Estimation Methods for Accelerated Failure Time Model

Zhezhen Jin

Department of Biostatistics
Mailman School of Public Health
Columbia University
722 West 168th Street
New York, NY 10032, USA
(e-mail: zj7@columbia.edu)

Abstract. In the literature, a lot of effort has been devoted to develop effective estimation and inference methods for the accelerated failure time (AFT) model for right censored data. In the talk, we will give a review on the recent development on the estimation and inference methods for the AFT model based on the work in [Jin *et al.*, 2003] and [Jin *et al.*, 2004].
Keywords: Accelerated failure time model, Right censoring, Rank estimation, Least squares method.

Right censored data are common in many scientific fields. The right censored data consist of (X_i, Y_i, δ_i), $i = 1, \cdots, n$, where X is a p-dimensional covariate, $Y = \min\{T, C\}$, with T being the response variable and C being the censoring variable, and $\delta = 1\{T \leq C\}$ being the indicator of censoring.

The accelerated failure time (AFT) model is of the same form as usual linear regression model:

$$\log T_i = X_i^T \beta_0 + \epsilon_i \tag{1}$$

where β_0 is the unknown true $p \times 1$ parameter of interest and ϵ_i $(i = 1, \cdots, n)$ are unobservable independent random errors with a common but completely unspecified distribution function. (Thus, the mean of ϵ is not necessarily 0). The AFT model is an attractive alternative to the popular Cox proportional regression model, [Cox, 1972].

Several approaches have been proposed for the estimation and inference on the AFT model in the literature. Rank-based methods were studied [Tsiatis, 1990], [Wei *et al.*, 1990], [Lai and Ying, 1991], [Lai and Ying, 1992], [Lin and Geyer, 1992], [Ying, 1993], [Fygenson and Ritov, 1994], among many others. Least squares based and M-estimation methods were investigated by [Miller, 1976], [Buckley and James, 1979], [Koul *et al.*, 1981], [Ritov, 1990] and [Lai and Ying, 1991], among many others. Despite theoretical advances, all these approaches are numerically complicated and difficult to implement, especially when the number of covariates is large. These are due to the non-differentiability and non-monotonicity of the estimating functions. Furthermore, the covariance matrices of the estimators are rather difficult

to obtain because they involve nonparametric estimation of the underlying unknown density function for ϵ.

Recently, we have developed new rank-based and least squares estimation and inference method for the AFT model [Jin *et al.*, 2003], [Jin *et al.*, 2004]. In [Jin *et al.*, 2003], a class of rank-based estimating functions are developed. The functions are monotone and can be easily solved by linear programming technique. The covariance matrix of the parameter estimators are obtained by a resampling method. In [Jin *et al.*, 2004], a numerically easy to implement least squares method is developed and a resampling method sharing the similar spirit in the rank-estimation is also proposed.

In the talk, we will give a review on the recent development on the estimation and inference methods for the AFT model.

References

[Buckley and James, 1979]I.V. Buckley and I. James. Linear regression with censored data. *Biometrika*, pages 429–436, 1979.

[Cox, 1972]D.R. Cox. Regression models and life-tables (with discussion). *J. R. Statist. Soc. Ser. B*, pages 187–220, 1972.

[Fygenson and Ritov, 1994]M. Fygenson and Y. Ritov. Monotone estimating equations for censored data. *Annals of Statist.*, pages 732–746, 1994.

[Jin *et al.*, 2003]Z. Jin, D.Y. Lin, L.J. Wei, and Z. Ying. Rank-based inference for the accelerated failure time model. *Biometrika*, pages 341–353, 2003.

[Jin *et al.*, 2004]Z. Jin, D.Y. Lin, and Z. Ying. On least-squares regression with censored data. *Manuscript*, 2004.

[Koul *et al.*, 1981]H. Koul, V. Susarla, and J. Van Ryzin. Regression analysis with randomly right-censored data. *Ann. Statist.*, pages 1276–1288, 1981.

[Lai and Ying, 1991]T.L. Lai and Z. Ying. Large sample theory of a modified buckley-james estimator for regression analysis with censored data. *Ann. Statist.*, pages 1370–1402, 1991.

[Lai and Ying, 1992]T.L. Lai and Z. Ying. Linear rank statistics in regression analysis with censored or truncated data. *Journal of Multivariate Statistics*, pages 13–45, 1992.

[Lin and Geyer, 1992]D.Y. Lin and C.J. Geyer. Computational methods for semiparamtric linear regression with censored data. *J. Computational and Graphical Stat.*, pages 77–90, 1992.

[Miller, 1976]R. G. Miller. Least squares regression with censored data. *Biometrika*, pages 449–464, 1976.

[Ritov, 1990]Y. Ritov. Estimation in linear regression model with censored data. *Ann. Statist.*, pages 354–372, 1990.

[Tsiatis, 1990]A.A. Tsiatis. Estimating regression parameters using linear rank tests for censored data. *Ann. Statist.*, pages 354–372, 1990.

[Wei *et al.*, 1990]L.J. Wei, Z. Ying, and D.Y. Lin. Linear regression analysis of censored survival data based on rank tests. *Biometrika*, pages 845–851, 1990.

[Ying, 1993]Z. Ying. A large sample study of rank estimation for censored regression data. *Ann. Statist.*, pages 76–99, 1993.