Simulating CDFs on a Computer

08/09/2018

We shall look at how any CDF can be simulated on a computer. Just to break the surprise, any CDF can be obtained from a uniform $[0,1]$ distribution. There are two questions that arise in mind:

a) What's so special about uniform $[0,1]$? Why not uniform $[1,2]$ or $[1,100]$ or $[a,b]$ for any $-\infty < a < b < \infty$?

b) If the computer can generate any CDF from uniform $[0,1]$, how does it generate uniform $[0,1]$ in the first place?

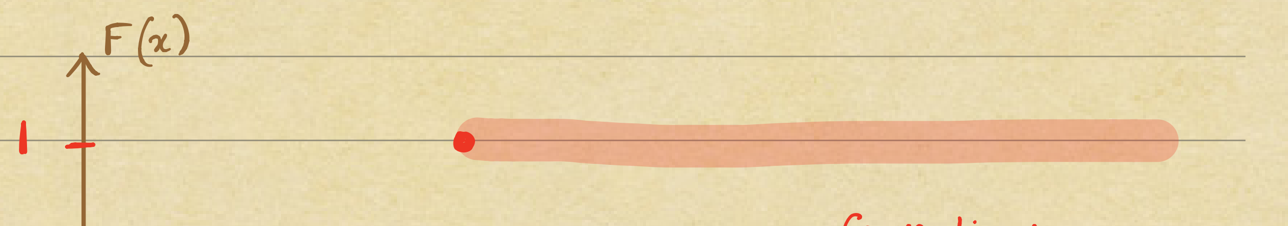
We shall first take up question a) above and answer it. The reason why only uniform $[0,1]$ distribution can be used as a starting point for generating CDFs is because any CDF $F = F(x)$, $x \in \mathbb{R}$, is a function of the form

$$F : \mathbb{R} \to [0,1],$$

always taking values in the interval $[0,1]$.

Let us look at some examples.

<u>Example 1</u>: Consider the CDF of a $Ber(3/4)$ distribution, whose sketch is as follows:



Mathematically, we can write

$$F(x) = \begin{cases} 0, & \text{if } x<0 \\ 1/4, & \text{if } 0 \leq x < 1 \\ 1, & \text{if } x \geq 1, \end{cases}$$

and if $(\Omega, \mathcal{F}, P)$ is a measurable space, with $X : \Omega \to \mathbb{R}$ being an $\mathcal{F}$-measurable random variable whose CDF is as above, then

$$P(X=0) = 1/4 = 1 - P(X=1).$$

<u>Question</u> : How do we generate F starting from uniform $[0,1]$ distribution ?

<u>Answer</u> : Consider a probability space $(\Omega, \mathcal{F}, P)$, and let

$$U : \Omega \to \mathbb{R}$$

be an $\mathcal{F}$-measurable random variable distributed according to uniform $[0,1]$ distribution. Let us define a new $\mathcal{F}$-measurable random variable X as follows :

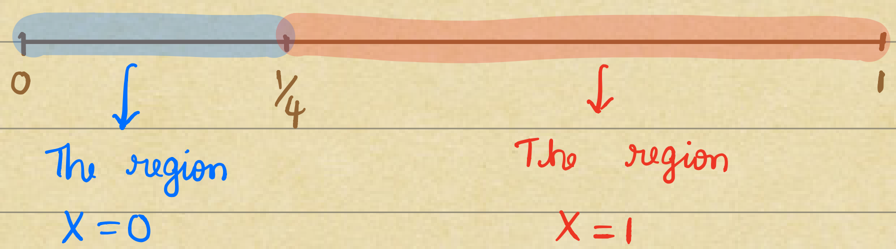
$$\forall\, \omega \in \Omega, \quad X(\omega) = \begin{cases} 0, & \text{if } U(\omega) \leq 1/4 \\ 1, & \text{if } U(\omega) > 1/4 . \end{cases}$$

Then, we see that $P(X=0) = P\left(U \leq \frac{1}{4}\right) = \frac{1}{4}$

$$\begin{aligned} P(X=1) &= P\left(U > \tfrac{1}{4}\right) \\ &= 1 - P\left(U \leq \tfrac{1}{4}\right) \\ &= \tfrac{3}{4}, \end{aligned}$$

hence giving us the CDF we desired to generate.

In other words, what we did was to partition the interval $[0,1]$ as follows:



<u>Example 2</u> :

<u>Question</u> : How do we generate a Binomial distribution with parameters $n=2$ and $p=\frac{1}{4}$ starting from a uniform distribution on $[0,1]$?

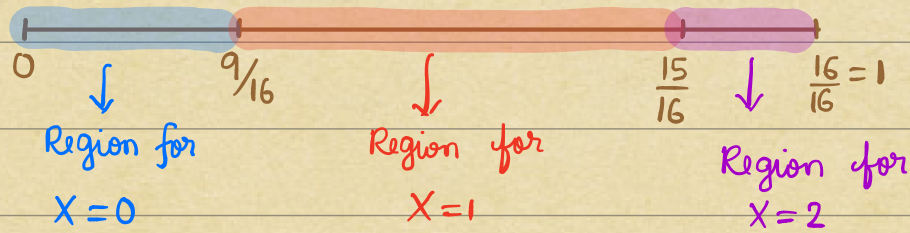
<u>Answer</u> : We want to generate the CDF of $\text{Bin}\left(2, \frac{1}{4}\right)$ distribution. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $X: \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable rv such that X follows $\text{Bin}\left(2, \frac{1}{4}\right)$ distribution, i.e.,

$$P(X=0) = {}^2C_0 \cdot \left(\tfrac{1}{4}\right)^0 \left(\tfrac{3}{4}\right)^{2-0} = \left(\tfrac{3}{4}\right)^2 = \frac{9}{16}$$

$$P(X=1) = {}^2C_1 \cdot (1/4)^1 \cdot (3/4)^{2-1} = 2 \cdot 1/4 \cdot 3/4 = \frac{6}{16}$$

$$P(X=2) = {}^2C_2 \cdot (1/4)^2 \cdot (3/4)^{2-2} = (1/4)^2 = 1/16.$$

Let $U : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable rv whose distribution is uniform $[0,1]$. Then, if we create a partition of the interval $[0,1]$ of the form



,

then we see that we get the desired $Bin(2, 1/4)$ CDF. Notice that the above partition means that for each $\omega \in \Omega$,

$$X(\omega) = \begin{cases} 0, & \text{if } 0 \leq U(\omega) \leq 9/16 \\ 1, & \text{if } 9/16 < U(\omega) \leq 15/16 \\ 2, & \text{if } 15/16 < U(\omega) \leq 1, \end{cases}$$

whence we get

$$P(X=0) = P\left(0 \leq U \leq 9/16\right) = 9/16$$

$$\begin{aligned} P(X=1) &= P\left(9/16 < U \leq 15/16\right) \\ &= P\left(U \leq 15/16\right) - P\left(U \leq 9/16\right) \\ &= \frac{15}{16} - \frac{9}{16} = \frac{6}{16} \end{aligned}$$

$$P(X=2) = P(15/16 < U \leq 1)$$
$$= P(U \leq 1) - P(U \leq 15/16)$$
$$= 1 - \frac{15}{16} = \frac{1}{16}.$$

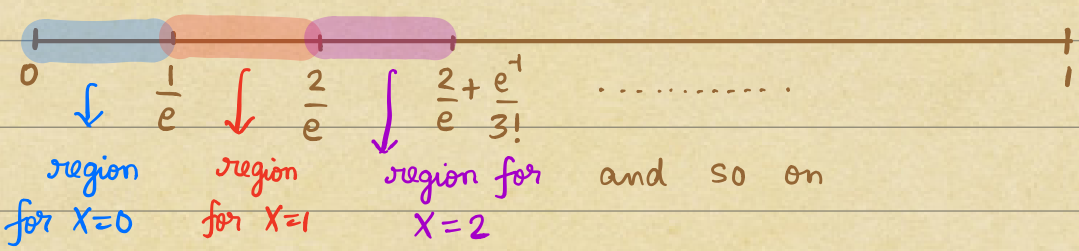
Example 3 :

Question : How do we generate a Poisson (1) distribution from a uniform [0,1] distribution ?

Answer : Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $X : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable rv with Poi(1) distribution. Then, we have

$$P(X=k) = \frac{e^{-1} \cdot 1^k}{k!} = \frac{e^{-1}}{k!}, \quad k = 0, 1, 2, \ldots$$

Thus, if we create a partition of the interval [0,1] of the form



then we see that we have generated the Poisson (1) CDF.

Since we know that

$$\sum_{k=0}^{\infty} e^{-1} \cdot \frac{1^k}{k!} = 1,$$

we can be sure that the sum of the lengths of the individual partitions will be equal to 1. What we did above was to assign for each $\omega \in \Omega$

$$X(\omega) = \begin{cases} 0, & \text{if } 0 \leq U(\omega) \leq 1/e \\ 1, & \text{if } 1/e < U(\omega) \leq 2/e \\ 2, & \text{if } 2/e < U(\omega) \leq 2/e + e^{-1}/3! \\ & \text{and so on,} \end{cases}$$

where $U : \Omega \to \mathbb{R}$ is an $\mathcal{F}$-measurable rv with uniform $[0,1]$ distribution.

This idea of partitioning works well for discrete random variables. The length of each sub-interval in the partition is equal to the probability of the random variable X taking a certain value. For example,

1. In the first example of Ber(3/4) distribution, the length of the first sub-interval is equal to $P(X=0)$.
2. In the second example of Bin$(2, 1/4)$, the length of the second sub-interval is $P(X=1)$.
3. In the third example of Poisson(1), the length of

the third sub-interval is $P(X=2)$.

However, if X is a continuous random variable, then we know that

$$P(X=x) = 0 \quad \forall\ x \in \mathbb{R}.$$

In such a case, the idea of partitioning doesn't seem to work (since in this case, the length of each sub-interval is zero).

How do we generate CDFs of continuous random variables from uniform $[0,1]$ distribution? Is there a general method to generate any CDF starting from uniform $[0,1]$ distribution that works even for discrete random variables?

The answer is yes! We shall now see this general method.

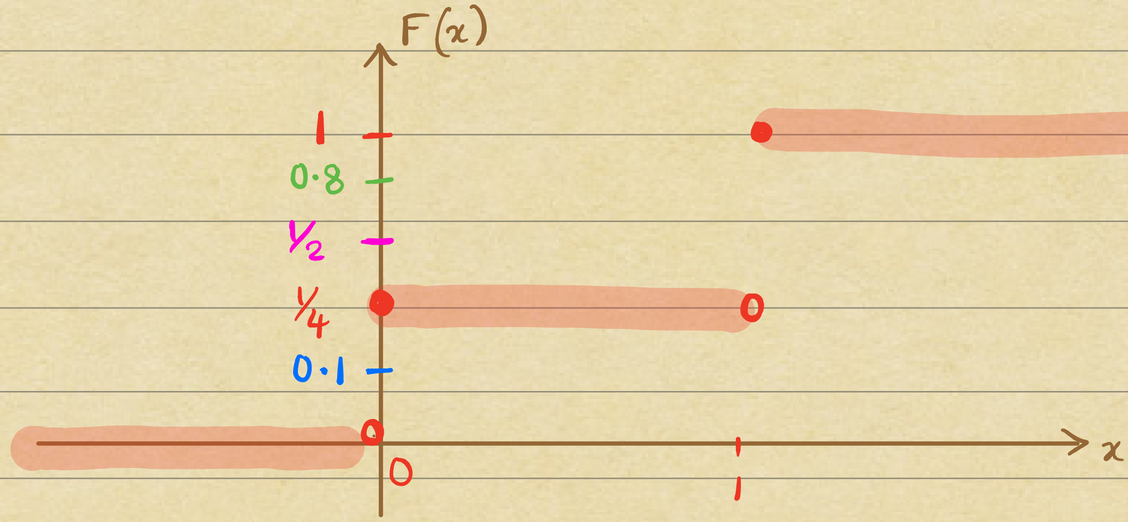
Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $U: \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable rv with uniform $[0,1]$ distribution. Suppose $F(x)$, $x \in \mathbb{R}$, is a CDF that we would like to simulate/generate starting from uniform $[0,1]$ distribution. Then, define a new $\mathcal{F}$-measurable rv $X: \Omega \to \mathbb{R}$ as:

$$\forall\ \omega \in \Omega, \quad X(\omega) = F^{-1}(U(\omega)) := \min\left\{ x \in \mathbb{R} : F(x) \geq U(\omega) \right\}.$$

Let's understand what $F^{-1}(\cdot)$ means by an example :

Example : 1

Take $F(\cdot)$ to be the CDF of a Ber(3/4) distribution. The sketch of F looks as follows :



Convention :

o – not included

• – included

Now,

$$F^{-1}(0.1) = \min\{x \in \mathbb{R} : F(x) \geq 0.1\} = 0.$$

$$F^{-1}(1/4) = \min\{x \in \mathbb{R} : F(x) \geq 1/4\} = 0$$

$$F^{-1}(1/2) = \min\{x \in \mathbb{R} : F(x) \geq 1/2\} = 1$$

$$F^{-1}(0.8) = \min\{x \in \mathbb{R} : F(x) \geq 0.8\} = 1$$

$$F^{-1}(1) = \min\{x \in \mathbb{R} : F(x) \geq 1\} = 1.$$

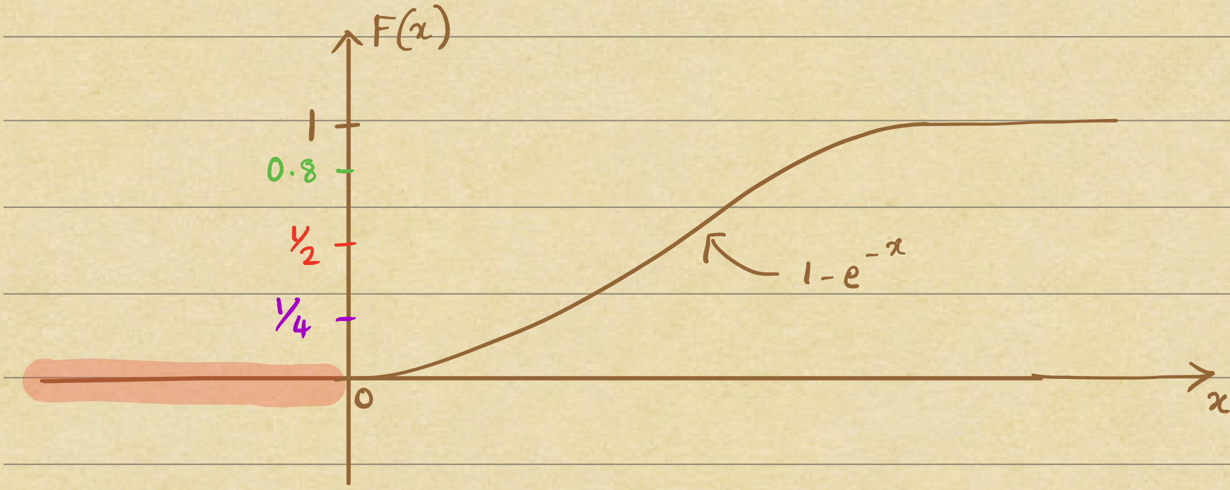
$$F^{-1}(0) = \min\{x \in \mathbb{R} : F(x) \geq 0\} = -\infty$$ (we take this to be $-\infty$).

Example : 2

Let $F(\cdot)$ be the CDF of an exponential distribution with parameter $\lambda = 1$. That is,

$$F(x) = \begin{cases} 1 - e^{-x}, & x \geq 0. \\ 0, & x < 0. \end{cases}$$

The sketch of F looks as follows:



In this case, we observe that F is a continuous function.

Here,

$$F^{-1}(0) = \min\{x \in \mathbb{R} : F(x) \geq 0\} = -\infty \text{ (we take this to be } -\infty\text{)}.$$

However, for $x \in [0, \infty)$, we notice that F is one-one and onto from $[0, \infty)$ to $[0, 1]$. Thus,

$$\begin{aligned} F^{-1}(1/4) &= \min\{x \in \mathbb{R} : F(x) \geq 1/4\} \\ &= \min\{x \in \mathbb{R} : 1 - e^{-x} \geq 1/4\} \\ &= \min\{x \in \mathbb{R} : e^{-x} \leq 3/4\} \\ &= \min\{x \in \mathbb{R} : x \geq \log_e 4/3\} \\ &= \log_e 4/3. \end{aligned}$$

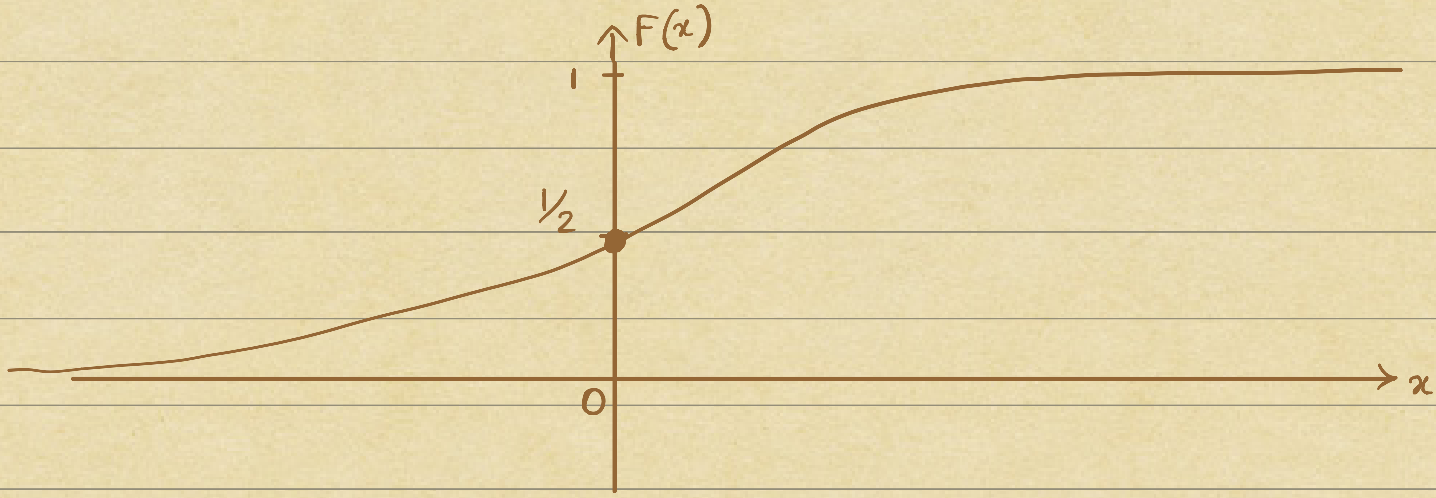
On similar lines, $F^{-1}(0.5)$ and $F^{-1}(0.8)$ can be computed.

We take $F^{-1}(1) = \infty$.

Example : 3

Let $F(x) = \begin{cases} \frac{1}{2} e^{x} , & x < 0 \\ 1 - \frac{1}{2} e^{-x} , & x \geq 0. \end{cases}$

The sketch of F looks as follows:



Notice that F is one-one and onto from $(-\infty, \infty)$ to $[0,1]$, and therefore the definition of $F^{-1}(\cdot)$ as above coincides with that of the inverse function of F.

Let us get back to our definition:

$$X(\omega) = F^{-1}(U(\omega))$$
$$= \min \{ x \in \mathbb{R} : F(x) \geq U(\omega) \}.$$

Now, in all of our earlier examples, X had the CDF that we wanted to simulate/generate. Is this the case even now? The answer is YES. We do not give a proof of this since it is quite involved.

We now come to the question of how a computer generates/simulates the CDF of uniform $[0,1]$ distribution. This is through what is known as "random number generators".

I will not write this here, but interested students may refer to the book by Scott Miller and Donald Childers titled "Probability and Random Processes", 2^{nd} edition. Look at chapter 12 of this book.