ADVANCED BAYESIAN MODELING

Bayesian Tools and Integration

Review: Notation and Concepts

Parameters (including hyperparameters): $\theta = (\theta_1, \ldots, \theta_d)$

Sampling density of data y (likelihood in θ): $p(y \mid \theta)$

Prior density: $p(\theta)$

Posterior density (by Bayes' rule):

$$p(\theta \mid y) \;=\; \frac{p(\theta)\,p(y \mid \theta)}{p(y)} \;\propto\; p(\theta)\,p(y \mid \theta)$$

Bayesian tools are based on $p(\theta \mid y)$:

- Posterior means and standard deviations
- Posterior probabilities
- Marginal posterior densities
- Posterior quantiles
- Posterior intervals
- Posterior predictions

Note: θ is almost always jointly continuous (even if y is discrete).

Therefore, most tools involve integration ...

Posterior Expectations

Many tools are posterior expectations of some scalar function $h(\theta)$,

$$\mathrm{E}\big(h(\theta) \mid y\big) \;=\; \int h(\theta)\, p(\theta \mid y)\, d\theta$$

which can be a multidimensional integral.

E.g., posterior mean of θ_j: $\qquad h(\theta) \;=\; \theta_j$

E.g., posterior variance of θ_j: $\qquad h(\theta) \;=\; \big(\theta_j - \mathrm{E}(\theta_j \mid y)\big)^2$

(Note: h can also depend on y)

A posterior probability is just a type of posterior expectation:

$$\begin{aligned}\Pr(\theta \in A \mid y) &= \int_A p(\theta \mid y)\, d\theta \\ &= \int 1_A(\theta)\, p(\theta \mid y)\, d\theta \;=\; \mathrm{E}\big(1_A(\theta) \mid y\big)\end{aligned}$$

where 1_A represents an indicator function:

$$1_A(\theta) = \begin{cases} 1, & \theta \in A \\ 0, & \text{otherwise} \end{cases}$$

Posterior Marginals

Let

$$\theta_{-j} \;=\; \theta \text{ without } \theta_j$$

Then the marginal posterior density of θ_j is

$$p(\theta_j \mid y) \;=\; \int p(\theta \mid y)\, d\theta_{-j}$$

Posterior Quantiles and Intervals

The posterior (lower) α quantile q_α of θ_j is defined implicitly in terms of an integral:

$$\alpha \;=\; \Pr(\theta_j \leq q_\alpha \mid y) \;=\; \int_{-\infty}^{q_\alpha} p(\theta_j \mid y)\, d\theta_j$$

For example, the median is $q_{0.5}$.

A central $(1-\alpha)$100% posterior interval for θ_j is just defined in terms of posterior quantiles:

$$(q_{\alpha/2},\; q_{1-\alpha/2})$$

Posterior Prediction

An observable (but unobserved) $\tilde{y}$ with sampling density

$$p(\tilde{y} \mid \theta) \qquad\qquad \text{(conditionally independent of } y\text{)}$$

has posterior predictive density

$$p(\tilde{y} \mid y) \;=\; \int p(\tilde{y} \mid \theta)\, p(\theta \mid y)\, d\theta$$

If $\tilde{y}$ is continuous, also use integration to find posterior predictive expectations (including probabilities).

Priors and Integration

There are also integrations based on the prior.

The normalizing factor is an example:

$$p(y) \;=\; \int p(y \mid \theta)\, p(\theta)\, d\theta$$

If the prior is proper, there are also prior expectations and probabilities:

$$\mathrm{E}\big(h(\theta)\big) \qquad\qquad \Pr(\theta \in A) \;=\; \mathrm{E}\big(1_A(\theta)\big)$$

These are defined by integrations involving the prior.