Journal of Sciences, Islamic Republic of Iran 34(3): 247 - 257 (2023) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Bootstrap Confidence Intervals for the Parameter of the Poisson-Sujatha Distribution and Their Applications to Agriculture

W. Panichkitkosolkul[1*] and Ch. Ponkaew[2]

[1] *Department of Mathematics and Statistics, Thammasat University, Pathumthani, 12120 Thailand*
[2] *Department of Mathematics, Faculty of Science and Technology, Phetchabun Rajabhat University, Phetchabun, 67000 Thailand*

Received: 13 May 2023 / Revised: 3 July 2023 / Accepted: 10 July 2023

Abstract

In a number of real-world situations, one encounters count data with over-dispersion such that the typical Poisson distribution does not suit the data. In the current situation, it is appropriate to employ a combination of mixed Poisson and Poisson-Sujatha (PS) distributions. The PS distribution has been investigated for count data, which is of primary interest to a number of disciplines, including biology, medicine, demography, and agriculture. However, no research has been conducted regarding generating bootstrap confidence intervals for its parameter. The coverage probabilities and average lengths of bootstrap confidence intervals derived from the percentile, basic, and biased-corrected and accelerated bootstrap methods were compared using Monte Carlo simulation. The results indicated that it was impossible to achieve the nominal confidence level using bootstrap confidence intervals for tiny sample sizes, regardless of the other settings. Furthermore, when the sample size was large, there was not much of a difference in the performance of the several bootstrap confidence intervals. The bias-corrected and accelerated bootstrap confidence interval demonstrated superior performance compared to the other methods in all of the cases examined. Moreover, the effectiveness of the bootstrap confidence intervals was proven through their application to agricultural data sets. The calculations offer significant evidence in favor of the suggested bootstrap confidence intervals.

Keywords: Interval Estimation; Poisson Distribution; Mixed Distribution; Count Data; Bootstrap Method.

Introduction

The Poisson distribution is frequently used to model the number of events that occur in a given time and/or location (1) .A Poisson distribution applies to data such as the number of thunderstorms per month, the number of orders a company will receive the next day, the number of calls received per hour at a call centre, the number of defects in a completed product, etc. (2). The Poisson distribution is an essential model for the

* Corresponding author: Tel: +660868929579; Email: wararit@mathstat.sci.tu.ac.th

analysis of count data, but its use is limited due to the equality of its mean and variance (equi-dispersion). In comparison to the Poisson distribution, count data frequently exhibit over-dispersion, with a variance larger than the mean (3, 4). The application of the Poisson distribution to data with over-dispersion can result in inaccurate analyses and incorrect conclusions (5). A possible solution for addressing over-dispersion in count data is to employ a mixed Poisson distribution. This approach assumes that the Poisson parameter, which governs the distribution, is a random variable characterized by a single parameterized distribution (6).

Shanker (7) investigated the mathematical and statistical properties of the Poisson-Sujatha (PS) distribution developed by combining the Poisson and Sujatha distributions. The PS distribution is derived from the Poisson distribution when the Poisson parameter λ (the average number of occurrences) follows a Sujatha distribution. The PS distribution was found to be more appropriate than the Poisson and Poisson-Lindley (8) distributions when applied to two actual data sets.

Shanker (9) introduced the Sujatha distribution as a one-parameter lifetime continuous distribution with a probability density function (pdf) defined as

$$f(x;\theta)=\frac{\theta^3}{2+\theta+\theta^2}\left(1+x+x^2\right)\exp(-\theta x), \qquad (1)$$

whereSujatha distribution The $\theta>0$. and $x>0$ is a continuous distribution that consists of a combination of three probability distributions: the exponential distribution, exp(θ), the gamma(2, θ) and the gamma(3, θ) distributions. These three distributions are weighted by their respective proportions $\frac{\theta^2}{2+\theta+\theta^2}$, $\frac{\theta}{2+\theta+\theta^2}$, and $\frac{2}{2+\theta+\theta^2}$, respectively**.** The aforementioned distribution has been utilized for the purpose of modeling lifetime data within the domains of engineering and biomedical science. Moreover, Shanker (9) demonstrated that the Sujatha distribution outperforms the exponential, Lindley (10), and Akash (11) distributions as a more suitable model. Shanker (9) had examined the important statistical properties of the Sujatha distribution. Figure 1 depicts the pdf plots of the Sujatha distribution with specified values of parameter θ.

The confidence interval, a fundamental component of statistical inference, is a range of values that is highly probable to encompass the real value of the population parameter of interest. It serves as a crucial outcome in numerous statistical analyses and plays a pivotal role in the interpretation of parameter estimations (12). According to our best knowledge, no studies have been done on calculating the confidence interval for a PS distribution parameter. Bootstrap confidence intervals

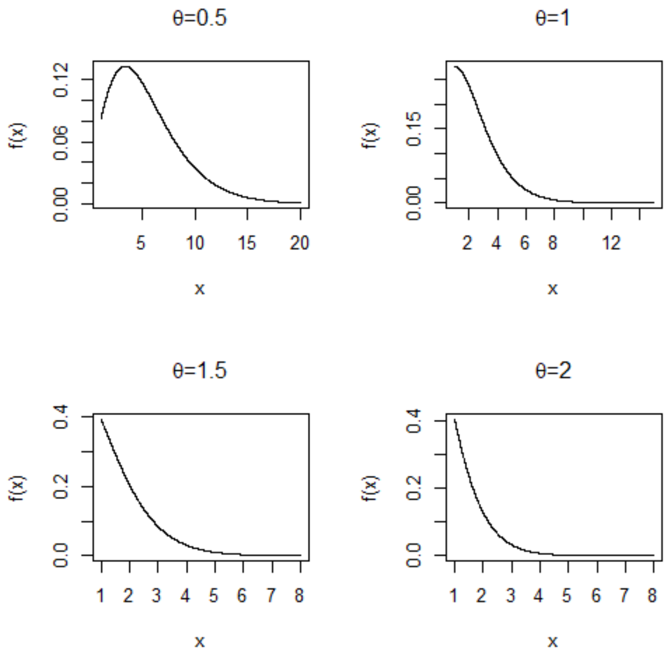

Figure 1. Plots of the pdf of the Sujatha distribution for θ = 0.5, 1, 1.5 and 2

for estimating the parameter quantify the uncertainty associated with statistical inference based on the sample data. The idea is to conduct a simulation study using real data in order to estimate the potential size of sampling error (13). The key objective of the current study is to evaluate the effectiveness of three different bootstrap confidence interval estimations, specifically the percentile bootstrap (PB), the basic bootstrap (BB), and the bias-corrected and accelerated (BCa) bootstrap, in estimating the parameter of the PS distribution. In addition, none of the bootstrap confidence intervals will be exact (i.e., the actual confidence level is precisely equal to the nominal confidence level $1-\alpha$), but they will all be consistent, with the confidence level approaching $1-\alpha$ as the sample size increases (14). Given the inherent limitations in conducting a theoretical comparison of bootstrap confidence intervals, we select to perform a simulated study to assess their respective advantages and disadvantages. In addition, the bootstrap methods were compared in a simulation investigation in a number of studies (see Reiser et al. (15), Flowers-Cano et al. (16), Mostajeran et al. (17)). In the current study, a Monte Carlo simulation was implemented to compare their performance. Based on the chance of coverage and the average length, the simulation results were used to find the method with the best performance.

Theoretical Background

In probability theory, the Poisson distribution is characterized by its probability mass function (pmf), which may be written as

$$p(y;\lambda)=\frac{\exp(-\lambda)\lambda^{y}}{y!}, \qquad (2)$$

where $y=0,1,2,\ldots,$ e is a constant equal to approximately 2.718282 and λ is a Poisson parameter; $\lambda>0$. Let X represent a random variable which follows the PS distribution with a parameter θ, which is commonly denoted as $X \sim \mathrm{PS}(\theta)$. According to Shanker (7), the pmf of the PS distribution is defined as

$$p(x;\theta)=\frac{\theta^{3}}{2+\theta+\theta^{2}}\frac{x^{2}+(\theta+4)x+(4+3\theta+\theta^{2})}{(\theta+1)^{x+3}}, \qquad (3)$$

where $x=0,1,2,\ldots,$ and $\theta>0$. Figure 2 shows the pmf plots of the PS distribution for a range of parameter values θ. According to the PS distribution, the expected value (mean) and variance of the random variable X are as follows:

$$E(X)=\mu=\frac{6+2\theta+\theta^{2}}{\theta(2+\theta+\theta^{2})}$$

and

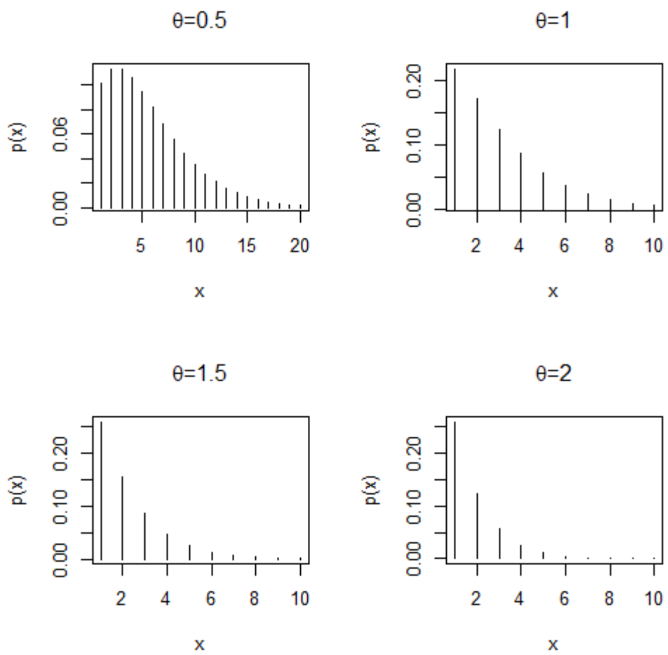

Figure 2. Plots of the pmf of the PS distribution for θ= 0.5, 1, 1.5 and 2

$$Var(X) = \sigma^2 = \frac{12 + 24\theta + 28\theta^2 + 14\theta^3 + 4\theta^4 + \theta^5}{\theta^2\left(2 + \theta + \theta^2\right)^2}.$$

Maximizing the log-likelihood function $\log L(x_i;\theta)$ or the logarithm of the joint pmf of $X_1,\ldots,X_n$ yields the point estimator of θ. Consequently, the derivation of the ML estimator for θ involves the following procedures:

$$\frac{\partial}{\partial\theta}\log L(x_i;\theta) = \frac{\partial}{\partial\theta}\begin{bmatrix} n\log\left(\dfrac{\theta^3}{2+\theta+\theta^2}\right) - \sum_{i=1}^{n}(x_i+3)\log(\theta+1) \\ +\sum_{i=1}^{n}\log\left[x_i^2+(\theta+4)x_i+(4+3\theta+\theta^2)\right] \end{bmatrix}$$

$$= \frac{n(6+2\theta+\theta^2)}{\theta(2+\theta+\theta^2)} - \frac{n(\bar{x}+3)}{\theta+1} + \sum_{i=1}^{n}\frac{x_i+(2\theta+3)}{\left[x_i^2+(\theta+4)x_i+(4+3\theta+\theta^2)\right]}.$$

We obtain the non-linear equation by solving for θ in the equation $\frac{\partial}{\partial\theta}\log L(x_i;\theta) = 0$:

$$\frac{n(6+2\theta+\theta^2)}{\theta(2+\theta+\theta^2)} - \frac{n(\bar{x}+3)}{\theta+1} + \sum_{i=1}^{n}\frac{x_i+(2\theta+3)}{\left[x_i^2+(\theta+4)x_i+(4+3\theta+\theta^2)\right]} = 0,$$

where $\bar{x} = n^{-1}\sum_{i=1}^{n}x_i$ denotes the sample mean. According to the inability to find a closed-form solution provided by the ML estimator for θ, numerical iteration techniques such as bisection, Newton-Raphson, and Ragula-Falsi methods can be used to solve the resulting non-linear problem. The paper utilized the maxLik package (18) with the Newton-Raphson method for ML estimation. The statistical software R (19) was employed for this purpose.

Bootstrap Confidence Intervals

The interval estimation, also known as confidence interval, is derived from an estimator that calculates the standard errors of a parameter, denoted as ϕ. Then, the standard error is multiplied by the critical value to add or subtract, giving the $(1-\alpha)100\%$ two-sided confidence interval for ϕ (for example, $\hat{\phi} \pm z_{1-(\alpha/2)}SE(\hat{\phi})$). This computation is based on the normality assumption of the estimator of ϕ (16). Nevertheless, there are a number of situations in which the normality assumption is not appropriate to make an estimation. In such instances, or when estimating the standard error is exceedingly challenging, it is reasonable to employ techniques based on the bootstrap method. The computationally intensive bootstrap methods described in this study offer an alternative to assuming the underlying distribution when constructing approximate confidence intervals (20). In this paper, we concentrate on the three bootstrap confidence intervals for the PS distribution parameter. The most frequently employed bootstrap confidence intervals in practice are the PB, BB, and BCa confidence intervals (14). With the use of the boot package (21) and the statistical software R (19), the bootstrap confidence intervals were calculated for this investigation.

Percentile Bootstrap (PB) Confidence Interval

The PB two-sided confidence interval is defined as the range bounded by the $(\alpha/2)\times 100$ and $(1-(\alpha/2))\times 100$ percentiles of the distribution for the estimated values of θ acquired from resampling or the distribution of $\hat{\theta}^*$, where θ denotes an important parameter and α denotes the level of significance (22). The procedure for constructing a PB confidence interval for the PS distribution parameter is as follows:

1) With a replacement, B random bootstrap samples of the underlying distribution are created, where B is the number of bootstrap replications,

2) From each bootstrap sample, a parameter estimate $\hat{\theta}^*$ is determined,

3) All parameter estimates from the B bootstrap samples are ranked from smallest to greatest, and

4) the $(1-\alpha)100\%$ PB two-sided confidence interval is created as follows:

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

where the notation $\hat{\theta}^*_{(r)}$ is the r^{th} quantile of a collection of the parameter estimate $\hat{\theta}^*$ arranged in ascending order, while $\hat{\theta}^*_{(s)}$ 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 α is the significance level. This study utilized α = 0.05 and B = 2,000; the two quantiles related to the lower and upper bounds of the PB two-sided confidence interval were $\hat{\theta}^*_{(r)} = \hat{\theta}^*_{(50)}$ (the 50th quantile) and $\hat{\theta}^*_{(s)} = \hat{\theta}^*_{(1950)}$ (the 1950th quantile).

Basic Bootstrap (BB) Confidence Interval

The BB method, often known as the simple bootstrap method, is just as straightforward to implement as the PB method. Consider the quantity of interest to be θ and the estimator of θ to be $\hat{\theta}$. The BB technique presumes that $\hat{\theta}-\theta$ and $\hat{\theta}^*-\hat{\theta}$ follow similar distributions (20). The $(1-\alpha)100\%$ BB two-sided confidence interval for θ is

$$CI_{BB}=\left[2\hat{\theta}-\hat{\theta}^*_{(s)},\,2\hat{\theta}-\hat{\theta}^*_{(r)}\right], \quad (5)$$

where the quantiles $\hat{\theta}^*_{(r)}$ and $\hat{\theta}^*_{(s)}$ represent the same percentile of the empirical distribution of bootstrap estimates $\hat{\theta}^*$ that are utilized in (4) to calculate the PB confidence interval.

Bias-Corrected and Accelerated (BCa) Bootstrap Confidence Interval

The calculation of BCa bootstrap confidence intervals commonly involves the utilization of influence statistics derived from jackknife simulations. However, incorporating jackknife simulation alongside ordinary bootstrapping is computationally expensive for the intended purposes. The BCa bootstrap confidence interval uses a bias-correction element and an acceleration element to correct for the bias and skewness of the bootstrap parameter estimates, mitigating the over-coverage problems seen with the PB confidence interval (22, 23, 24). Davison and Hinkley (25) and Chernick and LaBudde (14) described the mathematical particulars of the BCa adjustment. The bias-correction element $\hat{z}_0$ is calculated by

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

where Φ^{-1} is the inverse function of the standard normal distribution's cumulative distribution function. The acceleration element $\hat{a}$ is calculated via jackknife resampling, which entails generating n replicates of the initial set of data, where n is the sample sizes. The initial jackknife replication is obtained by omitting the first case $(i=1)$ from the initial sample, the second by omitting the second case $(i=2)$, etc., until a total of n samples, each with a size of $n-1$, are generated. Based on the jackknife resamples, we obtain the value of $\hat{\theta}_{(-i)}$. The acceleration factor $\hat{a}$ is given by

$$\hat{a}=\frac{\sum_{i=1}^{n}\left(\hat{\theta}_{(\cdot)}-\hat{\theta}_{(-i)}\right)^3}{6\left\{\sum_{i=1}^{n}\left(\hat{\theta}_{(\cdot)}-\hat{\theta}_{(-i)}\right)^2\right\}^{3/2}},$$

where $\hat{\theta}_{(\cdot)}=\sum_{i=1}^{n}\hat{\theta}_{(-i)}/n$. The computation of the values of α_1 and α_2 are as follows:

$$\alpha_1=\Phi\left\{\hat{z}_0+\frac{\hat{z}_0+z_{\alpha/2}}{1-\hat{a}\left(\hat{z}_0+z_{\alpha/2}\right)}\right\} \quad \text{and}$$

$$\alpha_2=\Phi\left\{\hat{z}_0+\frac{\hat{z}_0+z_{1-\alpha/2}}{1-\hat{a}\left(\hat{z}_0+z_{1-\alpha/2}\right)}\right\},$$

where $z_{\alpha/2}$ is the $\alpha/2$ quantile of the normal distribution with a mean of zero and standard deviation of 1. Then, the $(1-\alpha)100\%$ BCa bootstrap two-sided confidence interval for the PS distribution parameter is computed by

$$CI_{BCa}=\left[\hat{\theta}^*_{(j)},\hat{\theta}^*_{(k)}\right], \quad (6)$$

where $j=\lceil\alpha_1 B\rceil$ and $k=\lceil\alpha_2 B\rceil$. When the value $\hat{z}_0$ is set to 0 and $\hat{a}$ is also set to 0, it can be observed that the BCa confidence interval is equal to the PB confidence interval.

Simulation Study

In the current study, the interval estimation for the parameter of the PS distribution was estimated using three bootstrap two-sided confidence intervals. Due to the lack of a direct theoretical comparison, a Monte Carlo simulation investigation was conducted using R (19) version 4.3.1 to include cases with various sample sizes (n = 10, 30, 50, 100, and 500). To observe the effects of both large and small variances, the actual value of the parameter (θ) was chosen from 0.1, 0.5, 0.8, 1 and 1.5. Because Ukoumunne et al. (26) asserted that 2,000 bootstrap samples are enough to estimate the coverage probability for the 95% confidence intervals with a standard error of slightly under 0.5%, the number of bootstrap replications (B) was fixed at 2,000. From the initial pseudo-random sample, n-sized bootstrap samples were generated, and each simulation was repeated 1,000 times. Without sacrificing generality, the level of confidence $(1-\alpha)$ was set at 0.95. The study

Table 1 . Coverage probability and the average length of the 95 %bootstrap confidence intervals for θ of the PS distribution

n	θ	Coverage probability			Average length		
		PB	BB	BCa	PB	BB	BCa
10	0.1	0.894	0.889	0.886	0.0773	0.0773	0.0748
	0.5	0.872	0.862	0.882	0.4663	0.4666	0.4465
	0.8	0.894	0.893	0.903	0.8957	0.8998	0.8367
	1.0	0.914	0.882	0.915	1.2677	1.2641	1.1743
	1.5	0.893	0.883	0.893	2.8157	2.8582	2.4965
30	0.1	0.937	0.912	0.940	0.0433	0.0432	0.0425
	0.5	0.912	0.906	0.915	0.2529	0.2529	0.2474
	0.8	0.908	0.913	0.911	0.4397	0.4397	0.4273
	1.0	0.931	0.939	0.936	0.5904	0.5898	0.5702
	1.5	0.928	0.928	0.940	1.0862	1.0867	1.0256
50	0.1	0.947	0.933	0.948	0.0330	0.0331	0.0327
	0.5	0.942	0.935	0.942	0.1921	0.1924	0.1898
	0.8	0.929	0.931	0.932	0.3350	0.3351	0.3293
	1.0	0.937	0.934	0.947	0.4461	0.4465	0.4371
	1.5	0.935	0.924	0.939	0.7818	0.7812	0.7559
100	0.1	0.943	0.936	0.946	0.0237	0.0237	0.0236
	0.5	0.948	0.932	0.943	0.1343	0.1340	0.1332
	0.8	0.939	0.945	0.942	0.2346	0.2345	0.2324
	1.0	0.956	0.950	0.953	0.3082	0.3081	0.3051
	1.5	0.928	0.953	0.934	0.5359	0.5358	0.5279
500	0.1	0.941	0.948	0.943	0.0106	0.0106	0.0106
	0.5	0.948	0.948	0.946	0.0603	0.0603	0.0601
	0.8	0.953	0.949	0.951	0.1037	0.1037	0.1035
	1.0	0.949	0.949	0.946	0.1365	0.1365	0.1360
	1.5	0.961	0.945	0.958	0.2331	0.2331	0.2322

evaluated the performance of different bootstrap two-sided confidence intervals by examining their coverage probabilities and average lengths. The ideal confidence interval should have a coverage probability that is equal to or very close to the nominal confidence level, indicating that it contains the real parameter. Additionally, the proposed confidence interval with the shortest average length is preferred as it provides the most accurate estimation for the parameter of interest in a given scenario. Therefore, it is evident that in cases where the coverage probability is equal, a smaller average length signifies the suitability of the bootstrap confidence interval for that specific situation.

The study's findings are presented in Table 1. For the given value of n =10, the coverage probabilities of all three bootstrap confidence intervals exhibited a tendency to be below 0.90, indicating that they did not achieve the intended nominal confidence level. In these situations, however, the BCa bootstrap two-sided confidence interval outperformed the others. For a sample size of 30, we again find that no bootstrap confidence interval yields a probability of coverage in excess of the nominal confidence level of 0.95. For $n \geq$ 50, each bootstrap confidence interval achieved coverage probabilities close to the nominal level of confidence and had average lengths that were comparable. Nevertheless, the BCa bootstrap confidence interval exhibited a coverage probability that was more closely aligned with the specified nominal confidence level of 0.95. The coverage probabilities tended to rise along with the sample size, approaching the nominal confidence level of 0.95 as the sample size grew larger. Due to the relationship between variance and θ, the average lengths of the bootstrap confidence intervals increased as the value of θ was increased. Consequently, as the sample size increased, the average lengths of all three bootstrap two-sided confidence intervals decreased, with the BCa bootstrap confidence interval having the shortest average length for all situations examined. Furthermore, when the sample size was small (n =10), there was a statistically significant difference in the average length of the BCa bootstrap confidence interval compared to the others. It was observed that the PB and BB confidence intervals for the average length did not differ significantly across sample sizes. In brief, the BCa bootstrap two-sided confidence interval demonstrates superior performance in terms of estimated coverage probability and average length when applied to moderate and large sample sizes ($n \geq 50$).

Applications to Real-World Data

In this section, we demonstrate the applicability of bootstrap confidence intervals for estimating the PS distribution parameter using two real-world data sets.

Application to the Number of Red Mites

This example utilizes the number of red mites in a apple farm collected by Bliss and Fisher (27). Table 2 presents the dataset comprising 150 observations. The sample mean and standard deviation for this data set are 1.1500 and 1.4504, respectively. This study employs the chi-squared goodness-of-fit test to determine if the sample data are likely to be representative of a particular theoretical distribution (28). The chi-square statistic was 3.3889, while the p-value was 0.4950. Consequently, a PS distribution with $\hat{\theta}$ = 1.6533 is appropriate for this data set. Table 3 displays the 95% bootstrap confidence intervals for the PS distribution parameter. The results are consistent with those of the simulation because the average lengths of the BCa bootstrap confidence interval were shorter than those of the PB and BB intervals.

Application to the Number of Corn Borer Larvae Per Plant

McGuire et al. (29) recorded the number of corn borer larvae per plant in the field corn of Northwest Iowa, United States. Table 4 provides the dataset, with a total sample size of 324. The sample mean and standard deviation for this data set are 0.6481 and 0.9208, respectively. The chi-squared statistic for the chi-squared goodness-of-fit test (28) was 1.1743 and the p-value was 0.5559. Therefore, the data best fits a PS distribution with $\hat{\theta}$ = 2.4717. Table 5 shows the 95% bootstrap confidence intervals for the PS distribution parameter. Because the average lengths of the BCa bootstrap confidence interval were shorter than those of the PB and BB confidence intervals, the results were consistent with the simulation results.

Results and Discussion

For estimating the parameter of the Poisson-Sujatha distribution, the percentile bootstrap (PB), the basic bootstrap (BB), and the bias-corrected and accelerated (BCa) bootstrap methods were proposed. When the sample sizes were relatively small (n = 10 and 30), the coverage probabilities for all three bootstrap confidence intervals were significantly below the desired threshold of 0.95. When the sample size was large enough ($n \geq$ 50), the coverage probabilities and average lengths derived from the three bootstrap confidence intervals did not differ significantly. According to the results of our research, the BCa bootstrap confidence interval was the best for virtually all of the scenarios, both in the

Table 2 . The number of red mites on apple leaves

Number of red mites	0	1	2	3	4	≥ 5
Observed frequency	70	38	17	10	9	6
Expected frequency	66.4433	39.2898	21.7920	11.4538	5.7674	5.2537

Table 3 . The 95 %bootstrap two-sided confidence intervals and corresponding widths using all intervals for the parameter in the number of red mites

Methods	Confidence intervals	Widths
PB	)1.4312, 1.9408(	0.5096
BB	(1.3598, 1.8787)	0.5189
BCa	(1.4168, 1.9179)	0.5011

Table 4 . The number of corn borer larvae per plant

Number of corn borer larvae	0	1	2	≥ 3
Observed frequency	188	83	36	17
Expected frequency	193.6521	79.5626	31.6015	19.1838

Table 5 . The 95 %bootstrap two-sided confidence intervals and corresponding widths using all intervals for the parameter in the number of corn borer larvae per plant

Methods	Confidence intervals	Widths
PB	(2.2174, 2.7980)	0.5806
BB	(2.1590, 2.7292)	0.5702
BCa	(2.1903, 2.7586)	0.5683

simulation study and while utilizing real data sets. Our findings provided simulation results that correspond to Flowers-Cano et al.'s (16) research work. Using a Monte Carlo Simulation, they compared the coverage of several bootstrap confidence intervals. According to their findings, the coverage probabilities of the BCa bootstrap confidence interval were frequently greater than those of the other confidence intervals.

This study's limitation is that none of the bootstrap confidence intervals were exact, but they were consistent, indicating that the probability of coverage approaches 0.95 as sample sizes increase. In addition, the computation of three bootstrap confidence intervals is difficult and computationally intensive.

However, there are a number of R packages available for computing bootstrap confidence intervals, including the boot package (21), the bootstrap package (30), the semEff package (31), and the BootES package (32). Since R is an open-source programming language, users are allowed to install these packages. Future research could concentrate on the development of confidence intervals for parameter functions, such as the population mean, dispersion index, and coefficient of variation. Additionally, it is important to note that there is currently a lack of research available on the topic of hypothesis testing for the parameter of the PS distribution. The study of these issues can be explored in future research.

References

1. Andrew FS, Michael RW. Practical business statistics. 8th ed. San Diego: Academic Press; 2022.
2. Siegel AF. Practical business statistics. 7th ed. London: Academic Press; 2017.
3. Hougaard P, Lee M-LT, Whitmore GA. Analysis of overdispersed count data by mixtures of Poisson variables and Poisson processes. Biometrics. 1997;53(4):1225-1238.
4. Ong S-H, Low Y-C, Toh K-K. Recent developments in mixed Poisson distributions. ASM Sci J. 2021.
5. McElduff F.C. Models for discrete epidemiological and clinical data [PhD Thesis]. London: University College London; 2012.
6. Tharshan R, Wijekoon P. A new mixed Poisson distribution for over-dispersed count data: theory and applications. Reliab: Theory Appl. 2022;17(1):33-51.
7. Shanker R. On Poisson-Sujatha distribution and its applications to model count data from biological sciences. Biom Biostat Int J. 2016;3(4):1-8.
8. Sankaran M. The discrete Poisson-Lindley distribution. Biometrics. 1970;1926(1):145-149.
9. Shanker R. Sujatha distribution and its applications. Stat Transit New Series. 2016;17(3):391-410.
10. Lindley DV. Fiducial distributions and Bayes' theorem. J R Stat Soc Series B. 1958;20(1):102-107.
11. Shanker R. Akash distribution and its applications. Int J Prob Stat. 2015;4(3):65-75.
12. Tan SH, Tan SB. The correct interpretation of confidence intervals. Proc Singapore Healthc. 2010;19(3):276-278.
13. Wood M. Statistical inference using bootstrap confidence intervals. Signif. 2004;1(4):180-182.
14. Chernick MR, LaBudde RA. An introduction to bootstrap methods to R. 1st ed. Singapore: John Wiley & Sons; 2011.
15. Reiser M, Yao L, Wang X, Wilcox J, Gray S. A Comparison of bootstrap confidence intervals for multi-level longitudinal data using Monte-Carlo simulation. In: Chen DG, Chen J. (eds) Monte-Carlo simulation-based statistical modeling. Springer; 2017.
16. Flowers-Cano RS, Ortiz-Gómez R, León-Jiménez JE, Rivera RL, Cruz LAP. Comparison of bootstrap confidence intervals using Monte Carlo simulations. Water. 2018;10(2).
17. Mostajeran A, Iranpanah N, Noorossana R. A new bootstrap based algorithm for Hotelling's T2 multivariate control chart. J Sci I R I. 2016;27(3):269-278.
18. Henningsen A, Toomet O. maxLik: a package for maximum likelihood estimation in R. Comput Stat. 2011;26(3):443-458.
19. Ihaka R, Gentleman R. R: a language for data analysis and graphics. J Comput Graph Stat. 1996;5(3):299-314.
20. Meeker WQ, Hahn GJ, Escobar LA. Statistical intervals: a guide for practitioners and researchers. 2nd ed. New Jersey: John Wiley and Sons; 2017.
21. Canty A, Ripley B. boot: bootstrap R (S-Plus) functions. R package version 1.3-28.1, 2022.
22. Efron B. The jackknife, the bootstrap, and other resampling plans, in CBMS-NSF regional conference series in applied mathematics, Philadelphia: SIAM; 1982.
23. Efron B, Tibshirani RJ. An introduction to the bootstrap. 1st ed. New York: Chapman and Hall; 1993.
24. Efron B. Better bootstrap confidence intervals. J Am Stat Assoc. 1987;82(297):171-185.
25. Davison AC, Hinkley DV. Bootstrap methods and their application. 1st ed. Cambridge: Cambridge University Press; 1997.
26. Ukoumunne OC, Davison AC, Gulliford MC, Chinn S. Non-parametric bootstrap confidence intervals for the intraclass correlation coefficient. Stat Med. 2003;22(24):3805-3821.
27. Bliss CI, Fisher RA. Fitting the negative binomial distribution to biological data. Biometrics. 1953;9(2):176-200.
28. Turhan NS. Karl Pearson's chi-square tests. Educ Res Rev. 2020;15(9):575-580.
29. McGuire JU, Brindley TA, Bancroft TA. The distribution of european corn borer larvae pyrausta nubilalis (Hbn.), in field corn. Biometrics. 1957;13(1):65-78.
30. Kostyshak S. bootstrap: functions for the book "An introduction to the bootstrap". R package version, 2019.6, 2022.
31. Murphy MV. semEff: automatic calculation of effects for piecewise structural equation models. R package version 0.6.1, 2022.
32. Kirby KN, Gerlanc D. BootES: an R package for bootstrap confidence intervals on effect sizes. Behav Res Methods. 2013;45(4):905-927.