CS70 @ UC Berkeley, Spring 2026

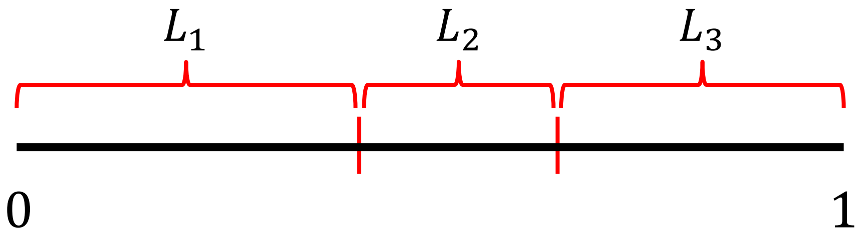
Lecture 25

Continuous Probability Distribution II

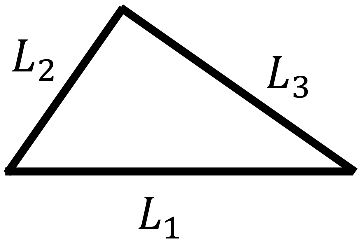
April 23, 2026

Random Breaks to Form a Triangle

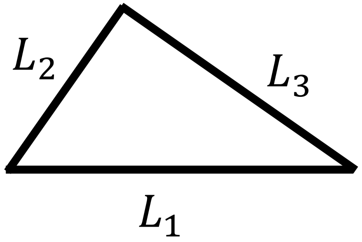
- Suppose Alice samples two points independently and uniformly at random from the interval [0,1], thereby obtaining 3 segments.



- What is the Probability that the three segments can form a triangle?



Random Breaks to Form a Triangle



$L_1 + L_2 + L_3 = 1$

Can form a triangle implies

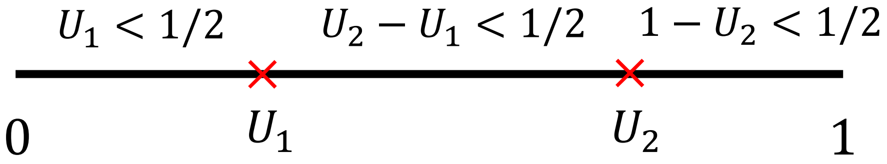
$L_1 < L_2 + L_3 \Rightarrow L_1 < 1 - L_1 \Rightarrow L_1 < \frac{1}{2}$

$L_2 < L_1 + L_3 \Rightarrow L_2 < 1 - L_2 \Rightarrow L_2 < \frac{1}{2}$

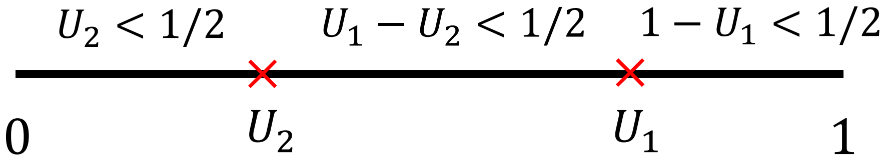
$L_3 < L_1 + L_2 \Rightarrow L_3 < 1 - L_3 \Rightarrow L_3 < \frac{1}{2}$

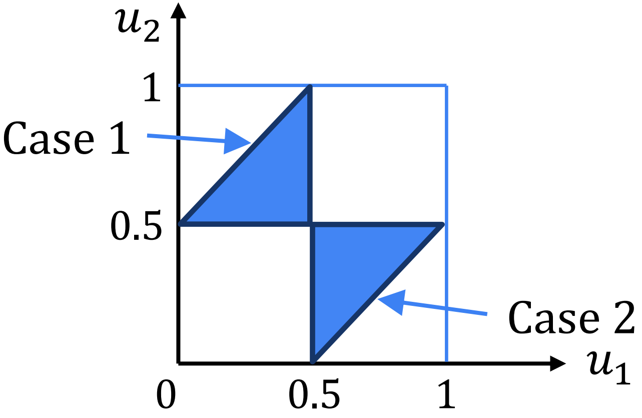
$U_1, U_2 \sim \text{Uniform}[0,1]$ and independent

Case 1



Case 2





$\mathbb{P}[(U_1, U_2) \in \text{Shared region})$

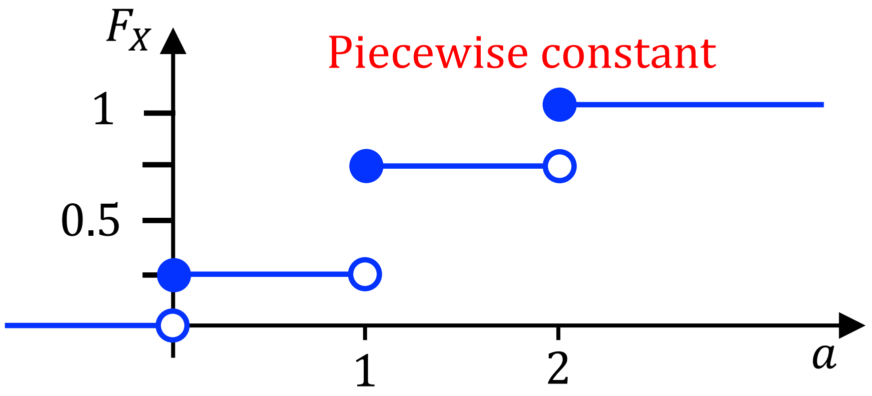
$= \text{ Area of shared region} = \frac{1}{4}$

Cumulative Distribution Function (CDF)

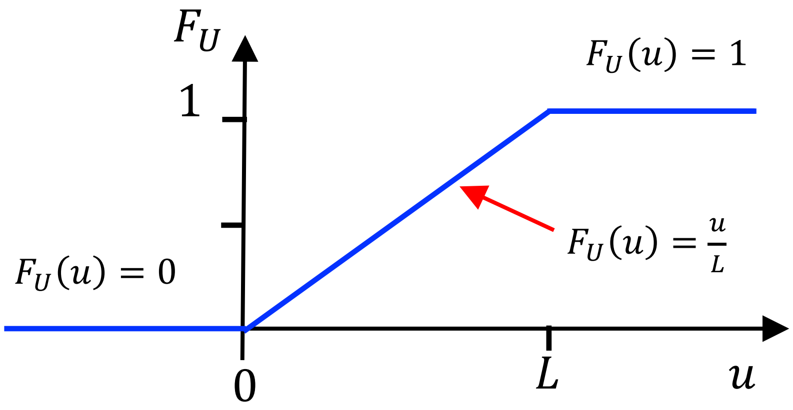
Definition (CDF). Given a random variable X, its cumulative distribution function F_X is defined as

$$F_X(a) = \mathbb{P}(X \leq a), \text{ for } a \in (-\infty, +\infty)$$

Example: Toss a fair coin twice.
$X(\omega) =$ Heads in $\omega \in \Omega$.



$U \sim$ Uniform$[0, L]$
Continuous CDF



Convergence in Distribution

CDF plays a central role in probability theory.

Definition (Convergence in distribution): $X_n \xrightarrow{d} X$ as $n \to \infty$
A sequence of random variables $X_1, X_2, X_3, \ldots$ is said to converge to another random variable X in distribution if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x),$$

for all $x \in \mathbb{R}$ where $F_X(x)$ is continuous.

Examples:

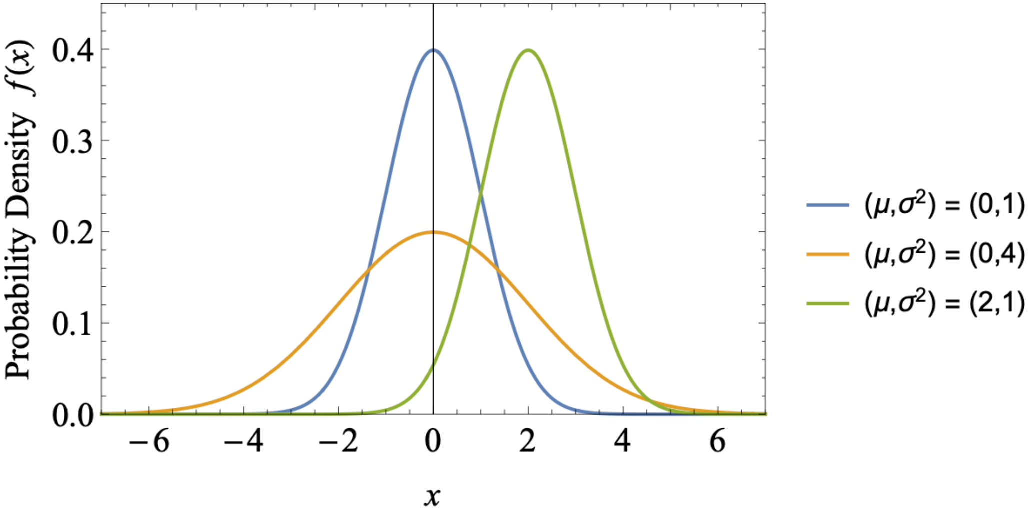
1. $X_n \sim \text{Binomial}(n, p_n)$, where $p_n = \lambda/n$. Then, $X_n \xrightarrow{d} X \sim \text{Poisson}(\lambda)$ as $n \to \infty$.
2. $Y_n \sim \text{Geometric}(p_n)$, where $p_n = \lambda/n$. Then, $\frac{Y_n}{n} \xrightarrow{d} Y \sim \text{Exp}(\lambda)$ as $n \to \infty$.

Normal Distribution

Definition (Normal Random Variable): $X \sim \mathrm{Normal}(\mu, \sigma^2)$
A continuous random variable X with probability density function

$$f_X(x) = \frac{1}{2\pi\sigma^2} e^{-(x-\mu)^2/(2\sigma^2)},$$

where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$.



The density becomes more spread out as σ^2 increases, while μ determines where the center is.

Normal Distribution

Definition (Normal Random Variable): $X \sim \text{Normal}(\mu, \sigma^2)$
A continuous random variable X with probability density function

$$f_X(x) = \frac{1}{2\pi\sigma^2} e^{-(x-\mu)^2/(2\sigma^2)},$$

where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$.

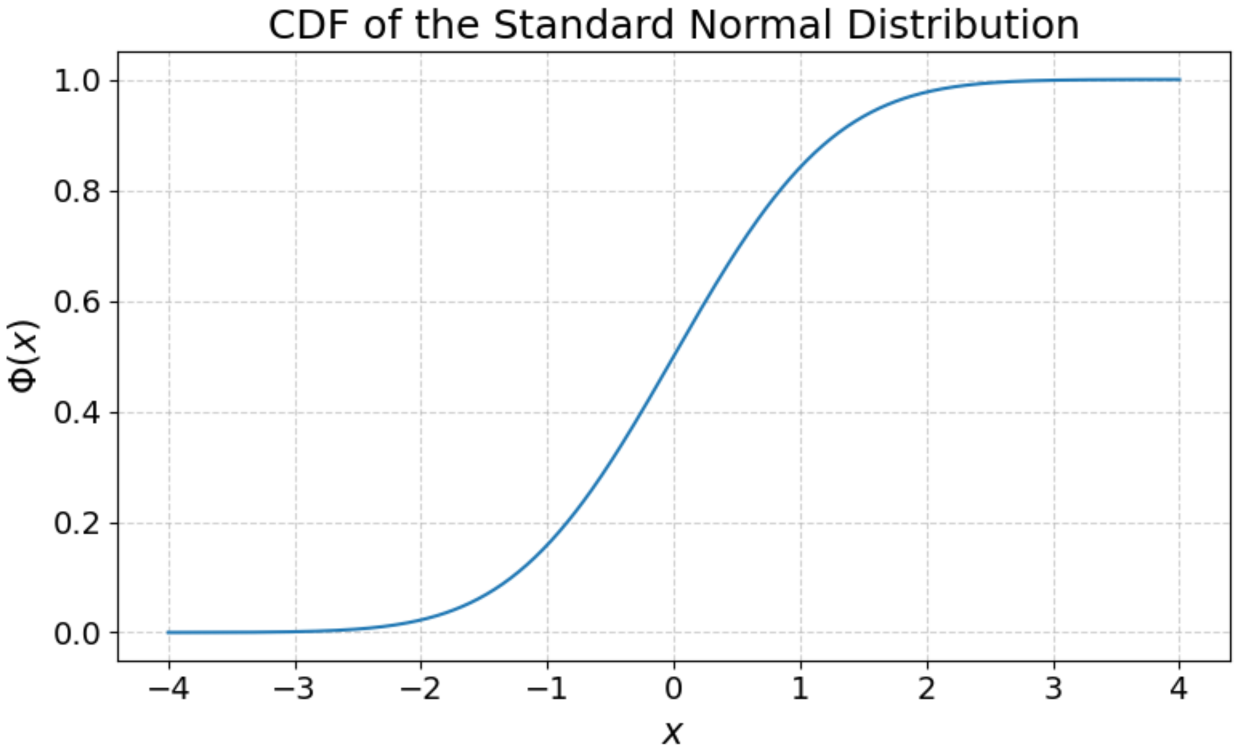
1. $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$
2. $\mathbb{E}[X] = \mu$
3. $\mathrm{Var}[X] = \sigma^2$

Proof of 2: Let $Y \sim \text{Normal}(0, \sigma^2)$. Then, $\mathbb{E}[Y] = 0$ by symmetry (i.e., $g(y) = y$ is an odd function of y, but $f_Y(y)$ is an even function of y, and the integration is over $(-\infty, +\infty)$.)

$$X = \mu + Y \sim \text{Normal}(\mu, \sigma^2) \Rightarrow \mathbb{E}[X] = \mu + \mathbb{E}[Y] = \mu.$$

CDF of the Standard Normal Distribution

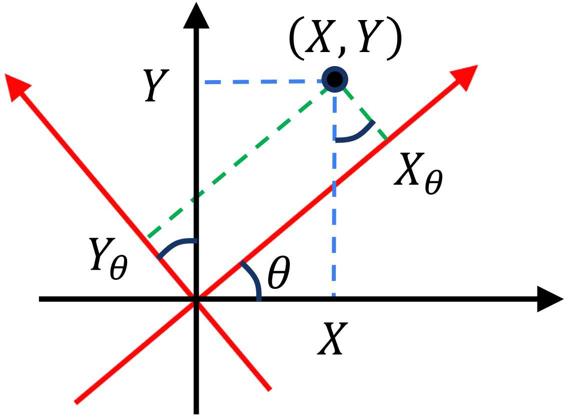
$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \mathrm{d}t$ denotes the c.d.f. of Normal(0,1).

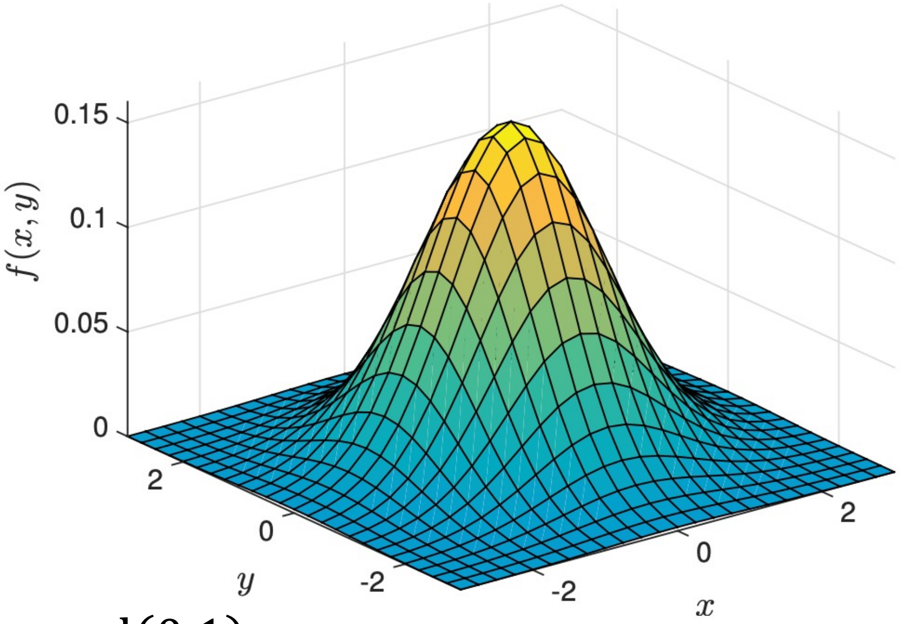


Sum of Independent Normal Random Variables

Suppose $X, Y \sim \text{Normal}(0,1)$ are independent.

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$





$$\boxed{\begin{aligned} X_\theta &= X\cos\theta + Y\sin\theta \\ Y_\theta &= -X\sin\theta + Y\cos\theta \end{aligned}}$$

Rotationally symmetric $\Rightarrow X_\theta, Y_\theta \sim \text{Normal}(0,1)$

Sum of Independent Normal Random Variables

Suppose $X, Y \sim \text{Normal}(0,1)$ are independent.

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

$$\boxed{\begin{aligned} X_\theta &= \quad X \cos\theta + Y \sin\theta \\ Y_\theta &= -X \sin\theta + Y \cos\theta \end{aligned}}$$

Rotationally symmetric $\Rightarrow$

$X_\theta, Y_\theta \sim \text{Normal}(0,1)$ and they are independent

$\Rightarrow \alpha X + \beta Y \sim \text{Normal}(0,1)$ if $\alpha^2 + \beta^2 = 1$

$\Rightarrow \sigma(\alpha X + \beta Y) \sim \text{Normal}(0, \sigma^2)$ if $\alpha^2 + \beta^2 = 1$ (follows from the p.d.f.)

Let $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}, \quad \alpha = \sigma_1/\sigma, \quad \beta = \sigma_1/\sigma$

Then, $U := \sigma\alpha X \sim \text{Normal}(0, \sigma_1^2)$, and $V := \sigma\beta Y \sim \text{Normal}(0, \sigma_2^2)$

U and V are independent since X and Y are independent.

$U + V \sim \text{Normal}(0, \sigma_1^2 + \sigma_2^2)$

More generally, thinking about $U' = U + \mu_1$ and $V' = \text{V} + \mu_2$, we conclude that if $U' \sim \text{Normal}(\mu_1, \sigma_1^2)$ and $V' \sim \text{Normal}(\mu_2, \sigma_2^2)$ are independent, then

$$U' + V' \sim \text{Normal}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Fun Fact about the Normal Distribution

Let X and Y be arbitrary random variables on the same probability space.

$$\begin{pmatrix} X_\theta \\ Y_\theta \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}}_{M} \begin{pmatrix} X \\ Y \end{pmatrix}$$

$$M^T M = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$M M^T = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

M is an orthogonal matrix

$$X_\theta^2 + Y_\theta^2 = (X_\theta \quad Y_\theta) \begin{pmatrix} X_\theta \\ Y_\theta \end{pmatrix} = (X \quad Y) M^T M \begin{pmatrix} X \\ Y \end{pmatrix} = (X \quad Y) \begin{pmatrix} X \\ Y \end{pmatrix} = X^2 + Y^2$$

$\Rightarrow$ Every orthogonal transformation of a vector preserves its length.

Theorem: Suppose X and Y are independent RVs. Then, X_θ, Y_θ are independent if and only if X, Y are both Normal RVs with the same variance.

Central Limit Theorem (CLT)

Theorem (Weak Law of Large Numbers, Lecture 23). Let $X_1, X_2, X_3, \ldots$ be a sequence of i.i.d. random variables with finite mean μ and finite variance σ^2. Let $S_n = X_1 + \cdots + X_n$. Then, $\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = 0, \ \forall \varepsilon > 0.$

WLLN is a statement about the convergence of the sample average $\frac{S_n}{n}$ to its mean as $n \to \infty$

Theorem (CLT): Let $X_1, X_2, X_3, \ldots$ be a sequence of i.i.d. random variables with finite mean μ and finite variance σ^2. Let $S_n = X_1 + \cdots + X_n$. Then,

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - \mu\right) \le x\right] = \Phi(x), \qquad \forall x \in \mathbb{R},$$

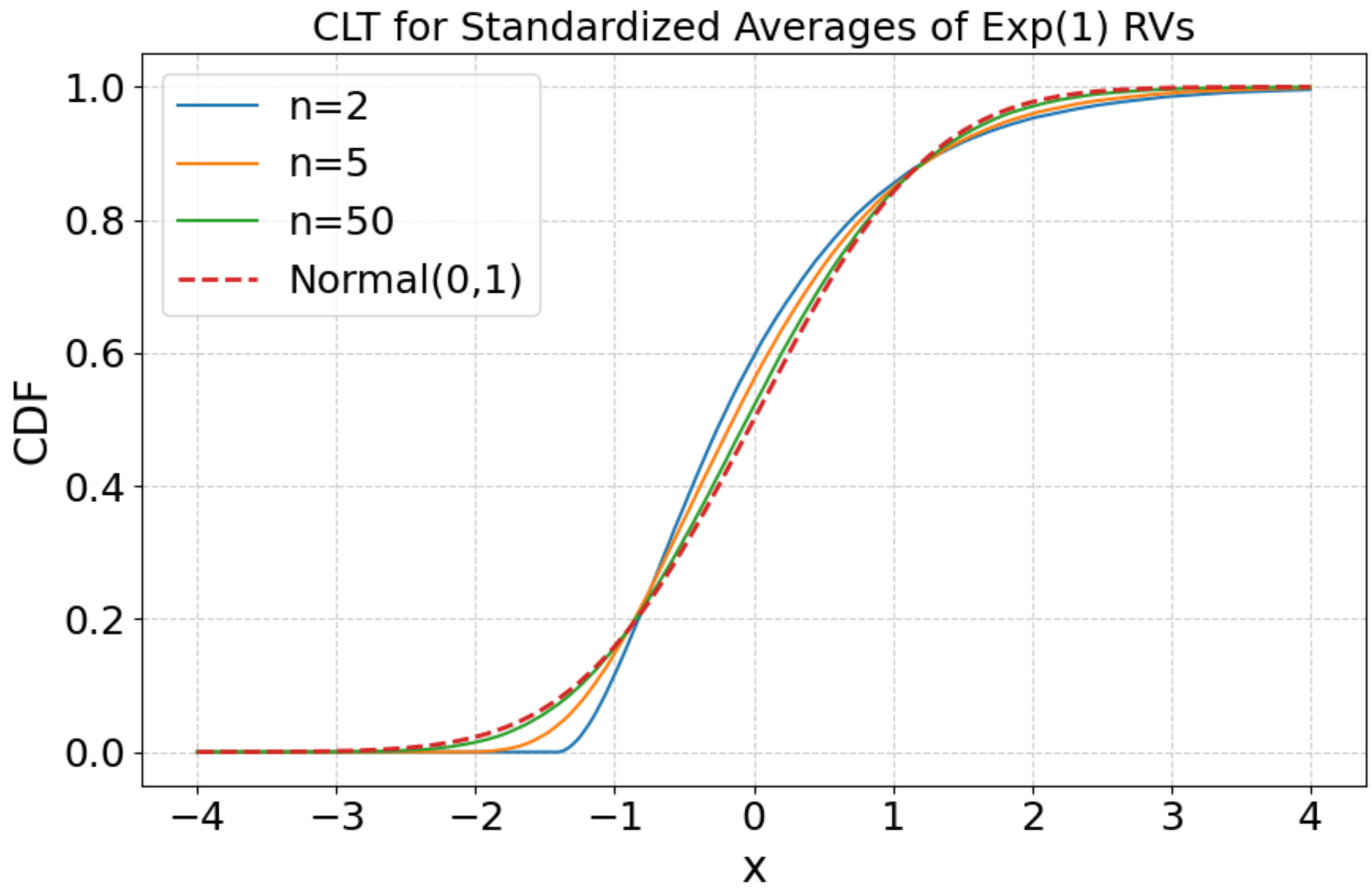
where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$ denotes the c.d.f. of Normal(0,1).

$\mathbb{E}\left[\frac{S_n}{n}\right] = \mu$ and $\mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n}$. Formally, we write $\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - \mu\right) \xrightarrow{d} X \sim \mathrm{Normal}(0,1)$ as $n \to \infty$.

CTL is a statement about the scaled fluctuation of $\frac{S_n}{n}$ about μ as $n \to \infty$.

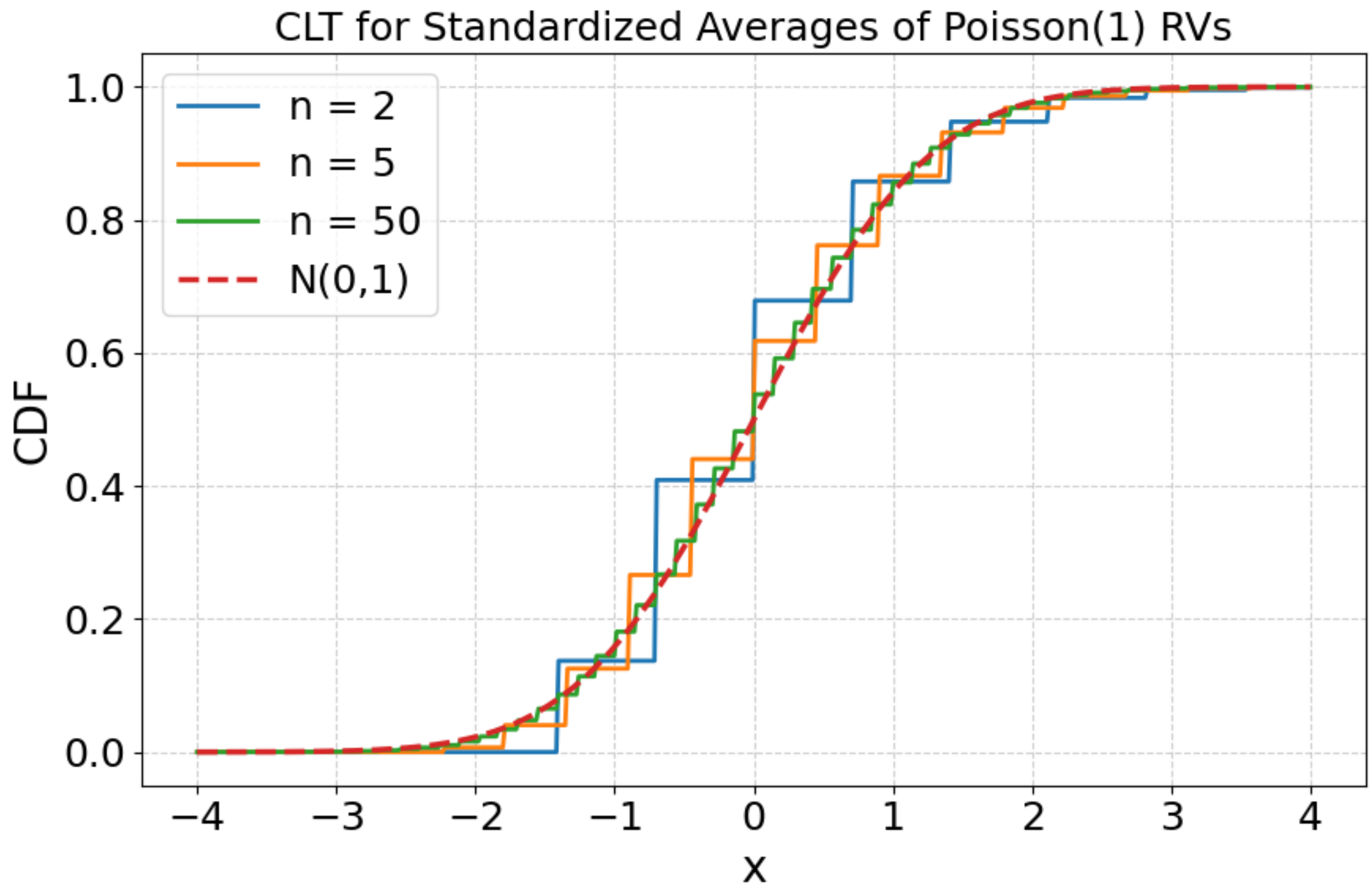
Application: For large n, $\frac{S_n}{n}$ is well approximated by $\mathrm{Normal}(\mu, \frac{\sigma^2}{n})$

CLT Demo



CLT for Standardized Averages of Exp(1) RVs
n=2
n=5
n=50
Normal(0,1)
CDF
x
1.0
0.8
0.6
0.4
0.2
0.0
−4
−3
−2
−1
0
1
2
3
4

CLT Demo



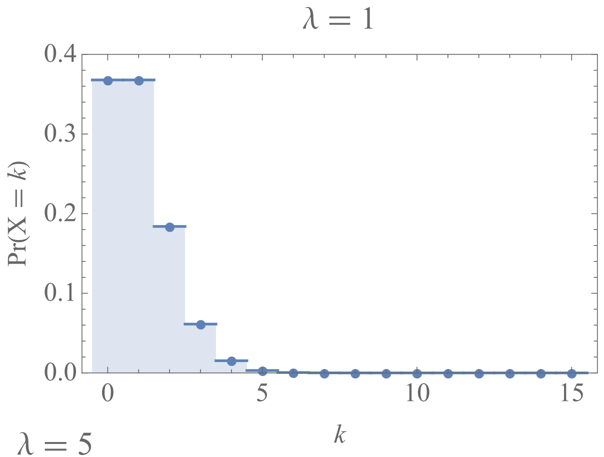
CLT for Standardized Averages of Poisson(1) RVs
n = 2
n = 5
n = 50
N(0,1)
CDF
0.0
0.2
0.4
0.6
0.8
1.0
−4
−3
−2
−1
0
1
2
3
4
x

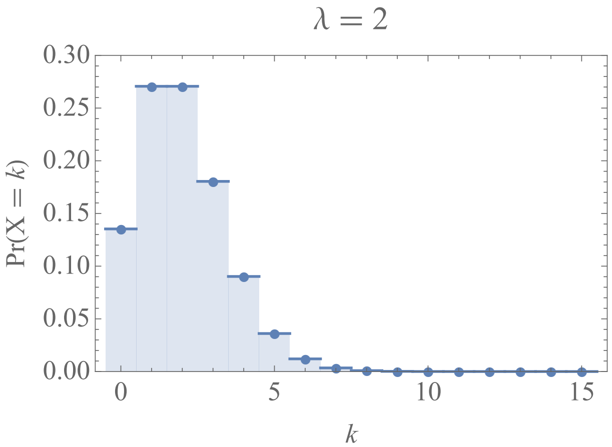
Can you now explain this behavior?

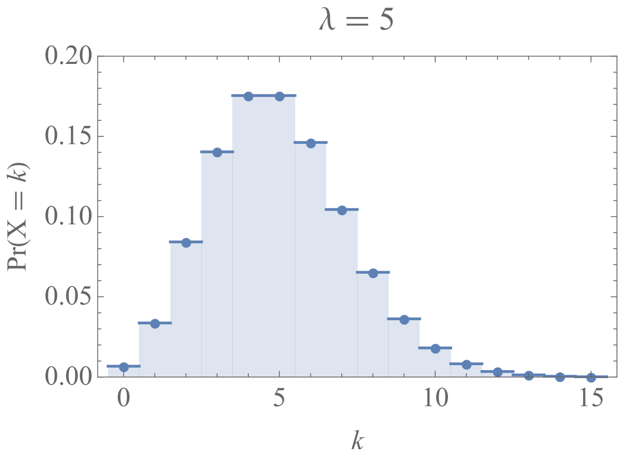
$N \sim \text{Poisson}(\lambda)$, where intensity $\lambda > 0$.

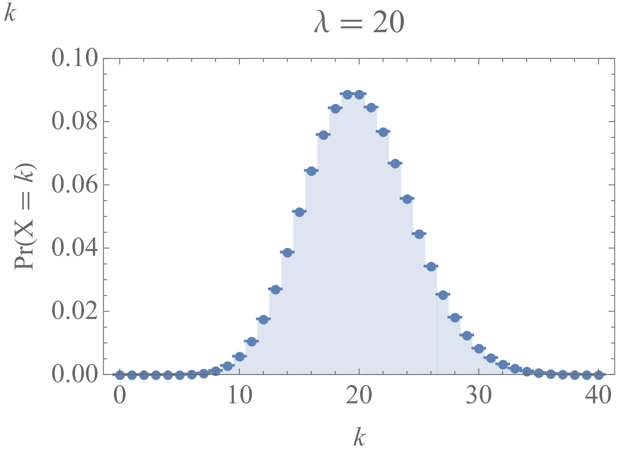
$\mathbb{P}(N = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, for $k \in \mathbb{N}$.

- # rain drops hitting a surface per second
- # radioactive particles emitted by radio active material during an interval of time





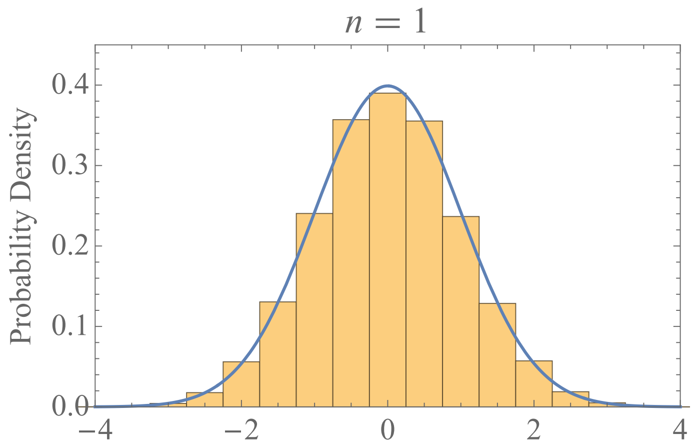




Looks more and more like a "bell curve" as λ gets large. We will see why this happens.

Poisson and CTL

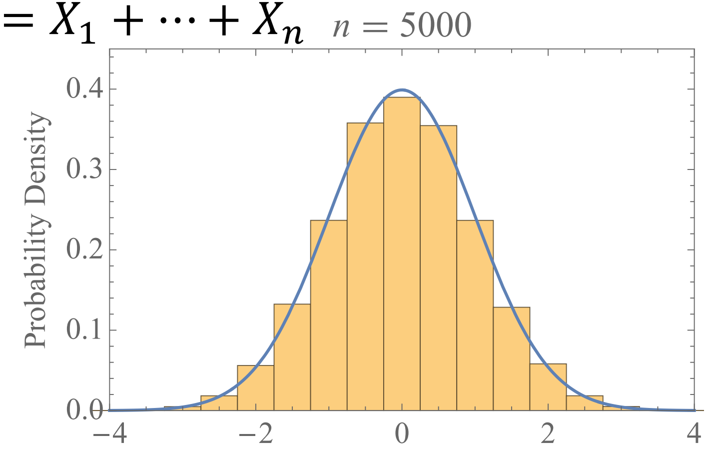
$$Y \sim \text{Poisson}(5000)$$



$$\frac{Y - \lambda}{\sqrt{\lambda}}$$

$$X_1, \ldots, X_{5000} \overset{\text{i.i.d.}}{\sim} \text{Poisson}(1)$$

$$S_n = X_1 + \cdots + X_n$$



$$\sqrt{n}\left(\frac{S_n}{n} - 1\right)$$

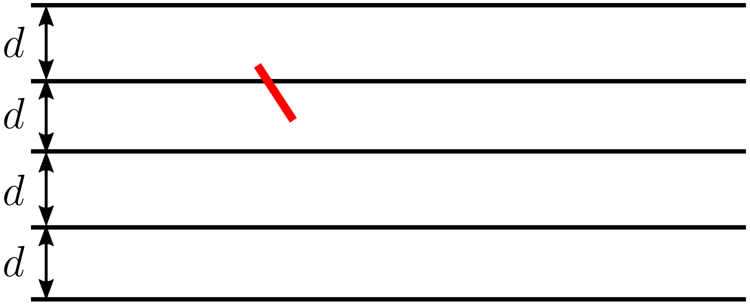
- Recall that if $Z_1 \sim \text{Poisson}(\lambda_1), \ldots, Z_n \sim \text{Poisson}(\lambda_n)$ are independent, then $Z_1 + \cdots + Z_n \sim \text{Poisson}(\lambda_1 + \cdots + \lambda_n)$
- So, in the above example, Y and $X_1 + \cdots + X_n$ are identically distributed.
- However, CLT implies $\sqrt{n}\left(\frac{S_n}{n} - 1\right)$ converges to Normal(0,1) as $n \to \infty$.

How to estimate π using a needle

The 18th century method (due to Buffon)

Assume an infinite array of parallel lines with gap size d. If a needle of length $\ell \leq d$ is thrown uniformly at random, what is the probability that the needle intersects the grid?

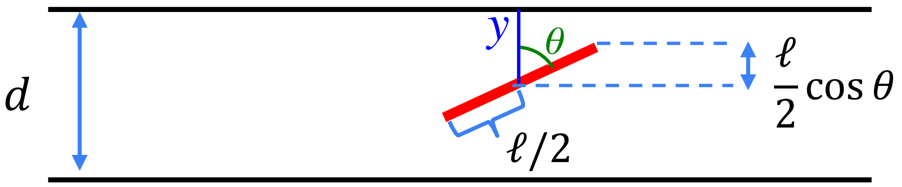
Hit



Miss

How to estimate π using a needle

$Y \sim \text{Uniform}\left[0, \frac{d}{2}\right]$ $\Theta \sim \text{Uniform}\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$



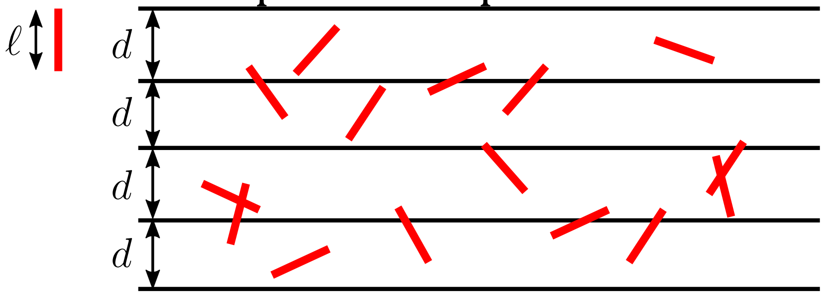
Y and Θ are independent RVs.

- Let Y be the distance between the center of the needle to the closest line.
- For a given θ, the needle will intersect the grid if and only if $y \le \frac{\ell}{2}\cos\theta$.

$$\mathbb{P}(\text{Needle intersects grid}) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{0}^{\frac{\ell}{2}\cos\theta} \frac{1}{\pi \frac{d}{2}} \, \mathrm{d}y \, \mathrm{d}\theta = \frac{2\ell}{\pi d}$$

Joint density $f_{Y,\Theta}(y,\theta)$ within the integration region

Independent repeated trials:



Let I_k be the indicator RV for the kth trial being successful, and let $S_n = I_1 + \cdots + I_n$.

Then, WLLN $\Rightarrow \frac{S_n}{n} \to \frac{2\ell}{\pi d}$ as $n \to \infty$.

So, $\frac{2n\ell}{S_n d} \to \pi$ as $n \to \infty$.

Is this a good estimator of π?

How to estimate π using a needle

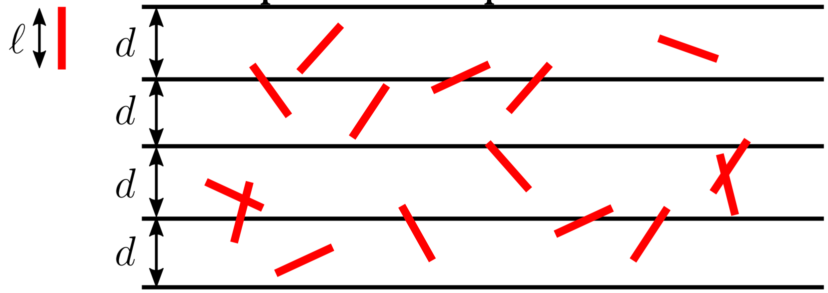
$\hat{\pi}(I_1, \ldots, I_n)$ from computer experiments for $d = \ell$

	Sample size n		
Experiment	100	1,000	10,000
1	2.778	3.257	3.120
2	3.571	3.091	3.136
3	3.226	3.160	3.137
4	4.167	3.200	3.152

Variability decreases as n increases.

CLT can be used to approximate the uncertainty of the estimator for a given sample size n.

Independent repeated trials:



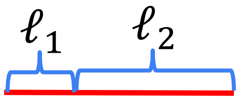
Let I_k be the indicator RV for the kth trial being successful, and let $S_n = I_1 + \cdots + I_n$.

Then, WLLN $\Rightarrow \frac{S_n}{n} \to \frac{2\ell}{\pi d}$ as $n \to \infty$.

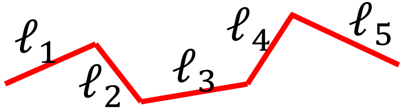
So, $\frac{2n\ell}{S_n d} \to \pi$ as $n \to \infty$.

Is this a good estimator of π?

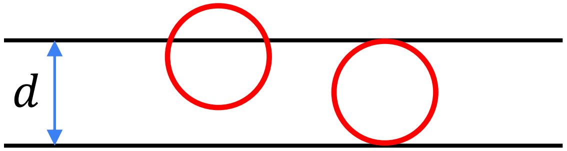
An Alternate Solution to Buffon's Needle Problem



1. Suppose the needle consists of two pieces of lengths ℓ_1 and ℓ_2
2. $p_\ell :=\ \mathbb{P}$(Needle of length ℓ intersects grid) = $p_{\ell_1} + p_{\ell_2}$, which is satisfied if $p_\ell = a\ell$, for some constant $a > 0$. **Need to find $\boldsymbol{a}$.**
3. Consider a polygonal chain consisting of n line segments of lengths $\ell_1, \ldots, \ell_n$, where $\sum_{k=1}^{n} \ell_k = \ell$ and $\ell_k < d$ for all $k = 1, \ldots, n$.

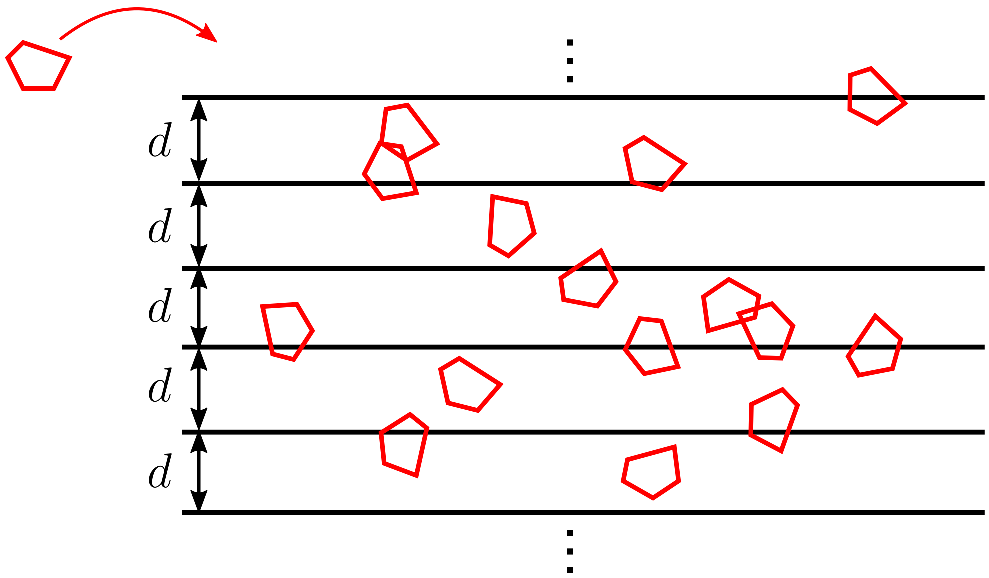


4. Let I_k be an indicator for the segment k intersecting the grid.
5. Let X = the total number of intersections. Then $X = \sum_{k=1}^{n} I_k$ and $\mathbb{E}[X] = \sum_{k=1}^{n} \mathbb{E}[I_k] = \sum_{k=1}^{n} a\,\ell_k = a\ell.$
6. A circle (infinitely many line segments) with diameter d (and hence length $\ell = \ \pi d$) always intersects the grid exactly twice (i.e., $\mathbb{P}(X = 2) = 1$).



So, $\mathbb{E}[X] = 2\mathbb{P}(X = 2) = 2 \Rightarrow a = \frac{2}{\pi d}$

Convex Polygon Throws



Consider a **sufficiently small** (so it cannot cross more than one line at once) convex polygon with circumference C.

X = the total number of intersections in a throw.

- Convexity of the polygon implies $\mathbb{P}(X = 2) + \mathbb{P}(X = 0) = 1$
- So, $\mathbb{E}[X] = 2\mathbb{P}(X = 2)$.
- From the previous slide, we have $\mathbb{E}[X] = \frac{2C}{\pi d}$
- Hence, $\mathbb{P}(\text{Polygon intersects grid}) = \mathbb{P}(X = 2) = \frac{C}{\pi d}$