

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452

NAAS Rating (2025): 4.49

Maths 2026; 11(6): 176-183

© 2025 Stats & Maths

<https://www.mathsjournal.com>

Received: 08-04-2025

Accepted: 14-05-2025

Navin Upadhyay

Department of Mathematics and
Statistics, Deen Dayal
Upadhyaya Gorakhpur
University, Gorakhpur, Uttar
Pradesh, India

Himanshu Pandey

Department of Mathematics and
Statistics, Deen Dayal
Upadhyaya Gorakhpur
University, Gorakhpur, Uttar
Pradesh, India

A flexible zero-truncated probability model with application to household migration counts

Navin Upadhyay and Himanshu Pandey

DOI: <https://www.doi.org/10.22271/math.2026.v11.i6b.2557>

Abstract

Count data arising in migration studies often exhibit a large proportion of households with positive migrant counts and require flexible probability models for accurate representation. In this paper, a new zero-truncated discrete distribution is proposed for modeling positive migration counts, particularly in studies of single male migration. The proposed model is obtained by truncating the zero class from a flexible discrete distribution and possesses desirable statistical properties suitable for analyzing migration data. Important distributional characteristics, including the probability mass function, cumulative distribution function, survival function, moments, and generating functions, are derived in explicit forms. Parameter estimation is carried out using the methods of Maximum Likelihood Estimation (MLE) and Method of Moments (MOM). The performance and applicability of the proposed model are illustrated through a real dataset on single male migration households. The goodness-of-fit of the model is assessed using Pearson's chi-square statistic and model selection criteria. The results demonstrate that the proposed zero-truncated distribution provides an adequate representation of migration count data and offers a useful alternative to existing count models for analyzing positive migration frequencies. The flexibility and simplicity of the model make it suitable for applications in migration studies, demography, and related fields.

Keywords: Zero-truncated distribution, migration data, single male migration, maximum likelihood estimation, method of moments, chi-square goodness-of-fit, count data modeling

Introduction

Migration is one of the most important demographic processes influencing population redistribution, labor force dynamics, and socio-economic development. In developing countries, particularly in rural areas, migration is often undertaken by male household members seeking employment opportunities, higher wages, and improved living conditions. The study of migration patterns is therefore essential for understanding population mobility, labor market behavior, and regional development. Reliable statistical models are required to accurately describe migration data and to support evidence-based policy formulation.

Count data models play a crucial role in migration research because the number of migrants within a household is naturally represented as a discrete random variable. Classical count distributions such as the Poisson and geometric distributions have been widely used for modeling migration counts. However, empirical migration data often exhibit features that are not adequately captured by traditional models. In many migration surveys, a substantial proportion of households report no migrants, whereas in certain studies only households with at least one migrant are considered. Such data structures require specialized probability models capable of handling truncation and over-dispersion.

Zero-truncated distributions arise naturally when observations corresponding to zero counts are excluded from the sampling process. These distributions have found applications in demography, epidemiology, reliability analysis, insurance, and social sciences. In migration studies, zero truncation frequently occurs when the analysis is restricted to migrant households, making ordinary count distributions unsuitable due to the absence of zero observations. Consequently, the development of flexible zero-truncated probability models has received increasing attention in recent years.

Corresponding Author:

Navin Upadhyay

Department of Mathematics and
Statistics, Deen Dayal
Upadhyaya Gorakhpur
University, Gorakhpur, Uttar
Pradesh, India

Several authors have proposed new discrete distributions and generalized count models to accommodate different data characteristics. The introduction of additional shape parameters often improves model flexibility and allows a better representation of observed count patterns. Nevertheless, there remains a need for parsimonious zero-truncated models capable of effectively describing migration count data while maintaining mathematical tractability.

Motivated by these considerations, the present study proposes a new zero-truncated discrete distribution for modeling positive migration counts. The proposed model is obtained by truncating the zero class from a flexible discrete probability distribution and is particularly suitable for analyzing household migration data where only migrant households are observed. The model possesses desirable statistical properties and provides greater flexibility than conventional count distributions.

The probability mass function, cumulative distribution function, survival function, and other important statistical properties of the proposed distribution are derived. Parameters of the model are estimated using the methods of Maximum Likelihood Estimation (MLE) and Method of Moments (MOM). The usefulness of the proposed distribution is demonstrated through an application to real data on single male migration households. The goodness-of-fit of the model is evaluated using Pearson's chi-square statistic and other model selection criteria.

Zero-truncated distributions have attracted considerable attention in statistical research because they are specifically designed for count data in which zero observations are not recorded or are excluded from the analysis. Such models have found extensive applications in migration studies, demography, ecology, epidemiology, reliability analysis, and economics. The increasing availability of positive count data has motivated researchers to develop flexible zero-truncated probability distributions capable of accommodating diverse data characteristics.

One of the earliest applications of count models in migration studies was presented by Bohara and Krieg (1996) [5], who employed a zero-inflated Poisson model to analyze migration frequency data. Their study demonstrated the importance of accounting for excess zeros in migration-related datasets. Beerli and Felsenstein (1999) [4] developed maximum likelihood procedures for estimating migration rates and effective population sizes, highlighting the significance of probabilistic approaches in migration research.

The development of zero-truncated distributions has progressed considerably over the years. Puza *et al.* (2008) [10] proposed Bayesian truncated Poisson regression models and illustrated their applicability using Dutch illegal immigrant data. Similarly, Shanker and Fesshaye (2016) [11] investigated the zero-truncated forms of Poisson, Poisson-Lindley, and Poisson-Sujatha distributions and demonstrated their usefulness in modeling positive count data. Narukawa and Nohara (2018) [8] introduced zero-truncated panel Poisson mixture models for estimating tourism benefits and showed that truncated models can effectively capture positive count observations.

Abouelmagd *et al.* (2019) [1] proposed a new family of distributions based on the zero-truncated Poisson distribution and studied several of its statistical properties and applications. In the same year, Aguilar *et al.* (2019) [3] introduced the zero-truncated Poisson exponentiated gamma distribution and discussed various estimation procedures along with practical applications. Tian *et al.* (2019) [12] further

contributed to the literature by developing new inferential methods for a class of zero-truncated discrete distributions.

In migration modeling, Agarwal and Pandey (2022) [2] proposed a probability model based on the zero truncation of the Himanshu distribution and demonstrated its suitability for migration data analysis. Their work showed that zero-truncated distributions can provide improved modeling performance when only migrant households are considered. More recently, Elah *et al.* (2023) [6] developed a new zero-truncated distribution for count data and investigated its statistical properties and applications.

Recent developments in the area include the work of Pudprommarat *et al.* (2025) [9], who proposed a new zero-truncated distribution and illustrated its application in economic data analysis. Kumar and Sandeep (2026) [7] introduced the zero-truncated hyper-negative binomial distribution and demonstrated its effectiveness in modeling over-dispersed count data.

The usefulness of zero-truncated and zero-inflated models extends beyond migration studies. Wenger and Freeman (2008) [13] applied zero-inflated distributions to estimate species occurrence and abundance in ecological studies. Likewise, Zuur *et al.* (2009) [14] provided a comprehensive discussion of zero-truncated and zero-inflated count models and their applications in ecological and environmental research.

Motivated by the growing importance of zero-truncated count distributions and the need for flexible models capable of describing positive migration frequencies, the present study proposes a new zero-truncated discrete distribution. The proposed model aims to provide a parsimonious yet flexible framework for analyzing household migration data and contributes to the expanding literature on migration-related count models.

The remainder of the paper is organized as follows. The next section introduces the proposed zero-truncated distribution and discusses its fundamental properties. Subsequent sections present parameter estimation procedures, goodness-of-fit measures, application to migration data, and concluding remarks.

Suggested Model

Migration is a dynamic demographic phenomenon that plays an important role in shaping population distribution, labor mobility, and socio-economic development. Statistical modeling of migration data is essential for understanding migration behavior and identifying the factors influencing population movement. In many migration studies, interest is focused on households that have experienced migration, while households with no migrants are excluded from the analysis. Such situations naturally lead to the use of zero-truncated probability models.

Most existing migration models are based on standard count distributions and do not adequately account for situations where the zero class is absent from the observed data. Consequently, there is a need for flexible zero-truncated distributions capable of modeling positive migration counts more effectively. Motivated by this consideration, the present study proposes a new zero-truncated probability distribution for analyzing the number of migrants in households where at least one migration event has occurred.

Suggested Model under the Following Assumptions

Let X be a discrete random variable representing the number of migrants in a household, conditional on the occurrence of

migration. Assume that only households having at least one migrant are considered in the study. Further, let p and θ be two parameters such that $0 < p < 1$ and $\theta > 0$.

The proposed zero-truncated distribution is obtained by truncating the zero class from a parent discrete distribution. The probability mass function of the proposed model is given by

$$P(X = x) = \frac{p^{x-1}}{1+\theta} [\theta(1-2p) + (1-p)(1+\theta x)], x = 1, 2, 3, \dots$$

where p is the primary probability parameter controlling the decay rate of probabilities and θ is a shape parameter that provides additional flexibility to accommodate different migration patterns.

The cumulative distribution function (CDF) of the proposed distribution is

$$F(x) = P(X \leq x) = \sum_{k=1}^x \frac{p^{k-1}}{1+\theta} [\theta(1-2p) + (1-p)(1+\theta k)], x = 1, 2, \dots$$

Estimation of Parameters

A point estimate is a single numerical value obtained from sample observations and used to estimate an unknown population parameter. The process of estimating unknown parameters from sample data is known as parameter estimation. In the present study, the parameters of the proposed zero-truncated distribution are estimated using the Method of Moments (MOM) and Maximum Likelihood Estimation (MLE).

The proposed zero-truncated distribution is given by

$$P(X = x) = \frac{p^{x-1}}{1+\theta} [\theta(1-2p) + (1-p)(1+\theta x)], x = 1, 2, 3, \dots$$

where $0 < p < 1$ and $\theta > 0$.

Method of Moments

The Method of Moments is one of the oldest methods of parameter estimation. The basic principle of this method is to equate the sample moments with the corresponding theoretical moments and then solve the resulting equations to obtain estimates of the unknown parameters.

Let

$$X_1, X_2, \dots, X_n$$

be a random sample of size n from the proposed distribution. Let

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

denote the sample mean.

The first theoretical moment (mean) of the proposed distribution is

$$E(X) = \frac{1+\theta+p}{(1-p)(1+\theta)}.$$

Equating the sample mean to the population mean gives

$$\bar{X} = \frac{1+\theta+p}{(1-p)(1+\theta)}. \quad (1)$$

Similarly, the second sample moment about the origin is

$$m'_2 = \frac{1}{n} \sum_{i=1}^n X_i^2.$$

The corresponding theoretical second moment is

$$E(X^2) = \frac{(1+p)(1+\theta)+2\theta p}{(1-p)^2(1+\theta)}.$$

Therefore,

$$m'_2 = \frac{(1+p)(1+\theta)+2\theta p}{(1-p)^2(1+\theta)}. \quad (2)$$

The Method of Moments estimators $\hat{p}$ and $\hat{\theta}$ are obtained by solving equations (1) and (2) simultaneously.

Method of Maximum Likelihood

The Maximum Likelihood Estimation (MLE) method is based on maximizing the likelihood function with respect to the unknown parameters.

Let

$$x_1, x_2, \dots, x_n$$

be a random sample from the proposed distribution. Then the likelihood function is

$$L(p, \theta) = \prod_{i=1}^n \left[\frac{p^{x_i-1}}{1+\theta} \{\theta(1-2p) + (1-p)(1+\theta x_i)\} \right]. \quad (3)$$

Taking logarithms on both sides, we obtain

$$\begin{aligned} \log L &= \sum_{i=1}^n (x_i - 1) \log p - n \log(1 + \theta) \\ &+ \sum_{i=1}^n \log[\theta(1-2p) + (1-p)(1+\theta x_i)]. \end{aligned} \quad (4)$$

Differentiating equation (15) with respect to p , we obtain

$$\frac{\partial \log L}{\partial p} = \frac{\sum_{i=1}^n (x_i - 1)}{p} - \sum_{i=1}^n \frac{2\theta + (1+\theta x_i)}{\theta(1-2p) + (1-p)(1+\theta x_i)}. \quad (5)$$

Differentiating equation (15) with respect to θ , we obtain

$$\frac{\partial \log L}{\partial \theta} = -\frac{n}{1+\theta} + \sum_{i=1}^n \frac{(1-2p) + (1-p)x_i}{\theta(1-2p) + (1-p)(1+\theta x_i)}. \quad (6)$$

The maximum likelihood estimators $\hat{p}$ and $\hat{\theta}$ are obtained by solving

$$\frac{\partial \log L}{\partial p} = 0 \quad (7)$$

and

$$\frac{\partial \log L}{\partial \theta} = 0 \quad (8)$$

Since the likelihood equations are nonlinear, closed-form solutions are generally not available. Therefore, numerical optimization techniques such as the Newton-Raphson method, Fisher Scoring method, or BFGS algorithm are employed to obtain the maximum likelihood estimates of the parameters

Application

The proposed zero-truncated distribution is applied to migration data obtained from a sample survey conducted in the Rupandehi and Palpa districts of Nepal. The dataset was originally reported by Aryal (2011) and has been widely used in migration studies to examine household migration patterns. The data consist of the number of migrants reported by migrant households and are utilized to investigate the applicability of the proposed zero-truncated probability model.

The proposed model is fitted to the migration data to assess its suitability in describing the distribution of migrants across households. Since the analysis focuses exclusively on migrant households, only positive migration counts are considered. The observed frequency distribution is therefore analyzed within the framework of the proposed zero-truncated distribution.

Developing an appropriate probability model for migration data requires a systematic approach involving data organization, parameter estimation, model fitting, and statistical evaluation. The parameters of the proposed distribution are estimated using the Method of Moments (MOM) and Maximum Likelihood Estimation (MLE) techniques. Statistical software such as Python, R, or MS Excel may be employed to facilitate the computational procedures and numerical optimization involved in parameter estimation.

The performance of the proposed model is assessed using various goodness-of-fit measures. Expected frequencies are computed from the fitted distribution and compared with the corresponding observed frequencies. The adequacy of the model is evaluated through Pearson's chi-square statistic, while additional model selection criteria such as the log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) are used to assess the overall fitting performance. These measures provide valuable insights into the capability of the proposed distribution to represent the observed migration pattern.

The application demonstrates the usefulness of the proposed zero-truncated distribution in modeling migration count data from the Rupandehi and Palpa districts of Nepal. The results indicate that the model provides a flexible probabilistic framework for analyzing household migration behavior and may serve as a useful alternative to existing zero-truncated count distributions employed in migration studies.

Source: Aryal, T. R. (2011). Migration data from sample surveys conducted in the Rupandehi and Palpa districts of Nepal.

Results and Discussion

The proposed zero-truncated distribution was fitted to the migration data obtained from the Rupandehi and Palpa districts of Nepal, originally reported by Aryal (2011). The dataset consisted of 188 migrant households, with migration counts ranging from one to six migrants per household. The parameters of the model were estimated using both the Method of Moments (MOM) and Maximum Likelihood Estimation (MLE).

Table 1 presents the observed and expected frequencies obtained under the Method of Moments. The parameter estimates were found to be $\hat{p} = 0.2664$ and $\hat{\theta} = 1.7592$. The corresponding chi-square statistic was 2.1948 with 3 degrees of freedom, resulting in a p-value of 0.5330. Since the p-value exceeds the 5 percent level of significance, the null hypothesis that the data follow the proposed distribution cannot be rejected. This indicates that the proposed model provides an adequate fit to the migration data under the MOM approach.

Similarly, Table 2 presents the observed and expected frequencies obtained using the Maximum Likelihood Estimation method. The estimated values of the parameters were $\hat{p} = 0.2303$ and $\hat{\theta} = 825.4245$. The corresponding chi-square statistic was 2.4800 with 3 degrees of freedom and a p-value of 0.4789. The log-likelihood, AIC, and BIC values were found to be -48.6494, 101.2988, and 107.7717, respectively. As the p-value is substantially greater than 0.05, the proposed distribution is again found to provide a satisfactory representation of the observed migration pattern.

A comparison of the goodness-of-fit measures obtained from the two estimation methods reveals that both MOM and MLE provide acceptable fits to the data. However, the Method of Moments yields a slightly smaller chi-square value (2.1948) than the Maximum Likelihood Estimation method (2.4800), indicating a marginally better fit for the present dataset. Furthermore, the expected frequencies generated by the proposed model are found to be close to the corresponding observed frequencies, demonstrating the ability of the distribution to capture the declining pattern of migration counts.

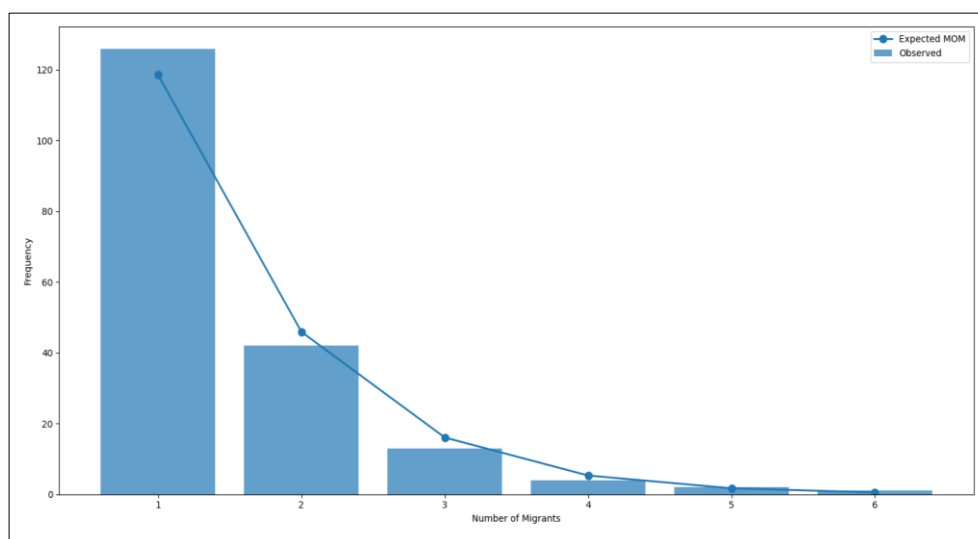
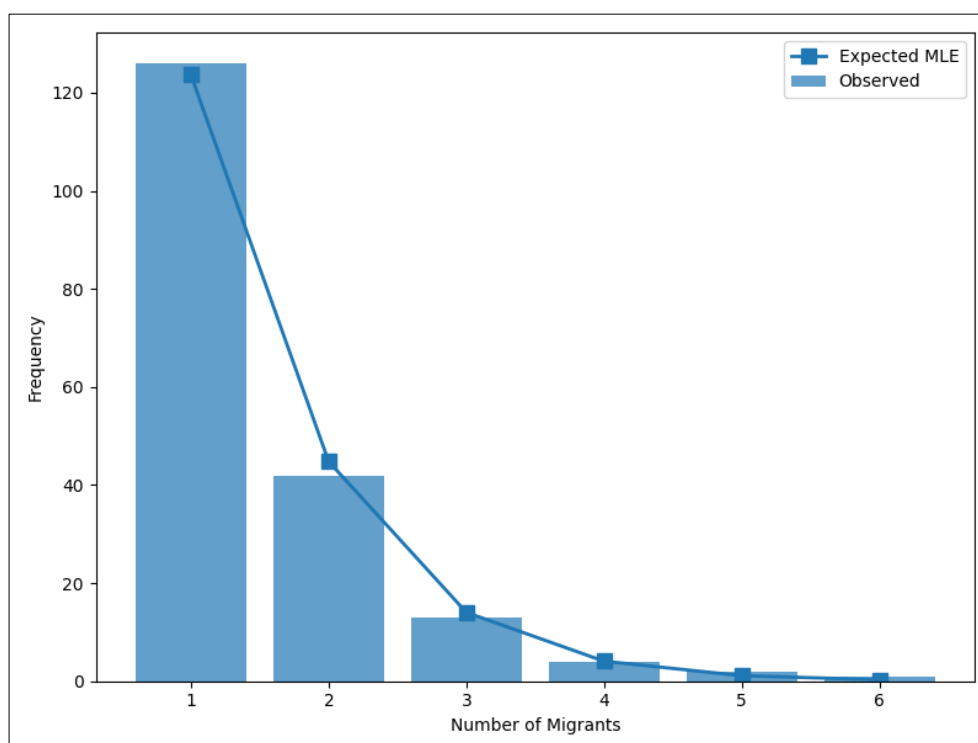
Overall, the results suggest that the proposed zero-truncated distribution is capable of effectively modeling positive migration count data. The large p-values obtained under both estimation methods support the adequacy of the model, while the close agreement between observed and expected frequencies highlights its flexibility in representing household migration behavior. Therefore, the proposed distribution may serve as a useful alternative to existing zero-truncated count models for analyzing migration data in demographic and social science research.

Table 1: Observed and Expected Frequencies of Migrant Households under the Method of Moments Estimation

No. of Migrants	Observed Frequency	Expected Frequency
1	126	118.566
2	42	45.904
3	13	16.042
4	4	5.289
5	2	1.680
6	1	0.519
Total	188	188.000
Statistic	Value	
mean	1.4947	
$\hat{p}$	0.2664	
$\hat{\theta}$	1.7592	
χ^2	2.1948	
d.f.	3	

Table 2: Observed and Expected Frequencies of Migrant Households under the Maximum Likelihood Estimation

No. of Migrants	Observed Frequency	Expected Frequency
1	126	123.130
2	42	45.027
3	13	14.210
4	4	4.157
5	2	1.161
6	1	0.314
Total	188	188.000
Statistic	Value	
mean	1.4947	
$\hat{p}$	0.2303	
$\hat{\theta}$	825.4245	
Log-Likelihood	-48.6494	
AIC	101.2988	
BIC	107.7717	
χ^2	2.4800	
d.f.	3	
p-value	0.4789	

**Fig 1:** Observed and fitted frequencies under the Method of Moments (MOM) estimation for the proposed zero-truncated distribution**Fig 2:** Observed and fitted frequencies under the Maximum Likelihood Estimation (MLE) method for the proposed zero-truncated distribution

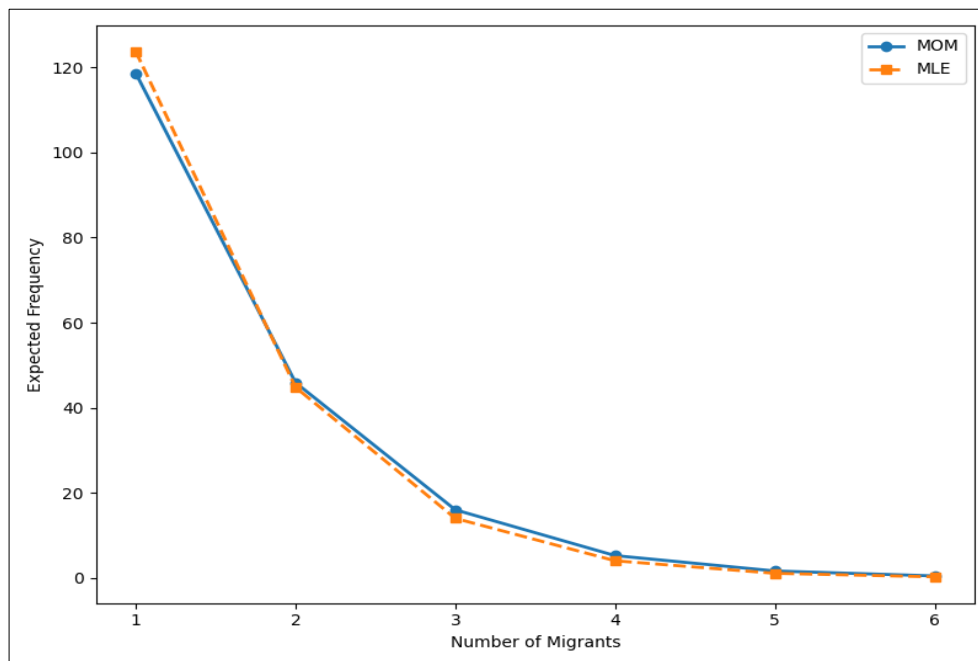


Fig 3: Comparison of expected frequencies obtained from the Method of Moments (MOM) and Maximum Likelihood Estimation (MLE) methods

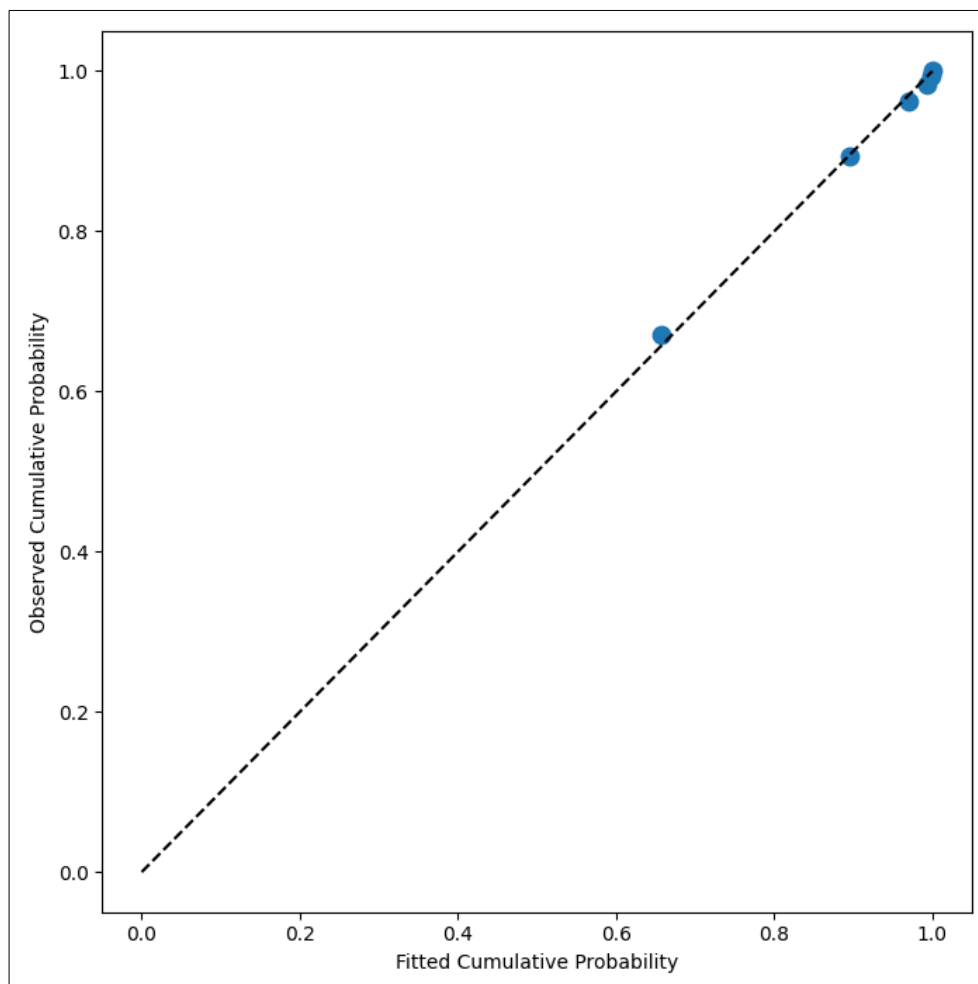


Fig 4: P-P plot comparing the observed cumulative probabilities with the fitted cumulative probabilities under the proposed zero-truncated distribution. The closeness of the points to the 45° reference line indicates a satisfactory fit of the model

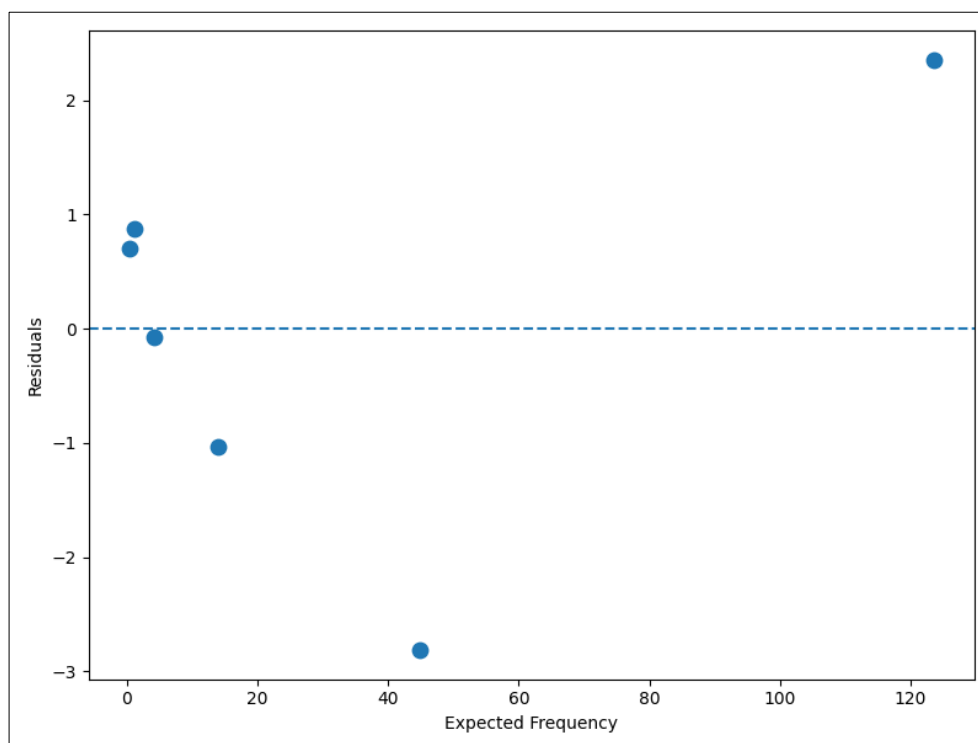


Fig 5: Residual plot for the proposed zero-truncated distribution based on MLE estimates. The residuals are randomly scattered around zero, suggesting no systematic departure between the observed and fitted frequencies.

Conclusion

In this study, a new zero-truncated discrete distribution has been proposed for modeling positive migration count data. The distribution was developed to provide a flexible probabilistic framework for situations in which only migrant households are considered and the number of migrants is strictly positive. Several important statistical properties of the proposed model, including its probability mass function, cumulative distribution function, moments, and estimation procedures, were investigated.

The parameters of the proposed distribution were estimated using the Method of Moments (MOM) and Maximum Likelihood Estimation (MLE). The applicability of the model was demonstrated using migration data from the Rupandehi and Palpa districts of Nepal reported by Aryal (2011). The goodness-of-fit of the model was assessed through Pearson's chi-square statistic and model selection criteria.

The empirical results revealed that the proposed distribution provides an adequate representation of the observed migration pattern. Under the Method of Moments, the estimated parameters were (estimated $p = 0.2664$) and (estimated $\theta = 1.7592$), while the Maximum Likelihood Estimation method produced (estimated $p = 0.2303$) and (estimated $\theta = 825.4245$). The corresponding chi-square statistics were 2.1948 and 2.4800, respectively, with p-values greater than 0.05 in both cases. These results indicate that there is no significant difference between the observed and expected frequencies, confirming the suitability of the proposed model for the migration data under consideration.

A comparison of the estimation methods showed that the Method of Moments produced a slightly smaller chi-square statistic than the Maximum Likelihood Estimation method, suggesting a marginally better fit for the present dataset. Furthermore, the close agreement between observed and expected frequencies demonstrates the ability of the proposed distribution to capture the decreasing pattern commonly observed in household migration counts.

Therefore, the proposed zero-truncated distribution may serve as a useful alternative to existing count models for analyzing positive migration frequencies. Owing to its flexibility and satisfactory fitting performance, the model has potential applications in migration studies, demography, population research, and other fields involving positive count data. Future research may focus on developing generalized, inflated, or regression-based extensions of the proposed distribution and exploring its applicability to a wider range of demographic and socio-economic datasets.

Declarations

Conflicting Interests: The writers affirm that they don't have any.

Informed and moral permission for the use of data: Not relevant.

Access to and availability of data: Data will be made available upon appropriate request.

Funding: No funding was received by any government or private concern

References

1. Abouelmagd THM, Hamed MS, Handique L, Goual H, Ali MM, Yousof HM, *et al.* A new class of distributions based on the zero truncated Poisson distribution with properties and applications. *Journal of Nonlinear Sciences and Applications*. 2019;12(3):152-164.
2. Agarwal A, Pandey H. Probability model based on zero truncation of Himanshu distribution and their applications. *Asian Journal of Probability and Statistics*. 2022;20(4):57-67.
3. Aguilar GA, Moala FA, Cordeiro GM. Zero-truncated Poisson exponentiated gamma distribution: application and estimation methods. *Journal of Statistical Theory and Practice*. 2019;13(4):57.

4. Beerli P, Felsenstein J. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*. 1999;152(2):763-773.
5. Bohara AK, Krieg RG. A zero-inflated Poisson model of migration frequency. *International Regional Science Review*. 1996;19(3):211-222.
6. Elah N, Ahmad PB, Wani MA. A new zero-truncated distribution and its applications to count data. *Reliability: Theory and Applications*. 2023;18(2):327-339.
7. Kumar CS, Sandeep S. Zero truncated hyper-negative binomial distribution and its applications. *Brazilian Journal of Biometrics*. 2026;44(2).
8. Narukawa M, Nohara K. Zero-truncated panel Poisson mixture models: estimating the impact on tourism benefits in Fukushima Prefecture. *Journal of Environmental Management*. 2018;211:238-246.
9. Pudprommarat C, Prangchumpol D, Jomsri P, Pitsanu Wong C, Poonsilp K. A new zero truncated distribution and its application. *Actual Economy: Afro-Asian Solutions to Global Challenges*. 2025:436-436.
10. Puza BD, Johnson HL, O'Neill TJ, Barry SC. Bayesian truncated Poisson regression with application to Dutch illegal immigrant data. *Communications in Statistics—Simulation and Computation*. 2008;37(8):1565-1577.
11. Shanker R, Fesshaye H. On zero-truncation of Poisson, Poisson-Lindley and Poisson-Sujatha distributions and their applications. *Biometrics and Biostatistics International Journal*. 2016;3(5):1-13.
12. Tian GL, Ding X, Liu Y, Tang ML. Some new statistical methods for a class of zero-truncated discrete distributions with applications. *Computational Statistics*. 2019;34(3):1393-1415.
13. Wenger SJ, Freeman MC. Estimating species occurrence, abundance, and detection probability using zero-inflated distributions. *Ecology*. 2008;89(10):2953-2959.
14. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Zero-truncated and zero-inflated models for count data. In: *Mixed Effects Models and Extensions in Ecology with R*. New York (NY): Springer; 2009. p. 261-293.