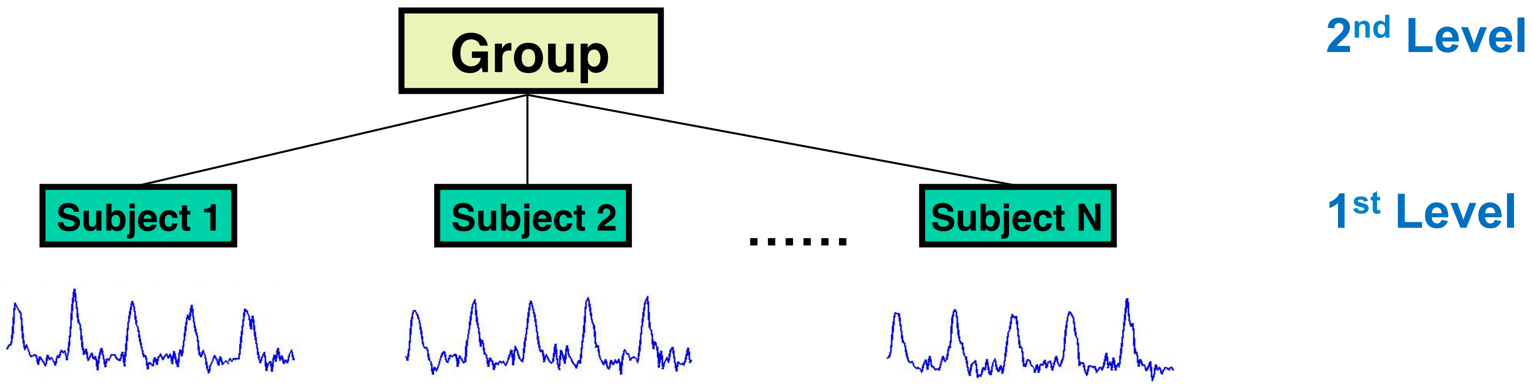
Module 17:
Group-level Analysis II

Statistical Analysis

- When applying statistics to real-world problems we need to separate between the model used to describe the data, the method of parameter estimation and the algorithm used to obtain them.
  - The model uses probability theory to describe the parameters of the unknown distribution thought to be generating the data.
  - The method defines the loss function that is minimized in order to find the unknown model parameters.
  - The algorithm defines the manner in which the chosen loss function is minimized.

Multi-level Model

- When performing group analysis we often use multi-level models. Often performed in two levels:
  - The first level deals with individual subjects.
  - The second level deals with groups of subjects.



- All inference typically performed in the ’massive univariate’ setting.

Preprocessing

- Several preprocessing steps should be applied prior to performing group analysis to ensure the validity of the results. They include:
  - Motion Correction
    - Intrasubject registration
  - Spatial normalization
    - Intersubject registration
  - Spatial Smoothing
    - Overcome limitations in the spatial normalization

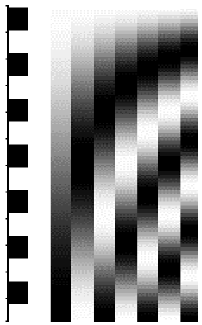
First level

- Suppose we have data from N different subjects, and for each subject *k*, we use the model:

where

$$\mathbf{Y}_k = \mathbf{X}_k \beta_k + \mathbf{e}_k$$

$$\mathbf{e}_k \sim N(0, \mathbf{V}_k)$$



- In the first level we have autocorrelated data with a relatively large number of observations

Let

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_N \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{X}_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{X}_N \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix} \qquad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix} \quad \text{and} \quad \mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{V}_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{V}_N \end{bmatrix}$$

First Level

- The full first level model can be written:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{e}$$

$$\mathbf{e} \sim N(0, \mathbf{V})$$

- Note that this model is separable, and it is possible to fit each subjects data individually.

Second Level

- The second level model can be written:

$$\beta = \mathbf{X}_g \beta_g + \boldsymbol{\eta}$$

$$\boldsymbol{\eta} \sim N(0, \mathbf{V}_g)$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{g0} \\ \beta_{g1} \end{bmatrix}$$

Here $\mathbf{X}_{\mathrm{g}}$ is the second level design matrix (e.g. separating cases from controls) and β_{g} the vector of second-level parameters.

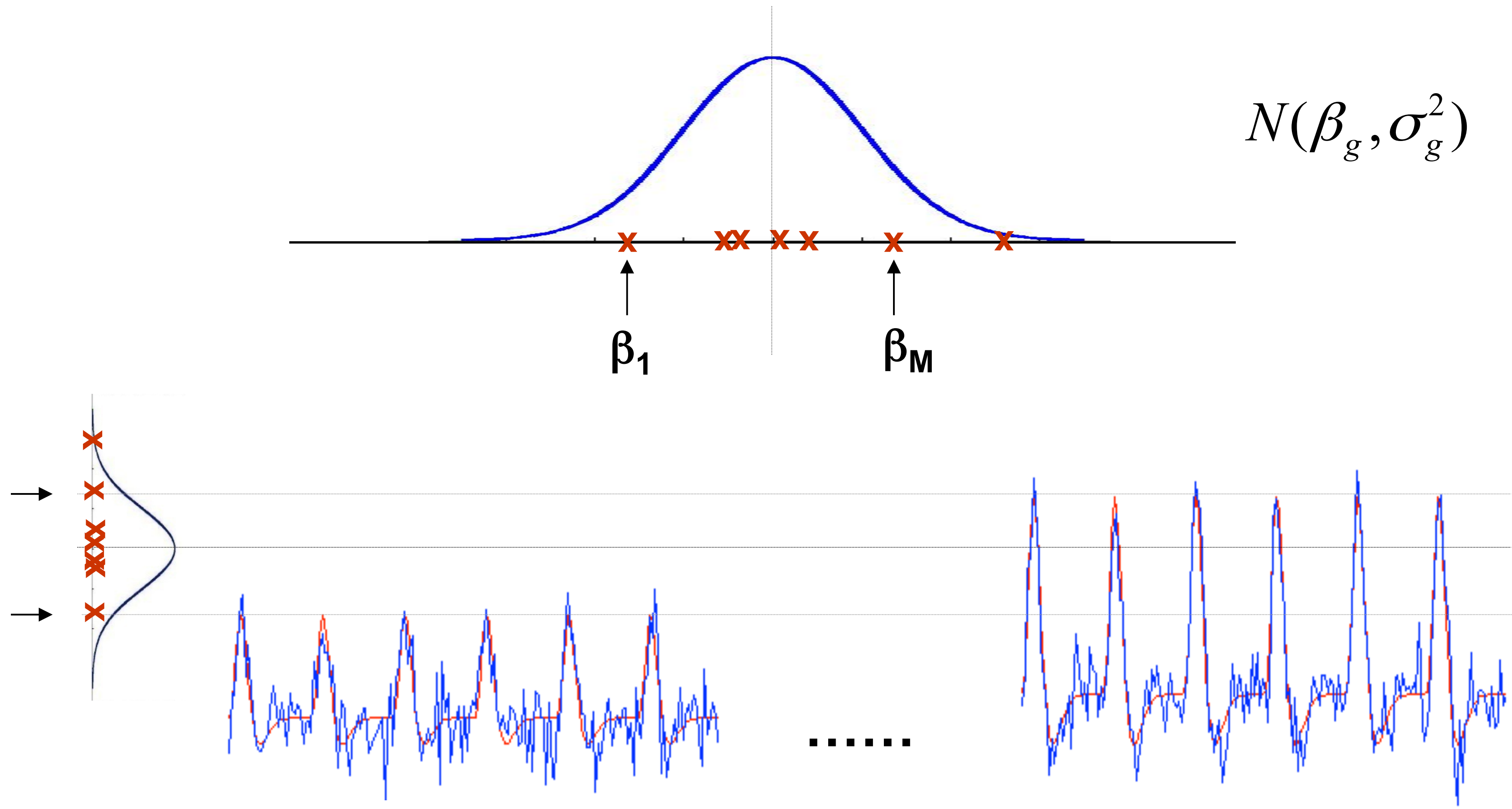
- In the second level we usually have IID data, but relatively few observations.

- The second level relates the subject specific parameters contained in β to the population parameters β_g.

- It assumes that the first level parameters are randomly sampled from a population of possible regression parameters.

- This assumption allows us to generalize the results to the whole population.

Illustration



Model Summary

- The model can be summarized as follows:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{e} \qquad \mathbf{e} \sim N(0, \mathbf{V})$$

$$\beta = \mathbf{X}_g \beta_g + \boldsymbol{\eta} \qquad \boldsymbol{\eta} \sim N(0, \mathbf{V}_g)$$

- Note that this model can be expanded further to incorporate more levels.

Mixed-effects Model

- The two-level model can be combined into a single level model:

$$
\begin{aligned}
\mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\
&= \mathbf{X}(\mathbf{X}_g\boldsymbol{\beta}_g + \boldsymbol{\eta}) + \boldsymbol{\varepsilon} \\
&= \mathbf{X}\mathbf{X}_g\boldsymbol{\beta}_g + \mathbf{X}\boldsymbol{\eta} + \boldsymbol{\varepsilon}
\end{aligned}
$$

Second-level Model

- We can write this:

$$\mathbf{Y} \sim N\left(\mathbf{X}\mathbf{X}_g\boldsymbol{\beta}_g, \mathbf{X}\mathbf{V}_g\mathbf{X}^T + \mathbf{V}\right)$$

Estimation

- Statistical techniques define the loss function that should be minimized in order to find the parameters of interest.
  - Commonly used techniques include: MLE & REML.

- Algorithms define the manner in which the chosen loss function is minimized.
  - Commonly used algorithms include: Newton-Raphson, Fisher-scoring, EM-algorithm, IGLS/RIGLS.

MLE vs REML

- Maximum likelihood method
  - Maximizes the likelihood of the data.
  - Produces biased estimates of the variance components.

- Restricted Maximum likelihood method
  - Maximizes the likelihood of the residuals.
  - Produces unbiased estimates of the variance components.

Algorithms

- Newton-Raphson
  - Iterative procedure that finds estimates using the derivatives at the current solution.
- Fisher Scoring
  - Iterative procedure that finds estimates using the Fisher Information.
  - Similar to Newton-Raphson.
- EM-algorithm
  - Iterative procedure that finds estimates from models that depend on unobserved latent variables (e.g., the second level error).

Software

- Different neuroimaging software packages have implemented mixed-effects models.
  - They differ in which method and algorithm they apply.

- However, a simple non-iterative two-stage least squares approach is used in most fMRI analysis.
  - The Summary Statistics Approach.

- Results from individual subject are reused in the second level, reducing the computational burden of fitting a full model.

Summary Statistics

- The summary statistics approach:
  - Fit a model to each subjects data.
  - Construct contrast images for each subject.
  - Conduct a t-test on the contrast images.

- Only the contrasts are recycled from the first level and not the variance components. Only one contrast can be estimated at a time.

- Assumptions:
  - Homogeneous intra-subject variance.
  - Balanced designs.

Group Analysis

- When using temporal basis sets at the first level it can be difficult to summarize the response with a single number, making group inference difficult.

- Here we can perform group analysis using
  - the “main” basis function,
  - all basis functions, or
  - re-parameterized fitted responses (Calhoun et al. (2004); Lindquist et al. (2009)).
    - Recreate the HRF and estimate the magnitude.
    - Use this information at the second level.

End of Module