REVSTAT – Statistical Journal
Volume 13, Number 3, November 2015, 245–262

BAYESIAN AND NON-BAYESIAN INTERVAL ESTIMATORS FOR THE POISSON MEAN

Authors: S. Nadarajah
– School of Mathematics, University of Manchester,
Manchester M13 9PL, UK
mbbsssn2@manchester.ac.uk

M. Alizadeh
– Department of Statistics, Shahid Bahonar University of Kerman
Kerman, Iran

S.F. Bagheri
– Department of Statistics, Razi University of Kermanshah
Kermanshah, Iran

Received: December 2013 Revised: May 2014 Accepted: May 2014

Abstract:

- Seventeen different confidence intervals for the Poisson mean are compared using simulations and a real data application. The interval estimators include the Wald interval estimator, the score interval estimator, the exact interval estimator, the bootstrap interval estimator, the equal tails credible interval estimator, Jeffreys prior credible interval estimator, the HPD credible interval estimator and the relative surprise credible interval estimator. Recommendations for choosing among the seventeen intervals are given for different sample sizes and different Poisson means.

Key-Words:

- *coverage length; coverage probability; Poisson distribution.*

AMS Subject Classification:

- 62E20.

1. INTRODUCTION

Poisson distribution has wide spread applications in almost every area of the sciences, engineering and medicine. Hence, it is important that accurate estimators are available for its rate parameter.

Many authors have studied estimation of the Poisson mean. A comparison of nine interval estimators for a Poisson mean when the expected number of events ≤ 5 is given in Barker [2]. An easy to use method to approximate Poisson confidence limits is discussed by Bégaud *et al.* [4]. Three interval estimators for linear functions of Poisson rates are given in Stamey and Hamilton [31]. Asymptotic interval estimators for Poisson regression are studied by Michael and Adam [24]. Swift [32] gives recommendations for choosing between twelve different confidence intervals for the Poisson mean. Bayesian interval estimation for the difference of two independent Poisson rates for under reported data is considered in Greer *et al.* [18]. Improved prediction intervals for binomial and Poisson distributions are given in Krishnamoorthy and Jie [20]. Simple approximate procedures for constructing binomial and Poisson tolerance intervals are given in Krishnamoorthy *et al.* [21]. Interval estimators for the difference between two Poisson rates are given in Li *et al.* [22]. Interval estimation for misclassification rate parameters in a complementary Poisson model is described by Riggs *et al.* [28]. Patil and Kulkarni [27] compare nineteen confidence intervals for the Poisson mean. See also Byrne and Kabaila [7] and Ng *et al.* [26].

Most studies we are aware of have compared the performances of only classical interval estimators for the Poisson mean: all of the nine estimators considered in Barker [2] are classical interval estimators; only one of the estimators considered in Swift [32] is a Bayesian credible interval estimator; all of the nineteen estimators considered in Patil and Kulkarni [27] are classical interval estimators; and so on. Also none of these papers have used a real data set to compare the performance of the estimators.

The aim of this note is a comparison study of classical interval estimators as well as Bayesian credible interval estimators for the Poisson mean. We consider equal numbers of classical interval estimators and Bayesian credible interval estimators with a range of priors considered for the latter. In total, we compare seventeen different interval estimators for the Poisson mean. Our comparison is based on simulations as well as a real data set.

The contents of this note are organized as follows. In Sections 2 and 3, several interval estimators are described for the Poisson mean. Section 2 describes the following classical interval estimators: the Wald interval estimator, the score interval estimator, the exact interval estimator, and the bootstrap interval estimator. Section 3 describes the following Bayesian credible estimators: the equal tails credible interval estimator, Jeffreys prior credible interval estimator, the HPD credible interval estimator and the relative surprise credible interval estimator. Each of the estimators in Section 3 was calculated under four different priors: uniform prior; exponential prior; gamma prior; chisquare prior. Section 4 performs a simulation study comparing the performance of the estimators of Sections 2 and 3. The performance is compared in terms of coverage probabilities and coverage lengths. A real data application is described in Section 5. Finally, some conclusions are noted in Section 6.

2. INTERVAL ESTIMATORS FOR POISSON MEAN

In this section, some methods to obtain interval estimators for the Poisson mean are described.

2.1. Approximate interval estimator

Here, we use some large sample methods for constructing interval estimators. Suppose $T\left(\underline{X}\right)$ is an estimator based on sample mean such that

$$\sqrt{n}\,\frac{T\left(\underline{X}\right)-\theta}{\sqrt{\nu(\theta)}} \stackrel{L}{\rightarrow} Z,$$

where $Z \sim N(0,1)$ and $\stackrel{L}{\rightarrow}$ means convergence in distribution (Chung [10]). Suppose further that there is a statistic $S\left(\underline{X}\right)$ so that $\nu(\theta) \stackrel{p}{\rightarrow} S\left(\underline{X}\right)$. Then, by Slutsky 's theorem,

$$\sqrt{n}\,\frac{T\left(\underline{X}\right)-\theta}{\sqrt{S\left(\underline{X}\right)}} \stackrel{L}{\rightarrow} Z\ .$$

We can obtain an approximate interval estimator for θ with confidence coefficient $1-\alpha$ by inverting the inequality (Rohatgi and Ehsanes Saleh [29]):

$$\left|\sqrt{n}\,\frac{T\left(\underline{X}\right)-\theta}{\sqrt{S\left(\underline{X}\right)}}\right| \leq z_{1-\alpha/2}\ .$$

In the following, we construct approximate interval estimators for the Poisson mean. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with mean λ. We consider two interval estimators.

a) The score interval. By using

$$Q = \frac{\overline{X} - \lambda}{\sqrt{\frac{\lambda}{n}}} \sim N(0,1) , \qquad n \to \infty ,$$

we can write

$$P\left(-z_{1-\frac{\alpha}{2}} < \frac{\overline{X} - \lambda}{\sqrt{\frac{\lambda}{n}}} < z_{1-\frac{\alpha}{2}} \right) = 1 - \alpha .$$

So, we have

$$\left| \frac{\overline{X} - \lambda}{\sqrt{\frac{\lambda}{n}}} \right| < z_{1-\frac{\alpha}{2}}$$

or

$$\lambda^2 - \lambda \left(2\overline{X} + \frac{z^2_{1-\frac{\alpha}{2}}}{n} \right) + \overline{X}^2 < 0 .$$

By solving this inequality, we see that

$$\left(\frac{\left(2\overline{X} + \frac{z^2_{1-\frac{\alpha}{2}}}{n} \right) - \sqrt{\Delta}}{2} , \frac{\left(2\overline{X} + \frac{z^2_{1-\frac{\alpha}{2}}}{n} \right) + \sqrt{\Delta}}{2} \right)$$

is an interval estimator for λ with confidence coefficient $1 - \alpha$, where

$$\Delta = \frac{z^2_{1-\frac{\alpha}{2}}}{n} \left(\frac{z^2_{1-\frac{\alpha}{2}}}{n} + 4\overline{X} \right) ,$$

see Shao [30].

b) The Wald interval. We know that $\widehat{\lambda} = \overline{X}$ is the maximum likelihood estimator for λ, so

$$Q = \frac{\overline{X} - \lambda}{\sqrt{\frac{\overline{X}}{n}}} \sim N(0,1) , \qquad n \to \infty .$$

So,

$$\left(\overline{X} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{X}}{n}} , \ \overline{X} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{X}}{n}} \right)$$

is an interval estimator for λ with confidence coefficient $1 - \alpha$. Sometimes $\overline{X} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{X}}{n}}$ can be less than zero. In this case, we use the interval estimator

$$\left(\max\left(0 , \ \overline{X} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{X}}{n}} \right) , \ \overline{X} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{X}}{n}} \right) .$$

2.2. Exact interval estimator

Let $X_1, X_2, ..., X_n$ be a random sample from a Poisson distribution with mean λ. Let $Y = \sum_{i=1}^{n} X_i$. We know that Y is a sufficient statistic for λ and $Y \sim$ Poisson$(n\lambda)$. Then, an exact interval estimator for λ with confidence coefficient $1-\alpha$ is given by

$$\left(\frac{1}{2n} \chi^2_{2y, \frac{\alpha}{2}}, \frac{1}{2n} \chi^2_{2(y+1), 1-\frac{\alpha}{2}} \right).$$

For $y = 0$, we take $\chi^2_{0, 1-\frac{\alpha}{2}} = 0$ (Casella and Berger [9]). Although an exact confidence interval estimator exists, it is still of interest to compare asymptotic and exact estimators. The readers are referred to Agresti and Coull [1]. They show that approximate approaches are better than the "exact" approach for interval estimation of the binomial distribution. Because Poisson distribution is a discrete distribution similar to the binomial distribution, it is of interest to investigate the performance of different interval estimators.

2.3. Bootstrap confidence intervals

Here, we use the percentile bootstrap method (see Davison and Hinkley [11] for details) to construct confidence intervals for λ. The percentile bootstrap method is popular: for example, Ibrahim and Kudus [19] used it to construct confidence intervals for the median of a three-parameter Weibull distribution.

The percentile bootstrap method can be applied as follows:

1. For a random sample $X_1, X_2, ..., X_n$ from a Poisson distribution with mean λ, compute the maximum likelihood estimate $\widehat{\lambda} = \overline{X}$.
2. Random select n observations from $X_1, X_2, ..., X_n$ with replacement.
3. Repeat step 2 B times to generate B bootstrap samples, say $X_1^j, X_2^j, ..., X_n^j$, $1 \leq j \leq B$.
4. Compute the maximum likelihood estimate $\widehat{\lambda}^j = \overline{X}^j$ of λ for each of the bootstrap samples in step 3.
5. Based on $\widehat{\lambda}^1, \widehat{\lambda}^2, ..., \widehat{\lambda}^B$, a $100(1-\alpha)$ percentile bootstrap confidence interval is

$$\left(2\widehat{\lambda} - \widehat{\lambda}^b, 2\widehat{\lambda} - \widehat{\lambda}^a \right),$$

where $a = (B+1)\frac{\alpha}{2}$ and $b = (B+1)\left(1 - \frac{\alpha}{2}\right)$.

3. BAYESIAN CREDIBLE INTERVALS

In this section, we discuss four Bayesian credible intervals for the Poisson mean. Bayesian credible intervals incorporate problem-specific contextual information from the prior distribution into estimates, whereas classical interval estimators are based solely on the data. In real applications, we should employ Bayesian approaches whenever strong prior information exist. This could provide good coverage and relatively narrow intervals for the parameter. We consider Bayesian credible intervals for the Poisson distribution under different priors.

3.1. Posterior distributions under different priors

The efficiency of Bayesian framework is largely dependent upon the choice of an appropriate prior distribution. The prior information is combined to the current information to update the belief regarding a particular characteristic of the data. The prior information can be of two types; informative and non-informative priors. Though, the choice of a prior depends upon the circumstances of the study, the search for a suitable prior is always of interest. We utilize both informative and non-informative priors for our posterior analysis.

Let $X_1, X_2, ..., X_n$ be a random sample from Poisson(λ). The prior and posterior distributions considered are as follows:

(**a**) For the uniform prior,

$$\pi(\lambda) \propto 1 \ , \qquad \lambda > 0 \ , \tag{3.1}$$

the posterior distribution is

$$\pi\left(\lambda|\mathbf{x}\right) = \frac{n^{\sum_{i=1}^{n} x_i + 1}}{\Gamma\left(\sum_{i=1}^{n} x_i\right)} \lambda^{\left(\sum_{i=1}^{n} x_i + 1\right) - 1} e^{-n\lambda} \ ,$$
$$\left[Gamma\left(\sum_{i=1}^{n} x_i + 1, n\right)\right] . \tag{3.2}$$

(**b**) For Jeffreys prior,

$$\pi(\lambda) \propto \lambda^{-\frac{1}{2}} \ , \qquad \lambda > 0 \ , \tag{3.3}$$

the posterior distribution is

$$\pi(\lambda|\mathbf{x}) = \frac{n^{\sum_{i=1}^n x_i + \frac{1}{2}}}{\Gamma\left(\sum_{i=1}^n x_i + \frac{1}{2}\right)} \lambda^{\left(\sum_{i=1}^n x_i + \frac{1}{2}\right)-1} e^{-n\lambda} ,$$

$$\left[Gamma\left(\sum_{i=1}^n x_i + \frac{1}{2}, n\right)\right] . \tag{3.4}$$

(**c**) For the exponential prior,

$$\pi(\lambda) = a e^{-a\lambda} , \qquad \lambda > 0 , \;\; a > 0 , \tag{3.5}$$

where a is a hyper parameter, the posterior distribution is

$$\pi(\lambda|\mathbf{x}) = \frac{(n+a)^{\sum_{i=1}^n x_i + 1}}{\Gamma\left(\sum_{i=1}^n x_i\right)} \lambda^{\left(\sum_{i=1}^n x_i + 1\right)-1} e^{-(n+a)\lambda} ,$$

$$\left[Gamma\left(\sum_{i=1}^n x_i + 1, \, n+a\right)\right] . \tag{3.6}$$

(**d**) For the gamma prior,

$$\pi(\lambda) = \frac{a^b}{\Gamma(b)} \lambda^{b-1} e^{-a\lambda} , \qquad \lambda > 0 , \;\; a > 0 , \;\; b > 0 , \tag{3.7}$$

where a and b are hyper parameters, the posterior distribution is

$$\pi(\lambda|\mathbf{x}) = \frac{(n+a)^{\left(\sum_{i=1}^n x_i + b\right)}}{\Gamma\left(\sum_{i=1}^n x_i\right)} \lambda^{\left(\sum_{i=1}^n x_i + b\right)-1} e^{-(n+a)\lambda} ,$$

$$\left[Gamma\left(\sum_{i=1}^n x_i + b, \, n+a\right)\right] . \tag{3.8}$$

(**e**) For the chisquare prior,

$$\pi(\lambda) = \frac{\lambda^{\frac{b}{2}-1} e^{-\frac{\lambda}{2}}}{\Gamma\left(\frac{b}{2}\right) 2^{\frac{b}{2}}} , \qquad \lambda > 0 , \;\; b > 0 , \tag{3.9}$$

where b is a hyper parameter, the posterior distribution is

$$\pi(\lambda|\mathbf{x}) = \frac{\left(n+\frac{1}{2}\right)^{\left(\sum_{i=1}^n x_i + \frac{b}{2}\right)}}{\Gamma\left(\sum_{i=1}^n x_i\right)} \lambda^{\left(\sum_{i=1}^n x_i + \frac{b}{2}\right)-1} e^{-\left(n+\frac{1}{2}\right)\lambda} ,$$

$$\left[Gamma\left(\sum_{i=1}^n x_i + \frac{b}{2}, \, n+\frac{1}{2}\right)\right] . \tag{3.10}$$

For more discussion, see Feroze and Aslam [15].

In the following, we find Bayesian credible intervals based on the derived posteriors.

3.2. Equal tails credible intervals

Table 1 presents the $1-\alpha$ equal tails credible intervals.

Table 1: The $1-\alpha$ equal tails credible intervals under the different priors and posteriors.

Priors	Pivotal quantity	Lower bound	Upper bound
Uniform	$2n\lambda$	$\frac{1}{2n}\,\chi^2_{2\left(\sum_{i=1}^n x_i+1\right),\,1-\frac{\alpha}{2}}$	$\frac{1}{2n}\,\chi^2_{2\left(\sum_{i=1}^n x_i+1\right),\,\frac{\alpha}{2}}$
Jeffreys	$2n\lambda$	$\frac{1}{2n}\,\chi^2_{2\left(\sum_{i=1}^n x_i+\frac{1}{2}\right),\,1-\frac{\alpha}{2}}$	$\frac{1}{2n}\,\chi^2_{2\left(\sum_{i=1}^n x_i+\frac{1}{2}\right),\,\frac{\alpha}{2}}$
Exponential	$2(n+a)\,\lambda$	$\frac{1}{2(n+a)}\,\chi^2_{2\left(\sum_{i=1}^n x_i+1\right),\,1-\frac{\alpha}{2}}$	$\frac{1}{2(n+a)}\,\chi^2_{2\left(\sum_{i=1}^n x_i+1\right),\,\frac{\alpha}{2}}$
Gamma	$2(n+a)\,\lambda$	$\frac{1}{2(n+a)}\,\chi^2_{2\left(\sum_{i=1}^n x_i+b\right),\,1-\frac{\alpha}{2}}$	$\frac{1}{2(n+a)}\,\chi^2_{2\left(\sum_{i=1}^n x_i+b\right),\,\frac{\alpha}{2}}$
Chisquare	$2\left(n+\frac{1}{2}\right)\lambda$	$\frac{1}{2\left(n+\frac{1}{2}\right)}\,\chi^2_{2\left(\sum_{i=1}^n x_i+\frac{b}{2}\right),\,1-\frac{\alpha}{2}}$	$\frac{1}{2\left(n+\frac{1}{2}\right)}\,\chi^2_{2\left(\sum_{i=1}^n x_i+\frac{b}{2}\right),\,\frac{\alpha}{2}}$

3.3. Jeffreys prior credible intervals

The non-informative Jeffreys prior plays a special role in the Bayesian analysis, see, for example, Berger [5]. In particular, Jeffreys prior is the unique first-order probability matching prior for a real-valued parameter with no nuisance parameter, see Ghosh [16]. In our setting, simple calculations show that the Fisher information about μ is $I(\mu)=n\left(\mu+b\mu^2\right)^{-1}$ and thus Jeffreys prior is proportional to

$$I^{\frac{1}{2}}(\mu) = n^{\frac{1}{2}}\left(\mu+b\mu^2\right)^{-\frac{1}{2}}.$$

Denoting the posterior distribution by J, the $(1-\alpha)$ Jeffreys credible interval for μ can be written as

$$(J_\alpha,\, J_{1-\alpha})\,, \tag{3.11}$$

where $J_{1-\alpha}$ and J_α are, respectively, the $1-\alpha$ and α quantiles of the posterior distribution based on n observations (Cai [8]). By (3.3) and (3.4), we can rewrite

(3.11) as

$$\left(G_{\alpha/2,\sum_{i=1}^{n} x_i+\frac{1}{2},n},\, G_{1-\alpha/2,\sum_{i=1}^{n} x_i+\frac{1}{2},n}\right).$$

For more discussion, see Brown *et al.* [6].

3.4. HPD credible intervals

The set $\{\theta : \pi(\theta|\underline{\mathbf{x}}) \geq k\}$ is called highest posterior density, where k is chosen so that

$$1-\alpha = \int_{\{\theta:\,\pi(\theta|\underline{\mathbf{x}})\geq k\}} \pi(\theta|\underline{\mathbf{x}})\, d\theta \ .$$

See Casella and Berger [9]. If the posterior pdf, $\pi(\theta|\underline{\mathbf{x}})$, is unimodal then the HPD set would be an interval, say $(\theta_{HL}, \theta_{HU})$ (Berger [5]). In this case, we construct HPD credible intervals for parameters of interest in the square:

$$\pi(\theta_{HL}|\underline{\mathbf{x}}) = \pi(\theta_{HU}|\underline{\mathbf{x}}) \ , \qquad \int_{\theta_{HL}}^{\theta_{HU}} \pi(\theta|\underline{\mathbf{x}})\, d\theta = 1-\alpha \ .$$

3.5. Relative surprise credible intervals

Relative surprise credible intervals for θ, as discussed in Evans [12], are based on a particular approach to assessing the null hypothesis $H_0 : \theta = \theta_0$. For this, we compute the observed relative surprise (ORS) defined by

$$\pi\left(\frac{\pi(\theta|\underline{\mathbf{x}})}{\pi(\theta)} > \frac{\pi(\theta_0|\underline{\mathbf{x}})}{\pi(\theta_0)}|\underline{\mathbf{x}}\right). \tag{3.12}$$

We see that (3.12) compares the relative increase in belief for θ, from a priori to a posteriori. Other approaches to measuring surprise are discussed in Good [17]. For estimation purposes, one may consider ORS in (3.12) as a function of θ_0 and select a value which minimizes this quantity as the estimator, called the least relative surprise estimator (LRSE). Moreover, to obtain a $1-\alpha$-credible region for θ, we simply invert (3.12) in the standard way to obtain the $(1-\alpha)$-relative surprise credible interval provided that

$$\pi\left(\frac{\pi(\theta|\underline{\mathbf{x}})}{\pi(\theta)} > \frac{\pi(\theta_0|\underline{\mathbf{x}})}{\pi(\theta_0)}\middle|\underline{\mathbf{x}}\right) \leq 1-\alpha \ .$$

It can be proved that if the posterior pdf $\pi(\theta|\underline{\mathbf{x}})$ is unimodal then the credible set is of the form $(\theta_{RL}, \theta_{RU})$ such that

$$\frac{\pi(\theta_{RL}|\underline{\mathbf{x}})}{\pi(\theta)} = \frac{\pi(\theta_{RU}|\underline{\mathbf{x}})}{\pi(\theta)} \ , \qquad \int_{\theta_{RL}}^{\theta_{RU}} \pi(\theta|\underline{\mathbf{x}})\, d\theta = 1-\alpha \ .$$

Relative surprise credible regions are shown to minimize, among Bayesian credible regions, the prior probability of covering a false value from the prior. Such regions are also shown to be unbiased in the sense that the prior probability of covering a false value is bounded above by the prior probability of covering the true value. Relative surprise credible regions are shown to maximize both the Bayes factor in favor of the region containing the true value and the relative belief ratio, among all credible regions with the same posterior content (Evans and Shakhatreh [13]).

3.6. Reparameterizations

A basic principle of inference is that inferences about a parameter of interest should be invariant under reparameterizations: for example, whatever rule we use to obtain a $(1-\alpha)$-credible region, $B_{1-\alpha}$, for a parameter of interest, θ, the rule should yield the region $\Psi\left(B_{1-\alpha}\right)$ for any one-to-one, sufficiently smooth, reparameterization $\psi = \Psi(\theta)$. Relative surprise credible inferences satisfy this principle. For greater detail, see Evans and Shakhatreh [13] and Baskurt and Evans [3].

4. COMPARISON OF CONFIDENCE INTERVALS

In this section, we compare the interval estimators of Sections 2 and 3: the Wald (WA) interval estimator, the score (SC) interval estimator, the exact (EX) interval estimator, Jeffreys (Jef) prior credible estimator, the bootstrap (Boot) interval estimator, the HPD credible interval estimator, the relative surprise (RS) credible interval estimator and the equal tails (EQ) credible interval estimator. Note that the HPD, RS and the EQ credible interval estimators depend on the chosen prior. Others do not depend on the chosen prior.

The comparison is based on coverage probabilities and coverage lengths computed by simulation. Each coverage probability and coverage length was computed over ten thousand replications of the simulated sample. Throughout, the level of significance was taken to be five percent.

The parameters of the priors can be chosen either arbitrarily or using empirical Bayes (EB) estimation. EB estimation is discussed in the Appendix. But our simulations showed that both arbitrary choice and EB estimation gave the same results. So, we choose the prior parameters arbitrarily as $a = 3$ and $b = 2$.

4.1. Comparison based on coverage probability

Here, we compare the interval estimators based on their coverage probabilities. Figures 1 to 9 in Nadarajah *et al.* [25] show how the coverage probabilities vary with respect to sample size and λ for the classical interval estimators and for the priors and posteriors given by (3.1)–(3.10). The following observations can be drawn from the figures:

- among the classical interval estimators, the WA, SC and EX estimators have the coverage probabilities acceptably close to the nominal level;
- among the classical interval estimators, the Boot estimator has the coverage probabilities unacceptably further away from the nominal level;
- among the Bayesian credible estimators with the uniform prior, the Jef, HPD and EQ estimators have the coverage probabilities acceptably close to the nominal level;
- among the Bayesian credible estimators with the uniform prior, the RS estimator has the coverage probabilities unacceptably further away from the nominal level;
- among the Bayesian credible estimators with other priors, the Jef estimator has the coverage probabilities acceptably close to the nominal level;
- among the Bayesian credible estimators with other priors, the RS, EQ and HPD estimators have the coverage probabilities unacceptably further away from the nominal level;
- the Boot estimator and the EQ credible interval estimator generally underestimate the coverage probability;
- the RS credible interval estimator generally overestimates the coverage probability;
- the HPD credible interval estimator sometimes underestimates and sometimes overestimates the coverage probability.

Although these observations are limited to the ranges of λ and n specified by Figures 1 to 9 in Nadarajah *et al.* [25], they held for other values too.

4.2. Comparison based on coverage length

Here, we compare coverage lengths of the interval estimators. Figures 10 to 18 in Nadarajah *et al.* [25] show how the coverage lengths vary with respect to sample size and λ for the classical interval estimators and for the priors and posteriors given by (3.1)–(3.10). The following observations can be drawn from the figures:

- the coverage lengths for each estimator generally increase with increasing λ;
- the coverage lengths generally decrease with increasing n except for the HPD and RS credible interval estimators;
- the coverage lengths for the HPD credible interval estimator sometime increase with n and sometimes decrease with n;
- also the coverage lengths for the RS credible interval estimator sometime increase with n and sometimes decrease with n;
- the coverage lengths appear largest for the HPD and RS credible interval estimators;
- the coverage lengths appear smallest for the WA, SC, EX, Jef, Boot and EQ estimators;
- among the Bayesian credible estimators, the coverage lengths appear largest for those with the exponential prior.

Although these observations are limited to the ranges of λ and n specified by Figures 10 to 18 in Nadarajah *et al.* [25], they held for other values too.

5. REAL DATA APPLICATIONS

Here, we present an analysis of the "Flying-bomb Hits in London During World War II" data reported by Feller [14]. The city was divided into five hundred and seventy six small areas of one-quarter square kilometers each, and the number of areas hit exactly k times was counted. There were a total of five hundred and thirty seven hits, so the average number of hits per area was 0.93. The observed frequencies in Table 2 are remarkably close to a Poisson distribution as we shall show now.

Table 2: Flying-bomb hits in London during World War II.

Hits	0	1	2	3	4	5^+
Observed	229	211	93	35	7	1

We fitted the Poisson, negative binomial and geometric distributions to the data in Table 2. The smallest chisquared statistic, the smallest Akaike information criterion and the smallest Bayesian information criterion were obtained for the Poisson distribution. The quantile–quantile plots for the three fits shown in Figure 19 in Nadarajah *et al.* [25] show that the Poisson distribution has the points closest to the straight line.

For the fit of the Poisson distribution, $\widehat{\lambda} = 0.9288194$ with standard error 0.04015632. Using these estimates, the confidence intervals of Sections 2 and 3 can be computed. They are shown in Table 3.

Table 3: Bayesian and non-Bayesian confidence intervals for the mean number of hits.

Intervals	Lower bound	Upper bound	Upper − Lower
WA	0.85338	1.01093	0.15755
SC	0.85011	1.00752	0.15741
EX	0.85343	1.00916	0.15573
Jeffreys	0.8526	1.01006	0.15746
Bootstrap	0.84375	1.01215	0.1684
HPD.u	0.46441	1.39323	0.92882
RS.u	0.46441	1.39323	0.92882
EQ.u	0.85343	1.01097	0.15754
HPD.e	0.46441	1.39323	0.92882
RS.e	0.46441	1.39323	0.92882
EQ.e	0.85048	1.00747	0.15699
HPD.g	0.46441	1.39323	0.92882
RS.g	0.46441	1.39323	0.92882
EQ.g	0.85379	1.01107	0.15728
HPD.c	0.46441	1.39323	0.92882
RS.c	0.46441	1.39323	0.92882
EQ.c	0.85352	1.01099	0.15747

We see that the coverage length is smallest for the EX estimator, second smallest for the EQ credible interval estimator, third smallest for the SC estimator, fourth smallest for the Jef credible interval estimator, fifth smallest for the WA estimator, sixth smallest for the bootstrap estimator and the largest for the HPD and RS credible interval estimators. These observations are consistent with the results in Section 4.2.

6. CONCLUDING REMARKS

The estimation of Poisson mean is of great importance because of wide spread applications of the Poisson distribution. We have compared seventeen different interval estimators for the Poisson mean. They were compared in terms of coverage probabilities and coverage lengths computed using simulations and a real data application. We have given various recommendations for choosing among the seventeen interval estimators. Some of them are: WA, SC and EX estimators are the best classical interval estimators in terms of coverage probabilities; Jef estimator is the best Bayesian credible interval estimator in terms of coverage probabilities; WA, SC, EX, Boot estimators are the best classical interval estimators in terms of coverage lengths; Jef and EQ estimators are the best Bayesian credible interval estimators in terms of coverage lengths.

ACKNOWLEDGMENTS

This research was partially supported by a research project grant from City University of Hong Kong (Project No. 9380058). An earlier version of this research was presented at APARM2012 in Nanjing, China. The authors would like to thank the Editor and the referee for careful reading and for their comments which greatly improved the paper.

APPENDIX: EMPIRICAL BAYES ESTIMATION

When the prior parameters are unknown, we may use another type of Bayesian estimation for estimating them without knowing or assessing the prior (subjective) distribution. The parameters of the (subjective) prior are estimated from the data. This method of estimation is called empirical Bayes (EB) estimation. For more details on EB estimation, see Maritz and Lwin [23].

In the following, we apply the EB method, in order to obtain an estimator for θ based on observed data. Suppose $X_1, X_2, ..., X_n \sim P(\lambda)$ is the observed data. Let $\lambda \sim \text{Gamma}(a, b)$ denote the prior distribution. Then, $\widehat{\lambda} = \overline{X}$. Using $S = n\widehat{\lambda} \sim P(n\lambda)$, we can write

$$f(s|\lambda) = \frac{e^{-n\lambda}(n\lambda)^s}{s!} \tag{6.1}$$

for $s = 0, 1, ...$ By using (6.1), we have

$$\begin{aligned} f(s) &= \int_0^\infty f(s|\lambda)\, g(\lambda)\, d\lambda \\ &= \int_0^\infty \frac{e^{-n\lambda}(n\lambda)^s}{s!} \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda}\, d\lambda \\ &= \frac{(s+a-1)!}{s!(a-1)!} \left(\frac{b}{b+n}\right)^a \left(\frac{n}{b+n}\right)^s, \end{aligned}$$

the probability mass function of a negative binomial random variable with parameters $p = \frac{b}{b+n}$ and $r = a$. The expectation and variance of a negative binomial random variable with parameters p and r are $\frac{rq}{p}$ and $\frac{rq}{p^2}$, respectively. Using these, the EB estimators of the prior parameters can be obtained as

- $\widehat{a} = \frac{\mu^2}{\sigma^2-\mu}$ and $\widehat{b} = \frac{n\mu}{\sigma^2-\mu}$ for the gamma prior;
- $\widehat{b} = \frac{n}{\mu}$ for the exponential prior;
- $\widehat{a} = \frac{\mu}{n}$ for the chisquare prior.

REFERENCES

[1] Agresti, A. and Coull, B.A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions, *The American Statistician*, **52**, 119–126.

[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**, 85–89.

[3] Baskurt, Z. and Evans, M. (2012). Hypothesis assessment and inequalities for Bayes factors and relative belief ratios, *Technical Report Number 1105*, 1 September 2011.

[4] Bégaud, B.; Martin, K.; Abouelfath, A.; Tubert-Bitter, P.; Moore, N. and Moride, Y. (2005). An easy to use method to approximate Poisson confidence limits, *European Journal of Epidemiology*, **20**, 213–216.

[5] Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, second edition, Springer Verlag, New York.

[6] Brown, L.D.; Cai, T.T. and Das Gupta, A. (2002). Confidence interval for a binomial proportion and asymptotic expansions, *The Annals of Statistics*, **30**, 160–201.

[7] Byrne, J. and Kabaila, P. (2005). Comparison of Poisson confidence intervals, *Communication in Statistics-Theory and Methods*, **34**, 545–556.

[8] Cai, T.T. (2005). One-sided confidence intervals in discrete distributions, *Journal of Statistical Planning and Inference*, **131**, 63–88.

[9] Casella, G. and Berger, R.L. (2002). *Statistical Inference*, Duxbury, United Press, Belmont, California.

[10] Chung, K.L. (2001). *A Course in Probability Theory*, third edition, Academic Press, New York.

[11] Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application*, Cambridge University Press, New York.

[12] Evans, M. (1997). Bayesian inference procedures derived via the concept of relative surprise, *Communications in Statistics-Theory and Methods*, **26**, 1125–1143.

[13] Evans, M. and Shakhatreh, M. (2008). Optimal properties of some Bayesian inferences, *Electronic Journal of Statistics*, **2**, 1268–1280.

[14] Feller, W. (1957). *An Introduction to Probability Theory and Its Applications*, second edition, John Wiley and Sons, New York.

[15] Feroze, N. and Aslam, M. (2012). Bayesian analysis of exponentiated gamma distribution under type II censored samples, *International Journal of Advanced Science and Technology*, **49**, 37–46.

[16] Ghosh, J.K. (1994). *Higher Order Asymptotics*, NSF-CBMS Regional Conference Series, Institute of Mathematical Statistics, Hayward, California.

[17] Good, I.J. (1988). Surprise index. In: "Encyclopaedia of Statistical Sciences" (S. Kotz, N.L. Johnson and C.B. Reid, Eds.), volume 7, John Wiley and Sons, New York.

[18] Greer, B.A.; Stamey, J.D. and Young, D.M. (2011). Bayesian interval estimation for the difference of two independent Poisson rates using data subject to under-reporting, *Statistica Neerlandica*, **65**, 259–274.

[19] Ibrahim, N.A. and Kudus, A. (2007). Bootstrap confidence interval for the median failure time of three-parameter Weibull distribution, *World Congress on Engineering*, **7**, 836–839.

[20] Krishnamoorthy, K. and Jie, P. (2011). Improved closed-form prediction intervals for binomial and Poisson distributions, *Journal of Statistical Planning and Inference*, **141**, 1709–1718.

[21] Krishnamoorthy, K.; Yanping, X. and Fang, X. (2011). A simple approximate procedure for constructing binomial and Poisson tolerance intervals, *Communication in Statistics-Theory and Methods*, **40**, 2243–2258.

[22] Li, H.Q.; Tang, M.L.; Poon, W.Y. and Tang, N.S. (2011). Confidence intervals for difference between two Poisson rates, *Communication in Statistics-Simulation and Computation*, **40**, 1478–1493.

[23] Maritz, J.L. and Lwin, T. (1989). *Empirical Bayes Methods*, second edition, Chapman and Hall, London.

[24] Michael, K. and Adam, K. (2007). Asymptotic confidence intervals for Poisson regression, *Journal of Multivariate Analysis*, **98**, 1072–1094.

[25] Nadarajah, S.; Alizadeh, M. and Bagheri, S.F. (2013). Bayesian and non-Bayesian interval estimators for the Poisson mean, *Technical Report*, School of Mathematics, University of Manchester, UK.

[26] Ng, H.K.T.; Filardo, G. and Zheng, G. (2008). Confidence interval estimating procedures for standardized incidence rates, *Computational Statistics and Data Analysis*, **52**, 3501–3516.

[27] Patil, V.V. and Kulkarni, H.V. (2012). Comparison of confidence intervals for the Poisson mean: Some new aspects, *REVSTAT*, **10**, 211–227.

[28] Riggs, K.; Young, D. and Stamey, J. (2011). Interval estimation for misclassification rate parameters in a complementary Poisson model, *Journal of Statistical Computation and Simulation*, **81**, 1145–1156.

[29] Rohatgi, V.K. and Ehsanes Saleh, A.K. (2001). *An Introduction to Probability and Statistics*, second edition, John Wiley and Sons, New York.

[30] Shao, J. (2003). *Mathematical Statistics*, second edition, Springer Verlag, New York.

[31] Stamey, J. and Hamilton, C. (2006). A note on confidence intervals for a linear function of Poisson rates, *Communications in Statistics-Simulation and Computation*, **35**, 849–856.

[32] Swift, M.B. (2009). Comparison of confidence intervals for a Poisson mean-further considerations, *Communication in Statistics-Theory and Methods*, **38**, 748–759.