Article

Not peer-reviewed version

Scaled-Invariant Extended Quasi-Lindley Model: Properties, Estimation and Application

Mohamed Kayid , Abdulrahman Abouammoh , Ghadah Alomani *

Posted Date: 25 July 2023

doi: 10.20944/preprints202307.1629.v1

Keywords: Quasi Lindley model; maximum likelihood estimator; expectation maximization algorithm

Article

Scaled-Invariant Extended Quasi-Lindley Model: Properties, Estimation and Application

Mohamed Kayid [1], Abdulrahman Abouammoh [1] and Ghadah Alomani [2,*]

[1] Department of Statistics and Operations Research, College of Science, King Saud University, Riyadh, Saudi Arabia, P.O. Box 2455, Riyadh 11451, Saudi Arabia
[2] Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
* Correspondence: gaalomani@pnu.edu.sa

Abstract: In many research fields, statistical probability models are often used to analyze real-world data. However, data from many fields, such as the environment, economics, and health care, do not necessarily fit traditional models. New empirical models need to be developed to improve their fit. In this paper, we explore a further extension of the quasi-Lindley model. Maximum likelihood, least square error, Anderson-Darling algorithm, and expectation-maximization algorithm are four techniques for estimating the parameters under study. All techniques provide accurate and reliable estimates of the parameters. However, the mean square error of the expectation maximization approach was lower. The usefulness of the proposed model was demonstrated by analyzing a dynamical systems data set, and the analysis shows that it outperforms the other models in all statistical models considered.

Keywords: Quasi Lindley model; maximum likelihood estimator; expectation maximization algorithm

1. Introduction

A Lindley model that is simple and remarkably flexible in application was proposed by [1]. It is characterized by the probability density function (pdf)

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

which is a mixture of two gamma models $G(1,\xi)$ and $G(2,\xi)$ with weights, $\xi/\xi+1$ and $1/\xi+1$, respectively. Numerous studies have been conducted on the Lindley model. For example, many properties, extensions, and applications of the model have been studied in [2–21]. A scale-invariant version of the Lindley model, namely the quasi Lindley (QL), with the pdf

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

which is a mixture of two gamma models $G(1,\xi)$ and $G(2,\xi)$ with weights $\alpha/\alpha+1$ and $1/\alpha+1$ respectively proposed by [22].

A family of models characterized by $f(x)$ is said to be scale-invariant if the transformation from x to kx lies within the family k pdf times a Jacobi associated with that transformation. Thus, if you change the scale of measurement or the unit of x, the fit remains invariant. For instance, a lifetime can be measured in days, hours, or minutes, and the unit of measurement does not affect inferences about lifetimes. Since scale invariance is an essential property of lifetime models, this model has attracted considerable interest. A comparison of the maximum likelihood estimator (MLE) and the expectation-maximization (EM) algorithm for estimating the parameters of the QL model studied by [23] and a new scale-invariant extension of the Lindley model proposed by [24].

Many data sets are composed of multiple populations or sources, and the subpopulation associated with each data is usually unknown or recorded. For example, the lifetime of a device or system may be available, but the manufacturer or an event associated with a living being without its geographic location may not be. Such data sets are mixtures because information about some covariates, such as the manufacturer or geographic location, that significantly affect the observations is unknown. For detailed information on mixture models, see [26,27]. The Lindley model and its extensions are examples of mixture models of the gamma distribution that can be useful for describing many real-world applications.

This study proposes a new extension of the scaled invariant QL model, a mixture of three gamma models, and is investigated. Some statistical and reliability properties, such as failure rate (FR), mean residual life (MRL), and p-quantile residual life (p-QRL) functions, are discussed. Four methods for estimating the model parameters are then discussed. It is examined that all methods provide consistent and efficient estimates of the parameters. However, the algorithm for maximizing expected values yields a lower mean square error.

The rest of the article is organized as follows. The scaled-invariant extended quasi-Lindley (EQL) model is explained in Section 2 along with some of its basic properties. Section 3 estimates the parameters of the model using the maximum likelihood (ML) method, the least squares error (LSE) method, the weighted LSE method, and the algorithm EM. A simulation study is then conducted in Section 4 to investigate and compare the behavior of the estimators. The proposed model is fitted to a dataset of intervals between successive air conditioning failures in a Boeing 720 aircraft to show how useful it could be in practice.

2. Scaled-invariant extended QL model

A random variable X follows from $EQL(\alpha,\xi)$ if its PDF equals

$$f(x)=\frac{\xi}{1+\alpha+\alpha^2}\left(1+\alpha\xi x+\frac{1}{2}\alpha^2\xi^2x^2\right)e^{-\xi x},\ \alpha\geq 0,\xi>0,x\geq 0. \tag{3}$$

It shows a mixture of $G(1,\xi)$, $G(2,\xi)$ and $G(3,\xi)$ with weights $1/1+\alpha+\alpha^2$, $\alpha/(1+\alpha+\alpha^2)$ and $\alpha^2/(1+\alpha+\alpha^2)$ respectively. When $\alpha=0$, it reduces to the exponential model. The reliability function is an important yet very simple measure in reliability theory and survival analysis. For the EQL model it is

$$R(x)=\frac{1}{(1+\alpha+\alpha^2)}\left(1+\alpha+\alpha^2+\alpha\xi x+\alpha^2\xi x+\frac{1}{2}\alpha^2\xi^2x^2\right)e^{-\xi x}. \tag{4}$$

The distribution function is simply related to the reliability function by $F(x)=1-R(x)$ and the quantile function which is in fact the inverse of the distribution function equals

$$q(p)=F^{-1}(p)=\min\{x:\ F(x)=p\},\quad 0<p<1.$$

The quantile function could be used for simulation random samples, estimating the parameters, and computing skewness of the model.

In addition, for the EQL the k-th moment is finite and equal to

$$E(X^k)=\frac{1}{1+\alpha+\alpha^2}\frac{1}{\xi^k}\left[\Gamma(k+1)+\alpha\Gamma(k+2)+\frac{\alpha^2}{2}\Gamma(k+3)\right]. \tag{5}$$

2.1. Statistics and Reliability Properties

The FR function at a time x expresses the instantaneous risk of fail at x given survival up to x. For EQL model we have

$$\lambda(x)=\frac{1+\alpha\xi x+\frac{1}{2}\alpha^2\xi^2x^2}{1+\alpha+\alpha^2+\alpha\xi x+\alpha^2\xi x+\frac{1}{2}\alpha^2\xi^2x^2}\xi. \tag{6}$$

Using simple algebra, we can see that the FR function increases from $\lambda(0)=\xi/1+\alpha+\alpha^2$ to $\lim_{x\to\infty}\lambda(x)=\xi$.

Figure 1 shows the shape of the pdf and the FR function for some parameter values. Two other useful and well-known measures in the reliability theory and survival analysis are MRL and p-QRL functions. At time x, they describe the mean and p-quantile of the remaining life for survival to x. For EQL, the MRL is obtained by

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2023 doi:10.20944/preprints202307.1629.v1

$$m(x) = \frac{1+2\alpha+3\alpha^2+(\alpha\xi+2\alpha^2\xi)x+\frac{1}{2}\alpha^2\xi^2x^2}{1+\alpha+\alpha^2+(\alpha\xi+\alpha^2\xi)x+\frac{1}{2}\alpha^2\xi^2x^2}\frac{1}{\xi}. \quad (7)$$

Since the FR function is increasing, it follows that the MRL function decreases from $m(0) = \frac{1+2\alpha+3\alpha^2}{1+\alpha+\alpha^2}\frac{1}{\xi}$ to $\frac{1}{\xi}$ at infinity.

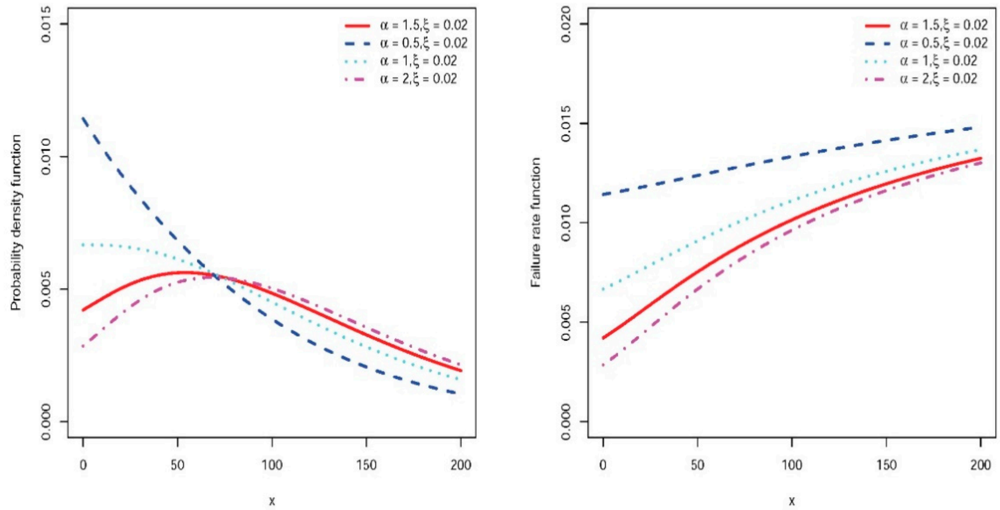

Figure 1. The PDF (left) and FR (right) of EQL for some parameter values.

The p-QRL reads

$$q_p(x) = R^{-1}((1-p)R(x)) - x, \quad (8)$$

which can be calculated numerically. Like the MRL, this measure is a decreasing function of x. When $p = 0.5$, it is called the median residual life, which is a good alternative to the MRL. In Figure 2, the MRL and the median residual lifetime are plotted for some parameter values and show their similar behavior.

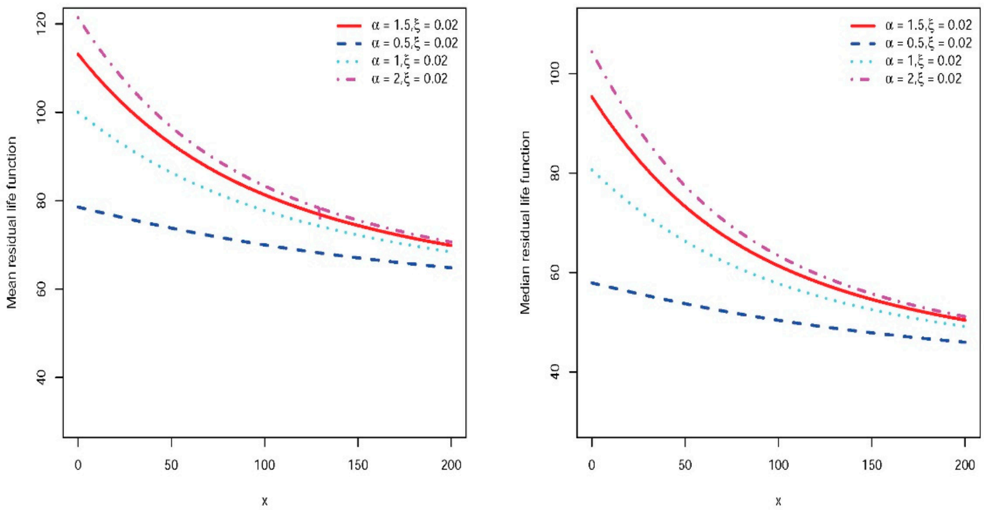

Figure 2. The MRL (left) and median residual life (right) of EQL for some parameter values.

An important concept in reliability theory and survival analysis is orderings between lifetimes. For two lifetimes X_1 and X_2 following reliability functions R_1 and R_2 respectively, we say that, X_2 is greater than X_1, $X_2 \geq X_1$, in stochastic if $R_2(x) \geq R_1(x)$ for every x. Equivalently we may write

$R_2 \geq R_1$ in stochastic. There are other useful orderings, e.g., by means of the FR function, $X_2 \geq X_1$, in FR if $h_1(x) \geq h_2(x)$ for every x. Moreover, $X_2 \geq X_1$, in MRL and p-QRL if $m_2(x) \geq m_1(x)$ and $q_{p,2}(x) \geq q_{p,1}(x)$ respectively for every x. The following result shows that EQL is internally ordered in terms of α.

Proposition 1. *Let X_i, $i = 1,2$ follows from $EQL(\alpha, \xi)$ and $\alpha_2 \geq \alpha_1$, then $X_2 \geq X_1$ in stochastic, FR, MRL and p-QRL.*

Proof. To show the FR ordering, the derivative of the FR function in terms of α is proportional to

$$\frac{d\lambda}{d\alpha} \propto -\left(\frac{1}{2}\alpha^2\xi^2x^2 + \alpha^2\xi x + 2\alpha\xi x + 2\alpha + 1\right) < 0.$$

So, the FR ordering follows. The stochastic, MRL and p-QRL orderings follows from FR ordering. See Lai and Xie [28] for relationship between orderings. □

3. Estimation

This section discusses some methods for estimating the parameters of the EQL model. In particular, the parameters are estimated using ML, LSE, weighted LSE methods, and an advanced EM algorithm.

3.1. ML method

Let $x_1, x_2, \ldots, x_n$ represent independent and identically distributed (iid) instances from $EQL(\alpha, \xi)$. Then, the log-likelihood function is

$$l(\alpha, \xi; \mathbf{x}) = n\ln\xi - n\ln(1 + \alpha + \alpha^2) + \sum_{i=1}^{n} \ln(1 + \alpha\xi x_i + \tfrac{1}{2}\alpha^2\xi^2x_i^2) - \xi\sum_{i=1}^{n} x_i. \quad (9)$$

The ML estimator of (α, ξ) denoted by $(\hat{\alpha}, \hat{\xi})$ maximizes the log-likelihood function and can be computed directly by numerical methods or by solving the following likelihood equations.

$$\frac{\partial}{\partial\alpha} l(\alpha, \xi; \mathbf{x}) = -n\frac{1 + 2\alpha}{1 + \alpha + \alpha^2} + \sum_{i=1}^{n} \frac{\xi x_i + \alpha\xi^2x_i^2}{1 + \alpha\xi x_i + \frac{1}{2}\alpha^2 x\xi^2x_i^2} = 0,$$

and

$$\frac{\partial}{\partial\xi} l(\alpha, \xi; \mathbf{x}) = \frac{n}{\xi} + \sum_{i=1}^{n} \frac{\alpha x_i + \alpha^2\xi x_i^2}{1 + \alpha\xi x_i + \frac{1}{2}\alpha^2\xi^2x_i^2} - \sum_{i=1}^{n} x_i = 0.$$

The observed Fisher information matrix can be calculated by replacing $\hat{\alpha}$ and $\hat{\xi}$ for α and ξ in the following Fisher information matrix.

$$O = \begin{bmatrix} -\frac{\partial^2}{\partial\alpha^2} & -\frac{\partial^2}{\partial\alpha\partial\xi} \\ -\frac{\partial^2}{\partial\xi\partial\alpha} & -\frac{\partial^2}{\partial\xi^2} \end{bmatrix} l(\alpha, \xi; \mathbf{x}). \quad (10)$$

Then the asymptotic distribution of $(\hat{\alpha}, \hat{\xi})$ is approximately the bivariate normal distribution with mean (α, ξ) and variance-covariance matrix O^{-1}.

3.2. LSE method

In this approach, we search for parameter values which minimize the sum of squared distances between the empirical distribution and the estimated distribution functions. More precisely, we minimize the following expression in terms of the parameters.

$$S^2 = \sum_{i=1}^{n} (F(x_i) - \hat{F}(x_i))^2.$$

By substituting the distribution function, we have

$$S^2 = \sum_{i=1}^{n} \left(\frac{1}{(1 + \alpha + \alpha^2)}\left(1 + \alpha + \alpha^2 + \alpha\xi x_i + \alpha^2\xi x_i + \frac{1}{2}\alpha^2\xi^2x_i^2\right)e^{-\xi x_i} - \frac{i}{n}\right)^2.$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2023 doi:10.20944/preprints202307.1629.v1

Then, the estimates could be computed as in the following.

$$(\hat{\alpha},\hat{\xi}) = \arg\min_{(\alpha,\beta,\lambda)} \sum_{i=1}^{n} \left(\frac{1}{(1+\alpha+\alpha^2)}\left(1+\alpha+\alpha^2+\alpha\xi x_i+\alpha^2\xi x_i+\frac{1}{2}\alpha^2\xi^2x_i^2\right)e^{-\xi x_i}-\frac{i}{n}\right)^2.$$

3.3. Weighted LSE method

A well-known weight which could improve the LSE estimate is $\frac{1}{F(x_i)(1-F(x_i))}$. With this idea, the weighted LSE estimate are computed by minimizing the following expression in terms of the parameters.

$$S^2 = \sum_{i=1}^{n} \frac{1}{F(x_i)(1-F(x_i))}(F(x_i)-\hat{F}(x_i))^2.$$

This method is well-known as the Anderson Darling (AD) method.

3.2. EM algorithm

Suppose that $X_i,\ i=1,2,\dots,n$ is an iid sample from $EQL(\alpha,\xi)$. For a short exposition, take $\theta=(\alpha,\xi)$. Since EQL is a mixture of three gamma models $G(j,\xi),\ j=1,2,3$, we consider a latent random variable V_i such that $V_i=j$, when X_i comes from $G(j,\xi)$. Thus, $(X_i|V_i=j,\theta)\sim G(j,\xi)$ and $P(V_i=j|\theta)=\frac{\alpha^{j-1}}{1+\alpha+\alpha^2}$, $j=1,2,3$. However, the latent variable V_i will not be observed, but applying it helps to improve the estimation of the parameters in an iterative process. With the evidences X_i and V_i, $i=1,2,\dots,n$, the likelihood function can be written as follows.

$$L(\theta;\mathbf{x},\mathbf{v}) = \prod_{i=1}^{n}\prod_{j=1}^{3}\left(g(x_i|\theta)P(V_i=j|\theta)\right)^{I(v_i=j)}, \quad (11)$$

where $I(v_i=j)$ equals 1 when $v_i=j$ and 0 otherwise, and $g_j(x_i|\theta)$ represents the PDF of gamma $G(j,\xi)$. Then, the log-likelihood function is

$$l(\theta;\mathbf{x},\mathbf{v}) = \sum_{i=1}^{n}\sum_{j=1}^{3} I(V_i=j)\ln\left(\frac{\xi^j x_i^{j-1}}{\Gamma(j)}e^{-\xi x_i}\frac{\alpha^{j-1}}{1+\alpha+\alpha^2}\right). \quad (12)$$

Since this function depends on the unobserved random variable V_i, we cannot estimate the parameters by maximizing them directly. One approach is to implement an iterative process with expectation (E) and maximization (M) steps. In the E step, the expected log-likelihood function is constructed with respect to the conditional latent variable. In the M step, the expected log-likelihood function is maximized to estimate the parameters.

E step:

Assume that the estimate of the parameters at iteration t, $\theta_t=(\alpha_t,\xi_t)$ is known. Then, by the well-known Bayes formula, the conditional probabilities of V_i is

$$\begin{aligned} p_{ij,t} &= P(V_i=j\mid X_i=x_i,\theta_t) = \frac{f(X_i=x_i\mid V_i=j,\theta_t)P(V_i=j\mid\theta_t)}{f(X_i=x_i|\theta_t)} \\ &= \frac{\frac{\xi_t^j}{\Gamma(j)}x_i^{j-1}e^{-\xi_t x_i}\alpha_t^{j-1}}{\sum_{j=1}^{3}\frac{\xi_t^j}{\Gamma(j)}x_i^{j-1}e^{-\xi_t x_i}\alpha_t^{j-1}},\quad i=1,2,\dots,n, j=1,2,3. \end{aligned} \quad (13)$$

So,

$$p_{i1,t} = \frac{1}{1+\alpha_t\xi x_i+\frac{1}{2}\alpha_t^2+\xi_t^2x_i^2}, \quad (14)$$

$$p_{i2,t} = \frac{\alpha_t\xi_t x_i}{1+\alpha_t\xi_t x_i+\frac{1}{2}\alpha_t^2+\xi_t^2x_i^2}, \quad (15)$$

and

$$p_{i3,t} = 1-p_{i1,t}+p_{i2,t}.$$

Now, applying these probabilities, we can write the expected log-likelihood function at iteration t.

$$Q(\theta|\theta_t) = E_{V|X,\theta_t}(l(\theta;\mathbf{x},\mathbf{V})) = \sum_{i=1}^{n} E_{V_i|X_i,\theta_t} \sum_{j=1}^{3} I(V_i = j)\ln\left(\frac{\xi^j x_i^{j-1}}{\Gamma(j)} e^{-\xi x_i} \frac{\alpha^{j-1}}{1+\alpha+\alpha^2}\right)$$

$$= \sum_{i=1}^{n} P(V_i = 1|X_i = x_i, \theta_t)\ln\left(\frac{\xi}{1+\alpha+\alpha^2} e^{-\xi x_i}\right)$$

$$+ \sum_{i=1}^{n} P(V_i = 2|X_i = x_i, \theta_t)\ln\left(\frac{\alpha \xi^2 x_i}{1+\alpha+\alpha^2} e^{-\xi x_i}\right)$$

$$+ \sum_{i=1}^{n} P(V_i = 3|X_i = x_i, \theta_t)\ln\left(\frac{1}{2}\frac{\alpha^2 \xi^3 x_i^2}{1+\alpha+\alpha^2} e^{-\xi x_i}\right)$$

$$= \sum_{i=1}^{n} (1 + p_{i2,t} + 2p_{i3,t})\ln\xi - \xi \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} (p_{i2,t} + 2p_{i3,t})\ln(\alpha x_i) - n\ln(1+\alpha+\alpha^2) + \sum_{i=1}^{n} p_{i3,t}\ln\frac{1}{2}. \quad (16)$$

Clearly, $Q(\theta|\theta_t)$ consists of three expressions

$$Q_1(\xi) = \sum_{i=1}^{n} (1 + p_{i2,t} + 2p_{i3,t})\ln\xi - \xi \sum_{i=1}^{n} x_i,$$

$$Q_2(\alpha) = \sum_{i=1}^{n} (p_{i2,t} + 2p_{i3,t})\ln(\alpha x_i) - n\ln(1+\alpha+\alpha^2), \quad (17)$$

depending solely on ξ and α respectively and $Q_3 = \ln\frac{1}{2}\sum_{i=1}^{n} p_{i3,t}$ which does not depend on ξ or α.

M step:

To estimate the parameters at $t+1$ iteration, we should maximize $Q(\theta|\theta_t)$ in terms of $\theta = (\alpha, \xi)$. Thus, for estimating ξ at iteration $t+1$, we could simply solve the likelihood equation $\frac{\partial Q_1(\xi)}{\partial \xi} = 0$ which gives $\hat{\xi}_{t+1}$ as in the following.

$$\hat{\xi}_{t+1} = \frac{\sum_{i=1}^{n} 1 + p_{i2,t} + p_{i3,t}}{\sum_{i=1}^{n} x_i}.$$

Similarly, by solving the likelihood equation $\frac{\partial Q_2(\alpha)}{\partial \alpha} = 0$, we could check that $\hat{\alpha}_{t+1}$ is the positive solution of the following quadratic equation in terms of α

$$\alpha^2(c - 2n) + \alpha(c - n) + c = 0,$$

where $c = \sum_{i=1}^{n} p_{i2,t} + 2p_{i3,t}$. The sequence θ_t converges to θ and we could stop the iterations when $Q(\theta|\theta_t)$ does not improve significantly, i.e. for a predefined small value $\epsilon > 0$, $Q(\theta|\theta_{t+1}) < Q(\theta|\theta_t) + \epsilon$, see Wu [29] for more information about convergence of the EM algorithm.

4. Simulations

The behavior of the estimators is examined and compared in a simulation study. We were able to generate a random sample of $EQL(\alpha, \xi)$ by the following steps:

1. First, drive one random instance from multinomial model with parameters (p_1, p_2, p_3, n) where $p_1 = 1/1 + \alpha + \alpha^2$, $p_2 = \alpha/1 + \alpha + \alpha^2$ and $p_3 = 1 - p_1 - p_2$. Assume the derived instance be (k_1, k_2, k_3).
2. Generate and mix three identical and independent (iid) random samples from $G(1, \xi)$, $G(2, \xi)$ and $G(3, \xi)$ with sizes k_1, k_2 and k_3 respectively.

In each simulation run, $r = 1000$ samples are generated with a size of $n = 80$ or 150. Then, the parameters are estimated for each instance using the ML, LSE, AD methods or EM algorithm. For the calculation of the optimum values of the parameters, the integrated function "optim" of R is used.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 July 2023 doi:10.20944/preprints202307.1629.v1

The initial values needed for computing all estimators are randomly generated from a uniform distribution, e.g., the initial values for α are randomly and uniformly derived from the interval (0.9α, 1.1α). Table 1 shows the bias (B) and mean square error (MSE) for estimators and for some parameter values calculated by the following relations:

$$B_\alpha = \frac{1}{r}\sum_{i=1}^{n}(\hat{\alpha}_i - \alpha),$$

and

$$MSE_\alpha = \frac{1}{r}\sum_{i=1}^{n}(\hat{\alpha}_i - \alpha)^2,$$

and similarly, for ξ. Small values of MSE reported in Table 1 show that all estimators are consistent and sufficiently efficient but EM algorithm outperforms others for all selected parameters.

Table 1. Simulation results for ML, LSE, AD and EM algorithm. The first and second lines of every cell corresponds to α and ξ respectively.

Method		n			
		80		150	
	α, ξ	B	MSE	B	MSE
MLE	0.1, 0.1	0.2093	0.1413	0.1486	0.0853
		0.0221	0.0015	0.0166	0.0010
	0.3, 0.5	0.1009	0.1222	0.0466	0.0771
		0.0415	0.0209	0.0168	0.0133
	0.8, 1	0.0598	0.1716	0.0273	0.0948
		0.0029	0.0326	-0.0043	0.0201
EM	0.1, 0.1	0.0833	0.0377	0.0463	0.0126
		0.0097	0.0004	0.0054	0.0002
	0.3, 0.5	0.1346	0.1095	0.0807	0.0554
		0.0530	0.0181	0.0329	0.0100
	0.8,1	0.1023	0.1856	0.0281	0.0728
		0.0223	0.0299	0.0015	0.0156
LSE	0.1, 0.1	0.2350	0.1970	0.1734	0.1185
		0.0306	0.0026	0.0230	0.0016
	0.3, 0.5	0.0906	0.1726	0.0466	0.1083
		0.0523	0.0315	0.0287	0.0206
	0.8,1	0.0609	0.2960	0.0143	0.1085
		0.0075	0.0500	-0.0038	0.0270
Weighted LSE (AD)	0.1, 0.1	0.0283	0.0592	0.0249	0.0392
		0.0097	0.0009	0.0084	0.0006
	0.3, 0.5	-0.1116	0.1157	-0.1373	0.0778
		-0.0233	0.0211	-0.0367	0.0121
	0.8,1	-0.3220	0.2795	-0.2426	0.1963
		-0.1497	0.0700	-0.1230	0.0466

5. Application

Table 2 shows 29-time intervals between successive air conditioning failures in a Boeing 720 aircraft. For more details about the experiment and the data, see Proschan [30].

 doi:10.20944/preprints202307.1629.v1

Table 2. Time interval between successive failures of air conditioner system of Boeing 720 aircraft.

10	60	186	61	49	14	24	56	20
84	44	59	29	118	25	156	310	76
44	23	62	130	208	70	101	208	

Total time on test time (TTT) is plotted in Figure 3 with an increasing FR function (left). The histogram of the data and the calculated PDF of the EQL are shown on the right.

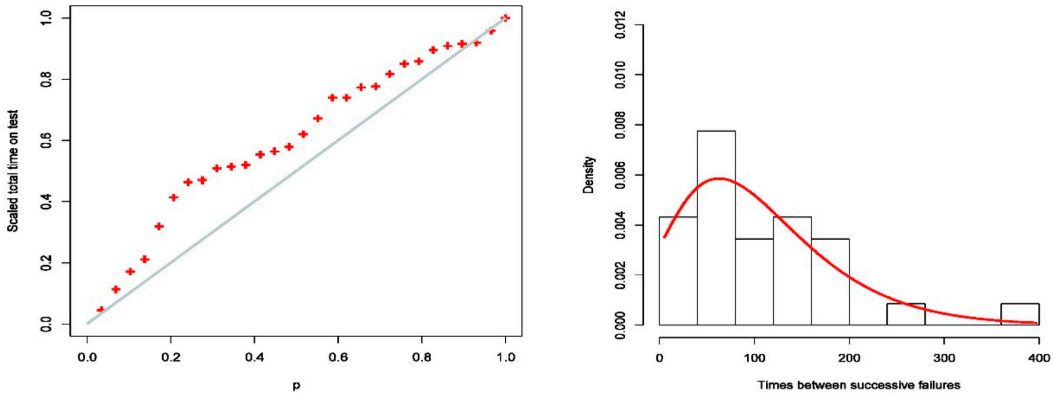

Figure 3. The TTT plot (left) and histogram along with estimated PDF (right) of times between failures of air conditioning system.

The dataset was fitted to the EQL and some alternative models as a comparative analysis. The parameters of the EQL are estimated using the MLE and EM methods. The estimates from EM and ML are approximately the same. The alternative models include the gamma, exponentiated gamma (EG), Lehmann gamma (LG), Marshal-Olkin gamma (MOG), and QL. For each model, the Akaike information criterion (AIC), Cramer-von Mises (CVM) statistics, Anderson-Darling (AD) statistics, and Kolmogorov-Smirnov (KS) statistics are calculated and summarized in Table 3.

Table 3. Fitting the successive times between failures.

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\xi}$	AIC	CVM p-value	AD p-value	KS p-value
EQL	1.9668	—	0.0215	331.22	0.0278 0.9843	0.1833 0.9944	0.0801 0.9923
Gamma	1.7195	—	0.0153	331.55	0.0363 0.9539	0.2399 0.9754	0.1028 0.9190
EG	2.8250	0.0823	0.1459	334.57	0.0647 0.7882	0.3836 0.8638	0.1308 0.7037
LG	1.4504	1.1997	0.0142	333.59	0.0373 0.9682	0.2454 0.9727	0.1041 0.9120
MOG	1.6439	1.2563	0.0161	333.37	0.0322 0.9705	0.2169 0.9851	0.0965 0.9498
QL	0.1382	—	0.0167	331.35	0.0320 0.9712	0.2057 0.9888	0.0965 0.9499

The analysis shows that the EQL performs better than the other models in all the statistics considered. In Figure 4, the empirical and fitted distribution function for EQL and some alternatives are plotted and provides a graphical investigation.

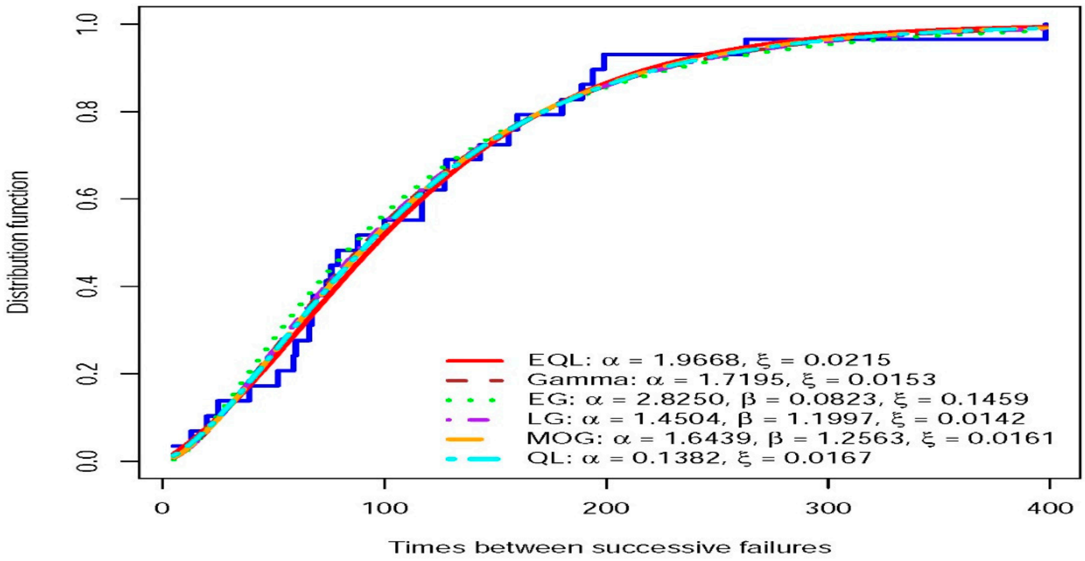

Figure 4. The empirical and estimated CDF for QL and some alternative models of times between failures of air conditioning system.

6. Conclusions

For data modeling and analysis, the right statistical model must be used to draw more accurate conclusions. The EQL model, which combines three gamma distributions, is an extension of QL, which can be used in various scientific disciplines. The model can be useful in practice, as shown by the analysis of a data set consisting of the intervals between successive air conditioning failures of a Boeing 720 aircraft. The ML approach and the EM algorithm provide accurate and consistent parameter estimates based on simulation results. However, the EM algorithm provides a more accurate approximation than the MLE.

Acknowledgments: This work is supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R226), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lindley, D.V. Fiducial distributions and Bayes' theorem. J. R. Stat. Soc. Ser. (Methodol.), **20** (1958), 102-107.
2. Ghitany, M.E.; Atieh, B.; Nadarajah, S. Lindley distribution and its application. Math. Comput. Simul. **78** (2008), 493–506.
3. Shanker, R.; Ghebretsadik, A.H. A New Quasi Lindley Distribution. Int. J. Stat. Syst. **8** (2013), 143-156.
4. Zakerzadeh, H.; Dolati, A. Generalized Lindley distribution. J. Math. Ext. **3** (2009), 13-25.
5. Shanker, R.; Shukla, K.K.; Shanker, R.; Leonida, T.A. A Three-Parameter Lindley Distribution. Am. J. Math. Stat. **7** (2017), 15-26.
6. Merovci, F.; Sharma, V.K. The Beta-Lindley Distribution: Properties and Applications. J. Appl. Math. **2014** (2014), 198951.
7. Ibrahim, E.; Merovci, F.; Elgarhy, M. A new generalized Lindley distribution. Math. Theory Model. **3** (2013), 30-47.
8. Sankaran, M. The discrete poisson-Lindley distribution. Biometrics, **26** (1970), 145-149.

9. Ghitany, M.E.; Al-Mutairi, D.K.; Balakrishnan, N.; Al-Enezi, L.J. Power Lindley distribution and associated inference. Comput. Stat. Data Anal. **64** (2013), 20-33.
10. Al-Mutairi, D.K.; Ghitany, M.E.; Kundu, D. Inferences on stress-strength reliability from Lindley distribution. Commun. Stat.-Theory Methods. **42** (2013), 1443-1463.
11. Zamani, H.; Ismail, N. Negative Binomial-Lindley Distribution and Its Application. J. Math. Stat. **6** (2010), 4-9.
12. Al-babtain, A.A.; Eid, H.A.; A-Hadi, N.A.; Merovci, F. The five parameter Lindley distribution. Pak. J. Stat. **31** (2014), 363-384.
13. Ghitany, M.E.; Al-Mutairi, D.K.; Aboukhamseen, S.M. Estimation of the reliability of a stress-strength system from power Lindley distributions. Commun. Stat.-Simul. Comput. **44** (2015), 118-136.
14. Abouammoh, A.M.; Alshangiti Arwa, M.; Ragab, I.E. A new generalized Lindley distribution. J. Stat. Comput. Simul. **85** (2015), 3662-3678.
15. Ibrahim, M.; Singh Yadav, A.; Yousof, H.M.; Goual, H.; Hamedani, G.G. A new extension of Lindley distribution: modified validation test, characterizations and different methods of estimation. Commun. Stat. Appl. Methods. **26** (2019), 473-495.
16. Marthin, P.; Rao, G.S. Generalized Weibull-Lindley (GWL) distribution in modeling lifetime Data. J. Math, **2020** (2020), 1-15.
17. Al-Babtain, A.A.; Ahmed A.H.N.; Afify, A.Z. A new discrete analog of the continuous Lindley distribution, with reliability applications. Entropy, **22** (2020), 1-15.
18. Joshi, R.K.; Kumar, V. Lindley Gompertz distribution with properties and applications. Int. J. Appl. Math. Stat. **5** (2020), 28-37.
19. Afify, A.Z.; Nassar, M.; Cordeiro, G.M.; Kumar, D. The Weibull Marshall and Olkin Lindley distribution: properties and estimation. J. Taibah Univ. Sci., **14** (2020), 192-204.
20. Chesneau, C.; Tomy, L.; Gillariose, J.; Jamal, F. The Inverted Modified Lindley Distribution. J Stat. Theory Pract. **14** (2020.
21. Algarni, A. On a new generalized lindley distribution: Properties, estimation and applications. PLoS ONE, **16** (2021), 1-18.
22. Shanker, R.; Mishra, A. A quasi Lindley distribution. Afr. J. Math. Comput. Sci. Res. **6** (2013) 64-71.
23. Kayid, M.; Nassr S. Al-Maflehi, EM Algorithm for Estimating the Parameters of Quasi-Lindley Model with Application, Journal of Mathematics, **2022** (2022), 1-9.
24. Kayid, M.; Alskhabrah, R.; Alshangiti, A.M. A New Scale-Invariant Lindley Extension Distribution and Its Applications, Mathematical Problems in Engineering, **2021**, Article ID 3747753, 1-16.
25. Alrasheedi, A., Abouammoh, A. and Kayid, M. A new flexible extension of the Lindley distribution with applications. Journal of King Saud University - Science, **34 2022**, 1-9.
26. Titterington, D.M.; Smith, A.F.M.; Makov, U.E. Statistical analysis of finite mixture distributions. Chichester, U.K.: John Wiley and Sons, 1985.
27. Ord, J.K. Families of frequency distributions, London, U.K.: Charles Griffin, 1972.
28. Lai, C.D.; Xie, M. Stochastic Aging and Dependence for Reliability. Springer: 2006. ISBN: 978-0-387-29742-2.
29. Wu, C.F.J. On the convergence properties of the EM algorithm. Ann. Stat., **11** (1983), 95-103.
30. Proschan, F. Theoretical Explanation of Observed Decreasing Failure Rate. Technometrics, **5** (1963), 375-383.