Journal of the Korean
Statistical Society
Vol. 24, No. 2, 1995

A Goodness-of-Fit Test for the Additive Risk Model with a Binary Covariate

Jinheum Kim[1] and Moon Sup Song[2]

ABSTRACT

In this article, we propose a class of weighted estimators for the excess risk in additive risk model with a binary covariate. The proposed estimator is consistent and asymptotically normal. When the assumed model is inappropriate, however, the estimators with different weights converge to nonidentical constants. This fact enables us to develop a goodness-of-fit test for the excess assumption by comparing estimators with different weights. It is shown that the proposed test converges in distribution to normal with mean zero and is consistent under the model misspecifications. Furthermore, the finite-sample properties of the proposed test procedure are investigated and two examples using real data are presented.

KEYWORDS: Goodness-of-fit test, Proportional hazards model, Additive risk model, Weighted estimator, Martingale, Excess risk.

[1]Statistical Research Institute, College of Natural Sciences, Seoul National University, Seoul, 151–742, Korea.

[2]Department of Computer Science and Statistics, Seoul National University, Seoul, 151–742, Korea.

1. INTRODUCTION

In investigating the association between covariates and the survival time, the proportional hazards and additive risk models can be considered. These two models are different in specifying the hazard function for the survival time T associated with a p vector of possibly time-varying covariates $Z(\cdot)$. The Cox proportional hazards model postulates the proportional effect of covariates $Z(t)$ on the baseline hazard function $\lambda_0(t)$,

$$\lambda(t; \boldsymbol{Z}) = \lambda_0(t) \exp(\boldsymbol{\gamma}_0' \boldsymbol{Z}(t)), \tag{1.1}$$

while the additive risk model postulates the additive effect,

$$\lambda(t; \boldsymbol{Z}) = \lambda_0(t) + \boldsymbol{\beta}_0' \boldsymbol{Z}(t), \tag{1.2}$$

where $\boldsymbol{\beta}_0$ and $\boldsymbol{\gamma}_0$ are p vectors of regression parameters (Lin and Ying (1994)).

In order to assess the adequacy of the proportional hazards assumption in model (1.1), many methods have been developed. Among them, Gill and Schumacher (1987) have considered two-sample model having $p = 1$ and an indicator covariate Z as a special case of model (1.1) and have constructed a test based on the difference of generalized rank estimators for the relative risk. The key idea behind the test procedure is that in nonproportional hazards situations, two different consistent estimators for the relative risk might give significantly different results. Lin (1991) extended the idea of Gill and Schumacher (1987) to model (1.1). In his article, a class of estimating functions for γ_0 has been introduced by incorporating the weights into the partial likelihood score function. When model (1.1) is inappropriate, the difference between two consistent estimators with different weights may be large.

The purpose of this article is to propose a test for checking the adequacy of model (1.2) with $p = 1$ and with a single binary covariate Z. In other words, we intend to construct a test procedure to assess whether the hazard difference between two groups in two-sample problem is constant or not. In the next section, we introduce some notations and assumptions. In Section

3, we propose a test based on the difference between two weighted estimators for the excess risk, which is a version of Gill and Schumacher's test. Also, the asymptotic and finite-sample properties of the proposed test are investigated. Finally, two examples using real data are provided in Section 4 as illustrations.

2. PRELIMINARIES

Let $X = \min(T, C)$, where T is the true survival time with an absolutely continuous distribution function and C is the censoring time corresponding to T. Let $\Delta = I(T \leq C)$, where $I(\cdot)$ is the indicator function. As mentioned in the previous section, let the covariate Z for a subject in group 1 be equal to 1 and Z in group 2 be equal to 0. Assume that T and C are independent and for each group i $(i = 1, 2)$, $(X_{ij}, \Delta_{ij}, Z_{ij})$ $(j = 1, \ldots, n_i)$ are independent and identically distributed replicates of (X, Δ, Z).

Furthermore, let us define some stochastic processes. For each $i = 1, 2$; $j = 1, \ldots, n_i$, let $N_{ij}(t) = I(X_{ij} \leq t \text{ and } \Delta_{ij} = 1)$ and $Y_{ij}(t) = I(X_{ij} \geq t)$. Also, for each $i = 1, 2$, let $N_i(t) = \sum_{j=1}^{n_i} N_{ij}(t)$ and $Y_i(t) = \sum_{j=1}^{n_i} Y_{ij}(t)$. Then, $N_i(t)$ means the number of deaths before or at t in group i and $Y_i(t)$ the number at risk at $t-$ in group i.

The additive risk model considered in this article can be represented as

$$\lambda(t; Z) = \lambda_0(t) + \beta_0 Z, \tag{2.1}$$

where β_0 is an unknown regression parameter. Then, under model (2.1), the counting processes $N_{ij}(t)$ have the intensity function $Y_{ij}(t)(\lambda_0(t) + \beta_0 Z_{ij})$. Also, according to the Doob-Meyer decomposition, $N_{ij}(t)$ can be uniquely decomposed so that $M_{ij}(t) = N_{ij}(t) - \int_0^t Y_{ij}(u)(\lambda_0(u) + \beta_0 Z_{ij})du$, where $M_{ij}(\cdot)$ is a square integrable martingale (Andersen, Borgan, Gill and Keiding (1993)).

In this article, we are interested in testing the additivity between two hazard functions. Therefore, our test problem is given by

$$\begin{gathered} H_0 : \lambda(t; Z = 1) - \lambda(t; Z = 0) = \beta_0 \quad \text{for some} \quad \beta_0 \\ vs. \\ H_1 : \lambda(t; Z = 1) - \lambda(t; Z = 0) \neq \beta_0 \quad \text{for any} \quad \beta_0. \end{gathered}$$

3. THE PROPOSED TEST AND ITS ASYMPTOTIC PROPERTIES

Under model (2.1), the regression parameter β_0 is interpreted as the excess risk, the difference between two hazard functions. Define a weighted estimator $\hat{\beta}_K$ for the excess risk β_0 as

$$\hat{\beta}_K = \int_0^\tau \frac{K(u)}{\int_0^\tau K(v)dv} \left[\frac{dN_1(u)}{Y_1(u)} - \frac{dN_2(u)}{Y_2(u)} \right],$$

where $K(\cdot)$ is a predictable weight function and $\tau = \inf\{t |\ Y_1(t)Y_2(t) = 0\}$. To investigate the asymptotic properties of the weighted estimator $\hat{\beta}_K$, let us assume that the following conditions hold:

(C1) $n_i/n \longrightarrow \rho_i$ $(i = 1, 2)$, where $n = n_1 + n_2$.

(C2) There exist functions y_1 and y_2 taking values in $(0, 1)$ such that under model (2.1), $\sup_{0 \le t \le \tau} |Y_i(t)/n_i - y_i(t)| \xrightarrow{p} 0$ $(i = 1, 2)$, where $\xrightarrow{p}$ denotes convergence in probability.

(C3) There exists a function g taking values in $[0, \infty)$ not everywhere zero such that under model (2.1), $\sup_{0 \le t \le \tau} |K(t) - g(t)| \xrightarrow{p} 0$.

Assuming that model (2.1) holds and the conditions (C1)–(C3) are satisfied, the consistency of weighted estimator $\hat{\beta}_K$ can be easily shown from Lenglart's inequality (Andersen, Borgan, Gill and Keiding (1993)) and the weak convergence of $n^{1/2}(\hat{\beta}_K - \beta_0)$ can be derived in the following theorem.

Theorem 1. Suppose that the conditions (C1)–(C3) are satisfied. Then, under model (2.1), $n^{1/2}(\hat{\beta}_K - \beta_0)$ converges in distribution to normal with mean zero and variance

$$\sigma_K^2 = \int_0^\tau \left[\frac{g(u)}{\int_0^\tau g(v)dv} \right]^2 \left[\frac{d\Lambda_1(u)}{\rho_1 y_1(u)} + \frac{d\Lambda_2(u)}{\rho_2 y_2(u)} \right],$$

where $\Lambda_i(\cdot)$ is the cumulative hazard function corresponding to group i. Furthermore, σ_K^2 can be consistently estimated by

$$\hat{\sigma}_K^2 = n \int_0^\tau \left[\frac{K(u)}{\int_0^\tau K(v)dv}\right]^2 \left[\frac{dN_1(u)}{Y_1(u)^2} + \frac{dN_2(u)}{Y_2(u)^2}\right].$$

Proof. Note that under model (2.1),

$$\sqrt{n}(\hat{\beta}_K - \beta_0) = \sqrt{n} \int_0^\tau \frac{K(u)}{\int_0^\tau K(v)dv} \left[\frac{dM_1(u)}{Y_1(u)} - \frac{dM_2(u)}{Y_2(u)}\right],$$

where $M_i(\cdot) = \sum_{j=1}^{n_i} M_{ij}(\cdot)$. According to the martingale central limit theorem and the conditions (C1)–(C3),

$$\sqrt{n} \int_0^\tau K(u) \left[\frac{dM_1(u)}{Y_1(u)} - \frac{dM_2(u)}{Y_2(u)}\right]$$

converges in distribution to normal with mean zero and variance

$$\int_0^\tau g(u)^2 \left[\frac{d\Lambda_1(u)}{\rho_1 y_1(u)} + \frac{d\Lambda_2(u)}{\rho_2 y_2(u)}\right].$$

Therefore, $n^{1/2}(\hat{\beta}_K - \beta_0)$ converges in distribution to normal with mean zero and variance σ_K^2.

On the other hand, we can guess from the consistency property of weighted estimators that the discrepancy between two consistent estimators $\hat{\beta}_{K_1}$ and $\hat{\beta}_{K_2}$ corresponding to weight functions K_1 and K_2, respectively, is small as long as model (2.1) is valid and some appropriate conditions are satisfied. Therefore, it is reasonable to consider $D_{K_1K_2}$, defined by $D_{K_1K_2} = \hat{\beta}_{K_1} - \hat{\beta}_{K_2}$, as a measure for checking the adequacy of model (2.1).

Theorem 2. Suppose that model (2.1) holds and the conditions (C1)–(C3) are satisfied. Then, $T_{K_1K_2} = n^{1/2}\hat{\sigma}_{K_1K_2}^{-1} D_{K_1K_2}$ converges in distribution to standard normal, where $\hat{\sigma}_{K_1K_2}^2$ is a consistent estimator of

$$\sigma_{K_1K_2}^2 = \int_0^\tau \left[\frac{g_1(u)}{\int_0^\tau g_1(v)dv} - \frac{g_2(u)}{\int_0^\tau g_2(v)dv}\right]^2 \left[\frac{d\Lambda_1(u)}{\rho_1 y_1(u)} + \frac{d\Lambda_2(u)}{\rho_2 y_2(u)}\right]$$

and is given by $\hat{\sigma}^2_{K_1K_2} = n(\hat{V}_{11} - \hat{V}_{12} - \hat{V}_{21} + \hat{V}_{22})$ with

$$\hat{V}_{ii'} = \int_0^\tau \frac{K_i(u)K_{i'}(u)}{\int_0^\tau K_i(v)dv \int_0^\tau K_{i'}(v)dv} \left[\frac{dN_1(u)}{Y_1(u)^2} + \frac{dN_2(u)}{Y_2(u)^2}\right] \quad (i, i' = 1, 2).$$

Proof. Under model (2.1), $D_{K_1K_2}$ can be represented as

$$D_{K_1K_2} = \int_0^\tau \left[\frac{K_1(u)}{\int_0^\tau K_1(v)dv} - \frac{K_2(u)}{\int_0^\tau K_2(v)dv}\right] \left[\frac{dM_1(u)}{Y_1(u)} - \frac{dM_2(u)}{Y_2(u)}\right].$$

From Fleming and Harrington (1991) and the conditions (C1)–(C3),

$$n^{-1/2}\left[\int_0^\tau \frac{nK_{i'}(u)}{Y_i(u)} dM_i(u) - \int_0^\tau \frac{g_{i'}(u)}{\rho_i y_i(u)} dM_i(u)\right] \xrightarrow{d} 0 \quad (i, i' = 1, 2),$$

where $\xrightarrow{d}$ denotes convergence in distribution and $g_{i'}(\cdot)$ is the limit of weight function $K_{i'}(\cdot)$ satisfying the condition (C3). Therefore, $n^{1/2}D_{K_1K_2}$ is asymptotically equivalent to

$$n^{1/2}\tilde{D}_{K_1K_2} = n^{-1/2}\int_0^\tau \left[\frac{g_1(u)}{\int_0^\tau g_1(v)dv} - \frac{g_2(u)}{\int_0^\tau g_2(v)dv}\right] \left[\frac{dM_1(u)}{\rho_1 y_1(u)} - \frac{dM_2(u)}{\rho_2 y_2(u)}\right].$$

Since $n^{1/2}\tilde{D}_{K_1K_2}$ is simply a sum of independent and identically distributed random variables, it follows by the martingale central limit theorem that $n^{1/2}\tilde{D}_{K_1K_2}$ converges in distribution to normal with mean zero. Also, the variance of the limiting distribution can be derived as follows:

$$\begin{aligned}
&E(n\tilde{D}^2_{K_1K_2}) \\
&= n^{-1}E\left\{\int_0^\tau \left[\frac{g_1(u)}{\int_0^\tau g_1(v)dv} - \frac{g_2(u)}{\int_0^\tau g_2(v)dv}\right] \left[\frac{dM_1(u)}{\rho_1 y_1(u)} - \frac{dM_2(u)}{\rho_2 y_2(u)}\right]\right\}^2 \\
&= n^{-1}E\left\{\int_0^\tau \left[\frac{g_1(u)}{\int_0^\tau g_1(v)dv} - \frac{g_2(u)}{\int_0^\tau g_2(v)dv}\right]^2 \left[\frac{Y_1(u)d\Lambda_1(u)}{\rho_1^2 y_1^2(u)} + \frac{Y_2(u)d\Lambda_2(u)}{\rho_2^2 y_2^2(u)}\right]\right\} \\
&\xrightarrow{p} \int_0^\tau \left[\frac{g_1(u)}{\int_0^\tau g_1(v)dv} - \frac{g_2(u)}{\int_0^\tau g_2(v)dv}\right]^2 \left[\frac{d\Lambda_1(u)}{\rho_1 y_1(u)} + \frac{d\Lambda_2(u)}{\rho_2 y_2(u)}\right].
\end{aligned}$$

Therefore, $n^{1/2}D_{K_1K_2}$ converges in distribution to normal with mean zero and variance $\sigma^2_{K_1K_2}$.

According to Theorem 2, if the absolute value of test statistic $T_{K_1K_2}$ is large, we may conclude that the additive risk assumption is not valid any longer. Furthermore, it can be shown that the proposed test statistic $T_{K_1K_2}$ is consistent against alternatives with monotone hazard differences provided $H(t) = (\int_0^\tau K_1(u)du)^{-1}K_1(t) - (\int_0^\tau K_2(u)du)^{-1}K_2(t)$ is monotone, too.

Theorem 3. Suppose that the conditions (C1)–(C3) are satisfied. Then, the goodness-of-fit test $T_{K_1K_2}$ is consistent against the model misspecification, $\lambda(t;Z) = \lambda_0(t) + \beta(t)Z$, where $\beta(t)$ is an unspecified monotone function of t provided $h(t) = (\int_0^\tau g_1(u)du)^{-1}g_1(t) - (\int_0^\tau g_2(u)du)^{-1}g_2(t)$, which is a limit of $H(t)$, is monotone in t.

Proof. Note that from Fleming and Harrington (1991) and the condition (C3),

$$\int_0^\tau K_{i'}(u)\frac{dN_i(u)}{Y_i(u)} \xrightarrow{p} \int_0^\tau g_{i'}(u)d\Lambda_i(u) \quad (i,i'=1,2). \tag{3.1}$$

Then, $D_{K_1K_2}$ in $T_{K_1K_2}$ converges in probability to

$$\int_0^\tau \left[\frac{g_1(u)}{\int_0^\tau g_1(v)dv} - \frac{g_2(u)}{\int_0^\tau g_2(v)dv}\right](d\Lambda_1(u) - d\Lambda_2(u)). \tag{3.2}$$

If both the hazard difference, $\lambda(t;Z=1) - \lambda(t;Z=0)$, and $h(\cdot)$ are monotone increasing or decreasing, then the limit (3.2) of $D_{K_1K_2}$ is positive. Similarly, if they are both monotone but in different directions, (3.2) is negative.

Now, consider the term $\hat\sigma_{K_1K_2}$ in $T_{K_1K_2}$. From the conditions (C1)–(C3) and (3.1), $n\hat V_{ii'}$ $(i,i'=1,2)$ converges in probability to

$$\int_0^\tau \frac{g_i(u)g_{i'}(u)}{\int_0^\tau g_i(v)dv \int_0^\tau g_{i'}(v)dv}\left[\frac{d\Lambda_1(u)}{\rho_1 y_1(u)} + \frac{d\Lambda_2(u)}{\rho_2 y_2(u)}\right]. \tag{3.3}$$

Thus, the limit (3.3) of $n\hat V_{ii'}$ is positive from the assumption that both g_1 and g_2 are positive. Finally, from the above two results, the test statistic

$T_{K_1K_2}$ is consistent against alternatives with monotone departures of excess risk provided $h(\cdot)$ is monotone.

4. FINITE-SAMPLE PROPERTIES OF $\hat{\beta}_K$ AND $T_{K_1K_2}$

4.1 Finite-Sample Properties of the Proposed Estimator

Monte Carlo experiments were carried out to assess the performance of $\hat{\beta}_K$ proposed in previous section for practical sample sizes. Table 1 shows some typical results. For this table, survival times were generated from the additive risk model, $\lambda(t;Z) = \lambda_0(t) + \beta_0 Z$, and censoring times from the uniform distribution $U(0,c)$, where $\lambda_0(\cdot)$ has a Weibull distribution and c is suitably chosen to ensure the desired censoring proportion. Also, three weight functions satisfying the condition (C3) were considered as follows: Gehan's weight, $K_G(t) = Y_1(t)Y_2(t)$; logrank weight, $K_L(t) = (Y_1(t)+Y_2(t))^{-1}Y_1(t)Y_2(t)$; Prentice-Wilcoxon's weight, $K_P(t) = (Y_1(t)+Y_2(t))^{-1}Y_1(t)Y_2(t)\hat{S}(t)$. Here, $\hat{S}(\cdot)$ denotes the Kaplan-Meier estimator of the survival function based on the combined sample of two groups (Kaplan and Meier (1958)). To investigate the differences between three weight functions, let us a weight process, $L(t) = W(t)Y_1(t)Y_2(t)(Y_1(t)+Y_2(t))^{-1}$. $K_G(t)$ with $W(t) = Y_1(t)+Y_2(t)$ gives relatively more weight to the early deaths than $K_L(t)$ with $W(t) = 1$. $K_P(t)$ with $W(t) = \hat{S}(t)$ is preferable to $K_G(t)$ when censoring is heavy. It is evident from Table 1 that our proposed estimators are nearly unbiased and also are better than the unweighted estimator.

4.2 Finite-Sample Properties of the Proposed Test Statistic

In our simulation studies, the empirical sizes and powers of test statistic $T_{K_1K_2}$ corresponding to all possible combinations of three different weight functions were considered. Hereafter, T_{GL} denotes the test statistic corresponding to a pair of Gehan's and logrank weight functions, T_{GP} a pair of Gehan's and Prentice-Wilcoxon's weight functions and T_{LP} a pair of logrank and Prentice-Wilcoxon's weight functions. According to Theorem 2, we can perform a two-tailed level α test by comparing the absolute value of the proposed test with the upper $\alpha/2$ quantile of the standard normal distribution. For brevity, $W(\rho,\tau)$ denotes the Weibull distribution with a scale parameter ρ and a shape parameter τ and $\text{Exp}(\lambda)$ the exponential distribution with a parameter λ.

Table 2 displays the empirical sizes of test statistics T_{GL}, T_{GP} and T_{LP}. As shown in Table 2, the sizes of the proposed test are well controlled and the proposed test is stable regardless of censoring distributions.

Table 3 displays the empirical powers of test statistics T_{GL}, T_{GP} and T_{LP}, for detecting the violation of the assumption of constant hazard difference. In order to generate alternatives, three pairs of two groups were considered as follows: $W(\sqrt{2},2)$ *vs.* $W(1,2)$, $W(2,2)$ *vs.* $W(1,2)$ and $W(\sqrt{6},2)$ *vs.* $W(1,2)$; therefore, the hazard differences for each pair are $2t$, $6t$ and $10t$, respectively. Table 3 shows that as N increases and CP decreases, the empirical powers of test statistics increase. Also, it is evident from Table 3 that the more departure from null hypothesis causes the powers of test procedures to increase. Being compared with the powers of tests T_{GL} and T_{LP}, those of test T_{GP} are always larger regardless of sample sizes, proportions of censoring and types of alternative hypothesis. These differences in powers, however, are not substantially great.

5. EXAMPLES

The goodness-of-fit test proposed in Section 3 is applied to two real data and the results are discussed. The first data is taken from Freireich *et al.* (Cox (1972)), which consist of the times to remission for two groups of leukemia patients. The p-values of T_{GL} and T_{LP} are 0.077 and 0.065 respectively. From these results, it is difficult to assume the additivity between hazard functions.

In fact, it has already been known that the Freireich *et al.*'s data satisfy significantly the proportional hazards assumption by Wei (1984) and Gill and Schumacher (1987).

As another example, we consider the small cell lung cancer(SCLC) data (Ying, Jung and Wei (1995)). This data set contains the survival times of 121 patients with SCLC. The 59 patients of 121 are administered etoposide followed by cisplatin and the remaining 62 patients cisplatin followed by etoposide. The p-values of T_{GL} and T_{LP} are 0.368 and 0.348, respectively. So, it is evident that it is reasonable to accept the additive risk assumption. Furthermore, when the same data are applied to Gill and Schumacher's test, the p-value of a pair of Gehan and logrank equals 0.029 and a pair of logrank and Prentice-Wilcoxon 0.027. Therefore, the assumption of proportional hazards is rejected.

Table 1. Empirical Estimators for the Mean and MSE of $\hat{\beta}_K$ Based on 1,000 Replications.

$\hat{\beta}_K$				$\hat{\beta}_{K_G}$		$\hat{\beta}_{K_L}$		$\hat{\beta}_{K_P}$		$\hat{\beta}_{K_*}$ [a]	
β_0	$\lambda_0(t)$	N[b]	CP[c]	mean	MSE	mean	MSE	mean	MSE	mean	MSE
0.5	1	50	0.2[d]	0.525	0.254	0.532	0.218	0.523	0.244	0.607	0.571
			0.5[e]	0.504	0.325	0.523	0.286	0.503	0.290	0.561	0.743
		100	0.2[d]	0.516	0.109	0.519	0.091	0.514	0.104	0.555	0.350
			0.5[e]	0.511	0.160	0.516	0.130	0.508	0.140	0.540	0.411
			0.7[f]	0.505	0.271	0.504	0.226	0.500	0.231	0.489	0.547
	2t	50	0.2[g]	0.504	0.107	0.499	0.148	0.501	0.108	0.497	0.940
			0.5[h]	0.518	0.132	0.517	0.168	0.511	0.135	0.524	0.848
		100	0.2[g]	0.498	0.056	0.502	0.076	0.497	0.057	0.546	0.740
			0.5[h]	0.491	0.067	0.493	0.088	0.488	0.071	0.521	0.676
			0.7[i]	0.526	0.070	0.521	0.089	0.518	0.075	0.518	0.511
2.0	1	50	0.2[j]	2.047	0.658	2.088	0.641	2.038	0.644	2.340	1.698
			0.5[k]	2.025	0.848	2.040	0.762	1.994	0.764	2.163	2.101
		100	0.2[j]	2.019	0.316	2.024	0.295	2.013	0.307	2.174	0.990
			0.5[k]	1.978	0.424	1.991	0.361	1.967	0.379	2.053	1.270
			0.7[l]	2.015	0.671	2.020	0.564	1.996	0.573	2.036	1.490
	2t	50	0.2[m]	2.066	0.368	2.075	0.397	2.056	0.363	2.216	1.387
			0.5[n]	2.086	0.490	2.078	0.514	2.053	0.467	2.156	1.814
		100	0.2[m]	2.042	0.181	2.048	0.207	2.037	0.181	2.140	1.116
			0.5[n]	2.016	0.211	2.015	0.227	2.001	0.207	2.061	1.165
			0.7[o]	2.015	0.244	2.010	0.257	1.991	0.237	2.047	1.160

NOTE: [a] Unweighted estimator with $K(t) = 1$. [b] Sample size in the combined sample. [c] Censoring proportion. [d] $U(0, 4.12)$. [e] $U(0, 1.30)$. [f] $U(0, 0.615)$. [g] $U(0, 3.92)$. [h] $U(0, 1.53)$. [i] $U(0, 0.944)$. [j] $U(0, 3.23)$. [k] $U(0, 0.908)$. [l] $U(0, 0.406)$. [m] $U(0, 3.16)$. [n] $U(0, 1.17)$. [o] $U(0, 0.633)$.

Table 2. Empirical Sizes of Test Statistics, T_{GL}, T_{GP} and T_{LP}, Based on 1,000 Replications.

α			0.01			0.05			0.10		
	N	CP	T_{GL}	T_{GP}	T_{LP}	T_{GL}	T_{GP}	T_{LP}	T_{GL}	T_{GP}	T_{LP}
(A)[a]	50	0.0	0.009	0.009	0.009	0.043	0.043	0.043	0.098	0.098	0.098
		0.2[b]	0.004	0.014	0.005	0.043	0.047	0.042	0.103	0.097	0.104
		0.5[c]	0.006	0.006	0.005	0.048	0.051	0.039	0.092	0.103	0.092
	100	0.0	0.006	0.006	0.006	0.040	0.040	0.040	0.080	0.080	0.080
		0.2[b]	0.008	0.006	0.009	0.051	0.054	0.052	0.109	0.097	0.103
		0.5[c]	0.009	0.008	0.009	0.053	0.051	0.054	0.104	0.099	0.106
	200	0.0	0.007	0.007	0.007	0.047	0.047	0.047	0.084	0.084	0.084
		0.2[b]	0.004	0.009	0.004	0.046	0.050	0.042	0.101	0.093	0.102
		0.5[c]	0.011	0.012	0.011	0.052	0.051	0.051	0.099	0.100	0.102
(B)[d]	50	0.0	0.010	0.010	0.010	0.046	0.046	0.046	0.099	0.099	0.099
		0.2[e]	0.011	0.017	0.011	0.048	0.059	0.048	0.097	0.108	0.097
		0.5[f]	0.008	0.015	0.009	0.057	0.052	0.046	0.096	0.102	0.091
	100	0.0	0.015	0.015	0.015	0.048	0.047	0.048	0.093	0.095	0.093
		0.2[e]	0.008	0.007	0.013	0.053	0.051	0.054	0.101	0.110	0.099
		0.5[f]	0.009	0.008	0.008	0.038	0.043	0.038	0.089	0.098	0.095
	200	0.0	0.009	0.009	0.009	0.049	0.049	0.049	0.098	0.098	0.098
		0.2[e]	0.009	0.009	0.010	0.059	0.047	0.060	0.107	0.106	0.107
		0.5[f]	0.012	0.011	0.014	0.051	0.063	0.055	0.109	0.119	0.108

NOTE: [a] Survival times of two groups were generated from $W(2, 1)$ and $W(1, 1)$ respectively. [b] $U(0, 3.69)$. [c] $U(0, 1.12)$. [d] Survival times were generated from $\lambda(t; Z) = 2t + 2Z$. [e] $Exp(0.376)$. [f] $Exp(1.36)$.

Table 3. Empirical Powers of Test Statistics, T_{GL}, T_{GP} and T_{LP}, for Detecting the Nonconstant Hazard Differences Based on 1,000 Replications.

α			0.01			0.05			0.10		
N	HD	CP	T_{GL}	T_{GP}	T_{LP}	T_{GL}	T_{GP}	T_{LP}	T_{GL}	T_{GP}	T_{LP}
50	$2t$	0.0	0.018	0.018	0.018	0.159	0.159	0.159	0.295	0.295	0.295
		0.2[a]	0.027	0.051	0.017	0.174	0.190	0.158	0.301	0.293	0.283
		0.5[b]	0.015	0.032	0.007	0.117	0.137	0.090	0.229	0.221	0.203

(*Continued*)

Table 3.(*Continued*) Empirical Powers of Test Statistics, T_{GL}, T_{GP} and T_{LP}, for Detecting the Nonconstant Hazard Differences Based on 1,000 Replications.

	α		0.01			0.05			0.10		
N	HD	CP	T_{GL}	T_{GP}	T_{LP}	T_{GL}	T_{GP}	T_{LP}	T_{GL}	T_{GP}	T_{LP}
50	$6t$	0.0	0.118	0.118	0.118	0.449	0.449	0.449	0.671	0.671	0.671
		0.2^c	0.086	0.129	0.066	0.408	0.372	0.375	0.638	0.537	0.612
		0.5^d	0.044	0.098	0.023	0.315	0.345	0.270	0.524	0.525	0.488
	$10t$	0.0	0.171	0.171	0.171	0.615	0.615	0.615	0.810	0.810	0.810
		0.2^e	0.148	0.165	0.121	0.576	0.487	0.561	0.793	0.659	0.773
		0.5^f	0.080	0.158	0.041	0.463	0.496	0.410	0.669	0.664	0.641
100	$2t$	0.0	0.128	0.128	0.128	0.421	0.421	0.421	0.609	0.609	0.609
		0.2^a	0.101	0.158	0.083	0.336	0.397	0.305	0.508	0.543	0.487
		0.5^b	0.063	0.100	0.038	0.233	0.292	0.185	0.395	0.428	0.350
	$6t$	0.0	0.560	0.560	0.560	0.883	0.883	0.883	0.944	0.944	0.944
		0.2^c	0.496	0.629	0.442	0.857	0.878	0.826	0.941	0.949	0.923
		0.5^d	0.354	0.458	0.252	0.717	0.773	0.629	0.859	0.866	0.803
	$10t$	0.0	0.764	0.764	0.764	0.962	0.962	0.962	0.991	0.991	0.991
		0.2^e	0.756	0.807	0.711	0.960	0.965	0.947	0.990	0.988	0.988
		0.5^f	0.532	0.653	0.432	0.845	0.886	0.802	0.934	0.946	0.905
200	$2t$	0.0	0.425	0.425	0.425	0.745	0.745	0.745	0.847	0.847	0.847
		0.2^a	0.365	0.491	0.322	0.665	0.750	0.629	0.802	0.844	0.758
		0.5^b	0.212	0.322	0.154	0.514	0.591	0.426	0.648	0.712	0.572
	$6t$	0.0	0.977	0.977	0.977	0.998	0.998	0.998	1.000	1.000	1.000
		0.2^c	0.954	0.986	0.926	0.993	0.998	0.992	0.995	0.999	0.995
		0.5^d	0.835	0.908	0.738	0.965	0.977	0.936	0.980	0.995	0.965
	$10t$	0.0	0.996	0.996	0.996	1.000	1.000	1.000	1.000	1.000	1.000
		0.2^e	0.999	0.999	0.997	1.000	1.000	1.000	1.000	1.000	1.000
		0.5^f	0.971	0.987	0.934	0.997	0.999	0.990	1.000	1.000	0.997

NOTE: [a] $U(0,3.78)$. [b] $U(0,1.48)$. [c] $U(0,3.32)$. [d] $U(0,1.26)$. [e] $U(0,3.12)$. [f] $U(0,1.16)$.

REFERENCES

(1) Andersen, P.K., Borgan, O., Gill, R.D., and Keiding N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, New York.

(2) Cox, D.R. (1972). Regression Models and Life-Tables (with Discussion). *Journal of the Royal Statistical Society*, **B 34**, 187–220.

(3) Fleming, T.R. and Harrington, D.P. (1991). *Counting Processes and Survival Analysis*. John Wiley and Sons, New York.

(4) Gill, R.D. and Schumacher, M. (1987). A Simple Test of the Proportional Hazards Assumption. *Biometrika*, **74**, 289–300.

(5) Kaplan, E.L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, **53**, 457–481.

(6) Lin, D.Y. (1991). Goodness-of-Fit Analysis for the Cox Regression Model Based on a Class of Parameter Estimators. *Journal of the American Statistical Association*, **86**, 725–728.

(7) Lin, D.Y. and Ying, Z. (1994). Semiparametric Analysis of the Additive Risk Model. *Biometrika*, **81**, 61–71.

(8) Wei, L.J. (1984). Testing Goodness-of-Fit for Proportional Hazards Model with Censored Observations. *Journal of the American Statistical Association*, **79**, 649–652.

(9) Ying, Z., Jung, S.H., and Wei, L.J. (1995). Survival Analysis with Median Regression Model. *Journal of the American Statistical Association*, **90**, 178–184.