Sankhyā : The Indian Journal of Statistics
1992, Volume **54**, Series A, Pt. 3, pp. 297–308.

EFFICIENT ESTIMATION WITH MANY NUISANCE PARAMETERS

(Part III)

By J. BHANJA and J. K. GHOSH

Indian Statistical Institute

SUMMARY. Part III contains the constructions of efficient estimates in a mixture model where we have in addition direct observations on G. Our method provides a more transparent solution than that of Ibragimov and Hasminskii, who first posed and solved this problem.

6*. Mixture models with observations on G

So far, we had only a single set of observations, which is used for estimating both θ and G. The question now arises what happens if another (independent) set of observations giving information only on G is also available ? Will the problem become simpler ? Hasminskii and Ibragimov (1983, § 3) has provided a positive answer to this question. In the following discussion we shall derive their results using the methods of Section 3 instead of the original method due to Hasminskii and Ibragimov (1983). The assumptions of direct observations on G allow a verification of the smoothness conditions of the optimal kernel. Less important, but still useful, is the fact that we also have a uniformly $\sqrt{n}$-consistent estimate of G so that the splitting technique can be avoided and the identifiability assumption becomes much weaker.

This problem is taken up mainly as a technically interesting case where the required smoothness of the optimal kernel can be demonstrated. However, it may also have some practical application as indicated in the following scenario.

Example 6.1. Suppose there is a source sending a signal θ over time. The signal as it comes out of the source at time t is distorted to

$$Z_t = \theta + \epsilon_t$$

where ϵ_t is the noise associated with the source, ϵ_t's are, say, i.i.d. $\mathcal{N}(0, \sigma^2)$, σ^2 known. As Z_t passes through a channel there is a further distortion ξ_t leading to the observation of

$$W_t = \theta + \epsilon_t + \xi_t.$$

Paper received. June 1988 ; revised July 1990.

AMS (1985) *subject classification*. 62G05, 62G20.

Key words and phrases. Estimation, semiparametric models, mixture models, fixed set-up models, asymptotic efficiency.

*Sections 1–5 appeared in February and June 1992 issues of Sankhyā.

ξ_t's are assumed to be i.i.d. with common distribution G. Clearly ξ_t's are not observable at the time a signal is being sent, since they are confounded with the signal. But the distribution G can be estimated by observing $W_t \equiv \xi_t$ when a signal of magnitude zero is being sent from another controlled source, i.e. when $Z_t \equiv 0$. Suppose that two independent sets of observations are recorded, the first one $(X_1, X_2, \ldots, X_n, \ldots)$ being recorded at time instances when a signal is being sent from the original source and the other one $(Y_1, Y_2, \ldots Y_n, \ldots)$ when sent from the other one.

More generally, one assumes given a signal θ and an "uncontrollable" noise ξ, the channel produces response X according to the density function $f(., \theta, \xi)$. Along with observations on X, one has observations on the distribution of ξ. Let us now write down the model explicitly.

Model III. In this model, observations $(\boldsymbol{X}_i, \boldsymbol{Y}_i)$ are i.i.d. random vectors in $(S_1, \mathcal{S}_1)^p \times (S_2, \mathcal{S}_2)^q (= (\boldsymbol{S}, \boldsymbol{\mathcal{S}}))$ with common distribution

$$P^{III}_{\theta_0, G_0} := \prod_{j=1}^{p} f(x_j, \theta_0, G_0) d\mu(x_j) \prod_{k=1}^{q} dG_0(y_k) \quad \text{for all } (\boldsymbol{x}, \boldsymbol{y}) \in S_1^p \times S_2^q$$
$$\text{for some } (\theta_0, G_0) \in \Theta \times \mathcal{G}$$

where $(S_1, \mathcal{S}_1)$ is an arbitrary measurable space, S_2 is a compact metric space, $\mathcal{S}_2 = \mathcal{B}(S_2)$, Θ is an open subset of $\boldsymbol{R}$ with compact closure $\bar{\Theta}$ and $\mathcal{G}$ is the set of all probability measures on $(S_2, \mathcal{S}_2)$. [As in Section 1, let us also assume that the probability measures are well-defined on $\bar{\Theta} \times \mathcal{G}$]

Note that

(1) Definitions 2.1—2.5 have obvious Model III-analogues which is obtained by replacing X_i's by $(\boldsymbol{X}_i, \boldsymbol{Y}_i)$'s and P_{θ_0, G_0} by $P^{III}_{\theta_0 G_0}$ in the relevant definition for Model II.

As in Section 2, we shall abbreviate the phrase 'in Model III' by '(III)'.

Notation. We shall denote the set of all Model III-kernels by $\boldsymbol{K}$.

We shall need the following definition.

Definition 6.1. A function $\psi : S_1 \times \bar{\Theta} \times \mathcal{G} \to \boldsymbol{R}$ $(S_2 \times \bar{\Theta} \times \mathcal{G} \to \boldsymbol{R})$ is called an $S_1(S_2)$-*kernel* if $\psi(., \theta, G) \in L_2^0(f(., \theta, G))$ $(L_2^0(G))$ for all (θ, G) in $\bar{\Theta} \times \mathcal{G}$ and the set of all $S_1 (S_2)$-kernels is denoted by $K_1(K_2)$.

Given an S_1-*kernel* ψ_1 and an S_2-kernel ψ_2, let us define a function Q_{ψ_1, ψ_2} from $\boldsymbol{S} \times \bar{\Theta} \times \mathcal{G}$ to $\boldsymbol{R}$ by

$$Q_{\psi_1, \psi_2}((\boldsymbol{x}, \boldsymbol{y}), \theta, G) := \sum_{j=1}^{p} \psi_1(x_j, \theta, G) + \sum_{k=1}^{q} \psi_2(y_k, \theta, G)$$
$$\text{for } ((\boldsymbol{x}, \boldsymbol{y}), \theta, G) \in \boldsymbol{S} \times \bar{\Theta} \times \mathcal{G} \qquad \ldots \quad (6.1)$$

Note that

(2) Relation (6.1) defines a bounded linear map from $K_1 \times K_2$ to $\boldsymbol{K}$.

As in relation (2.2), let us define the θ-score S_θ for Model III by

$$S_\theta((\boldsymbol{x}, \boldsymbol{y}), \theta, G) := \sum_{j=1}^{p} \frac{f'(x_j, \theta, G)}{f(x_j, \theta, G)}$$

$$= \sum_{j=1}^{p} s_\theta(x_j, \theta, G) \text{ for } ((\boldsymbol{x}, \boldsymbol{y}), \theta, G) \in \boldsymbol{S} \times \bar{\Theta} \times \mathcal{G} \qquad \dots \quad (6.2)$$

(3) $S_\theta \in \boldsymbol{K}$ if and only if $s_\theta \in K_1$ so that S_θ is well-defined under assumption (A3). Also for $p = 1$ and $q = 0$, $S_\theta \equiv s_\theta$.

Before proceeding further let us make the following convention and definition.

Convention : For any G in $\mathcal{G}$ and ϕ in $L_2^0(G)$, we shall denote the function $\int f(., ., \xi)\, \phi(\xi)\, dG(\xi)$ by $f(., ., \phi\, dG)$.

Definition 6.2. For any (θ, G) in $\bar{\Theta} \times \mathcal{G}$ and ϕ in $L_2^0(G)$, define $A_{\theta,G}(\phi)$ from S_1 to $\boldsymbol{R}$ by

$$A_{\theta,G}(\phi)(x) := \frac{f(x, \theta, \phi\, dG)}{f(x, \theta, G)} \;\forall\; x \in S_1$$

and for any ϕ in K_2, define $A(\phi)$ from $S_1 \times \bar{\Theta} \times \mathcal{G}$ to $\boldsymbol{R}$ by

$$A(\phi)(x, \theta, G) := A_{\theta,G}(\phi(., \theta, G))(x) \;\forall\; (x, \theta, G) \in S_1 \times \bar{\Theta} \times \mathcal{G}.$$

Then

(4) For any (θ, G) in $\bar{\Theta} \times \mathcal{G}$, $A_{\theta,G}$ is a bounded linear map from $L_2^0(G)$ to $L_2^0(f(., \theta, G))$ and A is a bounded linear map from K_2 to K_1.

For any (θ, G) in $\bar{\Theta} \times \mathcal{G}$, define

$$\boldsymbol{N}_{\theta,G} := \{\psi \in L_2^0(P_{\theta,G}^{III}) : \exists\, \phi \in L_2^0(G)$$

with

$$\psi(\boldsymbol{x}, \boldsymbol{y}) := \sum_{j=1}^{p} A_{\theta,G}(\phi)(x_j) + \sum_{k=1}^{q} \phi(y_k) \;\forall (\boldsymbol{x}, \boldsymbol{y})\}. \qquad \dots \quad (6.3)$$

(5) The elements of the space $\boldsymbol{N}_{\theta,G}$ may be thought of us 'directional scores' with respect to small variations in G. However, for the special case where $p = 1$ and $q = 0$, $\boldsymbol{N}_{\theta,G}$ is a proper subset of $N_{\theta,G}$ unless G has finite support so that (6.3) falls short of an analogue of (2.3).

As in Section 2, one can define an *optimal kernel* $\bar{\Psi}$ and the information I^{III} by the following analogue of relation (2.4).

$$\left.\begin{aligned} \bar{\Psi}(\,.\,, \theta, G) &:= \mathrm{Proj}_{N^{\perp}_{\theta,G}} (S_\theta(\,.\,, \theta, G)) \\ I^{III}\ (\theta, G) &:= \|\bar{\Psi}(\,.\,, \theta, G)\|^2_{L_2(P^{III}_{\theta,G})} \end{aligned}\right\} \forall(\theta, G) \qquad \dots \quad (6.4)$$

In order to get a simpler formula for evaluating $\bar{\Psi}$, we shall need the following definitions and lemma.

Definition 6.3. For any (θ, G) in $\bar{\Theta}\times\mathcal{G}$, the closed linear space in $L^0_2\,(P^{III}_{\theta,G})$ spanned by $S_\theta(\,.\,, \theta, G)$ and $N_{\theta,G}$ will be called *the tangent space at* (θ, G) and will henceforth be denoted by $T^{III}_{\theta,G}$.

Remark 6.1. Our tangent space coincides with that considered in Hasminskii and Ibragimov (1983, § 3).

Definition 6.4. Call a kernel Ψ a *tangent-space-kernel* if $\Psi(., \theta, G) \in T^{III}_{\theta,G}$ for any $(\theta, G) \in \bar{\Theta}\times\mathcal{G}$.

Observe that

(6) $T^{III}_{\theta,G}$ consists of functions of the form $\sum_{j=1}^{p} \phi_1(x_j) + \sum_{k=1}^{q} \phi_2(y_k)$ for some $\phi_1 \in L^0_2(f(., \theta, G))$ and $\phi_2 \in L^0_2(G)$ so that any tangent-space-kernel Ψ must be of the form Q_{ψ_1,ψ_2} for some $\psi_1 \in K_1$, $\psi_2 \in K_2$.

(7) A tangent-space-kernel Q_{ψ_1,ψ_2} is an optimal kernel if and only if

$$p(\textstyle\int \psi_1(., \theta, G)\, A_{\theta,G}(\phi)(.)\, f(., \theta, G) d\mu(.)) + q(\int \psi_2(., \theta, G)\phi(.)\, dG(.)) = 0 \quad \dots \quad (6.5)$$

for all (θ, G, ϕ) with $\phi \in L^0_2(G)$.

In Lemma 6.1(a) (vide relation (6.6)), we give a simpler sufficient condition for a tangent-space-kernel to be optimal. Lemma 6.1(b) gives a sort of converse which is a Model III-analogue of Lemma 2.1. Then in Lemmas 6.2—6.4 we show a smooth solution of (6.6) exists.

Lemma 6.1. (a) *If for some tangent-space-kernel* Q_{ψ_1,ψ_2},

$$p \textstyle\int \psi_1(., \theta, G) f(., \theta, G')\, d\mu\,(.) + q \int \psi_2(., \theta, G)\, dG'\,(.) = 0 \ \forall(\theta, G, G') \in \bar{\Theta}\times\mathcal{G}\times\mathcal{G} \quad \dots \quad (6.6)$$

then Q_{ψ_1,ψ_2} *is a version of the optimal kernel.*

(b) *Conversely, if* Q_{ψ_1,ψ_2} *is a version of the optimal kernel defined through* (6.2)—(6.4) *and for all* θ *the L.H.S. of* (6.6) *is continuous in* (G, G'), *then* (6.6) *holds for it.*

Proof. (a) In view of observations (4), (7) and the fact that the set of all bounded functions ϕ lying in $L_2^0\,(G)$ is dense in it, it is enough to show (6.5) for bounded ϕ's only.

Now, let ϕ be a bounded function in $L_2^0\,(G)$, then there is $\epsilon > 0$ such that $1+\eta\,\phi\,(y) \geqslant 0$ for all y in S_2 whenever $|\eta| < \epsilon$. Fix one such ϵ. Define the curve $G. : (-\epsilon, \epsilon) \rightarrow \mathcal{G}$ by

$$dG_\eta\,(.) = \{1+\eta\phi(.)\}\,dG\,(.) \text{ for all } \eta \in (-\epsilon, \epsilon).$$

The relation (6.6) with $G' = G_\eta$ implies relation (6.5) for ϕ. Since ϕ was arbitrary, this proves (a).

(b) An easy modification of the proof of Lemma 2.1 (b).

Let ψ_i be an S_i-kernel, $i = 1, 2$. Consider the following Model III-analogue of equation (3.1).

$$\sum_{\substack{i=1\\ i\ odd}}^{n} Q_{\psi_1,\,\psi_2}((\boldsymbol{X}_i, \boldsymbol{Y}_i), \theta, \hat{\hat{G}}_n^E) + \sum_{\substack{i=1\\ i\ even}}^{n} Q_{\psi_1,\,\psi_2}((\boldsymbol{X}_i, \boldsymbol{Y}_i), \theta, \hat{\hat{G}}_n^O) = 0 \qquad \ldots \quad (6.7)$$

where $\hat{\hat{G}}_n := F_{nq}\,(Y_{11}, \ldots, Y_{1q}, Y_{21}, \ldots, Y_{2q}, \ldots, Y_{n1}, \ldots, Y_{nq})$ and the suffixes E and O stand for the operations based on even and odd indices, respectively, as defined in (11) of Section 2.

(8) By Lemma C.1 of Appendix C, $\hat{\hat{G}}_n$ is a uniformly consistent (III) estimate of G_0 (vide Definition 2.1 and observation (1)).

Assume that

(E1) There is a uniformly $\sqrt{n}$-consistent (III) estimate U_n of θ_0 (vide Definition 2.2 and observation (1)).

Let $T_n(Q_{\psi_1, \psi_2})$ be the estimate defined through Definition 3.1 and observation (1). We are now going to give regularity conditions on f, ψ_1 and ψ_2 guaranteeing uniform asymptotic normality (III) (vide Definition 2.4 and observation (1)) of $T_n(Q_{\psi_1, \psi_2})$.

Fix (θ_0, G_0) in $\Theta \times \mathcal{G}$. Let δ_0 denote a positive real-number which will be chosen later. The following are the required regularity conditions.

U[i] Condition U(i) of Section 3 holds.

For any condition U(C) among U(ii), U(iv)-U(vi) of Section 3, U[C] denotes the condition that U(C) holds with ψ replaced by ψ_1 or the relevant parts of it hold with (ψ, f, μ) replaced by (ψ_2, u, G) where u denotes the function identically equal to one. The condition U[iii] is given below.

U[iii] For any $c > 0$ and $\varepsilon > 0$, the supremum over $\theta \epsilon B(\theta_0, \delta_0)$, $G \epsilon B(G_0, \delta_0)$ and $| \theta' - \theta | \leqslant c/\sqrt{n}$, of $(P^{III}_{\theta, G})^n$ $(\{ | \sqrt{n} p \int \psi_1 (., \theta, \hat{\hat{G}}_n) f(., \theta, G) \, d\mu(.) + \sqrt{n} \, q \int \psi_2(., \theta, \hat{\hat{G}}_n) dG(.) | > \epsilon \})$ tends to zero as $n \to \infty$.

The following is the Model III-analogue of Lemma 3.1a.

Lemma 6.2. *Assume* (*E*1). *If* f *satisfies condition* $U[i]$ *and* (ψ_1, ψ_2) *satisfies conditions* $U[ii] - U[vi]$, *then* $T_n(Q_{\psi_1, \psi_2})$ *is a uniformly* $\sqrt{n}$*-consistent solution* (*III*) (*vide Definition* 2.3 *and observation* (1)) *of* (6.7) *as well as UAN* (*III*) *wiih AV* $V(., ., \psi_1, \psi_2)$ *mhere*

$$V(\theta, G, \psi_1, \psi_2) := \frac{p \| \psi_1(., \theta, G) \|^2_{L_2(P_{\theta, G})} + q \| \psi_2(., \theta, G) \|^2_{L_2(G)}}{p < \psi_1(., \theta, G), s_\theta(., \theta, G) >^2_{L_2(P_{\theta, G})}} \; \forall \, (\theta, G).$$

Proof. An easy modification of the proof of the parts U(II)(B) and U(III)(B) of Lemma 3.1 proves the result.

From now on, we shall assume that q is positive.
Note that

(9) If there is ϕ in K_2 such that $Q_{s_\theta + A(\phi), \phi}$ satisfies relation (6.6), then

$$I^{III}(\theta, G) = p \| s_\theta(., \theta, G) + A(\phi)(., \theta, G) \|^2_{L_2(f(., \theta, G))} + q \| \phi(., \theta, G) \|^2_{L_2(G)} \forall (\theta, G).$$

Hence $I^{III}(\theta, G) > 0$ if and only if

(E0) $0 < \| s_\theta(., \theta, G) \|^2_{L_2(f(., \theta, G))} < \infty \; \forall \, (\theta, G)$.

(10) Assumption (D1) implies assumption (E0) and condition U[i].

Therefore, in view of observation (9), conditions U[ii]—U[vi] hold for $\psi_1 = s_\theta + A(\phi)$ and $\psi_2 = \phi$ provided there is ϕ belonging to $C_{0,1,0}(S_2 \times \bar{\Theta} \times \mathcal{G})$ such that Q_{ψ_1, ψ_2} satisfies relation (6.6).

Remark 6.2. In view of observations (9)—(10), it remains to prove the existence of ϕ lying in $C_{0,1,0}(S_2 \times \bar{\Theta} \times \mathcal{G})$ such that $Q_{s_\theta + A(\phi), \phi}$ satisfies the relation (6.6).

Let us now observe that, for any ψ_1 in K_1 and ψ_2 in K_2, relation (6.6) is equivalent to

$$p \int \psi_1(., G) f(., \theta, y) \, d\mu(.) + q \psi_2(y, \theta, G) = 0 \text{ for all } (\theta, G, y) \, \epsilon \, \bar{\Theta} \times \mathcal{G} \times S_2 \quad \ldots \quad (6.8)$$

Therefore, $Q_{s_\theta+A(\phi),\phi}$ satisfies relation (6.6) if and only if

$$p \int s_\theta(.,\theta,G)f(.,\theta,y)\,d\mu(.) = -p\int A(\phi)(.,\theta,G)f(.,\theta,y)\,d\mu(.) \quad \ldots \quad (6.9)$$
$$-q\phi(y,\theta,G) \quad \forall(\theta,G,y)$$

But for any $(\theta, G, y) \in \bar{\Theta}\times\mathcal{G}\times S_2$,

$$\int A(\phi)(.,\theta,G)f(.,\theta,y)\,d\mu(.) = \int \frac{\int\phi(y',\theta,G)f(x,\theta,y')\,dG(y')}{f(x,\theta,G)} f(x,\theta,y)\,d\mu(.)$$
$$= \int K(y,y',\theta,G)\,\phi(y',\theta,G)\,dG(y') \quad \ldots \quad (6.10)$$

where the function $K : S_2\times S_2\times\bar{\Theta}\times\mathcal{G}\to \boldsymbol{R}$ is defined by

$$K(y,y',\theta,G) := \int\frac{f(.,\theta,y)f(.,\theta,y')}{f(.,\theta,G)}\,d\mu(.) \text{ for } (y,y',\theta,G)\in S_2\times S_2\times\bar{\Theta}\times\mathcal{G}.$$

Assume that

(E2) (a) For any x in S_1, $f(x,.,.)\in C_{1,0}(\bar{\Theta}\times S_2)$

and (b) the following three families of functions

$$\left\{\frac{f(.,\theta,y)f(.,\theta,y')}{f(.,\theta,G)} : (y,y',\theta,G)\in S_2\times S_2\times\bar{\Theta}\times\mathcal{G}\right\},$$

$$\left\{\frac{f(.,\theta,y)f'(.,\theta,y')}{f(.,\theta,G)} : (y,y',\theta,G)\in S_2\times S_2\times\bar{\Theta}\times\mathcal{G}\right\}$$

and $\left\{\frac{f(.,\theta,y)f(.,\theta,y')f'(.,\theta,G)}{f(.,\theta,G)} : (y,y',\theta,G)\in S_2\times S_2\times\bar{\Theta}\times\mathcal{G}\right\}$

are uniformly integrable with respect to μ.

(11) Under assumption (E2), $K\in C_{0,0,1,0}(S_2\times S_2\times\bar{\Theta}\times\mathcal{G})$ and therefore the RHS of relation (6.10) defines a bounded operator B from CK_2 to CK_2 where CK_2 denotes the subspace $C_{0,1,0}(S_2\times\bar{\Theta}\times\mathcal{G})K_2$, of $C_{0,1,0}(S_2\times\bar{\Theta}\times\mathcal{G})$.

Define $\eta : S_2\times\bar{\Theta}\times\mathcal{G}\to\boldsymbol{R}$ by

$$\eta(y,\theta,G) := \int s_\theta(x,\theta,G)f(x,\theta,y)\,d\mu(x) \text{ for all } (y,\theta,G) \quad \ldots \quad (6.11)$$

(12) Under assumption (D1), η belongs to the set CK_2.

By (6.9)—(6.11) we see that for all (y,θ,G)

$$\eta(y,\theta,G) = -\left(B+\frac{q}{p}I\right)(\phi)(y,\theta,G)$$
$$= -C(\phi)(y,\theta,G) \quad \ldots \quad (6.12)$$

where C denotes the (bounded) operator $\left(B+\frac{q}{p}I\right)$ from CK_2 to itself.

In view of observations (9)—(12), Remark 6.2 and relations (6.10)—(6.12), it remains to prove that C is invertible. As a first step towards this goal let us show that

Lemma 6.3. *Under assumption* ($E2$), (a) C *is* 1-1 *and* (b) B *is a compact operator.*

Proof. (a) Let $\phi \in CK_2$ be such that $C\phi = 0$. Then,

$$\int \int \phi(y, \theta, G)\, K\,(y, y', \theta, G)\, \phi\,(y', \theta, G)\, dG(y)\, dG(y')$$

$$+\left(\frac{q}{p}\right) \|\phi\,(., \theta, G)\|^2_{L_2(G)} = 0 \quad \forall(\theta, G)$$

which, in turn, implies

$$\int \frac{\{\int\phi(y, \theta, G)f(x, \theta, y)dG(y)\}^2}{f(x, \theta, G)}\, d\mu(x)+\left(\frac{q}{p}\right) \|\phi(., \theta, G)\|^2_{L_2(G)} = 0 \quad \forall(\theta, G)$$

i.e., equivalently

$$\phi(., \theta, G) = 0 \quad \text{a.e. } [G] \quad \forall(\theta, G)$$

and hence

$$\phi \equiv 0 \text{ [Since } \phi \in CK_2]$$

(b) In the proofs available in standard literature, the measure G is fixed. For the sake of completeness a proof is given in Appendix D.

As a corollary to Lemma 6.3, one can show

Lemma 6.4. *Under assumptions* ($D1$) *and* ($E2$) *there is a function* $\tilde{\phi} \in CK_2$ *such that* $\psi_1 = s_\theta + \text{A}(\tilde{\phi})$ *and* $\psi_2 = \tilde{\phi}$ *together satisfy* (6.6) *and conditions* $U[ii]—U[vi]$.

Proof. In view of calculations done earlier it is enough to show that C is invertible.

Suppose not. Then using the fact that $q \neq 0$ and compactness of B conclude from Theorem 4.25(b) of Rudin (1974) that C is not 1—1, which contradicts Lemma 6.3(a).

In view of Lemmas 6.1—6.4 one can show

Theorem 6.5. *Assume* ($E1$), ($D1$) *and* ($E2$). *Let* $\tilde{\phi}$ *be as considered in Lemma* 6.4. *Then* $T_n(Q_{s_\theta, + A(\tilde{\phi}), \tilde{\phi}})$ *is UAN* (III) *with AV* ($1/1^{III}$).

Remark 6.3. It is worth pointing out that we do need the compactness of the operator B since it acts on a Banach space rather than a Hilbert space.

Note also that we make use of 'non-negative'-ness of B (vide Lemma 6.3) but the associated inner product will not give the norm of the Banach space.

Remark 6.4. We have verified assumptions (D1) and (E2) for the following two cases :

Case 1 (a) S is compact

(b) $f \in C_{0,2,0}(S \times \bar{\Theta} \times \mathcal{G})$

(c) $f(x, \theta, G) > 0 \;\forall\; (x, \theta, G)$

and (d) $P_{0,G}(\{|f'(., \theta, G)| > 0\}) > 0 \;\forall (\theta, G)$.

Case II We have Euclidean S and exponential f following assumptions (a) and (f) of Remark 3.4 and additional assumptions

$$(d)^* \sum_{j=1}^{k} (\pi_j')^2 (\theta, \xi) > 0 \;\forall\; (\theta, \xi)$$

and (e)* for any θ in Θ, both $2\pi_j(\{\theta\} \times S_2) - \bar{\pi}_j(\{\theta\} \times S_2)$ and $2\bar{\pi}_j(\{\theta\} \times S_2) - \pi_j(\{\theta\} \times S_2)$ belong to the interior of Ω.

In particular, Exampe 6.1 falls in Case II of the above remark.

Remark 6.5. To solve (6.7), we have to determine for various values of θ, $\tilde{\phi}(Y, \theta, \hat{\hat{G}}_n^O)$ and $\tilde{\phi}(Y, \theta, \hat{\hat{G}}_n^E)$ evaluated at all the observation Y_{ij}'s on Y.

Now, one can easily observe that for G's with finite support and y's restricted to the support of G, equation (6.12) can be rewritten in the form $A\boldsymbol{x} = \boldsymbol{y}$ where the matrix A is positive definite. Thus we have a unique solution for $\tilde{\phi}(Y_{ij}, \theta, \hat{\hat{G}}_n^O)$ for odd i's and for $\tilde{\phi}(Y_{ij}, \theta, \hat{\hat{G}}_n^E)$ for even i's.

To evaluate $\{\tilde{\phi}(Y_{ij}, \theta, \hat{\hat{G}}_n^O), 1 \leqslant j \leqslant q, i \text{ even}\}$ and $\{\tilde{\phi}(Y_{ij}, \theta, \hat{\hat{G}}_n^E), 1 \leqslant j \leqslant q, i \text{ odd}\}$, define $\hat{G}_{\varepsilon 1} := (1-\epsilon)\hat{\hat{G}}_n^O + \epsilon\, \hat{\hat{G}}_n^E$ and $\hat{G}_{\varepsilon 2} := \epsilon\, \hat{\hat{G}}_n^O + (1-\epsilon)\hat{\hat{G}}_n^E$ and solve the linear version of (6.12) with $G = \hat{G}_{\varepsilon 1}$ and $\hat{G}_{\varepsilon 2}$ for $\{\phi(Y_{ij}, \theta, \hat{G}_{\varepsilon 1}), 1 \leqslant j \leqslant q, i \text{ even}\}$ and $\{\hat{\phi}(Y_{ij}, \theta, \hat{G}_{\varepsilon 2}), 1 \leqslant j \leqslant q, i \text{ odd}\}$. Again the solutions are unique. Now let ϵ tend to zero.

In actual practice one would solve

$$\sum_{\substack{i=1 \\ i\, odd}}^{n} [\sum_{j=1}^{p} \{s_\theta(X_{ij}, \theta, \hat{\hat{G}}_n^E) + A(\tilde{\phi})(X_{ij}, \theta, \hat{\hat{G}}_n^E)\} + \sum_{k=1}^{q} \tilde{\phi}(Y_{ik}, \theta, \hat{G}_{\varepsilon 2})]$$

$$\sum_{\substack{i=1 \\ i\, even}}^{n} \left[\sum_{j=1}^{p} \{s_\theta(X_{ij}, \theta, \hat{G}_n^O) + A(\tilde{\phi})\, (X_{ij}, \theta, \hat{\hat{G}}_n^O)\} + \sum_{k=1}^{q} \tilde{\phi}(Y_{ik}\; \theta, \hat{G}_{\varepsilon 1}) \right] = 0$$

for various values of ϵ and stop when two consecutive values differ insignificantly.

Appendix D

Proof of Lemma 6.2(b). Want to show that $\overline{AS}$ is compact in $\| \|_{0,1,0}$-norm where

$$AS := \{A(\phi) : \phi \in S(C_{0,1,0}(S_2 \times \bar{\Theta} \times \mathcal{G}))\}$$

and $S(C_{0,1,0}(S_2 \times \bar{\Theta} \times \mathcal{G})) := \{\phi \in C_{0,1,0}(S_2 \times \bar{\Theta} \times \mathcal{G}) : \|\phi\|_{0,1,0} = 1\}$

Note that $\|\phi\|_{0,1,0} = \|\phi\|_{sup} + \left\| \frac{\partial}{\partial \theta} \phi \right\|_{sup}$ for $\phi \in C_{0,1,0}(S_2 \times \bar{\Theta} \times \mathcal{G})$.

Also note that $\overline{AS}$ is compact iff AS and $(AS)'$ are uniformly bounded and equicontinuous where

$$(AS)' := \left\{\frac{\partial}{\partial \theta}(A(\phi)) : \phi \in S(C_{0,1,0}(S_2 \times \bar{\Theta} \times \mathcal{G}))\right\}.$$

We shall only show that AS is uniformly bounded and equicontinuous. One can prove this fact for $(AS)'$ in a similar way.

Now let us observe that

$$\|A\phi\|_{sup} \leqslant \|K\|_{sup} \|\phi\|_{sup} \leqslant \|K\|_{sup} \|\phi\|_{0,1,0} \ \forall \ \phi \in C_{0,1,0}(S_2 \times \bar{\Theta} \times \mathcal{G}). \quad \ldots \quad (D.1)$$

Therefore AS is uniformly bounded.

Let us now fix $(y_0, \theta_0, G_0) \in S_2 \times \bar{\Theta} \times \mathcal{G}$, then

$$\begin{aligned}
&|A(\phi)(y, \theta, G) - A(\phi)(y_0, \theta_0, G_0)| \\
&= |\int K(y, y', \theta, G)\phi(y', \theta, G)\, dG(y') - \int K(y_0, y', \theta_0, G_0)\phi(y', \theta_0, G_0) dG_0(y')| \\
&\leqslant |\int \{K(y, y', \theta, G) - K(y_0, y', \theta_0, G_0)\}\phi(y', \theta, G) dG(y')| \\
&+ |\int K(y_0, y', \theta_0, G_0)\{\phi(y', \theta, G) - \phi(y', \theta_0, G_0)\}\, dG(y')| \\
&+ |\int K(y_0, y', \theta_0, G_0)\,\phi(y', \theta_0, G_0)\, d(G - G_0)(y')| \leqslant \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \quad \ldots \quad (D.2)
\end{aligned}$$

if $|\theta - \theta_0|$, $d(G, G_0)$ and $\rho(y, y_0) < \delta$ where ρ is some metric inducing topology on S_2 and $0 < \delta < \epsilon/(3\|K\|_{sup})$ is chosen in such a way that for any pair $(y, y', \theta, G), (\bar{y}, \bar{y}', \bar{\theta}, \bar{G})$; $|\theta - \bar{\theta}|$, $d(G, \bar{G})$, $\rho(y, \bar{y})$ and $\rho(y', \bar{y}') < \delta$ imply

$$|\phi(y, \theta, G) - \phi(\bar{y}, \bar{\theta}, \bar{G})| < \epsilon/(3\|K\|_{sup})$$

and

$$|K(y, y', \theta, G) - K(\bar{y}, \bar{y}', \bar{\theta}, \bar{G})| < \epsilon/3.$$

From (D.1) and (D.2) it follows that AS is uniformly bounded and equicontinuous.

Acknowledgement. The authors are grateful to Professor S. C. Bagchi, Dr. T. Samanta for many helpful discussions. They are indebted to Dr. G. Mishra for his help on the proof of compactness of the operator B (vide Lemma 6.1(b)). They would also like to thank the referees and the Co-Editor for many improvements in the presentation and pointing out an error in the original draft.

References

Amari, S. and Kumon, M. (1985). Optimal estimation in the presence of infinitely many nuisance parameters—Geometry of estimating functions. *Tech. Report* no. METR 85-2, Univ. Tokyo, Japan.

Bhanja, J. and Ghosh, J. K. (1987). Inference about a parameter in the presence of a large number of nuisance parameters. *Bull. Inst. Math. Stat.*, **16**, 73.

Bickel, P. J. (1982). On adaptive estimation. *Ann. Statist.*, **10**, 647-671.

Bickel, P. J. and Klaassen, C. A. J. (1986). Empirical Bayes estimation in functional and structural models and uniformly adaptive estimation of location. *Adv. Appl. Math.*, **7**, 55-69.

Chatterjee, S. K. and Das, K. (1983). Estimation of variance component for one-way classification with heteroscedastic error. *Calcutta Statist. Assoc. Bull.*, **32**, 57-78.

Hannan, J. and Huang, J. S. (1972). A stability of symmetrization of product measures with few distinct factors. *Ann. Math. Statist.*, **43**, 308–319.

Hannan, J. F. and Robbins, H. (1955). Asymptotic solutions of the compound decision problem for two completely specified distributions. *Ann. Math. Statist.*, **26**, 37-51.

Hasminskii, R. Z. and Ibragimov, I. A. (1983). On asymptotic efficiency in the presence of an infinite dimensional nuisance parameter. Probability Theory and Mathematical Statistics (Fourth USSR-Japan Symposium Proceedings, 1982. Edited by Itô and Prohorov) : *Lecture Notes in Math.*, 1021, 195-229, Springer-Verlag, New York.

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

Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimators in the presence of infinitely many nuisance parameters. *Ann. Math. Statist.*, **27**, 887-906.

Kumon, M. and Amari, S. (1984). Estimation of a structural parameter in the presence of a large number of nuisance parameters. *Biometrika*, **71**, 445-459.

Lindsay, B. G. (1980). Nuisance parameters, mixture models, and the efficiency of partial likelihood estimators. *Phil. Trans. Roy. Soc. London.*, Ser A, **296**, 639-662.

——— (1982). Conditional score functions : some optimality results. *Biometrika*, **69**, 503-512.

Loéve, M. (1963). *Probability Theory*, 3rd Edition. D. van-Nostard, Princeton.

Maitra, A. and Rao, B. V. (1975). Selection theorems and the reduction principle. *Trans. Amer. Math. Soc.*, **20**, 57-66.

NEYMANN, J. and SCOTT, E. (1948). Consistent estimates based on partially consistent observations. *Econometrika*, **16**, 1-32.

PARTHASARATHY, T. (1972). Selection theorems and their applications. *Lecture Notes in Mathematics*, **263**, Springer Verlag, New York.

PFANZAGL, J. (1982) (with W. Wefelmeyer). Contributions to a general asymptotic statistical theory. *Lecture Notes in Statistics*, **13**, Springer Verlag, New York.

ROBBINS, H. (1964). The empirical Bayes approach to statistical decision problems. *Ann. Math. Statist.*, **35**, 1-20.

RUDIN, W. (1974). *Functional Analysis*, Tata-Mc-Graw-Hill Edition.

SCHICK, A. (1986). On asymptotically efficient estimation in semiparametric models. *Ann. Statist.*, **14**, 1139-1151.

SRIVASTAVA, S. M. (1982). *Lecture Notes in Descriptive Set Theory*, Indian Statistical Institute.

VAN DER VAART, A. W. (1987). Statistical estimation in large parameter spaces. *Thesis*, University of Leiden.

——— (1988). Estimating a real parameter in a class of semiparametric models. *Ann. Statist.*, **16**, 1450-1474.

THEORETICAL STATISTICS AND MATHEMATICS DIVISION
INDIAN STATISTICAL INSTITUTE
203 B. T. ROAD
CALCUTTA 700 035
INDIA.