Jean-Baptiste Poline

Neurospin, I2BM, CEA
Saclay, France
Corrected $p$-values

Adjusted data

Random Field Theory

General Linear Model
Linear fit
$\rightarrow$ statistical image

Statistical Map
Uncorrected $p$-values

Corrected $p$-values

images

Design matrix

realignment & coregistration

smoothing

normalisation

Spatial filter

Anatomical Reference

Your question: a contrast

Statistical Map

Linear fit

Linear fit

$\rightarrow$ statistical image

Corrected $p$-values
Plan

- **REPEAT: model and fitting the data with a Linear Model**
- **Make sure we understand the testing procedures: t- and F-tests**
- **But what do we test exactly?**
- **Examples – almost real**
One voxel = One test (t, F, ...)

Temporal series
fMRI

amplitude

General Linear Model → fitting → statistical image

Statistical image (SPM)

voxel time course
Regression example...

\[ \text{voxel time series} = \beta_1 + \beta_2 + \text{box-car reference function} + \text{Mean value} \]

Fit the GLM
Regression example...

\[ \text{voxel time series} = \beta_1 + \beta_2 + \text{box-car reference function} + \text{Mean value} \]

\[ \beta_1 = 5 \]

\[ \beta_2 = 100 \]

Fit the GLM
...revisited: matrix form

\[ Y = \beta_1 \times f(t) + \beta_2 \times 1 + \varepsilon \]
Box car regression: design matrix...

\[ Y = X \times \beta + \varepsilon \]
Fact: model parameters depend on regressors scaling

Q: When do I care?

A: ONLY when comparing manually entered regressors (say you would like to compare two scores)

What if two conditions A and B are not of the same duration before convolution HRF?
What if we believe that there are drifts?
Add more reference functions / covariates ...

Discrete cosine transform basis functions
\[ \begin{align*}
Y &= X \times \beta + \varepsilon \\
&= \begin{pmatrix}
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4 \\
\vdots
\end{pmatrix}
\end{align*} \]
...design matrix

\[ Y = X \beta + \varepsilon \]

Y = data vector
X = design matrix
\( \beta \) = parameters = the betas (here: 1 to 9)
\( \varepsilon \) = error vector
Fitting the model = finding some **estimate** of the betas

How do we find the betas estimates? By minimizing the residual variance
Fitting the model = finding some **estimate** of the betas

\[
\begin{bmatrix}
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\end{bmatrix}
= \begin{bmatrix} 
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4 \\
\beta_5 \\
\beta_6 \\
\beta_7 \\
\vdots \\
\end{bmatrix}
+ \begin{bmatrix}
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\end{bmatrix}
\]

\[
Y = X \beta + \varepsilon
\]

**finding the betas = minimising the sum of square of the residuals**

\[
\|Y - X \beta\|^2 = \sum_i \left( y_i - \hat{y}_i \right)^2
\]

when \(\beta\) are estimated: let’s call them \(b\) (or \(\hat{\beta}\))

when \(\varepsilon\) is estimated : let’s call it \(e\)

estimated SD of \(\varepsilon\) : let’s call it \(s\)
Take home ...

- We put in our model regressors (or covariates) that represent how we think the signal is varying (of interest and of no interest alike)
  - WHICH ONE TO INCLUDE?
  - What if we have too many? Too few?

- Coefficients (= parameters) are estimated by minimizing the fluctuations, - variability – variance – of estimated noise – the residuals.

- Because the parameters depend on the scaling of the regressors included in the model, one should be careful in comparing manually entered regressors, or conditions of different durations.
Plan

- Make sure we all know about the estimation (fitting) part....
- Make sure we understand t and F tests
- But what do we test exactly?
- An example – almost real
A contrast = a weighted sum of parameters: \( c' \times b \)

\[
c' = 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0
\]

\( b_1 > 0 \) ?

Compute \( 1b_1 + 0b_2 + 0b_3 + 0b_4 + 0b_5 + \ldots = c'b \)

\[
c' = [1 \quad 0 \quad 0 \quad 0 \quad 0 \quad \ldots]
\]

divide by estimated standard deviation of \( b_1 \)

\[
T = \frac{c'b}{\sqrt{s^2c'(X'X)^{-1}c}}
\]

\( \text{contrast of estimated parameters} \)

\( \text{variance estimate} \)

\( \text{SPM} \{t\} \)
From one time series to an image

\[ Y : \text{data} = X \ast B + E \]

\[ \text{var}(E) = s^2 \]

\[ c' b \]

\[ T = \frac{c' b}{\sqrt{s^2 c'(X'X)^{-1} c}} \]

\[ c' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \]

\[ \text{spm}_\text{ResMS} \]

\[ \text{spm}_\text{con} \]

\[ \text{spm}_t \]
F-test: a reduced model

$H_0$: True model is $X_0$

$H_0$: $\beta_1 = 0$

$F \sim \frac{(S_0^2 - S^2)}{S^2}$

T values become F values. $F = T^2$

Both “activation” and “deactivations” are tested. Voxel wise p-values are halved.
Tests multiple linear hypotheses: Does $X_1$ model anything?

$H_0$: True (reduced) model is $X_0$

This (full) model? Or this one?

$F = \frac{(S_0^2 - S^2)}{S^2}$
F-test: a reduced model or ... multi-dimensional contrasts?

tests multiple linear hypotheses. Ex: does drift functions model anything?

\[ H_0: \text{True model is } X_0 \]

\[ H_0: \beta_{3-9} = (0 \ 0 \ 0 \ 0 \ ...) \]

This (full) model? Or this one?
<table>
<thead>
<tr>
<th>Convolution model</th>
<th>Design and contrast</th>
<th>SPM(t) or SPM(F)</th>
<th>Fitted and adjusted data</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="convolution_model.png" alt="Convolution model graph" /></td>
<td><img src="design_contrast.png" alt="Design and contrast" /></td>
<td><img src="spm_images.png" alt="SPM(t) or SPM(F) images" /></td>
<td><img src="fitted_data.png" alt="Fitted and adjusted data" /></td>
</tr>
</tbody>
</table>

For the convolution model, the graph shows the response over time, with peaks indicating the convolution of the stimulus and the system's impulse response. The design contrast includes a bar graph and a table showing the contrast values. SPM(t) and SPM(F) images illustrate the statistical significance of the contrasts, with red and blue highlighting the regions of interest. The fitted and adjusted data plot shows the relationship between time and response, with fitted lines and error bars for the adjusted data.
T and F test: take home ...

- T tests are simple combinations of the betas; they are either positive or negative (b1 – b2 is different from b2 – b1)

- F tests can be viewed as testing for the additional variance explained by a larger model wrt a simpler model, or

- F tests the sum of the squares of one or several combinations of the betas

- in testing “single contrast” with an F test, for ex. b1 – b2, the result will be the same as testing b2 – b1. It will be exactly the square of the t-test, testing for both positive and negative effects.
Plan

- Make sure we all know about the estimation (fitting) part....
- Make sure we understand t and F tests
- But what do we test exactly? Correlation between regressors
- An example – almost real
No correlation between green red and yellow
correlated regressors, for example
green: subject age
yellow: subject score
Testing for the red correlated contrasts
Very correlated regressors?

Dangerous!

Testing for the green
Testing for the green and yellow

If significant? Could be G or Y!
Completely correlated regressors? Impossible to test! (not estimable)

Testing for the green
Testing for first regressor: $T_{\text{max}} = 9.8$
Including the movement parameters in the model

Testing for first regressor: activation is gone! Right or Wrong?
Implicit or explicit (⊥) decorrelation (or orthogonalisation)

This generalises when testing several regressors (F tests)

cf Andrade et al., NeuroImage, 1999
Correlation between regressors: take home ...

- *Do we care about correlation in the design?*
  Yes, always

- *Start with the experimental design: conditions should be as uncorrelated as possible*

- *use F tests to test for the overall variance explained by several (correlated) regressors*
Plan

- Make sure we all know about the estimation (fitting) part ....
- Make sure we understand t and F tests
- But what do we test exactly? Correlation between regressors
- An example – almost real
Experimental Design → Design Matrix

Factorial design with 2 factors: modality and category
- 2 levels for modality (e.g., Visual/Auditory)
- 3 levels for category (e.g., 3 categories of words)
Asking ourselves some questions ...

- Design Matrix not orthogonal
- Many contrasts are non estimable
- Interactions MxC are not modelled

Test C1 > C2 : $c = [0 0 1 -1 0 0]$
Test V > A : $c = [1 -1 0 0 0 0]$
Test C1,C2,C3 ? (F) $c = [0 0 1 0 0 0]$
Test the interaction MxC ?

\[
\begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\]
Modelling the interactions
Test \( C_1 > C_2 \) : \( c = [1 \ 1 \ -1 \ -1 \ 0 \ 0 \ 0] \)

Test \( V > A \) : \( c = [1 \ -1 \ 1 \ -1 \ 1 \ -1 \ 0] \)

Test the category effect :
\[
\begin{bmatrix}
1 & 1 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & -1 & -1 & 0 \\
1 & 1 & 0 & 0 & -1 & -1 & 0
\end{bmatrix}
\]

Test the interaction \( M \times C \) :
\[
\begin{bmatrix}
1 & 1 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & -1 & -1 & 0 \\
1 & 1 & 0 & 0 & -1 & -1 & 0
\end{bmatrix}
\]

- Design Matrix orthogonal
- All contrasts are estimable
- Interactions \( M \times C \) modelled
- If no interaction ... ? Model is too “big”!
Test \( C_1 > C_2 \) ?
Test \( C_1 \) different from \( C_2 \) ?
from
\[
c = \begin{bmatrix}
1 & 1 & -1 & -1 & 0 & 0 & 0
\end{bmatrix}
\]
to
\[
c = \begin{bmatrix}
1 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
becomes an F test!

What if we use only:
\[
c = \begin{bmatrix}
1 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
OK only if the regressors coding for the delay are all equal
Toy example: take home ...

- use F tests when
  - Test for $>0$ and $<0$ effects
  - Test for more than 2 levels in factorial designs
  - Conditions are modelled with more than one regressor

- Check post hoc
Thank you for your attention!

jbpoline@cea.fr
\[
\begin{align*}
\text{Design Matrix} & \quad X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \\
\text{Parameters} & \quad \begin{aligned}
\hat{\beta}_1 &= \bar{y}_1 \\
\hat{\beta}_2 &= \bar{y}_2
\end{aligned} \\
\text{Contrasts} & \quad \begin{aligned}
(1, 0).\hat{\beta} &= \bar{y}_1 \\
(0, 1).\hat{\beta} &= \bar{y}_2 \\
(1, -1).\hat{\beta} &= \bar{y}_1 - \bar{y}_2 \\
(.5, .5).\hat{\beta} &= \text{mean}(\bar{y}_1, \bar{y}_2)
\end{aligned}
\end{align*}
\]

\[
P_X Y = X \beta
\]
Projector onto $X$

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$P_X Y = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} . Y = X \beta = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_2 \end{bmatrix}$$
\( (2) \quad X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \)

\[
\begin{align*}
\hat{\beta}_1 + \hat{\beta}_2 &= \bar{y}_1 \\
\hat{\beta}_2 &= \bar{y}_2
\end{align*}
\]

\[
(1, 1).\hat{\beta} = \bar{y}_1 \\
(0, 1).\hat{\beta} = \bar{y}_2 \\
(1, 0).\hat{\beta} = \bar{y}_1 - \bar{y}_2 \\
(.5, 1).\hat{\beta} = \text{mean}(\bar{y}_1, \bar{y}_2)
\]

\( (3) \quad X = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \)

\[
\begin{align*}
\hat{\beta}_1 + \hat{\beta}_3 &= \bar{y}_1 \\
\hat{\beta}_2 + \hat{\beta}_3 &= \bar{y}_2
\end{align*}
\]

\[
(1, 0, 1).\hat{\beta} = \bar{y}_1 \\
(0, 1, 1).\hat{\beta} = \bar{y}_2 \\
(1, -1, 0).\hat{\beta} = \bar{y}_1 - \bar{y}_2 \\
(.5, .5, 1).\hat{\beta} = \text{mean}(\bar{y}_1, \bar{y}_2)
\]
Main Effects and Interaction:
1. Main effect: 2 (A)
2. Main effect: 3 (B)
3. Interaction: 2 3 (A × B)

Contrast Weights
1. Main effect of A: 1 -1 0 0 0 ones(1,3)/3 -ones(1,3)/3
2. Main effect of B: 0 0 -1 0 1 [-1 0 1]*[1/2] [-1 0 1]*[1/2]
3. Interaction A × B: 0 0 0 0 0 -1 0 1 1 0 -1
4. Test for a single regressor in main effect of A (e.g. A1)
   1 0 ones(1,3)/3 ones(1,3)/3 zeros(1,3)
5. Test for a single regressor in main effect of B (e.g. B2)
   0.5 0.5 0 1 0 0 0.5 0 0 0.5 0
6. Test for a single regressor in interaction A x B (e.g. A1B3)
   1 0 0 0 1 0 0 1 0 0 0
\[ y_1 = \beta_1 + \beta_3 + \beta_6 + \varepsilon^{(1)} \]
\[ y_2 = \beta_1 + \beta_4 + \beta_7 + \varepsilon^{(2)} \]
\[ y_3 = \beta_1 + \beta_5 + \beta_8 + \varepsilon^{(3)} \]
\[ y_4 = \beta_2 + \beta_3 + \beta_9 + \varepsilon^{(4)} \]
\[ y_5 = \beta_2 + \beta_4 + \beta_{10} + \varepsilon^{(5)} \]
\[ y_6 = \beta_2 + \beta_5 + \beta_{11} + \varepsilon^{(6)} \]
\[ y_1 + y_2 + y_3 = 3\beta_1 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 + \beta_8 + \varepsilon^{1+2+3} \]
\[ y_4 + y_5 + y_6 = 3\beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_9 + \beta_{10} + \beta_{11} + \varepsilon^{4+5+6} \]
Estimation $[Y, X] [b, s]$

$Y = X \beta + \varepsilon$

$\varepsilon \sim \sigma^2 \mathcal{N}(0, I)$ (for one position)

$b = (X'X)^+ X'Y$

$e = Y - Xb$

$s^2 = (e' e / (n - p))$

Test $[b, s^2, c] [c'b, t]$

$\text{Var}(c'b) = s^2 c' (X'X)^+ c$

$t = c'b / \sqrt{s^2 c' (X'X)^+ c}$

$\text{under the null hypothesis } H_0 : t \sim \text{Student-}t(\ df)$

$df = n-p$
How is this computed? (F-test)

Estimation \([Y, X] [b, s]\)

\[ Y = X \beta + \varepsilon \]
\[ Y = X_0 \beta_0 + \varepsilon_0 \]

\[ \varepsilon \sim N(0, \sigma^2 I) \]
\[ \varepsilon_0 \sim N(0, \sigma_0^2 I) \]

\[ X_0 : X \text{ Reduced} \]

Test \([b, s, c] [\text{ess, } F]\)

\[ F \sim \frac{(s_0 - s)}{s^2} \]

\[ \rightarrow \text{image of } F \begin{cases} \text{spm_ess} \text{???} \end{cases} \]
\[ \rightarrow \text{image of } F \begin{cases} \text{spm_F} \text{???} \end{cases} \]

under the null hypothesis: \( F \sim F(p - p_0, n-p) \)