Nonparametric mixture modeling for mapping of brain activation using functional MRI data

R. R. Nandy[1], D. Cordes[2]

[1]Psychology, University of California, Los Angeles, California, United States, [2]Radiology, University of Washington, Seattle, Washington, United States

Introduction

In conventional hypothesis based fMRI analysis, usually the analysis is based purely on the null hypothesis. It is difficult to construct a specific alternate hypothesis and the distribution of the test statistic under the alternate hypothesis is usually difficult to compute since the activation level is unknown and need not be the same for all voxels. To address this problem, Everitt and Bullmore [1] have used a simple test statistic and a simple alternate hypothesis for which the distribution of the test statistic under the alternate hypothesis can be evaluated parametrically with the use of two parameters estimated from the data using maximum likelihood estimates. The fitted distributions of the test statistic under null and alternate hypotheses were then used in a mixture model to construct activation maps. This method has been further extended by Hartvig and Jensen [2] to account for spatial coherence in fMRI data. However, in both of these methods, it is necessary to make several assumptions, some of which are questionable. Instead, in this abstract, a nonparametric mixture model is introduced which make minimal assumptions whose validity are not in question. Furthermore, the proposed nonparametric method can easily extend to multivariate statistical methods, such as canonical correlation analysis (CCA) to take into account the spatial coherence without making any explicit assumption about the spatial structure in the data.

Methods

Formulation of the nonparametric mixture model - The nonparametric mixture model can be formulated in a very general framework with minimal assumptions. We assume that in the event of no activation in the brain related to the active task of interest, the inactive voxels can be classified into k homogeneous groups and the null density functions of the test statistic in these groups are denoted by $f_1(x),\ldots,f_k(x)$, respectively. The proportions of voxels in these groups are denoted by $p_1,\ldots,p_k$, respectively. A similar formulation is made for the distributions of the test statistic under the hypothesis of activation. Again, the active voxels can be classified into l homogeneous groups and the corresponding density functions and proportions are denoted by $g_1(x),\ldots,g_k(x)$ and $q_1,\ldots,q_k$, respectively.

Clearly, the null and activation density functions for a random voxel can be expressed as $f(x)=\sum_{i=1}^{k} p_i f_i(x)$ and $g(x)=\sum_{j=1}^{l} q_j g_j(x)$ respectively. Then in activation data, the mixture density function can be expressed as $h(x)=\beta f(x)+(1-\beta)g(x)=\beta\sum_{i=1}^{k} p_i f_i(x)+(1-\beta)\sum_{j=1}^{l} q_j g_j(x)$ where β is the proportion of non-activated voxels in the active brain.

Let A_n denote the indicator function at voxel n, where A_n=1 corresponds to the event that the voxel is active and vice versa. We want to estimate the posterior probability of the voxel being active given the observed sample for the activation data. A simple calculation yields the following expression for the posterior probability of activation at voxel n:

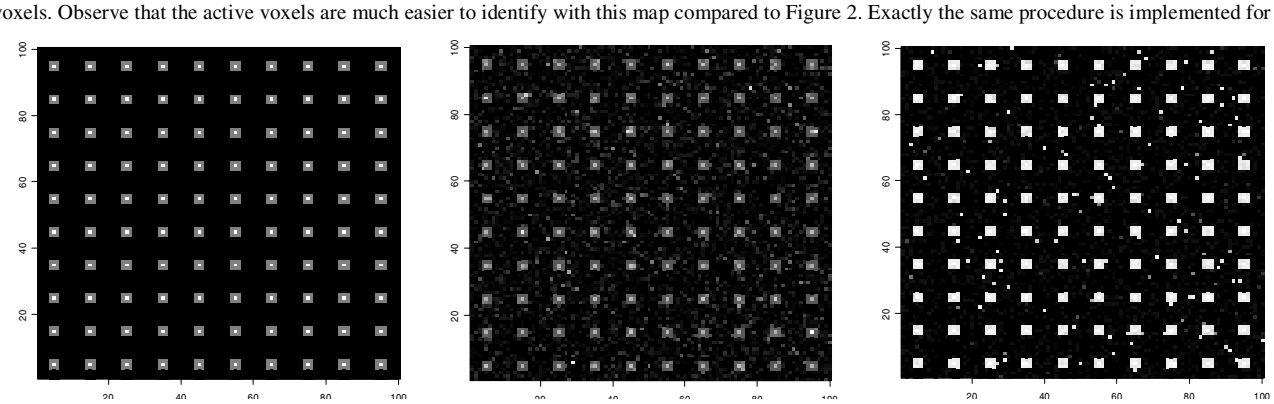
$$P(A_n=1\mid x_n)=\frac{(1-\beta)g(x_n)}{h(x_n)}=\frac{(1-\beta)\sum_{j=1}^{l} q_j g_j(x_n)}{\beta\sum_{i=1}^{k} p_i f_i(x_n)+(1-\beta)\sum_{j=1}^{l} q_j g_j(x_n)}=\frac{(1-\beta)g(x_n)}{\beta f(x_n)+(1-\beta)g(x_n)}.$$

Estimation of f, g and h - For the estimation process, it is necessary to acquire an additional set of resting state data with identical scanner protocol as in activation data. Since the resting-state data are not related to the paradigm used in the activation data, the resting-state data may be considered to be null relative to the paradigm of interest. The estimated empirical density function from the combined resting state data provides a direct estimate of $f(x)$. Furthermore, the activation data that we observe is already the mixed data and the estimated empirical density function from the activation data directly provides an estimate of $h(x)$. However, to obtain the estimates of the posterior probabilities, it is necessary to have an estimate of $g(x)$, which can be estimated from the above equation provided we can estimate β. β can be estimated using a method based on false positive fraction (*FPF*). For a fixed value of the test statistic, the corresponding *FPF* is calculated from the resting-state data. The *FPF* gives us an estimate of the number of voxels expected to be detected by error. Suppose the total number of voxels is N and m is the number of voxels detected to be active for a threshold α. It can be shown that a lower bound for β is $\frac{k-\alpha N}{(1-\alpha)N}$. The best estimate is obtained when the estimate attains the peak to take the maximum value.

Estimating the power of a test in the hypothesis based approach in fMRI - As mentioned previously, one of the weaknesses of the usual parametric hypothesis based formulation in fMRI data analysis is the absence of a proper alternate hypothesis which is specific. In effect, it is not possible to estimate the power of the statistical test in this parametric formulation. However, the ideas involved in the nonparametric mixture modeling can be used to estimate the power. In this setup the distribution of the test statistic under alternate hypothesis of activation is simply $g(x)$. Since we have already described the process to estimate $g(x)$, estimation of power for any chosen threshold is trivial.

Results

Due to space limitation, we only provide a simple simulation result for a single 100×100 slice. The preselected active voxels with activation levels are shown in Figure 1. The observed values of the F-statistic at these voxels after adding noise are shown in Figure 2. We also simulated a 100×100 slice with pure noise (equivalent to resting state for real data), which is used to estimate $f(x)$. In Figure 3, we present the map for posterior probabilities on the same slice, which nicely indicate the active voxels. Observe that the active voxels are much easier to identify with this map compared to Figure 2. Exactly the same procedure is implemented for activation data.



Figures 1-3

References

1. Everitt BS and Bullmore ET. Human Brain Mapping 7:1–14(1999).
2. Hartvig NV and Jensen JL. Human Brain Mapping 11:233–248(2000).