Group Analysis in AFNI

File: GroupAna.pdf

Gang Chen

SSCC/NIMH/NIH/HHS

FMRI Data Analysis

Experiment Design

↓

Scanning

↓

Pre-Processing

↓

Individual Subject Analysis: 1^{st} level

↓

Group Analysis: 2^{nd} level

↓

Post-Processing: clusterizing, ROI analysis, connectivity, …

• Overview

- Why do we need to do group analysis?
  - Cross-subject random effects
- Fixed-effects analysis
- Mixed-effects analysis
  - Nonparametric approach
    - **3dWilcoxon, 3dMannWhitney, 3dKruskalWallis, 3dFriedman**
  - Parametric approach
- Traditional parametric analysis
  - Effect size only: linear combination of regression coefficients (β)
    - **3dttest/3dttest++, 3ddot, 3dANOVA/2/3, 3dRegAna, GroupAna, 3dLME**
- New group analysis method
  - Effect size and precision: mixed-effects meta analysis (MEMA)
    - **3dMEMA**

• Group Analysis: Fixed-Effects Analysis

- Number of subjects $n < 6$
- Case study: difficult to generalize to whole population
- Model $\beta_i = b + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma_i^2)$, σ_i^2: within-subject variability
  - Fixed in the sense that cross-subject variability is not considered
- Direct fixed-effects analysis (**3dDeconvolve/3dREMLfit**)
  - Combine data from all subjects and then run regression
- Fixed-effects meta-analysis (**3dcalc**) : weighted least squares
  - $\beta = \sum w_i \beta_i / \sum w_i$, $w_i = t_i / \beta_i$ = weight for ith subject
  - $t = \beta \sum w_i / \sqrt{n} = \sum t_i / \sqrt{n}$

• Group Analysis: Mixed-Effects Analysis

- ❑ Non-parametric approach
  - ➢ $4 <$ number of subjects < 10
  - ➢ No assumption of data distribution (e.g., normality)
  - ➢ Statistics based on ranking
  - ➢ Individual and group analyses: separate
- ❑ Parametric approach
  - ➢ Number of subjects ≥ 10
  - ➢ Random effects of subjects: usually Gaussian distribution
  - ➢ Individual and group analyses: separate

• Mixed-Effects: Non-Parametric Analysis

- Programs: roughly equivalent to permutation tests
  - **3dWilcoxon** (~ paired *t*-test)
  - **3dFriedman** (~one-way within-subject with **3dANOVA2**)
  - **3dMannWhitney** (~ two-sample *t*-test)
  - **3dKruskalWallis** (~ between-subjects with **3dANOVA**)
- Pros: Less sensitive to outliers (more robust)
- Cons
  - Multiple testing correction **limited** to FDR (**3dFDR**)
  - Less flexible than parametric tests
    - Can't handle complicated designs with more than one fixed factor
    - Can't handle covariates

- **Mixed-Effects: Basic concepts in parametric approach**
  - Fixed factor/effect
    - Treated as a fixed variable (constant) in the model
      - Categorization of experiment conditions (modality: visual/audial)
      - Group of subjects (gender, normal/patients)
    - All levels of the factor are of interest
    - Fixed in the sense statistical inferences
      - apply only to the specific levels of the factor
      - don't extend to other potential levels that might have been included
  - Random factor/effect
    - Treated as a random variable in the model: exclusively subject in FMRI
      - average + effects uniquely attributable to each subject: *e.g.* $N(\mu, \sigma^2)$
    - Each individual subject is of NO interest
    - Random in the sense
      - subjects serve as random sample (representation) from a population
      - inferences can be generalized to a hypothetical population

- **Mixed-Effects: In case you love equations too much!!!**
  - ❑ Linear model for individual subject analysis
    - o $Y = X\beta + \varepsilon$, $\varepsilon \sim N_n(0, \sigma^2 I_n)$
    - o Only one random effect, residuals ε
    - o Individual subject analysis in FMRI
  - ❑ Linear mixed-effects (LME) model
    - o , $d_i \sim N(0, \psi)$, $e_i \sim N(0, \Lambda)$
    - o Two random effect components: cross-subject effect $Z_i d_i$ and within-subject effect ε
    - o Group analysis in FMRI: *t*-tests and ANOVAs are special cases of LME with idealized assumptions
    - o It is the cross-subject component $Z_i d_i$ that legitimizes the generalization at population level

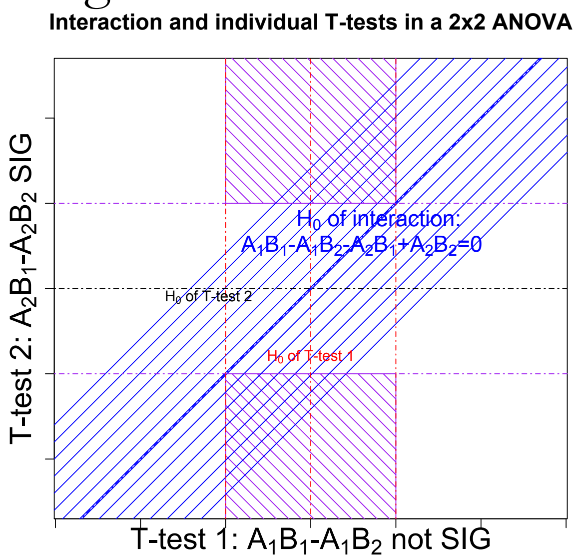
- **Mixed-Effects: Mixed-Effects Analysis**
  - Programs
    - **3dttest** (one-sample, two-sample and paired *t*)
    - **3dttest++** (one-sample, two-sample and paired *t*) + covariates (voxel-wise)
    - **3ddot** (correlation between two sets)
    - **3dANOVA** (one-way between-subject)
    - **3dANOVA2** (one-way within-subject, 2-way between-subjects)
    - **3dANOVA3** (2-way within-subject and mixed, 3-way between-subjects)
    - **3dRegAna** (regression/correlation, covariates)
    - **GroupAna** (Matlab package for up to 5-way ANOVA)
    - **3dLME** (R package for various kinds of group analysis)
    - **3dMEMA** (R package for meta analysis, t-tests plus covariates)

• Mixed-Effects: Which program should I use?

- Two perspectives: batch vs. piecemeal
  - Experiment design
    - Factors/levels, balancedness
      - * ANOVA: main effects, interactions, simple effects, contrasts, …
      - * Linear mixed-effects model
    - Most people are educated in this traditional paradigm!
    - Pros: get almost everything you want in one batch model
    - Cons: F-stat for main effect and interaction is difficult to comprehend
    - condensed/summarized test with vague information when levels/factors greater than 2 (**I don't like F-test personally!!! Sorry, Ronald A. Fisher…**),
    - with assumptions: homogeneity with multiple groups, and compound symmetry when a within-subject factor has more than 2 levels
  - Tests of interest
    - Simple/straightforward/piecemeal: focus on each individual test & handle one at a time
    - Mainly t-stat: one-sample, paired, two-sample
    - All main effects and interactions can be broken into multiple t-tests

• ANOVA vs t-tests: subtle differences

- ANOVA
  - Syntactic sugar for a special subgroup of regression
  - Used by researchers who are not statistician by training
  - Institutionalized; hard to convert them back to regression
- *F*-tests vs. *post-hoc* *t*-tests
  - Interaction *F* not significant; some individual *t*-tests significant
  - Interaction *F* significant; none of individual *t*-tests significant
  - Power issue
  - *F* for the main effect of a factor with two levels is essentially *t*
  - *F* for main effects and interactions of all factor with two levels are essentially *t*



- **Jack of All Trades** (well, almost): **`3dttest/3dttest++`**
  - Basic usage
    - One-sample *t*
      - One group: simple effect; Example: 10 subjects under condition *Vrel* with H_0: $\mu_V = 0$
    - Two-sample *t*
      - Two groups: Compare one group with another
      - ~ 1-way between-subject (**`3dANOVA2 -type 1`**)
      - Unequal sample sizes allowed
      - Homoskedasticity vs. heteroskedasticity: **`-unpooled`**
      - Example: 15 TD subjects vs. 13 autism subjects - H_0: $\mu_A = \mu_B$
    - Paired *t*
      - Two conditions of one group: Compare one condition with another
      - ~ one-way within-subject (**`3dANOVA2 -type 3`**)
      - ~ one-sample *t* on individual contrasts
      - Example: Difference of visual and auditory conditions for 10 subjects with H_0: $\mu_V = \mu_A$
  - Output: 2 values (effect and *t*)
  - Versatile program: Most tests can be done with **`3dttest`** - piecemeal vs. bundled
  - **`-mask`** option unavailable but desirable!

- **3dttest**: Example

- Paired t-test

```
3dttest -prefix ttest_V-A -paired \
     -set1                        \
          'OLSQ.FP.betas+tlrc[1]'    \
          'OLSQ.FR.betas+tlrc[1]'    \
          ......
          'OLSQ.GM.betas+tlrc[1]'    \
     -set2                        \
          'OLSQ.FP.betas+tlrc[0]'    \
          'OLSQ.FR.betas+tlrc[0]'    \
          ......
          'OLSQ.GM.betas+tlrc[0]'
```

Model type,

Input files for Arel condition

Input files for Vrel condition

- **ANVA program 1** : **`3dANOVA`**
  - Generalization of two-sample *t*-test
    - One-way between-subject: 2 or more groups of subjects
    - H_0: no difference across all levels (groups)
    - Examples of groups: gender, age, genotype, disease, *etc.*
    - Unequal sample sizes allowed
  - Assumptions
    - Normally distributed with equal variance across groups
  - Results: 2 values (% and *t*)
  - **`3dANOVA`** vs. **`3dttest`**
    - Equivalent with 2 levels (groups) if equal variance is assumed
    - More than 2 levels (groups): Can run multiple 2-sample *t*-test
    - **`3dttest`** allows heteroscedasticity (unequal variance across groups)

- **ANOVA program 2**: **`3dANOVA2`**
  - Designs: generalization of paired *t*-test
    - One-way within-subject (**`type 3`**)
      - Major usage
      - Compare conditions in one group
      - Extension and equivalence of paired *t*
    - Two-way between-subjects (**`type 1`**)
      - 1 condition, 2 classifications of subjects
      - Extension and equivalence two-sample *t*
      - Unbalanced designs disallowed: Equal number of subjects across groups
  - Output
    - Main effect (**`-fa`**): *F*
    - Interaction for two-way between-subjects (**`-fab`**): *F*
    - Contrast testing
      - Simple effect (**`-amean`**)
      - 1st level (**`-acontr, -adiff`**): among factor levels
      - 2nd level (interaction) for two-way between-subjects
      - 2 values per contrast: % and *t*

• 3dANOVA2: Example

- Two factors: A – condition (fixed, 2 levels); B – subject (random, 10 levels).
- Script s1.3dANOVA2 under ~/AFNI_data6/group_results/

```
3dANOVA2 -type 3 -alevels 2 -blevels 10                \
    -mask mask+tlrc                                    \
    -dset 1  1 'OLSQ.FP.betas+tlrc[Vrel#0_Coef]'       \
    -dset 2  1 'OLSQ.FP.betas+tlrc[Arel#0_Coef]'       \
    -dset 1  2 'OLSQ.FR.betas+tlrc[Vrel#0_Coef]'       \
    -dset 2  2 'OLSQ.FR.betas+tlrc[Arel#0_Coef]'       \
    ......
    -dset 1 10 'OLSQ.GM.betas+tlrc[Vrel#0_Coef]'       \
    -dset 2 10 'OLSQ.GM.betas+tlrc[Arel#0_Coef]'       \
    -amean 1 V                                         \
    -amean 2 A                                         \
    -adiff 1 2 VvsA                                    \
    -fa FullEffect                                     \
    -bucket anova.VA
```

Model type, Factor levels

Input for each cell in ANOVA table: Totally 2X10 = 20

t tests: one-sample type

t test: two-paired

F test: main effect

Output: bundled

All the F/t-tests here can be obtained with 3dttest!

• ANOVA program 3: `3dANOVA3`

- Designs
  - Two-way within-subject (**`type 4`**): Crossed design AXBXC
    - Generalization of paired *t*-test
    - One group of subjects
    - Two categorizations of conditions: A and B
  - Two-way mixed (**`type 5`**): Nested design BXC(A)
    - Two or more groups of subjects (Factor A): subject classification, e.g., gender
    - One category of condition (Factor B)
    - Nesting: balanced
  - Three-way between-subjects (**`type 1`**)
    - 3 categorizations of groups
- Output
  - Main effect (**`-fa and -fb`**) and interaction (**`-fab`**): *F*
  - Contrast testing
    - 1st level: **`-amean, -adiff, -acontr, -bmean, -bdiff, -bcontr`**
    - 2nd level: **`-abmean, -aBdiff, -aBcontr, -Abdiff, -Abcontr`**
    - 2 values per contrast : % and *t*

- **ANOVA program 4**: GroupAna
  - Pros
    - Matlab script package for up to 5-way ANOVA
    - Can handle both volume and surface data
    - Can handle following unbalanced designs (two-sample *t* type):
      - 3-way ANOVA type 3: BXC(A)
      - 4-way ANOVA type 3: BXCXD(A)
      - 4-way ANOVA type 4: CXD(AXB)
  - Cons
    - Use a commercial package: requires Matlab plus Statistics Toolbox
    - Difficult to test and interpret simple effects/contrasts
    - Complicated design, and compromised power
    - GLM approach (slow): heavy duty computation: minutes to hours
      - Input with lower resolution recommended
      - Resample with **`adwarp -dxyz #`** and **`3dresample`**
  - See http://afni.nimh.nih.gov/sscc/gangc for more info

- **Regression**: Group level
  - Correlation analysis
    - Between brain response and some continuous variables (covariates)
    - Continuous variables (covariates) are subject-level variables
      - behavioral data
      - physical atributes, e.g., age, IQ, brain volume, etc.
    - Correlation (spatial) between two sets of 3D data
      - 3ddot -demean
  - **`3dRegAna`**
    - One- or two-sample *t*-test + covariates
    - See http://afni.nimh.nih.gov/sscc/gangc/ANCOVA.html for more info

- **Regression: Group level**
  - Regression analysis at group level
    - Between brain response and some continuous variables (covariates)
    - Continuous variables (covariates) are subject-level variables
      - behavioral data
      - physical atributes, e.g., age, IQ, brain volume, etc.
      - Covariates can be voxel-wise values
  - **`3dttest++`** (new-ish program)
    - One- or two-sample *t*-test + covariates
    - Usage similar to 3dMEMA
    - More user-friendly than **`3dRegAna`**
    - More information can be found by typing the following at the terminal

      3dttest++ -help | less

• Linear Mixed-Effects Analysis: 3dLME

- Model $\hat{\mathbf{b}}_i = X_i\mathbf{a} + Z_i\mathbf{d}_i + \mathbf{e}_i$,
- Pros
  - R package: open source platform
  - Linear mixed-effects (LME) modeling
  - Versatile: handles almost all situations in one package
    - Unbalanced designs (unequal number of subjects, missing data, etc.)
    - ANOVA and ANCOVA, but unlimited factors and covariates
    - Able to handle HRF modeling with basis functions
    - Violation of sphericity: heteroscedasticity, variance-covariance structure
- Cons
  - High computation cost (lots of repetitive calculation)
  - Controversial regarding degrees of freedom
- See http://afni.nimh.nih.gov/sscc/gangc/lme.html for info

• Linear Mixed-Effects Analysis: 3dLME

❑ Running LME: HRF modeled with 6 tents

- Null hypothesis H_0: $\beta_1=\beta_2=\ldots=\beta_6=0$ (NOT $\beta_1=\beta_2=\ldots=\beta_6$)

```
Data:Volume                                   <-- either Volume or Surface
Output:test                                   <-- any string (no suffix
  needed)
MASK:Mask+tlrc.BRIK                           <-- mask dataset
FixEff:Time-1                                 <-- model formula for fixed
  effects
COV:                                          <-- covariate list
RanEff:   1                                   <-- random effect specification
VarStr:weights=varIdent(form=~1|Time)         <-- heteroscedasticity?
CorStr:correlation=corAR1(form=~Order|Subj) <-- correlation structure
SS:sequential                                 <-- sequential or marginal
Clusters:4
Subj      Time    TimeOrder   InputFile
Jim       t1         1    contrastT1+tlrc.BRIK
Jim       t2         2    contrastT2+tlrc.BRIK
......
Jim       t6         6    contrastT6+tlrc.BRIK
```

Mixed-Effects Meta Analysis: 3dMEMA

- Requirements
  - R installment, plus 'snow' package for parallel computing
- 4 running modes
  - Scripting: type '3dMEMA –help' at terminal to see usage
  - Sequential/interactive mode inside R: source("~/abin/3dMEMA.R")
  - Batch (if answers known): R CMD BATCH Cmds.R myDiary &
  - Command line: 3dMEMA command as a front end to R
- Pros
  - Makes more sense: better statistical properties, uses β **plus** *t*-statistic
  - Likely more statistically powerful
  - Less prone to outliers
  - Provides more diagnostic measures
  - Can include subject-level covariates in the analysis -- like **3dttest++**
- Cons
  - Longer runtime
  - Can't handle sophisticated situations: basis functions, ANOVAs, …

3dMEMA: example-scripting

Paired type test: visual-reliable vs. auditory-reliable

- script **s4.3dMEMA.V-A** under AFNI_data6/group_results/

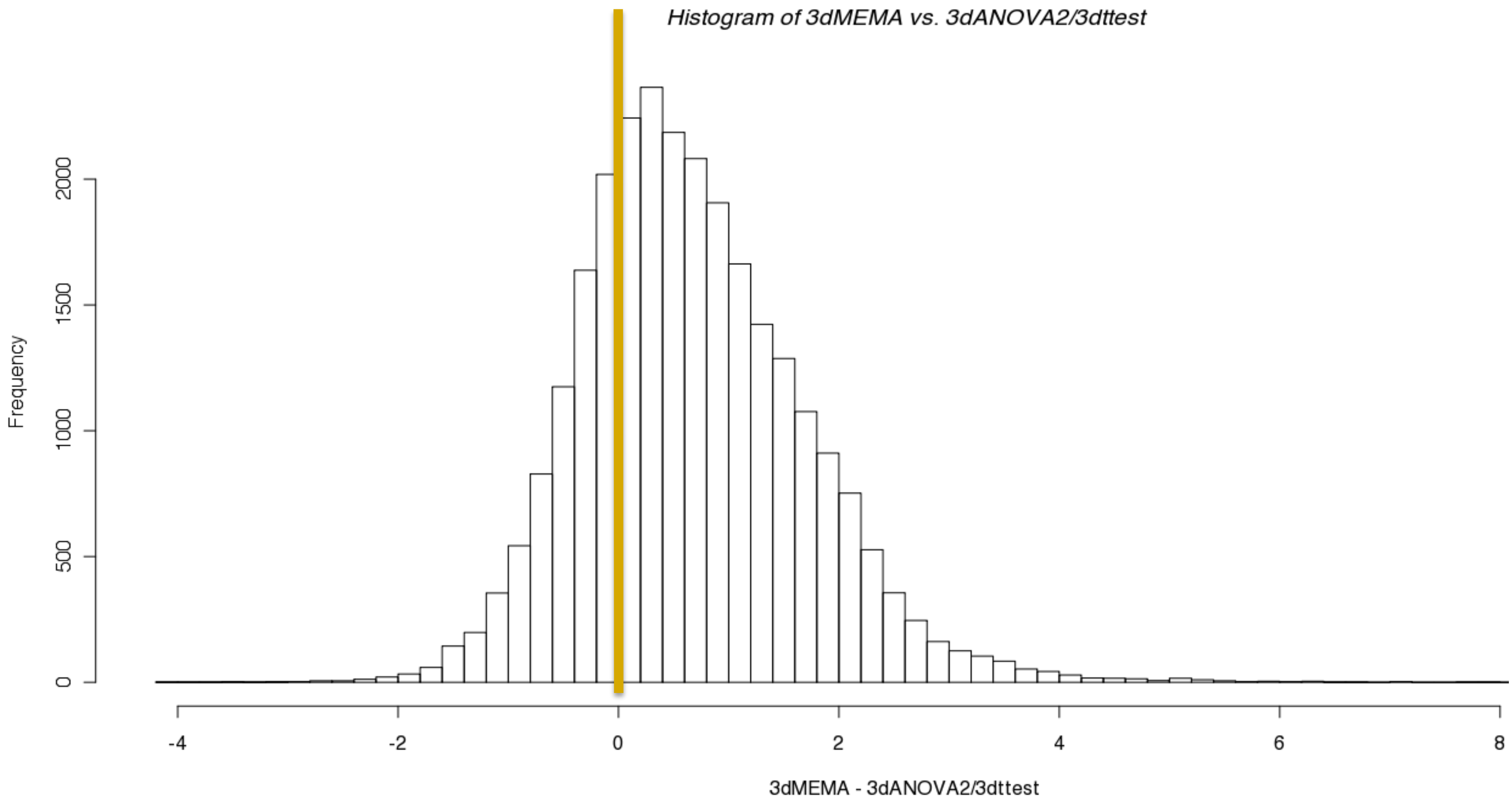
```
3dMEMA -prefix mema_V-A -mask mask+tlrc  -jobs 4 -max_zeros 3  \
    -conditions Vrel Arel –Hktest -model_outliers              \
    -set Arel                                                  \
        FP 'REML.FP.bt+tlrc[2]' 'REML.FP.bt+tlrc[3]'           \
        FR 'REML.FR.bt+tlrc[2]' 'REML.FR.bt+tlrc[3]'           \
        ……
        GK 'REML.GK.bt+tlrc[2]' 'REML.GK.bt+tlrc[3]'           \
        GM 'REML.GM.bt+tlrc[2]' 'REML.GM.bt+tlrc[3]'           \
    -set Vrel                                                  \
        FP 'REML.FP.bt+tlrc[0]' 'REML.FP.bt+tlrc[1]'           \
        FR 'REML.FR.bt+tlrc[0]' 'REML.FR.bt+tlrc[1]'           \
        ……
        GK 'REML.GK.bt+tlrc[0]' 'REML.GK.bt+tlrc[1]'           \
        GM 'REML.GM.bt+tlrc[0]' 'REML.GM.bt+tlrc[1]'
```

3dMEMA: example-interactive/batch

- One-sample test: visual-reliable
- Sequential/interactive mode on R prompt
  - Demo here
- Batch mode: R CMD BATCH scriptCMD.R myDiary.txt &
  - Remote running: nohup R CMD BATCH scriptCMD.R myDiary.txt &

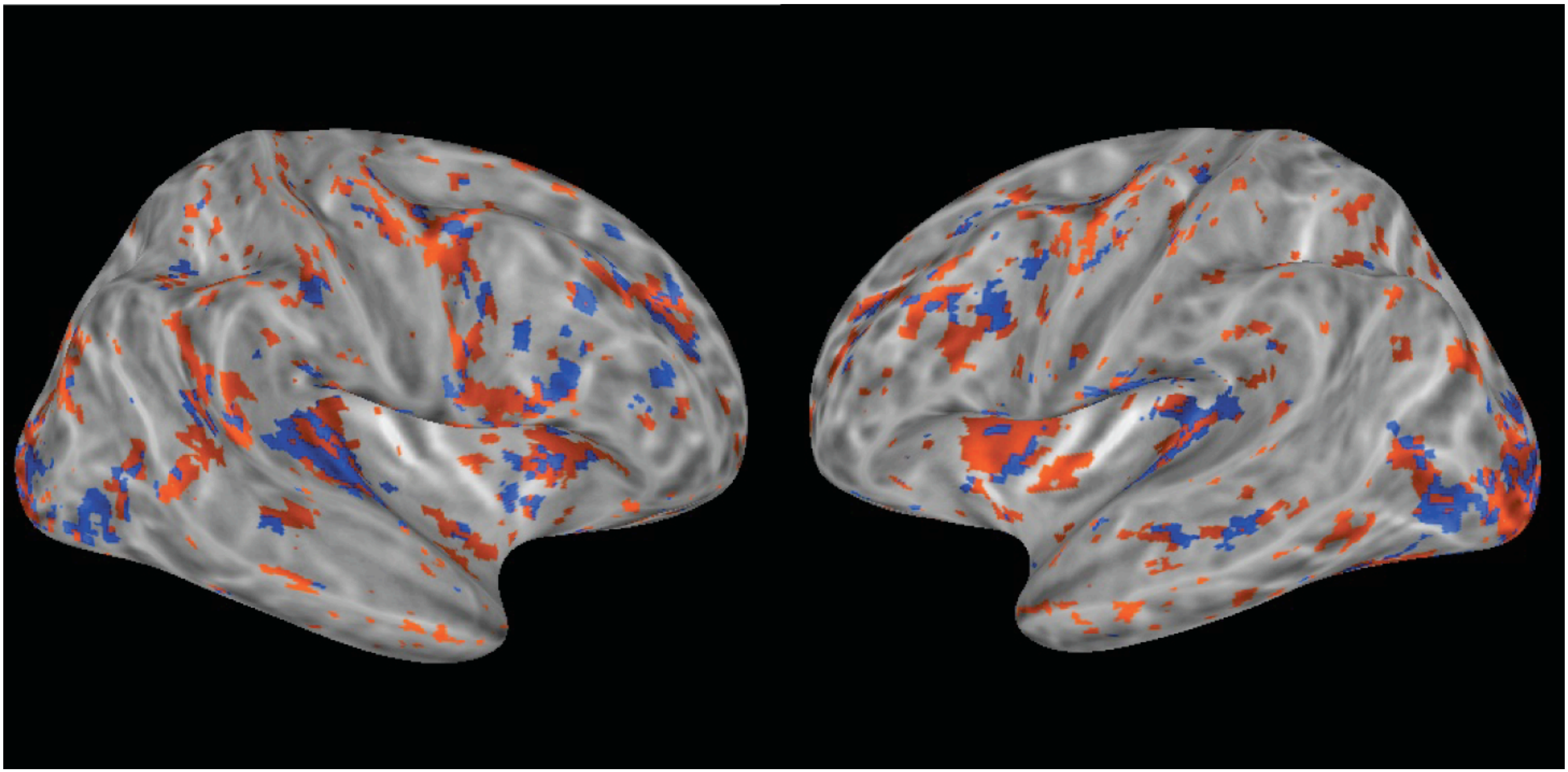
3dMEMA: comparison with 3dttest

- Majority of significant voxels with 3dMEMA gained power with a threshold of 2.0 for $t(30)$



3dMEMA: comparison with 3dttest

- Majority of significant voxels with 3dMEMA gained power (**red**: 3dMEMA higher; **blue**: 3dttest higher) with a threshold of 2.0 for *t*(9).



Why new group analysis approach?

- Our ultimate goal is not just to gain statistical power
- Old group analysis approach
  - Take β's from each subject, and run *t*-test, AN(C)OVA, LME
  - Three assumptions
    - Within/intra-subject variability (standard error, sampling error) is relatively small compared to cross/between/inter-subjects variability
    - Within/intra-subject variability roughly the same across subjects
    - Normal distribution for cross-subject variability (no outliers)
  - Violations seem everywhere: violating either can lead to suboptimal/invalid analysis
    - Common to see 40% up to 100% variability due to within-subject variability
    - Non-uniform within/intra-subject variability across subjects

How can we do it differently?

- For each effect estimate (β or linear combination of β' s)
  - Information regarding our confidence about the effect?
    - Reliability/precision/efficiency/certainty/confidence: standard error (SE)!
    - Smaller SE ➔ higher reliability/precision/efficiency/certainty/confidence
    - SE of an effect = estimated standard deviation (SD) of the effect
  - t-statistic of the effect
    - Signal-to-noise or effect vs. uncertainty: $t = \beta/\mathrm{SE}$
    - SE contained in t-statistic: $\mathrm{SE} = \beta/t$
  - Trust those β' s with high reliability/precision (small SE) through weighting/compromise
    - β estimate with high precision (lower SE) has more say in the final result
    - β estimate with high uncertainty gets downgraded

Weigh effects based on precision

- Dealing with outliers
  - Unreliable estimate (small *t*): small/big β + big SE
    - Will automatically be downgraded
    - May still slightly bias cross-subjects variability estimate to some extent, leading to unfavorable significance testing, but much better than conventional approach
  - Reliable estimate (big *t*): small/big β + small SE
    - Weighting only helps to some extent: if one subject has extremely small SE (big *t*), the group effect may be dominated by this subject
    - Needs delicate solutions: fundamentally why outliers?
      - Brain level: Considering covariate(s)? Grouping subjects?
      - Singular voxels: special modeling on cross-subject variance

Running 3dMEMA

- Currently available analysis types (+ covariates allowed)
  - One-sample: one condition with one group
  - Two-sample: one condition across 2 groups with homoskedasticity (same variability)
  - Paired-sample: two conditions with one group
  - Two-sample: one condition across 2 groups with heteroskedasticity (different variability)
  - Can also handle multiple between-subjects factors
- Output
  - Group level: % signal change + Z/t-statistic, τ^2 + Q
  - Individual level: λ + Z for each subject
- Modes
  - Scripting
  - Sequential mode on terminal
  - Batch mode: R CMD BATCH cmds.R diary.txt & *or* 3dMEMA

3dMEMA limitations

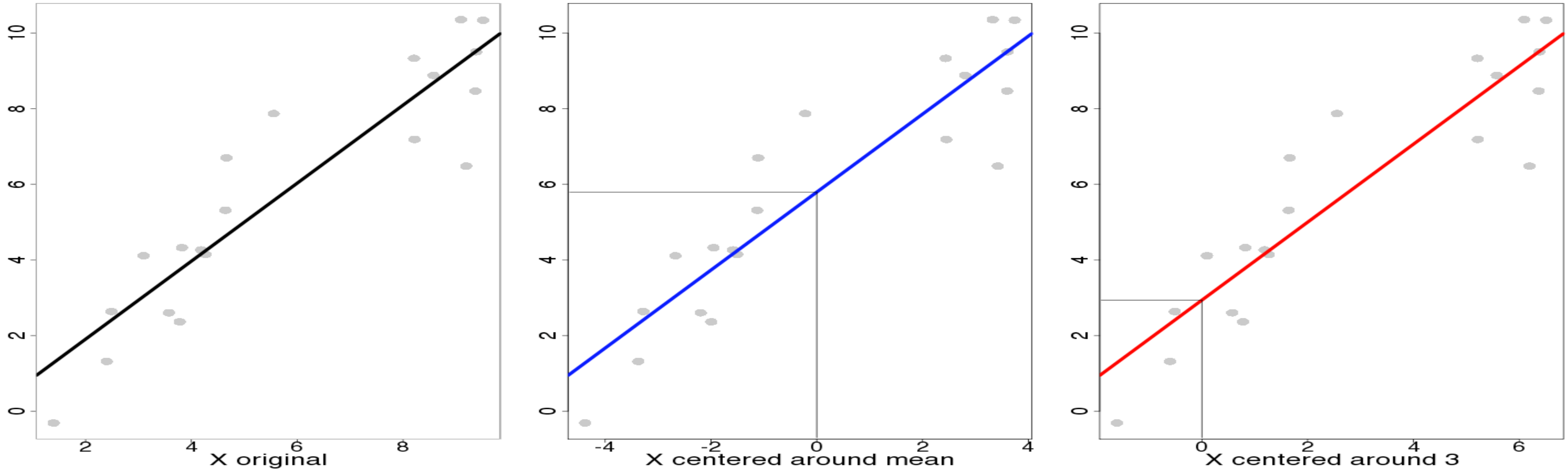
- Basis functions? Multiple βs per voxel?
  - Stick with 3dLME for now
- ANOVA?
  - Extension difficult
  - *t*-tests should be no problem
  - *F*-tests?
    - Some of them boil down to *t*-tests, for example:
      - *F*-test for interaction between A and B (both with 2 levels) with "3dANOVA3 -type 5…"
      - Equivalent to *t*-test for (A1B1-A1B2)-(A2B1-A2B2) or (A1B1-A2B1)-(A1B2-A2B2)
      - We can say more with *t* than *F*: a positive *t* shows A1B1-A1B2 > A2B1-A2B2 and A1B1-A2B1 > A1B2-A2B2
    - Do something for more complex *F* in the future?

Covariates

- Covariates
  - May or may not be of direct interest
  - Confounding, nuisance, or interacting variables
  - Subject-level (vs. trial-level: handled via amplitude modulation)
  - Controlling for variability in the covariate
  - Continuous or discrete?
  - One-sample model $y_i = \alpha_0+\alpha_1 x_i+\delta_i + \boldsymbol{\varepsilon}_i$, for ith subject
  - Two-sample model $y_i = \alpha_0+\alpha_1 x_{1i}+\alpha_2 x_{2i}+\alpha_3 x_{3i}+\delta_i + \boldsymbol{\varepsilon}_i$
- Examples
  - Age, IQ, brain volume, cortex thickness
  - Behavioral data

Handling covariates: one group

- Centering: tricky business (using age as an example)
  - $y_i = \alpha_0 + \alpha_1 x_i + \delta_i + \varepsilon_i$, for ith subject
  - Interested in group effect α_0 (x=0) while controlling (partialling out) x
  - α_1 - slope (change rate): % signal change per unit of x
  - Interpretability: group effect α_0 at what value of x: mean or any other value?

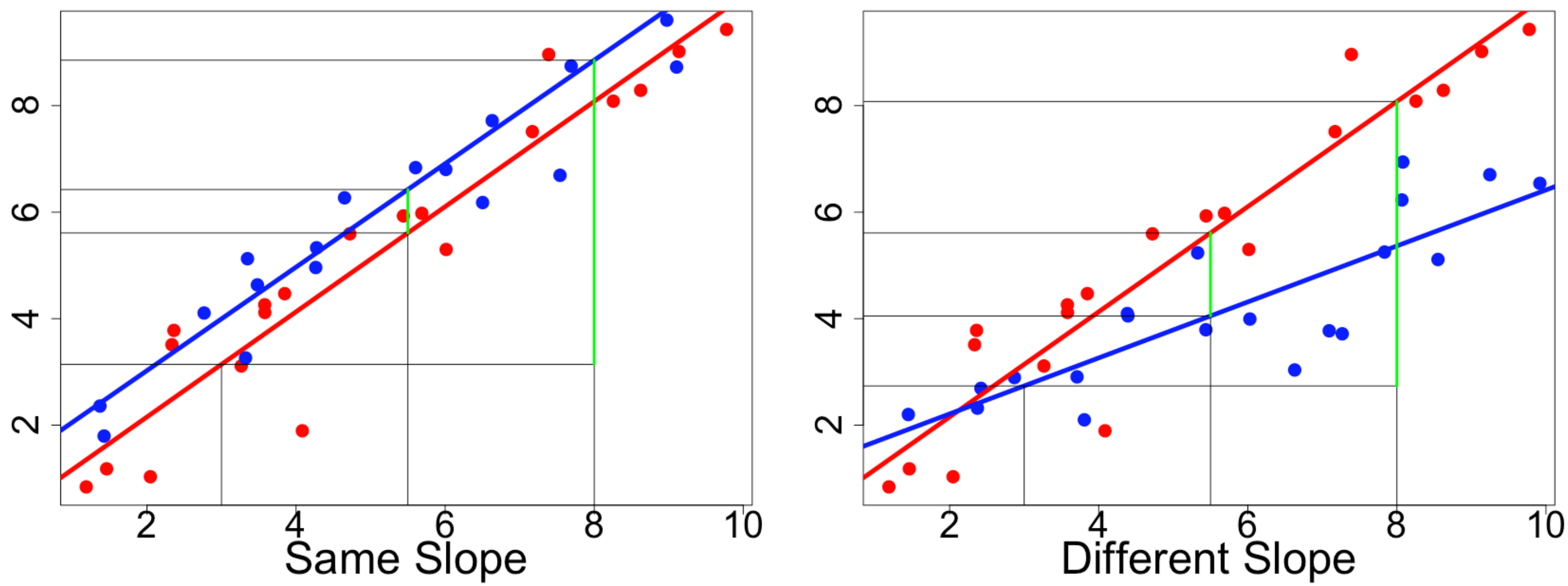


Covariates: trickier with 2 groups

- Center and slope
  - $y_i = \alpha_0+\alpha_1 x_{1i}+\alpha_2 x_{2i}+\alpha_3 x_{3i}+\delta_i + \varepsilon_i$, for ith subject
    - x_1: group indicator [0 or 1, say]
    - x_2: covariate
    - x_3: group effect in slope (interaction between group and covariate = $x_1 * x_2$)
  - What we're interested in
    - Group effects α_0 and α_1 while controlling covariate
  - Interpretability
    - Center
      - Group effect α_0 and α_1 at what covariate value?
      - Same or different center across groups?
    - Slope
      - same (α_3=0) or different ($\alpha_3 \neq 0$) slope across groups

Covariates: scenarios with 2 groups

- Center and slope (again using age as an example)
  - $y_i = \alpha_0+\alpha_1 x_{1i}+\alpha_2 x_{2i}+\alpha_3 x_{3i}+\delta_i + \varepsilon_i$, for ith subject
  - Interpretability
    - Same center and slope (α_3=0)
    - Different center with same slope (α_3=0)
    - Same center with different slope ($\alpha_3 \neq 0$)
    - Different center and slope ($\alpha_3 \neq 0$)



Start simple: one-sample test

- Random-effects: $y_i=\theta_i+\varepsilon_i=\alpha_0+\delta_i+\varepsilon_i$, for ith subject
  - y_i : β or linear combination (contrast) of β's from ith subject
  - $\theta_i=\alpha_0+\delta_i$: "true" individual effect from ith subject
  - α_0: group effect we'd like to find out
  - δ_i : deviation of ith subject from group effect α_0, $N(0, \tau^2)$
  - ε_i: sample error from ith subject, $N(0, \sigma_i^2)$, σ_i^2 known!
- Special cases
  - $\sigma_i^2=0$ reduced to conventional group analysis
    - One-sample t: $y_i = \alpha_0 + \delta_i$
  - $\delta_i=0$ ($\tau^2=0$) assumed in fixed-effects (FE) model: Ideally we could find out all possible explanatory variables so only an FE model is necessary!
- Mature meta analysis tools for this simple model
  - Broadly used in clinical trials/epidemiology in recent 20 yrs
  - A special case of linear mixed-effects model

MEMA with one-sample test

- Random-effects: $y_i = \alpha_0 + \delta_i + \varepsilon_i$, for ith subject
  - $\delta_i \sim N(0, \tau^2)$, $\varepsilon_i \sim N(0, \sigma_i^2)$
  - σ_i^2 known, τ^2 unknown = inter-subject variance (per-voxel)
  - What can we achieve?
    - Null hypothesis about group effect H_0: $\alpha_0 = 0$
    - Checking group heterogeneity H_0: $\tau^2 = 0$
    - Any outliers among the subjects? Adding some confounding variable(s)? Grouping subjects?
  - We know σ_i^2, and pretend we also knew τ^2, weighted least squares (WLS) gives $\hat{\alpha}_0 = \frac{\sum w_i y_i}{\sum w_i}, w_i = \frac{1}{\tau^2 + \sigma_i^2}$
    - The "best" estimate
    - **BLUE**: unbiased with minimum variance
  - **Wake up**: Unfortunately we don't know τ^2!!!
    - It must be estimated at the same time as α_0

Solving MEMA in one-sample case

- Estimating τ^2: a few approaches
  - Method of moment (MoM) - DSL
  - Maximum likelihood (ML)
  - Restricted/residual/reduced/marginal ML (REML): 3dMEMA
- Statistical testing
  - Group effect $\alpha_0=0$: $Z = \frac{\sum w_i y_i}{\sqrt{\sum w_i}} \cong N(0,1), w_i = \frac{1}{\tau^2 + \sigma_i^2}$
    - Wald or Z-test: assume enough subjects with normal distributions
    - Go with t-test when in doubt
  - Heterogeneity test $\tau^2=0$: $Q = \sum_{i=1}^{n} \frac{(y_i - \hat{\alpha}_0)^2}{\sigma_i^2} \sim \chi^2(n-1)$
  - Outlier identification for each subject through Z-statistic

We don't limit ourselves to simple case

- $y_i = \alpha_0 + \alpha_1 x_{i1} + \ldots + \alpha_{ip} x_{ip} + \delta_i + \varepsilon_i$, for ith subject
  - Mixed-effects model or meta regression
  - y_i: β or linear combination (contrast) of β's from ith subject
  - α_0: common group effect we'd like to find out
  - x_{ij}: an indicator/dummy variable showing, for example, group to which ith subject belongs, level at which a factor lies, or a continuous variable such as covariate (e.g., age, IQ) (j=1,…,p)
  - δ_i : deviation of ith subject from group effect α_0, $N(0, \tau^2)$
  - ε_i: sample error from ith subject, $N(0, \sigma_i^2)$, σ_i^2 known!
- Combine subjects into a concise model in matrix form
  - $\mathbf{y}_{n\times 1} = \mathbf{X}_{n\times p}\boldsymbol{\alpha}_{p\times 1} + \boldsymbol{\delta}_{n\times 1} + \boldsymbol{\varepsilon}_{n\times 1}$
  - $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\alpha}, \tau^2\mathbf{I}_n+\mathbf{V})$, $\mathbf{V} = \mathrm{diag}(\sigma_1,\ldots, \sigma_n)$ known, τ^2 unknown
  - Estimate $\boldsymbol{\alpha}$ and τ^2 simultaneously via maximizing REML

Dealing with outliers

- Detection
  - Ideally we wish to account for anything until having no cross-subject variability: $\tau^2 = 0$!
  - 4 quantities to check cross-subject variability
    - Cross subject variability (heterogeneity) τ^2
    - Q for H_0: $\tau^2 = 0$
    - Intra-class correlation (ICC): $\lambda = \sigma_i^2/(\sigma_i^2+\tau^2)$
    - Z statistic of $\boldsymbol{\varepsilon}_i$
- Modeling: how to handle outliers in the model?
  - Ignore those subjects with 2 s.d. away from mean?
    - Arbitrary: OK with data within 1.9 s.d.?
    - How about when outliers occur at voxel level?
    - If throwing away outliers at voxel level, varying DFs across brain?

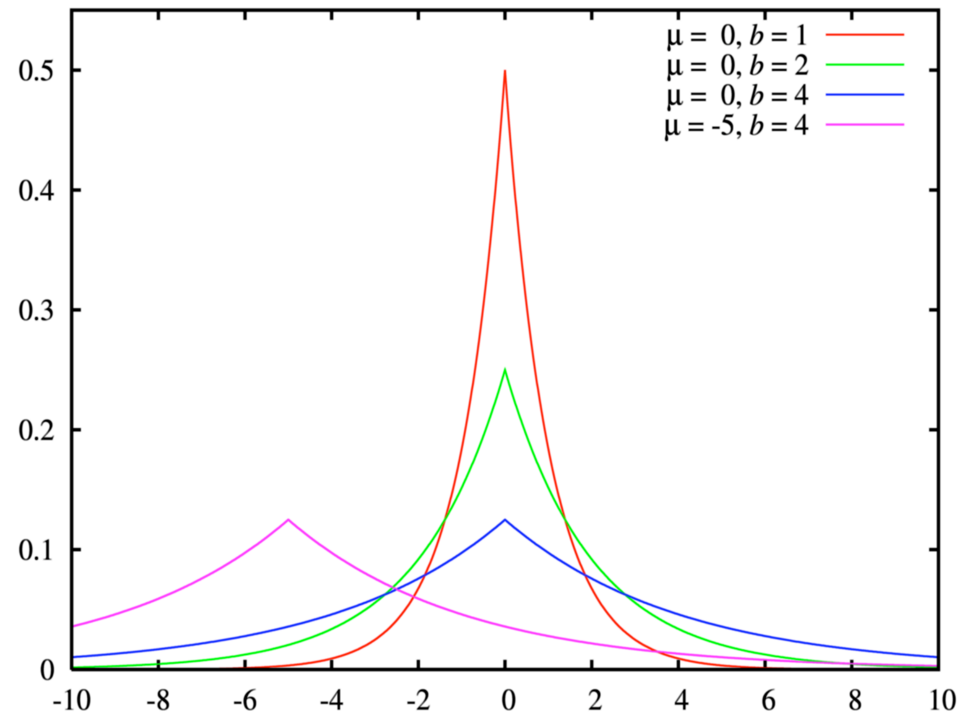
Modeling outliers

- Modeling: how to handle outliers in the model?
  - Typically a Gaussian for subject deviation: $\delta_i \sim N(0, \tau^2)$
  - With outliers, assume a Laplace (double exponential) distribution

$$f(x|\mu,b) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right)$$

- μ: location parameter
- b: scale parameter
- Mean=median=mode=μ
- Variance = $2b^2$
- Fatter tail but smaller Var
- Estimator of μ is sample median, and ML estimator of b

$$\hat{b} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{\mu}|$$



Modeling outliers

- Laplace distribution for outlier modeling
  - No REML form
  - Go with ML: variance estimate τ^2 might be slightly underestimated
  - Computation cost: higher
  - Generally higher statistical power

Moral of one investigator's story

- Story
  - Strong activation at individual level and in ROI analysis failed to show up at group level
  - Result with 3dMEMA showed consistency with individual and ROI analysis
  - Magic power of 3dMEMA? Relatively robust to some (unreliable) outliers
- Check brick labels for all input files

```
foreach subj (S1 S2 S3 …)
  3dinfo -verb ${subj}_file+tlrc  | grep 'sub-brick #0’
end
```

```
++ 3dinfo: AFNI version=AFNI_2008_07_18_1710 (Jul  8 2009) [32-bit]
 -- At sub-brick #0 ’contr_GLT#0_Coef' datum type is float:     -0.78438 to      0.867817
 -- At sub-brick #0 ’contr_GLT#0_Coef' datum type is float:    -0.444093 to      0.501589
…
```

Suggested preprocessing steps

- Input
  - β and *t*-statistic from each subject
  - One sub-brick per input file (3dbucket)
- Some suggestions
  - Slice timing correction and volume registration
  - Aligning/warping to standard space
    - Avoid troubling step of warping on *t*-statistic
  - Smoothing: 3dBlurToFWHM or 3dBlurInMask
  - Scaling
  - All input files, β and **more importantly** *t*-statistic, come from 3dREMLfit instead of 3dDeconvolve
  - No masking applied at individual level so that no data is lost at group level along the edge of (and sometimes inside) the brain

Comparisons among FMRI packages

Program	Language	Algorithm	Runtime	Group effect statistics	Covariates	Voxelwise outlier detection	Voxelwise outlier modeling
multistat (FMRIstat)	Matlab	EM for REML + spatial regularization	~1 min per test	*t*	✗	✗	✗
FLAME in FEAT (FSL)	C/C++	Bayesian + MCMC	45-200 min per test + threshold	fitted with *t*	✔	% subjects for group, *p* for each subject	mixture of two Gaussian
3dMEMA (AFNI)	R	ML/REML/MoM	3-15 min per test	*Z*/*t*	✔	τ^2 + Q for group, λ + *Z* for each subject	Laplace

Overview: 3dMEMA

- http://afni.nimh.nih.gov/sscc/gangc/MEMA.html
- Meta analysis: compromise between Bayesian and frequentist
  - Backbone: WLS + maximization of REML or ML of Laplace-Gauss
  - Currently available types
    - One-, two-, paired-sample test
    - Covariates allowed: careful with centering and interaction with groups
  - Output
    - Group level: group effect (% signal change) and statistics (Z/t), cross-subject heterogeneity τ^2 and Q (χ^2-test)
    - Individual level: λ + Z for each subject
  - Generally more powerful/valid than conventional approach
  - Relatively robust against most outliers
  - Moderate computation cost with parallel computing: 3-20 minutes
- Limitations
  - Can't handle sophisticated types: multiple basis functions; F-test
  - Computation cost