StatNews #63

Analyzing Longitudinal Data: Developmental Curves or Dynamic Models?
May 2004

Longitudinal studies collect data on multiple individuals over time. They are typically designed with visits at fixed intervals in time, although some designs may record events at any point in time. Longitudinal data are being collected more often because of several advantages. They provide a direct measurement of change and a clear temporal sequence, can help account for the effect of unmeasured variables, and can provide plausible evidence of causal relationships.

Despite the general appreciation for these advantages, there remains considerable confusion about how to analyze longitudinal data. Furthermore, this confusion is enhanced among researchers when experts seem themselves confused. For example, a recent book by Singer and Willett (2003) on the analysis of longitudinal data dismisses in one paragraph one of the two major conceptual approaches to analysis of longitudinal data, concluding erroneously that the approach does not capture change. The purpose of this newsletter is to describe these two approaches, and to explain why both approaches capture change but in different ways.

We will call the two approaches "developmental curve" and "dynamic". These names are not perfect, and several other names are used in the literature. Developmental curves are also called growth curves. Dynamic models are also called transitional or conditional models. In some literatures, however, "dynamic" means something different than is meant here.

In the developmental curve approach, the idea is to fit a curve to each individuals set of responses (i.e., dependent variables). Then, variation in the placement or shape of the curves is related to covariates (i.e., independent variables). This approach examines the effects of covariates on the pattern of responses described by the curves. In the dynamic approach, the response at one time is modeled as a function of the response at an earlier (i.e., lagged) time and covariates at an earlier time. This approach examines the effects of covariates on transitions in response from one time to another.

It is helpful to examine a simple situation to contrast these two approaches. Imagine that the response, a child's cognitive performance, is measured at two points in time, Y1 and Y2, and that we have a measure of parental caring practices at the first time, X1.

The developmental curve approach would be implemented as:

$$Y2 - Y1 = B0 + B1\ X1 + E \quad \text{(equation 1)}$$

B1 is interpreted as the effect of caring practices at time 1 (i.e., X1) on the change in cognitive performance from time 1 to time 2 (i.e., Y2 - Y1).

The dynamic approach would be implemented as:

$$Y2\ C0 + C1\ X1 + C2\ Y1 + E \quad \text{(equation 2)}$$

C1 is interpreted again as the effect of caring practices at time 1 on the change in cognitive performance from time 1 to time 2. The logic is that, after controlling for Y1 in the model, any variation in Y2 that is attributed to X1 must be differential change that occurred between Y1 and Y2. (This is the concept that Singer and Willett misunderstood.) Another way to understand this requires a little bit of algebra. If a term Y1 is subtracted from both sides of equation 2, then we end up with:

$$Y2 - Y1 = C0 + C1\ X1 + (C2-1)\ Y1 + E \quad \text{(equation 3)}$$

Now it is easier to see that C1 is the effect of X1 on the change from Y1 to Y2, while accounting for the initial value, Y1.

In this example, the difference in the two approaches is that the dynamic approach examines change accounting for (i.e., conditional on) the initial value of cognitive performance, whereas the developmental curve approach examines change without accounting for the initial status. The developmental curve approach captures change in cognitive performance and then determines how this change co-varies with caring practices. The dynamic approach estimates how caring practices influence the transition in cognitive performance from time 1 to 2.

Which of the two approaches is most useful in a given situation depends upon how the researcher prefers to view the problem as well as some technical considerations that are beyond the scope of this newsletter. It is also possible to combine aspects of both approaches. Further information can be found in Frongillo and Rowe (1999).

References:

Singer JD, Willett JB (2003) Applied Longitudinal Analysis: Modeling Change and Event Occurrence. Oxford: Oxford University Press.

Frongillo EA, Rowe EM. (1999) Challenges and solutions in using and analyzing longitudinal growth data. Chapter 5. In: Johnston FE, Eveleth P, Zemel B (eds). Human Growth in Context. London: Smith-Gordon, pp. 51-64.

Author: Edward Frongillo

(This newsletter was distributed by the Cornell Statistical Consulting Unit. Please forward it to any interested colleagues, students, and research staff. Anyone not receiving this newsletter who would like to be added to the mailing list for future newsletters should contact us at cscu@cornell.edu. Information about the Cornell Statistical Consulting Unit and copies of previous newsletters can be obtained at http://www.cscu.cornell.edu).