STATISTICAL SUPPORT

SWISS CENTER FOR AFFECTIVE SCIENCES

Caution with the use of GLMM/GLMER models

E-mail distributed on 04-12-2023

Dear all,

This week's Stat Support is targeted toward the users of **generalized linear mixed models (GLMM)**, or **GLMERs**, as they are known in the R package `lme4`. These are multilevel models that allow non-normal outcomes, such as binary or count data. At present, they appear to be the most popular choice for repeated measures logistic regression.

However, there is an important warning that comes with the interpretation of GLMMs. That is, their coefficients and effects have a subject-conditioned interpretation. For example, let us say we measured a binary outcome at 10 evenly-spaced time points within *N* subjects, and we wish to model the probability of success (1) versus failure (0) over time using, glmer(Y~Time+(1|Subject)), with a binomial distribution for Y. Then the interpretation of the resulting Time coefficient, β_{GLMM}, is as follows:

"Within a subject on average, for a 1-unit increase in time, the odds of success versus failure are expected to multiply by exp(β_{GLMM})."

This interpretation conditions the effect on individual subjects, rather than providing a "marginal" population-level interpretation (across subjects), as one would expect from the corresponding LMM with continuous outcome. In fact, if one were to fit, lmer(Y~Time+(1|Subject)), with a normal distribution for Y, the interpretation of the Time coefficient, β_{LMM}, would be as follows:

"Across subjects on average, for a 1-unit increase in time, the risk difference between success and failure increases by β_{LMM}."

This is an interpretation at the population-level! Whereas both the GLMM and the LMM incorporate subject-conditioning due to their inclusion of a random intercept, subject-variance is collapsed in the LMM because of the linear relationship with the outcome. By contrast, the GLMM has a non-linear link function, the log-odds transformation, which makes it mathematically impossible for β_{GLMM} to collapse across subjects for its interpretation. In fact, this property is shared by all GLMs and known as non-collapsibility. In simpler terms, for a $2 \times 2 \times K$ frequency table, the marginal 2×2 odds ratio cannot be expressed as a weighted sum of the *K* conditional odds ratios. As a consequence, the conditional odds ratio may be different and/or contradict the marginal one, **even in the absence of confounding**!

If the analyst wishes to estimate population-level effects for a repeated categorical outcome, it is recommended to use instead **Generalized Estimating Equations (GEE)**, which is the GLM equivalent of repeated measures (M)ANOVA. GEE has been implemented in the R packages `gee` and `geepack`. Like rm-ANOVA, it requires the specification of a correlation structure for the repeated measures, however for GEE this is only a "working correlation". The final standard errors are robust against model misspecification. The resulting coefficient, β_{GEE}, has the following interpretation:

"Across subjects on average, for a 1-unit increase in time, the odds of success versus failure are expected to multiply by exp(β_{GEE})."

So now we have a proper population effect. Notably, GEE will produce smaller logistic regression coefficients than GLMM for the same data. For the simplest model of one continuous predictor and a random intercept, one can even calculate exactly the factor with which β_{GEE} will shrink relative to β_{GLMM}, and which depends on the size of the estimated within-subject variance (see attached figure). For more general models with multiple covariates and random slopes, such equations are not available, however.

Users of GLMMs are strongly recommended to read up on the distinction between GEE and GLMM in the attached chapter from Fitzmaurice, Laird and Ware (2004), and decide whether they *really* need a model that performs subject-level, versus one that performs population-level inference.

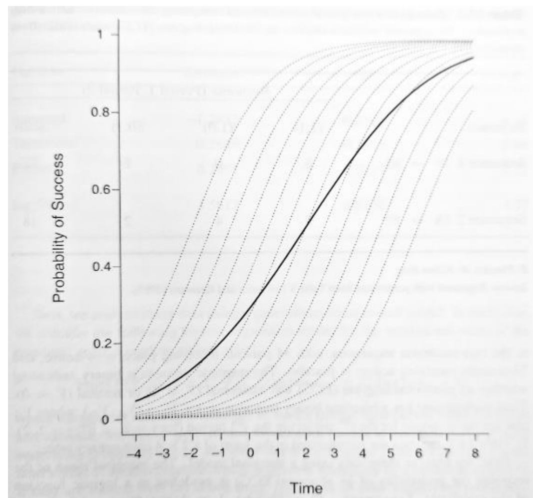

Fig. 13.1 Comparison of conditional probabilities of success (dotted lines) and marginal probability of success (solid line), averaged over the distribution of the random effects.

Figure 1. Marginal versus conditional GLMM coefficient (Fitmaurice, Laird & Ware, 2004)

The above should also make clear that GLMM is neither a straightforward extension of the GLM nor the LMM. Its introduction of random effects is odd when one considers that the standard logistic regression GLM does not have any variance parameter at all (e.g., residual variance), and the GLMM's target of inference is different from the LMM. Thus, introducing random effects to a GLM is not trivial. Such subtleties may be masked by the intuitive interface of functions such as `glmer` from `lme4`, but its users should carefully inform themselves of these models before using them.

Best,
Ben

--
Ben Meuleman, Ph.D.
Statistician
Swiss Center for Affective Sciences
University of Geneva | Campus Biotech
Chemin des Mines 9 | CH-1202 Genève
ben.meuleman@unige.ch | +41 (0)22 379 09 79