MATH547: Stochastic Processes

Matthew He

January 20, 2025

Abstract

1 *Preliminaries*

Definition 1 (Probability space)**.**

1.1 *Independence*

Probability theory is about finding some sort of independence.

Definition 2 (Independence)**.**
Independence of Events
Events $A_1, A_2, A_3, \dots, A_n$ are independent if and only if:
For any collection $I \subseteq \{1, 2, \dots, n\}$

$$P\left(\bigcap_{j\in I} A_j\right) = \prod_{j\in I} P(A_j).$$

Independence of Random Variables
For random variables $X_1, X_2, \dots X_n$, taking values in state spaces $S_1, S_2, \dots S_n$, these are independent if and only if:
for any $E_1 \subseteq S_1, E_2 \subset S_2, \dots E_n \subset S_n$ the events $A_j = \{X_j \in E_j\}$ for $1 \leq j \leq n$ are independent.

1.2 *Convergences and limit theorems*

Definition 3 (In-probability convergence)**.**
The sequence $(X_j : j \in \mathbb{N}_0)$ converges in-probability to X_0 if for all $\epsilon > 0$,

$$\lim_{j\to\infty} P(d(X_j - X_0) > \epsilon) = 0,$$

in which case we write $X_j \xrightarrow[j\to\infty]{\text{Pr}} X_0$.

Theorem 1 (Law of large number (weak))**.** *If $\{X_n\}_{n=0}^{\infty}$ are real-valued, independent, identically distributed random variables and $\mathbb{E}(X_i) = \mu < \infty$, then*

$$\frac{1}{n}\sum_{j=1}^{n} X_j \to^{\mathbb{P}}_{n\to\infty} = \mathbb{E}(Xi)$$

Definition 4 (Almost sure convergence).
The sequence $(X_j : j \in \mathbb{N}_0)$ *converges almost surely to* X_0 *if*

$$\Pr\left(\limsup_{j\to\infty} d(X_j, X_0) > 0\right) = 0,$$

in which case we write $X_j \xrightarrow[j\to\infty]{\text{a.s.}} X_0$.

In-probability convergence does not rule out rare events.

The intuition behind the limsup here? Limsup always exist.

It is generally the case that in-probability convergence is strictly weaker than almost sure convergence.

Theorem 2 (Law of large number (strong)). *If* $\{X_n\}_{n=0}^{\infty}$ *are real-valued, independent, identically distributed random variables and* $\mathbb{E}(X_i) = \mu < \infty$, *then*

$$\frac{1}{n}\sum_{j=1}^{n} X_j \to_{n\to\infty}^{a.s.} = \mathbb{E}(Xi)$$

Large number theorem tells us value will converge to the expected value. Then central limit theorem tells us how close it will be.

Example 1 (*In-probability v.s. almost sure convergence*)

Suppose $\{X_j\}^{\infty}$ *be independent* $\{0, 1\}$ *- valued random variables.*

$$P(X_j = 1) = 1/j$$

In this case, we have

$$\lim_{j\to\infty} P(|X_j - 0| > \epsilon) = P(X_j = 1) = 1/j = 0,$$

therefore $X_j \xrightarrow[j\to\infty]{\Pr} 0$. However,

$$P(\limsup_{j\to\infty} |X_j - 0| > 0) = P(\limsup_{j\to\infty} X_j = 1) = 1,$$

Definition 5 (Weak convergence). *The sequence* $(X_j : j \in \mathbb{N}_0)$ *converges in law to* X_0 *if for all bounded continuous functions* $\varphi : S \to \mathbb{R}$

$$\lim_{j\to\infty} \mathbb{E}\varphi(X_j) = \mathbb{E}\varphi(X_0),$$

in which case we write $X_j \xrightarrow[j\to\infty]{\text{law}} X_0$.

Theorem 3 (Central limit theorem). *If* $\{X_n\}_{n=0}^{\infty}$ *are iid. random variables, with* $\mathbb{E}(X_i) = \mu < \infty$ *and* $Var(X_i) = \sigma^2 < \infty$, *then* $\forall t \in \mathbb{R}$,

$$\lim_{n\to\infty} P\left(\frac{\sum_{j=1}^{n} X_j - n\mu}{\sigma\sqrt{n}} \leq t\right) = \Phi(t)$$

where $\Phi(t)$ *is the standard normal distribution function.*

1.3 Conditional probability and expectation

Definition 6 (Conditional expectation on random variables). *If X takes countably many values and Y is a random variable with either $Y \geq 0$ or $\mathbb{E}|Y| < \infty$, then,*

$$\begin{aligned}\mathbb{E}[Y|X] &= \sum_{X} \mathbb{E}[Y|X=x]\mathbb{1}(X=x) \\ &= \sum_{X} \frac{\mathbb{E}[Y|\mathbb{1}(X=x)]}{P(X=x)}\mathbb{1}(X=x).\end{aligned}$$

Or, in measure-theoretic terms, we have:

$$\mathbb{E}[Y|X] = \mathbb{E}[Y|\sigma(X)]$$

Another case where we can condition:

Definition 7 (Conditioning on families of independent random variables). *If $(X_1, X_2, \ldots, X_n)$ are independent random variables in some state space S, and $Y = f(X_1, X_2, \ldots, X_n)$ for $f : S^n \mapsto \mathbb{R}$*

2 Stochastic Processes

A stochastic process is a family of random variables indexed by natural numbers (time).

Definition 8 (Stochastic process). *A stochastic process is a collection of random variables $\{X_t\}_{t\in\mathbb{N}_0}$ where t represents time.*

Basic examples:
1. iid. randomHello Professor Masset! I was wondering if you would be free to meet up for a quick discussion about the things this term? We could meet whenever works best for you. If you're busy we can also just meet after the lab meeting starts
 Smile eyes variables (e.g. iid. dice rolls)
2. Random walk

Definition 9 (Filtration). *A filtration $(\mathcal{F}_j : j \geq j_0)$ is a sequence of σ-algebras with the property that they are increasing, so for all $j \geq j_0$, $\mathcal{F}_j \subseteq \mathcal{F}_{j+1}$. A stochastic process $(X_j : j \geq j_0)$ is adapted to a filtration if X_j is $\mathcal{F}_j$-measurable for all $j \geq j_0$. Any stochastic process also gives rise to a filtration, its natural filtration, just by setting $\mathcal{F}_j = \sigma(X_k : j_0 \leq k \leq j)$.*

Definition 10 (Countable state space markov chain). *A stochastic process $\{X_j\}_{j\in\mathbb{N}_0}$ satisfies the Markov property.*

σ-algebra is about information.

Why We Can Split the Event

We use the law of total probability (or the chain rule for conditional probabilities) to split the event $\bigcap_{i=1}^{m+1}\{X_{k+i} \in A_{k+i}\}$ at X_{k+1}. By conditioning on X_{k+1}, we factor the probability and then apply the induction hypothesis for the remaining terms.

Explanation of the Induction

The key is that for $n = m + 1$, we split the event $\bigcap_{i=1}^{m+1}\{X_{k+i} \in A_{k+i}\}$ at X_{k+1}, apply the induction hypothesis for the remaining m steps, and then use (a) to reduce the conditioning to X_k alone. This confirms that each additional step depends only on the immediately preceding state.

References