**Statistics for Data Science - Hands-on 1**

**Random Number Generation**

**What are Random Numbers?**

* **Random numbers** are sequences of numbers generated from a stochastic process, meaning their outcomes are unpredictable. In contrast, a deterministic process has outcomes that are fully predictable given initial conditions.
  + Example of a stochastic process: Rolling a die.
* A random number is selected using methods that give **equal probability** to all numbers within a specified distribution. To qualify as random, numbers must be **independent**, meaning there is no correlation between successive numbers.
* **Applications**: Random numbers are essential in various fields, including:
  + **Cryptography**: Used to generate cryptographic keys.
  + **Encryption algorithms**: Ensure secure communications.

**Random Number Generators (RNGs)**

A **Random Number Generator (RNG)** is a hardware device or software algorithm that produces numbers taken from a limited or unlimited distribution.

There are two main types:

1. **True Random Number Generators (TRNGs)**:
   * TRNGs rely on **unpredictable physical processes** to generate randomness.
   * These processes include atmospheric noise, thermal noise, radioactive decay, and other naturally occurring phenomena that are fundamentally random.
2. **Pseudo-Random Number Generators (PRNGs)**:
   * PRNGs use **mathematical algorithms** to generate a sequence of numbers that appear random.
   * Unlike true random numbers, PRNGs are not truly random because they depend on an initial **seed value**, making the sequence repeatable if the same seed is used.

**Random Number Seed**

A **random seed** is a number (or vector of numbers) used to initialize a PRNG. It acts as the starting point for generating a sequence of random numbers. Using the same seed ensures that the sequence of random numbers generated will be the same, which is useful for reproducibility in simulations.

**Random Variates**

A **random variate** is a variable generated from uniformly distributed pseudo random numbers, commonly used in simulations. These random variates serve as inputs to simulation models, allowing for the modelling of various stochastic processes.

**Random Variate Generation**

**Random variate generation** is critical in simulation modelling and analysis. Its objective is to produce observations that exhibit the same stochastic properties as a given random variable.

* To generate random variates, uniformly distributed random numbers are first produced within the interval **[0,1]**.
* From these uniformly distributed random numbers, other distributions can be derived using transformation techniques.

**Methods for Random Variate Generation**

There are several methods to generate random variates:

1. **Inverse Transform Method**:
   * This method transforms uniformly distributed random numbers to follow a desired distribution by applying the inverse of the cumulative distribution function (CDF).
2. **Acceptance-Rejection Method**:
   * This method generates candidate random variates from a simple distribution and accepts them based on a probability criterion.
   * This method has been discussed in the notebook.
3. **Composite Method**:
   * Combines multiple simple methods to generate random variates from complex distributions.
4. **Translations and Simple Transforms**:
   * Applies mathematical transformations to uniformly distributed random numbers to generate random variates from specific distributions.