The Annals of Statistics
1980, Vol. 8, No. 2, 377–398

COMPLETE CLASSES FOR SEQUENTIAL TESTS OF HYPOTHESES[1]

By L. D. Brown, Arthur Cohen, and W. E. Strawderman

Rutgers University

We consider problems of sequential testing when the loss function is the sum of a component due to an error in the terminal decision and a cost of observation component. In all cases we establish a characterization of a complete class or an essentially complete class. In order to obtain such results for testing a null hypothesis against an alternative hypothesis we establish complete class results for testing the closure of the null hypothesis against the closure of the alternative hypothesis. A complete class for testing closure of null against closure of alternative is an essentially complete class for testing null against alternative. Furthermore, a complete class for testing closure of null against closure of alternative is a complete class for testing null against alternative when the risks have certain continuity properties. Such continuity properties do hold in many cases.

Three models are treated. The first is when the closure of the null space is compact and the cost of the first observation is positive. Under very unrestrictive conditions it is shown that the Bayes tests form a complete class. This result differs considerably from most fixed sample analogues that have been studied.

The second model is when the closure of the null space is compact, the distributions are exponential family, and the cost of the first observation is zero. The third model is for the one dimensional exponential family case when the hypotheses are one sided.

1. Introduction and summary. We consider problems of sequential testing when the loss function is the sum of a component due to an error in the terminal decision and a cost of observation component. In all cases we establish a characterization of a complete class or an essentially complete class of tests. In order to obtain such results for testing a null hypothesis against an alternative hypothesis, we establish complete class results for testing the closure of the null hypothesis against the closure of the alternative hypothesis. For the problems treated here, a complete class for testing closure of null against closure of alternative is an essentially complete class for testing null against alternative. The results summarized below refer to testing the closure of null against closure of alternative. At the end of this section, some further remarks will be made about the distinction and importance of the two problems.

The first model treated is when the closure of the null hypothesis is compact and the cost of the first and all other observations is positive. Under very mild conditions it is proven that the Bayes tests form a complete class. The only

Received March 1978; revised November 1978.

[1]Research supported by National Science Foundation Grant No. MCS-75-23343

AMS 1970 *subject classifications.* Primary 62L10; secondary 62C07.

Key words and phrases. Sequential tests, complete classes, Bayes tests, generalized Bayes tests, admissibility, stopping rules.

assumptions required are that the joint density of the observations at every stage is, for almost every fixed value of the random vector, a well defined continuous function of the parameter on the closure of the null and closure of alternative spaces, and tends to zero as the norm of the parameter tends to infinity. We also assume here and in subsequent problems that the total cost of observations tends to infinity as the sample size tends to infinity. Multivariate and multiparameter problems are included in the treatment.

In cases where the Bayes tests are nonrandomized in both the stopping rule and terminal decision rule, the above complete class result implies that randomization should be eliminated. This will in fact be the case for problems involving the normal distribution with unknown mean vector and known covariance matrix. (More precisely, randomization must be allowed when no observations have yet been taken, but not thereafter.) Furthermore, in these cases it will follow that procedures which are not based only on sufficient statistics can be improved on.

The second model treated is when the observations are from an exponential family dominated by a measure which is absolutely continuous with respect to Lebesgue measure. The cost of the first observation is zero, although the cost of all other observations is positive. Again the closure of the null hypothesis space is assumed compact. If the closures of the null and alternative spaces are contained in the natural parameter space then each test in the complete class is as follows: There is a convex set A, such that if the first observation falls outside A, reject the null hypothesis. If the first observation falls inside A, then proceed according to a generalized Bayes test of the closure of the null vs. the closure of the alternative, where the generalized prior distribution is such that the integrated risk of the test is finite. Again under the normality assumption, randomization can be eliminated. If the closure of the null space is in the natural parameter space but the closure of the alternative space is not, then a similar characterization (not quite as above) will be given for an essentially complete class.

The third model treated is when the observations are from a one dimensional exponential family dominated by a nonatomic probability measure and the hypotheses are one sided and not necessarily bounded. (Precise assumptions will be given.) For this model we treat both the cases where the cost of the first observation is positive and where the cost of the first observation is zero. With the exception of what happens at stage zero, the results for the two cases are similar when both null and alternative spaces are unbounded. For the case where the cost of the first observation is zero and the closures of both null and alternative spaces are contained in the natural parameter space, the complete class is as follows: There exists an interval (a_1, a_2) such that if the first observation is less than a_1, stop and accept. If the first observation is greater than a_2, stop and reject. If the observation is in the interior of the interval, then proceed according to a generalized Bayes test of the closure of the null vs. the closure of the alternative. Randomization can be eliminated in all cases of this one dimensional, one sided model, where the cost of the first observation is zero and the dominating measure is nonatomic.

For the case where either closure of the null hypothesis space or closure of the alternative hypothesis space, but not both, are contained in the natural parameter space, a similar characterization will be given for an essentially complete class.

For this third model, it is known that an essentially complete class of tests consists of the monotone tests. See Sobel [7] and Brown, Cohen, and Strawderman [2]. The complete class which we describe here is much smaller than the class of monotone tests.

There are two important respects in which our results for the sequential problem are qualitatively different and superior to analogous results for the fixed sample size case. Farrell [4] studied the fixed sample size case. First, he required that the distributions come from an exponential family. In our first model, where the cost of the first observation is positive, we do not need such an assumption. Furthermore the Bayes procedures, rather than the generalized Bayes procedures, form the complete class. Second, he required that the closure of the null space be topologically separated from the closure of the alternative; and he gave an example to show why this assumption is needed in order to obtain a useful complete class theorem of this type. We do not make such an assumption.

At this point we will clarify the distinction between the fixed sample size case and the first model of the sequential case. In both cases the characterization is developed by describing the test procedure which is the limit (in some sense) of a sequence of Bayes tests. This is because it is known that such limits form an essentially complete class. Suppose $(\pi_{1k}, \Gamma_{1k}(\cdot), \Gamma_{2k}(\cdot))$ represents a sequence of prior probability distributions on the parameter space, where π_{1k} represents the probability that the parameter lies in the null space, $\Gamma_{1k}(\cdot)$ represents the conditional probability measure for the parameter given that it lies in the null space, and $\Gamma_{2k}(\cdot)$ is similarly defined for the alternative space. (It will be helpful at this point if the reader keeps in mind the concrete example of testing a one dimensional normal mean μ when the variance is one and the null hypothesis is $|\mu| \leqslant 1$, while the alternative is $|\mu| > 1$.) Now in the sequential case, if $\pi_{1k} \to 0$ as $k \to \infty$, the procedure which is the limit of the sequence of Bayes procedures stops at stage 0 (before taking any observations) and rejects the null hypothesis. This happens as soon as $\pi_{1k} < c_1$, the cost of the first observation. This fact is true regardless of what Γ_{1k} or Γ_{2k} do, and doesn't depend in any way on what the distribution of the observable random variables is. This latter fact explains why we need not assume an exponential family as does Farrell. In the fixed sample size case however, if $\pi_{1k} \to 0$ as $k \to \infty$, $\Gamma_{2k}(\cdot)$ may be gradually assigning its mass to larger values of the parameter (think of mass going to $\mu = \infty$ in the normal example) in such a way or at a certain rate, so that the limiting procedure does not necessarily always reject. In fact the sequence of ratios that determine the sequence of Bayes procedures may in fact be indeterminate in the form $0/0$. In such cases the limiting procedure depends on the distribution of the random variables (hence the exponential family assumption by Farrell.) Furthermore the limiting procedure need not be

Bayes, explaining the need for the introduction of non-Bayes procedures such as generalized Bayes or other non-Bayes procedures. Still further, these non-Bayes procedures can very well be admissible and must be contained in any complete class.

Next suppose that $\pi_{1k} \equiv \frac{1}{2}$ and that Γ_{1k} and Γ_{2k} send all their mass to a common boundary point. (In the normal example Γ_{1k} could put mass 1 at $\mu = 1 - 1/k$, and Γ_{2k} could put all its mass at $\mu = 1 + 1/k$.) The resulting limiting procedure in the sequential case would be a Bayes procedure for testing the closure of the null against the closure of the alternative when the prior puts probability $\frac{1}{2}$ on the common boundary point as a point in the closure of the null, and probability $\frac{1}{2}$ on the common boundary point as a point in the closure of the alternative. This resulting Bayes procedure would stop at stage zero and randomize between acceptance and rejection of the null hypothesis. This is because the first and future observations would incur costs while the posterior probabilities of error would remain at $\frac{1}{2}$ regardless of additional data points. This type of sequence of priors and others that send mass to the common boundary explain why we need to study the problem of testing the closure of the null hypothesis against the closure of the alternative.

Notice that in the fixed sample size case if one tests closure of null against closure of alternative and a prior puts all its mass on a common boundary point, equally as part of the null and as part of the alternative, then all procedures are Bayes against such a prior. Such a characterization is useless. This explains why Farrell assumes for the fixed sample size case that the closure of the null and closure of the alternative spaces are topologically separated. An example is given in Farrell where closure of null and closure of alternative are not topologically separated and one tests null against alternative with the result that all procedures are limits of Bayes procedures. Again this is a useless characterization.

To understand the above distinctions between the fixed sample size and sequential cases, and to understand the future development of the sequential case, and the difference in the results for the various sequential models, it may be helpful to think of a two stage procedure instead of a purely sequential procedure. The distinctions are true for the two stage case. For the first model the two stage case consists of stage 0, no observations, and stage 1, a choice of one observation at cost $c_1 > 0$. For the second model, the two stage case consists of stage 1, one observation at no cost which must be taken, and stage 2, where there is a choice of a second observation at a cost $c_2 > 0$. In the beginning of Section 4 we discuss the distinction between the first model, where $c_1 > 0$ and the second model, where $c_1 = 0$.

We now discuss further the significance of obtaining complete class results for the problem of testing the closure of the null against the closure of the alternative. As already mentioned, a complete class result for such a problem yields an essentially complete class for the problem of testing null against alternative. Furthermore, and most importantly, if the two risk functions, (one for parameter values in the closure of the null space and the second for parameter values in the closure of the alternative space) are continuous at parameter points in the common boundary, then the complete class for closure of null against closure of alternative is a complete class for null against alternative. In a subsequent paper, such a continuity property will be shown to hold under very general conditions. Of course when null and alternative are closed to begin with, then the complete class is obtained without further requirements. Also for truncated sequential problems, the desired continuity of the risks is valid under the conditions given in this paper and the complete class is obtained.

In some instances we find an essentially complete class for the closure of the null against the closure of the alternative. This class is also an essentially complete class for testing null against alternative.

In the next section we need preliminaries and definitions. In Section 3 through 5 we treat each of the three models in the order given above.

2. Definitions and preliminaries. The elements of the problem are as follows: Θ is the parameter space with typical element θ. The null space is $\Theta_1 \subset \Theta$ and the alternative space is $\Theta_2 \subset \Theta$. Assume Θ, Θ_1, and Θ_2 are measurable subsets of Euclidean space, R^q. The closures of Θ_1 and Θ_2 are denoted respectively as $\overline{\Theta}_1$ and $\overline{\Theta}_2$. We study the problem of testing the closure of the null hypothesis, $\theta \in \overline{\Theta}_1$, vs. the closure of the alternative hypothesis, $\theta \in \overline{\Theta}_2$. The reader may find it helpful to denote the parameter set of the null hypothesis by $\{(1, \theta)\}$ for $\theta \in \overline{\Theta}_1$ and the parameter set of the alternative by $\{(2, \theta)\}$ for $\theta \in \overline{\Theta}_2$. Then it will be easy to view the null hypothesis and alternative hypothesis sets as disjoint as is sometimes desired.

The action space $\mathcal{A}$ consists of pairs (n, τ) where n is the stopping time and τ is 1 or 2, depending on whether $\overline{\Theta}_1$ is accepted or rejected. The loss function, denoted by $L(\theta, (n, \tau))$ is given by

$$L_1(\theta, (n, \tau)) = C(n) + \begin{cases} 0 & \text{if } \tau = 1 \\ d_1 & \text{if } \tau = 2 \end{cases}, \quad \text{when } \theta \in \overline{\Theta}_1 \tag{2.1}$$

and

$$L_2(\theta, (n, \tau)) = C(n) + \begin{cases} 0 & \text{if } \tau = 2 \\ d_2 & \text{if } \tau = 1 \end{cases}, \quad \text{when } \theta \in \overline{\Theta}_2. \tag{2.2}$$

Equivalently, $L(\theta, (n, \tau)) = C(n)$ if $\theta \in \{(1, \theta)\}$, $\tau = 1$ or $\theta \in \{(2, \theta)\}$, $\tau = 2$; $L(\theta, (n, 2)) = C(n) + d_1$ if $\theta \in \{(1, \theta)\}$, and $L(\theta, (n, 1)) = C(n) + d_2$ if $\theta \in \{(2, \theta)\}$. Here $C(n)$ represents the cost of taking n observations. Let c_j represent the cost of taking the jth observation so that $C(n) = \sum_{j=1}^{n} c_j$. Also assume $c_j > 0$, all $j = 2, 3, \cdots$, and $C(n) \to \infty$ an $n \to \infty$.

The observation available to the statistician at stage i is an observation on a $p \times 1$ vector. This observation is denoted by x_i. We let $\bar{\mathbf{x}} = (x_1, x_2, \cdots)$ denote the $(p \times \infty)$ matrix of observations and $x_{(j)} = (x_1, x_2, \cdots, x_j)$. The corresponding random variables (measurable mappings) are denoted X_i, $\overline{\mathbf{X}}$, and $X_{(j)}$ respectively. We assume that there is a σ-finite measure μ which dominates the family $\{P_\theta(\cdot), \theta \in \Theta\}$ of probability measures for $\overline{\mathbf{X}}$ in the following sense: For each $j = 1, 2, \cdots$, over the σ-field generated by $X_{(j)}$, the measure P_θ is dominated by μ. Write $f_\theta^{(j)}(x_{(j)}) = dP_\theta/d\mu$ relative to this σ-field. When Θ is not closed we will usually, wish to assume that the family, $\{P_\theta, \theta \in \Theta\}$ can be extended to a family, $\{P_\theta, \theta \in \overline{\Theta}\}$ and that the families of densities $f_\theta^{(j)}(x_{(j)}), \theta \in \overline{\Theta}$ exist and have certain continuity properties, etc. to be specified later.

We will have occasion to state assumptions regarding $x_{(j)}$'s. When we write, for all $x_{(j)}$ or for every $x_{(j)}$ we mean for almost all or almost every, meaning, every $x_{(j)}$ except perhaps some $x_{(j)}$ which comprise a set of μ-measure zero.

When observations are independent and identically distributed from an exponential family, there is a sequence $\{S_i\}$, where S_i is a function of $X_{(i)}$ such that $\{S_i\}$ is a sufficient, transitive sequence for θ. (See Ferguson, page 334) for definition of transitive.) Write $S = (S_1, S_2, \cdots)$ so that S lies in an infinite product space. We define μ, P_θ, $f_\theta^{(j)}(s_{(j)})$ for this sample space as we did for the $\overline{\mathbf{X}}$ sample space.

A prior probability measure on $\Theta(\overline{\Theta})$, denoted by $\Gamma(\cdot)$, will be represented by a mixture expressed as $\pi_1\Gamma_1(\cdot) + \pi_1\Gamma_2(\cdot)$. Here, if T is a random variable with distribution Γ, then π_1 is the probability that $T \in \Theta_1(\overline{\Theta}_1)$ and Γ_1 represents the conditional distribution of T, given $T \in \Theta_1(\overline{\Theta}_1)$. Similarly for Γ_2. A prior $\Gamma(\cdot)$ may be represented as $(\pi_1, \Gamma_1(\cdot), \Gamma_2(\cdot))$.

A decision function δ may be expressed as a set of nonnegative functions $\delta_{ij}(x_{(j)} \geqslant 0$, $(i = 0, 1, 2; j = 1, 2, \cdots)$ defined for all $x_{(j)}$ such that $\Sigma_{i=0}^2\delta_{ij}(x_{(j)}) = 1$. The quantities $\delta_{ij}(x_{(j)})$, $i = 0, 1, 2$, represent respectively, the probability of taking another observation, accepting H_1, and accepting H_2 when j observations, $x_{(j)}$, have been taken. This definition of decision function for sequential tests is given in Sobel [7]. It is equivalent to the usual definition.

The risk function is denoted by $R(\theta, \delta) = E_\theta L(\theta, \delta)$ and the expected risk is $R(\Gamma, \delta) = ER(\theta, \delta)$. A Bayes procedure minimizes $ER(\theta, \delta)$. A generalized prior distribution $\Gamma(\cdot)$ has the properties of a prior distribution except that $\Gamma(\Theta)$ can be infinite.

Note that when $\tilde{\Theta}$ is a set in $\overline{\Theta}_1 \cap \overline{\Theta}_2$, then any prior $(\pi_1, \Gamma_1(\cdot), \Gamma_2(\cdot))$ specifies the amount of probability assigned to that set as part of $\overline{\Theta}_1$ or as part of $\overline{\Theta}_2$. For example, if $0 \in \overline{\Theta}_1 \cap \overline{\Theta}_2$, then $(1, \varepsilon\{0\}, \Gamma_2(\cdot))$ is distinguishable from $(\frac{1}{2}, \varepsilon\{0\}, \varepsilon\{0\})$, where $\varepsilon\{a\}$ is a degenerate probability distribution, all of whose mass is put at the point a. The latter prior assigns probability $\frac{1}{2}$ to 0 as a point in $\overline{\Theta}_1$, and probability $\frac{1}{2}$ to 0 as a point in $\overline{\Theta}_2$. In the alternative notation the latter prior assigns probability $\frac{1}{2}$ to parameter points (1, 0) and (2, 0).

We will refer to sequences of tests which converge regularly to a limit. For the definition of regular convergence we refer to Sobel [7], page 321. It is well known that tests which converge a.e., converge regularly. It is also well known that the closure of the class of Bayes tests with respect to regular convergence forms an essentially complete class. (See Le Cam [6].)

Some final notations and definitions. Let

$$(2.3)\qquad f_{(i)}^{(j)}(x_{(j)}) = \int_{\bar{\Theta}_i} f_\theta^{(j)}(x_{(j)})\Gamma_i(d\theta), \quad i = 1, 2; j = 1, 2, \cdots.$$

Note $f_{(i)}^{(j)}(x_{(j)})$ are densities equal to $dP_{(i)}/d\mu$ over the σ-field generated by $X_{(j)}$, where $P_{(i)}(\cdot) = \int_{\bar{\Theta}_i} P_\theta(\cdot)\Gamma_i(d\theta)$. Clearly $P_{(1)}(\cdot), P_{(2)}(\cdot)$ are dominated by $\nu(\cdot) = \pi_1 P_{(1)}(\cdot) + \pi_2 P_{(2)}(\cdot)$, and ν represents a marginal probability law for $\bar{\mathbf{X}}$. Let $\nu^{(n+1)|n}(\cdot|x_{(n)})$ denote the conditional distribution under ν of $x_{(n+1)}$ given $X_{(n)} = x_{(n)}$. Further let $g_{(i)}^{(j)}(x_{(j)}) = dP_{(i)}/d\nu, j = 1, 2, \cdots$, over the σ-field generated by $X_{(j)}$. It can be shown, (see [2], page 5) that

$$(2.4)\qquad g_{(i)}^{(n)}(x_{(n)}) = f_{(i)}^{(n)}(x_{(n)})/\left[\pi_1 f_{(1)}^{(n)}(x_{(n)}) + \pi_2 f_{(2)}^{(n)}(x_{(n)})\right].$$

Note that $\pi_1 g_{(1)}^{(j)} + \pi_2 g_{(2)}^{(j)} \equiv 1$. For the process defined by $P_{(i)}$, the conditional density of $X_{(n+1)}$ given $X_1 = x_1, X_2 = x_2 \cdots, X_n = x_n$ relative to $\nu^{(n+1)|n}(\cdot|x_{(n)})$ is given by $g_{(i)}^{(n+1)|n}(X_{(n+1)}|x_{(n)}) = g_{(i)}^{(n+1)}(X_{(n+1)})/g_{(i)}^{(n)}(x_{(n)})$.

3. Complete class for $\bar{\Theta}_1$ compact and $c_1 > 0$. In this section we assume $\bar{\Theta}_1$ is compact, $c_1 > 0$, for every j, $\theta \in \bar{\Theta}_1$, $\theta \in \bar{\Theta}_2$, $f_\theta^{(j)}(x_{(j)})$ is a well defined density, for every j and $x_{(j)}$, $f_\theta^{(j)}(x_{(j)})$ is continuous on $\bar{\Theta}_1$ and on $\bar{\Theta}_2$, and for every j and $x_{(j)}$, $f_\theta^{(j)}(x_{(j)}) \to 0$ as $\|\theta\| \to \infty$. This last requirement is needed for weak convergence arguments, as, for example, in Lemma 3.1. We obtain an essentially complete class of tests for Θ_1 vs. Θ_2 by obtaining a complete class of tests for $\bar{\Theta}_1$ vs. $\bar{\Theta}_2$ when the loss function is given by (2.1) and (2.2).

Let $(\pi_{1k}, \Gamma_{1k}(\cdot), \Gamma_{2k}(\cdot))$ be a sequence of prior distributions. Let $\delta(k; \bar{\mathbf{x}})$ be the Bayes tests for these priors and assume that the regular limit as $k \to \infty$ exists. Denote this limit by $\delta(\bar{\mathbf{x}})$. The existence of Bayes tests is a consequence of compactness of the space of decision rules. (See Le Cam [6].) One objective of this section is to show that δ is Bayes. To accomplish this objective we will consider the sequential testing problem truncated at M. That is, we must stop no later than at stage M. For this problem, if $(\pi_{1k}, \Gamma_{1k}(\cdot), \Gamma_{2k}(\cdot))$ is a prior distribution, we let $\beta_{n,k}^M(x_{(n)})$ be the minimum conditional expected risk given $X_{(n)}$ is observed at stage n and sampling continues at least to stage $(n + 1)$. For $n = 0, 1, 2, \cdots, (M - 2)$, use (2.3), (2.4) and Fubini's theorem to find

$$(3.1)\qquad \beta_{n,k}^M(x_{(n)}) = \int \min\left[\beta_{n+1,k}^M(x_{(n+1)}), C(n+1) + d_2\pi_{2k} g_{(2),k}^{(n+1)}(x_{(n+1)}),\right.$$
$$\left. C(n+1) + d_1\pi_{1k} g_{(1),k}^{(n+1)}(x_{(n+1)})\right]\nu^{(n+1)|n}(dx_{(n+1)}|x_{(n)}).$$

For $n = M - 1$,

$$(3.2)\qquad \beta_{M-1,k}^M(x_{(M-1)}) = C(M) + \int \min\left[d_2\pi_{2k} g_{(2),k}^{(M)}(x_{(M)}), d_1\pi_{1k} g_{(1),k}^{(M)}(x_{(M)})\right]$$
$$\cdot \nu^{M|(M-1)}(dx_{(M)}|x_{(M-1)}).$$

Note that $\min[\beta_{0,k}^M, \pi_{1k}d_1, \pi_{2k}d_2]$ is the minimum Bayes risk for the problem truncated at M. Also note that the Bayes procedure for the truncated problem is determined by $\beta_{n,k}^M(x_{(n)})$ as follows: Consider for $n = 0, 1, \cdots, M - 1$,

$$(3.3)\qquad D_{n,k}^M(x_{(n)}) = \beta_{n,k}^M(x_{(n)}) - C(n) - d_2\pi_{2k} g_{(2),k}^{(n)}(x_{(n)})$$

and

$$E^M_{n,k}(x_{(n)}) = \beta^M_{n,k}(x_{(n)}) - C(n) - d_1\pi_{1k}g^{(n)}_{(1),k}(x_{(n)}). \tag{3.4}$$

Continue sampling if $D^M_{n,k}(x_{(n)})$ and $E^M_{n,k}(x_{(n)})$ are both negative. Stop and reject if $E^M_{n,k}(x_{(n)}) > 0$ and $d_1\pi_{1k}g^{(n)}_{(1),k}(x_{(n)}) < d_2\pi_{2k}g_{(2),k}(x_{(n)})$, while the null hypothesis is accepted if $D^M_{n,k}(x_{(n)}) > 0$ and $d_1\pi_{1k}g^{(n)}_{(1),k}(x_{(n)}) > d_2\pi_{2k}g^{(n)}_{(2),k}(x_{(n)})$. At stage M, stop and accept if $d_1\pi_{1k}g^{(M)}_{(1),k}(x_{(M)}) > d_2\pi_{2k}g^{(M)}_{(2),k}(x_{(M)})$, and stop and reject if $d_1\pi_{1k}g^{(M)}_{(1),k}(x_{(M)}) < d_2\pi_{2k}g^{(M)}_{(2),k}(x_{(M)})$. If equalities occur at any stage from 0 to M, randomizations can be done.

For the untruncated sequential problem we define $\beta_{n,k}(x_{(n)})$ in an analogous way to $\beta^M_{n,k}(x_{(n)})$ and note that (3.1) is appropriate, as well as (3.3) and (3.4) in defining the Bayes procedure.

Next consider the modified truncated problem, which is the problem where we must stop no later than at stage M and if stage M is reached the loss is put equal to zero. For this modified truncated problem let $\boldsymbol{\beta}^M_{n,k}(x_{(n)})$ be the minimum conditional expected risk given $X_{(n)} = x_{(n)}$ is observed at stage n and sampling continues at least to stage $(n+1)$. For $n = M - 1$,

$$\boldsymbol{\beta}^M_{M-1,k}(x_{(M-1)}) = C(M), \tag{3.5}$$

while for $n = 0, 1, 2, \cdots, (M-2)$, $\boldsymbol{\beta}^M_{n,k}(x_{(n)})$ is defined recursively as $\beta^M_{n,k}(x_{(n)})$ is in (3.1), with $\boldsymbol{\beta}^M_{n+1,k}$ replacing $\beta^M_{n+1,k}$. The Bayes procedure for this modified truncated problem is determined by the analogues of (3.3) and (3.4) which we label as $\mathbf{D}^M_{n,k}(x_{(n)})$ and $\mathbf{E}^M_{n,k}(x_{(n)})$.

To show δ is Bayes will entail some lemmas. Before stating the first lemma let us treat a special case. Suppose the sequence $(\pi_{1k}, \Gamma_{1k}(\cdot), \Gamma_{2k}(\cdot))$ is such that the sequence $\{\pi_{1k}\}$ has a subsequence $\{\pi_{1k'}\}$ which converges to 0. Since $c_1 > 0$, for all sufficiently large k', it follows that $\delta(k'; \bar{\mathbf{x}})$ and hence $\delta(\bar{\mathbf{x}})$ will reject the null hypothesis with probability one without taking any observation. Thus in this special case, δ is trivially Bayes. In what follows then we will assume that any convergent subsequence of $\{\pi_{1k}\}$ goes to a positive limit. Now we give

LEMMA 3.1. *Given the sequence* $(\pi_{1k}, \Gamma_{1k}(\cdot), \Gamma_{2k}(\cdot))$, *there exists a subsequence* $(\pi_{1k'}, \Gamma_{1k'}(\cdot), \Gamma_{2k'}(\cdot))$ *and a prior distribution specified by* $(\pi_{1*}, \Gamma_{1*}(\cdot), \Gamma_{2*}(\cdot))$, *such that for every fixed n and* $x_{(n)}$, $\pi_{ik'}g^{(n)}_{(i),k'}(x_{(n)})$ *converges to* $\pi_{i*}g^{(n)}_{(i),*}(x_{(n)})$, $i = 1, 2$.

PROOF. The space of probability distributions is weakly compact where weak convergence is in the sense that, Γ_k, a sequence of probability distributions converges to Γ, if $\lim_{k\to\infty} \int h(\theta)\Gamma_k(d\theta) = \int h(\theta)\Gamma(d\theta)$, for all h such that, h is continuous and $h(\theta) \to 0$ as $\|\theta\| \to \infty$. The limit distribution need not be a probability distribution since $\Gamma(\Theta) \leqslant 1$. Hence, there exists a convergent subsequence whose limit distribution we denote by $(\tilde{\pi}_1, \tilde{\Gamma}_1(\cdot), \tilde{\Gamma}_2(\cdot))$. Note that since $\overline{\Theta}_1$ is compact, $\tilde{\Gamma}_1(\cdot)$ is in fact a probability measure but $\tilde{\Gamma}_2(\cdot)$ need not be. Define $(\pi_{1*}, \Gamma_{1*}(\cdot), \Gamma_{2*}(\cdot))$ as follows: If $\tilde{\Gamma}_2(\overline{\Theta}_2) = 0$, let $\pi_{1*} = 1$. Otherwise $\pi_{1*} = \tilde{\pi}_1/[\tilde{\pi}_1 +$

$\tilde{\pi}_2\tilde{\Gamma}_2(\overline{\Theta}_2)]$, $\pi_{2*} = 1 - \pi_{1*}$, $\Gamma_{1*}(\cdot) = \tilde{\Gamma}_1(\cdot)$, $\Gamma_{2*}(\cdot) = \tilde{\Gamma}_2(\cdot)/\tilde{\Gamma}_2(\overline{\Theta}_2)$. Recall that for every j and $x_{(j)}$, $f_\theta^{(j)}(x_{(j)})$ is continuous and tends to zero as $\|\theta\| \to \infty$. Now use this fact, (2.3), (2.4), the definition of π_{1*}, $\Gamma_{1*}(\cdot)$, and $\Gamma_{2*}(\cdot)$, and weak convergence to compete the proof of the lemma.

Note that the subsequence and limiting distribution in Lemma 3.1 did not depend on $x_{(n)}$ or n. Now we prove

LEMMA 3.2. *For fixed $x_{(n)}$, each $n = 0, 1, 2, \cdots, M-1$, each M, $D^M_{n,k'}(x_{(n)})$ converges to $D^M_{n,*}(x_{(n)})$, $E^M_{n,k'}(x_{(n)})$ converges to $E^M_{n,*}(x_{(n)})$, $\mathbf{D}^M_{n,k'}(x_{(n)})$ converges to $\mathbf{D}^M_{n,*}(x_{(n)})$ and $\mathbf{E}^M_{n,k'}(x_{(n)})$ converges to $\mathbf{E}^M_{n,*}(x_{(n)})$.*

PROOF. Note that $\beta^M_{n,k'}(x_{(n)})$, defined by (3.1) and (3.2) converges to $\beta^M_{n,*}(x_{(n)})$, for $n = 0, 1, 2, \cdots, M-1$, by virtue of Lemma 3.1 and the dominated convergence theorem. Now use this fact, (3.3), (3.4) and Lemma 3.1 again to complete the proof for $D^M_{n,k'}$ and $E^M_{n,k'}$. The proof for $\mathbf{D}^M_{n,k'}$ and $\mathbf{E}^M_{n,k'}$ is similar. This completes the proof of the lemma.

Before stating the next lemma we note expressions for $\beta_{n,k}(x_{(n)})$, $\beta^M_{n,k}(x_{(n)})$, and $\boldsymbol{\beta}^M_{n,k}(x_{(n)})$ as follows: Let $\delta(k; \overline{\mathbf{x}})$, $\delta^M(k; \overline{\mathbf{x}})$, and $\boldsymbol{\delta}^M(k; \overline{\mathbf{x}})$ be the Bayes tests for the untruncated problem, truncated problem, and modified truncated problem respectively. For $r = n+1, n+2, \cdots$ let $\psi_{r,k}(x_{(r-1)}|x_{(n)})$ be the conditional probability that the Bayes procedure goes to stage r, given that it has gone at least to stage $(n+1)$. Thus for $r = n+2, n+3, \cdots$, $\psi_{r,k}(x_{(r-1)}|x_{(n)}) = \Pi_{j=n+1}^{r-1}\delta_{0j}(k; x_{(j)})$ and $\psi_{n+1,k} = 1$. Thus

$$\beta_{n,k}(x_{(n)}) = \Sigma_{r=n+1}^{\infty}\int\psi_{r,k}(x_{(r-1)}|x_{(n)})(1 - \delta_{0r}(k; x_{(r)})) \cdot (C(r) + \min_i d_i\pi_{ik}g^{(r)}_{(i),k}(x_{(r)}))\nu^{|n}(d\overline{\mathbf{x}}|x_{(n)}), \tag{3.6}$$

where $\nu^{|n}$ denotes the conditional measure,

$$\beta^M_{n,k}(x_{(n)}) = \Sigma_{r=n+1}^{M}\int\psi^M_{r,k}(x_{(r-1)}|x_{(n)})(1 - \delta^M_{0r}(k; x_{(r)})) \cdot (C(r) + \min d_i\pi_i g^{(r)}_{(i),k}(x_{(r)}))\nu^{|n}(d\overline{\mathbf{x}}|x_{(n)}) \tag{3.7}$$

$$\boldsymbol{\beta}^M_{n,k}(x_{(n)}) = \Sigma_{r=n+1}^{M-1}\boldsymbol{\psi}^M_{r,k}(x_{(r-1)}|x_{(n)})(1 - \boldsymbol{\delta}^M_{0r}(k; x_{(r)})) \cdot (C(r) + \min d_i\pi_i g^{(r)}_{(i),k}(x_{(r)}))\nu^{|n}(d\overline{\mathbf{x}}|x_{(n)}) + C(M)\int\boldsymbol{\psi}^M_{M,k}(x_{(M-1)}|x_{(n)})\nu^{|n}(d\overline{\mathbf{x}}|x_{(n)}). \tag{3.8}$$

We now give

LEMMA 3.3. *For each M, each $n = 0, 1, \cdots, M$, each fixed $x_{(n)}$, each k',*

$$D^n_{n,k'}(x_{(n)}) \geqslant D^{n+1}_{n,k'}(x_{(n)}) \geqslant \cdots \geqslant D^M_{n,k'}(x_{(n)}) \geqslant D_{n,k'}(x_{(n)}) \geqslant \mathbf{D}^M_{n,k'}(x_{(n)}) \geqslant \cdots \geqslant \mathbf{D}^{n+1}_{n,k'}(x_{(n)}) \geqslant \mathbf{D}^n_{n,k'}(x_{(n)}). \tag{3.9}$$

Also

$$D^n_{n,*}(x_{(n)}) \geqslant D^{n+1}_{n,*}(x_{(n)}) \geqslant \cdots \geqslant D^M_{n,*}(x_{(n)}) \geqslant D_{n,*}(x_{(n)}) \geqslant D^M_{n,*}(x_{(n)}) \geqslant \cdots \geqslant \mathbf{D}^{n+1}_{n,*}(x_{(n)}) \geqslant \mathbf{D}^n_{n,*}(x_{(n)}). \tag{3.10}$$

As $M \to \infty$, $D^M_{n,k'}(x_{(n)}) \to D_{n,k'}(x_{(n)})$, $\mathbf{D}^M_{n,k'}(x_{(n)}) \to D_{n,k'}(x_{(n)})$, $D^M_{n,*}(x_{(n)}) \to D_{n,*}(x_{(n)})$, *and* $\mathbf{D}^M_{n,*}(x_{(n)}) \to D_{n,*}(x_{(n)})$. *Furthermore the same statements can be made with* E *replacing* D.

PROOF. From the definition of D and E in (3.3) and (3.4) we must prove the statements with β replacing D. Also the proof for $\beta_{n,k}$ and $\beta_{n,*}$ are the same. The inequalities in (3.9) and (3.10) are obvious. See, for example, Ferguson [5], page 318 and page 324. Next we use the arguments which are essentially the same as Ferguson [5], Theorem 3, page 318 and Theorem 6, page 324. To prove $\beta^M_{n,k'}(x_{(n)}) \to \beta_{n,k'}(x_{(n)})$, let $\tilde{\delta}^M(k'; \bar{\mathbf{x}})$ be $\delta(k'; \bar{\mathbf{x}})$ truncated at M with the best terminal decision at stage M. Then $\beta^M_{n,k'}(x_{(n)})$ is less than or equal to the expected risk using $\tilde{\delta}^M(k'; \bar{\mathbf{x}})$. Hence using (3.6) we have

$$\begin{aligned}(3.11)\quad \beta^M_{n,k'}(x_{(n)}) - \beta_{n,k'}(x_{(n)}) &\leqslant \int \psi_{M+1,k'}(x_{(M)}|x_{(n)})\left[\min d_i \pi_{ik'} g^{(M)}_{(i),k'}(x_{(M)})\right] \\ &\qquad \cdot \nu^{|n}(d\bar{\mathbf{x}}|x_{(n)}) \\ &\leqslant B\left\{\int \psi_{M+1,k'}(x_{(M)}|x_{(n)})\nu^{|n}(d\bar{\mathbf{x}}|x_{(n)})\right\},\end{aligned}$$

where $B = \max(d_1, d_2)$. The bracketed integral in (3.11), when multiplied by $C(M+1)$ is certainly less than the conditional Bayes risk, which in turn is less than or equal to $C(n) + B$. Hence $\int \psi_{M+1,k'}(x_{(M)}|x_{(n)})\nu^{|n}(d\bar{\mathbf{x}}|x_{(n)}) \leqslant [C(n) + B]/C(M+1)$. Since $C(M) \to \infty$ as $M \to \infty$, from (3.11) we have $\beta^M_n(x_{(n)}) \to \beta_n(x_{(n)})$. (The part of the lemma just proven is also given in Chow, Robbins, and Siegmund [3], Theorem 4.3, page 68.) To complete the proof of the lemma, note that $\beta^M_{n,k'}(x_{(n)})$ is bounded above by the conditional expected risk using the Bayes procedure corresponding to the modified truncated problem. Hence,

$$(3.12)\qquad \beta^M_{n,k'}(x_{(n)}) - \boldsymbol{\beta}^M_{n,k'}(x_{(n)}) \leqslant B\int \psi_M(x_{(M)}|x_{(n)})\nu^{|n}(d\bar{\mathbf{x}}|x_{(n)}).$$

The expression on the right-hand side of (3.12) tends to zero and thus the proof of the lemma is complete.

Now we are ready to prove

THEOREM 3.1. *The test* $\delta(\bar{\mathbf{x}})$, *which is the regular limit of the sequence* $\delta(k; \bar{\mathbf{x}})$, *is Bayes.*

PROOF. Let $\delta^*(\bar{\mathbf{x}})$ be the Bayes procedure with respect to the prior $(\pi_{1*}, \Gamma_{1*}(\cdot), \Gamma_{2*}(\cdot))$. Hence $\delta^*(\bar{\mathbf{x}})$ is determined for each n and each $\bar{\mathbf{x}}_n$ by $Q_{n,*}(x_{(n)}) = (D_{n,*}(x_{(n)}), E_{n,*}(x_{(n)}), d_1\pi_{1*}g^{(n)}_{(1),*}(x_{(n)}), d_2\pi_{2*}g^{(n)}_{(2),*}(x_{(n)}))$ through the rules given after (3.3) and (3.4). We show that $\delta(\bar{\mathbf{x}})$ is also Bayes with respect to $(\pi_{1*}, \Gamma_{1*}(\cdot), \Gamma_{2*}(\cdot))$. For suppose $\delta(\bar{\mathbf{x}})$ is not Bayes. Then there exists some integer n, and some set of $x_{(n)}$, say U, of positive ν measure, such that for these $x_{(n)}$, $(\delta_{0n}(x_{(n)}), \delta_{1n}(x_{(n)}), \delta_{2n}(x_{(n)})$ does not obey the rules given after (3.3) and (3.4).

At this point we partition the space of $x_{(n)}$ into seven sets. These sets include all the different possible combinations for $Q_{n,*}(x_{(n)})$ that have bearing on $(\delta^*_{0n}(x_{(n)}), \delta^*_{1n}(x_{(n)}), \delta^*_{2n}(x_{(n)}))$. For each set we also indicate the values of $(\delta^*_{0n}(x_{(n)}), \delta^*_{1n}(x_{(n)}), \delta^*_{2n}(x_{(n)}))$. The sets labeled T_r, $r = 1, 2, \cdots, 7$ are collections of $x_{(n)}$ for which the following hold: Let $(1) \equiv D_{n,*}(x_{(n)})$, $(2) \equiv E_{n,*}(x_{(n)})$, $(3) \equiv d_1\pi_{1*}\, g^{(n)}_{(1),*}(x_{(n)})$, $(4) \equiv d_2\pi_{2*}\, g^{(n)}_{(2),*}(x_{(n)})$. The letter A with a subscript indicates some arbitrary number between 0 and 1 inclusive.

(3.13)	$(1) > 0, (3) > (4).$	$: (0, 1, 0)$
(3.14)	$(2) > 0, (3) < (4).$	$: (0, 0, 1)$
(3.15)	$(1) > 0, (2) > 0, (3) = (4).$	$: (0, A_1, A_2), A_1 + A_2 = 1,$
(3.16)	$(1) < 0, (2) < 0.$	$: (1, 0, 0)$
(3.17)	$(1) = 0, (2) < 0, (3) > (4).$	$: (A_0, A_1, 0), A_0 + A_1 = 1$
(3.18)	$(1) < 0, (2) = 0, (3) < (4).$	$: (A_0, 0, A_2), A_0 + A_2 = 1$
(3.19)	$(1) = 0, (2) = 0, (3) = (4).$	$: (A_0, A_1, A_2), A_0 + A_1 + A_2 = 1$

Now suppose $\nu(T_1 \cap U) > 0$. Note that for any $x_{(n)} \in T_1 \cap U$, $D_{n,*}(x_{(n)}) > 0$. By Lemma 3.3 then, we have for $M \geqslant M_0$, say, $\mathbf{D}^M_{n,*}(x_{(n)}) > 0$. By Lemma 3.2 then we have for all $k' > k'(M_0)$, $\mathbf{D}^{M_0}_{n,k'}(x_{(n)}) > 0$. By the inequalities (3.9) of Lemma 3.3 then, we have that for all sufficiently large k', $D_{n,k'}(x_{(n)}) > 0$. Furthermore, for all sufficiently large k' we have $d_2\pi_{2,k'}\, g^{(n)}_{(2),k'}(x_{(n)}) < d_1\pi_{1,k'}\, g^{(n)}_{(1),k'}(x_{(n)})$, by Lemma 3.1. Hence for each $x_{(n)} \in T_1 \cap U$, $(\delta_{0n}(k'; x_{(n)}), \delta_{1n}(k'; x_{(n)}), \delta_{2n}(k', x_{(n)})) \to (0, 1, 0)$, as $k' \to \infty$. Since pointwise convergence implies regular convergence, and since $\delta(\bar{\mathbf{x}})$ is the regular limit of $\delta(k; \bar{\mathbf{x}})$, we must therefore have $(\delta_{0n}(x_{(n)}), \delta_{1n}(x_{(n)}), \delta_{2n}(x_{(n)})) = (0, 1, 0)$ a.e. ν on $(T_1 \cap U)$. This contradicts the claim that δ does not obey the rules for $x_{(n)} \in U$.

Next suppose $\nu(T_2 \cap U) > 0$. Referring to (3.13), the same argument using $\mathbf{E}$ instead of $\mathbf{D}$ shows that this cannot be true. Suppose $\nu(T_3 \cap U) > 0$. Similar reasoning implies that for all sufficiently large k', $\delta_{0n}(k'; (x_{(n)}) = 0$, i.e., the Bayes procedure must stop, which in turn implies that $\delta_{0n}(x_{(n)}) = 0$. Since $\delta^*_{in}(x_{(n)})$, $i = 1, 2$, are arbitrary, clearly $\delta_{in}(x_{(n)})$, $i = 1, 2$, obey the rules given after (3.3) and (3.4) no matter what they are. Hence this too is a contradiction. Suppose $\nu(T_4 \cap U) > 0$. Use the same argument as initially used, replacing $\mathbf{D}^M$ by D^M and replacing $\mathbf{E}^M$ by E^M, when appropriate. This also leads to a contradiction. Clearly the last three cases can be treated using the same ideas, which leads to the fact that $\delta(\bar{\mathbf{x}})$ is Bayes. This completes the proof of the theorem.

Next we prove

THEOREM 3.2. *The Bayes tests form a complete class.*

PROOF. It is well known that the closure (regular convergence) of the class of Bayes procedures is an essentially complete class. (See, Le Cam [6], page 78.) By

Theorem 3.1 we have that the regular limits of Bayes procedures are Bayes and hence the Bayes procedures are essentially complete. Now let $\delta(\bar{\mathbf{x}})$ be a procedure outside of this class, such that there is a procedure $\delta'(\bar{\mathbf{x}})$ in the class, whose risk function is the same as the risk of $\delta(\bar{\mathbf{x}})$. This would imply that $\delta(\bar{\mathbf{x}})$ would be Bayes with respect to the same prior for which $\delta'(\bar{\mathbf{x}})$ is Bayes. This contradicts the claim that $\delta(\bar{\mathbf{x}})$ lies outside of this class. Thus we have demonstrated that whenever the Bayes tests are essentially complete, they are in fact a complete class. This completes the proof of the theorem.

We now discuss some of the special cases for which Theorem 3.2 applies. As remarked earlier the theorem applies in great generality. It is not limited to tests concerned only with means. For example, all the conditions can be shown to hold for testing the null hypothesis that σ^2, the variance of a normal distribution (with known mean), is such that, $0 < a \leqslant \sigma^2 \leqslant b$, $b < \infty$, against the alternative $\{0 < \sigma^2 < a, b < \sigma^2 < \infty\}$. The conditions are easy to verify when working with the transformed parameter $\log \sigma^2$. Testing the variance of a normal distribution is a special case of a more general situation that can be handled. Namely a situation where the closure of the alternative space includes parameter values that do not lie in the natural parameter space determined by the exponential family, from which observations are taken. In such cases, if θ denotes the parameter then suppose $T(\theta)$ is a transformation from Θ onto R^q such that $T(\theta)$ maps points on the boundary of the natural parameter space to $\{\infty\}$. Then the theorem will apply provided that for all j and every $x_{(j)}$, $\lim_{\|T(\theta)\|\to\infty} f_\theta^{(j)}(x_{(j)}) \to 0$.

Discrete distributions such as binomial and Poisson can be treated. The verifications of conditions can be made in terms of the original variables and parameters. Although most nuisance parameter problems cannot be treated, the result will apply to classes of tests restricted by invariance. For example, in testing a normal mean with unknown variance, if only scale invariant tests are permitted, then the Bayes tests based on Student's t-statistic form a complete class among the class of invariant tests. (At stage one however, we must take two observations in this case.) Other types of invariance can be used to reduce other problems with nuisance parameters, so that the theorem would be appropriate for a restricted class of tests.

To shed some light on the result of Theorem 3.2 we give the following example which is somewhat surprising. Let (X_1, X_2) be bivariate normal with mean vector (μ_1, μ_2) and covariance matrix I. Consider the problem of testing $(\mu_1, \mu_2) = (0, 0)$ against the alternative that $(\mu_1, \mu_2) \neq (0, 0)$. Assume $c_1 > 0$, and c_2 is large, say $c_2 > \max(d_1, d_2)$. (This assumption on c_2 guarantees that the tests which terminate by stage 1 form a complete class.) Now consider the fixed sample size procedure, with sample size 1, which accepts the null hypothesis if $|x_1| < K$, $|x_2| < K$, for some $K > 0$. Theorem 3.2 implies that this fixed sample size test, which is admissible among fixed sample size tests, is inadmissible. The test which is better must stop at time zero with probability q, where q is strictly between 0 and 1, and make some terminal decision.

We now give an example which shows that Theorem 3.1 would not necessarily be true if $\lim_{n\to\infty} C(n) \neq \infty$. The example is one which violates Lemma 3.3 in the sense that $\beta_{0,*}^M \nrightarrow \beta_{0,*}$ as $M \to \infty$. The example is as follows. Let X be $N(\theta, 1)$. Let $\Theta_1 = [-1, 0]$, $\Theta_2 = (0, \infty)$. Let $d_1 = d_2 = 1$. Let $\lim_{n\to\infty} C(n) = \frac{1}{4}$. Choose $(\pi_{1k}, \Gamma_{1k}, \Gamma_{2k}) = (\frac{1}{2}, \varepsilon\{-1/k\}, \varepsilon\{1/k\})$. Then $\beta_{0,k} \leqslant \frac{1}{4}$, for every fixed k, since the Bayes risk is bounded by the expected risk of the procedure which takes an infinite sample and thereby makes an error in the terminal decision with probability zero. (Reject the null hypothesis if $\overline{X}_n > 0$ after choosing n sufficiently large, will insure that the risk is bounded by $\frac{1}{4}$.) On the other hand $(\pi_{1*}, \Gamma_{1*}, \Gamma_{2*}) = (\frac{1}{2}, \varepsilon\{0\}, \varepsilon\{0\})$ and the Bayes procedure for such a prior is not to sample and accept the null hypothesis with probability $\frac{1}{2}$. Thus $\beta_{0,*} = \frac{1}{2}$. Thus the regular limit of the sequence of Bayes procedures would not be Bayes against $(\pi_{1*}, \Gamma_{1*}, \Gamma_{2*})$. In fact it could be shown that the regular limit of the sequence of Bayes procedures could not be Bayes.

We now make

REMARK 3.1. Let X be a one dimensional random variable whose density belongs to the exponential family, (see (4.1)), and whose density is dominated by some measure ν, which is absolutely continuous with respect to Lebesgue measure. Consider the problem of testing $\overline{\Theta}_1 = [a, b]$, for $-\infty < a \leqslant b < \infty$, against $\overline{\Theta}_2 = (-\infty, \infty) - (a, b)$. Under the assumptions of this section the class of Bayes tests is minimal complete. This follows from Theorem 3.2 and the fact that Bayes tests for this problem would be unique a.e. ν. The uniqueness of the Bayes tests could be shown by arguments given in Brown, Cohen, and Strawderman [2]. For testing $\overline{\Theta}_1$ as above, against $\Theta_2 = (-\infty, \infty) - \overline{\Theta}_1$, the result is that the Bayes tests for $\overline{\Theta}_1$ vs. $\overline{\Theta}_2$ are minimally essentially complete. That is, all Bayes tests in the essentially complete class are admissible.

We next prove a theorem and a corollary. The theorem is concerned with elimination of randomization in the stopping rule and terminal decision rule, except at stage 0. The theorem will be given assuming that we sample from a multivariate normal distribution. It will be clear however that the proof would work for many other distributions in the exponential family. It is known that procedures based on a sufficient statistic are an essentially complete class. (See Ferguson, page 337). The corollary describes the case where tests based only on a sufficient statistic form a complete class as opposed to an essentially complete class.

Now we state

THEOREM 3.3. *Let $\{X_i\}$ be independent, identically distributed according to the multivariate normal distribution with mean vector θ and covariance matrix I. Then any test $\delta(\overline{\mathbf{x}})$, with components $(\delta_{0n}(x_{(n)}), \delta_{1n}(x_{(n)}), \delta_{2n}(x_{(n)})$ is inadmissible if for any $n \geqslant 1$ and any $i = 0, 1, 2$; $0 < \delta_{in}(x_{(n)}) < 1$ on a set of positive measure.*

PROOF. We will show that randomized tests are not Bayes. Once we show this, the theorem follows from Theorem 3.2.

Our first step is to recognize that given any test $\delta(\bar{\mathbf{x}})$ there exists a test based on the transitive sufficient statistic $S_n = \Sigma_{j=1}^n X_j$, say $\delta(s)$, which is just as good. (See Ferguson [5], Theorem 4, page 337). Furthermore $\delta(s)$ is a randomized test. This follows from the fact that S_n has a normal distribution and by the construction of the stopping rule (see Ferguson, page 336, equation (7.24)), and the construction of the terminal decision rule, (see Ferguson, page 120, equation (3.47)). Hence it suffices to show that the randomized test $\delta(s)$ cannot be Bayes among the class of all tests based on the transitive sufficient statistic S_n.

Next we note that Bayes tests based on S_n are such that the conditional Bayes risks at stage n, depend solely on s_n. That is, $\beta_n(s_1, s_2, \cdots, s_n)$ can be written as $\beta_n(s_n)$. To see this, observe from Brown, Cohen, Strawderman [2], equation (3.4) that $g_{(i)}^{(n)}(s_1, s_2, \cdots, s_n)$ can be written as $g_{(i)}^{(n)}(s_n)$. Also from Lemma 3.1 of the same reference $\nu^{((n+1)|n)}(\cdot|s_1, s_2, \cdots, s_n)$ is $\nu^{((n+1)|n)}(\cdot|s_n)$. Thus from (3.1) and (3.2) we have that β_n^M depends only on s_n, which in turn implies, by the proof of Lemma 3.3, that β_n depends only on s_n. Furthermore, from (3.2) (which is appropriate for β_n as well as β_n^M) the dependence on s_n enters the integrand of the expression on the right-hand side of (3.2), only through the conditional distribution of $S_{n+1}|S_n = s_n$.

Suppose we show that $\beta_n(s_n)$ is an analytic function of s_n. Then, since $g_{(i)}^{(n)}(s_n)$ is an analytic function of s_n, it will follow that D_n and E_n are analytic functions of s_n. This in turn will imply that for any fixed $n \geqslant 1$, the set of points for which $D_n = 0$, or $E_n = 0$, or $\pi_1 d_1 g_{(1)}^{(n)} = \pi_2 d_2 g_{(2)}^{(n)}$, is either a set of Lebesgue measure zero or is the n dimensional real space. See Farrell [4], Lemma 4.2, page 7. Suppose then that $\pi_1 d_1 g_{(1)}^{(n)}(s_n) = \pi_2 d_2 g_{(2)}^{(n)}(s_n)$ for every s_n. Since $f_\theta^{(n)}(s_n)$ are normal densities, it follows that $g_{(1)}^{(n)}(s_n) = g_2^{(n)}(s_n)$ for every s_n and $\pi_1 d_1 = \pi_2 d_2$. In turn this would imply that $\Gamma_1 = \Gamma_2$, by the uniqueness of LaPlace transforms. Under these conditions the Bayes test is to stop at time 0 and randomize between the two possible terminal decisions. Now suppose $D_n = 0$ for every s_n. Note from (3.1) and (3.3) that

$$\begin{aligned} D_n(s_n) &= C_{n+1} + \int \min\left[D_{n+1}(s_{n+1}), 0, d_1\pi_1 g_{(1)}^{(n+1)}(s_{n+1}) - d_2\pi_2 g_{(2)}^{(n+1)}(s_{n+1})\right] \\ &\quad \cdot \nu^{((n+1)|n)}(ds_{n+1}|s_n) \\ &= C_{n+1} + \int \min\left[D_{n+1}(s_{n+1}), 0, d_1\pi_1 g_{(1)}^{(n+1)}(s_{n+1}) - d_2\pi_2 g_{(2)}^{(n+1)}(s_{n+1})\right] \\ &\quad \cdot p(s_{n+1}|s_n)\, ds_{n+1}, \end{aligned} \tag{3.20}$$

where

$$p(s_{n+1}|s_n) \propto \left[e^{-s_n' s_n/2n} \Big/ \int e^{s_n'\theta} e^{-n\theta'\theta/2}\Gamma(d\theta)\right] \cdot e^{s_{n+1}' s_n} \cdot e^{-n s_{n+1}' s_{n+1}/2(n+1)} \int e^{-(n+1)(\theta - (S_{n+1}/n+1))'(\theta-(S_{n+1}/n+1))/2}\Gamma(d\theta). \tag{3.21}$$

From (3.21) we recognize that the distribution of $S_{n+1}|S_n = s_n$ is exponential family and is also a complete family. Hence if (3.20) is identically zero it follows that $\min\, [D_{n+1}(s_{n+1}), 0, d_1\pi_1 g_{(1)}^{(n+1)}(s_{n+1}) - d_2\pi_2 g_{(2)}^{(n+1)}(s_{n+1})] \equiv -c_{n+1}$. But

$[d_1\pi_1 g_{(1)}^{(n+1)}(s_{n+1}) - d_2\pi_2 g_{(2)}^{(n+1)}(s_{n+1})]$ cannot be identically constant on a set of positive Lebesgue measure as already argued and so $D_{n+1}(s_{n+1}) \equiv -c_{n+1}$. Continuing by induction this can be true only if $D_j(s_j) \equiv \Sigma_{i=n+1}^{j} c_i \to -\infty$ as $j \to \infty$. However $D_j(s_j) \geqslant -d_2$. This contradiction shows that D_n is not identically zero. Similarly for E_n. Thus we have that if Γ_1 and Γ_2 are distinct, the set of points for which $D_n = 0$ or $E_n = 0$ or $d_1\pi_1 g_{(1)}^{(n)}(s_n) = d_2\pi_2 g_{(2)}^{(n)}(s_n)$, is a set of measure zero. This in turn implies that any Bayes test cannot be randomized after stage 0.

To complete the proof of the theorem we note that $\beta_n(s_n)$ can be written as

$$\begin{aligned}\beta_n(s_n) = \int \min\big[\, &\beta_{n+1}(s_{n+1}), C(n+1) + d_2\pi_2 g_{(2)}^{(n+1)}(s_{n+1}),\\ &\times C(n+1) + d_1\pi_1 g_{(1)}^{(n+1)}(s_{n+1})\big] p(s_{n+1}|s_n)\, ds_{n+1},\end{aligned} \tag{3.22}$$

where $p(s_{n+1}|s_n)$ is given in (3.21). Clearly $\beta_n(s_n)$ is an analytic function of s_n. This completes the proof of Theorem 3.3.

Theorem 3.3 yields the following

COROLLARY 3.1. *Under the conditions of Theorem* 3.3, *any test procedure not based only on a sufficient statistic is inadmissible.*

PROOF. Given any test not based on a sufficient statistic there exists one based on a sufficient statistic which is just as good. As in the proof of Theorem 3.3 the matching test based on a sufficient statistic is randomized. Hence by Theorem 3.3 such a test is inadmissible, which implies the original test is inadmissible. This completes the proof of the corollary.

4. Exponential family and $c_1 = 0$. In this section we assume that the cost of the first observation is zero and that a first observation must be taken. Again we study limits of sequences of Bayes tests. We contrast the distinction between this model and the model of Section 3 where $c_1 > 0$. If we consider, as in the Introduction, the possibility that $\pi_{1k} \to 0$ as $\Gamma_{2k}(\cdot)$ send mass to $\{\infty\}$, then since we now have an observation, the determination of the limit of a sequence of Bayes tests depends on the distribution of the observable random variables. When $c_1 > 0$, this was not the case. Also when $c_1 > 0$, such a sequence of priors led to a limiting test which was Bayes. Now however the limiting procedure need not be Bayes (as in Farrell). Furthermore it is such a sequence of priors, plus the fact that we have an observation, that now requires σ-finite measures (as opposed to probability measures) and non-Bayes procedures, such as generalized Bayes tests for example, in the determination of the limiting test. We proceed.

The random $(p \times 1)$ vector X is distributed according to the exponential family if its distribution is

$$P_\theta(dx) = C(\theta) \exp \theta' x \mu(dx), \tag{4.1}$$

where θ is $(p \times 1)$. In this section we assume that the natural observations X_i are independent and identically distributed with distribution of the form (4.1). The sufficient transitive sequence $S_n = \Sigma_{j=1}^{n} X_j$, at stage n, then has distribution

$$P_\theta^{(n)}(ds_n) = C^n(\theta) \exp \theta' s_n \gamma^{(n)}(ds_n). \tag{4.2}$$

The measure $\gamma^{(n)}$ is assumed to be absolutely continuous with respect to Lebesgue measure. Let $\mathfrak{N}$ denote the natural parameter space. We also assume $C(\theta)e^{\theta' t} \to 0$ as $\theta \to \overline{\Theta} - \mathfrak{N}$, or as $\|\theta\| \to \infty$ for every t lying in the interior of the convex hull of the support of μ. This implies $C^n(\theta)e^{\theta' s_n} \to 0$ as $\theta \to \overline{\Theta} - \mathfrak{N}$, or as $\|\theta\| \to \infty$ for almost every s_n and every n. In accordance with our earlier convention we will ignore the set of values of s_n of measure zero for which the condition may not hold. These last assumptions are made so that we may treat cases involving distributions like the exponential, for example, where $\mathfrak{N} = (0, \infty)$ but $\overline{\Theta} = [0, \infty)$.

Now let the cost of the first observation $c_1 = 0$. For $j = 2, 3, \cdots$ however, assume $c_j > 0$ and $\lim_{n\to\infty} C(n) = \infty$. Let $\overline{\Theta}_1$ be compact and $\overline{\Theta}_1 \subset \mathfrak{N}$. As in the previous section we will test $H_1 : \theta \in \overline{\Theta}_1$ vs. $H_2 : \theta \in \overline{\Theta}_2$.

We start with

LEMMA 4.1. *Let z be a $p \times 1$ vector and let $Q_k(\cdot)$ be a sequence of σ-finite measures on Θ. Then there exists a subsequence and a σ-finite measure Q on $\overline{\Theta}$, with the following properties: There exists a convex set A such that for z lying in the interior of A, and for any bounded continuous function $g : \overline{\Theta} \to (-\infty, \infty)$,*

$$\textstyle\int_\Theta g(\theta)e^{z'\theta}Q_{k'}(d\theta) \to \int_{\bar\theta} g(\theta)e^{z'\theta}Q(d\theta) < \infty, \tag{4.3}$$

as $k' \to \infty$. For $z \in \bar{A}^c$,

$$\textstyle\int_\Theta e^{z'\theta}Q_{k'}(d\theta) \to \infty. \tag{4.4}$$

PROOF. The above lemma is known as the continuity theorem for multivariate LaPlace transforms. A proof is deducible from Brown [1], Theorem 2.2.1, page 864.

Now let $(\pi_{1k}, \Gamma_{1k}(\cdot), \Gamma_{2k}(\cdot))$ be a sequence of prior distributions. Let $\delta(k; \bar{\mathbf{x}})$ be the Bayes tests for these priors and let $\delta(\bar{\mathbf{x}})$ be the regular limit of the sequence $\{\delta(k; \bar{\mathbf{x}})\}$. Define an absorbing prior distribution on $\overline{\Theta}$ as a pair $(\Psi_1(\cdot), \Psi_2(\cdot))$, where $\Psi_1(\cdot)$ is a probability measure on $\overline{\Theta}_1$ and $\Psi_2(\cdot)$ is a σ-finite measure on $\overline{\Theta}_2$. A "posterior minimizer" test at s_1, with respect to $\Psi = (\Psi_1, \Psi_2)$ is defined as a test which minimizes for each s_1,

$$\begin{aligned}\textstyle\int_{\overline{\Theta}_1}\int_{\mathfrak{X}^*} L(\theta, \delta(\bar{\mathbf{x}}))P_\theta(d\bar{\mathbf{x}}^*)C(\theta)e^{s_1'\theta}\Psi_1(d\theta)& \\ +\textstyle\int_{\overline{\Theta}_2}\int_{\mathfrak{X}^*} L(\theta, \delta(\bar{\mathbf{x}}))P_\theta(d\bar{\mathbf{x}}^*)e^{s_1'\theta}\Psi_2(d\theta),&\end{aligned} \tag{4.5}$$

where $\bar{\mathbf{x}}^* = (x_2, x_3, \cdots) \in \mathfrak{X}^*$ and note $s_1 = x_1$. If $\overline{\Theta}_2 \subset \mathfrak{N}$ then the "posterior minimizer" relative to Ψ_1, Ψ_2 may also be described as a generalized Bayes test. If $\overline{\Theta}_2$ is not contained in $\mathfrak{N}$, then a "posterior minimizer" need not be generalized Bayes. (See the definition of generalized Bayes test and Remark 4.1 after the proof of Theorem 4.1 for clarifications.)

Define the class $\bar{B}$ of tests as follows: If the observed value s_1 lies in the complement of the closure of some given convex set A, stop and reject. If the observed value lies in the interior of the convex set, then the procedure is determined as a "posterior minimizer" test at s_1, against some $(\Psi_1(\cdot), \Psi_2(\cdot))$, where $\Psi_1(\cdot)$ is a probability measure on $\bar{\Theta}_1$ and $\Psi_2(\cdot)$ is a σ-finite measure on $\bar{\Theta}_2$, such that for $s_1 \in A, \int_{\bar{\Theta}_2} e^{s_1'\theta}\Psi_2(d\theta) < \infty$.

Before we state Theorem 4.1 we note another distinction between the model of Section 3 and the model of Section 4. In Section 3, if the sequence of priors converged to a common measure on the boundaries of $\bar{\Theta}_1$ and $\bar{\Theta}_2$, then the resulting Bayes procedure was to stop at time zero and randomize between acceptance and rejection of the null hypothesis. In this section such a sequence also leads to a Bayes procedure, one which is in $\bar{B}$, that requires stopping after the first observation and then doing anything. Hence in this situation, the Bayes test is not unique. These particular Bayes tests will be part of the essentially complete class determined by Theorem 4.1. In special cases of course, such as one dimensional exponential family, many of the tests would be obviously inadmissible and could be discarded. Nevertheless, this does not alter the statement of Theorem 4.1. In the proof of Theorem 4.1, when we say we proceed as in Theorem 3.1 it will be understood that the above distinction is made.

Now we state

THEOREM 4.1. *Let X be distributed according to* (4.1). *Let* $c_1 = 0$. *Then the class* $\bar{B}$ *is an essentially complete class.*

Let $\bar{\Theta}$ be a convex cone. Then the convex set A is the intersection of $\bar{\Theta}$ and half spaces whose outward normals are in $\bar{\Theta}$.

PROOF. The closure of the class of Bayes procedures is an essentially complete class. To establish a characterization of procedures in the essentially complete class we study $\delta(\bar{\mathbf{x}})$ which is the regular limit of the sequence $\delta(k; \bar{\mathbf{x}})$. Hence consider the sequence $[\Gamma_{1k}(\cdot), (\pi_{2k}/\pi_{1k})\Gamma_{2k}(\cdot)]$ and the sequence

(4.6)

$$d_1\pi_{1k}g_{(1),k}^{(1)}(s_1) = \frac{d_1\pi_{1k}\int_{\bar{\Theta}_1} e^{s_1'\theta}C(\theta)\Gamma_{1k}(d\theta)}{\pi_{1k}\int_{\bar{\Theta}_1} e^{s_1'\theta}C(\theta)\Gamma_{1k}(d\theta) + \pi_{2k}\int_{\bar{\Theta}_2} e^{s_1'\theta}C(\theta)\Gamma_{2k}(d\theta)}$$

$$= d_1\int_{\bar{\Theta}_1} e^{s_1'\theta}C(\theta)\Gamma_{1k}(d\theta)\Big/\Big[\int_{\bar{\Theta}_1} e^{s_1'\theta}C(\theta)\Gamma_{1k}(d\theta) + (\pi_{2k}/\pi_{1k})\int_{\bar{\Theta}_2} e^{s_1'\theta}C(\theta)\Gamma_{2k}(d\theta)\Big].$$

Since $\bar{\Theta}_1$ is compact, $\Gamma_{1k}(\cdot)$ has a weakly convergent subsequence to $\Psi_1(\cdot)$, where Ψ_1 is a probability measure. For the sequence $(\pi_{2k}/\pi_{1k})\Gamma_{2k}$ there are two possibilities. One is that there exists a compact subset $B \subset \bar{\Theta}_2$, and a subsequence with the property that $\lim_{k' \to \infty} (\pi_{2k'}/\pi_{1k'})\Gamma_{2k'}(B) = \infty$. In this case, in light of (4.6), the limiting $\delta(\bar{\mathbf{x}})$ stops at stage one and rejects $\bar{\Theta}_1$, for every x_1. This is in keeping with the theorem for $A = \{\phi\}$ or A a convex set of Lebesgue measure zero. If no such set B exists however, then by setting $Q_k = (\pi_{2k}/\pi_{1k})C(\theta)\Gamma_{2k}$ we can apply Lemma 4.1. We thus have that for s_1 lying in the interior of A, a subsequence of (4.6) converges pointwise to the limit

$$d_1\int_{\bar{\Theta}_1} e^{s_1'\theta}C(\theta)\Psi_1(d\theta)\Big/\Big[\int_{\bar{\Theta}_1} e^{s_1'\theta}C(\theta)\Psi_1(d\theta) + \int_{\bar{\Theta}_2} e^{s_1'\theta}\Psi_2(d\theta)\Big], \tag{4.7}$$

where Ψ_2 is the limit of Q_k over this subsequence. For $s_1 \in \bar{A}_1^c$, the limit of (4.6) over this subsequence is zero, which implies that for each such s_1, there exists a k^* sufficiently large so that for all $k' \geqslant k^*$, the expression in (4.6) is bounded by ε, $\varepsilon > 0$ a small arbitrary number. Since $\beta_{1k'}(s_1) \geqslant c_2 > \varepsilon$ for all k' sufficiently large, we have for such k', $\delta(k'; \bar{\mathbf{x}})$ is such that $(\delta_{01}(k'; s_1), \delta_{11}(k'; s_1), \delta_{21}(k'; s_1))$ is $(0, 0, 1)$, which implies that the regular limit is $(\delta_{01}(s_1), \delta_{11}(s_1), \delta_{21}(s_1)) = (0, 0, 1)$.

For s_1 lying in the interior of A we must proceed as was done in Theorem 3.1. That is, we want to study $\delta(\bar{\mathbf{x}})$ which is the regular limit of the sequence $\delta(k; \bar{\mathbf{x}})$. Note

$$(4.8)\quad \int_{\bar{\Theta}_2} e^{s_n'\theta} C^n(\theta)(\pi_{2k'}/\pi_{1k'})\Gamma_{2k'}(d\theta) = \int_{\bar{\Theta}_2} e^{(\Sigma_2^n x_j)'\theta} C^{n-1}(\theta) e^{s_1'\theta} C(\theta)(\pi_{2k'}/\pi_{1k'})\Gamma_{2k'}(d\theta).$$

Since $e^{(\Sigma_2^n x_j)'\theta} C^{(n-1)}(\theta) \to 0$ as $\theta \to \bar{\Theta}_2 - \mathfrak{N}$, for s_1 lying in the interior of A, Lemma 4.1 can be applied so that from (4.8) we have for every s_n,

$$(4.9)\qquad \int_{\bar{\Theta}_2} e^{s_n'\theta} C^n(\theta)(\pi_{2k'}/\pi_{1k'})\Gamma_{2k'}(d\theta) \to \int_{\bar{\Theta}_2} e^{s_n'\theta} C^{n-1}(\theta)\Psi_2(d\theta).$$

Also since $\bar{\Theta}_1$ is compact, the probability measure Ψ_1 is such that for all s_n,

$$(4.10)\qquad \int_{\bar{\Theta}_1} e^{s_n'\theta} C^n(\theta)\Gamma_{1k'}(d\theta) \to \int_{\bar{\Theta}_1} e^{s_n'\theta} C^n(\theta)\Psi_1(d\theta).$$

If

$$\tilde{g}_{(1)}^{(n)}(s_n) = \int_{\bar{\Theta}_1} e^{s_n'\theta} C^n(\theta)\Psi_1(d\theta) \Big/ \Big[\int_{\bar{\Theta}_1} e^{s_n'\theta} C^n(\theta)\Psi_1(d\theta) + \int_{\bar{\Theta}_2} e^{s_n'\theta} C^{n-1}(\theta)\Psi_2(d\theta)\Big],$$

and

$$\tilde{g}_{(2)}^{(n)}(s_n) = \int_{\bar{\Theta}_2} e^{s_n'\theta} C^{n-1}(\theta)\Psi_2(d\theta) \Big/ \Big[\int_{\bar{\Theta}_1} e^{s_n'\theta} C^n(\theta)\Psi_1(d\theta) + \int_{\bar{\Theta}_2} e^{s_n'\theta} C^{n-1}(\theta)\Psi_2(d\theta)\Big],$$

then for all $n = 1, 2, \cdots$, (4.9) and (4.10) imply that $g_{(i), k'}^{(n)}(s_n) \to \tilde{g}_{(i)}^{(n)}(s_n)$ pointwise, $i = 1, 2$. Next consider the truncated and modified truncated problems and proceed exactly as in Theorem 3.1 to argue that if $s_1 \in A$, the regular limit of the sequence $\delta(k; \bar{\mathbf{x}})$ is determined as a posterior minimizer procedure at s_1 against (Ψ_1, Ψ_2). This completes the proof that $\bar{B}$ is an essentially complete class.

The fact that A is the intersection of $\bar{\Theta}$ and half spaces whose outward normals in $\bar{\Theta}$ can be argued as in Farrell [4], page 21. This completes the proof of the theorem.

Before proceeding with the next theorem we clarify the distinction between "posterior minimizer" and generalized Bayes test. Note

DEFINITION 4.1. A generalized Bayes test at s_1, with respect to $\hat{\Gamma} = (\hat{\Gamma}_1(\cdot), \hat{\Gamma}_2(\cdot))$ is defined as a test which minimizes for each s_1,

$$\int_{\bar{\Theta}_1}\int_{\mathfrak{X}^*} L(\theta, \delta(\bar{\mathbf{x}}))P_\theta(d\bar{\mathbf{x}}^*)C(\theta)e^{s_1'\theta}\hat{\Gamma}_1(d\theta) + \int_{\bar{\Theta}_2}\int_{\mathfrak{X}^*} L(\theta, \delta(\bar{\mathbf{x}}))P_\theta(d\bar{\mathbf{x}}^*)C(\theta)e^{s_1'\theta}\hat{\Gamma}_2(d\theta),$$

where $\hat{\Gamma}_1(\cdot)$ is a probability measure on $\bar{\Theta}_1$ and $\hat{\Gamma}_2(\cdot)$ is a σ-finite measure on $\bar{\Theta}_2$.

REMARK 4.1. Note when $\bar{\Theta} \subset \mathfrak{N}$, then the "posterior minimizer" relative to (Ψ_1, Ψ_2) may also be described as the generalized Bayes test relative to $(\hat{\Gamma}_1, \hat{\Gamma}_2)$

where $\hat{\Gamma}_1 = \Psi_1$ and $\hat{\Gamma}_2(d\theta) = C^{-1}(\theta)\Psi_2(d\theta)$. The following is an example of a "posterior minimizer" which is not obviously generalized Bayes. Let X_i be independent, identically distributed random variables with an exponential density $f_\theta(x) = \theta e^{-x\theta}$, $x > 0$. Let $\overline{\Theta}_1 = \{1\}$, $\overline{\Theta}_2 = [0, \infty)$, $\Psi_2 = \varepsilon\{0\}$, $\Psi_1 = \varepsilon\{1\}$ and $c_j \equiv c$ where c is sufficiently small so that the test is truly sequential. The obvious construction of an equivalent generalized Bayes test would involve setting $\hat{\Gamma}_2(d\theta) = C^{-1}(\theta)\Psi_2(d\theta)$. Since $C(0) = 0$ such a $\hat{\Gamma}_2$ does not make sense.

Next we prove

THEOREM 4.2. *Let X be distributed according to* (4.1). *Let $c_1 = 0$ and let $\overline{\Theta}_2 \subset \mathfrak{N}$. Then the class $\overline{B}$ is a complete class. The integrated risk of any procedure in $\overline{B}$ is finite.*

PROOF. Use the same argument in the proof of Theorem 4.1, letting $\Psi_1 = \hat{\Gamma}_1$ and $\Psi_2(d\theta) = C(\theta)\hat{\Gamma}_2(d\theta)$, (since $C(\theta)$ is now well defined on all $\theta \in \overline{\Theta}_2$), to establish that $\overline{B}$ is an essentially complete class. The statement concerning the shape of A also follows from Theorem 4.1. Hence we need to prove that the integrated risk is finite and $\overline{B}$ is complete.

We proceed to show that the integrated risk of the limiting procedure is finite. The integrated risk of the procedure is less than or equal to the integrated risk of the procedure which stops after observing x_1 and then rejects if $s_1 \in \overline{A}^c$, while for $s_1 \in A$, the procedure rejects, accepts, or randomizes between acceptance and rejection, depending on whether $\Delta(s_1) > 0$, < 0, $= 0$, where

$$\Delta(s_1) = d_2 \int_{\overline{\Theta}_2} e^{s_1'\theta} C(\theta)\hat{\Gamma}_2(d\theta) - d_1 \int_{\overline{\Theta}_1} e^{s_1'\theta} C(\theta)\hat{\Gamma}_1(d\theta). \tag{4.11}$$

For this latter procedure the integrated risk is bounded by

$$d_1 \int_{\overline{\Theta}_1} \int e^{s_1'\theta} C(\theta)\mu(ds_1)\hat{\Gamma}_1(d\theta) + d_2 \int_{\overline{\Theta}_2} \int_{A \cap \{\Delta(s_1) < 0\}} e^{s_1'\theta} C(\theta)\mu(ds_1)\hat{\Gamma}_2(d\theta). \tag{4.12}$$

The first of the two terms in (4.12) is finite since $\hat{\Gamma}_1(\cdot)$ is a probability measure. For the second term use Fubini's theorem and (4.11) to find

$$\begin{aligned} &\int_{A \cap \{\Delta(s_1) < 0\}} \left\{ \int_{\overline{\Theta}_2} e^{s_1'\theta} C(\theta)\hat{\Gamma}_2(d\theta) \right\} \mu(ds_1) \\ &\qquad \leqslant (d_1/d_2) \int_{A \cap \{\Delta(s_1) < 0\}} \left\{ \int_{\overline{\Theta}_1} e^{s_1'\theta} C(\theta)\hat{\Gamma}_1(d\theta) \right\} \mu(ds_1) < \infty. \end{aligned} \tag{4.13}$$

Thus (4.12) is finite and so is the integrated risk of the procedure lying in the closure of the class of Bayes procedures.

We have already established that the procedures described in the theorem form an essentially complete class. We next show that they form a complete class. Let $\delta'(\overline{\mathbf{x}})$ be a procedure outside the class such that $\delta'(\overline{\mathbf{x}})$ is admissible. Then there is a procedure $\delta(\overline{\mathbf{x}})$ in the class, which is admissible, such that

$$R(\theta, \delta') = R(\theta, \delta). \tag{4.14}$$

Furthermore $\delta(\overline{\mathbf{x}})$ is the regular limit of some sequence $\delta(k; \overline{\mathbf{x}})$, where $\delta(k, \overline{\mathbf{x}})$, is Bayes with respect to $\Gamma_k = (\Gamma_{1k}, (\pi_{2k}/\pi_{1k})\Gamma_{2k})$, and

$$\int [R(\theta, \delta) - R(\theta, \delta(k))]\Gamma_k(d\theta) \to 0 \text{ as } k \to \infty. \tag{4.15}$$

The above assertions follow by arguments similar to those given in the proof of the Stein-Le Cam Theorem (See Farrell [4].) Consider

$$(4.16)\quad \int[R(\theta,\delta')-R(\theta,\delta(k))]\Gamma_k(d\theta) = \int\int_{\overline{\Theta}}[L(\theta,\delta')-L(\theta,\delta(k))]\Gamma_k(d\theta)P_\theta(d\bar{\mathbf{x}}),$$

the interchange of order of integration being permissible since Γ_k is finite. By (4.14) and (4.15), (4.16) $\to 0$ as $k\to\infty$. Rewrite (4.16) as

$$(4.17)\quad \int_{\mathfrak{X}_1}\int_{\mathfrak{X}^*}\int_{\overline{\Theta}}[L(\theta,\delta')-L(\theta,\delta(k))]\Gamma_k(d\theta)P_\theta(d\bar{\mathbf{x}}^*)e^{s_1'\theta}C(\theta)\mu(ds_1).$$

Let A be the convex set corresponding to the procedure $\delta(\bar{\mathbf{x}})$. That is, if

$$s_1(x_1)\in \bar{A}^c,\ \int_{\overline{\Theta}_2}e^{s_1'\theta}C(\theta)(\pi_{2k}/\pi_{1k})\Gamma_{2k}(d\theta)\to\infty.$$

Let $W=\{s_1:\delta'(s_1)=(\delta_{01}'(s_1),\delta_{11}'(s_1),\delta_{21}'(s_1))$ is such that either $\delta_{01}'(s_1)>0$, or $\delta_{11}'(s_1)>0\}$. Suppose $\mu(W\cap\bar{A}^c)>0$. Then note, since $\delta(k;(\bar{\mathbf{x}}))$ is Bayes that (4.17) is greater than or equal to

$$(4.18)\quad \int_{W\cap\bar{A}^c}\int_{\mathfrak{X}^*}\int_{\overline{\Theta}_1}[L(\theta,\delta')-L(\theta,\delta(k))]\Gamma_{1k}(d\theta)P_\theta(d\bar{\mathbf{x}}^*)e^{s_1'\theta}C(\theta)\mu(ds_1)$$
$$+\int_{W\cap\bar{A}^c}\int_{\mathfrak{X}^*}\int_{\overline{\Theta}_2}[L(\theta,\delta')-L(\theta,\delta(k))](\pi_{2k}/\pi_{1k})\Gamma_{2k}(d\theta)P_\theta(d\bar{\mathbf{x}}^*)e^{s_1'\theta}C(\theta)\mu(ds_1).$$

The first term in (4.18) is bounded for all k. Also

$$\int_{W\cap\bar{A}^c}\int_{\mathfrak{X}^*}\int_{\overline{\Theta}_2}L(\theta,\delta(k))(\pi_{2k}/\pi_{1k})\Gamma_{2k}(d\theta)P_\theta(d\bar{\mathbf{x}}^*)e^{s_1'\theta}C(\theta)\mu(ds_1)$$
$$<\int_{\mathfrak{X}_1}\int_{\mathfrak{X}^*}\int_{\overline{\Theta}_1}L(\theta,\delta(k))\Gamma_{1k}(d\theta)P_\theta(d\bar{\mathbf{x}}^*)e^{s_1'\theta}C(\theta)\mu(ds_1)<d_1,$$

since one could stop at stage one and reject for every x_1. On the other hand

$$\int_{W\cap\bar{A}^c}\int_{\mathfrak{X}^*}\int_{\overline{\Theta}_2}L(\theta,\delta')(\pi_{2k}/\pi_{1k})\Gamma_{2k}(d\theta)P_\theta(d\bar{\mathbf{x}}^*)e^{s_1'\theta}C(\theta)\mu(ds_1)$$
$$>c_2\int_{W\cap\bar{A}^c}\int_{\overline{\Theta}_2}\int_{\mathfrak{X}^*}(\pi_{2k}/\pi_{1k})\Gamma_{2k}(d\theta)P_\theta(d\bar{\mathbf{x}}^*)e^{s_1'\theta}C(\theta)\mu(ds_1)\to\infty.$$

Thus the second term in (4.18) tends to ∞. This contradicts the fact that (4.16) $\to 0$ unless $\delta'(\bar{\mathbf{x}})$ rejects whenever $s_1\in\bar{A}^c$. Hence suppose $\delta'(\bar{\mathbf{x}})$ rejects whenever $s_1\in\bar{A}^c$. Then conditionally on the event $\{s_1: s_1\in A\}$, $\delta'(\bar{\mathbf{x}})$ has the same risk as $\delta(\bar{\mathbf{x}})$. Since δ $(\bar{\mathbf{x}})$ is generalized Bayes at s_1 for $s_1\in A$, and since the integrated risk is finite it follows that $\delta'(\bar{\mathbf{x}})$ would also be generalized Bayes at s_1 for $s_1\in A$. Thus $\delta'(\bar{\mathbf{x}})$ would in fact be in the stated class of procedures. Thus the class of procedures is in fact a complete class and this completes the proof of the theorem.

REMARK 4.2. The analogues of Theorem 3.3 and Corollary 3.1 can be given for the model of this section.

5. One dimensional exponential family-one-sided hypotheses. In this section X is a one dimensional random variable whose distribution is (4.1) with μ a non-atomic probability measure. The sufficient transitive sequence has distribution (4.2) with γ absolutely continuous with respect to Lebesgue measure. We no longer require that $\overline{\Theta}_1$ be compact. We do assume though that every $\theta\in\overline{\Theta}_1$ is less than or equal to every $\theta\in\overline{\Theta}_2$. That is, the hypotheses are one-sided. Also we require that either $\bar{\Theta}_1$, $\bar{\Theta}_2$ or both are contained in $\mathfrak{N}$. For this model we can study the cases where $c_1 = 0$ and $c_1 > 0$. If $c_1 > 0$, we do have the option of stopping at time zero, or of randomizing at time zero, and this is decided for the limit of a sequence of Bayes procedures, depending on the limiting behavior of (π_{1k}, π_{2k}). The treatment of this case is similar to the case where $C_1 = 0$. Hence let us treat the case where $c_1 = 0$, both $\bar{\Theta}_1$ and $\bar{\Theta}_2$ are contained in $\mathfrak{N}$, and there exists a positive number K such that $[-K, K] \subset \mathfrak{N}$. The following procedures are said to lie in the class $\bar{B}$. At stage 1, there is an interval (a_1, a_2), $(-\infty \leqslant a_1 \leqslant a_2 \leqslant \infty)$ such that if $s_1 < a_1$, the procedure stops and accepts H_1; if $s_1 < a_2$, the procedure stops and rejects H_1; if $a_1 < s_1 < a_2$, then the procedure is generalized Bayes at s_1 with respect to a distribution $\tilde{\Gamma} = (\tilde{\Gamma}_1, \tilde{\Gamma}_2)$ where $\tilde{\Gamma}_i$, $i = 1, 2$, are σ-finite measures on $\bar{\Theta}_i$ respectively. We will show that $\bar{B}$ is a complete class of procedures. To start, let $\Gamma_k = (\pi_{1k}, \Gamma_{1k}(\cdot), \Gamma_{2k}(1))$ represent a sequence of prior distributions. Let $\delta(k; \bar{\mathbf{x}})$ be the Bayes tests for these priors and let $\delta(\bar{\mathbf{x}})$ be the regular limit of the sequence. Consider

$$\pi_{2k} g^{(1)}_{(2), k}(s_1) = \pi_{2k} \int_{\bar{\Theta}_2} e^{s_1\theta} C(\theta) \Gamma_{2k}(d\theta) \Big/ \Big[\pi_{1k} \int_{\bar{\Theta}_1} e^{s_1\theta} C(\theta) \Gamma_{1k}(d\theta) + \pi_{2k} \int_{\bar{\Theta}_2} e^{s_1\theta} C(\theta) \Gamma_{2k}(d\theta) \Big]. \tag{5.1}$$

Since $g^{(1)}_{(2), k}(\cdot)$ is a monotone function, there exists a subsequence for which $\lim_{k'\to\infty} \pi_{2k'} g^{(1)}_{(2), k'}(s_1)$ exists for all s_1. Let s_1^* be a point for which there exists a subsequence such that

$$\lim_{k'\to\infty} \pi_{2k'} g^{(1)}_{(2), k'}(s_1^*) \neq 0 \text{ or } 1. \tag{5.2}$$

If no such s_1^* exists, then $\delta(\bar{\mathbf{x}})$ stops at stage 1 and accepts or rejects H_1 as the left-hand side of (5.2) is 0 or 1. In this case, by the monotonicity of $\pi_{2k'} g^{(1)}_{(2), k'}(s_1)$, $a_1 = a_2$, and the procedure would be in $\bar{B}$. Hence let s_1^* satisfy (5.2). We state

LEMMA 5.1. *There exist numbers $a_1 \leqslant a_2$ and subprobability measures (Γ_1^*, Γ_2^*), such that if $s_1 < a_1$, $\pi_{2k'} g^{(1)}_{(2), k'}(s_1) \to 0$ as $k' \to \infty$; if $s_1 < a_2$, $\pi_{2k'} g^{(1)}_{(2), k'}(s_1) \to 1$ as $k' \to \infty$; if $a_1 < s_1 < a_2$,*

$$\pi_{2k'} g^{(1)}_{(2), k'}(s_1) \to \int_{\bar{\Theta}_2} e^{(s_1 - s_1^*)\theta} \Gamma_2^*(d\theta) \Big/ \Big[\int_{\bar{\Theta}_1} e^{(s_1 - s_1^*)\theta} \Gamma_1^*(d\theta) + \int_{\bar{\Theta}_2} e^{(s_1 - s_1^*)\theta} \Gamma_2^*(d\theta) \Big]. \tag{5.3}$$

Furthermore if $\Gamma_1^* = 0$, then $a_1 = s_1^*$. If $\Gamma_2^* = 0$, then $a_2 = s_1^*$.

PROOF. The proof is omitted.

Next we state

THEOREM 5.1. *The class of procedures $\bar{B}$ is a complete class.*

PROOF. The proof is omitted.

REMARK 5.1. In the case where either $\overline{\Theta}_1 \subset \mathfrak{N}$ or $\overline{\Theta}_2 \subset \mathfrak{N}$, but not both, an essentially complete class characterization is obtainable using the ideas of the proofs in Theorem 4.1 and Theorem 5.1.

REMARK 5.2. The analogues of Theorem 3.3 and Corollary 3.1 can be given for the models of this section. In fact in this situation the normality assumption can be dropped and we can simply require that μ is a nonatomic measure. The reason is that procedures in the complete class are uniquely determined. (See Brown, Cohen, Strawderman [2] or Sobel [7].)

REMARK 5.3. In Sobel [7] and Brown, Cohen, Strawderman [2], monotone procedures were shown to be an essentially complete class. The result here, in conjunction with those papers, not only proves that the monotone procedures are a complete class but also gives a complete class which is much smaller than the monotone procedures.

Acknowledgments. The authors would like to thank Professor Robert H. Berk for many helpful discussions.

REFERENCES

[1] BROWN, L. D. (1971). Admissible estimators, recurrent diffusions, and in soluble boundary values problems. *Ann. Math. Statist.* **42** 855–903.

[2] BROWN, L. D., COHEN, ARTHUR, STRAWDERMAN, W. E. (1979). Monotonicity of Bayes sequential tests. *Ann. Statist.* **7** 1222–1230.

[3] CHOW, Y. S., ROBBINS, H., and SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Houghton, Mifflin Company.

[4] FARRELL, R. H. (1968). Towards a theory of generalized Bayes tests. *Ann. Math. Statist.* **39** 1–22.

[5] FERGUSON, T. S. (1967). *Mathematical Statistics, A Decision Theoretic Approach*. Academic Press.

[6] LE CAM, L. (1955). An extension of Wald's theory of statistical decision functions. *Ann. Math. Statist.* **26** 69–81.

[7] SOBEL, M. (1952). An essentially complete class of decision functions for certain standard sequential problems. *Ann. Math. Statist.* **23** 319–337.

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903