STT 465

Monte Carlo Approximations

(G. de los Campos)

Motivation

⇒ In Bayesian analyses (and in in many other applications) we need to compute integrals.

⇒ Examples

- Marginal distribution $p(y_1) = \int p(y_1, y_2)\, dy_2 = \int p(y_1 \mid y_2)\, p(y_2)\, dy_2$
- Posterior mean (or other moments) $E(\theta \mid y) = \int \theta\, p(\theta \mid y)\, d\theta$
- Posterior probabilities. $p(\theta \subset \Omega \mid y) = \int_{\theta \subset \Omega} p(\theta \mid y)\, d\theta$

⇒ In most cases the integrals are difficult or even impossible to compute analytically.

⇒ **MC integration:** we can use sample from a distribution to approximate integrals.

Idea

⇒ We can use computers to draw samples from the target distribution

$$\{\theta_s\}_{s=1}^{S} \overset{iid}{\sim} p(\theta \mid y)$$

⇒ The empirical distribution approximates the target distribution

⇒ The quality of the approximation will simply depend on the number of samples (S)

<u>Example 1</u>: approximating a normal density

⇒ In most cases we are interested on some integral, e.g., $E\left[g(\theta) \mid y\right] = \int g(\theta) \times p(\theta \mid y) d\theta$

⇒ WLLN $\frac{1}{S}\sum_s g(\theta_s) \rightarrow E\left[g(\theta) \mid y\right]$

Examples

⇒ Example 2: Computing areas in a unit square

⇒ Example 3: Convergence (Figure 4.2)

- Running Mean
- Evaluating MC-Error by # of samples

⇒ Example 4: Evaluating functions of parameters (e.g., log-odds)