Functional Central Limit Theorem

Chiara Iannicelli

November 2023

1 Introduction

The Functional Central Limit Theorem (CLT), also known as the Donsker Invariance Principle, is an extension of the classical Central Limit Theorem to the space of real-valued functions.

While the traditional Central Limit Theorem deals with the convergence of the distribution of sums of independent and identically distributed random variables to a normal distribution, the Donsker Invariance Principle extends this idea to function spaces.

2 Key concepts

- **Empirical Process:** In the context of the Donsker Invariance Principle, the empirical process refers to the collection of sample paths obtained by evaluating a given real-valued function on a sequence of random variables. Each sample path represents the function applied to a different realization of the random variables.

- **Donsker's Theorem:** Donsker's Theorem states that under certain conditions, the empirical process converges in distribution to a limiting process known as the Brownian bridge. This limiting process is a continuous-time stochastic process with stationary, independent increments.

- **Skorokhod Space:** The Donsker Invariance Principle is often formulated in terms of the Skorokhod space, which is a space of cadlag (right-continuous with left limits) functions equipped with the Skorokhod topology. The Skorokhod topology is a topology that allows for the convergence of functions with respect to the supremum norm.

- **Applications:** The Donsker Invariance Principle has applications in various areas, including statistics and empirical processes theory. It is particularly useful in the analysis of statistical procedures involving functionals of empirical processes.

3 Formula

Donsker's Theorem can be stated as follows: Let $X_1, X_2, ...$ be a sequence of independent and identically distributed random variables with common cumulative distribution function F.

Consider the empirical process:

$$G_n(u) = \sqrt{n}(\hat{F}(u) - F(u))$$

where $\hat{F}(u)$ is the empirical distribution function based on the first n bservations. Then, as n approaches infinity, G_n converges in distribution in the Skorokhod space to the Brownian bridge.

4 Intuition

The Donsker Invariance Principle provides a way to study the limiting behavior of functionals of empirical processes.

It establishes a connection between empirical processes and continuous-time stochastic processes, facilitating the application of probabilistic tools to the analysis of statistical procedures involving function spaces.

In summary, the Donsker Invariance Principle extends the insights of the Central Limit Theorem to the realm of functional data analysis, offering a powerful framework for understanding the convergence behavior of empirical processes in statistical applications.

5 Simulation

Simulating the Donsker Invariance Principle involves generating a sequence of independent and identically distributed random variables, constructing the empirical process, and demonstrating its convergence to a Brownian bridge.
The code below uses NumPy for numerical operations and Matplotlib for plotting:

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

```

```python
def donsker_invariance_simulation(num_samples):
    # Generate a sequence of independent and identically distributed random variables
    random_variables = np.random.randn(num_samples)

    # Calculate the cumulative distribution function (CDF) of the sample
    empirical_cdf = np.cumsum(random_variables) / np.sqrt(num_samples)

    # Plot the empirical process
    plt.plot(empirical_cdf, label='Empirical Process')
    plt.title('Empirical Process and Convergence to Brownian Bridge')
    plt.xlabel('Sample Index')
    plt.ylabel('Value')

    # Plot the limiting Brownian bridge for comparison
    brownian_bridge = np.random.randn(num_samples) / np.sqrt(num_samples)
    plt.plot(brownian_bridge, label='Brownian Bridge', linestyle='dashed')

    plt.legend()
    plt.show()

# Set the number of samples for the simulation
num_samples = 1000

# Simulate the Donsker Invariance Principle
donsker_invariance_simulation(num_samples)
```

In this code:

- np.random.randn generates a sequence of independent and identically distributed standard normal random variables.

- np.cumsum calculates the cumulative sum of the random variables to obtain the empirical process.

- The empirical process is plotted alongside a dashed line representing a realization of the Brownian bridge for comparison.

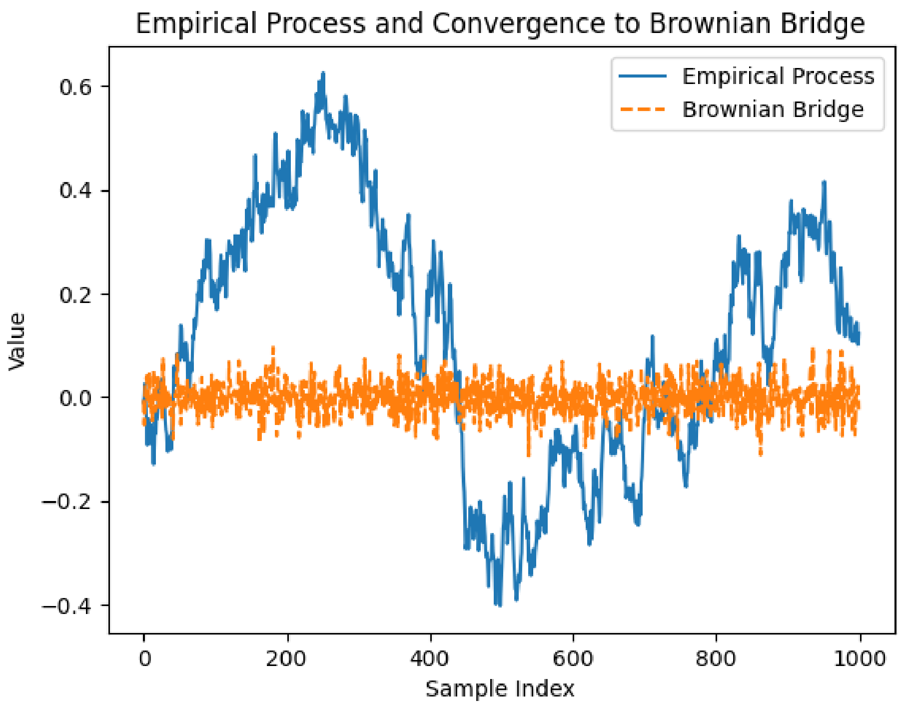

Figure 1: CLT simulation