REVSTAT – Statistical Journal
Volume 10, Number 2, June 2012, 211–227

COMPARISON OF CONFIDENCE INTERVALS FOR THE POISSON MEAN: SOME NEW ASPECTS

Authors: V.V. Patil
– Department of Statistics, Tuljaram Chaturchand College,
Baramati, Maharashtra, India
vaishu_kop@indiatimes.com

H.V. Kulkarni
– Department of Statistics, Shivaji University,
Kolhapur, Maharashtra, India
kulkarnih1@rediffmail.com

Received: February 2010 Revised: May 2011 Accepted: January 2012

Abstract:

- We perform a comparative study among nineteen methods of interval estimation of the Poisson mean, in the intervals (0,2), [2,4] and (4,50], using as criteria coverage, expected length of confidence intervals, balance of noncoverage probabilities, E(P-bias) and E(P-confidence). The study leads to recommendations regarding the use of particular methods depending on the demands of a particular statistical investigation and prior judgment regarding the parameter value if any.

Key-Words:

- *confidence intervals (CI); Poisson mean; comparison.*

AMS Subject Classification:

- 62-07, 62F25.

1. INTRODUCTION

Construction of CIs in discrete distributions is a widely addressed problem. The standard method of obtaining a $100 \times (1-\alpha)\%$ CI for the Poisson mean μ is based on inverting an equal tailed test for the null hypothesis $H_0 \colon \mu = \mu_0$. This is an "exact" CI, in the sense that it is constructed using the exact distribution.

Exact CIs are very conservative and too wide. A large number of alternate methods for obtaining CIs for μ based on approximations for the Poisson distribution are suggested in the literature to overcome these drawbacks. Desirable properties of those approximate CIs are:

- for $(1-\alpha)$ confidence interval the infimum over μ of the coverage probability should be equal to $(1-\alpha)$;
- confidence interval can not be shortened without the infimum of the coverage falling below $(1-\alpha)$.

We attempt to perform an exhaustive review of the existing methods for obtaining confidence intervals for the Poisson parameter and present an extensive comparison among these methods based on the following criterion:

1) Expected length of confidence intervals (E(LOC)),

2) Percent coverages (Coverage),

3) E(P-bias) and E(P-confidence),

4) Balance of right and left noncoverage probabilities.

Section 2 enumerates several methods for interval estimation of μ, giving appropriate references. Section 3 describes criteria used for comparison, Section 4 reports details of the comparative study and Section 5 presents concluding remarks.

2. A REVIEW OF THE EXISTING METHODS

Table 1 presented next reports 19 CIs for the Poisson mean. In the table, "Schwertman and Martinez" is abbreviated as SM, "Freeman and Tukey" by FT, "Wilson and Hilferty" by WH, "continuity correction" by CC, "Second Normal" by SN and "Likelihood Ratio" by LR. Furthermore, $\alpha_1 = \alpha/2$, $\alpha_2 = 1-\alpha/2$, and $x_c = x + c$ for any number c.

Table 1: Confidence limits for the nineteen methods.

Name and reference	Lower Limit	Upper Limit
1: Garwood (GW) (1936)	$(\chi^2_{(2x,\alpha_1)})/2$	$(\chi^2_{(2x_1,\alpha_2)})/2$
2: WH (WH) (1931)	$x\left(1-1/9x+Z_{\alpha_1}/3\sqrt{x}\right)^3$	$x_1\left(1-1/9x_1+Z_{\alpha_2}/3\sqrt{x_1}\right)^3$
3: Wald (W) SM (1994)	$x+Z_{\alpha_1}\sqrt{x}$	$x+Z_{\alpha_2}\sqrt{x}$
4: SN (SN) SM (1994)	$x+Z^2_{\alpha_1}/2+Z_{\alpha_1}\sqrt{x+Z^2_{\alpha_1}/4}$	$x+Z^2_{\alpha_1}/2+Z_{\alpha_2}\sqrt{x+Z^2_{\alpha_2}/4}$
5: Wald CC (FNCC) SM (1994)	$x_{-0.5}+Z_{\alpha_1}\sqrt{x_{-0.5}}$	$x_{0.5}+Z_{\alpha_2}\sqrt{x_{0.5}}$
6: SN CC (SNCC) SM (1994)	$x_{-0.5}+Z^2_{\alpha_1}/2+Z_{\alpha_1}\left(x_{-0.5}+Z^2_{\alpha_2}/4\right)^{.5}$	$x_{0.5}+Z^2_{\alpha_2}/2+Z_{\alpha_2}\left(x_{0.5}+Z^2_{\alpha_1}/4\right)^{.5}$
7: Molenaar (MOL) (1970)	$x_{-0.5}+\left(2Z^2_{\alpha_1}+1\right)/6+Z_{\alpha_1}\left(x_{-0.5}+\left(Z^2_{\alpha_1}+2\right)/18\right)^{.5}$	$x_{0.5}+\left(2Z^2_{\alpha_2}+1\right)/6+Z_{\alpha_2}\left(x_{0.5}+\left(Z^2_{\alpha_2}+2\right)/18\right)^{.5}$
8: Bartlett (BART) (1936)	$\left(\sqrt{x}+Z_{\alpha_1}/2\right)^2$	$\left(\sqrt{x}+Z_{\alpha_2}/2\right)^2$
9: Vandenbroucke (SR) (1982)	$\left(\sqrt{x_c}+Z_{\alpha_1}/2\right)^2$	$\left(\sqrt{x_c}+Z_{\alpha_2}/2\right)^2$
10: Anscombe (ANS) (1948)	$\left(\sqrt{x+3/8}+Z_{\alpha_1}/2\right)^2-3/8$	$\left(\sqrt{x+3/8}+Z_{\alpha_2}/2\right)^2-3/8$
11: FT (FT) (1950)	$0.25\left(\left(\sqrt{x}+\sqrt{x_1}+Z_{\alpha_1}\right)^2-1\right)$	$0.25\left(\left(\sqrt{x}+\sqrt{x_1}+Z_{\alpha_2}\right)^2-1\right)$
12: Hald (H) (1952)	$\left(\sqrt{x_{-.5}}+Z_{\alpha_1}/2\right)^2+.5$	$\left(\sqrt{x_{-.5}}+Z_{\alpha_2}/2\right)^2+.5$
13: Begaud (BB) (2005)	$\left(\sqrt{x_{.02}}+Z_{\alpha_1}/2\right)^2$	$\left(\sqrt{x_{.96}}+Z_{\alpha_2}/2\right)^2$
14: Modified Wald (MW) Barker (2002)	For $x=0$; 0 For $x>0$; Wald limit	For $x=0$; $-\log(\alpha_1)$ For $x>0$; Wald limit
15: Modified Bartlett (MB) Barker (2002)	For $x=0$; 0 For $x>0$; Bartlett limit	For $x=0$; $-\log(\alpha_1)$ For $x>0$; Bartlett limit
16: LR (LR) Brown *et al.* (2003)	No closed form	No closed form
17: Jeffreys (JFR) Brown *et al.* (2003)	$G\left(\alpha_1, x_{0.5}, 1/r\right)$	$G\left(\alpha_2, x_{0.5}, 1/r\right)$
18: Mid-P Lancaster (1961)	No closed form	No closed form
19: Approximate Bootstrap Confidence (ABC) Swift (2009)	$x+\dfrac{Z_0+Z_{\alpha_1}}{\left(1-a(Z_0+Z_{\alpha_1})\right)^2}\sqrt{x}$ where $a=Z_0=1/(6\sqrt{x})$	$x+\dfrac{Z_0+Z_{\alpha_2}}{\left(1-a(Z_0+Z_{\alpha_2})\right)^2}\sqrt{x}$

3. CRITERIA FOR COMPARISON

The criteria considered for the comparison among the above mentioned CIs are E(LOC) of CIs, coverage probability, ratio of the left to right noncoverage probabilities, E(P-confidence) and E(P-bias).

Here we explain the details of the three criterion for comparison mentioned in Section 1. Without loss of generality a sample of size $n = 1$ is considered. The comparisons are carried out over $\mu \in (0, 50]$.

The expected value of a function $g(x)$ is computed as $\sum_{x=0}^{\infty} g(x)\, p_\mu(x)$ where $p_\mu(x) = e^{-\mu}\mu^x/x!$. The infinite sums in the computation of these quantities were approximated by appropriate finite ones up to 0.001 margin of error.

The coverage probability $C(\mu)$, noncoverage probability on the left $L(\mu)$, noncoverage probability on the right $R(\mu)$, and corresponding expected length E(LOC) of a CI $\big(l(x), u(x)\big)$ are respectively computed by taking $g(x) = I\big(l(x) \le \mu \le u(x)\big)$, $I\big(\mu > u(x)\big)$, $I\big(\mu < l(x)\big)$ and $\big(u(x) - l(x)\big)$, where $I(\cdot)$ is the indicator function of the bracketed event.

3.1. Computation of E(P-confidence) and E(P-bias)

Let CI(x) be the CI obtained for the observation x having nominal level $(1-\alpha)\,100\%$. The P-bias and P-confidence are defined in terms of the standard equal tailed P-value function $P(\mu, x) = \min\big(2\,P_\mu(X \le x), 2\,P_\mu(X \ge x), 1\big)$. The P-confidence of the CI that measures how strongly the observation x rejects parameter values outside CI is defined as $C_p\big(\mathrm{CI}(x), x\big) = \big(1 - \sup_{\mu \notin \mathrm{CI}(x)} P(\mu, x)\big) \times 100\%$.

The P-bias of a CI which quantifies the largeness of P-values for values of μ outside the CI in comparison with those inside the CI is given by $b\big(\mathrm{CI}(x), x\big) = \max\big(0,\ \sup_{\mu \notin \mathrm{CI}(x)} P(\mu, x) - \inf_{\mu \in \mathrm{CI}(x)} P(\mu, x)\big) \times 100\%$. For the Poisson distribution $P(\mu, x)$ is continuous and a monotone function in μ in opposite directions to the left and right of the interval for each value of x. Hence the supremums and infimums occur at the upper or lower end points of the CIs. Consequently the formulae of P-bias and P-confidence are reduced to

$$C_p\big(\mathrm{CI}(x), x\big) = \Big(1 - \max\Big\{2\,P\big(X \ge x;\ \mu = l(x)\big), 2\,P\big(X \le x;\ \mu = u(x)\big)\Big\}\Big) \times 100\%\,,$$

$$b\big(\mathrm{CI}(x), x\big) = \max\Big\{0, \Big\{2\,P\big(X \ge x;\ \mu = l(x)\big) - 2\,P\big(X \le x;\ \mu = u(x)\big)\Big\}\Big\} \times 100\%\,.$$

Their expected values are computed as described above.

It was observed that when the actual value of μ is a fraction, the CI with their endpoints rounded to the nearest integer (for lower limit, rounding to an integer less than the limit and reverse for the upper limit) improved coverage probabilities to a very large extent at the cost of increasing E(LOC) at most by one unit. This is clearly visible from Figure 1 which displays the Box plot of coverages for the rounded and unrounded CIs obtained using Wald method. Similar pattern was observed for other methods.

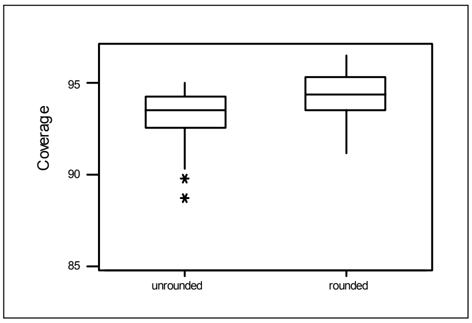


Figure 1: Impact of rounding on coverage of Wald CI.

Consequently the E(LOC) and percent coverages reported here correspond to these rounded intervals and the comparison carried out among the methods in the sequel is based on rounded intervals.

4. COMPARISON AMONG THE METHODS

4.1. Comparison based on coverages and E(LOC)

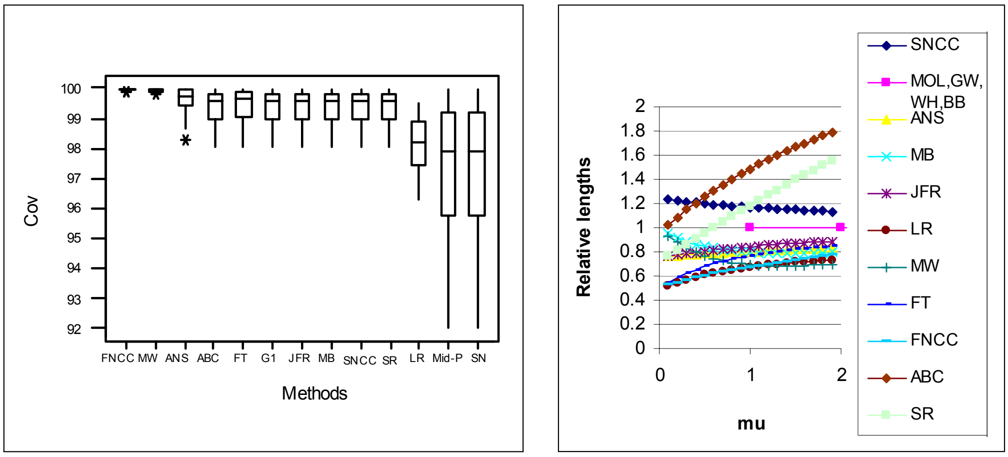
On careful examination revealed that different methods perform differently in certain subsets of the parameter space.

Consequently the performance of each method was studied separately on the three regions, namely (0,2), [2,4] and (4,50] in the parameter space. Panels (a) and (b) of Figures 2A to 4A display respectively the boxplots of coverages and graphs of relative E(LOC) of conservative methods (i.e. ratio of E(LOC) for the concerned method to the same for Garwood exact CI) for different regions defined above. Figures 2B to 4B display similar plots for nonconservative methods.

The observations from these graphs are tabulated in Table 2. The methods displayed in bold face have shortest length among the concerned group. Here $G1 = \{\mathrm{GW, MOL, WH, BB}\}$ and $G2 = \{\mathrm{BART, W, H}\}$.

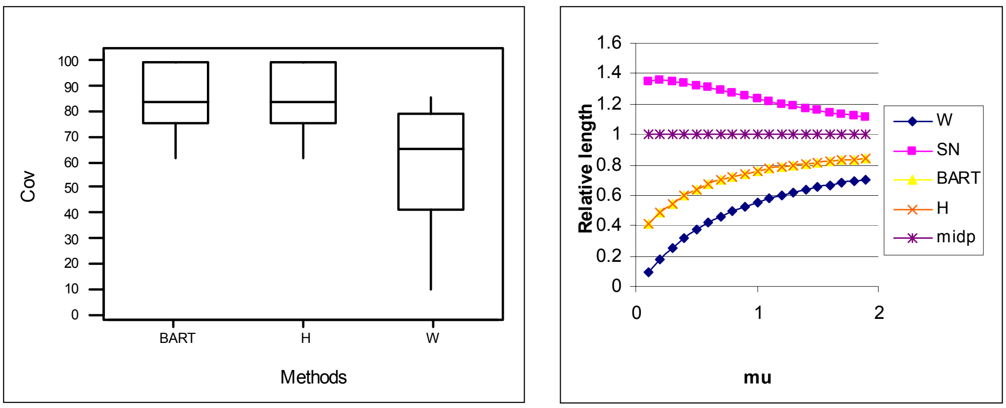
Table 2: Coverage performance of the nineteen methods.

Type	$\mu \in (0{,}2)$	$\mu \in [2{,}4]$	$\mu \in (4{,}50]$
Conservative	**FNCC**, **LR**, ANS, $G1$, FT, JFR, MB, MW, SN SNCC, ABC, SR, Mid-P	**MB**, **ANS**, SN, $G1$ ABC, SR, JFR SNCC, Mid-P	$\boldsymbol{G1}$, SNCC, ABC, LR H, BART, MW, ANS FT, MB, SN, Mid-P JFR, FNCC, W, SR
Non-Conservative	$G2$	$G2$, FNCC, FT, LR MW	—



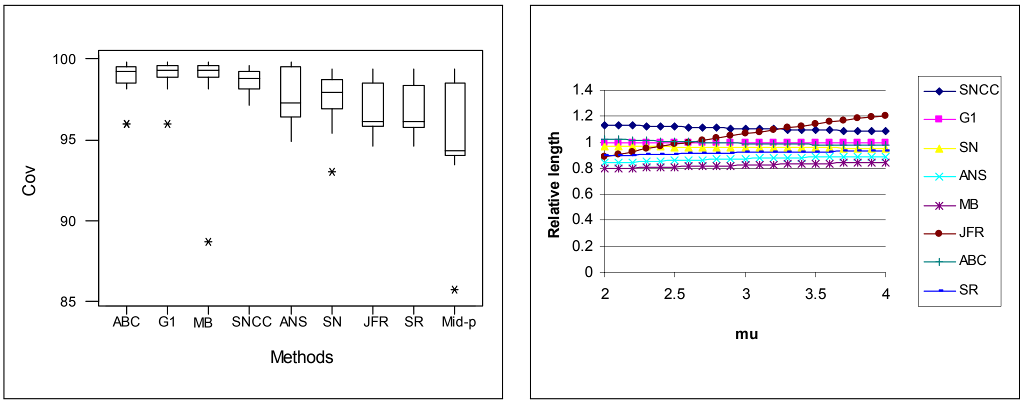
(a) Boxplot of coverages for conservative methods. **(b)** Relative lengths of conservative methods.

Figure 2A: Coverages and relative E(LOC) for conservative methods for parametric space (0,2), where $G1 = \{\text{GW}, \text{MOL}, \text{WH}, \text{BB}\}$.



(a) Boxplot of coverages for nonconservative methods. **(b)** Relative lengths of nonconservative methods.

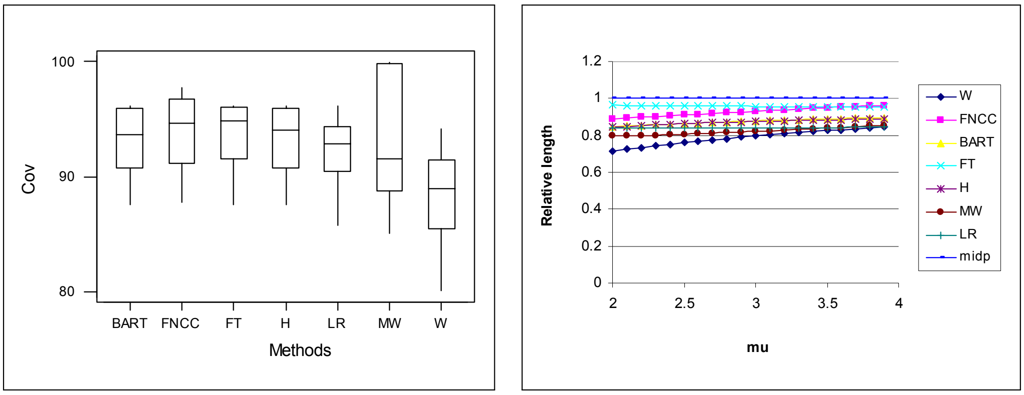
Figure 2B: Coverages and relative E(LOC) for nonconservative methods for parametric space (0,2).



(a) Boxplot of coverages for conservative methods.

(b) Relative lengths of conservative methods.

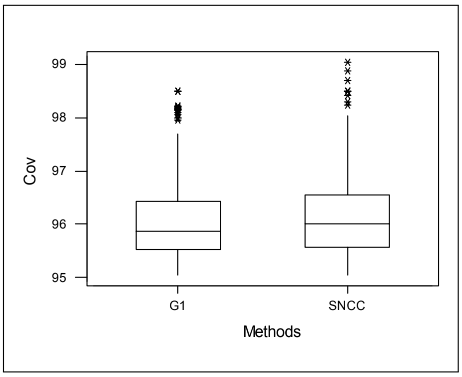
Figure 3A: Coverages and relative E(LOC) for conservative methods for parametric space [2,4], where $G1 = \{\mathrm{GW}, \mathrm{MOL}, \mathrm{WH}, \mathrm{BB}\}$.



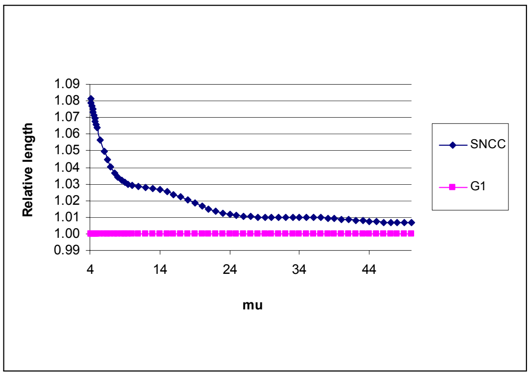
(a) Boxplot of coverages for nonconservative methods.

(b) Relative lengths of nonconservative methods.

Figure 3B: Coverages and relative E(LOC) for nonconservative methods for parametric space [2,4].

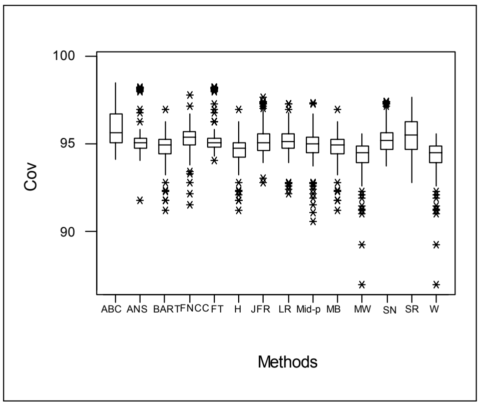


(a) Boxplot of coverages for conservative methods.

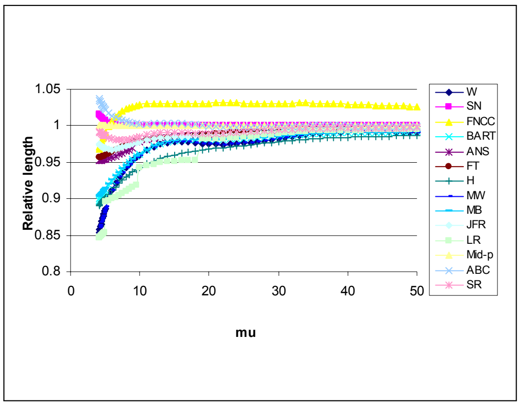


(b) Relative lengths of conservative methods.

Figure 4A: Coverages and relative E(LOC) for conservative methods for parametric space (4,50], where $G1 = \{\mathrm{GW}, \mathrm{MOL}, \mathrm{WH}, \mathrm{BB}\}$.



(a) Boxplot of coverages for nonconservative methods.



(b) Relative lengths of nonconservative methods.

Figure 4B: Coverages and relative E(LOC) for nonconservative methods for parametric space (4,50].

4.2. Comparison with respect to balance of noncoverage probabilities

For a two sided CI procedure it is desirable to have the right and left non-coverage probabilities to be fairly balanced. We plot the ratio of the left to right noncoverage probabilities as a function of Poisson mean for the nineteen methods in Figure 5A and 5B for regions (2,4) and (4,50). For balanced noncoverage, ratio should oscillate in the close neighborhood of 1. For region (0,2) all methods are well below 1, with the exception of Wald method.

A careful observation of figures leads to the following region wise performance of methods with respect to right-to-left noncoverage balance reported in Table 3.

Table 3: Performance based on right-to-left noncoverage balance.

Performance	$\mu \in$ (2,4)	$\mu \in$ (4,50)
Fairly balanced around 1	—	$G1$, ABC, LR, JFR, SR
Uniformly below 1	SNCC, SN, $G1$, ABC, MB	SN, SNCC, Mid-P
Uniformly above 1	LR, JFR, SR, Mid-P FT, ANS, FNCC, MW, $G2$	FT, MB, ANS, FNCC, MW, $G2$

4.3. Comparison based on E(P-bias) and E(P-confidence)

For comparison of methods on the basis of E(P-bias) and E(P-confidence), we consider three regions of sample space (0,2), (2,4) and (4,50). Three panels (a) to (c) of Figures 6 and 7 represent boxplots of E(P-confidence) and E(P-bias) for these three regions. Recommendations on the basis of E(P-bias) and E(P-confidence) for two regions tabulated in Table 4.

Table 4: Recommendations on the basis of E(P-bias) and E(P-confidence).

Performance	$\mu \in$ (0,2)	$\mu \in$ (2,4)	$\mu \in$ (4,50)
Smallest E(P-bias) Largest E(P-confidence)	FNCC, MW, W SNCC, SN, ABC, $G1$	FNCC, SNCC SNCC, SN, $G1$	SNCC, Mid-P, SN SNCC, $G1$, SN

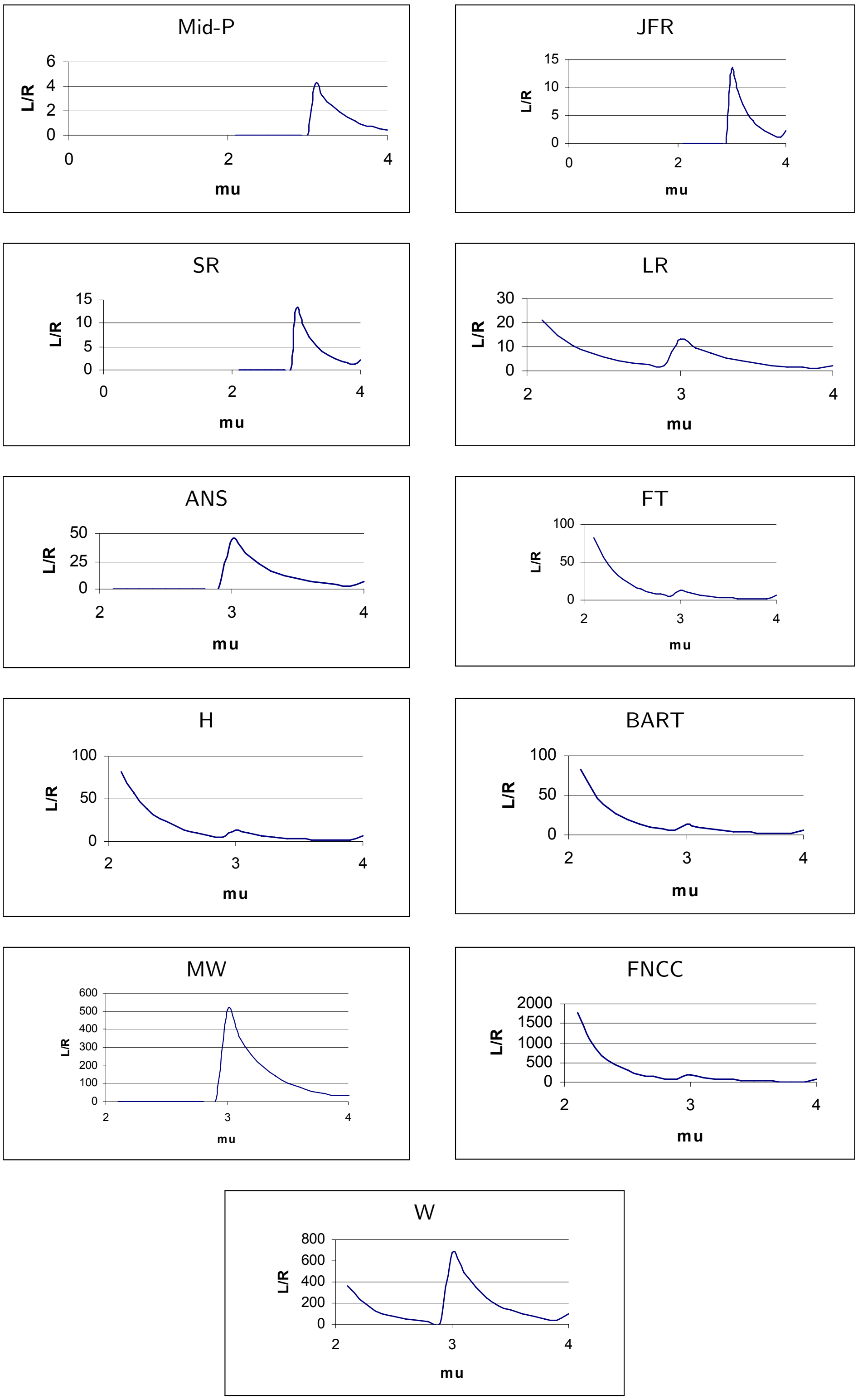
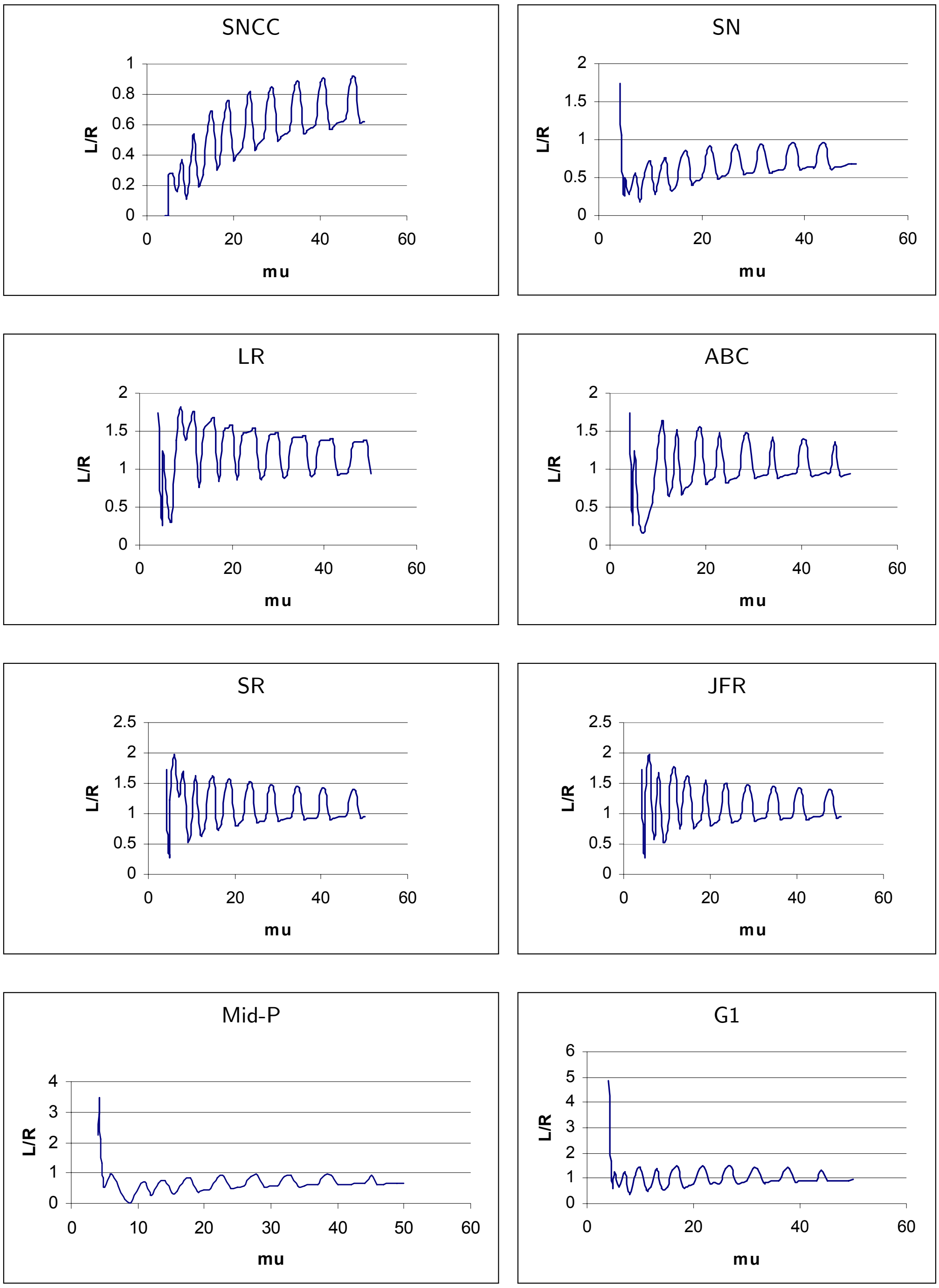


Figure 5A: Graph of ratio of noncoverage probabilities for parametric space (2,4]. The ratio of noncoverage probabilities for methods SNCC, SN, $G1$, ABC and MB are zero for parametric space (2,4].


SNCC
L/R
1
0.8
0.6
0.4
0.2
0
0 20 40 60
mu
SN
L/R
2
1.5
1
0.5
0
0 20 40 60
mu
LR
L/R
2
1.5
1
0.5
0
0 20 40 60
mu
ABC
L/R
2
1.5
1
0.5
0
0 20 40 60
mu
SR
L/R
2.5
2
1.5
1
0.5
0
0 20 40 60
mu
JFR
L/R
2.5
2
1.5
1
0.5
0
0 20 40 60
mu
Mid-P
L/R
4
3
2
1
0
0 10 20 30 40 50 60
mu
G1
L/R
6
5
4
3
2
1
0
0 20 40 60
mu

(continues)

(continued)

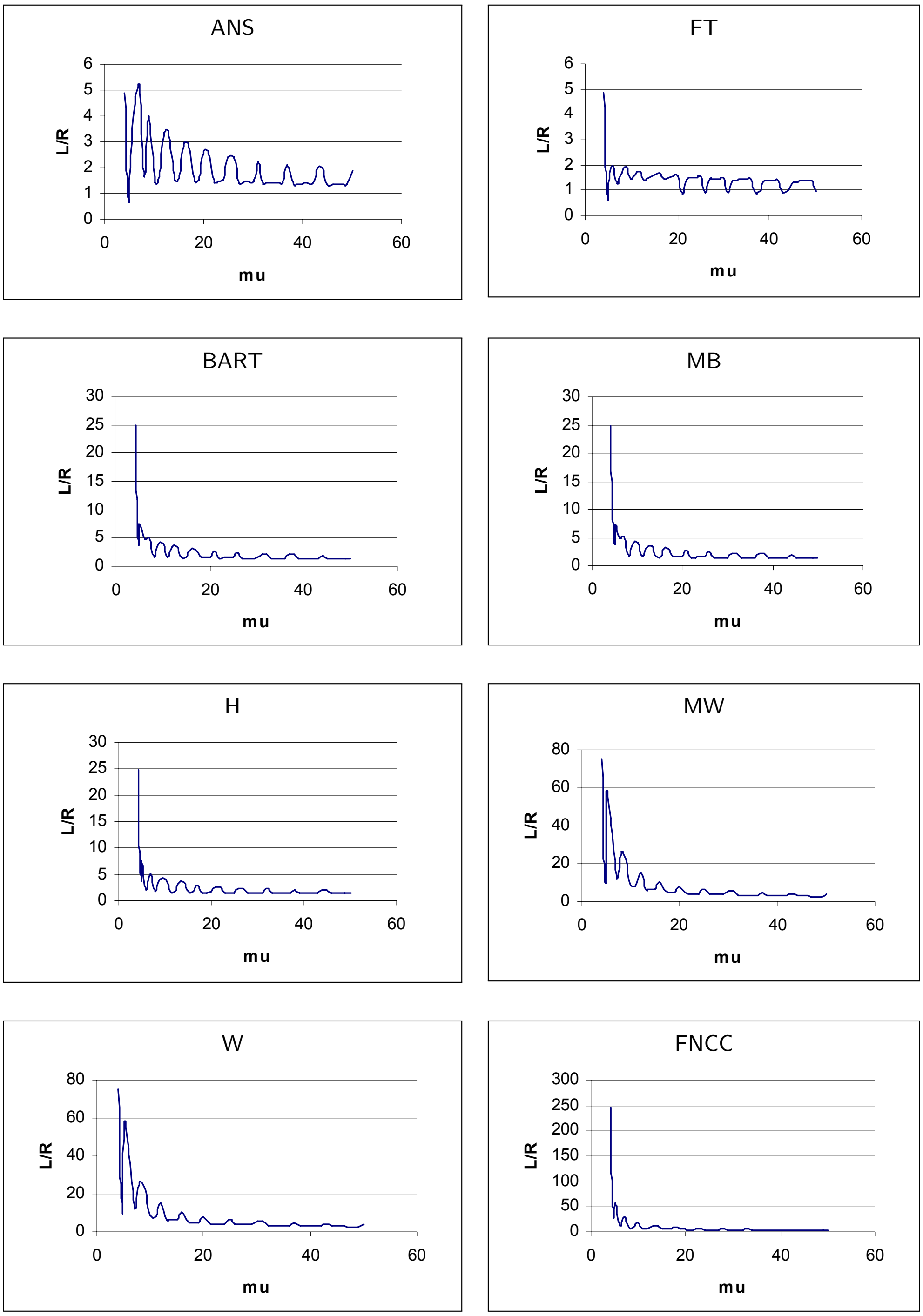


Figure 5B: Graph of ratio of non coverage probabilities for parametric space (4,50].

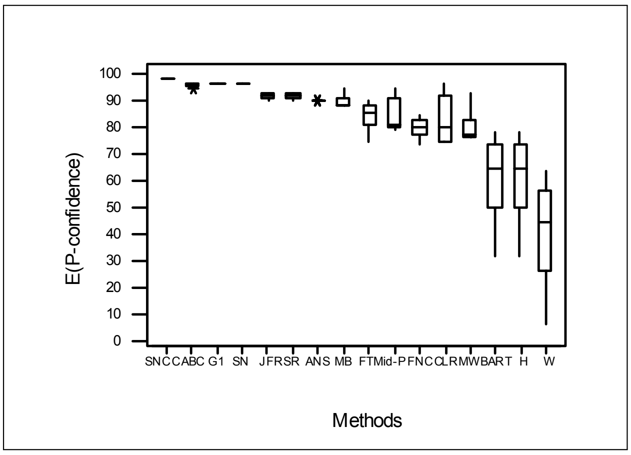


Figure 6(a): Boxplot of E(P-confidence) for parametric space (0,2].

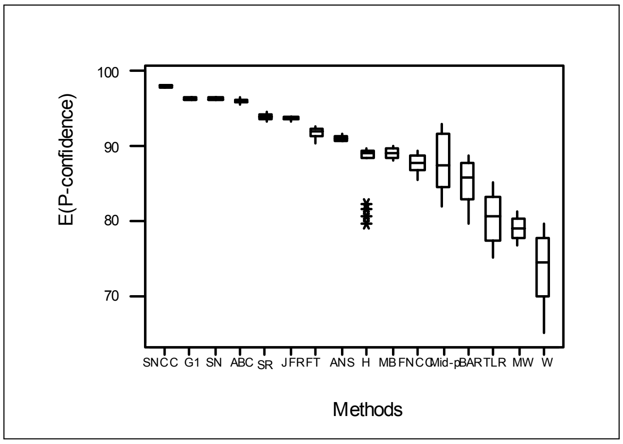


Figure 6(b): Boxplot of E(P-confidence) for parametric space (2,4].

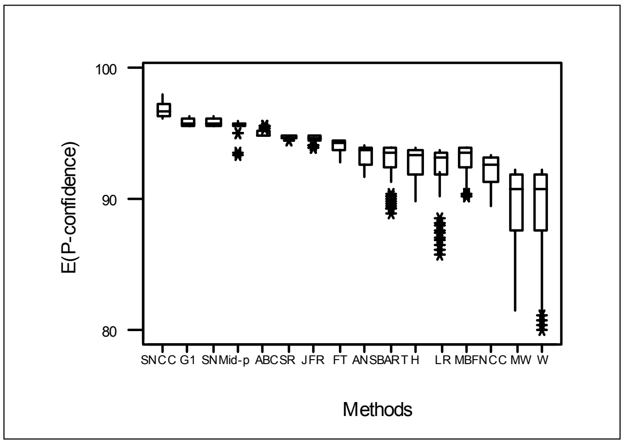


Figure 6(c): Boxplot of E(P-confidence) for parametric space (4,50).

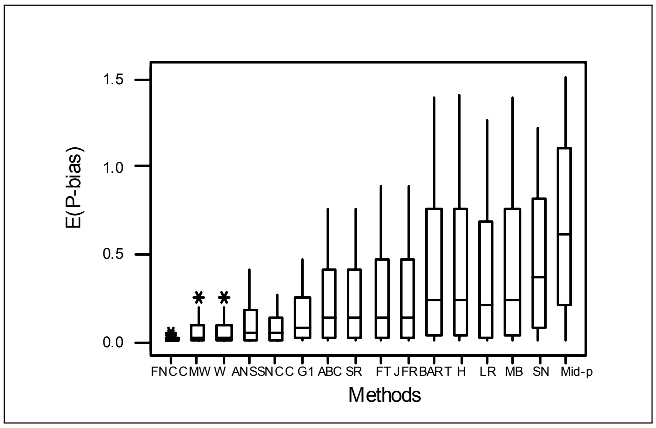


Figure 7(a): Boxplot of E(P-bias) for parametric space (0,2].

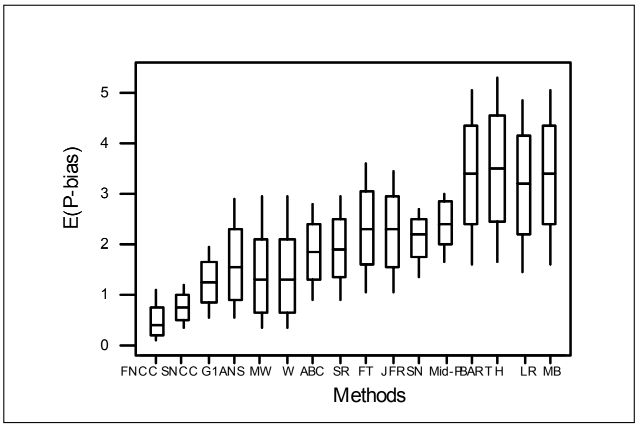


Figure 7(b): Boxplot of E(P-bias) for parametric space (2,4].

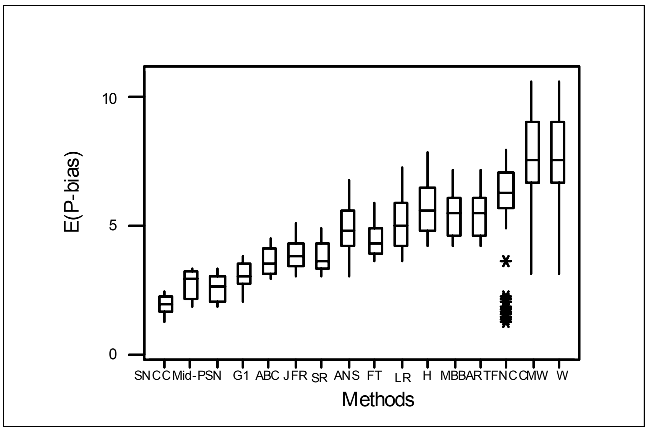


Figure 7(c): Boxplot of E(P-bias) for parametric space (4,50).

5. CONCLUDING REMARKS

Rounding of end points of CI considerably improves the coverages of CI. Our remarks are based on rounded intervals. A best choice for CI depends on the objectives of the underlying investigations and a broad prior knowledge about the underlying parameter if any.

Finally, our investigation suggests the following recommendations:

1) In the analysis of rare events where μ is expected to be very small *in between 0 to 2*, we recommend MW and FNCC method on the basis of highest coverage probabilities with shortest expected length and smallest expected P-bias and reasonable expected P-confidence. In this region LR is also recommendable on the basis of all the criteria except E(P-bias).

2) For the situations where the parameter is expected to be large *more than 4*, methods involved in $G1$ are the best choice. In fact the performance of methods in $G1$ is uniformly satisfactory (if not best) on the entire parameter space with respect to all the criteria, so in the absence of any knowledge regarding the underlying parameter, we recommend these methods for use.

3) We strongly recommend to avoid using W, BART, and H methods in all kinds of applications, since these are uniformly nonconservative for all parameter values, have large E(P-bias) and smallest E(P-confidence) and highly imbalanced noncoverage on the right and left side.

These recommendations are useful guidelines for consulting professionals, in data analysis, software development, and can be an interesting addition to the discussion of case studies in Applied Statistic courses.

ACKNOWLEDGMENTS

We are very much grateful to anonymous referees and the Associated Editor for many suggestions that greatly improved the manuscript. H.V. Kulkarni was supported by the grants received by Government of India, Department of Science and Technology, India, under the project Reference No. SR/S4/MS: 306/05.

REFERENCES

[1] ANSCOMBE, F.J. (1948). The transformation of Poisson, Binomial and Negative Binomial data, *Biometrika*, **35**, 246–254.

[2] BARKER, L. (2002). A comparison of nine confidence intervals for a Poisson parameter when the expected number of events is ≤ 5, *The American Statistician*, **56**(2), 86–89.

[3] BARTLETT, M.S. (1936). The square root transformation in the analysis of variance, *Journal of Royal Statistical Society Supplement*, **3**, 68–78.

[4] BEGAUD, B.; KARIN, M.; ABDELILAH, A.; PASCALE, T.; NICHOLAS, M. and YOLA, M. (2005). An easy to use method to approximate Poisson confidence limits, *European Journal of Epidemiology*, **20**(3), 213–216.

[5] BROWN, CAI and DASGUPTA (2003). Interval estimation in exponential families, *Statistica Sinica*, **13**, 19–49.

[6] FREEMAN, M.F. and TUKEY, J.W. (1950). Transformations related to the angular and square root, *Annals of Mathematical Statistics*, **21**, 606–611.

[7] GARWOOD, F. (1936). Fiducial limits for the Poisson distribution, *Biometrika*, **28**, 437–442.

[8] HALD, A. (1952). *Statistical Theory with Engineering Applications*, John Wiley and Sons, New York.

[9] LANCASTER, H.O. (1961). Significance tests in discrete distributions, *Journal of American Statistical Association*, **56**, 223–234.

[10] MOLENAAR, W. (1970). *Approximations to the Poisson, Binomial and Hypergeometric Distribution Functions*, Mathematical Center Tracts 31, Mathematisch Centrum, Amsterdam.

[11] SCHWERTMAN, N.C. and MARTINEZ, R.A. (1994). Approximate Poisson confidence limits, *Communication in Statistics — Theory and Methods*, **23**(5), 1507–1529.

[12] SWIFT, M.B. (2009). Comparison of confidence intervals for a Poisson Mean-Further considerations, *Communication in Statistics — Theory and Methods*, **38**, 748–759.

[13] VANDENBROUCKE, J.P. (1982). A shortcut method for calculating the 95 percent confidence interval of the standardized mortality ratio, (Letter), *American Journal of Epidemiology*, **115**, 303–304.

[14] WILSON, E.B. and HILFERTY, M.M. (1931). The distribution of chi-square, *Proceedings of the National Academy of Sciences*, **17**, 684–688.