MTH 511a - 2020: Lecture 6

Instructor: Dootika Vats

1 Generating continuous random variables

We will discuss three methods for generating continuous random variables:

1. Inverse transform
2. The accept-reject method
3. Ratio of uniforms

1.1 Inverse transform

The principles of the inverse transform method for discrete distributions, apply similarly to continuous random variables.

Consider a random variable X with probability density function $f(x)$ so that $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x) = 1$ and distribution function is

$$F(x) = \int_{-\infty}^{x} f(x)\, dx\,.$$

The following theorem will be the foundation for the inverse transform method.

Theorem 1. *Let $U \sim U[0,1]$. For any continuous distribution F, a random variable $X = F^{-1}(U)$ has distribution F.*

Proof. Let F_X be the distribution function of $X = F^{-1}(U)$. Then,

$$\begin{aligned} F_X(x) &= \Pr(X \leq x) \\ &= \Pr(F^{-1}(U) \leq x) \\ &= \Pr(F(F^{-1}(U)) \leq F(x)) \\ &= \Pr(U \leq F(x)) \\ &= F(x)\,. \end{aligned}$$

□

Example 1. **Exponential**(1)**:** For the Exponential(1) distribution, the cdf is$F(x) = 1 - e^{-x}$. Thus,

$$F^{-1}(u) = -\log(1-u)\,.$$

To generate $X \sim \text{Exp}(1)$ we can thus use the following algorithm:

Algorithm 1 Exponential(1) Inverse transform

1: Generate $U \sim U[0, 1]$

2: Set $X = -\log(1 - U) \sim \text{Exp}(1)$

Example 2. **Cauchy distribution:** Cauchy distribution has pdf

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

and

$$u = F(x) = \int_{\infty}^{x} f(y)dy = \frac{1}{\pi}\arctan(x) + \frac{1}{2}\,.$$

So, $F^{-1}(u) = \tan(\pi(u - .5))$.

Algorithm 2 Cauchy distribution

1: Generate $U \sim U[0, 1]$

2: Set $X = \tan(\pi(U - .5) \sim \text{Cauchy}$

Example 3. **Gamma distribution:** The CDF of a Gamma(n, λ) distribution is

$$F(x) = \int_0^x \frac{\lambda e^{-\lambda y}(\lambda y)^{n-1}}{\Gamma(n)}\,.$$

Thus, we don't know the CDF in closed form and cannot find the inverse. This is an example where the inverse transform method cannot work.

1.2 Accept-reject method

Suppose we cannot generate from distribution function $F(x)$ with pdf $f(x)$, like the Gamma distribution example. We can use accept-reject in a similar way as the discrete case.

Draw samples from a distribution with density $g(x)$, and accept or reject it based on certain probabilities.

Let the support of F be X and choose a proposal distribution G with density $g(x)$ *whose support is larger or the same as the support of F.* That is, if $\mathcal{Y}$ is the support of G then, $\mathsf{X} \subseteq \mathcal{Y}$. If we can fine c such that

$$\sup_{x \in \mathsf{X}} \frac{f(x)}{g(x)} \leq c\,,$$

then an accept-reject sampler can be implemented.

Algorithm 3 Accept-reject for continuous random variables

1: Draw $U \sim U[0,1]$

2: Draw proposal $Y \sim G$

3: **if** $U \leq \dfrac{f(Y)}{c\,g(Y)}$ **then**

4: Return $X = Y$

5: **else**

6: Go to Step 1.

Theorem 2. *Algorithm 3 returns $X \sim F$.*

Proof. Consider any set B in X. We will show that

$$\Pr(X \in B) = F(B)\,.$$

First, we consider the probability of acceptance:

$$\begin{aligned}
\Pr(\text{accept}) &= \Pr\left(U \leq \frac{f(Y)}{cg(Y)}\right) \\
&= \mathrm{E}\left[I\left(U \leq \frac{f(Y)}{cg(Y)}\right)\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[I\left(U \leq \frac{f(Y)}{cg(Y)}\right) \mid Y\right]\right] \\
&= \mathrm{E}\left[\Pr\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{E}\left[\frac{f(Y)}{cg(Y)}\right] \\
&= \int_{\mathcal{Y}} \frac{f(y)}{cg(y)} g(y)dy \\
&= \frac{1}{c}\int_{\mathcal{Y}} f(y)dy \\
&= \frac{1}{c}\int_{\mathcal{X}} f(y)dy + \frac{1}{c}\int_{\mathcal{Y}/\mathsf{X}} f(y)dy \\
&= \frac{1}{c}\,.
\end{aligned}
$$

Now that we have this established, consider

$$
\begin{aligned}
\Pr(X \in B) &= \Pr(Y \in B \mid \text{accept}) \\
&= \frac{\Pr\left(Y \in B, U < \dfrac{f(Y)}{cg(Y)}\right)}{\Pr(\text{accept})} \\
&= c \cdot \mathrm{E}\left[\mathrm{E}\left[I\left(Y \in B, U < \frac{f(Y)}{cg(Y)}\right) \mid Y\right]\right] \\
&= c \cdot \mathrm{E}\left[I\left(Y \in B\right)\mathrm{E}\left[I\left(U < \frac{f(Y)}{cg(Y)}\right) \mid Y\right]\right] \\
&= c \cdot \mathrm{E}\left[I\left(Y \in B\right)\frac{f(Y)}{cg(Y)}\right] \\
&= c \cdot \int_B \frac{f(y)}{cg(y)} g(y)dy \\
&= \int_B f(y) \\
&= F(B)\,.
\end{aligned}
$$

□

From the proof, we know that $\Pr(accept) = 1/c$, and just like the discrete example, the number of attempts it takes to generate an acceptance is distributed Geometric$(1/c)$. Thus

$$
\text{Mean number of loops for an acceptance is} = c\,.
$$

1.3 Questions to think about

- Can we use the inverse transform method to generate sample from a normal distribution?
- In A-R, do we want c to be large or small?
- Can we always find such a c?