Week 6 – First-level fMRI data analysis

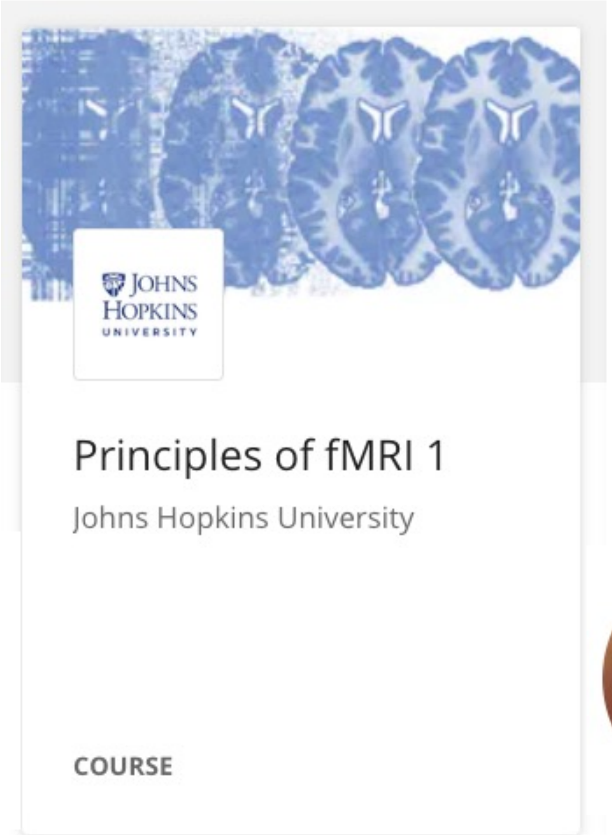
L06-01. Week 6 overview and General Linear Model

Byeol, Hongji, and Jungwoo

2 April 2021

Contents of this lecture

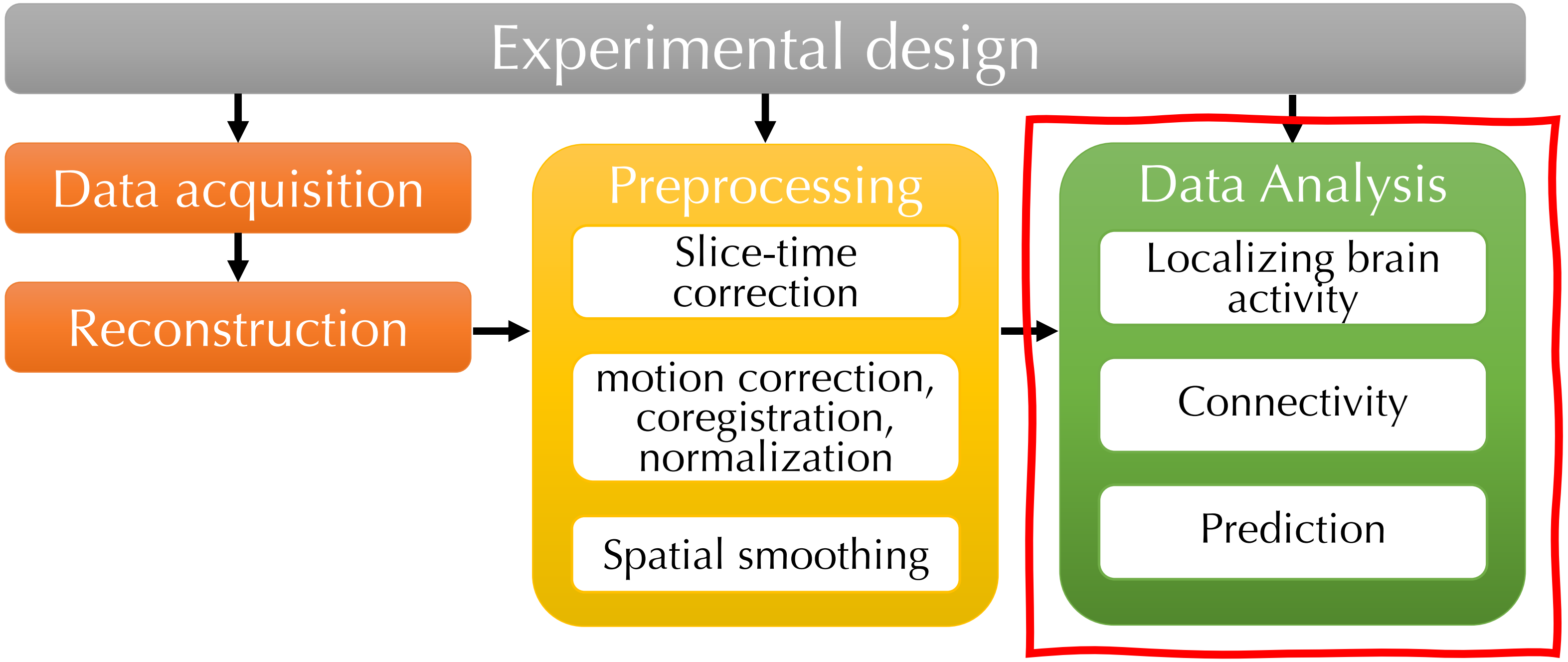
1. General linear model
2. BOLD & Canonical HRF
3. Overview of model building
4. Single-trial model (canlab_spm_fmri_model_job.m)
5. FIR model (canlab_spm_fmri_model_job.m)
6. Making custom regressors (onsets2fmridesign.m)
7. Residualize (canlab_connectivity_preproc.m)
8. Doing GLM with custom regressors (regress.m)
9. HRF modeling -
   Reason for using other Basis functions
10. HRF modeling -
    Example methods (IL & derivatives)



https://www.coursera.org/learn/functional-mri

1. General linear model - Data Processing Pipeline

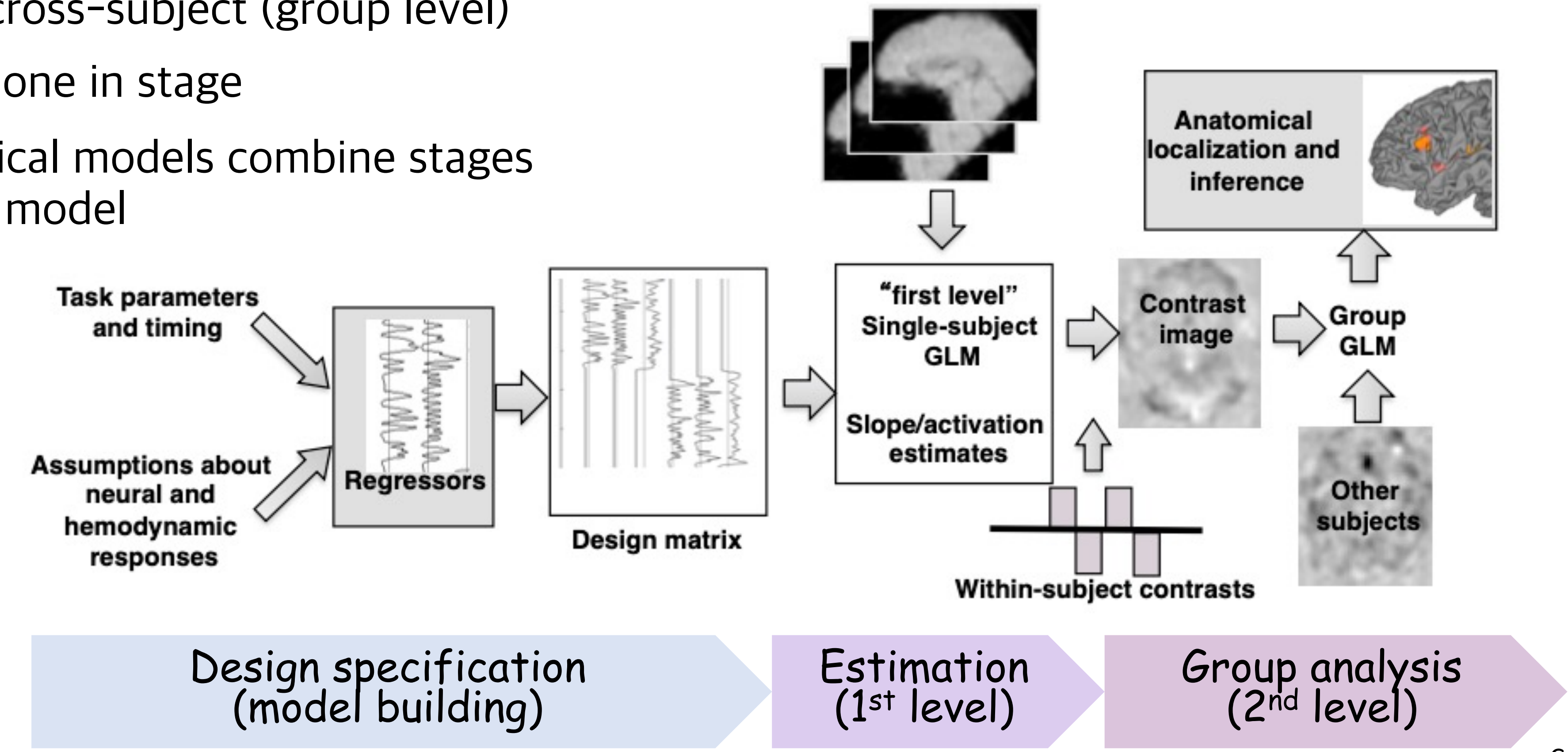
- Statistical analysis of fMRI data
  - Localizing brain areas activated by the task
  - Determining networks corresponding to brain function
  - Making predictions about psychological or disease states
- Data Processing Pipeline



Contents credit: Tor Wager

1. General linear model - Overview of the GLM analysis process

- Typically a two-level hierarchical analysis
    1. Within-subject (individual level)
    2. Across-subject (group level)
- Can be done in stage
- Hierarchical models combine stages into one model



Contents credit: Tor Wager

1. General linear model

- The **general linear model (GLM)** approach treats the data as a linear combination of model functions (predictors) plus noise (error).
- The model functions are assumed to have *known* shapes (i.e., straight line, or known curve), but their amplitudes (i.e., slopes) are *unknown* and need to be estimated.
- The GLM framework encompasses many of the commonly used techniques in fMRI data analysis (and data analysis more generally).
- Analysis
  - Regression
  - One-sample/Two-sample t-test
  - Analysis of variance (ANOVA)
  - Analysis of covariance (ANCOVA)
  - Multivariate analysis of variance (MANOVA)
  - …

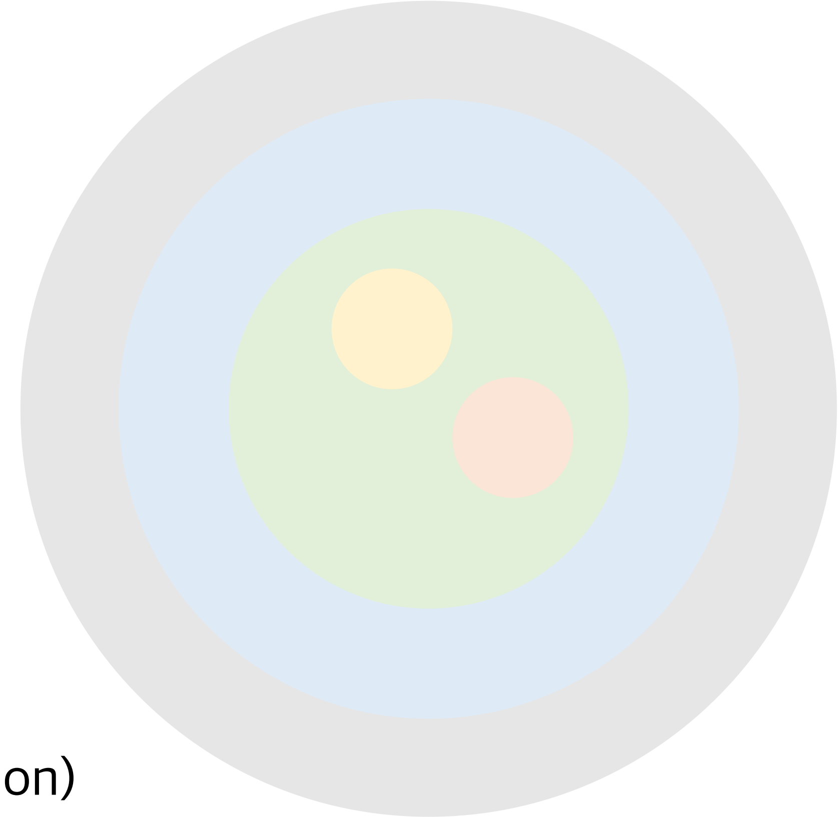
Contents credit: Tor Wager

1. General linear model - The GLM Family

Closed form solution

Iterative solutions

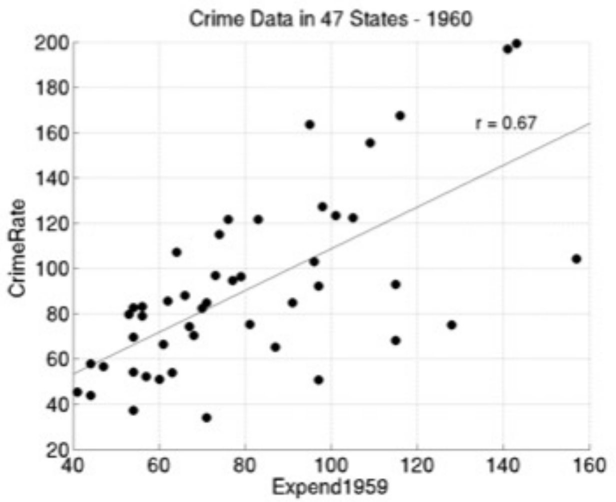
- Simple regression
- ANOVA
- Multiple regression
- General Linear Model
  - Mixed effects/hierarchical
  - Timeseries models (e.g., autoregressive)
  - Robust, penalized regression (LASSO, Ridge)
- Generalized Linear Model
  - Non-normal errors
  - Binary, categorical outcomes (logistic regression)



Contents credit: Tor Wager

1. General linear model

Specify model → Estimate → Statistical inference → Scientific interpretation



In case of the simple regression,

1. Simplification : linear relationship
2. Find slope and intercept
3. Test slope: finding p-value
4. What is the meaning of relationship?

$$y = ax + b + e$$

Outcome = slope x predictor + intercept + error
(DV) (constant) (residual)

Contents credit: Tor Wager

1. General linear model - The GLM Family

Dependent Variable	Predictors	Analysis
One continuous	Continuous 1+	Multiple regression
	Categorical 1+	ANOVA
	Categorical, continuous	ANCOVA
2 repeated measures	0 between 1 within	Paired t-test
k repeated measures	0 between k-1 categoraical within	One-way RM-ANOVA, MANOVA
4+ repeated measures	0 between 2+ categoraical within	Factorial RM-ANOVA
4+ repeated measures	1+ between 2+ categoraical within	"GLM"

Generalized least squares (iterative): Correlated errors (e.g., timeseries)

Correlated errors across repeated measures

Contents credit: Tor Wager

- Structural model for regression

- Variables
- Parameters

Dependent variable, predictor, predictor, predictor

$$\underbrace{y_i}_{\text{Dependent variable}} = \underbrace{\beta_0}_{\text{intercept}} + \beta_1 \overbrace{x_{i1}}^{\text{predictor}} + \underbrace{\beta_2}_{\text{slope}} \overbrace{x_{i2}}^{\text{predictor}} + \ \ldots + \underbrace{\beta_k}_{\text{slope}} \overbrace{x_{ik}}^{\text{predictor}} + \underbrace{\varepsilon_i}_{\text{error}}$$

- Solve for beta vector which minimizes sum of squared residuals
  - Estimating $\widehat{\beta_n}$
- Matrix notation $y = X\beta + \varepsilon$

Contents credit: Tor Wager

- With matrix notation

$$Y = X\beta + \varepsilon$$

as

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Observed data
Outcome data

Design matrix

Model parameters

Residuals

Contents credit: Tor Wager

1. General linear model - Regression model

$$Y = X\beta + \varepsilon$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Observed data
Outcome data

Design matrix

Model parameters

Residuals

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \beta_0 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \beta_1 \begin{bmatrix} X_{11} \\ X_{21} \\ \vdots \\ X_{1p} \end{bmatrix} + \cdots + \beta_p \begin{bmatrix} X_{1p} \\ X_{2p} \\ \vdots \\ X_{np} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Contents credit: Tor Wager

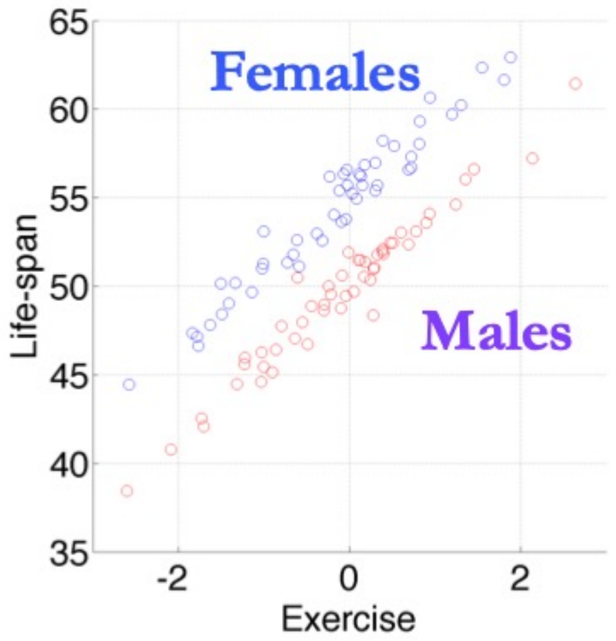
1. General linear model - Example of regression model

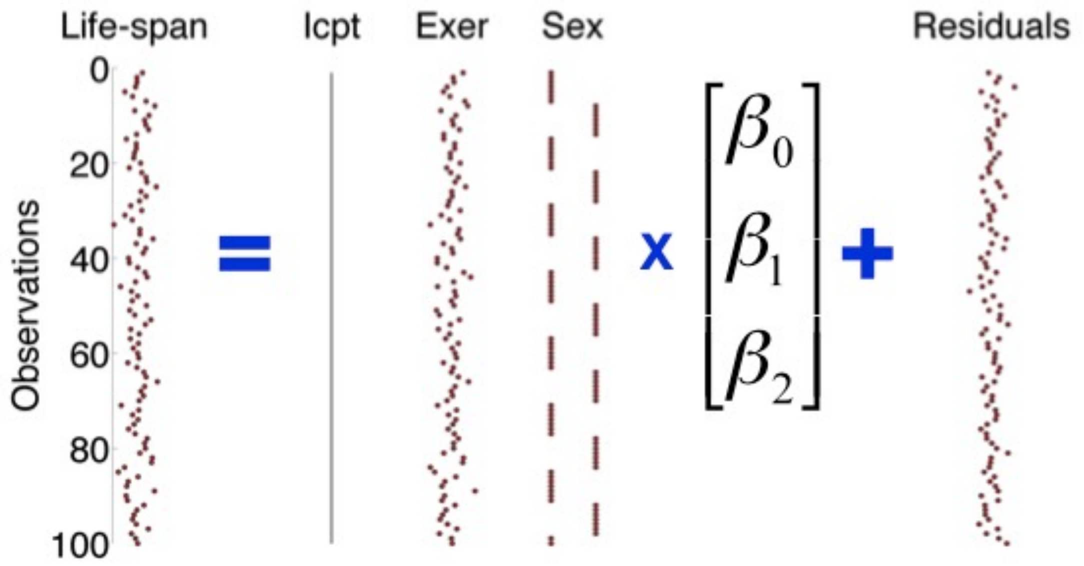
- Does exercise predict life-span? (not real data)
- *ANCOVA*
- Control for other variables that might be important, i.e., gender (M/F)

Observed data
Outcome data

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \dots & X_{1p} \\ 1 & X_{21} & \dots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \dots & X_{np} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Design matrix · Model parameters · Residuals





Contents credit: Tor Wager

1. General linear model - Design matrix

- In fMRI data, the design matrix specifies how the factors of the model change over time.
- The design matrix is an $n \times p$ matrix where n is the number of observations over time and p is the number of model parameters.

$$\mathrm{X} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \end{bmatrix}$$

- We'll see the example of design matrix in following videos!

Contents credit: Tor Wager

Cocoan 101

https://cocoanlab.github.io