A New Probability Model with Support on Unit Interval: Structural Properties, Regression of Bounded Response and Applications

Subrata Chakraborty \(^1\) · Seng Huat Ong \(^2,3\) · Choung Min Ng \(^3\)

Accepted: 20 September 2023 / Published online: 30 October 2023
© Grace Scientific Publishing 2023

Abstract
A new distribution on \((0, 1)\) known as the generalized Log–Lindley distribution is proposed by extending the Log–Lindley distribution. This new distribution is shown to be a weighted Log–Lindley distribution. Important probabilistic and statistical properties have been derived. An interesting characterization of the weighted distribution in terms of Kullback–Leibler distance and weighted entropy has also been obtained. A useful result in insurance for the distorted premium principle is presented and verified with numerical calculations. Application of the new regression models for bounded responses based on this distribution is illustrated and compared with the beta regression and Log–Lindley regression models for two real-life data sets on risk management and on outpatient health expenditure. Much better fits for both data sets justify the relevance of the new distribution in statistical modeling and analysis. Furthermore, this generalization, apart from adding flexibility for modeling, retains the compactness and tractability of statistical quantities required for statistical analysis, which is a feature of the Log–Lindley distribution. Thus, the generalized Log–Lindley distribution should be a useful addition to statistical models for practitioners.

Keywords Computer sampling · Distortion function · Log-concavity · Premium loading · Regression · Stochastic ordering · Weighted entropy · Unit gamma distribution

Mathematics Subject Classification 60EXX · 62E15 · 62F10 · 62J12
Abbreviations

LL  Log–Lindley
pdf  Probability density function
cdf  Cumulative distribution function
GLL  Generalized Log–Lindley
LR  Likelihood ratio
sf  Survival function

1 Introduction

In statistical research many attempts have been made to introduce alternatives to the classical beta distribution [see Mazucheli et al. [10] for details]. But most of these distributions involve special functions in their formulations except for the Kumaraswamy distribution [9]. Recently, a new probability density function (pdf) with bounded support in (0,1) was introduced by Gómez-Déniz et al. [6] as an alternative to the classical beta distribution. This new distribution is called the Log–Lindley (LL) distribution and was formulated by a suitable transformation of a particular case of the generalized Lindley distribution of Zakerzadeh and Dolati [21]. The distribution has compact expressions for the moments and the cumulative distribution function (cdf). Gómez-Déniz et al. [6] studied its important properties relevant to the insurance and inventory management applications. In the application to insurance premium loading, Gómez-Déniz et al. [6] showed that the risk-adjusted (distorted) premium based on the LL distribution falls between the net premium and the dual-power premium. The LL model is also shown to be an appropriate regression model for bounded responses and as an alternative to the beta regression model. The LL distribution of Gómez-Déniz et al. [6] is probably the latest in a series of proposed alternatives to the classical beta distribution. Unlike its predecessors the main attraction of the LL distribution is the nice compact forms for its pdf, cdf and moments, which do not involve any special functions. Recently, Jodrá and Jiménez-Gamero [8] derived the quantile function of the LL distribution in terms of the Lambert W function and proposed a method to sample from the LL distribution.

The pdf, cdf and moments of the LL distribution of Gómez-Déniz et al. [6] are, respectively, given by

\[ f(x; \theta, \lambda) = \frac{\theta^2}{1 + \lambda \theta} (\lambda - \log x) x^{\theta - 1}, \quad 0 < x < 1, \lambda \geq 0, \theta > 0 \] (1)

\[ F(x; \theta, \lambda) = \frac{x^{\theta} [1 + \theta (\lambda - \log x)]}{1 + \lambda \theta} \]

\[ E(X^r; \theta, \lambda) = \frac{\theta^2}{1 + \lambda \theta} \frac{1 + \lambda (r + \theta)}{(r + \theta)^2}, \quad r = \cdots, -2, -1, 1, 2, \cdots \]
In this paper, we extend the LL distribution to provide another distribution on \((0, 1)\). A probability distribution defined on \((0, 1)\) is useful in many ways, for instance, in insurance as a distortion function in the definition of a premium principle. The motivation for this extension is to add flexibility to the LL distribution, a potentially very useful alternative to the beta distribution, and at the same time, retain the compactness and tractability of expressions for the pdf, cdf and moments which facilitate its application in statistical analysis. Some important probabilistic properties have been derived, and the proposed distribution is shown to be a weighted LL distribution. We also justified the extended LL distribution by showing an improvement in fit to the same data used by Gómez-Déniz et al. [6] for their LL model. This data set was originally used in Schmit and Roth [13] and is about the cost effectiveness of risk management (measured in percentages) in relation to exposure to certain property and casualty losses, adjusted by several other variables such as size of assets and industry risk. Data were collected from a questionnaire survey of 374 risk managers (73 observations) of large US-based organizations. Further details are given in Sect. 6.1. A more recent data set from the Medical Expenditure Panel Survey (MEPS) conducted by the U.S. Agency of Health Research and Quality has also been fitted by the extended LL distribution. MEPS is a probability survey that gives nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian population. Detailed information on individuals of each medical care episode by type of services is collected and this allows the development of models of health care utilization to predict future expenditures. With the added flexibility of the extended LL model, a much better fit than the beta and Log–Lindley distributions is also obtained.

The paper has been organized as follows. Section 2 presents the extended LL distribution, generalized LL (GLL) distribution, and various properties like moments, mode, Shannon entropy, stochastic ordering and convexity and log-concavity. The distribution is derived as a weighted LL distribution. Computer sampling from the GLL distribution is also examined in Sect. 2, and as a special case, we propose a faster method of computer generation of LL samples. An interesting characterization of the weighted distribution in terms of the Kullback–Leibler distance and weighted entropy has been obtained. An application to insurance premium loading is considered in Sect. 3 where it is shown that the premium based on the GLL distribution lies between the proportional hazard premium [19, 20] and maximal premium. Section 4 gives a useful re-parameterization of the GLL distribution. Section 5 gives the score equations for maximum likelihood estimation and simple expressions for the elements of the information matrix. Illustrations of non-regression and regression data modeling with the risk management data set in Schmit and Roth [13] and Gómez-Déniz et al. [6] are examined in Sect. 6. A concluding discussion ends the paper.

2 Generalized Log–Lindley Distribution \(LL_p(\theta, \lambda)\)

The LL distribution of Gómez-Déniz et al. [6] was obtained from a two-parameter case of the three-parameter generalized Lindley (GL) distribution of Zakerzadeh and
Dolati [21] which has the pdf

\[ f(x) = \frac{\theta^2 (\theta x)^{\alpha-1} (\alpha + \lambda x) e^{-\theta x}}{(\lambda + \theta) \Gamma(\alpha + 1)}, \quad 0 < x < 1, \lambda \geq 0, \theta > 0. \]

Gómez-Déniz et al. [6] first substituted \( \alpha = 1 \) and then employed the transformation \( X = \log(1/Y) \) to obtain their LL distribution in (1).

### 2.1 Definition and Derivation as a Weighted Log–Lindley Distribution

In this section, we consider the same transformation of the GL distribution but without setting \( \alpha = 1 \) to derive an extension called the generalized Log–Lindley (GLL) distribution. This proposed three-parameter distribution can be seen as an extension of the LL distribution of Gómez-Déniz et al. [6] with an additional parameter ‘\( p \)’. We denote the new distribution by \( LL_p(\theta, \lambda) \). By setting \( p = \alpha - 1 \), the pdf and cdf of the proposed distribution are given, respectively, by

\[ f(x; \theta, \lambda, p) = \frac{\theta^{2+p}}{\Gamma(1+p)[1+p+\lambda \theta]}(-\log x)^p (\lambda - \log x)x^{\theta-1}, \quad 0 < x < 1, \]

\[ F(x; \theta, \lambda, p) = \frac{\left[x^\theta + (1 + p + \theta \lambda) Ei(-p, -\theta \log x)\right](-\theta \log x)^{1+p}}{\Gamma(1+p)[1+p+\lambda \theta]} \]

where

\[ Ei(n, z) = \int_1^\infty e^{-zt/t^n} dt = z^{n-1} \Gamma(1-n, z), \quad z > 0, \]

is the generalized exponential integral and \( \Gamma(1-n, z) \) is the upper incomplete gamma function. In particular, the recurrence relations

\[ n Ei(n+1, z) = e^{-z} - z Ei(n, z) \]

and

\[ \frac{d}{dz} Ei(n, z) = -z Ei(n-1, z) \]

can be utilized in conjunction with \( Ei(0, z) = e^{-z}/z \) to obtain expressions for higher integral values on \( n \). Moreover, as \( z \to \infty, Ei(n, z) = e^{-z}/z \) (see http://mathworld.wolfram.com/En-Function.html for more results).
Alternatively, we can write the cdf in (3) in terms of upper incomplete gamma function as

\[ F(x; \theta, \lambda, p) = \frac{x^{\theta}(-\theta \log x)^{1+p} + (1 + p + \theta \lambda) \Gamma(1 + p, -\theta \log x)}{\Gamma(1 + p)[1 + p + \lambda \theta]} . \]

Furthermore, by using

\[ \Gamma(p + 1, y) = p! e^{-y} \sum_{k=0}^{p} \frac{y^k}{k!} \]

for integer \( p \), the cdf is written in a compact form devoid of any special function as

\[ F(x; \theta, \lambda, p) = \frac{x^{\theta}(-\theta \log x)^{1+p} + (1 + p + \theta \lambda) x^{\theta} \sum_{k=0}^{p} (-\theta \log x)^k / k!}{(1 + p + \lambda \theta) p!} . \]

Throughout the article \( LL_p(\theta, \lambda) \) and GLL will be used synonymously.

We now formulate the generalized Log–Lindley distribution as a weighted Log–Lindley distribution. The general concept of weighted probability distributions formalized by Rao [11] has found applications in diverse areas. A weighted distribution is a result of modifying the original distribution by a weight function. This arises from the fact that the process of recording observations is done without equal probability for each observation. Let \( X \) be a random variable with pdf \( f(x) \) and \( w(x) \) be a weight function. The weighted pdf \( f_W(x) \) is given by

\[ f_W(x) = \frac{w(x) f(x)}{E[w(X)]} . \]

Let \( f(x) \) be the Log–Lindley pdf and \( w(x) = (-\log x)^p \). We find that

\[ E[w(X)] = \frac{\theta^p(1 + \lambda \theta)}{\Gamma(1 + p)(1 + p + \lambda \theta)} . \]

Then

\[ f_W(x) = \frac{\theta^p(1 + \lambda \theta)}{\Gamma(1 + p)(1 + p + \lambda \theta)} (-\log x)^p \frac{\theta^2}{1 + \lambda \theta} (\lambda - \log x)x^{\theta - 1}, \]

which is the GLL pdf.

2.2 Important Special Cases and Shape of \( LL_p(\theta, \lambda) \) Distributions

Here, we list some important special cases of the GLL distribution for different choices of parameter values:

1. In \( p = 0 \), \( LL \) distribution of Gómez-Déniz et al. [6] given by (1).
2. In \( \lambda = 0 \), \( LL_p(\theta, \lambda) \) reduces to a new extension of Unit Gamma distribution of Grassia [7] is obtained with pdf \( f(x) = \frac{\theta^{2+p}}{\Gamma(2+p)}(-\log x)^{\rho+1}x^{\theta-1} \).

3. In \( p = 0, \lambda = 0 \), \( LL_p(\theta, \lambda) \) reduces to the Unit Gamma distribution of Grassia [7] with pdf \( f(x) = \theta(-\log x)x^{\theta-1} \).

4. In \( p = 0, \theta = 1 \), as \( \lambda \to \infty \), \( LL_p(\theta, \lambda) \to \text{Uniform}(0, 1) \).

5. In \( p = 1 \), a new \( LL_1(\theta, \lambda) \) distribution with pdf, cdf and moments

\[
f(x; \theta, \lambda) = \frac{\theta^3}{[2 + \lambda \theta]}(-\log x)(\lambda - \log x)x^{\theta-1}, \quad 0 < x < 1, \lambda \geq 0, \theta > 0
\]

\[
F(x; \theta, \lambda) = \frac{x^\theta [2 + \theta \lambda - \theta \log x(2 + \theta \lambda - \theta \log x)]}{2 + \lambda \theta}
\]

\[
E(X^r; \theta, \lambda) = \frac{\theta^3 [(r + \theta)\lambda + 2]}{(r + \theta)^3 [2 + \lambda \theta]}, \quad r = \ldots, -2, -1, 1, 2, \ldots
\]

6. In \( p = 2 \), a new \( LL_2(\theta, \lambda) \) distribution with pdf, cdf and moments

\[
f(x; \theta, \lambda) = \frac{\theta^4}{6 + 2\lambda \theta}(-\log x)^2(\lambda - \log x)x^{\theta-1}, \quad 0 < x < 1, \lambda \geq 0, \theta > 0
\]

\[
F(x; \theta, \lambda) = \frac{x^\theta [6 + 2\theta \lambda - \theta \log x(6 + 2\theta \lambda - \theta \log x)]}{6 + 2\theta \lambda}
\]

\[
E(X^r; \theta, \lambda) = \frac{\theta^4 [2\lambda(r + \theta) + 6]}{(2\theta + 1)^3 [3 + \lambda \theta]}, \quad r = \ldots, -2, -1, 1, 2, \ldots
\]

Figure 1 illustrates the shapes of the \( LL_p(\theta, \lambda) \) distributions for different choices of parameters \( \theta, \lambda \) and index \( p \). It is observed that the distribution can be positively or negatively skewed and symmetrical. The pdf can also be increasing or decreasing.

### 2.3 Moments and Mode

The moments of \( LL_p(\theta, \lambda) \) distribution are given by

\[
E(X^r; \theta, \lambda, p) = \frac{\theta^{2+p}((r + \theta)\lambda + (1 + p))}{(r + \theta)^{2+p}[1 + p + \lambda \theta]}, \quad r = \ldots, -2, -1, 1, 2, \ldots
\]

In particular, the mean is

\[
E(X; \theta, \lambda, p) = \left( \frac{\theta}{1 + \theta} \right)^{2+p} \frac{1 + p + \lambda (1 + \theta)}{1 + p + \lambda \theta} = \mu. \quad \text{(4)}
\]

The mode plays an important role in the application of a distribution. For the \( LL_p(\theta, \lambda) \) distribution the mode occurs at
Fig. 1 Pdf plots of $LL_p(\theta, \lambda)$ when $p = 0$ (red), 1 (green), 3 (blue), 5 (brown), 7 (light green) for $\theta = 20, \lambda = 0.001$, $\theta = 5, \lambda = 0.1$, $\theta = 20, \lambda = 5$, $\theta = 2, \lambda = 0.1, e \theta = 0.5, \lambda = 10$, $f \theta = 0.9, \lambda = 0$

\[
\exp\left\{ \frac{1 + p + \lambda(1 - \theta) \pm \sqrt{4p\lambda(\theta - 1) + (1 + p + \lambda(1 - \theta))^2}}{2(1 - \theta)} \right\}, \text{ for } \theta \neq 1.
\]

2.4 Random Sampling from GLL Distribution

Jodrá and Jiménez-Gamero [8] derived the quantile function of the LL distribution in terms of the Lambert W function [4]. For more information on the Lambert W
function see https://mathworld.wolfram.com/LambertW-Function.html. This gives a
direct method of generating LL samples on the computer which only requires the com-
putation of the Lambert W function. Computation of special functions usually requires
relatively more computational effort compared to elementary functions. This clearly
impacts on the speed of the computer sampling which is crucial in a massive Monte
Carlo simulation experiment. We give a quick method of generating samples from the
LL distribution which is a particular case of the method for computer generation of
the generalized Log–Lindley distribution presented below.

Zakerzadeh and Dolati [21] expressed the pdf of their generalized Lindley distri-
bution as a finite mixture of two gamma pdf’s:

\[ f(x) = \frac{\theta}{\lambda + \theta} f_g(x; \alpha, \theta) + \frac{\lambda}{\lambda + \theta} f_g(x; \alpha + 1, \theta) \]

where

\[ f_g(x; \alpha, \theta) = \theta^\alpha x^{\alpha-1} e^{-\theta x} / \Gamma(\alpha). \]

Based on this finite mixture representation and the transformation \( X = -\log(Z) \),
we propose the following method to generate samples from the GLL distribution.
Let \( G(\alpha, \theta) \) denote the gamma random variable with parameters \((\alpha, \theta)\) and
\( U \) be the uniform random variable over \((0, 1)\).

Specify parameters \((\lambda, \alpha, \theta)\).

1. Generate \( U \).
2. If \( U \leq \theta / (\theta + \lambda) \),
   Generate \( X = G(\alpha, \theta) \);
   otherwise
   Generate \( X = G(\alpha + 1, \theta) \).
3. Accept \( Z = \exp(-X) \) as a GLL random variate (transform \( X \)).
   For a sample of size \( N \), repeat the three steps \( N \) number of times.

Algorithm 1 (Generation of GLL samples)

There are efficient methods for computer sampling from the gamma distribution.
For example, Cheng [3] gave an efficient algorithm based on the acceptance-rejection
method.

If \( \alpha = 1 \), Algorithm 1 provides a method for generating LL random samples. We
note that when \( \alpha = 1 \), \( G(1, \theta) \) is the exponential random variable and the quantile
function is \( Q(u) = -\theta^{-1} \log(1 - u), 0 < u < 1 \). Therefore, generation of an expo-
nential random variate is very quick, requiring only evaluation of the log function.
Since \( 1 - U \) is also uniform on \((0, 1)\), an exponential random variate is given by
\( X = -\theta^{-1} \log(U) \), and this saves one arithmetic operation. The random variable
\( X = G(2, \theta) = G(1, \theta) + G(1, \theta) \) may be taken as a sum of two independent expo-
nential random variables. We have the following algorithm for the generation of LL
samples.
Specify parameters $(\lambda, \theta)$. 

1. Generate $U$.
2. If $U \leq \theta/(\theta + \lambda)$, 
   \[ \text{Generate } X = G(1, \theta) = -\theta^{-1} \log(U); \]
   otherwise,
   \[ \text{Generate } X = G(2, \theta) = X_1 + X_2, \text{ where } X_i = -\theta^{-1} \log(U), i = 1,2. \]
   ($U_1 = U$ and $U_2$ is an additional uniform random variate.)
3. Accept $Z = \exp(-X)$ as a LL random variate (transform $X$).

Algorithm 2 (Generation of Log-Lindley samples)

**Remark 1** As an analytic comparison, Algorithm 2 requires two to three log function evaluations to generate one LL random variate, while the quantile approach of Jodrá and Jiménez-Gamero [8] needs one calculation of the Lambert W function.

### 2.5 Survival and Hazard Rate Functions

The survival and hazard rate functions are, respectively, given by

\[
S(x; \theta, \lambda, p) = \frac{(1 + p + \theta \lambda)\gamma(1 + p, -\theta \log x) - x^\theta (-\theta \log x)^{1+p}}{\Gamma(1 + p)[1 + p + \lambda \theta]}
\]

and

\[
r(x; \theta, \lambda, p) = \frac{\theta^2(\lambda - \log x)(-\theta \log x)^p}{(1 + p + \theta \lambda)\gamma(1 + p, -\theta \log x) - x^\theta (-\theta \log x)^{1+p}}
\]

where $\gamma(p, y) = \Gamma(p) - \Gamma(p, y)$, is the lower incomplete gamma function. It can be easily checked that for $p = 0$, the above results reduce to those for the $LL_0(\theta, \lambda)$ distribution of Gómez-Déniz et al. [6]. When $p$ is an integer, the expressions can be written in compact form using the result

\[
\Gamma(p + 1, y) = p! e^{-y} \sum_{k=0}^{p} y^k / k!
\]

Some illustrative plots of the $LL_p(\theta, \lambda)$ hazard function for different choices of parameters $\theta, \lambda$ and index $p$ are presented in Fig. 2, which reveal that the hazard function can be increasing and bath-tub shaped.

### 2.6 Shannon Entropy of $LL_p(\theta, \lambda)$ and Related Results

Entropy is a quantity that is often used in model selection. According to the maximum entropy principle, to make inference based on incomplete information, a distribution that best represents the current state of knowledge is the one with maximum entropy. Let $X$ be a random variable with pdf $f(x)$ over the domain $\Omega$. The Shannon entropy
of pdf $f(x)$ is defined by

$$H(X) = H(f) = -\int f(x) \log f(x) dx$$

where the integral is over the domain $\Omega$. We now derive the Shannon entropy of the GLL distribution, $LL_p(\theta, \lambda)$.

$$H_p(X) = E[-\log f(X; \theta, \lambda, p)]$$

$$= -E\left[ \log \left( \frac{\theta^{2+p}}{\Gamma(1+p)[1+p+\lambda\theta]}(-\log X)^p(\lambda - \log X)X^{\theta-1} \right) \right]$$

$$= -\log \left( \frac{\theta^{2+p}}{\Gamma(1+p)[1+p+\lambda\theta]} \right) - pE[\log(\log(1/X))]$$

$$- E\left[ \log \left( \lambda - \log X \right) X^{\theta-1} \right]$$

(5)

It can be checked that for $p = 0$, $H_p(X)$ reduces to $H_0(X)$, where $H_0(X)$ is the Shannon entropy of the LL distribution, $LL_0(\theta, \lambda)$ given by

$$H_0(X) = \frac{1}{\theta(1+\theta\lambda)} \left[ \theta(1-\lambda)(1-\theta) + \theta e^{\lambda\theta} Ei(-\lambda\theta) ight]$$

$$- \theta(1+\lambda\theta) \log \left( \frac{\lambda e^{\lambda\theta}}{1+\lambda\theta} \right) - 2$$

(see Proposition 2 in p. 51 of Gómez-Déniz et al. [6]). Here $Ei(z)$ is defined as

$$Ei(z) = -\int_{z}^{\infty} e^{-\omega}/\omega d\omega$$

(see http://mathworld.wolfram.com/En-Function.html for more results).

For integral values of index parameter $p$, exact expressions for $E[\log(\log(1/X))]$ and $E[\log((\lambda - \log X)X^{\theta-1})]$ as follows can be obtained using Mathematica:

For $p = 1$,

$$E[\log(\log(1/X))] = \frac{3 + \theta\lambda - (2 + \theta\lambda)(\gamma + \log \theta)}{2 + \theta\lambda}.$$
\[ E[\log(\{\lambda - \log X\} X^{\theta-1})] \\
= 6 - 3 \theta + 2 \theta \lambda (1 - \theta) - e^{\theta \lambda} (2 - \theta \lambda) Ei (-\theta \lambda) + \theta (2 + \theta \lambda) \log \lambda \]

For \( p = 2 \),
\[ E[\log(\log(1/X))] = \frac{11 + 3 \theta \lambda - 2(3 + \theta \lambda)(\gamma + \log \theta)}{6 + 2 \theta \lambda} \]

\[ E[\log(\{\lambda - \log X\} X^{\theta-1})] \\
= \frac{24 - \theta (13 - 6 \lambda + 7 \theta \lambda) - e^{\theta \lambda} \theta (6 + \theta \lambda (\theta \lambda - 4)) Ei (-\theta \lambda) + 2 \theta (3 + \theta \lambda) \log \lambda}{\theta (6 + 2 \theta \lambda)} \]

and so on.

In particular for \( p = 1 \), Shannon Entropy for \( LL_1(\theta, \lambda) \) using (5) is given by
\[ H_1(X) = \log \left[ \frac{(2 + \lambda \theta)}{(1 + \lambda \theta)^{2 \lambda}} \right] - \frac{3 + \theta \lambda - (2 + \theta \lambda)(\gamma + \log \theta)}{2 + \theta \lambda} - \frac{6 - 3 \theta + 2 \theta \lambda (1 - \theta) - e^{\theta \lambda} (2 - \theta \lambda) Ei (-\theta \lambda) + \theta (2 + \theta \lambda) \log \lambda}{\theta (2 + \theta \lambda)} \]

where \( \gamma \) is the Euler–Mascheroni constant.

Since the GLL distribution is a weighted LL distribution, it is of interest to examine the connection between weighted distributions and entropy. We first define weighted entropy (see, for example, [18] and Kullback–Leibler distance.

**Definition 1:**
(a) For a pdf \( f(x) \) over the domain \( \Omega \) and a nonnegative weight function \( w(x) \), the weighted entropy of \( f(x) \) is defined by
\[ H(w; f) = - \int w(x) f(x) \log f(x) dx \]

where the integral is over the domain \( \Omega \).

(b) The Kullback–Leibler distance between two pdf’s \( f(x) \) and \( g(x) \) is defined by
\[ KL(f; g) = \int f(x) \log(\frac{f(x)}{g(x)}) dx. \]

We state the connection between weighted distribution and entropy.
Theorem 1 Let \( f(x) \) and \( g(x) \) be two pdf’s. The pdf \( f(x) \) is a weighted \( g(x) \) with weight \( w(x) \), if and only if,

\[
H(f) = -KL(f; g) + \frac{1}{E[w(X)]} H(w; g).
\]

Proof If \( f(x) = \frac{w(x)g(x)}{E[w(X)]} \), the proof follows directly from the definition of Shannon entropy. Conversely, if the expression holds, straightforward manipulation leads to

\[- \int f(x) \log g(x) dx = \frac{1}{E[w(X)]} \int w(x)g(x) \log g(x) dx.\]

That is, pdf \( f(x) \) is a weighted \( g(x) \) with weight \( w(x) \).

Notes

(1) We may express Shannon entropy of \( LL_p(\theta, \lambda) \) in terms of weighted entropy of \( LL_{p-1}(\theta, \lambda) \) as follows:

\[
H_p(X) = E[-\log f(X; \theta, \lambda, p)]
\]
\[
= -E\left[\log\left(\frac{\theta(p + \lambda \theta)}{p(1 + p + \lambda \theta)}(-\log X) f(X; \theta, \lambda, p - 1)\right)\right]
\]
\[
= -\log\left(\frac{\theta(p + \lambda \theta)}{p(1 + p + \lambda \theta)}\right) - E[\log(-\log X)] - E[\log[f(X; \theta, \lambda, p - 1)]].
\]

Since

\[
f(X; \theta, \lambda, p) = \frac{\theta(p + \lambda \theta)}{p(1 + p + \lambda \theta)}(-\log X) f(X; \theta, \lambda, p - 1), \quad (*)
\]

\[
E[\log[f(X; \theta, \lambda, p - 1)]]
\]
\[
= \int_0^1 \log[f(X; \theta, \lambda, p - 1)] f(X; \theta, \lambda, p) dx
\]
\[
= \int_0^1 \log[f(X; \theta, \lambda, p - 1)] \frac{\theta(p + \lambda \theta)}{p(1 + p + \lambda \theta)}(-\log X) f(X; \theta, \lambda, p - 1) dx
\]
\[
= \frac{\theta(p + \lambda \theta)}{p(1 + p + \lambda \theta)} \int_0^1 (-\log X) \log[f(X; \theta, \lambda, p - 1)] f(X; \theta, \lambda, p - 1) dx
\]
\[
= -\log\left(\frac{\theta(p + \lambda \theta)}{p(1 + p + \lambda \theta)}\right) - E[\log(-\log X)]
\]
\[
- \frac{\theta (p + \lambda \theta)}{p(1 + p + \lambda \theta)} \int_0^1 (-\log x) \log[f(X; \theta, \lambda, p - 1)]f(X; \theta, \lambda, p - 1) dx
\]

\[
H_p(X) = \log \left[ \frac{\theta (1 + p + \lambda \theta)}{\theta (p + \lambda \theta)} \right] - \frac{\theta (p + \lambda \theta)}{p(1 + p + \lambda \theta)} H(-\log x; f(X; \theta, \lambda, p - 1)),
\]

where \( H(-\log x; f(X; \theta, \lambda, p - 1)) \) is the weighted entropy of \( LL_{p-1}(\theta, \lambda) \) with weight function \((-\log x)\).

(2) By using (*), we obtain interesting expressions for moments \( E[(\log X)^r] \) of \( LL_p(\theta, \lambda) \).

Replace \( p \) by \( p + 1 \) in (*) and integrate both sides. We get

\[
E(\log X; \theta, \lambda, p) = -\frac{(p + 1)(2 + p + \lambda \theta)}{\theta (1 + p + \lambda \theta)}.
\]

In general,

\[
E[(\log X)^r; \theta, \lambda, p] = -\frac{(p + 1)[r](1 + r + p + \lambda \theta)}{\theta^r (1 + p + \lambda \theta)}.
\]

For \( p = 0 \) gives the corresponding expression for LL distribution Gómez-Déniz et al. [6], Eq. (5), as

\[
E(\log X) = -\frac{(2 + \lambda \theta)}{\theta (1 + \lambda \theta)}.
\]

**Remark 2** Note that Eq. (5) in Gómez-Déniz et al. [6] contains a typo.

For \( n \) random variables \( X_1, X_2, ..., X_n \) from \( LL_p(\theta, \lambda) \), we get

\[
E(\log \prod_{i=1}^n X_i; \theta, \lambda, p) = -\frac{n(p + 1)(2 + p + \lambda \theta)}{\theta (1 + p + \lambda \theta)} = -\frac{(p + 1)(2 + p + n\lambda \theta \frac{\theta}{n})}{\theta \frac{\theta}{n} (1 + p + n\lambda \theta \frac{\theta}{n})},
\]

which is \( E(\log Y; \theta, \lambda, p) \), where \( Y \sim LL_p(\theta \frac{\theta}{n}, n\lambda, p) \). Additionally, we can get

\[
E(X \log X; \theta, \lambda, p) = -\left(\frac{\theta}{1 + \theta}\right)^{p+3} \frac{(p + 1)(2 + p + \lambda (1 + \theta))}{\theta (1 + p + \lambda \theta)}, \ldots,
\]

\[
E(X^r \log X; \theta, \lambda, p) = -\left(\frac{\theta}{r + \theta}\right)^{p+3} \frac{(p + 1)(2 + p + \lambda (r + \theta))}{\theta (1 + p + \lambda \theta)}.
\]
2.7 Stochastic Ordering

The comparison of random quantities through the notion of stochastic ordering has important applications in many areas, for instance, in risk theory and reliability theory; see Chapters 12, 15 and 16 in Shaked and Shantikumar [14]. In this section, we consider the likelihood ratio (LR), hazard rate and stochastic orderings for $LL_{p}(\theta, \lambda)$ random variables. The definitions of likelihood ratio, stochastic and hazard rate orders are as follows.

**Definition 2** Suppose two random variables $X$ and $Y$ have pdf’s $f$ and $g$, cdf’s $F$ and $G$, hazard rate $h_X$ and $h_Y$, respectively. Then

1. $X$ is said to be smaller than $Y$ in the **likelihood ratio order**, denoted by $X \leq_{LR} Y$, if $f(x)g(y) \geq f(y)g(x)$ for all $x \leq y$.
2. $X$ is said to be **stochastically** smaller than $Y$, denoted by $X \leq_{ST} Y$, if $F(x) \geq G(x)$ for all $x$.
3. $X$ is said to be smaller than $Y$ in **hazard rate order**, denoted by $X \leq_{HR} Y$, if $h_X(x) \leq h_Y(x)$ for all $x$.

**Theorem 2** Let $X_1$ and $X_2$ be random variables following $LL_{p_1}(\theta_1, \lambda_1)$ and $LL_{p_2}(\theta_2, \lambda_2)$ distributions, respectively. If $\theta_1 \leq \theta_2, \lambda_1 \leq \lambda_2$ and $p_2 \leq p_1$, then $X_1 \leq_{LR} X_2$.

**Proof** Consider the ratio

$$
\frac{f(x; \theta_2, \lambda_2, p_2)}{f(x; \theta_1, \lambda_1, p_1)} = \frac{\theta_2^{2+p_2} \Gamma(1 + p_1)[1 + p_1 + \lambda_1 \theta_1]}{\theta_1^{2+p_1} \Gamma(1 + p_2)[1 + p_2 + \lambda_2 \theta_2]} \left(-\log x\right)^{p_2-p_1} h(x),
$$

where

$$
h(x) = \frac{\lambda_2 - \log x}{\lambda_1 - \log x} \left(\frac{\theta_2}{\theta_1}\right).\]

Gómez-Déniz et al. [6] have shown that the function $h(x)$ is non-decreasing for $x \in (0, 1)$ if $\theta_1 \leq \theta_2, \lambda_1 \leq \lambda_2$. If $p_2 \leq p_1$, it is clear that $(-\log x)^{p_2-p_1}$ is non-decreasing for $x \in (0, 1)$. This implies that if $\theta_1 \leq \theta_2, \lambda_1 \leq \lambda_2$ and $p_2 \leq p_1$, then the ratio (6) is non-decreasing for $x \in (0, 1)$ and hence, $X_1 \leq_{LR} X_2$. □

Clearly, Theorem 2 may be applied in the context of monotone likelihood ratio to obtain uniformly most powerful tests for parameters of interest.

LR ordering is stronger than hazard rate and stochastic orderings and this leads to the following implications [6]: $X_1 \leq_{LR} X_2 \Rightarrow X_1 \leq_{HR} X_2 \Rightarrow X_1 \leq_{ST} X_2$. Therefore, as in Corollary 1 of Gómez-Déniz et al. [6] similar results can be shown for the GLL distribution as follows.

**Corollary 1** Let $X_1$ and $X_2$ be random variables following $LL_{p_1}(\theta_1, \lambda_1)$ and $LL_{p_2}(\theta_2, \lambda_2)$ distributions, respectively. If $\theta_1 \leq \theta_2, \lambda_1 \leq \lambda_2$ and $p_2 \leq p_1$, then the
(1) moments, $E[X_1^k] \leq E[X_2^k]$ for all $k > 0$;
(2) hazard rates, $r_1(x) \leq r_2(x)$ for all $x \in (0, 1)$.

**Remark 3** For a nonnegative random variable $X$, let $X^w$ be the corresponding weighted random variable derived using a nonnegative weight function $w(x)$ (see Sect. 2.1), then $X \geq X^w$ if $w(x)$ is decreasing Shaked and Shantikumar [15]. In view of the results in Sect. 2.1 that $LL_p(\theta, \lambda)$ is a weighted $LL_0(\theta, \lambda)$ with nonnegative decreasing weight function $w(x) = (-\log x)^p$ it is therefore obvious that $LL(\theta, \lambda) \geq_{LR} LL_p(\theta, \lambda), p > 0$. In fact, it can be further generalized to state that $LL_p(\theta, \lambda) \geq_{LR} LL_q(\theta, \lambda), q > p$.

The corresponding hazard rate and stochastic orderings follow as a consequence.

### 2.8 Convexity, Concavity and Log-Concavity for GLL Distribution

Gómez-Déniz et al. [6] have shown the LL cdf $F(x)$ to be convex but did not consider the property of log-concavity. In this section, we examine the convexity of the cdf and log-concavity of the pdf for the GLL distribution. For brevity, we abbreviate the cdf notation for $LL_p(\theta, \lambda)$ distribution in (3) as $F(x)$ in this section.

**Theorem 3** If $0 < \theta \leq 1$, $p \geq 0$ and $\lambda \geq 0$, then $F(x)$ is concave for $x \in (0, 1)$. Hence, for $0 < \theta \leq 1$, $F(x)$ is also log-concave for $x \in (0, 1)$ since $F(x) \geq 0$.

**Proof** If $0 < \theta \leq 1$, then $(-\log x)^p$, $(\lambda - \log x)$ and $x^{\theta-1}$ are decreasing in $x \in (0, 1)$ for $p \geq 0$ and $\lambda \geq 0$. This implies that the pdf

$$F'(x) = f(x; \theta, \lambda, p) = \frac{\theta^{2+p}}{\Gamma(1+p)[1+p+\lambda\theta]}(-\log x)^p(\lambda - \log x)x^{\theta-1},$$

is decreasing in $x \in (0, 1)$. Thus, $F(x)$ is concave for $x \in (0, 1)$.

Since $F(x) \geq 0$, concavity implies $F(x)$ is also log-concave for $x \in (0, 1)$.$\quad \Box$

**Theorem 4** The function $F(x)$ is neither convex nor concave for $x \in (0, 1)$ for any $\theta > 1, p < 0$ and $\lambda \leq 0$.

**Proof** See Appendix.$\quad \Box$

**Note:** $F(x)$ is convex for $x \in (0, 1)$ only for $p = 0$ (that is, when $LL_p(\theta, \lambda)$ distribution reduces to the LL distribution) and $\lambda(\theta - 1) \geq 1$ as shown in Theorem 4 of Gómez-Déniz et al. [6].

We next show that the $LL_p(\theta, \lambda)$ distribution is log-concave. A function $f(x)$ is log-concave if $\log f(x)$ is concave and log-convex if $\log f(x)$ is a convex function. Based on Definition 2 of Borzadaran and Borzadaran [2], the log-concavity of a function $f(x)$ on an interval $(a, b)$ is equivalent to $f'(x)/f(x)$ being monotonically decreasing in $(a, b)$ or $(\ln f(x))'' < 0$.

**Theorem 5** If $\theta > 1$, the $LL_p(\theta, \lambda)$ pdf is log-concave.

**Proof** See Appendix.$\quad \Box$
Remark 4 We have also shown that the LL \((LL_0(\theta, \lambda))\) pdf is log-concave provided \(\theta > 1\).

The log-concavity or log-convexity of pdfs implies many interesting properties of the distributions, especially reliability properties. There are many applications of this log-concavity property in diverse disciplines [1]. Many properties of the GLL distribution follow from the property of log-concavity [2], pp. 205–206). For instance,

(1) GLL pdf is strongly unimodal; a distribution \(F\) on \(\mathbb{R}\) is said to be strongly unimodal if the convolution of \(F\) with any unimodal distribution is again unimodal.

(2) GLL pdf is a Polya frequency density of order 2, that is,

\[
f(x - y) f(x' - y') - f(x - y') f(x' - y) \geq 0 \text{ for } x < x', y < y'
\]

(3) Hazard function is a non-decreasing function in \(x\).

(4) Distribution function and survival function are log-concave.

3 Application to Insurance Premium Loading

In this section, we apply the results of Sect. 2.8 to insurance premium loading. According to the basic premium principle, assuming common agreement on the risk distribution, the net premium \(P(X)\) to be charged on an insurance coverage for exposure to risk \(X\) is given by \(P(X) = E[X]\). If there is no agreement on the risk distribution, a loading is added to \(X\). One approach to add this loading is through transformation of the initial cumulative distribution function of \(X\) by a continuous and non-decreasing function \(h\) known as the distortion function. This transformation results in a new distribution corresponding to a random variable \(Y\). If the distortion function \(h\) is convex, this guarantees that \(X \leq_{ST} Y\) and which further implies that \(E[X] \leq E[Y]\), that is, loading is nonnegative (refer to Gómez-Déniz et al. [6] for further discussion and [12] and references therein).

Theorem 6 If \(F(x; \theta, \lambda, p)\) of \(LL_P(\theta, \lambda)\) given by (3) is concave, then \(1 - F(1 - x; \theta, \lambda, p)\) is a convex function from \((0, 1)\) to \((0, 1)\) for \(0 < \theta \leq 1\) and \(p \geq 0, \lambda \geq 0\).

Proof It is shown in Theorem 3 that the cdf \(F(x; \theta, \lambda, p)\) of \(LL_P(\theta, \lambda)\) is concave for \(0 < \theta \leq 1\) and \(p \geq 0, \lambda \geq 0\). Hence, \(-F\) is convex, and by considering the second derivative, it follows that \(1 - F(1 - x; \theta, \lambda, p)\) is a convex function from \((0, 1)\) to \((0, 1)\).

Remark 5 \(F(x; \theta, \lambda, p)\) can be used as a distortion function to distort survival function (sf) of a given random variable as stated in Corollary 2 next.

Corollary 2 If \(X\) is the risk with \(sf G(x)\) and let \(Z\) be a distorted random variable with \(sf F[G(x); \theta, \lambda, p]\) for \(0 < \theta < 1\) and \(p \geq 0, \lambda \geq 0\). Let \(E[Z] = \int_0^\infty [G(x); \theta, \lambda, p]dx\)
= P_{\theta,\lambda}(X) be the distorted premium and
P_n(X) = \int_0^\infty [\bar{G}(x)]^n dx, 0 < n \leq 1 be the
proportional hazard premium \[20\] of an insurance product, respectively. Then
P_{\theta,\lambda}(X) is a premium principle such that

1. \[ P_n(X) \leq P_{\theta,\lambda}(X) \leq \max(X), \text{ for all } n \geq \theta, \]
2. \[ P_{\theta,\lambda}(aX + b) = a P_{\theta,\lambda}(X) + b, \]
3. if \( \bar{G}_1(X_1) \) and \( \bar{G}_1(X_2) \) are sf of two non-negative risk random variables \( X_1 \) and \( X_2 \) with \( \bar{G}_1(X_1) \leq \bar{G}_2(X_2) \), that is, \( X_1 \) precedes \( X_2 \) under first stochastic dominance then \( P_{\theta,\lambda}(X_1) \leq P_{\theta,\lambda}(X_2) \),
4. if \( X_1 \) precedes \( X_2 \) under second stochastic dominance, that is, if \( \int_0^\infty \bar{G}_1(x_1)dx_1 \leq \int_0^\infty \bar{G}_2(x_2)dx_2 \) for all \( x \geq 0 \), then \( P_{\theta,\lambda}(X_1) \leq P_{\theta,\lambda}(X_2) \).

**Proof** See Appendix.

**Remark 6** This new distorted premium principle is a trade-off between the proportional hazard premium and maximal premium. For \( n \to 1 \), \( P_n(X) \to E(X) \) and for \( n \to 0 \), \( P_n(X) \to \max(X) \) (see \[19\]).

For a numerical confirmation of Corollary 2 (1) we present in Table 1, values of the proportional hazard premium \( P_n(X) \) and the distorted premium \( P_{\theta,\lambda}(X) \) obtained by using the GLL distribution with different parameter values and considering the exponential, Weibull, and Inverse Gaussian as the underlying risk distribution.

### 4 A Useful Re-parameterization of \( LL_p(\theta, \lambda) \)

Starting with a two-parameter Lindley distribution \[8, 16\] obtained a re-parameterized version of (1) with pdf

\[
f(x; \theta, \pi) = \theta(\pi + \theta(\pi - 1) \log x)x^{\theta-1}, \quad 0 < x < 1, \quad 0 \leq \pi \leq 1, \quad \theta > 0. \tag{7}
\]

In fact, this can be obtained by substituting \( \frac{\lambda \theta}{1+\lambda \theta} = \pi \), that is, \( \frac{1}{1+\lambda \theta} = 1 - \pi \) in the LL distribution given by (1).

The pdf in (7) overcomes the issue of unbounded parameter space of (1). Applying the same re-parameterization to \( LL_p(\theta, \lambda) \) distribution in (2) we obtain a re-parameterized version with pdf

\[
f(x; \theta, \pi, p) = \frac{\theta^{1+p}}{\Gamma(1+p)[1+(\pi - 1)p]}(-\log x)^p(\pi + \theta(\pi - 1) \log x)x^{\theta-1}, \tag{8}
\]

where \( 0 < x < 1, \quad 0 \leq \pi \leq 1, \quad \theta > 0, \quad p \geq 0. \)

Now the mean of the re-parameterized GLL can be easily derived by using (4):

\[
E(X; \theta, \pi, p) = \left(\frac{\theta}{1+\theta}\right)^{2+p} = \frac{\theta}{\theta + p(1-\pi)}/\frac{1+\theta}{1+p(1-\pi)}.
\]
Table 1 Values of proportional hazard premium, $P_n(X)$ and distorted premium, $P_{\theta,\lambda}(X)$ for fixed $p$ and for different distributions with given parameters

| $n$  | $P_n(X)$ | $P_{\theta,\lambda}(X), p = 1$ | $P_{\theta,\lambda}(X), p = 2$ |
|------|----------|---------------------------------|---------------------------------|
|      |          | $(\theta, \lambda)$           | $(\theta, \lambda)$           |
| 0.4  | 5.00     | (0.3, 0.5)                      | (0.3, 0.5)                      |
| 0.75 | 2.67     | (0.3, 1.5)                      | (0.3, 1.5)                      |
| 1.0  | 2.00     | (0.7, 0.5)                      | (0.7, 1.5)                      |

Exponential ($\lambda =$ rate)

| $\lambda$ | 0.5  | 2.0  | 0.5  | 2.0  | 0.5  | 2.0  |
|------------|------|------|------|------|------|------|
|            | 19.54| 4.88 | 4.00 | 128.68| 2.20 | 6.60 |
|            | 18.78| 4.69 | 2.04 | 121.09| 2.14 | 6.41 |
|            | 8.15 | 2.04 | 1.90 | 22.67 | 1.23 | 3.68 |
|            | 7.59 | 1.90 | 1.90 | 20.27 | 1.16 | 3.49 |
|            | 26.35| 6.59 | 6.59 | 217.99| 2.71 | 8.14 |
|            | 25.80| 6.45 | 2.78 | 210.63| 1.53 | 8.01 |
|            | 11.13| 2.67 | 1.48 | 39.11 | 4.58 | 4.45 |

Weibull ($\alpha =$ shape, $\beta =$ scale)

| $\alpha$ | 0.5, $\beta = 1.0$ | 1.5, $\beta = 0.5$ | 1.5, $\beta = 1.5$ |
|----------|---------------------|---------------------|---------------------|
|          | 12.50 | 0.83 | 2.49 |
|          | 3.56  | 0.55 | 1.64 |
|          | 2.00  | 0.45 | 1.35 |
|          | 128.68 | 2.20 | 6.60 |
|          | 121.09| 2.14 | 6.41 |
|          | 22.67 | 1.23 | 3.68 |
|          | 20.27 | 1.16 | 3.49 |
|          | 217.99| 2.71 | 8.14 |
|          | 210.63| 1.53 | 8.01 |
|          | 39.11 | 4.58 | 4.45 |

Inverse Gaussian ($\mu =$ mean, $\sigma =$ scale)

| $\mu$ | 0.5, $\sigma = 1.0$ | 2.5, $\sigma = 0.5$ | 2.0, $\sigma = 2.0$ |
|-------|---------------------|---------------------|---------------------|
|       | 1.05 | 3.99 | 3.99 |
|       | 0.62 | 3.83 | 1.63 |
|       | 0.50 | 1.63 | 1.52 |
|       | 5.42 | 5.30 | 2.21 |
|       | 5.42 | 5.30 | 2.21 |
|       | 15.09| 124.01| 187.97| 26.80| 25.06| 37.81|
|       | 4.42 | 117.48| 182.76| 25.61| 25.06| 36.92|
|       | 2.50 | 28.00 | 47.37 | 9.13 | 8.42 | 13.26|
|       | 124.01| 25.06 | 44.41 | 8.42 | 8.42 | 12.64|
For \( p = 0 \), we get back (1) from (7) and (2) from (8). Note that the new parameter \( \pi \) introduced in (8) is bounded.

### 5 Maximum Likelihood Estimation and Information Matrix

The likelihood function for a random sample of size \( n \) from the \( LL_p(\theta, \lambda) \) is

\[
L = \frac{\theta^{n(2+p)}}{(1 + p)(1 + p + \lambda \theta)^n} \left( \prod_{i=1}^{n} (\lambda - \log x_i) \right)^p \prod_{i=1}^{n} (\lambda - \log x_i)^{\theta^2 - 1}.
\]

The log-likelihood function is then given by

\[
l = \log L = n(2 + p) \log \theta - n \log \Gamma(1 + p) - n \log(1 + p + \lambda \theta) + p \sum_{i=1}^{n} \log(-\log x_i)
+ \sum_{i=1}^{n} \log(\lambda - \log x_i) + (\theta - 1) \sum_{i=1}^{n} \log x_i.
\]

The first- and second-order derivatives of the log-likelihood function needed to derive the information matrix are given in Appendix. For the information matrix, we obtain the following result.

\[
E \left[ \sum_{i=1}^{n} \frac{1}{(\lambda - \log x_i)^2} \right] = \frac{\theta^{2+p}\lambda^2}{\Gamma(1+p)[1+p+\lambda\theta]} \sum_{r=0}^{p} \binom{p}{r} (\frac{1}{\lambda \theta})^r (-\lambda)^p \Gamma(r, \lambda \theta).
\]

For \( p = 0 \), this reduces to the result given in Gómez-Déniz et al. [6].

The elements \((I_{j,k}), \ j, k = 1, 2, 3\) of the information matrix are given by

\[
I_{11} = \frac{n((2 + p)(1 + p + \lambda \theta)^2 - \lambda^2 \theta^2)}{\theta^2 (1 + p + \lambda \theta)^2},
\]

\[
I_{12} = -\frac{n \lambda \theta}{1 + p + \lambda \theta} + \frac{n}{1 + p + \lambda \theta}, \quad I_{13} = -\frac{n \lambda}{1 + p + \lambda \theta},
\]

\[
I_{22} = -\frac{n \theta^2}{(1 + p + \lambda \theta)^2} + \frac{\theta^{2+p} e^{\lambda \theta}}{\Gamma(1+p)[1+p+\lambda \theta]} \sum_{r=0}^{p} \binom{p}{r} \left( \frac{1}{\lambda \theta} \right)^r (-\lambda)^p \Gamma(r, \lambda \theta)
\]

\[
I_{23} = \frac{n \theta^2}{(1 + p + \lambda \theta)^2},
\]

\[
I_{33} = -n \left\{ \frac{\Gamma'(1 + p)}{\Gamma(1 + p)} \right\}^2 + n \frac{\Gamma''/(1 + p)}{\Gamma(1 + p)}.
\]

This matrix can be inverted to get the asymptotic variance–covariance matrix for the maximum likelihood estimates.

The derivatives of the gamma function \( \Gamma(\alpha) \) in \( I_{33} \) may be computed as follows:
Let
\[
\psi(\alpha) = \frac{d \log \Gamma(\alpha)}{d\alpha} \quad \text{and} \quad \psi'(\alpha) = \frac{d^2 \log \Gamma(\alpha)}{d\alpha^2}
\]
be the digamma and trigamma functions, respectively. We note the following formulas for the derivatives of the gamma function \(\Gamma(\alpha)\) in terms of digamma and trigamma functions:
\[
\Gamma'(\alpha) = \Gamma(\alpha) \psi(\alpha), \quad \Gamma''(\alpha) = \Gamma(\alpha) \left\{ \psi'(\alpha) + \psi^2(\alpha) \right\}.
\]

Spouge [17] has given efficient methods to compute the digamma and trigamma functions. In the programming language \(R\) the functions digamma(\(x\)) and trigamma(\(x\)) are available to compute these functions.

### 6 Data Modeling Applications

In this section, parameter estimation was performed using \(R\) packages. Beta regression is performed using the package betareg with logit link for the mean, while other regressions are performed through optimization using the function mle2 in the package bbmle. In what follows, for the beta regression \(\mu = \frac{a}{a+b}\) and \(\phi = a+b\) where \((a,b)\) are parameters of the beta distribution.

#### 6.1 Risk Management Data

In this section, we consider the modeling of the data set, originally used in Schmit and Roth [13] and considered by Gómez-Déniz et al. [6], about the cost effectiveness of risk management (measured in percentages) in relation to exposure to certain property and casualty losses, adjusted by several other variables such as size of assets and industry risk. Description of the data set may be found in Gómez-Déniz et al. [6]. We take the response variable to be \(Y = \text{FIRMCOST}/100\). Six other variables (covariates) are \(Y = \text{ASSUME} (X_1), \text{CAP} (X_2), \text{SIZELOG} (X_3), \text{INDCOST} (X_4), \text{CENTRAL} (X_5)\) and \(\text{SOPH} (X_6)\). We model response variable \(Y\) as well as its complementary \((1-Y)\) without and with covariates. To accommodate the covariates we introduce two regression models by first linking the covariates to the parameter \(\theta\) through the log link function, and then to the mean \(\mu\) through the logit link function for both the LL and GLL distributions. It may be noted that Jodrá and Jiménez-Gamero [8] also investigated the same data set using a re-parameterization of the LL distribution.

#### 6.1.1 Modeling the Centering Parameter \(\theta\) of \(LL_p(\theta, \lambda)\)

Here we first note that as \(\lambda \to \infty\), the mean of the \(LL_p(\theta, \lambda)\) distribution in (4),
\[
E(X; \theta, \lambda, p) = \left( \frac{\theta}{1+\theta} \right)^{2+p} \frac{1+p+\lambda(1+\theta)}{1+p+\lambda \theta} \to \left( \frac{\theta}{1+\theta} \right)^{1+p}.
\]
This implies that the parameter $\theta$ plays a certain “centering” role for the distribution. For the purpose of regression modeling of the parameter $\theta$, a suitable link function is required. Suppose that a random sample $Y_1, Y_2, \ldots, Y_n$ of size $n$ is obtained from the $LL_p(\theta, \lambda)$ distribution. For a set of $k$ covariates, the log link for the $LL_p(\theta, \lambda)$ regression model gives the parameter $\theta$ for each $Y_i$ as

$$
\theta_i = \exp\left(x_i^T \beta\right), \quad i = 1, 2, \ldots, n,
$$

where $x_i^T = (1, x_{i1}, \ldots, x_{ik})$ are the covariates with corresponding coefficients $\beta = (\beta_0, \beta_1, \ldots, \beta_k)$.

Here, the beta (with logit link to mean), LL (with log link to $\theta$) and $LL_p(\theta, \lambda)$ (with log link to $\theta$) regression models are considered and the log-likelihood values and parameter estimates for the models considered, without and with covariates, are presented in Table 2.

In terms of the log-likelihood values, it is clear from the results in Table 2 that the generalized Log–Lindley model fits the data best with or without covariates for the response variable $Y$. The plots of the Pearson residuals from the fitted GLL and LL regression models against $Y$ do not show any particular pattern apart from three potential outlying data points (Fig. 3), indicating that the models are adequate. It is also the best model for the regression of $1 - Y$, while the beta model is the best for this case without covariates. For the case of modeling $1 - Y$ with and without covariates, it is seen that the estimates for the parameter $p$ of the $LL_p(\theta, \lambda)$ model approaches 0 and hence, approaches the results for the LL distribution.

The estimates for Log–Lindley regression model may not be compared directly with that in Gómez-Déniz et al. [6] since different parameterization is used. The estimates for the Log–Lindley model with response $1 - Y$ obtained here are close to that obtained in Jodrá and Jiménez-Gamero [8]. For both responses $Y$ and $1 - Y$, it is clear that the covariates SIZELOG and INDCOST are statistically significant across level of significance $> 0.02$. Hence, the measure of the firm’s risk management cost effectiveness (FIRMCOST) is negatively associated with the logarithm of total assets (SIZELOG) but positively associated with the measure of the firm’s industry risk (INDCOST).

### 6.1.2 Modeling in Terms of Mean

Here, we have investigated the cases of $Y$ and $1 - Y$ without and with covariates for $LL_p(\theta, \lambda)$ re-parameterized in terms of its mean, $\mu$ in (4) together with two other new parameters $\varphi$ and $\gamma$ such that the parameters in (2) are replaced by:

$$
\theta = \frac{\mu \gamma (2 + \varphi) + \sqrt{\mu^2 \gamma^2 \varphi^2 + 4 \mu \gamma (1 + \varphi)}}{2(1 - \mu \gamma)(1 + \varphi)}, \quad p = \frac{\log \gamma}{\log [(1 + \theta)/\theta]}, \quad \lambda = \left(\frac{1 + p}{1 + \theta}\right) \varphi.
$$

The new parameters are such that $0 < \mu < 1, \varphi > 0, 1 \leq \gamma < 1/\mu$. The re-parametrized distribution will now be denoted in terms of the new parameters as
| Models | $Y$ | Log–Lindley, $LL_1(\theta, \lambda)$ | Generalized Log–Lindley, $LL_p(\theta, \lambda)$ |
|--------|-----|----------------------------------|----------------------------------|
|        |     |                                  |                                  |
| (a) Without covariates |     |                                  |                                  |
| Log-likelihood | 76.1175 | 76.6042 | 83.2511 | 76.1175 | 69.0196 | 69.0195 |
| Estimates | $a = 0.6125$ | $\lambda = 0.0343$ | $\lambda = 0.3824$ | $a = 3.7979$ | $\lambda = 4.16 \times 10^3$ | $\lambda = 2.65 \times 10^3$ |
| & | (0.08553) | (0.06115) | (0.4842) | (0.7154) | (3.99 \times 10^{-8}) | (1.57 \times 10^{-7}) |
| & | $b = 3.7979$ | $\theta = 0.6907$ | $\theta = 1.2694$ | $b = 0.6125$ | $\theta = 5.9076$ | $\theta = 5.9077$ |
| & | (0.7154) | (0.05884) | (0.2095) | (0.08553) | (0.6914) | (1.042) |
| & | $p = 1.7819$ | | | | | (0.1457) |
| (b) With covariates and log link for regression* |     |                                  |                                  |
| Log-likelihood | 87.7230 | 83.6526 | 98.2999 | 87.7230 | 96.7054 | 96.7624 |
| Estimates | (Intercept) | 1.8880 | 1.8422 | 2.7840* | −1.8880 | −2.7403 | −2.6807 |
| & | (1.1575) | (0.9491) | (0.6026) | (1.1575) | (1.3689) | (1.3786) |
| & | ASSUME | −0.01214 | −0.005083 | −4.98 \times 10^{-3} | 0.01214 | 0.03166 | 0.03191 |
| & | (0.01320) | (0.01117) | (6.83 \times 10^{-3}) | (0.01320) | (0.01517) | (0.01522) |
| & | CAP | 0.1780 | 0.05793 | 0.05638 | −0.1780 | −0.7586 | −0.7463 |
| & | (0.2374) | (0.1862) | (0.1138) | (0.2374) | (0.3303) | (0.3314) |
| Models | Explanatory variables | Beta ($\mu, \varphi$) | Log–Lindley, $LL_1(\theta, \lambda)$ | Generalized Log–Lindley, $LL_1(\theta, \lambda)$ | Beta ($\mu, \varphi$) | Log–Lindley, $LL_p(\theta, \lambda)$ | Generalized Log–Lindley, $LL_p(\theta, \lambda)$ |
| --- | --- | --- | --- | --- | --- | --- | --- |
| SIZELOG | | $-0.5115^*$ | $-0.2917^*$ | $-0.2876^*$ | $0.5115^*$ | $0.6962^*$ | $0.6963$ |
| | | $(0.1184)$ | $(0.09817)$ | $(0.06018)$ | $(0.1184)$ | $(0.1418)$ | $(0.1424)$ |
| INDCAST | | $1.2363^*$ | $0.7122$ | $0.6953^*$ | $1.2363^*$ | $3.6193^*$ | $-3.7226$ |
| | | $(0.4823)$ | $(0.3922)$ | $(0.2406)$ | $(0.4823)$ | $(0.9167)$ | $(0.9370)$ |
| CENTRAL | | $-0.01216$ | $-0.01971$ | $-0.01968$ | $0.01216$ | $5.53 \times 10^{-3}$ | $6.10 \times 10^{-3}$ |
| | | $(0.08953)$ | $(0.06939)$ | $(0.04243)$ | $(0.08953)$ | $(0.1108)$ | $(0.1109)$ |
| SOPH | | $-0.003721$ | $5.10 \times 10^{-4}$ | $3.82 \times 10^{-4}$ | $0.003721$ | $0.03671$ | $0.03558$ |
| | | $(0.02159)$ | $(0.01675)$ | $(0.01024)$ | $(0.02159)$ | $(0.02943)$ | $(0.02947)$ |
| | $\varphi = 6.331^*$ | $\lambda = 0.0199$ | $\lambda = 0.4190$ | $\varphi = 6.331^*$ | $\lambda = 108.58^*$ | $\lambda = 299.47$ |
| | $(1.114)$ | $(0.04646)$ | $(0.5182)$ | $(1.114)$ | $(5.392 \times 10^{-4})$ | $(8.87 \times 10^{-5})$ |
| | $p = 3.3748^*$ | $p = 3.3748^*$ | $p = 1.0 \times 10^{-6}$ | $p = 1.0 \times 10^{-6}$ | $(0.8506)$ | $(0.1458)$ |

*The estimated coefficient is significantly different from 0 with $p$ value < 0.02.

*Estimate approaches the boundary of parameter space and hence, inferences cannot be made for this set of coefficients.

*Deviances between Generalized Log–Lindley and Log–Lindley models for $Y$ and $1 - Y$ are, respectively, 29.295 and 0.114.
Fig. 3 Pearson residuals versus fitted values plots for (a) GLL and (b) LL models with covariates and log link to $\theta$ for the regression of $Y$ for the risk management data.

$LL_\gamma(\mu, \varphi)$. Clearly, when $\gamma = 1$, the re-parameterized Log–Lindley distribution of Gómez-Déniz et al. [6] is obtained. Now for a random sample $Y_1, Y_2, \ldots, Y_n$ of size $n$ from the $LL_\gamma(\mu, \varphi)$ distribution and a set of $k$ covariates, the logit link for the $LL_\gamma(\mu, \varphi)$ regression model gives the mean for each $Y_i$ as

$$\mu_i = \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}, \quad i = 1, 2, \ldots, n.$$  

Here, only LL and $LL_\gamma(\mu, \varphi)$ regression models with the logit link to mean as above are applied to the same data set used. The log-likelihood values and parameter estimates for the models considered, with and without covariates, are presented in Table 3.

Clearly, in terms of the log-likelihood values in Table 3 that $LL_\gamma(\mu, \varphi)$ model fits the data better for the response variable $Y$ as well as $1 - Y$. For the regression of $1 - Y$, the estimates of $\gamma$ approaches 1; this explains why results approach that of the LL regression model. From Fig. 4, there is no evidence to suggest that the GLL and LL regression models are inadequate.

6.2 Outpatient Health Expenditures Data

This data set is obtained with description of the variables from Frees [5]. We considered the outpatient health expenditures for 1,352 individuals in ‘000,000 USD, that is the response variable $Y = \text{EXPENDOP}/1,000,000$. Five other variables (covariates) considered are AGE, GENDER, insure, RACE and PHSTAT. We model the response variable ($Y$) without and with covariates. Similarly, the covariates are linked to the parameter $\theta$ for the Log–Lindley and generalized Log–Lindley regression models.
Table 3 Log-likelihood values and parameter estimates (standard errors) for re-parameterized Log–Lindley and generalized Log–Lindley models with covariates

| Models | Explanatory variables | $Y$ | $1 - Y$ |
|--------|-----------------------|-----|--------|
|        |                       | Log–Lindley, $LL_1(\mu, \varphi)$ | Generalized Log–Lindley, $LL_Y(\mu, \varphi)$ | Log–Lindley, $LL_1(\mu, \varphi)$ | Generalized Log–Lindley, $LL_Y(\mu, \varphi)$ |
| Log-likelihood* |                        | 83.7575 | 91.9525 | 96.7689 | 96.8252 |
| Estimates | (Intercept)           | 1.6767 (1.3365) | 0.2175 (1.0621) | -2.7200 (1.3749) | -2.7149 (4.13 × 10$^{-13}$) |
| ASSUME | $\times 7.57 \times 10^{-3}$ | 0.004907 (0.009199) | 0.03208 (0.01528) | 0.03221 (3.79 × 10$^{-11}$) |
| CAP     | 0.08903 (0.2632) | -0.1335 (0.2285) | -0.7378 (0.3317) | -0.7508 (3.75 × 10$^{-12}$) |
| SIZELOG | $-0.4281*$ (0.1396) | $-0.2644*$ (0.1102) | 0.7024* (0.1432) | 0.7049 (3.91 × 10$^{-14}$) |
| INDCOST | 0.9687 (0.5238) | 0.4451 (0.3873) | -3.7854* (0.9513) | -3.7945 (2.57 × 10$^{-13}$) |
| CENTRAL | $0.02318$ (0.09840) | $-0.07726$ (0.08120) | $6.86 \times 10^{-3}$ (0.1109) | 0.002473 (2.99 × 10$^{-10}$) |
| SOPH    | $2.91 \times 10^{-4}$ (0.02383) | 0.005268 (0.01864) | 0.03593 (0.02952) | 0.03570 (9.06 × 10$^{-11}$) |
| $\varphi = 0.03488$ (0.07875) | $\varphi = 0.1755$ (0.2182) | $\varphi = 67.212*$ (5.2 × 10$^{-5}$) | $\varphi = 65.0001$ (2.95 × 10$^{-15}$) |
| $\gamma = 2.6735*$ (0.7311) | $\gamma = 1.0021$# (-) | $\gamma = 1.0021$# (-) |

*The estimated coefficient is significantly different from 0 with p value < 0.02

#Estimate approaches the boundary of parameter space and hence, inferences cannot be made for this set of coefficients

*Deviances between generalized Log–Lindley and Log–Lindley models for $Y$ and $1 - Y$ are, respectively, 16.390 and 0.113 using the log link function in Sect. 6.1.1. The numerical results are presented in Table 4.

Note that, from Table 4, the estimated parameter $\lambda$ for the Log–Lindley distribution approached 0, which indicates that the model reduces to the Unit –Gamma distribution of Grassia [7]. In terms of the log-likelihood values in Table 4 it is clear that $LL_p(\theta, \lambda)$ model fits the data better for the response variable $Y$ for both the case without and with covariates. From the estimates of all regression models, the variables AGE, GENDER, insure and PHSTAT are statistically significant across level of significance $> 0.02$. 

\[ Springer \]
These variables are all positively associated with the amount of outpatient health expenditures $Y$. The variance for the residuals appears to increase slightly as the fitted value increases for the LL regression model (Fig. 5b), but no obvious pattern is observed for the plot for the GLL regression model (Fig. 5a), suggesting that the fit by GLL regression model may be better.

7 Concluding Discussion

A new distribution on $(0, 1)$ that nests the Log–Lindley distribution of Gómez-Déniz et al. [6] and retains compact expressions for cdf and moments has been proposed. This new distribution is shown to be a weighted Log–Lindley distribution which enhances its usefulness in statistical modeling. Many of its important structural properties like log-concavity and stochastic ordering are studied. An interesting characterization of the weighted distribution in terms of Kullback–Leibler distance and weighted entropy has been given. This is of utility for statistical inference with the proposed distribution as a weighted Log–Lindley distribution. A new class of distorted premium principle based on the proposed distribution is also introduced and some important results are stated. A re-parameterization of the proposed distribution results in parameters with bounded range and regression modeling is conducted through a Logit link function. Applications to two real-life data sets, with much better fits than the beta and Log–Lindley distributions, showed the relevance of the newly proposed distribution in modeling without covariates and also in regression analysis with covariates.
Table 4 Log-likelihood values and parameter estimates (standard errors) for beta, Log–Lindley and generalized Log–Lindley models, with and without covariates

(a) Without covariates

| Models | Beta, $\alpha, \beta$ | Log–Lindley, $LL_1(\theta, \lambda)$ | Generalized Log–Lindley, $LL_p(\theta, \lambda)$ |
|--------|----------------------|--------------------------------------|----------------------------------|
| Log-likelihood | 4197.526 | 3774.016 | 4333.53 |
| Estimates | $a = 0.5462$ | $\lambda = 0.0$ | $\lambda = 1.8340$ |
| | $(0.01749)$ | | $(1.9414)$ |
| $b = 27.965$ | $\theta = 0.3934$ | $\theta = 1.9236$ | $p = 8.0608$ |
| | $(1.3474)$ | | $(0.1686)$ |

(b) With covariates and log link for regression

| Models | Explanatory variables | Beta, $\mu, \phi$ | Log–Lindley, $LL_1(\theta, \lambda)$ | Generalized Log–Lindley, $LL_p(\theta, \lambda)$ |
|--------|----------------------|------------------|--------------------------------------|----------------------------------|
| Log-likelihood | 4270.168 | 3794.483 | 4435.117 |
| Estimates | (Intercept) | $-5.0302^*$ | $-1.3478$ | $0.3623^*$ |
| | | $(0.2034)$ | $(0.1619)$ | $(0.07745)$ |
| AGE | $0.01184^*$ | $0.004766$ | $0.004747^*$ | $0.04789^*$ |
| | $(1.8 \times 10^{-3})$ | $(1.5 \times 10^{-3})$ | $(6.5 \times 10^{-4})$ | |
| GENDER | $0.1170^*$ | $0.05170$ | $0.04789^*$ | $0.01710$ |
| | $(0.04945)$ | $(0.03971)$ | | |
| Insure | $0.3607^*$ | $0.1509$ | $0.1478^*$ | $0.02425$ |
| | $(0.07311)$ | $(0.05641)$ | | |
| RACE (Asian) | $-0.3493$ | $-0.1651$ | $-0.1984^*$ | $0.07130$ |
| | $(0.2120)$ | $(0.1677)$ | | |
| RACE (Black) | $-0.0130$ | $-0.02774$ | $-0.07190$ | $0.06205$ |
| | $(0.1815)$ | $(0.1463)$ | | |
| RACE (Native) | $0.3063$ | $0.1025$ | $0.1535$ | $0.09185$ |
| | $(0.2655)$ | $(0.2206)$ | | |
| RACE (White) | $0.02168$ | $-0.005097$ | $-0.04412$ | $0.05885$ |
| | $(0.1726)$ | $(0.1391)$ | | |
| PHSTAT (very good) | $0.09941$ | $0.04988$ | $0.06882^*$ | $0.02338$ |
| | $(0.06895)$ | $(0.05397)$ | | |
| PHSTAT (good) | $0.1925^*$ | $0.08882$ | $0.1090^*$ | $0.02381$ |
| | $(0.06977)$ | $(0.05498)$ | | |
| PHSTAT (fair) | $0.4272^*$ | $0.1877$ | $0.2018^*$ | $0.03154$ |
| | $(0.08931)$ | $(0.07295)$ | | |
| PHSTAT (poor) | $0.7788^*$ | $0.3069$ | $0.3660^*$ | $0.04043$ |
| | $(0.1073)$ | $(0.09514)$ | | |
| $\psi = 32.453^*$ | $\lambda = 0.0^*$ | $\lambda = 1.97 \times 10^{-4}$ | $p = 8.7734^*$ |
| | $(1.525)$ | $(0.08720)$ | $(0.3932)$ | |

*The estimated coefficient is significantly different from 0 with $p$ value < 0.02

#Estimate approaches the boundary of parameter space and hence, inferences cannot be made for this set of coefficients

'Deviance between generalized Log–Lindley and Log–Lindley models is 1281.27
Fig. 5 Pearson residuals versus fitted values plots for (a) GLL and (b) LL models with covariates and log link to $\theta$ for the regression of $Y$ for the outpatient health expenditures data

Acknowledgements  The authors would like to acknowledge the comments and suggestions of the Associate Editor and the reviewers which have greatly improved the quality of the work.

Author Contributions  All authors contributed equally to the paper.

Funding  No funding.

Data Availability  Presented in the paper.

Declarations

Conflict of interest  The authors declare that they have no conflict of interest.

Ethics approval and consent to participate  Not applicable.

Consent for publication  Not applicable.

Appendix

Proof of Theorem 4 For any $\theta > 1$, $p > 0$ and $\lambda \geq 0$, consider the second-order derivative of $F(x)$ given by

$$F''(x) = \frac{\theta^{2+p}}{\Gamma(1+p)[1+p+\lambda\theta]} \log x^{b-1} \left\{ (\log x)^2 - \left( \frac{\lambda}{\theta} - 1 \right) \log x - \frac{p\lambda}{\theta - 1} \right\}. $$
In order for $F(x)$ to be convex, $F''$ must be $\geq 0$ or that the term

$$\left\{ (\log x)^2 - \left( \lambda - \frac{p + 1}{\theta - 1} \right) \log x - \frac{p\lambda}{\theta - 1} \right\} \geq 0 \text{ for all } x \in (0, 1).$$

However, when $x \to 1$,

$$\left\{ (\log x)^2 - \left( \lambda - \frac{p + 1}{\theta - 1} \right) \log x - \frac{p\lambda}{\theta - 1} \right\} \to -\frac{p\lambda}{\theta - 1} \leq 0.$$

Thus, $F(x)$ is not convex for $x \in (0, 1)$ for any $\theta > 1$, $p > 0$ and $\lambda \geq 0$.

When $x \to 0$,

$$\left\{ (\log x)^2 - \left( \lambda - \frac{p + 1}{\theta - 1} \right) \log x - \frac{p\lambda}{\theta - 1} \right\} = (\log x)^2 \left\{ 1 - \frac{\left( \lambda - \frac{p + 1}{\theta - 1} \right)}{\log x} \frac{p\lambda}{(\theta - 1)(\log x)^2} \right\} \to \infty.$$

This implies that $F''$ cannot be $\leq 0$ for all $x \in (0, 1)$. Therefore, $F(x)$ is not concave for $x \in (0, 1)$ for any $\theta > 1$, $p > 0$ and $\lambda \geq 0$. \qed

**Proof of Theorem 5** It is sufficient to show that

$$\frac{f'(x)}{f(x)} = \frac{F''(x)}{F'(x)} = \frac{(\theta - 1)}{(- \log x)} \left\{ (\log x)^2 - \left( \lambda - \frac{p + 1}{\theta - 1} \right) \log x - \frac{p\lambda}{\theta - 1} \right\}$$

is monotonically decreasing in $x \in (0, 1)$. Rewrite the above as

$$\frac{f'(x)}{f(x)} = (\theta - 1) \left\{ (- \log x) - \left( \lambda - \frac{p + 1}{\theta - 1} \right) - \frac{p\lambda}{\theta - 1} \frac{1}{(- \log x)} \right\}.$$

Noting that $\log x$ is monotonically increasing in $(0, 1)$, $- \log x$ is monotonically decreasing. For $\theta > 1$, $f'(x)/f(x)$ is monotonically decreasing in $(0, 1)$. Thus $f(x)$ is log-concave. \qed

**Proof of Corollary 2** From Theorem 3 we know that $F(x; \theta, \lambda, p)$ is concave for $x \in (0, 1)$ when $0 < \theta \leq 1$, $p \geq 0$ and $\lambda \geq 0$. Also being a cdf, it is an increasing function of $x$ with $F(0; \theta, \lambda, p) = 0$ and $F(1; \theta, \lambda, p) = 1$.

The results (2), (3) and (4) follow immediately from Definition 6 of distortion premium principle and subsequent properties thereof in Wang [20]. For the result (1) from Wang [20] it follows that $P_{\theta, \lambda}(X) \leq \max(X)$.

Now we provide a proof of $P_n(X) \leq P_{\theta, \lambda}(X)$. We first prove that for any $p$, $(x; \theta, \lambda, p) \geq x^\theta$. \hfill \qed
Case I When $p$ is an integer, we get

$$F(x; \theta, \lambda, p) = \frac{x^\theta \theta^{1+p}(-\log x)^{1+p} + (1 + p + \theta \lambda) \Gamma(1 + p, -\theta \log x)}{(1 + p + \lambda \theta) \Gamma(1 + p)}$$

$$= x^\theta \left[ 1 + \sum_{k=1}^p \frac{(-\theta \log x)^k}{k!} + \frac{(-\theta \log x)^{1+p}}{(1 + p + \lambda \theta) \Gamma(1 + p)} \right] > x^\theta.$$ 

Case II For real $p$, we apply the integral representation (https://en.wikipedia.org/wiki/Incomplete_gamma_function; see section “Evaluation Formulae”)

$$\Gamma(1 + p, y) = e^{-y} y^{1+p} \int_0^\infty e^{-yu} (1 + u)^p du$$

$$> e^{-y} y^{1+p} \int_0^\infty e^{-yu} (u)^p du$$

$$= e^{-y} \int_0^\infty e^{-t (p+1)-1} dt = e^{-y} \Gamma(1 + p).$$

Hence,

$$F(x; \theta, \lambda, p) = \frac{x^\theta (-\theta \log x)^{1+p} + (1 + p + \theta \lambda) \Gamma(1 + p, -\theta \log x)}{\Gamma(1 + p)[1 + p + \lambda \theta]}$$

$$> \frac{x^\theta (-\theta \log x)^{1+p} + (1 + p + \theta \lambda) e^{\theta \log x} \Gamma(1 + p)}{\Gamma(1 + p)[1 + p + \lambda \theta]}$$

$$= x^\theta \left[ 1 + \frac{(-\theta \log x)^{1+p}}{\Gamma(1 + p)[1 + p + \lambda \theta]} \right] > x^\theta.$$ 

Therefore, for any $p$, $(x; \theta, \lambda, p) \geq x^\theta$. We have

$$F(x; \theta, \lambda, p) \geq x^\theta \Rightarrow x^\theta \leq F(x; \theta, \lambda, p)$$

for $x \in (0, 1)$. This implies

$$(\overline{G}(x))^\theta \leq F(\overline{G}(x); \theta, \lambda, p)$$

since $0 \leq \overline{G}(x) \leq 1$ for all $x \in \mathbb{R}$.

Now, for all $x \in \mathbb{R}$, we have

$$(\overline{G}(x))^n \leq F(\overline{G}(x); \theta, \lambda, p)$$
when $0 < \theta \leq n < 1$, that is,

$$
\int_0^{\infty} (G(x))^n \, dx \leq \int_0^{\infty} F(G(x); \theta, \lambda, p) \, dx.
$$

It follows that $P_n(X) \leq P_{\theta, \lambda}(X)$.

The first- and second-order derivatives of the log-likelihood function:

\[
\frac{\partial l}{\partial \theta} = \frac{n(2 + p)}{\theta} - \frac{n\lambda}{1 + p + \lambda \theta} + \sum_{i=1}^{n} \log x_i
\]
\[
\frac{\partial l}{\partial \lambda} = \frac{n\theta}{1 + p + \lambda \theta} + \sum_{i=1}^{n} \frac{1}{\lambda - \log x_i}
\]
\[
\frac{\partial l}{\partial p} = n \log \theta - n \frac{\Gamma'(1 + p)}{\Gamma(1 + p)} - \frac{n}{1 + p + \lambda \theta} + \sum_{i=1}^{n} \log[\log(1/x_i)]
\]
\[
\frac{\partial^2 l}{\partial \theta^2} = -\frac{n(2 + p)}{\theta^2} + \frac{n\lambda^2}{(1 + p + \lambda \theta)^2} = -\frac{n[(2 + p)(1 + p + \lambda \theta)^2 - \lambda^2 \theta^2]}{\theta^2(1 + p + \lambda \theta)^2}
\]
\[
\frac{\partial^2 l}{\partial \lambda^2} = -\frac{n\theta^2}{(1 + p + \lambda \theta)^2} - \sum_{i=1}^{n} \frac{1}{(\lambda - \log x_i)^2}
\]
\[
\frac{\partial^2 l}{\partial p^2} = n \left[ \frac{\Gamma'(1 + p)}{\Gamma(1 + p)} \right]^2 - n \frac{\Gamma''(1 + p)}{\Gamma(1 + p)}
\]
\[
\frac{\partial^2 l}{\partial p \partial \theta} = \frac{n}{\theta} + \frac{n\lambda}{(1 + p + \lambda \theta)^2}, \quad \frac{\partial^2 l}{\partial \lambda \partial \theta} = \frac{n\lambda \theta}{(1 + p + \lambda \theta)^2} - \frac{n}{1 + p + \lambda \theta}
\]
\[
\frac{\partial^2 l}{\partial p \partial \lambda} = -\frac{n\theta}{(1 + p + \lambda \theta)^2}.
\]
7. Grassia A (1977) On a family of distributions with argument between 0 and 1 obtained by transformation of the gamma and derived compound distributions. Aust J Stat 19(2):108–114
8. Jodrá P, Jiménez-Gamero MD (2016) A note on the Log–Lindley distribution. Insur Math Econ 71:189–194
9. Jones MC (2009) Kumaraswamy’s distribution: a beta-type distribution with some tractability advantages. Stat Methodol 6:70–81
10. Mazucheli J, Menezes AFB, Chakraborty S (2018) On the one parameter unit-Lindley distribution and its associated regression model for proportion data. J Appl Stat 4:700–714
11. Rao CR (1965) On discrete distributions arising out of methods of ascertainment. In: Patil GP (ed) Classical and contagious discrete distribution. Pergamon Press and Statistical Publishing Society, Calcutta, pp 320–332
12. Rolski T, Schmidli H, Schmidt V, Teugel J (1999) Stochastic processes for insurance and finance. Wiley
13. Schmit JT, Roth K (1990) Cost effectiveness of risk managements practice. J Risk Insur 57:455–470
14. Shaked M, Shantikumar JG (1994) Stochastic ordering and their applications. Academic Press, San Diego
15. Shaked M, Shantikumar JG (2007) Stochastic orders. Springer, New York (NY)
16. Shanker R, Sharma S, Shanker R (2013) A two-parameter Lindley distribution for modeling waiting time and survival time Data. Appl Math 49:363–368
17. Spouge JL (1994) Computation of the gamma, digamma, and trigamma functions. SIAM J Numer Anal 31:931–944
18. Suhov Y, Stuhl I, Sekeh S, Kelbert M (2016) Basic inequalities for weighted entropy. Aequ Math 90:817–848
19. Wang S (1995) Insurance pricing and increased limits rate making by proportional hazards transformations. Insur Math Econ 17:43–54
20. Wang S (1996) Premium calculation by transforming the layer premium density. ASTIN Bull 26:71–92
21. Zakerzadeh H, Dolati A (2010) Generalized Lindley distribution. J Math Ext 3:13–25

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.