# ADVANCED BAYESIAN MODELING

# Evaluating Monte Carlo Error

# Monte Carlo Approximation Error

**Monte Carlo error** is the deviation of the Monte Carlo approximation from the true value.

It is often measured by the **Monte Carlo standard error (SE)**: the standard deviation of the approximation that is due to random simulation.

Note: NOT the same as a posterior (or prior) standard deviation!

Approximations having a roughly normal simulation distribution will be within two SE of the true value about 95% of the time.

# Independent Simulation

When the variates are simulated independently, Monte Carlo error is easy to evaluate.

For now, assume independent simulation draws

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

from the posterior.

(Later: More complicated situations)

# For Expected Values

The Monte Carlo approximation

$$\frac{1}{S}\sum_{s=1}^{S} h(\theta^s) \qquad \text{for} \qquad \mathrm{E}\big(h(\theta) \mid y\big)$$

is correct on average:

$$\mathrm{E}\left(\frac{1}{S}\sum_{s=1}^{S} h(\theta^s) \mid y\right) \;=\; \frac{1}{S}\sum_{s=1}^{S} \mathrm{E}\big(h(\theta^s) \mid y\big) \;=\; \mathrm{E}\big(h(\theta) \mid y\big)$$

Moreover, the law of large numbers implies that the approximation converges in probability to $\mathrm{E}\big(h(\theta) \mid y\big)$ as $S \to \infty$.

Using independence, the variance of the approximation (if it exists) is

$$\operatorname{var}\left(\frac{1}{S}\sum_{s=1}^{S} h(\theta^s) \mid y\right) \;=\; \frac{1}{S^2}\sum_{s=1}^{S} \operatorname{var}\big(h(\theta^s) \mid y\big) \;=\; \frac{1}{S}\operatorname{var}\big(h(\theta) \mid y\big)$$

So the Monte Carlo SE is

$$\frac{1}{\sqrt{S}}\sqrt{\operatorname{var}\big(h(\theta) \mid y\big)}$$

Notes:

- Choosing $S$ sufficiently large can achieve high accuracy.
- Can estimate $\sqrt{\operatorname{var}\big(h(\theta) \mid y\big)}$ using simulated variates.

Let

$$\widehat{\mathrm{var}}(h) \quad = \quad \text{sample variance of draws } h(\theta^s)$$

Then the estimated Monte Carlo SE is

$$\frac{1}{\sqrt{S}} \sqrt{\widehat{\mathrm{var}}(h)}$$

Remember: This assumes draws are *independent*.

(When draws are NOT independent, this is called "`Naive SE`" in `rjags/coda`.)

# How Many Draws?

The number of independent draws needed to approximately attain a given SE is

$$\hat{S} \;\approx\; \frac{\widehat{\text{var}}(h)}{\mathsf{SE}^2}$$

A rule of thumb: Monte Carlo SE should be less than $1/20$ of the (posterior) standard deviation.

The number of independent draws needed for this would be

$$S \;>\; \frac{\widehat{\text{var}}(h)}{\left(\sqrt{\widehat{\text{var}}(h)}/20\right)^2} \;=\; 20^2 \;=\; 400$$

Remarks:

- If (posterior) variance exists, central limit theorem implies Monte Carlo approximate expected value is approximately normally distributed.
- If simulation draws are dependent, need different error estimates (later).
- Error estimates for approximated quantiles exist, but are less commonly used.