Week #2: Random Variables

Stochastic Methods Course

Notes by: Francisco Richter

25.02.2025

In many computational applications, pseudorandom number generators (PRNGs) produce integers in a large finite set, typically $\{0, 1, \dots, m-1\}$. To simulate a continuous uniform distribution, these integers are normalized to obtain numbers in the unit interval $[0, 1)$. Formally, if $\{X_n\}$ is a sequence produced by an LCG (or similar PRNG), we define

$$U_n = \frac{X_n}{m}.$$

Under ideal conditions, the sequence $\{U_n\}$ approximates a sequence of independent random variables uniformly distributed on $[0, 1)$.

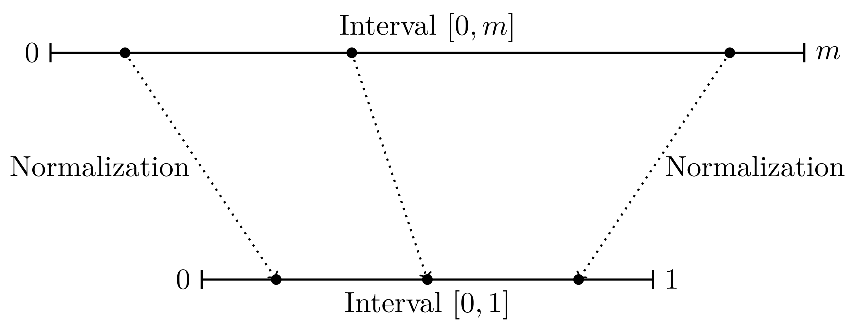

Figure 1: Normalization of PRNG outputs to the unit interval $[0, 1]$.

1 Mapping the Unit Interval to Discrete Outcome Spaces

Often the outcomes of an experiment belong to a discrete set. For instance:

- A fair die yields outcomes in $U_d = \{1, 2, 3, 4, 5, 6\}$.
- A coin toss yields outcomes in $U_c = \{H, T\}$.

A standard approach is to first generate a uniformly distributed number $U \in [0, 1)$ and then use a deterministic mapping to convert U into an outcome in the desired set. Such mappings are typically constructed using modular arithmetic or thresholding techniques.

Example 1.1 (Simulating a Coin Toss)**.** *Consider an LCG that generates integers uniformly in* $\{0, 1, \dots, 2^{32} - 1\}$. *Since exactly half of these values are less than* 2^{31}*, we define the outcome as follows:*

$$\textit{Outcome} = \begin{cases} H, & \text{if } X_n < 2^{31}, \\ T, & \text{if } X_n \geq 2^{31}. \end{cases}$$

Under the assumption of uniformity, this mapping yields $P(H) = P(T) = 0.5$.

2 A Rigorous Framework for Probability

Probability theory is founded on measure theory, which provides a precise mathematical description of uncertainty. The cornerstone of this framework is the concept of a *probability space*, defined as a triple $(\Omega, \mathcal{F}, P)$:

- Ω is the *sample space*—the set of all possible outcomes of a random experiment.
- $\mathcal{F}$ is a *σ-algebra* on Ω, meaning it is a collection of subsets of Ω (called *events*) that satisfies:
  - (i) $\Omega \in \mathcal{F}$,
  - (ii) If $E \in \mathcal{F}$, then its complement $E^c = \Omega \setminus E$ is also in $\mathcal{F}$,
  - (iii) $\mathcal{F}$ is closed under countable unions; that is, if $E_1, E_2, E_3, \dots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$.
- $P : \mathcal{F} \to [0,1]$ is a *probability measure* that assigns a number to each event in $\mathcal{F}$ with the properties:
  - (i) $P(\Omega) = 1$ (the probability that some outcome occurs is 1),
  - (ii) For any countable collection of disjoint events $E_1, E_2, \dots \in \mathcal{F}$, we have

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

A classic example is the unit interval $[0,1]$ with the *Lebesgue measure*. For any subinterval $[a,b] \subseteq [0,1]$, the measure is defined as

$$\mu([a,b]) = b - a.$$

Since $\mu([0,1]) = 1$, the Lebesgue measure naturally serves as a probability measure on $[0,1]$.

The *Borel σ-algebra* on $[0,1]$, denoted by $\mathcal{B}([0,1])$, is the collection of all sets that can be formed from open intervals by taking countable unions, intersections, and complements. This σ-algebra is crucial because it includes most sets encountered in practice, ensuring that the probability measure is defined on a rich collection of events.

Furthermore, if we have a measurable function (or mapping) $\phi : U \to [0,1]$ from another outcome space U, we can *induce* a probability measure on U by *pulling back* the Lebesgue measure. For any event $B \subseteq U$ (where B is measurable), the induced probability is defined as

$$P_U(B) = \mu(\phi(B)).$$

This construction guarantees that the essential properties of a probability measure—non-negativity, normalization, and countable additivity—are preserved.

3 Random Variables and Their Distributions

Once a probability measure P is established on a measurable space $(U, \mathcal{F})$, a *random variable* is defined as a measurable function mapping outcomes to real numbers.

Definition. Random Variable

Let $(U, \mathcal{F}, P)$ be a probability space. A function

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

is called a *random variable* if for every Borel set $B \subset \mathbb{R}$, the preimage $X^{-1}(B)$ is in $\mathcal{F}$. The set

$$R(X) = \{X(u) : u \in U\}$$

is called the *range* of X. If $R(X)$ is countable, X is said to be *discrete*.

Discrete Random Variables

For discrete random variables, the distribution is fully characterized by its probability mass function (pmf).

Definition. Probability Mass Function

For a discrete random variable X with range $R(X)$, the *probability mass function* f_X is defined by

$$f_X(x) = P\{X = x\}, \quad x \in R(X),$$

subject to:

1. $0 \le f_X(x) \le 1$ for all x,
2. $\displaystyle\sum_{x \in R(X)} f_X(x) = 1.$

Example 3.1 (Discrete Uniform Distribution)**.** *If X takes values in a finite set $\{a, a+1, \dots, b\}$ with equal probability, then for each $x \in \{a, a+1, \dots, b\}$,*

$$f_X(x) = \frac{1}{b - a + 1}.$$

Example 3.2 (Bernoulli Distribution)**.** *Let X represent a binary outcome such that*

$$X(u) = \begin{cases} 1, & \textit{if the outcome is a success,} \\ 0, & \textit{if the outcome is a failure.} \end{cases}$$

If $P(X = 1) = p$ and $P(X = 0) = 1 - p$, then

$$f_X(1) = p, \quad f_X(0) = 1 - p.$$

Example 3.3 (Binomial Distribution)**.** *If n independent Bernoulli trials are conducted with success probability p, and X denotes the number of successes, then*

$$f_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n.$$

Example 3.4 (Poisson Distribution)**.** *A random variable X is said to have a Poisson distribution with parameter $\lambda > 0$ if*

$$f_X(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots.$$

Continuous Random Variables

For continuous random variables, the distribution is described by a probability density function (pdf).

Definition. Probability Density Function

A continuous random variable X with range $R(X) \subset \mathbb{R}$ has a *probability density function* f_X satisfying

$$P(a \le X \le b) = \int_a^b f_X(x)\, dx, \quad \text{for any } [a, b] \subset R(X),$$

with the normalization condition

$$\int_{R(X)} f_X(x)\, dx = 1.$$

Example 3.5 (Exponential Distribution)**.** *If X follows an exponential distribution with rate parameter $\lambda > 0$, then its pdf is given by*

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0,$$

and $f_X(x) = 0$ for $x < 0$.

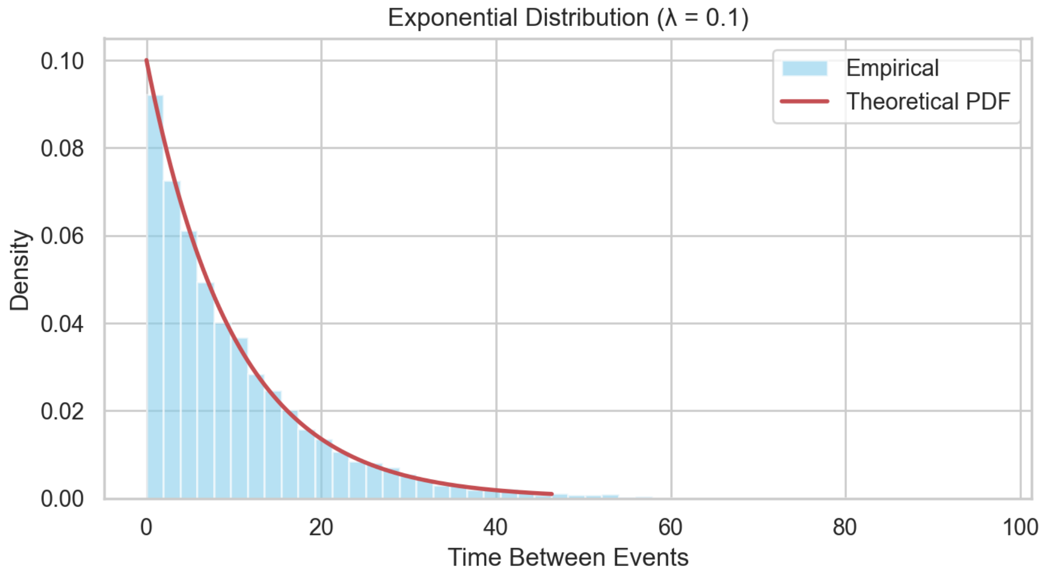

Figure 2: Exponential probability density function.

Example 3.6 (Normal Distribution)**.** *For a normal distribution with mean μ and variance σ^2, the pdf is*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}.$$

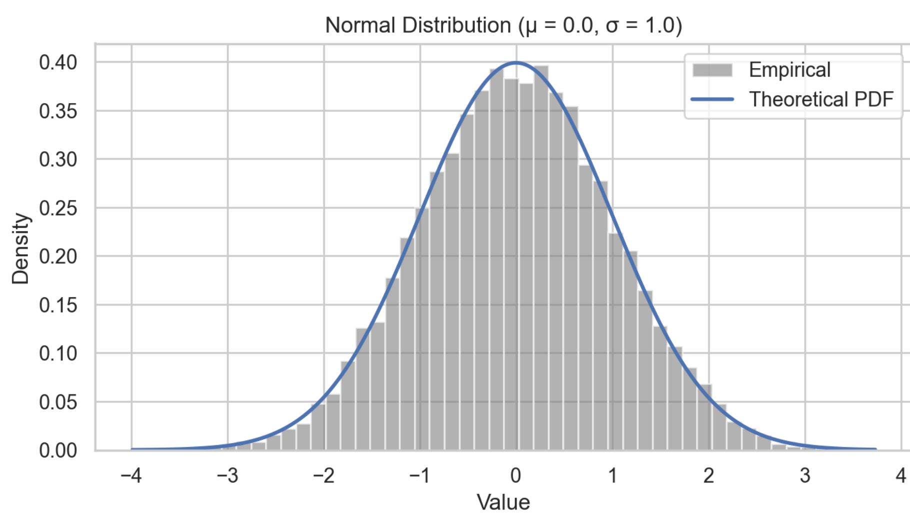

Figure 3: Normal probability density function.

4 The Cumulative Distribution Function

The cumulative distribution function (CDF) of a random variable X is defined as the probability that X takes a value less than or equal to x.

Definition. Cumulative Distribution Function

The *cumulative distribution function* of X is

$$F_X(x) = P\{X \leq x\}.$$

For a discrete random variable,

$$F_X(x) = \sum_{t \leq x} f_X(t),$$

and for a continuous random variable,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt.$$

Moreover, if the CDF $F_X(x)$ is invertible, one can generate samples from the distribution of X using the inverse transform method:

$$X = F_X^{-1}(U),$$

where U is uniformly distributed on $[0, 1]$.

5 Exercises

Exercise 1: Maximum of Five Uniform $(0, 1)$ Variables.

Let $Z_1, Z_2, \ldots, Z_5$ be independent random variables uniformly distributed on $(0, 1)$. Define

$$M = \max\{Z_1, Z_2, Z_3, Z_4, Z_5\}.$$

(a) Derive an expression for $P(M \leq x)$ for $0 \leq x \leq 1$.

(b) Differentiate $P(M \leq x)$ to obtain the pdf of M and interpret the result.

Exercise 2: Sum of Two Uniform $(0, 1)$ Variables.

Let W_1 and W_2 be independent random variables uniformly distributed on $(0, 1)$, and define

$$S = W_1 + W_2.$$

(a) Derive the probability density function of S.

(b) Derive the cumulative distribution function of S.

Exercise 3: Repeated Uniform $(0, 1)$ Picks Until Sum > 1.

Independently generate random numbers $V_1, V_2, \ldots$, each uniformly distributed on $(0, 1)$, until

$$V_1 + V_2 + \cdots + V_N > 1.$$

Let $X = N$ denote the number of picks required.

(a) Determine the pmf $p_X(n) = P(X = n)$.

(b) Determine the cdf $F_X(n) = P(X \leq n)$.

Exercise 4: From Binomial to Poisson.
Let $X_n \sim \text{Bin}(n, p)$ be a binomial random variable with parameters n and p. Define $\lambda = np$. Show that as $n \to \infty$ and $p \to 0$ in such a way that λ remains constant, the pmf of X_n converges to that of a Poisson random variable with mean λ.

Exercise 5: From Poisson to Exponential.
Consider a Poisson process with rate λ, where $N(t) \sim \text{Poisson}(\lambda t)$ denotes the number of events in time t. Let

$$T = \inf\{t > 0 : N(t) \geq 1\}$$

be the waiting time until the first event. Prove that

$$P(T > t) = e^{-\lambda t},$$

and conclude that $T \sim \text{Exp}(\lambda)$.

Exercise 6: Memoryless Property of the Exponential Distribution.
Let $X \sim \text{Exp}(\lambda)$. Prove that for any $s, t \geq 0$,

$$P(X > s + t \mid X > s) = P(X > t).$$