THE INSTITUTE

OF STATISTICS

THE CONSOLIDATED UNIVERSITY
OF NORTH CAROLINA

ON WEAK ADMISSIBILITY OF TESTS

by

Bennett Eisenberg
Lehigh University

and

Gordon Simons
UNC - Chapel Hill

DEPARTMENT OF STATISTICS
Chapel Hill, North Carolina

On Weak Admissibility of Tests

[Short title: Weakly Admissible Tests]

ABSTRACT

This paper lays the foundation for a general theory of weakly admissible tests, where a test consists of both an experiment (a σ field of information) and a decision function based on the experiment. A test is weakly admissible unless there is another test with as good a decision function, which uses no more information, and is strictly better in one of these respects. Weakly admissible tests are completely characterized. In the case where the experiment consists of a sequential sampling procedure, a method for constructing such tests is given.

ON WEAK ADMISSIBILITY OF TESTS

by

Bennett Eisenberg and Gordon Simons[†]

1. <u>Introduction</u>. A sequential test is weakly admissible unless there is
another test which has no greater error probabilities, never requires more
observations and is strictly better in one of these respects. Clearly, weak
admissibility is the minimal requirement for a good test. Eisenberg, Ghosh
and Simons (1976) show, for arbitrary underlying distributions on the data,
that a sequential probability ratio test is always weakly admissible. This
follows, of course, from the Wald-Wolfowitz optimality theorem when it is
applicable. On the other hand, Neyman-Pearson tests are not always weakly
admissible. This paper characterizes weak admissibility. Roughly speaking,
a test is weakly admissible if it is a sensible likelihood ratio test which
takes no unnecessary observations.

The theory of weakly admissible tests can be developed at a useful level
of abstraction that is somewhat removed from the trappings of sequential analysis.
Rather than discussing tests (N,D) where N is a stopping time and D is a
terminal decision (measurable over the σ-field E_N of "events up to time N"),
we shall discuss tests (E,D) where E is an "experiment" (a σ-field which
represents the information available to a decision maker) and D is a "decision
based on the experiment" (an E-measurable mapping from the basic space Ω into
the action space).[1] The decision maker must choose an experiment within a class

---

[1] A similar use of the term "experiment" has been made by Rényi (1970, page 3).

E of permissible experiments and then choose an appropriate decision for the experiment chosen. In the case of general sequential tests, E would be the class of σ-fields E_N, where N is a stopping time.

Properties of decisions are discussed in Section 2. The notion of level k equivalence of decisions is introduced here. This concept plays an important role in the general theory of hypothesis testing. In particular, it appears in a general version of the Neyman-Pearson lemma, appearing in the section. Many results concerned with the concept are proved throughout the paper.

A formal definition of weakly admissible tests is given in Section 3. Theorem 5 says that, under general conditions, such tests may be characterized as "curtailed," "proper" likelihood ratio tests. Theorems 6, 7 and 8 (appearing in Section 4) show, under mild assumptions, that a proper likelihood ratio test can always "be curtailed," and the resulting "curtailed version" is a proper likelihood ratio test as well.

Section 5 applies our general theory to several contexts arising in the area of sequential analysis with the bulk of the attention given to tests which permit general (unrestricted) sequential sampling. A method for constructing and identifying the curtailed version of a likelihood ratio test is described.

Generalizations of our theory beyond the context of two competing simple hypotheses could be considered but are not in this paper.

2. <u>On making a good decision for a given experiment</u>. Probability measures P and Q defined on a measurable space (Ω, F) represent two states of nature. The problem is to decide whether the "null state" P or the "alternative state" Q is the true state of nature. The decision maker must choose an experiment E from a family E of *possible experiments*, and on the basis of the information

obtained from this experiment, he must make a decision D in favor of P or Q. Formally, an *experiment* is a sub-σ-field of F and a *decision*, for a given experiment E, is an E-measurable mapping from the basic space Ω into the set $\{P,Q\}$. In accordance with conventional terminology, $\alpha = P(D=Q)$ denotes the *size* of D and $1-\beta$, where $\beta = Q(D=P)$, denotes the *power* of D. Also, α and β are called *error probabilities*. Occasionally, they will be written as a vector (α,β). Each pair (E,D), $E\epsilon E$, is called a *test*. Throughout the paper, R denotes the probability measure $\frac{1}{2}(P+Q)$, which clearly dominates both P and Q.

Associated with each experiment E, there is an extended nonnegative random variable λ, determined up to an R equivalence, which is E-measurable and satisfies the equations

$$(1) \qquad Q(E, \lambda \neq \infty) = \int_E \lambda \, dP, \quad E\epsilon E \ .$$

λ is called the *likelihood ratio* for E.[2] It always exists and satisfies the reverse relationships

$$P(E, \lambda \neq 0) = \int_E \lambda^{-1} dQ, \quad E\epsilon E \ ,$$

where $\lambda^{-1} = 0$ when $\lambda = \infty$.

For the remainder of this section, E will be a fixed experiment, λ will be its likelihood ratio and all decisions will be decisions for E.

Two decisions, D and D', will be called *equivalent* if

$$(2) \qquad\qquad R(D=D') = 1 \ ,$$

and will be called *level k equivalent (for E)*, $k\epsilon[0,\infty]$, if

$$(3) \qquad\qquad R(D \neq D', \lambda \neq k) = 0$$

---

[2] The reader is referred to Eisenberg, Ghosh and Simons (1976) for details about this notion beyond those described in the present paper. Essentially, λ is the Radon-Nikodym derivative dQ/dP relative to E.

and

(4) $R(D=P) = R(D'=P)$.

We will denote level k equivalence by $D \overset{k}{=} D'$ when the experiment E is understood from the context. When there is a possibility of confusion, we will write $D \overset{k}{=} D'$ for E.

It is easily checked that equivalent and level k equivalent decisions have the same error probabilities, and that both notions possess the reflexive, symmetric and transitive properties expected of an equivalence.

A decision will be called a (*level* k) *likelihood ratio decision* (*for* E) if it is equivalent to a decision which equals P when $\lambda < k$ and equals Q when $\lambda > k$. (Its value when $\lambda = k$ is unspecified in the definition.) If $D \overset{k}{=} D'$ and D is a level k likelihood ratio decision, then so is D'.

We are now in a position to state a quite general version of the:

Neyman-Pearson Lemma. *If there is a level* k *likelihood ratio decision* D *which is a most powerful decision of size* α, *then every other most powerful decision of size* α *is level* k *equivalent to* D. *In order that there exist such a decision* D *for some* $k \in [0, \infty]$, *it is sufficient that* P *has no atoms in* E.

An assumption that P has no atoms in E is quite mild and is more convenient in the present context than the usual (and equally mild) assumption that randomized decisions are permissible. In point of fact, both assumptions nearly amount to the same thing. Two observations should clarify the situation: (i) Suppose that U is a random variable which, under P and Q, is uniformly distributed on [0,1] and is independent of E. Then the information represented by the σ-field $\sigma(U)$ can be used, together with the information in E, to make "randomized decisions" relative to E. But the likelihood ratio λ' for the experiment $E' = E \vee \sigma(U)$ (the smallest σ-field containing E and $\sigma(U)$)

will be identical to λ. It follows that a likelihood ratio decision with randomization would just be a likelihood ratio decision relative to E'. Hence, there is little reason not to replace E by E' in the first place; the class of possible decisions is increased, but no unfair statistical advantage is conferred. (ii) If the σ-field E is large enough to support a uniformly distributed random variable under P, then P is non-atomic on E (and conversely).

It cannot be inferred from the Neyman-Pearson Lemma that every likelihood ratio decision is a most powerful decision of its size. This is not quite true. Neither would it be correct to conclude that every likelihood ratio decision is a least size decision for its power. The problem is that a level infinity likelihood ratio decision can foolishly choose P when $\lambda = \infty$, and a level zero likelihood ratio decision can foolishly choose Q when $\lambda = 0$. A decision D will be called *proper* (*for E*) if

$$R(D=Q, \ \lambda=0) \ = \ R(D=P, \ \lambda=\infty) \ = \ 0 \ .$$

Since $P(\lambda=\infty) = Q(\lambda=0) = 0$, it is easy to show that *a given error probability of a proper likelihood ratio decision can be improved upon (reduced) by another decision only at the expense of the other error probability*. In particular, every proper likelihood ratio decision is a most powerful decision of its size. Consequently, the Neyman-Pearson Lemma yields the following theorem.

Theorem 1. *If* D *is a proper likelihood ratio decision of level* k *and* D' *is a decision with no larger error probabilities, then* D' $\overset{k}{\equiv}$ D.

The proofs of the Neyman-Pearson Lemma and Theorem 1, which rely on equation (1), vary only slightly from standard textbook arguments and, hence will be omitted.

It should be emphasized that the notions of "level k equivalence," "likelihood ratio decision" and "proper decision" have meaning only within the context of a specific experiment. For instance,

$$D \overset{k}{=} D' \text{ for } E \not\Rightarrow D \overset{k}{=} D' \text{ for } E' \quad ,$$

even if the decisions D and D' are decisions for E' (i.e., are E'-measurable).

Example 1. Let X_1 and X_2 be independent Bernoulli variables under P and Q with $P(X_1=1) = Q(X_1=1) = 1/2$, and $P(X_2=1) = Q(X_2=0) = 1/3$. Further, let $D = Q$ iff $X_1 = 0$ and $D' = Q$ iff $X_1 = 1$. Then $D \overset{1}{=} D'$ for $E = \sigma(X_1)$ but $D \overset{1}{\neq} D'$ for $E' = \sigma(X_1, X_2)$. $\square$

In contrast to this example, there are several results which can be very roughly interpreted as saying that *a notion holds with respect to E' if it holds with respect to E, and $E' \subseteq E$.* The next three theorems give three examples where this rough statement can be made precise.

Theorem 2. *Suppose $D \overset{k}{=} D'$ for E and $E' \subseteq E$. Then $D \overset{k}{=} D'$ for E' providing only that D and D' are E'-measurable (making them decisions for E').*

Theorem 3. *Suppose D is a level k likelihood ratio decision for E. If D' is a decision for $E' \subseteq E$ and $D' \overset{k}{=} D$ for E, then D' is a level k likelihood ratio decision for E'.*

Theorem 4. *Suppose D is a proper decision for E. If D' is a decision for $E' \subseteq E$ and $D' \overset{k}{=} D$ for E and some $k \in [0,\infty]$, then D' is a proper decision for E'.*

The proofs of Theorems 2 and 3 depend on the following:

Lemma. *Let E and E' be two experiments with likelihood ratios λ and λ', respectively, and $k \in [0,\infty]$. If $E' \subseteq E$, then $R(E) = 0$ for every $E \in E'$ for which*

$E \subseteq [\lambda \leq k < \lambda'] \cup [\lambda \geq k > \lambda']$.

Remark 1. The statement of this lemma is made awkward by the fact that the

event $E^* = [\lambda \leq k < \lambda'] \cup [\lambda \geq k > \lambda']$ may not be in E' and $R(E^*) > 0$ is

possible. This is apparent from Example 1 if the roles of E and E' are reversed.

Proof: Without loss of generality, it can be assumed that $E \subseteq [\lambda \leq k < \lambda']$ or

$E \subseteq [\lambda \geq k > \lambda']$. (If it is not, consider $E_1 = E[\lambda' > k]$ and $E_2 = E[\lambda' < k]$.)

We shall only consider the case $E \subseteq [\lambda \leq k < \lambda']$ since the other case can be

handled similarly. If $P(E) > 0$, then $k < \infty$ and (using (1))

$$0 < \int_E (\lambda'-k)\,dP \leq \int_E (\lambda'-\lambda)\,dP = Q(E,\lambda'\neq\infty) - Q(E,\lambda\neq\infty) \leq Q(E) - Q(E) = 0,$$

a contradiction. Thus $P(E) = 0$. In turn, if $R(E) > 0$, then $k < \infty$ and

(using (1))

$$0 < Q(E) = Q(E,\lambda<\infty) = \int_E \lambda\,dP = 0,$$

a contradiction. Thus $R(E) = 0$. □

Proof of Theorem 2: By assumption (3) and (4) hold. Since (3) holds, it can

be assumed, without loss of generality that $\lambda = k$ when $D \neq D'$. Then it

follows directly from the lemma that $R(D\neq D',\lambda'\neq k) = 0$, which, together with

(4), implies $D \overset{k}{=} D'$ for E'. □

Proof of Theorem 3: Without loss of generality (cf., the proof of Theorem 2),

it can be assumed that the event $E = [D' = P,\lambda' > k] \cup [D' = Q,\lambda' < k]$ satis-

fies the requirements of the lemma. Then $R(E) = 0$ and the desired conclusion

follows. □

Proof of Theorem 4: If $k \neq \infty$, $R(D' = P,\lambda' = \infty) \leq R(D' = P,\lambda = \infty) = R(D = P,\lambda = \infty)$

$= 0$. If $k = \infty$, $R(D' = P, \lambda' = \infty) = R(D' = P) - R(D' = P,\lambda' < \infty) \leq R(D' = P)$

$- R(D' = P,\lambda < \infty) = R(D = P) - R(D = P,\lambda < \infty) = R(D = P,\lambda = \infty) = 0$. Thus

$R(D' = P, \lambda' = \infty) = 0$. Similarly, $R(D' = Q, \lambda' = 0) = 0$, and hence, D' is a proper decision for E'. □

3. <u>Weakly admissible tests</u>. Some of the terminology introduced in Section 2 to describe decisions will be adapted here, without formal announcement, to the descriptions of tests. E.g., (E,D) will be called a *likelihood ratio test* if D is a likelihood ratio decision for $E \in \mathcal{E}$.

Let (E,D) be a test with error probabilities (α,β), and let (E',D'), with error probabilities (α',β'), denote a general competitor to (E,D). The test (E,D) is called *weakly admissible* (relative to $\mathcal{E}$) unless there exists a test (E',D') for which

(5) $$\alpha' \leq \alpha, \quad \beta' \leq \beta, \quad E' \subseteq E ,$$

and which is strictly better in one of these respects, i.e., $\alpha' < \alpha$ or $\beta' < \beta$ or E' is a proper subset of E. Actually, it is undesirable to distinguish between two experiments whose sets are comparable up to a P and Q equivalence. Therefore, in order to avoid problems with the definition of weak admissibility given here, we shall insist that each of the experiments $E \in \mathcal{E}$ is complete under the probability measure $R = \frac{1}{2}(P+Q)$. I.e., for each $E \in \mathcal{E}$ and $E \in \mathcal{E}$, if $F \in \mathcal{F}$ and $R(E \triangle F) = 0$, then $F \in E$. (Alternatively, one could replace the experiments appearing in (5) by their completions under R. But this would needlessly complicate the notation and theory which follow.) With this convention, if (E,D) is a test and D' is equivalent to D, then D' is E-measurable. For this reason, we will no longer distinguish between equivalent decisions; equivalent decisions will be interpreted as equal.

A test (E,D) cannot be weakly admissible if there is another test (E',D') with D' equal to D and $E' \in \mathcal{E}$ strictly included in E. To be weakly admissible it is therefore necessary that a test be *minimal*. That is, there must be no

E' in E with E' strictly included in E and such that D is measurable over E'. However, even a minimal and proper likelihood ratio test need not be weakly admissible as the next example demonstrates.

Example 2. Let X_1 and X_2 be i.i.d. Bernoulli variables under P and Q with means 1/3 and 2/3, respectively. Further, let $E = \{E',E\}$ where $E' = \sigma(X_1)$ and $E = \sigma(X_1,X_2)$. Then (E,D) is a minimal and proper likelihood ratio test when D is the decision which chooses Q iff $X_2 = 1$. However, (E,D) is not weakly admissible because the test (E',D') is better, where D' is the decision which chooses Q iff $X_1 = 1$. $\qquad\qquad\square$

This example suggests that there is a more stringent property than that of minimality which every weakly admissible test must satisfy. A test (E,D) is said to be *curtailed* (relative to E) if there is no other test (E',D') with $E' \in E$ strictly included in E and with $D' \overset{k}{=} D$ (for E) for some $k \in [0,\infty]$. A curtailed test is minimal, but the converse is not true.

Theorem 5. Curtailed proper likelihood ratio tests are weakly admissible. Conversely, a weakly admissible test (E,D) is a curtailed proper likelihood ratio test if P is nonatomic on E.

Proof: If (E,D) is a proper likelihood ratio test of level k and (E',D') is any (competing) test which satisfies (5), then by Theorem 1, $D' \overset{k}{=} D$ for E, and hence $\alpha' = \alpha$ and $\beta' = \beta$. If, in addition, (E,D) is curtailed, then E' cannot be a proper subset of E. Thus (E,D) is weakly admissible. Conversely, if (E,D) is weakly admissible, it obviously is curtailed. Now consider the test (E',D'), where $E' = E$ and $D' = P$, D or Q as λ (the likelihood ratio for E) $= 0$, $\lambda \in (0,\infty)$ or $\lambda = \infty$, respectively. Since $P(\lambda = \infty) = Q(\lambda = 0) = 0$, the test (E',D') satisfies (5). Since (E,D) is weakly admissible, $\alpha' = \alpha$ and

$\beta' = \beta$. Then

$$R(D=P, \ \lambda=\infty) \ = \ Q(D=P, \ \lambda=\infty) \ = \ \alpha \ - \ Q(D=P, \ \lambda<\infty)$$

$$= \ \alpha' \ - \ Q(D'=P, \ \lambda<\infty) \ = \ Q(D'=P, \ \lambda=\infty) \ = \ 0 \ .$$

Similarly, $R(D=Q, \ \lambda=0) = 0$, and hence (E,D) is a proper test. Suppose, in addition, that P is nonatomic on E. Then by the Neyman-Pearson Lemma (described herein), there is a most powerful likelihood ratio decision D' for E (not to be confused with any previous D'), which has the same size as that for D. I.e., $\alpha' = \alpha$ and $\beta' \leq \beta$. Consequently (5) holds if E' is set equal to E. Since (E,D) is weakly admissible, β' must equal β. Thus D' and D are both most powerful decisions of size α (for E). The desired conclusion that D is a likelihood ratio decision (for E) then follows from the Neyman-Pearson Lemma. ☐

4. <u>Curtailed tests and curtailed versions of a test</u>. While it is a completely trivial matter, in most cases, to check whether a test is a proper likelihood ratio test, it is not always so easy to determine whether a test is curtailed. For instance, it is obvious that every sequential probability ratio test is a proper likelihood ratio test, but it is not obvious that all such tests are curtailed (with respect to the class of experiments that can be realized through sequential sampling). Nevertheless, this must be the case since Eisenberg, Ghosh and Simons (1976) show that every sequential probability ratio test is weakly admissible.

A curtailed test (minimal test) (E',D') will be called a curtailed version (minimal version) of the test (E,D) if $E' \subseteq E$ and $D' \overset{k}{=} D$ for E and some $k \in [0,\infty]$ (if $E' \subseteq E$ and D' is equal to D). In Example 2, the test (E',D') is a curtailed version of (E,D), while each of these two tests is a minimal version of itself.

Because of the importance of the notion of weak admissibility, it is apparent from the previous theorem that a study of curtailed versions should have priority over a study of minimal versions. But since a minimal version is sometimes a curtailed version, we shall first discuss an elementary theorem about minimal versions and then turn our attention to curtailed versions.

Theorem 6. _If_ E _is closed under the formation of intersections, then every test_ (E,D) _has a unique minimal version._

Proof: Let E' be the intersection of experiments in E for which D is a decision (i.e., for which D is measurable). Then $E' \epsilon E$, and D is E'-measurable. Clearly, (E',D) is a minimal version of (E,D) and no other minimal version is possible.

$\square$

Remark 2. There is an important situation when the minimal version promised by this theorem is a curtailed version for (E,D). This occurs when (E,D) is a proper likelihood ratio test of level k and $R(\lambda=k) = 0$, where λ is the likelihood ratio for E. For suppose (E',D') is a minimal version of (E,D) which is _not_ a curtailed test. Then there exists a test (E'',D'') with E'' properly contained in E' and $D'' \overset{k'}{=} D'$ for E' and some $k' \epsilon [0,\infty]$. But then D'' has the same error probabilities as D' and D, and hence, by Theorem 1, $D'' \overset{k}{=} D$ for E. Finally, since $R(\lambda=k) = 0$, it follows that D is E''-measurable, which is untenable with the assumption that (E',D') is a minimal version of (E,D).

The situation when $R(\lambda=k) > 0$ is much more complicated. In order for (E,D) to have a curtailed version, it is no longer sufficient that E be closed under the formation of intersections.

Theorem 7. _Suppose_ (E,D) _is a proper likelihood ratio test of level_ k _and_ $R(\lambda=k) > 0$. _If the class_ E _is closed under the formation of intersections and_ P _is nonatomic on (each experiment of)_ E, _then_ (E,D) _has a curtailed version_

but (in general) there is no assurance that it is unique.

Proof: Let E_o denote the class of experiments E' in E for which there is a decision D' that is level k equivalent to D for E. It is easily checked that a particular pair (E',D') is a curtailed version of (E,D) if E' is a minimal member of E_o in the sense of set inclusion. Now E_o possesses at least one minimal member, according to Zorn's Lemma, if every "chain" in E_o possesses a "lower bound" in E_o. I.e., if Γ is a linearly ordered index set and $\{E_\gamma, \gamma \in \Gamma\}$ is a non-increasing collection of experiments in E_o, then it must be shown that $E_o = \bigcap_{\gamma \in \Gamma} E_\gamma$ belongs to E_o, i.e., there is an E_o-measurable decision D_o that is level k equivalent to D. Briefly, this can be demonstrated as follows: Since P is nonatomic on E_γ, the range of values of $P(E)$, for $E \in E_\gamma$ satisfying $[\lambda > k] \subseteq E \subseteq [\lambda \geq k]$, is a closed interval (cf., Halmos (1950), page 174). Denote the left and right endpoints by $a_\gamma = P(A_\gamma)$ and $b_\gamma = P(B_\gamma)$, respectively. By assumption, there exists a decision D_γ for E_γ which is level k equivalent to D. This implies that the size of D_γ is the same as D; denote it by α. But $[\lambda > k] \subseteq [D_\gamma = Q] \subseteq [\lambda \geq k]$ and, hence, $a_\gamma \leq \alpha \leq b_\gamma$, for each $\gamma \in \Gamma$. Clearly, the intervals $[a_\gamma, b_\gamma]$ are nonincreasing and decrease down to some limiting interval $[a_o, b_o]$ containing α. Let γ_n be a nondecreasing sequence in Γ for which $a_{\gamma_n} \uparrow a_o$ and $b_{\gamma_n} \downarrow b_o$ as $n \to \infty$. Set $A_o = \limsup_{n \to \infty} A_{\gamma_n}$ and $B_o = \limsup_{n \to \infty} B_{\gamma_n}$. It is easy to check that the sets A_o and B_o are E_o-measurable, that they contain $[\lambda > k]$ and are contained in $[\lambda \geq k]$, and that they have probabilities $P(A_o) = a_o$ and $P(B_o) = b_o$. Since P is nonatomic on E_o, there exists a set $C_o \in E_o$ for which $[\lambda > k] \subseteq C_o \subseteq [\lambda \geq k]$ and $P(C_o) = \alpha$. Then the decision D_o which chooses Q on C_o is level k equivalent to D. Thus $E_o \in E_o$, as required.

The next example shows that the curtailed version (E',D') guaranteed by Zorn's Lemma does not have to be unique. A less artificial (though more complicated) example appears implicitly in Case 3 of Example 5 below. $\square$

Example 3. Let X_1 and X_2 be independent Bernoulli variables with common means $p = 1/3$ under P and $p = 2/3$ under Q. Further, let U be a uniform variable on $[0,1]$ which is independent of X_1 and X_2 under P and Q. Set $E = \sigma(X_1,X_2,U)$, $E_1' = \sigma(X_1,U)$, $E_2' = \sigma(X_2,U)$, $E'' = \sigma(U)$ and $E = \{E, E_1', E_2', E''\}$. P is nonatomic on E and E is closed under the formation of intersections. Consider the test (E,D) which chooses P if $X_1 = 0$ and Q if $X_1 = 1$. This is a proper likelihood ratio test of level one. It is easily checked that (E_1', D_1') and (E_2', D_2') are each curtailed versions of (E,D), where $D_1' \equiv D$ and $D_2' = P$ or Q as $X_2 = 0$ or 1, respectively. □

Theorem 8. A curtailed version of a proper level k likelihood ratio test is weakly admissible and a proper level k likelihood ratio test.

Proof: This follows immediately from Theorems 3, 4 and 5. □

Denote the likelihood ratios for the experiments E and E', appearing in Theorem 3, by λ and λ' respectively. It is easily seen from this theorem that (under the assumptions of the theorem)

(6) $\qquad$ $R(k \text{ is strictly between } \lambda \text{ and } \lambda') = 0$.

For the purposes of the next section, we shall establish a slightly stronger result:

Theorem 9. Assume D is a level k likelihood ratio decision for E and there exists a decision D' for $E' \subseteq E$ with $D' \overset{k}{=} D$ for E. Then
$$R(\lambda' \leq k < \lambda) = R(\lambda < k \leq \lambda') = 0 .$$
In particular, $\lambda = k$ on the set $[\lambda' = k]$ except on an R-null set.

Proof: We shall only show that $R(\lambda' \leq k < \lambda) = 0$, since the proof that $R(\lambda < k \leq \lambda')$ $= 0$ is similar. If $P(\lambda' \leq k < \lambda) > 0$, then $k < \infty$ and

$$0 < \int_{[\lambda'\leq k\leq\lambda]} (\lambda-k)dP = \left\{\int_{[\lambda'\leq k]} + \int_{[\lambda<k<\lambda']} - \int_{[\lambda<k]}\right\}(\lambda-k)dP$$

$$= \left\{\int_{[\lambda'\leq k]} - \int_{[\lambda<k]}\right\}(\lambda-k)dP \qquad \text{(by (6))}$$

$$= \left\{\int_{[\lambda'\leq k]} - \int_{[D=P]}\right\}(\lambda-k)dP$$

$$= \left\{\int_{[\lambda'\leq k]} - \int_{[D'=P]}\right\}(\lambda-k)dP \qquad \text{(by (3))}$$

$$= \int_{[\lambda'=k,D'=Q]} (\lambda-k)dP$$

$$= \int_{[\lambda'=k,D'=Q]} (\lambda-\lambda')dP$$

$$= Q(\lambda'=k,D'=Q,\lambda\neq\infty) - Q(\lambda'=k,D'=Q) \qquad \text{(by (1))}$$

$$= -Q(\lambda'=k,D'=Q,\lambda=\infty) \leq 0 \ ,$$

which contradicts the initial strict inequality. Thus $P(\lambda'\leq k<\lambda) = 0$ and it follows from the equations above that

$$(7) \qquad\qquad Q(\lambda'=k,D'=Q,\lambda=\infty) = 0 \ .$$

In turn, if $R(\lambda'\leq k<\lambda) > 0$, then $k < \infty$ and

$$0 < Q(\lambda'\leq k<\lambda) = Q(\lambda'=k<\lambda) \qquad \text{(by (6))}$$

$$= Q(\lambda'=k<\lambda,D=Q) = Q(\lambda'=k<\lambda,D'=Q) \qquad \text{(by (3))}$$

$$= Q(\lambda'=k<\lambda<\infty,D'=Q) \qquad \text{(by (7))}$$

$$\leq \int_{[\lambda' \leq k < \lambda]} \lambda dP \qquad \text{(by (1))}$$

$$= 0 \, ,$$

which contradicts the initial strict inequality. Thus $R(\lambda' \leq k < \lambda) = 0$. □

We complete this section with an interesting (though very artificial) example which shows that Theorem 7 would be false if the assumption that P is nonatomic on E were dropped.

Example 4. Let $X_1, X_2, \ldots$ be independent uniform variables on $[0,1]$ under *both* P and Q. Set $E = \{E_n : 1 \leq n \leq \infty\}$, where E_n, $1 \leq n < \infty$, is the closure under R of $\sigma(X_n, X_{n+1}, \ldots)$, and $E_\infty = \cap_{n=1}^\infty E_n$ is the closure of the tail σ-field of the X_i's. Clearly, E is closed under the formation of intersections, but P *is atomic on* E_∞ because of Kolmogorov's zero-one law. Also, $\lambda_n \equiv 1$ is the likelihood ratio of E_n, $1 \leq n \leq \infty$. Consider the level $k = 1$ proper likelihood ratio test (E_1, D), where $D = P$ or Q as $X_1 \leq \frac{1}{2}$ or $> \frac{1}{2}$, respectively. It has error probabilities $\alpha = \beta = \frac{1}{2}$. Clearly there is no curtailed version of this test: The only possible candidates are of the form (E_n, D') with $n < \infty$, and all such tests can be improved upon by increasing n and modifying D' in an obvious way. □

5. <u>Applications to sequential tests.</u> In applications, one is usually concerned with a nested sequence of experiments $E_1 \subseteq E_2 \subseteq \ldots$, where, typically, E_n represents the "information" obtainable from a sample of size n, i.e., $E_n = \sigma(X_1, \ldots, X_n)$, where $X_1, \ldots, X_n$ denote n observations. Fixed sample size tests take the basic form (E_n, D), while sequential tests take the basic form (E_N, D), where N is a stopping time adapted to $\{E_n, 1 \leq n \leq \infty\}$, $E_\infty = V_{n=1}^\infty E_n$ (the smallest σ-field which includes every E_n, $n \geq 1$) and E_N is the σ-field of

"events up to time N", i.e.,

$$E_N = \{E \epsilon F: E[N=n] \epsilon E_n, \quad 1 \le n \le \infty\}.$$

We shall not be concerned here with how the experiments E_n arise, but we shall insist, as we do of experiments in the previous sections, that they are complete under the probability measure R. It follows that the experiments E_N will be complete under R as well.

There are many different classes E of possible experiments that have been considered in the past, e.g.,

(i) *fixed sample size experiments*: $E = \{E_n, \ 1 \le n < \infty\}$,

(ii) *two-stage experiments*: $E = \{E_N\}$, where N is E_m-measurable for some $m \ge 1$ (depending on N) and $N \ge m$,

(iii) *general sequential experiments*: $E = \{E_N\}$, where $R(N < \infty) = 1$,

(iv) *truncated sequential experiments*: $E = \{E_N\}$, where $N \le n$ for some $n \ge 1$ (depending on N or on the class E itself),

(v) *experiments for power one tests*: $E = \{E_N\}$, where $Q(N < \infty) = 1$.

The tests referred to in (v) have been popularized by Robbins and others. They advocate using the decision D which chooses Q when $N < \infty$ and P when $N = \infty$ (i.e., the decision maker acts as if he believes P is correct until the experiment stops, if ever). The resulting test (E_N, D) has power $Q(D=Q) = Q(N < \infty) = 1$ and size $\alpha = P(N < \infty)$, which depends on how N is chosen.

The assumption in Theorems 6 and 7 that E is closed under the formation of intersections holds for all of the classes (i) - (v) except (ii), and even for that class Theorem 7's conclusion, namely that curtailed versions of proper likelihood ratio tests always exist, can be sustained. (The assumption that P is nonatomic on E is still required. Also, the conclusion of Theorem 6 is valid for the class, except for the uniqueness of the minimal version, which

seems unlikely.) Thus the problem of existence of curtailed versions is a relatively minor one. The problem of finding them can be more difficult.

Let us turn to this problem within the specific context of general sequential experiments (case (iii)). Theorem 9 points to a partial solution: Let (E_N, D) be a fixed general sequential test which (from now on) we suppose to be a proper likelihood ratio test of level k. Further, let λ_N denote the likelihood ratio for E_N and λ_n denote the likelihood ratio for E_n, $n \geq 1$. Finally, let

$$N' = \text{the first } n \geq 1 \text{ such that } R^{E_n}(\lambda_n \leq k < \lambda_N) = R^{E_n}(\lambda_N < k \leq \lambda_n) = 0 \ .$$

It is readily seen that N' is a well defined stopping time and, in particular, $N' \leq N$ a.s. (R). In fact, $N' \leq M$ a.s. (R) for every stopping time M such that

$$(8) \qquad R(\lambda_M \leq k < \lambda_N) = R(\lambda_N < k \leq \lambda_M) = 0 \ .$$

For on $[M=n]$,

$$R^{E_n}(\lambda_n \leq k < \lambda_N) = R^{E_n}(\lambda_M \leq k < \lambda_N) = 0 \quad \text{a.s.}$$

and

$$R^{E_n}(\lambda_N < k \leq \lambda_n) = R^{E_n}(\lambda_N < k \leq \lambda_M) = 0 \quad \text{a.s.}$$

Remark 3. Suppose (E_M, D_o) is any test for which $E_M \subseteq E_N$ and $D_o \overset{k}{=} D$ for E_N. By Theorem 9, (8) holds and, hence, $E_{N'} \subseteq E_M$. In particular, if (E_M, D_o) is a curtailed (or the unique minimal) version of (E_N, D), then $E_{N'} \subseteq E_M$.

Theorem 10. The following are equivalent:

(i) There exists a curtailed version of (E_N, D) of the form $(E_{N'}, D')$.

(ii) There is a decision D' for $E_{N'}$ which is level k equivalent to D for E_N.

(iii) There exists no curtailed version of (E_N, D) not of the form $(E_{N'}, D')$.

<u>Proof</u>: Condition (i) implies (ii) because of the definition of curtailed version. Likewise, condition (ii) implies (iii) because of the definition of a curtailed version, since every curtailed version (E_M, D_o) not of the form $(E_{N'}, D')$ would have to be such that $E_{N'}$ is a *proper* subset of E_M, on account of Remark 3. Finally, condition (iii) implies (i) because, according to Remark 2 and Theorem 7, (E_N, D) must have a curtailed version. $\quad \Box$

Theorem 10 is easiest to apply when the decision D' referred to in (ii) can be D itself. This occurs when $R(\lambda_N = k) = 0$ (see Remark 2) and, more generally, whenever $(E_{N'}, D)$ (is a test and) turns out to be the unique minimal version promised by Theorem 6.

There is a less compact way of describing the stopping time N', which is somewhat easier to intuit. Since

$$R^{E_n}(\lambda_n \leq k < \lambda_N) = I_{[\lambda_n \leq k]} R^{E_n}(\lambda_N > k) \quad \text{and} \quad R^{E_n}(\lambda_N < k \leq \lambda_n) = I_{[\lambda_n \geq k]} R^{E_n}(\lambda_N < k) \, ,$$

(where I_A denotes the indicator function for an event A), due to the smoothing properties of conditional expectations, it follows that N' is the first integer $n \geq 1$ such that (a) $\lambda_n > k$ and $R^{E_n}(\lambda_N < k) = 0$, or (b) $\lambda_n < k$ and $R^{E_n}(\lambda_N > k) = 0$, or (c) $\lambda_n = k$ and $R^{E_n}(\lambda_N \neq k) = 0$. For the experimenter, case (a) can be interpreted as saying: If sampling has continued up to time n and it is found that (i) $\lambda_n > k$ and (ii) λ_N will not be less than k, then stop. Cases (b) and (c) have similar interpretations.

Theorem 10 leaves open the possibility that there is no curtailed version of (E_N, D) of the form $(E_{N'}, D')$. Unfortunately, this "possibility" can be realized. The following example illustrates a basic complication that can occur: For a given experiment E_N, the class of experiments E_M which admit a curtailed version (E_M, D_o) of (E_N, D) depends on which likelihood ratio decision D is being considered.

This is significant, for it should be remembered that the stopping time N'
is defined with no reference to D.   It also will be apparent that the complica-
tion cannot be removed by simply admitting randomized decisions.

<u>Example 5.</u>   Let $E_n = \sigma(X_1,\ldots,X_n,U)$ $(n \geq 1)$, where the random variables
$X_1,X_2,\ldots$ are independent Bernoulli variables with common means equal to 1/3
and 2/3 under P and Q, respectively, and U is a uniform variable on [0,1] which
is independent of the X_i's under P and Q (that is available for making "random-
ized decisions," if desired).   Further let $N \equiv 4$.   Then

$$N' = 2 \quad \text{if } X_1 + X_2 = 0 \text{ or } 2 \; ,$$
$$= 3 \quad \text{if } X_1 + X_2 = 1 \; .$$

Suppose D = P when $\lambda_4 < 1$ and D = Q when $\lambda_4 > 1$, so that D is a likelihood
ratio decision for E_N of level k = 1.   We shall illustrate three different
situations that can arise depending on how D is defined on $[\lambda_4 = 1]$.

Case I: <u>D = Q on $[\lambda_4 = 1, U \leq 1/2]$.</u>   Here, there is a unique curtailed version
of the form $(E_{N'},D')$.   Either $\lambda_{N'} < 1$, in which case D' = P is mandated, or
$\lambda_{N'} > 1$, in which case D' = Q is mandated.

Case II: <u>D = Q on $[\lambda_4 = 1]$.</u>   Here there is a unique curtailed version (E_M,D_o),
and it is not of the form (E_N',D').

$$M = 2 \quad \text{if } X_1 + X_2 = 2 \; ,$$
$$= 3 \quad \text{if } X_1 + X_2 = 1 \text{ and } X_3 = 1 \; ,$$
$$= 3 \quad \text{if } X_1 + X_2 = 0 \text{ and } X_3 = 0 \; ,$$
$$= 4 \quad \text{if } X_1 + X_2 = 0 \text{ and } X_3 = 1 \; ,$$

and D_o = Q on $[\lambda_M \geq 1]$.   Obviously, E_M is strictly larger than E_N'.   The
decision D_o is actually the same as D, and, hence, (E_M,D_o) is also the unique

minimal version of (E_N, D).

Case III: $D = Q$ on $[\lambda_4 = 1, U \leq 2/3]$. Here, there is no curtailed version (E_M, D_O) of the form $(E_{N'}, D')$, but, in contrast to the previous case, there is more than one choice for (E_M, D_O). Two possible choices for M are as follows:

First Choice	Second Choice
M = 2 if $X_1 + X_2$ = 0 or 2,	M = 2 if $X_1 + X_2$ = 2 ,
= 3 if $X_1 + X_2$ = 1 and X_3 = 1 ,	= 3 if $X_1 + X_2$ = 1 ,
= 3 if $X_1 + X_2$ = 1, X_3 = 0 and $U \leq 1/2$,	= 3 if $X_1 + X_2$ = 0 and X_3 = 0 ,
= 4 if $X_1 + X_2$ = 1, X_3 = 0 and $U > 1/2$.	= 4 if $X_1 + X_2$ = 0 and X_3 = 1 .

For both choices, $D_O = Q$ on $[\lambda_M \geq 1]$ is mandated.   □

We can offer no algorithm which finds a curtailed version (E_M, D_O) for (E_N, D) in every possible situation.

References

[1] Eisenberg, Ghosh and Simons (1976). Properties of generalized sequential probability ratio tests. *Ann. Statist.*, *4*, 237-251.

[2] Halmos, Paul (1950). *Measure Theory*. New York: Van Nostrand.

[3] Rényi, Alfred (1970). *Foundations of Probability*. San Francisco: Holden-Day.