The Annals of Statistics
1979, Vol. 7, No. 3, 569–578

ON THE ADMISSIBILITY OR INADMISSIBILITY OF FIXED SAMPLE SIZE TESTS IN A SEQUENTIAL SETTING[1]

BY L. D. BROWN, ARTHUR COHEN, AND W. E. STRAWDERMAN
Rutgers University

Questions pertaining to the admissibility of fixed sample size tests of hypotheses, when sequential tests are available, are considered. For the normal case with unknown mean, suppose the risk function is a linear combination of probability of error and expected sample size. Then any fixed sample size test, with sample size $n > 2$, is inadmissible. On the other hand, suppose the risk function consists of the pair of components, probability of error and expected sample size. Then any optimal fixed sample size test for the one sided hypothesis is admissible.

When the variance of the normal distribution is unknown, t-tests are studied. For one-sided hypotheses and componentwise risk functions the fixed sample size t-test is inadmissible if and only if the absolute value of the critical value of the test is greater than or equal to one. This implies that for the most commonly used sizes, the fixed sample size t-test is inadmissible.

Other loss functions are discussed. Also an example for a normal mean problem is given where a nonmonotone test cannot be improved on by a monotone test when the risk is componentwise.

1. Introduction and summary. In this study we consider questions pertaining to the admissibility of fixed sample size tests of hypotheses, when sequential tests are also available. Most of the questions refer to specific distributions such as the normal, exponential family, and Student's t, although the proofs will work in other cases as well.

The first results are for the normal case with unknown mean and the hypothesis is one-sided or two-sided. If the risk function is a linear combination of probability of error and expected sample size, then any fixed sample size test is inadmissible. We will exhibit a better test. If, on the other hand the risk function is componentwise, that is, consists of the two components, probability of error and expected sample size, then any optimal fixed sample size test for the one-sided hypothesis is admissible. This latter result is true when the underlying distribution is exponential family and the support of the distribution is the entire real line. These results are true for other distributions and more general loss functions. See Remarks 2.2 and 2.3.

The next set of results is concerned with t-tests for the one-sided hypothesis, when the risk function is componentwise. The conclusion here is that the fixed sample size t-test is inadmissible if and only if the absolute value of the critical value for the test is greater than or equal to one. Hence for the most commonly

Received November 1977; revised May 1978.
[1]Research supported by N.S.F. Grant No. MCS-75-23343.
AMS 1970 *subject classifications.* Primary 62L10; secondary 62C15.
Key words and phrases. Sequential tests, admissibility, stopping rules, t-test, monotone tests.

used sizes, the fixed sample sized t-test is inadmissible. (These results are true for Hotelling's T^2 test as well. See Remark 3.2.) It will also be shown how to construct tests which improve on the fixed sample size t, that are admissible.

The first results for the normal distribution are in Section 2. The t-test results are in Section 3. In Section 4 we discuss an ordering for loss functions for sequential testing problems. Also we again consider the normal one-sided testing problem. This time we offer an example of a nonmonotone test procedure whose risk cannot be matched by a monotone test procedure when the risk function is componentwise. For the linear combination risk the monotone procedures are an essentially complete class. See Brown, Cohen and Strawderman [2].

2. Normal distribution with unknown mean. The model in this section is that X_i, $i = 1, 2, \cdots$ are a sequence of independent, identically distributed normal random variables with unknown mean θ and known variance which, without loss of generality, is taken to be one. We write $\mathbf{X} = (X_1, X_2, \cdots)$ so that $\mathbf{X}$ lies in an infinite product space.

The first situation considered is to test the one-sided hypothesis $H_1 : \theta < 0$ vs. $H_2 : \theta > 0$. We evaluate tests by a linear combination of the probability of error and expected sample size. Before stating the first theorem we need some preliminaries that will be appropriate for all remaining sections.

The parameter space Θ has typical element θ. We test $H_1 : \theta \in \Theta_1$ against $H_2 : \theta \in \Theta_2$. The action space consists of pairs (n, τ), where n is the stopping time and $\tau = 1$ or 2 depending on whether H_1 or H_2 is chosen. The loss for action $(n, 1)$ is nc for $\theta \in \Theta_1$ and $nc + 1$ for $\theta \in \Theta_2$. For action $(n, 2)$ the loss is $nc + 1$ for $\theta \in \Theta_1$ and nc for $\theta \in \Theta_2$. Here $c > 0$ represents the cost of taking an observation. A decision function $\delta(\mathbf{x})$ consists of a set of nonnegative functions $\delta_{ij}(x_1, x_2, \cdots, x_j)$, $i = 0, 1, 2$; $j = 1, 2, \cdots$, such that $\Sigma_{i=0}^{2}\delta_{ij}(x_1, x_2, \cdots, x_j) = 1$, where $\delta_{0j}(x_1, x_2, \cdots, x_j)$ represents the probability of taking the $(j + 1)$st observation given that $(x_1, x_2, \cdots, x_j)$ have been observed; $\delta_{1j}(x_1, x_2, \cdots, x_j)$ represents the probability of stopping at stage j and accepting H_1 given $(x_1, x_2, \cdots, x_j)$ have been observed; $\delta_{2j}(x_1, x_2, \cdots, x_j)$ represents the probability of stopping at stage j and accepting H_2 given $(x_1, x_2, \cdots, x_j)$ have been observed. Now for $j = 2, 3, \ldots$ let $\psi_j(x_1, x_2, \cdots, x_{j-1}) = \Pi_{k=1}^{j-1}\delta_{0k}(x_1, x_2, \cdots, x_k)$, represent the conditional probability of not stopping after the first $(j - 1)$ observations. (We let $\psi_1 \equiv 1$ without loss of generality for the problems treated here.) The risk function for the test procedure $\delta(\mathbf{x})$ can now be written as $\rho(\theta, \delta)$ where

$$
\begin{aligned}
(2.1)\quad \rho(\theta, \delta) &= \Sigma_{j=1}^{\infty}E_\theta\{\psi_j(X_1, X_2, \cdots, X_{j-1})[(1 - \delta_{0j}(X_1, X_2, \cdots, X_j))cj \\
&\quad + \delta_{2j}(X_1, X_2, \cdots, X_j)]\}, \qquad \text{for } \theta < 0; \\
&= \Sigma_{j=1}^{\infty}E_\theta\{\psi_j(X_1, X_2, \cdots, X_{j-1})[(1 - \delta_{0j}(X_1, X_2, \cdots, X_j))cj \\
&\quad + \delta_{1j}(X_1, X_2, \cdots, X_j)]\}, \qquad \text{for } \theta > 0.
\end{aligned}
$$

For nonrandomized test procedures, (i.e., $\delta_{ij}(x_1, x_2, \cdots, x_j) = 0$ or 1, for each $i, j, (x_1, x_2, \cdots, x_j)$) (2.1) may be written as

$$(2.2)\qquad \begin{aligned} \rho(\theta, \delta) &= P_\theta(W_\delta) + cE_\theta n_\delta \qquad \text{for } \theta < 0;\\ &= P_\theta(V_\delta) + cE_\theta n_\delta \qquad \text{for } \theta > 0, \end{aligned}$$

where W_δ is the critical region in the infinite product sample space, V_δ is the acceptance region, and n_δ is the random stopping time. Let $S_k = \Sigma_{i=1}^k X_i$ and let the fixed sample size test, denoted by $\delta(\mathbf{x})$, based on M observations be, reject if

$$(2.3)\qquad S_M > Mb,$$

for some constant b. We prove

THEOREM 2.1. *The test $\delta(\mathbf{x})$, based on* (2.3) *is inadmissible.*

PROOF. Consider the test $\delta'(\mathbf{x})$ which does the following: if $S_{M-1} < -(M-1)a$, then stop at stage $(M-1)$ and accept. If $S_{M-1} > M(a+b) - a$, then stop at stage $(M-1)$ and reject. Otherwise, stop at stage M and reject if $S_M > Mb$. Here a is a positive constant to be determined. We show for a, suitably chosen, that $\delta'(\mathbf{x})$ is better than $\delta(\mathbf{x})$. To do this first let $R_I = \{(S_{M-1}, X_M) : (S_{M-1} + X_M) > Mb\} \cap \{(S_{M-1}, X_M) : S_{M-1} < -(M-1)a\}$ and $R_{II} = \{(S_{M-1}, X_M) : (S_{M-1} + X_M) \leqslant Mb\} \cap \{S_{M-1}, X_M) : S_{M-1} > M(a+b) - a\}$. Now use (2.2) to note that the difference in risks, for $\theta < 0$, is

$$\begin{aligned} (2.4)\quad \rho(\theta, \delta) - \rho(\theta, \delta') &\\ &= [P_\theta(R_I) - P_\theta(R_{II})]\\ &\quad + c[P_\theta(S_{M-1} < -(M-1)a) + P_\theta(S_{M-1} > M(a+b) - a)]\\ &\geqslant cP_\theta(S_{M-1} < -(M-1)a)\\ &\quad - P_\theta(S_{M-1} > M(a+b) - a)P_\theta(X_M \leqslant -(M-1)a)\\ &= c\Phi\big(-(M-1)^{\frac{1}{2}}[a+\theta]\big)\\ &\quad - \Phi(-(M-1)a - \theta)P_\theta(S_{M-1} > M(a+b) - a)\\ &\geqslant \Phi(-(M-1)a - \theta)[c - P_\theta(S_{M-1} > M(a+b) - a)]. \end{aligned}$$

Clearly (2.4) will be positive for all $\theta < 0$ provided a is chosen so that

$$(2.5)\qquad P_\theta(S_{M-1} > M(a+b) - a) < c,$$

and

$$(2.6)\qquad \Phi(-(M-1)a - \theta) > 0.$$

For all sufficiently large a, the conditions in (2.5) and (2.6) are satisfied. If $\theta > 0$, a similar argument shows again that the difference in risks is positive for all sufficiently large a. This completes the proof of the theorem.

Next we consider the same model but test the two sided hypothesis $H_0 : \theta = 0$ vs. $H_1 \theta \neq 0$. The fixed sample size test studied is the UMPU test, denoted by $\delta(\mathbf{x})$, which rejects if

$$|S_M| > Mb. \tag{2.7}$$

We prove

THEOREM 2.2. *The test* $\delta(\mathrm{x})$ *based on* (2.7) *is inadmissible.*

PROOF. Consider the test $\delta'(\mathrm{x})$ which stops and rejects if $|S_{M-1}| > (M-1)a$, for a suitably chosen. Otherwise $\delta'(\mathrm{x})$ does as $\delta(\mathrm{x})$. For $\theta \neq 0$, since $\delta'(\mathrm{x})$ rejects more often and since $E_\theta n$ under δ' is clearly less than under δ, the risk for δ' is less than δ. Therefore the only case to study is when $\theta = 0$. Let $R_I = \{(S_{M-1}, X_M) : |S_{M-1} + X_M| < Mb\} \cap \{(S_{M-1}, X_M) : S_{M-1} < -(M-1)a\}$, $R_{II} = \{(S_{M-1}, X_M) : |S_{M-1} + X_M| < Mb\} \cap \{(S_{M-1}, X_M) : S_{M-1} > (M-1)a\}$. The difference in risks from (2.2) is

$$\begin{aligned} \rho(0, \delta) - \rho(0, \delta') &= cP_0\{|S_{M-1}| > (M-1)a\} - P_0(R_I) - P_0(R_{II}) \\ &\geqslant 2\{cP_0(S_{M-1} > (M-1)a) \\ &\quad - P_0(S_{M-1} > (M-1)a)P_0(X_M < Mb - (M-1)a)\}. \end{aligned} \tag{2.8}$$

Now proceed as in the proof of Theorem 2.1 to complete the proof of this theorem.

REMARK 2.1. Other fixed sample size tests (not necessarily unbiased ones) can be similarly shown to be inadmissible.

The last result of this section is concerned with the one-sided hypothesis, but now the risk is componentwise. That is, for a procedure $\delta(\mathrm{x})$, $\rho(\theta, \delta)$ for $\theta < 0$, consists of the pair

$$E_\theta n_\delta = \Sigma_{j=1}^{\infty} E_\theta\{\psi_j(X_1, X_2, \cdots, X_{j-1})(1 - \delta_{0j}(X_1, X_2, \cdots, X_j))j\}, \tag{2.9}$$

and

$$P_\theta(\text{Rejecting } H_1) = \Sigma_{j=1}^{\infty} E_\theta\{\psi_j(X_1, X_2, \cdots, X_{j-1})\delta_{2j}(X_1, X_2, \cdots, X_j)\}. \tag{2.10}$$

For $\theta > 0$, $\rho(\theta, \delta)$ consists of the pair (2.9) and

$$P_\theta(\text{Accepting } H_1) = \Sigma_{j=1}^{\infty} E_\theta\{\psi_j(X_1, X_2, \cdots, X_{j-1})\delta_{1j}(X_1, X_2, \cdots, X_j)\}. \tag{2.11}$$

We prove

THEOREM 2.3. *The fixed sample size test* $\delta(\mathrm{x})$ *based on* (2.3) *is admissible.*

PROOF. Suppose $\delta(\mathrm{x})$ is not admissible. Then there exists a test $\delta'(\mathrm{x})$ whose componentwise risk is better. In particular,

$$E_\theta n_{\delta'} \leqslant E_\theta n_\delta = M \tag{2.12}$$

and

$$\begin{aligned} &P_\theta(\text{Rejecting } H_1 \text{ under } \delta') \leqslant P_\theta(\text{Rejecting } H_1 \text{ under } \delta) \text{ for } \theta < 0 \\ &P_\theta(\text{Accepting } H_1 \text{ under } \delta') \leqslant P_\theta(\text{Accepting } H_1 \text{ under } \delta) \text{ for } \theta > 0, \end{aligned} \tag{2.13}$$

with strict inequality in (2.12) and/or (2.13) for some θ. The inequality in (2.12) implies that for $\delta'(\mathrm{x})$ there must exist a set of positive Lebesque measure on which the probability of stopping at stage $(M-1)$ or less is positive. If not, then it must be true that $\delta'(\mathbf{x})$ always stops exactly at stage M. Since $\delta(\mathbf{x})$ is UMP among fixed sample size tests, strict inequality could not hold in (2.13). In fact for $i = 1, 2$, let $A_{ij}^K = \{(x_1, x_2, \cdots, x_j) : (\prod_{k=1}^{j-1}\delta'_{0k}(x_1, x_2, \cdots, x_k))\delta'_{ij}(x_1, x_2, \cdots, x_j) > 1/K\}$, for $K = 1, 2, \cdots$. Then if μ denotes Lebesgue measure we must have $\mu(\cup_{i=1}^2 \cup_{j=1}^{M-1} \cup_{K=1}^{\infty} A_{ij}^K) > 0$. That is, for some i, $1 \leqslant j \leqslant M-1$, $K < \infty$, $\mu(A_{ij}^K) > 0$. For now let $\mu(A_{2j}^K) = \varepsilon > 0$ for some $j \leqslant M-1$ and some $K < \infty$. At this point we use the argument used by Stein [5] and Birnbaum [1]. That is, let E^j represent Euclidean j space, let $\gamma > 0$, $B > 0$, and consider

$$
\begin{aligned}
(2.14)\qquad &\left[P_\theta(\text{Rejecting } H_1 \text{ under } \delta')/P_\theta(\text{Rejecting } H_1 \text{ under } \delta)\right]\\
&\geqslant (1/K)P_\theta(A_{2j}^K)/P_\theta(S_M > Mb)\\
&= (1/K)\textstyle\int_{\{A_{2j}^K\times R^{(M-j)}\}} e^{-\frac{1}{2}[\Sigma_{i=1}^M (x_i-\theta)^2]}\\
&\quad\times \textstyle\prod_{i=1}^M dx_i / \int_{\{S_M > Mb\}} e^{-\frac{1}{2}[\Sigma_{i=1}^M (x_i-\theta)^2]}\prod_{i=1}^M dx_i\\
&\geqslant (1/K)\textstyle\int_{\{A_{2j}^K\times R^{(M-j)}\}\cap\{S_M < Mb-\gamma\}} e^{-\frac{1}{2}[\Sigma_{i=1}^M (x_i-\theta)^2]}\\
&\quad\times \textstyle\prod_{i=1}^M dx_i / \int_{\{S_M > Mb\}} e^{-\frac{1}{2}[\Sigma_{i=1}^M (x_i-\theta)^2]}\prod_{i=1}^M dx_i\\
&\geqslant (1/K)e^{-\theta\gamma}\textstyle\int_{\{A_{2j}^K\times R^{(M-j)}\}\cap\{S_M < Mb-\gamma\}} e^{-\frac{1}{2}\Sigma_{i=1}^M x_i^2}\\
&\quad\times \textstyle\prod_{i=1}^M dx_i / \int_{\{S_M > Mb\}} e^{-\frac{1}{2}\Sigma_{i=1}^M x_i^2}\prod_{i=1}^M dx_i\\
&= Be^{-\gamma\theta}.
\end{aligned}
$$

Clearly $\lim_{\theta\to-\infty} Be^{-\gamma\theta} = \infty$ which contradicts (2.13). A similar argument holds if $\mu(A_{1j}^K) = \varepsilon > 0$. This completes the proof of the theorem.

REMARK 2.2. Suppose the X_i are multivariate normal with unknown mean $\boldsymbol{\theta}$ and known covariance matrix and consider the fixed sample size chi-square test of $H_0 : \boldsymbol{\theta} = 0$. The theorems of this section are still true.

REMARK 2.3. Let the $0-1$ loss in the terminal decision be replaced by a more general function of the parameter, say $L(\theta, \tau)$. (An example would be linear loss, which for the one-sided problem would be $L(\theta, 1) = 0$ if $\theta < 0$, $L(\theta, 1) = k_1\theta$ if $\theta > 0$, $L(\theta, 2) = -k_2\theta$ if $\theta < 0$, $L(\theta, 2) = 0$ if $\theta > 0$, where k_1 and k_2 are positive constants.) Then, for the normal case, the theorems of this section would be true with the addition of conditions like $L(\theta, 2)\Phi(\theta) \to 0$ as $\theta \to -\infty$. For other distributions similar type conditions would be required.

3. Student's t-tests. The model in this section is that X_i, $i = 1, 2, \cdots$ are a sequence of independent, identically distributed normal random variables with unknown mean θ and unknown variance σ^2. The one-sided hypothesis $H_1 : \theta < 0$ vs. $H_2 : \theta > 0$ is considered when the risk is componentwise. Recall that the fixed sample size t-test based on a sample of M is to reject if

$$t = ((M-1)/M)^{\frac{1}{2}} S_M / \left[\Sigma_{i=1}^M (X_i - \bar{X})^2\right]^{\frac{1}{2}} > C. \tag{3.1}$$

For now let $C > 0$. Then (3.1) is equivalent to rejecting if $S_M > 0$, and

$$S_M^2 / (\Sigma_{i=1}^M X_i^2) > MC^2 / [(M-1) + C^2] = K. \tag{3.2}$$

We start by proving

LEMMA 3.1. *If $S_{M-1} \leqslant 0$, then $S_M/(\Sigma_{i=1}^M X_i^2)^{\frac{1}{2}} \leqslant 1$ for all $X_M > 0$.*

PROOF. Note that for fixed S_{M-1}, $S_{M-1} < 0$,

$$\begin{aligned} S_M / (\Sigma_{i=1}^M X_i^2)^{\frac{1}{2}} &= (S_{M-1} + X_M) / (\Sigma_{i=1}^{M-1} X_i^2 + X_M^2)^{\frac{1}{2}} \\ &\leqslant X_M / (\Sigma_{i=1}^{M-1} X_i^2 + X_M^2)^{\frac{1}{2}} \leqslant 1. \end{aligned} \tag{3.3}$$

Thus the lemma is proved.

Now we can prove

THEOREM 3.1. *The fixed sample size test of* (3.1) *is inadmissible for $M \geqslant 2$ provided $C \geqslant 1$.*

PROOF. Consider the test procedure which is the same as (3.1) except that when S_{M-1} is negative, the procedure stops and accepts. By Lemma 3.1 this new procedure chooses the identical terminal decision as the t-test and therefore has the same probabilities of error. However, the expected sample size is always less for the new procedure. This completes the proof of the theorem.

REMARK 3.1. The values of $C \geqslant 1$ cover most α levels of interest.

If $C \leqslant -1$, essentially the same argument given in Theorem 3.1 proves that the fixed sample size t-test is also inadmissible. For values of $|C| > 1$, similar arguments can be used to prove that the two sided t-test is inadmissible.

REMARK 3.2. Let X_i be multivariate normal with unknown mean vector $\boldsymbol{\theta}$ and unknown covariance matrix $\boldsymbol{\Sigma}$. Test $H_0 : \boldsymbol{\theta} = 0$ vs. $H_1 : \boldsymbol{\theta} \neq 0$ by the fixed sample Hotelling's T^2-test. The analogue of Theorem 3.1 is true. This can be shown by using the fact that Hotelling's T^2-test is derivable by the union-intersection procedure.

Next we prove

THEOREM 3.2. *The test in* (3.1) *is admissible provided* $|C| < 1$.

PROOF. For now assume $C > 0$ and let $\delta'(\mathbf{x})$ be a test which can potentially be better than the t-test. The stopping rule for $\delta'(\mathbf{x})$ must assign positive probability to stopping at some time sooner than M. Otherwise the expected sample size would be greater than or equal to M. If it were equal, without ever stopping sooner, $\delta'(\mathbf{x})$ could not beat t since t is known to be an admissible test among fixed sample size tests. (The t-test is UMPU.) Hence for $\delta'(\mathbf{x})$, there must exist a set E contained in

E^j, for some $j = 1, 2, \cdots, M-1$, $\mu(E) > 0$, such that $(\Pi_{k=1}^{j-1}\delta'_{0k}(x_1, x_2, \cdots, x_k))\delta'_{ij}(x_1, x_2, \cdots, x_j) = \varepsilon > 0$ for $i = 1$ or 2. For now let $i = 1$. For any point in E, note that the limit, as one X_k tends to ∞, $k = j + 1, \cdots, M$, of the quantity $(S_j + \Sigma_{k=j+1}^M X_k)/(\Sigma_{k=1}^j X_k^2 + \Sigma_{k=j+1}^M X_k^2)^{\frac{1}{2}}$, is one. This implies that there exists a set F contained in E^{M-j} such that $\mu(E \times F) > 0$ and such that for every $(x_1, x_2, \ldots, x_M)$ in $E \times F$, $t > C$ but δ' accepts H_1 with positive probability. At this point it is clear that the Stein argument used in Theorem 2.3 shows that the probability of the type II error for δ', will exceed the probability of the type II error for the t-test for some (θ, σ^2) points in the alternative space. Thus, such a δ' cannot beat t. Similar arguments work if δ' rejects instead of accepts sooner than M, and also if $C < 0$. This completes the proof of the theorem.

In the remainder of this section we demonstrate how to construct admissible tests which are better than the fixed sample size t-test. The better test will be determined by the critical value of the given t-test. We will need

LEMMA 3.2. *If* $\{S_{M-1}/[\Sigma_{i=1}^{M-1}X_i^2]^{\frac{1}{2}}\} = B$, *then the range of* $t(X_M) = \{(S_{M-1} + X_M)/[\Sigma_{i=1}^{M-1}X_i^2 + X_M^2]^{\frac{1}{2}}\}$ *is*

$$\left(-1, (B^2+1)^{\frac{1}{2}}\right] \quad \text{if } B > 0, \tag{3.4}$$

$$\left[-(B^2+1)^{\frac{1}{2}}, 1\right) \quad \text{if } B < 0. \tag{3.5}$$

PROOF. Consider $B > 0$. Note that $(dt/dX_M) > 0$ for $X_M < 0$. Hence $\inf_{X_M<0} t(X_M) = \lim_{X_M\to-\infty} t(X_M) = -1$. Since $\inf_{X_M>0} t(X_M) > 0$, it follows that the inf $t(X_M) = -1$ and the infimum is not attained. Now set $(dt/dX_M) = 0$ and find that the attainable global maximum of $t(X_M)$ is $(B^2+1)^{\frac{1}{2}}$. The proof for $B < 0$ is similar. This completes the proof of the lemma.

The lemma is the key to understanding the results of this section, since it gives precise bounds on the behavior of the statistic $\Sigma_{i=1}^M X_i/(\Sigma_{i=1}^M X_i^2)^{\frac{1}{2}}$ as M changes.

We first find admissible procedures which dominate the t-test for critical values in the range $1 \leqslant C < [2(M-1)/(M-2)]^{\frac{1}{2}}$ (which corresponds by (3.2) to $1 \leqslant K < 2$).

Let $\delta^{(1)}(\mathbf{x})$ be the test which stops at time $(M-1)$ and accepts H_0 if $[S_{M-1}/(\Sigma_{i=1}^{M-1}X_i^2)^{\frac{1}{2}} < (K-1)^{\frac{1}{2}}]$. Otherwise $\delta^{(1)}(\mathbf{x})$ stops at time M and accepts or rejects depending on whether $t < C$ or $t > C$. We prove

THEOREM 3.3. *Let* $M \geqslant 3$. *If* $1 < C < (2(M-1))^{\frac{1}{2}}/(M-2)^{\frac{1}{2}}$, *then* $\delta^{(1)}(\mathbf{x})$ *is better than the fixed sample size t-test and is admissible.*

PROOF. That $\delta^{(1)}(\mathbf{x})$ is better can be proved as in Theorem 3.1. Now suppose $\tilde{\delta}(\mathbf{x})$ is better than $\delta^{(1)}(\mathbf{x})$. Clearly $\tilde{\delta}(\mathbf{x})$ must stop with positive probability sometime before stage M. Suppose $\tilde{\delta}(\mathbf{x})$ stops at any stage before stage M and rejects. For fixed $(X_1, X_2, \cdots, X_{M-1})$, $\lim_{X_M\to-\infty} S_M/(\Sigma_{i=1}^M X_i^2)^{\frac{1}{2}} < 0$, which means that there will be a set of points of positive Lebesgue measure in E^M, which $\tilde{\delta}(\mathbf{x})$ will reject but $\delta^{(1)}(\mathbf{x})$ will accept. By the Stein argument this implies that the type I error for $\tilde{\delta}(\mathbf{x})$ will exceed the type I error for $\delta^{(1)}(\mathbf{x})$ for some (θ, σ^2) points in the parameter space. Suppose next that $\tilde{\delta}(\mathbf{x})$ stops for the first time at stage $(M-1)$ and accepts on a set of positive measure in $E^{(M-1)}$ for which $[S_{M-1}/(\Sigma_{i=1}^{M-1}X_i^2)^{\frac{1}{2}}] > (K-1)^{\frac{1}{2}}$. This would imply by Lemma 3.1 that there would be a set of positive measure in E^M for which $\tilde{\delta}(\mathbf{x})$ would accept but $\delta^{(1)}(\mathbf{x})$ would reject since for these points in $E^{(M-1)}$, $\sup_{X_M}[S_M/(\Sigma_{i=1}^{M}X_i^2)^{\frac{1}{2}}] > K^{\frac{1}{2}}$. The Stein argument would then yield that the type II error for $\tilde{\delta}(\mathbf{x})$ would exceed the type II error for $\delta^{(1)}(\mathbf{x})$ for some (θ, σ^2). Finally suppose $\tilde{\delta}(\mathbf{x})$ stops and accepts before stage $(M-1)$. It is easily seen, as in Lemma 3.2, that if $\tilde{\delta}(\mathbf{x})$ stops at stage j then

(3.6)
$$\sup_{(X_{j+1},\cdots,X_M)}\left[\left(S_j + X_{j+1} + \cdots + X_M\right)/\left(\Sigma_{i=1}^{j}X_i^2 + \Sigma_{i=j+1}^{M}X_i^2\right)^{\frac{1}{2}}\right] \geqslant (M-j)^{\frac{1}{2}}.$$

Since $(M-j)^{\frac{1}{2}} > K^{\frac{1}{2}}$, for $j = 1, 2, \cdots, (M-2)$ this means again that $\tilde{\delta}(\mathbf{x})$ would be accepting for points in E^M for which $\delta^{(1)}(\mathbf{x})$ would be rejecting. Again $\tilde{\delta}(\mathbf{x})$ would have a type II error larger than that for $\delta^{(1)}(\mathbf{x})$ for some (θ, σ^2) values. The above implies that $\tilde{\delta}(\mathbf{x})$ cannot be better than $\delta^{(1)}(\mathbf{x})$, which completes the proof of the theorem.

One can determine an admissible procedure $\delta^{(1)}(\mathbf{x})$ which would improve on the fixed sample t for any critical value $|C| > 1$. For now let $C > 0$. First determine the integer $k = 2, 3, \cdots,$ for which the interval

$$\left\{\left[(M-1)(k-1)/(M-k+1)\right]^{\frac{1}{2}}, \left[(M-1)k/(M-k)\right]^{\frac{1}{2}}\right\} \tag{3.7}$$

contains C. Note from (3.2) that $(k-1) < K \leqslant k$. Define $\delta^{(1)}(\mathbf{x})$ as follows: for $r = k$, stop, and accept if

$$\left[S_{M+1-r}/\left(\Sigma_{i=1}^{M+1-r}X_i^2\right)^{\frac{1}{2}}\right] < (K+1-r)^{\frac{1}{2}}. \tag{3.8}$$

Otherwise continue and stop and accept if for $r = k-1$, (3.8) holds. Repeat this process for $r = k-2, \cdots, 1$. When $r = 1$, stop and accept if (3.8) holds, but reject if (3.8) does not hold. Provided $M \geqslant (k+1)$, the proof of Theorem 3.3 can be extended to yield

THEOREM 3.4. *For $M \geqslant (k+1)$, $\delta^{(1)}(\mathbf{x})$ defined in* (3.8) *is admissible and is better than the fixed sample size t-test.*

4. Discussion of risk functions. There are four different risk functions that have been used to evaluate sequential tests of hypotheses. The four have an ordering from weakest to strongest. Each of the four considers probability of error as part of the risk function. In previous sections we have discussed the strongest of the four which is a linear combination of probability of error and expected sample size. We have also considered the next strongest which is componentwise with expected sample size as a component. The weakest of the four has been considered by Eisenberg, Ghosh and Simons [3]. The componentwise risk in this case includes the random sample size (almost everywhere) as opposed to expected sample size. Kiefer and Weiss [4] consider the probability distribution of sample size as a component and define one test to be better than the other on this count if the sample size distributions are stochastically ordered for all parameter points. Our results provide results for these other loss functions in some cases. For example, the proof of Theorem 3.1 can be used to establish inadmissibility of the fixed sample size t-test when $|C| \geqslant 1$ for all four loss functions. Theorem 3.2 establishes admissibility for $|C| < 1$ for three of the four loss functions. It would be desirable to further study properties of tests for each of these risk functions.

Brown, Cohen and Strawderman [2] prove that if the risk is linear combination (the strongest of the four mentioned above) then under certain conditions all Bayes tests are monotone. A monotone test for a one-sided hypothesis being one which, at stage j, for sufficient statistic S_j, stops and accepts for $S_j < a_j$, continues for $a_j \leqslant S_j \leqslant b_j$, and stops and rejects for $S_j > b_j$. We conclude this section with an example of a nonmonotone procedure which cannot be beaten by a monotone procedure when the risk is componentwise, with one component being expected sample size. All other conditions of the example satisfy the conditions required in the Brown, Cohen and Strawderman theorem.

EXAMPLE 4.1. Let X_i, $i = 1, 2, \cdots$ be independent, identically distributed normal variables with mean θ and variance 1. The hypothesis is $H_1 : \theta < 0$ vs. $H_2 : \theta > 0$. Let $\delta(\mathbf{x})$ be the test which stops and accepts at stage 1 if $X_1 \leqslant 0$, stops and rejects if $0 < X_1 \leqslant 1$, continues if $X_1 > 1$. Then $\delta(\mathbf{x})$ stops next at stage 2 and accepts or rejects as $\bar{X}_2 < 0$ or $\bar{X}_2 > 0$. Clearly $\delta(\mathbf{x})$ is a nonmonotone procedure.

Suppose $\delta'(\mathbf{x})$ is a monotone competitor to $\delta(\mathbf{x})$. In order for $\delta'(\mathbf{x})$ to compete on the component $E_\theta n_\delta$, $\delta'(\mathbf{x})$ must stop at stage 1 for all $X_1 \leqslant 1$. For if it did not $\lim_{\theta \to -\infty}(E_\theta n_{\delta'} - E_\theta n_\delta) > 0$. To see this, note that $E_\theta n_\delta = 1 + P_\theta(X_1 > 1)$, whereas $E_\theta n_{\delta'} \geqslant 1 + P_\theta$ (continuing). If there is an X_1 set of positive measure whose elements are less than 1, for which $\delta'(\mathbf{x})$ does not stop then the Stein type argument shows that the probability of continuing for δ goes to zero faster than for δ' as $\theta \to -\infty$. Hence δ' must stop for all $X_1 \leqslant 1$.

Next observe that δ' cannot stop and accept for $X_1 < C$ where $C > 0$. This is because as $\theta \to \infty$ the type II error for δ' would exceed the type II error for δ. To see this, note

$$(4.1) \quad P_\theta(\text{Accepting } H_1 \text{ under } \delta') - P_\theta(\text{Accepting } H_1 \text{ under } \delta) \geqslant \left[\Phi(C-\theta) - \Phi(-\theta) - \Phi(-2^{\frac{1}{2}}\theta)\right].$$

It is easily seen that as θ tends to infinity the right hand side of (4.1) approaches zero from above and thus is positive for large θ.

Thus δ', to preserve monotonicity, must stop and accept if $X_1 \leqslant C' \leqslant 0$ and reject for $X_1 > C'$. But again it is easily seen that as $\theta \to -\infty$, the type I error for δ' would exceed the type I error for δ. Thus we have shown that there does not exist a monotone test which is better than δ.

Acknowledgment. The authors are grateful to Professor Robert H. Berk for suggesting this area of study and for many helpful discussions.

REFERENCES

[1] Birnbaum, A. (1955). Characterizations of complete classes of tests of some multiparameter hypotheses with applications to likelihood ratio tests. *Ann. Math. Statist.* **26** 21–36.

[2] Brown, L. D., Cohen, Arthur and Strawderman, W. E. (1979). Monotonicity of Bayes sequential tests. To appear in *Ann. Statist.*

[3] Eisenberg, B., Ghosh, B. K. and Simons, G. (1976). Properties of generalized sequential probability ratio tests. *Ann. Statist.* **4** 237–251.

[4] Kiefer, J. and Weiss, L. (1957). Some properties of generalized sequential probability ratio tests. *Ann. Math. Statist.* **28**, 57–74.

[5] Stein, C. (1956). The admissibility of Hotelling's T^2-test. *Ann. Math. Statist.* **27** 616–623.

Department of Statistics
Rutgers University
New Brunswick, N.J. 08903