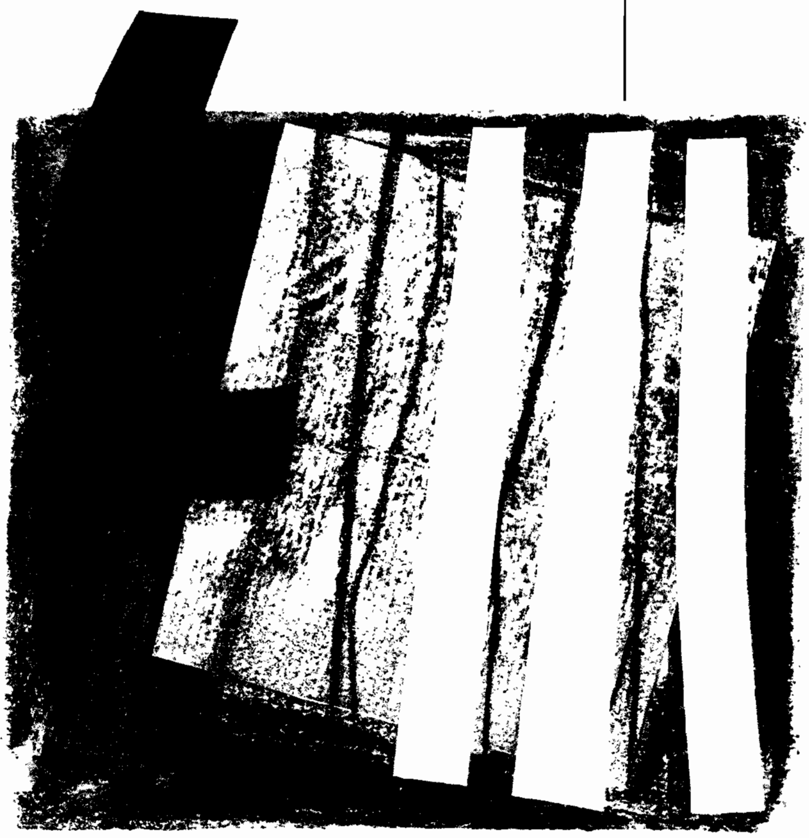
IDENTIFIABILITY OF DIFFERENTIABLE BAYES ESTIMATORS OF THE UNIFORM SCALE PARAMETER

Rosa E. Lillo and Miguel Martín

00-10



WORKING PAPERS

Universidad Carlos III de Madrid

Working Paper 00-10
Statistics and Econometrics Series (4)
February 2000

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 624-9849

IDENTIFIABILITY OF DIFFERENTIABLE BAYES ESTIMATORS OF THE UNIFORM SCALE PARAMETER.

Rosa E. Lillo and Miguel Martín*

Abstract

The problem of estimating the uniform scale parameter under the squared error loss function is investigated from a Bayesian viewpoint. A complete characterization of differentiable Bayes estimators and generalized Bayes estimators is given. The solution determines a family of prior measures both proper and improper, involving densities whose support is the whole parameter space, i.e, the interval $(0,\infty)$. Relations between degrees of smoothness of the estimators and the priors are investigated. We will also consider sequences, depending on the sample size, of Bayes (generalized Bayes) estimators with a fixed structure which are generated from a unique prior measure. They will be named *strong Bayes sequences* or *strong generalized Bayes sequences*. We characterize this type of Bayes estimation which is more restrictive than the usual one. As a consequence of the characterization results, we will prove that strong Bayes sequences of polynomial form are not possible for the uniform scale parameter. Moreover we will show that the sequence whose components are the *minimum risk equivariant* estimator for each sample size is the best strong generalized Bayes sequence of polynomial form.

Keywords: Bayesian analysis; characterization theorems; Bayes and generalized Bayes estimators; strong Bayes sequence; scale parameters.

*Lillo, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-91-624.98.57, Fax: 34-91-624.98.49, e-mail: lillo@est-econ.uc3m.es; Martín, Departamento de Estadística e Investigación Operativa, Universidad Complutense de Madrid, Avda. Complutense s/n, 28040 Madrid,. Spain, e-mail: martin@mat.ucm.es

IDENTIFIABILITY OF DIFFERENTIABLE BAYES ESTIMATORS OF THE UNIFORM SCALE PARAMETER

Rosa E. Lillo

Department of Statistics and Econometrics, Carlos III University

Madrid, 126. 28903 Getafe, Madrid, Spain

E–mail: lillo@est-econ.uc3m.es

and

Miguel Martín

Department of Statistics and Operations Research, Complutense University

Avda. Complutense s/n. 28040 Madrid, Spain

ABSTRACT

The problem of estimating the uniform scale parameter under the squared error loss function is investigated from a Bayesian viewpoint. A complete characterization of differentiable Bayes estimators and generalized Bayes estimators is given. The solution determines a family of prior measures both proper and improper, involving densities whose support is the whole parameter space, i.e, the interval $(0, \infty)$. Relations between degrees of smoothness of the estimators and the priors are investigated. We will also consider sequences, depending on the sample size, of Bayes (generalized Bayes) estimators with a fixed structure which are generated from a unique prior measure. They will be named *strong Bayes sequences* or *strong generalized Bayes sequences.* We characterize this type of Bayes estimation which is more restrictive than the usual one. As a consequence of the characterization results, we will prove that strong Bayes sequences of polynomial form are not possible for the uniform scale parameter. Moreover we will show that the sequence whose components are the *minimum risk equivariant* estimator for each sample size is the best strong generalized Bayes sequence of polynomial form.

Key words: Bayesian analysis, characterization theorems, Bayes and generalized Bayes estimators, strong Bayes sequence, scale parameters.

1 INTRODUCTION

Let $X_1, \ldots, X_n$ be a random sample from a uniform distribution on $(0, \theta)$, $\theta > 0$. The problem considered in this paper is the estimation of the scale parameter θ from a Bayesian viewpoint. The admissibility of estimators derived here is compared in terms of the squared error loss function. Since $Y = \max\{X_1, \ldots, X_n\}$ is a complete sufficient statistic for θ, this work focuses on estimators of the scale parameter based on Y. From now on, we will use $\hat{\theta}_n(Y)$ to denote an estimator of θ. A good estimator of θ should verify $\hat{\theta}_n(Y) \geq Y$.

The problem of admissibility is usually linked with the Bayesian estimation. Thus, given a point estimator of an unknown parameter, it is interesting to know if it is a Bayes estimator. In this context, results characterizing the analytic form of Bayes estimators (BE) are useful tools in mathematical inference. Theorems of characterization of linear BE have previously been given in the literature [see Johnson (1967), Kagan (1973), Goldstein (1975)]. Specifically, linear Bayes prediction has been used by the actuarial profession [see Kahn (1975)]. Some recent papers have focused on the problem of identifiability of a mixture model, connecting it with the question of unique determination of prior distributions in the Bayesian context. Cacoullos and Papageorgiou (1983) investigated the uniqueness of determination of the bivariate distribution for some discrete type conditionals and an arbitrary consistent regression function. A generalization of such results is provided in Wesolowski (1995a). Sapatinas (1995), in the case of a power–series mixture, obtained characterizations based on a regression function (posterior expectation). Other conditional specifications of similar nature are given in Wesolowski (1995b), Gupta and Wesolowski (1997) and Papageorgiou and Wesolowski (1997). Diaconis and Ylvisaker (1979) characterized linear Bayes estimation of the mean parameter of a random vector belonging to the exponential family. In this family, the conjugate priors typically used satisfy and are characterized by a relation of posterior linearity. As we will see in this paper, similar results do not apply to the problem of estimating the scale uniform parameter when the linear case is considered.

In spite of the importance of identiafibility of general BE, our objective is to characterize mathematically the form of (BE) of the uniform scale parameter with *good* analytic properties such as *to be k–times differentiable, $k \geq 1$* and *to be strictly increasing* within the parameter space, i.e., $(0, \infty)$. This approach leads us to consider absolutely continuous priors assuming additionally that their densities are *at least* $k-1$–times differentiable. Gupta and Wesolowski (1997) studied the problem of identification of uniform mixtures via posterior means when it is assumed that the prior is absolutely continuous and the sample size is unity. Although characterization problems applied to Bayesian estimation are not considered in this reference, it is interesting to note the coincidence of the results, specifically the agreement on the form of the prior density, which is derived by methods that are different from those developed here. In Lillo and Martín (1997), a complete characterization of BE without any additional conditions is given, i.e., absolute continuity of the priors involved is not assumed. However, we are now interested in the identifiability of BE with specifying precise forms. Hence, the aim of the current paper is to consider characterization of differentiable and strictly increasing BE of θ. Moreover, our approach relies on the identifiability of prior distributions associated with this type of BE, trying to obtain the prior density directly. The interest in studying this subject lies in characterizing BE with analytic properties that implies differentiability such as *polynomial form*, *convexity*, *concavity* and so on. Although the polynomial case is the only one studied in detail in this paper, other cases can be analyzed following arguments similar to those developed here.

The prior densities considered in this paper are positive for all $\theta > 0$. The reason for focusing on the estimation problem with these types of densities is that the parameter space is the interval $(0, \infty)$ and not any subset of it; therefore a positive probability should be assigned to every open interval. This restriction naturally involves prior measures based on positive densities within the parameter space and hence, increasing and differentiable BE can be derived. If we also assume that the prior is k–times differentiable with $k \geq 1$, $k+1$–times differentiable estimators can be obtained.

Generalized Bayes estimators (GBE) based on prior measures over the whole parameter space are also characterized in this paper. This class of estimators is important in this framework since many well known estimators such as *the minimum risk invariant* are GBE and not BE. Both problems of characterization are solved in Theorem 2.2 and Theorem 2.3, respectively. As a consequence of these two theorems, we will show that BE of the form aY are not possible; however, linear BE of the form $aY + b$, $b > 0$ are possible with some restriction on a.

A fundamental problem in Bayesian analysis consists of choosing an appropriate prior probability, proper or improper, that accurately reflects the available information on the unknown parameter. On the other hand, it would be advisable to consider sequences of estimators depending on the sample size whose components are BE or GBE with particular analytic form. With this comment in mind, we introduce the notion of *strong Bayes* sequence and *strong generalized Bayes* sequence as a sequence of estimators, one for each sample size n, derived from the same prior measure, proper or improper. We will also characterize this type of sequence mathematically in Theorem 2.4.

Applying this result, we will prove that strong Bayes sequences of polynomial form are not possible for θ and strong generalized Bayes sequences of polynomial form are only feasible with the structure $a_n Y$. Specifically, the sequence of *minimum risk invariant* estimators is a strong generalized Bayes sequence. This fact completes the list of good properties of this estimator.

Now we will introduce the notation to be used in this paper to refer to the two families of prior densities, proper and improper, respectively.

$$(1.1)\ \mathcal{D} = \left\{ f : f(\theta) > 0 \text{ for all } \theta > 0,\ \int_0^\infty f(\theta)\, d\theta = 1 \right\}$$

$$(1.2)\ \mathcal{D}_\infty = \left\{ f : f(\theta) > 0 \text{ for all } \theta > 0,\ \int_0^\infty f(\theta) d\theta = \infty,\ \int_x^\infty f(\theta) d\theta < \infty \text{ for all } x > 0 \right\}$$

Given $f \in \mathcal{D}$, the associated Bayes estimator under squared error loss is

$$\hat{\theta}_n(y) = \frac{\int_y^\infty \theta^{-(n-1)} f(\theta) d\theta}{\int_y^\infty \theta^{-n} f(\theta) d\theta} \tag{1.3}$$

We should bear in mind that BE related to improper prior densities are considered and are designated by the adjective *generalized*, provided that the integrals in (1.3) are finite. It is easy to see that for the uniform scale parameter this occurs if f belongs to class $\mathcal{D}_\infty$ defined above. From now on, the set of BE (GBE) derived from prior measures belonging to $\mathcal{D}$ ($\mathcal{D}_\infty$) and related to the sample size n is denoted by $\bar{\mathcal{D}}_n$ ($\bar{\mathcal{D}}_{n\infty}$).

2 CHARACTERIZATION THEOREMS

We will devote this section to the mathematical characterization of BE of the uniform scale parameter. Exploiting equation (1.3), Theorem 2.1 given below shows a general result concerning characterizations based on positive integrable functions and hence, characterization of BE and GBE are revealed as special cases. Let $\mathcal{I}$ be the set of positive functions integrable over $(0,\infty)$ with the weight $\omega(t) = t$, $t \in (0,\infty)$. For any $g \in \mathcal{I}$, consider now the basic equation,

$$m(y)\int_y^\infty g(t)\,dt = \int_y^\infty t g(t)\,dt, \qquad \text{for all } y > 0, \tag{2.4}$$

where $m : (0,\infty) \to I\!\!R$ is another function. In the sequel, let $\mathcal{S}$ denote the functions that verify equation (2.4) for some $g \in \mathcal{I}$. From characterization of a class of functions broader than $\mathcal{S}$, given in Zoroa et al. ((1990), Theorem 3.1), seven sufficient conditions for $m \in \mathcal{S}$ can be obtained, although some of them are not necessary. The first result of this paper deals with the study of analytic properties to typify the class $\mathcal{S}$.

Theorem 2.1 *Let $m : (0,\infty) \to I\!\!R$ be any real function. Then, $m \in \mathcal{S}$ if and only if the following four conditions hold.*

Condition (C1). $m(y) > y$, *for all* $y > 0$.
Condition (C2). $\lim_{y\to 0} m(y) = m(0^+) > 0$.
Condition (C3). m *is differentiable with positive derivative.*
Condition (C4). $m'(y)\,(m(y)-y)^{-2}\exp\left\{-\int_0^y (m(t)-t)^{-1}dt\right\}$ *is integrable over* $(0,\infty)$ *with the weight* $\omega(t) = t$.

Moreover, if these four conditions are fulfilled, a unique function $g \in \mathcal{I}$ exists given by

$$g(y) = \frac{m(0^+)m'(y)}{[m(y)-y]^2}\exp\left\{-\int_0^y \frac{dt}{m(t)-t}\right\}, \qquad \text{for all } y > 0, \tag{2.5}$$

with the property that the solution of (2.4) related to (2.5) is m.

Proof: First, we prove the necessity. Assume that $m \in \mathcal{S}$. Then, from (2.4),

$$m(y) = \frac{\int_y^\infty t g(t)\,dt}{\int_y^\infty g(t)\,dt}, \qquad \text{for some } g \in \mathcal{I} \tag{2.6}$$

Hence, (C1) follows easily and a simple sufficient condition for (C2) to hold is $g \in \mathcal{I}$. By standard calculations in (2.6), (C3) is obvious, that is, m is differentiable with positive derivative since

$$m'(y) \int_y^\infty g(t)\,dt = [m(y) - y]\, g(y), \qquad \text{for all } y > 0. \tag{2.7}$$

Hence the function $[m(y) - y]g(y)/m'(y)$ is differentiable and

$$\left\{ \ln \left[\int_y^\infty g(t)\,dt \right] \right\}' = -\frac{m'(y)}{m(y) - y}, \qquad \text{for all } y > 0.$$

Thus, we find

$$g(y) = \frac{m'(y)}{m(y) - y} \exp\left\{ - \int_0^y \frac{m'(t)}{m(t) - t}\,dt \right\}, \qquad \text{for all } y > 0.$$

Therefore

$$\int_0^y \frac{m'(t)}{m(t) - t}\,dt = \ln[m(y) - y] - \ln[m(0^+)] + \int_0^y \frac{dt}{m(t) - t}, \qquad \text{for all } y > 0 \tag{2.8}$$

and finally,

$$g(y) = \frac{m(0^+) m'(y)}{[m(y) - y]^2} \exp\left\{ - \int_0^y \frac{dt}{m(t) - t} \right\}, \qquad \text{for all } y > 0. \tag{2.9}$$

(C4) is verified by noting that g is integrable over $(0, \infty)$ with the weight $\omega(t) = t$. This completes the necessity conditions. To derive the sufficiency conditions, assume that m is a real function satisfying the four conditions stated in the statement of the theorem and consider the function g defined from m as in (2.9). Note that (C4) implies that $g \in \mathcal{I}$. To complete the proof, we only need to show that the function m defined as in (2.6) considering g and denoted by m_g is the initial m. Integrating by parts, we have

$$\begin{aligned} m_g(y) &= \frac{y\exp\left\{-\int_0^y \frac{m'(t)}{m(t)-t}\,dt\right\} - L_1 + \int_y^\infty \exp\left\{-\int_0^x \frac{m'(t)}{m(t)-t}\,dt\right\}\,dx}{\exp\left\{-\int_0^y \frac{m'(t)}{m(t)-t}\,dt\right\} - L_2} \\ &= \frac{y\exp\left\{-\int_0^y \frac{m'(t)}{m(t)-t}\,dt\right\} + \exp\left\{-\int_0^x \frac{d(t)}{m(t)-t}\right\}}{\exp\left\{-\int_0^y \frac{m'(t)}{m(t)-t}\,dt\right\}}. \end{aligned}$$

From (2.8) we immediately have $m_g = m$. Note that we have also proved the uniqueness of $g \in \mathcal{I}$ during both, our proof of sufficiency and the derivation of g in (2.9). Hence, the proof is complete. ∎

Lemma 2.1 *A sufficient condition for (C4) is*

Condition (C4'). $(m(y) - y)^{-1}$ *is not integrable over* $(0, \infty)$.

Proof: By routine calculations, we obtain

$$\begin{aligned} \int_0^\infty y\frac{m'(y)}{[m(y)-y]^2}\exp\left\{-\int_0^y \frac{dt}{m(t)-t}\right\}dy &= \int_0^\infty \exp\left\{-\int_0^y \frac{m'(t)}{m(t)-t}\,dt\right\}dy - L_1 \\ &= 1 - L_1 - L_2, \end{aligned}$$

where

$$\begin{aligned} L_1 &= \lim_{y\to\infty} y\exp\left\{-\int_0^y \frac{m'(t)}{m(t)-t}\,dt\right\}, \\ L_2 &= \lim_{y\to\infty} \exp\left\{-\int_0^y \frac{m'(t)}{m(t)-t}\,dt\right\}. \end{aligned}$$

Observe that $L_1 = L_2 = 0$ implies (C4). Since $L_1 = 0$ implies $L_2 = 0$, we only need to prove the first limit. Hence, it is easy to see that:

$$L_1 = \lim_{y\to\infty} \frac{m(y)}{m(y)-y}\exp\left\{-\int_0^y \frac{dt}{m(t)-t}\right\}.$$

Applying L'Hopital's rule and considering (C4'), it follows that

$$
\begin{aligned}
L_1 &= \lim_{y\to\infty} \frac{m'(y)}{[m'(y)-1]\exp\left\{\int_0^y \frac{dt}{m(t)-t}\right\} + \exp\left\{\int_0^y \frac{dt}{m(t)-t}\right\}} \\
&= \lim_{y\to\infty} \exp\left\{-\int_0^y \frac{dt}{m(t)-t}\right\} = 0.
\end{aligned}
$$

Since $L_1 = 0$ implies $L_2 = 0$, the proof is complete. ∎

Remark 2.1 (C4') is not a necessary condition in Theorem 2.1. For instance, consider the function $m(y) = y^2 + y + 1$ suggested by a referee. Then (C4') is not satisfied but (C1-C4) are and g as defined by (2.5) belongs to $\mathcal{I}$.

Remark 2.2 Note that if g is a density function with infinite support $(0,\infty)$, $m(y) - y$ in (2.4) is a *mean residual life function.* In the Reliability framework, characterizations of these functions have been given, (see Guess and Proschan (1988) for a review), although different from that obtained considering Theorem 2.1.

Now we return to our aim in this paper. Let $(X_1, \ldots, X_n)$ be a random sample from a uniform distribution over $(0,\theta)$, $\theta > 0$ and let $Y = \max\{X_1, \ldots, X_n\}$. The following results unify many standard, but cumbersome, Bayesian calculations. A characterization of BE of θ derived from prior density functions of class $\mathcal{D}$ follows easily from the preceding Theorem 2.1.

Theorem 2.2 *Let $\hat{\theta}_n$ be an estimator of θ. If $n = 1$, the necessary and sufficient condition for $\hat{\theta}_n \in \bar{\mathcal{D}}_n$ is $\hat{\theta}_n \in \mathcal{I}$; that is, (C1) to (C4) of Theorem 2.1 hold. If $n > 1$, the necessary and sufficient conditions are $\hat{\theta}_n \in \mathcal{I}$ and*

Condition (C5).

$$\text{(2.10)} \qquad \int_0^\infty t^n \frac{\hat{\theta}'_n(t)}{[\hat{\theta}_n(t)-t]^2} \exp\left\{-\int_0^t \frac{dz}{\hat{\theta}_n(z)-z}\right\} dt < \infty.$$

Hence, if these conditions are fulfilled, a unique prior density function $f_{\hat{\theta}_n} \in \mathcal{D}$ exists given by

$$\text{(2.11)} \qquad f_{\hat{\theta}_n}(\theta) \propto \theta^n \frac{\hat{\theta}'_n(\theta)}{[\hat{\theta}_n(\theta)-\theta]^2} \exp\left\{-\int_0^\theta \frac{dz}{\hat{\theta}_n(z)-z}\right\}, \qquad \text{for all } \theta > 0$$

with the property that the Bayes estimator related to (2.11) is $\hat{\theta}_n$.

Proof: Suppose $\hat{\theta}_n \in \bar{\mathcal{D}}_n$, then it follows from (1.3) taking $g(t) = t^n f(t)$ that $\hat{\theta}_n$ verifies equation (2.4). Hence, Conditions (C1)-(C4) of Theorem 2.1 provide an analytic characterization of $\hat{\theta}_n$. Function $h(y) = yg(y)$ is integrable over $(0, \infty)$ since g in (2.5) belongs to class $\mathcal{I}$. This implies that, if $n = 1$, the function f derived is a proper density and if $n > 1$, it is necessary to add condition (C5) (equation 2.10) in order to ensure that function f involved be a proper density. ■

Remark 2.3 The prior density given in (2.11) can be derived from the results of Gupta and Wesolowski (1997) [Equation (3.3), p. 178]. However, in this reference, necessary and sufficient conditions are not specified. Although the issue of identification of estimators is the same, the developments used to derive them are different.

We now consider the question of characterizing GBE of the uniform scale parameter; i.e. we want to explore when an estimator of θ belongs to the class $\bar{\mathcal{D}}_{n\infty}$.

Theorem 2.3 *Let $\hat{\theta}_n$ be an estimator of θ. Necessary and sufficient conditions for $\hat{\theta}_n \in \bar{\mathcal{D}}_{n\infty}$ are (C1), (C3), and*

Condition (C6). $\hat{\theta}_n(0^+) = 0$.

Condition (C7). *If $R(t)$ is a primitive function of $[\hat{\theta}_n(t) - t]^{-1}$, then*

$$\int_y^\infty t^n \frac{\hat{\theta}_n'(t)}{[\hat{\theta}_n(t) - t]^2} e^{-R(t)} dt < \infty, \qquad \text{for all } y > 0.$$

Condition (C8). *As $t \to 0$, $\frac{t^n \hat{\theta}_n'(t)}{[\hat{\theta}_n(t)-t]^2} e^{-R(t)} \longrightarrow \infty$*

Furthermore, if these conditions hold, a unique family of improper priors exists, denoted by $\mathcal{F}_{\hat{\theta}_n}$, such that if $f \in \mathcal{F}_{\hat{\theta}_n}$, then $f \in \mathcal{D}_\infty$ and there exists some $k \in I\!R^+$ such that

$$f(\theta) = k\theta^n \frac{\hat{\theta}_n'(\theta)}{[\hat{\theta}_n(\theta) - \theta]^2} e^{-R(\theta)}, \qquad \text{for all } \theta > 0 \tag{2.12}$$

Proof: Note that $f(x)/\int_y^\infty f(t)dt$ for $0 < y < x$, is the kernel of a density function taking values on (y, ∞). Conditions (C1), (C3) and (C7) are immediate by extending the result from the former characterization of Theorem 2.1 to the case of functions integrable over (y, ∞). Hence,

$$\frac{f(x)}{\left(\int_y^\infty f(t)dt\right)} = k_y x^n \frac{\hat{\theta}_n'(x)}{[\hat{\theta}_n(x) - x]^2} e^{-R(x)}, \qquad \text{for all } x > y.$$

Thus, it is easy to see that $\left(\int_y^\infty f(t)dt\right) k_y$ is independent of y and therefore, the equality (2.12) is clear. (C6) and (C8) follow easily considering that $f \in \mathcal{D}_\infty$. The sufficiency conditions are provided by using arguments parallel to Theorem 2.1. The uniqueness, except for proportional constants, follows naturally. ∎

Now, a relation between the orders of differentiability of the estimators and those of the corresponding priors is given.

Corollary 2.1 *Suppose that $\hat{\theta}_n \in \bar{\mathcal{D}}_n(\bar{\mathcal{D}}_{n\infty})$. Then $\hat{\theta}_n$ is k-times differentiable, if and only if, the corresponding prior is $k-1$-times differentiable.*

Proof: Taking $g(t) = t^{-n} f(t)$ in (2.7), we have

$$m'(y) = \left[\frac{m(y) - y}{t^n}\right] \frac{f(t)}{\int_y^\infty t^{-n} f(t)\, dt}, \qquad \text{for all } y > 0. \tag{2.13}$$

Hence, the conclusion yields easily observing (2.13), since m' and f have the same order of differentiabiliy. ∎

From (1.3), it is easy to see that the form of the Bayes decision depends on the sample size. However, it is of considerable interest to know if a prior density, proper or improper, can generate a finite family of BE or GBE for a fixed parameter; that is, if

$$\hat{\theta}(y) = E\left[\theta \,|\, y\right] < \infty, \text{ for all } n.$$

On the other hand, an interesting question is: given a family of BE depending on sample size with a particular structure, does any prior exist whose associated BE for all n are the components of this family?. With this idea in mind, we propose the following definition:

Definition 2.1 *Let $\left(\hat{\theta}_n(x_1, \ldots, x_n)\right)_{n \in \mathbb{N}}$ be a sequence of Bayes (generalized Bayes) estimators. Suppose that a prior (improper prior) density function $f \in \mathcal{D}$ $(\mathcal{D}_\infty)$ exists verifying that the Bayes estimator associated with f and related to the random sample $X_1, \ldots, X_n$ is $\hat{\theta}_n(x_1, \ldots, x_n)$. Then, a sequence of estimators satisfying this condition* *is named a* **strong–Bayes sequence** *(SBS), (***strong generalized Bayes sequence***), (SGBS).*

Now, to answer our last questions related to the identifiability of a given structure of sequences of BE, we obtain a characterization result for estimators of the uniform scale parameter which is restricted to measures provided by class $\mathcal{D}$ and $\mathcal{D}_\infty$.

Theorem 2.4 *Let $\hat{\theta}_n$ be a sequence of estimators for θ. Then, $\hat{\theta}_n$ is an SBS (SGBS) related to class $\mathcal{D}$ ($\mathcal{D}_\infty$), if and only if, $\hat{\theta}_n(y) \in \bar{\mathcal{D}}_n$, ($\bar{\mathcal{D}}_{n\infty}$) and $h_r(y) = h_1(y) = h(y)$, $r = 1, 2, \ldots$, with*

$$h_r(y) = \frac{M_r(y)y^{-r} - 1}{M_r'(y)} M_r(y), \qquad \text{for all } y > 0 \tag{2.14}$$

where $M_r(y) = \prod_{k=1}^{r} \hat{\theta}_k(y)$, $y > 0$. Then, the (generalized) prior density f associated to the sequence $\hat{\theta}_n$ takes the form

$$f(\theta) \propto \frac{1}{h(\theta)} \exp\left\{ - \int \frac{d\theta}{h(\theta)} \right\}, \qquad \text{for all } \theta > 0. \tag{2.15}$$

Proof: Observe that considering (2.4) for any n with the same f implies:

$$M_r(y) = \frac{\int_y^\infty f(t)\,dt}{\int_y^\infty t^{-r} f(t)\,dt}, \qquad \text{for all } y > 0$$

and by differentiating,

$$M_r'(y) = \frac{[y^{-r} M_r(y) - 1]}{\int_y^\infty t^{-r} f(t)\,dt} f(y), \qquad \text{for all } y > 0.$$

Thus, equation (2.14) holds for all $r \geq 1$ and the form of f is easily derived. ∎

On the other hand, as pointed out in the introduction, characterizations of k–times differentiable SBS, ($SGBS$) are useful to obtain estimators with good analytical properties. The following result deals with the case $k = 2$.

Theorem 2.5 *Let $(\hat{\theta}_n)_{n\in\mathbb{N}}$ be a SBS (SGBS) related to class $\mathcal{D}$. Then, $\hat{\theta}_n$, $n > 0$ is twice differentiable, if and only if, $\hat{\theta}_n(y) \in \bar{\mathcal{D}}_n$, $(\bar{\mathcal{D}}_{n\infty})$ and*

$$\xi_n(y) = \xi_1(y) + \frac{n-1}{y}, \qquad \text{for all } y > 0 \tag{2.16}$$

where

$$\xi_n(y) = \frac{2\hat{\theta}_n'(y)^2 - \hat{\theta}_n'(y) - \hat{\theta}_n''(y)\left[\hat{\theta}_n(y) - y\right]}{\hat{\theta}_n'(y)\left[\hat{\theta}_n(y) - y\right]}, \qquad \text{for all } y > 0. \tag{2.17}$$

Proof: First, note that a different equation for f is obtained by differentiating (2.13), whose solution is easy to obtain by routine calculations. Then, considering that f is differentiable and $(\hat{\theta}_n)_{n\in\mathbb{N}}$ is a SBS (SGBS), its density function has to hold for every $n > 0$

$$f(y) = y\exp\left\{-\int_0^y \xi_n(t)dt\right\},\ n > 0 \tag{2.18}$$

where ξ_n is defined in (2.17). Thus, from straightforward calculations using (2.18) with $n = 1$, the equation (2.16) is easily obtained. ∎

In the next section we will show how this definition and the previous characterization result may be applied in the context of the linear estimation of the scale uniform parameter.

3 LINEAR BAYES ESTIMATORS

As an application of results obtained in previous sections, we will study the case of linear estimators in detail. This example is also of interest since linear Bayes prediction has been widely used by the actuarial profession under the heading of the credibility theory, (see Kahn (1975)).

Corollary 3.1 *Let $X_1, \ldots, X_n$ be a random sample from a uniform distribution over $(0, \theta)$. Then, $\hat{\theta}_n(y) = ay$, $y > 0$ is a generalized linear Bayes decision belonging to class $\bar{\mathcal{D}}_{n\infty}$ iff*

$$(3.19) \qquad 1 < a < \frac{n}{n-1}, \quad \text{if } n > 1, \qquad \text{and} \qquad 1 < a, \quad \text{if } n = 1.$$

In this case, the prior density function is the improper Pareto distribution

$$(3.20) \qquad f(\theta) \propto \left(\frac{1}{\theta}\right)^{\beta+1}, \quad \text{for all } \theta > 0 \quad \text{with} \quad \beta = \frac{a}{a-1} - n.$$

Proof: Immediate from Theorem 2.3. ∎

Note that if the improper Pareto measure defined in (3.20) is added to the classic Pareto family (see DeGroot (1970)), the *natural* conjugate family for the uniform distribution is complete. Indeed the improper Pareto functions provide the most of the well known estimators of θ.

Corollary 3.2 *Let $X_1, \ldots, X_n$ be a random sample from a uniform distribution over $(0, \theta)$. For all $b > 0$, $\hat{\theta}_n(y) = y + b$ is a Bayes decision belonging to class $\bar{\mathcal{D}}_n$ related to* *the prior density function*

$$f(\theta) \propto \theta^n e^{-\frac{\theta}{b}}, \qquad \text{for all } \theta > 0. \tag{3.21}$$

Now, $\hat{\theta}_n(y) = ay + b$, $a > 1$, $b > 0$ *is a Bayes decision belonging to class* $\bar{\mathcal{D}}_n$, *if and only if, (3.19) is fulfilled. In this case, the prior density function is:*

$$f(\theta) \propto \frac{\theta^n}{(\theta+\alpha)^{\beta+1}}, \qquad \text{for all } \theta > 0, \quad \textit{with} \quad \beta = \frac{a}{a-1} \quad \alpha = \frac{b}{a-1}. \tag{3.22}$$

Proof: In both cases, the proof is based on on verifying the conditions of Theorem 2.2. Therefore, after straightforward manipulations, the p.d.f.s given in (3.21) and (3.22) are easily derived as proper prior distributions. ■

Note that (3.21) is the p.d.f. of a Gamma distribution for $n \geq 1$. Observe also that Bayes estimators for each sample size are not possible using (3.22) as a prior density. Hence, the idea of introducing the notion of *strong sequences.* Thus, our next objective is to look for Bayes or generalized Bayes sequences with simple analytic form generated by a unique prior distribution. However, the following result shows that such simple families are not feasible for the uniform scale parameter in comparison with the exponential family, (see Diaconis (1979)), where linear Bayes sequences are certainly possible using the natural conjugate prior measures for the location parameter.

Theorem 3.1 *Let* $P_n(y)$ *be a sequence of* k*-order polynomial functions such that* $P_n(y) > y$ *for* $y > 0$. *Then,*

1. $\{P_n(y)\}_{n\in\mathbb{N}} \notin SBS$

2. $\{P_n(y)\}_{n\in \mathbb{N}} \in$ *SGBS, if and only if,* $k = 1$ *and*

$$P_n(y) = \frac{\alpha+n+1}{\alpha+n}y, \quad \textit{for some} \quad \alpha > 0. \tag{3.23}$$

Proof: If $k = 1$, $P_n(y) = a_n y$ with $a_n > 1$, then (3.23) is attained by applying condition (2.16). If $k > 1$, assume that $P_n(y) = a_{0n} + a_{1n}y + \cdots + a_{kn}y^k$. It is easy to ensure that the independent coeficient is zero, i.e, $a_{0n} = 0$. Then, taking limits on both sides of (2.16) adapted to the polynomial case, we have

$$\begin{aligned} \lim_{y\to\infty} y\xi_n(y) &= k+1, \quad \text{for all } n \\ \lim_{y\to\infty} y\xi_1(y) + n - 1 &= k + 1 + n - 1. \end{aligned}$$

From this contradiction, it is obvious that polynomial sequences are not feasible with this Bayesian approach. ∎

Remark 3.1 As one of the referees pointed out, Theorem 3.1 can be proved in terms of Theorem 2.4. It follows that if the order of the polynomials is k then $h_r(y)$ if of the order $y^{(k-1)r+1}$ for $y \to \infty$. Consequently only $k = 1$ is possible. Taking $P_n(y) = a_n y + b_n$, and then taking $y \to 0$ it follows easily that $b_n = 0$ for any n. Now $P_n(y) = a_n y$ iff

$$\frac{\prod_{k=1}^{r} a_k - 1}{r} = a_1 - 1$$

and hence (3.23) is derived.

Most estimators used in the problem of estimating the scale parameter for the uniform distribution display the form described in (3.23). For example, remember that the *minimum risk equivariant* (MRE) estimator of the scale parameter for the quadratic error loss function is

$$\hat{\theta}_n^{(1)}(y) = \frac{n+2}{n+1}y, \qquad \text{for all } y > 0.$$

Observe that the MRE estimator of θ is a generalized strong–Bayes sequence for $\alpha = 2$. Setting $\alpha = 1$, the unbiased estimator is also obtained as a SGBS. The consequence of this fact is that usual estimators have this *good* property in relation to the new notion of Bayes estimation introduced here.

To sum up, as far as the linear estimation of the uniform scale parameter is concerned, we have proved that BE of the form ay are not possible but GBE are feasible. If $a > 1$ and $b > 0$, we have shown that BE are possible using prior distributions which do not generate a strong–Bayes sequence. Moreover, note that the prior distribution described in (3.22) only implies linear BE when the sample size is n, and the prior distribution described in (3.21) generates an SBS whose structure is not linear. In view of the fact that it is difficult to obtain strong Bayes estimators with simple structure for the uniform distribution, it is reasonable that the MRE estimator is the most used estimator.

4 FINAL NOTES

Remark 4.1 Consider now the error loss function typically used for scale parameter, i.e.

$$L(a,\theta) = \left(\frac{a-\theta}{\theta}\right)^2. \qquad (4.24)$$

The problem of characterizing BE or GBE generated for prior measures of $\mathcal{D}$ or $\mathcal{D}_\infty$ and under this loss function is also solved from results shown here. In this case the form of estimator is

$$\hat{\theta}_n(y) = \frac{\int_y^\infty \theta^{-(n+1)} f(\theta) d\theta}{\int_y^\infty \theta^{-(n+2)} f(\theta) d\theta}$$

Hence, observe that an estimator is Bayes or generalized Bayes for θ under the error loss function (4.24), if and only if, it belongs to class $\bar{\mathcal{D}}_{n+2}$ or $\bar{\mathcal{D}}_{n+2\infty}$, respectively.

Remark 4.2 It is well known that Bayesian estimation of the uniform scale parameter is usually linked with the Pareto family. Wesolowski and Ahsanullah (1995) considered Pareto density class of the form

$$f(\theta) = \frac{\beta\gamma^\beta}{(\gamma+\theta)^{\beta+1}}, \quad \theta > 0, \quad \beta, \gamma > 0. \tag{4.25}$$

Note that (4.25) belongs to class $\mathcal{D}$ since its support is $(0, \infty)$. In this paper, we have proved that the MRE estimator under squared error loss is a linear GBE derived from an improper measure described in (3.20). Observe that as $\gamma \to 0$, the proper density of (4.25) converges to the improper density of (3.20). Hence, we propose the use of these priors as a solution to the Bayesian estimation problem of the uniform scale parameter.

Remark 4.3 David (1981) and Lehmann (1983) describe other inference problems for the scale parameter of the uniform distribution with some additional restrictions on the parameter θ, such as

$$0 < \theta \le \theta_0, \quad \theta_0 \le \theta, \quad \theta_1 \le \theta \le \theta_2.$$

By methods similar to those developed here, the characterization of Bayes rules under restrictions on the scale parameter can be solved by considering priors with supports over these new parameter spaces.

ACKNOWLEDGEMENTS

We are indebted to two anonymous referees for their useful suggestions which led to a substantial improvement of the material.

BIBLIOGRAPHY

Cacoullos, T. and Papageorgiou, H. (1983). Characterizations of discrete distributions by a conditional distribution and regression function. *Ann. Ins. Statist. Math.*, **35**, 95–103.

David, H. A. (1981). *Order Statistics.* Wiley, New York.

Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential family. *Ann. Statist.*, **7** (2), 269–281.

DeGroot, M. H. (1970). *Optimal Statistical Decisions.* Mc–Graw Hill, New York.

Goldstein, M. (1975). Uniqueness relation for linear posterior expectations. *J. Roy. Statist. Soc. Series B*, **31**, 402–405.

Guess, F. and Proschan, F. (1988). Mean Residual Life: Theory and Applications. *Handbook of Statistics, Vol. 7*, (P. R. Krishnaiah and C. R. Rao, eds.), 215–224, Elsevier.

Gupta, A.K. and Wesolowski, J. (1997). Uniform mixtures via posterior means. *Ann. Inst. Statist. Math.*, **49(1)**, 171–180.

Johnson, W. L. (1967). Note on a uniqueness relation in certain accident proneness models. *J. Amer. Statist. Assoc.*, **62**, 288–289.

Kahn, P. M. (1975). *Credibility Theory and Application.* Academic Press, New York.

Kagan, A. M., Linnik, Yu. V. and Rao, C. R. (1973). *Characterization Problems in Mathematical Statistics.* Wiley, New York.

Lillo, R. E. and Martín, M. (1999). Characterization of Bayes estimators of the uniform scale parameter. *Statistics & Decisions*, **17**, 31–47.

Lehmann, E. L. (1983). *The Theory of Point Estimation.* Wiley, New York.

Papageorgiou, H. and Wesolowski, J. (1997). Posterior mean identifies the prior distribution in NB and related models. *Statist. Probab. Lett.*, **36**, 127–134.

Sapatinas, T. (1995). Identifiability of mixtures of power–series distributions and related characterizations. *Ann. Inst. Statist. Math.*, **47(3)**, 447–459.

Wesolowski, J. and Ahsanullah, M. (1995). Conditional specifications of the multivariate Pareto and Student distributions. *Commun. Statist–Theory Meth.*, **24**, 1023–1031.

Wesolowski, J. (1995a). Bivariate distributions via a Pareto conditional distribution and a regression function. *Ann. Inst. Statist. Math.*, **47(1)**, 177–183.

Wesolowski, J. (1995b). Bivariate discrete measures via a power series conditional distribution and a regression function. *J. Multivariate Anal.*, **55**, 219–229.

Zoroa, P., Ruiz, J. M. and Marín, J. (1990). A characterization based on conditional expectations. *Commun. Statist–Theory Meth.*, **19(8)**, 3127–3135.