Two-Side Confidence Intervals for the Poisson Means

Manlika Tanusit

***Abstract*—In this paper we consider interval estimation for the Poisson means. The following confidence intervals are considered: Wald CI, Score CI, Score continuity correction CI, Agresti and Coull CI, Bayes Wald CI, Bayes Score CI, and Bayes Score continuity correction CI. Each interval is examined for its coverage probability and its expected lengths. Based on this simulation; we recommend Score CI and the Wald CI for the small n and the larger n respectively.**

***Index Terms*—Bayes estimators, poisson distribution, interval estimation.**

I. Introduction

Estimation is one of the main branches of statistics. They refer to the process by which one makes inference about a population, based on information obtained from a sample. Statisticians use sample statistics to estimate population parameter. For example; sample means are used to estimate population mean; sample proportions, and to estimate proportions. An estimate of population parameters may be expressed in two ways; point and interval estimate. A point estimate of a population parameter is a single value of a statistic, for example, the sample mean $\bar{x}$ is a point estimate of the population mean μ. Similarly, the sample proportion p is a point estimate of the population proportion P. An interval estimate is defined by two numbers, between which a population parameter is to lie, for example, $a < \bar{x} < b$ is an interval estimate of the population mean μ. It indicates that the population mean is greater than a bit less than b. Statisticians use a confidence interval to express the precision and uncertainty associated with a particular sampling method, for example, we might say that we are 95% confidence that the true population mean falls within a specified range. This statement is a confidence interval. It mean that if we used the same sampling method to select different samples and compute different interval estimates, the true population mean would fall within a range defined by the sample statistic $\pm$ margin of error 95% of the time.

In this paper, we are compare seven method confidence interval (CIs) for the Poisson means, Poisson distribution is a discrete probability distribution. A poisson random variable (X) representing the number of successes occurring in a given time interval or a specified region of space.

The probability distribution is

$$P(X) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

where x=0, 1, 2,…, e=2.71828 (but use your calculator's e button), and λ =mean number of successes in the given time interval or region of space.

Confidence interval (CI) of Poisson means that popular is

$$\hat{\lambda} \pm z_{\frac{\alpha}{2}}\sqrt{\hat{\lambda}/n}$$

and $z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution. This formula, we call Wald method, it is easy to present and compute but it has poor coverage properties for small n. Brown [1] showed that Wald CI is actually far too poor and unreliable and the problem are not just for small n. [2] showed that Wald CI suffers from a series systematic bias in the coverage. Therefore, the other researcher are present any method for solve this problem; see [3]-[5], [6]. We have known that common method used maximum likelihood estimator. So, we require to developed the CIs by Bayes estimator with Wald CI, Score CI and Score continuity correction CI which the prior distribution of λ is assumed to be gamma distribution with α and β parameters. [7] showed that the optimal values of α and β are their value provide $\alpha\beta$ closed to the λ parameter. The α value is always higher than the β value, the α lies between 4.0 and 5.0.

II. The Confidence Intervals

1) Create random variable that has Poisson distribution.
2) Compute interval estimation for a Poisson distribution seven methods:

- Wald CI

$$\hat{\lambda} \pm Z\sqrt{\frac{\hat{\lambda}}{n}}. \tag{1}$$

- Score CI

$$\frac{\left(n\hat{\lambda} + z^2/2\right)}{n} \pm zn^{-1/2}\left(\hat{\lambda} + \frac{z^2}{4n}\right)^{1/2}. \tag{2}$$

- Score continuity correction CI

$$\frac{\left(2n\hat{\lambda} + Z^2 - 1\right) \mp Z\sqrt{Z^2 + 4n\hat{\lambda} \mp 2}}{2n}. \tag{3}$$

- Agresti and Coull CI

$$\frac{X + z^2/2}{n} \pm z\left(\frac{X + z^2/2}{n}\right)n^{-1/2}. \tag{4}$$

- Bayes Wald CI

$$\tilde{\lambda} \pm Z\sqrt{\frac{\tilde{\lambda}}{n}}. \tag{5}$$

Manuscript received July 10, 2012; revised August 16, 2012.
M. Tanusit is with the Major of Statistics, Faculty of Science, Maejo University, Chiang Mai 50290, Thailand (e-mail: manlika@mju.ac.th).

- Bayes Score CI

$$\frac{\left(\tilde{n}\tilde{\lambda}+z^2/2\right)}{\tilde{n}} \pm z\tilde{n}^{-1/2}\left(\tilde{\lambda}+\frac{z^2}{4\tilde{n}}\right)^{1/2}. \qquad (6)$$

- Bayes Score continuity correction CI

$$\frac{\left(2\tilde{n}\tilde{\lambda}+Z^2-1\right)\mp Z\sqrt{Z^2+4\tilde{n}\tilde{\lambda}\mp 2}}{2\tilde{n}}. \qquad (7)$$

Note that for (a)-(c) $\hat{\lambda}=\bar{x}$; $\hat{\lambda}$ is well known to be the MLE of λ and z is the $100(1-\alpha/2)$ percentile of the standard normal distribution

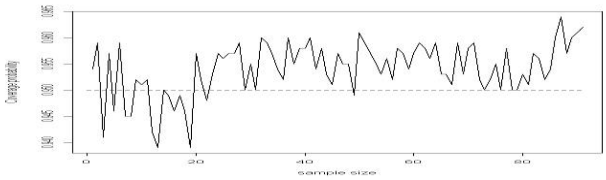
For (e)-(g) $\tilde{\lambda}=\frac{(\alpha+X)\beta}{n\beta+1}$ and $\tilde{n}=n\beta+1$.

$\tilde{\lambda}$ is Bayes' estimator of Poisson mean, and X is number of successes occurring in a given time interval or a specified region of space.

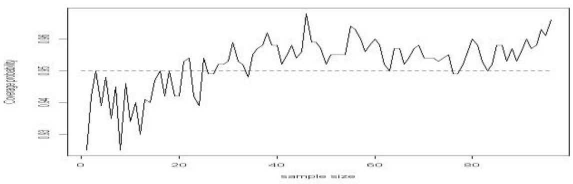
III. Simulation Results

The performance of the estimated coverage probabilities of the confidence interval (a)-(f) and their expected lengths were examined via Monte Carlo simulation. Data were generate from Poisson distribution with $\lambda=1,2,3,4,5$ and sample size n=10 to 100. All simulations were performed using programs written in the R Statistical software, repeated 1,000 times in each case at level of significance $\alpha=0.01, 0.05,$ and 0.10. The simulation results are shown in Figure 1-2 and Table 1.

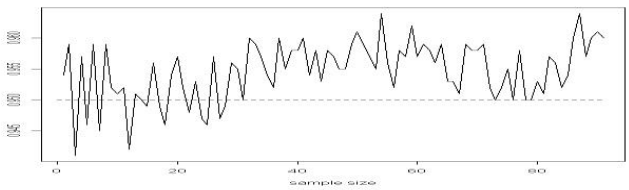
Fig. 1 show that the coverage probability of the Wald CI, Score CI, Agresti and Coull CI, and Score continuity correction CI for fixed $\lambda=3$ and variable n from 10 to 100. Naively, are may expect that the coverage probability gets systematically near the nominal level as the sample size n increases.



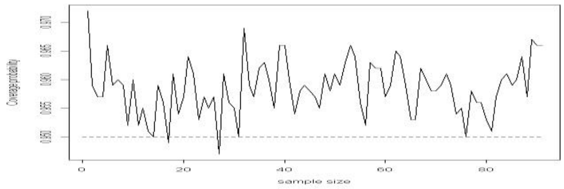
Wald CI



Score CI

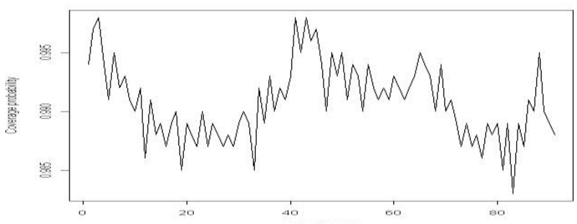


Agresti and Coull CI

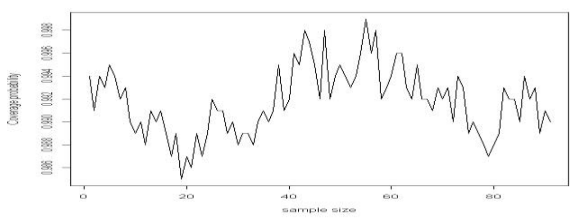


Score continuity correction CI

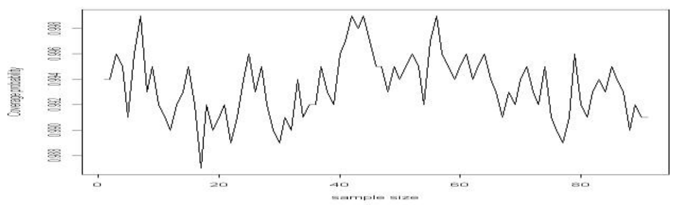
Fig. 1 Coverage probability of four intervals for a poisson mean with $\alpha=0.05$ and $n=10$ to 100.



Bayes Wald CI



Bayes Score CI



Bayes Score continuity correction CI

Fig. 2. Coverage probability of three intervals for a poisson mean with $\alpha=0.05$ and $n=10$ to 100.

Fig. 2. Show that the coverage probability of the alternative intervals with bayes estimator gives the coverage probability upper nominal level.

In addition, the expected lengths of the score CI and the Wald CI are much shorter than the other CIs when n are small and larger respectively. The expected length increase as the value λ gets larger (e.g. For Score CI, n=10 and $\alpha = 0.05$; 0.23284 for $\lambda = 1$; 0.327333 for $\lambda = 3$; and 0.482786 for $\lambda = 5$). Moreover, when the sample size increase, the expected length is shorter (e.g. For Score CI, $\lambda = 1$ and; 0.23284 for n=10; 0.102252 for n=50; and 0.072279 for n=100).

IV. Conclusions

Our main objective is to compare the CIs under different situation; we compare the performance of all the CIs described above under various situations and three different confidence levels $\alpha = 0.05, 0.01, 0.10$. We found that, Wald CI is below the nominal level, but it has smallest expected length for n are small (n<30), Score CI , Score corrected continuity CI , Agresti and Coull CI are above the nominal level and they have expected length slightly higher Wald CI. Three alternatives with Bayes estimator have a higher CI and they are similar with each other and do a good job, so, we recommend Score CI and the Wald CI for the small n and the larger n respectively**.**

Acknowledgment

This work was supported by Faculty of Science, Maejo University, Chiang Mai, Thailand.

References

[1] D. B. Lawrence, T. T. Cai, and A. Dasgupta, "Interval estimation for a binomial Proportion," *Statistica Sinica*, vol. 16, pp. 101-133, 2003.

[2] T. T. Cai, "One-sided confidence intervals in discrete distributions," *J. of Statistical planning and inference*, vol. 131, pp. 63-88, 2005.

[3] M. Tanusit, "A Comparative Study of Parameter Inference Methods for Poisson distribution: Small Sample Size," *Advance and Applications in Statistics*, vol. 14, pp. 49-55, 2010.

[4] T. D. Ross, "Accurate confidence intervals for binomial proportion and Poisson rate estimation," *Comput. Biol. Med*, vol. 33, pp. 509-531, 2003.

[5] J. D. Stamey, J. W. Seaman, and D. M. Young, "Bayesian analysis of complementary Poisson rate parameters with data subject to misclassification," *Statist. Plann. Inference*, vol. 134, pp. 36-48, 2005.

[6] G. D. Wardell, "Small Sample Interval Estimation of Bernoulli and Poisson Parameters," *The American Statistician*, vol. 51, pp. 321-325, 1997.

[7] W. Panichkitkosolkul, "The Optimal Parameters of Prior Distribution for Bayes' Estimator of Poisson Mean," *SWU Sci. J.*, vol. 23, pp. 23-31, 2007.

M. Tanusit received the B. S. degree in Statistics from Kasetsart University, Bangkok City, Thailand, in 2008. Currently, she is a lecturer at the Department of Statistics, Maejo University, Chiang Mai, Thailand. His research interests include stability and stabilization of dynamical systems, qualitative theory of differential/discrete-time systems.
Email: manlika@mju.ac.th