ON WALD'S COMPLETE CLASS THEOREMS[1]

By J. Kiefer

Cornell University

1. Summary. The purpose of this paper is to prove certain results concerning complete classes of strategies, some of which were announced in an abstract in *Bull. Am. Math. Soc.*, Vol. 57 (1951), p. 372.

2. Introduction. Except where explicitly stated to the contrary, we shall use the nomenclature and notation of Chapter 2 of [1] concerning zero-sum two-person games. Our considerations here do not require, however, that the payoff function $K(a, b)$ be bounded (or finite), but merely that it be bounded (by zero, without loss of generality) from below (because if unbounded in both directions, expectation relative to a mixed strategy might be undefined). This generalization is of use in some games and statistical work, as will be seen below. We remark without proof that such results as weakened forms of Theorems 1 and 4 of [2] may be proved under this set up. For example, we shall later use the following:

Theorem 1. *Suppose that* Ξ *and* H *are convex spaces of allowable strategies for players* 1 *and* 2, *respectively, that* $0 \leqq K(a, b) \leqq \infty$, *that* H *is weakly compact relative to* Ξ *in the sense of Wald* (i.e., *for any sequence* $\{\eta_i\}$ *in* H *there is a subsequence* $\{\eta_{i_j}\}$ *and an* η_0 *in* H *such that* $\liminf_{j\to\infty} K(\xi, \eta_{i_j}) \geqq K(\xi, \eta_0)$ *for all* ξ *in* Ξ), *and that there exists a sequence* $\{\xi_i\}$ *in* Ξ *such that for any* ξ *in* Ξ *and* η *in* H *there is a subsequence* $\{\xi_{i_j}\}$ (*all of whose elements may be the same*) *which may depend on* ξ *and* η *and is such that* $\lim_{j\to\infty} K(\xi_{i_j}, \eta) \geqq K(\xi, \eta)$. *Then the game is determined.*

The weak compactness assumption is enough to assure the existence of a minimax strategy for player 2. The above conditions may be weakened as in Theorem 4 of [2] or even further, and a generalization of Theorem 5 of [2] (which should be corrected there by assuming g_0 to be independent of ϵ) may similarly be proved.

3. Admissible strategies and complete classes. Wald considered two types of complete class theorems: those which give conditions under which the class of admissible strategies (e.g., of player 2) is complete, and those which give conditions under which the class of minimal strategies in the strict or wide sense is complete. The latter will occupy most of this paper. We remark, regarding the former, that the proof used by Wald in Theorem 2.22 of [1] actually suffices to prove the following:

Theorem 2. *Let* Ξ *and* H *be arbitrary spaces of mixed strategies with the property*

Received 3/4/52.

[1] Research sponsored by the Office of Naval Research.

that there exists a denumerable subset Ξ^* *of* Ξ *such that, if* η' *and* η'' *are any members of* H *for which* $K(\xi, \eta') \leqq K(\xi, \eta'')$ *for all* ξ *with strict inequality for some* ξ, *then there is a* ξ' *in* Ξ^* *with* $K(\xi', \eta') < K(\xi', \eta'')$. *Suppose also that* H *is weakly compact relative to* Ξ *in the sense of Wald. Then the class of all admissible strategies of player* 2 *is minimal complete.*

Note that $K(\xi, \eta)$ is not assumed bounded. An application of this theorem which indicates the usefulness of the hypothesis as stated herein over the stronger condition stated in Theorem 2.22 of [1], will be given in the next paragraph. We remark here that the condition of Theorem 2 is not necessary; for example, let A as well as B consist of all integers, Ξ and H consist of all probability measures on A and B, and $K(a, b) = 0$ if $a = b > 0$ and $K(a, b) = 2^{-|b|}$ otherwise; the class of all admissible strategies (those giving probability 1 to a single element $b > 0$) is then minimal complete, but H is not even weakly compact for every sequence of strategies for which each strategy is better than its predecessor (as is evidenced by the sequence of pure strategies $b = 0, -1, -2, \cdots$). On the other hand, the theorem does not remain valid if only weak compactness (but not the condition on Ξ^*) is assumed. For example, let A as well as B consist of all ordinals less than the first uncountable ordinal, let Ξ and H consist of all discrete probability measures on A and B, and let $K(a, b) = -1$ or 1 according to whether $a < b$ or $a \geqq b$, respectively. Then the condition of weak compactness is satisfied, but no strategy is admissible. (This example also illustrates why weak compactness alone is not enough to insure the determinateness of the game.) The above theorem may be generalized in an obvious manner by replacing the condition of weak compactness by a similar one on all well-ordered subsets of H whose power does not exceed that of some infinite Ξ^* with the stated property. (It is enough to consider only subsets of H whose members become "better" with increasing index.) It follows that the bicompactness condition used in Theorem 3 of [2], which implies such a condition for *every* subset of H, also implies the conclusion of Theorem 2 above.

As an important statistical application of Theorem 2, which also illustrates the advantage of using the condition on Ξ^* stated therein rather than that of the separability of Ξ^* in the sense of intrinsic metric (2.4) of [1], we shall now prove the following:

THEOREM 3. *Under Assumptions* 3.1 *to* 3.6 *of* [1], *the class of all admissible decision functions is minimal complete.*

This theorem extends the result of Theorem 2.22 of [1] to the setup of Chapter 3 of [1]. To prove it we let $\Xi^* = \bigcup_{i=1}^{\infty} \Xi_i$, where Ξ_i is a denumerable set of a priori distributions which is dense in Ξ in the sense of the metric $\rho_i(\xi_1, \xi_2) = \sup | r(\xi_1, \delta) - r(\xi_2, \delta) |$, the supremum being taken over all decision functions δ requiring at most i stages of experimentation. The existence of such Ξ_i follows from Theorems 3.3 and 2.16 of [1]. We shall show that Ξ^* satisfies the assumption of Theorem 2. Let δ_1 and δ_2 be two decision functions and ϵ a positive number such that $r(\xi) \geqq 0$ and $\sup_\xi r(\xi) > 2\epsilon$, where (using the notation of [1]) $r(\xi) = r(\xi, \delta_1) - r(\xi. \delta_2)$ (with the definition $\infty - \infty = 0$). We need only show that

$r(\xi') > 0$ for some ξ' in Ξ^*. Let $r_{1,m}(\xi) = \sup r(\xi, \delta^m)$, the supremum being over all δ^m requiring not more than m stages of observation and such that $r(\xi, \delta^m) \leqq r(\xi, \delta_1) + \epsilon$; if no such δ^m exists, we define $r_{1,m}(\xi) = 0$. Similarly, let $r_{2,m}(\xi) = \inf r(\xi, \delta^m)$, the infimum being over all δ^m requiring not more than m stages of observation and such that $r(\xi, \delta^m) \geqq r(\xi, \delta_2)$; if no such δ^m exist, we define $r_{2,m}(\xi) = +\infty$. Clearly, for each ξ, $r_{1,m}(\xi)$ is nondecreasing with m (we assume without loss of generality that the weight function W is nonnegative) and $r_{2,m}(\xi)$ is nonincreasing with m. Moreover, noting Lemma 3.3 of [1], and defining $r_m(\xi) = r_{1,m}(\xi) - r_{2,m}(\xi)$, we see that $r(\xi) \leqq \lim_{m\to\infty} r_m(\xi)$for every ξ for which $r(\xi, \delta_2)$ is finite. Moreover, $r_m(\xi) \leqq r(\xi) + \epsilon$ is nondecreasing in m, so that

$$\begin{aligned}\epsilon + \sup_{\xi\in\Xi^*} r(\xi) &\geqq \sup_{\xi\in\Xi^*} \lim_{m\to\infty} r_m(\xi) = \sup_{\xi\in\Xi^*} \sup_m r_m(\xi) \\ &= \sup_m \sup_{\xi\in\Xi^*} r_m(\xi) = \sup_m \sup_{\xi\in\Xi} r_m(\xi) = \sup_{\xi\in\Xi} \sup_m r_m(\xi) \\ &\geqq \sup_{\xi\in\Xi} r(\xi) > 2\epsilon,\end{aligned}$$

completing the proof. (It is essential here that $r_m(\xi)$ is increasing in m, so that the operations "lim" and "sup" may be interchanged.)

4. Minimal strategies and complete classes. We now turn to our main theorem, which generalizes Theorem 2.25 of [1]. The proof of the theorem is followed by two applications. The first of these is an essential strengthening of Theorems 3.17 and 3.20 of [1] regarding statistical decision functions. The second weakens the conditions of Theorem 2.25 of [1], even when $K(a, b)$ is bounded.

The idea of forming a new game with payoff function $K^*(a, b)$ is Wald's, and the proof of the first part of the conclusion of the theorem below is that of Theorem 2.25 of [1] if $K(a, b)$ is bounded. (The last part of the conclusion was proved under the stronger conditions that $K(a, b)$ is bounded and A and B are compact, so that minimality in the wide and strict senses are equivalent, in Theorem 3.10 of [4].) In the bounded case, any condition entailing the determinateness of the game and existence of a minimax strategy for player 2 and whose validity relative to K implies its validity relative to K^* (e.g., the condition of Theorem 2.25 of [1] or of Theorem 3 of [2]), also obviously entails the conclusion of the theorem below. When $K(a, b)$ is unbounded, one must be careful to use $K^*(\xi, \eta)$ only where $K(\xi, \eta)$ and $K(\xi, \eta_0)$ are not both infinite. Otherwise, K^* may not be properly defined. At the same time, it is useful to state the theorem in terms of the Ξ_N of the theorem rather than only in terms of A^*, since in many applications the Ξ_N may be chosen so that K^* is bounded from below on each Ξ_N (but not necessarily on A^*), so that in verifying condition (b) in applications one may use such results as that italicized in the first paragraph of this paper.

We recall (putting $\infty - \infty = 0$ in our case) that a strategy η' is minimal in the wide sense if

$$\inf_{\xi\in\Xi} [K(\xi, \eta') - \inf_{\eta\in H} K(\xi, \eta)] = 0. \tag{1}$$

THEOREM 4. *Suppose* $0 \leqq K(a, b) \leqq \infty$, $\Xi \supset A$, $\mathrm{H} \supset B$, *and that for any* η_0 *for which* $\inf_a K(a, \eta_0) < \infty$ *and which is not a member of the class* C_W *of all minimal strategies in the wide sense, there exists a sequence* $\{\Xi_i\}$ $(i = 1, 2, \cdots$, ad inf) *of subsets of* Ξ *such that*

(a) $\lim \inf_{N \to \infty} \Xi_N \supset A^* = \{a \mid K(a, \eta_0) < \infty\}$; *for every* N, $K(\xi, \eta_0) < \infty$ *for all* ξ *in* Ξ_N; *if* $\sup_a K(a, \eta_0) < \infty$, $\Xi_N \supset A$;

(b) *the game relative to* Ξ_N, H, *and* $K^*(\xi, \eta) = K(\xi, \eta) - K(\xi, \eta_0)$ *is determined and player* 2 *has a minimax strategy for this game.*

If $\sup_{a,b} K(a, b) = +\infty$, *suppose also that* H *is weakly compact relative to* A^* *for each* $\eta_0 \notin C_W$ *for which* $\inf_a K(a, \eta_0) < \sup_a K(a, \eta_0) = +\infty$. (*If* H *is weakly compact relative to* A, *this is automatically satisfied.*)

Then C_W *is complete. Moreover, for any* η_0 *not in* C_W *there is an* η_1 *in* C_W *and an* $\epsilon > 0$ *such that* $K(\xi, \eta_1) \leqq K(\xi, \eta_0) - \epsilon$ *for all* ξ *in* Ξ.

PROOF. We suppose $C_W \neq \mathrm{H}$, or the theorem is trivial; in particular, $\inf_{a,b} K(a, b) < \infty$, since otherwise $C_W = \mathrm{H}$. We now show that C_W is not empty. If there is an $\eta_0 \notin C_W$ with $\sup_a K(a, \eta_0) < R < \infty$, it follows from (b) that there is a minimax strategy η' relative to Ξ_N, H, and K^*. Since this game is determined and $\Xi \supset \Xi_N \supset A$ in this case, it is easy to verify that the game relative to Ξ, H, and K^* is determined, that η' is minimax for it, and hence that η' is minimal in the wide sense relative to Ξ, H, and K^* (since $0 \geqq K^*(a, \eta') \geqq -R$, the proof of Theorem 2.17 of [1] applies), and hence relative to Ξ, H, and K. On the other hand, if no such η_0 exists, the first sentence of the proof shows that there must exist an η_0 with non-empty A^* and (by the assumption following (b)) such that there exists a minimal strategy relative to any member of A^*. At any rate, C_W is not empty.

Let η_0 be any member of H which is not in C_W. If $K(a, \eta_0) = +\infty$ for all a, any η' in C_W (which is non-empty by the previous paragraph) is uniformly better than η_0 and is such that $K(\xi, \eta') \leqq K(\xi, \eta_0) - 1$ for all ξ. Hence, we may assume in what follows that $\inf_a K(a, \eta_0) < \infty$, and that the Ξ_N corresponding to this η_0 are non-empty for all N not less than some N_0. We now let $\Xi^* = \bigcup_{N=1}^{\infty} \Xi_N$, and define

$$\epsilon = \inf_{\xi \in \Xi^*} [K(\xi, \eta_0) - \inf_{\eta} K(\xi, \eta)]. \tag{2}$$

(It is clear that $\inf_\eta K(\xi, \eta) < \infty$ for all ξ. Otherwise, every η in H would be minimal and we would have $C_W = \mathrm{H}$.) Clearly, $\epsilon > 0$, or by (1) (with $\eta' = \eta_0$) η_0 would be minimal in the wide sense. Moreover, $\epsilon < \infty$, since Ξ_{N_0} is non-empty.

For any $N \geqq N_0$, let η_N be a minimax strategy for the game described in (b), so that

$$\sup_{\xi \in \Xi_N} \inf_{\eta} K^*(\xi, \eta) = \inf_{\eta} \sup_{\xi \in \Xi_N} K^*(\xi, \eta) = \sup_{\xi \in \Xi_N} K^*(\xi, \eta_N). \tag{3}$$

The common value of (3) is less than or equal to $-\epsilon$; for if, to the contrary, it were $-\epsilon + 2\rho$ for some $\rho > 0$, there would by (3) exist a ξ_0 in Ξ_N for which

$$-\epsilon + \rho \leqq \inf_{\eta} K^*(\xi_0, \eta) = \inf_{\eta} K(\xi_0, \eta) - K(\xi_0, \eta_0), \tag{4}$$

which would contradict (2). Hence, we must have

$$K(\xi, \eta_N) \leqq K(\xi, \eta_0) - \epsilon \qquad \text{for all } \xi \text{ in } \Xi_N . \tag{5}$$

Let the subsequence $\{N_j\}$ $(j = 1, 2, \cdots$, ad inf) of the positive integers and the strategy $\eta^* \epsilon \text{H}$ be (as guaranteed by weak compactness relative to A^* if $\sup_a K(a, \eta_0) = +\infty$, and putting $\eta^* = \eta_{N'}$ if $\sup_a K(a, \eta_0) < \infty$ and assuming without loss of generality that $\Xi_N = \Xi_{N'}$ for $N > N'$ in this case) such that

$$\liminf_{j\to\infty} K(a, \eta_{N_j}) \geqq K(a, \eta^*) \qquad \text{for all } a \text{ in } A^*. \tag{6}$$

It follows from (5) that

$$K(a, \eta^*) \leqq K(a, \eta_0) - \epsilon \qquad \text{for all } a \text{ in } A^*; \tag{7}$$

that is, $K(a, \eta^*) \leqq K(a, \eta_0) - \epsilon$ for all a for which $K(a, \eta_0) < \infty$. Since the latter set is nonempty, η^* is uniformly better than η_0, and in fact

$$K(\xi, \eta^*) \leqq K(\xi, \eta_0) - \epsilon \qquad \text{for all } \xi \text{ in } \Xi. \tag{8}$$

The minimality in the wide sense of η^* (i.e., the verification of (1) for $\eta' = \eta^*$) is a direct consequence of (8), the fact that Ξ^* is nonempty, and (2). This completes the proof of the theorem.

Application I. In the terminology of Chapter 3 of [1], let $\mathfrak{D}$ be the class of all decision functions and $\mathfrak{D}_b$ the class of all decision functions with bounded risk functions. Let C_s be the class of all Bayes solutions in the strict sense and C_W the class of all Bayes solutions in the wide sense. Wald showed that, under Assumptions 3.1 to 3.6 of [1], C_W is complete relative to $\mathfrak{D}_b$ (Theorem 3.17 of [1]), and that, under Assumptions 3.1 to 3.7 of [1], C_s is complete relative to $\mathfrak{D}_b$ (Theorem 3.20 of [1]). (These theorems were also proved by Wald under stronger conditions in [3], [4], and [5], and were stated under stronger conditions in [6]. In [3] and [4] (by Condition 7 of the latter) the risk function is always bounded. Theorems 2.6, 2.7, 3.5, and 3.6 of [5] are stated correctly, relative to $\mathfrak{D}_b$. The proofs of Theorems 2.5 and 3.4 of [5] are correct only if the statement of these theorems is interpreted relative to $\mathfrak{D}_b$; otherwise, the statement following equation (2.72) of [5] is false, since the W^* defined there need not satisfy Condition 2.2 of [5]). If, using Wald's notation and in particular putting δ_0 for η_0 and $r(F, \delta_0)$ for $K(a, \eta_0)$, one defines the Ξ_N of our theorem to consist of all ξ for which $\xi(A_N) = 1$, where $A_N = \{a \mid K(a, \eta_0) \leqq N\}$, it is easy to verify that A_N, the terminal decision space D^t, and the weight function $W^*(F, d) = W(F, d) - r(F, \delta_0)$ (when restricted to A_N) satisfy Assumptions 3.1 to 3.6 (and 3.7) of [1] whenever Ω, D^t, and $W(F, d)$ satisfy the corresponding assumptions. Hence, Theorems 3.4, 3.7, and 3.2 of [1] imply (putting $r^*(F, \delta) = r(F, \delta) - r(F, \delta_0)$ for our K^*) that condition (b) and the condition which follows it in our theorem are satisfied, so that the conclusion of Theorem 4 holds. Hence, we have proved the following:

Theorem 5. *In the statements of Theorem* 3.17 *and Theorem* 3.20 *of* [1], $\mathfrak{D}_b$ *may be replaced by* $\mathfrak{D}$.

The proof for Theorem 3.20 uses Theorem 3.15. The last part of the conclusion of the Theorem 4, when applied to the present case, yields a result not proved in [1] but proved under stronger conditions (e.g., all risk functions are bounded) in Theorem 4.11 of [4].

APPLICATION II. Suppose $0 \leqq K(a, b) \leqq \infty$, that Ξ and H are convex, that H is weakly compact relative to Ξ, and that there is a countable subset $\Xi^* = \{\xi_i\}$ of Ξ such that, given any ξ in Ξ, there is a subsequence $\{\xi_{i_j}\}$ of Ξ^* (whose elements are not necessarily different) such that $\lim_{j \to \infty} K(\xi_{i_j}, \eta) = K(\xi, \eta)$ for all η in H. We define $\Xi_N = \{\xi \mid K(\xi, \eta_0) < N; \xi \in \Xi\}$, and we note that only (b) need be verified to assure the applicability of Theorem 4. It is easy to verify that H is weakly compact relative to Ξ_N and the payoff function K^*. Moreover, for any ξ in Ξ_N there is by assumption a subsequence $\{\xi_{i_j}\}$ of Ξ^* with $\lim_{j \to \infty} K(\xi_{i_j}, \eta) = K(\xi, \eta)$ for all η in H. In particular, this holds for $\eta = \eta_0$, so that $K(\xi_{i_j}, \eta_0) < N$ for sufficiently large j. We conclude that $\Xi^* \cap \Xi_N$ satisfies relative to Ξ_N, H, K, and hence relative to Ξ_N, H, K^*, the same relationship that Ξ^* did to Ξ, H, K. From Theorem 1 stated in the first paragraph of this paper, we conclude that (b) is satisfied.

Even when $K(a, b)$ is bounded, the above condition of weak sequential separability is weaker than the strong separability condition used in Theorem 2.25 of [1].

REFERENCES

[1] A. WALD, *Statistical Decision Functions*, John Wiley and Sons, New York, 1950.

[2] S. KARLIN, "Operator treatment of minimax principle," *Ann. Mathematics Studies*, *no.* 24, Princeton University Press, (1950), pp. 133–154.

[3] A. WALD, "An essentially complete class of admissible decision functions," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 549–555.

[4] A. WALD, "Foundations of a general theory of sequential decision functions," *Econometrica*, Vol. 15 (1947), pp. 279–313.

[5] A. WALD, "Statistical decision functions," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 165–205.

[6] A. WALD, "Basic ideas of a general theory of statistical decision rules," *Proceedings of the International Congress of mathematicians, 1950*, American Mathematical Society, 1952, pp. 231–243.