REMARKS ON ADMISSIBILITY OF DECISION FUNCTIONS

TOKITAKE KUSAMA

(Received April 22, 1964; revised April 13, 1966)

1. Introduction

In the general theory of decision functions, we have a few necessary and sufficient conditions for admissibility. For example, C. Stein gave one in the game-theoretic formulation [1]. In the present paper, we shall give a criterion for admissibility of decision functions using a theorem concerning a topological vector space.

2. Notations and definitions

Let $(\mathscr{X}, \mathscr{B})$ be a sample space of points x, where $\mathscr{B}$ is a σ-field of subsets of $\mathscr{X}$, and suppose that x is distributed in $\mathscr{X}$ according to some unknown member of a given set $\mathscr{P}$ of probability measures p_θ $(\theta \in \Omega)$.

Throughout this paper we assume the following:

Assumption (A). The action space A is metrizable, locally compact and σ-compact.

Assumption (B). The loss function $w(\theta, a)$ is lower semicontinuous on A and $w(\theta, a) \geqq 0$ for every $\theta \in \Omega$ and $a \in A$.

Assumption (C). Every $p_\theta(\theta \in \Omega)$ is absolutely continuous with respect to a σ-finite measure μ. We denote the density $dp_\theta/d\mu$ by $f(\cdot, \theta)$.

Let $C_0(A)$ be the totality of bounded continuous functions on A vanishing outside a compact set of A with norm $\|\alpha\| = \sup_{a \in A} |\alpha(a)|$. Let $M(\mathscr{X})$ be the set of all bounded $\mathscr{B}$-measurable functions with norm $\|f\| = \text{ess.}\sup_{x \in \mathscr{X}} |f(x)|$ and $\mathscr{F}$ the set of all probability measures on A. A decision function δ is a mapping of $\mathscr{X}$ to $\mathscr{F}$: $x \to \delta(\cdot; x)$ satisfying $\int_A \alpha(a)\delta(da; x) \in M(\mathscr{X})$ for all $\alpha \in C_0(A)$. We denote this integral by $\delta \circ \alpha$. Let $\mathscr{D}$ be the space of decision functions and $E(\mathscr{P})$ be the linear subspace in $L^1(\mathscr{X})$ spanned by $\{f(x, \theta) : \theta \in \Omega\}$. We write $\int_{\mathscr{X}} s(x)f(x)d\mu(x)$ by $f \circ s$ for $f \in E(\mathscr{P})$.

A topology with neighborhoods

$$V\{\delta^* : \varepsilon, f_i, \alpha_i, i=1, 2, \cdots, n\}$$
$$=\{\delta : |f_i\circ\delta\circ\alpha_i - f_i\circ\delta^*\circ\alpha_i| < \varepsilon, i=1, 2, \cdots, n\},$$

where α_i runs through $C_0(A)$ and f_i through $E(\mathscr{P})$, will be called the regular topology π. Henceforward all terms on topology will be used in this sense unless the contrary is specified.

Under these assumptions, the risk function

$$r(\theta, \delta)=\int_{\mathscr{X}}\int_A w(\theta, a)\delta(da; x)f(x, \theta)d\mu(x)$$

can be defined for every θ and δ.

Let U be the space of all numerical functions on Ω topologized by the topology τ of pointwise convergence on Ω. We denote by $\mathscr{M}$ the set of all a priori distributions on Ω having as its carrier a finite subset of Ω. In particular we denote by ξ_θ the a priori distribution satisfying $\xi_\theta(\{\theta\})=1$.

By Assumptions $(A)\sim(C)$, $r(\theta, \delta)$ is a lower semicontinuous function of δ for every fixed θ and so, for any $\xi\in\mathscr{M}$, $\int_\Omega r(\theta, \delta)d\xi(\theta)$ is also lower semicontinuous with respect to δ. Hereafter we denote this integral by $r(\xi, \delta)$. An a priori distribution in $\mathscr{M}$ can be also considered as a continuous linear functional on (U, τ) putting $\xi(f)=\int_\Omega f(\theta)d\xi(\theta)$.

A decision function δ^* is called admissible if $r(\theta, \delta)\leqq r(\theta, \delta^*)$ for all θ implies $r(\theta, \delta)=r(\theta, \delta^*)$ for all θ.

Throughout this paper we denote by S^α the closure of $S(\subset\mathscr{D})$ for the regular topology π and by $\bar{H}$ the closure of $H(\subset U)$ for the topology τ.

3. Results

Let Y be a topological real vector space and L be a convex family of linear functionals on Y.

THEOREM 1. *Let C be a closed convex subset of Y and s_0 be an element of Y. Suppose that C and s_0 satisfy the following assumptions:*

Assumption (1. a). Every $f(\in L)$ is lower semicontinuous on C.

Assumption (1. b). There exists $f_0(\in L)$ such that a set $\{s : f_0(s)\leqq u\}\cap C$ is compact for all $u<\infty$.

Assumption (1. c). For any $s\in C$, there exists $f\in L$ such that $f(s_0)< f(s)$.

Then there exists $f^*(\in L)$ such that $f^*(s_0)<\inf_{s\in C} f^*(s)$.

PROOF. Put $H_f=\{s : f_0(s)\leqq f_0(s_0), f(s)\leqq f(s_0)\}\cap C$ for $f\in L$. At first suppose that $H_f\neq\phi$ for all f. Because of lower semicontinuity of f, H_f

is closed. Suppose that $\{H_f : f \in L\}$ has the finite intersection property. Then we have $\bigcap_{f\in L} H_f \neq \phi$ since H_{f_0} is compact by Assumption (1. b), and $H_f \subset H_{f_0}$ for all $f \in L$. $s \in \bigcap_{f\in L} H_f$ implies $f(s) \leqq f(s_0)$ for all $f \in L$, which contradicts to Assumption (1. c). So there must exist finite members of L^*, say $f_1, f_2, \cdots, f_n$ such that $\bigcap_{i=1}^{n} H_{f_i} = \phi$. Hence, for any $s \in C$, there exists f_j such that $s \notin H_{f_j}$, which shows

$$f_0(s_0) < f_0(s) \qquad \text{or} \qquad f_j(s_0) < f_j(s) \, . \tag{1.1}$$

To any $s \in Y$ there corresponds a point $T(s) = (f_0(s), f_1(s), \cdots, f_n(s))$ of R^{n+1}. The set $Q = \{T(s) : s \in C\}$ is convex since C is convex and f_i is linear for all $i = 0, 1, \cdots, n$, and by (1.1) the point $T(s_0)$ is a boundary or an exterior point of Q. If $T(s_0)$ is a boundary point of Q, then there exist $s_n \in C$ $(n = 1, 2, \cdots)$ such that $T(s_n) \in V_{1/n}(T(s_0)) \cap Q$, where $V_\varepsilon(p)$ is an ε-neighborhood of p. We have

$$\begin{aligned} s_n &\in \{s : |f_i(s) - f_i(s_0)| \leqq 1/n, \ i = 0, 1, \cdots, n\} \cap C \\ &\subset \{s : f_0(s) \leqq f_0(s_0) + 1/n\} \cap C \\ &\subset \{s : f_0(s) \leqq f_0(s_0) + 1\} \cap C \qquad \text{for all } n. \end{aligned}$$

Hence, by Assumption (1.b), the sequence $\{s_n : n = 1, 2, \cdots\}$ has a point of accumulation $s^* \in C$ and because of lower semicontinuity of f_i, we have $f_i(s^*) \leqq f_i(s_0)$ for all $i = 0, 1, 2, \cdots, n$, which contradicts (1.1). Hence $T(s_0)$ must be an exterior point of Q, that is to say, there exist $n+1$ reals $b_0, b_1, \cdots, b_n$ such that $\sum_{i=0}^{n} b_i = 1$, $b_i \geqq 0$ for all i, and $\sum_{i=0}^{n} b_i f_i(s_0) < \inf_{s \in C} \sum_{i=0}^{n} b_i f_i(s)$. $\sum_{i=0}^{n} b_i f_i$ is also included in L. Putting $f^* = \sum_{i=0}^{n} b_i f_i$, we obtain the desired result.

If there exists $\tilde{f}$ such that $H_{\tilde{f}} = \phi$, we have $f_0(s) > f_0(s_0)$, or $\tilde{f}(s) > \tilde{f}(s_0)$ for any $s \in C$. Putting $T(s) = (f_0(s), \tilde{f}(s))$, $Q = \{T(s) : s \in C\}$, the proof is immediately obtained by using the same method as in the case of $H_f \neq \phi$ for all f.

We note that this theorem is similar to the Mazur's theorem but a little different since we restrict the class of linear functionals and assume (1.b) [2].

Using the above theorem, we shall prove two theorems on admissibility.

For this purpose we assume the following:

Assumption (2. a). We consider the class $\mathcal{D}^b$ of decision functions satisfying $r(\theta, \delta) < \infty$ for all θ.

Assumption (2. b). $\lim_{n\to\infty} \inf_{a \in E_n^C} w(\theta, a) = \infty$ holds for all $\theta \in \Omega$, where

$E_1, E_2, \cdots$ is an increasing sequence of compact sets covering A.

Assumption (2. c). There exists $\theta_0 \in \Omega$ such that p_{θ_0} is equivalent[1] to μ.

THEOREM 2. *Under Assumptions (2. a)~(2. c), a necessary and sufficient condition for a decision function* δ_0 *to be admissible in* $\mathcal{D}^b$ *is that, for any closed convex subset* F *of* $\mathcal{D}^b$ *satisfying* $r(\cdot, \delta_0) \notin \{r(\cdot, \delta) : \delta \in F\}$, *there exists* $\xi \in \mathcal{M}$ *such that*

$$r(\xi, \delta_0) < \inf_{\delta \in F} r(\xi, \delta).$$

PROOF. The proof of necessity is as follows. $\mathcal{M}$ is considered as the convex set of continuous linear functionals on (U, τ) as was remarked in Section 2. We denote by K the set $\{r(\cdot, \delta) : \delta \in F\}$. Consider an arbitrary admissible δ_0. Now, putting $Y=(U, \tau)$, $C=\overline{K}$, $L=\mathcal{M}$, $s_0=\delta_0$, and $f_0=\xi_{\theta_0}$ in Theorem 1, we shall prove assumptions of the preceding theorem hold. Assumption (1. a) is clearly satisfied as was stated in Section 2. As for Assumption (1. b), it is sufficient to prove that $\overline{\{r(\cdot, \delta): r(\theta_0, \delta) \leqq u, \ \delta \in F\}}$ is compact for τ for all $u < \infty$ since we have for an arbitrary positive ε,

$$\begin{aligned} (2.1) \qquad \{h : h(\theta_0) = \xi_{\theta_0}(h) \leqq u, \ h \in \overline{K}\} & \\ \subset \overline{\{h : h(\theta_0) \leqq u+\varepsilon, \ h \in K\}} & \\ = \overline{\{r(\cdot, \delta) : r(\theta_0, \delta) \leqq u+\varepsilon, \ \delta \in F\}}. & \end{aligned}$$

For this purpose, first we shall prove that the set $\{\delta : r(\theta_0, \delta) \leqq u\}$ is relatively compact for the regular topology π. According to LeCam, a necessary and sufficient condition for $S(\subset \mathcal{D})$ to be relatively compact for π is that, for a given ε (>0) and $g \in E(\mathcal{P})$ with $g(x) \geqq 0$ for all x, and $\|g\|=1$, there exists $\alpha \in C_0(A)$ such that $0 \leqq \alpha(a) \leqq 1$ for all $a \in A$ and $g \circ \delta \circ \alpha > 1-\varepsilon$ for all $\delta \in S$ [3]. Now let $E_1, E_2, \cdots$ be compact subsets of A such that $E_i \subset E_{i+1}$ for all i and $\bigcup_{i=1}^{\infty} E_i = A$, and let $\alpha_i \in C_0(A)$, $i=1, 2, \cdots$ be such that $0 \leqq \alpha_i(a) \leqq 1$ and $\alpha_i(a)=1$ for all $a \in E_i$.

Suppose that $\{\delta : r(\theta_0, \delta) \leqq u\}$ is not relatively compact. Then, by the LeCam's criterion stated above, there exist a positive ε_0 and μ-integrable $g \in E(\mathcal{P})$ such that $g(x) \geqq 0$ for all x, $\|g\|=1$ and $\delta_n \in \{\delta : r(\theta_0, \delta) \leqq u\}$ $(n=1, 2, \cdots)$ satisfying

$$(2.2) \qquad g \circ \delta_n \circ \alpha_n < 1-\varepsilon_0 \quad \text{for all } n.$$

By (2.2) we have

$$(2.3) \qquad \int_{\mathcal{X}} \delta_n(E_n^c; x) g(x) d\mu(x) \geqq \varepsilon_0 \quad \text{for all } n.$$

[1] The term "equivalent" is used here in the sense of Halmos ([4] p. 126).

Put $d\nu = g d\mu$. From (2.3) and $\delta_n(E_n^c; x) \leqq 1$, we have

$$0 < \inf_n \nu(S_n) \quad (=\omega, \text{ say}), \tag{2.4}$$

where $S_n = \left\{x : \delta_n(E_n^c;\ x) \geqq \frac{\varepsilon_0}{2}\right\}$. By Assumption (2. c), ν is absolutely continuous with respect to p_{θ_0}. Hence, there exists $\lambda > 0$ such that $\nu(S) < \omega$ for every measurable set S for which $p_{\theta_0}(S) < \lambda$. So we have $\inf_n p_{\theta_0}(S_n) \geqq \lambda$ ([4] p. 125, Theorem B).

Hence we obtain

$$\begin{aligned} r(\theta_0, \delta_n) &\geqq \int_{S_n} \int_{E_n^c} w(\theta_0, a)\, \delta_n(da\,;\, x) f(x, \theta_0) d\mu(x) \\ &\geqq \int_{S_n} \int_{E_n^c} (\inf_{a \in E_n^c} w(\theta_0, a))\, \delta_n(da\,;\, x) f(x, \theta_0) d\mu(x) \\ &= \int_{S_n} (\inf_{a \in E_n^c} w(\theta_0, a))\, \delta_n(E_n^c\,;\, x) f(x, \theta_0) d\mu(x) \\ &\geqq \frac{\varepsilon_0}{2} \inf_{a \in E_n^c} w(\theta_0, a) p_{\theta_0}(S_n) \\ &\geqq \frac{\varepsilon_0}{2} \inf_{a \in E_n^c} w(\theta_0, a) \lambda. \end{aligned}$$

By Assumption (2. b) the right hand side tends to ∞ as n tends to ∞. Consequently we have

$$\lim_{n \to \infty} r(\theta_0, \delta_n) = \infty. \tag{2.5}$$

This contradicts $r(\theta_0, \delta_n) \leqq u$ for all n. The set $\{\delta : r(\theta_0, \delta) \leqq u\}$ is therefore relatively compact for π. Hence $\{\delta : r(\theta_0, \delta) \leqq u,\ \delta \in F\}$ is also relatively compact. Thus, by the lower semicontinuity of $r(\theta, \delta)$ as a function of δ, we have $\sup_{\delta \in Z} r(\theta, \delta) < \infty$ for all θ, where $Z = \{\delta : r(\theta_0, \delta) \leqq u,\ \delta \in F\}$. Hence the set $\{r(\cdot, \delta) : r(\theta_0, \delta) \leqq u,\ \delta \in F\}$ is relatively compact for τ since all its elements are contained in the product of the intervals $[0, \sup_{\delta \in Z} r(\theta, \delta)(<\infty)]$ $(\theta \in \Omega)$. Consequently $\overline{\{r(\cdot, \delta) : r(\theta_0, \delta) \leqq u,\ \delta \in F\}}$ is compact for τ. Hence by (2.1), $\{h : h(\theta_0) = \xi_{\theta_0}(h) \leqq u,\ h \in \overline{K}\}$ is also compact for τ.

To show that Assumption (1. c) holds, take an arbitrary $h \in \overline{K}$. If a subset H of U satisfies that, for every $g \in \overline{H}$, there exists a $g^* \in H$ such that $g^*(\theta) \leqq g(\theta)$ for all $\theta \in \Omega$, we shall say that H has the property (W) [3]. We shall show that K has the property (W). Let $V(h : \theta_1, \cdots, \theta_n, \varepsilon)$ be a neighborhood of h for the topology τ such that

$$V(h : \theta_1, \cdots, \theta_n, \varepsilon) = \{g : |h(\theta_i) - g(\theta_i)| < \varepsilon,\ i = 1, 2, \cdots, n\}.$$

For any finite subset $J=\{\theta_1, \theta_2, \cdots, \theta_n\}$ of Ω and positive integer m, we define $H(J, m)$ as follows:

$$H(J, m)=\left\{\delta : r(\cdot, \delta) \in V\left(h : \theta_0, \theta_1, \cdots, \theta_n, \frac{1}{m}\right)\right\}^a \cap F.$$

Since $h \in \overline{K}$, $H(J, m)$ is not empty and it is clear that $\bigcap_{i=1}^{n} H(J_i, m_i) \neq \phi$ for any $J_1, \cdots, J_k$ and $m_1, \cdots, m_k$. Since $H(J, m) \subset \{\delta : r(\theta_0, \delta) \leqq h(\theta_0)+1\}$, $H(J, m)$ is compact for π. Hence we have $\bigcap_{J,m} H(J, m) \neq \phi$. $\tilde{\delta} \in H(J, m)$ implies $r(\theta, \tilde{\delta}) \leqq h(\theta)$ for all $\theta \in \Omega$. This $\tilde{\delta}$ is clearly contained in F and therefore K has the property (W). Since δ_0 is admissible and $r(\cdot, \delta_0) \notin K$, there exists θ' such that $r(\theta', \delta_0) < r(\theta', \tilde{\delta})$. Thus we have $r(\theta', \delta_0) < h(\theta')$, that is to say, $\xi_{\theta'}(\delta_0) < \xi_{\theta'}(h)$. Assumption (1. c) is therefore satisfied. Hence, applying the preceding theorem we can prove that there exists $\xi \in \mathcal{M}$ such that $r(\xi, \delta_0) < \inf_{\delta \in F} r(\xi, \delta)$.

The proof of sufficiency is obvious. Let δ_0 be a decision function satisfying the statement of the theorem. For any δ $(r(\cdot, \delta_0) \neq r(\cdot, \delta))$, there exists $\xi \in \mathcal{M}$ such that $r(\xi, \delta_0) < r(\xi, \delta)$ since a set $\{\delta\}$ is convex. So δ_0 is admissible. This completes the proof of our theorem.

A theorem similar to Theorem 2 can be obtained when Ω is topologized. Let $\mathcal{M}^*$ be the totality of a priori distributions on Ω satisfying $\xi(O)>0$ for every open set O. $\mathcal{M}^*$ is clearly convex.

In addition to Assumptions (2. b) and (2. c), we assume

Assumption (3. a). The parameter space Ω is locally compact, σ-compact and metrizable,

Assumption (3. b). The set $\{r(\cdot, \delta) : \delta \in \mathcal{D}\}$ is equicontinuous as functions of θ, and instead of Assumption (2. a),

Assumption (3. c). We consider the class $\mathcal{D}^t$ of decision functions on Ω such that $\sup_\theta r(\theta, \delta) < \infty$.

Let U^* be the subset of U consisting of all continuous functions on Ω such that $\sup_\theta |f(\theta)| < \infty$. U^* is a topological vector subspace of (U, τ). Let τ^* be the relativization of τ to U^* and henceforth we shall denote by $\overline{H}$ the closure of H for the topology τ^*. By Assumption (3.c) we have $\overline{\{r(\cdot, \delta) : \delta \in \mathcal{D}^t\}} \subset U^*$.

If, for any $\xi \in \mathcal{M}^*$, we put $\xi(f)=\int_\Omega f(\theta)d\xi(\theta)$ for all $f \in U^*$, then ξ is additive on U^* and moreover we get the following lemma.

LEMMA 1. *$\xi(\in \mathcal{M}^*)$ is lower semicontinuous on* $\overline{\{r(\cdot, \delta) : \delta \in \mathcal{D}^t\}}$.

PROOF. We write $E=\{r(\cdot, \delta) : \delta \in \mathcal{D}^t\}$. Since $\xi(f_0) < \infty$ $(f_0 \in \overline{E})$, there exists a compact set $G(\subset \Omega)$ such that $\int_{G^c} f_0(\theta)d\xi(\theta) < \varepsilon$. As E is equicontinuous, so is also $\bar{E}$. Since the topology of pointwise convergence coincides with the topology of uniform convergence on compacta for an equicontinuous family [5], there exists a neighborhood $V(f_0)$ of f_0 for the topology τ^* satisfying

$$\int_G f_0(\theta)d\xi(\theta) - \int_G g(\theta)d\xi(\theta) < \varepsilon \quad \text{for all} \quad g \in V(f_0) \cap \bar{E}.$$

Hence we obtain

$$\begin{aligned}\xi(f_0)-\xi(g) &= \int_{G^c} f_0(\theta)d\xi(\theta) - \int_{G^c} g(\theta)d\xi(\theta) + \int_G f_0(\theta)d\xi(\theta) - \int_G g(\theta)d\xi(\theta) \\ &\leqq \int_{G^c} f_0(\theta)d\xi(\theta) + \varepsilon \leqq 2\varepsilon \qquad \text{for all} \quad g \in V(f_0) \cap \bar{E}.\end{aligned}$$

Since $g \in \bar{E}$ implies $g(\theta) \geqq 0$ for all θ, we have $\int_{G^c} g(\theta)d\xi(\theta) \geqq 0$ and so the second inequality follows. This shows that ξ ($\in \mathcal{M}^*$) is lower semicontinuous on $\bar{E}$.

THEOREM 3. *If we take $\mathcal{M}^*$ and $\mathcal{D}^t$ instead of $\mathcal{M}$ and $\mathcal{D}^b$, we have the same assertion as that of Theorem 2 under Assumptions (2. b), (2. c) and (3. a)~(3. c).*

PROOF. The proof of necessity is as follows. Let δ_0 be an admissible decision function and F be a closed convex set in $\mathcal{D}^t$ satisfying $r(\cdot, \delta_0) \notin \{r(\cdot, \delta) : \delta \in F\}$. Putting $Y=(U^*, \tau^*)$, $L=\mathcal{M}^*$, $C=\overline{\{r(\cdot, \delta) : \delta \in F\}}$, $s_0=\delta_0$ and $f_0=\xi_0$, where $\xi_0 \in \mathcal{M}^*$ is arbitrary, we shall show assumptions of Theorem 1 are satisfied.

It follows from Lemma 1 that Assumption (1. a) holds. As $\bar{E}=\overline{\{r(\cdot, \delta) : \delta \in \mathcal{D}^t\}}$ is the equicontinuous family and $\xi_0(O)>0$ for every open set O, there exists $\beta_{u,\theta}(<\infty)$, for each θ and $u(<\infty)$, such that $f(\theta) \leqq \beta_{u,\theta}$ for all $f \in \bar{E}$ for which $\xi_0(f) \leqq u$. Consequently the set $\{f : \xi_0(f) \leqq u\} \cap \bar{E}$ is compact for τ^*. Thus Assumption (1. b) is satisfied. As we showed in Theorem 2, the set $K=\{r(\cdot, \delta) : \delta \in F\}$ has the property (W). Hence, in the same way as Theorem 1, we can prove that Assumption (1. c) holds. Finally we remark that $\mathcal{M}^*$ is not empty. This is easily shown by virtue of Assumption (3. a).

The proof of sufficiency is completely the same as that of Theorem 2 and so is omitted.

WASEDA UNIVERSITY

REFERENCES

[1] C. Stein, "A necessary and sufficient condition for admissibility," *Ann. Math. Statist.*, 26 (1955), 518-522.
[2] N. Bourbaki, *Espaces Vectoriels Topologiques*, Paris, (1953), 70-74.
[3] L. LeCam, "An extension of Wald's theory of statistical decision functions," *Ann. Math. Statist.*, 26 (1955), 69-81.
[4] P. R. Halmos, *Measure Theory*, New York, 1950.
[5] J. L. Kelley, *General Topology*, New York, (1955), 232.