How are cross-validated decoding accuracies distributed across subjects?

Poster No:

3813

Submission Type:

Abstract Submission

Authors:

Joram Soch[1,2], John-Dylan Haynes[1,2,3,4,5,6,7,8]

Institutions:

[1]Bernstein Center for Computational Neuroscience, Berlin, Germany, [2]Berlin Center for Advanced Neuroimaging, Berlin, Germany, [3]Berlin School of Mind and Brain, Berlin, Germany, [4]Clinic for Neurology, Charité – Universitätsmedizin, Berlin, Germany, [5]Department of Psychology, Humboldt University, Berlin, Germany, [6]EXC NeuroCure, Charité – Universitätsmedizin, Berlin, Germany, [7]EXC Science of Intelligence, Technical University, Berlin, Germany, [8]CRC Volition and Cognitive Control, Technical University, Berlin, Germany

Introduction:

In multivariate pattern analysis (MVPA) for functional magnetic resonance imaging (fMRI) data, cross-validated decoding accuracy (cvDA) maps from first-level decoding analyses are typically subjected to second-level t-tests [1] in order to make population inference. This practice has been criticized as making questionable assumptions about cvDAs and effectively performing a fixed-effects analysis instead of a more appropriate random-effects analysis which is usually insinuated by the interpretation of the results [2]. The second-level t-test for cvDAs can also be challenged based on the nature of accuracies which, unlike the variates going into a t-test, (i) are neither in an *infinite range* (but bounded between 0 and 1), (ii) nor are they *normally distributed* (because they are proportions). Here, we estimate Beta distributions [5,6] from cvDAs across subjects to overcome those problems and provide a number of summary statistics to make statements about the observed accuracies.

Methods:

The cross-validated decoding accuracies r_i from subjects $i = 1,\ldots,N$ are assumed to be independent and come from an underlying Beta distribution [5]:

$p(r|\alpha,\beta) = \prod_i \mathrm{Bet}(r_i; \alpha, \beta)$

The parameters α and β can be obtained using maximum likelihood estimation (MLE) for the Dirichlet distribution [6], a generalization of the Beta distribution:

$(\alpha_{est},\beta_{est}) = \mathrm{argmax}_{(\alpha,\beta)} \log p(r|\alpha,\beta)$

From the estimated concentration parameters, one can then calculate (see Figure 1)

a) the *expected frequency*, i.e. the mean of the accuracy distribution

$EF = \alpha / (\alpha+\beta)$,

b) the *likeliest frequency*, i.e. the most likely decoding accuracy
$LF = (\alpha-1) / (\alpha+\beta-2)$,
c) the *exceedance probability* that the decoding accuracy is larger than chance level γ
$EP(\gamma) = Pr(r>\gamma|\alpha,\beta)$
d) and a *confidence interval* for the most likely accuracy, given a confidence level (1-α)
$CI(1-\alpha) = [CI_{low}, ..., LF, ..., CI_{upp}]$.

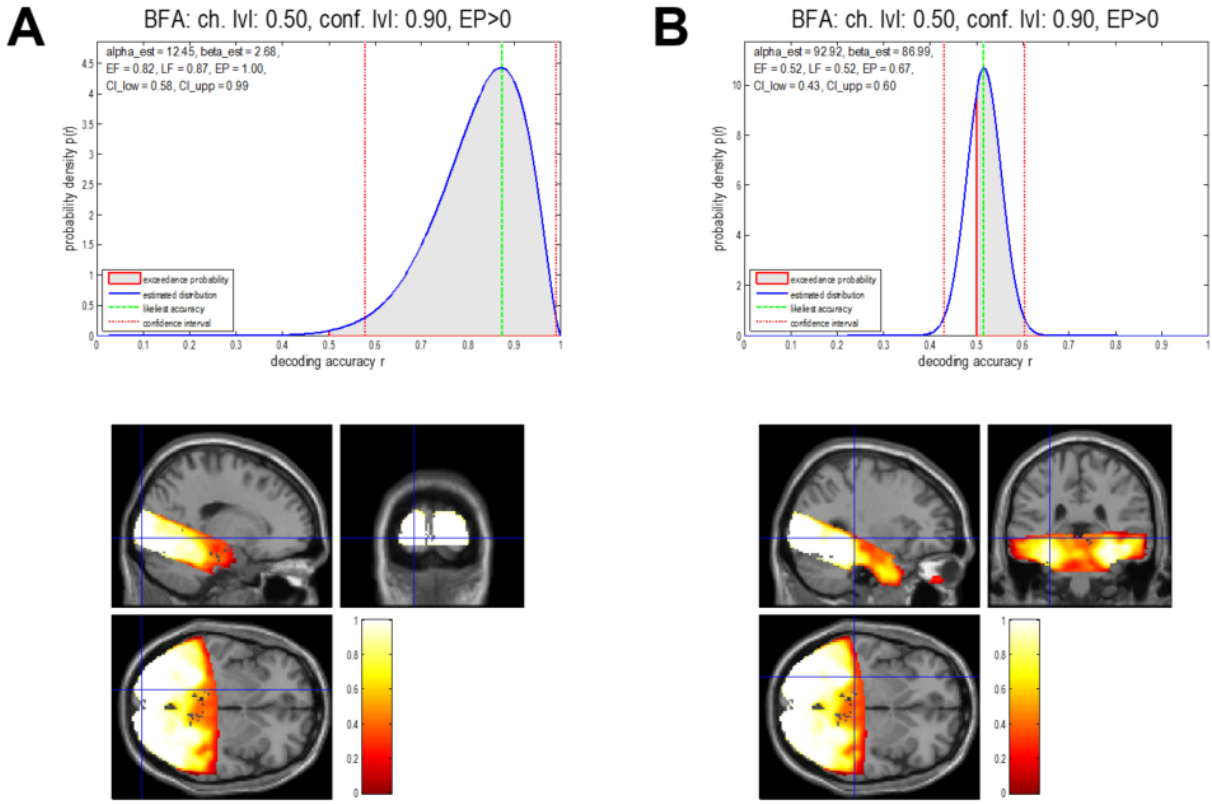

Figure 1. *Beta distributions for decoding accuracies: analysis of unpermuted data.* The lower panels show an unthresholded map of exceedance probability (chance level: 0.5) and the upper panels show estimated distributions and confidence intervals (confidence level: 0.9) for two selected voxels. **(A)** A voxel containing information about object category which was decoded in the present analysis. The distribution is centered on a high accuracy value ($\hat{\alpha}$ much larger than $\hat{\beta}$); EF, LF and EP are close to or equal to 1; the confidence interval is entirely above chance level. **(B)** A voxel with less information about object category. The distribution is much sharper (larger $\hat{\alpha}$ and $\hat{\beta}$); EF, LF and EP are closer to 0.5; and the confidence interval includes chance level.

Results:

We analyzed a set of cvDA maps [4] that comes from an earlier study on object recognition [3] and was previously used in methods development for second-level decoding inference [2]. We show that the above-mentioned summary statistics can give an intuitive understanding of the amount of information contained in a voxel or searchlight when decoding accuracies are significantly above (see Figure 1A) or when they are around chance level (see Figure 1B). Moreover, we observe that estimated concentration parameters across all voxels are compatible with a null distribution when labels were exchanged between classes (see Figure 2B), but not when unpermuted data were used (see Figure 2A).

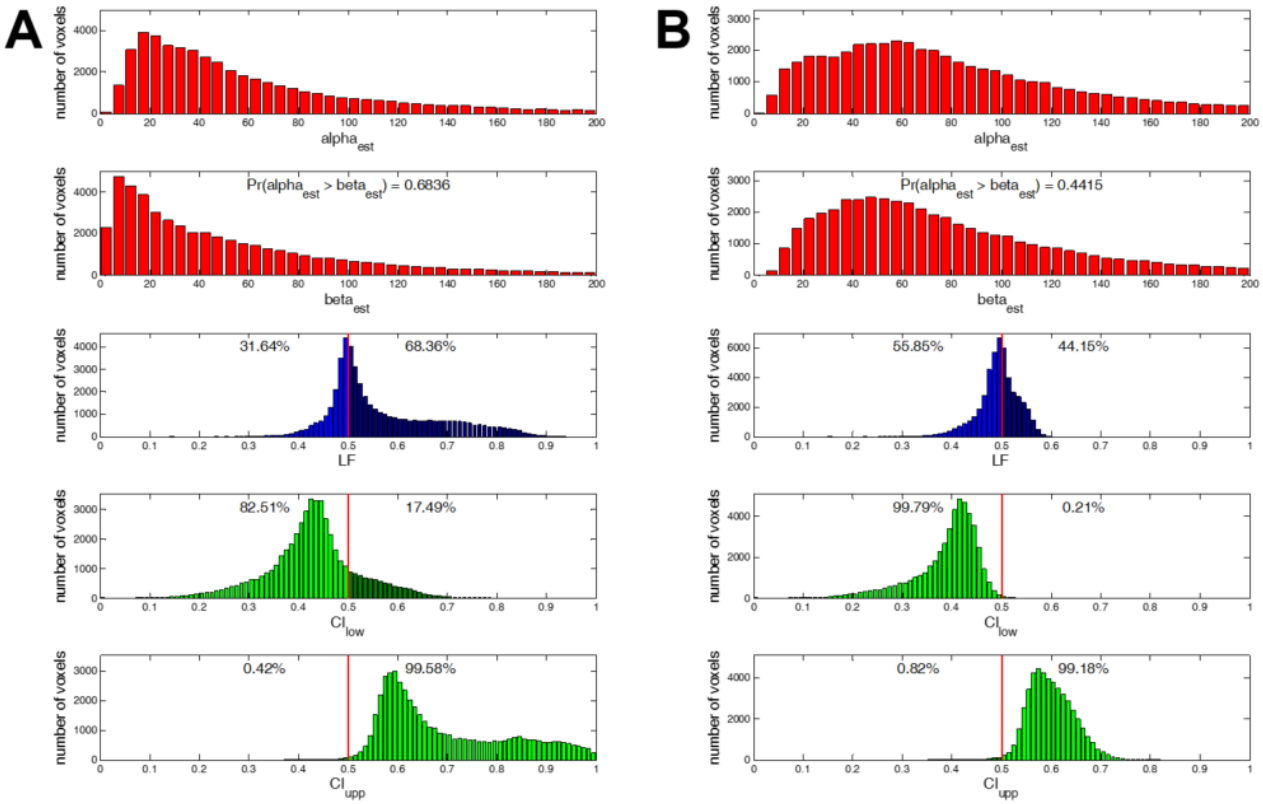

Figure 2. *Beta distributions for decoding accuracies: unpermuted vs. permuted data.* All panels show across-voxel histograms of Beta distributions estimated from decoding accuracies obtained when using **(A)** the original, unpermuted data vs. **(B)** permuted object category labels. The probability that $\hat{\alpha}$ is larger than $\hat{\beta}$ is equivalent to the proportion of voxels in which LF is larger than 0.5; this is much more often the case when using the unpermuted data characterized by an effect of object category in some of the voxels. The proportion of voxels in which CI_{low} is larger than 0.5 is equivalent to the probability that the confidence interval is entirely above chance level; this is almost never the case when using permuted labels which are intended to eradicate any effect.

Conclusions:

We have provided a method for second-level analysis of cvDA maps and validated it using empirical data. Unlike the currently received approach, i.e. the one-sample t-test, this type of analysis (i) acknowledges the domain of accuracies as proportions and (ii) allows the distribution of accuracies to be non-symmetric. A downside of the proposed method is that it, like the t-test, does not take into account that *true* accuracies cannot be below chance level and inference therefore only pertains to *observed* accuracies [2]. In the future, we want to evaluate asymptotic properties of this approach using simulated data and implement multiple comparison correction for the analysis of empirical data.

Modeling and Analysis Methods:

Classification and Predictive Modeling [1]
Methods Development [2]
Univariate Modeling

Neuroinformatics and Data Sharing:

Informatics Other

Perception, Attention and Motor Behavior:

Perception: Visual

Keywords:

Data analysis
Informatics
Machine Learning
Modeling
Multivariate
Perception
Statistical Methods
Univariate
Vision

[1|2]Indicates the priority used for review

My abstract is being submitted as a Software Demonstration.

Yes

Please indicate below if your study was a "resting state" or "task-activation" study.

Task-activation

Healthy subjects only or patients (note that patient studies may also involve healthy subjects):

Healthy subjects

Was any human subjects research approved by the relevant Institutional Review Board or ethics panel? NOTE: Any human subjects studies without IRB approval will be automatically rejected.

Yes

Was any animal research approved by the relevant IACUC or other animal research panel? NOTE: Any animal studies without IACUC approval will be automatically rejected.

Not applicable

Please indicate which methods were used in your research:

Functional MRI

For human MRI, what field strength scanner do you use?

3.0T

Which processing packages did you use for your study?

SPM
Other, Please list - MACS Toolbox (https://github.com/JoramSoch/MACS)

Provide references using author date format

[1] Reverberi, C. (2012), 'Compositionality of rule representations in human prefrontal cortex', Cerebral Cortex, vol. 22, iss. 6, pp. 1237-1246.

[2] Allefeld, C. (2016), 'Valid population inference for information-based imaging: From the second-level t-test to prevalence inference', NeuroImage, vol. 141, pp. 378-392.
[3] Cichy, R.M. (2011), 'Encoding the identity and location of objects in human LOC', NeuroImage, vol. 54, iss. 3, pp. 2297-2307.
[4] Allefeld, C. (2016), 'Test data based on Cichy et al. (2011)', GitHub; URL: https://github.com/allefeld/cichy-2011-category-smoothedaccuracy.
[5] Koch, K.R. (2007), 'Introduction to Bayesian Statistics', 2nd edition, Springer, Berlin/Heidelberg, ch. 2.4.4, p. 48.
[6] Minka, T. (2012), 'Estimating a Dirichlet distribution', technical report; URL: https://tminka.github.io/papers/dirichlet/minka-dirichlet.pdf.