J. Japan Statist. Soc.
Vol. 40 No. 2 2010 265–275

A POSTERIOR DENSITY FOR THE DIFFERENCE BETWEEN TWO BINOMINAL PROPORTIONS AND THE HIGHEST POSTERIOR DENSITY CREDIBLE INTERVAL

Youhei Kawasaki* ** and Etsuo Miyaoka***

The statistical inference concerning the difference between two independent binominal proportions has often been discussed in medical and statistical literature. This discussion is far more often based on the frequency theory of statistical inference than on the Bayesian theory. In this article, we propose the expression of the posterior probability density function (pdf) for the difference between two independent binominal proportions. In addition, we calculate the exact Highest Posterior Density (HPD) credible interval by using this expression. We also compare both the exact HPD credible interval and the approximate credible interval. We find that the former always has a narrower interval length than the latter.

Key words and phrases: Posterior density, binominal proportions, highest posterior density credible interval, hypergeometrics series.

1. Introduction

The statistical inference for the difference between two independent binominal proportions has been frequently discussed in papers. This interest arises from the fact that the actual confidence level does not correspond to the nominal confidence level (see Agresti and Caffo (2000)). The approximate confidence interval is constructed by using the asymptotical argument. Therefore, the finite sample size, particularly small sample sizes, creates difficulties. Many authors, such as Newcombe (1998), Brown and Li (2005) and Kawasaki *et al.* (2008), have shown that the actual confidence level is lower than the nominal confidence level.

Santner and Snell (1980) proposed the exact confidence interval that uses the relationship between the hypothesis test and confidence interval. However, Chan and Zhang (1999) pointed out that the exact confidence interval proposed by these two authors is too conservative. Yanakawa and Bian (2001) recommended an improvement for the case in which the observation value is extreme. Further, Agresti and Min (2001) and Andres and Tejedor (2003) provided a computing method for the confidence interval that uses a two-sided test. To compute their exact confidence interval, however, repeated calculations are necessary. Berger

Received May 31, 2010. Revised August 31, 2010. Accepted September 2, 2010.

*Biostatistics Group, Data Science Department, Development Division, Mitsubishi Tanabe Pharma Corporation, 2-2-6 Nihonbashi-Honcho, Chuo-ku, Tokyo 103-8405, Japan.

**Graduate School of Science, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan.

***Department of Math, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan.

and Boos (1994) and Kawasaki (2010) sought to improve the calculation efficiency by focusing on the restricted range of the nuisance parameter.

On the other hand, the Bayesian approach also has been applied to statistical inference of the binominal proportion. Draper and Guttman (1971) explored Bayesian estimation of the binomial sample size n based on r independent binomial observations. Using an independent beta distribution to the prior density, Novick and Grizzle (1965) focused on finding the posterior probability that $\pi_1 > \pi_2$ and discuss its application to sequential clinical trials. Cornfield (1966) also examined sequential trials from a Bayesian viewpoint, focusing on stopping rule theory. Irony and Pereira (1986) compared Fisher's exact test with a Bayesian test. Howard (1998) and Seneta and Phipps (2001) discussed the test that used the Bayesian inference. Especially, they showed the relation between Liebermeister's method and Bayesian method. Seneta (1994) discussed mathematics leading from the Bayesian method to Liebermeister's method. Hasemi *et al.* (1997) studied the Baysian inference for some associate measures, such as risk difference, relative risk, and the odds ratio. Agresti and Caffo (2000) and Pan (2002) attempted the improvement of the Wald interval by using the Bayes estimator. Agresti and Min (2005) studied the frequentist performance of Bayesian intervals for comparing the proportion of two independent binominal samples and found that the Jeffreys prior performance is as good as the score interval. Brown and Li (2005) argued that if one desires good coverage performance over the entire parameter space, it is best to use quite diffuse priors. However, even uniform priors are often too informative, and they recommended the Jeffreys prior.

However, these works have not shown an accurate posterior pdf for the difference between two independent binominal proportions. We give the expression of the accurate posterior pdf in Section 2. We calculate an HPD credible interval by using this expression. In addition, we calculate the approximate credible interval, and compare both the HPD credible interval and the approximate credible interval in Section 3 and Section 4. In Section 5, we conclude with a brief summary.

2. Difference of the posterior distribution for binominal proportion

Let X_1 and X_2 denote a binomial random variable for n_1 trials and n_2 trials and parameter π_1 and π_2, respectively. The conjugate prior density for π_i $(i = 1, 2)$ is the beta distribution with parameters α_i and β_i, where parameters $\alpha_i > 0$ and $\beta_i > 0$. The posterior pdf for π_i is given by

$$g_i(\pi_{i,\text{post}}) = \frac{1}{B(a_i, b_i)} \pi_{i,\text{post}}^{a_i - 1} (1 - \pi_{i,\text{post}})^{b_i - 1},$$

where $a_i = \alpha_i + x_i$, $b_i = n_i - x_i + \beta_i$ and $B(a, b)$ denotes the beta function. Let $\pi_{i,\text{post}}$ denote the binominal proportion in the posterior density. Using the posterior pdf, we can derive the exact posterior pdf for the difference between two independent binominal proportions.

THEOREM 1. *The exact posterior pdf $f_\Delta(\delta)$ for $\delta = \pi_{1,\text{post}} - \pi_{2,\text{post}}$ is proposed:*

$$f_\Delta(\delta) = \begin{cases} \dfrac{B(a_2,b_1)(1-\delta)^{a_2+b_1-1}}{B(a_1,b_1)B(a_2,b_2)} F_3(a_2,b_1,1-b_2,1-a_1;a_2+b_1;1-\delta,1-\delta) \\ \qquad\qquad \textit{for} \quad 0<\delta\le 1 \\ \dfrac{B(a_1,b_2)(1+\delta)^{a_1+b_2-1}}{B(a_1,b_1)B(a_2,b_2)} F_3(a_1,b_2,1-b_1,1-a_2;a_1+b_2;1+\delta,1+\delta) \\ \qquad\qquad \textit{for} \quad -1\le\delta<0 \\ \dfrac{B(a_1+a_2-1,b_1+b_2-1)}{B(a_1,b_1)B(a_2,b_2)} \qquad \textit{for} \quad \delta=0 \end{cases}$$

where

$$F_3(k_1,k_2,l_1,l_2;h;u_1,u_2) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(k_1)_i(k_2)_j(l_1)_i(l_2)_j}{(h)_{i+j}} \frac{u_1^i}{i!}\frac{u_2^j}{j!},$$

$\max\{|u_1|,|u_2|\}<1$, *denotes the Appell hypergeometric function, and $(k)_i$ is the Pochhammer symbol.*

PROOF. Pham-Gia and Turkkan (2008) led the pdf when X and Y are dependent. We give the proof similar to the approach in Pham-Gia and Turkkan (2008).

For $0<\delta\le 1$, we wish to find the pdf for δ. Using the change of variables $\delta = \pi_{1,\text{post}} - \pi_{2,\text{post}}$ and $\gamma = \pi_{2,\text{post}}$, the pdf can be written as follows,

$$\begin{aligned} f_\Delta(\delta) &= \int_0^{1-\delta} g_1(\delta+\gamma)g_2(\gamma)d\gamma \\ &= \frac{1}{B(a_1,b_1)B(a_2,b_2)} \int_0^{1-\delta} (\delta+\gamma)^{a_1-1}(1-\delta-\gamma)^{b_1-1}\gamma^{a_2-1}(1-\gamma)^{b_2-1}d\gamma \\ &= \frac{\delta^{a_1-1}(1-\delta)^{b_1-1}}{B(a_1,b_1)B(a_2,b_2)} \\ &\quad \times \int_0^{1-\delta} \left(1+\frac{\gamma}{\delta}\right)^{a_1-1} \left(1-\frac{\gamma}{1-\delta}\right)^{b_1-1} \gamma^{a_2-1}(1-\gamma)^{b_2-1}d\gamma. \end{aligned} \tag{2.1}$$

Additionally, using the change of variables $v=\gamma/(1-\delta)$, the pdf of δ can be written

$$\begin{aligned} f_\Delta(\delta) &= \frac{\delta^{a_1-1}(1-\delta)^{a_2+b_1-1}}{B(a_1,b_1)B(a_2,b_2)} \\ &\quad \times \int_0^1 v^{a_2-1}(1-v)^{b_1-1}(1-v(1-\delta))^{b_2-1}\left(1-\frac{v(\delta-1)}{\delta}\right)^{a_1-1} dv. \end{aligned} \tag{2.2}$$

Here, we use the Appell hypergeometric function, the expression (2.2) can be written as follows,

$$\begin{aligned} f_\Delta(\delta) &= \frac{B(a_2,b_1)\delta^{a_1-1}(1-\delta)^{a_2+b_1-1}}{B(a_1,b_1)B(a_2,b_2)} \\ &\quad \times F_1\left(a_2,1-b_2,1-a_1;a_2+b_1;1-\delta,\frac{\delta-1}{\delta}\right) \end{aligned} \tag{2.3}$$

(e.g., see Bailey (1972)). At the same time, we think that expression (2.3) does not fill the convergence condition, that is, the convergence condition is $|(\delta-1)/\delta|<1$. Srivastava and Karlsson (1985) proposed a transformation for the Appell hypergeometric function; using this technique we express that the Appell hypergeometric function is

$$\begin{aligned} &F_1\left(a_2, 1-b_2, 1-a_1; a_2+b_1; 1-\delta, \frac{\delta-1}{\delta}\right) \\ &\quad = \delta^{1-a_1} F_3(a_2, b_1, 1-b_2, 1-a_1; a_2+b_1; 1-\delta, 1-\delta). \end{aligned}$$

Hence, we can show that expression (2.3) becomes

$$f_\Delta(\delta) = \frac{B(a_2,b_1)(1-\delta)^{a_2+b_1-1}}{B(a_1,b_1)B(a_2,b_2)} F_3(a_2, b_1, 1-b_2, 1-a_1; a_2+b_1; 1-\delta, 1-\delta),$$

where $|1-\delta|<1$.

For $-1 \le \delta < 0$, we use the change of variables $\delta = \pi_{1,\text{post}} - \pi_{2,\text{post}}$ and $\nu = \pi_{1,\text{post}}$. The pdf of δ can be written as follows,

$$\begin{aligned} f_\Delta(\delta) &= \int_0^{1+\delta} g_1(\nu) g_2(\nu-\delta) d\nu \\ &= \frac{1}{B(a_1,b_1)B(a_2,b_2)} \int_0^{1+\delta} \nu^{a_1-1}(1-\nu)^{b_1-1}(\nu-\delta)^{a_2-1}(1-\nu+\delta)^{b_2-1} d\nu \\ &= \frac{(-\delta)^{a_2-1}(1+\delta)^{b_2-1}}{B(a_1,b_1)B(a_2,b_2)} \\ &\quad \times \int_0^{1+\delta} \nu^{a_1-1}(1-\nu)^{b_1-1}\left(1-\frac{\nu}{\delta}\right)^{a_2-1}\left(1-\frac{\nu}{1+\delta}\right)^{b_2-1} d\nu. \end{aligned}$$

Additionally, we use the change of variables $u = \nu/(1+\delta)$. The pdf of δ can be written as follows,

$$\begin{aligned} f_\Delta(\delta) &= \frac{(-\delta)^{a_2-1}(1+\delta)^{a_1+b_2-1}}{B(a_1,b_1)B(a_2,b_2)} \\ &\quad \times \int_0^1 u^{a_1-1}(1-u)^{b_2-1}(1-u(1+\delta))^{b_1-1}\left(1-\frac{u(1+\delta)}{\delta}\right)^{a_2-1} du \\ &= \frac{B(a_1,b_2)(-\delta)^{a_2-1}(1+\delta)^{a_1+b_2-1}}{B(a_1,b_1)B(a_2,b_2)} \\ &\quad \times F_1\left(a_1, 1-b_1, 1-a_2, a_1+b_2; 1+\delta; \frac{1+\delta}{\delta}\right) \\ &= \frac{B(a_1,b_2)(1+\delta)^{a_1+b_2-1}}{B(a_1,b_1)B(a_2,b_2)} F_3(a_1, b_2, 1-b_1, 1-a_2; a_1+b_2; 1+\delta; 1+\delta), \end{aligned}$$

where $|1+\delta|<1$.

Finally, from the expression (2.1), we give the pdf for $\delta = 0$,

$$\begin{aligned} f_\Delta(0) &= \frac{1}{B(a_1,b_1)B(a_2,b_2)} \int_0^1 \gamma^{a_1+a_2-2}(1-\gamma)^{b_1+b_2-2} d\gamma \\ &= \frac{B(a_1+a_2-1, b_1+b_2-1)}{B(a_1,b_1)B(a_2,b_2)}, \end{aligned}$$

where $a_1 + a_2 > 1$ and $b_1 + b_2 > 1$.

3. Credible interval

In this section, we give the method of construction for the exact HPD credible interval and the approximate credible interval.

3.1. Approximate credible interval

We can show the expectation of difference for the posterior density and the variance of difference for posterior density as follows,

$$E(\delta) = \mu_{1,\text{post}} - \mu_{2,\text{post}}, \qquad V(\delta) = \frac{\mu_{1,\text{post}}(1-\mu_{1,\text{post}})}{a_1+b_1+1} + \frac{\mu_{2,\text{post}}(1-\mu_{2,\text{post}})}{a_2+b_2+1},$$

where $\mu_{i,\text{post}} = a_i/(a_i+b_i)$ denotes the posterior mean of π_i, and by the Central Limit Theorem,

$$Z = \frac{\delta - E(\delta)}{\sqrt{V(\delta)}}$$

is approximately distributed as the standard normal distribution. Therefore, an approximate $100(1-\alpha)\%$ credible interval for δ is

$$\mu_{1,\text{post}} - \mu_{2,\text{post}} \pm z_{\alpha/2}\sqrt{v_{1,\text{post}} + v_{2,\text{post}}},$$

where z_α is the $100(1-\alpha)$-th percentile point of the standard normal density function, and

$$v_{i,\text{post}} = \frac{\mu_{i,\text{post}}(1-\mu_{i,\text{post}})}{a_i+b_i+1}$$

which is the posterior variance of π_i. The approximate credible interval is analogous to the simple confidence interval used in non-Bayesian inference (see Hauck and Anderson (1985)).

3.2. Highest posterior density credible interval

We do the inference about δ by using the exact posterior distribution, which we can obtain from the posterior distributions. Letting $f_\Delta(\delta)$ denote the posterior distribution of δ, we define a $100(1-\alpha)\%$ as any interval (cl, cu), where

$$\int_{cl}^{cu} f_\Delta(\delta)d\delta = 1-\alpha.$$

If for any $\delta_1 \in (cl, cu)$ and $\delta_2 \notin (cl, cu)$, $f_\Delta(\delta_1) \geq f_\Delta(\delta_2)$, then we call (cl, cu) a $100(1-\alpha)\%$ HPD credible interval (see, for example, Box and Tiao (1992)).

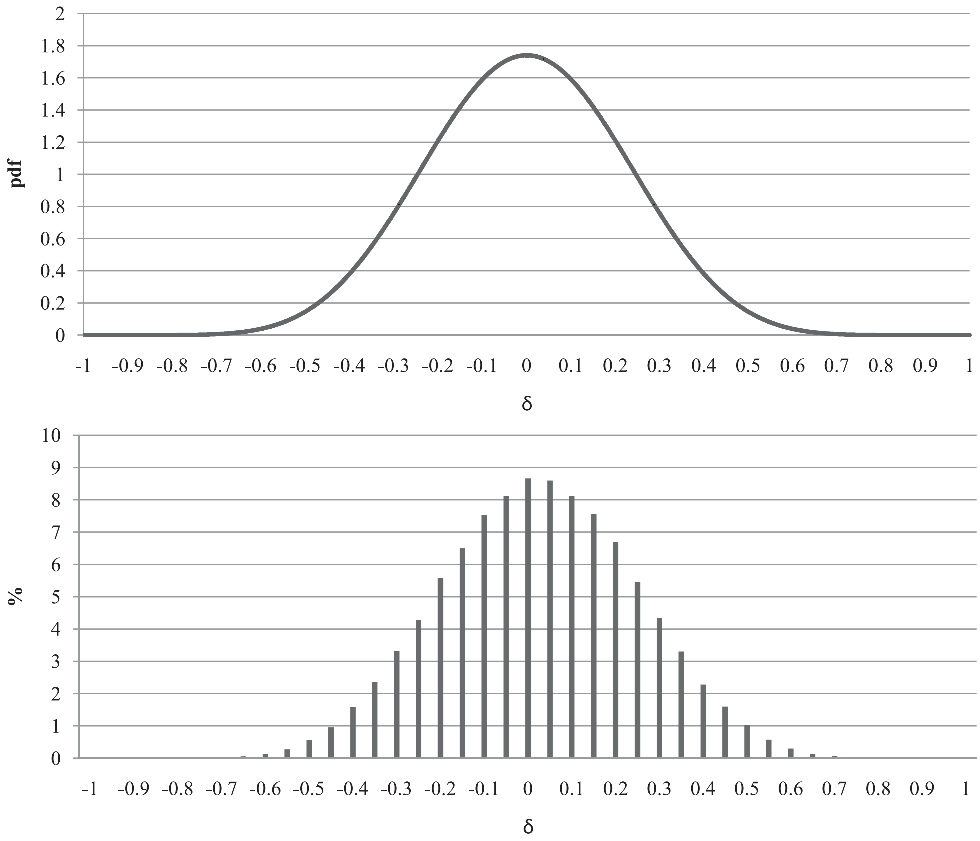

Figure 1. Posterior pdf and Histogram for difference of $\pi_{1,\text{post}} \sim \text{Beta}(5,4)$ $\pi_{2,\text{post}} \sim \text{Beta}(5,4)$.

4. Some numerical results

In this section, we show the posterior pdf of the difference, and we calculate the HPD credible interval and the approximate credible interval. Then we give some comparisons with the credible interval and the confidence interval.

We consider three cases. In case 1, we assume the difference of the mean of the posterior distribution is zero. The prior distribution is assumed to be non-informative prior. Let sampling results be $n_1 = 7$, $x_1 = 4$, $n_2 = 7$ and $x_2 = 4$. The posterior density of π_1 and π_2 are the $\text{Beta}(5,4)$. We show the posterior pdf in Figure 1. Moreover, to show that the calculated distribution is certain, we show the bar chart using 100,000 random numbers arising from posterior distribution. We are able to show that the shape of the posterior pdf and the bar chart are almost equal. Because the difference in both groups is assumed to be zero, the shape of posterior pdf shows symmetry and centers on 0. In case 2, the difference of the mean of the posterior distribution is assumed to be 0.222. The prior distribution is assumed to be non-informative prior. Let sampling results be $n_1 = 7$, $x_1 = 4$, $n_2 = 7$ and $x_2 = 2$. The posterior density of

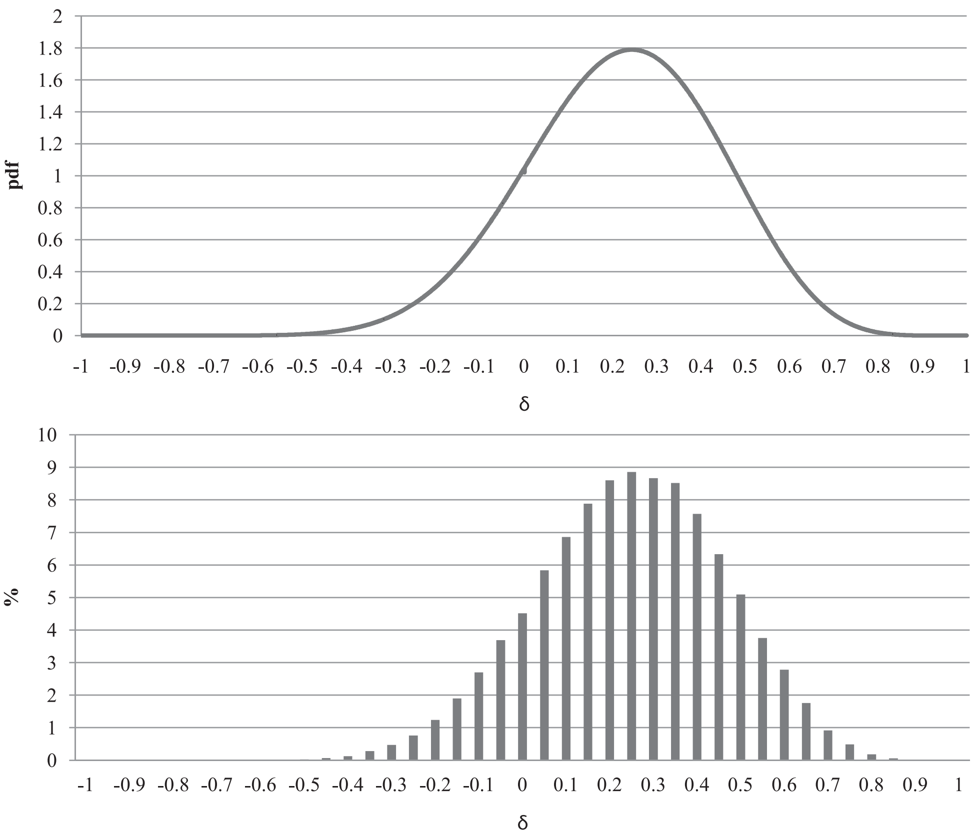

Figure 2. Posterior pdf and Histogram for difference of $\pi_{1,\text{post}} \sim \text{Beta}(5,4)$ $\pi_{2,\text{post}} \sim \text{Beta}(3,6)$.

π_1 and π_2 are the Beta$(5,4)$ and Beta$(3,6)$, respectively. We show the posterior pdf in Figure 2. The difference of the mean of the posterior pdf is 0.222. Hence the shape of the posterior pdf centers on 0.222. In case 3, the difference of the mean of the posterior distribution is assumed to be 0.667. The prior distribution is assumed to be non-informative prior. Let sampling results be $n_1 = 7$, $x_1 = 7$, $n_2 = 7$ and $x_2 = 1$. The posterior density of π_1 and π_2 are the Beta$(8,1)$ and Beta$(2,7)$, respectively. We show the posterior pdf in Figure 3. The shape of the posterior pdf has a large bias and shows asymmetry.

Next, we show the HPD credible interval and the approximate credible interval. In Table 1, we show the numerical example with the small sample. We show the numerical example with the large sample in Table 2. In addition, we took up some actual examples in Table 3. The first example is a vaccine study (Chan and Zhang (1999)). The second example was taken from StatXact (Cytel Software (1995)), and the third example was from an influenza vaccine study reported by Fries *et al.* (1993). Moreover, we calculate the Exact confidence interval and the Newcombe Hybrid confidence interval as references. Kawasaki (2010) explains the calculation of the first interval and Newcombe (1998) the second. In Table 1,

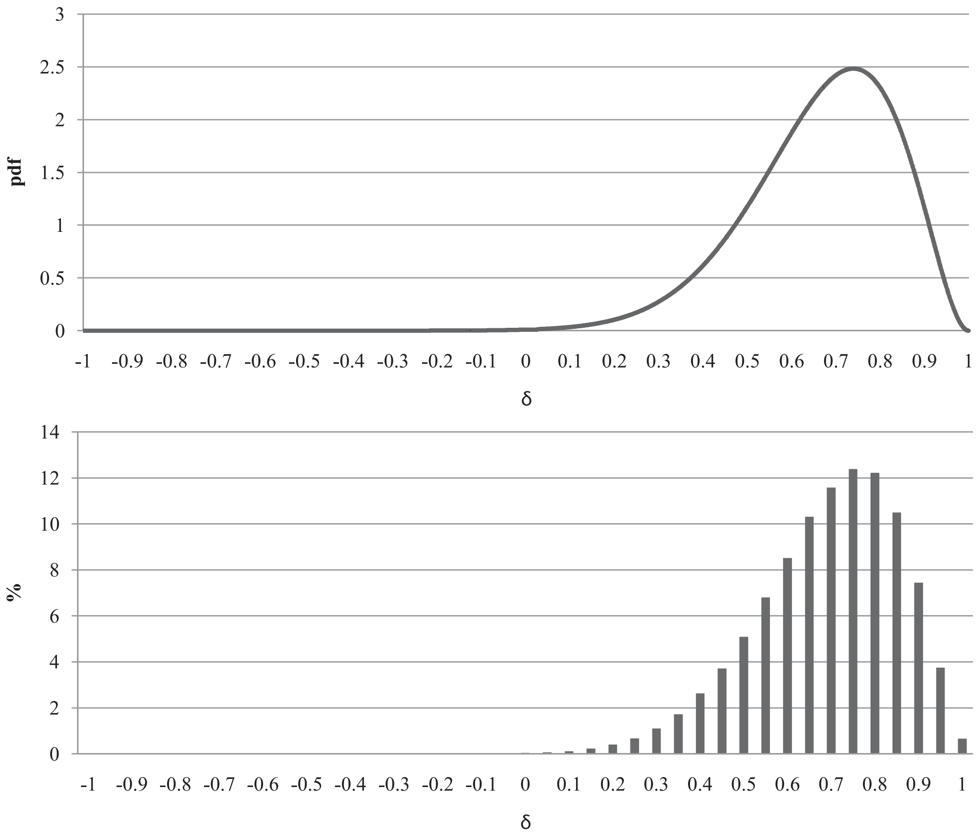

Figure 3. Posterior pdf and Histogram for difference of $\pi_{1,\text{post}} \sim \text{Beta}(8,1)$ $\pi_{2,\text{post}} \sim \text{Beta}(2,7)$.

the approximate credible interval exceeds the limit value 1. The approximate credible interval is calculated based on the normal approximation. Therefore, its result is not good when there is large bias in the distribution. In Tables 1 and 3, the HPD credible interval always has a narrow interval length, compared with the approximate credible interval. We think that it makes no sense to compare the credible interval with the confidence interval. However, the HPD credible interval and the approximate credible interval always have a narrow interval length compared with the Exact confidence interval. In Table 2, we confirmed that all intervals were almost the same. Finally, the widths in the HPD credible interval and the Newcombe Hybrid interval appear to be almost the same. In general, the Newcombe Hybrid interval is recommended. Hence, we believe that the HPD credible interval will also be recommended.

Table 1. 95% credible interval and confidence interval with small sample.

Sampling	Interval	Lower		Upper
$n_1 = 7, x_1 = 4$	HPD credible	-0.430	$\sim$	0.430
$n_2 = 7, x_2 = 4$	Approximate credible	-0.459	$\sim$	0.459
	Newcombe Hybrid	-0.420	$\sim$	0.420
	Exact	-0.541	$\sim$	0.541
$n_1 = 7, x_1 = 4$	HPD credible	-0.200	$\sim$	0.636
$n_2 = 7, x_2 = 2$	Approximate credible	-0.225	$\sim$	0.670
	Newcombe Hybrid	-0.193	$\sim$	0.624
	Exact	-0.299	$\sim$	0.746
$n_1 = 7, x_1 = 7$	HPD credible	0.348	$\sim$	0.950
$n_2 = 7, x_2 = 1$	Approximate credible	0.326	$\sim$	1.007
	Newcombe Hybrid	0.345	$\sim$	0.974
	Exact	0.321	$\sim$	0.996

Table 2. 95% credible interval and confidence interval with large sample.

Sampling	Interval	Lower		Upper
$n_1 = 70, x_1 = 40$	HPD credible	-0.160	$\sim$	0.160
$n_2 = 70, x_2 = 40$	Approximate credible	-0.162	$\sim$	0.162
	Newcombe Hybrid	-0.160	$\sim$	0.160
	Exact	-0.164	$\sim$	0.164
$n_1 = 70, x_1 = 40$	HPD credible	0.123	$\sim$	0.432
$n_2 = 70, x_2 = 20$	Approximate credible	0.123	$\sim$	0.433
	Newcombe Hybrid	0.122	$\sim$	0.429
	Exact	0.123	$\sim$	0.437
$n_1 = 70, x_1 = 70$	HPD credible	0.746	$\sim$	0.921
$n_2 = 70, x_2 = 10$	Approximate credible	0.746	$\sim$	0.921
	Newcombe Hybrid	0.744	$\sim$	0.921
	Exact	0.753	$\sim$	0.929

5. Conclusion

In this article, we propose the expression of a posterior pdf for the difference between two independent binominal proportions. We have assumed that the expression does not contain integration. The expression of the posterior pdf contains the hypergeometric series. The hypergeometric series is usually a built-in function of general software. Therefore, we can easily calculate the posterior pdf.

We calculate the HPD credible interval by using this expression. The HPD credible interval and the approximate credible interval always have narrow inter-

Table 3. 95% credible interval and confidence interval for the illustrative examples.

Sampling	Interval	Lower		Upper
$n_1 = 18, x_1 = 17$	HPD credible	0.042	~	0.543
$n_2 = 18, x_2 = 11$	Approximate credible	0.048	~	0.552
	Newcombe Hybrid	0.059	~	0.563
	Exact	−0.019	~	0.629
$n_1 = 10, x_1 = 9$	HPD credible	−0.005	~	0.671
$n_2 = 10, x_2 = 5$	Approximate credible	−0.020	~	0.686
	Newcombe Hybrid	−0.002	~	0.676
	Exact	−0.086	~	0.762
$n_1 = 15, x_1 = 12$	HPD credible	−0.013	~	0.595
$n_2 = 15, x_2 = 7$	Approximate credible	−0.017	~	0.605
	Newcombe Hybrid	−0.009	~	0.587
	Exact	−0.056	~	0.654

val length compared with the exact confidence interval. We effortlessly calculate the HPD credible interval by using the expression. Therefore, the biostatistician facilitates statistical inference by employing the HPD credible interval, thus furthering medical progress.

Acknowledgements

We are very grateful to the editor and two referees for their invaluable comments. The authors would also like to Dr. F. Zhang for his helpful comments and suggestion.

References

Agresti, A. and Caffo, B. (2000). Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures, *American Statistician*, **54**, 280–288.

Agresti, A. and Min, Y. (2001). On small-sample confidence intervals for parameters in discrete distributions, *Biometrics*, **57**, 963–971.

Agresti, A. and Min, Y. (2005). Frequentist performance of Bayesian confidence intervals for comparing proportions in 2 × 2 contingency tables, *Biometrics*, **61**, 515–523.

Andres, M. A. and Tejedor, H. I. (2003). Unconditional confidence interval for the difference between two proportions, *Biom. J.*, **45**, 426–436.

Bailey, W. N. (1972). *Generalized Hypergeometric Series*, Hafner, New York,

Berger, R. L. and Boos, D. D. (1994). P values maximized over a confidence set for the nuisance parameter, *J. Am. Stat. Assoc.*, **89**, 1012–1016.

Box, G. E. P. and Tiao, G. C. (1992). *Bayesian Inference in Statistical Analysis*, Wiley, New York.

Brown, L. and Li, X. (2005). Confidence intervals for two sample binomial distribution, *J. Stat. Plann. Inference*, **130**, 359–375.

Chan, I. S. F. and Zhang, Z. (1999). Test-based exact confidence intervals for the difference of two binomial proportions, *Biometrics*, **55**, 1202–1209.

Cornfield, J. (1966). A Bayesian test of some classical hypotheses—with applications to sequential clinical trials, *J. Am. Stat. Assoc.*, **61**, 577–594.

Cytel Software (1995). *StatXact Version 3*, Cytel Software, Cambridge.

Draper, N. and Guttman, I. (1971). Bayesian estimation of the binomial parameter, *Technometrics*, **13**, 667–673.

Fries, L. F., Dillon, S. B., Hildreth, J. E. K., Karron, R. A., Funkhouser, A. W., Friedman, C. J., Jones, C. S., Culleton, V. G. and Clements, M. L. (1993). Safety and immunogenicity of a recombinant protein influenza A vaccine in adult human volunteers and protective efficacy against wild-type HlNl virus challenge, *The Journal of Infectious Diseases*, **167**, 593–601.

Hauck, W. and Anderson, S. (1985). On testing for bioequivalence, *Biometrics*, **41**, 560–563.

Hasemi, L., Nandram, B. and Goldberg, R. (1997). Bayesian analysis for a single 2×2 table, *Stat. Med.*, **16**, 1311–1328.

Howard, J. V. (1998). The 2×2 table: a discussion from a Bayesian viewpoint, *Stat. Sci.*, **13**, 351–367.

Irony, T. Z. and Pereira, C. A. B. (1986). Exact tests for equality of two proportions: Fisher vs Bayes, *J. Stat. Comput. Sim.*, **25**, 93–114.

Kawasaki, Y. (2010). Comparison of exact confidence intervals for the difference between two independent binomial proportions, *Advances and Applications in Statistics*, **15**, 157–170.

Kawasaki, Y., Midorikawa, S. and Miyaoka, E. (2008). Comparisons of confidence intervals for the difference between two independent binomial proportions, Proceeding of the International Association for Statistical Computing, 829–838.

Newcombe, R. G. (1998). Interval estimation for the difference between independent proportions: Comparison of eleven methods, *Stat. Med.*, **17**, 873–890.

Novick, M. R. and Grizzle, J. E. (1965). A Bayesian approach to the analysis of data from clinical trials, *J. Am. Stat. Assoc.*, **60**, 81–96.

Pan, W. (2002). Approximate confidence intervals for one proportion and difference of two proportions, *Comput. Stat. Data Anal.*, **40**, 143–157.

Pham-Gia, T. and Turkkan, N. (2008). Exact small sample size determination in the Bayesian analysis of a 2×2 contingency table with dependent proportions, *Statistics*, **42**, 127–147.

Santner, T. J. and Snell, M. K. (1980). Small-sample confidence intervals for $p1 - p2$ and $p1/p2$ in 2×2 contingency tables, *J. Am. Stat. Assoc.*, **75**, 386–394.

Seneta, E. (1994). Carl lieebermeister's Hypergeometric tails, *Hist. Mat.*, **21**, 453–462.

Seneta, E. and Phipps, M. C. (2001). On the comparison of two observed frequencies, *Biom. J.*, **43**, 23–43.

Srivastava, H. M. and Karlsson, Per W. (1985). *Multiple Gaussian Hypergeometric Series*, Halsted Press, New York.

Yanakawa, T. and Bian, Q. (2001). Confidence interval in binomial distributions when 0 is observed, *The Biometric Society of Japan*, **76**, 117–120.