Lecture 01

Foundations of Randomness (Part I)

Probability & Statistics

Francisco Richter, Martina Boschi and Ernst Wit

Introduction

Our daily lives are filled with uncertainty, from the simple toss of a coin to the stock market's unpredictable rises and falls. This inherent unpredictability is often captured mathematically through random numbers and random variables. In many realms of science, engineering, and even entertainment, the ability to generate and harness randomness is invaluable. Random processes model phenomena as diverse as subatomic particle interactions, the growth of populations, and the strategies in a poker game.

1 Generating Randomness

In a world increasingly driven by computers — deterministic machines at their core — generating genuine randomness is challenging. Computers follow predefined instructions to produce predictable outcomes. So, how can such a machine create a random number or simulate the behavior of weather systems?

1.1 Random Numbers

The answer lies in algorithms that can mimic randomness, even if they aren't truly random. These algorithms are the heart of Random Number Generators (RNGs). By inputting an initial value or 'seed', they produce sequences of numbers that, for all practical purposes, seem random.

This 'pseudo-randomness' can be a desirable property. Imagine a scientist running a complex simulation of a galaxy. If something goes awry, it's ainvaluable to replay that simulation with the exact same sequence of 'random' events to debug and understand the problem. This reproducibility is only possible with pseudo-random numbers, not with truly random ones.

Before delving into how we generate these numbers, it's crucial to understand the nature of randomness and how we represent it mathematically.

Definition 1 (Random number) *A random number is an unpredictable value, generated independently from preceding or succeeding numbers. It lacks any discernible pattern or regularity, making it impossible to deduce without understanding the underlying random generation process.*

Moving from individual random numbers, it is essential to understand how we can generate a series of such numbers, leading us to the concept of a random number generator.

1.2 Random Number Generators (RNGs)

Definition 2 (Random number generator) *A Random Number Generator (RNG) is an algorithm that produces a sequence of numbers that lacks any pattern, i.e., appears random.*

More formally, an RNG is defined as a function:

$$R : S \to T$$

where:

- S *is the seed space , a finite set of initial states. An RNG is typically initialized with a value in* S*, known as the seed.*
- T *is the target space , typically the set of real numbers in the interval [0, 1) or a set of integer values.*

- *The function R maps each seed $s \in S$ to a target $t \in T$ in a manner that appears random.*

With the understanding of what random numbers and their generators are, we can now lay down some properties that a well-functioning random number generator should exhibit:

- **Unpredictability:** Without knowing the algorithm and seed, it should be impossible to predict future numbers.
- **Reproducibility:** Given the same seed, the RNG should produce the same sequence of numbers.
- **Representation of True Randomness:** The RNG should accurately represent true randomness, ensuring an equitable chance for all potential outcomes.
- **Long period:** The sequence of numbers should be long before repeating.
- **Efficiency:** The RNG should generate numbers quickly.

Whenever the RNG is such that these properties are fulfilled, then the target space can be conceived as a *aleatory universe*.

Definition 3 (Linear Congruential Generator) *The linear congruential generator generates a sequence of random numbers via the following linear recurrence relation:*

$$X_{n+1} = (aX_n + c) \mod m$$

where:

- *X_{n+1} is the next number in the sequence.*
- *X_n is the current number.*
- *a, c, and m are constants, known as the multiplier , increment , and modulus , respectively.*
- $\mod$ *denotes the modulus operation.*
- *The initial or seed value $X_0 = S$, is also required to start the sequence.*

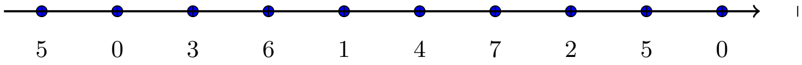

Figure 1: A sequence of realizations from a linear congruential generator (LCG). The x-axis represents the index in the sequence, and the y-coordinate of each point represents the value of the LCG at that index. The sequence is shown until it begins to repeat, including the first and second term of the repetition.

2 EXPLORING RANDOM VARIABLES

2.1 RANDOM VARIABLES

Definition 4 (Random Variable) *A random variable (RV) is a function that assigns a real number to a particular element of the aleatory universe U. A random variable X is defined as a function:*

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

which maps each possible outcome $u \in U$ to a real number $X(u) = x$. The set of all possible values of X, often denoted by $R(X)$ or simply the image of X, is a subset of the real numbers $\mathbb{R}$.

If $R(X)$ is a countable object, the RV is *discrete* .
Each random experiment may result in a collection of conceivable outcomes, named *sample space*, each of them corresponding to a value or a subset of values of the aleatory space. Let us consider one of the simplest random experiments, namely tossing a coin. If the coin is fair, we know that the probability of a head is equal to the probability of a tail and to 0.5 specifically. Given a RNG that produces uniform random numbers over its range, we can map the values below the half-period to `Head` (H) and those above to `Tail` (T). As an alternative approach, we can map even numbers to H and odd numbers to T. Experiment 1 is aimed at exploring in further detail this idea.

Experiment 1 (Probability and frequency) *Consider the Linear Congruential Generator (LCG), a type of RNG algorithm, given by the recurrence relation:*

$$X_{n+1} = (1664525 \times X_n + 1013904223) \mod 2^{32}$$

We consider the frequency approach to probability to determine the probability of either outcome (H or T). Given that the LCG produces uniform random numbers over its range, half of the numbers will be below $\frac{2^{32}}{2}$ and half will be above. Using this midpoint, numbers below this value can be mapped to `Head` *(H) and numbers above to* `Tail` *(T).*
Nevertheless, a RV assigns a real number to a particular element of the aleatory universe U. We claim that the generic variable $X_n = 0$ if we observe H and equals 1 in case of T. Therefore, as $n \to 2^{32}$:

$$P(u|X(u) = 1) = P(u|X(u) = 0) = 0.5$$

This ensures an almost equal probability for the LCG's output to represent either a H or a T for a sufficiently large number of trials.
The empirical validation of this theory can be observed in figure 2. As the sample size N increases, the proportions of H and T converge to the theoretical value of 0.5, reinforcing the uniformity and reliability of the LCG in mapping outcomes.

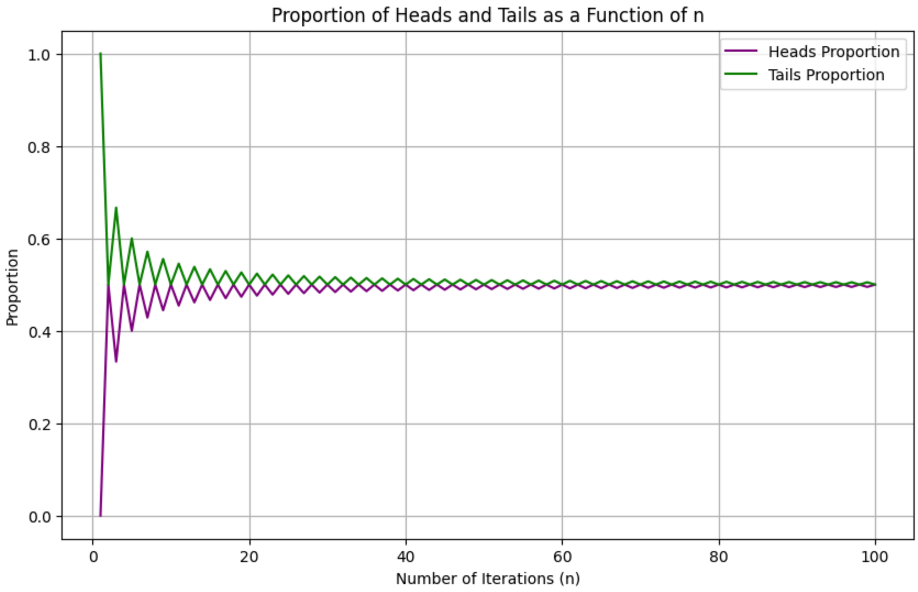

Figure 2: Empirical validation of the Head-Tail experiment.

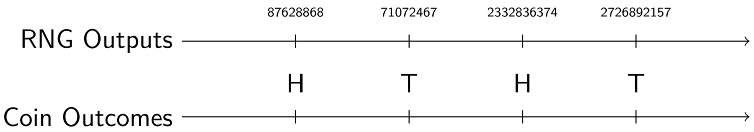

Figure 3: Mapping of LCG Outputs to Coin Toss Outcomes for First Four Iterations

2.2 PROBABILITY DISTRIBUTIONS

Definition 5 (Probability Mass Function) *The probability mass function (pmf) associated with a discrete random variable X is represented as a function $p_X : R(X) \to [0, 1]$, defined by*

$$p_X(x) = P(X = x) = P(u|X(u) = x)$$

i.e. the probability of observing an element of the aleatory universe U for which the variable assumes the value of interest. For every $x \in R(X)$, the following holds:

1. *For every outcome $x \in R(X)$, $0 \leq p_X(x) \leq 1$.*
2. *The sum of the probabilities for all possible outcomes is 1, i.e., $\sum_{x \in R(X)} p_X(x) = 1$.*

Let's consider the rolling dice process. Consider the scenario of rolling a fair six-sided dice. The sample space of this scenario is $\{1, 2, 3, 4, 5, 6\}$. Since the dice is fair, each outcome is equally likely. Therefore, we can divide the aleatory universe in 6 equal parts, each of them corresponding to one of the 6 potential outcomes of the random experiment. The pmf of this discrete random variable is thus:

$$p(X = x) = P(u|X(u) = x) = \begin{cases} \frac{1}{6} & \text{if } x \in \{1, 2, 3, 4, 5, 6\}, \\ 0 & \text{otherwise} \end{cases}$$

The pmf here characterizes a uniform distribution, each outcome from the dice roll has an equal chance of occurring, demonstrating the fairness of the dice.
One of the simplest forms of a discrete probability distribution is the discrete uniform distribution. As we just saw, if we let X represent the outcome of flipping a fair coin, X could take the values 0 and 1 (corresponding to each with probability $1/2$. Similarly, if X represents the outcome of rolling a fair six-sided die, X can take the values 1,2,3,4,5,6 each with probability $1/6$.

Example 1 (Discrete Uniform distribution) *Let X be a RV with outcomes equally likely in the set $\{a, a + 1, \ldots, b\}$. If X follows discrete uniform distribution, then the pmf can be defined as follows:*

$$p_X(k) = P(X = k) = P(u|X(u) = k) = \frac{1}{b - a + 1}, \quad \textit{for} \quad k = a, a + 1, \ldots, b.$$

The preceding examples primarily highlighted uniform distributions, where each outcome had an equal likelihood of occurrence. As we transition further, we will delve into some of the most renowned distributions that exhibit varied probabilities for different outcomes, breaking away from the uniformity principle.

Example 2 (Bernoulli Distribution) *Consider a binary experiment with outcomes that do not share the same probability. In this case the aleatory universe is divided in two equally sized components. The Bernoulli distribution provides a mathematical model for such situations.*
Let X be a RV describing the outcome of a binary experiment. X is said to follow a Bernoulli distribution if it takes on two possible outcomes: 1 (`success`) with probability p and 0 (`failure`) with probability $1 - p$. Formally, the pmf is defined as:

$$p_X(k) = P(X = k) = P(u|X(u) = k) = \begin{cases} p & \textit{if } k = 1, \\ 1 - p & \textit{if } k = 0. \end{cases}$$

Here, $0 \leq p \leq 1$ is the probability of success, and k can only take values 0 or 1.

The random variable describing the tossing of a fair coin can be also viewed as a particular case of Bernoulli RV, where the probability of `success` (`Tail`) is exactly equal to the probability of `failure` (`Head`).

Example 3 (Binomial Distribution) *Following the idea of the Bernoulli distribution, suppose we conduct a series of n Bernoulli trials. Each trial has two possible outcomes: `success` (with probability p) and `failure` (with probability $1 - p$). The probability of obtaining exactly k successes in these n trials is given by:*

$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

for $k = 0, 1, \ldots, n$. The random variable X representing the number of successes in n trials follows a binomial distribution.

Example 4 (Poisson Distribution) *A discrete random variable X is said to have a Poisson distribution with parameter $\lambda > 0$ (where λ is the average number of occurrences in a fixed interval or region) if its pmf is given by:*

$$p_X(k) = P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

for $k = 0, 1, 2, \ldots$.