Skewness and staging: Does the floor effect induce bias in multilevel AR(1) models?

Mohammadhossein Manuel Haqiqatkhah, Oisin Ryan & Ellen L. Hamaker

Utrecht University, Faculty of Social and Behavioural Sciences, Department of Methodology and Statistics

Keywords: Floor effect; Experience sampling method; Emotional inertia; Staging effect; Non-Gaussian time series

In the past two decades, the collection of intensive longitudinal data has become increasingly popular in psychological research. To study the dynamics in these data, the multilevel versions of the őrst-order autoregressive (AR(1)) model are often used. It has been suggested that individuals with more severe stages of mental disorders tend to have stronger autoregressions and cross-regressions among certain affective and psychopathology symptom measures. This phenomenon has been referred to as the staging effect.

Two of the main assumptions of the multilevel AR(1) model are level-1 and level-2 normality, which require that individual time series and sample means be normally distributed. However, these two assumptions are often violated in empirical data; importantly, healthier individuals, at many time instances, tend to score very low on negative emotions and symptoms, leading to the floor effect - that is, a high percentage of the responses are equal, or very close, to the lowest value on the scale - which is accompanied by less variability and high skewness.

Using a large-scale simulation study, we investigate the effect of skewness on the estimated autoregressive parameter in the multilevel AR(1) model. To do so, we first provide ways of detecting and characterizing the floor effect in empirical data. We then introduce three novel time series models that can generate skewed continuous and discrete valued responses (for Likert scales and counts data). Finally, we discuss the simulation study we preformed to answer our research question, in which we analyzed these data using the multilevel AR(1) model with fixed and random residual variance. The results indicate that using the more conventional model (with fixed residual variance) leads to negative bias, whereas using the more flexible model (with random residual variance) produces positive bias in the estimated autoregression. We discuss the implications of our study for choosing modeling approaches and data collection.