行政院國家科學委員會專題研究計畫　期中進度報告

後驗分配的近似計算(1/2)

計畫類別：個別型計畫
計畫編號：NSC92-2118-M-004-003-
執行期間：92 年 08 月 01 日至 93 年 07 月 31 日
執行單位：國立政治大學統計學系

計畫主持人：翁久幸

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 93 年 4 月 21 日

1 Introduction

In Weng [2], I derived a modified version of Stein's Identity and applied it to prove posterior normality of stochastic processes. This Stein's Identity first appeared in Woodroofe [3, 4] in the context of deriving a very weak type expansion. Weng [2] makes two contributions: (1) it relaxes the conditions for posterior normality when data come from a stochastic process, and (2) it is novel in including end point terms in an analysis based on Stein's Identity.

Let Γ denote a finite signed measure of the form $d\Gamma = fd\Phi$, where Φ denotes the standard normal distribution function and f is a real-valued function defined on $\mathcal{R}$. The difference between the modified and the original versions is the requirement on f. The original version (in Woodroofe [3, 4]) requires f to be continuously differentiable on $\mathcal{R}$, while the new one allows f to have jump discontinuities at both endpoints. Application of this result includes some nonhomogeneous Poisson processes and conditional exponential families.

2 Review of Stein's Identity

Now we review Stein's Identity. Write

$$\Phi h = \int h d\Phi$$

for functions h for which the integral is finite. Next let Γ denote a finite signed measure of the form $d\Gamma = fd\Phi$, where f is a real-valued function defined on $\mathcal{R}$ satisfying $\Phi|f| = \int |f| d\Phi < \infty$. For $p \geq 0$, denote by H_p the collection of all measurable functions $h: \mathcal{R} \to \mathcal{R}$ for which $|h(z)| \leq 1 + |z|^p$, and define $H = \cup_{p\geq 0} H_p$. Let

$$Uh(z) = e^{\frac{1}{2}z^2} \int_z^\infty [h(y) - \Phi h] e^{-\frac{1}{2}y^2} dy, \tag{1}$$

for $-\infty < z < \infty$.

Lemma 2.1 *There are (finite) positive constants $c_0, c_1, c_2, \ldots$ for which $UH_0 \subseteq c_0 H_0$ and $UH_p \subseteq c_p H_{p-1}$ for all $p = 1, 2, \ldots$.*

Proof. See Woodroofe [4, Lemma 1]. ■

Lemma 2.2 *Let r be a nonnegative integer. Suppose that $d\Gamma = f d\Phi$ as above. Suppose that*

$$\int_{\infty}^{\infty} |f| d\Phi + \int_{\infty}^{\infty} (1 + |z|^r)|f'(z)|\Phi(dz) < \infty.$$

Then

$$\Gamma h = \Gamma 1 \cdot \Phi h + \int_{\infty}^{\infty} Uh(z) f'(z)\Phi(dz), \tag{2}$$

for all $h \in H_r$.

Proof. See Woodroofe [3, Proposition 1]. ■

3 Current Progress

3.1 The Model

We are extending this method to multiparameter cases. Let X_t be a random vector distributed according to a family of probability densities $p_\theta(x_t)$, where t is a discrete or continuous parameter and $\theta \in \Theta$, an open subset of $\mathcal{R}^k$. Let P_θ and E_θ be the associated probability measure and expectation of p_θ. Assume that the log-likelihood function $\ell_t(\theta) = \log p_\theta(x_t)$ is twice continuously differentiable with respect to θ. Denote $\nabla \ell_t(\theta)$ as the vector of first-order partial derivatives, and $\nabla^2 \ell_t(\theta)$ as the matrix of second-order partial derivatives. Throughout let $\hat{\theta}_t$ be a root of the likelihood equation satisfying $\nabla \ell_t(\hat{\theta}_t) = 0$, where differentiation is with respect to θ. Whenever such a root exists and $-\nabla^2 \ell_t(\hat{\theta}_t)$ is positively definite, we define B_t and Z_t as

$$B_t' B_t = -\nabla^2 \ell_t(\hat{\theta}_t), \tag{3}$$

$$Z_t = B_t(\theta - \hat{\theta}_t); \tag{4}$$

otherwise, define B_t and Z_t arbitrarily (in a measurable way). Consider a Bayesian model in which θ has a prior density ξ. Then the posterior density of θ given data x_t is $\xi^t(\theta) \propto e^{\ell_t(\theta)}\xi(\theta)$, and the posterior density of Z_t is

$$\zeta^t(z) \propto \xi^t(\theta(z)) \propto e^{\ell_t(\theta) - \ell_t(\hat{\theta}_t)}\xi(\theta), \tag{5}$$

where the relation of θ and z is given in (4). Now a Taylor's expansion gives

$$\ell_t(\theta) = \ell_t(\hat{\theta}_t) + \frac{1}{2}(\theta - \hat{\theta}_t)' \nabla^2 \ell_t(\theta_t^*)(\theta - \hat{\theta}_t), \tag{6}$$

where θ_t^* lies between θ and $\hat{\theta}_t$. Letting $R_t(\theta)$ be such that

$$z_t' R_t(\theta) z_t = (\theta - \hat{\theta}_t)' [\nabla^2 \ell_t(\hat{\theta}_t) - \nabla^2 \ell_t(\theta_t^*)](\theta - \hat{\theta}_t), \tag{7}$$

it follows that

$$\ell_t(\theta) = \ell_t(\hat{\theta}_t) - ||z_t||^2/2 - z_t' R_t(\theta) z_t/2. \tag{8}$$

So (5) can be rewritten as

$$\zeta^t(z) \propto \phi_k(z) f_t(z), \tag{9}$$

where $f_t(z) = \xi(\theta(z))\exp[-z_t' R_t(\theta) z_t/2]$ and $\phi_k(z) = (1/\sqrt{2\pi}^k)\exp[-||z||^2/2]$. Some calculations are useful for later references.

$$\frac{\nabla f_t(Z_t)}{f_t(Z_t)} = B_t'^{-1}[\frac{\nabla \xi(\theta)}{\xi(\theta)} + \nabla \ell_t(\theta) - \nabla^2 \ell_t(\hat{\theta}_t)(\theta - \hat{\theta}_t)]. \tag{10}$$

3.2 Multiparameter Cases

The multivariate version of the modified Stein's Identity has been derived recently, but the details are omitted here.

We plan to use a two-parameter nonhomogeneous Poisson process as an application of our results. Let N_t, the number of events observed by time t, follow a nonhomogeneous Poisson process with time-dependent intensity function $\Lambda(t) = \lambda e^{\mu + \lambda t}$ over the time interval, where $\lambda > 0$ and μ are two unknown parameters. So, for each fixed t, N_t is a Poisson with mean $\int_0^t \lambda(s) ds$. Letting $\theta = (\lambda, \mu)'$, then the log-likelihood function is $\ell_t(\theta) = N_t(\log\lambda + \mu) + \lambda \sum_{i=1}^{N_t} x_i - e^{\mu}(e^{\lambda t} - 1)$, with derivatives

$$\nabla \ell_t(\theta) = (N_t/\lambda + \sum_{i=1}^{N_t} x_i - t e^{\mu + \lambda t}, N_t - e^{\mu}(e^{\lambda t} - 1))', \tag{11}$$

$$-\nabla^2 \ell_t(\theta) = \begin{pmatrix} N_t/\lambda^2 + t^2 e^{\mu + \lambda t} & t e^{\mu + \lambda t} \\ t e^{\mu + \lambda t} & e^{\mu}(e^{\lambda t} - 1). \end{pmatrix}. \tag{12}$$

Setting $\partial \ell_t / \partial \mu = 0$, we obtain

$$e^{\mu} = N_t/(e^{\lambda t} - 1). \tag{13}$$

Plugging this into $\partial \ell_t / \partial \lambda$ leads to

$$\frac{\partial \ell_t}{\partial \lambda} = N_t/\lambda + \sum_{i=1}^{N_t} x_i - t e^{\lambda t} N_t/(e^{\lambda t} - 1). \tag{14}$$

Given data by time t, $\lim_{\lambda\to\infty}\partial\ell_t/\partial\lambda = \sum_{i=1}^{N_t} x_i - tN_t$, which is negative because $x_{N_t} \leq t$ and $x_i < t \forall i < N_t$; and $\lim_{\lambda\to 0+}\partial\ell_t/\partial\lambda = \sum_{i=1}^{N_t} x_i - tN_t/2$, which tends to ∞ w.p.1 under P_{θ_0}. Moreover, since (14) is strictly decreasing in λ, it has a unique root $\hat{\lambda}_t > 0$ w.p.1 under P_{θ_0} when t is large enough. Then from (13), $\hat{\mu} = \log\{N_t/(e^{\hat{\lambda}t} - 1)\}$. From (12), it can be verified that $-\nabla^2\ell_t$ is positively definite at $\hat{\theta} = (\hat{\lambda}, \hat{\mu})'$; and therefore, $\hat{\theta}$ is the unique maximum likelihood estimator. In [1], it is shown that

$$\hat{\mu}_t - \mu_0 \xrightarrow{p} 0 \tag{15}$$

and

$$t(\hat{\lambda}_t - \lambda_0) \xrightarrow{p} 0. \tag{16}$$

The verification of $\det(B_t^{-1}) \xrightarrow{p} 0$ can be easily obtained by (12).

We are now verifying the other conditions for the asymptotic posterior normality holds. The condition regarding the prior may be considerably relaxed.

References

[1] N. Keiding. Estimation in the birth process. *Biometrika*, 61:71–80, 1974.

[2] R. C. Weng. On stein's identity for posterior normality. *Statistica Sinica*, 13:495–506, 2003.

[3] M. Woodroofe. Very weak expansions for sequentially designed experiments: linear models. *Ann. Statist.*, 17:1087–1102, 1989.

[4] M. Woodroofe. Integrable expansions for posterior distributions for one-parameter exponential families. *Statistica Sinica*, 2:91–111, 1992.