MTH 511a - 2020: Lecture 9

Instructor: Dootika Vats

1 Some insight into Accept-Reject proposals

1.1 Choosing proposal:

Sometimes it is difficult to find a good proposal or even one that works! That is, for a target density $f(x)$ it can sometimes be challenging to find a proposal density $g(x)$ such that

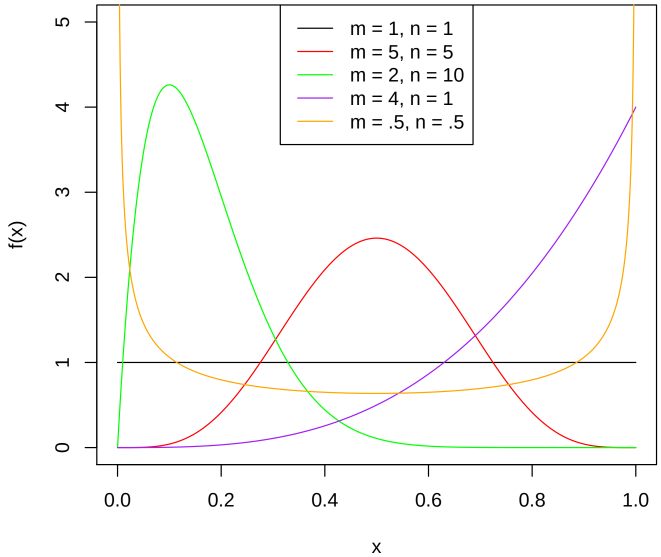
$$\sup_x \frac{f(x)}{g(x)} < \infty$$

Here are certain examples of when it may be difficult / impossible to implement accept-reject.

Example 1 (Beta). Consider a Beta(m, n)

$$f(x) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} x^{m-1}(1-x)^{n-1}$$

Depending on m and n, the Beta distribution can behave quite differently. Below is the figure for different choices of m and n. Particularly, note that when both $m, n < 1$ the Beta density function is unbounded!



When $m, n < 1$, if we use a uniform proposal distribution

$$\sup_{x \in (0,1)} \frac{f(x)}{g(x)} = \sup_{x \in (0,1)} \frac{f(x)}{1} = \infty! \,.$$

So a Uniform distribution will not work! In fact, any proposal density with $g(x) < \infty$ will not work. So this is an example of a distribution where it is difficult to find a good proposal distribution.

However, when say $n \geq 1$, then

$$\begin{aligned} f(x) &= \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} x^{m-1}(1-x)^{n-1} \\ &\leq \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} x^{m-1} \,. \end{aligned}$$

If we look at the upper bound, the function x^{m-1} on $x \in (0,1)$ can define a valid distribution if normalized. So, consider $g(x) = mx^{m-1}$, which is a proper density on $0 \leq x \leq 1$, and

$$\frac{f(x)}{g(x)} \leq m \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} := c$$

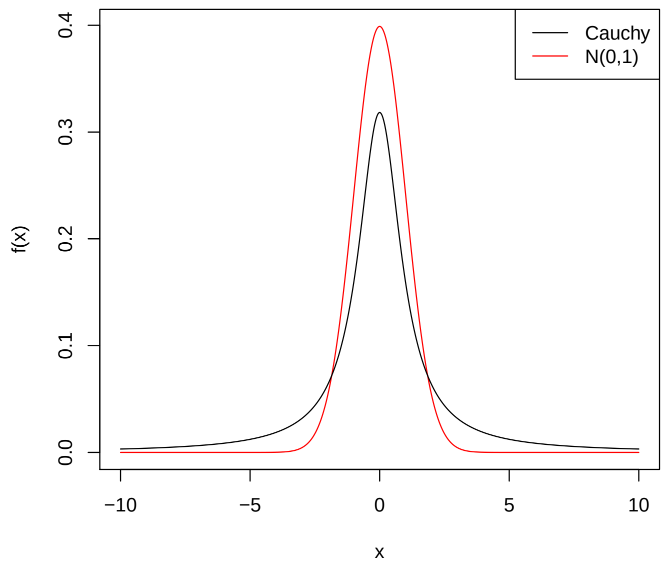
This Accept-Reject sampler can be implemented easily.

Example 2 (Accept-Reject for Cauchy target). Consider the target density

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2} \quad x \in \mathbb{R}\,,$$

The Cauchy distribution is known to have "fat tails" so that as $x \to \pm\infty$, the density function reduces to zero slowly. This means that it is very challenging to find $g(x)$ that "dominates" the density in the tails.

For example, let the proposal be $N(0, 1)$. The figure below shows the comparison of the tails and it is evident that the Normal distribution reduces dramatically compared to the Cauchy distribution.



So, if we look at

$$\frac{f(x)}{g(x)} = \frac{2\pi}{\pi} \frac{e^{x^2/2}}{1 + x^2} \to \infty \text{ as } x \to \pm\infty!$$

Thus Accept-reject can't be implemented here. In fact, as far we know, there are no possible standard accept-reject algorithms possible here.

1.2 Choosing parameters for proposal

If you have chosen a family of proposal distributions that you know give a finite c, it may be unclear what the best parameters for that proposal distribution is. That is, if the target $f(x)$ and the proposal density is $g(x|\theta)$, then you want to find a value of the parameter θ so that the resulting proposal is the "best".

Notice that the upper bound will be a function of θ, so that

$$\sup_x \frac{f(x)}{g(x|\theta)} \leq c(\theta)\,.$$

The value $c(\theta)$ is the expected number of loops for the accept-reject algorithm. Since we want this to be small, the best proposal density within this family would be the one that minimizes $c(\theta)$, so set

$$\theta^* := \arg\min_\theta c(\theta)\,.$$

Example 3 (Gamma distribution). Consider the target distribution Gamma(α, β)

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\,.$$

Further, suppose we want to use an Exp(λ) proposal. Then

$$g(x|\lambda) = \lambda e^{-\lambda x}\,.$$

We can now find $c(\lambda)$,

$$\begin{aligned} \frac{f(x)}{g(x)} &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{x^{\alpha-1} e^{-\beta x}}{\lambda e^{-\lambda x}} \\ &= \frac{\beta^\alpha}{\lambda\Gamma(\alpha)} x^{\alpha-1} e^{-x(\beta-\lambda)}\,. \end{aligned}$$

First note that no matter what λ is, if $0 < \alpha < 1$, then $f(x)/g(x) \to \infty$ as $x \to 0$. So accept-reject with this proposal won't work!

However, when $\alpha \geq 1$, then $x^{\alpha-1}$ increases, so we want to choose λ such that $e^{-x(\beta-\lambda)}$ decreases (since exponential decay is more powerful than polynomial increase) (of course, you should show this more mathematically). Thus we want $\beta > \lambda$!

Finally, we can now set $\alpha \geq 1$, $\lambda < \beta$,

$$c(\lambda) = \sup_x \frac{f(x)}{g(x)} = \sup_{x]} \frac{\beta^\alpha}{\lambda\Gamma(\alpha)} x^{\alpha-1} e^{-x(\beta-\lambda)}$$

which you can show, occurs at

$$x = \frac{\alpha-1}{\beta-\lambda}\,,$$

for which

$$c(\lambda) = \frac{\beta^\alpha}{\lambda\Gamma(\alpha)} \left(\frac{\alpha-1}{\beta-\lambda}\right)^{\alpha-1} e^{1-\alpha}\,,$$

which is minimized for

$$\lambda = \beta/\alpha.$$

Thus, the optimal exponential proposal for the Gamma$(\alpha, \beta), \alpha > 1$ is Exp(β/α).

2 Questions to think about

1. How would you implement accept-reject for Gamma(α, β) for $0 < \alpha < 1$?
2. Can we sample from $N(0, 1)$ using a Cauchy proposal?