Tatra Mt. Math. Publ. **17** (1999), 27–35

Tatra Mountains
Mathematical Publications

THE HÁJEK CONVOLUTION THEOREM AND PITMAN CLOSENESS FOR REGULAR BAYES ESTIMATORS

Pranab Kumar Sen

ABSTRACT. For possibly vector valued parameters, the Hájek (1970) convolution theorem provides asymptotic optimality properties of regular estimators under quite general regularity assumptions, and this basic result has been extended to more diverse setups during the past three decades. Allowing shrinking prior distributions, regular Bayes estimators are considered, and an analogous convolution theorem is incorporated in the study of (generalized) Pitman closeness of regular Bayes estimators. Related isomorphism results for asymptotically normal family of estimators, arising in robust statistical inference, are also appraised in the current context.

1. Introduction and preliminaries

Consider a family of probability measures $\{P_{\boldsymbol{\theta},n}\boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ on some measures space $\{(\mathcal{X}_n, \mathcal{A}_n) \colon n \geq 1\}$, where $\boldsymbol{\theta}$ is a finite-dimensional parameter lying in the interior of the parameter space $\boldsymbol{\Theta} \subseteq \mathbb{R}^k$, for some $k \geq 1$. Then following Le Cam (1960), we consider a bounded set $\mathcal{B}$ and a sequence $\{\delta_n\}$ of real numbers such that

$$\{P_{\boldsymbol{\theta}+\delta_n \mathbf{t}_n, n}\} \text{ is contiguous to } \{P_{\boldsymbol{\theta},n}\} \text{ for all } \mathbf{t}_n \in \mathcal{B}\,. \tag{1.1}$$

We denote the *log-likelihood ratio* statistic by

$$\Lambda(\boldsymbol{\theta} + \delta_n \mathbf{t}_n \colon \boldsymbol{\theta}) = \log\{dP_{\boldsymbol{\theta}+\delta_n \mathbf{t}_n, n}/dP_{\boldsymbol{\theta},n}\}\,. \tag{1.2}$$

Then the family $\mathcal{P}_n = \{P_{\gamma,n} \colon \gamma \in \boldsymbol{\Theta}\}$ is termed LAQ (*locally asymptotically quadratic*) at $\boldsymbol{\theta}(\in \boldsymbol{\Theta})$ if there exist a vector $\mathbf{S}_n$ and an almost surely (a.s.) nonnegative definite (nnd) matrix $\mathbf{K}_n$, both possibly stochastic, such that for every $\mathbf{t}_n \in \mathcal{B}$,

$$\Lambda(\boldsymbol{\theta} + \delta_n \mathbf{t}_n; \boldsymbol{\theta}) - \mathbf{t}_n' \mathbf{S}_n + \frac{1}{2} \mathbf{t}_n' \mathbf{K}_n \mathbf{t}_n \to 0\,, \tag{1.3}$$

2000 Mathematics Subject Classification: 62C12, 62J07.
Keywords: asymptotics, BAN estimators, FOADR, LAN, robust estimators, Stein-rule estimators, Zellner prior.

in $P_{\boldsymbol{\theta},n}$-probability. Whenever $\mathbf{K}_n$, though stochastic, has a limiting distribution that does not depend on $\mathbf{t}_n \in \mathcal{B}$, we have a LAMN (*locally asymptotically mixed normal*) family. Further, whenever $\mathbf{K}_n$ can be taken as a nonstochastic nnd matrix $\mathbf{K}$ (i.e., $\mathbf{K}_n$ has a limiting law degenerate at $\mathbf{K}$), we have the well known LAN (*locally asymptotically normal*) family. For simplicity of presentation, we confine ourselves to models that admit $\sqrt{n}$-*consistent* estimators, and for which in (1.1), $\delta_n \equiv n^{-1/2}$. In this case, we take $\boldsymbol{\theta}_n = \boldsymbol{\theta} + n^{-1/2}\mathbf{h}$, $\mathbf{h} \in \mathbb{R}^k$, and define the (Rao) *scores statistics* by

$$\mathbf{U}_n = \mathbf{U}_n(\boldsymbol{\theta}) = (\partial/\partial\mathbf{h})\Lambda(\boldsymbol{\theta} + n^{-1/2}\mathbf{h};\boldsymbol{\theta})|_{\mathbf{h}=\mathbf{0}}\,. \tag{1.4}$$

Also, we define the (Fisher) *information matrix* by

$$\mathbf{I}(\boldsymbol{\theta}) = \mathrm{E}_{\boldsymbol{\theta}}\big[\mathbf{U}_n(\boldsymbol{\theta})\big]\,\big[\mathbf{U}_n(\boldsymbol{\theta})\big]'\,. \tag{1.5}$$

Therefore, under the LAN setup, we have

$$\Lambda\big(\boldsymbol{\theta} + n^{-1/2}\mathbf{h};\boldsymbol{\theta}\big) = \mathbf{h}'\mathbf{U}_n(\boldsymbol{\theta}) - \frac{1}{2}\mathbf{h}'\mathbf{I}(\boldsymbol{\theta})\mathbf{h} + o_p(1)\,, \tag{1.6}$$

uniformly in $\mathbf{h} \in \mathcal{B}$, where under $\{P_{\boldsymbol{\theta},n}\}$,

$$\mathbf{U}_n(\boldsymbol{\theta}) \xrightarrow{\mathcal{D}} \mathcal{N}_k\big(\mathbf{0},\mathbf{I}(\boldsymbol{\theta})\big)\,. \tag{1.7}$$

Let $\{\mathbf{T}_n\}$ be a sequence of estimators of $\boldsymbol{\theta}$, and let

$$H_n(\,\cdot\,;\boldsymbol{\theta}) = \text{ probability law for } n^{1/2}(\mathbf{T}_n - \boldsymbol{\theta}) \text{ under } P_{\boldsymbol{\theta},n}\,. \tag{1.8}$$

Then $\{\mathbf{T}_n\}$ is said to be (*essentially*) *regular* at $\boldsymbol{\theta}$ if

$$H_n(\,\cdot\,;\boldsymbol{\theta}_n) \Rightarrow H(\,\cdot\,;\boldsymbol{\theta})\,, \quad \forall\boldsymbol{\theta}_n = \boldsymbol{\theta} + n^{-1/2}\mathbf{h}\,, \quad \mathbf{h} \in (\mathbf{D} \subset)\mathbb{R}^k\,, \tag{1.9}$$

where $\mathbf{D}$ is unique. Then the celebrated Hájek–Inagaki–Le Cam Convolution Theorem (Hájek 1970, 1972; Inagaki 1970, 1973, Le Cam 1972, 1979), for the simpler setup outlined above, can be presented as follows.

Suppose that $\{P_{\boldsymbol{\theta},n}\}$ *is* LAN *at* $\boldsymbol{\theta}_o(\in \boldsymbol{\Theta})$, *and* $\{\mathbf{T}_n\}$ *is a sequence of regular estimators at* $\boldsymbol{\theta}_o$. *Then, under* $P_{\boldsymbol{\theta}_o,n}$,

$$n^{1/2}(\mathbf{T}_n - \boldsymbol{\theta}_o) = \big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1}\mathbf{U}_n(\boldsymbol{\theta}_o) + \boldsymbol{\xi}_n + \mathbf{o}_p(1)\,, \tag{1.10}$$

where $\boldsymbol{\xi}_n$ *is (asymptotically) independent of* $\mathbf{U}_n(\boldsymbol{\theta}_o)$. *Thus,* $H(\cdot;\boldsymbol{\theta}_o)$, *the limiting probability law of* $n^{1/2}(\mathbf{T}_n - \boldsymbol{\theta}_o)$, *under* $P_{\boldsymbol{\theta}_o,n}$, *is given by*

$$H(\,\cdot\,;\boldsymbol{\theta}_o) = \Phi\Big(\mathbf{0}, \Big(\mathbf{I}(\boldsymbol{\theta}_o)\Big)^{-1}\Big) * G(\,\cdot\,;\boldsymbol{\theta}_o)\,, \tag{1.11}$$

where $G(\cdot;\boldsymbol{\theta}_o)$ *stands for the limit law for* $\boldsymbol{\xi}_n$ *at* $\boldsymbol{\theta}_o$, *and* $*$ *stands for the convolution.*

Next we recapitulate briefly the notion of *best asymptotically normal* (BAN) estimators that are *first-order efficient* (FOE), and note that under the aforesaid LAN condition, at $\boldsymbol{\theta}_o$, if $\{\mathbf{T}_n^*\}$ is BAN, then

$$n^{1/2}(\mathbf{T}_n^* - \boldsymbol{\theta}_o) = \big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1}\mathbf{U}_n + \mathbf{o}_p(1). \tag{1.12}$$

Moreover, for an estimator $\mathbf{T}_n$, not necessarily BAN (at $\boldsymbol{\theta}_o$), we say that a *first-order asymptotic distributional representation* (FOADR) holds, if there exists a suitable *score function* $\boldsymbol{\psi}(\cdot)$, such that at $\boldsymbol{\theta}_o$,

$$n^{1/2}(\mathbf{T}_n - \boldsymbol{\theta}_o) = n^{-1/2}\sum_{i=1}^{n} \boldsymbol{\psi}(X_i, \boldsymbol{\theta}_o) + \mathbf{o}_p(1), \tag{1.13}$$

where $\mathrm{E}_{\boldsymbol{\theta}_o}\boldsymbol{\psi}(X, \boldsymbol{\theta}_o) = \mathbf{0}$ and

$$\mathbf{V}_{\boldsymbol{\theta}_o}\big(\boldsymbol{\psi}(X, \boldsymbol{\theta}_o)\big) = \big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1} + \boldsymbol{\Gamma}, \tag{1.14}$$

where $\boldsymbol{\Gamma}$ is positive semidefinite (psd), and is a null matrix when $\mathbf{T}_n$ is BAN. Actually, $\boldsymbol{\Gamma}$ is the dispersion matrix for the distribution $G(\cdot, \boldsymbol{\theta}_o)$ of $\boldsymbol{\xi}_n$, in the convolution theorem (see (1.10)) when $\mathbf{T}_n$ is regular and admits a FOADR. Under additional regularity conditions, studied by Hájek (1972) and others, that validate (1.10) with the remainder term $\mathbf{o}_p(1)$ being $\mathbf{o}(1)$ in quadratic mean as well, the above representation also asserts the first-order *asymptotic sufficiency* of BAN estimators.

In the next section, we introduce Zellner's (1986) g-priors, and consider a class of Bayes estimators that are regular in the same sense as in the Hájek–Inagaki–Le Cam theorem, and this use the convolution theorem to characterize their asymptotic optimality properties. Section 3 deals with parallel results under the generalized Pitman closeness criterion. The concluding section deals with robust estimators and their Bayes versions.

2. Convolution theorem for regular Bayes estimators

Let us assume that the FOADR results and the convolution theorem remain valid for all $\boldsymbol{\theta}$ in a neighbourhood of the true parameter point $\boldsymbol{\theta}_o$. As in the work of Hájek (1970) and others, certain uniformity conditions are needed in this context, and they would be taken for granted here. Secondly, it is known that Bayes estimators and *maximum likelihood estimators* (MLE) are asymptotically equivalent (in a suitable norm) upto the order $o(n^{-1/2})$; typically, they differ by $O_p(n^{-1})$. The situation is different if one uses a shrinking prior on $\boldsymbol{\theta}$ that degenerates at $\boldsymbol{\theta}_o$ as n increases; this, of course, has a natural statistical interpretation that as the sample size becomes large, one should put less emphasis on the priors and more on the sample evidence. To facilitate the interplay of asymptotics and Bayes estimation, following Zellner (1986), we introduce a class of priors that is known as g-prior which in the context of a linear model relates to a Gaussian law whose dispersion matrix is tied down to the underlying incidence (design) matrix. In that way, it reflects the impact of the sample size as well, and results in a suitably normed shrinking prior in an asymptotic setup. Taking clues from the Zellner approach, we formulate the following prior law on $\boldsymbol{\theta}$.

$$(\boldsymbol{\theta} - \boldsymbol{\theta}_o) \stackrel{\mathcal{D}}{=} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_o)\,, \tag{2.1}$$

where we choose

$$\boldsymbol{\Sigma}_o = c_n\big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1}, \qquad c_n > 0\,,\ \lim_{n\to\infty} c_n = 0\,, \tag{2.2}$$

and we shall tune $\{c_n^* = nc_n\}$ in the light of a nondegenerate posterior distribution. Since the prior on $\boldsymbol{\theta}$ is Gaussian, using the FOADR for $(\mathbf{T}_n^* - \boldsymbol{\theta})$, along with the continuity of $\mathbf{I}(\boldsymbol{\theta})$ at $\boldsymbol{\theta}_o$, we obtain by convolution the following marginal law for $\mathbf{T}_n^* - \boldsymbol{\theta}_o$:

$$n^{1/2}(\mathbf{T}_n^* - \boldsymbol{\theta}_o) \xrightarrow{\mathcal{D}} \mathcal{N}\Big(\mathbf{0}\,, \big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1}(1 + c_n^*)\Big)\,. \tag{2.3}$$

The last result implies that the conditional law for $\boldsymbol{\theta}$, given $\mathbf{T}_n^*$ and $\boldsymbol{\theta}_o$, is asymptotically normal with the asymptotic mean

$$\boldsymbol{\theta}_o + \frac{c_n^*}{1 + c_n^*}(\mathbf{T}_n^* - \boldsymbol{\theta}_o) = \boldsymbol{\theta}_o + (1 - B_n)(\mathbf{T}_n^* - \boldsymbol{\theta}_o)\,, \tag{2.4}$$

where the (Bayes) shrinkage factor B_n is defined as

$$B_n = (1 + c_n^*)^{-1}\,, \quad \text{and} \quad B_n \in (0, 1)\,. \tag{2.5}$$

Note further that the asymptotic dispersion matrix of this posterior distribution is

$$n^{-1}(1 - B_n)\big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1}\,. \tag{2.6}$$

Therefore, in an asymptotic sense, the Bayes estimator of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_B^* = \boldsymbol{\theta}_o + (1 - B_n)\,(\mathbf{T}_n^* - \boldsymbol{\theta}_o)\,. \tag{2.7}$$

This resembles the familiar Stein-rule estimator, so that B_n can be interpreted as the shrinkage factor. It also reveals that if c_n^* converges to 0, i.e., $c_n = o(n^{-1})$, then $B_n \to 1$, so that $\hat{\boldsymbol{\theta}}_B^* \to \boldsymbol{\theta}_o$ with a dispersion matrix $\mathbf{o}(n^{-1})$. Similarly, if $c_n^* \to \infty$ as $n \to \infty$, i.e., $c_n \gg n$, then $B_n \to 0$, so that the Bayes estimator becomes asymptotically equivalent to $\mathbf{T}_n^*$. Hence, a limiting nondegenerate case arises only when $c_n^* \to c(> 0)$ as $n \to \infty$, i.e., the prior on $\boldsymbol{\theta}$ has the same convergence rate $(O(n^{-1/2}))$ as in the FOADR for $\mathbf{T}_n^*$. In the linear model, it is generally assumed that for the design matrix $\mathbf{T}_n^*$, as n increases, $n^{-1}\mathbf{T}_n^*$ converges to a p.d. limit, say $\mathbf{T}^*$, and that automatically links the Zellner g-prior to our $c_n^* = c$, for some $c > 0$. In this sense, we regard the chosen prior as g-prior. We shall find it more convenient to express $\hat{\boldsymbol{\theta}}_B^*$ as

$$\hat{\boldsymbol{\theta}}_B^* - \boldsymbol{\theta} = B_n(\boldsymbol{\theta}_o - \boldsymbol{\theta}) + (1 - B_n)(\mathbf{T}_n^* - \boldsymbol{\theta}), \tag{2.8}$$

where B_n, defined by (2.5), converges to $B = (1+c)^{-1}$. It is then easy to verify that

$$\boldsymbol{\nu}_B^* = \lim_{n\to\infty} n\mathrm{E}\big\{(\hat{\boldsymbol{\theta}}_B^* - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{nB}^* - \boldsymbol{\theta})'\big\} = (1 - B)\big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1}. \tag{2.9}$$

Let us now consider another regular estimator of $\boldsymbol{\theta}$, denoted by $\mathbf{T}_n$, and assume that it has the FOADR in (1.13), so that (1.14) is taken for granted. We convolute the g-prior on $\boldsymbol{\theta}$, and by similar arguments, we obtain that the conditional law for $\boldsymbol{\theta}$, given $\mathbf{T}_n, \boldsymbol{\theta}_o$, is asymptotically normal with mean

$$\boldsymbol{\theta}_o + \frac{c_n^*}{1 + c_n^*}\mathbf{I}^*(\mathbf{T}_n - \boldsymbol{\theta}_o) = \boldsymbol{\theta}_o + (1 - B_n)\mathbf{I}^*(\mathbf{T}_n - \boldsymbol{\theta}_o), \tag{2.10}$$

where B_n is defined as in (2.5), and

$$\mathbf{I}^* = \big(\mathbf{I} + B_n\boldsymbol{\Gamma}\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1}; \tag{2.11}$$

note that $\mathbf{I} - \mathbf{I}^*$ is psd, and it is a null matrix only when $\boldsymbol{\Gamma} = \mathbf{0}$ (or $B_n = 0$). We note that there might be some loss of information due to using a possibly less informative estimator $\mathbf{T}_n$ instead of $\mathbf{T}_n^*$, and that is reflected in the posterior mean computed above. In a similar fashion, it can be shown that the dispersion matrix of this asymptotic posterior distribution is

$$c_n\big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1}\big[\mathbf{I} - (1 - B_n)\,\mathbf{I}^*\big]. \tag{2.12}$$

The Bayes estimator obtained from the above posterior distribution may therefore be put as

$$\hat{\boldsymbol{\theta}}_B = \boldsymbol{\theta}_o + (1 - B_n)\,\mathbf{I}^*(\mathbf{T}_n - \boldsymbol{\theta}_o). \tag{2.13}$$

We rewrite this estimator as

$$\hat{\boldsymbol{\theta}}_B = \big(\mathbf{I} - (1 - B_n)\mathbf{I}^*\big)\boldsymbol{\theta}_o + (1 - B_n)\,\mathbf{I}^*\mathbf{T}_n. \tag{2.14}$$

We now incorporate the representation in (1.10) and (1.12) in the last expression, and obtain that

$$\begin{aligned}\hat{\boldsymbol{\theta}}_B - \boldsymbol{\theta}_o &= (1 - B_n)(\mathbf{T}_n - \boldsymbol{\theta}_o)\\ &= \mathbf{I}^*(\hat{\boldsymbol{\theta}}_B^* - \boldsymbol{\theta}_o) + (1 - B_n)n^{-1/2}\boldsymbol{\xi}_n + \mathbf{o}_p(n^{-1/2}),\end{aligned} \tag{2.15}$$

where $\boldsymbol{\xi}_n$ and $\mathbf{I}^*$ are defined as in (1.10) and (2.9), respectively. Moreover, using the convolution theorem, we conclude that $\boldsymbol{\xi}_n$ is (asymptotically) independent of $\mathbf{T}_n^*$ (and hence of $\hat{\boldsymbol{\theta}}_B^*$ as well). Therefore, neglecting terms that are $o_p(n^{-1/2})$, we rewrite $\hat{\boldsymbol{\theta}}_B$ as

$$\begin{aligned}\hat{\boldsymbol{\theta}}_B - \boldsymbol{\theta} &= \big[\mathbf{I} - (1-B_n)\mathbf{I}^*\big](\boldsymbol{\theta}_o - \boldsymbol{\theta}) + (1-B_n)\mathbf{I}^*(\mathbf{T}_n - \boldsymbol{\theta}) \\ &= \big[\mathbf{I} - (1-B_n)\mathbf{I}^*\big](\boldsymbol{\theta}_o - \boldsymbol{\theta}) + (1-B_n)\mathbf{I}^*(\mathbf{T}_n^* - \boldsymbol{\theta}) + (1-B_n)\mathbf{I}^* n^{-1/2}\boldsymbol{\xi}_n \\ &= [\mathbf{I} - \mathbf{I}^*](\boldsymbol{\theta}_o - \boldsymbol{\theta}) + \mathbf{I}^*(\hat{\boldsymbol{\theta}}_B^* - \boldsymbol{\theta}) + (1-B_n)\mathbf{I}^* n^{-1/2}\boldsymbol{\xi}_n\,.\end{aligned} \tag{2.16}$$

Let us denote the asymptotic dispersion matrices of $\sqrt{n}(\hat{\boldsymbol{\theta}}_B - \boldsymbol{\theta})$ by $\boldsymbol{\nu}_B$. Then from the above representations, namely (2.9) and (2.16), we obtain that

$$\begin{aligned}\boldsymbol{\nu}_B &= (1-B)\big(\mathbf{I}(\boldsymbol{\theta}_o)\big)^{-1} + \frac{(1-B)^2}{B}\big[(\mathbf{I} - \mathbf{I}^*)\mathbf{I}(\boldsymbol{\theta}_o)^{-1}(\mathbf{I} - \mathbf{I}^*)'\big] \\ &= \boldsymbol{\nu}_B^* + B(1-B)^2\big[\boldsymbol{\Gamma}\mathbf{I}(\boldsymbol{\theta}_o)\mathbf{I}^*\mathbf{I}(\boldsymbol{\theta}_o)^{-1}\mathbf{I}^*\mathbf{I}(\boldsymbol{\theta}_o)\boldsymbol{\Gamma}\big]\,,\end{aligned} \tag{2.17}$$

where by (2.5), we note that $\mathbf{I} - \mathbf{I}^*$ is psd; it is a null matrix when either $B_n = 0$ or $\boldsymbol{\Gamma} = \mathbf{0}$. Finally, we note that if $\mathbf{A}$ and $\mathbf{B}$ are psd (symmetric) matrices, then $(\mathbf{I} + \mathbf{A})\mathbf{B}(\mathbf{I} + \mathbf{A}) - \mathbf{B}$ is also psd. As such, we claim that

$$\boldsymbol{\nu}_B - \boldsymbol{\nu}_B^* = \text{ positive semi-definite}, \tag{2.18}$$

for all regular Bayes estimators, and the right hand side reduces to a null matrix when $\boldsymbol{\Gamma} = \mathbf{0}$ or $B_n = 0$, i.e., either $\mathbf{T}_n$ is itself FOE or the prior on $\boldsymbol{\theta}$ is degenerate in the sense that $\lim_{n\to\infty} c_n^* = \infty$. This extends the Hájek convolution theorem to regular Bayes estimators as well.

3. Generalized Pitman closeness of Bayes estimators

Let us appraise the situation in the light of the generalized Pitman closeness criterion (GPCC). For two rival estimators, say, $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$, of a common parameter $\boldsymbol{\theta}$, and a loss function $L(\mathbf{a}, \mathbf{b})$, the GPCC of $\hat{\boldsymbol{\theta}}_1$ with respect to $\hat{\boldsymbol{\theta}}_2$ is defined as

$$\begin{aligned}P(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2 | \boldsymbol{\theta}) = P\big\{L(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}) < L(\hat{\boldsymbol{\theta}}_2, \boldsymbol{\theta}) | \boldsymbol{\theta}\big\} \\ + \frac{1}{2} P\big\{L(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}) = L(\hat{\boldsymbol{\theta}}_2, \boldsymbol{\theta}) | \boldsymbol{\theta}\big\}\,.\end{aligned} \tag{3.1}$$

Then $\hat{\boldsymbol{\theta}}_1$ is regarded as closer to $\boldsymbol{\theta}$ than $\hat{\boldsymbol{\theta}}_2$, if the right hand side of (3.1) is $\geq 1/2$ for all $\boldsymbol{\theta}$, with the strict inequality sign holding for some $\boldsymbol{\theta}$. We refer to Keating, Mason and Sen (1993) for some discussion of the GPCC.

Ghosh and Sen (1991) extended the GPCC concept to a Bayesian setup and introduced the idea of the *Posterior Pitman closeness* (PPC) measure. Sen (1994) established an isomorphism of the GPCC and conventional asymptotic minimaxity under a matrix valued dominance that will be validated in the current context.

For two estimators $\mathbf{T}_{n1}, \mathbf{T}_{n2}$ with (asymptotic) dispersion matrices $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2$ respectively, if $\boldsymbol{\nu}_2 = \boldsymbol{\nu}_1 + \boldsymbol{\Gamma}$ for some psd $\boldsymbol{\Gamma}$, then $\mathbf{T}_{n1}$ is said to dominate $\mathbf{T}_{n2}$, in the matrix sense; we denote this by

$$\mathbf{T}_{n1} \succ_{\text{AMSEM}} \mathbf{T}_{n2} \,. \tag{3.2}$$

Let us consider a quadratic loss $L(\mathbf{a}, \mathbf{b}) = ||\mathbf{a} - \mathbf{b}||^2_{\mathbf{Q}}$ for a p.d. $\mathbf{Q}$. Suppose that the GPCC dominance of $\mathbf{T}_{n1}$ over $\mathbf{T}_{n2}$ holds for all $\mathbf{Q}$ belonging to a general class $\mathcal{Q}$. Then we say that $\mathbf{T}_{n1}$ dominates $\mathbf{T}_{n2}$ in the GPCC, uniformly in $\mathbf{Q} \in \mathcal{Q}$, and we denote this by

$$\mathbf{T}_{n1} \succ_{\text{APC},\mathcal{Q}} \mathbf{T}_{n2} \,. \tag{3.3}$$

For estimators admitting a FOADR, Sen (1994) established that

$$\{\succ_{\text{AMSEM}}\} \equiv \{\succ_{\text{APC},\mathcal{Q}}\} \,. \tag{3.4}$$

An important feature is that the APC ordering extends over the entire class $\mathcal{Q}$ of pd matrices that enter into the definition of the quadratic loss.

The AMSEM ordering of the regular Bayes estimators follows from the convolution theorem in the preceding section. As such, because of the above asymptotic equivalence result, we obtain the following GPCC result:

$$\hat{\boldsymbol{\theta}}^*_B \succ_{\text{APC},\mathcal{Q}} \hat{\boldsymbol{\theta}}_B \,, \qquad \text{for all regular Bayes estimators} \,. \tag{3.5}$$

Let us also look into the posterior Pitman closeness aspects of regular Bayes estimators. Let $\Pi(\boldsymbol{\theta})$ be a prior distribution defined on $\boldsymbol{\Theta}$, and let $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2$ be two Bayes estimators of $\boldsymbol{\theta}$ under the prior $\Pi(\boldsymbol{\theta})$. Then, as in Ghosh and Sen (1991), we define

$$\text{PPC}_\Pi(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2 | \mathbf{Y}) = P_\Pi\{L(\boldsymbol{\delta}_1, \boldsymbol{\theta}) < L(\boldsymbol{\delta}_2, \boldsymbol{\theta})\} + \frac{1}{2} P_\Pi\{L(\boldsymbol{\delta}_1, \boldsymbol{\theta}) = L(\boldsymbol{\delta}_2, \boldsymbol{\theta})\} \,. \tag{3.6}$$

Thus, $\boldsymbol{\delta}_1$ is said to be posterior Pitman closer to $\boldsymbol{\theta}$ than $\boldsymbol{\delta}_2$, under the prior $\Pi(\boldsymbol{\theta})$, provided

$$\text{PPC}_\Pi(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2 | \mathbf{Y}) \geq \frac{1}{2}, \qquad \forall\, \mathbf{Y} \text{ a.e.}, \tag{3.7}$$

with strict inequality holding for some $\mathbf{Y}$.

We confine ourselves to the Zellner g-prior as introduced in the earlier sections. Then noting that the posterior distributions are asymptotically multinormal, as obtained in Section 2, we conclude that under the adopted g-prior of $\boldsymbol{\theta}$,

$$\hat{\boldsymbol{\theta}}^*_B \succ_{\text{PPC}} \hat{\boldsymbol{\theta}}_B, \text{ for all regular Bayes estimators} \,. \tag{3.8}$$

These results put the Bayes estimator $\hat{\boldsymbol{\theta}}^*_B$ in a commandable position within the class of regular Bayes estimators as defined in Section 2.

4. Concluding remarks

Looking back at (1.10)–(1.12), we may observe that the class of regular Bayes estimators considered here is characterized by the Zellner g-prior (on $\boldsymbol{\theta}$) and regular estimators that admit FOADR (as in (1.13)). In this way, it is possible to include a large class of estimators which need not be of the maximum likelihood type. For example, if $\boldsymbol{\theta}$ is the location parameter of a (possibly multivariate) distribution, then we can consider various *robust estimators*; important subfamilies are the classes of the so called L-, M-, and *R-estimators* which are based on order statistics, robust statistics and rank statistics, respectively. There are also other estimators based on suitable statistical functionals, such as the U-statistics, von Mises' functions, and related Hadamard differentiable functionals. These estimators all admit FOADR, and often are amenable to higher order asymptotics too. Moreover, within each subclass, there exists a particular member that is BAN in the sense of (1.12). We refer to Jurečková and Sen (1996) for a unified account of such robust estimators of location, scale and regression parameters, and the related FOADR results. Therefore, we can construct regular Bayes estimators based on the Zellner g-prior and such robust estimators which admit FOADR. In this way, we have a class of Bayes estimators that include the conventional Bayes estimator as a notable member, and our asymptotic optimality properties relate to them.

It is also possible to consider some related empirical Bayes estimators, and to study the asymptotic optimality properties in the light of the Hájek convolution theorem. However, in general, an empirical Bayes estimator may not be (asymptotically) normal, though it could have some FOADR (without having a linear statistic that pertains to normality). This calls for some alternative formulations, and we report these findings in a future communication (Sen, 1999).

REFERENCES

[1] GHOSH, M.—SEN, P. K.: *Bayesian Pitman closeness*, Comm. Statist. Theory Methods **A20** (1991), 3659–3695.

[2] HÁJEK, J.: *A characterization of limiting distributions of regular estimates*, Z. Wahrsch. verw. Geb. **14** (1970), 323–330.

[3] HÁJEK, J.: *Local asymptotic minimax and admissibility in estimation*, in: Proc. 6th Berkeley Symp. Math. Statist. Prob., Vol. 1, 1972, pp. 175–194.

[4] INAGAKI, N.: *On the limiting distribution of a sequence of estimators with uniform property*, Ann. Inst. Statist. Math. **22** (1970), 1–13.

[5] INAGAKI, N.: *Asymptotic relations between the likelihood estimating function and the maximum likelihood estimator*, Ann. Inst. Statist. Math. **25** (1973), 1–26.

[6] JUREČKOVÁ, J.—SEN, P. K.: *Robust Statistical Inferences: Asymptotics and Interrelations*, John Wiley, New York, 1996.

[7] KEATING, J. P.—MASON, R. L.—SEN, P. K.: *Pitman's Measure of Closeness: A Comparison of Statistical Estimators*, SIAM, Philadelphia, 1993.

[8] Le CAM, L.: *Locally asymptotically normal families of distributions*, Univ. California Pub. Statist. **3** (1960), 37–98.

[9] Le CAM, L.: *Limits of experiments and a theorem of J. Hájek*, in: Proc. 6th. Berkeley Symp. Math. Statist. Prob., Vol. 1, 1972, pp. 245–261.

[10] Le CAM, L.: *On a theorem of J. Hájek*, in: Contributions to Statistics: J. Hájek Mem. Vol. (J. Jurečková, ed.), Academia Prague, 1979, pp. 119–137.

[11] SEN, P. K.: *Sequential Nonparametrics: Invariance Principles and Statistical Inference*, John Wiley, New York, 1981.

[12] SEN, P. K.: *Isomorphism of quadratic norm and PC ordering of estimators admitting first order AN representation*, Sankhyā **A56** (1994), 465–475.

[13] SEN, P. K.: *Semiparametric empirical Bayes estimation in linear models*, in: Empirical Bayes and Likelihood (E. Ahmad and N. Reid, eds.), Springer, Montreal, 1999.

[14] ZELLNER, A.: *On assessing prior distributions and Bayesian regression analysis with g-prior distributions*, in: Bayesian Inference and Decision Techniques (P. K. Goel and A. Zellner, eds.), North Holland, Amsterdam, 1986, pp. 233–243.

Received April 29, 1998

Departments of Biostatistics and Statistics
University of North Carolina
Chapel Hill, NC 27599–7400
U. S. A

E-mail: pksen@bios.unc.edu