Day 1 : Random Variables

Dain Kim

November 23rd

What is this seminar for?

Stochastic Processes are ubiquitous! So I hope this seminar is helpful in diverse contexts.
We will formalize various stochastic processes mathematically in depth first, and move on to some applications to ML.
Our tentative plan is as follows: Any feedback is welcome!

- Random Variables
- Martingales
- Markov Processes
- Brownian Motion
- Stochastic Integration
- Stochastic Differential Equation (SDE)
- Diffusion Processes
- (ML) Diffusion Models

Textbooks and Resources

- [B] Borodin, Andrei N. Stochastic processes. Cham: Birkhäuser, 2017.
- [CE] Cohen, Samuel N., and Elliott, Robert James. Stochastic calculus and applications. Vol. 2. New York: Birkhäuser, 2015.
- MIT OCW 18.445 https://ocw.mit.edu/courses/18-445-introduction-to-stochastic-processes-spring-2015/pages/lecture-notes/

Today's Reading

[B] Chapter 1.1. Random Variables

Random Variables

Ω : a set of outcomes or sample points of experiments.
$\mathcal{F}$: a σ-algebra of Ω.
$\mathbb{P}$: a probability measure on $\mathcal{F}$.
The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space*.

Random Variables

Ω : a set of outcomes or sample points of experiments.
$\mathcal{F}$: a σ-algebra of Ω.
$\mathbb{P}$: a probability measure on $\mathcal{F}$.
The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space*.

$\mathcal{B}(\mathbb{R})$: the Borel σ-algebra on $\mathbb{R}$.
X is *measurable* if for any Borel set Δ, its inverse is in $\mathcal{F}$.

Definition

A *random variable* is a measurable mapping $X\colon (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Distribution Functions

Given a random variable X, its *distribution* is defined as

$$\mathcal{P}_X(\Delta) = \mathbb{P}(\{\omega : X(\omega) \in \Delta\}) \quad \text{for all } \Delta \in \mathcal{B}(\mathbb{R}).$$

Distribution Functions

Given a random variable X, its *distribution* is defined as

$$\mathcal{P}_X(\Delta) = \mathbb{P}(\{\omega : X(\omega) \in \Delta\}) \quad \text{for all } \Delta \in \mathcal{B}(\mathbb{R}).$$

There is a unique corresponding *distribution function* defined as

$$F_X(x) = \mathbb{P}(\{\omega : X(\omega) \in (-\infty, x)\}) = \mathbb{P}(X < x) \quad \text{for all } x \in \mathbb{R}.$$

Distribution Functions

Given a random variable X, its *distribution* is defined as

$$\mathcal{P}_X(\Delta) = \mathbb{P}(\{\omega : X(\omega) \in \Delta\}) \quad \text{for all } \Delta \in \mathcal{B}(\mathbb{R}).$$

There is a unique corresponding *distribution function* defined as

$$F_X(x) = \mathbb{P}(\{\omega : X(\omega) \in (-\infty, x)\}) = \mathbb{P}(X < x) \quad \text{for all } x \in \mathbb{R}.$$

If there is a nonnegative measurable function $f_X : \mathbb{R} \to \mathbb{R}_{\geq 0}$ such that

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy \quad \text{for all } x \in \mathbb{R},$$

then f_X is called the *density* of X.

Expected Values

The *expectation* of a random variable X is

$$\mathbb{E}X = \int_\Omega X(\omega)\mathbb{P}(d\omega) = \int_{-\infty}^{\infty} x\mathcal{P}_X(dx)$$

where $\mathbb{E}|X| = \int_\Omega |X(\omega)|\mathbb{P}(d\omega) < \infty$.

Expected Values

The *expectation* of a random variable X is

$$\mathbb{E}X = \int_\Omega X(\omega)\mathbb{P}(d\omega) = \int_{-\infty}^{\infty} x\mathcal{P}_X(dx)$$

where $\mathbb{E}|X| = \int_\Omega |X(\omega)|\mathbb{P}(d\omega) < \infty$.
In terms of the distribution function F_X, the expectation is

$$\mathbb{E}X = \int_{-\infty}^{\infty} x dF_X(x)$$

as a Stieltjes integral.

Expected Values

The *expectation* of a random variable X is

$$\mathbb{E}X = \int_\Omega X(\omega)\mathbb{P}(d\omega) = \int_{-\infty}^{\infty} x\mathcal{P}_X(dx)$$

where $\mathbb{E}|X| = \int_\Omega |X(\omega)|\mathbb{P}(d\omega) < \infty$.
In terms of the distribution function F_X, the expectation is

$$\mathbb{E}X = \int_{-\infty}^{\infty} x dF_X(x)$$

as a Stieltjes integral.
Then the *expectation* of a measurable function g of X is given by

$$\mathbb{E}g(X) = \int_{-\infty}^{\infty} g(x) dF_X(x).$$

The Convergence of Random Variables

A sequence of r.v.s X_n converges to X

- in mean : $\mathbb{E}|X_n - X| \to 0$
- in mean square : $\mathbb{E}(X_n - X)^2 \to 0$
- in probability : for any $\epsilon > 0$, $\mathbb{P}(|X_n - X| \geq \epsilon) \to 0$
- with probability one (a.s.) : $\mathbb{P}(\omega : X_n(\omega) \to X(\omega)) = 1$

The Convergence of Random Variables

A sequence of r.v.s X_n converges to X

- in mean : $\mathbb{E}|X_n - X| \to 0$
- in mean square : $\mathbb{E}(X_n - X)^2 \to 0$
- in probability : for any $\epsilon > 0$, $\mathbb{P}(|X_n - X| \geq \epsilon) \to 0$
- with probability one (a.s.) : $\mathbb{P}(\omega : X_n(\omega) \to X(\omega)) = 1$

Proposition

$X_n \to X$ in probability if and only if for any sequence $\{n_m\} \subset \mathbb{N}$, there exists a subsequence n_{m_k} such that $X_{n_{m_k}} \to X$ a.s.

Borel-Cantelli Lemma, part 1

Let $A_1, A_2, \ldots$ be a sequence of events. Then the event

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

consists of those and only those sample points ω that belong to an infinite number of events A_n, $n = 1, 2, \ldots$.

Lemma (Borel-Cantelli Lemma, part 1)

If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 0.$$

Borel-Cantelli Lemma, part 1

Let $A_1, A_2, \ldots$ be a sequence of events. Then the event

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

consists of those and only those sample points ω that belong to an infinite number of events A_n, $n = 1, 2, \ldots$.

Lemma (Borel-Cantelli Lemma, part 1)

If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 0.$$

Proof. By the uniform bound,

$$\mathbb{P}(\limsup_{n\to\infty} A_n) \leq \mathbb{P}(\cup_{m=n}^{\infty} A_m) \leq \sum_{m=n}^{\infty} \mathbb{P}(A_m) \to 0 \text{ as } n \to \infty.$$

The Convergence of Random Variables

Proposition

$X_n \to X$ in probability if and only if for any sequence $\{n_m\} \subset \mathbb{N}$, there exists a subsequence n_{m_k} such that $X_{n_{m_k}} \to X$ a.s.

Proof. (Forward) For any $\epsilon_k \to 0$, set

$$n_{m_k} = \min \left\{ n_m : \mathbb{P}(|X_{n_m} - X| > \epsilon_k) \leq \frac{1}{2^k} \right\},$$

which is possible since $X_n \to X$ in probability. Thus

$$\mathbb{P}(|X_{n_{m_k}} - X| > \epsilon_k) \leq \frac{1}{2^k}.$$

The Convergence of Random Variables

Since $\sum \frac{1}{2^k} < \infty$, by the Borel-Cantelli Lemma, there exists $k_0 = k_0(\omega)$ such that

$$|X_{n_{m_k}} - X| \leq \epsilon_k \quad \text{for all } k \geq k_0.$$

This implies that $X_{n_{m_k}} \to X$ a.s.

The Convergence of Random Variables

Since $\sum \frac{1}{2^k} < \infty$, by the Borel-Cantelli Lemma, there exists $k_0 = k_0(\omega)$ such that

$$|X_{n_{m_k}} - X| \leq \epsilon_k \quad \text{for all } k \geq k_0.$$

This implies that $X_{n_{m_k}} \to X$ a.s.

(Backward) For the sake of contradiction, suppose X_n does not converge to X in probability. Then there exist $\epsilon > 0$, $\delta > 0$, and a sequence n_m, such that

$$\mathbb{P}(|X_{n_m} - X| > \epsilon) \geq \delta.$$

But then there are no subsequences n_{m_k} of n_m such that $X_{n_{m_k}} \to X$ in probability, a contradiction. $\square$

Uniformly Integrability

A family of random variables $\{X_\alpha\}_{\alpha\in A}$ is called *uniformly integrable* if

$$\lim_{c\to\infty}\sup_{\alpha\in A}\int_{|X_\alpha|\ge c}|X_\alpha|d\mathbb{P}=0.$$

Proposition

Let $\{X_n\}_{n\in\mathbb{N}}$ be a uniformly integrable family of random variables and let $X_n\to X$ in probability. Then the random variable X is integrable and $\mathbb{E}|X_n-x|\to 0$.