

Current Frontiers in Statistical Inference for Neuroimaging Data

Bertrand Thirion Organizer

inria

Palaiseau

France

Symposium

- Brain imaging relies extensively on statistical procedures
- Misuse of statistics is a widespread issue in the field
- A critical view on statistical procedures is always needed
- While basic brain mapping procedures were set up 25 years ago and have progressively improved since then, new approaches have appeared
- These novel approaches are an important opportunity for OHBM participants
- As part of the session, we will have a discussant that will launch the discussion and help framing questions

Objective

- Identify the factors that limit the power of neuroimaging studies
- Discover novel statistics procedures such as the control of false discovery proportion
- Discover the possibility of statistical control of multivariate models and associated caveats.

Target Audience

- Cognitive neuroscientists using statistical procedures and wishing to improve their way of handling inference (Cognitive neuroscience and population analyses).

Presentations

Cluster Failure or Power Failure? Towards a new level of inference for neuroimaging

Neuroimaging research has relied on the ability to map brain activity and networks through cluster-based inference, but the empirical validity of such popular methods is still being characterized. A landmark study in Eklund et al. 2016 showed that popular multiple comparison corrections based on cluster extent suffer from unexpectedly low specificity (i.e., high false positive rate). Yet that study's focus on specificity, while important, is incomplete. The validity of a method depends also on its sensitivity (i.e., true positive rate or power), yet the sensitivity of cluster correction remains poorly understood and converging evidence suggests more sensitive inferential procedures are needed to promote discovery. Here I will discuss recent work assessing sensitivity of gold-standard nonparametric cluster correction in task-based activation (Noble et al. 2018) and functional connectivity (Noble et al. 2020). This work relies on empirically derived measures of sensitivity derived by resampling real data from tasks in the Human Connectome Project and comparing results with those from the full “ground truth” datasets. I will then demonstrate how such pooling data across larger scales can substantially boost sensitivity. I will specifically highlight how a new tool—the Constrained Network-Based Statistic—achieves an optimal balance of sensitivity and spatial specificity relative to other approaches (Noble et al. 2020). Altogether, these results suggest that we have, until now, only measured the tip of the iceberg in much of the literature due to a reliance on limiting the familywise error rate through cluster extent-based inference. There is a need to revise our practices to improve sensitivity, and better accounting for the rich dependence in high dimensional neuroimaging data may be a good start.

Presenter

Stephanie Noble, Yale University, Radiology & Biomedical Imaging, New Haven, CT, United States

Putting It All Together

I feel we often find ourselves in data analysis pipeline “ruts”. We have a set of tools we may have been using for years, but then new methods are introduced and it isn’t clear how they could be incorporated into our existing pipelines to improve our research and what improvements would be expected. I will provide my thoughts about how we can move forward with this new information, how it may change how we formulate hypotheses and view our results.

Presenter

Jeannette Mumford, Pr., University of Wisconsin-Madison Madison, WI
United States

All Resolution Inference

The All Resolution Inference (ARI) framework of (Rosenblatt et al. 2018) allows two main things: (1) Estimate the proportion of active voxels in any voxel-subset in the brain. Contiguous or not. Supra-threshold or not. (2) Statistical guarantees of these estimates no matter how the voxel-subsets have been selected. Particularly, after endless circularity. These two possibilities are exciting for various reasons: Because of the “spatial specificity paradox” by which active clusters may contain only a single active voxel, so estimating the proportion of activation is of great interest. Because cluster selection in practice is not as “clean” as random-field theory would require. The fact that statistical guarantees are valid no matter how voxel-subsets are selected, is the ultimate insurance against dead-salmon brains. Because it turns out that the price paid, in terms of power, for such general guarantees, is not as large as one may intuitively expect.

Presenter

Jonathan Rosenblatt, Ben Gurion University Beer Sheva, Negev, Israel

Decoding with Confidence: Statistical control on decoder maps?

In brain imaging, decoding is widely used to infer relationships between brain and cognition, or to craft brain-imaging biomarkers of pathologies. Yet, standard decoding procedures do not come with statistical guarantees on coefficients, and thus do not give confidence bounds to interpret the pattern maps that they produce. Indeed, in whole-brain decoding settings, the number of explanatory variables is much greater than the number of samples, hence classical statistical inference methodology cannot be applied. Specifically, the standard practice that consists in thresholding decoding maps is not a correct inference procedure. We contribute a new statistical-testing framework for this type of inference. To overcome the statistical inefficiency of voxel-level control, we use a generalization of the Family Wise Error Rate (FWER) to account for a spatial tolerance δ , namely the δ -Family Wise Error Rate (δ -FWER). Then, we present a decoding procedure that can control the δ -FWER: the Ensemble of Clustered Desparsified Lasso (EnCluDL) (Chevalier et al. 2018), a procedure for multivariate statistical inference on high-dimensional structured data. We evaluate the statistical properties of EnCluDL with a thorough empirical study, along with three alternative procedures including decoder map thresholding. We show that EnCluDL exhibits the best recovery properties while ensuring the expected statistical control.

Presenter

Bertrand Thirion, Inria Palaiseau, France