Goodness-of-Fit Tests for Semiparametric Models with Multiple Event-Time Data

Suktae Choi
Food and Drug Adiministration
9201 Corporate blvd
Rockville, MD 20850, U.S.A.
chois@cder.fda.gov
Zhezhen Jin
Division of Biostatistics, Columbia University
600 West 168th street
NY, NY 10032, U.S.A.
Zhilian Ying
Department of Statistica, Columbia University
2990 Broadway
NY, NY 10032, U.S.A.

Abstract

A counting process approach to multiple event times modeled by an Andersen-Gill-type extension of the Cox proportional hazards regression model is considered. Tests for checking the validity of such a model against a general frailty model are proposed. These tests are derived from a class of statistics that are connected to Robbins' empirical Bayes estimation of Poisson means. We show that these tests are consistent against any alternative as specified by a nondegenerate frailty. A simple graphical method is introduced to visually check the appropriateness of model assumptions. Simulation studies are reported and a real life example is presented. A similar test for checking the gamma frailty assumption is also introduced.

Key words and phrases: Counting process, Cox regression, empirical Bayes, goodness of fit, martingale, multiple event times, Poisson process, random effect.

Introduction

Multiple event-time data have been the focus of many recent investigations in survival analysis. They represent a special kind of multivariate failure time data with each subject experiencing a series of ordered events. Such data arise naturally in biomedical research. Examples include studies of progression of AIDS and other diseases, and of recurrences of chronic diseases. Other examples arise in industrial life tests and software reliability studies.

In principle, one-sample multiple event-time data can be analyzed by nonparametric methods developed for more general multivariate survival time data. Nonparametric estimates of joint survival distribution can be found in Campbell (1981), Tsai, Leurgans and Crowley (1986), Dabrowska (1988), Prentice and Cai (1992) and van der Laan (1996), among others. These estimates are generally complicated, often involving non- parametric density-type estimation. They are also difficult to incorporate into analysis of regression problems.

By modeling inter-event gap times, Prentice, Williams and Peterson (1981) proposed a regression method for multiple event-time data. Parallel and subsequent developments for such data can be found in Gail, Santner and Brown (1980) and Dabrowska, Sun and Horowitz (1992). Wei, Lin and Weissfeld (1989), on the other hand, proposed a marginal proportional hazards regression model, in which each of the event times is modeled marginally through a Cox model. While their method has the advantage of requiring only mild model assumptions, it also has the disadvantage of not utilizing a dependency structure that can potentially increase efficiency.

A simple model that can handle multiple event-time data is provided by an extension of Cox regression to counting processes through application of a general approach due to Andersen and Gill (1982). This model can be handled in the same way as the Cox model, both numerically and theoretically. Results based on such a model are also easy to interpret. However, as will be elaborated in subsequent sections, the Andersen-Gill model does not take into account the dependence of events experienced by the same subject. To alleviate this shortcoming, one may introduce a random effect (frailty) into the Andersen-Gill model. The resulting formulation is a more general model that allows for suitable inference not only on regression effects but on associations within subject as well. Approaches in this direction can be found in Nielsen, Gill, Andersen and Sorensen (1992) and Oakes (1992).

The main objective of this investigation is to develop statistical methods for checking and testing the presence of a nondegenerate frailty. It is connected to a classical paper on the empirical Bayes methodology by Robbins (1955). By utilizing an interesting identity due to Robbins for Poisson random variables with mean parameter following a prior distribution, a class of statistics is derived. Th

e statistics are then used to construct tests for the Andersen-Gill model against alternatives of nonde generate frailty. It is shown that the tests are consistent against any such alternative. They are also u sed to construct simple graphical plots for model checking, as well as for visual inspection to check where deviations from the null model may occur. Simulation results indicate that the method perfor ms well in detecting the presence of frailty. A real life example of tumor occurrence in experimental mice (Gail, Santner and Brown (1980)) is used to illustrate the method. A similar method for check ing whether the frailty distribution belongs to the family of gamma distributions is also introduced.

Conclusion

We have proposed a class of statistics that could be useful to construct test statistics or checking the validity of an Andersen-Gill-type model for multiple event times against he presence of frailty. The statistics arise from comparing the standard Nelson-Aalen stimator with an alternative estimato r that is motivated by an interesting formula or the posterior mean due to Robbins (1955). We show that the resulting tests are consistent against any nondegenerate frailty alternative. Large sample just ification is given. Simulation results and a real example indicate that the proposed tests have reasonab le power. A useful by-product is a graphical tool for model checking.

When there is a nondegenerate frailty, we discussed briefly the problem of testing whether the fr ailty distribution belongs to the family of gamma distributions. The idea is to construct suitable stati stics that possess a certain pattern under the gamma frailty assumption. Lack of such a pattern woul d indicate a departure from this assumption. Modi_cations of test statistic is needed when there is ra ndom censoring. One possible approach is to use the same idea outlined in previous paragraph. Ho wever, the investigation of this approach is very preliminary and much work remains to be done.

REFERENCE

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, New York.

Andersen, P.K., and Gill, R.D. (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. *A nn. Statist*. **10**, 1100-1120.

Bremaud, P. (1981). *Point Processes and Queues: Martingale Dynamics.* Springer-Verlag, New York.

Casella, G. and Berger, R. L. (1990). *Statistical Inference.* Wadsworth & Brooks/Cole, Paci_c Grove, California.

Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika* **68**, 417-522.

Dabrowska, D. M. (1988). Kaplan-Meier estimate on the plane. *Ann. Statist.* **16**, 1475-1489.

Dabrowska, D. M., Sun, G. and Horowitz, M. M. (1992). Cox regression in a Markov renewal model with an appl ication to the analysis of bone marrow transplant data. *J. of Amer. Statist. Assoc.* **89**, 867{877.

Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis.* John Wiley, NewYork.

Gail, M. H., Santner, T. J. and Brown, C. C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics* **36**, 255-266.

Lawless, J.F. and Nadeau, C. (1995). Some Simple Robust Methods for the Analysis of Recurrent Events. *Techno metrics* **37**, 158-168.

Nielsen, G.G., Gill, R.D., Andersen, P.K. and Sorensen, T.I.A. (1992). A counting process approach to maximum l ikelihood estimation in frailty models. *Scan. J. Statist*. **19**, 25-43.

Oakes, D. (1992). Frailty models for multiple event-time data. In Klein, J.P. and Goel, P.K. (eds). *Survival Analysi s: State of the Art*, 471-480, Kluwer, Boston.

Parner, E. (1998). Asymptotic theory for the correlated Gamma-frailty model. *Ann. Statist.* **26**, 183-214.

Pepe, M. S. and Cai, J. (1993). Some Graphical Displays and Marginal Regression Analyses for Recurrent Failure Times and Time Dependent Covariates. *J. Amer. Statist. Assoc.* **88**, 811-820.

Prentice, R. L. and Cai, J. (1992). Covariance and survival function estimation using censored multivariate failure time data. *Biometrika* **79**, 495-512.

Prentice, R. L., Williams, B.J. and Peterson, A.V. (1981). On the regression analysis of multivariate failure time da ta. *Biometrika* **68**, 373-379.

Robbins, H. (1955). An empirical Bayes approach to statistics. *Proc. Third Berkeley Symposium on Math. Statist. and Prob.* **1**, 157-164.

Tsai, W-Y., Leurgans, S. and Crowley, J. (1986). Nonparametric estimation of a bivariate survival function in the presence of censoring. *Ann. Statist.* **14**, 1351-1365.

van der Laan, M. J. (1996). Efficient estimation in the bivariate censoring model and repairing NPMLE. *Ann. Stati st.* **24**, 596-627.

Wei, L. J., Lin, D. Y. and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Amer. Statist. Assoc.* **84**, 1065-1073.

Note

This paper excludes the body part to fit into 2 pages.
Copies in full version will be available during the presentation.