The functional CLT

Wendy Rosettini

November 2023

1 Introduction

In probability theory, Donsker's theorem[1] (also known as Donsker's invariance principle, or the functional central limit theorem), named after Monroe D. Donsker, is a functional extension of the central limit theorem.

Let $X_1, X_2, X_3, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean 0 and variance 1. Let $S_n = \sum_{i=1}^{n} X_i$. The stochastic process $\{S_n\}_{n \in N}$ is known as a random walk. Define the diffusively rescaled random walk (partial-sum process) by

$$W^{(n)}(t) = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}, \quad t \in [0, 1].$$

The central limit theorem asserts that $W^{(n)}(1)$ converges in distribution to a standard Gaussian random variable $W(1)$ as $n \to \infty$. Donsker's invariance principle extends this convergence to the whole function $\{W^{(n)}(t)\}_{t \in [0,1]}$. More precisely, in its modern form, Donsker's invariance principle states that: As random variables taking values in the Skorokhod space $\mathcal{D}[0, 1]$, the random function $\{W^{(n)}\}$ converges in distribution to a standard Brownian motion $\{W(t)\}_{t \in [0,1]}$ as $n \to \infty$.

1.1 Donsker-Skorokhod-Kolmogorov theorem for uniform distributions.

Formal statement: Let F_n be the empirical distribution function of the sequence of i.i.d. random variables $X_1, X_2, X_3, \ldots$ with distribution function F. Define the centered and scaled version of F_n by

$$G_n(x) = \sqrt{n}(F_n(x) - F(x)),$$

indexed by $x \in R$. By the classical central limit theorem, for fixed x, the random variable $G_n(x)$ converges in distribution to a Gaussian (normal) random variable $G(x)$ with zero mean and variance $F(x)(1 - F(x))$ as the sample size n grows.

[1] https://shorturl.at/oISWX

1.2 Theorem (Donsker, Skorokhod, Kolmogorov):

The sequence of $G_n(x)$, as random elements of the Skorokhod space $\mathcal{D}(-\infty, \infty)$, converges in distribution to a Gaussian process G with zero mean and covariance given by

$$\text{cov}[G(s), G(t)] = E[G(s)G(t)] = \min\{F(s), F(t)\} - F(s),$$

$F(t)$.

The process $G(x)$ can be written as $B(F(x))$ where B is a standard Brownian bridge on the unit interval.

2 Proof

Here there is an overview of the key steps involved in the proof.

1. Definition of the Empirical Process:

Define the empirical process $G_n(x)$ as the centered and scaled version of the empirical distribution function $F_n(x)$ based on the sample $X_1, X_2, \ldots, X_n$.

$$G_n(x) = \sqrt{n}(F_n(x) - F(x))$$

2. Central Limit Theorem (CLT) for Empirical Process:

Utilize the classical central limit theorem to establish that, for fixed x, the random variable $G_n(x)$ converges in distribution to a Gaussian (normal) random variable $G(x)$ as $n \to \infty$.

3. Extension to the Skorokhod Space:

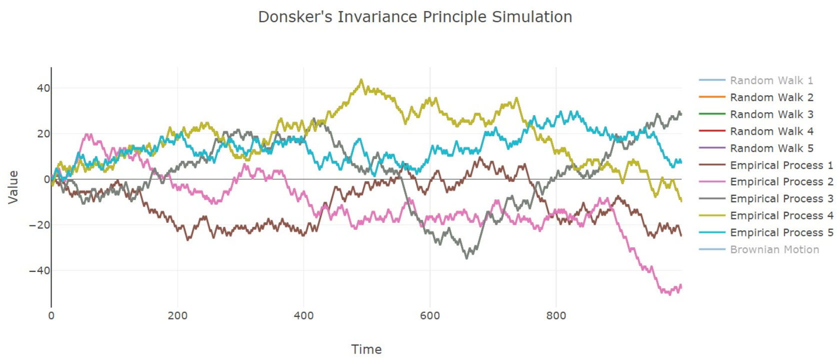
Extend the convergence result to the Skorokhod space $\mathcal{D}(-\infty, \infty)$, which is a space of functions equipped with the topology of convergence in distribution.

4. Identification of the Limit Process:

Identify the limit process $G(x)$ as a Gaussian process with zero mean and covariance given by $\text{cov}[G(s), G(t)] = \min\{F(s), F(t)\} - F(s)$, where $F(t)$ is the true distribution function.

5. Connection to Brownian Motion:

Show that the process $G(x)$ can be expressed as $B(F(x))$, where B is a standard Brownian bridge on the unit interval.



3 Simulations

The JavaScript simulation[2] provided is a visual representation of Donsker's Invariance Principle. Here's an explanation of the simulation:

1. Parameters:
   - `numSteps`: Number of steps in the random walk.
   - `numSimulations`: Number of simulations to run.
2. Simulation:
   - A simple random walk is generated for each simulation. At each step, a random value of +1 or -1 is added to the current value.
   - The empirical process is calculated, which is the cumulative sum of the random walk values. This process represents the partial sums of the random walk.
3. Plotting:
   - For each simulation, a line plot is created to show the trajectory of the random walk.
   - Another line plot is created to show the corresponding empirical process.
4. Brownian Motion: Additionally, a line plot for Brownian motion is included for comparison. Brownian motion is a continuous stochastic process and can be seen as a limit of the random walk when the number of steps becomes very large.
5. Interpretation:

[2] http://wendy.altervista.org/donskertheorem.html

- The simulation visually demonstrates how the trajectories of the random walks and their corresponding empirical processes evolve over time.
- As the number of steps increases, the empirical processes become smoother and converge toward a limiting continuous process, resembling Brownian motion. This aligns with Donsker's Invariance Principle.

6. Plot Interaction: The generated plot is interactive, can be zoomed in, pan, and hover over the lines to inspect specific points.

In summary, the simulation illustrates the convergence of the empirical process of a simple random walk to Brownian motion, showcasing a fundamental result in probability theory known as Donsker's Invariance Principle.