

## **A COMPARISON ON SMALL-SAMPLE POINT AND INTERVAL ESTIMATION OF POISSON DISTRIBUTION**

**MANLIKA TANUSIT**

Department of Mathematics and Statistics  
Faculty of Science  
Maejo University  
Chiang Mai 50290  
Thailand

### **Abstract**

The objective of this study is to compare point estimation methods and interval estimation methods for the Poisson distribution when sample sizes are small. Three methods of point estimation: Maximum likelihood method, Bayesian method, and Minimax method, and three methods of interval estimation: Normal method, Normal-Bayesian method, and Score-Bayesian method are considered. The lowest mean absolute error and the lowest average width are used as the criteria to selection for point estimation and interval estimation, respectively. The scopes of this study consist of sample sizes: 5, 6, 7, 8, 9, 10, and the parameter  $\lambda$  equal to 0.02, 0.04, 0.06, 0.08, and 0.1. Data are simulated 1,000 times generated by using the JAVA software. The results of this research are as follows. For point estimation, we recommend that for all sample sizes and parameter  $\lambda$ , the Bayesian method should be used. In the case of interval estimation, Normal-Bayesian method is recommended for sample sizes 5 to 8,  $\lambda$  between 0.2 to 0.4 and sample sizes 9 to 10,  $\lambda$  equal to 0.2, whereas, Score-Bayesian method should be considered for sample sizes 5 to 8, values of  $\lambda$  ranging from 0.6 to 1 and sample sizes 9 to 10,  $\lambda$  between 0.4 to 1.0.

---

2000 Mathematics Subject Classification: 60XX.

Keywords and phrases: point estimation methods, interval estimation methods, Poisson distribution.

Received August 24, 2009

## 1. Introduction

Estimation is one of the most important methods used in statistics. The value of the chosen parameters is depending on the estimation or conclusion from the sample data. There are 2 methods used to estimate value of parameters. They are Point Estimation and Interval Estimation. Point estimation is parameter estimation by focusing only on some particular value such as average estimation, variance estimation, and population proportion estimation. Interval estimation is the estimation of population parameters at some given point. For example, estimation of population mean within the range of  $A$  and  $B$  at 90% confidence level. For this method, the highest and lowest values for this estimation will be varied depend on level of confidence.

In this research, researcher is interested in comparative study of parameter inference methods for Poisson distribution when sample sizes are small. Point estimation: Maximum likelihood method, Bayesian method, and Minimax method are considered. Interval estimation: Normal method, Normal-Bayesian method, and Score-Bayesian method are considered to find appropriate method under the scope of same study.

## Materials and Methods

### Materials

JAVA Software.

### Methods

1. Create random variable that has Poisson distribution.
2. Compare the methods in point estimation by using the lowest mean absolute error as the selection criteria. The method that gives the lower means absolute error will be the most suitable method to be implemented for this situation. Methods of point estimation are as follows:

#### 2.1. Maximum likelihood method [4]

Maximum likelihood estimator for  $\lambda$  is  $\hat{\lambda} = \frac{x}{n}$ .

2.2. Bayesian method, in this case, determines prior distribution of  $\lambda$  is exponential distribution  $(x; \beta)$  [1]

$$\text{Bayesian estimator for } \lambda \text{ is } \hat{\lambda} = \frac{(x+1)\beta}{n\beta+1}, \text{ when } \beta = \frac{1}{\lambda}.$$

2.3. Minimax method [6]

$$\text{Minimax estimator of } \lambda \text{ is } \hat{\lambda} = \frac{2(x+1)}{3n}.$$

3. Compare the method of interval estimation by using the lowest average wide in making decision and the approximate confidence coefficient that was computed in each trial must be greater than 0.8930, 0.9449, and 0.9877 [5] at level of confidence 90%, 95%, and 99%, respectively, then we will consider that the methods of estimation give the approximate confidence coefficients not less than given confidence coefficient. After that the methods of estimation, that give the approximate confidence coefficients not less than given confidence coefficient will be chosen to do further comparison with the average width of interval estimation. Any methods that are able to give lowest average width of interval estimation, will be considered the method as the most suitable for that situation. Methods of interval estimation are as follows:

3.1. Normal method [2] with confidence interval of  $(1 - \alpha)100\% \lambda$  is

$$\left( \hat{\lambda} - Z_{1-\alpha/2} \sqrt{\frac{\hat{\lambda}(1-\hat{\lambda})}{n}}, \hat{\lambda} + Z_{1-\alpha/2} \sqrt{\frac{\hat{\lambda}(1-\hat{\lambda})}{n}} \right),$$

$$\text{when } \hat{p} = \frac{x}{n}.$$

3.2. Normal-Bayesian method with confidence interval of  $(1 - \alpha)100\% \lambda$  is

$$\left( \hat{\lambda} - Z_{1-\alpha/2} \sqrt{\frac{\hat{\lambda}(1-\hat{\lambda})}{n}}, \hat{\lambda} + Z_{1-\alpha/2} \sqrt{\frac{\hat{\lambda}(1-\hat{\lambda})}{n}} \right),$$

$$\text{when } \hat{\lambda} = \frac{(x+1)\beta}{n\beta+1}.$$

3.3. Score-Bayesian method [3] with confidence interval of  $(1 - \alpha)100\% p$  is

$$\left( \hat{\lambda} + \frac{Z_{1-\alpha/2}^2}{2n} - \frac{Z_{1-\alpha/2}}{n^{1/2}} \sqrt{\hat{\lambda} + \frac{Z_{1-\alpha/2}^4}{4n}}, \hat{\lambda} + \frac{Z_{1-\alpha/2}^2}{2n} + \frac{Z_{1-\alpha/2}}{n^{1/2}} \sqrt{\hat{\lambda} + \frac{Z_{1-\alpha/2}^4}{4n}} \right),$$

when  $\hat{\lambda} = \frac{2(x+1)}{3n}$ .

4. Summary.

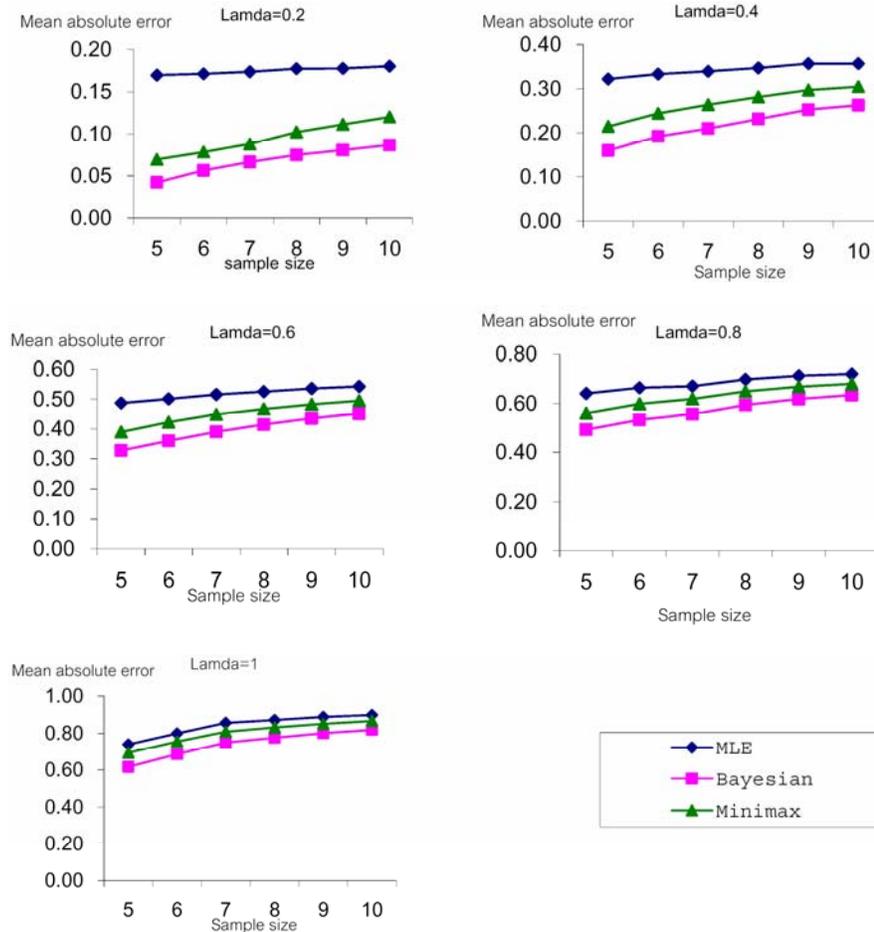
### Result and Discussion

**Point Estimation** from comparing the mean absolute error of parameter estimator estimated from parameter  $\lambda$  shown by 3 methods are summarized in the following table.

**Table 1.** Mean absolute error separate by sample size and parameter

$n$	$\lambda$	Maximum likelihood method	Bayesian method	Minimax method
5	0.2	0.1696	0.0420*	0.0693
	0.4	0.3220	0.1591*	0.2147
	0.6	0.4866	0.3273*	0.3895
	0.8	0.6410	0.4921*	0.5604
	1.0	0.7400	0.6167*	0.6933
6	0.2	0.1710	0.0561*	0.0782
	0.4	0.3331	0.1918*	0.2439
	0.6	0.5007	0.3593*	0.4222
	0.8	0.6645	0.5334*	0.5986
	1.0	0.8000	0.6857*	0.7556
7	0.2	0.1736	0.0664*	0.0879
	0.4	0.3396	0.2100*	0.2642
	0.6	0.5151	0.3905*	0.4482
	0.8	0.6286	0.5179*	0.5905
	1.0	0.8571	0.7500*	0.8095
8	0.2	0.1771	0.0748*	0.1021
	0.4	0.3469	0.2316*	0.2812
	0.6	0.5253	0.4142*	0.4668
	0.8	0.6994	0.5949*	0.6496
	1.0	0.8736	0.7766*	0.8324
9	0.2	0.1775	0.0804*	0.1113
	0.4	0.3571	0.2530*	0.2972
	0.6	0.5352	0.4352*	0.4827
	0.8	0.7143	0.6193*	0.6688
	1.0	0.8900	0.8010*	0.8526
10	0.2	0.1802	0.0865*	0.1201
	0.4	0.3571	0.2628*	0.3047
	0.6	0.5422	0.4511*	0.4948
	0.8	0.7217	0.6349*	0.6811
	1.0	0.9009	0.8190*	0.8673

- The lowest mean absolute error for each case.



**Figure 1.** Examples of mean absolute error separate by sample size and parameter.

From the Table 1 and Figure 1, we can conclude that for all sample sizes and parameter  $\lambda$ , Bayesian method should be used because this gives lowest mean absolute error.

**Interval Estimation** considering level of confidence at 90%, 95%, and 99%, we have found that methods that give the approximate confidence coefficient not less than given confidence coefficient yield are the following:

Normal-Bayesian method: sample sizes 5 to 8,  $\lambda$  between 0.2 to 0.4 and sample sizes 9 to 10,  $\lambda$  is equal to 0.2.

Score-Bayesian method with all parameter  $\lambda$  and all sample sizes.

When taking all given cases that give the approximate confidence coefficient not less than given confidence coefficient into comparison for average width, we have found that the methods that give the lowest average width are:

Normal-Bayesian method: sample sizes 5 to 8,  $\lambda$  between 0.2 to 0.4 and sample sizes 9 to 10,  $\lambda$  is equal to 0.2.

Score-Bayesian method: sample sizes 5 to 8, values of  $\lambda$  ranging from 0.6 to 1 and sample sizes 9 to 10,  $\lambda$  between 0.4 to 1.

### Conclusion

This research can be concluded as follow:

For point estimation, considering the mean absolute error of the Maximum likelihood, Bayesian and Minimax methods found for all sample sizes and parameter  $\lambda$ , Bayesian method should be used because there gave lowest mean absolute error.

For interval estimation, using the lowest average wide as method selection criteria, we conclude that we should use the appropriate methods for each situation as follows. Normal-Bayesian method is recommended for sample sizes 5 to 8,  $\lambda$  between 0.2 to 0.4 and sample sizes 9 to 10,  $\lambda$  equal to 0.2, whereas, Score-Bayesian method should be used for sample sizes 5 to 8, values of  $\lambda$  ranging from 0.6 to 1 and sample sizes 9 to 10,  $\lambda$  between 0.4 to 1.0.

### References

- [1] William M. Bolstad, Introduction to Bayesian Statistics, John Wiley and Sons, Inc., New York, 2004.
- [2] Lawrence D. Brown, T. Tony Cai and Anirban Dasgupta, Interval estimation in exponential families, *Statistica Sinica* 13 (2003), 19-49.

- [3] T. Tony Cai, One-side confidence interval in discrete distributions, *Journal of Statistical Planning and Inference* 131 (2005), 63-88.
- [4] G. Casella and R. L. Berger, *Statistical Inference*, John Wiley and Sons, Inc., New York, 2002.
- [5] B. K. Ghosh, Comparison of some approximate confidence intervals for binomial parameter, *J. Amer. Stat. Assoc.* 74 (1979), 894-900.
- [6] Roger V. Hogg and Elliot A. Tanis, *Probability and Statistical Inference*, Pearson Education, Inc., New Jersey, 2006.

