Stability and Convergence of the Posterior in Non-Regular Problems

Jayanta K. Ghosh*, Subhashis Ghosal and Tapas Samanta

Indian Statistical Institute, Calcutta

Abstract. It is shown that the posterior converges in a weak sense in a fairly general set up which includes the exponential, gamma and Weibull with a location parameter, and a reliability change point problem. A representation is obtained for the limiting posterior in case a stronger convergence, e.g. almost sure convergence, holds. This leads to a necessary condition under the above general set up and settles the question of almost sure convergence in the above examples. Under very general conditions it is also shown that the posterior is asymptotically free of the prior. Though primarily developed for non-regular problems, all the theorems apply to regular cases also.

1 Introduction

In Bayesian analysis, one starts with a prior and the resultant analysis is based on the posterior, given data. Since this involves a prior, naturally we are interested to know to what extent the Bayesian analysis is sensitive to the choice of a prior. Indeed, under mild conditions the prior has little or no effect if sample size is large, so that almost same conclusions will follow from almost any reasonable prior, i.e., we have almost complete prior robustness.

The next natural question is whether the posterior stabilizes as the sample size increases indefinitely. In this case the inference stabilizes and we want to know whether the posterior approaches a simple form. If so, the Bayesian analysis is remarkably simple and the approximate computation based on the simplified form is often quite accurate, even for moderately large sample sizes. (For an example, see Berger (1985, § [illegible])) Also, if one is interested in studying the frequentist coverage probabilities of Bayesian confidence intervals, the job is very much simplified.

The problem has been well investigated in the so called "regular" cases, where it has been observed that for a wide variety of priors the posterior, suitably normalized and centered at the maximum likelihood estimator (MLE), tends to the standard normal distribution. This fact was first observed by Laplace (1774) and more recently by Bernstein (1917) and von Mises (1931) and subsequently referred to as the Bernstein-von Mises Theorem or the Bayesian central limit theorem. Le Cam (1953, 1958) gave a rigorous proof of this result for independent and identically distributed (i.i.d.) observations. Various modifications and extensions of this result have been made by several authors including Bickel and Yahav (1969), Walker and Chao (1970), Dawid (1970), Heyde and Johnstone (1979), Kallianpur, Borwanker and Prakasa Rao (1971), Clarke and Barron (1990). A detailed discussion on the conditions underlying the Bernstein-von Moses theorem can be found in Le

*This work was completed while the first author was visiting Purdue University.

Cam (1970). Refinements of posterior normality are considered in Johnson (1970) and Ghosh, Sinha and Joshi (1982).

In this paper, we consider a general situation including a wide variety of non-regular cases. So far, the asymptotic behavior of posterior has not been studied in non-regular cases except in Samanta (1988), where a non-normal limit of posterior has been obtained for a particular type of discontinuous densities.

The main results of this paper are stated and proved in Sect. 2. It is shown that under mild conditions the posterior is asymptotically free of the prior.

Also we observe that a weak limit of posterior probability of any Borel set always exists under the general set up of Ibragimov and Hasminskii (1981). Indeed, it is shown that posterior densities as random processes in L^1 space converge and the limit is also identified. Although theoretically interesting, the weak limit itself is not quite useful, and hence we investigate whether a suitably centered posterior converges to a limit in a stronger sense. We confine ourselves to the i.i.d. case and observe that a limit, if it exists, will be free of the sample sequence. Finally, we exhibit a certain representation as a necessary condition for the existence of such limits.

In Sect. 3, we use the results of Sect. 2 in several important families of distributions, including the exponential, gamma and a reliability change point problem. The examples considered are regular cases, discontinuous densities and densities with singularities. In many of the non-regular examples, the representation theorem of Sect. 2 lead to non-existence of a.s. limit or limit in probability.

Our interest in these problems arose from the Bayesian study of the reliability problem of Ghosh, Joshi, and Mukhopadhyay (1992a,b) and the questions raised by Smith (1985), to which our attention was drawn by Louis Pericchi. Some of Smith's questions are answered here.

2 Main Results

Let $\{\mathcal{X}^{(n)}, \mathcal{A}^{(n)}, P_\theta^{(n)}; \theta \in \Theta\}$ be a sequence of statistical experiments generated by observation $X^{(n)} \in \mathcal{X}^{(n)}$ where $\Theta \subset \mathbb{R}^k$ is a Borel set with non-empty interior. We assume that for each $n \geq 1$, there exists a σ-finite measure $\nu^{(n)}$ on $(\mathcal{X}^{(n)}, \mathcal{A}^{(n)})$ dominating the family $\{P_\theta^{(n)}; \theta \in \Theta\}$ and let the Radon-Nikodym derivative

$$p^{(n)}(x^{(n)}, \theta) = \frac{dP_\theta^{(n)}}{d\nu^{(n)}}(x^{(n)}) \ .$$

Let $\pi(\cdot)$ be a prior density (possibly improper) on Θ. The posterior density given $X^{(n)}$, by Bayes' theorem, is given by

$$\pi(\theta|X^{(n)}) = \frac{\pi(\theta)p^{(n)}(X^{(n)}, \theta)}{\int_\Theta \pi(\theta)p^{(n)}(X^{(n)}, \theta)d\theta}$$

provided the integral converges. We assume that the posterior exists for (sufficiently large) n. The posterior probability of a set $A \subset \Theta$ is denoted by

$$\pi(\theta \in A|X^{(n)}) = \int_A \pi(\theta|X^{(n)})d\theta \ .$$

Unless otherwise indicated, integrals with respect to θ are taken over the whole of Θ.

The following theorem shows that under certain conditions, the posterior is insensitive to the choice of prior for large n.

Theorem 2.1 *Let θ_0 be an interior point of Θ and π_1, π_2 be two prior densities which are positive and continuous at θ_0. We assume that the posterior $\pi_i(\theta|X^{(n)})$ is consistent for $i = 1, 2$, i.e., given any $\eta > 0$ and a neighborhood V of θ_0, there exists $n_0 \geq 1$ such that for $i = 1, 2$*

$$\pi_i(\theta \in V|X^{(n)}) \geq 1 - \eta \text{ for all } n \geq n_0 \text{ a.s. } (\theta_0) \tag{2.3}$$

Then

$$\lim_{n\to\infty} \int |\pi_1(\theta|X^{(n)}) - \pi_2(\theta|X^{(n)})| d\theta = 0 \text{ a.s. } (\theta_0) \tag{2.4}$$

Proof. From consistency of posterior, it follows that for $n \geq n_0$

$$\int |\pi_1(\theta|X^{(n)}) - \pi_2(\theta|x^{(n)})| d\theta \leq \int_V \pi_1(\theta|X^{(n)})|1 - \frac{\pi_2(\theta|X^{(n)})}{\pi_1(\theta|X^{(n)})}| d\theta + 2\eta. \tag{2.5}$$

Let $\delta > 0$ be given. Then by continuity of π_i, get a neighborhood V of θ_0 so that on V

$$\pi_i(\theta_0)(1-\delta) \leq \pi_i(\theta) \leq \pi_i(\theta_0)(1+\delta) \tag{2.6}$$

and hence

$$(1-\delta)\pi_i(\theta_0)C_n \leq \int_V \pi_i(\theta)p^{(n)}(X^{(n)}, \theta)d\theta \leq (1+\delta)\pi_i(\theta_0)C_n \tag{2.7}$$

where

$$C_n = \int_V p^{(n)}(X^{(n)}, \theta)d\theta.$$

Clearly

$$\int_V \pi(\theta)p^{(n)}(x^{(n)}, \theta)d\theta \leq \int \pi_i(\theta)p^{(n)}(x^{(n)}, \theta)d\theta$$
$$\leq (1-\eta)^{-1} \int_V \pi_i(\theta)p^{(n)}(x^{(n)}, \theta)d\theta$$

and hence using (2.6)

$$\frac{(1-\eta)(1-\delta)p^{(n)}(x^{(n)}, \theta)}{(1+\delta)C_n} \leq \pi_i(\theta|X^{(n)}) \leq \frac{(1+\delta)}{(1-\delta)C_n} p^{(n)}(x^{(n)}, \theta) \ .$$

Thus we have

$$(1-\eta)\left(\frac{1-\delta}{1+\delta}\right)^2 \leq \frac{\pi_1(\theta|X^{(n)})}{\pi_2(\theta|X^{(n)})} \leq (1-\eta)^{-1}\left(\frac{1+\delta}{1-\delta}\right)^2 \ . \tag{2.8}$$

By choosing δ and η small enough and putting (2.8) in (2.5), we have the desired

Remark 2.1. If $\mathcal{F}$ is a family of priors such that $\pi(\theta_0) > 0$ for all π in $\mathcal{F}$, $\mathcal{F}$ is equicontinuous at θ_0 and (2.3) is satisfied uniformly in $\pi \in \mathcal{F}$, then

$$\lim_{n\to\infty} \sup\left\{\int |\pi_1(\theta|X^{(n)}) - \pi_2(\theta|X^{(n)})|d\theta : \pi_1, \pi_2 \in \mathcal{F}\right\} = 0 \quad \text{a.s.} \qquad (2.[illegible])$$

Remark 2.2. The assumption (2.3) in Theorem 2.1 about posterior consistency is generally a mild condition in the case of a finite dimensional parameter space. See, in this connection, Diaconis and Freedman (1986) and references therein. Condition (2.3) of Theorem 2.1 holds if Conditions (IH) below hold and $\Sigma||(\varphi_n)||^s < \infty$ for some $s > 0$. The latter condition holds in all common examples, where (φ_n) is a power of n.

We now investigate the existence of posterior limit, which, if it exists, is independent of the prior in view of Theorem 2.1. The first result, essentially due to Ibragimov-Hasminskii (1981) (henceforth abbreviated as IH) states that under quite general conditions, posterior probabilities of the normalized parameter converge weakly. The set up and assumptions are described below.

Let $\{\varphi_n\}$ be a sequence of $k \times k$ positive definite matrices converging to [illegible]. Let $\theta_0 \in \operatorname{int} \Theta$,

$$u = \varphi_n^{-1}(\theta - \theta_0)$$
$$\text{and } U_n = \{u \in \mathbb{R}^k : \theta + \varphi_n u \in \Theta\} \; .$$

Then U_n is a neighborhood of 0 in $\mathbb{R}^k$ tending to $\mathbb{R}^k$. Define the "likelihood [illegible] process"

$$Z_n(u) = Z_{n,\theta_0}(u) = \frac{p^{(n)}(x^{(n)}; \theta_0 + \varphi_n u)}{p^{(n)}(x^{(n)}; \theta_0)} \qquad [illegible]$$

considered as a random function in $u \in U_n$. For studying various asymptotic [illegible]erties, it is necessary to choose φ_n properly. See IH (1981) in this context [illegible] essentially unique choice of φ_n is shown.

From now onwards, expectations and probabilities refer to the 'true par[illegible] θ_0.

Conditions (IH).

1. For some $\alpha > 0$, $K > 0$, $a > 0$, $n_0 \geq 1$

   $$\sup\{||u_2 - u_1||^{-\alpha} E|Z_n^{1/2}(u_2) - Z_n^{1/2}(u_1)|^2 : ||u_1|| \leq R_1 ||u_2|| \leq R\} \leq K([illegible]$$

   for all $n \geq n_0$.
2. For all $u \in U_n, n \geq n_0$

   $$EZ_n^{1/2}(u) \leq \exp[-g_n(u)]$$

   where $\{g_n\}$ is a sequence of real valued functions on $[0, \infty)$ satisfying [illegible]

   (a) For fixed $n \geq 1, g_n(y) \uparrow \infty$ as $y \to \infty$

   (b) For any $N > 0$, $\lim_{\substack{y\to\infty \\ n\to\infty}} y^N \exp[-g_n(y)] = 0$

3. Let $\{Z(u) : u \in \mathbb{R}^k\}$ be a stochastic process not identically zero such [illegible] dimensional distribution of Z_n converge to that of Z.

Notation: Unless otherwise indicated, integral with respect to u is taken over the whole set where integrand is defined. We also write

$$\xi_n(u) = \frac{\pi(\theta_0 + \varphi_n u)Z_n(u)}{\int \pi(\theta_0 + \varphi_n u)Z_n(u)du} \text{ and } \xi(u) = \frac{Z(u)}{\int Z(u)du} \; .$$

Theorem 2.2 *Let Conditions (IH) be satisfied and π be a prior density continuous and positive at θ_0. Then for any Borel set A in $\mathbb{R}^k$,*

$$\pi(u \in A|X^{(n)}) \xrightarrow{d} \frac{\int_A Z(u)du}{\int Z(u)du} \; . \qquad (2.13)$$

The proof is implicit in the proof of Theorem I.10.2 of IH (1981).

Remark 2.3. Let $\mathcal{A}$ be any countable subcollection of $\mathcal{B}^k$, the Borel σ-field in $\mathbb{R}^k$. Then by a slight modification of the proof, it can be shown that the $\mathbb{R}^\infty$-valued process $(\int_A \xi_n(u)du : A \in \mathcal{A})$ converges in distribution to the $\mathbb{R}^\infty$-valued process $(\int_A \xi(u)du : A \in \mathcal{A})$ under Conditions (IH).

Let $\mathcal{P}$ denote the space of all absolutely continuous probabilities on $\mathbb{R}^k$ equipped with the total variation norm. $\mathcal{P}$ is isometrically identified with the space of all probability densities on $\mathbb{R}^k$ with L^1-norm.

The next result is somewhat technical and is a strengthened version of Theorem 2.2. This is used in Theorem 2.5 and is also of independent interest. Note that by Lemma A.4, one can consider ξ_n and ξ to be $\mathcal{P}$-valued or $L^1(\mathbb{R}^k)$-valued random variables.

Theorem 2.3 *The process ξ_n converges to the process ξ in $L^1(\mathbb{R}^k)$.*

Proof In view of Remark 2.3, it is enough to verify that $\{\xi_n\}$ forms a tight family. We use Lemma A.1. The first condition is trivial and it is enough to verify that

(a) $\sup\{\int_{||u|| \leq M} |\xi_n(u + x) - \xi_n(u)|du : n \geq n_0, ||x|| < \delta\}$

and

(b) $\sup\{\int_{||u|| > M} \xi_n(u)du : n \geq n_0\}$

are arbitrarily small with probability arbitrarily close to one if one chooses δ small and M large enough.

This has been verified for (b) in IH (1981). To verify this for (a) we note that

$$\int_{||u|| \leq M} |\xi_n(u + x) - \xi_n(u)|du$$
$$\leq (\int Z_n(u)du)^{-1} \int_{||u|| \leq M} |Z_n^{1/2}(u + x) - Z_n^{1/2}(u)||Z_n^{1/2}(u + x) + Z_n^{1/2}(u)|du$$
$$\leq 2(\int Z_n(u)du)^{-1} \{\int_{||u|| \leq M} |Z_n^{1/2}(u + x) - Z_n^{1/2}(u)|^2 du\}^{1/2}$$
$$\{\int_{||u|| \leq M} Z_n(u)du\}^{1/2} \; . \qquad (2.14)$$

by Cauchy-Schwartz inequality and the inequality $(a+b)^2 \le 2(a^2+b^2)$. Expression (2.14) is in turn less than or equal to

$$4\Big(\int Z_n(u)du\Big)^{-1/2}\Big\{\int_{||u||\le M}|Z_n^{1/2}(u+x)-Z_n^{1/2}(u)|^2du\Big\}^{1/2} .$$

Now by Condition (IH)(1) and by Chebyshev's inequality we have

$$P\{|Z_n^{1/2}(u+x)-Z_n^{1/2}(u)|>\varepsilon\} \le \frac{K\delta^{\alpha}(1+(M+\delta)^{\alpha})}{\varepsilon^2}$$

which can be made arbitrarily small whatever be ε and M. Thus with high probability the inequality

$$\int_{||u||\le M}|\xi_n(u+x)-\xi_n(u)|du \le 4\varepsilon\sqrt{M}\Big(\int Z_n(u)du\Big)^{1/2} \qquad (2.1$$

is satisfied. Clearly the last term can be made arbitrarily small uniformly in $n \ge$ if ε is small enough. Thus the result follows.

Remark 2.4. A result similar to Theorem 2.2 for the posterior with a random centering $\hat\theta$ also holds if $\hat\theta$ is such that

$$(\varphi_n^{-1}(\hat\theta-\theta_0),\xi_n(\cdot)) \xrightarrow{d} (W,\xi(\cdot)) \text{ in } \mathbb{R}^k \times L^1(\mathbb{R}^k) .$$

In this case the posterior probability of a set is given by (2.33).

The above technical results give only the weak limit of posterior probab However, we have familiar examples where the suitably centered posterior a limit almost surely. In regular case, the posterior centered at the MLE a normal distribution almost surely under certain conditions. Therefore one investigate whether the posterior, with suitable centering goes to a limit surely, or at least in probability. In the remaining portion of this section, necessary conditions for the existence of such a centering under a general

We confine ourselves only to the iid case where observation $X^{(n)}$ is $(X_1,\dots,X_n)$ and $p^{(n)}(X^{(n)},\theta)=\Pi_{i=1}^n f(X_i,\theta)$, $f(\cdot,\theta)$ is a probability den respect to a σ-finite measure ν on a standard Borel space $(\mathcal{X},\mathcal{A})$.

The next result shows that a limit, if it exists, must be free of the sample

Proposition 2.1 *Let $\hat\theta=\hat\theta(X_1,\dots,X_n)$ and $T=T(X_1,\dots,X_n)$ be functions of $X_1,\dots,X_n$ which may or may not involve θ_0. Put $v=T^{-1}($ let $A\in B^k$.*

Suppose for each sample sequence $\boldsymbol{X}$, there exists $c(\boldsymbol{X})$ such that

$$\pi(v\in A|X^{(n)}) \xrightarrow{P} c(\boldsymbol{X}) .$$

Then $c(\boldsymbol{X})$ does not depend on the sample sequence $\boldsymbol{X}$.

Proof. The posterior density of v is

$$\pi(v|X^{(n)})=\frac{\pi(\hat\theta+Tv)\Pi_{n=1}^n f(X_i,\hat\theta+Tv)}{\int\pi(\hat\theta+Tv)\Pi_{i=1}^n f(X_i,\hat\theta+Tv)dv} . \qquad (2.17)$$

This is a symmetric function of $X_1,\dots,X_n$ and so is

$$\pi(v\in A|X^{(n)})=\int_A \pi(v|X^{(n)})dv . \qquad (2.18)$$

By going through a subsequence, if necessary, (2.16) can be assumed to hold a.s. By an application of the Hewitt-Savage zero-one law (Chow-Teicher, 1988), it follows that $c(\boldsymbol{X})$ does not depend on $\boldsymbol{X}$.

Often it is true that

$$\varphi_n T \xrightarrow{P} \Sigma \qquad (2.19)$$

where Σ is a p.d. matrix. In such cases, one may assume that $T=\varphi_n^{-1}$.

The following definition will be useful.

Definition 2.1. An $\mathbb{R}^k$-valued random variable $\hat\theta=\hat\theta(X_1,\dots,X_n)$, symmetric in its arguments, is called a proper centering if for each $A\in B^k$, there exists a non-random $Q(A)$ such that

$$\sup\{|\pi(\varphi_n^{-1}(\theta-\hat\theta)\in A|X^{(n)})-Q(A)|: A\in B^k\} \xrightarrow{P} 0 . \qquad (2.20)$$

$\hat\theta$ is called a semiproper centering or wide sense proper centering if for each $A\in B^k$,

$$\pi(\varphi_n^{-1}(\theta-\hat\theta)\in A|X^{(n)}) \xrightarrow{P} Q(A) . \qquad (2.21)$$

Remark 2.5. If (2.20) is satisfied, it automatically follows that $Q(\cdot)$ is an absolutely continuous countably additive probability, and if (2.21) is satisfied, then $Q(\cdot)$ is a finitely additive probability. Proposition 2.1 makes clear why $Q(\cdot)$ must be non-random.

Proposition 2.2 *Let $\hat\theta$ be a proper centering such that $W_n=\varphi_n^{-1}(\hat\theta-\theta_0)$ converges weakly to a random variable W. Then for any countable subcollection $\mathcal{A}$ of B^k, the $\mathbb{R}^\infty$-valued process $\{\pi(\varphi_n^{-1}(\theta-\hat\theta)\in A|X^{(n)}): A\in\mathcal{A}\}$ converges weakly to the process $\{Q(A-W): A\in\mathcal{A}\}$.*

Proof. By weak convergence theory in $\mathbb{R}^\infty$, it is enough to prove the result for a finite collection $\{A_1,\dots,A_r\}$. Since Q is absolutely continuous, the mapping

$$x\mapsto (Q(A_1-x),\dots,Q(A_r-x))$$

is continuous by Lemma A.2 and hence

$$(Q(A_1-W_n),\dots,Q(A_r-W_n)) \xrightarrow{d} (Q(A_1-W),\dots,Q(A_r-W)) .$$

$$\max_{1\le i\le r}|\pi(\varphi_n^{-1}(\theta-\theta_0)\in A_i|X^{(n)})-Q(A_i-W_n)|$$
$$\le \sup\{|\pi(\varphi_n^{-1}(\theta-\hat\theta)\in A|X^{(n)})-Q(A)|: A\in B^k\}$$

goes to zero in probability. By Slutsky's theorem, the result now follows.

Proposition 2.3 *Assume Conditions (IH) and let $\hat{\theta}$ be a proper centering. Then $\varphi_n^{-1}(\hat{\theta} - \theta_0)$ is weakly convergent.*

Proof. We first show that $W_n := \varphi_n^{-1}(\hat{\theta} - \theta_0)$ is tight. If not, there exists $\epsilon > 0$ such that for any $\lambda > 0$, there is a subsequence $\{n'\}$ of $\{n\}$ for which

$$P\{\|W_{n'}\| > \lambda\} > \epsilon \quad \text{for all } n' . \tag{2.22}$$

Put

$$u = \varphi_n^{-1}(\theta - \theta_0)$$
$$v = \varphi_n^{-1}(\theta - \hat{\theta}) .$$

Then

$$\pi(v \in A | X^{(n)}) = \int_{A+W_n} \xi_n(u) du . \tag{2.23}$$

Fix a bounded set A and using arguments of IH (1981), find M large enough so that $\int_{\|u\|>M} \xi_n(u)du$ can be made as small as desired with probability close to one uniformly in $n \geq n_0$. Choose $\lambda > 0$ large enough such that $\|x\| > \lambda$ implies

$$A + x \subset \{\|u\| > M\} \tag{2.24}$$

combining (2.22) to (2.24), it follows that one must have $Q(A) = 0$. Clearly this cannot be true for every bounded set. So W_n is tight.

If W and W' are two subsequential limits, then by Proposition 2.2,

$$Q(A - W) \stackrel{d}{=} Q(A - W') \quad \text{for all } A \in B^k .$$

An application of Lemma A.3 completes the proof.

Remark 2.6. The conclusion of Proposition 2.3 is still valid even if Conditions (IH) are not satisfied, instead there exists a proper centering $\tilde{\theta}$ such that $\varphi_n^{-1}(\tilde{\theta} - \theta_0)$ is tight.

Theorem 2.4 *Assume Conditions (IH) and let $Z(u) = \exp[Y(u)]$. If a proper centering $\hat{\theta}$ exists, then there exists a random variable W such that*

$$\varphi_n^{-1}(\hat{\theta} - \theta_0) \xrightarrow{d} W$$

and for almost all $u \in \mathbb{R}^k$, $\xi(u - W)$ is non-random, equivalently for all $u_0, u_1 \in \mathbb{R}^k$,

$$Y(u_1 - W) - Y(u_0 - W) \text{ is non-random.}$$

Proof. By Proposition 2.3, such a W exists.

By Lemma A.3, $\zeta := \{Q(A - W): A \in B^k\}$ is an $\mathcal{M}_Q$-valued random element. Fix a countable field $\mathcal{A}$ which generates B^k. By Remark 2.3 and Proposition 2.2,

$$\left(\int_A \xi(u)du : A \in \mathcal{A}\right) \stackrel{d}{=} (Q(A - W): A \in \mathcal{A})$$

and hence $\zeta \stackrel{d}{=} \xi$.

Since $P\{\zeta \in \mathcal{M}_Q\} = 1$, we also have $P\{\xi \in \mathcal{M}_Q\} = 1$. Define ψ as in Lemma A.3 and hence by (2.26)

$$(\xi, \psi^{-1}(\xi)) \stackrel{d}{=} (\zeta, \psi^{-1}(\zeta)) = (\zeta, W) . \tag{2.27}$$

Define a map $\Lambda: \mathcal{P} \times \mathbb{R}^k \to \mathcal{P}$ by $\Lambda(Q, x) = Q^x$ where $Q^x(A) = Q(A + x)$ for all A in B^k.

By Lemma A.3, for a fixed Q, Λ is continuous in x whereas for fixed x, Λ is an isometry in Q. Thus (2.27) implies that

$$\Lambda(\xi, \psi^{-1}(\xi)) \stackrel{d}{=} \Lambda(\zeta, W) . \tag{2.28}$$

Putting $\psi^{-1}(\xi) = W^*$, we have $W^* \stackrel{d}{=} W$ and

$$\int_{A+W^*} \xi(u)du = Q(A) \quad \text{for all } A \in B^k ,$$

i.e.,

$$\int_A \xi(u - W^*)du = Q(A) \text{ for all } A \in B^k . \tag{2.29}$$

The conclusion is now immediate.

The next result shows that a proper centering, if it exists, is essentially unique.

Proposition 2.4 *Assume Conditions (IH) and let $\hat{\theta}$ and $\tilde{\theta}$ be two proper centerings. Then the associated probabilities and weak limits are shifts of each other.*

Proof. Let Q_1, Q_2 denote the associated measures and W_1, W_2 denote the weak limits of $\varphi_n^{-1}(\hat{\theta} - \theta_0)$ and $\varphi^{-1}(\tilde{\theta} - \theta_0)$ respectively (which exists by Proposition 2.3). By Proposition 2.2, it follows that $\mathcal{P}$-valued random process $\{Q_1(A - W_1): A \in B^k\}$ has the same distribution as $\{Q_2(A - W_2): A \in B^k\}$. Hence it follows that $Q_2 \in \mathcal{M}_{Q_1}$, say, $Q_2(A) = Q_1(A + c)$. Using arguments similar to Theorem 2.4, it follows that $W_2 \stackrel{d}{=} W_1 + c$.

Remark 2.7. Conditions (IH) are used only to guarantee the existence of the stated weak limits. So one can use Remark 2.6 instead of Proposition 2.3.

In the remaining part of this section, we give a partial answer to the question whether a semiproper centering exists.

Definition 2.2. A semiproper centering $\hat{\theta}$ is called regular if there exists a continuous function Λ from $\mathcal{P}$ to $\mathbb{R}^k$ such that

$$\varphi_n^{-1}(\hat{\theta} - \theta_0) = \Lambda(\xi_n) . \tag{2.30}$$

Theorem 2.5 *Assume Conditions (IH) and let $\hat{\theta}$ be a regular semiproper centering with associated measure Q. Then Q is countably additive and there exists a random variable W satisfying*

(a) $\varphi_n^{-1}(\hat{\theta} - \theta_0) \xrightarrow{d} W$

(b) $\xi(u - W)$ *is non-random for almost every* u .

Further, if $\tilde{\theta}$ is another regular semiproper centering with associated weak limit W', then W' is a shift of W.

Proof. By Theorem 2.3, it follows that (a) is satisfied with $W = \Lambda(\xi)$ and

$$(\xi_n, \varphi_n^{-1}(\hat{\theta} - \theta_0)) \xrightarrow{d} (\xi, W) \ . \tag{2.31}$$

For any Borel Set A in $\mathbb{R}^k$, the map $(f, x) \mapsto \int_{A+x} f(u)du$ from $L^1(\mathbb{R}^k) \times \mathbb{R}^k$ to $\mathbb{R}$ is continuous by lemma A.2. So (2.31) now implies that

$$\pi(\varphi_n^{-1}(\theta - \hat{\theta}) \in A | X^{(n)}) \xrightarrow{d} \int_{A+W} \xi(u)du \tag{2.32}$$

combining (2.21) with (2.32)

$$\int_A \xi(u - W)du = Q(A) \ . \tag{2.33}$$

Since the left hand side of (2.33) is countably additive, so is Q and this proves the first part.

Also (2.33) implies that

$$Q(A + W) = \int_A \xi(u)du$$

from which the second part follows as in Proposition 2.4.

The condition on $\hat{\theta}$ is likely to be satisfied if it is taken as a quantile of the posterior.

3 Examples

We now apply the results established in Sect. 2 to several important families of distributions.

Example 3.1. *Regular case.* It is well known that in the usual regular cases the posterior distribution centered at the MLE converges to a normal distribution in variation norm a.s. In these cases, the limiting likelihood ratio process is given by

$$Z(u) = \exp(u'\Delta - \frac{1}{2}u'Iu)$$

where I is Fisher's information matrix and Δ is a random vector having distribution $N_k(\underline{0}, I)$. Indeed, if one assumes conditions of Sect. III.3.1 of IH (1981), Conditions (IH) of Sect. 2 are satisfied. One can easily see that the necessary condition stated in Theorem 2.4 is satisfied with $W = I^{-1}\Delta$.

Example 3.2. *Non-Regular Case — Densities with Jumps.* We consider the set up and assumption of IH (1981, Chap. V, p. 242). We have a sequence of i.i.d. observations with values in $\mathbb{R}$ and common density $f(x, \theta)$ with respect to the Lebesgue measure where the parameter set Θ is an open interval (finite or infinite) in $\mathbb{R}$. Let $f(x, \theta)$ possess r jumps at $x_1(\theta), \ldots, x_r(\theta)$ and let

$$p_i(\theta) = \lim_{x \downarrow x_i(\theta)} f(x, \theta), q_i(\theta) = \lim_{x \uparrow x_i(\theta)} f(x, \theta), \quad i = 1, 2, \ldots, r \ .$$

We fix $\theta_0 \in \Theta$ and write p_i, q_i, x_i and x_i' in place of $p_i(\theta_0), q_i(\theta_0)$ $x_i(\theta_0)$ and $x_i'(\theta_0)$ respectively. It is shown in IH (1981) that Conditions (IH) of Sect. 2 are satisfied in this case with $\varphi(n) = n^{-1}$.

Whether a limit of the posterior exists or not depends on the nature of the jumps. Below we consider several important special cases.

Case 1. Assume that for each $i = 1, 2, \ldots, r$ one of the numbers p_i and q_i is zero. Set

$$\Gamma^+ = \{i : q_i = 0 \text{ and } x_i' > 0\} \bigcup \{i : p_i = 0 \text{ and } x_i' < 0\},$$

$$\Gamma^- = \{i : q_i = 0 \text{ and } x_i' < 0\} \bigcup \{i : p_i = 0 \text{ and } x_i' > 0\}$$

$$\text{and} \qquad c = \sum_{i=1}^{r} (p_i - q_i) x_i' \ .$$

Subcase 1A. Suppose that both Γ^+ and Γ^- are nonempty. In this case the limiting likelihood ratio process is given by

$$Z(u) = \begin{cases} e^{cu}, \text{ if } -\tau^- < u < \tau^+ \\ 0, \quad \text{otherwise} \end{cases}$$

where τ^- and τ^+ are independent exponentially distributed random variables with parameters $\alpha = \sum_{i \in \Gamma^-} (q_i - p_i) x_i'$ and $\beta = \sum_{i \in \Gamma^+} (p_i - q_i) x_i'$ respectively.

If $c = 0$, we have

$$\xi(u) = \begin{cases} (\tau^+ + \tau^-)^{-1}, \text{ if } -\tau^- < u < \tau^+ \\ 0, \quad \text{otherwise,} \end{cases}$$

and in case $c \neq 0$, we have

$$\xi(u) = \begin{cases} \frac{c e^{cu}}{e^{c\tau^+} - e^{-c\tau^-}}, \text{ if } -\tau^- < u < \tau^+ \\ 0, \quad \text{otherwise} \ . \end{cases}$$

In both cases, it is clear that the necessary condition of Theorem 2.4 is not satisfied and hence no a.s. limit of posterior exists. Simple examples of this kind are shifts of $U(0, 1)$ (with $c = 0$) and $U(\theta, 2\theta)$ (with $c \neq 0$).

Subcase 1B. Suppose that one of Γ^- and Γ^+ is empty. In case Γ^- is empty (the other case is similar), we have

$$\xi(u) = \begin{cases} c \exp\{c(u - \tau^+)\}, \text{ if } u < \tau^+ \\ 0, \quad \text{otherwise} \end{cases}$$

The necessary condition of Theorem 2.4 holds with $W = -\tau^+$.

Indeed a limit has been obtained in Samanta (1988) for a special situation of Subcase 1B, where the support of the density is an interval which is either increasing or decreasing in θ. Samanta (1988) assumed conditions similar to those of Weiss and Wolfowitz (1974, Chap. 5) and a uniform integrability type condition on $\log f$ and obtained an exponential limit. In these situations there exists a statistic Z_n such that the set $\{(X_1,\ldots,X_n)\colon f(X_i,\theta)>0 \text{ for all } i\}$ can be expressed as $\{Z_n(\underset{\sim}{X})>\theta\}$ or $\{Z_n(\underset{\sim}{X})<\theta\}$ according as the support is decreasing or increasing in θ. This Z_n acts as a proper centering.

Important examples of this case are shifts of exponential density, $U(0,\theta)$ etc.

Case 2. We now consider the case when both p_k and q_k are positive. We only consider the case with $r=1$ and $x_1'>0$.

In this case we have

$$Y(u)=\begin{cases}(p_1-q_1)u+\nu^+(u)\log\frac{q_1}{p_1}, & \text{if } u\ge 0\\ (p_1-q_1)u-\nu^-(u)\log\frac{q_1}{p_1}, & \text{if } u<0\end{cases}$$

where $\nu^+(u)$ and $\nu^-(-u)$ are independent homogeneous Poisson process with rates p_1x_1' and q_1x_1' respectively. One can show that the necessary condition of Theorem 2.4 is not satisfied.

An important example of this kind is the change point problem with

$$f(x,\theta)=\begin{cases}a\exp(-ax), & \text{if } 0<x<\theta\\ b\exp\{-a\theta-b(x-\theta)\}, & \text{if } x>\theta,\end{cases}$$

where $a>0$, $b>0$ $(a>b)$ are known constants and $\theta>0$ is the parameter of interest. See in this connection Basu, Ghosh and Joshi (1986) and Ghosh, Joshi and Mukhopadhyay (1992a,b).

Example 3.3. *Non-Regular Case — Densities with Singularities.* We consider a sequence of i.i.d. observations with density $f(x-\theta)$ where θ is a real parameter and $f(x)$ admits the representation

$$f(x)=\begin{cases}0, & \text{if } x<0,\\ p(x)|x|^\alpha & \text{if } x>0\end{cases}$$

in a neighborhood of zero, where $p(x)$ is a continuous function with $p=p(0)$ [illegible]. In this case we say that $f(x)$ has a singularity of order α at the point zero provided $-1<\alpha<1$. This example is a special case of singularity treated in IH (where it is shown that in presence higher order singularity (smaller value of α) singularities of lower orders do not affect the asymptotic analysis. In this case we assume that there is only one singularity of the highest order α and the following conditions hold:

(1) There exists a number $\lambda>1+\alpha$ such that for any neighborhood N of 0

$$\int_{N^c}|f^{1/\lambda}(x-h)-f^{1/\lambda}(x)|^\lambda dx=0(|h|^\lambda)\ .$$

(2)

$$\int|x|^\delta f(x)dx<\infty \text{ for some } \delta>0.$$

Under these assumptions, Conditions (IH) of Sect. 2 are satisfied with $\varphi_n = n^{-1/(1+\alpha)}$ and $Y(\cdot)=\log Z(\cdot)$ is given by

$$Y(u)=\begin{cases}\alpha\int\limits_0^\infty\log|1-\frac{u}{x}|(\nu(dx)-E\nu(dx))\\ -p\int\limits_0^\infty(|1-\frac{u}{x}|^\alpha-1-\alpha\log|1-\frac{u}{x}|)|x|^\alpha dx\\ +p\cdot\frac{u^{1+\alpha}}{1+\alpha}, & \text{if } u\ge\tau\\ -\infty, & \text{if } u<\tau\end{cases}$$

where ν is a non-homogeneous Poisson process with rate function $\lambda(x)=\frac{x^\alpha}{\Gamma(1+\alpha)}$ and τ is the first jump of the process ν.

We consider the case $\alpha\ne 0$. The case $\alpha=0$ is treated in Example 3.2. Let W be a random variable and $u_1>u_0$ be real numbers to be chosen later. In order that $Y(u_1-W)-Y(u_0-W)$ be non-random it must be true that the set where it is positive, i.e., the set $\{W+\tau<u_0\}$ is trivial. Using this fact for different u_0, one can show that $W+\tau$ must be constant, say $W=c-\tau$. We choose $u_1>u_0>c$. Putting $u_0'=u_0-c$, $u_1'=u_1-c$ we have

$$\begin{aligned}&Y(u_1-W)-Y(u_0-W)\\&=Y(u_1'+\tau)-Y(u_0'+\tau)\\&=-\alpha p\int_0^\tau[\log|1-\frac{u_1'+\tau}{x}|-\log|1-\frac{u_0'+\tau}{x}|]x^\alpha dx\\&\quad -p\int_0^\infty[g(u_1'+\tau,x)-g(u_0'+\tau,x)]x^\alpha dx\\&\quad +\frac{p}{1+\alpha}[(u_1'+\tau)^\alpha-(u_0'+\tau)^\alpha]\\&\quad +\alpha\int_\tau^\infty[\log|1-\frac{u_1'+\tau}{x}|-\log|1-\frac{u_0'+\tau}{x}|]\bar\nu(dx)\end{aligned}$$

where

$$g(u,x)=|1-\frac{u}{x}|^\alpha-1-\log|1-\frac{u}{x}|$$

and

$$\bar\nu(dx)=\nu(dx)-E\nu(dx)\ .$$

Clearly the first three terms are functions of τ only. If $Y(u_1-W)-Y(u_0-W)$ is non-random, then the conditional distribution of $Y(u_1-W)-Y(u_0-W)$ given τ is also degenerate. Therefore

$$\begin{aligned}&\int_\tau^\infty[\log|1-\frac{u_1'+\tau}{x}|-\log|1-\frac{u_0'+\tau}{x}|]\bar\nu(dx)\\&=\int_0^\infty\log|\frac{x-u_1'}{x-u_0'}\bar\mu(dx)\end{aligned}$$

has a degenerate conditional distribution given τ. Here

$$\bar\mu(dx)=\mu(dx)-E\mu(dx) \text{ and } \mu(x)=\nu(\tau+x).$$

However, given τ, μ is again a non-homogeneous Poisson process with rate function $p(\tau + x)^{\alpha}$. Thus the conditional variance of $\int_0^{\infty} \log|\frac{x-u_1'}{x-u_0'}|\mu(dx)$ is

$$p\int_0^{\infty} \log^2|\frac{x-u_1'}{x-u_0'}|(x+\tau)^{\alpha}dx > 0 \ .$$

This contradiction implies the non-existence of a limit of posterior.

Important examples of this kind are the gamma density

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)}e^{-x}x^{\alpha-1}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

and the Wiebull density

$$f(x) = \begin{cases} \alpha x^{\alpha-1}\exp(-x^{\alpha}), & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

where $0 < \alpha < 2$.

Another example is provided by

$$f(x) = \begin{cases} \frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}, & \text{if } 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

where $0 < a < 2$, $b > 0$ and $a < b$.

The case when $f(x)$ admits a representation

$$f(x) = \begin{cases} q(x)|x|^{\alpha}, & \text{if } x \leq 0 \\ 0, & \text{if } x > 0 \end{cases}$$

in a neighborhood of zero is exactly similar and hence omitted.

Example 3.4. *A Two Parameter Case.* Suppose that the observations are i.i.[illegible] a common density

$$f(x, \theta_1, \theta_2) = \begin{cases} \exp\{-\theta_2(x-\theta_1)\}, & x > \theta_1, \\ 0, & \text{otherwise,} \end{cases}$$

where $-\infty < \theta_1 < \infty$ and $\theta_2 > 0$ are two unknown parameters. As m[illegible] in Examples 3.1 and 3.2 (Case 1), if θ_1, is known, the posterior goes to a [illegible] distribution and on the other hand if θ_2 is known it has an exponential limit[illegible] θ_1 and θ_2 are unknown, the limiting likelihood ratio process is given by

$$Z(u_1, u_2) = \begin{cases} \exp\{cu_1 + u_2\Delta - \frac{1}{2}u_2^2 I\}, & \text{if } u_1 < \tau, \\ 0, & \text{otherwise,} \end{cases}$$

where $c > 0$ and $I > 0$ are constants depending on θ_1, θ_2 and Δ and τ are [illegible] dent random variables following $N(0, I)$ and exponential distribution with p[illegible] c respectively. It is interesting to note that (3.2) is the product of the likeli[illegible] processes obtained in the non-regular case when θ_2 is known and in the re[illegible] when θ_1 is known. It is easy to see that the necessary condition of Theo[illegible] satisfied in this case. We also expect, but don't have a proof yet, that Con[illegible] hold for this example. Indeed, proceeding in a manner similar to that in [illegible] (1988, Chap. 3 and 4) one can show that the posterior has an a.s. limit [illegible] product of exponential and normal.

Remark 3.1. There is a large class of non-regular cases like Examples 3.4 where in addition to a parameter θ_1 with respect to which the problem is nonregular there are other parameters θ_2 with respect to which the problem is regular. Smith (1985) studied a class of such examples in the context of maximum likelihood estimation. Another important example is the change point problem given by (3.1) with both a and b unknown. If Condition (IH) can be shown to be satisfied in all these cases, one may use our Theorem 2.4 to study the limiting behaviour of posterior.

Remark 3.2. For the change point problem with both a and b unknown, by direct calculations as in Chernoff and Rubin (1956), Ghosh, Joshi and Mukhopadhyay (1992b) show that the marginal of θ concentrates on $O(n^{-1})$-neighborhood of θ_0. This fact is then used to show that the conditional posterior of a and b given θ and $x_1, \ldots, x_n$ is approximately normal and free of θ (similar to the phenomenon noted in Example 3.4). One may then use this result to get a reference prior for a, b given θ. For θ a uniform distribution seems most natural. Also a similar method of investigation seems to work in other examples mentioned in Remark 3.1.

Remark 3.3. If no a.s. limit of posterior exists, one may try to find a simple approximation to the posterior which is free of prior. This may help in simulation and also make the computations simpler.

4 Appendix

Lemma A.1 *A subset Γ of $L^1(\mathbb{R}^k)$ has norm-compact closure iff*

(a) $\sup\{||f||: f \in \Gamma\} < \infty$

(b) $\lim_{||x|| \to 0} \sup\{\int_{\mathbb{R}^k} |f(x+y) - f(y)|dy: f \in \Gamma\} = 0$

(c) $\lim_{\lambda \to \infty} \sup\{\int_{||y|| > \lambda} |f(y)|dy: f \in \Gamma\} = 0$

For a proof, see Dunford and Schwartz (1957).

Lemma A.2 *For any $f \in L^1(\mathbb{R}^k)$, the mapping from $\mathbb{R}^k$ into $L^1(\mathbb{R}^k)$ which takes x to $f_x \equiv f(\cdot - x)$ is continuous.*

Lemma A.3 *Let θ be an absolutely continuous probability on $\mathbb{R}^k$ and let $\mathcal{M}_Q = \{Q_x: x \in \mathbb{R}^k\}$ where $Q_x(A) = Q(A - x)$.*

Then the mapping ψ sending x to Q_x is a homeomorphism of $\mathbb{R}^k$ onto $\mathcal{M}_Q$ and $\mathcal{M}_Q$ is closed in P.

Proof Clearly ψ is onto. To show it is one-to-one, let $\psi(x_1) = \psi(x_2)$ for some $x_1 \neq x_2$ and call $\delta = x_1 - x_2$. Then $Q(A) = Q(A + \delta)$, and so for any $n \geq 1$, $Q(A) = Q(A + n\delta)$.

If $A \in \mathcal{B}^k$ is bounded, $Q(A + n\delta) \to 0$ implying $Q(A) = 0$. Clearly this cannot happen for every bounded set A, and hence ψ is one-to-one.

ψ is continuous by Lemma A.2. If $\psi(x_n) \to \psi(x)$, then $\{x_n\}$ should be bounded, otherwise, $Q_x(A) = \lim_{n\to\infty} Q(A - x_n) = 0$ for all bounded set A, a contradiction. As ψ is one-to-one, all the subsequential limits should be equal to x. Hence $x_n \to x$. It is easy to show that $\mathcal{M}_Q$ is closed in $\mathcal{P}$.

Lemma A.4 *Let $(\Omega, \mathcal{E})$ be a measurable space and $\chi: \Omega \to \mathcal{P}$ be a map. Then χ is measurable iff for all A in B^k, the map χ_A defined by*

$$\chi_A(w) = \chi(w)(A)$$

is measurable.

Proof. Since $Q \mapsto Q(A)$ from $\mathcal{P}$ to $\mathbb{R}$ is continuous only if part is trivial.

For if part, let $\mathcal{F}$ be a countable field generating B^k. Then for any $Q, Q_0 \in \mathcal{P}$

$$\|Q - Q_0\| = \sup\{|Q(A) - Q_0(A)| : A \in \mathcal{F}\}$$

by a well known fact in measure theory, (see Halmos (1974, Theorem 13.D) for example).

Now $\{w: \|x(w) - Q_0\| \le x\} = \bigcap_{A \in \mathcal{F}} \{w: |x_A(w) - Q_0(A)| \le x\} \in \mathcal{E}$ by hypothesis.

References

Basu, A. P., Ghosh, J. K. and Joshi, S. N. (1988). On estimating change point in a failure rate. *Statistical Decision Theory and Related Topics IV* **2**, 239–252 (S. S. Gupta and J. O. Berger, Eds.). Springer-Verlag, New York.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis.* 2nd. ed. Springer-Verlag. New York.

Bernstein, S. (1917). *Theory of Probability* (in Russian).

Bickel, P. J. and Yahav, J. (1969). Some contributions to the asymptotic theory of Bayes solutions. *Z. Warsch. verw. Gabiete* **11**, 257–275.

Chao, M. T. (1970). The asymptotic behaviour of Bayes estimators. *Ann. Math. Statist.* **41**, 601–609.

Chernoff, H. and Rubin, H. (1956). The estimation of the location of a discontinuity density, *Proc. of the 3rd Berkeley Symposium on Mathematical Statistics and Probability* **1**, 19–37, University of California Press.

Chow and Teicher, H. (1988). *Probability Theory.* Springer-Verlag, New York.

Clarke, B. and Barron, A. R. (1990). Entropy risk and Bayesian central limit theorem (Preprint).

Dawid, A. P. (1970). On the limiting normality of posterior distributions. *Proc. Camb. Phil. Soc. B* **67**, 625–633.

Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *Ann. Statist.* **14**, 1–25.

Dunford, N. and Schwartz, J. T. (1957). *Linear Operators, Part I: General Theory.* Interscience, New York.

Ghosh, J. K., Sinha, B. K., and Joshi, S. N. (1982). Expansions for posterior probability and integrated Bayes risk. In *Statistical Decision Theory and Related Topics III* (S. S. Gupta and J. O. Berger eds.). Academic Press, New York, 403–456.

Ghosh, J. K., Joshi, S. N. and Mukhopadhyay, C. (1992a). A Bayesian Study of the Reliability Change Point Problem, under preparation.

Ghosh, J. K., Joshi, S. N. and Mukhopadhyay, C. (1992b). Bayesian Asymptotics in Reliability Change Point Problem, under preparation.

Halmos, P. (1974). *Measure Theory.* Springer-Verlag.

Heyde, C. C. and Johnstone, I. M. (1979). On asymptotic posterior normality for stochastic processes, *J.R.S.S., Series B* **41**, 184–189.

Ibragimov, I. A. and Hasminskii, R. Z. (1981). *Statistical Estimation: Asymptotic Theory.* Springer-Verlag, New York.

Johnson, R. A. (1970). Asymptotic expansions associated with posterior distributions, *Ann. Math. Statist.* **41**, 851–864.

Kallianpur, G. Borwanker, J. D. and Prakasa Rao, B. L. S. (1971). The Bernstein-von Mises theorem for stochastic processes. *Ann. Math. Statist.* **42**, 1241–1253.

Laplace, P. S. (1774). Memoire sur la probabilite des causes par les evenements. Memoires de mathematique et de physique presentes a l'academie 'rougate des sciences, and dus dans ses assembies, **6** 621–656. (Translated in *Statist. Sci.* **1**, 359–378.)

Le Cam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *Univ. Calif. Publ. Statist.* **1**, 277–330.

Le Cam, L. (1958). Les proprietes asymptotiques aes solutions de Bayes. *Publ. Inst. Statist. Univ. Paris* **7**, 17–35.

Le Cam, L. (1970). Remarks on the Bernstein-von Mises theorem. *Preprint.*

Lindley, D. V. (1961). The inference of prior probability distributions in statistic inference and decisions, *Proc. 4th Berkeley Symp.* **1**, 453–468.

Samanta, T. (1988). Some contributions to the asymptotic theory of estimation in non-regular case. *Ph.D. Thesis.* Indian Statistical Institute.

Smith, R. L. (1985). Maximum likelihood estimation in a class of non-regular cases. *Biometrika* **72**, 67–90.

von Mises, R. (1931). Wahrscheinlichkeitsrechnung. Springer, Berlin.

Walker, A. M. (1969). On the asymptotic behaviour of posterior distributions. *J. Roy. Statist. Soc., Ser. B* **31**, 80–88.

Wong, L. and Wolfowitz, J. (1974). *Maximum Probability Estimators and Related Topics.* Springer-Verlag. Berlin.

This article was processed using the LaTeX macro package with LLNCS style