Manual for Random Variable Simulation

We consider the problem of simulating a random variable X with some density function

1 Mixture Method

The mixture method consists in representing the target distribution as a mixture of simpler distributions. There are two variants:

Mixture of Multiple Distributions

Suppose we have a mixture of n distributions

$$f_1(x), f_2(x), \ldots, f_n(x)$$

with corresponding probabilistic weights

$$p_1, p_2, \ldots, p_n, \quad \sum_{i=1}^{n} p_i = 1.$$

1. Method via Uniform Distribution

The idea is to use a uniform random variable on the interval $[0, 1]$ to select a component.

Algorithm:

1. Generate $U \sim \text{Uniform}(0, 1)$.
2. Find the index k such that
$$\sum_{i=1}^{k-1} p_i < U \leq \sum_{i=1}^{k} p_i.$$
3. Choose $X \sim f_k(x)$.

In other words, the interval $[0, 1]$ is divided into segments of length p_i, and the uniform random number U determines the segment and thus the component.

Remark: Multinomial / Generalized Bernoulli Approach

For n components, one can also use a categorical (multinomial with one trial) random variable:

1. Generate $B \sim \text{Categorical}(p_1, \ldots, p_n)$.
2. If $B = k$, select $X \sim f_k(x)$.

This is a direct generalization of the Bernoulli method used for two components.

Notes

- Both methods are equivalent and valid for any number of components.
- For a large number of components, it is convenient to precompute the cumulative sums of weights for fast component selection using a uniform number.
- If all weights are equal ($p_i = 1/n$), selecting a component reduces to generating a uniform discrete random variable on $\{1, 2, \ldots, n\}$.

1.1 Mixture via Bernoulli Random Variable

If we can decompose the density into a mixture of simpler densities:

$$f(x) = pf_1(x) + (1-p)f_2(x),$$

where f_1, f_2 are densities easy to sample from, then:

1. Generate $B \sim \text{Bernoulli}(p)$
2. If $B = 1$, sample $X \sim f_1(x)$, else sample $X \sim f_2(x)$

For our example, a simple choice can be $f_1(x)$ uniform on $[-1, 1]$ and $f_2(x)$ uniform on $[-2, -1] \cup [1, 2]$ with appropriate weights.

2 Rejection Sampling Method (Acceptance-Rejection)

Suppose we want to generate samples from an arbitrary probability density $f(x)$ that may be difficult to sample from directly.

Algorithm

The rejection sampling method relies on a majorizing density $g(x)$ such that

$$f(x) \leq Mg(x) \quad \text{for all } x,$$

where $M \geq 1$ is a constant.

Algorithm:

1. Choose a majorizing density $g(x)$ that is easy to sample from.
2. Compute the constant
$$M = \sup_x \frac{f(x)}{g(x)}.$$
3. Repeat until the desired number of samples is obtained:
   (a) Generate $X \sim g(x)$
   (b) Generate $U \sim \text{Uniform}(0, 1)$
   (c) Accept X if
   $$U \leq \frac{f(X)}{Mg(X)}$$
   otherwise reject and repeat.

Notes

- The efficiency of the method depends on the choice of the majorizing function $g(x)$ and the constant M. A tighter bound (smaller M) results in fewer rejections.
- For multi-dimensional or complex distributions, $g(x)$ is often chosen as a simpler distribution that "envelops" $f(x)$.
- The accepted samples are distributed exactly according to $f(x)$.

Example
Density function proportional to

$$f(x) \propto a(1 - x^4), \quad x \in [-2, 2].$$

2.1 Algorithm

1. Choose a majorizing density $g(x)$ easy to sample from (for example, uniform $g(x) = 1/4$ on $[-2, 2]$)
2. Compute the constant $M = \max_{x\in[-2,2]} \frac{f(x)}{g(x)}$
3. Repeat until desired number of samples obtained:
   (a) Sample $X \sim g(x)$
   (b) Generate $U \sim \text{Uniform}(0, 1)$
   (c) Accept X if $U \leq \frac{f(X)}{Mg(X)}$, otherwise reject