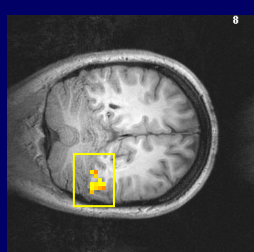
Extending the GLM

- So far, we have considered the GLM for one run in one subject
- The same logic can be applied to multiple runs and multiple subjects

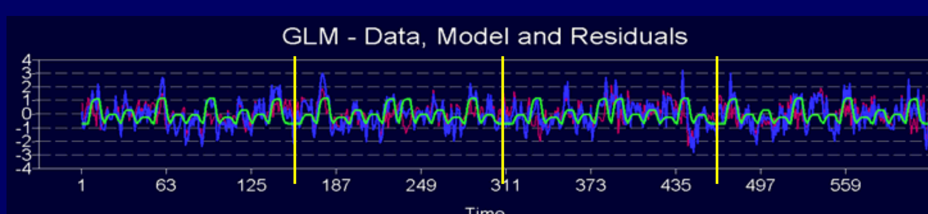
GLM Stats

For any given region, we can evaluate the GLM stats



Predictor	beta	se	t	p
faces left	1.793	0.132	13.539	0.000000
faces right	0.987	0.132	7.451	0.000000
faces fovea	1.848	0.132	13.956	0.000000
places left	0.672	0.132	5.075	0.000001
places right	0.429	0.132	3.237	0.001273
places fovea	0.631	0.132	4.769	0.000002

blue: original time course
=
green: best fitting model
+
red: residuals



total length of sequence = 4 runs * 155 volumes = 620 volumes

Outline

- Mixed effects motivation
- Evaluating mixed effects methods
- Three methods
  - – Summary statistic approach (HF) (SPM96,99,2)
  - – SPM2
  - – FSL3
- Conclusions

Overview

- Mixed effects motivation
- Evaluating mixed effects methods
- Three methods
  - Summary statistic approach (HF) (SPM96,99,2)
  - SPM2
  - FSL3
- Conclusions

Lexicon

Hierarchical Models

- Mixed Effects Models
- Random Effects (RFX) Models
- Components of Variance

... all the same

... all alluding to multiple sources of variation (in contrast to fixed effects)

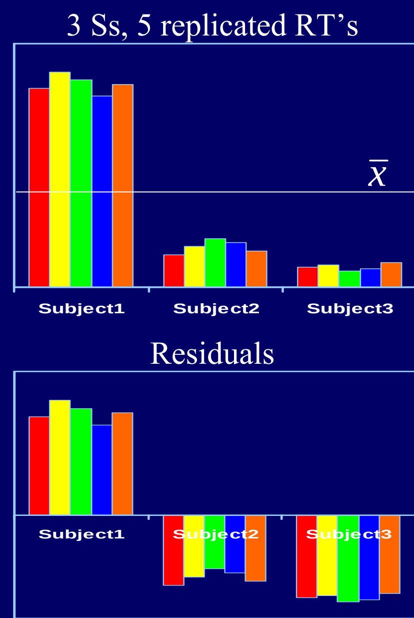
Random Effects Illustration

- Standard linear model

$$Y = X\beta + \varepsilon$$

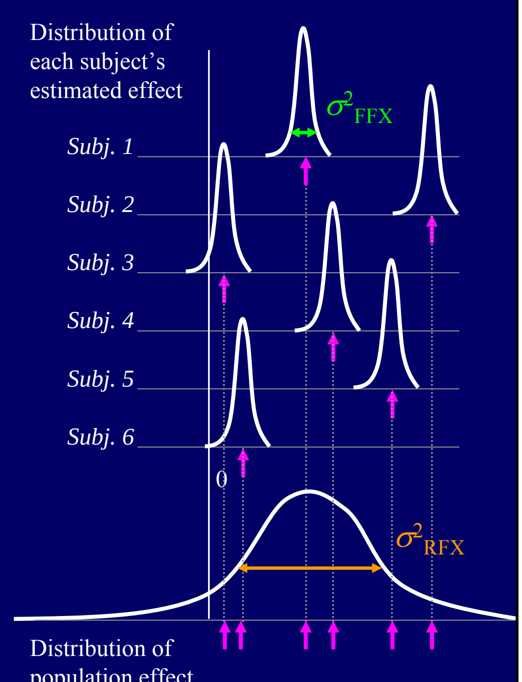
assumes only one source of *iid* random variation

- Consider this RT data
- Here, two sources
  - Within subject var.
  - Between subject var.
  - Causes dependence in ε

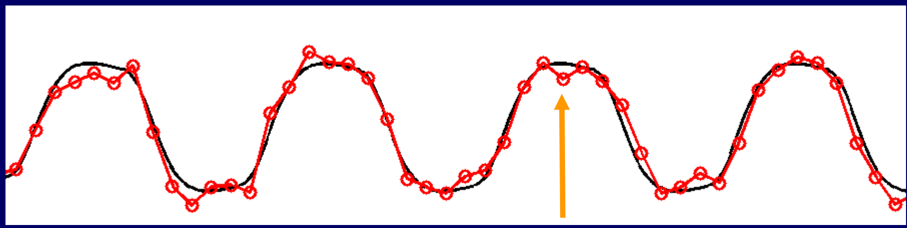


Fixed vs. Random Effects in fMRI

- Fixed Effects
  - Intra-subject variation suggests *all these subjects* different from zero
- Random Effects
  - Intersubject variation suggests *population* not very different from zero

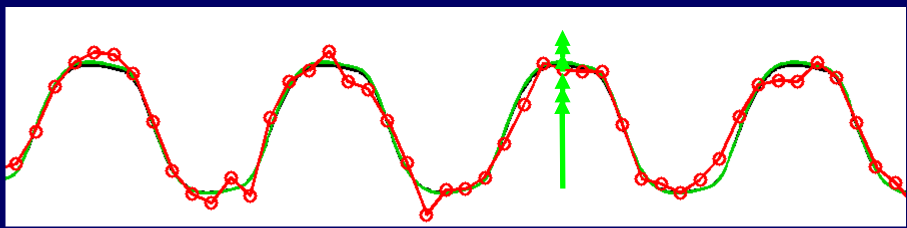


Fixed Effects



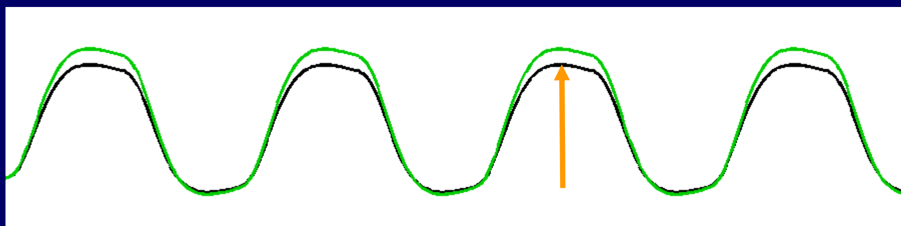
- Only variation (over subjects/sessions) is measurement error
- True Response magnitude is ***fixed***

Random/Mixed Effects



- Two sources of variation
  - Measurement error
  - Response magnitude
- Response magnitude is ***random***
  - Each subject/session has random magnitude
  -

Random/Mixed Effects



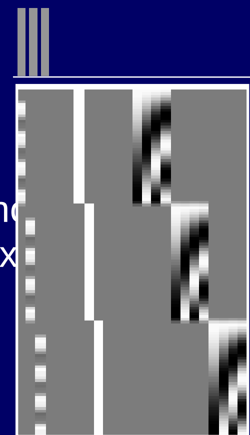
- Two sources of variation
  - Measurement error
  - Response magnitude
- Response magnitude is *random*
  - Each subject/session has random magnitude
  - But note, population mean magnitude is ***fixed***

Fixed vs. Random

- Fixed isn't "wrong," just usually isn't of interest
- Fixed Effects Inference
  - "I can see this effect in this cohort"
- Random Effects Inference
  - "If I were to sample a new cohort from the population I would get the same result"

Two Different Fixed Effects Approaches

- Grand GLM approach
  - Model all subjects at once
  - Good: Mondo DF
  - Good: Can simplify modeling
  - Bad: Assumes common varian
    over subjects at each vox
  - Bad: Huge amount of data



Two Different Fixed Effects Approaches

- Meta Analysis approach
  - Model each subject individually
  - Combine set of T statistics
    - mean(T)$\sqrt{n}$ ~ N(0,1)
    - sum(-logP) ~ χ^2_n
  - Good: Doesn't assume common variance
  - Bad: Not implemented in software
    Hard to interrogate statistic maps

Overview

- Mixed effects motivation
- Evaluating mixed effects methods
- Three methods
  - Summary statistic approach (HF) (SPM96,99,2)
  - SPM2
  - FSL3
- Conclusions

Assessing RFX Models Issues to Consider

- Assumptions & Limitations
  - What must I assume?
    - Independence?
    - “Sphericity”? (aka independence + homogeneous var.)
  - When can I use it
- Efficiency & Power
  - How sensitive is it?
- Validity & Robustness
  - Can I trust the P-values?
  - Are the standard errors correct?
  - If assumptions off, things still OK?

Issues: Assumptions

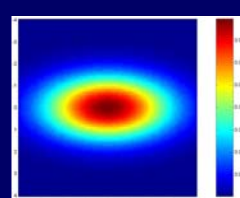
- Distributional Assumptions
  - Gaussian? Nonparametric?
- Homogeneous Variance
  - Over subjects?
  - Over conditions?
- Independence
  - Across subjects?
  - Across conditions/repeated measures
  - Note:
    - Nonsphericity = (Heterogeneous Var) or (Dependence)

What is (and isn't) sphericity?

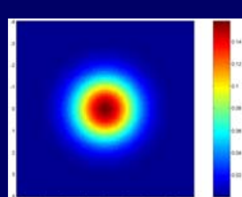
Sphericity $\leftrightarrow$ *iid* $\leftrightarrow$ $N(\boldsymbol{\mu},\boldsymbol{\Sigma}=\sigma^2\mathbf{I})$

$$\Downarrow$$

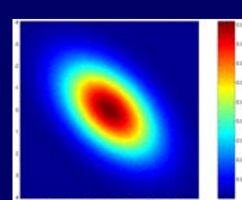
$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$



$$Cov(\varepsilon) = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$



$$Cov(\varepsilon) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



$$Cov(\varepsilon) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Issues: Soft Assumptions Regularization

- Regularization
  - Weakened homogeneity assumption
  - Usually variance/autocorrelation regularized over space
- Examples
  - fmristat - local pooling (smoothing) of $(\sigma^2_{RFX})/(\sigma^2_{FFX})$
  - SnPM - local pooling (smoothing) of σ^2_{RFX}
  - FSL3 - Bayesian (noninformative) prior on σ^2_{RFX}

Issues: Efficiency & Power

- Efficiency: 1/(Estimator Variance)
  - Goes up with n
- Power: Chance of detecting effect
  - Goes up with n
  - *Also* goes up with degrees of freedom (DF)
    - DF accounts for uncertainty in estimate of σ^2_{RFX}
    - Usually DF and n yoked, e.g. DF = n-p

Issues: Validity

- Are P-values accurate?
  - I reject my null when $P < 0.05$
    Is my risk of false positives controlled at 5%?
  - "Exact" control
    - FPR = α
  - Valid control (possibly conservative)
    - FPR $\leq \alpha$
- Problems when
  - Standard Errors inaccurate
  - Degrees of freedom inaccurate

Overview

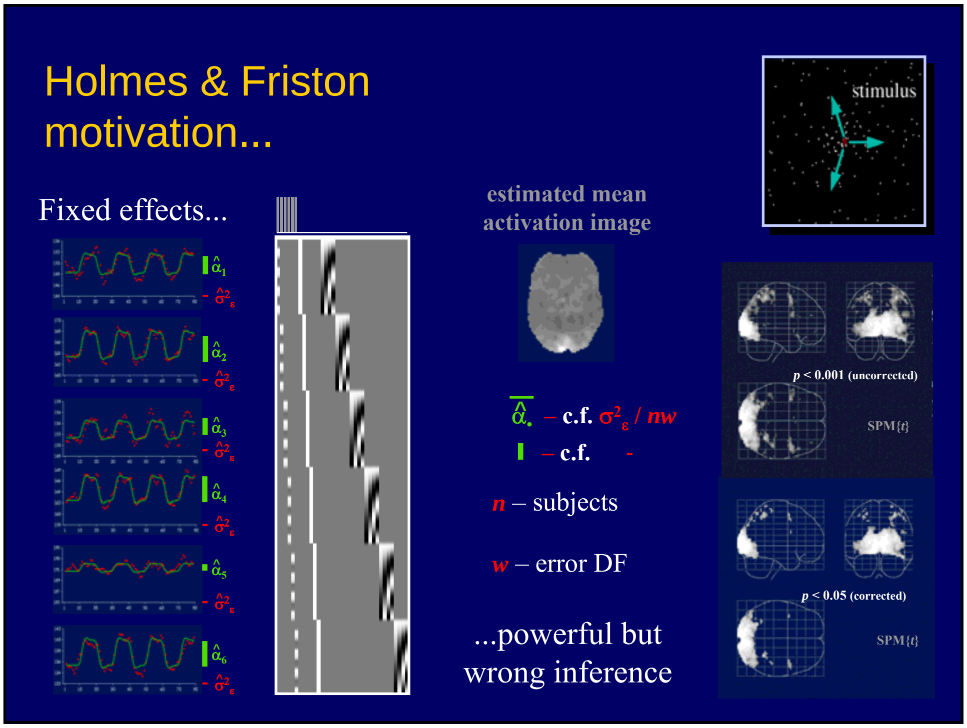
- Mixed effects motivation
- Evaluating mixed effects methods
- Three Methods
  - Summary statistic approach (HF) (SPM96,99,2)
  - SPM2
  - FSL3
- Conclusions

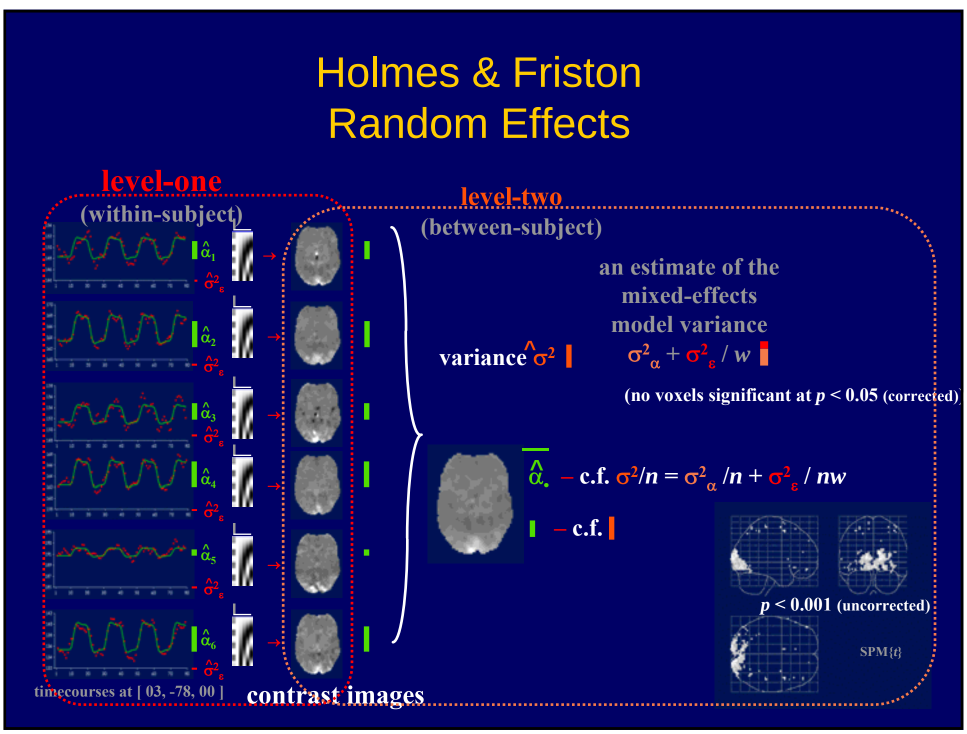
Overview

- Mixed effects motivation
- Evaluating mixed effects methods
- Three Methods
  - Summary statistic approach (HF) (SPM96,99,2)
  - SPM2
  - FSL3
- Conclusions

Method 1: Holmes & Friston

- Unweighted summary statistic approach
- 1- or 2-sample t test on contrast images
  - Intrasubject variance images not used (c.f. FSL)
- Procedure
  - Fit GLM for each subject i
  - Compute cb_i, contrast estimate
  - Analyze $\{cb_i\}_i$

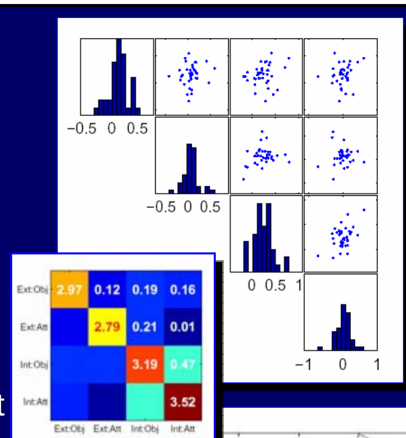

Holmes & Friston
motivation...
Fixed effects...
estimated mean
activation image
stimulus
– c.f. σ²ε / nw
– c.f.
n – subjects
w – error DF
...powerful but
wrong inference
p < 0.001 (uncorrected)
SPM{t}
p < 0.05 (corrected)
SPM{t}

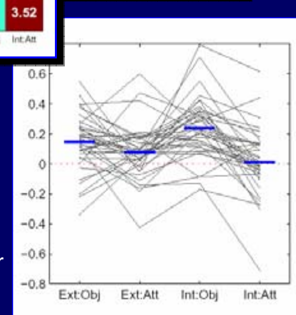

Holmes & Friston
Random Effects
level-one
(within-subject)
level-two
(between-subject)
an estimate of the
mixed-effects
model variance
variance σ²
σ²α + σ²ε / w
(no voxels significant at p < 0.05 (corrected)
– c.f. σ²/n = σ²α /n + σ²ε / nw
– c.f.
p < 0.001 (uncorrected)
SPM{t}
timecourses at [03, -78, 00]
contrast images

Holmes & Friston
Assumptions

- Distribution
  - Normality
  - Independent subjects
- Homogeneous Variance
  - Intrasubject variance homogeneous
    - σ^2_{FFX} same for all subjects
  - Balanced designs

Holmes & Friston
Limitations



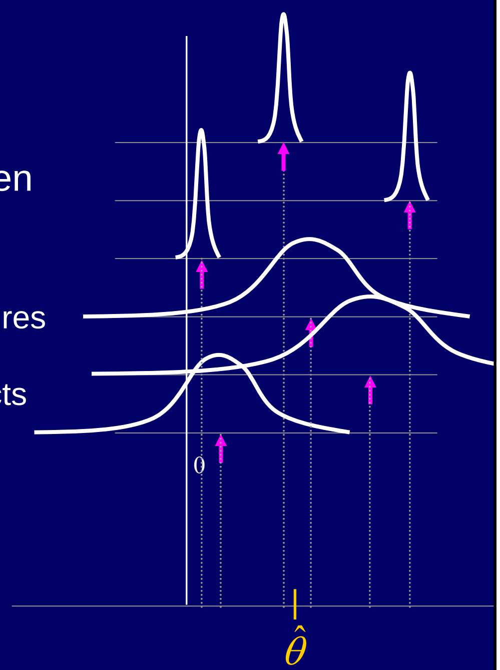


From HBM 2004 Poster WE 253

- Limitations
  - Only single image per subject
  - If 2 or more conditions, Must run separate model for each contrast
- Limitation a strength!
  - No sphericity assumption made on conditions
  - Though nonsphericity itself may be of interest...

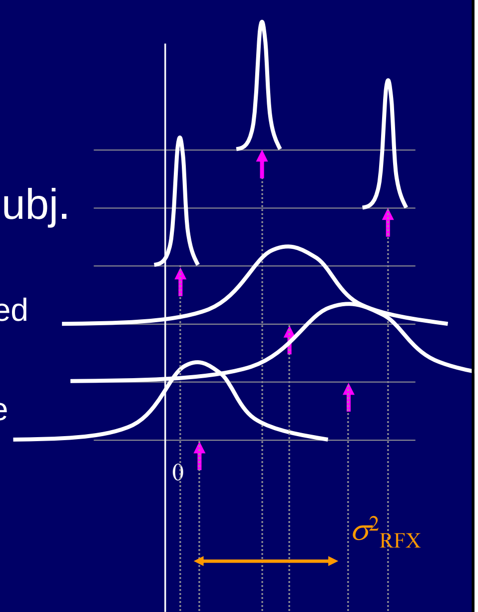
Holmes & Friston
Efficiency

- If assumptions true
  - Optimal, fully efficient
- If σ^2_{FFX} differs between subjects
  - Reduced efficiency
  - Here, optimal $\hat{\theta}$ requires down-weighting the 3 highly variable subjects



Holmes & Friston
Validity

- If assumptions true
  - Exact P-values
- If σ^2_{FFX} differs btw subj.
  - Standard errors OK
    - Est. of σ^2_{RFX} unbiased
  - DF not OK
    - Here, 3 Ss dominate
    - DF < 5 = 6-1



Holmes & Friston
Robustness

- Heterogeneity of σ^2_{FFX} across subjects...
  - How bad is bad?
- Dramatic imbalance *(rough rules of thumb only!)*
  - Some subjects missing 1/2 or more sessions
  - Measured covariates of interest having dramatically different efficiency
    - E.g. Split event related predictor by correct/incorrect
    - One subj 5% trials correct, other subj 80% trials correct
- Dramatic heteroscedasticity
  - A "bad" subject, e.g. head movement, spike artifacts

Overview

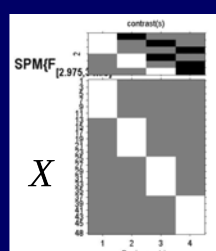
- Mixed effects motivation
- Evaluating mixed effects methods
- Three methods
  - Summary statistic approach (HF) (SPM96,99,2)
  - SPM2
  - FSL3
- Conclusions

Method 2: SPM2

- 1 effect per subject
  - Uses Holmes & Friston approach
- >1 effect per subject
  - Can't use SPM99; must use SPM2+
  - Variance basis function approach used...

SPM2 Notation: iid case

$$\underset{N\times 1}{y} = \underset{N\times p}{X}\ \underset{p\times 1}{\theta} + \underset{N\times 1}{\varepsilon}$$



$$\mathrm{Cor}(\varepsilon) = \lambda I$$

- 12 subjects,
  4 conditions
  - Use F-test to find differences btw conditions
- Standard Assumptions
  - Identical distn
  - Independence
  - "Sphericity"... but here not realistic!

Error covariance

N

N

Multiple Variance Components

$$\underset{N\times 1}{y} = \underset{N\times p}{X}\ \underset{p\times 1}{\theta} + \underset{N\times 1}{\varepsilon}$$

$$\mathrm{Cor}(\varepsilon) = \sum_k \lambda_k Q_k$$

Error covariance

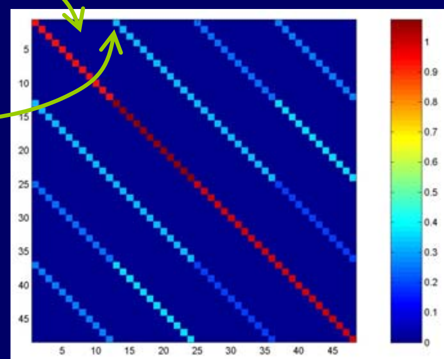
- 12 subjects, 4 conditions
- Measurements btw subjects uncorrelated
- Measurements w/in subjects correlated

Errors can now have different variances and there can be correlations

Allows for ‘nonsphericity’

N

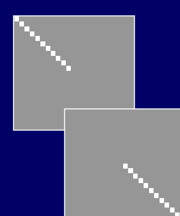
N



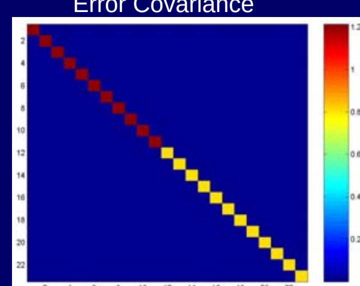
Non-Sphericity Modeling

- Errors are independent but not identical
  - Eg. Two Sample T
    Two basis elements

Q_k’s:



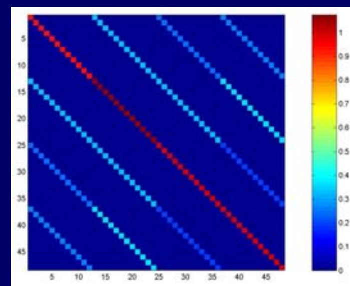
Error Covariance



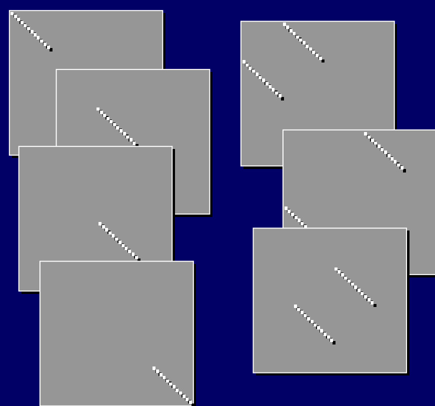
Non-Sphericity Modeling

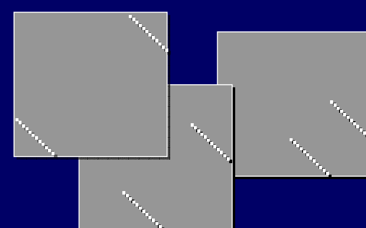
- Errors are not independent and not identical

Error Covariance



Q_k's:





Case Study: SPM2

- Assumptions & Limitations
  - $\mathrm{Cor}(\varepsilon) = \sum_k \lambda_k Q_k$ assumed to globally homogeneous
  - λ_k's only estimated from voxels with large F
  - Most realistically, Cor(ε) spatially heterogeneous
  - Intrasubject variance assumed homogeneous

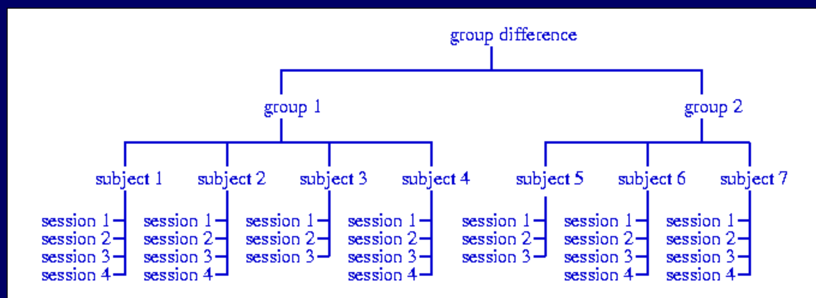
Case Study: SPM2

- Efficiency & Power
  - If assumptions true, fully efficient
- Validity & Robustness
  - P-values could be wrong (over or under) if local Cor(ε) very different from globally assumed
  - Stronger assumptions than Holmes & Friston

Overview

- Mixed effects motivation
- Evaluating mixed effects methods
- Three methods
  - Summary statistic approach (HF) (SPM96,99,2)
  - SPM2
  - FSL3
- Conclusions

FSL3: Full Mixed Effects Model

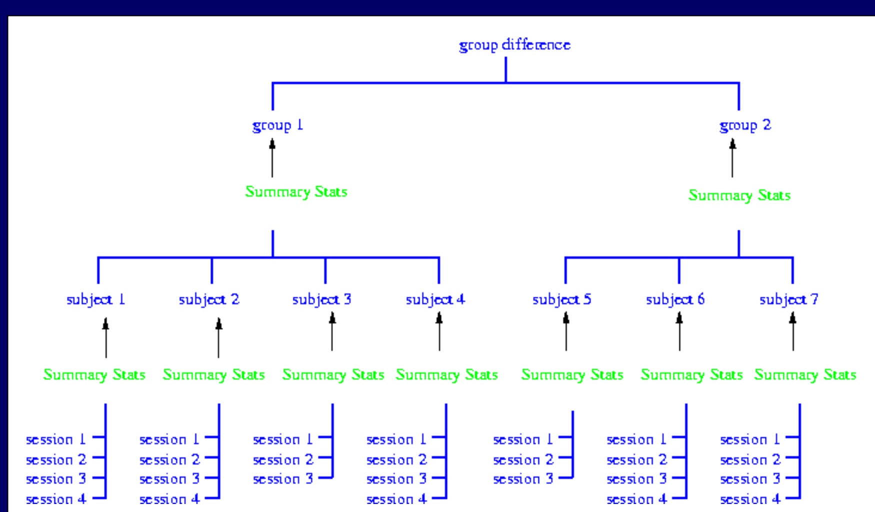


$Y = X_K \beta_K + \varepsilon_K$ First-level, combines sessions

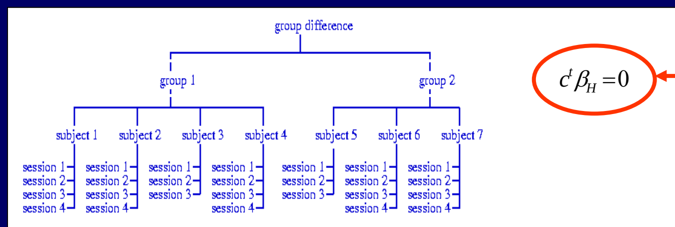
$\beta_K = X_G \beta_G + \varepsilon_G$ Second-level, combines subjects

$\beta_G = X_H \beta_H + \varepsilon_H$ Third-level, combines/compares groups

FSL3: Summary Statistics

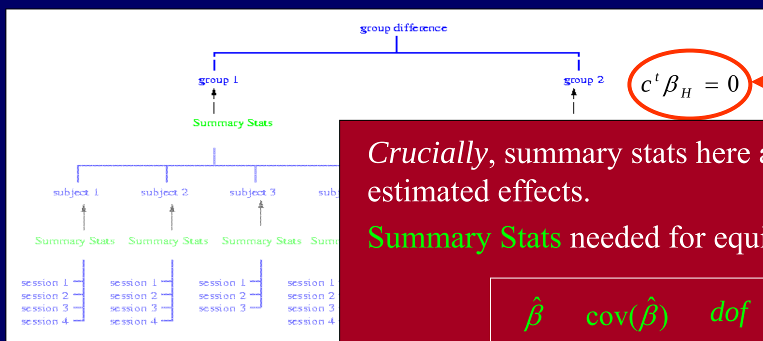


Summary Stats Equivalence



$c^t \beta_H = 0$

$\equiv$



$c^t \beta_H = 0$

Crucially, summary stats here are *not* just estimated effects.

Summary Stats needed for equivalence:

$\hat{\beta}$ $\quad$ $\mathrm{cov}(\hat{\beta})$ $\quad$ *dof*

Beckman *et al.*, 2003

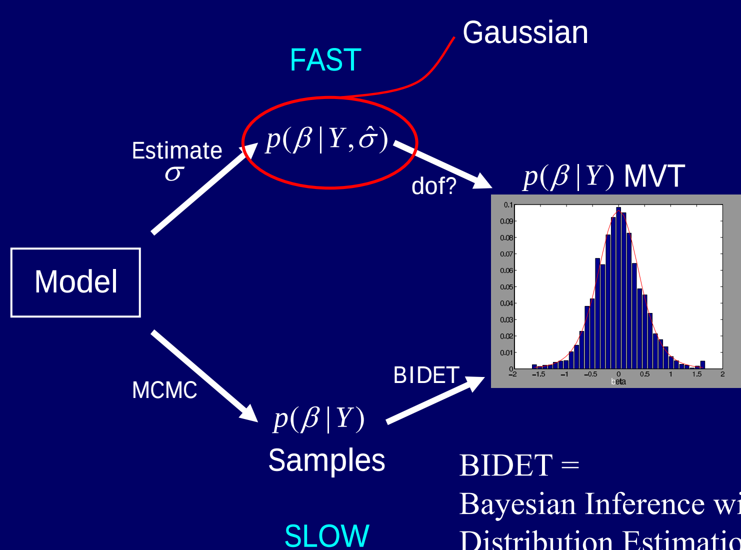
Case Study: FSL3's FLAME

- Uses summary-stats model equivalent to full Mixed Effects model
- Doesn't assume intrasubject variance is homogeneous
  - Designs can be unbalanced
  - Subjects measurement error can vary

Case Study: FSL3's FLAME

- Bayesian Estimation
  - Priors, priors, priors
  - Uses reference prior
- Final inference on posterior of β
  - $\beta \mid y$ has Multivariate T dist^n (MVT) but with unknown *dof*

Approximating MVTs



BIDET =
Bayesian Inference with
Distribution Estimation using T

Overview

- Mixed effects motivation
- Evaluating mixed effects methods
- Three methods
  - Summary statistic approach (HF) (SPM96,99,2)
  - SPM2
  - FSL3
- Conclusions

Conclusions

- Random Effects crucial for pop. inference
- Different methods available, each differing on...
  - Assumptions
  - Efficiency/Power
  - Validity & Robustness

References for *four* RFX Approaches in fMRI

- Holmes & Friston (HF)
  - Summary Statistic approach (contrasts only)
    - Holmes & Friston (HBM 1998). Generalisability, Random Effects & Population Inference. NI, 7(4 (2/3)):S754, 1999.
- Holmes *et al.* (SnPM)
  - Permutation inference on summary statistics
    - Nichols & Holmes (2001). Nonparametric Permutation Tests for Functional Neuroimaging: A Primer with Examples. HBM, 15;1-25.
    - Holmes, Blair, Watson & Ford (1996). Nonparametric Analysis of Statistic Images from Functional Mapping Experiments. JCBFM, 16:7-22.
- Friston *et al.* (SPM2)
  - Empirical Bayesian approach
    - Friston et al. Classical and Bayesian inference in neuroimaging: theory. NI 16(2):465-483, 2002
    - Friston et al. Classical and Bayesian inference in neuroimaging: variance component estimation in fMRI. NI: 16(2):484-512, 2002.
- Beckmann et al. & Woolrich et al. (FSL3)
  - Summary Statistics (contrast estimates *and* variance)
    - Beckmann, Jenkinson & Smith. General Multilevel linear modeling for group analysis in fMRI. NI 20(2):1052-1063 (2003)
    - Woolrich, Behrens et al. Multilevel linear modeling for fMRI group analysis using Bayesian inference. NI 21:1732-1747 (2004)