Multilevel Models

Philippe Rast

PSC 204B
UC Davis, Winter 2018

Topics

Single-level Regression:

Week 1 Linear Regression (G&H: 3,4)
Week 2 Multiple Regression
Week 3 Violation of Assumptions
Week 4 Logistic Regression and GLM (G&H: 5, 6)
Week 5 Over-fitting, Information Criteria and Model comparison (McE: 6)
Week 6 Regression inference via simulations (G&H: 7–10)

Multilevel Regression:

Week 7 Multilevel Linear Models (G&H: 11–13)
Week 8 Multilevel Models (G&H: 14, 15)
Week 9 Bayesian Inference (G&H: 18 / McE: 1, 2, 3)
Week 10 Fitting Models in Stan and brms (G&H: 16, 17 / McE: 11)

Overview

Sources of Variation

- Where is variation coming from?
- Between-person?
- Within-person?
- Within-person between measurement occasions?
  - e.g. Intensive measurement designs

Side Topic:
Mean induced associations in Cross-Sectional Data

Hofer & Sliwinski 2001: “mean induced associations”

- Correlations may reflect differences in means:
- The passage of time leads to “fake“ correlations

Problem Mean differences at the population level are confounded with changes at the individual level

e.g.: Speed Hypothesis, common cause hypotheses (sensory functioning, etc.)

We always knew it...

There has been no source more fruitful of fallacious statistical argument than the common influence of the time factor.
Cave and Pearson (1914, p. 354)

Biased correlations

Indices about rates of change based on cross-sectional age in age-heterogeneous samples are *always* biased. Why?
Individual values of two variables x and y of person i at a given age t corresponds to the sum of fixed and random effects:

$$\begin{aligned} x_{it} &= L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi} \\ y_{it} &= L_y + L_{yi} + S_y t_i + S_{yi} t_i + e_{yi} \end{aligned} \tag{1}$$

- L is the average level
- L_i is the individual departure from the fixed effect
- S is the average slope
- S_i is the individual departure from the fixed effect
- t_i age of person i at occasion t

Biased correlations

What is the relation among x_{it} and y_{it}?
Sample covariance:

$$cov_{xy} = \frac{1}{n-1} \sum_{1=1}^{N} (x_i - \overline{x})(y_i - \overline{y})$$

Population covariance:

$$COV(X, Y) = E[(X - \mu_x)(Y - \mu_y)] \tag{2}$$

Biased correlations

Inserting our linear model (1) into equation (2). Note that $E(X) = \mu_x$, and $E(Y) = \mu_y$. Hence:

$$\begin{aligned}
&COV(X,Y) \\
&= E[(L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi} - E(L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi})) \\
&\quad (L_y + L_{yi} + S_y t_i + S_{yi} t_i + e_{yi} - E(L_y + L_{yi} + S_y t_i + S_{yi} t_i + e_{yi}))] \\
&= E[(L_x + L_{xi} + S_x t_i + S_{xi} t_i + e_{xi} - (L_x + 0 + S_x E(t_i) + 0 + 0)) \\
&\quad (L_y + L_{yi} + S_y t_i + S_{yi} t_i + e_{yi} - (L_y + 0 + S_y E(t_i) + 0 + 0))] \;\mid E(t_i) = \bar{t} \\
&= E[(L_{xi} + S_x t_i + S_{xi} t_i + e_{xi} - S_x \bar{t}) \\
&\quad (L_{yi} + S_y t_i + S_{yi} t_i + e_{yi} - S_y \bar{t})] \;\mid \text{multiply all elements} \\
&= E[L_{xi} L_{yi} + L_{xi} S_y t_i + L_{xi} S_{yi} t_i \\
&\quad + S_x t_i L_{yi} + S_x t_i S_y t_i + S_x t_i S_{yi} t_i - S_x t_i S_y \bar{t} \\
&\quad + S_{xi} t_i L_{yi} + S_{xi} t_i S_y t_i + S_{xi} t_i S_{yi} t_i - S_{xi} t_i S_y \bar{t} \\
&\quad - S_x \bar{t} S_y t_i - S_x \bar{t} S_{yi} t_i + S_x \bar{t} S_y \bar{t}]
\end{aligned} \tag{3}$$

Biased correlations

The covariance of the fixed effects does not contain any random effects as their expected value is 0. Hence, (3) can be reduced to

$$\begin{aligned} &COV(X,Y) \\ &= E[S_x t_i S_y t_i - S_x t_i S_y \bar{t} - S_x \bar{t} S_y t_i + S_x \bar{t} S_y \bar{t}] \\ &= S_x S_y E(t_i^2 - 2\bar{t}t_i + \bar{t}^2) = S_x S_y E(t_i - \bar{t})^2 \\ &= S_x S_y Var(t_i) \end{aligned} \tag{4}$$

Conclusion

The covariance among x and y depends on the slope parameter and the variance in t_i. With increasing variance in t_i the correlation can take almost any value – even if there is no relation among both variables.

Simulation

Simulate Data with R which correspond to the models in (1).
Aim: Create dataset which contains three variables:

- Age
- x
- y

These values are to be generated given the linear models in (1)
Procedure:

1. Define own function which generates Data
   - Population model
2. Estimate correlation among x and y
3. manipulate t_i to understand its effect on the correlation

Simulation

Generate a function in R with `function()` and pass it to the object `score` which contains our function

```
score <- function(N,L, sdL, S, sdS, t.range){
  # N= number of subj, L= fixed Level; sdL= sd L;
  # t.range= e.g. agerange
  e <- rnorm(N, sd=abs(S)/2)
  ranL <- rnorm(N, sd=sdL)
  ranS <- rnorm(N, sd=sdS)
  vart <- runif(N, min=-(t.range/2), max=t.range/2)
  val <- matrix(ncol=2, nrow=N)
  for(i in 1:N){
  val[i,1] <- vart[i]
  val[i,2]<- L+ranL[i] + S*vart[i] + ranS[i]*vart[i] + e[i]}
  as.data.frame(val)
}
```

Simulation

Our object is `score`

Objects are assigned values and contents:

- Numbers, strings, datasets, functions, etc.

`R` uses the arrow `<-` to assign content to objects.
Here, our function is assigned to the object `score`:

```
score <- function(N,L, sdL, S, sdS, t.range)
```

The arguments `N, L, sdL, S, sdS` and `t.range` are called in our function.

Here: The arguments in the function are the population and simulation values for our model

After the call of `function()` its definition follows in curly brackets `{}`.

Simulation

- # is used to comment the code
- `rnorm(n, mean=0, sd=sdL)`; `n` random draws from a normal distribution $\sim N(0, sdL)$
- `runif(n, min=-(t.range/2), max=t.range/20`; `n` random draw from the uniform distribution
- `matrix()`; a Matrix with `N` rows and 2 columns is assigned to the object `val`
- `for(i in 1:N){...}`; A loop is defined
- In the first line, the i'th value of the vector `vart` is written into the first column and i'th row of `val`

Simulation

- `L+ranL[i] + S*vart[i] + ranS[i]*vart[i] + e[i]` is our model (cf. Equation 1)

  Second line of the loop
  - `L` and `S` are the fixed values
  - `ranL[i]`, `ranS[i]`, `vart[i]` and `e[i]` are random effects
- The index `i` increments after each loop
- `as.data.frame` transforms the matrix into a data set – not necessary but easier to handle later.

Simulation

Population values are taken from Lindenberger and Ghisletta (2009): *Digit Letter*.

- N = 500
- L= 48.29
- sdL= 0.43
- S= -0.81
- sdS= 0.05
- t.range= 10

Our function returns a dataset with 500 observations:

```
> score(500, 48.29, 0.43, -0.81, 0.05, 10)
            V1       V2
1   -4.44477040 53.25156
2   -3.18142497 51.20585
3    2.45730549 45.31237
.      .           .
    < omitted >
.      .           .
498  4.47916833 45.17909
499 -3.71434784 51.15925
500  2.75847019 45.88259
```

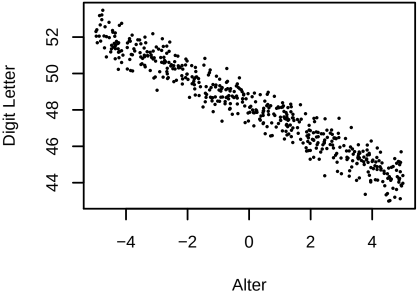
Simulation

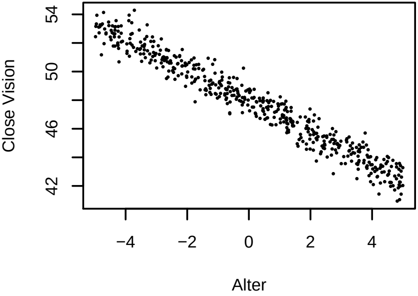
Two variables x and y are generated covering an age-range of 10 years (values from Lindenberger & Ghisletta, 2009):

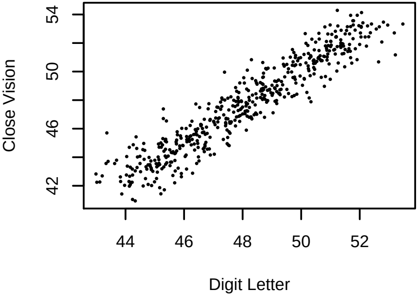
```
# Digit Letter
x <- score(1000, 48.29, 0.43, -0.81, 0.05, 10)
# Close Vision
y <- score(1000, 47.75, 0.37, -1.10, 0.09, 10)
```

These variables are sorted according to age and a new dataset `sim` is created.

```
# Data set with simulated data
sim <- data.frame(x[order(x[,1]),],y[order(y[,1]),2])
names(sim) <- c("Age", "x", "y")
> sim[1:3,]
          Age        x        y
851 -4.971680 51.70464 52.81497
187 -4.964126 52.19422 52.89674
520 -4.963927 51.52887 52.53314
```








```
plot(sim$Age, sim$x, xlab="Age",
     ylab="Digit Letter")
plot(sim$Age, sim$y, xlab="Age",
     ylab="Close Vision")
plot(sim$x,sim$y, xlab="Digit Letter",
     ylab="Close Vision")
```

Simulation

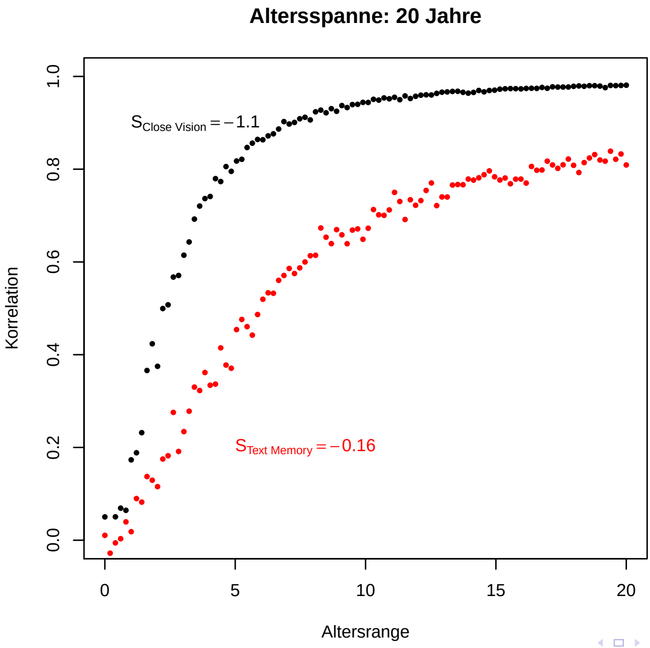
The correlation among x and y is:

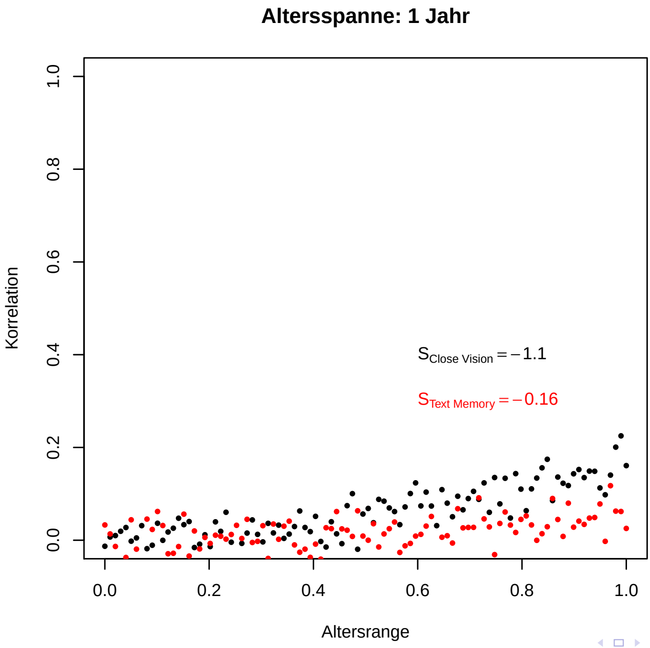
```
> cor.test(sim$x, sim$y)

Pearson's product-moment correlation

data:  sim$x and sim$y
t = 86.1402, df = 998, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9310502 0.9457991
sample estimates:
      cor
0.9388538
```

Simulation

Illustration of inflated correlation as function of $var(t_i)$.
Agerange: 0 to 20 years.

```
agerange <- 20
plot(0:agerange, seq(0,1, length.out=agerange+1), xlab="Agerange",
    ylab="Correlation", type="n")
for(a in seq(0,agerange, length.out=100)){
  x <- score(1000, 48.29, 0.43, -0.81, 0.05, a) # Digit Letter
  y <- score(1000, 47.75, 0.37, -1.10, 0.09, a) # close Vision
    # Dataframe with simulated data
  sim <- data.frame(x[order(x[,1]),],y[order(y[,1]),2])
  names(sim) <- c("Age", "x", "y")
  points(a, cor(sim$x, sim$y), cex=.8)
  }
```


Altersspanne: 20 Jahre
$S_{\text{Close Vision}} = -1.1$
$S_{\text{Text Memory}} = -0.16$
Korrelation
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
Altersrange



Altersspanne: 1 Jahr
Korrelation
0.0
0.2
0.4
0.6
0.8
1.0
Altersrange
$S_{\text{Close Vision}} = -1.1$
$S_{\text{Text Memory}} = -0.16$

Narrow Age Cohort Design

A Solution

Hofer and Sliwinski (2001) suggest the use of “narrow age cohorts” (NAC) designs:

- Narrow age bands
- Bias due to age-heterogeneity is reduced
- Remaining correlations are probably reflecting “true” interrelations
- NAC can be used post-hoc on existing data
- Poblem: Power issues with small age bands

Summary

- Different perspectives on same data
- Transformations can yield different interpretations
- Fundamental: Variance, covariance, correlation
- Correlation is index of consistency of individual differences
- Bias from multiple sources

- Why is this relevant?
- Mixing up sources of variation *is* relevant any application
- Also in longitudinal models

Confounding Sources of Variation

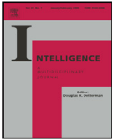
Recent Example

Intelligence 40 (2012) 260–268

Contents lists available at SciVerse ScienceDirect

Intelligence

journal homepage:



Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared

Paolo Ghisletta [a,b,*], Patrick Rabbitt [c,d], Mary Lunn [e], Ulman Lindenberger [f]

Confounding Sources of Variation

Recent Example

sents the MMLM, were Y_{aik} is the cognitive score at age a for individual i on the cognitive task k; I, lS, and qS are the Intercept and the linear and quadratic Slopes, respectively; $\beta_{1,2,3}$ are three retest effects; β_4 is the city effect; β_5 estimates sex' effects; $\beta_{6,7,8,9,10,11}$ are the socio-economic class' effects; and ε_{aik} is the error component.

$$Y_{aik} = I_{ik} + lS_{ik} \cdot A_{ai} + qS_k \cdot A_{ai}^2 + \beta_{(1,2,3)k} \cdot r_{(1,2,3)k} + \beta_{4k} \cdot city_i + \beta_{5k} \cdot sex_i + \beta_{(6,7,8,10,11)k} \cdot soc_{(1,2,3,4,5,6)ik} + \varepsilon_{aik} \tag{1}$$

Sources of Variation

- Multilevel perspective:
- *Level 1*: e.g. within-person
- *Level 2*: e.g. between-person

- Variables may operate at Level 1 or Level 2
- Some variables may actually contain a *mix* of both elements!
- Time-varying variables
  - Changes in stress may predict changes in affect
  - ▷ Daily stress might be positively related to daily reports of negative affect (NA)
  - How can this be conceived?
  - $\text{NA}_{ij} = \bar{\text{NA}}_i + \text{NA}^*_{ij}$
    Person average affect + plus (minus) a daily fluctuation
- Similar problem with age as time variable:
  - $\text{Age}_{ij} = \bar{\text{Age}}_i + \text{time}^*_{ij}$

Disaggregation of Within- and Between-Person Effects

- Lot's of attention in the last 10 years
- Recently, e.g. Wang & Maxwell, 2015
- Three options:
  1. No-centering
  2. Grand-mean centering
  3. Person or Group-mean centering
- Between-person relations are different from within-person relations conceptually and empirically.
- Not only could they have different magnitudes, but, in some cases, the two types of relations may even have different directions

▷ People who exercise more tend to have a lower risk of heart attack (negative between-person effect), whereas one is more likely to experience a heart attack while exercising (positive within-person effect).

Disaggregation of Within- and Between-Person Effects

Curran, Lee, Howard, Lane, and MacCallum (2012), "either confounding or mis-attributing" these two effects could lead to misleading results. Therefore, it is generally necessary to disaggregate between- and within-person effects. Fortunately, with longitudinal data, we can study both between- and within-person relations and try to disaggregate the two types of relations when both the outcome and the predictor are time varying.

- Disaggregating between- and within-person relations requires both design and analysis considerations
- consideration of two methodological issues: centering and detrending.
- The centering issue is relevant for the disaggregation, even when neither variable exhibits any trend over time,
- whereas the detrending issue is relevant only when at least one of the variables exhibits some trend over time.
- Centering:
    - a constant or a vector of constants has been subtracted from every value of the variable. In the context of disaggregation with longitudinal data, centering refers to possible redefinitions of the 0 point of the predictor variable
    - Three options are generally possible
    - No centering
    - Grand-mean centering
    - Person-mean centering

- Detrending:
  - Removing the trend from a time series
  - Detrending refers to controlling for the effect of time while examining the relation between the two variables.
  1. I detrending always needed when there is a trend over time?
  2. How should it be conducted?
  - Curran and Bauer (2011) suggest that we should *detrend the time-varying predictor*
  - Do we have to detrend the time-varying predictor or remove the time effect from the time-varying predictor to study within-person effects whenever the predictor may be related to time itself?
- Curran et al. (2012) and Hoffman and Stawski (2009) discussed how to use multilevel models to disaggregate between- and within-person effects
  - time is included as a covariate in the first-level model
  - for the purpose of detrending the *time-varying outcome*, not for detrending the time-varying predictor.
  - Curran and Bauer suggested detrending the predictor, whereas both Curran et al. and Hoffman and Stawski suggested detrending the outcome

Basic Multilevel Model for Disaggregating Between- and Within-Person Effects

- Data on two time-varying variables (Y and X)
- Y might represent self-concept and X might represent mood.
- Model M1:

$$
\begin{aligned}
y_{it} &= \gamma_{0i} + \gamma_{1i} x_{it} + e_{it} \\
\gamma_{0i} &= \gamma_{00} + u_{0i} \\
\gamma_{1i} &= \gamma_{10} + u_{1i}
\end{aligned}
$$

- y_{it} and x_{it} are the observed scores of Y and X of individual i at time t.
- Overall effect from Model M1 is a combination of both between- and within-person effects
- Between- and within-person effects are not disaggregated!

Basic Multilevel Model for Disaggregating Between- and Within-Person Effects

- When disaggregation of the two effects is considered
- Include respective predictors of between- and within-person effects
- Model M2:

$$
\begin{aligned}
y_{it} &= \gamma_{0i} + \gamma_{1i} xw_{it} + e_{it} \\
\gamma_{0i} &= \gamma_{00} + \gamma_{01} xb_i + u_{0i} \\
\gamma_{1i} &= \gamma_{10} + u_{1i}
\end{aligned}
$$

- x_{bi} and xw_{it} are between- and within-person characteristics of variable X.
- XB is a *time-invariant* predictor
- XW is a *time-varying* predictor
- γ_{01} is included in the model as a fixed effect only
- γ_{1i} is included as both a fixed effect, γ_{10} and as random effect, u_{1i}

Centering When There Are No Trends in Either X or Y

- Obtaining valid values for XB and XW in Equations 2 and 3 is the key to appropriately disaggregating between- and within- person effects.
- Scenario with no linear or nonlinear trends over time in the time-varying variables of interest, X and Y.
- In this case: At least two approaches we can use to center X:
  - grand-mean centering
  - person-mean centering

Centering When There Are No Trends in Either X or Y

- Grand-mean centering (M2)

$$
\begin{aligned}
y_{it} &= \gamma_{0i} + \gamma_{1i}(x_{it} - \bar{x}) + e_{it} \\
\gamma_{0i} &= \gamma_{00} + \gamma_{01}\bar{x_i} + u_{0i} \\
\gamma_{1i} &= \gamma_{10} + u_{1i}
\end{aligned}
$$

- Person-mean centering (M3)

$$
\begin{aligned}
y_{it} &= \gamma_{0i} + \gamma_{1i}(x_{it} - \bar{x}) + e_{it} \\
\gamma_{0i} &= \gamma_{00} + \gamma_{01}\bar{x_i} + u_{0i} \\
\gamma_{1i} &= \gamma_{10} + u_{1i}
\end{aligned}
$$

- The γ_{10} from the no centering model (M1) is conceptually different from above γ's which is neither the between-person effect nor the within-person effect but the composite effect without disaggregation.
- Composite effect is generally an uninterpretable blend of between- and within- person effects (Raudenbush and Bryk, 2002)

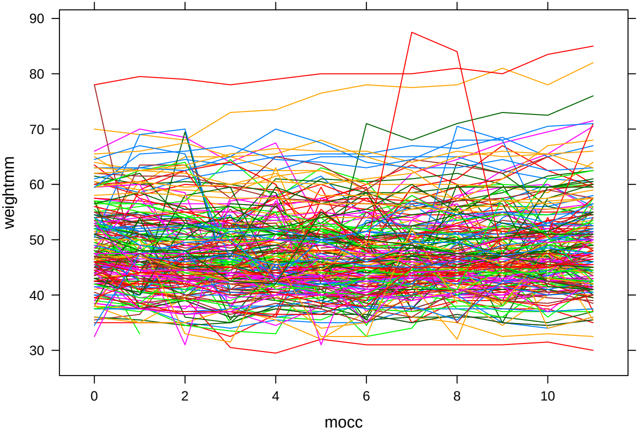
Centering When There Are No Trends in Either X or Y

- Grand-mean centering (M2) generally underestimates the variance of within- person effects
- Recommended: Person- mean centering over grand-mean centering (i.e., recommend M3 over M2)

Centering When There Are No Trends in Either X or Y

Example

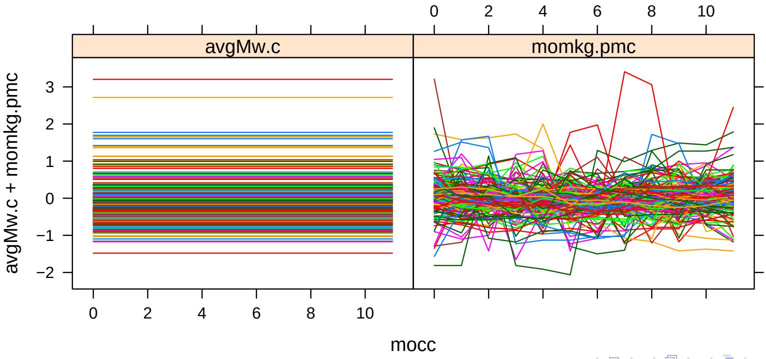
- Cebu: Mother's weight
- Fairly trend-free



Centering When There Are No Trends in Either X or Y

Example

```
## Person-mean centered
cebu$momkg.pmc <- (cebu$weightmm - cebu$avgMw)/10
## Center the average mom weight as well
cebu$avgMw.c <- scale(cebu$avgMw, scale = FALSE)/10
```



Centering When There Are No Trends in Either X or Y

```
Formula: weightbb ~ mocc + mocc2 + momkg.c +
                    mocc + mocc2 + momkg.c |id)

Random effects:
Groups   Name        Variance Std.Dev. Corr
id       (Intercept) 0.29616  0.5442
mocc        0.04386  0.2094   -0.04
mocc2       0.03506  0.1873   -0.13 -0.89
momkg.c     0.49270  0.7019   -0.07  0.30 -0.32
Residual             0.33359  0.5776
Number of obs: 2255, groups:  id, 188

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.37483    0.05482   98.05
mocc         0.70135    0.02080   33.71
mocc2       -0.27668    0.01839  -15.05
momkg.c      0.29078    0.06679    4.35
```

Centering When There Are No Trends in Either X or Y

```
Formula: weightbb ~ mocc + mocc2 + avgMw.c + momkg.pmc +
                   (mocc + mocc2 + momkg.pmc | id)

Random effects:
Groups   Name        Variance Std.Dev. Corr
id       (Intercept) 0.32426  0.5694
mocc        0.04604  0.2146   -0.10
mocc2       0.03618  0.1902    0.00 -0.89
momkg.pmc   0.73205  0.8556   -0.26  0.34 -0.36
Residual             0.32416  0.5694
Number of obs: 2255, groups:  id, 188

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.35310    0.05363   99.82
mocc         0.70314    0.02118   33.20
mocc2       -0.27713    0.01862  -14.88
avgMw.c      0.17480    0.06778    2.58
momkg.pmc    0.30313    0.08240    3.68
```

Centering When There Are No Trends in Either X or Y

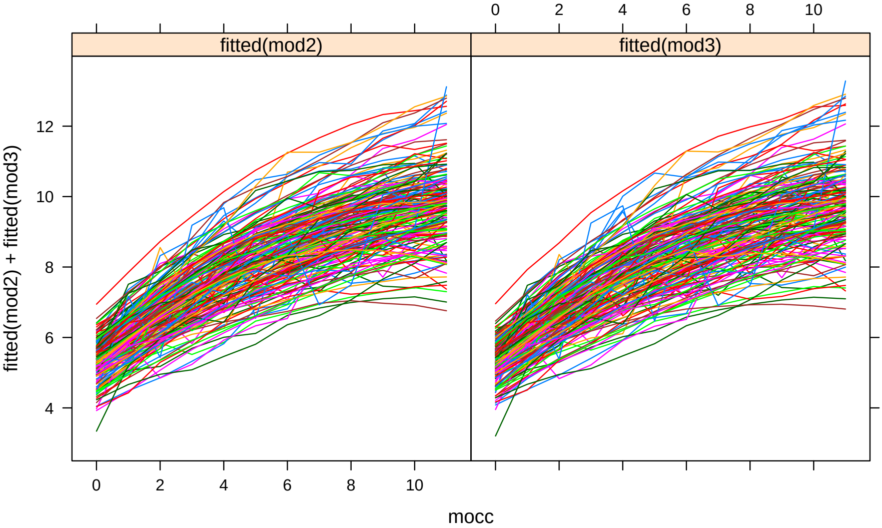
```
Grand-Mean:
Fixed effects:
Estimate Std. Error t value
(Intercept)  5.37483    0.05482   98.05
mocc         0.70135    0.02080   33.71
mocc2       -0.27668    0.01839  -15.05
momkg.c      0.29078    0.06679    4.35

Person-Mean
Fixed effects:
Estimate Std. Error t value
(Intercept)  5.35310    0.05363   99.82
mocc         0.70314    0.02118   33.20
mocc2       -0.27713    0.01862  -14.88
avgMw.c      0.17480    0.06778    2.58
momkg.pmc    0.30313    0.08240    3.68
```

Centering When There Are No Trends in Either X or Y

```
Grand-Mean:
Random effects:
Groups   Name        Variance Std.Dev. Corr
id       (Intercept) 0.29616  0.5442
mocc        0.04386  0.2094   -0.04
mocc2       0.03506  0.1873   -0.13 -0.89
momkg.c     0.49270  0.7019   -0.07  0.30 -0.32
Residual             0.33359  0.5776
Number of obs: 2255, groups:  id, 188

Person-Mean:
Random effects:
Groups   Name        Variance Std.Dev. Corr
id       (Intercept) 0.32426  0.5694
mocc        0.04604  0.2146   -0.10
mocc2       0.03618  0.1902    0.00 -0.89
momkg.pmc   0.73205  0.8556   -0.26  0.34 -0.36
Residual             0.32416  0.5694
Number of obs: 2255, groups:  id, 188
```

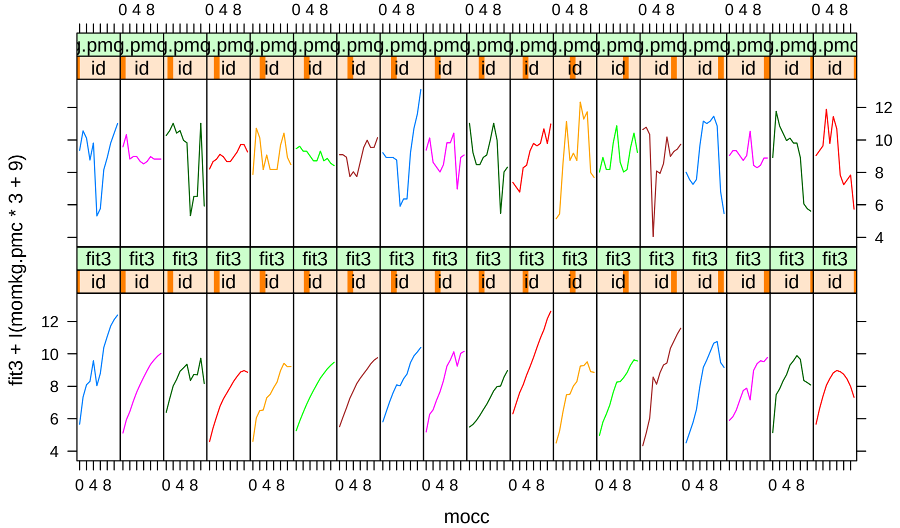
Centering When There Are No Trends in Either X or Y

Fitted: M2 and M3



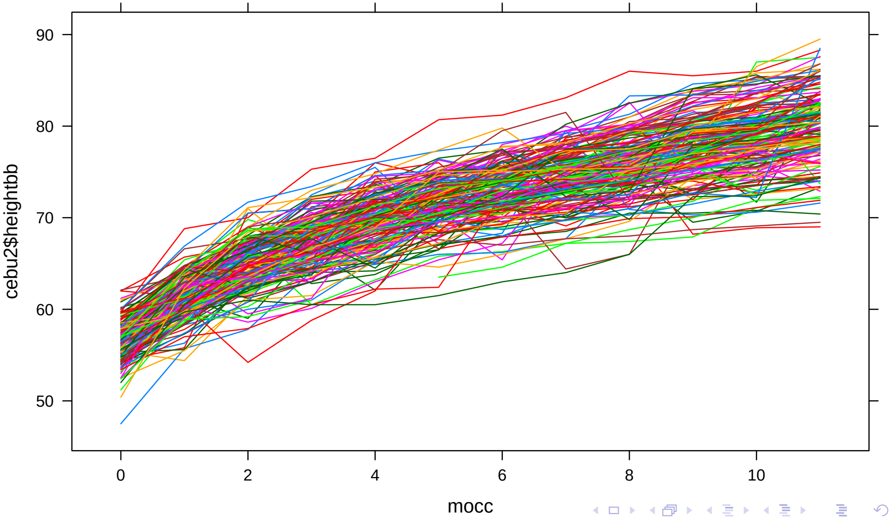
Centering When There Are No Trends in Either X or Y

Fitted: 18 Individuals



Detrending When There Are Trends in Either X or Y In

- Height of infant



Detrending When There Are Trends in Either X or Y In

- Two-step approach
- In the first step
  - OLS per individual
  - Mixed effects model
  - regressions of the time-varying predictor X on grand-mean centered time are fitted to each individual
  - In the second step, the estimated intercepts and residuals from the case-based regressions are used as observed data

$$\text{Step 1}\begin{cases} x_{it} = a_{0i} + a_{1i}(time_{it} - \bar{time}) + r_{xit} \\ r_{xit} = x_{it} - a_{0i} - a_{1i}(time_{it} - \bar{time}) \end{cases}$$

- In the second step
  - The estimated intercepts and residuals from the case-based regressions are used as observed data

$$\text{Step 2}\begin{cases} y_{it} = \gamma_{0i} + \gamma_{1i}(\hat{r}_{xit}) + e_{it} \\ \gamma_{0i} = \gamma_{00} + \gamma_{01}\hat{a}_{0i} + u_{0i} \\ \gamma_{1i} = \gamma_{10} + u_{1i} \end{cases}$$

Detrending When There Are Trends in Either X or Y In

- Via simulations, Curran and Bauer found that the between- and within-person effects were recovered with near-perfect accuracy in the balanced case
- Only modest bias in the unbalanced case
- When all the individuals have their average times equal to the grand mean of time, we have $\bar{time}_i = \bar{time}$ and, thus $\hat{a}_{0i} = \bar{x}_i$
- We can implement the two step approach in a single multilevel (Wang & Maxwell, 2015)

To Detrend or Not to Detrend

- Detrending is not always as necessary
- decision depends on whether the researchers want to control for the effect of time when looking into the within-person relation between two variables
- When the time effect is purposefully introduced by the study, one should preserve the time effects
- when the time effect is a result of factors that are not relevant to the study design or are not of research interest, one may want to control for the time effect via a detrending approach.
- "Basically, only when it is necessary to control for the time effect is detrending needed. In other cases, detrending may bring misleading results." (Wang & Maxwell, 2015)

Simulation Results

Wang & Maxwell, 2015

Table 8

Estimates of Variance Components and Their Standard Error Estimates Across Conditions From Different Centering Approaches

			No centering		Grand-mean centering		Person-mean centering	
True value	Nper	Ntime	Est.	*SE*	Est.	*SE*	Est.	*SE*
			Level-1 residual variance (σ_e^2)					
64	50	5	66.10	7.36	66.04	7.35	64.02	7.21
64	50	10	64.82	4.59	64.79	4.59	64.01	4.51
64	100	5	66.10	5.20	66.07	5.20	63.97	5.09
64	100	10	64.81	3.24	64.79	3.24	64.03	3.19
64	200	5	66.15	3.68	66.12	3.68	63.98	3.60
64	200	10	64.89	2.30	64.88	2.30	64.05	2.26
			Variance of within-person effects (σ_{11})					
4.00	50	5	2.79	1.54	2.81	1.55	4.03	2.02
4.00	50	10	3.27	1.14	3.28	1.15	4.00	1.34
4.00	100	5	2.68	1.06	2.69	1.06	4.03	1.42
4.00	100	10	3.27	0.80	3.27	0.80	4.01	0.94
4.00	200	5	2.65	0.74	2.65	0.74	4.02	0.99
4.00	200	10	3.24	0.56	3.24	0.56	3.99	0.66

Note. Nper = number of individuals; Ntime = number of time points; Est. = average estimate; *SE* = average standard error estimate.

- Strong effects on within-person variance
- Only person-mean centering is able to recover population parameters

Centering When Trends are Present

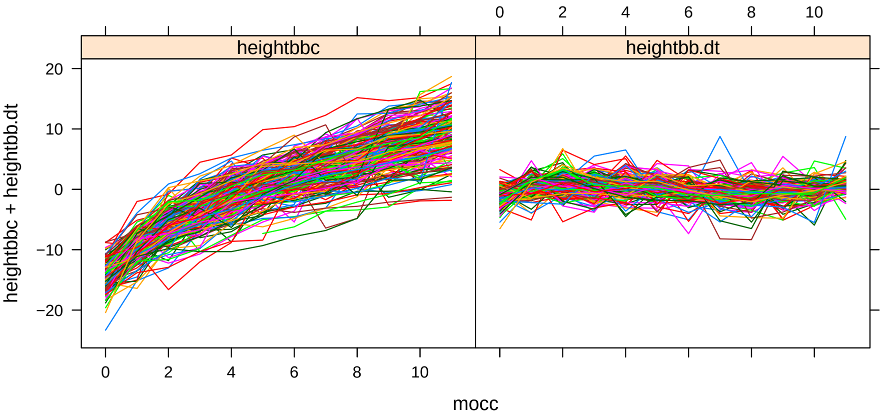
Example

```
> step1 <- lmer(heightbbc~I(mocc-5.5)+I((mocc-5.5)^2)+
                          (I(mocc-5.5)+I((mocc-5.5)^2)|id),
+                data=cebu2, na.action = na.exclude)
>
Linear mixed model fit by REML ['lmerMod']
Formula: heightbbc ~ I(mocc - 5.5) + I((mocc - 5.5)^2) +
                    (I(mocc - 5.5) + I((mocc - 5.5)^2) | id)
Random effects:
Groups   Name              Std.Dev. Corr
id       (Intercept)       2.43482
         I(mocc - 5.5)     0.34010   0.36
         I((mocc - 5.5)^2) 0.06374  -0.52  0.28
Residual                   1.80491
Number of obs: 2253, groups:  id, 188
Fixed Effects:
(Intercept)      I(mocc - 5.5)  I((mocc - 5.5)^2)
     1.4784             1.9511            -0.1221
```

Centering When Trends are Present

Example: Step 2, obtain residuals

```
cebu2$heightbb.dt <-  resid(step1)
xyplot(heightbbc + heightbb.dt~ mocc, groups = id, type = 'l', data
```



Centering When Trends are Present

Example: Step 2, obtain residuals

```
## Obtain intercept
hgt <- ranef(step1)$id[, '(Intercept)']
cebu2$avgHgt <- NA
for( i in 1:length(cebu2$id)){
   cebu2[cebu2$id == unique(cebu2$id)[i], 'avgHgt'] <- hgt[i]
}
```

Centering When Trends are Present

```
Formula: weightbb ~ mocc + mocc2 + heightbbc +
                   (mocc + mocc2 + heightbbc | id)
Random effects:
Groups    Name        Variance Std.Dev. Corr
id        (Intercept) 1.930323 1.3894
mocc         0.123174 0.3510   -0.89
mocc2        0.031443 0.1773    0.70 -0.92
heightbbc    0.009429 0.0971    0.95 -0.91  0.72
Residual              0.186571 0.4319
Number of obs: 2252, groups:  id, 188

Fixed effects:
             Estimate Std. Error t value
(Intercept)  8.237285   0.128135   64.29
mocc        -0.036736   0.033491   -1.10
mocc2        0.002683   0.017313    0.15
heightbbc    0.223428   0.009108   24.53
```

Centering When Trends are Present

```
Formula: weightbb ~ mocc + mocc2 + avgHgt + heightbb.dt +
                   (mocc + mocc2 + heightbb.dt | id)
Random effects:
Groups    Name        Variance Std.Dev. Corr
id        (Intercept) 0.30575  0.5529
mocc         0.04505  0.2123   -0.56
mocc2        0.03401  0.1844    0.34 -0.88
heightbb.dt 0.01415  0.1190    0.13 -0.01  0.00
Residual              0.18010  0.4244
Number of obs: 2252, groups:  id, 188

Fixed effects:
            Estimate Std. Error t value
(Intercept)  5.34657    0.04660  114.72
mocc         0.70120    0.01836   38.20
mocc2       -0.27294    0.01598  -17.08
avgHgt       0.20064    0.01398   14.35
heightbb.dt  0.21983    0.01127   19.51
```

Centering When Trends are Present

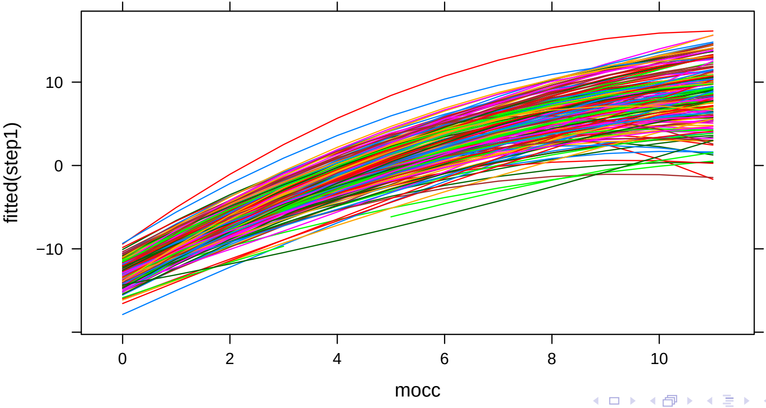
```
Not-Detrended:
Fixed effects:
Estimate Std. Error t value
(Intercept)  8.237285   0.128135   64.29
mocc        -0.036736   0.033491   -1.10
mocc2        0.002683   0.017313    0.15
heightbbc    0.223428   0.009108   24.53

Detrended:
Fixed effects:
Estimate Std. Error t value
(Intercept)  5.34657    0.04660  114.72
mocc         0.70120    0.01836   38.20
mocc2       -0.27294    0.01598  -17.08
avgHgt       0.20064    0.01398   14.35
heightbb.dt  0.21983    0.01127   19.51
```

Centering When Trends are Present

```
Not-Detrended:
Random effects:
Groups   Name        Variance Std.Dev. Corr
id       (Intercept) 1.930323 1.3894
mocc        0.123174 0.3510   -0.89
mocc2       0.031443 0.1773    0.70 -0.92
heightbbc   0.009429 0.0971    0.95 -0.91  0.72
Residual             0.186571 0.4319
Number of obs: 2252, groups:  id, 188

Detrended:
Random effects:
Groups   Name        Variance Std.Dev. Corr
id       (Intercept) 0.30575  0.5529
mocc        0.04505  0.2123   -0.56
mocc2       0.03401  0.1844    0.34 -0.88
heightbb.dt 0.01415  0.1190    0.13 -0.01  0.00
Residual             0.18010  0.4244
Number of obs: 2252, groups:  id, 188
```

Centering When Trends are Present

- Height and weight are obviously related
- Similar trend over time
- Different options of detrending
- Alternative, instead of using intercept, use predicted values of height.



Centering When Trends are Present

- Time (mocc) is dropped in favor of predicted weight

```
Formula: weightbb ~ heightpred + heightbb.dt +
                   (heightpred + heightbb.dt | id)
Random effects:
Groups   Name        Variance  Std.Dev. Corr
id       (Intercept) 0.2227203 0.4719
heightpred  0.0008525 0.0292   0.38
heightbb.dt 0.0137567 0.1173   0.15 0.15
Residual             0.2066667 0.4546
Number of obs: 2252, groups:  id, 188

Fixed effects:
            Estimate Std. Error t value
(Intercept) 8.038543   0.035823   224.4
heightpred  0.206313   0.002566    80.4
heightbb.dt 0.215669   0.011349    19.0
```

Summary

- Centering is needed when disaggregating between- and within- person effects
- Person-mean centering ap- proach on the time-varying predictor is recommended
- Detrending: Depends...
  - on study context and the substantive research question of interest
  - If one is interested in the relation after controlling for the time effect, detrending is needed.
  - If one believes the changes in the time-varying variables are attributable to study design or are of research interest, detrending may not be needed.
  - Detrending via multilevel models is recommended

A Bigger Picture

- Modeling also follows an evolution
- Interplay among what we can do in terms of modeling and what we have in terms of data
  - ▷ Interrelated
- Short overview of where we came and where we might be heading.

Development of Longitudinal Models in Psychology

Years	Popular Methods (Population ⟷ Individual)	Characteristic
1950	R-Technique — P-Technique	Cattell's Data Box
1960	ANOVA	Experimental Clusters
1970	Two-Stage	Few waves few datapoints
1980	Mixed-effects $\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ and variants	Increasing N few waves
1990	Marginal e.g. $h(\boldsymbol{X}_i\boldsymbol{\beta})$	Large N (e.g. HRS)
2000		Intensive Measurements
2010	Time-series e.g. $\alpha + \beta y_{t-1} + e_t$ dynamic models	Time Series continuous flow

Development of Longitudinal Models in Psychology

Years	Popular Methods (Population ⟷ Individual)	Characteristic
1950	R-Technique … P-Technique	Cattell's Data Box
1960	ANOVA	Experimental Clusters
1970	Two-Stage	Few waves few datapoints
1980	Mixed-effects $\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ and variants	Increasing N few waves
1990	Marginal e.g. $h(\boldsymbol{X}_i\boldsymbol{\beta})$	Large N (e.g. HRS)
2000	Time-series e.g. $\alpha + \beta y_{t-1} + e_t$ dynamic models	Intensive Measurements
2010		Time Series continuous flow

Development of Longitudinal Models in Psychology

Years	Popular Methods (Population ⟷ Individual)	Characteristic
1950	R-Technique ... P-Technique	Cattell's Data Box
1960	ANOVA	Experimental Clusters
1970	Two-Stage	Few waves few datapoints
1980	Mixed-effects $\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ and variants	Increasing N few waves
1990	Marginal e.g. $h(\boldsymbol{X}_i\boldsymbol{\beta})$	Large N (e.g. HRS)
2000		Intensive Measurements
2010	Time-series e.g. $\alpha + \beta y_{t-1} + e_t$ dynamic models	Time Series continuous flow

Development of Longitudinal Models in Psychology

Years	Popular Methods (Population ← → Individual)	Characteristic
1950	R-Technique … P-Technique	Cattell's Data Box
1960	ANOVA	Experimental Clusters
1970	Two-Stage	Few waves few datapoints
1980	Mixed-effects $\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ and variants	Increasing N few waves
1990	Marginal e.g. $h(\boldsymbol{X}_i\boldsymbol{\beta})$	Large N (e.g. HRS)
2000		Intensive Measurements
2010	Time-series e.g. $\alpha + \beta y_{t-1} + e_t$ dynamic models	Time Series continuous flow

Development of Longitudinal Models in Psychology

Popular Methods: Population ⟷ Individual

Years	Popular Methods	Characteristic
1950	R-Technique … P-Technique	Cattell's Data Box
1960	ANOVA	Experimental Clusters
1970	Two-Stage	Few waves few datapoints
1980	Mixed-effects $\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ and variants	Increasing N few waves
1990	Marginal e.g. $h(\boldsymbol{X}_i\boldsymbol{\beta})$	Large N (e.g. HRS)
2000	Time-series e.g. $\alpha + \beta y_{t-1} + e_t$ dynamic models	Intensive Measurements
2010		Time Series continuous flow

Development of Longitudinal Models in Psychology

Popular Methods: Population ⟷ Individual

Years	Popular Methods		Characteristic
1950	R-Technique	P-Technique	Cattell's Data Box
1960	ANOVA		Experimental Clusters
1970		Two-Stage	Few waves few datapoints
1980	Mixed-effects $\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ and variants		Increasing N few waves
1990	Marginal e.g. $h(\boldsymbol{X}_i\boldsymbol{\beta})$		Large N (e.g. HRS)
2000			Intensive Measurements
2010		Time-series e.g. $\alpha + \beta y_{t-1} + e_t$ dynamic models	Time Series continuous flow

Development of Longitudinal Models in Psychology

Years	Popular Methods (Population ⟷ Individual)	Characteristic
1950	R-Technique … P-Technique	Cattell's Data Box
1960	ANOVA	Experimental Clusters
1970	Two-Stage	Few waves few datapoints
1980	Mixed-effects $\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ and variants	Increasing N few waves
1990	Marginal e.g. $h(\boldsymbol{X}_i\boldsymbol{\beta})$	Large N (e.g. HRS)
2000		Intensive Measurements
2010	Time-series e.g. $\alpha + \beta y_{t-1} + e_t$ dynamic models	Time Series continuous flow

Some thought about...

- Nomothetic research: Search for general laws
- Aggregation: Typical approach to obtain nomothetic information
- Idiographic research: Focus on particularities of individual
- Ergodicity: Assumption that distribution of Variables in the population reflects distribution of variables in individual.

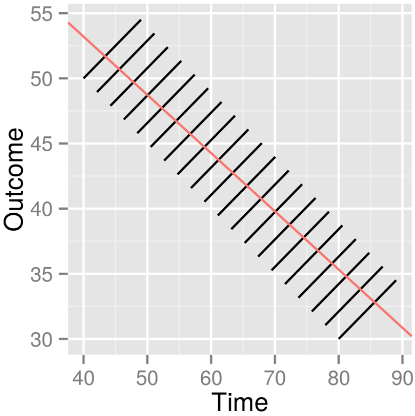
Aggregate $\neq$ Nomothetic

Aggregation

- Moments (Means, variances, proportions, correlations)
- Best guess for the aggregate
- ▷ Average across individuals

Nomothetic

- Pertains to extraction of general laws
- Individual parameters that are common among group/population
- ▷ All individuals increase over time



No Ergodicity? No Problem!

Idiographic does not mean that we can't obtain "general laws" or generalize.

- Modeling individuals may yield common parameters.
- Commonality is the nomothetic part
- ▷ Exploration of degree of commonality

Approximate nomothetic information with aggregation

- aggregation only yields nomothetic information if ergodicity holds

Nomothetic vs Idiographic

Assumption

- Results obtained at the population level reflect to some degree with-person processes
- Generalization from population to individual are meaningful

  Are these assumptions tenable?

Statement of Problem:

- Psychology focuses on individual variation between cases
- Results are commonly generalized to variation and explanation
  - in given populations
  - within individuals in these populations
- Ergodic theorem confines generalizability
- Ergodicity is hardly ever met in psychology

Nomothetic vs Idiographic

Consequence

- Results obtained from *inter*individual variation yield different results from studies based in *intra*individual variation
- we cannot blindly base statements about processes that take place within people on results that were obtained with standard large sample analyses.

Ergodicity

Assumption that distribution of Variables in the population reflects distribution of variables in individual.

- All population moments (e.g., means, variances, covariances) must be identical to the corresponding within-person moments
- Structural changes of time must be absent
- All developmental processes are by definition non-ergodic
- All within-person moments must be identical across individuals
- all subjects have to conform to the same statistical model

Approaching the Individual

Type of Information

Aggregate $\longleftrightarrow$ Idiographic

Regression models

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Mixed Effects

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$$

Mixed Effects Location and Scale

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i \text{ with } g(\sigma^2_{\nu_i}) = \boldsymbol{u}'_i\boldsymbol{\zeta} \qquad \text{and } g(\sigma^2_{\epsilon_{ij}}) = \boldsymbol{w}'_{ij}\boldsymbol{\tau}$$

Approaching the Individual

Type of Information

Aggregate ⟵ ⟶ Idiographic

Regression models

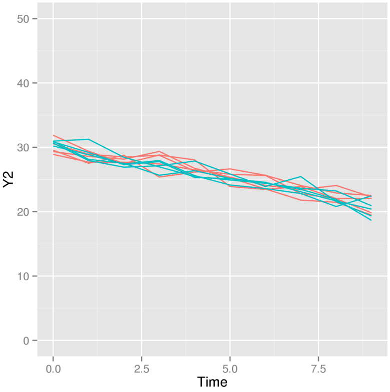
$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Mixed Effects

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$$

Mixed Effects Location and Scale

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i \text{ with } g(\sigma^2_{\nu_i}) = \boldsymbol{u}$$



Approaching the Individual

Type of Information

Aggregate ⟷ Idiographic

Regression models

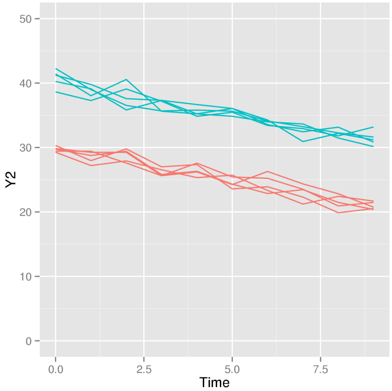
$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$$

Mixed Effects

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$$

Mixed Effects Location and Scale

$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ with $g(\sigma^2_{\nu_i}) = \boldsymbol{u}'_i$



Approaching the Individual

Type of Information

Aggregate ⟷ Idiographic

Regression models

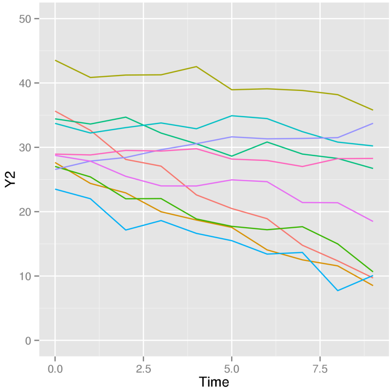
$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Mixed Effects

$$\boxed{\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i}$$

Mixed Effects Location and Scale

$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ with $g(\sigma^2_{\nu_i}) = \boldsymbol{u}_i'$



Approaching the Individual

Type of Information

Aggregate ⟵⟶ Idiographic

Regression models

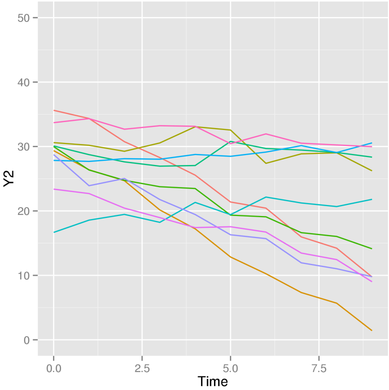
$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Mixed Effects

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$$

Mixed Effects Location and Scale

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i \text{ with } g(\sigma^2_{\nu_i}) = \boldsymbol{u}_i'\boldsymbol{\zeta}$$



Approaching the Individual

Type of Information

Aggregate ⟷ Idiographic

Regression models

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Mixed Effects

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$$

Mixed Effects Location and Scale

$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ with $g(\sigma^2_{\nu_i}) = \boldsymbol{u}_i'\boldsymbol{\zeta}$ and $g(\sigma^2_{\epsilon_{ij}}) = \boldsymbol{w}_{ij}'\boldsymbol{\tau}$

Approaching the Individual

Type of Information

Aggregate ⟷ Idiographic

Regression models

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Mixed Effects

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$$

Mixed Effects Location and Scale

$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i$ with $g(\sigma^2_{\nu_i}) = \boldsymbol{u}_i'\boldsymbol{\zeta}$ and $g(\sigma^2_{\epsilon_{ij}}) = \boldsymbol{w}_{ij}'\boldsymbol{\tau}$

...in a Nutshell

- Data are more intense and offer more information at individual level
- Methods need to be developed and explored that take advantage of individual information
- Move from aggregation to idiographic models (Castro-Schilo & Ferrer, 2013; Hamaker, 2012; Molenaar, 2004)

Features of Intensive Designs

- Observe change that occurs at different time scales
- ▷ Integration of different developmental trajectories within individuals

Introduction: Intensive Measurement Design

Classic multiwave design:

- Multiwave data do not have information on smaller scale
- ▷ Learning within waves: Taste of intensive designs
- Difficult to obtain micro level information
- No information on day-to-day variability
- Aimed at "slow" developmental processes

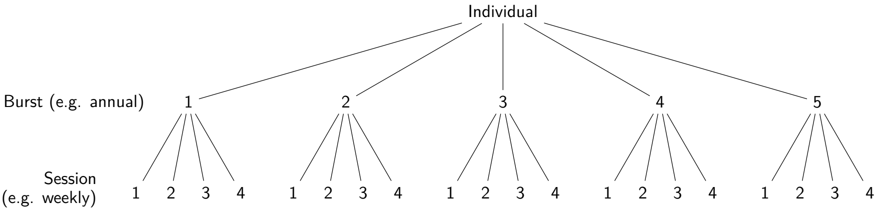
Ecological Momentary Assessments (EMA) & daily diary

- Measurement at small scale (hours, days)
- All information on micro level
- Aimed at short-term changes

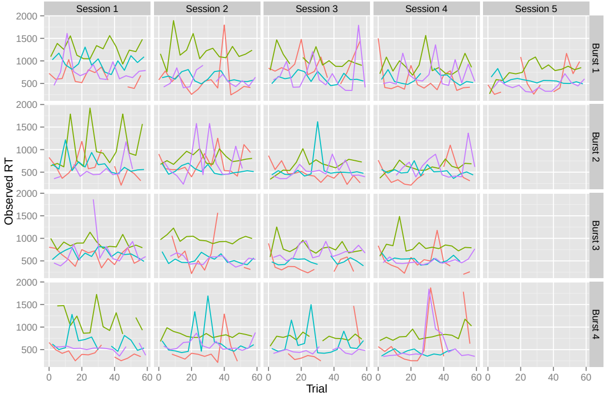
Introduction: Intensive Measurement Design

Intensive measurement designs (Nesselroade, 1991)

- Combination of multiwave design and EMA methods
- Multiple scales
- Suited for measuring change/variation in short and long term



Introduction: Intensive Measurement Design



- Variability is not noise but carries information*
  - ▷ systematic dynamic patterns of covariation

e.g. Cognition, affect, language use, perceived control† etc.

---

*Brose & Röcke, 2013; Fiske & Rice, 1955; Ram & Gerstorf 2009; Woodrow, 1932 etc.

†Eid & Diener, 1999; Eizenman, et al. 1997; van Geert et al. 2002; Rast et al. 2012; Siegler 1994 etc.

Typical Approach to Obtaining Index of IIV

Common approach to obtain within-person variability

- Extract IIV and compute an index (e.g. iSD)
- ▷ Step 1: De-trend data, print out residuals, compute index
- ▷ Step 2: Index$_i$ is used as predictor or dependent variable

Typically:

- Information about individual means is not retained
  - Especially problematic with heteroscedastic error terms
- Dependency among iM and iSD is not modeled
  - Interrelation between random effect terms of the means (location) part and the within-person variance (scale) is not independent.

Mixed Effects Location Scale Model

Alternative approach:
Mixed Effects Location Scale Model

(Hedeker et al., 2008; Rast et al. 2011, 2012, 2014, submitted)

- Model mean structure (location) and variability (scale) of the response
- Permits the use of explanatory variables for both
  - between-person variance as well as for
  - individual differences in within-person variance
- All correlations across both levels
  - interdependencies are all maintained
- All parameters are estimated **simultaneously**

Mixed Effects Location Scale Model

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\epsilon}_i,$$

- ▷ Standard: $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \sigma_\epsilon^2 \boldsymbol{\Psi}_i)$

LSM: IIV may fluctuate between individuals (i) and across time (j)

$$\sigma^2_{\epsilon_{ij}} = g(\mathbf{W}'_{ij}\boldsymbol{\tau} + \mathbf{V}'_{ij}\mathbf{t}_i).$$

- $\boldsymbol{\tau}$ defines the average WP variance: τ_0 (intercept), τ_1 (slope)
- Time-varying covariates $\mathbf{W}_{ij}$ for the fixed and $\mathbf{V}_{ij}$ for the random effect to influence the within-person variance estimate

- ▷ Different error distribution for each individual at each occasion

Mixed Effects Location Scale Model

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i,$$

▷ Standard: $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \sigma_\epsilon^2\boldsymbol{\Psi}_i)$

LSM: IIV may fluctuate between individuals (i) and across time (j)

$$\sigma^2_{\epsilon_{ij}} = g(\mathbf{W}'_{ij}\boldsymbol{\tau} + \mathbf{V}'_{ij}\mathbf{t}_i).$$

- $\boldsymbol{\tau}$ defines the average WP variance: τ_0 (intercept), τ_1 (slope)
- Time-varying covariates $\mathbf{W}_{ij}$ for the fixed and $\mathbf{V}_{ij}$ for the random effect to influence the within-person variance estimate

▷ Different error distribution for each individual at each occasion

"It is such variation from sitting to sitting, or from day to day, here designated by the term, 'quotidian variation,' that is to be considered. [...] .The responses on different days clearly are not all of the same category; they belong to different statistical populations." (Woodrow, 1932, pp. 246)

Application: Daily Reports of Stress and Affect

$$\text{Level 1:} \quad y_{ij} \sim N(\mu_{ij}, \sigma^2_{\epsilon_{ij}}) \quad \text{for } j = 1, ..., 7$$

$$\mu_{ij} = \beta_{0i} + \beta_{1i} \times \text{Session}_{ij}$$

$$\sigma^2_{\epsilon_{ij}} = \exp(\tau_{0i} + \tau_{1i} \times \text{Stress}_{ij})$$

$$\text{Level 2:} \quad \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\tau} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{\boldsymbol{\beta}} \\ \mu_{\boldsymbol{\tau}} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\boldsymbol{\beta}} & \sigma_{\boldsymbol{\beta\tau}} \\ \sigma_{\boldsymbol{\tau\beta}} & \sigma^2_{\boldsymbol{\tau}} \end{bmatrix} \right)$$

$$\text{Hyperpriors:} \quad \begin{bmatrix} \mu_{\boldsymbol{\beta}} \\ \mu_{\boldsymbol{\tau}} \end{bmatrix} \sim N(\mathbf{a}, \mathbf{B})$$

$$\begin{bmatrix} \sigma^2_{\boldsymbol{\beta}} & \sigma_{\boldsymbol{\beta\tau}} \\ \sigma_{\boldsymbol{\tau\beta}} & \sigma^2_{\boldsymbol{\tau}} \end{bmatrix}^{-1} \sim Wish(\mathbf{R}, k).$$

Rast, Hofer & Sparks (2012)

Application: Daily Reports of Stress and Affect

$$\text{Level 1:} \qquad y_{ij} \sim N(\mu_{ij}, \sigma^2_{\epsilon_{ij}}) \qquad \text{for } j = 1, ..., 7$$

$$\mu_{ij} = \beta_{0i} + \beta_{1i} \times \text{Session}_{ij}$$

$$\sigma^2_{\epsilon_{ij}} = \exp(\tau_{0i} + \tau_{1i} \times \text{Stress}_{ij})$$

$$\text{Level 2:} \qquad \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\tau} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{\boldsymbol{\beta}} \\ \mu_{\boldsymbol{\tau}} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\boldsymbol{\beta}} & \sigma_{\boldsymbol{\beta\tau}} \\ \sigma_{\boldsymbol{\tau\beta}} & \sigma^2_{\boldsymbol{\tau}} \end{bmatrix} \right)$$

$$\text{Hyperpriors:} \qquad \begin{bmatrix} \mu_{\boldsymbol{\beta}} \\ \mu_{\boldsymbol{\tau}} \end{bmatrix} \sim N(\mathbf{a}, \mathbf{B})$$

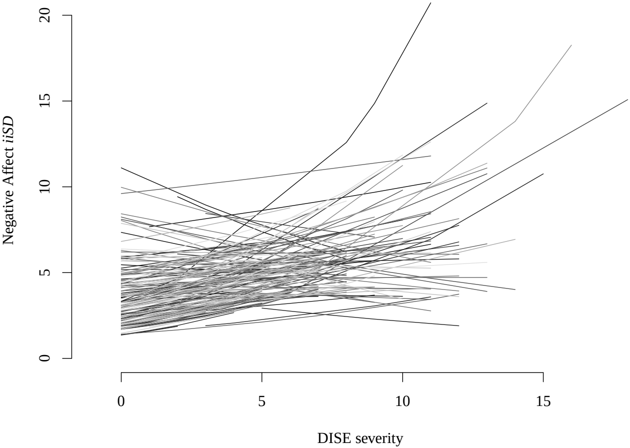
$$\begin{bmatrix} \sigma^2_{\boldsymbol{\beta}} & \sigma_{\boldsymbol{\beta\tau}} \\ \sigma_{\boldsymbol{\tau\beta}} & \sigma^2_{\boldsymbol{\tau}} \end{bmatrix}^{-1} \sim Wish(\mathbf{R}, k).$$

Rast, Hofer & Sparks (2012)

Results from Rast, Hofer and Sparks (2012)

Reactivity of negative affect: iSD by stress (DISEsev)

	Estimate	95% C.I.
β_0	22.94**	[21.9, 24.0]
β_1	-0.41**	[-0.56, -0.24]
$\sigma^2_{\beta_0}$	40.50**	[30.8, 52.5]
$\sigma^2_{\beta_1}$	0.13*	[0.04, 0.34]
r_{β_1,β_0}	$-.29^{ns}$	[-.69, .17]
r_{τ_0,β_0}	$.80$**	[.66, .90]
r_{τ_1,β_0}	$-.69$**	[-.84, -.48]
r_{τ_0,β_1}	$-.09^{ns}$	[-.58, .42]
r_{τ_1,β_1}	$.14^{ns}$	[-.41, .66]
r_{τ_1,τ_0}	$-.75$**	[-.87, -.55]
τ_0	2.55**	[2.28, 2.83]
τ_1	0.12**	[0.06, 0.17]
$\sigma^2_{\tau_0}$	1.27**	[0.80, 1.86]
$\sigma^2_{\tau_1}$	0.03**	[0.02, 0.06]



Application: LSM for Intensive Measurement Designs

Level 1:

$$
\begin{aligned}
y_{ijk} &\sim N(\mu_{ijk}, \sigma^2_{\epsilon_{ijk}}) \\
\mu_{ijk} &= \alpha_{0jk} + \alpha_{1jk}\text{Session}_{ijk} + \alpha_{2jk}\text{Burst}_{jk} + \beta\text{Session}_{ijk}\text{Burst}_{jk} \\
\sigma^2_{\epsilon_{ijk}} &= \exp(\tau_{0k} + \tau_{1k}\text{Session}_{ijk} + \tau_{2k}\text{Burst}_{ijk} + \lambda\text{Session}_{ijk}\text{Burst}_{jk})
\end{aligned}
$$

Random effects between individuals *and* between bursts within individuals.

Hyperpriors:

$$
\begin{bmatrix} \sigma^2_{\alpha_0} & \sigma_{\alpha_0\alpha_1} \\ \sigma_{\alpha_1\alpha_0} & \sigma^2_{\alpha_1} \end{bmatrix}^{-1} \sim \mathit{scaled\text{-}Wishart}\left(\mathbf{R1}, k1\right)
$$

$$
\begin{bmatrix} \sigma^2_{\boldsymbol{\alpha}} & \sigma_{\boldsymbol{\alpha\tau}} \\ \sigma_{\boldsymbol{\tau\alpha}} & \sigma^2_{\boldsymbol{\tau}} \end{bmatrix}^{-1} \sim \mathit{scaled\text{-}Wishart}\left(\mathbf{R2}, k2\right)
$$

Rast & MacDonald (2014)