stats

MDPI

Article

A Novel Generalization of Zero-Truncated Binomial Distribution by Lagrangian Approach with Applications for the COVID-19 Pandemic

Muhammed Rasheed Irshad [1], Christophe Chesneau [2,*], Damodaran Santhamani Shibu [3], Mohanan Monisha [3] and Radhakumari Maya [3]

[1] Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, India
[2] Department of Mathematics, Université de Caen Basse-Normandie, F-14032 Caen, France
[3] Department of Statistics, University College, Thiruvananthapuram 695 034, India
* Correspondence: christophe.chesneau@unicaen.fr

Abstract: The importance of Lagrangian distributions and their applicability in real-world events have been highlighted in several studies. In light of this, we create a new zero-truncated Lagrangian distribution. It is presented as a generalization of the zero-truncated binomial distribution (ZTBD) and hence named the Lagrangian zero-truncated binomial distribution (LZTBD). The moments, probability generating function, factorial moments, as well as skewness and kurtosis measures of the LZTBD are discussed. We also show that the new model's finite mixture is identifiable. The unknown parameters of the LZTBD are estimated using the maximum likelihood method. A broad simulation study is executed as an evaluation of the well-established performance of the maximum likelihood estimates. The likelihood ratio test is used to assess the effectiveness of the third parameter in the new model. Six COVID-19 datasets are used to demonstrate the LZTBD's applicability, and we conclude that the LZTBD is very competitive on the fitting objective.

Keywords: Lagrangian zero-truncated binomial distribution; index of dispersion; maximum likelihood method; generalized likelihood ratio test; COVID-19; simulation

MSC: 60E05; 62E10; 62F10

check for updates

Citation: Irshad, M.R.; Chesneau, C.; Shibu, D.S.; Monisha, M.; Maya, R. A Novel Generalization of Zero-Truncated Binomial Distribution by Lagrangian Approach with Applications for the COVID-19 Pandemic. *Stats* **2022**, *5*, 1004–1028. https://doi.org/10.3390/stats5040060

Academic Editor: Wei Zhu

Received: 21 September 2022
Accepted: 27 October 2022
Published: 30 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Certain discrete distributions whose support is a set of positive integers are known as zero-truncated discrete distributions (ZTDDs). ZTDDs are used in ecology to represent data relating to counts, such as the number of flower heads, fly eggs, European red mites, or the number of times snowshoe hares were collected over seven days. These distributions are also employed in sociology to simulate data such as the size of human groups in parks, beaches, and public locations. As a result, ZTDDs have applications in practically every discipline of study, including biology, medicine, psychology, demography, and political science. In particular, the zero-truncated Poisson distribution (ZTPD) was used in [1] to analyze the number of eggs and gall-cell counts in flower heads. The authors of [2] used the ZTPD to model deer hunting in California. The author of [3] employed the zero-truncated negative binomial distribution (ZTNBD) to model the number of children ever born to a sample of moms over 40 years old; additionally, the authors of [4] used the ZTNBD in a regression model to treat over-dispersed count data of ischemic stroke hospitalizations. The author of [5] analyzed stroke count data based on the ZTPD, ZTNBD, and zero-truncated generalized negative binomial distribution (ZTGNBD). The application of the ZTNBD in the investigation of rare species abundance and hospital stays was discussed in [6]. The authors of [7] considered the use of ZTBD as a randomization device.

Considering the health aspect, many different diseases, ranging from the ordinary cold to much more dangerous ailments like Middle East Respiratory Syndrome (MERS) and

Severe Acute Respiratory Syndrome (SARS), can be caused by the large family of viruses known as coronaviruses. The first cases of novel coronavirus (COVID-19) were found in Wuhan, China, in 2019 and the World Health Organization (WHO) has proclaimed it to be a pandemic. A coordinated international effort has been launched to halt the virus from spreading further, and the scientific community has contributed by starting various investigations. When it comes to model phenomena, statisticians play a critical role, and several attempts have already been made in the statistical literature. To estimate the daily new COVID-19 instances in China, the author of [8] used a mathematical model called SIR distribution. The authors of [9] developed a discrete version of the generalized Lindley distribution to model the daily new cases and deaths in the COVID-19 count data. A discrete type-2 half logistic exponential distribution was presented in [10] for estimating the number of COVID-19 deaths in Pakistan and Saudi Arabia. To model COVID-19 data in Singapore, the authors of [11] employed a discrete Marshall–Olkin inverted Topp–Leone distribution. Following the discovery of such a widespread epidemic, at least one new positive case is reported daily in practically all nations. To the best of our knowledge, ZTDDs are the most appropriate statistical model for such a situation. As far as we know, not even one statistician has attempted to model regularly occurring positive instances using ZTDDs. Hence, in this article, our aim is to propose a novel ZTDD to model the daily new positive cases. Furthermore, based on the same ZTDD, we also tried to model the number of deaths attributable to COVID-19 in a day.

On the other hand, Lagrangian distributions are a subclass of Lagrangian expansions, which were initially introduced in [12]. The authors of [13,14] introduced a discrete Lagrangian family (DLF) of probability distributions, which encompasses a vast and important class of probability distributions. It includes many families. Additionally, the authors of [14] showed that, under certain conditions, all discrete Lagrangian distributions converge to the normal and inverse Gaussian distributions. The author of [15], who discovered the Lagrangian negative binomial distribution, demonstrated its utility in a queuing process. The authors of [16] created the Lagrangian Katz family. The authors of [17] looked at how Lagrangian probability distributions can be employed to solve inferential difficulties in random mapping theory. The generalized Poisson gamma dependency model was developed in [18] using Lagrangian probability models. For collisional turbulent fluid-particle flows, the authors of [19] used the Lagrangian probability density function (pdf) models. The above-mentioned importance of the Lagrangian distributions immensely motivated us to propose a new ZTDD based on the Lagrangian approach. Therefore, based on the Lagrangian technique, we propose a unique ZTDD known as the Lagrangian zero-truncated binomial distribution (LZTBD) that can serve as a discrete model for a variety of count datasets.

The remaining parts of the paper are organized as follows: Section 2 presents some preliminaries of the Lagrangian probability distribution. In Section 3, we discuss the definition and properties of the LZTBD. The finite mixture of the new Lagrangian model is displayed in Section 4. In Section 5, we derive the maximum likelihood (ML) estimation method to estimate the unknown parameters of the LZTBD. The significance of the additional parameter is tested by using a generalized likelihood ratio test in Section 6. The finite sample performance of the ML estimation method is analyzed in Section 7 with a simulation study. Six real-world datasets are considered in Section 8 to demonstrate the usefulness of the proposed model. The concluding remarks are given in Section 9.

2. Some Preliminaries

In order to introduce the DLF, we consider the following Lagrange expansion presented in [20,21]:

$$\frac{g_2(z)}{1-\frac{zg_1'(z)}{g_1(z)}} = \sum_{r=0}^{\infty} b_r u^r, \tag{1}$$

where $u = \frac{z}{g_1(z)}$, $b_0 = g_2(0)$, $b_r = \frac{1}{r!}D^r\{(g_1(z))^r g_2(z)\}|_{z=0}$, with $D^r = \frac{d^r}{dz^r}$ and $r = 0, 1, 2, 3, \ldots$

Consider the Lagrange expansion given in (1), where $g_1(z)$ and $g_2(z)$ are successively differentiable analytic functions over $[-1, 1]$ such that $g_1(1) = g_2(1) = 1$, $g_1(0) \neq 0$, and $g_2(0) \geq 0$. A new type of probability mass function (pmf) was defined in [13,22], and it is indicated as follows:

$$P(X = r) = \frac{b_r}{\sum_{r=0}^{\infty} b_r}, \quad r = 0, 1, 2, 3, \ldots \tag{2}$$

provided that $\sum_{r=0}^{\infty} b_r$ is finite.

Putting $z = u = 1$ into (1), we obtain

$$\frac{g_2(1)}{1 - g_1'(1)} = \frac{1}{1 - g_1'(1)} = \sum_{r=0}^{\infty} b_r,$$

which gives, from (2),

$$P(X = r) = \frac{(1 - g_1'(1))D^r[\{(g_1(z))^r g_2(z)\}]|_{z=0}}{r!}, \quad r = 0, 1, 2, 3 \ldots \tag{3}$$

This pmf defined the DLF in the broad sense.

The corresponding probability generating function (pgf) is given by

$$\Psi(u) = \frac{(1 - g_1'(z))g_2(z)}{1 - \frac{z g_1'(z)}{g_1(z)}}, \tag{4}$$

where $z = u\, g_1(z)$.

Given the applications of the DLF built with $g_1(z)$ and $g_2(z)$ in (3), it is worthwhile to investigate additional horizon distributions using the new function $g_2(z)$. This is the basis for the study's updated distribution, which is shown below.

3. Construction of Lagrangian Zero-Truncated Binomial Distribution

The LZTBD is introduced in this section as a new member of the DLF.

Proposition 1. *Assume that the random variable (rv) X follows the LZTBD, in which* $0 < \alpha < \beta^{-1}, 0 < \beta < 1$ *and* $\gamma > 0$. *Then, the pmf of X is given by*

$$f(x) = P(X = x) = \frac{(1 - \alpha\beta)\left[\binom{\gamma + \alpha x}{x} - \binom{\alpha x}{x}\right]}{1 - (1 - \beta)^{\gamma}} \beta^x (1 - \beta)^{\gamma + \alpha x - x}, \quad x = 1, 2, 3 \ldots, \tag{5}$$

where $\binom{y}{x}$ *stands for the generalized binomial coefficient, that is* $\binom{y}{x} = \frac{y(y-1)\ldots(y-x+1)}{x!}$.

Proof. Let $g_1(z) = (1 - \beta + \beta z)^{\alpha}$ and $g_2(z) = \frac{(1-\beta+\beta z)^{\gamma} - (1-\beta)^{\gamma}}{1-(1-\beta)^{\gamma}}$ which satisfy the statements given in Section 2. Using the DLF given in (3), the pmf of the LZTBD can be derived as follows:

$$\begin{aligned} f(x) &= (1 - g_1'(1)) \frac{D^x[(g_1(z))^x g_2(z)]\Big|_{z=0}}{x!} \\ &= (1 - \alpha\beta) \frac{D^x\left[(1 - \beta + \beta z)^{\alpha x} \frac{(1-\beta+\beta z)^{\gamma} - (1-\beta)^{\gamma}}{1-(1-\beta)^{\gamma}}\right]\Big|_{z=0}}{x!} \\ &= \frac{(1 - \alpha\beta)\left[\binom{\gamma + \alpha x}{x} - \binom{\alpha x}{x}\right]}{1 - (1 - \beta)^{\gamma}} \beta^x (1 - \beta)^{\gamma + \alpha x - x}. \end{aligned}$$

Thus, the proof is completed. □

The distribution described in (5) is denoted as LZTBD(α, β, γ), and one can note $X \sim$ LZTBD(α, β, γ) to inform that an rv denoted by X follows the LZTBD with parameters α, β, and γ. Some special cases from the LZTBD are described below:

- For $\alpha \to 0$, the LZTBD(α, β, γ) reduces to the one-parameter ZTBD. In this sense, the LZTBD is a generalization of the ZTBD;
- For $\gamma = 1$, LZTBD(α, β, γ) reduces to the Lagrangian weighted Consul distribution given in [23].

Now, Figure 1 portrays the graphical representation of the LZTBD for different parameter values of α, β, and γ.

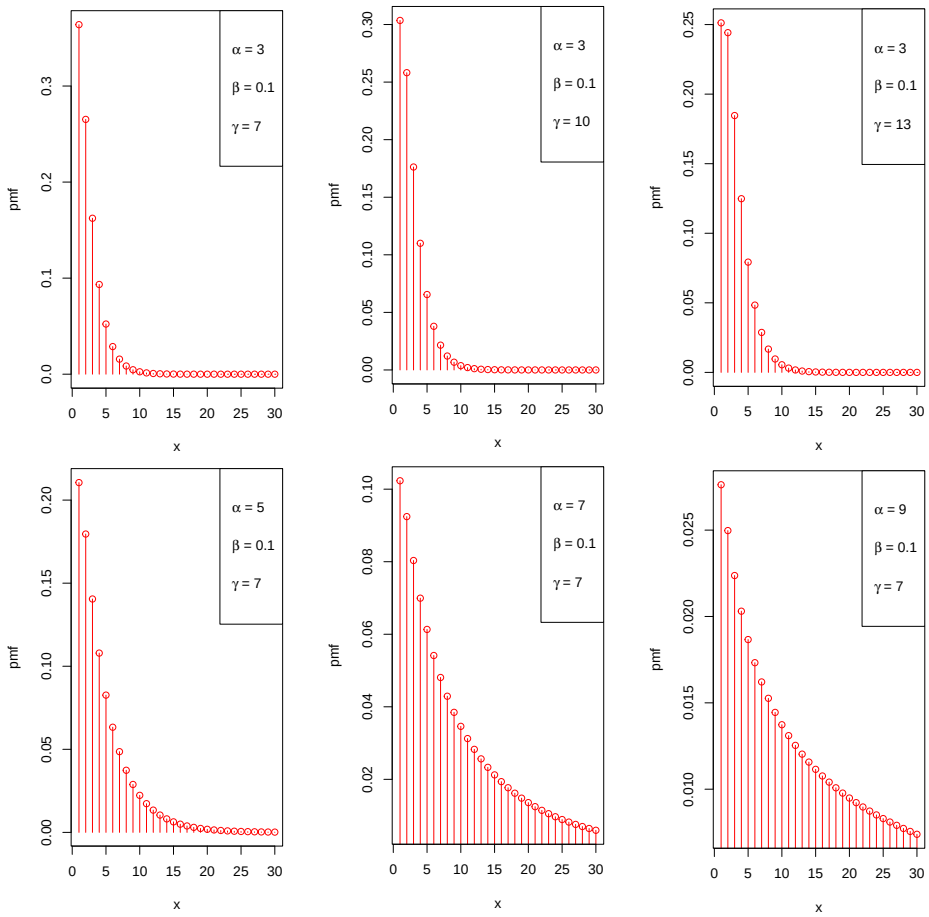


Figure 1. *Cont.*

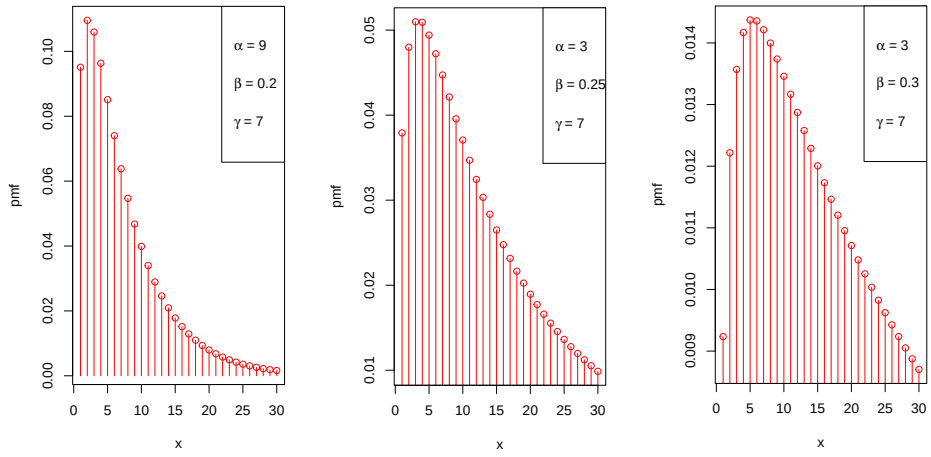


Figure 1. Various shapes of the probability mass function (pmf) of the LZTBD for different values of the parameters.

The hazard rate function (hrf) of the LZTBD is obtained by substituting the pmf in the following equation:

$$f_x = P(X = x | X \geq x) = \frac{f(x)}{\sum_{j=x}^{\infty} f(j)}, \; x = 1, 2, 3, \ldots \tag{6}$$

Following (6), it goes without saying that determining the closed form expression of the hrf is more difficult, although, in order to determine the shape of the hrf, we sketch its graph. Figure 2 demonstrates the following facts about the shapes of the hrf of the LZTBD, indicating that the LZTBD has all of the typical shapes, such as increasing, decreasing, and bathtub shapes for varying parameter values.

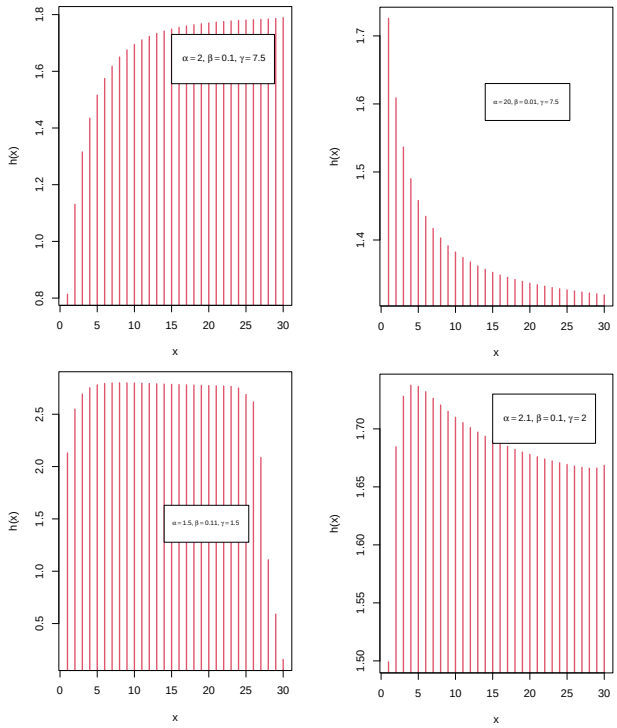


Figure 2. Various shapes of the hazard rate function (hrf) of the LZTBD for different parameter values.

Furthermore, the choice of various specific functions for $g_2(z)$ will provide various members of DLF. In the following, we list some DLF distributions available in the literature.

3.1. Poisson-Binomial Distribution

If we take $g_1(z) = (1-\beta+\beta z)^{\alpha}$ and $g_2(z) = e^{\gamma(z-1)}$, based on (3), the pmf of the considered distribution is obtained as

$$\begin{aligned} f_1(x) &= \frac{D^x\left[(1-\beta+\beta z)^{\alpha x} e^{\gamma(z-1)}\right]\Big|_{z=0}}{x!} \\ &= \frac{(1-\alpha\beta)e^{-\gamma}\gamma^x(1-\beta)^{\alpha x}}{x!}\,{}_2F_0\left(-x,-\alpha x;\frac{\beta}{\gamma(1-\beta)}\right),\ x=0,1,2,3\ldots, \end{aligned}$$

where ${}_2F_0$ is the hypergeometric function. Since their pmfs coincide, the corresponding distribution is identified as the Poisson-binomial distribution (see [23]).

3.2. Weighted Consul Distribution

If we take $g_1(z) = (1-\beta+\beta z)^{\alpha}$ and $g_2(z) = z$, based on (3), the pmf of the considered distribution can be derived as

$$\begin{aligned} f_2(x) &= (1-\alpha\beta)\frac{D^x[(1-\beta+\beta z)^{\alpha x} z]\Big|_{z=0}}{x!} \\ &= (1-\alpha\beta)\binom{\alpha x}{x-1}\beta^{x-1}(1-\beta)^{\alpha x-x+1},\ x=1,2,3\ldots, \end{aligned}$$

which is the pmf of the weighted Consul distribution (see [23]).

3.3. Weighted Delta Binomial Distribution

If we take $g_1(z) = (1-\beta+\beta z)^{\alpha}$ and $g_2(z) = z^{\gamma}$, based on (3), the pmf of the considered distribution is obtained as

$$\begin{aligned} f_3(x) &= (1-\alpha\beta)\frac{D^x[(1-\beta+\beta z)^{\alpha x} z^{\gamma}]\Big|_{z=0}}{x!} \\ &= (1-\alpha\beta)\binom{\alpha x}{x-\gamma}\beta^{x-\gamma}(1-\beta)^{\alpha x-x+\gamma},\ x=\gamma,\gamma+1,\gamma+2\ldots, \end{aligned}$$

which corresponds to the pmf of the weighted delta binomial distribution (see [23]).

3.4. Linear Function Binomial Distribution

If we take $g_1(z) = (1-\beta+\beta z)^{\alpha}$ and $g_2(z) = (1-\beta+\beta z)^{\gamma}$, based on (3), the pmf of the considered distribution can be derived as

$$\begin{aligned} f_4(x) &= (1-\alpha\beta)\frac{D^x[(1-\beta+\beta z)^{\alpha x}(1-\beta+\beta z)^{\gamma}]\Big|_{z=0}}{x!} \\ &= (1-\alpha\beta)\binom{\gamma+\alpha x}{x}\beta^{x}(1-\beta)^{\gamma+\alpha x-x},\ x=0,1,2\ldots, \end{aligned}$$

which is the pmf of the linear function binomial distribution (see [23]).

Proposition 2. *Let X be an rv following the LZTBD. Then, the median of X is defined by the smaller integer m greater or equal to* 1 *such that*

$$\sum_{x=1}^{m}\left[\binom{\gamma+\alpha x}{x}-\binom{\alpha x}{x}\right]\beta^x(1-\beta)^{\alpha x-x} \geq \frac{1-(1-\beta)^{\gamma}}{2(1-\beta)^{\gamma}(1-\alpha\beta)}. \tag{7}$$

Proof. By the definition, m is the smallest integer in the support of the rv, i.e., $\{1,2,\ldots\}$, such that $P(X \leq m) \geq \frac{1}{2}$, which is equivalent to the desired result. □

Proposition 3. *Let X be a rv following the LZTBD. Then, the mode of X, denoted by x_m, exists in $\{1, 2, \ldots\}$ and lies in the case:*

$$\frac{\eta(x_m+1)}{\eta(x_m)} \leq \frac{1}{\beta(1-\beta)^{\alpha-1}} \leq \frac{\eta(x_m)}{\eta(x_m-1)}, \tag{8}$$

where $\eta(x_m) = \binom{\gamma+\alpha\, x_m}{x_m} - \binom{\alpha\, x_m}{x_m}$.

Proof. We must find the integer $x = x_m$ for which $f(x)$ has the greatest value. That is, we aim to solve $f(x) \geq f(x-1)$ and $f(x) \geq f(x+1)$. First, note that $f(x)$ can also be written as

$$f(x) = \frac{(1-\alpha\beta)\beta^x(1-\beta)^{\gamma+\alpha x-x}\eta(x)}{1-(1-\beta)^\gamma},$$

where $\eta(x) = \binom{\gamma+\alpha x}{x} - \binom{\alpha x}{x}$.

Obviously, $f(x) \geq f(x-1)$ implies that

$$\frac{\eta(x)}{\eta(x-1)} \geq \frac{1}{\beta(1-\beta)^{\alpha-1}}. \tag{9}$$

Additionally, $f(x) \geq f(x+1)$ implies that

$$\frac{\eta(x+1)}{\eta(x)} \leq \frac{1}{\beta(1-\beta)^{\alpha-1}}. \tag{10}$$

By combining (9) and (10), we obtain (8). □

Proposition 4. *The pgf of an rv X following the LZTBD is expressed as*

$$\Psi(u) = E(u^X) = \frac{(1-\alpha\beta)\{(1-\beta+\beta z)^\gamma - (1-\beta)^\gamma\}}{(1-(1-\beta)^\gamma)\left(1-\frac{z\,\alpha\beta}{1-\beta+\beta z}\right)}, \tag{11}$$

where $z = u\,(1-\beta+\beta z)^\alpha$.

Proof. Using (4), the pgf of the LZTBD is of the following form:

$$\Psi(u) = \frac{(1-g_1'(1))g_2(z)}{\left(1-\frac{zg_1'(z)}{g_1(z)}\right)} = \frac{(1-\alpha\beta)\{(1-\beta+\beta z)^\gamma - (1-\beta)^\gamma\}}{(1-(1-\beta)^\gamma)\left(1-\frac{z\,\alpha\beta}{1-\beta+\beta z}\right)}.$$

Thus, the proof is complete. □

Corollary 1. *The moment generating function (mgf) of an rv X following the LZTBD is obtained by putting $z = e^s$ and $u = e^k$ in (11). That is,*

$$M(k) = E(e^{kX}) = \frac{(1-\alpha\beta)\{(1-\beta+\beta e^s)^\gamma - (1-\beta)^\gamma\}}{(1-(1-\beta)^\gamma)\left(1-\frac{\alpha\beta e^s}{1-\beta+\beta e^s}\right)}, \tag{12}$$

where $s = k + \alpha\log(1-\beta+\beta e^s)$.

Corollary 2. *The cumulant generating function (cgf) of an rv X following the LZTBD becomes*

$$C(k) = \log[M(k)] = \log\left[\frac{(1-\alpha\beta)\{(1-\beta+\beta e^k)^\gamma - (1-\beta)^\gamma\}}{(1-(1-\beta)^\gamma)\left(1-\frac{\alpha\beta e^s}{1-\beta+\beta e^s}\right)}\right], \tag{13}$$

where $s = k + \alpha \log(1 - \beta + \beta e^s)$.

Proposition 5. *Let $X_1, X_2, \ldots, X_n$ be n independently and identically distributed (iid) rvs following the LZTBD(α, β, γ). Then, the distribution of the random sum variable $V = \sum_{i=1}^{n} X_i$ has the following pgf:*

$$\Psi_1(u) = \frac{(1-\alpha\beta)^n\{(1-\beta+\beta z)^\gamma - (1-\beta)^\gamma\}^n}{(1-(1-\beta)^\gamma)^n\left(1-\frac{z\alpha\beta}{1-\beta+\beta z}\right)^n},$$

where $z = u\,(1-\beta+\beta z)^\alpha$.

Proof. Based on the pgf of the LZTBD given in (11), the pgf of the rv V becomes

$$\begin{aligned}\Psi_1(u) = \mathrm{E}(u^V) = \mathrm{E}(u^{X_1+X_2+\ldots+X_n}) = \prod_{i=1}^{n}\mathrm{E}(u^{X_i}) = \prod_{i=1}^{n}\Psi(u) = [\Psi(u)]^n \\ = \frac{(1-\alpha\beta)^n\{(1-\beta+\beta z)^\gamma - (1-\beta)^\gamma\}^n}{(1-(1-\beta)^\gamma)^n\left(1-\frac{z\alpha\beta}{1-\beta+\beta z}\right)^n}.\end{aligned}$$

This completes the proof. □

Proposition 6. *For any integer $r \geq 1$, the r^{th} factorial moment of an rv X following the LZTBD(α, β, γ) is given by*

$$\begin{aligned}\mu_{[r]} &= E[X(X-1)(X-2)\ldots(X-r+1)] \\ &= \left\{(1-(1-\beta)^\gamma)^{-1}D^r\left(1-\beta+\beta z)^\gamma\right)\right. \\ &\left. + \frac{\alpha\beta\sum_{i=1}^{r}(r-i+1)\mu_{[r-i]}D^i\left(u(1-\beta+\beta z)^{\alpha-1}\right)}{1-\alpha\beta}\right\}\Big|_{u=z=1},\end{aligned} \tag{14}$$

where $z = u(1-\beta+\beta z)^\alpha$.

Proof. By definition, the r^{th} factorial moment of the LZTBD(α, β, γ) is obtained by successively differentiating $\Psi(u)$ given in (11) r times with respect to (wrt) u and by putting $u = z = 1$. First, note that

$$(1 - ug_1'(z))\Psi(u) = (1 - g_1'(1))g_2(z).$$

Taking the first derivative wrt u on both sides, we obtain

$$\Psi(u)D^1(1 - ug_1'(z)) + \Psi'(u)(1 - ug_1'(z)) = (1 - g_1'(1))D^1(g_2(z)).$$

Taking the second derivative of the above equation wrt u on both sides, we obtain

$$\Psi(u)D^2(1-ug_1'(z)) + 2D^1(1-ug_1'(z))\Psi'(u) + (1-ug_1'(z))\Psi''(u) = (1-g_1'(1))D^2(g_2(z)).$$

Proceeding in a similar manner, the r^{th} derivative is of the following form:

$$D^r\Psi(u) = \frac{(1-g_1'(1))D^r(g_2(z)) - \sum_{i=1}^{r}(r-i+1)D^i(1-ug_1'(z))D^{r-i}\Psi(u)}{1-ug_1'(z)}. \tag{15}$$

Substitute $g_1(z) = (1-\beta+\beta z)^\alpha$, $g_2(z) = \frac{(1-\beta+\beta z)^\gamma-(1-\beta)^\gamma}{1-(1-\beta)^\gamma}$ and $z = u = 1$, in (15), we obtain (14). □

Proposition 7. *The mean (μ) and variance (σ^2) for the LZTBD are of the following forms, respectively,*

$$\mu = \frac{\gamma\beta}{(1-(1-\beta)^{\gamma})(1-\alpha\beta)} + \frac{\alpha\beta(1-\beta)}{(1-\alpha\beta)^2} \tag{16}$$

and

$$\begin{aligned}\sigma^2 = &\frac{\alpha\beta + (\alpha^2-3\alpha)\beta^2 + (2\alpha-\alpha^2)\beta^3 + 2\alpha^3\beta^4}{(1-\alpha\beta)^4} \\ &+ \frac{\gamma^2\beta^2 + \gamma\beta(1-\beta) - \gamma^2\alpha\beta^3}{(1-(1-\beta)^{\gamma})(1-\alpha\beta)^3} - \frac{\gamma^2\beta^2}{(1-(1-\beta)^{\gamma})^2(1-\alpha\beta)^2}.\end{aligned} \tag{17}$$

Proof. Using (14), we obtain

$$\begin{aligned}\mu = E(X) =& \frac{g_2'(1)}{(1-g_1'(1))} + \frac{g_1''(1) + g_1'(1) - (g_1'(1))^2}{(1-g_1'(1))^2} \\ =& \frac{\gamma\beta}{(1-(1-\beta)^{\gamma})(1-\alpha\beta)} + \frac{\alpha\beta(1-\beta)}{(1-\alpha\beta)^2}.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\sigma^2 &= E(X(X-1)) + E(X) - (E(X))^2 \\ &= \frac{g_2''(1) + g_2'(1) - (g_2'(1))^2}{(1-g_1'(1))^2} + \frac{(1+g_2'(1))g_1''(1) + g_1'(1) - (g_1'(1))^2)}{(1-g_1'(1))^3} \\ &+ \frac{g_1'''(1) + g_1'(1)g_1''(1) + 2g_1''(1)}{(1-g_1'(1))^3} + \frac{2(g_1''(1))^2}{(1-g_1'(1))^4} \\ &= \frac{\alpha\beta + (\alpha^2-3\alpha)\beta^2 + (2\alpha-\alpha^2)\beta^3 + 2\alpha^3\beta^4}{(1-\alpha\beta)^4} \\ &+ \frac{\gamma^2\beta^2 + \gamma\beta(1-\beta) - \gamma^2\alpha\beta^3}{(1-(1-\beta)^{\gamma})(1-\alpha\beta)^3} - \frac{\gamma^2\beta^2}{(1-(1-\beta)^{\gamma})^2(1-\alpha\beta)^2}.\end{aligned}$$

The desired expressions are obtained. □

A normalized measure of dispersion can be obtained by utilizing the variance-to-mean relationship. This measure is the well-known index of dispersion (IOD). The next result expressed it for the LZTBD, among others.

Proposition 8. *The IOD and coefficient of variation (CV) for the LZTBD are given as, respectively,*

$$\text{IOD} = \frac{\frac{\alpha\beta+(\alpha^2-3\alpha)\beta^2+(2\alpha-\alpha^2)\beta^3+2\alpha^3\beta^4}{(1-\alpha\beta)^4} + \frac{\gamma^2\beta^2+\gamma\beta(1-\beta)-\gamma^2\alpha\beta^3}{(1-(1-\beta)^{\gamma})(1-\alpha\beta)^3} - \frac{\gamma^2\beta^2}{(1-(1-\beta)^{\gamma})^2(1-\alpha\beta)^2}}{\frac{\gamma\beta}{(1-(1-\beta)^{\gamma})(1-\alpha\beta)} + \frac{\alpha\beta(1-\beta)}{(1-\alpha\beta)^2}} \tag{18}$$

and

$$\text{CV} = \frac{\sqrt{\frac{\alpha\beta+(\alpha^2-3\alpha)\beta^2+(2\alpha-\alpha^2)\beta^3+2\alpha^3\beta^4}{(1-\alpha\beta)^4} + \frac{\gamma^2\beta^2+\gamma\beta(1-\beta)-\gamma^2\alpha\beta^3}{(1-(1-\beta)^{\gamma})(1-\alpha\beta)^3} - \frac{\gamma^2\beta^2}{(1-(1-\beta)^{\gamma})^2(1-\alpha\beta)^2}}}{\frac{\gamma\beta}{(1-(1-\beta)^{\gamma})(1-\alpha\beta)} + \frac{\alpha\beta(1-\beta)}{(1-\alpha\beta)^2}}. \tag{19}$$

Proof. We have

$$\begin{aligned}\mathrm{IOD} &= \frac{\sigma^2}{\mu} \\ &= \frac{\frac{\alpha\beta+(\alpha^2-3\alpha)\beta^2+(2\alpha-\alpha^2)\beta^3+2\alpha^3\beta^4}{(1-\alpha\beta)^4}+\frac{\gamma^2\beta^2+\gamma\beta(1-\beta)-\gamma^2\alpha\beta^3}{(1-(1-\beta)^\gamma)(1-\alpha\beta)^3}-\frac{\gamma^2\beta^2}{(1-(1-\beta)^\gamma)^2(1-\alpha\beta)^2}}{\frac{\gamma\beta}{(1-(1-\beta)^\gamma)(1-\alpha\beta)}+\frac{\alpha\beta(1-\beta)}{(1-\alpha\beta)^2}}.\end{aligned}$$

Analogously, the CV is given by

$$\begin{aligned}\mathrm{CV} &= \frac{\sqrt{\sigma^2}}{\mu} \\ &= \frac{\sqrt{\frac{\alpha\beta+(\alpha^2-3\alpha)\beta^2+(2\alpha-\alpha^2)\beta^3+2\alpha^3\beta^4}{(1-\alpha\beta)^4}+\frac{\gamma^2\beta^2+\gamma\beta(1-\beta)-\gamma^2\alpha\beta^3}{(1-(1-\beta)^\gamma)(1-\alpha\beta)^3}-\frac{\gamma^2\beta^2}{(1-(1-\beta)^\gamma)^2(1-\alpha\beta)^2}}}{\frac{\gamma\beta}{(1-(1-\beta)^\gamma)(1-\alpha\beta)}+\frac{\alpha\beta(1-\beta)}{(1-\alpha\beta)^2}}.\end{aligned}$$

□

A probabilistic model's asymmetry degree and flatness are commonly assessed by their skewness and kurtosis coefficients, respectively. The third central moment, normalized by the variance raised to the power of $3/2$, can be used to calculate the first, whereas the fourth central moment divided by the square of the variance can be used to calculate the second. Mean, variance, CV, IOD, skewness, and kurtosis for selected values of parameters of the LZTBD(α, β, γ) are summarized in Table 1. From this table, it is evident that the LZTBD possesses both over-dispersion (IOD > 1) and under-dispersion (IOD < 1) for varying parameter values. It is also noted that the LZTBD is mainly right-skewed, and has several kurtosis levels.

Table 1. Values of some moment measures of the LZTBD for various values of parameters α, β, and γ.

β	γ	α	Mean	Variance	CV	IOD	Skewness	Kurtosis
0.5	1	0.3	1.2802	0.1501	0.3026	0.1172	3.8847	14.6003
		0.5	1.5555	0.4444	0.4285	0.2857	3.3511	10.1883
		0.7	1.9526	1.2205	0.5657	0.6250	2.8561	6.9567
		0.9	2.5619	3.4546	0.7254	1.3484	2.3669	4.3115
		1.1	3.5802	10.7636	0.9163	3.0063	1.8712	2.1372
0.4	3	0.3	1.8323	0.7348	0.4678	0.4010	2.9568	7.4378
		0.5	2.1007	1.1358	0.5073	0.5406	2.6612	5.7455
		0.7	2.4499	1.8497	0.5551	0.7550	2.3650	4.2161
		0.9	2.9189	3.2112	0.6139	1.1001	2.0624	2.8257
		1.1	3.5750	6.0267	0.6866	1.6857	1.7492	1.5815
0.3	5	0.3	2.0574	1.0607	0.5005	0.5155	2.7004	5.9584
		0.5	2.2665	1.4134	0.5242	0.6236	2.5057	4.9160
		0.7	2.5178	1.9288	0.5515	0.7660	2.3089	3.9376
		0.9	2.8245	2.7065	0.5824	0.9582	2.1084	3.0172
		1.1	3.2056	3.9230	0.6178	1.2237	1.9027	2.1568
0.2	7	0.3	1.9389	1.0021	0.5163	0.5168	2.8934	7.2006
		0.5	2.0671	1.2174	0.5337	0.5889	2.7471	6.3127
		0.7	2.2113	1.4917	0.5523	0.6746	2.6045	5.5028
		0.9	2.3745	1.8451	0.5720	0.7770	2.4639	4.7524
		1.1	2.5601	2.3060	0.5931	0.9006	2.3242	4.0507
0.1	9	0.3	1.5433	0.5853	0.4957	0.3792	3.7718	13.9025
		0.5	1.5963	0.6626	0.5099	0.4150	3.6638	13.0428
		0.7	1.6526	0.7503	0.5241	0.4540	3.5590	12.2257
		0.9	1.7123	0.8501	0.5384	0.4964	3.4572	11.4502
		1.1	1.7757	0.9639	0.5528	0.5428	3.3580	10.7145

4. Identifiability

Finite mixture models have received a lot of attention in recent years in real contexts. In astronomy, biology, genetics, medicine, psychiatry, marketing, and other fields, mixture models are widely utilized (see [24]). We derive finite mixtures of the LZTBD(α, β, γ) in this section. This mixed model may be appropriate in the context of future initiatives.

Let Y be a discrete rv with the pmf $h(y) = \sum_{i=1}^{g} l_i h_i(y)$, where $i = 1, 2, ...g$, $l_i > 0$ such that $\sum_{i=1}^{g} l_i = 1$, $h_i(y) \geq 0$ and $\sum_y h_i(y) = 1$. Then, we state that Y has a mixture distribution and $h(y)$ is a finite mixture of distributions. The constants $l_1, l_2, ..., l_g$ are known as mixing weights and $h_1(y), h_2(y), \ldots, h_g(y)$, the components of the mixture. We denote as Θ the collection of all distinct parameters in the components.

Let $\Sigma = \{U(y; \theta_i) : \theta_i \in \Theta\}$ be the class of pmf's from which mixtures are to be formed. Then, the class of finite mixtures of Σ with the appropriate class of pmf's is $\hat{\Delta} = \{\Delta(y) : \Delta(y) = \sum_{i=1}^{g} l_i U(y; \theta_i), l_i > 0, U(y; \theta_i) \in \Sigma, i = 1, 2, ...g\}$. So that $\hat{\Delta}$ is the convex hull of Σ.

Definition 1. *An intege-valued rv Y is said to have a g component mixture of the LZTBDs if it has the pmf $h(y) = P(Y = y)$ of the following form:*

$$h(y) = \sum_{i=1}^{g} l_i h_i(y), \tag{20}$$

where $0 \leq l_i \leq 1$, for each $i = 1, 2, 3 \ldots, g, \sum_{i=1}^{g} l_i = 1$,

$$h_i(y) = \frac{(1 - \alpha_i \beta_i)[\binom{\gamma_i + \alpha_i y}{y} - \binom{\alpha_i y}{y}] \beta_i^y (1 - \beta_i)^{\gamma_i + \alpha_i y - y}}{(1 - (1 - \beta_i)^{\gamma_i})}, \quad y = 1, 2, \ldots, \tag{21}$$

with $\gamma_i > 0, 0 \leq \alpha_i < \beta_i^{-1}$ and $0 < \beta_i < 1$ for each $i = 1, 2, \ldots, g$.

A distribution with pmf given in (20) is called the Lagrangian zero-truncated binomial mixture distribution with g components (LZTBMDg).

The following theorem from [25] is adopted to construct the identifiability conditions of the finite mixture model.

Theorem 1. *A necessary and sufficient condition for $\hat{\Delta}$ to be identifiable is that Δ should be linearly independent over the field of real numbers.*

Proof. The proof is stated in [25], hence, it is not included here. □

Next, applying Theorem 1, we outline the LZTBMDg's identifiability requirements.

Theorem 2. *The identifiability conditions for the LZTBMDg with the pmf $h(y)$ as given in (20) are $\alpha_i \neq \alpha_j$, $\beta_i \neq \beta_j$, and $\gamma_i \neq \gamma_j$ for $i, j \in \{1, 2, \ldots, g\}$, such that $i \neq j$.*

Proof. For the first step, take $g = 2$ and consider the following equation:

$$b_1 F_1(y) + b_2 F_2(y) = 0, \tag{22}$$

where b_1 and b_2 are any two arbitrary real numbers, $F_1(y) = \sum_{j=1}^{y} h(j)$ and $F_2(y) = \sum_{j=1}^{y} \phi(j)$ for $y = 1, 2, \ldots$, in which $\phi(j)$ is obtained from $h(j)$ by replacing α_j by τ_j , β_j by δ_j and γ_j by ω_j.

Assume that for each $i = 1, 2$ and $\alpha_i \neq \tau_i$, $\beta_i \neq \delta_i$ and $\gamma_i \neq \omega_i$. Thus, for $l_1 = l$, we have

$$\begin{aligned} F_1(y) = l \sum_{j=1}^{y} \frac{(1-\alpha_1\beta_1)\left[\binom{\gamma_1+\alpha_1 j}{j} - \binom{\alpha_1 j}{j}\right]\beta_1^j(1-\beta_1)^{\gamma_1+\alpha_1 j-j}}{1-(1-\beta_1)^{\gamma_1}} \\ + (1-l)\sum_{j=1}^{y} \frac{(1-\alpha_2\beta_2)\left[\binom{\gamma_2+\alpha_2 j}{j} - \binom{\alpha_2 j}{j}\right]\beta_2^j(1-\beta_2)^{\gamma_2+\alpha_2 j-j}}{1-(1-\beta_2)^{\gamma_2}} \end{aligned} \tag{23}$$

and

$$\begin{aligned} F_2(y) = l \sum_{j=1}^{y} \frac{(1-\tau_1\delta_1)\left[\binom{\omega_1+\tau_1 j}{j} - \binom{\tau_1 j}{j}\right]\delta_1^j(1-\delta_1)^{\omega_1+\tau_1 j-j}}{1-(1-\delta_1)^{\omega_1}} \\ + (1-l)\sum_{j=1}^{y} \frac{(1-\tau_2\delta_2)\left[\binom{\omega_2+\tau_2 j}{j} - \binom{\tau_2 j}{j}\right]\delta_2^j(1-\delta_2)^{\omega_2+\tau_2 j-j}}{1-(1-\delta_2)^{\omega_2}}. \end{aligned} \tag{24}$$

Now, from (22)–(24), we obtain the following equations:

$$\begin{aligned} b_1 \sum_{j=1}^{y} \frac{(1-\alpha_1\beta_1)\left[\binom{\gamma_1+\alpha_1 j}{j} - \binom{\alpha_1 j}{j}\right]\beta_1^j(1-\beta_1)^{\gamma_1+\alpha_1 j-j}}{1-(1-\beta_1)^{\gamma_1}} \\ + b_2 \sum_{j=1}^{y} \frac{(1-\tau_1\delta_1)\left[\binom{\omega_1+\tau_1 j}{j} - \binom{\tau_1 j}{j}\right]\delta_1^j(1-\delta_1)^{\omega_1+\tau_1 j-j}}{1-(1-\delta_1)^{\omega_1}} = 0 \end{aligned} \tag{25}$$

and

$$\begin{aligned} b_1 \sum_{j=1}^{y} \frac{(1-\alpha_2\beta_2)\left[\binom{\gamma_2+\alpha_2 j}{j} - \binom{\alpha_2 j}{j}\right]\beta_2^j(1-\beta_2)^{\gamma_2+\alpha_2 j-j}}{1-(1-\beta_2)^{\gamma_2}} \\ + b_2 \sum_{j=1}^{y} \frac{(1-\tau_2\delta_2)\left[\binom{\omega_2+\tau_2 j}{j} - \binom{\tau_2 j}{j}\right]\delta_2^j(1-\delta_2)^{\omega_2+\tau_2 j-j}}{1-(1-\delta_2)^{\omega_2}} = 0. \end{aligned} \tag{26}$$

Solving (25) and (26), we obtain

$$\begin{aligned} & b_1 \sum_{j=1}^{y} (1-\alpha_1\beta_1)(1-\tau_2\delta_2)(\beta_1\delta_2)^j(1-\beta_1)^{\gamma_1+\alpha_1 j-j}(1-\delta_2)^{\omega_2+\tau_2 j-j} \\ & \left\{ \frac{[\binom{\gamma_1+\alpha_1 j}{j} - \binom{\alpha_1 j}{j}][\binom{\omega_2+\tau_2 j}{j} - \binom{\tau_2 j}{j}]}{(1-(1-\beta_1)^{\gamma_1})(1-(1-\delta_2)^{\omega_2}} \right\} \\ & = b_1 \sum_{j=1}^{y} (1-\alpha_2\beta_2)(1-\tau_1\delta_1)(\beta_2\delta_1)^j(1-\beta_2)^{\gamma_2+\alpha_2 j-j}(1-\delta_1)^{\omega_1+\tau_1 j-j} \\ & \left\{ \frac{[\binom{\gamma_2+\alpha_2 j}{j} - \binom{\alpha_2 j}{j}][\binom{\omega_1+\tau_1 j}{j} - \binom{\tau_1 j}{j}]}{(1-(1-\beta_2)^{\gamma_2})(1-(1-\delta_1)^{\omega_1}} \right\}. \end{aligned} \tag{27}$$

Hence, by (27), we have $b_1 = 0$ and thus, $b_2 = 0$. Therefore, it may be inferred from Theorem 2 that $F_1(y)$ and $F_2(y)$ are linearly independent. Now that the argument may be applied to any positive integer g, the proof follows. □

Proposition 9. *The pgf of the LZTBMDg given in (20) is indicated as*

$$\Psi(u) = \sum_{i=1}^{g} \frac{(1-\alpha_i\beta_i)\{(1-\beta_i+\beta_i z_i)^{\gamma_i} - (1-\beta_i)^{\gamma_i}\}}{(1-(1-\beta_i)^{\gamma_i})(1-\frac{z_i\,\alpha_i\beta_i}{1-\beta_i+\beta_i\,z_i})}, \tag{28}$$

where $z_i = u\,(1-\beta_i+\beta_i\,z_i)^{\alpha_i}$.

Proof. The proof follows simply from Definition 1, given the pgf of the LZTBD mentioned in (11). □

5. Estimation of Parameters

In this section, we estimate the unknown parameters of the LZTBD by the ML estimation method.

It is worth mentioning that the model corresponding to the LZTBD(α, β, γ) is a tri-parametric model with parameters α, β, and γ. Let us have a random sample of size n from LZTBD and let the observed frequency be n_x, $x = 1, 2, \ldots, k$, so that $\sum_{x=1}^{k} n_x = n$, where k is the largest of the observed value having non-zero frequencies. Then, the likelihood function is given by

$$L = \prod_{x=1}^{k} \left\{ \frac{(1-\alpha\beta)\beta^x(1-\beta)^{\gamma+\alpha x-x}\left[\binom{\gamma+\alpha x}{x} - \binom{\alpha x}{x}\right]}{(1-(1-\beta)^{\gamma})} \right\}^{n_x}.$$

Therefore, the log-likelihood function is given by

$$\begin{aligned}\mathcal{L}_n = \text{logL} = {} & n\,\log(1-\alpha\beta) + n\,\overline{x}\,\log\beta + (n\gamma + n\overline{x}(\alpha-1))\,\log(1-\beta) \\ & - n\,\log(1-(1-\beta)^{\gamma}) + \sum_{x=1}^{k} n_x \left\{ \log\left[\prod_{i=0}^{x-1}(\gamma+\alpha x-i) - \prod_{i=0}^{x-1}(\alpha x-i) \right] \right\} \\ & - \sum_{x=1}^{k} n_x \log(x!),\end{aligned}$$

where $\overline{x} = \frac{1}{n}\sum_{x=1}^{k} x n_x$. The ML estimates (MLEs) are defined by maximizing $\mathcal{L}_n$ wrt the parameters. Let us denote by $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ the MLEs of α, β, and γ, respectively. On the computational side, the score vector is

$$S = \left(\frac{\partial \mathcal{L}_n}{\partial \alpha} \quad \frac{\partial \mathcal{L}_n}{\partial \beta} \quad \frac{\partial \mathcal{L}_n}{\partial \gamma} \right)^T,$$

where the partial derivatives of $\mathcal{L}_n$ wrt the parameters are

$$\frac{\partial \mathcal{L}_n}{\partial \alpha} = n\overline{x}\,\log(1-\beta) - \frac{n\,\beta}{1-\alpha\beta} + \sum_{x=1}^{k} n_x \frac{\frac{\partial}{\partial\alpha}\left[\prod_{i=0}^{x-1}(\gamma+\alpha x-i) - \prod_{i=0}^{x-1}(\alpha x-i)\right]}{\prod_{i=0}^{x-1}(\gamma+\alpha x-i) - \prod_{i=0}^{x-1}(\alpha x-i)},$$

$$\frac{\partial \mathcal{L}_n}{\partial \beta} = \frac{n\overline{x}}{\beta} - \frac{n\,\alpha}{(1-\alpha\beta)} - \frac{n\gamma + n(\alpha-1)\overline{x}}{1-\beta} - \frac{n\gamma(1-\beta)^{\gamma-1}}{1-(1-\beta)^{\gamma}}$$

and

$$\frac{\partial \mathcal{L}_n}{\partial \gamma} = n\log(1-\beta) + \frac{n\,\gamma\,(1-\beta)^{\gamma-1}}{1-(1-\beta)^{\gamma}} + \sum_{x=1}^{k} n_x \frac{\frac{\partial}{\partial\gamma}\left[\prod_{i=0}^{x-1}(\gamma+\alpha x-i)\right]}{\prod_{i=0}^{x-1}(\gamma+\alpha x-i) - \prod_{i=0}^{x-1}(\alpha x-i)}.$$

The MLEs can then be found by setting the score vector to zero, i.e., $S = 0$, and solving them concurrently. These equations cannot be solved analytically, and the R statistical software can be used to solve them numerically by means of iterative techniques such as the Newton–Raphson algorithm.

6. Likelihood Ratio Test

In this section, we test the significance of an additional parameter included in the LZTBD using the generalized likelihood ratio test (GLRT) (see [26]).

More precisely, to test the significance of the parameter α of the LZTBD(α, β, γ), here, we consider the GLRT procedure. The null hypothesis H_0 : "X follows the ZTBD" against the alternative hypothesis H_1 : "X follows the LZTBD". Here, the test statistic is given by

$$-2\log\lambda^* = 2\big(\mathcal{L}_n(\hat{\Theta}) - \mathcal{L}_n(\hat{\Theta}^*)\big), \tag{29}$$

where $\hat{\Theta}$ is the vector of MLEs of $\Theta = (\alpha, \beta, \gamma)$ with no constraints, and $\hat{\Theta}^*$ is the MLEs of Θ under H_0.

7. Simulation

We perform a simulation study by generating observations employing the R software to examine the asymptotic behavior of the MLEs of the parameters of the LZTBD. Here, we apply the inverse transformation method to simulate a LZTBD random sample (see [27]). The algorithm is as follows:

Step 1: Generate a random number from the uniform $U(0,1)$ distribution.

Step 2: $i = 1, P = (1-\alpha\beta)\ \beta\ \gamma\ (1-\beta)^{\gamma+\alpha-1}$, $F = P$.

Step 3: If $U < F$, set $X = i$ and stop.

Step 4: $P = P \times \frac{\beta(1-\beta)^{\alpha-1}(\gamma+\alpha i)\binom{\gamma+\alpha(i+1)}{i+1}}{(\gamma+\alpha(i+1))\binom{\gamma+\alpha i}{i}}$, $F = F + P, i = i+1$.

Step 5: Go to Step 3.

Conceptually, P is the probability that $X = i$, and F is the probability that X is less than or equal to i.

Additionally, indices such as MLEs, absolute biases, and mean squared errors (MSEs) are calculated using the following equations:

- Average value of MLEs: MLE$(\hat{a}) = \frac{1}{N}\sum_{i=1}^{N}\hat{a}_i$.
- Absolute average bias: Bias$(\hat{a}) = \frac{1}{N}\sum_{i=1}^{N}|\hat{a}_i - a|$.
- MSE: MSE$(\hat{a}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{a}_i - a)^2$.

Here, $a = \alpha$ or β or γ, and the index i represents the i^{th} generated sample. The simulation takes into account sample sizes of $n = 15, 50, 175, 500$, and 1000 for two different sets of parameter values of the LZTBD. We repeat the process $N = 1000$ times and report the estimates and MSEs in Table 2. From this table, one can infer that the estimates are quite stable and, more precisely, close to the true parameter values for these sample sizes. A decreasing trend is being observed in the absolute average bias and MSEs as we increase the sample size. Hence, the performance of the ML estimation is quite consistent and reliable.

Table 2. The maximum likelihood estimates (MLEs) simulation results for the parameters α, β, and γ.

Parameter Set	Sample Size	Paramters	Estimates	Absolute Bias	MSE
$\alpha = 0.5, \beta = 0.16, \gamma = 2.51$	$n = 15$	α	4.3060	3.8060	20.4138
		β	0.0838	0.0761	0.0138
		γ	1.0452	1.4647	2.7145
	$n = 50$	α	1.2245	0.7245	0.9985
		β	0.1302	0.0297	0.0046
		γ	1.3618	1.1481	2.0105
	$n = 175$	α	0.6913	0.1913	0.2249
		β	0.1582	0.0017	0.0019
		γ	1.9122	0.5977	1.6952

Table 2. *Cont.*

Parameter Set	Sample Size	Paramters	Estimates	Absolute Bias	MSE
	$n = 500$	α	0.4476	0.0523	0.0665
		β	0.1586	0.0013	0.0012
		γ	2.8752	0.3652	0.4870
	$n = 1000$	α	0.4918	0.0081	0.0243
		β	0.1609	0.0009	0.0007
		γ	2.5054	0.0049	0.3329
$\alpha = 1.1, \beta = 0.05, \gamma = 0.58$	$n = 15$	α	5.4405	4.3405	22.0280
		β	0.0429	0.0080	0.0016
		γ	0.7075	1.1275	2.5269
	$n = 50$	α	1.2827	0.0627	0.0805
		β	0.0614	0.0079	0.0004
		γ	0.9028	0.3228	2.1591
	$n = 175$	α	1.3129	0.0529	0.0802
		β	0.0421	0.0078	0.0003
		γ	0.2129	0.2639	1.8871
	$n = 500$	α	1.0494	0.0505	0.0754
		β	0.0566	0.0066	0.0003
		γ	0.6411	0.0611	0.2654
	$n = 1000$	α	1.0234	0.0302	0.0752
		β	0.0660	0.0030	0.0003
		γ	0.6112	0.0312	0.1333

8. Applications and Empirical Study

The aim of this section is to show the empirical importance of the LZTBD. We employ six genuine datasets to apply the superiority of the LZTBD fit to the more notable fields of COVID-19 with different nations, including Italy, Senegal, Pakistan, Saudi Arabia, Belgium, and Ethiopia. The graphical method used to determine the hrf of the data set is based on the Total Time on Test (TTT). Convex, concave, convex-then-concave, and concave-then-convex empirical TTT plots correspond to decreasing, increasing, bathtub shape, and upside-down bathtub shape for the corresponding hrf, respectively (see [28]). We employ the statistical software R to evaluate these datasets numerically. To show the possible benefit of the LZTBD, the distributions below are depicted:

- ZTBD with parameters β and γ, which has the following pmf:

$$f_5(x) = \binom{\gamma}{x} \frac{\beta^x (1-\beta)^{\gamma - x}}{1-(1-\beta)^x}, x = 1, 2, 3 \ldots$$

- Zero-truncated generalized binomial distribution (ZTGBD) with parameters α, β, and γ with the following pmf:

$$f_6(x) = \frac{\gamma}{\gamma + \alpha x} \binom{\gamma + \alpha x}{x} \frac{\beta^x (1-\beta)^{\gamma + \alpha x - x}}{1-(1-\beta)^\gamma}, x = 1, 2, 3 \ldots$$

- Zero-truncated discrete two parameter Poisson–Lindley distribution (ZTDTPPLD) with parameters γ and β (see [29]), which have the following pmf:

$$f_7(x) = \frac{\gamma^2}{\gamma^2 + 2\gamma\beta + \gamma + \beta} \frac{\beta x + \gamma + \beta + 1}{(\gamma + 1)^x}, \; x = 1, 2, 3 \ldots$$

- Zero-truncated Poisson–Lindley distribution (ZTPLD) with parameters α (see [30]), which has the following pmf:

$$f_8(x) = \frac{\alpha^2(x+\alpha+2)}{(\alpha^2+3\alpha+1)(\alpha+1)^x},\ x = 1,2,3,\ldots$$

- Intervened generalized Poisson distribution (IGPD) with parameters α, β, and γ (see [31]), which has the following pmf:

$$f_9(x) = \frac{\alpha\left\{(1+\gamma)((1+\gamma)\alpha+\beta x)^{x-1} - \gamma(\gamma\alpha+\beta x)^{x-1}\right\}}{e^{\alpha\gamma+\beta x}(e^{\alpha}-1)x!}, x = 1,2,3\ldots$$

8.1. COVID-19 Data Set from Italy

Italy's 61-day COVID-19 data collection, conducted from 13 June to 12 August 2021, is accessible in [11]. Daily newly reported cases are included in this data collection. The descriptive measures of the real data set, which include sample size (n), minimum (min), first quartile (Q_1), median (M_d), third quartile (Q_3), maximum (max), and inter-quartile range (IQR) are given in Table 3.

Table 3. Descriptive statistics for the COVID-19 data set of Italy.

Statistic	n	min	Q_1	M_d	Q_3	max	IQR
Values	61	3	13	21	28	63	15

In addition, Figure 3 shows an empirical TTT plot of the data, from which we deduce an increasing hrf.

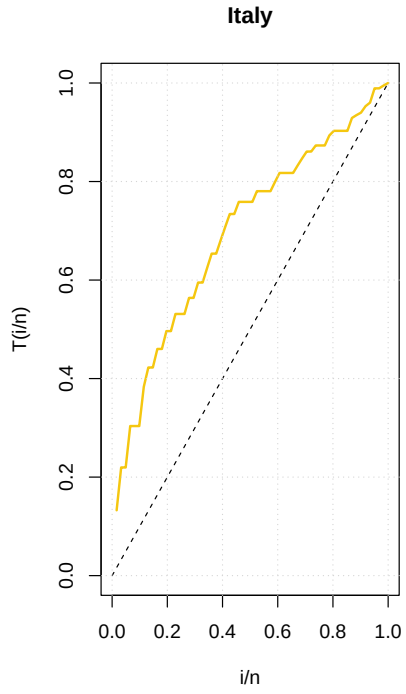


Figure 3. Total Time on Test (TTT) plot for the COVID-19 data set of Italy.

We compare the competitive distributions to the LZTBD using the statistical techniques provided, namely, the negative log-likelihood ($-\log L$), Akaike information criterion (AIC), Bayesian information criterion (BIC), andχ^2 statistic. Table 4 displays the corresponding MLEs, model adequacy measures, and χ^2 values. The LZTBD has lower model adequacy

measures and χ^2 values than the other distributions studied, as shown in Table 4. As a result, the suggested model is the most appropriate for modeling the given COVID-19 data. It is interesting to note that the empirical mean, variance, and IOD of this COVID-19 data set are 22.6229, 160.3388, and 7.0874, respectively, and the theoretical values for the mean, variance, and IOD measures of the LZTBD are 21.6234, 160.3248, and 7.4144, respectively. Thus, the empirical and theoretical means are almost the same, and the empirical and theoretical variances and IOD values are close to each other.

Table 4. Maximum likelihood estimates (MLEs), model adequacy measures and χ^2 values for the COVID-19 data set of Italy.

Model	ZTBD	ZTDTPPLD	ZTPLD	ZTGBD	IGPD	LZTBD
MLE	$\gamma = 63$ $\beta = 0.3587$	$\gamma = 0.0858$ $\alpha = 0.9999$	$\alpha = 0.0859$	$\alpha = 1.0276$ $\beta = 0.8142$ $\gamma = 4.5332$	$\alpha = 6.8936$ $\beta = 0.6298$ $\gamma = 0.2134$	$\alpha = 1.0150$ $\beta = 0.8302$ $\gamma = 3.1778$
$-\log L$	485.9380	239.8677	239.8677	234.5301	234.6429	234.5071
χ^2	12394.92	262.6835	262.6844	203.2570	205.0283	203.2323
df	7	6	7	5	5	4
AIC	973.8760	483.7375	481.7355	475.0602	475.2857	475.0141
BIC	975.9869	487.9572	483.8463	481.3928	481.6184	481.3468

In the case of GLRT, the calculated value based on the test statistic (29) is $2(-234.5071 + 485.9380) = 251.4309$ (p-value $= 0.0001$). As a result, at any level greater than 0.0001, the null hypothesis is rejected in favor of the alternative hypothesis. Hence, we conclude that the additional parameter α in the LZTBD is significant in light of the test procedure outlined in Section 6.

8.2. COVID-19 Data Set from Senegal

The LZTBD is fitted to another set of data for the COVID-19 in Senegal for 56 days of infection, which was recorded from 29 March 2021 to 23 May 2021. These data, which show the daily incidence of COVID-19 cases, were gathered by the World Health Organization (WHO) and are accessible at http://covid19.who.int/data, (accessed on 24 August 2022). Table 5 includes some information as well as descriptive statistics for these data.

Table 5. Descriptive statistics for the COVID-19 data set of Senegal.

Statistic	n	min	Q_1	M_d	Q_3	max	IQR
Values	56	15	30.75	45.50	62.50	107	31.75

In addition, Figure 4 shows an empirical TTT plot of the data from which an increasing hrf is revealed.

We compare the competitive distributions to the suggested distribution using the statistical techniques provided, specifically, the $-\log L$, AIC, BIC, and χ^2 values. Table 6 displays the corresponding MLEs, model adequacy measures, and χ^2 values of the LZTBD. The LZTBD's model adequacy measures and χ^2 values are less than those of the other examined models. As a result, the suggested model is the most appropriate for modeling the COVID-19 data from Senegal. It is worth noting that the empirical mean, variance, and IOD of these COVID-19 datasets are 46.54, 394.326, and 8.47, respectively, and the theoretical values for the mean, variance, and IOD measures of the LZTBD are 46.4, 394.324, and 8.49, respectively. Thus, the empirical and theoretical means are almost the same, and the empirical and theoretical variances and IOD values are very close to each other.

In the case of GLRT, the calculated value based on the test statistic (29) is $2(-244.0212 + 584.8307) = 340.8095$ (p-value $= 0.0004$). As a result, at any level greater than 0.0004, the null hypothesis is rejected in favor of the alternative hypothesis. Hence, we conclude that

the additional parameter α in the LZTBD is significant in light of the test procedure outlined in Section 6.

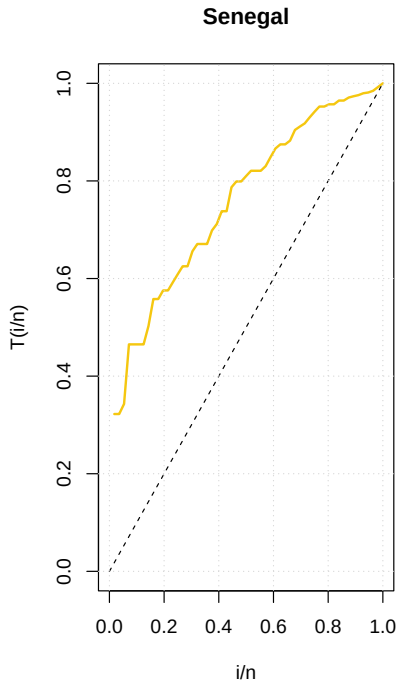


Figure 4. Total Time on Test (TTT) plot for the COVID-19 data set of Senegal.

Table 6. Maximum likelihood estimates (MLEs), model adequacy measures, and χ^2 value for the Senegal data set.

Model	ZTBD	ZTDTPPLD	ZTPLD	ZTGBD	IGPD	LZTBD
MLE	$\gamma = 107$ $\beta = 0.4349$	$\gamma = 0.9999$ $\alpha = 0.0422$	$\alpha = 0.0420$	$\alpha = 1.0000$ $\beta = 0.8838$ $\gamma = 6.1135$	$\alpha = 0.6689$ $\beta = 1.7954$ $\gamma = 5.5010$	$\alpha = 1.000$ $\beta = 0.8838$ $\gamma = 5.1175$
$-\log L$	584.8307	256.4383	256.6182	244.0218	244.2461	244.0212
χ^2	702.1028×10^4	547.6933	549.2564	396.7983	404.4258	396.7792
df	7	6	7	5	5	5
AIC	1171.661	516.8765	515.2363	494.0436	494.4922	494.0424
BIC	1173.687	520.9272	517.2617	500.1196	500.5682	500.1184

8.3. COVID-19 Data Set from Pakistan

The LZTBD is fitted to another set of data for the COVID-19 in Pakistan for 95 days of infection, which was recorded from 23 May 2021 to 25 August 2021. These data, which are available at http://covid19.who.int/data, (accessed on 24 August 2022), were acquired by the WHO and show the daily incidence of COVID-19 cases. Table 7 contains some information and descriptive statistics for these data.

Table 7. Descriptive statistics for the COVID-19 data set of Pakistan.

Statistic	n	min	Q_1	M_d	Q_3	max	IQR
Values	95	11	33	47	73	102	40

In addition, Figure 5 shows an empirical TTT plot of the data, showing an increasing hrf.

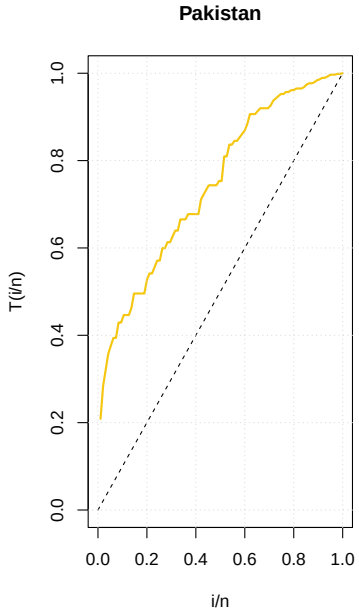


Figure 5. Total Time on Test (TTT) plot for the COVID-19 data set of Pakistan.

Using the statistical methods offered, specifically the $-\log L$, AIC, BIC, and χ^2 values, we compare the competing distributions to the suggested distribution. Table 8 displays the corresponding MLEs, model adequacy measures, and χ^2 values of the LZTBD. The LZTBD's model adequacy measures and χ^2 values are less than those of the other examined models. The suggested model is, therefore, the most suitable one to model the COVID-19 data from Pakistan. In addition, let us mention that the empirical mean, variance, and IOD of this COVID-19 dataset are 52.6842, 538.5801, and 10.2228, respectively, and the theoretical values for the mean, variance, and IOD measures of the LZTBD are 52.6704, 538.5509, and 10.2249, respectively. We thus observe that the empirical and theoretical means are almost equal, and the empirical and theoretical variances and IOD values are very close to each other.

In the case of GLRT, the calculated value based on the test statistic (29) is $2(-431.4451 + 1313.275) = 881.8299$ (p-value $= 0.0001$). As a result, at any level greater than 0.0001, the null hypothesis is rejected in favor of the alternative hypothesis. Hence, we conclude that the additional parameter α in the LZTBD is significant in light of the test procedure outlined in Section 6.

Table 8. Maximum likelihood estimates (MLEs), model adequacy measures, and χ^2 value for the Pakistan data set.

Model	**ZTBD**	**ZTDTPPLD**	**ZTPLD**	**ZTGBD**	**IGPD**	**LZTBD**
MLE	$\gamma = 102$ $\beta = 0.5165$	$\gamma = 0.9999$ $\alpha = 0.0373$	$\alpha = 0.0372$	$\alpha = 1.0000$ $\beta = 0.9107$ $\gamma = 5.1609$	$\alpha = 0.7132$ $\beta = 0.8605$ $\gamma = 8.1173$	$\alpha = 1.0000$ $\beta = 0.9109$ $\gamma = 4.1502$
$-\log L$	1313.275	441.8371	448.0789	431.447	432.5312	431.4451
χ^2	3454.6107×10^4	849.5348	851.525	647.5694	661.2751	647.3886
df	8	7	8	6	6	6
AIC	2628.551	899.6741	898.1577	868.8939	871.0625	868.8902
BIC	2631.104	904.7819	900.7116	876.5556	878.7241	876.5518

8.4. COVID-19 Data Set from Saudi Arabia

The LZTBD is fitted to another set of data of COVID-19 mortality numbers in Saudi Arabia for 83 days of infection, which was recorded from 30 May to 20 August 2020. The WHO gathered these data, which represent the number of deaths per day, and they are available at http://covid19.who.int/data, (accessed on 24 August 2022). Table 9 contains some information and descriptive statistics for these data.

Table 9. Descriptive statistics for the COVID-19 data set of Saudi Arabia.

Statistic	n	min	Q_1	M_d	Q_3	max	IQR
Values	83	17	32	37	41	58	9

In addition, Figure 6 shows an empirical TTT plot of the data and it shows an increasing hrf.

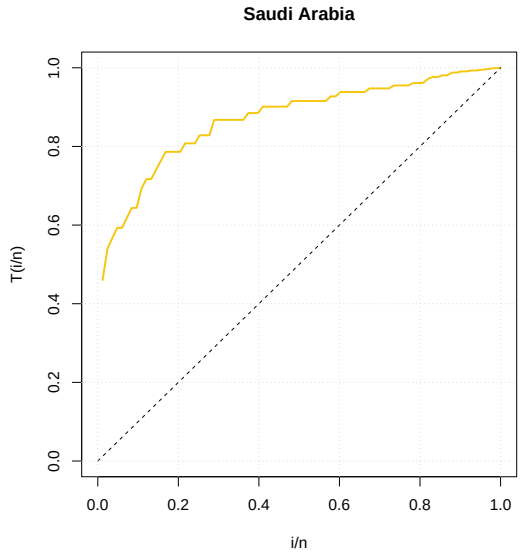


Figure 6. Total Time on Test (TTT) plot for the COVID-19 data set of Saudi Arabia.

We compare the competitive distributions to the suggested distribution using the statistical techniques provided, specifically, the $-\log L$, AIC, BIC, and χ^2 values. Table 10 displays the corresponding MLEs, model adequacy measures, and χ^2 values of the LZTBD. The LZTBD's model adequacy measures and χ^2 values are less than those of the other examined models. For modeling the COVID-19 data from Saudi Arabia, the suggested model is therefore the most suitable. Furthermore, the empirical mean, variance, and the IOD values of this COVID-19 dataset are 36.9277, 70.5313, and 1.9099, respectively, and the theoretical values for the mean, variance, and IOD measures of the LZTBD are 36.8724, 71.0567, and 1.9270, respectively. Hence, the empirical and theoretical means are almost the same, and the empirical and theoretical variances and IOD values are very close to each other.

In the case of GLRT, the calculated value based on the test statistic (29) is $2(-294.288 + 415.1392) = 120.8512$ (p-value $= 0.0002$). As a result, at any level greater than 0.0002, the null hypothesis is rejected in favor of the alternative hypothesis. Hence, we conclude that the additional parameter α in the LZTBD is significant in light of the test procedure outlined in Section 6.

Table 10. Maximum likelihood estimates (MLEs), model adequacy measures, and χ^2 value for the Saudi Arabia data set.

Model	ZTBD	ZTDTPPLD	ZTPLD	ZTGBD	IGPD	LZTBD
MLE	$\gamma = 58$ $\beta = 0.6365$	$\gamma = 0.9999$ $\alpha = 0.0530$	$\alpha = 0.0530$	$\alpha = 1.00008$ $\beta = 0.4774$ $\gamma = 40.4100$	$\alpha = 10.1672$ $\beta = 0.2780$ $\gamma = 1.6222$	$\alpha = 1.000026$ $\beta = 0.4797$ $\gamma = 39.0503$
$-\log L$	415.1392	356.3418	356.3418	294.289	294.3522	294.288
χ^2	2928.725	486.2035	486.2029	150.3025	150.2124	143.2237
df	7	6	7	5	5	4
AIC	832.2784	716.6837	714.6836	594.5770	594.7044	594.5763
BIC	834.6973	721.5213	717.1025	601.834	601.9609	601.8328

8.5. COVID-19 Data Set from Belgium

A different set of data on the COVID-19 infection in Belgium for 425 days (more than a year), which was recorded from 22 July 2021 to 19 September 2022, is fitted using the LZTBD. The WHO gathered these data, which represent the number of deaths per day, and are accessible at http://covid19.who.int/data, (accessed on 24 August 2022). Table 11 contains some information and descriptive statistics for these data.

Table 11. Descriptive statistics for the COVID-19 data set of Belgium.

Statistic	n	min	Q_1	M_d	Q_3	max	IQR
Values	425	1	7	13	24	66	17

In addition, Figure 7 shows an empirical TTT plot from which we can distinguish an increasing hrf.

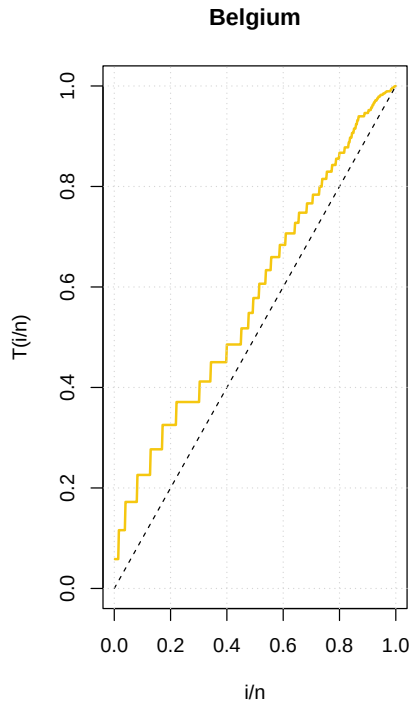


Figure 7. Total Time on Test (TTT) plot for the COVID-19 data set of Belgium.

We compare the competitive distributions to the suggested distribution using the statistical techniques provided, specifically, the -log L, AIC, BIC, and χ^2 values. Table 12 displays the corresponding MLEs, model adequacy measures, and χ^2 values of the LZTBD. The LZTBD's model adequacy measures and χ^2 values are less than those of the other examined models. As a result, the suggested model is the most appropriate for modeling the COVID-19 data from Belgium. It is worth noting that the empirical mean, variance, and IOD of this COVID-19 dataset are 17.122, 178.419, and 10.420, respectively, and the theoretical values for the mean, variance, and IOD measures of the LZTBD are 17.213, 178.412, and 10.365, respectively. Thus, the empirical and theoretical means are almost the same, and the empirical and theoretical variances and IOD values are very close to each other.

Table 12. Maximum likehood estimates (MLEs), model adequacy measures and χ^2 value for the Belgium data set.

Model	ZTBD	ZTDTPPLD	ZTPLD	ZTGBD	IGPD	LZTBD
MLE	$\gamma = 66$ $\beta = 0.2594$	$\gamma = 0.9568$ $\alpha = 0.1128$	$\alpha = 0.1109$	$\alpha = 1.2600$ $\beta = 0.6599$ $\gamma = 4.3279$	$\alpha = 0.7212$ $\beta = 0.5740$ $\gamma = 2.9265$	$\alpha = 1.1254$ $\beta = 0.7483$ $\gamma = 1.6206$
$-\log L$	3825.833	1604.421	1612.689	1602.449	1602.091	1601.074
χ^2	3338.99	1049.849	1049.988	1068.491	1079.062	1049.466
df	12	11	12	10	10	10
AIC	7653.665	3212.841	3227.379	3210.898	3210.181	3208.149
BIC	7657.717	3220.946	3231.431	3223.054	3222.338	3220.466

In the case of GLRT, the calculated value based on the test statistic (29) is $2(-1601.074 + 3825.833) = 2224.759$ (p-value $= 0.0012$). As a result, at any level greater than 0.0012, the null hypothesis is rejected in favor of the alternative hypothesis. Hence, we conclude that the additional parameter α in the LZTBD is significant in light of the test procedure outlined in Section 6.

8.6. COVID-19 Data Set from Ethiopia

The LZTBD is fitted to another set of data on the COVID-19 infection in Ethiopia for 301 days, which was recorded from 25 August 2020 to 21 June 2021. The WHO collected these data, which represent the number of deaths per day, and are accessible at http://covid19.who.int/data, (accessed on 24 August 2022). Table 13 contains some information and descriptive statistics for these data.

Table 13. Descriptive statistics for the COVID-19 data set of Ethiopia.

Statistic	n	min	Q_1	M_d	Q_3	max	IQR
Values	301	1	6	10	15	47	9

As an additional result, Figure 8 shows an empirical TTT plot of the data, where an increasing hrf can be seen.

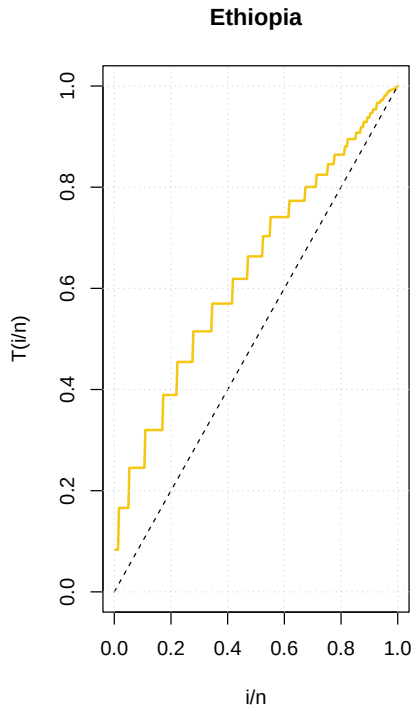


Figure 8. Total Time on Test (TTT) plot for the COVID-19 data set of Ethiopia.

We use the statistical techniques provided to compare the competitive distributions to the suggested distribution, specifically, the $-\log L$, AIC, BIC, and χ^2 values. Table 14 displays the corresponding MLEs, model adequacy measures, and χ^2 values of the LZTBD. The LZTBD's model adequacy measures and χ^2 values are less than those of the other examined models. The suggested model is therefore the most suitable one to model the Ethiopian COVID-19 data. In addition, it is found that the empirical mean, variance, and IOD of this COVID-19 dataset are 11.973, 67.00, and 5.5959, respectively, and the theoretical values for the mean, variance, and IOD measures of the LZTBD are 11.891, 67.02, and 5.6361, respectively. Thus, the empirical and theoretical means are almost equal, and the empirical and theoretical variances and IOD values are very close to each other.

In the case of GLRT, the calculated value based on the test statistic (29) is $2(1003.902 + 1680.734) = 676.832$ (p-value $= 0.0002$). As a result, at any level greater than 0.0002, the null hypothesis is rejected in favor of the alternative hypothesis. Hence, we conclude that the additional parameter α in the LZTBD is significant in light of the test procedure outlined in Section 6.

Table 14. Maximum likelihood estimates (MLEs), model adequacy measures, and χ^2 value for the Ethiopia data set.

Model	ZTBD	ZTDTPPLD	ZTPLD	ZTGBD	IGPD	LZTBD
MLE	$\gamma = 15$ $\beta = 0.2547$	$\gamma = 0.9999$ $\alpha = 0.1610$	$\alpha = 0.1554$	$\alpha = 2.0201$ $\beta = 0.3261$ $\gamma = 12.4252$	$\alpha = 0.5923$ $\beta = 0.4168$ $\gamma = 3.3590$	$\alpha = 1.1256$ $\beta = 0.6681$ $\gamma = 2.8021$
$-\log L$	1680.734	1011.388	1022.044	1004.179	1003.939	1003.902
χ^2	16363.17	326.5928	346.5871	298.3448	298.8601	294.8328
df	10	9	10	8	8	8
AIC	3363.467	2026.776	2046.087	2014.358	2013.878	2013.803
BIC	3367.175	2034.190	2049.794	2025.480	2025.00	2024.925

9. Conclusions

In this article, we used the Lagrange expansion to elaborate a new three-parameter distribution called the Lagrangian zero-truncated binomial distribution (LZTBD). It is worth noting that the proposed distribution is a generalized form of the well-known zero-truncated binomial distribution and the Lagrangian weighted Consul distribution. In particular, we paid close attention to the LZTBD. We investigated the shape properties of the probability mass and hazard functions. The expressions for the factorial moments, generating functions, mean, and median were derived. The identifiability of the LZTBD model was also proved. The LZTBD's model parameters are estimated using the maximum likelihood estimation method. A study employing the simulation technique was also performed to show how well the maximum likelihood estimates are performing. Six actual datasets were used to validate the applicability and demonstrate that the LZTBD offers a superior fit to the competing models.

Author Contributions: Conceptualization, M.R.I., C.C., D.S.S., M.M. and R.M.; methodology, M.R.I., C.C., D.S.S., M.M. and R.M.; software, M.R.I., C.C., D.S.S., M.M. and R.M.; validation, M.R.I, C.C., D.S.S., M.M. and R.M.; formal analysis, M.R.I., C.C., D.S.S., M.M. and R.M.; investigation, M.R.I., C.C., D.S.S., M.M. and R.M.; resources, M.R.I., C.C., D.S.S., M.M. and R.M.; data curation, M.R.I., C.C., D.S.S., M.M. and R.M.; writing—original draft preparation, M.R.I., C.C., D.S.S., M.M. and R.M.; writing—review and editing, M.R.I., C.C., D.S.S., M.M. and R.M.; visualization, M.R.I., C.C., D.S.S., M.M. and R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would also like to thank three reviewers for their thorough comments which led to improvement in the presentation of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Finney, D.; Varley, G. An example of the truncated Poisson distribution. *Biometrics* **1955**, *11*, 387–394. [CrossRef]
2. Creel, M.; Loomis, J. Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California. *Am. J. Agric. Econ.* **1990**, *72*, 434–441. [CrossRef]
3. Brass, W. Simplified methods of fitting the truncated negative binomial distribution. *Biometrika* **1959**, *45*, 59–68. [CrossRef]
4. Lee, A.; Wang, K.; Yau, K.; Somerford, P. Truncated Negative Binomial Mixed Regression Modelling of Ischaemic Stroke Hospitalizations. *Stat. Med.* **2003**, *22*, 1129–1139. [CrossRef]
5. Kennedy. Does Race Predict Stroke Readmission? An Analysis Using the Truncated Negative Binomial Model. *J. Natl. Med. Assoc.* **2005**, *97*, 699–713.
6. Phange, A.; Loh, E. Zero Truncated Strict Arcsine Model. *Int. J. Comput. Electr. Autom. Control. Inf. Eng.* **2013**, *7*, 989–991.
7. Zapata, Z.; Sedory, S.; Singh, S. Zero-truncated Binomial Distribution as a Randomization Device. *Sociol. Methods Res.* **2019**, *51*, 800–815. [CrossRef]
8. Nesteurk, I. Statistics-based predictions of coronavirus epidemic spreading in mainland China. *Innov. Biosyst. Bioeng.* **2020**, *4*, 8–13. [CrossRef]
9. El-Morshedy, M.; Altun, E.; Eliwa, M. A new statistical approach to model the counts of novel coronavirus cases. *Math. Sci.* **2021**, *16*, 37–50. [CrossRef]
10. Ahsan-ul Haq, M.; Babar, A.; Hashmi, S.; Alghamdi, A.S.; Afify, A.Z. The Discrete Type-II Half-Logistic Exponential Distribution with Applications to COVID-19 Data. *Pak. J. Stat. Oper. Res.* **2021**, *7*, 921–932. [CrossRef]
11. Almetwally, E.; Abdo, D.; Hafez, E.; Jawa, T.; Sayed-Ahmed, N.; Almongy, H. The new discrete distribution with application to COVID-19 Data. *Results Phys.* **2022**, *32*, 104987. [CrossRef] [PubMed]
12. Lagrange, J.L. *Mécanique Analytique*; Jacques Gabay: Paris, France, 1788.
13. Consul, P.C.; Shenton, L.R. Use of Lagrange expansion for generating generalized probability distributions. *SIAM J. Appl. Math.* **1972**, *23*, 239–248. [CrossRef]
14. Consul, P.C.; Shenton, L.R. Some interesting properties of Lagrangian distributions. *Commun. Stat.* **1973**, *2*, 263–272. [CrossRef]
15. Mohanty, S.G. On a generalized two-coin tossing problem. *Biom. Z.* **1966**, *8*, 266–272. [CrossRef]

16. Consul, P.C.; Famoye, F. Lagrangian Katz family of distributions. *Commun. Stat. Theory Methods* **1996**, *25*, 415–434. [CrossRef]
17. Berg, K.; Nowicki, K. Statistical inference for a class of modified power series distribution with applications to random mapping theory. *J. Stat. Plan. Inference* **1991**, *28*, 247–261. [CrossRef]
18. Li, S.; Black, D.; Lee, C.; Famoye, F. Dependence Models Arising from the Lagrangian Probability Distributions. *Commun. Stat.—Theory Methods* **2010**, *29*, 1729–1742. [CrossRef]
19. Innocenti, A.R.; Fox, O.; Chibbaro, S. A Lagrangian probability density-function model for collisional turbulent fluid-particle flows. *J. Fluid Mech.* **2019**, *862*, 449–489. [CrossRef]
20. Jensen, J.L. Sur une identité d' Abel et sur d' autres formules analogues. *Acta Math.* **1902**, *26*, 307–318. [CrossRef]
21. Riordan, J. *Combinatorial Identities*; John Wiley and Sons, Inc.: Newyork, NY, USA, 1968.
22. Janardan, K.G.; Rao, B.R. Lagrangian distributions of second kind and weighted distributions. *SIAM J. Appl. Math.* **1983**, *43*, 302–313. [CrossRef]
23. Consul, P.C.; Famoye, F. *Lagrangian Probability Distributions*; Birkhäuser: New York, NY, USA, 2006.
24. McLachlan, G.; Peel, D. *Finite Mixture Models*; Wiley: Hoboken, NJ, USA, 2000.
25. Titterington, D.M.; Smith, A.F.; Markov, U.E. *Statistical Analysis of Finite Mixture Distributions*; Wiley: Hoboken, NJ, USA, 1985.
26. Rao, C.R. Minimum variance and the estimation of several parameters. *Math. Proc. Camb. Philos. Soc.* **1947**, *43*, 280–283. [CrossRef]
27. Ross, S. *Simulation*, 5th ed.; Academic Press: Cambridge, MA, USA, 2013; pp. 5–38. [CrossRef]
28. Aarset, M.V. How to identify a bathtub hazard rate. *IEEE Trans. Reliab.* **1987**, *36*, 106–108. [CrossRef]
29. Shanker, R.; Shukla, K.K. Zero-Truncated Discrete Two-Parameter Poisson-Lindley Distribution with Applications. *J. Inst. Sci. Technol.* **2018**, *22*, 76–85. [CrossRef]
30. Shanker, R.; Fesshaye, H.; Selvaraj, S.; Yemane, A. On zero-truncation of Poisson and Poisson-Lindley distributions and their applications. *Biom. Biostat. Int. J.* **2015**, *2*, 168–181. [CrossRef]
31. Scollnik, D.M. On the Intervened Generalized Poisson Distribution. *Commun. Stat.—Theory Methods* **2006**, *35*, 953–963. [CrossRef]