Hashimoto, I. and Kudō, H.
Osaka J. Math.
5 (1968), 137-150

A UNIFIED CONCEPT OF ADMISSIBILITY OF STATISTICAL DECISION FUNCTIONS

ISAO HASHIMOTO AND HIROKICHI KUDŌ

(Received December 8, 1967)

The purpose of the present paper is to introduce a new and unified concept of admissibility in theory of statistical decision functions, and to give and discuss criteria for such an admissibility. The criteria proposed in this paper are not quite new, but rather a unification of criteria given by Karlin [6], Stein [10], [11], [12], Stone [13] and Takeuchi [14], and also of the equation defining the Bayes solution in the wide sense (see Wald [15] and LeCam [9]).

In Section 1 we give a generalized definition of admissibility which depends upon a family $\sum$ of subsets of the parameter space Θ. In case of $\sum$ being the family of all subsets of Θ our admissibility coincides with the usual one, whereas it is the negative of uniform improvability on Θ in case of $\sum$ being the single-element family $\{\Theta\}$. In general, we define the admissibility depending on $\sum$ as the negative of uniform improvability on some element of $\sum$. The concept of almost admissibility is one of the particular examples of such an admissibility.

A sufficient condition for such an admissibility is given and discussed also in this section. This condition has appeared in various forms (*e.g.*, [9], [10], [11], [14] and [15]), but they will be shown to be equivalent in this section. Especially, in [9] and [15], the procedure satisfying the condition is named as a Bayes solution in the wide sense provided that $\sum=\{\Theta\}$. In Section 2 we assume the subconvexity and the property (W) of the space D of available decision procedures. Then the sufficient condition given in Section 1 turns out to be also necessary and a complete class theorem similar to a theorem due to Wald and to LeCam ([9] and [15]) holds. The proofs contained in this section mostly follow LeCam's method in his paper [9].

In the first two sections we concern only with discrete prior probability measures with finite support. However they are not easy to treat with for the mathematical analysis, and so frequently we prefer to treat with continuous prior probability measures to discrete ones. In Section 3, we shall study the cases of general continuous prior measures and of continuous prior measures with densities. Under Assumptions of the subconvexity and the property (W) of D it is shown that there is no essential difference between the cases of general continuous prior measures and of discrete prior measures. However, for prior measures with densities, the situation is quite different. If we restrict ourselves to the case of continuous prior measures with densities, every sets of measure zero should be neglected, and hence only almost admissibility is to be concerned with, or the usual admissibility is concerned with only in the continuous risk function case. In his paper [12], Stein announced a variant of his criterion [10] for admissibility. Unfortunately it is stated in an ambiguous form (see Examples 2 and 3 in Section 3). The authors modify Stein's assertion and obtain a perfect criterion for almost admissibility (Thm. 9), on the basis of which we shall give proofs of admissibility of some examples (Examples 4 and 5). Takeuchi [14] gave a criterion (a sufficient condition) for admissibility in the continuous risk function case. It is shown that this criterion is also necessary (Thm. 10).

The concept of almost admissibility leads us properly to another concept of admissibility, the quasi admissibility. This is discussed briefly in Section 4.

The structure of the risk function is not taken into account in this paper, and hence the theorems contained here will be applied to zero-sum two-person games or to statistical games. However if the structure of risk function is given and used in the arguments, more profound results would be expected.

1. A general concept of admissibility

Consider a statistical game (Θ, D, r), where Θ is a parameter space, D a space of decision functions (d. f.'s) in which the statistician chooses his statistical strategy φ, and $r=r(\theta, \varphi)$ the risk function, bounded from below, uniformly in φ, when θ is the true parameter value. Let E be a nonempty subset of Θ, and ε a positive number.

DEFINITION 1. A strategy φ^1 is called an *(E, ε)-improvement* of a strategy φ^2 if $r(\theta, \varphi^1)\leqq r(\theta, \varphi^2)$ for all $\theta\in\Theta$, and $r(\theta, \varphi^1)\leqq r(\theta, \varphi^2)-\varepsilon$ for all $\theta\in E$. If such a strategy φ^1 exists, we say that φ^2 is *(E, ε)-improvable.*

Here and in the rest of this paper, the risk function $r(\theta, \varphi)$ may take the value ∞, and it is assumed for the convention that $\infty-\infty=0$, $\infty-a=\infty$ for any real number a and $a/0=\infty$ or $-\infty$ as $a\geq 0$ or <0.

Let $\sum$ be a family of nonempty subsets of Θ.

DEFINITION 2. A strategy φ is said to be *$\sum$-improvable*, if there is an E in $\sum$ and there is an $\varepsilon>0$ such that φ is (E, ε)-improvable. A strategy is said to be *$\sum$-admissible*, if it is not $\sum$-improvable.

Let Ξ_1 be the family of all prior probability measures on Θ, each of which assigns probability one to a finite subset of Θ. Here we shall introduce non-negative quantities $G_1(\varphi, E)$ and $H_1(\varphi, E)$, depending on a strategy φ and a nonempty subset E of Θ:

$$G_1(\varphi, E) = \inf_{\xi \in \Xi_1} \frac{r(\xi, \varphi) - \inf_{\varphi' \in D} r(\xi, \varphi')}{\xi(E)},$$

and

$$H_1(\varphi, E) = \sup_{\varphi' \in D} \inf_{\xi \in \Xi_1} \frac{r(\xi, \varphi) - r(\xi, \varphi')}{\xi(E)},$$

where $r(\xi, \varphi) = \sum_{i=1}^{n} \xi_i r(\theta_i, \varphi)$ with a prior measure ξ assigning ξ_i at a point $\theta_i \in \Theta$, $\xi_1 + \cdots + \xi_n = 1$, and $\xi_i > 0$. Evidently we have $0 \leqq H_1(\varphi, E) \leqq G_1(\varphi, E)$ for every $\varphi \in D$ and every nonempty $E \subset \Theta$.

Theorem 1. *If φ^0 is an (E, ε)-improvement of φ^*, then $G_1(\varphi^0, E) \leqq G_1(\varphi^*, E) - \varepsilon$ and $H_1(\varphi^0, E) \leqq H_1(\varphi^*, E) - \varepsilon$.*

Proof. By definition we have $r(\xi, \varphi^0) \leqq r(\xi, \varphi^*) - \varepsilon \cdot \xi(E)$ for every ξ in Ξ_1. Hence we have $\{r(\xi, \varphi^0) - r(\xi, \varphi)\}/\xi(E) \leqq \{r(\xi, \varphi^*) - r(\xi, \varphi)\}/\xi(E) - \varepsilon$ for every $\varphi \in D$, which implies our desired result.

Corollary 1. If φ is (E, ε)-improvable, then we have $\varepsilon \leqq H_1(\varphi, E) \leqq G_1(\varphi, E)$.

Corollary 2. If φ satisfies $H_1(\varphi, E) = 0$ or, in particular, $G_1(\varphi, E) = 0$ for each $E \in \sum$, then it is $\sum$-admissible.

Theorem 2. *If φ^0 satisfies $0 < \varepsilon < H_1(\varphi^0, E)$, then φ^0 is (E, ε)-improvable.*

Proof. From the assumption, there is a $\varphi \in D$ such that $\inf_{\xi \in \Xi_1} \{r(\xi, \varphi^0) - r(\xi, \varphi)\}/\xi(E) > \varepsilon$, or, in other words, that $\inf_{\theta \in \Theta} \{r(\theta, \varphi^0) - r(\theta, \varphi) - \chi_E(\theta) \cdot \varepsilon\} = \inf_{\xi \in \Xi_1} \{r(\xi, \varphi^0) - r(\xi, \varphi) - \varepsilon \cdot \xi(E)\} \geqq 0$, where $\chi_E(\theta)$ is the indicator function of E. The last inequality shows that φ is an (E, ε)-improvement.

Corollary. A strategy φ is $\sum$-admissible if and only if $H_1(\varphi, E) = 0$ for every $E \in \sum$.

Theorem 3. *The following four statements are equivalent each other:*

a) $G_1(\varphi^0, E) = 0$,

b) *There is a sequence $\{\xi_n\}$ of measures in Ξ_1 such that*

$$\lim_{n \to \infty} \frac{r(\xi_n, \varphi^0) - \inf_{\varphi \in D} r(\xi_n, \varphi)}{\xi_n(E)} = 0, \tag{1}$$

c) *For any $\varepsilon > 0$ there are a measure $\xi \in \Xi_1$ and a $\delta > 0$ such that*

$$\xi(E) > \delta, \quad \textit{and} \tag{2}$$

$$r(\xi, \varphi^0) - \inf_{\varphi \in D} r(\xi, \varphi) < \delta\varepsilon, \tag{3}$$

d) *There are a sequence* $\{\xi_n\}$ *of measures in* Ξ_1 *and a sequence* $\{A_n\}$ *of positive numbers such that*

$$\liminf_{n\to\infty} A_n\cdot\xi_n(E)>0\,, \quad \textit{and} \tag{4}$$

$$\liminf_{n\to\infty} A_n\{r(\xi_n,\varphi^0)-\inf_{\varphi\in D} r(\xi_n,\varphi)\}=0\,. \tag{5}$$

Proof. The equivalence of a) and b) is evident. Suppose a) holds. For any $\varepsilon>0$ there is a measure $\xi\in\Xi_1$ such that $\xi(E)>0$ and $r(\xi,\varphi^0)-\inf_{\varphi\in D}r(\xi,\varphi)<\varepsilon\xi(E)/2$ hold. Take a positive δ satisfying $\xi(E)>\delta>\xi(E)/2$, and we obtain (2) and (3). Suppose next that the statement c) holds for any $\varepsilon=1/n$. Put $\xi_n=\xi$, and $A_n=1/\delta$ for $\varepsilon=1/n$, $n=1, 2,\cdots$. Then we shall easily see that (4) and (5) hold. Lastly we shall assume that d) holds. From (4) there are a positive number M and a positive integer N such that $A_n\xi_n(E)>M$ for any $n>N$. Putting $A_n=M/\xi_n(E)$ in the formula (5), we obtain $G_1(\varphi^0, E)=0$.

By a small change of the above proof, we obtain a similar theorem about the quantity $H_1(\varphi, E)$.

Theorem 3′. *The following four statements are equivalent each other*:

a) $H_1(\varphi^0, E)=0$,

b) *For any* $\varphi\in D$ *there corresponds a sequence* $\{\xi_n\}$ *of measures in* Ξ_1 *such that* $\lim_{n\to\infty}\{r(\xi_n,\varphi^0)-r(\xi_n,\varphi)\}/\xi_n(E)\leqq 0$,

c) *For any* $\varepsilon>0$ *and* $\varphi\in D$ *there* are *a measure* $\xi\in\Xi_1$ *and a* $\delta>0$ *such that* $\xi(E)>\delta$ *and* $r(\xi,\varphi^0)-r(\xi,\varphi)<\delta\varepsilon$,

d) *For any* $\varphi\in D$ *there correspond a sequence* $\{\xi_n\}$ *of measures in* Ξ_1 *and a sequence* $\{A_n\}$ *of positive numbers such that* $\liminf_{n\to\infty} A_n\xi_n(E)>0$ *and* $\liminf_{n\to\infty} A_n\{r(\xi_n,\varphi^0)-r(\xi_n,\varphi)\}\leqq 0$.

Let $\sum_1$ and $\sum_2$ be two families of nonempty subsets of Θ, and write $\sum_1\leqq\sum_2$ if for any $E_2\in\sum_2$ there is a set $E_1\in\sum_1$ such that $E_1\subset E_2$. If $\sum_1\leqq\sum_2$, then $H_1(\varphi, E)=0$ for every $E\in\sum_1$, $G_1(\varphi, E)=0$ for every $E\in\sum_1$ and $\sum_1$-admissibility of φ imply $H_1(\varphi, E)=0$ for every $E\in\sum_2$, $G_1(\varphi, E)=0$ for every $E\in\sum_2$ and $\sum_2$-admissibility of φ, respectively.

Denote by $\sum_T$ and $\sum_\Theta$ the family of all nonempty subsets of Θ and the family consisting of a single element $\{\Theta\}$, respectively. And denote by $\sum_S$ the family of one-point sets of Θ. The $\sum_T$-admissibility and the $\sum_S$-admissibility coincide with each other, and are usually called as the admissibility. An element $\varphi\in D$ satisfying $G_1(\varphi, \Theta)=0$ is called as a Bayes solution in the wide sense [9], [15], or a Wald procedure. In the case of a topological space Θ with a countable base, denote by $\sum_G$ the family of nonempty subsets of Θ containing an open set. The $\sum_G$-admissibility will be called as the open admissibility. In the case of a parameter space Θ with a σ-field $\mathcal{A}$ and a σ-finite measure λ, denote by $\sum_\lambda$ the family of sets of positive measure. The $\sum_\lambda$-admissibility is called as λ-almost admissibility (see [5]).

DEFINITION 3. A family $\sum$ of nonempty subsets of Θ is said to be *separable*, if there is a countable subfamily $\sum'$ of $\sum$ such that $\sum' \leqq \sum$.

The family $\sum_G$ is evidently separable.

Theorem 4. *Suppose that $\sum$ be a separable family of nonempty subsets of Θ.*

(i) *$G_1(\varphi^0, E)=0$ for every $E \in \sum$ if and only if there are a sequence $\{\xi_n\}$ of measures in Ξ_1 and a sequence $\{A_n\}$ of positive numbers such that*

$$\liminf_{n\to\infty} A_n \cdot \xi_n(E) = 0 \quad \text{implies } E \notin \textstyle\sum, \text{ and} \tag{6}$$

$$\liminf_{n\to\infty} A_n\{r(\xi_n, \varphi^0) - \inf_{\varphi\in D} r(\xi_n, \varphi)\} = 0\,. \tag{7}$$

(ii) *A parallel statement to* (i) *holds when the quantity $G_1(\varphi^0, E)$ is replaced by the quantity $H_1(\varphi^0, E)$ and correspondingly the sequence $\{\xi_n\}$ and the sequence $\{A_n\}$ depend possibly on a choice of $\varphi \in D$, and further, instead of* (7),

$$\liminf_{n\to\infty} A_n\{r(\xi_n, \varphi^0) - r(\xi_n, \varphi)\} \leqq 0 \quad \textit{holds.}$$

Proof. Suppose that $G_1(\varphi^0, E)=0$ for every $E \in \sum$, namely that the statement d) of Theorem 3 holds for every $E \in \sum$. There is no loss of generality in assuming that $\sum = \{E_1, E_2, \cdots\}$, and that for any i there are a sequence $\{\xi_n^{(i)}\}$ of measures in Ξ_1 and a sequence $\{A_n^{(i)}\}$ of positive numbers such that

$$\lim_{n\to\infty} A_n^{(i)} \cdot \xi_n^{(i)}(E_i) = 1\,,$$

$$A_n^{(i)} \cdot \xi_n^{(i)}(E_i) > \frac{1}{2}\,, \quad \text{and}$$

$$A_n^{(i)}\{r(\xi_n^{(i)}, \varphi^0) - \inf_{\varphi\in D} r(\xi_n^{(i)}, \varphi)\} < \frac{1}{n^2}$$

for each $n=1, 2, \cdots$. Taking a probability measure $\eta_n = \sum_{i=1}^n A_n^{(i)} \cdot \xi_n^{(i)} / B_n$, where $B_n = \sum_{i=1}^n A_n^{(i)}$, we have the following inequalities:

$$\liminf_{n\to\infty} B_n \eta_n(E_i) \geqq \lim_{n\to\infty} A_n^{(i)} \xi_n^{(i)}(E_i) = 1\,, \quad \text{and}$$

$$\begin{aligned} &B_n\{r(\eta_n, \varphi^0) - \inf_{\varphi\in D} r(\eta_n, \varphi)\} \\ &= \sum_{i=1}^n A_n^{(i)} \cdot r(\xi_n^{(i)}, \varphi^0) - \inf_{\varphi\in D} \sum_{i=1}^n A_n^{(i)} \cdot r(\xi_n^{(i)}, \varphi) \\ &\leqq \sum_{i=1}^n A_n^{(i)}\{r(\xi_n^{(i)}, \varphi^0) - \inf_{\varphi\in D} r(\xi_n^{(i)}, \varphi)\} \\ &\leqq \sum_{i=1}^n \frac{1}{n^2} = \frac{1}{n}\,. \end{aligned}$$

These inequalities show that (6) and (7) hold. The other direction of implication is obvious. The latter half of the theorem can be seen similarly.

The statements a) and b) of Theorem 3 are the forms used by LeCam [9] and Wald [15] for the definition of a Bayes procedure in the wide sense. The statement b) of the same theorem is the form appeared in James and Stein's paper [5], and c) is the criterion proposed by Stein [10]. d) of Theorem 3 and (i) in Theorem 4 are similar to the form proposed by Stone [13], and d) of Theorem 3′ and (ii) in Theorem 4 are to Takeuchi [14] (see [7]).

2. Subconvex set D with the property (W)

In the preceding section we have observed that the quantity H_1 gives a complete criterion whereas $G_1=0$ is only a sufficient condition for admissibility. In the present section, we shall show the necessity of the condition $G_1=0$ under some additional assumptions on the class D of d.f.'s.

Assumption 1. D is subconvex, i.e., for any φ_1 and φ_2 in D and any real number $\alpha\in[0, 1]$ there is a φ_3 in D such that $r(\theta, \varphi_3)\leqq\alpha r(\theta, \varphi_1)+(1-\alpha)r(\theta, \varphi_2)$ for every $\theta\in\Theta$.

Let $\mathcal{F}$ be the family of all extended real nonnegative functions defined on Θ, equipped with Tychonov topology, that is, the direct product topology in a product space $[-\infty, \infty]^{\Theta}$ of replicas of the compact space $[-\infty, \infty]$. It is well known that $\mathcal{F}$ is compact in this topology.

Assumption 2. D possesses the property (W), i.e., for any accumulation point $f(\cdot)$ of the subset $R=\{r(\cdot, \varphi): \varphi\in D\}$ in $\mathcal{F}$, there exists a d.f. φ^* in D such that $r(\theta, \varphi^*)\leqq f(\theta)$ for all $\theta\in\Theta$.

For the reference to Assumption 2, see [8].

Naturally the property (W) holds for every subset of Θ whenever it does for Θ. Now we are going to prove the inverse statement of Theorem 1 under the Assumptions 1 and 2. To proceed to this end, we have to study the properties of a subconvex set possessing the property (W) in an extended Euclidean space $E^n=[-\infty, \infty]^n$.

For two points x and y in E^n, we shall write "$x\leqq y$" for the sentence "every coordinate of x is not larger than the corresponding coordinate of y", and also write "$x<y$" if $x\leqq y$ but $x\neq y$. Suppose that S be a set in the first quadrant in E^n and satisfies the Assumptions 1 and 2:

(Subconvexity) For any pair x_1 and x_2 in S and for any real number $\alpha\in[0, 1]$ there is a point x_3 in S such that $x_3\leqq\alpha x_1+(1-\alpha)x_2$.

(Property (W)) For any accumulation point x of S there is a point y in S such that $y\leqq x$.

Let H be the set of nonzero vectors $\xi=(\xi_1, \cdots, \xi_n)$ with nonnegative coordinates $\xi_1, \cdots, \xi_n$ and with $\sum_{i=1}^n\xi_i=1$.

Lemma 1. If S satisfies Assumptions 1 and 2, and a point x satisfies

$\xi\cdot x \geqq \inf_{y\in S}\xi\cdot y$ for every $\xi\in H$, then there is a point $y^*\in S$ such that $y^*\leqq x$.

Several propositions similar to this Lemma 1 are known (see *e.g.* [1]), and a slight modification of the proof of these propositions will give a proof in our case. So the proof will be omitted.

Let E be a subset of $\{1, 2, \cdots, n\}$, and $\mathbf{1}_E$ an n-vector whose i-th component equals one if $i\in E$ and zero if $i\notin E$.

Lemma 2. Suppose that S be a set of points in the first quadrant of E^n and satisfy Assumptions 1 and 2. If a point $x\in S$ satisfies

$$\inf_{\xi\in H}\{(\xi\cdot x-\inf_{y\in S}\xi\cdot y)/\xi\cdot\mathbf{1}_E\}\geqq\varepsilon>0\,, \tag{8}$$

then there is a point $y^*\in S$ such that $y^*\leqq x-\varepsilon\mathbf{1}_E$.

Proof.[1] From (8) it follows that, for every $\xi\in H$, $\xi\cdot(x-\varepsilon\mathbf{1}_E)>\inf_{y\in S}\xi\cdot y$. Therefore, by Lemma 1, there is $y^*\in S$ such that $y^*\leqq x-\varepsilon\mathbf{1}_E$.

Theorem 5. *Under Assumptions* 1 *and* 2, $G_1(\varphi^0, E)\geqq\varepsilon$ *is a necessary and sufficient condition for* φ^0 *being* (E, ε)*-improvable.*

Proof. The necessity has been shown in Corollary 1 of Theorem 1. For the sufficiency, let N be a finite subset of Θ. By Δ_N we denote the set of all real extended functions $f(\theta)$ of θ satisfying the following conditions: 1) $\inf_{\varphi\in D}r(\theta, \varphi)\leqq f(\theta)\leqq r(\theta, \varphi^0)$ for every $\theta\in\Theta$, 2) $f(\theta)\leqq r(\theta, \varphi^0)-\varepsilon$ for every $\theta\in N\cap E$ and 3) there exists a d.f. φ^* in D such that $f(\theta)=r(\theta, \varphi^*)$ for every $\theta\in N$. By Lemma 2, we can easily see that Δ_N is not empty. Since this holds for every finite set N, the family $\{\Delta_N\}$ has the finite intersection property. Therefore the intersection of the closures of Δ_N's in $\mathcal{F}$ has at least one element $g(\theta)$ which, by the property (W) of D, should be equal to or larger than the risk function $r(\theta, \varphi^*)$ of a certain (E, ε)-improvement φ^* of φ^0.

Theorem 5 is closely related to Mazur theorem in the theory of topological linear spaces (see *e.g.* [4]). As a direct implication of Theorem 5, we have

Corollary. If D satisfies Assumptions 1 and 2, then the $\sum$-admissibility of $\varphi\in D$ implies that $G_1(\varphi, E)=0$ for every $E\in\sum$.

Theorem 6. *Let E be a nonempty subset of* Θ, *and suppose that D satisfy Assumptions* 1 *and* 2. *Then the class of d.f.'s* $\varphi\in D$ *with* $G_1(\varphi, E)=0$ *is a complete class in D.*

Proof. Assume that a d.f. $\varphi^0\in D$ satisfies $0<G_1(\varphi^0, E)<\infty$. We can see easily that there is no loss of generality in doing so. By Theorem 5, there is a d.f.

1) This proof is due to the referee. The original proof given by the authers was longer than this.

$\varphi^* \in D$ such that $r(\theta, \varphi^*) \leqq r(\theta, \varphi^0) - G_1(\varphi^0, E)\chi_E(\theta)$, where $\chi_E(\theta)$ is the indicator function of the set E. And by Theorem 1, φ^* satisfies $G_1(\varphi^*, E) \leqq G_1(\varphi^0, E) - G_1(\varphi^0, E) = 0$, which shows that φ^* with $G_1(\varphi^*, E) = 0$ improves φ^0.

Theorem 7. *Let* $\sum$ *be a family of nonempty subsets of the parameter space* Θ. *Under Assumptions* 1 *and* 2, *the class* $\mathrm{A}(\sum)$ *of* $\sum$*-admissible d.f.'s is a complete class in D.*

Proof. Take a d.f. φ^0 outside of $A(\sum)$ and a finite number of sets $E_1, \cdots, E_n$ from $\sum$. Denote by $C(E)$, $E \in \sum$, the class of all d.f.'s φ satisfying 1) $r(\theta, \varphi) \leqq r(\theta, \varphi^0)$ for all $\theta \in \Theta$, 2) $r(\theta, \varphi) \leqq r(\theta, \varphi^0) - \varepsilon$ for every $\theta \in E$ with a certain $\varepsilon > 0$, and 3) $G_1(\varphi, E) = 0$. Suppose that $G_1(\varphi^0, E_1) > 0$. Then by Theorem 6 there is a d.f. $\varphi^1 \in C(E_1)$. If $G_1(\varphi^1, E_2) > 0$, then again by Theorem 6 we have a d.f. $\varphi^2 \in C(E_2) \cap C(E_1)$. Continuing this procedure, we get a d.f. φ^n belonging to $C(E_1) \cap \cdots \cap C(E_n)$. Therefore the family $\{C(E) \colon E \in \sum\}$ has the finite intersection property. Since $\mathcal{F}$ is compact, the intersection of the closures of $\{r(\theta, \varphi) \colon \varphi \in C(E)\}$ for all $E \in \sum$ contains at least one $f \in \mathcal{F}$. Since the set $R^0 = \{g \colon g(\theta) \leqq r(\theta, \varphi^0)$ for all $\theta\}$ is closed in $\mathcal{F}$, f belongs to R^0. By the property (W), there is a d.f. φ^* such that $r(\theta, \varphi^*) \leqq f(\theta)$ for all $\theta \in \Theta$. It is clear that $r(\theta, \varphi^*) \in R^0$, *i.e.*, $r(\theta, \varphi^*) \leqq r(\theta, \varphi^0)$ for every $\theta \in \Theta$ and that $G_1(\varphi^*, E) = 0$ for every $E \in \sum$. Thus we see that φ^* is a $\sum$-improvement of φ^0 and belongs to $A(\sum)$.

A modification of Stein's theorem [10] is:

Corollary 1. The class $A(\sum_S)$ is the minimal complete class in D, provided that D satisfies Assumptions 1 and 2.

Proof. It is enough to recall the Wald theorem [15] that if the class of admissible (in the usual sense) procedures is complete, then it is a minimal complete class.

A complete class theorem due to Wald [15] and LeCam [9] is a direct implication of Theorems 5 and 7:

Corollary 2. Under Assumptions 1 and 2, a d.f. φ with $G_1(\varphi, \Theta) > \varepsilon$ is uniformly improvable, and the class of Bayes solutions in the wide sense is a complete class in D.

3. The case of continuous prior distributions

In this section we consider the case where the parameter space Θ is a measurable space with a σ-algebra $\mathcal{A}$ of its subsets, and the risk function $r(\theta, \varphi)$ is $\mathcal{A}$-measurable for every $\varphi \in D$. Denote by Ξ the family of all probability measures on $\mathcal{A}$, and by $r(\xi, \varphi)$ the integral of $r(\theta, \varphi)$ with respect to a measure

$\xi\in\Xi$ over Θ. Define

$$G(\varphi^0, E)=\inf_{\xi\in\Xi}\frac{r(\xi,\varphi^0)-\inf_{\varphi\in D} r(\xi,\varphi)}{\xi(E)}, \quad \text{and}$$

$$H(\varphi^0, E)=\sup_{\varphi\in D}\inf_{\xi\in\Xi}\frac{r(\xi,\varphi^0)-r(\xi,\varphi)}{\xi(E)}.$$

Assumption 3. Every one-point set belongs to $\mathcal{A}$.

Under Assumption 3, it is clear that $G(\varphi, E)\leqq G_1(\varphi, E)$ for every $\varphi\in D$ and $E\in\mathcal{A}$, but the inverse inequality is not necessarily true, as is seen in the following example.

EXAMPLE 1. Let Θ be the interval (0, 1), and E a finite subset in Θ. Let a d.f. φ_E be such that its risk function $r(\theta, \varphi_E)$ coincides with $1-\chi_E(\theta)$, where $\chi_E(\theta)$ is the indicator function of E. Suppose that D be the class of all such d.f.'s. In this case, the d.f. φ_ϕ corresponding to the empty set ϕ satisfies $G(\varphi_\phi, \Theta)=0$ whereas $G_1(\varphi_\phi, \Theta)=1$. Notice that φ_ϕ is $\{\Theta\}$-admissible.

REMARK 1. When we replace Ξ_1 by Ξ and correspondingly $H_1(\varphi, E)$ and $G_1(\varphi, E)$ by $H(\varphi, E)$ and $G(\varphi, E)$, respectively, we have parallel statements to Section 1.

Thus Stone theorem [13; Thm. 1.3] is direct from Theorem 1, (i) of Theorem 4 and Remark 1.

Theorem 8. *Under Assumptions* 1–3, *the following three statements are all equivalent*:

a) φ *is* (E, ε)*-improvable*,

b) $G(\varphi, E)\geqq\varepsilon$,

c) $G_1(\varphi, E)\geqq\varepsilon$,

where E is $\mathcal{A}$-measurable, and $\varepsilon>0$.

Proof. By Theorem 1 and Remark 1, a) implies b). The statement c) is direct from b), and the implication of a) from c) is due to Theorem 5.

In the next stage let Θ be a Euclidean space of finite dimension. Let λ be the Lebesgue measure on Θ. In this case $\mathcal{A}$ should be considered as the σ-field of Lebesgue (or Borel) measurable subsets of Θ. In the rest of this paper, we shall assume, without explicitly stating, that the risk function $r(\theta, \varphi)$ is Lebesgue measurable for each d.f. $\varphi\in D$. Denote by F the family of all non-negative λ-integrable functions $f(\theta)$ with $\int_\Theta f(\theta)d\lambda=1$, and write $r(f, \varphi)=\int_\Theta r(\theta, \varphi)f(\theta)d\lambda$. Define

$$G_2(\varphi^0, E)=\inf_{f\in F}\frac{r(f,\varphi^0)-\inf_{\varphi\in D} r(f,\varphi)}{\int_E f(\theta)d\lambda}$$

for every measurable E of positive λ-measure. It is evident that $G_2(\varphi, E)=0$ for every E with $\lambda(E)>0$ implies the λ-almost admissibility of φ, because $G_2(\varphi, E)\geqq G(\varphi, E)$ and Theorem 1 holds (see Remark 1). However Example 2 below shows that the λ-almost admissibility of φ does not necessarily imply $G_2(\varphi, E)=0$, even if Assumptions 1–3 are satisfied.

EXAMPLE 2. Suppose that the parameter space Θ is the real line. Let E be a bounded open set in Θ, and N a non-empty set of λ-measure zero and disjoint of E. Suppose that $r(\theta, \varphi^0)=1$ constantly on Θ, and $r(\theta, \varphi^1)=0$ for $\theta\in E$, $=2$ for $\theta\in N$ and $=1$ for the other θ's. For any α in (0, 1) we define $r(\theta, \varphi^{\alpha})=\alpha r(\theta, \varphi^1)+(1-\alpha)r(\theta, \varphi^0)$ for every θ in Θ. Put $D=\{\varphi^{\alpha}: 0\leqq\alpha\leqq1\}$. Evidently Assumptions 1–3 are all satisfied. However we can easily observe that φ^0 is λ-almost admissible whereas $G_2(\varphi^0, E)=1$ and $G(\varphi^0, E)=0$.

By a similar method to the above, it is easy to construct an example in which φ^0 is λ-almost admissible but not admissible in the usual sense, and further $G_2(\varphi^0, E)\geqq G(\varphi^0, E)>0$ for a open set E with compact closure.

Example 2 is given for a counterexample to Stein's assertion in [12, p. 232]. Another counterexample to this assertion is given below:

EXAMPLE 3. Suppose that $\Theta=(-\infty, +\infty)$ and K be a given nowhere dense closed set of positive measure in Θ. Let $r(\theta, \varphi^0)$ equal one constantly on Θ, and $r(\theta, \varphi^1)$ coincide with $r(\theta, \varphi^0)$ everywhere on Θ except only for the set K where it takes value zero. For $D=\{\varphi^0, \varphi^1\}$, it is obvious that φ^0 is not λ-almost admissible but $G_2(\varphi^0, E)=0$ for any bounded open set E.

A modification of Stein's assertion in [12; p. 232] is as follows:

Theorem 9. *If for any finite interval J in Θ there are a sequence $\{f_n\}$ of elements of F and a sequence $\{A_n\}$ of positive numbers such that* a) $\liminf_{n\to\infty} A_n f_n(\theta)>0$ *for almost every θ in J and* b) $\liminf_{n\to\infty} A_n\{r(f_n, \varphi^0)-\inf_{\varphi\in D} r(f_n, \varphi)\}=0$, *then φ^0 is λ-almost admissible.*

Proof. Let E be a given set of positive measure contained in J and put $\alpha=\int_E \liminf_{n\to\infty} A_n f_n(\theta)d\lambda$. Then by Fatou lemma we have $\alpha\leqq\liminf_{n\to\infty} A_n \int_E f_n(\theta)d\lambda$. From the condition b), for any $\varepsilon>0$ we can choose a sequence $\{n_i\}$ of positive integers such that $A_{n_i}\{r(f_{n_i}, \varphi^0)-\inf_{\varphi\in D} r(f_{n_i}, \varphi)\}<\alpha\cdot\varepsilon$ for each $i=1, 2, \cdots$. On the other hand, for a sufficiently large i it holds that $\alpha/2<A_{n_i}\int_E f_{n_i}(\theta)d\lambda$. This shows that $G_2(\varphi^0, E)=0$, and so $G(\varphi^0, E)=0$ for every E of positive measure in J. Thus by Theorem 1 and Remark 1 φ^0 is λ-almost admissible in J. Since J is arbitrary, φ^0 is λ-almost admissible.

When the risk function $r(\theta, \varphi)$ of d.f. φ is continuous for every $\varphi\in D$, the concepts of the almost admissibility, the open admissibility and the admissibility in the usual sense coincide with each other.

EXAMPLE 4. Let the sample space X be the real line, and $\boldsymbol{B}$ the σ-algebra of Borel sets in X. Let the parameter space Θ be $(-\infty, \infty)$. The sample distribution $p(x|\theta)$ of x is normal $N(\theta, 1)$ when θ is true. We shall show, as an application of Theorem 9, the admissibility of the estimate $\varphi^0(x)=x$ of θ when the quadratic loss is considered, which has been already proved in [2], [3] and [6]. We take $A_n=\sqrt{2\pi}\,n$ and $f_n(\theta)=\dfrac{1}{\sqrt{2\pi}n}e^{-(\theta^2/2n^2)}$. It is not difficult to examine that the conditions a) and b) of Theorem 9 hold. Therefore $\varphi^0(x)$ is almost admissible. Since the risk function $r(\theta, \varphi)$ is continuous on Θ for each d.f. φ, φ is admissible in the usual sense.

EXAMPLE 5. Consider the sample space $X=1, 2, \cdots$, the parameter space $\Theta=(0, 1)$ and the sample distribution $p(x|\theta)=\theta^{x-1}(1-\theta)$. For this problem we shall show the admissibility of the m.l.e. $(x-1)/x$ of θ, when the loss is quadratic. We take $A_n=B(1/n, 1/n)=\int_0^1\theta^{(1/n)-1}(1-\theta)^{(1/n)-1}d\theta$ and $f_n(\theta)=\dfrac{1}{B(1/n, 1/n)}\theta^{(1/n)-1}(1-\theta)^{(1/n)-1}$ and we can easily seen that $\inf r(f_n, \varphi)=r(f_n, \varphi_n)$ is satisfied by $\varphi_n(x)=\dfrac{x+(1/n)-1}{x+(2/n)}$. From the fact that $\sup\limits_{x=1,2\cdots}|\varphi_n(x)-\varphi^0(x)|\leqq\dfrac{1}{n}$ and $A_n\int_0^1 f_n(\theta)d\theta\leqq 2n$, we have

$$A_n\{r(f_n, \varphi^0)-\inf r(f_n, \varphi)\}\leqq\frac{2}{n}.$$

Thus *a*) and *b*) of Theorem 9 are fulfiled. Therefore φ^0 is almost admissible. On the other hand, $r(\theta, \varphi)=\sum_{x=1}^{\infty}(\varphi(x)-\theta)^2p(x|\theta)$ is a uniformly convergent series for the bounded φ and hence it is a continuous function of θ. Since the class of bounded d.f.'s is a complete class, φ^0 is admissible.

Here we have to notice that φ^0 is a generalized Bayes solution for an improper prior measure $d\eta=\dfrac{d\theta}{\theta(1-\theta)}$ and the $A_nf_nd\theta$ tends to $d\eta$ in some sense.

Takeuchi gave the following condition of admissibility for the continuous risk function case [14; Thm. 2.6]:

Theorem 10. *Suppose that the risk function $r(\theta, \varphi)$ of $\varphi\in D$ is continuous on Θ for each φ. Then φ^0 is admissible if and only if there are a sequence $\{f_n\}$ of $f_n\in F$ and a sequence $\{A_n\}$ of positive numbers such that* 1) $\liminf_{n\to\infty}\int_E A_nf_n(\theta)d\lambda=0$

implies that for any open set G the difference $G-E$ is of positive λ-measure,
2) $\liminf_{n\to\infty} A_n\{r(f_n, \varphi^0)-r(f_n, \varphi)\} \leqq 0$ *for every* $\varphi \in D$.

Proof. The sufficiency of the condition is direct from (ii) of Theorem 4 and Remark 1 (see also Takeuchi [14]). For the proof of the necessity, it is enough to give sequences $\{f_n\}$ and $\{A_n\}$ for which 1) and 2) hold when φ^0 is admissible. Let N, M, k and m be all positive integers, and $\{\theta_1, \theta_2, \cdots\}$ a countable dense subset of Θ. Define

$$g_N(\theta) = \begin{cases} 1, & \text{if } -N<\theta<N\,, \\ 0, & \text{otherwise} \end{cases}$$

$$h_{Mkm}(\theta) = \begin{cases} M, & \text{if } \theta_k-\dfrac{1}{m}<\theta<\theta_k+\dfrac{1}{m}\,, \\ 0, & \text{otherwise, and} \end{cases}$$

$$f_{NMkm}(\theta) = \{g_N(\theta)+h_{Mkm}(\theta)\}/A_{NMm}, \quad \text{where} \quad A_{NMm} = 2\left(N+\frac{M}{m}\right).$$

Rearrange the four-fold sequences $\{f_{NMkm}\}$ and $\{A_{NMm}\}$ into simple sequences $\{f_n\}$ and $\{A_n\}$, respectively, in a common order. Then the sequences $\{f_n\}$ and $\{A_n\}$ satisfy the conditions 1) and 2) in the theorem. In fact, let φ be an arbitrary d.f. in D, and denote by G the set $\{\theta;\, r(\theta, \varphi^0)<r(\theta, \varphi)\}$. Clearly G contains an interval $I_{km}=(\theta_k-1/m, \theta_k+1/m)$ for certain k and m which are chosen as integers larger than an arbitrarily given integer. By V_{km} we denote the integral of $r(\theta, \varphi)-r(\theta, \varphi^0)$ over I_{km}. Let M be larger than $\int_{\Theta-G}\{r(\theta, \varphi^0)-r(\theta, \varphi)\}g_N(\theta)d\lambda/V_{km}$. Then we have

$$\begin{aligned} &A_{NMm}\{r(f_{NMkm}, \varphi^0)-r(f_{NMkm}, \varphi)\} \\ &\quad = r(g_N, \varphi^0)-r(g_N, \varphi)+r(h_{Mkm}, \varphi^0)-r(h_{Mkm}, \varphi) \\ &\quad \leqq \int_{\Theta-G}\{r(\theta, \varphi^0)-r(\theta, \varphi)\}g_N(\theta)d\theta-V_{km}M \leqq 0\,. \end{aligned}$$

This shows that 2) holds. The condition 1) is clear.

Takeuchi [14] gave several examples of applications of Theorem 10.

REMARK 2. The original form of Takeuchi's criterion is slightly different from ours, in such a way that the measures in our criterion in Theorem 10 are absolutely continuous with respect to Lebesgue measure whereas he does not assume the absolute continuity of the measures. Nevertheless, his criterion is also a necessary and sufficient condition for admissibility.

4. Quasi admissibility

Even for a λ-almost admissible d.f. φ^0 there might exist a d.f. φ^* such that $r(\theta, \varphi^*) \leqq r(\theta, \varphi^0)$ λ-almost everywhere on Θ, and $r(\theta, \varphi^*) < r(\theta, \varphi^0)$ on a set of positive λ-measure. Suppose that a statistician be satisfied with the almost admissibility of a d.f. There does remain no excuse for him in rejecting the increment of the risk upon a set of λ-measure zero if he obtains an improvement on a set of positive measure, and so he would prefer φ^* to φ^0. Such a consideration leads us to the following concept of admissibility.

DEFINITION 4. A d.f. $\varphi \in D$ is said to be *quasi admissible* if there is no d.f. $\varphi^* \in D$ such that $r(\theta, \varphi^*) \leqq r(\theta, \varphi^0)$ λ-almost everywhere on Θ, and $r(\theta, \varphi^*) < r(\theta, \varphi^0)$ on a set of positive λ-measure.

Let E be a member of $\sum_\lambda$, and denote by $I(f: E)$ the integral of $f(\in F)$ over E with respect to Lebesgue measure λ. Here we notice that the inequality $r(f, \varphi) \leqq r(f, \varphi^0) - \varepsilon I(f: E)$ for every $f \in F$ is equivalent to the statement that $r(\theta, \varphi) \leqq r(\theta, \varphi^0)$ λ-almost everywhere on Θ and $r(\theta, \varphi) \leqq r(\theta, \varphi^0) - \varepsilon$ holds on a set E' with $\lambda((E-E') \cup (E'-E)) = 0$.

Assumption 2′. The space D of d.f.'s possesses the property (W) when F is regarded as a parameter space and $r(f, \varphi)$, $f \in F$, as the risk function of $\varphi \in D$.

By a small modification of the proof of Theorem 1, we have

Theorem 11. *It is a necessary and sufficient condition for φ^0 being quasi admissible that*

$$\sup_{\varphi \in D} \inf_{f \in F} \frac{r(f, \varphi^0) - r(f, \varphi)}{I(f: E)} = 0$$

for every $E \in \sum_\lambda$. And, moreover, under Assumptions 1 and 2′ $G_2(\varphi^0, E) = 0$ for every $E \in \sum_\lambda$ is a necessary and sufficient condition for φ^0 being quasi admissible.

In the case of continuous risk functions, the quasi admissibility coincides with the admissibility in the usual sense. Hence the conditions for quasi admissibility become conditions for the admissibility in this case.

5. Acknowledgement

The authors are indebted to many people who attended to the Symposium on "The Admissibility of Statistical Decision Functions", held at the Institute of Mathematical Science, Kyoto University, and, in particular, to Goro Ishii, Keiiti Isii, Tokitake Kusama, Haruki Morimoto, Osamu Takenouchi and Kei Takeuchi. We wish to express our hearty thanks to those people for their kind comments. Thanks are also due to the referee for his helpful comments.

OSAKA UNIVERSITY
OSAKA CITY UNIVERSITY

References

[1] D. Blackwell and M.A. Girshick: Theory of Games and Statistical Decisions, New York, John Wiley and Sons, 1954.

[2] M.A. Girshick and L.J. Savage: *Bayes and minimax estimates for quadratic loss functions*, Proc. Second Berkeley Symp. on Math. Statist. and Prob., Vol. I, 1950, 53–73.

[3] J.L. Hodges and E.L. Lehmann: *Some applications of the Cramér-Rao inequality*, Proc. Second Berkeley Symp. on Math. Statist. and Prob., Vol. I, 1950, 13–22.

[4] K. Isii: *Inequalities of the types of Chebychev and Cramér-Rao and mathematical programming*, Ann. Inst. Statist. Math. **16** (1964), 277–293.

[5] W. James and C. Stein: *Estimation with quadratic loss*, Proc. Fourth Berkeley Symp. on Math. Statist. and Prob., Vol. I, 1961, 361–380.

[6] S. Karlin: *Admissibility for estimation with quadratic loss*, Ann. Math. Statist. **29** (1958), 406–454.

[7] H. Kudō: *Some problems in the complete class theorems*, Talk at New York University, 1965. (unpublished paper).

[8] H. Kudō: *On the property* (W) *of the class of statistical decision functions*, Ann. Math. Statist. **37** (1966), 1631–1642.

[9] L. LeCam: *An extension of Wald's theory of statistical decision functions*, Ann. Math. Statist. **26** (1955), 69–81.

[10] C. Stein: *A necessary and sufficient condition for admissibility*, Ann. Math. Statist. **26** (1955), 518–522.

[11] C. Stein: *The admissibility of Pitman's estimator of a single location parameter*, Ann. Math. Statist. **30** (1959), 970–979.

[12] C. Stein: *Approximation of improper prior measures by prior measures*, Bernoulli-Bayes-Laplace, Aniversary Volume, Proc. of Intern. Res. Sem. at Statist. Lab., Univ. of California, Edited by J. Neyman and L. LeCam, Berlin, Springer-Verlag, 1965.

[13] M. Stone: *Generalized Bayes decision functions, Admissibility and the exponential family*, Ann. Math. Statist. **38** (1967), 816–822.

[14] K. Takeuchi: *Theory of statistical estimations*, II. Sugaku **16** (1965), 139–149. (in Japanese).

[15] A. Wald: Statistical Decision Functions, New York, John Wiley and Sons, 1950.