Technical Report Accompanying

Monitoring a general class of two-sample survival statistics with applications

BY MINGGAO GU

Department of Mathematics and Statistics, McGill University, Montréal, Québec H3A 2K6, Canada

DEAN FOLLMANN AND NANCY L. GELLER

Office of Biostatistics Research, National Heart Lung and Blood Institute, 2 Rockledge Center, Bethesda, MD 20892-7938, U.S.A.

This is a technical report accompanying the paper "Monitoring a general class of two-sample survival statistics with applications". The main paper shall be refereed to as GFG in this report.

This report is composed of three parts. The first part provides the proof of (3.1) of GFG. The second part derives (3.3) and (3.4) of GFG. The third part formulates and proves an asymptotic convergence theorem for the sequential monitoring of the H-class of statistics proposed in GFG.

1. Proof of the asymptotic representation (3.1) of GFG

Conditions on $\hat{H}(u)$ and $H(u)$ for (3.1) of GFG to hold and therefore for the asymptotic normality of $\{n^{1/2}\hat{W}\}$ are:

Assumption 1. The function $H(u)$ is piecewise continuous with finitely many discontinuity points. Except on this finite set of discontinuity points of $[0, T_c]$, we have for all other $u \in [0, T_c]$,

$$\hat{H}(u) \to H(u) \quad \text{in probability as } n \to \infty. \tag{1.1}$$

Assumption 2. Let $V_{[a,b]}\{f(u)\}$ denote the total variation of the function $f(u)$ on the interval $[a, b]$. The functions $\hat{H}(u)$ and $H(u)$ are such that

$$V_{[0,T_c]}\{\hat{H}(u)\} + V_{[0,T_c]}\{H(u)\} \quad \text{is bounded in probability.}$$

Assumption 3. Denote by $Y_i(u)$ the risk set size at time u for group i. Then $E[Y_i(T_c)] \geq cn$, $i = 1, 2$ for some fixed constant $c > 0$.

If we assume that $\hat{H}(u) \to H(u)$ uniformly in probability, the theorem presented in this paper would exclude Case 2 introduced in Section 2 of GFG. Fortunately, (3.1) of GFG holds under the weaker Assumption 1.

The proof of (3.1) of GFG is based on Lemma 1 stated below. For a similar lemma see Proposition II.8.6 of Anderson *et al.* (1992, page 113) and the remark following that proposition. Lemma 1 is also a consequence of Helly-Bray lemma. We shall use the conventional notations of $D[a, b]$ and $C[a, b]$ to denote the spaces of right continuous functions with left hand limits and continuous functions on $[a, b]$ respectively.

LEMMA 1 *Suppose that there exists a sequence of processes $\{B_n(u), a \leq u \leq b\}$ defined on $D[a, b]$ which converges weakly to a process $\{B(u), a \leq u \leq b\}$, where $B(\cdot)$ takes values in $C[a, b]$. Also assume that $\hat{H}(u)$ and $H(u)$ defined on $[a, b]$ satisfy Assumptions 1 and 2, with interval $[0, T_c]$ replaced by $[a, b]$ and in addition, assume that H is continuous on $[a, b]$. Then $\int_a^s (\hat{H}(u) - H(u)) dB_n(u)$ for $s \in [a, b]$, defined by integration by parts if necessary, converges to zero in probability.*

Proof. First, by strong embbeding theory, we can assume that, possibly on a different probability space, B_n converges to B and $\hat{H}(u)$ converges to $H(u)$ almost surely.

Because of Assumption 2, for any $\epsilon > 0$ we may choose M large and Ω_1, a subset of the probability space with $\Pr\{\Omega_1\} \geq 1 - \epsilon/2$, such that $V_{[a,b]}[\hat{H}(u)] + V_{[a,b]}[H(u)] \leq M$ for $\omega \in \Omega_1$. Since B_n converges to B weakly and B takes values in $C[a, b]$, by the definition of the weak convergence, for the same ϵ and for all n, there exist N, $b_1, \ldots, b_N \in C[a, b]$ and Ω_2, a subset of the probability space with $\Pr\{\Omega_2\} \geq 1 - \epsilon/2$, such that for any $\omega \in \Omega_2$, there exists b_i with $\sup_u |B_n(u, \omega) - b_i(u)| \leq \epsilon/M$.

Now for $\omega \in \Omega_1 \cap \Omega_2$, upon integration by parts,

$$\begin{aligned}
& \left| \int_a^s \left\{ \hat{H}(u) - H(u) \right\} dB_n(u, \omega) - \int_a^s \left\{ \hat{H}(u) - H(u) \right\} db_i(u) \right| \\
\leq \; & 2 \sup_u |B_n(u, \omega) - b_i(u)| \; |\hat{H}(u) - H(u)| \\
& + \int_a^s |B_n(u, \omega) - b_i(u)| \left\{ |d\hat{H}(u)| + |dH(u)| \right\} \\
\leq \; & 3 \sup_u |B_n(u, \omega) - b_i(u)| \left[V_{[a,b]}\{\hat{H}(u)\} + V_{[a,b]}\{H(u)\} \right] \leq 3\epsilon.
\end{aligned} \tag{1.2}$$

For each ω and i, integration by parts and the Helly-Bray lemma imply that

$$\int_a^s \left\{ \hat{H}(u) - H(u) \right\} db_i(u) \to 0 \quad \text{as } n \to \infty. \tag{1.3}$$

(1.2) and (1.3) imply that when n is large $\int_a^s \{\hat{H}(u) - H(u)\} dB_n(u)$ is smaller than, say, 4ϵ when n is large for every $\omega \in \Omega_1 \cap \Omega_2$, and ϵ is arbitrary, we have proved the lemma. $\square$

LEMMA 2 *Suppose that the processes $\{B_n(u), a \le u \le b\}$ and $\{B(u), a \le u \le b\}$ satisfy the same conditions as in Lemma 1. Then for any $s \in (a, b)$,*

$$\lim_{\xi \to 0} \limsup_{n \to \infty} \left| \int_{s-\xi}^{s+\xi} \hat{H}(u) dB_n(u) \right| \to 0 \tag{1.4}$$

in probability. (1.4) also holds if $\hat{H}$ is replaced by H.

Proof. Since $\hat{H}(u)$ is a function of bounded variation, we have

$$\begin{aligned} \int_{s-\xi}^{s+\xi} \hat{H}(u) dB_n(u) &= \int_{s-\xi}^{s+\xi} \hat{H}(u) d\left\{B_n(u) - B_n(s-\xi)\right\} \\ = \quad & \hat{H}(s+\xi)\left\{B_n(s+\xi) - B_n(s-\xi)\right\} - \int_{s-\xi}^{s+\xi} \left\{B_n(u) - B_n(s-\xi)\right\} d\hat{H}(u), \end{aligned} \tag{1.5}$$

where the second equality is by integration by parts. So we have

$$\begin{aligned} & \left| \int_{s-\xi}^{s+\xi} \hat{H}(u) dB_n(u) \right| \\ \le \quad & 2 \sup_{u \in [s-\xi, s+\xi]} \left| B_n(u) - B_n(s-\xi) \right| V_{[0,T_c]}\left\{\hat{H}(u)\right\}. \end{aligned}$$

The claim of (1.4) follows by observing that B_n converges weakly to B, which implies that

$$\lim_{\xi \to 0} \limsup_{n \to \infty} \sup_{u \in [s-\xi, s+\xi]} \left| B_n(u) - B_n(s-\xi) \right| = 0 \tag{1.6}$$

in probability. The proof may be repeated with $H(u)$ replacing $\hat{H}(u)$. $\square$

Next we prove (3.1) of GFG under Assumptions 1-3 and the fact that the Kaplan-Meier estimator process $\sqrt{n}\{\hat{S}_i(u) - S_i(u)\}$ converges weakly. Define

$$\hat{W}_i = \int_0^{T_c} \hat{H}(u) d\left\{\hat{S}_i(u) - S_i(u)\right\}. \tag{1.7}$$

Then $\hat{W} = \hat{W}_1 - \hat{W}_2 + \int_0^{T_c} \hat{H}(u) d\{S_1(u) - S_2(u)\}$. It follows that (3.1) of GFG is equivalent to

$$\hat{W}_i = W_i + o_p(n^{-1/2}), \quad i = 1, 2, \tag{1.8}$$

where W_i is defined in (3.2). Only the proof of (1.8) for $i = 1$ will be given since the proof for $i = 2$ is similar.

We first note (see Gill, 1980) that the process

$$B_n(u) = n^{1/2}\left\{\hat{S}_1(u) - S_1(u)\right\} \tag{1.9}$$

converges weakly to a Gaussian process with continuous paths on the interval $[0, T_c]$.

Lemma 2 essentially shows that the contribution of the integral around a single point s, to $\hat{W}_1$ or W_1 is negligible. Since H is piecewise continuous, the rest of the interval is composed of subintervals like $[a, b]$, where H is continuous on $[a, b]$.

Applying Lemma 1 we have that for such an interval $[a, b]$, as $n \to \infty$

$$\sqrt{n}\int_a^b \left\{\hat{H}(u) - H(u)\right\} d\left\{\hat{S}_1(u) - S_1(u)\right\} \to 0 \tag{1.10}$$

in probability.

Lemma 2 and (1.10) together imply (1.8) for $i = 1$. Since the proof of (1.8) for $i = 2$ is the same, we have shown (3.1) of GFG.

2. Derivation of the variance formulae (3.3) and (3.4) of GFG

First we note that for $i = 1, 2$ and for $Y_i(s) > 0$,

$$\hat{S}_i(s) - S_i(s) = -S_i(s)\int_0^s \frac{\hat{S}_i(u-)}{S_i(u)}\frac{dM_i(u)}{Y_i(u)}, \tag{2.1}$$

where $M_i(u) = N_i(u) - \int_0^u Y_i(v)d\Lambda_i(v)$ is the martingale process associated with the counting process $N_i(u)$. By the definition of W_i (3.2) and (2.1), we have

$$W_i = -\int_0^{T_c} H(u)S_i(u)\frac{\hat{S}_i(u-)}{S_i(u)}\frac{dM_i(u)}{Y_i(u)} - \int_0^{T_c} H(u)\left\{\int_0^{u-}\frac{\hat{S}_i(v-)}{S_i(v)}\frac{dM_i(v)}{Y_i(v)}\right\}dS_i(u)$$

Changing the order of integration in the second term,

$$W_i = -\int_0^{T_c}\left\{H(u)S_i(u) + \int_u^{T_c} H(v)dS_i(v)\right\}\frac{\hat{S}_i(u-)}{S_i(u)}\frac{dM_i(u)}{Y_i(u)}.$$

The compensator of W_i^2 at T_c is

$$\langle W_i\rangle(T_c) = \int_0^{T_c}\left\{H(u)S_i(u) + \int_u^{T_c} H(v)dS_i(v)\right\}^2\frac{\hat{S}_i^2(u-)}{S_i^2(u)}(1 - \Delta\Lambda_i(u))\frac{d\Lambda_i(u)}{Y_i(u)}.$$

A standard argument (see Fleming and Harrington, 1991) using $1 - \Delta\Lambda_i(u) = S_i(u)/S_i(u-)$ shows that the variance of W_i can be estimated by the ith term of (3.3) of GFG. For the pooled variance estimator in (3.4) of GFG, we replace S_i by S in the above derivation and the same technique applies.

3. Weak convergence of the statistic W in sequential clinical trials

The assumptions needed for the weak convergence to hold are essentially the Assumptions 1-3 in Section 1 of this report, plus a regularity assumption (Assumption 3') concerning the administrative censoring caused by staggered entry. This assumption was also made for the weak convergence of the weighted log-rank statistic (Gu and Lai, 1991). Assumptions 1', 2' and 4' below are the counterparts of Assumptions 1-3 in Appendix A.

Assumption 1'. There exists a finite set $\mathbf{s}$, such that as a function of s, $H_t(s)$ is continuous except for $s \in \mathbf{s}$ for every t. As a function of t, $H_t(u)$ is continuous for every $u \notin \mathbf{s}$. Also we have for all $0 \leq s \leq T_c(t)$, $\tau_0 \leq t \leq \tau_1$ and $s \notin \mathbf{s}$, as $n \to \infty$,

$$\hat{H}_t(s) \to H_t(s) \to 0 \quad \text{in probability.} \tag{3.1}$$

Assumption 2'. Let $V_{[a,b]}[f(u)]$ denote the total variation of the function $f(u)$ on the interval $[a, b]$. The functions $\hat{H}_t(u)$ and $H_t(u)$ satisfy

$$\sup_{\tau_0 \leq t \leq \tau_1} \left[V_{[0,T_c(t)]}\{\hat{H}_t(u)\} + V_{[0,T_c(t)]}\{H_t(u)\} \right] \quad \text{is bounded in probability.}$$

Assumption 3'. Consider a sequence of clinical trials indexed by n, Assume that for $i = 1, 2$

$$\lim_{n\to\infty} \frac{2}{n} \sum_{j=1}^{n} 1_{[Z_j=i]} \Pr\{C_j^{(n)} \geq s, t - T_j^{(n)} \geq s\} = b_i(t, s) \tag{3.2}$$

for some function $b_i(t, s)$ with $t > s$, where $C_j^{(n)}$ and $T_j^{(n)}$ are the censoring and entry times for subject j in the nth clinical trial. We shall supress the superscript n whenever the dependence is obvious. In the case where (C_j, T_j), $j = 1, \cdots$, are i.i.d., we have $b_i(t, s) = \Pr\{C_1 \geq s, T_1 \leq t - s\}$.

Assumption 4'. For some $c > 0$ and any $t \in [\tau_0, \tau_1]$

$$S_i(T_c(t)) \cdot b_i(t, T_c(t)) \geq c \quad \text{for } i = 1, 2. \tag{3.3}$$

This essentially says that there should be enough information at the point $T_c(t)$ for all $t \in [\tau_0, \tau_1]$.

Assumption 5'. $T_c(t)$ is a continuous function in t.

THEOREM 1 *Under Assumptions 1'-5' and the hypothesis that $\sqrt{n}(S_1 - S_2) \to g(u)$, where g is a continuous function, as $n \to \infty$, the stochastic process $\{n^{1/2}\hat{W}(t), \tau_0 \leq t \leq \tau_1\}$ converges* *weakly to a Gaussian process* $\{W_\infty(t), \tau_0 \le t \le \tau_1\}$ *with mean* $\mu(t) = \int_0^{T_c(t)} H_t(u)dg(u)$ *and covariance*

$$E\left\{W_\infty(t)W_\infty(t')\right\} = \sigma(t,t'),$$

with

$$\sigma(t,t') = -2\sum_{i=1}^{2}\int_0^{\min(T_c(t),T_c(t'))} \xi(t,u)\xi(t',u)\frac{dS(u)}{S^2(u)b_i(\max(t,t'),u)}, \qquad (3.4)$$

where

$$\xi(t,u) = H_t(u)S(u) + \int_u^{T_c(t)} H_t(v)dS(v).$$

Morever, the covariance function $\sigma(t,t')$ *can be consistently estimated by the function (in the case* $t > t'$*)*

$$\hat\sigma(t,t') = \sum_{i=1}^{2}\int_0^{\min(T_c(t),T_c(t'))} \hat\xi(t,u)\hat\xi(t',u)\frac{\hat S_{i,t}(u-)}{\hat S_{i,t}(u)}\frac{dN_{i,t}(u)}{Y_{i,t}^2(u)},$$

where $\hat\xi(t,u)$ *is the sample version of* $\xi(t,u)$*.*

The proof of this theorem requires two lemmas. Lemma 3 is adapted from Corollary 1 of Gu and Lai (1991). Even though Corollary 1 of Gu and Lai (1991) was proved by assuming that $\eta_{n,t}(s)$ is predictable with respect to $\{\mathcal{F}_n(s)\}$ (defined in Gu and Lai, 1991), the condition was never used in the proof. The essence of the proof of that corollary is the Helly-Bray lemma and the proof of Lemma 3 below is thus similar to the proof of Lemma 1. We refer the reader to Gu and Lai (1991) for the proof of the lemma.

LEMMA 3 *Let Assumptions 1'-5' hold. For every* $t \in [\tau_0, \tau_1]$, *let* $\{\eta_{n,t}(s), 0 \le s \le T_c(t)\}$ *be a possibly random process. Suppose that there exists a nonrandom function* $\eta(t,s)$, $0 \le s \le T_c(t)$, $t \in [\tau_0,\tau_1]$ *such that* $\sup_{\tau_0\le t\le\tau_1} |\eta(t,0)| < \infty$, $\eta(t,s)$ *is continuous in* t *and as* $n \to \infty$

$$\sup_{0\le s\le T_c(t), t\in[\tau_0,\tau_1]} |\eta_{n,t}(s) - \eta(t,s)| \to_p 0,$$

$$\sup_{t\in[\tau_0,\tau_1]} \left[V_{0\le s\le T_c(t)}\left\{\eta_{n,t}(s)\right\} + V_{0\le s\le T_c(t)}\left\{\eta(t,s)\right\}\right] = O_p(1),$$

Let $M_{i,n,t}(s)$ *be the martingale as in (2.1) but constructed only with observations available at time* t*. Then* $\{n^{-1/2}\int_0^s \eta_{n,t}(u)M_{i,n,t}(du), 0 \le s \le T_c(t), t \in [\tau_0,\tau_1]\}$ *converges weakly as* $n \to \infty$ *to a zero-mean Gaussian process* W^* *with covariance function*

$$Cov(W^*(t,s), W^*(t',s')) = -\frac{1}{2}\int^{\min(s,s')} \eta(t,u)\eta(t',u)b_i(\min(t',t),u)dS_i(u).$$

The following consequence of Lemma 3 plays an important role in the proof of Theorem 1. It is also of importance in its own right.

LEMMA 4 *Assume the same conditions as for Lemma 3. Then for* $i = 1, 2$, $\{n^{1/2}(\hat{S}_{i,t}(s) - S_i(s)), 0 \le s \le T_c(t), t \in [\tau_0, \tau_1]\}$ *converges weakly to a mean-zero Gaussian process* W^* *with covariance function*

$$Cov(W^*(t,s), W^*(t',s')) = -2S_i(s)S_i(s') \int_0^{\min(s,s')} \frac{dS_i(u)}{b_i(\max(t',t),u)S_i^2(u)}.$$

Proof of Lemma 4. Consider the process $n^{1/2}(\hat{S}_{i,t}(s) - S_i(s))/S_i(s)$. By (2.1), this process can be written as $n^{-1/2}\int_0^s \eta_{n,t}(u)M_{i,n,t}(du)$ with $\eta_{n,t}(u) = -I_{(u\le s)}n\hat{S}_{i,t}(u-)/(S_i(u)Y_{i,t}(u))$. Conditions of Lemma 3 are satisfied with $\eta(t,u) = -2I_{(u\le s)}/S_i(u)b_i(t,u)$. The result follows from Lemma 3 easily. □

Proof of Theorem 1. We start by stating a sequential analogue of Lemma 2. Namely, if $s \in \mathbf{s}$, then we have

$$\lim_{\xi\to 0}\limsup_{n\to\infty}\sup_{t\in[\tau_0,\tau_1]}\left\{n^{1/2}\int_{s-\xi}^{s+\xi}\hat{H}_t(u)d[\hat{S}_{1,t}(u) - S_1(u)]\right\} \to 0 \tag{3.5}$$

in probability. Also (3.5) holds with $\hat{H}_t(u)$ replaced by $H_t(u)$. The proof parallels the proof of (1.4) with Assumption 2' replacing Assumption 2 and using Lemma 3 instead of the process (1.9) converges weakly. We omit the details.

Define for $i = 1, 2$

$$\begin{aligned}\eta_{i,n,t}(s) &= -n\left[\hat{H}_t(u)S_i(u) + \int_u^{T_c(t)}\hat{H}_t(v)dS_i(v)\right]\frac{\hat{S}_{i,t}(u-)}{S_i(u)}/Y_{i,t}(u),\\ \eta_i(t,s) &= -2\left[H_t(u)S_i(u) + \int_u^{T_c(t)}H_t(v)dS_i(v)\right]\frac{1}{S_i(u)b_i(t,u)}.\end{aligned}$$

Then

$$\hat{W}(t) = \hat{W}_1(t) - \hat{W}_2(t) + \int_0^{T_c(t)}\hat{H}_t(u)d[S_1(u) - S_2(u)],$$

where $n^{1/2}\hat{W}_i(t) = n^{-1/2}\int_0^{T_c(t)}\eta_{i,n,t}(s)dM_{i,t}(u)$. Because of (3.5), the above integration around the point $s \in \mathbf{s}$ is negligible. The rest of the interval is the union of such intervals $[a,b]$ where for all $s \in [a,b]$, (3.1) holds and $H_t(s)$ is continuous in both s and t, thus implies that the conditions of Lemma 3 hold (restricted to $s \in [a,b]$). Applying Lemma 3, we have that $n^{1/2}\hat{W}_i(t)$ converges weakly to a Gaussian process $W_{i,\infty}(t)$ where $\mathrm{E}[W_{i,\infty}(t)W_{i,\infty}(t')]$ is the same as the ith term on the right hand side of (3.4). Theorem 1 follows. □

Corollary 1 *If assumptions of Theorem 1 hold and if, in addition $T_c(t)$ is non-increasing in t, and*

$$\xi(t,u)\xi(t,u) = \xi^2(\max(t,t'),u) \text{ for } t,t' \in [\tau_0,\tau_1] \text{ and } u \leq \min(T_c(t),T_c(t')), \tag{3.6}$$

then $\{n^{-1/2}\hat{W}(t)/\hat{V}(t), \tau_0 \leq t \leq \tau\}$ converges weakly to a Gaussian process with independent increments.

Remark. Many statistics in the H-class of statistics represented by (4.2) of GFG have independent increments. For the difference in Kaplan-Meier estimates statistic in Case 1, if $\tau_0 > s$ then $\xi(t,u) = S(s)$ and (3.6) is satisfied. For the median test in Case 2, if $\tau_0 > m$ then $\xi(t,u) = S(m)$ so (3.6) is again satisfied. For the truncated Efron test, Case 3, we have $\xi(t,u) = \{S^2(T_c(t)) - S^2(u)\}/2$. Condition (3.6) is satisfied only if $\tau_0 > T_c(t) = T_c$. For the Pepe-Fleming class of statistics considered in Case 4, $\xi(t,u) = \int_u^{T_c(t)} S(v)\hat{w}(v)dv$, so (3.6) is satisfied with $\tau_0 > T_c(t) = T_c$. In general, if both $T_c(t)$ and $H_t(u)$ do not depend on t and $\tau_0 > T_c$, then the process has independent increments.

Proof of Corollary 1. If $T_c(t)$ is non-increasing and $\xi(t,u)\xi(t',u) = \xi^2(\max(t,t'),u)$ for $t,t' \in [\tau_0,\tau_1]$, then $\sigma(t,t') = \sigma^2(\max(t,t'))$, where

$$\sigma^2(t) = -2\sum_{i=1}^{2}\int_0^{T_c(t)} \xi(t,u)^2 \frac{dS(u)}{S^2(u)b_i(t,u)},$$

which is non-increasing and can be consistently estimated by $\hat{V}(t)$, where $\hat{V}(t)$ is defined in (4.3) or the corresponding pooled version variance estimator. In either case, we can show that $\sup_{\tau_0 \leq t \leq \tau_1} |n\hat{V}(t) - \sigma^2(t)| \to 0$ in probability.

From Theorem 1, we have $\{n^{1/2}\hat{W}(t)/\sigma^2(t), \tau_0 \leq t \leq \tau_1\}$ converges to a Gaussian process with independent increments since

$$\mathrm{E}\{n\hat{W}(t)\hat{W}(t')/(\sigma^2(t)\sigma^2(t'))\} \to 1/\sigma^2(\min(t,t')),$$

as $n \to \infty$. The corollary follows by noting that $n^{-1/2}\hat{W}(t)/\hat{V}(t) = n^{1/2}\hat{W}(t)/\sigma^2(t)(\sigma^2(t)/n\hat{V}(t))$ and using the functional form of Slutsky's lemma or the functional delta method. See Anderson et al. (1992) Theorem II.8.1. □

References

Anderson, P. K., Borgan, O., Gill, R. D. & Keiding, N. (1992). *Statistical Models Based on Counting Procsses.* New York : Springer-Verlag.

Fleming, T. R. & Harrington, D. P. (1991). *Counting Processes & Survival Analysis.* New York : Wiley.

Gill, R. (1980). *Censoring and Stochastic Integrals.* Math. Centre Tract **124**. Methematisch Centrum, Amsterdam.

Gu, M. & Lai, T. L. (1991). Weak convergence of time-sequential censored rank statistics with applications to sequential testing in clinical trials. *Ann. Statist.* **19**, 1403-33.