# Nonparametric Test For Current Status Data Under The Stratum Cox Proportional Hazards Model

Xiaodong Fan[a], Qingchun Zhang[b,*]

School of Science, Jilin Institute of Chemical Technology, Jilin City, Jilin Province, 132022, China

[a]fanxd14@mails.jlu.edu.cn, [b]48341846@qq.com

*corresponding author

**Keywords:** Current Status Data; Nonparametric Test; Cox Model; Stratum Effects

**Abstract:** This article develops the test methods of stratum interaction for current status data under the Cox proportional hazards model. We proposed a test procedure for this type of current status data and established the asymptotic properties of the test statistic by the martingale theory.

## 1. Introduction

Current status data occur when every study individual is observed only once at a particular observation time, and the observed information of whether the event interested has occurred no later than the observation time (Sun 2006; Tran et al. 2020; Xu et al. 2020). Sun and Yang (2000) discussed nonparametric tests for stratum effects of right-censored failure time data under the Cox proportional hazards model. Fan et al. (2019) discussed nonparametric test of stratum effects for interval-censored failure time data in the Cox proportional hazards model. In this paper, we considered nonparametric test for stratum effects of current status data from Cox proportional hazards model (Chen et al. 2012; Huang 1996). The remainder paper is organized as follows. In section 2, we introduce some notation and present the proposed test procedure, which is developed by the idea of Sun and Yang (2000). Some conclusions and remarks are given in Section 3.

## 2. Notation and Test Procedure

Assume that there are $q$ strata of subjects. Let $T_{ki}$ and $Z_{ki}$ denote the failure time and a vector of covariates, for subject $i$ in the $k$ th stratum, $i=1,2,\cdots,n_k, k=1,2,\cdots,q$, here $n_1+n_2+\cdots+n_q=n$. The hazards model of $T_{ki}$ given $Z_{ki}$ can be written as

$$\lambda_{ki}^t(t, Z_{ki}(s), s \le t) = \lambda_k(t) e^{\beta_0' Z_{ki}(t)}, \tag{1}$$

where $\lambda_k(t)$ is baseline hazard for the subjects $i$ in $k$ th stratum, and $\beta_0$ is regression parameters. Let $T_{ki}'s$ be the failure times, and $C_{ki}$ denote the censored times. The current status data consists of $\{(C_{ki}, \delta_{ki} = I(T_{ki} \ge C_{ki}), Z_{ki}(t), t \le C_{ki}); i=1,\cdots,n_k, k=1,\cdots,q\}$. In practice, $C_{ki}$ may depend on covariates too. For this, given $\{Z_{ki}(s), s \le t\}$, supposed that $C_{ki}$ follows the proportional hazards model

$$\lambda_{ki}^c(t, Z_{ki}(s), s \le t) = \lambda^c(t) e^{\gamma_0' Z_{ki}(t)}, \tag{2}$$

where $\gamma_0$ is regression parameters and $\lambda^c(t)$ is baseline hazard. In order to simplify the notation, it is assumed that the covariates have the same effects in models (1) and (2) for the subjects in all $q$ strata and the test procedure proposed in this paper is still right if the covariates have different effects. Next, we consider the following null hypothesis

DOI: 10.38007/Proceedings.0001574  ISBN: 978-1-80052-010-3

$$H_0 : \lambda_1(t) = \lambda_2(t) = \cdots = \lambda_q(t).$$

To test $H_0$, define $N_{ki}^c(t) = I(C_{ki} \le \min(t, T_{ki}))$, $\bar{N}_k^c(t) = \sum_{i=1}^{n_k} N_{ki}^c(\mathrm{t})$, and $\bar{N}(t) = \sum_{k=1}^{q} \bar{N}_k^c(\mathrm{t})$. Then $N_{ki}^c(t)$ is a counting process and has the intensity process

$$\lambda_{ki}(t, Z_{ki}(s), s \le t) = e^{\gamma_0' Z_{ki}(t) - \int_0^t e^{\beta_0' Z_{ki}(s)} ds} \lambda^c(t) e^{-\int_0^t \lambda_k(s) ds} = \tilde{\lambda}_k(t) e^{\gamma_0' Z_{ki}(t) - \int_0^t e^{\beta_0' Z_{ki}(s)} ds} \quad (3)$$

(Lin et al., 1998), where $\tilde{\lambda}_k(t) = \lambda^c(t) e^{-\int_0^t \lambda_k(s) ds}$. Note that equation (3) is the Cox proportional hazards model. Define $Y_{ki}^c(t) = I(C_{ki} \ge t)$, $Y_k^c(\mathrm{t}, \beta_0, \gamma_0) = \sum_{i=1}^{n_k} Y_{ki}^c(\mathrm{t}) e^{\gamma_0' Z_{ki}(t) - \int_0^t e^{\beta_0' Z_{ki}(s)} ds}$, and $J_k^c(\mathrm{t}) = I(\sum_{i=1}^{n_k} Y_{ki}^c(\mathrm{t}) > 0)$. Let $(\hat{\beta}, \hat{\gamma})$ is the partial likelihood estimator of the true parameter $(\beta_0, \gamma_0)$ under $H_0$ (Feng et al., 2015; Lin et al., 1998).

To test $H_0$, we propose the statistic $W(\hat{\beta}, \hat{\gamma}) = (W_1(\hat{\beta}, \hat{\gamma}), \cdots, W_q(\hat{\beta}, \hat{\gamma}))'$, where

$$W_k(\beta_0, \gamma_0) = \int_0^\tau L(t) Y_k^c(t, \beta_0, \gamma_0) \{ d\hat{\Lambda}_k(t, \beta_0, \gamma_0) - J_k^c(t) d\hat{\Lambda}_0(t, \beta_0, \gamma_0) \}. \quad (4)$$

Here $L(t)$ is weight function. In (4),

$$\hat{\Lambda}_k(\mathrm{t}, \beta_0, \gamma_0) = \int_0^t \frac{J_k^c(\mathrm{s})\, \mathrm{d}\bar{\mathrm{N}}_k^c(\mathrm{s})}{\sum_{i=1}^{n_k} Y_{ki}^c(\mathrm{s}) e^{\gamma_0' Z_{ki}(t) - \int_0^t e^{\beta_0' Z_{ki}(s)} ds}} \quad (5)$$

is the cumulative hazard function for the $k$ th stratum, and under $H_0$,

$$\hat{\Lambda}_0(\mathrm{t}, \beta_0, \gamma_0) = \int_0^t \frac{J(\mathrm{s})\, \mathrm{d}\bar{\mathrm{N}}(\mathrm{s})}{\sum_{k=1}^{q} \sum_{i=1}^{n_k} Y_{ki}^c(\mathrm{s}) e^{\gamma_0' Z_{ki}(t) - \int_0^t e^{\beta_0' Z_{ki}(s)} ds}} \quad (6)$$

is the common cumulative hazard function. The above test statistics were motivated by Sun and Yang (2000) about right censored data. Then under some mild regularity conditions and $H_0$, as $n \to \infty$ and $n_k / n \to q_k$, the distribution of $W(\hat{\beta}, \hat{\gamma})$ can be approximated by multivariate normal distribution $N(0, \Sigma(\hat{\beta}, \hat{\gamma}))$. The covariate matrix $\Sigma(\hat{\beta}, \hat{\gamma}) = (\hat{\sigma}_{k_1 k_2}(\hat{\beta}, \hat{\gamma}))_{q \times q}$, and

$$\hat{\sigma}_{k_1 k_2}(\hat{\beta}, \hat{\gamma}) = \sum_{l=1}^{q} \sum_{i=1}^{n_l} \{ \int_0^\tau L(\mathrm{t}) \mathrm{J}_{k_1}^c(\mathrm{t}) [\xi_{k_1 l} - \frac{n_{k_1} S_{k_1}^0(\mathrm{t}, \hat{\beta}, \hat{\gamma})}{n S^{(0)}(\mathrm{t}, \hat{\beta}, \hat{\gamma})}] dN_{li}^c(\mathrm{t})$$
$$\times \int_0^\tau L(\mathrm{t}) \mathrm{J}_{k_1}^c(\mathrm{t}) [\xi_{k_2 l} - \frac{n_{k_2} S_{k_2}^0(\mathrm{t}, \hat{\beta}, \hat{\gamma})}{n S^{(0)}(\mathrm{t}, \hat{\beta}, \hat{\gamma})}] dN_{li}^c(\mathrm{t}) \}, \quad (7)$$

$q_k > 0$ and $\sum_{l=1}^{q} q_k = 1$. In the above formula,

$$S_k^0(\mathrm{t}, \hat{\beta}, \hat{\gamma}) = \frac{1}{n_k} \sum_{i=1}^{n_k} Y_{ki}^c(\mathrm{t}) e^{\gamma_0' Z_{ki}(t) - \int_0^t e^{\beta_0' Z_{ki}(s)} ds}, \quad (8)$$

and

$$S^{(0)}(\mathrm{t}, \hat{\beta}, \hat{\gamma}) = \frac{1}{n} \sum_{l=1}^{q} \sum_{i=1}^{n_l} Y_{ki}^c(\mathrm{t}) e^{\gamma_0' Z_{ki}(t) - \int_0^t e^{\beta_0' Z_{ki}(s)} ds}. \quad (9)$$

Note that $\sum_{k=1}^{q} W_k(\hat{\beta}, \hat{\gamma}) = 0$. Let $W_0(\hat{\beta}, \hat{\gamma})$ denote the first $k-1$ constituents of $W_k(\hat{\beta}, \hat{\gamma})$, and $\hat{D}(\hat{\beta}, \hat{\gamma})$ denote $\Sigma(\hat{\beta}, \hat{\gamma})$ with the last row and column deleted. Then the $H_0$ can be test by using the test statistic $\chi_0^2 = W_0(\hat{\beta},\hat{\gamma})^T \hat{D}^{-1}(\hat{\beta},\hat{\gamma}) W_0(\hat{\beta},\hat{\gamma})$, which has an asymptotic the $\chi^2$ distribution with $q-1$ degrees of freedom. Using the method above, we need to choose the weight processes $L(t)$. For the comments on the selection of weight processes, see Sun and Yang (2000) and Andersen et al. (1993).

## 3. Conclusion

This paper discusses the test of the stratum effect in the stratified Cox proportional hazards model when current status failure time data are available. For the problem, a test approach was proposed and the asymptotic properties of the test procedures were established. A direction for future research is to generalize the proposed test procedure to Interval-censored data (Sun, 2006).

## References

[1] Chen, D., Sun, J., and Peace, K. (2012). Interval-Censored Time-to-Event Data: Methods and Applications. Chapman & Hall/CRC, London.

[2] Fan, X., Zhao, S., and Sun, J. (2019). Nonparametric Test for Stratum Effects in the Cox Model with Interval-censored Data. Communications in Statistics-Theory and Methods, in press 48, 300-310.

[3] Feng, Y., Ma, L., and Sun, J. (2015). Additive hazards regression with auxiliary covariates for case I interval-censored data. Scand J Stat 42, 118-136.

[4] Huang, J. (1996). Efficient estimation for the proportional hazard model with interval censoring. Ann Stat 24, 540-568.

[5] Lin, D., Oakes, D., and Ying, Z. (1998) Additive hazards regression with current status data. Biometrika 85, 289-298.

[6] Sun, J. (2006). Statistical Analysis of Interval-censored Failure Time Data. Springer, New York.

[7] Sun, J. and Yang, I. (2000). Nonparametric test for stratum effects in the cox model. Lifetime Data Analysis 6, 321-330.

[8] Tran T. M. P., Abrams, S. and Braekers, R. (2000). A general frailty model to accommodate individual heterogeneity in the acquisition of multiple infections: An application to bivariate current status data. Statistics in Medicine 39(12), 1695-1714.

[9] Xu, D., Zhao, S. and Sun, J. (2020). Regression analysis of dependent current status data with the accelerated failure time model. Communication in Statistics-Simulation and Computation. DOI: 10.1080/03610918.2020.1797795.