CS70 – Spring 2024

Lecture 22 – April 9

Summary of Last Lecture

- Markov's inequality (for non-neg. r.v.'s)

  $$\Pr[X \geq c] \leq \frac{1}{c} E[X]$$

- Chebyshev's inequality (for all r.v.'s)

  $$\Pr[|X - E[X]| \geq c] \leq \frac{1}{c^2} Var(X)$$

  $$\Pr[|X - E[X]| \geq c\sigma(X)] \leq \frac{1}{c^2}$$

Summary of Last Lecture (cont.)

- <u>Statistical estimation</u> :
  $X_1, X_2, \ldots, X_N$ are i.i.d. r.v.'s with expectation $E[X_i] = \mu$, variance $Var(X_i) = \sigma^2$

  Estimate of μ is : $\hat{\mu} = \frac{1}{N}(X_1 + \cdots + X_N)$

  <u>Thm</u> : If we take $\boxed{N \geq \frac{\sigma^2}{\mu^2} \cdot \frac{1}{\varepsilon^2 \delta}}$ samples, then

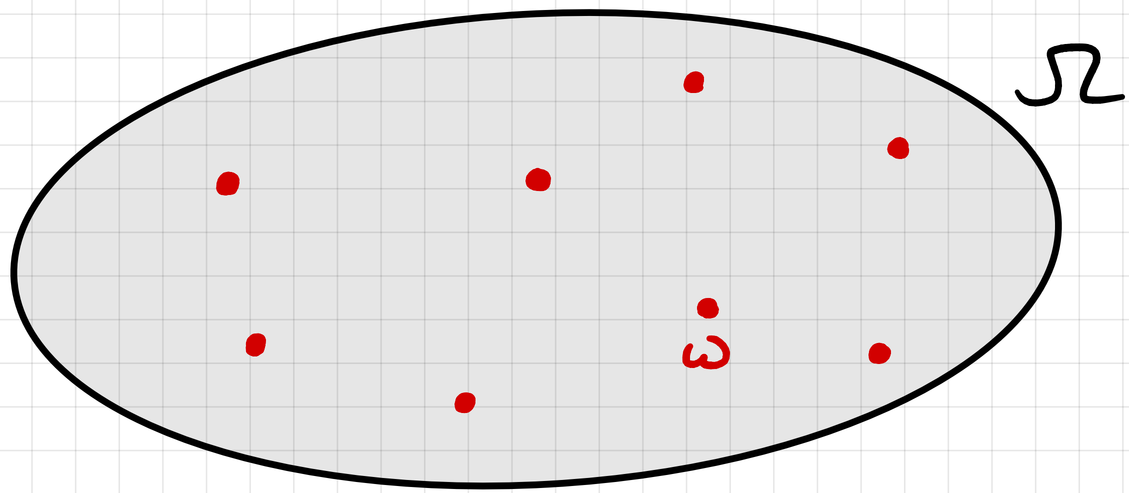
$$\Pr[|\hat{\mu} - \mu| \geq \varepsilon\mu] \leq \delta$$

- This is (a quantitative version of) the <u>Law of Large Numbers</u>

Continuous Probability

Up to now all our probability spaces were discrete i.e., finite or countably infinite

- Specify $\Pr[\omega]$ for each $\omega \in \Omega$
- $0 \le \Pr[\omega] \le 1$
- $\sum_{\omega \in \Omega} \Pr[\omega] = 1$



Note: This implies all random variables are also discrete (i.e., take on at most countably many values, e.g., 0, 1, 2, 3, ...)

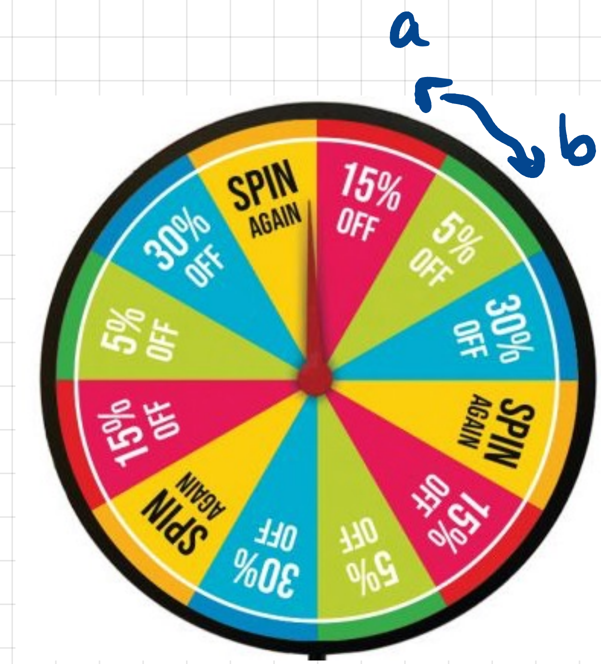
What if our prob. space is <u>uncountable</u>?

E.g. "wheel of fortune"

Pointer can end up at <u>any</u> position in $[0, \ell]$, where ℓ = circumference of wheel (or, equivalently, at any angle in $[0, 2\pi]$) $\longrightarrow$ <u>uncountably</u> many outcomes

Compare roulette wheel: only 38 outcomes

How do we assign probabilities to outcomes?



- For each $\omega \in [0, \ell]$,

  $\Pr[\omega] = \;??$

- $\displaystyle\sum_{\omega \in [0,\ell]} \Pr[\omega] = 1 \quad ??$

<u>Solution</u>: Instead assign probabilities to <u>intervals</u>:

for $0 \leq a < b \leq \ell$,

$$\boxed{\Pr\big[[a,b]\big] = \frac{\text{length of } [a,b]}{\text{length of } [0,\ell]} = \frac{b-a}{\ell}}$$

Solution: Instead assign probabilities to **intervals**: for $0 \le a < b \le \ell$,

$$\boxed{\Pr[[a,b]] = \frac{\text{length of } [a,b]}{\text{length of } [0,\ell]} = \frac{b-a}{\ell}}$$

These intervals are now our **atomic/basic events** (replacing sample points ω before)

Note that $\boxed{\Pr[[0,\ell]] = 1}$ and $\boxed{\Pr[a] = \Pr[[a,a]] = 0}$

We can then compute the probability of any event that can be expressed in terms of intervals – e.g. $\Pr[\bigcup I_i] = \sum_i \Pr[I_i]$ for disjoint intervals I_i

General theory of continuous prob. spaces $\longrightarrow$ $\boxed{\text{measure theory}}$

Continuous Random Variables

E.g. let X = position of pointer in wheel of fortune

Range of X is the continuous interval $[0, \ell]$

Again, $\Pr[X=a] = 0 \quad \forall a$

But we can define $\Pr[a \le X \le b] = \frac{b-a}{\ell}$

To make this more general, we need the idea of probability density

Definition: A *probability density function* (p.d.f.) for a continuous r.v. X is a function $f : \mathbb{R} \to \mathbb{R}$ satisfying:

- $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
- $\int_{-\infty}^{\infty} f(x)\,dx = 1$

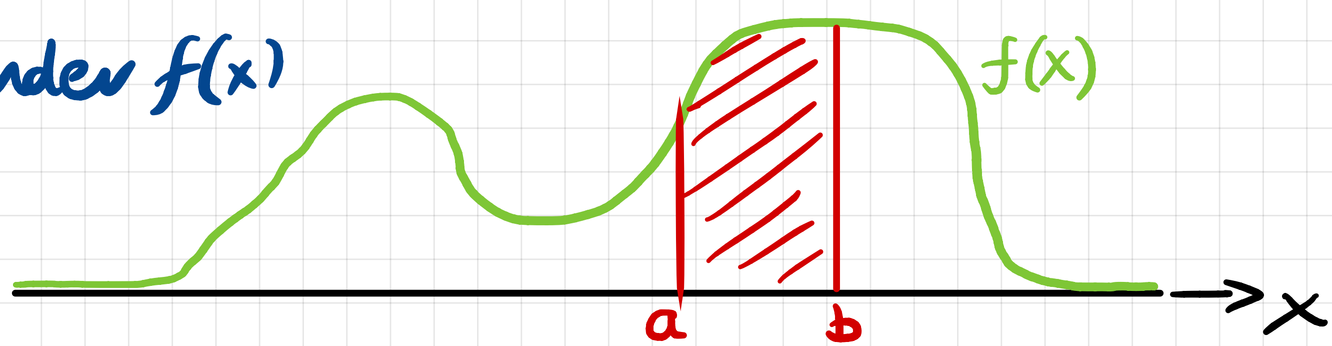
Then the *distribution* of X is defined by

$$\Pr[a \leq X \leq b] = \int_a^b f(x)\,dx \quad \forall a < b$$

total area under $f(x)$

$= \int_{-\infty}^{\infty} f(x)\,dx$

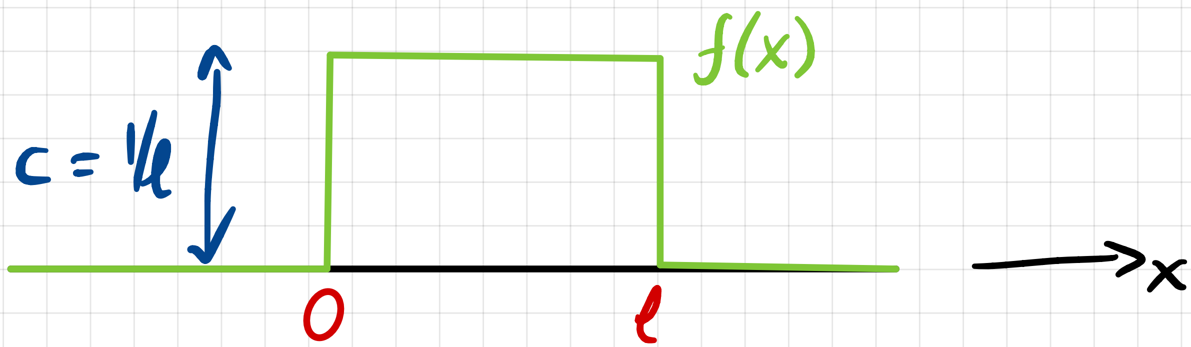
$= 1$



Example : Wheel of fortune

Here X is uniform on $[0,\ell]$, i.e., $\Pr[a \le X \le b] = \frac{b-a}{\ell}$

P.d.f. :

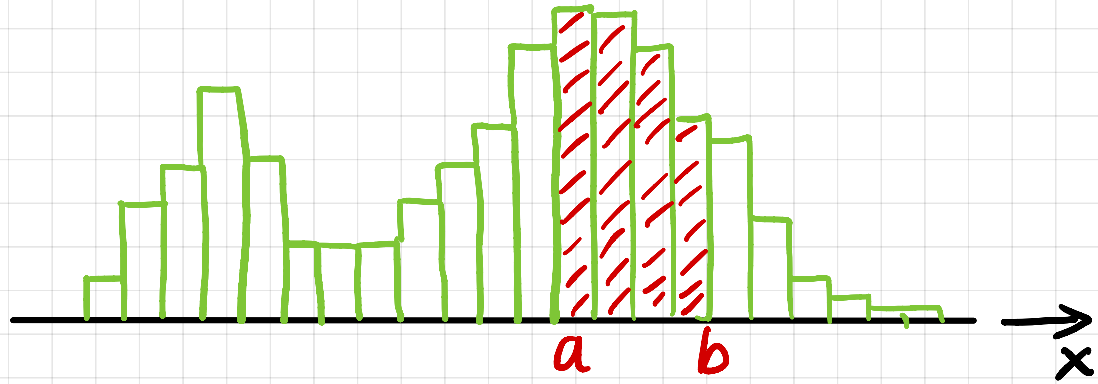


$$\boxed{f(x) = \begin{cases} 0 & x < 0 \\ c & 0 \le x \le \ell \\ 0 & x > \ell \end{cases}}$$

$$\int_{-\infty}^{\infty} f(x)\,dx = c\ell = 1 \quad \Rightarrow \quad \boxed{c = 1/\ell}$$

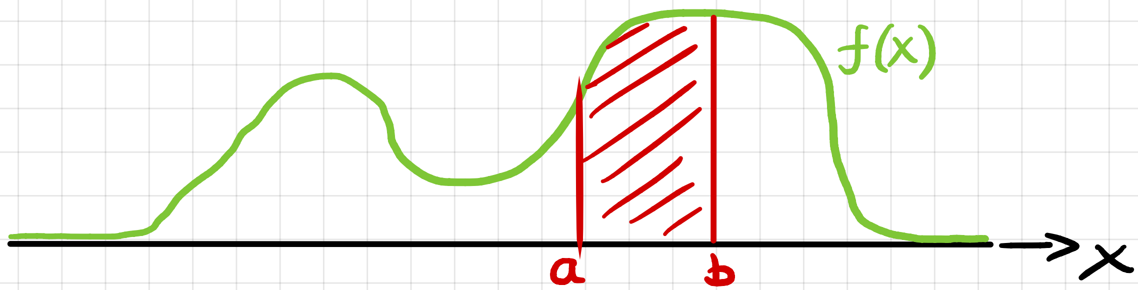
For $0 \le a < b \le \ell$: $\Pr[a \le X \le b] = \int_a^b f(x)\,dx = cx\Big|_a^b = \frac{b-a}{\ell}$

Comparison with discrete distributions



histogram

$$\Pr[a \le X \le b] = \sum_{a \le i \le b} \Pr[X = i]$$



p.d.f.

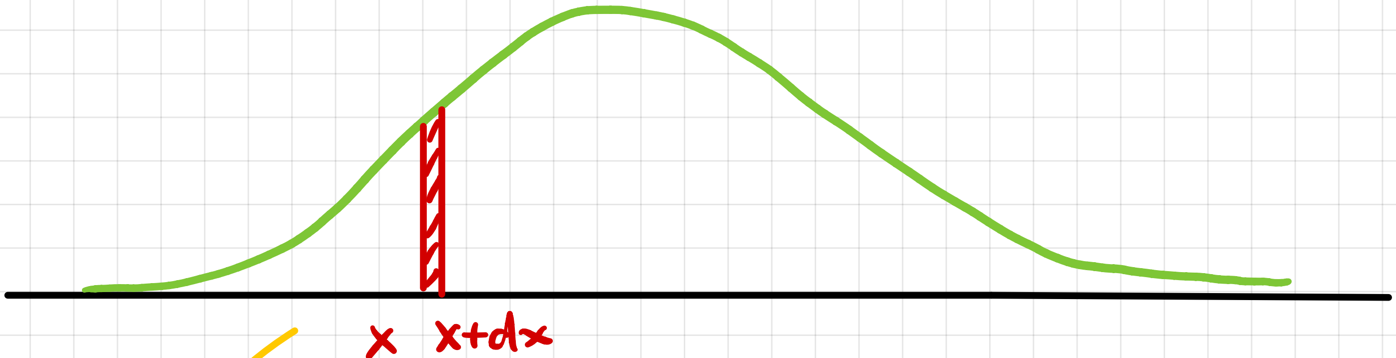
$$\Pr[a \le X \le b] = \int_a^b f(x)\,dx$$

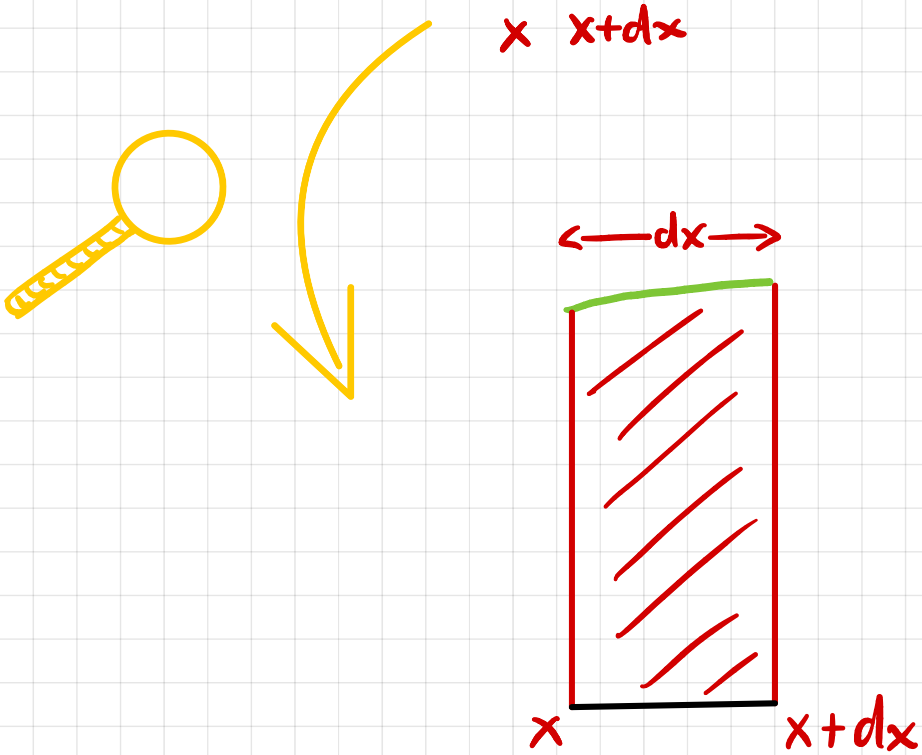
BUT NOTE: $f(x)$ is NOT a probability!

E.g. can have $f(x) > 1$!!

Instead, $f(x)$ is the probability density at x

Probability Density





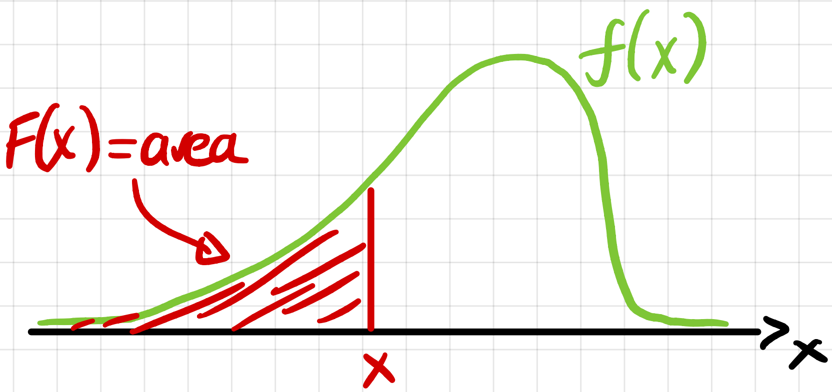
$$\Pr[x \le X \le x+dx] = \int_x^{x+dx} f(x)\,dx$$

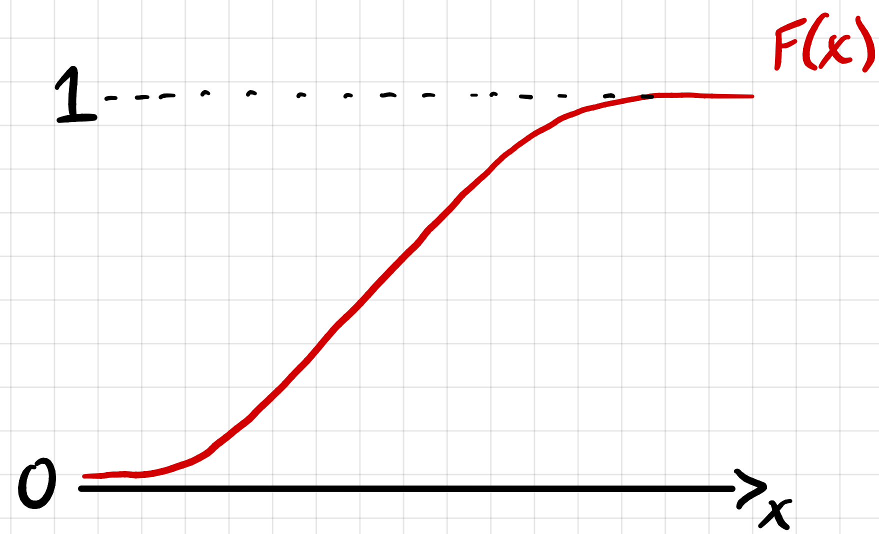
$$\approx f(x)\,dx$$

$f(x) =$ "probability per unit length" at x (density)

Definition: The cumulative distribution function (c.d.f.) of a continuous r.v. X is defined by

$$F(x) := \Pr[X \leq x] = \int_{-\infty}^{x} f(z)\, dz$$

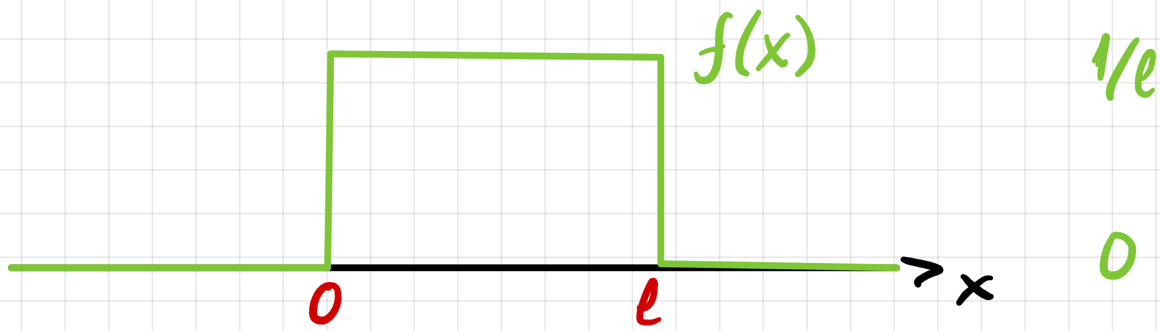




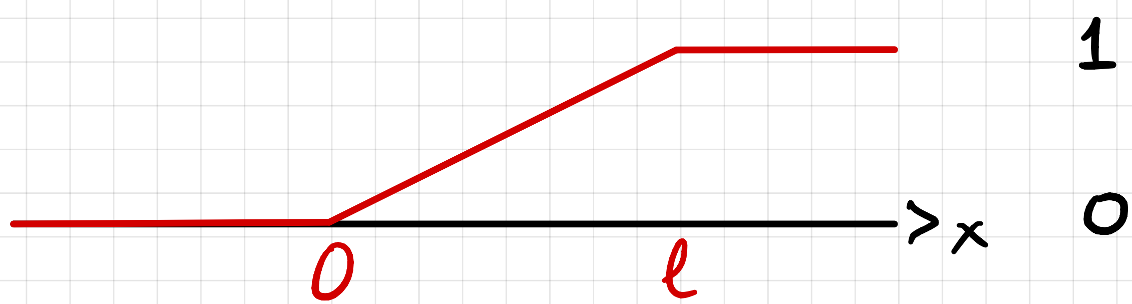
Note:

- $F(x)$ increases monotonically to 1 as $x \to \infty$
- $f(x) = \dfrac{dF(x)}{dx}$
- Can use either $f(x)$ or $F(x)$ to define r.v. X

Example : Wheel of fortune



$$f(x) = \begin{cases} 0 & x < 0 \\ 1/\ell & 0 \le x \le \ell \\ 0 & x > \ell \end{cases}$$



$$F(x) = \begin{cases} 0 & x < 0 \\ x/\ell & 0 \le x \le \ell \\ 1 & x > \ell \end{cases}$$

$$f(x) = \frac{dF(x)}{dx}$$

Expectation & Variance

Defn: The *expectation* of a continuous r.v. X with pdf f is

$$\boxed{E[X] := \int_{-\infty}^{\infty} x f(x)\, dx}$$

[Compare: $E[X] = \sum_a a \cdot \Pr[X=a]$]

Defn: The *variance* of X is

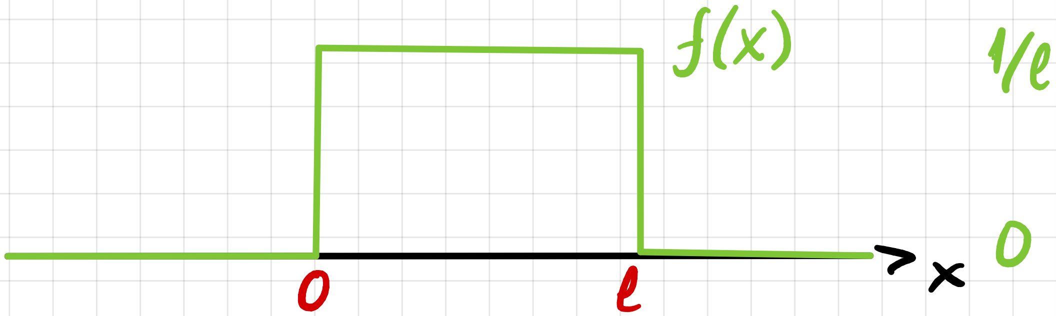
$$\mathrm{Var}(X) := E\left[(X - E[X])^2\right] = E[X^2] - E[X]^2$$

$$= \boxed{\int_{-\infty}^{\infty} x^2 f(x)\, dx - E[X]^2}$$

Generally: For a function $G : \mathbb{R} \to \mathbb{R}$,

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

Example: Wheel of fortune



$$f(x) = \begin{cases} 0 & x < 0 \\ 1/\ell & 0 \le x \le \ell \\ 0 & x > \ell \end{cases}$$

$$E[X] = \int_{-\infty}^{\infty} x f(x)\, dx = \int_0^{\ell} \frac{x}{\ell}\, dx = \left.\frac{x^2}{2\ell}\right|_0^{\ell} = \boxed{\frac{\ell}{2}}$$

$$Var(X) = \int_{-\infty}^{\infty} x^2 f(x)\, dx - E[X]^2$$

$$= \int_0^{\ell} \frac{x^2}{\ell}\, dx = \left.\frac{x^3}{3\ell}\right|_0^{\ell} = \frac{\ell^2}{3}$$

$$\Rightarrow Var(X) = E[x^2] - E[x]^2 = \frac{\ell^2}{3} - \frac{\ell^2}{4} = \boxed{\frac{\ell^2}{12}}$$

Compare: discrete uniform distribution on $[0, \ell-1]$
(assuming ℓ integer)

i.e., $\Pr[X=i] = \frac{1}{\ell}$ for $i = 0, 1, \ldots, \ell-1$

$$E[X] = \frac{1}{\ell}\left[0+1+2+\cdots+\ell-1\right] = \frac{1}{\ell}\cdot\frac{\ell(\ell-1)}{2} = \frac{\ell-1}{2}$$

$$\mathrm{Var}(X) = E[X^2] - E[X]^2$$

$$= \frac{1}{\ell}\left[0+1+4+\cdots+(\ell-1)^2\right]$$

$$= \frac{1}{\ell}\,\frac{(\ell-1)\ell(2\ell-1)}{6} = \frac{(\ell-1)(2\ell-1)}{6}$$

$$\Rightarrow \mathrm{Var}(X) = \frac{(\ell-1)(2\ell-1)}{12} - \frac{(\ell-1)^2}{4} = \frac{\ell^2-1}{12}$$

Markov's Inequality

Thm: For a continuous r.v. with p.d.f. f satisfying $f(x) = 0$ for $x < 0$:

$$\boxed{\Pr[X \geq c] \leq \frac{E[X]}{c}}$$

Chebyshev's Inequality

Thm: For a continuous r.v. X:

$$\boxed{\Pr\left[|X - E[X]| \geq c\right] \leq \frac{Var(X)}{c^2}}$$

Joint Distributions

Defn: A joint density function for two r.v.'s X, Y is a function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ satisfying:

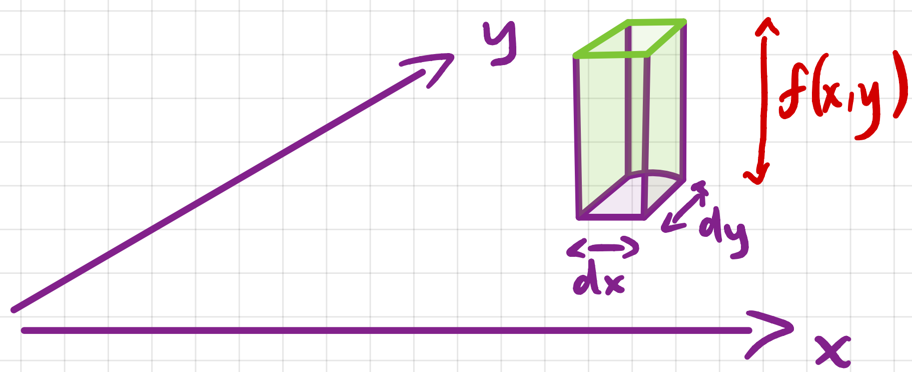
- $f(x,y) \geq 0 \quad \forall x, y \in \mathbb{R}$
- $\displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$

The joint distribution of X, Y is then

$$\boxed{\Pr[a \leq X \leq b,\ c \leq Y \leq d] = \int_c^d \int_a^b f(x,y)\,dx\,dy}$$

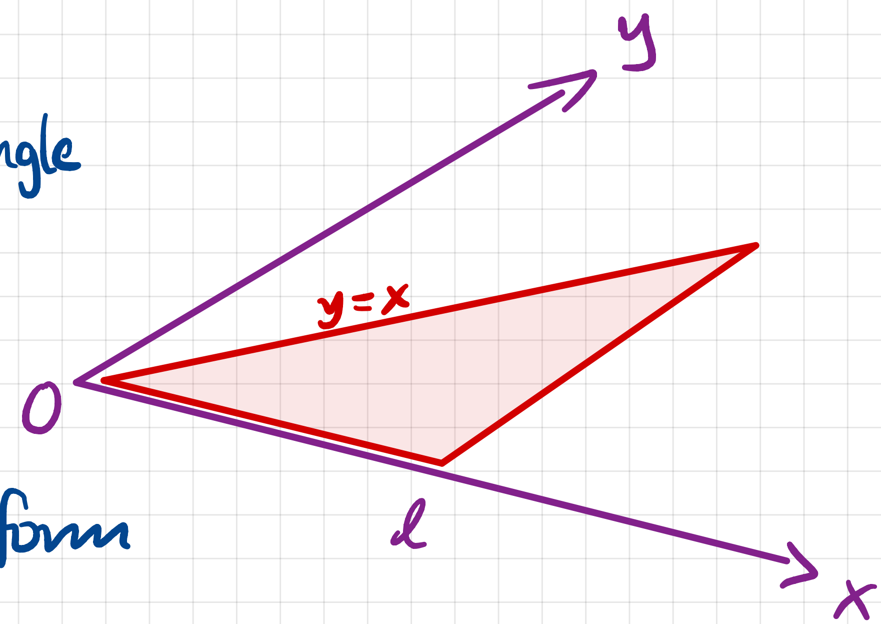
Interpretation of $f(x,y)$:

prob. density per unit area at (x,y)

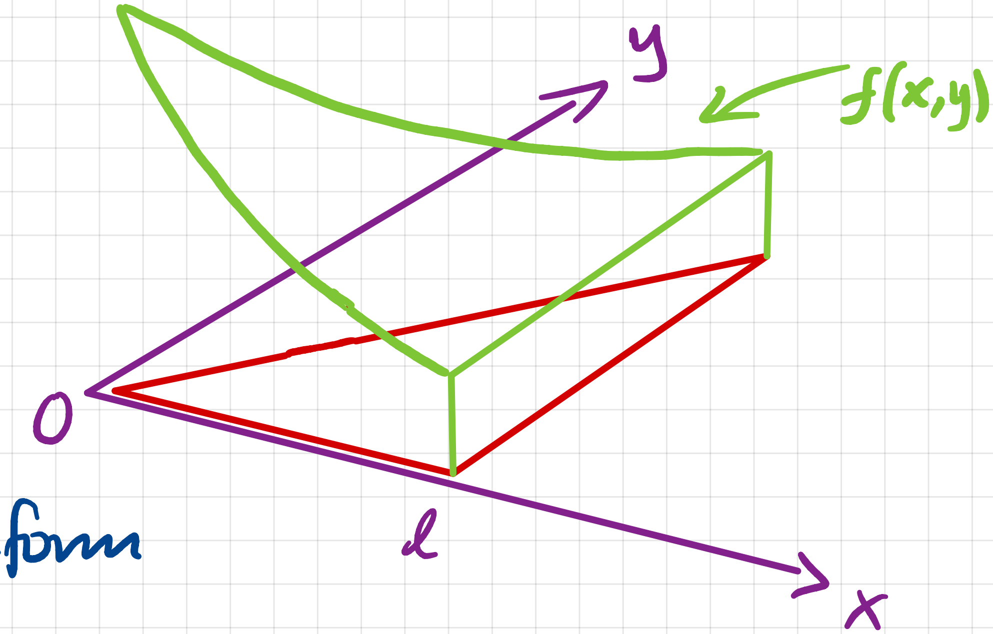


Example: Two-round game

- **Round 1:** You stake \$$\ell$ and win amount X uniform in $[0, \ell]$
- **Round 2:** You stake \$$X$ and win amount Y uniform in $[0, X]$



- $f(x,y) = 0$ outside red triangle
- Density of x is uniform on $[0, \ell]$
- Given x, density of y is uniform on $[0, x]$
- $$f(x,y) = \begin{cases} 1/\ell x & \text{for } (x,y) \in \triangle \\ 0 & \text{otherwise} \end{cases}$$



- $f(x,y) = 0$ outside red $\triangle$
- Density of x is uniform on $[0, \ell]$
- Given x, density of y is uniform on $[0, x]$
- $$f(x,y) = \begin{cases} 1/\ell x & \text{for } (x,y) \in \triangle \\ 0 & \text{otherwise} \end{cases}$$

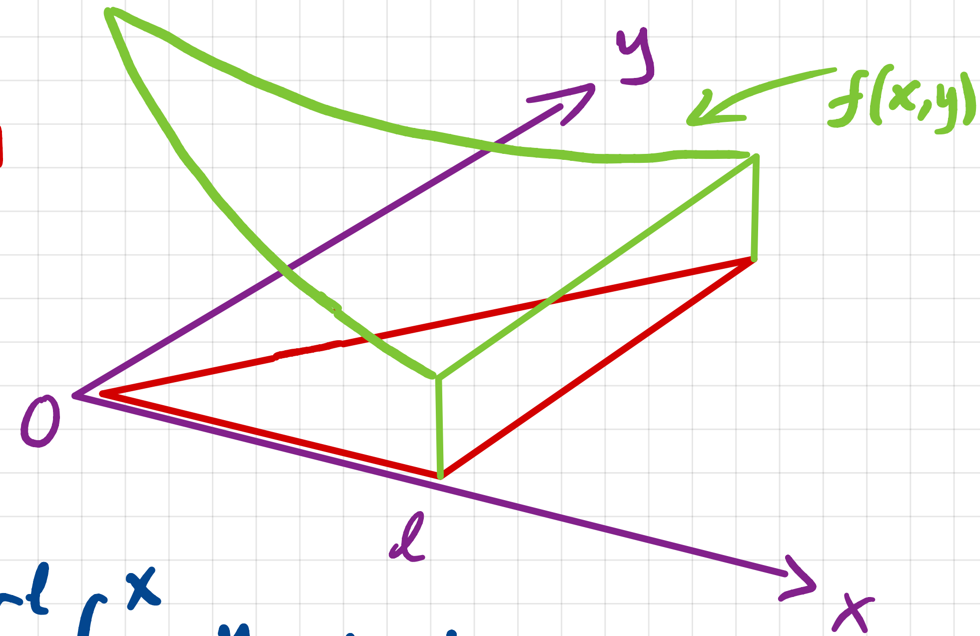
<u>Check</u>:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^\ell \left(\int_0^x \frac{1}{\ell x}\,dy \right) dx$$

$$= \int_0^\ell \left(\frac{y}{\ell x} \Big|_0^x \right) dx$$

$$= \int_0^\ell \frac{1}{\ell}\,dx = \frac{x}{\ell} \Big|_0^\ell = 1 \quad \checkmark$$

• $f(x,y) = \begin{cases} 1/\ell x & \text{for } (x,y) \in \Delta \\ 0 & \text{otherwise} \end{cases}$



$$E[y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\, dx\, dy = \int_0^{\ell}\int_0^{x} \frac{y}{\ell x}\, dy\, dx$$

$$= \int_0^{\ell} \left(\frac{y^2}{2\ell x} \Big|_0^x \right) dx$$

$$= \int_0^{\ell} \frac{x}{2\ell}\, dx$$

$$= \frac{x^2}{4\ell}\Big|_0^{\ell} = \boxed{\frac{\ell}{4}}$$

Independence

Defn: Continuous r.v.'s X, Y are <u>independent</u> if

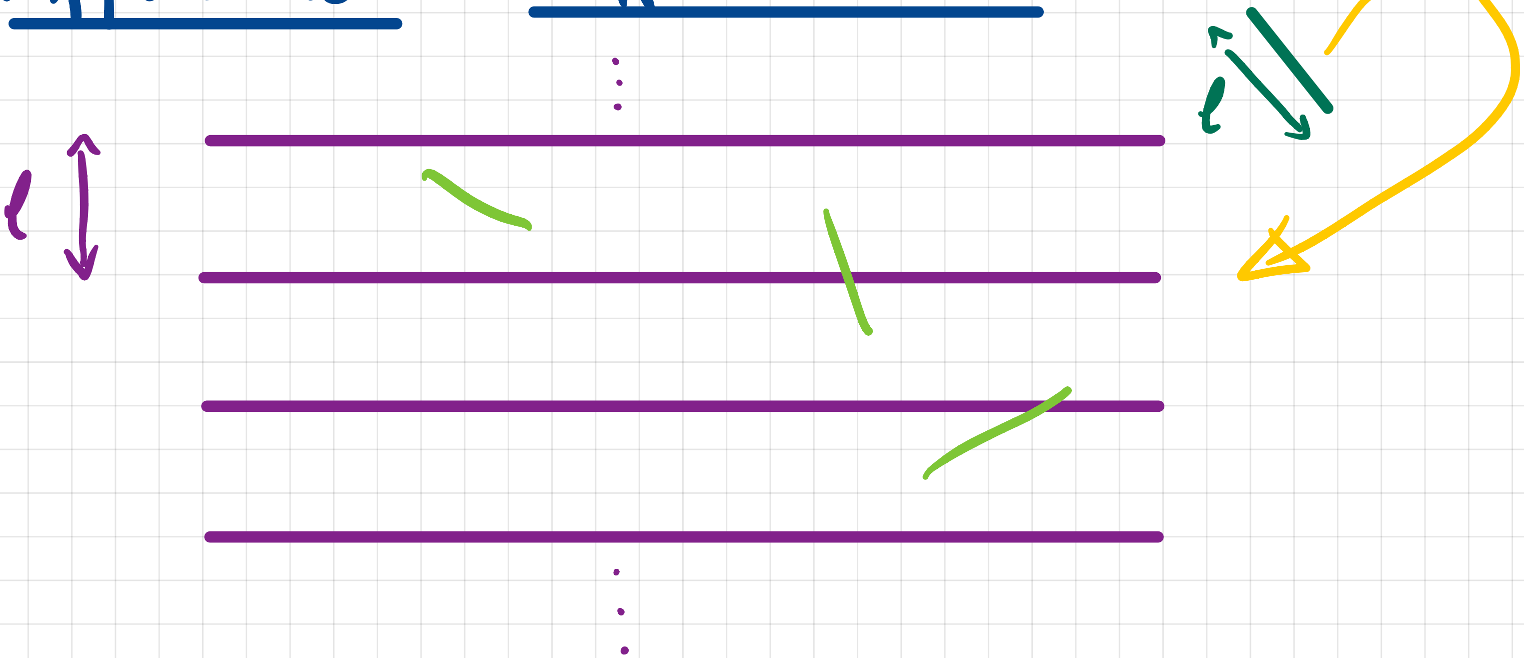
$$\Pr[a \le X \le b,\ c \le Y \le d] = \Pr[a \le X \le b]\,\Pr[c \le Y \le d]$$

$$\forall\ a < b,\ c < d$$

Thm: If X, Y are independent with pdf's $f(x)$, $g(y)$ respectively, then their joint density $h(x,y)$ is given by

$$\boxed{h(x,y) = f(x)g(y) \qquad \forall x, y \in \mathbb{R}}$$

Application: Buffon's Needle



- board with lines dist. ℓ apart
- needle length ℓ
- throw needle randomly onto board
- let $X = \begin{cases} 1 & \text{if needle hits a line} \\ 0 & \text{otherwise} \end{cases}$

Claim: $E[X] = \dfrac{2}{\pi}$

$$X = \begin{cases} 1 & \text{if needle hits a line} \\ 0 & \text{otherwise} \end{cases}$$

<u>Claim</u>: $E[X] = \frac{2}{\pi}$

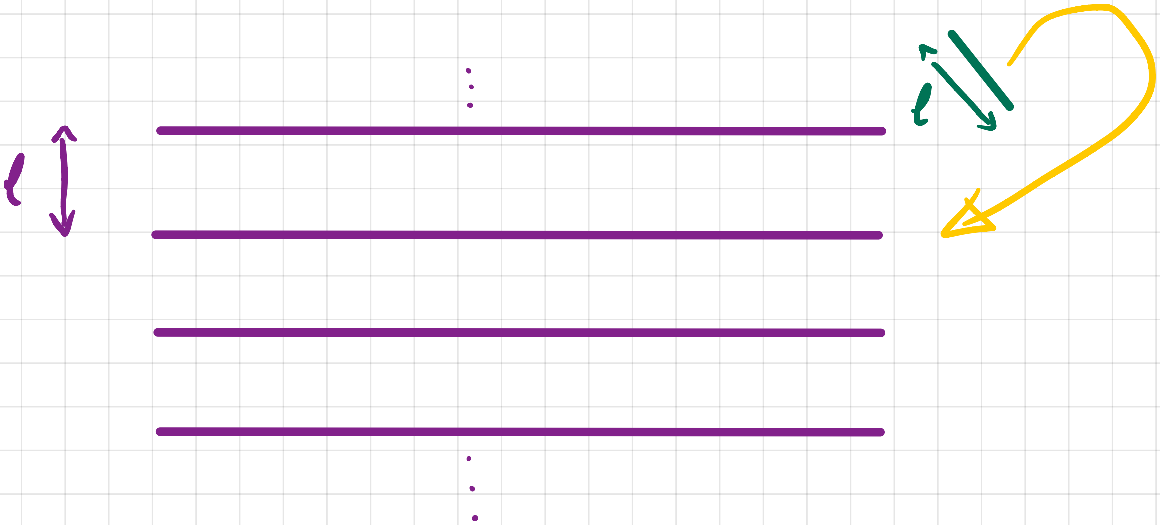
If Claim is true then we can estimate π as in previous lecture!

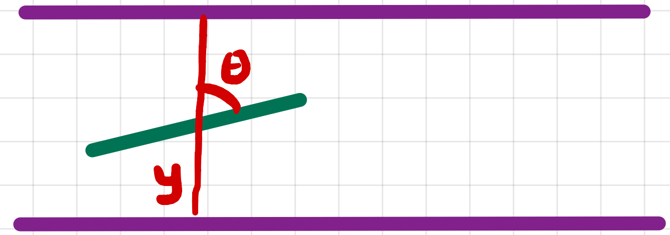
Perform experiment N times $\longrightarrow$ $X_1, \ldots, X_N$ (i.i.d.)

Output $\boxed{\hat{p} = \frac{1}{N}(X_1 + \ldots + X_N)}$

Then $E[\hat{p}] = \frac{2}{\pi}$ $\Rightarrow$ $\frac{2}{\hat{p}}$ estimates π

Number of trials needed for accuracy $(1 \pm \varepsilon)\pi$ with confidence $1-\delta$ is (by Chebyshev) $\leq \frac{\pi}{2} \cdot \frac{1}{\varepsilon^2 \delta} \leq \boxed{\frac{2}{\varepsilon^2 \delta}}$

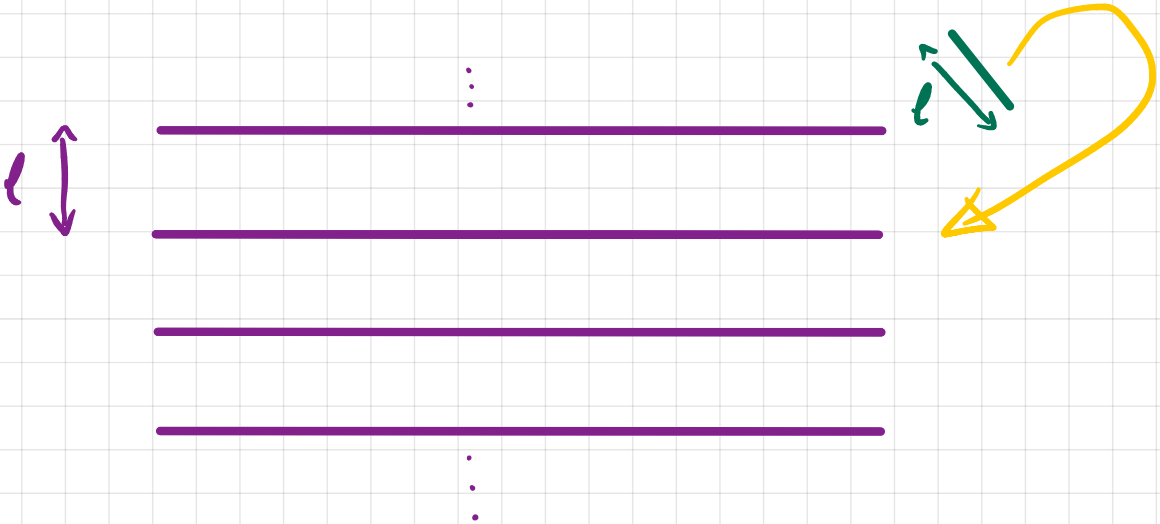


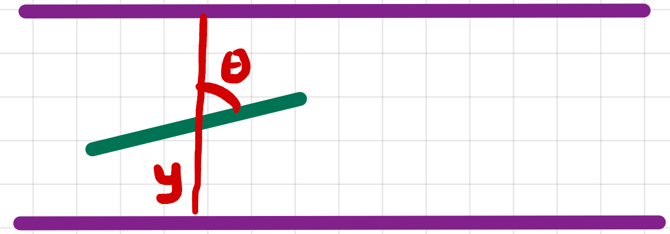


Outcome of throw described by 2 random variables:

y := dist. between needle midpoint & closest line $\quad 0 \leq y \leq \ell/2$

θ := angle between needle & vertical $\quad -\pi/2 \leq \theta \leq \pi/2$





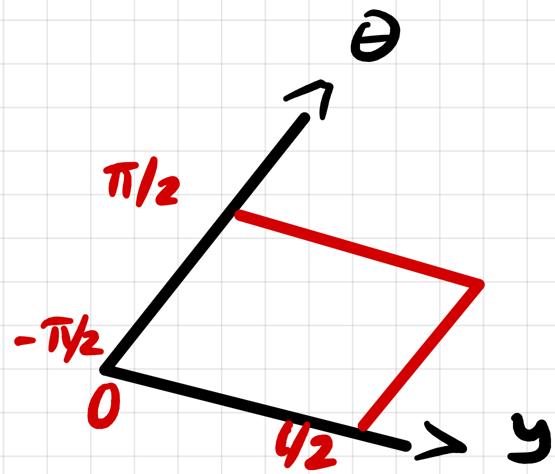
Outcome of throw described by 2 random variables:

y := dist. between needle midpoint & closest line $\quad 0 \leq X \leq \ell/2$

θ := angle between needle & vertical $\quad -\pi/2 \leq y \leq \pi/2$

Joint density $f(y, \theta)$ uniform over rectangle $[0, \ell/2] \times [-\pi/2, \frac{\pi}{2}]$

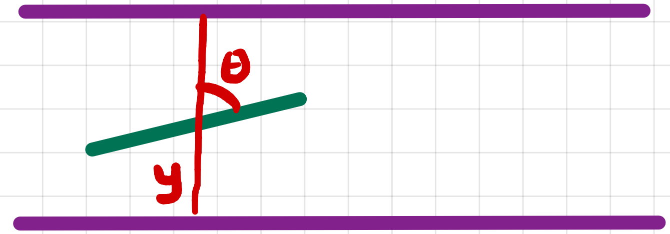
$$\Rightarrow f(y,\theta) = \begin{cases} 2/\pi\ell & (y,\theta) \in \square \\ 0 & \text{otherwise} \end{cases}$$



$$\left[\frac{\pi\ell}{2} = \text{area of } \square\right]$$

$$f(y,\theta) = \begin{cases} \frac{2}{\pi \ell} & (y,\theta) \in \square \\ 0 & \text{otherwise} \end{cases}$$

$$X = \begin{cases} 1 & \text{if needle hits a line} \\ 0 & \text{otherwise} \end{cases}$$



$$\boxed{\underline{\text{Claim}}: \ E[X] = \frac{2}{\pi}}$$

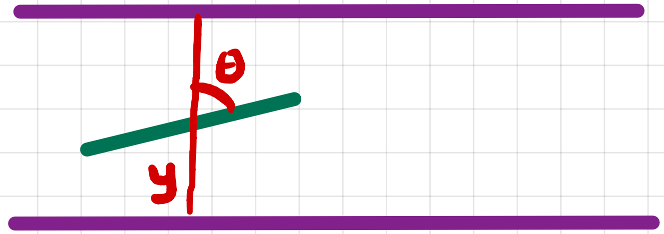
Note that $E[X] = \Pr[E]$ where E is event "needle hits line"

Q: When does E happen?

A: When $y \le \frac{\ell}{2} \cos\theta$

$$f(y,\theta) = \begin{cases} 2/\pi \ell & (y,\theta) \in \square \\ 0 & \text{otherwise} \end{cases}$$

$$X = \begin{cases} 1 & \text{if needle hits a line} \\ 0 & \text{otherwise} \end{cases}$$



$$\boxed{\underline{\text{Claim}}:\ E[X] = \frac{2}{\pi}}$$

Note that $E[X] = \Pr[E]$ where E is event "needle hits line"

Q: When does E happen?

A: When $y \leq \frac{\ell}{2}\cos\theta$

$$\text{So } \Pr[E] = \Pr\left[y \leq \tfrac{\ell}{2}\cos\theta\right] = \int_{-\pi/2}^{\pi/2} \int_0^{\ell/2 \cos\theta} f(y,\theta)\, dy\, d\theta$$

$$= \int_{-\pi/2}^{\pi/2} \left(\frac{2y}{\pi \ell} \Big|_0^{\ell/2 \cos\theta} \right) d\theta = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \cos\theta\, d\theta = \frac{1}{\pi} \sin\theta \Big|_{-\pi/2}^{\pi/2} = \boxed{\frac{2}{\pi}}$$