**Single-Session Analysis**

FMRI data

- Time
- Preprocessed data
- Motion correction
- High-pass filtering
- Spatial smoothing

Voxel-wise single-subject analysis

- Design matrix
- Stimulus/task timings
- Voxel time-series data
- GLM
- Single-subject effect size statistics

Effect size statistics

- Contrast
- Statistic Image
- Thresholding
- Significant voxels/clusters
Two different views of the data

A “smallish” number of volumes

A large number of time series
FMRI Modelling and Statistics

- An example experiment
- Multiple regression (GLM)
- T and F Contrasts
- Null hypothesis testing
- The residuals
- Thresholding: multiple comparison correction
FMRI Modelling and Statistics

- An example experiment
- Multiple regression (GLM)
- T and F Contrasts
- Null hypothesis testing
- The residuals
- Thresholding: multiple comparison correction
An example experiment

An FMRI adaptation of a classical PET experiment

• Three types of events
• 1st type: Word Generation

![Diagram of an FMRI adaptation experiment](image)
An example experiment

An FMRI adaptation of a classical PET experiment

• Three types of events
• 1st type: Word Generation

Noun is presented

Jellyfish

Screen

Verb is generated
Catch

Healthy Volunteer

Scanner

Bed
An example experiment

An FMRI adaptation of a classical PET experiment

- Three types of events
- 1st type: Word Generation

Noun is presented

Burger

Screen

Verb is generated

Fry

Healthy Volunteer

Scanner

Bed
An example experiment
An FMRI adaptation of a classical PET experiment

- Three types of events
  - 1st type: Word Generation
  - 2nd type: Word Shadowing

Verb is presented

Verb is repeated

Swim

Screen

Scanner

Healthy Volunteer

Bed
An example experiment

An FMRI adaptation of a classical PET experiment

- Three types of events
- 1st type: Word Generation
- 2nd type: Word Shadowing
An example experiment

An FMRI adaptation of a classical PET experiment

• Three types of events
• 1st type: Word Generation
• 2nd type: Word Shadowing
• 3rd type: Null event

Crosshair is shown
Screen

Healthy Volunteer
Scanner
Bed
An example experiment
An FMRI adaptation of a classical PET experiment

- Three types of events
- 1st type: Word Generation
- 2nd type: Word Shadowing
- 3rd type: Null event

Crosshair is shown

Scanner

Bed

Healthy Volunteer

Screen
An example experiment

An FMRI adaptation of a classical PET experiment

• Three types of events
• 1st type: Word Generation
• 2nd type: Word Shadowing
• 3rd type: Null event
• 6 sec ISI, random order
An example experiment
An FMRI adaptation of a classical PET experiment

- Three types of events
- 1st type: Word Generation
- 2nd type: Word Shadowing
- 3rd type: Null event
- 6 sec ISI, random order
- For 24 events of each type
FMRI Modelling and Statistics

• An example experiment
• **Multiple regression (GLM)**
• T and F Contrasts
• Null hypothesis testing
• The residuals
• Thresholding: multiple comparison correction
Building a model

Our task is now to build a model for that experiment.

What is our predicted response to the word generation events?
Building a model

Our task is now to build a model for that experiment.

What is our predicted response to the word generation events?

Stick-function at each occurrence of a “generation event”

Well, hardly like this...
Building a model

Our task is now to build a model for that experiment.

What is our predicted response to the word generation events?

That looks better!
Building a model

Our task is now to build a model for that experiment

What is our predicted response to the word generation events?

\[ \text{HRF} \]

And this is the prediction for the whole time-series
Building a model

Our task is now to build a model for that experiment.

What is our predicted response to the word generation events?

So, if we spot a time-series like this
Building a model

Our task is now to build a model for that experiment

What is our predicted response to the word generation events?

And then check it against our prediction, we can conclude that this pixel is into word generation.
Building a model

Our task is now to build a model for that experiment

And we can do the same for the word shadowing events?

This time we used the onset times for the shadowing events to get the predicted brain response for those
Building a model

Our task is now to build a model for that experiment

And we can do the same for the word shadowing events?

And we can look for voxels that match that...
Formalising it: Multiple regression

\[ \text{Observed data} \approx \beta_1 + \beta_2 \cdot \text{Predicted responses} \]

Word Generation

Word Shadowing

Unknown “parameters”

Observed data

Predicted responses

“Regressors”
Slight detour: Making regressors

Event timings at “high” resolution

Convolve with HRF

Predictions at “high” resolution

Sub-sample at $T_R$ of experiment

Regressor
Estimation: Finding the “best” parameter values

- The estimation entails finding the parameter values such that the linear combination "best" fits the data.

Let’s try these parameter values

\[
\approx \beta_1 \cdot 0.5 + \beta_2 \cdot 0.5
\]
Estimation:
Finding the “best” parameter values

- The estimation entails finding the parameter values such that the linear combination ”best” fits the data.

Hmm, not a great fit
Estimation: Finding the “best” parameter values

- The estimation entails finding the parameter values such that the linear combination ”best” fits the data.

Oh dear, even worse
Estimation:
Finding the “best” parameter values

• The estimation entails finding the parameter values such that the linear combination ”best” fits the data.

\[ \approx \beta_1 \cdot 1.04 + \beta_2 \cdot (-0.10) \]

But that looks good
Estimation:
Finding the “best” parameter values

- The estimation entails finding the parameter values such that the linear combination ”best” fits the data:

\[ \beta_1 \cdot 1.10 + \beta_2 \cdot 1.02 \]

And different voxels yield different parameters.
Estimation: Finding the “best” parameter values

The estimation entails finding the parameter values such that the linear combination ”best” fits the data.

\[ \beta_1 \cdot -0.04 + \beta_2 \cdot -0.03 \]

And different voxels yield different parameters
One model to fit them all

\[
\begin{bmatrix}
1.10 \\
1.02
\end{bmatrix}
\]

\[
\begin{bmatrix}
1.04 \\
-0.10
\end{bmatrix}
\]

\[
\begin{bmatrix}
-0.04 \\
-0.03
\end{bmatrix}
\]
And we can also estimate the residual error

Difference between data and best fit: “Residual error”

Residual errors
And we can also estimate the residual error

\[
\begin{bmatrix}
1.10 \\
1.02
\end{bmatrix}
\]

\[
\begin{bmatrix}
1.04 \\
-.10
\end{bmatrix}
\]

\[
\begin{bmatrix}
-.04 \\
-.03
\end{bmatrix}
\]
Summary of what we learned so far

- The “Model” consists of a set of “regressors” i.e. tentative time series that we expect to see as a response to our stimulus
- The model typically consists of our stimulus functions convolved by the HRF
- The estimation entails finding the parameter values such that the linear combination of regressors ”best” fits the data
- Every voxel has its own unique parameter values, that is how a single model can fit so many different time series
- We can also get an estimate of the error through the “residuals”
General Linear Model (GLM)

- This is placed into the General Linear Model (GLM) framework

\[
y = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \beta + \epsilon
\]

- \( y \): Data from a voxel
- \( X \): Design Matrix
- \( \beta \): Regression parameters, Effect sizes
- \( \epsilon \): Gaussian noise (temporal autocorrelation)
“Demeaning” and the GLM

- The mean value is uninteresting in an FMRI session
- There are two equivalent options:
  1. remove the mean from the data and don’t model it
  2. put a term into the model to account for the mean

In FSL we use option #1 for first-level analyses and #2 for higher-level analyses

A consequence is that the baseline condition in first-level analysis is **NOT** explicitly modelled (in FSL)
FMRI Modelling and Statistics

• An example experiment
• Multiple regression (GLM)
• T and F Contrasts
• Null hypothesis testing
• The residuals
• Thresholding: multiple comparison correction
**t-contrasts**

- A contrast of parameter estimates (COPE) is a linear combination of PEs:

\[
[1 \ 0]: \quad \text{COPE} = 1 \times \widehat{\beta}_1 + 0 \times \widehat{\beta}_2 = \widehat{\beta}_1 \\
[1 \ -1]: \quad \text{COPE} = 1 \times \widehat{\beta}_1 + -1 \times \widehat{\beta}_2 = \widehat{\beta}_1 - \widehat{\beta}_2
\]

- Test null hypothesis that COPE=0

\[
t = \frac{\text{COPE}}{\text{std}(\text{COPE})}
\]

\[
t-\text{statistic:}
\]
t-contrasts

\[ t = \frac{COPE}{\text{std}(COPE)} \]

Depends on

The Model, the Contrast, and the Residual Error
t-contrasts

\[
\begin{bmatrix}
1 \\
0
\end{bmatrix}
\]

const \times

The Model & the Contrast and the Residual Error
**t-contrasts**

- \([1 \ 0]\) : EV1 only (i.e. Generation vs rest)
- \([0 \ 1]\) : EV2 only (i.e. Shadowing vs rest)
t-contrasts

Contrast weight vector: $[1 \ 0]$

Asks the question: Where do we need this regressor to model the data, i.e. what parts of the brain are used when seeing nouns and generating related verbs?
t-contrasts

Contrast weight vector: $[1 \ 0]$

COPE = $1 \times 1.04 + 0 \times -0.10 = 1.04$

COPE = $\beta_1$
t-contrasts

\[ t = \frac{COPE}{\text{std}(COPE)} \]

Model

\[ \begin{bmatrix} 1.04 \\ -0.10 \end{bmatrix} \]

\[ \beta_1 \]

\[ \beta_2 \]

\[ \sigma \]
**t-contrasts**

- \([1 0]\) : EV1 only (i.e. Generation vs rest)
- \([0 1]\) : EV2 only (i.e. Shadowing vs rest)
- \([1 1]\) : EV1 + EV2 (Mean activation)
**t-contrasts**

Contrast weight vector: $[1 \ 1]$

$$COPE = 1 \times 1.10 + 1 \times 1.02 = 2.12$$

$$COPE = \beta_1 + \beta_2$$
t-contrasts

\[ \begin{bmatrix}
  1.10 \\
  1.02
\end{bmatrix} \]

\[ t = \frac{\text{COPE}}{\text{std(COPE)}} \]

Model
t-contrasts

- $[1 \ 0]$: EV1 only (i.e. Generation vs rest)
- $[0 \ 1]$: EV2 only (i.e. Shadowing vs rest)
- $[1 \ 1]$: EV1 + EV2 (Mean activation)
- $[-1 \ 1]$: EV2 - EV1 (More activated by Shadowing than Generation)
- $[1 \ -1]$: EV1 - EV2 (More activated by Generation than Shadowing (t-tests are directional))
t-contrasts

Contrast weight vector: \([1 \ -1]\)

\[
\text{COPE} = 1 \times 1.04 - 1 \times -0.10 = 1.14
\]

\[
\text{COPE} = \beta_1 - \beta_2
\]
t-contrasts

$$t = \frac{\text{COPE}}{\text{std}(\text{COPE})} = \frac{\beta_1 - \beta_2}{\sigma}$$
t-contrasts

Why $[1 \ -1]$ instead of $[1 \ 0]$?

$[1 \ 0]$

$\begin{bmatrix} 1.10 \\ 1.02 \end{bmatrix}$

$\beta_1$
$\beta_2$

$[1 \ -1]$
F-contrasts

We have two conditions:
Word Generation and Shadowing

We want to know:
Is there an activation to any condition?

First we ask: Is there activation to Generation?

\[
\begin{bmatrix}
1 & 0
\end{bmatrix}
\]
F-contrasts

We have two conditions: Word Generation and Shadowing.

We want to know: Is there an activation to any condition?

Then we ask: Is there activation to Shadowing?

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]
F-contrasts

We have two conditions: Word Generation and Shadowing

We want to know: Is there an activation to any condition?

Then we add the OR

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]
F-contrasts

We have two conditions: Word Generation and Shadowing.

We want to know: Is there an activation to any condition?

Is there an activation to any condition?

Is equivalent to:

Does any regressor explain the variance in the data?

Then we add the OR matrix:

$$\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$
F-contrasts

Full Model

Data
F-contrasts

Full Model → Fit Model → Estimate Residuals → SS_E
F-contrasts

\[ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

Full Model

\[ \begin{bmatrix} \vdots \\ \vdots \end{bmatrix} \]

Reduced Model

Fit Model

Estimate Residuals

\[ F = \frac{SS_R - SS_E}{SS_E} \]

\[ \updownarrow \]

\[ SS_R \]

\[ SS_E \]
F-contrasts

\[ F = \frac{SS_R - SS_E}{SS_E} = \frac{\text{Full Model}}{\text{Reduced Model}} = \]
F-contrasts

- Two conditions: A, B
- Is any condition significant?

- Set of COPEs form an F-contrast
- Or: “Is there a significant amount of power in the data explained by the combination of the COPEs in the F-contrast?”
- F-contrast is F-distributed
FMRI Modelling and Statistics

- An example experiment
- Multiple regression (GLM)
- T and F Contrasts
- Null hypothesis testing
- The residuals
- Thresholding: multiple comparison correction
Null Hypothesis Testing

$t$-statistic: \[ t = \frac{\hat{\beta}}{\text{std}(\beta)} \]

Under null hypothesis, $\beta = 0$, $t$ is $t$-distributed

(what are the chances of that?)
Null Hypothesis Testing

Under null hypothesis, $\beta=0$,

\[ t = \frac{\hat{\beta}}{\text{std}(\beta)} \]

$P$-value:

\[ P - Value = p(t > t' | \beta = 0) \]

Small $P$-Value = null hypothesis unlikely
If $P$-Value < $P$-threshold then voxel is “active”
P-threshold corresponds to False Positive Rate (FPR)
• FEAT performs spatial inference on z statistic maps
• Therefore, we convert t statistics to z statistics by equating probabilities under the tails of the distributions (t' -> p -> z')
Summary of what we learned so far

- Once you have fitted your model and estimated your parameters you can ask directed questions of your data using contrasts.
- A contrast can be a $t$-contrast, which asks relatively specific question, or an $F$-contrast, which can ask more general questions.
- A $t$- or $F$-value specifies the “degree of surprise” of the data (given that the null-hypothesis is true).
- A $t$- or $F$-value can be transformed to a $z$-score with the same $p$-value (degree of surprise).
FEAT Schematic

Stimulus Details

HRF

EV

Design Matrix
FEAT Schematic

Preprocessed Data

Stimulus Details

HRF

EV

Design Matrix

PEs

PE covariances
FEAT Schematic

Preprocessed Data

Stimulus Details

HRF

EV

Design Matrix

PEs

PE covariances

COPEs

VARCOPEs

t-stats

<table>
<thead>
<tr>
<th></th>
<th>Generation</th>
<th>Shadowing</th>
<th>Mean</th>
<th>Shad &gt; Gen</th>
<th>Gen &gt; Shad</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C2</td>
<td>0</td>
<td>1</td>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>C3</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>C4</td>
<td>-1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>C5</td>
<td>1</td>
<td>-1</td>
<td></td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>
FEAT Schematic

Preprocessed Data

Stimulus Details

HRF

EV

Design Matrix

C1 Generation 1 0
C2 Shadowing 0 1
C3 Mean 1 1
C4 Shad > Gen -1 1
C5 Gen > Shad 1 -1

PEs PE covariances

COPEs VARCOPEs

t-stats

f-stats
FEAT Schematic

Preprocessed Data

Stimulus Details  HRF

EV

Design Matrix

| C1  | Generation | 1  | 0 |
| C2  | Shadowing  | 0  | 1 |
| C3  | Mean       | 1  | 1 |
| C4  | Shad > Gen | -1 | 1 |
| C5  | Gen > Shad | 1  | -1 |

PEs  PE covariances

COPEs  VARCOPEs

t-stats  z-stats

f-stats  zf-stats
FMRI Modelling and Statistics

- An example experiment
- Multiple regression (GLM)
- T and F Contrasts
- Null hypothesis testing
  - The residuals
- Thresholding: multiple comparison correction
Choosing High-Pass Filter Cut-off

- Can use the tool `cutoffcalc` to determine a good cut-off value

Remember that MJ mentioned highpass filtering?

Temporal Filtering: Highpass

- Removes low frequency signals, including linear trend
- Must choose cutoff frequency carefully (lower than frequencies of interest = longer period)
Choosing High-Pass Filter Cut-off

- Can use the tool `cutoffcalc` to determine a good cut-off value
  OR
- Set by hand, but make sure model is not badly affected

Example: Boxcar EV with period 100s

Negligible effect on EV, so use cut-off of 100s

Example: Boxcar with period 250s

Substantial effect on EV, so need longer cut-off

Example: Boxcar with period 250s

Negligible effect on EV, so use cut-off of 250s
Non-independent/Autocorrelation/Coloured FMRI noise

Uncorrected, this causes:

- biased stats (increased false positives)
- decreased sensitivity

FSL fixes it for you in FEAT!

Cannot use randomise (see later) because of autocorrelation
FMRI Modelling and Statistics

- An example experiment
- Multiple regression (GLM)
- T and F Contrasts
- Null hypothesis testing
- The residuals
- Thresholding: multiple comparison correction
What happens when we apply “standard” statistical testing to imaging data?

z-map where each voxel $\sim N$. Null-hypothesis true everywhere, i.e. NO ACTIVATIONS

z-map thresholded at 1.64

16 clusters
288 voxels
$\sim 5.5\%$ of the voxels

That’s a LOT of false positives
What we really want

Let’s say we perform a series of identical studies

Each z-map is the end result of a study

Let us further say that the null-hypothesis is true

We want to threshold the data so that only once in 20 studies do we find a voxel above this threshold

There will be a whole talk on how to find such a threshold
Summary

• The GLM is used to summarise data in a few parameters that are pertinent to the experiment.
• GLM predicts how BOLD activity might change as a result of the experiment.
• We can test for significant effects by using t or f contrasts on the GLM parameters
• When thresholding the number of false positives needs to be controlled across the entire brain

That's all folks