

# Inference of statistical significance using Permutation Approach

Thomas Nichols, Ph.D.  
University of Warwick

Peter Kochunov  
University of Maryland

Habib Ganjgahi  
Statistics, Warwick

Anderson Winkler  
FMRIB, Oxford

SOLAR Workshop, IIGC 2015  
January 20, 2015

# Overview

- Background review
  - Voxel, cluster, TFCE tests
  - Multiple Testing Problem
    - FWE via permutation
    - FDR
- Permutation on non-independent subjects
  - Heritability Inference
  - Association Inference

# Different Ways to Test...

| Image Feature | Test Statistic                                                                                                                                                                                  |
|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Voxel         | 1. Statistic image value                                                                                                                                                                        |
| Cluster       | 1. Cluster size in voxels<br>2. Cluster size in units of resolution<br>3. Combination, Joint Peak-Cluster<br>4. Combination, Cluster Mass<br>5. Combination, Threshold-Free Cluster Enhancement |
| Set           | 1. Cluster count                                                                                                                                                                                |
| Peak          | 1. Statistic image value                                                                                                                                                                        |

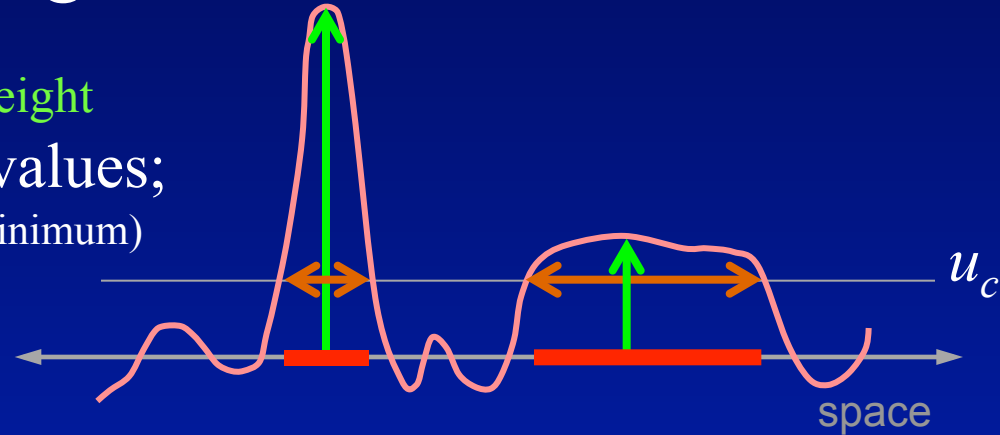
# Different Ways to Test...

| Image Feature | Test Statistic                                                                                                                                                                                  |
|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Voxel         | 1. Statistic image value                                                                                                                                                                        |
| Cluster       | 1. Cluster size in voxels<br>2. Cluster size in units of resolution<br>3. Combination, Joint Peak-Cluster<br>4. Combination, Cluster Mass<br>5. Combination, Threshold-Free Cluster Enhancement |
| Set           | 1. Cluster count                                                                                                                                                                                |
| Peak          | 1. Statistic image value                                                                                                                                                                        |

# Combining Cluster Size with Intensity Information

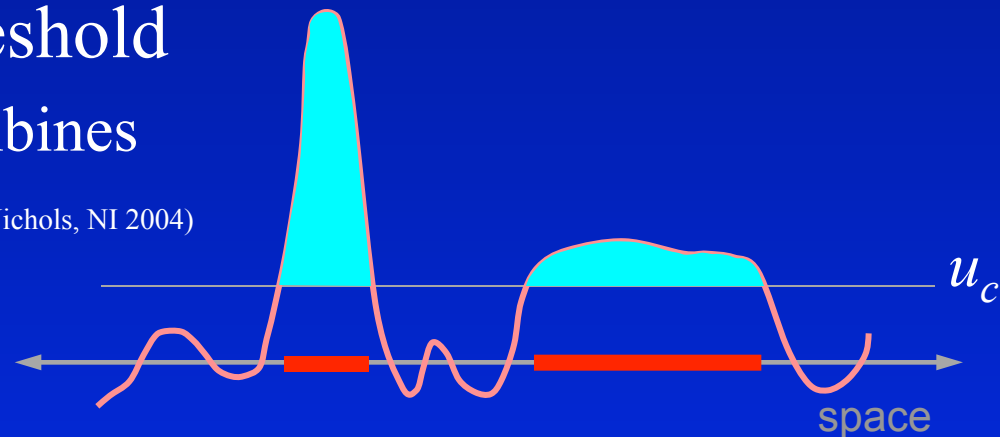
- Peak-Height combining Poline *et al.*, NeuroImage 1997

- Minimum  $P_{\text{extent}}$  &  $P_{\text{height}}$ 
  - Take better of two P-values;  
(use RFT to correct for taking minimum)
- Can catch small, intense clusters



- Cluster mass Bullmore *et al.*, IEEE Trans Med Img 1999

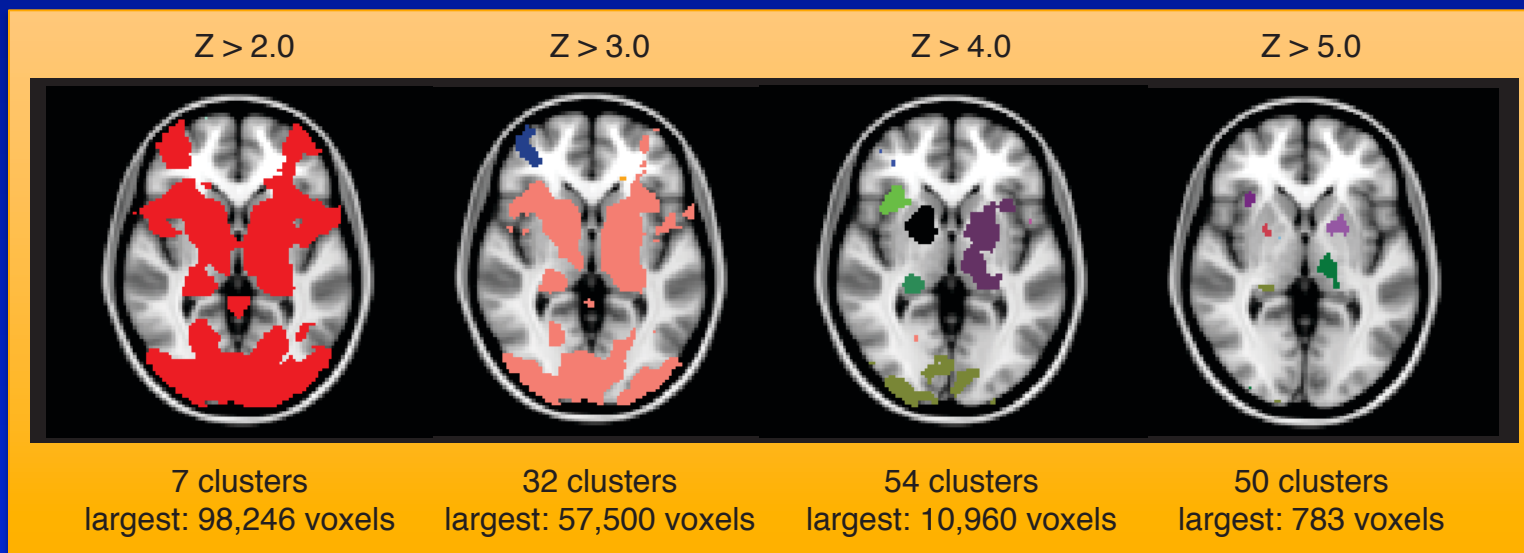
- Integral  $M$  above threshold
  - More powerfully combines peak & height (Hayasaka & Nichols, NI 2004)



- Both are still cluster inference methods!

# The Pesky Cluster Forming Threshold $u_c$

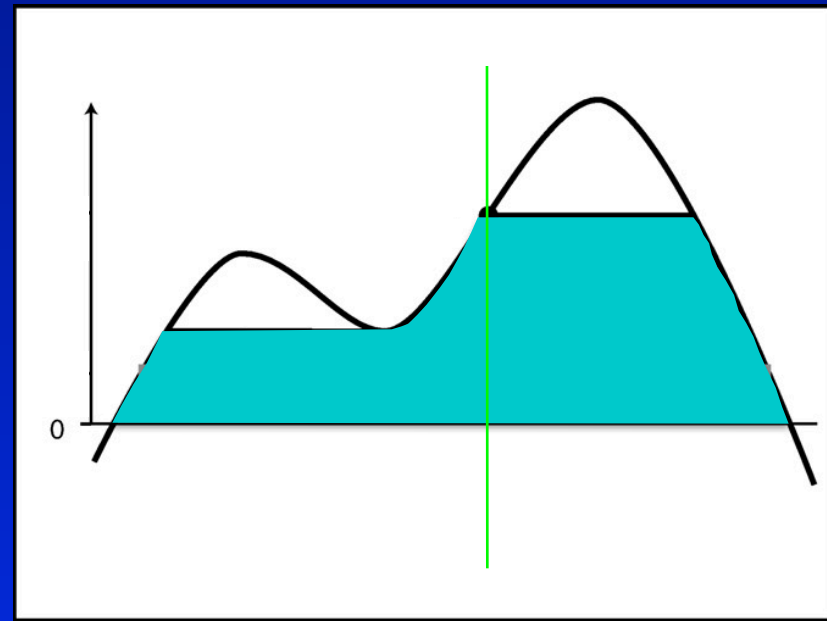
- Cluster inference is highly sensitive to cluster-forming threshold  $u_c$ 
  - Set too low, one big blob
  - Set too high, miss all the signal



# Threshold-Free Cluster Enhancement (TFCE)

- A cluster-informed voxel-wise statistic
- Consider cluster mass voxel-wise, for every  $u_c$ !
  - For a given voxel, sum up all clusters ‘below’
    - For all possible  $u_c$ , add up all clusters that contain that voxel
  - But this would give low  $u_c$ ’s too much weight
    - Low  $u_c$ ’s give big clusters just by chance

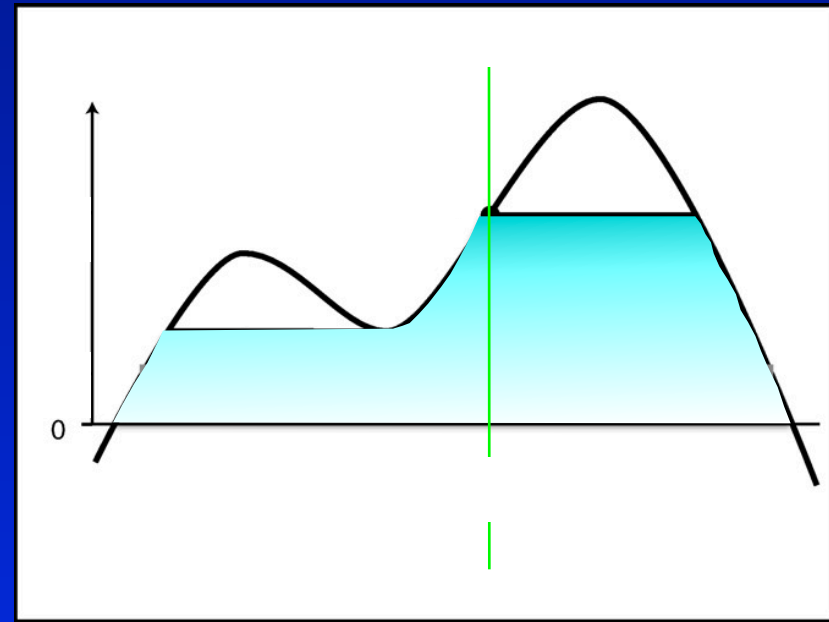
Smith & Nichols, NI 2009



# Threshold-Free Cluster Enhancement (TFCE)

- A cluster-informed voxel-wise statistic
- Consider cluster mass voxel-wise, for every  $u_c$ !
  - For a given voxel, sum up all clusters ‘below’
    - For all possible  $u_c$ , add up all clusters that contain that voxel
  - But this would give low  $u_c$ ’s too much weight
    - Low  $u_c$ ’s give big clusters just by chance
- Solution: Down-weight according to  $u_c$  !

Smith & Nichols, NI 2009

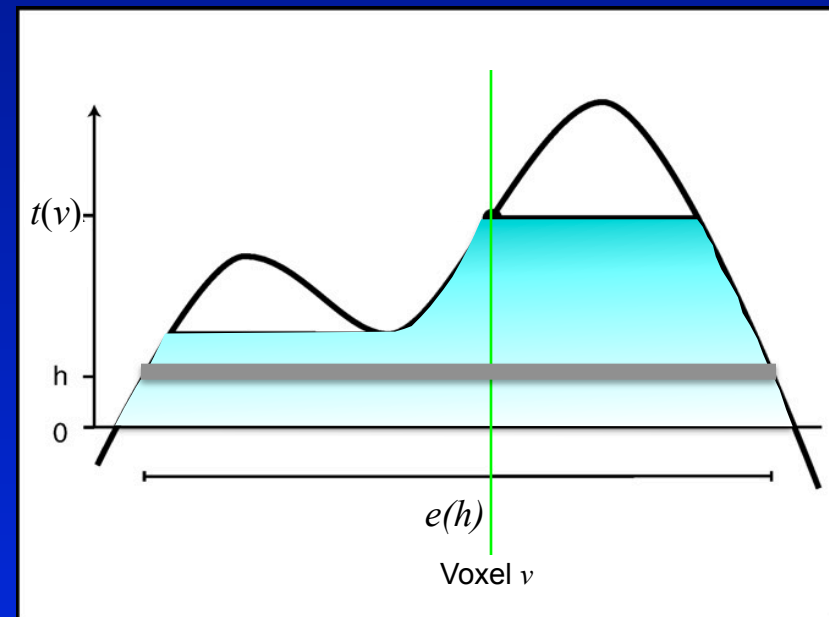


# Threshold-Free Cluster Enhancement (TFCE)

- TFCE Statistic for voxel  $v$

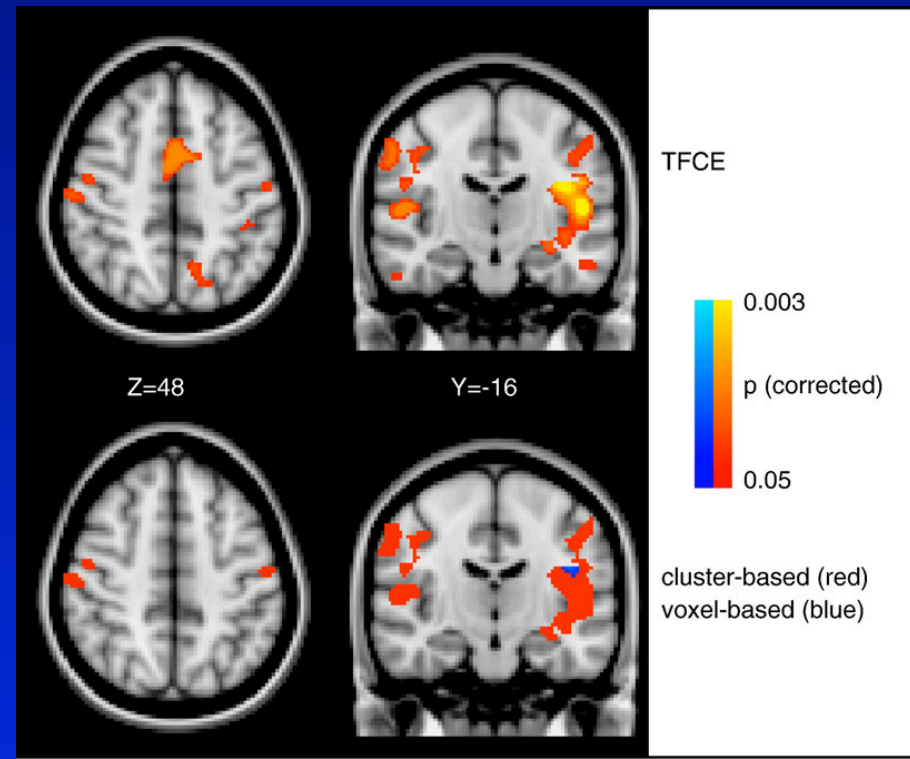
$$TFCE(v) = \int_0^{t(v)} h^H e(h)^E dh \approx \sum_{0, \delta, 2\delta, \dots, t(v)} h^H e(h)^E \delta$$

- Parameters  $H$  &  $E$   
control balance between  
cluster & height  
information
  - $H=2$  &  $E=1/2$  as  
motivated by theory



# TFCE Redux

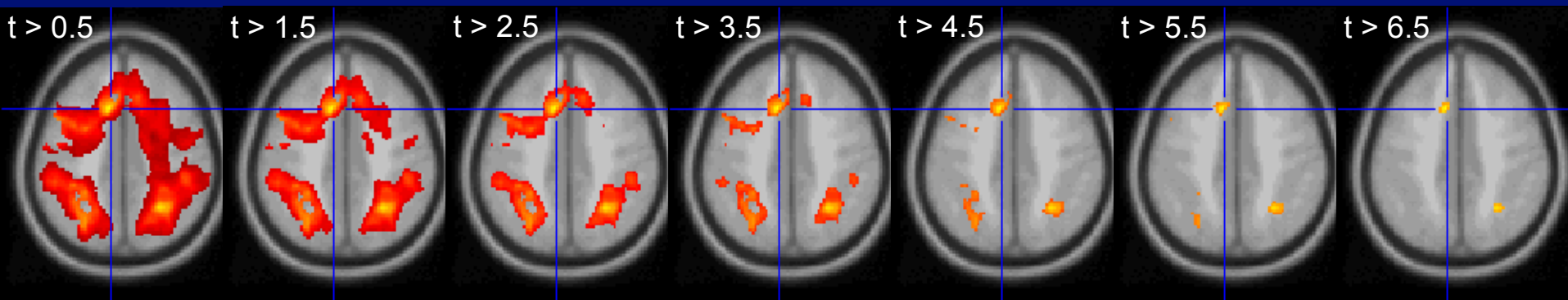
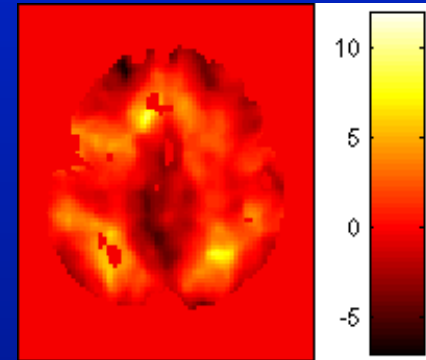
- Avoids choice of cluster-forming threshold  $u_c$
- Generally more sensitive than cluster-wise
- But yet less specific
  - Inference is on some cluster for some  $u_c$
  - “Support” of effect could extend far from significant voxels
- Implementation
  - Currently only FSL’s randomise



# Multiple testing problem...

# Multiple Testing Problem

- Massively Univariate Modeling
  - Fit model at each voxel
  - Create statistic images of effect
- Which of 100,000 voxels are sig?
  - $\alpha=0.05 \Rightarrow 5,000$  false positive voxels
- Which of (random number, say) 100 clusters significant?
  - $\alpha=0.05 \Rightarrow 5$  false positives clusters



# Solutions:

## Measuring False Positives

- Familywise Error Rate (FWER)
  - Familywise Error
    - Existence of one or more false positives
  - FWER is probability of familywise error
- False Discovery Rate (FDR)
  - R voxels declared active, V falsely so
    - Observed false discovery rate:  $V/R$
  - $FDR = E(V/R)$

# Other Solutions

- Bonferroni
- Maximum Distribution Methods
  - Random Field Theory
  - Permutation

# FEW Solutions: Bonferroni

- For a statistic image  $T...$ 
  - $T_i$   $i^{\text{th}}$  voxel of statistic image  $T$
- ...use  $\alpha = \alpha_0/V$ 
  - $\alpha_0$  FWER level (e.g. 0.05)
  - $V$  number of voxels
  - $u_\alpha$   $\alpha$ -level statistic threshold,  $P(T_i \geq u_\alpha) = \alpha$
- By Bonferroni inequality...

$$\begin{aligned}\text{FWER} &= P(\text{FWE}) \\ &= P(\cup_i \{T_i \geq u_\alpha\} \mid H_0) \\ &\leq \sum_i P(T_i \geq u_\alpha \mid H_0) \\ &= \sum_i \alpha \\ &= \sum_i \alpha_0 / V = \alpha_0\end{aligned}$$

Conservative under correlation

|              |           |
|--------------|-----------|
| Independent: | $V$ tests |
| Some dep.:   | ? tests   |
| Total dep.:  | 1 test    |

# FWER MTP Solutions

- Bonferroni
- Maximum Distribution Methods
  - Random Field Theory
  - Permutation

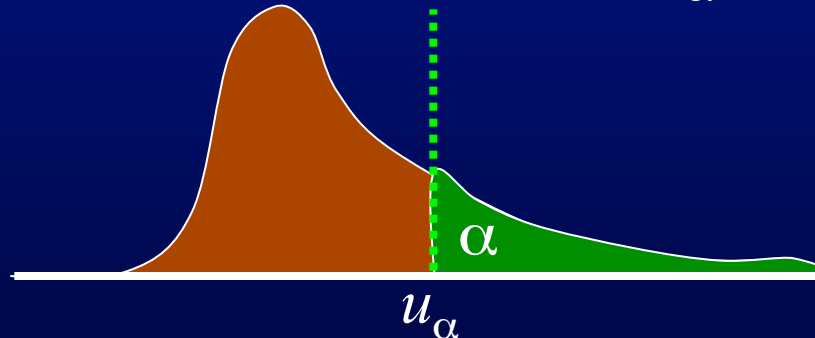
# FWER MTP Solutions: Controlling FWER w/ Max

- FWER & distribution of maximum

$$\begin{aligned}\text{FWER} &= P(\text{FWE}) \\ &= P(\text{One or more voxels} \geq u \mid H_o) \\ &= P(\text{Max voxel} \geq u \mid H_o)\end{aligned}$$

- 100(1- $\alpha$ )%ile of max dist<sup>n</sup> controls FWER

$$\text{FWER} = P(\text{Max voxel} \geq u_\alpha \mid H_o) \leq \alpha$$

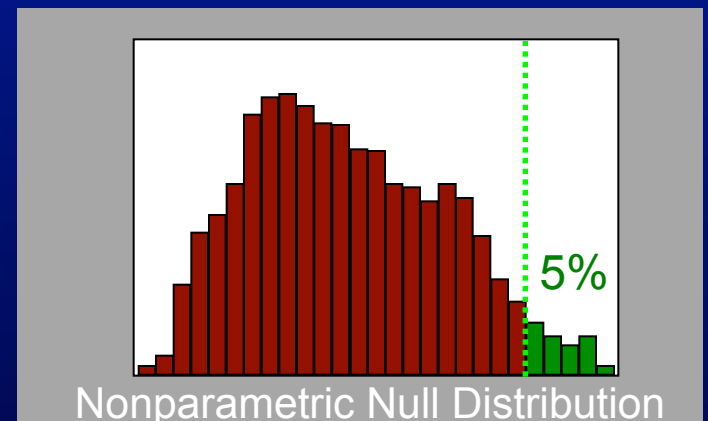
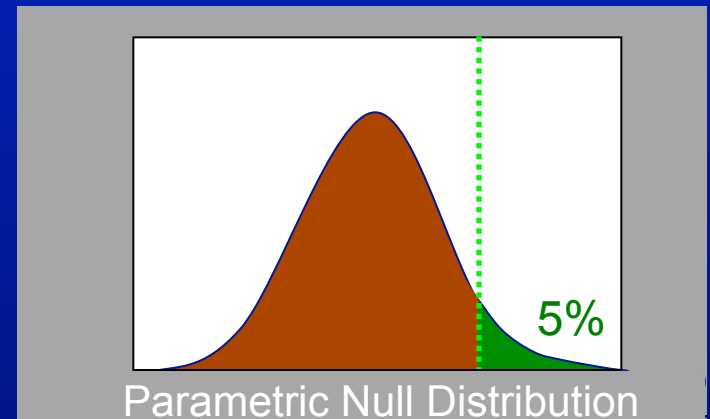


# FWER MTP Solutions

- Bonferroni
- Maximum Distribution Methods
  - Random Field Theory
  - Permutation

# Nonparametric Permutation Test

- Parametric methods
  - Assume distribution of statistic under null hypothesis
- Nonparametric methods
  - Use *data* to find distribution of statistic under null hypothesis
  - Any statistic!



# Permutation Test

## Toy Example

- Data from V1 voxel in visual stim. experiment  
A: Active, flashing checkerboard   B: Baseline, fixation  
6 blocks, ABABAB   Just consider block averages...

| A      | B     | A     | B     | A     | B     |
|--------|-------|-------|-------|-------|-------|
| 103.00 | 90.48 | 99.93 | 87.83 | 99.76 | 96.06 |

- Null hypothesis  $H_0$ 
  - No experimental effect, A & B labels arbitrary
- Statistic
  - Mean difference

# Permutation Test

## Toy Example

- Under  $H_o$ 
  - Consider all equivalent relabelings

|        |        |        |        |
|--------|--------|--------|--------|
| AAABBB | ABABAB | BAAABB | BABBAA |
| AABABB | ABABBA | BAABAB | BBAAAB |
| AABBAB | ABBAAB | BAABBA | BBAABA |
| AABBBA | ABBABA | BABAAB | BBABAA |
| ABAABB | ABBBAA | BABABA | BBBAAA |

# Permutation Test

## Toy Example

- Under  $H_o$ 
  - Consider all equivalent relabelings
  - Compute all possible statistic values

|        |       |        |       |        |       |        |       |
|--------|-------|--------|-------|--------|-------|--------|-------|
| AAABBB | 4.82  | ABABAB | 9.45  | BAAABB | -1.48 | BABBAA | -6.86 |
| AABABB | -3.25 | ABABBA | 6.97  | BAABAB | 1.10  | BBAAAB | 3.15  |
| AABBAB | -0.67 | ABBAAB | 1.38  | BAABBA | -1.38 | BBAABA | 0.67  |
| AABBBA | -3.15 | ABBABA | -1.10 | BABAAB | -6.97 | BBABAA | 3.25  |
| ABAABB | 6.86  | ABBBAA | 1.48  | BABABA | -9.45 | BBBAAA | -4.82 |

# Permutation Test

## Toy Example

- Under  $H_o$ 
  - Consider all equivalent relabelings
  - Compute all possible statistic values
  - Find 95%ile of permutation distribution

|              |              |              |              |
|--------------|--------------|--------------|--------------|
| AAABBB 4.82  | ABABAB 9.45  | BAAABB -1.48 | BABBAA -6.86 |
| AABABB -3.25 | ABABBA 6.97  | BAABAB 1.10  | BBAAAB 3.15  |
| AABBAB -0.67 | ABBAAB 1.38  | BAABBA -1.38 | BBAABA 0.67  |
| AABBBA -3.15 | ABBABA -1.10 | BABAAB -6.97 | BBABAA 3.25  |
| ABAABB 6.86  | ABBBAA 1.48  | BABABA -9.45 | BBBAAA -4.82 |

# Permutation Test

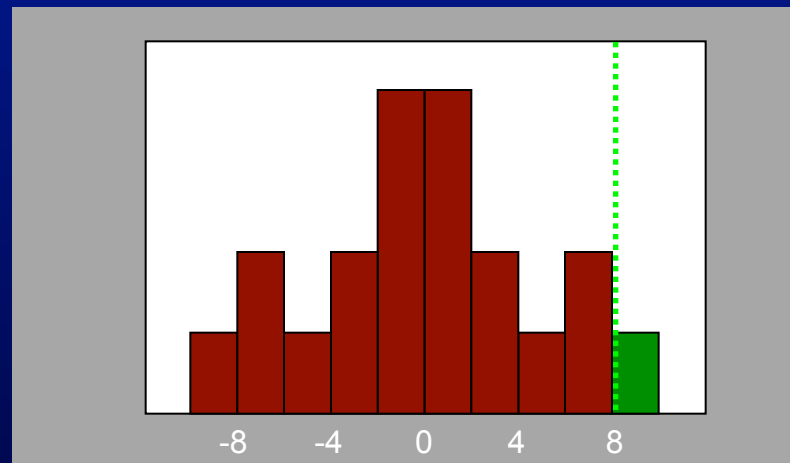
## Toy Example

- Under  $H_o$ 
  - Consider all equivalent relabelings
  - Compute all possible statistic values
  - Find 95%ile of permutation distribution

|              |              |              |              |
|--------------|--------------|--------------|--------------|
| AAABBB 4.82  | ABABAB 9.45  | BAAABB -1.48 | BABBAA -6.86 |
| AABABB -3.25 | ABABBA 6.97  | BAABAB 1.10  | BBAAAB 3.15  |
| AABBAB -0.67 | ABBAAB 1.38  | BAABBA -1.38 | BBAABA 0.67  |
| AABBBA -3.15 | ABBABA -1.10 | BABAAB -6.97 | BBABAA 3.25  |
| ABAABB 6.86  | ABBBA 1.48   | BABABA -9.45 | BBBAAA -4.82 |

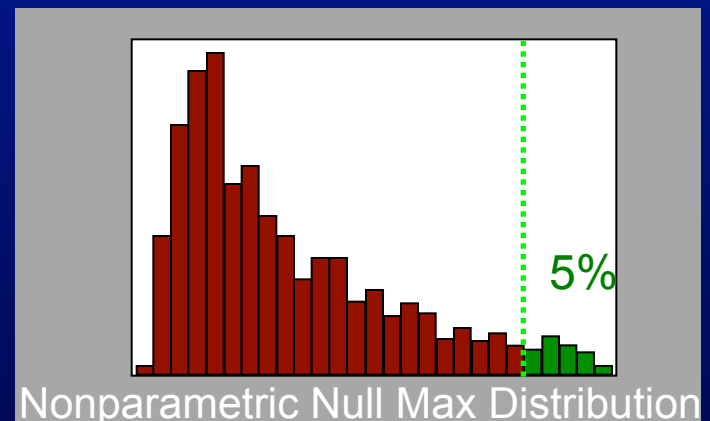
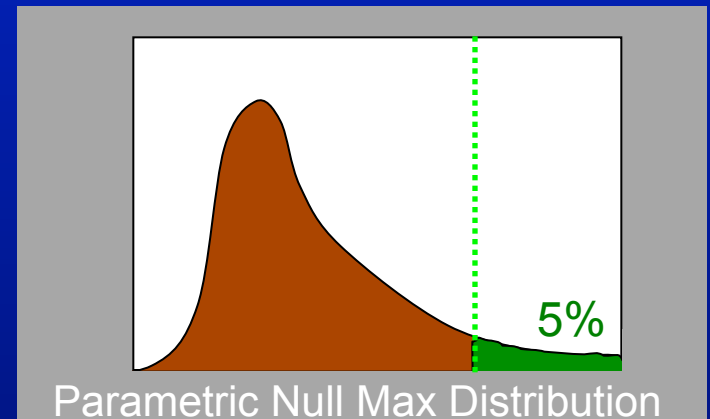
# Permutation Test Toy Example

- Under  $H_o$ 
  - Consider all equivalent relabelings
  - Compute all possible statistic values
  - Find 95%ile of permutation distribution



# Controlling FWER: Permutation Test

- Parametric methods
  - Assume distribution of *max* statistic under null hypothesis
- Nonparametric methods
  - Use *data* to find distribution of *max* statistic under null hypothesis
  - Again, any max statistic!



# Permutation Test Strengths

- Requires only assumption of exchangeability
  - Under  $H_0$ , distribution unperturbed by permutation
  - Allows us to build permutation distribution
- Subjects are exchangeable
  - Under  $H_0$ , each subject's A/B labels can be flipped
- fMRI scans not exchangeable under  $H_0$ 
  - Due to temporal autocorrelation

# Permutation Test Limitations

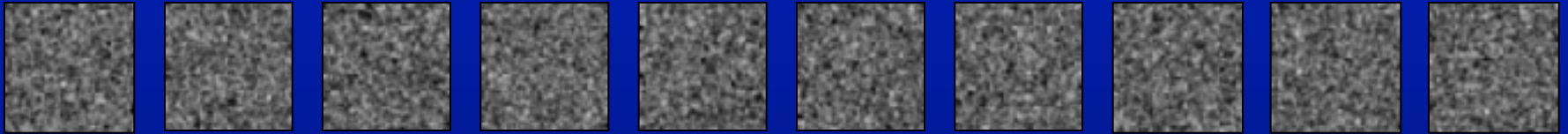
- Computational Intensity
  - Analysis repeated for each relabeling
  - Not so bad on modern hardware
- Implementation Generality
  - Each experimental design type needs unique code to generate permutations
    - Not so bad for population inference with t-tests
- Exchangeability
  - What about related subjects (see below)

# Measuring False Positives

- Familywise Error Rate (FWER)
  - Familywise Error
    - Existence of one or more false positives
  - FWER is probability of familywise error
- False Discovery Rate (FDR)
  - R voxels declared active, V falsely so
    - Observed false discovery rate:  $V/R$
  - $FDR = E(V/R)$

# False Discovery Rate Illustration:

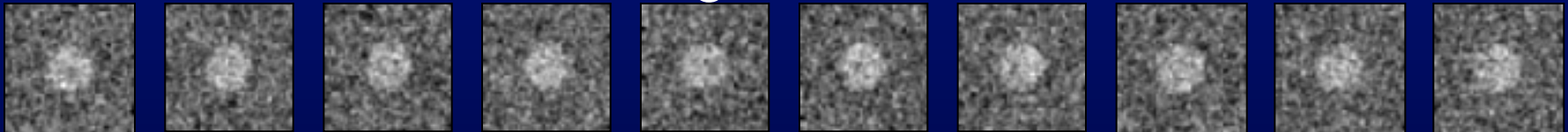
Noise



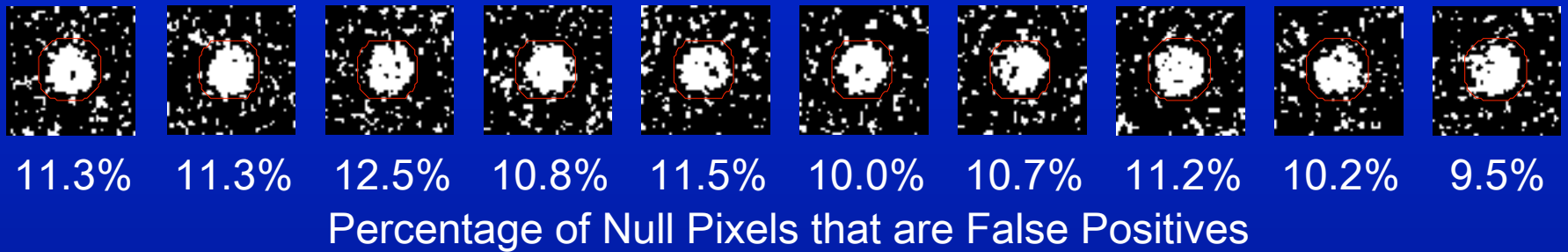
Signal



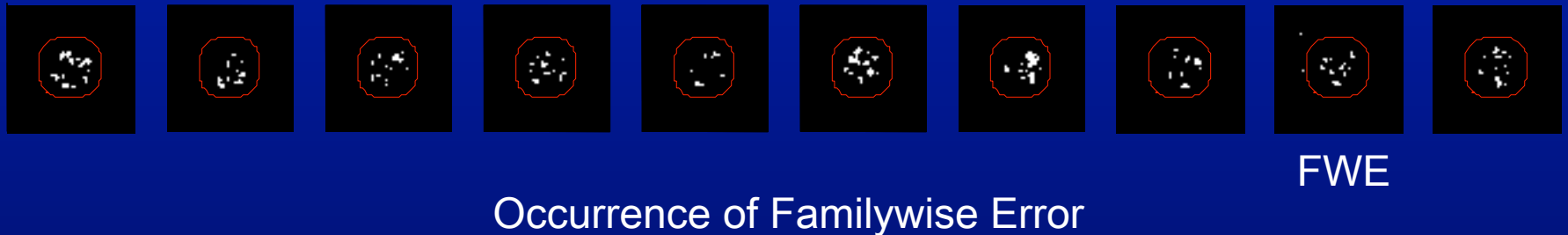
Signal+Noise



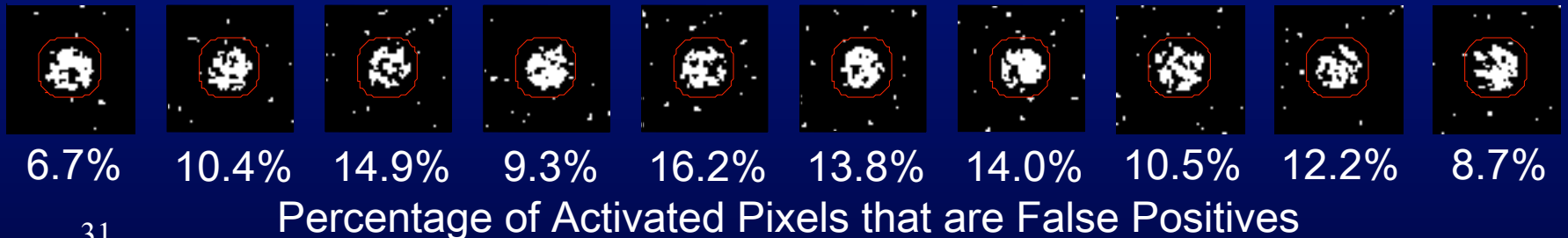
## Control of Per Comparison Rate at 10%



## Control of Familywise Error Rate at 10%



## Control of False Discovery Rate at 10%



# FDR: Conclusions

- New approach to multiple testing
  - Easy to apply
  - Adaptive to the amount of signal
- Limitations
  - Number of false positives can be large for large signals
  - Must be accepted/understood by the field

# Outline

- Background review
  - Voxel, cluster, TFCE tests
  - Multiple Testing Problem
    - FWE via permutation
    - FDR
- Permutation on non-independent subjects
  - Heritability Inference
  - Association Inference

# Modeling Related Individuals

- SOLAR's dependent GLM

$$Y \sim \mathcal{N}(X\beta, \Omega)$$

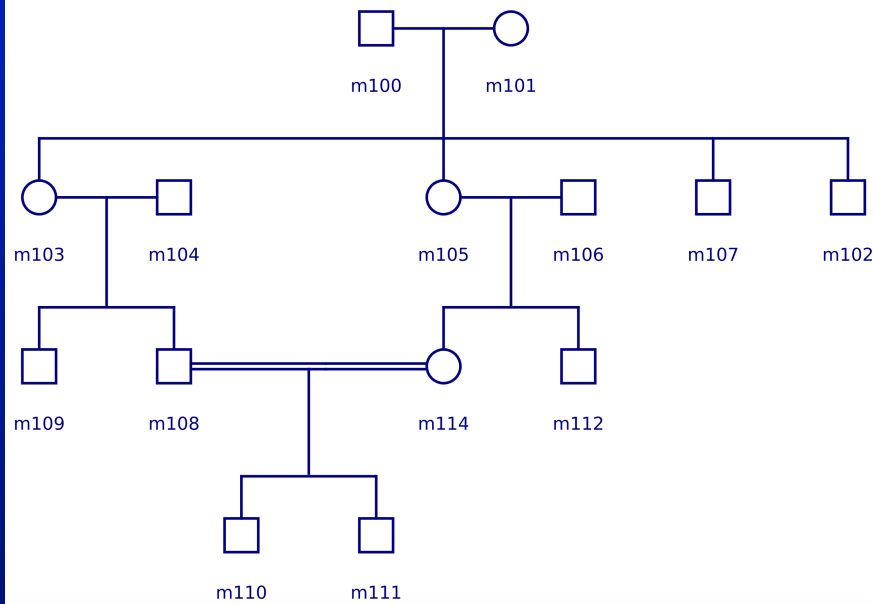
- $Y$  subject data at one voxel
  - FA, BOLD, GM, etc
- $X$  design matrix, coefficients  $\beta$ 
  - age, gender, etc
- $\Omega$  model for covariance

$$\Omega = 2\Phi\sigma_a^2 + I\sigma_e^2 \quad h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

- Dependence determined by kinship structure in  $\Phi$
- Iterative fitting to estimate:  $\beta, \sigma_a^2, \sigma_e^2$

# Pedigrees & Kinship Matrix

Pedigree tree



Kinship matrix  
 $2\Phi$

|      | m100 | m101 | m104 | m106 | m102 | m103 | m105 | m107 | m108 | m109 | m112 | m114 | m110 | m111 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| m100 | 1.00 | 0.00 | 0.00 | 0.00 | 0.50 | 0.50 | 0.50 | 0.50 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| m101 | 0.00 | 1.00 | 0.00 | 0.00 | 0.50 | 0.50 | 0.50 | 0.50 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| m104 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.50 | 0.50 | 0.00 | 0.00 | 0.25 | 0.25 |
| m106 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.50 | 0.50 | 0.25 | 0.25 |
| m102 | 0.50 | 0.50 | 0.00 | 0.00 | 1.00 | 0.50 | 0.50 | 0.50 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| m103 | 0.50 | 0.50 | 0.00 | 0.00 | 0.50 | 1.00 | 0.50 | 0.50 | 0.50 | 0.50 | 0.25 | 0.25 | 0.38 | 0.38 |
| m105 | 0.50 | 0.50 | 0.00 | 0.00 | 0.50 | 0.50 | 1.00 | 0.50 | 0.25 | 0.25 | 0.50 | 0.50 | 0.38 | 0.38 |
| m107 | 0.50 | 0.50 | 0.00 | 0.00 | 0.50 | 0.50 | 0.50 | 1.00 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| m108 | 0.25 | 0.25 | 0.50 | 0.00 | 0.25 | 0.50 | 0.25 | 0.25 | 1.00 | 0.50 | 0.12 | 0.12 | 0.56 | 0.56 |
| m109 | 0.25 | 0.25 | 0.50 | 0.00 | 0.25 | 0.50 | 0.25 | 0.25 | 0.50 | 1.00 | 0.12 | 0.12 | 0.31 | 0.31 |
| m112 | 0.25 | 0.25 | 0.00 | 0.50 | 0.25 | 0.25 | 0.50 | 0.25 | 0.12 | 0.12 | 1.00 | 0.50 | 0.31 | 0.31 |
| m114 | 0.25 | 0.25 | 0.00 | 0.50 | 0.25 | 0.25 | 0.50 | 0.25 | 0.12 | 0.12 | 0.50 | 1.00 | 0.56 | 0.56 |
| m110 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.38 | 0.38 | 0.25 | 0.56 | 0.31 | 0.31 | 0.56 | 1.06 | 0.56 |
| m111 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.38 | 0.38 | 0.25 | 0.56 | 0.31 | 0.31 | 0.56 | 0.56 | 1.06 |

- $2\Phi$  is proportion of shared genetic material

# Maximum Likelihood Estimate

Genetic linear model

$$\mathbf{y}_{n \times 1} = \mathbf{X}_{n \times j} \boldsymbol{\beta}_{j \times 1} + \mathbf{g} + \mathbf{e},$$

Variance

$$V[\mathbf{y}] = \boldsymbol{\Sigma} = \mathbf{K}h^2 + \mathbf{I}(1 - h^2),$$

K is the kinship kernel, h is heritability

# Ln-Likelihood

$$\ln L(\boldsymbol{\beta}, h^2 | \mathbf{y}, \mathbf{X}) = -\frac{1}{2} [N \ln 2\pi + \ln |\boldsymbol{\Sigma}| + \boldsymbol{\delta}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\delta}],$$

$$\boldsymbol{\Sigma} = \mathbf{K}h^2 + \mathbf{I}(1 - h^2)$$

The fitting approach compares goodness of fit of two models  
h=0 (sporadic model) and h>0 (polygenic model)

The second model needs to calculate  $\ln L()$  between 10-20 times  
Each time, a very large matrix (100,000x100,000) is inverted

# Inference on Heritability

- Null hypothesis is sporadic ( $h^2=0$ )

$$H_0 : \sigma_a^2 = 0$$

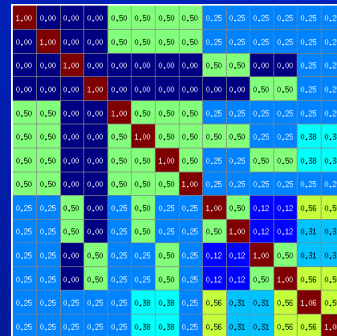
- Test for  $\sigma_a^2$  : Likelihood ratio test (LRT)
  - $\chi^2$ , uncorrected P-values
- Corrected P-values?
  - Voxel-wise FDR: Easy! Just uses  $P_{\text{uncorr}}$
  - Voxel or cluster or TFCE FWE: Not easy...
    - RFT works with GLM, not variance components
    - Permutation?

# Exchangeability failure

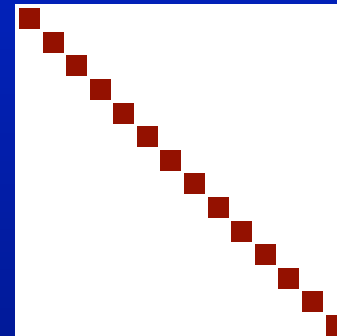
- Degree of dependence depends on  $h^2$

–  $h^2 > 1$

$\Omega = 2$



$\sigma_a^2 +$

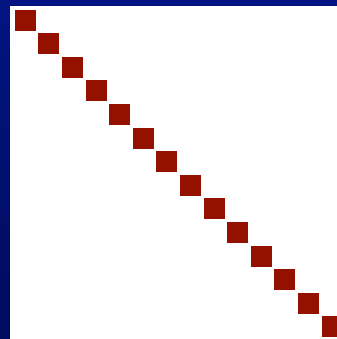


$\sigma_e^2$

- Polygenic matrix is *not* exchangeable

–  $h^2 = 0$

$\Omega =$

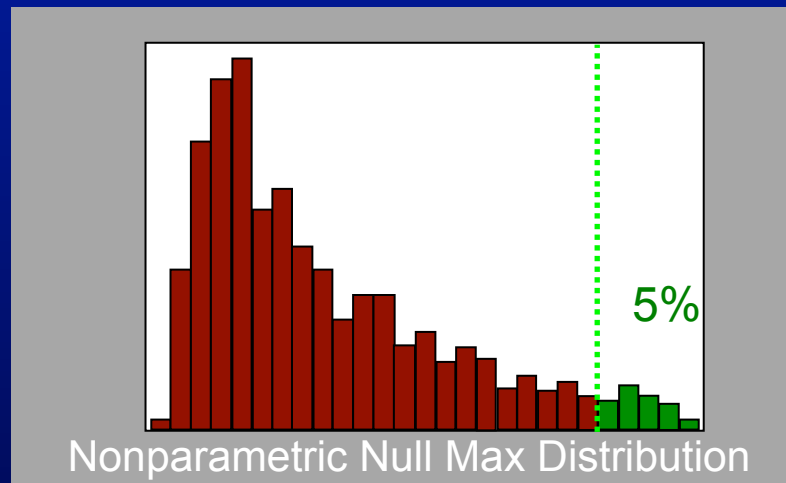


$\sigma_e^2$

- Sporadic matrix is exchangeable!

# FWE Inference with Permutation

- Same as before!
- Compute max LRT/ $\chi^2$  test over brain
- Use max perm. dist<sup>n</sup> to find  $u_{\text{FWE}}$ ,  $P_{\text{FWE}}$



- But wait!? What can we permute!?

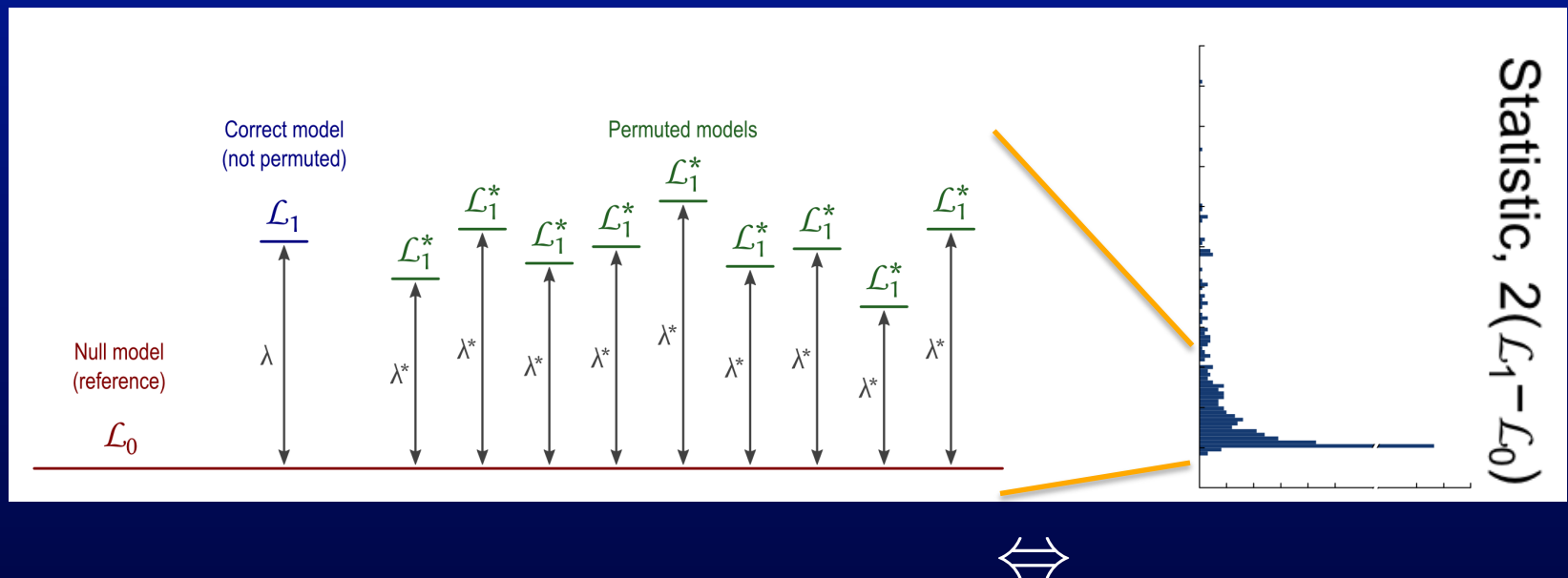
# Permutation Inference on Heritability

- Null hypothesis of no heritability ( $h^2=0$ )

$$H_0 : \sigma_a^2 = 0$$

- Under  $H_0$ , all subjects independent! ( $\Omega_0 = I\sigma_e^2$ )
- Randomly permute, refit, re-estimate  $\beta, \sigma_a^2, \sigma_e^2$

- Test for  $\sigma_a^2$  ....



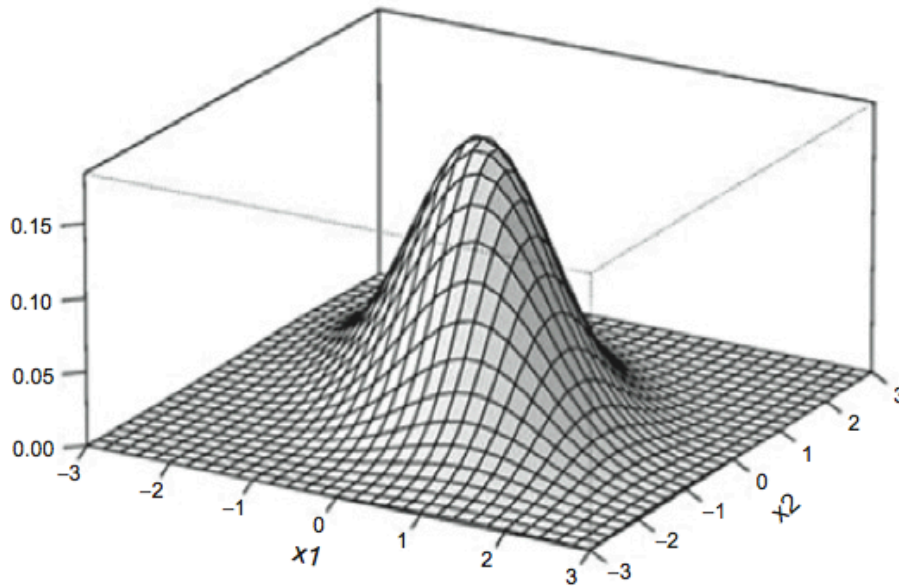
# Accelerating Heritability Inference with Score Test

- SOLAR-Eclipse's Maximum Likelihood Estimation
  - Highly optimized but requires 2 model fits
  - Full MLE is too slow for permutation
- Need approximation solutions that are fast and accurate!

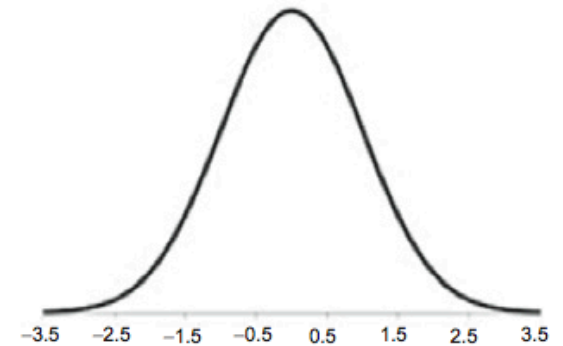
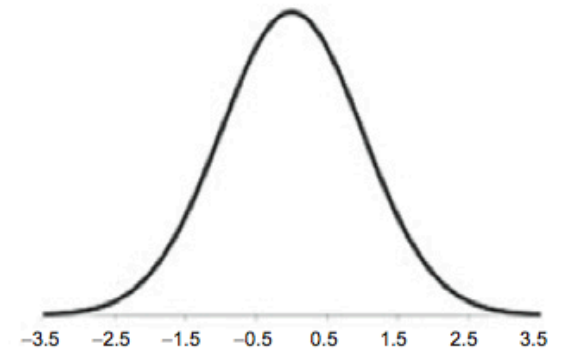
# Eigen Value Transform to Independence

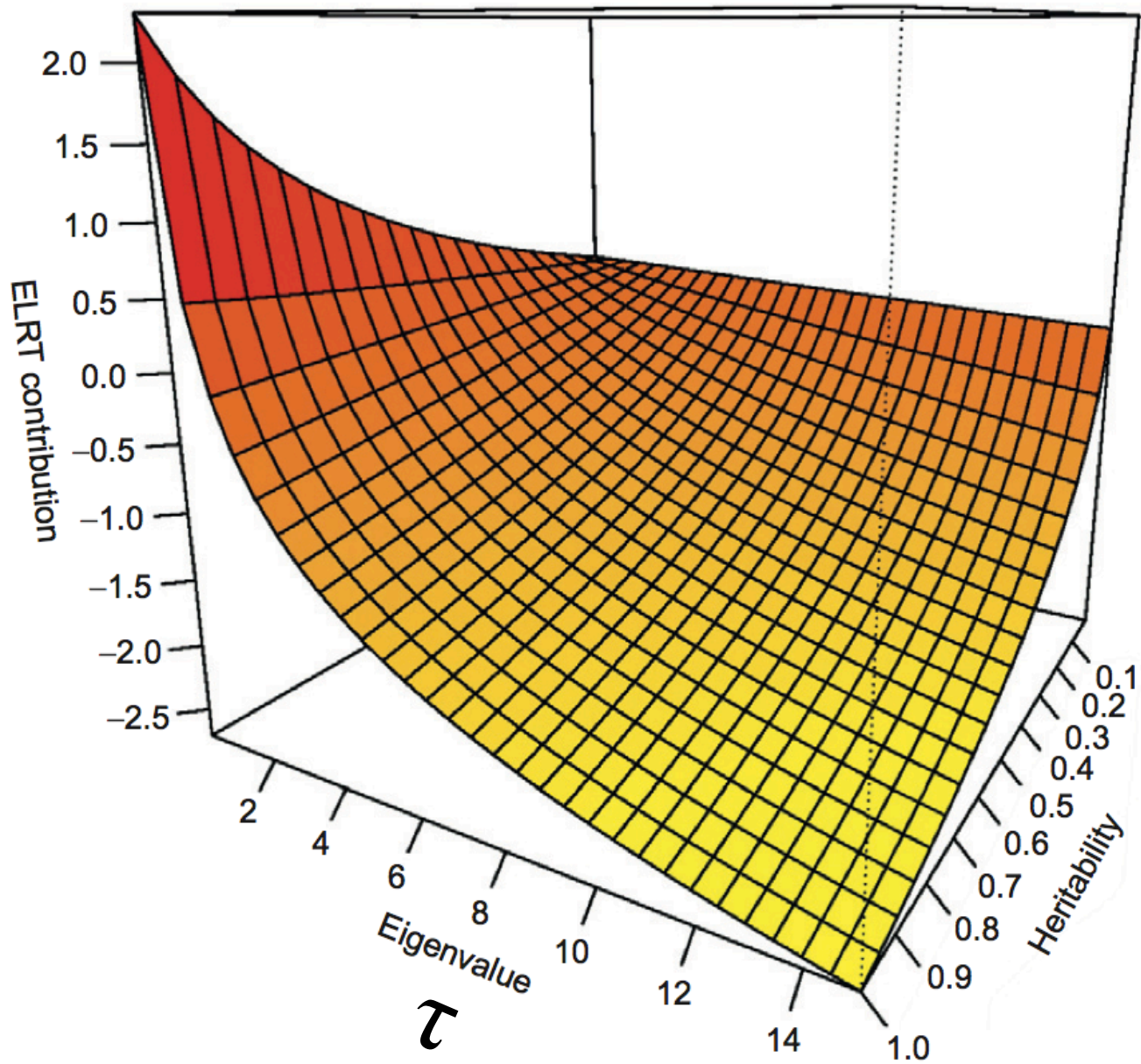
$$\mathbf{S}' : \delta \in V \rightarrow \tau \in W$$

Multivariate (bivariate)  
distribution



Two Orthogonal  
Univariate Distributions





# Eigen Value Decomposition

- Orthogonal transformation of phenotype data to independent space
- Likelihood before – needs inversion

$$\ln L(\boldsymbol{\theta}) = -\frac{1}{2}[N\ln 2\pi + \ln |\boldsymbol{\Sigma}| + \boldsymbol{\delta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}]$$

– Likelihood After – doesn't need inversion

$$\ln L(\boldsymbol{\theta}) = -\frac{1}{2} \left[ N\ln 2\pi + \sum \ln \lambda_i + \sum \frac{\tau_i^2}{\lambda_i} \right]$$

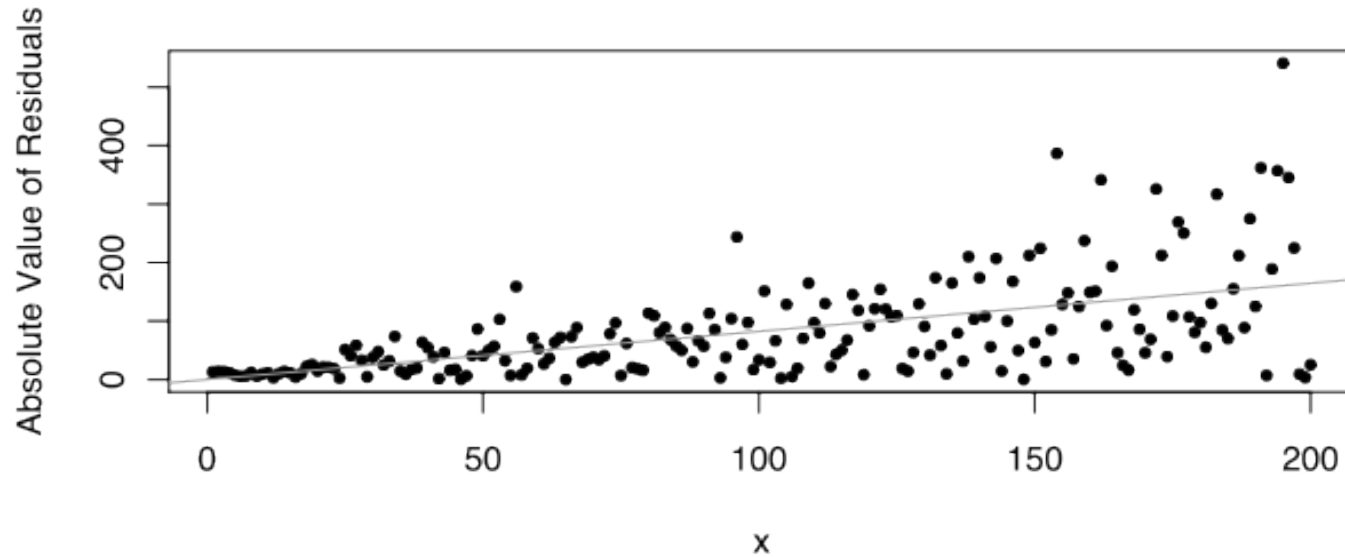
# Solar-Eclipse command

- Activate EVD use
  - option modeltype evd2
  - trait My\_trait
  - polygen
- SOLAR-Eclipse will save
  - My\_trait.evd
  - Contains evd-transformed trait and eigen values
- Greatly accelerates GWAS in family data
  - However, likelihood is still iterative!

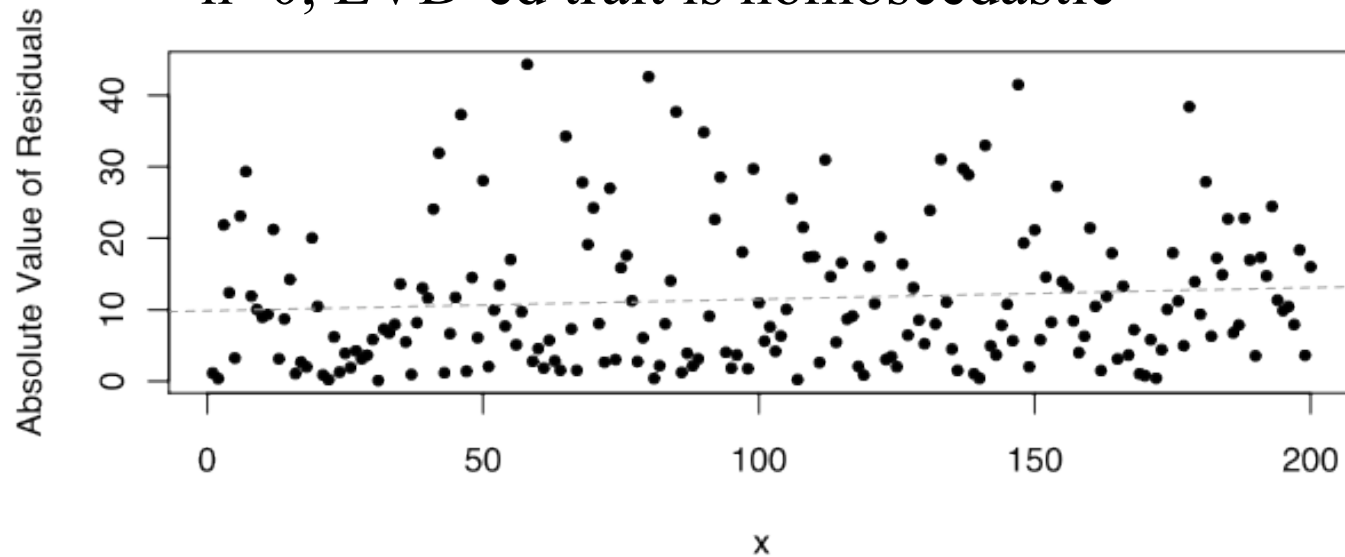
# EVD-transformed trait is

- Independent but heteroscedastic
- Heteroscedasticity
  - When random and independent variables follow a trend
  - Have non-zero regression line like vs. a predictor
    - In our case it is kinship coefficient

$h > 0$ , EVD-ed trait is heteroscedastic



$h = 0$ , EVD-ed trait is homoscedastic



# We can hack heteroscedasity for fast heritability estimates

- To approximate heritability!
- Need only to get EVD transform of the kinship
- Use single-step regression tests to measure heteroscedasity
  - Relate it to heritability
- Produce very fast permutation inference

# Hacked Heritability: Gory Details

- Transformed “EVD” model

$$Y^* = X^* \beta + \epsilon^*$$

- Independent, heteroscedastic error

$$\text{Var}(\epsilon^*) = \Sigma^* = \sigma_A^2 D_g + \sigma_E^2 I$$

- $D_g$  is diagonal matrix of eigenvalues of  $2\Phi$

$$D_g = \text{diag}\{\lambda_{gi}\}$$

- Parameters to estimate:

- Slopes  $\beta$ , genetic variance  $\sigma_A^2$  & error  $\sigma_E^2$

# Hacked Heritability: Regression!

- Newton's method updates

$$\begin{aligned}\hat{\beta}_{j+1} &= \left( X^{*'} (\hat{\Sigma}_j^*)^{-1} X^* \right)^{-1} X^{*'} (\hat{\Sigma}_j^*)^{-1} Y^* \\ \hat{\theta}_{j+1} &= \max \left\{ 0, \left( U' (\hat{\Sigma}_j^{*2})^{-1} U \right)^{-1} U' (\hat{\Sigma}_j^{*2})^{-1} f_j^* \right\}\end{aligned}$$

– Variance parameters collected  $\theta = (\sigma_A^2, \sigma_E^2)$

– Eigenvalue “Design matrix”  $U = [\mathbf{1}, \lambda_g]$

– Squared residuals  $f_i^* = \hat{\epsilon}_i^{*2}$

- Both updates are regression estimates!

–  $\beta$  is weighted LS regression of  $Y^*$  on  $X^*$

–  $\theta$  is weighted LS regression of  $f^*$  on  $U$

# Hacked Heritability: 3 Way!

- “0-step”, unweighted (OLS) regression

*for an  
initial  
guess*

$$\begin{aligned}\hat{\beta}_{\text{OLS}} &= (X^{*'}X^*)^{-1}X^{*'}Y^* \\ \hat{\theta}_{\text{OLS}} &= \max \left\{ 0, (U'U)^{-1}U'f_{\text{OLS}}^* \right\}\end{aligned}$$

- “1-step” of weighted (WLS) regression

*fast,  
remarkably  
accurate*

$$\begin{aligned}\hat{\beta}_{\text{WLS}} &= \left( X^{*'}(\hat{\Sigma}_{\text{OLS}}^*)^{-1}X^* \right)^{-1}X^{*'}(\hat{\Sigma}_{\text{OLS}}^*)^{-1}Y^* \\ \hat{\theta}_{\text{WLS}} &= \max \left\{ 0, \left( U'(\hat{\Sigma}_{\text{OLS}}^{*2})^{-1}U \right)^{-1}U'(\hat{\Sigma}_{\text{OLS}}^{*2})^{-1}f_{\text{OLS}}^* \right\}\end{aligned}$$

- Or, go all the way (fully converge), find

*The slow but  
optimal way,  
as in SOLAR*

$$\hat{\beta}_{\text{ML}}, \hat{\theta}_{\text{ML}}$$

$$\hat{h}_{\text{ML}}^2 = \hat{\sigma}_{\text{A,ML}}^2 / (\hat{\sigma}_{\text{A,ML}}^2 + \hat{\sigma}_{\text{E,ML}}^2)$$

# Fast Heritability: Tests

- 3 flavors of tests

- Likelihood ratio

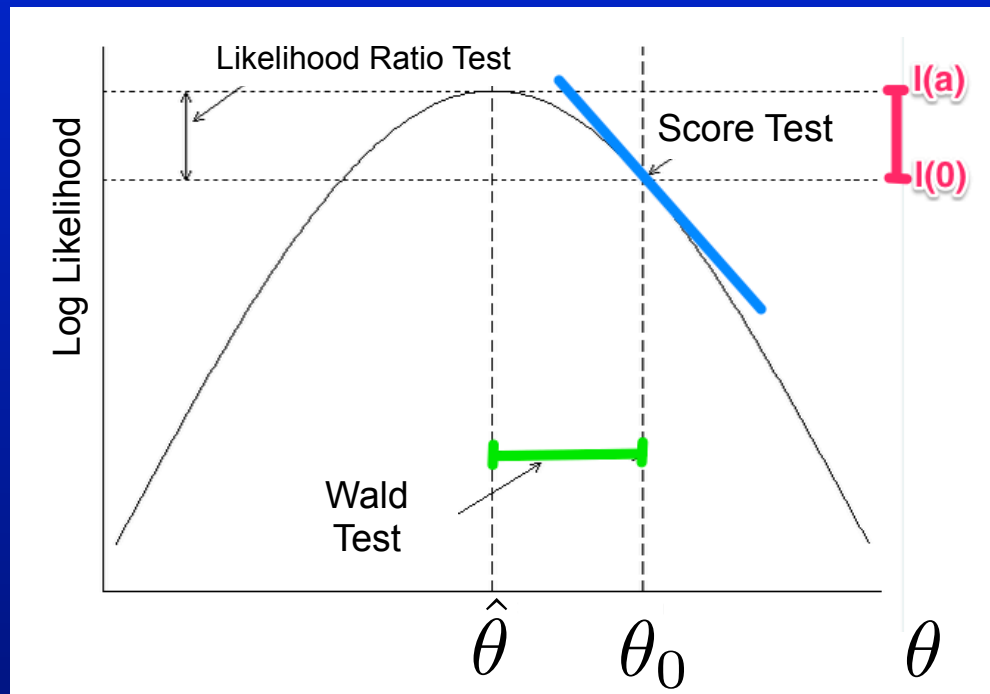
- Optimize model twice, under  $H_0$ , & under  $H_1$
    - Log ratio  $\approx \chi^2$

- Wald

- Estimate under  $H_1$ , compute how many SD's away from  $H_0$  value (e.g. 0)
    - Standardized difference  $\approx$  Normal

- Score

- Derivative of likelihood at null value  $\approx \chi^2$
    - No fitting of  $H_1$  model needed!



# Fast Heritability: Winner Tests

- Score test
  - Based on “0-step” estimate

$$T_S = \frac{1}{2} \left( \frac{\hat{\sigma}_{A,OLS}^2}{\hat{\sigma}_{OLS}^2} \right)^2 \lambda_g' \left( I - \frac{\mathbf{1}\mathbf{1}'}{N} \right) \lambda_g$$

- where  $\sigma_{OLS}^2$  is total variance

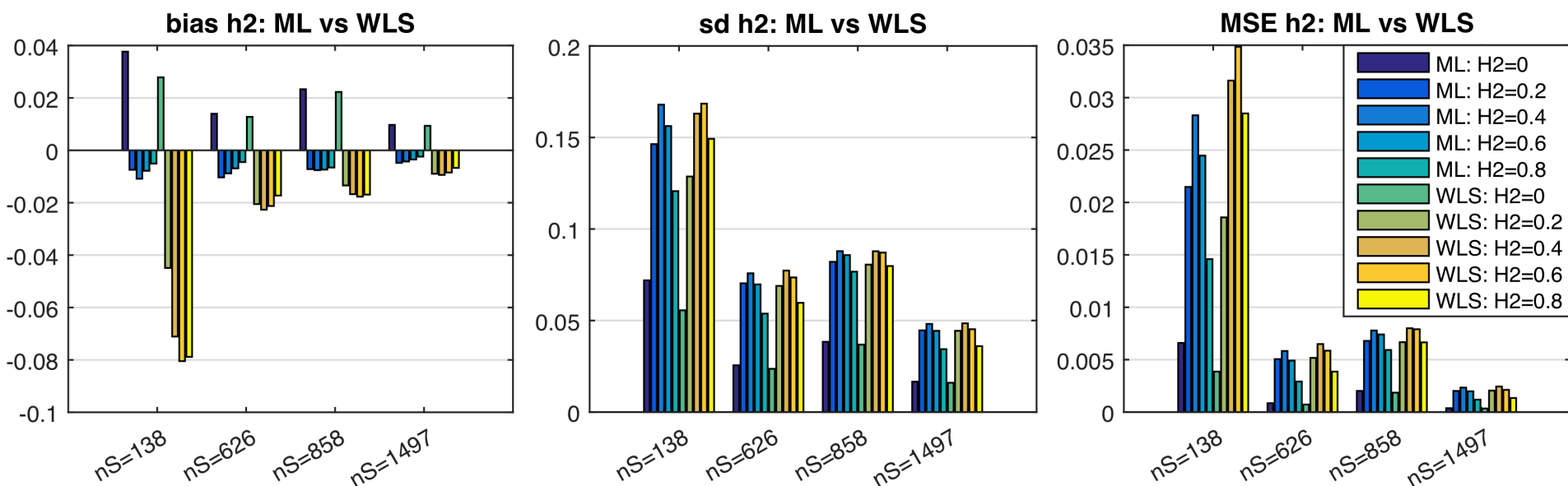
- Wald test
  - Based on “1-step” estimate

$$T_{W,WLS} = \frac{1}{2} (\hat{\sigma}_{A,WLS}^2)^2 [C(U' \hat{\Sigma}_{WLS}^{*-2} U)^{-1} C']^{-1}$$

- where  $C = [0 \ 1]$ , a contrast to pull  $\sigma_A^2$  out of  $\theta$

# Fast, accurate approximation

- Has some bias, but, comparable SD
- MSE also very similar



– These fast 1-step/0-step tools make heritability permutation tractable

# Example Permutation Results

- QTIM & GOBS data
- $P_{\text{uncorr}}$  from SOLAR
- $P_{\text{uncorr}}$  from permutation
  - Reduced significance
  - Weaker assumptions
    - Reduced power
    - Could be detecting artifactual result

|                       | <i>QTIM</i> | <i>GOBS</i>      |
|-----------------------|-------------|------------------|
| <b>Relatedness</b>    | Twins       | Pedigree         |
| <b>Race/Ethnicity</b> | Caucasian   | Mexican-American |
| <b>#Subjects</b>      | 292         | 859              |

| <i>Region of Interest</i> | <i>QTIM</i><br>$P_{\text{uncorr}}$<br>(LRT) | <i>QTIM</i><br>$P_{\text{uncorr}}$<br>(perm) | <i>GOBS</i><br>$P_{\text{uncorr}}$<br>(LRT) | <i>GOBS</i><br>$P_{\text{uncorr}}$<br>(perm) |
|---------------------------|---------------------------------------------|----------------------------------------------|---------------------------------------------|----------------------------------------------|
| <i>Skeleton</i>           | $1.30 \times 10^{-6}$                       | 0.359                                        | $6.48 \times 10^{-9}$                       | 0.137                                        |
| <i>GCC</i>                | $2.00 \times 10^{-7}$                       | 0.194                                        | $0.64 \times 10^{-9}$                       | <b>0.004</b>                                 |
| <i>BCC</i>                | $1.68 \times 10^{-8}$                       | 0.616                                        | $0.10 \times 10^{-10}$                      | <b>0.001</b>                                 |
| <i>SCC</i>                | $5.00 \times 10^{-5}$                       | 1.000                                        | $0.10 \times 10^{-10}$                      | <b>0.028</b>                                 |
| <i>FX</i>                 | $4.16 \times 10^{-8}$                       | 0.519                                        | $8.30 \times 10^{-4}$                       | <b>0.005</b>                                 |
| <i>CGC</i>                | $2.03 \times 10^{-9}$                       | 0.053                                        | $2.52 \times 10^{-8}$                       | <b>0.001</b>                                 |
| <i>CR</i>                 | $0.01 \times 10^{-9}$                       | <b>0.003</b>                                 | $0.10 \times 10^{-10}$                      | <b>0.001</b>                                 |
| <i>EC</i>                 | $0.01 \times 10^{-9}$                       | <b>0.046</b>                                 | $0.44 \times 10^{-9}$                       | <b>0.001</b>                                 |
| <i>IC</i>                 | $2.00 \times 10^{-7}$                       | 0.121                                        | $0.80 \times 10^{-9}$                       | <b>0.004</b>                                 |
| <i>IFO</i>                | $5.84 \times 10^{-9}$                       | <b>0.001</b>                                 | $1.60 \times 10^{-6}$                       | <b>0.001</b>                                 |
| <i>PTR</i>                | $0.86 \times 10^{-9}$                       | 0.415                                        | $3.26 \times 10^{-9}$                       | <b>0.001</b>                                 |
| <i>SFO</i>                | $0.76 \times 10^{-9}$                       | 0.525                                        | $1.00 \times 10^{-7}$                       | <b>0.001</b>                                 |
| <i>SLF</i>                | $0.10 \times 10^{-10}$                      | <b>0.007</b>                                 | $0.10 \times 10^{-10}$                      | <b>0.001</b>                                 |
| <i>SS</i>                 | $0.10 \times 10^{-10}$                      | <b>0.001</b>                                 | $1.07 \times 10^{-9}$                       | <b>0.001</b>                                 |
| <i>CST</i>                | $1.46 \times 10^{-1}$                       | 0.974                                        | $2.00 \times 10^{-6}$                       | 0.942                                        |

# Other cases:

## ACE Model for Twins

- ACE twin model: 3 variance terms:

$$\text{Cov}(\text{DZ}_1, \text{DZ}_2) = \sigma_a^2/2 + \sigma_c^2 + \sigma_e^2$$

$$\text{Cov}(\text{MZ}_1, \text{MZ}_2) = \sigma_a^2 + \sigma_c^2 + \sigma_e^2$$

- Null hypothesis of no heritability ( $h^2=0$ )

$$H_0 : \sigma_a^2 = 0$$

- Under  $H_0$ , not all subject independent!
  - Common environment correlates twin pairs
- Randomly permute MZ/DZ labels!
  - Twins stay together, but who is MZ/DZ changes

# Accelerated GWAS with Related Individuals

- GLM

$$Y \sim \mathcal{N}(X\beta, \Omega)$$

- $X$  design matrix – Now with nuisance (age, etc) and genotype/allelic dose!

- Testing now on slope,  $\beta_{\text{SNP}}$

- $H_0: \beta_{\text{SNP}} = 0$
- Must allow  $h^2 > 0$   
... Not independent!

- How to permute?

- Permute while respecting the kinship!

$X =$

| $\mu$ | Age | Sex | SNP |
|-------|-----|-----|-----|
| 1     | 44  | 0   | 0   |
| 1     | 39  | 1   | 2   |
| 1     | 21  | 0   | 1   |
| 1     | 36  | 1   | 2   |
| 1     | 42  | 0   | 1   |

# EVD Association Permutation (1)

- Untransformed model

$$Y = X\beta + \epsilon \quad \text{Var}(\epsilon) = 2\Phi\sigma_g^2 + I\sigma_e^2$$

- Diagonalize the SNP data just like covariates
  - S eigenvectors of kinship matrix  $\Phi$
  - S then makes data independent... but with het. variance

$$S'Y = S'X\beta + S'\epsilon \quad \text{Var}(S'\epsilon) = D_g\sigma_g^2 + I\sigma_e^2$$

– Where  $D_g$  is diagonal, with eigenvalues of  $2\Phi$

- With estimate of  $\sigma_g^2$  &  $\sigma_e^2$ , fully whiten

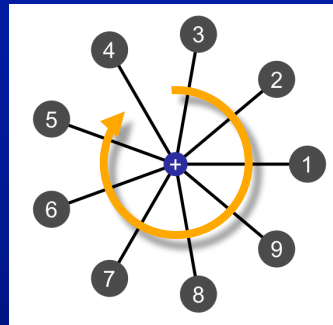
$$S^{*'}Y = S^{*'}X\beta + S^{*'}\epsilon \quad \text{Var}(S^{*'}\epsilon) \propto I$$

- Where  $S^* = S(D_g\sigma_g^2 + I\sigma_e^2)^{-1/2}$

# More Generally: Tree Permutation

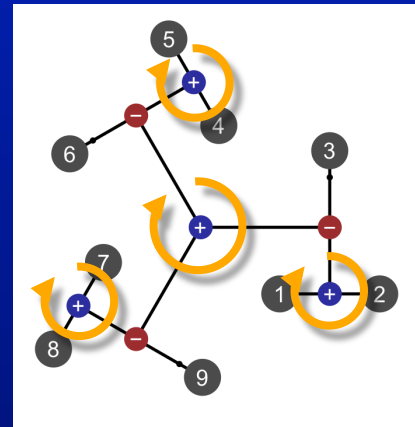
- Define “graph” of dependency
- Find permutations of arms of the graph that leave structure the same
- Human Connectome Project extended twin design
  - Twins plus non-twin (NT) siblings
  - Similar family sizes (mostly 3-4)
  - Graphs show permutation scheme that respects kinship

9 Unrelated subjects



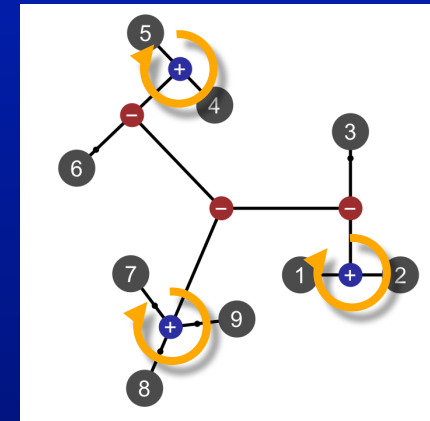
Any permutation of the 9 allowed.

3 families, each:  
1 MZ pair,  
1 NT sibling



Twin pairs can be flipped.  
Entire families can be exchanged.

3 families:  
1 MZ pair + NT  
1 DZ pair + NT  
3 NT



Twin pairs can be flipped.  
(Families not exchangeable.)

# Association Permutation Inference with General Pedigrees

- Tree approach limited
  - Great with highly regular pedigrees
  - But few permutations with arbitrary pedigrees
- Need more general approach: EVD Whitening!
  1. Estimate  $\beta$ ,  $\sigma^2_g$ ,  $\sigma^2_e$
  2. Use  $\sigma^2_g$ ,  $\sigma^2_e$  to decorrelate data (i.e. residuals)
  3. Permute decorrelate residuals
  4. Recorrelate, then re-estimate  $\beta$  (including  $\beta_{\text{SNP}}$ )

# EVD Association Permutation (2)

- The fully whitened model can then be permuted!

# Conclusions

- Different ways to make inference on images
  - Voxel, cluster, TFCE,...
- Must control false positive risk
  - FWE or FDR
- Permutation inference possible with kinship data
  - Heritability inference (relatively easy – just permute)
  - Association inference (trickier)
  - Accelerated 0-step/1-step approximations key for practical permutation inference

# References

- TE Nichols, S Hayasaka and TD Wager.  
Controlling the Familywise Error Rate in Functional  
Neuroimaging: A Comparative Review.  
Statistical Methods in Medical Research, *in press*.
- TE Nichols and AP Holmes.  
Nonparametric Permutation Tests for Functional  
Neuroimaging: A Primer with Examples.  
*Human Brain Mapping*, 15:1-25, 2002.
- <http://www.sph.umich.edu/~nichols>