8141

Table 1: Renyi Entropy of HGLE Distribution at $\delta=2.0, 2.5, 3.0$, $r=1.2$, $\beta_1=0.5$, $\beta_2=0.9$

6 Mean Residual Life

The expected additional life time given that a component has survived until time t is called mean residual life (MRL). The importance of MRL function in reliability and survival analysis is that it describes the aging process. Also MRL function uniquely determines its distribution function.

The MRL of a random variable X representing life of a component is given as follows.

$$M_R(t) = \frac{1}{\bar{G}(t)}\int_t^\infty \bar{G}(x)dx, \quad t > 0$$

The MRL function of a lifetime random variable X with HGLE distribution is given by

$$M_R(t) = \left(\frac{1-\bar{\theta}\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}t^2\right)\right)}{\theta\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}t^2\right)\right)}\right)^{\frac{1}{r}} \int_t^\infty \left(\frac{\theta\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)}{1-\bar{\theta}\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)}\right)^{\frac{1}{r}} dx$$

Table 2 displays MRL function for HGLE distribution at point t=0.2, 0.5, 0.9, $r = 1.2, \beta_1 = 0.5, \beta_2 = 0.9$ and for different choices of parameter θ.

θ	MRL,t=0.2	MRL,t=0.5	MRL,t=0.9
1.0	0.8087	0.6906	0.5735
1.5	0.9191	0.7676	0.6143
2.0	1.0034	0.8302	0.6502
2.5	1.0716	0.8829	0.6823
3.0	1.1288	0.9284	0.7114
3.5	1.1781	0.9685	0.7379

Table 2: Mean residual life of HGLE Distribution at $r = 1.2$, $\beta_1 = 0.5$, $\beta_2 = 0.9$

7 Order Statistics

Let $X_1, X_2, \ldots, X_n$ be a random sample taken from HGLE distribution and $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the corresponding order statistics. We derive the pdf of i^{th} order statistics $X_{i:n}$ which is denoted as $g_{i:n}(x)$, and express it as a linear combination of HGLE density function. We have the general formula for the pdf of the i^{th} order statistics as follows

$$g_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} g(x)\left(\bar{G}(x)\right)^{n-i}\left(1-\bar{G}(x)\right)^{i-1}$$

By using Binomial expansion, we get

$$g_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} g(x)\sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}\left(\bar{G}(x)\right)^{n+j-i}$$

On simplification $g_{i:n}(x)$ reduces to

$$g_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{i-1}\binom{i-1}{j}\frac{(-1)^j}{(n+j-i+1)}\theta^{\left(\frac{n+j-i+1}{r}\right)}$$

$$(n+j-i+1)(\beta_1+\beta_2 x)\frac{\exp\left(-(n+j-i+1)\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)r\right)}{\left(1-\bar{\theta}\exp\left(-\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)r\right)\right)^{1+\frac{(n+j-i+1)}{r}}}$$

$$= \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{i-1}\binom{i-1}{j}\frac{(-1)^j}{(n+j-i+1)} g_{n,j,i}(x)$$

where $g_{n,j,i}(x)$ is the density function of HGLE distribution with parameters $r(n+j-i+1)^{-1}$, θ, $(n+j-i+1)\beta_1$, $(n+j-i+1)\beta_2$.

The k^{th} moment of i^{th} order statistic of HGLE distribution can be derived by using Theorem 3 as the distribution of order statistics can be expressed as the linear combination of HGLE density function. Consider the asymptotic distribution of the first order statistic $X_{1:n}$ and n^{th} order statistic $X_{n:n}$. y using the asymptotic results for $X_{1:n}$ and $X_{n:n}$ (Arnold et al (1992), Kotz and Nadarajah (2001)), we can find the limiting distribution of extreme order statistics. We have

$$\lim_{n\to\infty} P(X_{1:n} \le a_n^* + b_n^* t) = 1 - \exp(-t^{-\alpha}),\ \ t > 0,\ \ \alpha > 0, \quad (7.1)$$

of Weibull type, where $a_n^* = G^{-1}(0)$ and $b_n^* = G^{-1}\left(\frac{1}{n}\right) - G^{-1}(0)$ if and only if $G^{-1}(0)$ is finite and, for all $t > 0$ and $c > 0$,

$$\lim_{\in\to 0^+} \frac{G(G^{-1}(0)+\in t)}{G(G^{-1}(0)+\in)} = t^{\alpha}. \quad (7.2)$$

For the maximal order statistics $X_{n:n}$, we have

$$\lim_{n\to\infty} P(X_{n:n} \le a_n + b_n t) = \exp(-e^{-t}),\ \ -\infty < t < \infty \quad (7.3)$$

of extreme value type, where $a_n = G^{-1}\left(1 - \frac{1}{n}\right)$ and $b_n = \frac{1}{ng(a_n)}$ if

$$\lim_{x\to\infty} \frac{d}{dx}\left(\frac{1}{h(x)}\right) = 0. \quad (7.4)$$

The selection of the norming constants a_n and b_n are not unique but depend on G. In general we use the result given in (Arnold et al (1992)) and the same used in (Ghitany and Kotz (2007)). The following theorem gives the limiting distributions of the smallest and largest order statistics from the HGLE distribution.

Theorem: 5

Let $X_{1:n}$ and $X_{n:n}$ be, respectively, the smallest and largest order statistics from $HGLE(\theta, r, \beta_1, \beta_2)$ distribution. Then

(i). $\lim_{n\to\infty} P\{X_{1:n} \le b_n^* t\} = 1 - \exp(-t),\ t > 0$, where $b_n^* = G^{-1}\left(\frac{1}{n}\right)$ and $G^{-1}(.)$ is given by (3.4).

(ii). $\lim_{n\to\infty} P\{X_{n:n} \le a_n + b_n t\} = \exp(-e^{-t}),\ -\infty < t < \infty$, where $a_n = G^{-1}\left(1 - \frac{1}{n}\right)$, $b_n = \frac{1}{ng(a_n)}$, and $g(.)$ and $G^{-1}(.)$, respectively are given by (3.2) and (3.4).

Proof:

(i). For HGLE distribution $G^{-1}(0) = 0$ which is finite and by using L Hospitals rule

$$\lim_{\in\to 0^+} \frac{G(G^{-1}(0)+\in t)}{G(F^{-1}(0)+\in)} = t \lim_{\in\to 0^+} \frac{g(\in t)}{g(t)} = t.$$

From (7.2) we have α=1 and the asymptotic distribution of $X_{1:n}$ is of Weibull type, where

$$b_n^* = G^{-1}\left(\frac{1}{n}\right) = \frac{\beta_1}{\beta_2}\left[\sqrt{1 + \frac{2\beta_2}{r\beta_1^2}\log\left(\bar{\theta} + \theta\left(1 - \frac{1}{n}\right)^{-r}\right)} - 1\right]$$

Hence (i) follows from (7.1) and (7.2).

(ii). For HGLE distribution, by using Von Mises sufficient condition for the weak convergence and the properties given in ([?]), we get,

$$\lim_{x\to G^{-1}(1)} \frac{d}{dx}\left(\frac{1}{h(x)}\right) = \lim_{x\to\infty}\left(\begin{array}{c} r(1-\theta)\exp\left(-r\left(\beta_1 x + \frac{\beta_2}{2}x^2\right)\right) \\ -\frac{\left(1-(1-\theta)\exp\left(-r\left(\beta_1 x+\frac{\beta_2}{2}x^2\right)\right)\beta_2\right)}{(\beta_1+\beta_2 x)^2} \end{array}\right) = 0$$

where $a_n = G^{-1}(1 - n^{-1})$, $G^{-1}(.)$ given in (3.4) and $b_n = \frac{1}{ng(a_n)}$, where b_n can be obtained by using a_n and (3.2). Hence, statement (ii) follows from (7.3) and (7.4).

8 Estimation

In this section we consider maximum likelihood estimation for a given sample $X_1, X_2, \ldots, X_n$. Then the log likelihood function is given by

$$\log L = \frac{n}{r}\log\theta + \sum_{i=0}^{n}\log(\beta_1 + \beta_2 x_i) - \sum_{i=0}^{n}\left(\beta_1 x_i + \frac{\beta_2}{2}x_i^2\right)$$
$$-\left(1 + \frac{1}{r}\right)\sum_{i=0}^{n}\log\left(1 - \bar{\theta}\exp\left(-r\left(\beta_1 x_i + \frac{\beta_2}{2}x_i^2\right)\right)\right)$$

The maximum likelihood estimates can be obtained by solving the equations $\frac{\partial logL}{\partial r} = 0, \frac{\partial logL}{\partial \beta_1} = 0, \frac{\partial logL}{\partial \beta_2} = 0,\ \frac{\partial logL}{\partial \theta} = 0$. These equations are non-linear and can be solved iteratively using nlm program in R software.

9 Data Analysis

Example 1

Consider the data set which gives the survival times for 121 breast cancer patients treated over the period 1929-1938 (Boag (1984), Lawless (2003)). We compare the HGLE distribution with two other distributions LE distribution with cdf given in (2.1) and MOLE distribution. The survival function of MOLE distribution can be derived by setting r=1 in (3.1). For the given data set, we estimate the unknown parameter of each distribution by the maximum likelihood method, with these obtained estimates we obtain the values of Kolmogrov Smirnov (K-S) statistics and p value. From the values given in Table 3, we observe that HGLE distribution is a competitive distribution compared with other two distributions.

Model	Parameters	Estimates	Log likelihood	K-S Statistics	p value
LE	β_1	0.0155	579.72	0.0826	0.3816
	β_2	0.000186			
MOLE	β_1	0.0029963			
	β_2	0.0000612	579.456	0.0563	0.8354
	θ	2.360			
HGLE	r	3.800			
	β_1	0.0199	578.780	0.0542	0.8698
	β_2	0.000135			
	θ	2.099			

Table 3: Fitting for the LE, MOLE and HGLE Distribution

Example 2

Consider the data set given in Chhikara and Folks (1989) and Lawless (2003), the data on repair times (in hours) for 46 failures of an airbone communications receiver and here we compare the HGLE distribution with MOLE distribution. Table 4 gives the MLE's of the fitted models to the current data with the K-S statistics and p- value. From the given table we can conclude that the HGlE distribution better fits the given data than the MOLE distribution.

Model	Parameters	Estimates	Log likelihood	K-S Statistics	p value
MOLE	β_1	0.00535			
	β_2	0.00324	103.5308	0.0955	0.785
	θ	0.0163			
HGLE	r	2.1513			
	β_1	0.000401	102.176	0.0676	0.9827
	β_2	0.00721			
	θ	0.00687			

Table 4: Fitting for the MOLE and HGLE Distribution

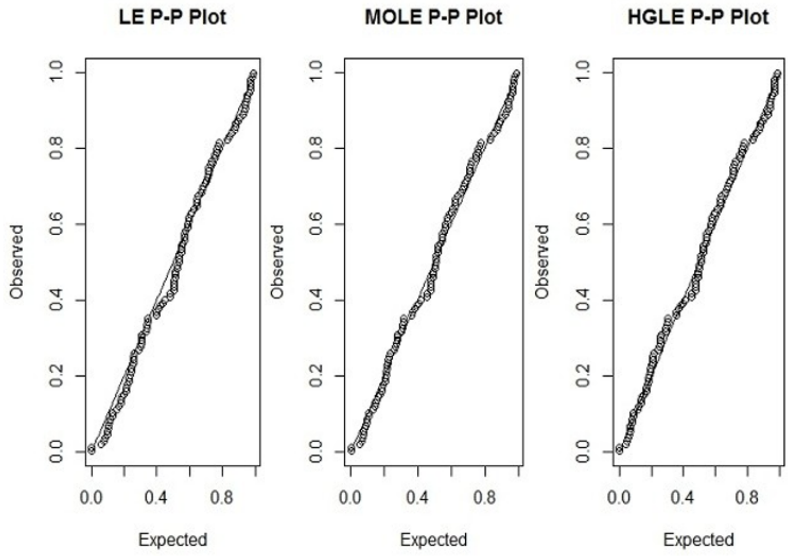

Figure 3: P P Plot for LE distribution, MOLE distribution and HGLE

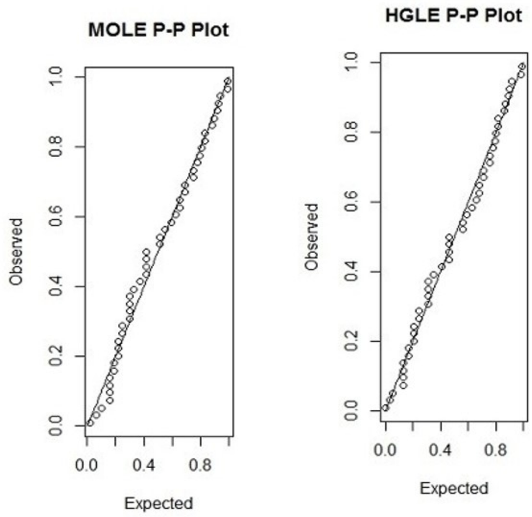

Figure 4: P P Plot for MOLE distribution and HGLE distribution

10 Conclusion

In this paper we have shown a generalization of LE distribution namely HGLE distribution. We have studied some of the statistical properties of the distribution such as probability density function, hazard rate function, moment generating function, distribution of order statistics, asymptotic distribution of extreme order statistics, Renyi entropy etc. The method of maximum likelihood estimation is also derived and also described two cases of real data application to show how HGLE distribution performs better than its baseline distributions.

References

[1] Aly and Benkherouf (2011) Aly and Benkherouf. L., A new family of distribution based on probability generating functions, *Sankhya B-Applied and Interdisciplinary Statistics*, **73**(1), 70-80, (2011).

[2] Arnold et al (1992) Arnold, B.C., Balakrishnan, N., and Nagaraja, H.N., A first course in order statistics. New York: Wiley, (1992).

[3] Azzalini (1985) Azzalini, A., A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12**, 171-178, (1985).

[4] Barreto-Souza et al 2013 Barreto-Souza, W., Lemonte, A.J. and Cordeiro, G.M., General results for the Marshall-Olkin's family of distributions, *Annals of the Brazilian Academy of Sciences*, **85**, 3-21, (2013).

[5] Batsidis and lemonte (2014) Batsidis, A. and Lemonte, A.J., On the Harris extended family of distributions, *Statistics, A Journal of Theoretical and Applied Statistics*, **49**, 1400-1421, (2014).

[6] Boag (1984) Boag, J. W., Maximum likelihood estimates of the proportion of patients cured by Cancer therapy, *Journal of Royal Statistical Society B*, **11**, 15-53, (1984).

[7] Broadbent (1958) Broadbent, S., Simple mortality rates, *Journal of Applied Statistics*, **7**, 86, (1958).

[8] Carbone et al (1967) Carbone, P., Kellerthouse, L. and Gehan, E., Plasmacytic Myeloma: A Study of the relationship of survival to various clinical manifestations and anomalous protein type in112 patients, *American Journal of medicine*, **42** 937-948, (1967).

[9] Embrechts et al (2013) Embrechts P., Klüppelberg, C., and Mikosch T. *Modelling Extremal Events for Insurance and Finance*, Springer, (2013).

[10] Chhikara and Folks (1989) Chhikara, R. S. and Folks, J. L., The Inverse Gaussian Distribution; Theory, Methodology and Applications, Marcel Dekker, New York, (1989).

[11] Cordeiro et al (2015)Cordeiro, G. M., Ortega, E. M. M., and Lemonte, A. J., The Poisson generalized linear failure rate model. *Communications in Statistics-Theory and Methods*, **44**(10), 2037-2058, (2015).

[12] Ferreira and Steel (2007) Ferreira, J.T.A.S. and Steel, M.F.J., A new class of skewed multivariate distributions with applications to regression analysis, *Statistica Sinica*, **17**, 505-529, (2007).

[13] Ghitany et al (2005) Ghitany, M.E., AL-Hussaini, E.K. and Al-Jarallah, R.A., Marshall-Olkin extended Weibull distribution and its application to censored data. *Journal of Applied Statistics*, **32**, 1025-1034, (2005).

[14] Ghitany and Kotz (2007) Ghitany, M.E. and Kotz, S., Reliability properties of Extended linear failure rate distributions. *Probability in the Engineering and Informational Sciences*, **21**, 441-450, (2007).

[15] Harris (1948) Harris, T.E., Branching processes . *The Annals of Mathematical Statistics*, **19**(4), 474-494, (1948).

[16] Jose et al (2010) Jose, K.K., Naik, S.R. and Ristic, M.M., Marshall- Olkin q- Weibull distribution and max-min process, *Statistical papers*, **51**, 837-851, (2010).

[17] JoseandPaul(2018), Jose, K.K. and Paul, A., Reliability Test Plans for Percentiles Based on the Harris Generalized Linear Exponential Distribution, *Stochastics and Quality Control*, 33(1), 61-70, (2018).

[18] Jose and Sivdas (2015) Jose, K. K. and Sivadas, R., Marshall-Olkin Exponentiated Generalized Fréchet Distribution and Its Applications, *Journal of Probability and Statistical Science*, **13** (2), 167-178, (2015).

[19] Kotz and Nadarajah (2001) Kotz, S. and Nadarajah, S., *Extreme Value Theory, Theory and Applications*. Singapore: World Scientific, (2001).

[20] Krishna et al (2013) Krishna, E., Jose, K. K., Alice, T. and Ristic, M.M., Marshall−Olkin Fréchet distribution, *Communications in Statistics - Theory and Methods*, **42**, 4091-4107, (2013).

[21] Lai et al (2006) Lai, Diew, C. and Xie, M., Bathtub Shaped Failure Rate life Distributions, Stochastic ageing and dependence for reliability. Springer, New York, (2006).

[22] Lawless (2003) Lawless, Statistical Models and Methods for Lifetime Data, Second Edition, John Wiley & Sons Inc., NewYork, (2003).

[23] Marshall and olkin (1997) Marshall, A. and Olkin, I., A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families, *Biometrika*, **84**, 641-652, (1997).

[24] Mahmoud and Alam (2010) Mahmoud, M.A.W. and Alam, F.M.A., The generalized linear exponential distribution, *Statistics and Probability Letters*, **80** 1005-1014, (2010).

[25] Nadarajah et al (2014) Nadarajah, S., Shahsanaei, F., and Rezaei, S., A new four-parameter lifetime distribution. *Journal of Statistical Computation and Simulation*, **84** 2, 248-263, (2014).

[26] Pinho et al (2015) Pinho, L.G.B., Cordeiro, G.M. and Nobre, J.S., The Harris extended exponential distribution, *Communications in Statistics- Theory and Methods*, **44** (16), 3486-3502, (2015).

[27] Zang et al (2005) Zhang, T., Xie, M., Tang, L.C., and Ng, S.H., Reliability and modeling of systems integrated with firmware and hard ware, *International Journal of Reliability, Quality and safety Engineering*, **12**(3), 227-239, (2005).