A Test Procedure for Right Censored Data under the Additive Model

Hyo-II Park\textsuperscript{1,a}, Seung-Man Hong\textsuperscript{b}

\textsuperscript{a}Dept. of Statistics, Chong-ju Univ., \textsuperscript{b}Dept. of Information Statistics, Korea Univ.

Abstract

In this research, we propose a nonparametric test procedure for the right censored and grouped data under the additive hazards model. For deriving the test statistics, we use the likelihood principle. Then we illustrate proposed test with an example and compare the performance with other procedure by obtaining empirical powers. Finally we discuss some interesting features concerning the proposed test.

Keywords: Additive hazards model, grouped data, log-rank test, score function.

1. Introduction

The proportional hazards model (PHM) has been one of the most frequently applied ones for the analysis of the life-time data. Since Cox (1972) has proposed the PHM, the PHM has been developed and modified successfully in many various situations. However when the proportionality among hazard functions may be suspicious, one may as well consider an alternative model rather than clinging to the PHM. Then the additive hazards model (AHM) may be a candidate for any possible alternatives. Let \( \lambda_0 \) be the baseline hazard function and \( z \), the regression vector, which is independent of the time \( t \). Then the hazard function \( \lambda(t, z) \) for the AHM can be represented with the \( p \times 1 \) regression coefficient vector \( \beta \) as follows:

\[
\lambda(t, z) = \lambda_0(t) + \beta'z, \tag{1.1}
\]

where the prime represents the transpose of a vector or matrix. Then the corresponding cumulative hazard function, \( \Lambda(t, z) \) and survival function, \( S(t, z) \) under the AHM (1.1) can be written as follows with the facts that \( \int_0^t \beta'zdx = t\beta'z \) and \( S(t) = \exp[-\Lambda(t)] \):

\[
\Lambda(t, z) = \int_0^t \left( \lambda_0(x) + \beta'z \right) dx = \Lambda_0(t) + t\beta'z
\]

and

\[
S(t, z) = \exp[-\Lambda_0(t)] \exp[-t\beta'z]. \tag{1.2}
\]

As an alternative model to the PHM, the AHM has not been widely used. The main reason for this may come from the fact that the conditional likelihood proposed by Cox (1972) can not be applied to the AHM because of the structure of the hazard function. The AHM (1.1) was initiated by Aalen (1980, 1989), who considered an inference procedure for \( \lambda_0 \) and \( \beta \) applying the least squares method.

\textsuperscript{1} This research was supported by the Korea Science and Engineering Foundation grant funded by the Korea government (MOST) (No. R01-2007-000-10666-0).

\textsuperscript{a} Corresponding author: Professor, Department of Statistics, Chong-ju University, Chong-ju, Choong-book 360-764, Korea. E-mail: hipark@cju.ac.kr
McKeague (1988) and Huffer and McKeague (1991) considered the weighted least squares estimates under some optimality consideration. Also Lin and Ying (1994) proposed an estimate procedure for using the counting process which has been used for the PHM as an ad hoc approach. McKeague and Sasieni (1994) developed partly parametric AHM. Also Scheike (2002) worked the AHM in this direction. For the multivariate data, Yin and Cai (2004) considered inferences based on the marginal AHM approach.

Sometimes one cannot help observing the objects whether they fail or not periodically or with time-schedule for some reasons. For example, after being exposed to the HIV virus, the observation must be carried out periodically since it usually takes several months for blood test results from HIV negative to HIV positive. In this case data set contains lots of tied value observations even though the underlying life-time distribution is continuous. This type of data set is called as the grouped data and can be analyzed by the data-specific method. Heitjan (1989) reviewed extensively the methodology and suggested several research directions. For the right censored data, Prentice and Gloeckler (1978) considered the inferences about $\beta$ under the PHM. Park (1993) proposed a class of nonparametric tests for the linear model whereas Neuhaus (1993) modified the so-called log-rank tests for the grouped data. In this study, we consider to propose a nonparametric test procedure for $\beta$ under the AHM (1.1) using the score function based on the likelihood principle for the grouped and right censored data. The scores will be derived using the discrete model approach (cf. Kalbfleisch and Prentice, 1980). First of all, we consider a simple score test statistic for the scalar case and then extend this procedure to the vector covariate. Then we illustrate our test with an example and compare our procedure with other one. Finally we discuss some interesting features about our test procedure.

2. A Simple Score Test

Suppose that we observe life time $T_i$ for the $i^{th}$ individual with some specific constant covariate, $z_i$, $i = 1, \ldots, n$. We assume that each subject is prone to be censored. In this way, the data set can be represented as $\{(T_i, \delta_i, z_i), i = 1, \ldots, n\}$, where $\delta_i$ stands for the censoring status with values 0 or 1 if censored or not. Since we are concerned with the grouped data, we assume that the positive half real line, $[0, \infty)$ is partitioned into $k$ number of sub-intervals such as $[0, \infty) = \bigcup_{i=1}^{k} [a_{i-1}, a_i)$, with $a_0 = 0$ and $a_k = \infty$. Then one can only have the information that $T_i$ is contained in one of the $k$ sub-intervals for all $i$. We denote $D_i$ and $C_i$ as the indicate sets for the uncensored and censored observations in the $l^{th}$ sub-interval $[a_{l-1}, a_l)$, respectively. Also we denote $R_l$ as the risk set of the $l^{th}$ sub-interval.

Finally we denote $d_l$ and $r_l$ as the sizes of $D_l$ and $R_l$, respectively, $l = 1, \ldots, k$. In this grouped continuous data, we assume that all the censorings occur at the end of a sub-interval and all the deaths proceed any censoring in the same sub-interval. Also we will assume that all the observations in the last sub-interval $[a_{k-1}, \infty)$ are censored at $a_{k-1}$ for some technical reason. Finally we assume that the survival function and censoring distribution function are independent to avoid the so-called identifiability problem. Then from the discrete model in Kalbfleisch and Prentice (1980) with all the assumptions and notation introduced up to now, we have with (1.2) that for $l = 1, \ldots, k - 1$,

$$\Pr \{T_i \in [a_{l-1}, a_l), \delta_i = 1, z_i \} \propto \exp[-\Lambda_0(a_{l-1})] \exp[-a_{l-1} \beta z_i] - \exp[-\Lambda_0(a_l)] \exp[-a_l \beta z_i]$$

and

$$\Pr \{T_i \in [a_{l-1}, a_l), \delta_i = 0, z_i \} \propto \exp[-\Lambda_0(a_l)] \exp[-a_l \beta z_i].$$

For $l = k$, we have that

$$\Pr \{T_i \in [a_{k-1}, \infty), \delta_i = 0, z_i \} \propto \exp[-\Lambda_0(a_{k-1})] \exp[-a_{k-1} \beta z_i].$$
Also we note that
\[
\exp[-\Lambda_0(a_{i-1})] \exp[-a_i \beta z_i] - \exp[-\Lambda_0(a_i)] \exp[-a_i \beta z_i] \\
= \left[ \exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})] \exp[(a_i - a_{i-1}) \beta z_i] - 1 \right] \exp[-\Lambda_0(a_i)] \exp[-a_i \beta z_i].
\]
Then under the AHM (1.1), the likelihood function for the discrete model becomes as
\[
L(\beta) = \prod_{t=1}^{k-1} \left[ \exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})] \exp[(a_i - a_{i-1}) \beta z_i] - 1 \right] \prod_{t=1}^{k-1} \left[ \exp[-\Lambda_0(a_i)] \exp[-a_i \beta z_i] \prod_{i \in C_i} \exp[-\Lambda_0(a_{k-1})] \exp[-a_{k-1} \beta z_i] \right] \times C(I),
\]
where \(C(I)\) denotes the portion of \(L(\beta)\) contributed by the censored observations. We assume that \(C(I)\) contains no information about \(\beta\) (i.e., non-informative censoring). Then by taking logarithm to \(L(\beta)\)
and differentiating the log-likelihood function \(l(\beta)\) with respect to \(\beta\), we have that
\[
\frac{\partial l(\beta)}{\partial \beta} = \sum_{i=1}^{k-1} \sum_{i \in D_i} \frac{\exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})] \exp[(a_i - a_{i-1}) \beta z_i] - \exp[\Lambda_0(a_i)] \exp[(a_i - a_{i-1}) \beta z_i] - 1}{\exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})] \exp[(a_i - a_{i-1}) \beta z_i] - 1} \sum_{i \in D_i}a_i z_i + \sum_{i \in C_i} a_{i} z_i.
\]
By substituting 0 for \(\beta\) in \(\frac{\partial l(\beta)}{\partial \beta}\), we have that
\[
W_n^0 = \sum_{i=1}^{k-1} \left\{ \sum_{i \in D_i} \frac{\exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})]}{\exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})] - 1} (a_i - a_{i-1}) z_i - \sum_{i \in D_i \cup C_i} a_i z_i \right\} + \sum_{i \in C_i} a_{i} z_i.
\]
Then one may use \(W_n^0\) for testing \(H_0: \beta = 0\) if the baseline hazard function \(\Lambda_0\) were fully known.
Then the resulting test would be optimal in the local sense. However since we have assumed that the baseline hazard function \(\Lambda_0\) is unknown, we consider to use a suitable estimate for \(\Lambda_0\) or \(\Lambda_0\). For this matter, first of all, we note that since under \(H_0: \beta = 0\),
\[
S(\tau) = \exp[-\Lambda_0(\tau)]
\]
we have that under \(H_0: \beta = 0\),
\[
\frac{\exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})]}{\exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})] - 1} = \frac{\exp[-\Lambda_0(a_{i-1})]}{\exp[-\Lambda_0(a_{i-1})] - \exp[-\Lambda_0(a_i)]} = \frac{S(a_{i-1})}{S(a_{i-1}) - S(a_i)}.
\]
Also we note that from the assumption for the relation between the censoring and death observations in the same sub-interval, the Kaplan-Meier estimate \(\hat{S}(a_i)\) of \(S(a_i)\) under \(H_0: \beta = 0\) is of the form
\[
\hat{S}(a_i) = \prod_{j=1}^{i} \left( 1 - \frac{d_j}{r_j} \right),
\]
where \(d_j\) and \(r_j\) are the sizes of \(D_j\) and \(R_j\) of the sub-interval \([a_{j-1}, a_j]\), \(j = 1, \ldots, k-1\). Then under \(H_0: \beta = 0\),
\[
\frac{\exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})]}{\exp[\Lambda_0(a_i) - \Lambda_0(a_{i-1})] - 1}
\]
can be consistently estimated by

$$\frac{\hat{S}(a_{l-1})}{\hat{S}(a_{l-1}) - \hat{S}(a_l)} = \frac{r_l}{d_i}$$  \hspace{1cm} (2.1)

Also we note that for each $l$, $l = 1, \ldots, k - 1$

$$a_l = (a_l - a_{l-1}) + (a_{l-1} - a_{l-2}) + \cdots + (a_1 - a_0) = \sum_{j=1}^{l}(a_j - a_{j-1}).$$

Therefore we see that

$$\sum_{l=1}^{k-1} \sum_{i \in D_l \cup C_i} a_l z_i + \sum_{i \in C_k} a_{k-1} z_i = \sum_{l=1}^{k-1} a_l \sum_{i \in D_l \cup C_l} z_i + a_{k-1} \sum_{i \in C_k} z_i = \sum_{l=1}^{k-1} \sum_{j=1}^{l} (a_j - a_{j-1}) \sum_{i \in D_j \cup C_l} z_i + \sum_{j=1}^{k-1} (a_j - a_{j-1}) \sum_{i \in C_k} z_i = \sum_{l=1}^{k-1} (a_l - a_{l-1}) \sum_{i \in R_l} z_i.$$  \hspace{1cm} (2.2)

Then from (2.1) and (2.2), we see that $W_n^0$ can be modified as

$$W_n = \sum_{l=1}^{k-1} \left\{ (a_l - a_{l-1}) \frac{r_l}{d_l} \sum_{i \in D_l} z_i - (a_l - a_{l-1}) \sum_{i \in R_l} z_i \right\}$$  \hspace{1cm} (2.3)

$$= \sum_{l=1}^{k-1} (a_l - a_{l-1}) \frac{r_l}{d_l} \left\{ \sum_{i \in D_l} z_i - \frac{d_l}{r_l} \sum_{i \in R_l} z_i \right\}.$$  \hspace{1cm} (2.4)

We note that under $H_0 : \beta = 0$, $W_n$ is a martingale with discrete compensators (cf. Fleming and Harrington, 1991). One may confirm this by re-expressing $W_n$ in (2.3) as a stochastic integral with identifying $w = (a_l - a_{l-1}) r_l/d_l$ in the Equation (4) of Jones and Crowley (1990). Therefore the expectation of $W_n$ is 0 under $H_0 : \beta = 0$.

Then for testing $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$, one may reject $H_0 : \beta = 0$ for large values of $|W_n|$. For any given significant level, in order to decide the critical value, we need the distribution of $W_n$ under $H_0 : \beta = 0$. However the derivation of the exact distribution of $W_n$ would be difficult because of the involvement of censoring distribution into the distribution of $W_n$ even under $H_0 : \beta = 0$. Therefore it is natural to consider the null distribution of $W_n$ in an asymptotic manner. In the following theorem, we state the asymptotic normality for $W_n$. One may find the proof in Jones and Crowley (1990) and Fleming and Harrington (1991), whose proofs use the martingale central limit theorem based on the counting process theory. Before stating the theorem, we provide a consistent estimate of the variance of $W_n$ (cf. Jones and Crowley, 1990) under $H_0 : \beta = 0$ in the following:

$$\delta_n^2 = \sum_{l=1}^{k-1} (a_l - a_{l-1})^2 \frac{r_l(r_l - d_l)}{(r_l - 1)^2 d_l^2} \left\{ \sum_{i \in R_l} (z_i - \bar{z}_l)^2 \right\},$$

where $\bar{z}_l = (1/r_l) \sum_{i \in R_l} z_i$. Also we note that $\delta_n^2$ is known to be unbiased (cf. Jones and Crowley, 1990).
Theorem 1. Under all the assumptions used up to now and with the following condition that
\[
\max \frac{1}{\sqrt{n}} \{z_1, \ldots, z_n\} \to 0, \tag{2.4}
\]
we have that under \(H_0: \beta = 0\)
\[
\frac{W_n}{\sqrt{\sigma_n^2}}
\]
converges in distribution to a standard normal distribution as \(n \to \infty\).

We note that the condition (2.4) is called Lindeberg-type condition (cf. Andersen and Gill, 1982) and is equivalent to the Noether’s condition (cf. Randles and Wolfe, 1979). When there is at most one uncensored observation in each sub-interval, we note that \(W_n\) becomes
\[
W_n = \sum_{l=1}^{k-1} (a_l - a_{l-1}) r_l \left( z_l - \frac{1}{r_l} \sum_{i \in R_l} z_i \right).
\]
Also we note that when the lengths of sub-intervals \([a_{l-1}, a_l)\) are all equal for all \(l, l = 1, \ldots, k - 1\), then the quantity \(a_l - a_{l-1}\) becomes a constant and so can be removed from the expression in \(W_n\) such as
\[
W_n = \sum_{l=1}^{k-1} \frac{r_l}{d_l} \left( \sum_{i \in D_l} z_i - \frac{d_l}{r_l} \sum_{i \in R_l} z_i \right). \tag{2.5}
\]
Especially, when each covariate \(z_i\) takes values only 0 or 1 as the indices of the populations for the two-sample problem, \(W_n\) has been called as a generalized (or weighted) log-rank statistic.

3. Vector Covariate Case
We now consider the extension to the \(p \times 1\) covariate vector case, \(p \geq 2\). Then for the \(i^{th}\) individual, the \(p \times 1\) covariate vector may be denoted as \(z_i = (z_{i1}, \ldots, z_{ip})\), \(i = 1, \ldots, n\). Also \(\beta = (\beta_1, \ldots, \beta_p)'\) denotes the corresponding regression coefficient vector. Then for the model (1.1), using the relation (1.2) with the same arguments for the scalar case, the likelihood function can be expressed as
\[
L(\beta) = \prod_{i=1}^{k-1} \prod_{i \in D_l} [\exp(\Lambda_0(a_i) - \Lambda_0(a_{i-1})] \exp((a_i - a_{i-1}) \beta z_i) - 1]
\]
\[
\prod_{i=1}^{k-1} \prod_{i \in D_l \cup C_l} [\exp(-\Lambda_0(a_i)] \exp(-a_i \beta z_i) \prod_{i \in C_k} [\exp(-\Lambda_0(a_{k-1})] \exp(-a_{k-1} \beta z_i) \times C(I),
\]
where \(C(I)\) is the portion of \(L(\beta)\) contributed by censoring. Also we assume the non-informative censoring scheme. Then for each \(j, j = 1, \ldots, p\), by differentiating partially the log-likelihood function, \(l(\beta)\), with respect to \(\beta_j\) and manipulating \(\partial l(\beta)/\partial \beta_j\) with the same arguments for the scalar case, one may obtain the following score statistic \(W_{jn}\):
\[
W_{jn} = \sum_{l=1}^{k-1} (a_l - a_{l-1}) \frac{r_l}{d_l} \left( \sum_{i \in D_l} z_{ij} - \frac{d_l}{r_l} \sum_{i \in R_l} z_{ij} \right).
\]
Then we note that for each for \( j, j = 1, \ldots, p \), \( W_{jn} \) is a martingale with discrete compensator under \( H_0 : \beta = 0 \). Therefore \( W_{jn} \) can be used as a test statistic for testing \( H_0 : \beta_j = 0 \). This fact in turn, suggests that we may consider a quadratic form based on \((W_{1n}, \ldots, W_{pn})'\) for a test statistic for testing \( H_0 : \beta = 0 \). To this end, we need a null consistent estimate, \( \hat{V}_n = (\hat{\sigma}_{jj'}n)_{j,j'=1,\ldots,p} \), of the covariance matrix of \((W_{1n}, \ldots, W_{pn})'\). In the sequel, let \( \tilde{z}_{ij} = (1/r_{ij}) \sum_{l \in R_i} z_{ij}, l = 1, \ldots, k \) and \( j = 1, \ldots, p \). Then from the previous section, it is obvious that for each \( j, j = 1, \ldots, p \), \( W_{jn} \), a consistent and unbiased null variance estimate \( \hat{\sigma}_{jjn}^2 = \hat{\sigma}_{jjn} \) for \( W_{jn} \) is

\[
\hat{\sigma}_{jjn}^2 = \hat{\sigma}_{jjn} = \sum_{l=1}^{k-1} (a_l - a_{l-1})^2 \frac{r_l(r_l - d_l)}{(r_l - 1)^2 d_l^2} \left\{ \sum_{l \in R_i} \left( \bar{z}_{ij} - \bar{z}_{ij} \right)^2 \right\}.
\]

Also a null covariance estimate \( \hat{\sigma}_{jj'n} \) of the covariance between \( W_{jn} \) and \( W_{jn'} \) for \( j \neq j' \) can be obtained by the same arguments used for the null variance estimate by noticing that the covariance between observations with \( z_{ij} \) and \( z_{i'j'} \) is 0 whenever \( i \neq i' \). Thus \( \hat{\sigma}_{jj'n} \) becomes of the form

\[
\hat{\sigma}_{jj'n} = \sum_{l=1}^{k-1} (a_l - a_{l-1})^2 \frac{r_l(r_l - d_l)}{(r_l - 1)^2 d_l^2} \left\{ \sum_{l \in R_i} \left( \bar{z}_{ij} - \bar{z}_{ij} \right) \left( \bar{z}_{i'j'} - \bar{z}_{i'j'} \right) \right\}.
\]

We note that \( \hat{\sigma}_{jj'n} \) is also a consistent estimate. Then with the assumption that \( \hat{V}_n \) is nonsingular, one may propose the following quadratic form for a test statistic for testing \( H_0 : \beta = 0 \)

\[
Q_n = \begin{pmatrix} W_{1n} \\ \vdots \\ W_{pn} \end{pmatrix}' \hat{V}_n^{-1} \begin{pmatrix} W_{1n} \\ \vdots \\ W_{pn} \end{pmatrix},
\]

where \( \hat{V}_n^{-1} \) is the inverse of \( \hat{V}_n \). Then one may reject \( H_0 : \beta = 0 \) in favor of \( H_1 : \beta \neq 0 \) for large values of \( Q_n \). Also in order to have critical value for any given significance level, we need the null distribution of \( Q_n \). Since the null distribution of \( Q_n \) contains the unknown censoring distribution, also we consider to obtain the limiting distribution of \( Q_n \) as for the scalar covariate case. Then with all the notations introduced up to now, we state the following main result.

**Theorem 2.** With the assumption that \( \hat{V}_n \) is nonsingular and the condition that

\[
\max \frac{1}{\sqrt{n}} \left\{ z_{ij}, \ldots, z_{nj} \right\} \to 0, \tag{3.1}
\]

for each \( j, j = 1, \ldots, p \), under \( H_0 : \beta = 0 \), \( Q_n \) converges to a central chi-square distribution with \( p \) degrees of freedom.

**Proof:** From Theorem 1, the Cramer-Wold device (cf. Billingsley, 1986) and the Slutsky's theorem with the fact that \( \hat{V}_n \) is a nonsingular consistent estimate under \( H_0 : \beta = 0 \), the result follows easily. \( \square \)

When \( \hat{V}_n \) is singular, \( i.e., \lvert \hat{V}_n \rvert = 0 \), Wei and Lachin (1984) recommended to add some number \( b_n \) such that \( b_n = o(n^{-1}) \) to each \( \hat{\sigma}_{jjn}^2, j = 1, \ldots, p \), where \( b_n = o(n^{-1}) \) means that \( nb_n \to 0 \) as \( n \to \infty \).
4. An Example and Simulation Results

In order to illustrate our test procedure, we consider the data reported by Embury et al. (1977) for the length of remission (in weeks) for the two groups (maintenance chemotherapy and control) with acute myelogenous leukemia patients. Since the length of remission for each patient was measured by week, the data set contains several tied observations. Therefore a sub-interval may be designated by each week. Then we note that the lengths of sub-intervals are all the same with unity. Thus we may use the statistic (2.5) rather than (2.3) for this problem with the corresponding variance estimate. The objective of the experiment was to see if the maintenance chemotherapy prolongs the length of remission. The data has been summarized as follows:

Control group: 5, 5, 8, 8, 12, 16+, 23, 27, 30, 33, 43, 45

Maintenance group: 9, 13, 13+, 18, 23, 28+, 31, 34, 45+, 48, 161+

where + indicates censored observation. We note that this is a two-sample problem. Therefore by allocating 0 or 1 to covariate $z_i$ for the $i$th individual according as from the control or maintenance chemotherapy group in (2.5), we obtain the following necessary quantities.

\[ W_n = 27.5 \quad \text{and} \quad \hat{\sigma}_n^2 = 253.5183. \]

Thus we have that

\[ \frac{W_n}{\hat{\sigma}_n^2} = 1.73. \]

The corresponding $p$-value is 0.042, which shows the strong evidence against $H_0 : \beta = 0$ in favor of $H_1 : \beta \neq 0$. In passing, we note that the procedure proposed by Prentice and Gloeckler (1978) gives 0.065 as its $p$-value.

The following table is the results of the simulation study, which are the empirical powers. In this study, we considered two tests: one is the proposed test (AHM) and the other, the test considered by Prentice and Gloeckler (1978) (PHM) under the proportional hazards model. For the survival function we consider the location-shift exponential distribution with the scale parameter $\lambda = 1$ such that for some $\beta > 0$

\[ f_\beta(t; \beta) = \begin{cases} 
\exp[-(t + \beta)], & \text{for } \beta \leq t < \infty, \\
0, & \text{otherwise.}
\end{cases} \tag{4.1} \]

We note that the survival function of (4.1) is $S(t; \beta) = \exp[-(t + \beta)]$ for $\beta < t$. Thus this can be considered to corresponds to the model (1.2) by varying the value of $\beta$. For the censoring distribution, we consider also the exponential distribution such that

\[ g(t) = \begin{cases} 
\left( \frac{1}{2} \right) \exp \left( -\frac{t}{2} \right), & \text{for } 0 < t < \infty, \\
0, & \text{otherwise.}
\end{cases} \]

We choose $\lambda = 1/2$ for the censoring distribution in order to avoid excessive censoring. The sample sizes are 20 for each sample and we vary the value of $\beta$ from 0 to 0.5 by 0.1 for the first sample while fixed as 0 for the second sample. Also we choose a partition of $[0, \infty)$ for grouping as $[0, 0.2), \ldots, [1.8, 2.0), [2.0, \infty)$, i.e., 11 sub-intervals. For each case, we obtained empirical power based on 1000 simulations. The simulations have been carried out by SAS/IML on PC version and the nominal significance level is 0.05. We note that for this case our proposed test shows better performance than that of Kallbiefisch and Prentice (1980).
### Table 1: Empirical powers based on simulation

<table>
<thead>
<tr>
<th></th>
<th>AHM</th>
<th>PHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.078</td>
<td>0.040</td>
</tr>
<tr>
<td>0.1</td>
<td>0.092</td>
<td>0.062</td>
</tr>
<tr>
<td>0.2</td>
<td>0.141</td>
<td>0.093</td>
</tr>
<tr>
<td>0.3</td>
<td>0.202</td>
<td>0.161</td>
</tr>
<tr>
<td>0.4</td>
<td>0.298</td>
<td>0.211</td>
</tr>
<tr>
<td>0.5</td>
<td>0.393</td>
<td>0.278</td>
</tr>
</tbody>
</table>

### 5. Some Concluding Remarks

Already we stated that the AHM (1.1) has not been so popular since it is not feasible to apply the conditional likelihood. In spite of this inconvenience, it would be worthwhile to have a statistical methodology for (1.1) in case that the PHM might be dubious. Especially, we note that the expression (4.1) under the AHM corresponds to the location translation model when the baseline hazard function is constant or the underlying distribution is exponential. In this case, the application of the log-rank test, which is optimal for the PHM would incur some loss of efficiency. Also we note that since we applied the likelihood principle for the derivation of the test statistic, the resulting test would be optimal in the light of power when (1.1) holds.

In Section 2, we assumed that all the observations in the last sub-interval \([a_{n-1}, \infty)\) are censored at \(a_{n-1}\), which is the beginning point of the last sub-interval. The reason for this is as follows. First of all, we note that the length of the last sub-interval is infinity. If there is any uncensored observation in the last sub-interval, then the length of the last sub-interval should be included in \(W_n\), which is an absurd expression. Also if we maintain the assumption that the censoring occurs at the end of each sub-interval, then the derivation of (2.2) becomes impossible for the censored observations in the last sub-interval. However in the real experiment, since always a researcher observes the objects during a finite time period, such an assumption becomes insignificant and cannot be applied for the real world.

For the null distribution, we derived the asymptotic normality using the large sample approximation. Also one may consider a re-sampling approach such as the permutation principle (cf. Good, 2000) to obtain a null distribution. Park (1993) and Neuhaus (1993) considered to apply the permutation principle for obtaining the null distribution of the test statistics for the right censored and grouped data. However if one applies the permutation principle for the censored data, then one must include the equality of unknown censoring distributions, which are of nuisance, in the null hypothesis. The resulting permutation test is known as exact but conditional. Also as another re-sampling method, one may use the bootstrap method (cf. Efron and Tibshirani, 1993). For the censored data, you may refer to Efron (1981) and Reid (1981). Unlike the permutation principle, the bootstrap method does not require the equality among censoring distributions for the null hypothesis. However because of the computational amount of work, the application of the re-sampling methods always take the Monte-Carlo approach.

We note that when there is at most one uncensored observation in each sub-interval, then this corresponds to the no tied-value case and the assumption for the allowance of discontinuity of hazard function disappears. Also in this research, only we considered the case that the covariate is independent of time. For the time-dependent case, the likelihood function would not be tractable because of the involvement of time into the cumulative covariate function such as \(Z(t) = \int_0^t z(x)dx\), which in turn requires some specific functional form of \(z(t)\). However in the light of applicability, this research should be done in the near future.
References


Received January 2009; Accepted January 2009