Research Article

Parameter Estimation of Poisson Distribution by Using Maximum Likelihood, Markov Chain Monte Carlo, and Bayes method

Autcha Araveeporn *
Department of Statistics, Faculty of Science,
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

Abstract

The objective of this research is to test a hypothesis that the means of Poisson parameter estimations obtained from Maximum Likelihood, Markov Chain Monte Carlo, and Bayes method were not different from the true parameters. Data was simulated from a Poisson distribution with the true λ parameter set at 0.5, 2, 5, 10, and 20 and the sample size at 5, 10, 30,50, 100, and 200. The results are as follows: the Maximum Likelihood method produced means of parameter estimations that were not perceivably different from the true parameters in all cases. On the other hand, the Markov Chain Monte Carlo and the Bayes methods produced dissimilar estimations to the true λ parameters when the sample sizes and the true λ parameters were small. Additionally, the maximum likelihood method produced minimum mean square errors when the sample sizes and the true parameters were small while the Markov Chain Monte Carlo and Bayes method did so when the sample sizes and the true parameters were large.

Keywords: Bayes Method; Maximum Likelihood Method; Markov Chain Monte Carlo Method; Poisson Distribution.

1. Introduction

Today, tremendous amount of new data is generated at an increasing rate. It is usually not possible to collect and measure all units of data in a population because of restrictions in time, budget, and labor. Any characteristics of a population are parameters of population. Since it is not possible to acquire and study all units of data in a population, we only collect and study a small part, a sample, of the population instead. We use a sample to estimate the true parameter of the population. We call it an estimator. In inferential statistics, a realized value of an estimator is used to describe the population.

Point estimation is a part of inferential statistics that uses a sample to get an estimator for interpreting the population. Some properties of a good point estimator are unbiasedness, sufficiency, completeness, and minimum variance unbiased estimator. There are several point estimation methods such as the moments method, the maximum likelihood method, the minimum chi-square method, the least squares method, and the Bayes method. We were interested in the maximum likelihood method because its estimator is a class of a minimum variance unbiased estimator. We were also interested in the Bayes method because it uses both a prior probability distribution and a posterior

***Correspondence** : kaautcha@hotmail.com

probability distribution to find an estimator. However, it is fairly difficult to demonstrate a posterior distribution from a probability distribution and a prior distribution. Thus, we also considered the Markov Chain Monte Carlo method [1]. It can overcome this particular problem because it uses only the posterior distribution to make statements about parameter estimation.

In this case, we were interested the experimental outcomes that occur randomly for the counts of events within intervals of time and space. The Poisson distribution is a discrete distribution that observe the counts of event in a given interval of time. The parameter of the Poisson distribution is the mean number of events per an interval of time. We set out to find Poisson parameter estimators by using the Maximum Likelihood (ML) method, the Markov Chain Monte Carlo (MCMC) method, and the Bayes method. The data we used were simulated from a Poisson distribution with varying true parameters and sample sizes.

We organize this paper as follows: Section 2 describes the method of parameter estimation. Section 3 describes the simulation study. Section 4 shows the test statistic and the criterion for data analysis. The results are discussed in Section 5. The conclusion is in Section 6.

2. Methods for Parameter Estimation

The parameter estimation of the Poisson distribution consist of the following three methods.

2.1 Maximum Likelihood (ML) Method

The ML method is the most popular techniques for deriving an estimator because it is simple to understand and to calculate the estimators. The solution of this method is established from the maximum of the likelihood function.

Let $X_1,...,X_n$ be independent and identically distributed (iid) random variables following a Poisson distribution with parameter λ, and $f(x_i \mid \lambda)$ denote the probability density function of $x_i \mid \lambda$. Then,

$$f(x_i \mid \lambda) = \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} \quad , \quad x_i = 0,1,2,... \quad .$$

Hence, the likelihood function is

$$L(\lambda) = \prod_{i=1}^{n} f\left(x_i \mid \lambda\right).$$

The ML estimator of the parameter λ is solved as follows:

$$L(\lambda) = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!},$$

$$\ln L(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i \ln\lambda - \ln\left(\prod_{i=1}^{n} x_i!\right),$$

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = -n + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0,$$

$$\lambda = \frac{\sum_{i=1}^{n} x_i}{n}, \text{ and}$$

$$\hat{\lambda} = \bar{x}.$$

The second derivative of $\ln L(\lambda)$ is

$$\frac{\partial^2 \ln L(\lambda)}{\partial \lambda^2} = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2} < 0.$$

Therefore, the ML estimator is $\hat{\lambda}_{MLE} = \bar{x}$.

2.2 Bayes Method

Let $X_1,...,X_n$ be iid random variables following a Poisson distribution with parameter λ, and $f(x_i \mid \lambda)$ denote the probability density function of $x_i \mid \lambda$. The likelihood function can be written as

$$L(\lambda) = \prod_{i=1}^{n} f(x_i \mid \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}.$$

The prior distribution of λ is a gamma distribution with parameters a,b defined as Gamma(a,b) or rewritten as

$$g(\lambda \mid a,b) = \frac{\lambda^{a-1}e^{-\frac{\lambda}{b}}}{\Gamma(a)b^a} \quad , \quad \lambda > 0.$$

The Poisson and gamma distributions are the conjugate distribution because the Poisson distribution lists the parameter λ which is similar to the format of a gamma distribution.

The posterior distribution of λ given x_i is

$$h(\lambda \mid x_i) = \frac{f(x_i \mid \lambda)g(\lambda \mid a,b)}{\int f(x_i \mid \lambda)g(\lambda \mid a,b)\,d\lambda}$$

$$= \frac{\dfrac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}\dfrac{\lambda^{a-1}e^{-\frac{\lambda}{b}}}{\Gamma(a)b^a}}{\int_0^\infty \dfrac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}\dfrac{\lambda^{a-1}e^{-\frac{\lambda}{b}}}{\Gamma(a)b^a}\,d\lambda}$$

$$= \frac{e^{-(n+\frac{1}{b})\lambda}\lambda^{n\bar{x}+a-1}}{\int_0^\infty e^{-(n+\frac{1}{b})\lambda}\lambda^{n\bar{x}+a-1}\,d\lambda}.$$

The term $\int_0^\infty e^{-(n+\frac{1}{b})\lambda}\lambda^{n\bar{x}+a-1}\,d\lambda$ is in the form of a gamma function and can be rewritten as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$$

$$= (\alpha-1)\Gamma(\alpha-1)$$

since

$$\int_0^\infty e^{-(n+\frac{1}{b})\lambda}\lambda^{n\bar{x}+a-1}\,d\lambda = \frac{1}{(n+\frac{1}{b})^{n\bar{x}+a-2}}\int_0^\infty e^{-(n+\frac{1}{b})\lambda}[(n+\frac{1}{b})\lambda]^{(n\bar{x}+a-1)}\,d(n+\frac{1}{b})\lambda$$

$$= \frac{\Gamma(n\bar{x}+a)}{(n+\frac{1}{b})^{n\bar{x}+a-2}}.$$

Hence, the posterior distribution is written as

$$h(\lambda \mid x_i) = \frac{e^{-(n+\frac{1}{b})\lambda}\lambda^{n\bar{x}+a-1}}{\dfrac{\Gamma(n\bar{x}+a)}{(n+\frac{1}{b})^{n\bar{x}+a-2}}}.$$

The posterior mean of λ is obtained through the following sequence of derivation:

$$E(\lambda \mid x_i) = \frac{\int_0^\infty \lambda\, e^{-(n+\frac{1}{b})\lambda}\lambda^{n\bar{x}+a-1}\,d\lambda}{\dfrac{\Gamma(n\bar{x}+a)}{(n+\frac{1}{b})^{n\bar{x}+a-2}}}$$

$$= \frac{\int_0^\infty e^{-(n+\frac{1}{b})\lambda}\lambda^{n\bar{x}+a}\,d\lambda}{\dfrac{\Gamma(n\bar{x}+a)}{(n+\frac{1}{b})^{n\bar{x}+a-2}}}$$

$$= \frac{\dfrac{\Gamma(n\bar{x}+a+1)}{(n+1/b)^{n\bar{x}+a-1}}}{\dfrac{\Gamma(n\bar{x}+a)}{(n+1/b)^{n\bar{x}+a-2}}}$$

$$= \frac{(n\bar{x}+a)\Gamma(n\bar{x}+a)\,(n+1/b)^{n\bar{x}+a-2}}{\Gamma(n\bar{x}+a)(n+1/b)^{n\bar{x}+a-1}}$$

$$= \frac{(n\,\bar{x}+a)b}{nb+1}.$$

The Bayes estimator is

$$\hat{\lambda}_{Bayes} = \frac{(n\,\bar{x}+a)b}{nb+1}.$$

The quality of the estimated λ is measured through a loss function. For example, the squared error loss function is

$$L(\lambda,\hat{\lambda}(x)) \;=\; \left[\lambda-\hat{\lambda}(x)\right]^{T}\left[\lambda-\hat{\lambda}(x)\right].$$

The posterior Bayes risk is

$$\int L(\lambda,\hat{\lambda}(x))\, f(\lambda\,|\,x)\, d\lambda$$

$$= \int\left[\lambda-\hat{\lambda}(x)\right]^{T}\left[\lambda-\hat{\lambda}(x)\right]f(\lambda\,|\,x)d\lambda.$$

The Posterior Bayes risk estimator is written as

$$\frac{\partial}{\partial\hat{\lambda}}\int L(\lambda,\hat{\lambda}(x))f(\lambda\,|\,x)d\lambda$$

$$= \frac{\partial}{\partial\hat{\lambda}}\int\left[\lambda-\hat{\lambda}(x)\right]^{T}\left[\lambda-\hat{\lambda}(x)\right]f(\lambda\,|\,x)d\lambda,$$

$$= \int\frac{\partial}{\partial\hat{\lambda}}\left[\lambda-\hat{\lambda}(x)\right]^{T}\left[\lambda-\hat{\lambda}(x)\right]f(\lambda\,|\,x)d\lambda,$$

$$= -2\int\left[\lambda-\hat{\lambda}(x)\right]f(\lambda\,|\,x)d\lambda \;=\; 0,$$

$$\Rightarrow \int \lambda f(\lambda\,|\,x)d\lambda \;=\; \int \hat{\lambda}(x)f(\lambda\,|\,x)d\lambda,$$

$$\Rightarrow \hat{\lambda}_{bayes} \;=\; \int \lambda(x)f(\lambda\,|\,x)d\lambda,$$

$$\Rightarrow \hat{\lambda}_{bayes} \;=\; E(\lambda\,|\,x).$$

The posterior Bayes risk estimator becomes the Bayes estimator.

2.3 Markov Chain Monte Carlo Method

Bayesian analysis treats all parameters as random, assigns prior distributions to characterize knowledge about parameter values, and uses the posterior distribution given the observed data as the basis of inference. Often the posterior distribution is a complicated models with many parameters, so statisticians have developed simulated methods to generate samples from the posterior distribution, namely the Markov Chain Monte Carlo (MCMC) method. Gibbs Sampling ([2]-[4]) is a popular MCMC method that generates values which are always moving to new values, and most importantly, does not require a specification of proposed distributions.

We carried out the Gibbs Sampling by means of a software package known as WinBUGS (Bayesian Inference Using Gibbs Sampling) introduced by Spiegelhalter *et al.*,[5]. We use the MCMC samples of the parameter obtained via WinBUGS to compute approximate posterior summaries as the posterior distribution.

Let $X_1,\ldots,X_n$ be iid random variables following a Poisson distribution with parameter λ, and let λ be a random variable of gamma distribution with parameters a and b. The estimated parameters are λ, a, and b.

The algorithm of Gibbs sampling from Markov Chain Monte Carlo [7] proceeds as follows.

1. Set initial values $a^{(t)}$ from an exponential distribution with parameters 1 and $b^{(t)}$ from the gamma distribution with parameter $(0.1,1)$.
   Notice that a,b are the parameters of the gamma distribution and that the values of a,b are greater than zero, which is supported the exponential and the gamma distribution.
2. For $t=1,2,\ldots,T$ update $a^{(t)}$ and $b^{(t)}$.
3. Generate $\lambda^{(t)}$ from the posterior distribution function based on the gamma distribution with the parameters $a^{(t)}$ and $b^{(t)}$ following 1.
4. Plot the density of the posterior distribution function.
5. Calculate the mean, the median, and the standard deviation from the posterior distribution function.
   For each chain, the first 2000 iterations were discarded and the last 5000

iterations were used to obtain the posterior distribution of the parameters.

Thus, the MCMC estimator is

$$\hat{\lambda}_{MCMC} = \frac{1}{T}\sum_{t=1}^{T}\lambda^{(t)}.$$

Moreover, the MCMC method is obtaining a and b and approximating as follows:

$$\hat{a}_{MCMC} = \frac{1}{T}\sum_{t=1}^{T}a^{(t)},\ \text{and}\ \hat{b}_{MCMC} = \frac{1}{T}\sum_{t=1}^{T}b^{(t)}.$$

The Bayes estimator is used $\hat{a}_{MCMC}$ and $\hat{b}_{MCMC}$ to compute the parameters λ as follows:

$$\hat{\lambda}_{Bayes} = \frac{(n\bar{x} + \hat{a}_{MCMC})\hat{b}_{MCMC}}{n\,\hat{b}_{MCMC} + 1}.$$

3. Research Scope

This section discusses a simulation study for investigating the performance of the ML method, the Markov Chain Monte Carlo method, and the Bayes method. In this case, the estimators are

$$\hat{\lambda}_{MLE} = \bar{x},$$

$$\hat{\lambda}_{MCMC} = \frac{1}{T}\sum_{t=1}^{T}\lambda^{(t)},$$

$$\hat{\lambda}_{Bayes} = \frac{(n\bar{x} + \hat{a}_{MCMC})\hat{b}_{MCMC}}{n\,\hat{b}_{MCMC} + 1}.$$

3.1 The procedure for the simulation study

In order to simulate the random variable X_i's that follow Poisson distributions with the true parameter λ of 0.5, 2, 5, 10, and 20, we proceed as follows.

3.1.1 Prior distribution is defined as the gamma distribution with parameters a and b.

3.1.2 The sample sizes are considered at n = 5, 10, 30, 50, 100, and 200.

3.1.3 The data is generated 500 times in each case with the R program [7].

3.2 A Test Statistic

A t statistic is used to whether the means of a parameters is different from the true values of the parameter. In this case, the hypotheses are

$H_0 : \mu_{\hat{\lambda}} = \lambda$ and $H_1 : \mu_{\hat{\lambda}} \neq \lambda$.

The t statistic is computed as follows:

$$t = \frac{\bar{\hat{\lambda}} - \lambda}{s_{\hat{\lambda}} / \sqrt{n}},$$

where $$s_{\hat{\lambda}} = \sqrt{\frac{\sum_{j=1}^{500}(\hat{\lambda}_j - \bar{\hat{\lambda}})^2}{n-1}},$$

and df = n-1.

For the level of significance at α, we will reject H_0 if $|t| > t_{\alpha/2,n-1}$.

Hence, a lower and upper bounds of (1-α)100% confidence interval are computed by $\lambda = \bar{\hat{\lambda}} \pm t_{\alpha/2,n-1}\frac{s_{\hat{\lambda}}}{\sqrt{n}}$.

3.3 The Criterion for data analysis

The Mean Square Error (MSE) is the criterion to indicate the performance of parameter estimation from these methods and is computed by

$$MSE = \frac{\sum_{j=1}^{500}(\lambda - \hat{\lambda}_j)^2}{500}.$$

4. The results

In this section, the parameter estimation with Poisson distribution is appeared from generated data with the true parameters and the sample sizes in previous section. Tables 1-3 showed the results in the form of tables and histograms. The first and the second columns of these tables represented the sample sizes and the true parameters from simulated data. A mean, a standard deviation, a lower and upper bounds of 95% confidence interval were given in the next four columns. The last two columns of these tables listed the t statistics and p-values for hypothesis testing.

By observing the p-values, the results appear as follows:

1. ML method

From Table 1, the ML method indicated that the means of the estimated parameters were not different from the true parameters in all cases. The histograms of the estimated parameters of all true parameters were presented in Figures 1-5. The histograms follow a normal distribution for large sample sizes.

2. MCMC Method

The p-values of the MCMC estimators from Table 2 indicated that the means of the estimated parameters were not different from the true parameters when $\lambda = 0.5$ at n = 5, 10, 30 ,and 50 and when $\lambda = 2$ at n = 10. The histograms of the estimated parameters of all true parameters were presented in Figures 6-10. Similar to the results from the ML method, the histograms follow a normal distribution.

3. Bayes Method

The p-values of the Bayes estimators from Table 3 show that the means of the estimated parameters were not different from the true parameters when $\lambda = 0.5$ at n = 5, 10 , and 30 and when $\lambda = 2$ at n = 5 and 10. The histograms of the estimated parameters of all of the true parameters were presented in Figures 11-15. Similar to the results from the previous two methods, the histograms follow a normal distribution.

Table 1. The mean, standard deviation (S.D.), lower confidence interval (LCI), upper confidence interval (UCI), t statistics (t), and p-values obtained via the ML method.

Sample sizes	λ	Mean	S.D.	LCI	UCI	t	p-values
n = 5	0.5	0.4924	0.3073	0.4653	0.5194	-0.5530	0.5805
	2	2.0228	0.6809	1.9629	2.0826	0.7487	0.4544
	5	5.0344	1.0808	4.9394	5.1293	0.7117	0.4770
	10	9.9056	1.4126	9.7814	10.0297	-1.4943	0.1357
	20	20.0956	1.8971	19.9289	20.2623	1.1268	0.2604
n= 10	0.5	0.5204	0.2337	0.4998	0.5409	1.9512	0.0515
	2	2.0186	0.4478	1.9792	2.0579	0.9287	0.3535
	5	5.0408	0.6998	4.9793	5.1022	1.3036	0.1930
	10	9.9790	1.0070	9.8905	10.0674	-0.4663	0.6412
	20	20.0206	1.4705	19.8913	20.1498	0.3132	0.7542
n= 30	0.5	0.5086	0.1355	0.4966	0.5205	1.4185	0.1567
	2	2.0006	0.2671	1.9771	2.0241	0.0558	0.9555
	5	4.9948	0.4209	4.9578	5.0318	-0.2727	0.7852
	10	10.0226	0.5786	9.9717	10.0734	0.8733	0.3829
	20	20.0502	0.7891	19.9809	20.1196	1.4243	0.1550
n= 50	0.5	0.5011	0.0984	0.4925	0.5098	0.2634	0.7924
	2	1.9960	0.2098	1.9776	2.0145	-0.4176	0.6764
	5	4.9969	0.3270	4.9682	5.0256	-0.2079	0.8354
	10	10.0187	0.4348	9.9802	10.0572	0.9508	0.3391
	20	20.0075	0.5807	19.9564	20.0585	0.2895	0.7723
n= 100	0.5	0.5016	0.0733	0.4952	0.5081	0.5063	0.6129
	2	2.0085	0.1371	1.9965	2.0206	1.3986	0.1626
	5	5.0163	0.2139	4.9975	5.0351	1.7090	0.0879
	10	10.0102	0.3063	9.9832	10.0371	0.7444	0.4571
	20	20.0006	0.4478	19.9612	20.0399	0.0300	0.9761
n = 200	0.5	0.5006	0.0497	0.4962	0.5050	0.2919	0.7705
	2	2.0042	0.0981	1.9956	2.0128	0.9709	0.3321
	5	5.0057	0.1569	4.9919	5.0195	0.8176	0.4140
	10	10.0140	0.2270	9.9904	10.0303	1.0270	0.3049
	20	19.9739	0.3169	19.9460	20.0017	-1.8404	0.0663

* indicates significance level at 5%

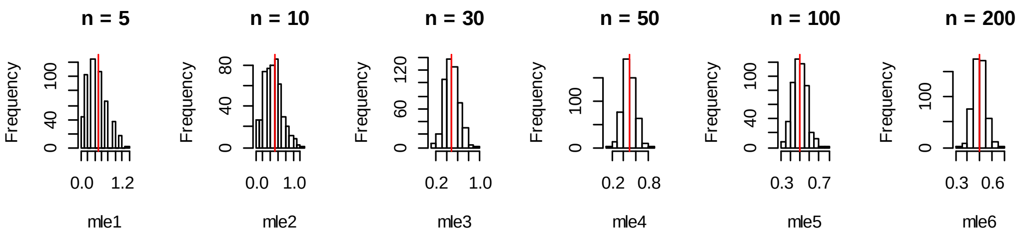

Fig.1. Histograms of estimated parameters λ with ML method when $\lambda = 0.5$.

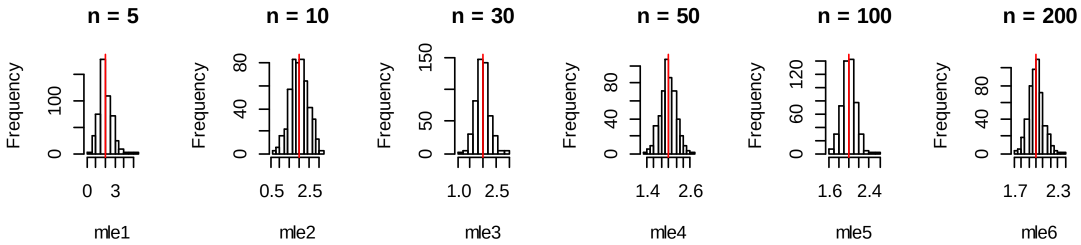

Fig.2. Histograms of estimated parameters λ with ML method when $\lambda = 2$.

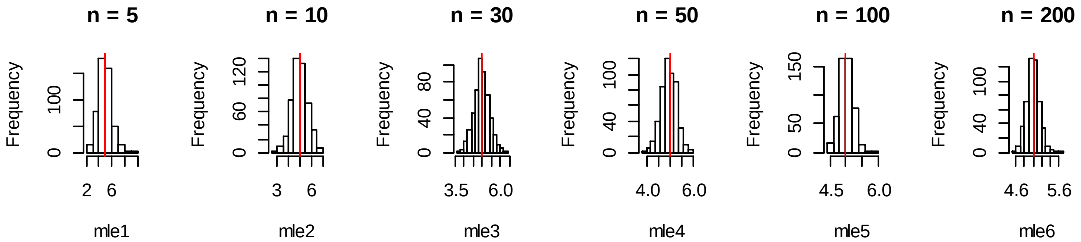

Fig.3. Histograms of estimated parameters λ with ML method when $\lambda = 5$.

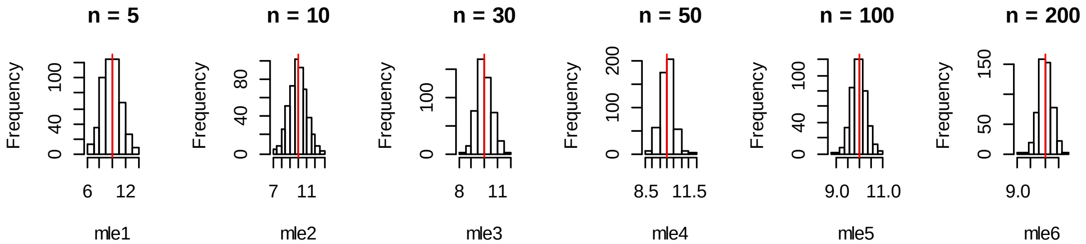

Fig.4. Histograms of estimated parameters λ with ML method when $\lambda = 10$.

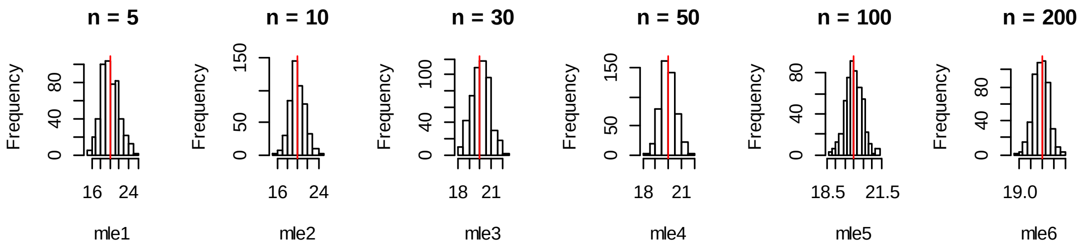

Fig.5. Histograms of estimated parameters λ with ML method when $\lambda = 20$.

Table 2. The mean, standard deviation (S.D.), lower confidence interval (LCI), upper confidence interval (UCI), t statistic (t), and p-values obtained via the MCMC method.

Sample sizes	λ	Mean	S.D.	LCI	UCI	t	p-values
n = 5	0.5	0.5567	0.3084	0.5295	0.5838	4.1100	0.0000*
	2	2.0567	0.6690	1.9980	2.1155	1.8981	0.0582
	5	5.0431	1.0791	4.9483	5.1379	0.8942	0.3717
	10	9.9058	1.4062	9.7822	10.0293	-1.4978	0.1348
	20	20.0675	1.9119	19.8995	20.2355	0.7896	0.4301
n= 10	0.5	0.5573	0.2327	0.5368	0.5778	5.5088	0.0000*
	2	2.0399	0.4470	2.0006	2.0792	1.9984	0.0462*
	5	5.0444	0.6975	4.9831	5.1057	1.4246	0.1549
	10	9.9756	1.0131	9.8866	10.0646	-0.5372	0.5924
	20	20.0145	1.4717	19.8851	20.1438	0.2202	0.8258
n= 30	0.5	0.5218	0.1361	0.5099	0.5338	3.5921	0.0003*
	2	2.0048	0.2671	1.9813	2.0283	0.4071	0.6841
	5	4.9987	0.4208	4.9618	5.0357	-0.0640	0.9490
	10	10.0242	0.5785	9.9733	10.0750	0.9356	0.3500
	20	20.0590	0.7912	19.9895	20.1285	1.6686	0.0950
n= 50	0.5	0.5094	0.0984	0.5007	0.5180	2.1387	0.0329*
	2	1.9956	0.2095	1.9772	2.0140	-0.4629	0.6436
	5	4.9997	0.3268	4.9710	5.0284	-0.0193	0.9846
	10	10.0177	0.4390	9.9791	10.0562	0.9018	0.3674
	20	20.0145	0.5827	19.9633	20.0657	0.5584	0.5768
n= 100	0.5	0.5053	0.0736	0.4988	0.5118	1.6620	0.1054
	2	2.0082	0.1367	1.9962	2.0202	1.3507	0.1774
	5	5.0178	0.2130	4.9990	5.0365	1.8702	0.0620
	10	10.0083	0.3062	9.9813	10.0352	0.6066	0.5444
	20	19.9993	0.4491	19.9598	20.0388	-0.0323	0.9742
n = 200	0.5	0.5019	0.0493	0.4976	0.5062	0.8785	0.3801
	2	2.0052	0.0987	1.9966	2.0139	1.1959	0.2323
	5	5.0042	0.1564	4.9905	5.0180	0.6096	0.5424
	10	10.0089	0.2266	9.9889	10.0288	0.8776	0.3806
	20	19.9729	0.3168	19.9450	20.0007	-1.9103	0.0560

* indicates significance level at 5%

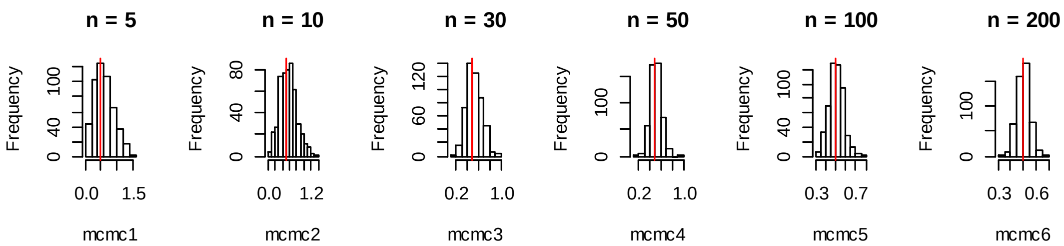

Fig.6. Histograms of estimated parameters λ with MCMC method when $\lambda = 0.5$.

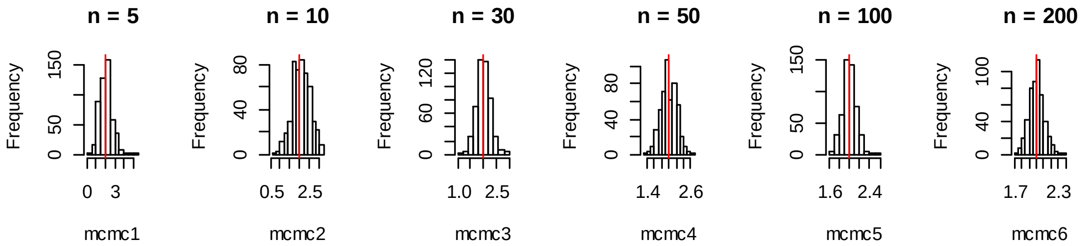

Fig.7. Histograms of estimated parameters λ with MCMC method when $\lambda = 2$.

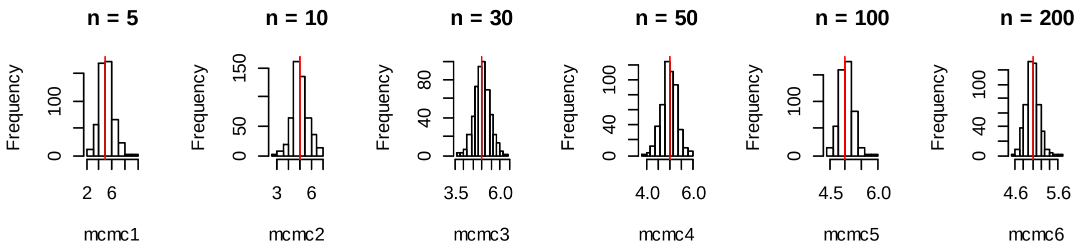

Fig.8. Histograms of estimated parameters λ with MCMC method when $\lambda = 5$.

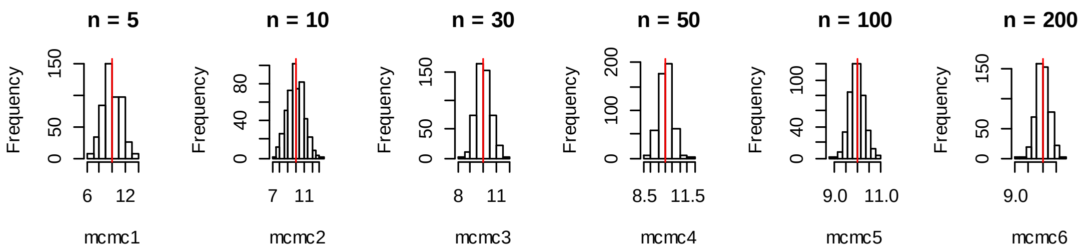

Fig.9. Histograms of estimated parameters λ with MCMC method when $\lambda = 10$.

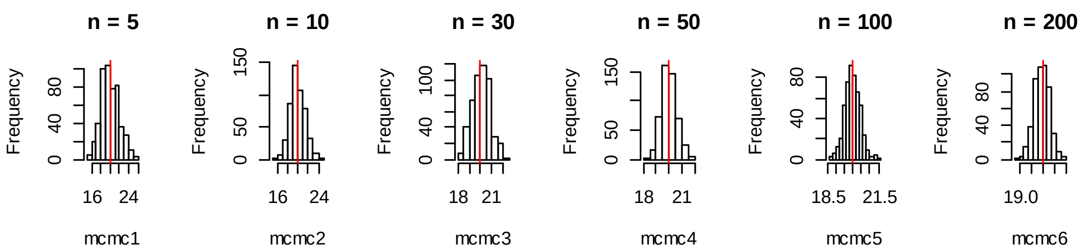

Fig.10. Histograms of estimated parameters λ with MCMC method when $\lambda = 20$.

Table 3. The mean, standard deviation (S.D.), lower confidence interval (LCI), upper confidence interval (UCI), t statistic (t), and p-values obtained via the Bayes method.

Sample sizes	λ	Mean	S.D.	LCI	UCI	t	p-values
n = 5	0.5	0.5566	0.3069	0.5297	0.5836	4.1290	0.0000*
	2	2.0625	0.6718	2.0035	2.1216	2.0825	0.0378*
	5	5.0482	1.0754	4.9537	5.1427	1.0023	0.3167
	10	9.9043	1.4101	9.7804	10.0282	-1.5268	0.1300
	20	20.0830	1.8977	19.9162	20.2497	0.9781	0.3285
n= 10	0.5	0.5554	0.2323	0.5349	0.5758	5.3332	0.0000*
	2	2.0408	0.4449	2.0017	2.0799	2.0535	0.0405*
	5	5.0495	0.6976	4.9882	5.1108	1.5866	0.1132
	10	9.9796	1.0064	9.8912	10.0680	-0.4523	0.6513
	20	20.0169	1.4697	19.8878	20.1461	0.2582	0.7964
n= 30	0.5	0.5207	0.1351	0.5088	0.5326	3.4304	0.0006*
	2	2.0080	0.2665	1.9846	2.0315	0.6787	0.4977
	5	4.9977	0.4204	4.9607	5.0346	-0.1205	0.9041
	10	10.0229	0.5785	9.9720	10.0737	0.8853	0.3764
	20	20.0484	0.7890	19.9791	20.1177	1.3734	0.1702
n= 50	0.5	0.5084	0.0983	0.4997	0.5170	1.9126	0.0563
	2	2.0004	0.2095	1.9820	2.0188	0.0442	0.9648
	5	4.9988	0.3268	4.9701	5.0275	-0.0810	0.9355
	10	10.0188	0.4382	9.9803	10.0573	0.9662	0.3364
	20	20.0067	0.5808	19.9557	20.0577	0.2600	0.7950
n= 100	0.5	0.5053	0.0732	0.4989	0.5117	1.6316	0.1034
	2	2.0107	0.1370	1.9987	2.0228	1.7596	0.0790
	5	5.0172	0.2139	4.9984	5.0360	1.8000	0.0724
	10	10.0101	0.3063	9.9832	10.0371	0.7436	0.4575
	20	20.0002	0.4478	19.9608	20.0395	0.0110	0.9913
n = 200	0.5	0.5024	0.0497	0.4980	0.5068	1.0995	0.2721
	2	2.0053	0.0980	1.9967	2.0139	1.2239	0.2216
	5	5.0061	0.1569	4.9923	5.0199	0.8755	0.3817
	10	10.0104	0.2270	9.9905	10.0304	1.0318	0.3027
	20	19.9737	0.3169	19.9459	20.0016	-1.8505	0.0648

* indicates significance level at 5%

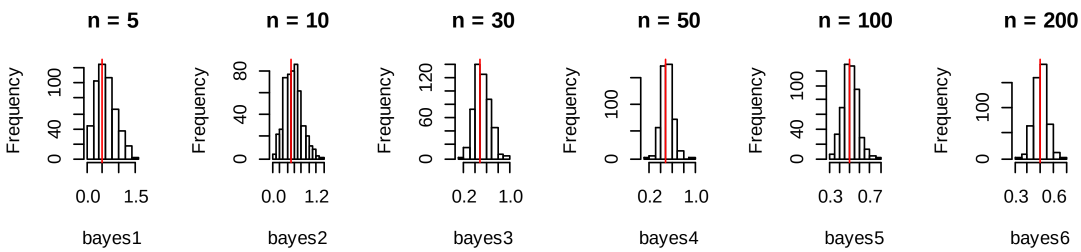

Fig.11. Histograms of estimated parameters λ with Bayes method when $\lambda = 0.5$.

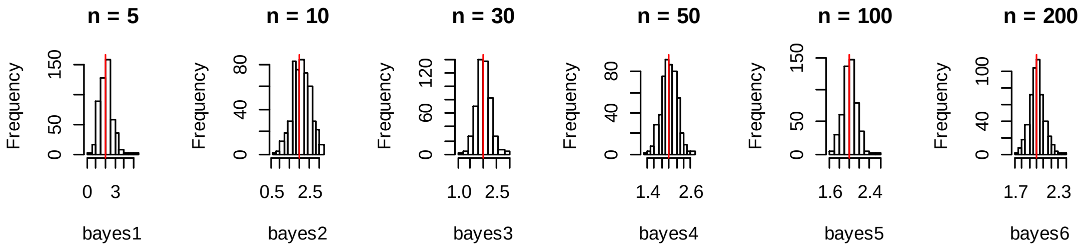

Fig.12. Histograms of estimated parameters λ with Bayes method when $\lambda = 2$.

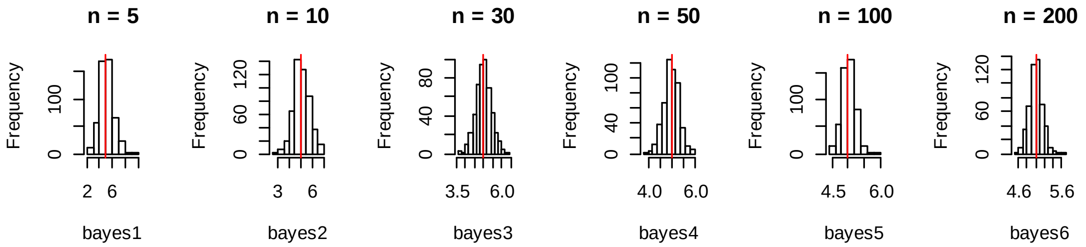

Fig.13. Histograms of estimated parameters λ with Bayes method when $\lambda = 5$.

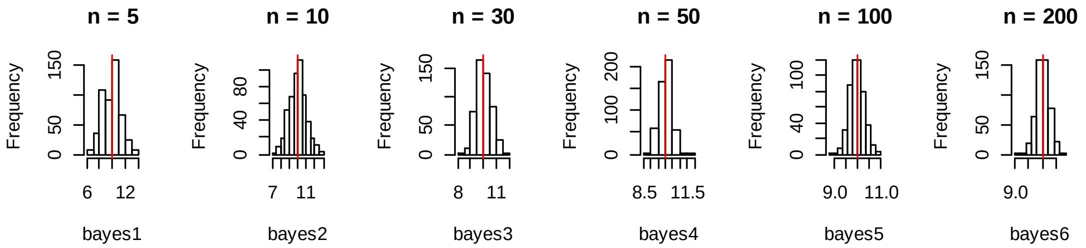

Fig.14. Histograms of estimated parameters λ with Bayes method when $\lambda = 10$.

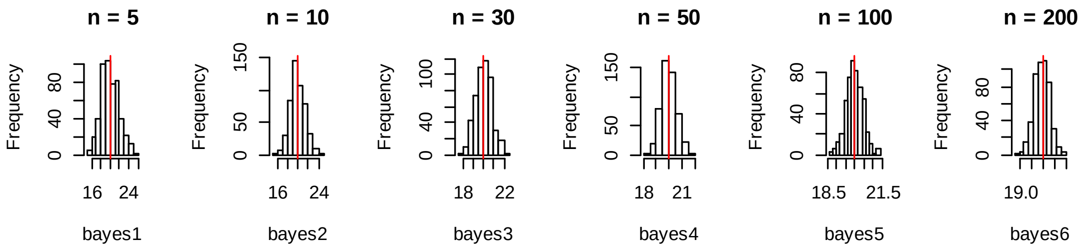

Fig.15. Histograms of estimated parameters λ with Bayes method when $\lambda = 20$.

Table 5 listed the method of parameter estimation when the MSEs is minimum depended on the sample sizes and the true parameters λ. As shown in Table 5, it appears that the following holds.

- For $\lambda = 0.5$, the ML method obtained a minimum MSE for all sample sizes.
- For $\lambda = 2$, the ML method yielded the lowest MSE when n = 5, 10, 30, and 200 whereas the MCMC method yielded the lowest MSE when n = 50 and 100 .
- For $\lambda = 5$, the ML method yielded the lowest MSE when n = 5, 10, 30, and 100 whereas the MCMC method yielded the lowest MSE when n = 200.
- For $\lambda = 10$, the ML method obtained a minimum MSE when n = 30. However, even for n = 10, 50, 100, and 200, the MCMC method produced minimum values of the MSEs. For the smallest sample sizes (n=5), the Bayes method yielded reasonably good estimates in terms of minimizing the MSE.
- For $\lambda = 20$, the ML method yielded the lowest MSE when n = 5, 10, 100, and 200 whereas the Bayes method yielded the lowest MSE when n = 30 and 50 .

Table 5. The minimum values for parameter estimation.

n	λ				
	0.5	**2**	**5**	**10**	**20**
n = 5	ML	ML	ML	Bayes	MCMC
n = 10	ML	ML	ML	MCMC	MCMC
n = 30	ML	ML	ML	ML	Bayes
n = 50	ML	MCMC	ML	MCMC	Bayes
n = 100	ML	MCMC	ML	MCMC	Bayes
n = 200	ML	ML	MCMC	MCMC	Bayes

5. Conclusion

In this paper, we analyzed the Poisson parameter estimation by testing hypothesis and computing a MSE from the ML, the MCMC, and the Bayes method. Through a simulation study, the means of estimated parameters from the ML method were not different from the true parameters in all cases. However, in some cases, the MSEs obtained via the ML method were not minimum compared to the other two mwthods, but it worked well for small sample sizes. For the MCMC and the Bayes methods, the hypothesis tests showed that the means of estimator were different from the true parameters when the sample sizes and the true parameters show large values that the results of the MSE are similar to the other two methods. We would recommend users to use the MCMC and the Bayes method where the sample sizes and parameters were large.

For parameter estimation with Poisson distribution, if we did not dentify the prior distribution, the ML method work reasonably well. On the other hand, when the prior distribution is identified, the MCMC and the Bayes method performed better.

6. References

[1] Gilk W., Richardson S., and Spiegelhalter D., *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London, 1996.

[2] Gemam S., and Geman D., Stochastic Relaxation, Gibbs Distribution and the Bayesian Restoration of Images, *IEEE Transactions on Pattern analysis and Machine Intelligence*, 1984 ;6: 721-741.

[3] Gelfand A., Hills S., Racine-Poon A., and Smith A., Illustration of Bayesian inference in normal data models using Gibbs sampling., *Journal of the American Statistical Association*, 1990; 85 : 972-985.

[4] Casella G., and George E I., Explaining the Gibbs Sampler, *The American Statistician*, 1992;46: 167-174.

[5] Spiegelhalter D., Thomas A., Best N., and Lunn D., *WinBUGS User Manual*, *Version 1.4* , MRC Biostatistics Unit, Institute of Public Health and Department of Epidemiology and Public Health, Imperial College School of Medicine, UK, 2003.

[6] Ntzoufras I., *Bayesian Modeling Using WinBUGS*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2009.

[7] R Development Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2004.