Simulation of Random Variables

In the following, we denote by U a random variable uniformly distributed on $[0,1]$.

We assume first that we know how to simulate U (see Appendix).

(a) Simulation of Bernoulli distribution of parameter p

Fix a number $p \in (0,1)$. Consider the random variable

$$X = 1_{(p,1)}(U) = \begin{cases} 0 & \text{if } U \leq p \\ 1 & \text{if } U > p \end{cases}.$$

Question: What is the distribution of X? Draw the CDF of X.

Answer: Let us compute the CDF (cumulative distribution function) of X:
Since $X \in [0,1]$, if $x < 0$, then $F_X(x) = 0$ and if $x \geq 1$, then $F_X(x) = 1$.
Now, let $x \in [0,1)$. One has

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(1_{(p,1)}(U) \leq x) = \mathbb{P}(U \leq p) = p,$$

the last equality comes from the fact that for U uniform on $[0,1]$, its CDF satisfies:

$$\forall x \in [0,1], \quad F_U(x) = x.$$

Conclusion: It follows that $\mathbb{P}(X = 0) = p$ and $\mathbb{P}(X = 1) = 1 - p$. We conclude that $X \sim \mathcal{B}(p)$ is a **Bernoulli distribution** of parameter p.

$\longrightarrow$ This process simulates **coin tossing** ("heads or tails").

(b) Simulation of uniform distribution on $\{1,2,3,4,5,6\}$

Consider the random variable

$$X = \sum_{i=1}^{6} i 1_{\left(\frac{i-1}{6}, \frac{i}{6}\right)} = \begin{cases} 1 & \text{if } 0 < U < \frac{1}{6} \\ 2 & \text{if } \frac{1}{6} < U < \frac{2}{6} \\ \vdots & \\ 6 & \text{if } \frac{5}{6} < U < 1 \end{cases}.$$

Question: What is the distribution of X? Draw the CDF of X.

Answer: Since $X \in \{1, 2, \ldots, 6\}$, if $k \notin \{1, 2, \ldots, 6\}$, then $\mathbb{P}(X = k) = 0$. Now, let $k \in \{1, 2, \ldots, 6\}$, then

$$\mathbb{P}(X = k) = \mathbb{P}\left(\frac{k-1}{6} < U < \frac{k}{6}\right) = \frac{1}{6}.$$

Conclusion: We conclude that $X \sim \text{Unif}\{1, 2, 3, 4, 5, 6\}$ is a **uniform distribution** on $\{1, 2, 3, 4, 5, 6\}$.

$\longrightarrow$ This process simulates **die rolling**.

(c) Simulation of distribution with bijective CDF

Let Y be a random variable such that F_Y is invertible (F_Y^{-1} exists). Consider the random variable

$$X = F_Y^{-1}(U).$$

Question: What is the distribution of X?

Answer: Let $x \in \mathbb{R}$. One has

$$\mathbb{P}(X \leq x) = \mathbb{P}(F_Y^{-1}(U) \leq x) = \mathbb{P}(U \leq F_Y(x)) = F_Y(x).$$

Conclusion: We conclude that $F_X = F_Y$, hence X and Y have same distribution.

- Example: Simulation of Cauchy distribution

Let Y be a standard Cauchy distribution, that is

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

The CDF of Y is

$$F_Y(x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+t^2}\, dt = \frac{1}{\pi}\left(\arctan(x) + \frac{\pi}{2}\right).$$

One has

$$F_Y(x) = u \Longleftrightarrow x = \tan\left(\pi(u - \frac{1}{2})\right).$$

Conclusion: Consider $X = \tan\left(\pi(U - \frac{1}{2})\right)$, then X has a Cauchy distribution.

(d) Simulation of arbitrary distribution (from a uniform on $[0, 1]$)

Let Y be a random variable. Let us define the generalized inverse of F_Y by

$$F_Y^{-1}(u) = \inf\{x \in \mathbb{R} : F_Y(x) > u\}, \quad u \in [0, 1].$$

Consider the random variable

$$X = F_Y^{-1}(U).$$

Question: What is the distribution of X?

Answer:

Lemma: For $x \in \mathbb{R}$ and $u \in [0, 1]$, one has

$$F_Y(x) > u \Longleftrightarrow x \geq F_Y^{-1}(u).$$

Proof: Exercise.

Let $x \in \mathbb{R}$. From Lemma above, one has

$$\mathbb{P}(X \leq x) = \mathbb{P}(F_Y^{-1}(U) \leq x) = \mathbb{P}(U < F_Y(x)) = F_Y(x).$$

Conclusion: We conclude that $F_X = F_Y$, hence X and Y have same distribution.

Consequence: To simulate an arbitrary random variable with CDF F, perform the following algorithm:

1. $\longrightarrow$ Compute F^{-1}.
2. $\longrightarrow$ Simulate U uniform on $[0, 1]$.
3. $\longrightarrow$ Output $X = F^{-1}(U)$.

From the above analysis, X is a random variable with CDF F.

Appendix. How to simulate a uniform random variable on $[0, 1]$?

It is **impossible** in practice to simulate "truly" random numbers in $[0, 1]$, as one would need to manipulate "infinity".

In practice, we use "pseudo-random numbers". Most random number generators start with an initial value X_0, called the seed, and then recursively compute values by specifying positive integers a, c and m, and then letting for $n \geq 0$,

$$X_{n+1} = (aX_n + c) \text{ modulo } m.$$

Thus each X_n is either $0, 1, \ldots, m-1$ and the quantity $\frac{X_n}{m}$ is taken as an approximation to a uniform random variable on $[0, 1]$. It can be shown that subject to suitable choices for a, c, m, the preceding gives rise to a sequence of numbers that looks as if it were generated from independent random variables uniform on $[0, 1]$.