

Modeling Group fMRI Data

Overview

- What is a mixed effects model
 - Fixed effects
 - Random effects
- 2-stage summary statistics approach
- How do different software packages work?
- Overview FSL modeling options

Overview

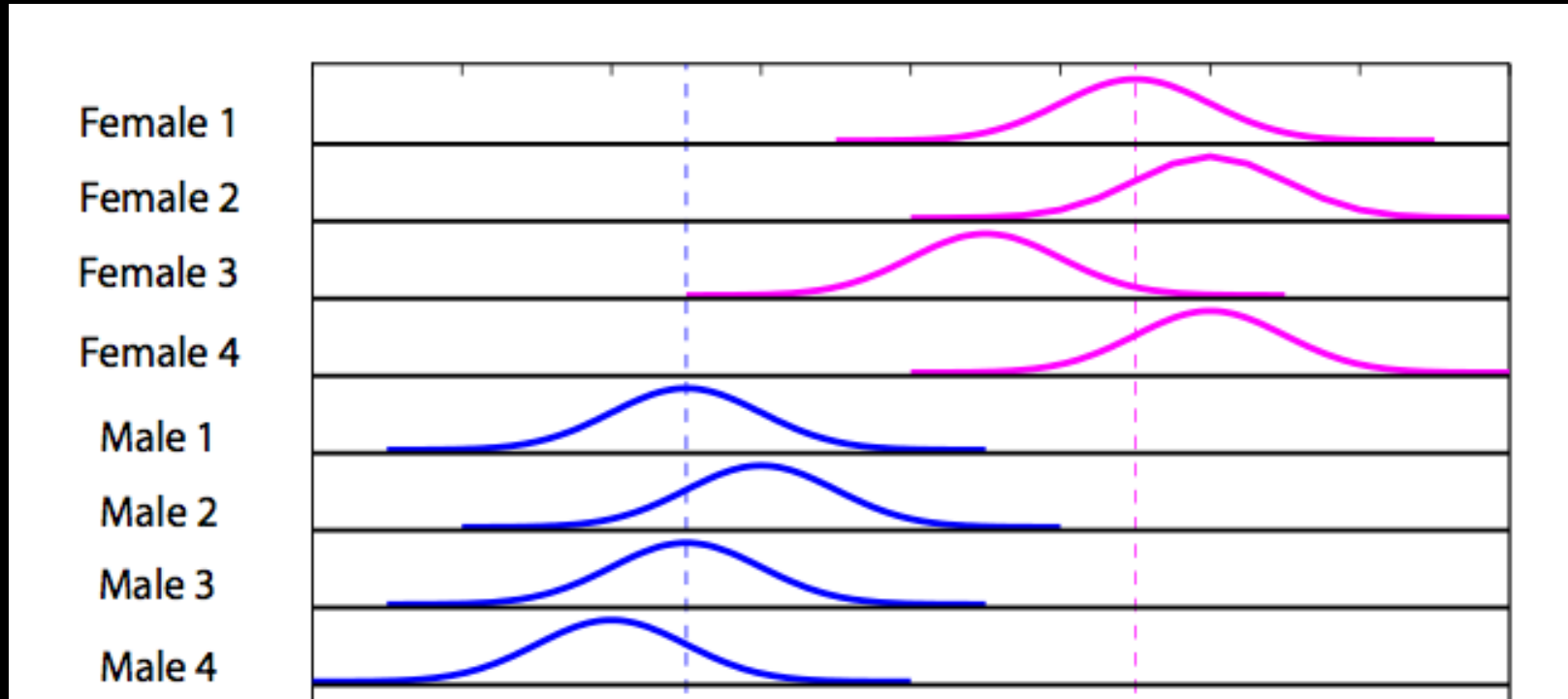
- What is a mixed effects model
 - Fixed effects
 - Random effects
- 2-stage summary statistics approach
- How do different software packages work?
- Overview FSL modeling options

Mixed Model Motivation

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

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

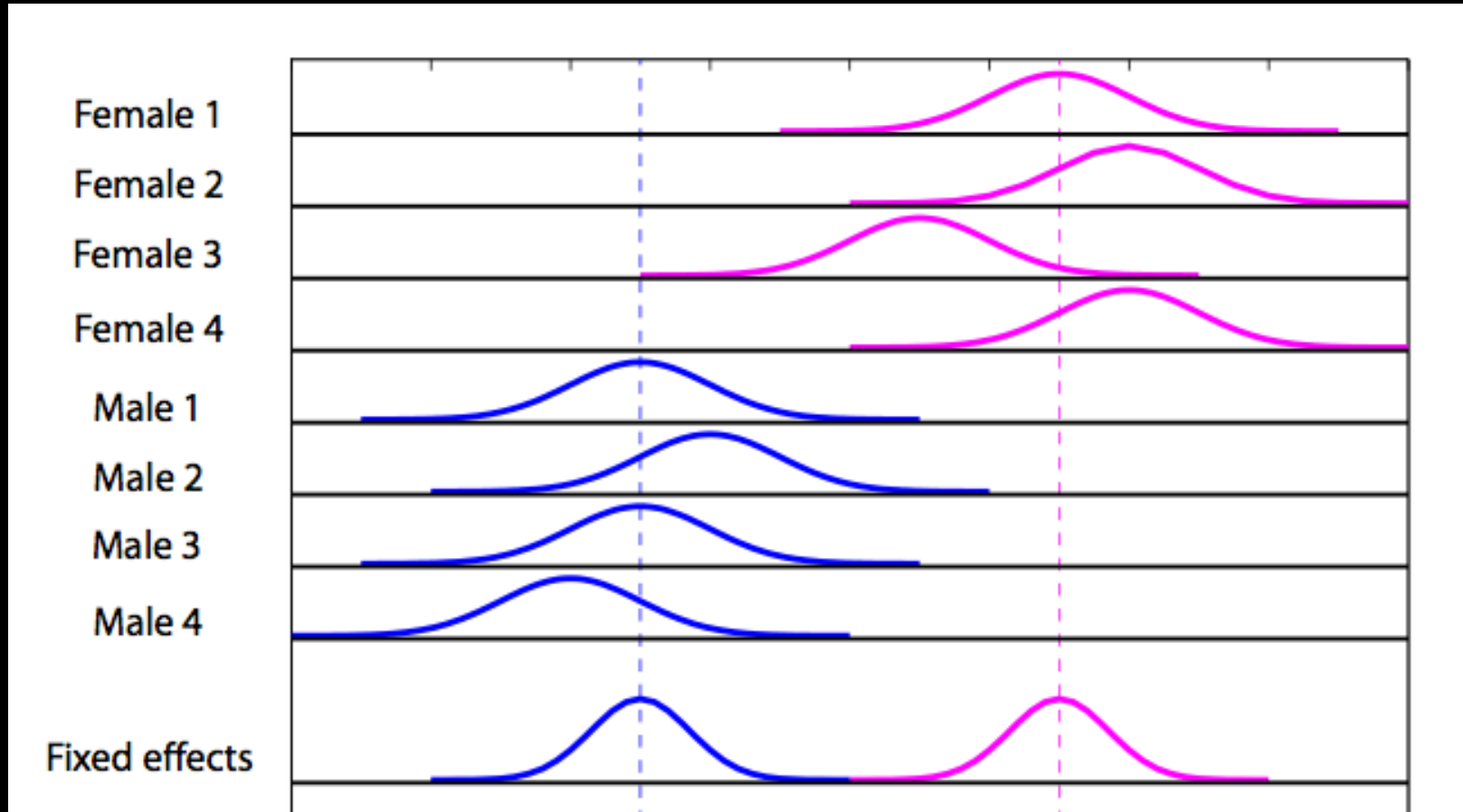


Each distribution has a variance of 1

Fixed effects analysis

- We're only interested in these exact 4 men and 4 women

- $\sigma_{\text{FFX}}^2 = \frac{1}{4}\sigma_{\text{W}}^2 = 0.25$

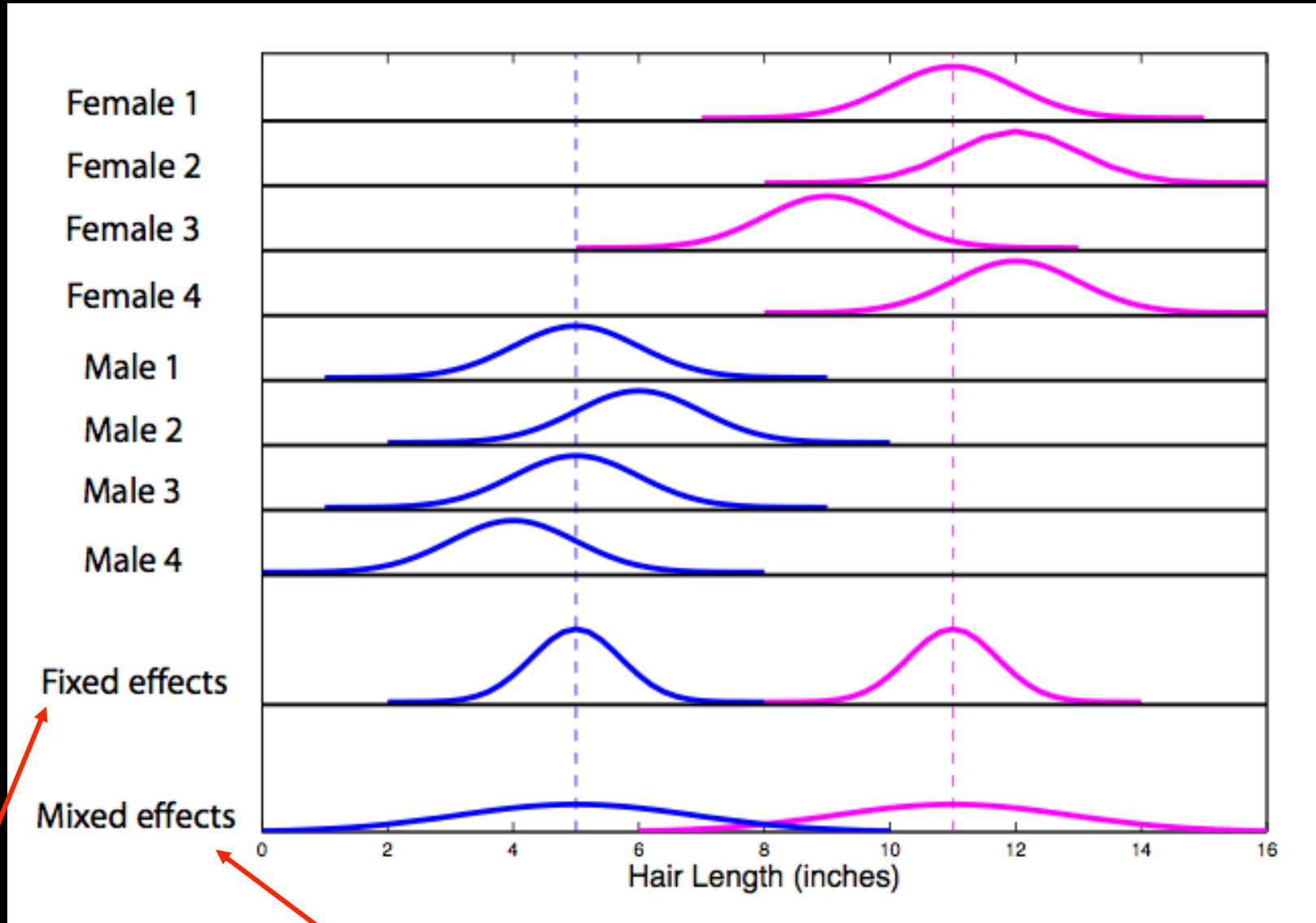


$$\sigma_{\text{FFX}}^2 = \frac{1}{4}\sigma_{\text{W}}^2 = 0.25$$

Mixed effects

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

$$\sigma_{\text{MFX}}^2 = \sigma_{\text{W}}^2/4 + \sigma_{\text{B}}^2/4 = 1/4 + 49/4 = 12.5$$



$$\sigma_{\text{FFX}}^2 = \frac{1}{4}\sigma_{\text{W}}^2 = 0.25$$

$$\sigma_{\text{MFX}}^2 = \sigma_{\text{W}}^2/4 + \sigma_{\text{B}}^2/4 = 1/4 + 49/4 = 12.5$$

Multiple hairs per subject

- Fixed effects variance

- $\sigma_{\text{FFX}}^2 = \frac{1}{4}\sigma_{\text{W}}^2/25 = 0.01$

- Mixed effects variance

- $\sigma_{\text{MFX}}^2 = \frac{1}{4}\sigma_{\text{W}}^2/25 + \frac{1}{4}\sigma_{\text{B}}^2 = 12.26$

Multiple hairs per subject

- Fixed effects variance

- $\sigma_{\text{FFX}}^2 = \frac{1}{4}\sigma_{\text{W}}^2/25 = 0.01$

- Mixed effects variance

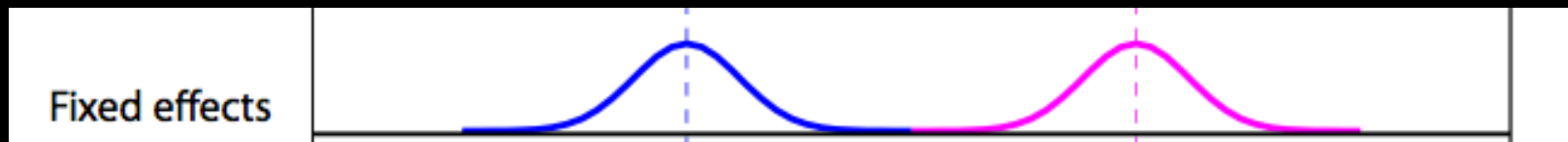
- $\sigma_{\text{MFX}}^2 = \frac{1}{4}\sigma_{\text{W}}^2/25 + \frac{1}{4}\sigma_{\text{B}}^2 = 12.26$

Between subject variance
typically dominates



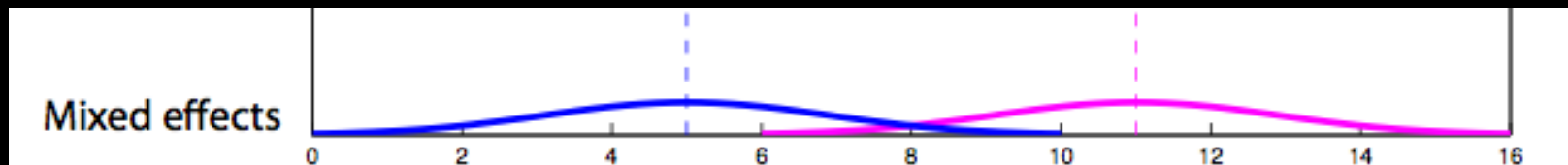
Wrong model leads to wrong conclusion

- 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

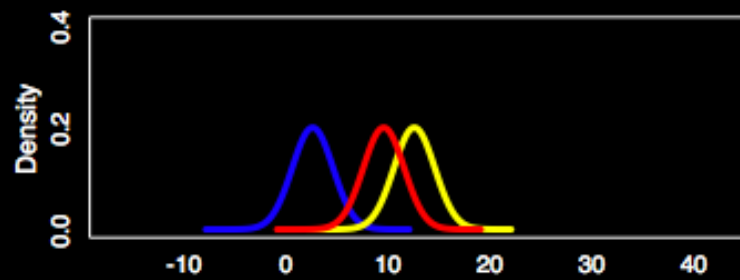


Mixed Model Comments

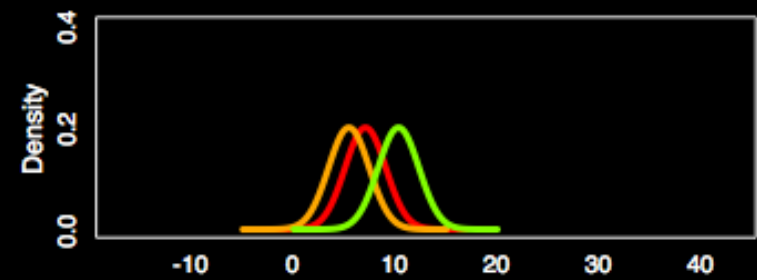
- If you fail to include a random effect when there is one
 - Results only apply to that data sample
 - P-values are smaller than mixed model p-values

Sample 1

Fixed



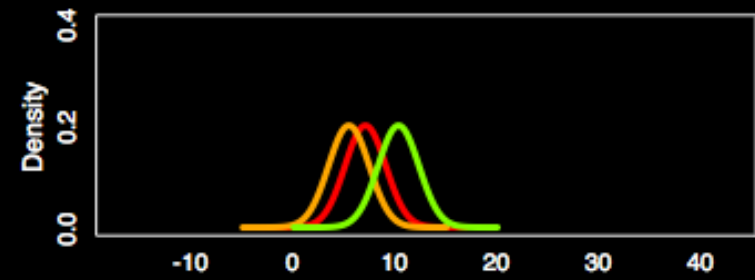
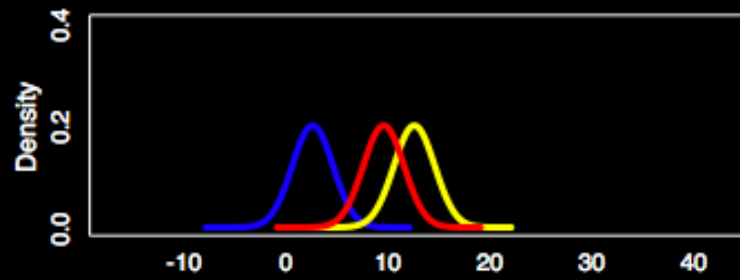
Mixed



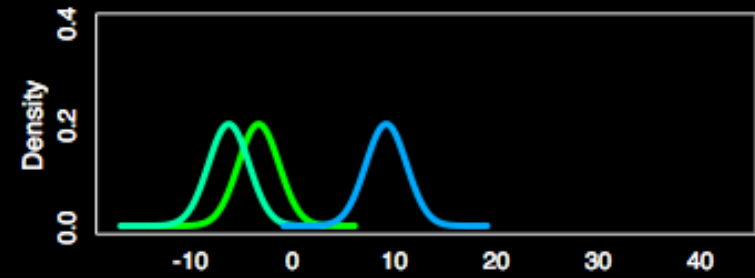
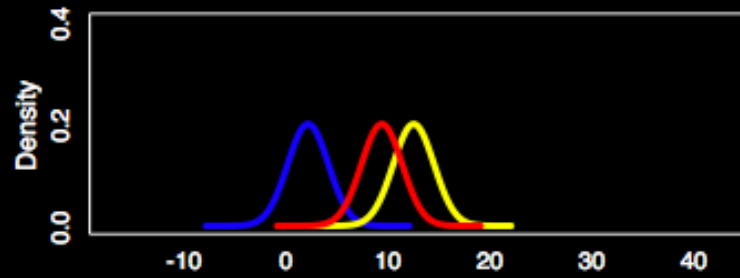
Fixed

Mixed

Sample 1



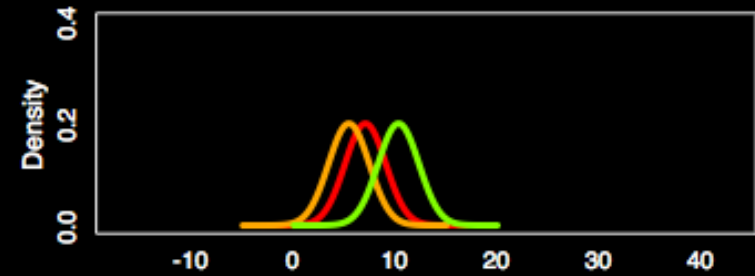
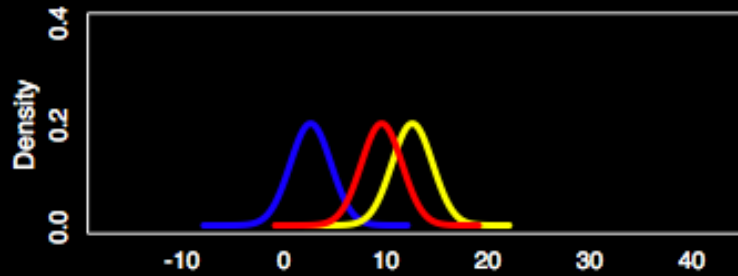
Sample 2



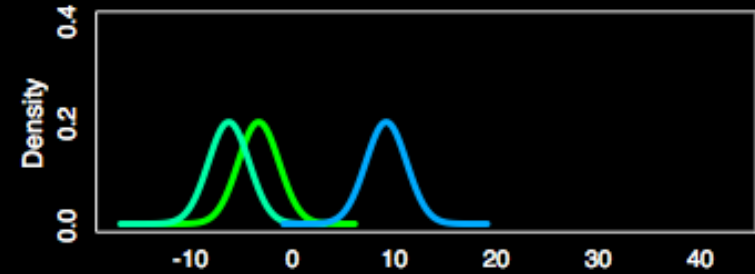
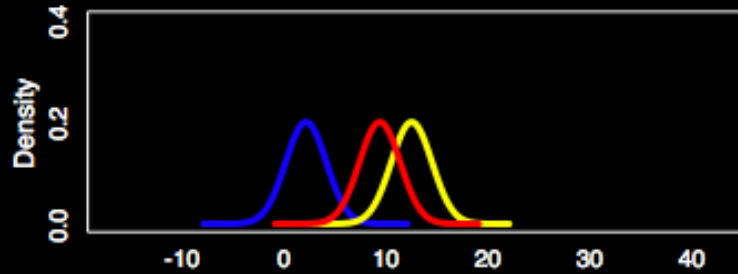
Fixed

Mixed

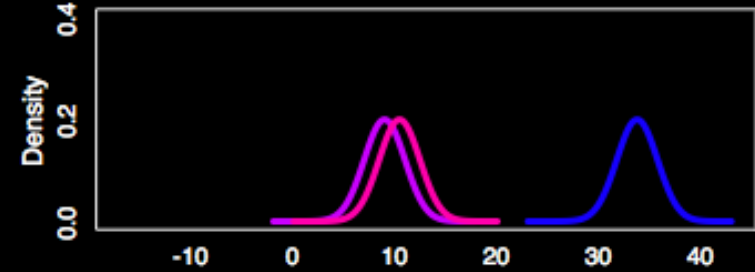
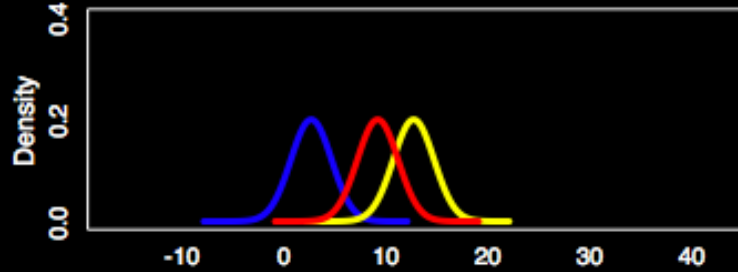
Sample 1



Sample 2



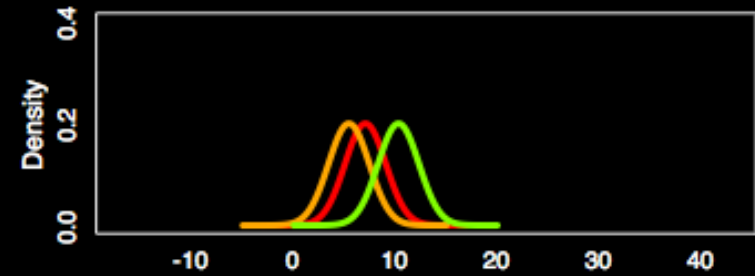
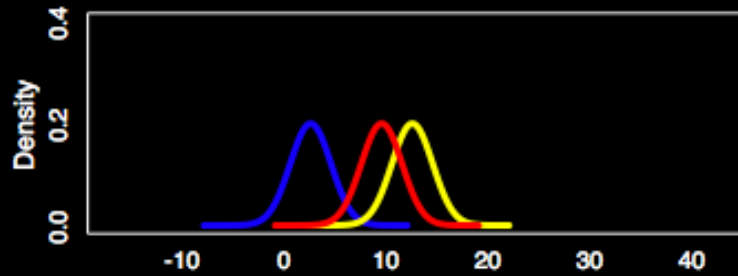
Sample 3



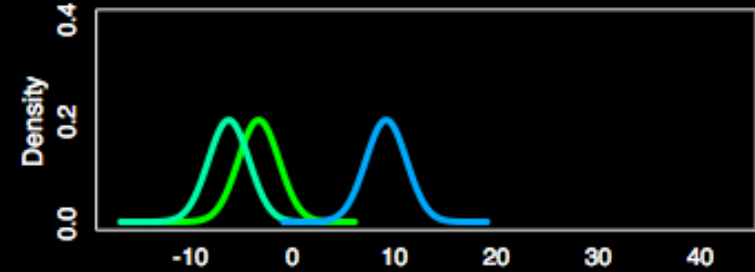
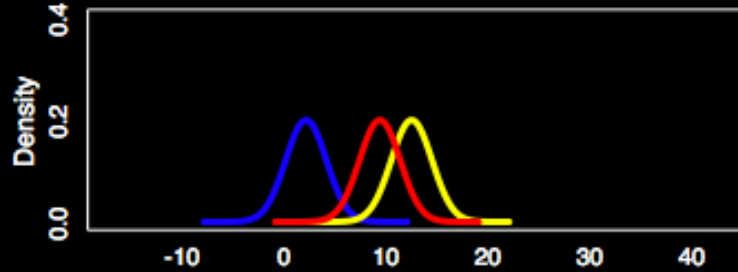
Fixed

Mixed

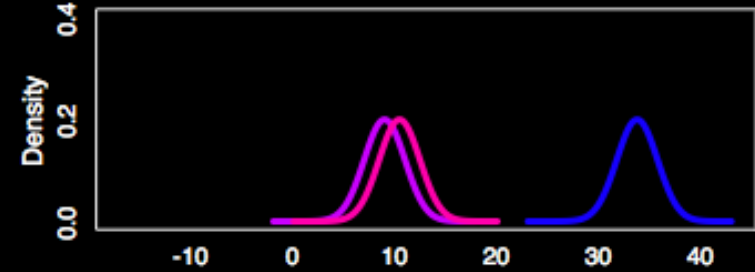
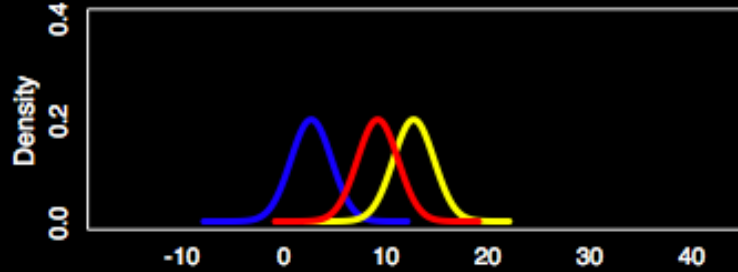
Sample 1



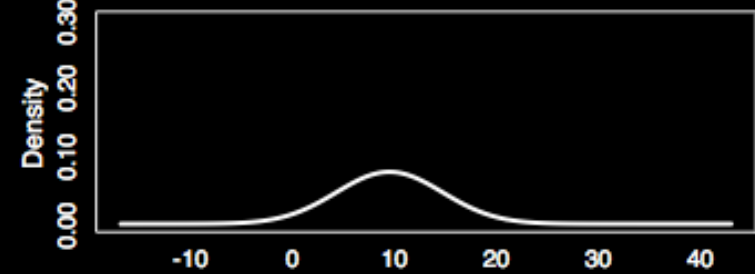
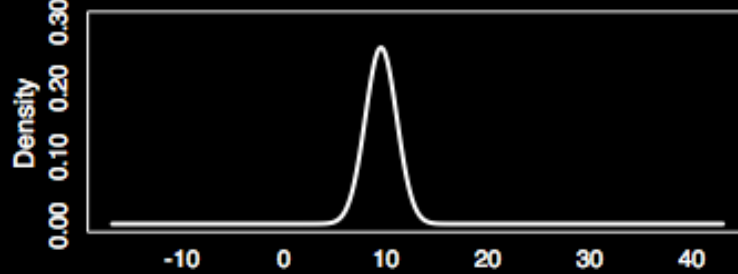
Sample 2



Sample 3



Group
Distribution



Important points so far...

- Ignoring random subject effect means that you're ignoring the fact that these subjects were randomly sampled
 - Inference only applies to sample you collected
- Including random subject effect *always* increases your variance

Important points so far...

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

Important points so far...

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

1 hair per subject

$$\sigma_{\text{MFX}}^2 = \sigma_{\text{W}}^2/4 + \sigma_{\text{B}}^2/4 = 12.5$$

25 hairs per subject

$$\sigma_{\text{MFX}}^2 = \frac{1}{4}\sigma_{\text{W}}^2/25 + \frac{1}{4}\sigma_{\text{B}}^2 = 12.26$$

Overview

- What is a mixed effects model
 - Fixed effects
 - Random effects
- 2-stage summary statistics approach
- How do different software packages work?
- Overview FSL modeling options

Fixed effects model:

modeling the mean of 3 females, 20 hairs

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

Fixed effect

Residual error

Mixed Effects Model

Stage 1

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

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

Stage 2

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

Random effect

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

Mixed Effects Model

Stage 1

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

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

Stage 2

$$\beta = X_g \beta_g + \eta \quad \text{Random effect}$$

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

Mixed Effects Model: All-In-One

$$Y = X X_g \beta_g + X \eta + \epsilon$$

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

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

$$\begin{array}{l}
 \text{Subject 1} \\
 \text{Subject 2} \\
 \vdots \\
 \text{Subject N}
 \end{array}
 \begin{pmatrix}
 \text{[wavy line]} \\
 \text{[wavy line]} \\
 \vdots \\
 \text{[wavy line]}
 \end{pmatrix}
 =
 \begin{pmatrix}
 \text{[checkerboard]} \\
 \text{[checkerboard]} \\
 \vdots \\
 \text{[checkerboard]}
 \end{pmatrix}
 \begin{pmatrix}
 \beta_{g1} \\
 \beta_{g2}
 \end{pmatrix}
 +
 \begin{pmatrix}
 \text{[checkerboard]} \\
 \text{[checkerboard]} \\
 \vdots \\
 \text{[checkerboard]}
 \end{pmatrix}
 \begin{pmatrix}
 \eta_{1,1} \\
 \eta_{1,2} \\
 \eta_{2,1} \\
 \eta_{2,2} \\
 \vdots \\
 \eta_{N,2}
 \end{pmatrix}
 +
 \begin{pmatrix}
 \epsilon_{1,1} \\
 \vdots \\
 \epsilon_{1,T} \\
 \epsilon_{2,1} \\
 \vdots \\
 \epsilon_{2,T} \\
 \vdots \\
 \epsilon_{N,1} \\
 \vdots \\
 \epsilon_{N,T}
 \end{pmatrix}$$

$$\begin{aligned}
 \text{Var}(\eta_{i,1}) &= \sigma_{btwn_1}^2 \\
 \text{Var}(\eta_{i,2}) &= \sigma_{btwn_2}^2
 \end{aligned}$$

$$\text{Cov} \begin{pmatrix} \epsilon_{i,1} \\ \epsilon_{i,2} \\ \vdots \\ \epsilon_{i,T} \end{pmatrix} = \sigma_{win_i}^2 V_i$$

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

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

Stage 2

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

Two-Stage Summary Statistics

Stage 1

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

$$\begin{pmatrix} Y_{2,1} \\ \vdots \\ Y_{2,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_2 + \begin{pmatrix} \epsilon_{2,1} \\ \vdots \\ \epsilon_{2,20} \end{pmatrix}$$

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

$$\epsilon_{i,j} \sim N(0, \sigma_{win}^2)$$

Two-Stage Summary Statistics

Stage 1

$$\begin{pmatrix} Y_{1,1} \\ \vdots \\ Y_{1,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_1 + \begin{pmatrix} \epsilon_{1,1} \\ \vdots \\ \epsilon_{1,20} \end{pmatrix}$$
$$\begin{pmatrix} Y_{2,1} \\ \vdots \\ Y_{2,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_2 + \begin{pmatrix} \epsilon_{2,1} \\ \vdots \\ \epsilon_{2,20} \end{pmatrix}$$
$$\begin{pmatrix} Y_{3,1} \\ \vdots \\ Y_{3,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_3 + \begin{pmatrix} \epsilon_{3,1} \\ \vdots \\ \epsilon_{3,20} \end{pmatrix}$$

$\epsilon_{i,j} \sim N(0, \sigma_{win}^2)$

Stage 2

Use first stage estimates

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

Two-Stage Summary Statistics

Stage 1

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

$$\begin{pmatrix} Y_{2,1} \\ \vdots \\ Y_{2,20} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta_2 + \begin{pmatrix} \epsilon_{2,1} \\ \vdots \\ \epsilon_{2,20} \end{pmatrix}$$

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

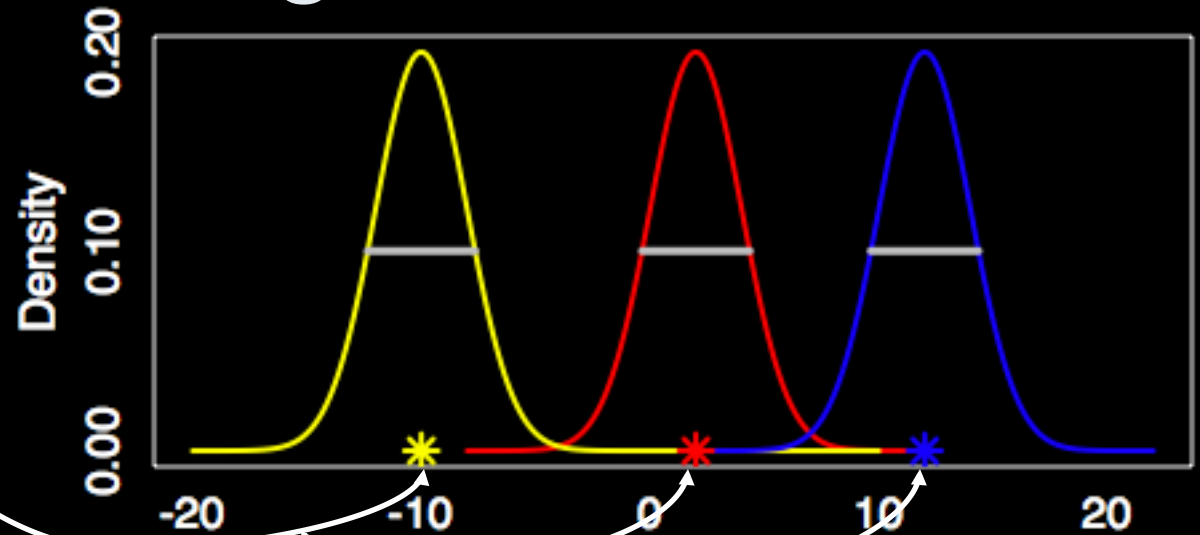
$$\epsilon_{i,j} \sim N(0, \sigma_{win}^2)$$

Stage 2

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

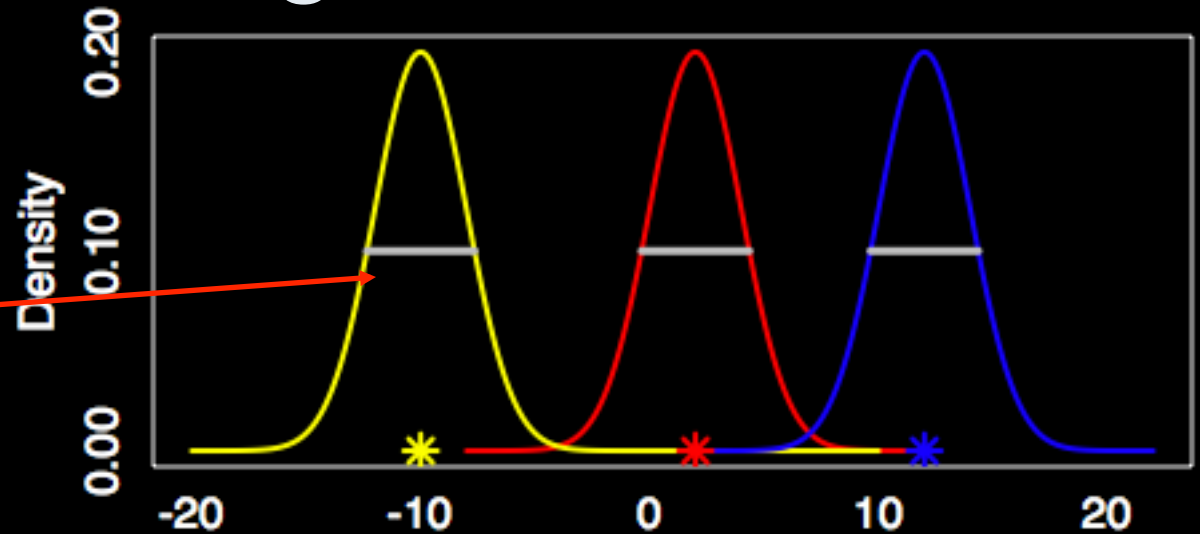
Two-Stage Model

- Stage 1
 - Means



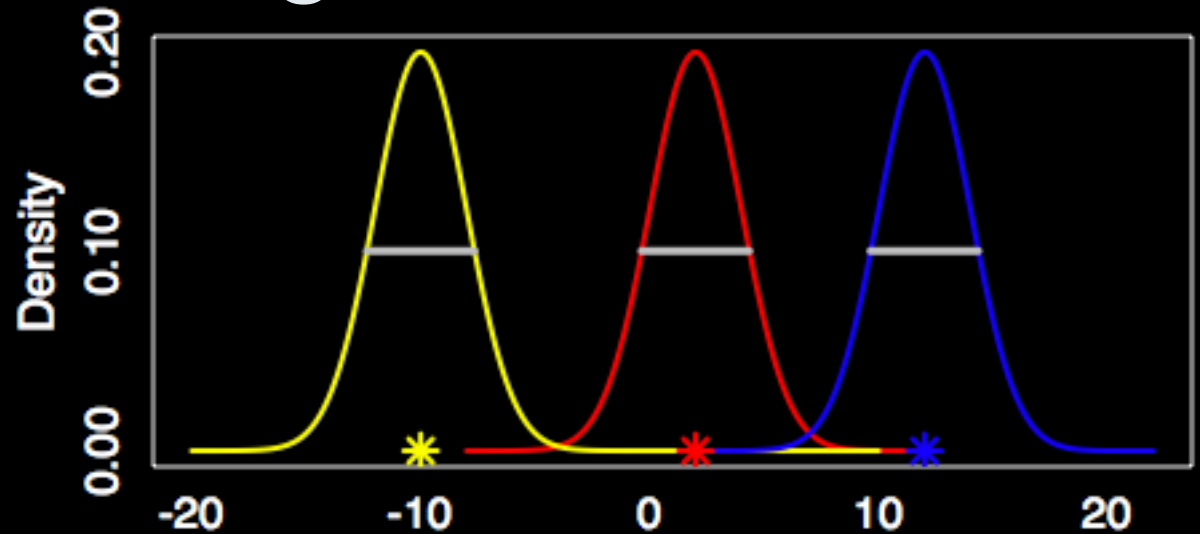
Two-Stage Model

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

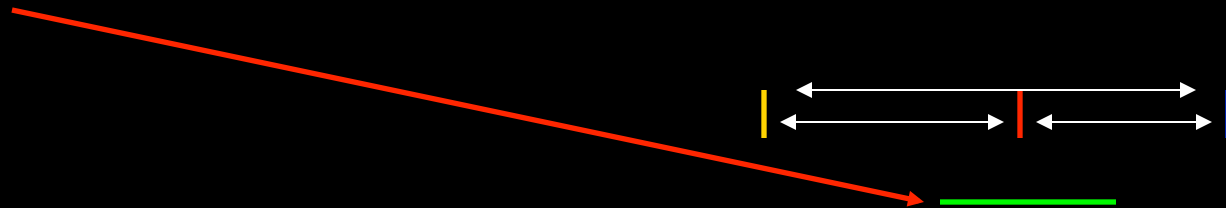


Two-Stage Model

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



- Stage 2
 - σ_{btwn}^2

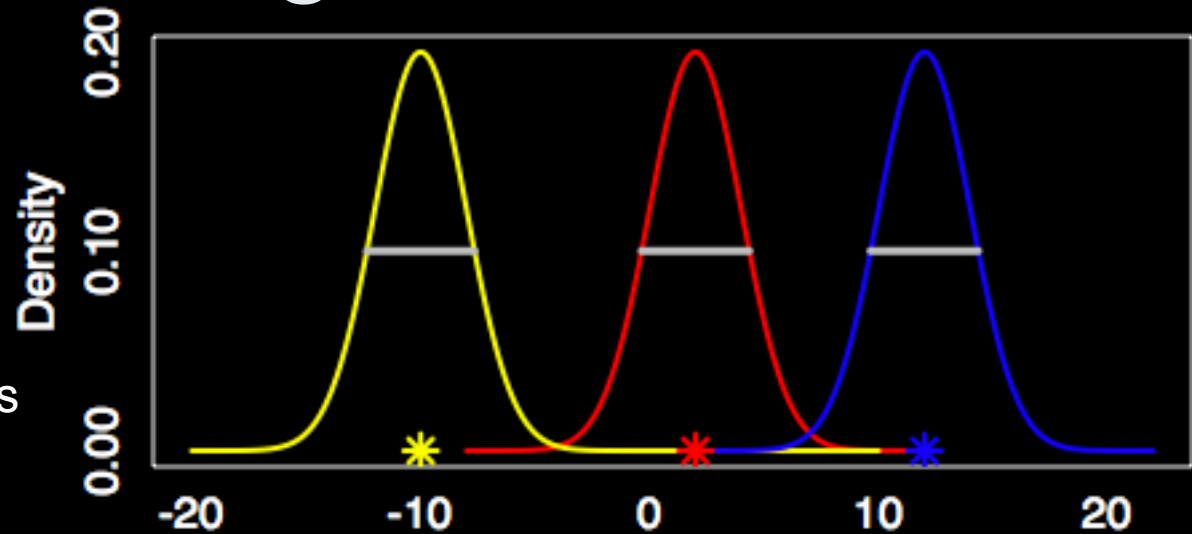


Two-Stage Model

- Stage 1

- Means

- σ_{win}^2
(same across subjects here)

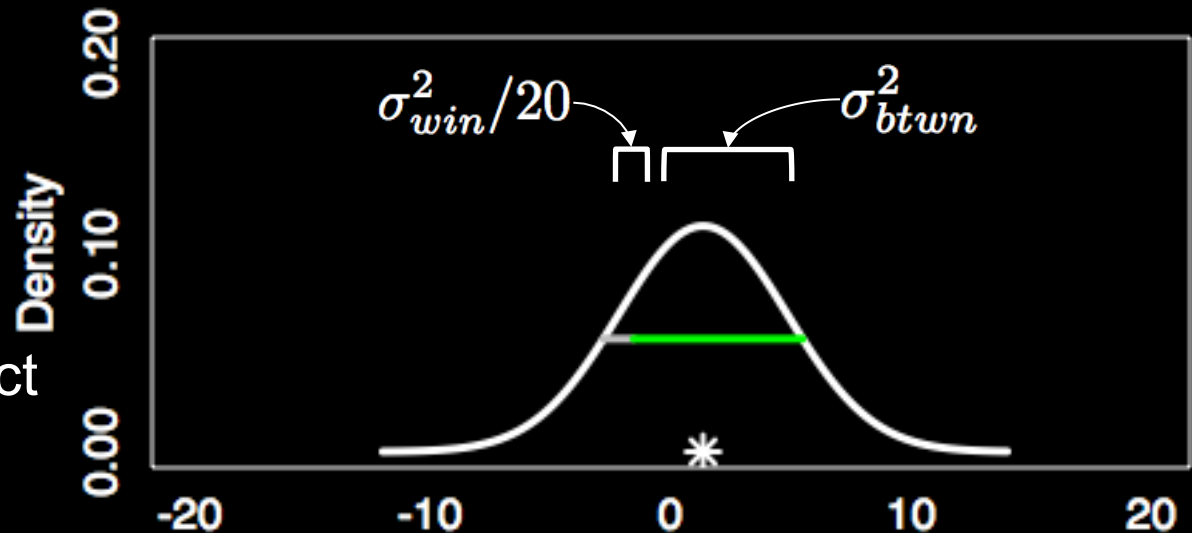


- Stage 2

- σ_{btwn}^2

- $\sigma_{mix}^2 = \frac{\sigma_{win}^2}{20} + \sigma_{btwn}^2$

- 20 hairs/subject



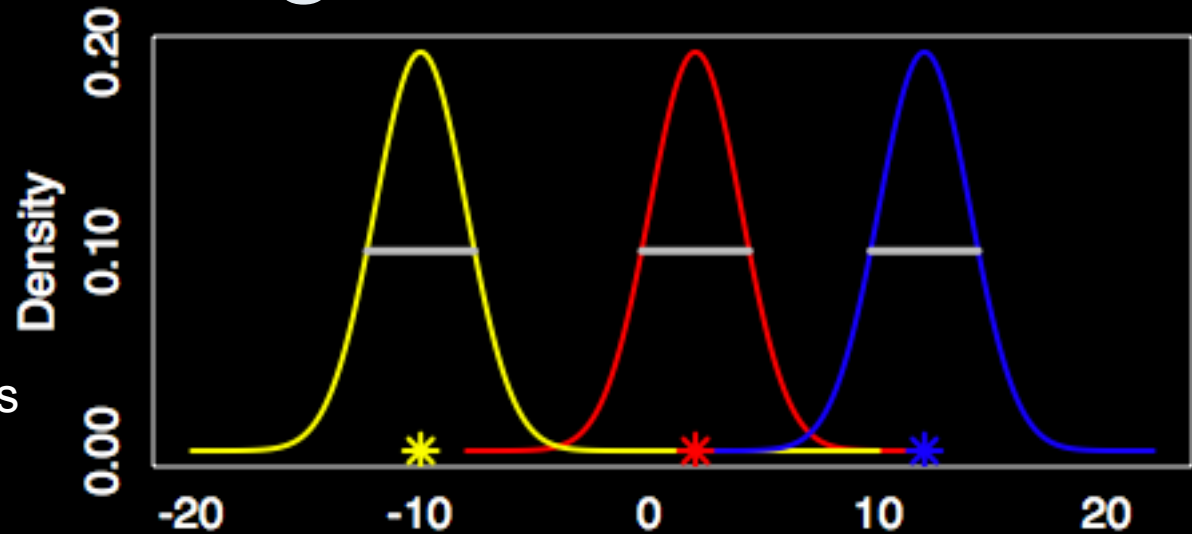
Two-Stage Model

- Stage 1

- Means

- σ_{win}^2

(same across subjects here)



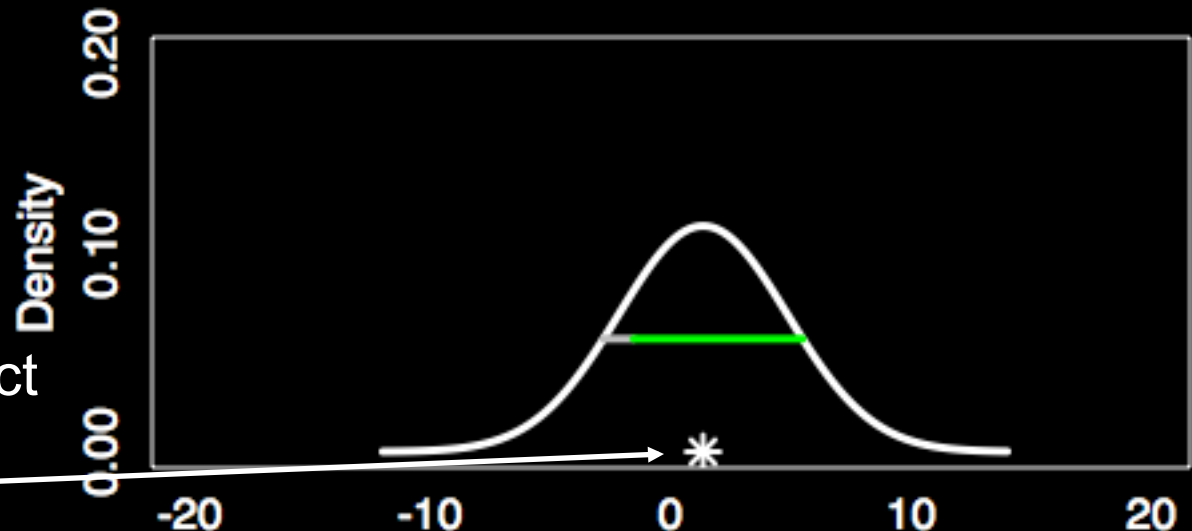
- Stage 2

- σ_{btwn}^2

- $\sigma_{mix}^2 = \frac{\sigma_{win}^2}{20} + \sigma_{btwn}^2$

- 20 hairs/subject

- Pop mean



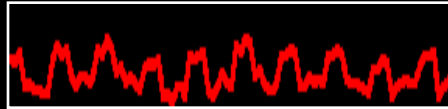
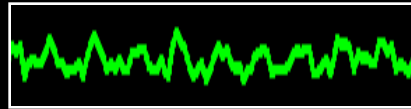
Two-Stage Model

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

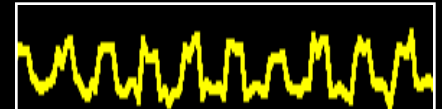
Two Stage Model fMRI

Stage 1

Estimate N
subject models



...



$$\begin{matrix} c\hat{\beta}_1 \\ \widehat{\text{Cov}}(c\hat{\beta}_1) \end{matrix}$$

$$\begin{matrix} c\hat{\beta}_2 \\ \widehat{\text{Cov}}(c\hat{\beta}_2) \end{matrix}$$

...

$$\begin{matrix} c\hat{\beta}_N \\ \widehat{\text{Cov}}(c\hat{\beta}_N) \end{matrix}$$

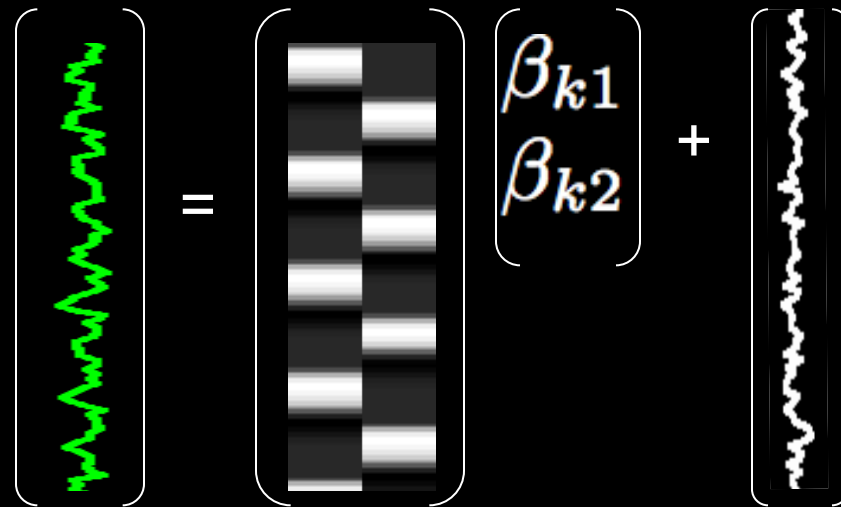
Stage 2

Estimate between
subject variance,
combine with
Stage 1 results

$$\begin{matrix} \hat{\beta}_g \\ \widehat{\text{Cov}}(\hat{\beta}_g) \end{matrix}$$

Stage 1: Subject Model

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


$$\begin{bmatrix} Y_k \end{bmatrix} = \begin{bmatrix} X_k \end{bmatrix} \begin{bmatrix} \beta_{k1} \\ \beta_{k2} \end{bmatrix} + \begin{bmatrix} \epsilon_k \end{bmatrix}$$

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

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

Stage 1: Estimation

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

Stage 1: Estimation

- W_k such that $W_k V_k W_k' = I_T$
- Whitenened 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 $W_k V_k W_k' = I_T$
- Whiten 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
 - $c\hat{\beta}_k = (X_k^{*'} X_k^*)^{-1} X_k^{*'} Y_k^*$
 - $\widehat{Cov}(c\hat{\beta}_k) = \hat{\sigma}_k^2 (X_k^{*'} X_k^*)^{-1}$

Stage 2: Group Model

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

$$\begin{pmatrix} c\hat{\beta}_1 \\ c\hat{\beta}_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \beta_g + \begin{pmatrix} \epsilon_g \\ \epsilon_g \\ \vdots \end{pmatrix}$$

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

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

Stage 2: Estimation

- W_g such that $W_g V_g W_g' = I_N$
- $W_g \hat{\beta}_{cont} = W_g X_g \beta_g + W_g \epsilon_g$
 $\hat{\beta}_{cont}^* = X_g^* \beta_g + \epsilon_g^*$
- $\hat{\beta}_g = \left(X_g^{*'} X_g^* \right)^{-1} X_g^{*'} \hat{\beta}_{cont}^*$
 $\widehat{Cov}(\hat{\beta}_g) = \left(X_g^{*'} X_g^* \right)^{-1}$
- $T = \hat{\beta}_g / \sqrt{\widehat{Cov}(\hat{\beta}_g)}$

Overview so far...

- The mixed model can be thought of as a two stage process
 - Stats software estimates both stages simultaneously
 - In fMRI it is computationally easier and more convenient to keep the stages separate
 - Easier to add new subjects
 - Not *exactly* the same as the stats software results

Overview so far

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

Overview so far

- When the model is estimated how is a subject with a high mfx variance treated differently than a subject with a low mfx variance?
 - Subjects with higher variability are down-weighted in the analysis!

Overview

- What is a mixed effects model
 - Fixed effects
 - Random effects
- 2-stage summary statistics approach
- How do different software packages work?
- Overview FSL modeling options

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

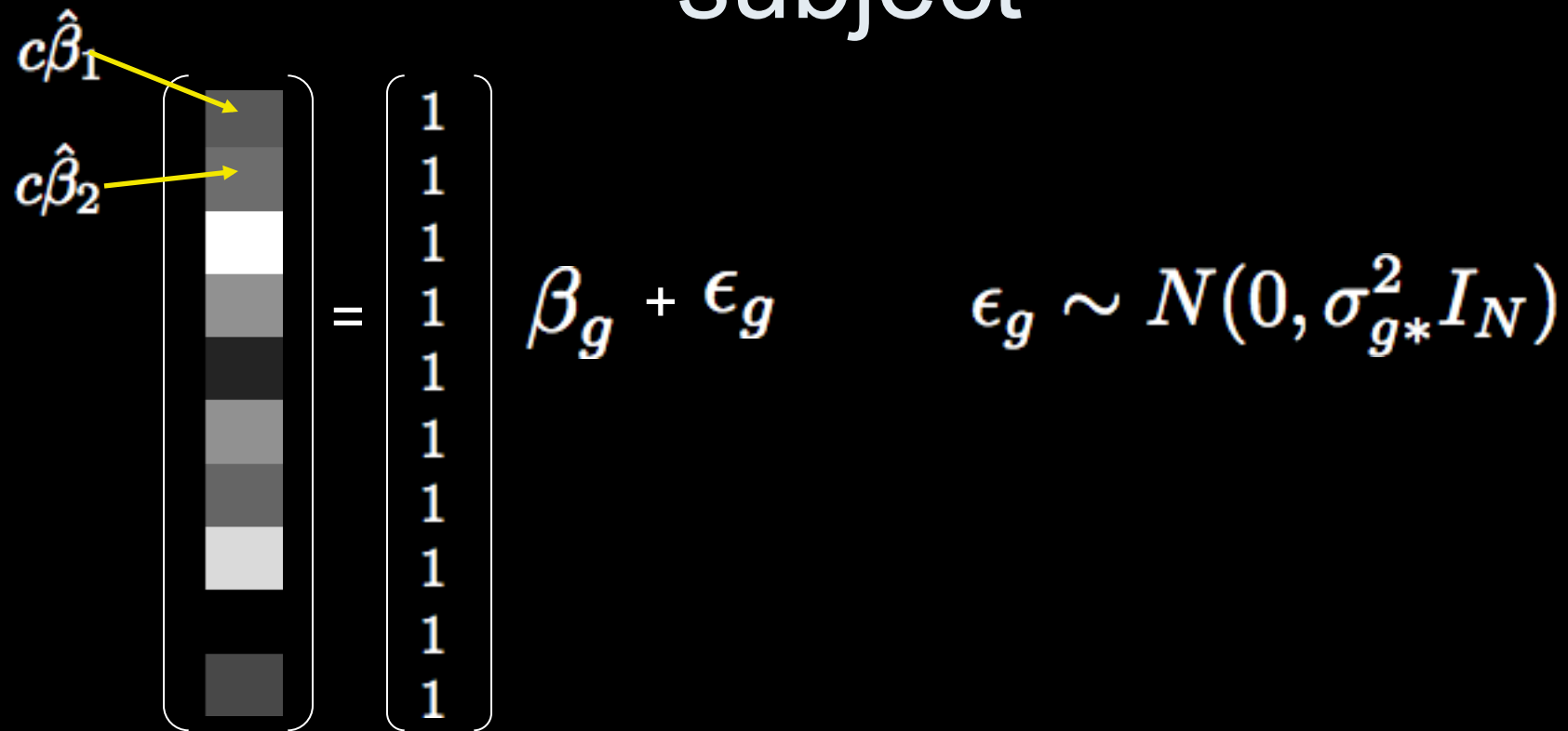
- Does not estimate σ_g^2
 - Assumes homogeneous variance across subjects
 - Assumes first level design is same across subjects

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

$$V_g = \sigma_{win_{all}}^2 I_N + \sigma_g^2 I_N = \sigma_{g*}^2 I_N$$

↑
OLS can be used

SPM2 : Single contrast per subject



$$\begin{bmatrix} c\hat{\beta}_1 \\ c\hat{\beta}_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \beta_g + \epsilon_g \quad \epsilon_g \sim N(0, \sigma_{g*}^2 I_N)$$

A one-sample T-test!

SPM2 : Multiple contrasts per subject

$$\begin{bmatrix} \hat{\beta}_{1,1} \\ \hat{\beta}_{1,2} \\ \hat{\beta}_{2,1} \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{g1} \\ \beta_{g2} \end{bmatrix} + \epsilon_g$$

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

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: FMRI 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

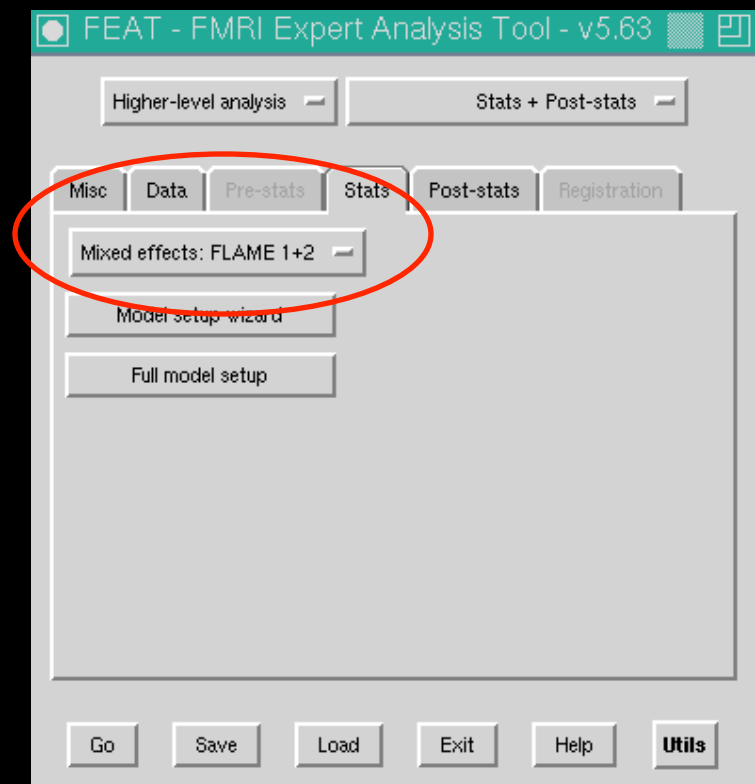
FSL

- 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?

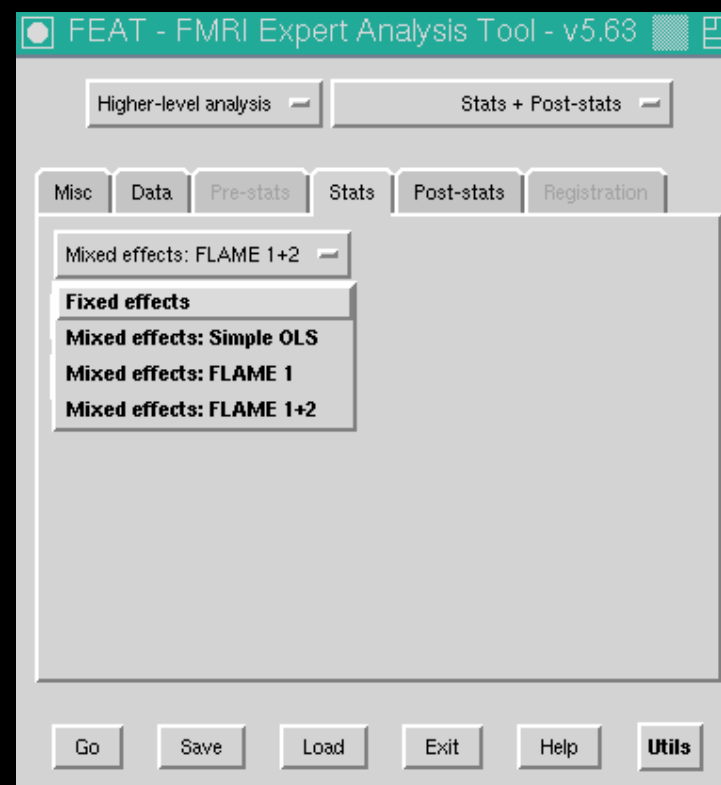
- FSL and FMRISTAT best for heteroscedastic variances
 - Different number of trials per subject
- SPM best for multiple correlated contrasts at group level
- Other differences in first level modeling may sway users one way or another

FSL Group Model Options



FSL Group Model Options

- Fixed effects
 - Only uses w/in sub variance
- Simple OLS
 - Assumes w/in sub variances are equal
- Flame 1& 2
 - w/in sub var and btwn sub var



W_g Matrix

- Recall we pre-multiply by W_g so our errors are uncorrelated and constant variance

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

$$V_g = \begin{pmatrix} \sigma_{win_1}^2 + \sigma_g^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{win_N}^2 + \sigma_g^2 \end{pmatrix} \rightarrow W_g = \begin{pmatrix} \frac{1}{\sqrt{\sigma_{win_1}^2 + \sigma_g^2}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{\sigma_{win_N}^2 + \sigma_g^2}} \end{pmatrix}$$

Act as weights

V_g Matrix Assumptions

- Fixed effects analysis

- Only appropriate for intermediate levels

- Assumes the between-run variability=0

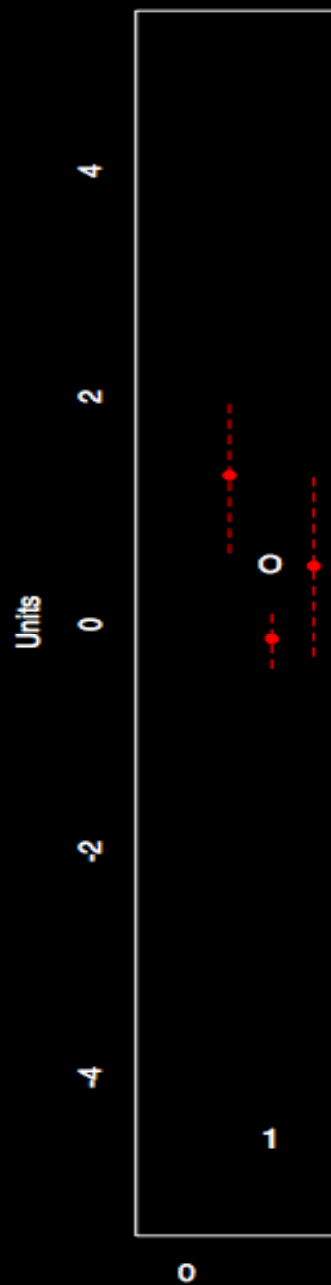
$$V_g = \begin{pmatrix} \sigma_{win_1}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{win_N}^2 \end{pmatrix}$$

- Why would we do this?

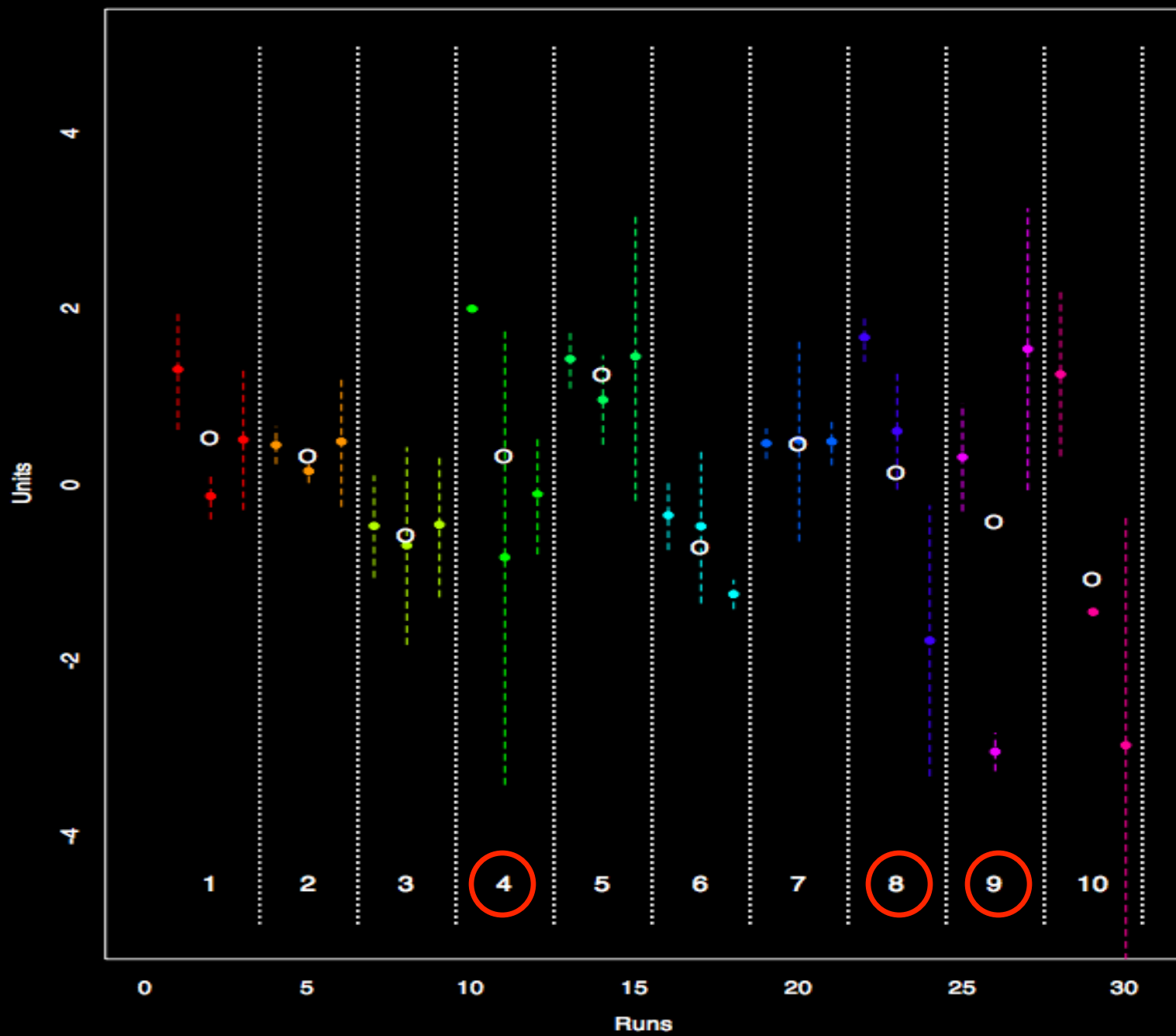
- What if df are low?

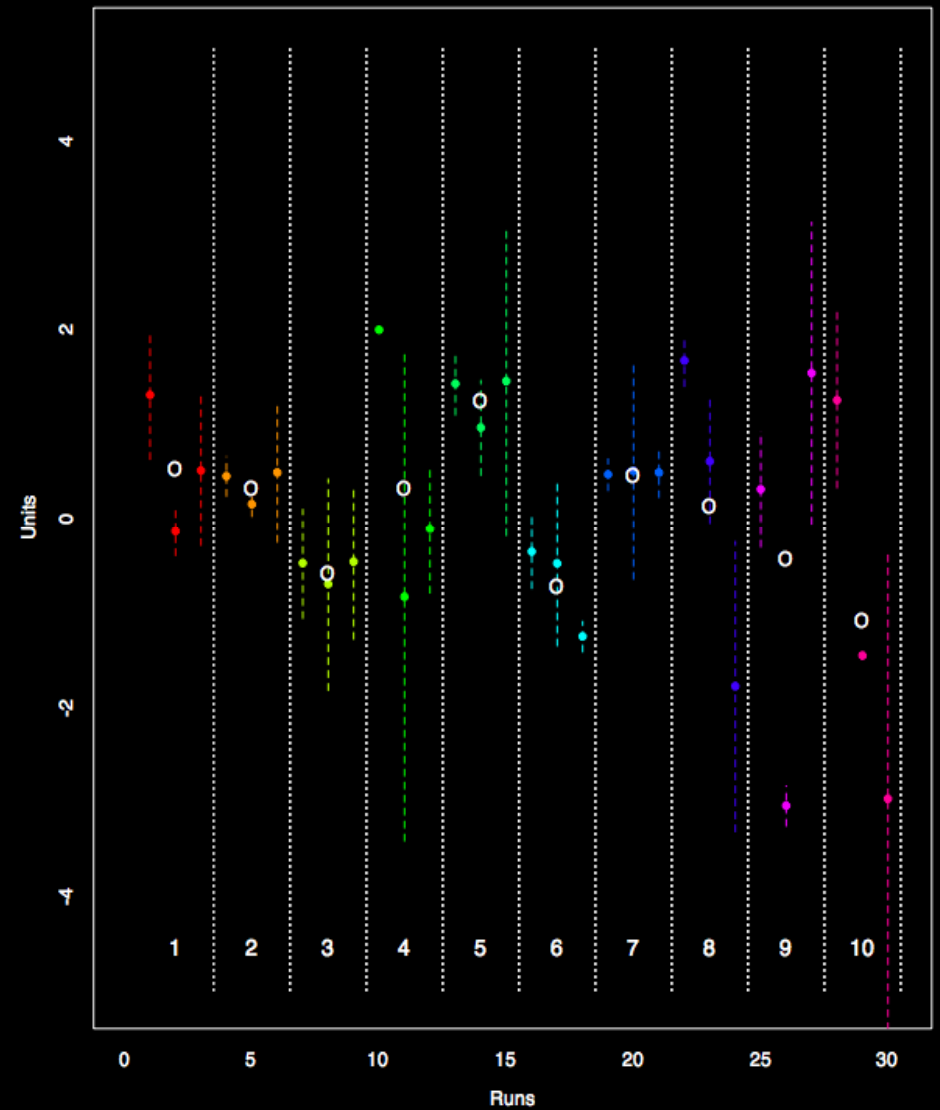
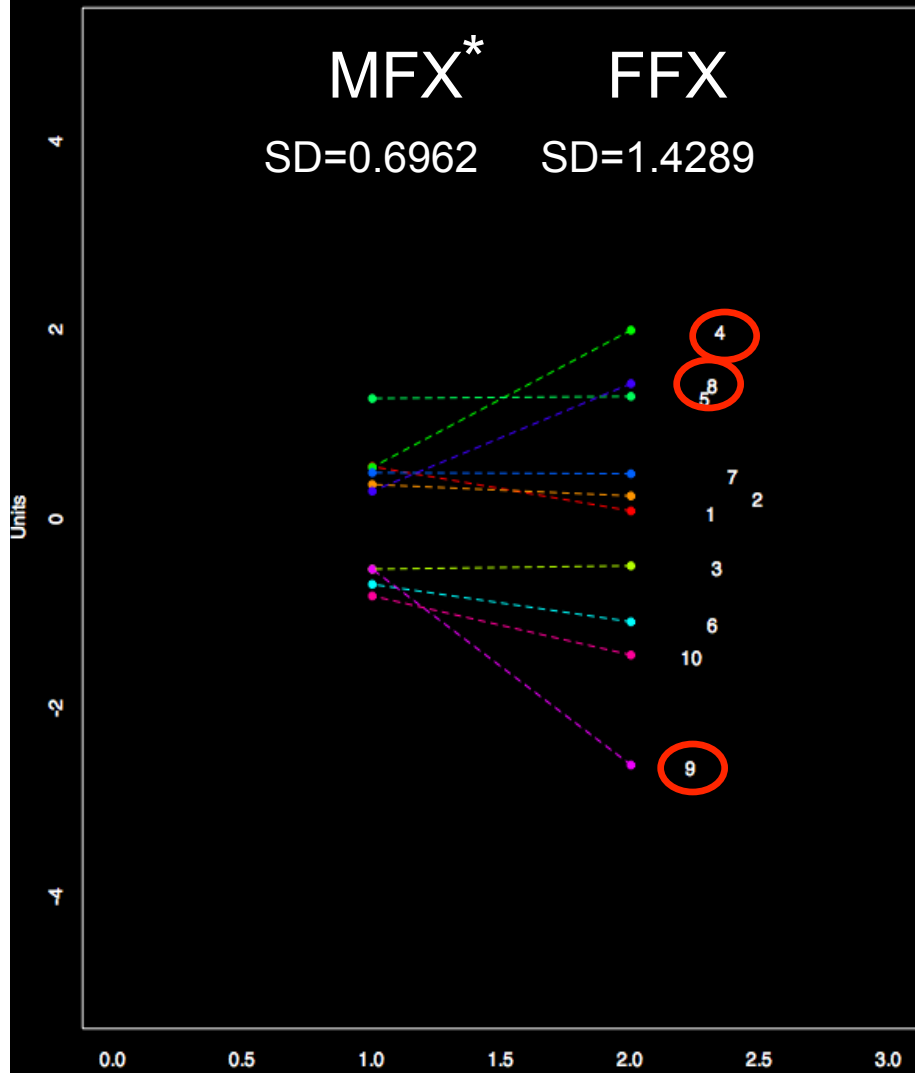
- $\hat{\sigma}_g^2$ has high variance

- If $\hat{\sigma}_g^2$ is too large, it will override differences in $\hat{\sigma}_{win_k}^2$



- 3 contrast estimates from 3 runs for 1 subject
- Dotted line indicates variance
- “o” marks unweighted mean





*Assuming overestimate of $\hat{\sigma}_g^2$

Fixed Effects

- Use to improve your mean estimates
 - eg correct trials
- Since variance is underestimated, you ***must*** only run this at an intermediate level
 - Higher level analysis soaks up rest of variance

Third Level Analysis

- Typically Flame and OLS have similar results
 - Flame is probably the best choice, since it adjusts for heterogeneous variance
 - OLS runs faster
 - OLS stats can be larger or smaller than Flame stats
- FE at level 3 is bad
 - Variance is underestimated
 - High risk of false positives

What is the “group” column for?

- This is the first column in the 2nd level Feat analysis design
- Assigns groups for estimating different *between* subject variances
- Must have a separable design
- Typically best if left to all 1's

Concluding Remarks

- Mixed models are appropriate for fMRI data
 - Include between-subject variance
 - Allows inference to be applied to entire population
- The two-stage summary statistics model
 - Computationally easier to estimate
 - Easier to add new subjects
- Software packages use the same basic model, but estimate σ_g^2 differently
- Use FE at intermediate levels and Flame at the top level in FSL