

0.0.1 Likelihood Distances

The likelihood distance is a global summary measure that expresses the joint influence of the subsets of observations, U , on all parameters in ϕ that were subject to updating. For classical linear models, the implementation of influence analysis is straightforward. ? points out the likelihood distance gives the amount by which the log-likelihood of the model fitted from the full data changes if one were to estimate the model from a reduced-data estimates. Importantly $LD(\psi_{(U)})$ is not the log-likelihood obtained by fitting the model to the reduced data set. It is obtained by evaluating the likelihood function based on the full data set (containing all n observations) at the reduced-data estimates.

$$LD((U)) = 2[l(\hat{\phi}) - l\hat{\phi}_{\omega}]$$

$$RLD((U)) = 2[l_R(\hat{\phi}) - l_R(\hat{\phi})_{\omega}]$$

An overall influence statistic measures the change in the objective function being minimized. For example, in OLS regression, the residual sums of squares serves that purpose. In linear mixed models fit by maximum likelihood (ML) or restricted maximum likelihood (REML), an overall influence measure is the likelihood distance (?). In LME models, fitted by either ML or REML, an important overall influence measure is the likelihood distance (?). The procedure requires the calculation of the full data estimates $\hat{\psi}$ and estimates based on the reduced data set $\hat{\psi}_{(U)}$. The likelihood distance is given by determining

$$LD_{(U)} = 2\{l(\hat{\psi}) - l(\hat{\psi}_{(U)})\} \tag{1}$$

$$RLD_{(U)} = 2\{l_R(\hat{\psi}) - l_R(\hat{\psi}_{(U)})\} \tag{2}$$

For classical linear models, the implementation of influence analysis is straightforward. However, for LME models, the problem is more complex. Update formulas for the fixed effects are available only when the covariance parameters are assumed to be known. A

measure of total influence requires updates of all model parameters. This can only be achieved in general is by omitting observations or cases, then refitting the model. This is a very simplistic approach, and computationally expensive.

? examines a group of methods that examine various aspects of influence diagnostics for LME models. For overall influence, the most common approaches are the *likelihood distance* and the *restricted likelihood distance*.

The logLik Function

logLik.lme returns the log-likelihood value of the linear mixed-effects model represented by object evaluated at the estimated coefficients. It is also possible to determine the restricted log-likelihood, if relevant, using this function. For the Blood Data Example, the loglikelihood of the JS.roy1 model can be computed as follows.

```
> logLik(JS.roy1)
'log Lik.' -2030.736 (df=8)
```