Statistics and Applications {ISSN 2454-7395 (online)}
Volume 21, No. 2, 2023 (New Series), pp 193–211
http://www.ssca.org.in/journal

The gLinear Failure Rate Distribution: A New Mixture with Bayesian and Non-Bayesian Analysis

R. M. Mandouh[1]
[1]*Department of Mathematical Statistics*
Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt

Received: 02 September 2022; Revised: 27 November 2022; Accepted: 13 March 2023

Abstract

A mixture of the gamma and linear failure rate distributions is constructed and studied. The model parameters are estimated using maximum likelihood and Bayesian based on real and simulated data.

Key words: Mixture Distribution; Maximum Likelihood Estimation; Laplace Approximation; MCMC; Variational Bayes.

AMS Subject Classifications: 62K05, 05B05

1. Introduction

In statistical analysis, the survival function and hazard function are used to model distribution of data representing lifetime or waiting time. The survival function or reliability function is the probability of survival of an item without failing until time t. Alternatively, we can describe the survival experience in term of hazard failure (instantaneous rate of death) which is the chance of death (failure) as a function of age. The hazard function or the instantaneous failure rate has many types which appeared in practice such as unimodal shaped; bathtub shaped and others. The main aim of this paper is to introduce a new distribution with two parameters. The hazard function of this distribution can be constant, unimodal (upside-down bathtub) or increasing-decreasing-increasing depending on the values of its two parameters. The shapes of the hazard function of the new distribution enables it to be a good model to fit various data sets.

The mixture distribution (Everitt (2013)) is one of the means can be utilized to construct these new distributions. The finite mixture is formed as follow:

$$f(x) = \sum_{i=1}^{c} p_i f_i(x)$$

where $\sum_{i=1}^{c} p_i = 1$ with $c = 2$ in our distribution. Many Papers dealing with two mixture models such as, Lindley (1958) introduced a one parameter distribution, now known as the

Corresponding Author: R. M. Mandouh
Email: rshmndoh@gmail.com, rshmndoh@cu.edu.eg

Lindley distribution. Ghitany *et al.* (2008) studied its properties in details. Shanker and Mishra (2013) added one extra parameter to Lindley distribution and introduced the quasi Lindley distribution. They studied some of its properties. Sen *et al.* (2016) proposed and studied another finite mixture distribution which is called the xgamma distribution.Sen and Chandra (2017) added one extra parameter to the xgamma distribution and introduced the quasi xgamma distribution. Moreover, many Papers dealing with three mixture models such as, Sarhan *et al.* (2014) introduced two lifetime distributions. They referred to these two distributions as $N(\beta)$ and $TN(\alpha, \beta)$ respectively and they discussed some properties of these two distribution such as the behavior of their hazard functions. Mahmoud *et al.* (2017) introduced two distributions based on mixing between different types of distributions.

2. The gLinear failure rate distribution

Now, we introduce a mixture density of two mixture components, one follows gamma (2, β) and the other follows linear failure rate (β, β^2) with mixing weights $\frac{\beta}{\alpha+\beta}$ and $\frac{\alpha}{\alpha+\beta}$. The pdf of the new mixture distribution will be as follows:

$$f(x) = \frac{\beta}{\alpha + \beta}(\beta^2 x + \alpha(1 + \beta x)e^{-\frac{\beta^2}{2}x^2})e^{-\beta x}, \quad x > 0,\ \beta, \alpha > 0. \tag{1}$$

We refer to this distribution as glfr (α, β). For $\alpha = 1$, we have the following new distribution as a special case

$$f(x) = \frac{\beta}{1 + \beta}(\beta^2 x + (1 + \beta x)e^{-\frac{\beta^2}{2}x^2})e^{-\beta x}, \quad x > 0,\ \beta > 0, \tag{2}$$

which is a mixture of gamma (2, β) and the other follows linear failure rate (β, β^2) with mixing weights $\frac{\beta}{1+\beta}$ and $\frac{1}{1+\beta}$ and we refer to this distribution as glfr (β). Figure (1) shows pdf of the glfr distribution for different parameter values. The corresponding cdf of (2.1) takes the following form

$$F(x) = \frac{1}{\alpha + \beta}(\beta + \alpha - e^{-\beta x}(\beta(1 + \beta x) + \alpha e^{-\frac{\beta^2}{2}x^2})), \quad x > 0,\ \beta, \alpha > 0. \tag{3}$$

Then the survival function is given by

$$S(x) = \frac{1}{\alpha + \beta}(e^{-\beta x}(\beta(1 + \beta x) + \alpha e^{-\frac{\beta^2}{2}x^2})), \quad x > 0,\ \beta, \alpha > 0, \tag{4}$$

and the hazard function is given by

$$h(x) = \frac{\beta(\beta^2 x + \alpha(1 + \beta x)e^{-\frac{\beta^2}{2}x^2})}{(\beta(1 + \beta x) + \alpha e^{-\frac{\beta^2}{2}x^2})}, \quad x > 0,\ \beta, \alpha > 0, \tag{5}$$

One can note that $h(x)$ is bounded, *i.e.* $\frac{\alpha\beta}{\alpha+\beta} < h(x) < \beta$. The hazard function of glfr distribution is plotted in Figure (2) for four different pairs of choices of α and β.

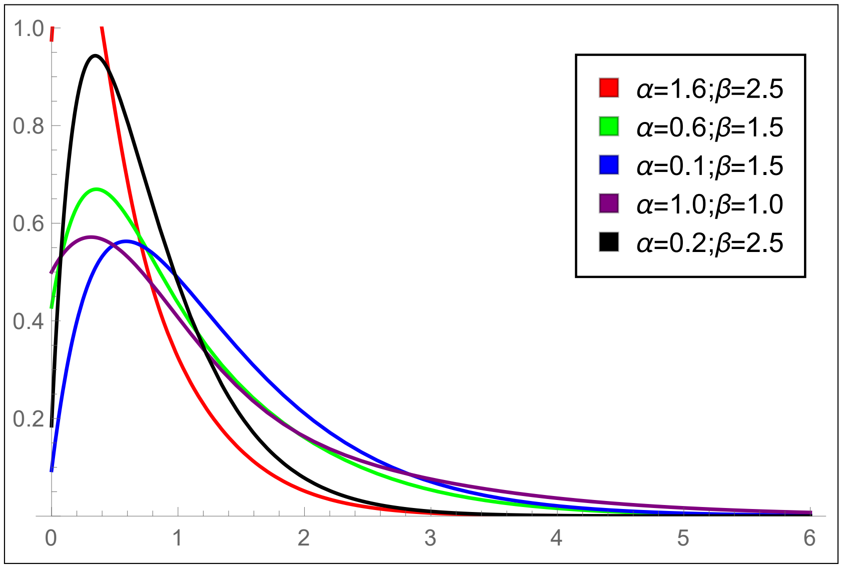

Figure 1: The gLinear failure rate pdfs for some parameter values

The moments and shape measures

Let X follow gLinear failure rate distribution. After some algebra, the rth moment of X is derived as

$$E(X^r) = \frac{\Gamma(r+2)}{\beta^{r-1}(\alpha+\beta)} + \frac{2\alpha\sqrt{e}}{\beta^r(\alpha+\beta)} \int_{1/\sqrt{2}}^{\infty} t(\sqrt{2}t-1)^r e^{-t^2} dt \tag{6}$$

Therefore, the expectation variance of the two parameter glfr distribution in terms of the error function (erf) and its complementary (erfc) are given by

$$E(X) = \frac{4\beta + \alpha\sqrt{2e\pi}(erfc(\frac{1}{\sqrt{2}}))}{2\beta(\alpha+\beta)},$$

and

$$Var(X) = \frac{6\beta + 2\alpha - \alpha\sqrt{2e\pi}(erfc(\frac{1}{\sqrt{2}}))}{\beta^2(\alpha+\beta)} - (\frac{4\beta + \alpha\sqrt{2e\pi}(erfc(\frac{1}{\sqrt{2}}))}{2\beta(\alpha+\beta)})^2,$$

where, $erfc(z) = 1 - erf(z)$ and $erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$
Also, one can use eq.(6) and the relation between the moments and the central moments to obtain skewness and kurtosis.

The mean residual life

One of special relevance in reliability and survival analysis is the analysis of the lifetime of a device after it has attained age x. Thus, if X is the lifetime with survival function given by (4), the corresponding residual lifetime after age x is the random variable $X_x = (X - x|X > x)$ and the mean residual life of X is defined as $m(x) = E(X - x|X > x)$. It is also called the expected additional lifetime given that a component has survived until time t is a function of t

$$\begin{aligned} m(x) &= E(X-x|X>x) \\ &= \frac{1}{S(x)}\int_x^{\infty} S(t)dt \\ &= \frac{\int_x^{\infty} e^{-\beta t}(\beta(1+\beta t)+\alpha e^{-\frac{\beta^2}{2}t^2})dt}{e^{-\beta x}(\beta(1+\beta x)+\alpha e^{-\frac{\beta^2}{2}x^2})} \\ &= \frac{\beta(2+\beta x)e^{-\beta x}+\sqrt{\frac{e\pi}{2}}erfc(\frac{1+\beta x}{\sqrt{2}})}{\beta(e^{-\beta x}(\beta(1+\beta x)+\alpha e^{-\frac{\beta^2}{2}x^2}))}, \end{aligned}$$

where, $erfc(z) = 1 - erf(z)$ and $erf(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt$.

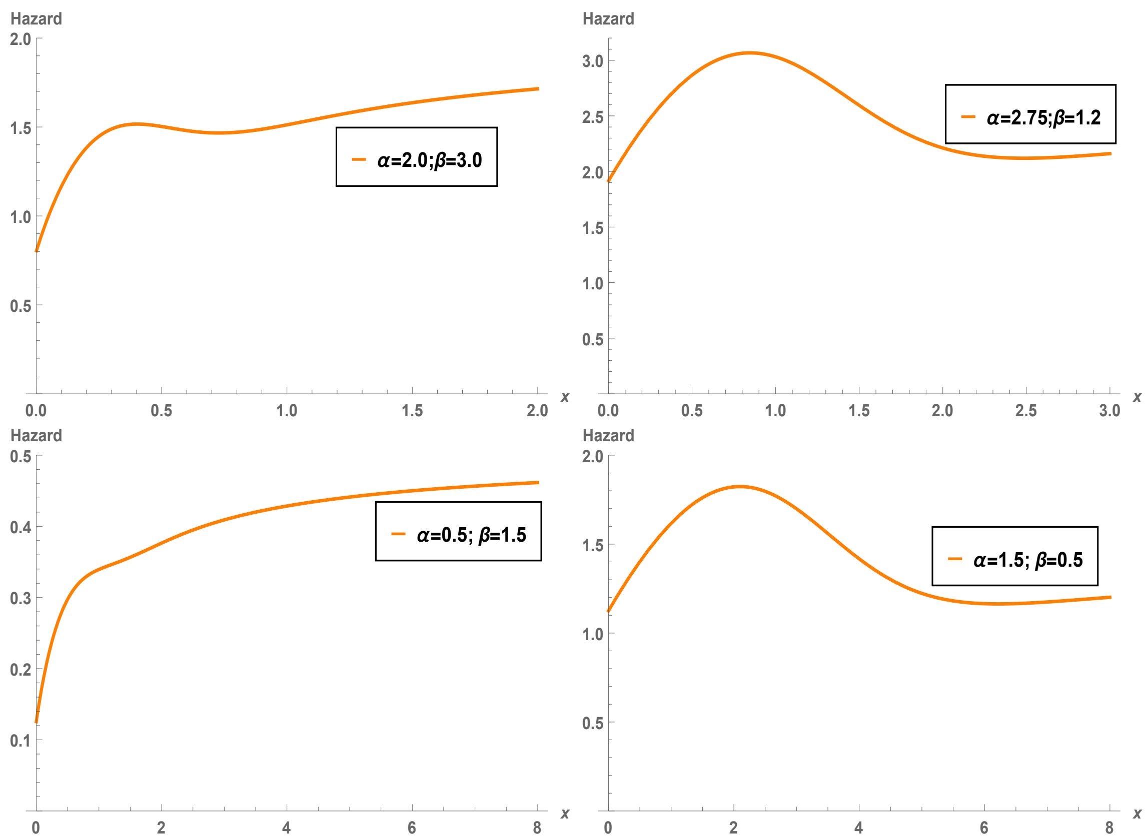

Figure 2: The hazard rate function of the gLinear failure rate for some parameter values

3. Maximum likelihood estimation (MLE)

For different statistical models, MLE is widely utilized to estimate the model parameters. Assume that n independent and identical items are put on a life test simultaneously. The lifetimes of these items are assume to have follow glinear failure rate distribution. Let $\mathbf{x} = (x_1, x_2, ..., x_n)$ be the failure times of the items. The Likelihood function for α, β is given by

$$L(\alpha,\beta;\mathbf{x}) = \frac{\beta^n}{(\alpha+\beta)^n}\prod_{i=1}^{n}(\beta^2 x_i + \alpha(1+\beta x_i)e^{-\frac{\beta^2}{2}x_i^2})e^{-\beta x_i} \tag{7}$$

The log-likelihood function is

$$\mathcal{L} = \mathcal{L}(\alpha,\beta;\mathbf{x}) = nln\beta - nln(\alpha+\beta) - \beta\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} ln\mathcal{A}_i(\alpha,\beta) \tag{8}$$

where $\mathcal{A}_i(\alpha,\beta) = \beta^2 x_i + \alpha(1+\beta x_i)e^{-\frac{\beta^2}{2}x_i^2}, \quad i = 1,2,...,n.$

Taking partial derivatives of the log-likelihood in (8) w.r.t. α and β, we have

$$\mathcal{L}_\alpha = -\frac{n}{\alpha+\beta} + \sum_{i=1}^{n}\frac{\mathcal{A}_{i,\alpha}(\alpha,\beta)}{\mathcal{A}_i(\alpha,\beta)} \tag{9}$$

$$\mathcal{L}_\beta = \frac{n}{\beta} - \frac{n}{\alpha+\beta} - \sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\frac{\mathcal{A}_{i,\beta}(\alpha,\beta)}{\mathcal{A}_i(\alpha,\beta)} \tag{10}$$

where

$$\begin{aligned}
\mathcal{A}_{i,\alpha}(\alpha,\beta) &= \frac{\partial\mathcal{A}_i(\alpha,\beta)}{\partial\alpha} = (1+\beta x_i)e^{-\frac{\beta^2}{2}x_i^2}, \\
\mathcal{A}_{i,\beta}(\alpha,\beta) &= \frac{\partial\mathcal{A}_i(\alpha,\beta)}{\partial\beta} = 2\beta x_i + \alpha x_i e^{-\frac{\beta^2}{2}x_i^2} - \alpha\beta(1+\beta x_i)x_i^2 e^{-\frac{\beta^2}{2}x_i^2}.
\end{aligned}$$

The second derivative of the log-likelihood are

$$\begin{aligned}
\mathcal{L}_{\alpha,\alpha} &= \frac{n}{(\alpha+\beta)^2} + \sum_{i=1}^{n}\frac{\mathcal{A}_i(\alpha,\beta)\mathcal{A}_{i,\alpha^2}(\alpha,\beta) - (\mathcal{A}_{i,\alpha}(\alpha,\beta))^2}{(\mathcal{A}_i(\alpha,\beta))^2} \\
\mathcal{L}_{\alpha,\beta} &= \frac{n}{(\alpha+\beta)^2} + \sum_{i=1}^{n}\frac{\mathcal{A}_i(\alpha,\beta)\mathcal{A}_{i,\alpha\beta}(\alpha,\beta) - (\mathcal{A}_{i,\alpha}(\alpha,\beta))(\mathcal{A}_{i,\beta}(\alpha,\beta))}{(\mathcal{A}_i(\alpha,\beta))^2} \\
\mathcal{L}_{\beta,\beta} &= -\frac{n}{\beta^2} + \frac{n}{(\alpha+\beta)^2} + \sum_{i=1}^{n}\frac{\mathcal{A}_i(\alpha,\beta)\mathcal{A}_{i,\beta^2}(\alpha,\beta) - (\mathcal{A}_{i,\beta}(\alpha,\beta))^2}{(\mathcal{A}_i(\alpha,\beta))^2}
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
\mathcal{A}_{i,\alpha^2}(\alpha,\beta) &= 0, \\
\mathcal{A}_{i,\alpha\beta}(\alpha,\beta) &= x_i e^{-\frac{\beta^2}{2}x_i^2}(1-\beta x_i(1+\beta x_i)), \\
\mathcal{A}_{i,\beta^2}(\alpha,\beta) &= 2x_i + \alpha x_i e^{-\frac{\beta^2}{2}x_i^2}(-\beta x_i - 3\beta x_i^2 + \beta^2 x_i^3(1+\beta x_i)).
\end{aligned}$$

To calculate the information matrix, the expectation of the following matrix is required

$$\mathcal{T}(\alpha,\beta) = -\begin{bmatrix}\mathcal{L}_{\alpha,\alpha} & \mathcal{L}_{\alpha,\beta} \\ \mathcal{L}_{\alpha,\beta} & \mathcal{L}_{\beta,\beta}\end{bmatrix}$$

Equating the derivatives in (9) and (10) to zero and solving them numerically to obtain the mle of α and β, say $\hat{\alpha}$ and $\hat{\beta}$ such that $\mathcal{T}(\hat{\alpha},\hat{\beta})$ is positive definite.

For Interval estimation of (α, β), the mle of parameters α and β are asymptotically normally distributed with means equal the true values of α and β and variances given by the inverse of the observed information matrix,$\mathcal{T}(\hat{\alpha}, \hat{\beta})$, i.e.

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} \sim N_2 \left[\begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \hat{\mathcal{T}}^{-1} \right] \tag{12}$$

where $\hat{\mathcal{T}}^{-1}$ is the inverse of $\mathcal{T}(\hat{\alpha}, \hat{\beta})$. Using (12), large sample $(1-\nu)100\%$ confidence intervals for α and β are $\hat{\alpha} \pm z_{\nu/2}\sqrt{var(\hat{\alpha})}$, $\quad \hat{\beta} \pm z_{\nu/2}\sqrt{var(\hat{\beta})}$, where $z_{\nu/2}$ is the upper $100\nu/2$ quantile of the standard normal distribution and $var(\hat{\alpha})$ and $var(\hat{\beta})$ are the main diagonal of $\hat{\mathcal{T}}^{-1}$.

4. Bayesian estimation

Let $x_1, x_2, ..., x_n$ be a random sample from glinear failure rate distribution. The likelihood of this sample is given by (7). Let the two parameters α and β are independent random variables with prior distributions gamma(a_1, b_1) and gamma(a_2, b_2), respectively. That is, the joint prior density of α and β is

$$g_0(\alpha, \beta) \propto \alpha^{a_1-1}\beta^{a_2-1}e^{-b_1\alpha+-b_2\beta}, \quad \alpha, \beta > 0 \tag{13}$$

where the hyperparameters a_i and $b_i, i = 1, 2.$ are assumed to be positive and known. Using the likelihood function (7) and the joint prior density function (13) and applying Bayes' theorem, we get the joint posterior density function of (α, β), given the data, as

$$g(\alpha, \beta|\boldsymbol{x}) \propto \frac{\alpha^{a_1-1}\beta^{a_2+n-1}}{(\alpha+\beta)^n}e^{-b_1\alpha+-b_2\beta}\prod_{i=1}^{n}(\beta^2 x_i + \alpha(1+\beta x_i)e^{-\frac{\beta^2}{2}x_i^2})e^{-\beta x_i}, \quad \alpha, \beta > 0 \tag{14}$$

Bayes estimators of the unknown parameters of any function of the unknown parameters, say $h(\boldsymbol{\theta})$, can be obtained as follows

$$E(h(\boldsymbol{\theta})|\boldsymbol{x}) = \frac{\int_0^\infty \int_0^\infty h(\boldsymbol{\theta})g_0(\alpha, \beta)exp(\mathcal{L})d\alpha d\beta}{\int_0^\infty \int_0^\infty g_0(\alpha, \beta)exp(\mathcal{L})d\alpha d\beta}, \tag{15}$$

Formula (15) involves a ratio of two multidimentional integrals and does not have analytical solution. Thus, some approximation methods were suggested to approximate these integrals and calculate the ratio of the integrals such as the methods discussed by Lindley (1958) and Tierney and Kadane (1986) . These methods work well for low dimensions. In this paper we will use Tierney and Kadane's approximation method. They approximate (15) by using Laplace method as follow

$$E(h(\boldsymbol{\theta})|\boldsymbol{x}) = \left(\frac{\det\boldsymbol{\Sigma}^*}{\det\boldsymbol{\Sigma}}\right)^{1/2} exp(n(\mathcal{L}(\hat{\boldsymbol{\theta}}^*) - \mathcal{L}(\hat{\boldsymbol{\theta}}))) \tag{16}$$

where $n\mathcal{L}(\hat{\boldsymbol{\theta}}^*) = lnh + lng_0 + \mathcal{L}$, $n\mathcal{L}(\hat{\boldsymbol{\theta}}) = lng_0 + \mathcal{L}$ and $\boldsymbol{\Sigma}^*$ and $\boldsymbol{\Sigma}$ are minus the inverse Hessian of $\mathcal{L}(\hat{\boldsymbol{\theta}}^*)$ and $\mathcal{L}(\hat{\boldsymbol{\theta}})$ evaluated at $\boldsymbol{\theta}^*$ and $\boldsymbol{\theta}$, respectively. For more details about Laplace approximation see Crawford (1994) and Tierney *et al.* (1989).

For many applications, Bayesian inference is performed using Markov Chain Monte Carlo (MCMC), which estimates expectations w.r.t. $g(\boldsymbol{\theta}|\boldsymbol{x})$ by sampling from it. One of MCMC, Metropolis-Hastings (MH) algorithm, is proposed here. MH algorithm requires a proposal distribution and a common choice of it is the multivariate normal distribution. Metropolis-Hastings algorithm steps are

1. Specify the size of the random draws, say m.
2. Choose an initial value of $\boldsymbol{\theta}$, say $\boldsymbol{\theta}^{(0)}$.
3. For $i = 1, 2, \ldots, m$, repeat the following steps:
   (a) Set $\theta^{(i)} = \theta^{(i-1)}$.
   (b) Generate a candidate value θ^* from a proposal distribution $p(\theta^{(*)}|\theta^{(i)})$.
   (c) Compute the ratio $\kappa = min(1, \frac{g(\boldsymbol{\theta}^{(*)}|data)/p(\boldsymbol{\theta}^{(*)}|\boldsymbol{\theta}^{(i)}}{g(\boldsymbol{\theta}^{(i)}|data)/p(\boldsymbol{\theta}^{(i)}|\boldsymbol{\theta}^{(*)})})$.
   (d) Generate a random value u from uniform distribution on (0, 1).
   (e) Put $\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^*, if \kappa \geq u,$ otherwise put $\theta^{(i)} = \theta^{(i-1)}$.
4. Return the values $\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(m)}$.

For more details about MH algorithm see Puza (2015). For other applications where θ is high dimensional or fast computation is of primary interest, variational Bayesian (VB) is an attractive alternative to MCMC. Yamaguchi *et al.* (2010) developed a VB approach for approximately computing posterior distributions of parameters of mixture of Erlang distribution and they investigated that computation speed of the VB becomes up to 200 times faster than that of MCMC. VB approximates the posterior distribution by a probability distribution with density $q(\boldsymbol{\theta})$ belonging to some tractable family of distributions Q such as Gaussians. The VB method treats an optimization to minimize the Kullback–Leibler (KL) divergence from an approximate posterior distribution to the exact posterior distribution, i.e. The best VB approximation $q^* \in Q$ is

$$q^* = \underset{q \in Q}{argmin} \left\{ KL(q||g(., \boldsymbol{x})) := \int q(\boldsymbol{\theta}) log \frac{q(\boldsymbol{\theta})}{g(\boldsymbol{\theta}|\boldsymbol{x})} d\boldsymbol{\theta} \right\}.$$

5. Simulation study

A simulation study was carried out to investigate the performance of the accuracy of point and interval estimates of the two parameters of the glfr(α, β) distribution. The following steps are carried out:

1. Specify the values of the parameters α and β.
2. Specify the sample size n.
3. Generate a random sample $(x_1, x_2, x_3, ..., x_n)$ with size n from glfr(α, β) distribution using the following algorithm:

- Generate $U \sim$ uniform$(0,1)$ with size n.
- Generate $V \sim$ gamma $(2,\beta)$ with size n.
- Generate $W \sim$ linear failure rate (β,β^2) with size n.
- If $u \leq \beta/(\alpha+\beta)$ set $x=u$, otherwise set $u=w$

4. Calculate the mle of the two parameters.

5. Repeat steps 2-4, N times.

6. Calculate the mean squared error (MSE), the average of the confidence interval widths, and the coverage probability for each parameter. The MSE associated with the MLE of the parameter θ, MSE_θ, is

$$\mathrm{MSE}_\theta = \frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}-\theta)^2,$$

where $\hat{\theta}$ is the MLE of θ. Coverage probability is the proportion of the N simulated confidence intervals which include the true parameter θ.

The simulation study is carried out using $N = 1000$. The sample sizes are 50, 75, 100, 150 and 200 and the selected parameter values are (α,β)= (0.8, 0.8), (0.8, 1.0), (1.0, 1.0), (1.0, 1.2), (1.2, 1.2) and (1.6, 2.5). Table 1 presents the MSE, coverage probability (CP_θ) and average width (AW_θ) of 95% confidence intervals of each parameter. This table shows that ,in the most cases, the MSEs and the average widths decrease as the sample size increases and the coverage probability are close to the nominal level of 95%.

6. Applications

In this section, to illustrate the applicability of the two new distributions proposed in this paper, we analyze three data sets. The first data set represents the remission times (in months) of a random sample of 128 bladder cancer patients. Bladder cancer is a disease in which abnormal cells multiply without control in the bladder. The most common type of bladder cancer recapitulates the normal histology of the urothelium and is known as transitional cell carcinoma.This data were studied by Zea *et al.* (2012).The second data represents a complete data with the exact times of failure. This data is considered a data set of the life of fatigue fracture of Kevlar 373/epoxy that are subject to constant pressure at the 90% stress level until all had failed. This data is considered by Ogunde *et al.* (2017). The three data set is provided in Murthy *et al.* (2004), page 278, about time between failures for repairable item.

We will refer to these data sets as data set 1, data set 2 and data set 3, respectively. For each data set, we fit the proposed distributions and other distributions such as the quasi xgamma (qxgamma), xgamma, quasi Lindley (qLinley), Lindley, linear failure rate (lfr) and gamma distributions. Goodness-of-fit tests are applied to verify which distribution better fits these data sets. The tests were carried out at 5% level of significance. We consider the common-known Kolmogorov-Smirnov (K-S) statistic, the Anderson-Darling (A-D), and Cramér-von Mises (C-M) statistics. Moreover, we consider some well-known measures such as the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the consistent Akaike information criterion (CAIC) and the Hannan-Quinn information criterion (HQIC). These criterions are defined by:

$$\begin{aligned}
\text{AIC} &= -2\mathcal{L}(\hat{\boldsymbol{\theta}}) + 2p; \\
\text{BIC} &= -2\mathcal{L}(\hat{\boldsymbol{\theta}}) + plog(n); \\
\text{CAIC} &= -2\mathcal{L}(\hat{\boldsymbol{\theta}}) + \frac{2pn}{n-p-1}; \\
\text{HQIC} &= -2\mathcal{L}(\hat{\boldsymbol{\theta}}) + 2log(log(n)).
\end{aligned}$$

where $\mathcal{L}(\hat{\boldsymbol{\theta}})$ denotes the log-likelihood function evaluated at the maximum likelihood estimates for parameters $\boldsymbol{\theta}$, p is the number of parameters and n is the sample size.Table 2 shows the MLE of the parameters of each model, the corresponding maximum log-likelihood value, the AIC, BIC, CAIC and HQIC for the three data sets. Table 3 presents the values of the statistics K-S, (A-D) (A*) and C-M (W*) for the three data sets using each model. The required numerical evaluations are carried out using R software.

For the first two data sets, glfr model has the smallest value of the Kolmogorov-Smirnov (largest P value), Anderson-Darling and the Cramér-von Mises gooness-of-fit tests statistics which indicate that the best fit is provided by glfr model for these data sets. For the third data set, gamma model is a better fit than glfr (α, β) model, see Table 3

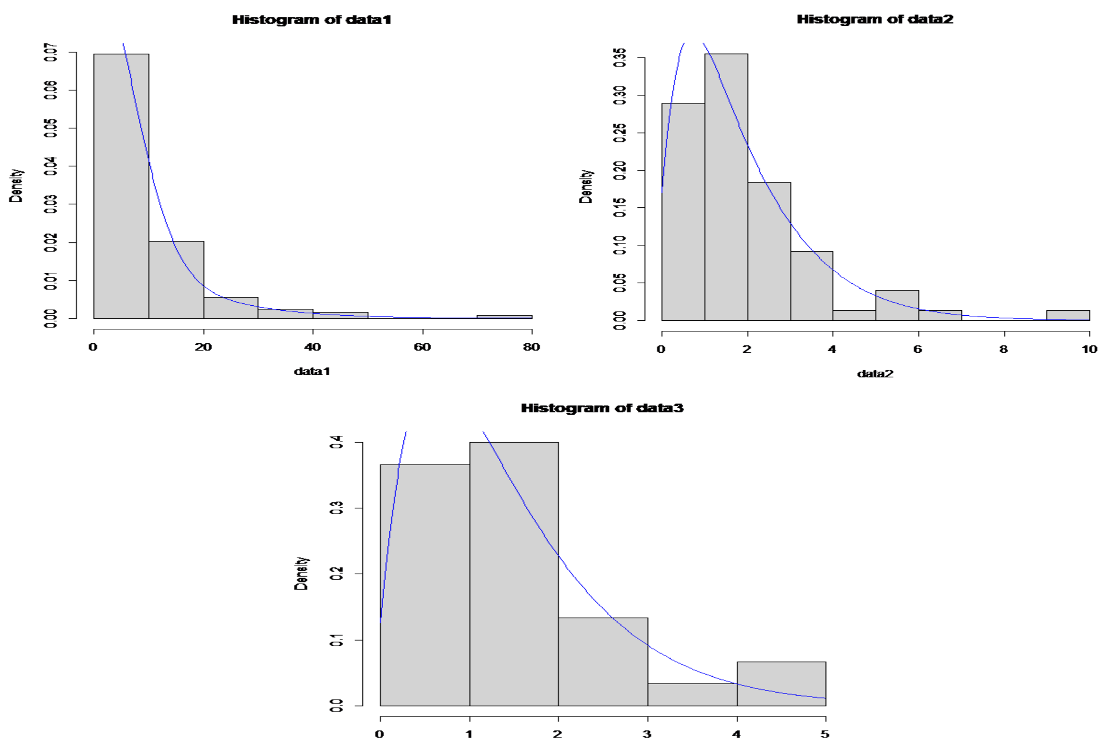

Figure 3: The histogram for the three data sets and fitted pdf of the gLinear failure rate distribution

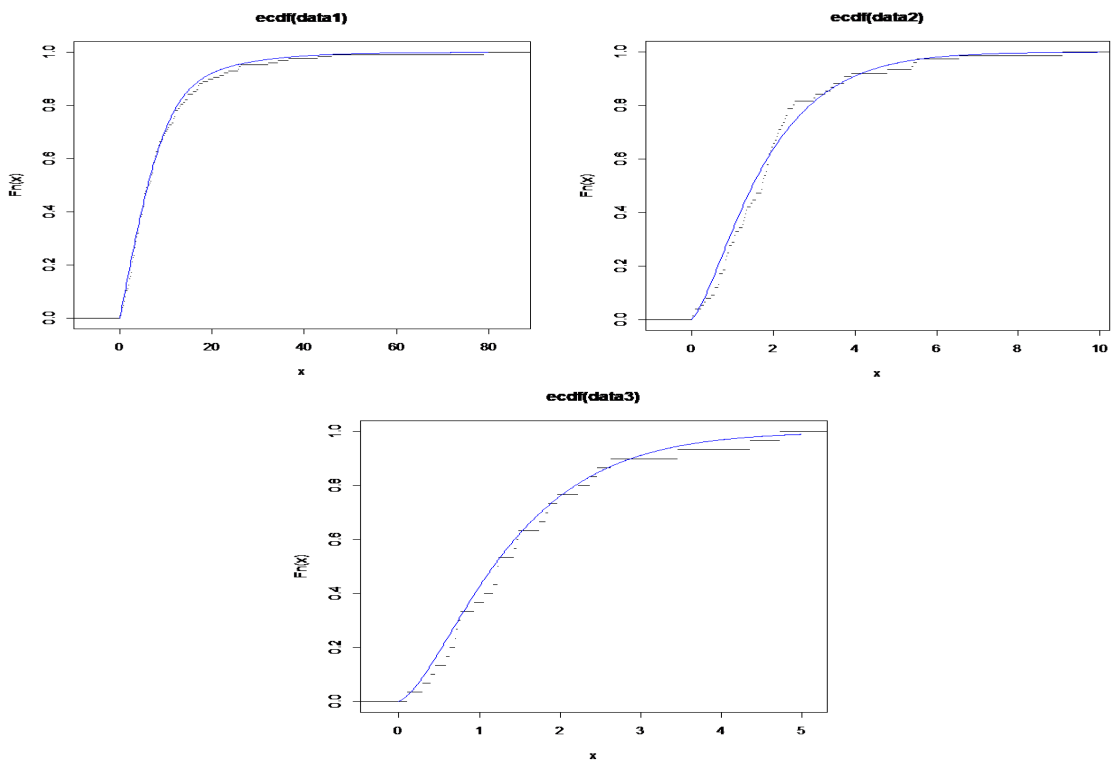

Figure 4: The empirical cdf for the three data sets and fitted cdf of the gLinear failure rate distribution

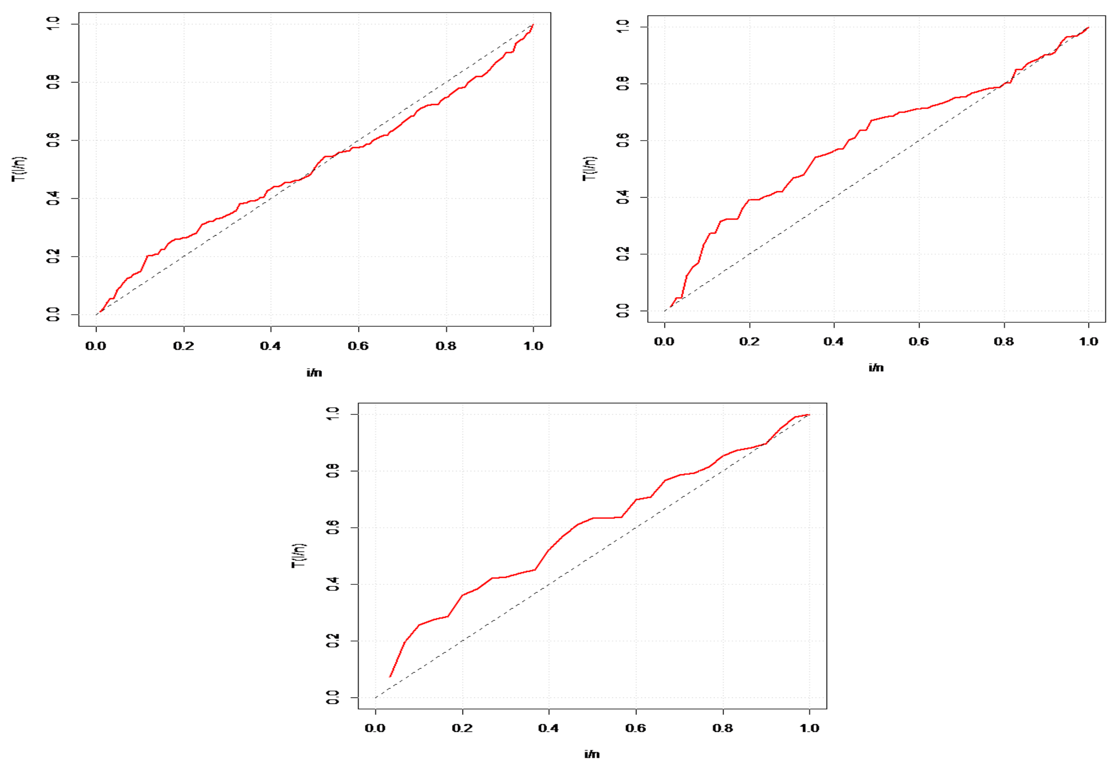

Figure 5: The TTT for the three data sets

From the results in Table 2, one can note that the values of AIC, BIC, CAIC and HQIC are smaller for the glfr distribution compared with those values of the other models, so the new distribution seems to be a very competitive model to the the first two data sets. For the third data set, gamma model has smaller values than glfr model.

Also, we plotted the scaled total time on test transform (TTT) which can help for selecting a model. The empirical scaled TTT transform (Aarset (1987)) can be used to identify the shape of the hazard function. As displayed in Figure 5 The TTT plot shows that the data set 1 has a unimodel hazard, while the rest of data sets have increasing hazards.

For Bayesian computations, we concern with three approaches; the Laplace approximation, MCMC and variational Bayes (VB). We obtain the approximate Bayes estimates of the two unknown parameters of the glfr distribution based on real data sets and simulated samples with true values $\alpha = 1.2, \beta = 1.2$. R package is used to compute these estimates. Using the first two data sets and gamma priors with different values of hyper parameters $((a_i, b_i) = (1, 0.001)$ and $(0.001, 0.001), i = 1, 2)$, the Laplace approximation, MCMC and variational Bayes are carried out and the results are shown in Tables 4-5. From these tables, one can note that the results are close for each other and for simulated data, the results get closer to the true values as the sample size increases. Also, the results are close to each other for different hyper parameters and Tables 4-5 display the results in the case of the gamma priors with $(a_1, b_1) = (1, 0.001)$. For MCMC, Figures 6- 7 show the trace, the approximated posterior density functions and autocorrelation plots of the two parameters of the glfr distribution. These Figures show that as the sample size increases, the chains look stationary, the kernel densities look Gaussian, and the ACF's or autocorrelation function plot show low autocorrelation.

7. Conclusion

A new mixture distribution named glinear failure rate distribution (glfr) is proposed in this paper. The glfr is a mixture of gamma and failure rate distributions. Based on some goodness of fit tests and some criteria for choosing the best fit among several, it is observed that the glfr gives a better fit than some common distributions. The maximum likelihood and Bayesian methods are applied to estimate the two unknown parameters of the glfr distribution. For Bayesian method, we used Laplace approximation, MCMC and Variational Bayes and the results are close to each other and for simulated data, the results get closer to the true values as the sample size increases.

Acknowledgements

I am very grateful to the Editors for their guidance and counsel. I am very grateful to the reviewer for valuable comments.

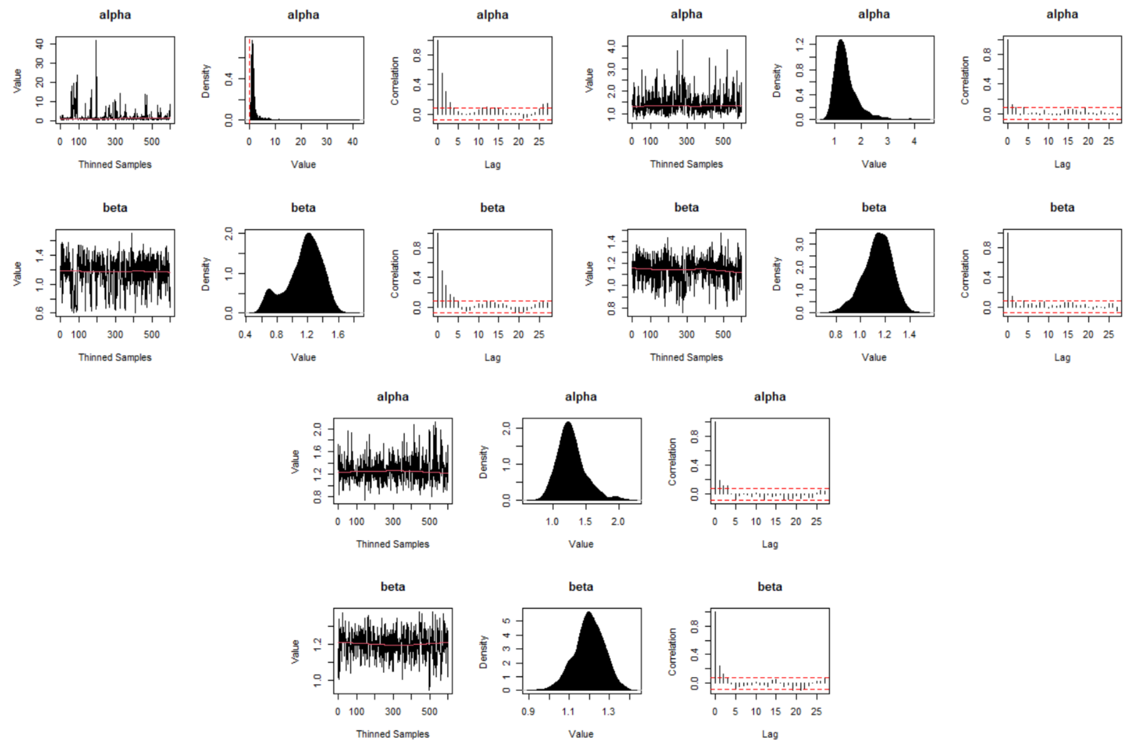

Figure 6: The trace, the approximated posterior density functions and autocorrelation plots of the two parameters for the simulated data

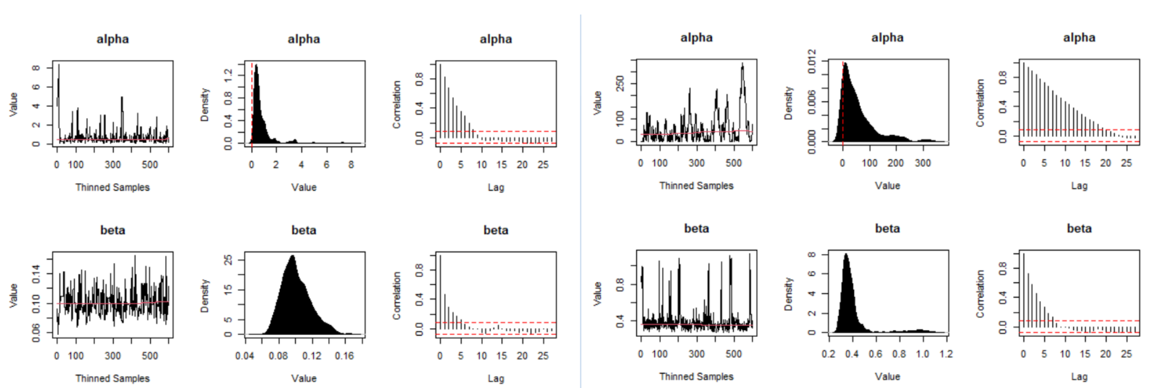

Figure 7: The trace, the approximated posterior density functions and autocorrelation plots of the two parameters for the two data sets

References

Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, **36**, 106–108.

Crawford, S. L. (1994). An application of the laplace method to finite mixture distributions. *Journal of the American Statistical Association*, **89**, 259–267.

Everitt, B. (2013). *Finite Mixture Distributions*. Springer Science & Business Media.

Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, **78**, 493–506.

Lindley, D. V. (1958). Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society, Series B (Methodological)*, **20**, 102–107.

Mahmoud, M. R., Mandouh, R., and Rasheedy, E. (2017). A new lifetime distribution based on finite mixtures. *International Journal of Applied Mathematics & Statistics*, **56**, 13–23.

Murthy, D. P., Xie, M., and Jiang, R. (2004). *Weibull Models*. John Wiley & Sons.

Ogunde, A., Ibraheem, A., and Audu, A. (2017). Performance rating of transmuted nadarajah and haghighi exponential distribution: an analytical approach. *Journal of Statistics: Advances in Theory and Applications*, **17**, 137–151.

Puza, B. (2015). *Bayesian Methods for Statistical Analysis*. ANU press.

Sarhan, A. M., Tadj, L., and Hamilton, D. C. (2014). A new lifetime distribution and its power transformation. *Journal of Probability and Statistics*, **2014**, 1–14.

Sen, S. and Chandra, N. (2017). The quasi xgamma distribution with application in bladder cancer data. *Journal of Data Science*, **15**, 61–76.

Sen, S., Maiti, S. S., and Chandra, N. (2016). The xgamma distribution: statistical properties and application. *Journal of Modern Applied Statistical Methods*, **15**, 38.

Shanker, R. and Mishra, A. (2013). A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research*, **6**, 64–71.

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, **81**, 82–86.

Tierney, L., Kass, R. E., and Kadane, J. B. (1989). Fully exponential laplace approximations to expectations and variances of nonpositive functions. *Journal of the American Statistical Association*, **84**, 710–716.

Yamaguchi, Y., Okamura, H., and Dohi, T. (2010). A variational bayesian approach for estimating parameters of a mixture of erlang distribution. *Communications in Statistics-Theory and Methods*, **39**, 2333–2350.

Zea, L. M., Silva, R. B., Bourguignon, M., Santos, A. M., and Cordeiro, G. M. (2012). The beta exponentiated pareto distribution with application to bladder cancer susceptibility. *International Journal of Statistics and Probability*, **1**, 8.

ANNEXURE

Table 1: MSE, coverage probability (CP) and average width (AW)

α	β	n	MSE_α	MSE_β	$CP_\alpha\%$	AW_α	$CP_\beta\%$	AW_β
0.8	0.8	50	0.8392	3.5144	99.5	3.2434	99.5	13.7551
		75	0.4296	2.5246	99.6	1.6422	99.7	9.8887
		100	0.3147	1.5077	98.9	1.1966	99.9	5.9097
		150	0.2112	0.1055	93.1	0.7744	94.7	0.3997
		200	0.1907	0.0972	93.1	0.6707	93.6	0.3628
0.8	1.0	50	0.5761	2.0099	99.5	2.2118	99.8	7.8758
		75	0.9416	1.3697	99.9	3.6092	99.9	5.3624
		100	0.3623	0.1921	90.4	1.2151	91.1	0.6762
		150	0.3131	0.1643	86.5	0.9748	90.8	0.5614
		200	0.2879	0.1544	88.1	0.8692	89.8	0.5184
1.0	1.0	50	2.5306	7.2913	98.8	9.8568	98.5	28.4180
		75	1.2069	4.7571	99.3	4.6747	99.3	18.5986
		100	0.7010	3.5606	99.6	2.7089	99.6	13.9390
		150	0.2758	0.1388	94.2	1.0408	95.8	0.5321
		200	0.2635	0.1197	93.7	0.9921	95.3	0.4594
1.0	1.2	50	2.26458	7.1844	99.5	8.8561	99.5	28.1341
		75	0.5315	3.0952	99.4	2.0152	99.8	12.1331
		100	0.3748	2.0887	93.4	1.3431	99.9	8.1863
		150	0.3503	0.1840	89.5	1.1865	92.9	0.6266
		200	0.3500	0.1709	89.6	1.1591	91.1	0.5976
1.2	1.2	50	1.8795	5.0542	99.6	7.2992	99.6	19.7939
		75	0.6501	2.8649	99.8	2.472	99.8	11.2279
		100	0.3362	0.1913	90.5	1.2115	94.8	0.7192
		150	0.3096	0.1483	92.3	1.1403	94.3	0.5555
		200	0.2896	0.1416	92.5	1.0543	93.9	0.5360
1.6	2.5	50	1.4358	0.8157	83.1	4.0621	87.0	2.5023
		75	1.4005	0.7625	76.4	3.3660	94.2	2.5460
		100	1.5233	0.5758	73.0	3.5535	89.8	1.8990
		150	1.3813	0.4837	61.7	2.7944	89.1	1.5026
		200	1.2944	0.4622	55.0	2.4079	85.8	1.3990

Table 2: The MLEs and some measures for the fitted models

Data set	Model	Parameter Estimates	$l(\hat{\boldsymbol{\theta}})$	AIC	BIC	CAIC	HQIC
Data 1	glfr(α, β)	$\hat{\alpha} = 0.1911$ and $\hat{\beta} = 0.1342$	-411.8	827.6	833.3	827.7	830
	glfr(β)	$\hat{\beta} = 0.0888$	-412.6	827.2	830.0	827.2	828.3
	qxgamma	$\hat{\alpha} = 2.8289$ and $\hat{\beta} = 0.1655$	-416.9	837.9	843.6	838.0	840.2
	xgamma	$\hat{\beta} = 0.2689$	-429.4	860.7	863.6	863.8	861.9
	qLindley	$\hat{\alpha} = 2.5292$ and $\hat{\beta} = 0.1381$	-414.9	833.7	839.4	833.8	836.1
	Lindley	$\hat{\beta} = 0.1960$	-419.5	841.1	843.9	841.1	842.2
	lfr	$\hat{\beta} = 0.0608$	-427.2	856.6	859.3	856.5	857.6
	gamma	$\hat{\beta} = 0.2135$	-426.8	855.6	858.4	855.6	856.8
Data 2	glfr(α, β)	$\hat{\alpha} = 0.2086$ and $\hat{\beta} = 0.9439$	-122.2	248.4	253.1	248.6	250.3
	glfr(β)	$\hat{\beta} = 0.5537$	-124.0	250.0	252.4	250.1	251
	qxgamma	$\hat{\alpha} = 0.2009$ and $\hat{\beta} = 1.3561$	-122.5	249.0	253.6	249.1	250.8
	xgamma	$\hat{\beta} = 1.0330$	-126.3	254.7	257.0	254.7	255.6
	qLindley	$\hat{\alpha} = 0.2947$ and $\hat{\beta} = 0.8823$	-122.0	248.0	252.7	248.2	249.9
	Lindley	$\hat{\beta} = 0.7948$	-123.7	249.4	251.7	294.4	250.3
	lfr	$\hat{\beta} = 0.3329$	-124.5	251.0	253.3	251.0	251.9
	gamma	$\hat{\beta} = 1.0210$	-123.2	248.4	250.7	248.4	249.3
Data 3	glfr(α, β)	$\hat{\alpha} = 0.1394$ and $\hat{\beta} = 1.3079$	-39.70	83.40	86.10	83.84	84.30
	glfr(β)	$\hat{\beta} = 0.8413$	-41.25	84.50	85.90	84.64	84.94
	qxgamma	$\hat{\alpha} = 0.1599$ and $\hat{\beta} = 1.156$	-40.52	85.04	87.84	85.48	85.94
	xgamma	$\hat{\beta} = 1.2690$	-42.14	86.28	87.69	86.42	86.73
	qLindley	$\hat{\alpha} = 0.4026$ and $\hat{\beta} = 1.2962$	-40.59	85.18	87.98	85.62	86.07
	Lindley	$\hat{\beta} = 0.9947$	-41.09	84.18	85.59	84.33	84.63
	lfr	$\hat{\beta} = 0.4408$	-40.73	83.46	84.86	83.60	83.91
	gamma	$\hat{\beta} = 1.3250$	-39.52	81.04	82.44	81.19	81.49

Table 3: Statistics K-S (P value), A* and W* for the three data sets

Data set	Model	K-S (P value)	A*	W*
Data 1	glfr(α, β)	0.059 (0.800)	0.3835	0.0607
	glfr(β)	0.055 (0.800)	0.6957	0.1236
	qxgamma	0.100 (0.100)	1.0160	0.1687
	xgamma	0.160(0.003)	2.2250	0.3787
	qLindley	0.074 (0.500)	0.9005	0.1510
	Lindley	0.120 (0.060)	1.0260	0.1717
	lfr	0.180 (6e-04)	2.2630	0.3849
	gamma	0.140 (0.010)	0.7260	0.1211
Data 2	glfr(α, β)	0.091 (0.500)	0.6425	0.1102
	glfr(β)	0.130 (0.200)	0.4919	0.8261
	qxgamma	0.110 (0.300)	0.7579	0.1266
	xgamma	0.150 (0.070)	0.9901	0.1698
	qLindley	0.120 (0.200)	0.6356	0.1082
	Lindley	0.120 (0.200)	0.6907	0.1173
	lfr	0.130 (0.200)	1.0470	0.1807
	gamma	0.098 (0.400)	0.7005	0.1182
Data 3	glfr(α, β)	0.099 (0.900)	0.1520	0.0200
	glfr(β)	0.130 (0.700)	0.1328	0.0188
	qxgamma	0.120 (0.800)	0.2934	0.0401
	xgamma	0.160 (0.400)	0.2481	0.03293
	qLindley	0.120 (0.800)	0.1930	0.0258
	Lindley	0.130 (0.700)	0.1843	0.2436
	lfr	0.110 (0.900)	0.3036	0.0465
	gamma	0.095 (0.900)	0.1496	0.0195

Table 4: Summary results for the posterior parameters in the case of the glfr distribution based on real data sets

Data set	Parameter	Estimate: Mode	Standard Deviation	LB	UB	Minutes of run-time
Laplace Approximation						
Data 1	α	0.3471	0.1829	0.0000	0.7128	0.00
	β	0.1061	0.0172	0.0718	0.1404	
Data 2	α	0.1157	0.1063	0.0000	0.3282	0.00
	β	0.9410	0.1049	0.7312	1.1509	
MCMC						
Data set	Parameter	Estimate: Mode	Standard Deviation	LB	UB	Minutes of run-time
Data 1	α	0.5968	0.5940	0.1383	2.1327	0.09
	β	0.10478	0.0186	0.0784	0.1475	
Data 2	α	56.4094	65.2242	0.0934	230.0292	0.07
	β	0.4108	0.1615	0.2878	0.9777	
Variational Bayesian						
Data set	Parameter	Estimate: Mean	Standard Deviation	LB	UB	Minutes of run-time
Data 1	α	0.4631	0.0996	0.2639	0.6622	0.02
	β	0.1028	0.0114	0.0799	0.1257	
Data 2	α	0.1763	0.9364	0.0000	0.3636	0.02
	β	0.9162	0.1047	0.7068	1.1255	

Table 5: Summary results for the posterior parameters in the case of the glfr distribution based on simulated samples with true values $\alpha = 1.2, \beta = 1.2$

		Laplace Approximation				
Samples	Parameter	Estimate: Mode	Standard Deviation	LB	UB	Minutes of run-time
n=300	α	1.2886	0.3798	0.6621	2.1081	0.00
	β	1.2314	0.1419	0.9973	1.5034	
n=500	α	1.2361	0.2786	0.6789	1.7934	0.00
	β	1.1732	0.1003	0.9726	1.3738	
n=1000	α	1.1895	0.1850	0.8196	1.5595	0.01
	β	1.2200	0.0742	1.0792	1.3608	
		MCMC				
Samples	Parameter	Estimate: Mode	Standard Deviation	LB	UB	Minutes of run-time
n=300	α	2.3272	3.2792	0.6940	12.1946	0.12
	β	1.1551	0.2263	0.6647	1.4986	
n=500	α	1.4436	0.4626	0.8844	2.6375	0.16
	β	1.1368	0.1152	0.8791	1.3386	
n=1000	α	1.2741	0.2175	0.9189	1.7781	0.26
	β	1.9835	0.0740	1.0441	1.3348	
		Variational Bayesian				
Samples	Parameter	Estimate: Mean	Standard Deviation	LB	UB	Minutes of run-time
n=300	α	1.3167	0.3329	0.6508	1.9826	0.02
	β	1.2242	0.13173	0.9607	1.4876	
n=500	α	1.3633	0.2822	0.7990	1.9276	0.04
	β	1.1455	0.0983	0.9490	1.3420	
n=1000	α	1.2414	0.1901	0.8613	1.6215	0.07
	β	1.2086	0.0703	1.0680	1.3492	