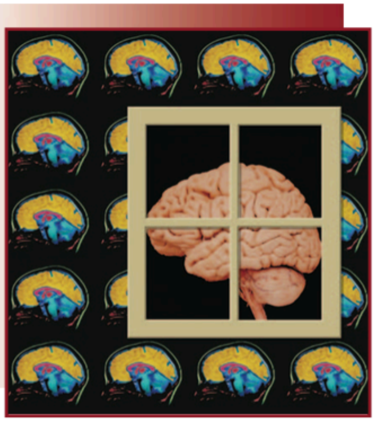
Group analysis

Wednesday, Lecture 1

Jeanette Mumford
University of Wisconsin - Madison

Paper



© DIGITAL STOCK

CTIONAL MAGNETIC RESONANCE IMAGING

Modeling and Inference of Multisubject fMRI Data

Using Mixed-Effects Models for Joint Analysis

BY JEANETTE A. MUMFORD AND THOMAS NICHOLS

Motivation

- How do you typically model repeated measures data?
  - Eg, behavioral study with 30 presentations of each of 6 stimuli?

Where we're going

- Fixed vs Mixed modeling
- 2 stage summary statistics approach to mixed model (using non-fmri example)
- 2 stage summary statistics approach with fMRI data
- Software differences (FSL and SPM)

Where we're going

- Fixed vs Mixed modeling
- 2 stage summary statistics approach to mixed model (using non-fmri example)
- 2 stage summary statistics approach with fMRI data
- Software differences (FSL and SPM)

Mixed Model Motivation

- Start with a simple ANOVA example
- Study: Is hair length different between males and females?

Mixed Model Motivation

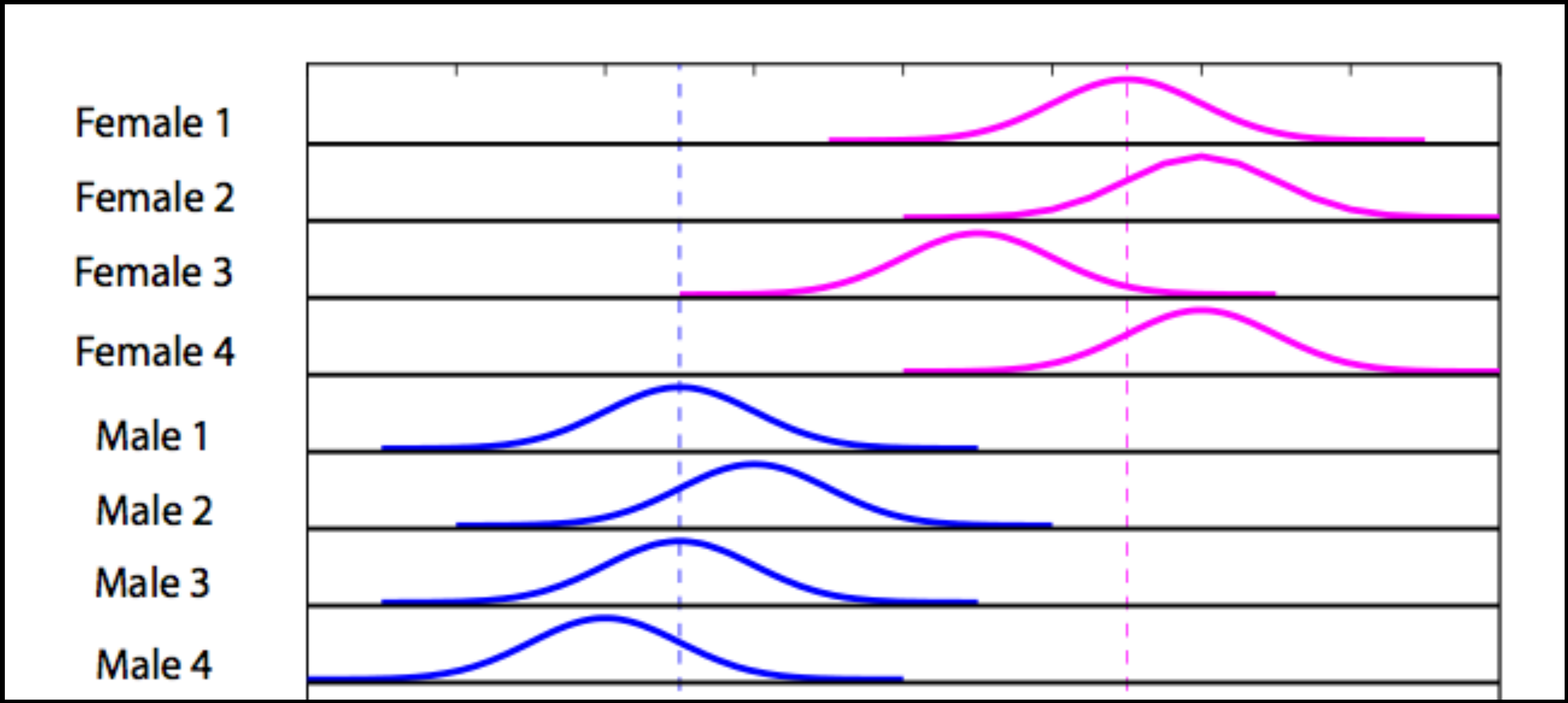
- Start with a simple ANOVA example
- Study: Is hair length different between males and females?

BTW, this example originally came from Friston and Holmes

Start: 1 hair per person

- Two sources of variability
  - Variance of hair length within person
  - Variance of hair length between people

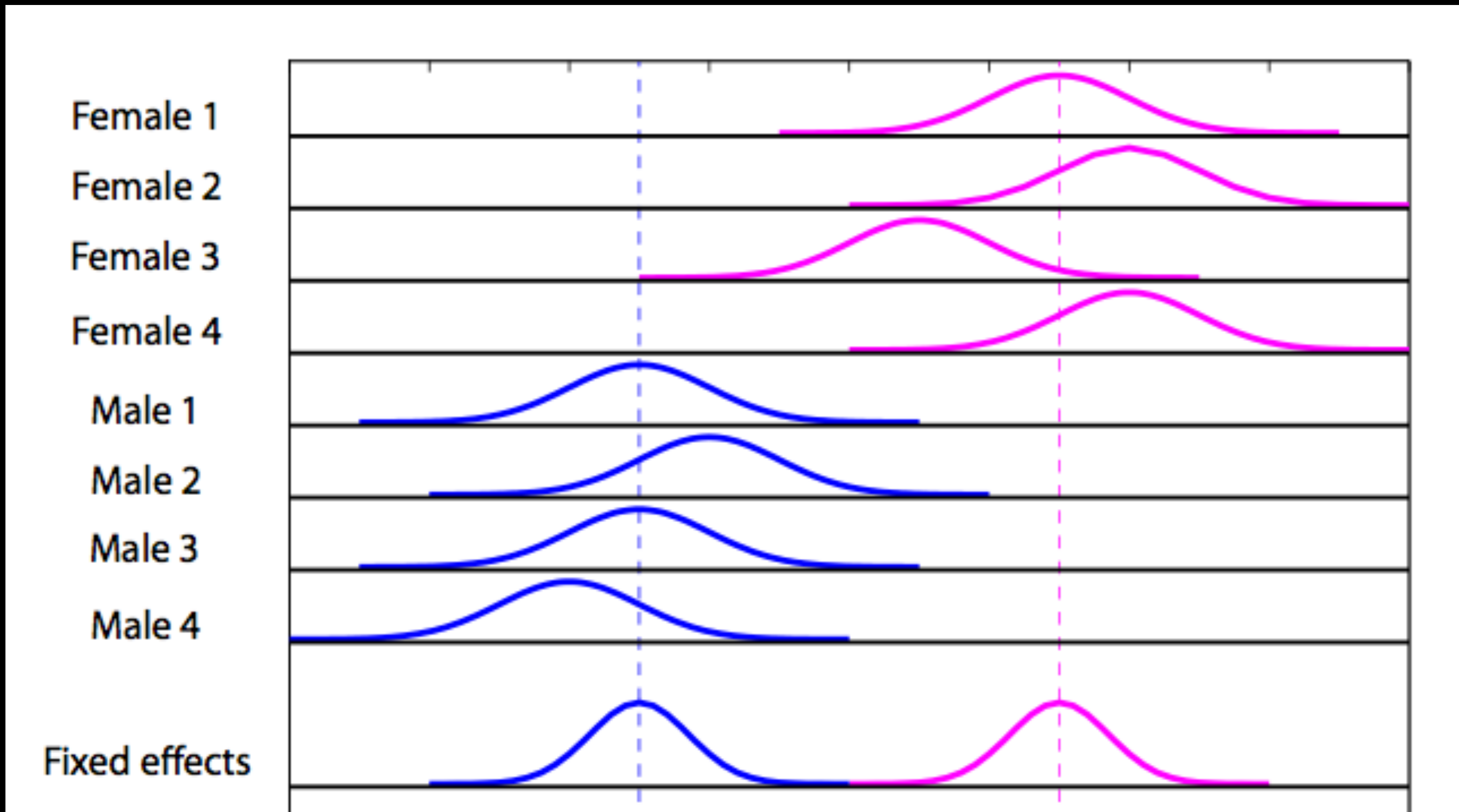
- Assume within-subject variance is 1



Each distribution has a variance of 1

Fixed effects analysis

- We're only interested in these exact 4 men and 4 women
  - $\sigma^2_{\mathrm{FFX}} = \frac{1}{4}\sigma^2_{\mathrm{W}} = 0.25$



$$\sigma^2_{\mathrm{FFX}} = \frac{1}{4}\sigma^2_{\mathrm{W}} = 0.25$$

Mixed effects

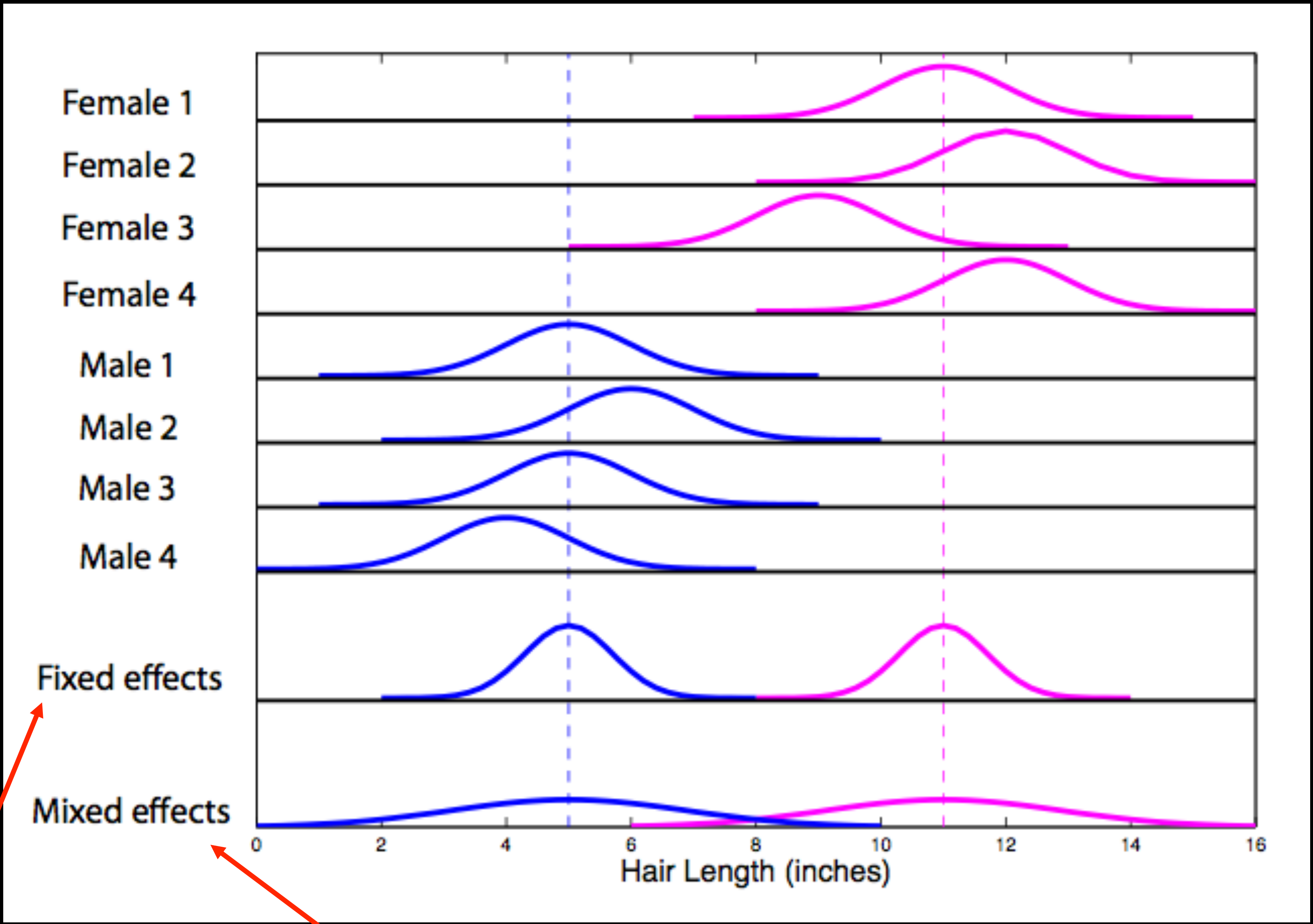
- Include both within *and* between subject variances
- Adding a between subject means subject is random

$$\sigma^2_{\mathrm{MFX}} = \sigma^2_{\mathrm{W}}/4 + \sigma^2_{\mathrm{B}}/4 = 1/4 + 49/4 = 12.5$$

Mixed effects

- Include both within *and* between subject variances
- Adding a between subject means subject is random
  - Anything with a variance is random!

$$\sigma^2_{\mathrm{MFX}} = \sigma^2_{\mathrm{W}}/4 + \sigma^2_{\mathrm{B}}/4 = 1/4 + 49/4 = 12.5$$



$$\sigma^2_{\mathrm{FFX}} = \frac{1}{4}\sigma^2_{\mathrm{W}} = 0.25$$

$$\sigma^2_{\mathrm{MFX}} = \sigma^2_{\mathrm{W}}/4 + \sigma^2_{\mathrm{B}}/4 = 1/4 + 49/4 = 12.5$$

Multiple hairs per subject

- Fixed effects variance
  - $\sigma^2_{\mathrm{FFX}} = \frac{1}{4}\sigma^2_{\mathrm{W}}/25 = 0.01$

- Mixed effects variance
  - $\sigma^2_{\mathrm{MFX}} = \frac{1}{4}\sigma^2_{\mathrm{W}}/25 + \frac{1}{4}\sigma^2_{\mathrm{B}} = 12.26$

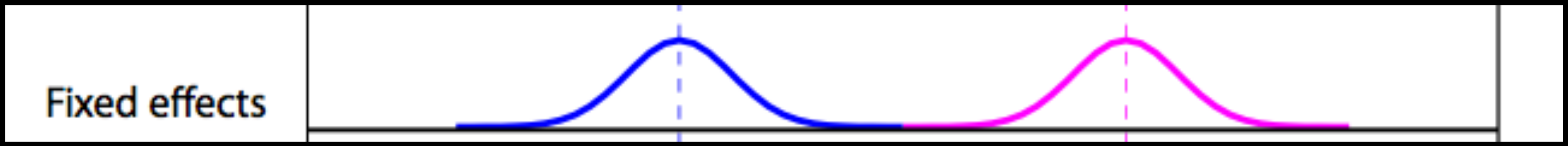
Multiple hairs per subject

- Fixed effects variance
  - $\sigma^2_{\text{FFX}} = \frac{1}{4}\sigma^2_{\text{W}}/25 = 0.01$
- Mixed effects variance
  - $\sigma^2_{\text{MFX}} = \frac{1}{4}\sigma^2_{\text{W}}/25 + \frac{1}{4}\sigma^2_{\text{B}} = 12.26$

Between subject variance typically dominates

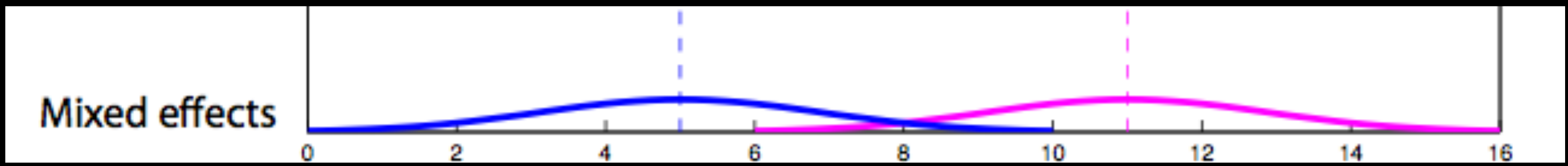
Wrong model leads to wrong conclusion

- Scenario 1: Fixed effects model
  - Significant difference in hair length
  - Result only applies to these 8 subjects



Wrong model leads to wrong conclusion

- Scenario 2: Mixed effects model
  - Cannot conclude there is a difference in hair length



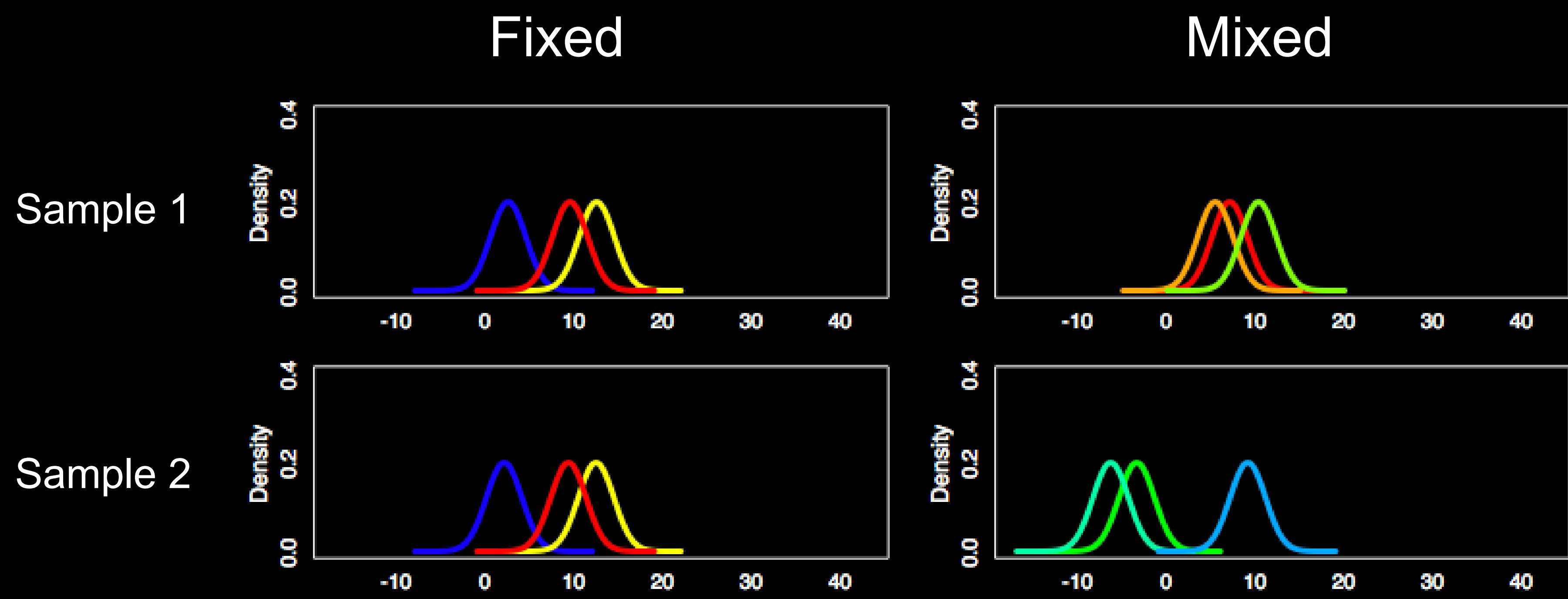
Mixed Model Comments

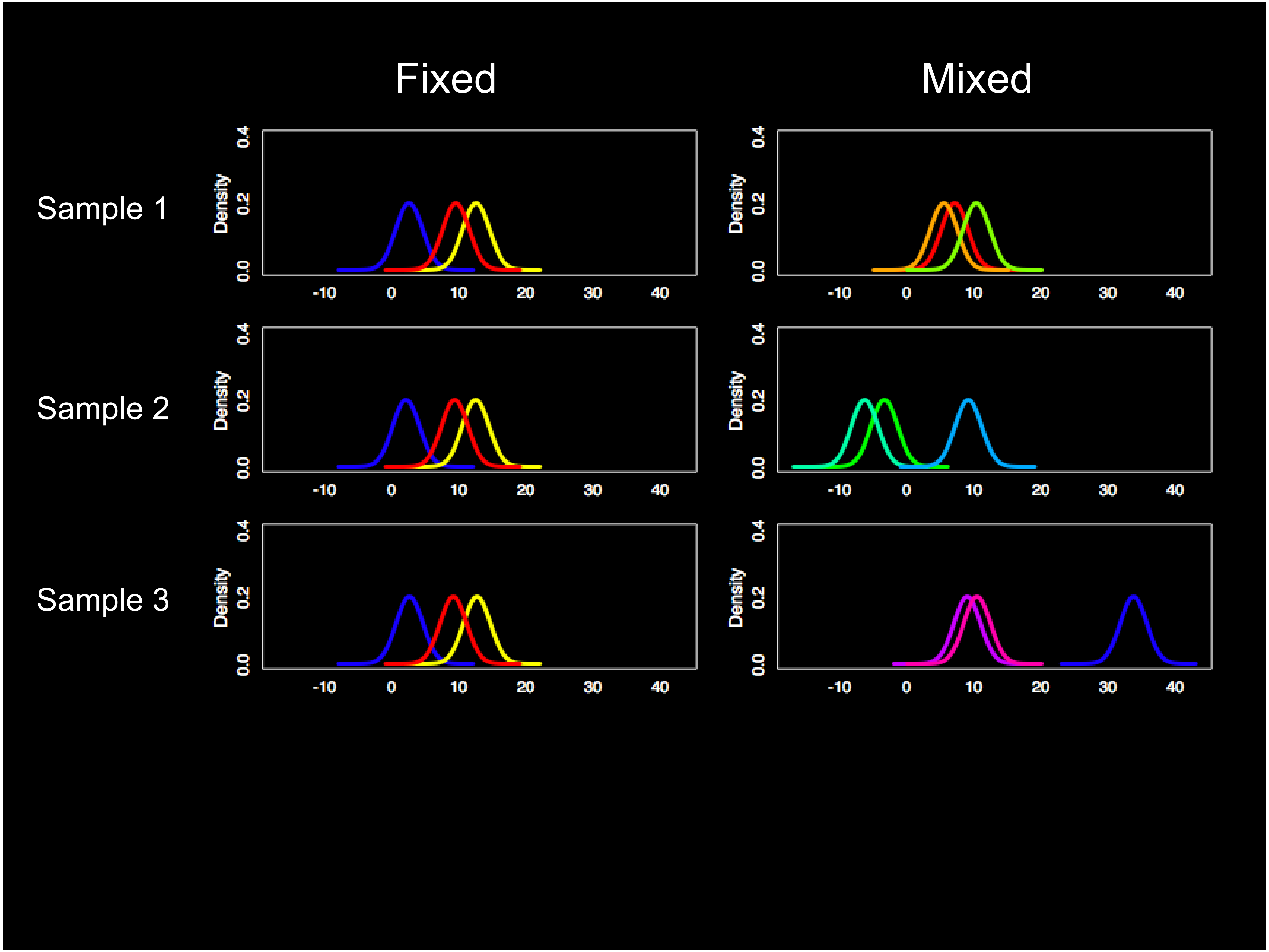
- If you fail to include a random effect when there is one
  - Results only apply to that data sample
  - P-values are smaller than mixed model p-values

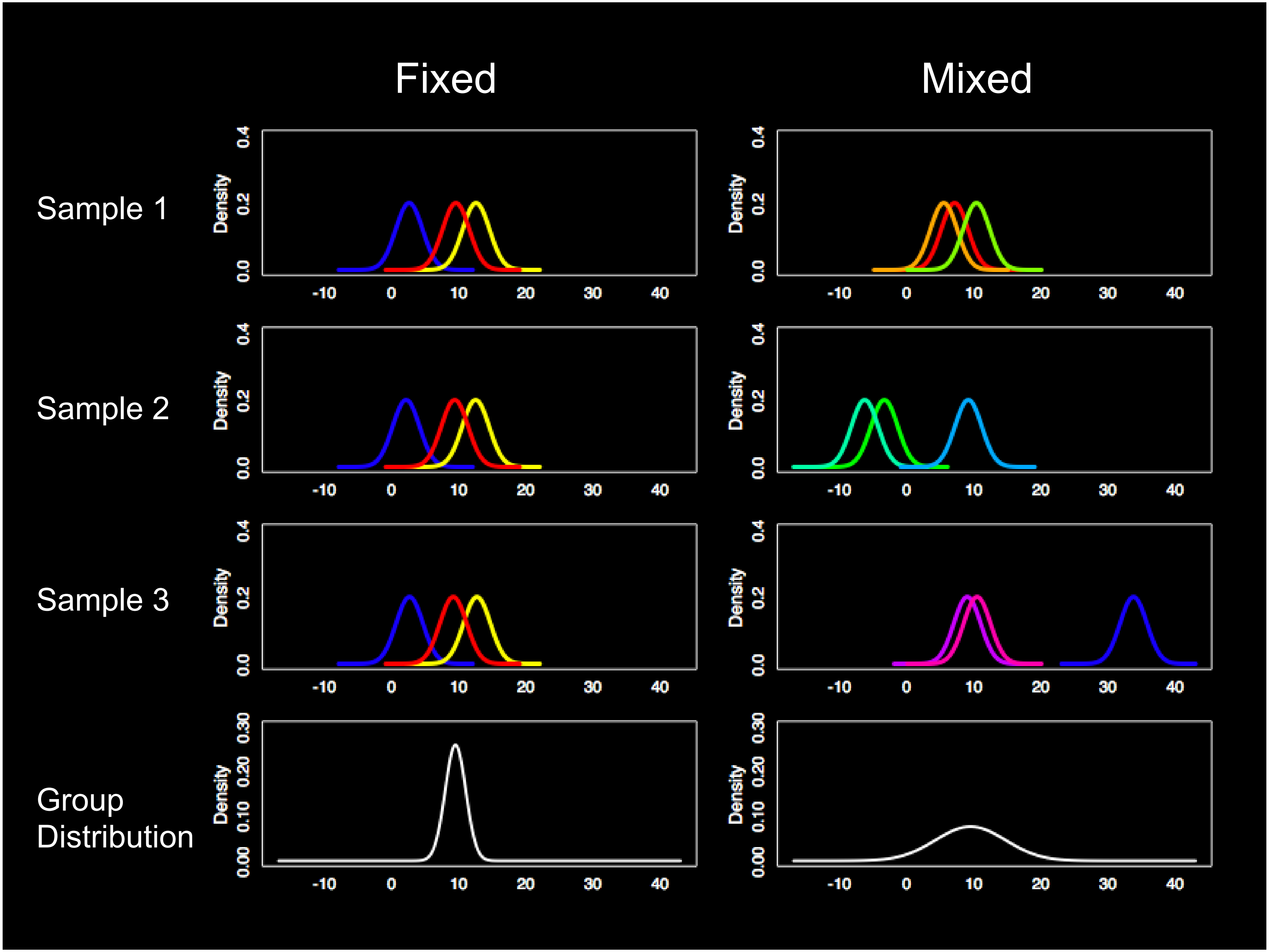
Fixed

Mixed

Sample 1


Fixed
Mixed
Sample 1
Sample 2
Density
0.0
0.2
0.4
-10
0
10
20
30
40


Fixed
Mixed
Sample 1
Sample 2
Sample 3
Density
0.0
0.2
0.4
-10
0
10
20
30
40



Fixed
Mixed
Sample 1
Sample 2
Sample 3
Group
Distribution
Density
0.0
0.2
0.4
0.00
0.10
0.20
0.30
-10
0
10
20
30
40

Take away

- What happens if you ignore the random subject effect ?

- What happens to the overall variance when you include a between-subject variance?

Take away

- What has a bigger impact in reducing variance?
  - Adding more hairs per subject?
  - Adding more subjects?

Where we're going

- Fixed vs Mixed modeling
- 2 stage summary statistics approach to mixed model (using non-fmri example)
- 2 stage summary statistics approach with fMRI data
- Software differences (FSL and SPM)

Moving forward with hair example

- Let’s focus on the hair distribution for 3 females with 20 hairs per person

Fixed effects model:

modeling the mean of 3 females, 20 hairs

Fixed effect

Residual error

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \\ Y_{2,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \epsilon_{1,1} \\ \epsilon_{1,2} \\ \epsilon_{1,3} \\ \epsilon_{2,1} \\ \vdots \\ \epsilon_{N,3} \end{pmatrix}, \qquad \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

Mixed Effects Model

Stage 1 $$Y = X\beta + \epsilon$$

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \\ Y_{2,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \epsilon_{1,1} \\ \epsilon_{1,2} \\ \epsilon_{1,3} \\ \epsilon_{2,1} \\ \vdots \\ \epsilon_{N,3} \end{pmatrix}, \ \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

Mixed Effects Model

Stage 1 $$Y = X\beta + \epsilon$$

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \\ Y_{2,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \epsilon_{1,1} \\ \epsilon_{1,2} \\ \epsilon_{1,3} \\ \epsilon_{2,1} \\ \vdots \\ \epsilon_{N,3} \end{pmatrix}, \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

Stage 2 $$\beta = X_g \beta_g + \eta$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}, \quad \eta_i \sim N(0, \sigma^2_{btwn})$$

Mixed Effects Model

Stage 1 $$Y = X\beta + \epsilon$$

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \\ Y_{2,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \epsilon_{1,1} \\ \epsilon_{1,2} \\ \epsilon_{1,3} \\ \epsilon_{2,1} \\ \vdots \\ \epsilon_{N,3} \end{pmatrix}, \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

Stage 2 $$\beta = X_g\beta_g + \eta$$

Random effect

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}, \quad \eta_i \sim N(0, \sigma^2_{btwn})$$

Mixed Effects Model

Stage 1 $$Y = X\beta + \epsilon$$

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \\ Y_{2,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \epsilon_{1,1} \\ \epsilon_{1,2} \\ \epsilon_{1,3} \\ \epsilon_{2,1} \\ \vdots \\ \epsilon_{N,3} \end{pmatrix}, \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

Stage 2 $$\beta = X_g\beta_g + \eta$$

Random effect

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}, \quad \eta_i \sim N(0, \sigma^2_{btwn})$$

Mixed Effects Model: All-In-One

$$Y = XX_g\beta_g + \quad X\eta \quad + \quad \epsilon$$

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \\ Y_{2,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \beta_g + \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} + \begin{pmatrix} \epsilon_{1,1} \\ \epsilon_{1,2} \\ \epsilon_{1,3} \\ \epsilon_{2,1} \\ \vdots \\ \epsilon_{N,3} \end{pmatrix}}_{\text{Variance Terms}}$$

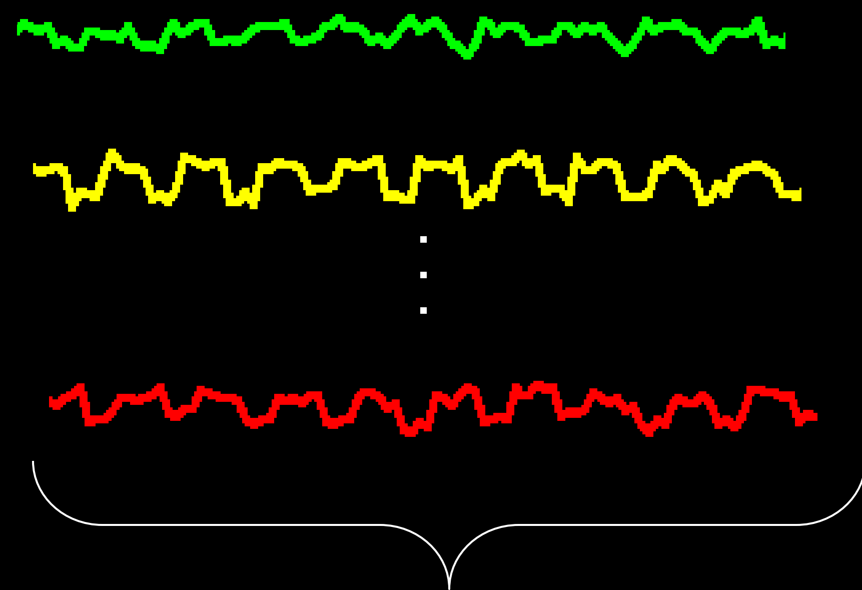
How does this relate to fMRI?

Subject 1

Subject 2

⋮

Subject N

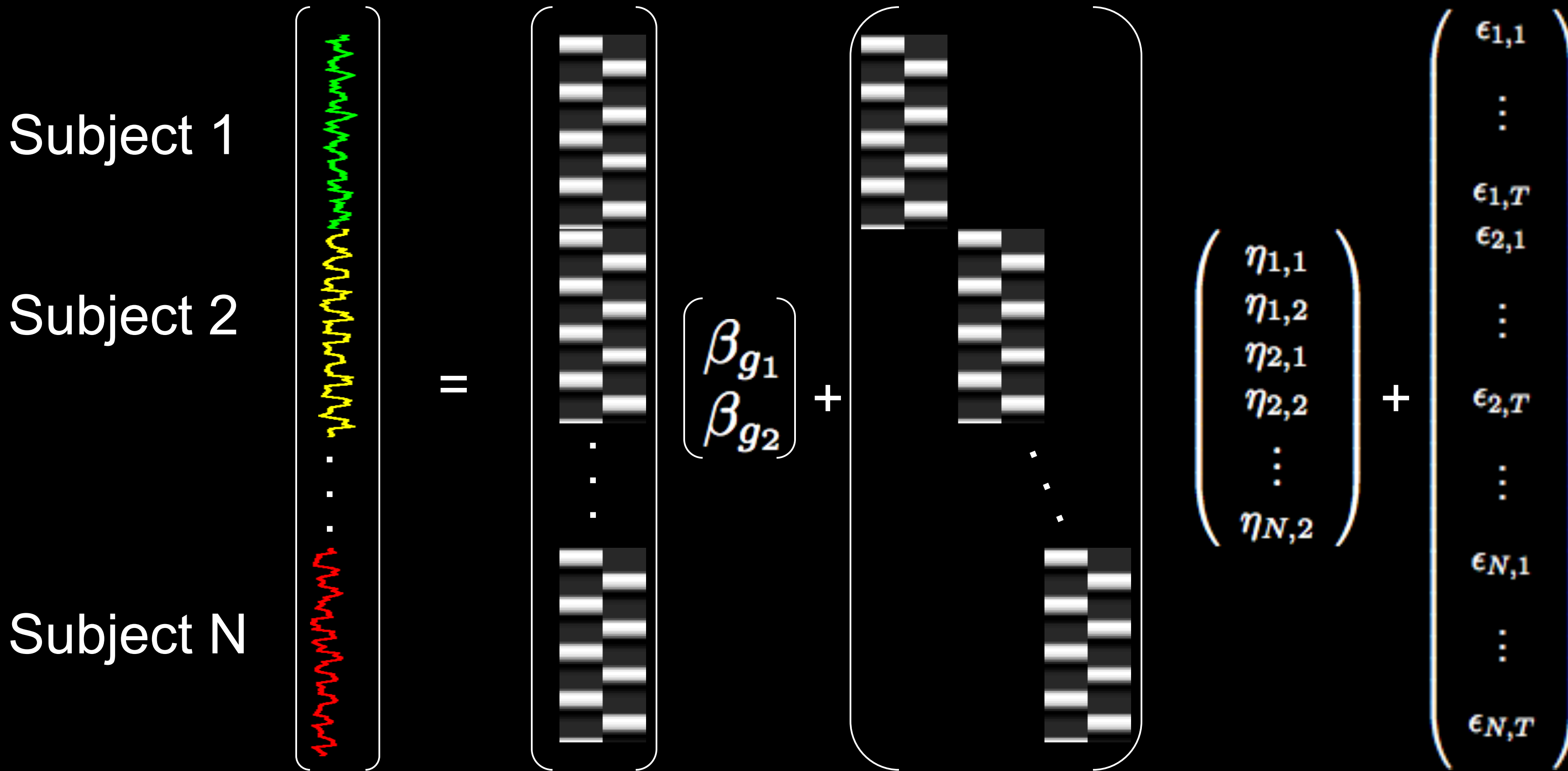


Each time series is a collection of data grouped by subject

A random subject effect is necessary to apply inference to total population

Mixed Model for fMRI Data

- fMRI data are more complicated than the hair length example
  - Not typically estimating an intercept
  - Time series are temporally autocorrelated
  - Time series can be quite long
- Let's take a look at the model!
  - A study with 2 stimuli of interest



$$\mathrm{Var}(\eta_{i,1}) = \sigma^2_{btwn_1}$$
$$\mathrm{Var}(\eta_{i,2}) = \sigma^2_{btwn_2}$$

$$\mathrm{Cov}\begin{pmatrix} \epsilon_{i,1} \\ \epsilon_{i,2} \\ \vdots \\ \epsilon_{i,T} \end{pmatrix} = \sigma^2_{win_i} V_i$$

Yuck!

- Computationally intensive
  - Large matrices that need to be inverted
- What if we add another subject?
  - Must estimate *whole* model for all subjects

Recall the two stages

Stage 1

$$Y \quad = \quad X\beta \quad + \quad \epsilon$$

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \\ Y_{2,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \epsilon_{1,1} \\ \epsilon_{1,2} \\ \epsilon_{1,3} \\ \epsilon_{2,1} \\ \vdots \\ \epsilon_{N,3} \end{pmatrix}, \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

Stage 2

$$\beta \quad = \quad X_g\beta_g \quad + \quad \eta$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}, \quad \eta_i \sim N(0, \sigma^2_{btwn})$$

Two-Stage Summary Statistics

Stage 1

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_1 + \begin{pmatrix} \epsilon_{1,1} \\ \vdots \\ \epsilon_{1,20} \end{pmatrix}$$

$$\begin{pmatrix} Y_{2,1} \\ \vdots \\ Y_{2,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_2 + \begin{pmatrix} \epsilon_{2,1} \\ \vdots \\ \epsilon_{2,20} \end{pmatrix} \qquad \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

$$\begin{pmatrix} Y_{3,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_3 + \begin{pmatrix} \epsilon_{3,1} \\ \vdots \\ \epsilon_{3,20} \end{pmatrix}$$

Two-Stage Summary Statistics

Stage 1

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_1 + \begin{pmatrix} \epsilon_{1,1} \\ \vdots \\ \epsilon_{1,20} \end{pmatrix}$$

$$\begin{pmatrix} Y_{2,1} \\ \vdots \\ Y_{2,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_2 + \begin{pmatrix} \epsilon_{2,1} \\ \vdots \\ \epsilon_{2,20} \end{pmatrix} \qquad \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

$$\begin{pmatrix} Y_{3,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_3 + \begin{pmatrix} \epsilon_{3,1} \\ \vdots \\ \epsilon_{3,20} \end{pmatrix}$$

Stage 2

Use first stage estimates

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \eta^*_1 \\ \eta^*_2 \\ \eta^*_3 \end{pmatrix}, \quad \mathrm{Var}(\eta^*_i) = \frac{\sigma^2_{win}}{W} + \sigma^2_{btwn}$$

Two-Stage Summary Statistics

Stage 1

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_1 + \begin{pmatrix} \epsilon_{1,1} \\ \vdots \\ \epsilon_{1,20} \end{pmatrix}$$

$$\begin{pmatrix} Y_{2,1} \\ \vdots \\ Y_{2,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_2 + \begin{pmatrix} \epsilon_{2,1} \\ \vdots \\ \epsilon_{2,20} \end{pmatrix} \qquad \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

$$\begin{pmatrix} Y_{3,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_3 + \begin{pmatrix} \epsilon_{3,1} \\ \vdots \\ \epsilon_{3,20} \end{pmatrix}$$

Stage 2

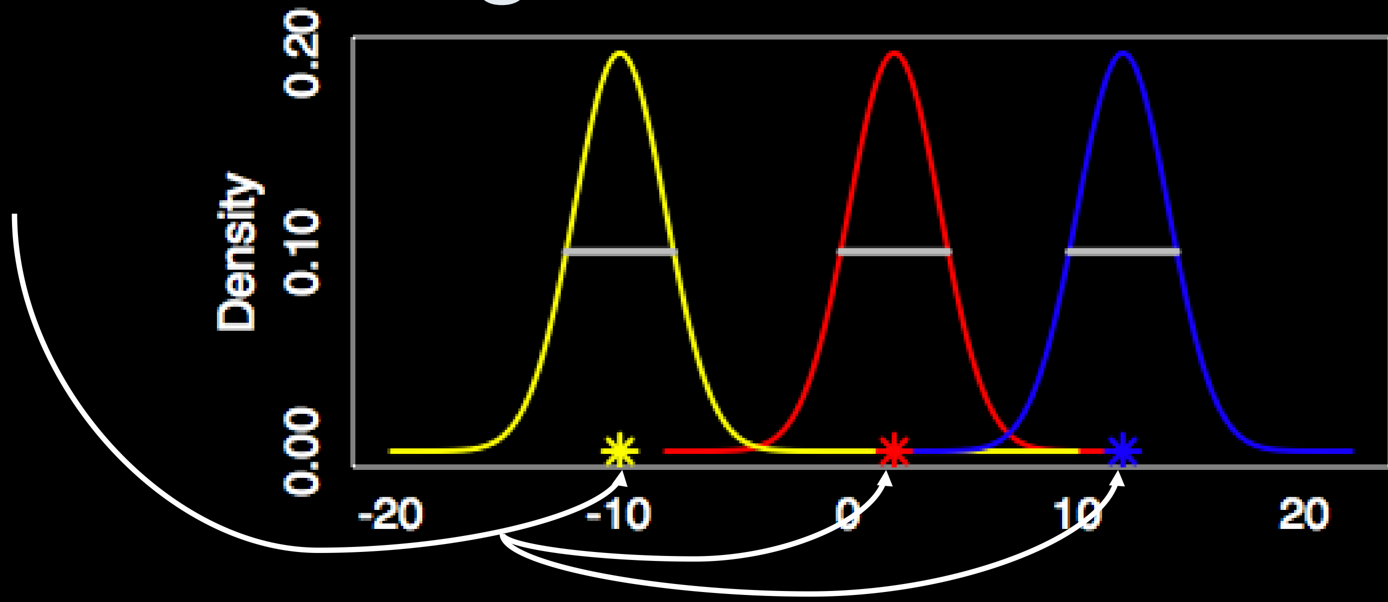
within

between

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \eta_1^* \\ \eta_2^* \\ \eta_3^* \end{pmatrix}, \quad \mathrm{Var}(\eta_i^*) = \frac{\sigma^2_{win}}{W} + \sigma^2_{btwn}$$

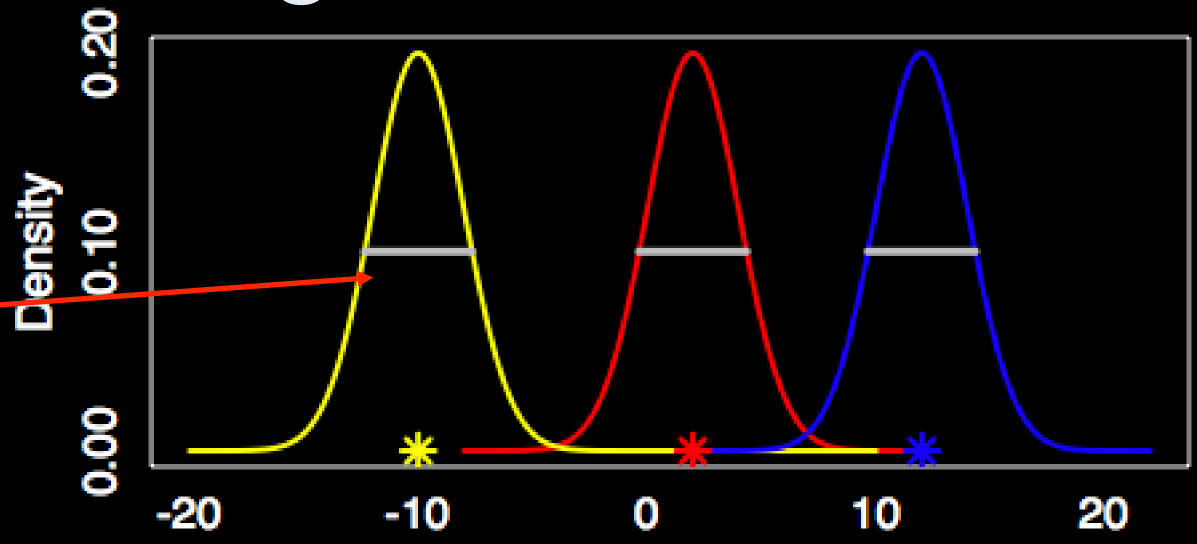
Two-Stage Model

- Stage 1
  - Means



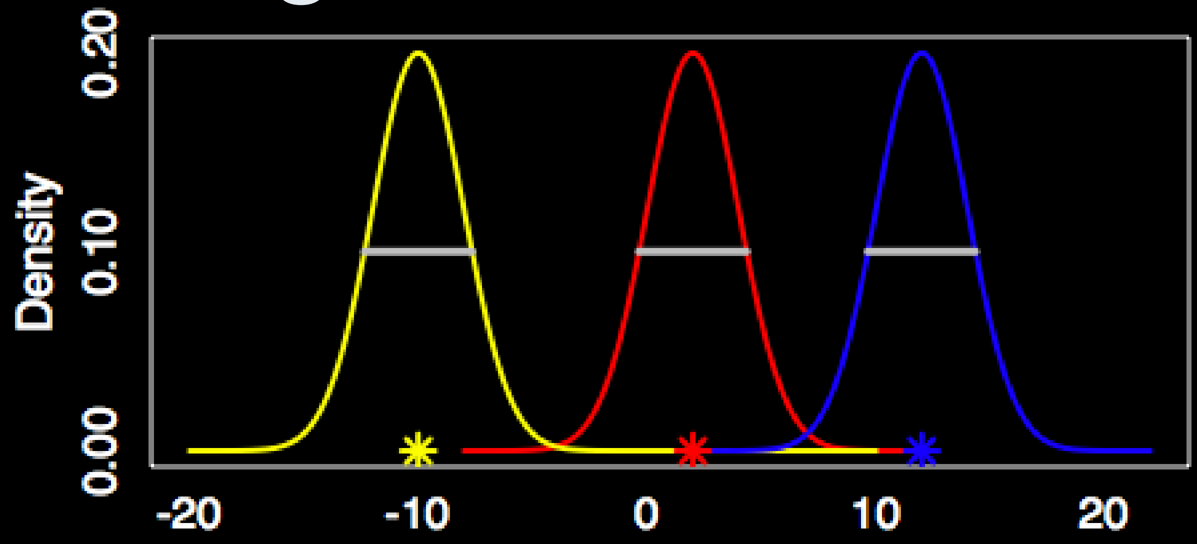
Two-Stage Model

- Stage 1
  - Means
  - σ^2_{win} (same across subjects here)

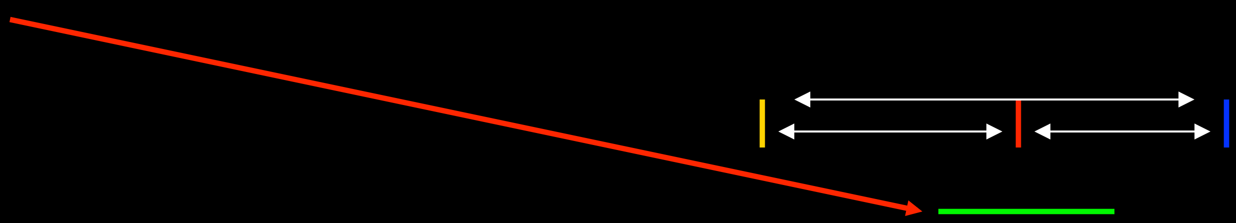


Two-Stage Model

- Stage 1
  - Means
  - σ^2_{win} (same across subjects here)

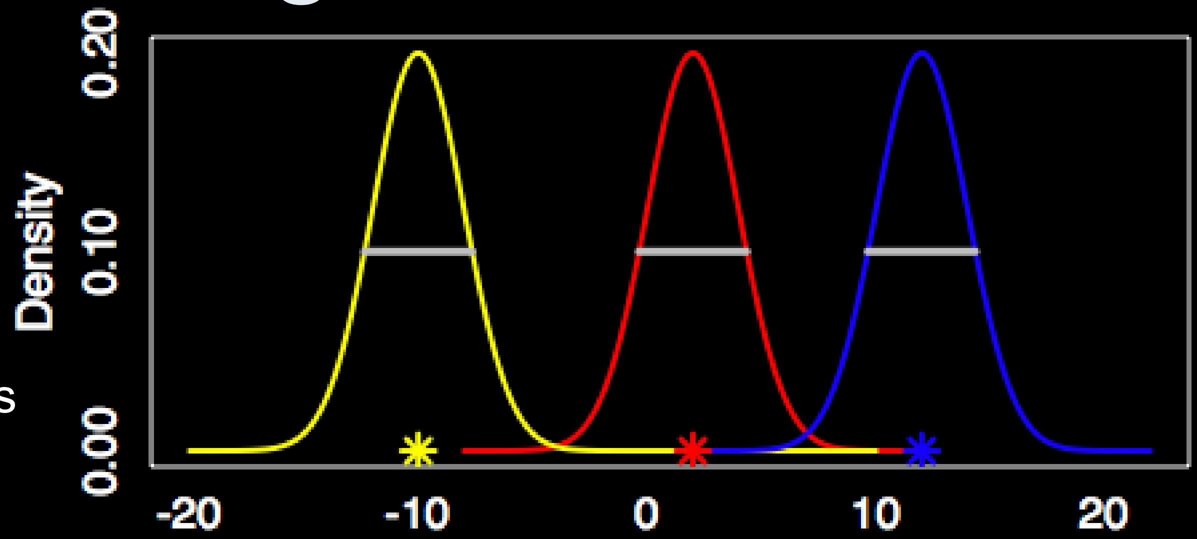


- Stage 2
  - σ^2_{btwn}

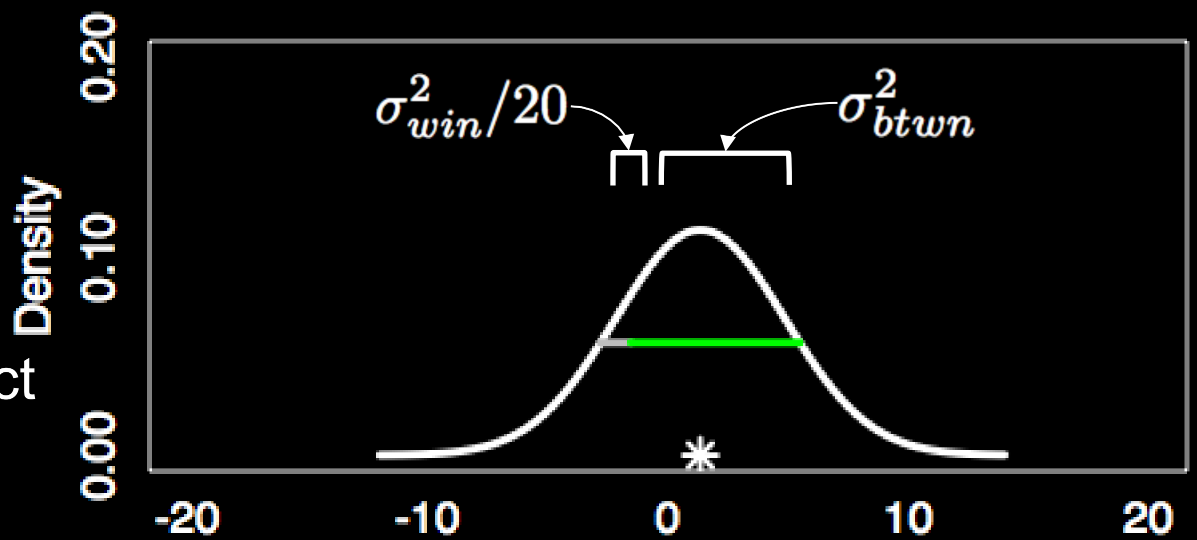


Two-Stage Model

- Stage 1
  - Means
  - σ^2_{win} (same across subjects here)

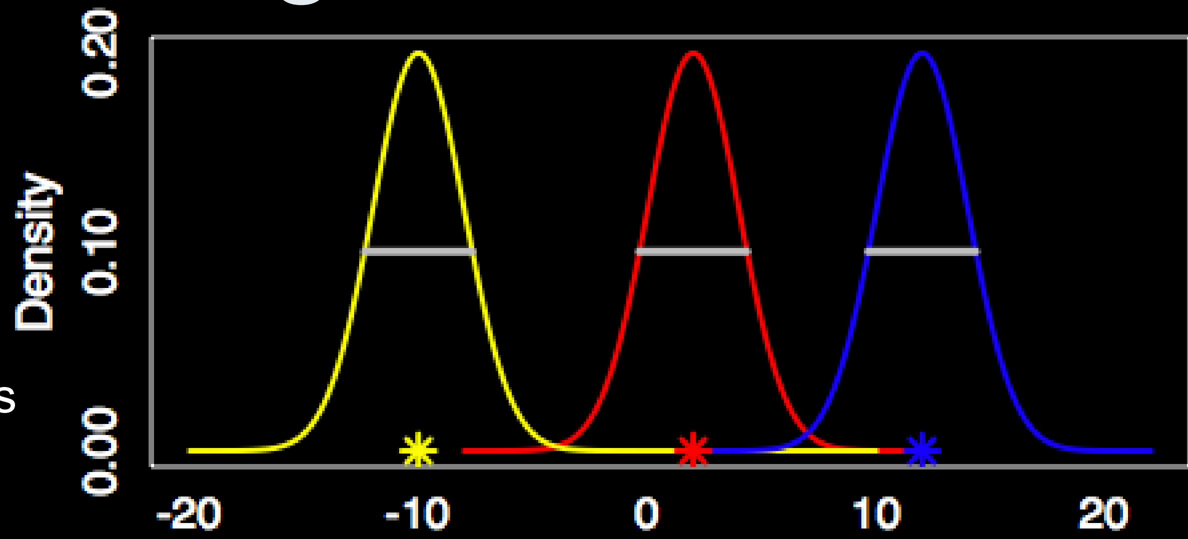


- Stage 2
  - σ^2_{btwn}
  - $\sigma^2_{mix} = \frac{\sigma^2_{win}}{20} + \sigma^2_{btwn}$
    - 20 hairs/subject

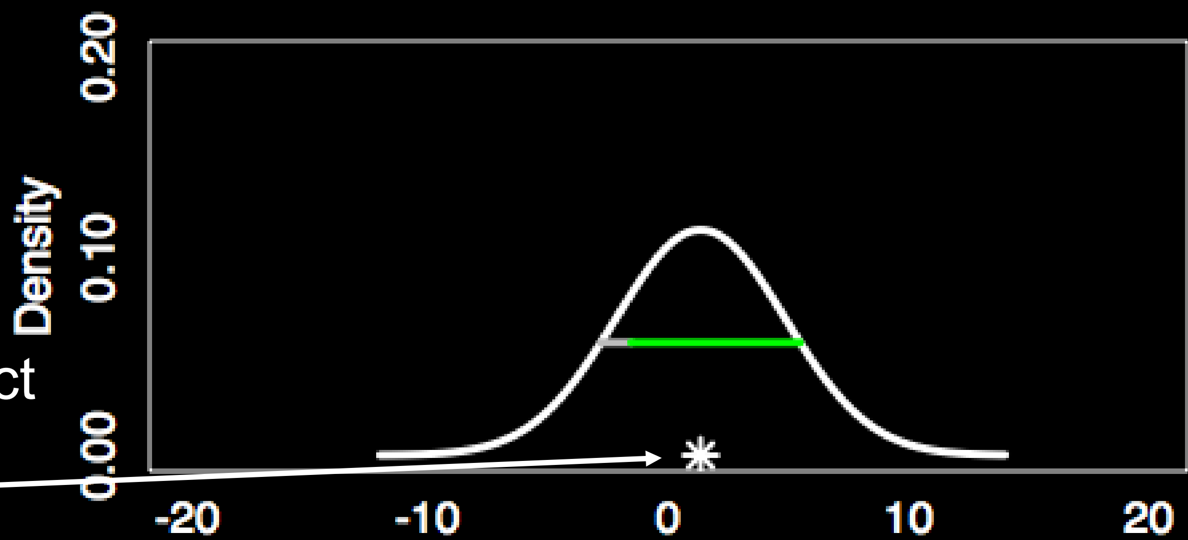


Two-Stage Model

- Stage 1
  - Means
  - σ^2_{win} (same across subjects here)



- Stage 2
  - σ^2_{btwn}
  - $\sigma^2_{mix} = \frac{\sigma^2_{win}}{20} + \sigma^2_{btwn}$
    - 20 hairs/subject
- Pop mean



Two-Stage Model

- $T = \dfrac{\sqrt{N}\hat{\beta}}{\sqrt{\sigma^2_{win}/W + \sigma^2_{btwn}}}$
  - N = # subjects
  - W = # measures within subject
- If new data are added, only run first stage for new data

Take away

- Do you understand the “cheat” we use for mixed models?

- Is the 2 stage summary statistics approach identical to a full mixed model?

Where we're going

- Fixed vs Mixed modeling
- 2 stage summary statistics approach to mixed model (using non-fmri example)
- 2 stage summary statistics approach with fMRI data
- Software differences (FSL and SPM)

Two-Stage Summary Statistics

Stage 1

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_1 + \begin{pmatrix} \epsilon_{1,1} \\ \vdots \\ \epsilon_{1,20} \end{pmatrix}$$

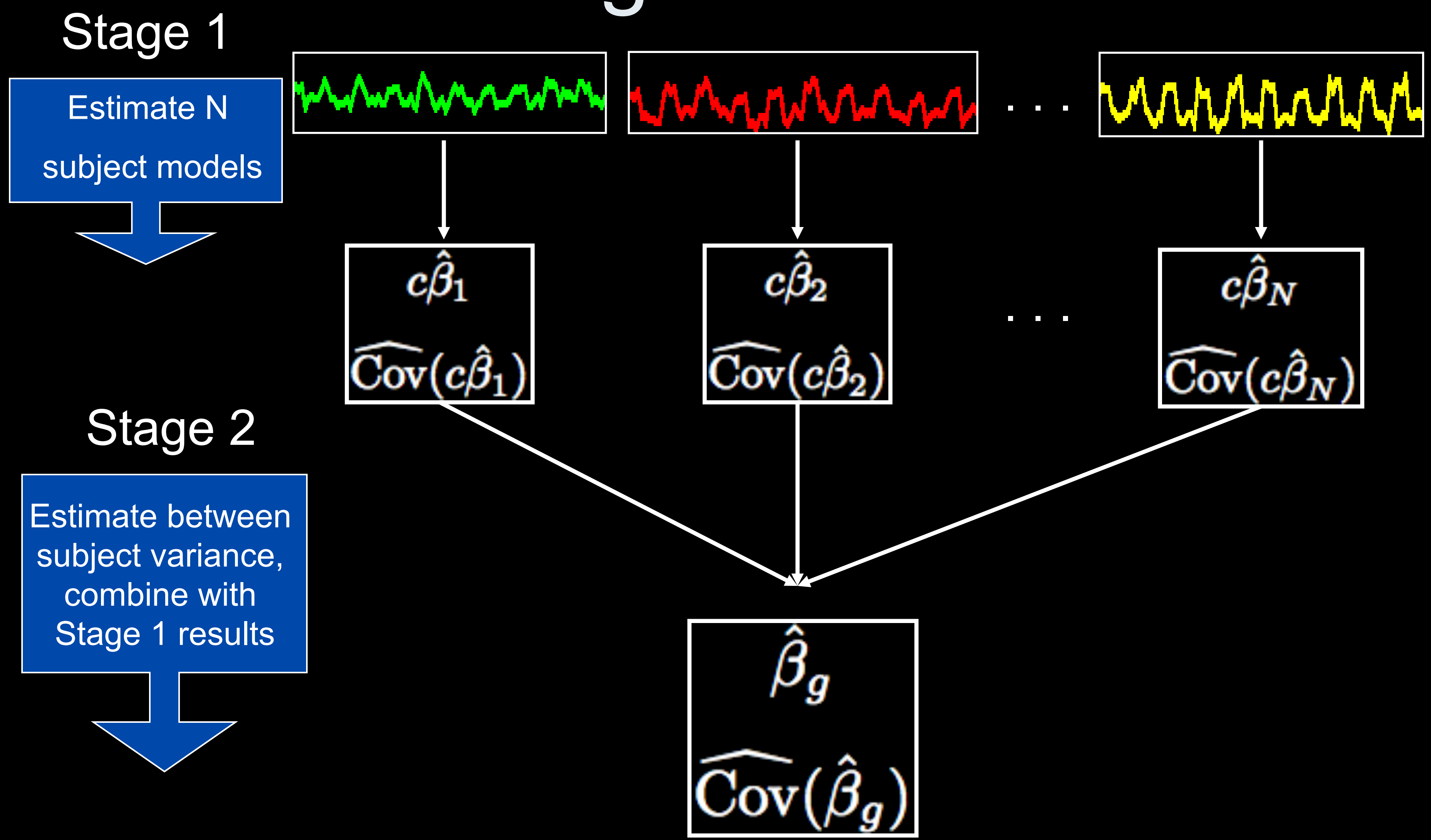
$$\begin{pmatrix} Y_{2,1} \\ \vdots \\ Y_{2,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_2 + \begin{pmatrix} \epsilon_{2,1} \\ \vdots \\ \epsilon_{2,20} \end{pmatrix} \qquad \epsilon_{i,j} \sim N(0, \sigma^2_{win})$$

$$\begin{pmatrix} Y_{3,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_3 + \begin{pmatrix} \epsilon_{3,1} \\ \vdots \\ \epsilon_{3,20} \end{pmatrix}$$

Stage 2

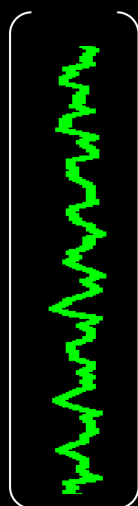
$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \eta^*_1 \\ \eta^*_2 \\ \eta^*_3 \end{pmatrix}, \quad \mathrm{Var}(\eta^*_i) = \frac{\sigma^2_{win}}{W} + \sigma^2_{btwn}$$

Two Stage Model fMRI

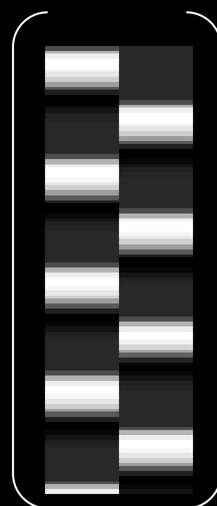


Stage 1: Subject Model

$$Y_k = X_k \beta_k + \epsilon_k$$

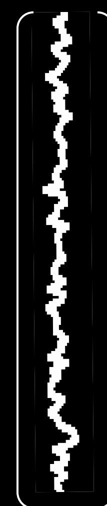


=



$$\begin{pmatrix} \beta_{k1} \\ \beta_{k2} \end{pmatrix}$$

+



$$\mathrm{Cov}(\epsilon_k) = \sigma_k^2 V_k$$

$$H_0 : \beta_{k1} - \beta_{k2} = 0$$

Stage 1: Estimation

- W_k such that $W_k V_k W_k' = I_T$

Stage 1: Estimation

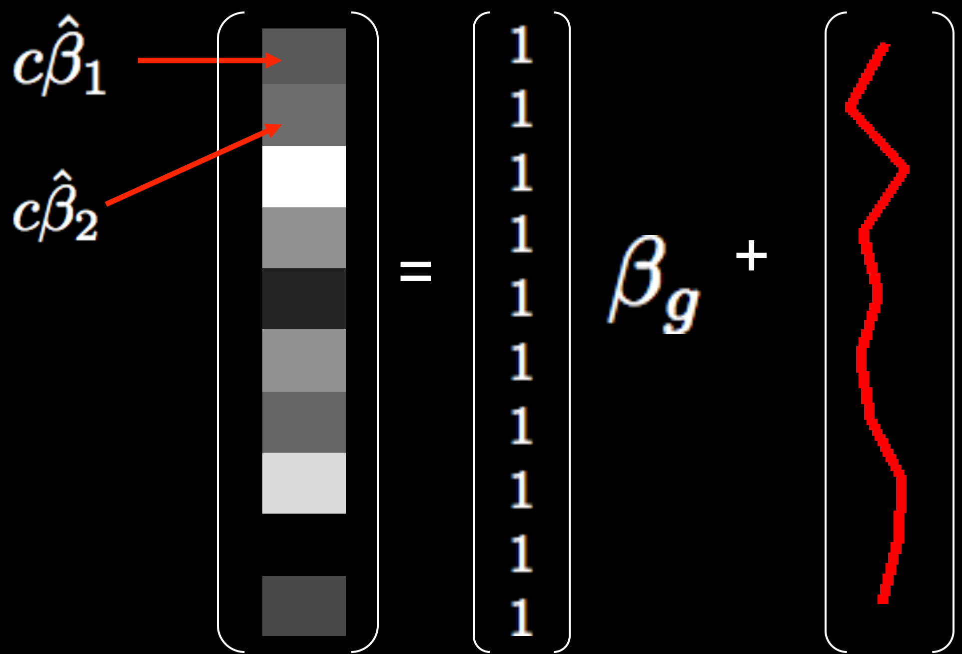
- W_k such that $W_k V_k W_k' = I_T$
- Whitened model
  - $W_k Y_k = W_k X_k \beta_k + W_k \epsilon_k$
  - $Y_k^* = X_k^* \beta_k + \epsilon_k^*$

Stage 1: Estimation

- W_k such that $W_k V_k W_k' = I_T$
- Whitened model
  - $W_k Y_k = W_k X_k \beta_k + W_k \epsilon_k$
  - $Y_k^* = X_k^* \beta_k + \epsilon_k^*$
- Use OLS on whitened model
  - $c\hat{\beta}_k = \left(X_k^{*\prime} X_k^* \right)^{-1} X_k^{*\prime} Y_k^*$
  - $\widehat{Cov}(c\hat{\beta}_k) = \hat{\sigma}_k^2 \left(X_k^{*\prime} X_k^* \right)^{-1}$

Stage 2: Group Model

$$\hat{\beta}_{cont} = X_g \beta_g + \epsilon_g$$



$$\mathrm{Cov}(\epsilon_g) = V_g = \begin{pmatrix} \sigma_1^2 c(X_1^{*'} X_1^*)^{-1} c' & & \\ & \ddots & \\ & & \sigma_N^2 c(X_N^{*'} X_N^*)^{-1} c' \end{pmatrix} + \sigma_g^2 I_N$$

Stage 2: Estimation

- W_g such that $W_g V_g W_g' = I_N$

Stage 2: Estimation

- W_g such that $W_g V_g W_g' = I_N$
- $W_g \hat{\beta}_{cont} = W_g X_g \beta_g + W_g \epsilon_g$
  $\hat{\beta}^*_{cont} = X^*_g \beta_g + \epsilon^*_g$

Stage 2: Estimation

- W_g such that $W_g V_g W_g' = I_N$
- $W_g \hat{\beta}_{cont} = W_g X_g \beta_g + W_g \epsilon_g$
  $\hat{\beta}^*_{cont} = X^*_g \beta_g + \epsilon^*_g$
- $\hat{\beta}_g = \left(X^{*\prime}_g X^*_g \right)^{-1} X^{*\prime}_g \hat{\beta}^*_{cont}$
  $\widehat{Cov}(\hat{\beta}_g) = \left(X^{*\prime}_g X^*_g \right)^{-1}$

Stage 2: Estimation

- W_g such that $W_g V_g W_g' = I_N$
- $W_g \hat{\beta}_{cont} = W_g X_g \beta_g + W_g \epsilon_g$
  $\hat{\beta}^*_{cont} = X^*_g \beta_g + \epsilon^*_g$

- $\hat{\beta}_g = \left(X^{*\prime}_g X^*_g \right)^{-1} X^{*\prime}_g \hat{\beta}^*_{cont}$
  $\widehat{Cov}(\hat{\beta}_g) = \left(X^{*\prime}_g X^*_g \right)^{-1}$
- $T = \hat{\beta}_g / \sqrt{\widehat{\mathrm{Cov}}(\hat{\beta}_g)}$

Take away

- What are the benefits of breaking the full mixed effects model into 2 stages for fmri?
  - You should be able to think of at least 2 reasons

Take away

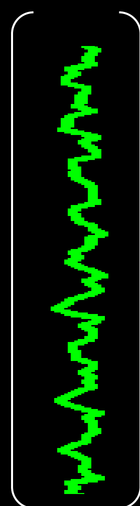
- When the model is estimated how is a subject with a high mfx variance treated differently than a subject with a low mfx variance?

Where we're going

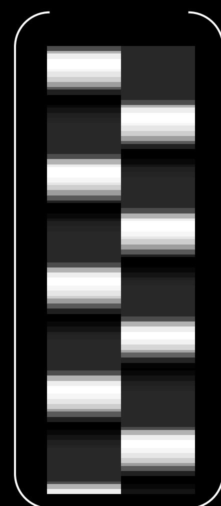
- Fixed vs Mixed modeling
- 2 stage summary statistics approach to mixed model (using non-fmri example)
- 2 stage summary statistics approach with fMRI data
- Software differences (FSL and SPM)

Stage 1: Subject Model

$$Y_k = X_k \beta_k + \epsilon_k$$

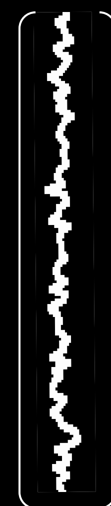


=



$$\begin{pmatrix} \beta_{k1} \\ \beta_{k2} \end{pmatrix}$$

+

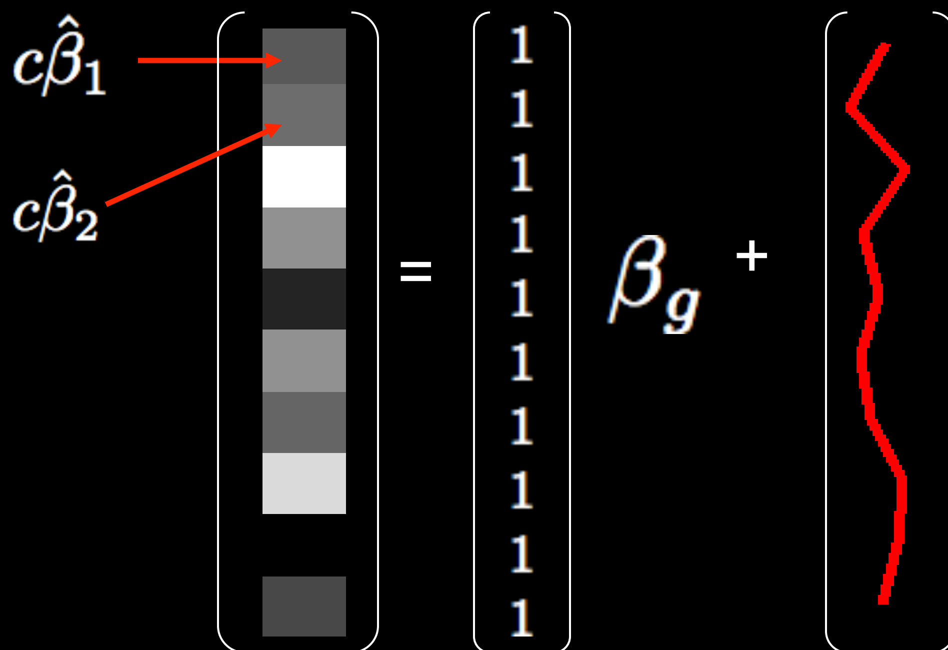


$$\mathrm{Cov}(\epsilon_k) = \sigma_k^2 V_k$$

$$H_0 : \beta_{k1} - \beta_{k2} = 0$$

Stage 2: Group Model

$$\hat{\beta}_{cont} = X_g \beta_g + \epsilon_g$$



$$\mathrm{Cov}(\epsilon_g) = V_g =$$

$$\begin{pmatrix} \sigma_1^2 c(X_1^{*'} X_1^*)^{-1} c' & & \\ & \ddots & \\ & & \sigma_N^2 c(X_N^{*'} X_N^*)^{-1} c' \end{pmatrix} + \sigma_g^2 I_N$$

How is the model estimated?

- Depends on software
  - SPM: Does not estimate σ_g^2
    - Due to a set of assumptions, estimation of is unnecessary
  - FSL: Bayesian approach to estimating σ_g^2

SPM2

- Does not estimate σ_g^2
  - Assumes homogeneous variance across subjects
  - Assumes first level design is same across subjects

$$\hat{\sigma}^2_{win_{all}} = \hat{\sigma}^2_1 c \left(X_1^{*'} X_1^* \right)^{-1} c' = \ldots = \hat{\sigma}^2_N c \left(X_N^{*'} X_N^* \right)^{-1} c'$$

$$V_g = \sigma^2_{win_{all}} I_N + \sigma^2_g I_N = \sigma^2_{g*} I_N$$

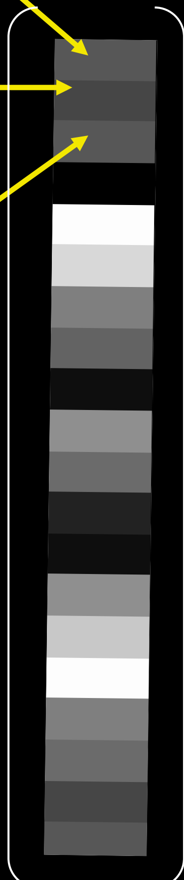
OLS can be used

SPM2 : Single contrast per subject

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \epsilon_g \qquad \epsilon_g \sim N(0, \sigma^2_{g*} I_N)$$

A one-sample T-test!

SPM2 : Multiple contrasts per subject



$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_{g_1} \\ \beta_{g_2} \end{pmatrix} + \epsilon_g$$

$$\epsilon_g \sim N(0, \sigma^2_{g*} \underline{\underline{V_{g^*}}})$$

Global correlation estimate

SPM2 : Summary

- Multiple contrasts per subject can enter second level
  – Contrasts can be correlated
  – T and F-tests are possible
- Special case
  – One contrast per subject…Reduces to T-test!

SPM2

- Pros
  - Model is easy to estimate
  - Model is easy to understand
  - Multiple contrasts can enter the group model and are *not* considered independent
- Cons
  - Global covariance estimate (same across voxels)
  - Assumes variance is homogeneous across subjects

FSL: FMRIB Software Library

- Bayesian approach to estimating model
- Inference is based on *posterior* distribution of the data
  - $P(\beta_g, \sigma_g^2, \nu_g | Y)$
  - Parameters of interest are treated as random

FSL : Second Level Estimation

- Flame 1: Maximum a posteriori (MAP) estimate of σ_g^2 found iteratively
  – Assumes degrees of freedom, $\nu_g = N - p$
- Flame 2: Slower MCMC method of estimation
    - Applied to voxels close to threshold in step 1
    - Fine tunes estimates of $\beta_g, \sigma_g^2, \nu_g$
- Details
  – Woolrich et al. NI (2004) 1732-47

FSL

- Pros
  - When single contrast is taken to the second level, ~equivalent to all-in-one model
  - Within-subject variances are carried to the second level
    - Heterogeneity across subjects is modeled
- Cons
  - Multiple contrasts in the group model are assumed to be independent

Which software?

- FSL best for heteroscedastic variances
  - Different number of trials per subject
- SPM best for multiple correlated contrasts at group level
- Other differences in first level modeling may sway users one way or another

Take away

- What’s the primary difference between FSL and SPM?

Questions?