On Discrimination Between the Lindley and xgamma Distributions

Subhradev Sen 1 · Hazem Al-Mofleh 2 · Sudhansu S. Maiti 3

Received: 10 October 2018 / Revised: 29 December 2019 / Accepted: 18 January 2020 / Published online: 29 January 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
For a given data set the problem of selecting either Lindley or xgamma distribution with unknown parameter is investigated in this article. Both these distributions can be used quite effectively for analyzing skewed non-negative data and in modeling time-to-event data sets. We have used the ratio of the maximized likelihoods in choosing between the Lindley and xgamma distributions. Asymptotic distributions of the ratio of the maximized likelihoods are obtained and those are utilized to determine the minimum sample size required to discriminate between these two distributions for user specified probability of correct selection and tolerance limit.

Keywords Asymptotic distributions · Kolmogorov–Smirnov distance · Likelihood ratio statistic · Life distributions

Mathematics Subject Classification 62E15 · 62E25 · 62F10

1 Introduction

It is an important problem in statistics to test whether some given observations, in view of modeling, follow one of the two probability distributions. If the two distribution possess similar structural, distributional and/or survival properties, then it is quite
reasonable to construct a test procedure to determine which particular distribution need to be selected in describing the data set coming from diverse areas of application.

For last one decade, Lindley distribution [15] has drawn attention of the researchers for modeling survival or reliability data sets. Its properties are explicitly studied by Ghitany et al. [12] and is shown quite flexible in modeling time-to-event data sets. Several authors have suggested significant extensions and variations of Lindley model, see for example, Sankaran [19], Gomez and Ojeda [13], Nadarajah et al. [17], Bakouch et al. [4], Ghitany et al. [11], Shanker et al. [24], Merovci and Sharma [16], Nedjar and Zeghdoudi [18], Shibu and Irshad [25], Asgharzadeh [1] and references therein for more details.

Recently, Sen et al. [23] introduced and studied xgamma distribution and applied it in describing survival/reliability data sets. The xgamma distribution has properties analogous to Lindley distribution and has similar mathematical form. For more better flexibility and ease of application, few extensions of xgamma distribution are been proposed in the literature (see Sen and Chandra [20]; Sen et al. [21]). However, the xgamma random variables are stochastically larger than those of Lindley (see Sen et al. [22]), both the distributions are the special finite mixtures of exponential and gamma distributions and both can effectively be utilized in analyzing positively skewed data sets.

We address the following problem in this article. Suppose an experimenter has observed \( n \) data points, say \( x_1, x_2, \ldots, x_n \) and he wants to use either one parameter Lindley model or one parameter xgamma model, which one will he prefer?

The problem of testing whether some given observations follow one of the two probability distributions, is quite old in the statistical literature (see Cox [7,8]; Atkinson [2,3]; Chambers and Cox [5]; Chen [6] and Dyer [10] for more details on this). We consider in this investigation the problem of discriminating between the Lindley and xgamma distributions. We use the ratio of maximized likelihood in discriminating between the two distribution functions. We obtain the asymptotic distribution of the natural logarithm of RML following the approach of White [26,27]. It is observed that the asymptotic distribution is normal and independent of unknown parameters. The asymptotic distribution can be utilized to compute the probability of correct selection (PCS), hence we also attempted to find minimum sample size required to discriminate between the two distributions for a given value of PCS.

The rest of the article is organized as follows. The method of likelihood ratio is described in Sect. 2. Asymptotic distributions of the logarithm of RML statistics under the null hypothesis are obtained in Sect. 3. Section 4 deals with the determination of minimum sample sizes. Some Monte-Carlo simulation studies are performed in Sect. 5 and real life data sets are analyzed as an illustration in Sect. 6. Finally, Sect. 7 concludes.

### 2 Ratio of Maximized Likelihoods

Let \( X_1, X_2, \ldots, X_n \) be independent and identically distributed (i.i.d) random variables from a Lindley or from an xgamma distribution function. We shall use the following notations throughout the article:
The probability density function (pdf) of a Lindley distribution with parameter $\lambda$ is denoted by

$$f_{LD}(x; \lambda) = \frac{\lambda^2}{(1 + \lambda)} (1 + x) e^{-\lambda x}, \quad x > 0, \lambda > 0. \quad (1)$$

We shall denote it by $X \sim LD(\lambda)$. The pdf of xgamma distribution with parameter $\theta$ is denoted by

$$f_{XG}(x; \theta) = \frac{\theta^2}{(1 + \theta)} \left(1 + \frac{\theta}{2}x^2\right) e^{-\theta x}, \quad x > 0, \theta > 0. \quad (2)$$

We shall denote it by $X \sim XG(\theta)$. Assuming the data coming from $LD(\lambda)$ or $XG(\theta)$, the likelihood functions are

$$L_{LD}(\lambda) = \prod_{i=1}^{n} f_{LD}(x_i; \lambda) \quad \text{and} \quad L_{XG}(\theta) = \prod_{i=1}^{n} f_{XG}(x_i; \theta),$$

respectively.

The ratio of maximized likelihood (RML) is defined as

$$L = \frac{L_{LD}(\hat{\lambda})}{L_{XG}(\hat{\theta})}, \quad (3)$$

where $\hat{\lambda}$ and $\hat{\theta}$ are the maximum likelihood estimators of $\lambda$ and $\theta$, respectively, based on $\{X_1, X_2, \ldots, X_n\}$. Taking natural logarithm of the RML, we have the log-likelihood ratio statistic as

$$T = \ln \left[ \frac{L_{LD}(\hat{\lambda})}{L_{XG}(\hat{\theta})} \right] = \ln \left( \frac{\hat{\lambda}}{\hat{\theta}} \right)^2 + \ln \left( \frac{1 + \hat{\theta}}{1 + \hat{\lambda}} \right) + (\hat{\theta} - \hat{\lambda}) \bar{X}$$

$$+ \frac{1}{n} \left[ \sum_{i=1}^{n} \ln(1 + X_i) - \sum_{i=1}^{n} \ln \left(1 + \frac{\hat{\theta}}{2}X_i^2\right) \right]. \quad (4)$$

Here $\bar{X}$ is the arithmetic mean of $\{X_1, X_2, \ldots, X_n\}$. It is to be noted that, in case of Lindley distribution,

$$\hat{\lambda} = \frac{-(\bar{X} - 1) + \sqrt{(\bar{X} - 1)^2 + 8\bar{X}}}{2\bar{X}} \quad (5)$$

and in case of xgamma distribution, $\hat{\theta}$ is the solution of the equation

$$\frac{(2 + \theta)}{\theta(1 + \theta)} + \frac{1}{n} \sum_{i=1}^{n} \frac{X_i^2}{2 (1 + \frac{\theta}{2}X_i^2)} = \bar{X}. \quad (6)$$

It is clear from (4) that distributions of $T$’s are independent of the parameters (see Dumonceaux et al. [9]).
The following procedure can be used to discriminate between Lindley and xgamma distributions.

1. Choose Lindley distribution if $T > 0$.
2. Choose xgamma distribution if $T < 0$.

### 3 Asymptotic Properties of RML

In this section we obtain the asymptotic distributions of RML for two different cases. Let us denote the almost sure convergence by $a.s.$ hereafter. The following notations are used.

For any Borel measurable function $h(\cdot)$, we denote,

\[ E_{LD}(h(Y)) : \text{mean of } h(Y) \text{ under the assumption that } Y \sim LD(\lambda), \]

and

\[ V_{LD}(h(Y)) : \text{variance of } h(Y) \text{ under the assumption that } Y \sim LD(\lambda). \]

Similarly, we define $E_{XG}(h(Y))$ and $V_{XG}(h(Y))$ as mean and variance of $h(Y)$, respectively, under the assumption that $Y \sim XG(\theta)$. Moreover, if $h(\cdot)$ and $g(\cdot)$ are two Borel measurable functions, then we define along the same line $\text{Cov}_{LD}[g(Y), h(Y)] = E_{LD}[g(Y)h(Y)] - E_{LD}[g(Y)]E_{LD}[h(Y)]$ and also $\text{Cov}_{XG}[g(Y), h(Y)]$ similarly, where $Y \sim LD(\lambda)$ and $Y \sim XG(\theta)$, respectively.

#### Case 1: Data Are Coming from Lindley Distribution

We assume that $X_1, X_2, \ldots, X_n$ are $n$ data points coming from $LD(\lambda)$. For proving the main result, we take the following lemma.

**Lemma 1** Under the assumption that the data are coming from $LD(\lambda)$, as $n \to \infty$, we have

(i) $\hat{\lambda} \to \lambda$ a.s., where

\[ E_{LD} \left[ \ln f_{LD}(X; \lambda) \right] = \max_{\tilde{\lambda}} E_{LD} \left[ \ln f_{LD}(X; \tilde{\lambda}) \right]. \]

(ii) $\hat{\theta} \to \tilde{\theta}$ a.s., where

\[ E_{XG} \left[ \ln f_{XG}(X; \tilde{\theta}) \right] = \max_{\theta} E_{XG} \left[ \ln f_{XG}(X; \theta) \right]. \]

We denote:

\[ T^* = \ln \left[ \frac{L_{LD}(\lambda)}{L_{XG}(\tilde{\theta})} \right]. \]

(iii) $\frac{1}{\sqrt{n}} \{T - E_{LD}(T)\}$ is asymptotically equivalent to $\frac{1}{\sqrt{n}} \{T^* - E_{LD}(T^*)\}$

**Proof** Applying the similar argument as in theorem 1 in White [27] we can easily get the proof and hence is omitted here. \qed

We have the following theorem.
Theorem 1 Under the assumption that the data are coming from $LD(\theta)$, $T$ is asymptotically normally distributed with mean $ELD(T)$ and variance $VLD(T) = VLD(T^*)$.

Proof By central limit theorem (CLT) and from part (ii) of Lemma 1, it is clear that $\frac{1}{\sqrt{n}} \{T^* - ELD(T^*)\}$ is asymptotically normally distributed with mean 0 and variance $VLD(T^*)$. Again by the part (iii) of Lemma 1, we have $T$ is asymptotically normally distributed with mean $ELD(T)$ and variance $VLD(T^*)$. Hence the proof of the theorem.

Now, we shall find $\tilde{\theta}$, $ELD(T)$ and $VLD(T)$. We define,

$$g(\theta) = ELD[\ln fXG(X; \theta)] = 2 \ln \theta - \ln(1 + \theta) + ELD[\ln \left(1 + \frac{\theta}{2}X^2\right)] - \frac{\theta(2 + \lambda)}{\lambda(1 + \lambda)}.$$

Therefore, $\tilde{\theta}$ can be obtained by maximizing $g(\theta)$ with respect to $\theta$ or as a solution of

$$2\tilde{\theta} - \frac{1}{1 + \tilde{\theta}} + ELD \left[ \frac{X^2}{1 + \frac{\tilde{\theta}}{2}X^2} \right] - \frac{2 + \lambda}{\lambda(1 + \lambda)} = 0.$$

It should be noted that, $\tilde{\theta}$ depends on $\lambda$, for brevity we do not make it explicit. For further development we compute $\tilde{\theta}$ for $\lambda = 1$, and we denote it by $\tilde{\theta}_1$. As it is seen that it is difficult to obtain an analytic solution, we solve it numerically.

It is to be noted that $\lim_{n \to \infty} ELD(T/n)$ and $\lim_{n \to \infty} VLD(T/n)$ exist.

We denote,

$$\lim_{n \to \infty} \frac{ELD(T)}{n} = AMLD(\lambda)$$

and

$$\lim_{n \to \infty} \frac{VLD(T)}{n} = AVLD(\lambda),$$

respectively.

Therefore, for large $n$,\n
$$\frac{ELD(T)}{n} \approx AMLD(\lambda) = ELD \left[ \ln fLD(X; \lambda) - \ln fXG(X; \tilde{\theta}) \right].$$

Hence,\n
$$AMLD(\lambda) = \ln \left(\frac{\lambda}{\tilde{\theta}}\right)^2 + \ln \left(1 + \frac{\tilde{\theta}}{1 + \lambda}\right) + (\tilde{\theta} - \lambda) \frac{(2 + \lambda)}{\lambda(1 + \lambda)} + ELD[\ln(1 + X)]$$

$$- ELD \left[ \ln \left(1 + \frac{\tilde{\theta}}{2}X^2\right) \right].$$

and we also have, for large $n$,
\[
\frac{V_{LD}(T)}{n} \approx AV_{LD}(\lambda) = V_{LD}\left[\ln f_{LD}(X; \lambda) - \ln f_{XG}(X; \tilde{\theta})\right] \\
= V_{LD}\left[\ln \left(\frac{\lambda}{\tilde{\theta}}\right)^2 + \ln \left(\frac{1 + \tilde{\theta}}{1 + \lambda}\right) + (\tilde{\theta} - \lambda)X + \ln(1 + X) - \ln \left(1 + \frac{\tilde{\theta}}{2}X^2\right)\right].
\]

Therefore,
\[
AV_{LD}(\lambda) = (\tilde{\theta} - \lambda)^2 \frac{(\lambda^2 + 4\lambda + 2)}{\lambda^2(1 + \lambda)^2} + V_{LD}[\ln(1 + X)] + V_{LD}\left[\ln \left(1 + \frac{\tilde{\theta}}{2}X^2\right)\right] \\
+ (\tilde{\theta} - \lambda) \left[Cov_{LD}\{X, \ln(1 + X)\} - Cov_{LD} \left\{X, \ln \left(1 + \frac{\tilde{\theta}}{2}X^2\right)\right\}\right] \\
- Cov_{LD}\left[\ln(1 + X), \ln \left(1 + \frac{\tilde{\theta}}{2}X^2\right)\right].
\]

(10)

**Case 2: Data are Coming from xgamma Distribution**

In this case, we assume that \(X_1, X_2, \ldots, X_n\) are \(n\) data points coming from xgamma distribution with parameter \(\theta\). We have the following lemma.

**Lemma 2** Under the assumption that the data are coming from \(XG(\theta)\), as \(n \to \infty\), we have

(i) \(\hat{\theta} \to \theta\) a.s., where
\[
E_{XG}\left[\ln f_{XG}(X; \theta)\right] = \max_{\hat{\theta}} E_{XG}\left[\ln f_{XG}(X; \hat{\theta})\right].
\]

(ii) \(\hat{\lambda} \to \tilde{\lambda}\) a.s., where
\[
E_{LD}\left[\ln f_{LD}(X; \tilde{\lambda})\right] = \max_{\lambda} E_{LD}\left[\ln f_{LD}(X; \lambda)\right].
\]

We denote
\[
T_* = \ln \left[\frac{L_{LD}(\tilde{\lambda})}{L_{XG}(\theta)}\right].
\]

(iii) \(\frac{1}{\sqrt{n}}[T - E_{XG}(T)]\) is asymptotically equivalent to \(\frac{1}{\sqrt{n}}[T_* - E_{XG}(T_*)]\).

**Proof** Proof comes applying the similar argument of White [27] in theorem 1, hence is omitted.  

We have the following theorem in this case.
Theorem 2  Under the assumption that data coming from $X \sim G(\theta)$, $T$ is asymptotically normally distributed with mean $E_{XG}(T)$ and variance $V_{XG}(T) = V_{XG}(T_*)$.

Proof  The proof comes with similar argument as given in Theorem 1.

Now, we shall obtain $\tilde{\lambda}$, $E_{XG}(T)$ and $V_{XG}(T)$.

Let us define,

$$h(\lambda) = E_{XG}[\ln f_{LD}(X; \lambda)] = 2\ln \lambda - \ln(1 + \lambda) + E_{XG}[\ln(1 + X)] - \lambda \frac{(3 + \theta)}{\theta(1 + \theta)}$$

Therefore, $\tilde{\lambda}$ can be obtained by maximizing $h(\lambda)$ with respect to $\lambda$ or as a solution of

$$\frac{2}{\tilde{\lambda}} - \frac{1}{1 + \tilde{\lambda}} - \frac{3 + \theta}{\theta(1 + \theta)} = 0. \quad (11)$$

It is noted that $\tilde{\lambda}$ depends on $\theta$, we do not make it explicit for brevity. For further development we compute $\tilde{\lambda}$ for $\theta = 1$, and we denote it by $\tilde{\lambda}_1$. It is difficult to obtain an analytic solution, so we solve it numerically.

We note that, $\lim_{n \to \infty} \frac{E_{XG}(T)}{n}$ and $\lim_{n \to \infty} \frac{V_{XG}(T)}{n}$ exist. We denote, with a similar fashion as before, $\lim_{n \to \infty} \frac{E_{XG}(T)}{n} = AM_{XG}(\theta)$ and $\lim_{n \to \infty} \frac{V_{XG}(T)}{n} = AV_{XG}(\theta)$, respectively.

We have, for large $n$,

$$\frac{E_{XG}(T)}{n} \approx AM_{XG}(\theta) = E_{XG}\left[\ln f_{LD}(X; \tilde{\lambda}) - \ln f_{XG}(X; \theta)\right],$$

so that

$$AM_{XG}(\theta) = \ln \left(\frac{\tilde{\lambda}}{\theta}\right)^2 + \ln \frac{1 + \theta}{1 + \tilde{\lambda}} + \frac{(3 + \theta)}{\theta(1 + \theta)} + E_{XG}[\ln(1 + X)]$$

$$- E_{XG}\left[\ln \left(1 + \frac{\theta}{2}X^2\right)\right]. \quad (12)$$

Again, for large $n$, we have

$$\frac{V_{XG}(T)}{n} \approx AV_{XG}(\theta) = V_{XG}\left[\ln f_{LD}(X; \tilde{\lambda}) - \ln f_{XG}(X; \theta)\right],$$

to obtain

$$AV_{XG}(\theta) = (\theta - \tilde{\lambda})^2 \frac{(\theta^2 + 8\theta + 3)}{\theta^2(1 + \theta)^2} + V_{XG}[\ln(1 + X)] + V_{XG}\left[\ln \left(1 + \frac{\theta}{2}X^2\right)\right]$$

$$+ (\theta - \tilde{\lambda}) \left\{ Cov_{XG}\{X, \ln(1 + X)\} - Cov_{XG}\left\{X, \ln \left(1 + \frac{\theta}{2}X^2\right)\right\} \right\}$$

$$- Cov_{XG}\left\{\ln(1 + X), \ln \left(1 + \frac{\theta}{2}X^2\right)\right\}. \quad (13)$$
Table 1  The values of $AM_{LD}(\lambda)$, $AV_{LD}(\lambda)$ and $\tilde{\theta}$ for Different $\lambda$

| $\lambda$ | $AM_{LD}(\lambda)$ | $AV_{LD}(\lambda)$ | $\tilde{\theta}$ |
|-----------|---------------------|---------------------|-------------------|
| 0.45      | 0.00794             | 0.01582             | 0.59983           |
| 0.55      | 0.00738             | 0.01450             | 0.72360           |
| 0.65      | 0.00699             | 0.01363             | 0.84547           |
| 0.70      | 0.00684             | 0.01330             | 0.90578           |
| 0.75      | 0.00670             | 0.01301             | 0.96574           |
| 0.78      | 0.00663             | 0.01286             | 1.00154           |
| 0.89      | 0.00639             | 0.01237             | 1.13186           |
| 0.90      | 0.00637             | 0.01233             | 1.14363           |
| 1.15      | 0.00593             | 0.01150             | 1.43473           |
| 1.16      | 0.00591             | 0.01147             | 1.44626           |
| 1.37      | 0.00561             | 0.01090             | 1.68648           |
| 1.38      | 0.00559             | 0.01087             | 1.69784           |

Table 2  The values of $AM_{XG}(\theta)$, $AV_{XG}(\theta)$ and $\tilde{\lambda}$ for Different $\theta$

| $\theta$ | $AM_{XG}(\theta)$ | $AV_{XG}(\theta)$ | $\tilde{\lambda}$ |
|----------|-------------------|-------------------|-------------------|
| 0.85     | −0.00718          | 0.01480           | 0.65520           |
| 0.90     | −0.00706          | 0.01456           | 0.69686           |
| 1.05     | −0.00674          | 0.01392           | 0.82302           |
| 1.10     | −0.00665          | 0.01372           | 0.86544           |
| 1.25     | −0.00639          | 0.01317           | 0.99369           |
| 1.26     | −0.00637          | 0.01314           | 1.00229           |
| 1.40     | −0.00616          | 0.01267           | 1.12328           |
| 1.50     | −0.00601          | 0.01236           | 1.21035           |
| 1.65     | −0.00580          | 0.01191           | 1.34185           |
| 1.80     | −0.00561          | 0.01149           | 1.47435           |
| 2.00     | −0.00536          | 0.01096           | 1.65242           |
| 2.05     | −0.00530          | 0.01084           | 1.69716           |

Note that $\tilde{\lambda}$, $\tilde{\theta}$, $AM_{LD}(\lambda)$, $AV_{LD}(\lambda)$, $AM_{XG}(\theta)$ and $AV_{XG}(\theta)$ are numerically computed by using R 3.5.1 programming language (R Core Team, 2018). Tables 1 and 2 display the values of $AM_{LD}(\lambda)$, $AV_{LD}(\lambda)$ and $\tilde{\theta}$ for different $\lambda$, and the values of $AM_{XG}(\theta)$, $AV_{XG}(\theta)$ and $\tilde{\lambda}$ for different $\theta$, respectively.

4 Sample Size Determination

In this section, we shall discuss a procedure to determine the minimum sample size required to discriminate between the Lindley and xgamma distributions, for a pre-specified user specific probability of correct selection (PCS).

It is important to know the distance (or closeness or proximity) between the two distributions, under consideration, prior to establish a discrimination strategy between them. There are several ways to measure the distance between the two probability
distributions, such as, Kolmogorov–Smirnov (K–S) distance or Helinger distance. It is quite natural that, for a given PCS, a very large sample is required to discriminate between the distributions that are very close.

On the other hand, a small sample may be adequate to discriminate between the two distributions that are not so close. However, it is also logical that, if the two distributions under consideration are very close then one may not need to differentiate between them from a practical angle. So, it is expected that a practitioner will pre-specify the PCS and a tolerance limit in view of the distance between the two probability distributions. The tolerance limit indicates that an user does not want to make a distinction between two probability distributions if the distance measure between them is less than the tolerance limit.

We can determine the minimum sample size required to discriminate between two distributions based on PCS and the tolerance limit. We apply here the very popular K–S distance to discriminate between two distribution functions.

We have observed in Sect. 4, that for a large $n$ RML statistics follow approximately normal distribution. We use this fact along with the help of K–S distance to determine the minimum sample size required such that PCS attains a certain protection level, say $p^*$, for a given tolerance level $D^*$. The K–S distance between two distribution functions, say $F(x)$ and $G(x)$, is defined as

$$\sup_{-\infty < x < \infty} |F(x) - G(x)|.$$  

(14)

The method is explained below for Case 1, the procedure for Case 2 follows exactly along the same line may be noted also. We have seen that $T$ is asymptotically normally distributed with mean $E_{LD}(T)$ and variance $V_{LD}(T)$, therefore, we have the probability of correct selection (PCS) as

$$PCS(\lambda) = Pr \{ T > 0 | \lambda \} \approx 1 - \Phi \left( \frac{-E_{LD}(T)}{\sqrt{V_{LD}(T)}} \right) = 1 - \Phi \left( \frac{-n \times AM_{LD}(\lambda)}{\sqrt{n \times AV_{LD}(\lambda)}} \right).$$  

(15)

Here $\Phi(.)$ denotes the distribution function of the standard normal variable. To determine the sample size required to attain at least a protection level $p^*$, we equate,

$$\Phi \left( \frac{-n \times AM_{LD}(\lambda)}{\sqrt{n \times AV_{LD}(\lambda)}} \right) = 1 - p^*$$

to solve for $n$ and it provides

$$n = \frac{z_{p^*}^2 \times AV_{LD}(\lambda)}{[AM_{LD}(\lambda)]^2}.$$  

(16)

Here $z_{p^*}$ is the 100$p^*$ percentile point of a standard normal distribution. $AM_{LD}(\theta)$ and $AV_{LD}(\lambda)$ are same as defined before. For $p^* = 0.90$ and for different values of $\lambda$, the possible values for $n$ is reported in Table 3.
Along with the similar argument, the sample size $n$ required for Case 2 is obtained as

$$n = \frac{z_{p^*}^2 \times AV_{XG}(\theta)}{[AM_{XG}(\theta)]^2}. \quad (17)$$

For $p^* = 0.90$ and for different values of $\theta$, the possible values for $n$ is reported in Table 4. From Tables 3 and 4 it is immediate that as $\lambda$ moves away from 0.78 and $\theta$ moves away from 1.26, for a given PCS, the required sample size decreases as expected.

Suppose that one would like to choose the minimum sample size needed for a given protection level $p^*$ when the distance between two distribution functions is greater than a given tolerance level $D^*$. We report K–S distance between $LD(\lambda)$ and $XG(\tilde{\theta})$ for different values of $\lambda$ is reported in Table 3. Here $\tilde{\theta}$ is same as defined in (8) and it has been reported in Table 1. Similarly, K–S distance between $XG(\theta)$ and $LD(\tilde{\lambda})$ for different values of $\theta$ is reported in Table 4. Here $\tilde{\lambda}$ is same as defined in (11) and it has been reported in Table 2.

Now, to determine the minimum sample size required to discriminate between Lindley and xgamma distributions for given $p^*$ and $D^*$. Suppose $p^* = 0.90$ and $D^* = 0.03$. Here tolerance level $D^* = 0.03$ means that the practitioner wants to discriminate between a Lindley distribution function and a xgamma distribution function only when their K–S distance is more than 0.03. From Table 3, it is clear that for Case 1, K–S distance will be more than 0.03 if $0.89 \leq \lambda \leq 1.38$. Similarly, from Table 4, it is clear that for Case 2, K–S distance will be more than 0.03 if $1.10 \leq \theta \leq 2.05$. Therefore, if the null distribution is Lindley, then for the tolerance level $D^* = 0.03$, one needs $n = \max\{143, 9\} = 143$ to meet the PCS when $p^* = 0.90$.

Similarly, if the null distribution is xgamma then one needs $n = \max\{137, 13\} = 137$ to meet the PCS when $p^* = 0.90$ and $D^* = 0.03$. Therefore, for the given tolerance level 0.03 one needs $n = \max\{143, 137\} = 143$ to meet the protection level $p^* = 0.90$ simultaneously for both the cases.

5 Simulation Studies

In this section, we present some simulation results to investigate the behavior of these asymptotic results derived in Sect. 4 for finite sample sizes. All simulations are performed using the R 3.5.1 programming language (R Core Team 2018). These R codes are available to the reader from the authors.

We consider $n = 20, 40, 60, 80, 100, 400$, and for Case 1, the null distribution is Lindley and the alternative is xgamma, we consider $\lambda = 0.45, 0.55, 0.65, 0.70, 0.75, 0.78, 0.89, 0.90, 1.15, 1.16, 1.37, 1.38$, for a fixed $\lambda$ and $n$, we generate a random sample of size $n$ from $LD(\lambda)$ and we replicate this process 25,000 times.

Similarly, for Case 2, the null distribution is xgamma and the alternative is Lindley, we consider $\theta = 0.85, 0.90, 1.05, 1.10, 1.25, 1.26, 1.40, 1.50, 1.65, 1.80, 2.00, 2.05$, for a fixed $\theta$ and $n$, we generate a random sample of size $n$ from $XG(\theta)$ and we replicate this process 25,000 times.
Table 3  The minimum sample size in (16), for $p^* = 0.90$, the null distribution is Lindley, the K–S distance between $LD(\lambda)$ and $XG(\tilde{\theta})$ for different values of $\lambda$

| $\lambda$ | 0.45 | 0.55 | 0.65 | 0.70 | 0.75 | 0.78 | 0.89 | 0.90 | 1.15 | 1.16 | 1.37 | 1.38 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 15 | 33 | 96 | 196 | 400 | 481 | 143 | 127 | 21 | 20 | 10 | 9 |
| $K – S$ | 0.02158 | 0.02520 | 0.02776 | 0.02867 | 0.02943 | 0.02983 | 0.03098 | 0.03106 | 0.03196 | 0.03193 | 0.03118 | 0.03086 |
Table 4  The minimum sample size in (17), for $p^* = 0.90$, the null distribution is xgamma, the K–S distance between $X G(\theta)$ and $LD(\tilde{\lambda})$ for different values of $\theta$

| $\theta$ | 0.85 | 0.90 | 1.05 | 1.10 | 1.25 | 1.26 | 1.40 | 1.50 | 1.65 | 1.80 | 2.00 | 2.05 |
|----------|------|------|------|------|------|------|------|------|------|------|------|------|
| $n$      | 23   | 30   | 87   | 137  | 525  | 532  | 191  | 89   | 41   | 24   | 14   | 13   |
| $K – S$  | 0.02674 | 0.02771 | 0.02964 | 0.03007 | 0.03099 | 0.03107 | 0.03104 | 0.03155 | 0.03175 | 0.03183 | 0.03164 | 0.03102 | 0.03087 |
Table 5 The PCS based on Monte-Carlo Simulations and the asymptotic results (in parenthesis) when the null distribution is Lindley, for different values of $\lambda$

| $\lambda \downarrow n \rightarrow$ | 20     | 40     | 60     | 80     | 100    | 400    |
|----------------------------------|--------|--------|--------|--------|--------|--------|
| 0.45                             | 0.62020| 0.65264| 0.68804| 0.71888| 0.73360| 0.89796|
|       (0.61118)                  |        | (0.65519)| (0.68763)| (0.71389)| (0.73613)| (0.89669)|
| 0.55                             | 0.61276| 0.65620| 0.68752| 0.71496| 0.72740| 0.89004|
|       (0.60797)                  |        | (0.65083)| (0.68248)| (0.70818)| (0.72999)| (0.88982)|
| 0.65                             | 0.61848| 0.65288| 0.67788| 0.71256| 0.73324| 0.88344|
|       (0.60560)                  |        | (0.64759)| (0.67866)| (0.70392)| (0.72540)| (0.88452)|
| 0.70                             | 0.61528| 0.65452| 0.68052| 0.70812| 0.72808| 0.88052|
|       (0.60459)                  |        | (0.64621)| (0.67703)| (0.70211)| (0.72344)| (0.88223)|
| 0.75                             | 0.61896| 0.65672| 0.67788| 0.70492| 0.72392| 0.88052|
|       (0.60366)                  |        | (0.64494)| (0.67553)| (0.70044)| (0.72164)| (0.88009)|
| 0.80                             | 0.61876| 0.65524| 0.67648| 0.70312| 0.72180| 0.87340|
|       (0.60313)                  |        | (0.64422)| (0.67468)| (0.69949)| (0.72061)| (0.87886)|
| 0.85                             | 0.61180| 0.65256| 0.67372| 0.70356| 0.72016| 0.87680|
|       (0.60135)                  |        | (0.64178)| (0.67179)| (0.69626)| (0.71712)| (0.87464)|
| 0.90                             | 0.61120| 0.65236| 0.67688| 0.70064| 0.72020| 0.87176|
|       (0.60120)                  |        | (0.64157)| (0.67154)| (0.69598)| (0.71682)| (0.87428)|
| 1.15                             | 0.61580| 0.64408| 0.67444| 0.69276| 0.70944| 0.86296|
|       (0.59767)                  |        | (0.63674)| (0.66581)| (0.68958)| (0.70988)| (0.86565)|
| 1.16                             | 0.61856| 0.64412| 0.67528| 0.69528| 0.71308| 0.86432|
|       (0.59754)                  |        | (0.63656)| (0.66560)| (0.68934)| (0.70962)| (0.86532)|
| 1.37                             | 0.60564| 0.63624| 0.67340| 0.68640| 0.70624| 0.86104|
|       (0.59489)                  |        | (0.63293)| (0.66128)| (0.68449)| (0.70436)| (0.85859)|
| 1.38                             | 0.60676| 0.64320| 0.66784| 0.68536| 0.70592| 0.85880|
|       (0.59477)                  |        | (0.63276)| (0.66108)| (0.68427)| (0.70412)| (0.85827)|

For each case and replication, we calculate the PCS based on asymptotic results derived in Sect. 4, furthermore, we calculate it as derived in Sect. 5. The results are reported in Tables 5 and 6.

It is clear from Tables 5 and 6 that as the sample size increases the PCS increases as expected. Also, for a small sample size, $n = 20$, the asymptotic results work quite well for both the cases for all possible parameter ranges. From the simulation study it is recommended that asymptotic results can be used quite effectively even when the sample size is as small as 20 for all possible choices of the shape parameters.

### 6 Real data illustration

In this section we analyze two data sets and use our method to discriminate between two distribution functions.
Table 6  The PCS based on Monte-Carlo Simulations and the asymptotic results (in parenthesis) when the null distribution is xgamma, for different values of $\theta$

| $\theta \downarrow n \rightarrow$ | 20  | 40  | 60  | 80  | 100 | 400 |
|-----------------------------------|-----|-----|-----|-----|-----|-----|
| 0.85                              | 0.59160 | 0.63696 | 0.67272 | 0.69856 | 0.71836 | 0.88056 |
|                                  | (0.60410) | (0.64554) | (0.67624) | (0.70122) | (0.72249) | (0.88110) |
| 0.90                              | 0.59040 | 0.63612 | 0.67144 | 0.69668 | 0.71684 | 0.87856 |
|                                  | (0.60319) | (0.64430) | (0.67478) | (0.69960) | (0.72073) | (0.87900) |
| 1.05                              | 0.58832 | 0.62940 | 0.66692 | 0.69312 | 0.71184 | 0.87252 |
|                                  | (0.60088) | (0.64114) | (0.67102) | (0.69541) | (0.71620) | (0.87351) |
| 1.10                              | 0.58948 | 0.62980 | 0.66380 | 0.68408 | 0.71192 | 0.86996 |
|                                  | (0.60020) | (0.64021) | (0.66992) | (0.69417) | (0.71486) | (0.87187) |
| 1.25                              | 0.58692 | 0.62796 | 0.66324 | 0.68860 | 0.70764 | 0.86732 |
|                                  | (0.59833) | (0.63764) | (0.66688) | (0.69077) | (0.71117) | (0.86728) |
| 1.26                              | 0.58712 | 0.62960 | 0.66472 | 0.68848 | 0.70440 | 0.86760 |
|                                  | (0.59821) | (0.63748) | (0.66669) | (0.69055) | (0.71094) | (0.86699) |
| 1.40                              | 0.58448 | 0.62688 | 0.65984 | 0.68536 | 0.70176 | 0.86180 |
|                                  | (0.59662) | (0.63531) | (0.66410) | (0.68766) | (0.70780) | (0.86301) |
| 1.50                              | 0.58004 | 0.62440 | 0.65500 | 0.68928 | 0.70168 | 0.86024 |
|                                  | (0.59555) | (0.63383) | (0.66235) | (0.68569) | (0.70567) | (0.86027) |
| 1.65                              | 0.5822 | 0.62712 | 0.65240 | 0.67924 | 0.70000 | 0.85336 |
|                                  | (0.59400) | (0.63170) | (0.65982) | (0.68285) | (0.70258) | (0.85626) |
| 1.80                              | 0.58444 | 0.62228 | 0.65968 | 0.67316 | 0.69392 | 0.84964 |
|                                  | (0.59251) | (0.62965) | (0.65737) | (0.68011) | (0.69960) | (0.85233) |
| 2.00                              | 0.58360 | 0.62036 | 0.64884 | 0.67656 | 0.69080 | 0.84968 |
|                                  | (0.59059) | (0.62700) | (0.65422) | (0.67656) | (0.69574) | (0.84717) |
| 2.05                              | 0.57592 | 0.62340 | 0.64804 | 0.67084 | 0.68944 | 0.84388 |
|                                  | (0.59012) | (0.62636) | (0.65345) | (0.67569) | (0.69479) | (0.84589) |

Illustration-I

As a first illustration, data set represents the number of revolution before failure of the 23 ball bearings in the life-test is considered. It was originally reported in Lawless [14]. When we use Lindley model, the MLE of the parameter and the corresponding log-likelihood value are obtained as 0.0273 and $-115.7356$, respectively. The K–S distance between the data and the fitted Lindley distribution function is 0.19283 and the corresponding $p$ value is 0.31739.

On the other hand, when we use xgamma model, the MLE of the parameter and the corresponding log-likelihood value are obtained as 0.04071 and $-113.9634$, respectively. The K–S distance between the data and the fitted xgamma distribution function is 0.13228 and the corresponding $p$ value is 0.76796. We also present the observed, expected frequencies for different groups and the corresponding $\chi^2$ statistics for both the distributions to the fitted data. We observe that for this data set, both the distributions provide quite good fit to the data. The results are presented in Table 7.
Table 7  Observed and expected frequencies for two distributions of ball bearing data

| Intervals | Observed frequency | Expected frequency |
|-----------|--------------------|--------------------|
|           |                    | Lindley | xgamma |
| 0–35      | 3                  | 5.93    | 4.50   |
| 35–55     | 7                  | 4.46    | 4.88   |
| 55–80     | 5                  | 4.52    | 5.45   |
| 80–100    | 3                  | 2.61    | 3.12   |
| Above 100 | 5                  | 5.48    | 5.05   |

Table 8  Observed and expected frequencies for two distributions of bank customer data

| Intervals | Observed frequency | Expected frequency |
|-----------|--------------------|--------------------|
|           |                    | Lindley | xgamma |
| 0–5       | 31                 | 29.73   | 26.88  |
| 5–10      | 31                 | 30.46   | 31.26  |
| 10–15     | 19                 | 19.36   | 22.03  |
| 15–20     | 10                 | 10.52   | 11.51  |
| Above 20  | 9                  | 9.93    | 8.32   |

The $\chi^2$ values are 3.0419 and 1.4667 for Lindley and xgamma distributions with corresponding $p$ values 0.3852 and 0.689 respectively. The logarithm of RML i.e., $T = -1.7722$ which is less than 0. Hence, it indicates to choose the xgamma model.

Illustration-II

Now, we represent another data set for second illustration. The second data set represents the waiting times (in min) before service of 100 bank customers (see Ghitany et al. [12]). For the second data set, when we use Lindley model, the MLE of the parameter and the corresponding log-likelihood value are obtained as 0.1866 and $-319.0374$, respectively. The K–S distance between the data and the fitted Lindley distribution function is 0.06768 and the corresponding $p$ value is 0.74946.

When we use xgamma model, the MLE of the parameter and the corresponding log-likelihood value are obtained as 0.2634 and $-321.0203$, respectively. The K–S distance between the data and the fitted xgamma distribution function is 0.06249 and the corresponding $p$ value is 0.82970.

In this case also we present the observed, expected frequencies for different groups and the corresponding $\chi^2$ statistics for both the distributions to the fitted data and we observe that, for this data set also, both the distributions provide quite good fit. The results are presented in Table 8. In this case, the $\chi^2$ values are 0.1833 and 1.3041 for Lindley and xgamma distributions with corresponding $p$ values 0.9802 and 0.7281 respectively. The logarithm of RML i.e., $T = 1.9829$ which is greater than 0. Hence, for this case the indication is to choose the Lindley model.
7 Concluding Remarks

In this article, problem of discriminating between the two families of probability distributions, viz., the Lindley and xgamma, is considered. The statistic based on the ratio of the maximized likelihoods is considered and we obtain the asymptotic distributions of the test statistics under null hypotheses. We compare the probability of correct selection using Monte-Carlo simulations with the asymptotic results and it is observed the asymptotic results work quite well for a wide range of the parameter space, even when the sample size is very small. Therefore, the asymptotic results can be utilized to estimate the probability of correct selection. To calculate the minimum sample size required for a user specified probability of correct selection, asymptotic results are used. The concept of tolerance level is used based on the distance between the two distribution functions. For a particular $D^*$ tolerance level the minimum sample size $n$ is obtained for a given user specified protection level. Tables are provided for the protection level 0.90, however, for the other protection level the tables can be easily generated, for example, if we need the protection level $p^* = 0.85$, then all the entries corresponding to the row of $n$, will be multiplied by $\frac{z_{0.85}}{z_{0.90}}$. Tables 3 and 4, therefore, can be used for any given protection level. This article further reveals potential scopes for future research all together. An approach considering a general metric for sample size determination could be an interesting area for future investigations.

Acknowledgements The authors would like to thank the anonymous referees and the Editor-in-Chief of the journal for their kind suggestions and valuable comments that have improved the article in its revised version.

References

1. Asgharzadeh A, Nadarajah S, Sharafi F (2017) Generalized inverse Lindley distribution with application to Danish fire insurance data. Commun Stat Theory Methods 46:5001–5021
2. Atkinson AC (1969) A test for discriminating between models. Biometrika 56(2):337–347
3. Atkinson AC (1970) A method for discriminating between models. J. R. Stat. Soc. Ser. B (Methodol.) 32:323–353
4. Bakouch HS, Al-Zahrani BM, Al-Shomrani AA, Marchi VA, Louzada F (2012) An extended Lindley distribution. J. Korean Stat. Soc. 41(1):75–85
5. Chambers EA, Cox DR (1967) Discrimination between alternative binary response models. Biometrika 54(3–4):573–578
6. Chen W (1980) On the tests of separate families of hypotheses with small sample size. J. Stat. Comput. Simul. 11(3–4):183–187
7. Cox, D.R.: Tests of separate families of hypotheses. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, pp. 105–123. University of California Press, Berkley (1961)
8. Cox DR (1962) Further results on tests of separate families of hypotheses. J. R. Stat. Soc. Ser. B (Methodol.) 24:406–424
9. Dumonceaux R, Antle CE, Haas G (1973) Likelihood Ratio Test for discriminating between two models with unknown location and scale parameters. Technometrics 15(1):19–27
10. Dyer AR (1973) Discrimination procedures for separate families of hypotheses. J. Am. Stat. Assoc. 68(344):970–974
11. Ghitany ME, Al-Mutairi DK, Balakrishnan N, Al-Enezi LJ (2013) Power Lindley distribution and associated inference. Comput. Stat. Data Anal. 64:20–33
12. Ghitany ME, Atieh B, Nadarajah S (2008) Lindley distribution and its application. Math. Comput. Simul. 78(4):493–506
13. Gomez DE, Ojeda EC (2011) The discrete Lindley distribution: properties and applications. J. Stat. Comput. Simul. 81(11):1405–1416
14. Lawless JF (1982) Statistical Models and Methods for Lifetime Data. Wiley, New York
15. Lindley DV (1958) Fiducial distributions and Bayes’ theorem. J. R. Stat. Soc. Ser. B (Methodol.) 20:102–107
16. Merovci, F., Sharma, V.K.: The beta-Lindley distribution: properties and applications. J. Appl. Math. (2014)
17. Nadarajah S, Bakouch HS, Tahmasbi R (2011) A generalized Lindley distribution. Sankhya B- Appl. Interdiscip. Stat. 73(2):331–359
18. Nedjar S, Zeghdoudi H (2016) On gamma Lindley distribution: properties and simulations. J. Comput. Appl. Math. 298:167–174
19. Sankaran M (1970) The discrete poisson-Lindley distribution. Biometrics 26(1):145–149
20. Sen S, Chandra N (2017a) The quasi xgamma distribution with application in bladder cancer data. J. Data Sci. 15(1):61–76
21. Sen S, Chandra N, Maiti SS (2017b) The weighted xgamma distribution: properties and application. J. Reliab. Stat. Stud. 10(1):43–58
22. Sen S, Chandra N, Maiti SS (2018) Survival estimation in xgamma distribution under progressively type-II right censored scheme. Model Assist. Stat. Appl. 13(2):107–121
23. Sen S, Maiti SS, Chandra N (2016) The xgamma distribution: statistical properties and application. J. Mod. Appl. Stat. Methods 15(1):774–788
24. Shanker R, Sharma S, Shanker R (2013) A two-parameter Lindley distribution for modeling waiting and survival times data. Appl. Math. 4:363–368
25. Shibu DS, Irshad MR (2016) Extended new generalized Lindley distribution. Statistica 76(1):41–56
26. White H (1982a) Maximum likelihood estimation of mis-specified models. Econometrica 50:1–25
27. White H (1982b) Regularity conditions for Cox’s test of non-nested hypotheses. J. Econom. 19:301–318

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.