

Score test for conditional independence between longitudinal outcome and time-to-event in the joint latent class model

Hélène Jacqmin-Gadda^{1,2}, Cécile Proust-Lima^{1,2}, Jeremy M.G. Taylor³

¹ INSERM, U897, 33076 Bordeaux, France.

² Université Victor Segalen, 33076 Bordeaux, France.

³ Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA

Latent class models have been recently developed for joint analysis of a longitudinal quantitative outcome and a time-to-event [1]. These models assume that the population is divided in G latent classes characterized by different evolution profiles of the marker described by a mixed model with mixture and different risk functions for the event. Compared with the alternative approach for joint modelling that uses shared random effects [2], latent class models have numerical advantages and take the heterogeneity of the population into account. However, selection of the number of classes and model checking are unresolved issues. Especially, the key assumption of conditional independence between the marker and the event given the latent classes is difficult to evaluate because the latent classes are not observed. As a posterior classification may be obtained from the posterior class-membership probabilities computed by the Bayes theorem, it has been proposed to check this assumption by testing the dependence between the two outcomes adjusting for the posterior classes [1] [3].

Using a joint model with latent classes and shared random effects, we develop a score test for the null hypothesis of independence between the marker and the outcome given the latent classes versus the alternative hypothesis that the risk of event depends on one or several random effects from the mixed model. A simulation study is performed to compare the score test performances with those of a test based on a mixed model adjusted on the posterior classes in two cases : when the alternative is correctly specified and when the data are simulated using a 3-class model while the estimated model includes only 2 classes. In both cases, the score test is much more powerful than the test based on the posterior classification. This shows that the score test should be preferred to check the conditional independence assumption and may be used to select the number of latent classes. The methodology is applied to develop a prognostic model for the recurrence of prostate cancer given the evolution of the Prostate Specific Antigen in a cohort of patients treated by radiation therapy.

References

- [1] Lin H, Turnbull BW, Mc Culloch CE, Slate EH (2002). Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association* 97: 53-65
- [2] Henderson R, Diggle P, Dobson A (2000). Joint modelling of longitudinal measurements and event time data 1:465-480
- [3] Bandeen-Roche K, Miglioretti DL, Zeger SL, Rathouz PJ (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association* 92: 1375-1386.