Donders Institute
for Brain, Cognition and Behaviour

Statistics
Beyond basic randomization tests

Robert Oostenveld

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen, The Netherlands

NatMEG, Karolinska Institutet, Stockholm

Radboud University Nijmegen

Talk outline

- Recap non-parametric randomization test
- Which hypothesis are we testing?
- Randomization approach
- Clustering procedure
- Organization of data and code in FT
- Implementing your own statistics
- Suggested further reading

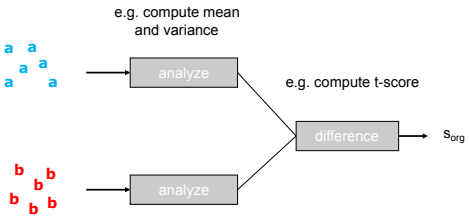
Randomization test: *general principle*

- Independent variable: condition
- Dependent variable: data

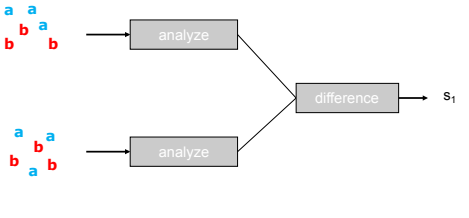
H0: the data is independent from the condition in which it was observed. *The data in the two conditions is **not** different*.

H1: the data depends on the condition in which it was observed. *The data in the two conditions is different*.

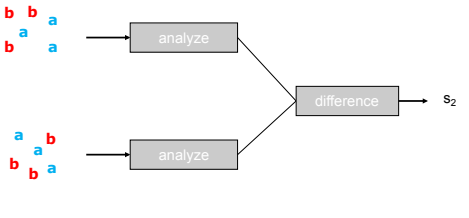
Randomization approach



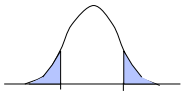
Randomization approach



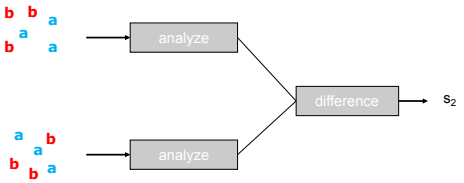
Randomization approach



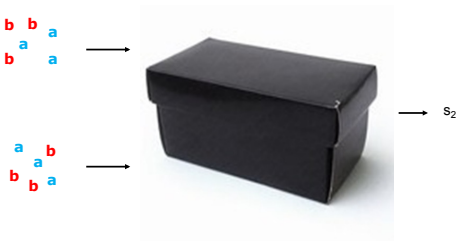
Distribution of “s” can take any shape



Randomization approach



Randomization approach

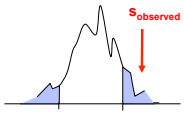


Non-parametric statistics

- Randomization of independent variable
- Hypothesis is about data, not about the specific parameter
- Randomization distribution of the statistic of interest “s” is approximated using Monte-Carlo approach
- H0 is tested by comparing observed statistic v.s. randomization distribution

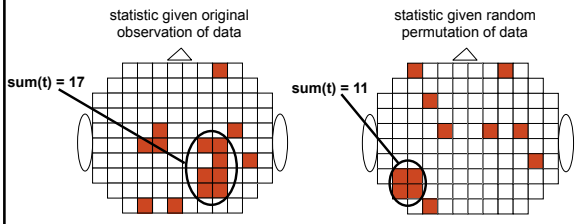
Avoid the MCP

- The statistic “s” can be anything
- Rather than testing everything, only test the most extreme observation *(i.e. the max statistic)*
- Compute the randomization distribution for the most extreme statistic
- Note that often we compute two extrema, one for each tail



Avoid the MCP and increase sensitivity

Accumulate evidence by clustering neighbouring observations (channels, voxels, time, frequency)

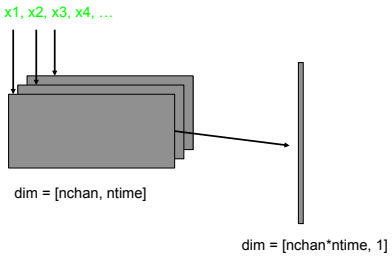


Increase the sensitivity

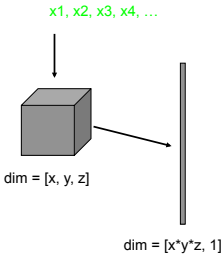
Use a massive univariate statistic that is sensitive to detecting the difference in the data **according to H1**.

- per channel-time-frequency (or voxel) point we compute a statistic, which we threshold, make clusters, and take maximum
- massive univariate can be T-statistic, F-statistic, beta weight from GLM, …

Organizing channel-level data

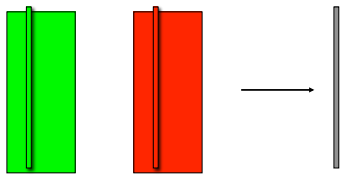


Organizing source-level data



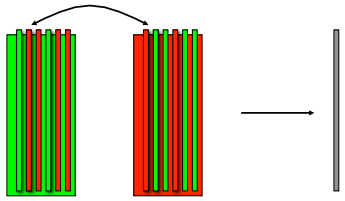
Organizing source-level data

[x1, x2, x3, x4, … x1, x2, x3, x4, …] ⟶ statistic



Compute massive univariate statistic

[x1, x2, x3, x4, … x1, x2, x3, x4, …] ⟶ statistic



FieldTrip implementation

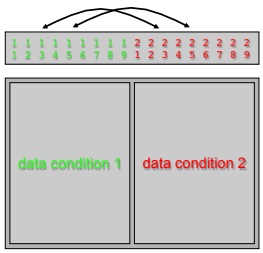
```
cfg  = [];
cfg.method   = 'analytic'
cfg.correctm = 'bonferroni'
ERPstats     = ft_timelockstatistics(cfg, ERP);
```

```
cfg  = [];
cfg.method   = 'analytic'
cfg.correctm = 'fdr'
ERPstats     = ft_timelockstatistics(cfg, ERP);
```

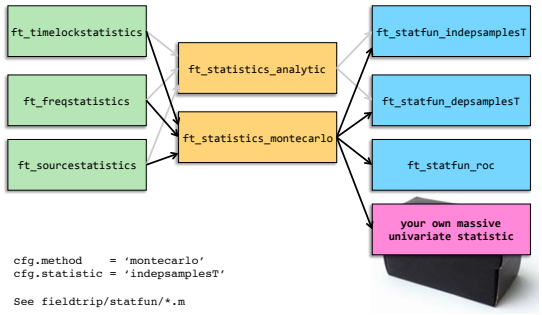
```
cfg  = [];
cfg.method   = 'montecarlo'
cfg.correctm = 'max'
TFRstats     = ft_sourcestatistics(cfg, source);
```

FieldTrip implementation - data

```
independent variable   cfg.ivar = 1 (i.e. row 1)
unit of observation    cfg.uvar = 2 (i.e. row 2)
```



FieldTrip implementation - code




```
cfg.method    = 'montecarlo'
cfg.statistic = 'indepsamplesT'

See fieldtrip/statfun/*.m
```


Summary

- Recap non-parametric randomization test
- Which hypothesis are we testing?
- Randomization approach
- Clustering procedure
- Organization of data and code in FT
- Implementing your own statistics

Suggested further reading

Tutorials

http://fieldtrip.fcdonders.nl/tutorial/eventrelatedstatistics
http://fieldtrip.fcdonders.nl/tutorial/cluster_permutation_timelock
http://fieldtrip.fcdonders.nl/tutorial/cluster_permutation_freq

http://fieldtrip.fcdonders.nl/faq/how_not_to_interpret_results_from_a_cluster-based_permutation_test

Papers

Nonparametric statistical testing of EEG- and MEG-data.
Maris E & Oostenveld R, J Neurosci Methods. 2007

Statistical testing in electrophysiological studies.
Maris E, Psychophysiology, 2011.