Statistical Analysis of fMRI data

Honors H195. Imaging from Molecules to Mind

Prof. Yassa | 213 Qureshey Research Lab | myassa@uci.edu
Group fMRI

- Individual subject findings do not generalize to the population
- Need a representative sample of the population tested → not always easy
- Selection Bias: e.g. young, healthy, cognitively normal college students
Normalization to Stereotactic Space

- Before we combine data from several subjects we have to account for inter-individual anatomical variability.
- We do this by normalization (warping to a standard template), also known as stereotaxis.
- For example, the space of Talairach and Tournoux
The Talairach Transformation

- A sequence of two transformations:

1. A **rigid alignment** of the commissures (AC-PC) (rotation, translation, 6 DOF, linear)

**5 markers:** AC superior edge, AC posterior margin, PC inferior edge, first mid-sag point, second mid-sag point.
2. A **piecewise linear deformation** of the brain
(piecewise linear scaling: 12 DOF, nonlinear)

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most anterior to AC</td>
<td>70 mm</td>
</tr>
<tr>
<td>AC to PC</td>
<td>23 mm</td>
</tr>
<tr>
<td>PC to most posterior</td>
<td>79 mm</td>
</tr>
<tr>
<td>Most inferior to AC</td>
<td>42 mm</td>
</tr>
<tr>
<td>AC to most superior</td>
<td>74 mm</td>
</tr>
<tr>
<td>AC to left or right</td>
<td>68 mm</td>
</tr>
<tr>
<td>Length of cerebrum</td>
<td>172 mm</td>
</tr>
<tr>
<td>Height of cerebrum</td>
<td>116 mm</td>
</tr>
<tr>
<td>Width of cerebrum</td>
<td>136 mm</td>
</tr>
</tbody>
</table>
The Talairach Atlas - Regional Labels

5 levels of structures outlined:

**Main structures**
(left, right cerebrum, cerebellum, brainstem...)

**Lobes**
(temporal, frontal, parietal, posterior, occipital, limbic, anterior; midbrain...)

**Gyri**
(temporal, precentral, fusiform; thalamus, ventricles...)

**Matter**
(white matter, Gray matter, CSF)

**Brodmann areas**
(areas 1-47, hippocampus, putamen...)

7 regions
12 regions
55 regions
3 regions
71 regions
Pros and Cons of Talairach Transform

- **Pros:**
  - In widespread use in the imaging community
  - Allows for averaging of fMRI data between subjects
  - Allows for comparison of activation foci (coordinates)
  - Easy to use

- **Cons:**
  - Based on an elderly alcoholic French woman’s brain (not representative)
  - Not appropriate for all brains (e.g. Japanese brains)
  - Activation foci can vary considerably
  - Alignment of subcortical structures is very rough (40%)

- Alternative: **ICBM template** (based on hundreds of scans)
Next step - smoothing (i.e., blurring)

- After normalization, often a smoothing step is employed.
- Each voxel’s intensity value is replaced by a weighted average of all neighboring voxels (convolution with a Gaussian kernel).
- Reduces inter-subject anatomical variability
Basics of group fMRI design

- Prepare data for analysis
  - Register to anatomical template using the most accurate method available
  - Use a sample-representative template
  - Smooth data to reduce residual anatomical variability
- Statistical Analysis
  - Hemodynamic modeling per subject (1st level)
  - Secondary model taking into account random effects
  - Correct for multiple comparisons appropriately
  - Include covariates in the model when appropriate
How to Analyze Group Data

- Three statistical approaches in group fMRI:
  - Fixed Effects
  - Random Effects
  - Conjunctions
How to Analyze Group Data

Source: Huettel et al., Functional Magnetic Resonance Imaging, p. 353
Fixed Effects

• Assume that the experimental manipulation is **fixed** and has the **same effect** on the BOLD response in all subjects
• Treats all subjects and sessions as part of the **same** experiment
• Only accounts for **within-subject variation**, and therefore is only valid for subjects in the experiment
• **No inferential power**
Fixed Effects

- **Time course addition**
  - Combining all data points from all subjects into a single t-test with very high df
  - E.g. fMRI experiment with 10 subjects and 20 data points each (10 during the active conditions, and 10 during the rest condition), we have 198 degrees of freedom.
  - Increased power (compare to fMRI analysis with 198 subjects!)

- **Time course averaging**
  - Also involves combining time courses from each subject but using a summary statistic approach
  - Fewer degrees of freedom, but more consistency within groups (due to averaging)
Fixed Effects

• Case Scenario:
  • Test 10 subjects
  • 5 subjects have a very large effect, while the other 5 subjects have almost no effect at all
  • After averaging, you find that you have a moderate effect in the average time course
  • Can you generalize the effect?
Random Effects

- This type of analysis incorporates information about the **distribution of the effect among subjects**.
- Each subject is considered as one of many possible subjects that could have participated in the study.
- Accounts for **subject-by-condition** effects (inter-subject variation).
Random Effects

- 2 stage procedure

Level 1

Level 2
Random Effects

- First, statistical maps are computed for each subject.
- Statistical maps are entered in a secondary test using the general linear model to evaluate significance.
- Can make *inferences at the population level*
- If secondary test is significant at the established alpha level, then we can conclude that the experimental condition would have the same effect on the population from which the subjects are drawn.
Random Effects

- Distributional Assumptions:
  - Normality
  - Error Homogeneity (Homoscedasticity)
  - Independence of observations
- Power and sensitivity go up with $N$ (more df)
- Results are valid, provided that assumptions hold true.
- Results can be generalized to the entire population.
Fixed vs. Random Effects

Individual subject statistical maps

FFX model

RFX model
Conjunctions

- The **joint refutation of** two or more null hypotheses
- Evaluates activation that is significant in **every subject-specific contrast**.
- **Typical** activation vs. average activation
- e.g. The *Anteater* tattoo
Conjunctions

- An alternative to fixed effects modeling
- Estimates the significance for activation that is **jointly present**
- Benefits from the power of fixed effects models (large df), but is less likely to be influenced by individual extremes
- Contrasts of interest from each subject are linearly combined to test the null hypothesis that activation is jointly zero.
- Answers the question: “Do ALL subjects activate X?”
Conjunctions

- Inference is still limited to sample but...
- Can infer “typical” activation (if sample is truly random)
Thresholding and Inference

- Statistical inference in fMRI is constrained by the need to exert control over Type I and Type II errors.
- We test statistics against the null hypothesis ($H_0$), with the null distribution (i.e. the distribution expected when there is no finding).
Thresholding and Inference

- **Type I error (alpha - $\alpha$) rate:** The probability of incorrectly rejecting a true null hypothesis (accepting a false positive).

- **Type II error (beta - $\beta$) rate:** The probability of incorrectly accepting a false null hypothesis (rejecting a true negative).
Thresholding and Inference

High Threshold
- Good Specificity
- Poor Sensitivity (risk of false negatives)

Med. Threshold

Low Threshold
- Poor Specificity (risk of false positives)
- Good Sensitivity
Thresholding and Inference

- Why do we threshold?
- Control **Type I error rate** (false positives)
- Since analysis is voxel-based, why not just take the X # of voxels with the highest intensity?
  - Problem 1: Voxel intensities are not independent
  - Problem 2: Statistical images are based on multiple comparisons (voxel-by-voxel), thus Type I error is inflated.

  Which of 100 voxels significant?
  \[ \alpha = 0.05 \implies 5 \text{ false positives clusters} \]

Which of **100,000** voxels are significant?
\[ \alpha = 0.05 \implies 5,000 \text{ false positive voxels} \]
Voxel-level inference

- Retain voxels above $\alpha$-level threshold $u_\alpha$
- Gives best spatial specificity
  - $H_0$ at a single voxel can be rejected
- Sensitive to high-intensity signals
Cluster-level inference

- Two step-process
  - Define clusters by arbitrary threshold $u_{\text{clus}}$
  - Retain clusters larger than $\alpha$-level threshold $k_\alpha$
- Sensitive to spatially extended signals