Statistica Sinica **12**(2002), 1027-1041

COMPUTATIONAL METHODS FOR EVALUATING SEQUENTIAL TESTS AND POST-TEST ESTIMATION VIA THE SUFFICIENCY PRINCIPLE

Xiaoping Xiong, Ming Tan and Michael H. Kutner

St. Jude Children's Research Hospital, University of Maryland and Emory Univeristy

Abstract: By the sufficiency principle, the probability density of a sequential test statistic under certain conditions can be factored into a known function that does not depend on the stopping rule and a conditional probability that is free of unknown parameters. We develop general theorems and propose a unified approach to analyzing and evaluating various properties of sequential tests and post-test estimation. The proposed approach is of practical value since it allows for effective evaluation of properties of special interest, such as the bias-adjustment of post-test estimation after a sequential test, and the probability of discordance between a sequential test and a nonsequential test.

Key words and phrases: Bias-adjusted estimation, eigenvalue function, probability of discordance, sequential clinical trial.

1. Introduction

Sequential hypothesis testing was first developed for use in traditional acceptance sampling and process control, with the goal of improving the efficiency of testing (Wald (1947)). For ethical and economical reasons, sequential testing now plays an important role in the design and analysis of clinical studies. The last two decades have witnessed the development of various sequential and group sequential procedures, some of which came into real applications (see, e.g., DeMets and Lan (1994), Whitehead (1997), Jennison and Turnbull (2000)). The difficulty in evaluating sequential tests and post-test estimation has limited the use of these methods (Siegmund (1985)). Analytical solutions are not generally available, especially for tests with nonlinear and discrete boundaries. Analytical solutions or asymptotic approximations of the classical operating characteristics (e.g., type I and II errors, average sample number (ASN)) have been obtained for sequential procedures of nontruncated linear boundaries (Wald (1947)), truncated linear boundaries (Anderson (1960), Samuel-Cahn (1974)), and classes of nonlinear boundaries (e.g., Lai and Siegmund (1977, 1979), Lai and Wijsman (1979)). Asymptotic approximations are usually difficult and are not generally applicable because the requirement of large sample size contradicts the very goal of early stopping in sequential tests. Numerical methods reported by Aroian (1968) and Armitage, McPherson and Rowe (1969), based on recursive convolutions of sequential probability distributions, provided numerical solutions for evaluating sequential tests. An alternative numerical method was provided by Jennison (1994) involving integration of a multivariate normal density. Although these methods can be used to evaluate basic characteristics of sequential tests, they cannot effectively evaluate the characteristics that have probabilistic complexity. The reader is referred to Lai (2001) for a recent comprehensive survey of sequential analysis.

Many sequential tests have virtually the same type I and II errors; thus, other properties must be examined to select the most appropriate sequential design. Properties of sequential tests other than the classical operating characteristics may also be of interest in practice. For example, a sequential test having a smaller ASN requires a larger (maximum) sample size. Therefore, minimizing ASN should not be the only criterion used to select sequential designs if both the ASN and the (maximum) sample size are to be minimized. It is also of interest to know the *probability of discordance*, which measures the probability that the sequential test does not agree with a nonsequential test to be performed through the planned end of the sequential test, or the probability that the sequential test does not agree with a nonsequential test of equivalent significance level and power. Such properties are of special interest in sequential clinical trials, where common sense dictates that sampling be stopped early if the interim conclusion is unlikely to be reversed should the trial have continued to its planned end. In general, to select a sequential test with the desired characteristics and to make valid inference, we must evaluate the properties of sequential designs and the post-test estimation of parameters. In this paper, using the sufficiency principle, we develop theorems for a class of sequential tests, based on which a unified approach is proposed to effectively evaluate various characteristics of interest.

In Section 2, we introduce a sufficiency identity and give a general equation for evaluating sequential procedures. We give recursive formulas for computing the fundamental *eigenvalue function* $l(n, s)$ which is free of unknown parameters. In Section 3, we develop methods for obtaining various characteristics of sequential procedures, such as the bias-adjusted estimate and its expectation, and the probability of discordance.

2. Main Results

The following setup is considered throughout this paper. Let $X_1, \ldots,$ be a sequence of variables, not necessarily independent or identically distributed. Assume that for any n, the joint distribution of $X_1, \ldots, X_n$ depends on the parameter θ, and that $S_n = g(X_1, \ldots, X_n)$ is a sufficient statistic for θ. Here X_n,

S_n, and θ can be scalars or vectors. Let N be a stopping time, i.e., for any n the event $(N \leq n)$ depends only on $X_1, \ldots, X_n$, not on $X_{n+1}, X_{n+2}, \ldots$. Assume $P_\theta(N < \infty) = 1$ and let $n_1 < n_2 < \cdots$ be such that, for each n_k, $P_\theta(N = n_k) > 0$ for some θ, and $P_\theta(N \notin \{n_1, n_2, \ldots\}) = 0$ for all θ.

2.1. Fundamental identity

Theorem 2.1. *Let $p_\theta(n, s) \equiv P_\theta(N = n, S_N = s)$ (or $\lim_{\Delta s \to 0} P_\theta(N = n, s < S_n < s + \Delta s)/\Delta s$) be the probability mass (or density) function of test statistic (N, S_N). Then for any (n, s), the mass (or density) function $p_\theta(n, s)$ can be factored as*

$$p_\theta(n, s) = f_\theta(n, s) l(n, s), \tag{1}$$

where $f_\theta(n, s) \equiv f_\theta^n(s) = P_\theta(S_n = s)$ and

$$l(n, s) = P(N = n | S_n = s), \tag{2}$$

which does not depend on θ.

Proof. For any (n, s) in the support of (N, S_N) (i.e., $P_\theta(N = n, S_N = s) > 0$) we have $P_\theta(N = n, S_N = s) = P_\theta(N = n, S_n = s) = P_\theta(S_n = s)P(N = n|S_n = s)$, because $S_N(\omega) = S_n(\omega)$ for any $\omega \in (N = n)$; hence (1) holds. For any (n, s) not in the support of (N, S_N) (i.e., in continuation region), $(N = n) \cap (S_n = s) = \emptyset$; thus, $p_\theta(n, s) = 0$ and $l(n, s) = 0$, which indicates that (1) still holds. Because S_n is a sufficient statistic for θ by assumption, then for any event A measurable to sigma field $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$, the conditional probability $P_\theta(A|S_n = s)$ does not depend on θ. In particular, the event $(N = n)$ is measurable to $\mathcal{F}_n$ by the definition of N and $l(n, s) = P(N = n|S_n = s)$ does not depend on θ.

Although this derivation is relatively straightforward using the well known sufficiency principle, (1) can lead to development of novel and effective methods for evaluating sequential tests and post-test estimation when combined with numerical algorithms. This approach was first proposed in Xiong (1991), and further developed in Xiong (1992, 1996). Here we establish a general framework upon which we develop methods applicable for solving difficult problems. The importance of (1) is that $f_\theta(n, s)$ depends on θ but not on the stopping rule, and that the conditional probability $l(n, s)$ depends on the stopping rule but not on θ when S_n is a sufficient statistic for θ. We refer to $l(n, s)$ as the eigenvalue function of the sequential statistic since it plays a role similar to the eigenvalue in linear algebra. From Theorem 2.1 we get the following corollary, by which calculation involving the sequential statistic (N, S_N) becomes one involving only non-sequential statistics $\{S_{n_k}\}_{k \geq 1}$.

Corollary 2.1. *For any function $H(n,s)$, if $E_\theta[H(n_k,S_{n_k})]$ exists for any possible values n_k of N, then*

$$h(\theta) \equiv E_\theta\{H(N,S_N)\} = \sum_{n_k} E_\theta\{H(n_k,S_{n_k})l(n_k,S_{n_k})\}, \tag{3}$$

where $l(n,s)$ is defined by (2)*. Moreover,*

$$h'(\theta) = E_\theta\left[H(N,S_N)\frac{\partial}{\partial\theta}\{\log f_\theta(N,S_N)\}\right]. \tag{4}$$

Equation (3) is a direct consequence of (1). Because $\frac{\partial}{\partial\theta}\{\log p_\theta(N,S_N)\} = \frac{\partial}{\partial\theta}\{\log f_\theta(N,S_N)\}$ by (1), we have (4), which can be evaluated by using (3) with $H^*(n,s) \equiv H(n,s)\frac{\partial}{\partial\theta}\{\log f_\theta(n,s)\}$. Having $h'(\theta)$ available is useful for solving the equation $h(\theta) = const$ using Newton-Ralphson, an application of which is given in Section 3.2.

2.2. Eigenvalue function

For practical computation, we exclude those points that can never be passed or reached by (N,S_N), which motivates the following definition.

Definition 2.1. For a sequential statistic (N,S_N), the stopping region and continuation region are defined, respectively, as $\mathcal{B} = \{(n,s) : P_\theta(N = n, S_n = s) > 0 \text{ for some } \theta\}$ and

$$\mathcal{C} = \{(n,s) : P_\theta(N > n, S_n = s) > 0 \text{ for some } \theta\}, \tag{5}$$

where $P_\theta(N = n, S_n = s)$ and $P_\theta(N > n, S_n = s)$ are interpreted, in the continuous case, as $\lim_{\Delta s\to 0} P_\theta(N = n, s < S_n < s+\Delta s)/\Delta s$ and $\lim_{\Delta s\to 0} P_\theta(N > n, s < S_n < s+\Delta s)/\Delta s$, respectively.

As $S_n = g(X_1,\ldots,X_n)$ is sufficient for θ, we restrict attention to those stopping rules N for which $\{N = n\}$ depends on $X_1,\ldots,X_n$ only through S_n given $\{N > n-1\}$. Such a stopping rule N (assuming values of $n_1, n_2, \ldots$ only) may be identified with a stopping set $\mathcal{B} = \cup_k\{n_k\}\times B_{n_k}$, where B_{n_k} is a subset of the range of S_{n_k}; more precisely $N = \inf\{n_k : (n_k, S_{n_k}) \in B\}$, the first hitting time of $\mathcal{B}$. Note that $\mathcal{B}\cap\mathcal{C} = \emptyset$ for this N and

$$P_\theta(N > n_{k-1}, S_{n_k} = s) = \begin{cases} P_\theta(N = n_k, S_{n_k} = s) & \text{if } (n_k,s)\in\mathcal{B} \\ P_\theta(N > n_k, S_{n_k} = s) & \text{if } (n_k,s)\in\mathcal{C}. \end{cases} \tag{6}$$

In addition, we assume for S_{n_k}'s that, for any event $A \in \sigma(S_{n_1},\ldots,S_{n_k})$,

$$P(A|S_{n_k}, S_{n_{k+i}}) = P(A|S_{n_k}) \quad \text{for any } k, i \ge 1. \tag{7}$$

It can be shown that (7) holds if the S_{n_k}'s form a Markov sequence.

Lemma 2.1. *Assume* (7) *holds for the* S_{n_k}*'s. For any* (n, s)*, let*

$$l^*(n, s) \equiv P(N > n | S_n = s). \tag{8}$$

Then for any $(n_k, s) \in \mathcal{B}$*,* $l^*(n_k, s) = 0$*. For any* $(n_k, s) \in \mathcal{C}$*, if* $k = 1$ *then* $l^*(n_1, s) \equiv 1$*; if* $k \geq 2$ *then*

$$l^*(n_k, s) = E\{l^*(n_{k-1}, S_{n_{k-1}}) | S_{n_k} = s\}. \tag{9}$$

Proof. For any $(n_k, s) \in \mathcal{B}$, if $l^*(n_k, s) > 0$, then $P_\theta(N > n_k, S_{n_k} = s) > 0$ for some θ, which contradicts $\mathcal{B} \cap \mathcal{C} = \emptyset$. Hence $l^*(n, s) = 0$. For $(n_k, s) \in \mathcal{C}$ and $k = 1$, $P_\theta(N = n_1, S_{n_1} = s) = 0$ for any θ and thus $P_\theta(N > n_1, S_{n_1} = s) = P_\theta(N \geq n_1, S_{n_1} = s) = P_\theta(S_{n_1} = s)$, which gives $l^*(n_1, s) = P(N > n_1 | S_{n_1} = s) = 1$. For $(n_k, s) \in \mathcal{C}$ and $k \geq 2$,

$$l^*(n_k, s) = E(1_{(N>n_{k-1})} | S_{n_k} = s) = E[E\{1_{(N>n_{k-1})} | S_{n_{k-1}}, S_{n_k}\} | S_{n_k} = s] \tag{10}$$

by (8) and (6), and by Theorem 34.4 in Billingsley (1986). By (7), we have $E\{1_{(N>n_{k-1})} | S_{n_{k-1}}, S_{n_k}\} = E\{1_{(N>n_{k-1})} | S_{n_{k-1}}\} = l^*(n_{k-1}, S_{n_{k-1}})$, and then (10) yields (9).

Theorem 2.2. *Assume* (7) *holds for the* S_{n_k}*'s. For any* $(n_k, s) \in \mathcal{C}$*,* $l(n_k, s) = 0$*. For any* $(n_k, s) \in \mathcal{B}$*, if* $k = 1$*, then* $l(n_1, s) \equiv 1$*; if* $k \geq 2$*, then*

$$l(n_k, s) = 1 - \sum_{i=1}^{k-1} E\{l(n_i, S_{n_i}) | S_{n_k} = s\}, \tag{11}$$

$$l(n_k, s) = E\{l^*(n_{k-1}, S_{n_{k-1}}) | S_{n_k} = s\}. \tag{12}$$

Proof. For any $(n_k, s) \in \mathcal{C}$, we have $P_\theta(N = n_k, S_{n_k} = s) = 0$ because $(n_k, s) \notin \mathcal{B}$, which implies $l(n_k, s) = 0$. If $(n_1, s) \in \mathcal{B}$, then $P_\theta(N > n_1, S_{n_1} = s) = 0$ for any θ because $(n_1, s) \notin \mathcal{C}$. Hence $P_\theta(N = n_1, S_{n_1} = s) = P_\theta(S_{n_1} = s)$, which gives $l(n_1, s) = 1$. For $(n_k, s) \in B$ and $k \geq 2$,

$$l(n_k, s) = P(N = n_k | S_{n_k} = s) = 1 - \sum_{i=1}^{k-1} E\{1_{(N=n_i)} | S_{n_k} = s\}. \tag{13}$$

As in Lemma 2.1, we have $E\{1_{(N=n_i)} | S_{n_k} = s\} = E\{E(1_{(N=n_i)} | S_{n_i}, S_{n_k}) | S_{n_k} = s\}$ and $E(1_{(N=n_i)} | S_{n_i}, S_{n_k}) = E(1_{(N=n_i)} | S_{n_i}) = l(n_i, S_{n_i})$, by which (13) yields (11). On the other hand,

$$\begin{aligned} l(n_k, s) &= P(N = n_k | S_{n_k} = s) = P\{N > n_{k-1} | S_{n_k} = s\} \\ &= E[E\{1_{(N>n_{k-1})} | S_{n_{k-1}}\} | S_{n_k} = s] \end{aligned} \tag{14}$$

by (6) and (7). By the definition of $l^*(n_{k-1}, S_{n_{k-1}})$, (14) yields (12).

The density of (N, S_N), $p_\theta(n, s)$, which can be obtained by (2) when $l(n, s)$ is known, can also be obtained directly from $l^*(n, s)$ when S_{n_k} has independent increment, as in the theorem below.

Theorem 2.3. *Suppose that for $k \geq 2$, $S_{n_k} - S_{n_{k-1}}$ is independent of $\mathcal{F}_{n_{k-1}}$, the sigma-field generated by $X_1, \ldots, X_{n_{k-1}}$, and let $q_\theta(d; n_{k-1}, n_k) \equiv P_\theta(S_{n_k} - S_{n_{k-1}} = d | S_{n_{k-1}}) = P_\theta(S_{n_k} - S_{n_{k-1}} = d)$. Then for $(n_k, s) \in \mathcal{B}$, the density of (N, S_N) at (n_k, s) is*

$$p_\theta(n_k, s) = E_\theta\{l^*(n_{k-1}, S_{n_{k-1}}) q_\theta(s - S_{n_{k-1}}; n_{k-1}, n_k)\}. \tag{15}$$

Proof. $(N > n_{k-1}) = (N \leq n_{k-1})^c$ and is measurable to $\mathcal{F}_{n_{k-1}}$, and thus independent of $S_{n_k} - S_{n_{k-1}}$ by assumption. For $k \geq 2$ and $(n_k, s) \in \mathcal{B}$, by (6),

$$\begin{aligned} p_\theta(n_k, s) &= P_\theta(N > n_{k-1}, S_{n_k} = s) = E_\theta\{P_\theta(N > n_{k-1}, S_{n_k} - S_{n_{k-1}} = s - S_{n_{k-1}} | S_{n_{k-1}})\} \\ &= E_\theta\{P_\theta(N > n_{k-1} | S_{n_{k-1}}) P_\theta(S_{n_k} - S_{n_{k-1}} = s - S_{n_{k-1}} | S_{n_{k-1}})\}. \end{aligned} \tag{16}$$

Then we obtain (15) by joining (16) and $P(N > n_{k-1} | S_{n_{k-1}}) = l^*(n_{k-1}, S_{n_{k-1}})$ and $P_\theta(S_{n_k} - S_{n_{k-1}} = s - S_{n_{k-1}} | S_{n_{k-1}}) = q_\theta(s - S_{n_{k-1}}; n_{k-1}, n_k)$.

Corollary 2.2. *Suppose that $S_{n_k} - S_{n_{k-1}}$ is independent of $\mathcal{F}_{n_{k-1}}$. Let $v_\theta(n_k, s) = P_\theta(N > n_k, S_{n_k} = s)$ and $\mathcal{C}_{n_k} = \{s : (n_k, s) \in \mathcal{C}\}$. We have*

$$p_\theta(n_k, s) = \int_{t \in \mathcal{C}_{n_{k-1}}} v_\theta(n_{k-1}, t) q_\theta(s - t; n_{k-1}, n_k) dt, \quad \textit{for } (n_k, s) \in \mathcal{B}, \tag{17}$$

$$v_\theta(n_k, s) = \int_{t \in \mathcal{C}_{n_{k-1}}} v_\theta(n_{k-1}, t) q_\theta(s - t; n_{k-1}, n_k) dt, \quad \textit{for } (n_k, s) \in \mathcal{C}. \tag{18}$$

Proof. Because $v_\theta(n_k, s) = l^*(n_k, s) f_\theta(n_k, s)$, we have (17) by (15), and (18) follows by multiplying (9) by $f_\theta(n_k, s)$.

2.2.1. Special cases: the mean parameter

Assume that X_i, $i = 1, 2, \ldots$, are observations from a population with mean $\theta = E_\theta(X_i)$ for any i, and that $S_n = \sum_{i=1}^n X_i$ is the partial sum. For testing

$$H_0 : \theta \leq \theta_0 \quad \text{vs.} \quad H_a : \theta > \theta_0 \tag{19}$$

sequentially, the statistic (N, S_N) usually has stopping and continuation regions, respectively, as

$$\begin{aligned} \mathcal{B} &= \{(n, s) : \ n = n_k, \ s \leq b_{n_k} \text{ or } s \geq a_{n_k} \quad \text{for} \quad k = 1, 2, \ldots\}, \\ \mathcal{C} &= \{(n, s) : n = n_k, \ b_{n_k} < s < a_{n_k} \quad \text{for} \quad k = 1, 2, \ldots\}, \end{aligned} \tag{20}$$

where the a_{n_k}'s and b_{n_k}'s are known constants, $-\infty \le b_{n_k} \le a_{n_k} \le \infty$.

For $n_j < n_k$, the *conditional distribution* of S_{n_j} given S_{n_k} will be denoted as

$$p_{n_j|n_k}(t|s) \equiv P(S_{n_j}=t|S_{n_k}=s). \tag{21}$$

This does not depend on θ if S_{n_k} is a sufficient statistic for θ.

Binomial or Hypergeometric distribution. Assume that $S_n \sim B(n,p)$ or $\sim \mathcal{H}(n; M = pN, N)$. Then

$$p_{n_j|n_k}(t|s) = h_{n_j|n_k}(t|s) \equiv \frac{\binom{s}{t}\binom{n_k-s}{n_j-t}}{\binom{n_k}{n_j}}. \tag{22}$$

Normal distribution. Assume $S_n \sim N(n\mu, n\sigma^2)$ for any n. Then

$$p_{n_j|n_k}(t|s) = \phi_{n_j|n_k}(t|s) \equiv \frac{\phi\left((t - \frac{n_j}{n_k}s)/\sqrt{n_j(1-\frac{n_j}{n_k})}\sigma\right)}{\sqrt{n_j(1-\frac{n_j}{n_k})}\sigma}, \tag{23}$$

where $\phi(x)$ is the density function of the standard normal distribution.

Poisson distribution. Assume $S_n \sim \mathcal{P}(n\lambda)$ for any n. Then

$$p_{n_j|n_k}(t|s) = b_{n_j|n_k}(t|s) \equiv \binom{s}{t}\left(\frac{n_j}{n_k}\right)^t\left(1-\frac{n_j}{n_k}\right)^{s-t}. \tag{24}$$

Example 2.1. Let N be the first exit time to stopping region $\mathcal{B}$ in (20). If $S_n \sim N(n\theta, n\sigma^2)$ for any n, then the conditional expectation in (11) can be calculated by the conditional distribution in (23). In the continuation region, or $s \in (b_{n_k}, a_{n_k})$, we have $l(n_k, s) = 0$ by Theorem 2.2. In the stopping region, or $s \in (b_{n_k}, a_{n_k})^c$,

$$l(n_k, s) = 1 - \sum_{i=1}^{k-1}\left\{\left(\int_{a_{n_i}}^{\infty} + \int_{-\infty}^{b_{n_i}}\right) l(n_i, t)\phi_{n_i|n_k}(t|s)dt\right\}$$

by (11) and (23). Similarly, by (12) and (23), $l(n_k, s) = \int_{b_{n_{k-1}}}^{a_{n_{k-1}}} l^*(n_{k-1}, t)\,\phi_{n_{k-1}|n_k}(t|s)dt$, where, for $t \in (b_{n_{k-1}}, a_{n_{k-1}})$, $l^*(n_{k-1}, t) = \int_{b_{n_{k-2}}}^{a_{n_{k-2}}} l^*(n_{k-2}, \tau)\,\phi_{n_{k-2}|n_{k-1}}(\tau|t)d\tau$ by (9) and (23).

Remark 2.1. By (1), the density of (N, S_N) for any θ is related to that for a given θ_0 by

$$p_\theta(n, s) = p_{\theta_0}(n, s)r(n, s; \theta, \theta_0), \tag{25}$$

where $r(n, s; \theta, \theta_0) = \frac{f_\theta(n,s)}{f_{\theta_0}(n,s)}$. If S_n is the sum of i.i.d. X_i's, then for $S_n \sim N(n\theta, n\sigma^2)$, $r(n, s; \theta, \theta_0) = \exp\{\frac{2s(\theta-\theta_0)-n(\theta^2-\theta_0^2)}{2\sigma^2}\}$; for $S_n \sim B(n, p)$, $r(n, s; p, p_0) = \left(\frac{p}{p_0}\right)^s\left(\frac{1-p}{1-p_0}\right)^{n-s}$; for $S_n \sim \mathcal{P}(n\lambda)$, $r(n, s; \lambda, \lambda_0) = \left(\frac{\lambda}{\lambda_0}\right)^s \exp\{-n(\lambda - \lambda_0)\}$.

Remark 2.2. The density of (N, S_N) can be obtained by four methods: First by (1), in which $l(n,s)$ can be obtained by (12); second by (15), in which $l^*(n,s)$ can be obtained by (9); third by (25), in which $p_{\theta_0}(n,s)$ can be obtained by (17) (or other methods, e.g., multivariate normal integration); and fourth, directly by equation (17). The first method does not require that $S_{n_k} - S_{n_{k-1}}$ be independent of $S_{n_{k-1}}$ for all n_ks of N, whereas the remaining methods do. The fourth method is traditional, and first used by Aroian (1968) for the binomial distribution and by Armitage (1969) for the normal distribution. However, because $v_\theta(n,s)$ in (17) depends on θ, the convolution (18) for $v_\theta(n,s)$ has to be evaluated for each θ. The third method improves on the fourth method by utilizing the relationship between $p_\theta(n,s)$ and $p_{\theta_0}(n,s)$, which has been used by Emerson and Fleming (1990) for the normal distribution. The ratio of the two densities in (25) is $r(n_k, s; \theta, 0) = \exp(s\theta - n_k\theta^2/2)$ (assume $\theta_0 = 0$ and $\sigma^2 = 1$), by which a small round-off error in computing $p_0(n_k, s)$ could result in a major error for $p_\theta(n_k, s)$. For example, in testing (19) at significance level 0.025 with 0.95 power to detect an alternative of $\theta = 0.5$, we have $p_{0.5}(22, 20.5)/p_0(22, 20.5) = 1808$ by (25) on point $(n_k, s) = (22, 20.5)$, which indicates that a round-off error for $p_0(n_k, s)$ would be amplified 1808 times for $p_\theta(n_k, s)$. As s increases, such an error increases quickly (e.g., if $s = 30$, the error is amplified $208,981$ times). This error could be fatal for those (n_k, s) at which $p_0(n_k, s)$ is close to 0 and $p_\theta(n_k, s)$ is large. In the first and second methods, $l(n_k, s)$ and $l^*(n_k, s)$ are between 0 and 1 and do not depend on θ, and $l(n_k, s)$ is small if and only if $p_\theta(n_k, s)$ is small for all θ. Thus, evaluation of $p_\theta(n_k, s)$ is more accurate and effective by the first and the second methods than by the third and fourth methods.

3. Evaluation of Sequential Tests and Post-Test Estimation

3.1. Classical operating characteristics

Assume N is the first exit time to $\mathcal{B}$ in (20) for testing the hypotheses in (19). Setting $H(n,s) = 1_{(s > a_n)}$ in (3) yields the *power function* of the sequential test

$$\beta(\theta) = P_\theta(S_N \geq a_N) = E_\theta\left\{1_{(S_N \geq a_N)}\right\} = \sum_{n_k} E_\theta\left\{1_{(S_{n_k} > a_{n_k})} l(n_k, S_{n_k})\right\}. \quad (26)$$

Setting $H(n,s) = n$ in (3) yields the *expected sample* size (average sample number (ASN), traditionally)

$$ASN(\theta) = E_\theta(N) = \sum_{n_k} E_\theta\left\{n_k l(n_k, S_{n_k})\right\}. \quad (27)$$

Since the distribution of S_{n_k} is known and $l(n_k, s)$ can be evaluated based on theorems in Section 2.2, calculation in (26) and (27) is straightforward.

3.2. Estimation after sequential tests

After a sequential test, θ is usually estimated by the maximum likelihood estimator $\hat{\theta}_{ml} = \hat{\theta}_{ml}(N, S_N)$ (e.g., $\hat{\theta}_{ml} = S_N/N$ in the normal case) which is well-known to be biased. A bias-adjusted estimator was suggested by Whitehead (1986) and by Chang, Wieand and Chang (1989). We propose to evaluate the bias-adjusted estimation by using the eigenvalue function, and demonstrate its efficiency in this application.

Let $B(\theta) = E_\theta\{\hat{\theta}_{ml}(N, S_N)\} - \theta$ be the bias of the estimator $\hat{\theta}_{ml}(N, S_N)$ for a given θ. An unbiased estimate of θ would be $(\hat{\theta}_{ml})_{observed} - B(\theta)$ which depends on the unknown θ. This motivates the equation $\tilde{\theta} = (\hat{\theta}_{ml})_{observed} - B(\tilde{\theta})$ that determines a bias-adjusted estimator $\tilde{\theta}$. By simple algebra, $\tilde{\theta}$ is the solution of the following equation for θ:

$$E_\theta(\hat{\theta}_{ml}) = (\hat{\theta}_{ml})_{observed}, \tag{28}$$

which may be viewed as estimation by the method of moments. Equation (28) can be solved for θ by straightforward computation by incorporating the eigenvalue function with the Newton-Ralphson method. By (3), the expectation of MLE as a function of θ is

$$h(\theta) = E_\theta\{\hat{\theta}_{ml}(N, S_N)\} = \sum_{n_k} E_\theta\left\{\hat{\theta}_{ml}(n_k, S_{n_k}) l(n_k, S_{n_k})\right\}. \tag{29}$$

By (3) and (4),

$$h'(\theta) = \sum_{n_k} E_\theta\left[\hat{\theta}_{ml}(n_k, S_{n_k}) \frac{\partial}{\partial\theta}\{\log f_\theta(n_k, S_{n_k})\} l(n_k, S_{n_k})\right]. \tag{30}$$

The bias-adjusted estimate, or the solution of equation (28), is the limit of $\tilde{\theta}_i$ where

$$\tilde{\theta}_i = \tilde{\theta}_{i-1} - \frac{h(\tilde{\theta}_{i-1}) - (\hat{\theta}_{ml})_{observed}}{h'(\tilde{\theta}_{i-1})}, \tag{31}$$

$i = 1, 2, \ldots$, and initial value $\tilde{\theta}_0 = (\hat{\theta}_{ml})_{observed}$. The bias-adjusted estimate $\tilde{\theta}$ is a function of $\hat{\theta}_{ml}$, or $\tilde{\theta} = \tilde{\theta}(\hat{\theta}_{ml})$. The bias-adjusted estimate $\tilde{\theta}$ still has a bias, $B^*(\theta) = E_\theta\{\tilde{\theta}(\hat{\theta}_{ml}(N, S_N))\} - \theta$. This tends to be substantially smaller than $B(\theta)$. From (3) we have

$$E_\theta\left\{\tilde{\theta}(\hat{\theta}_{ml}(N, S_N))\right\} = \sum_{n_k} E_\theta\left\{\tilde{\theta}(\hat{\theta}_{ml}(n_k, S_{n_k})) l(n_k, S_{n_k})\right\}. \tag{32}$$

Evaluation of $\tilde{\theta}(\hat{\theta}_{ml}(n_k, s))$ for each (n_k, s) in (32) requires a converging sequence in (31). Therefore, the efficiency for evaluating $B^*(\theta)$ is much improved by repeated use of $l(n_k, s)$ in (29), (30) and (32), as compared with the traditional method (i.e., the fourth method in Remark 2.2).

3.3. Probabilities of discordance

When early stopping of a clinical trial is considered, it is important to find the probability that a different decision would be reached should we have continued to collect data to the end (maximum information time) and then used a nonsequential test. We say that a sequential procedure and a nonsequential test are *comparable* if the probability of discordance $P(D)$ is negligible, where D is the event that the sequential test and the nonsequential test lead to different rejection/acceptance decisions when both are used on the same sequence of observations. Here we develop methods for deriving various probabilities of discordance by using the proposed approach.

Let $\mathcal{B}^a$ and $\mathcal{B}^r$ be the acceptance and rejection regions for (N, S_N) for testing hypotheses $H_0: \ \theta \in \Theta_0$ vs. $H_a: \ \theta \in \Theta_a$. Let $\mathcal{B}$ be the sequential stopping region for (N, S_N) as in (5), then $\mathcal{B} = \mathcal{B}^a \cup \mathcal{B}^r$. Let $\mathcal{R}^a$ and $\mathcal{R}^r$ be the acceptance and rejection regions for a nonsequential test based on S_m for testing the same hypotheses, where m, the sample size of S_m, is constant. Define events

$$D^a = \{(N, S_N) \in \mathcal{B}^r, S_m \in \mathcal{R}^a\} \quad \text{and} \quad D^r = \{(N, S_N) \in \mathcal{B}^a, S_m \in \mathcal{R}^r\}. \tag{33}$$

D^a is the event that the null hypothesis is accepted by the nonsequential test, but is rejected by the sequential test; similarly, D^r is the event with opposite actions. Hence $D = D^a \cup D^r$ and the probability of discordance between the test statistics (N, S_N) and S_m is

$$\rho(\theta) \equiv P_\theta(D) = P_\theta(D^a) + P_\theta(D^r). \tag{34}$$

Noting that $P_\theta(D) = P_\theta(D \cap (N \le m)) + P_\theta(D \cap (N > m))$, we present expressions for $P_\theta(D \cap (N \le m))$ in Theorem 3.1, and for $P_\theta(D \cap (N > m))$ in Theorem 3.2.

Theorem 3.1. *Let* $\mathcal{B}^a_{n_k} = \{s : (n_k, s) \in \mathcal{B}^a\}$ *and* $\mathcal{B}^r_{n_k} = \{s : (n_k, s) \in \mathcal{B}^r\}$. *Then*

$$P_\theta(D \cap (N \le m)) = E_\theta\{P(D \cap (N \le m) | S_m)\}, \tag{35}$$

$$P(D \cap (N \le m) | S_m) = \begin{cases} \sum_{n_k \le m} E\{1_{(S_{n_k} \in \mathcal{B}^a_{n_k})} l(n_k, S_{n_k}) | S_m\} & \textit{if} \ \ S_m \in \mathcal{R}^r, \\ \sum_{n_k \le m} E\{1_{(S_{n_k} \in \mathcal{B}^r_{n_k})} l(n_k, S_{n_k}) | S_m\} & \textit{if} \ \ S_m \in \mathcal{R}^a. \end{cases} \tag{36}$$

Proof. Equation (35) is clear. We need only show (36). For any $s \in \mathcal{R}^r$, we have $D^a \cap (S_m = s) = \emptyset$ by (33); hence $P(D^a | S_m = s) = 0$. Thus

$$P(D \cap (N \le m) | S_m = s) = P(D^r \cap (N \le m) | S_m = s)$$

$$
\begin{aligned}
&= P((N, S_N) \in \mathcal{B}^a, N \le m | S_m = s) = \sum_{n_k \le m} E\{1_{(S_{n_k} \in \mathcal{B}^a_{n_k})} 1_{(N=n_k)} | S_m = s\} \\
&= \sum_{n_k \le m} E[E\{1_{(S_{n_k} \in \mathcal{B}^a_{n_k})} 1_{(N=n_k)} | S_{n_k}\} | S_m = s] \\
&= \sum_{n_k \le m} E[1_{(S_{n_k} \in \mathcal{B}^a_{n_k})} E\{1_{(N=n_k)} | S_{n_k}\} | S_m = s]. \qquad (37)
\end{aligned}
$$

Joining (37) and $E\{1_{(N=n_k)}|S_{n_k}\} = l(n_k, S_{n_k})$, we have the first equation in (36). The second equation in (36) can be obtained similarly.

3.3.1. Discordance between sequential and nonsequential conclusions

When a sequential boundary is crossed, a natural question is whether the conclusion would be reversed if we did not stop but continued to the end. The chance of this event can be measured by the probability of discordance between a sequential test and the nonsequential test at the last stage of the sequential test. In a sequential test, we may want to ignore an early boundary crossing and continue to gather observations until a later stage, either to obtain a better estimate of the unknown parameter with a larger sample size, or to avoid early stopping caused by unexpected dependence among observations. If the probability of discordance is small, we will be less concerned with the possibility that the conclusion at early stopping will be reversed at later stages. Examples of sequential procedures that have a very small probability of discordance with their last stages are those based on stochastic curtailing (Lan, Simon, and Halperin (1982)) and those based on the sequential conditional probability ratio tests (Xiong (1995)).

Let $\rho_s \equiv P(D|S_m = s)$ be the *conditional probability of discordance* given $S_m = s$, where S_m is the observation at the final stage of the sequential test. Because $P(N > m) = 0$ by the definition of m, we have $P(D) = P(D \cap (N \le m))$ and

$$
\rho_s \equiv P(D|S_m = s) = \begin{cases} \sum_{n_k \le m} E\{1_{(S_{n_k} \in \mathcal{B}^a_{n_k})} l(n_k, S_{n_k}) | S_m = s\} & \text{if} \quad s \in \mathcal{R}^r, \\ \sum_{n_k \le m} E\{1_{(S_{n_k} \in \mathcal{B}^r_{n_k})} l(n_k, S_{n_k}) | S_m = s\} & \text{if} \quad s \in \mathcal{R}^a, \end{cases}
$$

by (36). Given a true θ, by averaging out the conditioning values of S_m, the *overall* probability of discordance between the sequential and nonsequential test is

$$
\begin{aligned}
\rho(\theta) = E_\theta &\left[1_{(S_m \in \mathcal{R}^r)} \sum_{n_k} E\{1_{(S_{n_k} \in \mathcal{B}^a_{n_k})} l(n_k, S_{n_k}) | S_m\} \right] \\
&+ E_\theta \left[1_{(S_m \in \mathcal{R}^a)} \sum_{n_k} E\{1_{(S_{n_k} \in \mathcal{B}^r_{n_k})} l(n_k, S_{n_k}) | S_m\} \right]. \qquad (38)
\end{aligned}
$$

The probabilities of discordance ρ_s and $\rho(\theta)$ measure the probability that the sequential test contradicts the nonsequential test. For ρ_s, the probability is conditional on $S_m = s$ and does not depend on θ. For $\rho(\theta)$, the probability averages out S_m, but depends on θ. Let $\rho = \max_s \rho_s$ and $\rho_{max} = \max_\theta \rho(\theta)$. Then $\rho_s \le \rho$ for any s; $\rho(\theta) \le \rho_{max} \le \rho$ for any θ. ρ and ρ_{max} can be used for design and evaluation of a sequential test.

3.3.2. Discordance with comparable non-sequential test

The significance level and/or the power of a sequential test could be different from those of the nonsequential test at the last stage of the sequential test. A nonsequential test that has the same significance level and power as a given sequential test may be more appropriate for comparison with the sequential test. This nonsequential test design is called the reference fixed sample size test (RFSST). The maximum sample size of a sequential procedure is usually larger than the sample size of its RFSST. Let m be the sample size of the nonsequential test and m^* be the maximum sample size of the sequential test design, then $m \le m^*$. Let D be the event that a sequential test and RFSST lead to a different rejection/acceptance decision when both tests are used on the same sequence of observations. Since $P_\theta(D) = P_\theta(D \cap (N \le m)) + P_\theta(D \cap (N > m))$, and $P_\theta(D \cap (N \le m))$ can be obtained by equations (35) and (36), we need only evaluate $P_\theta(D \cap (N > m))$, see Theorem 3.2 below. First a few lemmas.

Lemma 3.1. *Let $\mathcal{B}$ and $\mathcal{C}$ be defined as in* (5). *If $n_{k-1} \le m < n_k$, then for any set I in the range of S_m and for $(n_k, s) \in \mathcal{C}$,*

$$P(N>n_k, S_m\in I|S_{n_k}=s)=\begin{cases}E[1_{(S_m\in I)}E\{l^*(n_{k-1}, S_{n_{k-1}})|S_m\}|S_{n_k}=s] & \text{if } n_{k-1}<m,\\ E\{1_{(S_m\in I)}l^*(n_{k-1}, S_{n_{k-1}})|S_{n_k}=s\} & \text{if } n_{k-1}=m,\end{cases} \tag{39}$$

where $l^*(n_{k-1}, s)$ is defined as in (8). *For $(n_k, s) \in \mathcal{B}$,*

$$P(N=n_k, S_m\in I|S_{n_k}=s)=\begin{cases}E[1_{(S_m\in I)}E\{l^*(n_{k-1}, S_{n_{k-1}})|S_m\}|S_{n_k}=s] & \text{if } n_{k-1}<m,\\ E\{1_{(S_m\in I)}l^*(n_{k-1}, S_{n_{k-1}})|S_{n_k}=s\} & \text{if } n_{k-1}=m.\end{cases} \tag{40}$$

Proof. For any $(n_k, s)\in\mathcal{C}$, $P(N>n_k, S_m\in I|S_{n_k}=s)=P\{N>n_{k-1}, S_m\in I|S_{n_k}=s\}$. If $n_{k-1}<m$, then $P\{N>n_{k-1}, S_m\in I|S_{n_k}=s\}=E[1_{(S_m\in I)}E\{1_{(N>n_{k-1})}|S_m\}|S_{n_k}=s]$, which yields the first equation in (39) because $E\{1_{(N>n_{k-1})}|S_m\} = E\{l^*(n_{k-1}, S_{n_{k-1}})|S_m\}$. If $n_{k-1} = m$, then $P(N > n_{k-1}, S_m \in I|S_{n_k} = s) = E[1_{(S_m\in I)} E\{1_{(N>n_{k-1})}|S_{n_{k-1}}\}|S_{n_k} = s]$, which yields the second equation in (39) because $E\{1_{(N>n_{k-1})}|S_{n_{k-1}}\} = l^*(n_{k-1}, S_{n_{k-1}})$. Equations in (40) can be obtained similarly.

Lemma 3.2. *For $n_k > m$ and any (n_k, s), define*

$$l_a^*(n_k, s) \equiv P(N > n_k, S_m \in \mathcal{R}^a | S_{n_k} = s) \text{ and } l_r^*(n_k, s) \equiv P(N > n_k, S_m \in \mathcal{R}^r | S_{n_k} = s). \tag{41}$$

Then for $n_k > m$ and $(n_k, s) \in \mathcal{B}$, we have $l_a^*(n_k, s) = 0$ and $l_r^*(n_k, s) = 0$. For $n_k > m$ and $(n_k, s) \in \mathcal{C}$, if $n_{k-1} > m$, then

$$l_a^*(n_k, s) = E\{l_a^*(n_{k-1}, S_{n_{k-1}}) | S_{n_k} = s\} \text{ and } l_r^*(n_k, s) = E\{l_r^*(n_{k-1}, S_{n_{k-1}}) | S_{n_k} = s\}; \tag{42}$$

if $n_{k-1} \le m < n_k$, then for $(n_k, s) \in \mathcal{C}$, $l_a^*(n_k, s)$ and $l_r^*(n_k, s)$ are given by the right side of (39) *with $I = \mathcal{R}^a$ and $I = \mathcal{R}^r$, respectively.*

Proof. The proof is similar to that for (9), except for the case of $(n_k, s) \in \mathcal{C}$ and $n_{k-1} \le m < n_k$ which follows from Lemma 3.1.

Lemma 3.3. *For $n_k > m$ and any (n_k, s), define*

$$l_a(n_k, s) \equiv P(N = n_k, S_m \in \mathcal{R}^a | S_{n_k} = s) \text{ and } l_r(n_k, s) \equiv P(N = n_k, S_m \in \mathcal{R}^r | S_{n_k} = s). \tag{43}$$

For $n_k > m$ and $(n_k, s) \in \mathcal{C}$, we have $l_a(n_k, s) = 0$ and $l_r(n_k, s) = 0$. For $n_k > m$ and $(n_k, s) \in \mathcal{B}$, if $n_{k-1} > m$, then

$$l_a(n_k, s) = E\{l_a^*(n_{k-1}, S_{n_{k-1}}) | S_{n_k} = s\} \text{ and } l_r(n_k, s) = E\{l_r^*(n_{k-1}, S_{n_{k-1}}) | S_{n_k} = s\}, \tag{44}$$

where $l_a^*(n_k, s)$ and $l_r^*(n_k, s)$ are given as in (41). *If $n_{k-1} \le m < n_k$, then $l_a(n_k, s)$ and $l_r(n_k, s)$ are given by the right side of* (40) *with $I = \mathcal{R}^a$ and $I = \mathcal{R}^r$, respectively.*

Proof. The proof of this lemma is parallel to that of Lemma 3.2. by interchanging $(n_k, s) \in \mathcal{C}$ with $(n_k, s) \in \mathcal{B}$.

Theorem 3.2. *Let $\mathcal{B}^a_{n_k}$ and $\mathcal{B}^r_{n_k}$ be defined in Theorem* 3.1. *Then*

$$P_\theta\{D \cap (N > m)\} = \sum_{n_k > m} [E_\theta\{1_{(S_{n_k} \in \mathcal{B}^a_{n_k})} l_r(n_k, S_{n_k})\} + E_\theta\{1_{(S_{n_k} \in \mathcal{B}^r_{n_k})} l_a(n_k, S_{n_k})\}], \tag{45}$$

where $l_r(n_k, S_{n_k})$ and $l_a(n_k, S_{n_k})$ are given as in (44)*; and $\mathcal{B}^a_{n_k} = \{s : (n_k, s) \in \mathcal{B}^a\}$ and $\mathcal{B}^r_{n_k} = \{s : (n_k, s) \in \mathcal{B}^r\}$.*

Proof. By (34), we have

$$P\{D \cap (N > m)\} = P_\theta\{D^r \cap (N > m)\} + P_\theta\{D^a \cap (N > m)\}. \tag{46}$$

The first term on the right side of (46) is

$$P_\theta\{D^r \cap (N > m)\} = \sum_{n_k > m} P_\theta\{N = n_k, (n_k, S_{n_k}) \in \mathcal{B}^a, S_m \in \mathcal{R}^r\}$$

$$= \sum_{n_k > m} E_\theta[1_{\{(n_k, S_{n_k}) \in \mathcal{B}^a\}} E\{1_{(N = n_k, S_m \in \mathcal{R}^r)} | S_{n_k}\}] = \sum_{n_k > m} E_\theta\{1_{(S_{n_k} \in \mathcal{B}^a_{n_k})} l_r(n_k, S_{n_k})\}.$$

Similarly, the second term in (46) is $P_\theta\{D^a \cap (N > m)\} = \sum_{n_k > m} E_\theta\{1_{(S_{n_k} \in \mathcal{B}^r_{n_k})} l_a(n_k, S_{n_k})\}$. Hence, (45) holds.

Let n_{k^*} be the largest possible value of N less than or equal to the sample size m of the nonsequential test. Then the probability that the sequential test needs more samples than the nonsequential test is $P_\theta\,(N > n_{k^*}) = E_\theta\{P(N > n_{k^*}|S_{n_{k^*}})\} = E_\theta\{l^*(n_{k^*}, S_{n_{k^*}})\}$. Numerical examples for the probability of discordance and the above probability can be found in Tan, Xiong and Kutner (1998).

Acknowledgements

The first author thanks Steven Lalley as advisor of his Ph.D. research at Purdue University. The basic idea in this work originated from that research. The authors wish to thank the co-editor, referees, and James Boyett for helpful comments. The research work of Xiaoping Xiong and Ming Tan was supported in part by NIH grants R01HL61681 and CA 21765, and by the American Lebanese Syrian Associated Charities.

References

Aroian, L. A. (1968) Sequential analysis, direct method. *Technometrics* **10**, 125-132.

Armitage, P., McPherson, C. K. and Rowe, B. C. (1969). Repeated significance tests on accumulating data. *J. Roy. Statist. Soc. Ser. A* **132**, 235-244.

Anderson, T. W. (1960). A modification of the sequential probability ratio test to reduce sample size. *Ann. Math. Statist.* **31**, 165-197.

Billingsley. P. (1986). *Probability and Measure.* John Wiley, New York.

Chang, M. N., Wieand, H. S. and Chang, V. T. (1989). The bias of the sample proportion following a group sequential phase II clinical trial. *Statist. Medicine* **8**, 563-570.

DeMets, D. L. and Lan, K. K. G. (1994). Interim analysis: The alpha spending function approach. *Statist. Medicine* **13**, 1341-1352.

Emerson, S. S. and Fleming, T. R. (1990). Parameter estimation following group sequential hypothesis testing. *Biometrika* **77**, 875-892.

Jennison, C. (1994). Numerical computations for group sequential tests. In *Computing Science and Statistics* **25** (Edited by M. Tarter and M. D. Lock), 263-272. Interface Foundation of North America, Fairfax Station, VA.

Jennison, C. and Turnbull, B. W. (2000). *Group Sequential Methods with Applications to Clinical Trials.* Chapman and Hall, London.

Lai, T. L. and Siegmund, D. (1977). A nonlinear renewal theory with applications to sequential analysis I. *Ann. Statist.* **5**, 946-954, pp14-24.

Lai, T. L. and Wijsman, R. A. (1979). First exit time of a random walk from the bounds $f(n) \pm cg(n)$, with applications. *Ann. Statist.* **7**, 672-692.

Lai, T. L. (2001). Sequential analysis: some classical problems and new challenges. *Statist. Sinica* **11**, 303-350.

Lan, K. K. G., Simon, R. and Halperin, M. (1982). Stochastically curtailed tests in long-term clinical trials. *Sequential Anal.* **1**, 207-219.

Samuel-Cahn, E. (1974). Repeated significance test II, for hypotheses about the normal distribution. *Comm. Statist.* **3**, 711-733.

Siegmund, D. (1985). *Sequential Analysis.* Springer-Verlag, New York.

Tan, M., Xiong, X. and Kutner, M. H. (1998). Clinical trial designs based on sequential conditional probability ratio tests and reverse stochastic curtailing. *Biometrics* **54**, 682-695.

Wald, A. (1947). *Sequential Analysis.* Wiley, New York.

Whitehead, J. (1986). On the bias of maximum likelihood estimation following a sequential test. *Biometrika* **73**, 573-581.

Whitehead, J. (1997). *The Design and Analysis of Sequential Clinical Trials.* Wiley, New York.

Xiong, X. (1991). Sequential tests for hypergeometric distributions. Ph.D. Dissertation, Purdue University, West Lafayette, Indiana.

Xiong, X. (1992). Absorption probability distributions of random paths from finite populations. Technique Report #92-41C, Purdue University, West Lafayette, Indiana.

Xiong, X. (1995). A class of sequential conditional probability ratio tests. *J. Amer. Statist. Assoc.* **90**, 1463-1473.

Xiong, X. (1996). Absorption probability distributions of random paths from finite populations. *Sequential Anal.* **15**, 1-19.

Department of Biostatistics, St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis, TN 38105-2794, U.S.A.

E-mail: xiaoping.xiong@stjude.org

Division of Biostatistics, University of Maryland Greenbaum Cancer Center, 22 S. Greene St., Baltimore, Maryland 21201, U.S.A.

E-mail: mtan@umm.edu

Department of Biostatistics, Emory University, Atlanta, GA 30322, U.S.A.

E-mail: mkutner@sph.emory.edu

(Received February 2000; accepted February 2002)