pARI package: valid double-dipping via permutation-based All Resolutions Inference

Angela Andreella[1]   Livio Finos[2]   Jelle Goeman[3]

[1]Department of Statistical Sciences, University of Padua
[2]Department of Developmental Psychology and Socialisation, University of Padua
[3]Biomedical Data Sciences, Leiden University Medical Center

useR! - 6 July 2021

fMRI measures brain activation as changes in blood flow (BOLD) under a sequence of stimuli.

Cluster-wise method:

- Analyze set of contiguous voxels (S);

- H_S rejected means that S contains at least one active voxel
  $\rightarrow$ We don't know which ones and how many!!

- **Spatial specificity paradox**;

- **Double-dipping**.

Solution: All-resolutions inference (ARI) [1] → Inference on the number of truly active voxels.

- **pARI** is the Permutation-based version of **ARI**. The permutation structure permits to account for the correlation structure between tests unlike **ARI**;

- Both methods are based on the **closed testing procedure** for controlling the familywise error rate.

---

[1]Resenblatt, J. et al. (2018).

Every time that we want to infer inside a **data-driven (and not) cluster** (features set), we can use **ARI/pARI**:

- **Cluster fMRI data analysis** [2];

- **Gene expression cluster analysis** [3];

- **Cluster EEG data analysis** [4].

---

[2]Woo, C. et al. (2014).
[3]Berge, K. et al. (2017).
[4]Maris, E. et al. (2007).

```
devtools::install_github(angeella/pARI)
library(pARI)
```

- fMRI framework:

```
pARIbrain(copes, thr, mask, alpha, ...)
```

where `copes` is a list of **contrast** parameter estimates involving brain activation differences for each subject in NIfTI format.

- General framework:

```
pARI(data, ix, alpha, test.type, ...)
```

where `ix` is the **features set** of interest. It can be a vector of indices or a vector with length equals the number of features where different values indicate the different sets.

We analyzed the **Auditory data** collected by Pernet et al. (2015), i.e, people listening vocal and non-vocal sounds.

Group analysis on 140 subjects of the Vocal > Non-vocal **contrast** by the one sample t-test flipping the sign of 140 voxel-wise contrasts maps.

First, let download the data from the fMRIdata package:

```
devtools::install_github(angeella/fMRIdata)
library(fMRIdata)
data(Auditory_clusterTH3_2)
data(Auditory_copes)
data(Auditory_mask)
```

```
pARIbrain(copes = Auditory_copes,
          cluster = Auditory_clusterTH3_2,
          mask = Auditory_mask,
          alpha = 0.05)
```

Cluster S	Threshold t	Size $|S|$	% active $\bar{\pi}(S)$		P-Values p_{FWER}
			pARI	ARI	
Right STG/PT HG/IFG/T	3.2	11683	92.36%	84.98%	< 0.0001
Right STG/PT HG/IFG/T	4	8875	99.54%	98.5%	–
Right IFG	4	422	91.47%	83.18%	–
Right T	4	292	85.96%	64.04%	–
Right T	4	15	13.33%	0%	–

map_TDP: Create **true discovery proportion map** in nifti format.

Thanks to the amazing group that worked on the paper
Permutation-based true discovery proportions for fMRI cluster analysis! (in arXiv)

- Livio Finos   University of Padua;
- Jelle Goeman   Leiden University Medical Centre;
- Jesse Hemerik   Wageningen University;
- Wouter Weeda   Leiden University;

and thanks for your attention!

 https://github.com/angeella/useR_2021

 @aangeella

 angela.andreella@stat.unipd.it