

Generalized Linear Failure Rate Distribution

Ammar M. Sarhan^{a*} and Debasis Kundu^b

^a Department of Mathematics, Faculty of Science,
Mansoura University, Mansoura 35516, Egypt.
asarhan0@yahoo.com

^bDepartment of Mathematics and Statistics
Indian Institute of Technology Kanpur,
Kanpur, 208016, India
kundu@iitk.ac.in

Abstract

The exponential and Rayleigh are the two most commonly used distributions for analyzing lifetime data. These distributions have several desirable properties and nice physical interpretations. Unfortunately the exponential distribution only has constant failure rate and the Rayleigh distribution has increasing failure rate. The linear failure rate distribution generalizes both these distributions which may have non-increasing hazard function also. This paper introduces a new distribution, which generalizes the well known (1) exponential distribution, (2) linear failure rate distribution, (3) generalized exponential distribution, and (4) generalized Rayleigh distribution. The properties of this distribution are discussed in this paper. The maximum likelihood estimates of the unknown parameters are obtained. A real data set is analyzed and it is observed that the present distribution can provide a better fit than some other very well known distributions.

Key Words: Linear failure rate distribution, generalized exponential distribution, generalized Rayleigh distribution, maximum likelihood method.

*Current address: Department of Statistics and O.R., Faculty of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia. E-mail: asarhan@ksu.edu.sa

1 Introduction

In analyzing lifetime data one often uses the exponential, Rayleigh, linear failure rate or generalized exponential distributions. It is well known that exponential can have only constant hazard function whereas Rayleigh, linear failure rate and generalized exponential distribution can have only monotone (increasing in case of Rayleigh or linear failure rate and increasing/decreasing in case of generalized exponential distribution) hazard functions. Unfortunately, in practice often one needs to consider non-monotonic function such as bathtub shaped hazard function also, see, for example, Lai *et al.* [14]. In this paper we present a new simple distribution which may have bathtub shaped hazard function and it generalizes many well known distributions including the traditional linear failure rate distribution.

The linear failure rate distribution with the parameters $a > 0$ and $b > 0$, will be denoted by LFRD(a, b), has the following cumulative distribution function (CDF)

$$F_{LF}(x; a, b) = 1 - \exp \left\{ -ax - \frac{b}{2}x^2 \right\}, \quad x \geq 0. \quad (1)$$

It is easily observed that the exponential distribution (ED(a)) and the Rayleigh distribution (RD(b)) can be obtained from LFRD(a, b) by putting $b = 0$ and $a = 0$ respectively. Moreover, the probability density function (PDF) of the LFRD(a, b) can be decreasing or unimodal but the failure rate function is either constant or increasing only. See for example Bain [2], Sen and Bhattacharya [19], Lin *et al.* [16], Ghitany and Kotz [9] and the references cited their in this connection.

Recently, the two-parameter generalized exponential (GE) distribution has been introduced and studied quite extensively by Gupta and Kundu [11]. The two-parameter GE distribution with the parameters $a > 0, \theta > 0$, has the following distribution function

$$F_{GE}(x; a, \theta) = (1 - e^{-ax})^\theta; \quad x \geq 0. \quad (2)$$

It is observed that the GE(a, θ) can have decreasing or unimodal PDF and monotone (increasing/ decreasing) hazard functions, depending on the shape parameter θ . Unfortunately,

it can not have bathtub shaped hazard function. Surles and Padgett [20] recently introduced two-parameter Burr Type X distribution, also known as the generalized Rayleigh distribution, has increasing or bathtub shaped hazard function but it can not have decreasing hazard function.

In this paper we introduce a new three-parameter distribution function called as generalized linear failure rate distribution with three parameters a, b, θ and it will be denoted as $GLFRD(a, b, \theta)$. It is observed that the new distribution has decreasing or unimodal PDF and it can have increasing, decreasing and bathtub shaped hazard functions. We provide different statistical properties of this new distribution and some nice physical interpretations also. We provide the maximum likelihood estimates (MLEs) of the unknown parameters and it is observed that they can not be obtained in explicit forms. The MLEs can be obtained only by solving two non-linear equations. We analyze one real data set and it is observed that the present distribution provides better fit than many existing well known distributions.

The rest of the paper is organized as follows. In section 2 we present $GLFRD(a, b, \theta)$ and discuss its properties in Section 3. Section 4 discusses the distribution of order statistics of the $GLFRD$. The MLEs are provided in Section 5. Section 6 gives an illustrative example to explain how a real data set can be modeled by $GLFRD(a, b, \theta)$ and finally we conclude the paper in Section 7. A list of acronyms are provided in the Appendix A for quick references.

2 The GLFRD

Let X be a random variable with the following CDF for $a > 0, b > 0$ and $\theta > 0$ as follows;

$$F(x; a, b, \theta) = \left[1 - e^{-(ax + \frac{b}{2}x^2)}\right]^\theta, \quad x \geq 0. \quad (3)$$

Here θ is shape parameter. The distribution of this form is said to be a generalized linear failure rate distribution with parameters a, b, θ and will be denoted by $GLFRD(a, b, \theta)$. The PDF and the hazard function of $GLFRD(a, b, \theta)$ will be

$$f(x; a, b, \theta) = \theta(a + bx) \left[1 - e^{-(ax + \frac{b}{2}x^2)}\right]^{\theta-1} e^{-(ax + \frac{b}{2}x^2)}, \quad x \geq 0, \quad (4)$$

and

$$h(x; a, b, \theta) = \frac{\theta(a + bx) \left[1 - e^{-(ax + \frac{b}{2}x^2)}\right]^{\theta-1} e^{-(ax + \frac{b}{2}x^2)}}{1 - \left[1 - e^{-(ax + \frac{b}{2}x^2)}\right]^{\theta}}, \quad (5)$$

respectively. Recently, it is observed, see Gupta and Gupta [10], that the reversed hazard function plays an important role in the reliability analysis. The reversed hazard function of the GLFRD(a, b, θ) is

$$r(x; a, b, \theta) = \frac{f(x; a, b, \theta)}{F(x; a, b, \theta)} = \theta \frac{(a + bx)e^{-(ax + \frac{b}{2}x^2)}}{1 - e^{-(ax + \frac{b}{2}x^2)}} = \theta \frac{f(x; a, b, 1)}{F(x; a, b, 1)} = \theta r(x; a, b, 1). \quad (6)$$

It is well known that the hazard function or the reversed hazard function uniquely determines the corresponding probability density function. From (6) it is clear that the GLFRD(a, b, θ) is a proportional reversed hazard family. It may be mentioned that the reversed hazard function is a decreasing function.

Figure 1 and Figure 2 provide the PDFs and the failure rate functions of GLFRD(a, b, θ) for different parameter values. From the above figures it is immediate that the PDFs can be decreasing or unimodal and the hazard functions can be increasing, decreasing or bathtub shaped. It is immediate that from GLFRD(a, b, θ), the following special cases can be derived: (i) Linear failure distribution LFRD(a, b), when $\theta = 1$, (ii) Generalized exponential distribution GE(a, θ), when $b = 0, a > 0$, (iii) Generalized Rayleigh distribution GRD(b, θ), when $a = 0, b > 0$. One can easily verify that: (i) if $\theta = 1$, the hazard function is either increasing (if $b > 0$) or constant (if $b = 0$ and $a > 0$); (ii) when $\theta > 1$, the hazard function should be increasing only; and (iii) if $\theta < 1$, then the hazard function will be either decreasing if $b = 0$ or bathtub if $b > 0$.

COMMENTS: It is interesting to observe that when θ is a positive integer, the CDF of GLFRD(a, b, θ) represents the CDF of the maximum of a simple random sample of size θ from the linear failure rate distribution. Therefore, in that case the GLFRD(a, b, θ) provides the distribution function of a parallel system when each component has the linear failure rate distribution.

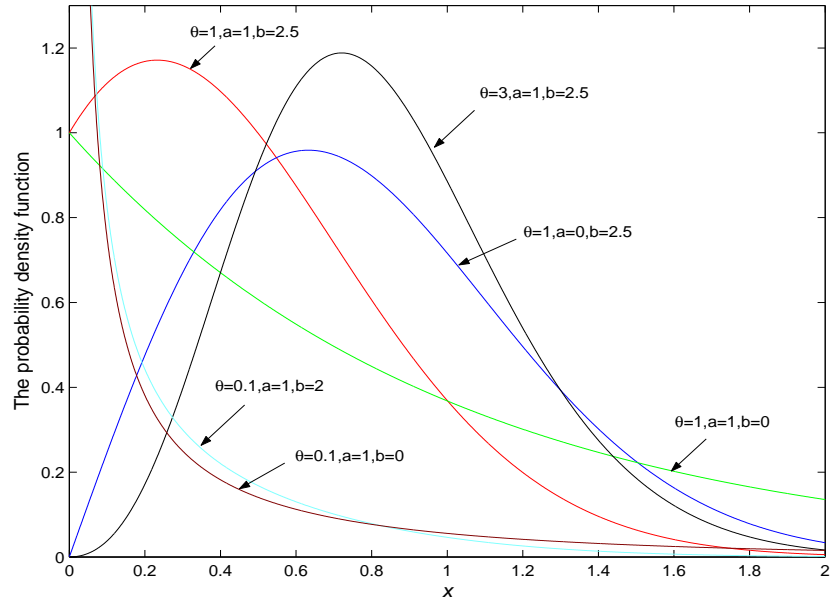


Figure 1. The probability density functions.

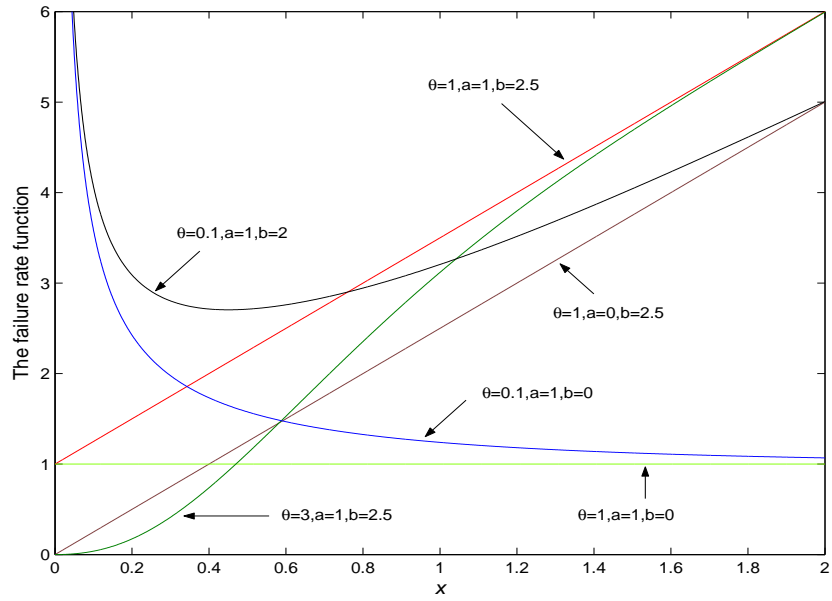


Figure 2. Failure rate function of the GLFRD(a, b, θ).

3 Statistical properties

3.1 Mean, median and mode

It is observed as expected that the mean of the GLFRD(a, b, θ) can not be obtained in explicit forms. It can be obtained as infinite series expansion. In general different moments of the GLFRD(a, b, θ) will be presented later.

The quantile x_q of the GLFRD(a, b, θ) is given by

$$x_q = \frac{1}{b} \left\{ -a + \sqrt{a^2 - 2b \ln \left(1 - q^{\frac{1}{\theta}} \right)} \right\}. \quad (7)$$

Using (7), the median of GLFRD(a, b, θ) can be obtained as

$$\text{med}(X) = \frac{1}{b} \left\{ -a + \sqrt{a^2 - 2b \ln \left(1 - 2^{-\frac{1}{\theta}} \right)} \right\}. \quad (8)$$

Moreover, the mode of GLFRD(a, b, θ) can be obtained as a solution of the following non-linear equation

$$\frac{b - (a + bx)^2}{a + bx} + (\theta - 1) \frac{(a + bx)}{\exp \left\{ ax + \frac{b}{2} x^2 \right\} - 1} = 0. \quad (9)$$

It is not possible to obtain the explicit solution in the general case. It has to be obtained numerically. For different special cases, the explicit forms may be obtained.

3.2 Moments

The following lemma gives the k^{th} moment of GLFRD(a, b, θ), when $\theta \geq 1$.

Lemma 3.1 *If X has GLFRD(a, b, θ), then the k^{th} moment of X , say $\mu^{(k)}$, is given as follows*

For $a = 0, b > 0$:

$$\mu^{(k)} = \frac{\theta \Gamma\left(\frac{k}{2} + 1\right)}{\left(\frac{b}{2}\right)^k} \sum_{i=0}^{\infty} \frac{(-1)^{-i} \binom{\theta-1}{i}}{(i+1)^{\frac{k}{2}+1}}, \quad (10)$$

For $a > 0, b \geq 0$:

$$\mu^{(k)} = \theta \sum_{i=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{(-1)^{-i} \binom{\theta-1}{i} \Gamma(k + \ell + 1) g_i^{(\ell)}(0)}{\ell! [a(i+1)]^{k+\ell+1}} \left[a + \frac{(k + \ell + 1)b}{(i+1)a} \right], \quad (11)$$

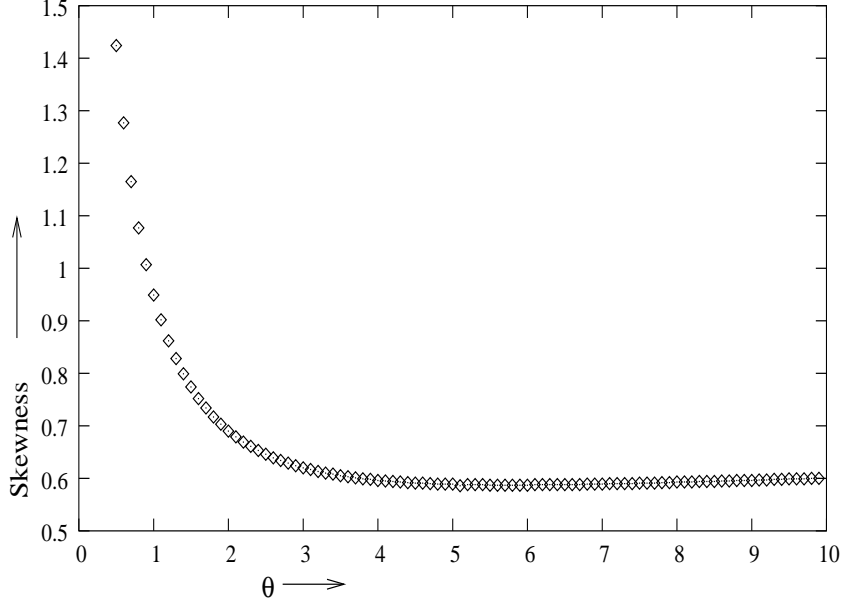


Figure 3. The skewness measure.

here $g_i^{(\ell)}(0) = \frac{d^\ell}{dx^\ell} \exp\{-\frac{1}{2}(i+1)bx^2\} \Big|_{x=0}$ and $\Gamma(\cdot)$ is the complete gamma function.

The proof of this lemma is provided in the Appendix B.

Based on the results, the measures of skewness and kurtosis of the GLFRD can be obtained as respectively;

$$\alpha = \frac{\mu^{(3)} - 3\mu\mu^{(2)} + 2\mu^3}{(\mu^{(2)} - \mu^2)^{\frac{3}{2}}} \quad (12)$$

and

$$\beta = \frac{\mu^{(4)} - 4\mu\mu^{(3)} + 6\mu^2\mu^{(2)} - 3\mu^4}{(\mu^{(2)} - \mu^2)^2}. \quad (13)$$

In Figure 3 and Figure 4 we provide α and β for different values of θ , when $a = 1$, $b = 2$. It is observed that α and β first decrease as θ increases and then start increasing.

3.3 Characterization

Let Y be an absolute continuous random variable with support only on the positive real axis. For $a > 0$, $b > 0$, for all real $t > 0$ and for all non-negative integer n , let us define

$$U(t) = -\ln\left(1 - e^{-(at + \frac{b}{2}t^2)}\right) \quad \text{and} \quad \alpha^{(n)}(t) = E(U^n(Y)|Y < t),$$

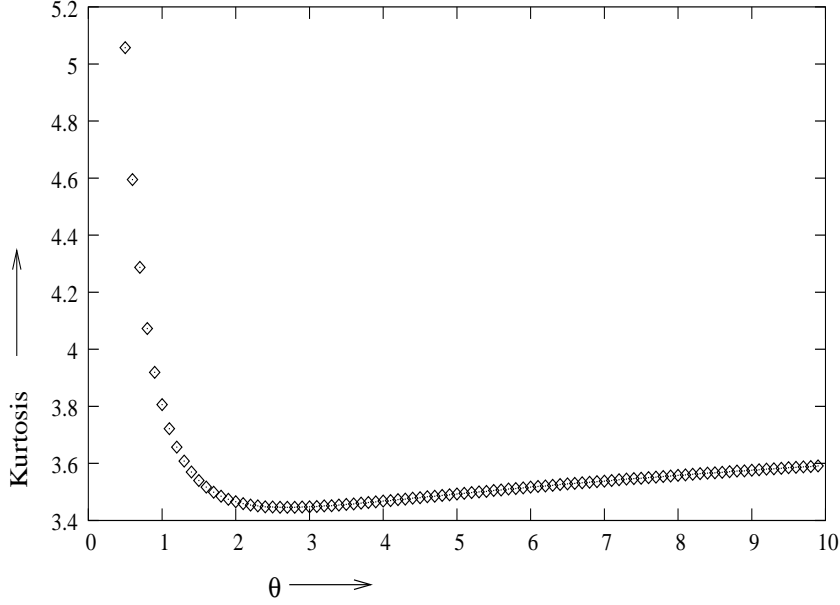


Figure 4. The Kurtosis measure.

here $U^n(\cdot)$ means the n -th power of $U(\cdot)$. Then we have the following results.

Theorem 3.1 The random variable Y follows $\text{GLFRD}(a, b, \theta)$ for $a > 0, b > 0, \theta > 0$, if and only if for all real $t > 0$, and for all non-negative integer n

$$\alpha^{(n)}(t) = U^n(t) + \frac{n}{\theta} \alpha^{(n-1)}(t). \quad (14)$$

Theorem 3.2 The random variable Y follows $\text{GLFRD}(a, b, \theta)$ for $a > 0, b > 0, \theta > 0$, if and only if for all real $t > 0$, and for all non-negative integer n

$$\text{Variance}(U(Y)|Y < t) = \frac{1}{\theta^2}. \quad (15)$$

Theorem 3.3 Suppose X_1, \dots, X_n are independent identically distributed random variables. Then the X_i follows GLFRD , if and only if $X_{(n)} = \max\{X_1, \dots, X_n\}$ follows GLFRD .

The proofs of Theorem 3.1 and Theorem 3.2 are provided in the appendix C. The proof of Theorem 3.3 is trivial and therefore it is omitted.

3.4 Maximum and Minimum

We have the following results regarding the maximum and minimum of the GLFRDs . All the proofs are quite straightforward and therefore they are omitted.

Theorem 3.4: If X_i s are independent random variables, and suppose X_i follows $\text{GLFRD}(a, b, \theta_i)$ for $i = 1, \dots, n$, then $X_{(n)}$ follows $\text{GLFRD}(a, b, \sum_{i=1}^n \theta_i)$

Theorem 3.5: If X_i s are independent and identically distributed random variables and X_i follows $\text{GLFRD}(a, b, \theta)$, then for all $-\infty < x < \infty$

$$\lim_{n \rightarrow \infty} P \left(\frac{X_{(n)} - b_n}{a_n} \leq x \right) = e^{-e^{-b\beta x}},$$

here $b_n = \frac{-2a + \sqrt{4a^2 + 8 \ln n}}{2b}$ and a_n is such that it satisfies the following two conditions;

$$\lim_{n \rightarrow \infty} a_n = 0, \quad \text{and} \quad \lim_{n \rightarrow \infty} a_n b_n = \beta.$$

Theorem 3.6: If X_i s are independent and identically distributed random variables and X_i follows $\text{GLFRD}(a, b, \theta)$, then for $X_{(1)} = \min\{X_1, \dots, X_n\}$ and for all $x > 0$

$$\lim_{n \rightarrow \infty} P \left(\frac{X_{(1)}}{c_n} \leq x \right) = 1 - e^{-(ax)^\theta},$$

where $c_n = n^{-\frac{1}{\theta}}$.

4 Distribution of order statistics

Let X_1, X_2, \dots, X_n be a simple random sample from $\text{GLFRD}(a, b, \theta)$ with PDF and CDF as in (4) and (3), respectively. Let $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ denote the order statistics obtained from this sample. In this section we provide the expressions for the PDFs and moments of order statistics for the $\text{GLFRD}(a, b, \theta)$. Also, the measures of skewness and kurtosis of the distribution of the r^{th} order statistic are presented. The PDF of $X_{(r)}$ is given by,

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)} [F(x; a, b, \theta)]^{r-1} [1 - F(x; a, b, \theta)]^{n-r} f(x; a, b, \theta), \quad (16)$$

where $f(x; a, b, \theta)$, $F(x; a, b, \theta)$ are the PDF and CDF given by (4) and (3), respectively.

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)} f(x; a, b, \theta) \sum_{j=0}^{n-r} \binom{n-r}{j} (-1)^j [F(x; a, b, \theta)]^{r+j-1} \quad (17)$$

substituting from (3) and (3) into (17), one gets

$$f_{r:n}(x) = \sum_{j=0}^{n-r} d_j(n, r) f(x; a, b, \theta_{r+j}), \quad (18)$$

where

$$\theta_i = i\theta, \quad d_j(n, r) = \frac{n(-1)^j \binom{n-r}{j} \binom{n-1}{r-1}}{(r+j)}.$$

The coefficients $d_j(n, r)$, $j = 1, 2, \dots, n - r$ do not depend on a, b, θ . Thus $f_{r:n}(x)$ is the weighted average of the generalized linear failure rate distributions with different shape parameters.

Theorem 4.1 *The k^{th} moment of order statistic $X_{(r)}$ is*

(i) if $a = 0, b > 0$:

$$\mu_{r:n}^{(k)} = \theta \left(\frac{2}{b}\right)^{\frac{k}{2}+1} \Gamma\left(\frac{k}{2} + 1\right) \sum_{j=0}^{n-r} \sum_{i=0}^{\infty} d_j^*(n, r) \frac{(-1)^i \binom{\theta_{r+j}-1}{i}}{(i+1)^{\frac{k}{2}+1}}. \quad (19)$$

(ii) if $a > 0, b \geq 0$:

$$\mu_{r:n}^{(k)} = \theta \sum_{j=0}^{n-r} \sum_{i=0}^{\infty} \sum_{\ell=0}^{\infty} d_j^*(n, r) \frac{(-1)^i \binom{\theta_{r+j}-1}{i} g_i^{(\ell)} \Gamma(k + \ell + 1)}{\ell! [(i+1)a]^{k+\ell+1}} \left[a + \frac{(k + \ell + 1)b}{(i+1)a} \right], \quad (20)$$

here $d_j^*(n, r) = (r+j) d_j(n, r)$.

The proof of this Theorem is given in the Appendix B.

Based on the results given in theorem (4.1), the measures of skewness and kurtosis of the distribution of the r^{th} order statistic can be evaluated from the following expressions

$$\alpha_{r:n} = \frac{\mu_{r:n}^{(3)} - 3\mu_{r:n} \mu_{r:n}^{(2)} + 2\mu_{r:n}^3}{\left(\mu_{r:n}^{(2)} - \mu_{r:n}^2\right)^{\frac{3}{2}}} \quad (21)$$

and

$$\beta_{r:n} = \frac{\mu_{r:n}^{(4)} - 4\mu_{r:n} \mu_{r:n}^{(3)} + 6\mu_{r:n}^2 \mu_{r:n}^{(2)} - 3\mu_{r:n}^4}{\left(\mu_{r:n}^{(2)} - \mu_{r:n}^2\right)^2}. \quad (22)$$

Extensive tables are available on request from the authors for different values of $\alpha_{r:n}$ and $\beta_{r:n}$ for the cases in which $n = 1, 2, 3, 4, 5$ and $\theta = 0.5, 1, 1.5, 2, 2.5, 3, 3.5$ with $1 \leq r \leq n$. From the table values the general findings are (i) The distribution of the r^{th} order statistic is positively skewed. (ii) For fixed θ and n , the values of $\beta_{r:n}$ increases as r increases.

5 Parameter estimations

In this section, we derive the maximum likelihood estimates of the unknown parameters a, b, θ of $\text{GLFRD}(a, b, \theta)$ based on a complete sample. Let us assume that we have a simple random sample X_1, X_2, \dots, X_n from $\text{GLFRD}(a, b, \theta)$. The likelihood function of this sample is

$$L = \prod_{i=1}^n f(x_i; a, b, \theta). \quad (23)$$

Substituting from (4) into (23), we get

$$L = \prod_{i=1}^n \left\{ \theta(a + bx_i) \left[1 - e^{-(ax_i + \frac{b}{2}x_i^2)} \right]^{\theta-1} e^{-(ax_i + \frac{b}{2}x_i^2)} \right\}. \quad (24)$$

It can be written as;

$$L = \theta^n \exp \{-aT_1 - bT_2\} \prod_{i=1}^n \left\{ (a + bx_i) \left[1 - e^{-(ax_i + \frac{b}{2}x_i^2)} \right]^{\theta-1} \right\}, \quad (25)$$

where $T_j = \frac{1}{j} \sum_{i=1}^n x_i^j$, $j = 1, 2$.

The log-likelihood function becomes

$$\mathcal{L} = n \ln \theta - aT_1 - bT_2 + \sum_{i=1}^n \ln(a + bx_i) + (\theta - 1) \sum_{i=1}^n \ln \left[1 - e^{-(ax_i + \frac{b}{2}x_i^2)} \right]. \quad (26)$$

Therefore, the normal equations are

$$\frac{\partial \mathcal{L}}{\partial a} = -T_1 + \sum_{i=1}^n \frac{1}{a + bx_i} + (\theta - 1) \sum_{i=1}^n \frac{x_i e^{-(ax_i + \frac{b}{2}x_i^2)}}{1 - e^{-(ax_i + \frac{b}{2}x_i^2)}} = 0, \quad (27)$$

$$\frac{\partial \mathcal{L}}{\partial b} = -T_2 + \sum_{i=1}^n \frac{x_i}{a + bx_i} + \frac{1}{2}(\theta - 1) \sum_{i=1}^n \frac{x_i^2 e^{-(ax_i + \frac{b}{2}x_i^2)}}{1 - e^{-(ax_i + \frac{b}{2}x_i^2)}} = 0, \quad (28)$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^n \ln \left[1 - e^{-(ax_i + \frac{b}{2}x_i^2)} \right] = 0. \quad (29)$$

The normal equations do not have explicit solutions and they have to be obtained numerically. Note that for a given a and b , the MLE of θ , say $\hat{\theta}(a, b)$ can be obtained as

$$\hat{\theta}(a, b) = -\frac{n}{\sum_{i=1}^n \ln \left[1 - e^{-(ax_i + \frac{b}{2}x_i^2)} \right]}.$$

Therefore, the MLEs of a and b can be obtained by solving two non-linear equations.

Asymptotic Confidence bounds. Since the MLEs of the unknown parameters a, b, θ can not be obtained in closed forms, it is not easy to derive the exact distributions of the MLEs. In this section, we derive the asymptotic confidence intervals of these parameters when $a > 0, b > 0$ and $\theta > 0$. The simplest large sample approach is to assume that the MLE $(\hat{a}, \hat{b}, \hat{\theta})$ are approximately trivariate normal with mean (a, b, θ) and covariance matrix I_0^{-1} , see Lawless [15], where I_0^{-1} is the inverse of the observed information matrix

$$I_0^{-1} = \begin{pmatrix} -\frac{\partial^2 \log L}{\partial a^2} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} & -\frac{\partial^2 \log L}{\partial a \partial b} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} & -\frac{\partial^2 \log L}{\partial a \partial \theta} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} \\ -\frac{\partial^2 \log L}{\partial b \partial a} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} & -\frac{\partial^2 \log L}{\partial b^2} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} & -\frac{\partial^2 \log L}{\partial b \partial \theta} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} \\ -\frac{\partial^2 \log L}{\partial \theta \partial a} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} & -\frac{\partial^2 \log L}{\partial \theta \partial b} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} & -\frac{\partial^2 \log L}{\partial \theta^2} \Big|_{\hat{a}, \hat{b}, \hat{\theta}} \end{pmatrix}^{-1} \quad (30)$$

$$= \begin{pmatrix} \text{Var}(\hat{a}) & \text{Cov}(\hat{a}, \hat{b}) & \text{Cov}(\hat{a}, \hat{\theta}) \\ \text{Cov}(\hat{a}, \hat{b}) & \text{Var}(\hat{b}) & \text{Cov}(\hat{b}, \hat{\theta}) \\ \text{Cov}(\hat{\theta}, \hat{a}) & \text{Cov}(\hat{\theta}, \hat{b}) & \text{Var}(\hat{\theta}) \end{pmatrix} \quad (31)$$

The derivatives in I_0 are given as follows

$$\frac{\partial^2 \mathcal{L}}{\partial a^2} = -\sum_{i=1}^n \frac{1}{(a + bx_i)^2} - (\theta - 1) \sum_{i=1}^n \frac{x_i^2 e^{-(ax_i + \frac{b}{2}x_i^2)}}{\left[1 - e^{-(ax_i + \frac{b}{2}x_i^2)}\right]^2}, \quad (32)$$

$$\frac{\partial^2 \mathcal{L}}{\partial a \partial b} = -\sum_{i=1}^n \frac{x_i}{(a + bx_i)^2} - \frac{1}{2}(\theta - 1) \sum_{i=1}^n \frac{x_i^3 e^{-(ax_i + \frac{b}{2}x_i^2)}}{\left[1 - e^{-(ax_i + \frac{b}{2}x_i^2)}\right]^2}, \quad (33)$$

$$\frac{\partial^2 \mathcal{L}}{\partial b^2} = -\sum_{i=1}^n \frac{x_i^2}{(a + bx_i)^2} - \frac{1}{4}(\theta - 1) \sum_{i=1}^n \frac{x_i^4 e^{-(ax_i + \frac{b}{2}x_i^2)}}{\left[1 - e^{-(ax_i + \frac{b}{2}x_i^2)}\right]^2}, \quad (34)$$

$$\frac{\partial^2 \mathcal{L}}{\partial \theta^2} = -\frac{n}{\theta^2}, \quad (35)$$

$$\frac{\partial^2 \mathcal{L}}{\partial a \partial \theta} = \sum_{i=1}^n \frac{x_i e^{-(ax_i + \frac{b}{2}x_i^2)}}{1 - e^{-(ax_i + \frac{b}{2}x_i^2)}}, \quad (36)$$

$$\frac{\partial^2 \mathcal{L}}{\partial b \partial \theta} = \frac{1}{2} \sum_{i=1}^n \frac{x_i^2 e^{-(ax_i + \frac{b}{2}x_i^2)}}{1 - e^{-(ax_i + \frac{b}{2}x_i^2)}}. \quad (37)$$

The above approach is used to derive the $100(1 - \alpha)\%$ confidence intervals of the parameters a, b, θ as in the following forms

$$\hat{a} \pm Z_{\alpha/2} \sqrt{\text{Var}(\hat{a})}, \quad \hat{b} \pm Z_{\alpha/2} \sqrt{\text{Var}(\hat{b})}, \quad \hat{\theta} \pm Z_{\alpha/2} \sqrt{\text{Var}(\hat{\theta})}. \quad (38)$$

Here, $Z_{\alpha/2}$ is the upper $(\alpha/2)$ th percentile of the standard normal distribution.

It should be mentioned here as it was pointed by a referee that if we do not make the assumption that the true parameter vector (a, b, θ) is an interior point of the parameter space then the asymptotic normality results will not hold. If any of the true parameter value is 0, then the asymptotic distribution of the maximum likelihood estimators is a mixture distribution, see for example Self and Liang [18] in this connection. In that case obtaining the asymptotic confidence intervals become quite difficult and it is not pursued here.

6 Data Analysis

In this section we provide a data analysis to see how the new model works in practice. The data have been obtained from Aarset [1] and it is provided below. It represents the lifetimes of 50 devices.

.1	.2	1	1	1	1	1	2	3	6	7	11	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				

We have used different distributions namely; Rayleigh distribution (RD), exponential distribution (ED), generalized exponential distribution (GED), linear failure rate distribution (LFRD), generalized Rayleigh distribution (GRD) and generalized linear failure rate distribution (GLFRD) to analyze the data. The MLE(s) of the unknown parameter(s) and the corresponding Kolmogorov-Smirnov (K-S) test statistic for 6 different models are given in Table 4.

Table 4. The MLE of the parameter(s) and the associated K-S values.

The model	MLE of the parameter(s)	K-S
RD(b)	$\hat{b} = 6.362 \times 10^{-4}$	0.315
ED(a)	$\hat{a} = 0.022$	0.171
GED(a, θ)	$\hat{a} = 0.019, \hat{\theta} = 0.78$	0.184
LFRD(a, b)	$\hat{a} = 0.014, \hat{b} = 2.4 \times 10^{-4}$	0.195
GRD(b, θ)	$\hat{b} = 3.093 \times 10^{-4}, \hat{\theta} = 0.352$	0.181
GLFRD(a, b, θ)	$\hat{a} = 3.822 \times 10^{-3}, \hat{b} = 3.074 \times 10^{-4}, \hat{\theta} = 0.533$	0.162

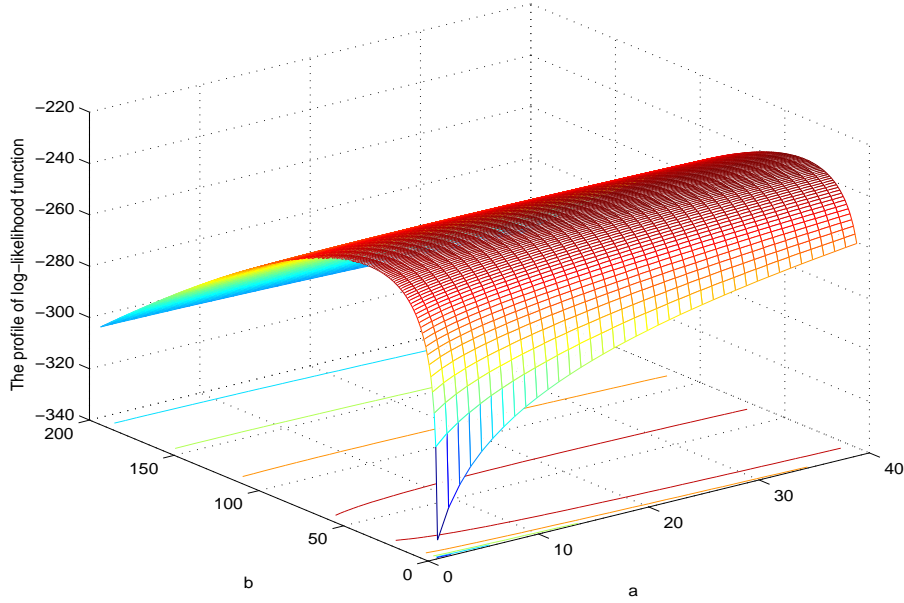


Figure 5. The profile of log-likelihood function for (a, b) .

Table 5. The MLE, the values of log-likelihood function and p-values.

The model	H_0	\mathcal{L}	Λ	d.f.	p-value
RD(β)	$a = 0, \theta = 0$	-264.053	61.816	2	3.775×10^{-14}
Exp(α)	$b = 0, \theta = 0$	-241.090	15.890	2	3.544×10^{-4}
GED(α, θ)	$b = 0$	-239.995	13.700	1	2.145×10^{-4}
LFRD(α, β)	$\theta = 0$	-238.064	09.838	1	1.709×10^{-3}
GRD(β, θ)	$a = 0$	-234.565	02.840	1	0.092

To show that the likelihood equations have a unique solution we plot the profile log-likelihood function of (a, b) in Figure 5.

Since RD(b), ED(a), GED(a, θ), LFRD(a, b) and GRD(b, θ) are special cases of GLFRD(a, b, θ), we perform the following testing of hypotheses; (i) $H_0 : a = 0, \theta = 1$, (ii) $H_0 : b = 0, \theta = 1$, (iii) $H_0 : b = 0$, (iv) $\theta = 1$ (v) $a = 0$. We present the log-likelihood values (\mathcal{L}), the values of the likelihood ratio test statistics (Λ) and the corresponding p -values in the following table. From the p values it is clear that we reject all the hypotheses when the level of significance is 0.1.

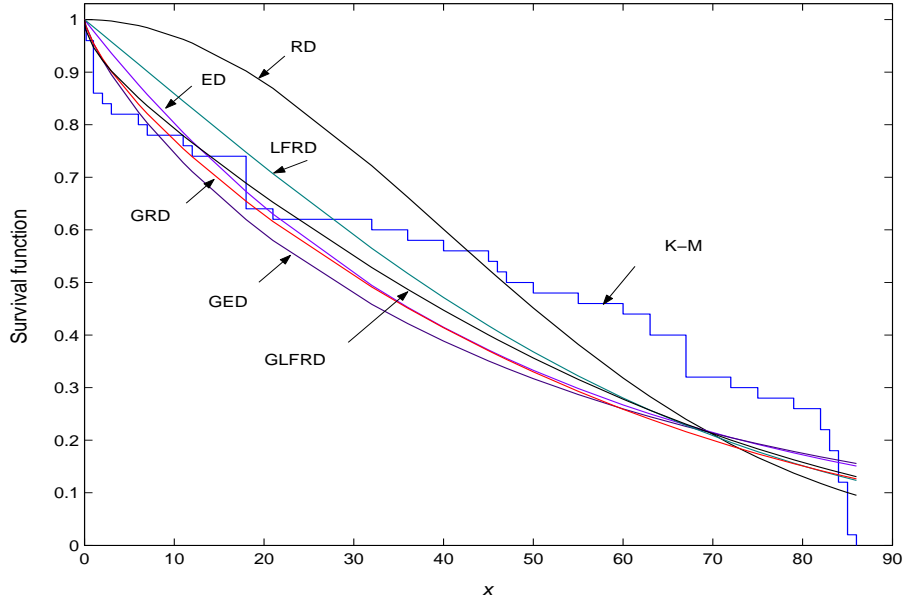


Figure 6. The Kaplan-Meier estimate of survival function and fitted survival functions.

For the comparison purpose, the empirical and fitted total time on test transforms (TTT-Transforms) using the GLFRD, GRD and GED are provided. For details about the TTT-Transform technique, we refer to [3], [5] and [6]. Figure 7 shows the plots of the empirical and fitted TTT-Transform of the data. Further, we computed: (i) the maximum of the distance between the fitted TTT-transform using each distribution and the empirical one; (ii) the mean squared error of the fitted TTT-Transforms using each suggested distribution. Table 6 gives the obtained results.

Table 6: The distance of the TTT-Transform.

Distribution	Maximum distance	MSE
GED	0.3332	0.0348
GRD	0.2286	0.0149
GLFRD	0.1940	0.0112

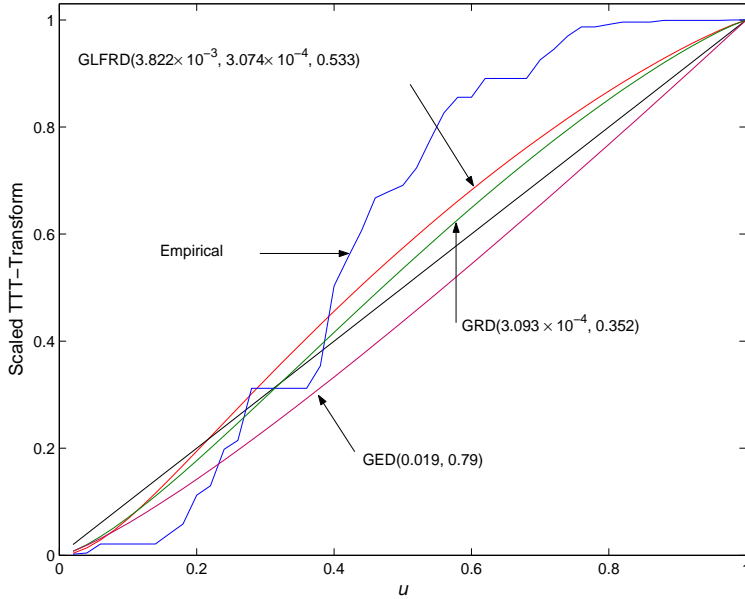


Figure 7. Empirical and fitted scaled TTT-Transform for the data.

Substituting the MLE of the unknown parameters in (30), we get estimation of the variance covariance matrix as

$$I^{-1} = \begin{bmatrix} 9.201 \times 10^{-6} & -7.050 \times 10^{-8} & 2.512 \times 10^{-4} \\ -7.050 \times 10^{-8} & 6.572 \times 10^{-9} & -6.976 \times 10^{-8} \\ 2.512 \times 10^{-4} & -6.976 \times 10^{-8} & 0.013. \end{bmatrix}. \quad (39)$$

The approximate 95% two sided confidence intervals of the parameters a , b and θ are $[0, 9.768 \times 10^{-3}]$, $[1.485 \times 10^{-4}, 4.663 \times 10^{-4}]$ and $[0.308, 0.757]$ respectively.

We have provided the nonparametric estimate of the survival function and the fitted survival functions in Figure 6. It is clear that the generalized linear failure rate distribution provides a good fit to the data set.

7 Conclusions

In this paper we have introduced a new three-parameter generalized linear failure rate distribution and study its different properties. It is observed that the proposed GLFRD has several desirable properties and several existing well known distributions can be obtained as special cases of this distribution. It is observed that the GLFRD can have increasing, decreasing and bathtub shaped hazard rate functions which are quite desirable for data analysis

purposes. We obtained the asymptotic properties of the MLEs of the unknown parameters but we have not studied the behavior of the MLEs for small sample sizes. Extensive simulations are required to study the rate of convergence of the MLEs and their various properties. More work is needed in this direction.

Acknowledgements

The authors would like to thank two referees and the associate editor for many constructive suggestions.

Appendix A

Acronyms:

LFRD:	Linear failure rate distribution
CDF:	Cumulative distribution function
ED:	Exponential distribution
RD:	Rayleigh distribution
PDF:	Probabibility density function
CDF:	Cumulative distribution function
GE:	Generalized exponential distribution
MLEs:	Maximum likelihood estimators
GLFRD:	Generalized failure rate distribution

Appendix B

The Proof of Lemma 3.1. Starting with

$$\mu^{(k)} = \int_0^{\infty} x^k f(x; a, b, \theta) dx$$

then substituting from (4) into the above relation we have

$$\mu^{(k)} = \theta \int_0^\infty x^k (a + bx) \left[1 - e^{-(ax + \frac{b}{2}x^2)} \right]^{\theta-1} e^{-(ax + \frac{b}{2}x^2)} dx.$$

Since $0 < e^{-(ax + \frac{b}{2}x^2)} < 1$ for $x > 0$, by using the binomial series expansion we have

$$\left[1 - e^{-(ax + \frac{b}{2}x^2)} \right]^{\theta-1} = \sum_{i=0}^{\infty} (-1)^i \binom{\theta-1}{i} e^{-i(ax + \frac{b}{2}x^2)}, \quad (40)$$

then

$$\mu^{(k)} = \theta \int_0^\infty x^k (a + bx) \sum_{i=0}^{\infty} (-1)^i \binom{\theta-1}{i} e^{-(i+1)(ax + \frac{b}{2}x^2)} dx. \quad (41)$$

Since the inner quantity of the summation is absolutely integrable, interchanging the integration and summation we get

$$\mu^{(k)} = \theta \sum_{i=0}^{\infty} (-1)^i \binom{\theta-1}{i} \int_0^\infty x^k (a + bx) e^{-(i+1)(ax + \frac{b}{2}x^2)} dx. \quad (42)$$

Now arises two cases. The first case arises when $a = 0$ and $b > 0$. In this case, the integral in (42) becomes

$$\int_0^\infty x^k (a + bx) e^{-(i+1)(ax + \frac{b}{2}x^2)} dx = \int_0^\infty bx^{k+1} e^{-\frac{b}{2}(i+1)x^2} dx = \frac{b}{2} \frac{\Gamma(\frac{k}{2} + 1)}{[\frac{b}{2}(i+1)]^{\frac{k}{2}+1}}. \quad (43)$$

Substituting from (43) into (42), one gets (10) which completes the first part of the lemma. The second case arises when $a > 0$ and $b \geq 0$. For this case, using the Taylor expansion of the function $e^{-\frac{b}{2}(i+1)x^2}$ given by

$$e^{-\frac{b}{2}(i+1)x^2} = \sum_{\ell=0}^{\infty} \frac{g_i^{(\ell)}(0)}{\ell!} x^\ell, \quad g_i^{(\ell)}(0) = \frac{d^\ell}{dx^\ell} e^{-\frac{b}{2}(i+1)x^2} \Big|_{x=0},$$

one can rewrite (42) as

$$\begin{aligned} \mu^{(k)} &= \theta \sum_{i=0}^{\infty} \sum_{\ell=0}^{\infty} (-1)^i \binom{\theta-1}{i} \frac{g_i^{(\ell)}(0)}{\ell!} \int_0^\infty x^{k+\ell} (a + bx) e^{-(i+1)ax} dx \\ &= \theta \sum_{i=0}^{\infty} \sum_{\ell=0}^{\infty} (-1)^i \binom{\theta-1}{i} \frac{g_i^{(\ell)}(0)}{\ell!} \left[\frac{a \Gamma(k + \ell + 1)}{[(i+1)a]^{k+\ell+1}} + \frac{b \Gamma(k + \ell + 2)}{[(i+1)a]^{k+\ell+2}} \right]. \end{aligned}$$

That completes the proof of the lemma.

The Proof of Theorem 4.1. The general definition of the k^{th} moment of order statistic $X_{(r)}$, is

$$\mu_{r:n}^{(k)} = \int_0^{\infty} x^k f_{r:n}(x) dx. \quad (44)$$

Substituting from (18) into (44), one gets

$$\mu_{r:n}^{(k)} = \sum_{j=0}^{n-r} d_j(n, r) \int_0^{\infty} x^k f(x; a, b, \theta_{r+j}) dx. \quad (45)$$

Since the integral in (45) is the k^{th} moment of $\text{GLFRD}(a, b, \theta_{r+j})$, then from (45) with the help of lemma (3.1) we get (19) and (20) which completes the proof.

Appendix C

Proof of Theorem 3.1

Necessary Part: If Y follows $\text{GLFRD}(a, b, \theta)$, then (14) follows using integration by parts from the expression of $\alpha^{(n)}(t)$.

Sufficient Part: Suppose (14) is true for the random variable Y . Let us denote the PDF and CDF of Y as $f_Y(\cdot)$ and $F_Y(\cdot)$ respectively. From (14) we obtain

$$\int_0^t U^n(t) f_Y(y) dy = U^n(t) F_Y(t) + \frac{n}{\theta} \int_0^t U^{n-1}(t) f_Y(y) dy. \quad (46)$$

Differentiating (46) with respect to t , we obtain

$$U^n(t) f_Y(t) = U^n(t) f_Y(t) - n U^{n-1}(t) F_Y(t) \frac{f(t; a, b, 1)}{F(t; a, b, 1)} + \frac{n}{\theta} U^{n-1}(t) f_Y(t). \quad (47)$$

Since $U^{n-1}(t) \neq 0$, from (47) we obtain

$$\frac{f_Y(t)}{F_Y(t)} = \theta \frac{f(t; a, b, 1)}{F(t; a, b, 1)}. \quad (48)$$

From the uniqueness property of the reversed hazard function and from (6), it is immediate that $f_Y(x)$ has to be of the form (4).

Proof of Theorem 3.2

Necessary Part: Suppose Y follows $\text{GLFRD}(a, b, \theta)$, then from Theorem 3.1, we get

$$\alpha^{(1)}(t) = U(t) + \frac{1}{\theta} \quad (49)$$

and

$$\alpha^{(2)}(t) = U^2(t) + \frac{2}{\theta}\alpha^{(1)}(t) = U^2(t) + \frac{2}{\theta} \left[U(t) + \frac{1}{\theta} \right] = (\alpha^{(1)}(t))^2 + \frac{1}{\theta^2}. \quad (50)$$

Therefore, the result follows.

Sufficient Part: Suppose (15) is true for all $t > 0$ for the random variable Y . Let us denote the PDF and CDF of Y as $f_Y(\cdot)$ and $F_Y(\cdot)$ respectively. Therefore, (15) can be written as

$$F_Y(t) \int_0^t U^2(y) f_Y(y) dy - \left[\int_0^t U(y) f_Y(y) dy \right]^2 = \frac{F_Y^2(t)}{\theta^2}. \quad (51)$$

Differentiating both sides of (51) with respect to t and cancelling $f_Y(t)$ from both sides, we get

$$\int_0^t U^2(y) f_Y(y) dy + F_Y(t)U^2(t) - 2U(t) \int_0^t U(y) f_Y(y) dy = 2\frac{F_Y(t)}{\theta^2}. \quad (52)$$

Differentiating both sides of (52) with respect to t and arranging terms, we obtain;

$$F_Y(t)U(t) - \int_0^t U(y) f_Y(y) dy = -\frac{f_Y(t)}{r(t; a, b, 1)\theta^2}. \quad (53)$$

After differentiating (53) with respect to t and rearranging we obtain the differential equation;

$$\theta^2 F_Y(t) = \frac{1}{r(t; a, b, 1)} \times \frac{d}{dt} \left(\frac{f_Y(t)}{r(t; a, b, 1)} \right). \quad (54)$$

A general solution of the differential equation (54) is of the form

$$F_Y(t) = c_1 e^{\alpha U(t)} + c_2 e^{-\alpha U(t)},$$

where c_1 and c_2 are arbitrary constants. Since $\lim_{t \downarrow 0} U(t) = \infty$ and $\lim_{t \rightarrow \infty} U(t) = 0$ and $F_Y(t)$ is a distribution function, then $c_1 = 0$ and $c_2 = 1$. Therefore, the result follows.

References

- [1] M. V. Aarset, "How to identify bathtub hazard rate," *IEEE Transactions on Reliability*, vol. R-36, 106 - 108, 1987.

- [2] L. J. Bain, "Analysis for the Linear Failure-Rate Life-Testing Distribution," *Technometrics*, 16(4), 551 - 559, 1974.
- [3] R. E. Barlow, R. Campo, "Total time on test processes and applications to failure data analysis", *Reliability and Fault Tree Analysis* (Barlow, Fussell, Singpurwalla, Eds), 451 - 481; SIAM, Philadelphia, 1975.
- [4] R.E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing. Probability Models.* Holt, Rinehart & Winston, New York, 1975.
- [5] B. Bergman, B. Klefsjo, "A graphical method applicable to agereplacement problems", *IEEE Transactions on Reliability*, vol R-31, 478 - 481, 1982.
- [6] B. Bergman, B. Klefsjo, "The total time on test concept and its use in reliability theory", *Operations Research*, vol 32, 596 - 606, 1984.
- [7] I.W. Burr, "Cumulative frequency function", *Annals of Mathematical Statistics*, 13, 215 - 232, 1942.
- [8] B.V. Genedenko, Y.K. Belyayev and A.D. Solovyev, *Mathematical Methods of Reliability Theory*, Academic Press, New York, 1969.
- [9] M.E. Ghitany, "Reliability properties of extended linear failure-rate distributions", *Probability in the Engineering and Information Sciences*, vol. 21, 441 - 450, 2006.
- [10] R.C. Gupta and R.D. Gupta, "Proportional reversed hazard model and its applications", *Journal of Statistical Planning and Inference*, vol. 137, no. 11, 3525 - 3536, 2007.
- [11] R.D. Gupta and D. Kundu, "Generalized exponential distribution", *Austral. & New Zealand Journal of Statistics*, vol. 41, 173 - 188, 1999.
- [12] D. Kundu and R.D. Gupta, "Characterization of the proportional (reversed) hazard model", *Communications in Statistics - Theory and Methods*, vol. 33, 3095 - 3102, 2004.

- [13] D. Kundu and M. Raqab, "Generalized Rayleigh distribution: different methods of estimations", *Computational Statistics & Data Analysis*, vol. 49, 187 - 200, 2005.
- [14] C. D. Lai, M. Xie, and D. N. P. Murthy, "Bathtub shaped failure rate distributions," *in Handbook in Reliability*, N. Balakrishnan and C. R. Rao, Eds., 20, 69 - 104, 2001.
- [15] J. F. Lawless, "Statistical Models and Methods for Lifetime Data," *John Wiley and Sons, New York*, 2003.
- [16] C.T. Lin, S.J.S. Wu and N. Balakrishnan, "Monte Carlo methods for Bayesian inference on the linear hazard rate distribution", *Communications in Statistics - Theory and Methods*, vol. 35, 575 - 590, 2006.
- [17] M.Z. Raqab, "Order statistics from the Burr type X model", *Computers Mathematics and Applications*, vol., 36, 111 - 120, 1998.
- [18] S.G. Self and K-Y Liang, "Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions", *Journal of the Americal Statistical Association*, vol. 82, 605 - 610, 1987.
- [19] A. Sen and G.K. Bhattacharya, "Inference procedure for the linear failure rate model", *Journal of Statistical Planning and Inference*,, vol. 46, 59 - 76, 1995.
- [20] J.G. Surles and W.J.Padgett, "Some properties of a scaled Burr type X distribution", *Journal of Statistical Planning and Inference*, 128, 271 - 280, 2005.