

5-1-2014

# Comparison of Three Calculation Methods for a Bayesian Inference of Two Poisson Parameters

Yohei Kawasaki

*Tokyo University of Science, Japan, yk\_sep10@yahoo.co.jp*

Etsuo Miyaoka

*Tokyo University of Science, Japan*

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

## Recommended Citation

Kawasaki, Yohei and Miyaoka, Etsuo (2014) "Comparison of Three Calculation Methods for a Bayesian Inference of Two Poisson Parameters," *Journal of Modern Applied Statistical Methods*: Vol. 13 : Iss. 1 , Article 26.

DOI: 10.22237/jmasm/1398918300

# Comparison of Three Calculation Methods for a Bayesian Inference of Two Poisson Parameters

**Yohei Kawasaki**

Tokyo University of Science  
Tokyo, Japan

**Etsuo Miyaoka**

Tokyo University of Science  
Tokyo, Japan

---

The statistical inference drawn from the difference between two independent Poisson parameters is often discussed in medical literature. Kawasaki and Miyaoka (2012) proposed an index  $\theta = P(\lambda_{1,\text{post}} < \lambda_{2,\text{post}})$ , where  $\lambda_{1,\text{post}}$  and  $\lambda_{2,\text{post}}$  denote Poisson parameters following posterior density. A new calculation method is proposed using MCMC and an approximate expression and exact expression for  $\theta$  are compared.

*Keywords:* Poisson distribution, Bayesian inference, MCMC method, Hypergeometric series

---

## Introduction

The statistical inference drawn from the difference between two independent Poisson parameters is often discussed in terms of the frequentist viewpoint rather than the Bayesian viewpoint. In this article, a Poisson parameter is assumed as the relapse rate of a wrong outcome and an adverse reaction rate. Therefore, a low value of the Poisson parameter is desirable.

Classical statistical analysis of outcomes observed in a randomized controlled clinical trial is based on the frequentist approach. The frequentist approach to hypothesis testing is based on the p-value. The inconvenience of using the p-value is well-known and has been documented by Lindley (1957) and Hwang, et al. (1992) among others.

A few different techniques for hypothesis testing have been developed under the Bayesian approach. Basu (1996) briefly showed the use of the Bayesian approach with respect to hypothesis testing. Let  $y$  be data from the probability density function; it is desired to test the null hypothesis against the alternative

---

*Dr. Kawasaki and Dr. Miyaoka are both Professors in the Department of Mathematics.  
Email Dr. Kawasaki at: [yk\\_sep10@yahoo.co.jp](mailto:yk_sep10@yahoo.co.jp).*

## COMPARISON OF THREE CALCULATION METHODS

hypothesis. One approach computes the posterior probability. Poisson parameters were applied to the posterior probability  $\theta$  that shows the difference between the posterior densities of the two independent Poisson parameters, which are considered as random variables. This index can be used to determine the probability that the Poisson parameter of a study drug is superior to that of a control drug.

Kawasaki and Miyaoka (2012) applied  $\theta$  to a one-side hypothesis based on a two-sample situation. They derived an exact and an approximate expression to determine  $\theta$ .

There are some pending issues with the above-mentioned method. An approximate method and exact method of  $\theta$  were adopted only while using a conjugate prior. The drawback of the approximate method is that it occasionally leads to a rough result in a small sample. The drawback of the exact method is that it is slightly complicated. In addition, the exact method requires extensive computing time with a large sample size. Hence, a Markov Chain Monte Carlo (MCMC) method for  $\theta$  is proposed as a solution to these problems.

### Methodology

If  $X_i$  is the number of events in a population of  $n_i$  patients (or over  $n_i$  units of time), and  $\lambda_i$  is the event rate, then the sampling distribution is

$$f_{X_i}(x_i | \lambda_i) = \frac{e^{-n_i \lambda_i} (n_i \lambda_i)^{x_i}}{x_i!}, \quad (1)$$

where  $i = 1, 2$ . The conjugate prior density of  $\lambda_i$  is the gamma distribution with parameters  $\alpha_i$  and  $\beta_i$  :

$$f_i(\lambda_i | \alpha_i, \beta_i) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda_i^{\alpha_i-1} e^{-\beta_i \lambda_i} \quad \lambda_i > 0 \quad (2)$$

where  $\alpha_i > 0$  and  $\beta_i > 0$ . The posterior density for  $\lambda_i$  is given as

$$g_i(\lambda_{i,post}) = \frac{b_i^{a_i}}{\Gamma(a_i)} \lambda_i^{a_i-1} e^{-b_i \lambda_i}, \quad (3)$$

where  $a_i = \alpha_i + x_i$ ,  $b_i = n_i + \beta_i$  and  $\Gamma(a)$  denotes the gamma function. Let  $\lambda_{i,post}$  denote the Poisson parameter in the posterior density.

### Approximate expression for $\theta$

$\theta$  can be calculated via an approximation using the standard normal table; assume that sample sizes,  $n_1$  and  $n_2$ , are large. It is necessary to determine a  $Z$ -test statistic. The expected difference in the posterior density and the variance in this difference can be expressed as:

$$E(\lambda_{1,post} - \lambda_{2,post}) = \mu_{1,post} - \mu_{2,post}, \quad (4)$$

$$V(\lambda_{1,post} - \lambda_{2,post}) = a_1 / b_1^2 + a_2 / b_2^2 \quad (5)$$

where  $\mu_{i,post} = a_i / b_i$  denote the posterior mean of  $\lambda_i$ . The  $Z_g$ -test statistic is

$$Z_g = \frac{(\lambda_{1,post} - \lambda_{2,post}) - E(\lambda_{1,post} - \lambda_{2,post})}{\sqrt{V(\lambda_{1,post} - \lambda_{2,post})}} \quad (6)$$

The  $Z_g$ -test statistic is approximately distributed according to the standard normal distribution. Therefore, the approximate probability of the index  $\theta$  is given as

$$\theta = P(\lambda_{1,post} < \lambda_{2,post}) \approx \Phi\left(\frac{-\mu_{1,post} + \mu_{2,post}}{\sqrt{a_1 / b_1^2 + a_2 / b_2^2}}\right) \quad (7)$$

where  $\Phi(\bullet)$  is the cumulative distribution function of the standard normal distribution. From this the approximate probability can be easily calculated.

### Exact method for $\theta$

Kawasaki and Miyaoka (2012) derived the exact expression for  $\theta$  using the posterior density. The exact expression for  $\theta$  is

## COMPARISON OF THREE CALCULATION METHODS

$$\begin{aligned}\theta &= P(\lambda_{1,post} < \lambda_{2,post}) \\ &= 1 - \frac{1}{a_2 \text{Beta}(a_1, a_2)} \left(\frac{b_2}{b_1 + b_2}\right)^{a_2} {}_2F_1\left(a_2, 1 - a_1, 1 + a_2, \frac{b_2}{b_1 + b_2}\right)\end{aligned}\quad (8)$$

where

$${}_2F_1(k_1, k_2; l; u) = \sum_{t=0}^{\infty} \frac{(k_1)_t (k_2)_t}{(l)_t} \frac{u^t}{t!} \quad (9)$$

is the hypergeometric series, and  $(k)_t$  is the Pochhammer symbol.

### MCMC Method for $\theta$

A computational procedure for  $\theta$  can be described using the MCMC method. The MCMC method is a means of sampling from a posterior density. A random-walk Metropolis-Hasting algorithm was used as the MCMC Method. Given that the samples come from two independent populations, the posterior joint distribution of  $\lambda_1$  and  $\lambda_2$  is a product of its marginal distributions. For this reason, samples can be obtained from the posterior distribution of  $\lambda_1 - \lambda_2$  by simulating  $k$  values from the posterior distribution of  $\lambda_1$  and  $\lambda_2$  using MCMC procedure of SAS, e.g.,  $\lambda_{1,post}^1, \lambda_{1,post}^2, \dots, \lambda_{1,post}^k$  and  $\lambda_{2,post}^1, \lambda_{2,post}^2, \dots, \lambda_{2,post}^k$ , respectively. By computing  $\lambda_{1,post}^1 - \lambda_{1,post}^1, \lambda_{1,post}^2 - \lambda_{1,post}^2, \dots, \lambda_{1,post}^k - \lambda_{1,post}^k$ , it is possible to obtain the simulated values from the posterior distribution of  $\lambda_1 - \lambda_2$ . The posterior samples obtained by the MCMC method after the burn-in period are  $\delta_1, \delta_2, \dots, \delta_k$ . Let  $\Delta_1, \Delta_2, \dots, \Delta_k$  be independent identically distributed random variables with distribution function  $F$ . The posterior sample is the observed value of  $\Delta_1, \Delta_2, \dots, \Delta_k$ . Note that  $\theta = P(\lambda_{1,post} < \lambda_{2,post})$  equals  $\theta = P(\lambda_{1,post} - \lambda_{2,post} < 0)$ , thus,  $\theta$  can be expressed as

$$\theta = P(\lambda_{1,post} < \lambda_{2,post}) = P(\lambda_{1,post} - \lambda_{2,post} < 0) \approx 1 - \hat{F}_k(0) \quad (10)$$

where

$$\hat{F}_k(s) = \frac{1}{k} \sum_{i=1}^k I(\Delta_i \leq s) \quad (11)$$

and

$$I(\Delta_i \leq s) = \begin{cases} 1 & \text{if } \Delta_i \leq s \\ 0 & \text{if } \Delta_i > s \end{cases} \quad (12)$$

is the empirical distribution function.

## Results

### Comparison of three methods

To compare the probabilities of the three methods for  $\theta$ , the difference between the sample rates (horizontal axis) where sample rate is calculated as  $X_i/n_i$ , were plotted against the difference between the probabilities of the MCMC and exact methods (vertical axis), as shown in Figures 1, 3, and 5. Similarly, the difference between the sample rates (horizontal axis) were plotted against the difference between the probabilities of the approximate and exact methods (vertical axis), as shown in Figures 2, 4, and 6. Figures 1, and 2 show situations that considered small sample sizes, i.e.,  $n_1 = n_2 = 10, 15, 20$ , and  $25$ ; Figures 3 and 4, show larger sample sizes, i.e.,  $n_1 = n_2 = 60, 70, 90$ , and  $100$ . Figures 5 and 6 consider groups of different sample sizes, that is,  $n_1 = 5, n_2 = 15$ ;  $n_1 = 5, n_2 = 25$ ;  $n_1 = 15, n_2 = 5$  and  $n_1 = 25, n_2 = 5$ .

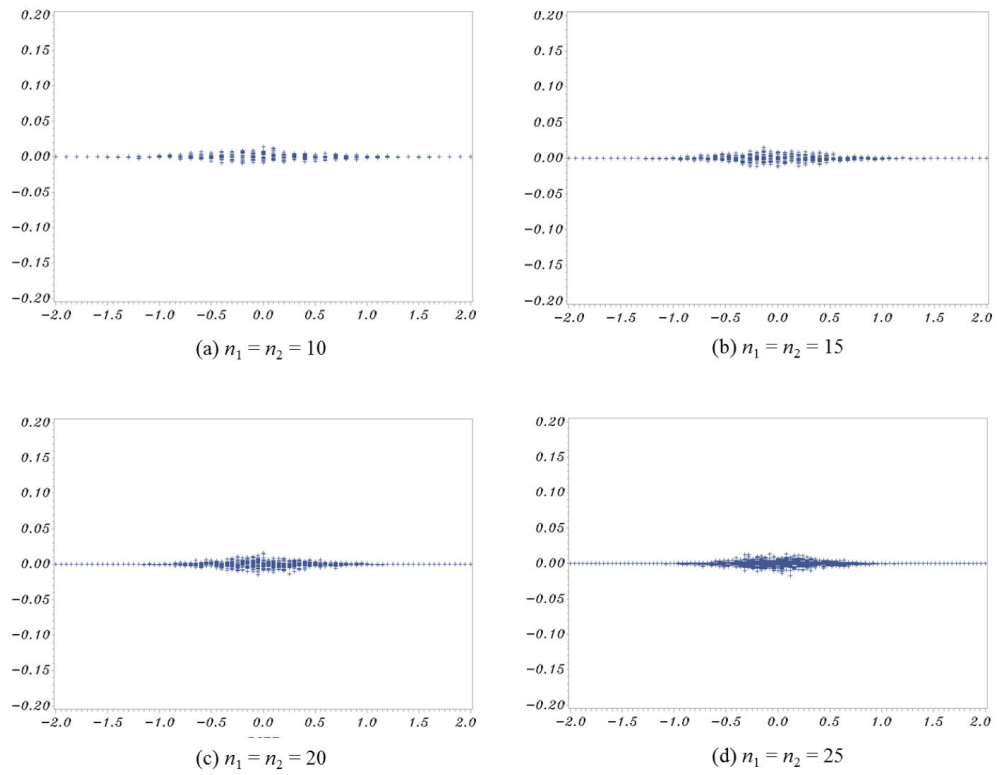
### Relationship between the difference in the probabilities and the difference in the sample rates.

In Figures 1(d) and 3(d), the probability of the MCMC method is approximately equal to that of the exact method when the difference between the sample rates is 1.0. Conversely, the difference between the probabilities of the MCMC and exact methods is around 0.01 when the difference between the sample rates is zero. Overall, when the difference between the sample rates is large, the probabilities of the MCMC and exact methods are roughly equal. By contrast, when the difference between the sample rates is small, the probability of the MCMC method is different from that of the exact method. This general pattern is similar for the difference in the probabilities of the approximation and exact methods.

### **Relationship between the sample size and the difference in the probabilities**

In Figure 2(a), the difference between the probabilities of the approximate and exact methods is around 0.15 when the difference between the sample rates is 0.01. For a slightly larger sample size (see Figure 2(d)), the difference between the probabilities of the approximate and exact methods is around 0.05 for the same difference between the sample rates. In addition, there is virtually no difference between the probabilities of the approximate and exact methods when the sample size is further increased, as shown in Figure 4(d). Thus, the sample size influences the accuracy of the probability of the approximate method. Also shown is the difference in the probabilities of the MCMC and exact methods. In Figure 1(a), the difference between the probabilities of the MCMC and exact methods is around 0.01 when the difference between the sample rates is zero. For a slightly larger sample size (see Figure 3(d)), the difference between the probabilities of the MCMC and exact method is around 0.01 for the same difference between the sample rates. Thus, the accuracy of the probability of the MCMC method always remains high even when the sample sizes are small.

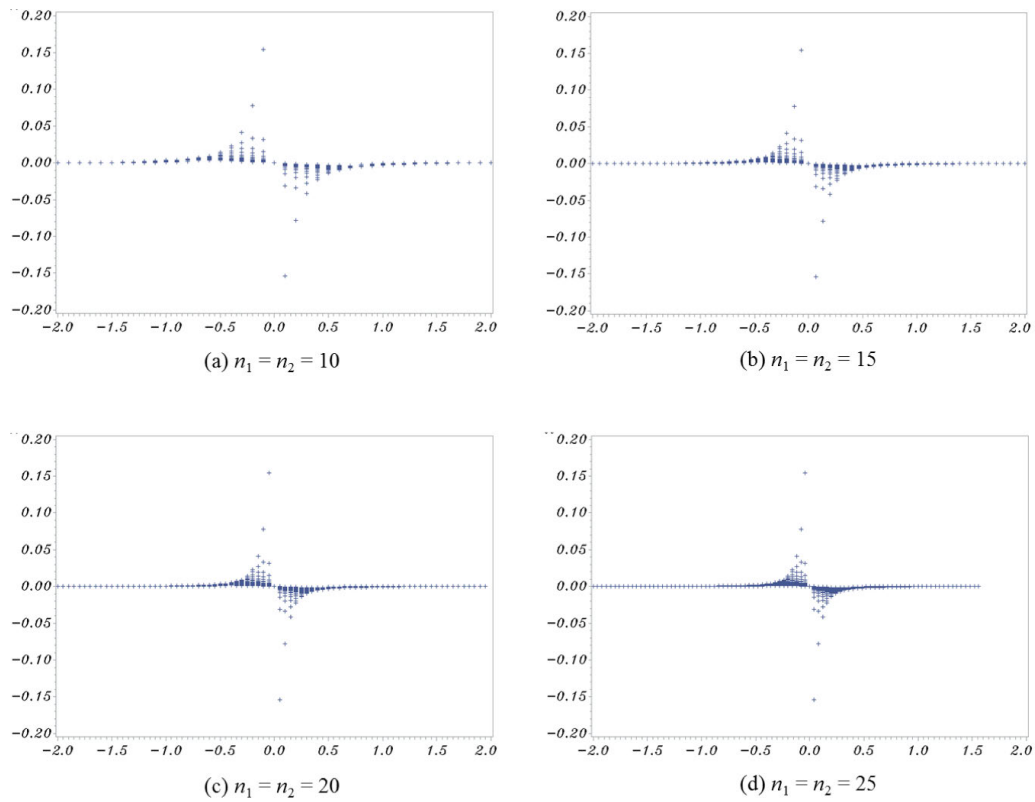
Finally, the difference between the probabilities when groups of different sample sizes are considered was investigated. In Figure 2(a), the difference between the probabilities of the approximate and exact methods is around -0.01 when the difference between the sample rates is 0.5. Conversely, in Figure 6(a), the difference between the probabilities of the approximate and exact methods is around -0.05 for the same difference between the sample rates. In both the cases, the total sample size ( $n_1 + n_2$ ) is the same. However, the difference between the probabilities of the approximate and exact methods is slightly greater in the case of groups with different sample sizes; the case of the MCMC method is also shown. In Figure 1(d), the difference between the probabilities of the MCMC and exact methods is around 0.01 when the difference between the sample rates is zero. Conversely, in Figure 5(d), the difference between the probability of the MCMC and exact methods is around 0.01 for the same difference between the sample rates. Therefore, the difference between the probabilities of the MCMC and exact methods is the same regardless of whether the sample sizes are equal or different.



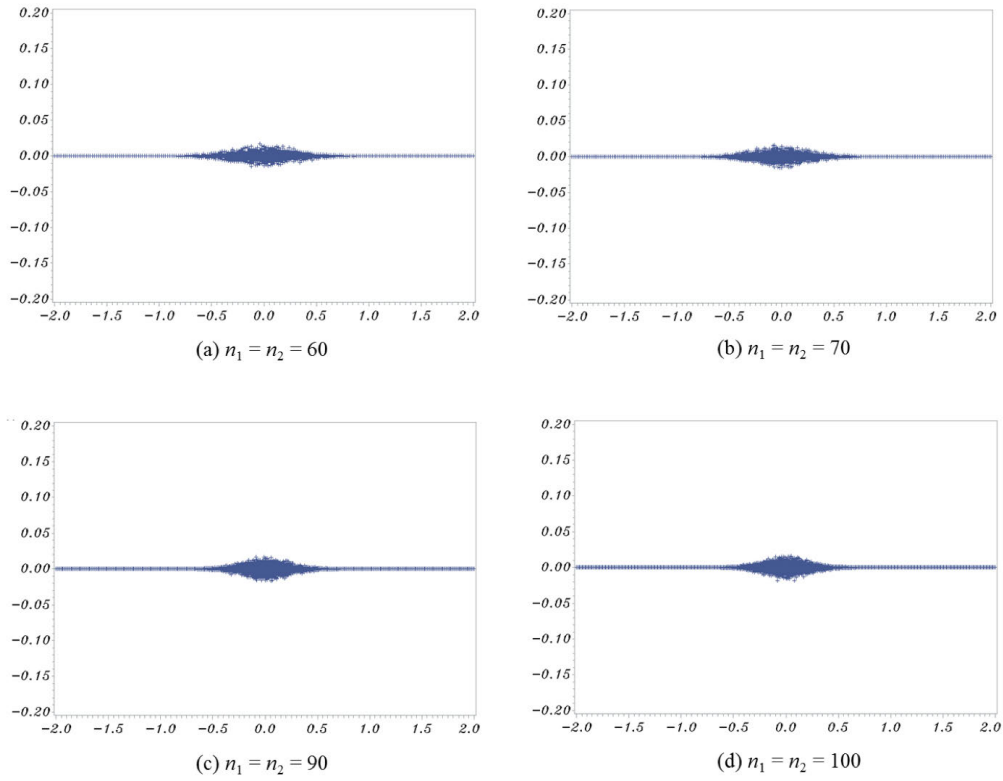
**Figure 1.** Comparison of the Exact and MCMC Method when sample sizes are small. (vertical axis : Differences of  $\theta$  in Exact and MCMC method. Prior distribution is Gamma(0.01,0.01). horizontal axis : Differences of two sample rates.



## COMPARISON OF THREE CALCULATION METHODS

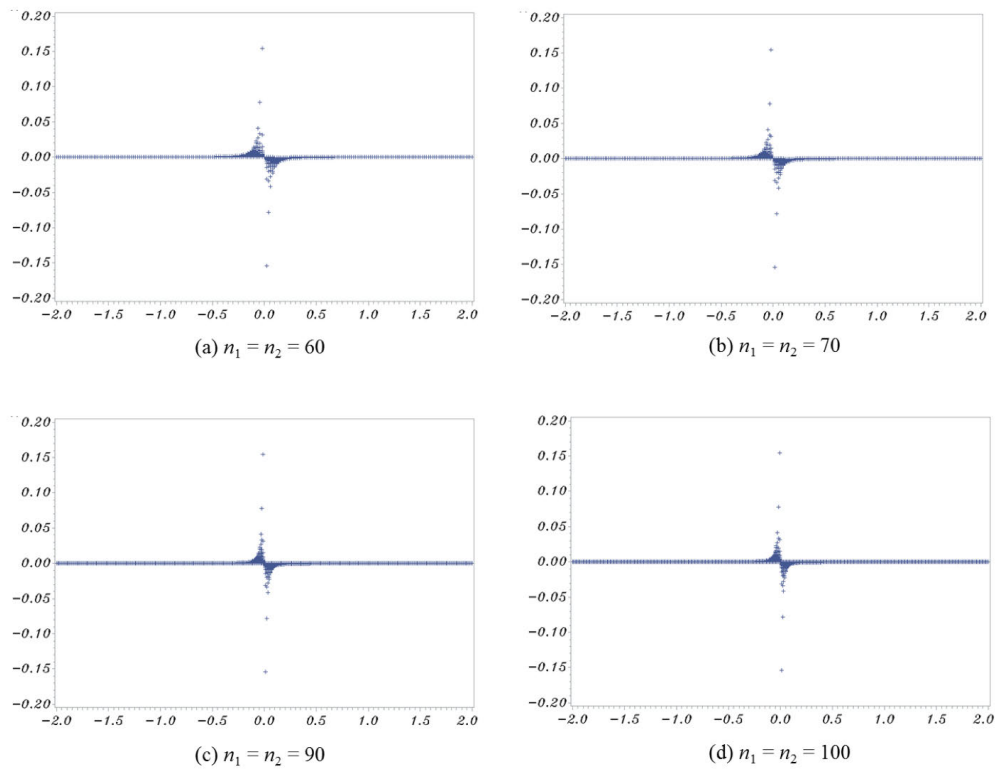


**Figure 2.** Comparison of the Exact and Approximate method when sample sizes are small. (vertical axis : Differences of  $\theta$  in Exact and Approximation method. horizontal axis : Differences of two sample rates.



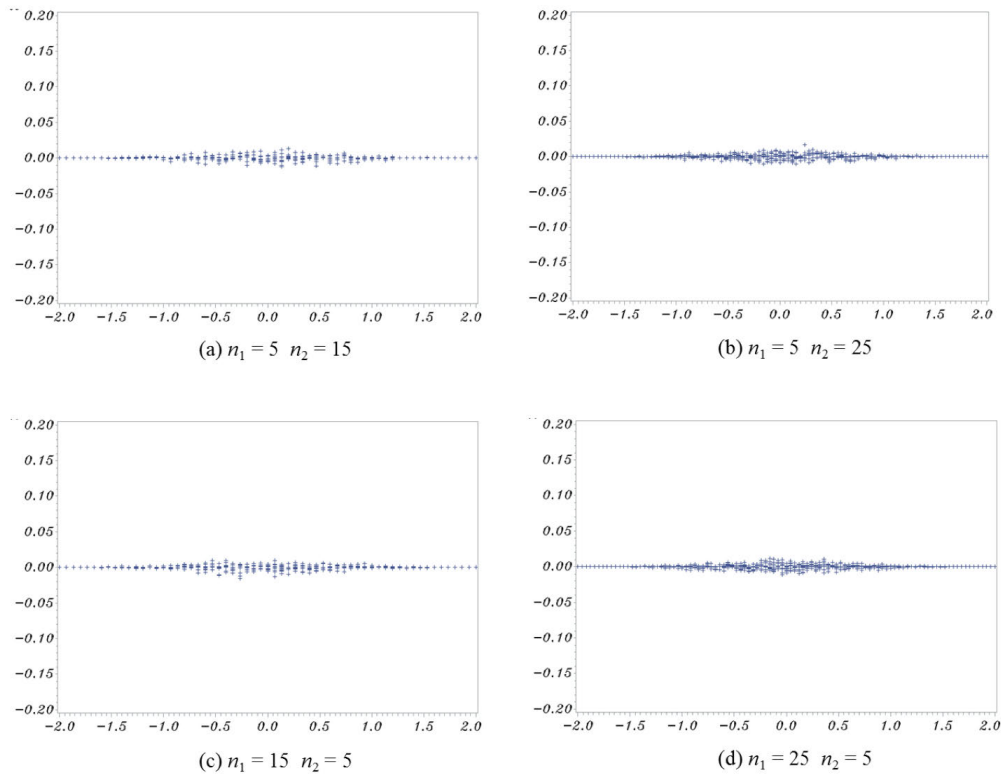
**Figure 3.** Comparison of the Exact and MCMC Method when sample sizes are large. (vertical axis : Differences of  $\theta$  in Exact and MCMC method. Prior distribution is Gamma(0.01,0.01). horizontal axis : Differences of two sample rates.

## COMPARISON OF THREE CALCULATION METHODS



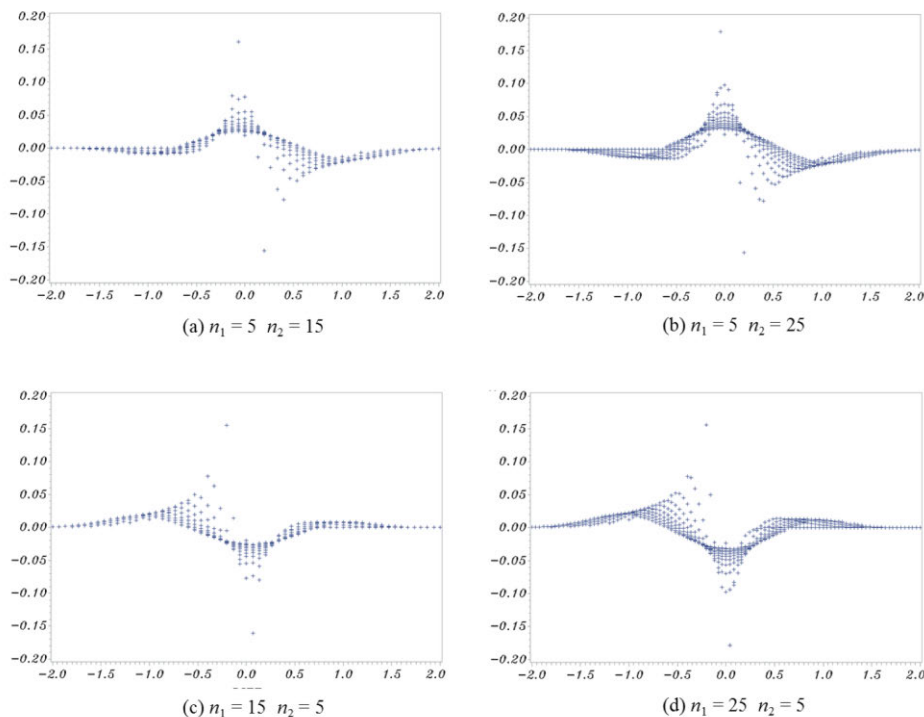
**Figure 4.** Comparison of the Exact and Approximate method when sample sizes are large. (vertical axis : Differences of  $\theta$  in Exact and Approximation method. horizontal axis : Differences of two sample rates.

---



**Figure 5:** Comparison of the Exact and MCMC Method when sample sizes are unbalanced. (vertical axis : Differences of  $\theta$  in Exact and MCMC method. Prior distribution is Gamma(0.01,0.01). horizontal axis : Differences of two sample rates.

## COMPARISON OF THREE CALCULATION METHODS



**Figure 6:** Comparison of the Exact and Approximate method when sample sizes are unbalanced. (vertical axis : Differences of  $\theta$  in Exact and Approximation method. horizontal axis : Differences of two sample rates.

---

## Conclusion

Three calculation methods were presented for the index  $\theta = P(\lambda_{1,post} < \lambda_{2,post})$ . A new procedure was described based on the MCMC method. The probabilities of these three methods were compared in order to test the relative effectiveness of each.

The expression for the exact method was presented, which includes a hypergeometric series, and it was speculated that this series causes the decrease in calculation efficiency when the sample size is very large. In addition, hypergeometric series are not built into SAS, which is a statistical software program frequently used in pharmaceutical development. Therefore, if SAS is used, a calculation program for hypergeometric series must be developed.

It is easy to calculate the probability for using the approximation method. This is an advantage when the approximate probability is used. Conversely, when

the difference in the sample rates is small and the sample sizes are unbalanced, the accuracy the approximation method is poor. That is, the accuracy of the probability of the approximation method depends on the sample size.

This study showed that the accuracy of the MCMC method was greater than that of the approximation method. Moreover, the probability of the MCMC method can be easily calculated using SAS. In addition, it is possible to use the non-conjugate prior for the prior distribution in the MCMC method. This is considered to be one of the advantages of the MCMC method.

## References

- Basu, S. (1996). Bayesian hypotheses testing using posterior density ratios. *Statistics and Probability Letters*, 30(1): 79-86.
- Hwang, J. T., Casella, G., Robert, C., Wells, M. T., & Farrell, R. H. (1992). Estimation of accuracy in testing. *The Annals of Statistics*, 20(1): 490-509.
- Lindley, D. V. (1957). A statistical paradox. *Biometrika*, 44(1/2): 187-192.
- Kawasaki, Y., & Miyaoka, E. (2012). A Bayesian inference of  $P(\lambda_1 < \lambda_2)$  for two Poisson parameters. *Journal of Applied Statistics*, 39(10): 2141-2152.