

Research Article

Power Loss of Stratified Log-Rank Test in Homogeneous Samples

Changyong Feng, Hongyue Wang, and Xin M. Tu

Department of Biostatistics and Computational Biology, University of Rochester, Rochester, NY 14642, USA

Correspondence should be addressed to Changyong Feng, feng@bst.rochester.edu

Received 25 September 2009; Accepted 7 February 2010

Academic Editor: Myong K. Jeong

Copyright © 2010 Changyong Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We study the loss of power of the stratified log-rank test (SLRT) compared to the unstratified log-rank test (ULRT) in the case of a large number of strata with relatively a small number of stratum sizes in terms of the asymptotic distributions of test statistics under local alternatives. The SLRT tends to lose information due to overstratification. It is better to test the homogeneity among strata before using the stratified log-rank test.

1. Introduction

It is well known in survival analysis that the (unstratified) log-rank test (ULRT) is the most efficient invariant test under contiguous alternatives in the proportional hazards model [1, 2]. Gill [3] gave a nice proof of this conclusion by using the Cauchy-Schwarz inequality.

In multicenter clinical trials with time-to-event as the primary outcome variable, we want to compare the treatment effects of two or more treatment methods. In the example in Section 6, patients were randomized in a 3 : 2 ration to two treatment groups. The primary outcome is the time to death from all causes during the study. Individuals from different centers are assumed to be independent. Even if the treatment effect can be assumed to be the same among centers, each center may have some factors which make the baseline hazard functions different from center to center. For this kind of data, the stratified log-rank test (SLRT) can be used to account for the baseline difference.

Some previous work has been developed to study the power loss of the log-rank test. Akazawa et al. [4] evaluated through simulation the loss of power of stratified log-rank test due to the heterogeneity in clinical trials. Generally, the power of the stratified log-rank test decreases due to two reasons: (i) the stratum size may be too small and (ii) the individuals in the same stratum are heterogeneous. The simulation shows that the loss of power is substantial when

the stratum size is very small and “the total number of failures and the treatment effect are fixed”. From the stratified Cox regression model in survival analysis, we can see that if the stratum size is small, its contribution to the overall test is small with censored data in the stratum. This may decrease the power of the stratified log-rank test. Note the stratified log-rank test is the score test from stratified partial likelihood (see Section 3).

In this paper, we consider the case where there is a large number of strata, but each stratum has a relatively small sample size. We assume that patients are homogeneous within each treatment group. For this kind of data, we can construct both the stratified and unstratified log-rank tests. We derive a variance relation between the SLRT and ULRT and quantify the power loss due to unnecessary stratification by this relation. We illustrate our approach with data from a multi-center clinical trial (MADIT II) to test the treatment effect of an implantable defibrillator on survival of patients with reduced left ventricular function after myocardial infarction.

This paper is organized as follows. Data and notation are described in Section 2. We derive the SLRT and ULRT and their local asymptotic distribution from the Cox proportional hazards regression model in Sections 3 and 4. An association between the SLRT and ULRT is developed in Section 5. We apply the approach to the MADIT II study data in Section 6, and offer concluding remarks in Section 7.

2. Data Structure and Notations

Suppose there are n centers with n_i patients randomized to treatment ($Z = 1$) or control group ($Z = 0$) in center i . Assume that $n_i, i = 1, \dots, n$, are independent and identically distributed (iid) positive integer random variables with finite second moments.

The underlying survival times $T_{ij}, i = 1, \dots, j = 1, \dots, n_i$, are subject to random censoring with censoring time C_{ij} . Here we assume that $\{C_{ij}\}$ are independent of $\{T_{ij}\}$ and $\{n_i\}$. We further assume that $\{C_{ij}\}$ are iid. Due to the censoring, the observable data are

$$X_{ij} = \min(T_{ij}, C_{ij}) = T_{ij} \wedge C_{ij}, \quad \delta_{ij} = \mathbf{1}\{T_{ij} \leq C_{ij}\}, \quad (1)$$

$$i = 1, \dots, n, \quad j = 1, \dots, n_i.$$

Define the stochastic processes $\{Y_{ij}\}$ and $\{N_{ij}\}$ by

$$Y_{ij}(t) = \mathbf{1}\{X_{ij} \geq t\}, \quad N_{ij}(t) = \mathbf{1}\{X_{ij} \leq t, \delta_{ij} = 1\}, \quad t \geq 0. \quad (2)$$

Therefore, $N_{ij}(t) = 1$ if and only if the event has happened before time t , and $Y_{ij}(t) = 1$ if and only if the patient is still at risk immediately before t . Note that

$$N_{ij}(t) = \delta_{ij} \mathbf{1}\{X_{ij} \leq t\} = \delta_{ij} \{1 - Y_{ij}(t+)\}, \quad (3)$$

$$\Delta N_{ij}(t) = N_{ij}(t) - N_{ij}(t-) = \delta_{ij} \{Y_{ij}(t) - Y_{ij}(t+)\}.$$

Note that $\Delta N_{ij}(t) = 1$ if and only if $\delta_{ij} = 1$ and $Y_{ij}(t) - Y_{ij}(t+) = 1$. The latter one means $X_{ij} = t$. Therefore, $\{Y_{ij}(t)\}^{\Delta N_{ij}(t)} = 1$ for all $t \geq 0$. Define

$$Y_i^{(1)}(t) = \sum_{j=1}^{n_i} Y_{ij}(t) Z_{ij},$$

$$Y_i^{(2)}(t) = \sum_{j=1}^{n_i} Y_{ij}(t) (1 - Z_{ij}),$$

$$Y^{(k)}(t) = \sum_{i=1}^n Y_i^{(k)}(t), \quad (4)$$

$$N_i^{(1)}(t) = \sum_{j=1}^{n_i} N_{ij}(t) Z_{ij},$$

$$N_i^{(2)}(t) = \sum_{j=1}^{n_i} N_{ij}(t) (1 - Z_{ij}),$$

$$N^{(k)}(t) = \sum_{i=1}^n N_i^{(k)}(t), \quad k = 1, 2.$$

Suppose that the hazard functions of T_{ij} are of the form

$$\lambda_{ij}(t | Z_{ij}) = \lim_{\Delta t \rightarrow 0} \frac{\Pr\{t \leq T_{ij} < t + \Delta t \mid T_{ij} \geq t, Z_{ij}\}}{\Delta t} \quad (5)$$

$$= \lambda_{i0}(t) \exp(Z_{ij}\beta).$$

For homogeneous samples, $\lambda_{i0}(t) = \lambda_0(t)$ for all i .

3. Stratified Log-Rank Test

The stratified log-rank test (SLRT) can be derived from the stratified Cox-proportional hazards model [5–7]. The contribution to the partial likelihood from stratum (center) i is

$$L_i(\beta) = \prod_{j=1}^{n_i} \prod_{t \geq 0} \left[\frac{\exp(Z_{ij}\beta)}{\sum_{l=1}^{n_i} Y_{il}(t) \exp(Z_{il}\beta)} \right]^{\Delta N_{ij}(t)}. \quad (6)$$

The log partial likelihood is

$$l_i(\beta) = \sum_{j=1}^{n_i} \int_0^\infty \left\{ Z_{ij}\beta - \log \left[\sum_{l=1}^{n_i} Y_{il}(t) \exp(Z_{il}\beta) \right] \right\} dN_{ij}(t). \quad (7)$$

Let

$$S_i^{(0)}(\beta, t) = \sum_{l=1}^{n_i} Y_{il}(t) \exp(Z_{il}\beta),$$

$$S_i^{(1)}(\beta, t) = \sum_{l=1}^{n_i} Y_{il}(t) Z_{il} \exp(Z_{il}\beta), \quad (8)$$

$$E_i(\beta, t) = \frac{S_i^{(1)}(\beta, t)}{S_i^{(0)}(\beta, t)}.$$

The first two order derivatives of the log partial likelihood are

$$s l_i(\beta) = \frac{\partial l_i(\beta)}{\partial \beta} = \sum_{j=1}^{n_i} \int_0^\infty \{Z_{ij} - E_i(\beta, t)\} dN_{ij}(t),$$

$$-I_i(\beta) = \frac{\partial^2 l_i(\beta)}{\partial \beta^2} = - \sum_{j=1}^{n_i} \int_0^\infty E_i(\beta, t) \{1 - E_i(\beta, t)\} dN_{ij}(t). \quad (9)$$

Under the null hypothesis $H_0 : \beta = 0$ (no treatment effect),

$$S_i^{(0)}(0, t) = Y_i^{(1)}(t) + Y_i^{(2)}(t),$$

$$S_i^{(1)}(0, t) = Y_i^{(1)}(t),$$

$$E_i(0, t) = \frac{Y_i^{(1)}(t)}{Y_i^{(1)}(t) + Y_i^{(2)}(t)},$$

$$s l_i(0) = \sum_{j=1}^{n_i} \int_0^\infty \{Z_{ij} - E_i(0, t)\} dN_{ij}(t) \quad (10)$$

$$= \sum_{j=1}^{n_i} \int_0^\infty \{Z_{ij} - E_i(0, t)\} dM_{ij}(t),$$

$$I_i(0) = \sum_{j=1}^{n_i} \int_0^\infty E_i(0, t) \{1 - E_i(0, t)\} dN_{ij}(t),$$

where $dM_{ij}(t) = dN_{ij}(t) - Y_{ij}(t)\lambda_0(t)dt$. Let

$$T_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^n s l_i(0). \quad (11)$$

The predictable variation of T_1 is

$$\begin{aligned} \langle T_1 \rangle &= \frac{1}{n} \sum_{i=1}^n \int_0^\infty \frac{Y_i^{(1)}(t) Y_i^{(2)}(t)}{Y_i^{(1)}(t) + Y_i^{(2)}(t)} \lambda_0(t) dt \\ &= \frac{1}{n} \sum_{i=1}^n I_i(0) + o_p(1). \end{aligned} \tag{12}$$

Under the null hypothesis, $T_1 \rightarrow N(0, \sigma_1^2)$, where σ_1^2 can be consistently estimated by $n^{-1} \sum_i^n I_i(0)$. The stratified log-rank test is defined as

$$SLRT = \frac{\sum_{i=1}^n sl_i(0)}{\sqrt{\sum_i^n I_i(0)}} = \frac{T_1}{\sqrt{n^{-1} \sum_i^n I_i(0)}}. \tag{13}$$

The asymptotic distribution of SLRT under local alternatives $\beta_n = \beta/\sqrt{n}$ is

$$SLRT \xrightarrow{\beta_n} N(\beta\sigma_1, 1), \tag{14}$$

which is derived in the appendix.

4. Unstratified Log-Rank Test

Similar to SLRT, the ULRT can be derived from the Cox proportional hazards model. The log partial likelihood function is

$$\begin{aligned} l(\beta) &= \sum_{i=1}^n \sum_{j=1}^{n_i} \int_0^\infty \left\{ \beta Z_{ij} - \log \left[\sum_{l=1}^n \sum_{m=1}^{n_l} Y_{lm}(t) \exp(\beta Z_{lm}) \right] \right\} dN_{ij}(t). \end{aligned} \tag{15}$$

Let

$$\begin{aligned} S^{(0)}(\beta, t) &= \sum_{i=1}^n S_i^{(0)}(\beta, t), \quad S^{(1)}(\beta, t) = \sum_{i=1}^n S_i^{(1)}(\beta, t), \\ E(\beta, t) &= \frac{S^{(1)}(\beta, t)}{S^{(0)}(\beta, t)}. \end{aligned} \tag{16}$$

The first-and the second-order derivatives of the log partial likelihood are

$$\begin{aligned} ul(\beta) &= \frac{\partial l(\beta)}{\partial \beta} = \sum_{i=1}^n \sum_{j=1}^{n_i} \int_0^\infty [Z_{ij} - E(\beta, t)] dN_{ij}(t), \\ -I(\beta) &= \frac{\partial^2 l(\beta)}{\partial \beta^2} = - \sum_{i=1}^n \sum_{j=1}^{n_i} \int_0^\infty E(\beta, t) [1 - E(\beta, t)] dN_{ij}(t). \end{aligned} \tag{17}$$

Under the null hypothesis $H_0 : \beta = 0$,

$$\begin{aligned} ul(0) &= \int_0^\infty \frac{Y^{(1)}(t) Y^{(2)}(t)}{Y^{(1)}(t) + Y^{(2)}(t)} \left\{ \frac{dN^{(1)}(t)}{Y^{(1)}(t)} - \frac{dN^{(2)}(t)}{Y^{(2)}(t)} \right\} \\ &= \sum_{i=1}^n \sum_{j=1}^{n_i} \int_0^\infty \{Z_{ij} - E(\beta, t)\} dM_{ij}(t), \end{aligned} \tag{18}$$

where (18) is the general form of log-rank used in literatures. Let

$$T_2 = \frac{1}{\sqrt{n}} ul(0). \tag{19}$$

The predictable variation of T_2 is

$$\langle T_2 \rangle = \frac{1}{n} \int_0^\infty \frac{Y^{(1)}(t) Y^{(2)}(t)}{Y^{(1)}(t) + Y^{(2)}(t)} \lambda_0(t) dt = \frac{1}{n} I(0) + o_p(1). \tag{20}$$

Under the null hypothesis, $T_2 \rightarrow N(0, \sigma_2^2)$, where σ_2^2 can be consistently estimated by $I(0)/n$. The unstratified log-rank test is defined as

$$ULRT = \frac{ul(0)}{\sqrt{I(0)}} = \frac{T_2}{\sqrt{I(0)/n}}. \tag{21}$$

With the same method as in the appendix, the asymptotic distribution of ULRT under local alternatives $\beta_n = \beta/\sqrt{n}$ is

$$ULRT \xrightarrow{\beta_n} N(\beta\sigma_2, 1). \tag{22}$$

5. A Relation between the Asymptotic Variances of T_1 and T_2

From martingale theory, the predictable covariation of T_1 and T_2 is

$$\langle T_1, T_2 \rangle = \frac{1}{n} \sum_{i=1}^n \int_0^\infty \frac{Y_i^{(1)}(t) Y_i^{(2)}(t)}{Y_i^{(1)}(t) + Y_i^{(2)}(t)} \lambda_0(t) dt = \langle T_1 \rangle. \tag{23}$$

This means that the asymptotic covariance of T_1 and T_2 satisfies

$$Acov(T_1, T_2) = Avar(T_1) = \sigma_1^2. \tag{24}$$

From Cauchy-Schwarz inequality,

$$\sigma_1^2 \leq \sigma_2^2. \tag{25}$$

Lemma 1. Suppose $Z \sim N(\theta, 1)$. Then for any $c > 0$, $\Pr\{|Z| \geq c\}$ is an increasing function of $|\theta|$.

This lemma is readily checked. From (22), (25), (A.4), and Lemma 1, the ULRT is always asymptotically more powerful than the SLRT in the homogeneous samples. The loss of power of SLRT is due to the loss of information from the unnecessary stratification.

6. A Real Example

We study the SLRT and ULRT in a multi-center clinical trial [8]. The Multicenter Automatic Defibrillator Implantation Trial II (MADIT-II) was designed to evaluate the potential survival benefit of a prophylactically implanted defibrillator in coronary patients with a prior myocardial infarction and advanced left ventricular (ejection fraction ≤ 0.3). The trial started in July 1997, and enrolled 1232 patients from 76 hospital centers (71 in US and 5 in Europe). The patients were randomized in a 3 : 2 ratio to receive either an

implantable defibrillator or conventional medical therapy. We first test the homogeneity of strata by the log-rank test (P -value = .06). The estimated variances of T_1 and T_2 are $\sigma_1^2 = 46.48$ and $\sigma_2^2 = 48.04$, respectively. This shows that the ULRT is asymptotically slightly more powerful than its SLRT counterpart. For those 76 centers, the number of patients ranges from 1 to 55 with mean value 16.4. There were 21 centers without any event by the end of the study. They did not contribute to the stratified log-rank test.

7. Discussion

In this paper, we studied the loss of power of stratified log-rank test in multi-center clinical trials with a large number of centers, but relatively small stratum size (assuming homogeneous strata). Our results show that asymptotic variance of SLRT is smaller than that of ULRT which makes the SLRT less powerful. Overstratification may incur loss of information compared to the unstratified log-rank test. However, there are some limitations in our study. First, we assumed that strata are homogeneous. In that case, the unstratified log-rank test should be the best choice. In practice, it is important to test homogeneity of strata before using the stratified log-rank test. Second, we considered the case with a large number of strata, but small stratum size. Another case of interest is a small number of strata with large stratum size. Although (23) is always true, the local alternatives cannot be specified in terms of n . We are currently investigating whether (25) is still true.

Appendix

The Asymptotic Distribution of SLRT Under Local Alternatives

To study the asymptotic distribution of SLRT under alternatives, we consider the following local alternatives: $\beta_n = \beta/\sqrt{n}$. Then from Taylor expansion,

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n sl_i(\beta_n) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n sl_i(0) - \beta \frac{1}{n} \sum_{i=1}^n I_i(0) \\ &\quad + \frac{\beta^2}{n} \frac{1}{2\sqrt{n}} \sum_{i=1}^n sl_i''(\theta_n), \end{aligned} \quad (\text{A.1})$$

where θ_n is between 0 and β_n . Note that $(1/2\sqrt{n}) \sum_{i=1}^n sl_i''(\theta_n)$ is bounded in probability uniformly. Then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n sl_i(\beta_n) = \frac{1}{\sqrt{n}} \sum_{i=1}^n sl_i(0) - \beta \frac{1}{n} \sum_{i=1}^n I_i(0) + o_p(1). \quad (\text{A.2})$$

Therefore,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n sl_i(0) \xrightarrow{\beta_n} N(\beta\sigma_1^2, \sigma_1^2). \quad (\text{A.3})$$

This means that

$$\text{SLRT} \xrightarrow{\beta_n} N(\beta\sigma_1, 1). \quad (\text{A.4})$$

Acknowledgments

The authors gratefully thank Dr. Author J. Moss (PI of MADIT-II) for allowing them to use the MADIT-II data in this paper. This research was supported by Grant 5U19AI056390-05 from the National Institutes of Health of USA.

References

- [1] R. Peto, "Rank tests of maximal power against Lehmann-type alternatives," *Biometrika*, vol. 59, pp. 472–475, 1972.
- [2] R. Peto and J. Peto, "Asymptotically efficient rank invariant test procedures (with discussion)," *Journal of the Royal Statistical Society. Series A*, vol. 135, no. 2, pp. 185–207, 1972.
- [3] R. D. Gill, *Censoring and Stochastic Integrals*, vol. 124 of *Mathematical Centre Tracts*, Mathematisch Centrum, Amsterdam, The Netherlands, 1980.
- [4] K. Akazawa, T. Nakamura, and Y. Palesch, "Power of logrank test and Cox regression model in clinical trials with heterogeneous samples," *Statistics in Medicine*, vol. 16, no. 5, pp. 583–597, 1997.
- [5] D. R. Cox, "Regression models and life-tables (with discussion)," *Journal of the Royal Statistical Society. Series B*, vol. 34, pp. 187–220, 1972.
- [6] D. R. Cox, "Partial likelihood," *Biometrika*, vol. 62, no. 2, pp. 269–276, 1975.
- [7] P. K. Andersen, Ø. Borgan, R. D. Gill, and N. Keiding, *Statistical Models Based on Counting Processes*, Springer Series in Statistics, Springer, New York, NY, USA, 1993.
- [8] A. J. Moss, W. Zareba, W. J. Hall, et al., "Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction," *The New England Journal of Medicine*, vol. 346, no. 12, pp. 877–883, 2002.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

