

Influence arises at two stages of the LME model. Firstly when \mathbf{V} is estimated by \hat{V} , and subsequent estimations of the fixed and random regression coefficients β and u , given \hat{V} .

0.1 Measures of Influence

The impact of an observation, or a case with multiple observations, on a regression fitting can be determined by the difference between the estimated regression coefficient of a model with all observations and the estimated coefficient when the particular observation is deleted. The measure DFBETA is the studentized value of this difference.

0.1.1 Influential Observations : DFBeta and DFBetas

The dfbeta refers to how much a parameter estimate changes if the observation or case in question is dropped from the data set. Cook's distance is presumably more important to you if you are doing predictive modeling, whereas dfbeta is more important in explanatory modeling.

0.1.2 DFBETA

DFBETAS (standardized difference of the beta) is a measure that standardizes the absolute difference in parameter estimates between a (mixed effects) regression model based on a full set of data, and a model from which a (potentially influential) subset of data is removed.

0.1.3 DFBETA

$$DFBETA_a = \hat{\beta} - \hat{\beta}_{(a)} \quad (1)$$

$$= B(Y - Y_{\bar{a}}) \quad (2)$$

In the case of method comparison studies, there are two covariates, and one can construct scatterplots of the pairs of dfbeta values accordingly, both for LOO and LSO calculations. Furthermore 95% confidence ellipse can be constructed around these scatterplots. Note that with k covariates, there will be $k + 1$ dfbetas (the intercept, β_0 , and 1 β for each covariate). For example there would be 2 sets of dfbeta, 510 values for each in the case of LOO, and 85 for LSO diagnostics.

0.1.4 DFFITS

DFFITS is a statistical measure designed to show how influential an observation is in a statistical model.

$$DFFITS = \frac{\widehat{y}_i - \widehat{y}_{i(k)}}{s_{(k)}\sqrt{h_{ii}}}$$

It is closely related to the studentized residual. For the sake of brevity, we will concentrate on the Studentized Residuals.

0.1.5 PRESS

The prediction residual sum of squares (PRESS) is a value associated with this calculation. When fitting linear models, PRESS can be used as a criterion for model selection, with smaller values indicating better model fits.

$$PRESS = \sum (y - y^{(k)})^2 \tag{3}$$

- $e_{-Q} = y_Q - x_Q\hat{\beta}^{-Q}$
- $PRESS_{(U)} = y_i - x_i\hat{\beta}_{(U)}$