Robust Linear Modelling of EEG data: the LIMO EEG plug-in

Arnaud Delorme

with Cyril Pernet, PhD, Edinburgh Imaging & Centre for Clinical Brain Sciences
Hierarchical Linear Model

1st level analysis:
GLM: $Y = X\beta + \epsilon$
$\rightarrow 1 \beta$ per column of $X$
(= within subject effects)

2nd level analysis:
Robust stats (Yuen t-tests, robust GLM,
robust Hotelling $T^2$)

Subject 1

Subject 2

Subject 3

Subject 4

Subject N

T-test
Regression
N-way ANOVA
N-way ANCOVA
Rep Measure ANOVA
Linear Modeling of EEG data

Diagram showing epoched data for different trials, with a design matrix for 1st level analysis.
Linear Modeling of EEG data
Random Effect Model

Model the data with fixed effects (the experimental conditions) and a random effect (subjects are allowed to have different overall values – considering subjects as a random variable)

Example: present stimuli from intensity -5 units to +5 units around the subject perceptual threshold and measure RT

→ Plot the data per intensity
Fixed effect = average across subjects → negative correlation?

$R^2 = 0.16$
$P < 0.03$
Random Effect Model

1\textsuperscript{st} level

2\textsuperscript{nd} level

Slope: [3.43, 3.51, 3.39, 3.41, 3.62, 3.55, ... ]

\( p < 10^{-12} \)
A regression is a linear model

**Varying factor:** Contrast of image

**Outcome:** Reaction time

A regression is a linear model

- We have an experimental measure $x$ (e.g. contrast)
A regression is a linear model

• We have an experimental measure $x$ (e.g. contrast)

• We then do the experiment and collect data $RT$ (e.g. reaction time)

$y = 2.7x + 23.6$
A regression is a linear model

- We have an experimental measure $x$ (e.g. contrast)
- We then do the experiment and collect data $RT$ (e.g. reaction time)
- Model: $RT = \beta_0 + x\beta_1 + \epsilon$
- Do some maths / run a software to find $\beta_1$ and $\beta_0$
- $RT^\wedge = 23.6 + 2.7x$
A regression is a linear model

For each trial

\[
RT_1 = \beta_0 + 10\beta_1 + \varepsilon_1 \\
RT_2 = \beta_0 + 5\beta_1 + \varepsilon_2 \\
RT_3 = \beta_0 + 7\beta_1 + \varepsilon_3 \\
\ldots
\]
A regression is a linear model

For each trial

\[
\begin{align*}
\text{RT}_1 &= \beta_0 + 10\beta_1 + \varepsilon_1 \\
\text{RT}_2 &= \beta_0 + 5\beta_1 + \varepsilon_2 \\
\text{RT}_3 &= \beta_0 + 7\beta_1 + \varepsilon_3 \\
\ldots
\end{align*}
\]

To test for significance compare the original regression model
\[
\text{RT}_i = \beta_0 + c_i\beta_1 + \varepsilon_i
\]
with the simplified model \( \text{RT}_i = \beta_0 + \varepsilon_i \)

Contrast level

Compare these errors
An ANOVA is a linear model

Varying factor: Type of image

Outcome: Reaction time (go/no-go)

\[ RT_{i,j} = \beta_0 + \beta_i + \epsilon_{i,j} \]

that is to say the data (e.g. RT) = a constant term (grand mean \( \beta_0 \)) + the effect of a treatment (\( \beta_1 \) for fishes 1 and \( \beta_2, \beta_3 \) for birds and reptiles) and the error term (\( \epsilon_{i,j} \))
\[ RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j} \]

that is to say the data (e.g. RT) = a constant term (grand mean \( \beta_0 \)) + the effect of a treatment (\( \beta_1 \) for fishes 1 and \( \beta_2, \beta_3 \) for birds and reptiles) and the error term (\( \varepsilon_{i,j} \))

For trial 4 (for example first trial of birds) we have

\[ RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,1} \]
\[ RT_{ij} = \beta_0 + \beta_i + \varepsilon_{ij} \]

that is to say the data (e.g. RT) = a constant term (grand mean \( \beta_0 \)) + the effect of a treatment (\( \beta_1 \) for fishes 1 and \( \beta_2, \beta_3 \) for birds and reptiles) and the error term (\( \varepsilon_{ij} \))

For trial 4 (for example first trial of birds) we have

\[ RT_{2,1} = \beta_0 + 0^*\beta_1 + 1^*\beta_2 + 0^*\beta_3 + \varepsilon_{2,1} \]

For trial 13 (for example second trial of birds) we have

\[ RT_{2,2} = \beta_0 + 0^*\beta_1 + 1^*\beta_2 + 0^*\beta_3 + \varepsilon_{2,2} \]
\[ RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j} \]

that is to say the data (e.g. RT) = a constant term (grand mean $\beta_0$) + the effect of a treatment ($\beta_1$ for fishes 1 and $\beta_2$, $\beta_3$ for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

For trial 4 (for example first trial of birds) we have

\[ RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,1} \]

For trial 13 (for example second trial of birds) we have

\[ RT_{2,2} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,2} \]

Statistics: if there is an effect of treatment then error of the simplified model $RT_{i,j} = \beta_0 + \varepsilon_{i,j}$ should be lower than the original model $RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$

Compare these errors

This is a GLM that is also strictly equivalent to running an ANOVA
The GLM can do both a Regression and an ANOVA

**Varying factor:** Type of image **AND** contrast

**Outcome:** Reaction time (go/no-go)

\[
RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + 0*\beta_4 + c_{2,1}*\beta_4 + \epsilon_{2,1}
\]

For example, for trial (first bird with contrast \( c_{2,1} \)) we have

Categorical var. ANOVA

Continuous var. REGRESSION
The design matrix

\[ y(1..3) = 1x\beta_1 + 0x\beta_2 + 0x\beta_3 + 0x\beta_4 + c + \text{error} \]
\[ y(4..6) = 0x\beta_1 + 1x\beta_2 + 0x\beta_3 + 0x\beta_4 + c + \text{error} \]
\[ y(7..9) = 0x\beta_1 + 0x\beta_2 + 1x\beta_3 + 0x\beta_4 + c + \text{error} \]
\[ y(10..12) = 0x\beta_1 + 0x\beta_2 + 0x\beta_3 + 1x\beta_4 + c + \text{error} \]

...
Design considerations

Illustration with a set of studies looking at the effect of stimulus amplitude and phase information

Factorial Designs: N*N*N*...

Categorical designs: Group level analyses of course but also Individual analyses with bootstrap

B1→B3 represents type of image, 2nd level ANOVA across subjects
B4→B6 represents amplitude, 2nd level ANOVA across subjects
Interaction between phase and amplitude can be assessed using 2-way ANOVA
Factorial Designs: N*N*N*...

Categorical designs: Group level analyses of course but also Individual analyses with bootstrap

Interaction between phase and amplitude can be assessed using 1-way ANOVA on B7 to B12. There is no interaction left between B1-B3 and B4-B6

Continuous designs

Note: just face and houses (no noise here)
What have we done: results

- Image all (R2, condition, covariate)
- Course plots, topoplots
Review group level results
Design questions!

• Let’s think how to analyse your data!
• Nb of conditions / covariates
• contrasts
• 1\textsuperscript{st} level covariates
• 2\textsuperscript{nd} level covariates