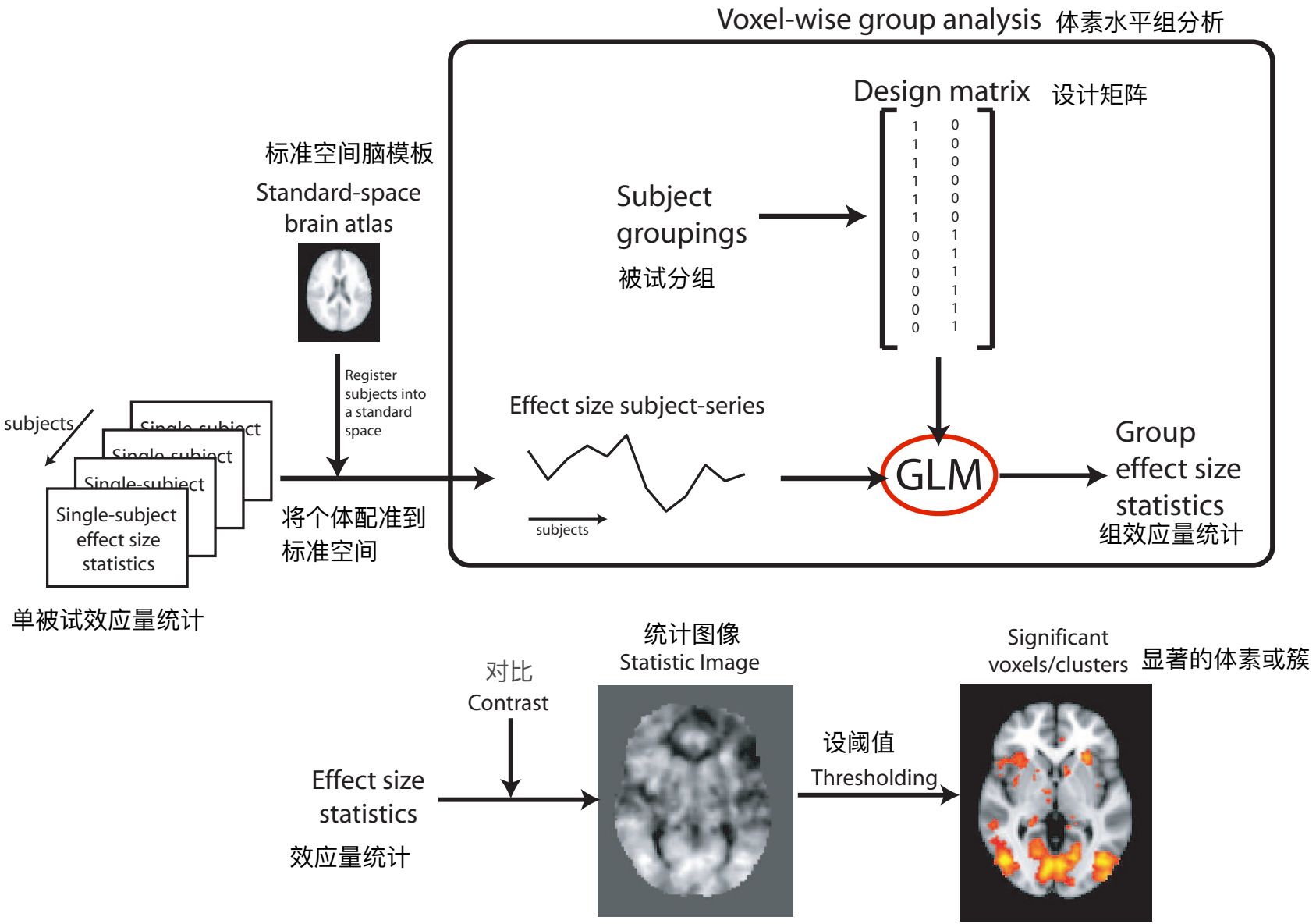
FMRI Group Analysis fMRI组分析



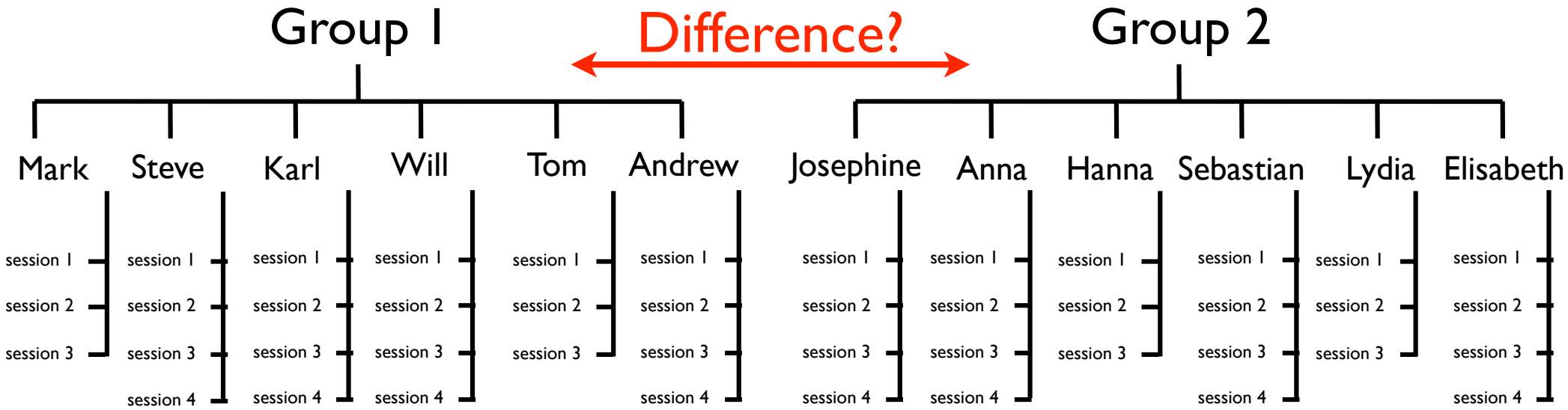
中文翻译：魏朝行 孔亚卓

Multi-Level FMRI analysis

多水平fMRI分析

- uses GLM at both lower and higher levels
  在高、低水平均使用GLM分析

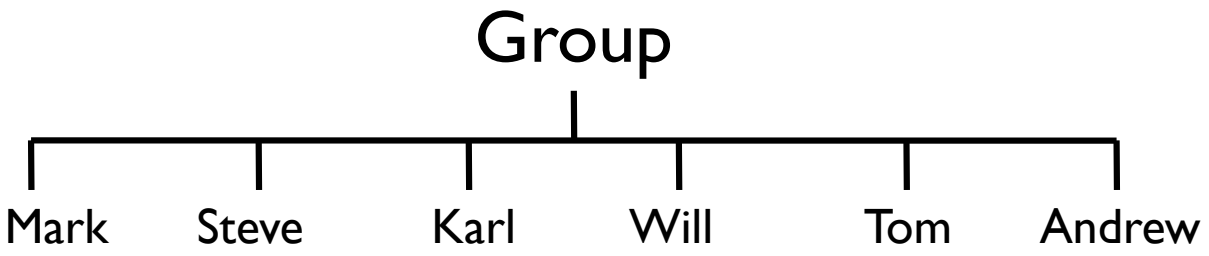
- typically need to infer across multiple subjects, sometimes multiple groups and/or multiple sessions
  通常需要跨多个被试、多个分组和/或多个节段



- questions of interest involve comparisons at the highest level 感兴趣的问题涉及最高水平的比较

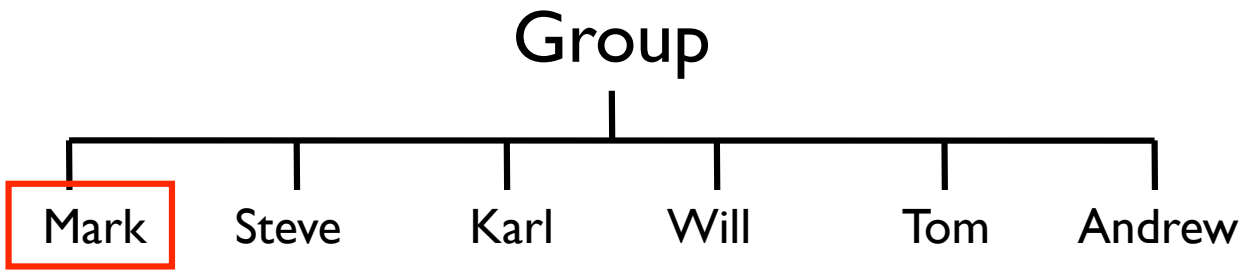
A simple example

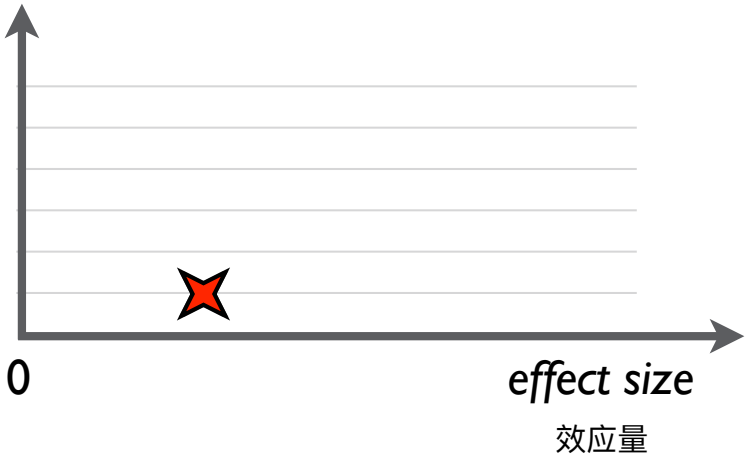
*Does the group activate on average?*组激活是平均吗?



A simple example

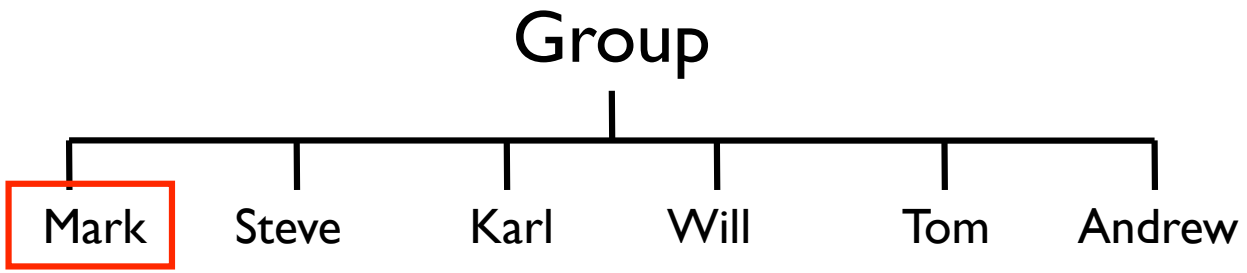
*Does the group activate on average?*组激活是平均吗?





A simple example

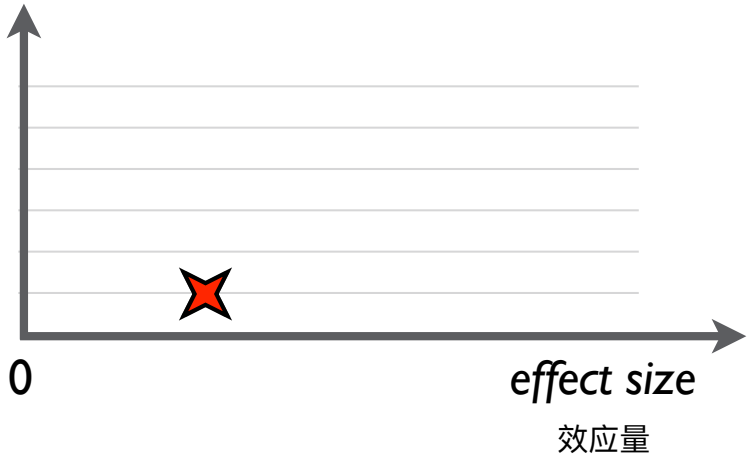
*Does the group activate on average?*组激活是平均吗?



$$Y_k = X_k \beta_k + \epsilon_k$$

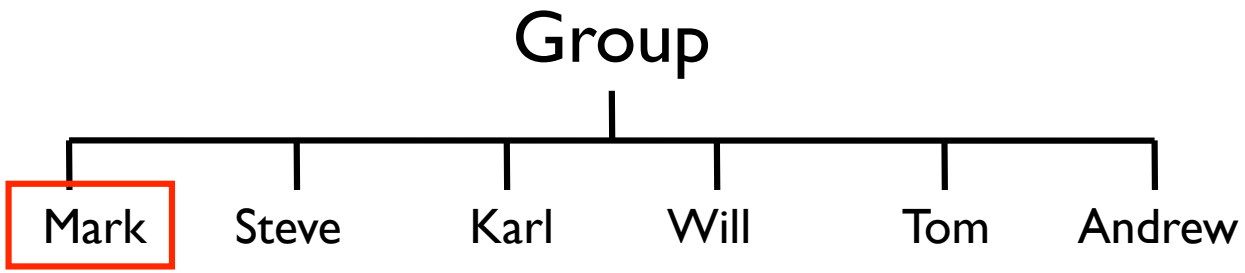
First-level GLM
on Mark's 4D FMRI
data set

基于被试Mark的4D fMRI的第一水平GLM分析

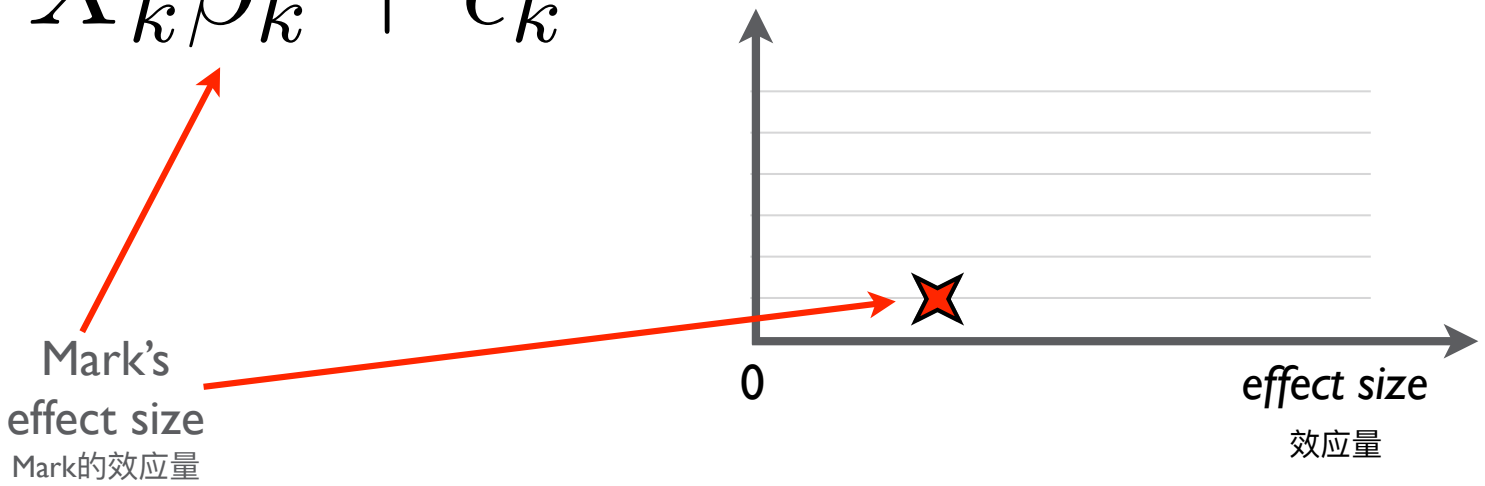


A simple example

*Does the group activate on average?*组激活是平均吗?

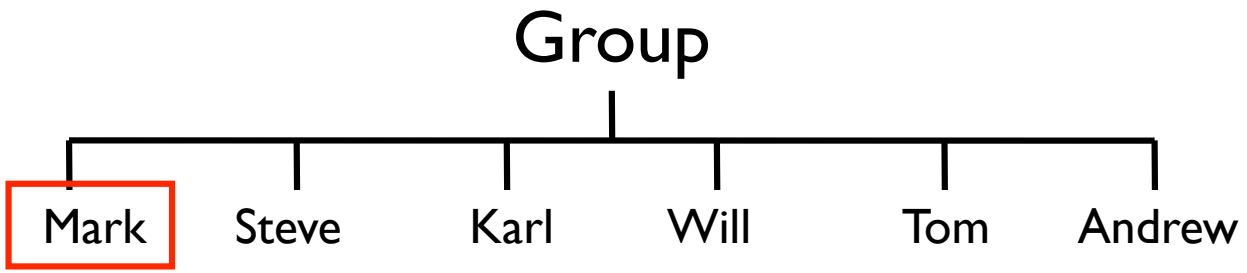


$$Y_k = X_k \beta_k + \epsilon_k$$

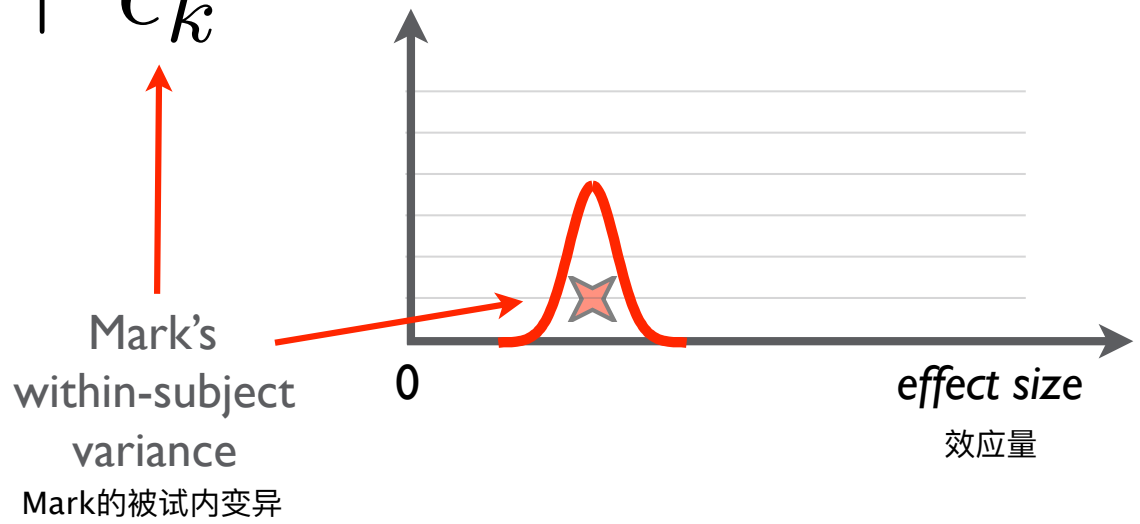


A simple example

*Does the group activate on average?*组激活是平均吗?

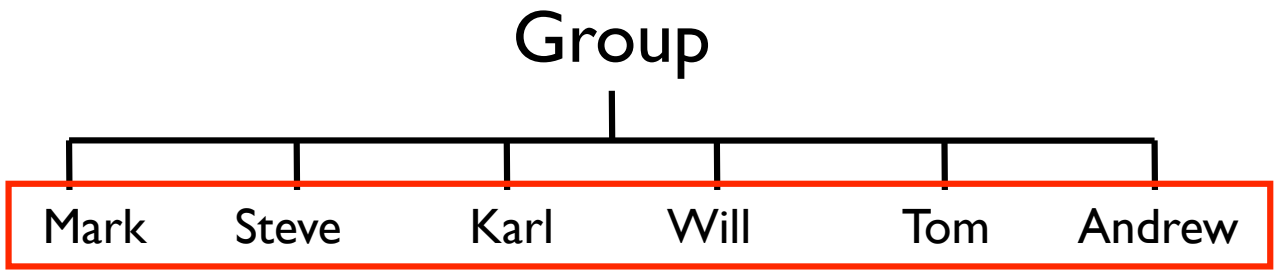


$$Y_k = X_k\beta_k + \epsilon_k$$



A simple example

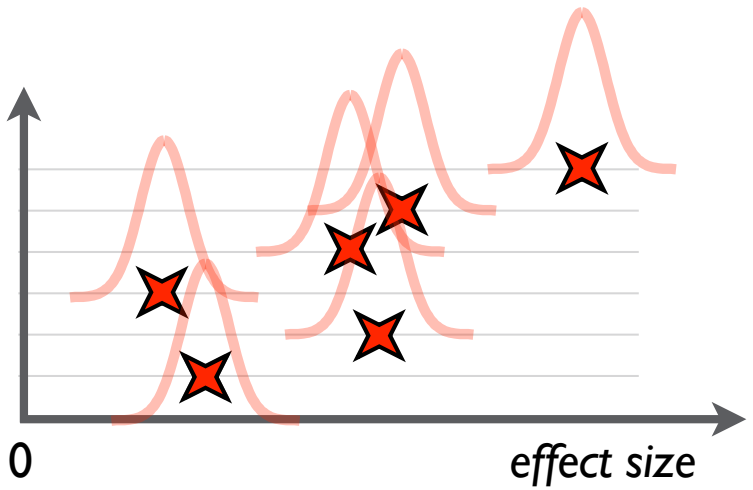
*Does the group activate on average?*组激活是平均吗?



$$Y_K = X_K \beta_K + \epsilon_K$$

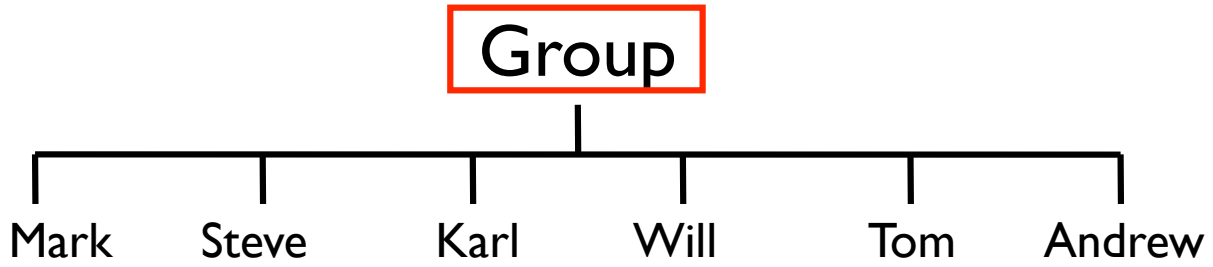
All first-level GLMs
on 6 FMRI data set

基于6个fMRI数据的第一水平GLM分析



A simple example

*Does the group activate on average?*组激活是平均吗?



What group mean are we after? Is it:

1. The group mean for those exact 6 subjects?
   精确的6个被试的组平均?
   Fixed-Effects (FE) Analysis **固定效应分析**

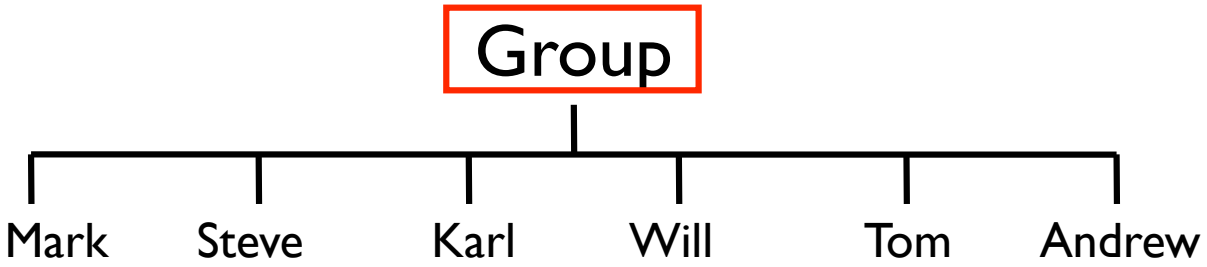
2. The group mean for the population from which these 6 subjects were drawn?6个被试所代表的群体的组平均?
   Mixed-Effects (ME) analysis **混合效应分析**

Fixed-Effects Analysis

固定效应分析

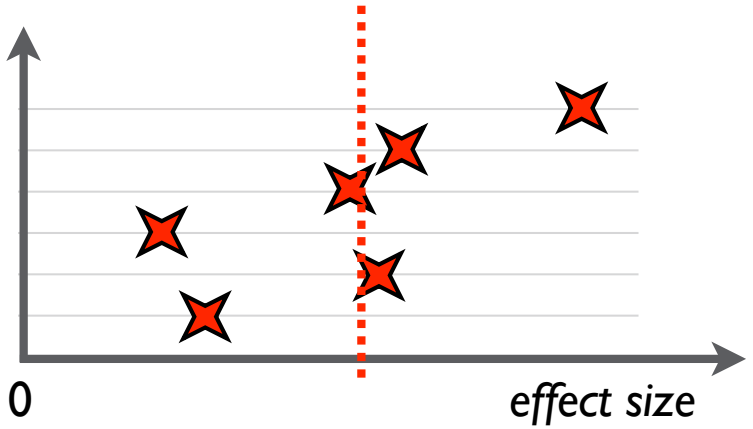
Do these exact 6 subjects activate on average?

这些是6个被试激活的精确平均值吗?



estimate group effect size as
straight-forward mean
across lower-level estimates

估计组效应的大小作为较低水平估计的直接均值



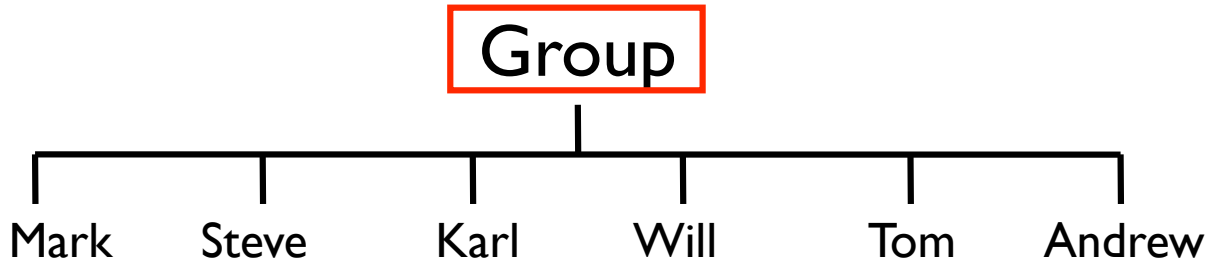
$$\beta_g = \frac{1}{6}\sum_{k=1}^{6}\beta_k$$

Fixed-Effects Analysis

固定效应分析

Do these exact 6 subjects activate on average?

这些是6个被试激活的精确平均值吗?

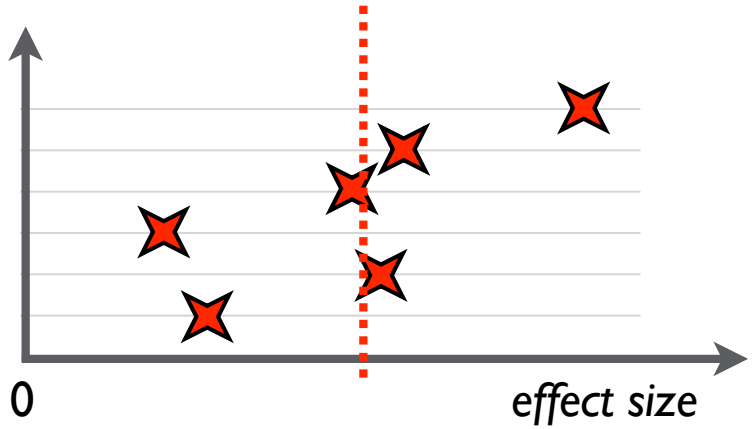


$$Y_K = X_K \beta_K + \epsilon_K$$

$$\beta_K = X_g \beta_g$$

$$X_g = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Group mean



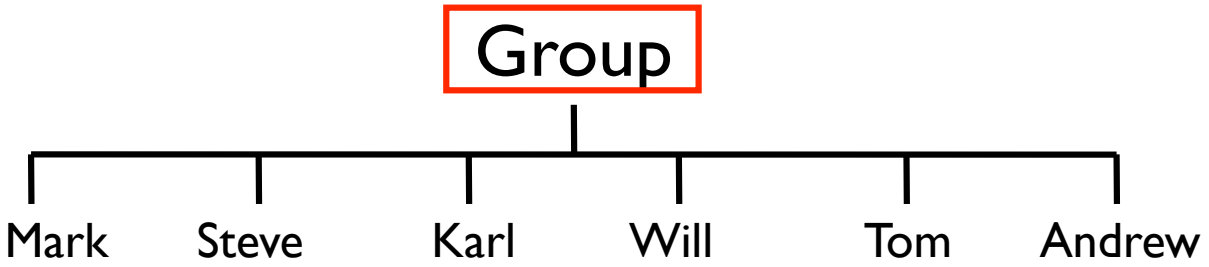
$$\beta_g = \frac{1}{6} \sum_{k=1}^{6} \beta_k$$

Fixed-Effects Analysis

固定效应分析

Do these exact 6 subjects activate on average?

是6个被试激活的精确平均吗?



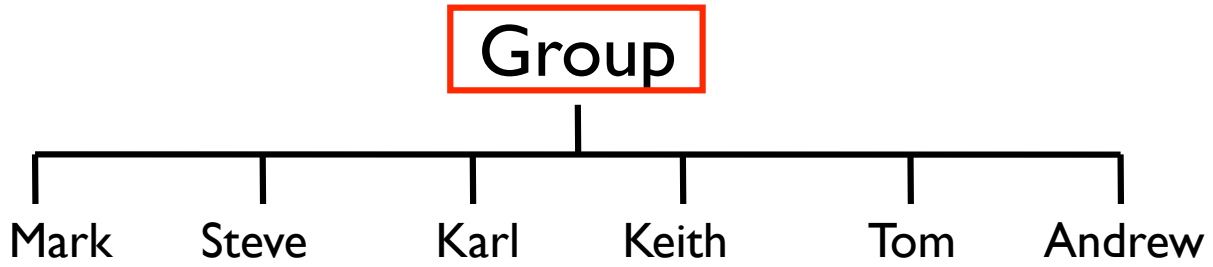
$$Y_K = X_K \beta_K + \epsilon_K$$
$$\beta_K = X_g \beta_g$$

Fixed Effects Analysis: 固定效应分析

- *Consider only these 6 subjects* 只考虑这6个被试
  - *estimate the mean across these subject* 根据这些被试估计均值
  - *only variance is within-subject variance* 只有被试内差异

A simple example

*Does the group activate on average?*组激活是平均吗?



What group mean are we after? Is it:

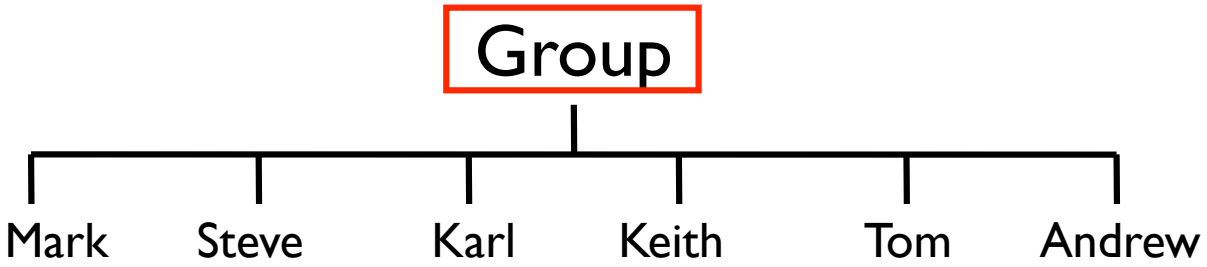
1. The group mean for those exact 6 subjects?
   精确的6个被试的组平均?
   Fixed-Effects (FE) Analysis **固定效应分析**

2. The group mean for the population from which these 6 subjects were drawn?6个被试所代表的群体的组平均?
   Mixed-Effects (ME) analysis **混合效应分析**

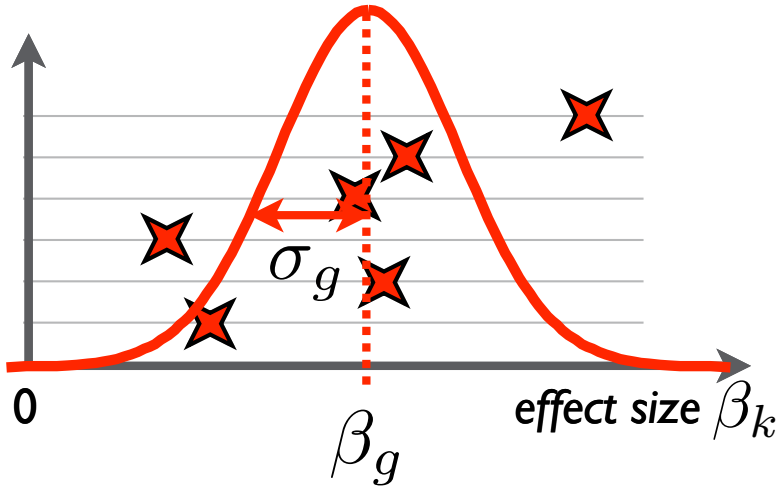
Mixed-Effects Analysis

混合效应分析

Does the population activate on average? 这些是群体的平均值吗?



$$Y_K = X_K \beta_K + \epsilon_K$$



Consider the distribution over the population from which our 6 subjects were sampled:

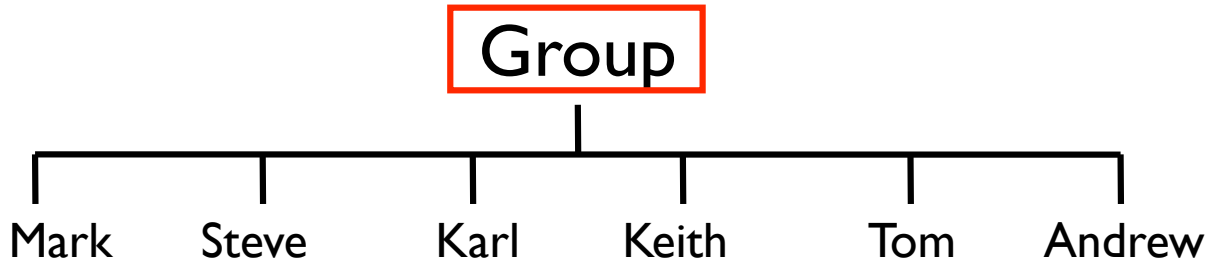
考虑到我们6个被试被抽样的人口分布情况

- σ_g^2 is the between-subject variance

Mixed-Effects Analysis

混合效应分析

*Does the population activate on average?*这些是总体的平均值吗?



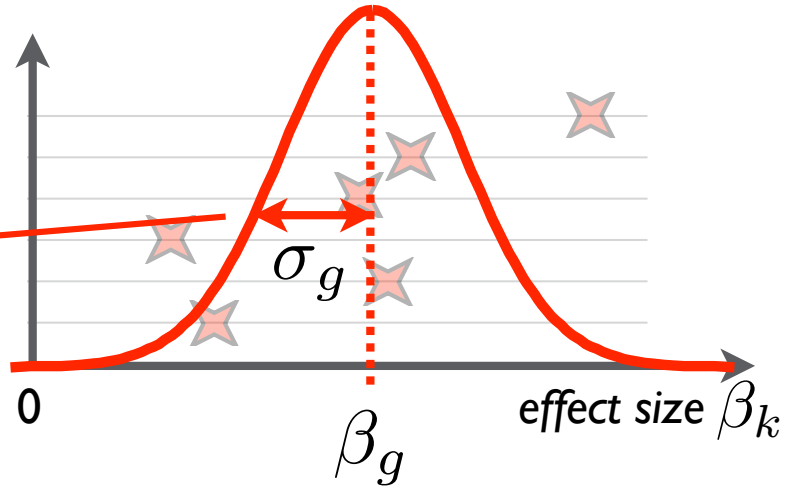
$$Y_K = X_K \beta_K + \epsilon_K$$

$$\beta_K = X_g \beta_g + \epsilon_g$$

$$X_g = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Population mean总体均值

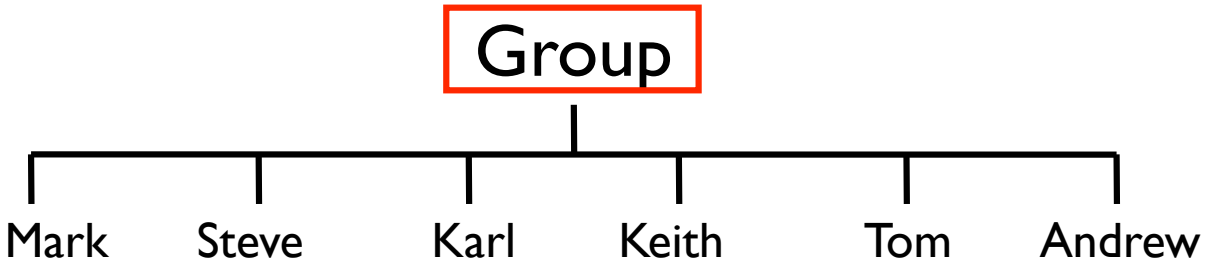
between-subject variation被试间变异



Mixed-Effects Analysis

混合效应分析

Does the population activate on average? 这些是总体的平均值吗?



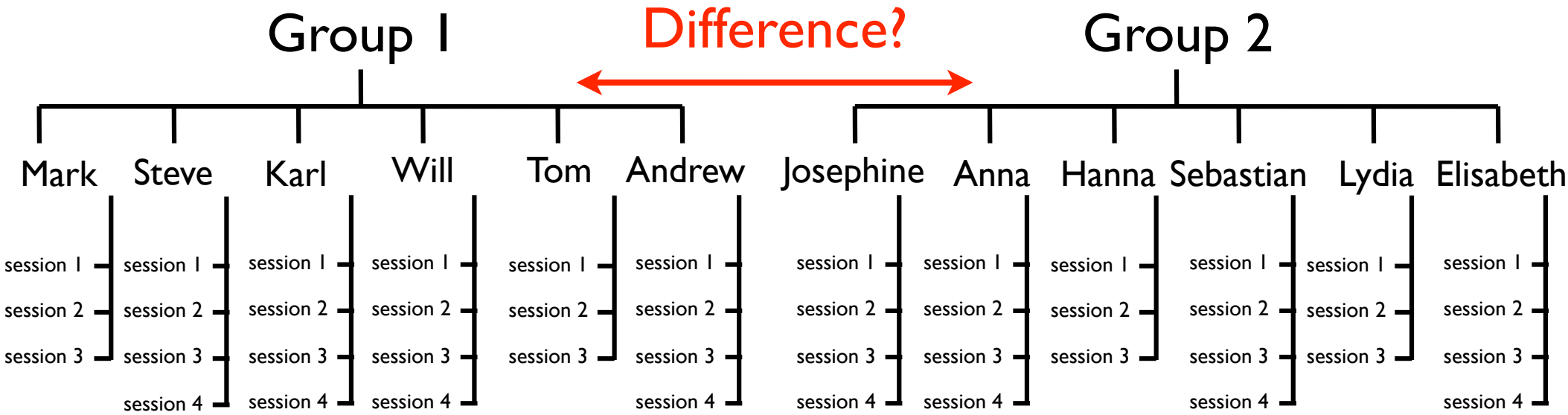
$$Y_K = X_K \beta_K + \epsilon_K$$
$$\beta_K = X_g \beta_g + \epsilon_g$$

Mixed-Effects Analysis:

- *Consider the 6 subjects as samples from a wider population*。
  认为这6个被试是来自更大总体的样本
  - *estimate the mean across the population* 估计总体的均值
  - *between-subject variance accounts for random sampling*
    有随机抽样导致的被试间差异

All-in-One Approach

一体式方法



- Could use one (huge) GLM to infer group difference
  可以使用一个（大的）GLM推断组间差异
  - difficult to ask sub-questions in isolation 难以单独提出子问题
  - computationally demanding 计算要求高
  - need to process again when new data is acquired
    在获取新数据时需要再次处理

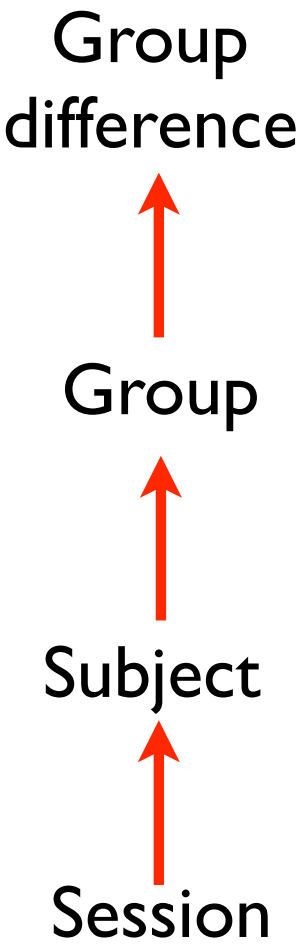
Summary Statistics Approach

概述统计方法

In FEAT estimate levels one stage at a time

在FEAT估计水平的每个步骤

- At each level:在每个水平上
  - Inputs are summary stats from levels below (or FMRI data at the lowest level)
    输入是来自较低级别的汇总统计信息（或最低级别的 FMRI 数据）
  - Outputs are summary stats or statistic maps for inference
    输出是用于推理的汇总统计数据或统计映射
- Need to ensure formal equivalence between different approaches!
  需要确保不同方法之间的形式相同

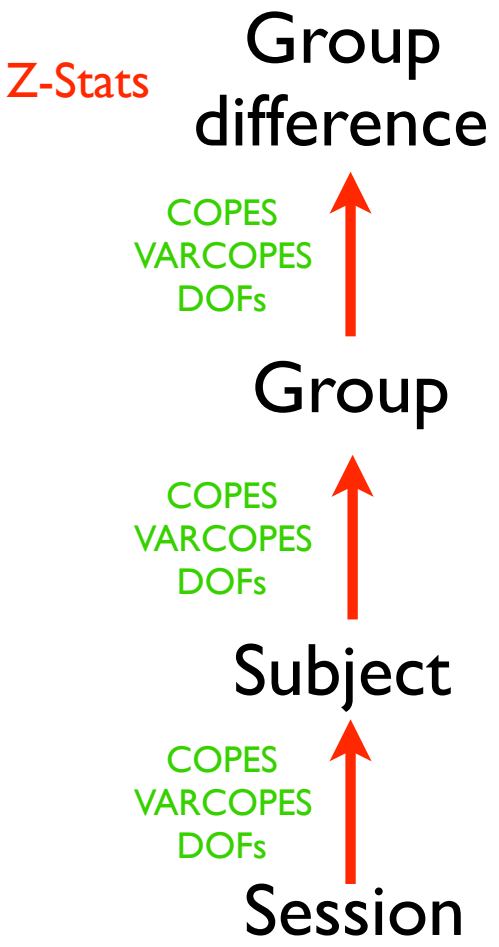


FLAME

FMRIB's Local Analysis of Mixed Effects

FMRIB的混合效应的局部分析方法

- Fully Bayesian framework完全贝叶斯框架
  - use non-central t-distributions: Input COPES, VARCOPES & DOFs from lower-level
    使用非中央 t 分布：输入较低级别的COPES、VARCOPES和DOFs
  - estimate COPES, VARCOPES & DOFs at current level
    在当前水平估计COPES、VARCOPES和DOFs
  - pass these up 向上传递
- Infer at top level在最高水平进行推断
- Equivalent to All-in-One approach
  相当于一体式的方法



FLAME Inference

- Default is:默认：
  - FLAME1: fast approximation for all voxels (using marginal variance MAP estimates)
    所有体素的快速近似值（使用边际方差 MAP 估计值）
- Optional slower, slightly more accurate approach:
  可选的较慢、稍微更准确的方法：
  - FLAME1+2:
    - FLAME1 for all voxels, FLAME2 for voxels close to threshold FLAME1 适用于所有体素，FLAME2 用于接近阈值的体素
    - FLAME2: MCMC sampling technique MCMC采样技术

Choosing Inference Approach

选择推断方法

1. Fixed Effects固定效应

   Use for intermediate/top levels
   用于中级/顶级水平

2. Mixed Effects - OLS 最小二乘

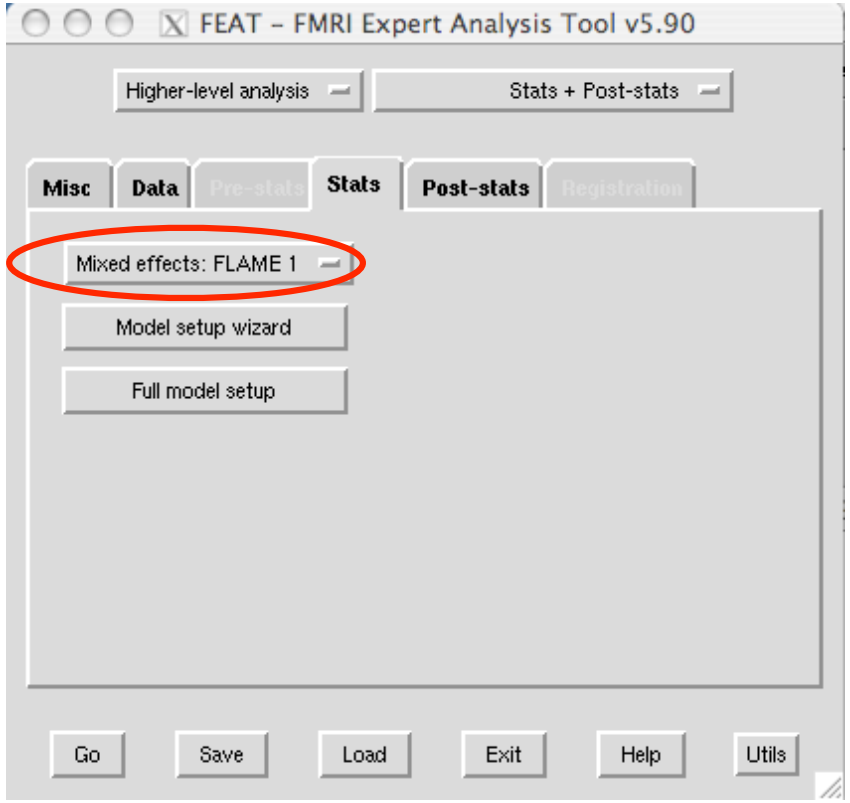
   Use at top level: quick and less accurate
   最高水平使用：快速但不太准确

3. Mixed Effects - FLAME 1

   Use at top level: less quick but more accurate
   最高水平使用：速度较慢，但更准确

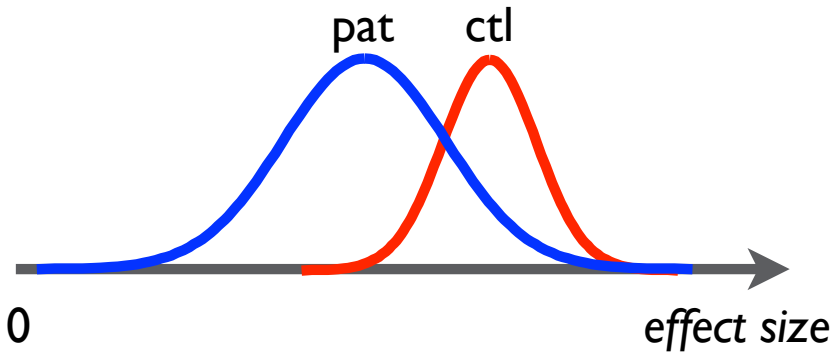
4. Mixed Effects - FLAME 1+2

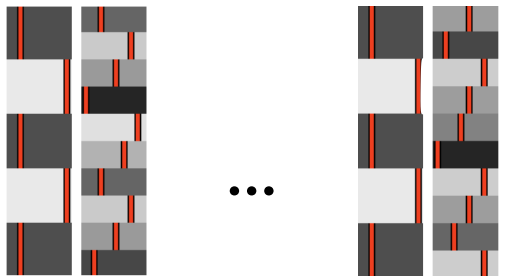
   Use at top level: slow but even more accurate
   最高水平使用：最慢最准确

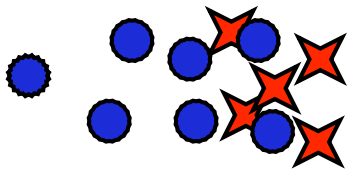


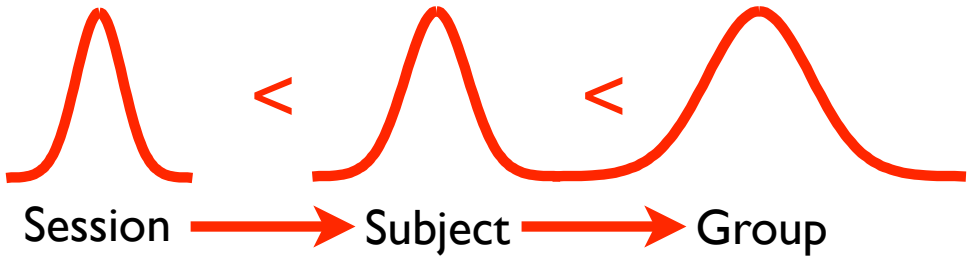
FLAME vs. OLS 对比正交最小二乘法

- allow different within-level variances (e.g. patients vs. controls)
  允许不同的组内差异（例如病人与对照组）
- allow non-balanced designs (e.g. containing behavioural scores)
  允许非平衡设计（例如包含行为评分）
- allow un-equal group sizes
  允许组间样本量不同
- solve the 'negative variance' problem
  解决"负方差"问题



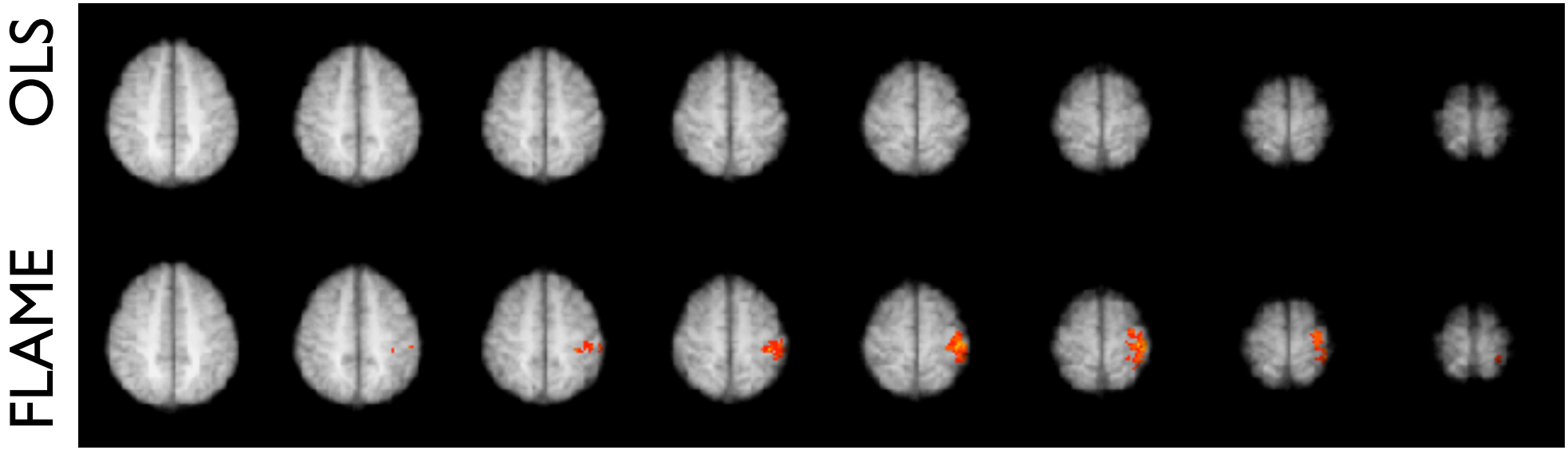






FLAME vs. OLS

- Two ways in which FLAME can give different Z-stats compared to OLS:与OLS相比，FLAME可以提供两种不同的Z检验方法：
  - *higher Z* due to increased efficiency from using lower-level variance heterogeneity
    由于使用较低级别的方差异构性而提高了效率，因此Z更高

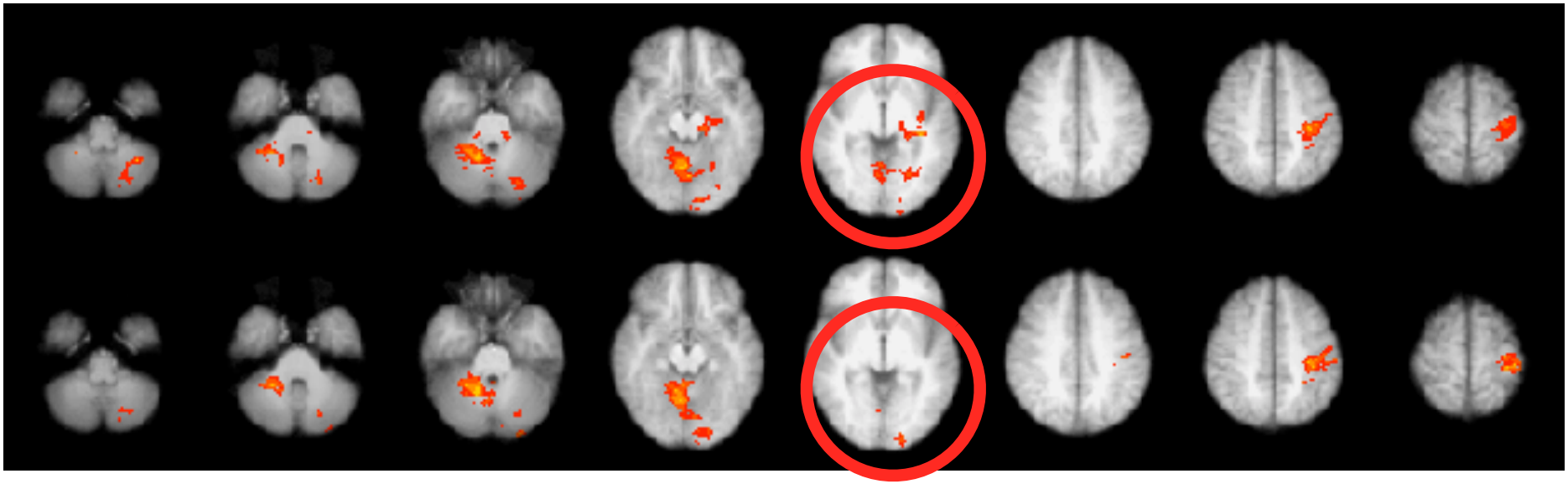


FLAME vs. OLS

- Two ways in which FLAME can give different Z-stats compared to OLS:与OLS相比，FLAME可以提供两种不同的Z检验方法：
  - *Lower Z* due to higher-level variance being constrained to be positive (i.e. solve the implied negative variance problem)
    由于较高水平的方差被限制为正值，因此Z值较低（即解决了负方差问题）

OLS

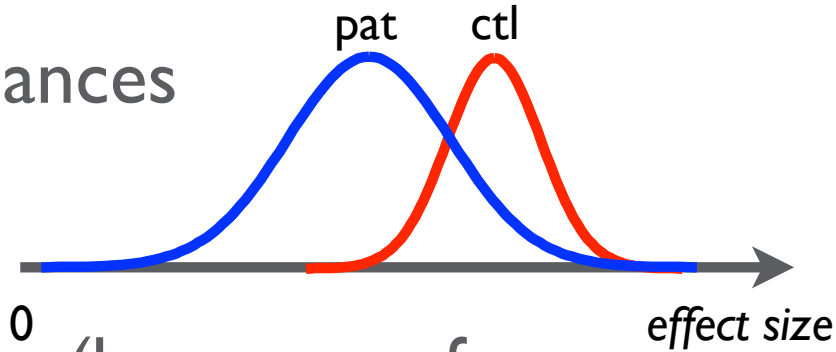
FLAME



Multiple Group Variances

多组差异

- can deal with multiple group variances
  可以处理多组差异



- separate variance will be estimated for each variance group (be aware of #observations for each estimate, though!)
  将为每个方差组估算单独的方差（不过请注意观察每个估算值）

- design matrices need to be 'separable', i.e. EVs only have non-zero values for a single group
  设计矩阵必须"可分离"，即，单个组的EV仅能为非零值

1	1.0	0
1	1.0	0
1	1.0	0
2	0	1.0
2	0	1.0
2	0	1.0

valid

1	1.0	1.0
1	1.0	1.0
1	1.0	1.0
2	1.0	-1.0
2	1.0	-1.0
2	1.0	-1.0

invalid

Examples

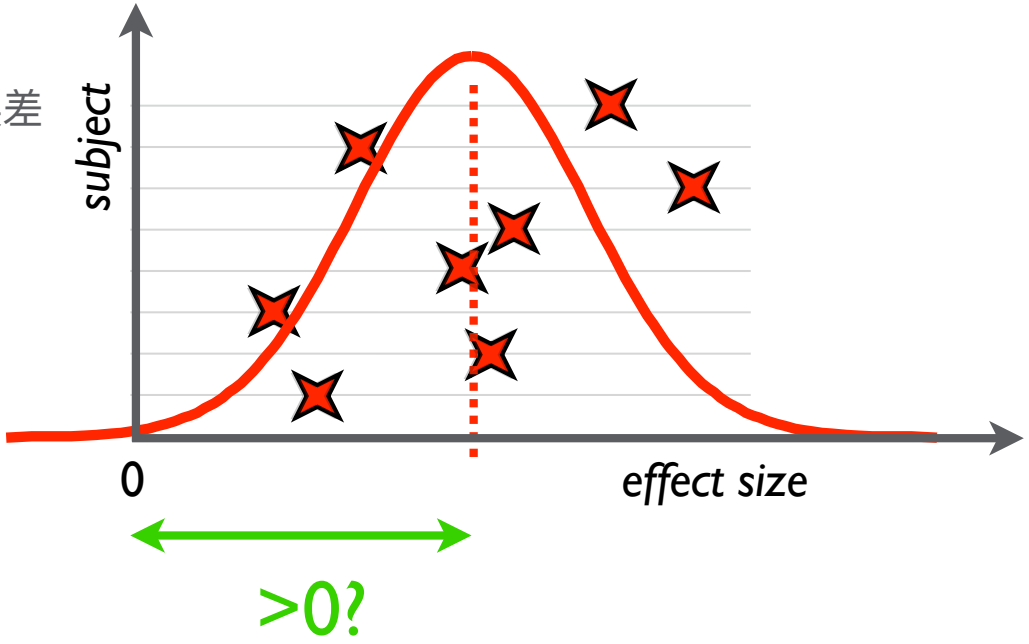
Single Group Average

单组平均值

- We have 8 subjects - all in one group - and want the mean group average:一组8个被试，求组平均

  *Does the group activate on average?*组激活是均值吗?

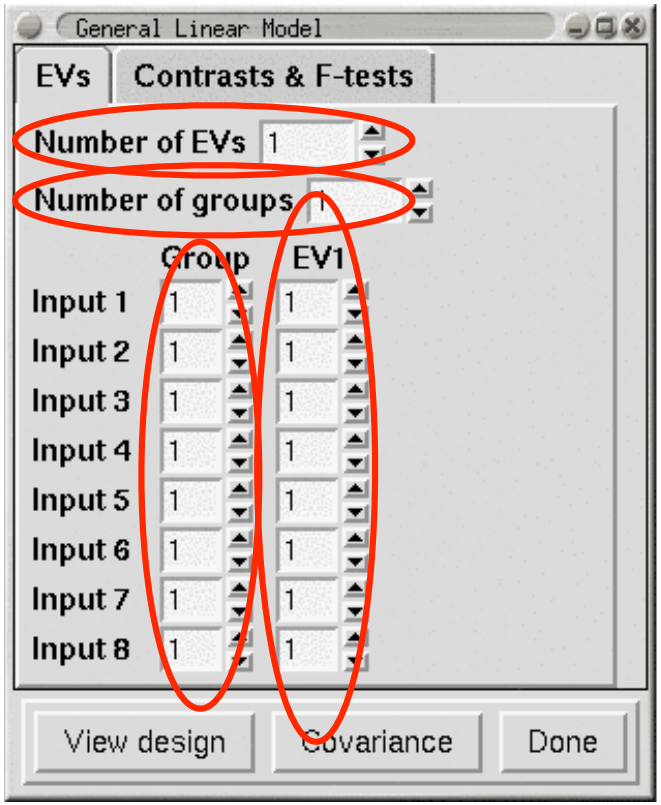
- estimate mean估计均值
- estimate std-error 估计标准误差 (FE or ME)
- test significance of mean > 0检测均值的显著性>0

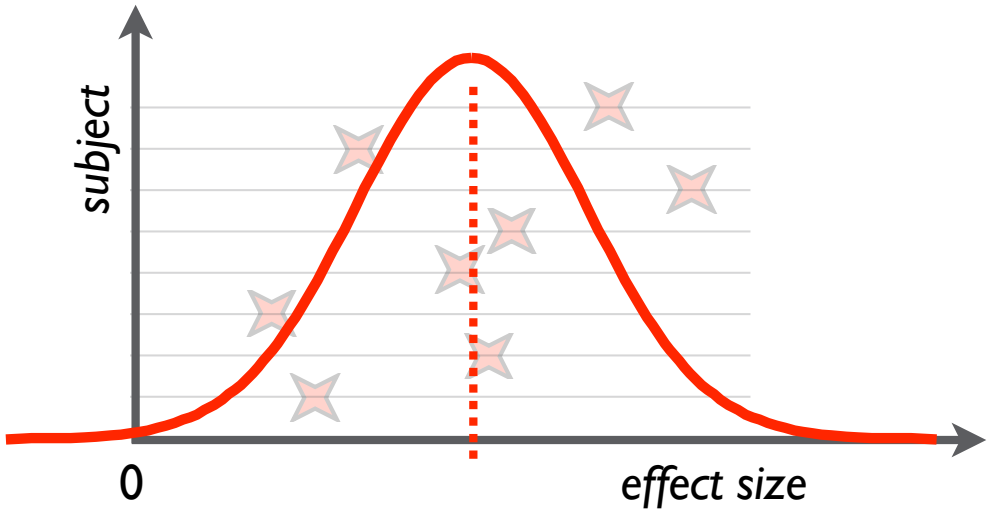


Single Group Average

单组平均值

*Does the group activate on average?*组激活是均值吗?

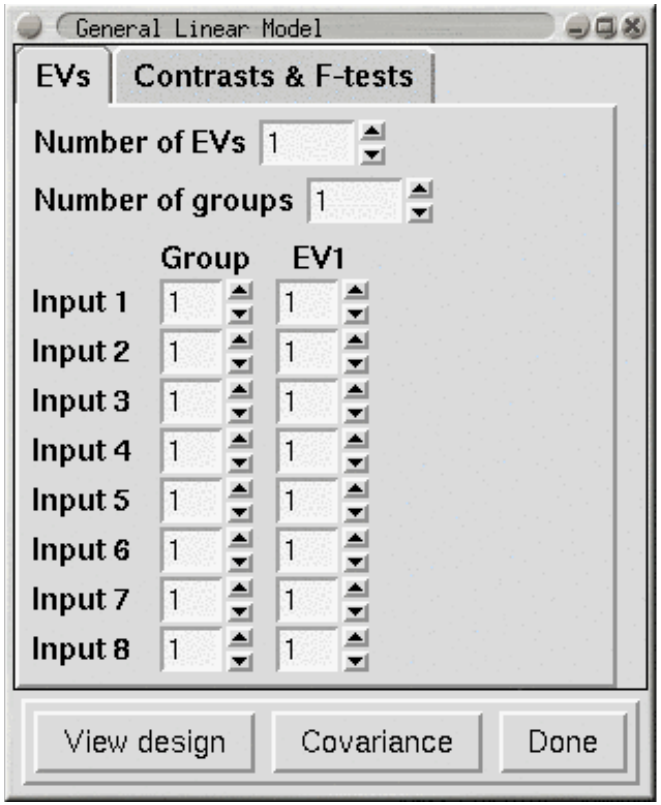


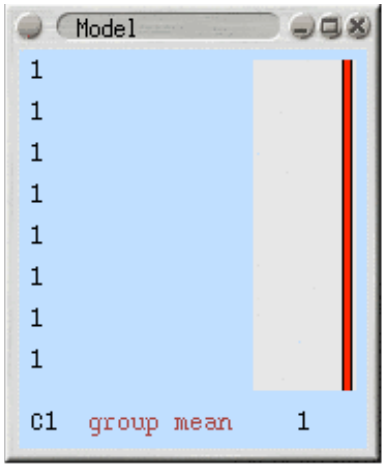


Single Group Average

单组平均值

*Does the group activate on average?*组激活是均值吗?





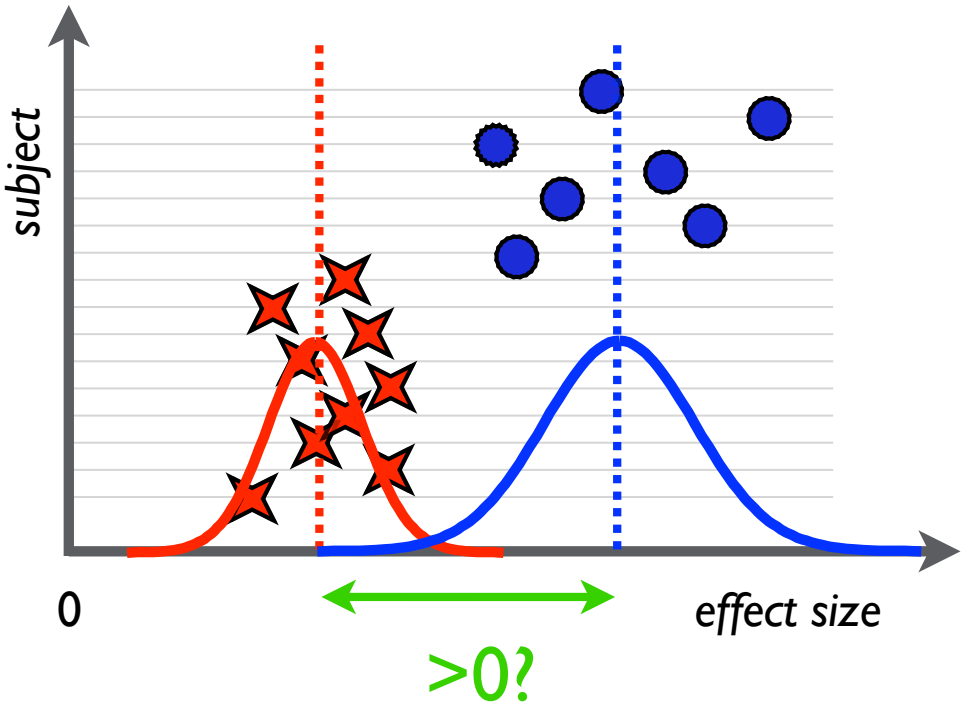
Unpaired Two-Group Difference

非配对的两组差异

- We have two groups (e.g. 9 patients, 7 controls) with different between-subject variance
  有两组（9病人，7控制组）组间方差不同

*Is there a significant group difference?*组间差异显著吗

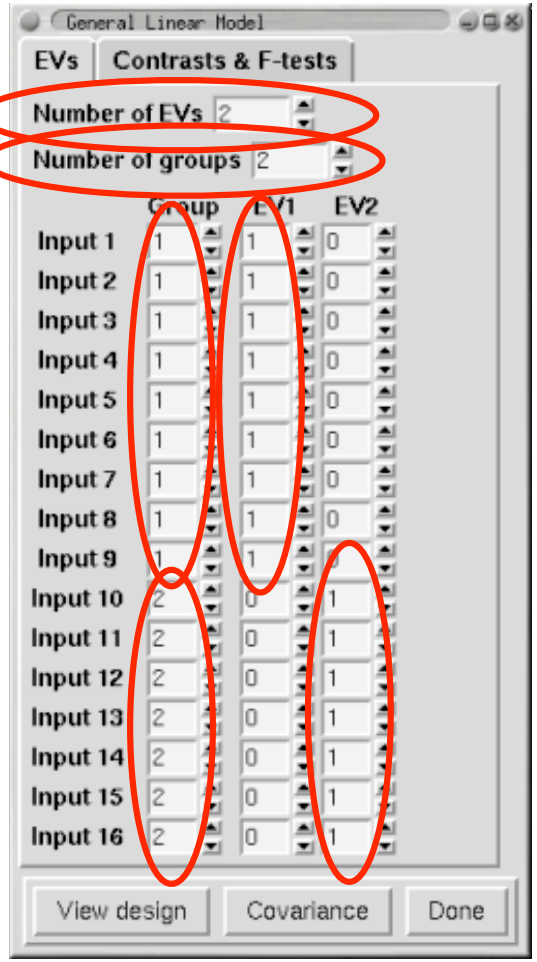
- estimate means估计均值
- estimate std-errors (FE or ME)估计标准误差
- test significance of difference in means
  检测均值差异的显著性

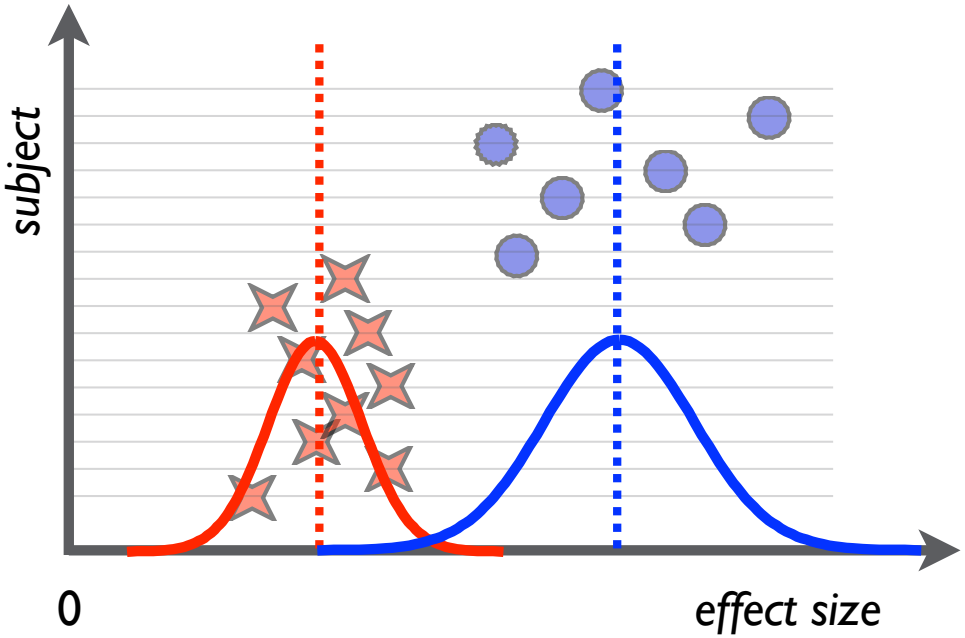


Unpaired Two-Group Difference

非配对的两组差异

*Is there a significant group difference?*组间差异显著吗

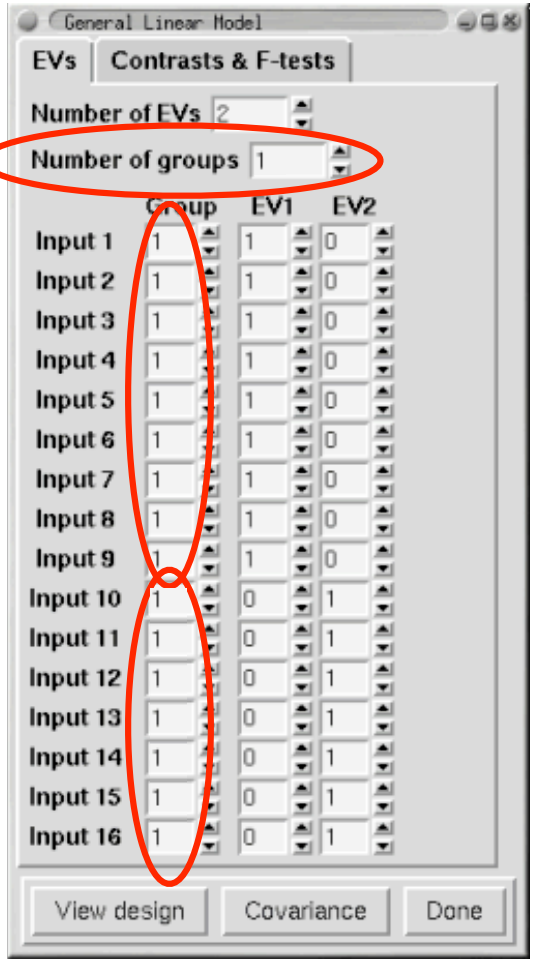


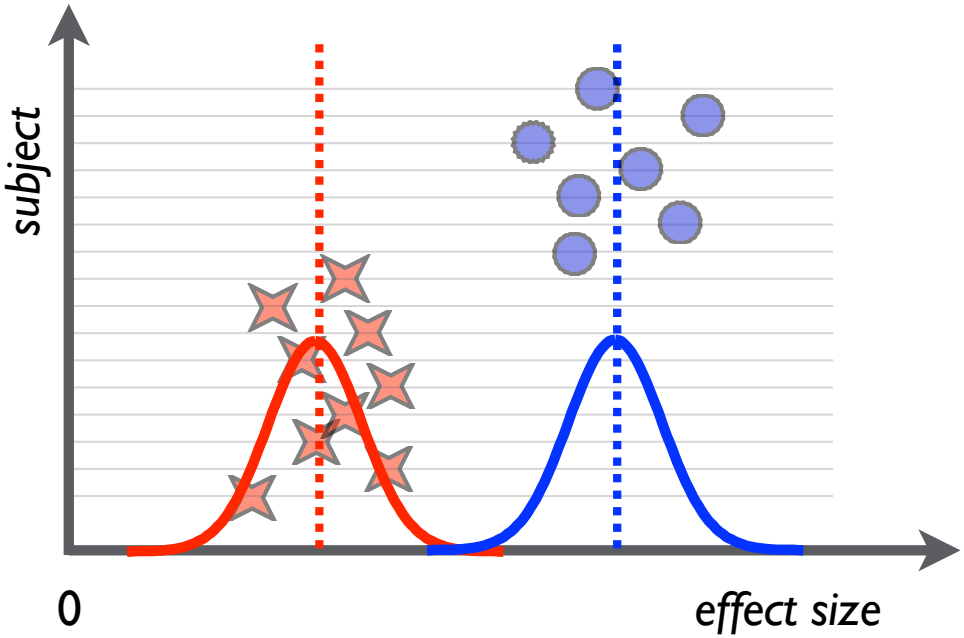


Unpaired Two-Group Difference

非配对的两组差异

*Is there a significant group difference?*组间差异显著吗

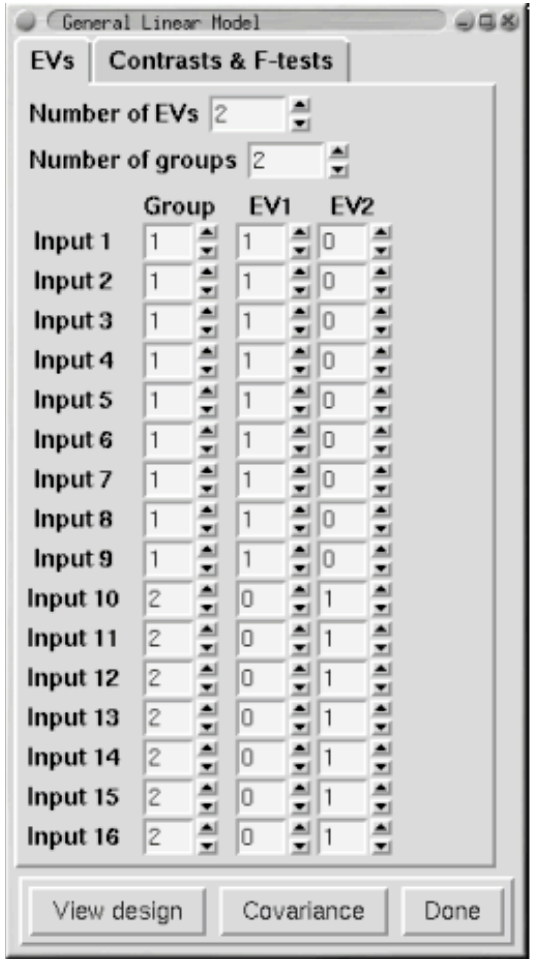


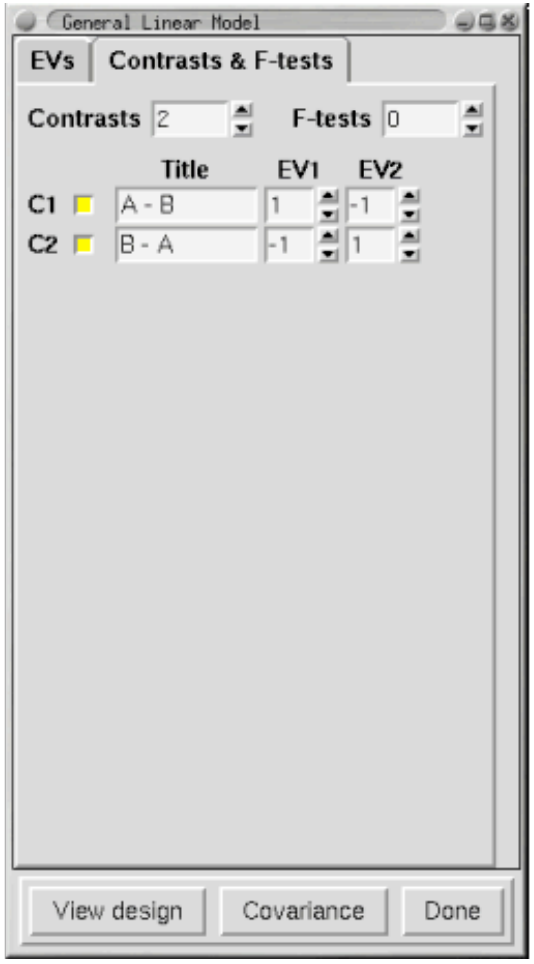


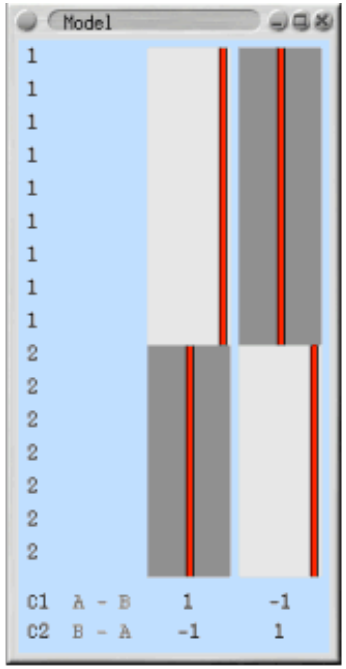
Unpaired Two-Group Difference

非配对的两组差异

*Is there a significant group difference?*组间差异显著吗







Paired T-Test

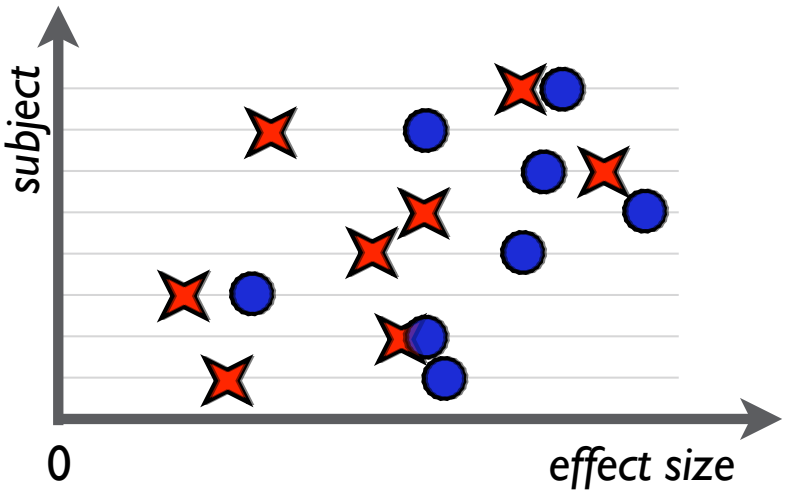
配对T检验

- 8 subjects scanned under 2 conditions (A,B)

  8个被试在AB两个条件下进行扫描

Is there a significant difference between conditions?

条件之间差异显著吗?



Paired T-Test

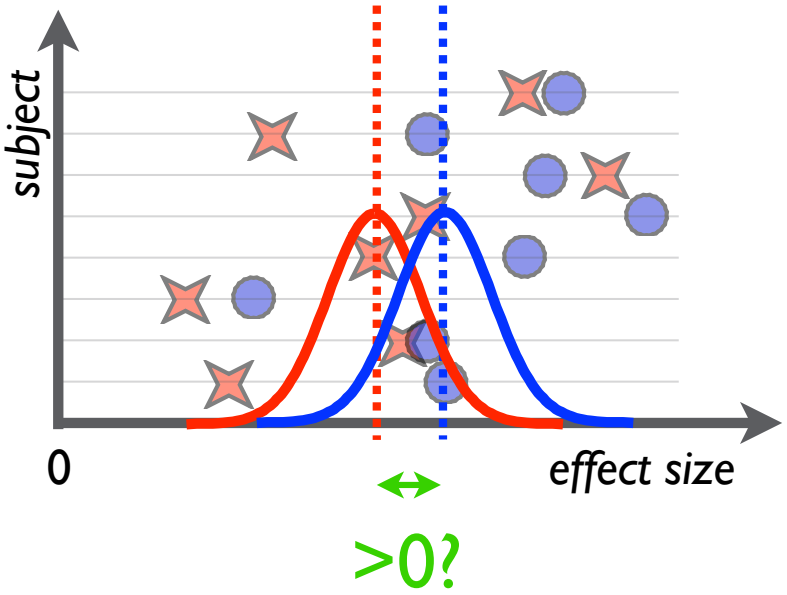
配对T检验

- 8 subjects scanned under 2 conditions (A,B)
  8个被试在AB两个条件下进行扫描

Is there a significant difference between conditions?

条件之间差异显著吗?

try non-paired t-test尝试进行非配对T检验



Paired T-Test

配对T检验

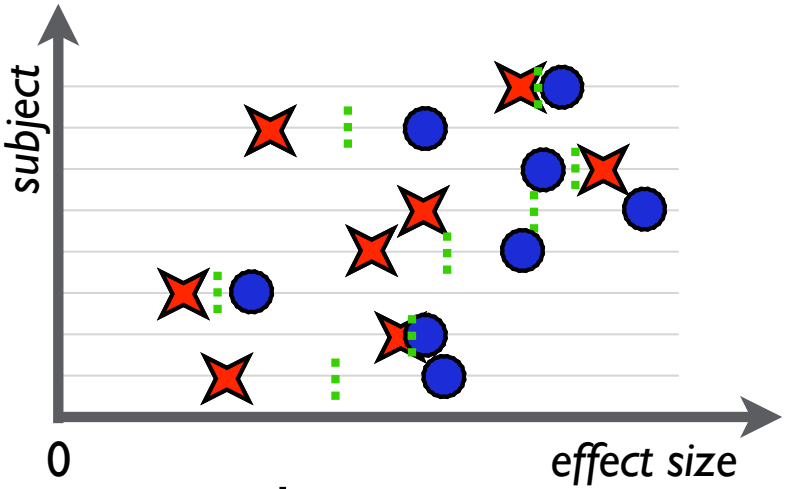
- 8 subjects scanned under 2 conditions (A,B)

  8个被试在AB两个条件下进行扫描

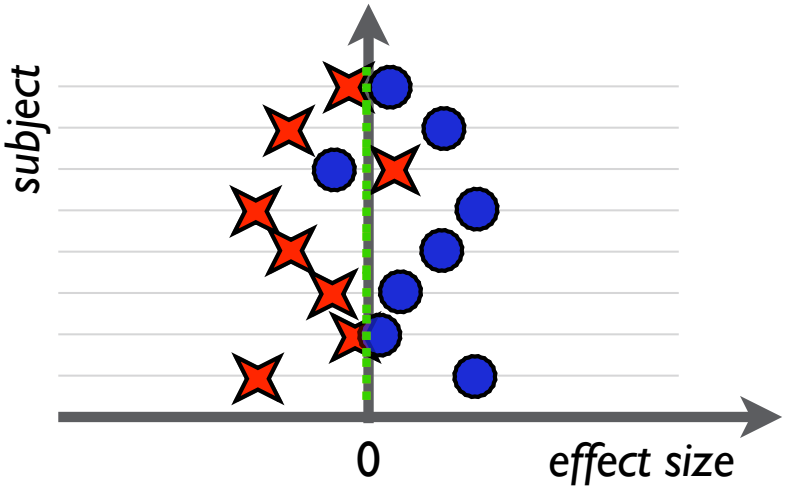
Is there a significant difference between conditions?

条件之间差异显著吗?

data



de-meaned data 去均值数据



subject mean accounts for large prop. of the overall variance

个体均值占总体方差的很大比例

Paired T-Test

配对T检验

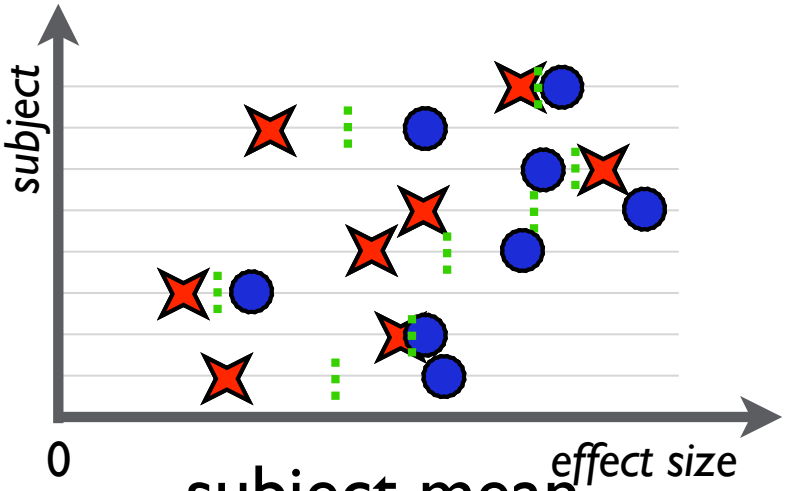
- 8 subjects scanned under 2 conditions (A,B)

  8个被试在AB两个条件下进行扫描

Is there a significant difference between conditions?

条件之间差异显著吗?

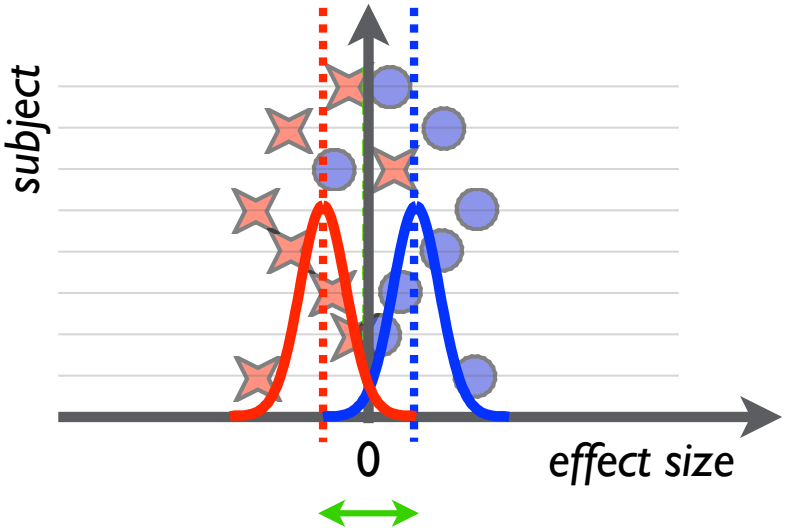
data



subject mean accounts for large prop. of the overall variance

个体均值占总体方差的很大比例

de-meaned data 去均值数据



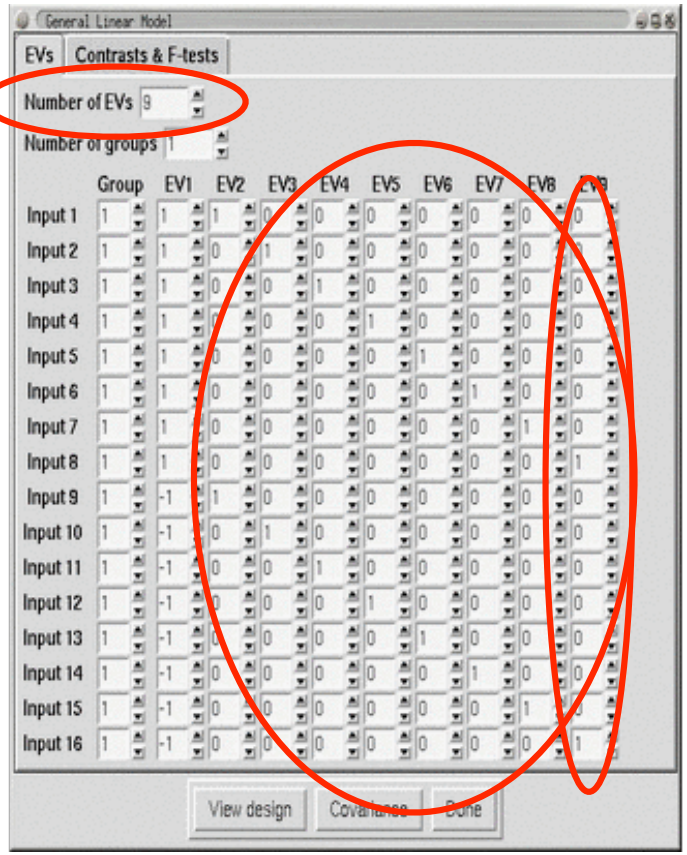
>0?

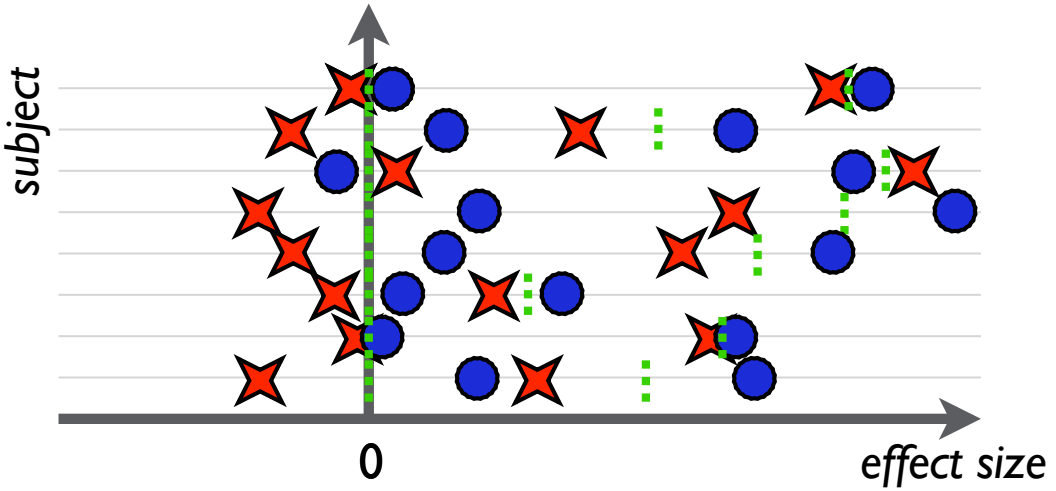
Paired T-Test

配对T检验

Is there a significant difference between conditions?

条件之间差异显著吗?



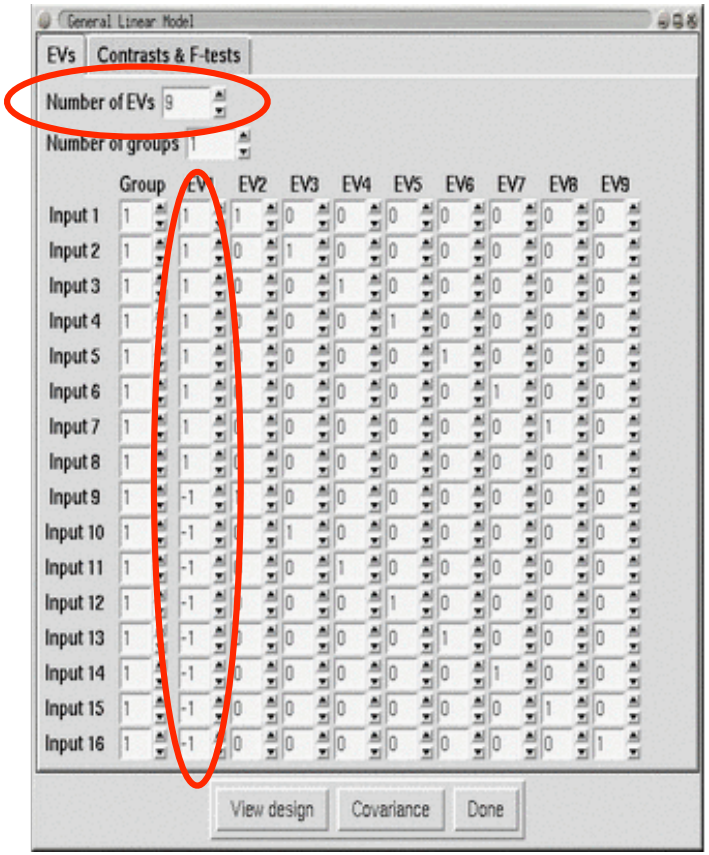


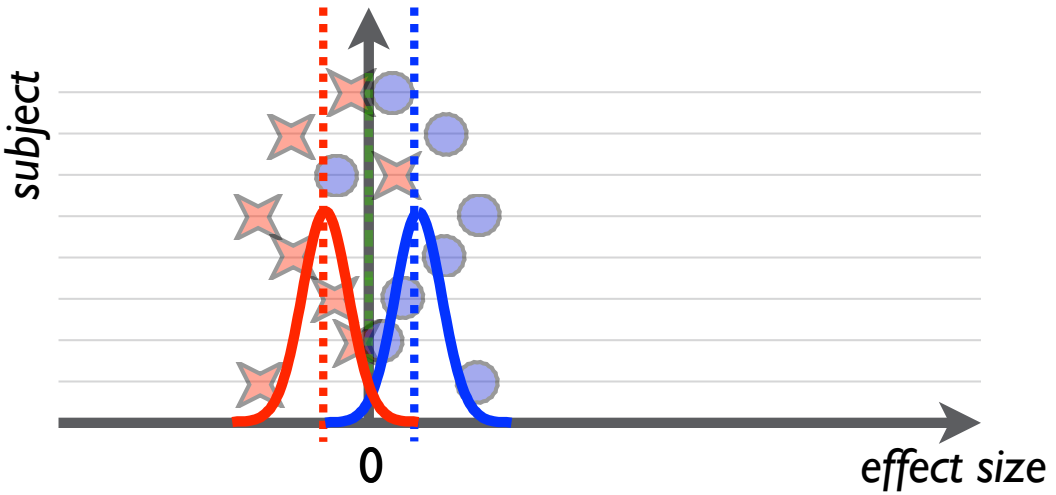
Paired T-Test

配对T检验

Is there a significant difference between conditions?

条件之间差异显著吗?



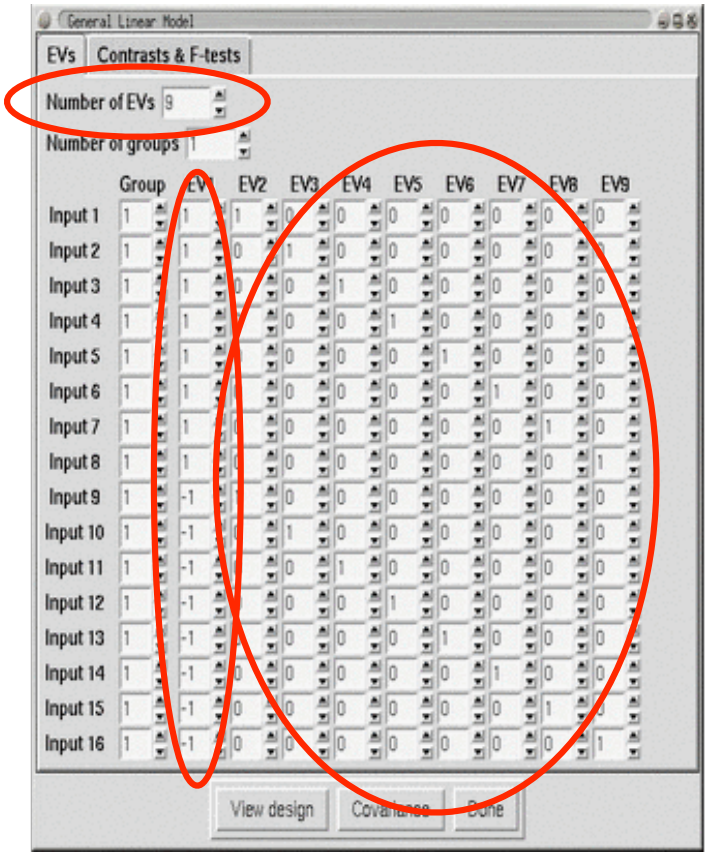


Paired T-Test

配对T检验

Is there a significant difference between conditions?

条件之间差异显著吗?



EV1models the A-B paired difference; EVs 2-9 are confounds which model out each subject's mean

EV1是A-B两配对组的差异建模；

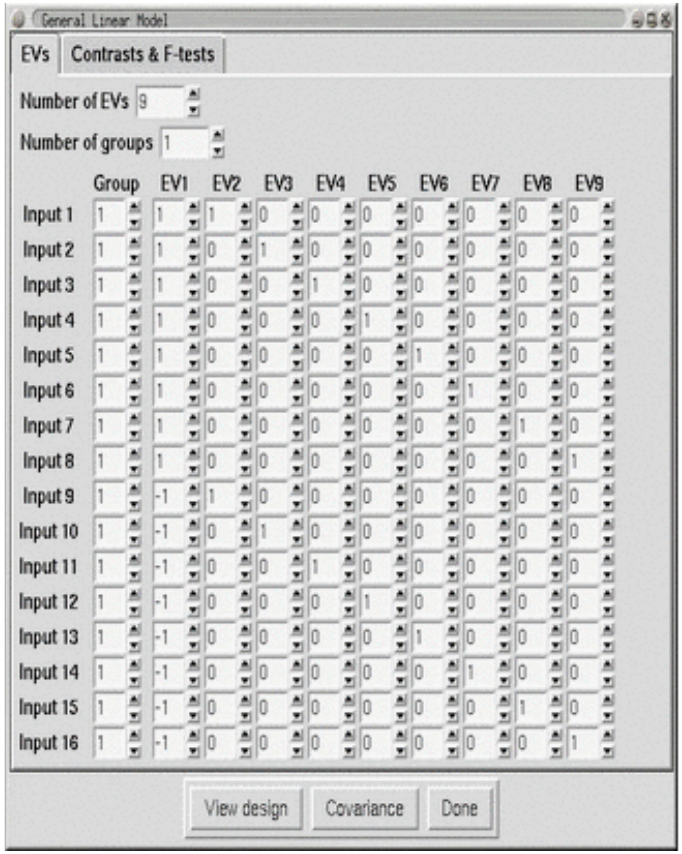
EV 2-9是混淆因素，为每个被试的均值建模

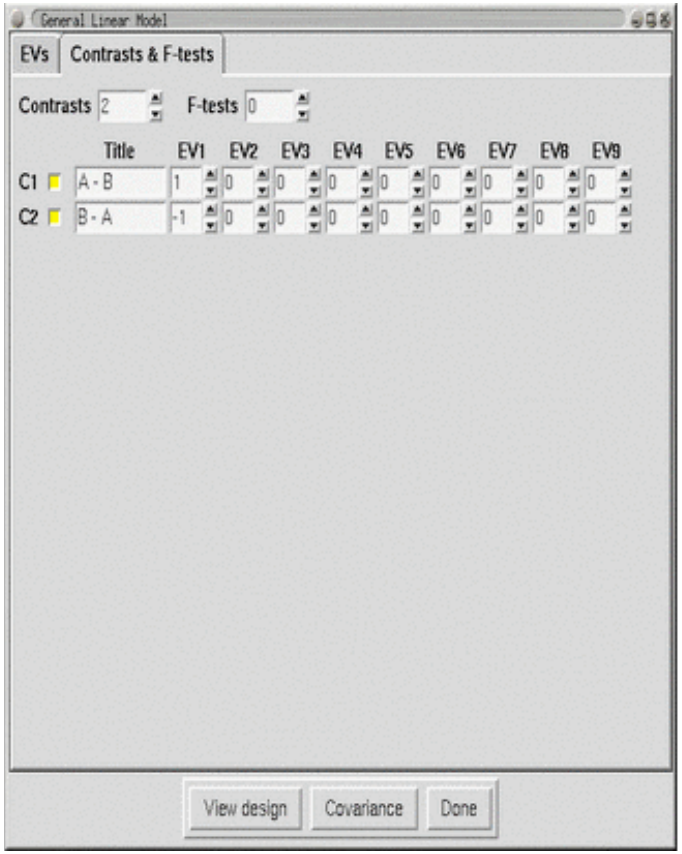
Paired T-Test

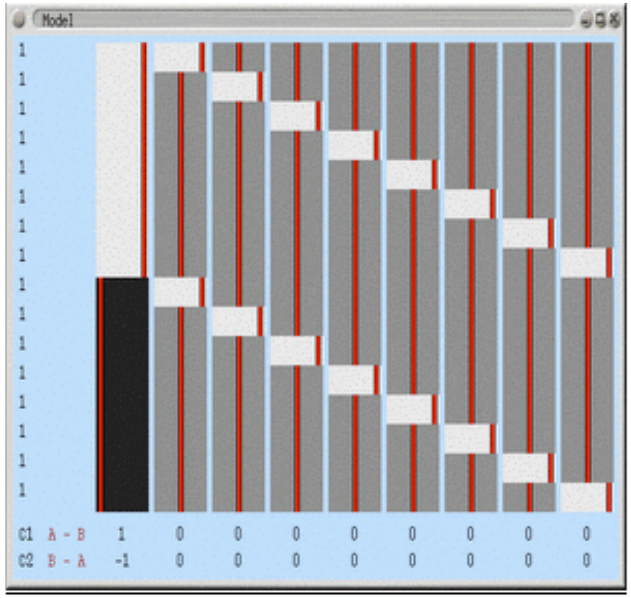
配对T检验

Is there a significant difference between conditions?

条件之间差异显著吗?







Multi-Session & Multi-Subject

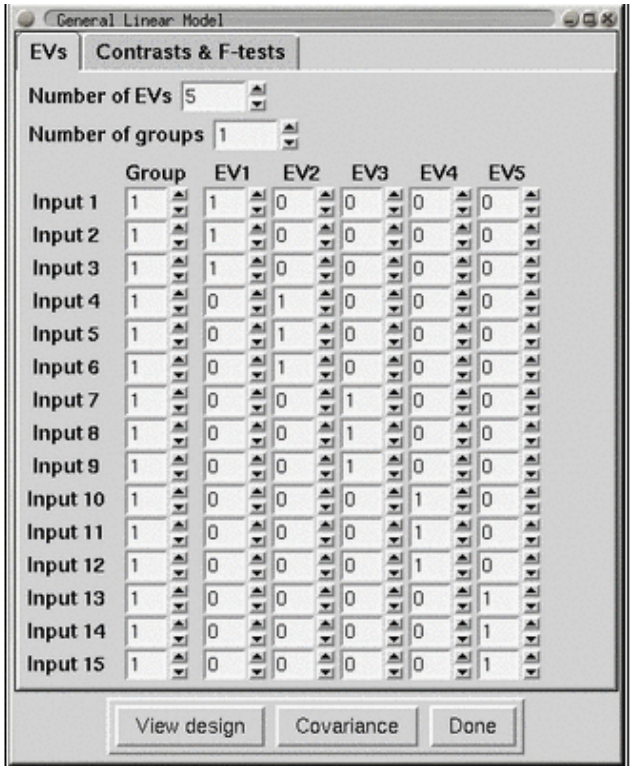
多阶段&多被试

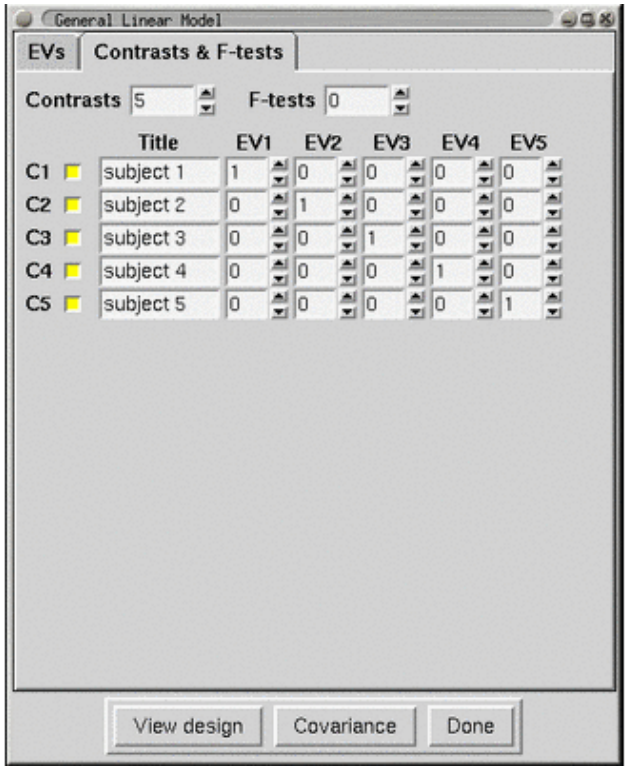
- 5 subjects each have three sessions.5个被试，每人3阶段

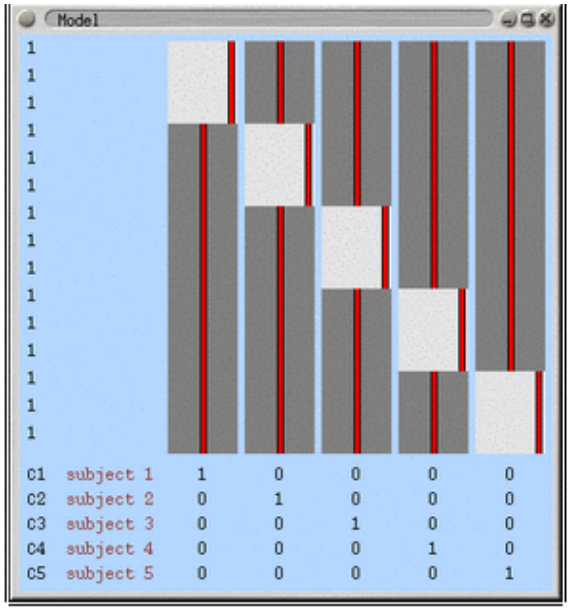
*Does the group activate on average?*组激活是平均值吗?

- Use three levels: in the **second level** we model the within-subject repeated measure

使用三个水平：在第二水平中，我们对被试内重复测量进行建模







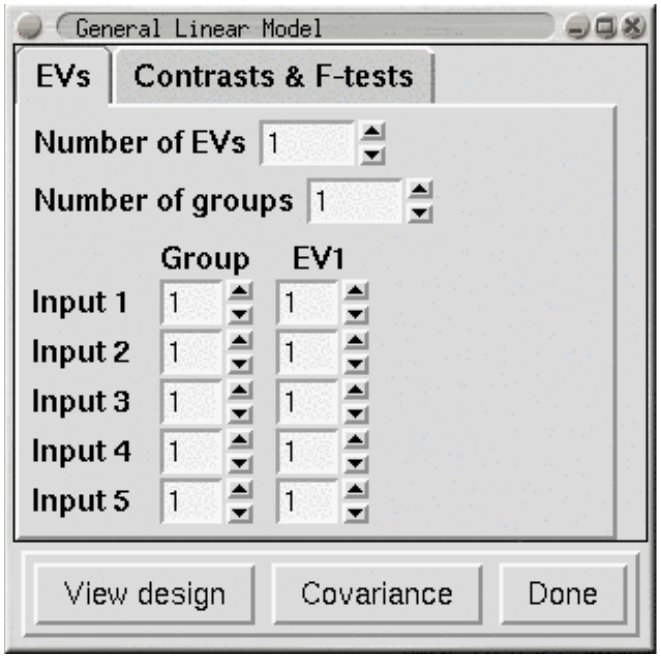
Multi-Session & Multi-Subject

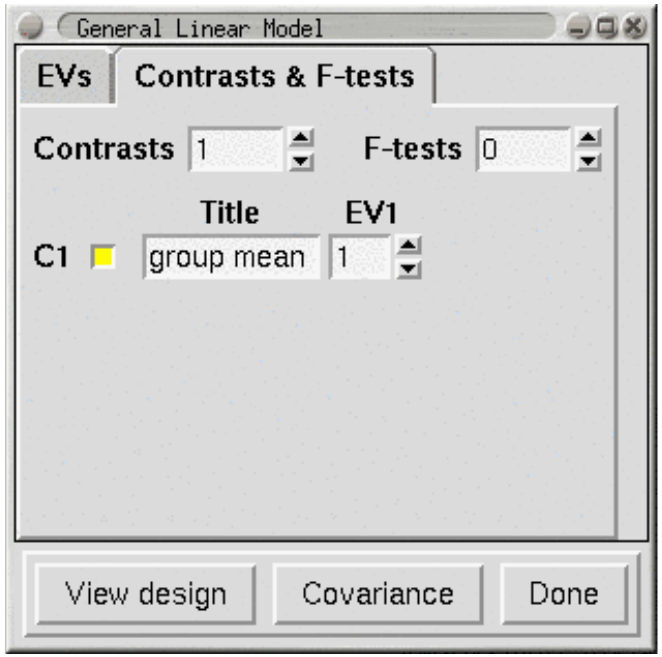
多阶段&多被试

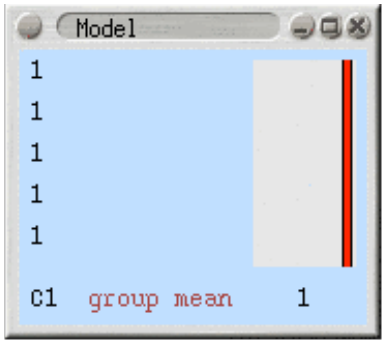
- 5 subjects each have three sessions.5个被试，每人3阶段

  *Does the group activate on average?*组激活是平均值吗?

- Use three levels: in the **third level** we model the between-subjects variance 在第三水平对被试间变异建模







Multi-Session & Multi-Subject

多阶段&多被试

- 5 subjects each have three sessions.5个被试，每人3阶段

  *Does the group activate on average?*组激活是平均值吗?

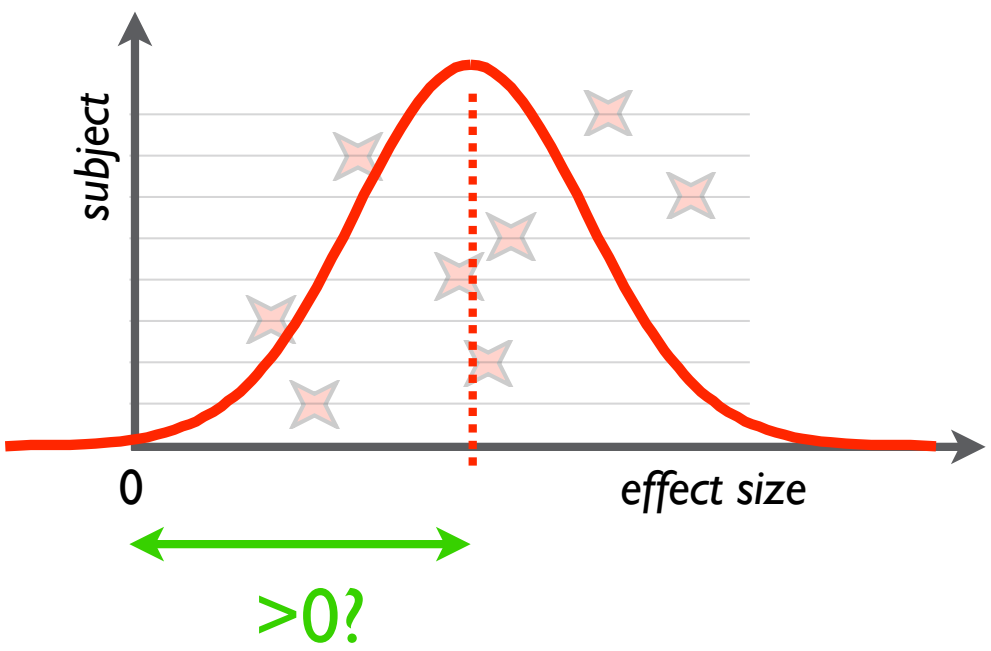
- Use three levels:
  - in the second level we model the within subject repeated measure typically using **fixed effects**(!) as #sessions are small

    在第二水平中，由于每阶段较小，我们通常使用固定效应对被试内的重复测量进行建模
  - in the third level we model the between subjects variance using fixed or mixed effects

    在第三水平使用固定效应或混合效应对被试间差异进行建模

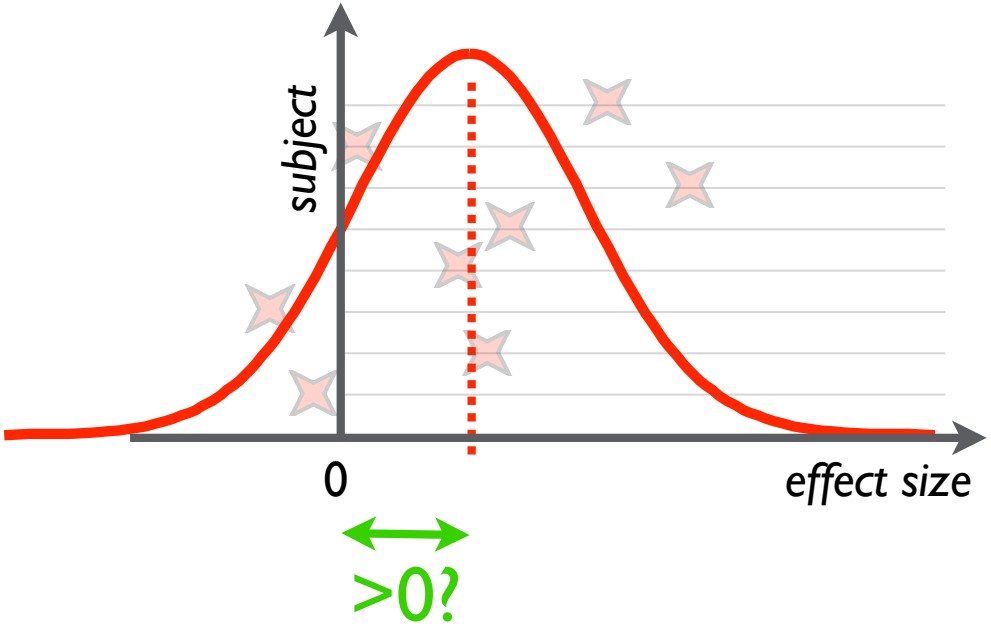
Reducing variance
减少方差

*Does the group activate on average?*组激活是平均值吗?



mean effect size large relative to std error

相对于标准误，平均效应量较大

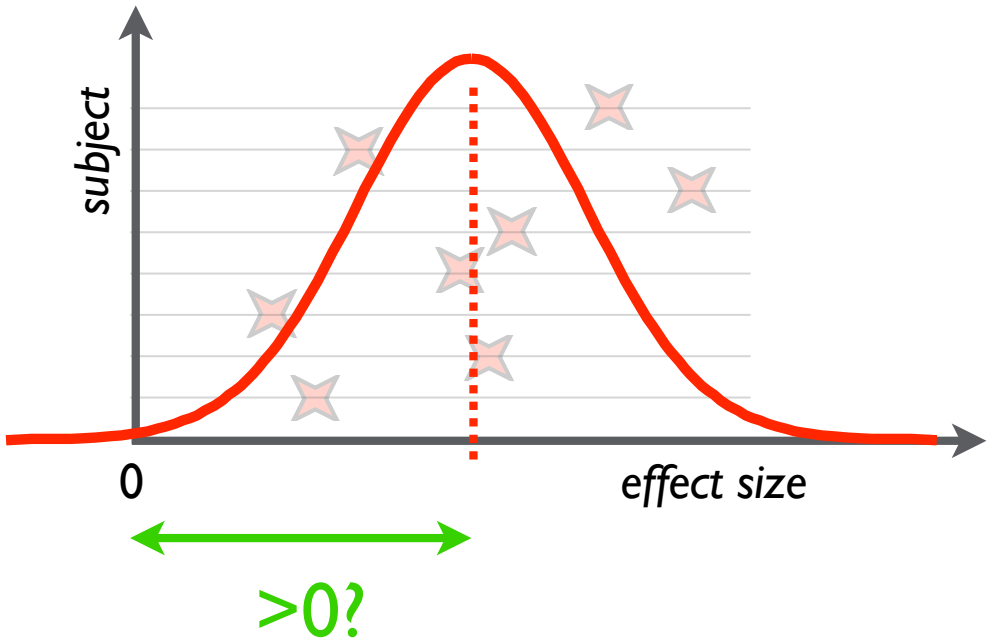


mean effect size small relative to std error

相对于标准误，平均效应量较小

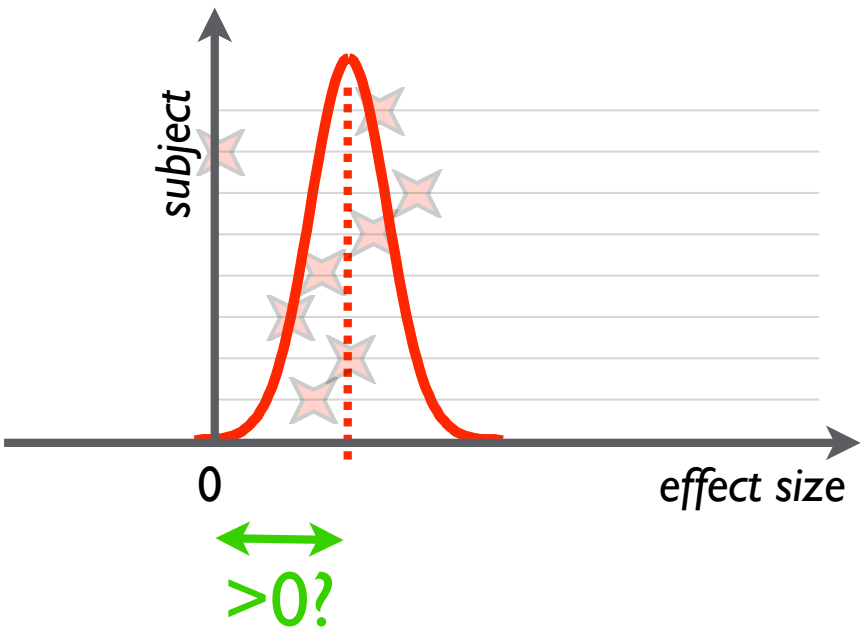
Reducing variance
减少方差

*Does the group activate on average?*组激活是平均值吗?



mean effect size large relative to std error

相对于标准误，平均效应量较大



mean effect size small relative to std error

相对于标准误，平均效应量较小

Single Group Average & Covariates

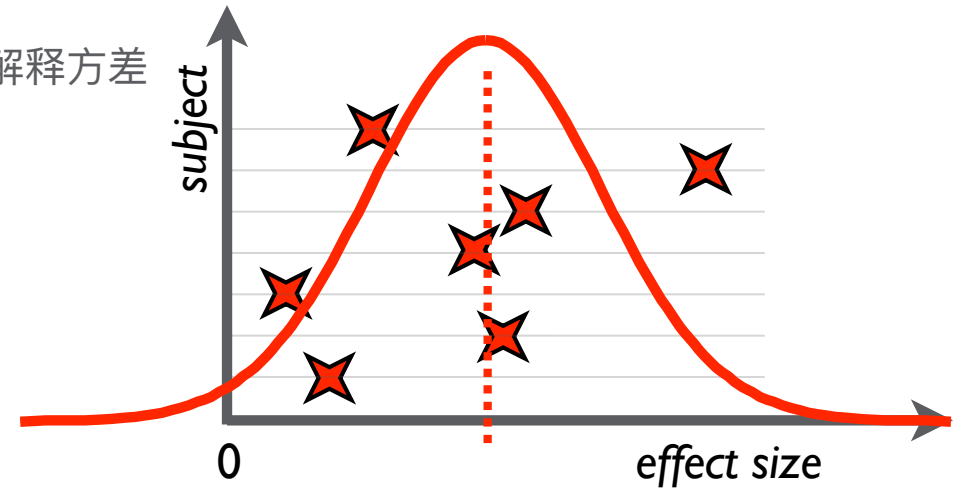
单组平均值和协变量

- We have 7 subjects - all in one group. We also have additional measurements (e.g. age; disability score; behavioural measures like reaction times):

  一组有7个被试，并且提供其他指标（例如年龄，残疾评分，行为指标（例如反应时间））

Does the group activate on average? 组激活是平均值吗?

- use covariates to 'explain' variation 用协变量解释方差
- estimate mean 估计均值
- estimate std-error (FE or ME) 估计标准误



Single Group Average & Covariates

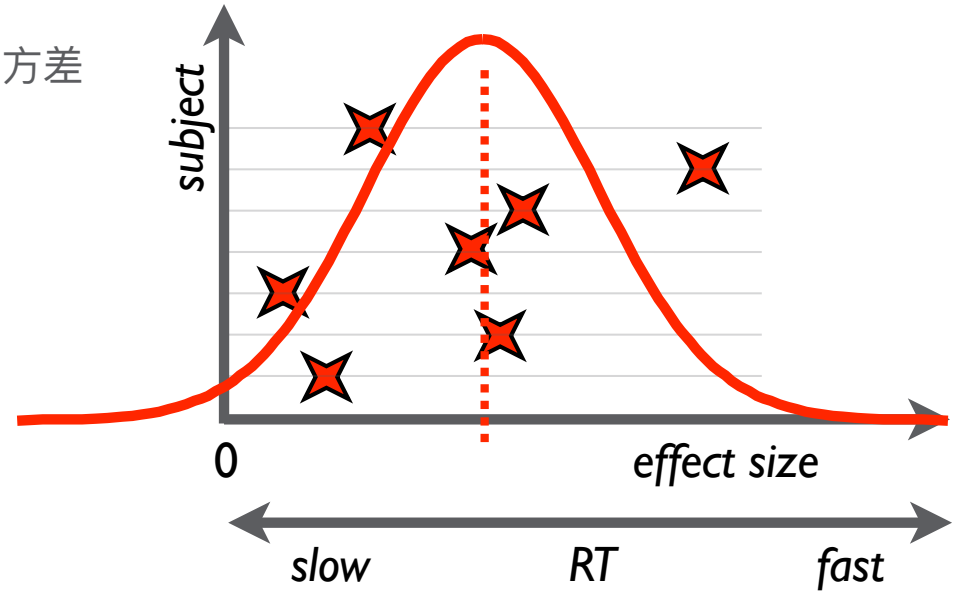
单组平均值和协变量

- We have 7 subjects - all in one group. We also have additional measurements (e.g. age; disability score; behavioural measures like reaction times):

一组有7个被试，并且提供其他指标（例如年龄，残疾评分，行为指标（例如反应时间））

*Does the group activate on average?*组激活是平均值吗?

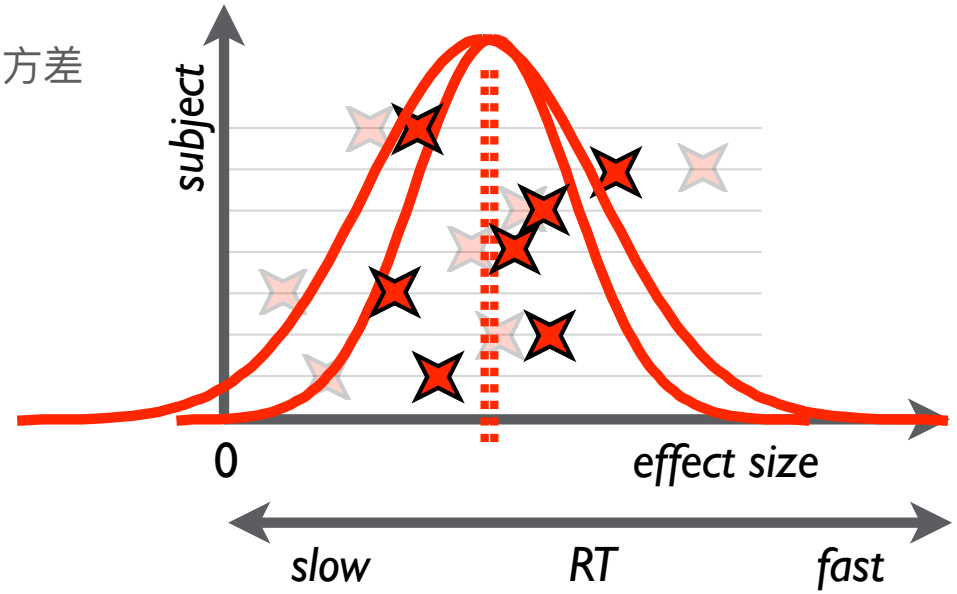
- use covariates to 'explain' variation用协变量解释方差
- estimate mean估计均值
- estimate std-error (FE or ME)估计标准误



Single Group Average & Covariates

单组平均值和协变量

- We have 7 subjects - all in one group. We also have additional measurements (e.g. age; disability score; behavioural measures like reaction times):
  一组有7个被试，并且提供其他指标（例如年龄，残疾评分，行为指标（例如反应时间））
- *Does the group activate on average?* 组激活是平均值吗?
- use covariates to 'explain' variation 用协变量解释方差
- estimate mean 估计均值
- estimate std-error (FE or ME) 估计标准误



Single Group Average & Covariates

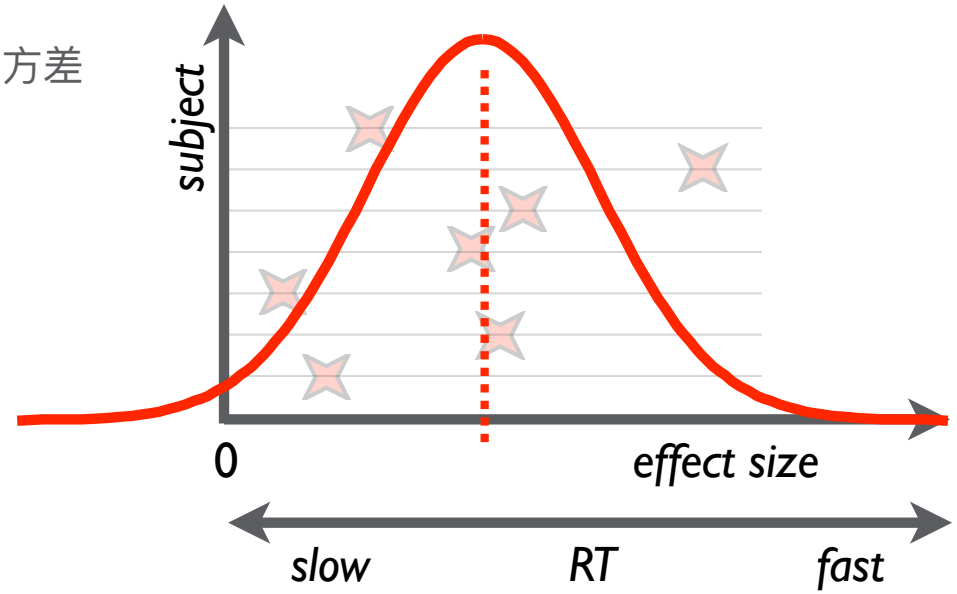
单组平均值和协变量

- We have 7 subjects - all in one group. We also have additional measurements (e.g. age; disability score; behavioural measures like reaction times):

一组有7个被试，并且提供其他指标（例如年龄，残疾评分，行为指标（例如反应时间））

*Does the group activate on average?*组激活是平均值吗?

- use covariates to 'explain' variation用协变量解释方差
- estimate mean估计均值
- estimate std-error (FE or ME)估计标准误



Single Group Average & Covariates

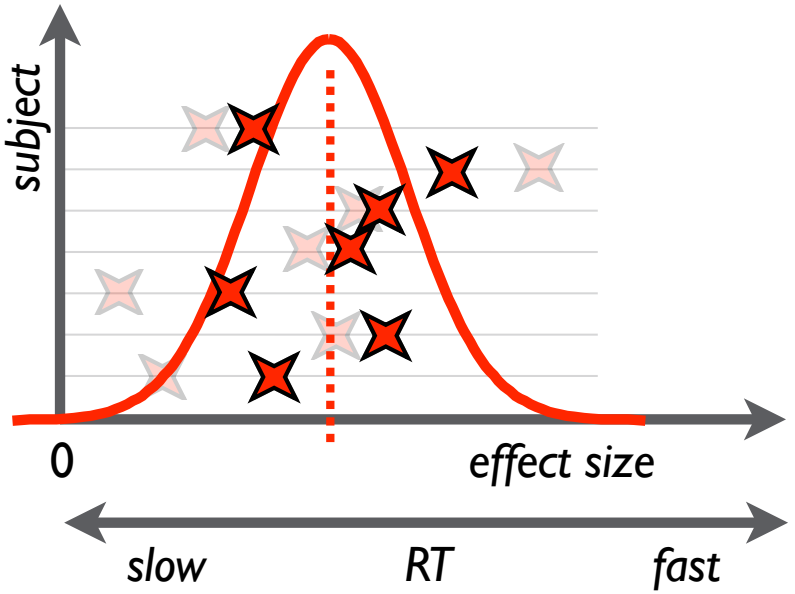
单组平均值和协变量

- We have 7 subjects - all in one group. We also have additional measurements (e.g. age; disability score; behavioural measures like reaction times):

一组有7个被试，并且提供其他指标（例如年龄，残疾评分，行为指标（例如反应时间））

*Does the group activate on average?*组激活是平均值吗?

- use covariates to 'explain' variation用协变量解释方差
- estimate mean估计均值
- estimate std-error (FE or ME)估计标准误

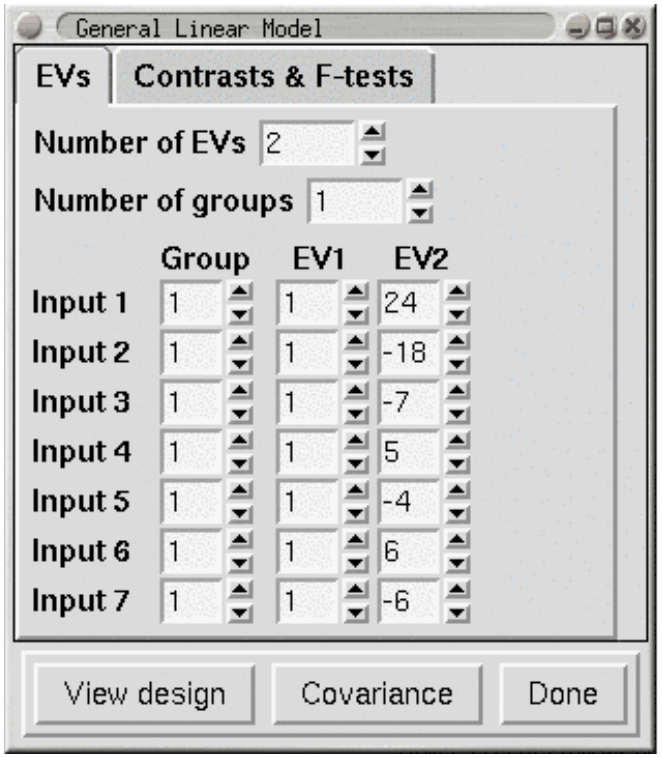


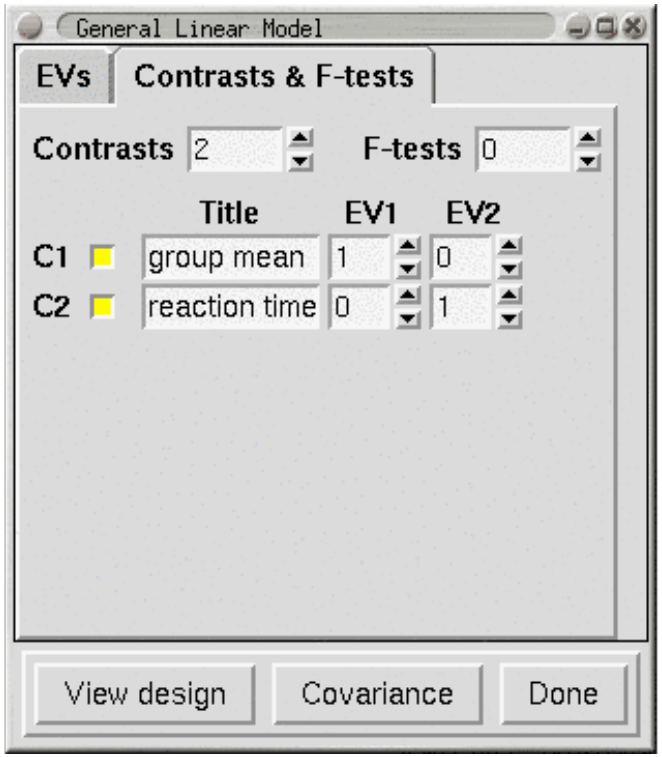
Single Group Average & Covariates

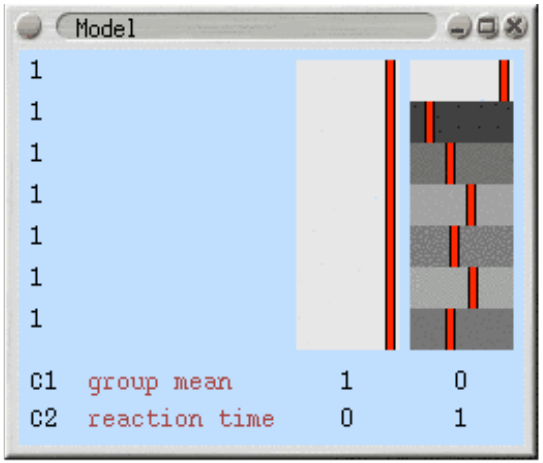
单组平均值和协变量

Does the group activate on average? 组激活是平均值吗?

- use covariates to 'explain' variation 用协变量解释方差
- need to de-mean additional covariates!
  需要为额外的协变量去均值!





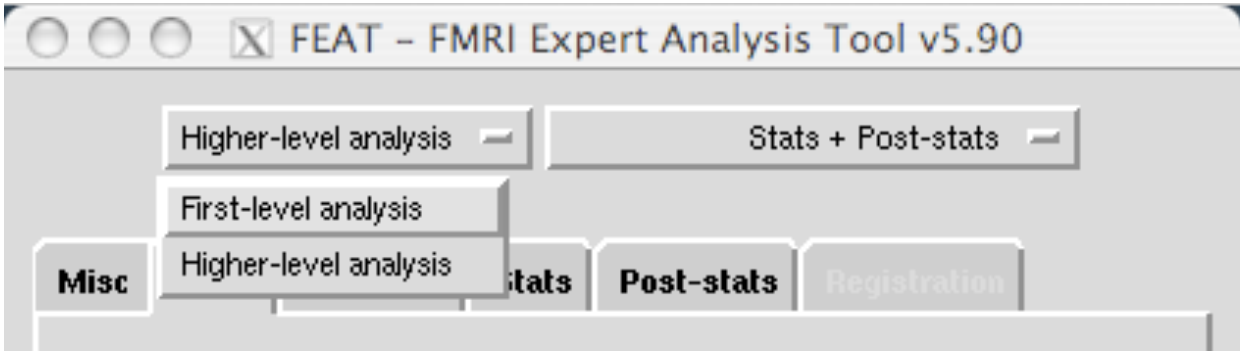


FEAT Group Analysis

FEAT组分析

- Run FEAT on raw FMRI data to get first-level .feat directories, each one with several (consistent) COPEs

  对原始FMRI数据运行FEAT以获取第一级.feat目录，每个目录均包含多个（一致的）COPE



  - low-res copeN/varcopeN → .feat/stats

    低分辨率 contrastN的参数估计值对比图像 /contrastN的方差图像 在.feat/stats文件夹下

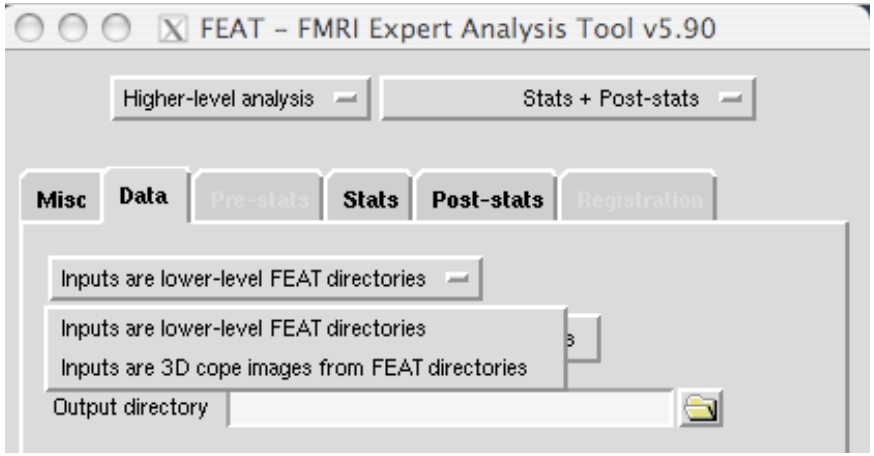
  - when higher-level FEAT is run, highres copeN/varcopeN → .feat/reg_standard

    当高水平FEAT运行时，高分辨率contrastN的参数估计值对比图像 /contrastN的方差图像 在.feat/reg_standard文件夹下

FEAT Group Analysis

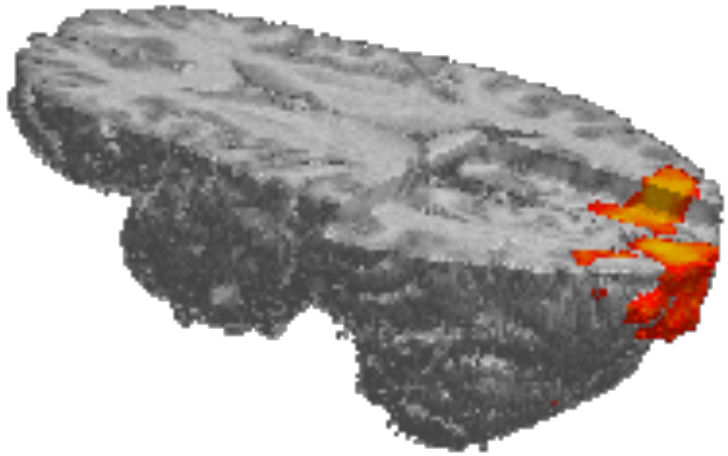
FEAT组分析

- Run second-level FEAT to get one .gfeat directory
  运行第二水平FEAT可以得到一个.gfeat文件夹
  - Inputs can be lower-level .feat dirs or lower-level COPEs
    可以输入低水平的.feat文件夹
    或者低水平的COPE文件



  - the second-level GLM analysis is run separately for each first-level COPE
    分别对每个第一水平的COPE图像运行第二水平的GLM分析
  - each lower-level COPE generates its own .feat directory inside the .gfeat dir
    每个低水平的COPE可以在.gfeat文件夹下生成一个单独的.feat文件夹

That's all folks



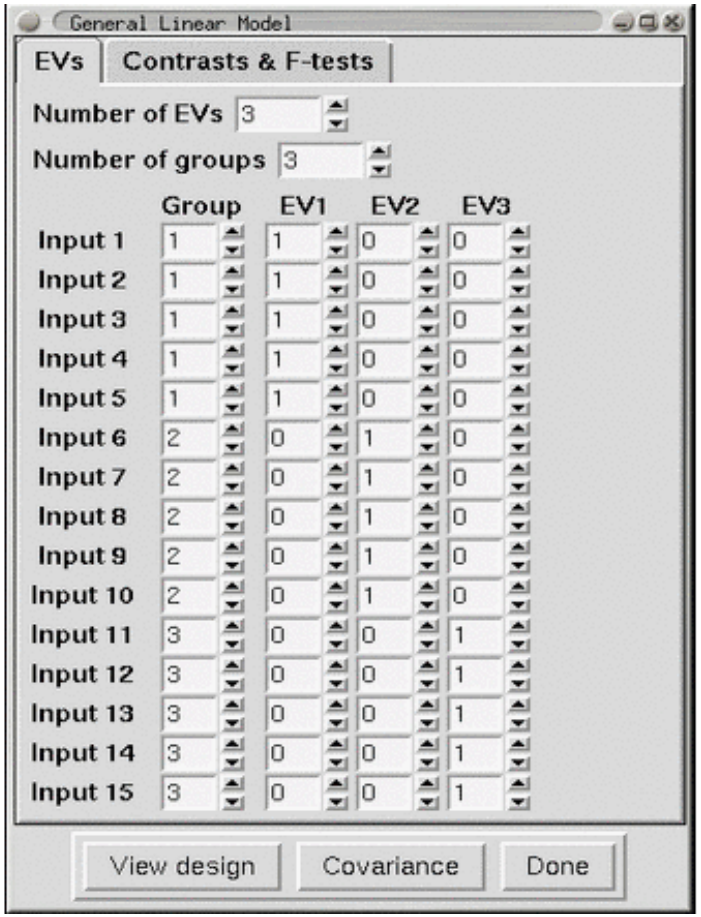
Appendix:

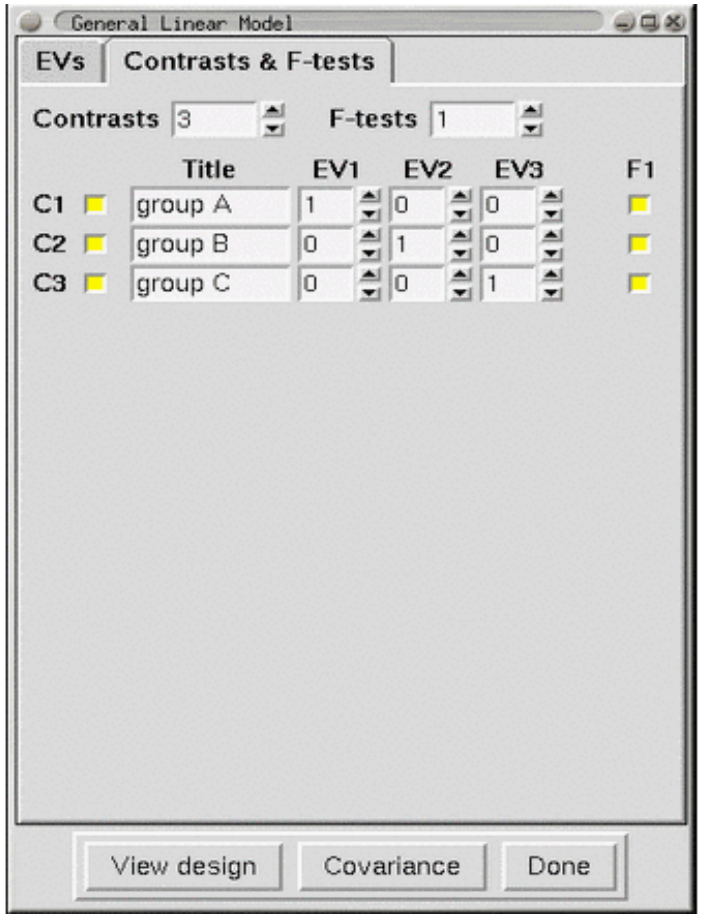
Group F-tests

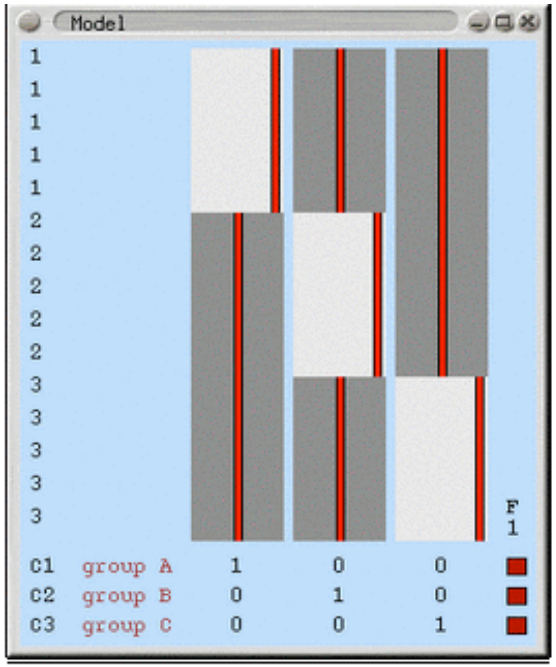
组水平F检验

- 3 groups of subjects3组被试

*Is any of the groups activating on average?*是否有组平均激活?







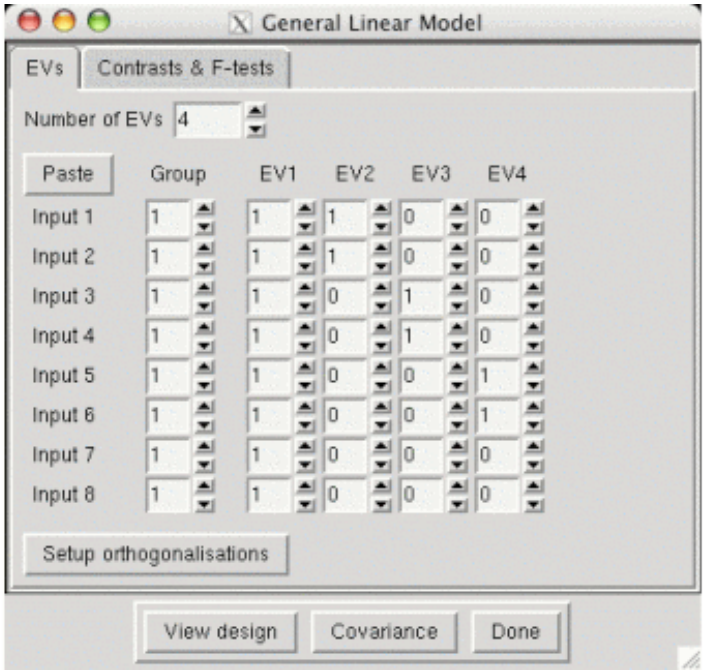
ANOVA: 1-factor 4-levels

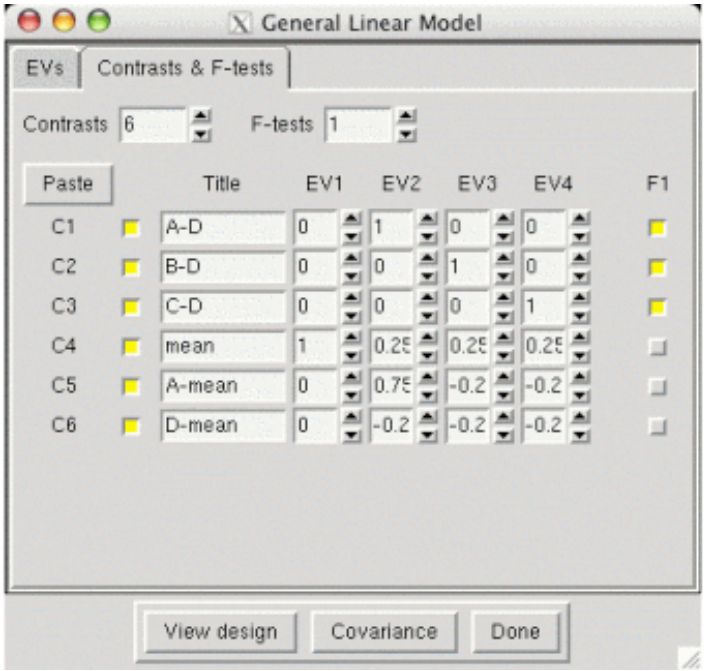
单因素4水平方差分析

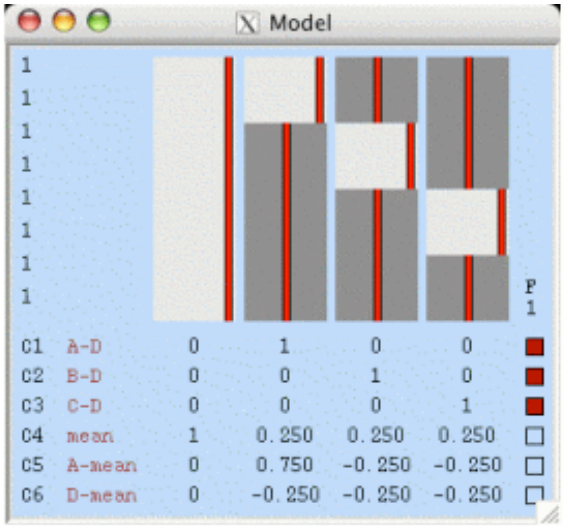
- 8 subjects, 1 factor at 4 levels
  8个被试，单因素4水平

Is there any effect? 是否存在什么效应?

- EV1 fits cond. D, EV2 fits cond A relative to D etc.
  EV1拟合条件D，EV2拟合条件A相对于D的差异等





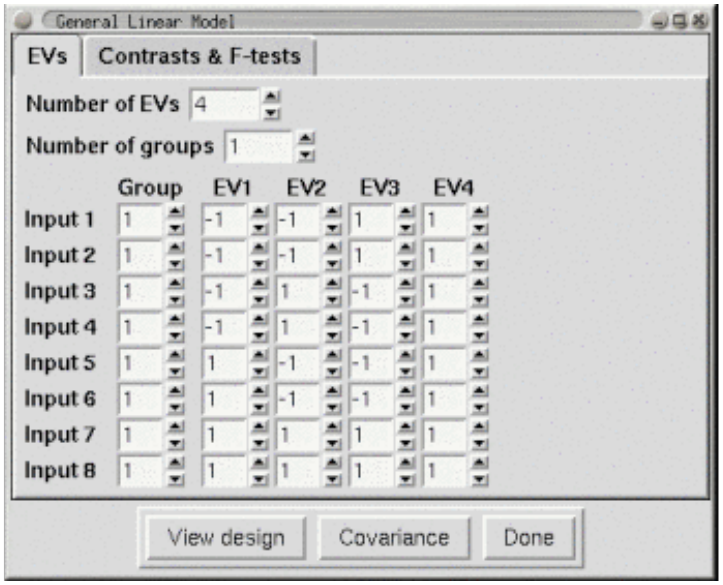


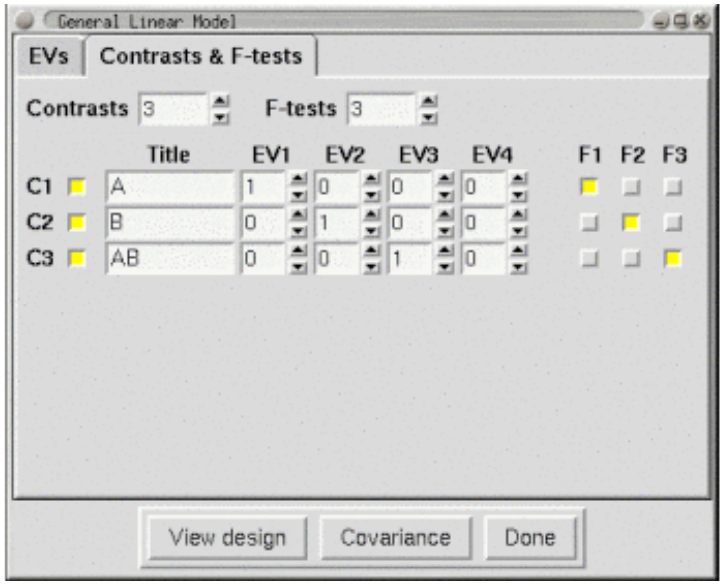
ANOVA: 2-factor 2-levels

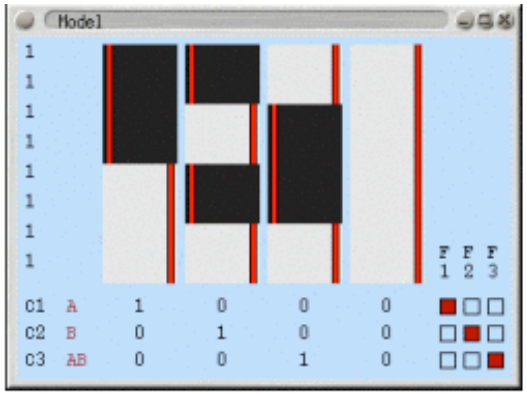
2因素2水平方差分析

- 8 subjects, 2 factor at 2 levels. FE Anova: 3 F-tests give standard results for factor A, B and interaction

  8个被试，2因素2水平。运行3次F检验可得到因素A,B以及交互作用的标准结果







- If both factors are random effects then Fa=fstat1/fstat3, Fb=fstat2/fstat3. 如果两个因素都有随机效应，则Fa=fstat1/fstat3,Fb=fstat2/fstat3

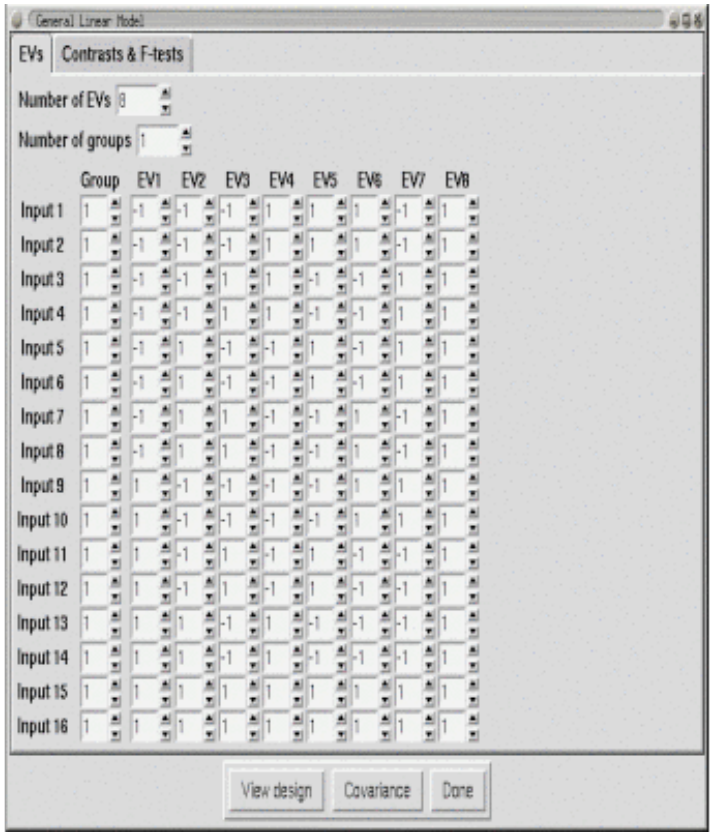
- ME: if fixed fact. is A, Fa=fstat1/fstat3

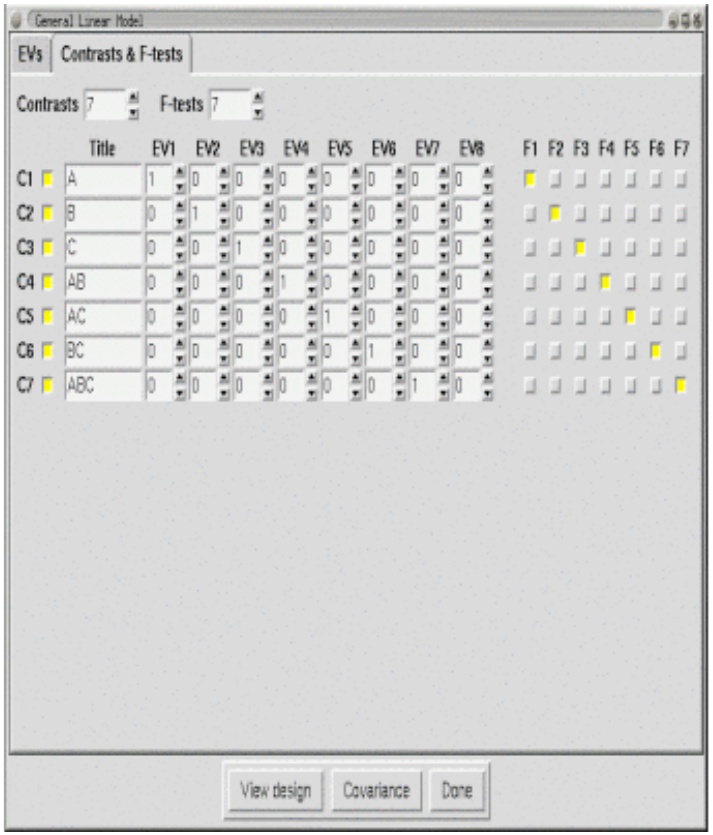
  如果固定效应是A，那么Fa=fstat1/fstat3

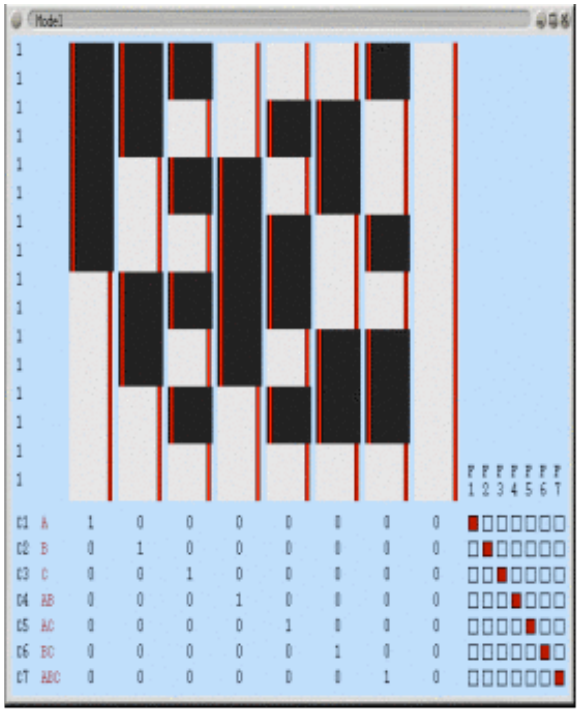
ANOVA: 3-factor 2-levels

3因素2水平方差分析

- 16 subjects, 3 factor at 2 levels. 16个被试，3因素2水平
- Fixed-Effects ANOVA:固定效应方差分析：







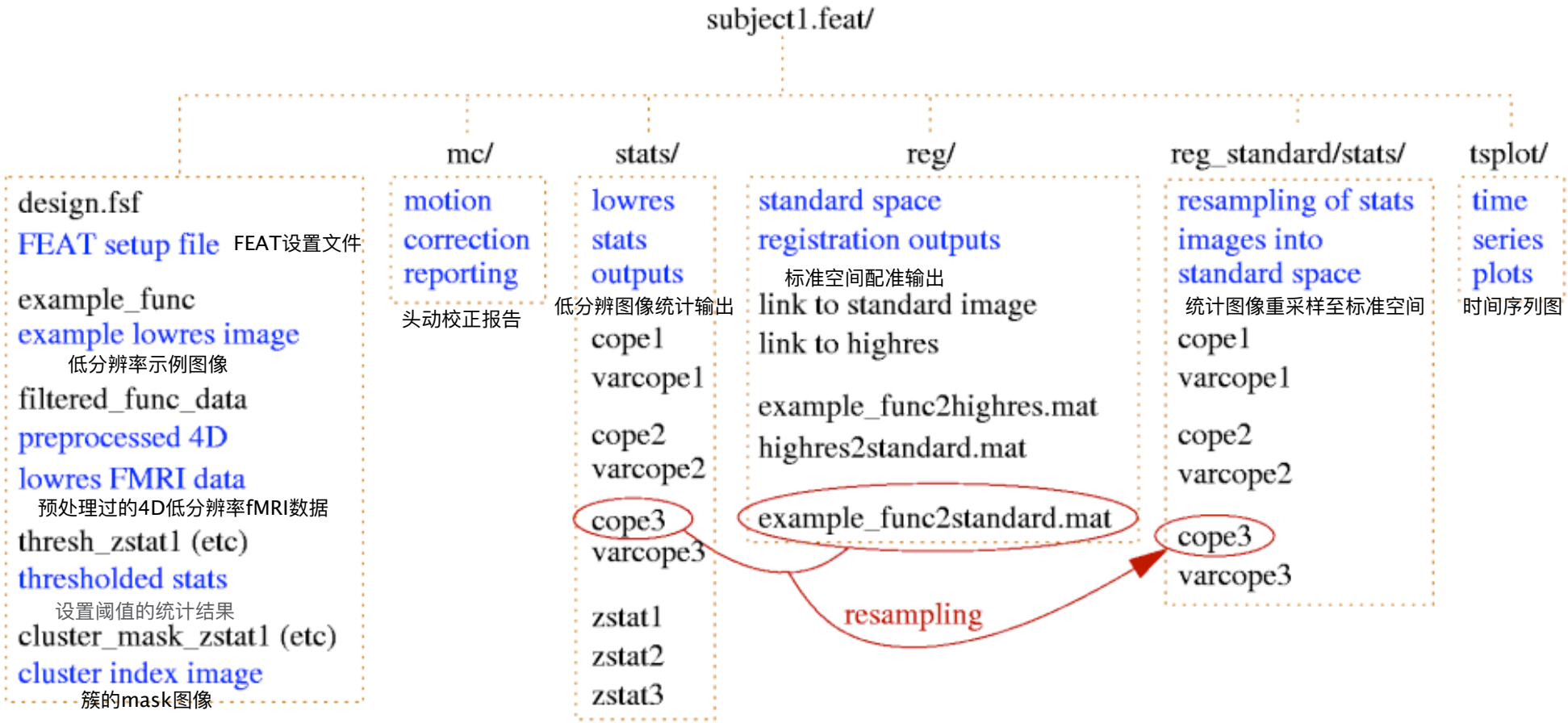
- For random/mixed effects need different Fs.

随机/混合效应需要不同的Fs

Understanding FEAT dirs

理解FEAT文件夹

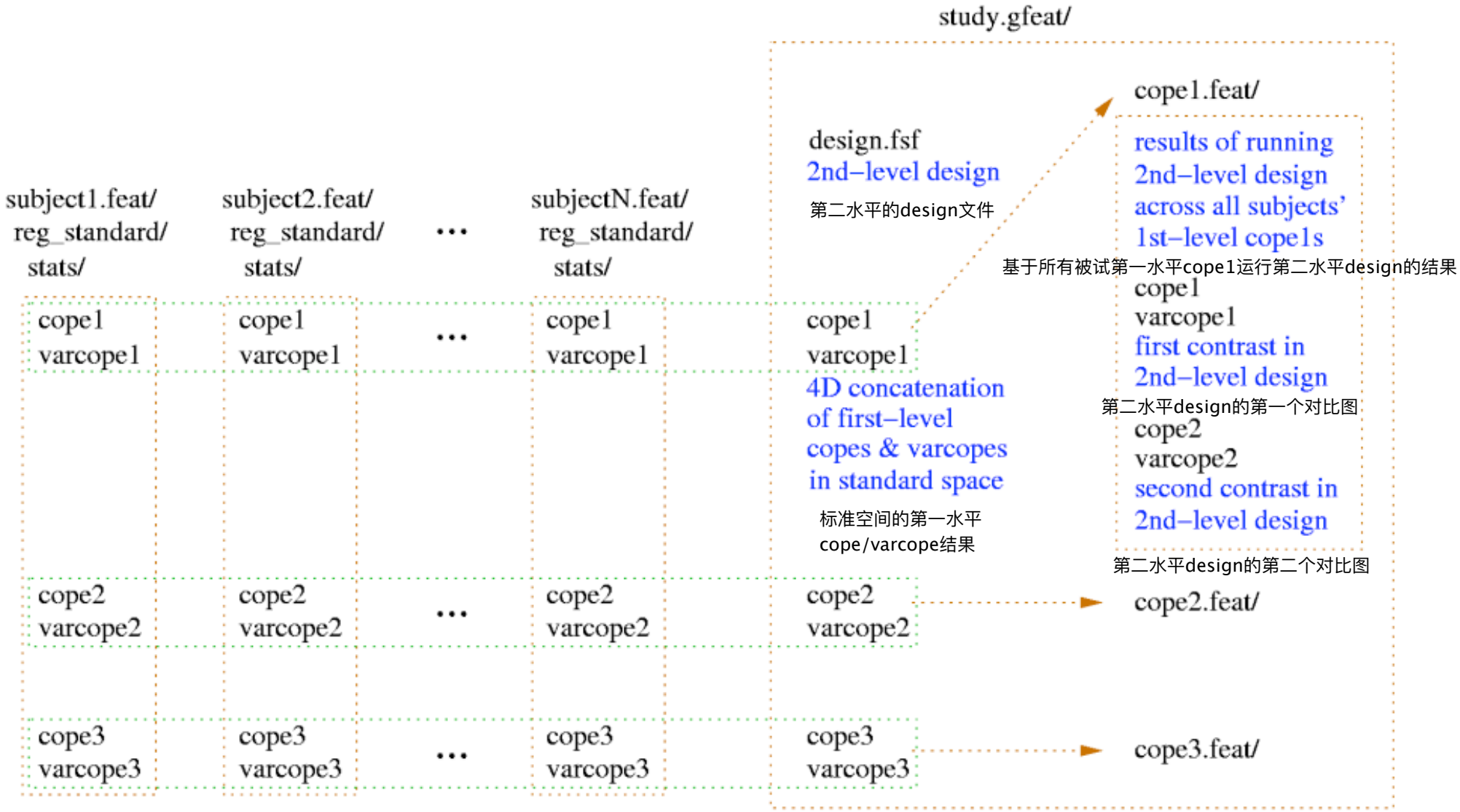
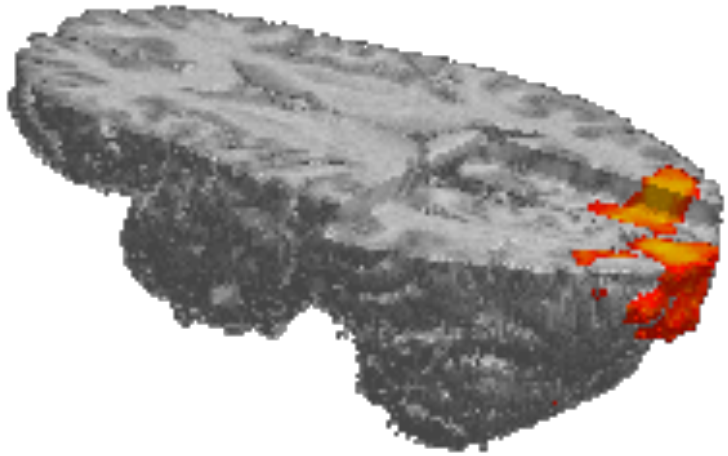
- First-level analysis:第一水平分析



Understanding FEAT dirs

理解FEAT文件夹

- Second-level analysis:



That's all folks