HW-3

Qn 1 How does the interpretation of fixed effect coefficients change in a mixed effect model over an 'only' fixed effect model?

Answer 1 *There are two components for any parameter estimate: 1) the estimated value and 2) the standard deviation.*

The fixed effect parameters still quantify the population average effects of different predictors. However, random effects change the covariance structure of the model. This results in the change in standard error of the fixed effect parameters. The estimates may or may not change depending on the model. Since it changes the standard error, the inferences may also change. In an 'only' fixed effect model, some parameters may turn out to be significant, but not in a random effect model or vice versa.

Qn 2 Let $y_{i,j,g}$ be the outcome of *i-th* subject in $t_{i,j}$-*th* time and being treated under the *g-th* treatment group. Consider the following two mixed-effect models:

- $y_{i,j,g} = \alpha + \theta g + \beta t_{i,j} + b_{1,i} + b_{2,i} t_{i,j} + b_{3,i} g + \epsilon_{i,j,g}$ with $b_{j,i} \sim \text{Normal}(0, \sigma_b^2)$ for $j = 1, 2, 3$ and $\epsilon_{i,j,g} \sim \text{Normal}(0, \sigma^2)$.
- $y_{i,j,g} = \alpha + \theta g + \beta t_{i,j} + b_{1,i} + b_{2,i} t_{i,j} + b_{3,g} + \epsilon_{i,j,g}$ with $b_{j,i} \sim \text{Normal}(0, \sigma_b^2)$ for $j = 1, 2$, $b_{3,g} \sim \text{Normal}(0, \sigma_c^2)$, and $\epsilon_{i,j,g} \sim \text{Normal}(0, \sigma^2)$.

Compare the two models in terms of their implications. Establish the correlations and covariances.

Answer 2 *One subject can belong to only one group. However, one group can contain multiple subjects. One of the major differences between the above two models is that the second model induces an across-subject dependence within a given group.*

- *Model 1:* $Cov(y_{i,j,g}, y_{i',j',g'}) = 0$ *for* $i \neq i'$ *for any* j, j' *and* g, g'.
- *Model 2:* $Cov(y_{i,j,g}, y_{i',j',g'}) = \sigma_c^2$ *for* $i \neq i'$ *for any* j, j' *and* $g = g'$.

Model 1 also sets individual-specific random effects for a time-invariant predictor 'group'. Hence, it is useful to change the covariance of the two groups, but the random effect estimates are not identifiable.

Qn 3 Let $y_{i,j}$ be the outcome observed for subject i at time $t_{i,j}$.

- What are the appropriate models for $y_{i,j}$ when $t_{i,j}$'s are not the same for all i?
- Write down the covariance $\text{cov}(y_{i,j}, y_{i,k})$ under the above-mentioned different model choices.

Answer 3 *The appropriate covariance models are Exponential or Gaussian kernel-based covariance models.*

Qn 4 How do the interpretations of a model change with time being a continuous predictor vs it being a factor-valued predictor? What are the major advantages of using factor-valued time as a predictor?

Answer 4 **In the balanced case:** *But the time points may not be gridded, just required to be uniform across all the subjects.*

Using time as a continuous predictor, the 'nature' of the time effect on the outcome will be specific to the choice of model. For example, in this case, the model will look like $y_{i,j} = \alpha + \beta t_{i,j} + \epsilon_{i,j}$ *then the effect of time on the outcome is linear. If* $y_{i,j} = \alpha + \beta t_{i,j} + \theta t_{i,j}^2 + \epsilon_{i,j}$, *then the effect is quadratic.*

But time as a factor can lead to any possible relation between time and the outcome. Here the model looks like $y_{i,j} = \alpha + \beta_j + \epsilon_{i,j}$. *Hence,* β_j *can have any relation with* j. *This type of model also belongs to the class of nonlinear function approximation using step functions.*

This is the reason why orthogonal contrasts are applied to the time as a factor-based model to identify an appropriate polynomial relation.

In case of subject-specific observation times

In this case, time as a continuous predictor is the only option. Otherwise, there will be too many predictors if time is considered as a factor.