ADVANCED BAYESIAN MODELING

Importance Sampling

Situation

Goal: Approximate posterior expectation $\mathrm{E}\big(h(\theta) \mid y\big)$.

Assume we can evaluate unnormalized continuous posterior density

$$q(\theta) \;\; \propto \;\; p(\theta \mid y)$$

Assume we can easily sample independently from positive continuous density $g(\theta)$, and also evaluate $g(\theta)$ (either exactly or up to a constant).

Importance Sampling Algorithm

1. Independently sample from $g(\theta)$ to produce

$$\theta^1, \quad \theta^2, \quad \ldots, \quad \theta^S$$

2. Approximate

$$\mathrm{E}\big(h(\theta) \mid y\big) \approx \frac{\sum_{s=1}^{S} h(\theta^s)\, w(\theta^s)}{\sum_{s=1}^{S} w(\theta^s)} \qquad \text{where} \qquad w(\theta^s) = \frac{q(\theta^s)}{g(\theta^s)}$$

Values $w(\theta^s)$ are the **importance weights**.

(Justification: BDA3, Sec. 10.4)

Remarks:

- Can reuse samples and weights many times (different functions h).
- Can estimate an effective sample size (BDA3, Sec. 10.4).
- What if an independent posterior sample of θ is still needed?

  Consider *importance resampling* (BDA3, Sec. 10.4).
- Worst case: Weights are highly skewed toward large values.