SKEW LOMAX DISTRIBUTION: PARAMETER ESTIMATION, ITS PROPERTIES, AND APPLICATIONS

Arjun Kumar Gaire[1,2] *

[1]Department of Science and Humanities, Khwopa Engineering College, Purbanchal University
[2]Central Department of Population Studies, Faculty of Humanities and Social Sciences, Tribhuvan University

Abstract

In the domain of the univariate distribution a large number of new distributions were introduced by using different generators. In this paper, a three-parameter distribution called the 'Skew-Lomax' distribution is proposed, which is the special case of the Azzalini distribution to generalize the Lomax distribution. The Lomax distribution is also called Pareto type II distribution, which is a heavy-tailed continuous probability distribution for a non-negative random variable. The statistical properties of the proposed Skew-Lomax distribution, including mean, variance, moments about the origin, cumulative distribution function, hazard rate function, quantile function, and random number generation have been derived. Also, the method of maximum likelihood and the method of moment to estimate the parameters of this distribution have been proposed. Three real data sets have been used to illustrate the usefulness, flexibility, and application of the proposed distribution. The coefficient of determination, chi-square test statistics, and the sum of the square of error depict that the proposed model is more flexible than the Lomax distribution.

Keywords: Age-specific fertility rate, Lomax, Peak flow, Reliability analysis, Waiting time.

1. Introduction

In the recent research on uni-variate probability distribution, different researchers, mathematicians, statisticians, and other social scientists formulate new family of distributions. Usually they extend, generalize and change the existing distribution by adding extra (shape, scale, location, or threshold) parameters. The primary aim of such modifications is the formation of new flexible distribution, which helps to better fit the data in a real-life situation. In the same line, the Lomax distribution has been modified by using the proposition of Azzalini and Capitanio (2013). Some structural properties of new distribution including mean, variance, cumulative distribution, reliability function, hazard rate function, cumulative hazard rate function, random number generation and the methods of parameter estimation by using method of moments and maximum likelihood method have been derived. Also, the model has been applied to three real data sets of the age of the mother at the birth of a child; the waiting time of customers in a bank before receiving the service, and the yearly peak flow of the river to show the flexibility and suitability of proposed model. The Chi-square test statistics, Sum of square of error (SSQ) and the coefficient of determination between observed and expected frequencies of three data sets have been used as validity tools of data fitting. Different test statistics and the graph of fitted results depict that the model is flexible enough to capture the real life data.

The remaining sections of this paper are organized as follows. Section 2 gives the brief introduction of Lomax distribution. In section 3, the PDF and the CDF of a three-parameter new distribution called Skew-Lomax (SL) distribution have been formulated and presented. Section 4 the some Mathematical and the statistical properties of the Skew Lomax distribution have been derived. Section 5 comprises the reliability analysis. Section 6 the proposed

*Corresponding author: Arjun Kumar Gaire
Department of Science and Humanities, Khwopa Engineering College
Email: arjun.gaire@gmail.com
(Received: November 10, 2020, Accepted: June 30, 2022)
https://doi.org/10.3126/jsce.v10i1.61011

methods of parameter estimation have been presented by using Method of Moments and method of likelihood and section 7 illustrates the applications, flexibility, and validity of the model by using three actual data sets of Age of Nepali mother at the birth of a child; yearly peak flow of the Bagmati river of Nepal; and waiting time customer at the bank before receiving the service. Finally, section 8 concludes the paper.

2. Lomax Distribution

Since Lomax distribution is not flexible enough to model and fit the real-life data (Alghamdi, 2018). So, the changed version of Lomax distribution is always desirable among the researchers. The Lomax distribution, a heavy-tailed probability distribution, is one of the continuous probability distribution functions for a non-negative random variable that is very useful in numerous fields of life testing, engineering, reliability analysis, and other various real-world situations. This distribution is named after K. S. Lomax (1954). It has been observed for adequately modeling income and wealth distributions, actuarial science, queuing theory, and Internet traffic modeling. However, Alghamdi (2018) states that the modeling data by using this distribution does not provide great flexibility. The probability density function (PDF) and the cumulative distribution function (CDF) of Lomax distribution with two parameters α and β is given in Equations (1-2) as

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

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

where $\alpha > 0$ is the shape parameter and $\beta > 0$ is the scale parameter. In many research, we have positive data greater than a certain number. For example, the age of the mother at the birth of a child in case of Age-specific fertility rates (ASFRs) which is greater than zero. In such cases, the researcher feels the threshold parameter. Therefore, if we add the third parameter $\gamma > 0$, the location parameter, to guarantee the time before which no failure occurs. Then, the PDF and the CDF of Lomax distribution with three parameters expressed in Equation (3) and Equation (4) are as:

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

$$G(x,\alpha,\beta,\gamma) = 1-\left(1+\frac{x-\gamma}{\beta}\right)^{-(\alpha)}, \text{for } x>\gamma \tag{4}$$

The k^{th} order moment of the Lomax distribution is defined only when the condition $\alpha > k$ and defined in Equation (5) as:

$$E(x^k) = \sum_{i=0}^{k}\binom{k}{i}\gamma^{k-i}\beta^i\frac{\Gamma(i+1)\Gamma(\alpha-i)}{\Gamma(\alpha+1)} \tag{5}$$

where $\Gamma(n)$ is a Gamma function defined by

$$\Gamma(n) = \int_0^\infty e^{-t}t^{n-1}dt$$

Here, Equations(3), (4), and (5) are the routine formula of Lomax distribution including the third parameter γ as a location parameter. In particular, the mean and variance of Lomax distribution with three parameters are: $E(X) = \gamma + \frac{\beta}{\alpha-1}$ and

$$\text{Variance}(X) = \begin{cases} \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}, \text{for } \alpha > 2 \\ \infty, \text{for } r1 < \alpha \le 2 \\ \text{Undefined , Otherwise} \end{cases}$$

Lomax distribution is usually used by researchers to model the data that are heavily tailed. It was used to model the income and wealth distribution (Harris, 1968; Atkinson and Harrison, 1978), to study the distributional patterns of computer files on the server by Holland et al. (2006), reliability analysis and life testing research by Hassan and Al-Ghamdi (2009) and many more.

3. Skew Lomax Distribution

Azzalini (1985, 2005) generalized Normal distribution by adding an extra asymmetry parameter $\lambda > 0$ and the PDF of the new Skew-Normal distribution by using the following relation as

$$f(z) = 2\phi(z)\ \Phi(\lambda Z),\ (z,\lambda)\ \epsilon\ R \tag{6}$$

Here, $\phi(z)$ and $\Phi(z)$ are the PDF and CDF of Standard Normal distribution. To construct the skew-symmetrical distribution other than the Standard Normal distribution, Azzalini and Capitanio (2013) used Azzalini's proposition in their book:

$$f(x) = 2\ g(x)\ G(x), x\ \epsilon R \tag{7}$$

where $g(x)$ and $G(x)$ are the PDF and CDF of any baseline distribution. By using this concept skew-uniform, skew- t, skew-Cauchy, skew-Laplace, and skew-logistic-distributions have been introduced by Gupta et al. (2002). Huang and Chen (2007) studied the generalized skew-Cauchy distribution. Later, Skew-Logistic distribution has been studied in detail (Nadarajah, 2009). The value of the CDF $G(x)$ and PDF $g(x)$ of the base distributions chosen in all the above cases were symmetrical about the origin. However, Shaw and Buckley (2009) suggested to choose any distribution other than the symmetrical one.

Based on this assumption Gaire et al. (2019) proposed and studied Skew-Log-Logistic distribution by choosing a non-symmetrical Log-Logistic distribution as a base distribution. In the same line, in this research, The Lomax distribution has been chosen as the base distribution, which is already right-skewed, aiming to construct distributions with flexible than the Lomax distribution.

In the literature on uni variate probability distribution, several generalized forms of different probability distribution have been found. Similarly, Lomax distribution is also extended and generalized by introducing additional shape, scale, or location parameters by different researchers. Further, they derived and presented the characteristics and the application in various fields. Some of the generalized versions of Lomax distributions found in the literature are Marshall-Olkin extended Lomax (M. Ghitany et al., 2007); Beta-Lomax (Rajab et al., 2013); Kumaraswamy Lomax, and McDonald Lomax (Lemonte and Cordeiro, 2013); Kumaraswamy-generalized Lomax (Shams, 2013); Transmuted Exponentiated Lomax (Ashour and Eltehiwy, 2013); Kumaraswamy Exponentiated Lomax (El-Batal and Kareem, 2014); Gamma-Lomax (Cordeiro et al., 2015); Exponential Lomax (El-Bassiouny et al., 2015); Weibull Lomax (Tahir et al., 2015); Gumbel-Lomax (Tahir et al., 2016); Power Lomax (Rady et al., 2016); Exponentiated Weibull-Lomax (Hassan and Abd-Allah, 2018) and Rayleigh Lomax (Fatima et al., 2018). At this juncture, here we introduce the Skew-Lomax distribution as a special case of Azzalini's proposition based on Azzalini and Capitanio (2013).

3.1. Probability Density Function of Skew-Lomax distribution

Let us suppose X is a random variable following Lomax distribution with PDF $g(x)$ and CDF $G(x)$ then substituting the values from Equations (3) and Equation (4) into Equation (7) the density function of the SL distribution is given as in Equation (8):

$$f(x) = 2\alpha\beta^{\alpha}\left(\frac{(x+\beta-\gamma)^{\alpha}-\beta^{\alpha}}{(x+\beta-\gamma)^{2\alpha+1}}\right), x \geq \gamma \quad (8)$$

Here, $f(x)$ is a probability density function since, $\int_0^{\infty} f(x)dx = 1$ with $f(x) \geq 0$ for every value of $x > \gamma$ as:

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

Let us suppose, $x+\beta-\gamma$ = t, then $dx = dt$,

When $x = \gamma, t = \beta$ and when $x = \infty, t = \infty$

$$\begin{aligned}\int_{\gamma}^{\infty} f(x)dx &= 2\alpha\beta^{\alpha}\int_{\beta}^{\infty}\left(\frac{t^{\alpha}-\beta^{\alpha}}{t^{2\alpha+1}}\right)dt \\ &= 2\alpha\beta^{\alpha}\int_{\beta}^{\infty}\left(t^{-\alpha-1}-\beta^{\alpha}t^{-2\alpha-1}\right)dt \\ &= 2\alpha\beta^{\alpha}\left(\frac{-1}{\alpha t^{\alpha}}+\frac{\beta^{\alpha}}{2\alpha t^{2\alpha}}\right)\Big|_{\beta}^{\infty} \\ &= 1\end{aligned}$$

The graph of the PDF of the SL distribution is presented in Figure 1 by using the Equation (8). Similarly, the graph of the CDF of the SL distribution is presented in Figure 2 by using Equation (9). Both graphs have been plotted using R-language software.

3.2. Cumulative Distribution Function of the SL distribution

The cumulative distribution function of the SL distribution is given by

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

Thus, the value of CDF of the Skew Lomax distribution as expressed in Equation (9) as,

$$F(x) = 1 + \frac{\beta^{2\alpha}}{(x+\beta-\gamma)^{2\alpha}} - \frac{2\beta^{\alpha}}{(x+\beta-\gamma)^{\alpha}} \quad (9)$$

4. Some Structural Properties

To provide a comprehensive description of the proposed model, this section deals with the computation of some statistical and mathematical properties of the SL distribution comprising quantile function, random number generation, Central Moments, mean, and variance as:

4.1. Quantile function and random number generation

The quantile function of distribution is obtained by inverting the CDF of the any distribution. Quantile function is widely applies in a statistical study for the calculation of quartiles and random number generation. For this, let us equate the CDF of the SL distribution with a number $'U'$ that follows uniform distribution in an interval $[0, 1]$.

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

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

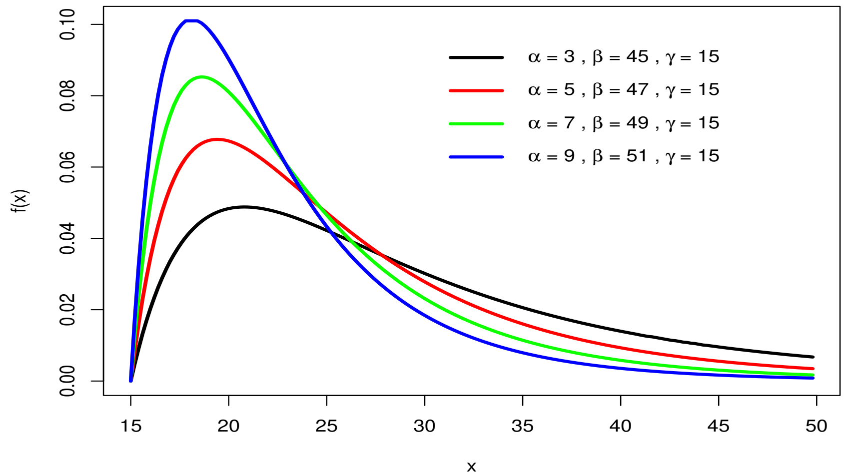

Figure 1. PDF of the SL Distribution

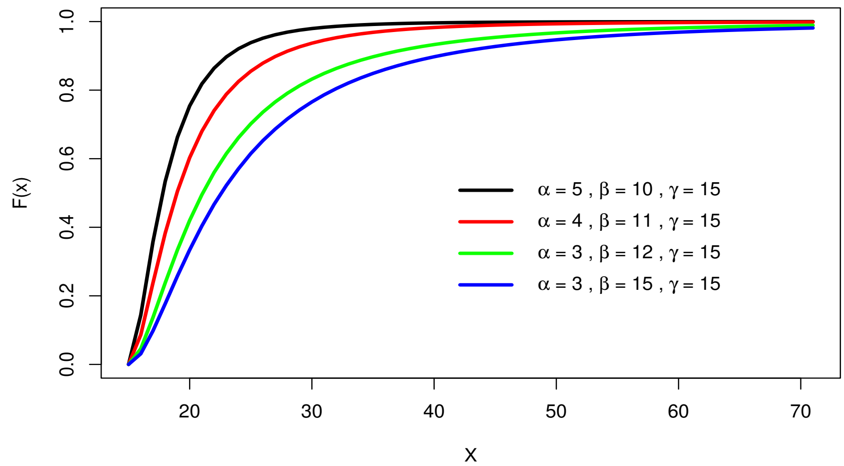

Figure 2. CDF of the SL Distribution

Let us suppose,

$$(x + \beta - \gamma)^\alpha = y$$

Then

$$(1 - U)y^2 - 2\beta^\alpha y + \beta^{2\alpha} = 0$$

which is quadratic in y. By solving this quadratic equation we get,

$$y = \frac{\beta^\alpha(1 \pm \sqrt{U})}{1 - U}$$

Thus, the value of a random variable X of the SL distribution is obtained as:

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

For the known value of parameters α, β and γ we can generate the random number X of the SL distribution by using Equation (10). Further, the First, second, and third quartiles are obtained by setting $u = 0.25, u = 0.5$, and $u = 0.75$ respectively, which is helpful for the computation of skewness and kurtosis of the distribution. In statistical analysis the concept of random number generation is used to generate the set of random number which follows certain distribution by setting the scenario in future. For example, if we are modeling the waiting time of customer in a bank before receiving the service assumed to follow the Skew Lomax distribution we can use this concept to generate set of random data of waiting time with different value of parameters and able to use simulation method to find the different scenarios in future so that we can make policy accordingly.

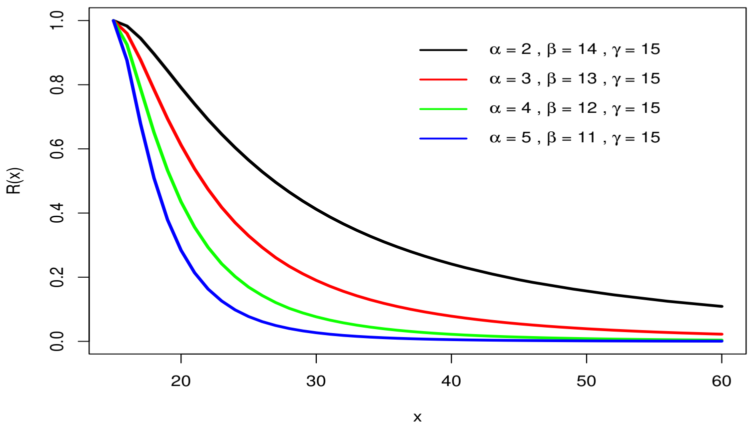

Figure 3. Reliability Function of the SL Distribution

4.2. Central moments

The k^{th} order moments of the SL distribution about the origin are defined as

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

Let us suppose $(x+\beta-\gamma) = t$, Then $dx = dt$, When $x = \gamma, t = \beta$ and when $x = \infty, t = \infty$

$$E(X^k) = 2\alpha\beta^\alpha \int_\beta^\infty (t-\beta+\gamma)^k \left(\frac{t^\alpha - \beta^\alpha}{t^{2\alpha+1}}\right) dt$$
$$= 2\alpha\beta^\alpha \int_\beta^\infty \left(\frac{(t-\beta+\gamma)^k}{t^{\alpha+1}} - \frac{\beta^\alpha (t-\beta+\gamma)^k}{t^{2\alpha+1}}\right) dt$$

Again, let us suppose $t = (\beta-\gamma)\ Sec^2\theta$ then
$dt = 2(\beta-\gamma)\ sec^2\theta\ tan\theta\ d\theta$,
When $t = \beta, \theta = 0$ and when $t = \infty, \theta = \frac{\pi}{2}$

$$E(X^k) = 4\alpha\beta^\alpha \int_0^{\frac{\pi}{2}} \frac{(\beta-\gamma)^{k+1}(sec^2\theta-1)^k sec^2\theta\ tan\theta}{(\beta-\gamma)^{\alpha+1}(sec^2\theta)^{\alpha+1}} d\theta$$
$$- 4\alpha\beta^{2\alpha} \int_0^{\frac{\pi}{2}} \frac{(\beta-\gamma)^{k+1}(sec^2\theta-1)^k sec^2\theta\ tan\theta}{(\beta-\gamma)^{2\alpha+1}(sec^2\theta)^{2\alpha+1}} d\theta$$

$$E(X^k) = 4\alpha\beta^\alpha(\beta-\gamma)^{k-\alpha} \int_0^{\frac{\pi}{2}} (sin\theta)^{2k+1}\ (cos\theta)^{2\alpha-2k-1}\ d\theta$$
$$- 4\alpha\beta^{2\alpha}(\beta-\gamma)^{k-2\alpha} \int_0^{\frac{\pi}{2}} (sin\theta)^{2k+1}\ (cos\theta)^{4\alpha-2k-1}\ d\theta$$

$$E(X^k) = 4\alpha\beta^\alpha(\beta-\gamma)^{k-\alpha} B(k+1, \alpha-k)$$
$$- 4\alpha\beta^{2\alpha}(\beta-\gamma)^{k-2\alpha} B(k+1, 2\alpha-k)$$

$$E(X^k) = 4\alpha\beta^\alpha(\beta-\gamma)^{k-\alpha} \frac{\Gamma(k+1)\Gamma(\alpha-k)}{\Gamma(\alpha+1)}$$
$$- 4\alpha\beta^{2\alpha}(\beta-\gamma)^{k-2\alpha} \frac{\Gamma(k+1)\Gamma(2\alpha-k)}{\Gamma(2\alpha+1)} \quad (11)$$

It is to be noted that moments of the SL distribution exist only for $\alpha > k$. In particular, the mean and variance of the SL distribution have been obtained as

$$E(X) = \frac{4\beta^\alpha(\beta-\gamma)^{1-\alpha}}{(\alpha-1)} - \frac{2\beta^{2\alpha}(\beta-\gamma)^{1-2\alpha}}{(2\alpha-1)} \quad (12)$$

$$E(X^2) = \frac{8\beta^\alpha(\beta-\gamma)^{2-\alpha}}{(\alpha-1)(\alpha-2)} - \frac{4\beta^{2\alpha}(\beta-\gamma)^{2-2\alpha}}{(2\alpha-1)(2\alpha-2)} \quad (13)$$

Now, The Variance of the SL distribution is given as

$$Var = \left\{\frac{8\beta^\alpha(\beta-\gamma)^{2-\alpha}}{(\alpha-1)(\alpha-2} - \frac{4\beta^{2\alpha}(\beta-\gamma)^{2-2\alpha}}{(2\alpha-1)(2\alpha-2)}\right\}$$
$$- \left\{\frac{4\beta^\alpha(\beta-\gamma)^{1-\alpha}}{(\alpha-1)} - \frac{2\beta^{2\alpha}(\beta-\gamma)^{1-2\alpha}}{(2\alpha-1)}\right\}^2$$

5. Reliability Analysis of the SL Distribution

The reliability or survival function of a random variable X is the probability of an event not failing before some time x. That is the probability that the system will survive beyond a specified time, and it is defined as $R(x) = 1 - F(x)$. Thus, the reliability function of the SL distribution is expressed in Equation (14). The plot of the reliability function of the SL distribution for different values of parameters are presented in Figure 3.

$$R(x) = \frac{2\beta^\alpha}{(x+\beta-\gamma)^\alpha} - \frac{\beta^{2\alpha}}{(x+\beta-\gamma)^{2\alpha}} \quad (14)$$

The hazard rate function of a random variable X is another characteristic of interest defined as $h(x) = \frac{f(x)}{1-F(x)}$. Thus, the hazard rate function for the SL distribution, which is the conditional probability of failure, given that it has survived up to the time x is given in Equation (15) as

$$h(x) = \frac{2\alpha\{(x+\beta-\gamma)^{\alpha} - \beta^{\alpha}\}}{(x+\beta-\gamma)\{2(x+\beta-\gamma)^{\alpha} - \beta^{\alpha}\}} \tag{15}$$

Similarly, the cumulative hazard rate function of the SL distribution is defined by $H(x) = -ln(R(x))$ is given in Equation (16) as

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

$$H(x) = ln\left\{\frac{(x+\beta-\gamma)^{2\alpha}}{2\beta^{\alpha}(x+\beta-\gamma)^{\alpha} - \beta^{2\alpha}}\right\} \tag{16}$$

The graphs of the hazard and cumulative hazard rate functions of the SL distribution for some selected values of the parameters have been presented in Figure 4, and Figure 5 respectively.

5.1. Entropy measure of the SL distribution

Entropy, which is used in various situations in science and engineering, measures the uncertainty of a continuous random variable X. In this section expressions of different entropy are discussed in detailed.

Renyi Entropy

The Renyi entropy (Rényi, 1961) of the SL distribution with PDF $f(x)$ is given as:

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

where $\rho > 0$ and $\rho \neq 1$. The integral in $I_R(\rho)$ of the SL distribution can be defined as:

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

Let us suppose $(x+\beta-\gamma) = t$, then $dx = dt$,
When $x = \gamma, t = \beta$ and $x = \infty, t = \infty$

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

Again, let us suppose $t = \beta Sec^2\theta$
then $dt = 2\beta sec^2\theta \; tan\theta \; d\theta$ when $t = \beta, \theta = 0$ and $t = \infty, \theta = \frac{\pi}{2}$

$$\int_{\gamma}^{\infty} (f(x))^{\rho} dx = 2^{\rho+1}\alpha^{\rho}\beta^{1-\rho}\int_{0}^{\frac{\pi}{2}} (sin\theta)^{2\rho+1}(cos\theta)^{2\alpha\rho-3} d\theta$$
$$= 2^{\rho+1}\alpha^{\rho}\beta^{1-\rho}B(\rho+1, \alpha\rho-1)$$

Therefore, The Renyi entropy of the SL distribution can be expressed as:

$$I_R(\rho) = \frac{1}{1-\rho} log\left(2^{\rho+1}\alpha^{\rho}\beta^{1-\rho}B(\rho+1, \alpha\rho-1)\right) \tag{17}$$

The q-Entropy

Similarly, the q- entropy introduced by Tsallis (1988) is defined by:

$$H_q(x) = \frac{1}{1-q} log\left(1 - \int (f(x))^{q} dx\right) \tag{18}$$

where $q > 0$ and $q \neq 1$. So, q-entropy of the SL distribution is given as:

$$H_q(x) = \frac{1}{1-q} log\left(1 - 2^{q+1}\alpha^{q}\beta^{1-q}B(q+1, \alpha q-1)\right) \tag{19}$$

6. Method of Estimation of Parameters

6.1. Method of moments

We can obtain the first three theoretical moments of the proposed distribution about the origin by putting $k = 1, 2,$ and 3 in the equation (11) expressed as.

$$E(X^k) = 4\alpha\beta^{\alpha}(\beta-\gamma)^{k-\alpha}\frac{\Gamma(k+1)\Gamma(\alpha-k)}{\Gamma(\alpha+1)}$$
$$- 4\alpha\beta^{2\alpha}(\beta-\gamma)^{k-2\alpha}\frac{\Gamma(k+1)\Gamma(2\alpha-k)}{\Gamma(2\alpha+1)}$$

The value of the parameters α, β and γ can be obtained by equating these theoretical moments with the corresponding sample moments.

6.2. Method of maximum likelihood

For the maximum likelihood estimates (MLEs) of the parameters involved in the SL distribution, let $X_1, X_2, ..., X_n$ be a set of n samples drawn from the SL distribution. Then the likelihood function of the SL distribution is given by

$$L = (2\alpha\beta^{\alpha})^n \prod_{i=1}^{n} \frac{((x_i+\beta-\gamma)^{\alpha} - \beta^{\alpha})}{(x_i+\beta-\gamma)^{2\alpha+1}} \tag{20}$$

Now, After taking the natural logarithm on both sides of Equation (20), it gives the log-likelihood function (lnL) of the SL distribution as

$$lnL = nln(2\alpha\beta^{\alpha}) - (2\alpha+1)\sum_{i=1}^{n} ln(x_i+\beta-\gamma)$$
$$+ \sum_{i=1}^{n} ln((x_i+\beta-\gamma)^{\alpha} - \beta^{\alpha})$$

The components of the score vector to estimate the parameter of the SL distribution are given by

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

$$\frac{\partial lnL}{\partial\beta} = \frac{n\alpha}{\beta} - (2\alpha+1)\sum_{i=1}^{n}\frac{1}{(x_i+\beta-\gamma)} + \alpha\sum_{i=1}^{n}\frac{(x_i+\beta-\gamma)^{\alpha-1} - \beta^{\alpha-1}}{(x_i+\beta-\gamma)^{\alpha} - \beta^{\alpha}} \tag{22}$$

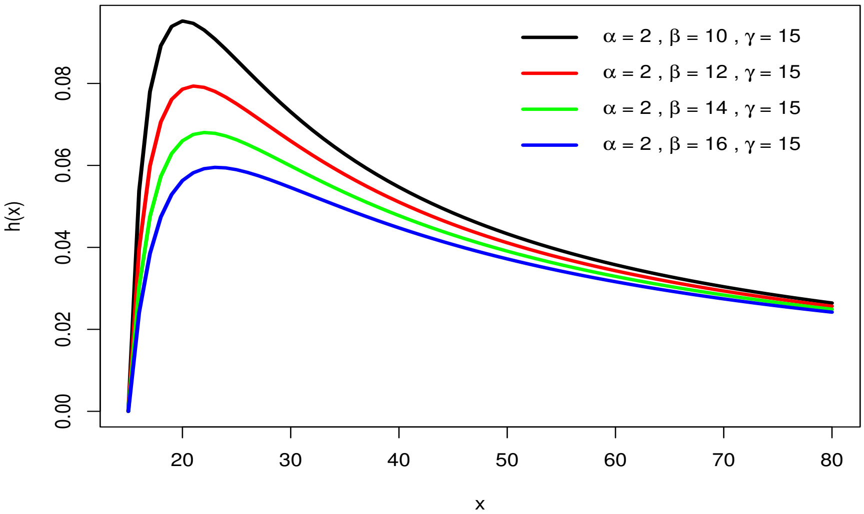

Figure 4. Hazard Rate Function of the SL Distribution

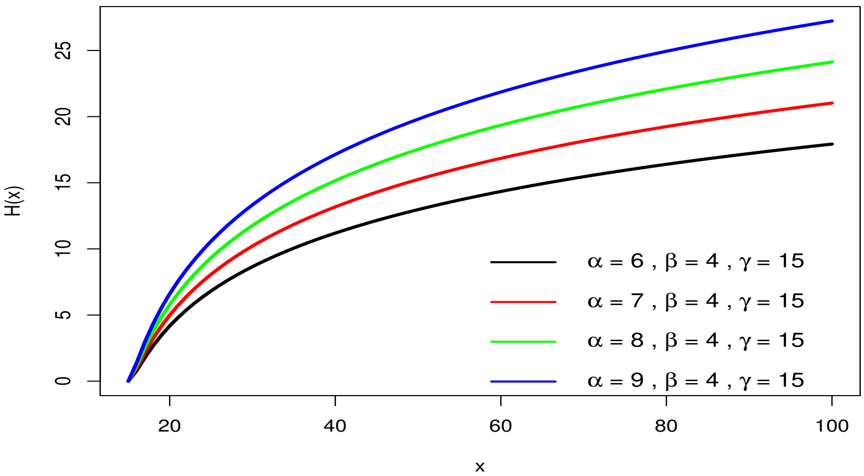

Figure 5. Cumulative Hazard Rate Function of SL Distribution

$$\frac{\partial lnL}{\partial \gamma} = (2\alpha + 1)\sum_{i=1}^{n} \frac{1}{(x_i + \beta - \gamma)} - \alpha \sum_{i=1}^{n} \frac{(x_i + \beta - \gamma)^{\alpha-1}}{(x_i + \beta - \gamma)^{\alpha} - \beta^{\alpha}} \quad (23)$$

By solving these non-linear system of equations by setting these score vectors to zero, we get the values of parameters α, β and γ. These non-linear system of equations have to be solve by using suitable numerical methods.

7. Numerial Application of Skew Lomax Distribution

In this section, we illustrate the goodness of fit of the proposed Skew Lomax distribution by using three real lifetime data sets; The First data set is taken from the Nepal Demographic and Health Survey (NDHS, 2017). To fit the age of the mother at the birth of a child in the case of ASFRs some of the right-skewed probability models and other models had been used in literature such as the Normal Mixture model (Peristera and Kostaki, 2007), flexible generalized skew Normal distribution (Mazzuco and Scarpa, 2011), Inverse Gaussian model (Gaire and Aryal, 2015), Polynomial Model (Gaire et al., 2022).

In this article, the proposed Skew Lomax distribution has been used to fit the Age of the mother at the birth of a child of Nepali mothers and the fitted results are compared with the Lomax distribution and the results are presented in Table 1.

Similarly, the second data set consists of the yearly peak flow ($1000m^3/sec$) of Bagmati River of Nepal at the Station Pandheradovan for 36 years (1979–2015) taken from the department of hydrology and meteorology of Nepal retrieved from DHM (2020) ($https : //www.dhm.gov.np$) as: 8.4, 2.89, 6.12, 5.08, 2.15, 7.63, 2.65, 2.89, 2.67, 4.41, 5.62, 3.09, 2.1, 16, 3.7, 3.5, 3.06, 3.3,

Table 1. Empirical and fitted value of ASFRs by Lomax and SL distribution

Age of Mother at the birth of a child	Observed ASFRs	Expected ASFRs (Lomax)	Expected ASFRs (SL)
15 – 19	88	83.802	84.987
20 – 24	172	177.770	168.759
25 – 29	124	95.674	122.807
30 – 34	59	51.516	45.618
35 – 39	18	27.753	19.772
40 – 44	6	14.958	9.921
45 – 49	2	8.066	5.541
	α	786.784	2.679
	β	6351.489	7.142
	γ	18.411	20.344
	χ^2	23.225	8.077
	R^2	0.955	0.991
	SSQ	1121.424	231.125

Table 2. Empirical and fitted value of yearly peak flow of Bagmati river of Nepal by using Lomax and the SL distribution

Peak flow ($'000m^3/sec.$)	Observed number of years	Expected number of year (Lomax)	Expected number of year (SL)
1 – 2	2	2.270	1.144
2 – 3	8	9.659	8.989
3 – 4	11	6.519	8.539
4 – 5	3	4.509	5.952
5 – 6	5	3.186	3.870
6 – 7	3	2.295	2.484
7 – 8	2	1.682	1.605
8 – 9	1	1.252	1.051
9 - 10	0	0.945	0.700
10 - 11	0	0.722	0.474
11 - 12	0	0.559	0.326
12 - 13	0	0.437	0.228
13 - 14	0	0.345	0.162
14 - 15	0	0.274	0.117
15 - 16	1	0.220	0.085
	α	4.951	6.841
	β	13.996	11.395
	γ	1.815	1.668
	χ^2	3.516	0.959
	R^2	0.790	0.869
	SSQ	31.848	19.933

4.66, 5.08, 6.2, 3.14, 8.0, 3.38, 6.85, 5.58, 3.5, 5.4, 2.18, 2.97, 4.78, 3.38, 3.38, 3.54, 1.16 and 1.8. The empirical and fitted frequencies of yearly peak flows are presented in Table 2 and the fitted results are compared with the Lomax distribution.

Further, the third data set comprises 100 observations on waiting time (minutes) of the customer at a bank before received the service which has been taken from M. E. Ghitany et al. (2008) as: 0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11, 11, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19, 19.9, 20.6, 21.3, 21.4, 21.9, 23, 27, 31.6, 33.1, 38.5. For this data set, the empirical and fitted frequencies are presented in Table 3 and the results are compared with Lomax distribution.

The data analysis for three sets is performed through the solver option available in Microsoft Excel. Parameters of the both distributions Lomax and Skew-Lomax have been estimated by minimizing the sum of square of error (SSQ) between observed and expected frequencies. The Lomax distribution is not enough flexible to fit the data. So, for fitting of Lomax distribution, it has been controlled by choosing the suitable initial guess of parameters.

Table 3. Empirical and fitted number of customer's waiting time (minutes) at the bank before receiving the service

Waiting time of customer at bank (minutes)	Observed number of customers	Expected frequency (Lomax)	Expected frequency (SL)
0 – 4	17	16.451	16.890
4 – 8	33	31.659	32.676
8 –12	21	19.569	21.277
12 – 16	12	12.123	12.110
16 – 20	8	7.543	6.815
20 – 24	5	4.703	3.908
24 – 28	1	2.940	2.303
28 – 32	1	1.843	1.396
32 – 36	1	1.159	0.870
36 – 40	1	0.731	0.556
	α	78.196	6.889
	β	653.194	35.259
	γ	2.497	1.184
	χ^2	0.376	0.092
	R^2	0.991	0.995
	SSQ	9.036	4.872

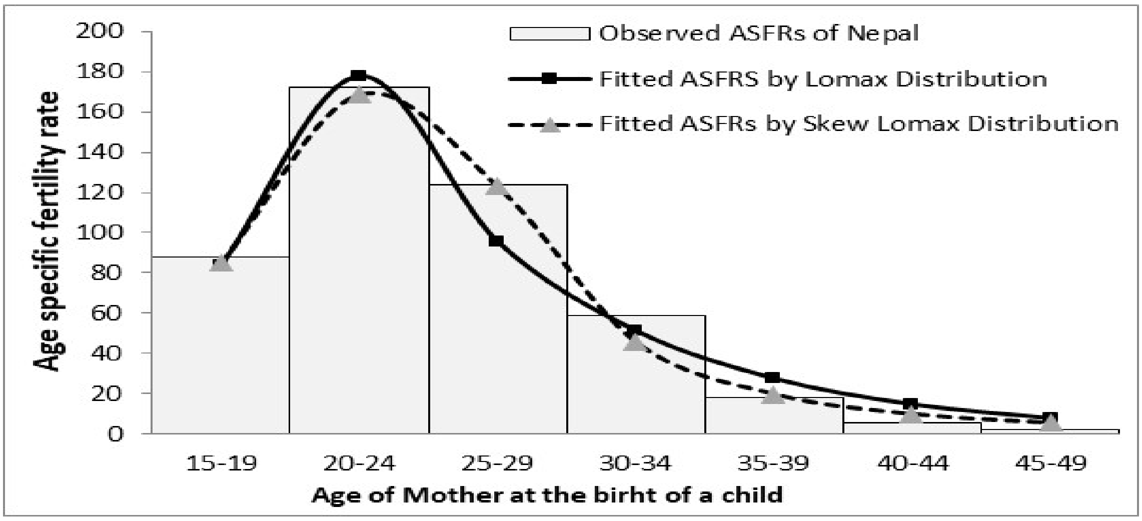

Figure 6. Observed and empirical age-specific fertility rate of Nepal-2016

The goodness of fit of the distribution has been performed by using Statistical tools: the coefficient of determination (R^2); the sum of the square of error (SSQ), and the Chi-square test statistics between the observed and the expected frequency of three real data sets. The distribution, which provides the least values of SSQ and chi-square as well as the maximum value of R^2, is considered the best fit the distributional pattern of data. From the analysis the Values of the SSQ and chi-square test statistics are found to be the least for each of the three data sets and the value of R^2 is the maximum for the proposed Skew Lomax distribution. So, the SL distribution model is claimed to more flexible and good to fit at least these data sets. The Observed and empirical values of the age of the mother at birth of a child for Nepali mothers (NDHS, 2017) has been presented in Figure 6. Similarly, the observed and fitted frequency of yearly peak flow of the Bagmati river of Nepal for 36 years has been illustrated in Figure 7. Furthermore, the observed and empirical frequency data of waiting time of customers in a bank before receiving the service for both Lomax and Sl distribution was depicted in Figure 8. The flexibility of Proposed Skew-Lomax distribution also has been observed in these figures of fitted results of three data sets.

8. Conclusion

A new three-parameter new probability distribution called Skew Lomax distribution has been formulated, which is the generalization of Lomax distribution as a sub-model of the distribution proposed by Azzalini and Capitanio (2013). The PDF and CDF of the proposed distribution have been formulated. The mathematical and statistical properties, including mean, variance, moments, reliability, hazard, cumulative hazard function, quantile function, random number generation, and entropy measure, have been studied.

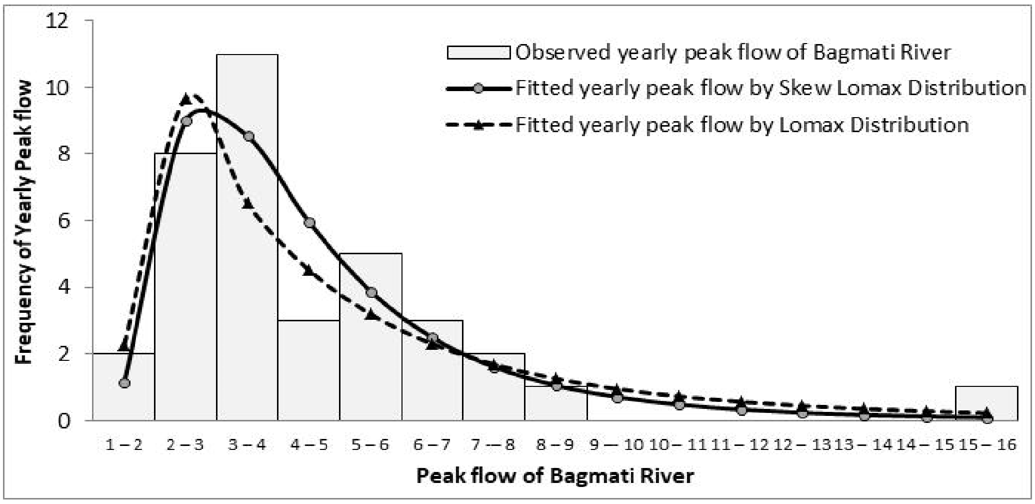

Figure 7. Observed and empirical data of yearly peak flow of Bagmati river

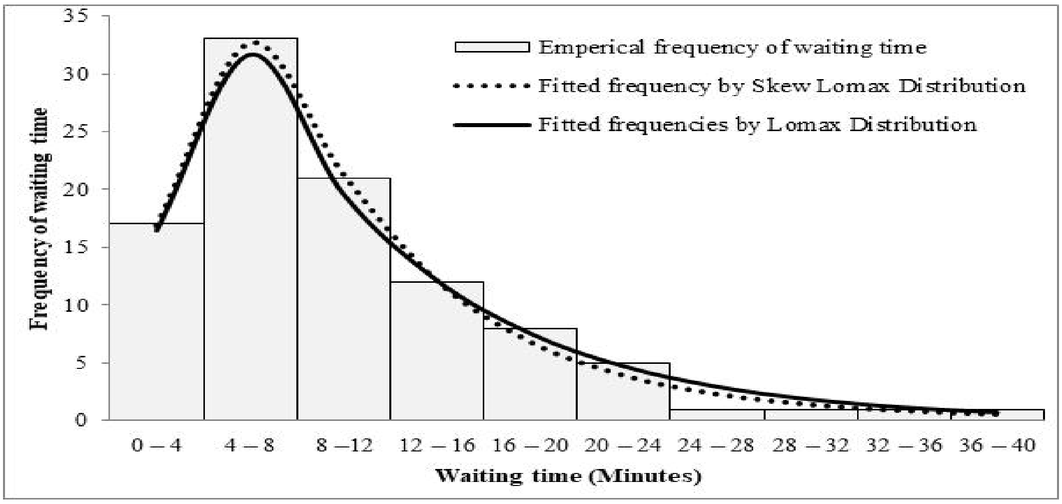

Figure 8. Observed and empirical data of waiting time of customers

The method of likelihood estimation and method of the moment to estimate the parameters associated have been formulated. To test the validity and suitability of the proposed model, three actual data sets of the age of the mother at the birth of a child; the waiting time of customers in a bank before receiving the service, and the yearly peak flow of the river have been applied. From the graphs and validity tools with minimum values of SSQ, Chi-square and maximum value of R^2 proposed model is a better fit at least these data than the Lomax distribution.

References

Alghamdi, A. S. (2018). *Study of generalized Lomax distribution and change point problem* (Doctoral dissertation). Bowling Green State University, Ohio, USA.

Ashour, S., & Eltehiwy, M. (2013). Transmuted exponentiated Lomax distribution. *Australian Journal of Basic and Applied Sciences*, *7*(7), 658–667.

Atkinson, M., & Harrison, A. J. (1978). *Distribution of personal wealth in Britain*. Cambridge.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, *12*(2), 171–178.

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

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

Cordeiro, G. M., Ortega, E. M., & Popović, B. V. (2015). The gamma-Lomax distribution. *Journal of Statistical Computation and Simulation*, *85*(2), 305–319.

DHM. (2020). *Yearly peak flow data of Bagmati river*. Department of Hydrology; Meteorology of Nepal.

El-Bassiouny, A., Abdo, N., & Shahen, H. (2015). Exponential Lomax distribution. *International Journal of Computer Applications*, *121*(13), 24–29.

El-Batal, I., & Kareem, A. (2014). Statistical properties of Kumaraswamy exponentiated Lomax distribution. *Journal of Modern Mathematics and Statistics*, *8*(1), 1–7.

Fatima, K., Jan, U., & Ahmad, S. (2018). Statistical properties of Rayleigh Lomax distribution with applications in survival analysis. *Journal of Data Science*, *16*(3), 531–548.

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), 80–83.

Gaire, A. K., Thapa, G. B., & 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*, 37–43.

Gaire, A. K., Thapa, G. B., KC, S., et al. (2022). Mathematical modeling of age-specific fertility rates of Nepali mothers. *Pakistan Journal of Statistics and Operation Research*, 417–426.

Ghitany, M., Al-Awadhi, F., & Alkhalfan, L. (2007). Marshall–Olkin extended Lomax distribution and its application to censored data. *Communications in Statistics—Theory and Methods*, *36*(10), 1855–1866.

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

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

Harris, C. M. (1968). The Pareto distribution as a queue service discipline. *Operations Research*, *16*(2), 307–313.

Hassan, A. S., & Al-Ghamdi, A. S. (2009). Optimum step stress accelerated life testing for Lomax distribution. *Journal of Applied Sciences Research*, *5*(12), 2153–2164.

Hassan, A. S., & Abd-Allah, M. (2018). Exponentiated Weibull-Lomax distribution: Properties and estimation. *Journal of Data Science*, *16*(2), 277–298.

Holland, O., Golaup, A., & Aghvami, A. (2006). Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration. *IEE Proceedings-communications*, *153*(5), 683–690.

Huang, W.-J., & Chen, Y.-H. (2007). Generalized skew-Cauchy distribution. *Statistics & probability letters*, *77*(11), 1137–1147.

Lemonte, A. J., & Cordeiro, G. M. (2013). An extended Lomax distribution. *Statistics*, *47*(4), 800–816.

Lomax, K. S. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American statistical association*, *49*(268), 847–852.

Mazzuco, S., & Scarpa, B. (2011). Fitting age-specific fertility rates by a skew-symmetric probability density function. *Padua: Department of Statistical Sciences, University of Padua (Working paper 10)*.

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

NDHS. (2017). *Nepal demographic and health survey 2016: Key indicators*. Ministry of Health, Nepal; New ERA; ICF.

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

Rady, E.-H. A., Hassanein, W., & Elhaddad, T. (2016). The power lomax distribution with an application to bladder cancer data. *SpringerPlus*, *5*, 1–22.

Rajab, M., Aleem, M., Nawaz, T., & Daniyal, M. (2013). On five parameter beta lomax distribution. *Journal of Statistics*, *20*(1), 102–118.

Rényi, A. (1961). On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, *4*, 547–562.

Shams, T. M. (2013). The Kumaraswamy-generalized Lomax distribution. *Middle-East Journal of Scientific Research*, *17*(5), 641–646.

Shaw, W. T., & Buckley, I. R. (2009). 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*.

Tahir, M. H., Cordeiro, G. M., Mansoor, M., & Zubair, M. (2015). The Weibull-Lomax distribution: Properties and applications. *Hacettepe Journal of Mathematics and Statistics*, *44*(2), 455–474.

Tahir, M. H., Hussain, M. A., Cordeiro, G. M., Hamedani, G., Mansoor, M., & Zubair, M. (2016). The Gumbel-Lomax distribution: Properties and applications. *Journal of Statistical Theory and Applications*, *15*(1), 61–79.

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