0 Introduction to Stochastic Processes

After introducing the basic elements we are going to work with, we can present the really interesting objects that are going to make use of the previously introduced building blocks: **stochastic process**. This chapter provides an introduction to stochastic processes and some of their properties.

0 Central Limit Theorem and Law of Large Numbers

Before actually defining stochastic processes, it is fundamental to introduce two key results in probability theory that are essential for understanding the behavior of stochastic processes: the **Central Limit Theorem** and the **Law of Large Numbers**. These theorems provide insights into how random variables behave when aggregated over time or across many trials.

0 Central Limit Theorem

Let us consider a sequence of random variables $X_1, X_2, X_3, \ldots$ and the **sum** of their first n elements. As we have seen in the previous chapter, this is again a random variable:

$$S_n = X_1 + X_2 + \ldots + X_n$$

We should already be familiar with the following facts:

- if all random variables X_i have a **common mean** $\mu = \mathbb{E}[\![X_i]\!]$, then the expected value of the sum is given by: $\mathbb{E}[\![S_n]\!]n\mu$

- if all random variables X_i are **independent** and have a **common variance** $\sigma^2 = \text{Var}\{X_i\}$, then the variance of the sum is given by: $\text{Var}\{S_n\} = n\sigma^2$

We would now like to investigate what we can say about the sequence of sums $S_1, S_2, S_3, \ldots$ as the number n of them grows. To work this out, let's consider the n standard normal random variables $X_i \overset{\text{i.i.d}}{\sim} N(0,1)$ for which we know that $\mathbb{E}[\![X_i]\!] = 0, \text{Var}\{X_i\} = 1$. We know that their sum S_n has expected value $\mathbb{E}[\![S_n]\!] = 0$ and variance $\text{Var}\{S_n\} = n$. The sum we have obtained is again a Normal random variable, but it's not standard anymore since its variance is not 1: $S_n \sim N(0,n)$.
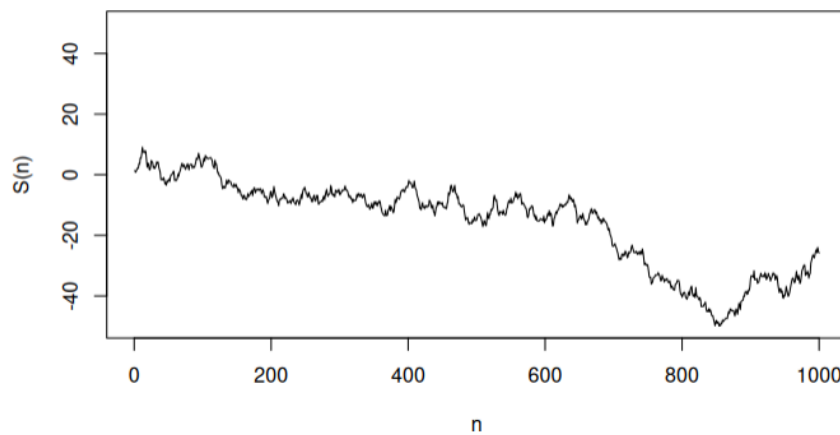

Figure 0.1: Sum of n i.i.d. standard normal random variables. $S_n \sim \mathcal{N}(0,n)$

Figure 0.1 shows the evolution of the random variable S_n as n increases: it is possible to notice that, the larger the n, the more the distribution of S_n spreads out.

Suppose we wanted to obtain a new value S_{n+1}, this would be given by the sum of S_n and a new independent standard normal random variable X_{n+1}. We know that S_n, X_{n+1} are independent, but the newly obtained sum S_{n+1} is not independent of S_n:

$$\mathbb{E}[\![S_{n+1} \mid S_n]\!] = \mathbb{E}[\![S_n]\!] + \mathbb{E}[\![X_{n+1}]\!] = S_n + 0 = S_n$$
$$\mathrm{Var}\{S_{n+1} \mid S_n\} = 1 \ \text{ since } \ S_n \ \text{ is now fixed}$$
$$\implies S_{n+1} \mid S_n \sim \mathcal{N}(S_n, 1)$$

In *time series analysis* this is called an **autoregressive process**. We discussed in some of the previous chapters that for sums of independent random variables we have that $S_{n+1} \mid S_n$ is independent of all the previous ones. This means we can write the **joint distribution** of $S_1, ..., S_n$ as follows:

$$f_{S_1, ..., S_n}(s_1, ..., s_n) = f_{S_1}(s_1) f_{S_2 \mid S_1}(s_2 \mid s_1) f_{S_3 \mid S_2}(s_3 \mid s_2) \, ...$$
$$= \mathcal{N}(0, 1) \cdot \mathcal{N}(s_1, 1) \cdot \mathcal{N}(s_2, 1) ... \mathcal{N}(s_{n-1}, 1)$$

0.1

This is not the same as having n independent random variables each with marginal distribution $\mathcal{N}(0, n)$. To better understand this, let's look at Figure 0.2 (left side) which shows some realizations of the sums of n i.i.d. standard normal random variables as the value of n increases. This is in strong contrast with what we have obtained in the right side of the figure, where we showcased the behavior of n independent random variables each with distribution $\mathcal{N}(0, n)$.
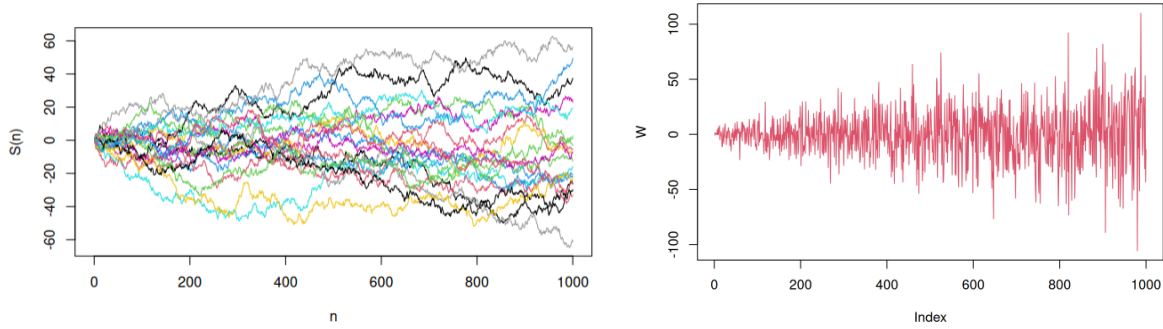

Figure 0.2: On the left: multiple realizations of the sum of n i.i.d. standard normal random variables. On the right: values obtained by simulating n a random variable $Y \sim \mathcal{N}(0, n)$ independently as the value of n increases.

It is evident that the left side of the figure represents a much more "tidy" behavior compared to the right side. This is because, in the left side, each new value is dependent on the previous one, while in the right side each value is completely independent of the previous ones and may strongly vary w.r.t. the previous one.