Werk

Titel: The Restricted Regularity of Admissible Minimax Criteria for many well known Para...

Autor: Brown, C.C.

Jahr: 1976

PURL: https://resolver.sub.uni-goettingen.de/purl?358794056_0023|log10

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

Metrika, Band 23, 1976, Seite 41–63. Physica-Verlag, Wien.

The Restricted Regularity of Admissible Minimax Criteria for many well known Parametric Statistical Problems

By *C.C. Brown*, Erlangen[1])

Summary: The stability of test selection criteria is a question that was raised by *LeCam* in 1964 [see *LeCam*, 1964], but which seems to have received very little subsequent attention. Roughly, the problem posed by *LeCam* was to decide whether a given criterion is regular at a given statistical problem, i.e. given that the criterion prescribes the test δ_0 for the statistical problem, (Λ_0, D_0, R_0), whether it prescribes a best δ which is nearly the same as δ_0 for any statistical problem, (Λ, D, R), near to (Λ_0, D_0, R_0).

A more precise formulation of *LeCam*'s question is one of the objectives of this paper. This involves defining a suitable topology on the domain of problems where the criterion in question is to be applied. Elementary minimax results of a general nature are proved, and counter examples are given to show that the assumptions used are not completely superfluous. The generality of these results seems to lend support to the authors opinion, that the method of formulation is suitable for treating minimax questions. We then consider the standard elementary problems of parametric statistics, (monotone likelihood ratio, general exponential family) and prove that the admissible minimax criterion is regular with respect to certain types of perturbations in these problems. The treatment is, in a sense, restricted and departs from *LeCam* in holding the risk function essentially constant while varying the parameter space. That the problem variations considered may be of some practical interest and that the general formulations used are not completely unrealistic is shown at the conclusion of the paper where we apply a restricted regularity theorem to obtain regularity results for the binomial case.

§ 1 Definitions and General Remarks

Let Θ be a parameter set in which, as we will always assume in what follows, a field of measurable sets is defined, and in which every one point subset is measurable. We denote by Θ^* the randomisation of Θ i.e. the set of all probability distributions on the measurable sets in Θ. Let, further, $\mathcal{D}$ be a set of decision functions δ on a sample space X and $R(\theta, \delta) \geqslant 0$ a finite risk function defined on $\Theta \times \mathcal{D}$, assumed to be measurable in θ for each $\delta \in \mathcal{D}$. If $\mathcal{D}$ is convex and, with respect to a topology T in $\mathcal{D}$, also compact, and if $R(\theta, \delta)$ is lower semicontinuous on $\mathcal{D}$ for each $\theta \in \Theta$, then the statistical problem $(\Theta^*, \mathcal{D}, R)$, considered as a two person game, is strictly determined, and there exists at least one minimax decision function $\delta_0 \in \mathcal{D}$, i.e. a δ_0 with

[1]) Dr. *C.C. Brown*, Mathematisches Institut der Universität Erlangen/Nürnberg, 852 Erlangen, Bismarckstr. 1 1/2.

$$\sup_{\theta} R(\theta, \delta_0) \leqslant \sup_{\theta} R(\theta, \delta) \tag{1.1}$$

for all $\delta \in \mathcal{D}$ (see, for example *Ferguson* [1967]). Since the set of minimax δ is closed in the $\mathcal{T}$-topology, an application of Zorn's lemma shows the existence of an admissible minimax $\delta_0 \in \mathcal{D}$, i.e. δ_0 is minimax and there is nothing in $\mathcal{D}$ that is better than δ_0; there is no $\delta \in \mathcal{D}, \theta_0 \in \Theta$ with

$$R(\theta, \delta_0) \geqslant R(\theta, \delta) \text{ for all } \theta \in \Theta, \tag{1.2}$$

$$R(\theta_0, \delta_0) \neq R(\theta_0, \delta).$$

Let $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ $(\Lambda_i \subset \Theta_i^*)$ be a family of statistical decision problems. A test selection criterion for this family can be considered to be a function Φ from the index set I into $\bigcup_{i \in I} \mathcal{D}_i$ with $\Phi(i) \in \mathcal{D}_i$ for each $i \in I$. For example, let $(\Theta^*, \mathcal{D}, R)$ be some fixed statistical problem. Define $\{(\Theta_i^*, \mathcal{D}_i, R_i)\}_{i \in I}$ to be the family of problems with $\Theta_i \subset \Theta$, $\mathcal{D}_i = \mathcal{D}$ and R_i equal to the restriction of R to $\Theta_i \times \mathcal{D}_i$. A selection criterion can then be any function from I into $\mathcal{D}$. Of special interest in this paper is the case where $\Phi(i)$ is an admissible minimax decision function for the problem $(\Lambda_i, \mathcal{D}_i, R_i)$. We will call Φ an admissible minimax criterion in this case. More generally, Φ is a minimax criterion if $\Phi(i)$ is minimax.

A useful test selection criterion (see *LeCam* [1964] § 7) should have a certain minimal degree of regularity in order to insure that two problems which are practically indistinguishable from each other have, as images under Φ, decision functions which are practically the same. A familiar problem $(\Theta^*, \mathcal{D}, R)$ where the family $\{(\tilde{\Theta}^*, \mathcal{D}, R) | \tilde{\Theta} \subset \Theta\}$ has many Minimax criteria, all of which are non-regular at $(\Theta^*, \mathcal{D}, R)$ is given in *Brown* [w.y.].

In the present paper we will show that there is at least one fairly natural way to define regularity of test criteria so that, for the well known traditional "solved" problems of parametric statistics, the solutions (they are invariably admissible minimax) correspond to admissible minimax criteria which are, at least in a restricted sense, regular at the problem in question. Furthermore, every admissible minimax criterion is regular there. More precisely, considering such a parametric problem $(\Theta^*, \mathcal{D}, R)$, we define $R(\lambda, \delta) : \Theta^* \times \mathcal{D} \to R^1$ by the usual formula,

$$R(\lambda, \delta) = \int R(\theta, \delta) \lambda(d\theta), \tag{1.3}$$

and let the family **F** of problems be of the form $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ where Λ_i as a convex subset of Θ^*, $\mathcal{D}_i = \mathcal{D}$ and R_i is R restricted to $\Lambda_i \times R_i$ $(\Lambda_i \times \mathcal{D})$. Then there is an admissible minimax criterion $\Phi : I \to \mathcal{D}$ which is regular at $(\Theta^*, \mathcal{D}, R)$, and every admissible minimax criterion Φ is regular there.

The family **F** defined here is admittedly rather restricted. Proofs of regularity in a larger family are equally desirable, but they will be postponed for a later paper. Even under the present restrictions on **F**, the results proved here might be practically useful. At the conclusion of this paper, we will show how they can be applied to a binomial

distribution problem to arrive, by means of a limiting process, at the decision function that is generally accepted as being appropriate. It will be seen that in this, as in many other problems, the mere requirement that Φ be regular minimax determines the decision function for the problem in question uniquely.

The final results proved here are of a special nature, but the initial ones will be more general, making use of far fewer assumptions on the family **F** of problems under consideration. We must define what we mean by regularity of a criterion. Technically we will mean nothing but continuity with respect to a topology on **F** which we will presently define. In doing so, we deal exclusively with families $\mathbf{F} = \{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ where the existence of some pair $(\Lambda, \mathcal{D})$ of uniform topological spaces such that $\Lambda_i \subset \Lambda$ and $\mathcal{D}_i \subset \mathcal{D}$ for all $i \in I$ is assured. This condition makes comparison of problems with one another possible. It appears to be a rather clumsy assumption, but there seems to be a justification for it, consisting mainly of a slightly more elegant formulation from which the condition can be proved. We indicate, briefly, how this formulation may look in practice.

Consider a family $\mathbf{F} = \{(\Theta_i^*, \mathcal{D}_i, R_i)\}_{i \in I}$, and let A be a family $\{A_{ij}\}$ of injections, $A_{ij} : \Theta_j \to \Theta_i$ with A_{ii} equal to the identity in Θ_i. We assume that the measurable sets in Θ_j are exactly the sets of the form $A_{ij}^{-1}(E)$ with E measurable in Θ_i. We require that A be transitive:

$$A_{ij} \in \mathsf{A}, A_{jk} \in \mathsf{A} \Rightarrow A_{ij} A_{jk} \in \mathsf{A}, \tag{1.4}$$

and we don't exclude the possibility that A contain only identity mappings. We define the partial ordering

$$j \leqslant i \Leftrightarrow \exists\, A_{ij} \in \mathsf{A}.$$

A special case of interest is one where the mapping A_{ij} exists iff $\Theta_j \subset \Theta_i$ and is equal to the identity on Θ_j. In general, the mappings of A induce a class of mappings, A^*, from Θ_j^* into Θ_i^* by means of the equation,

$$(A_{ij}^* \lambda)(f) = \lambda(f \cdot A_{ij}), \tag{1.5}$$

where $\lambda \in \Theta_j^*$ and f is any bounded measurable function on Θ_i. The mapping, A_{ij}, can be considered a positive linear mapping on the cone F_j^* generated by Θ_j^* into the cone F_i^* generated by Θ_i^*.

The assumption that there exist an all encompassing set Λ is now replaced by the assumption that the index set, I, be directed in the partial ordering induced by A. We construct something resembling an inductive limit. In the direct sum $\bigoplus_i F_i^*$, let the elements $\sum_{\alpha=1}^{n} \mu_{i_\alpha}$ and $\sum_{\alpha=1}^{n} \nu_{i_\alpha}$ be equivalent if there exists a j with $j \geqslant i_\alpha$ $(\alpha = 1, 2, 3, \ldots, n)$ such that

$$\sum_{\alpha=1}^{n} A_{ji_\alpha}^* \mu_{i_\alpha} = \sum_{\alpha=1}^{n} A_{ji_\alpha}^* \nu_{i_\alpha}.$$

We select, from $\bigoplus_i F_i^*$, the elements $\sum_{\alpha=1}^{n} \mu_{i_\alpha}$ which have the property that there exists a $j \geqslant i_\alpha$ $(\alpha = 1, 2, \ldots, n)$ with

$$\sum_{\alpha=1}^{n} (A^*_{ji_\alpha} \mu_{i_\alpha})(1) = 1. \tag{1.7}$$

The equivalence classes in $\bigoplus_i F_i^*$ of these elements are the elements of the desired maximal set Λ for the family $\{\Theta_i^*\}$, the natural mapping from Θ_i^*, into Λ being an imbedding since the A_{ij}^* are injections.

The introduction of a uniform structure into Λ can, under plausible conditions, be done using the uniform structures already decided upon for the Θ_i^*. Suppose that, for each $i, j \in I; j \leqslant i$, the uniform structure on Θ_j^* is exactly the set of entourages of the form,

$$A_{ij}^{*-1} u := \{(\lambda, \lambda') \in \Theta_j^* \times \Theta_j^* \mid (A_{ij}^* \lambda, A_{ij}^* \lambda') \in u\},$$

where u can be any entourage in $\Theta_i^* \times \Theta_i^*$. An entourage of Λ can be identified with a family $\{u_i\}_{i \in I}$ of entourages u_i of Θ_i^* $(i \in I)$ where we require that $u_j = A_{ij}^{*-1} u_i$ for $j \leqslant i$. The equivalence classes in Λ represented by

$$\sum_{\alpha=1}^{n} \mu_{i_\alpha} \quad \text{and} \quad \sum_{\alpha=1}^{n} \nu_{i_\alpha}$$

are close of order $\{u_i\}_{i \in I}$ if there is some $j \geqslant i_\alpha$ $(\alpha = 1, 2, \ldots, n)$ with

$$\left(\sum_{\alpha=1}^{n} A^*_{ji_\alpha} \mu_{i_\alpha}, \sum_{\alpha=1}^{n} A^*_{ji_\alpha} \nu_{i_\alpha} \right) \in u_j \subset \Theta_j^* \times \Theta_j^*.$$

The uniform structure in Λ thus defined induces in Θ_i^* $(i \in I)$ the uniform structure already assumed for it.

For the construction of $\mathcal{D}$, the mappings corresponding to the A_{ij}^* above would have a different origin, perhaps springing from a suitable transitive family of injections for the class $\{D_i\}$ of decision sets together with a suitable transitive family of measurable mappings B_{ij} from X_j, the sample space of $(\Lambda_j, \mathcal{D}_j, R_j)$, onto X_i, the sample space of $(\Lambda_i, \mathcal{D}_i, R_i)$.

There appear to be good reasons for constructing Λ (or $\mathcal{D}$) in the above way. For one, it is hard to see how a Λ for the general family $\{\Theta_i^*\}_{i \in I}$ could otherwise be defined so as to make suitable reference to the already existing set theoretical semi-ordering in the special case, $j \leqslant i \Leftrightarrow \Theta_j \subset \Theta_i$. The construction also appears to offer the possibility of identifying problems in with one another if they are, in a reasonable sense, equivalent.

We return, now, to the more immediate problems at hand and simply assume that the uniform spaces Λ and $\mathcal{D}$ have already been constructed.

Consider the general family, $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$. Each element of this family can be represented by the graph of R_i, considered as a subset of the product space $\Lambda \times \mathcal{D} \times R^1$, where R^1 denotes euclidean one-space. We say that two problems are

nearly the same when their graphs are. To give this convention a meaning, we define a topology on the power set of $E := \Lambda \times \mathcal{D} \times R^1$. Consider the product uniform structure for $\Lambda \times \mathcal{D} \times R^1$ generated by the euclidean structure of R^1 together with those of Λ and of $\mathcal{D}$. The entourages in $E \times E$ belonging to the product structure will be denoted by u, v, w etc. Denoting the subsets of E by capitals (S, T, M etc.), we define

$$U_u := \{(S, T) \mid S \subset u(T),\ T \subset u^{-1}(S)\}$$

to be the entourage in $\mathcal{P}(E) \times \mathcal{P}(E)$ induced by u in $E \times E$, and consider the class $\{U_u\}$ to be the system of entourages for $\mathcal{P}(E)$[2]). We say that S is close of order u to T if $(S, T) \in U_u$. That $\{U_u\}$ is a uniform structure follows from the fact that $U_u^{-1} = U_{u^{-1}}$ and the easily verified relations, $U_u \subset U_v$ when $u \subset v$, and $U_u U_v \subset U_{uv}$.

The topology generated by $\{U_u\}$ is not in general hausdorff, even if the system $\{u\}$ generates a hausdorff-topology in E.

There is already some precedent for choosing uniform structures on the Θ_i^*. For example, one can start with the distance function ρ_4 of *Wald* [1950, § 3.3] defined on Θ_i, and then consider the family L of continuous functions $\{f\}$ defined on Θ_i having the property

$$|f(\theta_1) - f(\theta_2)| \leqslant \rho_4(\theta_1, \theta_2) \tag{1.8}$$

for all θ_1 and θ_2 in Θ_i. Let the measurable sets in Θ_i be those belonging to the Borel field $\mathcal{B}$ generated by the ρ_4 topology. Since, for any closed set $E \subset \Theta_i$, the function

$$\rho_4(x, E) = \inf_{y \in E} \rho_4(x, y)$$

is in L, the Borel field $\mathcal{B}$ is also the smallest σ-field in Θ_i with respect to which the functions in L are measurable. If ρ_4 is a metric on Θ_i, it can be extended to the set Θ_i^* of all probability measures on $\mathcal{B}$ by the formula,

$$\rho_4(\lambda_1, \lambda_2) := \sup_{f \in L} \left| \int f(\theta)\, \lambda_1(d\theta) - \int f(\theta)\, \lambda_2(d\theta) \right| \tag{1.9}$$

for each pair (λ_1, λ_2) of elements from Θ_i. That $\rho_4(\lambda_1, \lambda_2) = 0$ implies that $\lambda_1 = \lambda_2$ is easily seen by using the uniqueness part of the theorem of *Daniell-Stone* applied to the *Stone* vector lattice of functions, g, for which there is a real number M_g with

$$|g(\theta_1) - g(\theta_2)| \leqslant M_g \rho_4(\theta_1, \theta_2)$$

for all $\theta_1, \theta_2 \in \Theta_i$.

---

[2]) A special case is the hausdorff metric. See also *Heyer* [1973, § 20].

We will call such an extension of a bounded metric on a set Θ to Θ^* its natural extension. Let Θ_0 be a subset of Θ and ρ a bounded metric on Θ. If ρ_0 is the restriction of ρ to Θ_0, it can be extended to Θ_0^* in more than one way. One possibility would be to extend ρ_0 directly as above, using the class L_0 of functions on Θ_0 having the property (1.8) with respect to ρ_0 on Θ_0. Another would be to imbed Θ_0 in Θ by means of the (measurable) identity mapping and then to restrict the natural extension of ρ (seen as a metric on Θ) to the set Θ_0^*. That these two possibilities give the same metric on Θ_0^* follows from the fact that any function satisfying (1.8) on Θ_0 can be extended to one satisfying (1.8) on Θ. If f_0 is such a function on Θ_0, then such an extension, f, on Θ is given by the formula,

$$f(\theta) = \sup_{\theta' \in \Theta_0} \{f_0(\theta') - \rho(\theta, \theta')\} \tag{1.10}$$

§ 2 Results of a General Nature

We return to the general situation where there are uniform topological spaces Λ and $\mathcal{D}$ containing the spaces Λ_i and $\mathcal{D}_i$ connected with the family $\mathbf{F} = \{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$. If $\nu \subset E \times E$ ($E = \Lambda \times \mathcal{D} \times R^1$) is an entourage in $E \times E$, then we will say that the problem, $(\Lambda_1, \mathcal{D}_1, R_1)$ is close of order ν to $(\Lambda_2, \mathcal{D}_2, R_2)$ if the graph of R_1 is close of order ν to the graph of R_2.

In many parametric problems, the risk function has the property,

$$|R(\theta_1, \delta) - R(\theta_2, \delta)| \leqslant M \rho_4(\theta_1, \theta_2) \tag{2.1}$$

for all θ_1, θ_2 in Θ and all δ, where $M \geqslant 0$ is fixed. When this is so, then it is easily verified that $R(\lambda, \delta)$ has a similar property in terms of the natural extension $\rho_4(\lambda_1, \lambda_2)$. This is a special case of the more general situation, now to be defined.

Definition 1: Let $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I} = \mathbf{F}$ be a family of problems with $\Lambda_i \subset \Lambda$, $\mathcal{D}_i \subset \mathcal{D}$ and R_i defined on $\Lambda_i \times \mathcal{D}_i$. We say that $\mathbf{F}$ is statistically regular at $(\Lambda_0, \mathcal{D}_0, R_0) \in \mathbf{F}$ if, for any $\epsilon > 0$ and any $\delta_0 \in \mathcal{D}$, there are entourages $w \subset E \times E$, $u_1 \subset \Lambda \times \Lambda$ and $u_2 \subset \mathcal{D} \times \mathcal{D}$ such that, for all $i \in I$ with $(\Lambda_i, \mathcal{D}_i, R_i)$ close of order w to $(\Lambda_0, \mathcal{D}_0, R_0)$,

$$\left.\begin{array}{l}(\lambda_1, \lambda_2) \in u_1 \cap \Lambda_i \times \Lambda_i \\ (\delta, \delta_0) \in u_2 \cap \mathcal{D}_i \times \mathcal{D}_i\end{array}\right\} \Rightarrow |R_i(\lambda_1, \delta) - R_i(\lambda_2, \delta)| < \epsilon.$$

We now prove

Theorem 1: Let $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I} = \mathbf{F}$ be a family, statistically regular at $(\Lambda_0, \mathcal{D}_0, R_0) \in \mathbf{F}$. If, for every entourage $\nu \subset E \times E$, the problem $(\Lambda_\nu, \mathcal{D}_\nu, R_\nu) \in \mathbf{F}$ is close of order ν to $(\Lambda_0, \mathcal{D}_0, R_0) \in \mathbf{F}$ and if all of the problems $(\Lambda_\nu, \mathcal{D}_\nu, R_\nu)$ and $(\Lambda_0, \mathcal{D}_0, R_0)$ are strictly determined with minimax values V_ν and V_0 respectively, then $\lim_\nu \sup V_\nu \leqslant V_0$ in the sense of the net $\{V_\nu\}$.

Proof: If the theorem were false, then there would be a $\Delta > 0$ and a subnet $\{V_{\bar{\nu}}\}$ with $V_{\bar{\nu}} \geqslant V_0 + \Delta$ for all $\bar{\nu}$. By strict determinateness, we have

$$\sup_{\Lambda_{\bar{\nu}}} \inf_{\mathcal{D}_{\bar{\nu}}} R_{\bar{\nu}}(\lambda, \delta) \geqslant \inf_{\mathcal{D}_0} \sup_{\Lambda_0} R_0(\lambda, \delta) + \Delta$$

for all $\bar{\nu}$. Letting $\epsilon > 0$, there exists a fixed $\delta_0 \in \mathcal{D}_0$ and, for each $\bar{\nu}$, a $\lambda'_{\bar{\nu}} \in \Lambda_{\bar{\nu}}$ with

$$(*) \quad R_{\bar{\nu}}(\lambda'_{\bar{\nu}}, \delta') + \epsilon \geqslant R_0(\lambda, \delta_0) - \epsilon + \Delta$$

for arbitrary $\delta' \in \mathcal{D}_{\bar{\nu}}$ and $\lambda \in \Lambda_0$. By regularity of **F**, let $w \subset E \times E, u_1 \subset \Lambda \times \Lambda$ and $u_2 \subset \mathcal{D} \times \mathcal{D}$ be the entourages of definition I. We can assume that $w = w_1 \times w_2 \times w_3$ is symmetric with $w_1^2 \subset u_1$, $w_2 \subset u_2$ and $w_3 \subset \{(r_1, r_2) \in R^2 \mid |r_1 - r_2| < \epsilon\}$. Let $\bar{\nu}$ be an entourage in $E \times E$ contained in w. Then the problem $(\Lambda_{\bar{\nu}}, \mathcal{D}_{\bar{\nu}}, R_{\bar{\nu}})$ is close of order w to $(\Lambda_0, \mathcal{D}_0, R_0)$. There exists, therefore, $\lambda_{\bar{\nu}} \in w_1(\lambda'_{\bar{\nu}}) \cap \Lambda_0$, and we let $\lambda = \lambda_{\bar{\nu}}$ in the inequality (*). Further, by w-proximity, let $(\lambda''_{\bar{\nu}}, \delta''_{\bar{\nu}}) \in \Lambda_{\bar{\nu}} \times \mathcal{D}_{\bar{\nu}}$ be an element of $(w_1 \times w_2)(\lambda_{\bar{\nu}}, \delta_0)$ with $R_0(\lambda_{\bar{\nu}}, \delta_0) \geqslant R_{\bar{\nu}}(\lambda''_{\bar{\nu}}, \delta''_{\bar{\nu}}) - \epsilon$. Then

$$R_{\bar{\nu}}(\lambda'_{\bar{\nu}}, \delta) + \epsilon \geqslant_{\nu} R_{\bar{\nu}}(\lambda''_{\bar{\nu}}, \delta''_{\bar{\nu}}) - 2\epsilon + \Delta$$

for all $\delta' \in \mathcal{D}_{\bar{\nu}}$. Choose $\delta' = \delta''_{\nu}$. Since $\lambda''_{\bar{\nu}} \in w_1(\lambda_{\bar{\nu}}) \subset w_1(w_1(\lambda'_{\bar{\nu}})) \subset u_1(\lambda'_{\bar{\nu}})$ and $\delta''_{\bar{\nu}} \in w_2(\delta_0) \subset u_2(\delta_0)$, statistical regularity gives

$$R_{\bar{\nu}}(\lambda''_{\bar{\nu}}, \delta''_{\bar{\nu}}) \geqslant R_{\bar{\nu}}(\lambda'_{\bar{\nu}}, \delta''_{\bar{\nu}}) - \epsilon$$

and hence

$$R_{\bar{\nu}}(\lambda''_{\bar{\nu}}, \delta''_{\bar{\nu}}) + 2\epsilon \geqslant R_{\bar{\nu}}(\lambda''_{\bar{\nu}}, \delta''_{\bar{\nu}}) - 2\epsilon + \Delta.$$

Therefore, $4\epsilon > \Delta$ for all $\epsilon > 0$ which contradicts $\Delta > 0$.

The hypothesis that **F** be statistically regular cannot be simply dropped altogether from the statement of theorem I, as the following example shows. Consider the family $\{P_\theta\}$ of normal distributions,

$$P_\theta(dx) = \frac{1}{\sqrt{2\pi}} \exp \quad -\frac{(x-\theta)^2}{2} \quad , \quad (\theta \geqslant 0)$$

on the real line. We define a parameter space, Θ, to be the union of $H_0 := \{\theta = 0\}$ and the open half line, $\{\theta \in R^1 \mid \theta > 0\}$. Correspondingly, we define a loss function, $L(\theta, i)$ on $\Theta \times \{0,1\}$ by

$$L(\theta, 0) = \begin{cases} 0 & \text{if } \theta \in H_0 \\ 1 & \text{otherwise} \end{cases}$$

$$L(\theta, 1) = \begin{cases} 1 & \text{if } \theta \in H_0 \\ l(\theta) & \text{otherwise,} \end{cases}$$

where the function $l(\theta)$ must be further specified. Considering the statistical test problem with Θ as parameter space and with set $\mathcal{D}$ of decision functions equal to the set of critical functions $\phi(x)$ on R^1, the risk function, $R(\theta, \phi)$, is given by

$$R(\theta, \phi) = \begin{cases} \int \phi P_\theta(dx) & \text{if } \theta \in H_0 \\ (1 - l(\theta)) \int (1 - \phi) P_\theta(dx) + l(\theta) & \text{otherwise.} \end{cases}$$

We introduce the topology on $\mathcal{D}$ generated by neighborhoods of the form,

$$\{\phi \in \mathcal{D} \mid \left| \int (\phi - \phi_0) P_\theta(dx) \right| < \epsilon\}, \quad (\phi_0 \in \mathcal{D}, \theta \in \Theta, \epsilon > 0).$$

A contradiction to the conclusion of theorem I is obtained by first choosing a distance ρ on Θ equal to *Wald*'s ρ_4, using constant $l(\theta)$ (= 0). On H_0^c, this is *Wald*'s ρ_1;

$$\rho_1(\theta, \theta') := \sup_E |P_\theta(E) - P_{\theta'}(E)|, \quad (\theta, \theta' \text{ in } H_0^c),$$

and $\rho(\theta, 0) = \rho_1(\theta, 0) + 1$ if $\theta \neq 0$. Now choose $l(\theta)$ equal to $\sum_{i=2}^{\infty} l_i(\theta)$ where, for integral $i \geqslant 2$,

$$l_i(\theta) = \max[1 - i\rho_1(\theta, i), 0], \quad (\theta \in H_0^c).$$

The function $l(\theta)$ so defined is continuous, but not uniformally continuous, on H_0^c with respect to the metric ρ. Correspondingly, the one problem family $\{(\Theta^*, \mathcal{D}, R)\}$ is not statistically regular at $(\Theta^*, \mathcal{D}, R)$. This latter fact is the basis for constructing an irregular family. For each integer $i > 0$, we define a problem $(\Theta_i^*, \mathcal{D}_i, R_i)$ as follows: $\mathcal{D}_i = \mathcal{D}$, $\Theta_1 = \Theta$, $\Theta_{i+1} = \Theta_i - \{\theta > 0 | l_{i+1} > 0\}$ for all $i > 0$, and $R_i(\theta, \phi)$ is $R(\theta, \phi)$ restricted to $\Theta_i \times \mathcal{D}_i$. We set $\Theta_0 = \bigcap_{i>0} \Theta_i$ (= $\{\theta > 0 | l(\theta) = 0\} \cup \{\theta = 0\}$), and define the problem, $(\Theta_0^*, \mathcal{D}_0, R_0)$ by $\mathcal{D}_0 = \mathcal{D}$ and $R_0 = R$ restricted to $\Theta_0 \times \mathcal{D}_0$. The one problem family $\{(\Theta_0^*, \mathcal{D}_0, R_0)\}$, is statistically regular at $(\Theta_0^*, \mathcal{D}_0, R_0)$, but the family, $\{(\Theta_i^*, \mathcal{D}_i, R_i)\}_{i \in I}$ where I is the set of non-negative integers, is not. The sequence, $\{(\Theta_i^*, \mathcal{D}_i, R_i)\}_{i>0}$, converges to $(\Theta_0^*, \mathcal{D}_0, R_0)$ in the sense already defined. All of the problems considered are strictly determined with values $V_i = 1$ for $i > 0$ and $V_0 = \frac{1}{2}$. This violates the conclusion of theorem I. Using similar ideas, one can construct a family **F** in which each problem has a risk function which uniformally continuous in θ, but for which the conclusion of theorem I is false.

In order to prove the next result, we assume regularity as seen from the point of view of statistician's opponant.

Definition II: Let $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I} = \mathbf{F}$ be a family of problems with $\Lambda_i \subset \Lambda$, $\mathcal{D}_i \subset \mathcal{D}$ and R_i defined on $\Lambda_i \times \mathcal{D}_i$. We say that **F** is opponant regular at $(\Lambda_0, \mathcal{D}_0, R_0) \in \mathbf{F}$ if,

for any $\epsilon > 0$ and any $\lambda_0 \in \Lambda$, there are entourages $w \subset E \times E$, $u_1 \subset \Lambda \times \Lambda$ and $u_2 \subset \mathcal{D} \times \mathcal{D}$ such that, for all $i \in I$ with $(\Lambda_i, \mathcal{D}_i, R_i)$ close of order w to $(\Lambda_0, \mathcal{D}_0, R_0)$,

$$\left.\begin{array}{l}(\lambda, \lambda_0) \in u_1 \cap \Lambda_i \times \Lambda_i \\ (\delta_1, \delta_2) \in u_2 \cap \mathcal{D}_i \times \mathcal{D}_i\end{array}\right. \Rightarrow |R_i(\lambda, \delta_1) - R_i(\lambda_0, \delta_2)| < \epsilon.$$

In this definition, we have just interchanged the roles of Λ and $\mathcal{D}$ in the definition of statistical regularity.

Theorem II: Let $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I} = \mathbf{F}$ be a family, opponant regular at $(\Lambda_0, \mathcal{D}_0, R_0) \in \mathbf{F}$. Let, for all $i \in I$, $\bar{V}_i$ be the minimax value of $(\Lambda_i, \mathcal{D}_i, R_i)$. Suppose that, for every entourage $v \subset E \times E$, the problem $(\Lambda_v, \mathcal{D}_v, R_v) \in \mathbf{F}$ is close of order v to $(\Lambda_0, \mathcal{D}_0, R_0)$. If $\delta_v \in \mathcal{D}_v$ for all v and if $\mathcal{D}_0$ is closed in $\mathcal{D}$, then every cluster point δ_0 of the net $\{\delta_v\}$ is in $\mathcal{D}_0$. If δ_v is minimax for $(\Lambda_v, \mathcal{D}_v, R_v)$, then $\lim_v \sup \bar{V}_v \leqslant \bar{V}_0$ implies that any cluster point, δ_0, of $\{\delta_v\}$ with $\delta_0 \in \mathcal{D}_0$ is minimax for $(\Lambda_0, \mathcal{D}_0, R_0)$.

Proof: The first conclusion of the theorem follows from $\delta_v \in \mathcal{D}_v \to \mathcal{D}_0$. To prove the second conclusion, suppose there were $\Delta > 0$ with

$$\sup_{\Lambda_0} R_0(\lambda, \delta_0) \geqslant \bar{V}_0 + \Delta.$$

Then, for $\epsilon > 0$, there is a fixed $\lambda_0 \in \Lambda_0$ with

$$R_0(\lambda_0, \delta_0) \geqslant \bar{V}_0 + \Delta - \epsilon.$$

By opponant regularity, let $w \subset E \times E$, $u_1 \subset \Lambda \times \Lambda$ and $u_2 \subset \mathcal{D} \times \mathcal{D}$ be the entourages of definition II. We can assume that $w = w_1 \times w_2 \times w_3$ is symmetric with $w_1 \subset u_1$, $w_2^2 \subset u_2$ and $w_3 \subset \{(r_1, r_2) \in R^2 \mid |r_1 - r_2| < \epsilon\}$. From $\lim_v \sup \bar{V}_v \leqslant \bar{V}_0$, it follows that there is an entourage $v \subset w$ such that $\bar{V}_v \leqslant \bar{V}_0 + \epsilon$ and $\delta_v \in w_2(\delta_0)$. The problem $(\Lambda_v, \mathcal{D}_v, R_v)$ is close of order w to $(\Lambda_0, \mathcal{D}_0, R_0)$. Therefore, there is a $(\lambda_v, \delta_v') \in \Lambda_v \times \mathcal{D}_v$ with $\lambda_v \in w_1(\lambda_0)$ and $\delta_v' \in w_2(\delta_0)$ such that

$$\begin{aligned} R_v(\lambda_v, \delta_v') + \epsilon &\geqslant R_0(\lambda_0, \delta_0) \\ &\geqslant \bar{V}_0 + \Delta - \epsilon \\ &\geqslant \bar{V}_v + \Delta - 2\epsilon \\ &= \sup_{\Lambda_v} R_v(\lambda, \delta_v) + \Delta - 2\epsilon \\ &\geqslant R_v(\lambda_v, \delta_v) + \Delta - 2\epsilon. \end{aligned}$$

Since $\delta_v \in w_2(\delta_0)$ and $\delta_v' \in w_2(\delta_0)$, then $\delta_v \in w_2^2(\delta_v') \subset u_2(\delta_v')$ with $\lambda_v \in w_1(\lambda_0) \subset u_1(\lambda_0)$. Opponant regularity then gives, $|R_v(\lambda_v, \delta_v') - R_v(\lambda_v, \delta_v)| < \epsilon$ and

$$R_\nu(\lambda_\nu, \delta_\nu) + 2\epsilon \geqslant R_\nu(\lambda_\nu, \delta_\nu) + \Delta - 2\epsilon$$

for all $\epsilon > 0$, which contradicts $\Delta > 0$.

Whether a given family of problems can be considered a subfamily of one that is opponant regular at $(\Lambda_0, \mathcal{D}_0, R_0)$ can often be answered affirmatively using the following remarks. In a polish space, Ω, every probability distribution λ defined on the Borel sets has a compact approximate carrier, i.e. a compact set $K \subset \Omega$ exists with $\lambda(K) \geqslant 1 - \epsilon$. If, as often seems to be the case in parametric statistics (see *Wald* [1950] theorem 3.3), the space Ω is only separable metric, then there is the weaker statement; for every $\epsilon > 0$, there exists a totally bounded Borel set $\overset{\circ}{K}$ with $\lambda(\overset{\circ}{K}) \geqslant 1 - \epsilon$. A case in point would be a parameter set in *Wald* [1950] with the metric ρ_4. The existance of such a set $\overset{\circ}{K}$ often makes a regularity proof possible. In order to see how this looks in detail, we return to the space Λ with metric ρ as constructed in § 1. We suppose that $\bar{\mathbf{F}} = \{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ is a family with $\Lambda_i \subset \Theta_i^* \subset \Lambda$, and that R_i is defined on $\Lambda_i \times \mathcal{D}_i$ by the formula

$$R_i(\lambda, \delta) = \int_{\Theta_i} R_i(\theta, \delta)\, \lambda(d\theta)$$

where $R_i(\theta, \delta) \geqslant 0$ is a θ-measurable function defined on $\Theta_i \times \mathcal{D}_i$. We state the following definitions.

Definition A: The family $\bar{\mathbf{F}}$ has property A at $(\Lambda_0, \mathcal{D}_0, R_0) \in \bar{\mathbf{F}}$ if there is an entourage $\nu \subset E \times E$ and an $M \geqslant 0$ such that, for every $i \in I$ with $(\Lambda_i, \mathcal{D}_i, R_i)$ close of order ν to $(\Lambda_0, \mathcal{D}_0, R_0)$,

$$\left.\begin{array}{l}(\theta_1, \theta_2) \in \Theta_i \times \Theta_i \\ \delta \in \mathcal{D}_i\end{array}\right. \Rightarrow |R_i(\theta_1, \delta) - R_i(\theta_2, \delta)| \leqslant M\rho(\theta_1, \theta_2).$$

Definition B: The family $\bar{\mathbf{F}}$ has property B at $(\Lambda_0, \mathcal{D}_0, R_0) \in \bar{\mathbf{F}}$ if, for each $\theta \in \Theta$ and $\epsilon > 0$, there are entourages $\nu \subset E \times E$ and $u_2 \subset \mathcal{D} \times \mathcal{D}$ such that, for all $i \in I$ with $(\Lambda_i, \mathcal{D}_i, R_i)$ close of order ν to $(\Lambda_0, \mathcal{D}_0, R_0)$,

$$\left.\begin{array}{l}\theta \in \Theta_i \\ (\delta_1, \delta_2) \in u_2 \cap \mathcal{D}_i \times \mathcal{D}_i\end{array}\right. \Rightarrow |R_i(\theta, \delta_1) - R_i(\theta, \delta_2)| < \epsilon.$$

Property A is, in general, stronger than statistical regularity at $(\Lambda_0, \mathcal{D}_0, R_0)$.

We call $\bar{\mathbf{F}}$ locally bounded at $(\Lambda_0, \mathcal{D}_0, R_0) \in \bar{\mathbf{F}}$ if there is an $N \geqslant 0$ and a $\nu \subset E \times E$ such that, for all $i \in I$ with $(\Lambda_i, \mathcal{D}_i, R_i)$ close of order ν to $(\Lambda_0, \mathcal{D}_0, R_0) \in \bar{\mathbf{F}}$, $R(\theta, \delta) \leqslant N$ on $\Theta_i \times \mathcal{D}_i$.

Lemma I: Let the family $\bar{\mathbf{F}}$ be such that, for each $i \in I$, Θ_i is a separable metric space in the restriction of ρ to Θ_i. If $\bar{\mathbf{F}}$ has the properties A and B at $(\Lambda_0, \mathcal{D}_0, R_0)$ and is locally bounded there, then $\bar{\mathbf{F}}$ is opponant regular there.

Proof: Choose ν and M as prescribed in Definition A, and assume that $R_i(\theta, \delta) \leqslant N$ for all $(\Lambda_i, \mathcal{D}_i, R_i)$ close of order u to $(\Lambda_0, \mathcal{D}_0, R_0)$. Let $\lambda_0 \in \Lambda$ and $u_1 = \{(\lambda_1, \lambda_2) | \rho(\lambda_1, \lambda_2) < \epsilon\} \subset \Lambda \times \Lambda$. Then, if $(\Lambda_i, \mathcal{D}_i, R_i)$ is close of order u to $(\Lambda_0, \mathcal{D}_0, R_0)$,

$$\text{i)} \quad \left.\begin{array}{l} (\lambda, \lambda_0) \in u_1 \cap \Lambda_i \times \Lambda_i \\ \delta \in \mathcal{D}_i \end{array}\right. \Rightarrow |R_i(\lambda, \delta) - R_i(\lambda_0, \delta)| \leqslant M\epsilon.$$

Since $\lambda_0 \in \Theta_j^*$ for some j, we can find a totally bounded $\overset{\circ}{K} \subset \Theta_j$ with $\lambda_0(\overset{\circ}{K}) \geqslant 1 - \epsilon$. Thus, by local boundedness, if $(\Lambda_i, \mathcal{D}_i, R_i)$ is close of order u to $(\Lambda_0, \mathcal{D}_0, R_0)$,

$$\text{ii)} \quad \delta \in \mathcal{D}_i \Rightarrow \left| \int_{\Theta_i} R_i(\theta, \delta)\, \lambda_0(d\theta) - \int_{\Theta_i \cap \overset{\circ}{K}} R_i(\theta, \delta)\, \lambda_0(d\theta) \right| \leqslant N\epsilon.$$

Let $\{\theta_\alpha\}$ $(\alpha = 1, 2, \ldots, n)$ be a finite subset of $\overset{\circ}{K}$ whose ϵ-neighborhoods cover $\overset{\circ}{K}$. For each α $(\alpha = 1, 2, \ldots, n)$ let $v^\alpha \subset E \times E$ and $u_2^\alpha \subset \mathcal{D} \times \mathcal{D}$ be such that

$$\left.\begin{array}{l} \theta_\alpha \in \Theta_i \\ (\delta_1, \delta_2) \in u_2^\alpha \cap \mathcal{D}_i \times \mathcal{D}_i \end{array}\right. \Rightarrow |R_i(\theta_\alpha, \delta_1) - R_i(\theta_\alpha, \delta_2)| < \epsilon$$

for all $(\Lambda_i, \mathcal{D}_i, R_i)$ close of order v^α to $(\Lambda_0, \mathcal{D}_0, R_0)$. Choosing $w = v \cap v^1 \cap v^2 \cap \ldots \cap v^n$ and $u_2^1 \cap u_2^2 \cap \ldots \cap u_2^n$, we have, for all α and $i \in I$ with $(\Lambda_i, \mathcal{D}_i, R_i)$ close of order w to $(\Lambda_0, \mathcal{D}_0, R_0)$,

$$\left.\begin{array}{l} \theta_\alpha \in \Theta_i \\ (\delta_1, \delta_2) \in u_2 \cap \mathcal{D}_i \times \mathcal{D}_i \end{array}\right. \Rightarrow |R_i(\theta_\alpha, \delta_1) - R_i(\theta_\alpha, \delta_2)| < \epsilon,$$

and therefore, for all $\theta \in \overset{\circ}{K}$,

$$\left.\begin{array}{l} \theta \in \Theta_i \\ (\delta_1, \delta_2) \in u_2 \cap \mathcal{D}_i \times \mathcal{D}_i \end{array}\right. \Rightarrow |R_i(\theta, \delta_1) - R_i(\theta, \delta_2)| < (2M + 1)\epsilon,$$

where we have used property A. Hence, if $(\Lambda_i, \mathcal{D}_i, R_i)$ is close of order w to $(\Lambda_0, \mathcal{D}_0, R_0)$, $(\delta_1, \delta_2) \in u_2 \cap \mathcal{D}_i \times \mathcal{D}_i$ implies

$$\text{iii)} \quad \left| \int_{\Theta_i \cap \overset{\circ}{K}} R_i(\theta, \delta_1)\, \lambda_0(d\theta) - \int_{\Theta_i \cap \overset{\circ}{K}} R_i(\theta, \delta_2)\, \lambda_0(d\theta) \right| \leqslant (2M + 1)\epsilon.$$

Combining i), ii) and iii) proves the lemma.

Definition: A family $\mathbf{F} = \{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ is regular at $(\Lambda_0, \mathcal{D}_0, R_0) \in \mathbf{F}$ if it is statistically regular and opponant regular there.

When regularity is assumed, then the conclusion of theorem I can be strengthened to read, $\lim_{v} V_v = V_0$. Perhaps more interesting are the following corollaries.

Corollary I: Let $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I} = \mathbf{F}$ be regular at $(\Lambda_0, \mathcal{D}_0, R_0) \in \mathbf{F}$. Suppose that every problem in $\mathbf{F}$ is strictly determined and that $\mathcal{D}_0$ is closed in $\mathcal{D}$. If, for every entourage $u \subset E \times E$, there is a problem $(\Lambda_u, \mathcal{D}_u, R_u) \in \mathbf{F}$ close of order u to $(\Lambda_0, \mathcal{D}_0, R_0)$ and if δ_u is minimax for $(\Lambda_u, \mathcal{D}_u, R_u)$, then every cluster point of $\{\delta_u\}$ is minimax for $(\Lambda_0, \mathcal{D}_0, R_0)$.

Corollary II: If, in corollary I, the set $\mathcal{D}_0$ is compact in $\mathcal{D}$ and if $(\Lambda_0, \mathcal{D}_0, R_0)$ has a unique minimax in $\mathcal{D}_0$, then every minimax criterion defined on $\mathbf{F}$ is regular at $(\Lambda_0, \mathcal{D}_0, R_0)$.

The proof of corollary II is based on the fact that, in any uniform space Ω, if K is a compact subset of Ω and if $\{x_i\}$ is a net of elements in Ω converging to K, then $\{x_i\}$ has a cluster point lying in K.

It doesn't seem to be very promising to search for results of the above kind when the problems in question are not strictly determined. Consider the following example. As in case just subsequent to the proof of theorem I, we consider the family $\{P_\theta\}$ of normal distributions with mean θ and unit variance. We choose, as parameter space Θ, the union of $\{\theta = 0\}$ with the half line $\{\theta \in R^1 \mid \theta \geqslant 1\}$. The loss function $L(\theta, i)$ is the restriction of the one we defined previously, but with $l(\theta) = 0$:

$$L(\theta, 0) = \begin{cases} 0 & \text{for } \theta = 0 \\ 1 & \text{for } \theta \geqslant 1 \end{cases}$$

$$L(\theta, 1) = 1 - L(\theta, 0).$$

We use the distance ρ_4 to define the uniform structure on Θ^* and let $\mathcal{D}$ be the class of critical functions on R^1 with, as in the previous example, the topology generated by the P_θ $(\theta \geqslant 0)$. We let $\mathcal{D}_0$ be equal to $\mathcal{D}$, $\Theta_0^* = \Theta^*$ and R_0 equal to the risk function,

$$R = \begin{cases} \int \phi \, P_\theta \, (dx) & \text{for } \theta = 0 \\ \int (1 - \phi) \, P_\theta \, (dx) & \text{for } \theta \geqslant 0. \end{cases}$$

The problem $(\Theta_0^*, \mathcal{D}_0, R_0)$ so defined is strictly determined with the essentially unique minimax

$$\phi_0(x) = \begin{cases} 0; & x \leqslant K \\ 1; & x \geqslant K \end{cases}$$

where K is an appropriately determined real number. Define, now, a set $\mathcal{D}_1$ as follows:

Let E_1 be the set of those critical functions ϕ on R^1 having the property that

$\phi(x) = 0$ for all sufficiently large $x \in R^1$. We define ϕ_1 by $\phi_1(x) \equiv \sum_{k=-\infty}^{\infty} I_k(x)$, where I_k is the indicator function for the interval $[2k, 2k+1]$. $\mathcal{D}_1$ is defined to be the convex hull of E_1 and the one point set $\{\phi_1\}$. The set $\mathcal{D}_1$ is dense in $\mathcal{D}$ with respect to the $\mathcal{D}$-topology. Thus, the statistical problem $(\Theta^*, \mathcal{D}_1, R_1)$, where $R_1 = R|_{\Theta^* \times \mathcal{D}_1}$, is close or order $u \subset E \times E$ to $(\Theta_0^*, \mathcal{D}_0, R_0)$ for arbitrary u. It also has the unique minimax ϕ_1 which is not equal to ϕ_0. It seems to be worth noticing, that $\mathcal{D}_1$ has the property that all of its extreme points are extreme points of $\mathcal{D}$. The example seems, therefore, to be a rather strong one, making it hard to see how conditions could be imposed on the problems of a family, sufficient for proving regularity, but not for proving strict determinateness.

§ 3 Applications of Theorems I and II

In this paragraph, we prove the restricted regularity (see § 1) of the admissible minimax criteria for certain of the best known problems of parametric statistics. The problems to be considered are, roughly speaking, the following:

i) monotone likelihood ration
ii) exponential family with one parameter and two sided alternative set
iii) exponential family with several parameters and two sided alternative set.

As already indicated in § 1, we will consider only solved problems. The *Behrens–Fisher* problem, for example, will remain open.

We consider the above mentioned problems in connection with the usual testtheoretical loss function which, when reasonally specified, gives a risk function

$$R(\theta, \phi) = \begin{cases} (l_1^0 - l_0^0 \int \phi P_\theta(dx) + l_0^0 & ;(\theta \in H_0) \\ (l_0^1 - l_1^1)(1 - \int \phi P_\theta(dx)) + l_1^1 ; (\theta \in H_0^c) \end{cases} \tag{3.1}$$

for each θ and each critical function ϕ, where $H_0 \subset \Theta$ is the nullhypothesis. We assume the following relations among the loss quantities l_i^j $(i, j = 0,1)$,

$$\begin{aligned} l_1^0 &> l_0^0 \qquad & l_0^1 &> l_1^1 \\ l_0^0 &\leqslant l_0^1 \qquad & l_1^0 &\geqslant l_1^1. \end{aligned} \tag{3.2}$$

The topology on the set $\mathcal{D}$ of critical functions ϕ will be that generated by the $\{P_\theta\}_{\theta \in \Theta}$. Thus, $\mathcal{D}$ can be considered to be the set of power functions, $\int \phi P_\theta(dx)$, on Θ. We use *Wald*'s ρ_4 to define a distance in Θ (and in Θ^* as prescribed in § 1) and thus a uniform structure on Θ^*. It can easily be seen that we cannot simply apply corollary II, § 2 to these problems since their minimax tests are not unique. There appear to be no useful easily recognizable conditions under which the uniqueness of the admissible

minimax for a problem insures regularity of admissible minimax criteria there. It is easy to construct examples in which the admissible minimax test is uniquely defined, but where no reasonable admissible minimax criterion is regular. We treat the above cases directly, therefore, without first searching for further general results.

We will often make use of a certain continuity property implicit in the generalized form of the *Neyman-Pearson* lemma. Suppose we are given a bounded convex subset N of $R^{n+1} = \{(x_1, \ldots, x_n, x_{n+1})\}$ with projection, M, on the $(x_1, \ldots, x_n)$ − plane. Let $\mathring{x} = (\mathring{x}_1, \ldots, \mathring{x}_n)$ be an interior point of M in $R^n = \{(x_1, \ldots, x_n)\}$ with respect to the euclidean norm, $\| \cdot \|$. If N has no interior point with respect to R^{n+1}, then N is a subset of a hyperplane Π in R^{n+1} with points satisfying

$$k_1 x_1 + \ldots + k_n x_n + k_{n+1} x_{n+1} = b = \text{const.}$$

Since $\mathring{x} \in M$ is interior to M, the coefficient k_{n+1} cannot be zero, and x_{n+1} is a continuous function of $(x_1, \ldots, x_n)$ on M. If N has an interior point, then there is some number, a, such that $(\mathring{x}_1, \ldots, \mathring{x}_n, a) \in N$ is interior to N. The function $y(x_1, \ldots, x_n)$ defined on M by

$$y(x_1, \ldots, x_n) \equiv \sup \{y \mid (x_1, \ldots, x_n, y) \in N\}$$

is then continuous at the point $\mathring{x} \in M$. This is seen is follows. Let $\epsilon > 0$ and let U_ϵ be an ϵ-neighborhood of $(\mathring{x}_1, \ldots, \mathring{x}_n, a)$ lying entirely in N. Then the function

$$y(\mathring{x})\left(1 - \frac{1}{\epsilon} \| x - \mathring{x} \|\right) + \frac{a}{\epsilon} \| x - \mathring{x} \|$$

defined on $\{x \in R^n \mid \|x - \mathring{x}\| < \epsilon\}$ is continuous, equal to $y(\mathring{x})$ at $x = \mathring{x}$ and nowhere greater than $y(x_1, \ldots, x_n)$ on $\{\|x - \mathring{x}\| < \epsilon\}$. The plane,

$$\sum_{i=1}^{n+1} k_i' x_i = b' = \text{const},$$

tangent to the set N at the point $(\mathring{x}_1, \ldots, \mathring{x}_n, y(\mathring{x}))$ represents x_{n+1} as a continuous function of $(x_1, \ldots, x_n)$ since $\mathring{x}$ is interior to M. Since this function is never less than $y(x_1, \ldots, x_n)$ on M and since it is equal to $y(\mathring{x})$ at $x = \mathring{x}$, y must be continuous at $x = \mathring{x}$. An easy corollary to these remarks is

Remark I (compare, *Lehmann* [1959] Theorem 5).
Let μ be a σ-finite measure on a measurable space $(\Omega, \mathcal{B}), f_1, \ldots, f_n, f_{n+1}$, integrable functions on $(\Omega, \mathcal{B}, \mu)$. Let $\mathcal{D}$ be the class of critical functions ϕ on Ω. Further, let M be the subset of R^n defined by

$$M \equiv \{(\int \phi f_1 \mu(d\omega), \ldots, \int \phi f_n \mu(d\omega) \mid \phi \in \mathcal{D}\},$$

and $\mathcal{C}(x) = \mathcal{C}(x_1, \ldots, x_n)$, the set of all $\phi \in \mathcal{D}$ with $\int \phi f_i \mu(d\omega) = x_i$ $(i = 1, 2, \ldots, n)$. If $(\overset{\circ}{x}_1, \ldots, \overset{\circ}{x}_n)$ is an interior point of M, then the function

$$y(x_1, \ldots, x_n) = \sup_{\phi \in \mathcal{C}(x)} \int \phi f_{n+1} \mu(d\omega)$$

is continuous at $x = \overset{\circ}{x}$ in M.

In the statement of remark I, we use the term "interior point of M" to denote a point in M, interior to M in the euclidean topology of the subspace of R^n spanned by M.

In the sequel, we consider a monotone likelihood ratio problem to be one having the following form: The parameter space Θ will be a subset of R^1. To each $\theta \in \Theta$ there corresponds a probability distribution P_θ on the sample space $(X, \mathcal{B})$ having density p_θ with respect to the fixed σ-finite measure μ on $(X, \mathcal{B})$. Further, there exists a finite valued random variable T on $(X, \mathcal{B})$ and a class $\{G^\theta_{\theta'}\}_{\theta' \geqslant \theta}$ of real valued, non-decreasing functions on the real line such that, for all θ and θ' in Θ with $\theta' \geqslant \theta$,

$$p_{\theta'}(x) = G^\theta_{\theta'}(T(x))\, p_\theta(x) \tag{3.3}$$

for all $x \in X$. The nullhypothesis, H_0, will be the (non empty proper) subset of all $\theta \in \Theta$ which are not greater than some fixed value $\theta_0 \in \Theta$. We demand that $P_\theta = P_{\theta'}$ only if $\theta = \theta'$ and consider a risk function of the form $R(\theta, \phi)$ given by (3.1).

The family $\mathbf{F}$ of problems we consider is of the form $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ where Λ_i is a convex subset of Θ^*, $\mathcal{D}_i = \mathcal{D}$ and R_i is R restricted to $\Lambda_i \times \mathcal{D}$. For $\lambda \in \Theta^*$, the symbol, $R(\lambda, \phi)$, is already defined by (1.3). Notice that the set Λ_i can also be considered the non-randomised parameter space for the problem $(\Lambda_i, \mathcal{D}_i, R_i)$ since it is convex and since, for every $\lambda \in \Lambda_i$, $R(\lambda, \phi)$ is continuous on $\mathcal{D}_i = \mathcal{D}$. Thus the problem $(\Lambda_i, \mathcal{D}_i, R_i)$ is strictly determined and has an admissible minimax ϕ [*Ferguson*, 1967]. We adopt the convention that $0 \in I$ and that $(\Lambda_0, \mathcal{D}_0, R_0)$ is the monotone likelihood ratio problem $(\Theta^*, \mathcal{D}, R)$.

We recapitulate the best known properties of the optimal α-level tests ϕ_α for this problem. If α is neither zero nor one, then ϕ_α can be chosen to be of the form

$$\phi_\alpha(x) = \begin{cases} 1 & \text{if } T(x) > C_0 \\ \gamma & \text{if } T(x) = C_0 \\ 0 & \text{if } T(x) < C_0 \end{cases} \tag{3.4}$$

where C_0 and γ $(0 \leqslant \gamma \leqslant 1)$ are constants satisfying

$$\begin{aligned} C_0 &= \sup \{C \mid P_{\theta_0}(T \leqslant C) \leqslant 1 - \alpha\} \\ &= P_{\theta_0}(T > C_0) + \gamma P_{\theta_0}(T = C_0). \end{aligned} \tag{3.5}$$

In case $\alpha = 0$, then $\phi_\alpha = 0$, and if $\alpha = 1$, then

$$\phi_1 = \begin{array}{l} 1 \text{ if } p_{\theta_0} > 0 \\ 0 \text{ if } p_{\theta_0} = 0. \end{array} \tag{3.6}$$

For all α $(0 \leqslant \alpha \leqslant 1)$ the tests ϕ_α have the following properties: Let $\beta(\theta)$ be the power function, $\int \phi_\alpha P_\theta (dx)$, of the test ϕ_α. Then

i) $\beta(\theta_0) = \alpha$

ii) $\beta(\theta)$ is non-decreasing in θ on Θ

iii) Let $\theta \leqslant \theta'$. If ϕ is a critical function with $\int \phi P_\theta (dx) = \beta(\theta)$, then $\beta(\theta') \geqslant \int \phi P_{\theta'} (dx)$

iv) Let $\theta_0 \geqslant \theta'$. If ϕ is a critical function with $\int \phi P_{\theta_0} (dx) = \alpha$, then $\beta(\theta') \leqslant \int \phi P_{\theta'} (dx)$.

It is clear that any test having the properties i) – iv) must also have the power function $\beta(\theta)$. Further, any admissible minimax test for the problem $(\Lambda_0, \mathcal{D}_0, R_0)$ must be one with the same power function as ϕ_α for a property chosen α. Using the fact that $P_{\theta'} \ll P_\theta$ for $\theta' \geqslant \theta_0 \geqslant \theta$, one can easily prove the uniquality,

$$\sup_{\theta \leqslant \theta_0} R_0(\theta, \phi) \leqslant \sup_{\theta > \theta_0} R_0(\theta, \phi) \tag{3.7}$$

for *any* minimax test ϕ for the problem. Thus the minimax value is given by $V = \sup_{\theta > \theta_0} R_0(\theta, \phi_\alpha)$ for α suitable chosen, and

$$V = (l_0^1 - l_1^1)\left(1 - \inf_{\theta > \theta_0} \int \phi_\alpha P_\theta (dx)\right) + l_1^1. \tag{3.8}$$

Now suppose that $\{\phi_{\alpha_i}\}$ is a net of α_i-optimal tests with the property

$$\lim_i \inf_{\theta > \theta_0} \int \phi_{\alpha_i} P_\theta (dx) = \inf_{\theta > \theta_0} \int \phi_{\alpha_0} P_\theta (dx)$$

where ϕ_{α_0} is α_0-optimal. Then, if $\inf_{\theta > \theta_0} \int \phi_{\alpha_0} P_\theta (dx) < 1$, one can prove, using properties i) – iii) above, that α_i converges to α_0. If $\inf_{\theta > \theta_0} \int \phi_{\alpha_0} P_\theta (dx) = 1$, then α_i may not converge to α_0 unless we specify more about $\{\phi_{\alpha_i}\}$ and ϕ_{α_0}.

Remark II: Let $(\Lambda_i, \mathcal{D}_i, R_i) \to (\Lambda_0, \mathcal{D}_0, R_0)$, $\{\phi_{\alpha_i}\}$ and ϕ_{α_0} be as above. Suppose that ϕ_{α_i} and ϕ_{α_0} are admissible for $(\Lambda_i, \mathcal{D}_i, R_i)$ and $(\Lambda_0, \mathcal{D}_0, R_0)$ respectively. Then $\inf_{\theta > \theta_0} \int \phi_{\alpha_i} P_\theta (dx) \to \inf_{\theta > \theta_0} \int \phi_{\alpha_0} P_\theta (dx)$ implies $\alpha_i \to \alpha_0$.

Proof: We can assume that $\int \phi_{\alpha_0} P_\theta (dx) = 1$ for all $\theta > \theta_0$. Non-convergence of α_i to α_0 implies the existence of a subnet $\{\phi_{\alpha_i}\}$ with one of the properties:

a) $\alpha_i > \alpha_0 + \epsilon$ for some $\epsilon > 0$ and all i,

b) $\alpha_i < \alpha_0 - \epsilon$ for some $\epsilon > 0$ and all i.

If b), then let $\bar{\phi}$ be a cluster point of $\{\phi_{\alpha_i}\}$ in $\mathcal{D}$. Then we have $R(\theta_0 \bar{\phi}) < R(\theta_0, \phi_{\alpha_0})$, $R(\theta, \bar{\phi}) = 1$ for all $\theta > \theta_0$ and, by property iv), $R(\theta, \bar{\phi}) = \lim_i R(\theta, \phi_{\alpha_i}) \leqslant R(\theta, \phi_{\alpha_0})$ for all $\theta \leqslant \theta_0$, so that ϕ_{α_0} would not be admissible. If a) then $\int \phi_{\alpha_i} P_\theta(dx) = 1$ for all i and $\theta > \theta_0$, otherwise the critical function

$$\frac{1-\alpha_i}{1-\alpha_0}\,\phi_{\alpha_0} + \frac{\alpha_i - \alpha_0}{1-\alpha_0}$$

would violate property iii) of ϕ_{α_i}. Thus, for all θ, $R(\theta, \phi_{\alpha_i}) \geqslant R(\theta, \phi_{\alpha_0})$ by property iv) of ϕ_{α_0}. We can show that ϕ_{α_i} must eventually become non-admissible for $(\Lambda_i, \mathcal{D}_i, R_i)$: There exists $M > 0$ such that

$$|R(\theta, \phi) - R(\theta', \phi)| \leqslant M \rho_4(\theta, \theta')$$

for all $\phi \in \mathcal{D}$ and θ, θ' in Θ. Since $(\Lambda_i, \mathcal{D}_i, R_i) \to (\Lambda_0, \mathcal{D}_0, R_0)$, we know that there is a $\{\lambda_i\}$ with $\lambda_i \in \Lambda_i$ and $\rho_4(\lambda_i, \delta_\theta) \to 0$, where δ_{θ_0} is the delta-measure at θ_0. Thus, for any $\epsilon' > 0$, there is i_0 with

$$\sup_\phi |R(\lambda_{i_0}, \phi) - R(\theta_0, \phi)| < \epsilon'.$$

Since $R(\theta_0, \phi_{\alpha_i}) > R(\theta_0, \phi_{\alpha_0}) + (l_1^0 - l_0^0) \in$ for all i, we find, for $\epsilon' > 0$ chosen sufficiently small, that $R(\lambda_{i_0}, \phi_{\alpha_{i_0}}) > R(\lambda_{i_0}, \phi_{\alpha_0})$ and $\phi_{\alpha_{i_0}}$ is non-admissible for $(\Lambda_{i_0}, \mathcal{D}_{i_0}, R_{i_0})$, which proves the remark.

There always exists a net of admissible ϕ_{α_i}. This is because, if ϕ is admissible for $(\Lambda, \mathcal{D}, R) \in \mathbf{F}$ then so is ϕ_α where $\alpha = \int \phi P_{\theta_0}(dx)$.

In the following, we make use of the regularity of the family $\mathbf{F}$ at $(\Lambda_0, \mathcal{D}_0, R_0)$. That it is statistically regular there is obvious (Definition B, following). Opponant regularity follows rather easily from Lemma I, § 2. Let ϕ_{α_i} be α_i-optimal for $(\Lambda_0, \mathcal{D}_0, R_0)$ and admissible minimax for $(\Lambda_i, \mathcal{D}_i, R_i)$. If $(\Lambda_i, \mathcal{D}_i, R_i) \to (\Lambda_0, \mathcal{D}_0, R_0)$ then we must have

$$\lim_i \sup_{\theta \leqslant \theta_0} R(\theta, \phi_{\alpha_i}) \leqslant \lim_i \sup_{\theta > \theta_0} R(\theta, \phi_{\alpha_i}).$$

If this were not so, then there would be a subnet and an $\epsilon > 0$ with

$$\sup_{\theta \leqslant \theta_0} R(\theta, \phi_{\alpha_i}) \geqslant \sup_{\theta > \theta_0} R(\theta, \phi_{\alpha_i}) + \epsilon.$$

A cluster point of $\{\phi_{\alpha_i}\}$ would be, by corollary I, theorem II, a minimax for $(\Lambda_0, \mathcal{D}_0, R_0)$ so that (3.8) would hold;

$$\sup_{\theta \leqslant \theta_0} R(\theta, \phi) \leqslant \sup_{\theta > \theta_0} R(\theta, \phi)$$

which would lead immediately to a contradiction. Therefore we have

$$\lim_i \sup_\theta R(\theta, \phi_{\alpha_i}) = \lim_i \sup_{\theta > \theta_0} R(\theta, \phi_{\alpha_i}).$$

Since $(\Lambda_i, \mathcal{D}_i, R_i) \to (\Lambda_0, \mathcal{D}_0, R_0)$ we have $V_i \to V_0$ (remarks preceding corollary I, theorem II, § 2) where V_i is the minimax value for $(\Lambda_i, \mathcal{D}_i, R_i)$. Further, there is an $M > 0$ such that

$$\sup_{\phi \in \mathcal{D}} |R(\theta, \phi) - R(\theta', \phi)| \leqslant M \rho_4(\theta, \theta')$$

for all θ, θ' in Θ, so that one can easily prove that

$$\lim_i \sup_{\Lambda_i} R(\lambda, \phi_{\alpha_i}) = \lim_i \sup_\theta R(\theta, \phi_{\alpha_i}),$$

and therefore that

$$\lim_i \sup_{\theta > \theta_0} R(\theta, \phi_{\alpha_i}) = V_0 = \sup_{\theta > \theta_0} R(\theta, \phi)$$

$$\lim_i \inf_{\theta > \theta_0} \int \phi_{\alpha_i} P_\theta(dx) = \inf_{\theta > \theta_0} \int \phi_{\alpha_0} P_\theta(dx),$$

and consequently, by admissibility of ϕ_{α_i}, that $\alpha_i \to \alpha$. Using remark I, above, in conjunction with properties i) – iv), it follows that, if α_0 is in the interior of the unit interval, $\int \phi_{\alpha_i} P_\theta(dx)$ converges to $\int \phi_{\alpha_0} P_\theta(dx)$ for all $\theta \in \Theta$. If α_0 is equal to one, then this is also true, as one can prove using ii) and iv). If $\alpha_0 = 0$, then the statement also holds. This is obvious for $\theta \leqslant \theta_0$. For $\theta > \theta_0$, it follows from $P_\theta \ll P_{\theta_0}$.

Let ϕ_i be admissible minimax for $(\Lambda_i, \mathcal{D}_i, R_i)$ and ϕ_{α_i} be α_i-optimal for $(\Lambda_0, \mathcal{D}_0, R_0)$ where $\alpha_i = \int \phi_i P_{\theta_0}(dx)$. The test ϕ_{α_i} is admissible minimax for $(\Lambda_i, \mathcal{D}_i, R_i)$, and the immediately preceding statements apply to the net $\{\phi_{\alpha_i}\}$ so defined. Furthermore, we have $R(\theta, \phi_i) \geqslant R(\theta, \phi_{\alpha_i})$ for all i and θ. For each θ, therefore,

$$R(\theta, \phi_i) - R(\theta, \phi_{\alpha_i}) \to 0,$$

since otherwise one could easily prove that ϕ_i must eventually become inadmissible for $(\Lambda_i, \mathcal{D}_i, R_i)$ as this problem approaches $(\Lambda_0, \mathcal{D}_0, R_0)$. It follows that $\int \phi_i P_\theta(dx) \to \int \phi_{\alpha_0} P_\theta(dx)$ for each θ. This proves

Restricted regularity theorem I: Let $(\Theta^*, \mathcal{D}, R)$ be a monotone likelihood ratio problem, **F** a family $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ of problems where $\Lambda_i \subset \Theta^*$, $\mathcal{D}_i \subset \mathcal{D}$, $R_i = R \mid_{\Lambda_i \times \mathcal{D}}$ and where Λ_i is convex for all $i \in I$. The admissible minimax test for $(\Theta^*, \mathcal{D}, R)$ is

unique and, if $(\Theta^*, \mathcal{D}, R) \in \mathbf{F}$, every admissible minimax criterion defined on a neighborhood of $(\Theta^*, \mathcal{D}, R)$ in $\mathbf{F}$ is regular there.

We now consider the two sided exponential family problem with one parameter. By this we will understand a problem $(\Theta^*, \mathcal{D}, R)$ with $\Theta \subset R^1$, $H_0 = \{\theta \in \Theta \mid \theta_1 \leqslant \theta \leqslant \theta_2\}$, θ_1 and θ_2 being interior points of Θ in the euclidean topology, and where the probability distribution P_θ of the family $\{P_\theta\}_{\theta \in \Theta}$ have densities, $C(\theta) e^{\theta T(x)}$, with respect to a fixed σ-finite measure μ on X, $T(x)$ being a fixed finite random variable on X. We demand that $P_\theta = P_{\theta'}$ only if $\theta = \theta'$.

As before, let $\mathbf{F} = \{\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ be a family of problems where Λ_i is a convex subset of Θ^*, $\mathcal{D}_i = \mathcal{D}$ and R_i is R restricted to $\Lambda_i \times \mathcal{D}$. For each $i \in I$, there is an admissible minimax for $(\Lambda_i, \mathcal{D}_i, R_i)$ and each problem is strictly determined. Similary to the monotone likelihood case, let $0 \in I$ and $(\Lambda_0, \mathcal{D}_0, R_0)$ be equal to $(\Theta^*, \mathcal{D}, R)$.

Let α_1 and α_2 be numbers in the closed unit interval, and $\theta_1 < \theta_2$.
If $C_{\alpha_1, \alpha_2} := \{\phi \in \mathcal{D} \mid \alpha_1 = E_{\theta_1}(\phi), \alpha_2 = E_{\theta_2}(\phi)\}$ is non empty, then there exists a test, $\phi_{\alpha_1, \alpha_2} \in C_{\alpha_1, \alpha_2}$ such that, for all $\phi \in C_{\alpha_1, \alpha_2}$,

$$\begin{aligned} &E_\theta(\phi_{\alpha_1, \alpha_2}) \geqslant E_\theta(\phi) \qquad (\theta \notin H_0) \\ &E_\theta(\phi_{\alpha_1, \alpha_2}) \leqslant E_\theta(\phi) \qquad (\theta \in H_0) \end{aligned} \tag{3.9}$$

[*Ferguson*, 1967]. Further, any other test in $\mathcal{D}$ having the properties (3.9) has the same power function as $\phi_{\alpha_2, \alpha_2}$.

Since $P_\theta \ll P_{\theta'}$ for all pairs of points θ and θ' in Θ we have, for any minimax ϕ,

$$\sup_{\theta \in H_0} R(\theta, \phi) = \sup_{\theta \notin H_0} R(\theta, \phi),$$

from which one easily proves that every minimax is unbiased. Since θ_1 and θ_2 are interior points of Θ we have, by continuity of the power function, $E_{\theta_1}(\phi) = E_{\theta_2}(\phi) = \alpha$ for any minimax ϕ, where

$$\alpha = \frac{V - 1_0^0}{l_1^0 - l_0^0},$$

V being the minimax value of $(\Theta^*, \mathcal{D}, R)$.

Let $(\Lambda_i, \mathcal{D}_i, R_i) \to (\Lambda_0, \mathcal{D}_0, R_0)$, and let ϕ_i be admissible minimax for $(\Lambda_i, \mathcal{D}_i, R_i)$. Then $\phi_{\alpha_1^i, \alpha_2^i}$ is also admissible minimax for $(\Lambda_i, \mathcal{D}_i, R_i)$ where $\alpha_1^i = E_{\theta_1}(\phi_i)$ and $\alpha_2^i = E_{\theta_2}(\phi_i)$. Further, similarly to the monotone likelihood case, admissibility gives $E_\theta(\theta_i) - E_\theta(\phi_{\alpha_1^i, \alpha_2^i}) \to 0$ for all $\theta \in \Theta$. If ϕ_∞ is a cluster point of ϕ_i, then ϕ_∞ is minimax for $(\Lambda_0, \mathcal{D}_0, R_0)$ and $E_{\theta_1}(\phi_\infty) = E_{\theta_2}(\phi_\infty) = \alpha$. Thus, for a suitable subnet $\{\phi_i\}$, $\alpha_1^i \to \alpha$ and $\alpha_2^i \to \alpha$. By remark I, if α is in the interior of the unit interval, then $E_\theta(\phi_{\alpha_1^i, \alpha_2^i}) \to E_\theta(\phi_{\alpha, \alpha})$ for all $\theta \in \Theta$. If α is either zero or one, then the latter statement follows from $P_\theta \ll P_{\theta'}$ for all θ, θ' in Θ, and we can conclude that every cluster point of the original net $\{\phi_i\}$ is essentially $\phi_{\alpha, \alpha}$. This proves regularity for $\theta_1 < \theta_2$. The case, $\theta_1 = \theta_2$ is treated similary, and we have

Restricted regularity theorem II: Let $(\Theta^*, \mathcal{D}, R)$ be a two sided exponential family problem with one parameter $\theta \in \Theta$, $\mathbf{F}$ a family $\{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$ of problems where $\Lambda_i \subset \Theta^*$, $\mathcal{D}_i = \mathcal{D}$, $R_i = R \mid_{\Lambda_i \times}$ and where Λ_i is convex for every $i \in I$. The admissible minimax test for $(\Theta^*, \mathcal{D}, R)$ is unique and, if $(\Theta^*, \mathcal{D}, R) \in \mathbf{F}$, every admissible minimax criterion defined on a neighborhood of $(\Theta^*, \mathcal{D}, R)$ in $\mathbf{F}$ is regular there.

The case of several parameter is more involved than that of only one. Here, we assume a parameter space Θ which is a subset of $R^k = \{(\theta^1, \theta^2, \ldots, \theta^k)\}$. The set Θ will be of the form $\Theta^1 \times S$ where S is a subset of $\{(\theta^2, \ldots, \theta^k)\}$ and $\Theta^1 \subset R^1$. The nullhypothesis H_0 will be a set of the form

$$H_0 = \Theta \cap \{\theta_1^1 \leqslant \theta^1 \leqslant \theta_2^1\}$$

where the non-empty sets,

$$\Theta \cap \{\theta^1 = \theta_1^1\} \qquad\qquad \Theta \cap \{\theta^1 = \theta_2^1\}$$

are interior to Θ with respect to the euclidean R^k topology. Letting θ denote the parameter vector $(\theta^1, \ldots, \theta^k)$ and T, a k-vector, $(T_1(x), \ldots, T_k(x))$ of finite random variables on X, the probability distributions $\{P_\theta\}_{\theta \in \Theta}$ have the densities, $C(\theta)\, e^{\theta T}$ with respect to the σ-finite measure μ on X (θT denotes the inner product). Because of the identity ($T \cdot \mu$ is σ-finite),

$$\int \phi\, e^{\theta T} \mu\,(dx) = \int E\,(\phi \mid T)\, e^{\theta t}\,(T \cdot \mu)\,(dt),$$

we can assume, without loss of generality, that $X = R^k$.

The problem, $(\Lambda_0, \mathcal{D}_0, R_0) = (\Theta^*, \mathcal{D}, R)$, so defined will be a member of the family. $\mathbf{F} = \{(\Lambda_i, \mathcal{D}_i, R_i)\}_{i \in I}$, where, as before, $\Lambda_i \subset \Theta^*$ is always convex, $\mathcal{D}_i = \mathcal{D}$ and $R_i = R \mid_{\Lambda_i \times \mathcal{D}_i}$. We remark that, since Θ is separable in the euclidean topology and since each function of the family $\{\int \phi\, P_\phi\,(dx)\}_{\phi \in \mathcal{D}}$ is continuous in θ with respect to that topology, the topology in $\mathcal{D}$ generated by $\{P_\theta\}_{\theta \in \Theta}$ is metrizable.

In order to construct the optimal (α-UMPU) test for the problem $(\Theta^*, \mathcal{D}, R)$, one can begin by defining

$$\tau(x) = \tau(x_1, \ldots, x_k) \equiv (x_2, \ldots, x_k) \in R^{k-1}.$$

Using the decomposition,

$$P_\theta(E) = \int P_\theta\,(E \mid \tau = t)\,(\tau \cdot P_\theta)\,(dt)\ ;\ E \subset X \text{ measurable},$$

the family $\{P_\theta(\cdot \mid \tau = t)\}$ can, for each t, be considered a one parameter exponential family on X with respect to a σ-finite measure ν_t on X. For each t, the corresponding one parameter problem has α-UMPU tests against the nullhypothesis H_0 : $\{\theta_1^1 \leqslant \theta^1 \leqslant \theta_2^1\}$. For fixed α, they can be considered to have the form $\theta(x_1, t)$. The so defined function of (x_1, t) can be chosen to be measurable in X, and α-UMPU for the problem $(\Theta^*, \mathcal{D}, R)$.

Let $(\Lambda_i, \mathcal{D}_i, R_i) \in \mathbf{F}$ and ϕ be an admissible minimax test for this problem. We define, for θ_1^1 and θ_2^1,

$$\alpha_1(t) = \int \phi P_{\theta_1^1}(dx|\tau = t)$$

$$\alpha_2(t) = \int \phi P_{\theta_2^1}(dx|\tau = t).$$

For each t, we consider an optimal two sided test ϕ_{α_1,α_2} for the family $\{P_{\theta^1}(\cdot \mid \tau = t)\}_{\theta^1 \in \Theta^1}$ against the nullhypotheses $\Theta^1 \cap \{\theta_1^1 \leqslant \theta^1 \leqslant \theta_2^1\}$, with $\alpha_1 = \alpha_1(t)$ and $\alpha_2 = \alpha_2(t)$. The family $\{\phi_{\alpha_1(t),\alpha_2(t)}\}_{t \in R^{k-1}}$ can be so chosen, that the function on X, defined by

$$\phi_{\alpha_1,\alpha_2}(x) := \phi_{\alpha_1(t),\alpha_2(t)}\ ; \qquad (t = \tau(x)),$$

is measurable. The proof of this uses the obvious measurability of α_1 and α_2 together with an argument differing very little from that for the case, $\alpha_1 = \alpha_2 = \text{const}$. Suppose that the net $\{(\Lambda_i, \mathcal{D}_i, R_i)\}$ converges to $(\Lambda_0, \mathcal{D}_0, R_0)$ and that ϕ^i is admissible minimax for $(\Lambda_i, \mathcal{D}_i, R_i)$ with

$$\begin{aligned} \alpha_1^i(t) &= \int \phi^i P_{\theta_1^1}(dx \mid \tau = t) \\ \alpha_2^i(t) &= \int \phi^i P_{\theta_2^1}(dx \mid \tau = t). \end{aligned} \tag{3.10}$$

The test, $\phi_{\alpha_1^i,\alpha_2^i}$, is minimax for $(\Lambda_i, \mathcal{D}_i, R_i)$ and, by compactness of $\mathcal{D}$, any subnet of $\{\phi_{\alpha_1^i,\alpha_2^i}\}$ has a cluster point ϕ in $\mathcal{D}$ which is minimax for $(\Lambda_0, \mathcal{D}_0, R_0)$. Therefore, $\int \alpha_1^i(t)\, \tau \cdot P_\theta(dt) \to \alpha$, where $\alpha = (V_0 - l_0^0)/(l_1^0 - l_0^0)$, for all $\theta \in \Theta$ with $\theta^1 = \theta_1^1$. Similarly for $\theta^1 = \theta_2^1$, $\int \alpha_2^i(t)(\tau P_\theta)(dt)$ converges to α. Further, by admissibility of ϕ^i, we can say, that $R(\theta, \phi^i) - R(\theta, \phi_{\alpha_1^i,\alpha_2^i})$ converges to zero and hence $\int (\phi^i - \phi_{\alpha_1^i,\alpha_2^i}) P_\theta(dx)$ also for every $\theta \in \Theta$.

Let ϕ_∞ be a cluster point of $\{\phi_{\alpha_1^i,\alpha_2^i}\}$, and define $\vartheta(\theta) = (\theta^2, \ldots, \theta^k)$. Since $\mathcal{D}$ is metrizable, we can find a countable subsequence $\{\phi_{\alpha_1^i,\alpha_2^i} \mid i = 1, 2 \ldots\} \subset \{\phi_{\alpha_1^i,\alpha_2^i}\}$ converging to ϕ_∞ and having the properties,

$$\int \alpha_1^i(t)(\tau \cdot P_\theta)(dt) \to \alpha \qquad (\theta^1 = \theta_1^1)$$

$$\int \alpha_2^i(t)(\tau \cdot P_\theta)(dt) \to \alpha \qquad (\theta^1 = \theta_2^1)$$

for all rational $\vartheta(\theta)$ in R^{k-1}.

The probability distributions $\tau \cdot P_\theta$ have densities, $C(\theta^1, \vartheta(\theta))\, e^{\vartheta(\theta)t}$ with respect to the σ-finite measure λ_{θ_1} on $T(X)$. For $\theta^1 = \theta_1^1$ we know that the family $\{\tau \cdot P_\theta\}$ is complete and generates a compact metrizable hausdorff topology on the set of all λ_{θ_1}-almost everywhere defined measurable functions $\alpha(t)$ with $0 \leqslant \alpha(t) \leqslant 1$. Similary

for $\theta^1 = \theta_2^1$. In these topologies, let α_1^∞ and α_2^∞ be cluster points for α_1^i and α_2^i respectively. By suitable choice of infinite subsequence of $\{\phi_{\alpha_1^i,\alpha_2^i}\}$, we can assume that $\alpha_1^i \to \alpha_1^\infty$ and $\alpha_2^i \to \alpha_2^\infty$. Then by bounded convergence,

$$\int \alpha_1^\infty(t)\,(\tau \cdot P_\theta)\,(dt) = \alpha \qquad (\theta^1 = \theta_1^1)$$

$$\int \alpha_2^\infty(t)\,(\tau \cdot P_\theta)\,(dt) = \alpha \qquad (\theta^2 = \theta_2^2)$$

for all rational $\vartheta(\theta)$, hence, by continuity with respect to ϑ, also for all $\vartheta(\theta)$. By completeness, then,

$$\alpha_1^\infty(t) = \alpha \qquad \{\tau \cdot P_\theta\}_{\theta^1 = \theta_1^1} \qquad -\text{a.e.}$$

$$\alpha_2^\infty(t) = \alpha \qquad \{\tau \cdot P_\theta\}_{\theta^1 = \theta_2^1} \qquad -\text{a.e.}$$

By $\tau \cdot P_\theta \ll \tau \cdot P_{\Theta'}$ for all θ, θ' in Θ, we have

$$\lim_i \alpha_1^i = \lim_i \alpha_2^i = \alpha \qquad \{\tau \cdot P_\theta\}_{\theta \in \Theta} \qquad -\text{a.e.}$$

By remark I, we can conclude that there is a subset M of R^{k-1} with $(\tau \cdot P_\theta)(M) = 1$ for all $\theta \in \Theta$ such that

$$\int \phi_{\alpha_1^i,\alpha_2^i} P_\theta\,(dx \mid \tau = t) \to \int \phi_{\alpha,\alpha} P_\theta\,(dx \mid \tau = t)$$

for all $\theta^1 \in \Theta^1$ and $t \in M$, where $\phi_{\alpha,\alpha}$ is the α-UMPU test for $\{P_\theta(\cdot \mid \tau = t)\}$. By bounded convergence and Fubini,

$$\int \phi_{\alpha_1^i,\alpha_2^i} P_\theta\,(dx) \to \int \phi_{\alpha,\alpha} P_\theta\,(dx),$$

and

$$\int \phi_\infty P_\theta\,(dt) = \int \phi_{\alpha,\alpha} P_\theta\,(dx),$$

where $\phi_{\alpha,\alpha}$ is α-UMPU (thus admissible minimax) for $\{P_\theta\}_{\theta \in \Theta}$. The regularity theorem II is thus also true for the case of several parameters.

§ 4 An Example With Binomial Distributions

The example to be discussed here is meant to demonstrate that the regularity theorems of § 3 can be used to prove meaningful limit theorems in special cases. The simplest non-trivial example is the problem of testing the hypothesis that the parameter p of a binomial distribution is equal to one half. Consider, first, the case where the parameter space Θ is the open unit interval. For the nullhypothesis, $H_0 := \{p = \frac{1}{2}\}$, there is essentially just one admissible minimax for the problem, and it is the test ϕ usually prescribed. It's a fact, however, that if we replace Θ by the union of the set $\{\frac{1}{2}\}$ and any non empty finite subset, Ω, of $\Theta - \{\frac{1}{2}\}$, then the new problem thus defined has essentially just one minimax test. By virtue of its uniqueness, the test must also be ad-

missible, and an obvious question concerns the behavior of a sequence $\{\phi_i\}$ of such tests as the finite subsets $\Omega_i \subset \Theta - \{\frac{1}{2}\}$ become increasingly dense in $\Theta - \{\frac{1}{2}\}$.

Under acceptable conditions on the sequence $\{\Omega_i\}$, we can show that the sequence $\{\phi_i\}$ converges to the test ϕ in the sense of the topology generated on $\mathcal{D}$ by the class of binomial distributions. It is therefore possible to arrive at the admissible minimax for the binomial problem by fairly well defined limiting processes. One might argue that if this were not so, any minimax solution to the problem would have to be regarded with considerable skepticism. There are many problems in parametric statistics, some of them being so far unsolved, where either a similar limiting process must yield an acceptable test or the suitability of the minimax as such must be placed in question.

We prove

Theorem III: Let $n > 0$ and $(\Theta^*, \mathcal{D}, R)$ be a problem where Θ is the open unit interval, $\mathcal{D}$ the class of critical functions on $X = \{k \mid 0 \leqslant k \leqslant n, k \text{ integer}\}$, and R is the function defined by (3.1) where $H_0 = \{p_0\} \subset \Theta$, and P_θ is equal to the binomial distribution on X with $p = \theta$. If $\epsilon > 0$, let $\Omega_\epsilon \subset \Theta - H_0$ be finite set such that, for any point $\theta \in \Theta - H_0$, there is a point θ' in Ω_ϵ with $|\theta - \theta'| < \epsilon$. Then, for each Ω_ϵ, there is a unique minimax test $\phi_\epsilon \in \mathcal{D}$ for the problem $([\Omega_\epsilon \cup H_0]^*, \mathcal{D}, R)$, and

$$\lim_{\epsilon \to 0} \int \phi_\epsilon P_\theta\,(dx) = \int \phi_\theta\,(dx)$$

for all $\theta \in \Theta$, where ϕ is the admissible minimax for $(\Theta^*, \mathcal{D}, R)$.

Proof: The conditions on $\{\Omega_\epsilon\}$ imply convergence of Ω_ϵ to the set $\Theta - H_0$ in the euclidean norm. On $\Theta - H_0$, *Wald*'s ρ_4 equals ρ_1, where, for p, p' in the closed unit interval,

$$\rho_1(p, p') = \sup_E |P_p(E) - P_{p'}(E)|,$$

where P_p is the binomial distribution with parameter p. Since $\lim_{|p-p'| \to 0} \rho_1(p, p') = 0$ in $[0,1]$, the metric ρ_1 is equivalent to the euclidean metric on $[0,1]$, thus also on $\Theta - H_0$, and $\Omega_\epsilon \to \Theta - H_0$ in the sense of ρ_1. Hence, $\Omega_\epsilon \cup H_0 \to \Theta$ and, using the separability of Θ, $[\Omega_\epsilon \cup H_0]^* \to \Theta^*$ in the sense of ρ_4. By the property (2.1) of R, therefore, the problems $([\Omega_\epsilon \cup H_0]^*, \mathcal{D}, R)$ converge to $(\Theta^*, \mathcal{D}, R)$. An application of regularity theorem II proves theorem III.

References

Bourbaki, N.: Topologie Générale Ch. 9 Livre III, Paris 1958.
Brown, C.C.: A conventional statistical problem with irregular minimax behavior.
Ferguson, T.S.: Mathematical statistics, New York–Frankfurt–Paris 1967.
Heyer, H.: Mathematische Theorie statistischer Experimente, Berlin–Heidelberg–New York 1973.
Kelley, J.L.: General Topology, London–New York 1955.
Köthe, G.: Topologische Lineare Räume I, Berlin–Heidelberg–New York 1960.
Le Cam, L.: Sufficiency and Approximate Sufficiency, Ann. Math. Stat. **35**, 1419–1455, 1964.
Lehmann, E.L.: Testing statistical hypotheses, New York 1959.
Wald, A.: Statistical Decision Functions, New York 1950.

Eingegangen am 29. August 1974
Revidierte Fassung am 18. Dezember 1974.