Important Random Variables (Discrete)

1) Bernoulli random variable.

- Toss a coin with $P(H) =: p$.
- $\underline{X} := \begin{cases} 1, & \text{if } H \text{ appears} \\ 0, & \text{else} \end{cases}$
- Prob. mass fn. $P(X=1) = P(H) = p$.

▷ $E[X] = P(X=1)\cdot 1 + P(X=0)\cdot 0 = p$.

(it's <u>not</u> 0 or 1 !)

2) Binomial random variable.

If we repeat the prior experiment n times, then how many times H appears? Say X.

- Prob. Mass fn.: $P(X=i) = \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$.

$\triangleright$ $E[X] = \sum_{i=0}^{n} P(X=i) \cdot i = \sum_{i=0}^{n} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \cdot i$

$= np \cdot \sum_{i=0}^{n} \binom{n-1}{i-1} \cdot p^{i-1} \cdot (1-p)^{n-i} = np \cdot (p+1-p)^{n-1} = np$.

Alternately, X = sum of Bernoulli trials X_i, $i \in [n]$:

$\triangleright$ $E[X] = \sum_{i=1}^{n} E[X_i] = \sum_i p = n \cdot p$.

3) Geometric Random Variable.

- Toss a coin with $P(H) =: p$, till you get H.
- What's $\underline{X_1} := \#(\text{tosses to get } H)$?

- Prob. mass fn.: $P(X_1 = k) = (1-p)^{k-1} \cdot p$

$\triangleright$ $E[X_1] = \sum_{k \geq 1} P(X_1 = k) \cdot k = \sum_{k \geq 1} (1-p)^{k-1} \cdot p \cdot k$.

- $\underline{\text{Simplify}}$: $= \sum_{k \geq 0} P(X_1 > k) = \sum_{k \geq 0} (1-p)^k = \frac{1}{1-(1-p)} = \frac{1}{p}$.

4) Negative binomial random variable.

- Toss a coin with $P(H) =: p$, till you get n H's.
- What's $X_n := \#(\text{tosses to get } n\ H\text{'s})$?
- Prob. mass fn.: $P(X_n = k) = \binom{k}{n} \cdot p^n \cdot (1-p)^{k-n}$

▷ $E[X_n] = \sum_{k \geq n} P(X_n = k) \cdot k = \sum_{k \geq n} \binom{k}{n} \cdot p^n \cdot (1-p)^{k-n} \cdot k$

- Better analysis: $E[X_n] = \sum_{i=1}^{n} E[i\text{-th } H \mid (i-1)\text{-th } H]$

$= n \cdot E[X_1] = n/p$.

Continuous Random Variables

- We could define mass function for X when its range is infinite, say $\mathbb{R}$.
  ↳ Naively $P(X=\varepsilon)=0, \forall \varepsilon \in \mathbb{R}$.
- Defn: • Continuous random variable X is defined by a fn. $f_X : \mathbb{R} \to \mathbb{R}$ called probability density function (pdf) s.t.
  - $\forall a, b \in \mathbb{R}$, $P(a \leq X \leq b) = \int_a^b f_X(x) \cdot dx$
  - $\int_{-\infty}^{\infty} f_X(x) \cdot dx = 1$.

infinitesimal

↳ Interpret "$f_X(x) \cdot dx$" as the prob. of X being "close to" x.

- With this interpretation, the expectation is:

Defn: $$E[X] := \int_{-\infty}^{\infty} x \cdot \left(f_X(x) \cdot dx\right).$$

1) Exponential Random Variable.

For parameter $\lambda > 0$, define $\underline{f_X(x)} := \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

▷ $$\int_0^{\infty} \lambda e^{-\lambda x} \cdot dx = \left(-e^{-\lambda x}\right)\Big|_0^{\infty} = \underline{1}.$$

▷ $$E[X] = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} \cdot dx = 1/\lambda$$

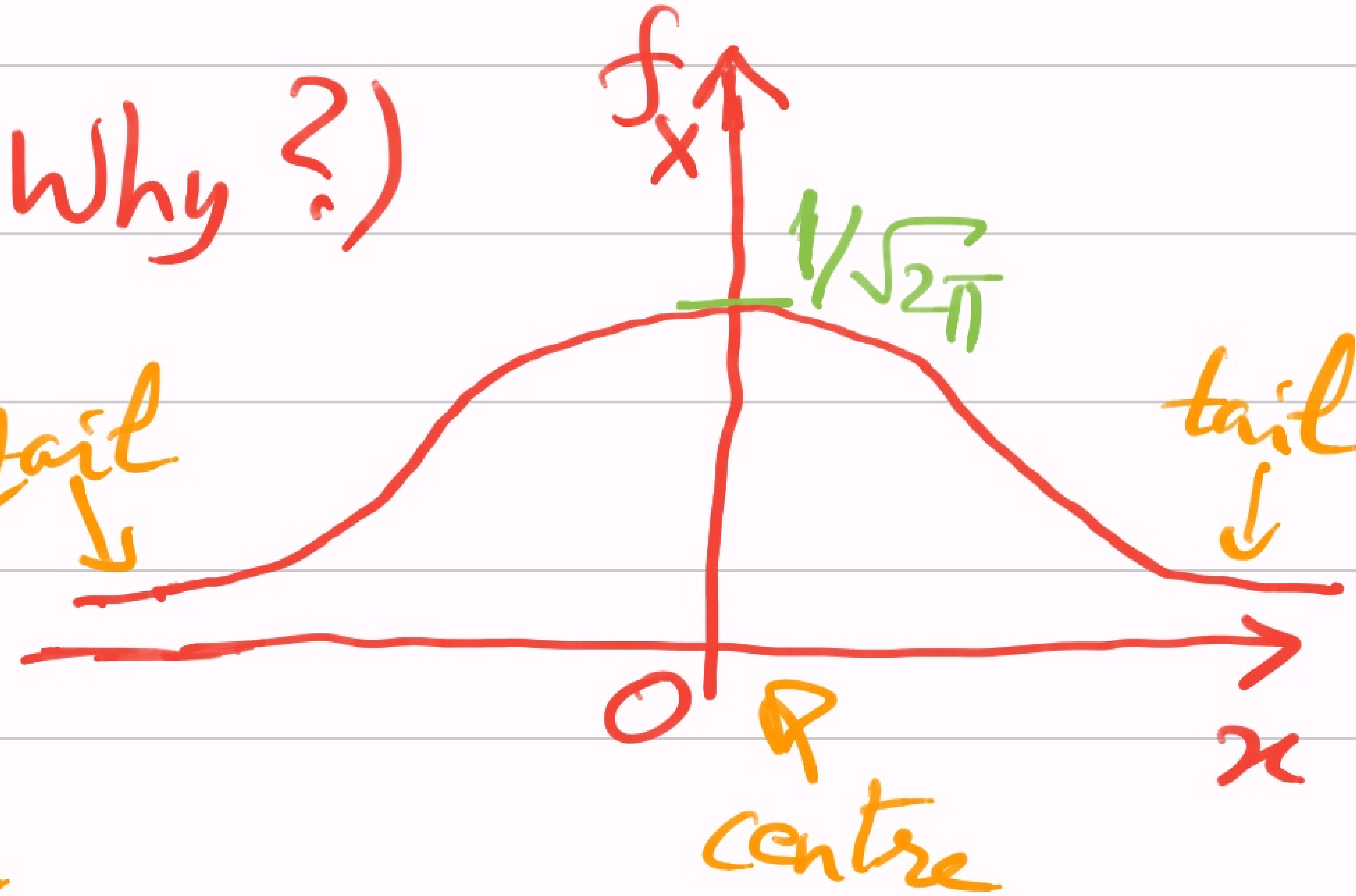
(Why?)

- Make it symmetric by using "x^2":

2) Normal / Gaussian random variable.

Define $f_X(x) := \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$, for $x \in \mathbb{R}$.

▷ $\int_{-\infty}^{\infty} f_X(x) \cdot dx = \underline{1}$ (Why?)



▷ $E[X] = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$

$= \frac{1}{\sqrt{2\pi}} \cdot \left(-e^{-x^2/2}\right)\Big|_{-\infty}^{\infty} = 0.$

- Gives the famous <u>bell-curved</u> Gaussian distribution (or <u>standard normal</u>).

- This is a fundamental distribution due to the <u>central limit theorem</u>:

Let $X_1, \ldots, X_n$ be a sequence of independent & identically distributed (<u>iid</u>) random variables drawn from a distribution with exp. $=: \underline{\mu}$.

Consider the <u>sample-average</u> $\overline{X}_n := \sum_{i\in[n]} X_i / n$.

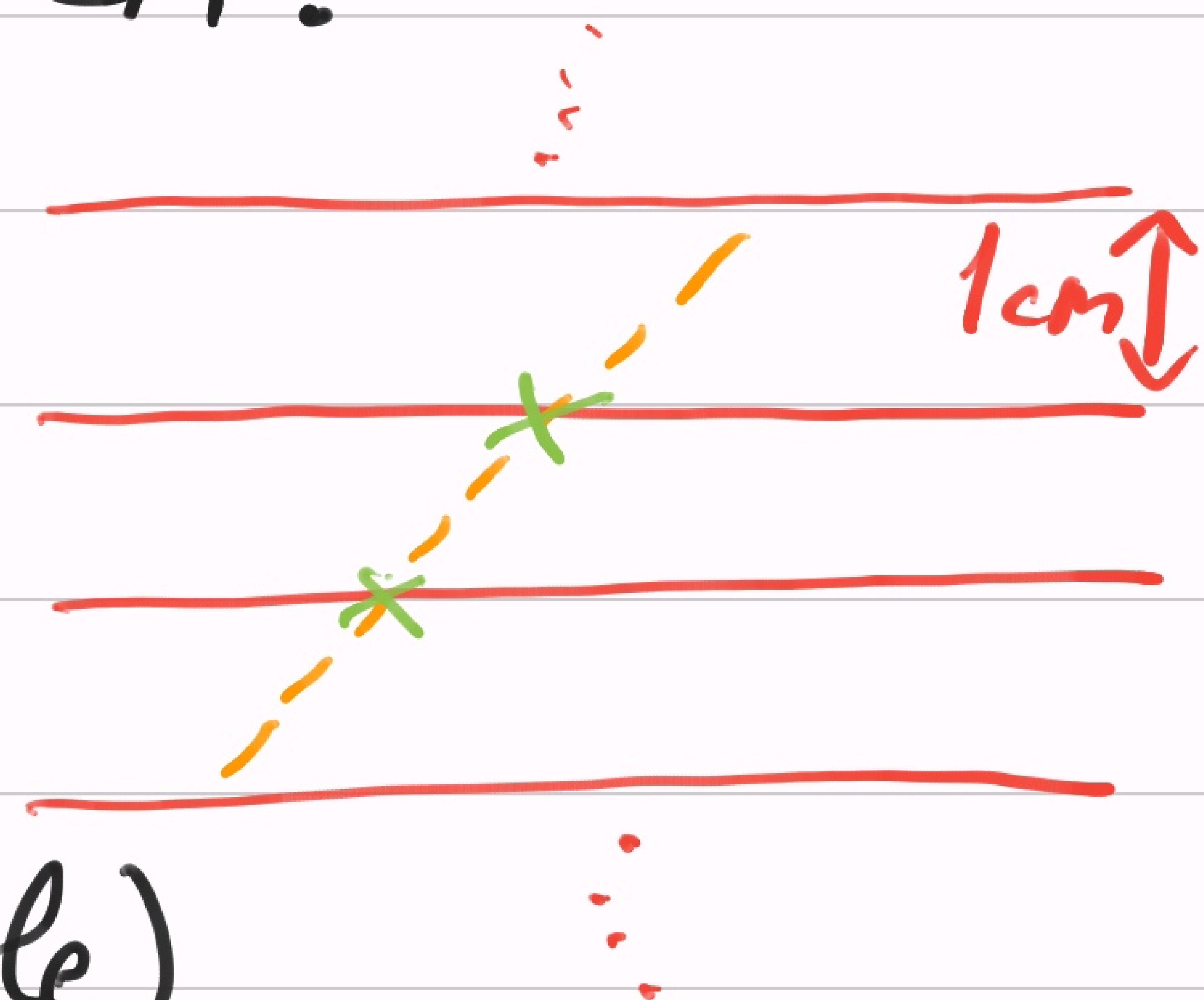
Then, $(\overline{X}_n - \mu)$, for <u>large</u> n, behaves like the standard-normal distribution!

- Let's see a stunning eg. in the continuous domain:

- eg. Buffon's Needle Problem:
  - Say on a floor, with parallel lines 1cm apart, you drop a needle of length 1cm.

    Let $X :=$ # intersections.

    What's $E[X] = ?$



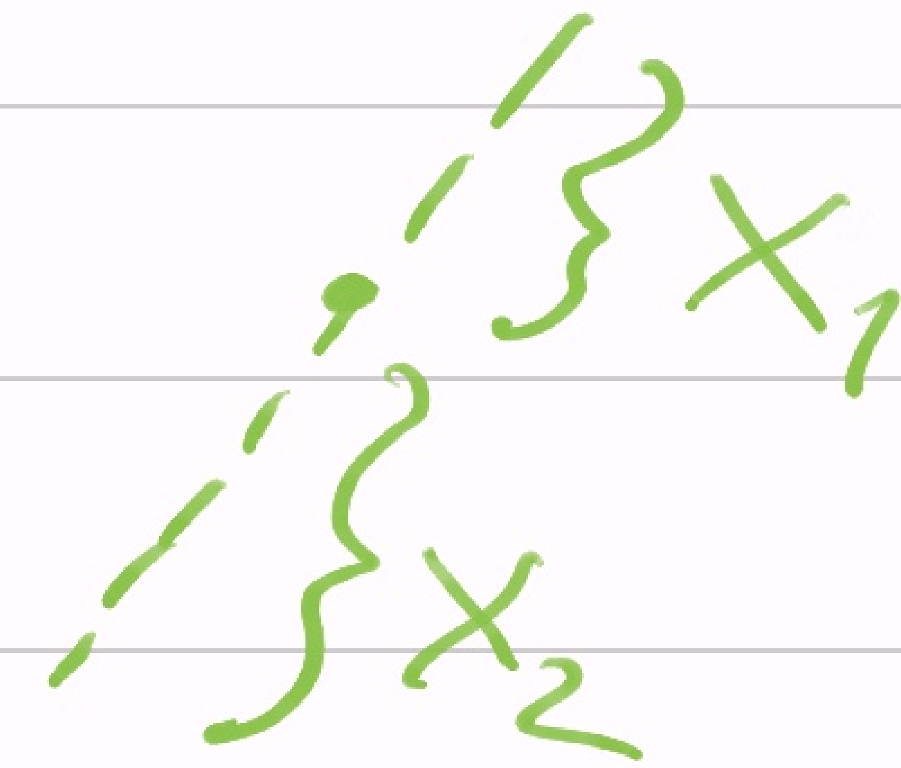
(Assume floor to be infinite.)

- Think of Ω as the set of (centre of the needle, orientation-angle).

Key insight: "Break" the needle into two parts. Consider #intersections in each; say X_1, X_2 resp.

▷ $X = X_1 + X_2$.

$\Rightarrow$ ▷ $E[X] = E[X_1] + E[X_2]$.



▷ This allows us to think of the two parts, being independently dropped.

• Doing this partition a large #times, we could transform the line into any shape; eg. circle.

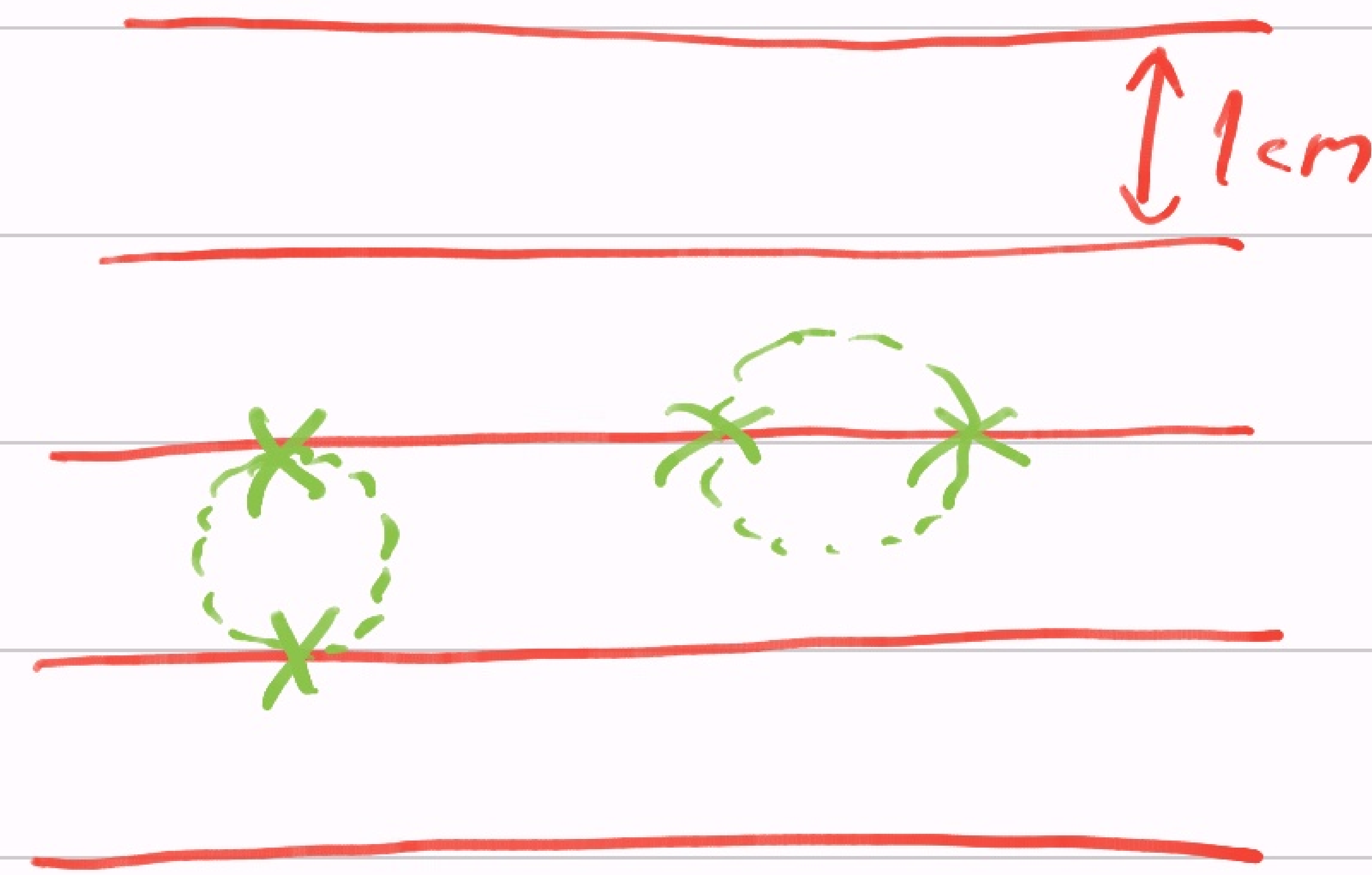
▷ Also, we can choose a length that suits our proof.

$\Rightarrow$ Pick a shape that gives the same # intersections wherever it drops!

▷ Circle of diameter = 1cm always gives 2 intersections!

▷ $E[X'] = 2$, when 'length' $= 2\pi \cdot \frac{1}{2} = \pi$ cm.

▷ $E[X] = \frac{2}{\pi} \cdot \ell$. $\square$



- Let's see an example that's very useful in computer science: