Ann. Inst. Statist. Math.
Vol. 52, No. 1, 28–42 (2000)

COVERAGE PROPERTIES OF ONE-SIDED INTERVALS IN THE DISCRETE CASE AND APPLICATION TO MATCHING PRIORS

JUDITH ROUSSEAU

Laboratoire de Statistique, CREST, Timbre J340, 92241 Malakoff Cedex, France and Laboratoire de Statistique Théorique et Appliquée, Université de Paris 6, France

(Received September 22, 1997; revised August 14, 1998)

Abstract. We consider asymptotic coverage properties of one-sided posterior confidence intervals for discrete distributions, with a unidimensional parameter of interest and a nuisance parameter of arbitrary dimension. In this case, no higher order asymptotic expansion of the frequentist coverage for these intervals is established, unless some randomization is added. We study here the existence of such frequentist expansions and propose simple continuity corrections based on a uniform random vector. This helps in determinig a family of matching priors for one sided intervals in the discrete case.

Key words and phrases: Asymptotic expansion, credible region, Edgeworth expansion, frequentist coverage, lattice distribution, posterior coverage.

1. Introduction

The past thirty years have seen a great number of papers dealing with approximate confidence intervals and, in particular, with the problem of determining confidence sets which would have both good frequentist and Bayesian coverage properties. In the continuous case this problem has been settled for different types of confidence sets. In particular, Welch and Peers (1963) and Peers (1965) consider one-sided confidence intervals with and without nuisance parameters, their results were improved by Tibshirani (1989), Mukerjee and Dey (1993) and Mukerjee and Ghosh (1997). They derive classes of priors such that both frequentist and Bayesian coverages of these confidence intervals agree to the order n^{-1}, i.e. matching priors for one-sided confidence intervals to the second order of approximation. They also prove that there is no matching prior to the order $n^{-3/2}$ except in some special cases. Other types of confidence regions have been considered as well, such as two-sided intervals, by Severini (1993) who provides asymptotic coverages to the order $O(n^{-3/2})$, Bayesian and frequentist, the latter being conditional on some ancillary statistics. Likelihood based regions and highest posterior density regions have also been studied; see for instance Ghosh and Mukerjee (1993).

However, these results do not extend to the discrete case, as far as the frequentist coverage is concerned, since they are based on formal (i.e. continuous) Edgeworth expansions. Only a few results have been established for lattice variables. Davison (1988) provides expansions for one-sided intervals, with implicit end-points, when the observations are univariate and on a lattice. There have been some advances, both in the binomial and in the Poisson cases; see Hall (1982) and Blyth (1986), on different possible approximations, to the order $O(n^{-1/2})$ of one-sided and two-sided intervals.

To derive higher order expansions, some continuity corrections have already been proposed. Lehmann (1986) has proposed to make each observation continuous by adding a uniform random variable on $[0,1]$ to every observation. Babu and Singh (1989) have established that the distribution of the convolution of a sum of n iid lattice variables with a uniform variable on $[-1/2\sqrt{n}, 1/2\sqrt{n}\,]$ has a continuous Edgeworth expansion to the order $O(n^{-1})$.

In this paper we study the asymptotic behaviour of one-sided intervals, both from a frequentist and a Bayesian point of view, when the observations are discrete, and we propose a method to smooth these intervals, so that they can be better controlled in terms of coverages, these results are obtained for multivariate as well as univariate models.

Let $X^n = (X_1, \dots, X_n)$ be a sample of random variables with distribution P_θ^n, where $\theta = (\theta_1, \dots, \theta_k)$ belongs to $\Theta \subset \mathbb{R}^k$. We consider inference about θ_1, $(\theta_2, \dots, \theta_k)$ being a nuisance parameter. Let π be a prior distribution on the parameter θ. In the following, $P^\pi[\cdot \mid X^n]$ denotes the posterior distribution of θ and C_α^π denotes the posterior one-sided confidence interval defined by

$$P^\pi[C_\alpha^\pi \mid X^n] = P^\pi[\theta_1 < k_n(\alpha) \mid X^n] = \alpha,$$

where $k_n(\alpha)$ depends on X^n. This is the one-sided posterior confidence interval considered by Peers (1965). As a first order approximation, it is well known that

$$P_\theta^n[C_\alpha^\pi] = \alpha + O(n^{-1/2}). \tag{1.1}$$

So, to this order, Bayesian and frequentist one-sided confidence coverages are equivalent. Peers (1965) has shown that when the random variables are continuous with respect to Lebesgue measure, $P_\theta^n[C_\alpha^\pi]$ has an asymptotic expansion of the form

$$P_\theta^n[C_\alpha^\pi] = \alpha + P_1(\theta,\pi)/\sqrt{n} + O(n^{-1}),$$

so that if the prior satisfies some differential equation (which leads to Jeffreys prior in the unidimensional case), (1.1) is correct to the order $O(n^{-1})$.

Unfortunately, these results do not apply in the discrete case. Indeed, in Section 2 we prove that expansions of the frequentist coverage of one-sided confidence intervals exist to the order $o(n^{-1/2})$, if and only if the statistic controlling the one-sided interval behaves like a nonlattice random variable, to the first order of approximation. Under this condition, denoted hypothesis (H) in Theorem 2.1, we can then determine a class of matching prior to this order of approximation. Moreover these matching priors are formally the same as those obtained in the continuous case by Peers (1965). To get a better approximation, we propose, in Section 3, a simple correction for one-sided Bayesian intervals, which is to add to the posterior α-quantiles a uniform perturbation of order $n^{-1/2}$. Then, the corrected one-sided intervals $C_\alpha^\pi(U)$ have correct coverage properties to the order $O(n^{-1})$, even when hypothesis (H) is not satisfied. This result is interesting from a theoretical point of view since it generates a class of matching priors, in the sense of Welch and Peers (1963), for a certain class of discrete distributions. An expansion to the third order is also established in Section 4 with applications to matching confidence sets. The corrections we propose give a great flexibility to the model, since we can then control the term of order $O(n^{-1})$ in the expansion. In particular we exhibit a correction allowing for a class of matching priors to the order $o(n^{-1})$. Besides, when the observations are related with a latent continuous process, this continuity correction has an interesting implication, since we show, in Subsection 5.2, that the impact of discretisation is of order $O(n^{-1/2})$. These results shed light on the asymptotic behaviour of lattice distributions and on the structure of discrete confidence intervals.

From a practical point of view, the corrected intervals also behave satisfactorily, in terms of their frequentist coverage. Simulations are proposed, in Subsection 5.1, for the Poisson and the binomial distributions, in order to compare our corrections with those already obtained in these two special cases.

2. On the existence of asymptotic expansions

In this section, we prove that when the model behaves like a nonlattice model, i.e. when hypothesis (H) is satisfied (see Theorem 1, below), there exists an asymptotic expansion to the order $o(n^{-1/2})$, which is fairly natural.

Let $f(X^n \mid \theta) = f_\theta(X^n)$ be the density of X^n with respect to some discrete measure. We note $l_n(\theta) = \log f(X^n \mid \theta)$. We assume that $X_1, \dots, X_n$ are independent and identically distributed. Let D denotes the differential operator with respect to θ; in particular, $D^t h$ represents the divergence of h, i.e. the sum of the derivatives of h. The following assumptions are standard in asymptotic expansions:

A1. $\log f_\theta(x)$ is 4 times differentiable in θ, and for all $\nu = (\nu_1, \dots, \nu_k) \in \mathbb{N}^k$ we have, if $|\nu| = \sum_{i=1}^k \nu_i = 3$,

$$E_\theta[|D^\nu \log f_\theta(X)|^3] < \infty,$$

and if $|\nu| = 4$, then $\forall K \subset \Theta$ compact, $\exists \epsilon > 0$ such that

$$\sup_{\theta \in K} E_\theta \left[\sup_{|\theta - \theta'| < \epsilon} |D^\nu \log f_{\theta'}(X)|^3 \right] < \infty.$$

We consider only the linearly independent components of $(D \log f_\theta(x), D^2 \log f_\theta(x))$ (as functions of x), so that their covariance matrix Σ is invertible.

A2. The Fisher information, $i(\theta)$, is positive definite and continuous.

A3. The prior π is positive and $\log \pi$ is twice continuously differentiable on Θ.
In the following let $\mathbb{R}$, $\mathbb{Q}$ and $\mathbb{Z}$ respectively denote, the sets of real numbers, rationals and integers. We assume that $D \log f_\theta(X)$ and $D^2 \log f_\theta(X)$ are random lattice vectors (matrix). Let $L_{(1)} = x_0 + \xi_1 \mathbb{Z} + \cdots + \xi_k \mathbb{Z}$ be the supporting lattice of $D \log f_\theta(X)$. For a matrix M, we denote $M^{i,j}$ the (i,j) component of its inverse, M^i the i-th row of its inverse and $M^{\cdot i}$ the i-th column of its inverse.

THEOREM 1. *When conditions* A1–A3 *hold, there exists a continuous asymptotic expansion of* $P_\theta^n[C_\alpha^\pi]$ *to the order* $o(n^{-1/2})$, *if and only if*, $\exists j_1, j_2 \le k$ *such that* $i(\theta)^1 \xi_{j_1} \neq 0$, *and*

$$\text{(H)} \qquad \frac{i(\theta)^1 \xi_{j_2}}{i(\theta)^1 \xi_{j_1}} \in \mathbb{R} - \mathbb{Q}.$$

The expansion is then the same as in the continuous case:

$$P_\theta^n[C_\alpha^\pi] = \alpha + \frac{\varphi(\Phi^{-1}(\alpha))}{\sqrt{n}} \left\{ \frac{i^1(\theta) D \log \pi(\theta)}{\sqrt{i^{1,1}(\theta)}} - D^t(i^1(\theta)/\sqrt{i^{1,1}(\theta)}) \right\} + o(n^{-1/2}).$$

Remark 2.1. This result implies that when (H) is satisfied, the frequentist coverage is equal to α to the order $o(n^{-1/2})$ if

$$[i(\theta)^{1,1}]^{-1/2} i(\theta)^1 D \log \pi - D^t\{i(\theta)^1 [i(\theta)^{1,1}]^{-1/2}\} = 0. \tag{2.1}$$

Peers (1965) gives further details on the form of the solutions of this partial differential equation.

The idea of the proof is to compare the probability $P_\theta^n[C_\alpha^\pi]$ with the formal expansion obtained in the continuous case. The differences appear near the boundaries of the confidence interval. See Appendix 1 for technical details.

Remark 2.2. Condition (H) implies that the leading term of the statistic related with the asymptotic expansion is discrete but nonlattice. Therefore it is natural that the order of the frequentist asymptotic expansion is the same as the order obtained when considering discrete nonlattice distributions. An interesting feature of this result is that it depends strongly on the parameterisation. To illustrate this, consider the following simple example.

Example 2.1. Let $(X_1, \dots, X_n)$ be drawn from a multinomial distribution $\mathcal{M}(n, p_1, p_2, p_3)$, with $\sum_{i=1}^3 p_i = 1$. The canonical parameterisaton is $p = (p_1, p_2)$. Then, condition (H) is not satisfied. However, if the parameter of interest is $\theta_1 = ap_1 + p_2$, with $a \in \mathbb{R} - \mathbb{Q}$ and the nuisance parameter remains p_2, then condition (H) is satisfied and equation (2.1) becomes, with $\pi(\theta) = e^{\psi(\theta)}\sqrt{i^{11}}$

$$(2.2) \qquad (a(\theta_1 - \theta_2) - \theta_1^2 + \theta_2)D_1\psi(\theta) + \theta_2(1 - \theta_1)D_2\psi(\theta) = -(a+1) + 3\theta_1.$$

There is no higher order expansion of the frequentist coverage of the one-sided confidence intervals, in particular when θ_1 is orthogonal to $(\theta_2, \dots, \theta_k)$ in the sense of Cox and Reid (1987), or in binomial and Poisson cases. In such cases, there is no answer to the problem of higher order asymptotics and in particular to the determination of a class of pertinent matching priors for such cases, unless some continuity corrections are added. Even though randomization contradicts the Likelihood Principle, it must be invoked to get an improvement in the frequentist coverage of these confidence intervals.

3. Randomized corrections and second order expansions

We propose in this section simple randomized corrections based on a uniform random variable, which lead to higher order expansions of the frequentist coverage of our intervals. Recall that from a frequentist point of view, randomization is necessary to reach a given level α for lattice distributions (Lehmann (1986)). Therefore we propose the following correction:

$$C_\alpha^\pi(U) = \{\theta_1 \leq k_n(\alpha) + J_n^1 U\},$$

where $U = \sum_{i=1}^k \xi_i U_i$ is based on k independent uniform random variables on $[-1/2, 1/2]$ independent of the observations, and J_n/n is an approximation of the Fisher information of the sample to the first order. It is then possible to obtain an asymptotic expansion of the frequentist as well as the posterior coverages of $C_\alpha^\pi(U)$ to the required order, i.e. $O(n^{-1})$. Moreover, we prove that this expansion is formally the same as in the continuous case. The term U can be seen as a 'smoothing' correction of the discrete $k_n(\alpha)$, so that $k_n(\alpha) + J_n^1 U$ approaches θ_1 to a higher order. The correction is necessarily of order $n^{-1/2}$, because it corresponds to the order of the 'jumps' in the likelihood, i.e. the distance between two values of the likelihood (θ being fixed). Note that to the first order, $C_\alpha^\pi(U) = C_\alpha^\pi$.

3.1 *Main result*

Set $J_n = -(D^2 l_n(\theta))_{\theta=\hat{\theta}}$. The expansions are expressed in terms of the maximum likelihood estimator, but we could use other consistent estimators such as Bayes estimators. This is in particular more coherent when considering confidence regions from a Bayesian point of view.

The corrected bound in $C^\pi_\alpha(U)$ then allows for an asymptotic expansion of the frequentist coverage of the Bayesian one-sided confidence region.

THEOREM 2. *If Θ is an open convex set and if the assumptions* A1–A3 *of Section* 2 *hold,*

$$(3.1)\quad P^n_\theta(\theta_1 < k_n(\alpha) + J^1_n U) = \alpha + \frac{\varphi(\Phi^{-1}(\alpha))}{\sqrt{n}} \left\{ \frac{i^1(\theta) D \log \pi(\theta)}{\sqrt{i^{1,1}(\theta)}} - D^t \left(i^1(\theta) / \sqrt{i^{1,1}(\theta)} \right) \right\} + O(n^{-1}),$$

where Φ and φ denote the standard normal distribution and its density w.r.t. the Lebesgue measure, respectively.

We thus obtain the same expansion as in the continuous case; see Peers (1965). The frequentist coverage is equal to α to the order $O(n^{-1})$ if and only if (2.1) is satisfied. In particular, if θ is real there is no asymptotic expansion of the uncorrected confidence interval and Jeffreys prior, $\pi(\theta) \propto \sqrt{i(\theta)}$, is the unique solution to (2.1) for the corrected one-sided interval $C^\pi_\alpha(U)$.

The posterior coverage of this corrected interval is still α to the order $O(n^{-1})$. Indeed,

$$\begin{aligned} P^\pi[\theta_1 < k_n(\alpha) + J^1_n U \mid X^n, U] \\ = P^\pi[\theta_1 < k_n(\alpha) \mid X^n] + J^1_n U \pi[k_n(\alpha) \mid X^n] + O_P(n^{-1}); \end{aligned}$$

taking the expectation with respect to U, leads to the required expansion.

This set is then, to the second order, a Bayesian confidence interval with coverage α and can be used as such. This implies in particular that there exists matching priors for corrected one-sided confidence intervals when the observations are random lattice vectors, and, moreover, that they satisfy the same formal differential equation as in the continuous case.

Consider, for instance, the binomial case where some bounds, for the one-sided intervals, have been studied (see Hall (1982), Blyth (1986)), using approximations to the order $O(n^{-1})$. Section 5 presents simulations to compare our results with theirs from a practical point of view. Theorem 1 implies that there is no expansion to the order $O(n^{-1})$ of $P^n_\theta[C^\pi_\alpha]$ and a randomized correction is needed.

Note that $X_1, \ldots, X_n$ need not be lattice random vectors. The only true requirement is that $D \log f_\theta(X)$ is a lattice random vector.

3.2 *Elements of proof*

The proof applies Babu and Singh's (1989) result to the derivatives of the log-likelihood. The expansions considered here are very similar to those of Welch and Peers (1963) and Peers (1965). Let

$$r(X^n, t_1) = \int_{\theta_1 \le t_1} \int_{\Theta_{k-1}} \pi[(\theta_1, \theta_2, \ldots, \theta_k) \mid X^n] d\theta_1 d\theta_2 \cdots d\theta_k,$$

where Θ_{k-1} denotes the set where $(\theta_2, \ldots, \theta_k)$ varies when $\theta_1 = t$ is fixed. We have

$$P_\theta^n[\theta_1 \le k_n(\alpha) + J_n^1 U] = P_\theta^n[r(X^n, \theta_1 - J_n^1 U) \le \alpha].$$

Considering the usual normal approximation of the posterior distribution and equation (19) in Welch and Peers (1963), we get

$$r(X^n, \theta_1 - J_n^1 U) = \Phi\left(z(X^n, \theta_1) - \frac{J_n^1 U}{\sqrt{J_n^{11}}}\right) + O(n^{-1}), \tag{3.2}$$

where $z(X^n, \theta_1) = \Phi^{-1}(r(X^n, \theta_1))$. Therefore,

$$P_\theta^n[C_\alpha^\pi(U)] = P_\theta^n[z(X^n, \theta_1) - J_n^1 U/\sqrt{J_n^{11}} \le \Phi^{-1}(\alpha)] + O(n^{-1}).$$

In the following, we omit θ in our notations whenever it does not induce ambiguities. Set $Z_{n,j}(\theta) = (D^j l_n(\theta) - E_\theta[D^j l_n(\theta)])/\sqrt{n}$, for $j = 1, 2$. Then, see Appendix 1, $z(X^n, \theta_1)$ can be expressed as, $z(X^n, \theta_1) = g(Z_n) + O_P(n^{-3/2})$, where Z_n is the largest sub-vector of $(Z_{n,1}, Z_{n,2})$ whose components are linearly independent as functions of X^n. Therefore, $z(X^n, \theta_1) - (J_n^{11})^{-1/2} J_n^1 U$ has exactly the same expansion as $z(X^n, \theta_1)$ in terms of the $Y_{n,j}$'s, where $Y_{n,1} = Z_{n,1} + U/\sqrt{n}$ and $Y_{n,2} = Z_{n,2} + U_{(2)}/\sqrt{n}$, and $U_{(2)}$ is a symmetric matrix such that the components of its upper triangular part are independent uniform random variables on $[-1/2, 1/2]$, independent of the observations and of U. Using Babu and Singh's result ((1989), Theorem 1) and considering $f(x) = \mathbb{I}_{g(x) \le \Phi^{-1}(\alpha)}$, where $\mathbb{I}$ represents the indicator function of a set, we finally establish (3.1).

Note that the above argument holds for any kind of likelihood-based confidence region, since this continuity correction allows for an expansion to the second order of the distribution of any statistic function of Z_n.

4. Randomized corrections and third order expansions

4.1 *Main result*

We now consider a higher order expansion for the frequentist coverage of this kind of one-sided confidence intervals. Babu and Singh's result (1989) does not hold any longer. However it is possible to find a class of functions $h_n(U)$ such that $P_\theta^n[\theta_1 + h_n(\bar{U}) \le k_n(\alpha)]$ has an asymptotic expansion. This leads to an asymptotic expansion of the frequentist coverage of the corrected confidence interval $C_\alpha^\pi(U) = \{\theta_1 + h_n(U) \le k_n(\alpha)\}$. We assume that the usual regularity conditions in the strongly non lattice case (see Bickel and Ghosh (1990)) hold for the density f_θ (w.r.t. the counting measure).

Let $\bar{U} = \sum_{i=1}^d \xi_i U_i$, where the U_i are uniform random variables on $[-1/2, 1/2]$; with $D^2 \log f_\theta(X) \in \sum_{i=k+1}^d \xi_i \mathbb{Z}$. Set T the linear operator such that $T Z_n = Z_{n,2}$ and $V = T\bar{U}$. In the following we note $\hat{b} = b(\hat{\theta})$ for any function b of θ and $\mu_{ijk} = E_\theta[D_{ijk} \log f_\theta(X)]$.

THEOREM 3. *If (H) holds and the U_i's are independent of the observations, then $P_\theta^n[\theta_1 + h_n(\bar{U}) \le k_n(\alpha)]$ has an asymptotic expansion to the order $o(n^{-1})$ if*

$$\begin{aligned} h_n(\bar{U}) = & -J_n^1 U_{(1)} + \frac{J_n^1 H_n(U)}{n^{1/2}} + \sqrt{n} J_n^1 V J_n^{-1} Z_{n,1} \\ & - \frac{\hat{\imath}^{1j} \hat{\mu}_{jkl} \hat{\imath}^{k} Z_{n,1} J_n^l U_{(1)}}{n^{1/2}} + \Phi^{-1}(\alpha) \frac{J_n^1 V J_n^{\cdot 1}}{2\sqrt{J_n^{11}}} \\ & + \frac{J_n^1 U_{(1)} \hat{\imath}^{1j} \hat{\imath}^{1k} \hat{\mu}_{jkl} J_n^l Z_{n,1}}{n^{1/2} J_n^{11}} - \Phi^{-1}(\alpha) \frac{\hat{\imath}^{1j} \hat{\imath}^{1k} \hat{\mu}_{jkl} J_n^l U_{(1)}}{2n\sqrt{J_n^{11}}} + O(n^{-2}), \end{aligned} \tag{4.1}$$

where $H_n(U)$ is a function of (Z_n, U), continuous in Z_n and differentiable in U, such that

$$\sum_{i=1}^{k} \int_{\bar{K}} \frac{\partial H_n^i(u)}{\partial u_i} du = 0 \tag{4.2}$$

and $E_\theta^n[|H_n(U)|] < \infty$ uniformly in U. The expansion of $P_\theta^n[C_\alpha^\pi(\bar{U})]$ is then

$$\begin{aligned} P_\theta^n[\theta_1 + h_n(\bar{U}) \le k_n(\alpha)] = \alpha &+ \frac{P_1(\theta,\alpha,\pi)}{\sqrt{n}} + \frac{P_2(\theta,\alpha,\pi)}{n} \\ &+ \sum_{j=1}^{k} (i^1\xi_j)^2 \frac{\varphi(\Phi^{-1}(\alpha))\Phi^{-1}(\alpha)}{8ni^{11}(\theta)} \\ &+ \varphi(\Phi^{-1}(\alpha)) \frac{R_n(\alpha,\theta,\pi,H_n)}{n} + o(n^{-1}), \end{aligned} \tag{4.3}$$

where P_1 and P_2 are the corresponding terms of the expansion in the continuous case and R_n is of order $O(1)$ and is due to H_n.

THEOREM 4. *If (H) is not satisfied, then $P_\theta^n[\theta_1 + h_n(\bar{U}) \le k_n(\alpha)]$ has an expansion to an order higher than $O(n^{-1})$ if $H_n(\bar{U})$ satisfies*

$$\sum_{i=1}^{k} \frac{\partial H_n^i(U)}{\partial U_i} = -U^t \hat{\Sigma}^{-1} x + O(n^{-1/2}), \tag{4.4}$$

where $H_n^i(U)$ is the i-th component of $H_n(U)$ in the basis $(\xi_1, \dots, \xi_d)$ and Σ is the asymptotic covariance matrix of Z_n. In this case the frequentist coverage of $C_\alpha^\pi(U)$ is

$$\alpha + \frac{P_1(\theta,\alpha,\pi)}{\sqrt{n}} + \frac{P_2(\theta,\alpha,\pi)}{n} + \sum_{j=1}^{k} (i^1\xi_j)^2 \frac{\varphi(\Phi^{-1}(\alpha))\Phi^{-1}(\alpha)}{24ni^{11}(\theta)} + O(n^{-3/2}). \tag{4.5}$$

The proof is given in Appendix 2.

Remark 4.1. The above conditions are close to being necessary conditions. In particular the U_i's need not be uniform random variables, but the density, $g_n(u)$ of $\bar{U}$ conditional on the observations must satisfy, see Appendix 2, for all t such that $-\sum_{j=1}^k |i^1\xi_j|/2 \le t \le \sum_{j=1}^k |i^1\xi_j|/2$,

$$\int_{\bar{K}} \mathbb{1}_{i^1 u_{(1)} \ge t}(g_n(u) - 1) du = 0.$$

In the case $k = 1$, this implies that lest U_1 be uniform on $[-1/2, 1/2]$, independent of the observations to the first order of approximation,

$$P_\theta^n[C_\alpha^\pi(\bar{U})] = \alpha + \frac{P_1(\theta,\alpha,\pi)}{\sqrt{n}} + s_n \frac{\gamma_2(\alpha,\theta,\pi)}{\sqrt{n}} + O(n^{-1}),$$

where s_n is a sequence of real numbers, which is dense in $[-1/2, 1/2]$.

Two particular choices of $H_n(U)$ are of interest:

(i) $H_n(U) = 0$. If $\pi(\theta)$ satisfies (2.1), i.e. π is a matching prior to the order $O(n^{-1})$, then, to the order $o(n^{-1})$, the frequentist coverage is equal to

$$P_\theta^n[C_\alpha^\pi(\bar{U})] = \alpha + \varphi(\Phi^{-1}(\alpha))\Phi^{-1}(\alpha)\frac{\bar{P}_2(\theta,\pi)}{n} + \varphi(\Phi^{-1}(\alpha))\Phi^{-1}(\alpha)\sum_{j=1}^{k}\frac{(i^1\xi_j)^2}{8ni^{11}},$$

where $\bar{P}_2(\theta,\pi)$ corresponds to the formal expansion (i.e. the continuous one).

(ii) $H_n(U) = H_{n,1}(U)$ is such that the frequentist and coverage is equal to α to the order $o(n^{-1})$, when π satisfies (2.1). There are many possible choices for $H_{n,1}(U)$, a simple one is for instance

$$H_{n,1}(U) = -\sqrt{n}(\theta_1 - \hat{\theta}_1)\sum_{j=1}^{k}\frac{h_j\hat{\xi}_j(U_j^2 - 1/4)}{2n} + \frac{\Phi^{-1}(\alpha)}{2n\sqrt{I_n^{11}}}\left(\bar{P}_2(\hat{\theta},\pi) + \sum_{j=1}^{k}\frac{(I_n^1\hat{\xi}_j)^2}{12I_n^{11}}\right)e_1 \tag{4.6}$$

where e_1 is the vector whose componants are $(1, 0, \ldots, 0)$ in the canonical basis of $\mathbb{R}^k$, $I_n = -D^2 l_n(\hat{\theta})/n$ and h_j is given by

$$\sum_{j=1}^{k} h_j I_n^1\hat{\xi}_j = -\sum_{j=1}^{k}\frac{(I_n^1\hat{\xi}_j)^2}{I_n^{11}} - 6\bar{P}_2(\hat{\theta},\pi).$$

This third order expansion sheds light on the asymptotic behaviour of likelihood based statistics for discrete distributions, see Appendix 2.

4.2 *Posterior coverage and matching priors*

Depending on the choice of $H_n(U)$, the posterior coverage of $C_\alpha^\pi(U)$ differs from α by a term of order $O(n^{-1})$.

PROPOSITION 1. *Under the same assumptions as in Theorem* 3,

$$P^\pi[C_\alpha^\pi(U) \mid X^n] = \alpha - \sum_{j=1}^{k}\frac{(\hat{\imath}^1\hat{\xi}_j)^2}{24J_n^{11}}\varphi(\Phi^{-1}(\alpha))\Phi^{-1}(\alpha) + \frac{\varphi(\Phi^{-1}(\alpha))\Gamma_n(H_n,\pi)}{n} + O(n^{-3/2}), \tag{4.7}$$

where Γ_n *is due to* $H_n(U)$.

This result is proved by first approximating $P^\pi[C_\alpha^\pi(U) \mid X^n, U]$, using the characteristic function and integrating over $\bar{U}$ this approximation.

There is no explicit form for Γ_n since $H_n(\bar{U})$ can be any continuous integrable function of Z_n.

COROLLARY 1. *When $H_n(U)$ is given by (4.6), and when π satisfies (2.1), the posterior and the frequentist coverages of the confidence interval are both equal to α to the order $o(n^{-1})$:*

$$P^\pi[C_\alpha^\pi(U) \mid X^n] = P_\theta^n[C_\alpha^\pi(U)] + O(n^{-3/2}) = \alpha + O(n^{-3/2}).$$

In the continuous case, studied by Welch and Peers (1963), no prior matches both coverages to the order $o(n^{-1})$, except for some special distributions. In the discrete case, our corrections imply matching priors to the order $o(n^{-1})$ as soon as (H) is satisfied. Therefore these corrections allow for greater accuracy.

Note that, in some cases, the corrected confidence interval can be in the form

$$\theta_1 + \frac{f_n(Z_{n,1}(\theta), U)}{n^{3/2}} \leq k_n(\alpha, \bar{U}), \qquad \text{where} \quad f_n(Z_{n,1}(\theta), U) = O_p(1).$$

This implies that, to obtain matching priors in the discrete case, it is, in some cases, necessary to slightly change the structure of the confidence region. This result is of interest, since it sheds light on the effects of discretization. From a practical point of view, however, it is better to approximate this correction term, so that the correction does not depend on the nuisance parameter any longer. A possibility is to approximate $Z_{n,1}(\theta)$ by a term in the form $Z_{n,1}(\tilde{\theta}(\theta_1)) + J_n\sqrt{n}(\tilde{\theta}(\theta_1) - \hat{\theta})$, where $\tilde{\theta}(\theta_1)$ is the constraint maximum likelihood estimator.

4.3 *Standard distributions*

We now consider two examples to illustrate the corrections of Sections 3 and 4.

4.3.1 *The binomial distribution*

Consider $X^n \sim B(n,p)$. Then $\hat{p} = X^n/n$, the Bayes estimator is $\delta^\pi = (X+0.5)/(n+1)$ when $\pi(p) \propto \sqrt{i(p)}$ and $\xi_1 = i(p)$. The correction term is

$$h_n(U) = -\frac{U_1}{n} - \frac{U_1(1-2\delta^\pi)\Phi^{-1}(\alpha)}{2n^{3/2}\sqrt{\delta^\pi(1-\delta^\pi)}} + (U_1^2 - 1/4)\frac{\Phi^{-1}(\alpha)}{2n^{3/2}\sqrt{\delta^\pi(1-\delta^\pi)}}.$$

The frequentist coverage of $C_\alpha^\pi(U)$ is then $P_\theta^n[C_\alpha^\pi(U)] = \alpha + O(n^{-3/2})$, while $P^\pi[C_\alpha^\pi(U) \mid X^n]$ is not equal to α, but to

$$P^\pi[C_\alpha^\pi(U) \mid X^n] = \alpha + \frac{\varphi(\Phi^{-1}(\alpha))\Phi^{-1}(\alpha)}{24n\hat{p}(1-\hat{p})} + O(n^{-3/2}).$$

The Poisson distribution has a similar behaviour.

4.3.2 *The multinomial distribution* (see Example 2.1)

If the parameter of interest is p_1 and the nuisance parameter is $(1-p_1)/p_2$, (H) is not satisfied. For any prior in the form $\pi(\theta_1, \theta_2) = [\theta_1(1-\theta_1)]^{-1/2}g(\theta_2)$, both the frequentist and the Bayesian coverages of the corrected confidence interval $C_\alpha^\pi(U) = \{\theta_1 \leq k_n(\alpha) + U_1/n\}$, where U_1 is uniform on $[-1/2, 1/2]$, are equal to α to the order $O(n^{-1})$. The following continuity correction allows for an expansion to the order $O(n^{-3/2})$ of its frequentist coverage.

$$\begin{aligned} h_n(U) = & -\frac{U_1}{n} - \frac{U_1^2 - 1/4}{2n^{3/2}(1-x_1/n)}\left\{\sqrt{n}(x_1/n - \theta_1) - \frac{\sqrt{n}(1 - x_1/n - \theta_2 x_2/n)}{(x_2/n)(x_3/n)}\right\} \\ & - \frac{U_1\sqrt{n}(x_1/n - \theta_1)(1-2x_1/n)}{n^{3/2}} \\ & - 3\frac{\Phi^{-1}(\alpha)U_1}{2n^{3/2}\sqrt{x_1/n(1-x_1/n)}}(1 - 2x_1/n(1-x_1/n)). \end{aligned}$$

5. Discussion

5.1 *Simulations*

The above corrected intervals have a theoretical interest since they allow for the determination of matching priors for discrete distributions and help in understanding the asymptotic behaviour of discrete distributions; see Subsection 4.1.

From a more practical point of view, we now report simulations in the binomial and in the Poisson cases and compare the frequentist coverages of our intervals with those of Hall (1982) and Blyth (1986), which are not randomized (with error of order $O(n^{-1/2})$).

We have simulated samples of size n for both distributions and derived the corresponding frequentist coverages based on $q = 10000$ replications. In these computations Welch and Peers' (1963) bound, as well as ours, are related to Jeffreys prior and expressed in terms of the associated Bayesian estimate.

Table 1 gives the bias between the observed and the nominal frequentist coverages for (a) the classical Welch and Peers one-sided interval, (b) our corrected interval to the order $O(n^{-1})$ ($H_n(U) = 0$), (c) our corrected interval to the order $O(n^{-3/2})$, (d) Hall's interval and (e) Blyth's intervals (A), (B), (C), as defined in Blyth (1986), for $n = 100$, 1000 and for various values of the parameter p, in the binomial case. Table 2 provides the same bias, for Welch and Peers' one-sided interval, ours (to orders $O(n^{-1})$ and $O(n^{-3/2})$) and Hall's in the Poisson case.

Both tables show that our bounds behave in a way similar to the alternatives. We note that in both cases, the bias of our intervals decrease faster than the others, but are greater for small values of n. This is due to the fact that the intercepts in the error terms are greater in our bounds. Therefore, the corrections proposed above have not only a theoretical interest, but have also good performances in practice.

5.2 *Discretisation and continuity correction*

When the observations are linked to a continuous process, the continuous correction is not so arbitrary any longer and Theorem 1 sheds some light on the distortion discretisation implies on inference. Let $Y^n = (Y_1, \ldots, Y_n)$ be a sample with density $f_\theta(Y^n)$ with respect to Lebesgue measure, consider $X^n = (X_1, \ldots, X_n)$ to be a discretised version of Y^n, i.e. $X_i = g(Y_i)$, where g takes its values on a discrete set, for instance $g(x) = \mathbb{I}_{x \leq s}$, where $s \in \mathbb{R}$ is fixed. Let g_θ be the density of X with respect to some discrete measure. Under the usual regularity conditions, when the prior $\pi \propto \sqrt{i(\theta)}$,

$$P_\theta^n[\theta \leq k_n(\alpha, Y^n)] = \alpha + O(n^{-1}).$$

The confidence one-sided interval can be approximated by: $C_\alpha(Y^n) = \{h_{n,1}(Y,\theta) \leq \Phi^{-1}(\alpha)\}$, and in the same way, the corrected discrete confidence set, based on X^n, can be expressed as: $C_\alpha(U, X^n) = \{h_{n,2}(X,\theta) - U\xi_1/\sqrt{n} \leq \Phi^{-1}(\alpha)\}$, to the order n^{-1}, where $h_{n,1}$ and $h_{n,2}$ have the same form. Since $P_\theta^n[C_\alpha(Y^n)] = P_\theta^n[C_\alpha(U, X^n)] = \alpha$, to the order n^{-1}, for all $\alpha \in (0,1)$, we have

$$h_{n,1}(Y,\theta) = h_{n,2}(X,\theta) - U\xi_1/\sqrt{n} + O(n^{-1}), \qquad \text{in distribution.}$$

Therefore discretisation implies an error of order $O(n^{-1/2})$ on the bounds of confidence intervals which can be modeled by a uniform random variable on $[-\xi_1/(2\sqrt{n}), \xi_1/(2\sqrt{n})]$. This result is quite close to the one obtained by Kolassa and McCullagh (1990) on the rounding effect, where they prove that it can asymptotically be modeled by a uniform random variable.

Table 1. Bias in the coverage for different one-sided confidence intervals in the binomial case.

	p	W.-P.	order 1	order 3/2	Hall	Blyth(A)	Blyth(B)	Blyth(C)
	0.1	−0.0667	−0.0667	−0.0654	−0.0076	−0.0076	−0.0076	+0.0272
	0.2	−0.0331	−0.0331	−0.0339	−0.0331	−0.0001	−0.0001	−0.0001
	0.3	−0.0273	−0.0227	−0.0217	−0.0273	+0.0004	+0.0004	+0.0004
$n = 100$	0.4	−0.0096	−0.0096	−0.0098	−0.0096	−0.0096	+0.0109	+0.0109
	0.5	−0.0190	−0.0084	−0.0073	+0.0045	+0.0045	+0.0045	+0.0045
	0.6	+0.0085	−0.0008	+0.0000	+0.0085	−0.0130	+0.0085	+0.0085
	0.7	−0.0019	+0.0037	+0.0043	−0.0019	−0.0019	+0.0181	+0.0181
	0.8	+0.0184	+0.0110	+0.0120	+0.0184	−0.0037	+0.0184	+0.0184
	0.9	+0.0121	+0.0201	+0.0209	+0.0121	−0.0222	+0.0295	+0.0121
	0.1	−0.0096	−0.0100	−0.0102	−0.0096	+0.0005	+0.0005	+0.0005
	0.2	−0.0091	−0.0045	−0.0041	+0.0006	+0.0006	+0.0006	+0.0080
	0.3	−0.0018	−0.0043	−0.0043	−0.0018	−0.0018	−0.0018	+0.0054
$n = 1000$	0.4	+0.0019	−0.0002	−0.0002	+0.0019	+0.0019	+0.0019	+0.0019
	0.5	−0.0021	+0.0008	+0.0008	−0.0021	−0.0021	+0.0036	+0.0036
	0.6	−0.0006	+0.0012	+0.0012	−0.0006	−0.0006	+0.0065	+0.0065
	0.7	+0.0038	+0.0027	+0.0028	+0.0038	−0.0040	+0.0038	+0.0038
	0.8	+0.0061	+0.0064	+0.0063	+0.0061	−0.0029	+0.0129	+0.0061
	0.9	+0.0076	+0.0084	+0.0083	+0.0076	−0.0021	+0.0174	+0.0076

Table 2. Bias in the coverage for different confidence intervals in the Poisson case.

$n = 100$					$n = 1000$				
λ	W.-P.	order 1	order 3/2	Hall	λ	W.-P.	order 1	order 3/2	Hall
0.5	−0.0132	−0.0043	−0.0011	+0.0033	0.5	+0.0002	+0.0004	−0.0002	+0.0002
1.0	−0.0061	−0.0033	−0.0011	+0.0054	1.0	−0.0003	+0.0003	−0.0000	−0.0003
1.5	−0.0009	−0.0009	−0.0007	−0.0009	1.5	−0.0024	−0.0024	−0.0026	−0.0024
2.0	−0.0043	−0.0041	−0.0030	−0.0043	2.0	−0.0007	−0.0010	−0.0011	−0.0007
2.5	−0.0034	−0.0051	−0.0039	−0.0034	2.5	+0.0017	+0.0023	+0.0021	+0.0017
3.0	−0.0024	−0.0051	−0.0039	−0.0038	3.0	−0.0008	−0.0005	−0.0006	−0.0008
3.5	−0.0025	−0.0111	−0.0032	−0.0025	3.5	+0.0029	+0.0016	+0.0015	+0.0029
4.0	−0.0008	−0.0035	−0.0020	−0.0008	4.0	−0.0006	−0.0006	−0.0006	−0.0006
4.5	+0.0019	+0.0006	+0.0011	+0.0019	4.5	+0.0040	+0.0030	+0.0029	+0.0040
5.0	+0.0000	−0.0026	−0.0021	+0.0000	5.0	+0.0013	+0.0008	+0.0007	+0.0013

Acknowledgements

The author is grateful to Professor C. P. Robert for many valuable comments. She also thanks Professors G. Casella, T. J. DiCiccio and M. T. Wells for many helpful discussions and their warm hospitality, as well as the referees for their remarks. The visit at Cornell University which initiated this work was partially supported by a CNRS-NSF grant.

Appendix 1: Proof of Theorem 1

We consider the same notations as in Subsection 3.2. $g(Z_n)$ has the form

$$g(Z_n) = -\frac{i(\theta)^1 Z_{n,1}}{\sqrt{i^{1,1}(\theta)}} + \frac{A(Z_{n,2})^t Z_{n,1}}{\sqrt{n}} + \frac{Z_{n,1}^t B_1 Z_{n,1}}{\sqrt{n}} - \frac{\delta^\pi}{\sqrt{n}} + \frac{M_n(Z_n)}{n}, \tag{A.1}$$

where $A(Z_{n,2})$ is linear in $Z_{n,2}$, B is a matrix of order $O(1)$, δ^π is a constant and M_n is polynomial in Z_n. Let $q_n(x)$ be the density of the formal Edgeworth expansion of Z_n, with $\tilde{L}_n = \tilde{L}/\sqrt{n}$ the supporting lattice of Z_n, i.e. $\tilde{L} = \sum_{j=1}^d \xi_j \mathbb{Z} + \tilde{x}_o$. Set $C(x) = \{y = y_1\xi_1 + \cdots + y_d\xi_d; -1/(2\sqrt{n}) < y_i - x_i \le 1/(2\sqrt{n})\}$. We compare $P_\theta^n[C_\alpha^\pi]$ with its formal expansion:

$$\begin{aligned}\Delta &= \sum_{x\in\tilde{L}_n} f(x)p_n(x) - \int_{\mathbb{R}^d} f(y)q_n(y)dy = \sum_{x\in\tilde{L}_n}\int_{C(x)}\left(\frac{f(x)p_n(x)}{\lambda(C(x))} - f(y)q_n(y)\right)dy \\ &= \sum_{x\in\tilde{L}_n} f(x)\int_{C(x)}(q_n(x)-q_n(y))dy + \sum_{x\in\tilde{L}_n}\int_{C(x)} q_n(y)(f(x)-f(y))dy + O(n^{-3/2}).\end{aligned}$$

We have

$$B(x) = \int_{C(x)} q_n(y)(f(x)-f(y))dy \neq 0, \quad \Leftrightarrow$$

$$\Phi^{-1}(\alpha) - \sum_{i=1}^k \frac{|i(\theta)^1\xi_i|}{2\sqrt{n i^{11}(\theta)}} + O(n^{-1}) \le g(x) \le \Phi^{-1}(\alpha) + \sum_{i=1}^k \frac{|i(\theta)^1\xi_i|}{2\sqrt{n i^{11}(\theta)}} + O(n^{-1}).$$

Let δ be the vector whose coordinates are $\delta_r = i(\theta)^1\xi_r$. Then, see Bhattacharya and Rao (1986),

$$\begin{aligned}\sum_{x\in\tilde{L}_n} B(x) = &\int_{\mathbb{R}^d}\int_{[-1/2,1/2]^d} q_n(x)dxdt \\ &\times\left(-\mathbb{1}_{0\le g(x)-\Phi^{-1}(\alpha)\le -\delta^t t/\sqrt{ni^{11}}+h_n/n} + \mathbb{1}_{0\ge g(x)-\Phi^{-1}(\alpha)\ge \delta^t t/\sqrt{ni^{11}}+h_n/n}\right) \\ &+\sum_{r=1}^d n^{-(d-r+1)/2}\sum_{L_{r+1}}\cdots\sum_{L_d}\int_{\mathbb{R}^r}\int_{[-1/2,1/2]^d} d_r[S_1(\sqrt{n}x_r)\varphi_\Sigma(x)]dx_{(\sim r)}dt \\ &\times\left(-\mathbb{1}_{0\le g(x)-\Phi^{-1}(\alpha)\le -\delta^t t/\sqrt{ni^{11}}+h_n/n} + \mathbb{1}_{0\ge g(x)-\Phi^{-1}(\alpha)\ge \delta^t t/\sqrt{ni^{11}}+h_n/n}\right) \\ &+\frac{R_n}{n},\end{aligned}$$

where $S_1(y)$ is a real function, discontinuous at every integer, defined by Bhattacharya and Rao ((1986), Appendix 2), d_r denotes the differentiation with respect to the r-th component of x, $dx_{(\sim r)} = \prod_{l<r} dx_l$, h_n is of order $O(1)$ and φ_Σ denotes the density of $\mathcal{N}(0,\Sigma)$. If $r > k$, the term

$$\begin{aligned}I_r = n^{-(d-r+1)/2}&\sum_{L_{r+1}}\cdots\sum_{L_d}\int_{\mathbb{R}^r}\int_{[-1/2,1/2]^d} \\ &\times \mathbb{1}_{0\le g(x)-\Phi^{-1}(\alpha)\le -\delta^t t/\sqrt{ni^{11}}+h_n/n}\, d_r[S_1(\sqrt{n}x_r)\varphi_\Sigma(x)]dx_{(\sim r)}dt\end{aligned}$$

is of order $O(n^{-1})$. If $\forall (r,s) \le k$ such that $i(\theta)^1\xi_r \neq 0$, $\delta_s/\delta_r \in \mathbb{Q}$, then the the following transformation, $y_j = \sum_{k=1}^k i(\theta)^1\xi_k x_k$, $y_k = x_k$, $k \neq j$, takes its values on a lattice, where $\delta_j \neq 0$. Therefore, calculations lead to

$$\sum_{x\in \bar{L}_n} B(x) = \sum_{y\in \bar{L}_n} B'(y) = \sum_{r=1}^{d} I_r + \frac{R_n}{n}$$

with $I_r = O(n^{-1})$ if $r \neq j$, and $I_j = H_n/\sqrt{n}$, where $H_n = O(1)$ has no limit, when n goes to infinity.

If $\exists s \neq j$ such that $\delta_s/\delta_j \in \mathbb{R} - \mathbb{Q}$, then there exists a continuous expansion to the order $o(n^{-1/2})$. Indeed, suppose without lack of generality that $s > j$ and $\forall r < j$, $\delta_r = 0$. Set $e(t) = \delta^t t - \lfloor \delta^t t \rfloor$, $e_j(t) = e(t)/\delta_j - \lfloor e(t)/\delta_j \rfloor$ and F_j the distribution function of $e_j(t)$. Then

$$\begin{aligned} I_j = &- n^{-(d-j+1)/2} \sum_{\underline{p}\in Z^{d-j}} \int_{\mathbb{R}^j} \varphi_\Sigma(\underline{p}/\sqrt{n}, t_j(\underline{p}/\sqrt{n}))\mathbb{1}_{\delta^t t\ge 0} \\ &\times \Big[E[e_j(t)] - 1 + F_j(\alpha_n + \beta(\underline{p}_{(1)}) - \underline{p}^t B_1 \underline{p} - \lfloor \sqrt{n} t_j(\underline{p}/\sqrt{n}) - \bar{p}^t B_1 \bar{p}/n \rfloor) \\ &\quad - F_j(1 - \alpha_n - \beta(\underline{p}_{(1)}) + \underline{p}^t B_1 \underline{p} + \lfloor \sqrt{n} t_j(\underline{p}/\sqrt{n}) - \bar{p}^t B_1 \bar{p}/n \rfloor) + \frac{R_n}{n} \Big], \end{aligned}$$

where $\alpha_n = \sqrt{n}\Phi^{-1}(\alpha) - \lfloor \sqrt{n}\Phi^{-1}(\alpha) \rfloor$, and $\beta(\underline{p}_{(1)}) = \sum_{i\neq s}^k \delta_i p_i - \lfloor \sum_{i\neq s}^k \delta_i p_i \rfloor$. Considering the same kind of calculations as Pólya and Szegö, ((1972), Chapter 4), we obtain, that $\forall k > 0$,

$$\frac{1}{n^{-(k-j)/2}} \sum_{|\underline{p}_{(1)}|\le n} \varphi_\Sigma(\underline{p}_{(1)}/\sqrt{n} \mid X) e^{2i\pi k \underline{p}^t B_1 \underline{p}} e^{2i\pi k \underline{p}_{(1)}} = 0.$$

Therefore

$$\begin{aligned} \int_0^1 F_j(t)dt = &\lim_{n\to\infty} n^{-(k-j)/2} \sum_{\underline{p}_{(1)}\in \mathbb{Z}^{k-j}} \varphi_\Sigma(\underline{p}_{(1)} \mid k_r, r \notin (j,\ldots,k)) \\ &\times F_j(\alpha_n + \beta(\underline{p}_{(1)}) - \underline{p}^t B_1 \underline{p} - \lfloor \sqrt{n} t_j(\underline{p}/\sqrt{n}) - \bar{p}^t B_1 \bar{p}/n \rfloor) \\ = &E[e_j(y)\mathbb{1}_{\delta^t t > 0}], \end{aligned}$$

and Theorem 1 is proved.

Appendix 2: Proofs of Theorems 3 and 4

The idea of the proof is to replace, in expansion (25) of Peers (1965), $Z_{n,i}$ by $Y_{n,i} = Z_{n,i} + W_{(i)}/\sqrt{n}$, where $W_{(i)}$ are continuous random vectors whose distribution is given below. It leads to

$$\begin{aligned} \bar{Z}(X,\theta) &= g(Y_n) \\ &= z(X,\theta) - \frac{i(\theta)^1 W_{(1)}}{\sqrt{n i^{1,1}}} + \frac{A(W_{(2)})^t B_2 Z_{n,1}}{n\sqrt{i^{1,1}}} \\ &\quad + \frac{A(Z_{n,2})^t B_2 W_{(1)}}{n\sqrt{i^{1,1}}} + \frac{W_{(1)}^t B_1 Z_{n,1} + Z_{n,1}^t B_1 W_{(1)}}{n\sqrt{i^{1,1}}} + O(n^{-3/2}). \end{aligned}$$

Let P_n be the distribution of Z_n and F_n the distribution of $W/\sqrt{n}$ conditional on $Z_n = x$, with density h_n. Let $f(x) = \mathbb{I}_{g(x)\leq\Phi^{-1}(\alpha)}$. In the following, we note M_n any quantity of order $O(1)$ which can be neglected. We choose H_n so that $\int f d(P_n * F_n - Q_n)$ can be approximated to a higher order, where Q_n denotes the formal Edgeworth expansion of a sum of continuous iid vectors with density q_n. Set $\tilde{K} = \{t = \sum_{i=1}^{d} \xi_i t_i ; |t_i| < 1/2\}$. $\lambda(\tilde{K})$ denotes its Lebesgue measure. Then

$$\int f d(P_n * F_n - Q_n) = O(n^{-3/2}) + \sum_{x\in L_n} \int_{C(x)} f(y)\{q_n(x)h_n(y-x)\lambda(\tilde{K})n^{-d/2} - q_n(y)\}dy.$$

Two cases occur:

(i) $\forall t \in \tilde{K}$, $g(x + t/\sqrt{n}) \leq \Phi^{-1}(\alpha)$, which implies that

$$\begin{aligned} B(x) &= \int_{C(x)} f(y)\{q_n(x)h_n(y-x)\lambda(\tilde{K})/n^{3/2} - q_n(y)\}dy \\ &= -n^{-d/2}\left(\frac{1}{24n}\sum_{i=1}^{d} \xi_i^t q_n''(x)\xi_i + O(n^{-3/2})\right). \end{aligned}$$

(ii) $f(x + t/\sqrt{n})$ varies when t runs the set $\tilde{K}$, for which

$$B(x) = n^{-d/2}\int_{\tilde{K}} \mathbb{I}_{i(\theta)^1 u_{(1)}\geq t_{n,1}} \left\{q_n(x)(g_n(u)\lambda(\tilde{K}) - 1) - \frac{u^t q_n'(x)}{\sqrt{n}} - \frac{u^t q_n''(x)u}{2n}\right\} du.$$

Set B as the set of such x's and $t_{n,1} = \sqrt{n i^{1,1}}(-\Phi^{-1}(\alpha) + g(x))$. To get a higher order expansion than in the pure lattice case (see Theorem 1), it is necessary that

$$\int_{\tilde{K}} \mathbb{I}_{i(\theta)^1 u_{(1)}\geq t_{n,1}} (g_n(u)\lambda(\tilde{K}) - 1)du = O(n^{-1/2}).$$

Therefore, if

$$g_n(u)\lambda(\tilde{K}) = 1 + \frac{f_n(u)}{\sqrt{n}},$$

with $\int_{\tilde{K}} f_n(u)du = 0$, then

$$B(x) = n^{-(d+1)/2} q_n(x)\int_{\tilde{K}} \mathbb{I}_{i(\theta)^1 u_{(1)}\geq t_{n,1}} \left\{f_n(u) - u^t j_n'(x) + \frac{M_n}{\sqrt{n}}\right\} du,$$

where $j_n'(x) = q_n'(x)/q_n(x)$. Considering Theorem 1, if (H_2) is satisfied, then

$$\sum_{x\in B} B(x) = \int_{\mathbb{R}^d}\int_{\tilde{K}} \varphi_\Sigma(x)\mathbb{I}_{B(t)}(x)(f_n(u \mid x) - u^t j_n'(x))dudx + o(n^{-1}).$$

If (H_2) does not hold, there is no asymptotic expansion to orders higher than $O(n^{-1})$, unless

$$\int_{\tilde{K}} \mathbb{I}_{i(\theta)^1 u_{(1)}\geq t_{n,1}} \left\{f_n(u) - u^t j_n'(x) + \frac{M_n}{\sqrt{n}}\right\} dt = O(n^{-1/2}).$$

This is satisfied in particular if

$$f_n(u) = \sum_{i=1}^{k} u_i \xi_i^t j_n'(x).$$

Besides, if the conditional density of W with respect to Lebesgue measure is

$$g_n(w) = \lambda(\tilde{K})^{-1}\mathbb{1}_{\tilde{K}}(w)\left(1 + \frac{f_n(w)}{\sqrt{n}}\right),$$

then, $U = W(1 + H_n(W)/\sqrt{n})$, where

$$\sum_{i=1}^{d} \frac{\partial H_n^i(u)}{\partial u_i} = f_n(u)$$

is a random vector whose coordinates in the basis $(\xi_1, \ldots, \xi_d)$ are $\mathcal{U}[-1/2, 1/2]$ r.v.'s.

References

Babu, J. G. and Singh, K. (1989). Note on Edgeworth expansion for lattice distributions, *J. Multivariate Anal.*, **30**, 27–33.

Bickel, P. and Ghosh, J. K. (1990). A decomposition for the likelihood ratio statistic and the Bartlett correction—A Bayesian argument, *Ann. Statist.*, **18**, 1070–1090.

Bhattacharya, R. and Rao, R. (1986). *Normal Approximations and Asymptotic Expansions*, 2nd ed., Wiley, New York.

Blyth, C. R. (1986). Approximate binomial confidence limits, *J. Amer. Statist. Assoc.*, **81**, 843–855.

Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference, *J. Roy. Statist. Soc. Ser. B*, **49**, 1–39.

Davison, A. C. (1988). Approximate conditional inference in generalized linear models, *J. Roy. Statist. Soc. Ser. B*, **50**, 445–461.

Ghosh, J. K. and Mukerjee, R. (1993). Frequentist validity of highest posterior density regions in the multiparameter case, *Ann. Inst. Statist. Math.*, **45**, 293–302.

Hall, P. (1982). Improving the normal approximation when constructing one-sided confidence intervals for binomial or Poisson parameters, *Biometrika*, **69**(3), 647–652.

Kolassa, J. E. and McCullagh, P. (1990). Edgeworth series for lattice distributions, *Ann. Statist.*, **18**, 981–985.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, Wiley, New York.

Mukerjee, R. and Dey, D. K. (1993). Frequentist validity of posterior quantiles in the presence of a nuisance parameter: Higher order asymptotics, *Biometrika*, **80**, 499–505.

Mukerjee, R. and Ghosh, M. (1997). Second order probability matching priors, *Biometrika* (to appear).

Pólya, G. and Szegö, G. (1972). *Problems and Theorems in Analysis*, Vol. 1, Springer, New York.

Peers, W. H. (1965). On confidence points and Bayesian probability points in the case of several parameters, *J. Roy. Statist. Soc. Ser. B*, **27**, 9–16.

Severini, T. A. (1993). Bayesian estimates which are also confidence intervals, *J. Roy. Statist. Soc. Ser. B*, **55**, 533–540.

Tibshirani, R. (1989). Noninformative priors for one parameters of many, *Biometrika*, **76**, 604–608.

Welch, B. L. and Peers, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods, *J. Roy. Statist. Soc. Ser. B*, **25**, 318–329.