Journal of Multivariate Analysis **83**, 1–21 (2002)
doi:10.1006/jmva.2001.2037

Asymptotic Properties of HPD Regions in the Discrete Case

Judith Rousseau

Laboratoire de Statistique, CREST; and Laboratoire de Statistiques Théoriques et Appliquées, Paris VI, Paris, France

Received March 6, 1998; published online January 11, 2002

This paper obtains asymptotic expansions of the frequentist distributions of modified likelihood ratio statistics when the observations are discrete. An upper bound of the uncertainty due to the discrete nature of the observations is obtained, which is slightly larger than Yarnold's result (obtained in the case of elliptic confidence regions). Higher order results are also derived from continuity corrections. They are instrumental in the determination of matching priors for HPD regions to the orders $o(n^{-1})$ and $O(n^{-3/2})$.

AMS 1991 subject classifications: 62E20; 62F25; 62H10.

Key words and phrases: asymptotic expansion; Bartlett correction; continuity correction; lattice distribution; likelihood ratio statistic.

1. INTRODUCTION

Asymptotic coverage properties of confidence regions have been widely studied, from both a frequentist and a Bayesian point of view. An interesting issue in such studies is the determination of classes of matching priors associated with those confidence regions. Welch and Peers (1963), Peers (1965, 1968), Nicolaou (1993), Tibshirani (1989), and Dey and Mukerjee (1993) have studied one-sided and two-sided intervals. DiCiccio and Stern (1993, 1994) and Ghosh and Mukerjee (1993) have studied likelihood based regions, i.e. different types of highest modified likelihood confidence regions, with and without nuisance parameters. These studies are mainly concerned with the frequentist and Bayesian asymptotic distributions of modified likelihood ratio statistics. The above authors show that these statistics have frequentist and Bayesian Bartlett corrections to the order $O(n^{-3/2})$ or $O(n^{-2})$. All these studies are restricted to nondiscrete observations.

When the observations are not discrete, i.e. when they satisfy the Cramer condition, the methods used in these studies for the expansion of likelihood based statistics are now well established. They are based on Laplace

0047-259X/02 $35.00

approximations for Bayesian coverages and on Edgeworth or saddlepoint continuous expansions for frequentist coverages. When the observations are discrete random variables, however, there is no continuous Edgeworth or saddlepoint expansion, as shown by Bhattacharya and Rao (1986). Therefore, when the observations are discrete, we cannot apply the results on the asymptotic expansions of the frequentist coverages of confidence regions obtained in the strongly nonlattice case. Few results exist in discrete setups. Yarnold (1972) obtains an expansion of the frequentist probability that a sum of iid lattice random vectors, $X \in \mathbb{R}^p$, belongs to smooth convex sets of the form

$$C = \{X^t \Sigma^{-1} X \leqslant c\},$$

where Σ is the covariance matrix of X and $c > 0$. He proves that the probability of C is equal to the cth quantile of a chi-square random variable with p degrees of freedom, $Q_p(c)$, plus a term proportional to the difference between the number of lattice points in C and its volume, when properly renormalized. This term was proved to be of order $O(n^{-p/(p+1)})$ by Esseen (1945). Bentkus and Götze (1995) improve on this result by showing that this term is in fact of order $O(n^{-1})$ when the dimension of X is large enough, i.e. larger than 9. Siotani and Fujikoshi (1984) obtain the same kind of result as Yarnold (1972) in the case of HPD regions for multinomial observations, but do not exhibit the order of their error term. Frydenberg and Jensen (1989) show, through simulations, that the Bartlett corrections do not improve the accuracy of the chi-square adjustment to the frequentist distribution of the likelihood ratio statistic for multinomial observations.

In this paper we study HPD regions, with and without nuisance parameters. These regions are important Bayesian confidence regions, in particular from a decision theoretic point of view; see Robert (1994). We establish, in Section 2, the existence of a continuous expansion of the frequentist distribution of the posterior likelihood ratio statistic in the general case, to orders smaller than $O(n^{-1/2})$. These orders depend on the dimension of the parameter of interest p, i.e. if $p \geqslant 3$ we prove that the uncertainty due to the discrete nature of the observations can be bounded by a term of order $O(n^{-p/(p+2)} \log n)$. To get higher order expansions in order to make higher order studies such as the determination of a class of matching priors, we propose in Section 3 continuity corrections based on uniform random variables, which induce greater accuracy for the frequentist distribution of the HPD regions and we obtain asymptotic expansions of the frequentist coverage of HPD regions to the order $o(n^{-1})$ and $O(n^{-3/2})$.

Bayesian calculations are obviously still valid when the observations are discrete. We thus use the results obtained by DiCiccio and Martin (1991) on the Bayesian distribution of modified likelihood ratio statistics to obtain the approximate expression to the order $O(n^{-3/2})$ of HPD regions and to determine the Bayesian coverages of the corrected HPD regions. It appears that generally speaking, corrected HPD regions allow for both Bayesian and frequentist Bartlett corrections. We deduce from these results matching priors to the order $o(n^{-1})$.

This determination of a class of matching priors is important since it gives a way of comparing Bayesian and frequentist approaches. It is also a technique to determine priors in a noninformative setup. Moreover this study sheds light on the structure of HPD regions in the discrete case, in particular by comparing the impact of the continuity corrections on the frequentist and on the Bayesian coverages of the HPD regions; see Section 3.3.

2. ASYMPTOTIC EXPANSIONS IN THE DISCRETE CASE

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables on $\mathbb{R}^k$, with a common distribution P_θ, where $\theta = (\theta_{(1)}, \theta_{(2)}) \in \Theta \subset \mathbb{R}^k$; $\theta_{(1)} \in \mathbb{R}^p$ is the parameter of interest. We denote $X^n = (X_1, \ldots, X_n)$ and P_θ^n the distribution of X^n. Assume that P_θ has a density $f(x; \theta)$ with respect to a discrete measure μ and denote $l_n(\theta)$ the log-likelihood. Let π be a prior density with respect to the Lebesgue measure on Θ. HPD regions are defined by

$$C_\alpha^\pi = \{\pi_1(\theta_{(1)} \mid X^n) \geqslant g_n(\alpha)\},$$

and $P^\pi[C_\alpha^\pi \mid X^n] = \alpha$, where $\pi_1(\theta_{(1)} \mid X^n)$ is the marginal posterior density of $\theta_{(1)}$ and $P^\pi[\,.\mid X^n]$ is the posterior distribution of θ. Kass *et al.* (1989) showed, using Laplace expansions on $\pi_1(\theta_{(1)}) = \int e^{l_n(\theta)}\pi(\theta)\, d\theta_{(2)} / (\int e^{l_n(\theta)}\pi(\theta)\, d\theta)$, that C_α^π is approximated, to the order $O(n^{-2})$, by

$$\{\theta_{(1)}; W_n(\theta_{(1)}) \leqslant k_n(\alpha)\},$$

where $W_n(\theta_{(1)})$ is the adjusted posterior likelihood ratio statistic:

$$W_n(\theta_{(1)}) = 2[l_n(\tilde\theta) + \log \bar\pi(\tilde\theta) - l_n(\tilde\theta(1)) - \log \bar\pi(\tilde\theta(1))].$$

$\tilde\theta$ denotes the posterior mode and $\tilde\theta(1)$ the constrained posterior mode, i.e. when $\theta_{(1)}$ is fixed, $\bar\pi$ is

$$\bar\pi(\theta) = \pi(\theta)\, det[J_{22}(\theta)]^{1/2},$$

$J_{22}(\theta)$ being minus the $(k-p)\times(k-p)$ matrix of second order derivatives of $l_n(\theta)/n$ taken with respect to the nuisance parameter $\theta_{(2)}$.

It is well known that the adjusted posterior likelihood ratio is asymptotically distributed as a chi-square with p degrees of freedom to the first order of approximation, both from a frequentist and from a Bayesian point of view. In this section, we study the asymptotic behaviour of $P_\theta^n[C_\alpha^\pi]$ and more generally the asymptotic distribution of modified likelihood ratio statistics such as the ones considered by DiCiccio and Stern (1994). We first introduce some notations used throughout the paper. Let D denote the differentiation with respect to θ. More precisely, $Dg(\theta)$ is the gradiant vector of g, $D^2 g(\theta)$ is the matrix of second derivatives of g, and so on. Set

$$Z_{n,1} = n^{-1/2} D \log l_n(\theta), \qquad Z_{n,j} = n^{-1/2}(D^j l_n(\theta) - \mu_{n,j}(\theta)), \qquad j = 2, 3,$$

where $\mu_{n,j}(\theta) = E_\theta[D^j l_n(\theta)]$. We denote $\mu_{ijl} = E_\theta[\partial^3 \log f_\theta(X)/\partial\theta_i\,\partial\theta_j\,\partial\theta_l]$ and I the Fisher information matrix per observation. We use matricial notations throughout the paper; for any matrix A, we denote A^{rs}, the (r, s)th component of the inverse of A. $A^{\cdot 1}$ is the vector whose rth component is A^{r1}, $r \leqslant k$ and $A^{1\cdot} = (A^{\cdot 1})^t$ is the transpose of $A^{\cdot 1}$. $I^{(1,1)}$ is the upper left part of size p of the inverse of I and S is the inverse of $I^{(1,1)}$. Similarly, $I^{(\cdot,1)}$ denotes the $k \times p$ matrix whoses components are the I^{jl}, $j \leqslant k$ and $l \leqslant p$ and $I^{(1,\cdot)}$ is its transpose. Set $T = I^{(\cdot,1)} S I^{(1,\cdot)}$ and $N = I^{-1} - T$.

We denote Z_n, the vector whose components are the components of $(Z_{n,1}, Z_{n,2}, Z_{n,3})$, which are linearly independent, as functions of X. In other words, Z_n is such that its covariance matrix is definite-positive. Let d be the size of Z_n. We assume that Z_n is a lattice random vector and we denote L_n its supporting lattice, $L_n = n^{-1/2}[\sum_{j=1}^d \xi_j \mathbb{Z} + x_o]$, so $(\xi_1, \dots, \xi_d)$ is the basis of the lattice support of Z_n, when $n = 1$; let l be the determinant of $(\xi_1, \dots, \xi_d)$; see Bhattacharya and Rao (1986). Σ is the asymptotic covariance matrix of Z_n. Throughout this paper intervals in $\mathbb{R}$, will be denoted $[a, b]$ and hypercubes in $\mathbb{R}^q$ will be denoted $[a, b]^q$, for any $q \in \mathbb{N} - \{0\}$ and any $a, b \in \mathbb{R}$. Hence $[a, b] = [a, b]^1$.

We assume usual regularity conditions, as in Bickel and Ghosh (1990).

THEOREM 1. *Under the above assumptions,*

(1) *if* $p \geqslant 3$,

$$\sup_{w \in \mathbb{R}^+} |P_\theta^n[W_n(\theta_{(1)}) \leqslant w] - Q_p(w)| = O(n^{-p/p+2} \log n), \tag{1}$$

where Q_p *is the* χ_p^2 *cdf.*

(2) *If* $p = 1$, *there exists a continuous expansion to the order* $o(n^{-1/2})$, *if and only if* $\exists r, s \leqslant k$ *such that*

$$\frac{I^{(1,\cdot)}\xi_s}{I^{(1,\cdot)}\xi_r} \in \mathbb{R} - Q. \qquad [H]$$

Note that the result (1) above is uniform in w. It is quite likely that a better order of approximation could be obtained without this constraint. In particular, we conjecture that better results can be obtained when w is not too close to 0.

Condition $[H]$ is quite natural. Indeed, when $p=1$, C_α^π can be approximated by a confidence region in the form:

$$-\sqrt{w/S} \leqslant I^{1\cdot} Z_{n,1} + O_P(n^{-1/2}) \leqslant \sqrt{w/S}.$$

Therefore, condition $[H]$ simply implies that if the main term $(I^{1\cdot} Z_{n,1})$ is discrete but nonlattice (i.e. $[H]$ is satisfied) then there is a frequentist expansion to the order $o(n^{1/2})$, whereas if it is lattice (i.e. $[H]$ is not satisfied) then there is no expansion to orders higher than $O(n^{-1/2})$.

The proof of the first part of Theorem 1 is based on Prawitz's inequality (1972).

Proof of Theorem 1. The adjusted posterior likelihood ratio statistic can be approximated, in terms of the random vector Z_n; see DiCiccio and Stern (1994).

$$W_n(\theta_{(1)}) = Z_{n,1}^t T Z_{n,1} + \frac{Z_{n,1}^t M(Z_n)\, Z_{n,1}}{\sqrt{n}} + 2\frac{Z_{n,1}^t T\beta}{\sqrt{n}} + O(n^{-1}), \tag{2}$$

where $M(Z_n)$ is linear in Z_n and $\beta = D\log\pi(\theta) + D\log det[I_{22}(\theta)]/2$. $W_n(\theta_{(1)})$ can then be considered, to the order $O(n^{-3/2})$, as a function of Z_n and we shall denote indifferently, $W_n(Z_n)$ or $W_n(\theta_{(1)})$. Let p_n be the density associated with the formal Edgeworth expansion of Z_n (corresponding to the continuous case) to the order $O(n^{-3/2})$, i.e.

$$p_n(x) = \sum_{r=0}^{2} n^{-r/2} P_r(-\phi\colon \{\bar{\chi}_{v,n}\})\,(x),$$

where the functions $P_r(-\phi\colon \{\bar{\chi}_{v,n}\})$ are defined in Bhattacharya and Rao (1986, Chap. 2, Lemma 7.2). In particular, $P_0(-\phi\colon \{\bar{\chi}_{v,n}\})$ is the normal density with null expectation and covariance matrix Σ, the asymptotic covariance matrix of Z_n. Continuous Edgeworth expansions imply that when Cramer's condition is satisfied, the frequentist asymptotic distribution of W_n, is, see Bhattacharya and Ghosh (1978),

$$P_\theta^n[W_n \leqslant w] = \int_{\mathbb{R}^d} p_n(x)\, \mathbb{I}_{W_n(x)\leqslant w}\, dx + O(n^{-2}).$$

However, Cramer's condition is not satisfied in the discrete case. To approximate $P_\theta^n[W_n \leqslant w]$, we thus use Prawitz's (1972) inequality: let $K(t) = K_1(t) + iK_2(t)/(\pi t)$, with

$$K_1(t) = 1 - |t|, \qquad K_2(t) = \pi t(1-|t|)\frac{\cos \pi t}{\sin \pi t} + |t|, \qquad \text{when} \quad |t| \leqslant 1,$$

and $K(t) = 0$ when $|t| > 1$; then $\forall r > 0$, for any distribution function F

$$F(x+) \leqslant \frac{1}{2} + V.P. \int_{\mathbb{R}} \exp(-ixt)\,\frac{K(t/r)}{r}\, f(t)\, dt, \tag{3}$$

$$F(x-) \geqslant \frac{1}{2} - V.P. \int_{\mathbb{R}} \exp(-ixt)\,\frac{K(-t/r)}{r}\, f(t)\, dt, \tag{4}$$

where $f(t)$ is the characteristic function associated with F, and V.P. denotes Cauchy's principal value (see Prawitz, 1972), i.e.

$$V.P. \int_{\mathbb{R}} l(t)\, dt = \lim_{h \downarrow 0} \int_{|t| \geqslant h} l(t)\, dt,$$

for a function l.

We denote $F_n(w) = P_\theta^n[W_n \leqslant w]$, $f_n(t)$ its characteristic function. Let $\bar{V} = \sum_{i=1}^d \xi_i U_i$ and $U = \sum_{i=1}^k \xi_i U_i$, where $U_1, \ldots, U_d$ are iid uniform random variables on $[-1/2, 1/2]$. Then $Z_n + \bar{V}/\sqrt{n}$ is a continuous random vector, and $G_n(w)$, the distribution function of $W_n(Z_n, +\bar{V}/\sqrt{n})$, satisfies

$$G_n(w) - \int_{\mathbb{R}^d} p_n(x)\, \mathbb{I}_{W_n(x) \leqslant w}\, dx = O(n^{-1});$$

see Theorem 2.

Let $g_n(t)$ be the characteristic function associated with $G_n(w)$. Since $G_n(w)$ is continuous, we obtain

$$\begin{aligned} F_n(w) - G_n(w) \leqslant{}& \frac{i}{2\pi r} \int_{|t| \leqslant r} e^{-itw} \left(1 - \frac{|t|}{r}\right) \frac{\cos \pi t/r}{\sin \pi t/r} (g_n(t) - f_n(t))\, dt \\ &+ \frac{1}{2r} \int_{|t| \leqslant r} e^{-itw} \left(1 - \frac{|t|}{r}\right) (g_n(t) + f_n(t))\, dt \\ &+ \frac{i}{2\pi r} \int_{|t| \leqslant r} e^{-itx}\, |t|\, (f_n(t) - g_n(t))/t\, dt, \end{aligned}$$

$$F_n(w)-G_n(w)\geqslant\frac{i}{2\pi r}\int_{|t|\leqslant r}e^{-itw}\left(1-\frac{|t|}{r}\right)\frac{\cos\pi t/r}{\sin\pi t/r}(f_n(t)-g_n(t))\,dt$$
$$-\frac{1}{2r}\int_{|t|\leqslant r}e^{-itx}\left(1-\frac{|t|}{r}\right)(f_n(t)+g_n(t))\,dt$$
$$+\frac{i}{2\pi r}\int_{|t|\leqslant r}e^{-itw}\,|t|\,(f_n(t)-g_n(t))/t\,dt,$$

Note that the terms on the right hand side of the above inequalities are real. We prove in Appendix 1 that when $|t|\,\rho_n=O(n/\log n)$, $|t|>n^{1/3}$ and when $\rho_n=o(\sqrt{n})$,

$$|f_n(t)|=O(\rho_n^{-p/2}).$$

Let $p\geqslant 3$, $\forall\,|t|\leqslant n^{1/3}$; we have

$$\begin{aligned}f_n(t)-g_n(t)&=E\{e^{itW_n(Z_n)}[1-\exp(itU^tT^{1/2}R_n/\sqrt{n}+itU^tTU/n+M_n/n^{3/2})]\}\\&=-\frac{t}{n}E\left[e^{itW_n(Z_n)}\left(\frac{iU^tTU}{2}-\frac{t(U^tT^{1/2}R_n)^2}{2}\right)\right](1+O(t^3n^{-1}))\\&=-\frac{t(itr(T)-tE_\theta^n[W_ne^{itW_n}]}{24n}(1+O(t^3n^{-1}))\\&\leqslant\frac{|t|\,|tr(T)+|tM||}{24n}(1+O(t^3n^{-1})).\end{aligned}$$

Decomposing $[-H,H]$ into $\{-n^{1/3},n^{1/3}\}$, $\{t;n^{1/3}<|t|<r/2\}$ and $\{t;r/2<|t|<r\}$, we obtain that

$$\begin{aligned}|F_n(w)-G_n(w)|&\leqslant\Gamma\left(\rho_n^{-p/2}\int_{n^{1/3}}^{r/2}t^{-1}\,dt+\rho_n^{-p/2}+\frac{1}{r}\right)+O(n^{-1})\\&\leqslant\Gamma'\,(\rho_n^{-p/2}\log n+r^{-1})+O(n^{-1}),\end{aligned}$$

where Γ and Γ' are independent of n and w. Conditions on r and ρ_n imply that the tightest choice is $\rho_n=n^{2/(p+2)}$ and $r=n^{p/p+2}\log n$.

When $p=1$, the HPD region can be approximately expressed as a two-sided interval to the order $O(n^{-3/2})$. The term of order $O(1)$ in $W_n(\theta_{(1)})$ is then $\sum_{r=1}^k I^{(1,.)}Z_{n,1}/\sqrt{I^{11}}$, which is the same as the term of order $O(1)$ for the statistic involved in the confidence one-sided or two-sided intervals. In the case of one-sided intervals, Rousseau (2000) proved that continuous expansions to the order $o(n^{-1/2})$ exist if and only if $I^{(1,.)}Z_{n,1}$ is not a lattice random variable, that is, if and only if $[H]$ is satisfied. The author's argument can be applied to this case, which implies the second part of Theorem 1. ∎

Note 2.1. When $p=2$, it is known that at worst,

$$|F_n(w)-G_n(w)|=O(n^{-1/2}).$$

It is, however, likely that when $p=2$ we obtain better results than when $p=1$ and that $|F_n(w)-G_n(w)|=o(n^{-1/2})$ rather than $O(n^{-1/2})$.

Note 2.2. Yarnold (1972) proved that for X_i, $i=1, \ldots, n$ iid random vectors in $\mathbb{R}^p$ having a lattice distribution with mean 0 and covariance matrix V,

$$Pr\left[\sum_{i=1}^{n} X_i/\sqrt{n}\in B_c\right]-Q_p(c)=O(n^{-p/p+1}),$$

where $B_c=\{x; x^tV^{-1}x\leqslant c\}$. His approximation is better than ours, but he studied the case of bounded smooth convex sets, whereas the HPD regions is only approximately so. Indeed, when considered as a set of Z_n's, it is neither bounded nor convex when we consider terms of order higher than $o(1)$. Even though this difference seems minor, the techniques used for that kind of confidence regions are then much more involved.

Note 2.3. This result can be applied to any modified likelihood ratio statistics such as those defined by DiCiccio and Stern (1994).

Note 2.4. Even though we have no proof on the nonexistence of continuous expansion to the order $o(n^{-1})$, Theorem 1 and previous results on the subject (Yarnold, 1972; Bentkus and Götze, 1995) seem to indicate that no such expansion exists. When there exists an asymptotic expansion to the order $o(n^{-1/2})$, it is equal to the formal expansion derived from Q_n (i.e. in the continuous case), therefore DiCiccio and Stern (1994) and Ghosh and Mukerjee (1993)'s results on matching priors for HPD regions in the strongly non lattice case imply that no meaningful matching prior can be obtained for HPD regions in the discrete case. To address this problem of higher order expansions of the frequentist coverages of HPD regions, we propose, in Section 3, continuity corrections.

3. CONTINUITY CORRECTIONS: A GENERAL METHOD

Since there is no result on continuous asymptotic expansions of the frequentist coverage of HPD regions to orders higher than $o(n^{-1/2})$, it is of interest to consider a family of corrected HPD regions such that an asymptotic expansion of its frequentist coverage can be obtained to the order $o(n^{-1})$, at least. This would allow us, in particular, to determine matching priors to this order of approximation, i.e. meaningful matching priors.

The idea is to add a continuity correction based on a uniform random vector to smooth the jumps in the likelihood.

3.1. *Asymptotic Expansion of the Frequentist Coverage of HPD Regions*

In the continuous case, Ghosh and Mukerjee (1993) and DiCiccio and Stern (1994) have proved that penalized likelihood ratio statistics can be Bartlett corrected, in terms of both their frequentist and their Bayesian coverages, i.e. $\exists b_1(\theta)$, b_n^π such that

$$P_\theta^n[W_n(\theta_{(1)}) \leqslant w] = Q_p(w) + \frac{wq_p(w)\, b_1(\theta)}{n} + O(n^{-2})$$

and

$$P^\pi[W_n(\theta_{(1)}) \leqslant w \mid X^n] = Q_p(w) + \frac{wq_p(w)\, b_n^\pi}{n} + O_P(n^{-2}),$$

which is equivalent to the fact that the frequentist distribution of $W_n/(1+B(\theta)/n)$ and the Bayesian distribution of $W_n/(1+b_n^\pi/n)$ are chi-square with p degrees of freedom to the order $O(n^{-2})$. $b_1(\theta)$ and b_n^π are respectively the frequentist and the Bayesian Bartlett corrections of W_n. We consider in this section corrected posterior likelihood ratio statistics that have frequentist Bartlett correction to the order $o(n^{-1})$.

Let q_p be the density of the distribution Q_p of a chi-square random variable with p degrees of freedom. Set A the linear operator such that $A(Z_n) = Z_{n,2}$. Recall that $Z_{n,2}$ is a $k \times k$ symmetrical matrix, set $V = A(\bar{V})$, where $\bar{V} = \sum_{i=1}^d \xi_i U_i$, U_i, $i = 1 \ldots d$ are d independent uniform random variables on $[-1/2, 1/2]$ and let $U = \sum_{i=1}^k \xi_i U_i$. In other words, $\bar{V}$ is a uniform random vector on the cube $\{v = \sum_{i=1}^d v_i \xi_i, |v_i| < 1/2\}$. Let $H_n(v, z)$: $[-1/2, 1/2]^d \times \mathbb{R}^d \to \mathbb{R}^d$ be continuous in z and continuously differentiable in v such that the support of $\bar{V} = \{\sum_{i=1}^d v_i \xi_i, |v_i| \leqslant 1/2\}$ is the same as the support of $\bar{V} + H_n(\bar{V}, z)/\sqrt{n}$ for all $|z| \leqslant \log n^2$ (to the order $O(n^{-1})$). This condition is satisfied in particular, when:

(i) $H_n(v, Z) = 0$ for all v on the boundary of $[-1/2, 1/2]^d$ and all $|z| \leqslant \log n^2$.

(ii) $H_n(v, Z)$ is continuously differentiable in v.

Assume also that

$$\sum_{i=1}^{d} \int_{[-1/2,\,1/2]^d} \frac{\partial H_{n,i}(v,z)}{\partial v_i}\,dv = \int_{[-1/2,\,1/2]^d} \bar{h}_n(v,z)\,dv = 0, \qquad \forall z \in \mathbb{R}^d,$$

where $H_{n,i}$ is the ith component of H_n, in the basis $(\xi_1, \ldots, \xi_d)$ and

$$\bar{h}_n(v,z) = \sum_{j=1}^{k} v_j \xi_j^t [\nabla p_n(z)]/p_n(z) + h_n(v, z_1^t T z_1) + O_p(n^{-1/2}), \tag{5}$$

∇p is the gradient of p and $v.\xi = \sum_{i=1}^{d} v_i \xi_i$. DiCiccio and Stern (1994) prove that $W_n = R_n^t R_n + O(n^{-3/2})$, where

$$R_n = T^{1/2} Z_{n,1} + \frac{T^{1/2} Z_{n,2}(2I^{-1} - T)\,Z_{n,1}}{2\sqrt{n}} + \frac{T^{1/2} M(Z_{n,1})}{\sqrt{n}} + \frac{T^{1/2}\beta}{\sqrt{n}} + O(n^{-1}),$$

with $T^{1/2} = S^{1/2} I^{(1,.)} \in \mathbb{R}^{p\times k}$, and $M(Z_{n,1})$ is a matrix in $\mathbb{R}^{k\times k}$ whose components are linear functions of $Z_{n,1}$. Let

$$\bar{Z}_n = Z_n + \bar{V}/\sqrt{n} + H_n(\bar{V}, Z_n)/n;$$

corrected modified likelihood ratio statistics are then defined by

$$\bar{W}_n(\bar{V}) = W_n(\bar{Z}_n) + O_P(n^{-3/2}) \tag{6}$$

$$= W_n + 2\frac{U^t T^{1/2} R_n}{\sqrt{n}} + 2\frac{H_{n,1}(\bar{V}, Z_n)^t\, T Z_{n,1}}{n} + \frac{U^t T U}{n}$$
$$+ \frac{Z_{n,1}^t T V (2I^{-1} - T)\,Z_{n,1}}{n} + \frac{Z_{n,1}^t T (M_1(U) + M_2(U)/3)\,Z_{n,1}}{n}, \tag{7}$$

where $H_{n,1}(\bar{V}, Z_n)$ represents the first k components of $H_n(\bar{V}, Z_n)$ so that $\bar{Z}_{n,1} = Z_{n,1} + U/\sqrt{n} + H_{n,1}(\bar{V}, Z_n,)/n$. Under the same assumptions as in Section 2, we have the following result:

THEOREM 2. (1) *If $p \geqslant 2$, then the corrected HPD regions*

$$C_\alpha^\pi(\bar{V}) = \{\bar{W}_n(\bar{V}) \leqslant k_n(\alpha)\}$$

satisfy

$$P_\theta^n[C_\alpha^\pi(\bar{V})] = \alpha + \frac{b(\theta)\, w q_p(w)}{n} + \frac{w q_p(w) \sum_{j=1}^{d} \sum_{r,s=1}^{k} \xi_{j,r} \xi_{j,s} T_{rs}}{12pn} + o(n^{-1}), \tag{8}$$

where $b(\theta) = b_1(\theta) - b_n^\pi + O_p(n^{-1/2})$, w is such that $Q_p(w) = \alpha$ and $\xi_{j,r}$ denotes the rth component of ξ_j in the canonical basis.

(2) *If $p=1$, and if $[H]$ is not satisfied then the frequentist coverage of $C_\alpha^\pi(\bar{V})$ has a continuous expansion to the order $o(n^{-1})$ if*

$$h_n(v, z_1^t T z_1) = O_P(n^{-1/2}). \tag{9}$$

The expansion is then

$$P_\theta^n[C_\alpha^\pi(\bar{V})] = \alpha + \frac{b(\theta)\, wq_p(w)}{n} + \frac{wq_p(w) \sum_{j=1}^d \sum_{r,s=1}^k \xi_{j,r}\xi_{j,s}T_{rs}}{12pn} + O(n^{-3/2}). \tag{10}$$

Note that this result is valid for any modified likelihood ratio statistics such as those studied by DiCiccio and Stern (1994).

Note also that the frequentist coverage of these corrected statistics allows for Bartlett corrections, which are equal to those obtained in the continuous term, plus a term due to the discreteness of the observations.

The condition on H_n, is not so restrictive as it seems, and there are many solutions to (5). In particular, functions such as

$$H_n(v, Z_n) = -\sum_{j=1}^k \frac{[v_j^2 - 1/4]\, \xi_j^t\, \Sigma^{-1} Z_n + v_j^2 g_j(z_1^t T z_1)}{2}\, \xi_j \tag{11}$$

satisfy (5), for any continuously differentiable functions g_j's. These are not the only solutions.

Proof of Theorem 2. It is equivalent, to the order $O(n^{-3/2})$, to express $\bar{Z}_n$ as $Z_n + Y/\sqrt{n}$, where $Y = \sum_{i=1}^d \xi_i Y_i$ and Y_i, $i = 1 \ldots d$ are continuous random variables in $[-1/2, 1/2]$, we denote $\tilde{h}_n$ the joint density of $(Y_1, \ldots, Y_d)$ conditional on Z_n, and $y.\xi = \sum_{i=1}^d y_i \xi_i$, then

$$\begin{aligned}
& P_\theta^n[W_n(\bar{Z}_n) \leqslant w] - \int_{\mathbb{R}^d} \mathbb{I}_{W_n(z) \leqslant w}\, p_n(z)\, dz \\
&\quad = \frac{l}{n^{d/2}} \sum_{z \in L_n} \int_{[-1/2,1/2]^d} \mathbb{I}_{W_n(z + y.\xi/\sqrt{n})}\, (q_n(z)\, \tilde{h}_n(y) \\
&\qquad - q_n(z + y.\xi/\sqrt{n}))\, dy\, dz + O(n^{-/2}) \\
&\quad = -\frac{l}{n^{d/2}} \sum_{x \in L_n} \mathbb{I}_{W_n(z) \leqslant w} \sum_{j=1}^d \frac{\xi_j^t p_n''(z)\, \xi_j}{24n} \\
&\qquad + \frac{l}{n^{d/2}} \sum_{x \in L_n} p_n(z) \int_{[-1/2,1/2]^d} \mathbb{I}_{B(z,y)} \left(\tilde{h}_n(y) - 1 - \frac{y.\xi^t \nabla p_n(z)}{\sqrt{n}\, p_n(z)} \right) dy \\
&\qquad + O(n^{3/2})
\end{aligned}$$

$$= -\sum_{j=1}^{d} \int_{\mathbb{R}^d} \mathbb{I}_{W_n(z) \leqslant w} \frac{\xi_j^t p_n''(z)\, \xi_j}{24n}$$
$$+ \frac{l}{n^{d/2}} \sum_{x \in L_n} p_n(z) \int_{[-1/2,\, 1/2]^d} \mathbb{I}_{B(z,\, y)} \left(\tilde{h}_n(y) - 1 - \frac{y.\xi^t \nabla p_n(z)}{\sqrt{n}\, p_n(z)} \right) dy$$
$$+ O(n^{3/2}), \tag{12}$$

where

$$B(z, y) = \{W_n(z + y.\xi/\sqrt{n}) \leqslant w < \sup_{y \in [-1/2,\, 1/2]^d} W_n(z + y.\xi/\sqrt{n})\}.$$

Here $p''n(z)$ denotes the matrix of second derivatives of $p_n(z)$ (wrt z). The term defined by (12) has a continuous expansion to the order $o(n^{-1/2})$, if

$$\int_{[-1/2,\, 1/2]^d} \mathbb{I}_{B(z,\, y)} (\tilde{h}_n(y) - 1)\, dy = O(n^{-1/2}),$$

that is if $\tilde{h}_n(y) = \mathbb{I}_{[-1/2,\, 1/2]^d}(1 + \bar{h}_n(y, z)/\sqrt{n})$, where $\int_{[-1/2,\, 1/2]^d} \bar{h}_n(y, z)\, dy = 0$. If $p \geqslant 3$ or if $[H]$ is satisfied, we prove in Appendix B that

$$\frac{l}{n^{d/2}} \sum_{x \in L_n} p_n(z) \int_{[-1/2,\, 1/2]^d} \mathbb{I}_{B(z,\, y)} \left(\bar{h}_n(y, z) - y.\xi^t \frac{\nabla p_n(z)}{p_n(z)} \right) dy$$
$$= \int_{\mathbb{R}^d} p_n(z) \int_{[-1/2,\, 1/2]^d} \mathbb{I}_{B(z,\, y)} \left(\bar{h}_n(y, z) - y.\xi^t \frac{\nabla p_n(z)}{p_n(z)} \right) dy + o(n^{-1/2}) \tag{13}$$

where

$$\bar{h}_n(y, z) = \sum_{j=1}^{k} y_j \xi_j^t \frac{\nabla p_n(z)}{p_n(z)} + h_n(y, z_1^t T z_1).$$

If $p = 2$, since no results are obtained in this case, see Note 2.1, we use the corrections defined by (9), which leads to the same expansion of the frequentist coverage of $C_\alpha^\pi(\bar{V})$, as in (10). ▮

We now consider the impact of the corrections on the posterior coverages of the HPD regions.

3.2. *Bayesian Coverage of $C_\alpha^\pi(\bar{V})$ and Matching Priors*

The Bayesian coverage of the corrected HPD region depends on the function $H_n(\bar{V}, Z_n)$. We recall that this function depends on θ through $Z_n = Z_n(\theta)$. We prove that for the most natural choices of H_n, i.e. those defined by (11), there exist Bartlett corrections for the Bayesian coverage. Let $\hat{T} = T(\hat{\theta})$, $\hat{S} = S(\hat{\theta})$, and $\hat{S}^{1/2}$ be the definite positive square root of $\hat{S}$.

THEOREM 3. *If H_n is defined by* (11), *the posterior coverage satisfies*

$$P^{\pi}[\bar{W}_n(\bar{V}) \leqslant w \mid X^n] = Q_p(w) - w q_p(w) \frac{-pb_n^{\pi} + B + \sum_{j=1}^{d} \sum_{r,s=1}^{k} \hat{\xi}_{j,r} \hat{\xi}_{j,s} \hat{T}_{rs}/12}{pn}. \quad (14)$$

B is the term corresponding to $H_n(\bar{V}, Z_n)$, which is given in Appendix C when H_n satisfies (11).

The corrected modified likelihood ratio statistics allow for Bayesian Bartlett corrections, which are equal to

$$n^{-1}\left(b_n^{\pi} - B/p - \sum_{j=1}^{d} \sum_{r,s=1}^{k} \hat{\xi}_{j,r} \hat{\xi}_{j,s} \hat{T}_{rs}/12p\right).$$

Proof. To prove this result, we first determine the conditional probability $P^{\pi}[\bar{W}_n(\bar{V})\,(1 - b_n^{\pi}/n) \mid X^n, \bar{V}]$. Let $y = \sqrt{n}\,(\theta - \hat{\theta})$ and $\hat{Z}_{n,i} = Z_{n,i}(\hat{\theta})$, $i \leqslant 3$. Then, if J_n denotes the empirical Fisher information matrix (calculated at $\hat{\theta}$),

$$Z_{n,1}(\theta)_j = -(J_n y)_j + \sum_{i,l=1}^{k} \frac{D_{jil} l_n(\hat{\theta})\, y_i\, y_l}{n^{3/2}} + O_P(n^{-1}),$$

and

$$(Z_{n,2})_{i,j} = (\hat{Z}_{n,2})_{i,j} + \sum_{l=1}^{k} y_l (D_{i,jl} l_n(\hat{\theta})/n + D_l I(\hat{\theta})_{i,j}) + O_P(n^{-1/2}).$$

Set

$$R_n^1(U) = R_n + \frac{\hat{T}^{1/2} U}{\sqrt{n}} + \frac{1}{n} \sum_{i=1}^{k} \left(\frac{\partial T^{1/2} U}{\partial \theta_i}\right)_{\theta = \hat{\theta}}.$$

Then,

$$\bar{W}_n(\bar{V}) = R_n^1(U)^t\, R_n^1(U) + \frac{P(y, \bar{V})}{n} - 2\, \frac{H_n(\bar{V}, y)\, \hat{T} \hat{I} y}{n},$$

where $P(y, \bar{V})$ is polynomial in y with degree 2 and is linear in $\bar{V}$. Simple calculations prove that $R_n^1(U)^t\, R_n^1(U)$ is distributed from a noncentral chi-square distribution with p degrees of freedom and non-centrality parameter

$$\frac{\delta}{n} = \frac{1}{n} \|\tilde{T}^{1/2} U + B_R\|^2$$

to the order $O(n^{-3/2})$, where B_R is the posterior expectation of R_n; in other words, $|B_R|^2/n$ is the noncentrality parameter of R_n. Algebra on the conditional characteristic function $\varphi_{(1)}(t)$ leads to

$$\varphi_{(1)}(t) = E\left[e^{itR_n^1(U)^t T_n^1(U)} \left(1 + it\,\frac{P(y, \bar{V})}{n}\right) \Big| X^n, \bar{V}\right]$$

$$= \frac{e^{it\delta/n(1-2it)^{-1}}}{(1-2it)^{p/2}}\left[1 - \frac{a_1(\bar{V} + a_1'(\bar{V})}{n}\right] + \frac{2a_1(\bar{V})}{2n(1-2it)}$$

$$- \frac{(1-2it)\,a_1'(\bar{V})}{n} + O(n^{-3/2}),$$

where $a_1(\bar{V})$ and $a_1'(\bar{V})$ are of order $O(1)$ and linear in $\bar{V}$, therefore those terms disappear, when integrating over $\bar{V}$. There only remains the noncentrality term. Assume that H_n is defined by (11), in other words, that it is a linear function of Z_n plus a function of W_n. Then we prove, in Appendix C that there exists a function of $\bar{V}$, $B = O_P(1)$, such that

$$\varphi_{(2)}(t) = \frac{2it}{n}\, E[e^{itR_n^1(U)\,tR_n^1(U)} H_{n,1}(\bar{V}, y)\,\hat{T}^{1/2} R_n \mid X^n, \bar{V}]$$

$$= \frac{itB(\bar{V}))}{n(1-2it)^{(p+1)/2}}.$$

Finally, integrating over v, we obtain that the characteristic function of $W_n(\bar{Z}_n)$ can be approximated by

$$\varphi_n(t) = \frac{1}{(1-2it)^{p/2}} + \frac{it(b_n^\pi + \sum_{j=1}^k \hat{\xi}_j^t \hat{T} \hat{\xi}_j/12 + B)}{2n(1-2it)^{p/2+1}},$$

where B is given in Appendix C. Inverting the characteristic function, we finally obtain (14). ∎

Note 3.2. When H_n is defined by (11), matching priors are defined by

$$\frac{\sum_{j=1}^k \xi_j^t T \xi_j}{6p} + \frac{B}{p} = b(\theta) + O_{P_\theta}(n^{-1/2})$$

(see Ghosh and Mukerjee (1993) and DiCiccio and Stern (1994) for the expression of $b(\theta)$).

3.3. *The Multinomial Case*

Let $X = (X_1, \ldots, X_k)$ be a multinomial random vector, with cell probabilities $\theta = (\theta_1, \ldots, \theta_k)$ ($\sum_{i=1}^k \theta_i = 1$); the parameters of interest are $k-1$ independent smooth functions of θ:

$$\eta_1 = g_1(\theta), \ldots, \eta_{k-1} = g_{k-1}(\theta).$$

We have

$$Z_{n,1}(\eta) = n^{-1/2} \sum_{i=1}^{k-1} \bar{e}_i \sum_{j=1}^{k-1} \frac{\partial \theta_j}{\partial \eta_i} \left[\frac{X_j - n\theta_l}{\theta_k} + \sum_{l=1}^{k-1} \frac{X_l - n\theta_l}{\theta_k} \right],$$

where $\bar{e}_1, \ldots, \bar{e}_{k-1}$ is the canonical basis in $\mathbb{R}^{k-1}$. Thus, $Z_{n,1}$ is a lattice random vector, whose support is defined by

$$L_n = \sum_{i=1}^{k-1} \xi_i \mathbb{Z}/\sqrt{n} + x_{0,n}$$

and

$$\xi_i = \sum_{j=1}^{k-1} \left(\frac{1}{\theta_i} \frac{\partial \theta_i}{\partial \eta_j} + \frac{1}{\theta_k} \sum_{l=1}^{k-1} \frac{\partial \theta_l}{\partial \eta_j} \right) \bar{e}_j,$$

for $i = 1, \ldots, k-1$ and

$$x_{0,n} = -n \sum_{i=1}^{k-1} \left(\frac{\partial \theta_j}{\partial \eta_i} + \frac{\theta_j}{\theta_k} \sum_{l=1}^{k-1} \frac{\partial \theta_l}{\eta_i} \right) \bar{e}_i.$$

It is very easy to check that $Z_{n,2}$ is linearly dependent on $Z_{n,1}$, as a function of X, for any $1-1$ transformation η. Thus in the multinomial case, $Z_n = Z_{n,1}$ and $d = k-1$. The Fisher information matrix has the form

$$I(\eta) = \frac{d\theta^t}{d\eta} I(\theta) \frac{d\theta}{d\eta}.$$

We obtained, see Theorem 1, in the purely discrete case a continuous expansion to the order $O(\log n n^{-(k-1)/(k+1)})$. To obtain expansions to the order $o(n^{-1})$, we thus consider a continuity correction in the form (11), and more precisely,

$$\begin{aligned} H_n(v, Z_{n,1}(\eta)) &= -\sum_{j=1}^{k-1} \frac{[v_j^2 - 1/4]\, \xi_j^t I(\eta)^{-1} Z_{n,1}(\eta)}{2} \xi_j \\ &= -\sum_{j=1}^{k-1} \frac{[v_j^2 - 1/4]}{2\sqrt{n}} \left(\frac{X_j}{\theta_j} - \frac{X_k}{\theta_k} \right) \xi_j(\eta). \end{aligned}$$

The corrected posterior likelihood ratio statistic is equal to

$$\begin{aligned} \bar{W}_n(\bar{Z}_n) = W_n(\theta) + 2n^{-1} \sum_{i=1}^{k-1} v_i \left(\frac{X_i}{\theta_i} - \frac{X_k}{\theta_k} \right) \\ + \sum_{i=1}^{k-1} \frac{[v_i^2 - 1/4]}{2n} \left[1 - n^{-1} \left(\frac{X_i}{\theta_i} - \frac{X_k}{\theta_k} \right)^2 \right], \end{aligned}$$

where

$$W_n(\theta) = 2\left(\sum_{i=1}^{k} X_i \log[\tilde{\theta}_i/\theta_i] + \log[\pi(\tilde{\theta})/\pi(\theta)]\right).$$

Matching priors are then solutions of the following partial differential equation:

$$b(\eta) = \frac{1}{3p} \sum_{j=1}^{k-1} \left(\frac{1}{\theta_j} + \frac{1}{\theta_k}\right).$$

For the canonical parametrisation (i.e. θ), the above equation becomes:

$$\frac{1}{3} \sum_{j=1}^{k-1} \theta_j^{-1} + \frac{(k-1)}{3\theta_k} = 2(k-3)\,\theta_k + k^2 - 2k - 6$$
$$- \sum_{r=1}^{k-1} \frac{\pi_{rr}\theta_r(1-\theta_r)}{\pi} + \sum_{r \neq s} \frac{\pi_{rs}\theta_r\theta_s}{\pi} + \sum_{u=1}^{k-1} \frac{\pi_u}{\pi}(-1 + k\theta_u).$$

Note in particular that the corrected likelihood ratio statistics has a frequentist Bartlett correction which is different from $b_1(\theta)$ which is the one obtained formally from the continuous case. If we understand the continuity corrections as a representation of the effect of discretisation, it might be an explanation for the nonefficiency of $b_1(\theta)$ observed by Frydenberg and Jensen (1989).

APPENDIX A

Proof of Theorem 1*: Upper Bound of* $|f_n(t)|$ *when* $|t|$ *Is Large*

$Z_n = \sum_{i=1}^d \xi_i z_i$, where $z_i \in \mathbb{Z}/\sqrt{n}$, i.e. $z = (z_1, \ldots, z_d)$ represents the coordinates of Z_n in the basis $(\xi_1, \ldots, \xi_d)$. Recall that $f_n(z)$ is in the form

$$\begin{aligned} f_n(t) &= E[e^{itW_n}] \\ &= E[e^{itZ_{n,1}^t TZ_{n,1} + itZ_{n,1}^t TM(Z_n)\, Z_{n,1}/\sqrt{n}}] \end{aligned}$$

where M is a matrix in the form

$$M = M_1(Z_{n,1}) + M_2 Z_{n,2} + \frac{M_n(Z_n)}{\sqrt{n}},$$

M_1 is a linear function of $Z_{n,1}$ independent of n, M_2 is a matrix independent of Z_n, and n and M_n are polynomial functions of Z_n whose coefficients are polynomial functions in $n^{-1/2}$. We first study

$$f_{n,1}(t) = E[e^{itZ_{n,1}^t TZ_{n,1}}].$$

Denote $X_i(\theta) = D\log f_\theta(X_i)$ and $Y_1 = \sum_{i=1}^{\rho_n^2} X_i(\theta)$, $Y_2 = \sum_{i=\rho_n^2+1}^{(n+\rho_n^2)/2} X_i(\theta)$ and $Y_3 = \sum_{i=(n+\rho_n^2)/2+1}^{n} X_i(\theta)$. Then, we the following symmetrization inequality as in Bentkus and Götze (1995),

$$\begin{aligned}
|f_{n,1}(t)|^2 &\leqslant E(|E[e^{itY_1^tTY_1/n+2itY_1^tT(Y_2+Y_3)/n} \mid Y_2, Y_3]|^2)\\
&= E[e^{itY_1^tTY_1/n-it\bar{Y}_1^tT\bar{Y}_1/n+2it\tilde{Y}_1^tT(Y_2+Y_3)/n}]\\
&\leqslant E\,|E[e^{2it\tilde{Y}_1^tT[Y_2+Y_3]/n} \mid \tilde{Y}_1]|\\
&= E[e^{2it/n\tilde{Y}_3^tT\tilde{Y}_1}],
\end{aligned}$$

where $\tilde{Y}_i$ is the symmetrization of Y_i, i.e. $\tilde{Y}_i = Y_i - \bar{Y}_i$, with Y_i and $\bar{Y}_i$, independent and identically distributed. $\tilde{Y}_3$ and $\tilde{Y}_1$ are now independent and

$$E[e^{2it/n\tilde{Y}_3^tT\tilde{Y}_1}] = E[E[\exp(2it/n(T\tilde{Y}_1)^t\,\tilde{X}(\theta)) \mid \tilde{Y}_1]^{n/2-\rho_n^2/2}],$$

with $\tilde{X} = X(\theta) - \bar{X}(\theta)$, $X(\theta)$ and $\bar{X}(\theta)$ iid with the same distribution as the $X_j(\theta)$'s. Therefore the characteristic function of $\tilde{X}(\theta)$, $\tilde{f}_1$, satisfies for all $v \in \mathbb{R}^k$,

$$\begin{aligned}
\tilde{f}_1(v) &= \sum_{\tilde{x}} \tilde{p}(\tilde{x})\, e^{iv^t\tilde{x}} = \sum_{z\in L_1} p(z)\, E_X[e^{iv^t(X-z)}]\\
&= f_1(v)\, f_1(-v) = |f_1(v)|^2.
\end{aligned}$$

Using a Taylor expansion around zero of $\log|f_1(z)|^2 = \log f_1(z) + \log f_1(-z)$, where the logarithm of a complex number $re^{i\theta}$, $\theta\in(-\pi,\pi]$ is defined by

$$\log r + i\theta,$$

we obtain, as in Bhattacharya and Rao (1986, Chap. 2)

$$n\log|f_1(2t(T\tilde{Y}_1)/n)|^2 = -\tilde{Y}_1^tT\tilde{Y}_1t^2 + \sum_{|v|=2r;\, r=2}^{\lfloor s+1/2\rfloor} \frac{(-1)^r\,\chi_v v_n^v}{(2r)!\,n^{2r-1}} + R_n,$$

where $v_n = 2t(T\tilde{Y}_1) \in \mathbb{R}^k$, the χ_v's denote the cumulants of X, for $v = (v_1, \ldots, v_k)$, $|v| \geqslant s+2$ and $v_n^v = v_{n,1}^{v_1} \cdots v_{n,k}^{v_k}$. Since there exists $\alpha > 0$ such that

$$v_n^t v_n \geqslant \frac{1}{8}\left|\sum_{|v|=2r;\, r=2}^{\lfloor s+1/2\rfloor} \frac{(-1)^r\,\chi_v v_n^v}{(2r)!\,n^{2r-1}} + R_n\right|, \qquad \text{when} \quad |v_n| < \alpha n,$$

we have

$$|f_1(2t(T\tilde{Y}_1)/n)|^{n-\rho_n^2} \leqslant \exp\left[-\frac{\tilde{Y}_1^t T \tilde{Y}_1 z^2}{2}\right], \qquad \text{when} \quad |v_n| < \alpha n.$$

This implies in particular that if $|t| \leqslant \delta n/\log n$, for some $\delta > 0$,

$$\begin{aligned}
|E[e^{2it/n\tilde{Y}_3^t T\tilde{Y}_1}]| &\leqslant E\left\{\exp\left[-\frac{\tilde{Y}_1^t T \tilde{Y} t^2}{2}\right]\right\} + P_\theta^n[|v_n| \geqslant \alpha n] \\
&= \frac{l^2}{\rho_n^{2k}} \sum_{\underline{m} \in \mathbb{Z}^k} \varphi_{\tilde{I}}(\underline{m}/\rho_n) \sum_{\underline{m}' \in \mathbb{Z}^k} \varphi_{\tilde{I}}(\underline{m}'/\rho_n)\, e^{-t^2(\underline{m}-\underline{m}')^t \tilde{T}(\underline{m}-\underline{m}')/2} \\
&\quad \times (1 + O(\rho_n^{-1})) + O(n^{-2}) \\
&= \frac{l}{\rho_n^p}(1 + O(\rho_n^{-1})) + O(n^{-2}) + O(n^{-2}).
\end{aligned}$$

Therefore, when $|t| \leqslant \delta n/\log n$

$$f_{n,1}(t) = O(\rho_n^{-p/2}).$$

Using the same kind of calculation as in the symmetrization argument, see also Bentkus and Götze (1995), we obtain that

$$|E[e^{itW_n}]|^2 \leqslant [E[e^{2it/n\tilde{Y}_3^t T\tilde{Y}_1}] + E[e^{2it/n\tilde{Y}_3^t T\tilde{Y}_2}]]\,(1 + o(1)),$$

and

$$f_n(t) = O(\rho_n^{-p/2}),$$

when $|t|$ is large enough.

APPENDIX B

Proof of Expansion (13)

Let

$$\begin{aligned}
\Delta_n = \frac{l}{n^{d/2}} \sum_{z \in L_n} p_n(z) \int_{[-1/2,\,1/2]^d} \mathbb{I}_{B(z,\,v)} h_n(v, z_1^t T z_1)\, dz \\
- \int_{\mathbb{R}^d} \int_{[-1/2,\,1/2]^d} p_n(z)\, \mathbb{I}_{B(z,\,v)} h_n(v, z_1^t T z_1)\, dv + o(n^{-1/2}).
\end{aligned}$$

Moreover, when $z \in B(z, y)$, $W_n(z) = w + O_P(n^{-1/2}) = z_1^t T z_1 + O_p(n^{-1/2})$, therefore

$$h_n(v, z_1^t T z_1) = h_n(v, w) + (z_1^t T z_1 - w)\left(\frac{\partial h_n(v, x)}{\partial x}\right)_{x=w} + O_P(n^{-1}).$$

We thus have:

$$\Delta_n = \int_{[-1/2, 1/2]^d} h_n(v, w)\left(P_\theta^n[W_n(Z_n + v\xi/\sqrt{n}) \leqslant w] - \int_{\mathbb{R}^d} \mathbb{I}_{B(z, v)}\, p_n(z)\, dz\right) dv$$
$$+ \int_{[-1/2, 1/2]^d} \left(\frac{\partial h_n(v, x)}{\partial x}\right)_{x=w} E_\theta^n\,[(z_1^t T z_1 - w)\,\mathbb{I}_{B(Z_n, v)}]\, dv + o(n^{-1}). \quad (15)$$

On $B(Z_n, v)$, when $|Z_n| \leqslant \log n^2$, $|z_1^t T z_1 - w| \leqslant M \log n^2/\sqrt{n}$; therefore

$$E_\theta^n\,[(z_1^t T z_1 - w)\,\mathbb{I}_{B(Z_n, v)}] \leqslant M\,\frac{\log n^2}{\sqrt{n}}\, P_\theta^n[B(Z_n, v)]$$
$$\leqslant M'\,\frac{\log n^2}{n},$$

where M and M' are positive constants. The first term of the right hand side of (15) is dealt with as in Theorem 1 and is therefore a $o(n^{-1/2})$. Hence, $\Delta_n = o(n^{-1/2})$ and Eq. (13) is proved.

APPENDIX C

Calculation of $\varphi_2(z)$

We have

$$\varphi_2(z) = \frac{2iz}{n}\, E[e^{izR_n^1(U)\, tR_n^1(U)} H_{n,1}(\bar{V}, y)\, \hat{T}^{1/2} R_n \mid X^n, \bar{V}]$$
$$= \frac{iz}{n} \sum_{j=1}^{k} [v_j^2 - 1/4]\, \hat{\xi}_j^t \hat{\Sigma}^{-1} E\,[Z_n \hat{\xi}_j^t \hat{T} J_n\, y e^{ity^t J_n \hat{T} J_n y} \mid X^n, \bar{V}]$$
$$- \frac{iz}{n} \sum_{j=1}^{k} E\,[g_j(\bar{V}, y^t J_n \hat{T} J_n\, y)\, e^{ity^t J_n \hat{T} J_n y} \hat{\xi}_j^t \hat{T} J_n\, y \mid X^n, \bar{V}] + O(n^{-31/2}).$$

The latter term of the right hand side is null due to symmetry arguments. We thus obtain

$$\varphi_2(z) = \frac{iz}{n} \sum_{r=1}^{d} \sum_{l,j=1}^{k} \sum_{a=1}^{p} [v_j^2 - 1/4]\, \hat{\xi}_j^t \hat{\Sigma}^{\cdot r} D_l z_r(\hat{\theta})$$

$$\times \hat{\xi}_j^t \hat{I}^{\cdot 1} \hat{S}_{.a} E[y_l\, y_a e^{ity_{(1)}^t S y_{(1)}} \mid X^n, \bar{Z}]$$

$$= \frac{iz}{n(1-2iz)^{p/2+1}} \sum_{r=1}^{d} \sum_{l,j=1}^{k} [v_j^2 - 1/4]\, \hat{\xi}_j^t \hat{\Sigma}^{\cdot r} D_l z_r(\hat{\theta})\, \hat{\xi}_j^t \hat{T}^{\cdot l},$$

because $Z_n(\theta) = Z_n(\hat{\theta}) + \sum_{l=1}^{k} D_l Z_n(\hat{\theta})\,(\theta - \hat{\theta})$. Integrating over $\bar{V}$ leads to

$$E[e^{izR_n^1(U)\, tR_n^1(U)} H_{n,1}(\bar{V}, y)\, \hat{T}^{1/2} R_n \mid X^n]$$

$$= -\frac{iz}{6n(1-2iz)^{p/2+1}} \sum_{r=1}^{d} \sum_{l,j=1}^{k} \hat{\xi}_j^t \hat{\Sigma}^{\cdot r} D_l z_r(\hat{\theta})\, \hat{\xi}_j^t \hat{T}^{\cdot l}$$

and B is equal to

$$B = -\frac{1}{6} \sum_{r=1}^{d} \sum_{l,j=1}^{k} \hat{\xi}_j^t \hat{\Sigma}^{\cdot r} D_l z_r(\hat{\theta})\, \hat{\xi}_j^t \hat{T}^{\cdot l}.$$

ACKNOWLEDGMENTS

This paper is based on part of the author's Ph.D. dissertation of the University of Paris 6. I thank Professor Christian Robert, my advisor, for helpful comments and discussions and the subject material and I thank the referee's for their comments on the paper.

REFERENCES

R. N. Bhattacharya and J. K. Ghosh, On the validity of the formal Edgeworth expansion, *Ann. Statist.* **6** (1978), 434–451.

R. N. Bhattacharya and R. Rao, "Normal Approximations and Asymptotic Expansions," 2nd ed., Wiley, New York, 1986.

V. Bentkus and F. Götze, On the number of lattice points in a large ellipsoids, *Dokl. Akad. Nauk* **343** (1995), 439–440.

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

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

T. J. DiCiccio and A. M. Martin, Approximations of marginal tails probabilities for a class of smooth functions with applications to Bayesian and conditional inference, *Biometrika* **78** (1991), 731–740.

T. J. DiCiccio and S. E. Stern, On Bartlett adjustments for approximate Bayesian inference, *Biometrika* **80** (1993), 731–740.

T. J. DiCiccio and S. E. Stern, Frequentist and Bayesian Bartlett correction of test statistics based on adjusted profile likelihoods, *J. Roy. Statist. Soc. Ser. B* **56** (1994), 397–448.

C. G. Esseen, Fourier analysis of distribution functions, *Acta. Math.* **77** (1945), 1–125.

M. Frydenberg and J. L. Jensen, Is the "improved likelihood ratio statistic" really improved in the discrete case? *Biometrika* **76** (1989), 655–661.

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

R. E. Kass, L. Tierney, and J. B. Kadane, Approximate methods for assessing influence and sensitivity in Bayesian analysis, *Biometrika* **76** (1989), 863–874.

H. Prawitz, Limits for a distribution, if the characteristic function is given in a finite domain, *Skand. Aktuar Tidskr.* (1972), 138–154.

C. P. Robert, "The Bayesian Choice," Springer-Verlag, New York, 1994.

J. Rousseau, Coverage properties of one-sided intervals in the discrete case and applications to matching priors, *Ann. Inst. Statist. Math.* **52** (2000), 28–42.

M. Siotani and Y. Fujikoshi, Asymptotic approximations for the distribution of multinomial goodness-of-fit statistics, *Hiroshima Math. J.* **14** (1984), 115–124.

J. K. Yarnold, Asymptotic approximations for the probability that a sum of a lattice random vectors lies in a convex set, *Ann. Math. Statist.* **43** (1972), 1566–1580.