Bayesian Statistics

Fabio Sigrist

ETH Zurich, Autumn Semester 2019

Today's topics

- Empirical Bayes methods
- Examples

Empirical Bayes methods

Empirical Bayes method

- Recall that the marginal posterior can be computed as

$$\pi(\theta \mid x) \propto \int \pi(\theta \mid x, \xi) f(x \mid \xi) \pi(\xi) d\xi$$

- Instead of approximating this integral, the **empirical Bayes method** uses

$$\pi(\theta \mid x) \approx \pi(\theta \mid x, \widehat{\xi}(x)) = \frac{f(x|\theta)\pi(\theta|\widehat{\xi}(x))}{f(x \mid \widehat{\xi}(x))}$$

where

$$\widehat{\xi}(x) = \arg\max_{\xi} f(x \mid \xi)$$

and

$$f(x \mid \xi) = \int f(x|\theta)\pi(\theta|\xi) d\theta$$

Empirical Bayes method

- $\widehat{\xi}(x)$ is the marginal maximum likelihood estimator of the hyperparameter
- Instead of taking a weighted average, one takes the value with maximal weight (assuming that $\pi(\xi)$ is flat around $\widehat{\xi}(x)$)

Comments on the empirical Bayes method

Advantage

- This method avoids not only the computation of the integral, but also the choice of a hyperprior $\pi(\xi)$

Drawbacks

- The data x is used twice:
  1. First, to select the prior $\pi(\theta \mid \widehat{\xi}(x))$
  2. Then, to compute the posterior according to Bayes formula
  - This is somewhat undesirable from a conceptual point of view
- In general, the uncertainty is underestimated

Conclusion

- Due to this, Bayesians often avoid empirical Bayes methods
- From a pragmatic point of view, empirical Bayes methods can be useful and have good frequentist properties

Examples of the application of empirical Bayes methods

- Normal means
- Hierarchical Poisson model
- One-way ANOVA model

Normal means example

Assume the following model:

- Likelihood: $X_1, \dots X_n$ i.i.d. $\sim \mathcal{N}(\theta, 1)$
- Hierarchical prior:
  - $\theta \sim \mathcal{N}(\mu, \tau^2)$, μ known
  - $\tau^{-2} \sim \text{Gamma}(\gamma, \lambda)$

See blackboard for more details

Clicker question

Hierarchical Poisson model example

Assume the following model:

- Likelihood: X_j independent $\sim$ Poisson(θ_j), $j = 1, 2, \ldots, J$
- Hierarchical prior:
  - θ_j i.i.d. $\sim$ Gamma(γ, λ)
  - Some prior $\pi(\gamma, \lambda)$

See blackboard for more details

Clicker question

One-way ANOVA model example

Assume the following model:

- Likelihood: $y_{ij} = \theta_i + \varepsilon_{ij}, \quad \varepsilon_{ij}$ i.i.d $\sim \mathcal{N}(0, \sigma_\varepsilon^2)$, $j = 1, \dots, n_i, i = 1, \dots, I$
- Hierarchical prior:
  - θ_i i.i.d $\sim \mathcal{N}(\mu, \tau^2)$
  - For simplification, σ_ε is assumed to be known
  - $\pi(\mu, \tau^2) = \pi(\mu)\pi(\tau^2)$ (specified later)

See blackboard for more details

Clicker question