Statistica Sinica **9**(1999), 913-922

REPEATED SIGNIFICANCE TESTS OF A MULTI-PARAMETER IN SURVIVAL ANALYSIS

I-Shou Chang, Chao A. Hsiung and Leng-Cheng Hwang

National Central University, Academia Sinica, Taipei and Tamkang University

Abstract: This paper illustrates a martingale method of constructing repeated significance tests for a multi-dimensional parameter in survival analysis by reducing it to tests for one-dimensional parameters. This illustration is made with parametric survival data with staggered entry. A simulation study is included to indicate its numerical performance.

Key words and phrases: Multi-dimensional parameter, orthogonal estimating functions, repeated significance tests, survival analysis.

1. Introduction

The purpose of this paper is to illustrate a martingale method of constructing repeated significance tests for a multi-dimensional parameter in survival analysis by reducing it to tests for one-dimensional parameters. We will make the illustration by considering parametric survival data with staggered entry times. The parameter space is assumed to be two-dimensional for simplicity, although it can be extended to higher dimensions without difficulty.

For parametric survival data with staggered entry, we know the likelihood score process is a martingale relative to the calendar time filtration. In case the parameter space is one-dimensional, we can make a random time change to obtain a Brownian motion approximation to the likelihood score process, which paves the way to proposing repeated significance tests of the parameter. (cf. Andersen, Borgan, Gill and Keiding (1993), p.397).

The situation is different in the multi-dimensional case. Although components of the likelihood score process are martingales, they are not orthogonal. The operating characteristics of sequential tests defined in terms of these likelihood processes would be difficult to calculate. We will consider martingale transforms of these components so as to get orthogonal martingale estimating functions and then apply the strong representation theorem to obtain the standard R^2-valued Brownian motion approximation to them. With this preparation, we propose repeated significance tests and study their asymptotic properties.

We would like to remark here that, although the model of Tsiatis, Boucher and Kim (1995) involves a multi-dimensional parameter, the hypothesis they considered concerns only the value of a one-dimensional parameter. In fact, their work can be extended with the method presented in this paper.

This paper is organized as follows. Section 2 contains the theoretical developments of the repeated significance tests and Section 3 presents a numerical study. The numerical study indicates that our theory is satisfactory. We would like to note that this method may also be useful in other models when the hypothesis of interest is multi-dimensional. One example is the case of paired survival data with staggered entry (cf. Chang, Hsiung and Chuang (1997)). Another is the case of Cox regression model with simultaneous entry.

2. Repeated Significance Tests for Parametric Survival Data

Let (Y_j, Z_j, X_j, C_j) be an i.i.d. sequence of random vectors with Y_j denoting the entry time, Z_j the covariate, X_j the survival time and C_j the censoring time of the jth person in a clinical trial. Assume that Y_j and X_j are independent; conditional on Y_j and Z_j, X_j and C_j are independent. Assume the hazard function of X_j given Z_j is of the form

$$\{\lambda(\cdot, Z_j, \theta) \mid \theta \in \Theta\} \tag{2.1}$$

for some open set $\Theta \subset R^2$.

We are interested in proposing repeated significance tests for the true parameter p in the situation that the data available at calendar time t is

$$\{Y_j \wedge s, Z_j, (Y_j + X_j \wedge C_j) \wedge s, 1_{[(Y_j+X_j)\wedge s \le (Y_j+C_j)\wedge s]} \mid s \le t, j = 1, \ldots, J\}. \tag{2.2}$$

Let

$$M_j(t,\theta) = 1_{[Y_j+X_j,\infty)}(t\wedge(Y_j+C_j)) - \int_0^t \lambda(s-Y_j, Z_j, \theta) 1_{(Y_j, Y_j+X_j\wedge C_j]}(s)\, ds. \tag{2.3}$$

The assumption (2.1) implies that $M_j(t,\theta)$ is a calendar time martingale relative to the probability for θ.

In order to write down the log-likelihood for (2.2), we introdue the stochastic processes

$$N_j^t(u) = 1_{[X_j,\infty)}(u \wedge C_j \wedge (t - Y_j)^+),$$
$$H_j^t(u) = 1_{(0, X_j\wedge C_j \wedge (t-Y_j)^+]}(u).$$

We note that $N_j^t(u) - \int_0^u \lambda(s, Z_j, \theta) H_j^t(s)\, ds$ is a martingale in u for each $t > 0$. Then, according to Chang and Hsiung (1988), the log-likelihood of the data (2.2) at time t is

$$L_J(t,\theta) = \sum_{j=1}^J \int_0^t \log \lambda(u, Z_j, \theta)\, dN_j^t(u) - \sum_{j=1}^J \int_0^t \lambda(u, Z_j, \theta) H_j^t(u)\, du. \tag{2.4}$$

Assume that $\lambda(t, Z_j, \theta)$ has bounded third derivatives in θ and is bounded away from 0. Then, the likelihood score process

$$\begin{aligned} U_{J,l}(t,\theta) &\equiv \frac{\partial}{\partial\theta_l} L_J(t,\theta) \\ &= \sum_{j=1}^{J} \int_0^t \frac{\frac{\partial\lambda}{\partial\theta_l}(u-Y_j, Z_j, \theta)}{\lambda(u-Y_j, Z_j, \theta)} 1_{(Y_j,\infty)}(u)\, dM_j(u,\theta) \end{aligned} \tag{2.5}$$

is a calendar time martingale relative to the probability for θ. (cf. Chang and Hsiung (1988), Tsiatis, Boucher and Kim (1995)). Here $l = 1, 2$.

Let

$$\tilde{U}_{J,1}(t,\theta) = U_{J,1}(t,\theta),$$

$$\tilde{U}_{J,2}(t,\theta) = U_{J,2}(t,\theta) - \sum_{j=1}^{J} \int_0^t a_J(u,\theta) \frac{\frac{\partial\lambda}{\partial\theta_1}(u-Y_j, Z_j, \theta)}{\lambda(u-Y_j, Z_j, \theta)} 1_{(Y_j,\infty)}(u)\, dM_j(u,\theta). \tag{2.6}$$

Here

$$\begin{aligned} a_J(u,\theta) = &\Big(\sum_{j=1}^{J} \frac{(\frac{\partial\lambda}{\partial\theta_1}(u-Y_j, Z_j, \theta))(\frac{\partial\lambda}{\partial\theta_2}(u-Y_j, Z_j, \theta))}{\lambda(u-Y_j, Z_j, \theta)} 1_{(Y_j, Y_j+X_j\wedge C_j]}(u) \Big) \\ &\cdot \Big(\sum_{j=1}^{J} \frac{(\frac{\partial\lambda}{\partial\theta_1}(u-Y_j, Z_j, \theta))^2}{\lambda(u-Y_j, Z_j, \theta)} 1_{(Y_j, Y_j+X_j\wedge C_j]}(u) \Big)^{-1}. \end{aligned} \tag{2.7}$$

In this paper, we adopt the convention that $\frac{0}{0} = 0$. We note that a_J is well-defined because of the Schwartz inequality for a mutual variation process. (cf. Elliott (1982), p.101). Let $a(u,\theta)$ be the limit of $a_J(u,\theta)$ as J tends to infinity, whose existence is implied by the Law of Large Numbers.

It follows from (2.6) and (2.7) that, relative to the probability for θ, the mutual predictable variation process

$$\begin{aligned} &< \tilde{U}_{J,1}(\cdot,\theta), \tilde{U}_{J,2}(\cdot,\theta) >_t \\ = &\sum_{j=1}^{J} \int_0^t \frac{\frac{\partial\lambda}{\partial\theta_1}(u-Y_j, Z_j, \theta)(\frac{\partial\lambda}{\partial\theta_2}(u-Y_j, Z_j, \theta) - a_J(u,\theta)\frac{\partial\lambda}{\partial\theta_1}(u-Y_j, Z_j, \theta))}{\lambda(u-Y_j, Z_j, \theta)} \\ &\cdot 1_{(Y_j, Y_j+X_j\wedge C_j]}(u)\, du \\ = &\,0, \end{aligned}$$

which says that $\tilde{U}_{J,1}(\cdot,\theta)$ and $\tilde{U}_{J,2}(\cdot,\theta)$ are orthogonal martingales relative to the probability for θ.

Let $P^{J(\phi)}$ denote the probability for the parameter $\theta = p + \frac{\phi}{\sqrt{J}}$. Applying the Martingale Central Limit Theorem, we get

Proposition 2.1. *Under* $P^{J(\phi)}$, $\frac{1}{\sqrt{J}}\binom{\tilde U_{J,1}(t,p+\frac{\phi}{\sqrt{J}})}{\tilde U_{J,2}(t,p+\frac{\phi}{\sqrt{J}})}$ *converges to* $\binom{G_1(t)}{G_2(t)}$ *as* J *goes to infinity, where* $G_1(\cdot)$ *and* $G_2(\cdot)$ *are two continuous independent mean* 0 *Gaussian martingales with variances*

$$g_1(t) = E_p\int_0^t \frac{(\frac{\partial\lambda}{\partial\theta_1}(u-Y_1,Z_1,p))^2}{\lambda(u-Y_1,Z_1,p)} 1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du, \tag{2.8}$$

$$\begin{aligned} g_2(t) = E_p\int_0^t &\Big(\frac{(\frac{\partial\lambda}{\partial\theta_2}(u-Y_1,Z_1,p))}{\lambda(u-Y_1,Z_1,p)} - a(u,p)\frac{(\frac{\partial\lambda}{\partial\theta_1}(u-Y_1,Z_1,p))}{\lambda(u-Y_1,Z_1,p)}\Big)^2 \\ &\cdot\lambda(u-Y_1,Z_1,p)1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du, \end{aligned} \tag{2.9}$$

respectively.

Straight-forward calculations (cf. Chang and Hsiung (1988)) lead to

Proposition 2.2. *Under* $P^{J(\phi)}$, *as* J *goes to infinity,*

i) $\frac{1}{J}\frac{\partial}{\partial\theta_1}\tilde U_{J,1}(t,p)$ *converges to* $-g_1(t)$,

ii) $\frac{1}{J}\frac{\partial}{\partial\theta_2}\tilde U_{J,1}(t,p)$ *converges to* $-g_{12}(t)$,

iii) $\frac{1}{J}\frac{\partial}{\partial\theta_1}\tilde U_{J,2}(t,p)$ *converges to* 0,

iv) $\frac{1}{J}\frac{\partial}{\partial\theta_2}\tilde U_{J,2}(t,p)$ *converges to* $-g_{22}(t)$,

where

$$g_{12}(t) = E_p\int_0^t \frac{\frac{\partial\lambda}{\partial\theta_1}(u-Y_1,Z_1,p)\frac{\partial\lambda}{\partial\theta_2}(u-Y_1,Z_1,p)}{\lambda(u-Y_1,Z_1,p)} 1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du,$$

$$\begin{aligned} g_{22}(t) = {}& E_p\int_0^t \frac{(\frac{\partial\lambda}{\partial\theta_2}(u-Y_1,Z_1,p))^2}{\lambda(u-Y_1,Z_1,p)} 1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du \\ &+E_p\int_0^t a(u,p)\frac{\frac{\partial\lambda}{\partial\theta_1}(u-Y_1,Z_1,p)\frac{\partial\lambda}{\partial\theta_2}(u-Y_1,Z_1,p)}{\lambda(u-Y_1,Z_1,p)} 1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du. \end{aligned}$$

We note that $g_{22}(t)$ also equals

$$\begin{aligned} &E_p\int_0^t \frac{(\frac{\partial\lambda}{\partial\theta_2}(u-Y_1,Z_1,p))^2}{\lambda(u-Y_1,Z_1,p)} 1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du \\ &+E_p\int_0^t a^2(u,p)\frac{(\frac{\partial\lambda}{\partial\theta_1}(u-Y_1,Z_1,p))^2}{\lambda(u-Y_1,Z_1,p)} 1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du, \end{aligned}$$

and $g_2(t)$ also equals

$$\begin{aligned} &E_p\int_0^t \frac{(\frac{\partial\lambda}{\partial\theta_2}(u-Y_1,Z_1,p))^2}{\lambda(u-Y_1,Z_1,p)} 1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du \\ &-E_p\int_0^t a^2(u,p)\frac{(\frac{\partial\lambda}{\partial\theta_1}(u-Y_1,Z_1,p))^2}{\lambda(u-Y_1,Z_1,p)} 1_{(Y_1,Y_1+X_1\wedge C_1]}(u)\,du. \end{aligned}$$

Applying the Mean-Value Theorem, we get

$$\begin{aligned}\frac{1}{\sqrt{J}}\begin{pmatrix}\tilde{U}_{J,1}(t,p)\\ \tilde{U}_{J,2}(t,p)\end{pmatrix} &= \frac{1}{\sqrt{J}}\begin{pmatrix}\tilde{U}_{J,1}(t,p+\frac{\phi}{\sqrt{J}})\\ \tilde{U}_{J,2}(t,p+\frac{\phi}{\sqrt{J}})\end{pmatrix}\\ &\quad -\frac{1}{J}\begin{pmatrix}\frac{\partial}{\partial\theta_1}\tilde{U}_{J,1}(t,\theta_*) & \frac{\partial}{\partial\theta_2}\tilde{U}_{J,1}(t,\theta_*)\\ \frac{\partial}{\partial\theta_1}\tilde{U}_{J,2}(t,\theta_{**}) & \frac{\partial}{\partial\theta_2}\tilde{U}_{J,2}(t,\theta_{**})\end{pmatrix}\phi, \end{aligned} \tag{2.10}$$

for some θ_* and θ_{**} lying between p and $p+\frac{\phi}{\sqrt{J}}$.
It follows from (2.10), Proposition 2.1 and Proposition 2.2 that we have

Theorem 2.1. *Under $P^{J(\phi)}$, $\frac{1}{\sqrt{J}}\begin{pmatrix}\tilde{U}_{J,1}(t,p)\\ \tilde{U}_{J,2}(t,p)\end{pmatrix}$ converges weakly to $G^\phi(t)=\begin{pmatrix}G_1^\phi(t)\\ G_2^\phi(t)\end{pmatrix}$, where $G_1^\phi(t)-\phi_1 g_1(t)-\phi_2 g_{12}(t)$ and $G_2^\phi(t)-\phi_2 g_{22}(t)$ are two continuous independent mean 0 Gaussian martingales with variances $g_1(t)$ and $g_2(t)$ respectively.*

It follows from (2.6) that predictable variation processes are

$$< \frac{1}{\sqrt{J}}\tilde{U}_{J,1}(\cdot,p) >_t = \frac{1}{J}\sum_{j=1}^{J}\int_0^t \frac{(\frac{\partial\lambda}{\partial\theta_1}(u-Y_j,Z_j,p))^2}{\lambda(u-Y_j,Z_j,p)}1_{(Y_j,Y_j+X_j\wedge C_j]}(u)\,du, \tag{2.11}$$

$$\begin{aligned}< \frac{1}{\sqrt{J}}\tilde{U}_{J,2}(\cdot,p) >_t &= \frac{1}{J}\sum_{j=1}^{J}\int_0^t \frac{(\frac{\partial\lambda}{\partial\theta_2}(u-Y_j,Z_j,p)-a_J(u,p)\frac{\partial\lambda}{\partial\theta_1}(u-Y_j,Z_j,p))^2}{\lambda(u-Y_j,Z_j,p)}\\ &\quad \cdot 1_{(Y_j,Y_j+X_j\wedge C_j]}(u)\,du.\end{aligned} \tag{2.12}$$

Let

$$\tilde{\tau}_{J,i}(s)=\inf\{t\ge 0 \mid < \frac{1}{\sqrt{J}}\tilde{U}_{J,i}(\cdot,p) >_t > s\}. \tag{2.13}$$

Because $< \frac{1}{\sqrt{J}}\tilde{U}_{J,i}(\cdot,p) >_t$ converges almost surely to $g_i(t)$, we know $\tilde{\tau}_{J,i}(s)$ converges almost surely to $g_i^{-1}(s)$. From this together with Theorem 2.1 and the strong representation theorem (cf. Pollard (1984), p.71), we get

Theorem 2.2. *Under $P^{J(\phi)}$, $\frac{1}{\sqrt{J}}\begin{pmatrix}\tilde{U}_{J,1}(\tilde{\tau}_{J,1}(t),p)\\ \tilde{U}_{J,2}(\tilde{\tau}_{J,2}(t),p)\end{pmatrix}$ converges weakly to $B^\phi(t)=\begin{pmatrix}B_1^\phi(t)\\ B_2^\phi(t)\end{pmatrix}$, where $B_1^\phi(t)-\phi_1 t-\phi_2 g_{12}(g_1^{-1}(t))$ and $B_2^\phi(t)-\phi_2 g_{22}(g_2^{-1}(t))$ are two independent standard Brownian motions.*

Theorem 2.2 is useful in providing repeated significance tests for the parameter θ. For example, for the hypothesis

$$H_0:\theta=p,$$

we introduce the stopping time

$$T_i=\inf\Big\{t\geq 0 \;|< \frac{1}{\sqrt{J}}\tilde{U}_{J,i}(\cdot,p)>_t\geq m_0, |\frac{1}{\sqrt{J}}\tilde{U}_{J,i}(t,p)|\geq d_i\sqrt{<\frac{1}{\sqrt{J}}\tilde{U}_{J,i}(\cdot,p)>_t}\Big\}, \tag{2.14}$$

truncated as soon as $<\frac{1}{\sqrt{J}}\tilde{U}_{J,i}(\cdot,p)>_t\geq m$ for some $m>m_0$. The repeated significance test is to stop sampling at $(T_1\wedge\eta_1)\wedge(T_2\wedge\eta_2)$ and reject H_0 if and only if $T_1<\eta_1$ or $T_2<\eta_2$, where

$$\eta_i=\inf\{t\geq 0 \;|< \frac{1}{\sqrt{J}}\tilde{U}_{J,i}(\cdot,p)>_t\geq m\}. \tag{2.15}$$

Here m_0, m and d_i are given constants. Because of the independence guaranteed by Theorem 2.2, the significance level and some of the powers can be approximated by the related results for one-dimensional Brownian motions. (cf. Siegmund (1985), pp.73-81). In Section 3, some of the detailed computation will be illustrated.

3. Numerical Studies

3.1. Complete sequential tests

The distributions of Y_j, Z_j, X_j and C_j for this simulation study are described as follows. Let Y_j be uniformly distributed in the interval $[0,b]$, $C_j=\infty$, the covariate $Z_j=(Z_{j1},Z_{j2})$ and Z_{ji} be i.i.d. with $P(Z_{ji}=0)=P(Z_{ji}=1)=\frac{1}{2}$. Conditional on Z_j, X_j has intensity $\exp(\theta_1Z_{j1}+\theta_2Z_{j2})$. The simulation results presented in the following are based on sample size $J=100$.

Let $T'_{(d)}=\inf\{t\geq m_0 \mid |W_\mu(t)|\geq d\sqrt{t}\}$, where W_μ is a Brownian motion with drift μ. From Corollary 4.19 and Theorem 4.21 of Siegmund (1985), we know

$$P_0(T'_{(d)}<m)=(d-d^{-1})\phi(d)\log(\frac{m}{m_0})+4d^{-1}\phi(d)+o(d^{-1}\phi(d)), \tag{3.1}$$

$$P_\mu(T'_{(d)}<m)=1-\Phi[m^{\frac{1}{2}}(dm^{-\frac{1}{2}}-\mu)]+\{\phi[m^{\frac{1}{2}}(dm^{-\frac{1}{2}}-\mu)]/(\mu m^{\frac{1}{2}})\}\cdot(1+o(1)), \tag{3.2}$$

where P_μ is the probability corresponding to W_μ, ϕ and Φ are respectively the density function and distribution function of a standard normal variable.

We now consider the hypothesis $H_0:\theta=p=(0,0)$. Let $m_0=0.05$ and $m=0.3$, d_i be specified by (3.1) so as to make $P_0(T'_{(d_i)}<m)$ approximately α_i, for i=1, 2. Here α_1, α_2 are non-negative numbers and satisfy $\alpha_1+\alpha_2-\alpha_1\alpha_2=0.05$. With these constants, we will consider the tests defined in (2.14) and (2.15).

Straightforward calculations give $a(u,p) = \frac{1}{2}$, $g_{12}(t) = \frac{1}{2}g_1(t)$ and $g_{22}(t) = \frac{5}{3}g_2(t)$. Let $\mu_1 = \sqrt{J}(\theta_1 + \frac{\theta_2}{2})$, $\mu_2 = \frac{5}{3}\sqrt{J}\theta_2$. Theorem 2.2 indicates that

$$\begin{aligned} & P^{(\theta_1,\theta_2)}(T_1 < \eta_1 \text{ or } T_2 < \eta_2) \\ = & P^{(\theta_1,\theta_2)}(T_1 < \eta_1) + P^{(\theta_1,\theta_2)}(T_2 < \eta_2) - P^{(\theta_1,\theta_2)}(T_1 < \eta_1,\ T_2 < \eta_2) \end{aligned} \tag{3.3}$$

is approximately

$$P_{\mu_1}(T'_{(d_1)} < m) + P_{\mu_2}(T'_{(d_2)} < m) - P_{\mu_1}(T'_{(d_1)} < m)P_{\mu_2}(T'_{(d_2)} < m), \tag{3.4}$$

where $P^{(\theta_1,\theta_2)}$ is the probability corresponding to the parameter (θ_1, θ_2).

When $\theta_1 = \theta_2 = 0$, (3.4) shows that the nominal significance level of the test is 0.05. For other values of θ_1, θ_2, we can use (3.1), (3.2) and (3.4) to find asymptotic power, denoted by β. We denote the empirical power (3.3) by $\beta^{(0)}$ for $b = 0$ and $\beta^{(1)}$ for $b = 1$. We recall that $b = 0$ is the simultaneous entry case.

Tables 1-4 present, for different values of θ_1 and θ_2, the empirical powers $\beta^{(0)}$ and $\beta^{(1)}$ by calculating the proportion of 10,000 replicates for which $T_1 < \eta_1$ or $T_2 < \eta_2$.

Table 1. Significance levels, empirical and asymptotic powers $\beta^{(0)}$ and β for $\alpha_1 = \alpha_2 = 0.02532$ and for various values of (θ_1, θ_2) and $\triangle t$.

$(\theta_1,\theta_2) \setminus \triangle t$	1	0.1	0.01	0.001	0.0001	β
(0.0,0.0)	0.0078	0.0284	0.0473	0.0472	0.0438	0.0500
(0.1,0.0)	0.0233	0.0532	0.0787	0.0914	0.0858	0.0648
(0.0,0.1)	0.0257	0.0578	0.0850	0.0924	0.0922	0.0982
(0.1,0.1)	0.0552	0.1074	0.1341	0.1477	0.1460	0.1175

Table 2. Significance levels, empirical and asymptotic powers $\beta^{(1)}$ and β for $\alpha_1 = \alpha_2 = 0.02532$ and for various values of (θ_1, θ_2) and $\triangle t$.

$(\theta_1,\theta_2) \setminus \triangle t$	1	0.1	0.01	0.001	0.0001	β
(0.0,0.0)	0.0106	0.0294	0.0418	0.0461	0.0509	0.0500
(0.1,0.0)	0.0248	0.0617	0.0810	0.0879	0.0897	0.0648
(0.0,0.1)	0.0229	0.0536	0.0757	0.0848	0.0799	0.0982
(0.1,0.1)	0.0523	0.1028	0.1327	0.1389	0.1369	0.1175

To simulate empirical power $\beta^{(0)}$ and $\beta^{(1)}$, we need to simulate the score process and then check whether it has crossed the boundary from information time m_0 to m. Since the score process changes its value continuously, we consider $t \in [0, 3]$ and divide the interval $[0, 3]$ into equal spaced subintervals and only check whether it has crossed the boundary at the endpoints of these subintervals.

Table 3. Significance levels, empirical and asymptotic powers $\beta^{(0)}$ and β for $\alpha_1 = 0.005, \alpha_2 = 0.04523$ and for various values of (θ_1, θ_2) and $\triangle t$.

$(\theta_1, \theta_2) \setminus \triangle t$	1	0.1	0.01	0.001	0.0001	β
(0.0,0.0)	0.0093	0.0299	0.0428	0.0470	0.0473	0.0500
(0.1,0.0)	0.0161	0.0391	0.0621	0.0692	0.0642	0.0540
(0.0,0.1)	0.0313	0.0653	0.0826	0.0902	0.0919	0.1084
(0.1,0.1)	0.0444	0.0833	0.1153	0.1211	0.1223	0.1151

Table 4. Significance levels, empirical and asymptotic powers $\beta^{(1)}$ and β for $\alpha_1 = 0.005, \alpha_2 = 0.04523$ and for various values of (θ_1, θ_2) and $\triangle t$.

$(\theta_1, \theta_2) \setminus \triangle t$	1	0.1	0.01	0.001	0.0001	β
(0.0,0.0)	0.0100	0.0303	0.0451	0.0486	0.0449	0.0500
(0.1,0.0)	0.0199	0.0427	0.0666	0.0638	0.0683	0.0540
(0.0,0.1)	0.0267	0.0572	0.0750	0.0791	0.0825	0.1084
(0.1,0.1)	0.0428	0.0855	0.1005	0.1128	0.1109	0.1151

In Tables 1-4, $\triangle t$ is the length of the subintervals. The first column is the value of the parameter (θ_1, θ_2), the 2nd, 3rd, 4th and 5th columns display the empirical powers $\beta^{(0)}$ or $\beta^{(1)}$ when $\triangle t = 1, 0.1, 0.01, 0.001, 0.0001$ respectively, and the last column is the asymptotic power β. These tables indicate that the empirical powers get close to the asymptotic powers as the subintervals $\triangle t$ gets smaller, and (3.4) provides quite good approximation to the power of the test.

Remark. With the previous discussion, we now indicate an additional advantage of our approach as follows.

Given the significance level α, and parameter values θ_1, θ_2 satisfying $\theta_1^2+\theta_2^2 = 1$, for example, we can use (3.1) and (3.2) to choose α_1, α_2 and d_1, d_2 so that the test has a large power in the direction (θ_1, θ_2).

3.2. Group sequential tests

Since one typically performs group sequential tests in medical applications, we present a simulation study for such tests based on the work in Section 2 and Subsection 3.1. In particular, the data in this simulation study is the same as described in the first paragraph of Subsection 3.1.

For given $d_1, d_2, \triangle t$ and K, we define

$$\bar{T}_i = \inf\{t \geq 0 | t = k \cdot \triangle t, k = 1, \ldots, K, |\frac{1}{\sqrt{J}} \tilde{U}_{J,i}(\tilde{\tau}_{J,i}(t), p)| \geq d_i \sqrt{t}\}. \quad (3.5)$$

Let

$$\bar{T}'_{(d)} = \inf\{t \geq 0 | t = k \cdot \triangle t, k = 1, \ldots, K, |W_\mu(t)| \geq d\sqrt{t}\}. \quad (3.6)$$

It follows from Theorem 2.2 that the significance level and power of the group sequential test

$$P^{(\theta_1,\theta_2)}(\bar{T}_1 \le K \ \text{ or } \ \bar{T}_2 \le K) \tag{3.7}$$

is approximately

$$P_{\mu_1}(\bar{T}'_{(d_1)} \le K) + P_{\mu_2}(\bar{T}'_{(d_2)} \le K) - P_{\mu_1}(\bar{T}'_{(d_1)} \le K)P_{\mu_2}(\bar{T}'_{(d_2)} \le K), \tag{3.8}$$

where $P^{(\theta_1,\theta_2)}$, P_{μ_1} and P_{μ_2} are the same as described in Subsection 3.1.

Let $\bar{\beta}^{(0)}$ and $\bar{\beta}^{(1)}$ denote respectively the empirical significance levels and empirical powers of (3.7) for $b = 0$ and $b = 1$. Let $\bar{\beta}$ denote the empirical significance levels and empirical powers of (3.8).

Tables 5-6 present $\bar{\beta}^{(0)}$, $\bar{\beta}^{(1)}$ and $\bar{\beta}$ for $K = 5$, $d_1 = d_2 = 2.68$ and $\triangle t = 0.05$ and for different sample sizes. Both $\bar{\beta}^{(0)}$ and $\bar{\beta}^{(1)}$ are the proportion of 10,000 replicates for which $\bar{T}_1 \le K$ or $\bar{T}_2 \le K$. Since (3.6) involves really only multivariate normal distributions, $\bar{\beta}$ is the proportion of 10,000 replicates for which $\bar{T}'_{(d_1)} \le K$ or $\bar{T}'_{(d_2)} \le K$. These simulation results indicate that (3.8) approximates (3.7) reasonably well.

Table 5. Significance levels and empirical powers $\bar{\beta}^{(0)}$ and $\bar{\beta}^{(1)}$ for $J = 100$ and $\bar{\beta}$ for various values of (θ_1, θ_2).

(θ_1,θ_2)	$\bar{\beta}^{(0)}$	$\bar{\beta}^{(1)}$	$\bar{\beta}$
(0.0,0.0)	0.0485	0.0465	0.0491
(0.1,0.0)	0.0813	0.0847	0.0614
(0.0,0.1)	0.0862	0.0726	0.0870
(0.1,0.1)	0.1340	0.1287	0.1135

Table 6. Significance levels and empirical powers $\bar{\beta}^{(0)}$ and $\bar{\beta}^{(1)}$ for $J = 200$ and $\bar{\beta}$ for various values of (θ_1, θ_2).

(θ_1,θ_2)	$\bar{\beta}^{(0)}$	$\bar{\beta}^{(1)}$	$\bar{\beta}$
(0.0,0.0)	0.0499	0.0505	0.0506
(0.1,0.0)	0.1053	0.1026	0.0771
(0.0,0.1)	0.1095	0.0949	0.1333
(0.1,0.1)	0.1971	0.1866	0.1898

Acknowledgements

Thanks are due to Y. C. Yao and Z. Ying for helpful conversations during the preparation of this work. In particular, the simplified proof for Theorem 2.2

using strong representation was suggested by Ying. We are also grateful to the referee/Associate Editor for their valuable suggestions on the numerical studies.

This research is partially supported by the National Science Council of R.O.C.

References

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Process.* Springer, New York.

Chang, I.-S. and Hsiung, C. A. (1988). Likelihood process in parametric model of censored data with staggered entry–Asymptotic properties and applications. *J. Multivariate Anal.* **24**, 31-45.

Chang, I.-S., Hsiung, C. A. and Chuang, Y.-C. (1997). Applications of a frailty model to sequential survival analysis. *Statist. Sinica.* **7**, 127-138.

Elliott, R. Y. (1982). *Stochastic Calculus and Applications.* Springer, New York.

Pollard, D. (1984). *Convergence of Stochastic Process.* Springer-Verlag, New York.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals.* Springer, New York.

Tsiatis, A. A., Boucher, H. and Kim, K. (1995). Sequential methods for parametric survival models. *Biometrika* **82**, 165-173.

Department of Mathematics, National Central University, Chung-Li, Taiwan.

E-mail: chang@math.ncu.edu.tw

Division of Biostatistics, National Health Research Institutes and Institute of Statistical Science, Academia Sinica, Taipei, Taiwan.

E-mail: hsiung@nhri.org.tw

Department of Mathematics, Tamkang University, Tamsui, Taipei, Taiwan.

E-mail: lchwang@math.tku.edu.tw

(Received September 1997; accepted September 1998)