Journal of Current Science and Technology, September-December 2023 Vol. 13 No. 3, pp. 725-733

ISSN 2630-0656 (Online)

Cite this article: Suriyakat, W., & Panichkitkosolkul, W. (2023). Nonparametric bootstrap confidence intervals for the population mean of a Zero-Truncated Poisson-Lindley distribution and their application. *Journal of Current Science and Technology, 13*(3), 725-733. https://doi.org/10.59796/jcst.V13N3.2023.1077

Journal of Current Science and Technology

Journal homepage: https://jcst.rsu.ac.th

Nonparametric Bootstrap Confidence Intervals for The Population Mean of A Zero-Truncated Poisson-Lindley Distribution and Their Application

Wannaphon Suriyakat[1] and Wararit Panichkitkosolkul[2*]

[1]Faculty of Science and Technology, Pibulsongkram Rajabhat University, Phitsanulok 65000, Thailand
[2]Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathum Thani 12120, Thailand

*Corresponding author; Email: wararit@mathstat.sci.tu.ac.th

Received 22 March, 2023; Revised 21 June, 2023; Accepted 21 June, 2023;
Published online 30 August, 2023.

Abstract

Recently, the zero-truncated Poisson-Lindley distribution has been proposed for studying count data containing non-zero values. However, the nonparametric bootstrap confidence interval estimation of the population mean has not yet been studied. In this study, confidence interval estimation based on percentile, simple, and biased-corrected bootstrap methods was compared in terms of coverage probability and average interval length via Monte Carlo simulation. The true values of parameter (θ) were set as 0.25, 0.5, 1, 1.5, and 2, and the population means μ are approximate 7.7586, 4.0909, 2.4000, 1.8817, and 1.6364, respectively. The bootstrap samples $(B = 1{,}000)$ of size n were generated from the original sample, and each simulation was repeated 1,000 times. The results indicate that attaining the nominal confidence level using the bootstrap confidence intervals was impossible for small sample sizes regardless of the other settings. Moreover, when the sample size was large, the performance of the nonparametric bootstrap confidence intervals was not substantially different. Overall, the bias-corrected bootstrap confidence interval outperformed the others, even for small sample sizes. Last, the nonparametric bootstrap confidence intervals were used to calculate the confidence interval for the population mean of the zero-truncated Poisson-Lindley distribution via two numerical examples, the results of which match those from the simulation study.

Keywords: *bootstrap interval; count data; interval estimation; Lindley distribution, simulation*

1. Introduction

Poisson distribution is a discrete distribution that measures the probability of a given number of events happening in specific regions of time or space (Kissell & Poserina, 2017; Siegel & Wanger, 2022). Data such as the number of errors in a computer program submitted for the first time to a mainframe computer by each student, the number of people who will apply for a job tomorrow, the number of defects in a finished product, the number of heat stroke patients per day in summer, the number of bacteria in a higher organism, etc., follow a Poisson distribution (Siegel & Wanger, 2022; Lefebvre, 2000).

The probability mass function (pmf) of a Poisson distribution is defined as

$$p(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0,1,2,\ldots, \ \lambda > 0, \qquad (1)$$

where e is a constant approximately equal to 2.71828 and λ is the mean number of events within a given interval of time or space. This probability model can be used to analyze data containing zeros and positive values that have low occurrence probabilities within a predefined time or area range (Sangnawakij, 2021). 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 analyze count data without zeros. David & Johnson (1952) developed the zero-truncated (ZT) Poisson (ZTP) distribution, which has been applied to datasets of the length of stay in hospitals, the number of published journal articles in various disciplines, the number of children ever born to a sample of mothers over 40 years old, and the number of passengers in cars (Hussain, 2020). A ZT distribution's pmf can be derived as follows.

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

where $p_0(x)$ and $p_0(0)$ are the pmf of the un-truncated distribution for any value of x and $x = 0,$ respectively. Sankaran (1970) proposed the Poisson-Lindley (PL) distribution which has the pmf:

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

The mathematical and statistical properties of the PL distribution for modeling count data were established by Sankaran (1970). The PL distribution arises from the Poisson distribution when parameter λ follows the Lindley distribution proposed by Lindley (1958) with probability density function (pdf):

$$f(\lambda;\theta) = \frac{\theta^2}{1+\theta}(1+\lambda)e^{-\theta\lambda}, \ \ \lambda > 0, \theta > 0. \qquad (4)$$

Shanker et al. (2015a) showed that the pdf in (4) is a better model than the exponential distribution for modeling lifetime data. Several distributions have been introduced as an alternative to the ZTP distribution for handling over-dispersion in data, such as ZT Poisson-Amarendra (ZTPA) (Shanker, 2017a), ZT Poisson-Akash (Shanker, 2017b) and ZT Poisson-Ishita (Shukla et al., 2020) distributions.

Ghitany et al. (2008) proposed the ZT Poisson-Lindley (ZTPL) distribution and its properties, such as the moment, coefficient of variation, skewness, kurtosis, and the index of dispersion. The method of moments and the maximum likelihood have also been derived for estimating its parameter. Furthermore, when the ZTPL distribution was applied to real data sets, it was more suitable than the ZTP distribution.

In statistics, the population refers to a group or collection of various entities such as numbers, people, or objects. The population mean represents the average value within this group. To the best of our knowledge, no research has been conducted on estimating the confidence interval for the population mean of the ZTPL distribution. It is essential to note that the score function of ZTPL distribution is complicated, and the maximum likelihood estimator has no closed form. Therefore, likelihood-based, score, and Wald-type confidence intervals have no closed forms. In such cases, finding these confidence intervals can be challenging; alternative methods, such as numerical techniques or resampling methods like the nonparametric bootstrap method, can be utilized. Nonparametric bootstrap confidence intervals provide a way of quantifying the uncertainties in statistical inferences based on a sample of data. The concept is to run a simulation study based on the actual data to estimate the likely extent of sampling error (Wood, 2004).

2. Objective

The objective of the current study is to assess the efficiencies of three nonparametric bootstrap confidence intervals, namely the percentile bootstrap (PB), the simple bootstrap (SB), and the bias-corrected (BC) bootstrap methods, to estimate the population mean of the ZTPL distribution. Additionally, none of the nonparametric bootstrap confidence intervals will be exact (i.e., the actual confidence level is exactly equal to the nominal confidence level $1-\alpha$) but they will all be consistent, meaning that the confidence level approaches $1-\alpha$ as the sample size gets large (Chernick & LaBudde, 2011). In light of the impossibility of a theoretical comparison of these nonparametric bootstrap confidence intervals, we conducted a simulated study to evaluate their relative merits. Moreover, several studies have compared the nonparametric bootstrap confidence intervals through simulation studies (see Reiser et al., 2017; Flowers-Cano et al., 2018). In this study, a Monte Carlo simulation study is conducted to compare their performance and use the results to determine the best-performing method based on the coverage probability and the average length.

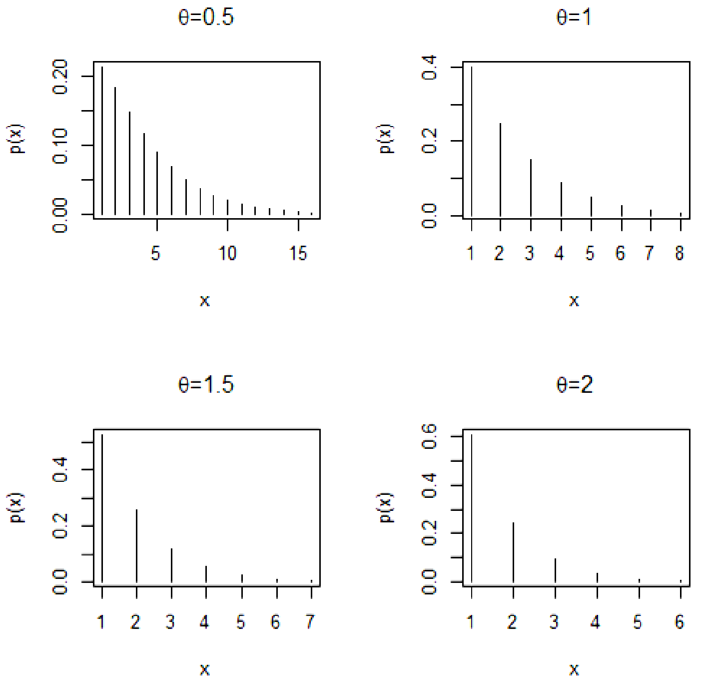

Figure 1 The plots of the pmf of the ZTPL distribution with $\theta = 0.5$, 1, 1.5 and 2.

3. Methodology

3.1 Theoretical background

To obtain new probability distributions, compounding of probability distributions is an innovative approach to fit data sets inadequately fit by common distributions. Ghitany et al. (2008) proposed a novel mixture distribution by compounding Poisson distribution with the Lindley distribution, as there is a need to find a more flexible model for analyzing statistical data. The pmf of the PL distribution is given in Eq.(3).

Let X be a random variable which follow the ZTPL distribution (Ghitany et al., 2008) with parameter θ, it is denoted as *X~ ZTPL(θ)*. Using Eqs. (2) and (3), the pmf of the ZTPL distribution can be obtained as:

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

The plots of the pmf of the ZTPL distribution with some specified parameter values θ are shown in Figure 1. It is clear that the shape of Figure 1 depends on the parameter values θ. From Figure 1, when the parameter value is larger, the $p(1)$ is larger and $p(x) \to 0$ for the large value of x.

The expected value and variance of X are as follows:

$$E(X) = \mu = \frac{(\theta+1)^2(\theta+2)}{\theta(\theta^2+3\theta+1)} \tag{5}$$

and

$$Var(X) = \sigma^2 = \frac{(\theta+1)^2(\theta^3+6\theta^2+10\theta+2)}{\theta^2(\theta^2+3\theta+1)^2}.$$

The point estimator of θ is obtained by maximizing the log-likelihood function $\log L(x_i;\theta)$ or the logarithm of joint pmf of $X_1,\ldots,X_n$. Therefore, the maximum likelihood (ML) estimator for θ of the ZTPL distribution is derived by the following processes:

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

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

Solving the equation $\frac{\partial}{\partial\theta}\log L(x_i;\theta) = 0$ for θ, we have the 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 $\bar{x} = \sum_{i=1}^{n} x_i / n$ denotes the sample mean.

Since the ML estimator for θ does not provide a closed-form, the non-linear equation can be solved by the numerical iteration methods such as Newton-Raphson method, bisection method and Ragula-Falsi method. In this research, we use the maxLik package (Henningsen & Toomet, 2011) with the Newton-Raphson method for ML estimation in the statistical software R.

The point estimator of the population mean μ can be estimated by replacing the parameter θ with the ML estimator for θ shown in Eq. (5). Therefore, the point estimator of the population mean μ is given by:

$$\hat{\mu} = \frac{(\hat{\theta}+1)^2(\hat{\theta}+2)}{\hat{\theta}\left(\hat{\theta}^2+3\hat{\theta}+1\right)},$$

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

3.2 Nonparametric bootstrap confidence intervals

In this study, we focus on the three nonparametric bootstrap confidence intervals for the population mean of the ZTPL distribution. In practice, the popular nonparametric bootstrap confidence intervals are the percentile bootstrap, the simple bootstrap, and the bias-corrected bootstrap methods. The computer-intensive bootstrap methods described in this study provide an alternative for constructing approximate confidence intervals for the population mean without having to make an assumption about the underlying distribution (Meeker et al., 2017).

3.2.1 Percentile Bootstrap (PB) confidence interval

The percentile bootstrap confidence interval is the interval between the $(\alpha/2)\times 100$ and $(1-(\alpha/2))\times 100$ percentiles of the distribution of μ estimates obtained from resampling or the distribution of $\hat{\mu}^*$, where μ represents a parameter of interest and α is the level of significance (e.g., α = 0.05 for 95% confidence intervals) (Efron, 1982). A percentile bootstrap confidence interval for μ can be obtained as follows:

1) B random bootstrap samples are generated,
2) a parameter estimate $\hat{\mu}^*$ is calculated from each bootstrap sample,
3) all B bootstrap parameter estimates are ordered from the lowest to highest, and
4) the $(1-\alpha)100\%$ percentile bootstrap confidence interval is constructed as follows:

$$CI_{PB} = \left[\hat{\mu}^*_{(r)}, \hat{\mu}^*_{(s)}\right], \tag{6}$$

where $\hat{\mu}^*_{(\alpha)}$ denotes the α^{th} percentile of the distribution of $\hat{\mu}^*$ and $0 \le r < s \le 100$. For example, a 95% percentile bootstrap confidence interval with 1,000 bootstrap samples is the interval between the 2.5 percentile value and the 97.5 percentile value of the 1,000 bootstrap parameter estimates.

3.2.2 Simple Bootstrap (SB) confidence interval

The simple bootstrap confidence interval, sometimes called the basic bootstrap confidence interval, is a method as easy to apply as the percentile bootstrap method. Suppose that the quantity of interest is μ and that the estimator of μ is $\hat{\mu}$. The simple bootstrap method assumes that the distributions of $\hat{\mu}-\mu$ and $\hat{\mu}^*-\hat{\mu}$ are approximately the same (Meeker et al., 2017). The $(1-\alpha)100\%$ simple bootstrap confidence interval for μ is:

$$CI_{SB} = \left[2\hat{\mu}-\hat{\mu}^*_{(s)}, 2\hat{\mu}-\hat{\mu}^*_{(r)}\right], \tag{7}$$

where the quantiles $\hat{\mu}^*_{(r)}$ and $\hat{\mu}^*_{(s)}$ are the same percentile of empirical distribution of bootstrap estimates $\hat{\mu}^*$ used in (6) for the percentile bootstrap method.

3.2.3 Bias-Corrected (BC) bootstrap confidence interval

To overcome the over-coverage issues in percentile bootstrap confidence intervals (Efron & Tibshirani, 1993), the BC bootstrap method corrects for the bias of the bootstrap parameter estimates by incorporating a bias-correction factor (Efron, 1987; Efron & Tibshirani, 1993). The bias-correction factor $\hat{z}_0$ is estimated as the proportion of the bootstrap estimates less than the original parameter estimate $\hat{\mu}$,

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

where Φ^{-1} is the inverse function of a standard normal cumulative distribution function (e.g., $\Phi^{-1}(0.975) \approx 1.96$). With the value of $\hat{z}_0$, the values α_1 and α_2 are calculated,

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

where $z_{\alpha/2}$ is the α quantile of the standard normal distribution (e.g. $z_{0.05/2} \approx -1.96$). Then, the $(1-\alpha)100\%$ BC bootstrap confidence interval for μ is as follows:

$$CI_{BC} = \left[\hat{\mu}^*_{(\alpha_1)}, \hat{\mu}^*_{(\alpha_2)}\right], \tag{8}$$

where $\hat{\mu}^*_{(\alpha)}$ denotes the α^{th} percentile of the distribution of $\hat{\mu}^*$.

4. Result

4.1 Simulation study

The nonparametric bootstrap confidence intervals for the population mean of a ZTPL distribution was considered in this study. Due to the unavailability of a direct theoretical comparison, a Monte Carlo simulation study was designed using R (Ihaka & Gentleman, 1996) version 4.2.2 to cover cases with different sample sizes (n = 10, 30, 50, 100, 200, and 500). To observe the effect of small and large variances, the true values of parameter (θ) were set as 0.25, 0.5, 1, 1.5, and 2 and the population means μ are approximate 7.7586, 4.0909, 2.4000, 1.8817 and 1.6364, respectively. $B = 1{,}000$ bootstrap samples of size n were generated from the original sample and each simulation was repeated 1,000 times. Without loss of generality, the nominal confidence level $(1-\alpha)$ was set at 0.95. The performance of the nonparametric bootstrap confidence intervals was compared in terms of their empirical coverage probabilities and average lengths; the one with a coverage probability greater than or close to the nominal confidence level (meaning that it contains the true value). From this study, we conclude that the coverage probability is greater than or equal to the nominal confidence level when the empirical coverage probability exceeds 0.936, as determined using the one-proportion z-test with a significance level of 0.95. Additionally, using the shortest average length enables a more accurate estimation of the nonparametric bootstrap confidence interval for the population mean.

The simulation results of the study were reported in Table 1. For $n = 10$, 30 and 50, the empirical coverage probabilities of all three bootstrap confidence intervals tended to be less than 0.95 and thus did not reach the nominal confidence level. Nevertheless, the BC bootstrap confidence interval outperformed the other in these situations. For $n = 100$, all bootstrap methods provided empirical coverage probabilities less than the nominal confidence level in several cases. For $n \geq$ 200, all of the nonparametric bootstrap confidence intervals attained empirical coverage probabilities close to the nominal confidence level and provided similarly average length. Therefore, as the sample size was increased, the empirical coverage probabilities of the confidence intervals tended to increase and approach the nominal confidence level of 0.95.

Moreover, the average lengths of the confidence intervals decreased when the value of μ was decreased because of the relationship between the variance and θ. Unsurprisingly, as the sample size was increased, the average lengths of all three bootstrap confidence intervals decreased. Although the average length of the PB and SB confidence intervals was the shortest when the sample size was small, it provided a poor empirical coverage probability value significantly below the nominal confidence level. In summary, the BC bootstrap confidence interval performs best in terms of empirical coverage probability even with small sample sizes as long as the population mean of the ZTPL distribution is not too large.

4.2 Empirical application of the nonparametric bootstrap confidence intervals

We used two real-world count data sets to demonstrate the applicability of the nonparametric bootstrap confidence intervals for estimating the population mean of the ZTPL distribution.

4.2.1 Immunogold assay example

The number of counts of sites with particles from immunogold assay data collected by Cullen et al. (1990) was used for this example. The data consisted of 198 observations were reported in Table 2. For this dataset, the sample mean, and the standard deviation were 1.576 and 0.891, respectively. For the Chi-squared goodness-of-fit test (Turhan, 2020), the Chi-squared statistic was 0.5467 and the p-value was 0.7608. Thus, a ZTPL distribution with $\hat{\theta} = 2.1831$ was suitable for this dataset. The point estimator of the population mean was 1.5765. Table 3 reported the 95% non-parametric bootstrap confidence intervals for the population mean of a ZTPL distribution. The estimated parameter $\hat{\theta}$ was near to 2. The results corresponded with the simulation results because the average lengths of the PB and SB confidence intervals were shorter than those of the BC bootstrap confidence interval.

4.2.2 Demographic example

Table 4 showed the demographic data on the number of fertile mothers who have experienced at least one child death (Shanker et al., 2015b). The total sample size was 135. For the Chi-squared goodness-of-fit test (Turhan, 2020), the Chi-squared statistic was 3.3797 and the p-value was

0.1845. Thus, a ZTPL distribution with $\hat{\theta} = 2.0891$ was suitable for this dataset. The point estimator of the population mean was 1.6058. The 95% nonparametric bootstrap confidence intervals for the population mean of a ZTPL distribution are reported in Table 5. The results correspond with the simulation results for $n = 100$ and $\theta = 2$ because the average lengths of the PB confidence interval were shorter than those of the BC bootstrap confidence interval.

Table 1 Empirical coverage probability and average length of the 95% nonparametric bootstrap confidence intervals for the population mean of a ZTPL distribution.

n	θ	μ	Coverage probability			Average length		
			PB	**SB**	**BC**	**PB**	**SB**	**BC**
10	0.25	7.7586	0.868	0.843	0.874	6.6393	6.6272	6.7193
	0.5	4.0909	0.871	0.844	0.878	3.4392	3.4355	3.4970
	1	2.4000	0.883	0.862	0.902	1.9051	1.9044	1.9561
	1.5	1.8817	0.882	0.862	0.925	1.3753	1.3731	1.4347
	2	1.6364	0.864	0.824	0.930	1.0913	1.0899	1.1535
30	0.25	7.7586	0.928	0.914	0.934	4.1799	4.1815	4.2155
	0.5	4.0909	0.927	0.912	0.928	2.2176	2.2201	2.2391
	1	2.4000	0.911	0.908	0.917	1.2031	1.2033	1.2121
	1.5	1.8817	0.910	0.899	0.928	0.8495	0.8504	0.8603
	2	1.6364	0.924	0.895	0.938*	0.6708	0.6709	0.6812
50	0.25	7.7586	0.932	0.925	0.935	3.3074	3.3059	3.3264
	0.5	4.0909	0.937*	0.936*	0.944*	1.7466	1.7434	1.7546
	1	2.4000	0.929	0.922	0.924	0.9532	0.9547	0.9631
	1.5	1.8817	0.923	0.916	0.925	0.6743	0.6726	0.6774
	2	1.6364	0.926	0.912	0.935	0.5364	0.5378	0.5432
100	0.25	7.7586	0.953*	0.956	0.952*	2.3512	2.3489	2.3498
	0.5	4.0909	0.930	0.927	0.934	1.2428	1.2435	1.2446
	1	2.4000	0.928	0.907	0.921	0.6641	0.6646	0.6667
	1.5	1.8817	0.940*	0.930	0.945*	0.4830	0.4835	0.4866
	2	1.6364	0.933	0.924	0.933	0.3857	0.3854	0.3874
200	0.25	7.7586	0.934	0.939*	0.939*	1.6788	1.6770	1.6812
	0.5	4.0909	0.949*	0.944*	0.953*	0.8790	0.8789	0.8802
	1	2.4000	0.943*	0.945*	0.942*	0.4823	0.4820	0.4834
	1.5	1.8817	0.939*	0.939*	0.940*	0.3439	0.3440	0.3451
	2	1.6364	0.947*	0.943*	0.942*	0.2750	0.2749	0.2757
500	0.25	7.7586	0.941*	0.944*	0.945*	1.0692	1.0685	1.0707
	0.5	4.0909	0.938*	0.940*	0.942*	0.5598	0.5589	0.5596
	1	2.4000	0.938*	0.931	0.943*	0.3030	0.3031	0.3030
	1.5	1.8817	0.947*	0.947*	0.944*	0.2177	0.2178	0.2184
	2	1.6364	0.945*	0.946*	0.942*	0.1758	0.1753	0.1757

* Represents the empirical coverage probability is greater than the nominal confidence level by using the one-proportion z-test.

Table 2 The number of counts of sites with particles from immunogold assay data

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 The 95% nonparametric bootstrap confidence intervals and corresponding lengths using all intervals for the population mean in the immunogold assay example

Methods	**Confidence intervals**	**Lengths**
PB	(1.4556, 1.7034)	0.2478
SB	(1.4556, 1.6974)	0.2418
BC	(1.4647, 1.7243)	0.2596

Table 4 The 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 The 95% nonparametric bootstrap confidence intervals and corresponding lengths using all intervals for the population mean in the demographic example

Methods	Confidence intervals	Lengths
PB	(1.4423, 1.7777)	0.3354
SB	(1.4207, 1.7752)	0.3545
BC	(1.4434, 1.7914)	0.3480

5. Discussion and conclusion

Herein, we propose three nonparametric bootstrap confidence intervals, namely the PB, SB, and BC bootstrap methods, to estimate the population mean of a ZTPL distribution. When the sample size was 10, 30, or 50, the coverage probabilities of all three were substantially lower than 0.95. When the sample size was large enough $(n \geq 100)$, the coverage probabilities and average lengths using three nonparametric bootstrap confidence intervals were not markedly different. According to our findings, the BC bootstrap confidence interval performed the best even for small sample sizes and parameter settings tested in both the simulation study and using real data sets. Confidence intervals are obtained from a parametric estimator of the standard errors of a quantity of interest μ. Then, the $(1-\alpha)100\%$ confidence interval for μ is obtained by adding or subtracting the standard error multiplied by a critical value (for example, $\hat{\mu} \pm z_{1-(\alpha/2)} SE(\hat{\mu})$). This calculation assumes that the distribution of the estimator of μ is approximately normal (Flowers-Cano et al., 2018). However, there are several situations in which the assumption of normality is violated. In these cases, or when the standard error is very difficult to be estimated, an alternative is to use techniques based on the nonparametric bootstrap method. The nonparametric bootstrap methods described in this study provide an alternative for constructing approximate confidence intervals without assuming the underlying distribution (Meeker et al., 2017). This is an advantage of this study.

On the other hand, the limitation of this study is the fact that none of the nonparametric bootstrap confidence intervals were exact, but they would be consistent, meaning that the coverage probability approaches 0.95 as the sample sizes get larger. In addition, three bootstrap confidence intervals are not easy to compute and are computer intensive. However, there are numerous available packages in R for computing the bootstrap confidence intervals such as boot package (Canty & Ripley, 2021), bootstrap package (Kostyshak, 2019), semEff package (Murphy, 2022) and BootES package (Kirby & Gerlanc, 2013). Since R is open-source, users are free to download these packages. Future research could focus on the construction of the confidence intervals for the parameter functions such as the population mean, variance, and coefficient of variation. Moreover, there is no research on hypothesis testing for the parameter of the ZTPL distribution. These issues can be studied in future work.

6. Acknowledgement

The authors are appreciative of the anonymous referees whose insightful comments significantly improved the paper's quality and presentation.

7. References

Canty, A., & Ripley, B. (2021). *boot: Bootstrap R (S-Plus) functions. R package version 1.3-28.1*. Retrieved from https://cran.rproject.org/web/packages/boot/index.html.

Chernick, M. R. & Labudde, R. A. (2011). *An introduction to bootstrap methods with applications to R* (1st ed.). New Jersey: John Wiley and Sons.

Cullen, M. J., Walsh, J., Nicholson, L. B., & Harris, J. B. (1990). Ultrastructural localization of dystropin in human muscle by using gold immunolabelling. *Proceedings of the Royal Society of London. B. Biological Sciences*, *240*(1297), 197-210. https://doi.org/10.1098/rspb.1990.0034.

David, F. N., & Johnson, N. L. (1952). The truncated Poisson. *Biometrics*, *8*(4), 275-285. https://doi.org/10.2307/3001863.

Efron, B. (1982). *The Jackknife, the bootstrap, and other resampling plans* [Conference presentation]. CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia: SIAM.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*(397), 171-185. https://doi.org/10.1080/01621459.1987.10478410.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap* (1st ed.). London: Chapman and Hall.

Flowers-Cano, R. S., Ortiz-Gómez, R., León-Jiménez, J. E., Rivera, R. L., & Cruz, L. A. P. (2018). Comparison of bootstrap confidence intervals using Monte Carlo Simulations. *Water*, *10*(2), Article 166. https://doi.org/10.3390/w10020166.

Ghitany, M. E., Al-Mutairi, D. K., & Nadarajah, S. (2008). Zero-truncated Poisson-Lindley distribution and its application. *Mathematics and Computers in Simulation*, *79*(3), 279-287. https://doi.org/10.1016/j.matcom.2007.11.021.

Henningsen, A., & Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, *26*(3), 443-458. https://doi.org/10.1007/s00180-010-0217-1.

Hussain, T. (2020). A zero truncated discrete distribution: Theory and applications to count data. *Pakistan Journal of Statistics and Operation Research*, *16*(1), 167-190. https://doi.org/10.18187/pjsor.v16i1.2133.

Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, *5*(3), 299-314. https://doi.org/10.1080/10618600.1996.10474713.

Kirby, K. N., & Gerlanc, D. (2013). BootES: An R package for bootstrap confidence intervals on effect sizes. *Behavior Research Methods*, *45*(4), 905-927. https://doi.org/10.3758/s13428-013-0330-5.

Kissell, R., & Poserina, J. (2017). *Optimal sports math, statistics, and fantasy* (1st ed.). Massachusetts, US: Academic Press.

Kostyshak, S. (2019). *bootstrap: Functions for the book An introduction to the bootstrap. R package version 2019.6*. Retrieved from https://cran.rproject.org/web/packages/bootstrap/index.html.

Lefebvre, M. (2000). *Applied probability and statistics* (1st ed.). Springer Sciences.

Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society: Series B (Methodological)*, *20*(1), 102-107. https://doi.org/10.1111/j.2517-6161.1958.tb00278.x

Meeker, W. Q., Hahn, G. J., & Escobar, L. A. (2017). *Statistical intervals*: *A guide for practitioners and researchers* (2nd ed.). Toronto: John Wiley and Sons.

Murphy, M. V. (2022). *semEff: Automatic calculation of effects for piecewise structural equation models. R package version 0.6.1*. Retrieved from https://cran.rproject.org/web/packages/semEff/index.html.

Reiser, M., Yao, L., Wang, X., Wilcox, J. & Gray, S. (2017). A Comparison of bootstrap confidence intervals for multi-level longitudinal data using Monte-Carlo simulation. In Chen, D. G., Chen, J. (Eds), *Monte-Carlo Simulation-Based Statistical Modeling*. New York: Springer. https://doi.org/10.1007/978-981-10-3307-0_17

Sankaran, M. (1970). The discrete Poisson-Lindley distribution. *Biometrics*, *26*(1), 145-149. https://doi.org/10.2307/2529053.

Sangnawakij, P. (2021). Confidence interval for the parameter of the zero-truncated Poisson distribution. *The Journal of Applied Science*, *20*(2), 13-22. https://doi.org/10.14416/j.appsci.2021.02.002

Shanker, R., Fesshaye, H., Selvaraj S. & Yemane, A. (2015a). On zero-truncation of Poisson and Poisson-Lindley distributions and their applications, *Biometrics & Biostatistics International Journal*, *2*(6), 168-181. https://doi.org/10.15406/bbij.2015.02.00045

Shanker, R., Fesshaye, H., Selvaraj, S. (2015b). On modeling of lifetimes data using exponential and Lindley distributions. *Biometrics & Biostatistics International Journal*, *2*(5), 140-147. https://doi.org/10.15406/bbij.2015.02.00042.

Shanker, R. (2017a). A zero-truncated Poisson-Amarendra distribution and its application. *International Journal of Probability and Statistics*, *6*(4), 82-92. https://doi.org/10.15406/bbij.2016.04.00099

Shanker, R. (2017b). Zero-truncated Poisson-Akash distribution and its applications. *American Journal of Mathematics and Statistics*, *7*(6), 227-236.

Shukla, K. K., Shanker, R., & Tiwari, M. K. (2020). Zero-truncated Poisson-Ishita distribution and its application. *Journal of Scientific Research*, *64*(2), 287-294. https://doi.org/10.37398/JSR.2020.640240.

Siegel, A. F., & Wanger, M. R. (2022). *Practical business statistics* (8th ed.). Academic Press.

Turhan, N. S. (2020). Karl Pearson’s chi-square tests. *Educational Research Review*, *15*(9), 575-580. https://doi.org/10.5897/ERR2019.3817.

Wood, M. (2004). Statistical inference using bootstrap confidence intervals. *Significance*, *1*(4), 180-182. https://doi.org/10.1111/j.1740-9713.2004.00067.x