The Annals of Statistics
1977, Vol. 5, No. 5, 909–921

LEAST FAVORABLE PAIRS FOR SPECIAL CAPACITIES

BY HELMUT RIEDER

University of Freiburg

The least favorable pair (LFP) that Huber (1965), (1968) wrote down when he considered minimax test problems between neighborhoods of single probability measures P_0, P_1 defined in terms of ε-contamination and total variation is a canonical but only one possible choice of an LFP.

We treat these neighborhoods by means of special capacities. The minimax test statistic is obtained by explicitly solving a minimization program, all LFP's are characterized by their $(P_0 + P_1)$-densities, another LFP is given explicitly.

The technique is similar to that used by Huber and Strassen (1973), but is simpler and more constructive in this special situation.

1. Introduction. In their famous paper Huber and Strassen (1973) generalized the classical Neyman–Pearson lemma to 2-alternating capacities v_0, v_1 on a polish sample space Ω by proving the existence of a minimax test statistic π and an LFP (Q_0, Q_1). The Neyman–Pearson tests between Q_0 and Q_1, based on π, then constitute a minimal essentially complete class of minimax tests for all fixed sample sizes between $\mathfrak{P}_{v_0}$ and $\mathfrak{P}_{v_1}$, the classes of all probability measures (p.m.'s) setwise dominated by v_0 and v_1, respectively.

Due to the generality of the admitted capacities, however, their proofs are nonconstructive and mere existence proofs. Furthermore, because of a continuity property of their capacities, Ω has to be compact metric in order that the required assumptions are fulfilled for neighborhoods of p.m.'s P_0, P_1 in terms of ε-contamination and total variation.

On the other hand, under somewhat shifted assumptions, Huber (1965), (1968) just wrote down a minimax test statistic and an LFP for these special neighborhoods and an arbitrary sample space.

We treat this case in a constructive way by means of special capacities (without the mentioned continuity property; see Section 3 below). The minimax test statistic is obtained by explicitly solving a minimization program (Theorem 5.1). Then all LFP's can be characterized by their $(P_0 + P_1)$-densities (Theorem 5.2). Huber's LFP turns out to be a canonical choice, but another LFP can be given explicitly (Section 6). For middle values of dP_1/dP_0, however, the densities are unique.

Beforehand, in Section 2, we restate in a slightly generalized and strengthened form the role of LFP's for minimax test problems and a uniqueness property of the minimax test statistic—independently of any structure of the classes of

Received November 1975; revised March 1977.

AMS 1970 *subject classifications*. Primary 62G35; Secondary 62G35.

Key words and phrases. Least favorable pairs, minimax test problems, ε-contamination, total variation, capacity.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics*.
www.jstor.org

distributions. The uniqueness property reappears in a sharpened version when the special neighborhoods are considered.

2. Minimax test problems and least favorable pairs. Let $\mathfrak{M}$ denote the set of all p.m.'s on an arbitrary sample space Ω, endowed by the σ-field $\mathfrak{B}$, and let Φ be the set of all tests on Ω. Then a *minimax test problem* (MTP) is formally described by a hypothesis $\mathfrak{P}_0 \subset \mathfrak{M}$ and an alternative $\mathfrak{P}_1 \subset \mathfrak{M}$, the set of all admitted tests and a risk function. Whereas the classes $\mathfrak{P}_j$, $j = 0, 1$, are not yet further specified, we confine ourselves from the first to risk functions R that can be derived from the errors of the first and second kind by assigning to each test φ between $Q' \in \mathfrak{P}_0$ and $Q'' \in \mathfrak{P}_1$ the risk

$$R(\varphi; Q', Q'') = r(\textstyle\int \varphi \, dQ', \int (1 - \varphi) \, dQ'') \,, \tag{2.1}$$

where $r\colon [0, 1] \times [0, 1] \to [0, \infty]$ is a continuous function increasing in each argument.

Accordingly, for $\alpha \in [0, 1]$, the class Φ_α of all admitted tests is defined by

$$\Phi_\alpha = \{\varphi \in \Phi\colon \sup_{Q' \in \mathfrak{P}_0} \textstyle\int \varphi \, dQ' \leqq \alpha\} \,. \tag{2.2}$$

Then $\varphi^* \in \Phi_\alpha$ is a solution of the MTP $(\mathfrak{P}_0, \mathfrak{P}_1, r, \alpha)$, called a minimax test, iff it minimizes $\sup \{R(\varphi; Q', Q'')\colon Q' \in \mathfrak{P}_0, Q'' \in \mathfrak{P}_1\}$ for $\varphi \in \Phi_\alpha$. This setup obviously includes the MTP's considered by Huber (1965) as simple special cases.

Such MTP's can be reduced to simple test problems if there is a pair $(Q_0, Q_1) \in \mathfrak{P}_0 \times \mathfrak{P}_1$ which has the following property:

$$\begin{aligned} Q_0(\pi > t) &= \sup \{Q'(\pi > t)\colon Q' \in \mathfrak{P}_0\} \,, \\ Q_1(\pi > t) &= \inf \{Q''(\pi > t)\colon Q'' \in \mathfrak{P}_1\} \,, \qquad 0 < t < \infty \end{aligned} \tag{2.3}$$

where π is a suitable version of the Radon–Nikodym derivative dQ_1/dQ_0,

$$\frac{dQ_1}{dQ_0} = \left\{\frac{q_1}{q_0} : q_j \in \frac{dQ_j}{d(Q_0 + Q_1)} \,, \; q_j \geqq 0, \, j = 0, 1, \, q_0 + q_1 > 0\right\} . \tag{2.4}$$

(Q_0, Q_1) is called a *least favorable pair* (LFP) for $(\mathfrak{P}_0, \mathfrak{P}_1)$. If the special version π should be stressed, we also write $(Q_0, Q_1 \,|\, \pi)$.

If follows at once that each Neyman–Pearson test φ^* between Q_0 and Q_1, based on π, satisfies

$$R(\varphi^*; Q', Q'') \leqq R(\varphi^*; Q_0, Q_1) \,, \quad \forall \, (Q', Q'') \in \mathfrak{P}_0 \times \mathfrak{P}_1 \,. \tag{2.5}$$

Now the simple MTP $(\{Q_0\}, \{Q_1\}, r, \alpha)$ is easily seen (by the weak compactness of Φ and the continuity and isotony of r) to be solved by a Neyman–Pearson test φ^* which automatically lies in Φ_α. In view of (2.5), φ^* is also a solution of the MTP $(\mathfrak{P}_0, \mathfrak{P}_1, r, \alpha)$.

Thus we get the following generalization of Huber (1965), Theorem 1:

PROPOSITION 2.1. *An* LFP $(Q_0, Q_1 \,|\, \pi)$ *is least favorable for any* MTP $(\mathfrak{P}_0, \mathfrak{P}_1, r, \alpha)$.

There is some ambiguity about the notion of an LFP. Whereas definition (2.3) as stated underlies Huber's and Strassen's respective papers, the original definition of a least favorable pair of a priori distributions by Lehmann (1959), page 328, involves power functions only (denoted by β subsequently). Restricting a priori distributions to one point measures in his definition and requiring independence of the level α we arrive at the following condition for (Q_0, Q_1) to be an LFP:

$$(2.6) \qquad \beta_{(Q_0,Q_1)}(\alpha) \leqq \beta_{(Q',Q'')}(\alpha) \qquad \text{for all} \quad \alpha\,, \quad (Q', Q'') \in \mathfrak{P}_0 \times \mathfrak{P}_1\,.$$

Obviously, property (2.3) implies (2.6). The converse holds if there exists an LFP in the sense of (2.3). This is the reason why we need not differ in the sequel and can work with condition (2.3).

Proposition 2.2. *Let* (Q_{00}, Q_{10}) *be an* LFP *in the sense of* (2.6). *If there exists an* LFP $(Q_{01}, Q_{11} \mid \pi_1)$ *in the sense of* (2.3), *then* $(Q_{00}, Q_{10} \mid \pi_1)$ *too is an* LFP *in the sense of* (2.3).

Proof. Both pairs have the same (minimal) power function β. But then π_1 and $\pi_0 \in dQ_{10}/dQ_{00}$ are equally distributed:

$$(2.7) \qquad \mathfrak{L}_{Q_{00}}\pi_0 = \mathfrak{L}_{Q_{01}}\pi_1 \qquad \text{and} \qquad \mathfrak{L}_{Q_{10}}\pi_0 = \mathfrak{L}_{Q_{11}}\pi_1\,.$$

For one has, for example,

$$(2.8) \qquad Q_{00}(\pi_0 > t) = \inf\{\alpha : t\alpha - \beta(\alpha) = \inf_{\alpha'} t\alpha' - \beta(\alpha')\}\,, \qquad 0 < t < \infty\,,$$

since $I_{\{\pi_0 > t\}}$ has minimum level among all Bayes tests at the a priori distribution $(t/(1+t), 1/(1+t))$ for (Q_{00}, Q_{10}). It follows that

$$(2.9) \qquad \begin{aligned} Q_{00}(\pi_0 > t) &= Q_{01}(\pi_1 > t) \geqq Q_{00}(\pi_1 > t)\,, \\ Q_{10}(\pi_0 > t) &= Q_{11}(\pi_1 > t) \leqq Q_{10}(\pi_1 > t)\,, \qquad 0 < t < \infty\,, \end{aligned}$$

showing that the real measures $t \cdot Q_{00} - Q_{10}$, which are obviously minimized by $\{\pi_0 > t\}$, are also minimized by $\{\pi_1 > t\}$. Hence one must have

$$(2.10) \qquad \pi_0 = \pi_1 \quad (Q_{00} + Q_{10})\text{-a.e.}$$

and in particular: $\pi_1 \in dQ_{10}/dQ_{00}$.

Finally, by (2.10), (2.9) and (2.3),

$$(2.11) \qquad \begin{aligned} Q_{00}(\pi_1 > t) &= Q_{01}(\pi_1 > t) \geqq Q'(\pi_1 > t)\,, \\ Q_{10}(\pi_1 > t) &= Q_{11}(\pi_1 > t) \leqq Q''(\pi_1 > t)\,, \end{aligned}$$

holds for $(Q', Q'') \in \mathfrak{P}_0 \times \mathfrak{P}_1$, $0 < t < \infty$, which is the assertion. □

By the way, in view of (2.10), we have proved the following uniqueness property of the likelihood ratio of LFP's.

Proposition 2.3. *Let* $(Q_{00}, Q_{10} \mid \pi_0)$ *and* $(Q_{01}, Q_{11} \mid \pi_1)$ *be* LFP's. *Then*

$$(2.12) \qquad \pi_0 = \pi_1 \quad (Q_{00} + Q_{01} + Q_{10} + Q_{11})\text{-a.e.}$$

This proposition generalizes Theorem 5.1 of Huber and Strassen (1973), whose proof is bound to capacities.

3. Special capacities. In the sequel we investigate neighborhoods $\mathfrak{P} \subset \mathfrak{M}$ of $P \in \mathfrak{M}$ of the following structure:

$$\mathfrak{P} = \{Q \in \mathfrak{M} : (1-\varepsilon)P(B) - \delta \leqq Q(B) \leqq (1-\varepsilon)P(B) + \varepsilon + \delta, \quad \forall\, B \in \mathfrak{B}\}, \tag{3.1}$$

where

$$0 \leqq \varepsilon, \delta \leqq 1, \qquad 0 < \varepsilon + \delta < 1. \tag{3.2}$$

For $\delta = 0$ one has the ε-contamination model, for $\varepsilon = 0$ the total variation model. $\mathfrak{P}$ can also be described by

$$\mathfrak{P} = \{Q \in \mathfrak{M} : Q(B) \leqq v(B), \ \ \forall\, B \in \mathfrak{B}\}, \tag{3.3}$$

where $v : \mathfrak{B} \to [0, 1]$ is defined by

$$\begin{aligned} v(B) &= ((1-\varepsilon)P(B) + \varepsilon + \delta) \wedge 1, && \text{if } B \neq \emptyset \\ &= 0, && \text{if } B = \emptyset. \end{aligned} \tag{3.4}$$

One immediately verifies:

$$v(\emptyset) = 0, \qquad v(\Omega) = 1, \tag{3.5}$$

$$B' \subset B'' \Rightarrow v(B') \leqq v(B''), \tag{3.6}$$

$$B_n \uparrow B \Rightarrow v(B_n) \uparrow v(B), \tag{3.7}$$

$$v(B' \cup B'') + v(B' \cap B'') \leqq v(B') + v(B''). \tag{3.8}$$

Hence v has all properties of a 2-alternating Choquet-capacity except for the continuity property (4) in Huber and Strassen (1973), page 252, which would require a compact Ω.

The conjugate u to v satisfies $u(B) = 1 - v(\Omega \backslash B)$, and is given by

$$\begin{aligned} u(B) &= ((1-\varepsilon)P(B) - \delta) \vee 0, && \text{if } B \neq \Omega \\ &= 1, && \text{if } B = \Omega. \end{aligned} \tag{3.9}$$

For the following definition let v_0, v_1 be induced by $P_j \in \mathfrak{M}$, ε_j, δ_j ($j = 0, 1$) according to (3.2), (3.4). Generalizing the Radon–Nikodym derivative of two probability measures, the RND dv_1/dv_0 is defined as the set of all measurable $\pi : \Omega \to [0, \infty]$, such that

$$w_t(A_t) = \inf w_t(\mathfrak{B}) \quad \forall\, t \in (0, \infty), \qquad \text{where} \quad w_t = tv_0 - u_1, \quad A_t = \{\pi > t\}. \tag{3.10}$$

4. Technicalities.

LEMMA 4.1. *The function $t \to w_t(A_t)$, see (3.10), is absolutely continuous on $(0, \infty)$, with derivative (in measure) $t \to v_0(A_t)$.*

PROOF. Take $0 < s < t < \infty$. Then by (3.10):

$$0 \leqq (t-s)v_0(A_t) = w_t(A_t) - w_s(A_t) \leqq w_t(A_t) - w_s(A_s) \leqq w_t(A_s) - w_s(A_s) = (t-s)v_0(A_s),$$

and absolute continuity follows, as well as the expression for the (right-hand) derivative, if we let $t \downarrow s$, and observe (3.7). □

The second lemma corresponds to Huber and Strassen (1973), Lemma 2.5.

LEMMA 4.2. $\forall B \in \mathfrak{B}\ \exists Q \in \mathfrak{P}$: $Q(B) = v(B)$. (See (3.3), (3.4).)

PROOF. If $B = \emptyset$ or $P(B) = 1$ take $Q = P$. If $B \neq \emptyset$ and $P(B) = 0$ take $Q = (1-(\varepsilon+\delta))P + (\varepsilon+\delta)\delta_x$, where δ_x denotes the one point mass in $x \in B$.
Now $0 < P(B) < 1$. Then also $0 < Q_1(B) < 1$, where $Q_1 = (1-\varepsilon)P + \varepsilon\delta_x$, and we can take

$$Q(A) \equiv \frac{v(B)}{Q_1(B)} \cdot Q_1(A \cap B) + \frac{1-v(B)}{1-Q_1(B)} \cdot Q_1(A \backslash B) \quad (A \in \mathfrak{B}). \qquad \square$$

LEMMA 4.3. *Let* $\mathfrak{P}_j$ $(j = 0, 1)$ *be defined according to* (3.1), (3.2). *Then*

$$\mathfrak{P}_0 \cap \mathfrak{P}_1 \neq \emptyset \tag{4.1}$$

is equivalent to

$$(1-\varepsilon_1)P_1(B) + \varepsilon_1 + \delta_1 \geqq (1-\varepsilon_0)P_0(B) - \delta_0 \quad \forall B \in \mathfrak{B}. \tag{4.2}$$

PROOF. One direction is trivial. For the other one, take $\mu \in \mathfrak{M}$ dominating P_0 and P_1 with respective densities p_0 and p_1. Then, by (4.2), if $\delta = \delta_0 + \delta_1$,

$$\int ((1-\varepsilon_0)p_0 - (1-\varepsilon_1)p_1)^+ \, d\mu \leqq \varepsilon_1 + \delta. \tag{4.3}$$

Therefore: $\gamma_1 \geqq 0$, where

$$\gamma_1 = 1 - \frac{1}{\varepsilon_1 + \delta} \int ((1-\varepsilon_0)p_0 - (1-\varepsilon_1)p_1)^+ \, d\mu. \tag{4.4}$$

Hence

$$h_1 = \frac{1}{\varepsilon_1 + \delta}((1-\varepsilon_0)p_0 - (1-\varepsilon_1)p_1)^+ + \gamma_1 \tag{4.5}$$

defines the μ-density of a $H_1 \in \mathfrak{M}$ with the property

$$\varepsilon_1 H_1(B) \geqq (1-\varepsilon_0)P_0(B) - (1-\varepsilon_1)P_1(B) - \delta \quad \forall B \in \mathfrak{B}. \tag{4.6}$$

In the same way (after taking complements in (4.6)) we manufacture a $H_0 \in \mathfrak{M}$ such that

$$(1-\varepsilon_1)P_1 + \varepsilon_1 H_1 + \delta_1 \geqq (1-\varepsilon_0)P_0 + \varepsilon_0 H_0 - \delta_0 \quad \text{on } \mathfrak{B}. \tag{4.7}$$

Finally, a suitable convex combination of the thus contaminated P_j's is seen to lie in $\mathfrak{P}_0 \cap \mathfrak{P}_1$. □

In the last lemma we summarize some facts essentially contained in Huber (1965), (1968). Given $0 < \varepsilon_j + \delta_j < 1$, $P_j \in \mathfrak{M}$, $j = 0, 1$, $\Delta \in dP_1/dP_0$. Define new parameters

$$\nu_j = \frac{\varepsilon_j + \delta_j}{1 - \varepsilon_j}, \qquad \omega_j = \frac{\delta_j}{1 - \varepsilon_j} \tag{4.8}$$

and functions $\psi_0, \psi_1 \colon (0, \infty) \to [0, \infty)$ by

$$\psi_0(t) = \frac{1}{\nu_1 + \omega_0 t}(t \cdot P_0(\Delta < t) - P_1(\Delta < t)), \tag{4.9}$$

$$\psi_1(t) = \frac{1}{\nu_0 t + \omega_1}(P_1(\Delta > t) - t \cdot P_0(\Delta > t)). \tag{4.10}$$

Furthermore, two equations in $\Delta_0, \Delta_1 \in (0, \infty)$ are given by

$$\Delta_0 \cdot P_0(\Delta < \Delta_0) - P_1(\Delta < \Delta_0) = \nu_1 + \omega_0 \Delta_0, \tag{4.11}$$

$$P_1(\Delta_1 < \Delta) - \Delta_1 \cdot P_0(\Delta_1 < \Delta) = \nu_0 \Delta_1 + \omega_1. \tag{4.12}$$

LEMMA 4.4. *ψ_0 is strictly increasing on $\{\psi_0 > 0\}$, ψ_1 is strictly decreasing on $\{\psi_1 > 0\}$, and ψ_j characterizes the solution Δ_j, which uniquely exists, by $\psi_j(\Delta_j) = 1$ $(j = 0, 1)$.*

5. Main theorems. These technical preparations enable us to construct now the RND dv_1/dv_0 and to characterize all LFP's for the special capacities introduced in Section 3.

The following notation is used: $P_j \in \mathfrak{M}$, $\mathfrak{P}_j$, ε_j, δ_j, v_j, u_j according to (3.1), (3.2), (3.4), (3.9), Δ_j according to (4.11), (4.12), $j = 0, 1$. Let p_j denote the density of P_j with respect to some dominating, σ-finite μ. Furthermore, we may assume that $\mathfrak{P}_0 \cap \mathfrak{P}_1 = \emptyset$ since otherwise the problem of LFP's becomes trivial.

Now, if $(Q_0, Q_1 \mid \pi)$ is an LFP for $\mathfrak{P}_0$ and $\mathfrak{P}_1$, Lemma 4.2 shows that we must necessarily have

$$Q_0(\pi > t) = v_0(\pi > t), \qquad Q_1(\pi > t) = u_1(\pi > t) \quad \forall\, t \in (0, \infty). \tag{5.1}$$

Since $\pi \in dQ_1/dQ_0$ and $Q_0 \leqq v_0$, $Q_1 \geqq u_1$ one concludes

$$\begin{aligned} &t \cdot v_0(\pi > t) - u_1(\pi > t) \\ &\qquad = t \cdot Q_0(\pi > t) - Q_1(\pi > t) \leqq t \cdot Q_0(B) - Q_1(B) \\ &\qquad \leqq t \cdot v_0(B) - u_1(B) \quad \forall\, B \in \mathfrak{B}. \end{aligned} \tag{5.2}$$

This means (remember (3.10)) that

$$\pi \in \frac{dv_1}{dv_0}. \tag{5.3}$$

In order to derive the form of π we determine dv_1/dv_0.

THEOREM 5.1.

$$\frac{dv_1}{dv_0} = \left\{\frac{1 - \varepsilon_1}{1 - \varepsilon_0}(\Delta_0 \vee \Delta \wedge \Delta_1) \colon \Delta \in \frac{dP_1}{dP_0}\right\}. \tag{5.4}$$

Then all LFP's can be characterized by $(P_0 + P_1)$-densities.

THEOREM 5.2. *If* $(Q_0, Q_1 | \pi)$ *is an* LFP *for* $(\mathfrak{P}_0, \mathfrak{P}_1)$ *then it has the following properties*:

(5.5) μ *dominates* Q_0, Q_1 — *with densities* q_0, q_1, *say.*

(5.6) $$q_1/q_0 = \frac{1-\varepsilon_1}{1-\varepsilon_0}(\Delta_0 \vee \Delta \wedge \Delta_1) \qquad (Q_0 + Q_1)\text{-a.e. } \textit{for some} \quad \Delta \in \frac{dP_1}{dP_0}.$$

(5.7) $$q_0 = (1-\varepsilon_0)p_0 \qquad \mu\text{-a.e. } \textit{on} \quad \{\Delta_0 \leqq \Delta \leqq \Delta_1\}.$$

(5.8) $$(1-\varepsilon_0)\frac{p_1}{\Delta_0} \leqq q_0 \leqq (1-\varepsilon_0)p_0 \qquad \mu\text{-a.e. } \textit{on} \quad \{\Delta < \Delta_0\}.$$

(5.9) $$(1-\varepsilon_0)p_0 \leqq q_0 \leqq (1-\varepsilon_0)\frac{p_1}{\Delta_1} \qquad \mu\text{-a.e. } \textit{on} \quad \{\Delta_1 < \Delta\}.$$

(5.10) $$Q_0(\Delta < \Delta_0) = (1-\varepsilon_0)P_0(\Delta < \Delta_0) - \delta_0.$$

On the other hand, if (5.5)—(5.10) *are satisfied by* $Q_0, Q_1 \in \mathfrak{M}$, *then* $(Q_0, Q_1 | \pi)$ *is an* LFP *for* $(\mathfrak{P}_0, \mathfrak{P}_1)$ *exactly with each* $\pi = ((1-\varepsilon_1)/(1-\varepsilon_0))(\Delta_0 \vee \Delta \wedge \Delta_1)$, $\Delta \in dP_1/dP_0$.

REMARKS. At first glance these conditions seem to be asymmetric in Q_0 and Q_1. But in view of (5.6), condition (5.7) is equivalent to

(5.11) $$q_1 = (1-\varepsilon_1)p_1 \qquad \mu\text{-a.e. on} \quad \{\Delta_0 \leqq \Delta \leqq \Delta_1\},$$

and the first respective second inequality in (5.8) respective (5.9) mean

(5.12) $$q_1 \geqq (1-\varepsilon_1)p_1 \qquad \mu\text{-a.e. on} \quad \{\Delta < \Delta_0\},$$

(5.13) $$q_1 \leqq (1-\varepsilon_1)p_1 \qquad \mu\text{-a.e. on} \quad \{\Delta_1 < \Delta\}.$$

Similarly, by (4.11), (4.12), condition (5.10) is equivalent to

(5.14) $$Q_1(\Delta < \Delta_0) = (1-\varepsilon_1)P_1(\Delta < \Delta_0) + \varepsilon_1 + \delta_1,$$

(5.15) $$Q_0(\Delta_1 < \Delta) = (1-\varepsilon_0)P_0(\Delta_1 < \Delta) + \varepsilon_0 + \delta_0,$$

(5.16) $$Q_1(\Delta_1 < \Delta) = (1-\varepsilon_1)P_1(\Delta_1 < \Delta) - \delta_1.$$

PROOF OF THEOREM 5.1. Take a $\pi \in dv_1/dv_0$, w_t and A_t as in (3.10). At first we show the equivalence of

(5.17) $$w_t(A_t) = t - 1 \qquad \text{and}$$

(5.18) $$t \leqq \frac{1-\varepsilon_1}{1-\varepsilon_0}\Delta_0, \quad \Delta_0 \text{ given by (4.11).}$$

But (5.17) is equivalent to

(5.19) $$t^{-1} \geqq \sup\left\{\frac{u_0(B)}{v_1(B)} : B \in \mathfrak{B}, B \neq \emptyset\right\}$$

and this sup is not smaller than

(5.20) $$\sup\left\{\frac{(1-\varepsilon_0)P_0(B) - \delta_0}{(1-\varepsilon_1)P_1(B) + \varepsilon_1 + \delta_1} : B \in \mathfrak{B}, B \neq \emptyset\right\}.$$

By (4.9) and Lemma 4.4, (5.20) equals

$$\frac{1-\varepsilon_0}{1-\varepsilon_1}\Delta_0^{-1}. \tag{5.21}$$

Now, the assumption, that the sup in (5.19) is strictly greater than (5.21) implies the existence of a $B \in \mathfrak{B}$, $B \neq \emptyset$, such that

$$\frac{u_0(B)}{v_1(B)} > \frac{1-\varepsilon_0}{1-\varepsilon_1}\Delta_0^{-1} \tag{5.22}$$

and therefore

$$u_0(B) > (1-\varepsilon_0)P_0(B) - \delta_0, \tag{5.23}$$

or

$$v_1(B) < (1-\varepsilon_1)P_1(B) + \varepsilon_1 + \delta_1, \tag{5.24}$$

and in either case

$$v_1(B) = 1. \tag{5.25}$$

Because of $\mathfrak{P}_0 \cap \mathfrak{P}_1 = \emptyset$ we conclude from Lemma 4.3, that

$$\psi_0\left(\frac{1-\varepsilon_0}{1-\varepsilon_1}\right) > 1, \tag{5.26}$$

where ψ_0 is given by (4.9); hence by Lemma 4.4,

$$\Delta_0 < \frac{1-\varepsilon_0}{1-\varepsilon_1}. \tag{5.27}$$

Together with (5.22), (5.25) this leads to the contradiction: $u_0(B) > 1$. The same chain of arguments proves the equivalence

$$w_t(A_t) = 0 \Leftrightarrow t \geqq \frac{1-\varepsilon_1}{1-\varepsilon_0}\Delta_1, \tag{5.28}$$

where Δ_1 is given by (4.12).

Define

$$t_0 = \frac{1-\varepsilon_1}{1-\varepsilon_0}\Delta_0, \quad t_1 = \frac{1-\varepsilon_1}{1-\varepsilon_0}\Delta_1. \tag{5.29}$$

Now Lemma 4.1 and (3.7) come into play and show, by differentiating (5.17):

$$v_0(A_t) = 1 \quad \forall t < t_0, \qquad \text{hence also} \tag{5.30}$$

$$u_1(A_t) = 1 \quad \forall t < t_0, \qquad \text{therefore} \tag{5.31}$$

$$A_t = \Omega \quad \forall t < t_0. \tag{5.32}$$

Since $w_t(A_t) \neq t - 1$, for $t > t_0$ we have, on the other hand,

$$A_t \neq \Omega \quad \forall t > t_0. \tag{5.33}$$

From (5.28), Lemma 4.1 and (3.7) we similarly derive the equivalence

$$A_t = \emptyset \Leftrightarrow t \geqq t_1. \tag{5.34}$$

Next, we observe that for $t_0 \leqq t < t_1$:

$$v_0(A_t) = (1 - \varepsilon_0)P_0(A_t) + \varepsilon_0 + \delta_0 , \tag{5.35}$$

$$u_1(A_t) = (1 - \varepsilon_1)P_1(A_t) - \delta_1 . \tag{5.36}$$

Indeed, Lemma 4.1, (5.17), (5.18) show that

$$t \cdot v_0(A_t) - u_1(A_t) \to t_0 - 1 , \qquad \text{as} \quad t \downarrow t_0 . \tag{5.37}$$

Therefore, the assumption $v_0(A_{t_0}) = 1$ implies $u_1(A_t) \to 1$ as $t \downarrow t_0$, contradicting (3.2) and (5.33). $\inf_{t<t_1} u_1(A_t) = 0$ can be excluded by the same argumentation.

Since $w_{t_0}(A_{t_0}) = t_0 - 1$ and $v_0(A_{t_0}) < 1$ we furthermore must have $u_1(A_{t_0}) < 1$. Finally: $A_t \neq \varnothing \ \forall\, t < t_1$, by (5.34), completes the proof of (5.35), (5.36).

These equations now imply that

$$A_t \quad \text{minimizes} \quad \mu_t \qquad\qquad t_0 \leqq t < t_1 \tag{5.38}$$

where

$$\mu_t = t \cdot (1 - \varepsilon_0)P_0 - (1 - \varepsilon_1)P_1 .$$

If we take a $\Delta \in dP_1/dP_0$ and define $B_t = \{((1 - \varepsilon_1)/(1 - \varepsilon_0))\Delta > t\}$, (5.38) remains true with B_t instead of A_t, therefore

$$|\mu_t|(A_t \triangle B_t) = 0 \qquad\qquad t_0 \leqq t < t_1 \tag{5.39}$$

(where $|\mu_t|$ denotes the total variation of μ_t), hence also

$$(P_0 + P_1)\left(\pi \leqq t < \frac{1 - \varepsilon_1}{1 - \varepsilon_0}\,\Delta \ \text{ or } \ \pi \geqq t > \frac{1 - \varepsilon_1}{1 - \varepsilon_0}\,\Delta\right) = 0 \qquad t_0 \leqq t < t_1 . \tag{5.40}$$

In view of (5.32) and (5.34) we conclude

$$\pi = \frac{1 - \varepsilon_1}{1 - \varepsilon_0}\,(\Delta_0 \vee \Delta \wedge \Delta_1) \tag{5.41}$$

if we modify Δ suitably within $(P_0 + P_1)$-equivalence.

The other inclusion follows by (5.1)—(5.3), since in Section 6 each $((1 - \varepsilon_1)/(1 - \varepsilon_0))(\Delta_0 \vee \Delta \wedge \Delta_1)$, $\Delta \in dP_1/dP_0$, turns out to be a suitable version of the RND of some LFP. □

PROOF OF THEOREM 5.2. Suppose first that (Q_0, Q_1) has the properties (5.5)—(5.10). In order to show that $Q_0 \in \mathfrak{P}_0$ we consider the decomposition

$$Q_0(B) = Q_0(B, \Delta < \Delta_0) + Q_0(B, \Delta_0 \leqq \Delta \leqq \Delta_1) + Q_0(B, \Delta_1 < \Delta) \quad (B \in \mathfrak{B}) .$$

By (5.8)

$$Q_0(B, \Delta < \Delta_0) \leqq (1 - \varepsilon_0)P_0(B, \Delta < \Delta_0) , \tag{5.42}$$

and by (5.7)

$$Q_0(B, \Delta_0 \leqq \Delta \leqq \Delta_1) = (1 - \varepsilon_0)P_0(B, \Delta_0 \leqq \Delta \leqq \Delta_1) , \tag{5.43}$$

and by (5.9)

$$Q_0(\Omega\backslash B, \Delta_1 < \Delta) \geqq (1 - \varepsilon_0)P_0(\Omega\backslash B, \Delta_1 < \Delta) , \tag{5.44}$$

which in combination with (5.15) gives

$$Q_0(B, \Delta_1 < \Delta) \leqq (1 - \varepsilon_0)P_0(B, \Delta_1 < \Delta) + \varepsilon_0 + \delta_0 . \tag{5.45}$$

Adding up yields: $Q_0 \in \mathfrak{P}_0$. $Q_1 \in \mathfrak{P}_1$ is obtained similarly.

Next by (5.15), (5.7) we have

$$Q_0(\Delta > t) = (1 - \varepsilon_0)P_0(\Delta > t) + \varepsilon_0 + \delta_0 \quad \forall\, t \in [\Delta_0, \Delta_1) , \tag{5.46}$$

and by (5.16), (5.11)

$$Q_1(\Delta > t) = (1 - \varepsilon_1)P_1(\Delta > t) - \delta_1 \quad \forall\, t \in [\Delta_0, \Delta_1) . \tag{5.47}$$

Now, in view of (5.6), $\pi = ((1 - \varepsilon_1)/(1 - \varepsilon_0))(\Delta_0 \vee \Delta \wedge \Delta_1)$, $\Delta \in dP_1/dP_0$, is a version of dQ_1/dQ_0. Therefore, (5.46) and (5.47) establish $(Q_0, Q_1 \mid \pi)$ as an LFP. By (5.3), (5.4), this version π of dQ_1/dQ_0 is also necessary. Note that in general these π's do not constitute all possible versions of dQ_1/dQ_0!

On the other hand, let $(Q_0, Q_1 \mid \pi)$ be an LFP. By (5.3), (5.4) we have

$$\pi = \frac{1 - \varepsilon_1}{1 - \varepsilon_0}(\Delta_0 \vee \Delta \wedge \Delta_1) \quad \text{with some} \quad \Delta \in \frac{dP_1}{dP_0} \tag{5.48}$$

such that especially $Q_1(\Delta < \Delta_0) = ((1 - \varepsilon_1)/(1 - \varepsilon_0))\Delta_0 Q_0(\Delta < \Delta_0)$. Now, a strict inequality either in

$$Q_0(\Delta < \Delta_0) \geqq (1 - \varepsilon_0)P_0(\Delta < \Delta_0) - \delta_0 \quad \text{or in} \tag{5.49}$$

$$Q_1(\Delta < \Delta_0) \leqq (1 - \varepsilon_1)P_1(\Delta < \Delta_0) + \varepsilon_1 + \delta_1 \tag{5.50}$$

would imply: $\psi_0(\Delta_0) < 1$, contradicting Lemma 4.4. Thus (5.10) and (5.14) must hold, as well as (5.15), (5.16), which are proved by the same argumentation.

Using (5.10) and $Q_0 \in \mathfrak{P}_0$ again, we have for $B \subset \{\Delta < \Delta_0\}$:

$$\begin{aligned} Q_0(B) &= Q_0(\Delta < \Delta_0) - Q_0(\Omega\backslash B, \Delta < \Delta_0) \\ &\leqq (1 - \varepsilon_0)P_0(\Delta < \Delta_0) - \delta_0 - ((1 - \varepsilon_0)P_0(\Omega\backslash B, \Delta < \Delta_0) - \delta_0) , \quad \text{hence} \end{aligned}$$

$$Q_0(B) \leqq (1 - \varepsilon_0)P_0(B) \quad \forall\, B \subset \{\Delta < \Delta_0\} . \tag{5.51}$$

The same argumentation, using (5.14)—(5.16), yields

$$Q_1(B) \geqq (1 - \varepsilon_1)P_1(B) \quad \forall\, B \subset \{\Delta < \Delta_0\} \tag{5.52}$$

$$Q_0(B) \geqq (1 - \varepsilon_0)P_0(B) \quad \forall\, B \subset \{\Delta_1 < \Delta\} \tag{5.53}$$

$$Q_1(B) \leqq (1 - \varepsilon_1)P_1(B) \quad \forall\, B \subset \{\Delta_1 < \Delta\} . \tag{5.54}$$

For $B \subset \{\Delta_0 \leqq \Delta \leqq \Delta_1\}$ define $B' = B \cup \{\Delta < \Delta_0\}$, $B'' = B \cup \{\Delta_1 < \Delta\}$, then by (5.10) and $Q_0 \in \mathfrak{P}_0$:

$$\begin{aligned} Q_0(\Delta < \Delta_0) + Q_0(B) = Q_0(B') &\geqq (1 - \varepsilon_0)P_0(B') - \delta_0 \\ &= (1 - \varepsilon_0)P_0(\Delta < \Delta_0) - \delta_0 + (1 - \varepsilon_0)P_0(B) \\ &= Q_0(\Delta < \Delta_0) + (1 - \varepsilon_0)P_0(B) . \end{aligned}$$

Using (5.15) and B'' instead, yields $Q_0(B) \leqq (1 - \varepsilon_0)P_0(B)$, such that

$$Q_0(B) = (1 - \varepsilon_0)P_0(B) \quad \forall\, B \subset \{\Delta_0 \leqq \Delta \leqq \Delta_1\} \tag{5.55}$$

and analogously

$$Q_1(B) = (1 - \varepsilon_1)P_1(B) \quad \forall\, B \subset \{\Delta_0 \leqq \Delta \leqq \Delta_1\}\,. \tag{5.56}$$

In any case, Q_0 and Q_1, which are equivalent measures, are dominated by $(P_0 + P_1)$, hence by μ. If we express (5.51)—(5.56) by μ-densities, we get (5.7)—(5.9), (5.11)—(5.13). □

6. Construction of LFP's. In order to construct LFP's one has to solve the system of inequalities (5.8), (5.9), (5.12), (5.13) under the side conditions (5.10), (5.14)—(5.16). Because of (5.6) we may confine ourselves to the construction of a Q_0, hence to the solution of (5.8), (5.9), (5.10) and (5.15).

One way to fulfill (5.8) is to try a convex combination with a constant coefficient $\alpha \in [0, 1]$:

$$q_0 = (1 - \varepsilon_0)\left((1 - \alpha)p_0 + \alpha \frac{p_1}{\Delta_0}\right) \qquad \text{on} \quad \{\Delta < \Delta_0\}\,. \tag{6.1}$$

In view of (5.10), α should satisfy

$$(1 - \alpha)P_0(\Delta < \Delta_0) + \alpha\Delta_0^{-1}P_1(\Delta < \Delta_0) = P_0(\Delta < \Delta_0) - \omega_0\,. \tag{6.2}$$

Because of (4.11)

$$\alpha = \frac{\omega_0 \Delta_0}{\nu_1 + \omega_0 \Delta_0} \tag{6.3}$$

does it.

After a similar argument for q_0 on $\{\Delta_1 < \Delta\}$ we arrive at

$$\begin{aligned} q_0 &= \frac{1 - \varepsilon_0}{\nu_1 + \omega_0 \Delta_0}(\nu_1 p_0 + \omega_0 p_1) && \text{on} \quad \{\Delta < \Delta_0\} \\ &= (1 - \varepsilon_0)p_0 && \text{on} \quad \{\Delta_0 \leqq \Delta \leqq \Delta_1\} \\ &= \frac{1 - \varepsilon_0}{\nu_0 \Delta_1 + \omega_1}(\omega_1 p_0 + \nu_0 p_1) && \text{on} \quad \{\Delta_1 < \Delta\}\,. \end{aligned} \tag{6.4}$$

This is Huber's definition of an LFP, see Huber (1965), (1968).

Another definition is possible if, instead of a convex combination, we essentially take the bounds themselves in (5.8), (5.9). Determine $\Delta_{00} \in (0, \Delta_0)$ and $\Delta_{11} \in [\Delta_1, \infty)$ by

$$\Delta_{00} \cdot P_0(\Delta < \Delta_{00}) - P_1(\Delta < \Delta_{00}) = \omega_0 \Delta_{00} \tag{6.5}$$

$$P_1(\Delta_{11} < \Delta) - \Delta_{11} \cdot P_0(\Delta_{11} < \Delta) = \nu_0 \Delta_{11}\,. \tag{6.6}$$

Note that only the parameters of $\mathfrak{P}_0$ enter here, such that our Q_0 will be independent of ε_1, δ_1. (6.5) is equivalent to

$$\psi_0(\Delta_{00}) = \frac{\omega_0 \Delta_{00}}{\nu_1 + \omega_0 \Delta_{00}} \tag{6.7}$$

or

$$\tilde{\psi}_0(\Delta_{00}) = 1 \tag{6.8}$$

where $\tilde{\psi}_0$ is defined as ψ_0 according to (4.9) but with parameters $\tilde{\nu}_1 = 0$, $\tilde{\omega}_0 = \omega_0$. (6.6) is equivalent to

$$\psi_1(\Delta_{11}) = \frac{\nu_0 \Delta_{11}}{\nu_0 \Delta_{11} + \omega_1} \tag{6.9}$$

or

$$\tilde{\psi}_1(\Delta_{11}) = 1 \tag{6.10}$$

where $\tilde{\psi}_1$ is defined as ψ_1 according to (4.10) but with parameters $\tilde{\nu}_0 = \nu_0$, $\tilde{\omega}_1 = 0$.

It follows from (6.7), (6.9) and the monotonicity of ψ_0, ψ_1 that such solutions must satisfy $\Delta_{00} < \Delta_0$, $\Delta_{11} \geqq \Delta_1$. As $\lim_{t\to 0} \tilde{\psi}_1(t) = \infty$, $\lim_{t\to\infty} \tilde{\psi}_1(t) = 0$, Δ_{11} exists uniquely by the intermediate value theorem and the continuity and monotonicity properties of $\tilde{\psi}_1$, according to Lemma 4.4. As $\lim_{t\to 0} \tilde{\psi}_0(t) = 1/\omega_0 \cdot P_0(\Delta = 0)$, $\lim_{t\to\infty} \tilde{\psi}_0(t) = 1/\omega_0$, we obtain the unique existence of $\Delta_{00} \in (0, \Delta_0)$ under the assumption

$$P_0(\Delta = 0) < \omega_0\,. \tag{6.11}$$

In this case we define

$$\begin{aligned} q_0 &= (1-\varepsilon_0)\frac{p_1}{\Delta_{00}} && \text{on} \quad \{\Delta < \Delta_{00}\} \\ &= (1-\varepsilon_0)p_0 && \text{on} \quad \{\Delta_{00} \leqq \Delta \leqq \Delta_{11}\} \\ &= (1-\varepsilon_0)\frac{p_1}{\Delta_{11}} && \text{on} \quad \{\Delta_{11} < \Delta\} \end{aligned} \tag{6.12}$$

and (5.8), (5.9), (5.10), (5.15) are easily seen to be satisfied.

If

$$P_0(\Delta = 0) \geqq \omega_0 > 0 \tag{6.13}$$

we take

$$\begin{aligned} q_0 &= \left(1 - \varepsilon_0 - \frac{\delta_0}{P_0(\Delta = 0)}\right)p_0 && \text{on} \quad \{\Delta = 0\} \\ &= (1-\varepsilon_0)p_0 && \text{on} \quad \{0 < \Delta \leqq \Delta_{11}\} \\ &= (1-\varepsilon_0)\frac{p_1}{\Delta_{11}} && \text{on} \quad \{\Delta_{11} < \Delta\}\,. \end{aligned} \tag{6.14}$$

And if

$$\omega_0 = 0 \tag{6.15}$$

it suffices to take

$$\begin{aligned} q_0 &= (1-\varepsilon_0)p_0 && \text{on} \quad \{\Delta \leqq \Delta_{11}\} \\ &= (1-\varepsilon_0)\frac{p_1}{\Delta_{11}} && \text{on} \quad \{\Delta_{11} < \Delta\}\,. \end{aligned} \tag{6.16}$$

Acknowledgment. I would like to thank Professor H. Witting and A. Buja for their valuable comments and suggestions.

REFERENCES

HUBER, P. J. (1965). A robust version of the probability ratio test. *Ann. Math. Statist.* **36** 1753–1758.

HUBER, P. J. (1968). Robust confidence limits. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **10** 269–278.

HUBER, P. J. and STRASSEN, V. (1973). Minimax tests and the Neyman–Pearson lemma for capacities. *Ann. Statist.* **1** 251–263.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
DER ALBERT-LUDWIGS-UNIVERSITÄT
78 FREIBURG IM BREISGAU
HERMANN-HERDER-STRASSE 10
GERMANY