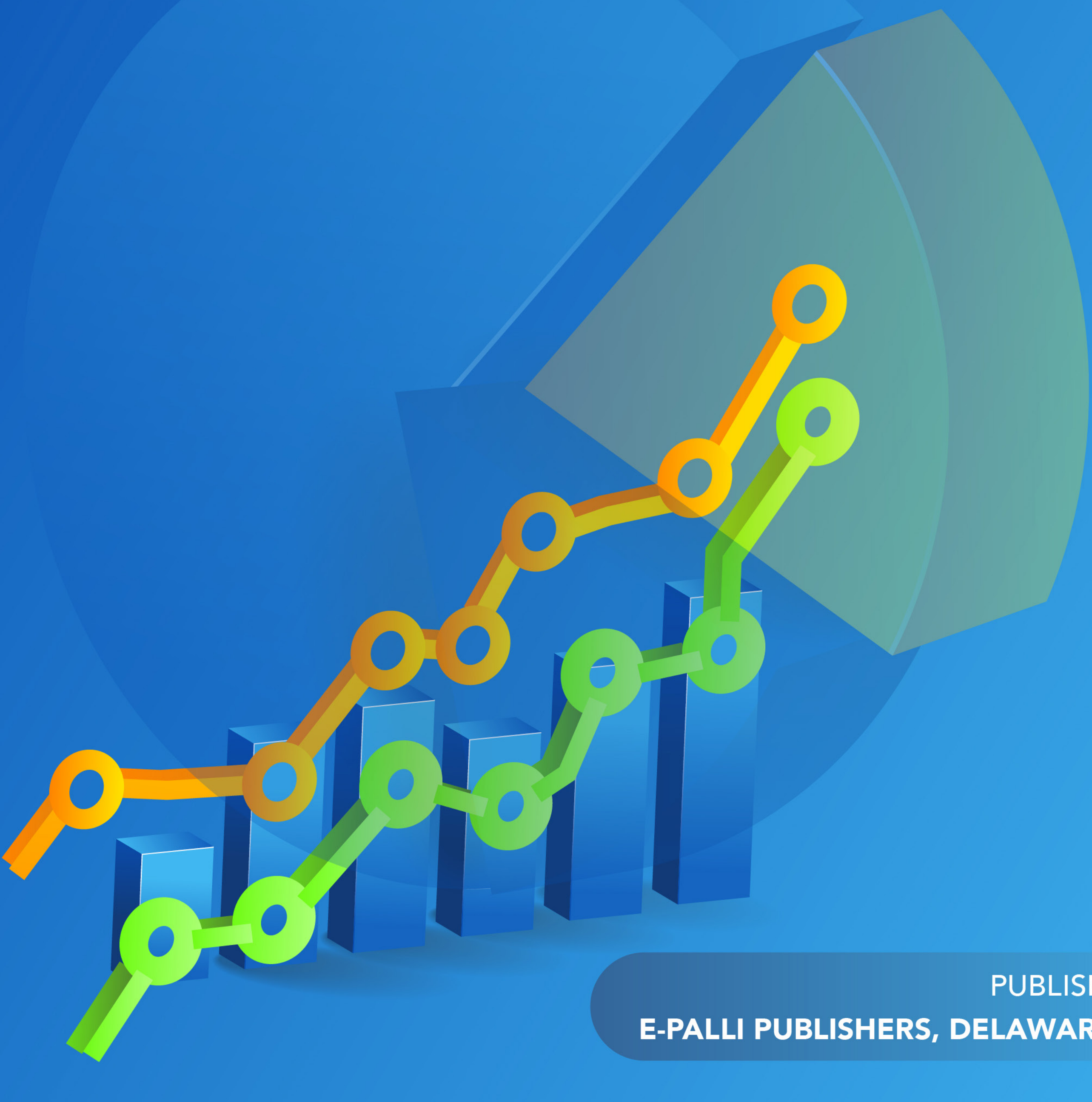
American Journal of Applied Statistics and Economics (AJASE)

ISSN: 2992-927X (ONLINE)

VOLUME 4 ISSUE 1 (2025)



PUBLISHED BY

E-PALLI PUBLISHERS, DELAWARE, USA

American Journal of Applied Statistics and Economics (AJASE)

Volume 4 Issue 1, Year 2025
ISSN: 2992-927X (Online)
DOI: https://doi.org/10.54536/ajase.v4i1.4235
https://journals.e-palli.com/home/index.php/ajase

Advancing Statistical Modelling: A Comparative Study of Zero-Truncated Distributions in Economic Analysis

B. E. Omokaro[1], C. O. Aronu[2*]

Article Information

Received: December 19, 2024

Accepted: January 21, 2025

Published: July 03, 2025

Keywords

AIC, Economic Analysis, Geometric-Zero Truncated Poisson, Maximum Likelihood Estimation, Model Fit, MSE, Zero-Truncated Poisson Pareto

ABSTRACT

This study explores the application of two zero-truncated distributions, Geometric-Zero Truncated Poisson (GZTP) and Zero-Truncated Poisson Pareto (ZTPP), in modelling economic datasets, with a particular focus on Nigeria's key economic indicators. Secondary data from the Central Bank of Nigeria's Statistical Bulletin (2021), spanning from 1989 to 2020, was used, covering variables such as Real Gross Domestic Product (RGDP), export and import goods, money supply, and Brent crude oil prices. The objectives were to: Introduce and describe the mathematical properties of the GZTP and ZTPP distributions; compare the performance of these distributions across datasets using metrics such as AIC, BIC, and MSE; and recommend the most efficient distribution for modelling economic variables and predicting trends. The study employs the Maximum Likelihood Estimation (MLE) method for parameter estimation, implemented in R programming. Model performance was evaluated using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Mean Squared Error (MSE). The results indicate that the ZTPP distribution outperforms the GZTP distribution across all datasets, with significantly lower AIC, BIC, and MSE values. Specifically, the ZTPP achieved an average AIC of -1422.54, BIC of -1422.55, and MSE of 30,161.94, compared to the GZTP's -766.39, -762.82, and 31,163.2, respectively. These findings highlight the superior model fit and predictive accuracy of the ZTPP distribution, making it a more robust tool for economic analysis, particularly in cases where zero occurrences are impossible. This comparative study underscores the importance of choosing the right distribution for economic modelling to achieve high accuracy and reliable results.

INTRODUCTION

Considering the growing significance of zero-truncated distributions in econometrics and economic forecasting. The Zero-truncated distributions, including the Zero-Truncated Poisson (ZTP), Gamma Zero-Truncated Poisson (GZTP), and Zero-Truncated Poisson Pareto (ZTPP) distributions, have garnered attention for their ability to model count data that exclude zero observations, a common occurrence in economic datasets. Recent advancements in these models, such as the work by Niyomdecha and Srisuradetchai (2023), who introduced the complementary gamma zero-truncated Poisson (CGZTP) distribution, have demonstrated the effectiveness of these distributions in modelling lifetime and economic data. The research by Niyomdecha *et al.* (2023) proposed the gamma zero-truncated Poisson (GZTP) distribution, combining gamma and zero-truncated Poisson distributions using the minimum function. Their study explored the distribution's characteristics, including hazard function and MLE estimation, with simulation tests confirming its adaptability in modelling lifetime data. Ngamkham and Panta (2023) addressed the estimation challenges of the Zero-Truncated Poisson (ZTP) distribution, proposing a delta method for parameter estimation, which was applied to a real dataset on unrest events in southern Thailand. Similarly, the development of the ZTPP distribution by Badr *et al.* (2023) has provided superior fit metrics, including Akaike Information Criteria (AIC) and Bayesian Information Criterion (BIC), when applied to economic datasets. The work by Panichkitkosolkul (2023) proposed the zero-truncated Poisson-Ishita distribution and evaluated various bootstrap methods for estimating confidence intervals, concluding that the simple bootstrap approach was most efficient for larger sample sizes. Irshad *et al.* (2023) developed the Lagrangian Intervened Poisson Distribution (LIPD), a generalized approach for over-dispersed and under-dispersed datasets, and showcased its application using MLE and simulations. Akdogan *et al.* (2019) introduced the geometric-zero truncated Poisson (GZTP) distribution, demonstrating its usefulness for discrete systems with increasing hazard rates. Shukla *et al.* (2020) adapted the Poisson-Ishita distribution to the zero-truncated Poisson-Ishita distribution (ZTPID), showing its superior fit in count data without zero values. Agarwal and Pandey (2024) introduced the inflated zero-truncated Poisson Ailamujia distribution (IMZTPAD), which demonstrated a better fit in modelling child mortality and genetic count data. Panichkitkosolkul (2024) further investigated the zero-truncated Poisson-Lindley (ZTPL) distribution and found that non-parametric bootstrap methods performed best for larger sample sizes. Pankaj *et al.* (2023) examined system reliability using a dual repair technique and regenerative methodologies, providing insights into improving system reliability. Ghosh *et al.* (2023) introduced a bivariate geometric distribution for negatively correlated count data, and Abbas (2023) proposed a bivariate generalized geometric distribution

[1] Department of Statistics, Delta State Polytechnic, Otefe, Oghara, Delta State, Nigeria
[2] Department of Statistics, Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State, Nigeria
[*] Corresponding author's e-mail: amaro4baya@yahoo.com

(BGGD) for correlated count data, utilizing a Bayesian approach for analysis. Shang *et al.* (2023) presented a novel method for predicting aero-engine coaxiality using geometric distribution error modelling and deep learning, achieving high prediction accuracy.

These advancements underline the potential of zero-truncated models to enhance the precision and efficiency of economic analysis, particularly when dealing with data that exhibits truncation at zero, such as income, expenditure, or event counts. Despite these advancements, a notable gap remains in the comprehensive comparison of these zero-truncated distributions in the context of economic analysis. While individual studies have demonstrated the utility of these distributions, there is a lack of systematic evaluation across various economic datasets, with limited attention to performance metrics like AIC, BIC, and Mean Squared Error (MSE) in this domain. This study aims to fill this gap by introducing and comparing the mathematical properties of the GZTP and ZTPP distributions, alongside their performance in economic modelling. By applying these distributions to real-world economic datasets, the study seeks to identify the most efficient distribution for modelling economic variables and predicting trends. Through this comparison, the study will provide valuable insights into the suitability of these distributions for addressing specific economic modelling challenges, thereby contributing to the advancement of statistical methods in economic analysis. Hence, the aim of this study was to advance the understanding of zero-truncated statistical distributions and their application in economic data modelling. The specific objectives are to: Introduce and describe the mathematical properties of the GZTP and ZTPP distributions; compare the performance of these distributions across datasets using metrics such as AIC, BIC, and MSE; and recommend the most efficient distribution for modelling economic variables and predicting trends.

MATERIALS AND METHODS

Source of Data collection for the study

The study utilized secondary data, sourced from reliable publications such as the Central Bank of Nigeria's Statistical Bulletin for 2021, included key economic indicators from 1989 to 2020. These indicators were Real Gross Domestic Product (RGDP), export and import goods, money supply, and Brent crude oil prices. This data set provided crucial insights into Nigeria's economic performance, trade, monetary policy, and the impact of oil prices, forming the basis for detailed econometric analysis.

MATERIALS AND METHODS

Table 1 presents the Probability Mass Functions (PMFs) of two advanced distributions the Geometric-Zero Truncated Poisson (GZTP) and Zero-Truncated Poisson Pareto (ZTPP)

Table 1: The PMF of the GZTP and ZTPP distribution

S/No.	Distribution	PMF	Source
1.	GZTP distribution	$e^{-\lambda}((1-p)^x-(1-p)^{x-1}),/(1-e^{-\lambda}),\lambda>0,0\leq p\leq 1,x\in\{1,2,3,\ldots\}$	Niyomdecha *et al.* (2023)
2	ZTPP distribution	$(\lambda^n e^{-\lambda})/n!(1-e^{-\lambda})$ $,\lambda>0,n\in\{0,1,2,3,\ldots\}$	Badr *et al.* (2023)

The Geometric-Zero Truncated Poisson (GZTP) and Zero-Truncated Poisson Pareto (ZTPP) distributions in Table 1 exhibit notable flexibility and unique properties, making them suitable for modelling diverse real-world phenomena. The GZTP distribution incorporates a geometric component through the parameter p, allowing it to capture overdispersion and varying probabilities for successive events, which is crucial in applications with decaying probabilities. Its range of p between 0 and 1 further enhances its adaptability to different datasets. On the other hand, the ZTPP distribution, characterized by its Poisson-based structure truncated at zero, is particularly effective in handling count data where zero occurrences are impossible. Its dependency solely on the rate parameter λ simplifies its application while maintaining robustness in capturing event frequencies. Both distributions are zero-truncated, addressing scenarios where non-zero occurrences are mandatory, and their closed-form PMFs enable straightforward parameter estimation and interpretation. These properties underscore their utility in fields such as ecology, reliability analysis, and actuarial science.

Parameter Estimation

Parameter estimates for each distribution were derived using the Maximum Likelihood Estimation (MLE) method, which maximizes the likelihood function L(θ|x) given by:

$$L(\theta \mid x) = \prod_{i=1}^{n} f(x_i;\theta) \qquad (1)$$

where $f(x_i;\theta)$ represents the Probability Mass Function (PMF) of the distribution, θ is the vector of parameters, and x_i are the observed data points (Casella & Berger, 2002). MLE implementation was performed in the R programming language (R Core Team, 2023).

Model Performance Measures of the distributions

The model performance was evaluated using the following criteria:

i. Akaike Information Criterion (AIC):

$$AIC=-2\ln(L)+2k \qquad (2)$$

where L Where likelihood of the model, and k is the number of estimated parameters (Akaike, 1974).

ii. Bayesian Information Criterion (BIC):

$$BIC=-2\ln(L)+k\ln(n) \qquad (3)$$

where n is the sample size (Schwarz, 1978).

iii. Mean Squared Error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (4)$$

where y_i represents the observed values, and yˆ_i are the predicted values.

These metrics were computed for both the GZTP and ZTPP distributions across datasets, as shown in Table 3. To ensure a comprehensive evaluation, the average values of AIC, BIC, and MSE for each distribution were also computed and summarized in Table 4.

Comparative Analysis

The comparative analysis focused on the ability of the GZTP and ZTPP distributions to minimize AIC, BIC, and MSE. The ZTPP distribution demonstrated superior performance across the datasets, as evidenced by its consistently lower average AIC, BIC, and MSE values compared to the GZTP distribution. This suggests that the ZTPP distribution provides a better fit for the economic datasets under consideration.

RESULTS AND DISCUSSIONS

Table 2 provides an overview of the descriptive statistics for key economic variables employed in the study, highlighting their central tendencies, variability, and distributional characteristics. The Real Gross Domestic Product (RGDP) exhibits a mean of 365 units with a standard deviation of 99, indicating moderate variability around the average, and ranges from 237 to 569 units, with a median of 330. The skewness of 0.72 suggests a slight rightward skew, while the negative kurtosis (-0.72) indicates a relatively flat distribution compared to the normal curve. Exported goods have a mean value of 40 units and a standard deviation of 30, with values spanning from 10 to 99 units. Its skewness (0.76) and kurtosis (-0.78) reflect a similar pattern to RGDP, indicating asymmetry and platykurtic tendencies. Import goods average 26 units with a standard deviation of 19, ranging from 6 to 66 units, showing a slightly higher skewness (0.77) and comparable kurtosis (-0.78). The Money Supply (M2) variable, measured in millions, has a mean of 59,694,188 and exhibits substantial variability (standard deviation of 20,216,127), ranging from 32,262,332 to 109,951,956. Its skewness (0.57) and kurtosis (-0.36) indicate a mild rightward skew and a distribution closer to normal. Lastly, Brent Crude (BRT) prices show the highest variability, with a mean of 47, a standard deviation of 32, and a range from 11 to 133. Its skewness (0.89) highlights a more pronounced rightward skew, while the kurtosis (-0.32) remains slightly platykurtic. These statistics collectively reveal distinct patterns and variability across the variables, providing insights into their economic implications and informing further analyses.

Table 2: Descriptive Statistics of dataset

Variable	Mean	St. Dev	Minimum	Median	Maximum	Skewness	Kurtosis
RGDP	365	99	237	330	569	0.72	-0.72
Export_Goods	40	30	10	31	99	0.76	-0.78
Import_Goods	26	19	6	20	66	0.77	-0.78
Money_Supply(M2)	59694188	20216127	32262332	56782642	109951956	0.57	-0.36
Brent Crude (BRT)	47	32	11	35	133	0.89	-0.32

The result presented in Table 3 summarizes the performance metrics of the GZTP and ZTPP distributions across various economic datasets, providing insights into their parameter estimates, model fit, and predictive accuracy. For RGDP, the GZTP distribution achieved a lower AIC (-2258.62) and BIC (-2255.87) compared to the ZTPP (-2150.62 and -2154.22, respectively), but the ZTPP showed a marginally lower Mean Squared Error (MSE) of 143,807.4 versus 148,809.8 for the GZTP. Similarly,

Table 3: Performance Comparison of GZTP and ZTPP Distributions across Economic Datasets

Dataset	Distributions	Parameter estimates	AIC	BIC	MSE
RGDP	GZTP	λ=0.1000,p=1.0000	-2258.61900	-2255.8700	148809.8000
	ZTPP	λ=1.0000	-2150.6200	-2154.2200	143807.4000
Export_Goods	GZTP	λ=1.0000,p=0.1000	-158.61900	-155.8170	2488.2640
	ZTPP	λ=33.5437	-791.1342	-791.733	2487.7510
Import_Goods	GZTP	λ=1.0000,p=0.1000	-158.6300	-155.8170	1009.2800
	ZTPP	λ=25.4986	-541.2217	-543.8205	1009.6800
Log(Money_ Supply)	GZTP	λ=1.0000,p=0.1000	-167.8410	-164.273	318.6976
	ZTPP	λ=17.8489	-202.3526	-198.5684	315.3769
BRT	GZTP	λ=1.0000,p=0.1000	-1088.2200	-1082.3400	3189.9470
	ZTPP	λ=38.1595	-3427.3550	-3424.4130	3189.4930

for Export_Goods, ZTPP demonstrated superior fit metrics (AIC = -791.13, BIC = -791.73) and a slightly better MSE (2487.75) than GZTP (AIC = -158.62, BIC = -155.82, MSE = 2488.26). In the Import_Goods dataset, GZTP slightly outperformed ZTPP in MSE (1009.28 vs. 1009.68), but ZTPP had better AIC and BIC values (-541.22 and -543.82, respectively). For Log(Money_Supply), ZTPP consistently outperformed GZTP with a lower MSE (315.38 vs. 318.70), AIC (-202.35 vs. -167.84), and BIC (-198.57 vs. -164.27). Lastly, in the Brent Crude (BRT) dataset, ZTPP exhibited a significant advantage in AIC (-3427.36) and BIC (-3424.41) while maintaining a marginally lower MSE (3189.49 vs. 3189.95). These results underscore the flexibility and robustness of ZTPP across datasets, particularly in achieving better model fit and predictive accuracy.

Table 4 presents the average performance metrics of the Geometric-Zero Truncated Poisson (GZTP) and Zero-Truncated Poisson Pareto (ZTPP) distributions across multiple datasets, highlighting their comparative effectiveness. On average, the ZTPP distribution significantly outperformed the GZTP in terms of model fit, as indicated by its lower Average AIC (-1422.54) and Average BIC (-1422.55), compared to the GZTP's Average AIC (-766.39) and Average BIC (-762.82). Furthermore, the ZTPP demonstrated superior predictive accuracy, with a lower Mean Squared Error (MSE) of 30,161.94, as opposed to 31,163.2 for the GZTP. These results suggest that the ZTPP distribution offers a more robust modelling framework, with consistently better performance metrics across datasets, making it a preferable choice for applications requiring high accuracy and reliable fit.

Table 4: Comparative Analysis of Average Performance Metrics for GZTP and ZTPP Distributions

Distributions	Average AIC	Average BIC	Average MSE
GZTP	-766.386	-762.823	31163.2
ZTPP	-1422.54	-1422.55	30161.94

CONCLUSION

This study highlights the flexibility and robustness of the GZTP and ZTPP distributions in modelling real-world phenomena, particularly in economic datasets. The GZTP distribution, with its geometric component, is well-suited for capturing overdispersion and varying probabilities, making it adaptable to datasets with decaying probabilities. Conversely, the ZTPP distribution, relying solely on the rate parameter λ, excels in handling count data where zero occurrences are impossible, providing a simpler yet effective approach. The descriptive statistics of key economic variables, such as Real GDP (RGDP), Exported Goods, Import Goods, Money Supply (M2), and Brent Crude (BRT), reveal distinct patterns and variability, which inform further analyses. The comparative performance of the two distributions indicates that, on average, the ZTPP outperforms the GZTP in terms of model fit, with significantly lower AIC, BIC, and Mean Squared Error (MSE) values across multiple datasets. Specifically, the ZTPP achieved lower AIC and BIC values, such as -1422.54 and -1422.55, respectively, compared to the GZTP's -766.39 and -762.82, and demonstrated superior predictive accuracy with an average MSE of 30,161.94 versus 31,163.2 for the GZTP. These findings underscore the ZTPP distribution's superior capability in providing reliable model fit and predictive performance, making it a more suitable choice for applications that demand high accuracy in economic modelling.

REFERENCES

Abbas, N. (2023). On classical and Bayesian reliability of systems using bivariate generalized geometric distribution. *Journal of Statistical Theory and Applications, 22*(3), 151–169.

Agarwal, A., & Pandey, H. (2024). An inflated modelling of zero truncated Poisson Ailamujia distribution and its application to child mortality and genetics count data. *Journal of Scientific Research, 16*(1), 89–96.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Akdogan, Y., Kus, C., Bidram, H., & Kinaci, I. (2019). Geometric-zero truncated Poisson distribution: Properties and applications. *Gazi University Journal of Science, 32*(4), 1339–1354.

Badr, A. M. M., Hassan, T., El Din, T. S., & Ali, F. A. M. (2023). Zero truncated Poisson-Pareto distribution: Application and estimation methods. *WSEAS Transactions on Mathematics, 22,* 132–138.

Casella, G., & Berger, R. L. (2002). *Statistical inference* (2nd ed.). Duxbury Press.

Ghosh, I., Marques, F., & Chakraborty, S. (2023). A bivariate geometric distribution via conditional specification: Properties and applications. *Communications in Statistics: Simulation and Computation, 52*(12), 5925–5945.

Irshad, M. R., Monisha, M., Chesneau, C., Maya, R., & Shibu, D. S. (2023). A novel flexible class of intervened Poisson distribution by Lagrangian approach. *Stats, 6*(1), 150–168.

Ngamkham, T., & Panta, C. (2023). On the normal approximations to the method of moments point estimators of the parameter and mean of the zero-truncated Poisson distribution. *Lobachevskii Journal of Mathematics, 44*(11), 4790–4797.

Niyomdecha, A., & Srisuradetchai, P. (2023). Complementary gamma zero-truncated Poisson distribution and its application. *Mathematics, 11*(11), 2584.

Niyomdecha, A., Srisuradetchai, P., & Tulyanitikul, B. (2023). Gamma zero-truncated Poisson distribution with the minimum compounded function. *Thailand Statistician, 21*(4), 863–886.

Panichkitkosolkul, W. (2023). Bootstrap confidence intervals for the index of dispersion of zero-truncated Poisson-Ishita distribution. *Science and Technology Asia, 28*(2), 9–17.

Panichkitkosolkul, W. (2024). Non-parametric bootstrap confidence intervals for index of dispersion of zero-truncated Poisson-Lindley distribution. *Maejo International Journal of Science and Technology, 18*(1), 1–12.

R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2), 461–464.

Shang, K., Wu, T., Jin, X., Zhang, Z., Li, C., Liu, R., Wang, M., Dai, W., & Liu, J. (2023). Coaxiality prediction for aeroengines precision assembly based on geometric distribution error model and point cloud deep learning. *Journal of Manufacturing Systems, 71,* 681–694.

Shukla, K. K., Shanker, R., & Tiwari, M. K. (2020). Zero-truncated Poisson-Ishita distribution and its applications. *Journal of Scientific Research, 64*(2), 287–294.