Measuring Individual Change: A Gentle Introduction to the Pros and Cons of Modern Models

Lesa Hoffman

Associate Professor of Quantitative Methods,
Child Language Program

Associate Scientist,
Research Design and Analysis Unit

Schiefelbusch Institute for Life Span Studies
University of Kansas

Overview

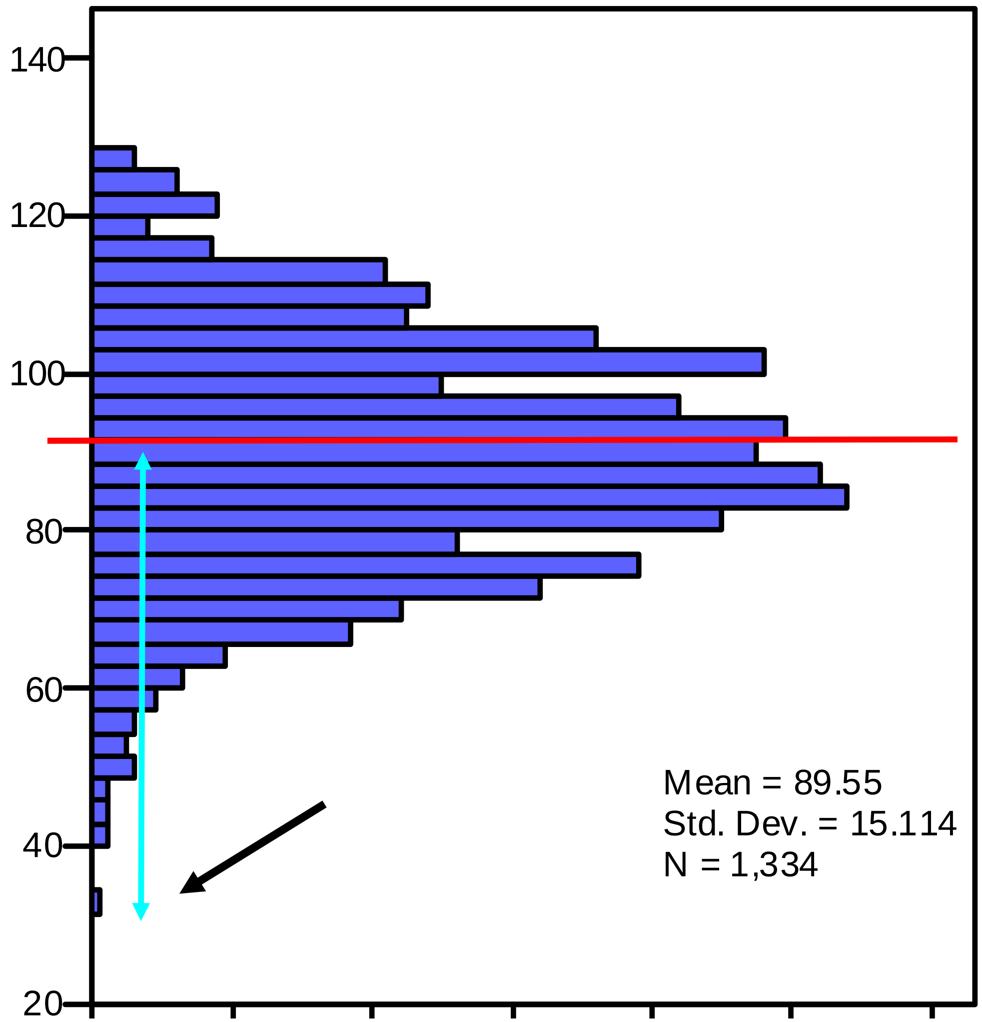
- Organizing principles
  - From one to many kinds of variance
- From cross-sectional to longitudinal (growth) models
  - Multilevel/hierarchical linear/latent growth curve models
- Models for exploring "heterogeneous populations"
  - Latent class/profile/transition/mixture models
- Models for confirming/testing heterogeneous groups
  - Heterogeneous variance longitudinal models
  - Confirmatory longitudinal mixture models

The Two Sides of Any Model

- **Model for the Means:**
  - *Aka* **Fixed Effects**, Structural Part of Model
  - What you are used to **caring about for testing hypotheses**
  - How the expected outcome for a given observation varies as a function of values for **known** predictor variables

- **Model for the Variance:**
  - *Aka* **Random Effects and Residuals**, Stochastic Part of Model
  - What you are used to **making assumptions about** instead
  - How residuals are distributed and related across observations (groups, persons, time, etc.) → these relationships are called "dependency" and ***this is the primary way that longitudinal models differ from general linear models*** *(e.g., regression)*

An Empty Between-Person Model (i.e., Single-Level)



$$y_i = \beta_0 + e_i$$

Filling in values:

$$32 = \underbrace{90}_{y_{pred}} + -58$$

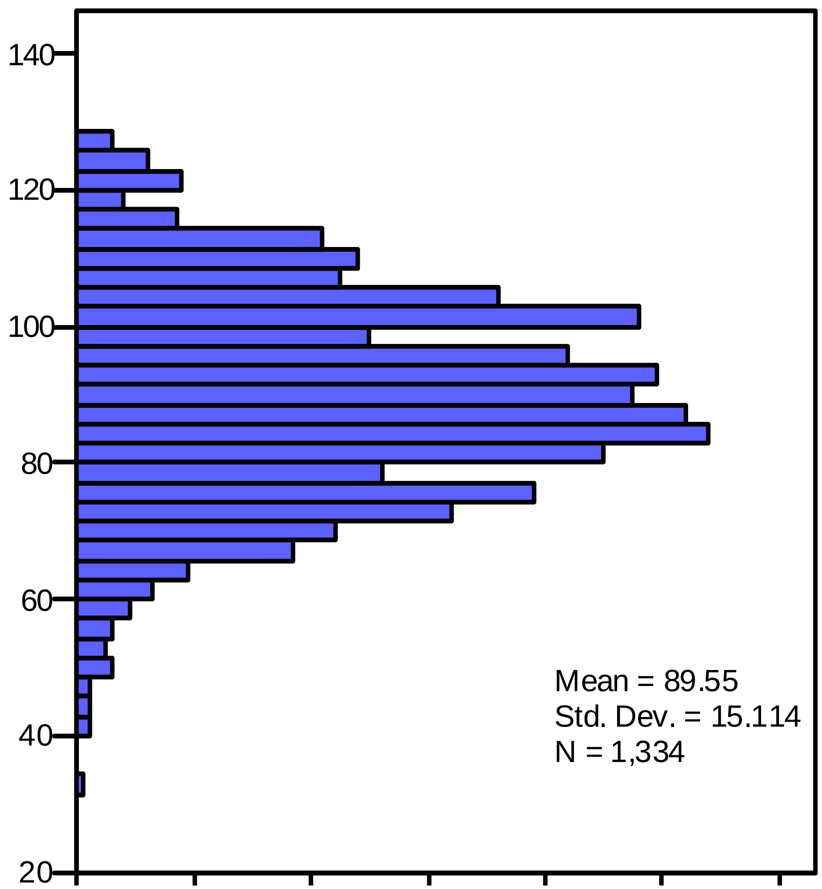
Model for the Means

y_i error variance:

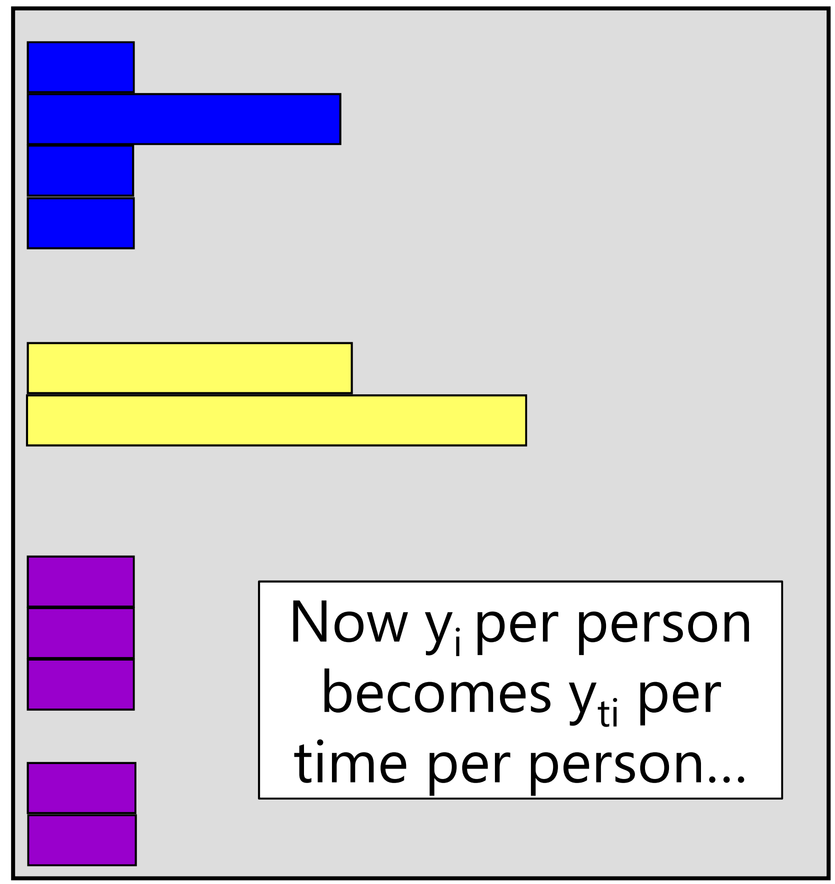
$$\frac{\sum\left(y_i - y_{pred}\right)^2}{N - 1}$$

Adding Within-Person Information… (i.e., to become a Multilevel Model)

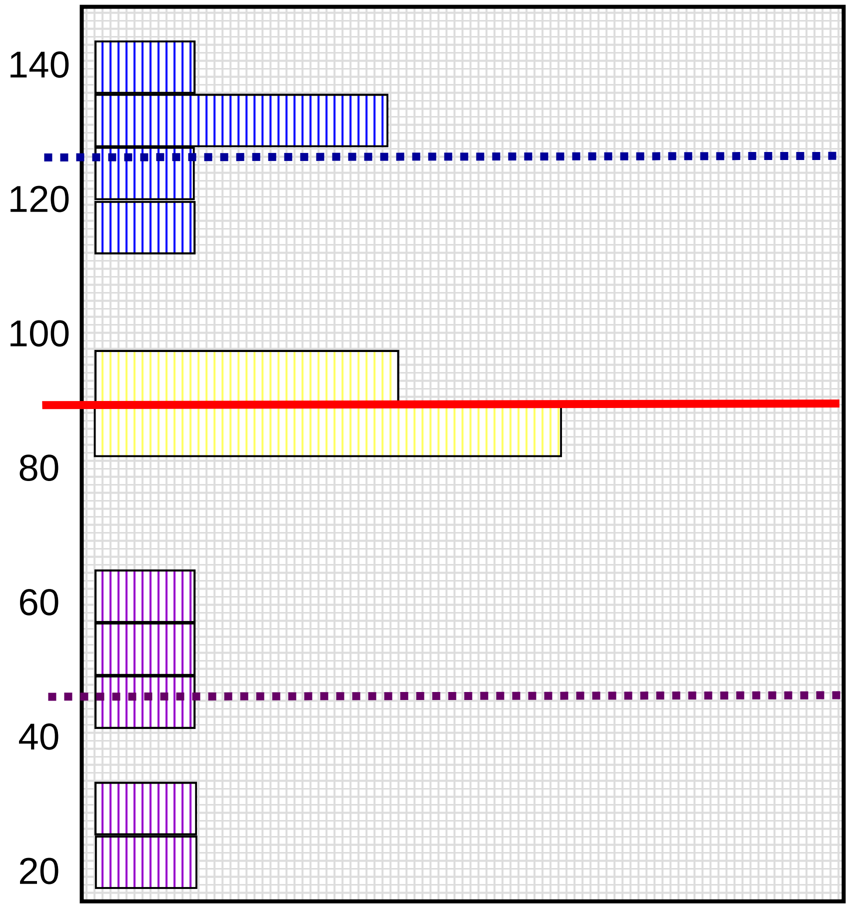
Full Sample Distribution



3 People, 5 Occasions each



Empty +Within-Person Model for y_{ti}



Start off with mean of y_{ti} as "best guess" for any value:

= Grand Mean

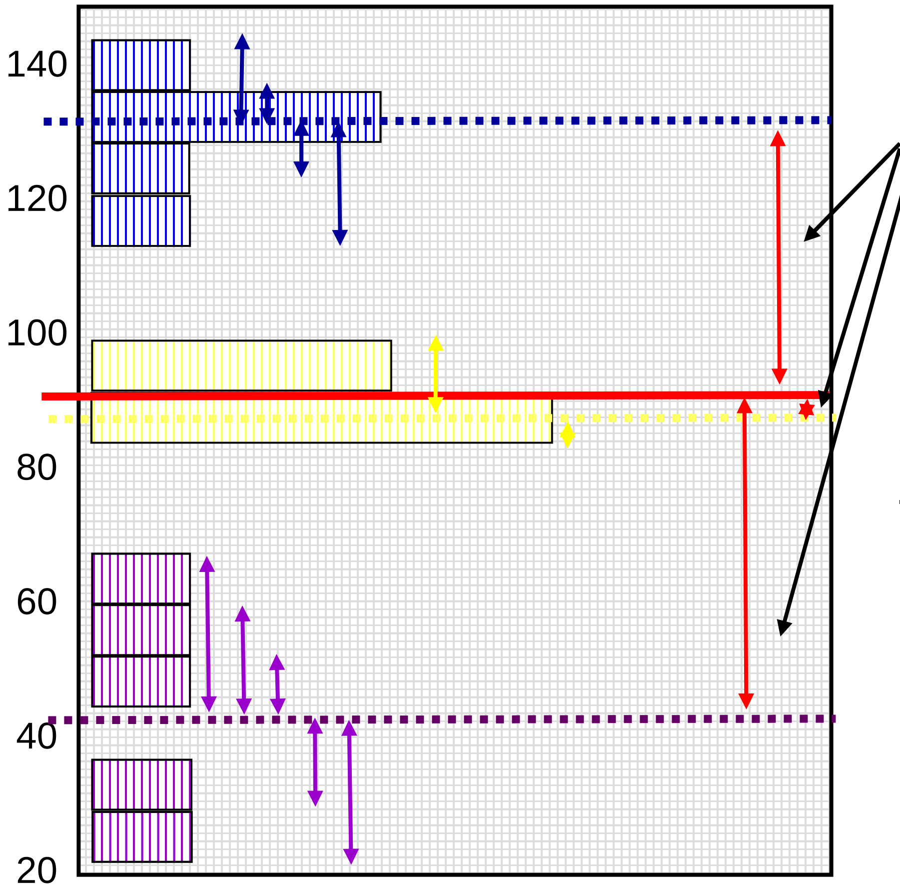
= Fixed Intercept

Can make better guess by taking advantage of repeated observations:

= Person Mean

→ Random Intercept

Empty +Within-Person Model



y_{ti} variance → 2 sources:

Between-Person (BP) Variance:

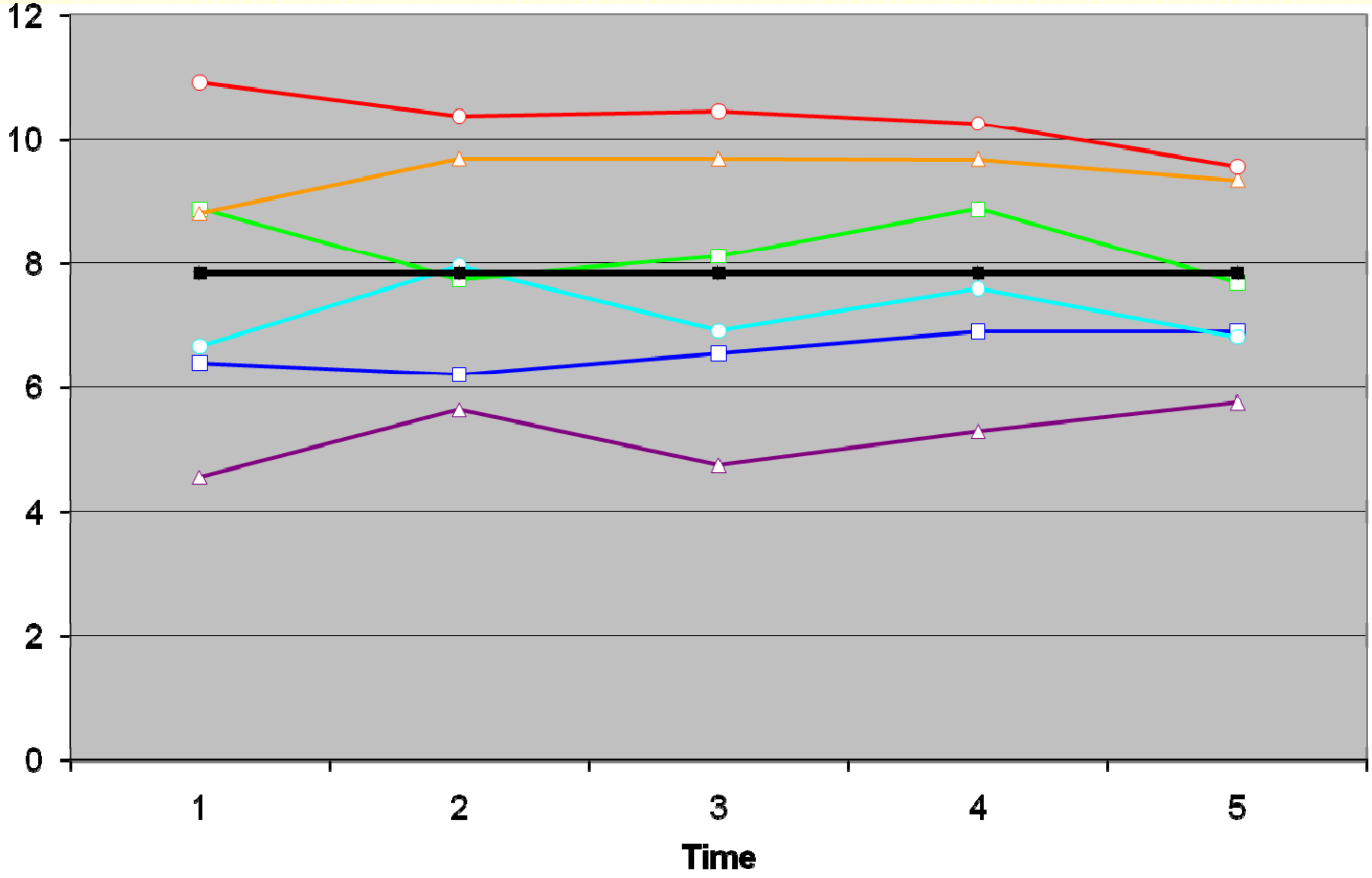
- → Differences from **GRAND** mean
- → **INTER**-Individual Differences

Within-Person (WP) Variance:

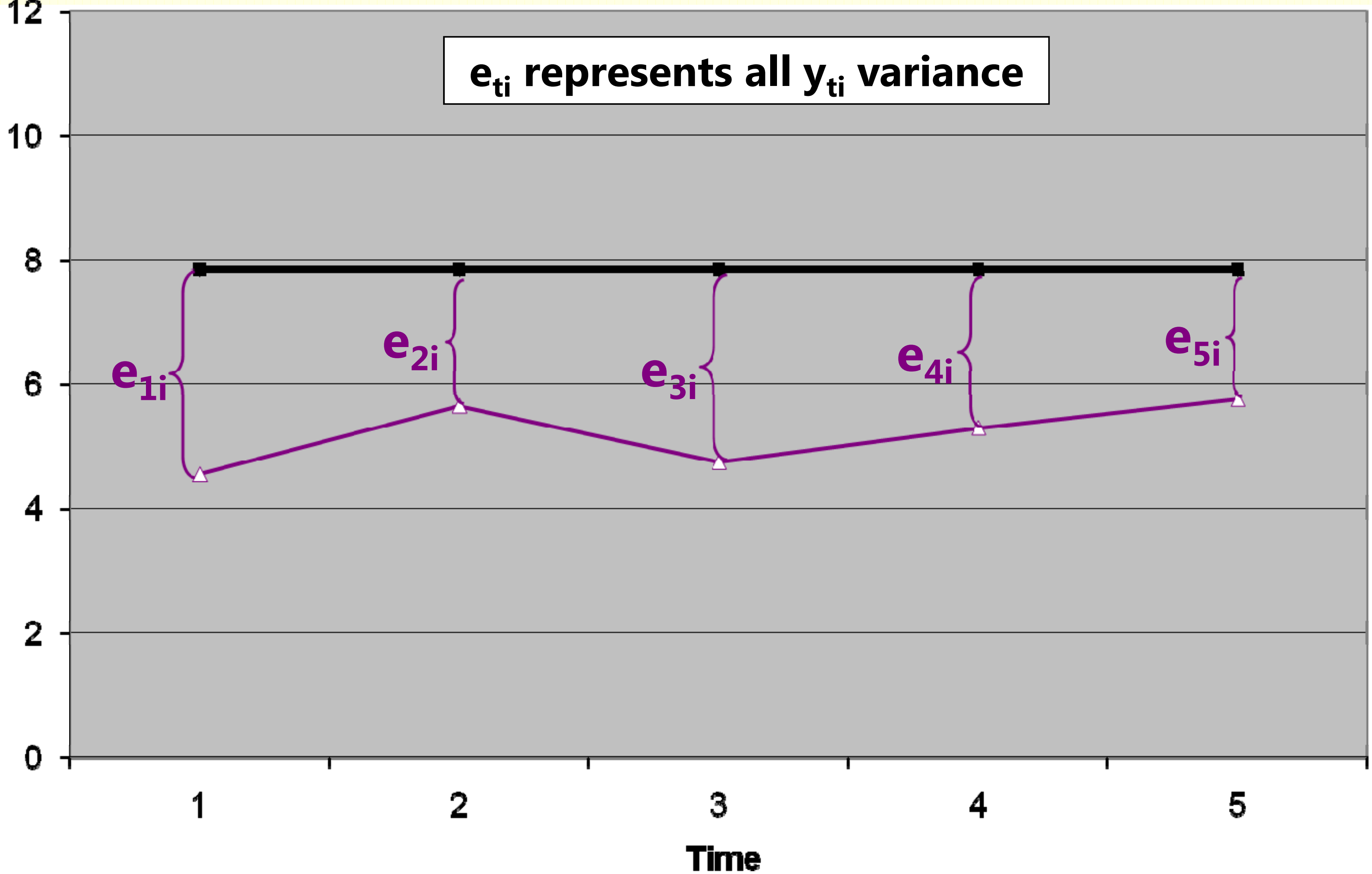
- → Differences from **OWN** mean
- → **INTRA**-Individual Differences
- → This part is only observable through longitudinal data.

Now we have 2 piles of variance in y_{ti} to predict.

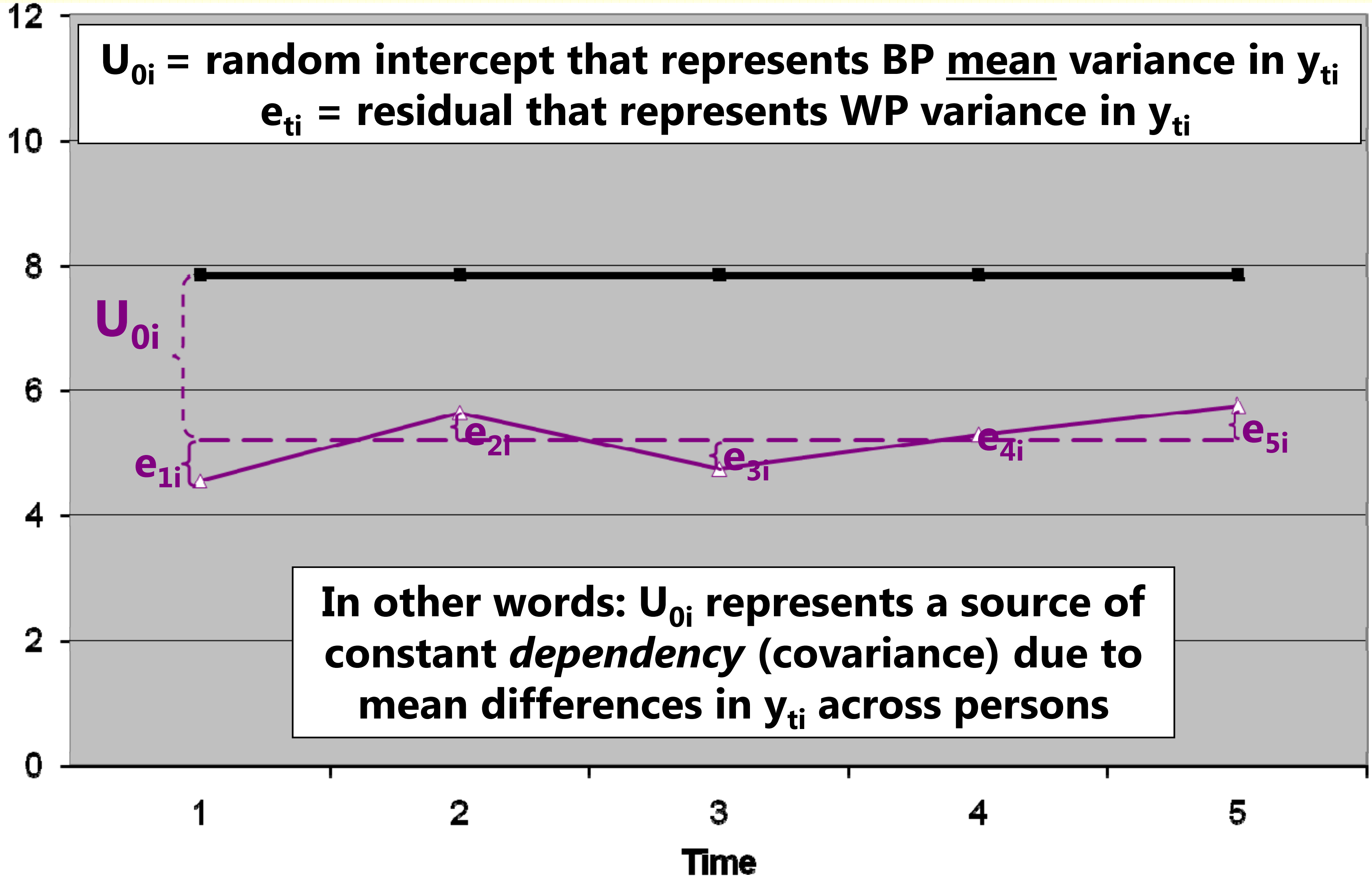
Hypothetical Longitudinal Data (black line = sample mean)



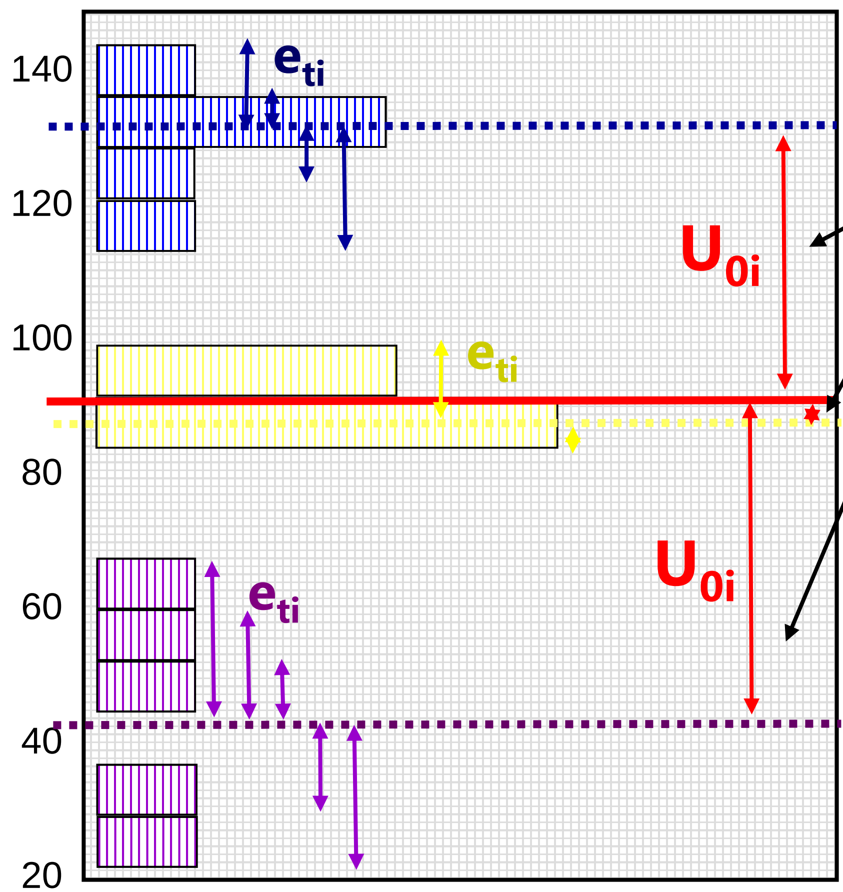
"Error" in a BP Model for the Variance: Single-Level Model



“Error” in a +WP Model for the Variance: Multilevel Model



Empty +Within-Person Model



y_{ti} variance → 2 sources:

Level 2 Random Intercept Variance (of U_{0i}, as $\tau^2_{U_0}$):

→ **Between**-Person Variance

→ Differences from **GRAND** mean

→ **INTER**-Individual Differences

Level 1 Residual Variance (of e_{ti}, as σ^2_e):

→ **Within**-Person Variance

→ Differences from **OWN** mean

→ **INTRA**-Individual Differences

BP and +WP Conditional Models

- Multiple Regression, **Between-Person** ANOVA: **1 PILE**
  - $y_i = (\beta_0 + \beta_1 X_i + \beta_2 Z_i \ldots) + e_i$
  - e_i → ONE residual, assumed uncorrelated with equal variance across observations (here, just persons) → "**BP (all) variation**"

- Repeated Measures, **Within-Person** ANOVA: **2 PILES**
  - $y_{ti} = (\beta_0 + \beta_1 X_i + \beta_2 Z_i \ldots) + U_{0i} + e_{ti}$
  - U_{0i} → A random intercept for differences in person means, assumed uncorrelated with equal variance across persons → "**BP (mean) variation**" = $\tau^2_{U_0}$ is now "leftover" after predictors
  - e_{ti} → A residual that represents remaining time-to-time variation, usually assumed uncorrelated with equal variance across observations (now, persons and time) → "**WP variation**" = σ^2_e is also now "leftover" after predictors

Repeated Measures (RM) ANOVA

- Models with a random intercept to predict a constant correlation of outcomes from the same person are also called:
  - Hierarchical linear models, multilevel models, general linear mixed models, growth curve models, random coefficient models…
- The key to how these latter models can extend beyond traditional RM ANOVA lies in maximum likelihood estimation (ML in MLM) instead of least squares (LS in ANOVA)
  - Options for other types of outcomes: Normal vs. not-normal
  - Options for uncooperative participants: Missing or unbalanced data
  - Options for extension: What if a random intercept is not enough to describe all sources of between-person differences?

Addressing Uncooperative Participants

- ML allows **incomplete** and **unbalanced** responses…

RM ANOVA via LS: uses **multivariate** (wide) data structure:

ID	Sex	T1	T2	T3	T4
100	0	5	6	8	12
101	1	4	7	.	11

People missing any data are excluded (data from ID 101 are not included at all)

MLM via ML: uses **stacked** (long) data structure:

Only occasions with missing data are excluded

ID 100 uses 4 cases
ID 101 uses 3 cases

ID	Sex	Time	Y
100	0	1	5
100	0	2	6
100	0	3	8
100	0	4	12
101	1	1	4
101	1	2	7
101	1	3	.
101	1	4	11

Time can also be **unbalanced** across people such that each person can have his or her own measurement schedule: Time "0.9" "1.4" "3.5" "4.2"…

Moving Beyond a Random Intercept

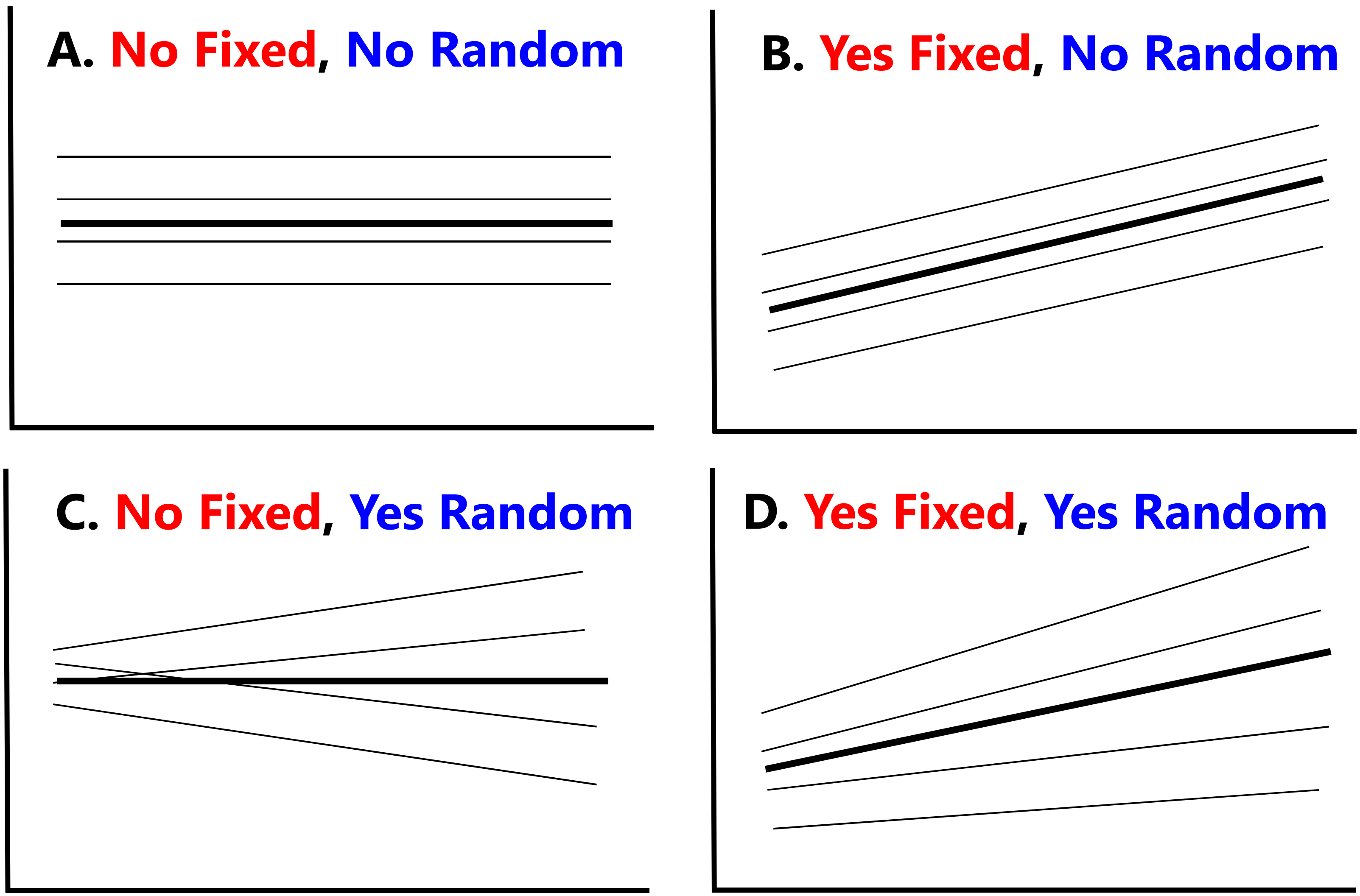
- 2 questions about the possible effects of *time*:

1. **Is there an effect of time on average?**
   - If the line describing the sample means not flat?
   - Significant **FIXED** effect of time

2. **Does the average effect of time vary across individuals?**
   - Does each individual need his or her own line?
   - Significant **RANDOM** effect of time

Fixed and Random Effects of **Time**

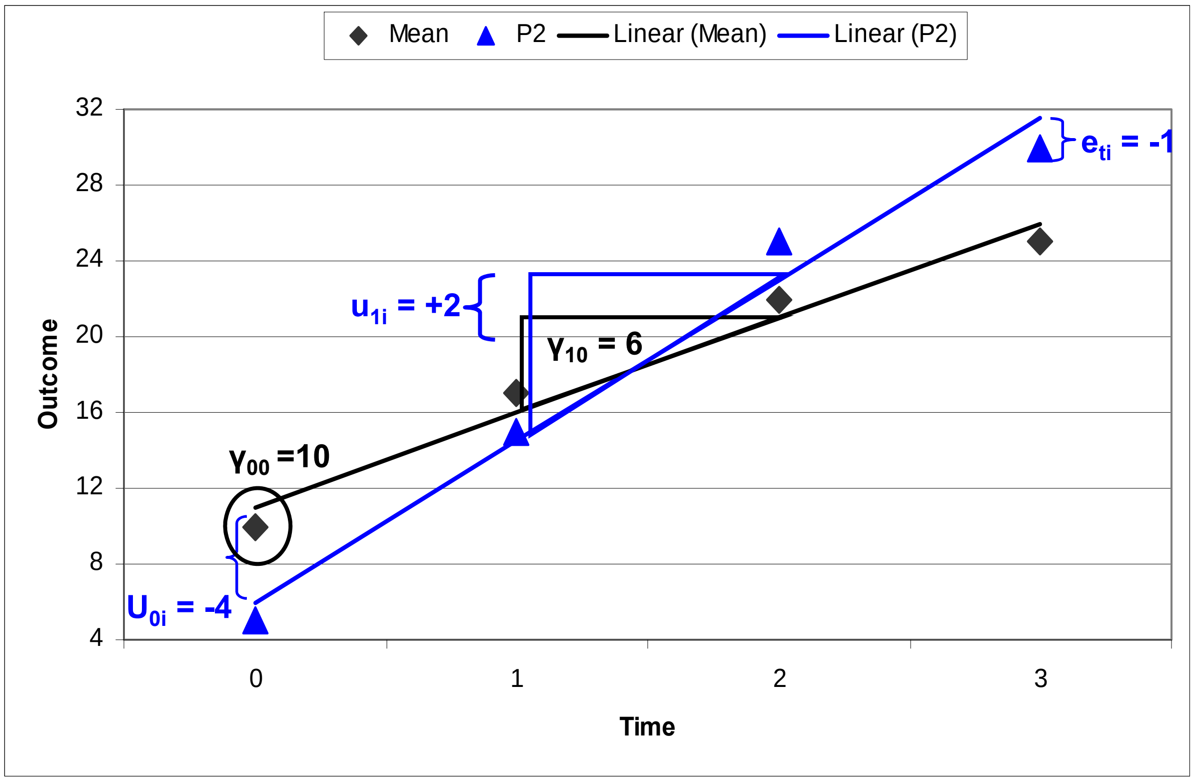
(Note: The intercept is random in every figure)



A "Random Linear Time" Model

$$y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i})(Time_{ti}) + e_{ti}$$

- γ_{00}: Fixed Intercept
- U_{0i}: Random Intercept Deviation
- γ_{10}: Fixed Slope
- U_{1i}: Random Slope Deviation
- e_{ti}: error for person i at time t



6 Parameters:

2 Fixed Effects:

γ_{00} Intercept, γ_{10} Slope

2 Random Effects Variances:

U_{0i} Intercept Variance $= \tau^2_{U_0}$

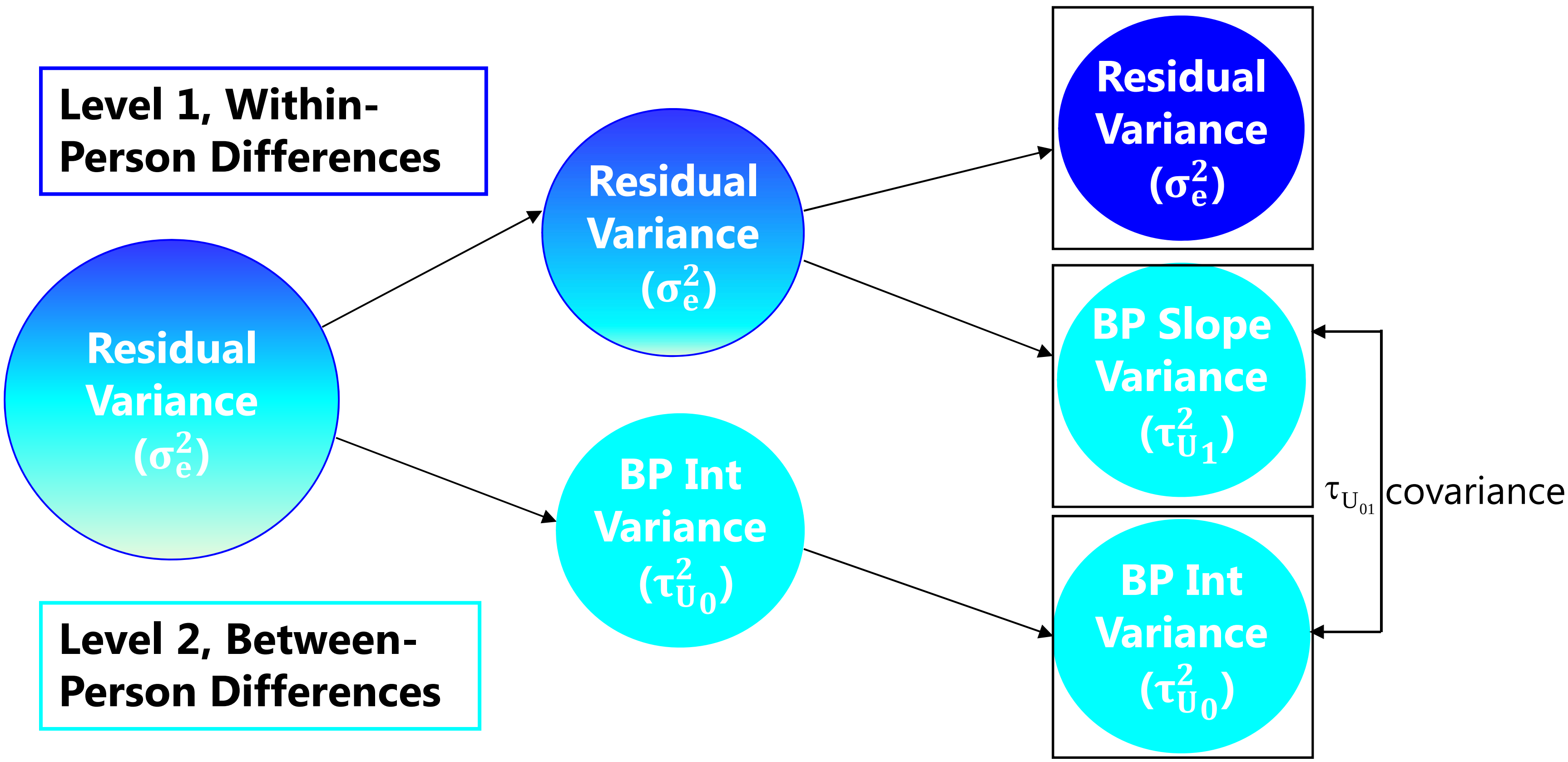
U_{1i} Slope Variance $= \tau^2_{U_1}$

Int-Slope Covariance $= \tau_{U_{01}}$

1 e_{ti} Residual Variance $= \sigma^2_e$

MLM "Handles" Dependency

- Where does each kind of person dependency go? Into a new random effects variance component (or "pile" of variance):



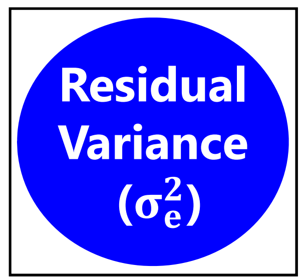
Piles of Variance (as Random Effects)

- By adding a random slope, we **carve up** our total variance into 3 piles:
  - BP (error) variance around intercept
  - BP (error) variance around slope
  - WP (error) residual variance

These 2 piles are just 1 pile of "error variance" in Univariate RM ANOVA

- **But making piles does NOT make error variance go away…**

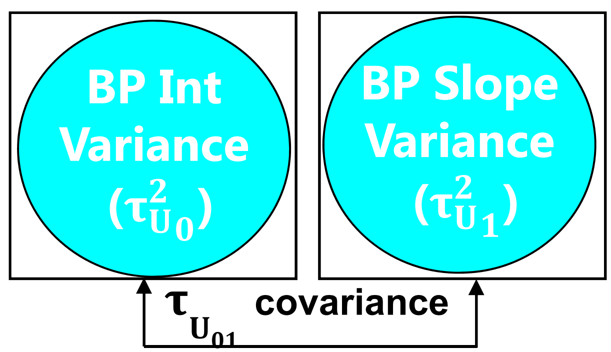
Level 1 (one source of) <u>Within-Person Variation</u>: gets accounted for by time-level predictors



FIXED effects make variance go away (explain variance).

RANDOM effects just make a new pile of variance.

Level 2 (two sources of) <u>Between-Person Variation</u>: gets accounted for by person-level predictors



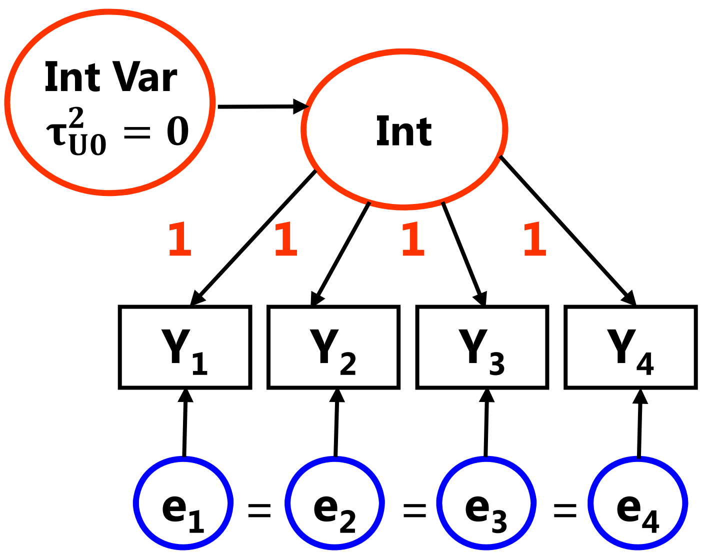
Options for Longitudinal Models

- Although models and software are separate, longitudinal data can be analyzed via multiple analytic frameworks:
  - **"Multilevel/Mixed Models"**
    - Dependency over time, persons, groups, etc. are modeled via random effects (multivariate → univariate through "levels" of stacked/long data)
    - Builds on GLM, generalizes more easily to additional levels of analysis and crossed sampling (e.g., if people change groups over time)
  - **"Structural Equation Models"**
    - Dependency over time *only* is modeled via latent variables (single-level analysis using multivariate/wide data)
    - Generalizes easier to broader analysis of latent constructs, mediation, and multivariate multilevel models in general (aka, "Multilevel SEM")

  - Because random effects and latent variables are the same thing, many longitudinal models can be specified/estimated either way

Random Effects as Latent Variables

- **BP model: e_{ti}-only model for the variance**
  - **$y_{ti} = \gamma_{00} + e_{ti}$**



Mean of the intercept factor = fixed intercept γ_{00}

Loadings of intercept factor = 1 (all occasions contribute equally)

Item intercepts = 0 (always)

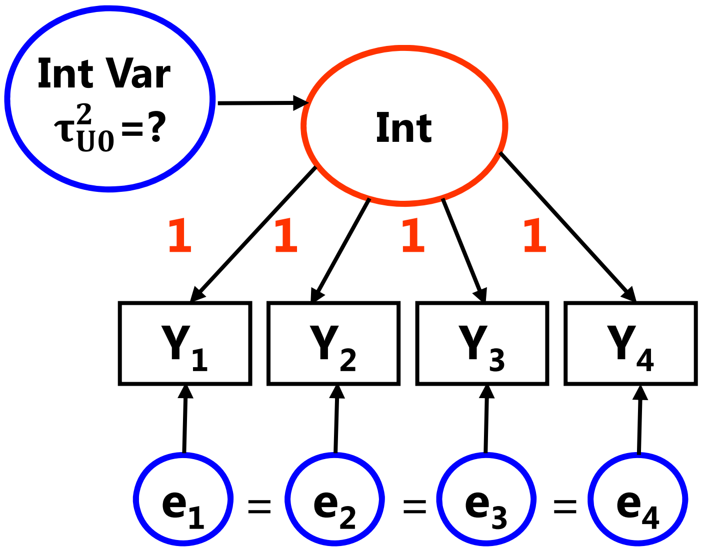
Variance of intercept factor = 0 so far

Residual variance (e) is assumed to be equal across occasions

  - After controlling for the *fixed* intercept, residuals are assumed uncorrelated

Random Effects as Latent Variables

- **+WP model: U_{0i} + e_{ti} model for the variance**
  - **$y_{ti} = \gamma_{00} + U_{0i} + e_{ti}$**



Mean of the intercept factor = fixed intercept γ_{00}

Loadings of intercept factor= 1 (all occasions contribute equally)

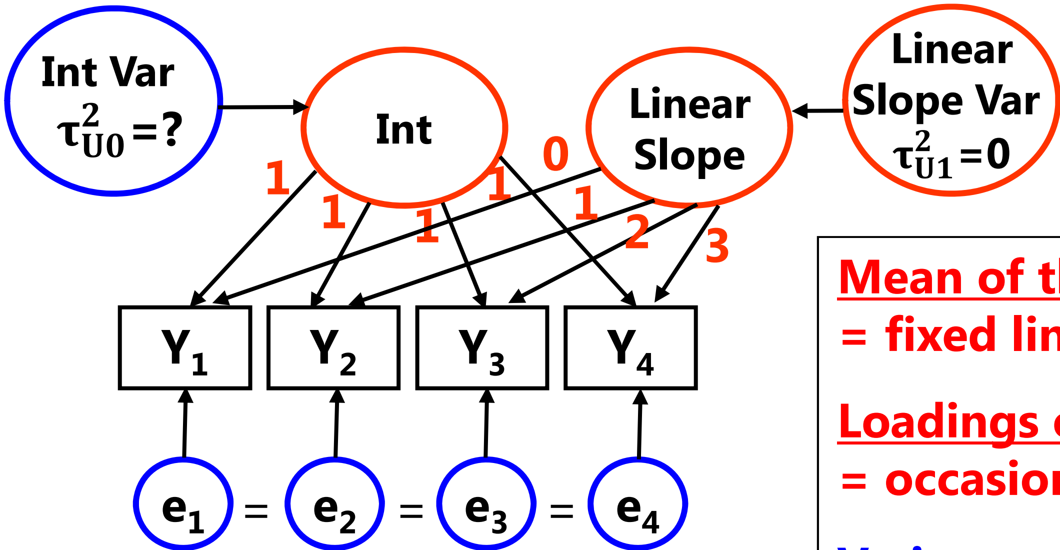
Variance of intercept factor = random intercept variance

Residual variance (e) is assumed to be equal across occasions

  - After controlling for the *random* intercept, residuals are assumed uncorrelated

Random Effects as Latent Variables

- **Fixed linear time, random intercept model:**
  - $y_{ti} = \gamma_{00} + (\gamma_{10}Time_{ti}) + U_{0i} + e_{ti}$



Mean of the linear slope factor = fixed linear slope γ_{10}

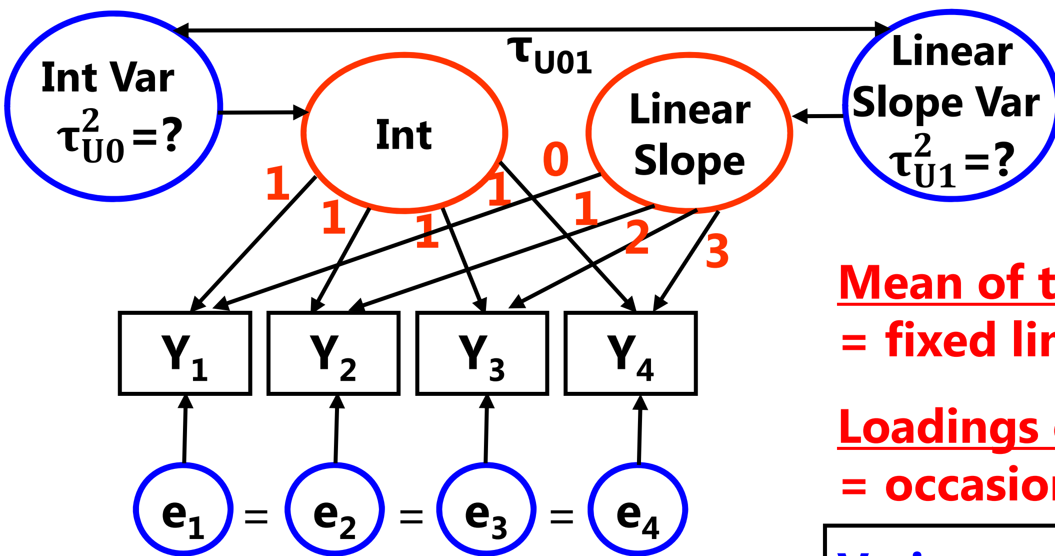
Loadings of linear slope factor = occasions (keep real time)

Variance of linear slope factor = 0

- After controlling for the *fixed linear slope and random intercept*, residuals are assumed uncorrelated

Random Effects as Latent Variables

- **Random linear time model:**
  - $y_{ti} = \gamma_{00} + (\gamma_{10}\text{Time}_{ti}) + U_{0i} + (U_{1i}\text{Time}_{ti}) + e_{ti}$



Mean of the linear slope factor = fixed linear slope γ_{10}

Loadings of linear slope factor = occasions (keep real time)

Variance of linear slope factor = random slope variance

  - After controlling for the *random linear slope and random intercept*, residuals are assumed uncorrelated

Intermediate Summary

- Longitudinal models use random effects/latent variables to quantify and predict sources of variability:
  - Between persons (BP) in intercept and (aspects of) change over time
    - Why do people start at different places?
    - Why do people change at different rates?
  - Within persons (WP) after controlling for individual change
    - Why are you off your line today?

- Individuals are conceptualized as **continuously varying** from one another in each of the between-person dimensions
  - If so, then one set of variances describes the entire sample
  - What if that's not the case?
  - Enter the "heterogeneous population model" variants…

Models for Finding "Hidden" Groups

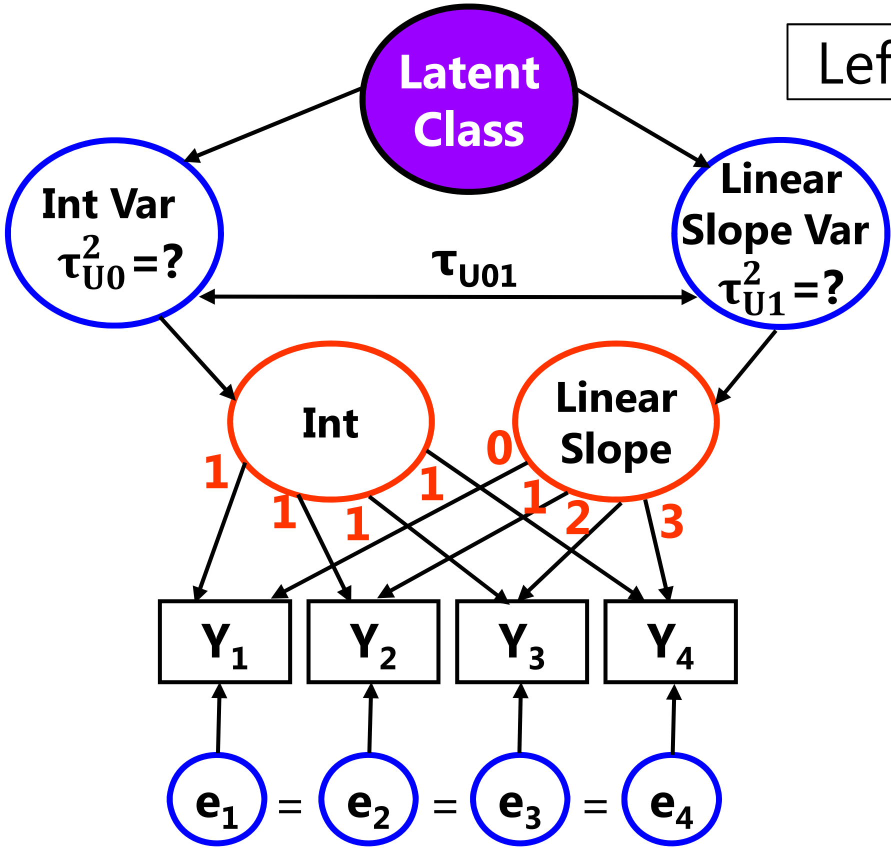
- Related to traditional cluster analysis (using least squares)
  - ML variants were popularized by Lazarsfeld and Henry (1968)

- Instead of continuous individual variation, models postulate existence of qualitatively different latent (hidden) subgroups

- More generally known as "finite mixture models," specific model names depend on type of outcomes to be classified:
  - Categorical, cross-sectional outcomes? "Latent class analysis"
  - Continuous, cross-sectional outcomes? "Latent profile analysis"
  - Change in group status over time? "Latent transition analysis"
  - Change in longitudinal outcomes over time? "Growth mixture models"
  - All have similar limitations, but we'll focus on **growth mixture models**

Growth Mixture Models (GMM)

Left: typical depiction of a GMM



- **Latent Class** = categorical unobserved variable that predicts probabilistic membership in c classes

GMMs are advertised being able to detect differing latent trajectories across people, but as used in practice, they have significant limitations:

1. Completely exploratory
2. Sensitive to non-normality
3. Distort individual variability
4. Classes can only predict existing random effects
5. Classes are not needed to examine prediction

1. GMMs are exploratory

- **How many classes?** ***????***
  - Programs provide relative goodness of fit info, but simulation results suggest these are problematic in practice
    - Information criteria (AIC, BIC) are inadequate for determining # classes
    - Entropy based on classification is only valid if the model fits…
- **What are the classes? How should they differ?** ***????***
  - Nature of the classes is determined entirely by the program
  - Get probability of membership to each class for each person, but this will likely change after predicting class membership
    - Should NEVER use the most likely class as an observed variable!

2. GMMs predict non-normality

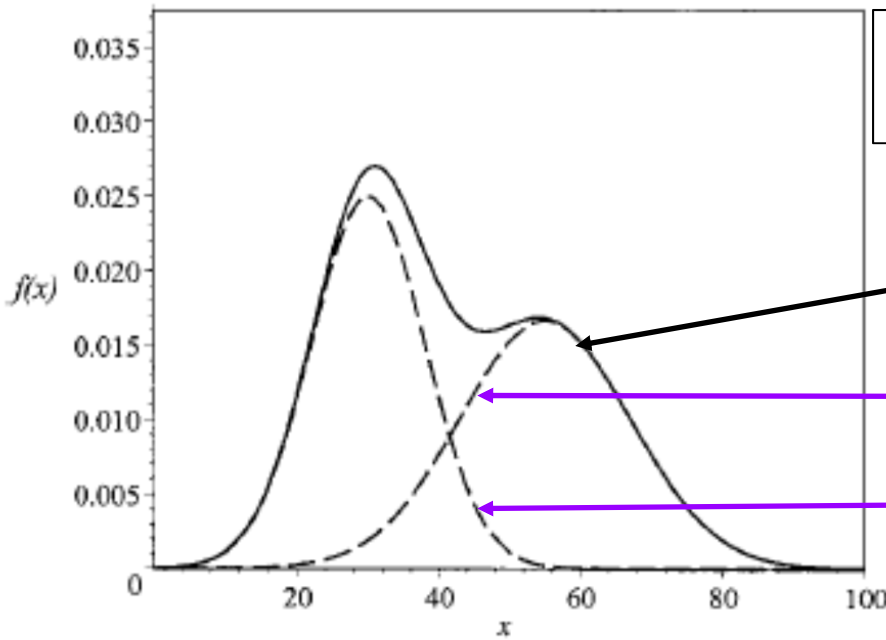


Figure 2 from Bauer & Curran (2003, *Psychological Methods*)

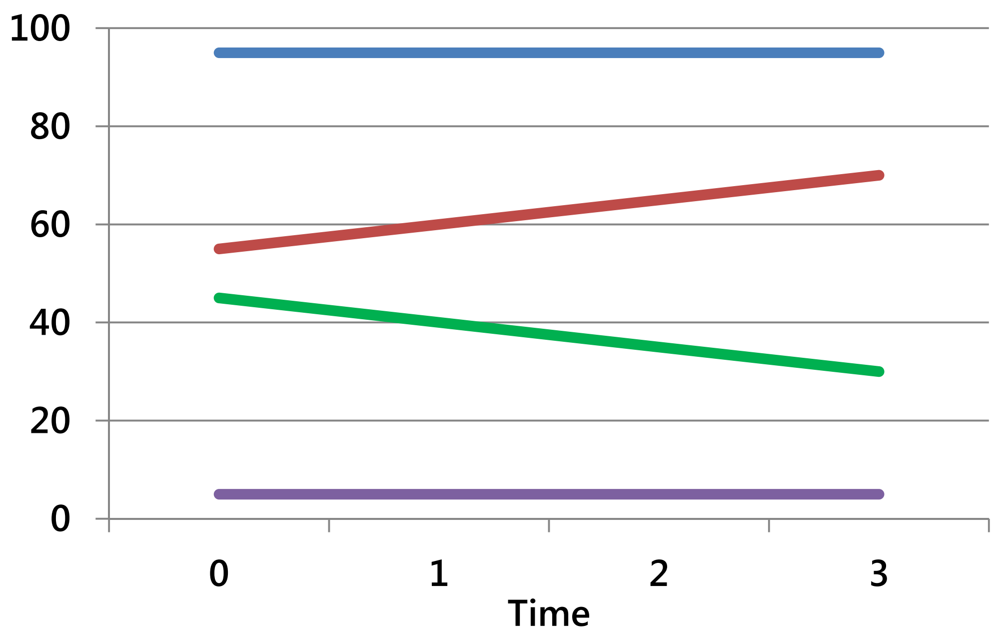
The overall non-normal distribution in sample…

… is described by a mixture of two normal distributions instead

- A lesser-known but statistically indistinguishable purpose of GMMs is to approximate a non-normal overall distribution
  - **So if you fit a GMM erroneously assuming conditional normality, you WILL find two or more latent classes for that reason alone**

3. GMMs distort individual variability

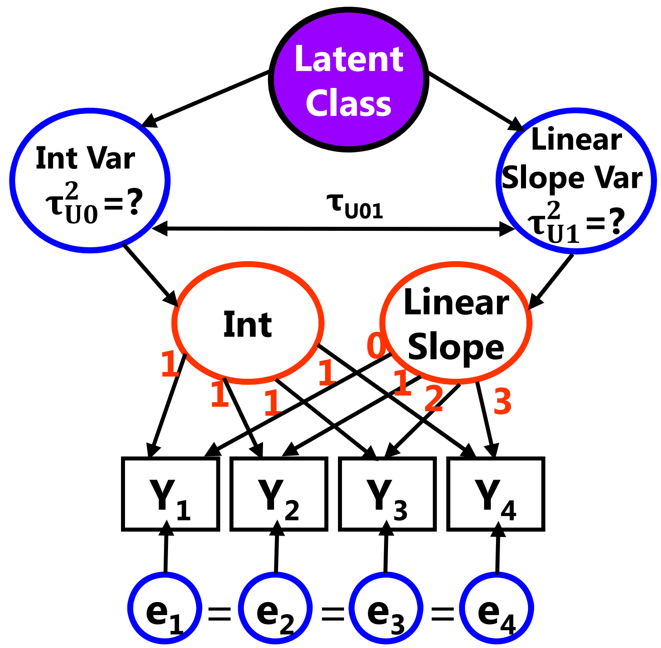
- What about individual differences within classes?
  Well, that depends on the program, too:
  - SAS PROC TRAJ: What variability?
  - Mplus: Variability is equal across classes, which is likely to be logically impossible… (but freeing this constraint leads to estimation problems)

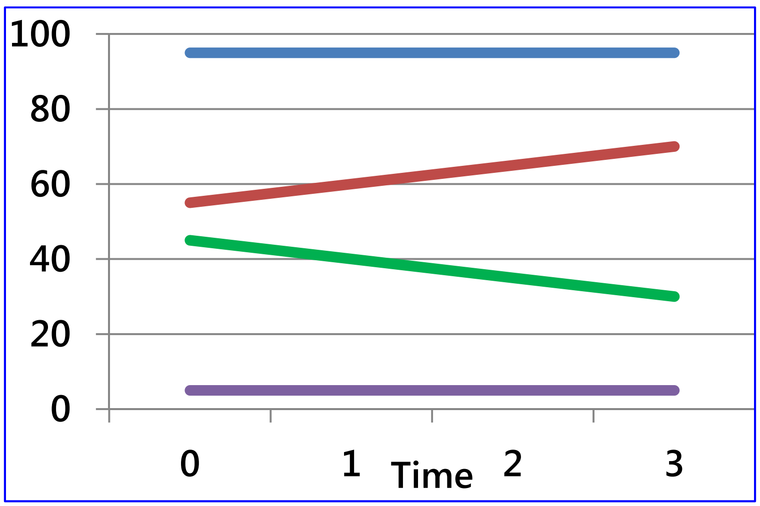


For example: The "high stable" and "low stable" groups cannot possibly have the same intercept and slope variability as the other groups…

4. GMMs can only predict model-specified random effects

- Latent classes serve to categorize existing intercepts and slopes…
  - For example, given the specification of a random linear time slope model, latent groups may only differ in level and kinds of linear change…

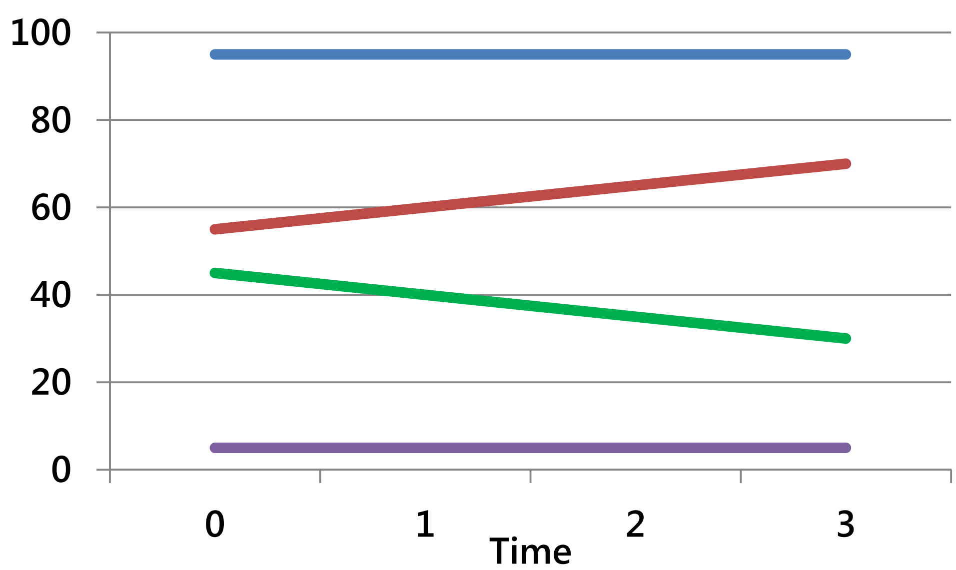


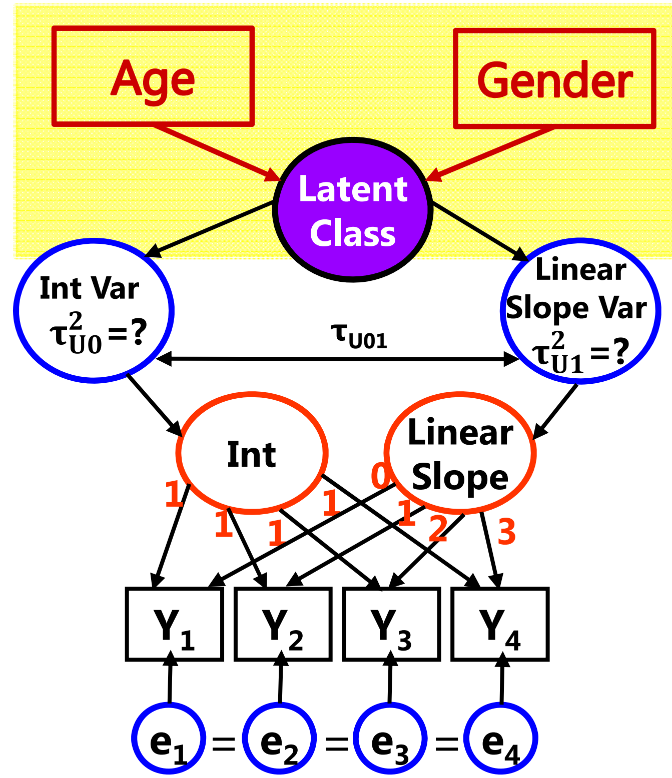


… just as people *already* do in the random linear growth model!

Latent classes just get in the way.

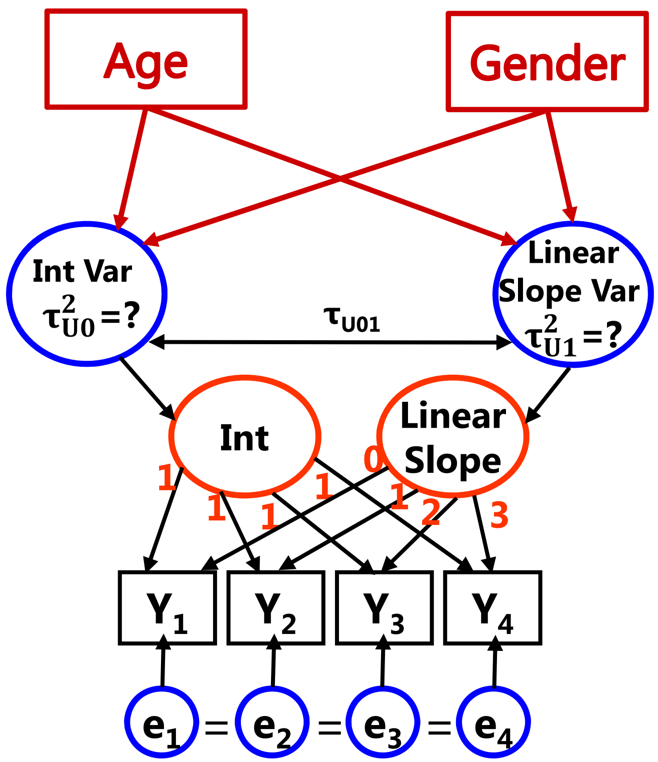
5. GMMs are not needed to examine prediction

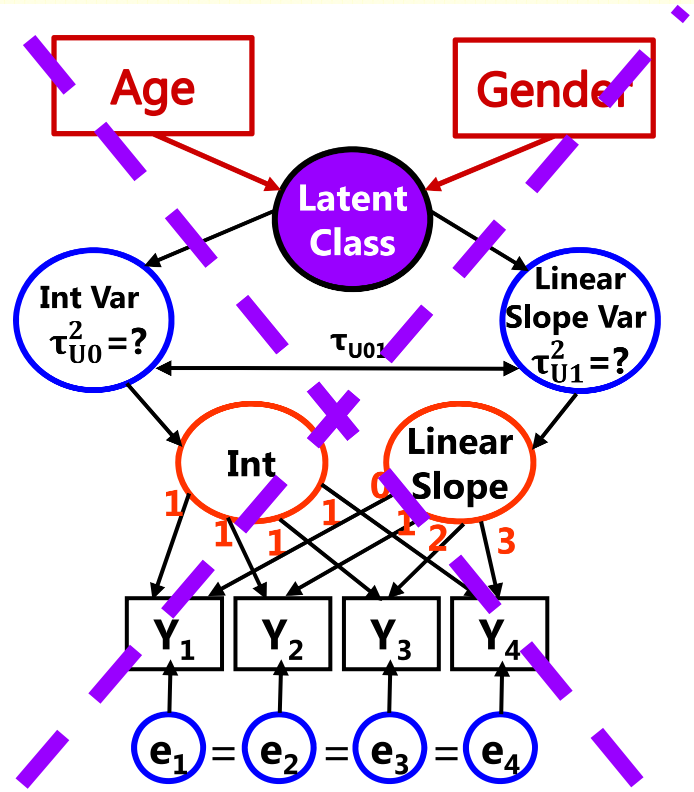




- After fitting a GMM, it is often of interest to then **predict class membership from covariates**…

5. GMMs are not needed to examine prediction





- …but the covariates should directly predict the random intercept and slopes themselves instead!

So what should we do instead?

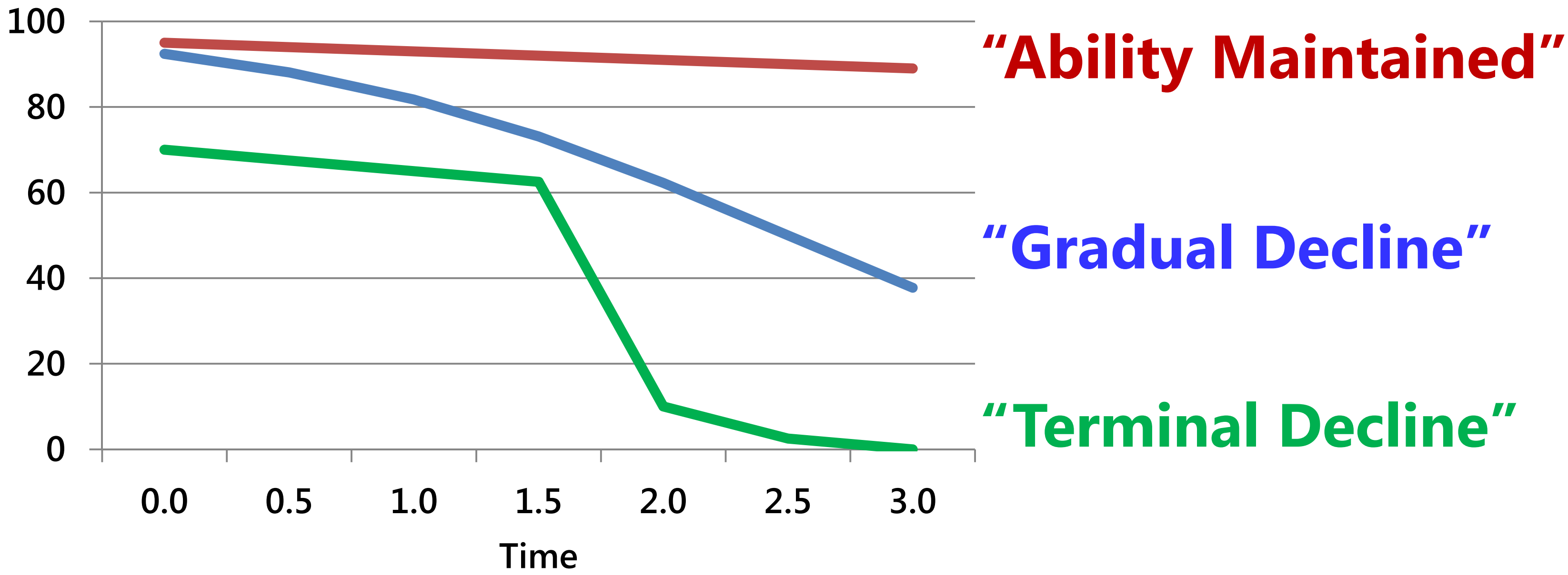
- Before fitting a typical GMMs, specify the most **appropriate conditional outcome distribution**
  - Account for floor/ceiling effects of observed measures

- Determine if groups are really necessary to answer your questions… for prediction of differences, probably not!

- If differences due to known predictors are of interest, consider **location-scale longitudinal models** instead (i.e., heterogeneous variance models; see Don Hedeker's work)
  - Allows for prediction of mean differences in intercepts and slopes, as well as prediction of differences in their **amount of variability**
  - Also very useful to intra-individual variability (IIV) designs

So what should we do instead?

- Specify a **confirmatory, hypothesis-driven model** that defines the different group trajectories *a priori*
- Most useful given **qualitatively different** kinds of change
  - Fit different model of change within each group
  - Constrain parameters as needed to ensure order/interpretability



Conclusions

- Longitudinal models with random effects/latent variables expand on traditional RM ANOVA via ML:
  - Multiple sources of between-person differences (from random intercepts only to random slopes for change)
  - Individuals vary continuously from another in growth terms

- Typical uses of growth mixture models try to describe these continuous differences via latent groups instead
  - But are completely exploratory, sensitive to violations of distributional assumptions, inflexible with respect to forms of change, and get in the way of predicting individual differences
  - Confirmatory models may remedy these problems but are seldom used in longitudinal applications