The Multilevel First-Order Autoregressive Model: A Bayesian Look at Stability and Sensitivity.

Joran Jongerling and Ellen L. Hamaker
Department of Methods and Statistics
Utrecht University

July 5, 2012

Abstract

The autoregressive (AR) model is a model for longitudinal data that quantifies the stability of constructs, like mood, by regressing the current value of the construct on previous ones. In this study, the focus is on a multilevel extension of this model, the Multilevel First-Order Autoregressive (ML-AR(1)) model. This model allows for the analysis of nested data, that is, data in which observations are nested within objects of study (e.g., amount of (negative) affect nested within individuals). On the first, or observation level of the ML-AR(1) model, separate autoregressive relationships are specified for each object of study. Specifically, if $y_{i,t}$ is object i's score at time-point t and $y_{i,t-1}$ is object i's score on time-point $t-1$, then the equation on level 1 can be written as,

$$y_{i,t} = \phi_i y_{i,t-1} + \epsilon_{i,t}, \tag{1}$$

where ϕ_i is object i's AR-parameter, used to regress its current value of y on the previous one, and $\epsilon_{i,t}$ is a random error term that represents variance in $y_{i,t}$ that cannot be predicted from the previous measurement occasion. These random error terms, called *innovation* in time-series literature, are independent and normally distributed with 0 mean and variance $\sigma^2_{\epsilon,i}$.

On the second, or object level of the ML-AR(1) model, between-object differences in ϕ_i and $\sigma^2_{\epsilon,i}$ are modeled. For example, if we assume that both the AR-parameter and the innovation variance are normally distributed across study objects, the equations on the second level become,

$$\phi_i \sim N(\mu_\phi, \sigma^2_\phi), \text{ and} \tag{2}$$

$$\sigma^2_{\epsilon,i} \sim N(\mu_{\sigma^2_\epsilon}, \sigma^2_{\sigma^2_\epsilon}). \tag{3}$$

The advantage of the ML-AR(1) model is that it allows for these between-object differences in the parameters of the AR model, since they can provide important information. For instance, if researchers investigate the amount of depression of several individuals, then differences in the AR-parameters across these individuals indicate that the amount of depression is more stable in some individuals than in others. In addition, differences in innovation variance are indicative of differential exposure and/or sensitivity to unmodelled internal and external stimuli that influence individuals' amount of depression. The causes for these differences in stability and sensitivity/exposure can subsequently be further investigated by incorporating predictors for ϕ_i and $\sigma^2_{\epsilon,i}$ into the model. For example, some individuals may be more sensitive to the occurrence of negative events or receiving social support than others. And perhaps this difference in sensitivity is linked to personality characteristics.

Although between-unit differences in ϕ_i and $\sigma^2_{\epsilon,i}$ are obviously important sources of information, they cannot be investigated with all statistical software. In fact, the Maximum Likelihood (ML) estimation methods available in standard statistical software can only incorporate between-unit differences in ϕ into a model. We therefore introduce two Bayesian estimation methods for the ML-AR(1) model that allow for more freedom in model specification and that can investigate differences in σ^2_ϵ. The performance of these Bayesian methods, as indicated by the bias and coverage of the parameter estimates, will be compared to three standard ML methods using an extensive simulation study. In addition the minimum requirements regarding number of observations and number of study objects will be determined for the five different methods.