University of Tripoli

Faculty of Science
Department of Statistics

A Comparison of Some Confidence Intervals for Estimating the Poisson Coefficient of Variation: A Simulation Study

Najla Nouri Alkshaik

Supervisor: **Prof. Yousef M. Emhemmed**

Thesis was submitted in partial fulfillment of the requirements for the
Degree of Master in Statistics
Spring 2022

Declaration

I am *Najla Nuri Ibrahim Alkshaik* the undersigned hereby confirm that the work contained in this thesis / dissertation, unless otherwise referenced is the researcher's own work, and has not been previously submitted to meet requirements of an award at this University or any other higher education or research institution, I furthermore, cede copyright of this thesis dissertation in favor of University of Tripoli.

Student name: *Najla Nuri Ibrahim Alkshaik.*

Signature: ………………………..

Date:12/3/2023

Abstract

This thesis is aiming to demonstrate and compare the different methods of constructing a two-sided confidence intervals for the coefficient of variation (*CV*), or that, for signal–to-noise ratio (*SNR*) of a Poisson distribution based on bootstrapping simulating techniques.

As an introductory chapter of this thesis, chapter one is designed to include a comprehensive introduction. Chapter two is devoted to review aspects and definitions of all seven adopted methods of confidence intervals for the Poisson coefficient of variation (*CV*) as well as the confidence intervals for the signal–to-noise ratio (*SNR*) as a reciprocal of the *CV*.

Chapter three is devoted to the application part where simulation study is to be conducted to compare the performance of the four considered methods namely: the Wald with Continuity Correction (*WCC*), Wald Bootstrap method (*WaldB*), WaldZ (*WaldZ*) and Wald with Continuity Correction Bootstrap method (*WCCZ*)) of constructing a 95% two-sided confidence intervals for the Poisson coefficient of variation with different sample sizes and varying parameter values.

As last chapter of this thesis, chapter four is devoted to the final part of this study that illustrates the most important discussions and conclusions and then an outline of possible future work by which this study could be extended. Finally, the Matlab functions for the four confidence interval methods is to be given in subsequent chapters.

This thesis considered several confidence intervals for estimating the Poisson population coefficient of variation. A simulation study will be conducted to compare the performance of the four proposed confidence interval methods.

Data will be generated from Poisson distribution for CV = 0.05, 0.1, 0.15, …, 0.45, and 0.50 using Matlab software. The coverage probability and interval length for each confidence interval method will be calculated and reported.

Keywords: coefficient of variation (*CV*), confidence interval, coverage probability, expected length, Poisson distribution, bootstrap samples mean, and Signal-to-Noise Ratio (SNR).

بسم الله الرحمن الرحيم

إهداء خاص

نهدي هذا العمل المتواضع الى روح (الدكتور البهلول عمر شلابي) الذي إقترح هذا البحث وأشرف على تنفيذ جزء هام منه ولكن داهمه المرض وهو في خضم هذا البحث ولم يمهله القدر على تكملته. ندعوا له بواسع الرحمة والمغفرة وندعوا من الله أن يجعل هذا العمل وسائر الاعمال في مجال التعليم العالي التى قضى المرحوم فيها جل حياته أن تكون صدقة جارية على روحه الطاهرة

﴿ إِنَّا لِلَّهِ وَإِنَّا إِلَيْهِ رَاجِعُونَ ﴾

الأهداء

إلى من قاد البشرية وعقولهم إلى مرفأ الأمان ، معلم البشرية الأول

سيدنا محمد ﷺ ،،

إلى من وفاتهم المنيَّة ، وكانوا خير مثال لرب الأسرة
أبي وزوجي رحمهما الله وغفر لهم

إلى من وضع المولى – سبحانه وتعالى – الجنة تحت قدميها ،
ووقرها في كتابه العزيز...
أمي الحبيبة

إلى بذرة الفؤاد وأمل الغد ، بناتي
" ريان ، مرام ، جنان "

إلى إخوتي وأخواتي مصدر فخري وأبنائهم وبناتهم ،،
إلى من ربطني بهم علاقة صداقة .. وورد المحبة
إلى إخوة و أخوات جمعتني بهم الدراسة وميدان العمل ..
زملائي وزميلاتي الكرام ،،
إلى كل يدٍ وقلبٍ سارَ معي درب الإنجاز لأكون ،،،
إلى كل هؤلاء أهدي هذه الدراسة ، راجيا من الله أن تكون نافذة علم
وبطاقة معرفة .. وأن ينفعنا وينفع بنا ،،

شكر وتقدير

قال الله تعالى : (وَمَنْ يَشْكُرْ فَإِنَّمَا يَشْكُرُ لِنَفْسِهِ) لقمان 12

وقال رسوله الكريم ﷺ : (من لم يشكر الناس ، لم يشكر الله عز وجل)

أحمد الله تعالى حمداً كثيراً طيباً مباركاً ملئ السموات والأرض على ما أكرمني به من إتمام هذه الدراسة التي أرجو أن تنال رضاه .

ثم أتوجه بجزيل الشكر وعظيم الأمتنان إلى :

الدكتور الفاضل : يوسف محمد امحمد الأصيبعي

حفظه الله وأطال في عمره لتفضله الكريم بالأشراف على هذه الدراسة بعد وفاة الدكتور البهلول عمر شلابي رحمه الله وتكرمه بنصحي وتوجيهي حتى إتمام هذه الدراسة.

كما أتقدم بالشكر للدكتورة نادية القريقني لمساعدتها لي في العرض الأول وجميع أعضاء هيئة التدريس بقسم الأحصاء لما قدموه من دعم معنوي ووفقني الله واياكم

الباحثة

نجلاء نوري الكشيك

Contents

List of Tables

List of Figures

Chapter One

Introduction

Chapter 1

1.0 Introduction

The coefficient of variation (*CV*) for numerical measurements, also known as relative standard deviation (*RSD*), is defined as the ratio of the standard deviation (*SD*) to the mean (μ) of these measurements. The *CV* gives the standard deviation as proportion of the mean, and it is sometimes an informative quantity. For example, a value $SD = 10$ has little meaning unless we can compare it to something else. If SD is observed to be 10 and μ is observed to be 1000, then the amount of variation is small relative to the mean. However, if SD is observed to be 10 and μ is 5, then the variation is quite large relative to the mean. If we were studying the precision (variation in spread measurements) of a measuring instrument, the first case ($CV = 0.01$) might give quite acceptable precision but the second case ($CV = 2$) would be quite unacceptable (Wackerly, et al. 2008).

The *CV* is a unit-free measure of relative spread found to be very useful in descriptive studies (Panichchkikosolkul, 2009), also it is useful in finance and actuarial science by using it to measure the relative risks (Miller and Karson, 1977). The *CV* can be used to make comparisons across several populations that have different units of measurement, such as, for example, the variability of the weights of newborns, measured in grams, with the height of adults, measured in centimeters. The *CV* is not defined for a mean equal to zero, and it is unreliable for small means relative to the standard deviation. Even if the mean of the measurements is not zero, but the measurements contain both positive and negative values and the mean is close to zero, then the *CV* can be misleading.

It is very well known that the standard deviation represents noise and other interference and in some cases the standard deviation is not important in itself, but only in comparison to the mean. This gives rise to the term: Signal-to-Noise Ratio (*SNR*), which is equal to the converse of the coefficient of variation (i.e., *SNR*=1/*CV*).
It is commonly used in image processing (for examples, Tania, 2008; Jitendra, 2009; John, 2007), where the *SNR* of an image is usually calculated as the ratio of the mean pixel value to the standard deviation of the pixel values over a given neighborhood (F. George and B.M. G. Kibria, 2012). SNR measures how much signal has been corrupted by noise (McGibney and Smith, 1993).

In real life examples like image processing, Signal-to-Noise Ratio (*SNR*) describes the quality of a measurement. It is the ratio of the measured signal to the overall measured noise at that pixel. High *SNR* is particularly important in applications requiring precise light measurement. The detected photons in a CCD (Charged Coupled Device) camera or a photomultiplier tube follow a Poisson distribution, which is responsible forthe Photon Noise and determines the Signal to Noise Ratio of the acquired image (Lee, 2009; Willkinson and Schut, 1998). Since the *SNR* of Poisson distribution has special interest in imaging, we will be discussing different methods of confidence interval for *SNR* of Poisson distribution.

To test the significance of the *SNR*, a hypothesis test can be conducted and a confidence interval can be generated to reject or not reject the null hypothesis. Confidence intervals associated with point estimates provide more specific knowledge about the population characteristics than the p-values in the test of hypothesis (Visintainer and Tejani, 1998). The precision of a confidence interval can be determined through the width and coverage probability of the interval. Given constant coverage, as the width of the $(1-\alpha)$ 100% confidence interval decreases, the accuracy of the estimate increase (Kelley, 2007). The coverage level is the probability that the estimated interval will capturethe true *CV* or *SNR* value (Banik and Kibria, 2010).

There are various methods available for estimating the confidence interval for a population *CV* or *SNR*, such as, parametric, nonparametric, modified and bootstrapping (Banik and Kibria, 2010). The bootstrap approach is anon-parametric and computer-intensive tool used for estimating and making inferences about the parameters thatwas introduced by Efron (1979). It will be especially usefulbecause, unlike other methods, this technique does not require any assumptions to be made about the underlying population of interest (Banik and Kibria, 2010). Therefore,bootstrapping technique can be applied for estimating or hypothesis testing to all situations. This method is implemented by simulating an original data set then randomly selecting data several times with replacement to estimate the parameter of a distribution. For more information on the confidence interval for the *CV*, we refer Miller (1991); Sharma and Krishna (1994); Vangel (1996) Banik and Kibria (2011) and recently F. George and B.M. G. Kibria (2012) among others.

The literature on the confidence intervals for the *CV* or *SNR* of a Poisson distribution is very limited. The objective of this thesis is to propose some confidence interval estimators for *SNR* and find some good estimators for the practitioners. Six confidence intervals that already exist in literature for *CV* and they are considered for *SNR* by using the inverse relationship between *CV* and *SNR*. Since a theoretical comparison is not possible, a simulation study has been conducted to compare the performance of the interval estimators. Finally, based on the simulation results, the intervals with high coverage probability and smaller width were recommended for practitioners.

1.1 Basic Notation

Let $X_1, X_2, \cdots, X_n$ be an independently and identically distributed (iid) random sample of size n from a distribution with finite mean μ and standard deviation σ so that $CV = \sigma/\mu$ is the population coefficient of variation. When the parameters μ and σ or the underlying distribution are unknowns, the parameters μ and σ are estimated from the observed data. The estimated γ is then defined as $\widehat{CV} = \hat{\sigma}/\hat{\mu}$ where $\hat{\mu}$ and $\hat{\sigma}$ are the estimated values of μ and σ respectively.

1.2 Literature Review

The researchers often calculate $\widehat{CV}$ using sample values but they rarely do construct two-sided confidence intervals for CV (Mahmoud and Hassani, 2009).
Confidence interval estimation allows the researcher to have an idea about the precision of the point estimate rather than only a p-value for rejection or no rejection of a specified null hypothesis (Albatineh, et al., 2017). To do interval estimation on CV, one needs to make assumption about the shape of the population distribution. One also needs to know the distribution of $\widehat{CV}$. The exact distribution of the sample coefficient of variation from a normally distributed population is not easy and obtaining a two-sided confidence interval for CV in this case would require using the non-central *t* distribution and sequential techniques (Koopmans, et al., 1964). However, many researchers have already done good job for the inference of CV.

When the underlying population is normally distributed, several two-sided confidence intervals for CV are constructed and modified by many authors since 1932. For more information on the two-sided confidence intervals for CV, we refer to McKay (1932), Fieller (1932), Pearson (1932), Hendricks and Robey (1936), Koopmans, et al.

(1964), Umphrey (1983), Gregoire (1984), Bhoj and Ahsanullah (1993), Reid (1996), Vangel (1996), Tian, (2006), Rahim, et al. (2007), Panichchkikosolkul (2009), Banik and Kibria (2011), and recently Panichchkikosolkul (2017).

Sharma and Krishna (1994) developed two-sided confidence interval for the reciprocal of CV without making an assumption about the population distribution. The reciprocal of CV is called the population Signal-to-Noise Ratio (*SNR*), i.e., $SNR = \mu/\sigma = CV^{-1}$. In digital communications the *SNR* is a measure of the signal strength relative to background noise, while in quality control, the *SNR* represents the magnitude of the mean of a process compared to its variation. The *SNR* measures how much signal has been corrupted by noise, see McGibney and Smith (1993) for a discussion. Several two-sided confidence intervals that already exist in literature for *SNR* and they can be considered for CV by using the inverse relationship between *SNR* andγ, for example, see Sharma and Krishna (1994), Banik and Kibria (2010), George and Kibria (2012) and Albatineh, et al. (2014).

Panichkitkosolkul (2010) proposed four new two-sided confidence intervals for the coefficient of variation of the Poisson distribution based on obtaining two-sided confidence intervals for the Poisson mean.

1.3 Contribution of the Thesis

1. The thesis provides a comprehensive review of methods for setting confidence intervals for the Poisson coefficient of variation.
2. This thesis proposes three new confidence intervals for the coefficient of variation of a Poisson distribution based on obtaining confidence interval methods for the *SNR* of a Poisson distribution that reviewed by George and Kibria (2011).

Chapter Two

Statistical Methodology

Chapter 2

2.0 Statistical Methodology

2.1 Introduction

Let $X_1, X_2, \cdots, X_n$ be an independently and identically distributed random sample of size n from a Poisson distribution with mean λ. Then the population coefficient variation for Poisson distribution is $CV = 1/\sqrt{\lambda}$. In this chapter, a $(1-\alpha)100\%$ confidence intervals methods for CV will be reviewed, four existing methods and three proposed new methods. Following 7 methods will be considered.

2.2 Existing Methods

Wararit Panichkitkosolkul proposed four new confidence intervals for the coefficient of variation of Poisson distribution based on obtaining confidence intervals for the Poisson mean (Panichchkikosolkul, 2010). The four confidence intervals for the coefficient of variation of a Poisson distribution based on Wald (*W*), Wald with continuity correction (*Wcc*), Scores (*S*) and Variance stabilizing (*VS*) confidence interval are as follows:

$$CI_W = \left[\left(\sqrt{\bar{X} - z_{1-\alpha/2}\sqrt{\frac{\bar{X}}{n}}}\right)^{-1}, \left(\sqrt{\bar{X} + z_{1-\alpha/2}\sqrt{\frac{\bar{X}}{n}}}\right)^{-1}\right],$$

$$CI_{Wcc} = \left[\left(\sqrt{\bar{X} - z_{1-\alpha/2}\sqrt{\frac{\bar{X}+0.5}{n}}}\right)^{-1}, \left(\sqrt{\bar{X} + z_{1-\alpha/2}\sqrt{\frac{\bar{X}+0.5}{n}}}\right)^{-1}\right],$$

$$CI_S = \left[\left(\sqrt{\bar{X} + \frac{\left(z_{1-\alpha/2}\right)^2}{2n} + z_{1-\alpha/2}\sqrt{\frac{4\bar{X}+\frac{\left(z_{1-\alpha/2}\right)^2}{n}}{4n}}}\right)^{-1}, \left(\sqrt{\bar{X} + \frac{\left(z_{1-\alpha/2}\right)^2}{2n} - z_{1-\alpha/2}\sqrt{\frac{4\bar{X}+\frac{\left(z_{1-\alpha/2}\right)^2}{n}}{4n}}}\right)^{-1}\right],$$

$$CI_{VS} = \left[\left(\sqrt{\bar{X} + \frac{\left(z_{1-\alpha/2}\right)^2}{4n} + z_{1-\alpha/2}\sqrt{\frac{\bar{X}}{n}}} \right)^{-1}, \left(\sqrt{\bar{X} + \frac{\left(z_{1-\alpha/2}\right)^2}{4n} - z_{1-\alpha/2}\sqrt{\frac{\bar{X}}{n}}} \right)^{-1} \right].$$

The simulation results showed that the (*WCC*) based trust period is more appropriate than the other three confidence intervals in terms of coverage probability and hence this method will be adopted along with other three proposed methods to form the four methods considered in the application part of this thesis.

2.3 Proposed Methods

The proposed new confidence intervals methods which are based on three confidence interval methods for the *SNR* of a Poisson distribution that reviewed by George and Kibria (2011) are as follows:

$$CI_{WaldB} = \left[\left(\sqrt{\bar{X} - Z_{\alpha/2}\sqrt{\frac{\bar{X}_B}{n}}} \right)^{-1}, \left(\sqrt{\bar{X} + Z_{\alpha/2}\sqrt{\frac{\bar{X}_B}{n}}} \right)^{-1} \right],$$

$$CI_{Waldz} = \left[\left(\sqrt{\bar{X} - T^*_{\alpha/2}\sqrt{\frac{\bar{X}}{n}}} \right)^{-1}, \left(\sqrt{\bar{X} + T^*_{1-\alpha/2}\sqrt{\frac{\bar{X}}{n}}} \right)^{-1} \right],$$

$$CI_{Wccz} = \left[\left(\sqrt{\bar{X} - T^*_{\alpha/2}\sqrt{\frac{\bar{X}_B + 0.5}{n}}} \right)^{-1}, \left(\sqrt{\bar{X} + T^*_{1-\alpha/2}\sqrt{\frac{\bar{X}_B + 0.5}{n}}} \right)^{-1} \right],$$

where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $T^*_{\alpha/2}$ and $T^*_{1-\alpha/2}$ are the $(\alpha/2)^{th}$ and $(1-\alpha/2)^{th}$ sample quintiles of $T_i^* = \frac{\bar{X}_i - \bar{X}_B}{\hat{\sigma}_B}$, where$\bar{X}_i^*$ is the i^{th} bootstrap sample mean, $\hat{\sigma}_B = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}(\bar{X}_i^* - \bar{X}_B)^2}$, $\bar{X}_B = \frac{1}{B}\sum_{i=1}^{B}\bar{X}_i^*$ and B is the number of bootstrap samples. The number of bootstrap samples is typically between 1000 and 2000.

Chapter Three

Demonstration of Simulated Data

Chapter3

3.0 Demonstration of Simulated Data

3.1 Introduction

Our attention is now turned to the application part of the thesis by applying the above mentioned confidence intervals for estimating the Poisson population coefficient of variation to different simulated data sets each with different sample sizes and different Poisson population parameter. This study considered several confidence intervals for estimating the Poisson population coefficient of variation. A simulation study will be conducted to compare the performance of the four adopted confidence interval methods. The generated data are from a Poisson distribution with a varying valus of the parameter λ i.e., with a varying valus of the Poisson population coefficient of variation $CV = 1/\sqrt{\lambda}$= 0.05, 0.1, 0.15, …,0.45, and 0.50 using Matlab software. The coverage probability and interval length for each confidence interval method will be calculated and reported.

Often, it is necessary to investigate the properties of a statistical procedure using simulation techniques. However, since a theoretical comparisons between confidence interval methods for the Poisson coefficient of variation is not possible, bootstrapping technique can be applied by simulating an original data set then randomly selecting data several times with replacement to estimate the unknown parameters. The simulation study will be conducted to compare the performance of only four considered confidence interval methods which are obtained under the same simulation conditions. The four two-sided confidence interval methods are: (1) Wald with Continuity Correction (*WCC*), (2) Wald Bootstrap method (*WaldB*), (3) (*WaldZ*), and (4) Wald with Continuity Correction Bootstrap method (*WccZ*).

3.2 Simulation Steps and Criteria

A “good” confidence interval should have coverage close to the nominal confidence level and short length. The performance of the estimated coverage probabilities of the confidence intervals (1) to (4) and their estimated expected lengths will be examined by Bootstrap simulation.

To study the performance of the adopted four confidence interval methods, the coverage probability (*CP*) and the length (*L*) are to be considered. For each one of the considered four methods, a $(1-\alpha)100\%$confidence interval denoted by (*L*,*U*) which based on M replicates, the estimated coverage probability (ECP) and the expected length (EL) are to be obtained, where,

$$ECP = (\widehat{1-\alpha}) = \#(L \leq \gamma \leq U)/M$$

$$EL = \sum_{i=1}^{M} (U_i - L_i) \Big/ M.$$

3.2.1 The Estimated Lengths (EL) of The Confidence Intervals

To enable comparisons, data will be generated from Poisson distribution with λ = 400, 100, 44.44, 25, 16, 11.11, 8.16, 6.25, 4.93, 4 and sample sizes; *n* = 10, 15, 25, 50, 75, 100. The corresponding coefficient of variations are *CV*= 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50. The estimated 95 % two-sided confidence intervals of each coefficient of variation are then calculated along with their corresponding lengths. The estimated lengths (*EL*) are given in Tables 3.1 to 3.3 whereas their corresponding plots are given in Figs 3.1 to 3.3. All simulations are performed using programs written in the Matlab software, repeated 100,000 times in each case at confidence level 0.95. All calculations for this simulation study are based on Monte-Carlo Simulation and then double precision computations are adopted (i.e., 16 digit accuracy).

Table 3.1 below gives the estimated lengths (*EL*) of the confidence intervals for $\lambda = 400$ and confidence coefficient 0.95 of the four considered methods for sample sizes varying from 10 to 100.

Table 3.1 The Estimated Lengths (*EL*) for *CV*= 0.05 and Confidence Coefficient 0.95.

$CV = 0.05$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.00130000	0.00130000	0.00130000	0.00130000
15	0.00126650	0.00126580	0.00126260	0.00126330
25	0.00098095	0.00098034	0.00097919	0.00097996
50	0.00069351	0.00069308	0.00069284	0.00069320
75	0.00056623	0.00056588	0.00056569	0.00056606
100	0.00049034	0.00049003	0.00048994	0.00049023

As can be seen from Table 3.1, the estimated lengths of all confidence intervals of the four considered methods are very close to each other for each sample size and it is really difficult to distinguish between them, with a very slight preferable of the confidence interval that based on *WaldZ*, especially for large sample sizes (25 or more). This interpretation can also be seen clearly from Fig 3.1.Furthermore, it is very nature to notice that the expected lengths are decreased dramatically with the increase of sample sizes.

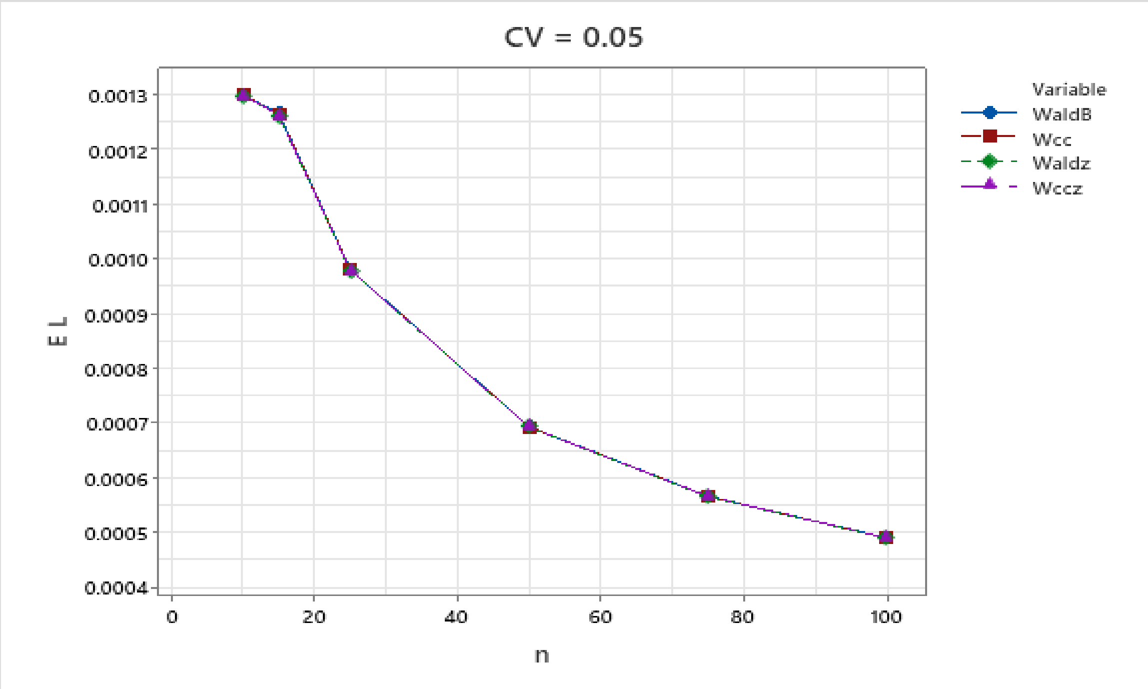


Fig 3.1 The Estimated Lengths (EL); $(\lambda = 400)$.

This behavior of the estimated lengths (EL) of the confidence intervals, when plotted against the sample sizes, will remain almost the same for $\lambda = 100$, 44.44, 25, 16 with confidence coefficient 0.95 of the four considered methods for sample sizes varying from 10 to 100. These simulated results are summarized in Table 3.2 below.

The estimated lengths are very tide to each other and really hard to distinguish between them, with a slight preferable of the confidence interval that based on *WaldZ*, especially for large sample sizes.

Table 3.2 The Estimated Lengths (*EL*) for *CV* = 0.10, 0.15, 0.20, 0.25 and $1 - \alpha = 0.95$.

$CV = 0.10$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.0051000	0.0051000	0.0051000	0.0051000
15	0.0050851	0.0050723	0.0050583	0.0050712
25	0.0039349	0.0039251	0.0039198	0.0039302
50	0.0027807	0.0027738	0.0027721	0.0027792
75	0.0022700	0.0022643	0.0022635	0.0022690
100	0.0019655	0.0019606	0.0019601	0.0019648
$CV = 0.15$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.0115000	0.0114000	0.0114000	0.0115000
15	0.0115100	0.0114450	0.0114090	0.0114720
25	0.0088960	0.0088461	0.0088318	0.0088820
50	0.0062816	0.0062465	0.0062419	0.0062772
75	0.0051262	0.0050976	0.0050951	0.0051236
100	0.0044383	0.0044135	0.0044118	0.0044364
$CV = 0.20$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.0206000	0.0204000	0.0204000	0.0206000
15	0.0206380	0.0204310	0.0203520	0.0205570
25	0.0159270	0.0157680	0.0157360	0.0158960
50	0.0112280	0.0111170	0.0111060	0.0112170
75	0.0091613	0.0090707	0.0090649	0.0091560
100	0.0079291	0.0078508	0.0078461	0.0079249
$CV = 0.25$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.032600	0.032100	0.031900	0.032400
15	0.032599	0.032087	0.031935	0.032440
25	0.025104	0.024714	0.024647	0.025038
50	0.017672	0.017400	0.017378	0.017650
75	0.014406	0.014184	0.014171	0.014392
100	0.012468	0.012277	0.012269	0.012458

The corresponding Fig 3.2 will also enhancing this interpretations, i.e., as the Poisson parameter λ varying from 16 to 400 the expected lengths are very tidy to each other for all the four methods. Furthermore, these expected lengths are seen to be decreased dramatically with the increase of sample sizes for $\lambda = 400, 100$, 44.44, 25, 16.

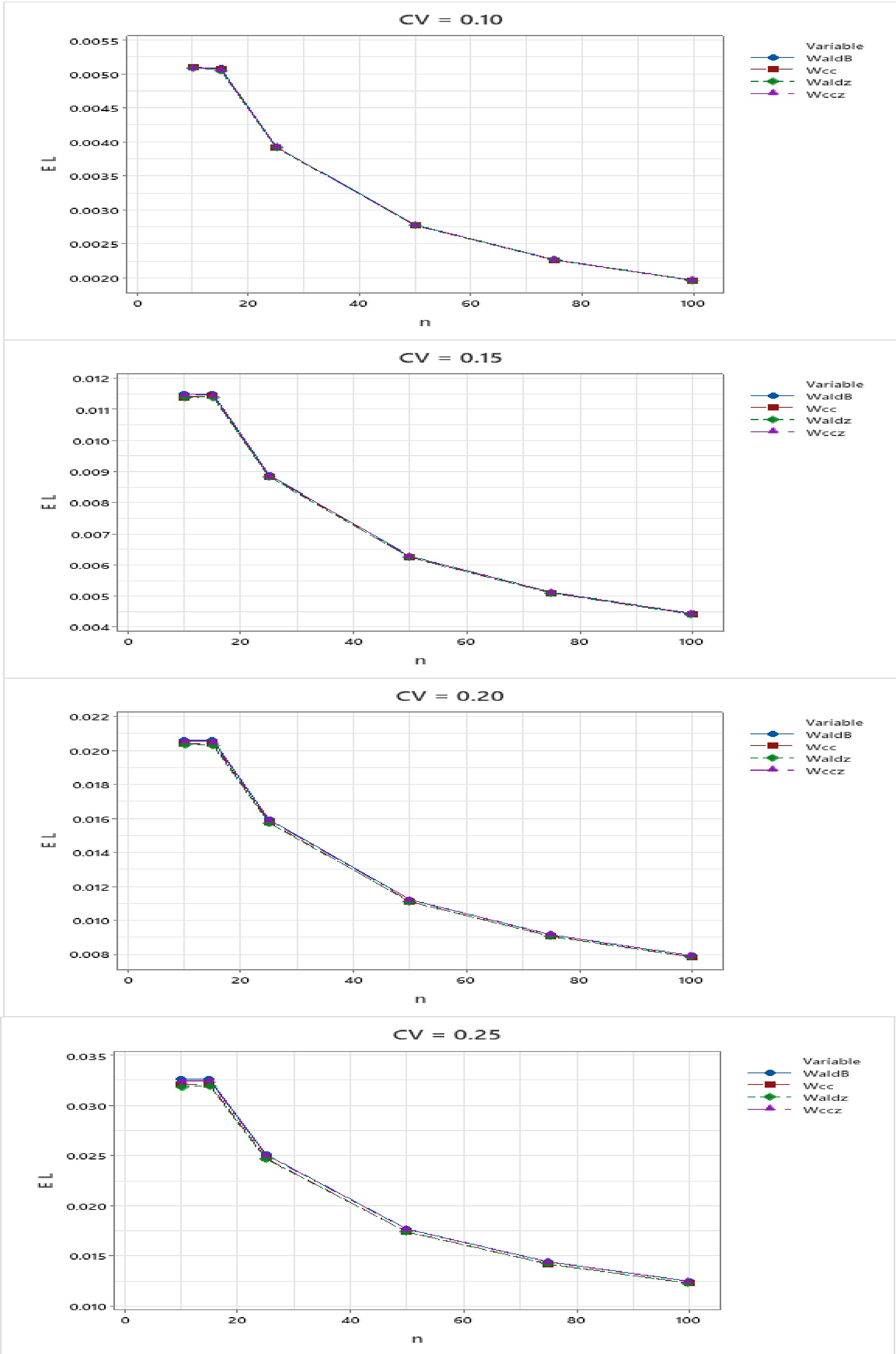


Fig 3.2 The Estimated Lengths (EL); $(\lambda = 100, 44.44, 25, 16)$.

Table 3.3 below gives the simulated results of the estimated lengths (EL) of the confidence intervals, with varying sample sizes from 10 to 100, for $\lambda = 11.11,\ 8.16,\ 6.25,\ 4.93,\ 4$ with confidence coefficient 0.95 of all considered methods. Again the estimated lengths are much closed to each other with a superiority of the confidence intervals based on *WaldZ* and *WaldB* methods, especially for sample sizes as large as 25 or more.

Table 3.3 The Estimated Lengths (*EL*) for *CV* = 0.30, 0.35, 0.40, 0.45, 0.50 and$1-\alpha=0.95$.

$CV=0.30$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.047600	0.046500	0.046200	0.047300
15	0.047572	0.046492	0.046224	0.047293
25	0.036527	0.035712	0.035593	0.036404
50	0.025673	0.025107	0.025063	0.025629
75	0.020913	0.020454	0.020431	0.020888
100	0.018090	0.017694	0.017678	0.018074
$CV=0.35$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.065900	0.063800	0.063400	0.065400
15	0.065850	0.063807	0.063363	0.065379
25	0.050370	0.048844	0.048635	0.050155
50	0.035292	0.034241	0.034169	0.035215
75	0.028725	0.027875	0.027834	0.028684
100	0.024845	0.024111	0.024084	0.024814
$CV=0.40$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.087600	0.084000	0.083300	0.086800
15	0.087583	0.084020	0.083300	0.086815
25	0.066737	0.064104	0.063776	0.066376
50	0.046643	0.044843	0.044722	0.046515
75	0.037931	0.036478	0.036412	0.037861
100	0.032780	0.031529	0.031484	0.032732
$CV=0.45$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.113500	0.107600	0.106500	0.112300
15	0.113480	0.107590	0.106490	0.112270
25	0.085965	0.081680	0.081171	0.085398
50	0.059802	0.056907	0.056721	0.059596
75	0.048592	0.046261	0.046161	0.048479
100	0.041971	0.039968	0.039898	0.041894
$CV=0.50$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.143600	0.134300	0.132700	0.141700
15	0.143580	0.134310	0.132650	0.141710
25	0.108120	0.101490	0.100690	0.107240
50	0.074939	0.070503	0.070224	0.074630
75	0.060777	0.057221	0.057060	0.060608
100	0.052472	0.049420	0.049316	0.052359

The corresponding Fig 3.3 clearly support our interpretations, i.e., as the Poisson parameter varying from 4 to 11.11 the expected lengths are very tidy to each other except for the methods of *WaldZ* and *WaldB* which show a slight superiority over the other two methods, especially for sample sizes as large as 25 or more. Again, these expected lengths are seen to be decreased dramatically with the increase of the sample size for all the given Poisson parameter values.

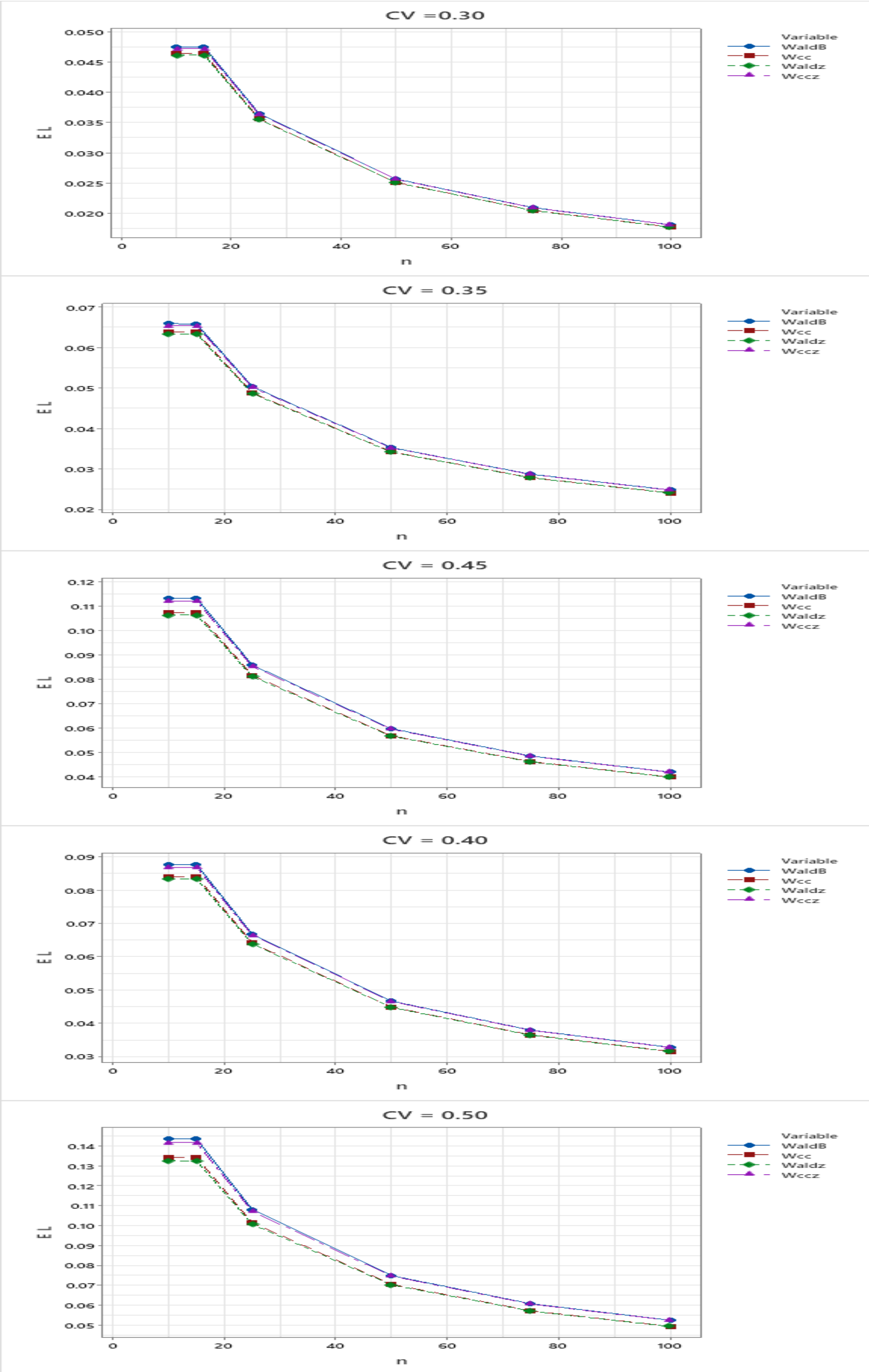


Fig 3.3 Estimated Lengths (EL) ; $(\lambda = 11.11, 8.16, 6.25, 4.93, 4)$.

3.2.2 The Estimated Coverage Probability (*ECP*) For Different Values of *n*

We now turn our attention to the estimated coverage probability (*ECP*) for different values of the Poisson parameter and varying sample sizes. This simulated study has been conducted to know the effect of the parameter value to the estimated coverage probability (*ECP*), of the estimated 95 % two-sided confidence intervals of the Poisson coefficient of variation, for varying sample sizes. Table 3.4 below gives the simulated values of the coverage probability (*ECP*) for the Poisson parameter $\lambda = 400$ (or, *CV*= 0.05) With varying sample sizes from 10 to 100 of the four considered methods.

Table 3.4 The Estimated Coverage Prob. (*ECP*) for *CV*= 0.05 and Confidence Coefficient 0.95.

$CV = 0.05$				
n	$WaldB$	WCC	$WaldZ$	$WccZ$
10	0.9506	0.9499	0.9494	0.9491
15	0.9506	0.94987	0.94944	0.94906
25	0.94957	0.94906	0.94808	0.94839
50	0.95044	0.95044	0.9498	0.94994
75	0.94992	0.94992	0.94939	0.94977
100	0.94923	0.94901	0.94864	0.94899

The corresponding Fig 3.4 below shows that the method of *WaldB* provide us with the best estimated coverage probability followed by the estimated method of *WCC*. These two methods show their superiority over the other two methods for all sample sizes.

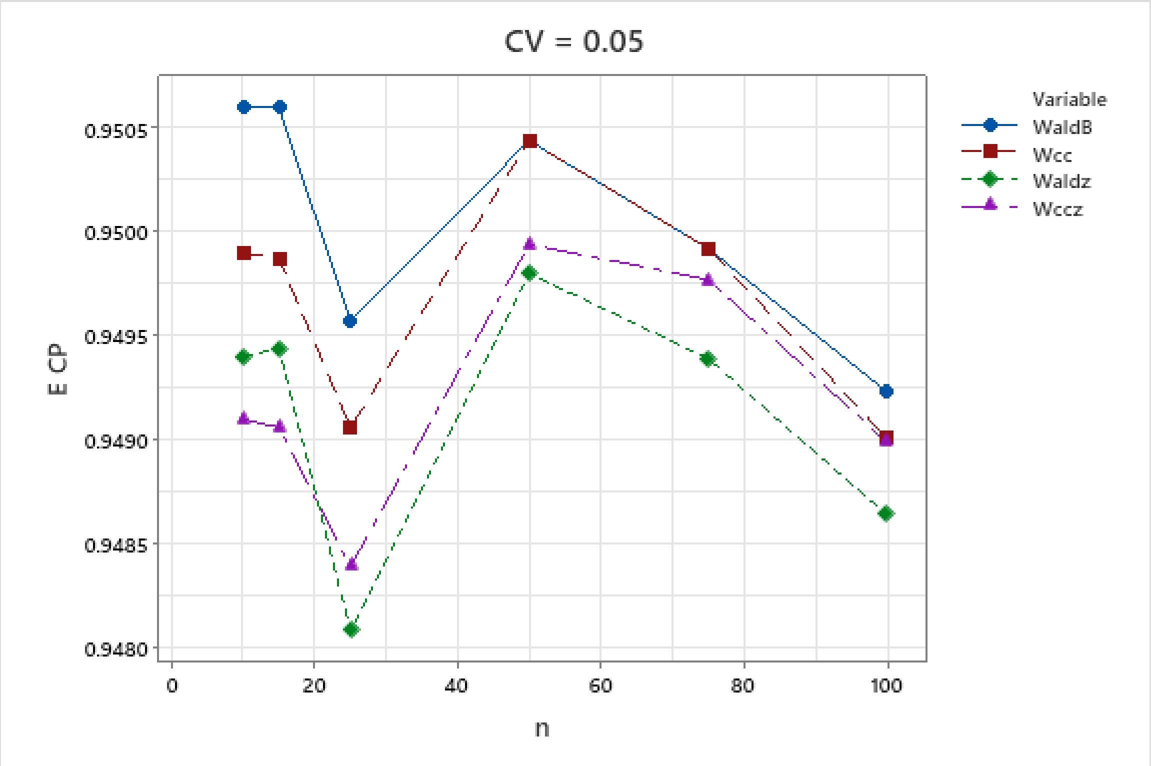


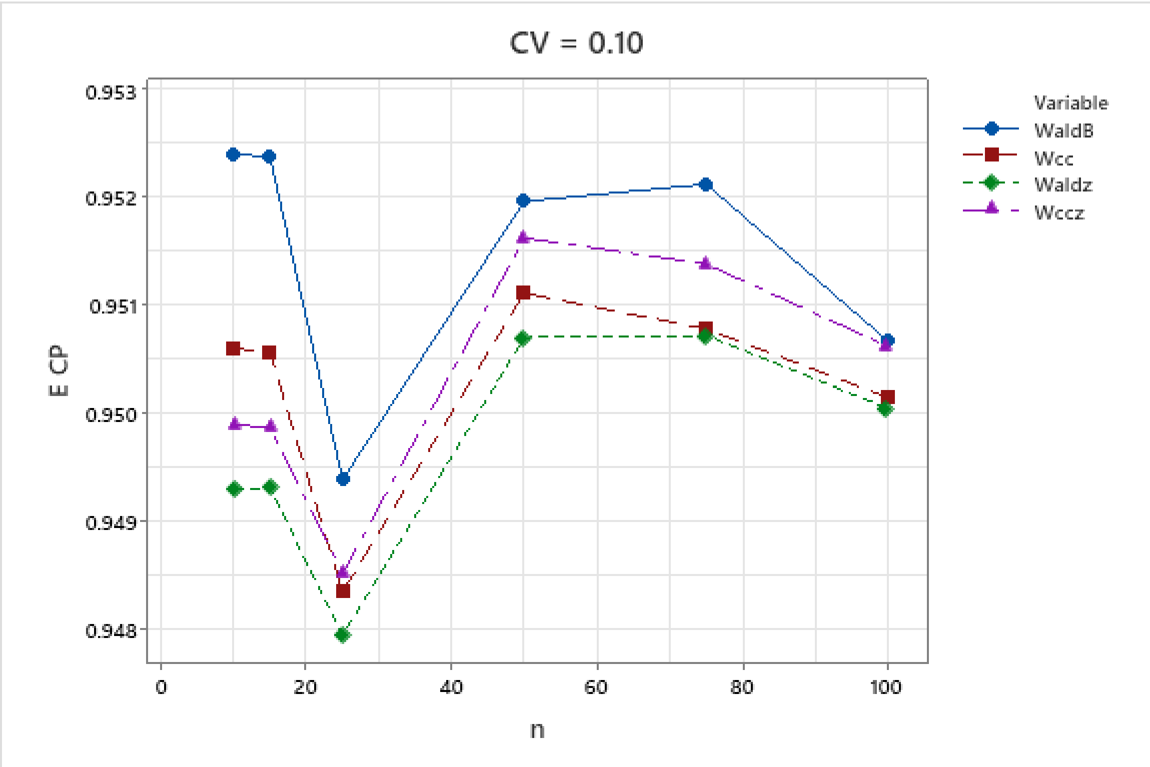
Fig 3.4 The Estimated Coverage Prob. (ECP); $(\lambda = 400)$.

Table 3.5 below gives the simulated values of the coverage probability (*ECP*) for the Poisson parameters $\lambda = 100$, 44.44 (or, *CV*=0.10, 0.15) With varying sample sizes from 10 to 100 of the four considered methods. The obtained estimated values for the coverage probability are very similar and really difficult to distinguish between them, the corresponding figure is then more important to visually judge the performance of these estimation methods.

Table 3.5 The Estimated Coverage Prob. (*ECP*) for *CV*= 0.10, 0.15 and Confidence Coeff. 0.95.

$CV = 0.10$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.9524	0.9506	0.9493	0.9499
15	0.95237	0.95056	0.94931	0.94987
25	0.94939	0.94835	0.94795	0.94852
50	0.95196	0.95111	0.9507	0.95162
75	0.95212	0.95078	0.95072	0.95138
100	0.95067	0.95015	0.95005	0.95062
$CV = 0.15$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.9497	0.9496	0.9469	0.9481
15	0.94966	0.94965	0.94691	0.94814
25	0.95034	0.95034	0.94907	0.95041
50	0.95206	0.95085	0.95032	0.95144
75	0.95164	0.94951	0.95008	0.95073
100	0.9514	0.95051	0.94977	0.95119

The corresponding Fig 3.5 below shows again that the method of *WaldB* provide us with the best estimated coverage probability followed by the estimated method of *WccZ* except for sample sizes 10 and 15 where *WCC* is doing better than the method *WccZ*. The *WaldB* method is again show its superiority over the other three methods for all sample sizes. The *WaldZ* method is the worst method for estimating the coverage probability of the Poisson parameters $\lambda = 400, 100$, 44.44 (or, *CV* = 0.05, 0.10, 0.15) With varying sample sizes from 10 to 100 of the four considered methods.



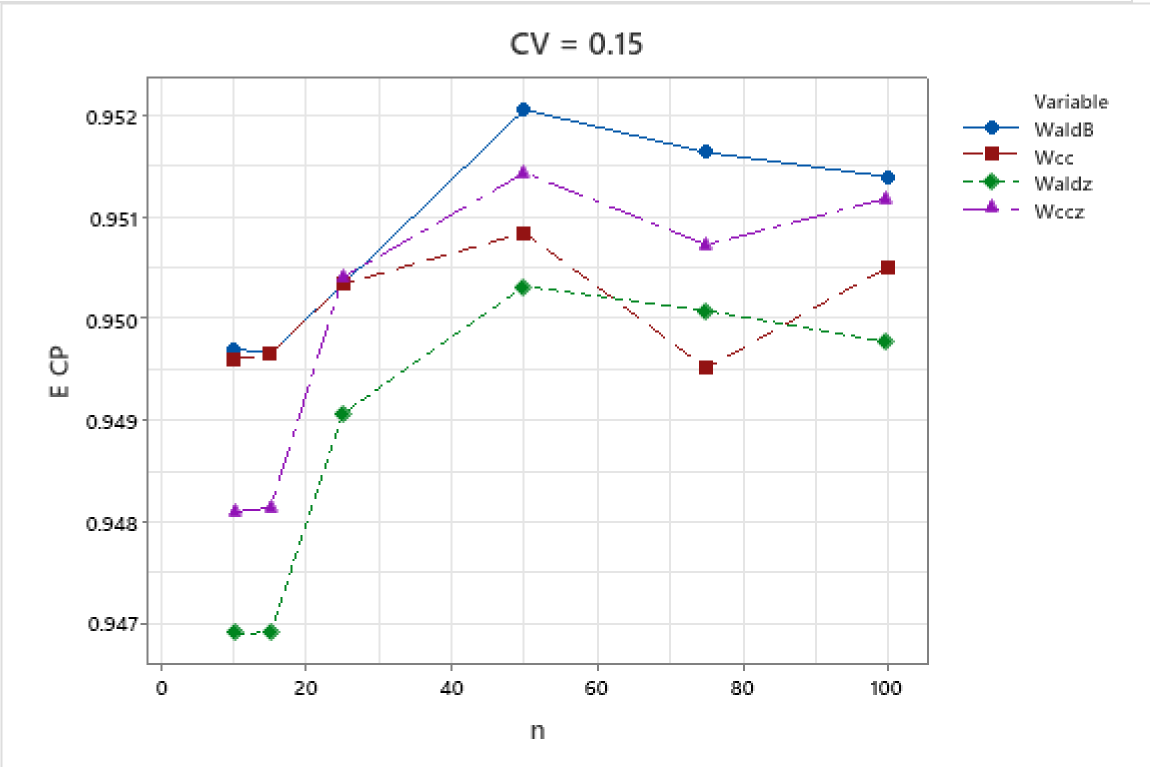


Fig 3.5 The Estimated Coverage Prob. (ECP) ; $(\lambda = 100, 44.4)$.

Table 3.6 below presents the simulated values of the coverage probability (*ECP*) for the Poisson parameters $\lambda = 25, 16, 11.11, 8.16$ (or, *CV*=0.20, 0.25, 0.30, 0.35) With varying sample sizes from 10 to 100 of the four considered methods. At all considered cases, the estimated values obtained by the methods of *WaldB* and *WccZ* are close to each other and seem to be, in general, better than the estimated values obtained by the other two methods of *WCC* and *WaldZ*.

Table 3.6 The Estimated Coverage Prob. (*ECP*) for *CV*= 0.2, 0.25, 0.3, 0.35 and Confidence Coefficient 0.95.

$CV = 0.20$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.95200	0.94940	0.94800	0.95030
15	0.95200	0.94936	0.94801	0.95032
25	0.95101	0.94893	0.94836	0.95033
50	0.95185	0.95011	0.94913	0.95162
75	0.95046	0.94818	0.9487	0.9511
100	0.95127	0.94907	0.94875	0.95122
$CV = 0.25$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.95020	0.95020	0.94910	0.95250
15	0.95024	0.95024	0.94914	0.95250
25	0.95127	0.95127	0.94933	0.95299
50	0.95317	0.94902	0.94879	0.95219
75	0.95337	0.94937	0.94959	0.95302
100	0.95289	0.95012	0.94948	0.95289
$CV = 0.30$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.95480	0.94640	0.94860	0.95300
15	0.95476	0.94639	0.94855	0.95305
25	0.95443	0.95178	0.94938	0.95412
50	0.95323	0.94805	0.94934	0.95410
75	0.95526	0.94991	0.94988	0.95445
100	0.95273	0.9495	0.94874	0.95362
$CV = 0.35$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.95690	0.95030	0.94780	0.95430
15	0.95686	0.95026	0.94785	0.95426
25	0.95274	0.95270	0.94913	0.95543
50	0.95496	0.94972	0.95026	0.95628
75	0.95672	0.95012	0.94980	0.95683
100	0.95505	0.94867	0.94862	0.95548

For more precise comparisons, we again relay on the figures of the estimated coverage probability when plotted against the sample size for different values of the Poisson parameter λ. Fig 3.6 below shows again that the performance of the methods of *WaldB* and *WccZ* are the best in estimating the coverage probability compared to the other two methods of *WCC* and *WaldZ* which showed again that they are under estimating the coverage probability and that for all Poisson parameters $\lambda = 25, 16, 11.11, 8.16$ and all sample sizes *n* = 10, 15, 25, 50, 75, 100.

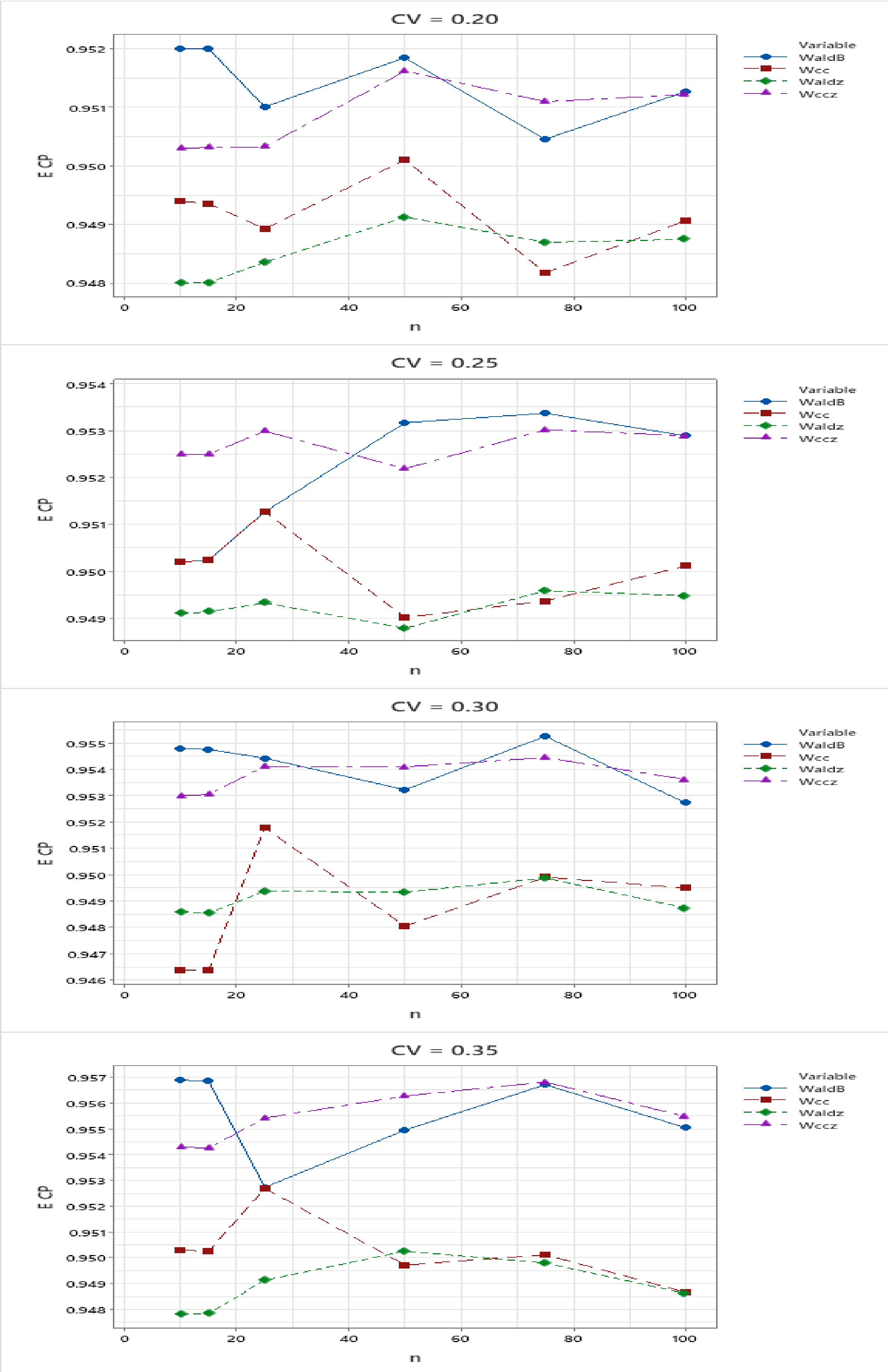


Fig 3.6 The Estimated Coverage Prob. (ECP) ; $(\lambda = 25, 16, 11.11, 8.16)$.

The simulated values of the coverage probability (*ECP*) for the Poisson parameters $\lambda = 6.25, 4.93, 4$ (or, *CV*=0.40, 0.45, 0.50) With varying sample sizes from 10 to 100 of the four considered methods are presented respectively in Table 3.7 below. At all considered cases, the estimated values obtained by the methods of *WaldB* and *WccZ* are close to each other and seem to be, in general, better than the estimated values obtained by the methods of *WCC* and *WaldZ.*

Table 3.7 The Estimated Coverage Prob. (ECP) for CV= 0.40, 0.45, 0.50 and Confidence Coefficient 0.95.

$CV = 0.40$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.96130	0.94960	0.94930	0.95830
15	0.96127	0.94961	0.94926	0.95827
25	0.95780	0.94891	0.94848	0.95729
50	0.95743	0.95179	0.94939	0.95812
75	0.95896	0.95081	0.95091	0.95899
100	0.95881	0.95050	0.95024	0.95825
$CV = 0.45$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.95320	0.94110	0.94500	0.95640
15	0.95316	0.94112	0.94504	0.95642
25	0.95867	0.94879	0.94778	0.95831
50	0.96039	0.95014	0.94905	0.95981
75	0.95872	0.94981	0.94895	0.95934
100	0.95752	0.94776	0.94813	0.95849
$CV = 0.50$				
n	*WaldB*	*WCC*	*WaldZ*	*WccZ*
10	0.96200	0.94780	0.94490	0.95830
15	0.96202	0.94776	0.94495	0.95825
25	0.95619	0.94427	0.9469	0.96015
50	0.96104	0.94634	0.94872	0.96113
75	0.96163	0.94921	0.9486	0.96134
100	0.96047	0.95051	0.94881	0.96114

For more precise comparisons, we again relay on the figures of the estimated coverage probability when plotted against the sample size for different values of the Poisson parameter λ. Fig 3.7 below shows again that the performance of the methods of *WaldB* and *WccZ* are the best in estimating the coverage probability compared to the other two methods of *WCC* and *WaldZ* which showed again that they are under estimating the coverage probability and that for all Poisson parameters $\lambda = 6.25, 4.93, 4$ and all sample sizes *n* = 10, 15, 25, 50, 75, 100.

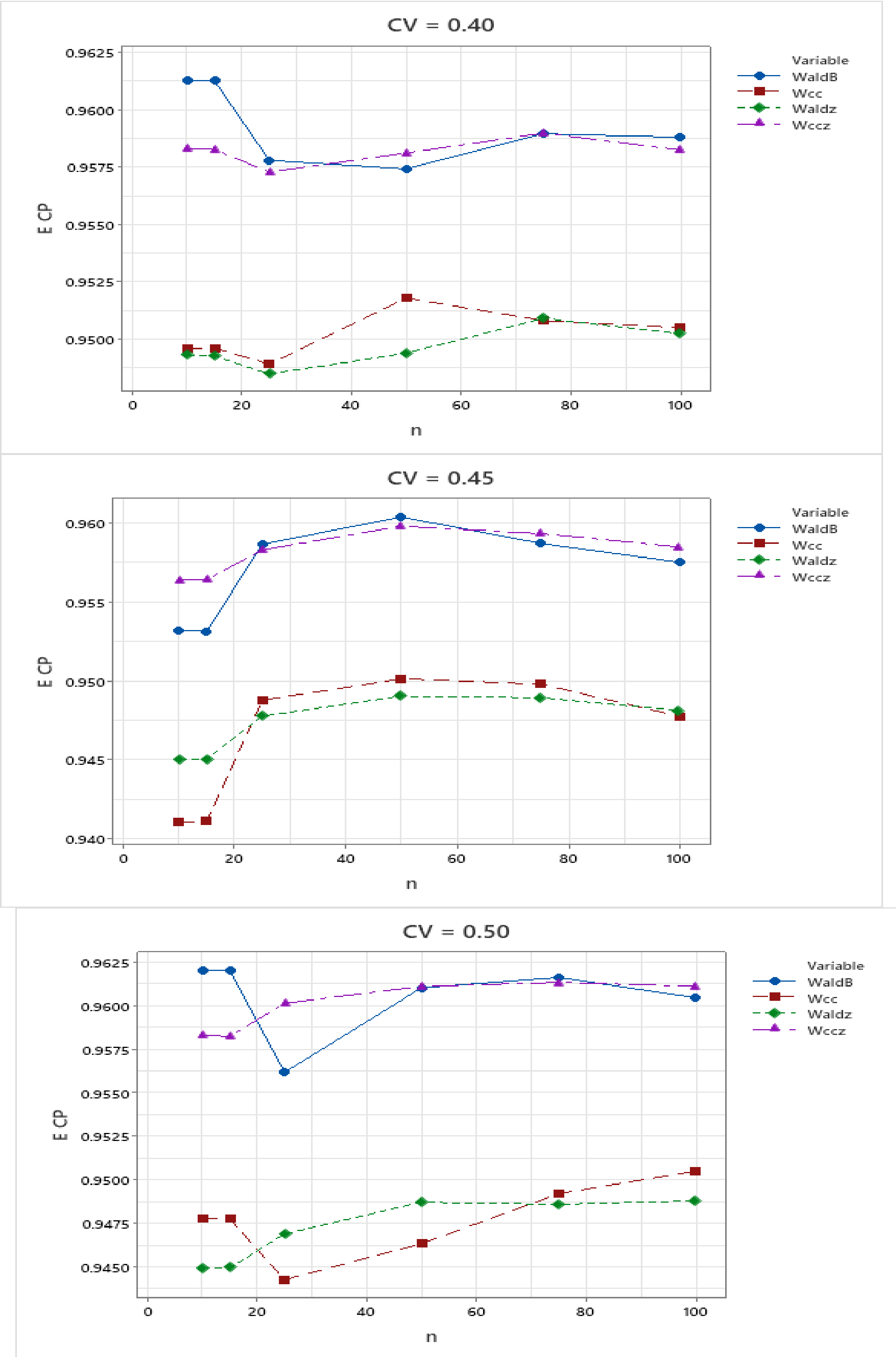


Fig 3.7 The Estimated Coverage Prob. (ECP) ; $(\lambda = 6.25, 4.93, 4)$.

3.2.3 The Estimated Coverage Probability for Different Values of λ.

For more understanding to the behavior of the estimated coverage probability (*ECP*), we may need to look at it from different angle. This can be done by Sketching the estimated coverage probability, for a fixed sample size, against the Poisson parameter values λ = 400, 100, 44.44, 25, 16, 11.11, 8.16, 6.25, 4.93, 4 instead of sketching it against the varying sample size values n = 10, 15, 25, 50, 75, 100 (see subsection 3.2.2 above) seem to provide us with more insight to its behavior. For this we now turn our attention to the estimated coverage probability (*ECP*) for different values of the Poisson parameter with a specific value of the sample size each time. This simulated study has been conducted to see the effect of the parameter values to the estimated coverage probability (*ECP*) for a fixed sample size.

Table 3.8 below presents the simulated values of the coverage probability (*ECP*) for different Poisson parameter values and that for each fixed sample size (n = 10, 15, 25, 50, 75, 100) of the four considered methods. Although it is difficult to study the effect of the Poisson parameter values to the estimating coverage probability for each sample size by just looking to the figures of Table 3.8 below. This is due to the closeness of these estimating coverage probability values for all the considered methods. But for large Poisson parameter values (λ = 400, 100, 44.44, 25) one can recognize that the four methods are doing similar job when estimating the coverage probability for each sample size. As the Poisson parameter values get smaller (λ = 16, 11.11, 8.16, 6.25, 4.93, 4), both *WaldB* and *WccZ* methods seem to do much better, in estimating the coverage probability values, compared with the other two methods of *WaldZ* and *WCC* which are under estimating the coverage probability for all sample sizes.

For more precise comparisons, we cannot only rely on the figures of Table 3.8, but it is more important to have a visual inspection to the sketched figures of the estimated coverage probability against the Poisson parameter values λ, for a given fixed sample size.
The corresponding Fig 3.8 (a) and (b) below shows clearly, and as we said earlier, that the performance of the four methods, in estimating the coverage probability values, for large Poisson parameter values (λ = 400, 100, 44.44, 25) is very similar with a little favored job of the *WaldB* and *WccZ* methods. As the Poisson parameter values get smaller (λ = 16, 11.11, 8.16, 6.25, 4.93, 4), both *WaldB* and *WccZ* methods are very clearly doing a much better estimating job, compared with the other two methods of *WaldZ* and *WCC* which are clearly under estimating the coverage probability for all given sample sizes.

Table 3.8 The Estimated Coverage Prob. *(ECP)* against the Coefficient of Variation (*CV*).

Estimated Coverage Probability (*ECP*) for Confidence Coefficient 0.95.										
Method	(a) Coefficient of Variation (CV) for *n*=10									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.95060	0.95240	0.94970	0.9520	0.95020	0.95480	0.95690	0.96130	0.95320	0.9620
WCC	0.94990	0.95060	0.94960	0.94940	0.95020	0.94640	0.95030	0.94960	0.94110	0.94780
WaldZ	0.94940	0.94930	0.94690	0.9480	0.94910	0.94860	0.94780	0.94930	0.94500	0.94490
WccZ	0.94910	0.94990	0.94810	0.95030	0.95250	0.95300	0.95430	0.95830	0.95640	0.95830
Method	(b) Coefficient of Variation CV) for *n*=15									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.95060	0.95237	0.94966	0.9520	0.95024	0.95476	0.95686	0.96127	0.95316	0.96202
WCC	0.94987	0.95056	0.94965	0.94936	0.95024	0.94639	0.95026	0.94961	0.94112	0.94776
WaldZ	0.94944	0.94931	0.94691	0.94801	0.94914	0.94855	0.94785	0.94926	0.94504	0.94495
WccZ	0.94906	0.94987	0.94814	0.95032	0.9525	0.95305	0.95426	0.95827	0.95642	0.95825
Method	(c) Coefficient of Variation (CV) for *n*=25									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.94957	0.94939	0.95034	0.95101	0.95127	0.95443	0.95274	0.9578	0.95867	0.95619
WCC	0.94906	0.94835	0.95034	0.94893	0.95127	0.95178	0.95270	0.94891	0.94879	0.94427
WaldZ	0.94808	0.94795	0.94907	0.94836	0.94933	0.94938	0.94913	0.94848	0.94778	0.94690
WccZ	0.94839	0.94852	0.95041	0.95033	0.95299	0.95412	0.95543	0.95729	0.95831	0.96015
Method	(d) Coefficient of Variation (CV) for *n*=50									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.95044	0.95196	0.95206	0.95185	0.95317	0.95323	0.95496	0.95743	0.96039	0.96104
WCC	0.95044	0.95111	0.95085	0.95011	0.94902	0.94805	0.94972	0.95179	0.95014	0.94634
WaldZ	0.9498	0.95070	0.95032	0.94913	0.94879	0.94934	0.95026	0.94939	0.94905	0.94872
WccZ	0.94994	0.95162	0.95144	0.95162	0.95219	0.95410	0.95628	0.95812	0.95981	0.96113
Method	(e) Coefficient of Variation (CV) for *n*=75									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.94992	0.95212	0.95164	0.95046	0.95337	0.95526	0.95672	0.95896	0.95872	0.96163
WCC	0.94992	0.95078	0.94951	0.94818	0.94937	0.94991	0.95012	0.95081	0.94981	0.94921
WaldZ	0.94939	0.95072	0.95008	0.94870	0.94959	0.94988	0.94980	0.95091	0.94895	0.94860
WccZ	0.94977	0.95138	0.95073	0.95110	0.95302	0.95445	0.95683	0.95899	0.95934	0.96134
Method	(f) Coefficient of Variation (CV) for *n*=100									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.94923	0.95067	0.9514	0.95127	0.95289	0.95273	0.95505	0.95881	0.95752	0.96047
WCC	0.94901	0.95015	0.95051	0.94907	0.95012	0.94950	0.94867	0.95050	0.94776	0.95051
WaldZ	0.94864	0.95005	0.94977	0.94875	0.94948	0.94874	0.94862	0.95024	0.94813	0.94881
WccZ	0.94899	0.95062	0.95119	0.95122	0.95289	0.95362	0.95548	0.95825	0.95849	0.96114

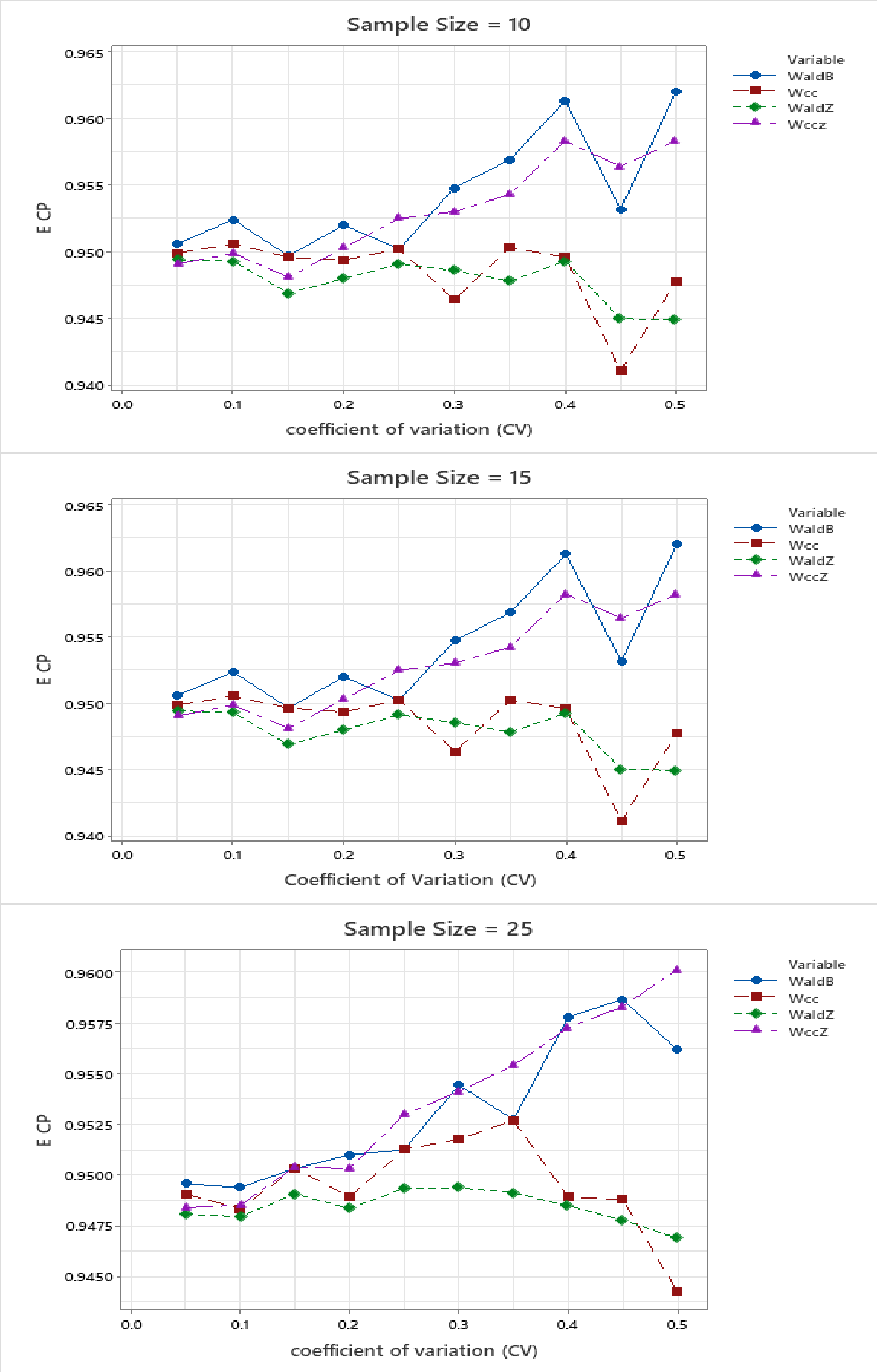


Fig 3.8 (a) The Estimated Coverage Prob. Vs. The Coefficient of Variation.

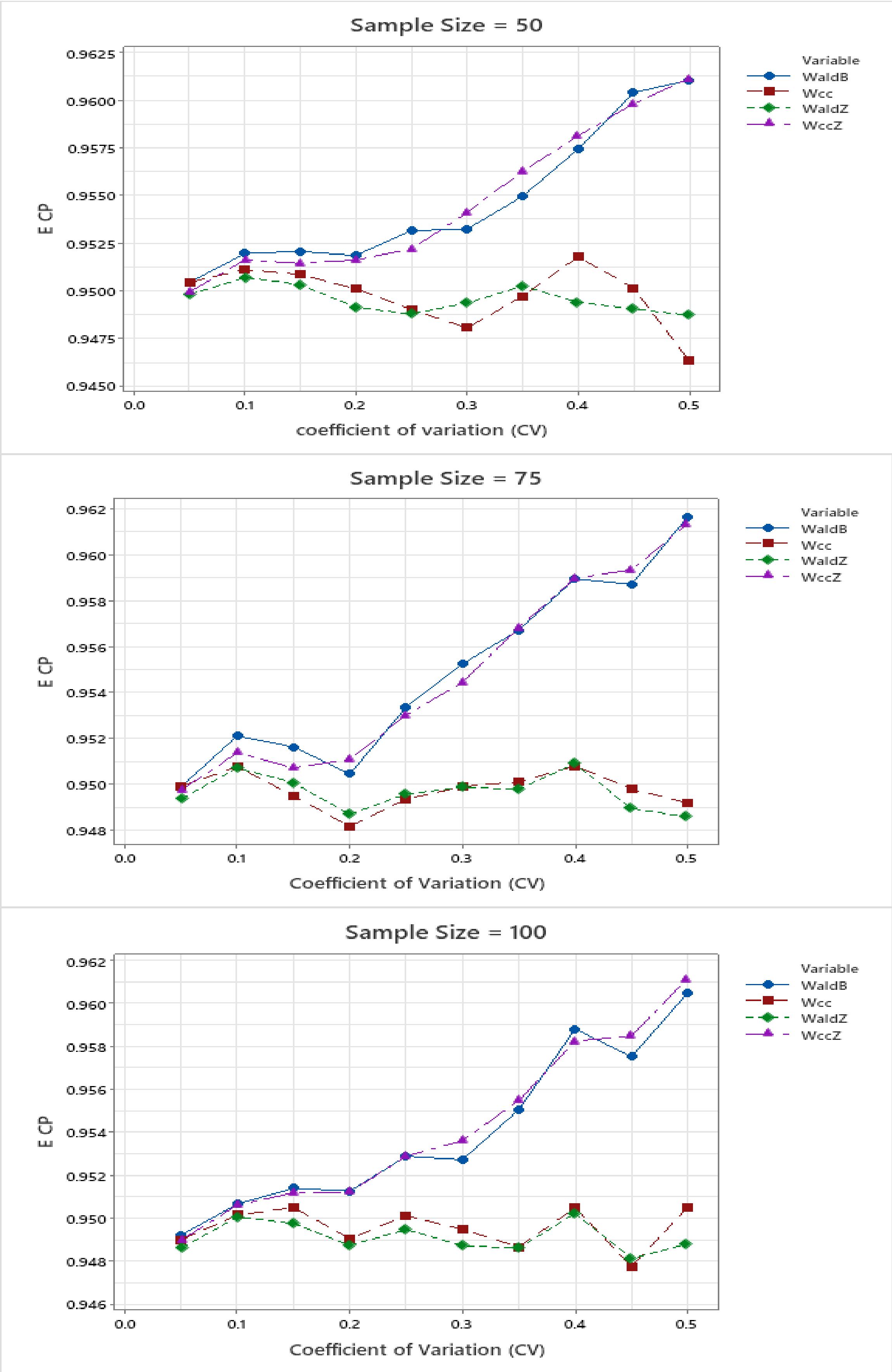


Fig 3.8 (b) The Estimated Coverage Prob. Vs.The Coefficient of Variation.

3.2.4 The Estimated Lengths (EL) For Different Values of λ.

To be able to understand more the behavior of the estimated lengths (*EL*) of the two sided confidence intervals of the coefficient of variation of the Poisson distribution. This can be done by Sketching the estimated lengths (*EL*), for a fixed given sample size, against the Poisson parameter values λ = 400, 100, 44.44, 25, 16, 11.11, 8.16, 6.25, 4.93, 4 (or, equivalently against the coefficient of variation values *CV* = 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50) instead of sketching it against the varying sample size values *n* = 10, 15, 25, 50, 75, 100 (see subsection 3.2.1 above) seem to provide us with more insight to its behavior. For this we now turn our attention to the estimated lengths (*EL*) for different values of the Poisson parameter with a specific value of the sample size each time. This simulated study has been conducted to see the effect of the parameter values (or, equivalently the effect of the coefficient of variation values) on the estimated lengths (*EL*) of the obtained two sided confidence intervals of the coefficient of variation of the Poisson distribution, for a fixed sample size.

Table 3.9 below presents the estimated values of the estimated lengths (*EL*) for different Poisson parameter values (or, equivalently, for different coefficient of variation values) and that for each fixed sample size (*n* = 10, 15, 25, 50, 75, 100) of the four considered methods. Although it is difficult to study the effect of the Poisson parameter values on the estimating lengths for each sample size by just staring to the figures of Table 3.9 below.
But for large Poisson parameter values (λ = 400, 100, 44.44, 25, 16, 11.11, 8.16) one can recognize that the four methods are doing very similar job when estimating the lengths for each sample size. As the Poisson parameter values get smaller (λ = 6.25, 4.93, 4), both *WaldZ* and *WCC* methods seem to do much better, in estimating the coverage probability values, compared with the other two methods of *WaldB* and *WccZ* which gives a wider confidence intervals (with larger lengths) for all sample sizes.

To perform a more precise comparisons, a visual inspection to the sketched figures of the estimated lengths against the Poisson parameter values λ (or, equivalently against the coefficient of variation values), for a given fixed sample size. The corresponding Fig 3.9 (a) and (b) below shows clearly, that the performance of the four methods, in estimating the lengths, for large Poisson parameter values (λ = 8 or more) but as the Poisson parameter values get smaller (λ = 6 or less), the methods of *WaldZ* and *WCC* are very clearly doing a much better estimating job, compared with the other two methods of *WaldB* and *WccZ* which are clearly gives a wider confidence intervals (with larger lengths) for all considered sample sizes.

Table 3.9 The Estimated Lengths (EL) against the Coefficient of Variation (CV).

Estimated Length (EL) for Confidence Coefficient 0.95.										
Method	(a) Coefficient of Variation (CV) for *n*=10									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.00130	0.00510	0.01150	0.02060	0.03260	0.04760	0.06590	0.08760	0.11350	0.14360
WCC	0.00130	0.00510	0.01140	0.02040	0.03210	0.04650	0.06380	0.0840	0.10760	0.13430
WaldZ	0.00130	0.00510	0.01140	0.02040	0.03190	0.04620	0.06340	0.08330	0.10650	0.13270
WccZ	0.00130	0.00510	0.01150	0.02060	0.03240	0.04730	0.06540	0.08680	0.11230	0.14170
Method	(b) Coefficient of Variation (CV) for *n*=15									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.00127	0.00509	0.01151	0.02064	0.03260	0.04757	0.06585	0.08758	0.11348	0.14358
WCC	0.00127	0.00507	0.01144	0.02043	0.03209	0.04649	0.06381	0.08402	0.10759	0.13431
WaldZ	0.00126	0.00506	0.01141	0.02035	0.03193	0.04622	0.06336	0.08330	0.10649	0.13265
WccZ	0.00126	0.00507	0.01147	0.02056	0.03244	0.04729	0.06538	0.08682	0.11227	0.14171
Method	(c) Coefficient of Variation (CV) for *n*=25									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.00098	0.00393	0.00890	0.01593	0.02510	0.03653	0.05037	0.06674	0.08597	0.10812
WCC	0.00098	0.00393	0.00885	0.01577	0.02471	0.03571	0.04884	0.06410	0.08168	0.10149
WaldZ	0.00098	0.00392	0.00883	0.01574	0.02465	0.03559	0.04863	0.06378	0.08117	0.10069
WccZ	0.00098	0.00393	0.00888	0.0159	0.02504	0.03640	0.05016	0.06638	0.08540	0.10724
Method	(d) Coefficient of Variation (CV) for *n*=50									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.00069	0.00278	0.00628	0.01123	0.01767	0.02567	0.03529	0.04664	0.05980	0.07494
WCC	0.00069	0.00277	0.00625	0.01112	0.01740	0.02511	0.03424	0.04484	0.05691	0.07050
WaldZ	0.00069	0.00277	0.00624	0.01111	0.01738	0.02506	0.03417	0.04472	0.05672	0.07022
WccZ	0.00069	0.00278	0.00628	0.01122	0.01765	0.02563	0.03521	0.04652	0.05960	0.07463
Method	(e) Coefficient of Variation (CV) for *n*=75									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.00057	0.00227	0.00513	0.00916	0.01441	0.02091	0.02873	0.03793	0.04859	0.06078
WCC	0.00057	0.00226	0.00510	0.00907	0.01418	0.02045	0.02787	0.03648	0.04626	0.05722
WaldZ	0.00057	0.00226	0.00510	0.00906	0.01417	0.02043	0.02783	0.03641	0.04616	0.05706
WccZ	0.00057	0.00227	0.00512	0.00916	0.01439	0.02089	0.02868	0.03786	0.04848	0.06061
Method	(f) Coefficient of Variation (CV) for *n*=100									
	0.05	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**
WaldB	0.00049	0.00197	0.00444	0.00793	0.01247	0.01809	0.02485	0.03278	0.04197	0.05247
WCC	0.00049	0.00196	0.00441	0.00785	0.01228	0.01769	0.02411	0.03153	0.03997	0.04942
WaldZ	0.00049	0.00196	0.00441	0.00785	0.01227	0.01768	0.02408	0.03148	0.03990	0.04932
WccZ	0.00049	0.00196	0.00444	0.00792	0.01246	0.01807	0.02481	0.03273	0.04189	0.05236

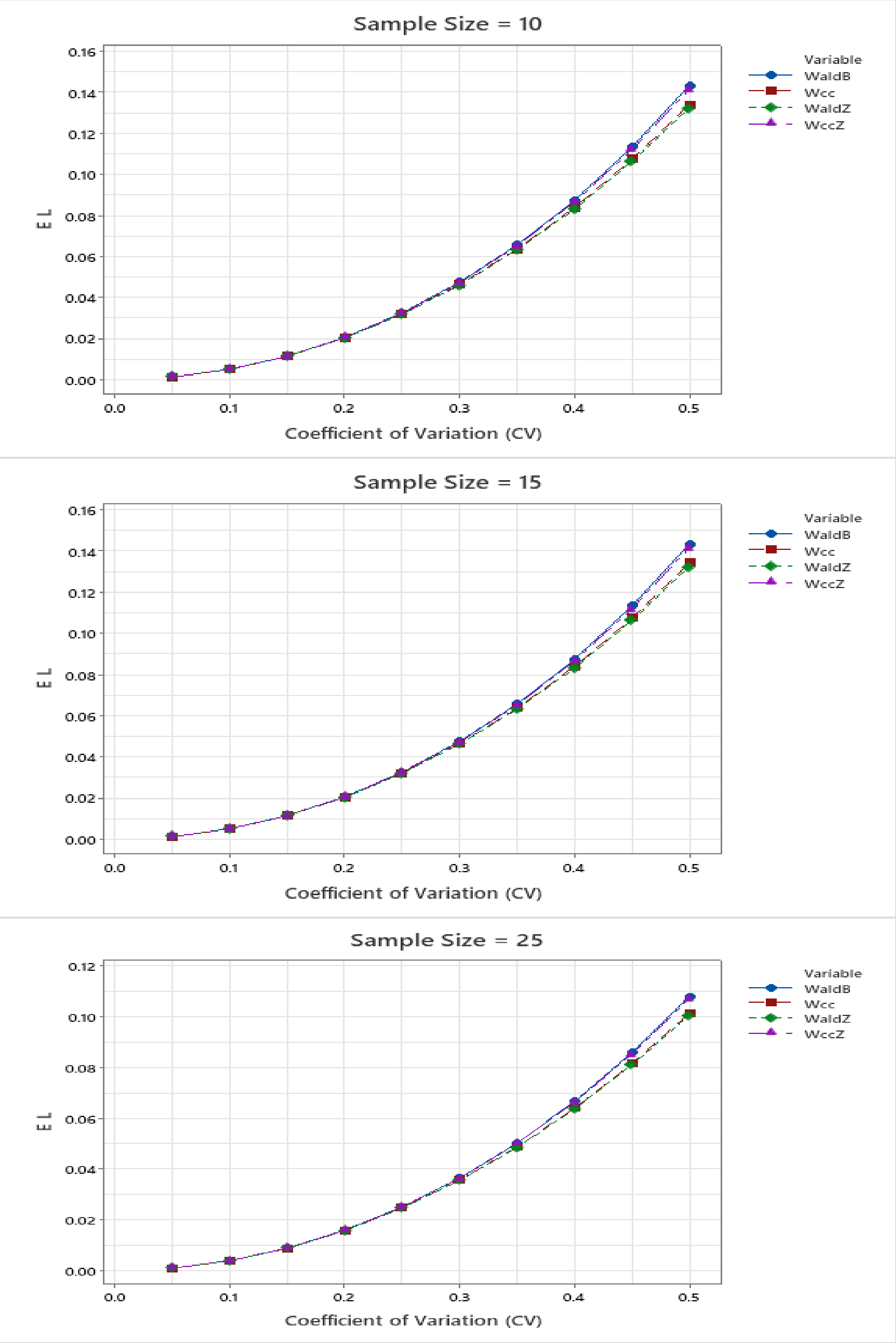


Fig 3.9 (a) The Estimated Lengths (*EL*) against the Coefficient of Variation (*CV*).

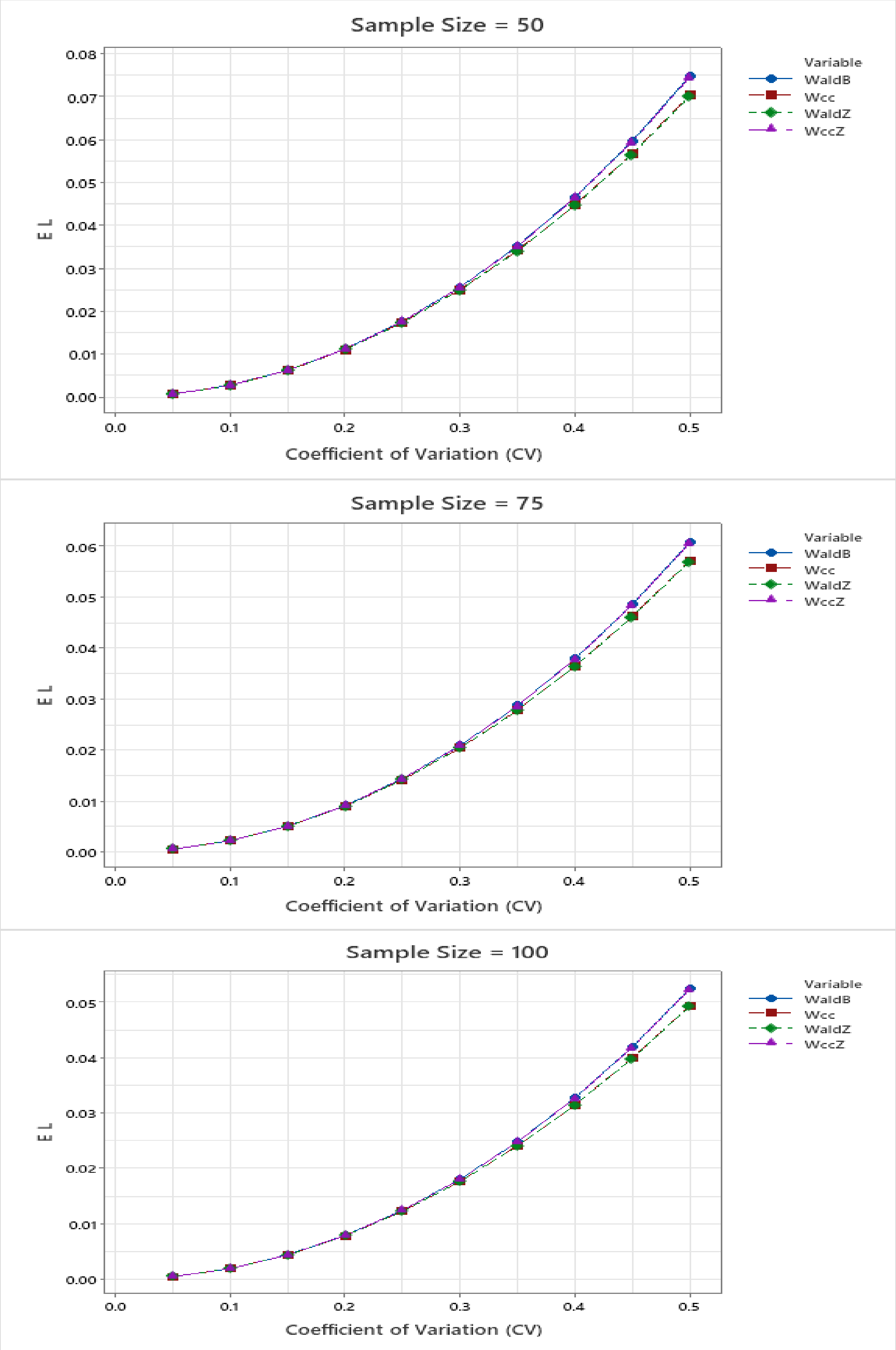


Fig 3.9 (b) The Estimated Lengths (*EL*) against the Coefficient of Variation (*CV*).

Chapter Four

Discussion and Conclusions

Chapter 4

4.0 Discussion and Conclusions

4.1 Introduction

The main aim of this final chapter is to summarize all the above results and to judge the performance of the above mentioned methods that have been used to estimate the confidence interval of the coefficient of variation of a Poisson population of parameter λ. This is done, as given in chapter 3 above, by applying the four methods adopted by this thesis to different simulated data sets with different sample sizes and different Poisson population parameter values. This study considered four different confidence interval methods for estimating the Poisson population coefficient of variation (or equivalently, the SNR). A simulation study has been conducted to compare the performance of four proposed confidence interval methods. The used simulated data have been generated from the Poisson distribution with a varying vales of the parameter λ, i.e., with a varying valus of the Poisson population coefficient of variation using Matlab software. The coverage probability and interval length of each confidence interval method will be calculated and reported.

To investigate the properties of a statistical procedure, a simulation techniques are good choice, especially where a theoretical study is not possible. Bootstrapping technique has been applied to compare between confidence interval methods for the Poisson coefficient of variation. This is done by simulating an original data set then randomly selecting data several times with replacement to estimate the unknown parameters. The simulation study has been conducted to compare the performance of the four considered confidence interval methods under the same simulation conditions. The generated data sets are performed using programs written in the Matlab software, repeated 100,000 times in each case at confidence level 0.95. All calculations for this simulation study are based on Monte-Carlo Simulation and then double precision computations are adopted.

4.2 The Overall Conclusions

We employ the four two-sided confidence interval methods, the Wald with Continuity Correction (*WCC*) method, the Wald Bootstrap method (*WaldB*), the (*WaldZ*) method and the Wald with Continuity Correction Bootstrap method (*WccZ*) to many obtained simulated data sets. Throughout this thesis many important points and useful results are obtained and the main features that could be drawn are summarized below:

(1) Estimated Lengths (EL) Against Sample sizes (n) of confidence Intervals of data generated from Poisson distribution with parameter values (λ) varying from 4 to 400 and sample sizes (*n)* varying from 10 to 100. The estimated 95 % two-sided confidence intervals of each coefficient of variation are then calculated along with their corresponding lengths.

As can be seen from the obtained tables and figures that the estimated lengths of all confidence intervals of the four considered methods are very close to each other for each sample size and it is really difficult to distinguish between them, with a very slight preferable of the confidence interval that based on *WaldZ*, especially for large sample sizes (25 or more) and as the Poisson parameter λ varying from 16 to 400 the expected lengths are very tidy to each other for all the four methods. Furthermore, these expected lengths are seen to be decreased dramatically with the increase of sample sizes for $\lambda = 400, 100$, 44.44, 25, 16. But for $\lambda = 11.11,\ 8.16,\ 6.25,\ 4.93,\ 4$ the estimated lengths are again much closed to each other with a superiority to the confidence intervals based on *WaldZ* and *WaldB* methods, especially for sample sizes as large as 25 or more.

(2) Estimated Coverage Probability (ECP) Against Sample sizes (n) for different values of the Poisson parameter and varying sample sizes. This is to know if there is any effect of the parameter value to the estimated coverage probability (*ECP*), of the estimated 95 % two-sided confidence intervals of the Poisson coefficient of variation of varying sample sizes.

For the Poisson parameter $\lambda = 400$ (or, *CV*= 0.05) with varying sample sizes from 10 to 100 of the four considered methods, it has be shown that the method of *WaldB* provide us with the best estimated coverage probability followed by the estimated method of *WCC*. These two methods show their superiority over the other two methods for all sample sizes. For $\lambda = 100$, 44.44 (or, *CV*=0.10, 0.15) With varying sample sizes, the obtained estimated values for the coverage probability are very similar and really difficult to distinguish between them and again that the method of *WaldB* provide us with the best estimated coverage probability followed by the estimated method of *WccZ* except for sample sizes 10 and 15 where *WCC* is doing better than the method *WccZ*. The *WaldB* method is again show its superiority over the other three methods for all sample sizes. The *WaldZ* method is the worst method for estimating the coverage probability of the Poisson parameters $\lambda = 400, 100\ , 44.44$ (or, *CV*= 0.05, 0.10, 0.15). For $\lambda = 25, 16, 11.11, 8.16, 6.25, 4.93, 4$ (or, *CV*=0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50) for all sample sizes we can see that the methods of *WaldB* and *WccZ* are close to each other and seem to be, in general, better than the estimated values obtained by the other two methods of *WCC* and *WaldZ*.

(3) The Estimated Coverage Probability (ECP) for Different Values of $\boldsymbol{\lambda}$ with a specific value of the sample size (n) at each time. This simulated study has been conducted to see the effect of the parameter values to the estimated coverage probability (*ECP*) for a fixed sample size. For large Poisson parameter values (λ = 400, 100, 44.44, 25) one can recognize that the four methods are doing similar job when estimating the coverage probability for each sample size, but as the Poisson parameter values get smaller (λ = 16, 11.11, 8.16, 6.25, 4.93, 4), both *WaldB* and *WccZ* methods seem to do much better, in estimating the coverage probability values, compared with the other two methods of *WaldZ* and *WCC* which are under estimating the coverage probability for all sample sizes.

(4) The Estimated Lengths (EL) against Different Values of λ for a specific fixed sample size (n) each time. This simulated study has been conducted to see the effect of the parameter values (or, equivalently the effect of the coefficient of variation values) on the estimated lengths (EL) of the obtained two sided confidence intervals of the coefficient of variation of the Poisson distribution, for a fixed sample size. It has been shown that for large Poisson parameter values (λ = 400, 100, 44.44, 25, 16, 11.11, 8.16) one can recognize that the four methods are doing very similar job when estimating the lengths for each sample size. As the Poisson parameter values get smaller (λ = 6.25, 4.93, 4), both WaldZ and WCC methods seem to do much better, in estimating the coverage probability values, compared with the other two methods of WaldB and WccZ which gives a wider confidence intervals (with larger lengths) for all sample sizes.

Directions for Future Work

There are a number of possible extensions to the work presented in this thesis. Here we shall give some ideas of possible future work.

1. Throughout this thesis our attention has been restricted only to four cases of finding the estimated confidence interval. This could be extended to consider more estimated cases.
2. All calculations for this simulation study are based on Monte-Carlo Simulation so one can try another simulation method.

References

(1) Albatineh, A. N., Boubakari, I, and Kibria, B. M. G. (2017), New Confidence Interval Estimator of the Signal-to-Noise Ratio Based on Asymptotic Sampling Distribution. *Communication in Statistics-Theory and Methods.* **46**(2):574-590.

(2) Albatineh, A. N., George, F., Kibria B. M. G. and Wilcox, M. L. (2014). Confidence interval estimation of the signal-to-noise ratio using ranked set sampling: a simulation study. *Thailand Statistician*, **12**(1):55-69.

(3) Albatineh, A. N., Kibria, B. M. and Zogheib, B. (2014), Asymptotic sampling distribution of inverse coefficient of variation and its applications: revisited, *International Journal of Advanced Statistics and Probability*, **2**:15-20.

(4) Banik, S. and B.M.G. Kibria (2010). Comparison of some parametric and nonparametric type one sample confidence intervals for estimating the mean of a positively skewed distribution. *Communications in Statistics-Simulation and Computation*, **39**:361-389.

(5) Banik, S. and Kibria, B. M. G. (2011), Estimating the population coefficient of variation by confidence intervals. *Communications in Statistics-Simulation and Computation*, **40**:1236–1261.

(6) Bhoj, D. S. and Ahsanullah, M. (1993), Testing equality of coefficients of variation of two populations. *Biometrical Journal*, **35**(3):355–359.

(7) Fieller, E.C. (1932), A numerical test of the adequacy of A.T. McKay's approximation. *Journal of the Royal Statistical Society*, **95**(4):699-702.

(8) George, F. and Kibria, B. M. G. (2011). Confidence Intervals for Signal to Noise Ratio of a Poisson distribution. *American Journal of Biostatistics*,**2**(2) 44-55.

(9) George, F. and Kibria, B. M. G. (2012). Confidence intervals for estimating the population signal-to-noise ratio: a simulation study. *Journal of Applied Stats*, **39**:1225-1240.

(10) Gregoire, T. (1984), Confidence intervals for the coefficient of variation. School of Forestry and Environmental Studies, Yale University.

(11) Hendricks, W. A. and Robey, K W. (1936), The Sampling Distribution of the Coefficient of Variation. *Annals of Mathematical Statistics*, **7**(3):129-132.

(12) Koopmans, L.H., Owen, D.B., and Rosenblatt, J.I. (1964), Confidence intervals for the coefficient of variation for the normal and log normal distributions. *Biometrika*, **51**:25-32.

(13) Mahmoud, R. and Hassani, H. (2009), Two new confidence intervals for the coefficient of variation in a normal distribution, *Journal of Applied Statistics*, **36**: 429-442.

(14) McGibney, G. and Smith, M. R. (1993). An unbiased signal-to-noise ratio measure for magnetic resonance images, *Medical Physics*, **20**:1077-1078.

(15) McKay, A. T. (1932), Distribution of the coefficient of variation and the extended t-distribution. *Journal of Royal Statistical Society*, **B 95**:695-698.

(16) Miller, E.G. and Karson, M.J. (1977), Testing the equality of two coefficients of variation. *American Statistical Association: Proceedings of the Business and Economics Section*, **Part I**: 278-283.

(17) Panichchkikosolkul, W. (2009), Improved confidence intervals for a coefficient of variation of a normal distribution. *Thailand Statistician*, **7**(2):193–199.

(18) Panichchkikosolkul, W. (2010), A Simulation Comparison of New Confidence Intervals for the Coefficient of Variation of a Poisson Distribution. *Silpakorn University Science and Technology Journal*, **4**(2):14-20.

(19) Panichchkikosolkul, W. (2013), Confidence Intervals for the Coefficient of Variation in a Normal Distribution with a Known Population Mean. *Journal of Probability and Statistics*, **2013**,1-12.

(20) Panichchkikosolkul, W. (2017), Confidence Intervals for the Coefficient of Variation in a Normal Distribution with a Known Mean and a Bounded Standard Deviation. KMUTND, *International Journal of Applied. Sciences and. Technology*, **10**(2):1-10.

(21) Pearson, E.S. (1932), Comparison of A.T. McKay approximation with experimental sampling results, *Journal of the Royal Statistical Society*, **95**(4):703-704.

(22) Rahim, M., Hossein, H., and Rob, W. (2007), Is the sample coefficient of variation a good estimator for the population coefficient of variation? Cardiff University, UK. http://mpra.ub.uni-muenchen.de/6106/.

(23) Reid, N. (1996), The roles of conditioning in inference. *Journal of the Royal Statistical Society*, **B 54**:701-738.

(24) Sharma, K. K. and Krishna, H. (1994), Asymptotic sampling distribution of inverse coefficient of variation and its applications. *IEEE Trans. Reliability*, **43**(4):630-633.

(25) Tian, L. (2006), Inferences on the within-subject coefficient of variation. *Statistics in Medicine*, **25**:2008-2017.

(26) Umphrey, G.J. (1983), A comment on McKay's approximation for the coefficient of variation. *Communications in Statistics-Simulation and Computation*, **12**(5):629-635.

(27) Vangel, M.G. (1996), Confidence intervals for a coefficient of variation. *The American Statistician*, **15**(1):21-26.

(28) Wackerly, D. D., Mendenhall, W., and Scheaffer, R. L. (2008). *Mathematical statistics with applications*. Belmont, CA: Thomson Brooks/Cole.

```
function [CP1,EL1,CP2,EL2,CP3,EL3,CP4,EL4]=simcpel(cv,n,alpha,M);
cp1=0;cp2=0;cp3=0;cp4=0;
el1=0;el2=0;el3=0;el4=0;
for i=1:M
X=poissrnd((1/cv)^2,n,1);
ci1=ciwcc(X,alpha);
ci2=ciwb(X,alpha);
ci3=ciwz(X,alpha);
ci4=ciwccz(X,alpha);
if (cv>=ci1(:,1)&cv<=ci1(:,2));
   cp1=cp1+1;
end,
el1=el1+(ci1(:,2)-ci1(:,1));
 if (cv>=ci2(:,1)&cv<=ci2(:,2));
   cp2=cp2+1;
end,
el2=el2+(ci2(:,2)-ci2(:,1));
 if (cv>=ci3(:,1)&cv<=ci3(:,2));
   cp3=cp3+1;
end,
el3=el3+(ci3(:,2)-ci3(:,1));
 if (cv>=ci4(:,1)&cv<=ci4(:,2));
   cp4=cp4+1;
end,
el4=el4+(ci4(:,2)-ci4(:,1));
end
CP1=cp1/M;EL1=el1/M;
CP2=cp2/M;EL2=el2/M;
CP3=cp3/M;EL3=el3/M;
CP4=cp4/M;EL4=el4/M;
```

```
function CIwcc=ciwcc(X,alpha);
n=length(X);
c=1-alpha/2;
Zc=norminv(c,0,1);
Xb=mean(X);
L=(Xb+Zc*sqrt((Xb+0.5)/n))^(-1/2);
U=(Xb-Zc*sqrt((Xb+0.5)/n))^(-1/2);
CIwcc=[L,U];
```

```
function CIwB=ciwb(X,alpha);
n=length(X);
c=1-alpha/2;
Zc=norminv(c,0,1);
Xb=mean(X);
B=2000;
SB=bootstrp(B,@mean,X);
XB=mean(SB);
L=(Xb+Zc*sqrt(XB/n))^(-1/2);
U=(Xb-Zc*sqrt(XB/n))^(-1/2);
```

```
CIwB=[L,U];
```

```
function CIwz=ciwz(X,alpha);
n=length(X);
c=alpha/2;
Xb=mean(X);
B=2000;
SB=bootstrp(B,@mean,X);
XB=mean(SB);
sigB=std(SB);
T=(SB-XB)/sigB;
t1_c=quantile(T,1-c);
tc=quantile(T,c);
L=(Xb+t1_c*sqrt(Xb/n))^(-1/2);
U=(Xb+tc*sqrt(Xb/n))^(-1/2);
CIwz=[L,U];
```

```
function CIwccz=ciwccz(X,alpha);
n=length(X);
c=alpha/2;
Xb=mean(X);
B=2000;
SB=bootstrp(B,@mean,X);
XB=mean(SB);
sigB=std(SB);
T=(SB-XB)/sigB;
t1_c=quantile(T,1-c);
tc=quantile(T,c);
L=(Xb+t1_c*sqrt((Xb+0.5)/n))^(-1/2);
U=(Xb+tc*sqrt((Xb+0.5)/n))^(-1/2);
CIwccz=[L,U];
```

كلية العلوم

جامعة طرابلس

قسم الاحصاء

دراسة محاكاة للمقارنة بين فترات الثقة لتقدير معامل الاختلاف للتوزيع بواسون

نجلاء نوري إبراهيم الكشيك

الأستاذ المشرف : أ.د. يوسف محمد امحمد الاصبيعي

قدمت هذه الرسالة استكمالا لمتطلبات الإجازة العليا (الماجستير) في الإحصاء

ربيع 2022

الملخص :

تهدف هذه الأطروحة إلى توضيح ومقارنة الطرق المختلفة لإنشاء فترات ثقة من الطرفين لمعامل الاختلاف (CV) أو لنسبة الإشارة إلى الضوضاء (SNR) لتوزيع بواسون بناءً على تقنية محاكاة مونتي كارلو.

كفصل تمهيدي لهذه الأطروحة ، تم تصميم الفصل الأول ليشمل مقدمة شاملة. تم تخصيص الفصل الثاني لاستعراض الجوانب والتعريفات لجميع الطرق السبع المعتمدة لفترات الثقة لمعامل بواسون للتغير (CV) وكذلك فترات الثقة لنسبة الإشارة إلى الضوضاء (SNR) كمقلوب لمعامل الاختلاف.

الفصل الثالث مخصص لجزء التطبيق حيث سيتم إجراء دراسة المحاكاة لمقارنة أداء الطرق الأربع المدروسة وهي: والد مع تصحيح الاستمرارية (WCC) وطريقة البوتستراب ($WaldB$) و والوالد زد ($WaldZ$) و والد مع الاستمرارية طريقة البوتستراب ($WCCZ$) لإنشاء فترات الثقة من الطرفين بنسبة 95٪ لمعامل الاختلاف لتوزيع بواسون بأحجام عينات مختلفة وقيم معلمات مختلفة.

والفصل الأخير من هذه الأطروحة ، تم تخصيص الفصل الرابع للجزء الأخير من هذه الدراسة التي توضح أهم المناقشات والاستنتاجات ثم مخططًا للعمل المستقبلي المحتمل الذي يمكن من خلاله توسيع هذه الدراسة.

أخيرًا ، سيتم إعطاء وظائف ماتلاب لأساليب فترات الثقة الأربعة في الفصول اللاحقة.

تناولت هذه الأطروحة عدة فترات ثقة لتقدير معامل الاختلاف لتوزيع بواسون. سيتم إجراء دراسة محاكاة لمقارنة أداء الطرق الأربعة لفترات الثقة المقترحة.

سيتم إنشاء البيانات من توزيع بواسون لمعامل الاختلاف = 0.05 ، 0.1 ، 0.15 ، ... ، 0.45 ، و 0.50 باستخدام برنامج ماتلاب سيتم حساب احتمال التغطية وطول فترة الثقة لكل طريقة من طرق فترات الثقة.