Higher-order approximations for interval estimation in binomial settings

Ana-Maria Staicu

Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK

ARTICLE INFO

Article history:
Received 12 September 2007
Received in revised form 22 February 2009
Accepted 23 March 2009

Keywords:
Binomial distribution
Confidence intervals
Edgeworth expansion
First-order matching intervals
Modified likelihood root

In this paper we revisit the classical problem of interval estimation for one-binomial parameter and for the log odds ratio of two binomial parameters. We examine the confidence intervals provided by two versions of the modified log likelihood root: the usual Barndorff-Nielsen’s $r^*$ and a Bayesian version of the $r^*$ test statistic.

For the one-binomial problem, this work updates the findings of Brown et al. [2003. Interval estimation in exponential families. Statistica Sinica 13, 19–49; 2002. Confidence intervals for a binomial proportion and asymptotic expansion. The Annals of Statistics 30, 160–201] and Cai [2005. One-sided confidence intervals in discrete distributions. Journal of Statistical Planning and Inference 131, 63–88] to higher-order methods. For the log odds ratio of two binomial parameters we show via Edgeworth expansion that both versions of the $r^*$ statistics give confidence intervals which nearly completely eliminate the systematic bias in the unconditional smooth coverage probability. We also give expansions for the length of the confidence intervals.

© 2009 Elsevier B.V. All rights reserved.

Contents

1. Introduction ............................................................................................................................... 2
2. Confidence intervals for the binomial parameter ........................................................................................................ 2
  2.1. Accurate confidence intervals ................................................................................................. 3
  2.2. Edgeworth expansion of the coverage probability ................................................................. 4
  2.3. Expansion of the expected length ............................................................................................. 5
3. Confidence intervals for the log odds ratio parameter .................................................................... 5
  3.1. Accurate confidence intervals ................................................................................................. 6
  3.2. Comparison of the coverage probability .................................................................................. 7
  3.3. Comparison of the expected length .......................................................................................... 7
  3.4. Numerical study ...................................................................................................................... 8
    3.4.1. Example 1 ........................................................................................................................ 9
    3.4.2. Example 2 ........................................................................................................................ 9
4. Discussion ............................................................................................................................... 10
Acknowledgments ......................................................................................................................................... 10
Appendix A ............................................................................................................................................... 10
References ................................................................................................................................................. 12

E-mail address: A.Staicu@bristol.ac.uk

0378-3758/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2009.03.021

1. Introduction

We consider two basic, but very important, problems in statistical practice, namely interval estimation for a binomial proportion and comparison between two proportions with respect to one of the commonly used measures, the log odds ratio. For the first problem, the unreliability of the Wald statistic and the adequacy of alternative intervals are well established. For the second one there is little conclusive study on a best approach. In this paper we study one-sided interval estimation for the binomial proportion and for the log odds ratio of two binomial proportions obtained by inversion of highly accurate asymptotic test statistics.

For one-binomial proportion, Brown et al. (2002, 2003) give a comprehensive treatment of various two-sided confidence intervals including Wald and score, and recommend the Jeffreys intervals for their excellent coverage and length properties. Cai (2005) discusses one-sided interval estimation and proves that the score and the likelihood root intervals exhibit serious bias, while the Jeffreys intervals or second-order corrected intervals still have good coverage probability. The present work is an extension of these papers to the intervals obtained by inversion of the modified likelihood root test statistic, also known as the Barndorff-Nielsen’s $r^*$ statistic. For continuous settings, this test statistic has an asymptotically normal approximation with an error of $O(n^{-3/2})$ (Barndorff-Nielsen and Cox, 1994, Chapter 6.6) as opposed to the Wald, or score or the signed likelihood root which have $O(n^{-1/2})$ error in the normal approximation.

The properties of the $r^*$ confidence intervals are investigated analytically. For the coverage probability we use Edgeworth approximation, which contains smooth terms as well as oscillation terms, due to the lattice nature of the binomial variable. The smooth coverage probability of the one-sided $r^*$ intervals equals the nominal level with error $O(n^{-3/2})$. As well, their expected length is only slightly larger than the length of Jeffreys intervals, which are credible intervals derived from Jeffreys prior. Therefore, taking the view of Brown et al. (2002), that the smooth terms are the most important in the two-term Edgeworth expansion, the $r^*$ confidence intervals provide remarkable coverage accuracy and length properties. In addition, this accurate approach has the advantage that it naturally provides one with a test statistic in a hypothesis testing scenario for the binomial parameter (Cai, 2005).

In the setting of two independent binomials, we discuss interval estimation for the log odds ratio, using the traditional conditional approach. The score-based intervals as well as the likelihood-based intervals have been remarked to work well for large samples, but to show over conservativeness for small sample sizes (Agresti and Min, 2002). Agresti (2001) surveys various interval estimation techniques in light of the continuing controversy between the conditional and the unconditional approach. His general recommendation is to use confidence intervals formed by inversion of the test using the mid-$P$ value (Lancaster, 1961), acknowledging that ordinary $p$-values obtained with “approximate conditioning” techniques may be equally good (see also Pierce and Peters, 1999). In this article we study confidence interval estimation from the approximate conditioning perspective; see Davison and Wang (2002).

We examine theoretically and numerically the confidence intervals for the log odds ratio based on two modifications of the signed likelihood root: the Barndorff-Nielsen’s $r^*$ statistic (Barndorff-Nielsen and Cox, 1994, Chapter 6) and a Bayesian version of this, referred to as $r^*_{B}$, which is obtained by using approximate Bayesian inference with first-order matching priors. Like in the one-parameter problems, in multiparameter problems with scalar component of interest, the accuracy of the $r^*$ confidence intervals is theoretically justified only in continuous settings (Barndorff-Nielsen, 1986). In discrete settings, various works have commented on the accuracy of the $r^*$ confidence intervals. There are some references which adjust for the discrete nature of the problem: for instance, Pierce and Peters (1992) discuss the accuracy of the $r^*$ confidence limits, obtained by considerations of continuity corrections, in approximating the exact confidence limits for a scalar component of interest. Davison and Wang (2002) and Brazzale et al. (2007, Chapter 2) examine numerically the $r^*$ confidence limits without continuity corrections, which accurately approximate the limits derived from a continuous embedding model of the discrete problem. We take the latter view and investigate the smooth coverage probability of both versions of $r^*$ intervals for the log odds ratio, by a two-term Edgeworth expansion. Our findings corroborate that both methods are highly accurate, with $r^*$ statistic giving slightly wider confidence intervals, which have somewhat better coverage than $r^*_{B}$.

One important feature of this work is that all the investigated methods are used without continuity correction; ordinary $p$-values calculated in this fashion can be viewed as an approximation to the mid-$P$ value; see Pierce and Peters (1999) and Davison and Wang (2002). The difference between comparing an exact approach with a continuity corrected approximation, and comparing the mid-$P$ approach with the approximation is mentioned also in Strawderman and Wells (1998) for the case of several $2 \times 2$ tables. In general this view entails a trade-off between the control over the Type I error rate and the decrease in the conservativeness of the corresponding confidence intervals.

The paper is organized as follows. Section 2 introduces the confidence intervals for the binomial proportion based on the Barndorff-Nielsen’s $r^*$ test statistic and studies its coverage probability and expected length. Section 3 presents the intervals for the log odds ratio of two binomial proportions, that are derived using the two versions of the Barndorff-Nielsen’s $r^*$ statistic. The confidence intervals are analyzed following roughly the same methodology, but adjusted for the nuisance parameters. The results are illustrated by several examples throughout the paper. Section 4 summarizes our conclusions.

2. Confidence intervals for the binomial parameter

Let $Y_1, Y_2, ..., Y_n$ be independently, identically distributed observations from the binomial model $\text{Bin}(1, \theta)$. Exact inference for the parameter $\theta$ uses the distribution of the sum $Y = \sum_{i=1}^{n} Y_i$, where $Y \sim \text{Bin}(n, \theta)$ and the corresponding Clopper–Pearson
intervals, although unnecessarily wider, guarantee a coverage probability larger than the nominal level. There are, however, other confidence region definitions besides the Clopper–Pearson intervals: for example Sterne (1954) which are shorter than the Clopper–Pearson intervals, and have the confidence level guaranteed to be at least as advertised. An alternative approach is to use the so-called mid-$P$-value, $p_{\text{mid}}(\theta) = \Pr(Y \leq y - 1; \theta) + 0.5 \Pr(Y = y; \theta)$, which may be considered as a continuity correction to the “exact” $p$-value function, $p(\theta) = \Pr(Y \leq y; \theta)$. This mid-$P$-value no longer guarantees the bounds on the Type I error probabilities; however, it has the appealing property, as some authors argue that the resulting intervals have coverage approximately equal to the nominal level (Vollset, 1993; Agresti and Coull, 1998).

Inference for $\theta$ in the form of $p$-values and confidence intervals based on the mid-$P$-value can be directly compared to those based on approximate methods, and no continuity correction is needed; for discussion see Davison and Wang (2002) and Davison et al. (2006). Naturally, by ignoring the continuity correction for the approximate methods, the coverage probability of the related confidence intervals goes below the nominal levels, which may have serious consequences, in particular for one-sided intervals. Consequently, one is faced with the following dilemma. To use an approach that provides narrower confidence intervals for which the actual coverage probability could be less than the nominal level, but usually close to it. Or to use an approach which controls the Type I error rates, but for which the actual coverage probability could be much larger than the nominal level. In this paper, we take the former view and discuss interval estimation obtained by highly accurate approximate methods without the need of continuity correction.

Approximate inference for $\theta$ includes either standard first-order methods, derived from Wald, score or signed likelihood root, or more accurate higher-order methods. For one-sided interval estimation, the extensive study of Cai (2005) points out that the score and the signed likelihood root, usually recommended for two-sided interval estimation, are outperformed by the Jeffreys intervals and the second-order corrected intervals. This corroborates the general view that the accuracy of the approximate methods plays a far more important role in one-sided interval estimation than in two-sided, where typically an excess in one end might compensate excess on the other end with the resulting coverage close to the nominal. In the next section we investigate the confidence intervals provided by a modified likelihood root statistic; see Barndorff-Nielsen (1986), Barndorff-Nielsen and Cox (1994, Chapter 6.6).

2.1. Accurate confidence intervals

Throughout this section let $\psi = \log(\theta/(1 - \theta))$ be the log reparameterization and denote by $\ell(\psi)$ the log-likelihood function and by $\ell(\psi) = n^{-1}E[-d^2\ell(\psi)/d\psi^2]$ the expected Fisher information per observation for the binomial model in this new reparameterization. Denote further by $W_n = \sqrt{n(\psi - \psi_0)\{\ell(\psi_0)\}^{1/2}}$, $Z_n = \sqrt{n\ell'(\psi_0)\{\ell(\psi_0)\}^{1/2}}$ and $R_n = \text{sign}(\hat{\psi} - \psi_0)\{2(\ell'(\hat{\psi}) - \ell(\psi_0))\}^{1/2}$ the usual first-order test statistics, namely the Wald statistic in the log parameterization, the score statistic and the signed likelihood root statistic. Here $\psi$ designates the maximum likelihood estimate of $\psi$ and $\ell'(\psi) = d\ell(\psi)/d\psi$ the score function. The confidence intervals obtained by inversion of the acceptance regions of any of these first-order test statistics have a coverage probability equal to the nominal level with an error of order $O(n^{-1/2})$, and have been thoroughly studied in the literature; see Brown et al. (2002, 2003). We introduce now the confidence intervals obtained by the inversion of the so-called third-order modified likelihood root statistic.

The modified likelihood root statistic, also known as the Barndorff-Nielsen’s $r^*$ after Barndorff-Nielsen (1986), is a modified version of the signed likelihood root that decreases the relative error in an asymptotic normal approximation to order $O(n^{-1})$ in discrete settings, where $n$ is the sample size. The adjustment uses, beside the signed likelihood root, a reparameterization invariant test statistic that involves sample space derivatives; in exponential models, this test statistic simplifies to the usual Wald in the canonical parameterization. The modified likelihood root statistic is defined by

$$r^* = R_0 + R_n^{-1}\log(W_n/R_n) \overset{\mathcal{D}}{\sim} N(0, 1),$$

suppressing the dependence on both $\psi$ and $n$, where $W_n$ and $R_n$ are given above. Although calculated in the log parameterization, the modified likelihood root statistic is invariant to reparameterization, and thus it make sense to use it for inference about the parameter $\theta$. An upper limit $r^*$ interval for $\theta$, $C_{\theta}^{1-\alpha} = \{0, \theta_{\alpha}^{\text{up}}\}$, is constructed by inversion of the test which accepts the null hypothesis $H_0 : \theta \leq \theta_0$ if $r^* \geq -\kappa$, where $r^* = r^*(\psi_0)$, $\psi_0 = \log(\theta_0/(1 - \theta_0))$ and $\kappa = 100(1 - \alpha)$th percentile of the standard normal distribution. In discrete models, the upper limit $\theta_{\alpha}^{\text{up}}$ defined in this way approximates the correct $100(1 - \alpha)$% upper limit to order $O(n^{-1})$; see Davison and Wang (2002).

Similar to the score or the signed likelihood root, the $r^*$ test statistic yields confidence limits that cannot be expressed in closed form. Therefore, in order to calculate the analytical expression for the $r^*$ confidence limits, we need a two-step approach. First, we find a stochastic expansion of $r^*$ with respect to the standard Wald statistic $W_n$, which is the Wald statistic taken in the initial parameterization $\theta$, $W_n = \sqrt{n(\theta - \theta_0)\{\ell(\theta_0)\}^{1/2}}$. Then, we revert this expansion to obtain the analytical expansion of the $100(1 - \alpha)$% upper limit $r^*$ to the desired order of accuracy. Using this procedure we obtain the following expansion for the $100(1 - \alpha)$% upper limit, $\hat{\theta}_{\alpha}^{\text{up}}$, to order $O(n^{-2})$:

$$\hat{\theta}_{\alpha}^{\text{up}} = \hat{\theta} + \kappa \hat{\sigma} n^{-1/2} + v_0(1 - 2\hat{\theta})n^{-1} + \kappa(\gamma_1 \hat{\sigma} + \gamma_2 \hat{\sigma}^{-1})n^{-3/2},$$

where \( \hat{\sigma}^2 = \hat{\theta}(1 - \hat{\theta}) \), \( v_0 = \frac{1}{2} \kappa^2 + \frac{1}{6} \), \( v_1 = -\left( \frac{13}{36} \kappa^2 + \frac{17}{36} \right) \) and \( v_2 = \frac{1}{36} \kappa^2 + \frac{7}{72} \). Note that a 100(1 - \( \alpha \))% lower limit \( r^* \), \( \hat{\theta}_L \), has a similar analytical expansion to (2) with \( \kappa \) replaced by \( -\kappa \).

Calculation of the \( r^* \) limits has one inconvenience: \( r^* \) is not well-defined at the boundaries, \( y = 0 \) or \( n \). For these extreme values we make an ad hoc choice to perturb the data by the correction \( \pm 1/n \). This choice seems to give \( r^* \) limits that are very close to the corresponding mid-\( P \) limits, even for small sample sizes. Fig. 1 illustrates the comparison of the confidence limits obtained by inverting the ordinary \( P \)-value determined by \( r^* \), \( Pr(\theta | r^*) = \Phi(r^*, \theta) \), the ones obtained by using the fourth-order approximation to the \( r^* \) limits as given by (2) and the confidence limits obtained by inverting the mid-\( P \)-value for two modest sample sizes \( n = 7 \) and \( n = 15 \). In the next section we prove theoretical results for the coverage probability of the 100(1 - \( \alpha \))% upper limit \( r^* \) intervals.

### 2.2. Edgeworth expansion of the coverage probability

In continuous settings the two-term Edgeworth expansion gives an error of approximation of order \( O(n^{-3/2}) \). Such an Edgeworth approximation is a continuous distribution function, and thus if used in the lattice case, it approximates the continuous part in the distribution and the error cannot be less than \( O(n^{-1/2}) \); for details see Severini (2000, Chapter 2.3). We will use this type of Edgeworth expansion in Section 3.2. There are, however, versions of the Edgeworth expansion that have an error of approximation of order \( O(n^{-3/2}) \); see Bhattacharya and Rao (1976, Chapter 6). We use such a version next, in studying the coverage probability of the modified likelihood root confidence intervals.

To obtain the Edgeworth expansion for the confidence probability of the one-sided \( r^* \) confidence intervals it requires a two-step approach. First, we write the stochastic expansion of \( r^* \) in powers of the score statistic, \( Z_n \). Then we expand the coverage probability of the corresponding upper limit confidence intervals, by using the expansion for the cumulative distribution function of \( Z_n \); see the Appendix for more details. Alternatively one could use the expansion of \( r^* \) with respect to \( \tilde{W}_n \), found in Section 2.1, followed by the corresponding Edgeworth expansion of the distribution of \( \tilde{W}_n \). Continuing the theorems presented in Brown et al. (2002) and Cai (2005) we state the following proposition.

**Proposition 2.1.** The coverage probability of the 100(1 - \( \alpha \))% upper limit confidence interval CI\( _u \), satisfies

\[
Pr(\theta \in CI_u) = (1 - \alpha) + Q_3(\theta, z^u_+ \sigma^{-1} \phi(\kappa)n^{-1/2} + (1/2)(1 - 2\theta)Q_3(\theta, z^u_+ \sigma^{-1} \phi(\kappa)n^{-1} + O(n^{-3/2})),
\]

where \( \sigma^2 = \theta(1 - \theta) \), \( z^u_+ \) is defined by (A.3) in the Appendix, \( Q_3(\theta, z) = g(\theta, z) - \frac{1}{2} \) and \( Q_2(\theta, z) = \frac{1}{2} g^2(\theta, z) - \frac{1}{2} g(\theta, z) + \frac{1}{12} \). Here \( g(\theta, z) \) denotes the fractional part of \( n\theta + n^{1/2} z \sigma \) and we assume that \( n\theta + n^{1/2} z \sigma \) is not an integer.

Note that the coverage probability of the \( r^* \) intervals even contains terms of \( O(n^{-1/2}) \). Although this is a discrete setting, this may still seem surprising, because the density function of \( r^* \) statistic agrees with the standard normal density function up to a multiplicative error of \( O(n^{-3/2}) \), in continuous settings. A closer inspection at this coverage probability reveals that the coefficients of \( n^{-1/2} \) and \( n^{-1} \) terms involve the functions \( Q_1(\theta, z) \) and \( Q_2(\theta, z) \), which are in fact oscillatory functions. These functions capture the oscillation in the coverage and, although depending also on \( n \), do not vanish in spite of how large \( n \) is. Nevertheless, in “a very weak” sense (Woodroofe, 1986) the oscillatory part in the above coverage probability is of order \( O(n^{-3/2}) \). More specifically, using Brown et al. (2002) it is straightforward to show that the average of each of the \( n^{-1/2} \) and \( n^{-1} \) oscillatory terms with respect to a smooth compactly supported function on a subinterval \( [a, b] \subset (0,1) \) satisfying the Lipschitz condition is of order lower than \( O(n^{-3/2}) \). As one of the referees points out, this argument crucially requires that the interest parameter is simultaneously bounded and integrated out, and practitioners may feel uncomfortable with this step. On the other hand, one can argue that such average provides a simpler and better picture of the performance of the oscillatory part over a set than at a single parameter value; see also Sweeting (2001) who considers average coverage probabilities with a view to evaluating the performance in repeated use. From this viewpoint, it is sensible to evaluate the intervals based on the non-oscillating part. The sum of the \( O(n^{-1/2}) \) and \( O(n^{-1}) \) non-oscillatory terms measures the systematic bias in the coverage, which when added to the nominal level, it provides a smooth.
approximation to the coverage probability. In particular, there is no systematic bias in the coverage of the $r^*$ confidence intervals to order $O(n^{-3/2})$.

Although these results hold asymptotically, the $r^*$ intervals show excellent coverage properties in settings of small sample sizes as well. We investigated numerically the coverage probability of the $r^*$ one-sided intervals for all sample sizes in the range $n = 10–50$. Here we present only the results for $n = 10$ and $n = 30$: Fig. 2 shows the comparison between the coverage probability of the 99% upper limit $r^*$ intervals and the upper limit mid-$P$ intervals. For $n = 30$ these plots are very similar to the ones presented in Fig. 4 of Cai (2005).

2.3. Expansion of the expected length

We now give the expansion of the expected distance of the upper confidence limit from the true value $\theta$. This expansion differs qualitatively from the Edgeworth expansion in the sense that, in the expansion for the length, the terms of order $O(n^{-1})$ and $O(n^{-3/2})$ are the relevant ones; the term of order $O(n^{-1/2})$ is common to all the asymptotic intervals.

**Proposition 2.2.** Let $\text{CI}^{u}_{r^*}$ be the $100(1 - \alpha)\%$ upper limit $r^*$ interval. The expected distance of the upper limit of $\text{CI}^{u}_{r^*}$ from $\theta$ admits the expansion

$$E[\mathcal{L}^{u}_{r^*}] = \kappa \sigma n^{-3/2} + \delta_1(\kappa, \theta) n^{-1} + \kappa \delta_2(\kappa, \theta) n^{-3/2} + O(n^{-2}),$$

where $\mathcal{L}^{u}_{r^*} = \hat{\theta}_{r^*} - \theta$, and

$$\delta_1(\kappa, \theta) = (\frac{1}{6} \kappa^2 + \frac{1}{6} \kappa (1 - 2 \theta)), \quad \delta_2(\kappa, \theta) = \frac{1}{12} (2 \kappa^2 - 2) \sigma^{-1} - \frac{1}{72}(26 \kappa^2 + 34) \sigma.$$

It is interesting to note that (3) agrees with the corresponding expansion for the so-called second-order corrected interval, derived in Cai (2005), to $O(n^{-2})$. Such a result is expected to hold for continuous models: the second-order corrected intervals by definition provide a coverage probability equal to the nominal level in a two-term Edgeworth expansion. For discrete settings this result seems to be new.

3. Confidence intervals for the log odds ratio parameter

Next, we consider two independent samples $\{Y_{11}, Y_{12}, \ldots, Y_{1n_1}\}$ and $\{Y_{21}, Y_{22}, \ldots, Y_{2n_2}\}$ from the binomial models $\text{Bin}(1, \theta_1)$ and $\text{Bin}(1, \theta_2)$, respectively. This section discusses interval estimation for the log odds ratio of the binomial parameter,
\[ \psi = \log(\theta_1(1 - \theta_2)/\theta_2(1 - \theta_1)) \]

Exact inference for \( \psi \) is usually achieved through the conditional distribution of \( Y_1|Y \),

\[ \Pr(Y_1 = y_1|Y = y; \psi) = \frac{ \binom{n_1}{y_1} \binom{n_2}{y - y_1} e^{\psi y_1}}{\sum_{j=0}^{\min(n_1, y)} \binom{n_1}{j} \binom{n_2}{y - j} e^{\psi j}}. \]

where \( Y_1 = \sum_{i=1}^{n_1} Y_{1i}, Y_2 = \sum_{i=1}^{n_2} Y_{2i} \) and \( Y = Y_1 + Y_2, \) \( \text{lo} = \max(0, y - n_2) \) and \( \text{up} = \min(n_1, y) \). Following the approach used in the one-binomial setting we consider approximate inference, as given by higher-order asymptotic methods based on this conditional distribution (see also Davison and Wang, 2002) and we study the confidence intervals obtained by inverting them.

This section expands the methodology and the results described in the previous section to account for a nuisance parameter in the multiparameter setting of comparing two binomial parameters by their log odds ratio. Besides the notational and computational effort, the main challenge is to evaluate unconditionally the theoretical performance of the confidence intervals for \( \psi \), derived by a conditional approach. We discuss the results after introducing some additional notation.

### 3.1. Accurate confidence intervals

We consider \( \gamma = (\psi, \eta) \) a one-to-one reparameterization to \( \theta = (\theta_1, \theta_2) \), where \( \psi \) is the parameter of interest defined above and \( \eta = \log(\theta_2/(1 - \theta_2)) \) is the nuisance parameter and let \( \ell(\gamma) = \ell(\gamma; Y_1, Y_2) \) be the log-likelihood for the combined binomial model expressed in the new reparameterization. We write \( \ell_0(\gamma) = \frac{1}{n} \ell(\gamma; Y_1, Y_2) \) for the observed information matrix and \( \ell_0(\gamma) = \ell(\gamma; Y_1, Y_2) \tilde{c}_0 \) for the Fisher information matrix per observation, where \( n = n_1 + n_2 \) is the size of the combined sample, and use subscript notation to indicate the partition of these matrices in accordance to the partition of the parameter; for example \( \ell_0(\gamma) = \ell(\gamma; Y_1, Y_2) \tilde{c}_0 \). Denote further by \( \ell_0(\gamma) = \ell(\psi, \eta_0) \) the profile log-likelihood function, where \( \eta_0 \) is the maximum likelihood estimate of \( \eta \) for a specified value of \( \psi \) and by \( \ell_0(\psi) = \ell(\psi, \eta) \) the corresponding profile observed information, where \( \ell_0(\psi) \) is the second derivative of \( \ell(\psi) \). The first-order test statistics, Wald, score and the signed likelihood root can be written then as \( W_n = \sqrt{n} \tilde{c}_0 (\psi - \eta(\psi)) \) and \( Z_n = \sqrt{n} \ell_0(\psi)(\psi - \eta(\psi)) \) and \( R_n = \text{sign}(\psi - \eta(\psi)) 2 \ell(\psi) - \ell_0(\psi) \) for one parameter of interest is given by (see Barndorff-Nielsen, 1986)

\[
\begin{align*}
 r^* &= R_n + R_n^{-1} \log \rho \frac{W_n}{R_n},
\end{align*}
\]

where \( \rho = |\ell_0(\psi)|/|\ell_0(\psi, \eta)| \) is an adjustment for the nuisance parameter \( \eta \). For a complete derivation of \( r^* \) see Barndorff-Nielsen and Cox (1994, Chapter 6.6). In multiparameter exponential models with a scalar parameter of interest, the statistic \( r^* \) is derived from a saddlepoint approximation to the conditional distribution of the maximum likelihood estimator \( \hat{\psi}; \hat{\eta} \); see Pierce and Peters (1992). As in the one-parameter case, \( r^* \) is invariant to interest respecting reparameterization and it has a density function which agrees with the standard normal density function with relative error \( O(n^{-3/2}) \), for continuous models. The modified likelihood root interval is constructed by inversion of the \( r^* \) test statistic: \( \text{CI}^u = (-\infty, \hat{\psi}^u) \) is the 100(1 - \( \alpha \))% upper limit \( r^* \) interval, if the 100(1 - \( \alpha \))% upper limit \( \hat{\psi}^u \) satisfies \( \Phi(r^*; \hat{\psi}^u) = \alpha \). Note that again the \( r^* \) limits cannot be expressed in closed form, and thus the resulting confidence intervals, in spite of their frequently remarked excellent performance in terms of coverage probability, have not benefitted from the same popularity that other less accurate intervals like Wald and score for instance have received. Expression (A.9) provided in the Appendix gives an \( O(n^{-2}) \) approximation to the \( r^* \) limit.

Bayesian procedures with non-informative priors have been long known to provide posterior intervals with good coverage properties. DiCiccio and Martin (1993) propose to use matching priors in approximate Bayesian inference as an alternative to more complicated frequentist formulas. Using the prior \( \pi(\gamma) \propto \ell_0^{1/2}(\gamma) i_{P}(\gamma) \), where \( i_{P}(\gamma) = i_{P}(\gamma) - i_{P}(\gamma) i_{P}(\gamma)^{-1} i_{P}(\gamma) \), the resulting tail area approximation is \( \Pr(\Psi \geq \psi|y) = \Phi(r_B^u) \) with an error of \( O(n^{-3/2}) \), where \( r_B^u \) is

\[
\begin{align*}
 r_B^u &= R_n + R_n^{-1} \log \rho \frac{Z_n}{R_n},
\end{align*}
\]

(Brazzale et al., 2007, Chapter 8) and is referred to as the Bayesian version of the modified likelihood root, \( r_B^u \). The prior \( \pi(\gamma) \) is in the Tibshirani–Peers class of matching priors (Tibshirani, 1989; Peers, 1965) and is discussed in more detail in Staicu and Reid (2008).

Using the asymptotic properties of \( r_B^u \), the Bayesian modified likelihood root interval is defined in analogy to the \( r^* \) interval: \( \hat{\psi}^u \) is the 100(1 - \( \alpha \))% upper limit \( r_B^u \) if \( \Phi(r_B^u, \hat{\psi}^u) = \alpha \). These limits are analyzed in further detail in Section 3.3. In the next section we discuss the coverage probability properties of the confidence intervals based on the two versions of modified likelihood root.
3.2. Comparison of the coverage probability

In spite of the conditional approach used in deriving the proposed confidence limits, the confidence intervals are evaluated by their unconditional coverage and length properties, which is fairly common in the $2 \times 2$ tables literature (see for instance Lloyd and Moldovan, 2000). Denote by $CI$ a generic interval for $\psi$. The coverage probability of $CI$ is defined as

$$C(\gamma, n_1, n_2) = Pr(\psi \in CI; \gamma) = \sum_{y_1=0}^{n_1} \sum_{y_2=0}^{n_2} Pr(Y_1 = y_1; \gamma)Pr(Y_2 = y_2; \gamma)1(\psi; y_1, y_2), \quad \text{(6)}$$

where $Pr(Y_i = y_i; \gamma)$ is the binomial probability corresponding to the binomial distribution with size $n_i$ and parameter $\theta_i$, and $1(\psi; y_1, y_2)$ is the indicator function that equals 1 if the interval $CI$ contains $\psi$ for $Y_i = y_i$ for $i = 1, 2$ and 0 if not.

To study the coverage accuracy of the upper limit intervals based on the $r^*$ and $r^*_u$ statistics, we use again the Edgeworth expansion, with the limitation that only the smooth component of the coverage is retained in the expansion. We conjecture that the oscillatory component of the coverage is identical for the two intervals and furthermore that is of order $O(n^{-3/2})$ in the “very weak” sense of Brown et al. (2002). This is argued by the high similarity of the test statistics in the one and two binomial settings (see Section 4), as well as by using Theorem 23.1 of Bhattacharya and Rao (1976, Chapter 6.3), but a formal proof will be carried out elsewhere. In view of these points we focus only on the smooth terms which capture the main regularity in the coverage. The advantage of ignoring the oscillatory component of the coverage is that the Edgeworth expansion can be actually obtained by the same techniques that are used for continuous models.

A direct Edgeworth approach involves a stochastic expansion for the Wald statistic and then evaluation of the corresponding approximate cumulants. Alternatively, this tedious algebra can be avoided by using an elegant shrinkage argument in a Bayesian analysis due to Dawid (1991), which is nicely presented in Datta and Mukerjee (2004). We give next the Edgeworth expansion for the coverage probability of both $CI^u_r$ and $CI^u_g$, obtained by using the shrinkage argument.

**Proposition 3.1.** Let $\psi$ be the log odds ratio of two binomial parameters. The smooth coverage probability of the upper limit intervals $CI^u_r$ and $CI^u_g$ admits the following expansions, correct to $O(n^{-3/2})$:

$$P_r(\theta) = Pr(\psi \in CI^u_r) = (1 - \alpha), \quad \text{(7)}$$

$$P_g(\theta) = Pr(\psi \in CI^u_g) = (1 - \alpha) - \frac{3a_2^2 + 2b_2}{24} \kappa \phi(\kappa)n^{-1}, \quad \text{(8)}$$

where $a_2 = a_2(\theta) = (A(1 - A))^{-1/2}(A^2(\sigma_1/\sigma))^2 \rho_2^2 + (1 - A)^2(\sigma_2/\sigma)^3 \rho_1^3$ and $b_2 = b_2(\theta)$ has a rather complicated expression and it is given by (A.11) in the Appendix and $A = n_1/n$. Here we use $\sigma_i$ for the variance and $\rho_i$ for the third standardized cumulant of the Bin$(1, \theta_i)$, for $i = 1, 2$, while $\sigma^2 = A\sigma_1^2 + (1 - A)\sigma_2^2$.

The derivation of these expansions makes use of two important results: (1) the $r^*_u$ limits agree with the exact posterior limits with an error of $O(n^{-2})$ and (2) Theorem 1 of Mukerjee and Dey (1993) which gives the Edgeworth expansion for the frequentist coverage probability of the exact limits. For the first expansion we use the inversion of the Edgeworth approximation described in Hall (1983). The algebra involved is rather heavy and we omit it here to save space. The proof of Proposition 3.1 is briefly sketched in the Appendix.

Both expressions (7) and (8) provide an expansion of the smooth coverage probability only: a complete Edgeworth expansion should include oscillation terms of order $O(n^{-3/2}), O(n^{-1})$. As an additional note, here and in the results that follow all the quantities are obtained in the parameterization $\gamma = (\psi, \eta)$. The parameterization $\theta$ is used only for simplicity of the exposition; writing $a_2$ in the parameterization $\gamma$ requires further the substitution $\theta_1 = \exp(\psi + \eta)/(1 + \exp(\psi + \eta))$ and $\theta_2 = \exp(\eta)/(1 + \exp(\eta))$.

3.3. Comparison of the expected length

In this section we provide expansions for the expected value of the distance between each of the two upper limits and $\psi$, correct to $O(n^{-3/2})$. The theoretical calculations are fairly technical, but essentially expand the methodology from Section 2.3 to account for the nuisance parameter $\eta$; see Staicu (2007) for details. Here we present the main conclusions from these calculations.

First, let $CI^u_r = (-\infty, \hat{\psi}^u_r)$ and $CI^u_g = (-\infty, \hat{\psi}^u_g)$ be the 100$(1 - \alpha)$% upper limit intervals for $\psi$ and denote the distance of the upper limits $\hat{\psi}^u_r$ and $\hat{\psi}^u_g$ from $\psi$ by $L_r$ and $L_g$, respectively.

**Proposition 3.2.** The expected distance of the upper limits of $CI^u_r$ and $CI^u_g$ from $\psi$ admits the following expansions to $O(n^{-2})$:

$$E[L_r] = \kappa(A(1 - A))^{-1/2} \frac{\sigma_1 \sigma_2}{\sigma} n^{-1/2} + \delta_1(\kappa, \theta)n^{-1} + \kappa \delta_2(\kappa, \theta)n^{-3/2},$$

$$E[L_g] = \kappa(A(1 - A))^{-1/2} \frac{\sigma_1 \sigma_2}{\sigma} n^{-1/2} + \delta_1(\kappa, \theta)n^{-1} + \kappa \delta_2(\kappa, \theta)n^{-3/2},$$
where

\[
\delta_1(\kappa, \theta) = \{A(1 - A)\}^{-1/2} \sigma_1 \sigma_2 \left\{ \frac{a_2}{6} \kappa^2 - \frac{a_2 + 3a_0}{6} \right\},
\]

\[
\delta_2(\kappa, \theta) = \{A(1 - A)\}^{-1/2} \sigma_1 \sigma_2 \left\{ \frac{5a_2^2 + 3b_2}{72} \kappa^2 + \left( \frac{c_0}{8} - \frac{8a_2^2 + 3b_2}{72} - \frac{3b_0 + 3a_0^2 + 3a_2a_0}{12} \right) \right\},
\]

\[
\delta_2(\kappa, \theta) = \{A(1 - A)\}^{-1/2} \sigma_1 \sigma_2 \left\{ \frac{5a_2^2 + 3b_2}{72} \kappa^2 + \left( \frac{c_0}{8} - \frac{17a_2^2 + 9b_2}{72} - \frac{3b_0 + 3a_0^2 + 3a_2a_0}{12} \right) \right\}.
\]

\[
a_0 = a_0(\theta) = \{A(1 - A)\}^{1/2} (\sigma_1 \sigma_2 / \sigma^2) / (\sigma_1 / \sigma) \rho_1^2 - (\sigma_1 / \sigma) \rho_1^2 \}
\]

and \(a_2\) and \(b_2\) are defined in Proposition 3.1. The expressions for \(b_0 = b_0(\theta)\)

\[
\text{and } c_0 = c_0(\theta) \text{ involve the fourth standardized cumulants of the two binomial variables and are given by (A.12) and (A.13), respectively, in the Appendix.}
\]

The proof follows from the expression of the expectation of the \(r^*\) limits and \(r^*_y\) limits provided in (A.9) and (A.10), using \(f_\hat{\varphi}(\hat{\psi}) = i_{\varphi\varphi}(\hat{\psi})\) and the consistency of the maximum likelihood estimator. The coefficients of the \(n^{-1/2}\) and \(n^{-1}\) terms are the same in the two intervals. It is interesting to note that for all values of \(\psi, Cl_r\) is always slightly wider than \(Cl_y\).

3.4. Numerical study

We examine now some particular examples, in order to illustrate the performance of the coverage and the length of the two intervals in finite sample sizes. It should be noted first that the same inconvenience for the boundaries \(y_1 = 0\) or \(n_1 = 0\) or \(n_2\), noted in the one-binomial setting, extends to this two binomial situation. Pierce and Peters (1992) refer to these data points as uninformative and ignore them. Unfortunately, these cases are needed in the calculation of the actual coverage probability, when the purpose is to compare the coverage probability of the confidence intervals with the corresponding nominal level. For this reason, we use the tests with the ad hoc correction for \(y_1\) and/or \(y_2\) of either \(n_1^{-1/2}\), \(n_1 - n_1^{-1}\) and/or \(n_2^{-1}\), \(n_2 - n_2^{-1}\), to be consistent with the approach proposed in Section 2.1. In general, this choice does not affect the coverage probability; it does have
some impact when at least one-binomial parameter is close to 0 or 1; see Staicu (2007) for alternative approaches to handle this boundary values problem.

3.4.1. Example 1

Fig. 3 displays the coverage probability of the 95% upper limit intervals based on $r^*$ and $r^*_B$ statistics for two different sample sizes. The left panels give the comparison for $\theta_1 = 0.4$ and $-3 \leq \psi \leq 3$, and the right panels for $\theta_1 = 0.9$ and $-1.3 \leq \psi \leq 5$. The upper panels show the coverage for equal sample sizes $n_1 = n_2 = 10$ while the lower panels show the coverage for different sample sizes $n_1 = 10$ and $n_2 = 30$. Clearly, the unconditional coverage probability of the intervals for $\psi$ varies with both $\psi$ and $\theta_1$, or alternatively $\eta$, since $\eta = -\psi + \log(\theta_1/(1-\theta_1))$. But in both cases, the coverage for the two confidence intervals wiggles around the nominal level, with more jiggling for the case where the combined sample size is larger. In fact we performed an extensive numerical study of the coverage probability for various other sample sizes $n_1$ in the range 10–30 and for different values of the complementary parameter of $\psi$. Our findings suggest some indication of a slightly larger coverage probability for $r^*$ intervals than the coverage for $r^*_B$ intervals, but this difference vanishes as the combined sample sizes increases.

3.4.2. Example 2

We examine to some detail the case $n_1 = n_2 = 10$. Fig. 4 presents the coverage probabilities and the expected length of the two-sided 95% confidence intervals derived from the mid-$P$, $r^*$ and $r^*_B$ for $0 \leq \theta_1 \leq 1$ and several values of $\psi$. It is interesting to compare the upper left plot to the analogous one presented in Fig. 2 of Agresti and Min (2002): the two-sided unconditional intervals and the mid-$P$ confidence intervals appear to give similar coverage probability when $\psi = 0$. More interestingly, the left panels reveal that the coverage probabilities of the 95% confidence intervals based on $r^*$, $r^*_B$ and mid-$P$ are identical and above the nominal level for all values of $\theta_1$, with $r^*_B$ confidence intervals having the smallest expected length. In fact, when the true log odds ratio equals $\psi = 0$, we find that for all the sample sizes $n_1$ in the range of 10–30, with $n_1 = n_2$, both $r^*$-type confidence intervals result in identical coverage probability; the conservativeness suggested by the common coverage probability decreases with larger sample sizes.

4. Discussion

This work extends the comprehensive study of the confidence intervals based on the first-order asymptotic methods to higher-order asymptotic ones. In the context of a single binomial variable, our findings update the results of Cai (2005) and Brown et al. (2002, 2003). The theoretical results proved for the Barndorff-Nielsen’s $r^*$ statistic therefore correctly reinforce previous numerical studies regarding the performance of the $r^*$ confidence intervals. We should also mention that in our investigation another modification of the likelihood root showed surprising properties. More specifically, the modified likelihood root $R_n - R_n^{-1} \log(R_n/Z_n)$ provides confidence intervals of identical coverage probability and expected length to Jeffreys credible intervals, discussed in Cai (2005), to order $O(n^{-3/2})$. A detailed study of these intervals is given in Staicu (2007). Perhaps, also interesting is that this new test statistic for testing one-binomial parameter, also invariant to reparameterizations, seems to be the equivalent of the Bayes version of $r^*$ statistic, in the context of testing two independent binomial parameters. However, the findings for this new confidence interval hold for the binomial setting only and one should not expect to hold in more general settings.

For the log odds ratio in two independent samples our results are limited to providing Edgeworth expansions only for the smooth coverage probability. Overall, both the $r^*$ intervals and the $r^*_0$ intervals almost entirely eliminate the systematic bias in the coverage probability evaluated unconditionally.

To summarize, Barndorff-Nielsen’s $r^*$ statistic-based interval is recommended for interval estimation of the binomial proportion and/or the log odds ratio of two proportions, especially for small and moderate sample sizes. The interval has excellent properties in terms of coverage and length.

Acknowledgments

The work is part of my Ph.D. thesis and was supported by the Natural Sciences and Engineering Research Council of Canada. I am extremely grateful to my supervisor Nancy Reid, for inspiring suggestions, for reading previous drafts of this work and very helpful comments and to Lidia Nicola. I would like to thank two anonymous referees for very constructive comments which have led to a much improved paper.

Appendix A.

Derivation of the Edgeworth expansion for the coverage of $r^*$-intervals for the binomial proportion. Brown et al. (2002) show that by using Theorem 23.1 of Bhattacharya and Rao (1976, Chapter 6.3), the cumulative distribution function of $Z_n$, $F_n(z) = \Pr(Z_n \leq z)$, can be approximated by a two-term Edgeworth expansion as

$$F_n(z) = \Phi(z) + p_1(z)\phi(z)n^{-1/2} + p_2(z)\phi(z)n^{-1} - Q_3(\theta,z)\sigma^{-1}\phi(z)n^{-1/2} + (Q_1(\theta,z)\sigma p_3(z) - Q_3(\theta,z)z)\sigma^{-2}\phi(z)n^{-1}$$

(A.1)

to $O(n^{-3/2})$, for values of $z$ such that $n\theta + n^{1/2}\sigma z$ is not an integer, where $\sigma^2 = \theta(1 - \theta)$. This expansion uses the following quantities:

$$p_1(z) = -\frac{1}{6}(z^2 - 1)(1 - 2\theta)\sigma^{-1},$$

$$p_2(z) = \frac{1}{36}(2z^5 - 11z^3 + 3z) - \frac{\sigma^{-2}}{72}(z^5 - 7z^3 + 6z),$$

$$p_3(z) = -\frac{1}{6}(z^2 - 3z)(1 - 2\theta)\sigma^{-1}.$$  

Note that expression (A.1) involves the usual smooth terms specific to an Edgeworth expansion, $p_1(z)$ and $p_2(z)$, which can be expressed using Hermite polynomials, and also contains oscillatory terms $Q_1(\theta,z)$ and $Q_3(\theta,z)$. These oscillatory terms are a consequence of the discrete nature of the binomial variable and they are present at both orders $O(n^{-1/2})$ and $O(n^{-1})$. When $z = z(n)$ depends on $n$, say $z = z_0 + c_1n^{-1/2} + c_2n^{-1} + O(n^{-3/2})$ where $z_0, c_1, c_2 \in \mathbb{R}$ are constants, expression (A.1) can be simplified further, just by rearranging the terms of appropriate order.

We first expand $r^*$ with respect to the score statistic $Z_n$ by using Taylor expansions and obtain

$$r^* = Z_n + (-\frac{1}{6}(1 - 2\theta)\sigma^{-1}Z_n^2 + \frac{1}{6}(1 - 2\theta)\sigma^{-1}n^{-1/2} + \frac{1}{72}(5\sigma^{-2} - 14)Z_n^3 - \frac{1}{72}(5\sigma^{-2} - 2)Z_n)n^{-1}. \tag{A.2}$$

The expansion can be reverted by series inversion techniques, leading to an expansion of $Z_n$ in terms of powers of $r^*$. Let $\kappa$ be the $100(1 - x)$th percentile of the standard normal distribution, and denote by $z_{\kappa}^n$, the $100(1 - x)$ quantile of the $r^*$ statistic; $z_{\kappa}^n$, can be written to $O(n^{-3/2})$ as

$$z_{\kappa}^n = -\kappa - \frac{1}{6}(1 - \kappa^2)(1 - 2\theta)\sigma^{-1}n^{-1/2} + \kappa(\frac{1}{12}(7\kappa^2 + 5) + \frac{1}{72}(k^2 - 1)\sigma^{-2})n^{-1}. \tag{A.3}$$

We now use the above Edgeworth expansion to calculate the distribution function at the point $z_{\kappa}^n$, and the result of Proposition 2.1 follows.
Expansion for $r^*$ and $r^*_u$-based intervals. We take a similar approach to Section 2.1 and first expand $r^*$ and $r^*_u$ test statistic with respect to $W_n$ and then invert the expansion using series inversion techniques. For $r^*$ for example, the first expansion is given by

$$r^* = W_n + \left( \frac{\hat{a}_2}{6} W_n^2 - \frac{\hat{a}_2 + 3\hat{a}_0}{6} \right) n^{-1/2} + \left\{ \frac{\hat{a}_2 + 3\hat{b}_2}{72} W_n^3 + \left( \frac{4\hat{a}_2 + 3\hat{b}_2}{72} + \frac{3\hat{b}_0 + 3\hat{a}_0^2}{12} + \frac{3\hat{a}_0}{6} \right) W_n \right\} n^{-1},$$

(A.4)

where

$$\hat{a}_2 = \left\{ \frac{\partial^3}{\partial\psi^3} \ell_\nu(\psi) \right\} (\nu_0(\psi))^{-3/2},$$

(A.5)

$$\hat{b}_2 = \left\{ \frac{\partial^4}{\partial\psi^4} \ell_\nu(\psi) \right\} (\nu_0(\psi))^{-2},$$

(A.6)

$$\hat{a}_0 = \frac{\partial}{\partial\psi} \ell_\nu(\psi, \nu_0) (\nu_0(\psi))^{-1/2},$$

(A.7)

$$\hat{b}_0 = \frac{\partial^2}{\partial\psi^2} \ell_\nu(\psi, \nu_0) (\nu_0(\psi))^{-1}.$$

(A.8)

Note that the right side of (A.5) when expressed with respect to the binomial proportions $\hat{v}_1$ and $\hat{v}_2$ gives $a_2(\hat{v})$ with $a_2(\hat{v})$ introduced by Proposition 3.1; same is true for all the other three quantities, for $a_0(\hat{v})$, $b_2(\hat{v})$ and $b_0(\hat{v})$ introduced by Proposition 3.2. A similar expansion is obtained for $r^*_u$, with the difference that the first term of the coefficient of $W_n$ in (A.4) is $(13\hat{a}_2^4 + 9\hat{b}_2)/72$. The inversion of expansion (A.4) and the counterpart corresponding to $r^*_u$ give the following approximations for the upper limits, accurate to $O(n^{-2})$:

$$\hat{\psi}^u = \nu + \kappa (\nu_0(\hat{\nu}))^{-1/2} n^{-1/2} + \left( \frac{\hat{a}_2}{6} \kappa^2 - \frac{\hat{a}_2 + 3\hat{a}_0}{6} \right) (\nu_0(\hat{\psi}))^{-1/2} n^{-1}$$

$$+ \left\{ \frac{5\hat{a}_2^2 + 3\hat{b}_2}{72} \kappa^3 - \frac{8\hat{a}_2^2 + 3\hat{b}_2}{72} + \frac{3\hat{b}_0 + 3\hat{a}_0^2}{12} + \frac{3\hat{a}_0}{6} \right\} (\nu_0(\hat{\psi}))^{-1/2} n^{-3/2}.$$

(A.9)

$$\hat{\psi}^u = \nu + \kappa (\nu_0(\hat{\psi}))^{-1/2} n^{-1/2} + \left( \frac{\hat{a}_2}{6} \kappa^2 - \frac{\hat{a}_2 + 3\hat{a}_0}{6} \right) (\nu_0(\hat{\psi}))^{-1/2} n^{-1}$$

$$+ \left\{ \frac{5\hat{a}_2^2 + 3\hat{b}_2}{72} \kappa^3 - \frac{17\hat{a}_2^2 + 9\hat{b}_2}{72} + \frac{3\hat{b}_0 + 3\hat{a}_0^2}{12} + \frac{3\hat{a}_0}{6} \right\} (\nu_0(\hat{\psi}))^{-1/2} n^{-3/2}.$$

(A.10)

Proof of Proposition 3.1. The algebra involved is rather heavy and we omit most of the calculations. We first discuss the Edgeworth expansion for the coverage of $\text{CI}^u_0 = (-\infty, \hat{\psi}^u)$. Under the prior $\pi(\gamma)$, the $r^*_u$ statistic approximates the marginal posterior distribution with error of $O(n^{-3/2})$. This means that the $r^*_u$ limit, given above, equals the posterior limit of $\psi$, up to $O(n^{-2})$. We use $100(1 - \alpha)$ upper limit $r^*_u$ as given by (A.10) in Theorem 1 of Mukerjee and Dey (1993). Since the prior $\pi(\gamma)$ is a first-order matching prior it implies $\text{Pr}(\psi \leq \hat{\psi}_u; \theta) = (1 - \alpha) + \kappa \psi(\kappa) T_2(\pi, \gamma) n^{-1} \approx O(n^{-3/2})$, where $T_2(\pi, \gamma)$ is given by (3.5b) in the cited paper. The expression for $T_2$ reduces to $-3(\hat{a}_2^2 + 2\hat{b}_2)/24$, where $b_2 = b_2(\hat{v})$ is equal to

$$b_2 = 4(1 - \Delta)^{-1}(-2\Delta(\sigma_1(\sigma_1)\rho^2 - (1 - \Delta)^3(\sigma_2(\sigma_1)\rho_3^2 + \sigma_1(\sigma_1)(\sigma_2(\sigma_1)\rho_3^2)\rho_2^3 + 3\sigma_2^2\rho_2^3),$$

(A.11)

where $\alpha$ and $b_2$ are defined in Propositions 3.1 and 3.2, and $\rho^4_1$ and $\rho^4_2$ are the fourth standardized cumulant of the Bin(1, $\theta_1$) and Bin(1, $\theta_2$), respectively. The representation in parameterization $\theta$ was used for simplicity purposes only.

The Edgeworth expansion for the coverage probability of $r^*$ interval is obtained by first writing the $r^*$ limit with respect to the $r^*_u$ limit and then by inverting the Edgeworth expansion found in (7).

Additional notation needed by Proposition 3.2. In the calculation of the expected length of both $r^*$ and $r^*_u$ one-sided intervals we used the following notation:

$$b_0 = \frac{\sigma_1^2\sigma_2^2}{\sigma_4^2}(1 - \Delta)\rho_1^4 + A_0\rho_2^4 - \frac{\sigma^4_1}{\sigma^4_2}(1 - \Delta)^2(\sigma_2(\sigma_1)\rho_1^2 + A(1 - \Delta)(\sigma_1(\sigma_1)\rho_3^2)\rho_2^2,,$$

(A.12)

$$c_0 = 38\frac{\sigma_1^2\sigma_2^2}{\sigma_4^2}(1 - \Delta)^2(\sigma_2(\sigma_1)\rho_1^2 + \Delta(1 - \Delta)(\sigma_1(\sigma_1)\rho_3^2)\rho_2^2 + \frac{\sigma^2_1\sigma_2^2}{\sigma^4_4}(1 - \Delta)(\rho_2^2)^2 + A(\rho_2^2)^2)$$

$$+ \frac{3}{A(1 - \Delta)}((1 - \Delta)^2(\sigma_2(\sigma_1)\rho_1^2)^2 + A^2(\sigma_1(\sigma_1)\rho_3^2)\rho_2^2) + 2b_0 + 54\rho_2^6.$$

(A.13)
where $a_0$ is given in Proposition 3.1. For example $c_0 = c_0(\theta)$ is actually defined by the expression below, after substituting $\theta_1$ for $\exp(\psi + \eta)/(1 + \exp(\psi + \eta))$ and $\theta_2$ for $\exp(\eta)/(1 + \exp(\eta))$:

$$c_0(\gamma) = \frac{3}{8} \frac{\hat{c}_1}{\hat{c}_1} \left( \frac{\hat{c}_0(\gamma)}{\hat{c}_0(\gamma)} \right)^{-1} \frac{\hat{c}_2}{\hat{c}_2} \left( \frac{\hat{c}_0(\gamma)}{\hat{c}_0(\gamma)} \right) - \frac{2}{8} \left[ \frac{\hat{c}_2^2}{\hat{c}_2^2} \left( \frac{\hat{c}_0(\gamma)}{\hat{c}_0(\gamma)} \right)^2 \hat{R}(\gamma) + 2 \frac{\hat{c}_2}{\hat{c}_2} \left( \frac{\hat{c}_0(\gamma)}{\hat{c}_0(\gamma)} \right) \hat{R}(\gamma) + \frac{\hat{c}_2^2}{\hat{c}_2^2} \left( \frac{\hat{c}_0(\gamma)}{\hat{c}_0(\gamma)} \right) \hat{R}(\gamma) \right].$$

where we use superscripts to indicate the partition of the inverse of the Fisher information matrix in accordance to the partition of the parameter: for example $\hat{R}(\gamma)$ denotes the $(\psi, \eta)$ component of the $(\hat{c}(\gamma))^{-1}$.

References


