Journal of the Korean
Data & Information Science Society
1999, Vol. 10, No. 1, pp. 103 ∼ 110

Asymptotic Relative Efficiencies of the Nonparametric Relative Risk Estimators for the Two Sample Proportional Hazard Model

Kil Ho Cho [1] · In Suk Lee [2] · Jeen Kap Choi [3] · Seong Hwa Jeong [4]
Dal Woo Choi [5]

Abstract

In this paper, we summarize some relative risk estimators under the two sample model with proportional hazard and examine the relative efficiencies of the nonparametric estimators relative to the maximum likelihood estimator of a parametric survival function under random censoring model by comparing their asymptotic variances.

Key Words and Phrases: Nonparametric estimator, Relative efficiency, Relative risk.

1. Introduction

We consider the problem of the estimation of the proportional hazard ratio for clinical trial situations. In the two sample problem, the proportional hazard model is specified that

$$\lambda_2(t) = \theta\lambda_1(t),$$

where θ is called relative risk. Thus we can see that the ratio of hazard function has the interpretation of relative risk and has an intuitive appeal as a descriptive statistics.

[1]Professor, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.
[2]Professor, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.
[3]Professor, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.
[4]Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.
[5]Lecturer, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.

In the proportional hazard model, if we know the lifetime and censoring distributions, then the relative risk is easily estimated by maximum likelihood methods. But if the distributions are not specified, then it can be obtained only one by nonparametric methods. The nonparametric proportional hazard model was derived by Cox(1972) for the analysis of survival data. In the Cox's proportional hazard model, if $p = 1$ and $\mathbf{z}$ is the indicator function for treatment group, then the model is reduced to

$$\lambda_2(t) = e^{\beta}\lambda_1(t),$$

where $\theta = e^{\beta}$ is called the relative risk. Cox(1975) proposed the partial likelihood method that can be used for inference about relative risk, and Efron(1977) showed that maximum partial likelihood is asymptotically efficient.

On the other hand, Begun and Reid(1983) and Andersen(1983) proposed the generalized rank estimator of relative risk, and proved that each test in these papers is equivalent to a consistent and asymptotically normally distributed estimator of the hazard ratio in the two sample model with proportional hazard.

In this paper, we summarize some relative risk estimators under the two sample model with proportional hazard and examine the asymptotic efficiencies of the nonparametric estimators relative to the maximum likelihood estimator of a parametric survival function under random censoring model by comparing their asymptotic variances.

2. Estimation of Relative Risk

Let $T_{11}, T_{12}, \cdots, T_{1n_1}$ and $T_{21}, T_{22}, \cdots, T_{2n_2}$ be nonnegative random lifetimes with absolutely continuous distribution function F_1 and F_2, respectively, and let $C_{11}, C_{12}, \cdots, C_{1n_1}$ and $C_{21}, C_{22}, \cdots, C_{2n_2}$ be nonnegative random censoring times with absolutely continuous distribution function G_1 and G_2, respectively. By random censorship model, the true lifetimes T_{ij}'s are censored on the right by the censoring times C_{ij}'s, so that we only observe

$$(X_{ij}, \epsilon_{ij}), \quad i = 1, 2, \quad j = 1, 2, \cdots, n_i,$$

where

$$X_{ij} = \min(T_{ij}, C_{ij}), \quad \epsilon_{ij} = I(T_{ij} \le C_{ij}),$$

and the censoring times C_{ij} are assumed to be independent of the lifetimes T_{ij}.

Let $n = n_1 + n_2$ and $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_n)$ be the combined ordered survival times, without regard to censoring. Then we define the sample indicators of the corresponding $\mathbf{Y}$,

$$\mathbf{Z} = (Z_1, Z_2, \cdots, Z_n),$$

where

$$Z_j = \begin{cases} 1 & \text{if} \quad Y_j \text{ from sample 1,} \\ 0 & \text{if} \quad Y_j \text{ from sample 2,} \end{cases}$$

and a vector of censoring indicators

$$\delta = (\delta_1, \delta_2, \cdots, \delta_n),$$

where

$$\delta_j = \begin{cases} 1 & \text{if} \quad Y_j \text{ is uncensored observation,} \\ 0 & \text{if} \quad Y_j \text{ is censored observation.} \end{cases}$$

Further, we also define the number at risk in the each sample at Y_j by

$$n_{1j} = \sum_{k=j}^{n}(1 - Z_k), \quad n_{2j} = \sum_{k=j}^{n} Z_k.$$

Following sections deal with the several estimators of relative risk.

2.1 Maximum Likelihood Estimator

Let $T_{1\cdot}$ and $T_{2\cdot}$ be independent identically distributed (i.i.d.) exponential random variables with parameter α_1 and α_2, respectively, and let $C_{1\cdot}$ and $C_{2\cdot}$ be i.i.d. exponential random variables with parameter α_1' and α_2', respectively. Then the likelihood is proportional to

$$L \propto \alpha_i^{n_{u_i}} exp\left[-\alpha_i \sum_{j=1}^{n_i} X_{ij}\right]$$

where n_{u_i} is the number of uncensored observations in each sample. Therefore, the maximum likelihood estimator (MLE) of α_i is

$$\widehat{\alpha}_i = \frac{n_{u_i}}{\sum_{j=1}^{n_i} X_{ij}}, \quad i = 1, 2, \tag{1}$$

and the equation (1) can be written in terms of notations of the combined sample ;

$$\widehat{\alpha}_1 = \frac{\sum_{j=1}^{n} \delta_j (1 - Z_j)}{\sum_{j=1}^{n} Y_j (1 - Z_j)}, \quad \widehat{\alpha}_2 = \frac{\sum_{j=1}^{n} \delta_j Z_j}{\sum_{j=1}^{n} Y_j Z_j}.$$

Since α_i is the hazard function of the exponential distribution, the MLE of θ is given by

$$\widehat{\theta}_{MLE} = \frac{\left(\sum_{j=1}^{n} \delta_j Z_j\right) / \left(\sum_{j=1}^{n} Y_j Z_j\right)}{\left(\sum_{j=1}^{n} \delta_j (1 - Z_j)\right) / \left(\sum_{j=1}^{n} Y_j (1 - Z_j)\right)},$$

and $\sqrt{n}(\widehat{\theta}_{MLE} - \theta)$ is asymptotically normal with mean 0 and variance

$$\frac{\theta^2}{\eta_1\eta_2}\left(\frac{\int_0^\infty [\eta_1(1-H_1(t)) + \eta_2(1-H_2(t))\theta]\lambda_1(t)dt}{\int_0^\infty (1-H_1(t))\lambda_1(t)dt \int_0^\infty (1-H_2(t))\lambda_2(t)dt}\right).$$

Here $\eta_i = \lim_{n\to\infty}\frac{n_i}{n}$ and $1 - H_i(t)$ is the survival function of $X_{i\cdot}$, that is

$$\begin{aligned} 1 - H_i(t) &= [1 - F_i(t)][1 - G_i(t)] \\ &= \exp[-(\alpha_i + \alpha_i')t]\ , \quad i = 1, 2. \end{aligned}$$

2.2 Maximum Partial Likelihood

We consider the special case of Cox's proportional hazard model in which $p = 1$ and covariate vector is the sample indicator function. Then the Cox's proportional hazard model is reduced to

$$\lambda_2(t) = e^{\beta}\lambda_1(t),$$

where $\theta = e^{\beta}$ is called the relative risk. This model is equivalent to the two sample problem with proportional hazard. Thus the estimator of relative risk is obtained by the following Cox's partial likelihood with $\log\theta = \beta$,

$$L(\beta) = \prod_{j=1}^{n}\left(\frac{e^{\beta z_j}}{n_{1j} + n_{2j}e^{\beta}}\right)^{\delta_j}.$$

Then the partial likelihood estimator $\widehat{\theta}_{COX}$ of θ is $e^{\widehat{\beta}}$, where $\widehat{\beta}$ is the solution to

$$\begin{aligned} 0 &= \frac{\partial}{\partial\beta}\log L(\beta) \\ &= \sum_{j=1}^{n}\delta_j\left[z_j - \frac{n_{2j}e^{\beta}}{(n_{1j} + n_{2j}e^{\beta})}\right]. \end{aligned} \tag{2}$$

Solving the equation (2), which usually requires a iterative method. Therefore, if $\widehat{\beta}^0$ is an initial value, then $\widehat{\beta}$ is obtained by the Newton-Raphson's method,

$$\widehat{\beta}^1 = \widehat{\beta}^0 + i^{-1}(\widehat{\beta}^0)\frac{\partial}{\partial\widehat{\beta}^0}\log L(\widehat{\beta}^0),$$

where

$$\begin{aligned} i(\beta) &= -\frac{\partial^2}{\partial\beta^2}\log L(\beta) \\ &= \sum_{j=1}^{n}\delta_j\frac{n_{1j}n_{2j}e^{\beta}}{(n_{1j} + n_{2j}e^{\beta})^2}. \end{aligned}$$

Cox asserted that $\widehat{\beta}$ is asymptotically normal with mean β and variance $i^{-1}(\beta)$, if $\widehat{\beta}$ is the solution of the equation (2). Thus the estimated variance of $\widehat{\theta}_{COX}$ is obtained by multiplying $i^{-1}(\beta)$ by $(e^{\beta})^2$, using the delta method, and that $\sqrt{n}(\widehat{\theta}_{COX} - \theta)$ is asymptotically normal with mean 0 and variance

$$\frac{\theta^2}{\eta_1\eta_2}\left(\int_0^\infty \frac{(1-H_1(t))(1-H_2(t))\theta\lambda_1(t)dt}{\eta_1(1-H_1(t))+\eta_2(1-H_2(t))\theta}\right)^{-1}.$$

2.3 Generalized Rank Estimator

Begun and Reid(1983) and Andersen(1983) derived the generalized rank estimator of relative risk given by

$$\widehat{\theta}_K = \frac{\int_0^\infty K(t)d\widehat{\Lambda}_2(t)}{\int_0^\infty K(t)d\widehat{\Lambda}_1(t)},$$

where $K(t)$ is a predictive random weight function that depending only on the observations up to t^-, just before time t, and $\widehat{\Lambda}_i(t)$ is the Nelson-Aalen estimator (Nelson(1972), Aalen(1978)) of the cumulative hazard function for each sample.

Let $U_i(t)$ is defined by the subdistribution function of the observed death from each sample:

$$dU_i(t) = (1-H_i(t))\lambda_i(t).$$

Then

$$\begin{aligned} d\widehat{\Lambda}_1(t) &= \frac{dU_1(t)}{1-H_1(t)} = \frac{\delta_j(1-z_j)}{n_{1j}}, \\ d\widehat{\Lambda}_2(t) &= \frac{dU_2(t)}{1-H_2(t)} = \frac{\delta_j z_j}{n_{2j}}. \end{aligned}$$

Thus $\sqrt{n}(\widehat{\theta}_K - \theta)$ is asymptotically normal with mean 0 and variance

$$\theta^2\left(\frac{\int_0^\infty K^2(t)\frac{\eta_1(1-H_1(t))+\eta_2(1-H_2(t))\theta}{\eta_1(1-H_1(t))\eta_2(1-H_2(t))\theta}\lambda_1(t)dt}{[\int_0^\infty K^2(t)\lambda_1(t)dt]^2}\right)$$

(Andersen(1983)).

We now introduce the two types generalized rank estimators by the choice of $K(t)$.

First, we consider the generalized rank estimator using the weight function, $K(t)$, given by

$$K(t) = \frac{n_{1j}n_{2j}}{n_{1j}+n_{2j}}.$$

Then the estimator $\widehat{\theta}_{MH}$ can be obtained as follows :

$$\widehat{\theta}_{MH} = \frac{\sum_{j=1}^{n}(\delta_j z_j n_{1j})/(n-j+1)}{\sum_{j=1}^{n}(\delta_j(1-z_j)n_{2j})/(n-j+1)},$$

and $\sqrt{n}(\widehat{\theta}_{MH} - \theta)$ is asymptotically normal with mean 0 and variance

$$\frac{\theta^2}{\eta_1\eta_2}\left(\frac{\int_0^\infty \frac{(1-H_1(t))(1-H_2(t))}{(1-H(t))^2}[\eta_1(1-H_1(t)) + \eta_2(1-H_2(t))\theta]\lambda_1(t)dt}{\theta[\int_0^\infty \frac{(1-H_1(t))(1-H_2(t))}{1-H(t)}\lambda_1(t)dt]^2}\right).$$

Secondly, taking the weight function as follows:

$$K(t) = n_{1j}n_{2j}.$$

Then the generalized rank estimator, namely $\widehat{\theta}_{GH}$, is given by

$$\widehat{\theta}_{GH} = \frac{\sum_{j=1}^{n}\delta_j z_j n_{1j}}{\sum_{j=1}^{n}\delta_j(1-z_j)n_{2j}},$$

and $\sqrt{n}(\widehat{\theta}_{GH} - \theta)$ is asymptotically normal with mean 0 and variance

$$\frac{\theta^2}{\eta_1\eta_2}\left(\frac{\int_0^\infty (1-H_1(t))(1-H_2(t))[\eta_1(1-H_1(t)) + \eta_2(1-H_2(t))\theta]\lambda_1(t)dt}{\theta[\int_0^\infty (1-H_1(t))(1-H_2(t))\lambda_1(t)dt]^2}\right).$$

3. Relative Efficiency

In this section, we consider the asymptotic relative efficiencies of the relative risk estimators. The relative efficiencies are obtained by comparing their asymptotic variances which are presented in the previous section.

Since $\widehat{\theta}_{MLE}$ is a maximum likelihood estimator of relative risk, the $\widehat{\theta}_{MLE}$ always has a smaller asymptotic variance than any other estimators of θ, and the asymptotic variances of generalized rank estimators, $\widehat{\theta}_K$, can be written

$$\frac{\theta^2}{n}\left(\frac{\int_0^\infty K^2(t)\frac{\eta_1(1-H_1(t))+\eta_2(1-H_2(t))\theta}{\eta_1(1-H_1(t))\eta_2(1-H_2(t))\theta}\lambda_1(t)dt}{[\int_0^\infty K^2(t)\lambda_1(t)dt]^2}\right).$$

Thus we can see that Cox's estimator of θ is at least as efficient as any estimator $\widehat{\theta}_K$ by the inequality

$$\int_0^\infty K^2(t)\frac{\eta_1(1-H_1(t)) + \eta_2(1-H_2(t))\theta}{\eta_1(1-H_1(t))\eta_2(1-H_2(t))\theta}\lambda_1(t)dt$$
$$\times \int_0^\infty \frac{\eta_1(1-H_1(t))\eta_2(1-H_2(t))\theta}{\eta_1(1-H_1(t)) + \eta_2(1-H_2(t))\theta}\lambda_1(t)dt$$
$$\geq \left(\int_0^\infty K(t)\lambda_1(t)dt\right)^2.$$

Here the last inequality holds from the Cauchy-Schwarz inequality.

In particular, if $\theta = 1$, then the efficiency of the estimator $\widehat{\theta}_{MH}$ is that of $\widehat{\theta}_{COX}$. Moreover, if the distribution in each sample is exponentially distributed with the same parameter and censoring scheme, then the asymptotic variances of the estimators, $\widehat{\theta}_{MLE}$, $\widehat{\theta}_{COX}$, and $\widehat{\theta}_{MH}$, are the same.

To demonstrate some results of the above mentioned, we examine asymptotic efficiencies of the nonparametic estimators relative to the maximum likelihood estimator, which are computed for several exponential lifetime distributions and censoring scheme with various censoring rate and for true value of relative risk $\theta = 1(0.5)3$ and sample ratio $\eta_i = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}$, $i = 1, 2$, and $\eta_1 + \eta_2 = 1$.

Table 3.1 summarizes the relative efficiencies of nonparametric relative risk estimators relative to maximum likelihood estimator, such as Eff_1, Eff_2, and Eff_3, where

$$\begin{aligned} Eff_1 &= (A.Var\widehat{\theta}_{COX}/A.Var\widehat{\theta}_{MLE}), \\ Eff_2 &= (A.Var\ \widehat{\theta}_{MH}/A.Var\ \widehat{\theta}_{MLE}), \\ Eff_3 &= (A.Var\ \widehat{\theta}_{GH}/A.Var\ \widehat{\theta}_{MLE}). \end{aligned}$$

From Table 3.1, we can observe the following facts :

1. When the lifetime distribution in each sample is exponentially distributed with the same parameter and censoring scheme, the $\widehat{\theta}_{MLE}, \widehat{\theta}_{COX}$, and$\widehat{\theta}_{MH}$ have the same efficiencies.

2. In most cases, the Eff_2 are similar to Eff_1 and the Eff_3 are slightly larger then Eff_1 and Eff_2.

3. Obviously, the Eff_1, Eff_2 and Eff_3 get larger as the censoring rate and θ increase, regardless of sample ratio.

References

1. Aalen, O. O. (1978). Nonparametric inference for a family of counting processes. *Annals of Statistics*, 6, 701-726.

2. Andersen, P. K. (1983). Comparing survival distributions via hazard ratio estimates. *Scandinavian Journal of Statistics*, 10, 77-85.

3. Begun, J. M. and Reid, N. (1983). Estimating the relative risk with censored data. *Journal of American Statistical Association*, 78, 337-341.

4. Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of Royal Statistical Society B*, 34, 187-202.

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

6. Crowley, J., Liu, P. Y. and Voelkel, G. (1982). *Estimation of the ratio of hazard functions.* IMS Lecture notes-monograph series 2, 56-73.

7. Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. *Journal of American Statistical Association*, 72, 557-565.

8. Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14, 945-965.

Table 3.1 Asymptotic Relative Efficiencies of the Nonparametric Estimators

	Sample Ratio	Eff_1	Eff_2	Eff_3
$\theta = 1.0$	$\eta_1 = 0.25$ $\eta_2 = 0.75$	1.00000	1.00000	1.33333
	$\eta_1 = 0.33$ $\eta_2 = 0.67$	1.00000	1.00000	1.33333
Censoring Rate	$\eta_1 = 0.50$ $\eta_2 = 0.50$	1.00000	1.00000	1.33333
Sample 1 : 0%	$\eta_1 = 0.67$ $\eta_2 = 0.33$	1.00000	1.00000	1.33333
Sample 2 : 0%	$\eta_1 = 0.75$ $\eta_2 = 0.25$	1.00000	1.00000	1.33333
$\theta = 1.5$	$\eta_1 = 0.25$ $\eta_2 = 0.75$	1.03602	1.03733	1.46948
	$\eta_1 = 0.33$ $\eta_2 = 0.67$	1.03940	1.04079	1.43848
Censoring Rate	$\eta_1 = 0.50$ $\eta_2 = 0.50$	1.03918	1.04036	1.37648
Sample 1 : 10%	$\eta_1 = 0.67$ $\eta_2 = 0.33$	1.03174	1.03243	1.31448
Sample 2 : 10%	$\eta_1 = 0.75$ $\eta_2 = 0.25$	1.02576	1.02620	1.28348
$\theta = 2.0$	$\eta_1 = 0.25$ $\eta_2 = 0.75$	1.21301	1.22810	1.77736
	$\eta_1 = 0.33$ $\eta_2 = 0.67$	1.21186	1.22587	1.68706
Censoring Rate	$\eta_1 = 0.50$ $\eta_2 = 0.50$	1.18407	1.19405	1.51773
Sample 1 : 10%	$\eta_1 = 0.67$ $\eta_2 = 0.33$	1.13525	1.14058	1.36195
Sample 2 : 30%	$\eta_1 = 0.75$ $\eta_2 = 0.25$	1.10544	1.10869	1.28865
$\theta = 2.5$	$\eta_1 = 0.25$ $\eta_2 = 0.75$	1.09514	1.10984	1.63902
	$\eta_1 = 0.33$ $\eta_2 = 0.67$	1.10363	1.11865	1.59153
Censoring Rate	$\eta_1 = 0.50$ $\eta_2 = 0.50$	1.10518	1.11769	1.49061
Sample 1 : 30%	$\eta_1 = 0.67$ $\eta_2 = 0.33$	1.08883	1.09648	1.38092
Sample 2 : 10%	$\eta_1 = 0.75$ $\eta_2 = 0.25$	1.07404	1.07902	1.32242
$\theta = 3.0$	$\eta_1 = 0.25$ $\eta_2 = 0.75$	1.28547	1.32843	1.95217
	$\eta_1 = 0.33$ $\eta_2 = 0.67$	1.28922	1.33097	1.85364
Censoring Rate	$\eta_1 = 0.50$ $\eta_2 = 0.50$	1.26392	1.29706	1.65656
Sample 1 : 30%	$\eta_1 = 0.67$ $\eta_2 = 0.33$	1.20549	1.22464	1.45947
Sample 2 : 30%	$\eta_1 = 0.75$ $\eta_2 = 0.25$	1.16517	1.17718	1.36094