Nonparametric Thresholding Methods
(FWE inference w/ SnPM)

Thomas Nichols, Ph.D.
Assistant Professor
Department of Biostatistics
University of Michigan

http://www.sph.umich.edu/~nichols

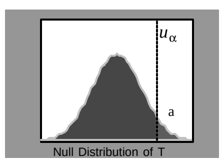
Introduction to fMRI
August 17, 2005

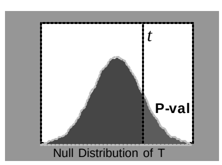
Overview

- Nonparametric Permutation test
- Multiple Comparisons Problem
  - Which of my 100,000 voxels are "active"?
- SnPM
  - Permutation test to find threshold
  - Control chance of any false positives (FWER)

Hypothesis Testing

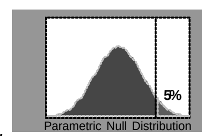
- Null Hypothesis H_0
- Test statistic T
  - t observed realization of T
- α level
  - Acceptable false positive rate
  - $P(T>u_\alpha \mid H_0) = \alpha$
- P-value
  - Assessment of t assuming H_0
  - $P(T > t \mid H_0)$
    - Prob. of obtaining stat. as large or larger in a new experiment
  - P(Data|Null) <u>*not*</u> P(Null|Data)



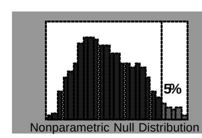


Nonparametric Inference

- Parametric methods
  - Assume distribution of statistic under null hypothesis
  - Needed to find P-values, u_α



- 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_o
  - 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

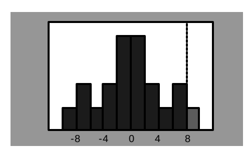
Permutation Test Toy Example

- 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	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



Nonparametric Inference: Permutation Test

- Assumptions
  - Null Hypothesis Exchangeability
- Method
  - Compute statistic t
  - Resample data (without replacement), compute t^*
  - $\{t^*\}$ permutation distribution of test statistic
  - P-value = $\#\{ t^* > t \} \;/\; \#\{ t^* \}$
- Theory
  - Given data and H_0, each t^* has equal probability
  - Still can assume data randomly drawn from population

Permutation Test Strengths

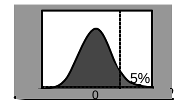
- Requires only assumption of exchangeability
  - Under Ho, distribution unperturbed by permutation
  - Allows us to build permutation distribution
- Subjects are exchangeable
  - Under Ho, each subject's A/B labels can be flipped
- fMRI scans not exchangeable under Ho
  - Due to temporal autocorrelation

Permutation Test Limitations

- Computational Intensity
  - Analysis repeated for each relabeling
  - Not so bad on modern hardware
    - No analysis discussed below took more than 3 hours
- Implementation Generality
  - Each experimental design type needs unique code to generate permutations
    - Not so bad for population inference with t-tests

13

Multiple Comparisons Problem (MCP)



- Standard Hypothesis Test
  - Controls Type I error of each test, at say 5%
  - But what if I have 100,000 voxels?
    - 5,000 false positives on average!
- Must control false positive rate
  - What false positive rate?
  - Chance of 1 or more Type I errors
  - Chance of 50 or more?

14

MCP 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)

15

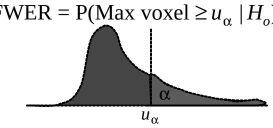
FWER MCP Solutions: Controlling FWER w/ Max

- FWER & distribution of maximum

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

- 100(1-α)%ile of max dist^n controls FWER

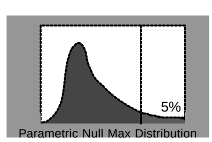
$$\text{FWER} = \text{P(Max voxel} \geq u_\alpha \mid H_o) \leq a$$



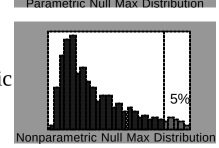
16

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!



Parametric Null Max Distribution



Nonparametric Null Max Distribution

17

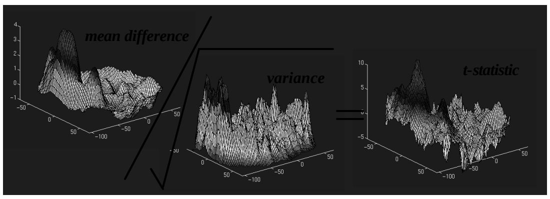
Permutation Test Other Statistics

- Collect max distribution
  - To find threshold that controls FWER
- Consider smoothed variance *t* statistic
  - To regularize low-df variance estimate

18

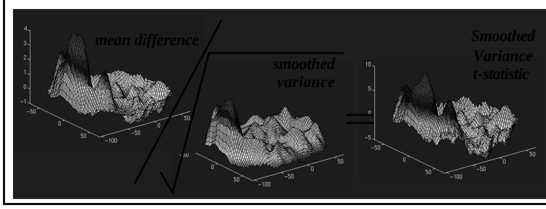
Permutation Test Smoothed Variance *t*

- Collect max distribution
  - To find threshold that controls FWER
- Consider smoothed variance *t* statistic



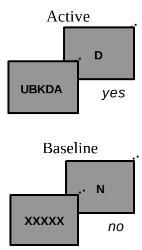
Permutation Test Smoothed Variance *t*

- Collect max distribution
  - To find threshold that controls FWER
- Consider smoothed variance *t* statistic



Permutation Test Example

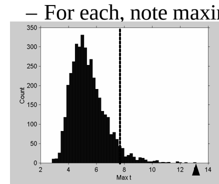
- fMRI Study of Working Memory
  - 12 subjects, block design Marshuetz et al (2000)
  - Item Recognition
    - Active: View five letters, 2s pause, view probe letter, respond
    - Baseline: View XXXXX, 2s pause, view Y or N, respond
- Second Level RFX
  - Difference image, A-B constructed for each subject
  - One sample, smoothed variance *t* test



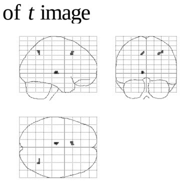
21

Permutation Test Example

- Permute!
  - 2^{12} = 4,096 ways to flip 12 A/B labels
  - For each, note maximum of *t* image



Permutation Distribution Maximum *t*



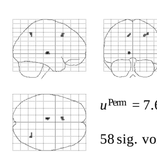
Maximum Intensity Projection Thresholded *t*

22

Permutation Test Example

- Compare with Bonferroni
  - $\alpha = 0.05/110{,}776$
- Compare with parametric RFT
  - 110,776 2×2×2mm voxels
  - 5.1×5.8×6.9mm FWHM smoothness
  - 462.9 RESELs

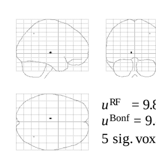
23



$u^{Perm} = 7.67$

58 sig. vox.

t_{11} Statistic, Nonparametric Threshold

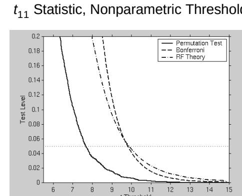


$u^{RF} = 9.87$

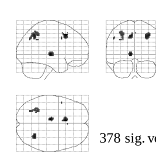
$u^{Bonf} = 9.80$

5 sig. vox.

t_{11} Statistic, RF & Bonf. Threshold



Test Level vs. t_{11} Threshold



378 sig. vox.

Smoothed Variance *t* Statistic, Nonparametric Threshold

24

Does this Generalize? RFT vs Bonf. vs Perm.

RFT vs Bonf. vs Perm.

Monte Carlo Evaluations

- What's going wrong?
  - Normality assumptions?
  - Smoothness assumptions?
- Use Monte Carlo Simulations
  - Normality strictly true
  - Compare over range of smoothness, df
- Previous work
  - Gaussian (Z) image results well-validated
  - t image results hardly validated at all!

27

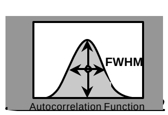
Monte Carlo Evaluations Challenges

- Accurately simulating *t* images
  - Cannot directly simulate smooth *t* images
  - Need to simulate *n* smooth Gaussian images
    (*n* = degrees of freedom)
- Accounting for all sources of variability
  - Most M.C. evaluations use known smoothness
  - Smoothness not known
  - We estimated it residual images

28

Monte Carlo Evaluations

- Simulated One Sample T test
  - 32x32x32 Images (32767 voxels)
  - Smoothness: 0, 1.5, 3, 6,12 FWHM
  - Degrees of Freedom: 9, 19, 29
  - Realizations: 3000



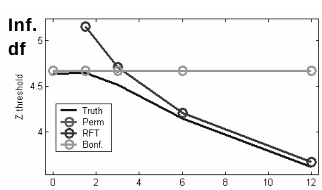
- Permutation
  - 100 relabelings
  - Threshold: 95%ile of permutation dist^n of maximum
- Random Field
  - Threshold: { u : E(χ_u | H_o) = 0.05 }
- Also Gaussian

29

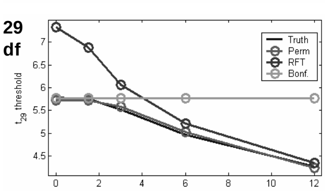
Familywise Error Thresholds

- RFT valid but conservative
- Gaussian not so bad (FWHM >3)
- t_{29} somewhat worse

Inf. df



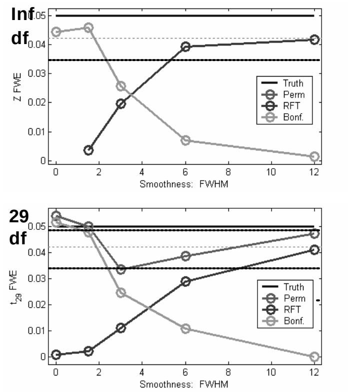
29 df



more

Familywise Rejection Rates

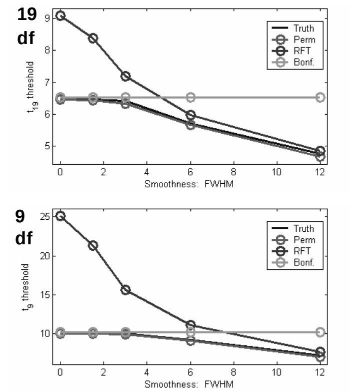
- Need > 6 voxel FWHM



more

Familywise Error Thresholds

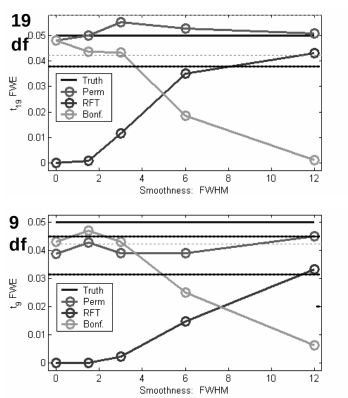
- RF & Perm adapt to smoothness
- Perm & Truth close
- Bonferroni close to truth for low smoothness



more

Familywise Rejection Rates

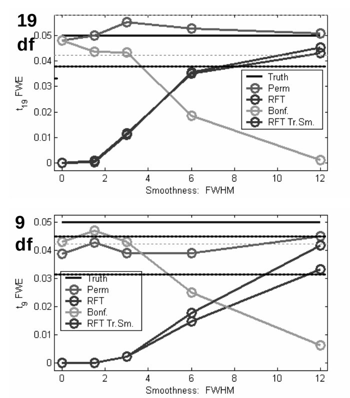
- Bonf good on low df, smoothness
- Bonf bad for high smoothness
- RF only good for high df, high smoothness
- Perm exact



more

Familywise Rejection Rates

- Smoothness estimation is not (sole) problem



cont

Performance Summary

- Bonferroni
  - Not adaptive to smoothness
  - Not so conservative for low smoothness
- Random Field
  - Adaptive
  - Conservative for low smoothness & df
- Permutation
  - Adaptive (Exact)

35

Understanding Performance Differences

- RFT Troubles
  - Multivariate Normality assumption
    - True by simulation
  - Smoothness estimation
    - Not much impact
  - Smoothness
    - You need lots, more at low df
  - High threshold assumption
    - Doesn't improve for α_0 less than 0.05 (not shown)

36

HighThr

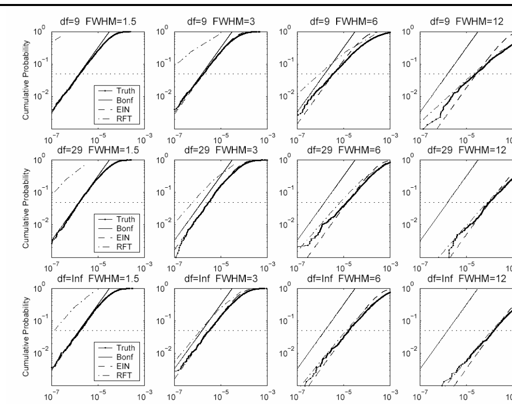
Conclusions

- *t* random field results conservative for
  - Low df & smoothness
  - 9 df & ≤12 voxel FWHM; 19 df & < 10 voxel FWHM
- Bonferroni surprisingly satisfactory for low smoothness
- Nonparametric methods perform well overall
- More data and simulations needed
  - Need guidelines as to when RF is useful
  - Better understand what assumption/approximation fails

References

- 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

Data ThrRslt **MC ThrRslt** **MC P Rslt** **EstSmCf**



Back