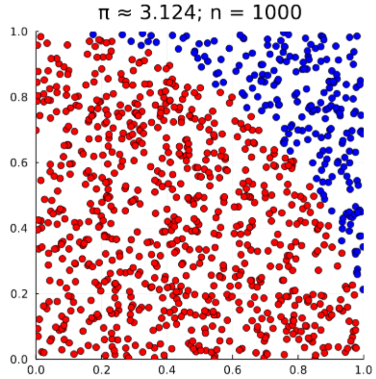
Anatomy of a Random Variable

Outline

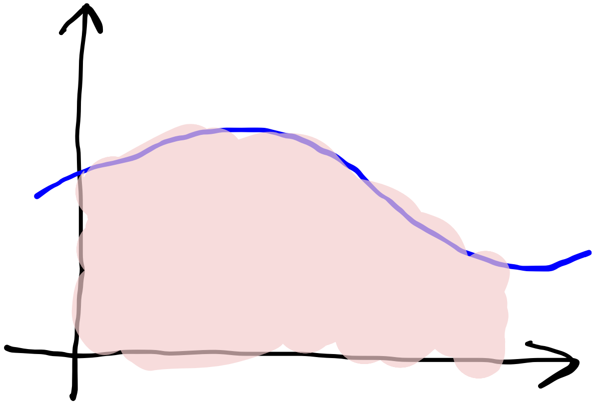


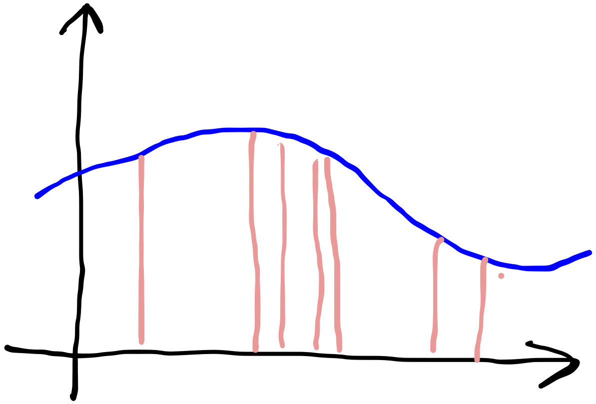
- A Motivating Example: Monte Carlo Integration
- Rigorous Definitions of a Random Variable
- Law of large numbers and the Central Limit Theorem

$$X : \Omega \to E$$

$$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right)$$

Monte Carlo Integration





$$X_i \sim U(\Omega)$$

$$I = \int_\Omega f(x)dx$$

$$I \approx Q_N \equiv \frac{\int_\Omega dx}{N} \sum_{i=1}^{N} f(X_i)$$

Monte Carlo Integration

Special Case: Expectation

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x\, p(x)\, dx$$

$$\approx \frac{1}{N} \sum_{i=1}^{N} X_i$$

How accurate is this?

Random Variables

Why are probability distributions not enough?

~~Consider this definition: *Two random variables are equal if their probability distributions are the same.*~~

$$A \sim \mathrm{Bernoulli}(0.5)$$

$$B = \neg A$$

$$\cancel{B \stackrel{?}{=} A}$$

Random Variables

Sample space Event Space Probability Measure

Given a **probability space** $(\Omega, \mathcal{F}, P)$, and a **measurable space** $(E, \mathcal{E})$, an E-valued **random variable is a measurable function** $X : \Omega \to E$.

$$\omega \in \Omega$$

$$X(\omega) \in E$$

What is this function?

What is $\omega \in \Omega$?

Example: Coin World

$$\Omega = \{H, T\}$$

$$E = [0, 1]$$

$$X(\omega) = \mathbf{1}_{\{H\}}(\omega)$$

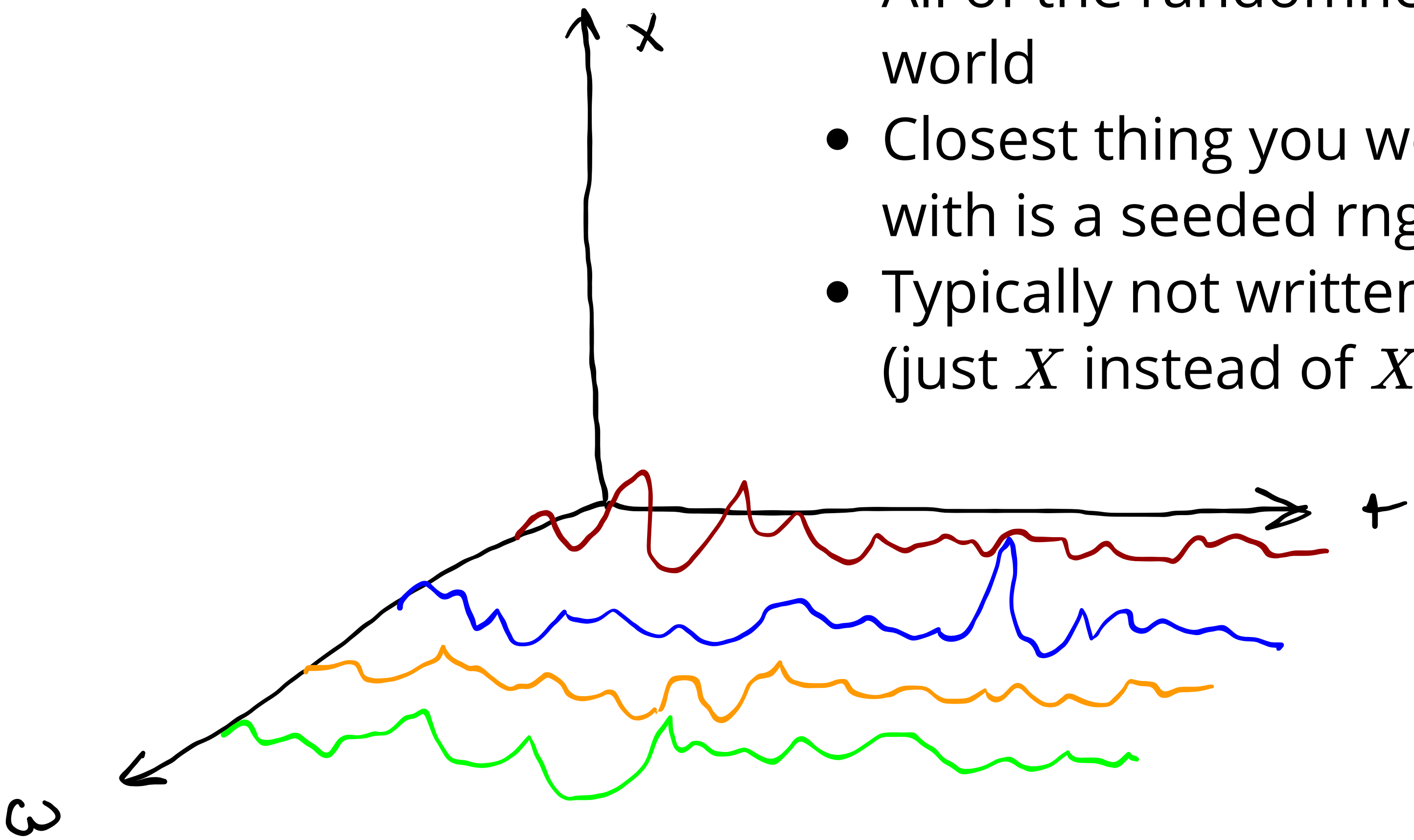
Example: Many coins

$$\Omega = \{H, T\}^\infty$$

$$E = [0, 1]$$

$$X_i(\omega) = \mathbf{1}_{\{H\}}(\omega_i)$$

What is $\omega \in \Omega$?



- All of the randomness in the world
- Closest thing you work regularly with is a seeded rng
- Typically not written expressly (just X instead of $X(\omega)$)

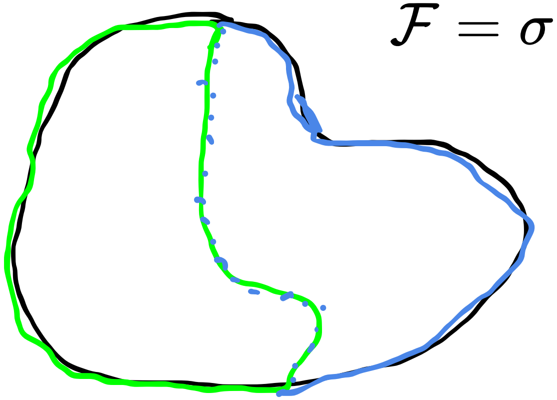
What is $\mathcal{F}$ (and $\mathcal{E}$)?

- "σ-algebra" or "σ-field"
- Subset of subsets of Ω (that is, $\mathcal{F} \subseteq 2^{\Omega}$)
- Three requirements to be a σ-field
  - $\Omega \in \mathcal{F}$
  - If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$ (where $A^c = \Omega \backslash \mathcal{F}$)
  - If $A_i \in \mathcal{F}$ for $i \in \mathbb{N}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$
- $\sigma(\cdot)$ creates a σ-field from a set of generators

$\Omega: \{1,2,3\}$

$2^{\Omega} = \{\{1\},\{2\},\{3\},\emptyset,$
$\{1,2\},\{1,3\},\{2,3\}$
$\{1,2,3\}\}$

Ω



$$\mathcal{F} = \sigma(\{\ \}) = \{\Omega, \ , \ , \emptyset\}$$

Borel Sigma Algebra

The Borel σ-algebra for a topological space Ω is the σ-field generated by all open sets in Ω.

The Borel σ-field on $\mathbb{R}$ is $\mathcal{B} = \sigma(\{(a, b) : a, b \in \mathbb{R}\})$

$(-\infty, 1)$ $(2, \infty)$

- Is $[1, 2]$ in $\mathcal{B}$?
- Is 1 in $\mathcal{B}$?
- Is π in $\mathcal{B}$?

- $\Omega \in \mathcal{F}$
- If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$ (where $A^c = \Omega \backslash \mathcal{F}$)
- If $A_i \in \mathcal{F}$ for $i \in \mathbb{N}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$

What is P?

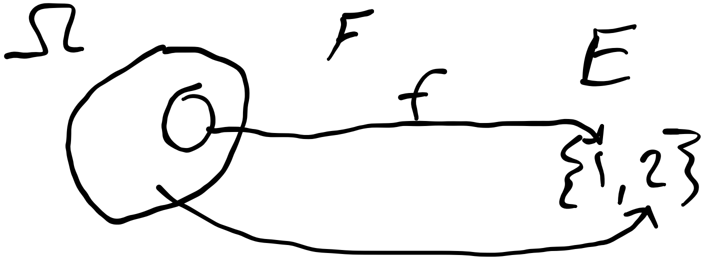
A probability measure P is a function $P : \mathcal{F} \to [0, 1]$ having the following properties:

1. $0 \leq P(A) \leq 1 \quad \forall\, A \in \mathcal{F}$.
2. $P(\Omega) = 1$.
3. (Countable additivity) $P(A) = \sum_{n=1}^{\infty} P(A_n)$ whenever $A = \cup_{n=1}^{\infty} A_n$ is a countable union of disjoint sets $A_n \in \mathcal{F}$

Random Variables

Given a **probability space** $(\Omega, \mathcal{F}, P)$, and a **measurable space** $(E, \mathcal{E})$, an E-valued **random variable is a measurable function** $X : \Omega \to E$.

A function $f : \Omega \to E$ is measurable if for every $A \in \mathcal{E}$, the pre-image of A under f is in $\mathcal{F}$. That is, for all $A \in \mathcal{E}$



$$f^{-1}(A) \in \mathcal{F}$$

Are there functions that are not Borel-measurable?

Advantages over pdf definition

- Rigorous treatment of deterministic outcomes
- More sophisticated convergence concepts
- Better way of thinking about related random variables (personally, I think)

$$
\begin{tikzcd}
& \omega \arrow[dl] \arrow[dr] & \\
A \arrow[rr, leftrightarrow] & & B
\end{tikzcd}
$$

Break

Exercise 1.2.5. *Let* $\Omega = \{1, 2, 3\}$. *Find a* σ*-field* $\mathcal{F}$ *such that* $(\Omega, \mathcal{F})$ *is a measurable space, and a mapping* X *from* Ω *to* $\mathbb{R}$, *such that* X *is not a random variable on* $(\Omega, \mathcal{F})$.

A function $f : \Omega \to E$ is measurable if for every $A \in \mathcal{E}$, the pre-image of A under f is in $\mathcal{F}$. That is, for all $A \in \mathcal{E}$

$$f^{-1}(A) \in \mathcal{F}$$

$$\mathcal{F} = \{\Omega, \emptyset, \{1\}, \{2, 3\}\}$$

$$X = \mathbf{1}_{\{1,2\}}$$

https://timer.onlineclock.net/

Convergence

Review: For a (deterministic) sequence $\{x_n\}$, we say

$$\lim_{n \to \infty} x_n = x$$

or

$$x_n \to x$$

if, for every $\epsilon > 0$, there exists an N such that $|x_n - x| < \epsilon$ for all $n > N$.

Convergence

In what senses can we talk about random variables converging?

- Sure ("pointwise")
- Almost Sure
- In Probability
- Weak ("in distribution"/"in law")

When are two R.V.'s the same?

X = Y if $X(\omega) = Y(\omega) \quad \forall \omega \in \Omega$

In practice, there are often unimportant ω where this is not true.

We say that X is *almost surely* the same as Y if $P(\{\omega : X(\omega) \neq Y(\omega)\}) = 0.$

This is denoted $X \overset{a.s.}{=} Y$ and the terms *almost everywhere* (a.e.) and *with probability 1* (w.p.1) mean the same thing.

Sure Convergence

$$X_n(\omega) \to X(\omega) \quad \forall\, \omega \in \Omega$$

Almost Sure Convergence

$X_n \stackrel{a.s.}{\to} X$ if there exists $A \in \mathcal{F}$ with $P(A) = 1$ such that $X_n(\omega) \to X(\omega)$ for each fixed $\omega \in A$.

Does sure convergence imply almost sure convergence?

Convergence in Probability

$X_n \to_p X$ if $P(\{\omega : |X_n(\omega) - X(\omega)| > \epsilon\}) \to 0$ for any fixed $\epsilon > 0$.

Does $X_n \overset{a.s}{\to} X$ imply $X_n \to_p X$?

Yes.

Convergence in Probability

Does $X_n \to_p X$ imply $X_n \overset{a.s}{\to} X$?

No.

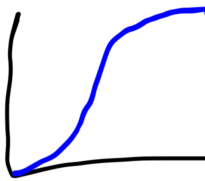
PROOF. Consider the probability space $\Omega = (0,1)$, with Borel σ-field and the Uniform probability measure U of Example 1.1.11. Suffices to construct an example of $X_n \to_p 0$ such that fixing each $\omega \in (0,1)$, we have that $X_n(\omega) = 1$ for infinitely many values of n. For example, this is the case when $X_n(\omega) = \mathbf{1}_{[t_n, t_n+s_n]}(\omega)$ with $s_n \downarrow 0$ as $n \to \infty$ slowly enough and $t_n \in [0, 1-s_n]$ are such that any $\omega \in [0,1]$ is in infinitely many intervals $[t_n, t_n + s_n]$. The latter property applies if $t_n = (i-1)/k$ and $s_n = 1/k$ when $n = k(k-1)/2 + i$, $i = 1, 2, \ldots, k$ and $k = 1, 2, \ldots$ (plot the intervals $[t_n, t_n + s_n]$ to convince yourself). ∎

But there exists a subsequence n_k such that $X_{n_k} \overset{a.s.}{\to} X$.

Weak Convergence

Let $F_X : \mathbb{R} \to [0, 1]$ be the cumulative distribution function of real-valued random variable X.

$F_X(x) = P(X \leq x)$



$X_n \xrightarrow{D} X$ if $F_{X_n}(\alpha) \to F_X(\alpha)$ for each fixed α that is a continuity point of F_X.

"Weak convergence", "convergence in distribution", and "convergence in law" all mean the same thing.

Convergence

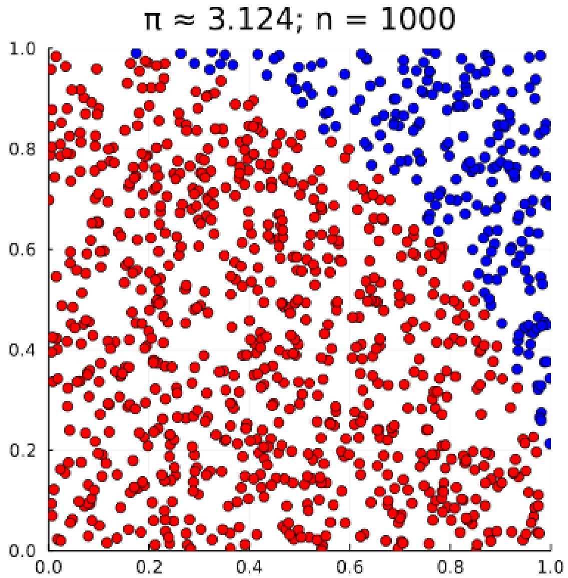
In what senses can we talk about random variables converging?

Stronger

- Sure ("pointwise")
- Almost Sure
- In Probability
- Weak ("in distribution"/"in law")

Implies

Convergence of MC integration

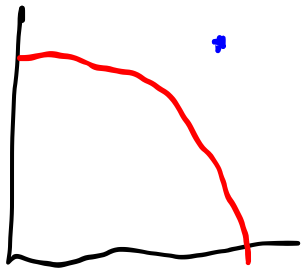


Let X_i be independent, identically distributed random variables with mean μ, and $Q_N \equiv \frac{1}{N}\sum_{i=1}^{N} X_i$.

$$Q_N \stackrel{?}{\to} \mu?$$

Convergence of MC integration

$\exists \omega \in \Omega$ where you always sample the same point.



$Q_N \to \mu$ (sure)?

$Q_N \xrightarrow{a.s.} \mu$? Strong law of large numbers

Probability that there are enough measurements off in one direction to keep $|Q_N - \mu| > \epsilon$ decays with more samples.

$Q_N \to_p \mu$? Weak law of large numbers

$Q_N \xrightarrow{D} \mu$?

Convergence *Rate* of M.C. Integration

How do you quantify $|Q_N - \mu|$?

Run M sets of N simulations and plot a histogram of Q_N^j for $j \in \{1, \ldots, M\}$.

Central Limit Theorem

Lindeberg-Levy CLT: If $\mathrm{Var}[X_i] = \sigma^2 < \infty$, then

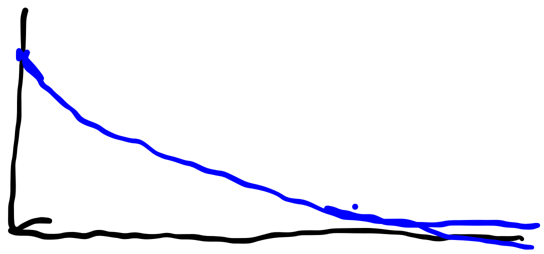
$$\sqrt{N}(Q_N - \mu) \xrightarrow{D} \mathcal{N}(0, \sigma)$$

After many samples Q_N starts to look distributed like $\mathcal{N}(\mu, \frac{\sigma}{\sqrt{N}})$

Central Limit Theorem

Two somewhat astounding takeaways:

1. Error decays at $\frac{1}{\sqrt{N}}$ *regardless of dimension*.



2. You can estimate the "standard error" with

$$SE = \frac{s}{\sqrt{N}}$$

where s is the sample standard deviation.