Probab. Th. Rel. Fields 81, 159–184 (1989)

Probability Theory and Related Fields

Multinomial Approximations for Nonparametric Experiments which Minimize the Maximal Loss of Fisher Information *

S. Luckhaus [1] and W. Sauermann [2]

[1] Sonderforschungsbereich 123, Universität Heidelberg, Im Neuenheimer Feld 293, D-6900 Heidelberg 1, Federal Republic of Germany
[2] Gödecke AG, Biometrics Department, Mooswaldallee 1–9, D-7800 Freiburg, Federal Republic of Germany

Summary. Certain nonparametric product experiments $\mathscr{P}_n^n$ can asymptotically be approximated by multinomial experiments obtained by a finite interval partition of the sample space, the real line. For specific families $\mathscr{P}_n$ defined in terms of bounded Fisher information and monotone likelihood ratios with bounded derivatives we study the problem to calculate a partition which is optimal in the sense that it minimizes the maximal loss of Fisher information caused by the discretization. This leads to a minimax problem. By considering partitions of the sample space into k intervals which have equal probability under a density h and then letting $k \to \infty$ we obtain an expansion for the quantity "loss of Fisher information" which is of order k^{-2} under regularity conditions. The corresponding minimax problem for the first order term of this expansion is shown to be the unique solution of a free boundary problem.

1. Introduction

We shall study the problem of approximating certain nonparametric experiments (models) by finite dimensional (parametric) experiments. The dimension of their parameter spaces will have to increase with the desired degree of approximation. This provides a possibility to treat questions concerning the behaviour of statistical procedures by examining the approximating experiments for which the whole theory of parametric inference is available.

The nonparametric experiments considered here will be families of distributions on the real line. Several ways of approximating these experiments are possible, we choose to work with multinomial experiments derived by a partition of the sample space into intervals. The quality of approximation will be measured by the maximal loss of Fisher information.

By restricting to partitions generated by a continuous probability density it is possible to extend methods of Freedman, Diaconis (1981) and thus obtain

* This work has been supported by the Deutsche Forschungsgemeinschaft

an analytic expression for the quantity "loss of Fisher information". This further enables one to solve the minimax problem mentioned above. Although this is only possible for very special nonparametric experiments, the methods developed here should be of some independent interest, since approximations of this type are common in other areas of statistics.

We shall start by describing more closely the asymptotic framework. Suppose that for $n=1,2,\dots$ we are given a family $\mathscr{P}_n$ of probability measures on the real line endowed with the field of Borel sets and that each $\mathscr{P}_n$ contains a fixed nulldistribution P_0. Conditions on the families $\mathscr{P}_n$ under which the product experiments $\mathscr{P}_n^n=(P^n\colon P\in\mathscr{P}_n)$ can be approximated (in the sense of LeCam's deficiency distance Δ) by multinomial experiments $\mathscr{M}_n$ which are generated by a fixed interval partition of the sample space were investigated by D.W. Müller (1979). The experiments $\mathscr{M}_n$ can be obtained as follows. Let $\omega=(A_1,\dots,A_k)$ be a partition of $\mathbb{R}^1$ into k intervals. The set of all such partitions will be denoted by $\mathscr{I}(k)$. Let $i(x)$ be the index i for which $x\in A_i$. If X has distribution P then $i(X)$ has distribution $\bar{P}=(P(A_1),\dots,P(A_k))$ on $\{1,2,\dots,k\}$. The product experiments $(\bar{P}^n\colon P\in\mathscr{P}_n)$ are then equivalent to multinomial experiments $\mathscr{M}_n$.

For describing the conditions on the families $\mathscr{P}_n$ which guarantee that the experiments $\mathscr{M}_n$ and $\mathscr{P}_n^n$ have similar statistical attributes, let $h^2(P,Q)=\int(\sqrt{dP}-\sqrt{dQ})^2$ and $h_A^2(P,Q)=\int_A(\sqrt{dP}-\sqrt{dQ})^2$ denote the Hellinger and conditional Hellinger distance. Assume

(1.1) there exists a constant C, such that $n\,h^2(P_0,P)\leqq C$ for all n and $P\in\mathscr{P}_n$.

Under this condition the product measures P_0^n and P^n $(P\in\mathscr{P}_n)$ do not completely separate as $n\to\infty$. Further assume that there is no information in rate events:

(1.2) for every $\varepsilon>0$ there exist $\delta>0$ and n_0 such that $n\,h_A^2(P_0,P)\leqq\varepsilon$ for all $P\in\mathscr{P}_n$, $n\geqq n_0$, if $P_0(A)<\delta$,

and that

(1.3) the likelihood ratio dP/dP_0 is monotone for $P\in\mathscr{P}_n$.

Then (cf. D.W. Müller (1979)) for each $\varepsilon>0$ we find a number k and $\omega\in\mathscr{I}(k)$ such that $\Delta(\mathscr{P}_n^n,\mathscr{M}_n)\leqq\varepsilon$ for large n, where $\mathscr{M}_n$ is the multinomial experiment associated with ω as described above.

For any probability measure P on $\mathbb{R}^1$ define the function $\zeta_{P\cdot n}\in L^2(P_0)$ by $\zeta_{P\cdot n}=n^{\frac{1}{2}}((dP/dP_0)^{\frac{1}{2}}-1)$ where by convention $\zeta_{P\cdot n}=-n^{\frac{1}{2}}$ on $\{dP_0=0\}$. If for a sequence (P_n) $\zeta_{P_n\cdot n}$ converges to g in $L^2(P_0)$, we shall call the function g the asymptotic direction of the sequence (P_n). In this case g determines most of the asymptotic behaviour of the binary product experiments (P_0^n,P^n); in particular $4\|g\|_2^2=4\int g^2\,dP_0$ is the Fisher information of the sequence (P_0,P^n). Similarly (cf. D.W. Müller (1979), Lemma 4) $4\|\pi_\omega g\|_2^2$ is the Fisher information of the corresponding sequence $(\bar{P}_0,\bar{P}_n)$. Here $\pi_\omega g$ denotes the orthogonal projection of g onto the linear space spanned by the indicator functions I_{A_i} of the intervals A_i, i.e.

$$\pi_\omega g = \sum_{i=1}^{k} I_{A_i} \cdot \int_{A_i} g \, dP_0 / P_0(A_i).$$

So for the sequence (P_n) the quantity

$$4(\|g\|_2^2 - \|\pi_\omega g\|_2^2) = 4(\|g - \pi_\omega g\|_2^2)$$

is the loss of Fisher information due to using ω.

For asymptotic purposes we may therefore define the families $\mathscr{P}^n$ to be considered by specifying the set of possible asymptotic directions, which will be a subset $\hat{\mathscr{D}}$ of $L^2(P_0)$. Since

$$|\int \zeta_{P \cdot n} \, dP_0| \leqq \frac{C}{2} n^{-\frac{1}{2}}$$

and

$$\int \zeta_{P \cdot n}^2 \, dP_0 \leqq n \, h^2(P_0, P) \leqq C,$$

every $g \in \hat{\mathscr{D}}$ has to satisfy

(1.4) $$\int g \, dP_0 = 0$$

and

(1.5) $$\int g^2 \, dP_0 \leqq C$$

(or, in other words, the Fisher information is uniformly bounded). Also g should be monotone. Here we shall assume

(1.6) g is nondecreasing.

We replace the information condition (1.2) by the following stronger condition (1.7), which is better tractable mathematically. It was intended not to exclude the normal shift model ($g' = \text{const}$).

(1.7) g is absolutely continuous and $g' \leqq M$.

Now define

(1.8) $$\hat{\mathscr{D}} := \{ g \in L^2(P_0) \mid g \text{ fulfills (1.4)–(1.7)} \}.$$

For this model we want to find a partition $\omega \in \mathscr{I}(k)$ which minimizes the maximal loss of Fisher information, that is we want to solve the minimax problem

(1.9) $$\inf_{\omega \in \mathscr{I}(k)} \sup_{g \in \hat{\mathscr{D}}} \|g - \pi_\omega g\|_2^2.$$

Clearly the size k of the partitions has to be held fixed, because for $k \to \infty$ the loss of Fisher information will tend to zero.

2. Expansion of the Approximation Error

In this section we derive an expansion for $\|g-\pi_{\omega} g\|_2^2$ when the number of intervals of ω tends to infinity. Thereby we extend methods of Freedman and Diaconis (1981) and Ghurye and Johnson (1980). Proposition 2.7 of the former reference treats the case that P_0 is the Lebesgue measure λ^1 restricted to the unit interval (0, 1), and that ε_k is the "equidistant" partition of (0, 1), i.e.,

(2.1) $$\varepsilon_k=(B_{1k}, \ldots, B_{kk}),$$

where

$$B_{ik}=\left(\frac{i-1}{k}, \frac{i}{k}\right)$$

(for ease of notation we neglect the points i/k which form a P_0-nullset). If g is absolutely continuous and $g' \in L^2(P_0)$, we may apply Proposition 2.7 of Freedman and Diaconis (1981) to obtain

(2.2) $$\|g-\pi_{\varepsilon_k} g\|_2^2=\frac{l_k^2}{12} \int_0^1 g'^2\, dP_0+o(k^{-2})$$

as $k \to \infty$, where $l_k=$length of $B_{ik}=k^{-1}$. In order to give an idea of the proof of (2.2), note that we have

$$\|g-\pi_{\varepsilon_k} g\|_2^2=\frac{l_k^2}{12} \int_0^1 g'^2\, dP_0$$

if g' is constant on each of the intervals B_{ik}. (2.2) will follow by approximation.

We want to generalize (2.2) to the case of general partitions. For this purpose we now assume that h is a probability density on $\mathbb{R}^1$ and that $\omega_k \in (k)$ is defined by

(2.3) $\omega_k=(A_{1k}, \ldots, A_{kk})$, where

$A_{ik}=(z_{i-1}, z_i), \quad -\infty=z_0<z_1<\ldots<z_k=\infty,$

and each A_{ik} has probability k^{-1} under the probability measure with density h.

The length $l_k(x)$ of the interval A_{ik} which covers $x \in \mathbb{R}^1$ is approximately $k^{-1} h(x)^{-1}$, so one would expect

$$\|g-\pi_{\omega_k} g\|_2^2 \approx \frac{k^{-2}}{12} \int_{\infty}^{\infty} g'^2(x)\, h^{-2}(x)\, P_0(dx).$$

The following proposition shows that this is true for rather arbitrary h and P_0. We assume that h and ω_k are given as above and that P_0 has a density f.

Proposition 2.1. *Under the following conditions*

(i) $\int_{-\infty}^{\infty} g'^2 h^{-2} f d\lambda^1 < \infty$, *and* $\int_{-\infty}^{\infty} g^2 f d\lambda^1 < \infty$,

(ii) *f and h are continuous and strictly positive,*

(iii) *there exists* $\alpha > 0$ *such that* f/h *is nondecreasing on* $(-\infty, -\alpha)$ *and nonincreasing on* (α, ∞), *we have*

$$\|g - \pi_{\omega_k} g\|_2^2 = \frac{k^{-2}}{12} \int_{-\infty}^{\infty} g'^2 h^{-2} f d\lambda^1 + o(k^{-2}).$$

Proof. Let H be the distribution function of h. By Definition (2.3) $H(z_i) = i/k$ and therefore the image of ω_k under the map H: $\mathbb{R}^1 \to (0, 1)$ is the equidistant partition ε_k: $H(A_{ik}) = B_{ik}$. Therefore if we substitute $z = H^{-1}(t)$ in the integral

$$\|g - \pi_{\omega_k} g\|_2^2 = \int (g(z) - \pi_{\omega_k} g(z))^2 P_0(dz)$$

we obtain

$$\|g - \pi_{\omega_k} g\|_2^2 = \|\bar{g} - \bar{\pi}_{\varepsilon_k} \bar{g}\|_{2,Q}^2, \tag{2.4}$$

where $\bar{g}(t) = g \circ H^{-1}(t)$, Q is the image measure of P_0 under the map H and $\bar{\pi}_{\varepsilon_k}$ is the orthogonal projection of $L^2(Q)$ onto the linear span of $\{I_{B_{ik}} \mid i = 1, \ldots, k\}$. One easily verifies that Q has density $q = f \circ H^{-1}/h \circ H^{-1}$ with respect to λ^1 and that $g \in L^2(P_0)$ if and only if $\bar{g} \in L^2(Q)$. Let

$$\Delta(k) = \|\bar{g} - \bar{\pi}_{\varepsilon_k} \bar{g}\|_{2,Q}^2 - \frac{k^{-2}}{12} \int_0^1 \bar{g}'^2 dQ.$$

We are going to show

$$\Delta(k) = o(k^{-2}); \tag{2.5}$$

since $\int_0^1 \bar{g}'^2 dQ = \int_{-\infty}^{\infty} g'^2 h^{-2} f d\lambda^1$, this will prove the proposition.

Now we have an equidistant partition, but since Q need not be the Lebesgue measure, the methods of Freedman and Diaconis (1981) which we want to use have to be modified for our purpose. With the abbreviations

$$m_{ik} = \int_{B_{ik}} (\bar{g}(t) - \bar{\pi}_{\varepsilon_k} \bar{g}(t))^2 Q(dt) \quad \text{and} \quad n_{ik} = \frac{k^{-2}}{12} \int_{B_{ik}} \bar{g}'^2 dQ$$

$$\Delta(k) = \sum_{i=1}^{k} m_{ik} - n_{ik}, \tag{2.6}$$

and writing

$$\bar{g}(t)=\bar{g}\left(\frac{i-1}{k}\right)+\int_{\frac{i-1}{k}}^{t}\bar{g}'(u)\,du \tag{2.7}$$

we get

$$\begin{aligned}
m_{ik}&=\int_{B_{ik}}\left(\bar{g}(t)-\frac{1}{Q(B_{ik})}\int_{B_{ik}}\bar{g}(s)\,Q(ds)\right)^2 Q(dt)\\
&=\int_{B_{ik}}\left(\int_{\frac{i-1}{k}}^{t}\bar{g}'(u)\,du-\frac{1}{Q(B_{ik})}\int_{B_{ik}}\int_{\frac{i-1}{k}}^{s}\bar{g}'(u)\,du\,Q(ds)\right)^2 Q(dt)\\
&=\int_{B_{ik}}\left(\int_{\frac{i-1}{k}}^{t}\bar{g}'(u)\,du\right)^2 Q(dt)-\frac{1}{Q(B_{ik})}\left(\int_{B_{ik}}\int_{\frac{i-1}{k}}^{t}\bar{g}'(u)\,du\,Q(dt)\right)^2.
\end{aligned}$$

Now

$$\begin{aligned}
\int_{B_{ik}}\left(\int_{\frac{i-1}{k}}^{t}\bar{g}'(u)\,du\right)^2 Q(dt)&=\int_{B_{ik}}\int_{B_{ik}}\int_{B_{ik}}\bar{g}'(u)\,\bar{g}'(v)\,I_{\left(\frac{i-1}{k},t\right)}(u)\,I_{\left(\frac{i-1}{k},t\right)}(v)\,du\,dv\,Q(dt)\\
&=\int_{B_{ik}}\int_{B_{ik}}\bar{g}'(u)\,\bar{g}'(v)\int_{B_{ik}}I_{\left(\frac{i-1}{k},t\right)}(u)\,I_{\left(\frac{i-1}{k},t\right)}(v)\,Q(dt)\,du\,dv\\
&=\int_{B_{ik}}\int_{B_{ik}}\bar{g}'(u)\,\bar{g}'(v)\,Q\left(\left(u\vee v,\frac{i}{k}\right)\right)du\,dv,
\end{aligned}$$

and

$$\begin{aligned}
\int_{B_{ik}}\int_{\frac{i-1}{k}}^{t}\bar{g}'(u)\,du\,Q(dt)&=\int_{B_{ik}}\int_{B_{ik}}\bar{g}'(u)\,I_{\left(\frac{i-1}{k},t\right)}(u)\,du\,Q(dt)\\
&=\int_{B_{ik}}\bar{g}'(u)\,Q\left(\left(u,\frac{i}{k}\right)\right)du,
\end{aligned}$$

and therefore

$$\begin{aligned}
m_{ik}&=\int_{B_{ik}}\int_{B_{ik}}\bar{g}'(u)\,\bar{g}'(v)\left[Q\left(\left(u\vee v,\frac{i}{k}\right)\right)\quad\frac{Q\left(\left(u,\frac{i}{k}\right)\right)Q\left(\left(v,\frac{i}{k}\right)\right)}{Q(B_{ik})}\right]du\,dv \qquad (2.8)\\
&=\iint\bar{g}'(u)\,\bar{g}'(v)\;q(u)^{\frac{1}{2}}\,q(v)^{\frac{1}{2}}\,\psi_{iK}(u,v)\,du\,dv
\end{aligned}$$

where

$$\psi_{ik}(u,v)=(q(u)\,q(v))^{-\frac{1}{2}}\left[Q\left(\left(u\vee v,\frac{i}{k}\right)\right)-\frac{Q\left(\left(u,\frac{i}{k}\right)\right)Q\left(\left(v,\frac{i}{k}\right)\right)}{Q(B_{ik})}\right]I_{B_{ik}}(u)\cdot I_{B_{ik}}(v).$$

It is necessary to examine the tail behaviour of the approximation. Let $\beta = H(-\alpha)$; according to (iii) q is nondecreasing on $(0,\beta)$ and nonincreasing on $(1-\beta, 1)$. Let $0<\beta^*<\beta$ and split the sum in (2.6) into three parts:

$$\begin{aligned}\Delta(k) &= \sum_{i=1}^{[k\beta^*]+1} (m_{ik}-n_{ik}) + \sum_{i=[k\beta^*]+2}^{k-[k\beta^*]-1} (m_{ik}-n_{ik}) + \sum_{i=k-[k\beta^*]}^{k} (m_{ik}-n_{ik}) \\ &=: \sum_{i\in I_k^{(1)}(\beta^*)} (m_{ik}-n_{ik}) + \sum_{i\in I_k^{(2)}(\beta^*)} (m_{ik}-n_{ik}) + \sum_{i\in I_k^{(3)}(\beta^*)} (m_{ik}-n_{ik})\end{aligned}$$

(here $[x]$ denotes the integral part of x).

Consider the third sum. If $i\in I_k^{(3)} = I_k^{(3)}(\beta^*)$, then $i \geqq k-[k\beta^*]$ and

$$\frac{i-1}{k} \geqq 1 - \frac{[k\beta^*]+1}{k} \geqq 1-\beta^* - \frac{1}{k} > 1-\beta$$

if k is large, and therefore q is nonincreasing on B_{ik}. Thus for $u, v\in B_{ik}$ and $i\in I_k^{(3)}$,

$$\frac{Q\left(\left(u\vee v, \frac{i}{k}\right)\right)}{(q(u)\,q(v))^{\frac{1}{2}}} \leqq \frac{q(u\vee v)\left(\frac{i}{k} - u\vee v\right)}{(q(u)\,q(v))^{\frac{1}{2}}} \leqq k^{-1}\left(\frac{q(u\vee v)}{q(u\wedge v)}\right)^{\frac{1}{2}} \leqq k^{-1}.$$

Also

$$\frac{Q\left(\left(u, \frac{i}{k}\right)\right)}{(q(u)\,Q(B_{ik}))^{\frac{1}{2}}} = \left(\frac{Q\left(\left(u,\frac{i}{k}\right)\right)}{q(u)}\right)^{\frac{1}{2}} \left(\frac{Q\left(\left(u,\frac{i}{k}\right)\right)}{Q(B_{ik})}\right)^{\frac{1}{2}} \leqq \left(\frac{q(u)\left(\frac{i}{k}-u\right)}{q(u)}\right)^{\frac{1}{2}} \leqq k^{-\frac{1}{2}},$$

and therefore we have the estimate

$$|\psi_{ik}(u,v)| \leqq 2k^{-1} I_{B_{ik}}(u)\, I_{B_{ik}}(v), \qquad \text{if } i\in I_k^{(3)}.$$

Consequently

$$\begin{aligned}\Big|\sum_{i\in I_k^{(3)}} (m_{ik}-n_{ik})\Big| &\leqq \sum_{i\in I_k^{(3)}} |m_{ik}| + \sum_{i\in I_k^{(3)}} |n_i| \\ &\leqq 2k^{-1} \sum_{i\in I_k^{(3)}} \int_{B_{ik}} \int_{B_{ik}} |\bar g'(u)|\,|\bar g'(v)|\, q(u)^{\frac{1}{2}} q(v)^{\frac{1}{2}}\, du\, dv \\ &\quad + \frac{k^{-2}}{12} \sum_{i\in I_k^{(3)}} \int_{B_{ik}} \bar g'^2\, dQ \\ &= 2k^{-1} \sum_{i\in I_k^{(3)}} \Big(\int_{B_{ik}} |\bar g'(u)|\, q(u)^{\frac{1}{2}}\, du\Big)^2 + \frac{k^{-2}}{12} \int_{1-\frac{[k\beta^*]+1}{k}}^{1} \bar g'^2\, dQ \\ &\leqq 2k^{-2} \sum_{i\in I_k^{(3)}} \int_{B_{ik}} \bar g'^2\, dQ + \frac{k^{-2}}{12} \int_{1-\frac{[k\beta^*]+1}{k}}^{1} \bar g'^2\, dQ \\ &= (2+\tfrac{1}{12})\, k^{-2} \int_{1-\frac{[k\beta^*]+1}{k}}^{1} \bar g'^2\, dQ\end{aligned}$$

by the Cauchy-Schwarz inequality. We conclude that for each $\varepsilon>0$

$$\Big|\sum_{i\in I_k^{(3)}(\beta^*)}(m_{ik}-n_{ik})\Big|\leqq\varepsilon k^{-2}$$

if $\beta^*>0$ is small enough (use (i)). We similarly can conclude

$$\Big|\sum_{i\in I_k^{(1)}(\beta^*)}(m_{ik}-n_{ik})\Big|\leqq\varepsilon k^{-2}$$

if $\beta^*>0$ is sufficiently small.

So in order to prove (2.5) we only need to show that for each $\beta^*>0$

$$\sum_{i\in I_k^{(2)}(\beta^*)}(m_{ik}-n_{ik})=o(k^{-2}). \tag{2.9}$$

First note that $\bigcup_{i\in I_k^{(3)}(\beta^*)} B_{ik}\subset[\beta^*,1-\beta^*]$, and as q is positive and continuous there exist $0<\gamma<B$ such that $\gamma\leqq q(t)\leqq B$ for $\beta^*\leqq t\leqq 1-\beta$. Therefore we find step functions q_k which are constant on the intervals B_{ik} and approximate q uniformly on $[\beta^*,1-\beta^*]$:

$$\sup_{\beta^*\leqq t\leqq 1-\beta^*}|q(t)-q_k(t)|\to 0 \qquad \text{as } k\to\infty.$$

We may assume that $\gamma\leqq q_k(t)\leqq B$. Let

$$\varphi_{ik}(u,v):=\left[\lambda^1\left(\left(u\vee v,\frac{i}{k}\right)\right)-\frac{\lambda^1\left(\left(u,\frac{i}{k}\right)\right)\lambda^1\left(\left(v,\frac{i}{k}\right)\right)}{k^{-1}}\right]\cdot I_{B_{ik}}(u)\,I_{B_{ik}}(v).$$

For estimating $\psi_{ik}(u,v)-\varphi_{ik}(u,v)$ consider for $u,v\in B_{ik}$, $i\in I_k^{(2)}(\beta^*)$

$$\left|\frac{Q\left(\left(u\vee v,\frac{i}{k}\right)\right)}{(q(u)\,q(v))^{\frac{1}{2}}}-\lambda^1\left(\left(u\vee v,\frac{i}{k}\right)\right)\right|$$

$$=\left|\frac{\int_{u\vee v}^{i/k} q(t)\,dt}{(q(u)\,q(v))^{\frac{1}{2}}}-\frac{\int_{u\vee v}^{i/k} q_k(t)\,dt}{(q_k(u)\,q_k(v))^{\frac{1}{2}}}\right|$$

$$\leqq\frac{\int_{u\vee v}^{i/k}|q(t)-q_k(t)|\,dt}{(q(u)\,q(v))^{\frac{1}{2}}}+\int_{u\vee v}^{i/k} q_k(t)\,dt\,|(q(u)\,q(v))^{-\frac{1}{2}}-(q_k(u)\,q_k(v))^{-\frac{1}{2}}|$$

$$\leqq\gamma^{-1}k^{-1}\sup_{\beta^*\leqq t\leqq 1-\beta^*}|q(t')-q_k(t)|+B\,k^{-1}\gamma^{-2}\,2\sqrt{B}\sup_{\beta^*\leqq t\leqq 1-\beta^*}|q(t)^{\frac{1}{2}}-q_k(t)^{\frac{1}{2}}|.$$

A similar estimate holds for

$$\left|\frac{Q\left(\left(u,\frac{i}{k}\right)\right)Q\left(\left(v,\frac{i}{k}\right)\right)}{(q(u)\,q(v))^{\frac{1}{2}}\,Q(B_{ik})}-\frac{\lambda^1\left(\left(u,\frac{i}{k}\right)\right)\lambda^1\left(\left(v,\frac{i}{k}\right)\right)}{k^{-1}}\right|,$$

therefore, since $\sup\limits_{\beta^*\leqq t\leqq 1-\beta^*}|q(u)^{\frac{1}{2}}-q_k(u)^{\frac{1}{2}}|$ converges to zero, there exist a constant C_0 and a sequence δ_n, $\delta_n\to 0$, such that

$$|\psi_{ik}(u,v)-\varphi_{ik}(u,v)|\leqq C_0\,k^{-1}\delta_n$$

if $i\in I_k^{(2)}(\beta^*)$, $u, v\in B_{ik}$.

This entails (compare (2.8)) that

$$\begin{aligned}(2.10)\qquad & \Big|\sum_{i\in I_k^{(2)}} m_{ik}-\sum_{i\in I_k^{(2)}}\iint \bar g'(u)\,q(u)^{\frac{1}{2}}\,\bar g'(v)^{\frac{1}{2}}\,q(v)^{\frac{1}{2}}\,\varphi_{ik}(u,v)\,du\,dv\Big|\\ &\leqq C_0\,k^{-1}\delta_n\sum_{i\in I_k^{(2)}}\int_{B_{ik}}\int_{B_{ik}}|\bar g'(u)\,q(v)^{\frac{1}{2}}\,\bar g'(v)\,q(v)^{\frac{1}{2}}|\,du\,dv\\ &\leqq C_0\,k^{-2}\delta_n\sum_{i\in I_k^{(2)}}\int_{B_{ik}}\bar g'^2\,dQ\\ &\leqq C_0\,k^{-2}\delta_n\int_{\beta^*}^{1-\beta^*}\bar g'^2\,dQ=o(k^{-2})\end{aligned}$$

(use the Cauchy-Schwarz inequality).

It follows from (i) that we find step functions t_k taking the value t_{ik} on B_{ik}, which approximate $\bar g' q^{\frac{1}{2}}$ in $L^2([\beta^*, 1-\beta^*])$:

$$\int_{\beta^*}^{1-\beta^*}(\bar g'\,q^{\frac{1}{2}}-t_k)^2\,d\lambda^1\to 0\qquad\text{as } k\to\infty.$$

Therefore (writing $I_k^{(2)}$ for $I_k^{(2)}(\beta^*)$)

$$\begin{aligned}(2.11)\qquad & \sum_{i\in I_k^{(2)}}\iint \bar g'(u)\,\bar g'(v)\,q(u)^{\frac{1}{2}}\,q(v)^{\frac{1}{2}}\,\varphi_{ik}(u,v)\,du\,dv\\ &=\sum_{i\in I_k^{(2)}}\iint t_k(u)\,t_k(v)\,\varphi_{ik}(u,v)\,du\,dv+R_k,\end{aligned}$$

and

$$\begin{aligned}R_k=&\sum_{i\in I_k^{(2)}}\iint[\bar g'(u)\,q(u)^{\frac{1}{2}}\,\bar g'(v)\,q(v)^{\frac{1}{2}}-t_k(u)\,t_k(v)]\,\varphi_{ik}(u,v)\,du\,dv\\ =&\sum_{i\in I_k^{(2)}}\iint(\bar g'(u)\,q(u)^{\frac{1}{2}}-t_k(u))\,\bar g'(v)\,q(v)^{\frac{1}{2}}\,\varphi_{ik}(u,v)\,du\,dv\\ &+\sum_{i\in I_k^{(2)}}\iint(\bar g'(v)\,q(v)^{\frac{1}{2}}-t_k(v))\,t_k(u)\,\varphi_{ik}(u,v)\,du\,dv\end{aligned}$$

and since $|\varphi_{ik}(u,v)| \leqq 2k^{-1} I_{B_{ik}}(u) I_{B_{ik}}(v)$, we have

$$\begin{aligned}
|R_k| &\leqq 2k^{-1} \sum_{i\in I_k^{(2)}} \int_{B_{ik}} |\bar{g}'(v)\, q(v)^{\frac{1}{2}}|\, dv \int_{B_{ik}} |\bar{g}'(u)\, q(u)^{\frac{1}{2}} - t_k(u)|\, du \\
&\quad + 2k^{-1} \sum_{i\in I_k^{(2)}} \int_{B_{ik}} |t_k(u)|\, du \int_{B_{ik}} |\bar{g}'(v)\, q(v)^{\frac{1}{2}} - t_k(v)|\, dv \\
&\leqq 2k^{-1} \sum_{i\in I_k^{(2)}} \Big(k^{-1} \int_{B_{ik}} (\bar{g}'\, q^{\frac{1}{2}} - t_k)^2\, d\lambda^1 \cdot k^{-1} \int_{B_{ik}} \bar{g}'^2\, dQ\Big)^{\frac{1}{2}} \\
&\quad + 2k^{-1} \sum_{i\in I_k^{(2)}} \Big(k^{-1} \int_{B_{ik}} t_k^2\, d\lambda^1 \cdot k^{-1} \int_{B_{ik}} (\bar{g}'\, q^{\frac{1}{2}} - t_k)^2\, d\lambda^1\Big)^{\frac{1}{2}} \\
&\leqq 2k^{-2} \left(\int_{\beta^*}^{1-\beta^*} \bar{g}'^2\, dQ \cdot \int_{\beta^*}^{1-\beta^*} (\bar{g}'\, q^{\frac{1}{2}} - t_k)^2\, d\lambda^1 \right)^{\frac{1}{2}} \\
&\quad + 2k^{-2} \left(\int_{\beta^*}^{1-\beta^*} t_k^2\, d\lambda^1 \cdot \int_{\beta^*}^{1-\beta^*} (\bar{g}'\, q^{\frac{1}{2}} - t_k)^2\, d\lambda^1 \right)^{\frac{1}{2}} \\
&= o(k^{-2}),
\end{aligned}$$

where we repeatedly applied the Cauchy-Schwarz inequality.

Finally consider

$$\begin{aligned}
(2.12) \qquad \sum_{i\in I_k^{(2)}} \iint t_k(u)\, t_k(v)\, \varphi_{ik}(u,v)\, du\, dv &= \sum_{i\in I_k^{(2)}} t_{ik}^2 \int_{B_{ik}} \int_{B_{ik}} \varphi_{ik}(u,v)\, du\, dv \\
&= \sum_{i\in I_k^{(2)}} t_{ik}^2 \frac{k^{-2}}{12} \\
&= \frac{k^{-2}}{12} \sum_{i\in I_k^{(2)}} \int_{B_{ik}} t_k^2\, d\lambda^1 \\
&= \frac{k^{-2}}{12} \sum_{i\in I_k^{(2)}} \int_{B_{ik}} \bar{g}'^2\, dQ + o(k^{-2}) \\
&= \sum_{i\in I_k^{(2)}} n_{ik} + o(k^{-2}).
\end{aligned}$$

The second equality is obtained by evaluating the integral. Now combining (2.10), (2.11) and (2.12) yields (2.9) and the proof is completed.

Remark. One might be interested if for a fixed P_0 the expansion stated in Proposition 2.1 holds uniformly for a given set of pairs (g, h) of functions on $\mathbb{R}^1$. It is more natural to discuss this question for the corresponding expansion (2.5) which involves the transformed functions $\bar{g}$ and q. So assume $\mathscr{V}$ is a set of pairs $(\bar{g}, q)$, where $\bar{g}$ is an absolutely continous function on $(0,1)$ and q is a probability density on $(o, 1)$. Let

$$\mathscr{V}_1 = \{\bar{g} \mid (\bar{g}, q) \in \mathscr{V} \text{ for some } q\}$$

and

$$\mathscr{V}_2 = \{q \mid (\bar{g}, q) \in \mathscr{V} \text{ for some } \bar{g}\}.$$

Assume

($\mathrm{i_u}$) The functions $\{\bar{g}'^2 q \mid (\bar{g}, q) \in \mathscr{V}\}$ are uniformly integrable in $L^1((0,1))$, and $\int \bar{g}^2 \, dQ < \infty$ for $(\bar{g}, q) \in \mathscr{V}$,

($\mathrm{ii_u}$) for each compact interval $K \subset (0,1)$ the functions $\bar{g}'|_K$ ($\bar{g} \in \mathscr{V}_1$) and the functions $q|_K$ ($q \in \mathscr{V}_2$) are equicontinuous ($f|_K$ means the restriction of a function f to K).

($\mathrm{iii_u}$) there exists $0 < \beta \leqq 1/2$ such that each $q \in \mathscr{V}_2$ is nondecreasing on $(0, \beta)$ and nonincreasing on $(1-\beta, 1)$,

($\mathrm{iv_u}$) for each compact interval $K \subset (0,1)$ there exist $\gamma > 0$ and B such that $\gamma \leqq q \leqq B$ on K, if $q \in \mathscr{V}_2$.

Then we have

$$k^2 \left(\|\bar{g} - \bar{\pi}_{\varepsilon_k} \bar{g}\|_{2,Q}^2 - \frac{k^{-2}}{12} \int \bar{g}'^2 \, dQ \right) \to 0$$

as $k \to \infty$, uniformly in $(\bar{g}, q) \in \mathscr{V}$.

The proof of this is similar to the proof of Proposition 2.1. One has to approximate $\bar{g}'$ and q uniformly by step functions. Because of the equicontinuity assumption ($\mathrm{ii_u}$) the quality of this approximation does not depend on $(\bar{g}, q)$. As for the tail behaviour note that ($\mathrm{i_u}$) implies that $\int_0^{\beta^*} \bar{g}'^2 \, dQ + \int_{1-\beta^*}^1 \bar{g}'^2 \, dQ \to 0$ uniformly in $(\bar{g}, q) \in \mathscr{V}$ if $\beta^* \to 0$.

3. The Minimax Problem

Motivated by Proposition 2.1 we replace the term $\|g - \pi_\omega g\|_2^2$ in (1.9) by its first order approximation $\frac{k^{-2}}{12} \int g'^2 h^{-2} f$ (from now on we shall omit the symbol $d\lambda^1$ in the integrals, since integration will always be understood with respect to Lebesgue measure) and solve the corresponding minimax problem find $h_1 \in \mathscr{M}$ such that $\inf_{h \in \mathscr{M}} \sup_{g \in \hat{\mathscr{D}}} \int g'^2 h^{-2} f = \sup_{g \in \hat{\mathscr{D}}} \int g'^2 h_1^{-2} f$, where $\mathscr{M} = \{h \in L^1(\mathbb{R}^1) \mid h \geqq 0, \int h = 1\}$ and $\hat{\mathscr{D}}$ is defined by (1.8). Here we make the general assumption on f that there exists a probability density h_0 with $\int h_0^{-2} f < \infty$. The next lemma shows that we can replace the set $\hat{\mathscr{D}}$ by the sets $\mathscr{D}$ or $\bar{\mathscr{D}}$ which are defined in the following way:

$$\mathscr{D} = \{g \in L^2(P_0) \mid g \text{ is absolutely continuous}, 0 \leqq g' \leqq M, \text{and} \int g^2 f \leqq C\},$$

$$\bar{\mathscr{D}} = \{g \in \hat{\mathscr{D}} \mid \int g^2 f = C\}.$$

Lemma 3.1. *We have*

$$\sup_{g \in \bar{\mathscr{D}}} \int g'^2 h^{-2} f = \sup_{g \in \mathscr{D}} \int g'^2 h^{-2} f.$$

If $\bar{\mathscr{D}}$ *is not the empty set, then*

$$\sup_{g\in\bar{\mathscr{D}}} \int g'^2 h^{-2} f = \sup_{g\in\mathscr{D}} \int g'^2 h^{-2} f.$$

Proof. The first assertion of Lemma 3.1 is an immediate consequence of the inequality $\int (g-(\int gf))^2 f \leqq \int g^2 f$. To prove the second assertion we first show that $\mathscr{D} \neq \emptyset$ if and only if for the function $g_*(x) = Mx - \int Mzf(z)\,dz$ we have $\int g_* f \geqq C$. The "if"-part is obvious, set $\bar{g} = \alpha g_*$ for a suitable $\alpha \in [0,1]$. Now assume that there is a $\bar{g}$ in $\bar{\mathscr{D}}$, then $\int \bar{g} f = 0$, $\int \bar{g}^2 f = C$ and $0 \leqq \bar{g}' \leqq M$ by definition. Let $\bar{g}_s(x) = (1-s)\,\bar{g} + s\,M(x-\bar{x}_0)$, where $\bar{g}(\bar{x}_0) = 0$.

$$\bar{g}_s^2 \geqq \bar{g}^2 \quad \text{and} \quad \int \bar{g}_s f = s \int M(x-\bar{x}_0)\, f(x)\,dx.$$

Therefore

$$\begin{aligned}\int (\bar{g}_s - \int \bar{g}_s f)^2 f &= \int \bar{g}_s^2 f - (\int \bar{g}_s f)^2\\ &= \int \bar{g}_s^2\, f - s^2 (\int M(x-\bar{x}_0)\, f(x)\,dx)^2\\ &\geqq \int \bar{g}^2 f - s^2 \cdot p, \qquad p > 0.\end{aligned}$$

On the other hand

$$M(x-\bar{x}_0) - \int M(x-\bar{x}_0)\, f(x)\,dx = M\,x - \int M\,x\, f(x)\,dx = g_*(x),$$

so

$$\bar{g}_s - \int \bar{g}_s f = (1-s)\,\bar{g} - s\,g_*$$

and

$$\begin{aligned}\int (\bar{g}_s - \int \bar{g}_s f)^2 &\leqq (1-s) \int \bar{g}^2 f + s \int g_*^2 f\\ &= \int \bar{g}^2 f - s(\int \bar{g}^2 f - \int g_*^2 f)\\ &= \int \bar{g}^2 f - s(C - \int g_*^2 f).\end{aligned}$$

Thus $\int g_*^2 f \geqq C$ since otherwise the two inequalities would contradict each other for small s.

Now it is easy to prove the second part of the lemma. If $g \in \mathscr{D}$, $\int g^2 f < C$, we set $g_s = s\,g + (1-s)\,g_*$ $g_s \in \mathscr{D}$ and $g_s' \geqq g'$. As $\int g_s^2 f$ is continuous in s and $\int g_0^2 f \geqq C > \int g_1^2 f$ we find s_0 with $\int g_{s_0}^2 f = C$.

From now on we shall assume $\bar{\mathscr{D}} \neq \emptyset$. Otherwise g_* solves the maximum problem.

Lemma 3.2. *Suppose* $\int h^{-2} f < \infty$ *and* $\int x^2 f(x)\,dx < \infty$. *Then*

$$\sup_{g\in\mathscr{D}} \int g'^2 h^{-2} f = \max_{g\in\mathscr{D}} M \int g' h^{-2} f. \tag{3.1}$$

Proof. The right hand side of (3.1) is indeed a maximum because the functional $g' \to \int g' h^{-2} f$ is weak*-continuous on the dual $L^1(\mathbb{R}^1)^*$ of $L^1(\mathbb{R}^1)$ and the set $\mathscr{D}_d$ consisting of the derivatives g' with $g - \int gf \in \mathscr{D}$ is weak*-compact. The last assertion follows from the weak*-compactness of the set

$$\{h \in L^1(\mathbb{R}^1)^* \mid 0 \leqq h \leqq M\}$$

because the map $g' \to \int \left(\int_0^x g' \right)^2 f(x)\,dx - \left(\int \left(\int_0^x g' \right) f(x)\,dx \right)^2$ is continuous on $\mathscr{D}_d$. To see this let $g'_n \in \mathscr{D}_d$ and $g'_n \to h \in \mathscr{D}_d$ weakly*. Define $g_n(x) := \int_0^x g'_n$. $g_n(x) \leqq Mx$ and $g_n(x) = \int I_{[0,x]} g'_n \to \int I_{[0,x]} h = \int_0^x h =: H(x)$. Therefore by Lebesgue's dominated convergence theorem

$$\int g_n^2 f - \left(\int g_n f\right)^2 \to \int H^2 f - \left(\int H f\right)^2.$$

Next we are going to construct, for a given $g \in \mathscr{D}$, an approximating $g_\varepsilon \in \mathscr{D}$ such that g'_ε is 0 or M a.e. $[\lambda^1]$ and $g'_\varepsilon \to g'$ weakly*. We partition $I_1 = \{x \mid g(x) < 0\}$ and $I_2 = \{x \mid g(x) > 0\}$ into intervals $I_1^j = (a_1^j, b_1^j)$ and $I_2^j = (a_2^j, b_2^j)$ of length at most ε. Set

$$g_\varepsilon(x) = \begin{cases} g(a_1^j) + M(x - a_1^j), & \text{for } a_1^j \leqq x \leqq a_1^j + M^{-1} \int_{I_1^j} g' \\ g(b_1^j), & \text{for } a_1^j + M^{-1} \int_{I_1^j} g' \leqq x \leqq b_1^j \\ 0, & \text{for } x \in \mathbb{R}^1 - (I_1 \cup I_2) \\ g(a_2^j), & \text{for } a_2^j \leqq x \leqq b_2^j - M^{-1} \int_{I_2^j} g' \\ g(b_2^j) + M(x - b_2^j), & \text{for } b_2^j - M^{-1} \int_{I_2^j} g' \leqq x \leqq b_2^j. \end{cases}$$

We note that $g_\varepsilon \in \mathscr{D}$ and $|g_\varepsilon - g| \leqq M\varepsilon$. By compactness we find $\varepsilon_n \to 0$ and $h \in \mathscr{D}_d$ such that $g'_{\varepsilon_n} \to h$ weakly* and clearly $g' = h$. Since $M g'_\varepsilon = g'^2_\varepsilon$ we have

$$M \int g' h^{-2} f = M \lim_n \int g'_{\varepsilon_n} h^{-2} f = \lim_n \int g'^2_{\varepsilon_n} h^{-2} f$$

and therefore

$$\max_{g \in \mathscr{D}} M \int g' h^{-2} f \leqq \sup_{g \in \mathscr{D}} \int g'^2 h^{-2} f.$$

Now equality is immediate since

$$g'^2 \leqq M g' \qquad \text{for all } g \in \mathscr{D}.$$

Remark. In general the supremum on the left hand side of (3.1) will not be attained. This is because a possible maximum g_0 of the quadratic functional is also a maximum of the linear functional and has to fulfill $\int g_0'^2 h^{-2} f = M \int g'_0 h^{-2} f$. Since $g_0'^2 \leqq M g'_0$, we can conclude $g_0'^2 = M g'_0$ on $\{h^{-2} f > 0\}$. So on this set g'_0 takes on only the two values 0 and M. On the other hand the maximum of the linear functional can easily be shown to be unique. In general its derivative will have values strictly between 0 and M. Therefore g_0 cannot be the maximum of the linear functional, which is a contradiction.

Though this is not needed in the sequel, let us point out now to calculate for fixed h, $h^{-2}f\in L^1(\mathbb{R}^1)$, the maximal g for the linear functional.

Proposition 3.1. *Suppose*

$$\int g_1' h^{-2}f=\sup_{g\in\mathcal{D}}\int g' h^{-2}f.$$

Then $g_1=-v_\lambda'/f$, *where* v_λ *is the unique solution of the free boundary problem with constraint*

$$-\left(\frac{1}{f}v_\lambda'\right)'-M\cdot H(h^{-2}f-\lambda v_\lambda)\ni 0, \tag{3.2}$$

$$v_\lambda(-\infty)=v_\lambda(\infty)=0, \quad \textit{and } \lambda>0$$

such that $\int\frac{1}{f}(v_\lambda')^2=C$.

(Here we denote by H the Heaviside function

$$H(u)=\begin{cases}\{0\}, & \text{for } u<0\\ [0,1], & \text{for } u=0\\ \{1\}, & \text{for } u>0,\end{cases}$$

and for $z\in\mathbb{R}^1$ and $A\subset\mathbb{R}^1$ the expression $z+A$ has to be interpreted as $\{z\}+A=\{z+a\,|\,a\in A\}$.)

Proof. (a) Derivation of the Formula. We denote by $U_\delta(x)$ the closed δ-neighbourhood of x. According to Lebesgue's density theorem (see e.g., Hewitt-Stromberg (1965), Theorem 18.2), for almost all points x_1 and x_2 with $g_1'(x_1)>0$ and $g_1'(x_2)<M$ there exist $\varepsilon>0$ and $\gamma>0$ such that

$$\lambda^1(U_\delta(x_1)\cap\{g_1'>\varepsilon\})>\gamma\delta$$

and

$$\lambda^1(U_\delta(x_2)\cap\{g_1'<M-\varepsilon\})>\gamma\delta$$

for small $\delta>0$.

Setting

$$\varphi_{\delta,x_1}(x):=\int_{-\infty}^{x} I_{U_\delta(x_1)\cap\{g_1'>\varepsilon\}}(t)\,dt/\lambda^1(U_\delta(x_1)\cap\{g_1'>\varepsilon\}) \tag{3.3}$$

$$\varphi_{\delta,x_2}(x):=\int_{-\infty}^{x} I_{U_\delta(x_2)\cap\{g_1'<M-\varepsilon\}}(t)\,dt/\lambda^1(U_\delta(x_2)\cap\{g_1'<M-\varepsilon\})$$

we conclude

$$\int(g_1-\lambda_1\delta\varphi_{\delta,x_1}+\lambda_2\delta\varphi_{\delta,x_2})^2f-\int g_1^2f=-2\lambda_1\delta\int\varphi_{\delta,x_1}g_1f+2\lambda_2\delta\int\varphi_{\delta,x_2}g_1f+R$$

where $|R| \leqq \text{const} \cdot (\lambda_1^2 + \lambda_2^2)\,\delta^2$. Now for $\delta \to 0$

$$\int \varphi_{\delta, x_i}\, g_1 f \to \int_{x_i}^{\infty} g_1 f \qquad (i = 1, 2)$$

and this is strictly positive by the monotonicity of g_1 and by $g_1 \not\equiv 0$, $\int g_1 f = 0$. So we find sequences $\lambda_\nu^1 \to 0$ $(i = 1, 2)$, $\delta_\nu \to 0$ of positive numbers such that

$$(3.4) \qquad \frac{\lambda_\nu^1}{\lambda_\nu^2} \to \frac{\int_{x_2}^{\infty} g_1 f}{\int_{x_1}^{\infty} g_1 f} \quad \text{and} \quad g_1 - \lambda_\nu^1\, \delta_\nu\, \varphi_{\delta_\nu, x_1} + \lambda_\nu\, \delta_\nu\, \varphi_{\delta_\nu, x_2} \in \mathscr{D}.$$

To prove this note first that the condition on the derivative for this function to belong to $\mathscr{D}$ is fulfilled for small λ_1, λ_2 and δ. Secondly $\int g_1^2 f \leqq C$, and $-2\lambda_1\, \delta \int \varphi_{\delta, x_1}\, g_1 f + 2\lambda_2\, \delta \int \varphi_{\delta, x_2}\, g_1 f = 0$ is equivalent to

$$\frac{\lambda_1}{\lambda_2} = \frac{\int \varphi_{\delta, x_2}\, g_1 f}{\int \varphi_{\delta, x_2}\, g_1 f}.$$

So (3.4) follows from the estimation of R. By the maximality of g_1

$$-\lambda_\nu^1\, \delta_\nu \int \varphi'_{\delta_\nu, x_1}\, h^{-2} f + \lambda_\nu^2\, \delta_\nu \int \varphi'_{\delta_\nu, x_2}\, h^{-2} f \leqq 0$$

and therefore for almost all x_1, $g_1'(x_1) > 0$, and x_2, $g_1'(x_2) < M$ we get

$$\frac{h^{-2} f(x_2)}{\int_{x_2}^{\infty} g_1 f} \leqq \frac{h^{-2} f(x_1)}{\int_{x_1}^{\infty} g_1 f}.$$

So we find a Lagrange parameter $\lambda > 0$ such that

(i) $h^{-2} f(x_1) > \lambda \int_{x_1}^{\infty} g_1 f$ implies $g_1'(x_1) = M$

and

(ii) $h^{-2} f(x_2) < \lambda \int_{x_2}^{\infty} g_1 f$ implies $g_1'(x_2) = 0$.

Setting $v_\lambda(x) := \int_x^{\infty} g_1 f$ we get the formula.

(b) The Monotonicity of $\lambda \to v_\lambda$. Suppose $\lambda_1 \geqq \lambda_2$ and v_i is a solution of (3.2) for $\lambda = \lambda_i$ $(i=1,2)$. Multiplying the difference of the Eqs. (3.2) with $(v_1 - v_2)^+$ we get

$$-\left(\frac{1}{f}(v_1 - v_2)'\right)'(v_1 - v_2)^+$$
$$-M(H(h^{-2}f - \lambda_1 v_1) - H(h^{-2}f - \lambda_2 v_2))(v_1 - v_2)^+ \ni 0.$$

Now $M(H(h^{-2}f - \lambda_1 v_1) - H(h^{-2}f - \lambda_2 v_2))(v_1 - v_2)^+ \subset [-\infty, 0]$. This is clear for $v_1 \leqq v_2$. For $v_1 > v_2$ we have $-\lambda_1 v_1 < -\lambda_2 v_2 \leqq -\lambda_2 v_2$ and the statement follows from the monotonicity of H. As a consequence we get

$$-\left(\frac{1}{f}(v_1 - v_2)'\right)'(v_1 - v_2)^+ \leqq 0$$

and hence

$$\int_{-\infty}^{\infty} \left(\frac{1}{f}(v_1 - v_2)'\right)'(v_1 - v_2)^+ \geqq 0.$$

Integration by parts yields

$$-\int_{-\infty}^{\infty} \frac{1}{f}(v_1 - v_2)'((v_1 - v_2)^+)' + \frac{1}{f}(v_1 - v_2)'(v_1 - v_2)^+\Big|_{-\infty}^{\infty} \geqq 0,$$

Since the boundary terms vanish we get $(v_1 - v_2)^+ = 0$ and $v_1 \leqq v_2$.

(c) Uniqueness of the Solution of (3.2). For fixed λ, v_λ is unique by (b). Consequently it is the solution of the minimum problem

$$(*) \qquad \int \frac{1}{f}(v')^2 + 2M \int \left(\frac{1}{\lambda} h^{-2} f - v\right)^+ \to \min$$

whose Euler equation is (3.2). Now take $\lambda_1 > \lambda_2$, by (b) $v_{\lambda_1} \leqq v_{\lambda_2}$. Suppose $\int \frac{1}{f}(v'_{\lambda_1})^2 \geqq \int \frac{1}{f}(v'_{\lambda_2})^2$, then by the minimality of v_{λ_1}

$$\int \left(\frac{1}{\lambda_1} f/h^2 - v_{\lambda_1}\right)^+ \leqq \int \left(\frac{1}{\lambda_2} f/h^2 - v_{\lambda_2}\right)^+.$$

Since on the other hand $-v_{\lambda_1} \geqq -v_{\lambda_2}$ we have equality

$$\int \left(\frac{1}{\lambda_1} f/h^2 - v_{\lambda_1}\right)^+ = \int \left(\frac{1}{\lambda_2} f/h^2 - v_{\lambda_2}\right)^+$$

and since

$$\int \frac{1}{f}(v'_{\lambda_1})^2 = \int \frac{1}{f}(v'_{\lambda_2})^2,$$

v_{λ_2} solves the minimum problem (*) for λ_1 as well.

By uniqueness of the solution of (*) we therefore have $v_{\lambda_1}=v_{\lambda_2}$. This proves that for $\lambda_1>\lambda_2$ either $\int \frac{1}{f}(v'_{\lambda_1})^2<\int \frac{1}{f}(v'_{\lambda_2})^2$ or $v_{\lambda_1}=v_{\lambda_2}$ and therefore the solution of (3.2) is unique.

After this digression let us resume the investigation of the minimax problem. We want to interchange the order of sup and inf. The only obstacle to apply a minimax theorem is the noncompactness of $\mathcal{M}$ in the weak*-topology. So instead of $\mathcal{M}$ consider for $L>0$

$$\mathcal{M}_L:=\{h\in L^1(\mathbb{R}^1)\,|\,0\leqq h\leqq L,\ \int h\geqq 1\}.$$

$\mathcal{M}_L$ is a convex set, compact in $L^1(\mathbb{R}^1)^*$. The functional $h\to\int g'f/h^2$ is lower semi continuous on

$$\mathcal{M}'_L:=\mathcal{M}_L\cap\{h\,|\int h^{-2}f<\infty\}.$$

For $\lambda\in\mathbb{R}^1$

$$\{h\in\mathcal{M}'_L\,|\int g'h^{-2}f\leqq\lambda\}$$

is compact.

So we may apply the minimax theorem (see e.g., Kindler (1979)) to get

$$\inf_{h\in\mathcal{M}'_L}\sup_{g\in\mathcal{D}}\int g'h^{-2}f=\sup_{g\in\mathcal{D}}\inf_{h\in\mathcal{M}'_L}\int g'h^{-2}f.$$

For $h\in\mathcal{M}':=\mathcal{M}\cap\{h\,|\int h^{-2}f<\infty\}$ denote by h_L the function $h_L:=h\wedge L\in\mathcal{M}'_L$. For $g\in\mathcal{D}$ and $h\in\mathcal{M}'$

$$\begin{aligned}\int g'h_L^{-2}f&=\int_{\{h\leqq L\}}g'h_L^{-2}f+\int_{\{h>L\}}g'h_L^{-2}f\\&\leqq\int g'h^{-2}f+ML^{-2}.\end{aligned}$$

Hence

$$\begin{aligned}\inf_{h\in\mathcal{M}'_L}\int g'h^{-2}f&\leqq\inf_{h\in\mathcal{M}'}\int g'h_L^{-2}f\\&\leqq\inf_{h\in\mathcal{M}'}\int g'h^{-2}f+ML^{-2}\end{aligned}$$

and therefore

$$\begin{aligned}\inf_{h\in\mathcal{M}'}\sup_{g\in\mathcal{D}}\int g'h^{-2}f&\leqq\inf_{h\in\mathcal{M}'_L}\sup_{g\in\mathcal{D}}\int g'h^{-2}f\\&=\sup_{g\in\mathcal{D}}\inf_{h\in\mathcal{M}'_L}\int g'h^{-2}f\\&\leqq\sup_{g\in\mathcal{D}}\inf_{h\in\mathcal{M}'}\int g'h^{-2}f+ML^{-2}.\end{aligned}$$

As L was arbitrary we have

$$\inf_{h\in\mathcal{M}'}\sup_{g\in\mathcal{D}}\int g'h^{-2}f\leqq\sup_{g\in\mathcal{D}}\inf_{h\in\mathcal{M}'}\int g'h^{-2}f.$$

The converse inequality is trivial, so

$$\inf_{h\in\mathcal{M}'}\sup_{g\in\mathcal{D}}\int g' h^{-2} f=\sup_{g\in\mathcal{D}}\inf_{h\in\mathcal{M}'}\int g' h^{-2} f.$$

We are looking for a solution h_1 of

$$\inf_{h\in\mathcal{M}'}\sup_{g\in\mathcal{D}}\int g' h^{-2} f=\sup_{g\in\mathcal{D}}\int g' h_1^{-2} f.$$

By the preceding argument

$$\sup_{g\in\mathcal{D}}\int g' h_1^{-2} f=\sup_{g\in\mathcal{D}}\inf_{h\in\mathcal{M}'}\int g' h^{-2} f=:\alpha.$$

Therefore we are going to calculate h_g:

$$\inf_{h\in\mathcal{M}'}\int g' h^{-2} f=\int g' h_g^{-2} f$$

and g_1:

$$\sup_{g\in\mathcal{D}}\int g' h_g^{-2} f=\int g_1' h_{g_1}^{-2} f.$$

As $\int g_1' h_1^{-2} f\leqq\alpha=\int g_1' h_{g_1}^{-2} f\leqq\int g_1' h_1^{-2} f$ by the strict convexity of $h\to h^{-2} f$

$$h_1=h_{g_1}\qquad \text{on } \{g_1'>0\} \text{ (later we shall show: } g_1'>0 \text{ a.e.).}$$

So let h_g be the (unique) minimum of $\int g' h^{-2} f$. Take as a variation

$$h_\lambda(x)=[h_g(x)+I_{\{h_g\geqq\varepsilon\}}(x)\,\lambda\,\varphi(x)]/C(\lambda,\varphi),$$
$$C(\lambda,\varphi)=\int[h_g+I_{\{h_g\geqq\varepsilon\}}\lambda\,\varphi]=1+\lambda\int_{\{h_g\geqq\varepsilon\}}\varphi.$$

By the minimality of h_g

$$\begin{aligned}0&\geqq\int g' h_g^{-2} f-\int g' h_\lambda^{-2} f\\&=\int g' f(h_g^{-2}-C(\lambda,\varphi)^2 h_g^{-2})+\int g' f(C(\lambda,\varphi)^2 h_g^{-2}-h_\lambda^{-2})\\&=\int_{\{h_g\geqq\varepsilon\}}2\lambda\,\varphi(x)\left(-\int g' h_g^{-2} f+g'(x)\,h_g^{-3}(x)\,f(x)\right)dx+o(\lambda),\end{aligned}$$

and therefore

$$\int_{\{h_g\geqq\varepsilon\}}\varphi(x)\left(-\int g' h_g^{-2} f+g'(x)\,h_g^{-3}(x)\,f(x)\right)dx=0.$$

Taking as a variation $h_\lambda=(h_g+\lambda\,\varphi)/\int(h_g+\lambda\,\varphi)$ with $\lambda>0$, $\varphi>0$ we get

$$\int\varphi(x)\left(-\int g' h_g^{-2} f+g'(x)\,h_g^{-3}(x)\,f(x)\right)dx\leqq 0.$$

So we have

$$h_g = \text{const} \cdot (g' f)^{\frac{1}{3}} = (g' f)^{\frac{1}{3}} / \int (g' f)^{\frac{1}{3}}$$

(incidentically this proves: the existence of $h_0 \in L^1(\mathbb{R}^1)$ such that $\int h_0^{-2} f < \infty$ is equivalent to $f^{\frac{1}{3}} \in L^1(\mathbb{R}^1)$).

Now we can calculate g_1. As

$$\int g' h_g^{-2} f = (\int (g' f)^{\frac{1}{3}})^3$$

g_1 fulfills

$$\int (g_1' \mathrm{f})^{\frac{1}{3}} = \sup_{g \in \mathscr{D}} \int (g' f)^{\frac{1}{3}}.$$

We proceed as in the proof of Proposition 3.1. Take as a variation

$$g_1 - \lambda^1 \delta_\nu \varphi_{\delta_\nu, x_1} + \lambda^2 \delta_\nu \varphi_{\delta_\nu, x_2}$$

with

$$\frac{\lambda_\nu^1}{\lambda_\nu^2} \to \frac{\int_{x_2}^{\infty} g_1 f}{\int_{x_1}^{\infty} g_1 f} \qquad \text{(compare (3.4))},$$

where φ_{δ, x_i} are defined by (3.3). We calculate

$$(*) \qquad \frac{f(x_1)(g_1' f)^{-\frac{2}{3}}(x_1)}{\int_{x_1}^{\infty} g_1 f} \geqq \frac{f(x_2)(g_1' f)^{-\frac{2}{3}}(x_2)}{\int_{x_2}^{\infty} g_1 f}$$

if $g_1'(x_1) > 0$ and $g_1'(x_2) < M$. An immediate conclusion is that $g_1' > 0$ a.e. or $g_1' = 0$ a.e. Also if $0 < g_1'(x) < M$ then

$$l(x) = \frac{f(x)(g_1' f)^{-\frac{2}{3}}(x)}{\int_{x}^{\infty} g_1 f} = \lambda > 0$$

(independent of x). Therefore if $l(x) > \lambda$ we must have $g_1'(x) = M$ and if $l(x) < \lambda$ then $g_1'(x) = 0$ (compare with $(*)$).

Now assume $g_1' > 0$ a.e., then $l \geqq \lambda$ a.e. We shall prove that

$$(**) \qquad g_1'(x) = \inf \left\{ M, \left(\frac{f^{\frac{1}{3}}(x)}{\lambda \int_{x}^{\infty} g_1 f} \right)^{\frac{3}{2}} \right\}.$$

Case I: $M<\left(\frac{f^{\frac{1}{3}}(x)}{\lambda\int\limits_x^\infty g_1 f}\right)^{\frac{3}{2}}$. Since $g_1'(x)\leqq M$, we then have $g_1'(x)<\left(\frac{f^{\frac{1}{3}}(x)}{\lambda\int\limits_x^\infty g_1 f}\right)^{\frac{3}{2}}$ or (equivalently) $l(x)>\lambda$, and hence $g_1'(x)=M$.

Case II: $M\geqq\left(\frac{f^{\frac{1}{3}}(x)}{\lambda\int\limits_x^\infty g_1 f}\right)^{\frac{3}{2}}$. Under this hypothesis we get $l(x)=\lambda$, since $l(x)>\lambda$ would imply $g_1'(x)=M$ and $\frac{f^{\frac{1}{3}}(x)M^{-\frac{2}{3}}}{\int\limits_x^\infty g_1 f}=l(x)>\lambda$ contrary to the hypothesis. But $l(x)=\lambda$ is equivalent to

$$g_1'(x)=\left(\frac{f^{\frac{1}{3}}(x)}{\lambda\int\limits_x^\infty g_1 f}\right)^{\frac{3}{2}}=\lambda^*\frac{f^{\frac{1}{2}}(x)}{\left(\int\limits_x^\infty g_1 f\right)^{\frac{3}{2}}}\qquad(\text{where } \lambda^*=\lambda^{-\frac{3}{2}}).$$

Hence (**) is proved. Setting $v_{\lambda^*}(x)=\int\limits_x^\infty g_1 f$ yields

$$-\left(\frac{1}{f}v_{\lambda^*}'\right)'-\inf\{M,\lambda^* f^{\frac{1}{2}} v_{\lambda^*}^{-\frac{3}{2}}\}=0. \tag{3.5}$$

So we have proved the major part of

Proposition 3.2. *The minimax solution* h_1 *is given by* $h_1=h_{g_1}=(g_1'f)^{\frac{1}{3}}/\int(g_1'f)^{\frac{1}{3}}$, *and* g_1 *is determined by* $g_1=-v_{\lambda^*}'/f$, *where* $\lambda^*>0$ *and* v_{λ^*} *is the unique solution of* (3.5) *with the constraints*

(i) $v_{\lambda^*}(-\infty)=v_{\lambda^*}(\infty)=0$,

(ii) $\int(v_{\lambda^*}')^2/f=C$.

Proof. We still have to prove uniqueness.

(a) Let $\lambda_1\leqq\lambda_2$ and v_{λ_i} be solutions of (3.5) with constraint (i). We claim that $v_{\lambda_1}\leqq v_{\lambda_2}$. Testing the difference of the Eqs. (3.5) with $(v_{\lambda_1}-v_{\lambda_2})^+$ we get

$$\int\frac{1}{f}[(v_{\lambda_1}-v_{\lambda_2})^+]'^2+\int(\inf\{M,\lambda_2 f^{\frac{1}{2}}v_{\lambda_2}^{-\frac{3}{2}}\}-\inf\{M,\lambda_1 f^{\frac{1}{2}}v_{\lambda_1}^{-\frac{3}{2}}\})(v_{\lambda_1}-v_{\lambda_2})^+=0.$$

But

$$\lambda_1 v_{\lambda_1}^{-\frac{2}{3}}(x)\leqq\lambda_2 v_{\lambda_2}^{-\frac{2}{3}}(x)\qquad\text{if } v_{\lambda_1}(x)\geqq v_{\lambda_2}(x)$$

so $\int\frac{1}{f}[(v_{\lambda_1}-v_{\lambda_2})^+]'^2=0$ and consequently $v_{\lambda_1}\leqq v_{\lambda_2}$.

(b) Let v and $\bar{v}$ be solutions of (3.5) with constraint (i) for Lagrange parameters λ and $\bar{\lambda}$. Assume that

$$\int\frac{1}{f}(v')^2=C\geqq\bar{C}=\int\frac{1}{f}(\bar{v}')^2.$$

We already know that either $v \leqq \bar{v}$ or $v \geqq \bar{v}$. Furthermore v and $\bar{v}$ are the unique solutions of the minimum problem

$$\int \frac{1}{f}(v')^2 - 2\int \Phi(\lambda, v) \to \min$$

for their respective Lagrange parameters. Here

$$\Phi(\lambda, v) = \inf\{M v, \lambda^{\frac{2}{3}} M^{\frac{1}{3}} f^{\frac{1}{3}}\} + [2 M^{\frac{1}{3}} \lambda^{\frac{2}{3}} f^{\frac{1}{3}} - 2\lambda f^{\frac{1}{2}} v^{-\frac{1}{2}}]^{+}.$$

(3.5) is the Euler equation of this minimum problem.

In case that $\bar{v} \geqq v$ we obtain $-\Phi(\lambda, \bar{v}) \leqq -\Phi(\lambda, v)$, and therefore

$$\int \left(\frac{1}{f}(\bar{v}')^2 - \Phi(\lambda, \bar{v})\right) \leqq \int \left(\frac{1}{f}(\bar{v}')^2 - \Phi(\lambda, v)\right) = \bar{C} - \int \Phi(\lambda, v)$$

$$\leqq C - \int \Phi(\lambda, v) = \int \left(\frac{1}{f}(v')^2 - \Phi(\lambda, v)\right).$$

By the uniqueness of the minimal v in this case we get $v = \bar{v}$. This proves that in any case we have $\bar{v} \leqq v$.

Now according to (a) condition (i) uniquely determines the solution of (3.5) for each λ^*, and according to (b) condition (ii) uniquely determines λ^*. Thus we have proved Proposition 3.2.

We turn to the interpretation of the results. Assume that f is continuous, $\int x^2 f(x)\,dx < \infty$, and $\int f^{\frac{1}{3}} < \infty$. Then $\int g^2\,dP_0$ is finite for $g \in \mathscr{D}$ and there exists $h_0 \in \mathscr{M}$ with $\int h_0^{-2} f < \infty$. Assume further that $f^{\frac{1}{3}}(x) \Big/ \int_x^{\infty} t f(t)\,dt \to \infty$ for $|x| \to \infty$, then it follows from the differential equation (3.5) that for $|x|$ large enough $g_1'(x) = M$ and therefore $h_1(x) \sim f^{\frac{1}{3}}(x)$ and $f(x)/h_1(x) \sim f^{\frac{2}{3}}(x)$; hence the pair (g, h_1) fulfills the conditions of Propositon 2.1 for every $g \in \mathscr{D}$. Compare h_1 with any $\bar{h} \in \mathscr{M}$ such that $(g, \bar{h})$ fulfills these conditions for every $g \in \mathscr{D}$. We then have

$$\int g_1' h_1^{-2} f = \sup_{g \in \mathscr{D}} \int g' h_1^{-2} f \leqq \sup_{g \in \mathscr{D}} \int g' \bar{h}^{-2} f$$

and according to Lemma 3.2

$$\sup_{g \in \mathscr{D}} \int g'^2 h_1^{-2} f \leqq \sup_{g \in \mathscr{D}} \int g'^2 \bar{h}^{-2} f.$$

What does this mean for the respective loss of Fisher information? Let ω_k^1 and $\bar{\omega}_k$ be the partitions associated with h_1 and $\bar{h}$ according to (2.3). h_1 is better than $\bar{h}$ in the sense that for each $g \in \mathscr{D}$ we find $\hat{g} \in \mathscr{D}$ and k_0, such that

$$\|g - \pi_{\omega_k^1}(g)\|_2^2 < \|\hat{g} - \pi_{\bar{\omega}_k}(\hat{g})\|_2^2, \quad \text{if } k > k_0.$$

Our results have to be interpreted in this way, since the expansion of Proposition 2.1 is not uniform on $\mathscr{D} \times \mathscr{M}$.

Example. We solved the free boundary problem (3.5) numerically for $f(x) = (2\pi)^{-1}\exp(-x^2/2)$ and $M=1$ (by scaling we may always standardize in this way). For a choice of λ^*'s we obtained the results:

λ^*	M	C	$\int g_1' h_1^{-2} f$
0.01	1.0	0.0934	13.15
0.05	1.0	0.2944	21.08
0.1	1.0	0.4689	25.19
0.2	1.0	0.7213	29.38
0.3	1.0	0.8974	31.58

The following exhibits show the functions g_1 and h_1. In each case g_1 has tails of the form constant$+M\cdot x$, and h_1 has tails proportional to $\exp(-x^2/6)$, that is proportional to the normal density with mean 0 and variance 3.

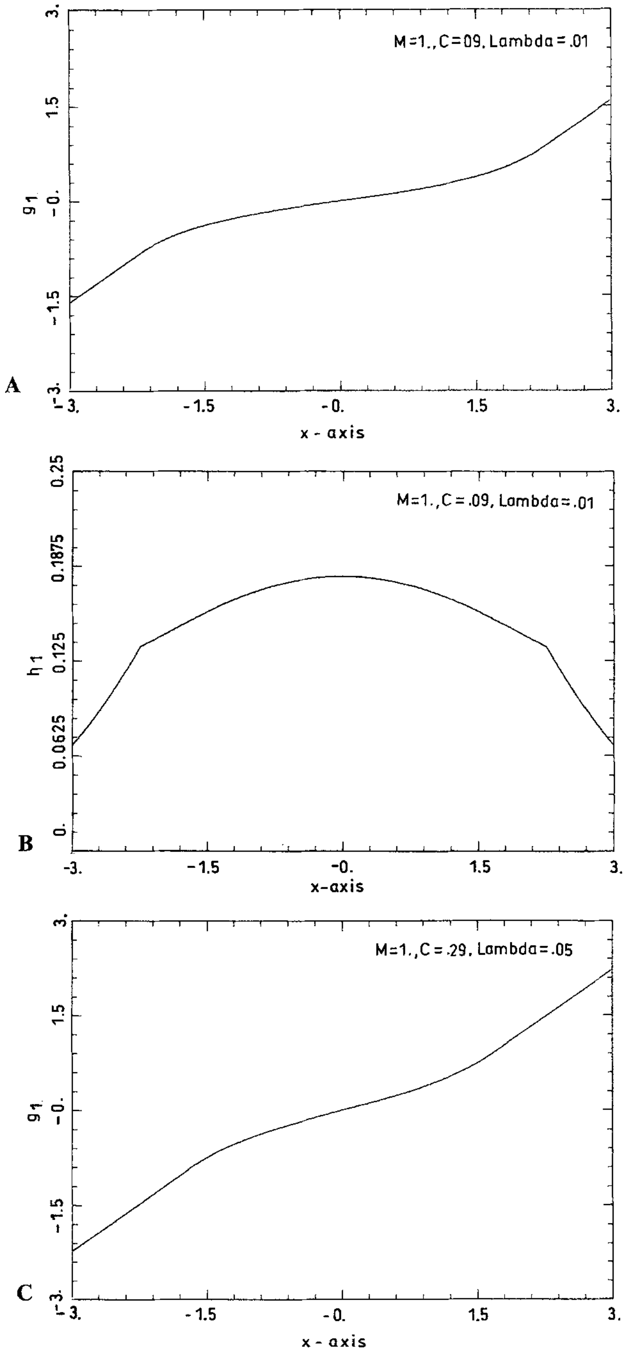
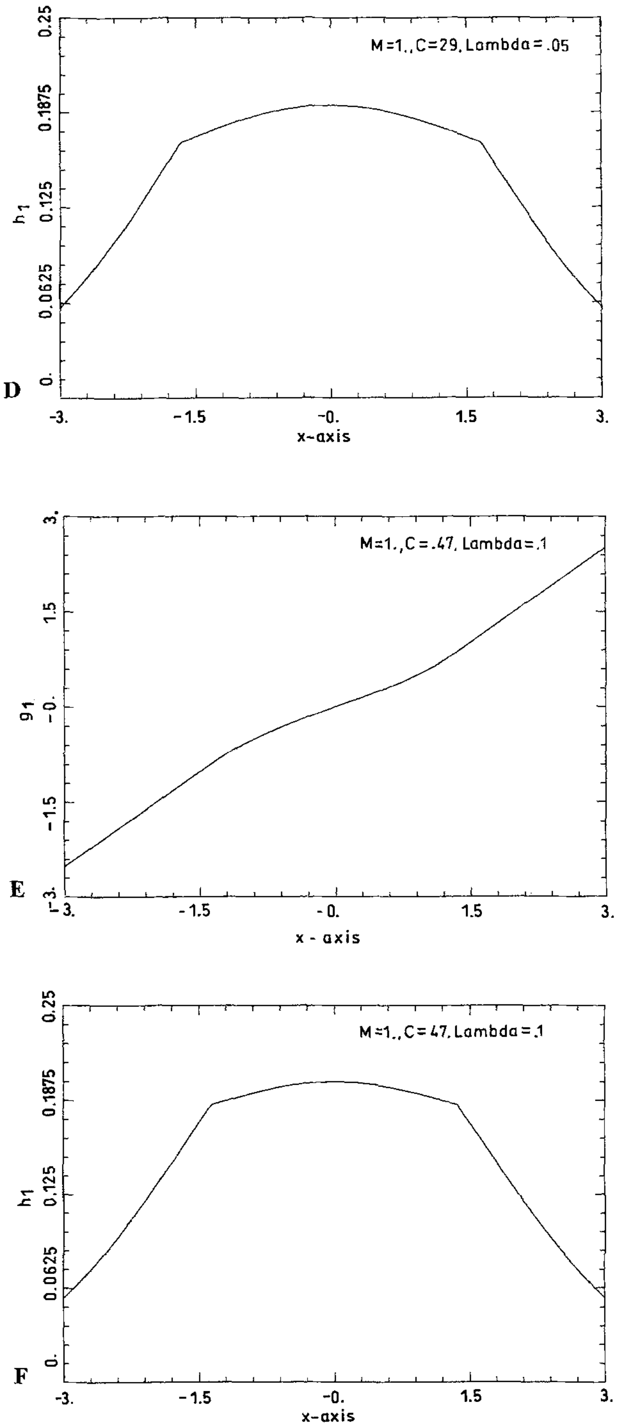


Fig. 1A–C

Acknowledgement. We thank W. Ehm and D.W. Müller who contributed some of the main ideas, and also J. Franke, G. Kersting and E. Mammen for helpful discussions.



D
M=1., C=29, Lambda=.05
h_1
0.25
0.1875
0.125
0.0625
0.
-3.
-1.5
-0.
1.5
3.
x-axis
E
M=1., C=.47, Lambda=.1
g_1
3.
1.5
-0.
-1.5
-3.
-3.
-1.5
-0.
1.5
3.
x-axis
F
M=1., C=47, Lambda=.1
h_1
0.25
0.1875
0.125
0.0625
0.
-3.
-1.5
-0.
1.5
3.
x-axis

Fig. 1D–F

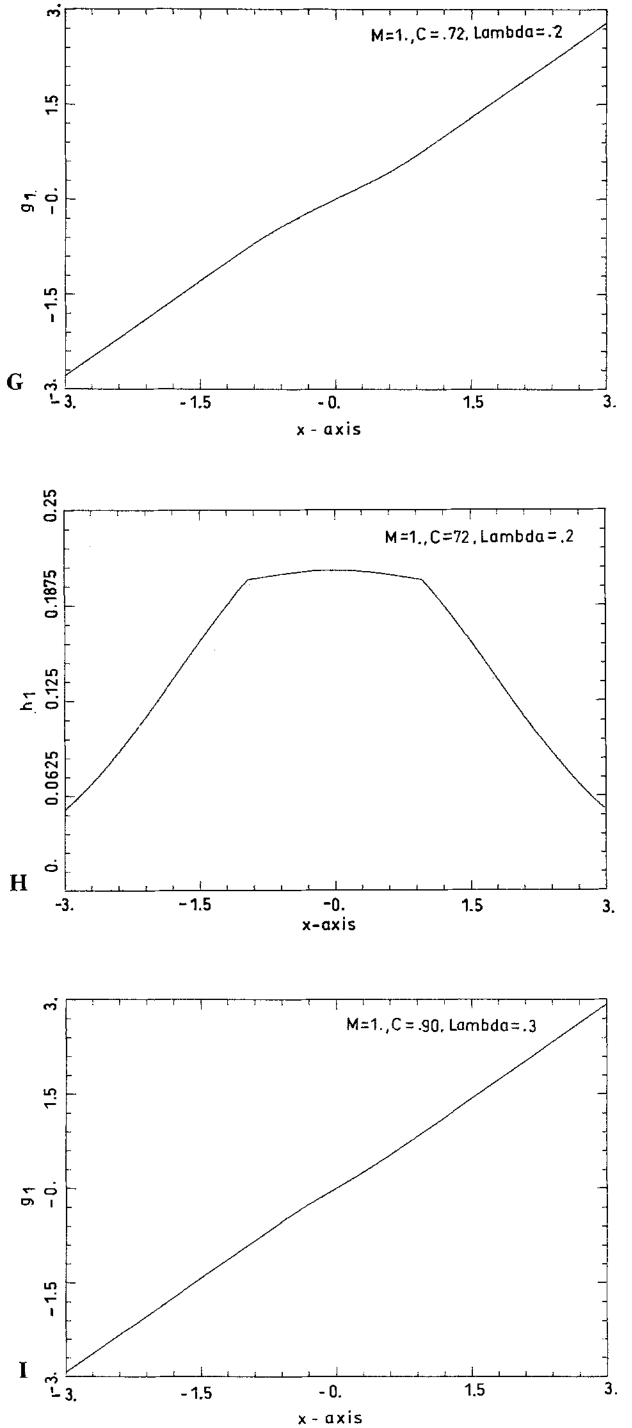

G
M=1.,C=.72, Lambda=.2
g1
x - axis
H
M=1.,C=72, Lambda=.2
h1
x-axis
I
M=1.,C=.90, Lambda=.3
g1
x - axis

Fig. 1 G–I

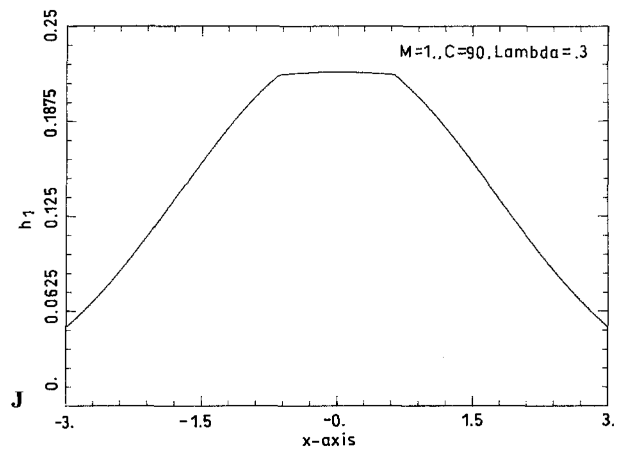


Fig. 1 J

References

Freedman, D., Diaconis, P.: On the histogram as a density estimator: L_2 theory. Z. Wahrscheinlichkeitstheor. Verw. Geb. **57**, 453–476 (1981)

Ghurye, S.G., Johnson, B.R.: Discrete approximations to the information integral. Preprint (1980)

Hewitt, E., Stromberg, K.: Real and abstract analysis. New York, Heidelberg, Berlin: Springer 1965

Kindler, J.: Minimaxtheoreme und das Integraldarstellungsproblem. Manuscripta Math. **29**, 277–294 (1979)

Müller, D.W.: Asymptotically multinomial experiments and the extension of a theorem of Wald. Z. Wahrscheinlichkeitstheor. Verw. Geb. **50**, 179–204 (1979)

Received December 21, 1987; in revised form July 22, 1988