Econ 551: Note on A/R Sampling[1]

Professor John Rust, Department of Economics, Yale University

A/R sampling (short for *acceptance-rejection sampling*) is a way of generating random draws $\tilde{\theta}$ from a density $p(\theta)$ that is hard to sample from by drawing from another *envelope* or *majorizing* density $g(\theta)$ that is easy to generate random draws from. The majorizing density g must satisfy the following condition:

$$\sup_{\theta \in C} \frac{p(\theta)}{g(\theta)} \leq a < \infty \tag{1}$$

where $C = \text{support}(p)$. Note that to do A/R method we must know some constant a satisfying inequality (1), since the A/R procedure works as follows

1. Draw a uniform random draw $\tilde{u} \sim U(0,1)$.
2. Draw a random draw $\tilde{z} \sim g(z)$, i.e. a random draw from the majorizing density.
3. if $\tilde{u} > p(\tilde{z})/[ag(\tilde{z})]$ reject the draw and go back to step 1. Otherwise **stop,** accept the draw, and set $\tilde{\theta} = \tilde{z}$ as a random draw from $p(\theta)$.

Claim: $\tilde{\theta}$ is a random draw from $p(\theta)$.

Proof: We need to show that

$$P\{\tilde{\theta} \leq x | \text{stop}\} = \int_{-\infty}^{x} p(\theta')d\theta' \tag{2}$$

where the probability on the left hand side of equation (2) is the conditional probability that $\tilde{\theta} \leq x$ given that the A/R simulation procedure has stopped. However that probability can be written as:

$$P\{\tilde{\theta} \leq x | \text{stop}\} = \frac{P\{\tilde{\theta} \leq x \cap \text{stop}\}}{P\{\text{stop}\}} \tag{3}$$

The probability in the numerator of (3) can be written as

$$P\{\tilde{\theta} \leq x \cap \text{stop}\} = \int_{-\infty}^{x} P\{\text{stop}|\theta\}g(\theta)d\theta \tag{4}$$

But the probability of stopping given a draw $\tilde{\theta} = \tilde{z}$ is just

$$P\{\text{stop}|\tilde{z}\} = P\left\{\tilde{u} \leq \frac{p(\tilde{z})}{ag(\tilde{z})}\right\} = \frac{p(\tilde{z})}{ag(\tilde{z})} \tag{5}$$

since $\tilde{u}$ is uniformly distributed and independent of $\tilde{z}$. Substituting equation (5) into equation (4) we get:

$$P\{\tilde{\theta} \leq x \cap \text{stop}\} = \int_{-\infty}^{x} \frac{p(\theta)}{ag(\theta)}g(\theta)d\theta = \int_{-\infty}^{x} \frac{p(\theta)}{a}d\theta \tag{6}$$

[1] This discussion closely follows the treatment in section 3.2 of John Geweke's (1996) article, "Monte Carlo Simulation and Numerical Integration" in the *Handbook of Computational Economics* H. Amman, D. Kendrick and J. Rust (eds.), Elsevier.

Letting $x \to \infty$ we get

$$P\{\text{stop}\} = P\{\tilde{\theta} \leq +\infty \cap \text{stop}\} = \int_{-\infty}^{+\infty} \frac{p(\theta)}{ag(\theta)} g(\theta) d\theta = \int_{-\infty}^{+\infty} \frac{p(\theta)}{a} d\theta = \frac{1}{a} \tag{7}$$

Substituting equations (6) and (7) into equation (3) we get the result in equation (2).

Note that $p(\theta)$ does not have to integrate to 1 for the A/R method to work: the normalizing factor necessary for $p(\theta)$ to be a valid probability density, $\int_{-\infty}^{+\infty} p(\theta)d\theta$, can be absorbed into the constant a. However since we need to know an actual value of a for the A/R method to work, we must at least have an estimate of the upper bound on this normalizing factor to determine an appropriate constant a satisfying inequality (1). Note how the normalizing factor ends up multiplying a:

$$\sup_{\theta \in C} \frac{p(\theta)}{\int_{-\infty}^{+\infty} p(\theta')d\theta' g(\theta)} \leq a \Longrightarrow \sup_{\theta \in C} \frac{p(\theta)}{g(\theta)} \leq a \int_{-\infty}^{+\infty} p(\theta')d\theta'$$

So that even if we don't know the normalizing factor but can guess an upper bound for it, then we can still do A/R sampling to draw from the density computed by dividing p by its normalizing factor. We can always use a very large value of a, but the cost is a low probability of acceptance (probability of stopping) so the A/R sampling method can be very inefficient and time consuming. This is why Markov Chain Monte Carlo methods such as Gibbs sampling and the Metropolis-Hastings methods are typically used instead of A/R in order to simulate draws from a posterior density in Bayesian econometrics.