A new approximation of the posterior distribution of the log-odds ratio

Marc Fredette[*]   Jean-François Angers[†‡]

CRM-2643

December 1999

[*]Dept. of Statistics and Actuarial Sciences, University of Waterloo, Waterloo, ON; fredette@xxx.uwaterloo.ca

[†]Dép. de mathématiques et de statistique, Université de Montréal, Montréal, QC H3C 3J7; angers@dms.umontreal.ca

[‡]This research has been partially supported by NSERC of Canada.

Abstract

In this paper, the posterior density of the log-odds ratio is studied. It is assumed that the observations have a multinomial distribution and that the prior on the multinomial parameters is a Dirichlet density. Several approximations currently available are reviewed. When the prior parameters of the Dirichlet density are integers, it is shown that the posterior moments can be computed exactly. A new approximation, similar to the Edgeworth expansion is also proposed. Using a numerical example, the different methods of approximation of posterior density are compared.

Résumé

Dans cet article, nous étudions la densité *a posteriori* du rapport de cote. Nous faisons l'hypothèse que les observations proviennent d'une densité multinomiale et que les paramètres de celle-ci ont une loi *a priori* de Dirichlet. Nous commençons par revoir les approximations de la densité *a posteriori* employées habituellement dans la littérature. Par la suite, lorsque les paramètres de la densité de Dirichlet sont entiers, nous montrons que les moments *a posteriori* peuvent être calculés exactement. Une nouvelle approximation, similaire au développement d'Edgeworth, est aussi présentée. Nous terminons en comparant les différentes méthodes d'approximation à l'aide d'un exemple numérique.

1 Introduction

The assessment of the independence of two categorical variables with only two possible values is a very important problem in applied statistics. This problem arises frequently in area such as epidemiology, psychiatry, clinical trial, *etc.* In a Bayesian setting, it is usually assumed that the components of the log-odds ratio have a Dirichlet prior. Under such model, the asymptotic behavior of the posterior has been studied by several authors such as Lindley (1964), Bloch and Watson (1967), Altham (1969), Aitchison and Shen (1980), Latorre (1982) and O'Hagan (1994, Chapter 10).

In the first section of this paper, several approximations of the posterior mean and variance of the log-odds ratio are reviewed. If the prior parameters are integers, it is shown that the posterior moments are given in an analytic form. Several new approximations are then given. The first type is based on the normal distribution. The second one, discussed in Section 3, is based on the Latorre's expansion (*cf.* Latorre, 1982) and the last one is similar to the Edgeworth expansion. Finally, in Section 4, a numerical example is given to illustrate the different approximations.

2 Normal approximations

Let $\underset{\sim}{N}$ be a multinomial random vector with parameters $(n, \underset{\sim}{\theta})$ where $n = \sum_{j=1}^{p} N_j$ is a fixed known number. However, $\underset{\sim}{\theta}$ is unknown and it takes values in the simplex $\mathcal{S}^p = \{\theta_j \geq 0,\ i = 1, \ldots, p;\ \sum_{j=1}^{p} \theta_j = 1\}$. It is assumed that the prior on $\underset{\sim}{\theta}$ is a Dirichlet with parameter $\underset{\sim}{a}$, denoted by $\underset{\sim}{\theta} \sim D(\underset{\sim}{a})$. Hence

$$\pi(\underset{\sim}{\theta}) \quad = \quad \begin{cases} B(\underset{\sim}{a})^{-1} \prod_{j=1}^{p} \theta_j^{a_j - 1} & \text{if } \underset{\sim}{\theta} \in \mathcal{S}^p, \\ 0 & \text{otherwise,} \end{cases}$$

where $\underset{\sim}{a} = (a_1, \ldots, a_p) \in (\mathbb{R}^+)^p$ and $B(\underset{\sim}{a}) = \prod_{j=1}^{p} \Gamma(a_j) \,/\, \Gamma(\sum_{j=1}^{p} a_j)$.

Several values of $\underset{\sim}{a}$ have been used in the litterature (*cf.* Walley, 1997). For example, if one chooses $a_j \equiv 1$, then the prior corresponds to the uniform density on $\mathcal{S}^p$ while if one chooses $a_j \equiv 1/2$, the prior corresponds to the Jeffrey's noninformative prior.

The main advantage of the Dirichlet priors is that they form a conjugate family. Consequently, to obtain the posterior density one has only to replace a_j by $a_j + n_j$ for all j. To ease the notation, the prior and the posterior parameters will both be denoted by a_j unless it is important to distinguish the prior from the posterior density.

Other priors can be used to model the *a priori* information on $\underset{\sim}{\theta}$. Aitchison and Shen (1980) used the logistic-normal density. Instead of using a prior on $\underset{\sim}{\theta}$, Albert and Gupta (1983a) suggests to put a prior on a reparametrisation of $\underset{\sim}{\theta}$, *i.e.* on $(\theta_1 + \theta_2, \theta_1 + \theta_3, \theta_1 - [\theta_1 + \theta_2][\theta_1 + \theta_3], \theta_4)$. Albert and Gupta (1982, 1983b) and Good (1976) proposed hierarchical model for $\underset{\sim}{\theta}$ based on the Dirichlet density.

2.1 Approximation of the moments

Let ϕ be the log-odds ratio. Hence, ϕ can be written as $\phi = \sum_{i=1}^{4} c_j \log \theta_j$ where $\underset{\sim}{c} = (1, -1, -1, 1)$. The normal density is often used to approximate the posterior of ϕ. O'Hagan (1994, Chapter 10) shows that if all the a_j are large, the posterior of ϕ can be approximated using a normal density with mean and variance given by:

$$\mu = \sum_{j=1}^{4} c_j \left(\log(a_j) - \frac{1}{2a_j} \right), \tag{1}$$

$$\sigma^2 = \sum_{j=1}^{4} \frac{1}{a_j}. \tag{2}$$

Other approximations for the moments have been proposed. Lindley (1964) and Bloch and Watson (1967) suggest the following approximations:

$$\mu = \sum_{j=1}^{4} c_j \log(a_j - k_j),$$

$$\sigma^2 = \sum_{j=1}^{4} \frac{1}{a_j - k_j},$$

where $k_j = \frac{1}{2}$ in Lindley (1964) while $k_j = \frac{1}{2} - \frac{1}{24a_j}$ in Bloch and Watson (1967). It can be shown that the approximated moments are $O(a_{min}^{-2})$ where $a_{min} = \min_{1 \le j \le 4}(a_j)$, for Lindley (1964) and O'Hagan (1994, Chapter 10). For the method proposed in Bloch and Watson (1967), the approximation error is $O(a_{min}^{-3})$.

2.2 Analytic evaluation of moments

In order to be able to compute the exact moments of ϕ, its characteristic function is required. It can be shown to be equal to

$$\varphi_\phi(t) = \mathbb{E}[e^{it\phi}] = \prod_{j=1}^{4} \frac{\Gamma(a_j + itc_j)}{\Gamma(a_j)},$$

where $i = \sqrt{-1}$ (*cf.* Blotch and Watson, 1967). In the next theorem, an analytic expression for the moments of ϕ is given. Its proof is given in the appendix.

Theorem 1. *Let ϕ be the log-odds ratio from a 2×2 contingency table. Let $\zeta(k, a)$ be the Rieman's zeta function, that is $\zeta(k, a) = \sum_{n=a}^{\infty} 1/(a+n)^k$. If $\underset{\sim}{\theta}$ has a Dirichlet prior with parameter $\underset{\sim}{a}$, then*

$$\mathbb{E}[\phi^k] = \sum_{l=1}^{k} \binom{k-1}{l-1} b_l(\underset{\sim}{a}) \mathbb{E}[\phi^{k-l}],$$

where $b_l(\underset{\sim}{a}) = (l-1)! \sum_{j=1}^{4} (-c_j)^l \zeta(l, a_j)$.

Consequently,

$$\mathbb{E}[\phi] = b_1(\underset{\sim}{a}) \text{ and } \mathbb{V}[\phi] = b_2(\underset{\sim}{a}). \tag{3}$$

Using Abramovitz and Stegun (1965, Section 6.4), it can be shown that the mean and the variance proposed in O'Hagan (1994, Chapter 10) (*cf.* equations (1) and (2)) are indeed equal to the first terms of the asymptotic expansion of $b_1(\underset{\sim}{a})$ and $b_2(\underset{\sim}{a})$ respectively.

If $a_j \in \mathbb{N}^*$ for $j = 1, 2, 3$ and 4, then the $b_l(\underset{\sim}{a})$ can be computed exactly. Let

$$\gamma(a,b) = \begin{cases} \operatorname{sign}(b-a)\sum_{l=l_-}^{l_+} 1/l^k & \text{if } l_- \neq l_+, \\ 0 & \text{otherwise,} \end{cases}$$

where $l_- = \min(a, b)$ and $l_+ = \max(a, b)$. Using the Parseval's identity, it can then be shown that,

$$b_l(\underset{\sim}{a}) = \begin{cases} (l-1)!\,[\gamma(a_2,a_1)+\gamma(a_3,a_4)] & \text{if } l \text{ is odd,} \\ (l-1)!\left[\frac{2(2\pi)^l(-1)^{\frac{l+2}{2}}B_l}{l!} - \sum_{j=1}^4 \gamma(1,a_j-1)\right] & \text{if } l \text{ is even,} \end{cases}$$

where B_l is the l^{th} Bernouilli's number (*cf.* Edwards, 1974, Chapter 1).

In Aitchison and Shen (1980), it is shown that the values of μ and σ^2 which minimise the Kullback's distance between the density of ϕ and a normal density are the mean and the variance of ϕ given by equation (3). Since, we are now able to compute the exact moment of ϕ when the a_j's parameters are integer, the Aitchison and Shen (1980) method can be used to approximate the posterior density of ϕ. It should be noted that, because of the constraint on the a_j's, the Jeffreys prior cannot be used.

3 A new approximation

In Latorre (1982), it is shown that the exact posterior density of ϕ can be written as an infinite sum. This density is given in the next theorem.

Theorem 2. *Let ϕ be the log-odds ratio. If the density of $\underset{\sim}{\theta}$ is a Dirichlet with parameter $\underset{\sim}{a}$, then the density of ϕ is given by*

$$\pi_\Phi(\phi) = \begin{cases} Ke^{\phi a_1}F(a_1+a_3, a_1+a_2, a_+; 1-e^{\phi}) & \text{if } \phi \le 0, \\ Ke^{-\phi a_2}F(a_1+a_2, a_2+a_4, a_+; 1-e^{-\phi}) & \text{if } \phi > 0, \end{cases}$$

where $K = B(a_1+a_2, a_3+a_4)/[B(a_1,a_3)B(a_2,a_4)]$ and

$$F(a,b,c;x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{i=0}^{\infty}\frac{\Gamma(a+i)\Gamma(b+i)x^i}{\Gamma(c+i)i!}.$$

In Latorre (1982), the exact cumulative distribution is also given. However, it is more complicated to compute since it involves the evaluation of incomplete beta function. (Note that, in Latorre (1982), $I_t(a_2, j+1)$ should be replaced by $I_{t^{-1}}(a_2, j+1)$ for $t > 1$.)

3.1 An alternative approximation

Using a similar argument as in the Edgeworth expansion (see Hall, 1992, Chapter 2), a new approximation for the posterior density of ϕ can be found. It is given in the next theorem whose proof can be found in the appendix.

Theorem 3. *Let ϕ be the log-odds ratio and suppose that $\underset{\sim}{\theta}$ has a Dirichlet prior with parameter $\underset{\sim}{a}$. Let also $Z = b_2(\underset{\sim}{a})^{-1/2}[\phi - b_1(\underset{\sim}{a})]$. If*

$$g(t) = \exp\left\{\sum_{k=3}^{m}\frac{b_k(\underset{\sim}{a})}{k!}\left(\frac{it}{\sqrt{b_2(\underset{\sim}{a})}}\right)^k + \frac{b_{m+1}(\underset{\sim}{a}+itp_1\underset{\sim}{c})}{(m+1)!}\left(\frac{it}{\sqrt{b_2(\underset{\sim}{a})}}\right)^{m+1}\right\}, \tag{4}$$

where $0 < p_1 < b_2(\underset{\sim}{a})^{-1/2}$, *then the posterior density of* Z *can be written as*

$$\pi_Z(z) = \sum_{k=0}^{m} D_k(\underset{\sim}{a}) M_k(z) + E_m(z),$$

where

$$\begin{aligned} D_k(\underset{\sim}{a}) &= \frac{i^k g^{(k)}(0)}{\sqrt{2\pi} k!}, \\ M_k &= \frac{\partial^k}{\partial z^k}(e^{-z^2/2}), \\ E_m(z) &= \frac{1}{\sqrt{2\pi}(m+1)!} \int_{-\infty}^{\infty} e^{-itz} f(t) g^{(m+1)}(p_2 t) t^{m+1} dt, \end{aligned}$$

where $0 < p_2 < b_2(\underset{\sim}{a})^{-1/2}$ *and* $f(t)$ *represents the standard normal density.*

Neglecting the error term $E_m(z)$, the new approximation is given by

$$\widehat{\pi}_{(\Phi,m)}(\phi) = \frac{\widehat{\pi}_{(Z,m)} \left(b_2(\underset{\sim}{a})^{-1/2} [\phi - b_1(\underset{\sim}{a})] \right)}{b_2(\underset{\sim}{a})^{1/2}},$$

where

$$\widehat{\pi}_{(Z,m)}(z) = \sum_{k=0}^{m} D_k(\underset{\sim}{a}) M_k(z). \tag{5}$$

It should be noted that, this approach has been previously mentioned in Bloch and Watson (1967). However, it was introduced with approximation of the moments of ϕ rather than the exact form given in Theorem 1. Using equation (5), the posterior distribution can be approximated by

$$\widehat{\Pi}_{(Z,m)}(z) = \Phi(z) + \sum_{k=1}^{m} D_k(\underset{\sim}{a}) M_{k-1}(z).$$

The evaluation of this approximation and of equation (5) can be done using the formulae given in the next subsection. Note that only the $M_k(z)$'s functions depend on z. In the next two subsections, the components of equation (5) are studied.

3.2 The $D_k(\underset{\sim}{a})$ and $M_k(z)$ coefficients

In this subsection, recursive equations to compute $D_k(\underset{\sim}{a})$ and $M_k(z)$ are given. The rate of convergence of $D_k(\underset{\sim}{a})$ as $a_{min} \to \infty$ is also discussed.

Even though, it is impossible to write down analytically the $M_k(z)$ function, a recursive equation to evaluate it is available. This is due to the fact that $M_k(z)$ can be written using the Hermite's polynomials.

If $H_k(z)$ represents the k^{th} Hermite polynomial, then

$$M_k(z) = H_k(z/\sqrt{2})(-\sqrt{2})^{-k} e^{-z^2/2}. \tag{6}$$

Using the recursive equation for the Hermite polynomials, (*cf.* Gradshteyn and Ryzhik, 1980, equation 8.952), and equation (6), the following recursive equation for the $M_k(z)$ can be obtained

$$M_k(z) = \begin{cases} e^{-z^2/2} & \text{if } k = 0, \\ -z e^{-z^2/2} & \text{if } k = 1, \\ -z M_{k-1}(z) - (k-1) M_{k-2}(z) & \text{if } k \geq 2, \end{cases}$$

if $|z| < \infty$.

Even though it is more complicated, a recursive equation exists to compute the $D_k(\underset{\sim}{a})$ and it is given in the next proposition. The proof of this proposition is similar to the one of Theorem 1, so it is omitted but it can be found in Fredette (1999).

Proposition 1. *Let* $D_k(\underset{\sim}{a}) = i^k g^{(k)}(0)/[\sqrt{2\pi}k!]$, *where* $g(t)$ *is given in equation (4). If* $k \leq m$, *then*

$$D_k(\underset{\sim}{a}) = \begin{cases} \frac{1}{\sqrt{2\pi}} & \text{if } k = 0, \\ 0 & \text{if } k = 1, 2, \\ \frac{(-1)^l}{k} \sum_{l=3}^{k} \frac{D_{k-l}(\underset{\sim}{a}) b_l(\underset{\sim}{a})}{(l-1)!\sigma^l} & \text{if } 3 \leq k \leq m, \end{cases}$$

where $\sigma^2 = b_2(\underset{\sim}{a})$.

The proposed approximation depends on the a_j's only through $D_k(\underset{\sim}{a})$. Hence, all $D_k(\underset{\sim}{a})$ but $D_0(\underset{\sim}{a})$ should go to 0 as $a_{min} \to \infty$. If this is true, then

$$\lim_{a_j \to \infty} \widehat{\pi}_{(Z,m)}(z) = D_0 M_0(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2},$$

which is consistent with the normal approximations discussed in Section 2. In Fredette (1999), it is shown that

$$\lim_{a \to \infty} a^{l-1} \zeta(l, a) = 1/[l-1].$$

Using this asymptotic property of $\zeta(l, a)$, the following proposition can be shown.

Proposition 2. *If* $k = 3q + r$ *where* $q \geq 1$ *and* $r \in \{0, 1, 2\}$, *then* $D_k(\underset{\sim}{a}) = O(a_{min}^{-(q+r)/2})$.

It should be noted that the normal approximation neglects the terms starting at $D_3(\underset{\sim}{a})$ which is $O(a_{min}^{-1/2})$.

3.3 The error term

In this subsection, the asymptotic behavior of the error term, $E_m(z)$ is investigated and it is given in the next proposition.

Proposition 3. *Let*

$$E_m(z) = \frac{\int_{-\infty}^{\infty} e^{-itz} f(t) g^{(m+1)}(p_2 t) t^{m+1} dt}{\sqrt{2\pi}(m+1)!},$$

where $f(t)$ *is the standard normal p.d.f.,* $0 < p_2 < 1/\sigma$, $\sigma = b_2(\underset{\sim}{a})^{1/2}$ *and where* $g(t)$ *is given by equation (4). Then*

$$|E_m(z)| \leq \mathcal{O}(a_{min}^{-(m-1)/2}).$$

The proof of this proposition is mostly technical and depends heavily on the choice of m. It can be found in Fredette (1999) for the case of $m \bmod 4 = 2$.

4 Numerical example

In this section, a numerical example is presented. Using this example, the different approximations given in this paper are illustrated. The data come from Chin *et al.* (1961). Among 18 inhabitants of Des Moines, Iowa, between 15 and 19 years of age, suffering from poliomyelitis, one is interested to know if the Salk's vaccine increases or not the probability of being paralysed. The results of this study are given in Table 1.

Put Table 1 around here

Let (N_1, N_2, N_3, N_4) be a multinomial random vector with parameter $\underset{\sim}{\theta}$ and suppose that n is kwown. Since we have no prior information about $\underset{\sim}{\theta}$, the prior is chosen to be a Dirichlet density with $a_j \equiv 1$.

Using these data, we computed the posterior mean and variance along with the 95% credibility region for each of the approximation methods proposed in Sections 2 and 3. Note that the credibility regions used are not HDP but they are defined to be on the form (a, b) where

$$\int_{-\infty}^{a} \pi_\Phi(\phi) d\phi = \int_{b}^{\infty} \pi_\Phi(\phi) d\phi = \frac{\alpha}{2} = 0.025.$$

The results are given in Table 2. We also simulated 10 000 values of ϕ from a Dirichlet density with parameters $n_j + 1$ for $j = 1, 2, 3$ and 4. From these values, we computed its mean and variance and a 95% empirical credible interval. These results are also given in Table 2. From this table, one can see that the new approximation and the normal approximation with exact moments give better estimates of the moments than the other methods. The O'Hagan (1994) approximation underestimates the variance. For all the other techniques, the obtained credibility regions are similar. Note that 100 000 terms in the Latorre's expansion were required in order to compute the credibility region with a precision of three decimal places.

In Figure 1, different approximations of the posterior density of ϕ are given along with the histogram of the simulated values of ϕ. In order to make the graphic easier to interpret, only the Bloch and Watson (1967) approximation, the normal approximation with the exact moments, the Latorre (1982) approximation and the one given in Section 3 are presented. The O'Hagan (1994) and Lindley (1964), being similar to the Bloch and Watson (1967), are omitted. From Figure 1, it can be seen that the proposed approximation is closed to the Latorre (1982) "exact" density (computed with 1000 terms). In Figure 2, one can see that the number of terms required to obtain a good approximation using the Latorre (1982) technique can be quite large. Using the proposed approximation, equivalent approximation is obtained using only 5 terms.

Put Table 2 around here

Put Figures 1 and 2 around here

5 Conclusion

In conclusion, one can say that the normal based approximation of the posterior density of ϕ are very easy to use and they approximate well the posterior when the n_j's are large. However, this approximation always provides symmetric density which might not be the case for the "true" posterior. When the n_j's are moderate or small, the normal based approximations do not do that well. One way to improve upon it is to do as in Section 3 and to increase the number of terms used in the approximation. If the a_j's are integer, the normal approximation is better if one uses the exact moments. If not, the approximation method of Bloch and Watson (1967) is the best normal approximation.

The method discussed in Latorre (1982) provides an exact form for the posterior density of ϕ. However, the convergence rate of the summation can be quite slow for some patterns of $a_j + n_j$. Given sufficient computer time, this method will always converge toward the "true" posterior. The main advantage of the approximation technique proposed in Section 3 is that a limited number of terms can be used to provide a closed approximation of the posterior density of ϕ.

Appendix

Proof of Theorem 1. Let $f(t) = \sum_{j=1}^{4} \log(\Gamma(a_j + itc_j))$. Then $f^{(k)}(t) = \sum_{j=1}^{4}(ic_j)^k \psi^{(k-1)}(a_j + itc_j)$ for $k = 1, 2, \ldots$. Using equation (8.362) of Gradshteyn and Ryzhik (1980) it can be shown that

$$\psi^{(l)}(x) = \begin{cases} \psi(1) - \sum_{k=0}^{\infty}\left(\frac{1}{x+k} - \frac{1}{k+1}\right) & \text{if } l = 1, \\ (-1)^{n+1} n! \sum_{k=0}^{\infty} \frac{1}{(x+k)^{n+1}} & \text{if } l \geq 2. \end{cases}$$

Consequently

$$f^{(l)}(t) = i^l b_l(\underset{\sim}{a} + it\underset{\sim}{c}).$$

By induction, it can be shown that the k^{th} derivative of the characteristic function of ϕ is given by

$$\frac{\partial^k}{\partial t^k}\varphi_\phi(t) = \sum_{l=1}^{k} \binom{k-1}{l-1} f^{(l)}(t)\varphi_\phi^{(k-l)}(t). \tag{7}$$

Using the previous equation, Theorem 1 directly follows.

First, let check if equation (7) is true when $k = 1$.

$$\begin{aligned} \varphi_\phi^{'}(t) &= i\sum_{j=1}^{4}\left(\prod_{l\neq j}^{4}\frac{\Gamma(a_l + itc_l)}{\Gamma(a_l)}\right)\left(\frac{\Gamma(a_j + itc_j)}{\Gamma(a_j)}\right) c_j \psi(a_j + itc_j) \\ &= i\sum_{j=1}^{4} c_j \psi(a_j + itc_j)\varphi_\phi(t). \end{aligned}$$

Now, suppose that equation (7) is true for $l = 1, 2, \ldots, k-1$, and let shows it is still true for $l = k$. The k^{th} derivative of $\varphi_\phi(t)$ is given by

$$\begin{aligned} \varphi_\phi^{(k)}(t) &= \frac{\partial}{\partial t}\left(\varphi_\phi^{(k-1)}(t)\right) \\ &= \sum_{l=1}^{k-1}\binom{k-2}{l-1} f^{(l)}(t)\varphi_\phi^{(k-l)}(t) + \sum_{l=2}^{k}\binom{k-2}{l-2} f^{(l)}(t)\varphi_\phi^{(k-l)}(t) \\ &= f^{'}(t)\varphi_\phi^{(k-1)}(t) + \sum_{l=2}^{k-1}\left\{\binom{k-2}{l-1} + \binom{k-2}{l-2}\right\} f^{(l)}(t)\varphi_\phi^{(k-l)}(t) \\ &\quad + f^{(k)}(t)\varphi_\phi(t) \\ &= \sum_{l=1}^{k}\binom{k-1}{l-1} f^{(l)}(t)\varphi_\phi^{(k-l)}(t). \end{aligned}$$

□

Proof of Theorem 3. The first step of the proof is to write the log of the characteristic function as a finite sum of $m+1$ terms. Using formula 8.326 of Gradshteyn and Ryzhik (1980), it can be shown that

$$\begin{aligned}
\log(\varphi_\phi(t)) &= \sum_{j=1}^{4} \log\left[\frac{\Gamma(a_j + itc_j)}{\Gamma(a_j)}\right] \\
&= \sum_{j=1}^{4} \log\left[\frac{a_j e^{i\psi(1)c_j t}}{a_j + ic_j t} \prod_{n=1}^{\infty} \left(\frac{e^{ic_j t/n}(a_j+n)}{a_j+n+ic_j t}\right)\right] \\
&= \sum_{j=1}^{4} \left[\sum_{n=0}^{\infty} \log\left(\frac{a_j+n}{a_j+n+ic_j t}\right) + c_j\left(\sum_{n=1}^{\infty} \frac{it}{n} + it\psi(1)\right)\right] \\
&= -\sum_{j=1}^{4}\sum_{n=0}^{\infty} \log\left(1 + \frac{ic_j t}{a_j+n}\right) \\
&= \sum_{j=1}^{4}\sum_{n=0}^{\infty} \left[\sum_{k=1}^{m} \frac{(-1)^k}{k}\left(\frac{ic_j t}{a_j+n}\right)^k + \frac{(-1)^k}{(m+1)}\left(\frac{ic_j p_1 t}{a_j + ic_j p_1 t}\right)^{m+1}\right] \\
&= \sum_{k=1}^{m} \frac{(it)^k}{k!} b_k(\underset{\sim}{a}) + \frac{(it)^{m+1}}{(m+1)!} b_{m+1}(\underset{\sim}{a} + itp_1\underset{\sim}{c}).
\end{aligned}$$

The density of Z can then be computed using the inversion theorem. Using Theorem 5.2.1 of Walker(1991, Chapter 5), it can be shown

$$\begin{aligned}
\pi_Z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itz}\varphi_Z(t)dt \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-itz} g(t) f(t) dt \\
&= \int_{-\infty}^{\infty} e^{-itz}\left(\sum_{k=0}^{m} i^{-k} D_k(\underset{\sim}{a}) t^k + \frac{g^{(m+1)}(p_2 t) t^{m+1}}{(m+1)!}\right) f(t) dt \\
&= \sum_{k=0}^{m} i^{-k} D_k(\underset{\sim}{a}) \int_{-\infty}^{\infty} e^{-itz} t^k f(t) dt \\
&\quad + \frac{1}{\sqrt{2\pi}(m+1)!}\int_{-\infty}^{\infty} e^{-itz} f(t) g^{(m+1)}(p_2 t) t^{m+1} dt \\
&= \sum_{k=0}^{m} D_k(\underset{\sim}{a})\left(\frac{\partial^k}{\partial z^k}(e^{-z^2/2})\right) + E_m(z) \\
&= \sum_{k=0}^{m} D_k(\underset{\sim}{a}) M_k(z) + E_m(z).
\end{aligned}$$

□

References

Abramovitz, M. and Stegun, L.A. (1965), *Handbook of mathematical functions* (Dover, New-York).

Aitchison, J. and Shen, S.M. (1980), Logistic-normal distributions: some properties and uses, *Biometrika* **67**, 261-272.

Albert, J.H. and Gupta, A.K. (1982), Mixtures of Dirichlet distributions and estimation in contingency tables, *The Annals of Statistics* **4**, 1261-1268.

Albert, J.H. and Gupta, A.K. (1983a), Estimation in contingency tables using prior information, *Journal of the Royal Statistical Society, Series **B***, **45**, 60-69.

Albert, J.H. and Gupta, A.K. (1983b), Bayesian estimation methods for 2×2 contingency tables using mixtures of Dirichlet distributions, *Journal of the American Statistical Association* **78**, 708-717.

Altham, P.M.E. (1969), Exact Bayesian analysis of a 2×2 contingency table and Fisher's exact significance test, *Journal of the Royal Statistical Society, Series B* **31**, 261-269.

Bloch, D.A. and Watson, G.S. (1967), A Bayesian study of the multinomial distribution, *The Annals of Mathematical Statistics* **38**, 1423-1435.

Chin, T., Marine, W., Hall, E., Gravelle, C., and Speers, J. (1961), The influence of Salk vaccination on the epidemic pattern and the spread of the virus in the community, *The American Journal of Hygiene* **73**, 67-94.

Edwards, H.M. (1974), *Rieman's Zeta function* (Samuel Eilenberg and Hyman Bass, New-York).

Fredette, M. (1999), *Estimation du rapport de cotes dans un tableau de contingence* 2×2 (Master thesis, Département de mathématiques et de statistique, Université de Montréal).

Good, I.J. (1976), On the application of symmetric Dirichlet distributions and their mixtures to contingency tables, *The Annals of Statistics* **4**, 1159-1189.

Gradshteyn, I.S. and Ryzhik, I.M. (1980), *Table of integrals, series and products* (Academic press Inc. LTD, Londres, 2nd ed.).

Hall, P. (1992), *The bootstrap and Edgeworth expansion* (Springer-Verlag, New-York).

Latorre, G. (1982), The exact posterior disribution of the cross-ratio of a 2×2 contingency table, *Journal of Statistical Computation and Simulation* **16**, 19-24.

Lindley, D.V. (1964), The Bayesian analysis of contingency tables, *The Annals of Mathematical Statistics* **35**, 1622-1643.

O'Hagan, A. (1994), *Kendall's advanced theory of statistics (Volume 2B)* (Halsted Press, New-York).

Walker, J.S. (1991), *Fast Fourier Transforms* (CRC PRESS, Boca Raton).

Walley, P. (1997), Inferences from multinomial data: learning about a bag of marbles, *Journal of the Royal Statistical Society, Series B* **58**, 3-57.

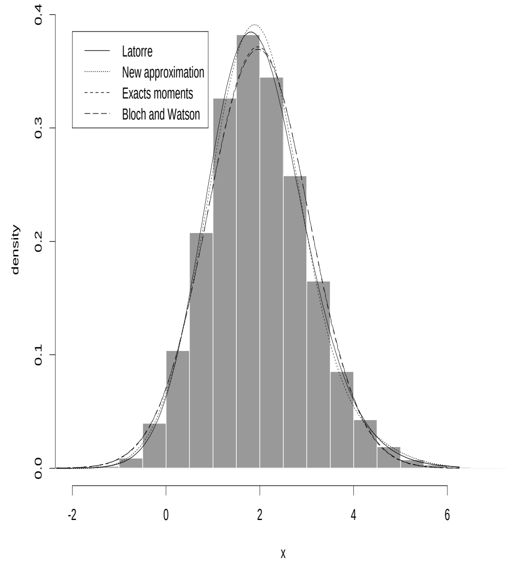

Figure 1: Approximations of the posterior density of ϕ.

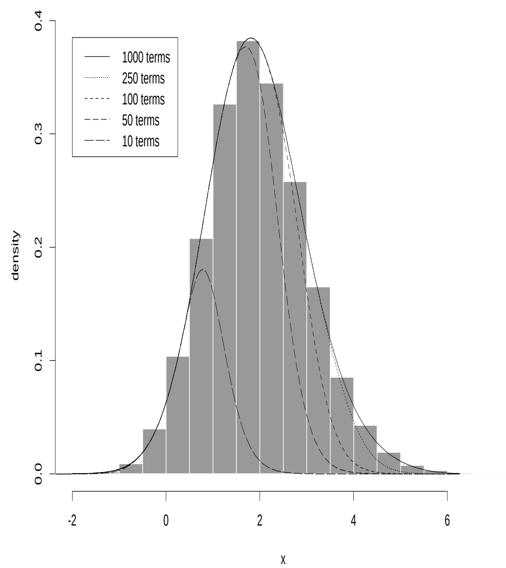

Figure 2: Convergence of the Latorre's expansion.

Table 1: Data from Chin *et al.* (1961)

	Paralysed	Nonparalysed	Total
Vaccined	7	4	11
unvaccined	1	6	7
Total	8	10	18

Table 2: Comparison of the different approximations

Methods	$\widehat{E}[\phi]$	$\widehat{V}[\phi]$	Credibility region
O'Hagan	1.939	0.968	(0.011,3.867)
Lindley	1.977	1.176	(-0.148,4.103)
Bloch and Watson	1.963	1.166	(-0.154,4.080)
Exact moments	1.960	1.153	(-0.145,4.064)
$\widehat{\pi}_{(\phi,5)}$	1.960	1.153	(-0.037,4.276)
$\widehat{\pi}_{(L,1000)}$	1.963	1.156	(-0.007,4.242)
Simulation	1.960	1.132	(0.016,4.205)