1 Introduction

Multivariate classification algorithms are powerful tools for predicting cognitive or pathophysiological states from neuroimaging data [1]. However, there are unresolved questions regarding the statistical evaluation of classification performance in group studies.

Inference in group studies requires models that explicitly account for fixed-effects (within-subjects) and random-effects (across-subjects) variance components. This principle is standard practice in mass-univariate analyses [2-5] but has not yet been adopted in classification studies.

2 Variational mixed-effects inference

We recently introduced Bayesian models for mixed-effects inference in multivariate classification studies [6,7]. Until now, their practical utility was limited by the high computational complexity of the underlying Markov chain Monte Carlo (MCMC) sampling algorithms.

Here, we solve this remaining problem by introducing a variational Bayes (VB) approach to inference [8].

Conventional maximum-likelihood estimation

Bayesian mixed-effects inference (slow)

Variational Bayes approximation (fast)

Iterative optimization of posterior moments using EM

4 Illustrative examples

In a first setting, we generated synthetic classification outcomes for a group of 20 subjects with 200 trials each. VB inferences were practically indistinguishable from an MCMC reference implementation. Fixed-effects approaches, by contrast, were over- or underconfident.

In a second simulation, we considered a smaller, more heterogeneous group. VB and MCMC posterior means agreed nicely. Fixed-effects methods were over- or underconfident, and a conventional, random-effects t-test provided an uninterpretable confidence interval that included accuracies above 100%.

5 Advantages over conventional methods

1) Inference on the population mean

In contrast to the posterior Bayesian posterior interval (blue), the conventional confidence interval (red) is biased, and a conventional fixed-effects analysis (orange) is overconfident.

2) Inference on individual accuracies

Unlike conventional sample accuracies (blue), the proposed Bayesian posterior means of individual accuracies (black) are informed by data from (or ‘shrunk towards’) the entire group. This prevents overfitting and yields more precise estimates.

3) Inference on the balanced accuracy

The balanced accuracy (green) provides a more useful measure of classification performance than the accuracy (blue), especially in the context of imbalanced data. The example illustrates this in a single-subject setting.

4) Computational efficiency

While MCMC ultimately achieves a marginally lower error, VB is considerably more efficient by approximately 4 orders of magnitude.

3 Model assumptions and algorithm

p(\(\theta\), y | x) = p(y | x, \(\theta\)) p(\(\theta\))

Variational inference

p(\(\theta\) | y) = \(\arg\max\) p(y | x, \(\theta\)) p(\(\theta\))

Mean-field approximation

p(\(\theta\) | y) = \(\prod\) p(\(\theta_i\) | y)

Parameter assumptions

\(\theta_{ij} = x_{ij} \exp(\eta_{ij}) \) and \(\eta_{ij} \sim \mathcal{N}(\mu, \sigma)\)

In the second setting, we considered a smaller, more heterogeneous group. VB and MCMC posterior means agreed nicely. Fixed-effects methods were over- or underconfident, and a conventional, random-effects t-test provided an uninterpretable confidence interval that included accuracies above 100%.

6 Conclusions

- Bayesian mixed-effects inference for group studies (I) models within-subject and across-subjects uncertainty, (II) enables conclusions in terms of intuitive probability statements, and (III) can be used with various performance measures, e.g., the balanced accuracy. In contrast to previous sampling algorithms, our variational approximation is computationally highly efficient.

- We hope that our approach will improve the precision and interpretability of statistical inference in future classification studies.

References