09/28/21 Agenda

- review key probability definitions
- LLN / CLT
- why Monte Carlo (MC) methods are ill-suited for when we don't know if a random variable has finite expectation
- why simple MC has convergence like $O(1/\sqrt{n})$
- ex 1. (oh-so tangentially related)
  MC integration (polytope ← stochastic
  estimate π reachability)
- online code examples of propagating uncertainty through a PDE/ODE

REVIEW

Def. a continuous random variable X is a (continuous) map $X: \Omega \longrightarrow \mathbb{R}$ where

- Ω is any nonempty set (called the event space).

The Ω represents "abstract elementary events" (something in the actual world that we observe that doesn't immediately have an associated numeric value).

Ex. $\Omega = \{\text{"Heads"}, \text{"Tails"}\}$ } the literal events observed in the world; for this class, it's not important to consider them

stars hearts

What's so random?

$$X : \Omega \longrightarrow \mathbb{R}$$

- "randomness" comes from not knowing the input "event" $\omega \in \Omega$
- we can assign a probability to certain "outcomes" occuring

$\mathbb{P}(X \leq a) =$ shorthand for: what is the chance that $X(\omega) \leq a$ for an unknown $\omega \in \Omega$?

side: measure theory

Def. A probability density function (PDF) of a random variable X is a map $f_X : \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0}$ that satisfies certain properties:

non-neg. reals

* Unless otherwise stated, all given random vars have a prob. density (or a prob. mass function)

(1) For $a, b \in \mathbb{R}$ $(a < b)$, we have

$$\mathbb{P}(a < X < b) = \int_a^b f_X(s)\,ds,$$

(2) $f_X \geq 0$

(3) f_X is piecewise constant

(4) $\int_{-\infty}^{\infty} f_X(s)\,ds = 1$

Ex. $X \sim N(0,1) \Longrightarrow f_X : \mathbb{R} \to \mathbb{R}_{>0}$

$$f_X(s) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right)$$

$$\mathbb{P}(X \leq a) = \int_{-\infty}^{a} f_X(s)\,ds$$

Ex. $X \sim U[-1,1]$

$$f_X(s) = \frac{1}{2}$$

(in general, $X \sim U[c,d]$, $f_X(s) = \frac{1}{d-c}$, $c \neq d$)

Def. (First order statistics)

For a continuous RV X with PDF f_X,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} s\, f_X(s)\,ds = \text{mean of } X \quad \text{and}$$

$$\mathbb{V}[X] = \int_{-\infty}^{\infty} \left(s^2 - \mathbb{E}(X)^2\right) f_X(s)\,ds = \text{variance of } X$$

Thm. (law of large numbers)

Consider some i.i.d. random variables $Y_1, \ldots, Y_n$ with respect to some random variable $\underline{Y}$. Then if $\mathbb{E}[Y] < \infty$, we have that

independent and identically distributed

$$S_n := \frac{1}{n} \sum_{i=1}^{n} Y_i$$

(sample mean)

converges in some sense to the "true" mean $\mathbb{E}[Y]$. In particular,

the weak law of large numbers states that $S_n \longrightarrow \mathbb{E}(Y)$ converges in probability; that is,

$$\text{for all } \epsilon > 0, \quad \lim_{n \to \infty} \mathbb{P}(|S_n - \mathbb{E}(Y)| > \epsilon) = 0.$$

↳ some number

The strong law of large numbers says $S_n \longrightarrow \mathbb{E}(Y)$ almost surely ("with probability 1", "convergence almost everywhere"); that is,

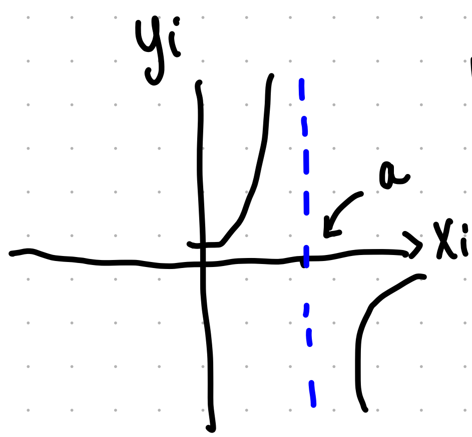
$$\mathbb{P}\left[\lim_{n \to \infty} S_n = \mathbb{E}(Y)\right] = 1.$$

Who cares if $\mathbb{E}(Y) < \infty$?

You do!

Let $X_i \sim N(0,1)$. Let $Y_i = \frac{1}{a - X_i}$ $a \in \mathbb{R}$ (fixed).

"distributed as" or "is a RV whose PDF is dictated by the RHS expression"



$\mathbb{E}(Y)$ is undefined when $X_i = a$

- We might run into trouble when our RVs involve fractions. (Importance sampling....)

How about that var(Y)?

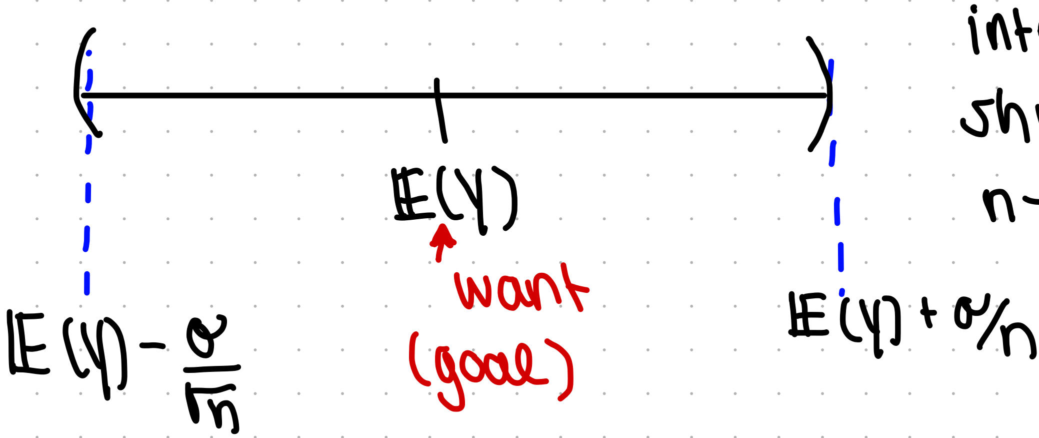
When $var(Y) < \infty$, then the Central Limit Theorem gives that for i.i.d. Y_i,

$$\sqrt{n}\left(\frac{|S_n - \mathbb{E}(Y)|}{var(Y)}\right) \xrightarrow{dist.} N(0,1).$$

("weaker" than convergence in probability)

wikipedia

$$S_n \in \left[\mathbb{E}(Y) - \frac{\sigma}{\sqrt{n}}, \mathbb{E}(Y) + \sigma/\sqrt{n}\right] \quad \sigma = \sqrt{var(Y)}$$



interval length shrinks as $n \to \infty$

$$S_n \xrightarrow{\text{dist.}} N\left(\mathbb{E}(Y), \frac{\sigma^2}{n}\right)$$

$$\mathbb{E}(Y) = \mathbb{E}(Y_i)$$

$$\underline{|S_n - \mathbb{E}(Y)|} \leq \frac{\propto \sigma}{\sqrt{n}}$$

$\propto$ proportionality const. $(\propto > 0)$

Thus,

$$S_n \to \mathbb{E}(Y) \text{ in } \underline{\mathcal{O}(1/\sqrt{n})}.$$

How does this relate to MC explicity?

input parameter X	model function f	response (QoI) $y \approx f(X)$

- We care about

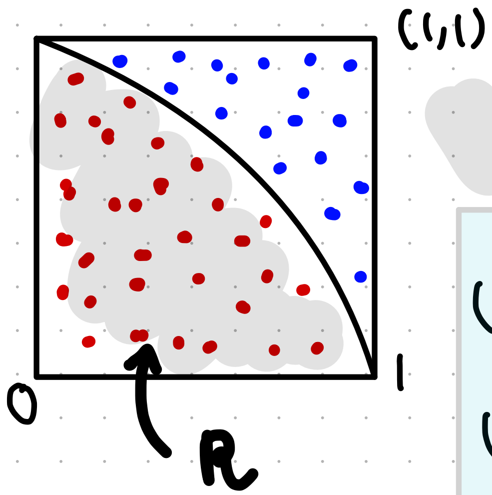
$$\mathbb{E}(Y) = \mathbb{E}(\underline{f}(X))$$

estimating this quantity with

$$S_n(f) := \frac{1}{n} \sum_{i=1}^{n} f(X_i)$$

Ex. MC INTEGRATION

Goal: estimate π



$= \frac{1}{4}\pi =$ area of region R

(1) Take $\vec{X}_i \sim U(0,1)$

(2) If $\vec{X} \in R$, color the dot red

else color blue

Check if $X \in R$ via checking if $\underbrace{(X_{i,1})^2 + (X_{i,2}^2)}_{\text{components of } \vec{X} \in \mathbb{R}} \leq 1$.

Define

$$Y_i = \mathbb{1}_{\text{ith point in } R} = \begin{cases} 1 & \text{if } \vec{X}_i \in R \\ 0 & \text{else.} \end{cases}$$

(indicator random var.)

Consider $S_n = \frac{1}{n} \sum_{i=1}^{n} Y_i$ and $\mathbb{E}(Y)$.

Then $S_n = \frac{1}{n} \sum_{i=1}^{n} y_i$ only nonzero if $\vec{X}_i \in R$

$$= \frac{\sum_{i \text{ s.t. } X_i \in R} (1)}{n} = \frac{\# \text{ of points in } R}{\text{total } \# \text{ of points}} = \underbrace{\left(\frac{\text{red dots}}{\text{red} + \text{blue dots}} \right)}_{\text{ratio}}$$

$$\mathbb{E}(Y_i) \overset{\text{law of total expectation}}{=} \underbrace{\mathbb{E}[Y_i \mid X_i \in R]}_{1}\, \mathbb{P}(X_i \in R) + \overbrace{\mathbb{E}[Y_i \mid X_i \notin R]}^{0}\, \mathbb{P}(X_i \notin R)$$

$$= \mathbb{P}(X_i \in R) + 0$$

$$= \mathbb{P}(X_i \in R)$$

Y_i are i.i.d. $\Longrightarrow \mathbb{E}(Y) = \mathbb{P}(\underline{X_i} \in R)$

$$S_n \longrightarrow \mathbb{E}(Y)$$

$$\begin{array}{ccc} S_n & & \\ \| & & \\ \dfrac{\#\text{ of red pts}}{\text{total }\#\text{ of points } n} & \xrightarrow{\text{LLN as } n\to\infty} & \mathbb{P}(X_i \in R) \\ & & \| \\ & & \displaystyle\int_{s\in R} \underbrace{f_{X_i}(s)}_{\text{density}}\, ds \\ & & \| \\ & & \text{area}(R) \end{array}$$

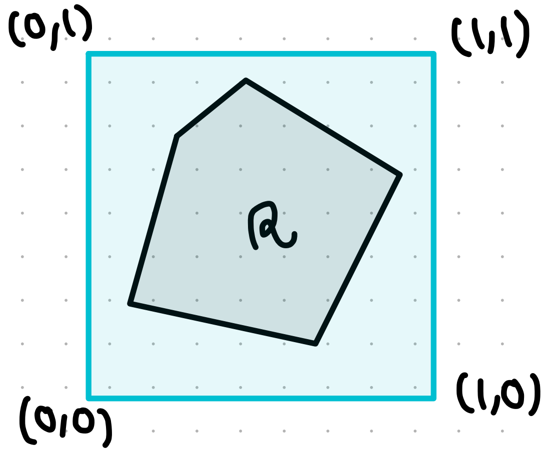
$X_i \sim U(0,1)$

$$\overbrace{\left(\frac{\#\text{ of red pts}}{\text{total }\#\text{ of points } n}\right)}^{S_n} \xrightarrow[n\to\infty]{\text{LLN}} \text{area}(R) = \pi/4$$

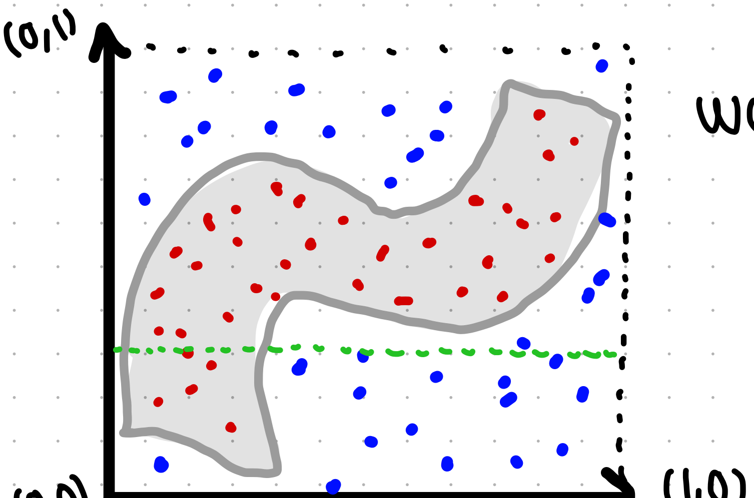
(optional lil' HW problem)

so $\pi \approx 4 S_n$ for large n.

Using MC to find the area of stranger regions



R = polytope (e.g. representing feasible initial states)



weird region

$$S_n = \text{ratio of red pts to total points} \xrightarrow{n \to \infty} \text{area of weird region}$$

Take $X_i \sim U(0,1)$.

Let $Y_i = \underline{\mathbb{1}\{X_i \in \text{weird region}\}}$. How to efficiently evaluate if a given $X_i \in R$?

Ex:

- Inside or outside alg.

(count parity of # of times .... line intersects boundary of R to check if $\vec{X}_i \in R$)