Post-hoc power estimation for topological inference in fMRI

Joke Durnez[a], Beatrijs Moerkerke[a] and Thomas Nichols[b]

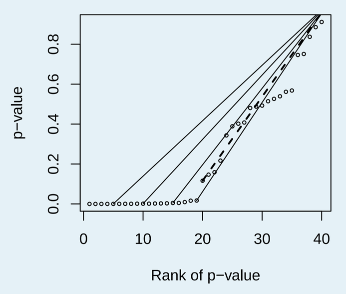
[a] Department of Data Analysis, Ghent University, Belgium

[b] Department of Statistics & Warwick Manufacturing Group, University of Warwick, United Kingdom

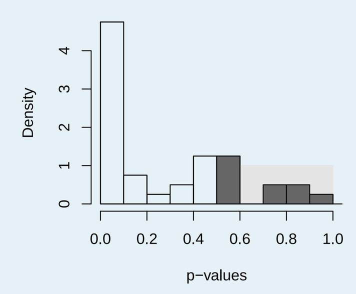
Introduction

- Topological inference on clusters or peaks is increasingly being used in neuroimaging.
- Uncorrected topological p-values can be submitted to any multiple testing procedure to control an excess of type I errors.
- More stringent threshold $\Rightarrow$ drop in power.
- To meet an acceptable level of power: compute power (and sample size) *a priori*; however, these techniques depend on many unknown parameters.
- Zehetmayer et al. (2010) study post-hoc power in the context of statistical genetics.
- **Our goal: estimate the post-hoc power of an fMRI study post-hoc using estimates for the proportion of non-activated features π_0.**
- π_0 can be estimated using the distribution uncorrected p-values.
- Topological p-values can be found using RFT and permutations, both for clusters and for peaks.
- We use permutation p-values for clusters, and RFT p-values for peaks.

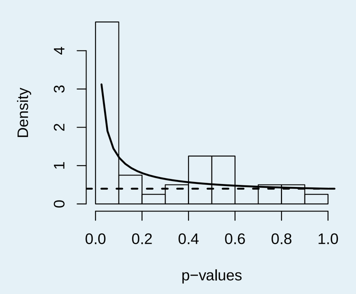
Estimating π_0, the proportion of null tests



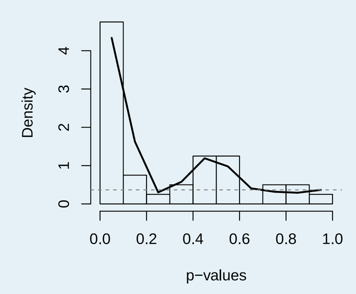
Benjamini,2000 - BH: After sorting the p-values, the slope S_i is calculated between the ith p-value and the point $(m+1,1)$ for ascending i (solid lines). The first i for which $S_i < S_{i-1}$ is considered the first p-value from the alternative hypothesis (dashed line). Subsequently $\hat{m}_0$ is estimated as $\min(m,(1/S-i+1))$.



Storey, 2003 - ST; Storey, 2001 - S : For a certain λ (here: 0.5), π_0 is estimated as the density of p-values greater than λ (dark grey) divided by the expected density under H_0, $1-\lambda$ (light grey). Storey, 2003 and Storey, 2001 differ in the way they optimize the choice of λ.



Pounds, 2003 - PM: The probability density function is assumed to be a mixture of a uniform distribution (dashed line) and a beta-distribution (solid line), with weight π_0 to the uniform distribution. Through maximal likelihood estimation, the optimal weights and parameters for the beta-distribution are estimated.



Pounds, 2004 - PC: The probability density function $\hat{f}(p_i)$ is estimated by applying a loess smoother through the histogram of p-values (solid line). π_0 is then estimated as $\hat{f}(\max(p_i))$.

Estimating post-hoc power

	Declared active	Declared inactive	Total
Truly non-active	$F = \alpha m_0$	$m_0 - \alpha m_0$	m_0
Truly active	$T = S - \alpha m_0$	$m_1 - (S - \alpha m_0)$	m_1
Total	S	$m - S$	m

- "Power" True Positive Rate: $TPR = E(T)/m_1 \leftrightarrow FPR = E(F)/m_0$
- False Nondiscovery Rate: $FNR = E[(m_1 - T)/(m - S)]$ if $S \neq m$ and 0 otherwise $\leftrightarrow FDR = E(F)/S$

Simulations

- fMRI data ($40 \times 40 \times 40 \times 400$) are simulated using `neuRosim` with spatial ($\sigma = 2.5$) and temporal ($\rho = 0.2$) noise.
- 5 regions of $7 \times 7 \times 7$ voxels are related to the blocked design (20 blocks) with 400 timepoints.
- The signal-to-noise ratio: 0.015.
- A GLM is fit to the data using FSL.
- Excursion threshold: $P(T_{398} > u) = 0.01$
- Peak and cluster p-values are calculated.
- 500 simulations have been performed.
- Simulations have been repeated under different conditions for smoothness, excursion threshold, number of activated regions and signal-to-noise ratio's with equal conclusions.

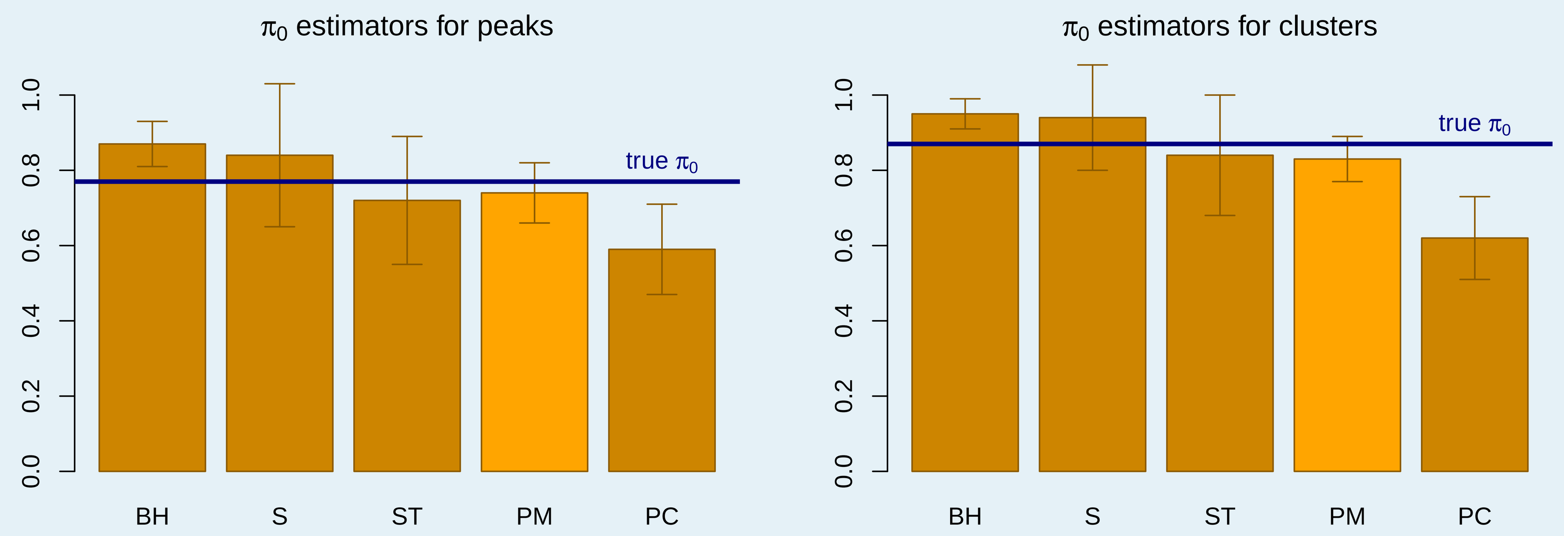
References and acknowledgements

Benjamini and Hochberg (2000). *Journal of the royal statistical society. Series B, statistical methodology*, 57(1), 289-300.
Storey (2001). *Journal of the royal statistical society. Series B, statistical methodology*, 64(30), 479-498.
Storey and Tibshirani (2003). *Proceedings of the national academy of sciences of the united states of America*, 100(16), 9440-9445.
Pounds and Morris (2003). *Bioinformatics*, 19(10), 1236-1242.
Pounds and Cheng (2004). *Bioinformatics*, 20(11), 1735-1745.
Henson, Shallice, Gorno-Tempini and Dolan (2002). *Cerebral cortex*, 12(2), 178-186.
Zehetmayer and Posch (2010). *Bioinformatics*, 26(8), 1050-1056.

This work was carried out using the STEVIN Supercomputer Infrastructure at Ghent University, funded by Ghent University, the Flemish Supercomputer Center (VSC), the Hercules Foundation and the Flemish Government â department EWI.

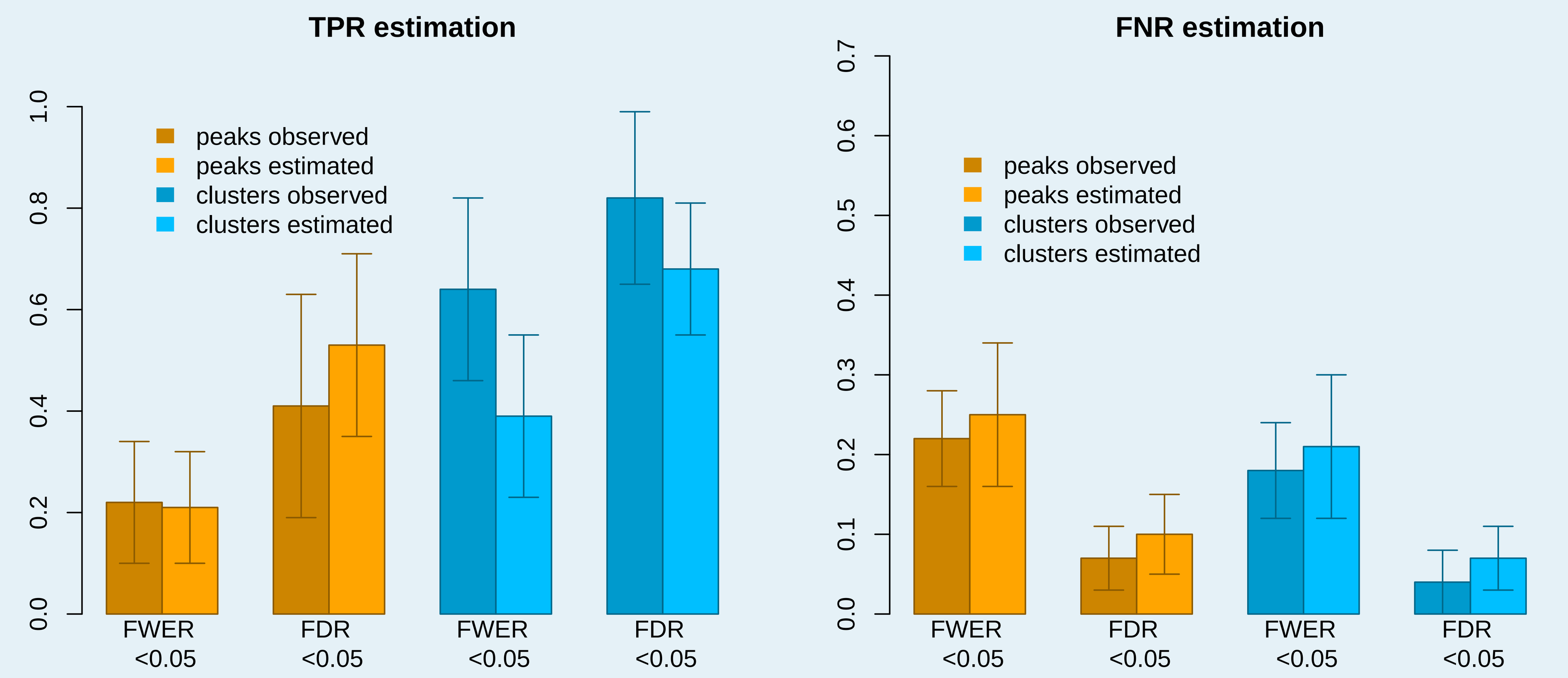
Evaluation π_0 estimators on simulated data

Average estimates of $\widehat{\pi}_0$ for the different estimators with their standard deviation.



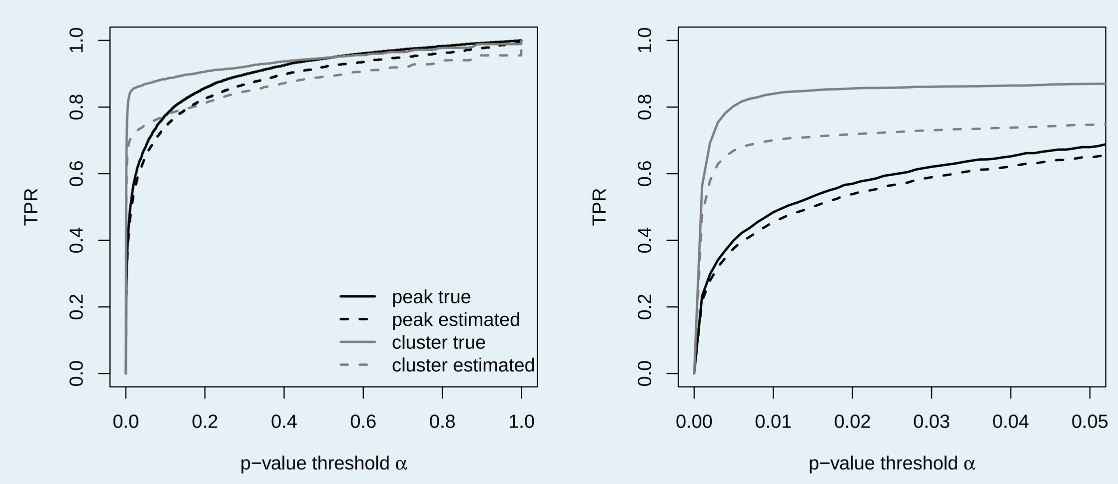
- PM produces best overall performance.

Evaluation of post-hoc power estimation on simulated data



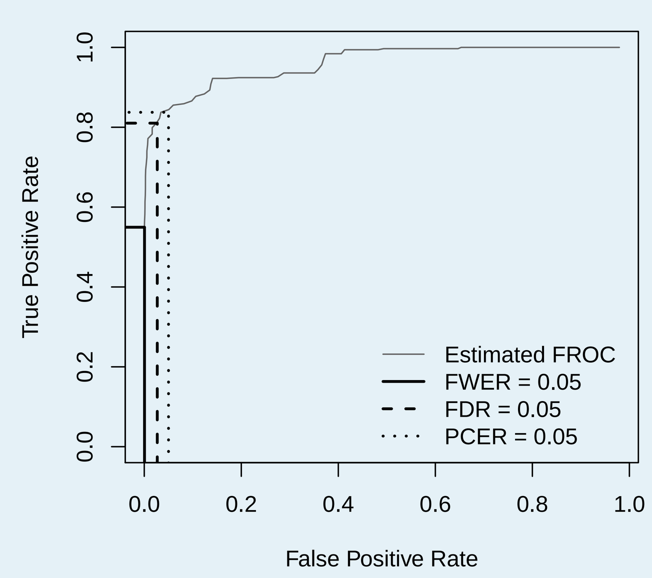
Estimating the free-receiver operator curve

FROC: TPR is plotted against the FPR for a range of p-value thresholds



Real data example

- study by Henson et al., 2002: single-subject event-related design. Average effect of presenting faces.
- 97 peaks discovered
- FWER control $\alpha = 0.05$
  - 35 significant peaks
  - $\widehat{TPR} = 55\%$
  - $\widehat{FNR} = 44\%$
- FDR control $\alpha = 0.05$
  - 52 significant peaks
  - $\widehat{TPR} = 81\%$
  - $\widehat{FNR} = 26\%$



Discussion and conclusion

- Pounds and Morris (2003) best estimator under different fMRI conditions
- Post-hoc power calculations: For a univariate tests ill-advised (e.g. power always greater than 50% for significant test). For multiple testing, provides useful estimate of proportion of true positives.
- Procedure can be used on any collection of uncorrected p-values: voxel-based morphometry, diffusion tensor imaging,...
- Henson study: very low power using FWER control $\Rightarrow$ **call for better balance between sensitivity and specificity.**

Joke.Durnez@UGent.be