A MULTIVARIATE WEAK CONVERGENCE RESULT WITH APPLICATION TO GROUP SEQUENTIAL CLINICAL TRIALS (Version 1a)

Michael R. Kosorok, PhD
Department of Biostatistics

MADISON, WISCONSIN
A MULTIVARIATE WEAK CONVERGENCE RESULT WITH APPLICATION TO GROUP SEQUENTIAL CLINICAL TRIALS (version 1c)

Department of Biostatistics Technical Report 108

BY MICHAEL R. KOSOROK

University of Wisconsin–Madison

Weak convergence results for the two-sample function-indexed weighted log-rank statistics of Kosorok (1998) and Kosorok and Lin (1998) are extended to the group sequential setting with staggered patient entry. The generality of these results allows for multivariate combinations of a wide variety of asymptotically Gaussian processes, and an example is given which combines a function-indexed weighted log-rank with a binary response statistic. Monte Carlo simulation is used to accurately determine the null distribution of measurable continuous mappings which can, in turn, be used to construct group sequential boundaries.

AMS 1991 subject classifications. Primary 60F05, 62L10; secondary 60G44, 65C05.

Key words and phrases. Binary data, censored data, contiguous alternatives, function-indexed log-rank, martingales, Monte Carlo methods, multivariate failure time data, staggered entry.
1. Introduction. Many large scale clinical trials incorporate group sequential interim monitoring to permit appropriate early stopping in the presence of a clear treatment effect so that unnecessary patient discomfort and monetary expense can be avoided (Fleming and DeMets, 1993). Frequently, staggered entry of patients into these trials is an operational necessity which can complicate the distributional properties of the test statistics used (Tsiatis, 1982). The function-indexed log-rank statistics of Kosorok (1998) and Kosorok and Lin (1998) have been shown to be have potentially greater power for certain collections of ordered hazards alternatives than previous weighted log-rank statistics. The primary focus of this paper is to obtain weak convergence results for these function-indexed log-rank statistics in the group sequential setting with staggered patient entry so that the associated increase in power can be translated into more efficient clinical trial designs.

The surprising generality of these results allows for multivariate combinations of a wide variety of asymptotically Gaussian processes, including several well-known statistics for which the asymptotic distribution in the group sequential staggered entry setting has not been previously described. For example, the proposed methodology can be applied to Renyi-type supremaums of weighted log-rank statistics (see Fleming, Harrington, and O’Sullivan, 1987) which are—in this setting—a continuous mapping of a multivariate Brownian motion process with a complex correlation structure. In addition, a simple method of Monte Carlo simulation is also presented and shown to be an accurate method of determining the null distribution of such statistics, thereby permitting the construction of group sequential boundaries during clinical trial execution.

Section 2 gives the background and overall strategy for obtaining the desired weak convergence results. The main theoretical results are then presented in section 3, and specific examples of applications for both failure times and binary outcomes are given in section 4. The paper concludes with a brief discussion in section 5.

by an element $f$ in a compact set $H$, where, for $u \in [0, \infty]$, $D[0, u]$ is the space of bounded real cadlag functions (right-continuous with left-hand limits) on $[0, u]$, where the limit at infinity exists if $u = \infty$ and where $D[0, 0]$ is defined to be the set of reals. The generality of this indexing structure encompasses many other families of stochastic processes, including many non-failure time statistics. A univariate statistic—as a trivial example—can be formulated as such a process by letting $u = 0$ and letting $H$ consist of a single point. Denote $A(F, G)$ to be the space of continuous mappings from a metric space $F$ to another metric space $G$.

Let $n$ be the number of individual patients and $m$ the number of statistics employed, where each statistic satisfies the following:

(S1) **Stochastic process appropriateness:**

$$
\bar{X}_j^n(f, t) = \sum_{i=1}^{n} X^n_{ij}(f, t)
$$

is a stochastic process doubly indexed by $f \in H_j$ and $t \in [0, u_j]$ and taking values in $A(H_j, D[0, u_j])$, where $\{H_j, \rho_j\}$ is a compact metric space, $u_j \in [0, \infty]$, and where $\bar{X}_j^n(f, t)$ is adapted to the right-continuous filtration $\mathcal{F}_t^{in}$ (which is allowed to depend on $j$) for all $f \in H_j$, $j = 1 \ldots m$.

The dimension $m$ will usually be the number of statistics employed at each interim analysis time multiplied by the number of interim analyses.

In order to obtain weak convergence of the above multivariate collections, we will

1. approximate these statistics by sums of independent random vectors to obtain convergence of finite dimensional distributions;

2. utilize tightness of the marginal distributions, in the appropriate subspaces of $D[0, \infty]$, to obtain tightness over finite collections of function-indices; and

3. utilize the results of Kosorok (1998), to obtain tightness, and therefore weak convergence, over the entire function-indexing space.
In order to accomplish part (3), we will need to make use of the following theorem of Kosorok (1998). Before proceeding, let \( \phi \) be the metric for \( F \) and \( \gamma \) the metric for \( G \) in the space \( \mathbf{A}(F,G) \) defined previously. For \( x, y \in \mathbf{A}(F,G) \), define the metric

\[
\alpha(x,y) = \sup_{f \in F} \gamma(x(f), y(f)).
\]

**Theorem 1** (hereafter referred to as KT1) Suppose we have a sequence of stochastic processes \( \{X_n(\cdot)\} \) in \( \mathbf{A}(F,G) \), with \( \{F, \phi\} \) compact and \( \{G, \gamma\} \) complete and separable, and where the following conditions obtain:

(i) All finite dimensional distributions of the form \( \{X_n(f_1), X_n(f_2), \ldots, X_n(f_M)\} \), for \( f_i \in F, i = 1 \ldots M \), and for \( M < \infty \), converge weakly on \( (\{G, \gamma\})^M \) to some \( \{X(f_1), X(f_2), \ldots, X(f_M)\} \) as \( n \to \infty \).

(ii) \( \forall f, g \in F, \)

\[
\gamma(X_n(f), X_n(g)) \leq q(\phi(f, g)) U_n,
\]

where

(ii.a) \( \gamma \) is a bounded (not depending on \( f \) and \( g \)),

(ii.b) \( q \) is continuous and non-decreasing with \( q(0) = 0 \), and

(ii.c) \( \{U_n\} \) is a stochastically bounded sequence of real random variables, i.e., \( \forall \epsilon > 0, \exists \tau < \infty: P[U_n \leq \tau] > 1 - \epsilon \ \forall n \geq 1 \), not depending on \( f \) and \( g \).

Then \( X_n(\cdot) \) converges weakly on \( \{\mathbf{A}(F,G), \alpha\} \) to \( X(\cdot) \) as \( n \to \infty \).

We now need to specify the metric spaces we will be considering. Define \( \mathbf{H} = H_1 \times H_2 \times \cdots \times H_m \) (a product space), and for any \( f \in \mathbf{H} \), admit the decomposition \( f = \{f^1, f^2, \ldots, f^m\} \), where \( f^j \in H_j \) and \( j = 1 \ldots m \). For any \( f, g \in \mathbf{H} \), let \( \rho(f,g) = \max_{1 \leq j \leq m} \rho_j(f^j, g^j) \). Define \( \mathbf{D} = D[0, u_1] \times D[0, u_2] \times \cdots \times D[0, u_m] \) (another product space), and for any \( a \in \mathbf{D} \), admit the decomposition \( a = \{a_1, a_2, \ldots, a_m\} \), where \( a_j \in D[0, u_j] \), \( j = 1 \ldots m \). For any \( a, b \in \mathbf{D} \), let \( d(a,b) = \max_{1 \leq j \leq m} d_j(a_j, b_j) \), where \( d_j \) is a bounded
and complete version of the Skorohod metric on $D[0,u_j]$ for $u_j > 0$ and is the usual Euclidean metric if $u_j = 0$, $j = 1 \ldots m$. Now, for any $x, y \in A(H, D)$, define the metric
\[ \alpha(x, y) = \sup_{f \in H} d(x(f), y(f)). \]
The choice of $u_j$, $j = 1 \ldots m$, is dictated by the application of interest; for example, the choice of $u_j$ might be the maximum possible length of follow-up at the $j$'th interim analysis, $j = 1 \ldots m$.

3. Main results. In this section, we obtain multivariate weak convergence of the collections $\{ \overline{X}^{n}_{j}(\cdot, \cdot), j = 1 \ldots m \}$ described in the previous section, uniform consistency of an appropriate covariance estimator, and weak convergence for a Monte Carlo approximation of the corresponding asymptotic null distribution.

3.1 Multivariate weak convergence. Conditions (A1), (B1), (J1), and (J2) below will be needed to obtain convergence of all finite dimensional distributions:

(A1) Asymptotic independence: For each $f \in H_j$,
\[
\sup_{t \in [0, u_j]} \left| \overline{X}^{n}_{j}(f, t) - \sum_{i=1}^{n} X^{*n}_{ij}(f, t) - \mu_{j}(f, t) \right| \to 0
\]
in probability, as $n \to \infty$, where $\mu_{j}(f, \cdot)$ is a constant, bounded cadlag function and
\[
\overline{X}^{n}_{j}(f, t) \equiv \sum_{i=1}^{n} X^{*n}_{ij}(f, t) \in D[0, u_j]
\]
is a mean zero $\mathcal{F}^{n}_{t}$-adapted martingale for all $f \in H_j$, and where the collection $\{X^{*n}_{ij}(\cdot, \cdot), i = 1 \ldots n\}$ is stochastically independent across $i$, $j = 1 \ldots m$, but possibly dependent across $j$.

(B1) Boundedness: $X^{*n}_{ij}(f, t)$ defined in (A1) satisfies
\[
E \left[ \sqrt{n} X^{*n}_{ij}(f, t) \right]^{d} \leq K_{j}, \quad i = 1 \ldots n, \text{ for some } K_{j} < \infty \text{ and all } n \geq 1, \text{ all } f \in H_j, \text{ and all } t \in [0, u_j], j = 1 \ldots m.
\]

(J1) Joint independence: The processes $\{X^{*n}_{i1}(\cdot, \cdot), X^{*n}_{i2}(\cdot, \cdot), \ldots, X^{*n}_{im}(\cdot, \cdot)\}$, $i = 1 \ldots n$, are stochastically independent across $i$. 

5
(J2) **Joint covariance convergence:**

\[
\lim_{n \to \infty} \sum_{i=1}^{n} E \left[ X_{i_j}^n(f^i, t_j) X_{i_k}^n(f^k, t_k) \right] = V_{jk}(f^j, f^k; t_j, t_k),
\]

for all \( f^j \in H_j, f^k \in H_k, t_j \in [0, u_j], \) and \( t_k \in [0, u_k], \) over \( j, k = 1 \ldots m. \)

The following condition will be needed to obtain tightness of finite collections of function-indices:

(T1) **Tightness:** For each \( f \in H_j, \overline{X}_j^n(f, \cdot) \) is tight (over \( n \)) in the Skorohod topology on \( D[0, u_j], j = 1 \ldots m. \)

Finally, tightness over the entire function-indexed space requires the following continuity condition:

(C1) **Continuity:** For all \( f, g \in H_j, \)

\[
\sup_{t \in [0, u_j]} \left| \overline{X}_j^n(f, t) - \overline{X}_j^n(g, t) \right| \leq q_j(p_j(f,g)) U_j^n,
\]

where

(i) \( q_j(\cdot) \) is continuous and non-decreasing with \( q_j(0) = 0; \) and

(ii) \( \{U_j^n, n \geq 1\} \) is a stochastically bounded sequence of real random variables, i.e.,

\[
\forall \epsilon > 0, \exists \tau < \infty: P \left[ U_j^n \leq \tau \right] > 1 - \epsilon \ \forall n \geq 1, \text{ not depending on } f \text{ and } g;
\]

\( j = 1 \ldots m. \)

**Remark 1** Only conditions \( (J1) \) and \( (J2) \) restrict the joint behavior of the \( m \) statistics considered: all of the remaining conditions restrict only the marginal behavior of each statistic individually, \( j = 1 \ldots m. \)

**Theorem 2** Suppose we have the collection of statistics

\[
\overline{X}^n(f) \equiv \{ \overline{X}_1^n(f^1, \cdot), \overline{X}_2^n(f^2, \cdot), \ldots, \overline{X}_m^n(f^m, \cdot) \},
\]
where each $\bar{X}_j(f^j, \cdot)$ satisfies condition (S1), $j = 1 \ldots m$, and conditions (A1), (B1), (J1), (J2), (T1), and (C1) are also satisfied. Then $\bar{X}^n(f)$ converges weakly on $\{A(H, D, \alpha)\}$ to a Gaussian process $Z(f)$ with mean function

$$\mu(f) \equiv \{\mu_1(f^1, \cdot), \mu_2(f^2, \cdot), \ldots, \mu_m(f^k, \cdot)\}$$

and covariance function

$$(3.1) \quad \left[V_{f, g}(s, t)\right]_{jk} = V_{jk}(f^j, g^k, s_j, t_k),$$

for $f^j, g^j \in H_j$, $s_j, t_j \in [0, u_j]$, $j, k = 1 \ldots m$, and where $[A]_{jk}$ is the $jk$'th element of the matrix $A$.

Proof. For every finite collection of indices $\{f_k \in H, k = 1 \ldots M\}$, let $t^w_j \in [0, u_j]$ and $c_{jkw}$ be real, $j = 1 \ldots m$, $k = 1 \ldots M$, $w = 1 \ldots L$. Then

$$B^2_n \equiv \text{var} \left[ \sum_{j=1}^{m} \sum_{k=1}^{M} \sum_{w=1}^{L} c_{jkw} \sum_{i=1}^{n} X^*_n(i_j, f^w_k, t^w_j) \right]$$

converges to

$$B^2 = \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \sum_{k=1}^{M} \sum_{k_2=1}^{M} \sum_{w=1}^{L} \sum_{w_2=1}^{L} c_{j_1,k,w_1} c_{j_2,k_2,w_2} V_{j_1,j_2}(f^{j_1}_{k_1}, f^{j_2}_{k_2}, t^{w_1}_{j_1}, t^{w_2}_{j_2})$$

by (J2). If $B^2 = 0$, then, trivially,

$$\sum_{j=1}^{m} \sum_{k=1}^{M} \sum_{w=1}^{L} c_{jkw} \sum_{i=1}^{n} X^*_n(i_j, f^w_k, t^w_j)$$

converges in distribution, as $n \to \infty$, to a normal random variable with mean zero and variance zero. If $B^2 > 0$, then

$$\sum_{i=1}^{n} E \left[ \left( \sum_{j=1}^{m} \sum_{k=1}^{M} \sum_{w=1}^{L} c_{jkw} \sum_{i=1}^{n} X^*_n(i_j, f^w_k, t^w_j) \right)^4 \right]$$

$$\leq (mML)^4 \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{M} \sum_{w=1}^{L} c_{jkw}^4 E \left[ (X^*_n(i_j, f^w_k, t^w_j))^4 \right]$$

$$\leq K/n,$$
for some \( K < \infty \), which, by corollary 1.9.3 of Serfling (1980) and the Cramer-Wold device, implies that all finite-dimensional distributions converge. Since tightness of the marginal distributions implies tightness of the joint distribution, we obtain tightness over \( \{ f_k \in H, k = 1 \ldots M \} \), by (T1), satisfying condition (i) of KT1.

Now, by (C1), we have \( \forall f, g \in H \) that
\[
\max_{1 \leq j \leq m} \sup_{t \in [0,u_j]} |X^n_j(f^j, t) - \overline{X}_j^n(g^j, t)| \wedge 1 \leq \left\{ \max_{1 \leq j \leq m} q_j(\rho_j(f^j, g^j)) \right\} \left\{ \max_{1 \leq j \leq m} U^n_j \right\}
\leq \left\{ \max_{1 \leq j \leq m} q_j(\rho(f, g)) \right\} \left\{ \max_{1 \leq j \leq m} U^n_j \right\},
\]
and condition (ii) of KT1 is then satisfied by noting that \( \forall a, b \in D[0,u_j], \sup_{t \in [0,u_j]} |a(t) - b(t)| \)

is a continuous mapping, \( j = 1 \ldots m \); and, therefore, that \( \forall a, b \in D \),
\[
\max_{1 \leq j \leq m} \sup_{t \in [0,u_j]} |a_j(t) - b_j(t)| \wedge 1
\]
is also continuous as well as being bounded. Hence by KT1, we obtain the desired weak convergence. \( \square \)

3.2 A uniformly consistent covariance estimator. The covariance estimator we propose requires, for \( i = 1 \ldots n \), an approximation of \( X^{*n}_{ij}(\cdot, \cdot) \) given in (A1)—for each of the \( m \) statistics \( (j = 1 \ldots m) \) given in (S1)—of the following form:

(S2) Stochastic process appropriateness: Each \( \overline{X}^{*n}_{ij}(f, t) \) is a mean zero \( \mathcal{F}^m_t \)-adapted martingale on \( D[0,u_j], i = 1 \ldots n \), \( f \in H_j, j = 1 \ldots m \).

In order to obtain uniform consistency over all time axes for finite collections of function-indices, this approximation will also need to satisfy the following conditions:

(A2) Asymptotic independence:
\[
\lim_{n \to \infty} \sum_{i=1}^{n} E \left[ (\overline{X}^{*n}_{ij}(f, u_j) - X^{*n}_{ij}(f, u_j))^2 \right] = 0,
\]

for all \( f \in H_j \), where \( X^{*n}_{ij}(\cdot, \cdot) \) is defined in (A1), \( j = 1 \ldots m \).
(B2) **Boundedness:**

\[
E \left[ \left( \sqrt{n} \tilde{X}_{ij}^{*n}(f_i, u_i) \right)^2 \right] \leq K_j
\]

for some \( K_j < \infty, i = 1 \ldots n, \) and for all \( f \in H_j, j = 1 \ldots m. \)

Finally, uniform consistency over the entire function-indexed space requires the following continuity and boundedness conditions:

(C2) **Continuity:** For all \( f, g \in H_j, \) if \( a_1 \ldots a_n \) are a sequence of real numbers, then

\[
\left( \sum_{i=1}^{n} a_i \left( \tilde{X}_{ij}^{*n}(f_i, t) - \tilde{X}_{ij}^{*n}(g_i, t) \right) \right)^2 \leq q_j^2 (\rho_j(f, g)) \left( \sum_{i=1}^{n} a_i \tilde{Y}_{ij}^{*n}(t) \right)^2,
\]

where

(i) \( q_j^2 (\cdot) \) satisfies (C1.1); and

(ii) \( \tilde{Y}_{ij}^{*n}(t), i = 1 \ldots n, \) are uncorrelated (across \( i \)) mean zero \( \mathcal{F}_t \)-martingales with

\[
\sum_{i=1}^{n} E \left[ \left( \tilde{Y}_{ij}^{*n}(u_j) \right)^2 \right] \leq K_j
\]

for some \( K_j < \infty \) and all \( n \geq 1, j = 1 \ldots m. \)

(B3) **Boundedness:** For all \( f \in H_j \) and \( i = 1 \ldots n, \)

\[
\left( \tilde{X}_{ij}^{*n}(f_i, t) \right)^2 \leq K_j \left( \tilde{Y}_{ij}^{*n}(t) \right)^2,
\]

for some \( K_j < \infty \) not depending on \( f \) and for all \( n \geq 1; \) and (C2.ii) is also satisfied, \( j = 1 \ldots m. \)

The unusual generality of condition (C2) will be particularly useful when we examine the asymptotic properties of the Monte Carlo method described later.

**Theorem 3** Suppose the conditions of theorem 2 obtain, and, in addition, conditions (S2), (A2), (B2), (C2), and (B3) hold. Then, the variance estimator \( \hat{V}_{f, g}(s, t), \) where

\[
\hat{V}_{f, g}(s, t)_{jk} = \hat{V}_{jk}(f^j, g^k; s_j, t_k)
\]

\[
\equiv \sum_{i=1}^{n} \tilde{X}_{ij}^{*n}(f_i, s) \tilde{X}_{il}^{*n}(g_i, t)
\]

(3.2)
is uniformly consistent over all \( f, g \in H \) and \( s, t \in \{[0, u_1] \times [0, u_2] \times \cdots \times [0, u_m]\} \) for \( V_{f,g}(s,t) \) given in (3.1).

Proof. We first obtain uniform consistency over all time axes for finite collections of function-indices. For any \( f, g \in H \) and \( j, l = 1 \ldots m, \)

\[
\begin{align*}
(3.3) \quad & \sup_{s \in [0, u], t \in [0, u]} \left| \frac{1}{n} \sum_{i=1}^{n} \left( X_{ij}^{*n}(f^j, s)X_{il}^{*n}(g^l, t) - X_{ij}^{*n}(f^j, s)X_{il}^{*n}(g^l, t) \right) \right| \\
& \leq \left( \sup_{s \in [0, u], t \in [0, u]} \left( \sum_{i=1}^{n} X_{ij}^{*n}(f^j, s) \right)^2 \right)^{1/2} \left( \sup_{t \in [0, u]} \left( \sum_{i=1}^{n} X_{il}^{*n}(g^l, t) \right)^2 \right)^{1/2} \\
& \quad + \left( \sup_{t \in [0, u]} \left( \sum_{i=1}^{n} X_{il}^{*n}(g^l, t) \right)^2 \right)^{1/2} \left( \sup_{s \in [0, u]} \left( \sum_{i=1}^{n} X_{ij}^{*n}(f^j, s) \right)^2 \right)^{1/2}.
\end{align*}
\]

By (S2), \( \sum_{i=1}^{n} \left[ X_{ij}^{*n}(f^j, s) \right]^2 \) is a submartingale and, by proposition 2.7.16(a) of Ethier and Kurtz (1986)—hereafter abbreviated by EKP—(B2) implies that \( \sup_{s \in [0, u], t \in [0, u]} \sum_{i=1}^{n} \left[ X_{ij}^{*n}(f^j, s) \right]^2 \) is a stochastically bounded sequence (over \( n \)) of random variables. By (A1), \( \sum_{i=1}^{n} \left[ X_{il}^{*n}(g^l, t) \right]^2 \) is also a submartingale and we have by (B1) that \( \sup_{t \in [0, u]} \sum_{i=1}^{n} \left[ X_{il}^{*n}(g^l, t) \right]^2 \) is likewise a stochastically bounded sequence (over \( n \)). The result now follows by noting that both \( \sum_{i=1}^{n} \left[ X_{ij}^{*n}(f^j, s) - X_{ij}^{*n}(f^j, s) \right]^2 \) and \( \sum_{i=1}^{n} \left[ X_{il}^{*n}(g^l, t) - X_{il}^{*n}(g^l, t) \right]^2 \) are also submartingales and that (A2) can therefore be applied to establish that both are stochastically bounded sequences (over \( n \)).

Now, \( \forall f, g, f_*, g_* \in H, \)

\[
\begin{align*}
\max_{1 \leq j, k \leq m} \sup_{s_j \in [0, u_j]} \sup_{t_k \in [0, u_k]} \left| \frac{1}{n} \sum_{i=1}^{n} X_{ij}^{*n}(f^j, s_j)X_{ik}^{*n}(g^k, t_k) - \frac{1}{n} \sum_{i=1}^{n} X_{ij}^{*n}(f^j, s_j)X_{ik}^{*n}(g^k, t_k) \right| \\
& \leq \max_{1 \leq j, k \leq m} \left[ \left( \sup_{s_j \in [0, u_j]} \sum_{i=1}^{n} \left[ X_{ij}^{*n}(f^j, s_j) \right]^2 \right)^{1/2} \left( \sup_{t_k \in [0, u_k]} \sum_{i=1}^{n} \left[ X_{ik}^{*n}(g^k, t_k) \right]^2 \right)^{1/2} \right] \\
& \quad + \left( \sup_{t_k \in [0, u_k]} \sum_{i=1}^{n} \left[ X_{ik}^{*n}(g^k, t_k) \right]^2 \right)^{1/2} \left( \sup_{s_j \in [0, u_j]} \sum_{i=1}^{n} \left[ X_{ij}^{*n}(f^j, s_j) \right]^2 \right)^{1/2} \\
& \leq \max_{1 \leq j, k \leq m} \left[ \left( K_j \sum_{s_j \in [0, u_j]} \left[ X_{ij}^{*n}(s_j) \right]^2 \right)^{1/2} \left( q_j(\rho_k(g^k, g_*)) \sup_{t_k \in [0, u_k]} \sum_{i=1}^{n} \left[ X_{ik}^{*n}(t_k) \right]^2 \right)^{1/2} \right] \\
& \quad + \left( K_k \sum_{t_k \in [0, u_k]} \left[ X_{ik}^{*n}(t_k) \right]^2 \right)^{1/2} \left( \rho_j(\rho_j(f^j, f_*)) \sup_{s_j \in [0, u_j]} \sum_{i=1}^{n} \left[ X_{ij}^{*n}(s_j) \right]^2 \right)^{1/2} \\
& \quad + \left( \rho_j(\rho_j(f^j, f_*)) \sup_{s_j \in [0, u_j]} \sum_{i=1}^{n} \left[ X_{ij}^{*n}(s_j) \right]^2 \right)^{1/2} \\
& = 10.
\end{align*}
\]
\[
\leq \left( \max_{1 \leq j \leq k} K_j \right)^{1/2} \left( \max_{1 \leq j \leq m} q_j(\rho_j(f^j, f^j)) + \max_{1 \leq j \leq m} q_j(\rho_j(g^j, g^j)) \right)^{1/2} \\
\times \left( \max_{1 \leq j \leq m} \sup_{s_j \in [0, u_j]} \sum_{i=1}^n \left[ \hat{Y}_{ij}^n(s_j) \right]^2 \right)^{1/2},
\]
for some \( K_j, j = 1 \ldots m \), by (C2) and (B3); but
\[
\max_{1 \leq j \leq m} \sup_{s_j \in [0, u_j]} \sum_{i=1}^n \left[ \hat{Y}_{ij}^n(s_j) \right]^2
\]
is a stochastically bounded sequence (over \( n \)) by (C2.ii). We can now combine this result with KT1 to obtain that
\[
\sup_{f \in H^1, g \in H^1} \max_{s_j \in [0, u_j]} \sup_{s_j \in [0, u_j]} \left| \hat{V}_{ij}(f^i, g^i; s_j, t_k) - V_{ij}(f^i, g^i; s_j, t_k) \right|
\]
converges to zero in probability as \( n \to \infty \).

3.3 A Monte Carlo approximation. In this section, we demonstrate that a simple Monte Carlo procedure applied to the approximations given in (S2) yields approximate replicates of the asymptotic null distribution of \( \bar{X}(f) \), described in Theorem 2. To accomplish this, we need to add the following tightness condition:

(T2) Tightness: If \( Z_1 \ldots Z_n \) is a sequence of independent standard normal deviates, then \( \sum_{i=1}^n Z_i \bar{X}_{ij}^n(f, \cdot) \) is tight (over \( n \)) in the Skorohod topology on \( D[0, u_j] \), for all \( f \in H_j, j = 1 \ldots m \).

**Theorem 4** Suppose the conditions of theorem 2 obtain, and, in addition, conditions (S2), (A2), (T2), and (C2) hold. Let \( Z_{iq}, i = 1 \ldots n \) and \( q = 1 \ldots Q \), be independent standard normal random variables. Define
\[
\bar{X}_{jq}^n(f, t) \equiv \sum_{i=1}^n Z_{iq} \bar{X}_{ij}^n(f, t),
\]
for all \( f \in H_j, t \in [0, u_j], j = 1 \ldots m, \) and
\[
\bar{X}_q^n(f) \equiv \left\{ \bar{X}_{1q}^n(f^1, \cdot), \bar{X}_{2q}^n(f^2, \cdot), \ldots, \bar{X}_{mq}^n(f^m, \cdot) \right\},
\]
\( q = 1 \ldots Q \). Then the collection \( \left\{ \bar{X}_q^n(\cdot), q = 1 \ldots Q \right\} \) converges weakly on \( \{A(H, D), \alpha\}^Q \) to \( Q \) independent replicates of a mean zero Gaussian process on \( A(H, D) \) with the same distribution as \( Z(\cdot) - \mu(\cdot) \), where \( Z \) and \( \mu \) are as given in Theorem 2.
Proof. First note that

$$X^n_{jq}(f,t) - \sum_{i=1}^n Z_{iq} X^{*n}_{ij}(f,t) = \sum_{i=1}^n Z_{iq} \left[ \bar{X}^{*n}_{ij}(f,t) - X^{*n}_{ij}(f,t) \right]$$

is an $\bar{F}^{jn}_t$-martingale, where

$$\bar{F}^{jn}_t \equiv \sigma \{ F_t^{jn}, Z_{iq}, i = 1 \ldots n, q = 1 \ldots Q \}$$

and $\sigma \{ A \}$ is the smallest $\sigma$-field making all of $A$ measurable; hence

$$\left( \sum_{i=1}^n Z_{iq} \left[ \bar{X}^{*n}_{ij}(f,t) - X^{*n}_{ij}(f,t) \right] \right)^2$$

is a submartingale with

$$E \left[ \left( \sum_{i=1}^n Z_{iq} \left[ \bar{X}^{*n}_{ij}(f,t) - X^{*n}_{ij}(f,t) \right] \right)^2 \right] \leq \sum_{i=1}^n E \left[ \left( \bar{X}^{*n}_{ij}(f,t) - X^{*n}_{ij}(f,t) \right)^2 \right],$$

which converges to zero as $n \to \infty$, and EKP thus implies that

$$\sup_{t \in [0,u]} \left| X^n_{jq}(f,t) - \sum_{i=1}^n Z_{iq} X^{*n}_{ij}(f,t) \right|$$

converges to zero in probability as $n \to \infty$. Arguments similar to those used in theorem 2 can now be used on the processes $\sum_{i=1}^n Z_{iq} X^{*n}_{ij}(f,t)$, $q = 1 \ldots Q$, to establish the necessary convergence of finite dimensional distributions, and (T2) now yields weak convergence for finite collections of function-indices.

Condition (C2) now implies that for $q = 1 \ldots Q$,

$$\sup_{t \in [0,u]} \left( \sum_{i=1}^n Z_{iq} \left[ \bar{X}^{*n}_{ij}(f^i,t) - \bar{X}^{*n}_{ij}(g^i,t) \right] \right)^2 \wedge 1 \leq q_j^*(\rho(f^i,g^i)) \sup_{t \in [0,u]} \left( \sum_{i=1}^n Z_{iq} \bar{Y}^{*n}_{ij}(t) \right)^2,$$

$j = 1 \ldots m$, and all $f, g \in H$; but $\left( \sum_{i=1}^n Z_{iq} \bar{Y}^{*n}_{ij}(t) \right)^2$ is an $\bar{F}^{jn}_t$-submartingale with

$$E \left[ \left( \sum_{i=1}^n Z_{iq} \bar{Y}^{*n}_{ij}(t) \right)^2 \right] = \sum_{i=1}^n E \left[ \left( \bar{Y}^{*n}_{ij}(u_j) \right)^2 \right] \leq K_j,$$

for some $K_j < \infty$, and hence is a stochastically bounded sequence (over $n$), $j = 1 \ldots m$. We can now apply arguments similar to those used in the proof of Theorem 2 in order to utilize KT1 to obtain the desired results. □
4. **Two examples.** The two examples that follow show how the previous results apply to statistics assessing both a failure time and a binary outcome in a group sequential clinical trial with staggered patient entry.

4.1. A *failure time outcome.* The following model is an adaptation of the staggered entry failure time model of Tsiatis (1982) and will be the basis for our study:

**MODEL 1** For \( i = 1 \ldots n \), let \( U_i \) denote the calendar time of entry, \( T_i \) the failure time measured from study entry, \( C_i^* \) the time to loss of follow-up measured from study entry, and \( W_i \) the indicator of treatment 1, for patient \( i \). Assume that the pair \( \{U_i,C_i^*\} \) is statistically independent of \( T_i \) conditional on \( W_i \), and that the random vectors \( \{U_i,T_i,C_i^*,W_i\}^T \) (where superscript \( T \) denotes transpose), \( i = 1 \ldots n \), are independent and identically distributed, with the distribution of \( T_i \) determined by the hazard function \( \Lambda_{W_i}(\cdot) \), the remaining components of the distribution fixed, and with \( E[W_i] = p \in (0,1) \). Let \( A_1 \ldots A_J \) be \( J \) fixed calendar times for interim analysis, and define

\[
C_{ij} \equiv [(A_j - U_i) \lor 0] \land C_i^* ,
\]

for \( j = 1 \ldots J \). Suppose also that for some \( \phi : [0, \infty) \mapsto [-c_\phi,c_\phi] \), where \( 0 < c_\phi < \infty \), some finite \( \beta \), and some baseline hazard \( \Lambda_0(t) \),

\[
d\Lambda_{(1)}^\phi(t) = \frac{\exp \left[ \frac{\beta \phi(t)}{2\sqrt{n}} \right] d\Lambda_0(t)}{1 + \left( \exp \left[ \frac{\beta \phi(t)}{2\sqrt{n}} \right] - 1 \right) \Delta \Lambda_0(t)}
\]

and

\[
d\Lambda_{(2)}^\phi(t) = \frac{\exp \left[ -\frac{\beta \phi(t)}{2\sqrt{n}} \right] d\Lambda_0(t)}{1 + \left( \exp \left[ -\frac{\beta \phi(t)}{2\sqrt{n}} \right] - 1 \right) \Delta \Lambda_0(t)}
\]

This is the proportional odds contiguous alternative model—which becomes the proportional hazards contiguous alternative model when the baseline hazard is continuous—and is quite useful for studying the null distribution and power of weighted log-rank tests.

For \( i = 1 \ldots n \) and \( j = 1 \ldots J \), we also need to define

\[
N_{ij}(t) = I(T_i \leq t, C_i \geq T_j),
\]
where $I_{\{B\}}$ is the indicator function of $B$, and

$$Y_{ij}(t) = I_{\{T_i \wedge C_{ij} \geq t\}},$$

the observed failure counting process and at risk process, respectively, for patient $i$ and interim analysis time $A_j$. For $j = 1 \ldots J$, define

$$\overline{Y}_j^{(1)}(t) = \sum_{i=1}^{n} W_i Y_{ij}(t),$$

$$\overline{Y}_j^{(2)}(t) = \sum_{i=1}^{n} (1 - W_i) Y_{ij}(t),$$

$$\mathcal{F}_i^j = \sigma\{N_{ij}(s), Y_{ij}(s^+), W_i, i = 1 \ldots n, s \leq t\},$$

and $\mathcal{F}_i^n = \sigma\{\mathcal{F}_i^j, t \geq 0\}$, where $\sigma\{A\}$ is the smallest $\sigma$-field making all of $A$ measurable.

Under the assumptions of model 1, the following limits exist:

$$\pi_j^{(1)}(t) = \lim_{n \to \infty} \frac{\overline{Y}_j^{(1)}(t)}{\sum_{i=1}^{n} W_i}$$

and

$$\pi_j^{(2)}(t) = \lim_{n \to \infty} \frac{\overline{Y}_j^{(2)}(t)}{\sum_{i=1}^{n} (1 - W_i)}$$

uniformly over $t \in [0, \infty]$. We also need to define $I_j = \{t : \pi_j^{(1)}(t) \wedge \pi_j^{(2)}(t) > 0\}$ and $u_j^* = \sup I_j$.

Now, for $j = 1 \ldots J$, define $B(\{\mathcal{F}_i^n, n \geq 1\})$ to be the class of all random sequences of functions $\{b^n : [0, \infty) \to [0, 1], n \geq 1\}$ such that for each $n \geq 1$ and $t \in [0, \infty)$, $b^n(t)$ is $\mathcal{F}_i^n$-predictable and, for each closed subinterval of $I_j$, $I^* \subset I_j$, the following holds:

$$\sup_{s \in I^*} |b^n(s) - b(s)| \to 0$$

in probability, as $n \to \infty$, for some $b : [0, \infty) \to [0, 1]$, where $b$ is left-continuous with right hand limits and satisfies the remaining conditions of Definition 2 of Kosorok (1998). Also define $G^+_1(K)$ to be the set of absolutely continuous functions mapping from $[0, 1]^*$ to $[0, 1]$ for which the total of the $L_2$ norms of all first cross-partial derivatives are bounded by $K$ as described in Definition 2 of Kosorok (1998).
The statistic we will use for assessing the treatment effect at analysis time $A_j, j = 1 \ldots J,$ will be a standardized version of the function-indexed weighted log-rank statistic $X^n_{ij}(f, t) = \sum_{i=1}^n X^n_{ij}(f, t)$, where

$$X^n_{ij}(f, t) = n^{-1/2} \int_0^t f(b^n_{ij}(s)) \left( W_i - \frac{\bar{Y}_{ij}^{(1)}(s)}{\bar{Y}_{ij}^{(1)}(s) + \bar{Y}_{ij}^{(2)}(s)} \right) dN_{ij}(s),$$

$f \in H_j, \{b^n_{ij} \equiv (b^n_{ij}, \ldots, b^n_{ij})^T, n \geq 1\}, \{b^n_{ij}, n \geq 1\} \in B(\{F^n, n \geq 1\})$, each $H_j$ is either equal to or is a closed subset of some $G^+(K_l), K_l \subset \infty, l = 1 \ldots r, j = 1 \ldots J$, and where $r$ is finite. For a discussion of how to choose $\{b^n_{ij}, n \geq 1\}$ and $H_j, j = 1 \ldots J$, see Kosorok and Lin (1998).

The approximation processes—of the form (S2)—we will be using for the covariance and Monte Carlo estimation are

$$\tilde{X}^n_{ij}(f, t) \equiv n^{-1/2} \int_0^t f(b^n_{ij}(s)) \left( W_i - \frac{\bar{Y}_{ij}^{(1)}(s)}{\bar{Y}_{ij}^{(1)}(s) + \bar{Y}_{ij}^{(2)}(s)} \right) d\tilde{M}^{n(2-W_i)}_{ij}(s),$$

where

$$\tilde{M}^{n(l)}_{ij}(t) = \int_0^{t_{\tau^n_j}} \left\{ 1 - \frac{1}{\bar{Y}_{ij}^{(l)}(s)} \right\}^{-1/2} \left\{ dN_{ij}(s) - Y_{ij}(s) \frac{d\bar{N}^{n(l)}_{ij}(s)}{\bar{Y}_{ij}^{(l)}(s)} \right\},$$

$$M^{n(l)}_{ij}(t) \equiv N_{ij}(t) - \int_0^t Y_{ij}(s) dN_{ij}(s),$$

for $i = 1 \ldots n$ and $l = 1, 2$, and where

$$\tau^n_j = \sup \left\{ t : \bar{Y}_{ij}^{(1)}(t) \wedge \bar{Y}_{ij}^{(2)}(t) > 1 \right\},$$

$$\bar{N}^{(1)}_{ij}(t) = \sum_{i=1}^n W_i N_{ij}(t),$$

$$\bar{N}^{(2)}_{ij}(t) = \sum_{i=1}^n (1 - W_i) N_{ij}(t),$$

$$\bar{M}^{n(1)}_{ij}(t) = \sum_{i=1}^n W_i M^n_{ij}(t),$$

$$\bar{M}^{n(2)}_{ij}(t) = \sum_{i=1}^n (1 - W_i) M^n_{ij}(t),$$

for $j = 1 \ldots J$. 

15
Theorem 5 Under Model 1, the $J$ statistics with components of the form (4.1) satisfy condition (S1) and the approximations of the form (4.2) satisfy condition (S2), for all $u_j \in [0, \infty]$ (although $u_j^*$ may not necessarily be equal to $\infty$), and

$$
\mu_j(f_j, t) = \beta \int_0^\infty f_j^j(b_j(s)) \phi(s) \left\{ \frac{p(1-p) \pi_j^{(1)}(s) \pi_j^{(2)}(s)}{p \pi_j^{(1)}(s) + (1-p) \pi_j^{(2)}(s)} \right\} \{1 - \Delta_{0}(s)\} d\Delta_{0}(s),
$$

where $b_j$ is the in-probability limit of $\{b_j^n, n \geq 1\}, j = 1 \ldots J$, and where $f \in H$. Furthermore, conditions (A1), (B1), (J1), (J2), (T1), (C1), (A2), (B2), (C2), and (B3) are all satisfied, and therefore the conclusions of Theorems 2, 3, and 4 are valid.

Proof. For $j = 1 \ldots J$, condition (S1) clearly holds for $X^n_j(\cdot, \cdot)$. Since the limiting Gaussian process given in Theorem 3 of Kosorok (1998) (hereafter referred to as KT3), is continuous except for possibly cadlag jumps at a known countable collection of points, it is tight, and thus (T1) is satisfied. By Theorem 4 of Kosorok (1998), condition (C1) is also satisfied. For each $f \in H_j$,

$$
\overline{X}_j^n(f, t) = \overline{M}_j^n(f, t) + \mu_j^n(f, t),
$$

where

$$
(4.3) \quad \overline{M}_j^n(f, t) = \sum_{i=1}^n n^{-1/2} \int_0^t f(b_j^n(s)) \left( W_i - \frac{\overline{Y}_j^{(1)}(s)}{\overline{Y}_j^{(1)}(s) + \overline{Y}_j^{(2)}(s)} \right) dM_j^n(s),
$$

and

$$
\mu_j^n(f, t) = \sum_{i=1}^n n^{-1/2} \int_0^t f(b_j^n(s)) \left( W_i - \frac{\overline{Y}_j^{(1)}(s)}{\overline{Y}_j^{(1)}(s) + \overline{Y}_j^{(2)}(s)} \right) Y_i(s) d\Lambda_{loc}^{1/2-W_i}(s).
$$

By KT3, $\mu_j^n(f, t)$ converges uniformly over $t \in [0, \infty)$ to $\mu_j(f, t)$, as $n \to \infty$.

Define

$$
X_{ij}^n(f, t) = n^{-1/2} \int_0^t f(b_j(s)) \left( W_i - \frac{\pi_j^{(1)}(s)}{p \pi_j^{(1)}(s) + (1-p) \pi_j^{(2)}(s)} \right) dM_j^n(s),
$$

let $\overline{X}_j^n(f, t) = \sum_{i=1}^n X_{ij}^n(f, t)$, and notice that

$$
E \left[ \{ \overline{M}_j^n(f, t) - \overline{X}_j^n(f, t) \}^2 \right]
$$

16
\[
\leq 2E \left[ n^{-1} \sum_{i=1}^{n} \int_{0}^{t} \left\{ f\left( b_{j}^{n}(s) \right) \right\}^{2} \left( \frac{Y_{j}^{(1)}(s)}{Y_{j}^{(1)}(s) + Y_{j}^{(2)}(s)} - \frac{p_{i}^{1}(s)}{p_{i}^{1}(s) + (1-p)\pi_{j}^{2}(s)} \right)^{2} dN_{ij}(s) \right] \\
+ 2E \left[ n^{-1} \sum_{i=1}^{n} \int_{0}^{t} \left\{ f\left( b_{j}^{n}(s) \right) - f\left( b_{j}(s) \right) \right\}^{2} \left( \frac{p_{i}^{1}(s)}{p_{i}^{1}(s) + (1-p)\pi_{j}^{2}(s)} \right)^{2} dN_{ij}(s) \right].
\]

The arguments on the right-hand side can be shown to go to zero, as \( n \to \infty \), for any \( t \in [0, \infty] \) by using arguments along the lines used in the proof of KT3, and condition (A1) is thus established.

Since

\[
\sqrt{n}X_{ij}^{*n}(f,t) = \int_{0}^{t} G_{ij}(s) dM_{ij}^{n}(s),
\]

where \( |G_{ij}(\cdot)| \) is bounded by 2 and \( N_{ij}(\infty) \) is bounded by 1, we have by Lemma 1—given in the appendix—that

\[
E \left[ \left| \sqrt{n}X_{ij}^{*n}(f,t) \right|^{4} \right] \leq 720.
\]

Thus (B1) is satisfied. Furthermore, the conditions of Model 1 readily imply both (J1) and (J2).

For \( j = 1 \ldots J \), condition (S2) clearly holds for the approximations given in (4.2). If, for \( i = 1 \ldots n \) and \( k = 1 \ldots J \), we let

\[
G_{ij}^{n}(t) \equiv n^{-1/2} f\left( b_{j}^{n}(t) \right) \left( W_{i} - \frac{Y_{j}^{(1)}(t)}{Y_{j}^{(1)}(t) + Y_{j}^{(2)}(t)} \right)
\]

and

\[
G_{ij}(t) \equiv n^{-1/2} f\left( b_{j}(t) \right) \left( W_{i} - \frac{p_{i}^{1}(t)}{p_{i}^{1}(t) + (1-p)\pi_{j}^{2}(t)} \right)
\]

then

\[
E \left[ \left\{ \bar{X}_{ij}^{*n}(f,t) - X_{ij}^{*n}(f,t) \right\}^{2} \right] \\
= E \left[ \int_{0}^{t} G_{ij}^{n}(s) \left[ I\left\{ Y_{j}^{(1)}(s) Y_{j}^{(2)}(s) > 1 \right\} \right] \left( 1 - \frac{1}{Y_{j}^{(2-W_{i})}(s)} \right)^{-1/2} \right. \\
\times \left. \left( dM_{ij}^{n}(s) - Y_{ij}(s) \frac{dM_{ij}^{n(2-W_{i})}(s)}{Y_{j}^{(2-W_{i})}(s)} \right) - dM_{ij}^{n}(s) \right] + \left[ G_{ij}^{n}(s) - G_{ij}(s) \right] dM_{ij}^{n}(s) \right]^{2} \\
\leq 2E \left[ \int_{0}^{t} \left\{ \left[ G_{ij}(s) \right]^{2} \delta_{j}^{(2-W_{i})n}(s) + \left[ G_{ij}^{n}(s) - G_{ij}(s) \right]^{2} \right\} Y_{ij}(s) d\Lambda_{j}^{n(2-W_{i})}(s) \right],
\]

17
where
\[
\delta_j^{(i)n}(s) = 1 - \frac{1}{\left\lfloor \frac{1}{\overline{Y}_j^{(i)}(s) \wedge Y_j^{(2)}(s)} > 1 \right\rfloor} \left\{ 2 \left(1 - \frac{1}{\overline{Y}_j^{(i)}(s)} \right)^{1/2} - 1 \right\}
\]
for \( l = 1, 2 \); and thus for any \( t \in I_j \),
\[
\sum_{i=1}^{n} E \left[ \left\{ \overline{X}^{*n}_{ij}(f, \infty) - X^{*n}_{ij}(f, \infty) \right\}^2 \right] 
\leq 2 E \left[ \sum_{i=1}^{n} \int_0^t \left\{ \left[ G^{n}_{ij}(s) \right]^2 \delta_j^{(2-W_i)n}(s) + \left[ G^{n}_{ij}(s) - G_{ij}(s) \right]^2 \right\} Y_{ij}(s) d\Lambda^{n}_{2-W_j}(s) \right] 
+ 2 E \left[ \sum_{i=1}^{n} \int_t^\infty \left\{ \left[ G^{n}_{ij}(s) \right]^2 \delta_j^{(2-W_i)n}(s) + \left[ G^{n}_{ij}(s) - G_{ij}(s) \right]^2 \right\} d\Lambda^{n}_{2-W_j}(s) \right] 
\leq 2 E \left[ \max_{i=1,2} \sup_{s\in[0,t]} \delta_j^{(i)n}(s) \right] + 2 E \left[ \max_{i=1,2} \sup_{s\in[0,t]} \left| G^{n}_{ij}(s) - G_{ij}(s) \right|^2 \right] 
+ 10 \sum_{l=0,1} E \left[ \int_t^\infty \frac{\overline{Y}_{ij}^{(1)}(s)\overline{Y}_j^{(2)}(s)}{\overline{Y}_j^{(1)}(s) + \overline{Y}_j^{(2)}(s)} \left\{ \frac{\overline{Y}_j^{(i+1)}(s)}{n} \right\} d\Lambda^{n}_{2-W_j}(s) \right],
\]
which can be shown to converge to zero by letting \( t = u^*_j \) if \( I_j \) is right-closed, or by letting \( t \uparrow u^*_j \) if \( I_j \) is right-open, using arguments similar to those used in the proof of KT3. Thus (A2) is satisfied, and it is clear from the preceding arguments that (B2) is also satisfied.

Arguments based on the reproducing kernel Hilbert space properties of \( G^+_*(K) \), similar to those given in the proofs of Theorems 2 and 4 of Kosorok (1998), can be used to establish (C2) and (B3). Finally, (T2) follows from Theorem 2 of Kosorok and Lin (1998). □

**Corollary 1** Suppose the conditions of Theorem 5 obtain and
\[
\inf_{i \leq j \leq J} \inf_{f \in H_j} V_{jj}(f^i, f^j; \infty, \infty) > 0,
\]
where \( V_{jk}(f^j, g^k; s_j, t_k) \), \( j, k = 1 \ldots J \), is as defined in Theorem 2. Then, for all \( \{u_1, \ldots, u_J\}^T \in ([0, \infty])^J \),

(i) The standardized statistics
\[
\left\{ \left[ \overline{V}_{jj}(f^i, f^j; \infty, \infty) \right]^{-1/2} \overline{X}_{ij}^{n}(f^j, \cdot), j = 1 \ldots J \right\}
\]

18
converge weakly on \(\{A(H, D), \alpha\}\) to a Gaussian process \(Z_0(f)\), with mean function

\[
\nu(f) \equiv \left\{ \left[ V_{jj}(f^j, f^j; \infty, \infty) \right]^{-1/2} \mu_j(f^j, \cdot), j = 1 \ldots J \right\}
\]

and covariance function

\[
(4.5) \quad \left[ V_{jj}(f^j, f^j; \infty, \infty) V_{kk}(g^k, g^k; \infty, \infty) \right]^{-1/2} V_{jk}(f^j, g^k; s_j, t_k),
\]

for \(f^j, g^j \in H_j, s_j, t_j \in [0, u_j], j, k = 1 \ldots J\).

(ii) The covariance estimator

\[
\left\{ \left[ \hat{V}_{jj}(f^j, f^j; \infty, \infty) \hat{V}_{kk}(g^k, g^k; \infty, \infty) \right]^{-1/2} \hat{V}_{jk}(f^j, g^k; s_j, t_k) \right\}
\]

is uniformly consistent for (4.5) over all \(f^j, g^j \in H_j\) and \(s_j, t_j \in [0, \infty], j = 1 \ldots J\).

(iii) The standardized Monte Carlo replicates

\[
\left\{ \left[ \hat{V}_{jj}(f^j, f^j; \infty, \infty) \right]^{-1/2} \hat{X}^n_{jq}(f^j, \cdot), j = 1 \ldots J, q = 1 \ldots Q \right\}
\]

converge weakly on \(\{A(H, D), \alpha\}^Q\) to \(Q\) independent replicates of a mean zero Gaussian process on \(A(H, D)\) with the same distribution as \(Z_0(\cdot) - \nu(\cdot)\) given above.

Proof. These results follow directly from Theorem 5 and the version of Slutsky’s theorem given as Corollary 3.3 in Chapter 3 of Ethier and Kurtz (1986).

4.2. Combination with a binary outcome. The following is a simple contiguous alternative model for a binary outcome in a group sequential trial with staggered patient entry:

MODEL 2 For \(i = 1 \ldots n\), let \(U_i^*\) denote the calendar time of observation, \(E_i \in \{0, 1\}\) the observed binary response (with \(E_i = 1\) indicating failure), and \(W_i\) the indicator of treatment \(1\), for patient \(i\). Assume that \(U_i^*\) is statistically independent of \(E_i\) conditional on \(W_i\), and that the random vectors \(\{U_i^*, E_i, W_i\}^T, i = 1 \ldots n\), are independent and identically distributed,
with $P[E_i = 1] = \gamma_{(2-W_i)}^n$, the remaining components of the distribution fixed, and $E[W_i] = p \in (0,1)$. Let $A_1 \ldots A_J$ be $J$ fixed calendar times for interim analysis, and define

$$\delta_{ij} \equiv I_{\{U_i^* \geq A_j\}},$$

for $j = 1 \ldots J$. Suppose also that for some $\gamma_0 \in (0,1)$ and some finite $\zeta$,

$$\gamma_{(1)}^n = \left[ \left( \gamma_0 + \frac{\zeta}{2\sqrt{n}} \right) \lor 0 \right] \land 1$$

and

$$\gamma_{(2)}^n = \left[ \left( \gamma_0 - \frac{\zeta}{2\sqrt{n}} \right) \lor 0 \right] \land 1.$$

For $i = 1 \ldots n$ and $j = 1 \ldots J$, define

$$\bar{\xi}_j^n = \sum_{i=1}^n \xi_{ij}^n,$$

where

$$\xi_{ij}^n = \left( W_i - \frac{\overline{d}_j^{(1)}}{\overline{d}_j^{(1)} + \overline{d}_j^{(2)}} \right) \frac{\delta_{ij} E_i}{\sqrt{n}},$$

$$\overline{d}_j^{(1)} = \sum_{i=1}^n W_i \delta_{ij}, \text{ and}$$

$$\overline{d}_j^{(2)} = \sum_{i=1}^n (1 - W_i) \delta_{ij}.$$

The statistic $\bar{\xi}_j^n$ is essentially the difference in failure proportions between treatments observed at time $A_j$.

Define also

$$\tilde{\xi}_{ij}^n = I_{\left\{ \frac{\overline{d}_j^{(1)}}{\overline{d}_j^{(1)} + \overline{d}_j^{(2)}} > 1 \right\}} \left( 1 - \frac{1}{\overline{d}_j^{(2-W_i)}} \right)^{-1/2} \left( W_i - \frac{\overline{d}_j^{(1)}}{\overline{d}_j^{(1)} + \overline{d}_j^{(2)}} \right) \left( E_i - \frac{\overline{E}_j^{(2-W_i)}}{\overline{d}_j^{(2-W_i)}} \right) \frac{\delta_{ij} E_i}{\sqrt{n}},$$

for $i = 1 \ldots n$, where

$$\overline{E}_j^{(1)} = \sum_{i=1}^n W_i \delta_{ij} E_i; \text{ and}$$

$$\overline{E}_j^{(2)} = \sum_{i=1}^n (1 - W_i) \delta_{ij} E_i;$$

20
and construct the variance estimator $\tilde{U}$, where
\[
[\tilde{U}]_{jk} = \sum_{i=1}^{n} \hat{\xi}_{ij} \hat{\xi}_{ik},
\]
for $j, k = 1 \ldots J$.

The following corollary establishes the large sample results for model 2 which parallel the results of subsection 4.1:

**Corollary 2** Suppose model 2 obtains. Then

(i)
\[
\bar{\xi}^n = \left\{\bar{\xi}_1^n, \bar{\xi}_2^n, \ldots, \bar{\xi}_J^n\right\}^T
\]
converges in distribution, as $n \to \infty$, to a multivariate normal random variable with mean vector
\[
\gamma^0 = \left\{\gamma_1^0, \gamma_2^0, \ldots, \gamma_J^0\right\}^T,
\]
where
\[
\gamma_j^0 = \frac{p(1 - p)\phi_j^{(1)}\phi_j^{(2)}}{p\phi_j^{(1)} + (1 - p)\phi_j^{(2)}}
\]
$\phi_j^{(1)}$ is the limiting value of
\[
\frac{d_j^{(1)}}{\sum_{i=1}^{n} W_i},
\]
and $\phi_j^{(2)}$ is the limiting value of
\[
\frac{d_j^{(2)}}{\sum_{i=1}^{n}(1 - W_i)},
\]
for $l = 1, 2$ and $j = 1 \ldots J$, and with covariance $U$, with elements $[U]_{jk}$, $j, k = 1 \ldots J$;

(ii) the covariance estimator $\tilde{U}$ is consistent for $U$; and

(iii) if $Z_{iq}$, $i = 1 \ldots n$ and $q = 1 \ldots Q$, are independent standard normal random variables, and if
\[
\bar{\xi}_q^n = \left\{\bar{\xi}_{1q}^n, \bar{\xi}_{2q}^n, \ldots, \bar{\xi}_{Jq}^n\right\}^T,
\]
where

\[ \xi_{ij}^n = \sum_{i=1}^{n} Z_{ij} \xi_{ij}, \]

for \( j = 1 \ldots J \), then the collection \( \{ \xi_{ij}^n, q = 1 \ldots Q \} \) converges in distribution to \( Q \) independent replicates of a mean zero multivariate normal distribution with covariance \( U \).

Proof. Standard asymptotic arguments could be used for all steps of this proof, but it will be useful to follow arguments established in previous sections. Clearly, \( \xi_j^n \) will satisfy (S1) if \( H_j = \{0\} \) and \( u_j = 0 \). (T1) is also clearly satisfied. Standard arguments can be used to show that \( \xi_j^n \) is asymptotically equivalent to

\[ \gamma_j^0 + \sum_{i=1}^{n} \xi_{ij}^n, \]

where

\[ \xi_{ij}^n = \left( W_i - \frac{p \phi_j^{(1)}}{p \phi_j^{(1)} + (1 - p) \phi_j^{(2)}} \right) \left( E_i - \gamma_{2-W_i} \right) \frac{\delta_{ij}}{\sqrt{n}}, \]

and (A1) is satisfied for \( j = 1 \ldots J \). Conditions (B1), (J1), and (J2) are also easily satisfied. Condition (S2) is readily satisfied for \( \hat{\xi}_{ij}^n \), \( i = 1 \ldots n \) and \( j = 1 \ldots J \); and so are conditions (A2), (B2), and (T2). Conditions (C1), (C2), (C3), and (B3) are also easily satisfied by the triviality of \( H_j \) and \( u_j \), \( j = 1 \ldots J \), and the results of Theorems 2, 3, and 4 now apply. \( \square \)

The following model combines models 1 and 2 to form a bivariate model with both a failure time and a binary outcome:

**Model 3** Assume that both models 1 and 2 hold and that the random vectors

(4.6)

\[ \{ U_i, T_i, C_i^*, U_i^*, E_i, W_i \}^T, \]

\( i = 1 \ldots n \), are independent and identically distributed, with the distribution appropriately depending on \( n \). Assume also that the vectors given in (4.6) jointly converge in distribution as \( n \to \infty \).
Assume the notation used in both sections 4.1 and 4.2. Define

\[ \mathcal{X}^n(\vec{f}) \equiv \left\{ \bar{X}^n_1(f^1, \cdot), \ldots, \bar{X}^n_J(f^J, \cdot), \bar{X}^n_{J+1}(0,0), \ldots, \bar{X}^n_{2J}(0,0) \right\}, \]

where, for \( j = J + 1, \ldots, 2J \), \( \bar{X}^n_j(0,0) \equiv \bar{e}^n_j \) and

\[ \tilde{f} \equiv \left\{ f^1, \ldots, f^{2J} \right\}^T \in \tilde{H} \equiv H^J \times \{0\}^J, \]

where \( H \) is defined in section 4.1. Also, for any \( \{u_1, \ldots, u_J\}^T \in ([0, \infty])^T \), define \( \tilde{D} \equiv \{D[0, u_1] \times \cdots \times D[0, u_J]\} \times \mathbb{R}^J \), where \( \mathbb{R} \) is the space of real numbers endowed with the Euclidean metric; and for any \( x, y \in A(\tilde{H}, \tilde{D}) \), define the metric

\[ \hat{\alpha} \equiv \sup_{\tilde{f} \in \tilde{H}} \tilde{d}\left(x(\tilde{f}), y(\tilde{f})\right), \]

where \( \tilde{d} \) is the appropriate Skorohod metric for \( \tilde{D} \).

Define \( \bar{X}^n_{i,j}(f^j, t_j) \), for \( j = 1 \ldots 2J \), such that

\[ \bar{X}^n_{i,j}(f^j, t_j) \equiv \bar{X}^n_{i,j}(f^j, t_j) \]

for \( j = 1 \ldots J \) and

\[ \bar{X}^n_{i,j}(f^j, t_j) \equiv \bar{e}^n_{i,[j]}, \]

for \( j = J + 1 \ldots 2J \), where \([j] \equiv j - J \) and

\[ \tilde{t} \equiv \{t_1, \ldots, t_{2J}\}^T \in ([0, \infty])^J \times \{0\}^J. \]

The following corollary combines the results of subsection 4.1 with corollary 2 to establish the large sample distributional results for model 3:

**Corollary 3** Suppose model 3 obtains. Then, for any \( \{u_1, \ldots, u_J\}^T \in ([0, \infty])^J \),

(i) \( \bar{X}^n(\cdot) \) converges weakly on \( \left\{ A(\tilde{H}, \tilde{D}), \hat{\alpha} \right\} \) to a Gaussian process \( \bar{Z}(\tilde{f}) \) with mean function

\[ \bar{\mu}(\tilde{f}) \equiv \{\mu_1(f^1, \cdot), \ldots, \mu_J(f^J, \cdot), \gamma_1^0, \ldots, \gamma_J^0\}, \]

\[ \bar{\mu}(\tilde{f}) \equiv \{\mu^1(f^1, \cdot), \ldots, \mu^J(f^J, \cdot), \gamma_1^0, \ldots, \gamma_J^0\}, \]

23
where \( \mu_j(\cdot, \cdot) \) and \( \gamma_j^0 \) are defined in corollaries 1 and 2, and with covariance function
\[
\left[ V_{\tilde{f}, \tilde{g}}(\tilde{s}, \tilde{t}) \right]_{jk} = V_{jk}(f^j, g^k; s_j, t_k),
\]
\( j, k = 1 \ldots 2J, \) where
\[
\tilde{g} = \{ g^1, \ldots, g^{2J} \} \in \tilde{\mathcal{H}}
\]
and
\[
\tilde{s} = \{ s_1, \ldots, s_{2J} \}^T \in ([0, \infty])^J \times \{ 0 \}^J ;
\]
\( (ii) \) the covariance estimator
\[
\left[ \tilde{V}_{\tilde{f}, \tilde{g}}(\tilde{s}, \tilde{t}) \right]_{jk} = \sum_{i=1}^n \tilde{X}_{ij}^n(f^j, s_j, \tilde{X}_{ik}^n(g^k, t_k),
\]
for \( j, k = 1 \ldots 2J, \) is uniformly consistent for \( V_{\tilde{f}, \tilde{g}}(\tilde{s}, \tilde{t}) ; \) and
\( (iii) \) if \( Z_{iq}, i = 1 \ldots n \) and \( q = 1 \ldots Q, \) are independent standard normal random variables,
and if
\[
\tilde{X}_q^n(\bar{f}) \equiv \left\{ \bar{X}_q^n(\bar{f}), [\xi_q^n]^T \right\},
\]
then the collection \( \{ \bar{X}_q^n(\cdot), q = 1 \ldots Q \} \) converges weakly on \( \{ A(\tilde{\mathcal{H}}, \tilde{\mathcal{D}}), \tilde{\alpha} \}^Q \) to \( Q \) independent replicates of \( \tilde{Z}(\cdot) - \tilde{\mu}(\cdot) \) (the null distribution of the limiting process given in part \( (i) \) above).

Proof. Conditions (S1), (T1), (A1), (B1), (S2), (A2), (B2), (T2), (C1), (C2), (C3), and (B3) are satisfied by Theorem 5 and Corollary 2. Since condition (J1) is satisfied by assumption and, by the convergence in distribution of the underlying data as \( n \to \infty \) (J2) is also satisfied, the results follow. \( \square \)

**Remark 2** The version of Slutsky's theorem given as Corollary 3.3 in Chapter 3 of Ethier and Kurtz (1986) can be easily applied to Corollary 3 to obtain results for the standardized statistics of the form given in Corollary 1.
5. **Discussion.** The generality of the forgoing large sample results allows for the utilization of several complex continuous mappings of asymptotically Gaussian statistics for several outcome measures in a group sequential clinical trial with staggered patient entry. As an example, a Renyi-type statistic—such as the supremum of the absolute value over time of the log-rank—could be used for a primary failure time outcome while a simple comparison of proportions could be used for a binary toxicity outcome: in this case, the underlying statistics are asymptotically Gaussian, but the supremum versions of the statistics are not (see chapter 2, section 11 of Billingsley, 1968, for example). The results of Section 4.2 apply to this setting (as well as to the more complex function-indexed setting) and allow an accurate estimation of the corresponding null distribution so that critical regions can be constructed during clinical trial execution.

A simple approach to constructing such critical regions would be as follows. First a decision as to how much of the experiment-wise type I error rate \( c \) to "spend" at each of the \( J \) looks must be made so that \( c_1 + c_2 + \cdots + c_J = c \). At the first look, \( N \) Monte Carlo replicates of the null distribution at look 1 can be made based on the available data. Based on these replicates, either a rectangular or elliptical critical region can then be formed such that \( c_1 N \) replicates are rejected. At the second look (if rejection has not yet occurred), \( N \) more replicates of the null distribution at both looks 1 and 2 can be made and the distribution, conditional on not having exceeded the critical region at look 1, can be estimated such that \( c_2 N \) replicates are rejected at look 2. This process can be repeated in a similar manner until the conclusion of the clinical trial.

Many other combinations and possibilities of outcomes can be incorporated into the scheme presented in this paper. Once the marginal conditions given in section 2 have been established for each of the individual statistics at each interim analysis time, very little additional work is required to obtain joint convergence and asymptotic validity of the Monte Carlo replicates.
Acknowledgments. The funding for this research was provided by a University of Wisconsin-Madison Graduate School Research Award. I would also like to thank David L. DeMets and Richard J. Chappell for a number of enlightening discussions on multivariate outcomes in group sequential clinical trials.

APPENDIX

Lemma 1 Let \( N(t) \) be a bounded counting process on \([0, \infty]\) adapted to the right-continuous filtration \( \mathcal{F}_t \) such that \( N(\infty) \leq K_N < \infty \) almost surely (a.s.) and \( N(0) = 0 \) a.s. with right-continuous predictable compensator \( A(t) \) such that \( A(0) = 0 \) a.s.. Let \( H(t) \) be a bounded \( \mathcal{F}_t \)-predictable process such that \( \sup_{s \in [0, \infty]} |H(s)| \leq K_H < \infty \) a.s.. Let \( M(t) = N(t) - A(t) \). Then, \( \forall t \in [0, \infty] \),

\[
E \left[ \int_0^t H(s) dM(s) \right]^4 \leq 45 K_H^4 K_N^2.
\]

Proof.

\[
\left[ \int_0^t H(s) dM(s) \right]^4 = 4 \int_0^t \left( \int_0^s H(x) dM(x) \right)^3 H(s) dM(s) + 6 \int_0^t \left( \int_0^s H(x) dM(x) \right)^2 H^2(s) [dM(s)]^2 + 4 \int_0^t \left( \int_0^s H(x) dM(x) H^3(s) [dM(s)]^3 + \int_0^s H^4(s) [dM(s)]^4 \right),
\]

by Itô’s formula, where

\[
[dM(s)]^2 = \{1 - \Delta A(s)\} dN(s) - \Delta A(s) dM(s),
\]

\[
[dM(s)]^3 = \{1 - \Delta A(s)\} \{1 - 2\Delta A(s)\} dN(s) + [\Delta A(s)]^2 dM(s),
\]

\[
[dM(s)]^4 = \{1 - \Delta A(s)\} \{1 - 3\Delta A(s) + [\Delta A(s)]^2\} dN(s) - [\Delta A(s)]^3 dM(s).
\]

Thus,

\[
(A.1) \quad E \left[ \int_0^t H(s) dM(s) \right]^4 \leq 6E \left[ \int_0^t \left( \int_0^s H(x) dM(x) \right)^2 H^2(s) \{1 - \Delta A(s)\} dN(s) \right]
\]

26
\[ + 4E \left[ \int_0^t \left[ \int_0^{s^-} H(x) dM(x) \right] H^3(s) \{1 - \Delta A(s)\} \{1 - 2\Delta A(s)\} dN(s) \right] + E \left[ \int_0^t H^4(s) \{1 - \Delta A(s)\} \left\{1 - 3\Delta A(s) + [\Delta A(s)]^2\right\} dN(s) \right]. \]

Now, for some predictable \(B : |B(s)| \leq b < \infty\) for all \(s \geq 0\),

\[
\int_0^t \left[ \int_0^{s^-} H(x) dM(x) \right] B(s) dN(s) = \int_0^t H(s) dM(s) \times \int_0^t B(s) dN(s) - \int_0^t \left[ \int_0^{s^-} B(x) dN(x) \right] H(s) dM(s) - \int_0^t B(s) H(s) \{1 - \Delta A(s)\} dN(s),
\]

which implies that

\[
(A.2) \quad \left| E \left[ \int_0^t \left\{ \int_0^{s^-} H(x) dM(x) \right\} B(s) dN(s) \right] \right| \\
\leq \sqrt{E \left[ \left\{ \int_0^t H(s) dM(s) \right\}^2 \left\{ \int_0^t B(s) dN(s) \right\}^2 \right] + bK_H K_N} \\
\leq bK_H \left( K_N^{3/2} + K_N \right).
\]

Next,

\[
\int_0^t \left[ \int_0^{s^-} H(x) dM(x) \right]^2 dN(s) \\
= \left[ \int_0^t H(x) dM(x) \right]^2 N(t) - \int_0^t N(s-) d \left[ \int_0^s H(x) dM(x) \right]^2 \\
- \int_0^t \Delta N(s) d \left[ \int_0^s H(x) dM(x) \right]^2 \\
= \left[ \int_0^t H(x) dM(x) \right]^2 N(t) \\
- \int_0^t N(s-) \left\{ 2H(s) \int_0^{s^-} H(x) dM(x) - H^2(s) \Delta A(s) \right\} dM(s) \\
- 2 \int_0^t \int_0^{s^-} H(x) dM(x) H(s) \{1 - \Delta A(s)\} dN(s) \\
- \int_0^t \{ N(s-) + 1 - 2\Delta A(s) \} H^2(s) \{1 - \Delta A(s)\} dN(s),
\]

thus

\[
\left| E \left[ \int_0^t \left\{ \int_0^{s^-} H(x) dM(x) \right\}^2 dN(s) \right] \right|
\]

27
\[
\begin{align*}
&\leq 2K_H^2 K_N^2 + 2 \left| E \left[ \int_0^t \int_0^{x^-} H(x) dM(x) H(s) \{1 - \Delta A(s)\} dN(s) \right] \right| \\
&\leq 2K_H^2 K_N^2 + 2K_H^2 \left( K_N^{3/2} + K_N \right) \\
&\leq 6K_H^2 K_N^2
\end{align*}
\]

by A.2. Hence,

\[
(A.1) \leq 36K_H^4 K_N^4 + 4K_H^4 \left( K_N^{3/2} + K_N \right) + K_H^4 K_N,
\]

and the results follows. \(\Box\)
REFERENCES


