

lcmm: an R package for estimation of latent class mixed models and joint latent class models

Cécile Proust-Lima^{1,2,*}, Benoit Liqueur^{1,2}

1. INSERM U897, Bordeaux, France

2. ISPED, Bordeaux Segalen University, France

*Contact author: cecile.proust@isped.u-bordeaux2.fr

Keywords: classification, joint models, mixed models, growth mixture model, longitudinal data, non-Gaussian data, time-to-event data.

The linear mixed model is routinely used to describe change over time of a quantitative outcome in longitudinal studies. However, it is limited to the analysis of a quantitative Gaussian outcome, in a homogeneous population, and does not handle association with a time-to-event although it is frequent in practice. The R package **lcmm** extends the linear mixed model to (1) the study of heterogeneous populations through the estimation of latent class mixed model, and (2) the joint analysis of longitudinal and time-to-event data through the estimation of joint latent class models. In each case, both Gaussian or non Gaussian quantitative and ordinal outcomes can be analysed.

Latent class mixed models consist in exploring the latent profiles of trajectories in heterogeneous population. They combine the mixed models theory to account for the individual correlation in repeated measures, and the latent class models theory to discriminate homogeneous latent groups when modelling trajectories. Despite a large interest in this approach also known as growth mixture models, implementation in free softwares is very limited. Within the **lcmm** package, the function `hlme` estimates latent class mixed models assuming a quantitative Gaussian outcome and `lcmm` extends this approach to handle non Gaussian quantitative and ordinal outcomes that are very frequent especially in psychological and quality of life studies.

Joint models to analyse jointly longitudinal and time-to-event data have also become increasingly popular in statistics. There exist two kinds of joint models, shared random-effect models in which functions of the random-effects from mixed model are included in the survival model, and joint latent class models which assume the population is constituted of latent classes with a specific longitudinal outcome trajectory and a specific risk of event. While an R package **JM** was recently developed to estimate shared random-effect models, no free software exists for joint latent class models. We propose in **lcmm** the function `jointlcmm` to estimate such joint models, both for Gaussian or non Gaussian quantitative and ordinal outcomes.

Whatever the models, estimation in **lcmm** is performed using a maximum likelihood method using a modified Marquardt algorithm with strict convergence criteria. Only for ordinal outcomes, a numerical integration is required. `hlme`, `lcmm` and `jointlcmm` allow any shape of trajectory, and covariates as predictors of the latent class structure as well as of class-specific trajectories. The joint model includes different baseline functions and common or class-specific covariate effects. Finally, posterior classification and goodness-of-fit measures are provided within post estimation functions.

References

- Proust, Jacqmin-Gadda (2005). Estimation of linear mixed models with a mixture of distribution for the random-effects. *Computer Methods and Programs in Biomedicine*, **78**, 165-73
- Proust-Lima, Joly et al. (2009). Joint modelling of multivariate longitudinal outcomes & a time-to-event: a nonlinear latent class approach, *CSDA*, **53**, 1142-54
- Proust-Lima & Taylor (2009). Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: a joint modeling approach, *Biostatistics*, **10**, 535-49