Personality and Individual Differences xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Analyzing dynamic data: A tutorial☆

William Revelle[a,*], Joshua Wilt[b]

[a] Northwestern University, Evanston, IL, United States
[b] Case Western University, Cleveland, OH, United States

ARTICLE INFO

Keywords:
Open source
R
Dynamic data
Repeated measures

ABSTRACT

Modern data collection techniques allow for intensive measurement within subjects. Analyzing this type of data requires analyzing data at the within subject as well as between subject level. Although sometimes conclusions will be the same at both levels, it is frequently the case that examining within subject data will show much more complex patterns of results than when they are simply aggregated. This tutorial is a simple introduction to the kind of data analytic strategies that are possible using the open source statistical language, R.

The study of personality has traditionally emphasized how people differ from each other and the reliability and validity of these differences. This has been reflected in the many publications in this journal and others emphasizing the structure of personality, scale construction, and validation. The typical data collected emphasized the "R" approach of Cattell's data box (Cattell, 1946a, 1966), that is, correlating how participants differ across items/tests. Cattell's data box also included the possibility of studying how one person varied over time ("P"). Sometimes the approach would consider stabilities across time as measured by the correlation of measures taken at two different time points ("S"). One of the more impressive stabilities is the correlation of .56 over 79 years of IQ scores from age 11 to age 90 (Deary, Pattie, & Starr, 2013). An example of what Cattell referred to as a diagonal in his data box would be the correlation across time of individuals taken on different measures. A powerful example of this is the prediction of health related outcomes in middle age from teacher ratings of students in grades 1–6 (Hampson & Goldberg, 2006).

In the past 30 years or so, we have seen an exciting change in the way we collect data, in that we now can study how individuals vary over time (Cattell's P approach). To Cattell, this was "the method for discovering trait unities" (Cattell, 1946b, p 95). The emphasis is now upon individual variability with the added complexity of how these patterns of individual change differ across participants (e.g., Bolger & Laurenceau, 2013; Mehl & Conner, 2012; Wilt, Funkhouser, & Revelle, 2011; Wilt, Bleidorn, & Revelle, 2016). Although the methods were originally developed to examine data with a nested structure (e.g., students nested within classes nested within schools Bryk & Raudenbush, 1992), the use of these techniques across many occasions within individuals has been labeled *Intensive Longitudinal Methods* (Walls & Schafer, 2006) and "captures life as it is lived" (Bolger, Davis, & Rafaeli, 2003). We refer to data that show systematic variation over time as dynamic to distinguish them from static cross sectional data. Formal models that distinguish between dynamic patterns versus stochastic variation (Revelle & Condon, 2015) are beyond the scope of this paper. Although it is possible to examine group patterns over time, it is more typical to consider how individuals differ in their patterning across time.

Analytic strategies for analyzing such multi-level data have been given different names in a variety of fields and are known by a number of different terms such as the random effects or random coefficient models of economics, multi-level models of sociology and psychology, hierarchical linear models of education or more generally, mixed effects models (Fox, 2016). Although frequently cautioned not to do so, some psychologists continue to use a repeated measures analysis of variance approaches rather than the more accurate mixed effects models.

The analysis of data at multiple levels presents at least two challenges, one is that of interpretation, the other is that of statistical inference. It has long been known (Yule, 1903) that relationships found within groups are not necessarily the same as those between groups. Although when aggregating across British health districts, it appeared that increased mortality was associated with increases in vaccinations, when examined at the within district level, it was clear that vaccinations reduced mortality (Yule, 1912). Variously known as Simpson's paradox (Simpson, 1951), or the ecological fallacy (Robinson, 1950), the observation is that relationships of aggregated data do not imply the same relationship at the disaggregated level. Such results are examples of non-ergodic relationships, that is, relationships that differ from the individual to the group level (Molenaar, 2004; Nesselroade & Molenaar, 2016).

☆ Preparation of this manuscript was funded in part by grant SMA-1419324 from the National Science Foundation to WR. This is the authors' version as submitted to PAID. We gratefully acknowledge Aaron Fisher for making his data set publicly available.

* Corresponding author at: Department of Psychology, Northwestern University, Evanston, IL 60208, United States.
E-mail address: revelle@northwestern.edu (W. Revelle).

http://dx.doi.org/10.1016/j.paid.2017.08.020
Received 13 May 2017; Received in revised form 7 August 2017; Accepted 11 August 2017

More importantly, when the effect of levels is ignored, structural relationships are difficult to interpret. The correlation between two variables (x and y) when x and y are measured within individuals is a function of the correlation between the individual means ($r_{xy_{between}}$), the pooled within individual correlations ($r_{xy_{wihin}}$) and the relationships between the data and the between group means $\eta_{between}$ as well as the the correlation of the data within the within subject means η_{within}.

$$r_{xy} = \eta_{x_{within}} * \eta_{y_{within}} * r_{xy_{within}} + \eta_{x_{within}} * \eta_{y_{between}} * r_{xy_{between}}. \quad (1)$$

Classic examples of this phenomenon other than Yule's vaccination data include bias in graduate admissions as well as effective tax rates. While the overall admissions rate at the University of California suggested a bias against women, when the data were disaggregated and examined at the department level, this effect actually reversed (Bickel, Hammel, & O'Connell, 1975); tax rates can decrease across all income groups even though total taxes increase (Wagner, 1982) as and tutorials have started appearing. Bolger and Laurenceau (2013) provide an excellent book reviewing methods for analyzing this kind of data and include examples in four of the standard data processing systems (MPLUS, SPSS, SAS, and R). Of these four, only the last one is not proprietary and advances the concept of open source software. More importantly in this era of conducting reproducible research (Leek & Jager, 2017) R facilitates the dissemination of reproducible statistical code.

If not already, R is well on its way to becoming the lingua franca of statistical analysis. It is open source, free, and extraordinarily powerful. Most importantly, more and more *packages* are being contributed to core R (R Core Team, 2017). As of this writing there are at least 11,000 packages that add to the functionality of R. Given our commitment to open science and the use of open source software, we devote this tutorial to how to use R for simulating and analyzing the intensive longitudinal data that is frequently found in the study of individual differences. We rely heavily on the work of Bolger and Laurenceau (2013) as well as the software manuals for four very powerful R packages (Bates, Mächler, Bolker, & Walker, 2015; Bliese, 2016; Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2016; Revelle, 2017). We use a "toy" data set of Shrout and Lane (2012), an open data set released by Fisher (2015), as well as some simulations using the `sim.multi` function. We emphasize an exploratory data approach using graphical displays and a confirmatory approach using a few of the more commonly used R packages.

What is R and how to use it? R is a data analysis system that is both open source and is also extensible. By open source, we mean that the actual computer code behind all operations is available to anyone to examine and to reuse, within the constraints of the GPL 2.0 (GNU General Public License, 1991). It is free software in the meaning of free speech in that everyone can use it, everyone can examine the code, everyone can distribute it, and everyone can add to it. R may be downloaded for free from the Comprehensive R Archive Network (CRAN which may be found at https://cran.r-project.org) and is available for PCs, MacOS, and Linux/Unix operating systems. For purposes of speed, much of core-R is written in Fortran or C++, but most of the packages for R are written in R itself. For R is more than a statistical system, it is a programming language. This means that R is extensible in that anyone can add *packages* to the CRAN as well as other repositories such as GitHub or BioConductor (http://bioconductor.org). CRAN has certain quality assurance tests that guarantee that the contributed programs have consistent documentation, including examples, and will not fail while running these examples. CRAN does not check the validity or utility of submitted packages, that is up to the contributor as well as the users of the packages. As of this writing, several thousand contributors have added on more than 11,000 *packages* to core-R and this number increases daily.

R was originally developed between 1992 and 1995 by Ross Ihaka and Robert Gentleman at the University of Auckland as a way to implement the S computer language for MacIntosh computers. They were soon joined by others around the world to enhance the development and distribution of R. There are about 20 primary program developers of "Core R" (R Core Team, 2017) who take responsibility for maintaining and distributing the basic system. This is a very eclectic group in that its members come from all over the world.

What makes R so powerful is the programming philosophy of core-R as well as the packages. Rather than give voluminous output for each function, the functions display only the most important aspects of the analysis, and save additional results as elements of the returned object. These objects may then be processed by additional functions. The power of this implementation is that specialized packages can take advantage of the more general core-R features. Thus, the correlation function (`cor`) can be used by functions that do factor analysis (`fa`) and the `mean` function can be used for a function to basic descriptive statistics (`describe`), which can be combined with the `by` function to do statistics broken down by groups (`describeBy`) or be combined again with functions that do correlations, to provide some basic multilevel statistics (`statsBy`). Without much effort, standard functions such as `aov` which does ANOVA, or `lme` to do linear mixed effects models can be integrated into other functions to find, for instance, intra class correlations (`ICC`) or multilevel reliability (`multilevel.reliability`). These functions in turn, may be used by the end user by just giving one or two commands. Sometimes terse and sometimes extensive, "help" files for each function are included for all functions for all packages. In the appendix to this article, we include the specific commands for example that we give. In the text we prefer to give a more high level summary of the necessary operations. Because there are so many useful texts and web-based tutorials on R it is hard to suggest any particular one. A very short introduction to R is the *Introduction to R* by Venables, Smith, and the R development core team (2017) which is available as a book for a fee, or as a pdf to download from the web for free.

1. The basic model

A typical psychological research problem that requires multilevel modeling is the study of how people differ in the pattern of their feelings, thoughts and behaviors over time and place. That people differ is not the question, but rather are these differences systematic and how best to describe them. The analysis could be examining patterns of affect or behavior over time (Fisher, 2015; Fisher & Boswell, 2016), or how people differ in the emotional responses as a function of the situation (Wilt & Revelle, 2017a; Wilt, 2017b) or how couples relationships change over time (Rubin & Campbell, 2012).

The basic concept of multilevel modeling of dynamics is to decompose variation between individuals and within individuals. While the within individual variability is usually treated as error in conventional analysis of variance, it is this within subject variability that is the essence of multilevel modeling: the analysis of how individuals differ in their pattern of responses over time and how these differences may, in turn, be modeled. For, if we measure individuals over multiple occasions, we can also find the within person mean and variance over time, the within subject correlation of measures over time, and the within person correlation of multiple measures. Thus, we can describe each individual's unique signature over time and space (Hamaker, Ceulemans, Grasman, & Tuerlinckx, 2015; Hamaker, Grasman, & Kamphuis, 2016; Hamaker & Wichers, 2017).

Let **X** represent our data, with an individual observation $x_{ijk}$with subscripts *i,j,k* to represent subjects, measures, and time. We can find the overall mean μ and variance σ^2, and decompose these into a function of the within person mean over time for each variable $\mu_{ij.}$ and variance $\sigma^2_{ij.}$. The *between* subject covariances $\sigma_{.j_1,j_2}$ represent the