Paper ###-2018

Modeling Heterogeneous Within-Subject Variability

Madeline Craft, University of California, Davis; Shelley Blozis, University of California, Davis

ABSTRACT

The development of smart phones, which can conveniently collect intensive longitudinal data, called ecological momentary assessment (EMA) data, has allowed researchers to better explore dynamic within-subject (WS) and between-subject (BS) processes. Typically, multilevel models are applied to EMA data to account for the hierarchical nature of longitudinal data. Whereas multiple regression estimates a single average trajectory for all subjects, multilevel models account for the unique influence of an individual on his or her repeated measurements by estimating an average trajectory for each subject.

Multilevel models can be extended to not only summarize differences in average trajectories but also differences in variances about the average trajectories. This extension of the basic multilevel model, which has been called the location scale model in recent years (see Hedeker, Mermelstein & Demirtas, 2008, 2012; Rast, Hofer & Sparks, 2012), utilizes a log-linear representation of the residual error variance to capture the non-negative nature of variances. Although log-linear models of residual error variances have been around for some time (see Aitkin, 1987; Harvey, 1976), there has been resurging interest in these models due to the recent abundance of EMA data.

SAS® PROC MIXED can be used to estimate models of the WS variance, but SAS® PROC NLMIXED is necessary for models of the BS variance. In addition, SAS® PROC NLMIXED is necessary to fit a WS subject variance model with a random scale effect. Fitting the full location scale model in SAS® PROC NLMIXED is done in stages. This paper is intended for an audience with any level of SAS proficiency and some interest in longitudinal data methods, however it is expected that the audience has at least a basic knowledge of multilevel models. It will not be assumed that the audience has experience with SAS® PROC MIXED or NLMIXED.

INTRODUCTION

The development of smart phones, which can conveniently collect intensive longitudinal data, called Ecological Momentary Assessment (EMA) data, has allowed researchers to better explore dynamic within-subject (WS) and between-subject (BS) processes. Longitudinal data consists of three or more repeated measurements from each individual (Singer & Willett, 2003), while EMA data typically consists of 30 to 40 repeated measurements from each individual (Hedeker, Mermelstein & Demirtas, 2012). Researchers often fit multilevel models to longitudinal data, however an extension of the multilevel model, which has been called a location scale model in recent years (see Hedeker, Mermelstein & Demirtas, 2008, 2012; Rast, Hofer & Sparks, 2012), can be fit to EMA data.

THE MULTILEVEL MODEL

The multilevel model can accommodate as many levels as the data require. In this paper we will represent the multilevel model as a simple two-level model. Level one of a multilevel model with a single time-varying predictor can be written as:

$$Y_{ij} = \beta_{0i} + \beta_{1i}X_{ij} + \varepsilon_{ij}.$$

Y_{ij} represents a longitudinal outcome where repeated measures j are nested within individual i. β_{0i} is a random intercept and has a subscript i which allows individuals to deviate from the average initial status. When subscripted, we say that the model includes a random intercept effect. β_{1i} represents the effect of X_{ij} on the outcome and has a subscript i which allows X_{ij} to have a different effect on each individual. When subscripted, we say that the model includes a random slope effect.

The WS errors ε_{ij} represent individual deviations from each individual's respective fitted trajectory and are assumed to be normally distributed with a mean of zero and homoscedastic variance:

$$\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2).$$

The assumption of homoscedastic variance can be relaxed so that, within an individual, variances are correlated and heteroscedastic across time. This is done by fitting an alternative error covariance structure and is demonstrated later in the paper. The SAS documentation refers to the matrix of WS variances and covariances as the **R** matrix and may refer to the level one residual effects as the "R-side effects". All R-side modeling is specified in the REPEATED statement of SAS® PROC MIXED.

Level two of a multilevel model with a single time-invariant predictor can be written as:

$$\beta_{0i} = \gamma_{00} + \gamma_{01}Y_i + \zeta_{0i}$$
$$\beta_{1i} = \gamma_{10} + \gamma_{11}Y_i + \zeta_{1i}$$

The intercept and slope parameters with random effects from level one become the outcome variables at level two. γ_{00} represents the average intercept for all individuals. γ_{01} represents the average effect of some person-level predictor on the individual's intercept. γ_{10} represents the average slope for all individuals. γ_{11} represents the average effect of some person-level predictor on the individual's slope. The γ symbols are referred to as the fixed effects. ζ_{0i} and ζ_{1i} are the random effects and represent BS deviations from the average intercept and slope, respectively. The random effects are assumed to be multivariate normally distribution with means of zero and a correlated and heteroscedastic covariance matrix:

$$\begin{bmatrix}\zeta_{0i}\\ \zeta_{1i}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\ 0\end{bmatrix}, \begin{bmatrix}\sigma^2_{\zeta 0} & \sigma_{01}\\ \sigma_{10} & \sigma^2_{\zeta 1}\end{bmatrix}\right).$$

The SAS documentation refers to the matrix of BS variances and covariances as the **G** matrix and may refer to the level two random effects as the "G-side effects". All G-side modeling is specified in the RANDOM statement of SAS® PROC MIXED.

ALTERNATIVE ERROR COVARIANCE STRUCTURES

Although the default assumption in SAS is that the **R** matrix contains independent and homogeneous variances, this is an unlikely assumption with longitudinal data where we would expect that repeated measurements within an individual are influenced by the individual over time. As mentioned earlier, one way to deal with this heterogeneity is to specify an alternative error covariance structure. A complete list of built-in alternative structures can be found in the SAS documentation. Some of the structures on that list include the autoregressive structures, Toeplitz, and compound symmetry.

An autoregressive structure is appropriate for homoscedastic or heteroscedastic error variances with correlations that decrease exponentially with time. A Toeplitz structure is appropriate for homoscedastic or heteroscedastic error variances with correlations that also decrease with time, but do not decrease as a function of the time period before. A compound symmetry structure is appropriate for homoscedastic or heteroscedastic error variances with constant correlations across time. The default assumption of independent and homogeneous error variances is called the variance component structure.

Appropriate specification of an alternative error covariance structure can improve model fit and produce unbiased parameter estimates (Blozis, S.A & Harring, J.R., 2014). Specification is carried out in SAS® PROC MIXED within the RANDOM statement using the TYPE= option.

THE LOCATION SCALE MODEL

'Location' and 'scale' are terms that have been used by statisticians to refer to the mean and standard deviation, respectively (see Huber, 1972; Oja 1981). Therefore, multilevel models summarize differences in location, and the location scale model summarizes differences in location and scale. A location scale model can also be adjusted to allow subjects to vary in the magnitude of the error variance beyond just the influence of the predictors through the inclusion of a random scale effect. While the idea has been presented in the past, Hedeker and his colleagues (Hedeker, Mermelstein & Demirtas, 2008; Hedeker, Demirtas & Mermelstein, 2009) have been among the first to implement random scale effects in the residual error variance models (Li & Hedeker, 2012).

The BS portion of a location scale model with a single time-invariant predictor can be written as:

$$\sigma^2_{\zeta i} = \exp(\alpha_0 + \alpha_1 U_i).$$

$\sigma^2_{\zeta i}$ is the variance of the BS errors (or the random effects). α_0 is an intercept term, and $\alpha_1$represents the effect of U_i on the outcome. Predictors of the BS error variance can only be time-invariant. This log-linear representation of the WS error variance is adopted in order to model the distribution of the variances, which are never allowed to be negative (Aitkin, 1987; Cook & Weisberg, 1983; Harvey, 1976; Hoffman, 2007; Leonard, 1975; Lin et al., 1997).

The WS portion of the location scale model with a single time-varying predictor can be written as:

$$\sigma^2_{\varepsilon ij} = \exp(\tau_0 + \tau_1 Z_{ij}).$$

$\sigma^2_{\varepsilon ij}$ is the variance of the WS errors. τ_0 is an intercept term, and τ_1 represents the effect of Z_{ij} on the outcome. Predictors of the WS error variance can be either time-varying or time-invariant. The model of the variance of the WS errors has even been extended in recent years to include a random scale effect, which is much like a random location effect in that it allows individuals to differ from the average WS error variance beyond just the influence of the predictors (Hedeker & Mermelstein, 2007; Hedeker, Mermelstein, Berbaum & Campbell, 2009; Hedeker, Mermelstein & Demirtas, 2008, 2012). This extended model can be written as:

$$\sigma^2_{\varepsilon ij} = \exp(\tau_0 + \tau_1 Z_{ij} + \omega_i).$$

PAPER OUTLINE

The remainder of this paper outlines a scenario in which a location scale model would be appropriate and summarizes the steps one would take to fit such a model using an applied example. Our example explores trajectories of negative affect using a sample of 2,022 individuals who completed the second wave of the National Study of Daily Experiences (NSDE), which is a project from the Midlife Development in the United States Survey (MIDUS). These individuals provided negative affect measures on eight consecutive days. Instructions and code are provided for many of the methods described above.

EXAMPLE

We begin by specifying the following unconditional multilevel model using SAS® PROC MIXED:

$$NA_{ij} = \beta_0 + \zeta_{0i} + \epsilon_{ij}.$$

The unconditional model includes a random intercept effect ζ_{0i} to capture the influence of an individual i on his or her repeated measurements of negative affect. The intraclass correlation coefficient (ICC) gives a sense of the degree of association among the repeated measurements specific to each individual. We calculate an ICC equal to 0.52, which can be interpreted to mean that 52% of the total variability in negative affect is due to the influence of individuals on their repeated measurements.

After estimating the unconditional model to obtain the ICC, sex is added as a predictor of negative affect. The code used to fit this model is as follows:

```
proc mixed data = mydata method = ML;
  class ID;
  model NA = FEM / solution;
  random intercept / sub = ID type = un;
  repeated / sub = ID type = vc;
run;
```

ID is a variable denoting the subject identifier, NA is a scale score for negative affect, and FEM is a binary variable where 1 denotes a female subject and 0 denotes a male subject. The TYPE= option on the RANDOM statement is set to “un” for unrestricted, and the TYPE= option on the REPEATED statement is set to “vc” for variance component. Each of these TYPE= options align with the basic multilevel model assumptions of correlated and heterogeneous BS random effects and independent and homogeneous WS residual effects, respectively. As mentioned before, modeling of the G-side effects is always specified in the RANDOM statement, and modeling of the R-side effects is always specified in the REPEATED statement.

The following code is used to fit the same model, but with an autoregressive WS error covariance structure rather than a variance component structure:

```
proc mixed data = mydata method = ML;
  class ID;
  model NA = FEM / solution;
  random intercept / sub = ID type = un;
  repeated / sub = ID type = AR(1);
run;
```

In the model with the variance component structure the AIC and BIC are equal to 2492.5 and 2503.7, respectively. In the model with the AR(1) structure the AIC and BIC are equal to 2055.1 and 2071.9, respectively. The adoption of more appropriate assumptions regarding the error covariance structure improved the fit of the model.

Next, we supply figures to aid in the visualization of the WS and BS variability. Figure 1 shows trajectories of negative affect across eight days for a random subsample of six individuals. The variability (the dots) about the individual fitted trajectory (the straight line) is the WS error variance that can be modeled by a location scale model. Figure 2 shows overlaid trajectories of negative affect across the eight days for a random subsample of eight individuals. The variability in intercepts (the dots at day one) about the average intercept (the straight line) is the BS random intercept effect variance that can be modeled by a location scale model. Note that the trajectories in Figure 2 are not fitted trajectories, and the average intercept is just an average of the intercepts and not a fitted intercept, but the figure, nevertheless, gives the reader a visual sense of the BS variability.

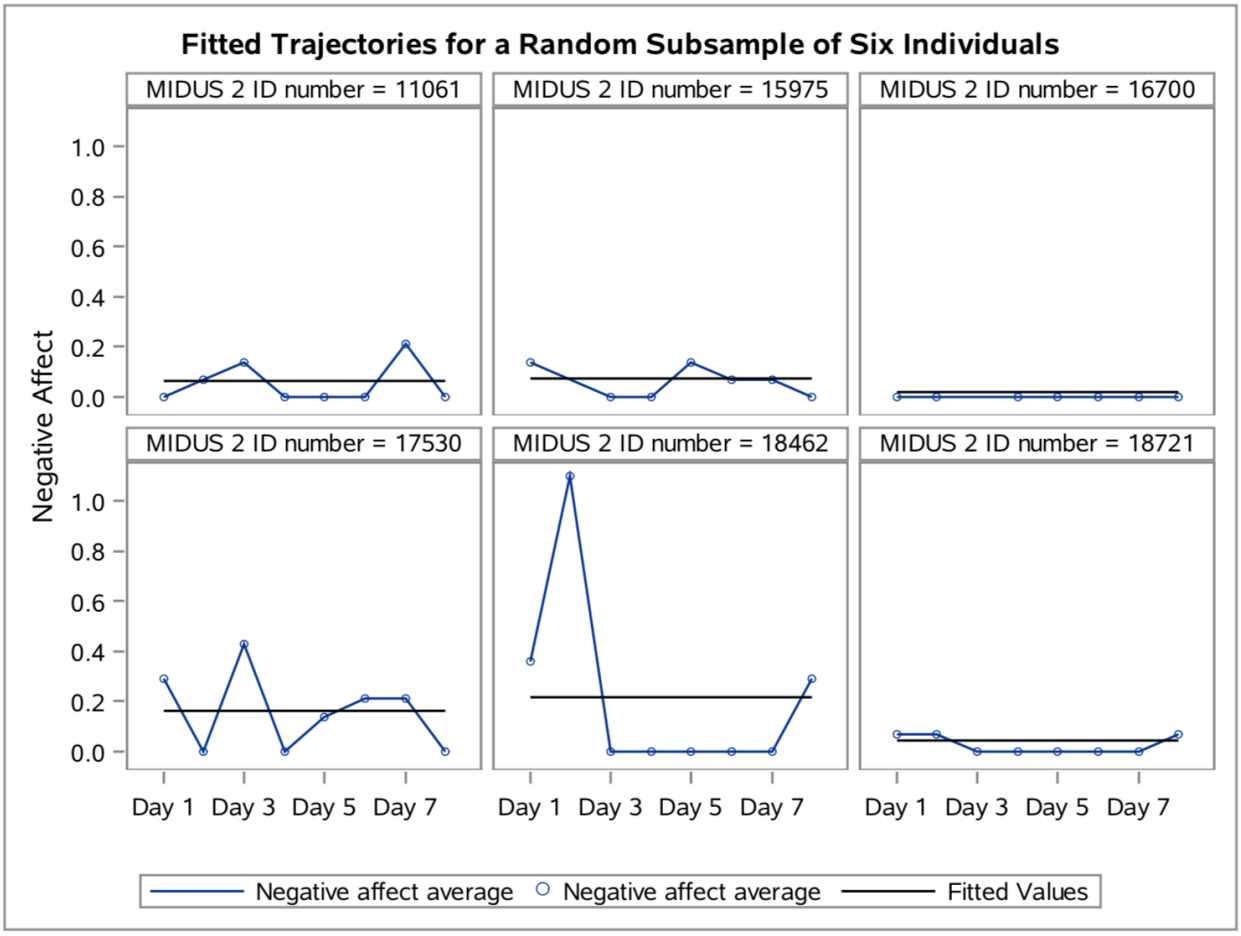

Figure 1. Trajectories of Negative Affect Across the Eight Days for a Random Subsample of Six Individuals.

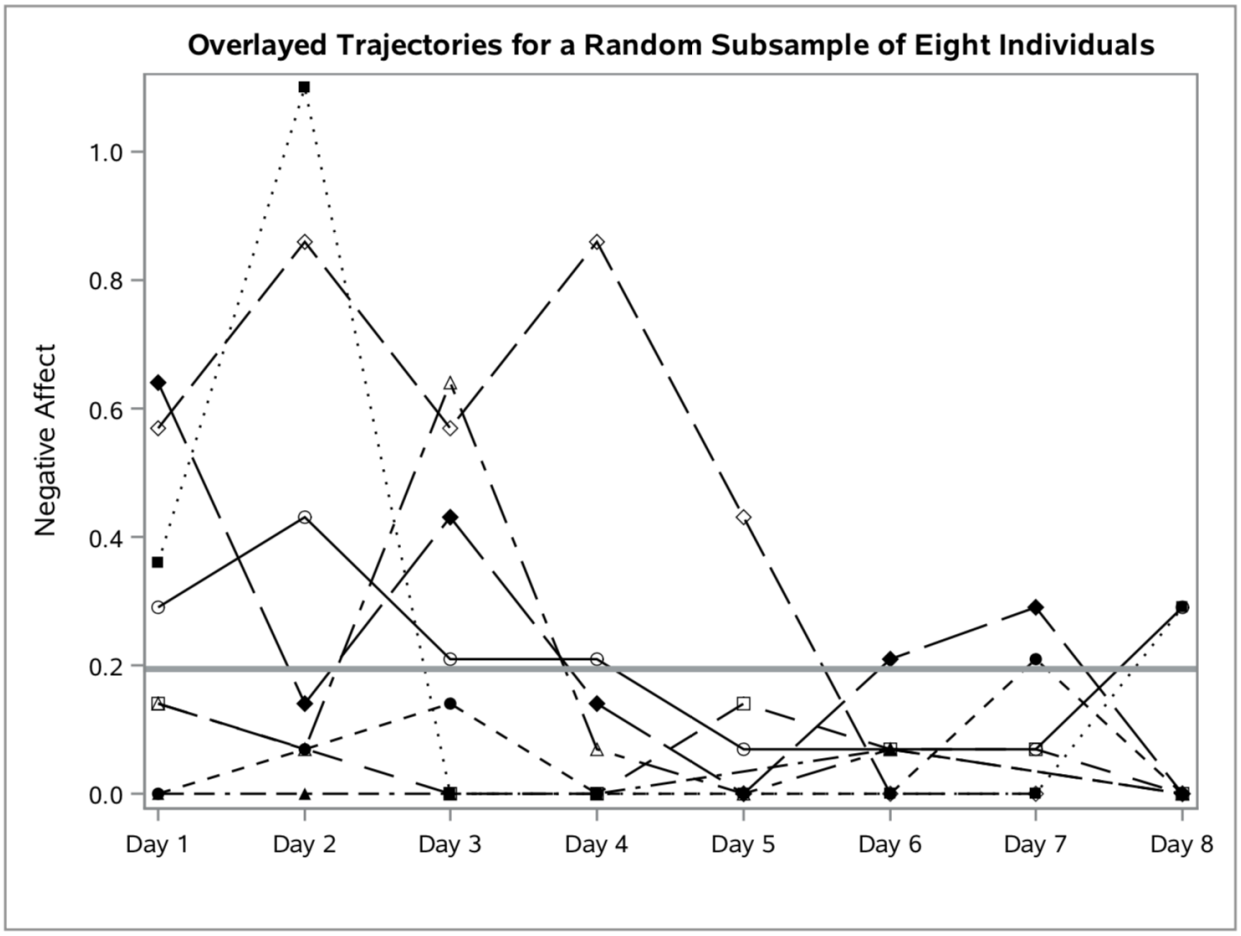

Figure 2. Overlaid Trajectories of Negative Affect Across the Eight Days for a Random Subsample of Eight Individuals.

The following code is used to fit a model of negative affect predicted by sex along with a model of the WS error variance predicted by sex:

```
proc mixed data = mydata method = ML;
  class ID FEM;
  model NA = FEM / solution;
  random intercept / sub = ID type = un;
  repeated / sub = ID type = AR(1)local = EXP(FEM);
run;
```

The above code produces a location scale model with the results presented in Table 1. The fixed effects are interpreted in the usual way, but the effect of sex on the WS error variance must be exponentiated to return the units to an interpretable scale. Being female, in comparison to being male, results in a 2.26 unit increase in the amount of WS error variance.

To add a random effect to the previous model we must work up to the full model in stages using SAS® PROC NLMIXED. Starting values are required by SAS® PROC NLMIXED, so we start by fitting the previous model using the estimates from Table 1, but this time in SAS® PROC NLMIXED. The following code is used to specify this model:

```
proc nlmixed data = mydata gconv = 1e-12 maxiter = 10000 noad qpoints = 1;
  parms gamma00 = 0.1834 gamma01 = 0.0356 lnvare = -3.0285 tau1 = -0.2031;
  z = gamma00 + gamma01*FEM;
  vare = EXP(lnvare + tau1*FEM);
  model NA ~ Normal (z, vare);
run;
```

Using the estimates obtained from this code as starting values, we can now specify the full model using the following code:

```
proc nlmixed data = mydata gconv = 1e-12 maxiter = 10000 noad qpoints = 1;
  parms gamma00 = 0.4708 gamma01 = 0.0019 lnvare = -3.1424 tau1 = 0.3621
  varu1 = 0.0652 varw = 6.7494;
  z = gamma00 + gamma01*FEM + u1;
  vare = EXP(lnvare + tau1*FEM + w);
  model NA ~ Normal (z, vare);
  random u1 w ~ Normal ([0, 0], [varu1, 0, varw]) subject = ID;
  bounds varw >= 0;
run;
```

To aid in the interpretation of the above code, we present the equation for the model below:

$$NA_{ij} = \gamma_{00} + \gamma_{01}FEM_i + \delta_{0i} + \varepsilon_{ij}$$
$$\sigma^2_{\varepsilon_{ij}} = \exp(\tau_0 + \tau_1 FEM_i + \omega_i).$$

In the code, γ_{00} is represented by "gamma00", $\gamma_{01}$is represented by "gamma01", $\ln(\sigma^2_{\varepsilon_{ij}})$ is represented by "lnvare", τ_1 is represented by "tau1", $\sigma^2_{\zeta i}$ is represented by "varu1", and $\sigma^2_{\omega i}$ is represented by "varw". The random effects δ_{0i} and ω_i are assumed to be normally distributed with means of zero, and independent and heteroscedastic variances. The BOUNDS statement ensures that the estimate of the variance of the random effect ω_i remains positive. The results of this model are presented in Table 2. All estimates are interpreted as they were before, but now individuals are allowed to differ in their WS error variance beyond just the influence of sex.

Fixed Effect	***Estimate***	***P value***
γ_{00}	0.1834	<.0001
γ_{01}	0.0356	.0024

Random Effect	***Estimate***	***P value***
$\sigma^2_{\delta_{0i}}$	0.0611	<.0001
τ_1	0.8162	<.0001
$\sigma^2_{\varepsilon_{ij}}$	0.0484	<.0001

Table 1. Fixed and Random Effect Estimates from a Model of the WS Error Variance.

Fixed Effect	***Estimate***	***P value***
γ_{00}	0.1369	<.0001
γ_{01}	0	.7318

Random Effect	***Estimate***	***P value***

$\sigma^2_{\delta_{0i}}$	0	<.0001
τ_1	0.2257	<.0001
$\sigma^2_{\varepsilon_{ij}}$	0.0135	<.0001
$\sigma^2_{\omega_i}$	6.5081	<.0001

Table 2. Fixed and Random Effect Estimates from a Model of the WS Error Variance with a Random Effect.

CONCLUSION

The goals of this paper were to present methods for dealing with heterogeneous variance, outline a scenario in which a location scale model would be appropriate, and summarize the steps one would take to fit such a model using SAS. There are many potentially useful applications of this model. In the case of bipolar depression, for example, where palliative rather than curative strategies may be the best treatment option, perhaps the scale of a subject's trajectory is more relevant than the location. We can include time-varying and person-level predictors in the WS and BS variance models to study the relationship between predictors and the magnitude of the WS and BS variances. For example, a treatment may be effectively stabilizing extreme fluctuations in a patient's depressive symptoms even if it is not effectively reducing the overall level of the symptoms. If the WS variance is not modeled, there may be a relationship between a treatment and the magnitude of variation about the average depressive symptoms trajectory that is overlooked if the average trajectory does not show improvement. In other words, a treatment may be effectively stabilizing the extremes in a patient's symptoms even if it is not effectively reducing the average level of the symptoms. A stabilizing treatment such as this would surely be a valuable addition to the treatment options available for individuals with bipolar depression.

While background information on specifying a BS variance model was presented, an example was not provided. A BS model would be specified in much the same way as a WS model, but with only time-invariant predictors, using SAS® PROC NLMIXED (see Hedeker, Mermelstein, & Demirtas, 2012). Location scale models of both the WS and BS variances can provide rich information that basic multilevel models of the mean cannot.

REFERENCES

Aitkin, M. (1987). Modelling variance heterogeneity in normal regression using GLIM. Applied Statistics. 36(3), 332.

Cook, D. R., & Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. Biometrika. 70(1), 1-10.

Harring, J. R., & Blozis, S. A. (2014). Fitting correlated residual error structures in nonlinear mixed-effects models using SAS PROC NLMIXED. Behavior research methods, 46(2), 372-384.

Harvey, A. C. (1976). Estimating regression models with multiplicative heteroscedasticity. Econometrica, 44(3), 461.

Hedeker, D., Demirtas, H., & Mermelstein, R.J. (2009). A mixed ordinal location scale model for analysis of ecological momentary assessment (EMA) Data. Statistics and its Interface. 2, 391-401.

Hedeker, D., & Mermelstein, R. J. (2007). Mixed-effects regression models with heterogeneous variance: analyzing ecological momentary assessment data of smoking. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), Modeling Ecological and Contextual Effects in Longitudinal Studies of Human Developmen. (pp. 183–206). Mahwah, NJ: Erlbaum.

Hedeker, D., Mermelstein, R. J., Berbaum, M. L., & Campbell, R. T. (2009). Modeling mood variation associated with smoking: an application of a heterogeneous mixed-effects model for analysis of ecological momentary assessment (EMA) data. Addiction, 104(2), 297-307.

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) Data. Biometrics. 64(2), 627–634.

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2012). Modeling between- and within-subject variance in ecological momentary assessment (EMA) data using mixed-effects location scale models. Statistics in Medicine. 31(27), 3328-3336.

Hoffman, L. (2007). Multilevel models for examining individual differences in within-person variation and covariation over time. Multivariate behavioral research. 42(4), 609-629.

Huber, P. J. (1972). The 1972 wald lecture robust statistics: A review. The Annals of Mathematical Statistics, 43(4), 1041-1067.
Leonard T. (1975). A bayesian approach to the linear model with unequal variances. Technometrics. 17(1), 95–102.
Li, X. & Hedeker, D. (2012). A three-level mixed-effects location scale model with an application to ecological momentary assessment (EMA) data. Statistics in Medicine. 31(26), 3192-3210.
Lin, X., Raz, J., Harlow, S.D., (1997). Linear mixed models with heterogeneous within-cluster variances. Biometrics. 53(3), 910-23.
Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. Scandinavian Journal of statistics, 154-168.
Rast, P., Hofer, S. M., & Sparks, C. (2012). Modeling individual differences in within-person variation of negative and positive affect in a mixed effects location scale model using BUGS/JAGS. Multivariate Behavioral Research, 47(2), 177-200.
Singer, J. D., & Willett, J. B. (2003). Applied longitudinal data analysis: modeling change and event occurrence. New York: Oxford University Press.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Madeline Craft
University of California, Davis
mcraft@ucdavis.edu