schabenberger examines the use and implementation of influence measures in LME models.

Influence is understood to be the ability of a single or multiple data points, through their presences or absence in the data, to alter important aspects of the analysis, yield qualitatively different inferences, or violate assumptions of the statistical model (*schabenberger*).

Outliers are the most noteworthy data points in an analysis, and an objective of influence analysis is how influential they are, and the manner in which they are influential.

schabenberger describes a simple procedure for quantifying influence. Firstly a model should be fitted to the data, and estimates of the parameters should be obtained. The second step is that either single of multiple data points, specifically outliers, should be omitted from the analysis, with the original parameter estimates being updated. This is known as 'leave one out leave k out' analysis. The final step of the procedure is comparing the sets of estimates computed from the entire and reduced data sets to determine whether the absence of observations changed the analysis.

A residual is the difference between an observed quantity and its estimated or predicted value. In LME models, there are two types of residuals, marginal residuals and conditional residuals. A marginal residual is the difference between the observed data and the estimated marginal mean. A conditional residual is the difference between the observed data and the predicted value of the observation. In a model without random effects, both sets of residuals coincide.

schabenberger notes that it is not always possible to derive influence statistics necessary for comparing full- and reduced-data parameter estimates.

Abstract

This paper reviews the use of diagnostic measures for LME models in SAS. This text has been widely cited by texts that don't deal with SAS implementations.

Schabenberger: Summary and Conclusions

- Standard residual and inuence diagnostics for linear models can be extended to linear mixed models. The dependence of xed-effects solutions on the covariance parameter estimates has important ramications in perturbation analysis.

- To gauge the full impact of a set of observations on the analysis, covariance parameters need to be updated, which requires retting of the model.

- The experimental INFLUENCE option of the MODEL statement in the MIXED procedure (SAS 9.1) enables you to perform iterative and noniterative inuence analysis for individual observations and sets of observations.

- The conditional (subject-specic) and marginal (population-averaged) formulations in the linear mixed model enable you to consider conditional residuals that use the estimated BLUPs of the random effects, and marginal residuals which are deviations from the overall mean.

- Residuals using the BLUPs are useful to diagnose whether the random effects components in the model are specied correctly, marginal residuals are useful to diagnose the xed-effects components.
- Both types of residuals are available in SAS 9.1 as an experimental option of the MODEL statement in the MIXED procedure.
- It is important to note that influence analyses are performed under the assumption that the chosen model is correct. Changing the model structure can alter the conclusions. Many other variance models have been t to the data presented in the repeated measures example. You need to see the conclusions about which model component is affected in light of the model being fit.
- For example, modeling these data with a random intercept and random slope for each child or an unstructured covariance matrix will affect your conclusions about which children are inuential on the analysis and how this influence manifests itself.