Report on the doctoral thesis "Neyman's Smooth Tests in Survival Analysis" by Mr. David Kraus, Department of Probability and Mathematical Statistics, Charles University, Prague.

Main ideas

The thesis deals with the idea of smooth tests of goodness of fit due to Neyman (1937). The main feature is that the null hypothesis is embedded in a 'smooth' alternative model determined by a k-dimensional parameter and a set of k linearly independent base functions. The goodness of fit testing problem then translates to a score test for this parameter (zero versus different from zero).
The second main idea is that of data-driven choice of k due to Ledwina (1994). This selection rule is based on maximizing the penalised score statistic.

Main Contribution

The main contribution of the present work is the development of the above ideas in the context of survival data. An extra complication is that the observations are possibly right censored. Neyman's smooth goodness of fit test is reformulated to models which are specified through hazard rate functions.
The idea is not entirely new: an important forerunner is Peña (1998) who dealt with parametric models. The results in the present thesis are new since they deal with nonparametric and semiparametric models in survival analysis.
Chapter 1 deals with testing the null hypothesis of equality of the hazard rates in the two-sample problem. Chapter 2 deals with equality of subdistribution hazard rates in the two-sample competing risk model. Chapter 3 deals with the semiparametric null hypothesis of proportional rates in the two-sample problem (e.g. proportional hazards or proportional odds). Chapters 4-6 deal with data-driven smooth tests for the proportional hazards assumption in the Cox regression model with time varying covariates.

Methodology

The candidate has correctly worked out the asymptotic properties of the proposed tests. He fully masters the high level mathematics of the counting process theory. He is able to find limiting distributions and to use bootstrap techniques in case of complicated covariance structures. He has also consistently done the computational aspects of simulations for comparing the new tests with other existing tests. It should also be noted that the thesis goes together with two software packages.

Interest

The obtained results in this thesis are very relevant for statistical inference in the two sample problem. The applications are in the area of survival analysis but also in the area of reliability theory. The obtained results will lead to publications in good quality international journals. One paper appeared in Lifetime Data Analysis (2007). Another one is accepted for Communications in Statistics-Theory and Methods (2008) and three more papers are submitted.

Conclusion

With this thesis, the candidate has proved that he has the necessary qualifications to become a doctor. He worked creatively on important statistical problems with applications in survival analysis and reliability theory. He masters the mathematical, statistical and computational background and has written up the results in a scientific way. This thesis and the publications that will come out of it will be the definite starting point of a promising career for this junior scientist.

Noël Veraverbeke
Professor
Center for Statistics
University Hasselt
Belgium
noel.veraverbeke@uhasselt.be