The functional CLT (Donsker's invariance principle): proof and simulations

Pietro Colaguori 1936709

November 2023

1 Proof

Donsker's Invariance Principle, also known as the Functional Central Limit Theorem (CLT), is a mathematical result in probability theory that establishes a convergence result for certain types of stochastic processes to a particular limiting process. This principle is named after the mathematician Michael Donsker.

Donsker's theorem is a generalization of the central limit theorem (CLT) to the case of stochastic processes, rather than numbers. The idea is that if we have a sequence of independent and identically distributed random variables, we can construct a function from their partial sums and then make it converge in distribution to a standard Brownian motion, which is a continuous stochastic process with Gaussian increments.

The thesis is the following: Let $X_1, X_2, ...$ be i.i.d. RVs with $\mu = 0$ and variance σ^2 and let

$$S_n(t) = \frac{\sum_{i=1}^{[nt]} X_i}{\sqrt{n}}$$

be the partial-sum process. Then, by Donsker's theorem we know that:

$$S_n(t) \to^d \sigma B_t$$

in $D[0, 1]$ as $n \to \infty$. Which means that the partial-sum process converges in distribution to a Brownian motion in the space of càdlàg function on $[0, 1]$. Keeping in mind that càdlàg functions are those functions that are right-continous and have left limits, also called RCLL.

Regarding the proof, it can be divided into several steps that aim at showing how the characteristic function of $S_n(t)$ converges to the characteristic function of a Brownian motion.

1. The family $\{S_n(t)\}$ is tight in the space of càdlàg functions in $[0, 1]$.
2. Any weak limit of a subsequence of $\{S_n(t)\}$ is a standard Brownian motion.

3. The limit is unique, implying that the entire sequence $\{\{S_n(t)\}$ converges to a Brownian motion.

In summary, Donsker's theorem establishes a convergence result that connects discrete-time random walks to continuous-time Brownian motion, *de facto* extending the CLT to stochastic processes, which is particularly important in the field of finance.

2 Simulations

I created the following Python script, it generates random samples, calculates and plots an empirical distribution function, then we can see that as we increase the value of the sample size n the S_n converges to a Brownian motion.

```python
import numpy as np
import streamlit as st
import matplotlib.pyplot as plt

# number of processes
mean = 0
variance = 1
n = 10000

# Generate n i.i.d. variables
X_i = np.random.normal(mean, variance, n)

# Compute the sum
S_n = np.zeros(n)
for i in range(n):
    S_n[i] = np.sum(X_i[:i+1]) / np.sqrt(i+1)

# Plot the graph
st.line_chart(S_n)
```

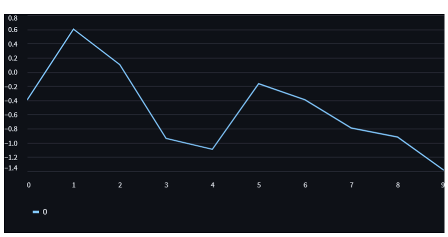

Figure 1: Simulation run with $n = 10$.

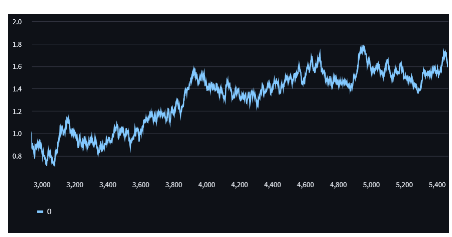

Figure 2: Simulation run with $n = 10000$.