Rejection Sampling

PUBH 8442: Bayes Decision Theory and Data Analysis

Eric F. Lock
UMN Division of Biostatistics, SPH
elock@umn.edu

03/31/2021

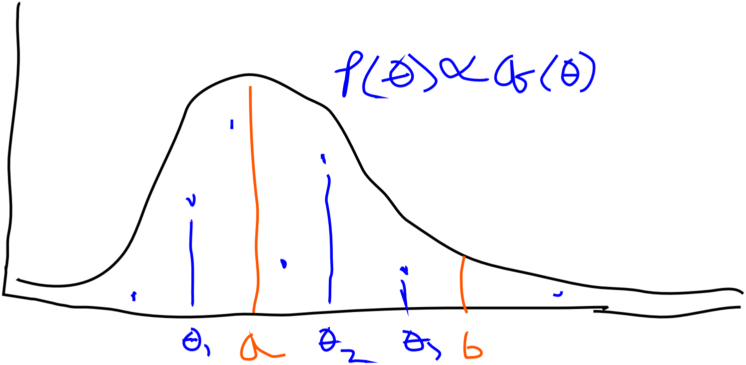
Background

- We would like to sample from a (potentially unknown) density p
- This is equivalent to sampling uniformly from the area under p
- If $p \propto q$, this is also equivalent to sampling uniformly from area under q
- In our context, p is a posterior distribution and q is the unnormalized version

$$q(\overset{\theta}{\cancel{x}}) = p(\mathbf{y} \mid \theta) p_\theta(\theta).$$

- How to sample uniformly underneath q?

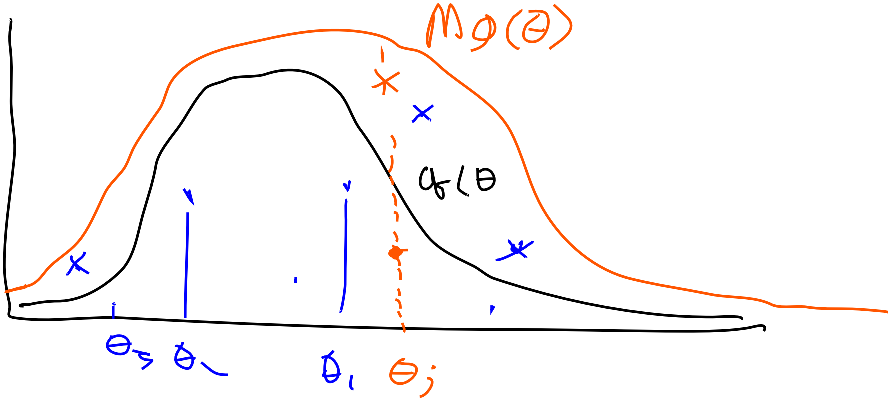
$$p(\theta) \propto q(\theta)$$



$$P(\theta \in (a,b)) = \int_a^b p(\theta)\,d\theta$$

Background

- Choose an *enveloping function* Mg that you can sample underneath
$$q(\theta) \leq Mg(\theta)$$
- Ignore those samples that are above q
- Those samples underneath q are uniform underneath q. Use these.
- In practice, g is a density function and M is a constant to assure Mg envelopes q


$Mq(\theta)$
$q(\theta$
θ_3
θ_2
θ_1
θ_j

Rejection sampling

- Algorithm to simulate from posterior distribution $p(\theta \mid \mathbf{y})$:
  - Generate $\theta_j \sim g(\theta)$
  - Generate $U \sim \text{Uniform}(0, 1)$
    - Note $(\theta_j, UMg(\theta_j))$ is a random point under $Mg(\theta_j)$

    $q(\theta_j)$
  - If $MUg(\theta_j) < p(\mathbf{y} \mid \theta_j)p(\theta_j)$ accept θ_j; otherwise reject θ_j.
- The accepted θ_j correspond to draws from $p(\theta \mid \mathbf{y})$
- Repeat above process until desired number of (accepted) samples N is obtained

Rejection sampling

- Equivalent algorithm:

  1. Generate $\theta_j \sim g(\theta)$
  2. Accept with probability

$$p_j = \frac{\overbrace{p(\mathbf{y} \mid \theta_j)p(\theta_j)}^{q(\theta_j)}}{Mg(\theta_j)}$$

  otherwise reject θ_j.

$$\to E p_j = \int \frac{q(\theta)}{M\cancel{g(\theta)}} \cancel{g(\theta)}\, d\theta = \frac{\int q(\theta)\, d\theta}{M}$$

Comments

- Similar to importance sampling
  - Simulations from known density $g(\theta)$ used to approximate unknown $p(\theta \mid \mathbf{y})$
  - Importance sampling is analogous to using p_j as weights instead of acceptance probabilities
- Rejection sampling yields direct samples from $p(\theta \mid \mathbf{y})$ (not weighted samples)
- Envelope function $Mg(\theta)$ MUST satisfy $Mg(\theta) \geq p(\cancel{\theta \mid \mathbf{y}})p(\theta)\ \forall\theta$

  $y \mid \theta$
  - Otherwise results are inconsistent
- Ideally, M is as small as possible and $Mg(\theta)$ is close to $p(\theta \mid \mathbf{y})p(\theta)$

Simple example: normal density

- Goal: simulate samples from $p(\theta) = N(0, 1)$ using rejection sampling[1]

- Use $g(\theta) = \text{Cauchy}(0, 2)$ as envelope function.

$$g(\theta) = \frac{1}{2\pi\left[1 + \left(\frac{\theta}{2}\right)^2\right]}$$

$M g(0) \geq p(0)$

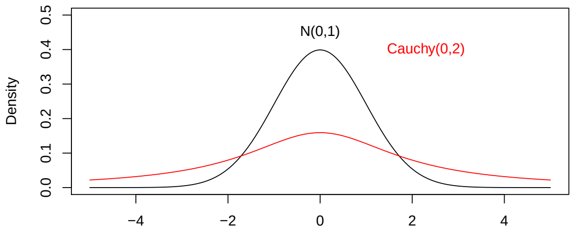
$M \frac{1}{2\pi} \geq \frac{1}{\sqrt{2\pi}} \frac{e^0}{1}$

$M \geq \sqrt{2\pi}$

[1] Example source: http://glau.ca/?p=227

Simple example: normal density

- Consider $M = 1$:

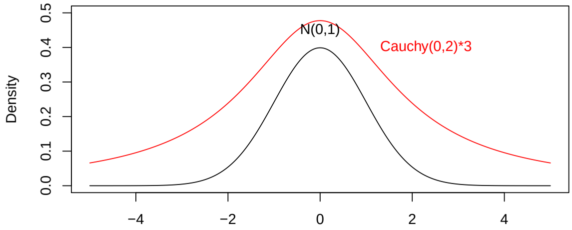


http://www.ericfrazerlock.com/Rejection_Sampling_Rcode1.r

- Not a proper envelope

Simple example: normal density

- Try $M = 3$:

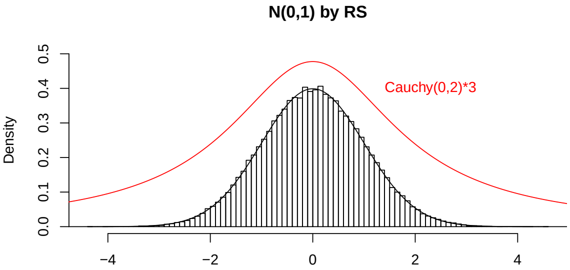


http://www.ericfrazerlock.com/Rejection_Sampling_Rcode1.r

- This is a proper envelope (though somewhat inefficient)

Simple example: normal density

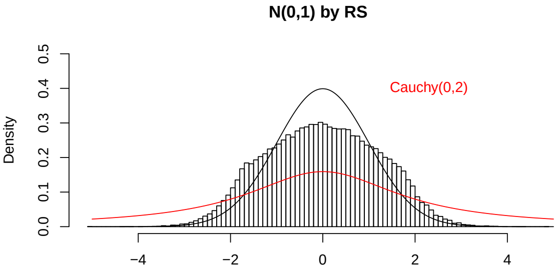
- Histogram of accepted samples using $M = 3$:



http://www.ericfrazerlock.com/Rejection_Sampling_Rcode1.r

Simple example: normal density

- Histogram of accepted samples using $M = 1$:



http://www.ericfrazerlock.com/Rejection_Sampling_Rcode1.r