บทความวิจัย

ช่วงความเชื่อมั่นของสัมประสิทธิ์การแปรผันสำหรับการแจกแจงปัวซงที่มีศูนย์มาก

ผึ้งพร ทองชมภู และ ธิดาเดียว มยุรีสวรรค์*
ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 08 2151 5595 อีเมล: tidadeaw@kku.ac.th DOI: 10.14416/j.kmutnb.2019.04.007
รับเมื่อ 1 มิถุนายน 2561 แก้ไขเมื่อ 19 มีนาคม 2562 ตอบรับเมื่อ 29 มีนาคม 2562 เผยแพร่ออนไลน์ 18 เมษายน 2562

บทคัดย่อ

งานวิจัยนี้เสนอช่วงความเชื่อมั่นของสัมประสิทธิ์การแปรผันสำหรับการแจกแจงปัวซงที่มีศูนย์มาก (ZIP) โดยใช้หลักการประมาณค่าพารามิเตอร์ของการแจกแจง ZIP ด้วยตัวประมาณภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimators; MLEs) และนำมาสร้างเป็นช่วงความเชื่อมั่นของสัมประสิทธิ์การแปรผันสำหรับการแจกแจง ZIP ประสิทธิภาพของช่วงความเชื่อมั่นพิจารณาจากความน่าจะเป็นคุ้มรวมและความกว้างเฉลี่ยของช่วงความเชื่อมั่น โดยใช้โปรแกรม R จำลองข้อมูลซ้ำเป็นจำนวนมาก ที่สัมประสิทธิ์การแปรผัน (κ) มีค่าตั้งแต่ 0.39 ถึง 3.32 กำหนดพารามิเตอร์ของการแจกแจง ZIP คือ ค่าเฉลี่ยของจำนวนครั้งของเหตุการณ์ที่สนใจในขอบเขตที่กำหนด (λ) เท่ากับ 5(5)25 สัดส่วนของข้อมูลที่เป็นศูนย์ (ω) เท่ากับ 0.1(0.1)0.9 และขนาดตัวอย่าง (n) เท่ากับ 30, 40, 50, 100 และ 200 ที่ระดับความเชื่อมั่น 0.90, 0.95 และ 0.99 ผลการวิจัยแสดงให้เห็นว่าที่ระดับความเชื่อมั่น 0.90 และ 0.95 ช่วงความเชื่อมั่นที่เสนอให้ค่าความน่าจะเป็นคุ้มรวมตรงตามเกณฑ์ที่กำหนดทุกกรณีของ λ เมื่อ n = 50, 100, 200 และ ω = 0.2–0.9 สำหรับที่ระดับความเชื่อมั่น 0.99 พบว่า ค่าความน่าจะเป็นคุ้มรวมตรงตามเกณฑ์ที่กำหนดทุกกรณีของ ω เมื่อ n = 200 และ λ = 20, 25 และผลการวิจัยแสดงให้เห็นว่าค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นลดลงเมื่อ ω เพิ่มขึ้น อย่างไรก็ตาม เมื่อ n เพิ่มขึ้น พบว่า ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นลดลงทุกระดับของ λ, ω

คำสำคัญ: ช่วงความเชื่อมั่น, สัมประสิทธิ์การแปรผัน, การแจกแจงปัวซงที่มีศูนย์มาก, ความน่าจะเป็นคุ้มรวม

การอ้างอิงบทความ: ผึ้งพร ทองชมภู และ ธิดาเดียว มยุรีสวรรค์, “ช่วงความเชื่อมั่นของสัมประสิทธิ์การแปรผันสำหรับการแจกแจงปัวซงที่มีศูนย์มาก,” *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 29, ฉบับที่ 4, หน้า 652–666, ต.ค.–ธ.ค. 2562.

Research Article

The Confidence Interval of the Coefficient of Variation for a Zero - Inflated Poisson Distribution

Phuengporn Thongchomphu and Tidadeaw Mayureesawan*
Department of Statistics, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand

* Corresponding Author, Tel. 08 2151 5595, E–mail: tidadeaw@kku.ac.th
DOI: 10.14416/j.kmutnb.2019.04.007
Received 1 June 2018; Revised 19 March 2019; Accepted 29 March 2019; Published online: 18 April 2019

Abstract

This research was to present the confidence interval of the coefficient of variation for Zero-Inflated Poisson (ZIP) distribution by using the principle parameter estimation of ZIP distribution with Maximum Likelihood Estimator and to create the confidence interval of the coefficient of variation for ZIP distribution. The efficiency of the interval was considered by the coverage probability and the average width of the confidence interval. The R program was employed to simulate the repetition of numerous data set, which coefficient of variation (κ) was at 0.39 to 3.32. The parameters of ZIP distribution were average number of events in a specified region (λ), equal to 5(5)25. The proportions of observed zero (ω) were 0.1(0.1)0.9 and the sample sizes (n) were 30, 40, 50, 100 and 200 with the confidence levels of 0.90, 0.95 and 0.99. The results showed that at the confidence levels of 0.90 and 0.95, the proposed confidence interval gave the coverage probability as criteria defined in all cases of λ when n = 50, 100, 200 and ω = 0.2–0.9. For the confidence level of 0.99, found that the coverage probability meets the defined criteria in all cases of ω when n = 200 and λ = 20, 25 In addition, the results showed that the average width of the proposed confidence interval decreases when ω increase. However, as n increases, the average width of the confidence interval decreases for all levels of λ and ω.

Keywords: Confidence Interval, Coefficient of Variation, Zero-Inflated Poisson Distribution, Coverage Probability

Please cite this article as: P. Thongchomphu and T. Mayureesawan, "The confidence interval of the coefficient of variation for a zero - inflated Poisson distribution," *The Journal of KMUTNB*, vol. 29, no. 4, pp. 652–666, Oct.–Dec. 2019.

1. Introduction

Poisson distribution is a discrete probability distribution of a given number of events occurring in a fixed interval of time or space such as the number of customers visiting a bank per minute on Monday during 10.00–10.30 am or the number of typo words in a page made by a typist. The probability mass function (pmf) is given as in Equation (1).

$$f(x\ ;\ \lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}\ ;\ x = 0,\ 1,\ 2,\ \dots, \qquad (1)$$

where X is a random variable which Poisson distribution has both mean and variance equal to λ; and $\lambda>0$. In the case that the interested events do not occur frequently, mean and variance are not equal. In the study of the traffic incident considering factors from the driving license database of the Department of Motor Vehicles (DMV) in California [1], it is found that the mean of number of accident occurrence per driver was less than the variance; this is called Overdispersion which is not in congruence with Poisson distribution. It is caused by the abundant samples of zero number. Poisson distribution has only one parameter of λ which is not sufficient for explaining the characteristics of the population so the distribution has to be used to enable to explain Overdispersion occur, which is a Zero-Inflated Poisson distribution (ZIP distribution).

ZIP distribution is mixed distribution between Bernoulli distribution and Poisson distribution. The probability mass function is given as in Equation (2).

$$f(x\ ;\ \lambda\ ,\ \omega) = \begin{cases} \omega + (1-\omega)e^{-\lambda}\ ;\ x = 0 \\ \dfrac{(1-\omega)e^{-\lambda}\lambda^{x}}{x!}\ ;\ x = 1, 2, 3, \dots \end{cases} \qquad (2)$$

The X is a random variable which is ZIP distribution with the mean = $(1-\omega)\,\lambda = \mu$ and the variance = $(1-\omega)\lambda + \left(\frac{\omega}{1-\omega}\right)((1-\omega)\lambda)^2$ when $\lambda > 0$ and $0 \le \omega \le 1$, the distribution with the characteristic of Overdispersion. ZIP distribution has been studied in abundant of research such as in medical field. Vandenbroek [2] conducted a study of the distribution of the number of times on bladder infection of male AIDS patients in Utrecht University Hospital and found that the data had the ZIP distribution. Bohning *et al*. [3] have done the research on dental epidemiology in Belo Horizonte city by studying the amount of the decayed, teeth extraction, and teeth filling in both deciduous and succedaneous teeth of children with DMFT (The decayed, missing and filled teeth index). The result showed that DMFT index shows ZIP distribution. Regarding the industrial filed, Xie *et al*. [4] have studied the nonconforming control chart based on ZIP distribution (c_{zip}-chart) and the findings revealed that c_{zip}-chart can well detect the changes when the manufacturing process shows the Zero-Inflated nonconformity. Tidadeaw and Vilasinee [5] have studied the nonconforming control chart based on ZIP distribution (c_{zip}-chart) by using Maximum Likelihood Estimator (MLE) which considers the upper edge of the data set only. Besides, Ketmee and Mayureesawan [6] have developed three new control charts for processes with non-conformities based on a ZIP distribution. In agricultural field, Yip [7] has studied the number of insects per leaf and this data set demonstrated ZIP distribution. Also, in actuarial science or insurance field, Boucher *et al*. [8] have adopted the model of ZIP distribution to develop the analytical model of insurance data.

When analyzing statistical data and observing its congruence with distribution, mean and variance are explained about the characteristics of data set. Mean is the one measure of the central tendency of a variable in the data set. Variance is the measure how far set of numbers are spread out from its mean in the set of data. However, to compare the dispersion of two sets of data, the variance cannot be used to explain especially when mean of each set of data is different. Then the ratio of absolute variation of the mean is employed or so-called the relative measure of dispersion. The absolute variation with most commonly used is the coefficient of variation.

Coefficient of variation (κ) is a relative measure of dispersion calculated from standard deviation divided by mean for the population data that mean= μ and variance = σ^2. Therefore, $\kappa = \sigma/\mu$ when $\mu \neq 0$. Coefficient of variation is used to compare two or more sets of dispersion, both data with different and same units. However, the data set with many coefficients of variation would have more dispersion than the one with less coefficient of variation. Coefficient of variation is broadly used such as coagulation test of the patients [9], a measure of stock risks [10], and a measure of ceramic hardening [11]. As well as, it is applied to weather forecasts on variations in velocity [12], rainfall trend forecast [13], earthquake prediction [14], and used to measure the accuracy of 3 categories of Micropipettes which pass the standard criteria of NCCLS, EP5-A2 and ISO 8655-6 [15]. Even though the coefficient of variation is beneficial in many studies, it is mostly point estimation which involves the use of sample data to calculate a single value. The estimate can be equal to the coefficient of variation of population or not. Also, there is the tendency of error from coefficient of variation of population depending on the random samples.

By this reason, there are many researchers interested in studying the confidence interval of the coefficient of variation for a Zero-Inflated Poisson distribution, for example;

Vangel [16] developed the confidence interval of the coefficient of variation for the normal distribution of McKay [17], or Modified McKay with the use of Pivotal Quantity Method, $Q = \frac{K^2(1+\kappa^2)}{(1+\theta K^2)\,\kappa^2}$ with θ function when $\theta = \frac{\nu}{\nu+1}$ and $\nu = n-1$. However, McKay's approach is effective when the sample size (n) is more than or equal to 10 and the population coefficient of variation (κ) is less than or equal to 0.33 [18]. Therefore, Vangel studied the small sample size by improving the interval of the coefficient of McKay's with θ function when $\theta = \frac{\nu}{\nu+1}\left[\frac{2}{\chi^2_{\nu,\alpha/2}}+1\right]$ and $\nu = n-1$. Then, This method was compared with McKay method with n = 5 and κ = 0.04 at the confidence coefficient of 0.95. It showed that Modified McKay method was more effective than McKay method, but the average width of the confidence intervals is similar.

Wararit [19] improved the interval of Vangel's coefficient by replacing sample's coefficient of variation with Maximum Likelihood Estimator: MLE for normal distribution. Then this method was compared to McKay and Vangel (Modified McKay) method with n = 10, 15, 25, 50 and 100, κ = 0.1, 0.2 and 0.3 at the confidence level of 0.90, 0.95. The findings revealed that Vangel and Panichkitkosolkul's method generated the coverage probability not less than the confidence level in every event. Panichkitkosolkul method had slightly higher coverage probability than Vangel. Meanwhile, McKay provided the coverage probability less than the specified confidence level in some events. When considering the average width, Panichkitkosolkul had the shortest average width of interval in every event.

Wararit [20] introduced the confidence intervals for the coefficient of variation of Poisson distribution developed from 4 methods of the confidence intervals of Poisson mean, namely; Wald: W, Wald Interval with Continuity Correction: WCC, Score: S, and Variance stabilizing: VS by identifying the sample size in the study as 10, 15, 25, 50 and 100, κ = 0.1, 0.2 and 0.3 at the confidence coefficient of 0.90 and 0.95. The coverage probability of the confidence intervals developed by WCC method was close to the identified confidence level the most. Moreover, all 4 methods provide a slight difference in average width. Therefore, if considering only the coverage probability, the confidence interval developed by WCC method has more effectiveness than W, S and VS methods.

For the study of the above-mentioned research, many researchers are interested in studying the confidence interval of the coefficient of variance for different distributions such as Normal distribution, Poisson distribution by bringing interesting and effective techniques to create a confidence of interval. So, the aim of this study is to present the confidence interval of the coefficient of variation for a ZIP distribution by using the parameter estimation principle of ZIP distribution with Maximum Likelihood Estimators.

The paper outline is as follows, in Section 2, we presented materials and methods. A simulation results are conducted in Section 3. Finally, some discussions and conclusions are suggested in Section 4.

2. Materials and Methods

In this section, we have reviewed the ZIP approximation confidence interval for the probability of the data with zero in the ZIP distribution presented by Numna [21] in Section 2.1. Later, Section 2.2 the confidence interval of the coefficient of variation for ZIP distribution.

2.1 The ZIP approximation confidence interval for the probability of the data with zero

Saranya and Jansakul [21] has studied the confidence interval of the proportions of observed zero (ω) for random variables with ZIP distribution in order to compare the coverage probability for the ZIP distribution based on the basis of the confidence interval of ω, the sample sizes (n) is defined as 25, 50, 100 and 200. The parameter of the ZIP distribution with the average number of events in a specified region (λ) is equal to 1.25, 1.50, 2.00 and 2.25 and is equal to 0.25, 0.35, 0.45, 0.55, 0.65 and 0.75. The confidence interval (1–α)100% of ω is given as in inequality (3) or (4).

$$1-\frac{\bar{X}-Z_{\alpha/2}\sqrt{\left\{\frac{1}{n}\left(\frac{(1-\omega)\ \mu+\omega\mu^2}{1-\omega}\right)\right\}}}{\lambda} \le \omega \le 1-\frac{\bar{X}+Z_{\alpha/2}\sqrt{\left\{\frac{1}{n}\left(\frac{(1-\omega)\ \mu+\omega\mu^2}{1-\omega}\right)\right\}}}{\lambda} \quad (3)$$

where $\mu=(1-\omega)\lambda$, $Z_{\alpha/2}$ is upper (1–α)100 percentile of the Standard Normal distribution.

$$1-\frac{\bar{X}-Z_{\alpha/2}\sqrt{\frac{E(X)+E(X)(\lambda-E(X))}{n}}}{\lambda} \le \omega \le 1-\frac{\bar{X}+Z_{\alpha/2}\sqrt{\frac{E(X)+E(X)(\lambda-E(X))}{n}}}{\lambda} \quad (4)$$

Practically, only the upper confidence interval would be calculated by substituting λ and $E(X)$ with Maximum Likelihood Estimator of $\lambda(\hat{\lambda})$ and the average of the samples $\bar{x}$ respectively, so the confidence interval $(1-\alpha)100\%$ of ω is given as in Equation (5).

$$CI_{\omega} = 1 - \frac{\bar{x} \pm Z_{\alpha/2}\sqrt{\dfrac{\bar{x} + \bar{x}(\hat{\lambda} - \bar{x})}{n}}}{\hat{\lambda}}. \quad (5)$$

When considering the sample size in each level, it would be found that the coverage probability is increased when ω is increases.

2.2 Proposed confidence interval of the coefficient of variation

The confidence interval of the coefficient of variation for Zero-Inflated Poisson (ZIP) distribution by using the principle parameter estimation of ZIP distribution with Maximum Likelihood Estimates create the confidence interval of the coefficient of variation for ZIP distribution. Therefore, we can derive the confidence interval of the coefficient of variation for a ZIP distribution based on the above confidence interval for the proportions of observed zero (ω).

Define X as the random variable with ZIP distribution, the population coefficient of variation is given as in Equation (6).

$$\kappa = \frac{\sigma}{\mu} = \frac{\sqrt{(1-\omega)\lambda + \left(\dfrac{\omega}{1-\omega}\right)((1-\omega)\lambda)^2}}{(1-\omega)\,\lambda} = \sqrt{\frac{1+\omega\lambda}{(1-\omega)\lambda}} \quad (6)$$

The sample estimate of κ is given as in Equation (7).

$$\hat{\kappa} = \frac{s}{\bar{x}} = \frac{\sqrt{(1-\tilde{\omega})\tilde{\lambda} + \left(\dfrac{\tilde{\omega}}{1-\tilde{\omega}}\right)((1-\tilde{\omega})\tilde{\lambda})^2}}{(1-\tilde{\omega})\,\tilde{\lambda}} = \sqrt{\frac{1+\tilde{\omega}\tilde{\lambda}}{(1-\tilde{\omega})\tilde{\lambda}}}, \quad (7)$$

when $\tilde{\omega}$ and $\tilde{\lambda}$ are the Maximum Likelihood Estimators (MLEs) for λ and ω, respectively.

The MLEs of the parameters λ and ω are given as in Equations (8) and (9).

$$\tilde{\lambda} = \bar{x}^{+}(1 - e^{\tilde{\lambda}}) \quad (8)$$

$$\tilde{\omega} = 1 - \frac{\bar{x}}{\tilde{\lambda}} \quad (9)$$

From the Equation (9), there would be [22] in Equation (10).

$$\bar{x} = (1-\tilde{\omega})\,\tilde{\lambda}, \quad (10)$$

where $\hat{\lambda}$ is the mean of primary samples for calculating $\tilde{\lambda}$, which is equal to $\sum_{i=1}^{n} X_i \big/ n$,

$\bar{x}^{+}$ is the sample mean which is greater than zero ($x > 0$),

$\bar{x}$ is the sample mean.

From inequality (3), the confidence interval of the coefficient of variation for a ZIP distribution can be derived as follows:

$$1-\alpha = P\left(1 - \frac{\bar{X} - Z_{\alpha/2}\sqrt{\left\{\dfrac{1}{n}\left(\dfrac{(1-\omega)\mu + \omega\mu^2}{1-\omega}\right)\right\}}}{\lambda} \le \omega \le 1 - \frac{\bar{X} + Z_{\alpha/2}\sqrt{\left\{\dfrac{1}{n}\left(\dfrac{(1-\omega)\mu + \omega\mu^2}{1-\omega}\right)\right\}}}{\lambda}\right). \quad (11)$$

Multiplying inequality (11) by ω, we have

$$= P\left(\lambda\left[1-\frac{\bar{X}-Z_{\alpha/2}\sqrt{\left\{\frac{1}{n}\left(\frac{(1-\omega)\mu+\omega\mu^2}{1-\omega}\right)\right\}}}{\lambda}\right] \le \omega\lambda \le \lambda\left[1-\frac{\bar{X}+Z_{\alpha/2}\sqrt{\left\{\frac{1}{n}\left(\frac{(1-\omega)\mu+\omega\mu^2}{1-\omega}\right)\right\}}}{\lambda}\right]\right). \quad (12)$$

Adding 1 to inequality (12). We get

$$= P\left(1+\lambda-\bar{X}-Z_{\alpha/2}\sqrt{\left\{\frac{1}{n}\left(\frac{(1-\omega)\mu+\omega\mu^2}{1-\omega}\right)\right\}} \le 1+\omega\lambda \le 1+\lambda-\bar{X}+Z_{\alpha/2}\sqrt{\left\{\frac{1}{n}\left(\frac{(1-\omega)\mu+\omega\mu^2}{1-\omega}\right)\right\}}\right). \quad (13)$$

Taking the square root to inequality (13). We have

$$= P\left(\sqrt{\frac{1+\lambda-\bar{X}-Z_{\alpha/2}\sqrt{\frac{1}{n}\left\{(1-\omega)\lambda+(1-\omega)\lambda\left(\lambda-(1-\omega)\lambda\right)\right\}}}{(1-\omega)\lambda}} \le \sqrt{\frac{1+\omega\lambda}{(1-\omega)\lambda}} \le \sqrt{\frac{1+\lambda-\bar{X}+Z_{\alpha/2}\sqrt{\frac{1}{n}\left\{(1-\omega)\lambda+(1-\omega)\lambda\left(\lambda-(1-\omega)\lambda\right)\right\}}}{(1-\omega)\lambda}}\right). \quad (14)$$

Thus, the confidence interval (1–α)100% of the coefficient of variation for a ZIP distribution [inequality (14)] is obtained, which is

$$\sqrt{\frac{1+\lambda-\bar{X}-Z_{\alpha/2}\sqrt{\frac{1}{n}\left\{(1-\omega)\lambda+(1-\omega)\lambda\left(\lambda-(1-\omega)\lambda\right)\right\}}}{(1-\omega)\lambda}} \le \kappa \le \sqrt{\frac{1+\lambda-\bar{X}+Z_{\alpha/2}\sqrt{\frac{1}{n}\left\{(1-\omega)\lambda+(1-\omega)\lambda\left(\lambda-(1-\omega)\lambda\right)\right\}}}{(1-\omega)\lambda}}, \quad (15)$$

where $\kappa = \sqrt{\frac{1+\omega\lambda}{(1-\omega)\lambda}}$.

Practically, from inequality (15), the parameter λ and ω are unknown, therefore they are estimated by the Maximum Likelihood Estimators of λ and ω, respectively.

The proposed confidence interval (1–α)100% of the parameter κ for the data with ZIP distribution created by this research is

$$\sqrt{\frac{1+\tilde{\omega}\tilde{\lambda}-Z_{\alpha/2}\sqrt{\frac{1}{n}\left\{(1-\tilde{\omega})\,\tilde{\lambda}+(1-\tilde{\omega})\,\tilde{\lambda}\,(\tilde{\lambda}-(1-\tilde{\omega})\,\tilde{\lambda})\right\}}}{(1-\tilde{\omega})\,\tilde{\lambda}}} \le \kappa \le \sqrt{\frac{1+\tilde{\omega}\tilde{\lambda}+Z_{\alpha/2}\sqrt{\frac{1}{n}\left\{(1-\tilde{\omega})\,\tilde{\lambda}+(1-\tilde{\omega})\,\tilde{\lambda}\,(\tilde{\lambda}-(1-\tilde{\omega})\,\tilde{\lambda})\right\}}}{(1-\tilde{\omega})\,\tilde{\lambda}}}. \quad (16)$$

From inequality (16), the lower limit will be set under the following condition $P(L \le \kappa \le U) = 1-\alpha$ when

$$\frac{1+\tilde{\omega}\tilde{\lambda}-Z_{\alpha/2}\sqrt{\frac{1}{n}\left\{(1-\tilde{\omega})\tilde{\lambda}+(1-\tilde{\omega})\tilde{\lambda}\left(\tilde{\lambda}-(1-\tilde{\omega})\tilde{\lambda}\right)\right\}}}{(1-\tilde{\omega})\tilde{\lambda}} \ge 0.$$

Then,

$$\left(1+\tilde{\omega}\tilde{\lambda}\right)^2 \ge (Z_{\alpha/2})^2\frac{1}{n}\left\{(1-\tilde{\omega})\tilde{\lambda}+(1-\tilde{\omega})\tilde{\lambda}\left(\tilde{\lambda}-(1-\tilde{\omega})\tilde{\lambda}\right)\right\}.$$

We have in inequality (17).

$$n \ge \left(\frac{Z_{\alpha/2}}{\hat{\kappa}}\right)^2 \quad (17)$$

where the estimated coefficient of variation $\hat{\kappa} = \sqrt{(1+\tilde{\omega}\tilde{\lambda})/((1-\tilde{\omega})\tilde{\lambda})}$ and $Z_{\alpha/2}$ is upper (1–α)100 percentile of the Standard Normal distribution.

2.3 Performances of estimators

The criterions in comparing the confidence intervals of the coefficient of variation for Zero-

Inflated Poisson distribution, are the coverage probability and the average width of the confidence interval. The coverage probability and the average width of the confidence interval are calculated as follows:

2.3.1 Coverage probability

Estimation of the Coverage Probability (CP) would be considered whether or not the confidence interval generated can cover actual coefficient of variation value. The coverage probability is calculated as following [Equation (18)]

$$CP = \frac{\sum_{i=1}^{m} C_i}{m} \ ; \ i = 1,\ 2,\ ...,\ m \ , \tag{18}$$

whereas

$$C_i \begin{cases} 1 \ ; \ \mathrm{L_i} \le \kappa \le U_i \ ; \ i \ = \ 1,\ 2,\ ...,\ m \\ 0 \ ; \ \text{otherwise} \end{cases}$$

when C_i is an index showing the confidence interval covering the parameters of cycle i,

U_i is upper limit of the confidence interval on cycle i,

L_i is lower limit of the confidence interval on cycle i,

m is the number of cycles in the simulations,

κ is the population coefficient of variation.

The reliability of the confidence interval depends on the coverage probability whether or not it is equal to the defined confidence coefficient. If the coverage probability is greater than the confidence coefficient, the confidence interval gained is too wide which results in an exaggerated conclusion. If the coverage probability is lower than the defined confidence coefficient, the confidence interval gained is too narrow affecting the result which distorts the truth. Therefore, the coverage probability should be simultaneously considered on average width. As the good confidence interval, the coverage probability should be similar to the defined confidence level. Therefore, the researcher has defined the consideration criteria using the confidence interval criteria (1–α)100% of p when p is the proportion of the interesting population for the study and $\hat{p}$ is the proportion of the interesting samples. The confidence interval (1–α)100% of p is given as

$$\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \ \le \ p \ \le \ \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}. \tag{19}$$

Applying inequality (19) to the research, having the confidence interval (1–α)100% of CP when CP is the proportion of the confidence interval covering the parameters or the coverage probability, $\hat{CP}$ is the estimation of the proportion of the confidence interval covering the parameters or confidence coefficient, and m is the number of cycles in the repeated simulations which the confidence interval (1–α)100% of CP is given as in inequality (20).

$$\hat{CP} - Z_{\alpha/2}\sqrt{\frac{\hat{CP}\ (1-\hat{CP})}{m}} \ \le \ CP \ \le \ \hat{CP} + Z_{\alpha/2}\sqrt{\frac{\hat{CP}\ (1-\hat{CP})}{m}} \tag{20}$$

The efficiency consideration of the created confidence interval of the coefficient of variation for ZIP distribution would be considered with the coverage probability of parameter κ and the average width of the confidence by the consideration criteria of the coverage probability of the parameter κ which meets the criteria. At the confidence level 0.90, 0.95 and 0.99, the coverage probability has to be between [0.8984, 0.9016], [0.9486, 0.9514] and [0.9892, 0.9908], respectively.

2.3.2 Average width of the confidence interval

Estimation of the average width of the confidence interval (AW) is calculated as following [Equation (21)]:

$$AW = \frac{\sum_{i=1}^{m}(\mathrm{U}_i - \mathrm{L}_i)}{m} \;;\; i = 1, 2, ..., \mathrm{m}. \qquad (21)$$

3. Simulation Results

In this section, we have reported the results of using Monte Carlo simulations to investigate the estimated coverage probabilities and their average width of the confidence interval. The R program is employed to simulate the repetition of 100,000 data set. The coefficient of variation (κ) range from 0.39 to 3.32. The average number of events in a specified space (λ) are studied in 5, 10, 15, 20 and 25. The proportions of observed zero (ω) are in range of 0.1-0.9: every single addition was 0.1. The sample sizes (n) are 30, 40, 50, 100 and 200 with the confidence coefficient are 0.90, 0.95 and 0.99. Table 1–3 show estimated coverage probabilities of the confidence interval and their average width of the confidence interval for a ZIP distribution at 1–α = 0.90, 0.95 and 0.99, respectively.

From the simulation results in Table 1–2, when the average number of events in a specified region (λ) is equal to 5, the estimated coverage probabilities meet the criteria in many cases; such as n = 30, ω = 0.2(0.1)0.4; n = 40, ω = 0.2(0.1)0.8; as n = 50, 100, 200, ω = 0.2(0.1)0.9. For λ = 10 and 15, the estimated coverage probabilities meet the criteria when n = 30, ω =0.2(0.1)0.7; n =40, 50, 100, 200, ω = 0.2(0.1)0.9. In addition, when λ = 20 and 25, the estimated coverage probabilities meet the criteria in all cases of n and ω = 0.2(0.1)0.9.

Table 1 The coverage probabilities (CP) and the average widths of the confidence interval (AW) (in parenthesis) at the confidence coefficient of 0.90

λ	κ	ω	n = 30	40	50	100	200
5	0.58	0.1	0.9533*	0.9567	0.9576	0.9571	0.9586
			(0.3126)**	(0.2676)	(0.2379)	(0.1662)	(0.1169)
	0.71	0.2	***0.8992***	***0.9011***	***0.9005***	***0.8998***	***0.9015***
			(0.3081)	(0.2651)	(0.2360)	(0.1656)	(0.1167)
	0.85	0.3	***0.8990***	***0.9014***	***0.9014***	***0.8996***	***0.8998***
			(0.3055)	(0.2634)	(0.2350)	(0.1653)	(0.1166)
	1.00	0.4	***0.8989***	***0.9004***	***0.8999***	***0.9003***	***0.9014***
			(0.3039)	(0.2624)	(0.2343)	(0.1651)	(0.1165)
	1.18	0.5	0.9077	***0.9011***	***0.8991***	***0.9008***	***0.9006***
			(0.3028)	(0.2617)	(0.2338)	(0.1649)	(0.1165)
	1.41	0.6	0.9075	***0.9016***	***0.8986***	***0.8996***	***0.8993***
			(0.3021)	(0.2612)	(0.2334)	(0.1648)	(0.1164)
	1.73	0.7	0.9040	***0.9007***	***0.8999***	***0.9012***	***0.8991***
			(0.3015)	(0.2608)	(0.2332)	(0.1647)	(0.1164)
	2.24	0.8	0.9096	***0.9001***	***0.8992***	***0.8998***	***0.8986***
			(0.3010)	(0.2605)	(0.2330)	(0.1646)	(0.1163)
	3.32	0.9	0.9053	0.9043	***0.8987***	***0.9010***	***0.9013***
			(0.3006)	(0.2603)	(0.2328)	(0.1645)	(0.1163)
10	0.47	0.1	0.9322	0.9302	0.9384	0.9389	0.9401
			(0.3214)	(0.2723)	(0.2410)	(0.1672)	(0.1172)
	0.61	0.2	***0.8989***	***0.9005***	***0.8999***	***0.9005***	***0.9014***
			(0.3109)	(0.2669)	(0.2373)	(0.1661)	(0.1169)
	0.76	0.3	***0.9014***	***0.9000***	***0.8994***	***0.9011***	***0.8994***
			(0.3068)	(0.2642)	(0.2355)	(0.1655)	(0.1167)
	0.91	0.4	***0.9014***	***0.9016***	***0.9002***	***0.9010***	***0.8993***
			(0.3046)	(0.2628)	(0.2346)	(0.1652)	(0.1165)
	1.10	0.5	***0.8989***	***0.9011***	***0.8998***	***0.9001***	***0.9008***
			(0.3032)	(0.2620)	(0.2340)	(0.1650)	(0.1165)
	1.32	0.6	***0.9004***	***0.8999***	***0.8994***	***0.9008***	***0.9015***
			(0.3023)	(0.2614)	(0.2335)	(0.1648)	(0.1164)
	1.63	0.7	***0.9003***	***0.9012***	***0.9004***	***0.8998***	***0.8991***
			(0.3016)	(0.2609)	(0.2332)	(0.1647)	(0.1164)
	2.12	0.8	0.9044	***0.9005***	***0.9001***	***0.8985***	***0.9014***
			(0.3011)	(0.2606)	(0.2330)	(0.1646)	(0.1164)
	3.16	0.9	0.9050	***0.9009***	***0.9004***	***0.8989***	***0.9007***
			(0.3006)	(0.2603)	(0.2328)	(0.1645)	(0.1163)

Table 1 The coverage probabilities (*CP*) and the average widths of the confidence interval (*AW*) (in parenthesis) at the confidence coefficient of 0.90 (*Continued*)

λ	κ	ω	n = 30	40	50	100	200
15	0.43	0.1	0.9723	0.9051	0.9310	0.9162	0.9228
			(0.3302)	(0.2753)	(0.2432)	(0.1678)	(0.1174)
	0.58	0.2	***0.9002***	***0.9008***	***0.9005***	***0.9000***	***0.9000***
			(0.3120)	(0.2678)	(0.2379)	(0.1663)	(0.1169)
	0.72	0.3	***0.9007***	***0.9010***	***0.9000***	***0.9005***	***0.9003***
			(0.3074)	(0.2646)	(0.2358)	(0.1656)	(0.1167)
	0.88	0.4	***0.9004***	***0.9012***	***0.9000***	***0.9003***	***0.9009***
			(0.3049)	(0.2630)	(0.2347)	(0.1652)	(0.1166)
	1.06	0.5	***0.9008***	***0.9010***	***0.8999***	***0.9002***	***0.8990***
			(0.3034)	(0.2621)	(0.2340)	(0.1650)	(0.1165)
	1.29	0.6	***0.9004****	***0.9013***	***0.8994***	***0.9010***	***0.8995***
			(0.3024)**	(0.2614)	(0.2336)	(0.1648)	(0.1164)
	1.60	0.7	***0.9005***	***0.9011***	***0.8996***	***0.8988***	***0.9000***
			(0.3017)	(0.2609)	(0.2332)	(0.1647)	(0.1164)
	2.08	0.8	***0.8992***	***0.8996***	***0.9013***	***0.8999***	***0.9011***
			(0.3011)	(0.2606)	(0.2330)	(0.1646)	(0.1164)
	3.11	0.9	0.9065	***0.9001***	***0.8999***	***0.8992***	***0.8995***
			(0.3007)	(0.2603)	(0.2328)	(0.1645)	(0.1163)
20	0.41	0.1	0.9680	0.9022	0.8852	0.9061	0.9116
			(0.3288)	(0.2780)	(0.2443)	(0.1682)	(0.1175)
	0.56	0.2	***0.9014***	***0.8990***	***0.9001***	***0.9000***	***0.8995***
			(0.3134)	(0.2682)	(0.2383)	(0.1664)	(0.1170)
	0.71	0.3	***0.8988***	***0.8998***	***0.8990***	***0.9012***	***0.9007***
			(0.3078)	(0.2648)	(0.2360)	(0.1656)	(0.1167)
	0.87	0.4	***0.9001***	***0.9010***	***0.9005***	***0.9010***	***0.8990***
			(0.3051)	(0.2631)	(0.2348)	(0.1652)	(0.1166)
	1.05	0.5	***0.8998***	***0.9002***	***0.9008***	***0.8991***	***0.8999***
			(0.3035)	(0.2621)	(0.2341)	(0.1650)	(0.1165)
	1.27	0.6	***0.9009***	***0.9007***	***0.9003***	***0.8997***	***0.8995***
			(0.3025)	(0.2615)	(0.2336)	(0.1648)	(0.1164)
	1.58	0.7	***0.8993***	***0.8994***	***0.8992***	***0.8991***	***0.8989***
			(0.3017)	(0.2610)	(0.2333)	(0.1647)	(0.1164)
	2.06	0.8	***0.9000***	***0.8998***	***0.8994***	***0.8993***	***0.9005***
			(0.3011)	(0.2606)	(0.2330)	(0.1646)	(0.1164)
	3.08	0.9	***0.9005***	***0.9000***	***0.9012***	***0.8994***	***0.9003***
			(0.3007)	(0.2603)	(0.2328)	(0.1645)	(0.1163)
25	0.39	0.1	0.9679	0.9398	0.8672	0.9033	0.9050
			(0.3320)	(0.2803)	(0.2451)	(0.1685)	(0.1176)
	0.55	0.2	***0.9009***	***0.9007***	***0.9014***	***0.9001***	***0.8992***
			(0.3141)	(0.2684)	(0.2386)	(0.1664)	(0.1170)
	0.70	0.3	***0.9006***	***0.9011***	***0.9001***	***0.9008***	***0.9006***
			(0.3079)	(0.2650)	(0.2360)	(0.1657)	(0.1167)
	0.86	0.4	***0.9006***	***0.8999***	***0.9006***	***0.9001***	***0.9010***
			(0.3052)	(0.2632)	(0.2348)	(0.1653)	(0.1166)
	1.04	0.5	***0.8993***	***0.9007***	***0.9000***	***0.9008***	***0.9014***
			(0.3036)	(0.2622)	(0.2341)	(0.1650)	(0.1165)
	1.26	0.6	***0.9001***	***0.9000***	***0.9011***	***0.8998***	***0.8993***
			(0.3025)	(0.2615)	(0.2336)	(0.1648)	(0.1164)
	1.57	0.7	***0.9010***	***0.8995***	***0.9000***	***0.8995***	***0.8995***
			(0.3017)	(0.2610)	(0.2333)	(0.1647)	(0.1164)
	2.05	0.8	***0.9015***	***0.9010***	***0.9010***	***0.9010***	***0.8990***
			(0.3011)	(0.2606)	(0.2330)	(0.1646)	(0.1164)
	3.07	0.9	***0.8989***	***0.9013***	***0.9000***	***0.9009***	***0.9004***
			(0.3007)	(0.2603)	(0.2328)	(0.1645)	(0.1163)

Note: 1. * is *CP* and ** is *AW*

2. Bold and italic refer to that the coverage probability meets the criteria.

Table 2 The coverage probabilities (*CP*) and the average widths of the confidence interval (*AW*) (in parenthesis) at the confidence coefficient of 0.95

λ	κ	ω	n = 30	40	50	100	200
5	0.58	0.1	0.9846*	0.9845	0.9843	0.9854	0.9848
			(0.3805)**	(0.3234)	(0.2864)	(0.1990)	(0.1396)
	0.71	0.2	***0.9511***	***0.9497***	***0.9512***	***0.9493***	***0.9510***
			(0.3720)	(0.3185)	(0.2830)	(0.1980)	(0.1393)
	0.85	0.3	***0.9488***	***0.9492***	***0.9504***	***0.9492***	***0.9509***
			(0.3669)	(0.3156)	(0.2812)	(0.1974)	(0.1391)
	1.00	0.4	***0.9513***	***0.9492***	***0.9512***	***0.9509***	***0.9512***
			(0.3640)	(0.3138)	(0.2799)	(0.1970)	(0.1389)
	1.18	0.5	0.9534	***0.9491***	***0.9508***	***0.9506***	***0.9492***
			(0.3621)	(0.3127)	(0.2792)	(0.1967)	(0.1388)
	1.41	0.6	0.9465	***0.9492***	***0.9500***	***0.9504***	***0.9512***
			(0.3608)	(0.3118)	(0.2785)	(0.1965)	(0.1388)
	1.73	0.7	0.9531	***0.9490***	***0.9508***	***0.9509***	***0.9500***
			(0.3598)	(0.3112)	(0.2781)	(0.1963)	(0.1387)
	2.24	0.8	0.9522	***0.9496***	***0.9502***	***0.9508***	***0.9502***
			(0.3590)	(0.3107)	(0.2777)	(0.1962)	(0.1387)
	3.32	0.9	0.9454	0.9548	***0.9500***	***0.9514***	***0.9511***
			(0.3584)	(0.3102)	(0.2774)	(0.1961)	(0.1386)

Table 2 The coverage probabilities (*CP*) and the average widths of the confidence interval (*AW*) (in parenthesis) at the confidence coefficient of 0.95 (*Continued*)

λ	κ	ω	n = 30	40	50	100	200
10	0.47	0.1	0.9918	0.9802	0.9662	0.9725	0.9761
			(0.4001)	(0.3335)	(0.2924)	(0.2007)	(0.1402)
	0.61	0.2	***0.9492***	***0.9511***	***0.9494***	***0.9500***	***0.9505***
			(0.3785)	(0.3214)	(0.2855)	(0.1987)	(0.1395)
	0.76	0.3	***0.9500***	***0.9505***	***0.9509***	***0.9500***	***0.9502***
			(0.3696)	(0.3170)	(0.2823)	(0.1977)	(0.1392)
	0.91	0.4	***0.9499***	***0.9504***	***0.9502***	***0.9504***	***0.9505***
			(0.3654)	(0.3146)	(0.2805)	(0.1972)	(0.1390)
	1.10	0.5	***0.9491***	***0.9503***	***0.9506***	***0.9508***	***0.9500***
			(0.3629)	(0.3131)	(0.2795)	(0.1968)	(0.1389)
	1.32	0.6	***0.9510***	***0.9507***	***0.9498***	***0.9497***	***0.9503***
			(0.3612)	(0.3121)	(0.2787)	(0.1965)	(0.1388)
	1.63	0.7	***0.9504***	***0.9504***	***0.9506***	***0.9510***	***0.9513***
			(0.3600)	(0.3113)	(0.2782)	(0.1964)	(0.1387)
	2.12	0.8	0.9568	***0.9505***	***0.9495***	***0.9498***	***0.9509***
			(0.3591)	(0.3107)	(0.2778)	(0.1962)	(0.1387)
	3.16	0.9	0.9422	***0.9502***	***0.9514***	***0.9504***	***0.9504***
			(0.3584)	(0.3103)	(0.2774)	(0.1961)	(0.1386)
15	0.43	0.1	0.9900	0.9881	0.9615	0.9663	0.9650
			(0.4074)	(0.3406)	(0.2969)	(0.2018)	(0.1405)
	0.58	0.2	***0.9498***	***0.9488***	***0.9505***	***0.9505***	***0.9505***
			(0.3803)	(0.3231)	(0.2863)	(0.1990)	(0.1396)
	0.72	0.3	***0.9509***	***0.9506***	***0.9510***	***0.9500***	***0.9507***
			(0.3704)	(0.3177)	(0.2827)	(0.1979)	(0.1392)
	0.88	0.4	***0.9490***	***0.9489***	***0.9510***	***0.9502***	***0.9504***
			(0.3659)	(0.3150)	(0.2808)	(0.1972)	(0.1390)
	1.06	0.5	***0.9511***	***0.9495***	***0.9510***	***0.9503***	***0.9496***
			(0.3632)	(0.3133)	(0.2796)	(0.1968)	(0.1389)
	1.29	0.6	***0.9490****	***0.9501***	***0.9509***	***0.9510***	***0.9504***
			(0.3614)**	(0.3122)	(0.2788)	(0.1966)	(0.1388)
	1.60	0.7	***0.9511***	***0.9512***	***0.9508***	***0.9501***	***0.9505***
			(0.3601)	(0.3114)	(0.2782)	(0.1964)	(0.1387)
	2.08	0.8	***0.9501***	***0.9502***	***0.9501***	***0.9494***	***0.9506***
			(0.3592)	(0.3108)	(0.2778)	(0.1962)	(0.1387)
	3.11	0.9	0.9522	***0.9499***	***0.9505***	***0.9490***	***0.9503***
			(0.3584)	(0.3103)	(0.2775)	(0.1961)	(0.1386)
20	0.41	0.1	0.9864	0.9835	0.9905	0.9644	0.9539
			(0.4007)	(0.3466)	(0.3008)	(0.2026)	(0.1407)
	0.56	0.2	***0.9512***	***0.9504***	***0.9500***	***0.9501***	***0.9510***
			(0.3808)	(0.3241)	(0.2869)	(0.1992)	(0.1397)
	0.71	0.3	***0.9504***	***0.9514***	***0.9511***	***0.9501***	***0.9501***
			(0.3711)	(0.3181)	(0.2829)	(0.1980)	(0.1393)
	0.87	0.4	***0.9499***	***0.9495***	***0.9506***	***0.9502***	***0.9501***
			(0.3663)	(0.3152)	(0.2809)	(0.1973)	(0.1390)
	1.05	0.5	***0.9509***	***0.9506***	***0.9505***	***0.9506***	***0.9496***
			(0.3634)	(0.3134)	(0.2797)	(0.1969)	(0.1389)
	1.27	0.6	***0.9510***	***0.9498***	***0.9505***	***0.9503***	***0.9500***
			(0.3615)	(0.3123)	(0.2789)	(0.1966)	(0.1388)
	1.58	0.7	***0.9501***	***0.9509***	***0.9504***	***0.9506***	***0.9502***
			(0.3602)	(0.3114)	(0.2783)	(0.1964)	(0.1387)
	2.06	0.8	***0.9505***	***0.9514***	***0.9506***	***0.9510***	***0.9504***
			(0.3592)	(0.3108)	(0.2778)	(0.1962)	(0.1387)
	3.08	0.9	***0.9496***	***0.9502***	***0.9509***	***0.9502***	***0.9502***
			(0.3584)	(0.3103)	(0.2775)	(0.1961)	(0.1386)
25	0.39	0.1	0.9864	0.9809	0.9904	0.9437	0.9521
			(0.4059)	(0.3454)	(0.3046)	(0.2031)	(0.1409)
	0.55	0.2	***0.9510***	***0.9508***	***0.9504***	***0.9506***	***0.9502***
			(0.3820)	(0.3249)	(0.2874)	(0.1994)	(0.1397)
	0.70	0.3	***0.9492***	***0.9506***	***0.9498***	***0.9500***	***0.9491***
			(0.3715)	(0.3184)	(0.2831)	(0.1980)	(0.1393)
	0.86	0.4	***0.9503***	***0.9508***	***0.9498***	***0.9507***	***0.9506***
			(0.3664)	(0.3153)	(0.2810)	(0.1973)	(0.1391)
	1.04	0.5	***0.9503***	***0.9511***	***0.9494***	***0.9509***	***0.9505***
			(0.3635)	(0.3135)	(0.2797)	(0.1969)	(0.1389)
	1.26	0.6	***0.9502***	***0.9505***	***0.9512***	***0.9508***	***0.9500***
			(0.3616)	(0.3123)	(0.2789)	(0.1966)	(0.1388)
	1.57	0.7	***0.9505***	***0.9503***	***0.9501***	***0.9506***	***0.9500***
			(0.3602)	(0.3114)	(0.2783)	(0.1964)	(0.1387)
	2.05	0.8	***0.9506***	***0.9508***	***0.9502***	***0.9506***	***0.9509***
			(0.3592)	(0.3108)	(0.2778)	(0.1962)	(0.1387)
	3.07	0.9	***0.9500***	***0.9507***	***0.9492***	***0.9508***	***0.9508***
			(0.3585)	(0.3103)	(0.2775)	(0.1961)	(0.1386)

Note: 1. * is *CP* and ** is *AW*

2. Bold and italic refer to that the coverage probability meets the criteria.

Table 3 The coverage probabilities (*CP*) and the average widths of the confidence interval (*AW*) (in parenthesis) at the confidence coefficient of 0.99

λ	κ	ω	n = 30	40	50	100	200
5	0.58	0.1	0.9992*	0.9991	0.9989	0.9987	0.9986
			(0.5355)**	(0.4435)	(0.3877)	(0.2648)	(0.1845)
	0.71	0.2	***0.9905***	***0.9896***	***0.9893***	***0.9893***	***0.9902***
			(0.5087)	(0.4282)	(0.3786)	(0.2623)	(0.1837)
	0.85	0.3	***0.9903***	***0.9895***	***0.9897***	***0.9893***	***0.9896***
			(0.4921)	(0.4207)	(0.3737)	(0.2608)	(0.1832)
	1.00	0.4	***0.9895***	***0.9907***	***0.9904***	***0.9899***	***0.9892***
			(0.4853)	(0.4166)	(0.3708)	(0.2598)	(0.1829)
	1.18	0.5	0.9887	***0.9892***	***0.9898***	***0.9896***	***0.9898***
			(0.4806)	(0.4137)	(0.3688)	(0.2592)	(0.1827)
	1.41	0.6	0.9938	***0.9896***	***0.9902***	***0.9898***	***0.9905***
			(0.4772)	(0.4117)	(0.3674)	(0.2587)	(0.1825)
	1.73	0.7	0.9891	***0.9901***	***0.9902***	***0.9901***	***0.9899***
			(0.4748)	(0.4102)	(0.3663)	(0.2583)	(0.1824)
	2.24	0.8	0.9885	***0.9895***	***0.9897***	***0.9905***	***0.9902***
			(0.4729)	(0.4090)	(0.3655)	(0.2580)	(0.1823)
	3.32	0.9	0.9836	***0.9903***	***0.9895***	***0.9905***	***0.9902***
			(0.4715)	(0.4081)	(0.3648)	(0.2578)	(0.1822)
10	0.47	0.1	0.9984	0.9984	0.9989	0.997	0.9965
			(0.5636)	(0.4675)	(0.4076)	(0.2692)	(0.1858)
	0.61	0.2	***0.9892***	***0.9902***	***0.9894***	***0.9895***	***0.9894***
			(0.5258)	(0.4368)	(0.3844)	(0.2639)	(0.1843)
	0.76	0.3	***0.9894***	***0.9897***	***0.9903***	***0.9903***	***0.9903***
			(0.4985)	(0.4243)	(0.3761)	(0.2615)	(0.1835)
	0.91	0.4	***0.9905***	***0.9896***	***0.9901***	***0.9907***	***0.9900***
			(0.4889)	(0.4186)	(0.3721)	(0.2602)	(0.1831)
	1.10	0.5	***0.9905***	***0.9900***	***0.9902***	***0.9902***	***0.9901***
			(0.4824)	(0.4148)	(0.3696)	(0.2594)	(0.1828)
	1.32	0.6	***0.9902****	***0.9901***	***0.9896***	***0.9896***	***0.9903***
			(0.4782)**	(0.4123)	(0.3679)	(0.2588)	(0.1826)
	1.63	0.7	***0.9901***	***0.9906***	***0.9895***	***0.9902***	***0.9906***
			(0.4753)	(0.4105)	(0.3666)	(0.2584)	(0.1824)
	2.12	0.8	***0.9902***	***0.9898***	***0.9899***	***0.9899***	***0.9896***
			(0.4732)	(0.4092)	(0.3656)	(0.2581)	(0.1823)
	3.16	0.9	0.9910	***0.9897***	***0.9906***	***0.9903***	***0.9903***
			(0.4716)	(0.4081)	(0.3649)	(0.2578)	(0.1822)
15	0.43	0.1	0.9971	0.9975	0.9990	0.9920	0.9941
			(0.5735)	(0.4745)	(0.4208)	(0.2722)	(0.1867)
	0.58	0.2	***0.9904***	***0.9924***	***0.9904***	***0.9905***	***0.9906***
			(0.5360)	(0.4455)	(0.3876)	(0.2648)	(0.1845)
	0.72	0.3	***0.9901***	***0.9905***	***0.9902***	***0.9898***	***0.9904***
			(0.5019)	(0.4262)	(0.3773)	(0.2619)	(0.1836)
	0.88	0.4	***0.9901***	***0.9900***	***0.9900***	***0.9901***	***0.9902***
			(0.4905)	(0.4195)	(0.3727)	(0.2604)	(0.1831)
	1.06	0.5	***0.9902***	***0.9893***	***0.9905***	***0.9900***	***0.9904***
			(0.4833)	(0.4153)	(0.3699)	(0.2595)	(0.1828)
	1.29	0.6	***0.9900***	***0.9897***	***0.9904***	***0.9901***	***0.9899***
			(0.4787)	(0.4126)	(0.3680)	(0.2589)	(0.1826)
	1.6	0.7	***0.9896***	***0.9900***	***0.9901***	***0.9903***	***0.9902***
			(0.4756)	(0.4107)	(0.3667)	(0.2584)	(0.1824)
	2.08	0.8	***0.9903***	***0.9897***	***0.9900***	***0.9895***	***0.9905***
			(0.4733)	(0.4093)	(0.3657)	(0.2581)	(0.1823)
	3.11	0.9	***0.9897***	***0.9904***	***0.9906***	***0.9901***	***0.9907***
			(0.4716)	0.4082	0.3649	(0.2578)	(0.1822)
20	0.41	0.1	0.9917	0.9971	0.9988	0.9910	0.9903
			(0.5502)	(0.4902)	(0.4237)	(0.2745)	(0.1872)
	0.56	0.2	***0.9908***	***0.9916***	***0.9905***	***0.9898***	***0.9893***
			(0.5298)	(0.4504)	(0.3894)	(0.2653)	(0.1847)
	0.71	0.3	***0.9902***	***0.9899***	***0.9901***	***0.9901***	***0.9894***
			(0.5040)	(0.4274)	(0.3780)	(0.2621)	(0.1837)
	0.87	0.4	***0.9900***	***0.9907***	***0.9895***	***0.9896***	***0.9906***
			(0.4914)	(0.4200)	(0.3730)	(0.2606)	(0.1832)
	1.05	0.5	***0.9903***	***0.9902***	***0.9900***	***0.9902***	***0.9904***
			(0.4837)	(0.4156)	(0.3701)	(0.2596)	(0.1828)
	1.27	0.6	***0.9907***	***0.9903***	***0.9898***	***0.9906***	***0.9898***
			(0.4789)	(0.4127)	(0.3682)	(0.2589)	(0.1826)
	1.58	0.7	***0.9908***	***0.9903***	***0.9905***	***0.9904***	***0.9903***
			(0.4757)	(0.4108)	(0.3668)	(0.2584)	(0.1824)
	2.06	0.8	***0.9898***	***0.9903***	***0.9904***	***0.9899***	***0.9906***
			(0.4734)	(0.4093)	(0.3657)	(0.2581)	(0.1823)
	3.08	0.9	**0.9900**	**0.9898**	**0.9896**	**0.9898**	**0.9904**
			(0.4717)	(0.4082)	(0.3649)	(0.2578)	(0.1822)

Table 3 The coverage probabilities (CP) and the average widths of the confidence interval (AW) (in parenthesis) at the confidence coefficient of 0.99 (*Continued*)

λ	κ	ω	n = 30	40	50	100	200
25	0.39	0.1	0.9891*	0.9956	0.9982	0.9923	0.9897
			(0.5593)**	(0.4761)	(0.4188)	(0.2765)	(0.1877)
	0.55	0.2	***0.9905***	***0.9905***	***0.9903***	***0.9903***	***0.9895***
			(0.5350)	(0.4501)	(0.3910)	(0.2657)	(0.1848)
	0.70	0.3	***0.9905***	***0.9902***	***0.9899***	***0.9904***	***0.9900***
			(0.5053)	(0.4282)	(0.3785)	(0.2623)	(0.1837)
	0.86	0.4	***0.9902***	***0.9905***	***0.9901***	***0.9896***	***0.9898***
			(0.4921)	(0.4203)	(0.3733)	(0.2606)	(0.1832)
	1.04	0.5	***0.9904***	***0.9905***	***0.9907***	***0.9907***	***0.9907***
			(0.4840)	(0.4158)	(0.3702)	(0.2596)	(0.1828)
	1.26	0.6	***0.9908***	***0.9899***	***0.9907***	***0.9904***	***0.9898***
			(0.4791)	(0.4128)	(0.3682)	(0.2589)	(0.1826)
	1.57	0.7	***0.9907***	***0.9903***	***0.9901***	***0.9902***	***0.9906***
			(0.4758)	(0.4108)	(0.3668)	(0.2585)	(0.1824)
	2.05	0.8	***0.9904***	***0.9903***	***0.9901***	***0.9901***	***0.9900***
			(0.4734)	(0.4093)	(0.3657)	(0.2581)	(0.1823)
	3.07	0.9	***0.9906***	***0.9896***	***0.9907***	***0.9906***	***0.9904***
			(0.4717)	(0.4082)	(0.3649)	(0.2578)	(0.1822)

Note: 1. * is CP and ** is AW

2. Bold and italic refer to that the coverage probability meets the criteria.

From the simulation results in Table 3, when the average number of events in a specified region (λ) is equal to 5, the estimated coverage probabilities meet the criteria in many cases; such as n = 30, ω = 0.2(0.1)0.4; n = 40, 50, 100, 200, ω = 0.2(0.1)0.9. For λ = 10 the estimated coverage probabilities meet the criteria when n = 30, ω = 0.2(0.1)0.8; as n = 40, 50, 100, 200, ω = 0.2(0.1)0.9. For λ = 15 the estimated coverage probabilities meet the criteria when n = 40, ω = 0.3(0.1)0.9; as n = 30, 50, 100, 200, ω = 0.2(0.1)0.9. For λ = 20 the estimated coverage probabilities meet the criteria when n = 40, ω = 0.3(0.1)0.9; n = 30, 50, 100, ω = 0.2(0.1)0.9; as n = 200, ω = 0.1(0.1)0.9. In addition, λ = 25 the estimated coverage probabilities meet the criteria when n = 30, 40, 50, 100, ω = 0.2(0.1)0.9; as n = 200, all cases of ω.

Considering the Average Widths of the confidence interval (AW) from the simulation results in Table 1–3, we found that when λ is increased, the AW is slightly reduced at the same level of λ and n. In case of the same level of λ and ω, as n is increased, AW is decreased.

4. Discussion and Conclusions

We presented the confidence interval of the coefficient of variation for a ZIP distribution by using the principle parameter estimation of a ZIP distribution with the Maximum Likelihood Estimators. The coverage probability and the average width of confidence interval are considered as a good estimator in accordance with the criterion. We observed that the performance of the estimators depend on λ, ω and n. The confidence coefficient of 0.90 and 0.95, the confidence interval is good efficiency in all cases of λ when ω = 0.2(0.1)0.9 and n = 50, 100, 200. However, it is not good efficiency when ω = 0.1 and n = 30, 40. The confidence coefficient of 0.99, the confidence interval is good efficiency in all case of ω when = 0.2(0.1)0.9 and n = 50, 100. As λ = 20, 25, the confidence interval is good efficiency in all case of ω when n = 200. In Additionally, we found that the average width of the proposed confidence interval is decreased when ω are increased. The cause of the average width of the confidence interval decreased is; when sample's variance is increased, so the lower limit and upper limit of the confidence interval are approximate. However, as n is increased, the average width of this one is decreased for all level of λ and ω. Therefore, at the same level of λ and ω. n has a great influence on the average width of the proposed confidence interval. Also, in each level of λ, at the level of the same n, we found that ω have little effect on the average width.

5. Acknowledgments

The authors gratefully acknowledge Khon Kaen University for financial support of this research.

References

[1] J. Kuan, R. C. Peck, and M. K. Janke, "Statistical methods for traffic accident research," presented at the 1990 Taipei Symposium in Statistics, Taipei, Taiwan, June 28–30, 1990.

[2] J. Vandenbroek, "A score test for zero inflation in a Poisson-distribution," *International Biometric Society*, vol. 51, no. 2, pp. 738–743, 1995.

[3] D. Bohning, E. Dietz, P. Schlattmann, L. Mendonca, and U. Kirchner, "The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology," *Journal of the Royal Statistical Society Series A-Statistics in Society*, vol. 162, no. 2, pp. 195–209, 1999.

[4] M. Xie, B. He, and TN. Goh, "Zero-inflated Poisson model in statistical process control," *Computational Statistics & Data Analysis*, vol. 38, pp. 191–201, 2001.

[5] V. Peerajit and T. Mayureesawan, "Nonconforming control charts for zero-inflated Processes," M.S. thesis, Department of Applied Statistics, Faculty of Applied Statistics King Mongkut's University of Technology North Bangkok, 2009 (in Thai).

[6] N. Katemee and T. Mayureesawan, "Control charts for zero-inflated Poisson models," *Applied Mathematical Sciences*, vol. 6, no. 56, pp. 2791-2803, 2012.

[7] P. Yip, "Inference about the mean of a Poisson distribution in the presence of a nuisance parameter," *Australian Journal of Statistics*, vol. 30, pp. 299–306, 1988.

[8] J. P. Boucher, M. Denuit, and M. Guillen, "Number of accidents or number of claims? An approach with zero-inflated Poisson models for panel data," *Journal of Risk and Insurance*, vol. 76, no. 4, pp. 821–845, 2009.

[9] J. J. V. Veen, A. Gatt, A. E. Bowyer, P. C. Cooper, S. Kitchen, and M. Makris, "Calibrated automated thrombin generation and modified thromboelastometry in haemophilia A," *Journal of Thrombosis Research*, vol. 123, pp. 895–901, 2009.

[10] E. G. Miller and M. J. Karson, "Testing the equality of two coefficients of variation," in *American Statistical Association: Proceedings of the Business and Economics Secti on, Part I*, 1977, pp. 278–283.

[11] J. Gong and Y. Li, "Relationship between the estimated Weibull modulus and the coefficient of variation of the measured strength for ceramics," *Journal of the American Ceramic Society*, vol. 82, pp. 449–452, 1999.

[12] K. Ko, K. Kim, and J. Huh, "Variations of wind speed in time on Jeju Island, Korea," *Journal of Energy*, vol. 35, pp. 3381–3387, 2010.

[13] K. N. Krishnakumar, G. S. L. H. V. Prasadarao, and C. S. Gopakumar, "Rainfall trends in

twentieth century over Kerala, India." *Journal of Atmospheric Environment*, vol. 43, pp. 1940–1944, 2009.

[14] A. B. Lashak, M. Zare, H. Abedi, and M. Y. Radan, "The application of coefficient of variations in earthquake forecasting," *Journal of Seismology and Earthquake Engineering; Tehran*, vol. 11, no. 2, pp. 55–62, 2009.

[15] H. A. Majd, J. Hoseini, H. Tamaddon, and A. A. Baghban, "Comparison of the precision of measurements in three types of micropipettes according to NCCLS EP5-A2 and ISO 8655-6," *Journal of Paramedical Sciences (JPS)*, vol. 1, no. 3, pp. 2–8, 2010.

[16] M. G. Vangel, "Confidence intervals for a normal coefficient of variation," *American Statistician*, vol. 50, pp. 21–26, 1996.

[17] A. T. MaKay, "Distribution of the coefficient of variation and the extended "t" distribution," *Journal of the Royal Statistical Society*, vol. 95, no. 4, pp. 695–698, 1932.

[18] B. Iglewicz and R.H. Myers, "Comparisons of approximations to the percentage points of the sample coefficient of variation," *Technometrics*, vol. 12, pp. 166–169, 1970.

[19] W. Panichkitkosolkul, "Improved confidence intervals for a coefficient of variation of a normal distribution," *Thailand Statistician*, vol. 7, no. 2, pp. 193–199, 2009 (in Thai).

[20] W. Panichkitkosolkul, "A simulation comparision of new confidence intervals for the coefficient of variation of Poisson distribution," *Silpakorn University Science and Technology Journal*, vol. 4, no. 2, pp. 14–20, 2010 (in Thai).

[21] S. Numna and J. Naratip, "Analysis of extra zero counts using zero-inflated Poisson models," M.S. thesis, Department of Mathematics and Statistics, Faculty of Science, Prince of Songkla University, 2009 (in Thai).

[22] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate Discrete Distributions*, 3rd ed. New York: Wiley, 2005.