Maejo Int. J. Sci. Technol. **2024**, *18*(01), 1-12

Maejo International Journal of Science and Technology

ISSN 1905-7873; e- ISSN: 2697-4746
Available online at www.mijst.mju.ac.th

Full Paper

Non-parametric bootstrap confidence intervals for index of dispersion of zero-truncated Poisson-Lindley distribution

Wararit Panichkitkosolkul

Department of Mathematics and Statistics, Faculty of Science and Technology,
Thammasat University, Pathumthani 12120, Thailand

E-mail: wararit@mathstat.sci.tu.ac.th

Received: 4 October 2023 / Accepted: 17 December 2023 / Published: 5 January 2024

Abstract: The Poisson distribution may not fit the data in several real-life circumstances. In this case the zero-truncated Poisson-Lindley (ZTPL) distribution has been proposed as a statistical model for counting data that do not include zero values. The index of dispersion (IOD) is a valuable tool for evaluating the suitability of the distribution in modelling observed count data. Nevertheless, the examination of the non-parametric bootstrap method for estimating confidence intervals (CIs) of the IOD of the ZTPL distribution has not been conducted. The study of the non-parametric bootstrap CI for the IOD can provide a more nuanced and informative understanding of data variability. This is crucial for various applications including comparisons between groups, risk assessment, decision-making, and ensuring the robustness of statistical conclusions. This study aims to investigate the performance of non-parametric bootstrap CIs derived from percentile, simple, and bias-corrected bootstrapping methods. Coverage probability and average length are evaluated using Monte Carlo simulation. The simulation results demonstrate that achieving the designated confidence level using non-parametric bootstrap CIs is unattainable for small sample sizes, irrespective of the other parameters. In addition, the performance of the non-parametric bootstrap CIs does not differ significantly when the sample size is large. The bias-corrected bootstrap CI demonstrates superior performance compared to other methods, even when dealing with limited sample sizes. Using two numerical examples, non-parametric bootstrap methods are utilised to calculate the CI for the IOD of a ZTPL distribution. The results match those of the simulation study.

Keywords: bootstrap interval, count data, index of dispersion, interval estimation, Lindley distribution

---

Maejo Int. J. Sci. Technol. **2024**, *18*(01), 1-12

INTRODUCTION

The Poisson distribution is a discrete probability distribution that quantifies the likelihood of a specified number of events occurring within designated temporal or spatial intervals [1, 2]. The Poisson distribution is applied to data such as the number of lightning flashes in a thunderstorm, the number of vehicles passing a checkpoint, the number of residents who will need the cardiac machine tomorrow, the number of fumbles a team makes during a game, and so on [3]. The Poisson probability model can be utilised in the analysis of data sets that comprise both zero values and positive integer values with low probability of occurrence within a predetermined temporal or spatial range [4]. The Poisson distribution is commonly employed as a fundamental model for analysing count data. However, its applicability is limited by the constraint of equi-dispersion which refers to the equality of its mean and variance. When count data exhibit over-dispersion, meaning that the variance is greater than the mean [5], a commonly used approach is to employ a mixed Poisson distribution. This distribution assumes that the Poisson parameter is a random variable [6].

Sankaran [7] investigated the mathematical and statistical properties of Poisson-Lindley (PL) distribution, which he created by combining the Poisson and Lindley distributions. The PL distribution is derived from the Poisson distribution when the Poisson parameter, denoted as λ, follows a Lindley distribution as proposed by Lindley [8] in 1958. The mathematical and statistical characteristics of the PL distribution were established by Sankaran [7]. Two estimation methods were used to estimate the parameter of the PL distribution, and when applied to two real-world data sets, the PL distribution proved to be more appropriate than the Poisson distribution [9]. The probability mass function (pmf) of the PL distribution is defined in equation (1):

$$p_0(x;\theta) = \frac{\theta^2(\theta+2+x)}{(\theta+1)^{x+3}},\ x = 0,1,2,...,\ \theta > 0. \tag{1}$$

Studies in the past have shown that the zero-modification Poisson distributions are different versions of the traditional Poisson distribution that take into account situations where the chance of a zero event happening is either higher or lower than what the standard Poisson model would predict. In this case the traditional Poisson distribution can lead to an inappropriate fit for both the counts of zero and non-zero values [10]. The zero-inflated Poisson distribution is a model that accommodates excess zeros in count data by combining a Poisson distribution for non-zero counts with a separate process that models the probability of observing zero counts [11]. In the case of the Poisson hurdle distribution, this distribution is a two-component model that addresses count data with excess zeros. It is suitable for situations where a distinct process affects the occurrence of zeros because it incorporates a hurdle component to model the decision to have any counts and a Poisson count component to model the distribution of non-zero counts [12, 13]. Zero-inflated Poisson and Poisson hurdle distributions have been much studied in research. For example, the zero-inflated mixed Poisson transmuted exponential distribution and its properties were proposed by Adetunji and Sabri [14]. Argawu and Mekebo [15] applied the zero-inflated Poisson regression analysis to the factors associated with under-five mortality in Ethiopia. Zou et al. [16] used generalised fiducial inference to construct confidence intervals (CIs) for the means of zero-inflated Poisson and Poisson hurdle models.

However, probability models can become truncated when a range of possible values for the variables is either disregarded or impossible to observe. Indeed, zero-truncation is often enforced when one wants to analyse count data without zeros. David and Johnson [17] developed the zero-truncated (ZT) Poisson (ZTP) distribution. Hussain [18] applied the ZTP distribution to data sets of the number of red mites per leaf, the remission times in the week for twenty Leukomia patients, and the number of goals scored by any team. A ZT distribution's pmf can be derived as

$$p(x;\theta)=\frac{p_0(x;\theta)}{1-p_0(0;\theta)},\ \ x=1,2,3,..., \tag{2}$$

where $p_0(x;\theta)$ and $p_0(0;\theta)$ are the pmf of the un-truncated distribution for any value of x and $x=0$ respectively.

Numerous distributions have been proposed as substitutes for the ZTP distribution in order to address the issue of over-dispersion in data. These include the ZT Poisson-Amarendra distribution [19], ZT Poisson-Akash [20] distribution, and ZT Poisson-Ishita distribution [21]. Ghitany et al. [22] introduced the ZT Poisson-Lindley (ZTPL) distribution and examined its many features including moments, coefficient of variation, skewness, kurtosis and index of dispersion (IOD). Both the maximum likelihood estimation and method of moments have been derived for the purpose of estimating the parameter of interest. Moreover, the ZTPL distribution exhibits superior performance compared to the ZTP distribution when applied to real-world data sets.

This paper focuses on the IOD, which is the ratio of variance to the mean. It is a normalised measure of the dispersion of a probability distribution. When the probability distribution of the number of occurrences in an interval is a Poisson distribution, the IOD has a value of 1. Consequently, the measure can be used to determine if observed count data can be modelled with a Poisson distribution. When the IOD is less than 1, it indicates that a data set demonstrates under-dispersion. On the other hand, when the IOD surpasses 1, a data set demonstrates the phenomenon of over-dispersion [23].

Researchers employ IOD in a multitude of disciplines such as epidemiology, finance and ecology to gain insights into the distribution patterns of events or values. As an illustration, in the field of epidemiology the IOD could be employed to evaluate the presence of a random distribution of a disease or the existence of clusters of cases. Anderson and Siddiqui [24] investigated the sampling distribution of the IOD under each of Poisson, negative binomial and binomial distributions. Panichkitkosolkul [25, 26] recently studied bootstrap CIs for the IOD of the ZT Poisson-Ishita and ZT Poisson-Amarendra distributions respectively. However, there is currently lack of research on the estimation of the CI for the IOD of the ZTPL distribution. Non-parametric bootstrap CIs offer a means of quantifying the uncertainties associated with statistical inferences derived from a sample of data. The idea is to conduct a simulation study using actual data to estimate the likely extent of sampling error [27]. Determining the IOD of a ZTPL distribution necessitates the evaluation of three bootstrap CIs, namely the percentile bootstrap (PB), the simple bootstrap (SB) and the bias-corrected (BC) bootstrap. We undertake a simulated study to assess the relative benefits of these non-parametric bootstrap CIs because a theoretical comparison of them is not possible. In addition, the non-parametric bootstrap CIs have been compared in several simulation experiments (see Reiser et al. [28] and Flowers-Cano et al. [29]). This study does a Monte Carlo simulation to assess the effectiveness of different methods and thereafter determines the ideal approach based on the coverage probability and the average length.

METHODS

Point Parameter Estimation for ZTPL Distribution

A novel method for fitting data sets that are not well fit by typical parametric distributions is the compounding of probability distributions. Ghitany et al. [22] introduced a new mixed

Maejo Int. J. Sci. Technol. **2024**, *18*(01), 1-12

distribution that combines the Poisson distribution with the Lindley distribution. This approach was motivated by the requirement for a more adaptable model to analyse statistical data. The pmf of the PL distribution is provided in equation (1).

Let X be a random variable which follows the ZTPL distribution [22] with parameter θ; it is denoted as $X \sim \text{ZTPL}(\theta)$. Using equations (1) and (2), the pmf of the ZTPL distribution can be obtained as

$$p(x;\theta) = \frac{\theta^2}{\theta^2+3\theta+1}\frac{(\theta+2+x)}{(\theta+1)^x}, \; x = 1,2,3,..., \; \theta > 0.$$

The pmf plots of the ZTPL distribution with specific values for the parameters θ are depicted in Figure 1. The expected value, variance and IOD of X are as follows:

$$E(X) = \mu = \frac{(\theta+1)^2(\theta+2)}{\theta(\theta^2+3\theta+1)}, \quad Var(X) = \sigma^2 = \frac{(\theta+1)^2(\theta^3+6\theta^2+10\theta+2)}{\theta^2(\theta^2+3\theta+1)^2},$$

and

$$IOD(X) = \gamma = \frac{\theta^3+6\theta^2+10\theta+2}{\theta(\theta+2)(\theta^2+3\theta+1)}. \quad (3)$$

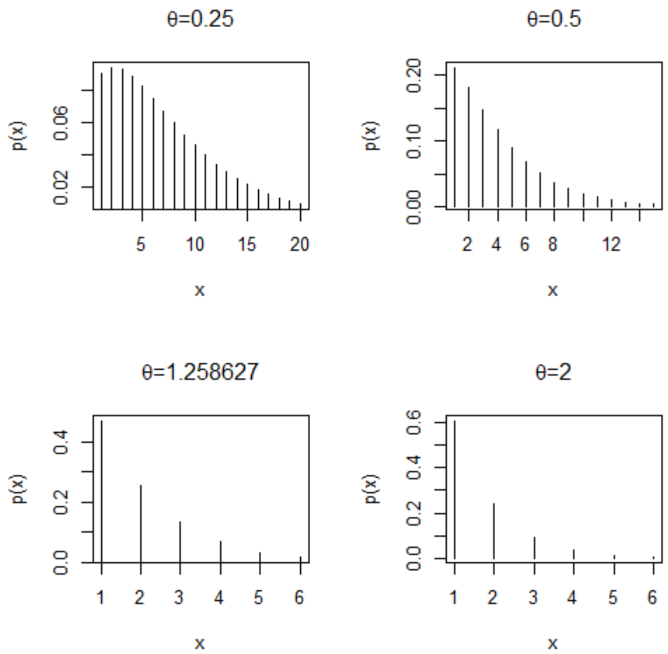

Figure 1. Pmf plots of ZTPL distribution with θ = 0.25, 0.5, 1.258627 and 2

The log-likelihood function $\log L(x_i;\theta)$ is maximised to obtain the point estimator of θ. Therefore, the maximum likelihood (ML) estimator for θ of the ZTPL distribution is derived by the following processes:

$$\frac{\partial}{\partial\theta}\log L(x_i;\theta) = \frac{\partial}{\partial\theta}\left[n\log\left(\frac{\theta^2}{\theta^2+3\theta+1}\right) - \sum_{i=1}^{n} x_i\log(\theta+1) + \sum_{i=1}^{n}\log(x_i+\theta+2)\right]$$

Maejo Int. J. Sci. Technol. **2024**, *18*(01), 1-12

$$= \frac{2n}{\theta} - \frac{n(2\theta+3)}{(\theta^2+3\theta+1)} - \frac{n\bar{x}}{\theta+1} + \sum_{i=1}^{n} \frac{1}{x_i+\theta+2}.$$

By solving the equation $\frac{\partial}{\partial\theta}\log L(x_i;\theta)=0$ for θ, we obtain a non-linear equation

$$\frac{2n}{\theta} - \frac{n(2\theta+3)}{(\theta^2+3\theta+1)} - \frac{n\bar{x}}{\theta+1} + \sum_{i=1}^{n} \frac{1}{x_i+\theta+2} = 0,$$

where $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$. As a closed-form solution for the ML estimator cannot be found, numerical iteration methods such as Newton-Raphson method, bisection method, fixed-point iteration method and secant method can be employed to solve the non-linear problem. The research utilises maxLik package [30] in statistical software R [31] for ML estimation using the Newton-Raphson method.

The point estimator of the IOD can be estimated by replacing the parameter θ with the ML estimator for θ shown in equation (3). Therefore, the point estimator of the IOD is given by

$$\hat{\gamma} = \frac{\hat{\theta}^3+6\hat{\theta}^2+10\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+2)(\hat{\theta}^2+3\hat{\theta}+1)},$$

where $\hat{\theta}$ is the ML estimator for θ.

Non-parametric Bootstrap CIs

Percentile bootstrap (*PB*) *CI*

The PB CI is a non-parametric method that estimates the uncertainty surrounding a population parameter by resampling the original sample. It is particularly useful when the underlying distribution of the data is unknown or complicated [32]. The procedure for acquiring a PB CI for θ is outlined as follows:

1) Collect the Sample Data: Start with the data from the initial sample, which represents a subset of the population. Consider that there are n observations in the sample.

2) Resample with Replacement: The bootstrap method entails resampling with replacement from the initial sample. The observations are selected from the original sample, with the prospect of multiple selections of the same observation.

3) Calculate the Statistic: The statistic of interest (parameter, mean, median, etc.) is computed for each bootstrap sample. A distribution of the statistic is obtained through repetitive resampling.

4) Generate Confidence Interval: In order to construct a CI, it is important to arrange the bootstrap statistics in ascending order and subsequently select the appropriate percentiles. For example, if we want a 95% two-sided CI, we would select the 2.5$^{\text{th}}$ percentile as the lower bound and 97.5$^{\text{th}}$ percentile as the upper bound. The $(1-\alpha)100\%$ two-sided PB CI for γ is constructed as

$$CI_{PB} = \left[\hat{\gamma}^*_{(r)}, \hat{\gamma}^*_{(s)}\right], \tag{4}$$

where $\hat{\gamma}^*_{(r)}$ is the r^{th} quantile of a collection of the parameter estimate $\hat{\gamma}^*$ arranged in ascending order, while $\hat{\gamma}^*_{(s)}$ is the s^{th} quantile of the aforementioned collection, $r=\lceil(\alpha/2)B\rceil$, $s=\lceil(1-(\alpha/2))B\rceil$, where $\lceil x\rceil$ stands for the ceiling function of x, and $1-\alpha$ is the confidence level. This study utilises α = 0.05 and B = 2,000; the two quantiles related to the lower and upper bounds of the PB two-sided CI are $\hat{\gamma}^*_{(r)} = \hat{\gamma}^*_{(50)}$ (the 50th quantile) and $\hat{\gamma}^*_{(s)} = \hat{\gamma}^*_{(1950)}$ (the 1950th quantile).

Simple bootstrap (*SB*) *CI*

The SB CI, also known as the basic bootstrap CI, is a straightforward and easily applicable method for constructing a CI, similar to the PB CI. Consider the quantity of interest to be γ and the estimator of γ to be $\hat{\gamma}$. The SB CI implies that the distributions of $\hat{\gamma} - \gamma$ and $\hat{\gamma}^* - \hat{\gamma}$ are comparable [33]. The $(1-\alpha)100\%$ two-sided SB CI for γ is

$$CI_{SB} = \left[2\hat{\gamma} - \hat{\gamma}^*_{(s)}, 2\hat{\gamma} - \hat{\gamma}^*_{(r)}\right], \tag{5}$$

where the quantiles $\hat{\gamma}^*_{(r)}$ and $\hat{\gamma}^*_{(s)}$ correspond to the same percentile of empirical distribution of bootstrap estimates $\hat{\gamma}^*$ employed in equation (4) for the PB CI.

Bias-corrected (*BC*) *bootstrap CI*

To surmount the over-coverage issues of PB CI [34], the BC bootstrap CI incorporates a bias-correction factor to correct the bias of the bootstrap parameter estimates [35]. The estimation of the bias-correction factor $\hat{z}_0$ involves determining the proportion of bootstrap estimates smaller than the original parameter estimate $\hat{\gamma}$,

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\gamma}^* \le \hat{\gamma}\}}{B}\right),$$

where Φ^{-1} is defined as the inverse of the cumulative function of standard normal distribution, $\#\{\cdot\}$ is the indicator function and B is the bootstrap replication. The values α_1 and α_2 are computed using the value of $\hat{z}_0$. They are given by

$$\alpha_1 = \Phi\{2\hat{z}_0 + z_{\alpha/2}\} \text{ and } \alpha_2 = \Phi\{2\hat{z}_0 + z_{1-\alpha/2}\},$$

where $z_{\alpha/2}$ is the α quantile of the standard normal distribution. Then the $(1-\alpha)100\%$ two-sided BC bootstrap CI for γ is

$$CI_{BC} = \left[\hat{\gamma}^*_{(j)}, \hat{\gamma}^*_{(k)}\right], \tag{6}$$

where $j = \lceil \alpha_1 B \rceil$ and $k = \lceil \alpha_2 B \rceil$, while $\lceil x \rceil$ is the ceiling function of x.

RESULTS AND DISCUSSION

Simulation Study and Results

The non-parametric bootstrap two-sided CIs for the IOD of a ZTPL distribution was considered in this study. Using R [31] version 4.3.1 and boot package [36], a Monte Carlo simulation study was designed to encompass cases with different sample sizes (n = 20, 40, 60, 80, 100 and 200). The true values of parameter (θ) were set as 0.25, 0.5, 1.258627, 2 and 3, and the IODs were 4.7969, 2.5091, 1.0000, 0.6136 and 0.3965 respectively. The bootstrap replication (B) was set at 2,000. A set of bootstrap samples, each of size n, was generated from the original sample. The process of generating these samples was repeated 1,000 times. The nominal confidence level $(1-\alpha)$ was chosen at 0.95, without any loss of generality. We compared the performance of non-parametric bootstrap CIs by assessing their coverage probabilities and average lengths. In this study we can conclude that the coverage probability is greater than or equal to the nominal confidence level when the estimated coverage probability is greater than or equal to 0.939 by using the one-proportion z-test with a significance level of 0.05. Moreover, the bootstrap CI with the minimum average length can be used to estimate the parameter more precisely. The R source code for the simulation study is available in the literature [37].

The results of the simulation study are organised and presented in Table 1. For values of n equal to 20, 40 and 60, the coverage probabilities of all three non-parametric bootstrap CIs exhibit a

Table 1. Coverage probability and average length of 95% non-parametric bootstrap CIs for IOD of ZTPL distribution

n	θ	γ	Coverage probability			Average length		
			PB	SB	BC	PB	SB	BC
20	0.25	4.7969	0.902	0.901	0.910	3.0111	3.0086	3.0342
	0.5	2.5091	0.910	0.900	0.911	1.7938	1.7945	1.8056
	1.258627	1.0000	0.909	0.909	0.918	0.9851	0.9857	0.9943
	2	0.6136	0.896	0.881	0.924	0.7212	0.7223	0.7365
	3	0.3965	0.899	0.892	0.938	0.5640	0.5644	0.5887
40	0.25	4.7969	0.936	0.934	0.938	2.1718	2.1732	2.1826
	0.5	2.5091	0.944*	0.935	0.944*	1.3280	1.3269	1.3309
	1.258627	1.0000	0.931	0.913	0.937	0.7296	0.7294	0.7331
	2	0.6136	0.913	0.892	0.918	0.5412	0.5406	0.5460
	3	0.3965	0.919	0.919	0.940*	0.4202	0.4203	0.4261
60	0.25	4.7969	0.922	0.925	0.926	1.7745	1.7783	1.7804
	0.5	2.5091	0.942*	0.938	0.939*	1.0771	1.0765	1.0798
	1.258627	1.0000	0.919	0.919	0.925	0.5952	0.5958	0.5978
	2	0.6136	0.936	0.940*	0.938	0.4469	0.4468	0.4489
	3	0.3965	0.937	0.920	0.943*	0.3528	0.3528	0.3557
80	0.25	4.7969	0.940*	0.943*	0.941*	1.5534	1.5524	1.5583
	0.5	2.5091	0.949*	0.946*	0.944*	0.9398	0.9398	0.9411
	1.258627	1.0000	0.939*	0.931	0.947*	0.5237	0.5245	0.5253
	2	0.6136	0.949*	0.936	0.952*	0.3926	0.3923	0.3943
	3	0.3965	0.939*	0.923	0.952*	0.3011	0.3006	0.3023
100	0.25	4.7969	0.935	0.935	0.936	1.3828	1.3835	1.3849
	0.5	2.5091	0.939*	0.941*	0.942*	0.8536	0.8538	0.8555
	1.258627	1.0000	0.944*	0.946*	0.950*	0.4696	0.4697	0.4711
	2	0.6136	0.946*	0.935	0.946*	0.3513	0.3508	0.3529
	3	0.3965	0.931	0.922	0.931	0.2741	0.2737	0.2751
200	0.25	4.7969	0.948*	0.950*	0.951*	0.9836	0.9842	0.9849
	0.5	2.5091	0.951*	0.946*	0.951*	0.6008	0.6014	0.6017
	1.258627	1.0000	0.947*	0.946*	0.944*	0.3335	0.3334	0.3336
	2	0.6136	0.936	0.934	0.934	0.2515	0.2510	0.2515
	3	0.3965	0.948*	0.946*	0.951*	0.1922	0.1923	0.1926

* Empirical coverage probability is greater than or equal to nominal confidence level.

Maejo Int. J. Sci. Technol. **2024**, *18*(01), 1-12

tendency to be below 0.95, thereby failing to attain the expected nominal confidence level. Nevertheless, the BC bootstrap CI exhibits superior performance compared to the others under these conditions in terms of coverage probability. For values of n equal to 80, 100 and 200, the non-parametric bootstrap CIs achieve coverage probabilities that are in close proximity to the specified confidence level. Additionally, these intervals exhibit similar average lengths. As a result, the coverage probabilities of the CIs tend to rise with increasing sample sizes and become closer to the nominal confidence level of 0.95.

Moreover, as the value of the IOD is decreased, the average length of the CIs decreases due to the relationship between the IOD and θ. As expected, the average lengths of all three bootstrap CIs decrease as the sample size increases. Even though the PB and SB CIs' average lengths are shorter when the sample size is small ($n = 20$), this results in a poor coverage probability value that is much lower than the nominal confidence level. The BC bootstrap CI exhibits superior performance in terms of coverage probability even when dealing with small sample sizes, provided that the IOD of the ZTPL distribution is not excessively large.

Empirical Applications of Non-parametric Bootstrap CIs

Two actual count data sets are utilised to illustrate the suitability of non-parametric bootstrap CIs in estimating the IOD of a ZTPL distribution.

Immunogold assay example

This example utilises the count of sites containing particles obtained from the immunogold assay data collected by Cullen et al. [37]. The sample mean and standard deviation for this data set of 198 observations, shown in Table 2, are 1.576 and 0.891 respectively. The Chi-squared statistic for the Chi-squared goodness-of-fit test [38] is 0.5467, and the p-value is 0.7608. Consequently, a ZTPL distribution with $\hat{\theta} = 2.1831$ is appropriate for this data set. The point estimator of the IOD is 0.5586. The data set demonstrates the phenomenon of under-dispersion because the IOD is less than 1. The 95% non-parametric bootstrap CIs for the IOD of a ZTPL distribution are presented in Table 3. Because the average length of the BC bootstrap CI is shorter than that of the PB and SB CIs, the results are consistent with the simulation results.

Table 2. Number of counts of particle-containing sites found in immunogold assay

Number of particles	1	2	3	≥ 4
Observed frequency	122	50	18	8
Expected frequency	124.7689	46.7604	17.0663	9.4044

Table 3. 95% Non-parametric bootstrap CIs and corresponding lengths using all intervals for IOD in immunogold assay example

Method	CI	Length
PB	(0.4501, 0.6754)	0.2253
SB	(0.4421, 0.6723)	0.2302
BC	(0.4488, 0.6733)	0.2245

Maejo Int. J. Sci. Technol. **2024**, *18*(01), 1-12

Demographic example

Shanker et al. [39] present in Table 4 the demographic data regarding the number of fertile mothers who have experienced at least one infant death. The total sample size is 135. The Chi-squared statistic is 3.3797, whereas the p-value for the Chi-squared goodness-of-fit test is 0.1845 [38]. As a result, a ZTPL distribution with $\hat{\theta}$ equal to 2.0891 is an appropriate choice for this data set. The point estimator of the IOD is 0.6039. The data set exhibits under-dispersion as indicated by an IOD value of less than 1. Table 5 contains the 95% non-parametric bootstrap CIs for the IOD of a ZTPL distribution that have been calculated using the bootstrap method. The BC bootstrap CI is shorter than the PB and SB CIs; therefore the findings match the simulated results for $n =$ 100 and $\theta =$ 2.

Table 4. Number of fertile mothers who have experienced at least one child death

Number of child deaths	1	2	3	≥ 4
Observed frequency	89	25	11	10
Expected frequency	83.4486	32.3222	12.1818	7.0474

Table 5. 95% Non-parametric bootstrap CIs and corresponding lengths using all intervals for IOD in demographic example

Method	CI	Length
PB	(0.4279, 0.7477)	0.3198
SB	(0.4127, 0.7372)	0.3245
BC	(0.4408, 0.7595)	0.3187

Discussion

Based on the simulation results, all three bootstrap CIs functioned well in all scenarios with large sample sizes ($n \geq 80$). The average lengths remained largely unchanged and the coverage probabilities were close to the nominal confidence level. However, when dealing with small sample sizes ($n =$20, 40 and 60), all three bootstrap CIs displayed coverage probabilities that were below the nominal confidence level. In addition, when both the sample size and the IOD value increased, the average lengths of all the bootstrap CIs decreased. The findings of this research do not exhibit any substantial differences when compared to previous studies because there are theoretical reasons to generally prefer BC bootstrap CIs. The present methodology has the potential to assist scientists in comparing the IOD across different samples or experimental conditions, allowing them to assess variations in the spatial distribution of labelled entities. For example, scientists may compare IOD values between healthy and diseased tissues or between treated and untreated samples to understand changes in the distribution pattern. Furthermore, the IOD can serve as a quality control metric for immunogold assay experiments. Consistent IOD values across replicates indicate reliable and reproducible results, while large variations may signal issues with experimental procedures or sample preparation.

One limitation of this study is that the non-parametric bootstrap CIs are not exact, but they exhibit consistency, indicating that the coverage probability tends to approach 0.95 as the sample sizes increase. Furthermore, the computation of three non-parametric bootstrap CIs is difficult and requires significant computational resources. However, there exist several R packages that can be utilised for the computation of bootstrap CIs. These packages include the boot package [36], bootstrap package [40], semEff package [41] and BootES package [42]. Users have the freedom to download these packages as R is an open-source software. It would be advantageous to focus the comparative examination of alternative CI estimations in relation to the bootstrap CIs reported in this research in future investigations. The construction of CIs for functions of parameters, such as the difference and the ratio of the IOD, is of interest. Additionally, there is a lack of statistical theoretical research regarding hypothesis testing for the IOD of the ZTPL distribution. The bootstrap CIs examined in this study can be utilised for alternative distributions. These topics may require additional examination in future studies.

CONCLUSIONS

The IOD quantifies the measure of variability or dispersion within a probability distribution. Within the framework of the ZTPL distribution, the IOD can provide valuable insights into the extent of spread within the distribution. Bootstrap CIs offer several benefits including their robustness, flexibility and capacity to make conclusions without relying on a particular data distribution. They exhibit strong performance with non-Gaussian data and in scenarios where traditional parametric approaches are unsuitable. In addition, they offer a method that does not rely on a specific model to estimate the distribution of statistics obtained by sampling, and they are relatively simple to put into practice. The results indicate that the size of the sample has a significant influence on the performance of the bootstrap CIs. The coverage probabilities of all non-parametric bootstrap CIs are significantly below the desired confidence level of 0.95 when the sample sizes are 20, 40 and 60. For sample sizes greater than or equal to 80, there are no significant differences observed in the coverage probabilities and average lengths obtained from all non-parametric bootstrap CIs. Based on the present research outcomes, the BC bootstrap CI exhibits superior performance across a wide range of scenarios.

ACKNOWLEDGEMENTS

The author would like to thank the Editor-in-Chief and the anonymous reviewers for their insightful suggestions and careful reading of the manuscript.

REFERENCES

1. R. V. Hogg, J. W. McKean and A. T. Craig, "Introduction to Mathematical Statistics", 6th Edn., Pearson Prentice Hall, Upper Saddle River, **2005**, pp.143-147.
2. A. F. Siegel and M. R. Wanger, "Practical Business Statistics", 8th Edn., Academic Press, London, **2022**, pp.190-192.
3. R. J. Larsen and M. L. Marx, "An Introduction to Mathematical Statistics and Its Applications", 4th Edn., Pearson Prentice Hall, Upper Saddle River, **2006**, pp.275-286.
4. P. Sangnawakij, "Confidence interval for the parameter of the zero-truncated Poisson distribution", *J. Appl. Sci.*, **2021**, *20*, 13-22.
5. S. H. Ong, Y. C. Low and K. K. Toh, "Recent developments in mixed Poisson distributions", *ASM Sci. J.*, **2021**, *14*, doi:10.32802/asmscj.2020.464.
6. R. Tharshan and P. Wijekoon, "A new mixed Poisson distribution for over-dispersed count data: Theory and applications", *Reliab. Theory Appl.*, **2022**, *17*, 33-51.
7. M. Sankaran, "The discrete Poisson-Lindley distribution", *Biometr.*, **1970**, *26*, 145-149.

Maejo Int. J. Sci. Technol. **2024**, *18*(01), 1-12

8. D. V. Lindley, "Fiducial distributions and Bayes' theorem", *J. R. Stat. Soc. B*, **1958**, *20*, 102-107.
9. R. Shanker and H. Fesshaye, "On Poisson-Lindley distribution and its applications to biological sciences", *Biom. Biostat. Int. J.*, **2015**, *2*, 103-107.
10. S. E. Perumean-Chaney, C. Morgan, D. McDowall and I. Aban, "Zero-inflated and overdispersed: What's one to do?", *J. Stat. Comput. Simul.*, **2013**, *83*, 1671-1683.
11. D. Lambert, "Zero-inflated Poisson regression, with an application to defects in manufacturing", *Technometr.*, **1992**, *34*, 1-14.
12. J. Mullahy, "Specification and testing of some modified count data models", *J. Econ.*, **1986**, *33*, 341-365.
13. D. C. Heilbron, "Zero-altered and other regression models for count data with added zeros", *Biometr. J.*, **1994**, *36*, 531-547.
14. A. A. Adetunji and S. R. M. Sabri, "On zero-inflated mixed Poisson transmuted exponential distribution: Properties and applications to observation with excess zeros", *Maejo Int. J. Sci. Technol.*, **2023**, *17*, 68-80.
15. A. S. Argawu and G. G. Mekebo, "Zero-inflated Poisson regression analysis of factors associated with under-five mortality in Ethiopia using 2019 Ethiopian mini demographic and health survey data", *PLoS One*, **2023**, *18*, Art.no.e0291426.
16. Y. Zou, J. Hannig and D. S. Young, "Generalized fiducial inference on the mean of zero-inflated Poisson and Poisson hurdle models", *J. Stat. Distrib. Appl.*, **2021**, *8*, Art.no.5
17. F. N. David and N. L. Johnson, "The truncated Poisson", *Biometr.*, **1952**, *8*, 275-285.
18. T. Hussain, "A zero truncated discrete distribution: Theory and applications to count data", *Pak. J. Stat. Oper. Res.*, **2020**, *16*, 167-190.
19. R. Shanker, "A zero-truncated Poisson-Amarendra distribution and its application", *Int. J. Probab. Stat.*, **2017**, *6*, 82-92.
20. R. Shanker, "Zero-truncated Poisson-Akash distribution and its applications", *Am. J. Math. Stat.*, **2017**, *7*, 227-236.
21. K. K. Shukla, R. Shanker and M. K. Tiwari, "Zero-truncated Poisson-Ishita distribution and its applications", *J. Sci. Res.*, **2020**, *64*, 287-294.
22. M. E. Ghitany, D. K. Al-Mutairi and S. Nadarajah, "Zero-truncated Poisson-Lindley distribution and its application", *Math. Comput. Simul.*, **2008**, *79*, 279-287.
23. D. R. Cox and P. A. W. Lewis, "The Statistical Analysis of Series of Events", 1st Edn., Methuen, London, **1966**, pp.71-72.
24. J. C. Anderson and M. M. Siddiqui, "The sampling distribution of the index of dispersion", *Commun. Stat. Theory Meth.*, **1994**, *23*, 897-911.
25. W. Panichkitkosolkul, "Bootstrap confidence intervals for the index of dispersion of zero-truncated Poisson-Ishita distribution", *Sci. Technol. Asia*, **2023**, *28*, 9-17.
26. W. Panichkitkosolkul, "Bootstrap methods for estimating the confidence interval for the index of dispersion of the zero-truncated Poisson-Amarendra distribution", *Interdiscip. Res. Rev.*, **2023**, *18*, 13-22.
27. M. Wood, "Statistical inference using bootstrap confidence intervals", *Signif.*, **2004**, *1*, 180-182.
28. M. Reiser, L. Yao, X. Wang, J. Wilcox and S. Gray, "A Comparison of bootstrap confidence intervals for multi-level longitudinal data using Monte-Carlo simulation", in "Monte-Carlo

Maejo Int. J. Sci. Technol. **2024**, *18*(01), 1-12

Simulation-Based Statistical Modeling" (Ed. D. G. Chen and J. D. Chen), Springer, Singapore, **2017**, Ch.3.

29. R. S. Flowers-Cano, R. Ortiz-Gómez, J. E. León-Jiménez, R. L. Rivera and L. A. P. Cruz, "Comparison of bootstrap confidence intervals using Monte Carlo simulations", *Water*, **2018**, *10*, Art.no.166.
30. A. Henningsen and O. Toomet, "maxLik: A package for maximum likelihood estimation in R", *Comput. Stat.*, **2011**, *26*, 443-458.
31. R. Ihaka and R. Gentleman, "R: A language for data analysis and graphics", *J. Comput. Graph. Stat.*, **1996**, *5*, 299-314.
32. B. Efron, "The Jackknife, the Bootstrap, and Other Resampling Plans", 1st Edn., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, **1982**, pp.27-36.
33. W. Q. Meeker, G. J. Hahn and L. A. Escobar, "Statistical Intervals: A Guide for Practitioners and Researchers", 2nd Edn., John Wiley and Sons, New York, **2017**, pp.245-266.
34. B. Efron and R. J. Tibshirani, "An Introduction to the Bootstrap", 1st Edn., Chapman and Hall, New York, **1994**, pp.325-328.
35. B. Efron, "Better bootstrap confidence intervals", *J. Am. Stat. Assoc.*, **1987**, *82*, 171-185.
36. A. Canty and B. Ripley, "Boot: Bootstrap R (S-Plus) functions", **2022**, https://cran.r-project.org/web/packages/boot/index.html (Accessed: October 2023).
37. W. Panichkitkosolkul, "R source code for the simulation study: non-parametric bootstrap confidence intervals for index of dispersion of zero-truncated Poisson-Lindley distribution", **2023**, https://codeocean.com/capsule/7192289/tree (Accessed: December 2023).
38. M. J. Cullen, J. Walsh, L. V. Nicholson and J. B. Harris, "Ultrastructural localization of dystrophin in human muscle by using gold immunolabelling", *Proc. R. Soc. Lond. B Biol. Sci.*, **1990**, *240*, 197-210.
39. N. S. Turhan, "Karl Pearson's chi-square tests", *Educ. Res. Rev.*, **2020**, *15*, 575-580.
40. R. Shanker, H. Fesshaye, S. Selvaraj and A. Yemane, "On zero-truncation of Poisson and Poisson-Lindley distributions and their applications", *Biom. Biostat. Int. J.*, **2015**, *2*, 168-181.
41. S. Kostyshak, "Bootstrap: Functions for the Book – An Introduction to the Bootstrap", **2022**, https://cran.r-project.org/web/packages/bootstrap/index.html (Accessed: October 2023).
42. M. V. Murphy, "semEff: Automatic calculation of effects for piecewise structural equation models", **2022**, https://cran.r-project.org/web/packages/semEff/index.html (Accessed: October 2023).
43. K. N. Kirby and D. Gerlanc, "BootES: An R package for bootstrap confidence intervals on effect sizes", *Behav. Res. Meth.*, **2013**, *45*, 905-927.