STATISTICS IN TRANSITION new series, March 2024
Vol. 25, No. 1, pp. 43–62, https://doi.org/10.59170/stattrans-2024-003
Received – 10.06.2020; accepted – 06.06.2023

Skew Log-Logistic distribution: properties and application

Arjun Kumar Gaire[1], Yogendra Bahadur Gurung[2]

Abstract

This paper introduces a novel three-parameter skew-log-logistic distribution. The research involves the development of a new random variable based on Azzalini and Capitanio's (2013) proposition. Additionally, various statistical properties of this distribution are explored. The paper presents a maximum likelihood method for estimating the distribution's parameters. The density function exhibits unimodality with heavy right tails, while the hazard function exhibits rapid increase, unimodality, and slow decrease, resulting in a right-skewed curve. Furthermore, four real datasets are utilized to assess the applicability of this new distribution. The AIC and BIC criteria are employed to assess the goodness of fit, revealing that the new distribution offers greater flexibility compared to the baseline distribution.

Key words: Log-Logistic, skew, marriage, menarche, age-specific fertility rate.

1. Introduction

Different families of distribution created from the baseline distribution by using different mathematical techniques have attracted the interest of statisticians and other scholars. In the literature, univariate probability distributions have been modified by adding extra parameters such as shape, scale, or location in the existing distribution, the primary aim of such extension, generalization, and modification of the existing distribution is to generate a more flexible distribution. Such new distributions have been applied to fit a distribution pattern of real-world problems such as in finance, economics, physics, biostatistics, actuarial science, reliability analysis, engineering, and many more fields. In this study, a new random variable from the application of Azzalini and Capitanio's proposition (Azzalini and Capitanio, 2013) has been introduced. For this, the Log-Logistic (LLog) distribution is chosen as a base distribution. Heavy-tailed distribution is always desired by the researcher to capture the right-tailed skewed data. This research is motivated to find the distribution to capture the unusual data or outliers present in the real dataset. Four real data sets of the age of the Nepalese mother at the birth of a child, the waiting time of customers at the bank before receiving the service, the age at first marriage of Nepalese females, and the age at

[1] Khwopa Engineering College, Bhaktapur, Nepal. E-mail: arjun.gaire@gmail.com,
ORCID: https://orcid.org/0000-0002-1958-9797.
[2] Central Department of Population Studies, Kirtipur, Nepal. E-mail: gurungyb@gmail.com.

menarche of Nepalese girls have been applied to test the suitability and flexibility of the proposed distribution.

The rest of the paper is organized as follows. In Section 2, a brief review of LLog distribution has been presented. In Section 3 a new distribution, called hereafter, 'Skew-Log-Logistic' (SLLog) distribution, is formulated and some statistical properties have been derived. Section 4 includes the methods of parameter estimation and Section 5 illustrates the application and validity of a model by using the four real data sets. Finally, Section 6 concludes the paper.

2. Log Logistic Distribution

The LLog distribution is a popular logistic distribution, which was initially developed to model population growth by Verhulst (1838) as cited in (Tahir et al., 2014). It is a continuous distribution with a uni-model failure rate function for a non-negative random variable. If T has a logistic distribution, then $X = e^T$ has LLog distribution. It is popularly known as Fisk-distribution in economics (Fisk, 1961). This distribution is applicable for modeling in various real-world situations, viz.: wealth and income (Fisk, 1961); economics and actuarial sciences (Kleiber and Kotz, 2003); flow data in hydrology (Ashkar and Mahdi, 2006) and 'time following a heart transplantation' in biostatistics (Collet, 2015). Similarly, Yilmaz et al. (2011) used it to estimate the seismic risk and earthquake occurrence probabilities. Further, Tahir et al. (2014) applied it to study the reliability analysis. Furthermore, it is used by Surendran and Tota-Maharaj, (2015) for modeling daily water consumption, estimation, and forecasting. So, LLog is a widely applicable model in different walks of life.

The probability density function (PDF) and the cumulative distribution function (CDF) of the three-parameter LLog distribution are given as

$$g(x) = \frac{\alpha}{\beta} \frac{\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{2}}, \text{ for x} > \gamma \tag{1}$$

$$G(x) = \frac{\left(\frac{x-\gamma}{\beta}\right)^{\alpha}}{1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}} \tag{2}$$

where $\alpha > 0$ is a shape parameter, $\beta > 0$ is a scale parameter and γ is a threshold or location parameter. The random variables under study in the different situations have positive values and the minimum cutoff value of these random variables is greater than zero, such as the minimum age of the mother at the birth of a child. Here, we consider the third threshold or location parameter of the LLog distribution.

The basic properties of this distribution are studied by Kleiber and Kotz (2003), Lawless (2003), and Ashkar and Mahdi (2006). The k^{th} order moments of two-parameter LLog distribution are derived and studied by Tadikamalla (1980) for $\alpha > k$ as

$$E(x^k) = \beta^k\, B\left(1+\frac{k}{\alpha}, 1-\frac{k}{\alpha}\right) = \beta^k \frac{\Gamma\left(1+\frac{k}{\alpha}\right)\Gamma\left(1-\frac{k}{\alpha}\right)}{\Gamma(2)} = \frac{k\pi\beta^k}{\alpha sin\frac{k\pi}{\alpha}} \tag{3}$$

where $B(a,b)$ is the Beta function defined as $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$. Also, the value of the Beta function is computed by using the relation as $B(a,b) = \frac{\Gamma a\ \Gamma b}{\Gamma(a+b)}$.

In particular, the mean and variance of the two-parameter LLog distribution are given as

$$\text{Mean} = \frac{\pi\beta}{\alpha sin\frac{\pi}{\alpha}} \text{ for } \alpha > 1 \text{ and Variance } = \frac{2\pi\beta^2}{\alpha sin\frac{2\pi}{\alpha}} - \left(\frac{\pi\beta}{\alpha sin\frac{\pi}{\alpha}}\right)^2 \text{for } \alpha > 2.$$

3. Skew Log-Logistic Distribution

The Normal distribution was extended to the Skew-Normal distribution by adding an asymmetry parameter $\lambda > 0$ (Azzalini, 1985, 2005). The PDF of a Skew-Normal distribution was derived by using the relation expressed in Equation (4).

$$f(z) = 2\,\phi(z)\,\Phi(\lambda z),\ z \in R,\ \lambda \in R \tag{4}$$

Here, $\phi(z)$ and $\Phi(z)$ are the PDF and CDF of Standard Normal distribution. The general formula for the construction of a skew-symmetrical distribution other than the Standard Normal distribution proposed by Azzalini and Capitanio (2013) is as:

$$f(x) = 2\,g(x)\,G(x),\ x \in R \tag{5}$$

where g(x) and G(x) are the PDF and CDF of any baseline distribution. Gupta et al. (2002) introduced Skew-uniform, Skew-t, Skew-Cauchy, Skew-Laplace, and Skew-logistic distributions. Later, Nadarajah (2009) studied in detail about the Skew-Logistic distribution. The base distributions chosen in all of these cases were a symmetrical distribution about the origin. However, Shaw and Buckley (2007) claimed to choose any distribution other than the symmetrical one (p. 15). Thus, in this research, the LLog distribution is chosen as a base distribution that is already positively skewed. Since the distributions proposed by different researchers are unable to catch the extreme value of data. We hope this construction of a heavy-tailed distribution could catch the unusual extreme value that exists in the data. [Note: Some or part of this research is published as a preliminary result in proceeding (Gaire et al., 2019)].

The LLog distribution is chosen as the base distribution because it has been preferred by different researchers in their generalization, modification and extension due to the flexible nature of both PDF and hazard rate functions. Different forms of generalization of the LLog distributions are found in literature used by different scholars. Some of frequently used distributions are exponentiated LLog distribution (Rosaiah et al., 2006); Beta LLog distribution (Lemonte, 2012) by using the generator introduced by Eugene et al. (2002) and Jones (2004); Kumaraswamy LLog distribution proposed by De-Santana et al. (2012) by using the relationship provided by Cordeiro and Castro (2011); Transmuted LLog distribution introduced by Aryal (2013) using the concept of the quadratic transmutation rank map of Shaw and Buckley (2007); Marshall-Olkin Extended LLog distribution proposed and studied by Gui (2013) using the concept of Marshall and Olkin (1997); Zografos-Balakrishnan LLog distribution introduced and studied by Hamedani (2013) based on the concept of Zografos-Balakrishnan generalized distribution (Zografos and Balakrishnan, 2009); McDonald LLog distribution proposed and studied by Tahir et al. (2014) using the concept of Alexander et al. (2012); Extended LLog distribution studied and presented by Lima and Cordeiro (2017) using an exponentiated generalized class of distribution of Cordeiro et al. (2013). Similarly, Additive Weibull LLog distribution has been introduced by Hemeda (2018) using the concept suggested by Hassan and Hemeda (2016); Transmuted generalized LLog distribution was studied by Adeyinka and Olapade (2019). At this juncture, the SLLog distribution is introduced and formulated; further some structural properties of the distribution are derived, along with a method of parameter estimation, and applied to four real data sets for model validity.

3.1. Probability Density Function of the SLLog Distribution

In this section, the PDF of the SLLog distribution is introduced. After substituting the values of $g(x)$ and $G(x)$ of the LLog distribution in Equation (5) we obtained the PDF of the SLLog distribution in Equation (6) as:

$$f(x) = \frac{2\alpha}{\beta} \frac{\left(\frac{x-\gamma}{\beta}\right)^{2\alpha-1}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{3}}, \text{ for x} > \gamma \tag{6}$$

Here, $f(x)$ is a probability density function since the total probability under a given range is unity. Figure 1 depicts the plots of PDF of the distribution for the selected values of parameters. The graph shows that the PDF is right-skewed for selected values of parameters.

3.2. Cumulative Distribution Function of SLLog Distribution

The CDF of the SLLog distribution is defined as:

$$F(x) = \int_{\gamma}^{x} f(x)dx = \int_{\gamma}^{x} \frac{2\alpha}{\beta} \frac{\left(\frac{x-\gamma}{\beta}\right)^{2\alpha-1}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{3}}\, dx$$

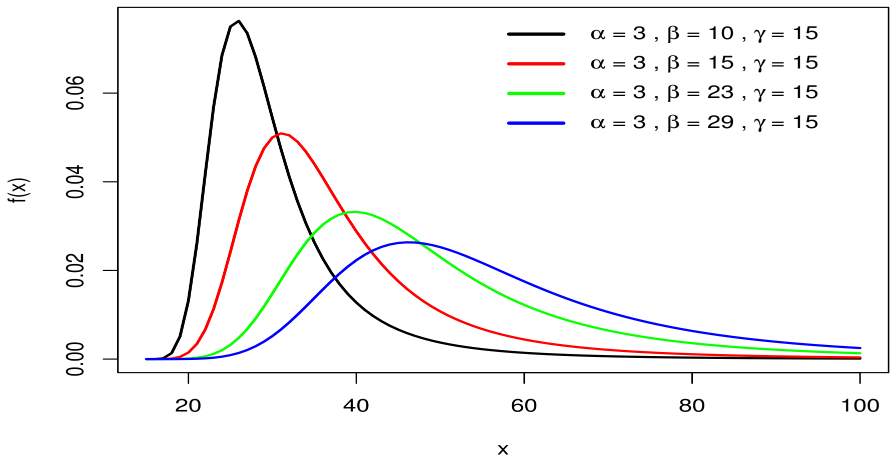

Figure 1: Plots of PDF of SLLog distribution for selected values of parameters

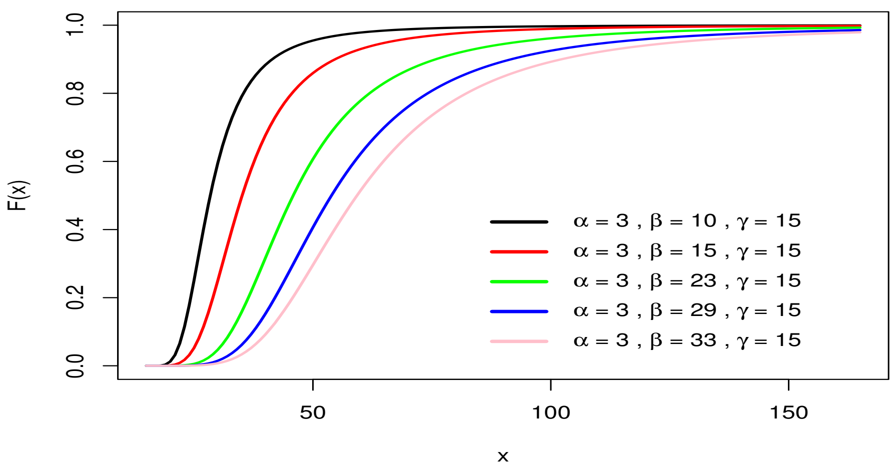

Figure 2: Graph of CDF for the selected values of parameters

On simple calculation, it gives the value of F(x) as:

$$F(x) = 1 - \frac{2}{1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}} + \frac{1}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{2}}, \text{ for x} > \gamma \tag{7}$$

The graph of CDF of SLLog distribution has been presented in Figure 2 for the selected values of the parameters α, β, and fixed value of $\gamma = 15$. The graph is monotonically increasing and the maximum value is 1 for a different set of parameters selected.

3.3. Moments About Origin of the SLLog Distribution

To calculate the k^{th} order moments of the SLLog distribution about the origin, first, consider the third parameter $\gamma = 0$, then the density function (6) becomes,

$$f(x) = \frac{2\alpha}{\beta} \frac{\left(\frac{x}{\beta}\right)^{2\alpha-1}}{\left(1+\left(\frac{x}{\beta}\right)^{\alpha}\right)^{3}}, \text{ for x} > 0$$

Now, the k^{th} order moments of the SLLog distribution about the origin is

$$E(x^k) = \int_0^\infty X^k f(x)dx = \int_0^\infty x^k \frac{2\alpha}{\beta} \frac{\left(\frac{x}{\beta}\right)^{2\alpha-1}}{\left(1+\left(\frac{x}{\beta}\right)^{\alpha}\right)^{3}} dx = 2\,\beta^k B\left(2+\frac{k}{\alpha}, 1-\frac{k}{\alpha}\right)$$

By using the relation of integration from Gradshteyn and Ryzhik (2000),

$$E(x^k) = 2\,\beta^k \frac{\Gamma\left(2+\frac{k}{\alpha}\right)\Gamma\left(1-\frac{k}{\alpha}\right)}{\Gamma 3}$$

Here, it is to be noted that moments of the SLLog distribution are only defined for $\alpha > k$ as.

$$E(x^k) = \frac{\pi(k+\alpha)\beta^k}{\alpha^2 sin\frac{k\pi}{\alpha}} \tag{8}$$

In particular, the mean and variance of the SLLog Distribution are given in Equation (9).

$$\text{Mean} = \frac{\pi(\alpha+1)\beta}{\alpha^2 sin\frac{\pi}{\alpha}} \text{ and Variance } = \frac{\pi(\alpha+2)\beta^2}{\alpha^2 sin\frac{2\pi}{\alpha}} - \left(\frac{\pi(\alpha+1)\beta}{\alpha^2 sin\frac{\pi}{\alpha}}\right)^2 \tag{9}$$

Thus, the value of the mean of SLLog distribution for $\alpha = 3$ and $\beta = 10$ is 16.1252. Table 1 gives the value of the first four moments of distribution about the origin for different values of parameters. These moments can be used to compute the value of skewness and kurtosis of the distribution. The values of moments are increased with the increase in the value of parameters.

Table 1: Value of first four moments about the origin for different values of parameters

Parameters	$K=1$	$K=2$	$K=3$	$K=4$
$\alpha=8,\ \beta=10$	11.54	138.84	1753.35	23571.43
$\alpha=9,\ \beta=11$	12.47	160.62	2146.27	29992.96
$\alpha=10,\ \beta=12$	13.42	184.72	2617.31	38367.2
$\alpha=11,\ \beta=13$	14.37	211.02	3170.38	48922.32

3.4. Random Number Generation and Quantile Function of the SLLog Distribution

A set of random numbers can be generated by using the method of inversion from the CDF of the SLLog distribution. For this, let, $F(x) = U$, where, the function U follows the uniform distribution in an interval $[0,1]$ as

$$1 - \frac{2}{1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}} + \frac{1}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{2}} = U$$

$$\frac{2}{1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}} - \frac{1}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{2}} = 1-U$$

Let, $\left(\frac{x-\gamma}{\beta}\right)^{\alpha}$ = Z. Then, it leads to: $Z^2(1-U)-2UZ-U=0$

This is quadratic in Z. After simple calculation, this becomes

$$Z = \frac{U \pm \sqrt{U}}{1-U}$$

Here, $U < 1$ so the term becomes negative and this negative term is not included in further analysis. Thus, the value of the random variable is given as

$$X = \gamma + \beta\left(\frac{U+\sqrt{U}}{1-U}\right)^{\frac{1}{\alpha}} \tag{10}$$

For the known value of parameters α, β and γ, one can generate a set of random numbers X by using Equation (10). Similarly, by choosing the suitable value of U in Equation (10) one can also get the different values of quantiles such as the first, second, and third quartiles obtained by setting $U = \frac{1}{4}$, $U = \frac{1}{2}$, and $U = \frac{3}{4}$ respectively.

3.5. Reliability Analysis of the SLLog Distribution

The reliability function $R(x)$ as defined by Rodriguez (2010) is simply the complement of the CDF. It is also the probability that a random variable X will take a value greater than a number x or the probability of an item not failing before some time x. So, it is defined as $R(X) = Prob(X > x) = 1 - Prob(X \leq x) = 1 - F(x)$

The graph of the reliability function of the SLLog distribution is presented in Figure 3 and the expression is given in Equation (11). The graph of the reliability function is decreasing with respect to increase in the value of variable X.

$$R(x) = \frac{2}{1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}} - \frac{1}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{2}} = \frac{1+2\left(\frac{x-\gamma}{\beta}\right)^{\alpha}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{2}} \tag{11}$$

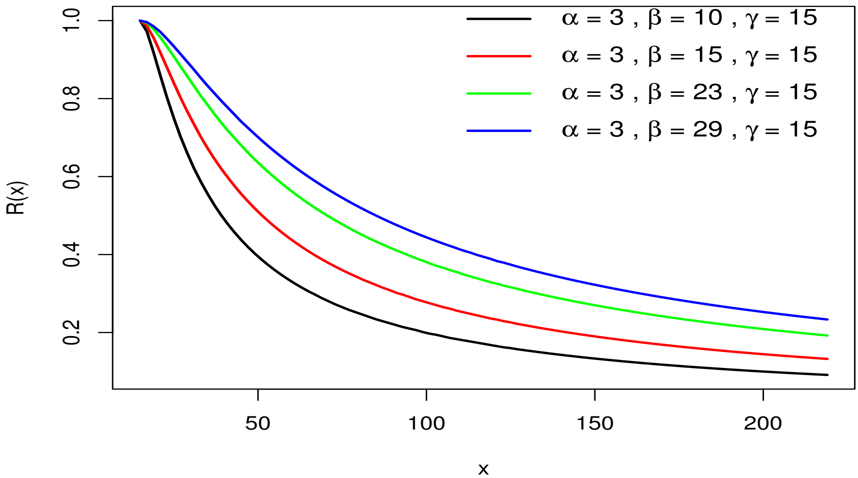

Figure 3: The plot of the reliability function for selected values of parameters

3.6. Hazard, Inverse Hazard, and Cumulative Hazard Rate Function

The other characteristics of interest of a random variable are the hazard and inverse hazard rate function defined as $h(x) = \frac{f(x)}{1-F(x)}$, and $rh(x) = \frac{f(x)}{F(x)}$. Thus, the hazard rate function for the SLLog distribution, which is the conditional probability of failure, given that it has survived up to the time x is given in Equation (12), and the graph of the hazard rate function is presented in Figure 4. Similarly, the inverse hazard rate function defined by (Barlow et al., 1963) for SLLog is present in Equation (13). The hazard function increases fast along with the uni-modality and decreases slowly creating a right skew curve.

$$h(x) = \frac{2\alpha}{\beta} \frac{\left(\frac{x-\gamma}{\beta}\right)^{2\alpha-1}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)\left(1+2\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)} \tag{12}$$

Similarly, the Inverse hazard rate function of the SLLog distribution is given as:

$$rh(x) = \frac{2\alpha}{\beta} \frac{1}{\left(\left(\frac{x-\gamma}{\beta}\right)+\left(\frac{x-\gamma}{\beta}\right)^{\alpha+1}\right)} \tag{13}$$

Furthermore, the cumulative hazard rate function of the SLLog distribution is defined by $H(x) = -ln(R(x))$ as given in Equation (14) and the graph is increasing with respect to the increase in values of variable X, which has been depicted in Figure 5.

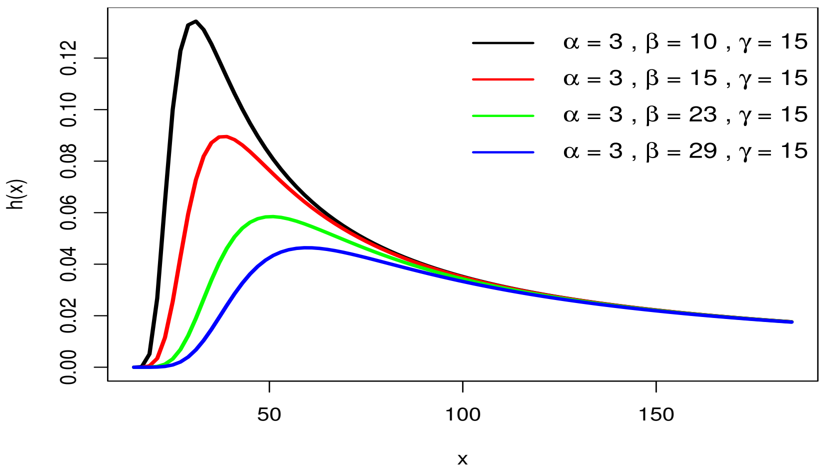

Figure 4: Plot of hazard rate function for the selected value of parameters

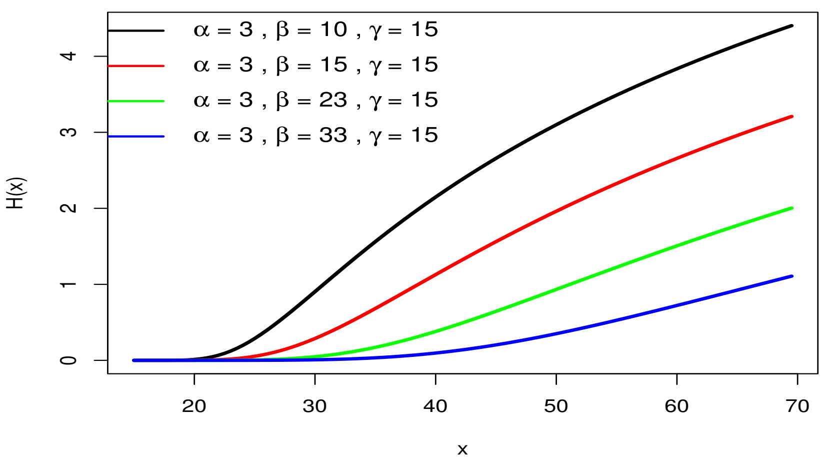

Figure 5: Graph of cumulative hazard rate function for selected values of parameters

$$H(x) = -ln(R(x)) = -ln\left(\frac{1+2\left(\frac{x-\gamma}{\beta}\right)^{\alpha}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{2}}\right)$$

$$H(x) = 2\,ln\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right) - ln\left(1+2\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right) \tag{14}$$

3.7. Entropy Measure of SLLog Distribution

Entropy is defined as the measure of the variation of the uncertainty of a random variable which is used in various situations in science and engineering. Different forms of the entropy are studied and compared, here we only derived the expression of two types of entropy Renyi entropy and q-entropy of the SLLog distribution.

3.7.1 Reny Entropy

First of all, for the SLLog random variable X with PDF $f(x)$, the Renyi entropy as defined by (Renyi, 1961) which has a similar role of kurtosis to measure and compare the shapes of densities is given as

$$I_R(\rho) = \frac{1}{1-\rho}\, ln\left(\int (f(x))^{\rho} dx\right)$$

Where $\rho > 0$ and $\rho \neq 1$ and ρ is a real non-integer. And the integral is computed as,

$$\begin{aligned}\int_{\gamma}^{\infty} (f(x))^{\rho} dx &= \int_{\gamma}^{\infty} \left(\frac{2\alpha}{\beta}\right)^{\rho} \frac{\left(\frac{x-\gamma}{\beta}\right)^{(2\alpha-1)\rho}}{\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{3\rho}}\, dx \\ &= 2^{\rho}\left(\frac{\alpha}{\beta}\right)^{\rho-1} B\left(\frac{2\rho\alpha-\rho+1}{\alpha}, \frac{4\rho-2\rho\alpha-1}{\alpha}\right)\end{aligned}$$

$$\int_{\gamma}^{\infty} (f(x))^{\rho} dx = 2^{\rho}\left(\frac{\alpha}{\beta}\right)^{\rho-1} \frac{\Gamma\left(\frac{2\rho\alpha-\rho+1}{\alpha}\Gamma\frac{4\rho-2\rho\alpha-1}{\alpha}\right)}{\Gamma(3\rho)}$$

Therefore, the Renyi entropy of the SLLog distribution can be expressed as

$$I_R(\rho) = \frac{1}{1-\rho} ln\left(2^{\rho}\left(\frac{\alpha}{\beta}\right)^{\rho-1} \frac{\Gamma\left(\frac{2\rho\alpha-\rho+1}{\alpha}\Gamma\frac{4\rho-2\rho\alpha-1}{\alpha}\right)}{\Gamma(3\rho)}\right) \tag{15}$$

3.7.2 q-Entropy

For the SLLog random variable X with PDF $f(x)$, the q-entropy as defined and introduced by Havarda and Charvat (1967) and later applied to physical problems by Tsallis (1988) is defined as

$$I_R(q) = \frac{1}{1-q}\left(1-\int (f(x))^{q} dx\right)$$

Where $q > 0$ and $q \neq 1$ and q is a real non-integer.

Therefore, after using the expression of Equation (15) with replace of ρ by q the q-entropy of the SLLog distribution can be expressed as

$$I_R(q) = \frac{1}{1-q}\left(1-2^q\left(\frac{\alpha}{\beta}\right)^{q-1}\frac{\Gamma\left(\frac{2q\alpha-q+1}{\alpha}\Gamma\frac{4q-2q\alpha-1}{\alpha}\right)}{\Gamma(3q)}\right) \tag{16}$$

Entropy is the average amount of information conveyed by an event when considering all possible outcomes or events drawn from the probability distribution. It is also used to measure disorder. It is also used to measure the variation of the uncertainty of a random variable in various situations in science and engineering.

4. Method of Parameter Estimation

To estimate the parameters involved in the SLLog distribution, the expression is derived by using the maximum likelihood estimates (MLEs) method. Let $X_1, X_2, \ldots, X_n$ be a set of n samples drawn from a SLLog distribution. Then the likelihood function of this distribution is given by

$$L = \left(\frac{2\alpha}{\beta}\right)^n \prod_{i=1}^{n}\left(\frac{\left(\frac{x_i-\gamma}{\beta}\right)^{2\alpha-1}}{\left(1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\right)^3}\right) \tag{17}$$

Therefore, the log-likelihood function of the SLLog distribution becomes

$$lnL = n\, ln\left(\frac{2\alpha}{\beta}\right) + (2\alpha-1)\sum_{i=1}^{n} ln\left(\frac{x_i-\gamma}{\beta}\right) - 3\sum_{i-1}^{n} ln\left(1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\right) \tag{18}$$

The components of the score vector to estimate the parameters associated with the SLLog distribution are given by

$$\frac{\partial \ln L}{\partial \alpha} = \frac{n}{\alpha} + 2\sum_{i=1}^{n} ln\left(\frac{x_i-\gamma}{\beta}\right) - 3\sum_{i-1}^{n}\left(1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\right)^{-1}\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha} ln\left(\frac{x_i-\gamma}{\beta}\right) \tag{19}$$

$$\frac{\partial \ln L}{\partial \beta} = -\frac{n}{\beta} - n\left(\frac{2\alpha-1}{\beta}\right) + \frac{3\alpha}{\beta}\sum_{i-1}^{n}\left(1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\right)^{-1}\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha} \tag{20}$$

$$\frac{\partial \ln L}{\partial \gamma} = -\left(\frac{2\alpha-1}{\beta}\right)\sum_{i=1}^{n}\left(\frac{x_i-\gamma}{\beta}\right)^{-1} + \frac{3\alpha}{\beta}\sum_{i-1}^{n}\left(1+\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha}\right)^{-1}\left(\frac{x_i-\gamma}{\beta}\right)^{\alpha-1} \tag{21}$$

By solving the nonlinear system of equations simultaneously using suitable numerical methods by setting the score vector to zero we obtain the value of parameters α, β and γ of the SLLog distribution.

5. Application of SLLog Distribution

To test the potentiality of the proposed SLLog distribution, four real data sets are presented. To test the validity and suitability of the proposed models, Akaike's Information Criteria (AIC), and Bayesian Information Criteria (BIC) at the maximum value of Negative Log-likelihood (NLL) of probability distributions have been applied. The formulas of AIC and BIC for the fitted models are given as

$$AIC = 2k - 2\,lnL \tag{22}$$

$$BIC = k\,ln(n) - 2\,lnL \tag{23}$$

where k is the number of parameters associated with the probability distribution. n is the number of observations and lnL is the log-likelihood function at the maximum likelihood estimate of that distribution.

The first data set is taken from the Nepal Demographic and Health Survey (NDHS, 2022). Different demographers and researchers used different right-skewed probability distribution models to test the goodness of fit of Age-Specific Fertility Rates (ASFRs) of different countries viz. Peristera and Kostaki (2007) used the Normal mixture model to capture both traditional and modern distorted ASFRs. Mazzuco and Scarpa (2011) applied a flexible generalized skew Normal distribution to fit the fertility pattern of countries that experienced a bimodal-fertility schedule eg. the USA, the UK, Ireland, and countries that keep a classic fertility pattern viz. Italy and the Czech Republic. Gaire and Aryal (2015) applied inverse Gaussian model to describe the distribution pattern of ASFRs of Nepalese mothers. Asili et al. (2014) used skew-logistic probability to fit ASFRs of Italy and the same model was applied to fit the ASFRs of India by Mishra et al. (2017). A polynomial model was used by Gaire et al. (2022). In this paper, the proposed SLLog model is applied to the age of the mother at the birth of a child to fit ASFRs of Nepal, and the results are compared with baseline distribution which are presented in Table 2.

Table 2: Parameter estimation and different test statistics for the age of the mother at the birth of a child

Distribution	Parameters			NLL	AIC	BIC
	α	β	γ			
LLog	24.612	0.321	17.445	-602.39	1210.77	1210.61
SLLog	5.958	22.914	0.000	-28.280	60.360	60.252

The second data set consists of 100 observations of the waiting time (minutes) of a customer at the bank before receiving the service and it has been taken from Ghitany et al. (2008) and recently the same data was applied to Skew-Lomax distribution by Gaire (2022). The values of parameters and the result of test statistics have been presented in Table 3.

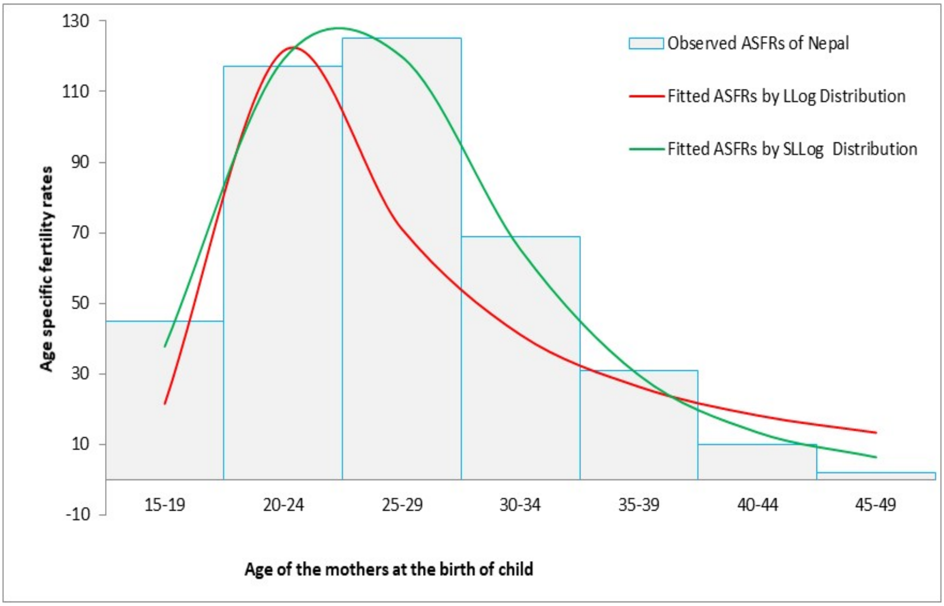

Figure 6: Empirical and fitted ASFRs of Nepalese Mothers

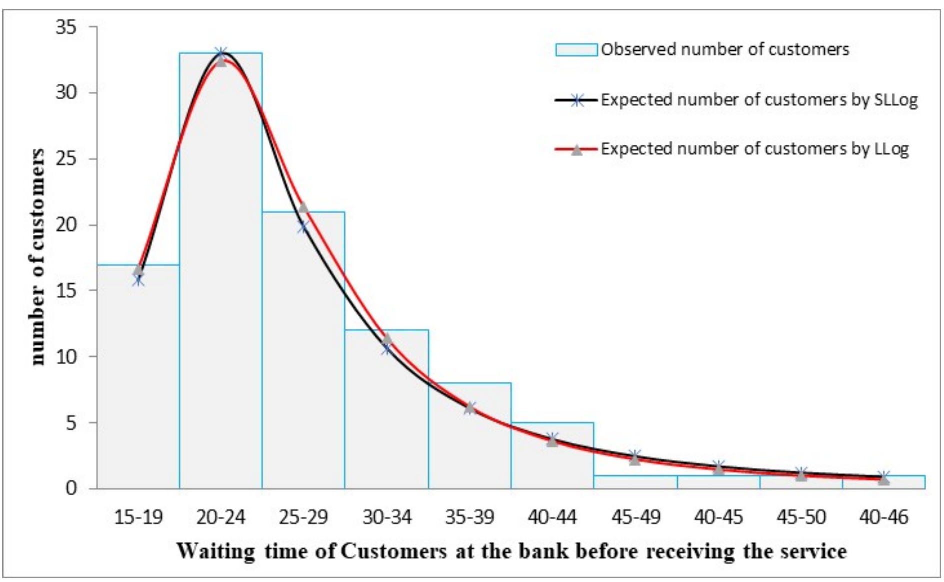

Figure 7: Empirical and fitted number of customers waiting for a service at the bank

The third data set consists of 10,631 data of age at first marriage of Nepalese women taken from (NDHS, 2022). The values of the estimated parameters along with the test statistics have been presented in Table 4.

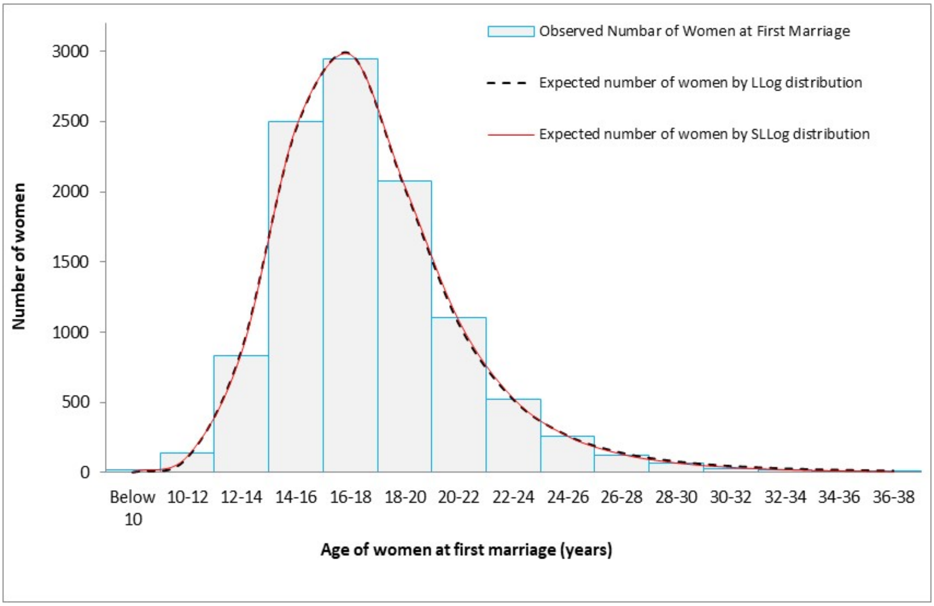

Figure 8: Empirical and fitted number of Nepalese women with age at first marriage

Table 3: Parameter Estimation and different test statistics for waiting time of customers

Distribution	Parameters			NLL	AIC	BIC
	α	β	γ			
LLog	2.185	7.935	0.198	-44.135	94.270	95.178
SLLog	1.811	5.031	0.000	-43.703	93.407	94.315

Finally, the fourth data set consists of 14,349 data on the age of girls at menarche and has been taken from (NDHS, 2022). The value of the estimated parameter along with the test statistics have been presented in Table 5.

Table 4: Parameter estimation and different test statistics for age at first marriage

Distribution	Parameters			NLL	AIC	BIC
	α	β	γ			
LLog	4.694	8.349	8.883	-80.837	167.674	169.799
SLLog	7.994	14.946	0.522	-74.264	154.528	156.446

In general smaller values of NLL, AIC, and BIC values of goodness of fit of the probability distribution suggest the best fit to the data. The values of AIC and BIC at the maximum likelihood estimate for the proposed SLLog distribution are lower than that of the LLog distribution for all four data sets. This clearly showed that the proposed model is flexible enough to fit the data better than that of the base distribution.

Table 5: Parameter estimation and different test statistics for the age of girls at menarche

Distribution	Parameters			NLL	AIC	BIC
	α	β	γ			
LLog	7.420	6.471	7.677	-63.30	132.60	134.77
SLLog	13.644	13.269	0.000	-49.985	105.970	107.88

6. Conclusions

A new three-parameter skew probability distribution model has been formulated as the SLLog distribution. Some of the statistical properties of the distribution have been studied. The parameter estimation method is discussed by using maximum likelihood. To test the suitability and validity of the proposed model four real data sets, viz. age of the Nepalese mother at the birth of a child, the waiting time of the customer before receiving the service, the age at first marriage of Nepalese female, and the age of Nepalese girls at menarche have been used. The AIC and BIC test criteria have been applied to test the validity and suitability of the model obtained at the maximum value of negative log-likelihood of the probability distribution. The observed values of AIC and BIC show that the proposed distribution is more flexible than the baseline distribution to fit the pattern of these real data sets.

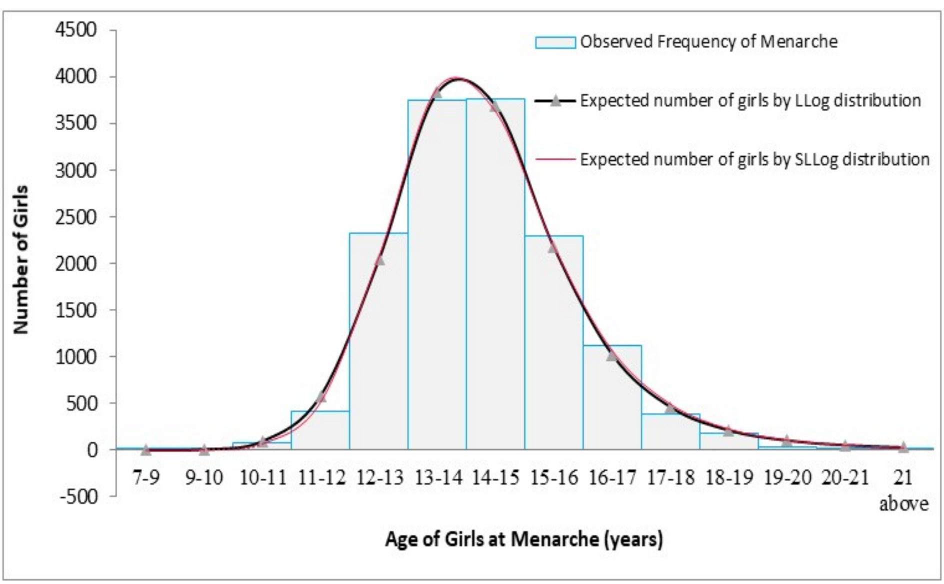

Figure 9: Empirical and fitted number of Nepalese girls at the age of menarche

Acknowledgment

The first author would like to acknowledge the travel grant support 2019 by UGC Nepal to present the research at an international conference. Both authors wish to thank the Chief Editor, the Editors, and anonymous referees for their comments and suggestions which have greatly improved our manuscript.

References

Adeyinka, F. S., and Olapade, A. K., (2019). On transmuted four parameters generalized Log-Logistic distribution. *International Journal of Statistical Distributions and Applications*, 5(2), pp. 32–37.

Alexander, C., Cordeiro, G. M., Ortega, E. M. and Sarabia, J. M. (2012). Generalized Beta-generated distributions. *Computational Statistics and Data Analysis*, 56(6), pp. 1880–1897.

Aryal, G. R., (2013). Transmuted log-logistic distribution. *Journal of Statistics Applications and Probability*, 2(1), pp. 11–20.

Ashkar, F., Mahdi, S., (2006). Fitting the log-logistic distribution by generalized moments. *Journal of Hydrology*, 328(3-4), pp. 694–703.

Asili, S., Rezaei, S. and Najjar, L., (2014). Using skew-logistic probability density function as a model for age-specific fertility rate pattern. *BioMed Research International*, 10, pp. 1–5.

Azzalini A., (1985). A class of distributions that includes the normal ones. *Scandinavian Journal of Statistics*, pp. 171–178.

Azzalini A., (2005). The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, 32(2), pp. 159–188.

Azzalini, A., Capitanio, A., (2013). *The skew-normal and related families* (Vol. 3), London: Cambridge University Press.

Barlow, R. E., Marshall, A. W. and Proschan, F., (1963). Properties of probability distributions with monotone hazard rate. *The Annals of Mathematical Statistics*, 34(2), pp. 375–389.

Collett, D., (2015). *Modeling survival data in medical research.* Boca Raton, Florida USA: CRC press.

Cordeiro, G. M., De-Castro M., (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81, pp. 883–898.

Cordeiro, G. M., Ortega, E. M. and Da-Cunha, D. C., (2013). The exponentiated generalized class of distributions. *Journal of Data Science*, 11(1), pp. 1–27.

De-Santana, T. V. F., Ortega, E. M., Cordeiro, G. M. and Silva, G. O., (2012). The Kumaraswamy-log-logistic distribution. *Journal of Statistical Theory and Applications*, 11(3), pp. 265–291.

Eugene, N., Lee, C. and Famoye, F., (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, 31(4), pp. 497–512.

Fisk, P. R., (1961). The graduation of income distribution. *Econometrica*, 29(2), pp. 171–185.

Gaire, A. K., Aryal, R., (2015). Inverse Gaussian model to describe the distribution of age-specific fertility rates of Nepal. *Journal of Institute of Science and Technology*, 20(2), pp. 80–83.

Gaire, A. K., Thapa G. B. and KC, S., (2019). *Preliminary results of Skew Log-logistic distribution, properties, and application. Proceeding of the 2nd International Conference on Earthquake Engineering and Post Disaster Reconstruction Planning*, 25–27 April 2019, Bhaktapur, Nepal, pp. 37–43.

Gaire, A. K., Thapa, G. B. and KC, S., (2022). Mathematical modeling of age-specific fertility rates of Nepali mothers. *Pakistan Journal of Statistics and Operation Research*, 18(2), pp. 417–426. https://doi.org/10.18187/pjsor.v18i2.3319.

Gaire, A. K., (2022). Skew Lomax distribution, parameter estimation, its properties, and applications. *Journal of Science and Engineering*, 10, pp. 1–11.

Ghitany, M.E., Atieh, B. and Nadarajah, S., (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78(4), pp. 493–506.

Gradshteyn, I. S., Ryzhik, I. M., (2000). *Table of integrals, series, and products*, San Diego, CA: Academic Press.

Gupta, A. K., Chang, F. C. and Huang, W. J., (2002). Some skew-symmetric models. *Random Operators and Stochastic Equations*, 10(2), pp. 133–140.

Gui, W., (2013). Marshall-Olkin extended log-logistic distribution and its application in minification processes. *Applied Mathematical Science*, 7(80), pp. 3947–3961.

Harvda, J., Charvat, F., (1967). Quantification method of classification processes. Concept of structural a-entropy. *Kybernetika*, 3(1), pp. 30–35.

Hassan, A. S., Hemeda, S. E., (2016). The additive Weibull-G family of probability distributions. *International Journals of Mathematics and Its Applications*, 4(2), pp. 151-164.

Hemeda, S., (2018). Additive Weibull Log Logistic distribution: Properties and application. *Journal of Advanced Research in Applied Mathematics and Statistics*, 3(4), pp. 8–15.

Hamedani, G., (2013). The Zografos-Balakrishnan log-logistic distribution: Properties and applications. *Journal of Statistical Theory and Applications*, 12(3), pp. 225–244.

Jones, M., (2004). Families of distributions arising from distributions of order statistics. *Test*, 13(1), pp. 1–43.

Kleiber, C., Kotz, S., (2003). *Statistical size distributions in economics and actuarial sciences* (Vol. 470). New York: John Wiley and Sons.

Lawless, J. F., (2003). *Statistical models and methods for lifetime data*. Vol. 362. New York: John Wiley and Sons.

Lemonte, A. J., (2014). The Beta log-logistic distribution. *Brazilian Journal of Probability and Statistics*, 28(3), pp. 313–332.

Lima, S. R., Cordeiro, G. M., (2017). The extended Log-Logistic distribution: Properties and application. *Anais da Academia Brasileira de Ciências*, 89(1), pp. 3–17.

Marshall, A. W., Olkin, I., (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84(3), pp. 641–652.

Mazzuco, S., Scarpa, B., (2011). *Fitting an age-specific fertility rate by Skew-symmetrical probability density function, the University of Padova, Working paper Series*, Italy, 10, pp. 1–18.

Mishra, R., Singh, K. K. and Singh, A., (2017). A model for age-specific fertility rate pattern of India using skew-logistic distribution function. *American Journal of Theoretical and Applied Statistics*, 6(1), pp. 32–37.

Nadarajah, S., (2009). The skew logistic distribution. *Advances in Statistical Analysis*, 93(2), pp. 187–203.

NDHS, (2022). *Nepal Demographic and Health Survey 2022: Key Indicators Report.* Kathmandu, Nepal: Ministry of Health and Population; New ERA; and ICF., Nepal.

Peristera, P., Kostaki, A., (2007). Modeling fertility in modern populations. *Demographic Research*, 16, pp. 141–194.

Renyi, A., (1961). *On measures of entropy and information, Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, 1, pp. 547–561, Barkeley: The University of California Press.

Rodrıguez, G., (2010). *Parametric survival models.* New Jersey: Rapport technique, Princeton University.

Rosaiah, K., Nagarjuna, K. M., Kumar, D. C. U. S. and Rao, B. S., (2014). Exponential-log-logistic additive failure rate model. *Int J Sci Res Publ.*, 4(3), pp. 1–5.

Shaw, W. T., Buckley, I. R., (2007). *The alchemy of probability distributions: Beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map*. arXiv preprint arXiv:0901.0434.

Surendran, S., Tota-Maharaj, K., (2015). *Log logistic distribution to model water demand data*. Procedia Engineering, 119, pp. 798–802.

Tadikamalla, P. R., (1980). A look at the Burr and related distributions. *International Statistical Review/Revue Internationale de Statistique*, pp. 337–344.

Tahir, M. H., Mansoor, M., Zubair, M. and Hamedani, G., (2014). McDonald log-logistic distribution with an application to breast cancer data. *Journal of Statistical Theory and Applications*, 13(1), pp. 65–82.

Tsallis, C., (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52, pp. 479–487.

Yilmaz, V., Erişoğlu, M. and Çelik, H. E., (2011). *Probabilistic prediction of the next earthquake in the NAFZ (North Anatolian Fault Zone)*, Turkey: Doğuş Üniversitesi Dergisi, 5(2), pp. 243–250.

Zografos, K., Balakrishnan, N., (2009). On families of Beta-and generalized Gamma-generated distributions and associated inference. *Statistical Methodology*, 6(4), pp. 344–362.