Log likelihood values and negative deviances

Definition of the deviance

The deviance is defined as $-2\ln(likelihood)$. Let T denotes the number of level-1 records, and f the number of fixed effects. In all the programs, there is a constant added to the likelihood of $\left(\frac{X}{2}\right)\ln(2\pi)$, where $X = T$ for HLM2's full ML, and HLM3 (which is also full ML), and $T - f$ for HLM2's normal REML.

Interpretation of negative deviances

The value of the log likelihood depends on the scale of the data. It is defined as the product of the probability density functions, evaluated at the estimated parameter values. Although the total area under a probability density function is scaled to be equal to 1, this does not imply that the probability density function evaluated at a certain point in parameter space has to be less than 1. The likelihood function can therefore exceed 1, in which case $-2\ln(L)$ (the deviance) will be negative.

Consider the following example:

Suppose x has the normal probability density function

$$f(x) = k \times \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$$

where $k = \frac{1}{\sqrt{2\pi\sigma^2}}$.

If $x = \mu$, then $f(x) = k$. If σ^2 is greater than $\frac{1}{2\pi}$, that is 0.2831, then k will be less than 1. For all other cases, k will be larger than 1. From this follows that $-2\ln(L)$ can be either positive or negative, depending on how well the model fits the data and on the scale of the data.

In terms of real numbers, a model with a deviance of -100 will be evaluated as fitting the data better if the chi-square for comparison with a model with a deviance of, for example, 100 is statistically significant.