MAT 653: Statistical Simulation

Instructor: Dr. Wei Li Scribe: Kyle Beiter

Oct 7th , 2021

The Accept-Reject Method

Motivation:

Suppose we wish to generate random variables following a distribution f, however we can only generate random variables from another distribution g.

Assumptions:

i. For all x, if $f(x) > 0$ then $g(x) > 0$, or equivalently $supp(f) \subseteq supp(g)$
ii. For all x, $f(x) \leq Mg(x)$, for some M and $M \geq 1$

Algorithm to generate $X \sim f$:

1. Generate $Y \sim g$ and $U \sim unif(0,1)$ independently
2. Accept $X = Y$ if $U \leq \frac{f(Y)}{Mg(Y)}$, otherwise repeat 1.

Proof idea: Consider the cumulative distribution function as $P(Y \leq x \mid U \leq \frac{f(Y)}{Mg(Y)})$ and show this is equal to $P(X \leq x)$ by first rewriting using formula for conditional probability, then direct computation using the integral over the correct ranges to calculate the relevant probabilities.

Properties:

- When f, g are probability density functions and $M \geq 1$ then $P(U \leq \frac{f(Y)}{Mg(Y)}) = \frac{1}{M}$, so the acceptance probability is $\frac{1}{M}$, as such we should choose the smallest M possible in order to minimize the number of draws that are rejected. $M = max(f(x)/g(x))$ will work.
- When only the kernels of some probability density functions$\tilde{f}, \tilde{g}$ are known (but still we can draw from g), then the same method can still be used to generate from the actual f by applying the algorithm with $\tilde{M} > 0$ when $\tilde{f} \leq \tilde{M}\tilde{g}$.
- Drawback of Accept-Reject method is that when proposals are rejected, there is no use for the generated values, wasting computational time.

The Fundamental Theorem of Simulation

If f is a probability density function in $\mathbb{R}^d$ (supported on $\mathcal{X}$), then we can write $f(x) = \int_0^{f(x)} 1du$ and furthermore 1 is a pdf of the uniform random variable. Hence we can consider f as the marginal distribution of X for the joint distribution $(X, U) \sim \mathrm{unif}(\{(x, u) : 0 < u < f(x), x \in \mathcal{X}\})$.

Then being able to generate $X \sim f$ is equivalent to being able to generate $(X, U) \sim \text{unif}(\{(\text{x}, \text{u}) : 0 < \text{u} < \text{f}(\text{x})\})$ where we then collect the generated X's. So this is essentially the same as sampling uniformly from the area under the curve of f.

Questions:

How do we sample uniformly from the area under the curve of f, i.e., generate $(X, U) \sim \text{unif}(\{(\text{x}, \text{u}) : 0 < \text{u} < \text{f}(\text{x})\})$?

1. Can we generate $X \sim f$ then generate $U \mid f(X) = f(x) \sim \text{unif}(0, f(x))$? This would indeed generate $(X, U) \sim \text{unif}(\{(\text{x}, \text{u}) : 0 < \text{u} < \text{f}(\text{x})\})$ (can prove it). But the answer is no, since the purpose of this simulation is to make simulating f simpler, so simulating from f directly is pointless.
2. Can we generate U from its marginal distribution and then generate $f \mid U$? No, this would require knowing how to generate the marginal distribution of U, and generate f given U, both may be not feasible.

What We Can Do

Suppose $f(\cdot) \leq Mg(\cdot)$ where f, g are two probability distribution functions, so we can see that $\text{supp}(f) \subseteq \text{supp}(g)$. Also suppose we can draw from some larger set $(Y, U) \sim \text{unif}\{(y, u) : 0 < u < g(y)\}$. Then iteratively draw from this distribution and keep the pairs of (y, u)'s such that y that satisfy $u < f(y)$. Then we collect these y's–the samples we keep will satisfy $y \sim f$.

Corollary

If $X \sim f$ and we have another distribution g so that $f(\cdot) \leq Mg(\cdot)$, for some $M \geq 1$ and $f, g : \mathbb{R}^d \to \mathbb{R}$. Then to simulate f it is sufficient to generate:

- $Y \sim g$ and $U \mid Y = y \sim \text{unif}(0, Mg(y))$, collect $\{(y, u)\}$
- Keep those pairs (y, u) that satisfy $u < f(y)$, call these qualified y as x

Then each $x \sim f$.

Proof:

Intuitive argument

The **first step** of the corollary creates a uniform distribution on $E = \{(y, u) : 0 < u < Mg(y)\} \subseteq \mathbb{R}^{d+1}$ (can prove it). Further, the uniform distribution on E gives another uniform distribution when restricted onto a subset of E (intuitively correct, but why?). Note that $\{(x, u) : 0 < u < f(x)\}$ is a subset of E. The **second step** of the corollary practically enforces this restriction of E onto $\{(x, u) : 0 < u < f(x)\}$, thus retaining samples that are essentially uniform draws on $\{(x, u) : 0 < u < f(x)\}$. Finally, invoking the Fundamental Theorem of Simulation, the uniform drawing on the restricted set $\{(x, u) : 0 < u < f(x)\}$ has a marginal pdf for x which is the target pdf f, giving the conclusion of the corollary.

Another view of Accept-rejct method from the Fundamental Theorem of Simulation

The above corollary is equivalent to the Accept-reject method. The algorithm also works for $x \propto \tilde{f}$ where $\tilde{f}$ is the kernel function of some pdf f.

Example:

Let $f(x) = x^2$ on $\{x \in [0, 3^{1/3}]\}$. Then choose the rectangle to be $[0, 3^{1/3}] \times [0, 3^{2/3}]$ as this is the smallest suitable rectangle. Randomly generate points in this rectangle, which can be done by independently simulating (y, u) with $y \sim \text{unif}([0, 3^{1/3}]), v \sim \text{unif}([0, 3^{2/3}])$ and accept $y \sim f$ if $f(y) > u$.