

Extensões do Modelo de Cox para Acontecimentos Múltiplos: Implementação em R

Ivo Sousa-Ferreira

Departamento de Matemática, FCEE, Universidade da Madeira, ivo.ferreira@staff.uma.pt

Ana Maria Abreu

Departamento de Matemática, FCEE, Universidade da Madeira e

Centro de Investigação em Matemática e Aplicações (CIMA), abreu@staff.uma.pt

Palavras-chave: Análise de Sobrevivência, acontecimentos recorrentes, estatística computacional, modelos semiparamétricos, *software R*, simulação de dados.

Resumo:

O modelo de Cox [2] foi desenvolvido para estudar o tempo de vida até à ocorrência de um único acontecimento. No entanto, a ocorrência de vários acontecimentos para um mesmo indivíduo, algo que é cada vez mais frequente, deu origem a um certo tipo de extensões deste modelo [9]. As mais usuais foram propostas por Prentice, Williams e Peterson [6], Andersen e Gill [1], Wei, Lin e Weissfeld [10] e Lee, Wei e Amato [4].

Neste trabalho, pretende-se exemplificar o modo como estas quatro extensões do modelo de Cox são implementadas com recurso ao *software* estatístico R [7], mais concretamente através do *package survival* [8]. Para isso é necessário dispor de um conjunto de dados, donde recorreu-se ao *package survsim* [5] para simular acontecimentos recorrentes [3]. Por conseguinte, serão feitas algumas considerações acerca do cuidado que é necessário ter na construção e organização da base de dados, onde a definição do intervalo de risco se revela uma componente decisiva. Uma vez que cada modelo foi proposto para tratar dados com determinadas características, será ainda apresentado um guia prático que, para além de destacar as diferenças entre estes modelos, permite apontar para a escolha do modelo mais adequado a cada caso de estudo.

Agradecimentos

O presente trabalho foi parcialmente suportado pela Fundação para a Ciência e a Tecnologia (FCT), através de fundos nacionais, no âmbito do projeto UID/MAT/04674/2013.

Referências

- [1] Andersen, P. K. e Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *The Annals of Statistics*, Vol. 10, No. 4, pp. 1100–1120.
- [2] Cox, D.R. (1972). Regression Models and Life-Tables (with discussion). *Journal of the Royal Statistical Society B*, Vol. 34, No. 2, pp. 187-220.
- [3] Ferreira, I. M. S. (2016). *Modelos para acontecimentos múltiplos*. (Dissertação de Mestrado). Universidade da Madeira, Funchal, Portugal.
- [4] Lee, E. W., Wei, L. J. e Amato, D. A. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In Klein, J. P. and Goel, P. K. (Eds.): *Survival Analysis: State of the Art*, pp. 237–247, Kluwer Academic Publisher, Dordrecht.

- [5] Moriña, D. e Navarro, A. (2014). The R package survsim for the simulation of simple and complex survival data. *Journal of Statistical Software*, Vol. 59, No. 2, pp. 1–20.
- [6] Prentice, R. L., Williams, B. J. e Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika*, Vol. 68, No. 2, pp. 373–379.
- [7] R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org/>.
- [8] Therneau, T. M. (2017). *survival: A package for survival analysis in S*. Package do R versão 2.41-3. URL: <http://CRAN.R-project.org/package=survival>.
- [9] Therneau, T. M. e Grambsch, P. M. (2000). *Modeling survival data: extending the Cox model*. Springer-Verlag, New York.
- [10] Wei, L. J., Lin, D. Y. e Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, Vol. 84, No. 408, pp. 1065–1073.