

# Introduction to Multivariate (Repeated Measures) Models

- Today's Class:
  - Where we've been so far: GLM for univariate outcomes
  - From univariate models to multivariate models
  - Maximum likelihood and model comparisons
  - Repeated measures models for the variance

# The Two Sides of \*Any\* Model

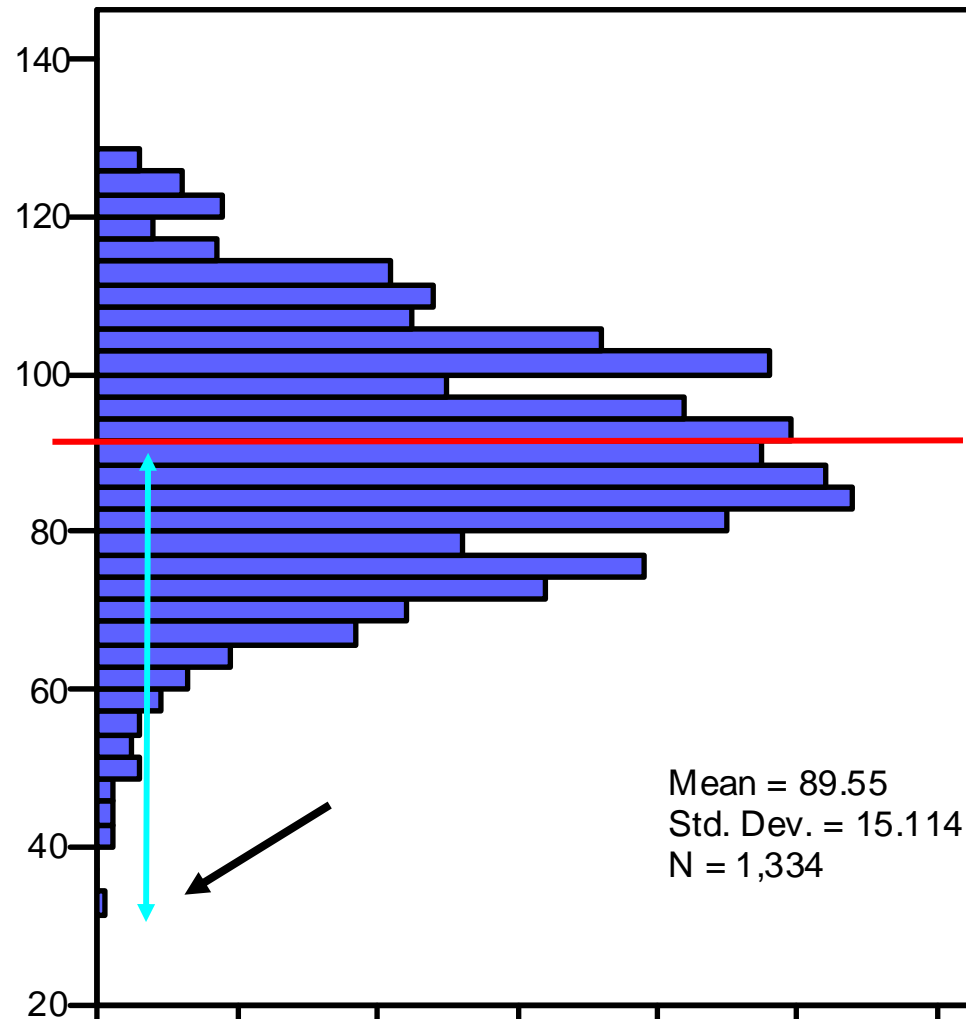
- Model for the Means:

- *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 on predictor variables

- Model for the Variance:

- *Aka* **Random Effects and Residuals**, Stochastic Part of Model
- How residuals are distributed and related across observations
- What you are used to **making assumptions about** instead...
- For general linear models, that residuals come from a **normal** distribution, are **independent** across persons, and have **constant variance** across persons and predictors (“identically distributed”)

# The Simplest General Linear Model: The “Empty” Model



$$y_i = \beta_0 + e_i$$

**Filling in values:**

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

$y_{\text{pred}}$  is also called  $\hat{y}$  ("y-hat")

Model for the Means

**$y_i$  error variance:**

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

# The Two Sides of a *General Linear Model*

$$y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i \dots + e_i$$

Our new focus

- **Model for the Means (Predicted Values):**

- A weighted linear function of predictor values creates each person's SINGLE expected (predicted) outcome (i.e., this is a univariate model)
- Weights =  $\beta$  fixed effects  $\rightarrow$  can be linear slopes, nonlinear slopes, piecewise slopes, group intercept differences, group slope differences...
- Significance of single  $\beta$  fixed effects given via univariate Wald tests (Est / SE =  $t$  or  $z \rightarrow p$ -value; TEST/ESTIMATE/LINCOM  $\rightarrow$  model-implied)
- Significance of multiple  $\beta$  fixed effects tested using multivariate Wald test (all effects=0?) via TEST/CONTRAST
  - Test of overall  $R^2$ ? Include \*all\* model predictor fixed effects
  - Test of change to  $R^2$ ? Include just \*new\* predictor fixed effects
  - "Omnibus" (ANOVA) group differences? Include all group-specific effects

# The Two Sides of a *General Linear Model*

$$y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i \dots + e_i$$

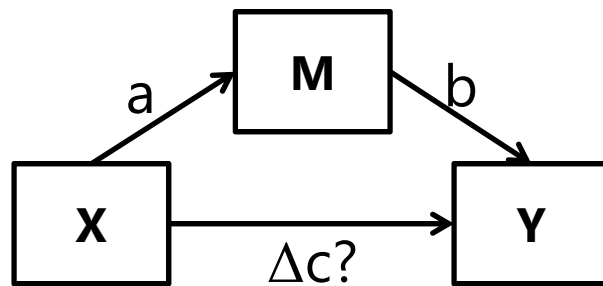
Our new focus

## • Model for the Variance:

- $e_i \sim N(0, \sigma_e^2) \rightarrow$  ONE residual (unexplained) deviation
- $e_i$  has a mean of 0 with some estimated constant variance  $\sigma_e^2$ , is normally distributed, is unrelated to predictors, and is unrelated across observations (across all people here)
- **Estimated parameter is residual variance** (not each  $e_i$ )
- What happens when each person has more than one  $y_i$ ?  
A single independent  $e_i$  will not be sufficient because:
  - Each outcome may have a different amount of residual variance
  - Residuals of outcomes from the same person will be correlated
  - So we need multivariate models a new data structure...

# Uses of “Multivariate” Models:

- **When  $y_i$  is still a single outcome conceptually, but:**
  - You have more than one outcome per person created by multiple conditions (e.g., longitudinal or repeated measures designs)
  - When your  $y_i$  comes from people nested or clustered in groups, such that you really have multivariate outcomes of a group (e.g., children nested in teachers, people nested in families)
- **When your hypotheses involve more than one  $y_i$ :**
  - To compare predictor effect sizes across outcomes (e.g., is a treatment effect bigger on outcome A than outcome B?)
  - You want to test indirect effects among them (i.e., mediation):



In this “path model” M is an outcome of X and a predictor of Y

Indirect effect =  $\Delta c? = a*b=0?$

# The General Linear Model

- The general linear model incorporates many different labels of related analyses under one unifying umbrella term:

	Categorical X's	Continuous X's	Both Types of X's
Univariate (one outcome)	"ANOVA"	"Regression"	"ANCOVA"
Multivariate (2+ outcomes)	"MANOVA"	"Multivariate Regression"	"MANCOVA"

- What these models all have in common is the use of a **normal conditional distribution** (for the residuals that remain after creating conditional outcomes from the model predictors)
- The use of these words almost always means **estimation** using "**least squares**" (LS), *aka* "**ordinary least squares**" (OLS)

# From One Outcome to Many...

- Least squares (LS) has a “closed form” solution (its “sums of squares” formulae) when used for GLM for single outcomes
- For GLM for multiple outcomes, LS quickly becomes useless...
  - Cannot handle missing outcomes (listwise-deletes entire person instead)
  - Only two options for modeling residual correlation between outcomes
  - Requires balanced data (same number of outcomes per higher unit)
- This is part of the reason why we’ve been using (restricted) maximum likelihood via **MIXED** instead of LS GLM:
  - **Convenience:** MIXED can produce fixed effects and contrasts that are model-implied, but not directly given (e.g., simple slopes of interactions)
  - **Model comparisons:** can be requested in a single step for any combination of effects (e.g., test for change in  $R^2$  from any new effects)
  - **Generalizability:** We can estimate **univariate or multivariate** models for normal outcomes using the same MIXED routine

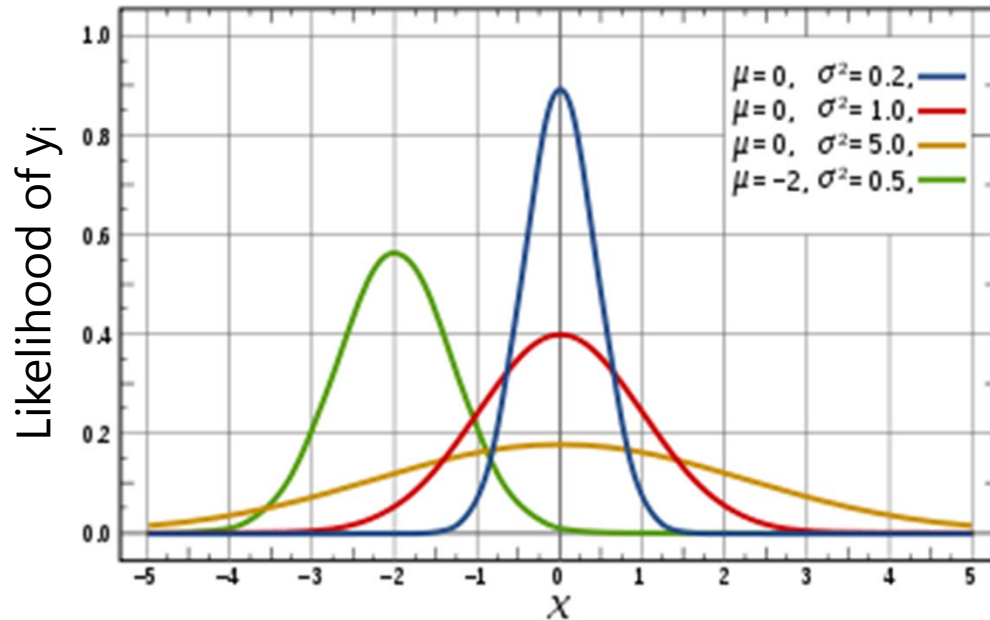


# End Goals of Maximum Likelihood Estimation

1. Obtain “most likely” values for each unknown model parameter (fixed effects, variances of residuals, and any other sources of variance and covariance) → **the estimates**
2. Obtain an index as to how likely each parameter value actually is (i.e., “really likely” or pretty much just a guess?) → **the standard error (SE) of the estimates**
3. Obtain an index as to how well the model we’ve specified actually describes the data → **the model fit indices**

**How does all of this happen? Probability distributions!**  
(i.e., probability density functions, or PDFs)

# Univariate Normal Distribution



Univariate Normal PDF:

$$f(y_i) = \frac{1}{\sqrt{2\pi\sigma_e^2}} * \exp\left[-\frac{1}{2} * \frac{(y_i - \hat{y}_i)^2}{\sigma_e^2}\right]$$

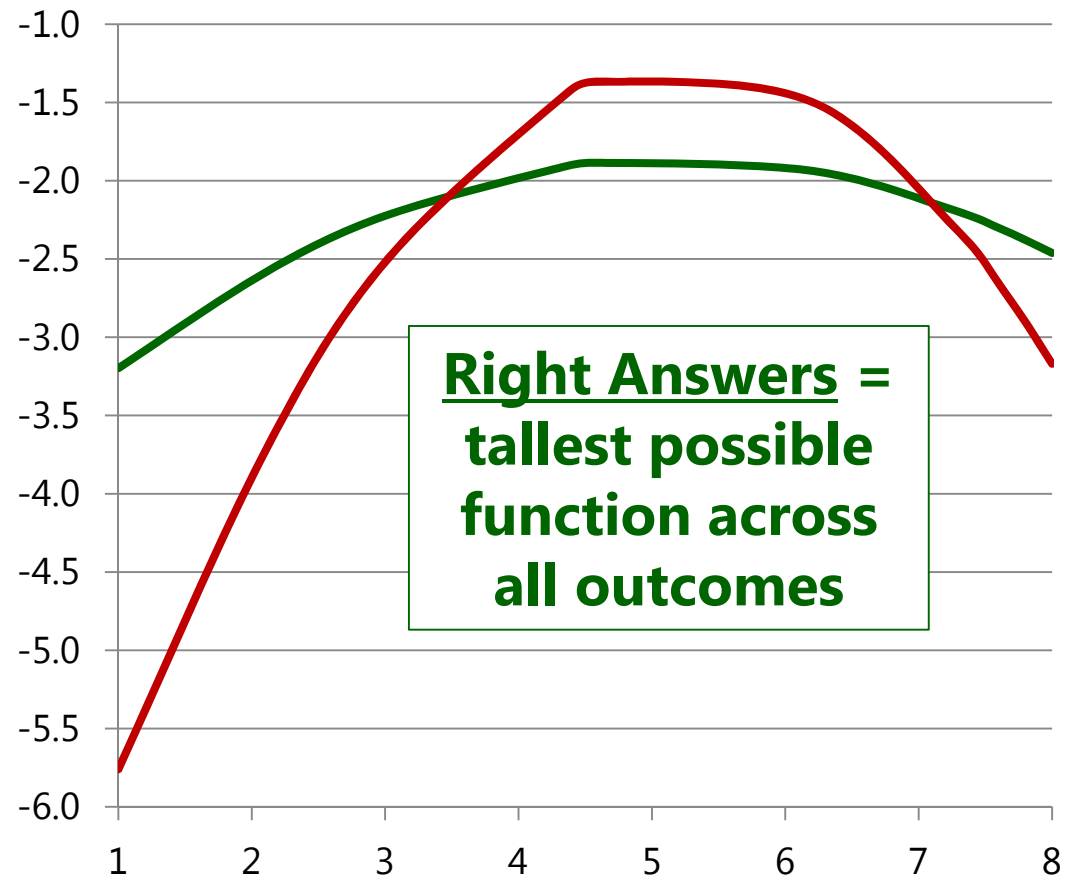
Sum over persons for log of  $f(y_i)$  =  
Model Log-Likelihood  $\rightarrow$  Model Fit

- This PDF tells us how **likely** any value of  $y_i$  is given two pieces of info:
  - Conditional mean  $\hat{y}_i$
  - residual variance  $\sigma_e^2$
- We can see this work using the NORMDIST function in excel!
  - Easiest for empty model:  
 $y_i = \beta_0 + e_i$
- We can check our math via MIXED!

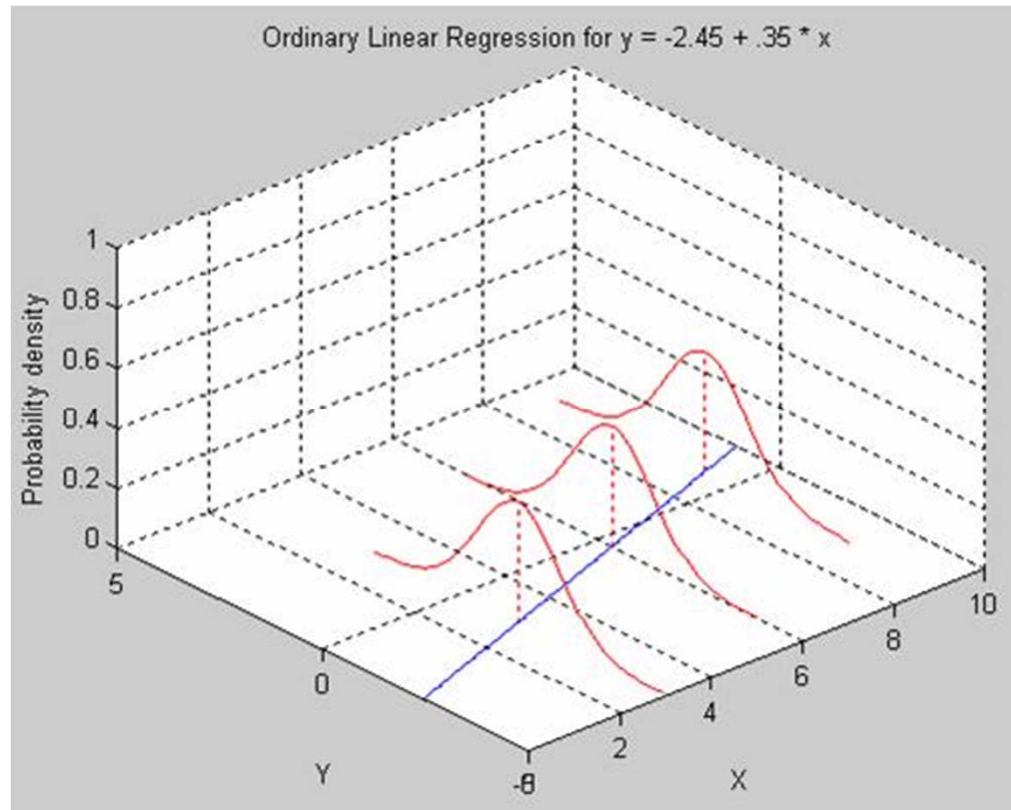
# ML via Excel “NORMDIST”

**Key idea: Normal Distribution formula → data height**

Mean	5.19	5.24
Variance	6.56	2.00
	<b>Right</b>	<b>Wrong</b>
<b>Outcome</b>	<b>Log(Height)</b>	<b>Log(Height)</b>
1.0	-3.20	-5.76
2.1	-2.59	-3.73
3.0	-2.22	-2.52
4.3	-1.92	-1.49
4.6	-1.89	-1.37
6.2	-1.94	-1.50
7.3	-2.20	-2.33
7.6	-2.30	-2.66
7.8	-2.38	-2.90
8.0	-2.46	-3.17
<b>SUM = Model LL = taller is better</b>	<b>-23.09</b>	<b>-27.42</b>



# Conditional Univariate Normal



This function applies for any value of  $X$ , such as in GLM:

- Fixed effects (intercept, predictor slopes) create a conditional mean for each person,  $\hat{y}_i$
- We assume the same residual variance  $\sigma_e^2$  holds for all values of  $\hat{y}_i$

Univariate Normal PDF:

$$f(y_i) = \frac{1}{\sqrt{2\pi\sigma_e^2}} * \exp\left[-\frac{1}{2} * \frac{(y_i - \hat{y}_i)^2}{\sigma_e^2}\right]$$

$$\begin{aligned} y_i &= \beta_0 + \beta_1 X_i + e_i \\ \hat{y}_i &= \beta_0 + \beta_1 X_i \\ e_i &= y_i - \hat{y}_i \quad \sigma_e^2 = \frac{\sum_{i=1}^N e_i^2}{N-2} \end{aligned}$$

# Review: Variances and Covariances

## Variance:

Dispersion of  $y$

$$\text{Variance } (y_d) = \frac{\sum_{i=1}^N (y_{di} - \hat{y}_{di})^2}{N - k}$$

## Covariance:

How  $y$ 's go together,  
unstandardized

$$\text{Covariance } (y_1, y_2) = \frac{\sum_{i=1}^N (y_{1i} - \hat{y}_{1i})(y_{2i} - \hat{y}_{2i})}{N - k}$$

## Correlation:

How  $y$ 's go together,  
standardized (-1 to 1)

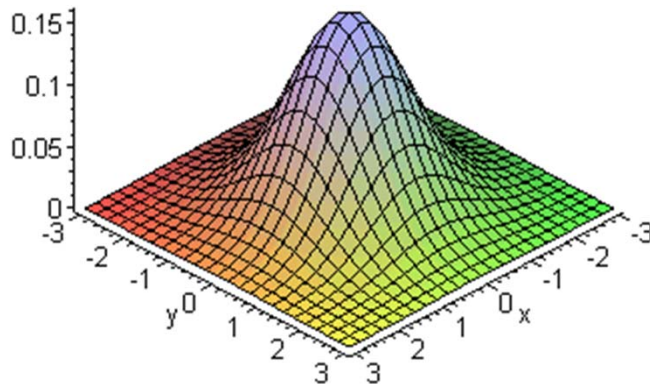
$$\text{Correlation } (y_1, y_2) = \frac{\text{Covariance}(y_1, y_2)}{\sqrt{\text{Variance}(y_1)} * \sqrt{\text{Variance}(y_2)}}$$

$N$  = # higher units,  $d$  = which DV (outcome),  $i$  = which unit,  
 $k$  = # fixed effects,  $\hat{y}_{di} = y_{di}$  predicted from fixed effects

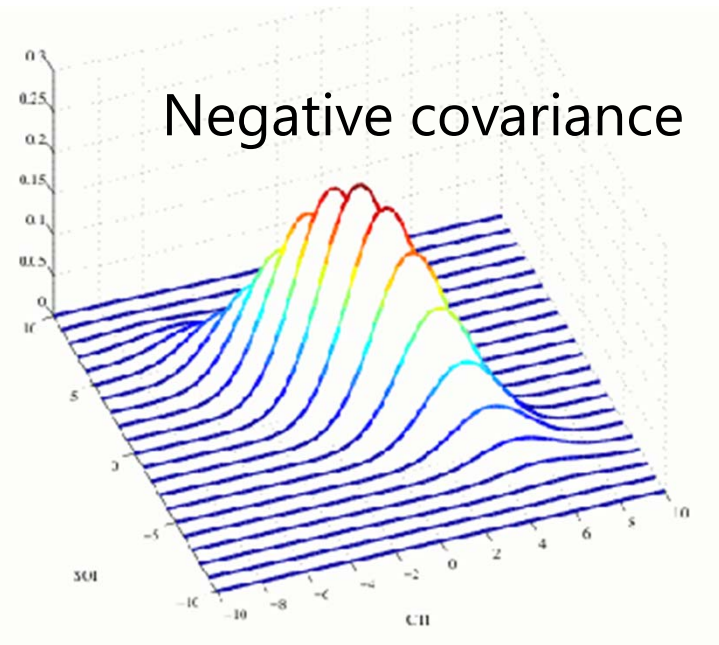
# Multivariate Normal Distribution

- Same principle as univariate, but log-likelihood is calculated for **each person's SET of outcomes** (then LL is summed over persons)
- Model parameters to be found include parameters that predict **EACH outcome's residual variance and their residual covariances**
- So each outcome's likelihood height has its own dimension, but the joint shape depends on the covariance between outcomes:

No covariance



Negative covariance



# Try, try, then try again...

- The best possible answers for the model parameters (e.g., fixed effects and residual variance) can be calculated via least squares given certain ideal circumstances:
  - Complete data, normally distributed residuals with constant variance, and only one dimension of sampling (i.e., single-level, univariate model)
- For almost all other analyses, the best possible estimates of these parameters have to be searched for iteratively
  - Different algorithms are used to decide which values to try given that each parameter has its own distribution of possible values → like an uncharted mountain in which each parameter gets its own dimension
  - Calculus helps the program scale this multidimensional mountain
    - At the top, all first partial derivatives (linear slopes at that point)  $\approx 0$
    - Positive first partial derivative? Too *low*, try again.
    - Negative first partial derivative? Too *high*, try again.
    - Matrix of partial first derivatives = “score function” = “gradient”

# End Goals 1 and 2: Model Estimates and SEs

- Process terminates (the model “converges”) when the next set of tried parameter values don’t improve the LL very much...
  - e.g., SAS MIXED default convergence criteria = .00000001
  - Those are the values for the parameters that, relative to the other possible values tried, are “most likely” → the estimates
- But we need to know how trustworthy those estimates are...
  - Precision is indexed by the steepness of the multidimensional mountain, where steepness → more negative partial second derivatives
  - Matrix of partial second derivatives = “Hessian matrix”
  - Hessian matrix  $\ast -1$  = “information matrix”
  - So steeper function = more information = more precision = smaller SE

$$\text{Each parameter SE} = \frac{1}{\sqrt{\text{information}}}$$



# End Goal #3: How well does the model fit?

- Multivariate models require assessment of **relative model fit**: how well does the model fit relative to other possible models?
- Relative fit is indexed by overall model **log-likelihood (LL)**:
  - Log of likelihood for each person's outcomes given model parameters
  - Sum log-likelihoods across all independent persons = **model LL**
  - Two flavors: Maximum Likelihood (ML) or Restricted ML (REML)
- What you get for this on your output varies by software...
- Given as  $-2 \times \log$  likelihood ( $-2LL$ ) in SAS or SPSS MIXED:  
–**2LL** gives BADNESS of fit, so **smaller** value = better model
- Given as just log-likelihood (LL) in STATA MIXED and Mplus:  
**LL** gives GOODNESS of fit, so **bigger** value = better model

# Assessing Relative Model Fit

- Nested models (i.e., in which one is a subset of the other) can now differ from each other in two important ways
- **Model for the Means** → which predictors and which fixed effects of them are included in the model
  - Does not require assessment of relative model fit using LL or  $-2LL$  (can still use univariate or multivariate Wald tests for this)
- **Model for the Variance** → what the pattern of variance and covariance of residuals from the same unit should be
  - DOES require assessment of relative model fit using LL or  $-2LL$
  - Cannot use the Wald test  $p$ -values that show up on the output for testing significance of variances because those  $p$ -values are use a two-sided sampling distribution for what the variance could be (but variances cannot be negative, so those  $p$ -values are not valid)

# Why should we care?

- What happens if we have the wrong model for the variance (i.e., our model predicts no covariance and/or equal variance across outcomes when this is not true)?
- Validity of the tests of the predictors depends on having the “most right” model for the variance
  - Estimates should be ok → come from model for the means
  - Standard errors (and thus p-values) can be inaccurate
- Unfortunately the model with the “right answer” is only possible to estimate in balanced data: When everyone has the same discrete set of multivariate outcomes
  - Tests of “absolute” fit are available in this case as well (stay tuned)
  - Otherwise we aim for “least wrong” among plausible alternatives

# Comparing Models for the Variance

- **Two main questions in choosing a model for the variance:**
  - How does the residual variance differ across outcomes?
  - How are the residuals from the same unit correlated?
- Nested models are compared using a **“likelihood ratio test”**:  
**–2ΔLL test** (aka, “ $\chi^2$  test” in SEM; “deviance difference test” in MLM)

“fewer” = from model with fewer parameters  
“more” = from model with more parameters

Results of 1. & 2. must be positive values!

1. Calculate **–2ΔLL**: if given  $-2LL$ , do  $-2\Delta LL = (-2LL_{\text{fewer}}) - (-2LL_{\text{more}})$   
if given  $LL$ , do  $-2\Delta LL = -2 * (LL_{\text{fewer}} - LL_{\text{more}})$
2. Calculate **Δdf** = (# Parms<sub>more</sub>) – (# Parms<sub>fewer</sub>)
3. **Compare –2ΔLL to  $\chi^2$  distribution with df = Δdf**
4. Get  $p$ -value from CHIDIST in excel or LRTEST option in STATA

# Comparing Models for the Variance

- What your  $p$ -value for the  $-2\Delta LL$  test means:
  - If you **ADD** parameters, then your model can get **better** (if  $-2\Delta LL$  test is significant) or **not better** (not significant)
  - If you **REMOVE** parameters, then your model can get **worse** (if  $-2\Delta LL$  test is significant) or **not worse** (not significant)
- Nested or non-nested models can also be compared by **Information Criteria** that also reflect model parsimony
  - No significance tests or critical values, just "smaller is better"
  - **AIC** = Akaike IC =  $-2LL + 2 * (\#parameters)$
  - **BIC** = Bayesian IC =  $-2LL + \log(N) * (\#parameters)$
  - What "parameters" means depends on flavor (except in stata):
    - ML = ALL parameters; REML = variance model parameters only

# Flavors of Maximum Likelihood

- Remember that Maximum likelihood comes in 2 flavors:
- **“Restricted (or residual) maximum likelihood”**
  - Only available for general linear models or general linear mixed models (that assume normally distributed residuals)
  - Is same as LS given complete outcomes, but it doesn't require them
  - Estimates variances the same way as in LS (accurate)  $\rightarrow \frac{\sum(y_i - y_{\text{pred}})^2}{N - k}$
- **“Maximum likelihood” (ML; also called FIML\*)**
  - Is more general, is available for the above plus for non-normal outcomes and latent variable models (CFA/SEM/IRT)
  - Is NOT the same as LS: it under-estimates variances by not accounting for the # of estimated fixed effects  $\rightarrow \frac{\sum(y_i - y_{\text{pred}})^2}{N}$
- *\*FI = Full information  $\rightarrow$  it uses all original data (they both do)*

# Flavors of Full-Information Maximum Likelihood

- Restricted maximum likelihood (REML; used in MIXED)
  - Provides unbiased variances  $\frac{\sum(y_i - y_{\text{pred}})^2}{N - k}$
  - Especially important for small N (< 100 units)
  - **-2ΔLL test** cannot be used to compare models differing in fixed effects (no biggee; we can do this using univariate or multivariate Wald tests)
  - **-2ΔLL test** MUST be used to compare different models for the variance
- Maximum likelihood (ML; also used in MIXED)  $\frac{\sum(y_i - y_{\text{pred}})^2}{N}$ 
  - Variances (and SEs) are too small in small samples
  - Is only option in most software for path models and SEM
  - **-2ΔLL test** can be used to compare **any** nested model; must be used to compare different models for the variance

# ML vs. REML in a nutshell

Remember "population" vs. "sample" formulas for calculating variance?

**"Population"**

$$\frac{\sum(y_i - y_{\text{pred}})^2}{N}$$

**"Sample"**

$$\frac{\sum(y_i - y_{\text{pred}})^2}{N - k}$$

<b>All comparisons must have same N!!!</b>	<b>ML</b>	<b>REML</b>
To select, type...	METHOD=ML (-2 log likelihood)	METHOD=REML <i>default</i> (-2 res log likelihood)
In estimating variances, it treats fixed effects as...	<b>Known</b> (df for having to also estimate fixed effects is not factored in)	<b>Unknown</b> (df for having to estimate fixed effects is factored in)
So, in small samples, L2 variances will be...	<b>Too small</b> (less difference after N=30-50 or so)	<b>Unbiased</b> (correct)
But because it indexes the fit of the...	<b>Entire model</b> (means + variances)	<b>Variances model only</b>
You can compare models differing in...	<b>Fixed and/or random effects</b> (either/both)	<b>Random effects only</b> (same fixed effects)



# Rules for Comparing Models

**All observations must be the same across models!**

**Compare Models Differing In:**

Type of Comparison:	Means Model (Fixed) Only	Variance Model (Random) Only	Both Means and Variances Model (Fixed and Random)
<b><u>Nested?</u></b> YES, can do significance tests via...	Fixed effect $p$ -values from ML or REML -- OR -- ML $-2\Delta LL$ only (NO REML $-2\Delta LL$ )	NO $p$ -values  REML $-2\Delta LL$ (ML $-2\Delta LL$ is ok if big N)	ML $-2\Delta LL$ only (NO REML $-2\Delta LL$ )
<b><u>Non-Nested?</u></b> NO signif. tests, instead see...	ML AIC, BIC (NO REML AIC, BIC)	REML AIC, BIC (ML ok if big N)	ML AIC, BIC only (NO REML AIC, BIC)

Nested = one model is a direct subset of the other

Non-Nested = one model is not a direct subset of the other

# Models for Repeated Measures (RM)

- We will start with repeated measures data in which  $y_i$  is still a single outcome conceptually, but you have more than one outcome per person created by multiple conditions
  - Traditionally RM data are analyzed using some kind of ANOVA
- If the RM data come from truly one outcome per condition, then so-called “ANOVA” models are potentially problematic restrictions of more general multivariate models (up next)
- But if the RM data really come from multiple outcomes per condition, they should NOT be aggregated into one outcome in order to use an ANOVA model (e.g., ANOVA on condition means created from trials of the same kind is a bad idea)
  - Stay tuned for what to do instead...

# ANOVA for RM or longitudinal data?

- There are 3 possible “kinds” of ANOVAs we could use:
  - Between-Persons/Groups, Univariate RM, and Multivariate RM
- **NONE OF THEM ALLOW:**
  - **Missing outcomes** (do listwise deletion due to least squares)
  - **Time-varying predictors** (covariates are person predictors only)
- Each includes the same “saturated” model for the means for condition: fixed effects create all possible mean differences (e.g., 4 parameters to create 4 distinct time condition means)
  - **“Saturated means model”**:  $\beta_0 + \beta_1(T_1) + \beta_2(T_2) + \beta_3(T_3)$
  - **The *time condition* variable must be balanced and discrete !**
- Types of ANOVAs differ by what they predict the variance and covariance across residuals from the same person to be...

# From a “Multivariate” to “Stacked” Data

**New data structure so that  $y_i$  is still a single outcome....**

RM ANOVA uses “wide”  
**multivariate** data structure:

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

A row = a case = a person

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

ML/REML in MIXED

uses “long” or  
**stacked** data structure instead:

A case is now one outcome per person

Only cases missing data are excluded

ID 100 uses 4 cases

ID 101 uses 3 cases

ID	Girl	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”...

# Telling MIXED you have repeated measures (dependent) data:

- New command to specify the multivariate residual distribution:
  - **REPEATED** in SAS/SPSS; “**residuals**” in STATA
  - Requires ID variables (what is repeated over what?) and TYPE option that specifies the pattern of residual variance and covariance you predict

## **R Matrix** (here, of 4 conditions):

diagonal = residual variances

off-diagonals = residual covariances

$$\begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22}^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}^2 \end{bmatrix}$$

## **RCORR Matrix** (SAS and STATA):

diagonal = standardized variance=1

off-diagonals = residual correlations

$$\begin{bmatrix} 1 & r_{12} & r_{13} & r_{14} \\ r_{21} & 1 & r_{23} & r_{24} \\ r_{31} & r_{32} & 1 & r_{43} \\ r_{41} & r_{42} & r_{43} & 1 \end{bmatrix}$$

# 1. Between-Groups (Independent) ANOVA

- **Uses a single estimate of residual variance  $\sigma_e^2$**  that is predicted to be the same across all outcomes
- **Predicts NO covariance of residuals from same person**
- Will usually be **very, very wrong** for repeated measures data
  - RM effects tested against wrong residual variance (significance tests will often be way too conservative)
- Predicts a variance-covariance matrix (here, for 4 outcomes) like this, called "**Variance Components**" (**R matrix** is TYPE=**VC** on SAS/SPSS REPEATED, or "independent" in STATA):  
$$\begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ 0 & 0 & \sigma_e^2 & 0 \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

# 2a. Univariate Repeated Measures

- Separates total variance into two sources:
  - **Between-Person mean differences** (called "CS" for now, stay tuned)
  - **Within-Person residual** (remaining variance due to  $\sigma_e^2$ )

- Predicts a variance-covariance matrix (here, for 4 outcomes) like this, called "**Compound Symmetry**" (**R matrix** is TYPE=**CS** on SAS/SPSS REPEATED, "exchangeable" in STATA):

$$\begin{bmatrix} \sigma_e^2 + CS & CS & CS & CS \\ CS & \sigma_e^2 + CS & CS & CS \\ CS & CS & \sigma_e^2 + CS & CS \\ CS & CS & CS & \sigma_e^2 + CS \end{bmatrix}$$

- Person mean differences are the only reason why occasions are correlated, so **RCORR** will look like this, in which ICC = intraclass correlation (stay tuned):

$$\begin{bmatrix} 1 & ICC & ICC & ICC \\ ICC & 1 & ICC & ICC \\ ICC & ICC & 1 & ICC \\ ICC & ICC & ICC & 1 \end{bmatrix}$$

- Will usually be at least somewhat wrong for repeated measures data, especially longitudinal data

# The Problem with Univariate RM ANOVA

- Univ. RM ANOVA ( $CS + \sigma_e^2$ ) predicts **compound symmetry**:
  - All residual variances and all covariances are equal across outcomes
  - In other words, the amount of error observed should be the same at any occasion, so a single, pooled error variance term makes sense
  - If not, tests of fixed effects may be biased (i.e., sometimes tested against too much or too little error, if error is not really constant over time)
  - **COMPOUND SYMMETRY RARELY FITS FOR LONGITUDINAL DATA**
- But to get the correct tests of the fixed effects, the data must only meet a less restrictive assumption of **sphericity**:
  - In English → **pairwise differences** between adjacent occasions have equal variance and covariance (satisfied by default with only 2 occasions)
  - If compound symmetry is satisfied, so is sphericity (but see above)
  - Significance test provided in ANOVA for where data meet sphericity assumption
  - **Other RM ANOVA approaches are used when sphericity fails...**



# The Other Repeated Measures ANOVAs...

- 2b. **Univariate RM ANOVA with sphericity corrections**

- Based on  $\epsilon$  → how far off sphericity (from 0-1, 1=spherical)
- Applies an overall correction for model df based on estimated  $\epsilon$ , but it doesn't really address the problem that data  $\neq$  model

- 3. **Multivariate Repeated Measures ANOVA**

- All variances and covariances are estimated separately (here, for 4 outcomes), called "Unstructured" (**R matrix** is TYPE=UN on SAS/SPSS REPEATED, "unstructured" in STATA)

$$\begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22}^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}^2 \end{bmatrix}$$

- **It's not a model, it IS the data!**

- Because it can never be wrong, UN can be useful for **complete and balanced multivariate data** with few outcomes (e.g., 2-4)

- Parameters =  $\frac{\#outcomes * (\#outcomes + 1)}{2}$  so can be hard to estimate

- Every other model for the variance is nested within Unstructured, so we can do model comparisons to see if all other models are NOT WORSE

# What Else is There? Lots!

- There are MANY other kinds of variance-covariance patterns allowed within MIXED (especially in SAS)
- Here's one more that can be useful in repeated measures data (although it's not available in STATA as near as I can tell)

## • Compound Symmetry Heterogeneous: TYPE=CSH

➤ Separate residual variance per outcome in **R matrix** →

➤ Because variances differ by outcome, then **R matrix** covariances differ too →

$$\begin{bmatrix} \sigma_{T1}^2 & CSH\sigma_{T1}\sigma_{T2} & CSH\sigma_{T1}\sigma_{T3} & CSH\sigma_{T1}\sigma_{T4} \\ CSH\sigma_{T2}\sigma_{T1} & \sigma_{T2}^2 & CSH\sigma_{T2}\sigma_{T3} & CSH\sigma_{T2}\sigma_{T4} \\ CSH\sigma_{T3}\sigma_{T1} & CSH\sigma_{T3}\sigma_{T2} & \sigma_{T3}^2 & CSH\sigma_{T3}\sigma_{T4} \\ CSH\sigma_{T4}\sigma_{T1} & CSH\sigma_{T4}\sigma_{T2} & CSH\sigma_{T4}\sigma_{T3} & \sigma_{T4}^2 \end{bmatrix}$$

➤ But covariance is created from a constant correlation: CSH as shown in **RCORR** →

$$\begin{bmatrix} 1 & CSH & CSH & CSH \\ CSH & 1 & CSH & CSH \\ CSH & CSH & 1 & CSH \\ CSH & CSH & CSH & 1 \end{bmatrix}$$

# Nesting of RM Models for the Variance

- **Worst –2LL: variance components** (TYPE=VC; “independent”)
  - Residual variances are equal across outcomes; no covariances
- **Best –2LL: unstructured** (TYPE=UN; “unstructured”)
  - Residual variances and covariances all estimated separately; no pattern
- Intermediate alternatives and their comparisons:
- **Compound Symmetry** (TYPE=CS; “exchangeable”)
  - Equal residual variances and covariances (and equal correlations)
  - If fits better than VC, there IS residual covariance among outcomes
  - If fits worse than UN, residual variances and covariances are not constant
- **Compound Symmetry Heterogeneous** (TYPE=CSH in SAS/SPSS)
  - Different residual variances but equal residual correlations
  - If fits better than CS, residual variances are not constant
  - If fits worse than UN, residual correlations are not constant

# Wrapping Up...

- So far in general linear models for univariate outcomes, we've focused solely on the model for the means
  - How to include and interpret fixed effects for any kind of predictor (continuous, semi-continuous, categorical...)
- But when analyzing data in which each sampling unit has more than one outcome...
  - ... repeated measures, longitudinal, or nested/clustered data...
- ...We also have to address the model for the variance
  - How does the residual variance differ across outcomes?
  - How are the residuals from the same unit correlated?
  - We have to use  $-2\Delta LL$  tests to compare nested models to decide which fits best, otherwise our fixed effect SEs are suspect