LETTER TO THE EDITOR

BAYESIAN ANALYSIS OF CASE-CONTROL STUDIES


From: Markku Nurminen and Pertti Mutanen
Department of Epidemiology and Biostatistics
Institute of occupational Health
Topeliuksenkatu 41 a A
SF-00250 Helsinki, Finland

Marshall presents the exact posterior density of the odds ratio (OR) and shows examples of its graphical presentation. The paper gives rise to some critical remarks.

Marshall writes the expression for the posterior density function of OR \( f(\text{OR}) \) (formula (3)), under the model of independent prior beta distributions for the comparison of binomial proportions ('risks') in case-referent ('control') studies with binary exposure-response data. In this context Marshall omits mention of the previous derivation of the exact posterior cumulative (integral) distribution function \( F(\text{OR}) \) with a general beta prior given by Nurminen and Mutanen\(^1\) (p. 68), who also derive a closed-form series representation of \( F(\text{OR}) \), explicitly for a rectangular prior. Marshall states (p. 1225) 'In fact, their formula is quite general', but fails to recognize that the density function \( f(\text{OR}) \) is obtainable directly from the integral form of the cumulative function \( F(\text{OR}) \) upon differentiation. We show in the Appendix that \( F'(\text{OR}) = f(\text{OR}) \) and that these functions are identical. Incidentally, the distribution function given by Nurminen and Mutanen is general also in that it is formulated in terms of a general risk parameter, either risk difference, risk ratio or risk-odds ratio. Independently, Zelen and Parker\(^2\) (formula (7)) derived an expression for the posterior distribution of \( \log(\text{OR}) \), which after reparameterization is interchangeable with formula (3) of Marshall. Yet contrary to the statement made by Zelen and Parker\(^2\) (p. 263), the posterior distribution of OR does exist in closed form as shown by Nurminen and Mutanen\(^1\) (Appendix A.3). In actual application, Nurminen and Mutanen resorted both to numerical integration procedures and to series expansions; however, these become tedious with large numbers. Fortunately, the computations can be expedited by use of the recursive relation between the beta-binomial function and the hypergeometric function.\(^3\)

Marshall proceeds to consider the approximate method based on the statistic \( \log(\text{OR}) \), and notes that the Bayesian statistic, which is due to Lindley,\(^4\) is formally equivalent to the frequentist asymptotic statistic given by Woolf.\(^5\) But Marshall neglects to state that this equivalence was earlier also pointed out by Nurminen and Mutanen\(^1\) (p. 76).

Finally, Marshall (p. 1226) obtains the formula for the prior density of OR in the particular case assuming rectangular priors for the two proportions, but again ignores that this distribution was originally described by Nurminen and Mutanen\(^1\) (p. 74) as follows: 'The prior density is a monotonically decreasing function of OR from infinity as \( OR=0 \) to zero as \( OR=\infty \); the analytic form of the density function being \( [(OR+1)\log(OR)-2(OR-1)]/(OR-1) \) when \( OR\neq 1 \) and assuming the value of 1/6 just at \( OR=1 \).'

APPENDIX: PROOF THAT \( F'(\text{OR}) = f(\text{OR}) \)

The posterior cumulative distribution function of the odds ratio (OR) assuming beta prior densities for the proportions is (Nurminen and Mutanen\(^1\))

\[
F(\text{OR}) = \int_0^1 B_1 \{ [(\text{OR})R_0/[1+(\text{OR}-1)R_0]] \} B_0(R_0) dR_0, \tag{1}
\]

where the beta functions \( B_1 \) and \( B_0 \) are the cumulative and density functions of the beta distribution for sample
LETTER TO THE EDITOR

\[ F'(OR) = \left\{ R_0(1 - R_0)[1 + (OR - 1)R_0] \right\} \beta_1 \left\{ (OR)R_0/[1 + (OR - 1)R_0] \right\} \beta_2(R_0) dR_0. \]  \hspace{1cm} (2)

Substitution of the beta densities (including the beta function B) in (2) yields (with Marshall’s notation for the constants \(e, f, g\) and \(h\) defining the prior and the data for the two samples):

\[ F'(OR) = \left\{ R_0(1 - R_0)[1 + (OR - 1)R_0] \right\} \]
\[ \times \int_0^1 B(e,f)^{-1} \{ (OR)R_0/[1 + (OR - 1)R_0] \}^{e-1} \{ 1 - (OR)R_0/[1 + (OR - 1)R_0] \}^{f-1}
\times B(g,h)^{-1} R_0^{g-1}(1 - R_0)^{h-1} dR_0
\]
\[ = B(e,f)^{-1} B(g,h)^{-1} (OR)^{-1} \int_0^1 \left\{ R_0^{g-1}(1 - R_0)^{h-1}/[1 + (OR - 1)R_0]^{f+1} \right\} dR_0. \]  \hspace{1cm} (3)

The above formula (3) is identical with Marshall’s formula (3) for \( f(OR) \).

REFERENCES


AUTHOR’S REPLY

I accept, and acknowledge, that Nurminen and Mutanen have done important work on derivation of the posterior distribution. However, the emphasis in my paper was not on derivation of the exact density, but on the log-normal approximation to it and operational use of the Bayesian method. In fact, their paper only came to my notice after I had done this work.

The posterior density can, of course, be obtained by differentiation of their formula (it would be worrying if otherwise). This, I felt, did not need mentioning. That I should note the equivalence with Woolf’s method without reference to Nurminen and Mutanen seems, to me, relatively unimportant. However, it may perhaps be more serious that Nurminen and Mutanen’s own derivation of the formula for rectangular priors was not acknowledged. I apologize for this omission.

ROGER J. MARSHALL
Department of Community Health
School of Medicine
University of Auckland
Private Bag
Auckland, New Zealand