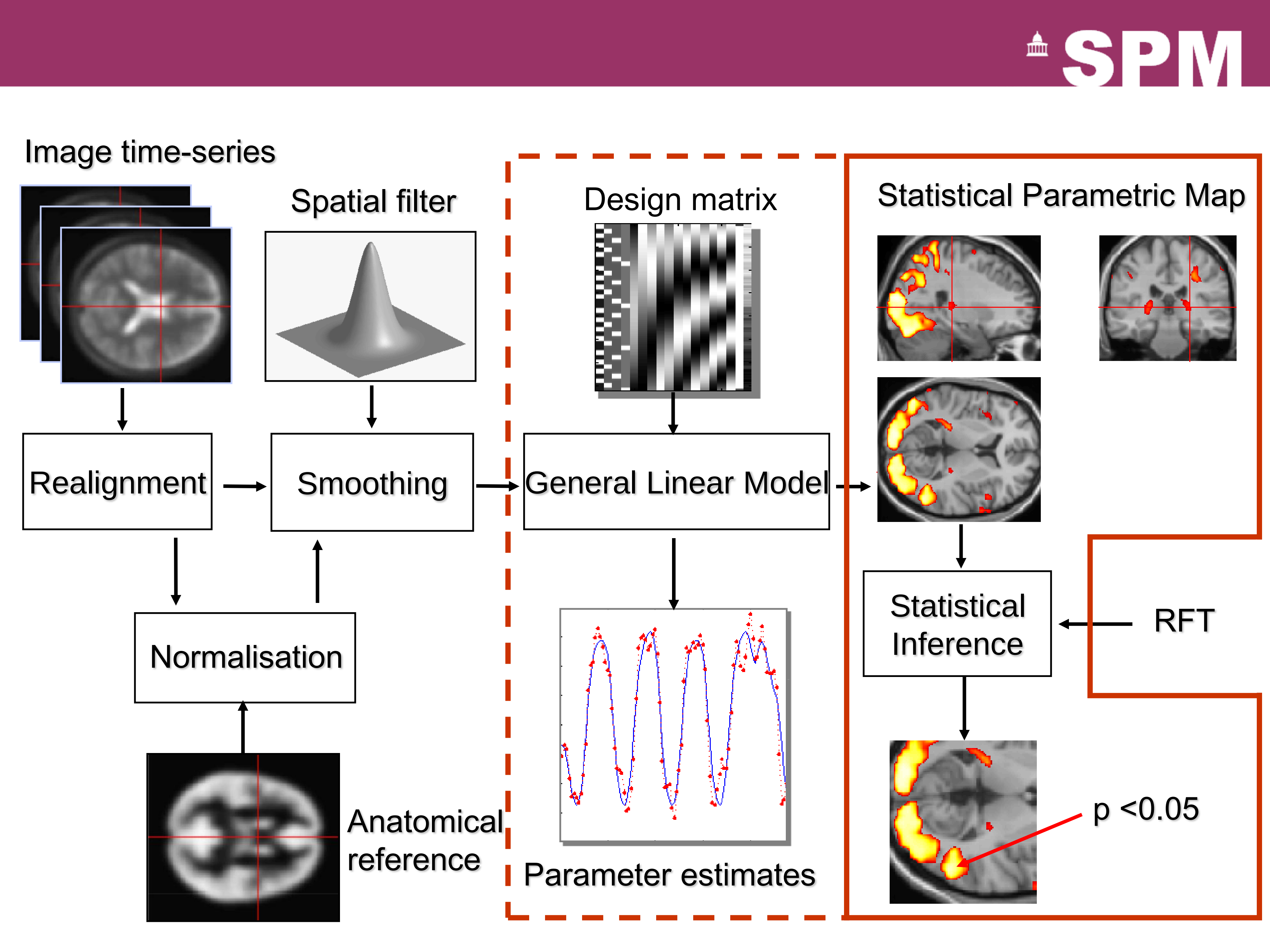
Contrasts & Statistical Inference

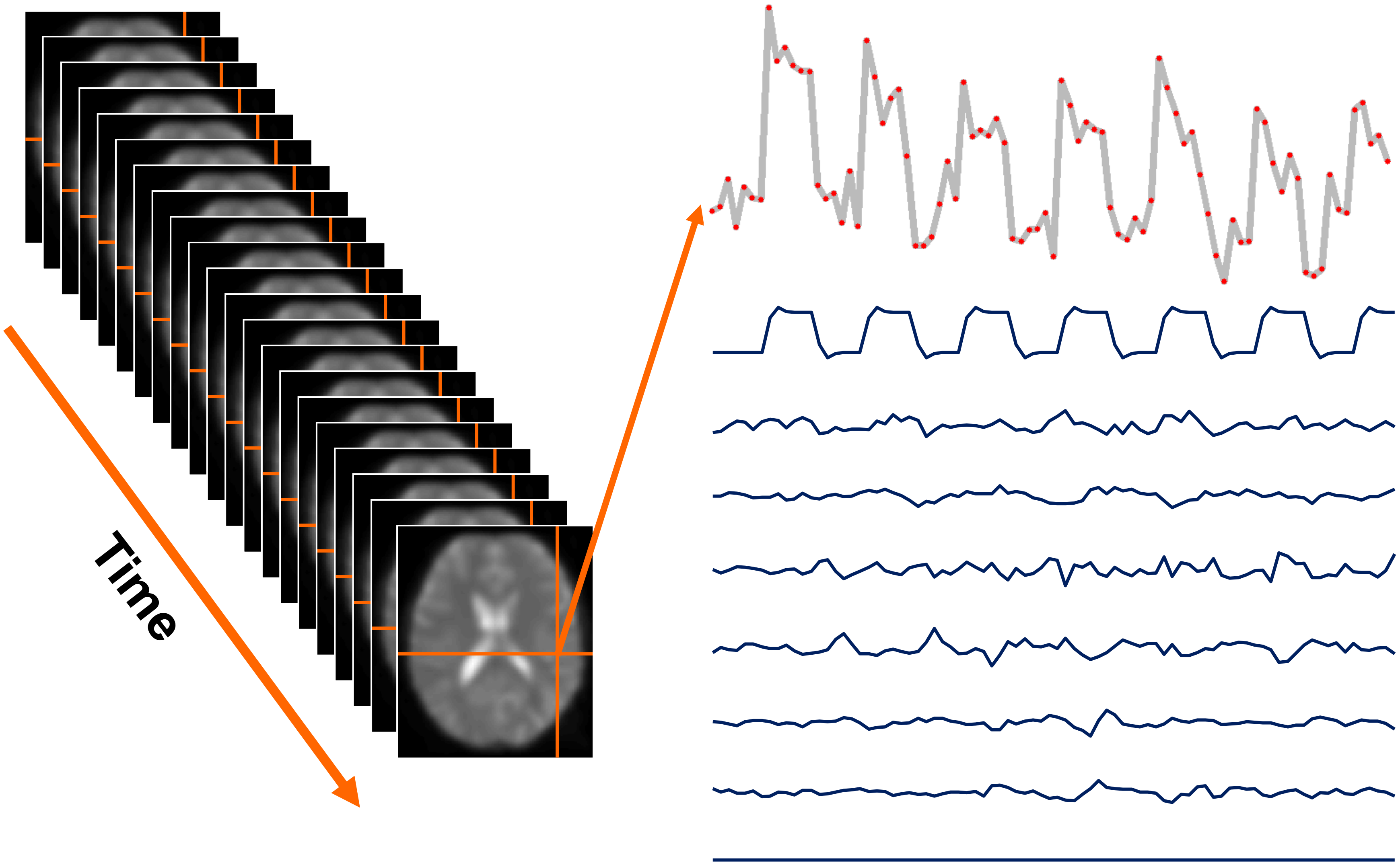
Christophe Phillips
Slides: Guillaume Flandin

GIGA – Cyclotron Research Centre *in vivo* imaging
University of Liège

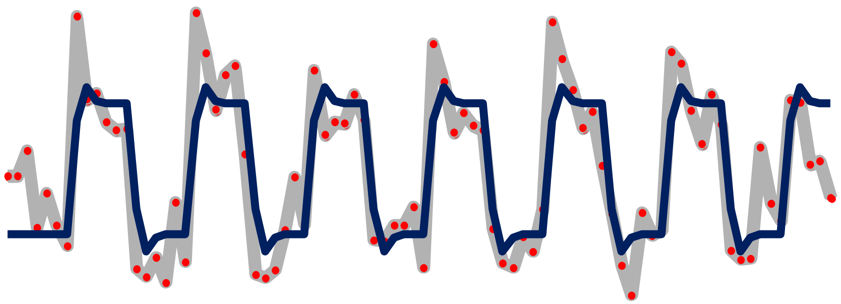
SPM Course
Woluwe, 28 November 2019


SPM
Image time-series
Spatial filter
Design matrix
Statistical Parametric Map
Realignment
Smoothing
General Linear Model
Statistical Inference
RFT
Normalisation
Anatomical reference
Parameter estimates
p <0.05

A mass-univariate approach



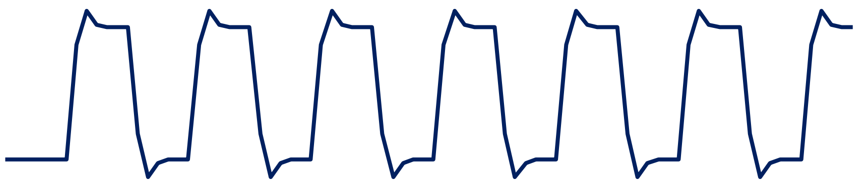
Estimation of the parameters



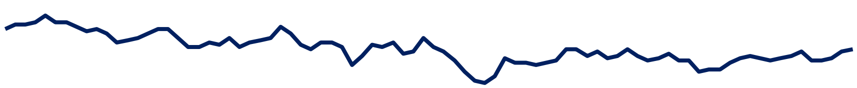
i.i.d. assumptions: $\varepsilon \sim N(0, \sigma^2 I)$

OLS estimates: $\hat{\beta} = (X^T X)^{-1} X^T y$

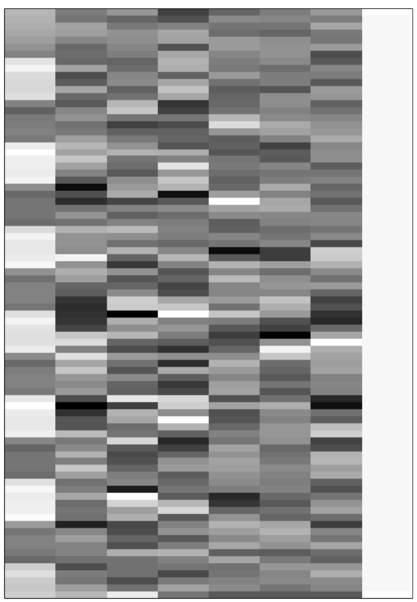
$\hat{\beta}_1 = 3.9831$



$\hat{\beta}_{2-7} = \{0.6871, 1.9598, 1.3902, 166.1007, 76.4770, -64.8189\}$



$\hat{\beta}_8 = 131.0040$

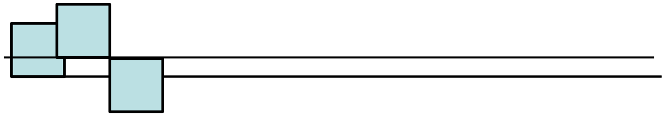
$y =$  β $+\varepsilon$

$\hat{\varepsilon} =$

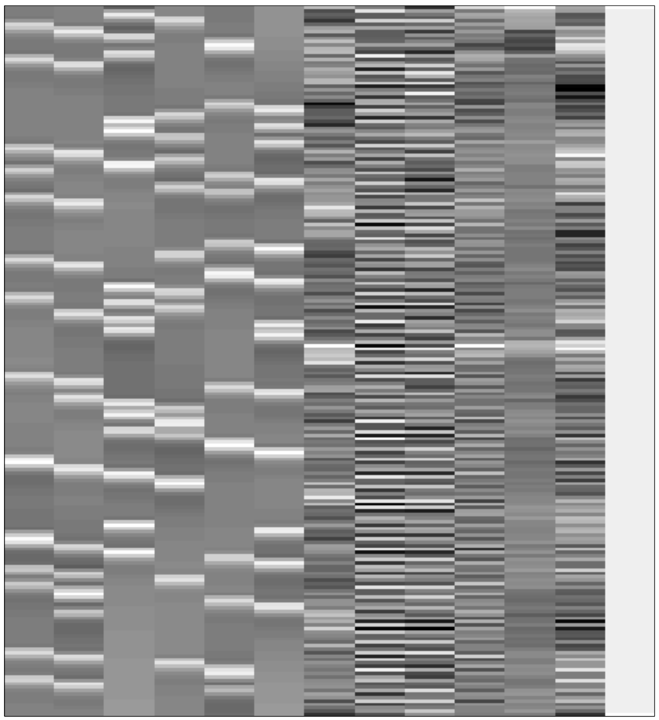
$\boxed{\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})}$

$\boxed{\hat{\sigma}^2 = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{N-p}}$

Contrasts



[0 1 -1 0 0 0 0 0 0 0 0 0 0 0]



- A contrast selects a specific effect of interest.
  - ⇨ A contrast c is a vector of length p.
  - ⇨ $c^T\beta$ is a linear combination of regression coefficients β.

$$c = [1\ 0\ 0\ 0\ \ldots]^T$$

$$c^T\beta = \mathbf{1} \times \beta_1 + \mathbf{0} \times \beta_2 + \mathbf{0} \times \beta_3 + \mathbf{0} \times \beta_4 + \cdots$$
$$= \boldsymbol{\beta_1}$$

$$c = [0\ 1\ -1\ 0\ \ldots]^T$$

$$c^T\beta = \mathbf{0} \times \beta_1 + \mathbf{1} \times \beta_2 + \mathbf{-1} \times \beta_3 + \mathbf{0} \times \beta_4 + \cdots$$
$$= \boldsymbol{\beta_2 - \beta_3}$$

$$\boxed{c^T\hat{\beta} \sim N(c^T\beta, \sigma^2 c^T (X^T X)^{-1} c)}$$

Hypothesis Testing

To test a hypothesis, we construct “test statistics”.

- **Null Hypothesis H_0**

  Typically what we want to disprove (no effect).

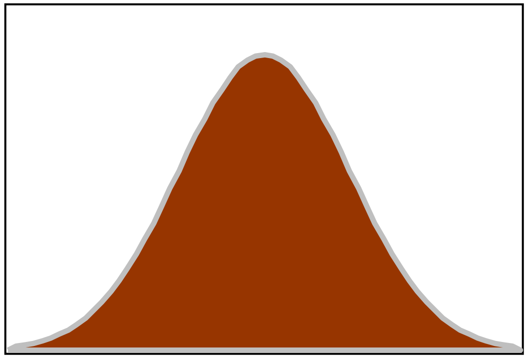
  ⇨ The Alternative Hypothesis H_A expresses outcome of interest.

- **Test Statistic T**

  The test statistic summarises evidence about H_0.

  Typically, test statistic is small in magnitude when the hypothesis H_0 is true and large when false.

  ⇨ We need to know the distribution of T under the null hypothesis.



Null Distribution of T

Hypothesis Testing

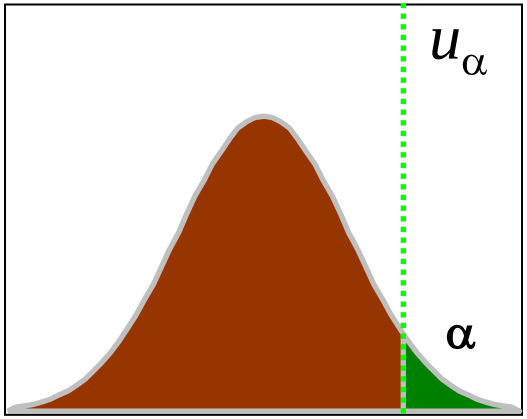
- ***Significance level α:***

  Acceptable *false positive rate α*.

  ⇨ threshold u_α

  Threshold u_α controls the false positive rate

$$\alpha = p(T > u_\alpha \mid H_0)$$



Null Distribution of T

- **Conclusion about the hypothesis:**

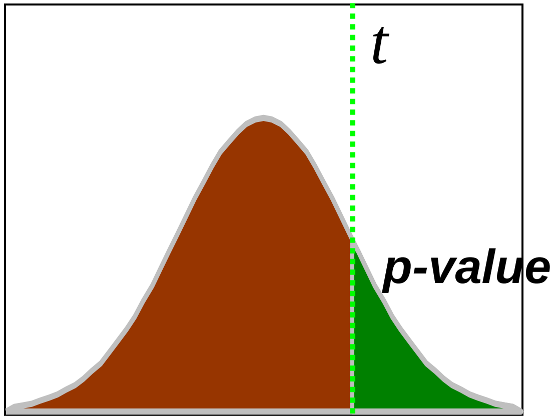
  We reject the null hypothesis in favour of the alternative hypothesis if $t > u_\alpha$

- ***p-value*:**

  A *p-value* summarises evidence against H_0.

  This is the chance of observing value more extreme than *t* under the null hypothesis.

$$\boxed{p(T > t|H_0)}$$

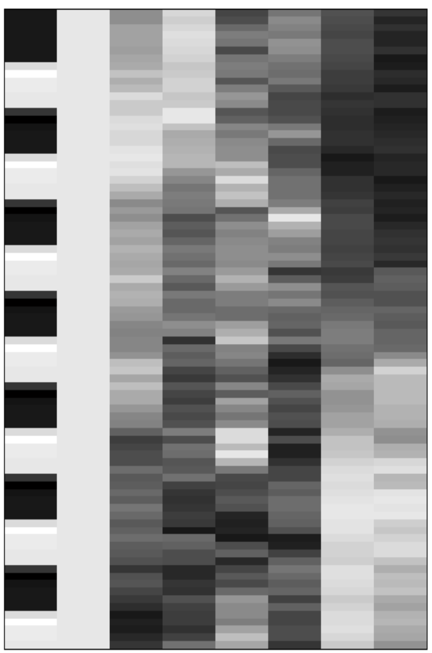


Null Distribution of T

***T*-test** - one dimensional contrasts – SPM{*t*}

c^T **= 1 0 0 0 0 0 0 0**

$\beta_1\ \beta_2\ \beta_3\ \beta_4\ \beta_5$ ...



Question: box-car amplitude > 0 ?

=

$\beta_1 = c^T\beta > 0$?

Null hypothesis: $H_0: c^T\beta = 0$

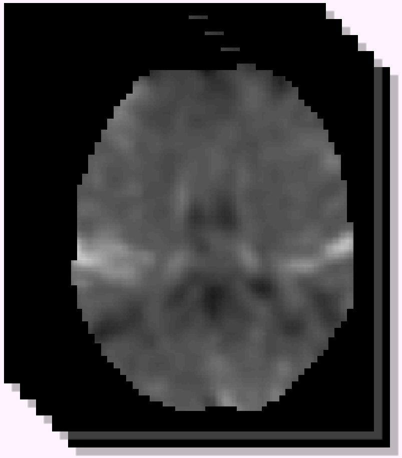
Test statistic:

$$T = \frac{\textbf{contrast of estimated parameters}}{\sqrt{\textbf{variance estimate}}}$$

$$T = \frac{c^T\hat{\beta}}{\sqrt{\mathrm{var}(c^T\hat{\beta})}} = \frac{c^T\hat{\beta}}{\sqrt{\hat{\sigma}^2 c^T\left(X^T X\right)^{-1} c}} \sim t_{N-p}$$

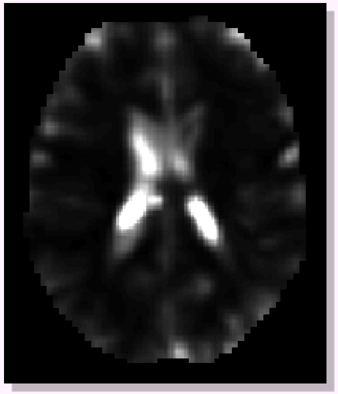
T-contrast in SPM

- For a given contrast *c*:



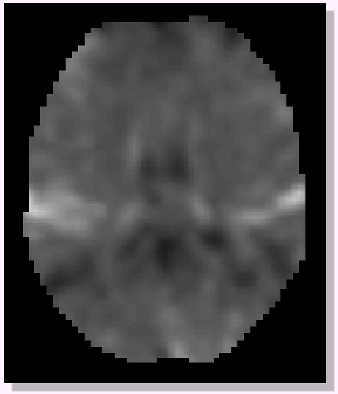
beta_???? images

$$\hat{\beta} = (X^T X)^{-1} X^T y$$



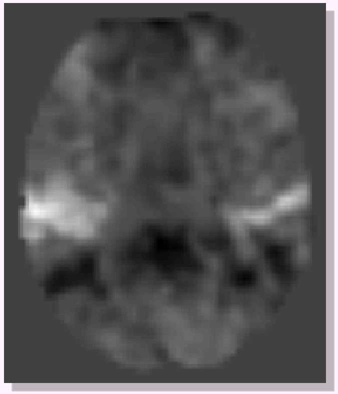
ResMS image

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{N - p}$$



con_???? image

$$c^T \hat{\beta}$$



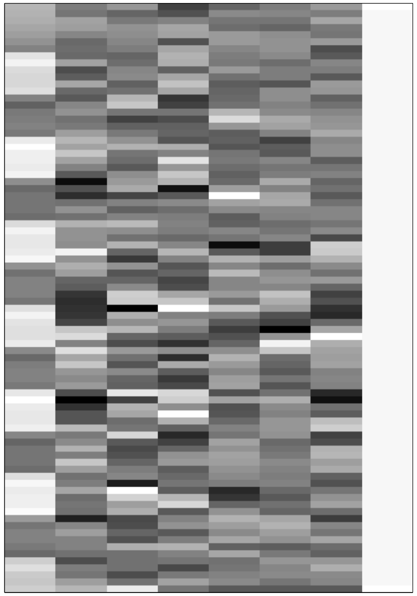
spmT_???? image

SPM{*t*}

T-test: a simple example

- Passive word listening versus rest

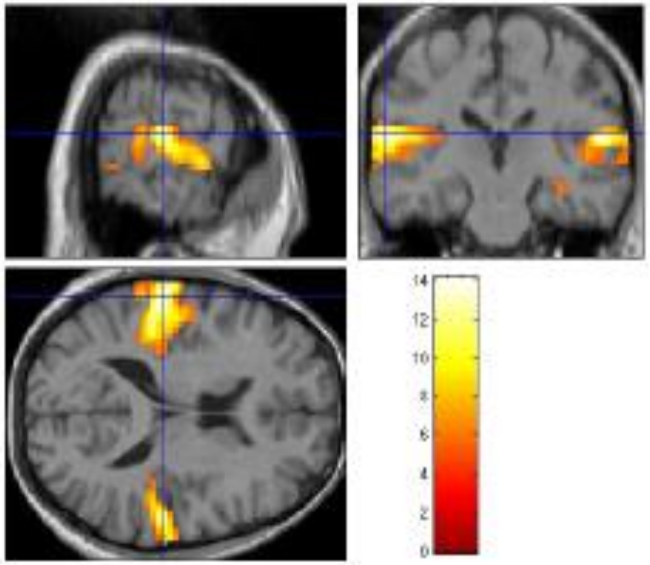
$c^T = [\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$



Q: activation during listening ?

Null hypothesis: $\beta_1 = 0$

$$t = \frac{c^T \hat{\beta}}{\sqrt{\mathrm{var}(c^T \hat{\beta})}}$$



SPMresults:

Height threshold T = 3.2057 {p<0.001}

voxel-level

T	($Z_{\equiv}$)	$p_{\text{uncorrected}}$	mm	mm	mm
13.94	**Inf**	**0.000**	**-63**	**-27**	**15**
12.04	Inf	0.000	-48	-33	12
11.82	Inf	0.000	-66	-21	6
13.72	**Inf**	**0.000**	**57**	**-21**	**12**
12.29	Inf	0.000	63	-12	-3
9.89	7.83	0.000	57	-39	6
7.39	**6.36**	**0.000**	**36**	**-30**	**-15**
6.84	**5.99**	**0.000**	**51**	**0**	**48**
6.36	**5.65**	**0.000**	**-63**	**-54**	**-3**
6.19	**5.53**	**0.000**	**-30**	**-33**	**-18**
5.96	**5.36**	**0.000**	**36**	**-27**	**9**
5.84	**5.27**	**0.000**	**-45**	**42**	**9**
5.44	**4.97**	**0.000**	**48**	**27**	**24**
5.32	**4.87**	**0.000**	**36**	**-27**	**42**

T-test: summary

- *T*-test is a *signal-to-noise* measure (ratio of estimate to standard deviation of estimate).

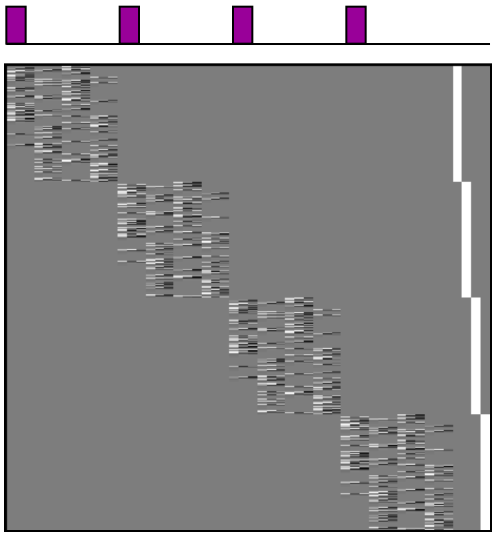
- Alternative hypothesis:

$$H_0:\quad c^T\beta = 0 \qquad \text{vs} \qquad H_A:\quad c^T\beta > 0$$

- *T*-contrasts are simple combinations of the betas; the T-statistic does not depend on the scaling of the regressors or the scaling of the contrast.

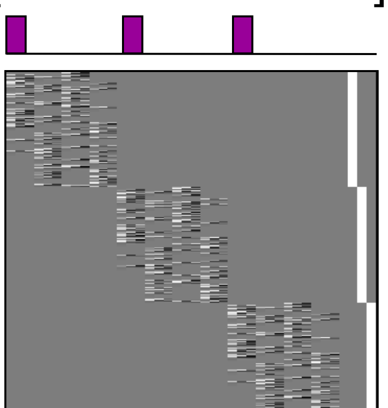
Scaling issue

[1 1 1 1] **/ 4**



Subject 1

[1 1 1] **/ 3**



Subject 5

$$T = \frac{c^T \hat{\beta}}{\sqrt{\mathrm{var}(c^T \hat{\beta})}} = \frac{c^T \hat{\beta}}{\sqrt{\hat{\sigma}^2 c^T \left(X^T X\right)^{-1} c}}$$

- The *T*-statistic does not depend on the scaling of the regressors.
- The *T*-statistic does not depend on the scaling of the contrast.
- Contrast $c^T \hat{\beta}$ depends on scaling.

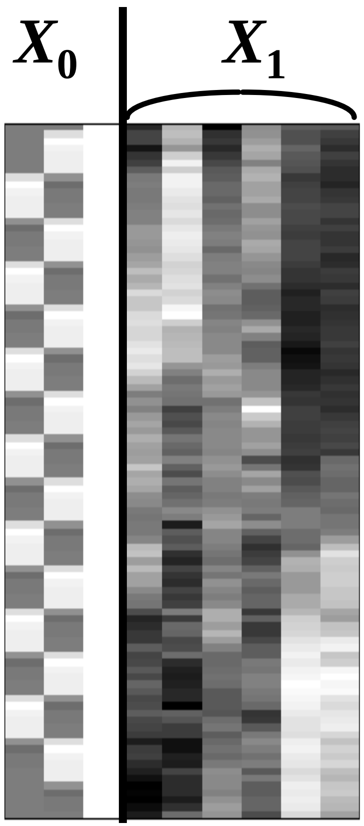
➢ Be careful of the interpretation of the contrasts $c^T \hat{\beta}$ themselves (eg, for a second level analysis):

sum ≠ average

F-test - the extra-sum-of-squares principle

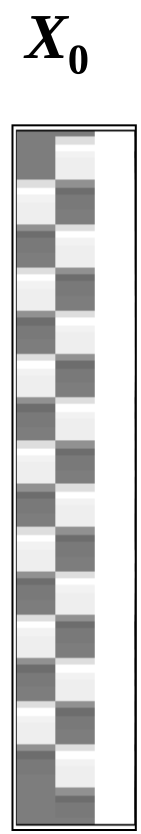
- Model comparison:

Null Hypothesis H$_0$: True model is X_0 (reduced model)



Full model ?

$\longrightarrow$ **RSS** $\sum \hat{\varepsilon}^2_{full}$



or Reduced model?

$\longrightarrow$ **RSS$_0$** $\sum \hat{\varepsilon}^2_{reduced}$

Test statistic: ratio of explained variability and unexplained variability (error)

$$F \propto \frac{RSS_0 - RSS}{RSS}$$

$$F \propto \frac{ESS}{RSS} \sim F_{\nu_1,\nu_2}$$

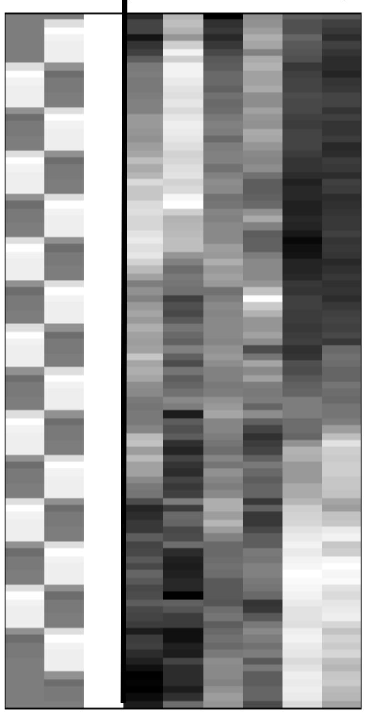
ν_1 = rank(X) – rank(X_0)
ν_2 = N – rank(X)

***F*-test** - multidimensional contrasts – SPM{*F*}

- Tests multiple linear hypotheses:

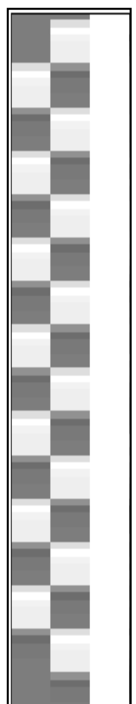
H$_0$: True model is X_0	**H$_0$:** $\beta_4 = \beta_5 = \ldots = \beta_9 = 0$	**test H$_0$:** $c^T\beta = 0$?

X_0 | X_1 $(\beta_{4\text{-}9})$



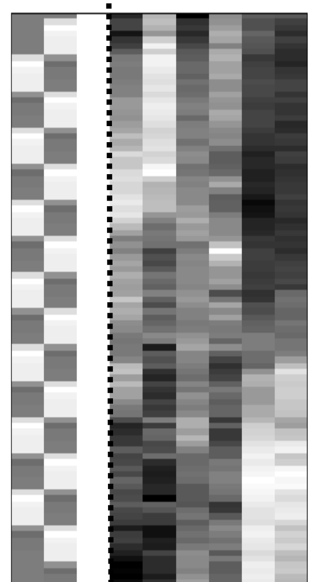
Full model?

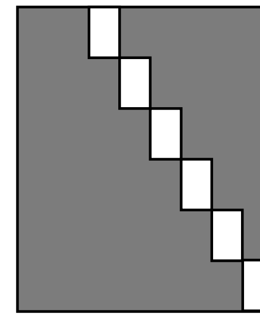
X_0

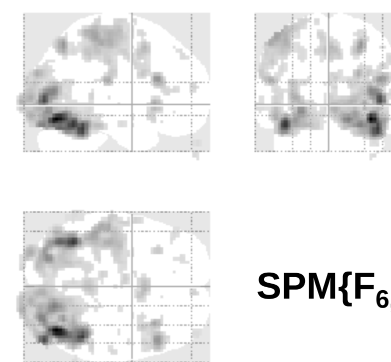


Reduced model?

$$c^T = \begin{array}{ccc|cccccc} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

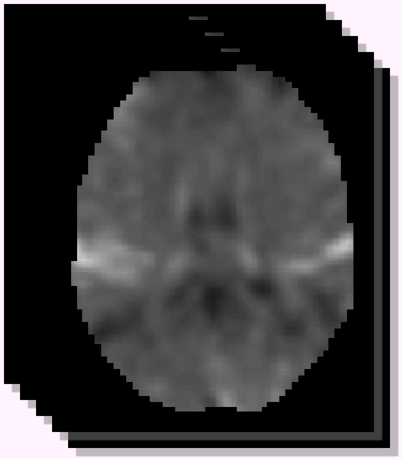






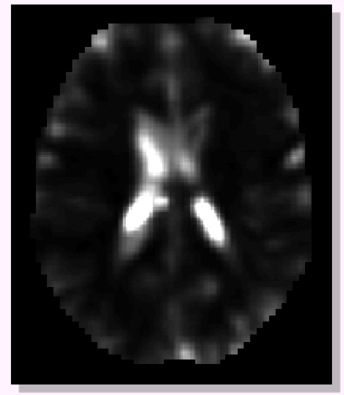
SPM{$F_{6,322}$}

F-contrast in SPM



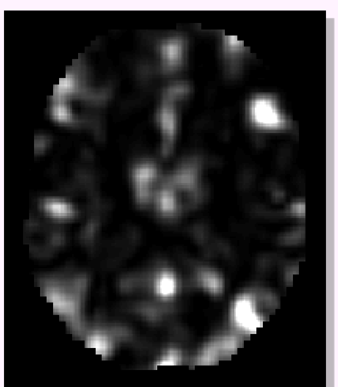
beta_???? images

$$\hat{\beta} = (X^T X)^{-1} X^T y$$



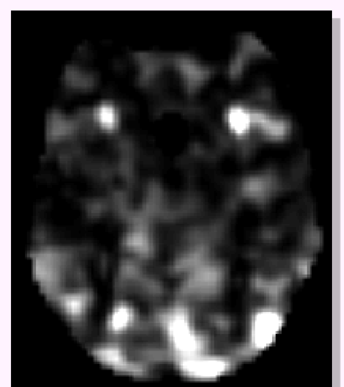
ResMS image

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{N - p}$$



ess_???? images

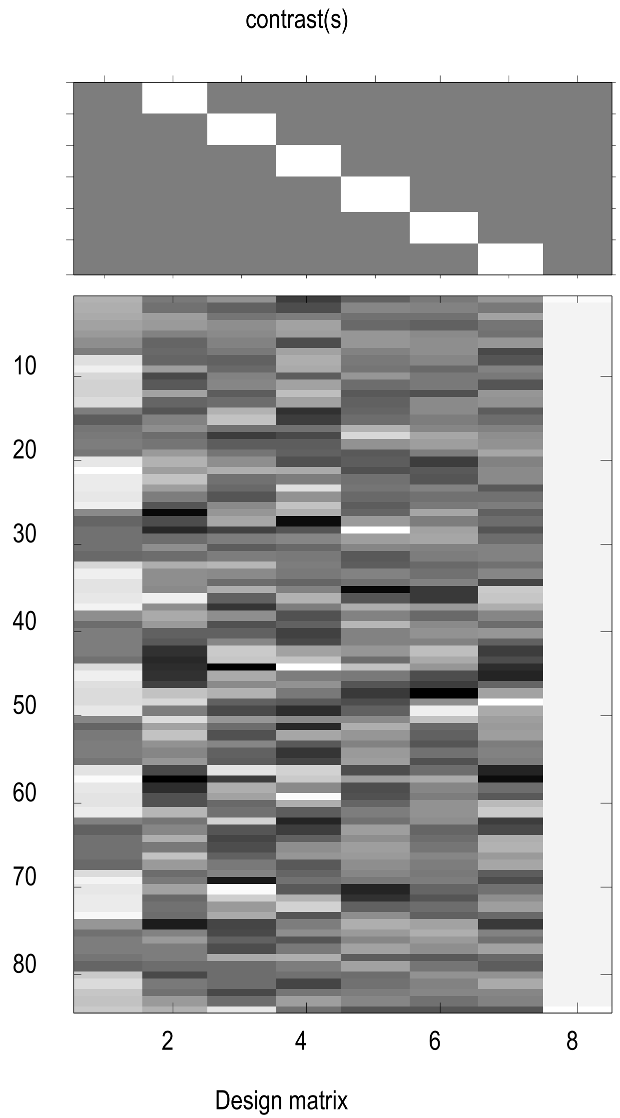
$$(\, RSS_0 \;-\; RSS \,)$$

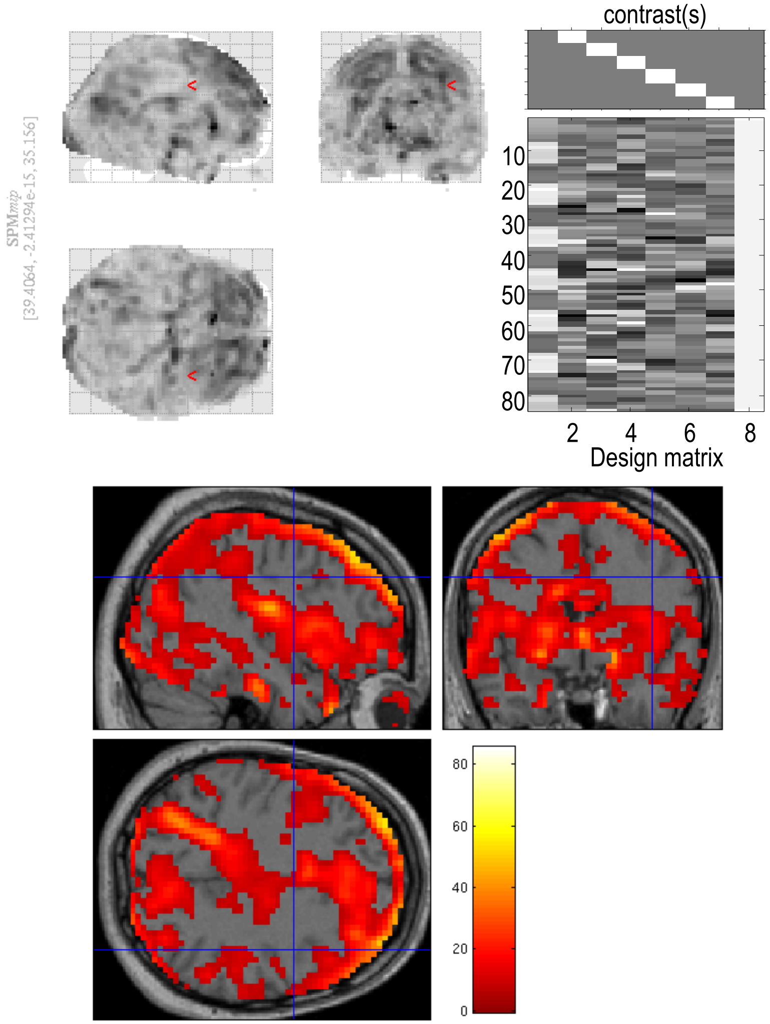


spmF_???? images

SPM{F}

F-test example: movement related effects





F-test: summary

- F-tests can be viewed as testing for the additional variance explained by a larger model wrt a simpler (***nested***) model ⇨ ***model comparison***.
- F tests a weighted **sum of squares** of one or several combinations of the regression coefficients β.
- In practice, we don't have to explicitly separate *X* into $[X_1X_2]$ thanks to **multidimensional contrasts**.
- Hypotheses:

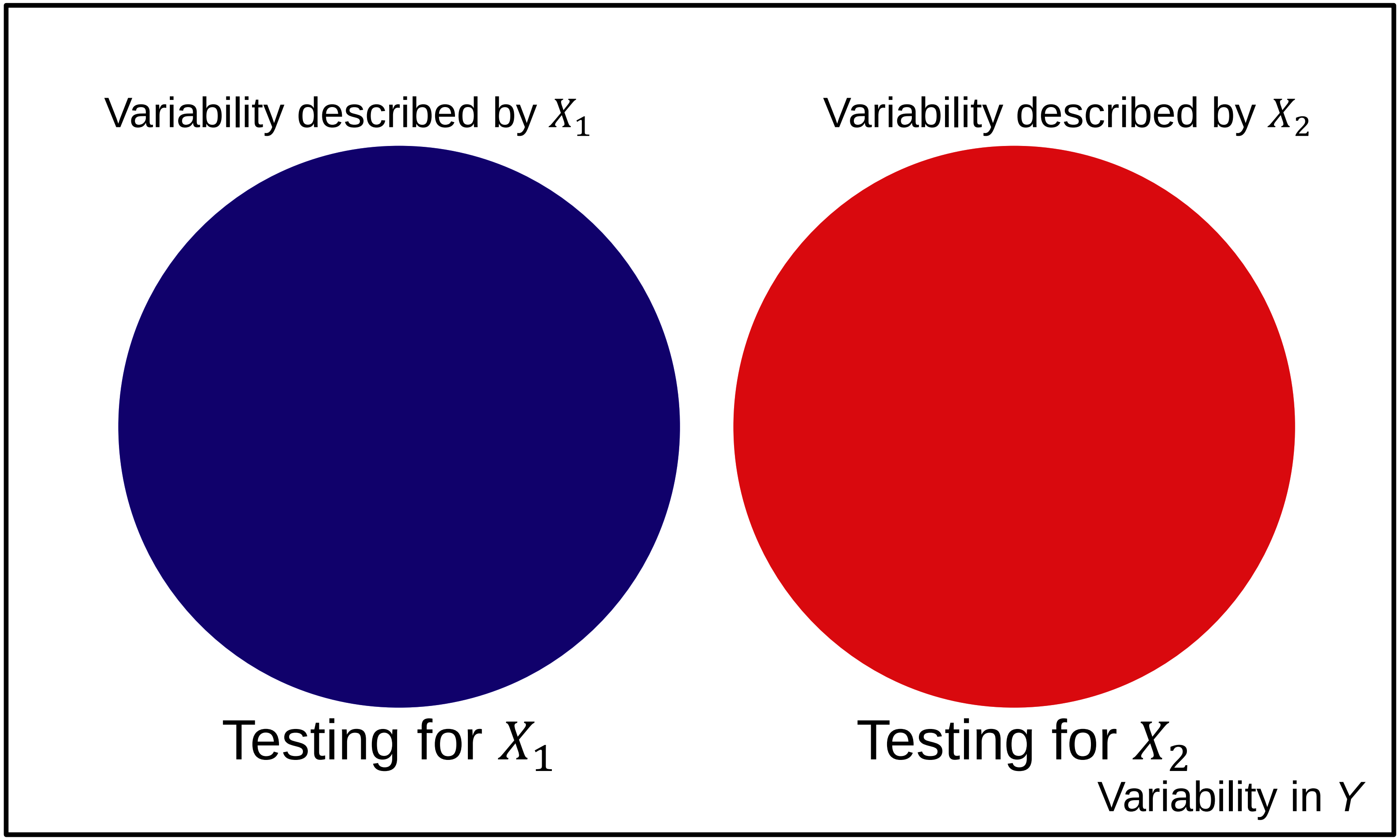
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{Null Hypothesis } H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

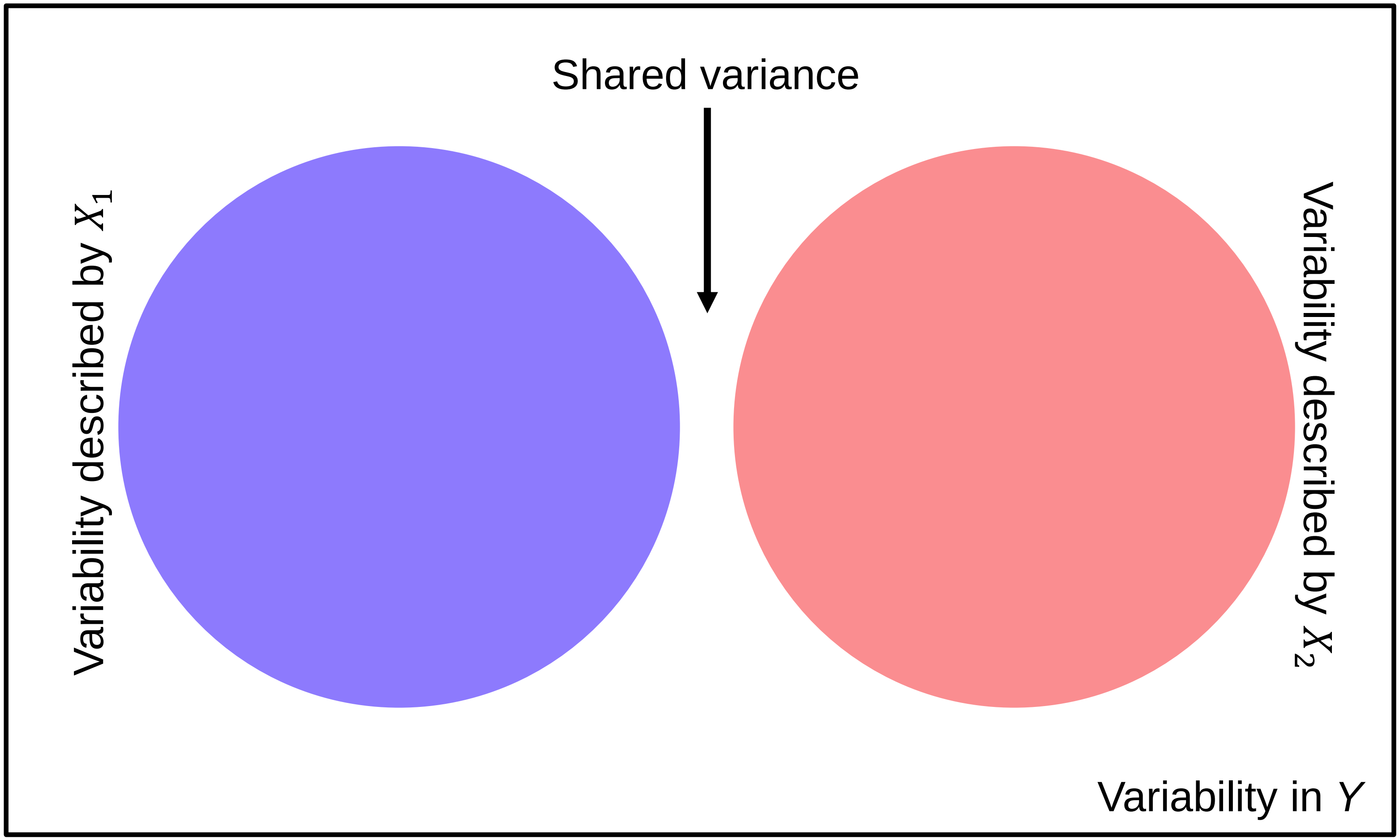
$$\text{Alternative Hypothesis } H_A: \text{at least one } \beta_k \neq 0$$

- In testing uni-dimensional contrast with an *F*-test, for example $\beta_1 - \beta_2$, the result will be the same as testing $\beta_2 - \beta_1$. It will be exactly the square of the *t*-test, testing for both positive and negative effects.

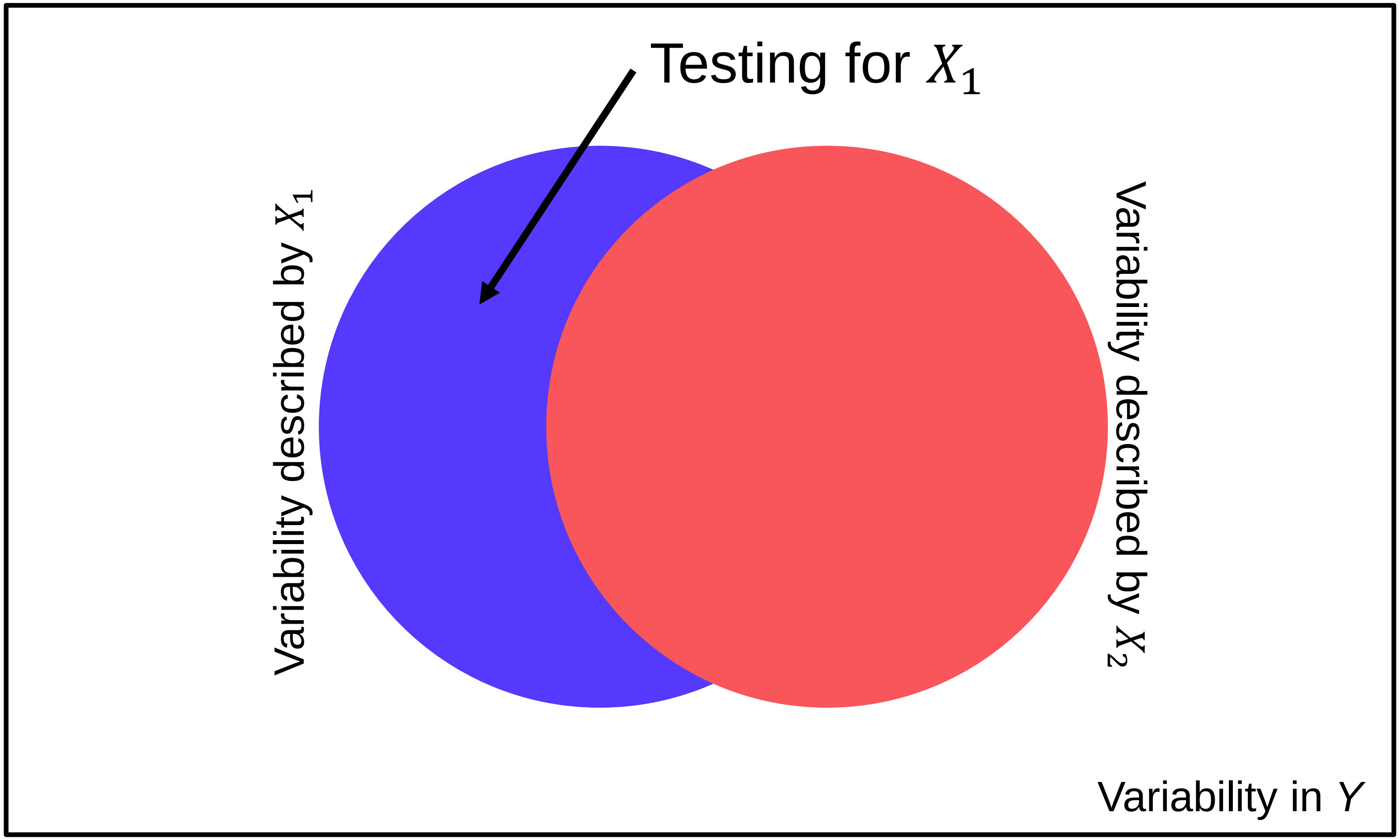
Orthogonal regressors



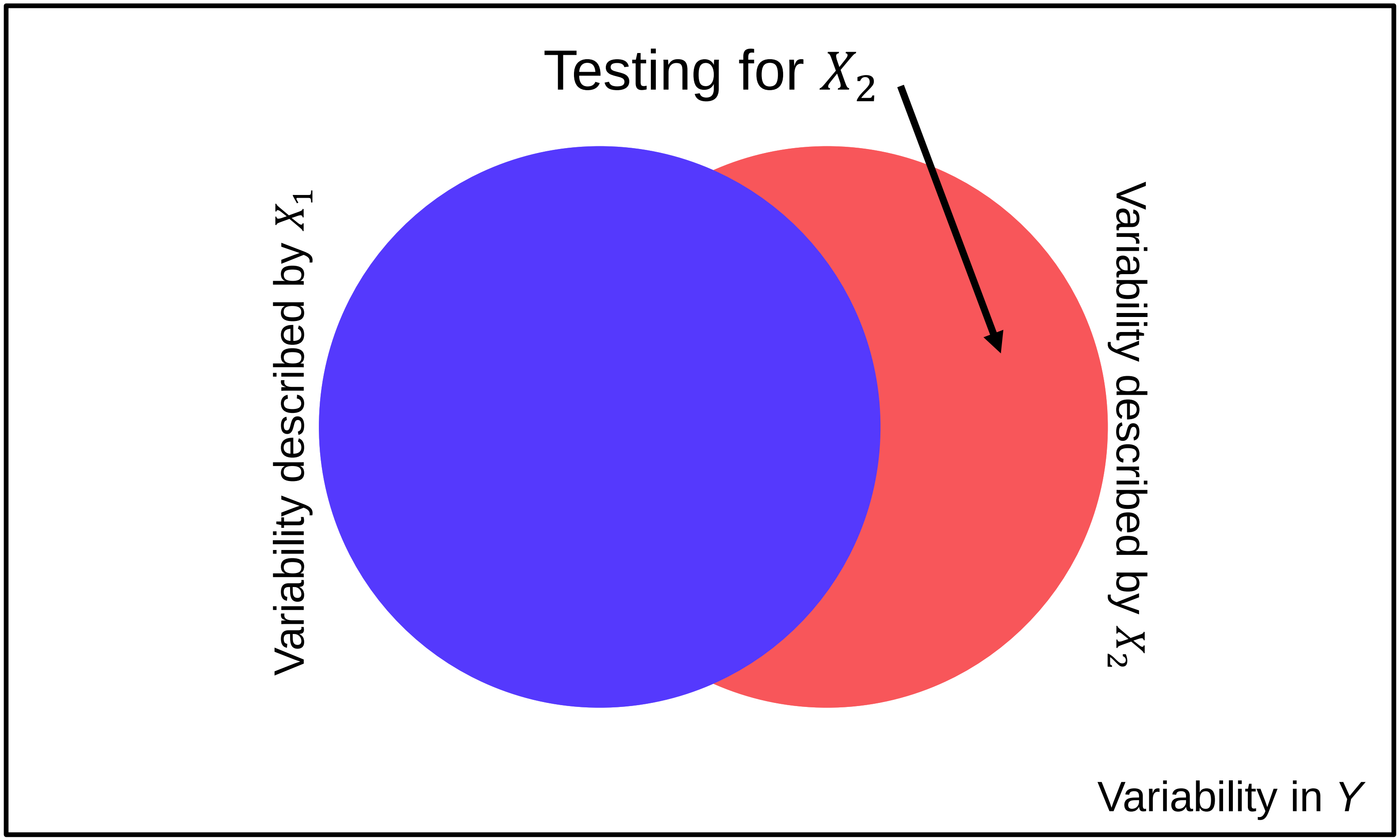
Correlated regressors



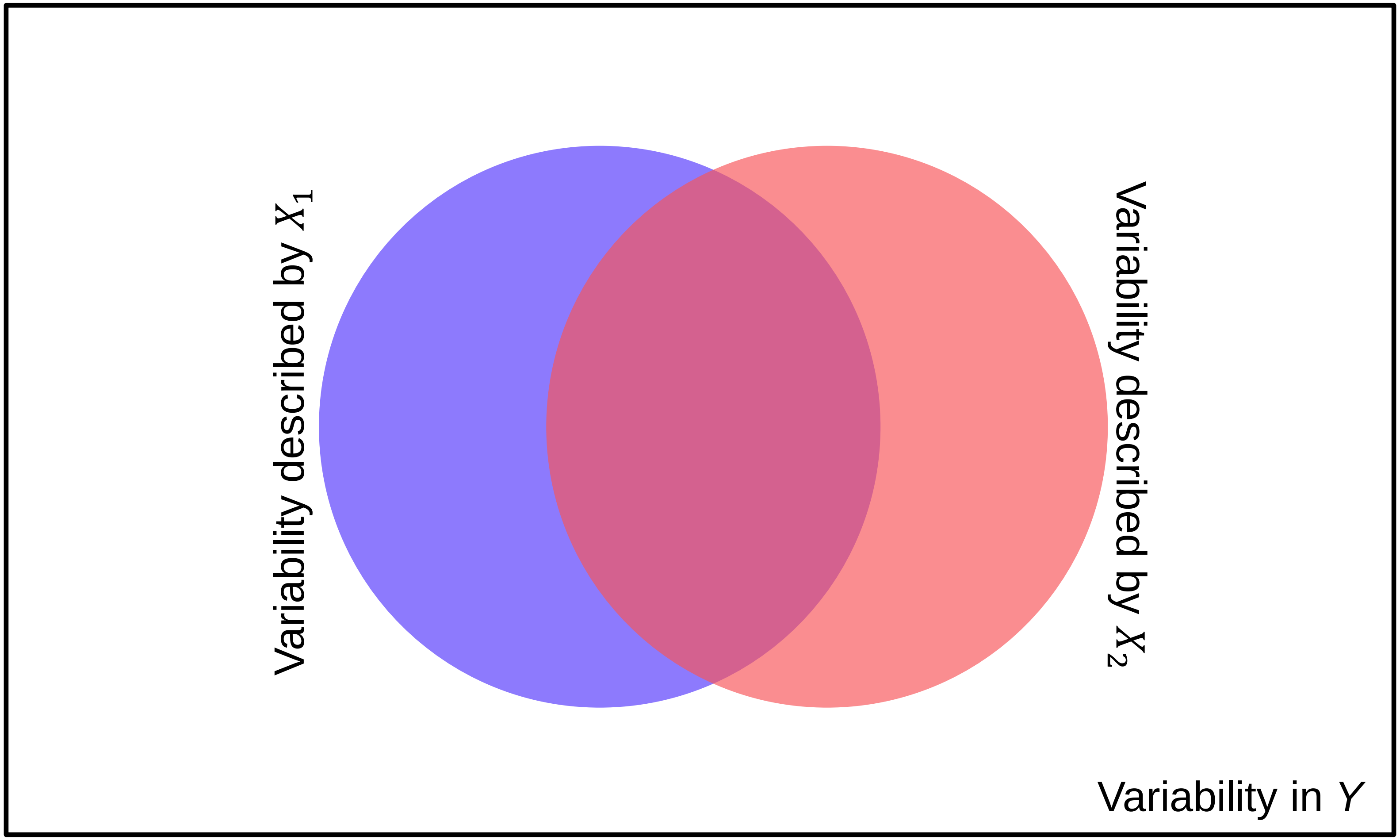
Correlated regressors



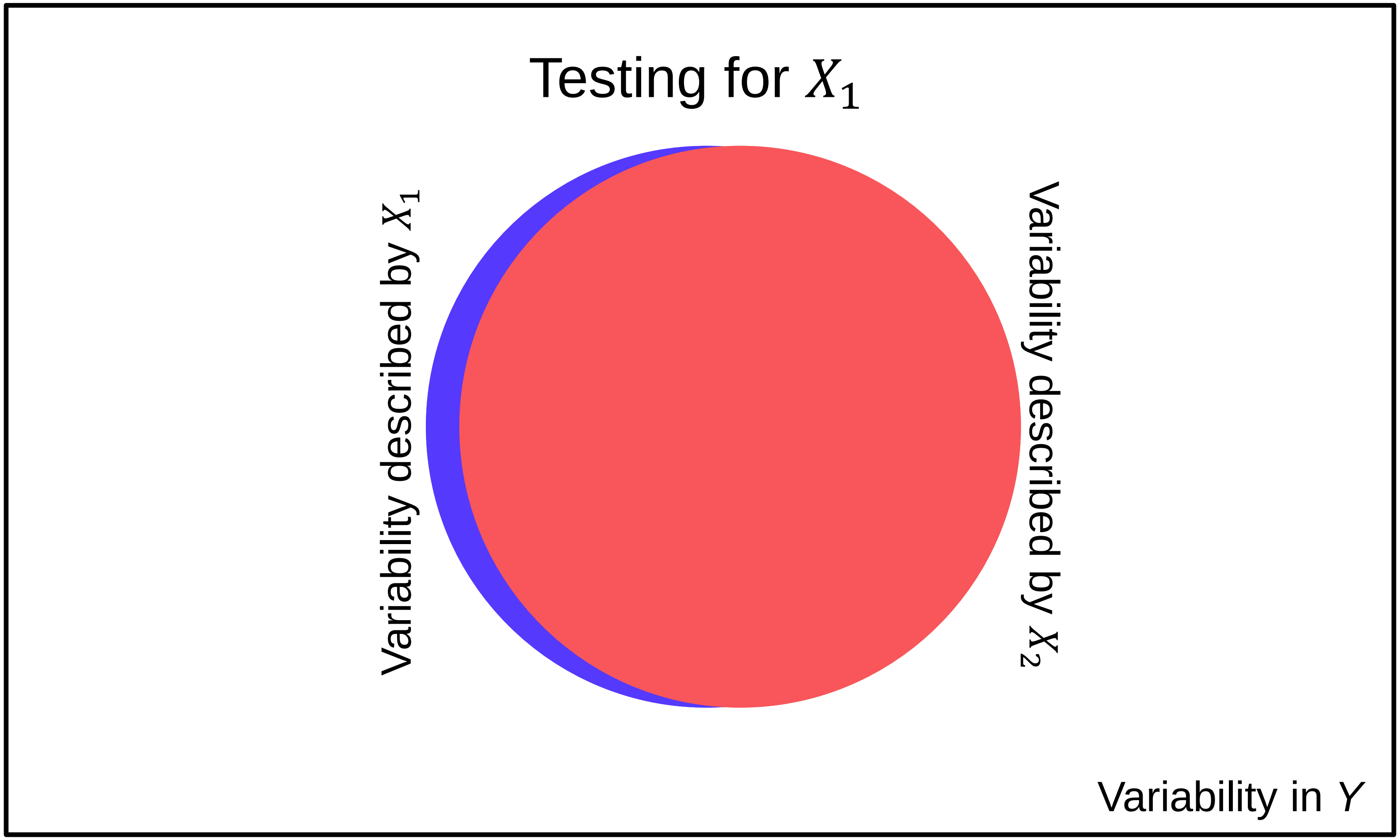
Correlated regressors



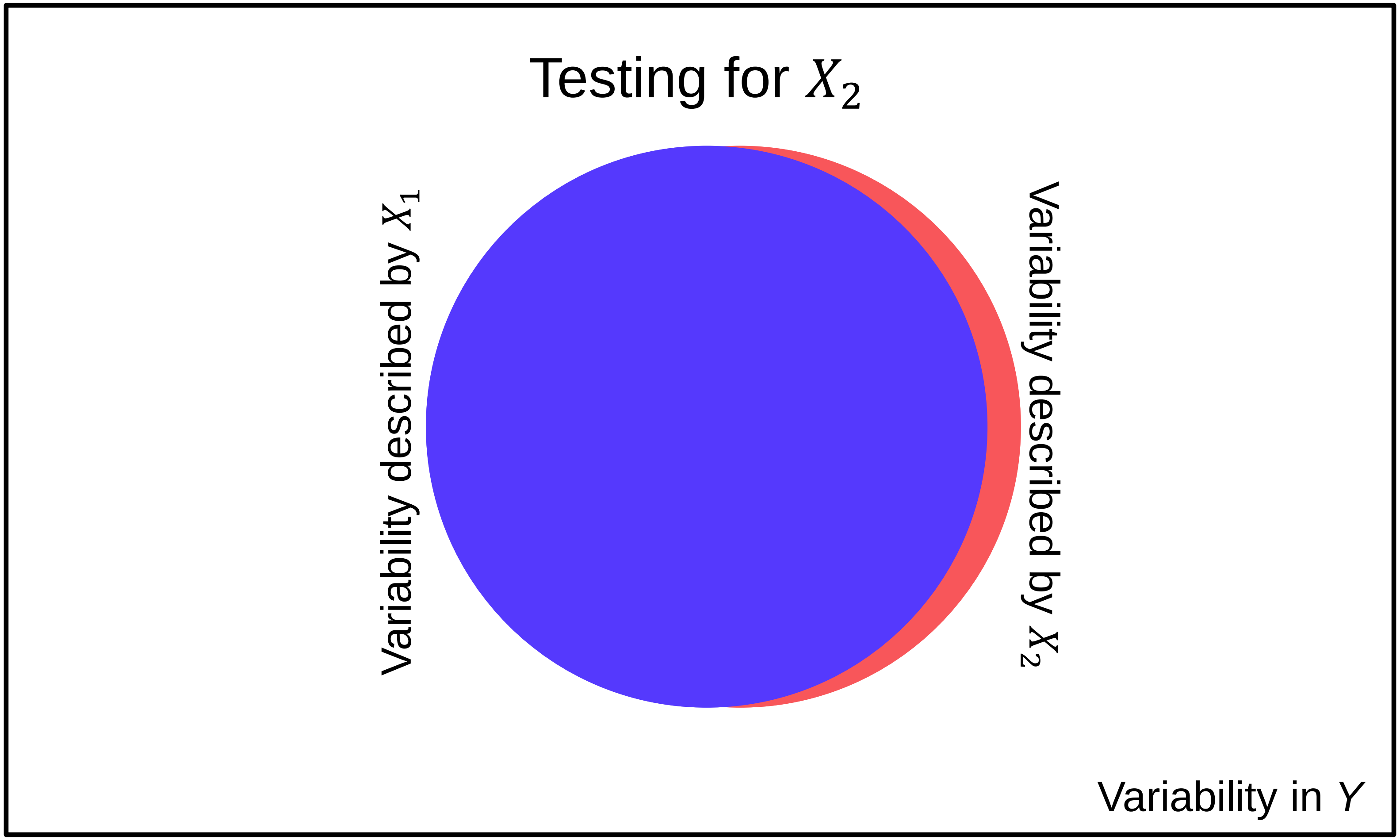
Correlated regressors



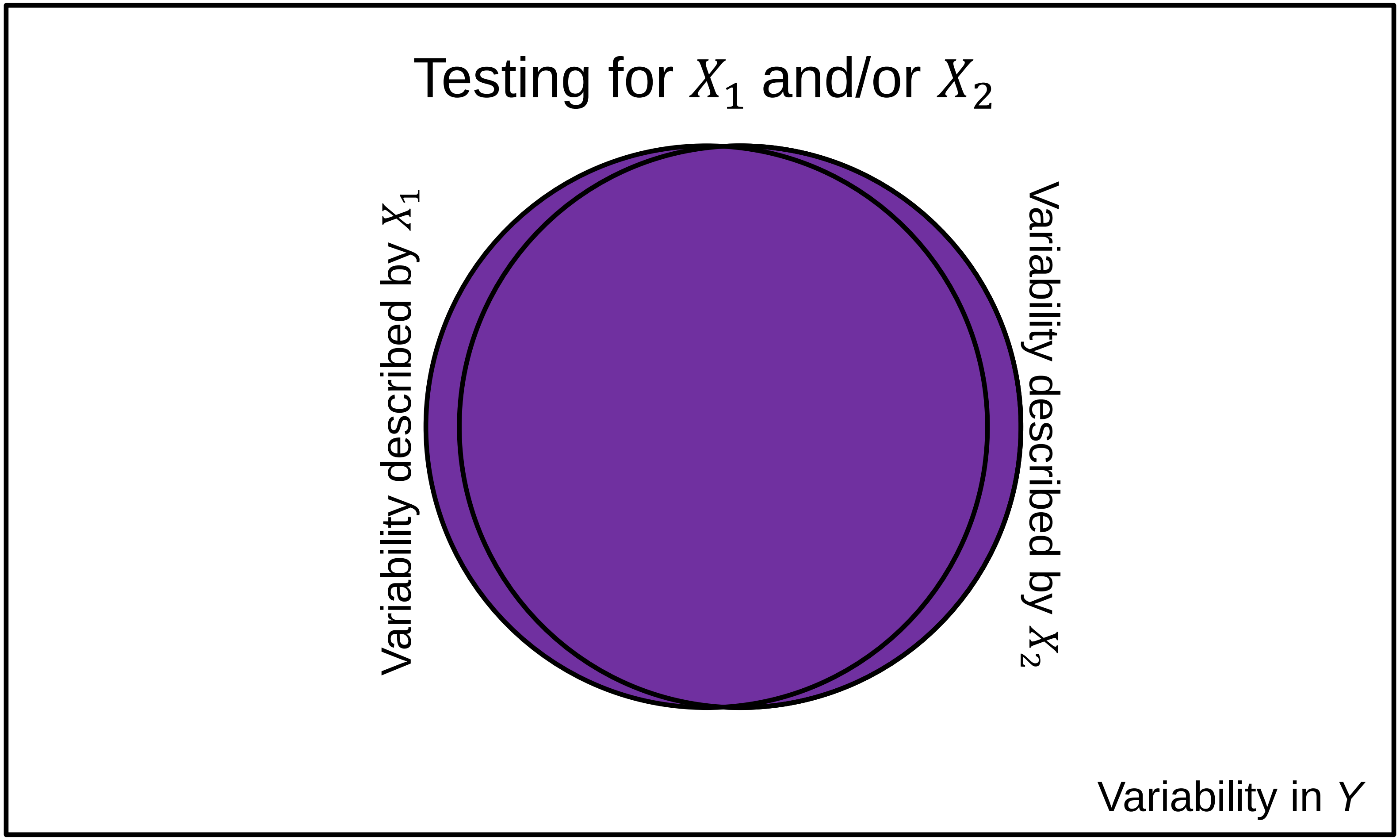
Correlated regressors



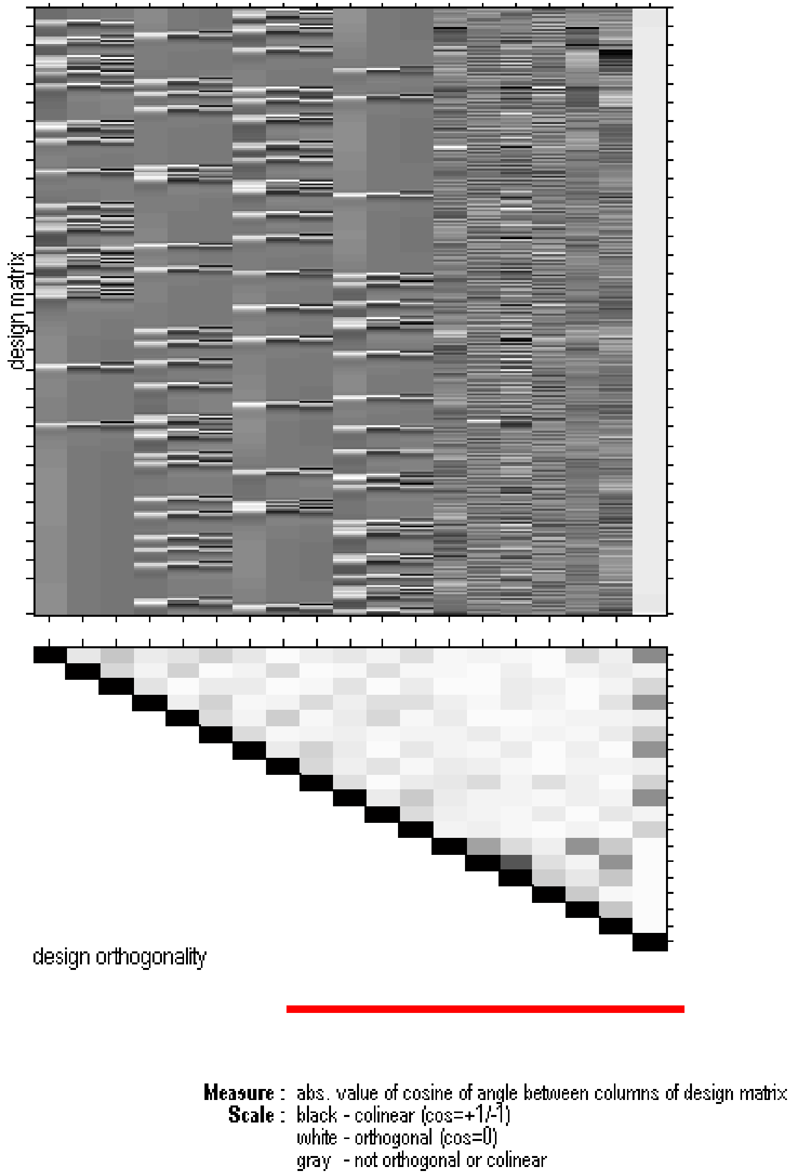
Correlated regressors



Correlated regressors



Design orthogonality

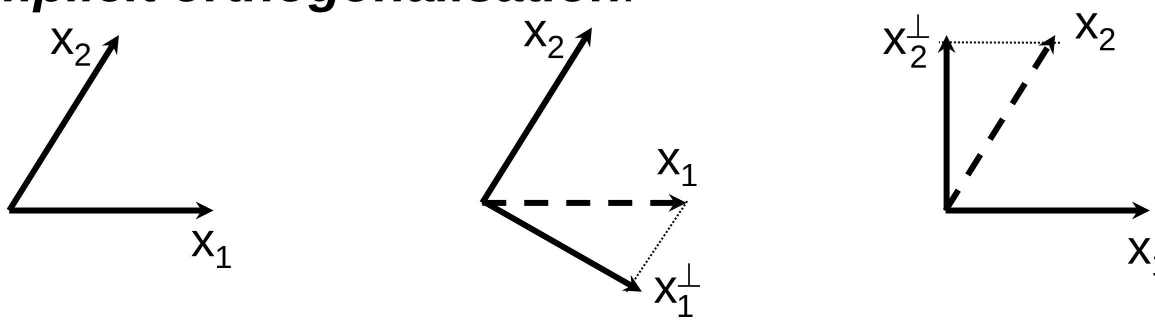


- For each pair of columns of the design matrix, the orthogonality matrix depicts the magnitude of the **cosine of the angle** between them, with the range 0 to 1 mapped from white to black.

- If both vectors have **zero mean** then the cosine of the angle between the vectors is the same as the **correlation** between the two variates.

Correlated regressors: summary

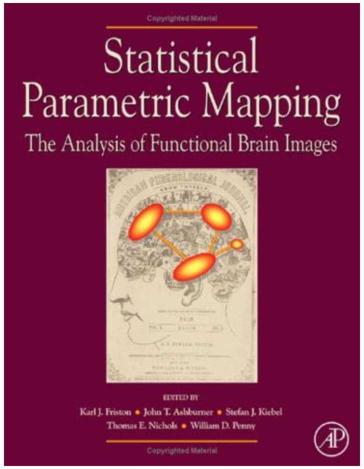
- We implicitly test for an **additional** effect only. When testing for the first regressor, we are effectively removing the part of the signal that can be accounted for by the second regressor:
  ⇨ ***implicit orthogonalisation***.



$$x_2^{\perp} = x_2 - x_1.x_2\ x_1$$

- Orthogonalisation = decorrelation. Parameters and test on the non modified regressor change.
  Rarely solves the problem as it requires assumptions about which regressor to uniquely attribute the common variance.
  ⇨ change regressors (i.e. design) instead, e.g. factorial designs.
  ⇨ use F-tests to assess overall significance.
- Original regressors may not matter: it's the contrast you are testing which should be as decorrelated as possible from the rest of the design matrix

Bibliography:



- *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier, 2007.
- *Plane Answers to Complex Questions: The Theory of Linear Models*. R. Christensen, Springer, 1996.
- *Statistical parametric maps in functional imaging: a general linear approach*. K.J. Friston et al, Human Brain Mapping, 1995.
- *Ambiguous results in functional neuroimaging data analysis due to covariate correlation*. A. Andrade et al., NeuroImage, 1999.
- Estimating efficiency a priori: a comparison of blocked and randomized designs. A. Mechelli et al., NeuroImage, 2003.