ELSEVIER

Journal of Statistical Planning and Inference 46 (1995) 293–309

journal of statistical planning and inference

Empirical Bayes testing procedures in some nonexponential families using asymmetric Linex loss function

Su-Yun Huang

Institute of Statistical Science, Academia Sinica, Nankang, Taipei 11529, Taiwan, ROC

Received 14 November 1992

Abstract

This paper deals with the empirical Bayes testing problem in some nonexponential families using asymmetric Linex error loss. The asymptotic optimality of the proposed empirical Bayes testing procedures is studied. For a certain class of discrete priors, the convergence rate is of order $O(n^{-1}(\log n)^{1+\varepsilon})$ for arbitrarily small $\varepsilon > 0$. For a certain class of continuous priors such that f_G is decreasing, the convergence rate is of order $O(n^{-2/3+\delta^*})$ for arbitrarily small $\delta^* > 0$.

AMS Subject Classification: 62C12

Key words: Asymptotic optimality; Empirical Bayes testing procedures; Linex loss function

1. Introduction

Empirical Bayes methods have received considerable attention since Robbins (1956). There is vast literature on empirical Bayes methods dealing with exponential families. In recent years, there have been a growing interest in nonexponential families as well. (See Van Houwelingen, 1987; Nogami, 1988; Liang, 1990 for a uniform distribution $U(0, \theta)$, and Datta, 1991 for a more general setup. See Singh and Prasad, 1989 and Prasad and Singh, 1990 for a truncated exponential distribution, and Tiwari and Zalkikar, 1990; Liang, 1993, for a Pareto distribution.) In the above references, the authors have considered only using either squared error loss for estimation problems, or symmetric linear error loss for testing problems. Under these symmetry considerations, equal seriousness for overestimation and underestimation is assumed. However in some cases, symmetric loss function may be inappropriate and unrealistic. Varian (1975) gives such an example in a Bayesian approach to a real estate assessment. He has pointed out that an underestimation of a house's market value would

SSDI 0378-3758(94)00112-X

only result in the loss of tax revenues and an overestimation would lead both the house owner and the assessment office into lengthy and expensive court procedures. Varian has found the usual squared error loss inappropriate, and introduced an asymmetric loss function called Linex, which addresses different seriousness to overestimation and underestimation respectively. Later, Zellner (1986) employs the Linex loss in the Bayesian analysis of several central statistical estimation and prediction problems. Also Kuo and Dey (1990) consider estimation of a Poisson mean using the Linex loss. The model is useful in the software reliability assessment. Basu and Ebrahimi (1991) have used the Linex loss in a life time testing and reliability estimation. There are also other typical empirical Bayes problems including estimation or testing on defectives proportion, mean worker accident rate, bone loss rate for an osteoporosis high risk group, etc. (Berger, 1985). The Linex loss can be useful for these problems. We all know that the loss function plays an important role in Bayesian analysis, consequently it also severely influences the empirical approach to approximating the Bayes rule. However, there is lack of such studies on asymmetric loss in the empirical Bayes analysis. The author feels that this area deserves more attention.

In empirical Bayes analysis dealing with nonexponential families, there are two commonly seen models. One is, given θ, the random variables X has a p.d.f. of the form

$$f(x\,|\,\theta)=\frac{a(x)}{A(\theta)}I_{(0,\theta)}(x)$$

with $a(x)$ a known function and $A(\theta)$ so determined that $f(x\,|\,\theta)$ is a density. Examples are uniform distribution and truncated exponential. The other model is a conditional p.d.f. of the form

$$f(x\,|\,\theta)=\frac{a(x)}{A(\theta)}I_{(\theta,\infty)}(x).$$

Examples are translated exponential and Pareto distribution. Note that the second model has a much thicker right tail than the first one. In this paper, we will be dealing with the second model. The second model has greater difficulty in obtaining asymptotic optimality results and is more challenging than is the first model.

Suppose that, given θ, the random variable X has a distribution with p.d.f. of the form:

$$f(x\,|\,\theta)=\frac{a(x)}{A(\theta)}I_{(\theta,\infty)}(x), \tag{1.1}$$

where $a(x)>0$ is a known function and $A(\theta)=\int_\theta^\infty a(x)\,\mathrm{d}x<\infty$ for all $\theta>0$. Assume the parameter θ is a realization of a random variable Θ having a prior distribution G on $[0,\infty)$. Let θ_0 be a known positive constant. We are interested in testing H_0: $\theta\geqslant\theta_0$ against H_1: $\theta<\theta_0$. The following asymmetric Linex loss function is employed.

$$L_0(\theta) = c(\mathrm{e}^{b(\theta_0-\theta)} - b(\theta_0-\theta) - 1)I_{(\theta<\theta_0)}, \quad b \neq 0, \ c > 0,$$

$$L_1(\theta) = c(\mathrm{e}^{b(\theta_0-\theta)} - b(\theta_0-\theta) - 1)I_{(\theta \geqslant \theta_0)},$$

where $L_i(\theta)$ indicates the loss when the decision is in favor of H_i, while θ is the true state of the parameter. The constant c serves to scale the loss function. Since it does not affect the Bayes decision rule nor the empirical Bayes study, c is assumed to be 1. The constant b determines the shape of the loss function. Varian (1975) and Zellner (1986) have discussed the behavior of the loss function and their various applications. When $b > 0$, as $|\theta_0 - \theta| \to \infty$ the loss increases almost exponentially for wrongly deciding in favor of H_0, and almost linearly for wrongly deciding in favor of H_1. That is, overestimation is more serious than underestimation. When $b < 0$, the linear-exponential increases are interchanged. For small value of $|b|$, the loss function is close to the squared error loss.

In this paper, we propose certain empirical Bayes testing procedures and show their asymptotic optimality. We also study the convergence rates for certain classes of priors.

2. Bayes rule and some properties

Given θ, let X be a random variable having p.d.f. $f(x|\theta)$ of the form (1.1). Throughout this paper, we assume that $a(x)$ is positive, Lipschitz continuous and decreasing in x. A decision rule d is defined to be a mapping from the sample space of X into $[0,1]$ such that $d(x)$ is the probability of accepting H_0 given $X = x$. That is, $d(x) = P\{\text{accept } H_0 \,|\, X = x\}$. Let $L(\theta, d)$ denote the loss associated with the decision rule d. Then

$$L(\theta, d(x)) = (1 - d(x))\ell(\theta)I_{(\theta \geqslant \theta_0)} + d(x)\ell(\theta)I_{(\theta<\theta_0)}, \tag{2.1}$$

where $\ell(\theta) = \mathrm{e}^{b(\theta_0-\theta)} - b(\theta_0-\theta) - 1$, $b > 0$. Here we may assume $x \geqslant \theta_0$, otherwise the action to take is obvious to be in favor of H_1 with zero loss. Therefore, $x \geqslant \theta_0$ is assumed throughout this paper. The expected posterior loss of $d(x)$ is given by

$$\int_0^x L(\theta, d(x))\,\mathrm{d}G(\theta|x) = d(x)\varphi_G(x) + c(x), \tag{2.2}$$

where

$$\varphi_G(x) = \frac{a(x)}{f_G(x)}\left(\int_0^{\theta_0} \frac{\ell(\theta)}{A(\theta)}\,\mathrm{d}G(\theta) - \int_{\theta_0}^x \frac{\ell(\theta)}{A(\theta)}\,\mathrm{d}G(\theta)\right) \tag{2.3}$$

with $f_G(x) = \int_0^x f(x|\theta)\,\mathrm{d}G(\theta)$, the marginal p.d.f. of X, and where

$$c(x) = \int_{\theta_0}^x \ell(\theta)\,\mathrm{d}G(\theta|x).$$

Therefore the Bayes decision rule is

$$d_G(x) = \begin{cases} 1 & \text{if } \varphi_G(x) \leqslant 0, \\ 0 & \text{if } \varphi_G(x) > 0, \end{cases}$$

or equivalently

$$d_G(x) = \begin{cases} 1 & \text{if } H_G(x) \leqslant 0, \\ 0 & \text{if } H_G(x) > 0, \end{cases}$$

where $H_G(x) = \varphi_G(x) f_G(x)$.

There are some interesting properties in this Bayesian study, which will be needed later.

Lemma 1. *Let* $\psi_1(x) = \int_0^x (1/A(\theta))\,\mathrm{d}G(\theta)$. *Then*

$$f_G(x) = a(x)\psi_1(x).$$

Proof. Lemma follows immediately from the definition of $\psi_1(x)$. □

Lemma 2. *Let* $\psi_2(x) = \int_0^x \psi_1(t)\,\mathrm{d}t$ *and* $\psi_3(x) = \int_0^x \mathrm{e}^{-bt}\psi_1(t)\,\mathrm{d}t$. *Then*

$$\varphi_G(x) = -\ell(x) + \frac{a(x)}{f_G(x)}\{b(\psi_2(x) - 2\psi_2(\theta_0)) - b\mathrm{e}^{b\theta_0}(\psi_3(x) - 2\psi_3(\theta_0))\}.$$

Proof. By integration by parts,

$$\begin{aligned} &\int_0^{\theta_0} \frac{\ell(\theta)}{A(\theta)}\mathrm{d}G(\theta) - \int_{\theta_0}^{x} \frac{\ell(\theta)}{A(\theta)}\mathrm{d}G(\theta) \\ &= \ell(\theta_0)\psi_1(\theta_0) - \int_0^{\theta_0} \ell'(\theta)\psi_1(\theta)\,\mathrm{d}\theta - \ell(x)\psi_1(x) + \ell(\theta_0)\psi_1(\theta_0) \\ &\quad + \int_{\theta_0}^{x} \ell'(\theta)\psi_1(\theta)\,\mathrm{d}\theta \\ &= -\ell(x)\psi_1(x) + b(\psi_2(x) - 2\psi_2(\theta_0)) - b\mathrm{e}^{b\theta_0}(\psi_3(x) - 2\psi_3(\theta_0)), \end{aligned}$$

since $\ell(\theta_0) = 0$. By (2.3), Lemma 1 and the above computation, Lemma 2 follows. □

The minimum Bayes risk is derived below. Let

$$A_G = \{x \geqslant \theta_0 : \varphi_G(x) > 0\} = \{x \geqslant \theta_0 : H_G(x) > 0\}$$

and

$$B_G = \{x \geqslant \theta_0 : \varphi_G(x) < 0\} = \{x \geqslant \theta_0 : H_G(x) < 0\}.$$

Define

$$\alpha_G = \begin{cases} \sup A_G & \text{if } A_G \neq \phi, \\ \theta_0 & \text{if } A_G = \phi \end{cases}$$

and

$$\beta_G = \begin{cases} \inf B_G & \text{if } B \neq \phi, \\ \infty & \text{if } B = \phi. \end{cases}$$

From (2.3) it is obvious that $\alpha_G \leqslant \beta_G$, as

$$\int_0^{\theta_0} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta) - \int_{\theta_0}^{x} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta)$$

is decreasing in x. The minimum Bayes risk is

$$r(G, d_G) = \int_{x=\theta_0}^{\infty} \int_{\theta=0}^{x} L(\theta, d_G(x))\,\mathrm{d}G(\theta \mid x) f_G(x)\,\mathrm{d}x$$

$$= \int_{x=\theta_0}^{\infty} (d_G(x)\,\varphi_G(x) + c(x)) f_G(x)\,\mathrm{d}x. \tag{2.4}$$

Let C_Ω be the class of all prior distributions on $[0, \infty)$ such that (2.4) is finite and let $C_{[0,m]}$ be the class of all prior distributions on $[0, m]$. (The number m can be an arbitrary upper bound of the prior support. It reflects the prior knowledge of the support and is not necessarily to be sharp.) We will consider two major subclasses of C_Ω. Let $C_1 = \{G \in C_\Omega$: G is a discrete distribution with finitely many discontinuity points$\}$ and let $C_2 = \{G \in C_{[0,m]}$: f_G is decreasing and G satisfies condition (2.5)$\}$:

$$G(y) - G(x) \geqslant c_G(y - x) \tag{2.5}$$

for those x and y such that $\alpha_G - h_0 \leqslant x \leqslant y \leqslant \alpha_G$ or $\beta_G \leqslant x \leqslant y \leqslant \beta_G + h_0$, where $c_G > 0$ is some constant.

Lemma 3. *Assume that $G \in C_1$. Then $f_G(x)$ is decreasing in x on any interval $[m_i, m_{i+1})$, $i = 1, \ldots, N$, where $m_1 < \cdots < m_N$ are support points of G and $m_{N+1} = \infty$.*

Proof. From definition of $\psi_1(x)$ in Lemma 1, $\psi_1(x)$ is constant on $[m_i, m_{i+1})$. Since $f_G(x) = a(x)\psi_1(x)$ from Lemma 1 and since $a(x)$ is decreasing, $f_G(x)$ is decreasing on $[m_i, m_{i+1})$. □

Lemma 4. *Assume that $G \in C_2$. For $x \geqslant m$, we have*

$$d_G(x) = d_G(m).$$

Proof. For $x \geqslant m$, we have

$$\int_0^{\theta_0} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta) - \int_{\theta_0}^{x} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta) = \int_0^{\theta_0} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta) - \int_{\theta_0}^{m} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta).$$

Thus $\varphi_G(x)$ does not change sign on $[m, \infty)$. Therefore, $d_G(x) = d_G(m)$ for $x \geqslant m$. □

3. Empirical Bayes testing procedures with rate

In the empirical Bayes framework, we consider i.i.d. copies $(X_1, \theta_1), \ldots, (X_n, \theta_n)$ of (X, θ), where θ has some unknown prior distribution G, and given θ, X has a distribution with p.d.f. $f(x \mid \theta)$. The Xs are observable but θs are not. Let $X_1, \ldots, X_n$ be a previously seen sequence of i.i.d. random variables with p.d.f. $f_G(x)$. At the current stage, stage $n+1$, let θ be a realization of Θ and X be the associated current observation. An empirical Bayes procedure for hypothesis testing concerning the current status of the parameter, H_0: $\theta \geqslant \theta_0$ against H_1: $\theta < \theta_0$, can be obtained by first estimating $H_G(x)$ by

$$H_n(x) = H_n(X_1, \ldots, X_n; X = x),$$

and then adopting the following empirical procedures. For G in C_Ω, we adopt

$$d_n(x) = \begin{cases} 0 & \text{if } H_n(x) > 0 \\ 1 & \text{if } H_n(x) \leqslant 0 \end{cases} \quad \text{for } x \geqslant \theta_0. \tag{3.1}$$

For G in $C_{[0,m]}$, we adopt

$$d_n(x) = \begin{cases} 0 & \text{if } H_n(x) > 0 \\ 1 & \text{if } H_n(x) \leqslant 0 \end{cases} \quad \text{for } \theta_0 \leqslant x \leqslant m \tag{3.2}$$

or

$$d_n(x) = d_n(m) \quad \text{for } x > m. \tag{3.3}$$

We shall obtain $H_n(x)$ by estimating $f_G(x)$, $\psi_2(x)$ and $\psi_3(x)$ in $H_G(x)$. Note that

$$\psi_2(x) = \int_0^x \frac{f_G(t)}{a(t)} \mathrm{d}t \quad \text{and} \quad \psi_3(x) = \int_0^x \frac{\mathrm{e}^{-bt} f_G(t)}{a(t)} \mathrm{d}t.$$

There are unbiased estimators given by

$$\hat{\psi}_2(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{a(X_i)} I_{(X_i \leqslant x)} \quad \text{and} \quad \hat{\psi}_3(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{\mathrm{e}^{-bX_i}}{a(X_i)} I_{(X_i \leqslant x)}.$$

We use

$$f_n(x) = \frac{F_n(x+h) - F_n(x)}{h}$$

to estimate $f_G(x)$, where F_n is the empirical distribution function based on $\{X_1, \ldots, X_n\}$. The estimator $f_n(x)$ is a kernel estimator with a left-sided uniform kernel $K(t) = I_{[-1,0]}(t)$. The use of a left-sided kernel instead of a symmetric one is for the consistency reason to avoid dominant bias at left boundary. The expected Bayes risk of the empirical Bayes decision rule d_n is

$$Er(G, d_n) = E \int_{x=\theta_0}^{\infty} (d_n(x) \varphi_G(x) + c(x)) f_G(x) \, \mathrm{d}x, \tag{3.4}$$

where the expectation is taken with respect to the joint distribution of $X_1, \ldots, X_n$.

A sequence of empirical Bayes testing procedures $\{d_n\}_{n=1}^{\infty}$ is said to be asymptotically optimal if $Er(G,d_n) - r(G,d_G) \to 0$ as $n \to \infty$. Moreover, if $Er(G,d_n) - r(G,d_G) = O(\alpha_n)$, where $\{\alpha_n\}_{n=1}^{\infty}$ is a sequence of positive numbers such that $\lim_{n\to\infty} \alpha_n = 0$, then $\{d_n\}_{n=1}^{\infty}$ is said to be asymptotically optimal with convergence rate of order $\{\alpha_n\}_{n=1}^{\infty}$. Here we shall discuss the asymptotic optimality property for G in C_Ω, the most general class of priors, and also various convergence rates of $Er(G,d_n) - r(G,d_G)$ for G in restricted classes C_1 and C_2, respectively.

3.1. Asymptotic optimality for C_Ω

By (2.4) and (3.4), we have

$$0 \leqslant Er(G,d_n) - r(G,d_G)$$
$$= \int\int_{x=\theta_0}^{\infty} (d_n(x) - d_G(x))\,\varphi_G(x) f_G(x)\,\mathrm{d}x\,\mathrm{d}\mu_\infty, \tag{3.5}$$

where μ_∞ is the product measure on the space of sequence $(x_1, x_2, x_3, \ldots)$ resulting from the joint distribution of $(X_1, X_2, X_3, \ldots)$. For any prior G in C_Ω, it is easy to see from the finiteness of (2.4) that $\varphi_G(x) f_G(x)$ is integrable with respect to the Lebesgue measure, and hence is integrable with respect to $\mathrm{d}x\,\mathrm{d}\mu_\infty$. By the Lebesgue-dominated convergence theorem applied to (3.5), we have

$$0 \leqslant \lim_{n\to\infty} Er(G,d_n) - r(G,d_G)$$
$$= \int\int_{x=\theta_0}^{\infty} \left(\lim_{n\to\infty} d_n(x) - d_G(x)\right)\varphi_G(x) f_G(x)\,\mathrm{d}x\,\mathrm{d}\mu_\infty.$$

Thus the asymptotic optimality can be obtained by showing

$$\lim_{n\to\infty} d_n(x) = d_G(x) \quad \text{a.s.,}$$

which is easily seen from that

$$H_n(x) \to H_G(x) \quad \text{a.s.}$$

with respect to the product measure of Lebesgue measure and μ_∞.

3.2. Convergence rate over C_1

Suppose G has discontinuity at $\{m_i\}_{i=1}^{N}$. It is easy to see that $H_G(x)$, $\varphi_G(x)$ and $f_G(x)$ have jumps at $\{m_i\}_{i=1}^{N}$, and are continuous on $[m_i, m_{i+1})$, $i = 1, \ldots, N$.

By (2.4) and (3.4), we have

$$0 \leqslant Er(G,d_n) - r(G,d_G)$$
$$= \int_{\theta_0}^{\infty} (Ed_n(x) - d_G(x))\,\varphi_G(x) f_G(x)\,\mathrm{d}x. \tag{3.6}$$

From the definitions of α_G, β_G and d_G, and from Eqs. (3.2), (3.3) and (3.6), we have

$$\begin{aligned} &Er(G,d_n) - r(G,d_G) \\ &= \int_{\theta_0}^{\alpha_G} P\{d_n(x)=1\}\,\varphi_G(x) f_G(x)\,\mathrm{d}x + \int_{\beta_G}^{m_N} P\{d_n(x)=0\}\,|\varphi_G(x)| f_G(x)\,\mathrm{d}x \\ &\quad + \int_{m_N}^{\infty} P\{d_n(x)=0\}\,|\varphi_G(x)| f_G(x)\,\mathrm{d}x \\ &= \int_{\theta_0}^{\alpha_G} P\{H_n(x)\leqslant 0\}\,\varphi_G(x) f_G(x)\,\mathrm{d}x + \int_{\beta_G}^{m_N} P\{H_n(x)>0\}\,|\varphi_G(x)| f_G(x)\,\mathrm{d}x \\ &\quad + \int_{m_N}^{\infty} P\{H_n(x)>0\}\,|\varphi_G(x)| f_G(x)\,\mathrm{d}x, \end{aligned} \tag{3.7}$$

if $m_N \geqslant \beta_G$, where F_G is the distribution function of f_G. As for $\alpha_G \leqslant m_N \leqslant \beta_G$, the second term above drops; and as for $m_N < \alpha_G$, the second term above drops and the upper bound of the integral in the first term is replaced by m_N. However, without loss of generality, we may assume $m_N \geqslant \beta_G$. The convergence rate of (3.6) is investigated through the following lemmas.

Lemma 5. *For $m_i \leqslant x < m_{i+1} - h \leqslant \alpha_G$, we have*

$$P\{H_n(x)\leqslant 0\} = \mathrm{O}(\mathrm{e}^{-\gamma_1^{(i)} nh}) \quad \textit{for some constant } \gamma_1^{(i)} > 0$$

uniformly in x.

Proof.

$$\begin{aligned} P\{H_n(x)\leqslant 0\} &= P\{H_G(x) - H_n(x) \geqslant H_G(x)\} \\ &= P\{\ell(x)(f_n(x) - f_G(x)) + a(x)b(\psi_2(x) - \hat{\psi}_2(x)) \\ &\quad + 2a(x)b(\hat{\psi}_2(\theta_0) - \psi_2(\theta_0)) + a(x)b\mathrm{e}^{b\theta_0}(\hat{\psi}_3(x) - \psi_3(x)) \\ &\quad + 2a(x)b\mathrm{e}^{b\theta_0}(\psi_3(\theta_0) - \hat{\psi}_3(\theta_0)) \geqslant H_G(x)\} \\ &\leqslant P\left\{\ell(x)(f_n(x) - f_G(x)) \geqslant \frac{H_G(x)}{5}\right\} \\ &\quad + P\left\{a(x)b(\psi_2(x) - \hat{\psi}_2(x)) \geqslant \frac{H_G(x)}{5}\right\} \\ &\quad + P\left\{2a(x)b(\hat{\psi}_2(\theta_0) - \psi_2(\theta_0)) \geqslant \frac{H_G(x)}{5}\right\} \\ &\quad + P\left\{a(x)b\mathrm{e}^{b\theta_0}(\hat{\psi}_3(x) - \psi_3(x)) \geqslant \frac{H_G(x)}{5}\right\} \\ &\quad + P\left\{2a(x)b\mathrm{e}^{b\theta_0}(\psi_3(\theta_0) - \hat{\psi}_3(\theta_0)) \geqslant \frac{H_G(x)}{5}\right\}. \end{aligned} \tag{3.8}$$

The first term in (3.8) is

$$P\left\{\ell(x)(f_n(x)-f_G(x)) \geqslant \frac{H_G(x)}{5}\right\}$$

$$= P\left\{F_n(x+h)-F_n(x)-F_G(x+h)+F_G(x) \geqslant \frac{hH_G(x)}{5\ell(x)} + hf_G(x) - F_G(x+h)+F_G(x)\right\}. \tag{3.9}$$

Now

$$hf_G(x) - F_G(x+h) + F_G(x) = hf_G(x) - \int_x^{x+h} f_G(t)\,\mathrm{d}t. \tag{3.10}$$

By Lemma 3, $f_G(t)$ being decreasing on $[m_i, m_{i+1})$, (3.10) is nonnegative for $x \in [m_i, m_{i+1} - h)$. Also note that $H_G(x) > 0$ for $x < \alpha_G$. Therefore the right-hand side of the inequality in (3.9) is positive. For h small enough, we also have

$$|I_{(x<X_i<x+h)} - F_G(x+h) + F_G(x)| \leqslant 1 - F_G(x+h) + F_G(x).$$

Hence by Bernstein's (Bennett, 1962, p. 34) inequality, the tail probability in (3.9) can be bounded as follows.

$$(3.9) \leqslant \exp\left\{-\frac{t^2(x)}{2n(F_G(x+h)-F_G(x)) + t(x)(1-F_G(x+h)+F_G(x))/3}\right\},$$

where $t(x) = (nhH_G(x)/5\ell(x)) + nhf_G(x) - nF_G(x+h) + nF_G(x)$. Continued from the above inequality, we have

$$(3.9) \leqslant \exp\left\{-\frac{t^2(x)}{2n(F_G(x+h)-F_G(x)) + t(x)/3}\right\}$$

$$\leqslant \exp\left\{-\frac{(nhH_G(x)/25\ell(x))^2}{2n(F_G(x+h)-F_G(x)) + nhH_G(x)/15\ell(x) + nhf_G(x)/3}\right\}$$

$$\leqslant \exp\left\{-\frac{nhH_G^2(x)/25\ell^2(x)}{7f_G(x)/3 + H_G(x)/15\ell(x)}\right\}$$

$$\leqslant \exp\left\{-\frac{nhf_G(x)\varphi_G^2(x)/25\ell^2(x)}{7/3 + \varphi_G(x)/15\ell(x)}\right\},$$

$$\leqslant \mathrm{e}^{-\gamma_1^{(i)}nh}, \tag{3.11}$$

with $\gamma_1^{(i)} = f_G(m_{i+1}^-)\varphi_G^2(m_i)/(175\ell^2(m_{i+1})/3 + 5\ell(m_{i+1})\varphi_G(m_i)/3)$, where $f_G(m_{i+1}^-) = \lim_{\varepsilon\to 0^+} f_G(m_{i+1}-\varepsilon) = a(m_{i+1})\psi_1(m_i)$. The last inequality of (3.11) holds because, on $[m_i, m_{i+1})$, $\varphi_G(x)$ is constant, $f_G(x)$ is decreasing and $\ell(x)$ is increasing.

The second term in (3.8) is

$$P\left\{a(x)b(\psi_2(x)-\hat{\psi}_2(x)) \geqslant \frac{H_G(x)}{5}\right\} = P\left\{\psi_2(x)-\hat{\psi}_2(x) \geqslant \frac{H_G(x)}{5ba(x)}\right\}. \tag{3.12}$$

Note that

$$\hat{\psi}_2(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{a(X_i)} I_{(X_i \leqslant x)}$$

and that, for $m_i \leqslant x < m_{i+1}$,

$$0 \leqslant \frac{1}{a(X_i)} I_{(X_i \leqslant x)} \leqslant \frac{1}{a(x)}.$$

By Theorem 2 of Hoeffding (1963), we have

$$\begin{aligned}(3.12) &\leqslant \exp\{-2n(H_G(x)/5b)^2\}\\ &= \exp\{-\gamma_2^{(i)} n\},\end{aligned} \tag{3.13}$$

where $\gamma_2^{(i)} = 2\varphi_G^2(m_i) f_G^2(m_{i+1}^-)/25b^2$. By applying Theorem 2 of Hoeffding (1963) in a similar way to the rest terms in (3.8), we have

$$P\{2a(x)b(\hat{\psi}_2(\theta_0) - \psi_2(\theta_0)) \geqslant H_G(x)/5\} \leqslant e^{-\gamma_3^{(i)} n}, \tag{3.14}$$

$$P\{a(x)be^{b\theta_0}(\hat{\psi}_3(x) - \psi_3(x)) \geqslant H_G(x)/5\} \leqslant e^{-\gamma_4^{(i)} n}, \tag{3.15}$$

$$P\{2a(x)be^{b\theta_0}(\psi_3(\theta_0) - \hat{\psi}_3(\theta_0)) \geqslant H_G(x)/5\} \leqslant e^{-\gamma_5^{(i)} n} \tag{3.16}$$

for $m_i \leqslant x < m_{i+1} \leqslant \alpha_G$, where $\gamma_3^{(i)}$, $\gamma_4^{(i)}$ and $\gamma_5^{(i)}$ are some positive constants. By (3.11) and (3.13) to (3.16), Lemma 5 follows. □

Lemma 6. *For $\beta_G \leqslant m_i \leqslant x < m_{i+1} - h \leqslant m$, we have*

$$P\{H_n(x) > 0\} = O(e^{-\gamma_6^{(i)} nh}) \quad \textit{for some constant } \gamma_6^{(i)} > 0$$

uniformly in x.

Proof.

$$\begin{aligned} P\{H_n(x) > 0\} &= P\{H_G(x) - H_n(x) < H_G(x)\}\\ &= P\left\{F_n(x+h) - F_n(x) - F_G(x+h) + F_G(x) < \frac{hH_G(x)}{5\ell(x)} + hf_G(x)\right.\\ &\quad \left. - F_G(x+h) + F_G(x)\right\}\\ &\quad + P\left\{a(x)b(\psi_2(x) - \hat{\psi}_2(x)) < \frac{H_G(x)}{5}\right\}\\ &\quad + P\left\{2a(x)b(\hat{\psi}_2(\theta_0) - \psi_2(\theta_0)) < \frac{H_G(x)}{5}\right\}\\ &\quad + P\left\{a(x)be^{b\theta_0}(\hat{\psi}_3(x) - \psi_3(x)) < \frac{H_G(x)}{5}\right\}\\ &\quad + P\left\{2a(x)be^{b\theta_0}(\psi_3(\theta_0) - \hat{\psi}_3(\theta_0)) < \frac{H_G(x)}{5}\right\}. \end{aligned} \tag{3.17}$$

Note that $H_G(x) = \varphi_G(m_i) f_G(x) \leqslant \varphi_G(m_i) f_G(m_{i+1}^-) < 0$ and that, when h is sufficiently small

$$0 \leqslant h f_G(x) - \int_x^{x+h} f_G(t)\,\mathrm{d}t$$
$$\leqslant h\psi_1(x)(a(x) - a(x+h)) = \mathrm{O}(h^2).$$

Therefore

$$\frac{hH_G(x)}{5\ell(x)} + h f_G(x) - F_G(x+h) + F_G(x) < 0,$$

when h is sufficiently small. Then we can apply Bernstein's inequality to the first term in (3.17) and get

$$P\left\{F_n(x+h) - F_n(x) - F_G(x+h) + F_G(x) < \frac{hH_G(x)}{5\ell(x)}\right.$$
$$\left. + h f_G(x) - F_G(x+h) + F_G(x)\right\}$$
$$\leqslant \exp\left\{-\frac{t^2(x)}{2n(F_G(x+h) - F_G(x)) - t(x)(1 - F_G(x+h) + F_G(x))/3}\right\},$$
$$= \mathrm{O}\left(\exp\left\{-\frac{nhH_G^2(x)}{50 f_G(x)\ell^2(x)}\right\}\right)$$
$$= \mathrm{O}(\mathrm{e}^{-\gamma_6^{(i)}nh}),$$

where $\gamma_6^{(i)} = f_G(m_{i+1}^-)\,\varphi_G^2(m_i)/50\ell^2(m_{i+1})$.

Note that as $m_i \leqslant x < m_{i+1}$, by the decreasing property of $a(x)$,

$$0 \leqslant \frac{1}{a(X)} I_{(X \leqslant x)} \leqslant \frac{1}{a(x)},$$
$$0 \leqslant \frac{\mathrm{e}^{-bX}}{a(X)} I_{(X \leqslant x)} \leqslant \frac{1}{a(x)}.$$

Again apply Theorem 2 of Hoeffding (1963) to the rest four terms in (3.17). Similar to the proof of Lemma 5, they can easily be shown to be of order $\mathrm{O}(\mathrm{e}^{-\gamma_7^{(i)}n})$, for some constant $\gamma_7^{(i)} > 0$, uniformly in x. Proof is completed. □

Note that the constants $\gamma_1^{(i)}, \gamma_6^{(i)}$, $i = 1, \ldots, N-1$, can be bounded below simultaneously by

$$\inf_{\theta_0 \leqslant x < \alpha_G,\, \beta_G \leqslant x < m_N} \frac{H_G^2(x)/25\ell^2(x)}{7 f_G(x)/3 + H_G(x)/15\ell(x)},$$

which is greater than zero by the definition of α_G and β_G and the discreteness of the prior G. (Both α_G and β_G must be one of the support points.)

Lemma 7.

$$\int_{m_N}^{\infty} P\{H_n(x) > 0\} |\varphi_G(x)| f_G(x)\,dx = O(e^{-\gamma_8 nh})$$

for some constant $\gamma_8 > 0$.

Proof. Follow the argument in the proof of Lemma 6, one can get

$$\int_{m_N}^{\infty} P\{H_n(x) > 0\} |\varphi_G(x)| f_G(x)\,dx$$
$$= O\left(\int_{m_N}^{\infty} \exp\left\{ - \frac{nhH_G^2(x)}{50 f_G(x)\ell^2(x)} \right\} |\varphi_G(x)| f_G(x)\,dx \right).$$

Note that $\varphi_G^{(x)}$ and $\psi_1(x)$ are both constants on $[m_N, \infty)$. Therefore,

$$\frac{H_G^2(x)}{f_G(x)\ell^2(x)} = \frac{\psi_1(m_N)\varphi_G(m_N)^2\, a(x)}{\ell^2(x)}$$

is continuous on $[m_N, \infty)$ and then one can apply the mean-value theorem to the above integral.

$$\int_{m_N}^{\infty} P\{H_n(x) > 0\} |\varphi_G(x)| f_G(x)\,dx$$
$$= O\left(\exp\left\{ - \frac{nh\psi_1(m_N)\,\varphi_G(m_N)^2\, a(m^*)}{50\ell^2(m^*)} \right\} \int_{m_N}^{\infty} |\varphi_G(x)| f_G(x)\,dx \right)$$
$$= O(e^{-\gamma_8 nh}), \quad m^* \in (m_N, \infty),$$

where $\gamma_8 = \psi_1(m_N)\varphi_G(m_N)^2 a(m^*)/50\ell^2(m^*)$ and that $\varphi_G(x) f_G(x)$ is integrable by the finiteness of the Bayes risk in (2.4). □

Theorem 1. *Let* $\{d_n\}_{n=1}^{\infty}$ *be the sequence of empirical Bayes testing procedures constructed through* (3.2) *and* (3.3) *with* $h = O(n^{-1}(\log n)^{1+\varepsilon})$ *for arbitrarily small* $\varepsilon > 0$. *Then* $\{d_n\}_{n=1}^{\infty}$ *has the following asymptotic optimality*:

$$Er(G, d_n) - r(G, d_G) = O(n^{-1}(\log n)^{1+\varepsilon})$$

for all $G \in C_1$.

Proof. Note that

$$\int_{m_i - h}^{m_i} (Ed_n(x) - d_G(x))\,\varphi_G(x) f_G(x)\,dx = O(h)$$

for $i = 1, \ldots, N$. Therefore, from (3.7), Lemmas 5–7, we have

$$0 \leqslant Er(G, d_n) - r(G, d_G)$$
$$= O(e^{-\gamma nh}) + O(h),$$

for some constant $\gamma > 0$. By choosing $h = O(n^{-1}(\log n)^{1+\varepsilon})$, we have

$$Er(G, d_n) - r(G, d_G) = O\left(\frac{(\log n)^{1+\varepsilon}}{n}\right). \qquad \square$$

3.3. Convergence rate over C_2

Lemma 8. *Any prior distribution G in C_2 is continuous.*

Proof. Suppose G is discontinuous with a jump at $m_0 \in [0, m]$. By definition of $\psi_1(x)$

$$\psi_1(x) = \int_0^x \frac{1}{A(\theta)} \mathrm{d}G(\theta),$$

$\psi_1(x)$ is increasing and has a jump at m_0. Thus, $f_G(x) = a(x)\psi_1(x)$ has a jump at m_0 too. Since $a(x)$ is continuous by assumption, $f_G(m_0^-) < f_G(m_0)$, which contradicts the condition $f_G(x)$ being decreasing. Hence G has to be continuous. Proof is completed. $\square$

Since G is continuous, the function $\varphi_G(x)$ is continuous too. Thus

$$\varphi_G(\alpha_G) = \varphi_G(\beta_G) = 0.$$

The following lemmas are useful to study the asymptotic behavior of $Er(G, d_n) - r(G, d_G)$.

Lemma 9. (a) *For $\alpha_G - h_0 \leqslant x \leqslant \alpha_G$, we have*

$$\varphi_G(x) \geqslant c_\varphi(\alpha_G - x) \quad \textit{for some constant } c_\varphi > 0.$$

(b) *For $\beta_G \leqslant x \leqslant \beta_G + h_0$, we have*

$$|\varphi_G(x)| \geqslant c_\varphi(x - \beta_G).$$

Proof. We will give the proof for the first statement only. The proof for the second statement is similar. Since $\varphi_G(\alpha_G) = 0$, we have

$$\begin{aligned}\varphi_G(x) &= \varphi_G(x) - \frac{\varphi_G(\alpha_G)\psi_1(\alpha_G)}{\psi_1(x)} \\ &= \frac{1}{\psi_1(x)}\left(\int_0^{\theta_0} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta) - \int_{\theta_0}^x \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta)\right) \\ &\quad - \frac{1}{\psi_1(x)}\left(\int_0^{\theta_0} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta) - \int_{\theta_0}^{\alpha_G} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta)\right)\end{aligned}$$

$$= \frac{1}{\psi_1(x)} \int_x^{\alpha_G} \frac{\ell(\theta)}{A(\theta)} \mathrm{d}G(\theta)$$

$$\geqslant \frac{\ell(\alpha_G - h_0)}{\psi_1(\alpha_G) A(\alpha_G - h_0)} c_G(\alpha_G - x),$$

by condition (2.5). □

Lemma 10.

$$\int_{\theta_0}^{\alpha_G} (Ed_n(x) - d_G(x)) \varphi_G(x) f_G(x) \mathrm{d}x = \mathrm{O}\left(\left(\frac{1}{nh} \right)^{1-\delta/2} \right) + \mathrm{O}(h^{2-\delta})$$

for any arbitrarily small $\delta > 0$.

Proof.

$$\int_{\theta_0}^{\alpha_G} (Ed_n(x) - d_G(x)) \varphi_G(x) f_G(x) \mathrm{d}x = \int_{\theta_0}^{\alpha_G} P\{H_n(x) \leqslant 0\} H_G(x) \mathrm{d}x$$

$$\leqslant \int_{\theta_0}^{\alpha_G} P\{|H_n(x) - H_G(x)| \geqslant H_G(x)\} H_G(x) \mathrm{d}x$$

$$\leqslant \int_{\theta_0}^{\alpha_G} E|H_n(x) - H_G(x)|^{2-\delta} (H_G(x))^{-1+\delta} \mathrm{d}x,$$

by Chebyshev's inequality, for any arbitrarily small $\delta > 0$. The asymptotic behavior of $H_n(x)$ is studied below.

By the c_r-inequality (Loève, 1962, p. 155) and then Jensen's inequality, we have

$$E|H_n(x) - H_G(x)|^{2-\delta} \leqslant 2E|H_n(x) - EH_n(x)|^{2-\delta} + 2|EH_n(x) - H_G(x)|^{2-\delta}$$

$$\leqslant 2(\operatorname{Var} H_n(x))^{1-\delta/2} + 2|EH_n(x) - H_G(x)|^{2-\delta}.$$

Now

$$\operatorname{Var} H_n(x) \leqslant 5\ell^2(x) \operatorname{Var} f_n(x) + 5b^2 a^2(x) (\operatorname{Var} \hat{\psi}_2(x) + 4 \operatorname{Var} \hat{\psi}_2(\theta_0) + \mathrm{e}^{2b\theta_0} \operatorname{Var} \hat{\psi}_3(x) + 4\mathrm{e}^{2b\theta_0} \operatorname{Var} \hat{\psi}_3(\theta_0)),$$

where

$$\operatorname{Var} f_n(x) \leqslant \frac{1}{nh^2} (F_G(x+h) - F_G(x))$$

$$\leqslant \frac{1}{nh} f_G(x) \quad \text{(since } f_G \text{ is decreasig)}$$

$$\operatorname{Var} \hat{\psi}_2(x) = \frac{1}{n} \operatorname{Var} \frac{I_{(X \leqslant x)}}{a(X)} \leqslant \frac{1}{n} \int_0^x \frac{f_G(t)}{a^2(t)} \mathrm{d}t \leqslant \frac{1}{a^2(x) n}$$

and

$$\operatorname{Var}\hat{\psi}_3(x) = \frac{1}{n}\operatorname{Var}\frac{e^{-bX}I_{(X\leqslant x)}}{a(X)} \leqslant \frac{1}{a^2(x)n}.$$

Therefore,

$$\operatorname{Var}H_n(x) = \mathrm{O}\left(\frac{\sup_{\theta_0\leqslant x\leqslant\alpha_G}\ell^2(x)f_G(x)}{nh}\right) = \mathrm{O}\left(\frac{1}{nh}\right) \tag{3.18}$$

uniformly in $x\in[\theta_0,\alpha_G]$. Since $\hat{\psi}_2(x)$ and $\hat{\psi}_3(x)$ are both unbiased estimators of $\psi_2(x)$ and $\psi_3(x)$ respectively,

$$\begin{aligned} EH_n(x) - H_G(x) &= -\ell(x)(Ef_n(x) - f_G(x)) \\ &= -\frac{\ell(x)}{h}(F_G(x+h) - F_G(x) - hf_G(x)). \end{aligned}$$

And since f_G is decreasing,

$$\begin{aligned} 0 &\leqslant EH_n(x) - H_G(x) \\ &\leqslant \ell(x)(f_G(x) - f_G(x+h)) \\ &\leqslant \sup_{\theta_0\leqslant x\leqslant\alpha_G}\left\{\frac{\ell(x)(f_G(x)-f_G(x+h))}{h}\right\}h \\ &= \mathrm{O}(h), \end{aligned} \tag{3.19}$$

as $a(x)$ is Lipschitz continuous.

Now we only need to check that the following integral is finite to complete the proof.

$$\begin{aligned} \int_{\alpha_G-h_0}^{\alpha_G}(H_G(x))^{-1+\delta}\mathrm{d}x &= \int_{\alpha_G-h_0}^{\alpha_G}(\varphi_G(x)f_G(x))^{-1+\delta}\mathrm{d}x \\ &\leqslant (f_G(\alpha_G))^{-1+\delta}\int_{\alpha_G-h_0}^{\alpha_G}c_\varphi^{-1+\delta}(\alpha_G-x)^{-1+\delta}\mathrm{d}x \\ &\qquad \text{by Lemma 9(a)} \\ &< \infty. \qquad \square \end{aligned}$$

Lemma 11. *We have*

$$\begin{aligned} &\int_{\beta_G}^{m}(Ed_n(x) - d_G(x))\varphi_G(x)f_G(x)\mathrm{d}x \\ &\quad = \mathrm{O}\left(\left(\frac{\sup_{\beta_G\leqslant x\leqslant m}\ell^2(x)f_G(x)}{nh}\right)^{1-\delta/2}\right) \\ &\qquad + \mathrm{O}\left(\left(\sup_{\beta_G\leqslant x\leqslant m}\frac{\ell(x)(f_G(x)-f_G(x+h))}{h}h\right)^{2-\delta}\right) \end{aligned}$$

and

$$(Ed_n(m)-1)\varphi_G(m)(1-F_G(m))$$
$$=\mathrm{O}\left(\left(\frac{\sup_{\beta_G\leqslant x\leqslant m}\ell^2(x)f_G(x)}{nh}\right)^{1-\delta/2}\right)$$
$$+\mathrm{O}\left(\left(\sup_{\beta_G\leqslant x\leqslant m}\frac{\ell(x)(f_G(x)-f_G(x+h))}{h}h\right)^{2-\delta}\right).$$

Proof. Follow the argument in the proof of Lemma 10 and replace $\sup_{\theta_0\leqslant x\leqslant\alpha_G}$ by $\sup_{\beta_G\leqslant x\leqslant m}$ in (3.18) and (3.19). □

Suppose that, as x goes to infinity, $f_G(x)=\mathrm{O}(x^{-q})$, $q>1$ to insure integrability of f_G, and that $(f_G(x)-f_G(x+h))/h=\mathrm{O}(x^{-q-1})$. Then we have the following result.

Lemma 12.

$$Er(G,d_n)-r(G,d_G)=\begin{cases}\mathrm{O}((m^{2-q/nh})^{1-\delta/2})+\mathrm{O}(h^{2-\delta}) & \text{for } 1<q<2,\\ \mathrm{O}((nh)^{-1+\delta/2})+\mathrm{O}(h^{2-\delta}) & \text{for } q\geqslant 2.\end{cases}$$

Proof. As x goes to infinity, we have

$$\ell^2(x)f_G(x)=\mathrm{O}(x^{2-q})\quad\text{and}\quad \ell(x)(f_G(x)-f_G(x+h))/h=\mathrm{O}(x^{-q}).$$

Therefore, as $m\to\infty$,

$$\sup_{\beta_G\leqslant x\leqslant m}\ell^2(x)f_G(x)=\begin{cases}\mathrm{O}(m^{2-q}) & \text{for } 1<q<2,\\ \mathrm{O}(1) & \text{for } q\geqslant 2\end{cases}$$

and

$$\sup_{\beta_G\leqslant x\leqslant m}\ell(x)(f_G(x)-f_G(x+h))/h=\mathrm{O}(1).\qquad\square$$

Theorem 2. *Let $\{d_n(x)\}_{n=1}^{\infty}$ be the sequence of empirical Bayes testing procedures constructed through* (3.2) *and* (3.3) *with*

$$h=\begin{cases}\mathrm{O}(m^{(2-q)/3}n^{-1/3}) & \text{for } 1<q<2,\\ \mathrm{O}(n^{-1/3}) & \text{for } q\geqslant 2.\end{cases}$$

Then $\{d_n\}_{n=1}^{\infty}$ has the following asymptotic optimality:

$$Er(G,d_n)-r(G,d_G)=\begin{cases}\mathrm{O}(m^{(2-q)(2/3-\delta^*)}n^{-2/3+\delta^*}) & \text{for } 1<q<2,\\ \mathrm{O}(n^{-2/3+\delta^*}) & \text{for } q\geqslant 2\end{cases}$$

for any arbitrarily small $\delta^*>0$.

Proof. Theorem 2 follows immediately from the discussion above with $\delta^*=\delta/3$. □

Theoretically, for a fixed value m the convergence rate is $O(n^{-2/3+\delta^*})$ as $n \to \infty$. In addition to the rate, Theorem 2 also tells us the effect of the values of m, small versus large, on the convergence rate. For $1 < q < 2$, the large m is, the larger the sample size n is required for the asymptotic optimality result to take effect. However, when $f_G(x)$ decays to zero at a rate not slower than x^{-2}, the value m has no effect on the rate.

Acknowledgements

The author would like to express her gratitude to Professor TaChen Liang for introducing her into the empirical Bayes methods research area and for his constant encouragement and helpful discussion.

The author would also like to thank a referee for some very insightful comments, in particular for the one which leads to the consideration of the effect on rate by the value m.

References

Basu, A.P. and N. Ebrahimi (1991). Bayesian approach to life testing and reliability estimation using asymmetric loss function. *J. Statist. Plann. Inference* **29**, 21–31.

Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57**, 33–45.

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer, New York.

Datta, S. (1991). Nonparametric empirical Bayes estimation with $O(n^{-1/2})$ rate of a truncation parameter. *Statist. Decisions* **9**, 45–61.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30.

Kuo, L. and D.K. Dey (1990). On the admissibility of the linear estimators of the Poisson mean using Linex loss functions. *Statist. Decisions* **8**, 201–210.

Liang, T. (1990). On the convergence rates of a monotone empirical Bayes test for uniform distribution. *J. Statist. Plann. Inference* **26**, 25–34.

Liang, T. (1993). Convergence rates for empirical Bayes estimation of the scale parameter in a Pareto distribution. *Comput. Statist. Data Anal.* **16**, 35–46.

Loève, M. (1962). *Probability Theory*, 3rd ed. D. Van Nostrand, New York.

Nogami, Y. (1988). Convergence rates for empirical Bayes estimation in the uniform $U(0, \theta)$ distribution. *Ann. Statist.* **16**, 1335–1341.

Prasad, B. and R.S. Singh (1990). Estimation of prior distribution and empirical Bayes estimation in a nonexponential family. *J. Statist. Plann. Inference* **24**, 81–86.

Robbins, H. (1956). An empirical Bayes approach to statistics. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **I**, 157–164.

Singh, R.S. and B. Prasad (1989). Uniformly strongly consistent prior distribution and empirical Bayes estimators with asymptotic optimality and rates in a non-exponential family. *Sankhya A* **51**, 334–342.

Tiwari, R.C. and J.N. Zalkikar (1990). Empirical Bayes estimation of the scale parameter in a Pareto distribution. *Comput. Statist. Data Anal.* **10**, 261–270.

Van Houwelingen, J.C. (1987). Monotone empirical Bayes tests for uniform distribution using the maximum likelihood estimator of a decreasing density. *Ann. Statist.* **15**, 875–879.

Varian, H.R. (1975). A Bayesian approach to real estate assessment. In: S.E. Feinberg and A. Zellner, Eds., *Studies in Bayesian Econometrics and Statistics in Honor of L.J. Savage.* North-Holland, Amsterdam, 195–208.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss function. *J. Amer. Statist. Assoc.* **81**, 446–451.