PROCEDURES FOR SERIAL TESTING IN CENSORED SURVIVAL DATA

D.P. Harrington

Department of Applied Mathematics and Computer Science
University of Virginia, Charlottesville, Virginia

T.R. Fleming and S.J. Green

Department of Medical Statistics and Epidemiology
Mayo Clinic, Rochester, Minnesota

1. Introduction

Prospective studies, such as those carried out in many cancer centers throughout the world, need to be carefully monitored and subjected to interim analyses to satisfy important ethical considerations. Typically, the therapeutic efficacy and resulting survival distribution for an experimental treatment regimen are compared to the efficacy and survival obtained from a currently accepted standard regimen. These studies often give rise to the dual need to terminate as soon as possible any trial in which it is sufficiently clear either that (1) the experimental treatment yields better results than the standard treatment or (2) the data strongly contradict the hypothesis of some minimally acceptable treatment difference. In this paper, we examine the problem of constructing closed sequential experimental designs allowing for hypothesis tests at multiple points in time when the data gathered are censored failure time data. The tests we study are useful for examining various forms of dependence of an underlying survival function $S(x)$ on a random scalar covariate Z.

2. Model and Notation

In this manuscript we will adopt the excellent notation proposed by Tsiatis (this volume) for this problem. In particular, suppose that in a prospective study the following variables are associated with the i^{th} study subject in a sample of n such independent and identically distributed subjects:

Y_i: the entry time (measured from the beginning of the study)

X_i: the time from study entry to a specified endpoint

W_i: the time from study entry until that subject is lost to follow-up

Z_i: random scalar valued covariate.

Since X,Y and W generally are stochastically dependent on Z (which may, for instance, denote sample membership in a two sample comparative study), the following notation will be used to denote the conditional distributions:

$$H(x|z) = P(Y \leq x|Z = z),$$
$$S(x|z) = P(X > x|Z = z),$$
$$\overline{G}(x|z) = P(W > x|Z = z).$$

We shall assume throughout that $P(Y \leq y,\ X > x,\ W > w|Z = z) = H(y|z)\ S(x|z)\ G(w|z)$ for all (y,x,w). Assume $S(x|z)$ is continuous. Let $\lambda(x|z) \equiv \frac{d}{dx} \ell n\ S(x|z)$.

If data of this sort are analyzed at calendar time t, that is, t units of time after the beginning of the study, the available data for subject i would include $\{X_i(t),\ \Delta_i(t),\ Z_i I\{Y_i \leq t\}\}$ where

$$X_i(t) \equiv \max\{\min(X_i,\ t - Y_i,\ W_i),\ 0\}$$

and

$$\Delta_i(t) \equiv I\{X_i \leq \min(t - Y_i,\ W_i)\} \quad .$$

Here I{E} is the indicator random variable for the event E. In many cases, these data are used to test hypotheses about the dependence of $S(x|z)$ on the covariate values z. In the next section we will review some hypotheses of interest and appropriate test statistics when one is conducting only a single test.

3. A Class of Single Stage Rank Statistics

3.1 Tests of H_0: $S(x|z) = S(x)$. The G^ρ Family

The operating characteristics of nonparametric procedures used in survival theory are most clearly understood when hypotheses are tested only once. Thus, in this section, we will temporarily assume that the data will be analyzed only at calendar time t. Without loss of generality then, we can assume in Section 3 that $H(t|z) \equiv 1$.

The most common hypothesis in this situation is $H_0{:}S(x|z) = S(x)$, $0 \leq x \leq t$, where S(x) is unspecified. The form of the statistic used here of course depends on the way in which possible covariate dependence is modeled in the survivor function. If for the scalar covariate Z one assumes $S(x|z) = \{S_0(x)\}^{\exp(z\beta)}$, then the partial likelihood score statistic for testing the equivalent hypothesis $H_0{:}\beta = 0$ yields the logrank test (Mantel 1966, Cox 1972, Peto and Peto 1972). In our current notation, this statistic is

$$\sum_{i=1}^{n} \Delta_i(t) \left[Z_i - \frac{\sum_{\ell=1}^{n} Z_\ell \, I\{X_\ell(t) \geq X_i(t)\}}{\sum_{\ell=1}^{n} I\{X_\ell(t) \geq X_i(t)\}} \right].$$

A number of authors (Tarone and Ware, 1977; Prentice and Marek, 1979; and Harrington and Fleming, 1981) have proposed generalizing the above test of H_0 by incorporating weights into the terms in the above sum, yielding statistics of the form

$$(1) \qquad S_n(t) = \sum_{i=1}^{n} \hat{Q}\{t, X_i(t)\} \, \Delta_i(t) \left[Z_i - \frac{\sum_{\ell=1}^{n} Z_\ell \, I\{X_\ell(t) \geq X_i(t)\}}{\sum_{\ell=1}^{n} I\{X_\ell(t) \geq X_i(t)\}} \right].$$

The changes induced on the operating characteristics of the logrank test are clearly understood when $\hat{Q}(t,x) = \{\hat{S}(t,x)\}^{\rho}$, $\rho \geq 0$, where $\hat{S}(t,x)$ is the value at survival time $x \leq t$ of the left continuous version of the Kaplan-Meier product limit estimator computed from the pooled sample at calendar time t. This family of tests has been called the G^{ρ} family, and was proposed and studied for the k-sample problem in Harrington and Fleming (1981). Of course, when $\rho = 0$ the test is the logrank, and when $\rho = 1$, the test is essentially equivalent to the generalized Wilcoxon statistic proposed by Peto and Peto (1972) and by Prentice (1978).

For the two sample problem, where Z is 0 or 1, the following theorem indicates the types of departures against which each G^{ρ} test procedure is fully efficient. The proof relies on Corollary 5.3.1 in Gill (1980) and is given in detail in Harrington and Fleming (1981).

THEOREM 1:

Let $S_j(x) \equiv S(x|Z=j)$ and $\lambda_j(x) \equiv -\frac{d}{dx} \ell n\, S_j(x)$ for $j = 0,1$. Fix $\rho \geq 0$. The test based on the statistic G^{ρ} is fully efficient for testing $H_0: \beta = 0$ against $H_A: \beta \neq 0$ for the Lehmann (1953) family of alternatives

(2) $$S_1(x) = S_0(x)\ [\{S_0(x)\}^{\rho} + [1 - \{S_0(x)\}^{\rho}]e^{\beta}]^{-1/\rho},\ 0 \leq x \leq t$$

or, equivalently,

(3) $$\lambda_1(x) = \lambda_0(x)\ e^{\beta}\ [\{S_0(x)\}^{\rho} + [1 - \{S_0(x)\}^{\rho}]e^{\beta}]^{-1},\ 0 \leq x \leq t$$

if and only if

$$H(t-x|Z=1)\ \bar{G}(x|Z=1) = H(t-x|Z=0)\ \bar{G}(x|Z=0),\ x \leq t \quad .$$

Interestingly, for $\rho = 0$ (resp. $\rho = 1$) we simply recover the result that the logrank test (resp. Peto and Peto - Wilcoxon test) is fully efficient for time transformed location alternatives for the extreme value (resp. logistic) distribution.

If Z is an arbitrary scalar covariate, the relationships (2) and (3) can be recast, for $\rho > 0$, as

$$S(x|z) = S_0(x) \; [\{S_0(x)\}^\rho + [1 - \{S_0(x)\}^\rho] e^{\beta z}]^{-1/\rho} \quad , \tag{4}$$

$$\lambda(x|z) = \lambda_0(x) \; e^{\beta z} \; [\{S_0(x)\}^\rho + [1 - \{S_0(x)\}^\rho] e^{\beta z}]^{-1} \quad . \tag{5}$$

The G^ρ tests discussed above are applicable to testing $H_0 : \beta = 0$ in this setting as well.

One may often be interested in testing $H_0 : \beta = \beta_0$, β_0 not necessarily zero. This situation may arise, for instance, in cancer clinical trials when a more toxic experimental treatment is being tested against a standard treatment, and one wishes to assess whether data gathered contain significant evidence against a minimally acceptable treatment difference, say a 25% decrease in the underlying hazard. In the next two sub-sections, we will examine statistics $\tilde{S}_n(t)$ appropriate for testing the more general hypothesis $H_0 : \beta = \beta_0$, at calendar time t. We will begin by considering the specific case of testing $H_0 : \beta = \beta_0$ under the proportional hazards model, which warrants special consideration due to its wide applicability. This model, of course, is given in (5) when $\rho = 0$.

3.2 Tests of $H_0 : \beta = \beta_0$ under Proportional Hazards

The proper form for $\tilde{S}_n(t)$ under the model $\lambda(x|z) = \lambda_0(x) \; e^{\beta z}$ can be seen from both a heuristic and a formal point of view. Tsiatis (1981 and this volume) has pointed out that expression (1) for $S_n(t)$, useful in testing

$H_0: \lambda(x|z) = \lambda_0(x)$, is equal to

(6)
$$S_n(t) = \sum_{i=1}^{n} \int_0^t \hat{Q}(t,x) \left[Z_i - \frac{\sum_{\ell=1}^{n} Z_\ell I\{X_\ell(t) \geq x\}}{\sum_{\ell=1}^{n} I\{X_\ell(t) \geq x\}} \right] [dN_i(t,x) - I\{X_i(t) \geq x\}\, \lambda_0(x)dx],$$

where $N_i(t,x) = I\{X_i(t) \leq x,\ \Delta_i(t) = 1\}$. For testing the general hypothesis $H_0: \beta = \beta_0$ under the model $\lambda(x|z) = e^{\beta z}\ \lambda_0(x)$, one might set $Q(t,x) \equiv 1$ and then replace $\lambda_0(x)$ in (6) with $e^{z_i \beta}\ \lambda_0(x)$, where, in turn, $\lambda_0(x)$ must be estimated from the data. Under $H_0: \beta = \beta_0$, and at calendar time t, a natural estimator for $\Lambda_0(x) \equiv \int_0^x \lambda_0(u)du$ is

$$\hat{\Lambda}_0(t,x) = \sum_{i=1}^{n} \int_0^x \left[\sum_{j=1}^{n} e^{\beta_0 z_j}\ I\{X_j(t) \geq u\} \right]^{-1} dN_i(t,u) \quad .$$

The following lemma, which has a simple algebraic proof, demonstrates that this heuristic approach yields a statistic which is easy to calculate.

LEMMA 1:

$$\tilde{S}_n(t) \equiv \sum_{i=1}^{n} \int_0^t \left[Z_i - \frac{\sum_{\ell=1}^{n} Z_\ell I\{X_\ell(t) \geq x\}}{\sum_{\ell=1}^{n} I\{X_\ell(t) \geq x\}} \right] \left[dN_i(t,x) - I\{X_i(t) \geq x\} e^{\beta_0 Z_i}\ d\hat{\Lambda}_0(t,x) \right] ,$$

(7)
$$= \sum_{i=1}^{n} \int_0^t \left[Z_i - \frac{\sum_{\ell=1}^{n} Z_\ell\ e^{\beta_0 Z_\ell}\ I\{X_\ell(t) \geq x\}}{\sum_{\ell=1}^{n} e^{\beta_0 Z_\ell}\ I\{X_\ell(t) \geq x\}} \right] dN_i(t,x) \quad .$$

Although the above approach is only intuitively reasonable, the following lemma indicates that the statistic just derived to test $H_0:\beta = \beta_0$, under the model $\lambda(x|z) = \lambda_0(x)\ e^{\beta z}$, has a more formal justification.

LEMMA 2:

Let $L(\beta,t)$ be the Cox (1975) partial likelihood constructed at calendar time t from the proportional hazards model $\lambda(x|z) = \lambda_0(x)\ e^{z\beta}$. That is,

$$L(\beta,t) = \prod_{i=1}^{n} \left[\frac{e^{Z_i\beta}}{\sum_{\ell=1}^{n} e^{Z_\ell\beta}\ I\{X_\ell(t) \geq X_i(t)\}} \right]^{\Delta_i(t)} .$$

Then $\tilde{S}_n(t)$ in expression (7) is simply the score test statistic

$$\frac{\partial}{\partial\beta}\ \ell n\ L(\beta,t)\Big|_{\beta=\beta_0} .$$

In this sub-section, we have discussed testing the hypothesis that

$$H_0:\ \lambda(x|z) = \lambda_0(x)\ e^{\beta_0 z} . \tag{8}$$

It should be observed that the alternative of interest to H_0 may not always satisfy the proportional hazards assumption, i.e., may not be specified by (8) with β_0 replaced by β. More generally, if data analysis occurs at calendar time t, one may wish to test $H_0:\alpha = 0$ vs $H_A:\alpha \neq 0$ where

$$\lambda(x|z) = \lambda_0(x)\ \exp\{z(\beta_0 + \alpha Q(t,x))\} \tag{9}$$

for some function $Q(t,x)$ continuous in x. We assume that $Q(t,x)$ is independent of α but may be a function of $\lambda_0(x)$. Clearly (9) reduces to (8) when $\alpha = 0$.

With analysis occurring at calendar time t, let $\hat{Q}(t,x)$ be a consistent estimator of $Q(t,x)$ under H_0. Forming Cox's partial likelihood $L(\beta_0,\alpha,t)$ based upon the relationship (9), the resulting expression can be used to formulate a score-type test statistic $\frac{\partial}{\partial\alpha} \ell n\, L(\beta_0,\alpha,t)\Big|_{\alpha=0}$. Replacing $Q(t,x)$ by $\hat{Q}(t,x)$ yields

$$(10) \qquad \tilde{S}_n(t) = \sum_{i=1}^{n} \int_0^t \hat{Q}(t,x) \left[Z_i - \frac{\sum_{\ell=1}^{n} Z_\ell\, e^{\beta_0 Z_\ell}\, I\{X_\ell(t) \geq x\}}{\sum_{\ell=1}^{n} e^{\beta_0 Z_\ell}\, I\{X_\ell(t) \geq x\}} \right] dN_i(t,x) \quad .$$

We propose $\tilde{S}_n(t)$ as defined in (10) be employed to test $H_0{:}\alpha = 0$ vs $H_A{:}\alpha \neq 0$ for the hazard relationship specified by (9). Its distribution under H_0 will be derived in §4. Observe that expression (10) reduces to (7) when $\hat{Q}(t,x) \equiv 1 \equiv Q(t,x)$ and it reduces to (1) when $\beta_0 = 0$.

Setting $\hat{Q}(t,x) = \{\hat{S}(t,x)\}^\rho$ in expression (10) would yield a generalized G^ρ statistic for testing $H_0{:}\alpha = 0$ in relationship (9) with

$$(11) \qquad Q(t,x) = \exp\left\{ -\rho \int_0^x \left[\frac{E\{e^{Z\beta_0}\, \bar{G}(u|Z)\, H(t-u|Z)\, \{S_0(u)\}^{\exp(Z\beta_0)}\}}{E\{\bar{G}(u|Z)\, H(t-u|Z)\, \{S_0(u)\}^{\exp(Z\beta_0)}\}} \right] \lambda_0(u)du \right\} .$$

When $H(u|z) = H(u)$ and $\bar{G}(u|z) = \bar{G}(u)$, equation (11) reduces to $Q(t,x) = [E\{(S_0(x))^{\exp(Z\beta_0)}\}]^\rho$. If instead one assumes $\beta_0 = 0$, then (11) reduces to $Q(t,x) = \{S_0(x)\}^\rho$. It follows that the G^ρ test procedure, specified by (10) when $\hat{Q}(t,x) = \{\hat{S}(t,x)\}^\rho$ and $\beta_0 = 0$, has been derived as being appropriate for testing $H_0{:}\alpha = 0$ in hazard relationship (9) with $\beta_0 = 0$ and $Q(t,x) = \{S_0(x)\}^\rho$ or, as noted earlier, for testing $H_0{:}\beta = 0$ in hazard relationship (5). As would be expected, these two hazard relationships are very similar.

3.3 Tests Under More General Models

We observed in the previous sub-section that the statistic in (10) arises as a score-type test statistic appropriate for testing $H_0: \alpha = 0$ under the model $\lambda(x|z) = \lambda_0(x)\ \exp\{z(\beta_0 + \alpha Q(t,x))\}$.

More generally, one could be interested in a test of the hypothesis $H_0: \alpha = 0$ vs $H_A: \alpha \neq 0$ where

$$(12) \qquad \lambda(x|z) = \lambda_0(x)\ \exp\{z(Q_2(t,x) + \alpha Q_1(t,x))\}$$

for continuous functions $Q_i(t,x)$ which are independent of α but may be functions of $\lambda_0(x)$.

For $i = 1,2$, let $\hat{Q}_i(t,x)$ be a consistent estimator under H_0 of $Q_i(t,x)$, where t continues to represent the calendar time of analysis. Forming Cox's partial likelihood based upon the relationship (12), one can again obtain a score-type statistic to test $H_0: \alpha = 0$. Replacing $Q_i(t,x)$ by $\hat{Q}_i(t,x)$ yields

$$(13) \qquad \tilde{S}_n(t) = \sum_{i=1}^{n} \int_0^t \hat{Q}_1(t,x) \left[Z_i - \frac{\sum_{\ell=1}^{n} Z_\ell I\{X_\ell(t) \geq x\}\ e^{Z_\ell \hat{Q}_2(t,x)}}{\sum_{\ell=1}^{n} I\{X_\ell(t) \geq x\}\ e^{Z_\ell \hat{Q}_2(t,x)}} \right] dN_i(t,x) \quad .$$

Motivated by the frequent need, described in the Introduction, to perform interim analyses of the data, we will examine in the next section the distributions of the statistics which we have just discussed and how they can be employed when performing repeated significance testing in censored survival data. Unfortunately, it appears that the techniques to be used are only applicable when $\hat{Q}_2(t,x)$ in (13) is non-random. As a result, we will restrict our attention hereafter to statistics of the form appearing in expression (10).

4. Repeated Significance Testing in Censored Survival Data

4.1 Critical Regions for Repeated Tests

The structure of our repeated testing critical regions will be essentially that proposed by Slud and Wei (1982).

1. We will assume one will perform up to K tests based on up to n individuals. The j^{th} test will be performed at time t_j using $\tilde{S}_n(t_j)$ as defined in (10). $\tilde{S}_n(t_j)$ is a mean zero statistic under H_0 specified by (8).

2. For a fixed overall significance level α, we will choose $\pi_1, \pi_2, \ldots, \pi_K$ such that $0 < \pi_j$ and $\sum_{j=1}^{K} \pi_j = \alpha$.

3. Critical values $\{a_1, \ldots, a_K\}$ will be recursively determined, i.e., having chosen $a_1, \ldots, a_{j-1}$, we will choose a_j so that

$$P\{\tilde{S}_n(t_1) < a_1, \ldots, \tilde{S}_n(t_{j-1}) < a_{j-1},\ \tilde{S}_n(t_j) \geq a_j | H_0\} = \pi_j .$$

4. We will reject H_0 if and only if one observes the event

$$R = \bigcup_{j=1}^{K} \{\tilde{S}_n(t_1) < a_1, \ldots, \tilde{S}_n(t_{j-1}) < a_{j-1},\ \tilde{S}_n(t_j) \geq a_j\},$$

   which is the nuion of K mutually exclusive components. Thus $P(R|H_0) = \alpha$.

To carry out this approach one needs to determine the joint distribution of $n^{-1/2}\{\tilde{S}_n(t_1),\ \tilde{S}_n(t_2), \ldots, \tilde{S}_n(t_j)\}$ for any $t_1 \leq t_2 \leq \cdots \leq t_j$, $j = 1, \ldots, K$. We will indicate how the asymptotic joint distribution is obtained in the special case when Z_i assumes finitely many levels. Specifically, we will assume $P(Z_i = c_k) = p_k$ for $k = 1, \ldots, m$, where m is finite and $p_1 + p_2 + \cdots + p_m = 1$.

The following, which is an alternative form for $\tilde{S}_n(t)$ defined in (10) and which is the direct analogue of (6) for $\beta_0 \neq 0$, is easy to establish and will be useful in what follows.

LEMMA 3: $\tilde{S}_n(t) =$

$$\sum_{i=1}^{n}\int_0^t \hat{Q}(t,x)\left[Z_i - \frac{\sum_{\ell=1}^{n} Z_\ell e^{\beta_0 Z_\ell} I\{X_\ell(t)\geq x\}}{\sum_{\ell=1}^{n} e^{\beta_0 Z_\ell} I\{X_\ell(t)\geq x\}}\right] d\left[N_i(t,x) - \int_0^x e^{\beta_0 Z_i} I\{X_i(t)\geq u\}\ \lambda_0(u)du\right].$$

We note that $\tilde{S}_n(t)$ is of the form $\sum_{i=1}^{n}\int_0^t h_i(t,x)\ dM_i(t,x)$. For fixed t, $M_i(t,x)$ is a square integrable martingale with respect to the basis $\mathcal{F}_t^i = \{\mathcal{F}_{t,x}^i; 0\leq x\leq t\}$, where $\mathcal{F}_{t,x}^i$ is the sigma sub-field generated by the random variables $\{I\{Y_i\leq t\},\ Z_i I\{Y_i\leq t\},\ Y_i I\{Y_i\leq t\},\ I\{X_i\leq \min(u, t-Y_i, W_i)\},\ I\{W_i\leq \min(u, t-Y_i, X_i)\}: 0\leq u\leq x\}$. The martingale structure will be important in the following two lemmas.

LEMMA 4:

For each n and each t, let $\bar{M}_{k,n}(t,x) \equiv \sum_{i=1}^{n} M_i(t,x)\ I\{Z_i = c_k\};\ k=1,\dots,m;$ and let $\mathcal{B}_{n,t}$ be a basis containing $\{\mathcal{F}_t^i : i=1,\dots,n\}$. Let $\pi(t,x|Z) \equiv H(t-x|Z)\ S(x|Z)\ \bar{G}(x|Z)$, where we assume $\pi(t,x|Z) > 0$ for $0 < x < t$. Define

$$\hat{\mu}(t,x) = \sum_{i=1}^{n} e^{\beta_0 Z_i} Z_i I\{X_i(t)\geq x\} \Big/ \sum_{i=1}^{n} e^{\beta_0 Z_i} I\{X_i(t)\geq x\}$$

and

$$\mu(t,x) = E\left\{e^{\beta_0 Z}\ Z\,\pi(t,x|Z)\right\} \Big/ E\left\{e^{\beta_0 Z}\ \pi(t,x|Z)\right\},$$

where $\hat{\mu}(t,x) \equiv 0$ if $\sum_{i=1}^{n} I\{X_i(t)\geq x\} = 0$, and $\mu(t,t) \equiv 0$ if $\pi(t,t|z) = 0$. Define

$$\hat{\hat{S}}_n(t) = \sum_{i=1}^{n}\int_0^t Q(t,x)\ \{Z_i - \mu(t,x)\}\ d\{N_i(t,x) - \int_0^x e^{\beta_0 Z_i}\lambda_0(u)\ I\{X_i(t)\geq u\}du\}.$$

Assume that

1) $\sup_{0 \leq x \leq \tau} |Q(t,x) - \hat{Q}(t,x)| \overset{P}{\to} 0$ for all $\tau < t$, where $\overset{P}{\to}$ means convergence in probability,

2) $\hat{Q}$ is bounded over $[0,t]$, left continuous with right hand limits, and adapted to $\mathcal{B}_{n,t}$.

Then, under H_0,

$$n^{-1/2} \{\hat{\tilde{S}}_n(t) - \tilde{S}_n(t)\} \overset{P}{\to} 0 \quad .$$

PROOF:

$$n^{-1/2} \{\tilde{S}_n(t) - \hat{\tilde{S}}_n(t)\}$$

$$\doteq \sum_{k=1}^{m} n^{-1/2} \int_0^t \{\hat{Q}(t,x) - Q(t,x)\} \{c_k - \hat{\mu}(t,x)\} \, d\bar{M}_{k,n}(t,x)$$

$$+ \sum_{k=1}^{m} n^{-1/2} \int_0^t Q(t,x) \{\mu(t,x) - \hat{\mu}(t,x)\} \, d\bar{M}_{k,n}(t,x)$$

$$\equiv \sum_{k=1}^{m} n^{-1/2} E^1_{k,n}(t) + \sum_{k=1}^{m} n^{-1/2} E^2_{k,n}(t) \quad .$$

We will establish that the above expression converges to zero in probability under H_0 by appealing to the central limit theorem (Gill 1980, §2.4) for stochastic integrals with respect to counting process martingales.

Fix ε and t. Since $\bar{M}_{k,n}$ is a square integrable martingale with respect to $\mathcal{B}_{n,t}$ (Fleming and Harrington, 1981) and $(\hat{Q} - Q)$ is bounded and predictable, $n^{-1/2} E^1_{k,n}$ is a square integrable martingale with zero expectation and predictable covariation process

$$\int \{\hat{Q}(t,x) - Q(t,x)\}^2 \{c_k - \hat{\mu}(t,x)\}^2 \, d < \bar{M}_{k,n}(t,x), \bar{M}_{k,n}(t,x) > =$$

$$\int \{\hat{Q}(t,x) - Q(t,x)\}^2 \{c_k - \hat{\mu}(t,x)\}^2 \, d \int_0^x e^{\beta_0 c_k} \lambda_0(u) \, n^{-1} \sum_{i:Z_i = c_k} I\{X_i(t) \geq u\} du \ .$$

Since sup $|\hat{Q}-Q|$ is bounded, τ_ε can be chosen such that

$$\int_{\tau_\varepsilon}^t e^{\beta_0 c_k} \lambda_0(u) du \cdot \sup_{0 \leq x \leq t} [\{Q(t,x) - \hat{Q}(t,x)\}^2 \{c_k - \hat{\mu}(t,x)\}^2] < \varepsilon/2 \ .$$

Further, since $\sup_{0 \leq x \leq \tau_\varepsilon} |\hat{Q}(t,x) - Q(t,x)| \overset{P}{\to} 0$, n_1 can be chosen such that for $n \geq n_1$,

$$P(e^{\beta_0 c_k} \Lambda_0(t) \sup_{0 \leq x \leq \tau_\varepsilon} [\{\hat{Q}(t,x) - Q(t,x)\}^2 \{c_k - \hat{\mu}(t,x)\}^2] < \varepsilon/2) > 1-\varepsilon \ .$$

Then $P(< n^{-1/2} E^1_{k,n}(t), n^{-1/2} E^1_{k,n}(t) > < \varepsilon) > 1-\varepsilon$ for $n > n_1$. The martingale central limit theorem (Gill, Theorem 2.4.1) then implies that $n^{-1/2} E^1_{k,n}(t) \overset{P}{\to} 0$. We can show in a similar fashion that $n^{-1/2} E^2_{k,n}(t) \overset{P}{\to} 0$, and thus $n^{-1/2} \{\hat{\tilde{S}}_n(t) - \tilde{S}_n(t)\} \overset{P}{\to} 0$.

By Lemma 4 and an application of the Cramér-Wold device it is now sufficient to find the asymptotic joint distribtuion of $n^{-1/2}\{\hat{\tilde{S}}_n(t_1), \hat{\tilde{S}}_n(t_2), \ldots, \hat{\tilde{S}}_n(t_j)\}$. If we first define the stochastic processes $M_i(x)$, $i = 1,2,\ldots,n$

$$M_i(x) = \Delta_i(x) - \int_0^x I\{Y_i \leq u \leq Y_i + \min(X_i, W_i)\} \lambda(u - Y_i | Z_i) du \ ,$$

it then follows by a simple time transformation that

$$\hat{\tilde{S}}_n(t) = \sum_{i=1}^n \int_0^t Q(t, x - Y_i) \{Z_i - \mu(t, x - Y_i)\} \, dM_i(x) \equiv \sum_{i=1}^n A_i(t) \ ,$$

a sum of independent and identically distributed random variables. That $n^{-1/2}\{\hat{\tilde{S}}_n(t_1), \hat{\tilde{S}}_n(t_2),\ldots,\hat{\tilde{S}}_n(t_j)\}$ converges to a multivariate normal distribution follows from the central limit theorem. As with $M_i(t,x)$, $M_i(x)$ is a square integrable martingale, but with respect to the basis $\{\tilde{F}_x^i\ ;\ 0 \leq x \leq \infty\}$ where $\tilde{F}_x^i$ is the sigma sub-field generated by the random variables $\{I\{Y_i \leq u\}$, $Z_i I\{Y_i \leq u\}$, $I\{X_i \leq \min(u - Y_i, W_i)\}$, $I\{W_i \leq \min(u - Y_i, X_i)\}$; $0 \leq u \leq x\}$. Asymptotic moments given in the lemma below follow from results of Meyer (1976) for stochastic integrals with respect to martingales.

LEMMA 5:

Assume conditions given in Lemma 4 and let $0 \leq t \leq t' \leq t_j$. Then

$$E\ \hat{\tilde{S}}_n(t) = E\ A_i(t) = 0, \text{ and } \operatorname{cov}\{n^{-1/2}\ \hat{\tilde{S}}_n(t),\ n^{-1/2}\ \hat{\tilde{S}}_n(t')\} =$$

$$\operatorname{cov}\{A_i(t),A_i(t')\} = \int_0^t Q(t,x)\ Q(t',x)\, E[\{Z-\mu(t,x)\}^2\ e^{\beta_0 Z}\ \lambda_0(x)\ \pi(t,x|Z)]\, dx \quad .$$

When $\beta_0 = 0$, one obtains results presented by Tsiatis (this volume). However, application of martingale stochastic integral results simplifies the covariance calculation he made for this special case.

Since $\operatorname{cov}\{A_i(t),A_i(t')\}$ depends upon t' only through $Q(t',x)$, it follows that $\{n^{-1/2}\ \tilde{S}_n(t): t \geq 0\}$ converges to a limit process having independent increments whenever $Q(t,x)$ is independent of t. Such is the case for the generalized G^ρ family when either $S(x|Z)$ or $H(x|Z)$ is independent of Z. This can be seen by observing $Q(t,x) = \{S^*(t,x)\}^\rho$ when $\hat{Q}(t,x) = \{\hat{S}(t,x)\}^\rho$, where

$$S^*(t,x) \equiv \exp[- \int_0^x \left\{ \sum_{k=1}^m p_k\ \pi(t,u|Z=c_k)\,\lambda(u|Z=c_k) \Big/ \sum_{k=1}^m p_k\ \pi(t,u|Z=c_k) \right\} du] \quad .$$

Although the derivation of the asymptotic distribution of $n^{-1/2}\{\tilde{S}_n(t_1),\ldots, \tilde{S}_n(t_k)\}$ assumes independent, identically distributed covariates Z_i, it can be extended to studies in which the covariate values are balanced through forced randomization. In particular, assume that in a study of n subjects, n_k of those will have covariate value c_k, $k=1,\dots,m$ and that $n_k \equiv np_k$. The term $\sum_{i=1}^{n} A_i(t)$ may then be viewed as a sum of independent, but non-identically distributed terms. One may still show that $n^{-1/2}\{\tilde{S}_n(t_1),\dots,\tilde{S}_n(t_k)\}$ is asymptotically multivariate normal, with zero mean and with

$$\lim_{n\to\infty} \mathrm{Cov}\{n^{-1/2}\,\tilde{S}(t),\ n^{-1/2}\,\tilde{S}_n(t')\}$$

$$= \sum_{k=1}^{m} p_k \int_0^t Q(t,x)\ Q(t',x)\ \{c_k - \mu(t,x)\}^2\ e^{\beta_0 c_k}\ \lambda_0(x)\ \pi(t,x|c_k)dx\ ,$$

where $t<t'$. Note that this expression agrees with the covariance formula in Lemma 5.

Recall, by Lemma 5, that $\mathrm{cov}\{n^{-1/2}\,\tilde{S}_n(t),\ n^{-1/2}\,\tilde{S}_n(t')\}$ converges to

$$\sigma(t,t') \equiv \int_0^t Q(t,x)\ Q(t',x)\ E[\{Z-\mu(t,x)\}^2\ e^{\beta_0 Z}\ \pi(t,x|Z)]d\Lambda_0(x)\quad .$$

If $\hat{\sigma}(t,t')$ denotes a consistent estimator of $\sigma(t,t')$, then the actual test statistics employed at t and t' are $n^{-1/2}\{\hat{\sigma}(t,t)\}^{-1/2}\ \tilde{S}_n(t)$ and $n^{-1/2}\{\hat{\sigma}(t',t')\}^{-1/2}\ \tilde{S}_n(t')$, which are positively correlated. One consistent estimator is given by $\hat{\sigma}(t,t') \equiv$

$$\int_0^t \hat{Q}(t,x)\ \hat{Q}(t',x)\left[\frac{1}{n}\sum_{j=1}^{n}\{Z_j-\hat{\mu}(t,x)\}^2\ e^{\beta_0 Z_j}\ I\{X_j(t)\geq x\}\right] d\hat{\Lambda}_0(t,x) =$$

$$\frac{1}{n}\sum_{i\in R(t)}\left[\Delta_i(t)\ \hat{Q}(t,X_i(t))\ \hat{Q}(t',X_i(t))\right.$$

$$\left.*\left\{\sum_{j\in R(t,X_i(t))} e^{\beta_0 Z_j}\left\{Z_j - \frac{\sum_{\ell\in R(t,X_i(t))} Z_\ell\, e^{\beta_0 Z_\ell}}{\sum_{\ell\in R(t,X_i(t))} e^{\beta_0 Z_\ell}}\right\}^2 \Big/ \sum_{j\in R(t,X_i(t))} e^{\beta_0 Z_j}\right\}\right],$$

where R(t) denotes the set of indices $\{j=1,\ldots,n\}$ such that $Y_j \leq t$, and R(t,x) denotes the set such that $\{X_j(t) \geq x\}$. That $n^{-1/2}\{\hat{\sigma}(t,t)\}^{-1/2}\ \tilde{S}_n(t)$ does not require knowledge of n is important for applications.

4.2 Selection of π_i, $i=1,\ldots,K$.

Several different approaches exist for choosing π_i; $i=1,\ldots,K$. One approach of particular interest would be to select $\pi_1, \pi_2,\ldots,\pi_{K-1}$ small, with $\pi_K \approx \alpha$, where α is the size of the procedure. Procedures discussed by Haybittle (1971) and O'Brien and Fleming (1979) are conceptually related to this. The resulting serial testing procedure would then allow early testing to detect substantial departures from H_0, satisfying ethical considerations. In addition, the critical value for the statistic employed at the K^{th} and final stage of the procedure would be nearly identical to the critical value which is appropriate when a single procedure is based upon that statistic. Such a sequential procedure would have power nearly identical to that of the corresponding single stage procedure. On the other hand, the serial testing procedure resulting from repeated use of a statistic will have operating characteristics considerably different from those of the corresponding single stage procedure if one chooses π_i, $i=1,\ldots,K$, such that $\pi_K \ll \alpha$. In heavily censored data, the sequential procedure will give much ligher weight to "later" occurring departures from H_0 than the corresponding single stage procedure. A careful theoretical consideration of the power and efficiencies of these types of serial testing procedures seems to be difficult.

ACKNOWLEDGEMENT

The research of the first author was supported by Eagles Grant 23 and NSF Grant MSC-80-02887. The research of the second and third authors was supported by NIH Grant CA-24089.

REFERENCES

Cox, D.R. (1972). Regression models and life tables (with discussion). Journal of the Royal Statistical Society B 34, 187-220.

Cox, D.R. (1975). Partial likelihood. Biometrika 62, 269-276.

Fleming, T.R. and Harrington, D.P. (1981). A class of hypothesis tests for one and two sample censored survival data. Communications in Statistics A 10, 763-794.

Gill, R.D. (1980). Censoring and Stochastic Integrals. Mathematical Centre Tracts 124, Mathematische Centre, Amsterdam.

Haybittle, J.L. (1971). Repeated assessment of results in clinical trials of cancer treatment. British Journal of Radiology 44, 793-797.

Harrington, D.P. and Fleming, T.R. (1981). A class of rank test procedures for censored survival data. Technical Report Series No-12, Section of Medical Research Statistics, Mayo Clinic. To appear in Biometrika, 1983.

Lehmann, E.L. (1953). The power of rank tests. Annals of Mathematical Statistics 24, 23-43.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemotherapy Reports 50, 163-170.

Meyer, P.A. (1976). Un cours sur les integrales stochastiques, p. 245-400. In: Seminaire de Probabilities X, Lecture Notes in Mathematics 511, Springer-Verlag, Berlin.

O'Brien, P.C. and Fleming, T.R. (1979). A multiple testing procedure for clinical trials. Biometrics 35, 549-556.

Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant test procedures (with discussion). Journal of the Royal Statistical Society A 135, 185-206.

Prentice, R.L. (1978). Linear rank tests with right censored data. Biometrika 65, 167-179.

Prentice, R.L. and Marek, P. (1979). A qualitative discrepancy between censored data rank tests. Biometrics **35**, 861-867.

Slud, E.V. and Wei, L.J. (1982). Two-sample repeated significance tests based on the modified Wilcoxon statistic. Journal of the American Statistical Association. To appear.

Tarone, R.E. and Ware, J. (1977). On distribution free tests for equality of survival distributions. Biometrika **64**, 156-160.

Tsiatis, A.A. (1981a). The asymptotic joint distribution of the efficient scores test for the proportional hazards model calculated over time. Biometrika **68**, 311-315.