ON CERTAIN CLASSES OF STATISTICAL DECISION PROCEDURES

By H. S. Konijn

University of California, Berkeley

Summary. The paper considers classes of decision procedures which in certain ways put a bound on the associated losses of incorrect terminal decision or cost of experimentation. Conditions are given under which these classes fulfill the conditions that Wald imposes on classes of decision procedures in his general theory.

1. Introduction. We investigate a problem arising in Wald's general theory of statistical decision functions (all references to Wald are to [1]), and freely use his notations.

Wald considers the general case, in which sampling may be done by stages, by means of a decision procedure δ fixed in advance. This procedure determines at each state $k \geqq 0$ of experimentation whether to continue experimentation or not, on the basis of the observations obtained thus far (and perhaps also with the aid of an additional randomization experiment). In case the procedure indicates that a $(k + 1)$st stage is to take place, it also determines, on the same basis, which subset d_{k+1}^e of the possible collection D^e of variables is to be observed next. In case the procedure leads to termination of experimentation, it will also designate a particular final decision d^t contained in the set D^t of preadmitted, possible terminal decisions. Thus, for any given procedure δ, the experimental and terminal decisions to be taken are random variables, which depend on the sample point $X = \{X_1, X_2, \cdots\}$ to be observed, and, in case of a randomized procedure, on the randomization experiments to be performed; we may denote them by $\delta^e(X) = \{\delta_1^e(X), \delta_2^e(X), \cdots\}$ and $\delta^t(X)$, respectively.

Let $W(F, d^t)$ denote the loss due to taking the terminal decision d^t when F is the true distribution, and $c(d_1^e, \cdots, d_k^e, x)$ the cost of observing stagewise the k sets of variables $d_1^e, \cdots, d_k^e$ when the observed sample point equals $x = \{x_1, x_2, \cdots\}$. Let

$$P\{F, y \mid \delta\} = \Pr\{W(F, \delta^t(X)) > y \mid F, \delta\}$$

and

$$Q\{F, z \mid \delta\} = \Pr\{c(\delta^e(X), X) > z \mid F, \delta\},$$

where F and δ after the verical bar indicate that the probabilities are to be computed under the assumption that F is the true distribution and δ the adopted decision procedure.

We adopt Wald's assumptions 3.1 to 3.5 which are as follows.

Received 5/29/52, revised 2/14/53.

(a) The class of stochastic processes F consists either of discrete processes only, or of absolutely continuous processes only.

(b) The loss W is a bounded function, uniformly continuous in its second argument, with the modulus of continuity independent of the first argument.

(c) The cost c cannot decrease when a stage of observation is added, is a nonnegative unbounded function of the number of observations (uniformly in the other variables), and for any k, $d_1^e, \cdots, d_k^e$, is either a bounded function of the values of the observations or identically equals ∞.

(d) The space D^t of possible terminal decisions is compact (in the uniform topology with respect to W).

It may be noted that in the greater part of Wald's book it is assumed that W is nonnegative, but that in a few places this assumption is violated; however, the assumption that W is bounded is not violated, and this allows one to make adjustments in all proofs which make use of the nonnegativeness of W. The expected loss from incorrect terminal decisions and the expected cost of experimentation equal

$$r_1(F, \delta) = \int_0^\infty P\{F, y \mid \delta\}\, dy$$

and

$$r_2(F, \delta) = \int_0^\infty Q\{F, z \mid \delta\}\, dz,$$

respectively.

Speaking generally, Wald's theory is concerned with the search for procedures among a given class $\mathfrak{D}$, which in some sense minimize the "risk" $r(F, \delta) = r_1(F, \delta) + r_2(F, \delta)$. It may in various cases be desirable not to admit to the competition those procedures of an otherwise naturally arising class for which (1) r_1, or r_2, is larger than a given number, (2) the chance that W, or c, exceeds a given amount is larger than some small number, (3) r is not a bounded function of F, or, (4) the number of stages of experimentation has a positive probability of being unbounded. If so, it may be convenient to know whether the assumptions on $\mathfrak{D}$ under which Wald derives his general results still hold for the restricted class if they hold for $\mathfrak{D}$. The present paper is addressed to that question.

2. Statement of results. For given numbers α, β, $y = y(\alpha)$, $z = z(\beta)$, y_0, and z_0, we define the following subclasses of any given class $\mathfrak{D}$ of decision procedures

$$\mathfrak{D}\{\alpha, y(\alpha); -\} = \{\delta \,\varepsilon\, \mathfrak{D}: P\{F, y(\alpha) \mid \delta\} \leqq \alpha \text{ for all } F\},$$

$$\mathfrak{D}\{-; \beta, z(\beta)\} = \{\delta \,\varepsilon\, \mathfrak{D}: Q\{F, z(\beta) \mid \delta\} \leqq \beta \text{ for all } F\},$$

$$\mathfrak{D}_{y_0} = \{\delta \,\varepsilon\, \mathfrak{D}: r_1(F, \delta) \leqq y_0 \text{ for all } F\},$$

$$\mathfrak{D}^{z_0} = \{\delta \,\varepsilon\, \mathfrak{D}: r_2(F, \delta) \leqq z_0 \text{ for all } F\}.$$

The principles of notation adopted here are these: preceding the semicolon are conditions in terms of $P\{F, y \mid \delta\}$, following the semicolon are conditions in terms of $Q\{F, z \mid \delta\}$; the subscripts denote restrictions on r_1, and the superscripts restrictions on the cost of experimentation not stated directly in terms of $Q\{F, z \mid \delta\}$.

We inquire whether, if Wald's general assumptions 3.6 on the class of decision procedures hold for $\mathfrak{D}$, they also hold for these subclasses. Using the notation given below, these assumptions can be stated as follows.

(i) $\mathfrak{D}$ is convex.

(ii) $\mathfrak{D}$ is closed (in the regular sense of convergence defined by Wald).

(iii) For any k, there exists c_k such that $p(d^e_{(k)} \mid x_{[k-1]}, \delta)$ vanishes whenever $d^e_{[k]}$ is not contained in $(1, \cdots, c_k)$.

(iv) $p(d^e_{(k)}, \mid x_{[k-1]}, \delta)$ vanishes whenever $c(d^e_{(k)}, x_{[k-1]}, x^k) = \infty$ identically in x^k.

(v) Given $k \geqq 0$ and δ in $\mathfrak{D}$, there exists δ^k in $\mathfrak{D}$ such that $\delta^k\{D^t \mid d^e_{(k)}, x_{[k]}\} = 1$, whatever be $d^e_{(k)}$ or $x_{[k]}$, while

$$\delta^k\{\bar{D} \mid d^e_{(h)}, x_{[h]}\} = \delta\{\bar{D} \mid d^e_{(h)}, x_{[h]}\},$$

whatever be $\bar{D} \subset D^e \cup D^t$, $h < k$, $d^e_{(h)}$, or $x_{[h]}$. (Wald's formulation of this assumption, which permits truncation of the decision procedure at any stage $k \geqq 0$, is unnecessarily strong.)

We also examine from this point of view

$$\mathfrak{D}^{(b)} = \{\delta \, \varepsilon \, \mathfrak{D}: r(F, \delta) \text{ is a bounded function of } F\},$$

(called $\mathfrak{D}_b$ in Wald, p. 100), and

$$\mathfrak{D}^{(c)} = \{\delta \, \varepsilon \, \mathfrak{D}: \text{for any given } F \text{ the number of stages of experimentation is bounded almost certainly}\}.$$

We find that if $\mathfrak{D}$ satisfies Wald's assumptions 3.6, then for every set of numbers y_0, z_0, α, β, $z = z(\beta)$, and all except an at most denumerable collection of choices of $y = y(\alpha)$, this assumption holds also for

$$\mathfrak{D}_{y_0}, \mathfrak{D}^{z_0}, \mathfrak{D}\{-; \beta, z(\beta)\}, \mathfrak{D}\{\alpha, y(\alpha); 0, z(0)\}, \mathfrak{D}^{(b)}, \mathfrak{D}^{(c)},$$
$$\mathfrak{D}^{(b)}\{\alpha, y(\alpha); -\}, \mathfrak{D}^{(c)}\{\alpha, y(\alpha); -\}, \mathfrak{D}^{z_0}\{\alpha, y(\alpha); -\},$$

and thus also for

$$\mathfrak{D}^{(b)}\{\alpha, y(\alpha); \beta, z(\beta)\}, \mathfrak{D}^{(c)}\{\alpha, y(\alpha); \beta, z(\beta)\}, \mathfrak{D}^{z_0}\{\alpha, y(\alpha); \beta, z(\beta)\}, \text{etc.}$$

Let $\mathfrak{D}^{(0)}$ denote the class of all possible decision procedures subject to (iii) only, and $\mathfrak{D}^{(1)}$ the subclass of $\mathfrak{D}^{(0)}$ which satisfies (iv). Wald remarks that $\mathfrak{D}^{(1)}$ satisfies all of his assumptions 3.6, so that all classes mentioned in the preceding paragraph satisfy them if $\mathfrak{D} = \mathfrak{D}^{(1)}$. Moreover, it is easy to see that all these classes except $\mathfrak{D}_{y_0}$, $\mathfrak{D}^{(c)}$, and $\mathfrak{D}^{(c)}\{\alpha, y(\alpha); -\}$ already satisfy them if $\mathfrak{D} = \mathfrak{D}^{(0)}$. Note also that property (v) is not needed for any of Wald's proofs when $\mathfrak{D}$ is contained in $\mathfrak{D}^{(0)(b)}$, $\mathfrak{D}^{(0)(c)}$ or $\mathfrak{D}^{(0)z_0}$.

3. Notation. We occasionally make Wald's notation a little more explicit or condensed, and also introduce some new notation. Represent D^e by $(1, 2, \cdots)$ and let

$$d^e_{(k)} = (d^e_1, \cdots, d^e_k),\ d^e_{[k]} = d^e_1 \cup \cdots \cup d^e_k .$$

Write $x_{[k]} = \{x_i(i \,\varepsilon\, d^e_{[k]})\}$, $M_{[k]}$ for the set of all possible values of $x_{[k]}$, $\bar{M}_{[k]}$ for a bounded measurable subset of $M_{[k]}$; write $x^k = \{x_i(i \,\varepsilon\, d^e_k)\}$, M^k for the set of all possible values of x^k.

When F is the true distribution, let

$$F(x^k \mid x_{[k-1]}) = \Pr\{X_i \leqq x_i(i \,\varepsilon\, d^e_k) \mid X_j = x_j(j \,\varepsilon\, d^e_{[k-1]})\},$$
$$F(x_{[k]}) = \Pr\{X_j \leqq x_j(j \,\varepsilon\, d^e_{[k]})\}.$$

We recall, in a somewhat more specific form, the following definitions of Wald's Sections 1.2 and 3.1.4.

$$p(\bar{D}^t \mid 0, \delta) = \delta(\bar{D}^t \mid 0),$$
$$p(d^e_{(k)} \mid x_{[k-1]}, \delta) = \delta(d^e_1 \mid 0)\delta(d^e_2 \mid d^e_1, x_{[1]}) \cdots \delta(d^e_k \mid d^e_{(k-1)}, x_{[k-1]}),$$
$$p(d^e_{(k)}, \bar{D}^t \mid x_{[k]}, \delta) = p(d^e_{(k)} \mid x_{[k-1]}, \delta)\delta(\bar{D}^t \mid d^e_{(k)}, x_{[k]}),$$
$$q(d^e_{(k)}, \bar{D}^t \mid F, \delta) = \int_{M_{[k]}} p(d^e_{(k)}, \bar{D}^t \mid x_{[k]}, \delta)\, dF(x_{[k]}).$$

Let

$$M^k(z \mid d^e_{(k)}, x_{[k-1]}) = \{x^k \,\varepsilon\, M^k : c(d^e_{(k)}, x_{[k-1]}, x^k) \leqq z \mid d^e_{(k)}, x_{[k-1]}\},$$
$$\bar{D}^t\{F, y\} = \{d^t \,\varepsilon\, D^t : W(F, d^t) > y \mid F\}.$$

Then

$$P\{F, y \mid \delta\} = \sum_{k=0}^{\infty} \sum_{d^e_k} \cdots \sum_{d^e_1} q(d^e_{(k)}, \bar{D}^t\{F, y\} \mid F, \delta),$$

$$1 - Q\{F, z \mid \delta\}$$
$$= \sum_{k=1}^{\infty} \sum_{d^e_k} \cdots \sum_{d^e_1} \int_{M_{[k-1]}} \{p(d^e_{(k)} \mid x_{[k-1]}, \delta) \int_{M^k(z \mid d^e_{(k)}, x_{[k-1]})} dF(x^k \mid x_{[k-1]}\}\, dF(x_{[k-1]}).$$

4. Convexity.

THEOREM 1. *Let $\mathfrak{D}$ be convex. Then $\mathfrak{D}\{\alpha, y(\alpha); -\}$ and $\mathfrak{D}\{-; \beta, z(\beta)\}$ are convex for every α, β, $y = y(\alpha)$, $z = z(\beta)$; $\mathfrak{D}_{y_0}$ and $\mathfrak{D}^{z_0}$ are convex for every y_0, z_0; $\mathfrak{D}^{(b)}$ and $\mathfrak{D}^{(c)}$ are convex.*

PROOF. Let δ_1 and δ_2 be elements of $\mathfrak{D}\{\alpha, y(\alpha); -\}$ and $0 < \theta < 1$, $\rho = 1 - \theta$. To show that there exists $\delta \,\varepsilon\, \mathfrak{D}\{\alpha, y(\alpha); -\}$ such that 3.14a, b (Wald) are satisfied, note that there exists such a $\delta \varepsilon \mathfrak{D}$ since $\mathfrak{D}$ is convex, and that this δ actually is an element of $\mathfrak{D}\{\alpha, y(\alpha); -\}$ as

$$P\{F, y(\alpha) \mid \delta\} = \sum_{k=0}^{\infty} \sum_{d_k^e} \cdots \sum_{d_1^e} \int_{M[k]} p(d_{(k)}^e, \bar{D}^t\{F, y(\alpha)\} \mid x_{[k]}, \theta\delta_1 + \rho\delta_2)\, dF(x_{[k]})$$

$$= \theta P\{F, y(\alpha) \mid \delta_1\} + \rho P\{F, y(\alpha) \mid \delta_2\} \leqq \alpha.$$

Quite similarly we show $Q\{F, z(\beta) \mid \delta\} \leqq \beta$.

Then

$$r_1(F, \delta) = \theta \int_0^\infty P\{F, y \mid \delta_1\}\, dy + \rho \int_0^\infty P\{F, y \mid \delta_2\}\, dy$$

$$= \theta r_1(F, \delta_1) + \rho r_1(F, \delta_2),$$

so that if δ_1 and δ_2 are elements of $\mathfrak{D}_{y_0}$ with δ satisfying 3.14a, b (Wald), $\delta \varepsilon \mathfrak{D}_{y_0}$; and similarly we obtain the convexity of $\mathfrak{D}^{z_0}$, and thus of $\mathfrak{D}^{(b)}$.

The convexity of $\mathfrak{D}^{(c)} = \{\delta \varepsilon \mathfrak{D}$: for each F there exists k' with

$$\sum_{k=0}^{k'} \sum_{d_k^e} \cdots \sum_{d_1^e} q(d_{(k)}^e, D^t \mid F, \delta) = 1\}$$

is immediate.

5. Closure.

THEOREM 2. *Let $\mathfrak{D}$ be closed, and a subset of $\mathfrak{D}^{(0)}$. Then $\mathfrak{D}_{y_0}$ and $\mathfrak{D}^{z_0}$ are closed for every y_0, z_0; and $\mathfrak{D}^{(b)}$ is closed.*

PROOF. By Wald's Theorem 3.2, which holds for any subset of $\mathfrak{D}^{(0)}$, we have, for $t = 1, 2$, that $\delta_i \rightarrow \delta_0$ (in Wald's regular sense) implies that there exists a subsequence $\{\delta_{i_j}\}$ such that

$$\liminf_{j=\infty} r_t(F, \delta_{i_j}) \geqq r_t(F, \delta_0);$$

thus, if for $i = 1, 2, \cdots, r_t(F, \delta_i) \leqq r_t^0 (= y_0$ for $t = 1, = z_0$ for $t = 2)$, then $r_t(F, \delta_0) \leqq r_t^0$ for $t = 1, 2$.

That $\mathfrak{D}^{(b)}$ is closed follows similarly: $r(F, \delta_{i_j}) \leqq M_{i_j}$ (say) $< \infty$; therefore, for given $\epsilon > 0$ there is a j such that $r(F, \delta_0) \leqq r(F, \delta_{i_j}) + \epsilon \leqq M_{i_j} + \epsilon < \infty$.

THEOREM 3. *Let $\{\delta_i\}$ be a sequence of decision procedures converging to a decision procedure δ_0 in the regular sense as defined by Wald. Then for all z and F, $\lim_{i=\infty} Q\{F, z \mid \delta_i\} = Q\{F, z \mid \delta_0\}$.*

PROOF. By Wald's assumption 3.5, since

$$\int_{M^k(z \mid d_{(k)}^e, x[k-1])} dF(x^k \mid x_{[k-1]})$$

vanishes when $d_{[k]}^e$ contains a sufficiently large number of elements, there exists for any F and z an integer k' such that the terms in the expression for

$1 - Q\{F, z \mid \delta\}$ with $k > k'$ do not make any contribution. We shall therefore discuss a finite sum of terms, write

$$1 - Q\{F, z \mid \delta\} = \sum_{k=1}^{k'} \sum_{d_k^e} \cdots \sum_{d_1^e} T\{d_1^e, \cdots, d_k^e ; F, z \mid \delta\},$$

and show that if $\delta_i \to \delta_0$, $\lim_{i=\infty} T\{d_{(k)}^e ; F, z \mid \delta_i\} = T\{d_{(k)}^e ; F, z \mid \delta_0\}$. Let f be the elementary law corresponding to F, and fix F, z, k, $d_1^e, \cdots, d_k^e$, setting for simplicity in notation $d_{[k-1]}^e = (1, \cdots, r)$.

I) Case where F is discrete. The case where F has a finite number of jumps is immediate, so suppose it has an infinite number of jumps. Let $T\{d_{(k)}^e ; F, z \mid \delta\} = \lim_{n=\infty} s_n(\delta)$ with

$$s_n(\delta) = \sum_{t_r=1}^{n_r(n)}, \cdots, \sum_{t_1=1}^{n_1(n)} \{p(d_{(k)}^e \mid x_{1t_1}, \cdots, x_{rt_r}, \delta) \cdot \sigma_{f,z}(x_{1t_1}, \cdots, x_{rt_r}) f(x_{1t_1}, \cdots, x_{rt_r})\}.$$

Here

$$\sigma_{f,z}(x_{1t_1}, \cdots, x_{rt_r}) = \int_{M^k(z \mid d_{(k)}^e, x_{[k-1]})} f(x^k \mid x_{[k-1]})\, dx^k \leqq 1,$$

and for any $j = 1, \cdots, r$, $n_j(n)$ is a nondecreasing integer-valued function of n.

Since $p(d_{(k)}^e \mid x_{1t_1}, \cdots, x_{rt_r}, \delta_i) \leqq 1$ and

$$\lim_{n=\infty} \sum_{t_r=1}^{n_r(n)} \cdots \sum_{t_1=1}^{n_1(n)} f(x_{1t_1}, \cdots, x_{rt_r}) = 1,$$

$\lim_{n=\infty} s_n(\delta_i)$ exists and is finite for each i. Now from the definition of regular convergence, $\lim_{i=\infty} p(d_{(k)}^e \mid x_{1t_1}, \cdots, x_{rt_r}, \delta_i) = p(d_{(k)}^e \mid x_{1t_1}, \cdots, x_{rt_r}, \delta_0)$, so by Weierstrass' rule

$$T\{d_{(k)}^e; F, z \mid \delta_0\} = \lim_{n=\infty} s_n(\delta_0) = \lim_{i=\infty} \lim_{n=\infty} s_n(\delta_i) = \lim_{i=\infty} T\{d_{(k)}^e ; F, z \mid \delta_i\}.$$

II) Case where F is absolutely continuous.

$$\varphi_z(x_{[k-1]} \mid d_{(k)}^e) = f(x_{[k-1]}) \int_{M^k(z \mid d_{(k)}^e, x_{[k-1]})} f(x^k \mid x_{[k-1]})\, dx^k,$$

which is not greater than $f(x_{[k-1]})$, satisfies the conditions of Wald's Lemma 3.1, so (using the notation of Wald's sections 3.9 and following)

$$T\{d_{(k)}^e; F, z \mid \delta_i\} = \int_{M_{[k-1]}} p(d_{(k)}^e \mid x_{[k-1]}, \delta_i) \varphi_z(x_{[k-1]} \mid d_{(k)}^e)\, dx_{[k-1]}$$

$$= \int_{M_{[k-1]}} \varphi_z(x_{[k-1]} \mid d_{(k)}^e)\, dP(d_{(k)}^e \mid \bar{M}_{[k-1]}, \delta_i).$$

But by the definition of regular convergence and Lemma 3.1 of Wald, the limit as $i \to \infty$ of the last expression is

$$\int_{M[k-1]} \varphi_z(x_{[k-1]} \mid d^e_{(k)})\, dP(d^e_{(k)} \mid \bar{M}_{[k-1]}, \delta_0),$$

which equals $T\{d^e_{(k)} ; F, z \mid \delta_0\}$.

COROLLARY. *If $\mathfrak{D}$ is closed, $\mathfrak{D}\{-; \beta, z(\beta)\}$ is closed for every β and $z = z(\beta)$.*

THEOREM 4. *Let $\{\delta_i\}$ be a sequence of decision procedures converging to a decision procedure $\delta_0 \varepsilon \mathfrak{D}^{(c)}$ in the regular sense as defined by Wald. Then, for any F, the chance variables $W(F, \delta^t_i(X))$ converge to $W(F, \delta^t_0(X))$ in distribution as $i \to \infty$, so that* $\lim_{i=\infty} P\{F, y \mid \delta_i\} = P\{F, y \mid \delta_0\}$ *for all y except an at most denumerable collection (depending on δ_0).*

PROOF. Fix F, k, $d^e_1, \cdots, d^e_k$, and show that

$$\lim_{i=\infty} q(d^e_{(k)}, \bar{D}^t\{F, y\} \mid F, \delta_i) = q(d^e_{(k)}, \bar{D}^t\{F, y\} \mid F, \delta_0)$$

for all y except an at most denumerable collection.

I) Case where F is discrete. Since $W(F, d^t)$ is continuous in d^t, $\bar{D}^t\{F, y\}$ is open for every y. If $\bar{D}^t\{F, y\} = \bar{D}^t\{F, y'\}$ and $y' < y$, $\bar{D}^t\{F, y\}$ has an empty boundary. Now consider the ordered collection $\{\bar{D}^t\{F, y\}\}$ of shrinking sets for increasing y. Only at an at most denumerable collection of y can the boundary $B\{F, y\}$ of $\bar{D}^t\{F, y\}$ be nonempty with $q(d^e_{(k)}, B\{F, y\} \mid F, \delta_0)$ positive; on the other y, $\lim_{i=\infty} P\{F, y \mid \delta_i\} = P\{F, y \mid \delta_0\}$, by the definition of regular convergence.

II) Case where F is absolutely continuous. For each integer m, consider the finite sequence $\{\bar{D}^t_{k_1 \cdots k_m}\}$ $(k_j = 1, \cdots, r_j ; j = 1, \cdots, m)$ of Wald's Section 3.1.4. As

$$\begin{aligned}\operatorname{Sup} \sum_{\bar{D}^t_{k_1 \cdots k_m} \subset \bar{D}^t\{F,y\}} & P(d^e_{(k)}, \bar{D}^t_{k_1 \cdots k_m} \mid \bar{M}_{[k]}, \delta_i) \\ & \leqq P(d^e_{(k)}, \bar{D}^t\{F, y\} \mid \bar{M}_{[k]}, \delta_i) \\ & \leqq \operatorname{Inf} \sum_{\bar{D}^t_{k_1 \cdots k_m} \supset \bar{D}^t\{F,y\}} P(d^e_{(k)}, \bar{D}^t_{k_1 \cdots k_m} \mid \bar{M}_{[k]}, \delta_i),\end{aligned}$$

we have by the definition of convergence in the regular sense

$$\begin{aligned}&\operatorname{Sup} \sum_{\bar{D}^t_{k_1 \cdots k_m} \subset \bar{D}^t\{F,y\}} P(d^e_{(k)}, \bar{D}^t_{k_1 \cdots k_m} \mid \bar{M}_{[k]}, \delta_0) \\ &\leqq \liminf_{i=\infty} P(d^e_{(k)}, \bar{D}^t\{F, y\} \mid \bar{M}_{[k]}, \delta_i) \leqq \limsup_{i=\infty} P(d^e_{(k)}, \bar{D}^t\{F, y\} \mid \bar{M}_{[k]}, \delta_i) \\ &\leqq \operatorname{Inf} \sum_{\bar{D}^t_{k_1 \cdots k_m} \supset \bar{D}^t\{F,y\}} P(d^e_{(k)}, \bar{D}^t_{k_1 \cdots k_m} \mid \bar{M}_{[k]}, \delta_0).\end{aligned}$$

For all except an at most denumerable set of y-intervals (which may degenerate to points) the left- and right-hand sides can be made to differ by less than any preassigned positive number, by making m large enough. But if, for $y' < y$,

$\bar{D}^t\{F, y\} = \bar{D}^t\{F, y'\}$, the left- and right-hand sides have the same value provided we take m so large that the diameter of each set in $\{\bar{D}^t_{k_1 \cdots k_m}\}$ is less than $y - y'$, so that for all $\bar{M}_{[k]}$ we have

$$\lim_{i=\infty} P(d^e_{(k)}, \bar{D}^t\{F, y\} \mid \bar{M}_{[k]}, \delta_i) = P(d^e_{(k)}, \bar{D}_t\{F, y\} \mid \bar{M}_{[k]}, \delta_0)$$

for all except an at most denumerable set of y.

We now note that Wald's Lemma 3.1 remains valid if we write T^y_i for $T_i (i = 0, 1, \cdots)$, insert "except for an at most denumerable set D_S of y" after (3.50) and "except for an at most denumerable set D of y" after (3.51); this is seen by letting ϵ and $1/c$ approach 0 through a denumerable set of values and considering in (3.58) the complement of $D^c_\epsilon = \bigcup_j (D^c_{S_j})$, with $D = \bigcup_c \bigcup_\epsilon (D^c_\epsilon)$. Consequently, as, for $i = 0, 1, \cdots$,

$$q(d^e_{(k)}, \bar{D}^t\{F, y\} \mid F, \delta_i) = \int_{M_{[k]}} f(x_{[k]}) dP(d^e_{(k)}, \bar{D}^t\{F, y\} \mid \bar{M}_{[k]}, \delta_i),$$

where f is the density corresponding to F, we have for all except an at most denumerable set of y

$$\lim_{i=\infty} q(d^e_{(k)}, \bar{D}^t\{F, y\} \mid F, \delta_i) = q(d^e_{(k)}, \bar{D}^t\{F, y\} \mid F, \delta_0).$$

COROLLARY. *If $\mathfrak{D}$ is closed, $\mathfrak{D}^{(c)}$ is closed, and*

$$\mathfrak{D}^{(c)}\{\alpha, y(\alpha); -\}, \mathfrak{D}\{\alpha, y(\alpha); 0, z(0)\}, \text{ and } \mathfrak{D}^{z_0}\{\alpha, y(\alpha); -\}$$

are closed for every choice of $z = z(0)$, z_0, and α, and all choices of $y = y(\alpha)$ except an at most denumerable collection.

PROOF. As above we prove that

$$\lim_{i=\infty} q(d^e_{(k)}, D^t \mid F, \delta_i) = q(d^e_{(k)}, D^t \mid F, \delta_0)$$

and obtain the closure of $\mathfrak{D}^{(c)}$ from the remark at the end of the proof of Theorem 1.

For any $z = z(0)$, if $\delta_0 \varepsilon \mathfrak{D}\{-; 0, z(0)\}$, then the number of stages of experimentation is bounded almost certainly under the procedure δ_0; it follows from Theorem 4 that for all α and all choices of $y = y(\alpha)$ except an at most denumerable collection $\mathfrak{D}^{(c)}\{\alpha, y(\alpha); -\}$ as well as $\mathfrak{D}\{\alpha, y(\alpha); 0\ z(0)\}$ are closed.

Let $P(F, y, k' \mid \delta)$ be the probability (under F, δ) that $W(F, \delta^t) > y$ and the number k of stages of experimentation $\leqq k'$; $P(F, y \mid k', \delta)$ the probability that $W(F, \delta^t) > y$, given that $k \leqq k'$; $P_{k'}(F, y \mid \delta)$ the probability that $W(F, \delta^t) > y$ and $k > k'$; and $R(F, k' \mid \delta)$ the probability that $k > k'$. For $i = 0, 1, \cdots$,

$$P\{F, y \mid \delta_i\} = P(F, y, k' \mid \delta_i) + P_{k'}(F, y \mid \delta_i) \leqq P(F, y, k' \mid \delta_i) + R(F, k' \mid \delta_i),$$

and for all y except an at most denumerable set $P(F, y \mid k', \delta_i)$ and

$$1 - R(F, k' \mid \delta_i) = \sum_{k=0}^{k'} \sum_{d^e_k} \cdots \sum_{d^e_1} q(d^e_{(k)}, D^t \mid F, \delta_i)$$

converge to $P(F, y \mid k', \delta_0)$ and $1 - R(F, k' \mid \delta_0)$ respectively when $\delta_i \rightarrow \delta_0$, so that $\lim_{i=\infty} P(F, y, k' \mid \delta_i) = P(F, y, k' \mid \delta_0)$ for this set of y. Therefore we have for this set of y

$$P(F, y, k' \mid \delta_0) \leqq \liminf_{i=\infty} P\{F, y \mid \delta_i\}$$
$$\leqq \limsup_{i=\infty} P\{F, y \mid \delta_i\} \leqq P(F, y, k' \mid \delta_0) + R(F, k' \mid \delta_0)$$

with $\lim_{k'=\infty} R(F, k' \mid \delta_0) = 0$, for, since $\delta_0 \in \mathfrak{D}^{(b)}$, (3.40) of Wald's holds. Consequently, for all y except at an most denumerable set, $\lim_{i=\infty} P\{F, y \mid \delta_i\} = P\{F, y \mid \delta_0\}$ for $\delta_0 \in \mathfrak{D}^{(b)}$. Since $\mathfrak{D}^{z_0}$ is a subset of $\mathfrak{D}^{(b)}$, and $\mathfrak{D}^{z_0}$ and $\mathfrak{D}^{(b)}$ are closed, this gives the closure of both $\mathfrak{D}^{(b)}\{\alpha, y(\alpha); -\}$ and $\mathfrak{D}^{z_0}\{\alpha, y(\alpha); -\}$ for all z_0, α, and all $y = y(\alpha)$ except perhaps a denumerable set.

REFERENCE

[1] A. Wald, *Statistical Decision Functions*, John Wiley and Sons, 1950.