Earthline Journal of Mathematical Sciences
E-ISSN: 2581-8147
Volume 10, Number 2, 2022, Pages 385-407
https://doi.org/10.34198/ejms.10222.385407

Theory and Applications of the Transmuted Continuous Bernoulli Distribution

Christophe Chesneau[1], Festus C. Opone[2,*] and Ngozi O. Ubaka[3]

[1] Department of Mathematics, LMNO, University of Caen, 14032 Caen, France
e-mail: christophe.chesneau@unicaen.fr

[2] Department of Statistics, University of Benin, Benin City, Nigeria
e-mail: festus.opone@physci.uniben.edu

[3] Department of Statistics, Federal University of Oyo-Ekiti, Ekiti State, Nigeria
e-mail: obiaderi.oddafi@fuoye.edu.ng

Abstract

Modern applied statistics naturally give rise to the continuous Bernoulli distribution (data fitting, deep learning, computer vision, etc). On the mathematical side, it can be viewed as a one-parameter distribution corresponding to a special exponential distribution restricted to the unit interval. As a matter of fact, manageable extensions of this distribution have great potential in the same fields. In this study, we motivate a transmuted version of the continuous Bernoulli distribution with the goal of analyzing proportional data sets. The feature of the created transmuted continuous Bernoulli distribution is an additional parameter that realizes a linear tradeoff between the min and max of two continuous random variables with the continuous Bernoulli distribution. The standard study process is respected: we derive some mathematical properties of the proposed distribution and adopt the maximum likelihood estimation technique in estimating the unknown parameters involved. A Monte Carlo simulation exercise was conducted to examine and confirm the asymptotic behavior of the obtained estimates. In order to show the applicability of the proposed

Received: June 29, 2022; Accepted: August 2, 2022

2020 Mathematics Subject Classification: 62E15, 62J05.

Keywords and phrases: transmuted scheme, continuous Bernoulli distribution, moments, quantile.

*Corresponding author

distribution, three proportional data sets are analyzed and the results obtained are compared with competitive distributions. Empirical findings reveal that the transmuted continuous Bernoulli distribution promises more flexibility in fitting proportional data sets than its competitors.

1 Introduction

In order to motivate the findings of this study, we need a retrospective on the continuous Bernoulli distribution and the transmuted scheme, which are the subjects of the two coming subsections.

1.1 The continuous Bernoulli distribution

The basic definition of the continuous Bernoulli distribution is recalled below.

Definition 1. *The continuous Bernoulli distribution with parameter* $\theta \in (0,1)$*, also denoted as* $\mathcal{CB}(\theta)$*, is defined with the pdf given by*

$$f(x) = \begin{cases} C(\theta)\theta^x(1-\theta)^{1-x}, & x \in (0,1), \\ 0, & x \notin (0,1), \end{cases} \tag{1.1}$$

where $C(\theta)$ *is the following constant:*

$$C(\theta) = \begin{cases} 2, & \theta = \dfrac{1}{2}, \\ \dfrac{2\tanh^{-1}(1-2\theta)}{1-2\theta}, & \theta \in (0,1)/\left\{\dfrac{1}{2}\right\}, \end{cases} \tag{1.2}$$

where $\tanh^{-1}(x)$ *denotes the inverse hyperbolic tangent defined by* $\tanh^{-1}(x) = (1/2)\ln[(1+x)/(1-x)]$.

Therefore, the $\mathcal{CB}(\theta)$ distribution can be presented as a one-parameter continuous distribution with support of $(0,1)$; it thus belongs to the family of unit distributions. It is found useful in several branches of probability theory, statistics, and machine learning. In particular, it is used quite efficiently to simulate the pixel intensities of natural images in deep learning and computer vision, notably in the setting of variational autoencoders. As a result, it establishes a precise probabilistic equivalent for the widely used binary cross entropy loss, which is frequently utilized in nonlinear systems with values in $[0,1]$. We refer the reader to [13] and [9] for more information on this topic. Furthermore, the $\mathcal{CB}(\theta)$ distribution belongs to the family of truncated exponential distributions.

There are several mathematical results concerning the $\mathcal{CB}(\theta)$ distribution. The most manageable of them are recalled below.

First, the cdf of the $\mathcal{CB}(\theta)$ distribution is expressed as

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x, & \theta = \dfrac{1}{2} \text{ and } x \in (0,1), \\ \dfrac{\theta^x (1-\theta)^{1-x} + \theta - 1}{2\theta - 1}, & \theta \in (0,1)/\left\{\dfrac{1}{2}\right\} \text{ and } x \in (0,1), \\ 1, & x \geq 1. \end{cases}$$

The hazard rate function (hrf) is obtained as $h(x) = f(x)/[1 - F(x)]$, $x \in \mathbb{R}$, (the expression is omitted for the sake of place).

By inverting $F(x)$, the quantile function (qf) of the $\mathcal{CB}(\theta)$ distribution is indicated as

$$Q(u) = \begin{cases} u, & \theta = \dfrac{1}{2} \text{ and } u \in (0,1), \\ \dfrac{\ln[(2\theta - 1)u + 1 - \theta] - \ln(1-\theta)}{\ln(\theta) - \ln(1-\theta)}, & \theta \in (0,1)/\left\{\dfrac{1}{2}\right\} \text{ and } u \in (0,1). \end{cases}$$

As the main central parameter, the mean of a random variable X following the $\mathcal{CB}(\theta)$ distribution is given by

$$E(X) = \begin{cases} \dfrac{1}{2}, & \theta = \dfrac{1}{2}, \\ \dfrac{\theta}{2\theta - 1} + \dfrac{1}{2\tanh^{-1}(1-2\theta)}, & \theta \in (0,1)/\left\{\dfrac{1}{2}\right\}. \end{cases}$$

As the main dispersion parameter, the variance of a random variable X following the $\mathcal{CB}(\theta)$ distribution is expressed as

$$V(X) = \begin{cases} \frac{1}{12}, & \theta = \frac{1}{2}, \\ \frac{(1-\theta)\theta}{(1-2\theta)^2} + \frac{1}{[2\tanh^{-1}(1-2\theta)]^2}, & \theta \in (0,1)/\left\{\frac{1}{2}\right\}. \end{cases}$$

More theory is given in [13], but emphasis will be put on the moment measures above in this study.

1.2 Transmuted scheme

The transmuted generated family of distributions finds its origin in the article of [3]. It is defined below.

Definition 2. *The transmuted generated family of distributions, with parameter* $\lambda \in [-1, 1]$*, also denoted as* $\mathcal{TG}(\lambda)$*, is defined with the cdf given by*

$$F_G(x) = 1 - [1 - \lambda G(x)][1 - G(x)], \quad x \in \mathbb{R},$$

where $G(x)$ *denotes any cdf of a chosen (absolutely) continuous distribution that may depend on several parameters (in other words,* $G(x)$ *is the cdf of a baseline continuous distribution).*

The transmuted scheme was introduced by [18]. The aim of the $\mathcal{TG}(\lambda)$ family of distributions is to add more functionalities to a baseline distribution by using the parameter λ. The transmuted scheme is elaborated to realize a linear tradeoff between the baseline distribution, and the distributions of the min and max of two independent random variables also having this baseline distribution.

The pdf of the $\mathcal{TG}(\lambda)$ family of distributions is given as

$$f_G(x) = g(x)[1 + \lambda - 2\lambda G(x)], \quad x \in \mathbb{R},$$

where $g(x)$ refers to the pdf related to $G(x)$.

The qf of the $\mathcal{TG}(\lambda)$ family of distributions is given as

$$Q_G(u) = \begin{cases} G^{-1}(u), & \lambda = 0 \text{ and } u \in (0,1), \\ G^{-1}\left\{\dfrac{1+\lambda-\sqrt{(1+\lambda)^2-4\lambda u}}{2\lambda}\right\}, & \lambda \neq 0 \text{ and } u \in (0,1). \end{cases}$$

Clearly, both $f_G(x)$ and $Q_G(x)$ are possibly manageable, depending on the degree of complexity of the baseline distribution.

Many authors have introduced different generalizations of classical distributions based on the transmuted scheme. These generalizations include the transmuted Lindley distribution by [14], transmuted Weibull distribution by [1], transmuted modified inverse Weibull distribution by [6], transmuted generalized Lindley distribution by [7], transmuted two-parameter Lindley distribution by [10], etc. Recently, [17] considered the transmuted version of the Marshall-Olkin extended Topp-Leone distribution proposed by [16].

1.3 Motivation and organization

The goal of this study is to broaden the scope of the $\mathcal{CB}(\theta)$ distribution using a transmuted scheme; we consider the $\mathcal{CB}(\theta)$ distribution as a baseline distribution of the $\mathcal{TG}(\lambda)$ family. Thus, we intend to add some new functional perspectives of the $\mathcal{CB}(\theta)$ distribution without adding too much complexity. The benefits and details are described in detail in this study from both the theoretical and practical viewpoints, with the use of real-life data.

The remaining sections of the paper are organized as follows: Section 2 introduces the formulation of the transmuted continuous Bernoulli distribution and the derivation of its mathematical properties. The parameter estimation and a Monte Carlo simulation study are conducted in Section 3. Section 4 presents the real-life data fitting of the proposed distribution together with some competitive distributions. The concluding remark is presented in Section 5.

2 The Transmuted Continuous Bernoulli Distribution

As previously stated, we introduce the transmuted continuous Bernoulli distribution by considering the $\mathcal{CB}(\theta)$ distribution as a baseline of the $\mathcal{TG}(\lambda)$ family. The precise definition is as follows.

Definition 3. *The transmuted continuous Bernoulli distribution, with parameters* $\lambda \in [-1, 1]$ *and* $\theta \in (0, 1)$*, also denoted as* $\mathcal{TCB}(\lambda, \theta)$*, is defined with the cdf given by*

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - (1 - \lambda x)(1 - x), & \theta = \frac{1}{2} \text{ and } x \in (0, 1), \\ 1 - \left[1 - \lambda \frac{\theta^x (1-\theta)^{1-x} + \theta - 1}{2\theta - 1}\right]\left[1 - \frac{\theta^x (1-\theta)^{1-x} + \theta - 1}{2\theta - 1}\right], & \theta \in (0, 1)/\left\{\frac{1}{2}\right\} \text{ and } x \in (0, 1), \\ 1, & x \geq 1. \end{cases}$$

The pdf of the $\mathcal{TCB}(\lambda, \theta)$ distribution is given by

$$f(x) = \begin{cases} 1 + \lambda - 2\lambda x, & \theta = \frac{1}{2} \text{ and } x \in (0, 1), \\ C(\theta)\theta^x (1-\theta)^{1-x}\left[1 + \lambda - 2\lambda \frac{\theta^x (1-\theta)^{1-x} + \theta - 1}{2\theta - 1}\right], & \theta \in (0, 1)/\left\{\frac{1}{2}\right\} \text{ and } x \in (0, 1), \end{cases} \tag{2.1}$$

where $C(\theta)$ refers to (1.2).

The hrf is obtained as

$$h(x) = \begin{cases} \frac{1 + \lambda - 2\lambda x}{(1 - \lambda x)(1 - x)}, & \theta = \frac{1}{2} \text{ and } x \in (0, 1), \\ \frac{C(\theta)\theta^x (1-\theta)^{1-x}\left[1 + \lambda - 2\lambda \frac{\theta^x (1-\theta)^{1-x} + \theta - 1}{2\theta - 1}\right]}{\left[1 - \lambda \frac{\theta^x (1-\theta)^{1-x} + \theta - 1}{2\theta - 1}\right]\left[1 - \frac{\theta^x (1-\theta)^{1-x} + \theta - 1}{2\theta - 1}\right]}, & \theta \in (0, 1)/\left\{\frac{1}{2}\right\} \text{ and } x \in (0, 1). \end{cases}$$

Figure 1 presents the pdf and hrf plots of the $\mathcal{TCB}(\lambda, \theta)$ distribution at different values of the parameters.

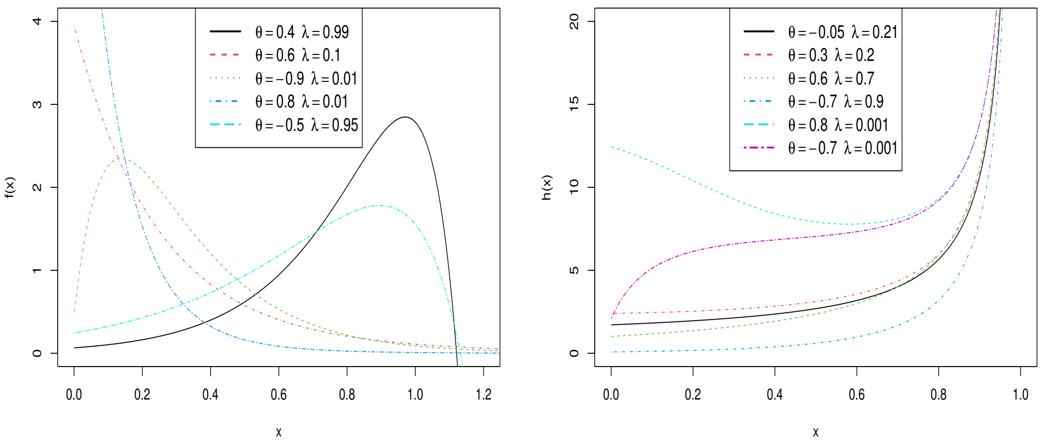

Figure 1: The pdf(left) and hrf(right) plots of the $\mathcal{TCB}(\lambda, \theta)$ distribution.

According to Figure 1, the $\mathcal{TCB}(\lambda, \theta)$ distribution can accommodate exponentially decreasing (reversed-J), left-skewed, and right-skewed unimodal shapes, whereas the hrf can exhibit a bathtub-shape or increase hazard property. These hazard properties show that the $\mathcal{TCB}(\lambda, \theta)$ distribution can feature in the analysis of increasing failure rate data sets.

The qf of the $\mathcal{TCB}(\lambda, \theta)$ distribution is expressed as

$$Q(u) = \begin{cases} u, & \lambda = 0, \theta = \frac{1}{2} \text{ and } u \in (0,1), \\ \dfrac{\ln[(2\theta-1)u + 1 - \theta] - \ln(1-\theta)}{\ln(\theta) - \ln(1-\theta)}, & \lambda = 0, \theta \in (0,1)/\left\{\frac{1}{2}\right\} \text{ and } u \in (0,1), \\ \dfrac{1+\lambda-\sqrt{(1+\lambda)^2 - 4\lambda u}}{2\lambda}, & \lambda \neq 0, \theta = \frac{1}{2} \text{ and } u \in (0,1), \\ \dfrac{\ln\left[(2\theta-1)\frac{1+\lambda-\sqrt{(1+\lambda)^2-4\lambda u}}{2\lambda} + 1 - \theta\right] - \ln(1-\theta)}{\ln(\theta) - \ln(1-\theta)}, & \lambda \neq 0, \theta \in (0,1)/\left\{\frac{1}{2}\right\} \text{ and } u \in (0,1), \end{cases} \tag{2.2}$$

where $C(\theta)$ is given as (1.2).

In order to provide a moment study, let us consider the following lemma, which can be of independent interest.

Lemma 1. *Let $a \in (0,1)$, $b > 0$ and c be a positive integer. Then we have*

$$\int_0^1 x^c a^{bx}(1-a)^{b(1-x)}dx = \frac{(1-a)^b}{[2b\tanh^{-1}(1-2a)]^{c+1}}\gamma(c+1, 2b\tanh^{-1}(1-2a)),$$

where $\gamma(x,u) = \int_0^u t^{x-1}e^{-t}dt$ is the lower incomplete gamma function. If we restrict the result to $a \in (0,1/2)$, we can assume that c is a positive real number, instead of a positive integer.

Proof. After a rewriting of the main term, and the change of variable $y = 2b\tanh^{-1}(1-2a)x$, we obtain

$$\begin{aligned}\int_0^1 x^c a^{bx}(1-a)^{b(1-x)}dx &= \int_0^1 x^c e^{bx\ln(a)+b(1-x)\ln(1-a)}dx \\ &= (1-a)^b \int_0^1 x^c e^{-x[2b\tanh^{-1}(1-2a)]}dx \\ &= \frac{(1-a)^b}{[2b\tanh^{-1}(1-2a)]^{c+1}} \int_0^{2b\tanh^{-1}(1-2a)} y^c e^{-y}dy \\ &= \frac{(1-a)^b}{[2b\tanh^{-1}(1-2a)]^{c+1}}\gamma(c+1, 2b\tanh^{-1}(1-2a)).\end{aligned}$$

This ends the proof. □

The next proposition gives the exact expression of the moments related to the $\mathcal{TCB}(\lambda,\theta)$ distribution.

Proposition 1. *Let X be a random variable with the $\mathcal{TCB}(\lambda,\theta)$ distribution and m be an integer. Then the m-th moment of X is given by*

$$E(X^m) = \begin{cases} \dfrac{m(1-\lambda)+2}{(m+1)(m+2)}, & \theta = \dfrac{1}{2}, \\ \dfrac{2\lambda}{(1-2\theta)^2}\dfrac{(1-\theta)^2}{[2\tanh^{-1}(1-2\theta)]^m}\left\{\left[\dfrac{(1+\lambda)(1-2\theta)}{2\lambda(1-\theta)} - 1\right]\gamma(m+1, 2\tanh^{-1}(1-2\theta)) \right. & \\ \left. + \dfrac{1}{2^{m+1}}\gamma(m+1, 4\tanh^{-1}(1-2\theta))\right\}, & \theta \in (0,1)/\left\{\dfrac{1}{2}\right\}. \end{cases}$$

Proof. For the case $\theta = 1/2$, we have

$$\begin{aligned}E(X^m) &= \int_{-\infty}^{+\infty} x^m f(x)dx = \int_0^1 x^m(1+\lambda-2\lambda x)dx = \frac{1+\lambda}{m+1} - 2\lambda\frac{1}{m+2} \\ &= \frac{m(1-\lambda)+2}{(m+1)(m+2)}.\end{aligned}$$

For the case $\theta \in (0,1)/\{1/2\}$, we have

$$\begin{aligned} E(X^m) &= \int_{-\infty}^{+\infty} x^m f(x)dx \\ &= C(\theta)\int_0^1 x^m\theta^x(1-\theta)^{1-x}\left[1+\lambda-2\lambda\frac{\theta^x(1-\theta)^{1-x}+\theta-1}{2\theta-1}\right]dx \\ &= C(\theta)\left[1+\lambda-2\lambda\frac{\theta-1}{2\theta-1}\right]\int_0^1 x^m\theta^x(1-\theta)^{1-x}dx \\ &- C(\theta)\frac{2\lambda}{2\theta-1}\int_0^1 x^m\theta^{2x}(1-\theta)^{2(1-x)}dx. \end{aligned}$$

By applying Lemma 1 with adequate configurations, we obtain

$$\begin{aligned} E(X^m) &= C(\theta)\left[1+\lambda-2\lambda\frac{\theta-1}{2\theta-1}\right]\frac{1-\theta}{[2\tanh^{-1}(1-2\theta)]^{m+1}}\gamma(m+1, 2\tanh^{-1}(1-2\theta)) \\ &- C(\theta)\frac{2\lambda}{2\theta-1}\frac{(1-\theta)^2}{[4\tanh^{-1}(1-2\theta)]^{m+1}}\gamma(m+1, 4\tanh^{-1}(1-2\theta)) \\ &= \frac{2\lambda}{(1-2\theta)^2}\frac{(1-\theta)^2}{[2\tanh^{-1}(1-2\theta)]^m}\left\{\left[\frac{(1+\lambda)(1-2\theta)}{2\lambda(1-\theta)}-1\right]\gamma(m+1, 2\tanh^{-1}(1-2\theta))\right. \\ &\left.+\frac{1}{2^{m+1}}\gamma(m+1, 4\tanh^{-1}(1-2\theta))\right\}. \end{aligned}$$

This ends the proof. □

Based on the moments established in Proposition 1, we can easily derive the mean of a random variable X following the $\mathcal{TCB}(\lambda, \theta)$ distribution; it is given as

$$\mu = E(X) = \begin{cases} \dfrac{3-\lambda}{6}, & \theta = \dfrac{1}{2}, \\ \dfrac{2\lambda}{(1-2\theta)^2}\dfrac{(1-\theta)^2}{2\tanh^{-1}(1-2\theta)}\left\{\left[\dfrac{(1+\lambda)(1-2\theta)}{2\lambda(1-\theta)}-1\right]\gamma(2, 2\tanh^{-1}(1-2\theta))\right. & \\ \left.+\dfrac{1}{4}\gamma(2, 4\tanh^{-1}(1-2\theta))\right\}, & \theta \in (0,1)/\left\{\dfrac{1}{2}\right\}. \end{cases}$$

It is worth mentioning that $\gamma(2,u) = 1-(1+u)e^{-u}$, which can be used in the expression above. The variance can be expressed in a similar way through the standard formula $V(X) = E(X^2) - \mu^2$. The m-th central moment of X is given by

$$E[(X-\mu)^m] = \sum_{k=0}^{m}\binom{m}{k}E(X^k)(-1)^{m-k}\mu^{m-k}.$$

We can express it by using Proposition 1. Hence, the moment skewness coefficient is given by $S = E[(X-\mu)^3]/V(X)^{3/2}$ and the moment kurtosis coefficient is given by $K = E[(X-\mu)^4]/V(X)^2$. Similarly, we can express them via Proposition 1.

3 Parameter Estimation and Simulation Study

We now investigate the statistical configuration of the $\mathcal{TCB}(\lambda, \theta)$ distribution, assuming that the parameters λ and θ are unknown. We thus planned to estimate them, for data fitting purposes mainly.

3.1 Estimation

In this study, the maximum likelihood method is considered. The interest of this method is to provide very efficient estimates, called the maximum likelihood estimates (MLEs), that are able to approach the true values of the parameters under some concrete circumstances (see [4]). Let $x_1, x_2, \ldots, x_n$ be n independent observations from a random variable X with the $\mathcal{TCB}(\theta, \lambda)$ distribution. The values of this sample represent possible data on the unit intervals (proportion, percentage, etc.). The standard likelihood function associated with the pdf in (2.1) at $x = (x_1, \ldots, x_n)$ is defined as

$$L(x) = \prod_{i=1}^{n} f(x_i),$$
$$= \begin{cases} \displaystyle\prod_{i=1}^{n} (1 + \lambda - 2x_i), & \theta = \dfrac{1}{2}, \\ C^n(\theta)\theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i} \displaystyle\prod_{i=1}^{n} \left[1 + \lambda - \dfrac{2\lambda\theta^{x_i}(1-\theta)^{1-x_i} + \theta - 1}{2\theta - 1}\right], & \theta \in (0,1)/\left\{\dfrac{1}{2}\right\}. \end{cases} \tag{3.1}$$

Taking the logarithm of (3.1), we obtain the log-likelihood function as

$$\ell(x)=\begin{cases}\sum_{i=1}^{n}\ln\left(1+\lambda-2x_i\right), & \theta=\frac{1}{2},\\ n\ln C(\theta)+\ln(\theta)\sum_{i=1}^{n}x_i+\ln\left(1-\theta\right)\left(n-\sum_{i=1}^{n}x_i\right) & \\ +\sum_{i=1}^{n}\ln\left[1+\lambda-\frac{2\lambda\theta^{x_i}(1-\theta)^{1-x_i}+\theta-1}{2\theta-1}\right], & \theta\in(0,1)/\left\{\frac{1}{2}\right\}.\end{cases} \tag{3.2}$$

The MLEs of θ and λ, say $\hat{\theta}$ and $\hat{\lambda}$, are obtained by maximizing (3.2) with respect to θ and λ. They can be obtained through the first partial derivatives of $\ell(x)$ according to the parameters, derived as

$$\frac{\partial\ell(x)}{\partial\theta}=\frac{n\left(1-2\theta\right)\left[\frac{2\tanh^{-1}(1-2\theta)}{(1-2\theta)^2}-\frac{2}{(1-2\theta)\left[1-(1-2\theta)^2\right]}\right]}{\tanh^{-1}\left(1-2\theta\right)}+\frac{n-\sum_{i=1}^{n}x_i}{1-\theta}+\frac{\sum_{i=1}^{n}x_i}{\theta}$$
$$+\sum_{i=1}^{n}\frac{2\left(2\lambda\theta^{x_i}(1-\theta)^{1-x_i}+\theta-1\right)-(2\theta-1)\left[1+2\lambda x_i\theta^{x_i-1}(1-\theta)^{1-x_i}+2\lambda\theta^{x_i}\left((1-\theta)^{-x_i}(x_i-1)\right)\right]}{(2\theta-1)\left[(1+\lambda)(2\theta-1)-2\lambda\theta^{x_i}(1-\theta)^{1-x_i}+\theta-1\right]},$$

and

$$\frac{\partial\ell(x)}{\partial\lambda}=\sum_{i=1}^{n}\frac{(2\theta-1)-2\left(1-\theta\right)^{1-x_i}\theta^{x_i}}{(2\theta-1)\left(1+\lambda\right)-2\lambda\left(1-\theta\right)^{1-x_i}\theta^{x_i}+\theta-1}.$$

As a result, the MLEs are obtained by solving the following system of non-linear equations: $\frac{\partial\ell(x)}{\partial\theta}=0$ and $\frac{\partial\ell(x)}{\partial\lambda}=0$. The solutions to these non-linear equations can be obtained using the "**bbmle**" package in the R statistical software program. It is well-known that the random versions of these MLEs are asymptotically unbiased and normal. In practice, this implies that, for n large enough, the true values of the parameters are well approached by the MLEs, in both the punctual and (normal) confidence interval senses. For more information on these properties, we refer to [4].

In addition, for data fitting purposes, it is important to present the parametric plug-in method. Based on the MLEs, we can estimate all the underlying functions of the $\mathcal{TCB}(\theta,\lambda)$ distribution. As the main examples, $\widehat{F}(x)=F(x;\widehat{\theta},\widehat{\lambda})$ gives an estimate of the cdf $F(x)=F(x;\theta,\lambda)$, and $\widehat{f}(x)=f(x;\widehat{\theta},\widehat{\lambda})$ gives an estimate of the pdf $f(x)=f(x;\theta,\lambda)$.

3.2 Monte Carlo simulation study

In this subsection, we investigate the performance and asymptotic behavior of the MLEs through a Monte Carlo simulation study. To accomplish this, we use (2.2) to generate random samples of observations from a random variable with the $\mathcal{TCB}(\theta, \lambda)$ distribution. At the following fixed parameter values: $(\theta = 0.6, \lambda = 0.2), (\theta = 0.8, \lambda = 0.1)$ and $(\theta = 0.9, \lambda = 0.5)$, the simulation is repeated 1000 times for different sample sizes $n = (50, 100, 150, 200, 500)$. For $\phi \in \{\theta, \lambda\}$, the asymptotic behavior of the MLEs is investigated using the following quantities:

1. Average bias (AB) defined as $\text{AB} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\phi}_i - \phi)$,

2. Mean square error (MSE) specified by $\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\phi}_i - \phi)^2$,

3. Coverage probability (CP) of the 95% confidence interval (CI) for ϕ defined by

$$\text{CP} = \frac{1}{N}\sum_{i=1}^{N} I\left(\hat{\phi}_i - u_*\sqrt{var(\hat{\phi}_i)} < \phi < \hat{\phi}_i + u_*\sqrt{var(\hat{\phi}_i)}\right),$$

4. Average width (AW) of the 95% CI for ϕ given by

$$\text{AW} = \frac{2u_*}{N}\sum_{i=1}^{N}\sqrt{var(\hat{\phi}_i)},$$

where $I(.)$ is the indicator function, $var(\hat{\phi}_i)$ is the empirical variance of $\hat{\phi}_i$ and $z_* = 1.959964$.

Table 1 presents the simulation results for the AB, MSE, CP and AW for the MLEs of θ and λ.

Table 1: Simulation results for the MLEs in the context of the $\mathcal{TCB}(\theta, \lambda)$ distribution.

		AB		MSE		CP		AW	
Parameters	n	θ	λ	θ	λ	θ	λ	θ	λ
	50	0.0449	0.1576	0.0290	0.1706	0.808	0.844	1.0536	2.3168
$\theta = 0.6$	100	0.0211	0.0839	0.0189	0.1005	0.868	0.862	0.9359	1.9821
$\lambda = 0.2$	150	0.0196	0.0645	0.0159	0.0806	0.852	0.886	0.8415	1.7705
	200	0.0054	0.0318	0.0128	0.0573	0.902	0.916	0.8409	1.7576
	500	0.0027	0.0164	0.0095	0.0430	0.868	0.896	0.6682	1.3839
	50	0.0103	0.1498	0.0131	0.1773	0.790	0.850	0.07801	2.2742
$\theta = 0.8$	100	0.0028	0.0737	0.0096	0.1033	0.820	0.900	0.6303	1.8034
$\lambda = 0.1$	150	-0.0045	0.0356	0.0079	0.0765	0.864	0.896	0.5995	1.6852
	200	-0.0082	0.0166	0.0071	0.0686	0.874	0.916	0.5852	1.6176
	500	-0.0100	0.0002	0.0051	0.0438	0.886	0.916	0.4792	1.2943
	50	-0.0329	-0.0154	0.0109	0.1755	0.828	0.862	0.5252	1.8157
$\theta = 0.9$	100	-0.0209	-0.0224	0.0070	0.1010	0.858	0.908	0.3676	1.3452
$\lambda = 0.5$	150	-0.0201	-0.0316	0.0053	0.0744	0.912	0.920	0.2949	1.0974
	200	-0.0179	-0.0444	0.0039	0.0581	0.912	0.932	0.2510	0.9676
	500	-0.0057	-0.0092	0.0011	0.0185	0.938	0.958	0.1169	0.5284

From Table 1, we note the following observations:

- The ABs decrease as n increases,
- both MLEs can be positively as well as negatively biased,
- the MSEs of both MLEs decrease as n increases,
- the CPs of both parameter estimates approach 0.95 as n increases,
- the AWs of both MLEs decrease as n increases.

These results reveal the consistency property of the MLEs.

4 Data Fitting

4.1 Framework

To illustrate the potentiality of the $\mathcal{TCB}(\theta, \lambda)$ distribution in real-life data analysis, we fitted the $\mathcal{TCB}(\theta, \lambda)$ distribution together with some well-known unit-distributions in literature to three real data sets. Their fits are obtained by considering the maximum likelihood method. They are compared based on some statistical model selection tools. The pdfs of the competitor distributions are defined as follows:

1. Unit-Burr XII distribution (UBXIID) by [12]:
   $f(x; \alpha, \beta) = \alpha\beta x^{-1}(-\ln x)^{\beta-1}\left(1 + (-\ln x)^{\beta}\right)^{-(\alpha+1)}$;
2. Unit-Burr III distribution (UBIIID) by [15]:
   $f(x; \lambda, \beta) = \lambda\beta x^{-2}\left(x^{-1} - 1\right)^{\beta-1}\left(1 + (x^{-1} - 1)^{\beta}\right)^{-(\lambda+1)}$;
3. Continuous Bernoulli distribution reported in [19] (see (1.1)).

Data set 1: The first set of data consists of the time to infection of kidney dialysis patients during the months reported in [11]. The data set is: 2.5, 2.5, 3.5, 3.5, 3.5, 4.5, 5.5, 6.5, 6.5, 7.5, 7.5, 7.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 12.5, 13.5, 14.5, 14.5, 21.5, 21.5, 22.5, 22.5, 25.5, 27.5. Employing a similar idea in [2], we transform the data set to lie within the interval [0,1] by dividing its values by an arbitrary number slightly higher than the maximum values in the data set. Having the highest number in the data set as 27.5, we divide by 30, yielding the following unit data set: 0.08333333, 0.08333333, 0.11666667, 0.11666667, 0.11666667, 0.15000000, 0.18333333, 0.21666667, 0.21666667, 0.25000000, 0.2500000, 0.2500000, 0.2500000, 0.28333333, 0.31666667, 0.35000000, 0.3833333, 0.41666667, 0.41666667, 0.4500000, 0.48333333, 0.4833333, 0.71666667, 0.71666667, 0.75000000, 0.75000000, 0.85000000, 0.91666667.

Data set 2: The second data set is the records of 72 exceedances of flood peaks (in m^3/s) of the Wheaton river near Carcross in the Yukon Territory, Canada for the years 1958-1984. The data set was first used by [5] to illustrate the potential of the generalized Pareto distribution. The same transformation technique is performed on the data set to obtain a unit data set as: 0.026153846 0.033846154 0.221538462 0.016923077 0.006153846, 0.316923077, 0.081538462, 0.010769231, 0.029230769, 0.200000000, 0.184615385, 0.143076923, 0.021538462, 0.287692308, 0.130769231, 0.392307692, 0.178461538, 0.216923077, 0.340000000, 0.016923077, 0.038461538, 0.221538462, 0.026153846, 0.578461538, 0.009230769, 0.033846154, 0.600000000, 0.004615385, 0.230769231, 0.169230769, 0.112307692, 0.352307692, 0.026153846, 0.001538462, 0.016923077, 0.009230769, 0.138461538, 0.026153846, 0.107692308, 0.309230769, 0.006153846, 0.043076923, 0.216923077, 0.152307692, 0.160000000, 0.164615385, 0.461538462, 0.055384615, 0.086153846, 0.473846154, 0.204615385, 0.064615385, 0.392307692, 0.052307692, 0.183076923, 0.330769231, 0.424615385, 0.560000000, 0.041538462, 0.984615385, 0.023076923, 0.038461538, 0.421538462, 0.015384615, 0.416923077, 0.310769231, 0.258461538, 0.081538462, 0.149230769, 0.423076923, 0.038461538, 0.415384615.

Data set 3: The third data set represents the waiting times (in minutes) before service of 100 bank customers. [8] used the data set to show that the Lindley distribution is a better model than the exponential distribution in real life data fitting. Again, after the transformation, we obtain a unit data set as: 0.0200, 0.0200, 0.0325, 0.0375, 0.0450, 0.0475, 0.0475, 0.0525, 0.0650, 0.0675, 0.0725, 0.0775, 0.0800, 0.0825, 0.0875, 0.0900, 0.1000, 0.1025, 0.1050, 0.1050, 0.1075, 0.1075, 0.1100, 0.1100, 0.1150, 0.1175, 0.1175, 0.1200, 0.1225, 0.1225, 0.1250, 0.1325, 0.1375, 0.1425, 0.1425, 0.1525, 0.1550, 0.1550, 0.1550, 0.1575, 0.1675, 0.1725, 0.1775, 0.1775, 0.1775, 0.1775, 0.1850, 0.1900, 0.1925, 0.2000, 0.2050, 0.2150, 0.2150, 0.2150, 0.2200, 0.2200, 0.2225, 0.2225, 0.2375, 0.2400, 0.2425, 0.2450, 0.2675, 0.2725, 0.2750, 0.2750, 0.2775, 0.2800, 0.2800, 0.2875, 0.2975, 0.3100, 0.3125, 0.3225, 0.3250, 0.3275, 0.3325, 0.3400, 0.3425, 0.3475, 0.3525, 0.3850, 0.3850, 0.4325, 0.4325, 0.4525, 0.4550, 0.4600, 0.4725, 0.4750, 0.4975, 0.5150, 0.5325, 0.5350, 0.5475, 0.5750, 0.6750, 0.7900, 0.8275, 0.9625.

4.2 Statistical results

The MLEs, estimated Log-likelihood ($LogL$), Akaike information criterion (AIC), Kolmogorov-Smirnov ($K-S$) and Cramér-von-Mises (W^*) test statistics of the distributions for the three data sets under study are presented in Tables 2 - 4. These model selection criteria will be used to ascertain the model that best fits the data sets. Hereafter, we shall denote the $\mathcal{TCB}(\theta, \lambda)$ distribution as TCBD for convenience's sake.

Table 2: Summary statistics for data set 1.

Distributions	*MLEs*	*LogL*	*AIC*	$K-S$ ($p-value$)	W^* ($p-value$)
TCBD	$\theta = 0.9630$	3.0344	-2.0688	0.1280	0.0954
	$\lambda = 0.6088$			(0.7479)	(0.6109)
UBXIID	$\alpha = 1.1075$	1.3736	1.2526	0.1663	0.2023
	$\beta = 2.0977$			(0.4204)	(0.2638)
UBIIID	$\lambda = 0.6266$	2.8522	-1.7044	0.1592	0.1746
	$\beta = 1.5534$			(0.4766)	(0.3237)
CBD	$\theta = 0.1783$	2.5733	-1.1466	0.1526	0.1136
				(0.5314)	(0.5249)

Table 3: Summary statistics for data set 2.

Distributions	*MLEs*	*LogL*	*AIC*	$K-S$ ($p-value$)	W^* ($p-value$)
TCBD	$\theta = 0.2657$	49.1144	-94.2288	0.1331	0.2029
	$\lambda = 0.0118$			(0.1557)	(0.2621)
UBXIID	$\alpha = 0.4235$	34.8783	-65.7567	0.1593	0.5366
	$\beta = 2.9667$			(0.0516)	(0.0319)
UBIIID	$\lambda = 0.2167$	43.0243	-82.0486	0.1361	0.3801
	$\beta = 2.0738$			(0.1386)	(0.0813)
CBD	$\theta = 0.0056$	48.8068	-93.6136	0.1464	0.2351
				(0.0910)	(0.2087)

Table 4: Summary statistics for data set 3.

Distributions	*MLEs*	*LogL*	*AIC*	$K-S$ ($p-value$)	W^* ($p-value$)
TCBD	$\theta = 0.9491$	43.0725	-82.1450	0.1360	0.3951
	$\lambda = 0.2212$			(0.0594)	(0.0742)
UBXIID	$\alpha = 0.5417$	31.5278	-59.0556	0.1649	0.8902
	$\beta = 3.4227$			(0.0087)	(0.0044)
UBIIID	$\lambda = 0.2897$	33.3206	-62.6481	0.1907	1.1323
	$\beta = 2.3714$			(0.0013)	(0.0012)
CBD	$\theta = 0.0251$	41.9787	-81.9574	0.1544	0.5365
				(0.0169)	(0.0340)

4.3 Discussion of the results

The applicability of a statistical model in real-life data analysis is judged based on how small the goodness-of-fit test statistics values are. According to Tables 2 - 4, the $\mathcal{TCB}(\theta, \lambda)$ distribution has the highest estimated log-likelihood value and the lowest value in terms of AIC, $K-S$, and W^* test statistics with the highest corresponding $p-value$, making the model the most appropriate model to be considered in analyzing the data sets under study. It is the only distribution with p-values superior to 0.05 in data set 3, validating its distributional adequacy. We further accessed the performance of the $\mathcal{TCB}(\theta, \lambda)$ distribution over the competitor distributions based on a graphical representation. Figures 2 - 4 display the probability-probability (P-P) plots of the distributions for the three data sets, respectively.

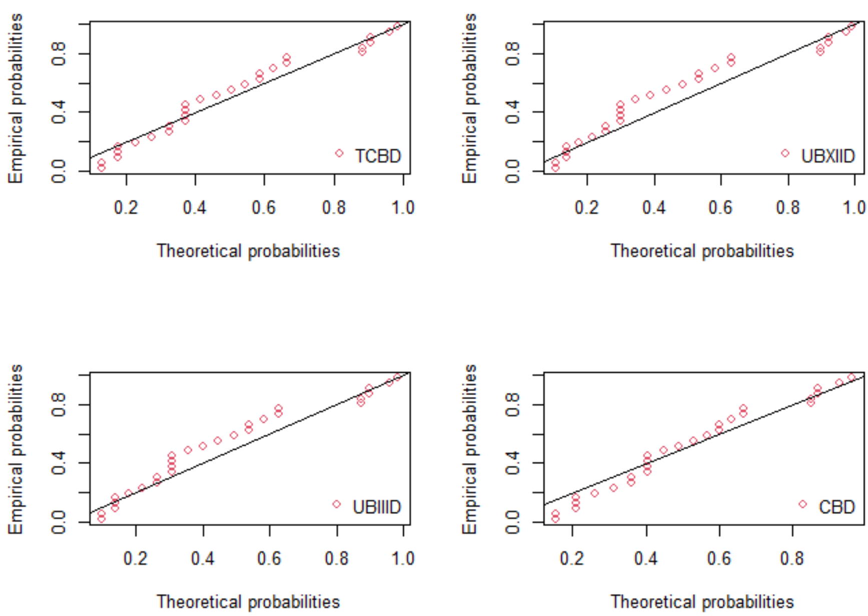

Figure 2: P-P plots of the distributions for data set 1.

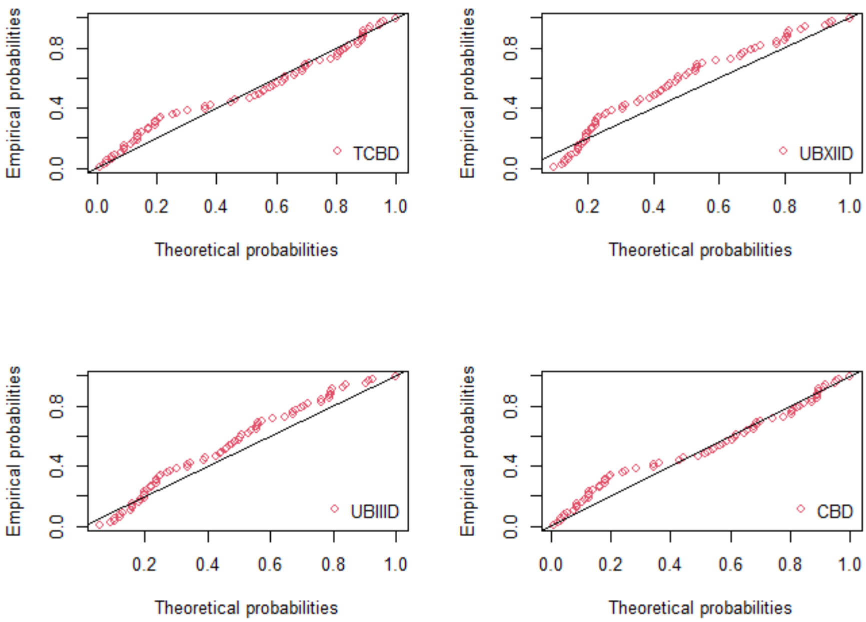

Figure 3: P-P plots of the distributions for data set 2.

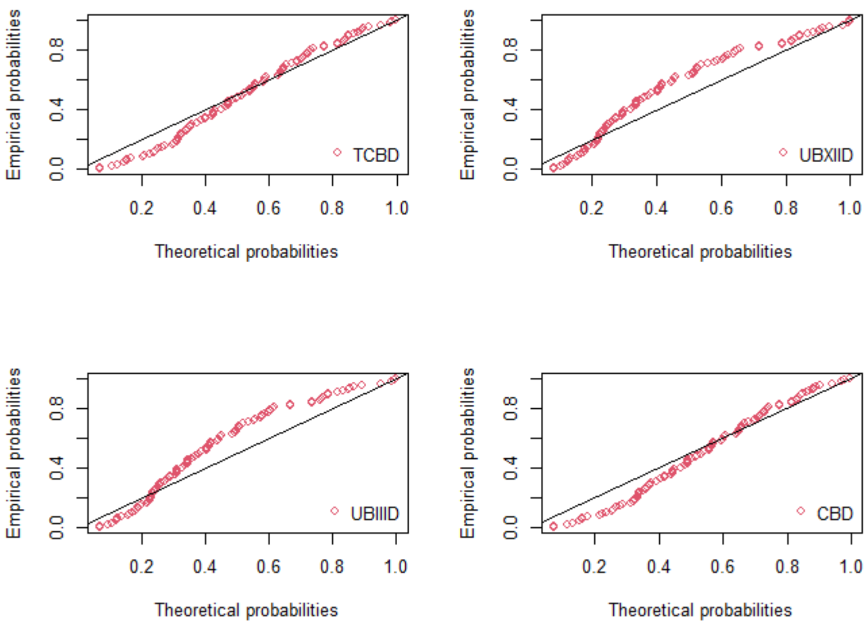

Figure 4: P-P plots of the distributions for data set 3.

According to these figures, the fit of the $\mathcal{TCB}(\theta, \lambda)$ distribution matches the fit of the data sets better than the other distributions under consideration. Thus, validating the claim that the $\mathcal{TCB}(\theta, \lambda)$ distribution is the best model to fit the three data sets under study.

5 Concluding Remark

In this paper, we have introduced the two-parameter transmuted continuous Bernoulli distribution applicable for fitting proportional data sets. It was represented by $\mathcal{TCB}(\theta, \lambda)$, with θ and λ as distribution parameters. Plots of the probability density function reveal that the $\mathcal{TCB}(\theta, \lambda)$ distribution accommodates exponentially decreasing (reversed-J), left-skewed and right-skewed unimodal shapes. On the other hand, the hazard rate function can exhibit bathtub-shapes or increase in shapes. These unique features are essential in analyzing increasing failure rate data sets. A Monte Carlo simulation study has been conducted to illustrate the performance of the maximum likelihood estimates of the unknown parameters of the $\mathcal{TCB}(\theta, \lambda)$ distribution. Finally, the potentiality of the $\mathcal{TCB}(\theta, \lambda)$ distribution in real-life data fittings was examined using three proportional data sets, and the results obtained reveal that the proposed distribution provides consistently better fits than the competitive distributions. The perspectives for probability and statistics applications of the $\mathcal{TCB}(\theta, \lambda)$ distribution, such as deep data fitting, regression modeling and machine learning applications, are numerous, and this paper provides the first steps in that direction.

References

[1] G. R. Aryal and C. P. Tsokos, Transmuted Weibull distribution: A generalization of the Weibull probability distribution, *European Journal of Pure and Applied Mathematics* 4(2) (2011), 89-102.

[2] R. A. R. Bantan, C. Chesneau, F. Jamal, M. Elgarhy, M. H. Tahir, A. Aqib, M. Zubair and S. Anam, Some new facts about the unit-Rayleigh distribution with applications, *Mathematics* 8(11) (2020), 1954. `https://doi.org/10.3390/math8111954`

[3] M. Bourguignon, I. Ghosh and G. M. Cordeiro, General results for the transmuted family of distributions and new models, *Journal of Probability and Statistics* 2016 (2016), Article ID 7208425, 12 pp. `https://doi.org/10.1155/2016/7208425`

[4] G. Casella and R. L. Berger, *Statistical Inference*, Brooks/Cole Publishing Company: Bel Air, CA, USA, 1990.

[5] V. Choulakian, and M. A. Stephens, Goodness-of-fit for the generalized Pareto distribution, *Technometrics* 43(4) (2001), 478-484. `https://doi.org/10.1198/00401700152672573`

[6] I. Elbatal, Transmuted modified inverse Weibull distribution, *International Journal of Mathematical Archive* 4(8) (2013), 117-129. https://doi.org/10.12988/ams.2014.44267

[7] M. Elgarhy, M. Rashed and A. W. Shawki, Transmuted generalized Lindley distribution, *International Journal of Mathematics Trends and Technology* 29(2) (2016), 145-154. https://doi.org/10.14445/22315373/IJMTT-V29P520

[8] M. Ghitany, B. Atieh and S. Nadarajah, Lindley distribution and its applications, *Mathematics and Computers in Simulation* 78 (2008), 493-506. https://doi.org/10.1016/j.matcom.2007.06.007

[9] E. Gordon-Rodriguez, G. Loaiza-Ganem and J. P. Cunningham, The continuous categorical: a novel simplex-valued exponential family, 36th International Conference on Machine Learning, ICML 2020, International Machine Learning Society (IMLS), 2020.

[10] S. A. Kemaloglu and M. Yilmaz, Transmuted two-parameter Lindley distribution, *Communications in Statistics - Theory and Methods* 46(23) (2017), 11866-11879. https://doi.org/10.1080/03610926.2017.1285933

[11] J. P. Klein and M. L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*, Springer, Berlin, 2006.

[12] M. Korkmaz and C. Chesneau, On the unit Burr-XII distribution with the quantile regression modeling and applications, *Computational and Applied Mathematics* 40(1) (2021), 1-26. https://doi.org/10.1007/s40314-021-01418-5

[13] G. Loaiza-Ganem and J. P. Cunningham, The continuous Bernoulli: fixing a pervasive error in variational autoencoders, *Advances in Neural Information Processing Systems* (2019), 13266-13276.

[14] F. Merovci, Transmuted Lindley distribution, *International Journal of Open Problems in Computer Science and Mathematics* 6(2) (2013), 63-72. https://doi.org/10.12816/0006170

[15] K. Modi and V. Gill, Unit Burr-III distribution with application, *Journal of Statistics and Management Systems* 23 (2020), 579-592. https://doi.org/10.1080/09720510.2019.1646503

[16] F. C. Opone, and B. N. Iwerumor, A new Marshall-Olkin extended family of distributions with bounded support, *Gazi University Journal of Science* 34(3) (2021), 899-914. https://doi.org/10.35378/gujs.721816

[17] F. C. Opone and J. E. Osemwenkhae, The transmuted Marshall-Olkin extended Topp-Leone distribution, *Earthline Journal of Mathematical Sciences* 9(2) (2022), 179-199. https://doi.org/10.34198/ejms.9222.179199

[18] W. T. Shaw and I. R. C. Buckley, The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map, UCL Discovery Repository, 2007. http://discovery.ucl.ac.uk/id/eprint/643923

[19] K. Wang and M. Lee, Continuous Bernoulli distribution: simulator and test statistic, 2020. http://dx.doi.org/10.13140/RG.2.2.28869.27365