A Crame^^ ´ r-von Mises Type Test of the Proportional Hazards Assumption

Yamaguchi, Kazunori
Department of Information Systems, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

https://doi.org/10.15017/17714

出版情報：九州大学大学院総合理工学報告. 10 (1), pp.67-70, 1988-06-30. Interdisciplinary Graduate School of Engineering Sciences, Kyushu University
バージョン：

A Cramér-von Mises Type Test of the Proportional Hazards Assumption

Kazunori YAMAGUCHI*

(Received Fed. 29, 1988).

The proportional hazards model has been widely applied for censored observations in the fields of survival analysis and reliability theory, which was proposed by D. R. Cox. In the practical application, it is essentially important for the mathematical model to have an assumption of proportionality of hazards. The purpose of the present paper is to propose a Cramér-von Mises type test of the proportional hazards assumption for randomly censored data.

1. Introduction

The proportional hazard model is popular and widely available to survival analysis with censored data. An important problem arising in applying this model is to assess the proportionality of the hazards. However, the analitical results are mostly discussed on the existence of the proportionality. In fact, Andersen [1] has mentioned that a number of worked examples of analyses of survival data using this model have been published, but surprisingly little attention has been paid to the problem of model checking. In view of such a necessity of model checking, we propose a test of the proportionality of the hazards.

The observations from sample i $(i=1, 2)$ are (X_{ij}, d_{ij}) $j=1,\cdots, n_i$, where $X_{ij}=\min\{X^0_{ij}, U_{ij}\}$, $d_{ij}=I(X_{ij}=X^0_{ij})$, $I(\cdot)$ is the indicator function, X^0_{ij} is the true survival time of j-individual of i-sample with a continuous distribution function F_i, and U_{ij} is the corresponding censoring variable. We assume that U_{ij} are independent identically distributed random variables with distribution function L_i, and that the variables U_{ij} are independent of the variables X_{ij}.

Furthermore, we assume that $n_i/n \to r_i$, $0<r_i<1$, as $n\to\infty$, where $n=n_1+n_2$. Let the cumulative hazard function of F_i be denoted by $H_i(t)$, that is $H_i(t)=-\log(1-F_i(t))$. Also let

$$N_i(t)=\#\{j;\ x_{ij}\leqq t \text{ and } d_{ij}=1\},$$

$$Y_i(t)=\#\{j;\ x_{ij}\leqq t\},$$

$$y_i(t)=(1-F_i(t))(1-L_i(t-)).$$

We assume that the supports of $y_i(t)$ are $[0, \infty]$.

We intend to test the null hypothesis that $H_1(t)=\theta H_2(t)$, for $t\geqq 0$ and some unknown constant θ.

*Department of Information System. Graduate Student

Now we can write the logarithm of the Cox's partial likelihood as follows,

$$\int_0^\infty \log\theta\, dN_1(s) - \int_0^\infty \log\{Y_1(s)\theta + Y_2(s)\}\, d\{N_1(s)+N_2(s)\},$$

and let

$$C_n(\theta;t) = \int_0^t \log\theta\, dN_1(s) - \int_0^t \log\{Y_1(s)\theta + Y_2(s)\}\, d\{N_1(s)+N_2(s)\}.$$

The estimator $\hat{\theta}$ is defined as the solution to the likelihood equation $(\partial/\partial\theta)\ C_n(\theta;\infty) = 0$. The asymptotic properties of this estimator have been studied by many workers, for example Tsiatis [4] and Wong [6].

Furthermore let $B_n(\theta;t)$ be the product of the parameter θ and the derivative of $C_n(\theta;t)$ with respect to θ. Then,

$$B_n(\theta;t) = \int_0^t dN_1(s) - \int_0^t Y_1(s)\theta/\{Y_1(s)\theta + Y_2(s)\}\, d\{N_1(s)+N_2(s)\}.$$

This can be interpreted as the residual at the time t. The first term of left hand side is the observed number of deaths from sample 1 until time t. The second term is the corresponding expected number under the null hypothesis.

Because the parameter θ is unknown, we replace θ by $\hat{\theta}$ in B_n and let,

$$W_n(t) = n^{-1/2} B_n(\hat{\theta}; t).$$

It is clear by definition that $W_n(0) = W_n(\infty) = 0$.

In the next section, we propose a test by using $W_n(t)$.

2. A newly proposed test

At first we give two lemmas.

Lemma 1

Under $H_1(t) = \theta_0 H_2(t)$ for any $t \geqq 0$ and some θ_0, the estimator $\hat{\theta}$ is strongly consistent.

(proof)

$B_n(\theta;t)$ is clearly a non-increasing function of θ. For fixed θ, $1/n\ B_n(\theta;t) \to B(\theta;\infty)$ a. s., where,

$$B(\theta;t) = \int_0^t \frac{r_1 y_1(s)\ r_2 y_2(s)}{\theta r_1 y_1(s) + r_2 y_2(s)}\, d\{H_1(s) - \theta H_2(s)\}.$$

If $H_1(t) = \theta_0 H_2(t)$ for any $t \geqq 0$, then $B(\theta_0;\infty) = 0$. $B(\theta;t)$ is also a strictly decreasing function on neighborhood of θ_0. Therefore, $Pr\ (lim\ \hat{\theta} = \theta_0) = 1$.

Let $V_n(t)=\{\hat{\theta}Q_n(\infty)\}^{-1/2}W_n(t)$,

$$Q_n(t)=n^{-1}\int_0^t \frac{Y_1(s)\,Y_2(s)}{\{\hat{\theta}Y_1(S)+Y_2(s)\}^2}\,d\{N_1(s)\,N_2(s)\},$$

$$Q(t)=r_1 r_2\int_0^t \frac{y_1(s)\,y_2(s)}{\theta r_1 y_1(s)+r_2 y_2(s)}\,dH_2(s),$$

$G_n(t)=Q_n(t)/Q_n(\infty)$ *and*

$G(t)/Q(t)/Q(\infty)$.

Lemma 2

Under $H_1(t)=\theta_0 H_2(t)$ for any $t>0$ and some θ_0, the process $\{V_n(t);\,0\leqq t\leqq\infty\}$ converges in law to the process $\{W^0(G(t));\,0\leqq t\leqq\infty\}$, where $W^0(\cdot)$ is the tied-down Brownian motion process.

(proof) See Theorem 1 of Wei [5].

Let $T_n=\int_0^\infty V_n(s)^2\,dG_n(s)$. This is our proposed test statistic with the properties which are represented by following two theorems.

The above two lemmas yield the following theorem.

Theorem 1

Under the null hypothesis, the asymptotic distribution of T_n is same to the distribution of $\int_0^1 W^0(t)^2\,dt$.

The consistent property of this test is obtained by the next theorem.

Theorem 2

When the null hypothesis is not hold,

for any $c>0$, $\lim_{n\to\infty} Pr\,\{T_n>c\}=1$.

(proof)

From the proof of lemma 1, the estimator $\hat{\theta}$ converges to a constant θ_A, even if the assumption of proportionality of hazards is violated. On the other hand, for fixed t,

$$n^{-1/2}W_n(t)\underset{\text{a. s.}}{\rightarrow}\int_0^t \frac{r_1 y_1(s)\,r_2 y_2(s)}{\theta_a r_1 y_1(s)+r_2 y_2(s)}\,d\{H_1(s)-\theta_A H_2(s)\}.$$

Clearly there exists some t^* such that

$$\int_0^{t^*} \frac{r_1 y_1(s)\, r_2 y_2(s)}{\theta_A r_1 y_1(s) + r_2 y_2(s)}\, d\{H_1(s) - \theta_A H_2(s)\} \neq 0.$$

Therefore $W_n(t^*) \to \infty$ in P for some t^* and $T_n \to \infty$ in P.

3. Remarks

The tables of the critical points of the distribution of $\int_0^1 W^0(t)^2\, dt$ were given in some papers, for example, Schumacher [3].

A general form of the integral-type test statistics of this problem can be represented as follows,

$T_{n,\psi} = \int_0^\infty \psi(s)\, V_n(s)^2\, dG_n(s)$, where ψ is a weight function. We can obtain various test statistics by selection of the weight function ψ. For example, when the estimated variance function of $V_n(t)$ is selected as $\psi(t)^{-1}$, the test statistic $T_{n,\psi}$ is an Anderson-Darling type test statistic. The asymptotic distributions of these statistics were studied in Durbin [2].

References

1) Andersen, P. K. (1982) Testing goodness-of-fit of Cox's regression model. Biometrics 38, 67−77.

2) Durbin, J. (1973) Distribution theory for tests based on the sample distribution function. CBMS-NSF 9.

3) Schumacher, M. (1984) Two-sample tests of Cramér-von Mises and Kolmogorov-Snirnov type for randomly censored data. Int. Stati. Rev. 52, 263−281.

4) Tsiatis, A. A. (1981) A large sample study of Cox's regression model. Annals of statistics 9, 93−108.

5) Wei, L. J. (1984) Testing goodness of fit for proportional hazard model with censored observations. JASA 79, 649−652.

6) Wong, W. H. (1986) Theory of partial likelihood. Annals of statistics 14, 88−123.