Assessing the Significance of Excursion Regions in Functional Brain Imagery via Spatial Scan Analysis and Importance Sampling

Carey E. Priebe
Johns Hopkins University, Department of Mathematical Sciences
Baltimore, MD 21218–2682 USA
cep@jhu.edu

Daniel Q. Naiman
Johns Hopkins University, Department of Mathematical Sciences
Baltimore, MD 21218–2682 USA

Jiang Hu
Johns Hopkins University, Department of Mathematical Sciences
Baltimore, MD 21218–2682 USA

1. Introduction

We consider the problem of assessing the significance of excursion regions, or local regions of high intensity, in functional brain imagery (PET, fMRI, etc.). For example, one application of PET involves collecting volumetric brain image data from each of a number of subjects during two different states, A and B, and subtracting these to produce a contrast image $C = B - A$. Excursion regions in the contrast image C are considered to be regions with increased blood flow associated with state B relative to state A. This in turn is considered to be indicative of increased neural activity representing regionally specific effects attributable to state B. Assessing the significance of such regions is a key initial stage in the attempt to understand the workings of the brain.

We present a methodology based on spatial scan analysis and importance sampling which provides a computationally efficient unbiased estimate of the p-value for the test of homogeneity against the alternative of nonhomogeneity (the existence of regions with increased blood flow). Our methodology is indicated in cases where the true p-value is not so small as to allow for approximations based on extreme value theory, yet small enough to prohibit estimation via naive Monte Carlo simulation.

2. Statistical Parametric Mapping

A common approach to the analysis of functional brain imagery in general, and to the assessment of excursion region significance in particular, involves the use of statistical parametric maps (SPM) [see, e.g., Friston, et al. (1995)]. After registration, normalization, and transformation procedures have been applied, the subtractive application described above allows for the contrast image C to be modelled as a Gaussian random field (GRF) under the null hypothesis of "no change."

Let $C = \{C_x : x \in X\}$ be the (finitely indexed) observed GRF under consideration. (For PET applications the domain X is a $65 \times 87 \times 26$ voxel cube representing Talairach & Tournoux space. The actual search volume considered in the analysis consists of a subset of 47428 voxels in the Talairach & Tournoux cube; these are the voxels considered to actually represent locations in the brain.) The spatial covariance function (correlation structure) is assumed to be known a priori, or is estimated from the observed field.

An excursion region is defined to be a connected or contiguous subset of voxels above some specified threshold τ. To address the question of whether a particular excursion region is statistically significant, two methods for statistical inference are included in the SPM software: p-values based on the largest value in the region, and p-values based on the size of the region. (The latter may be more valuable in some situations due to increased specificity.) These p-value estimates are obtained based on extreme value theory approximations to the Euler (or Hadwiger) characteristic.

3. Spatial Scan Analysis

An alternative approach to the desired inference is provided by spatial scan analysis. Consider a scan window of some given geometry (size & shape) g. (In fact, multiple scan window sizes and shapes can be considered simultaneously by considering a set of geometries $\mathcal{G} = \{g_1, \cdots, g_K\}$.) At each location $x \in X$ the scan region $R(x; g) \subset X$ is defined. An averaging operation is performed on the field observations C_y associated with locations y in the scan region $R(x; g)$, thus yielding for each location $x \in X$ and geometry $g \in \mathcal{G}$ the locality statistic $\Psi_{x,g}$. By judicious choice of the geometries $\mathcal{G}$, sensitivity to specific anticipated excursion region characteristics can be achieved. The scan statistic is then defined to be $S := \max_{g \in \mathcal{G}} \max_{x \in X} \Psi_{x,g}$. For a particular excursion region in the observed field, an observed value $S = s$ is obtained. The significance of this value ($P[S \geq s]$) is, like that of the SPM statistics, unavailable in closed form.

Two computational approaches are available for obtaining an unbiased estimate of $p = P[S \geq s]$; naive Monte Carlo simulation and the importance sampling approach of Namian & Priebe (2001).

4. Comparison

Given a specified GRF model and a true but unknown p-value p we consider three approaches to assessing the significance of local excursions: extreme value theory (EVT), naive Monte Carlo (NMC), and importance sampling (IS). Our interest is in the practical performance of these three methods – accuracy of estimation and computational requirements.

Extreme value theory is, prima facie, based on limit theorems. As such, the accuracy of the estimates obtained is a function of the actual p-value itself; extreme value theory approximations can be accurate in the tails of the null distribution (for small p-values) but badly biased for moderate or large p-values. (Of course, the threshold τ and the GRF model itself – more precisely, the correlation structure – also effect the accuracy of the various p-value estimates.) However, the calculation of the EVT estimate is effectively instantaneous, and thus when the estimate is sufficiently accurate this approach is indicated.

Note, however, that the accuracy of EVT's computationally intensive competitors, NMC and IS, is a function of available time and resources for computation, and of their implementation details. A particular application – say, an SPM session in which preprocessing has been completed and the user clicks to request a p-value – has the specified operational parameter T – the maximum acceptable time to response, fixed computational resources – cpu speed, memory, etc., and a given implementation of NMC and IS. In this case, appealing to the mean squared error criterion, we define $p^\star$ to be the p-value at which the performance of NMC is equal to the performance of IS, and $p^{\star\star}$ to be the p-value at which the performance of EVT is equal to the performance of IS. Since for a particular model EVT yields a scalar function of the true p-value while both NMC and IS yield unbiased estimates, this equates to

$$p^\star := p \;\; such \;\; that \;\; \mathrm{V}ar[\hat{p}_{NMC}(p,T)] = \mathrm{V}ar[\hat{p}_{IS}(p,T)] \tag{1}$$

and

$$p^{\star\star} := p \;\; such \;\; that \;\; \mathrm{B}ias^2[\hat{p}_{EVT}(p)] = \mathrm{V}ar[\hat{p}_{IS}(p,T)]. \tag{2}$$

For $p < p^{\star\star}$ EVT is superior to IS, while for $p > p^\star$ NMC is superior to IS. Thus, if $p^{\star\star} < p^\star$ then IS is indicated for $p \in (p^{\star\star}, p^\star)$. If, furthermore, the interval $(p^{\star\star}, p^\star)$ can be demonstrated to encompass an operationally significant range of p-values (say, for instance, if $(p^{\star\star}, p^\star) \cap (.01, .1) \neq \emptyset$) then this importance sampling methodology should be an element of the practitioner's toolbox. Preliminary studies indicate that this is indeed the case.

Furthermore, if T remains unchanged, as available computational resources increase $p^{\star\star}$ will decrease (the time required for each importance sampling replication will decrease, allowing for more importance sampling replications and thus a decrease in $\mathrm{V}ar[\hat{p}_{IS}(p,T)]$) so that the proposed importance sampling methodology will come to dominate even more so as technology progresses.

5. Edge Effects

The orignial importance sampling implementation [Naiman & Priebe (2001)] ignored edges in the GRF. This introduces a bias into $\hat{p}_{IS}$ (and $\hat{p}_{NMC}$ as well). A new version has been implemented which considers the Talairach & Tournoux cube and the on-brain voxels therein directly, thereby incorporating edge effects. Since incorporating edge effects increases the time required for each importance sampling replication, the mean squared error of the p-value estimates may be smaller if edge effects are ignored. Note, however, that as the available computational resources increase (alternatively, as T increases) $p^{\star\star} \searrow 0$ if edge effects are incorporated (the unbiased case) while $p^{\star\star} \searrow c > 0$ when ignoring edge effects.

ACKNOWLEDGMENTS

This work is partially supported by Office of Naval Research Grants N00014-95-1-0777 and N00014-01-1-0011 and by the Acheson J. Duncan Fund for the Advancement of Research in Statistics. The authors thank Dr. Henry Holcomb and Ning Ma of the Maryland Psychiatric Research Center for PET data and expertise. The SPM software is developed by members & collaborators of the Wellcome Department of Cognitive Neurology, Institute of Neurology, University College London, and is available at <http://www.fil.ion.bpmf.ac.uk/spm/>.

REFERENCES

Naiman, D.Q. and Priebe, C.E. (2001). Computing scan statistic p-values using importance sampling, with applications to genetics and medical image analysis. *Journal of Computational and Graphical Statistics*, to appear.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D. and Frackowiak, R.S.J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, **2**, 189-210.

RÉSUMÉ

Ce papier présente une méthodologie pour adresser le problème d'évaluation de l'importance de régions d'excursion en cerveau imagerie.