Z. Wahrscheinlichkeitstheorie verw. Gebiete
58, 397–405 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete
© Springer-Verlag 1981

On Solutions of Minimax Test Problems for Special Capacities

Tadeusz Bednarski

Institute of Mathematics, Polish Academy of Sciences, Wroclaw Branch

Summary. Solutions to minimax test problems between neighbourhoods generated by specially defined capacities are discussed. The capacities are superpositions of probability measures and concave functions, so the paper covers most of the earlier results of Huber and Rieder concerning minimax testing between ε-contamination and total variation neighbourhoods. It is shown that the Neyman-Pearson lemma for 2-alternating capacities, proved by Huber and Strassen, can be applied to test problems between noncompact neighbourhoods of probability measures. It turns out that the Radon-Nikodym derivative between the special capacities is usually a nondecreasing function of the truncated likelihood ratio of some probability measures.

1. Introduction

Strassen (1964, 1965) generalized the classical Neyman-Pearson lemma to 2-alternating Choquet's capacities defined on finite spaces. The well known extension of this result to polish spaces was given by Huber and Strassen (1973). In spite of its generality the result does not cover the important cases of ε-contamination, total variation and Prokhorov neighbourhoods if the probability space is not compact. Minimax problems of testing between ε-contamination and between total variation neighbourhoods were solved by Huber (1965, 1968) and then generalized by Rieder (1977). Huber states the solutions to his minimax test problems justifying their correctness. Proceeding in a more constructive way Rieder first finds a Radon-Nikodym derivative between his special capacities and then constructs least favourable pairs of distributions. This second step would be unnecessary if we knew that such pairs, as in Theorem 4.1 of Huber and Strassen (1973), exist. One should notice however that characterizations of least favourable pairs of distributions may be useful in certain problems, see Rieder (1980).

0044-3719/81/0058/0397/$01.80

The aim of this paper is to show that both the construction of the derivative and the existence of the least favourable pair of distributions can actually be inferred from Huber and Strassen (1973) results. We consider a large class of set functions, called here special capacities, which are superpositions of probability measures and concave functions. The combination of ε-contamination and total variation neighbourhoods is therefore covered. Only the pure ε-contamination case is not covered by our technique. Assumptions about capacities and probability measures we make are not restrictive from the practical point of view.

In Sect. 3 we define our special capacities and show that they are upper probabilities of some convex sets of probability measures. The Neyman-Pearson lemma for special capacities is proved in Sect. 4. It turnes out that usually the Radon-Nikodym derivative between special capacities is a nondecreasing function of truncated likelihood ratio of probability measures. This result is proved in Sect. 5. The fact may be of practical importance. A large class of neighbourhoods of probability measures considered in this paper varies in topological and measure theoretic properties while there is no essential difference between the structure of minimax solutions. This suggests that the technical convenience is a well justified factor in defining departures from hypothetical parametric models, at least as far as testing is concerned. The last section extends this statement and gives examples. One of them corresponds to Rieder's model.

2. Basic Facts and Definitions

In the sequel we shall always assume that Ω is a polish space, $\mathscr{B}$ is its Borel σ-field and $\mathscr{M}$ is the set of all probability measures on $\mathscr{B}$. For two sets of probability measures $\mathscr{P}_0 \subset \mathscr{M}$ and $\mathscr{P}_1 \subset \mathscr{M}$ we define their upper probabilities by

$$v_i(A) = \sup\{P(A)\colon P \in \mathscr{P}_i\}, \qquad i=0,1.$$

Following Huber and Strassen (1973) we shall say that a measurable function π is a Radon-Nikodym derivative between v_1 and v_0 if for every $t \geqq 0$ we have

$$t v_0(\pi > t) + v_1(\pi \leqq t) = \inf\{t v_0(A) + v_1(A^c)\colon A \in \mathscr{B}\}.$$

If in addition there exist a pair of distributions $(Q_0, Q_1) \in \mathscr{P}_0 \times P_1$ such that $\pi = dQ_1/dQ_0$ and $Q_0(\pi > t) = v_0(\pi > t)$, $Q_1(\pi \leqq t) = v_1(\pi \leqq t)$, it is called least favourable. Usually π cannot be uniquely determined via the pair (Q_0, Q_1). The existence of π and (Q_0, Q_1) in particular ensures that Neyman-Pearson tests for Q_0 against Q_1 are minimax for $\mathscr{P}_0$ against $\mathscr{P}_1$ with the same level and minimum power. In the case v_0, c_1 are 2-alternating capacities then π and (Q_0, Q_1) always exist (Huber and Strassen (1973)). Bednarski (1978) proves that the pair (Q_0, Q_1) forms then the least informative binary experiment in $\mathscr{P}_0 \times \mathscr{P}_1$ and the existence of least informative binary experiments is sufficient for a set of probability measures to be generated by a 2-alternating capacity. The 2-alternating capacity is a set function v from $\mathscr{B}$ to $[0,1]$, which satisfies the following conditions:

i) $0=v(\emptyset)\leqq v(A)\leqq v(B)\leqq v(\Omega)=1$ for $A\subset B$,

ii) $A_n\uparrow A\Rightarrow v(A_n)\uparrow v(A)$ for $\{A_n\}\subset\mathscr{B}$,

iii) $F_n\downarrow F\Rightarrow v(F_n)\downarrow v(F)$ for F_n closed,

iv) $v(A\cup B)+v(A\cap B)\leqq v(A)+v(B)$ for all A, B.

Basic properties of v are described in Choquet (1953/54), Huber and Strassen (1973) and in Bednarski (1978). The upper probability of total variation and ε-contamination neighbourhoods considered by Rieder (1977) satisfies in general all the conditions of 2-alternating capacity except iii). This fact precludes the possibility of direct application of Huber and Strassen (1973) main results.

3. Special Capacities

For a probability measure P Rieder (1977) defines his special capacity by the formula $u_0(A)=[(1-\varepsilon)P(A)+\varepsilon+\delta]\wedge 1$ if $A\neq\emptyset$ and $u_0(\emptyset)=0$, where $\varepsilon,\delta\geqq 0$, $\varepsilon+\delta<1$. The set function u_0 is therefore given by the superposition of P and a particular concave function. As we shall see the superposition of probability measures and concave functions gives set functions which are upper probabilities of sets of probability measures. Moreover many interesting minimax test problems between two such sets can be solved via the Neyman-Pearson lemma for 2-alternating capacities.

Let f be a concave function f: $[0,1]\to[0,1]$ such that $f(1)=1$. Denote by $\mathscr{F}$ the class of all such functions and by $\mathscr{F}_0$ the subclass of $\mathscr{F}$ which contains all continuous functions on $[0,1]$ vanishing in 0.

Definition 3.1. A set function $v\colon\mathscr{B}\to[0,1]$, $v(\emptyset)=0$ is said to be a special capacity if there is $f\in\mathscr{F}$ and $P\in\mathscr{M}$ so that $v(A)=f[P(A)]$ for all $A\varepsilon\mathscr{B}$, $A\neq\emptyset$.

The concavity of f implies (Theorem 10.1 of Rockafellar (1970)) that f is continuous on $(0,1)$. Since $f(1)=1$ we infere $f(x)\geqq x$ for all $x\in[0,1]$, f is nondecreasing and it is continuous in 1. Thus the special capacity always satisfies the conditions i) and ii), and by concavity of f also iv) is fulfilled. Rieder (1977) point out for his u_0 that it has all the properties of 2-alternating capacities except iii). This is also the case with our special capacities. The property iii) would require either a compact Ω and continuity of f on $[0,1]$ or $f\in\mathscr{F}_0$.

Lemma 3.1. *Let* $g\colon[0,1]\to[0,1]$. *For* $P\in\mathscr{M}$ *and* $A\neq\emptyset$ *define* $v_{gP}(A)=g[P(A)]$ *and put* $v_{gP}(\emptyset)=0$. *Then* v_{gP} *satisfies the conditions* i), ii), iv) *for every polish* Ω *and every* $P\in\mathscr{M}$ *iff* $g\in\mathscr{F}$.

Proof. The sufficiency was explained above. For the necessity take $\Omega=[0,1]$ and $P=\lambda$, the Lebesque measure on Ω. By i) and ii) we have $g(1)=1$, g is nondecreasing and left continuous. Applying iv) to the sets $A=[0,(x+y)/2]$, $B=[0,y]\cup[(x+y)/2,x]$, where $x,y\in[0,1]$, $x>y$, we obtain $[g(x)+g(y)]/2\leqq g[(x+y)/2]$. Therefore g is concave.

Similarly we prove,

Lemma 3.2. *The set function v_{gP} is Choquet's 2-alternating capacity for every polish Ω and every $P\in\mathcal{M}$ iff $g\in\mathcal{F}_0$.*

Now we shall consider sets of probability measures majorized by the special capacities. Further if necessary we shall index special capacities by functions from $\mathcal{F}$ and probability measures defining them.

Lemma 3.3. *Assume v is a special capacity defined by $f\in\mathcal{F}$ and $P\in\mathcal{M}$. Then for every $B\in\mathcal{B}$ there is $P_0\in\mathcal{M}$ such that $P_0(A)\leqq v(A)$ for all $A\in\mathcal{B}$ and $P_0(B)=v(B)$.*

Proof. If $f\in\mathcal{F}_0$ then v is a 2-alternating capacity and the lemma is a consequence of Lemma 2.5 of Huber and Strassen (1973). Assume then $f\notin\mathcal{F}_0$. In the case $P(B)=0$ and $B\neq\emptyset$ take any $\omega_0\in B$ and define

$$P_0(A)=\begin{cases} f[P(B)] & \text{if } A=\{\omega_0\}\\ (1-f[P(B)])\cdot P(A) & \text{if } A\subset\Omega\setminus\{\omega_0\},\end{cases}$$

which has the required property. If $P(B)>0$ take $f_0\in\mathcal{F}_0$ defined by the formula

$$f_0(x)=\begin{cases} f(x) & \text{if } x\in[P(B),1]\\ xf[P(B)]/P(B) & \text{if } x\in[0,P(B)].\end{cases}$$

Then $v_{fP}\geqq v_{f_0P}$ and $v_{fP}(B)=v_{f_0P}(B)$. Since v_{f_0P} is 2-alternating capacity, Lemma 2.5 of Huber and Strassen (1973) completes the proof.

Using again Lemma 2.4 of Huber and Strassen and arguing as above we obtain,

Lemma 3.4. *Let v_{fP} be a special capacity and let $F_1\subset F_2\subset\ldots\subset F_n$ be closed sets such that $P(F_1)>0$. Then there is a probability measure $P_0\leqq v_{fP}$ such that $P_0(F_i)=v_{fP}(F_i)$ for $i=1,2,\ldots,n$.*

4. The Neyman-Pearson Lemma for Special Capacities

Every special capacity defines a set $\mathcal{P}_v$ of all probability measures majorized by v,

$$\mathcal{P}_v=\{P\in\mathcal{M}: P(A)\leqq v(A) \text{ for all } A\in\mathcal{B}\}.$$

We shall say $\mathcal{P}_v$ is generated by v. By Lemma 3.3 the upper probability of $\mathcal{P}_v$ coincides with v. In this section we shall consider minimax test problems between sets generated by special capacities. Let $\mathcal{P}_0\subset\mathcal{M}$ and $\mathcal{P}_1\subset\mathcal{M}$ be generated by $v_{f_0P_0}$ and $v_{f_1P_1}$ respectively.

Theorem 4.1. *Assume P_0, $P_1\in\mathcal{M}$ are mutually absolutely continuous and let f_0, $f_1\in\mathcal{F}\setminus\mathcal{F}_0$ be such that there is $0<x_0<1$ so that $f_0(x_0)=f_1(x_0)=1$. Then there exist Radon-Nikodym derivative $\pi=dv_{f_1P_1}/dv_{f_0P_0}$ and a pair of probability measures $(Q_0,Q_1)\in\mathcal{P}_0\times\mathcal{P}_1$ such that for all $t\geqq0$*

$$Q_0(\pi>t)=v_{f_0P_0}(\pi>t),\qquad Q_1(\pi\leqq t)=v_{f_1P_1}(\pi\leqq t)$$

and π is a version of dQ_1/dQ_0.

Remarks. The absolute continuity requirement is not restrictive from the practical point of view for the commonly used parametric models have this property. The same corresponds to the second assumption which is fulfilled in particular by the combination of ε-contamination and total variation neighbourhoods.

Proof. The absolute continuity of P_0, P_1 implies that there is ε, $x_0 > \varepsilon > 0$ such that

$$P_0(A) \leqq \varepsilon \Rightarrow P_1(A^c) \geqq x_0$$

and

$$P_1(A) \leqq \varepsilon \Rightarrow P_0(A^c) \geqq x_0.$$

Define

$$g_i(x) = \begin{cases} f_i(x) & \text{if } x \in [\varepsilon, 1] \\ x f_i(\varepsilon)/\varepsilon & \text{if } x \in [0, \varepsilon) \end{cases}$$

and consider the following expressions, crucial for the construction of minimax solutions (see Huber and Strassen (1973)):

$$T_0(t) = \inf_{A \in \mathscr{B}} [t f_0[P_0(A)] + f_1[P_1(A^c)]],$$

$$T_1(t) = \inf_{A \in \mathscr{B}} [t g_0[P_0(A)] + g_1[P_1(A^c)]].$$

By the definition of g_i we always have $T_0(t) \geqq T_1(t)$ for $t \geqq 0$. In fact there is equality.

Suppose, by contradiction, that for a fixed t, $T_0(t) > T_1(t)$. Since $v_{g_0 P_0}$, $v_{g_1 P_1}$ are 2-alternating capacities, Lemma 3.1 of Huber and Strassen (1973) implies that there is $A_t \in \mathscr{B}$ for which the infimum in $T_1(t)$ is attained. The sharp inequality implies that either $P_0(A_t) < \varepsilon$ or $P_1(A_t^c) < \varepsilon$. If $P_0(A_t) < \varepsilon$ then $g_1[P_1(A_t^c)] = 1$ and we infer $P_0(A_t) = 0$. Thus taking $A_t = \emptyset$ we obtain the equality $T_0(t) = T_1(t)$ which is contradictory to our previous assumption. The same argument applies if $P_1(A_t^c) < \varepsilon$.

Let then $\{A_t\}_{t \geqq 0}$ be a decreasing family of measurable sets for which the infima in $T_1(t)$ are attained. Such a family, by Lemmas 3.1 and 3.2 of Huber and Strassen (1973), exists. Put $A_t = \Omega$ for all t for which $P_0(A_t) = 1$ and $A_t = \emptyset$ for all t for which $P_0(A_t) = 0$. This modification gives again a decreasing minimizing family which is also good for $v_{f_0 P_0}$, $v_{f_1 P_1}$. Therefore the variable $\pi(x) = \inf\{t : x \notin A_t\}$ and an arbitrary least favourable pair (Q_0, Q_1) for $v_{g_0 P_0}$, $v_{g_1 P_1}$ satisfies, by Theorem 4.1 of Huber and Strassen (1973), the statement of our theorem.

The above proof gives,

Corollary 4.1. *Under the assumption of Theorem* 4.1 *there exist 2-alternating capacities* $v'_0 \leqq v_{f_0 P_0}$ *and* $v'_1 \leqq v_{f_1 P_1}$ *and a variable* $\pi = dv'_1/dv'_0$ *such that* $\pi = dv_{f_1 P_1}/dv_{f_0 P_0}$. *Moreover every least favourable pair of distributions* (Q_0, Q_1) *for sets generated by* v'_0, v'_1 *is least favourable for the sets generated by* $v_{f_0 P_0}$ *and* $v_{f_1 P_1}$. *Every Radon-Nikodym derivative* $dv_{f_1 P_1}/d_{f_0 P_0}$ *equals* dv'_1/dv'_0.

One should notice that the minimax test problems between sets of probability measures generated by special capacities can always, under the assumptions of Theorem 4.1, be "reduced" to some minimax test problems between weakly compact sets of probability measures. The set $\mathscr{P}_{v_{fP}}$ must be weakly compact only if $f \in \mathscr{F}_0$.

The statement of Theorem 4.1 is trivially satisfied if the sets of probability measures generated by the special capacities $v_{f_0 P_0}$ and $v_{f_1 P_1}$ are not disjoint. Once the theorem is used for the purpose of minimax testing we would like to know whether the sets are disjoint or not. The following lemma gives a necessary and sufficient condition for this.

Lemma 4.1. *Let* P_0, P_1, f_0, f_1 *fulfill the conditions of Theorem* 4.1 *and let* $\mathscr{P}_0$ *and* $\mathscr{P}_1$ *be generated by the special capacities* $v_0 = v_{f_0 P_0}$ *and* $v_1 = v_{f_1 P_1}$ *respectively. Then* $\mathscr{P}_0 \cap \mathscr{P}_1 \neq \emptyset$ *if and only if* $v_0(A) \geqq 1 - v_1(A^c)$ *for every* $A \in \mathscr{B}$.

Proof. To prove the necessity let us, by contradiction, assume that there exist $B \in \mathscr{B}$ so that $v_0(B) < 1 - v_1(B^c)$. Then for every $P \in \mathscr{P}_0$ and $Q \in \mathscr{P}_1$ we obtain $P(B) < Q(B)$ which contradicts $\mathscr{P}_0 \cap \mathscr{P}_1 \neq \emptyset$. On the other hand it is easily seen that the statement of Theorem 4.1 implies

$$\inf_{A \in \mathscr{B}} [v_0(A) + v_1(A^c)] = \inf_{A \in \mathscr{B}} [Q_0(A) + Q_1(A^c)],$$

where (Q_0, Q_1) is the least favourable pair of distributions. By the assumption the above expression equals 1 therefore $Q_0 = Q_1$.

Using the Neyman-Pearson lemma for 2-alternating capacities and arguing as above we can see that the lemma holds for 2-alternating capacities without the additional assumptions needed for the special capacities.

5. The Radon-Nikodym Derivative for the Special Capacities

The construction of the derivative is important when one considers minimax test problems between sets of probability measures generated by capacities, see Huber (1965, 1968) and Rieder (1977, 1978). For the total variation and ε-contamination model the derivative is equal to the truncated likelihood ratio between the central probability measures. The result proved below describes the relation between the Radon-Nikodym derivative for special capacities $v_{f_0 P_0}$, $v_{f_1 P_1}$ and their central probability measures P_0 and P_1. As in Theorem 4.1 we assume the mutual absolute continuity of P_0 and P_1 and the existence of $x_0 \in (0, 1)$ so that $f_0(x_0) = f_1(x_0) = 1$. Take $v_0 = v_{f_0 P_0}$ and $v_1 = v_{f_1 P_1}$.

Theorem 5.1. *Assume that for every significance level* $\alpha \in (0, 1)$ *there is a non-randomized Neyman-Pearson test for testing* P_0 *against* P_1. *Then there is a non-decreasing function* $h \geqq 0$ *such that* $dv_1/dv_0 = h(dP_1/dP_0)$.

Proof. The existence of the derivative $\pi = dv_1/dv_0$ is guarantied by Theorem 4.1. By the same theorem for every $t \geqq 0$ the set $\{\pi > t\}$ minimizes

$$t f_0[P_0(\cdot)] + f_1[P_1(\cdot^c)]. \tag{5.1}$$

This expression is also minimized by the critical region of every Neyman-Pearson test of the level $P_0(\pi > t)$ for P_0 against P_1. Let $\{B_t\}_{t \geqq 0}$ denote a family of such critical regions and denote by D the set of all $t \geqq 0$ for which $f_0[P_0(B_t)] < 1$ and $f_1[P_1(B_t^c)] < 1$. Then for $t \in D$ we have $\{\pi > t\} = B_t$ $P_0 + P_1$ a.e. and so for such t's $\{\pi > t\} = \{dP_1/dP_0 > w(t)\}$ $P_0 + P_1$ a.e. for a function w. Therefore for $t_1, t_2 \in D$, $t_1 < t_2$, the condition $P_0 + P_1\{t_2 \geqq \pi > t_1\} > 0$ yields $w(t_1) < w(t_2)$. This implies that w can be modified to a nondecreasing function w_0 so that

$$\{\pi > t\} = \{dP_1/dP_0 > w_0(t)\} \quad P_0 + P_1 \text{ a.e. for every } t \in D$$

and every version dP_1/dP_0.

Let us take $w_0(t) = 0$ if $f_0[P_0(\pi > t)] = 1$ and $w_0(t) = \infty$ if $f_1[P_1(\pi \leqq t)] = 1$. Then if dP_1/dP_0 is an arbitrary everywhere finite version one can easily see that $A_t = \{dP_1/dP_0 > w_0(t)\}$ minimizes (5.1) for every $t \geqq 0$ and the family is decreasing. Hence the function $\pi_0(\omega) = \inf\{t\colon \omega \notin A_t\}$ is equal dv_1/dv_0. Since w_0 is nondecreasing we infere that π_0 is a nondecreasing function of dP_1/dP_0.

Remarks. Since π_0 and π can be approximated from below by simple functions equal $P_0 + P_1$ a.e. we obtain $\pi_0 = \pi$ $P_0 + P_1$ a.e. However the equality $\pi_1 = dv_1/dv_0$ $P_0 + P_1$ a.e. does not in general imply that π_1 is the derivative of v_1 with respect to v_0 (compare Rieder (1977) p. 918).

From the proof of the last theorem one can notice that the variable dv_1/dv_0 is usually a nondecreasing function of truncated likelihood ratio dP_1/dP_0. In the particular case when the functions f_0, f_1 are linear for these $x \in [0, 1]$ for which $f_i(x) < 1$, the derivative is proportional to the truncated likelihood ratio. This is seen with the models considered by Huber and Rieder. Truncation of the likelihood ratio is connected with the shape of f_0 and f_1. If there is $x_0 \in (0, 1)$ such that $f_0(x_0) = f_1(x_0) = 1$ then the derivative between the capacities is the function of the truncation of dP_1/dP_0.

In the very "regular" cases when f_0, f_1 are differentiable and the distribution functions $G(t) = P_1(dP_1/dP_0 \leqq t)$, $F(t) = P_0(dP_1/dP_0 \leqq t)$ are also differentiable the solution w_0 can be constructed by differentiation of $tf_0[1 - F(s)] + f_1[G(s)]$ with respect to s.

6. Examples, Discussion

First it will be demonstrated how the results of the previous sections can be applied to the ε-contamination and total variation model explored by Rieder (1977). For this we take $f_i(x) = [(1 - \varepsilon_i)x + \varepsilon_i + \delta_i] \wedge 1$, where $i = 0, 1$, $\varepsilon_i \geqq 0$, $\delta_i > 0$ and $\varepsilon_i + \delta_i < 1$.

Theorem 6.1. *Assume $v_0 = v_{f_0 P_0}$ and $v_1 = v_{f_1 P_1}$ generate disjoint sets of probability measures $\mathscr{P}_0$ and $\mathscr{P}_1$ and let P_0, P_1 be mutually absolutely continuous. Then:*

1) *for some uniquely determined constants $\Delta_0 \leqq \Delta_1$ and arbitrary version dP_1/dP_0 we have*

$$dv_1/dv_0 = \Delta_0 \vee (1 - \varepsilon_1)/(1 - \varepsilon_0)\, dP_1/dP_0 \wedge \Delta_1,$$

2) *there exists a pair* $(Q_0, Q_1) \in \mathscr{P}_0 \times \mathscr{P}_1$ *such that* dv_1/dv_0 *is a version of* dQ_1/dQ_0 *and* $Q_0(dv_1/dv_0 > t) = v_0(dv_1/dv_0 > t)$,

$$Q_1(dv_1/dv_0 \leqq t) = v_1(dv_1/dv_0 \leqq t) \quad \textit{for all } t \geqq 0.$$

Proof. Existence of the pair and the Radon-Nikodym derivative is guarantied by Theorem 4.1. To construct dv_1/dv_0 we shall use Corollary 4.1.

Consider the Bayes test problems between $\mathscr{P}_0$ and $\mathscr{P}_1$ and denote $T_\alpha(v_0, v_1) = \inf_{A \in \mathscr{B}} [\alpha v_0(A) + (1-\alpha) v_1(A^c)]$, $\alpha \in [0,1]$. Let for every $\alpha \in [0,1]$ $A_\alpha \in \mathscr{B}$ denote the set for which the infimum is attained. We always have $T_\alpha(v_0, v_1) < \alpha \wedge (1-\alpha)$. Let D be the set of all these α for which $T_\alpha(v_0, v_1) < \alpha \wedge (1-\alpha)$. Since $\mathscr{P}_0 \cap \mathscr{P}_1 = \emptyset$, by Lemma 4.1 we have that $1/2 \in D$ so that $D \neq \emptyset$. It is easily seen that D is a subinterval of $[0,1]$ with some endpoints we shall denote by $\alpha_0 \leqq \alpha_1$. For every $\alpha \in D$ the set A_α defines a Bayes test for testing P_0 against P_1 with priors

$$a = \alpha(1-\varepsilon_0)[\alpha(1-\varepsilon_0) + (1-\alpha)(1-\varepsilon_1)], \qquad 1-a.$$

Therefore $A_\alpha = \{(1-\varepsilon_1)/(1-\varepsilon_0)\, dP_1/dP_0 > \alpha/(1-\alpha)\}$ $P_0 + P_1$ a.e. It is straightforward to see that the family $\{A_\alpha\}$ defined by

$$A_\alpha = \begin{cases} \emptyset & \text{for } 1 \geqq \alpha \geqq \alpha_1 \\ (1-\varepsilon_1)/(1-\varepsilon_0)\, dP_1/dP_0 > \alpha/(1-\alpha) & \text{for } \alpha \in [\alpha_0, \alpha_1) \\ \Omega & \text{for } 0 \leqq \alpha < \alpha_0 \end{cases}$$

where dP_1/dP_0 is an arbitrary version of the Radon-Nikodym derivative, minimizes $\alpha v_0(A) + (1-\alpha) v_1(A^c)$ for every $\alpha \in [0,1]$. Hence

$$\pi(\omega) = \inf\{\alpha/(1-\alpha) : \omega \notin A_\alpha\} = [\Delta_0 \vee (1-\varepsilon_1)/(1-\varepsilon_0)\, dP_1/dP_0 \wedge \Delta_1]$$

is by Corollary 4.1 equal the derivative dv_1/dv_0 for which there exists a least favourable pair of distributions.

The constants $\Delta_i = \alpha_i/(1-\alpha_i)$, i=0,1, are determined by the equations

$$\alpha v_0[(1-\varepsilon_1)/(1-\varepsilon_0)\,\pi > \Delta_0] + (1-\alpha) v_1[(1-\varepsilon_1)/(1-\varepsilon_0)\,\pi \leqq \Delta_0] = 1-\alpha,$$
$$\alpha v_0[(1-\varepsilon_1)/(1-\varepsilon_0)\,\pi > \Delta_1] + (1-\alpha) v_1[(1-\varepsilon_1)/(1-\varepsilon_0)\,\pi \leqq \Delta_1] = \alpha,$$

equivalent to those obtained by Rieder (1977). The uniqueness of the solutions Δ_0, Δ_1 is provided by the fact that $T_\alpha(P_0, P_1)$ is strictly increasing on the interval $[0, 1/2]$ and on $[1/2, 1]$ it is strictly decreasing.

The neighbourhood generated by v_0 is not weakly compact if Ω is not compact and the probability measures majorized by v_0 need not be absolutely continuous with respect to P_0. The minimax test problem between $\mathscr{P}_0$ and $\mathscr{P}_1$ can be, by Corollary 4.1, reduced to a test problem between sets generated by 2-alternating capacities. These sets are weakly compact and probability measures belonging to them are absolutely continuous with respect to the central probability measures (see the proof of Theorem 4.1). There is a measurable function which is the Radon-Nikodym derivative for both problems.

Another example of a natural special capacity is given by the function $f(x) = cx \wedge 1$, $c > 1$. The set of probability measures generated by the 2-alternating capacity $v_{fP}(A)=cP(A)\wedge 1$ is weakly compact and the probability measures are absolutely continuous with respect to P. The neighbourhood $\mathscr{P}_{v_{fP}}$ can be viewed as a set of possible departures from P. These departures are all bounded by the function v_{fP} directly proportional to P. The construction of the Radon-Nikodym derivative for such capacities, carried as in the proof of Theorem 6.1, shows that the derivative is the same as in 1).

It should be noted that proofs to all the theorems given in this paper use as the main tool Lemma 3.2 of Huber and Strassen (1977). Considerations similar as those given to proofs of Theorems 5.1 and 6.1 can also be found in Rieder (1977), pages 916, 917.

Theorem 5.1 and remarks following it indicate that the effect of truncation of the likelihood dP_1/dP_0 is common to all the used capacities. The superposition of the "truncation" and a nondecreasing function does not seem to be so influential, at least for small samples. This may throw some light on a matter how to choose departures from postulated parameteric models. The capacities considered here do not cover the case of Prohorov neighbourhoods though, in the author's view, the approach presented may be helpful to solve minimax test problems for Prohorov neighbourhoods. A particular case of minimax testing between the Prohorov neighbourhoods of probability measures is considered by Osterreicher (1978).

References

Bednarski, T.: Binary experiments, minimax test and 2-altering capacities. Preprint **131** of the Institute of Mathematics of the Polish Academy of Sciences (1978). To appear in Ann. Stat.

Choquet, G.: Theory of capacities. Ann. Inst. Fourier **5**, 131–292 (1953/54)

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

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

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

Osterreicher, F.: On the contruction of least favourable pairs of distributions. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete. **43**, 49–55 (1978)

Rieder, H.: Least favourable pairs for special capacities. Ann. Statist. **5**, 909–921 (1977)

Rieder, H.: A robus asymptotic testing model. Ann. Statist. **6**, 1080–1094 (1978)

Rieder, H.: On local asymptotic minimaxity and admissibility in robust estimation. [Unpublished manuscript 1980]

Rockafellar, R.T.: Convex Analysis. Princeton, New Jersey: Princeton University Press 1970

Strassen, V.: Messfehler und Information. Z. Wahrscheinlichkeitstheorie und verw. Gebiete. **2**, 423–439 (1964)

Strassen, V.: The existence of probability measures with given marginals. Ann. Math. Statist. **36**, 423–439 (1965)

Received December 20, 1980; in revised form June 25, 1981