Bayesian Inference

Will Penny

Wellcome Department of Imaging Neuroscience, University College London, UK

SPM Course, London, May 2004
Bayesian Inference

- Gaussians
- Posterior Probability Maps (PPMs)
- Hemodynamic Models (HDMs)
- Comparing models (DCMs)
One parameter

Likelihood and Prior

\[ p(y \mid \theta^{(1)}) = N(\theta^{(1)}, \lambda^{(1)}) \]
\[ p(\theta^{(1)}) = N(\theta^{(2)}, \lambda^{(2)}) \]

Posterior

\[ p(\theta^{(1)} \mid y) = N(m, P) \]
\[ P = \lambda^{(1)} + \lambda^{(2)} \]
\[ m = \frac{\lambda^{(1)}}{P} \theta^{(1)} + \frac{\lambda^{(2)}}{P} \theta^{(2)} \]

Relative Precision Weighting
Two parameters
Posterior Probability Maps - PPMs

Least squares parameter estimation

Bayesian parameter estimation

No Priors

Shrinkage priors
Prior:

\[ p(\mu_i) = N(0, \alpha) \]

\[ \frac{1}{\alpha} = 1 \]
Data

Likelihood:

\[ p(y_{i,n} \mid \mu_i) = N(\mu_i, \beta_i) \]

i=1..V=20 voxels
n=1..N=10 scans
Least Squares Estimates

\[ \hat{\mu}_i = \frac{1}{N} \sum_{n=1}^{N} y_{i,n} \]

\[ \frac{1}{\beta_i} = \frac{1}{N-1} \sum_{n=1}^{N} (y_{i,n} - \hat{\mu}_i)^2 \]
Least Squares Estimates

\[ \hat{\mu}_i = \frac{1}{N} \sum_{n=1}^{N} y_{i,n} \]

\[ \frac{1}{\beta_i} = \frac{1}{N-1} \sum_{n=1}^{N} (y_{i,n} - \hat{\mu}_i)^2 \]

Error=7.10
Bayesian Estimates

E-Step:
\[ \hat{\mu}_i = \frac{\gamma_i}{N} \sum_{n=1}^{N} y_{i,n} \]

M-Step:
\[ \frac{1}{\hat{\beta}_i} = \frac{1}{N - \gamma_i} \sum_{n=1}^{N} (y_{i,n} - \hat{\mu}_i)^2 \]
\[ \hat{\alpha} = 0 \quad \gamma_i = 1 \]

Iteration 1   Error=7.10
Bayesian Estimates

E-Step:

$$\hat{\mu}_i = \frac{\gamma_i}{N} \sum_{n=1}^{N} y_{i,n}$$

M-Step:

$$\frac{1}{\hat{\beta}_i} = \frac{1}{N - \gamma_i} \sum_{n=1}^{N} (y_{i,n} - \hat{\mu}_i)^2$$

$$\gamma_i = \frac{N \hat{\beta}_i}{N \hat{\beta}_i + \hat{\alpha}}$$

$$1 = \frac{1}{\hat{\alpha}} \sum_i \gamma_i \sum_{i=1}^{V} \hat{\mu}_i^2 = 1.09$$

Iteration 2  Error=2.95
Bayesian Estimates

E-Step:

\[ \hat{\mu}_i = \frac{\gamma_i}{N} \sum_{n=1}^{N} y_{i,n} \]

M-Step:

\[ \frac{1}{\hat{\beta}_i} = \frac{1}{N - \gamma_i} \sum_{n=1}^{N} \left( y_{i,n} - \hat{\mu}_i \right)^2 \]

\[ \gamma_i = \frac{N \hat{\beta}_i}{N \hat{\beta}_i + \hat{\alpha}} \]

\[ \frac{1}{\hat{\alpha}} = \frac{1}{\sum_i \gamma_i} \sum_{i=1}^{V} \hat{\mu}_i^2 = 1.03 \]
SPMs and PPMs

PPMs: Show activations of a given size

SPMs: show voxels with non-zero activations
Hemodynamic Models - HDMs

\[ \dot{s} = \varepsilon_1 u_1(t) + \ldots + \varepsilon_f u_f(t) - s/\tau_s - (f_{in} - 1)/\tau_f \]

Photic

Motion

Attention

fMRI study of attention to visual motion
PPMs

Advantages

One can infer a cause
DID NOT elicit a response

SPMs conflate effect-size and effect-variability

Disadvantages

Use of shrinkage priors over voxels is computationally demanding

Utility of Bayesian approach is yet to be established
Comparing Dynamic Causal Models

Bayesian Evidence:

\[ p(y|m) = \int p(y|m, \theta) p(\theta|m) d\theta \]

Bayes factors:

\[ B_{12} = \frac{p(y|m = 1)}{p(y|m = 2)} \]