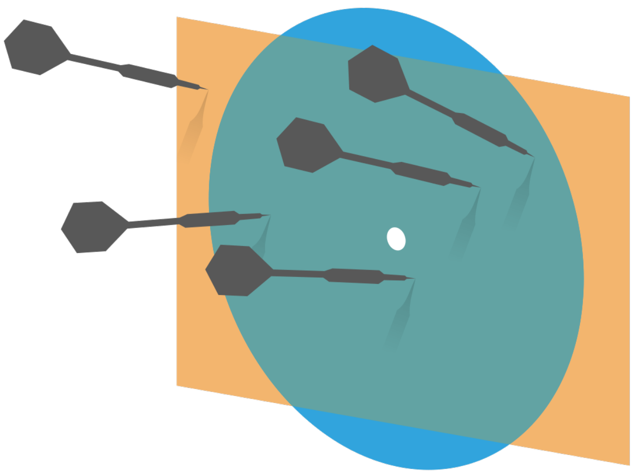
Probability and Statistics for Computer Science



Credit: wikipedia

Can we call e the exciting e ?

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

Hongye Liu, Teaching Assistant Prof, CS361, UIUC, 9.29.2021

What is the number?

$$e^x = \sum_{k=0}^{\infty} a_k x^k$$

$$a_k = ? \quad \frac{1}{k!}$$

Taylor $\quad f(x)$

What is the number?

$$e^x = \sum_{k=0}^{\infty} a_k x^k$$

$$a_k = ? \quad \frac{1}{k!}$$

$$e^0 = 1$$

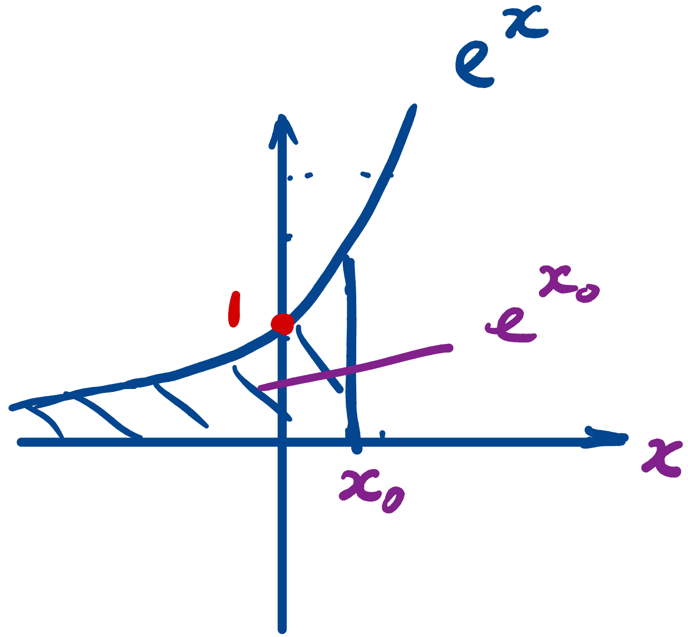
$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

$$(e^x)' = e^x \checkmark$$

What is the number?

$$e^x = \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

$$a_k = ? \quad \frac{1}{k!}$$



How many empty slots?

Hashing N items to k slots ($N \geq k$) collisions are allowed, and will be handled by linked list. What is the expected number of empty slots?

$$X_i = \begin{cases} 1 & \text{slot } i \text{ remains empty after hashing} \\ 0 & \text{otherwise} \end{cases}$$

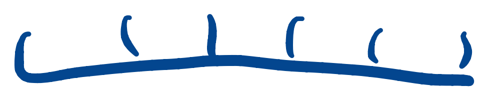
$P(X_i = 1) = ?$

$E[X_i] = ?$ ← $P(X_i = 1)$

$E[X_{ES}] = ?$

$$E\left[\sum_{i=1}^{k} X_i\right] = k \cdot \left(1 - \frac{1}{k}\right)^N$$

$$\left(1 - \frac{1}{k}\right)^N$$



$\rightarrow \frac{1}{e}$

Last time

The classic discrete distributions

Bernoulli:

Binomial

Geometric

H, T

H $k=1$

TH $k=2$

$$P(X=k)=(1-p)^{k-1}p$$

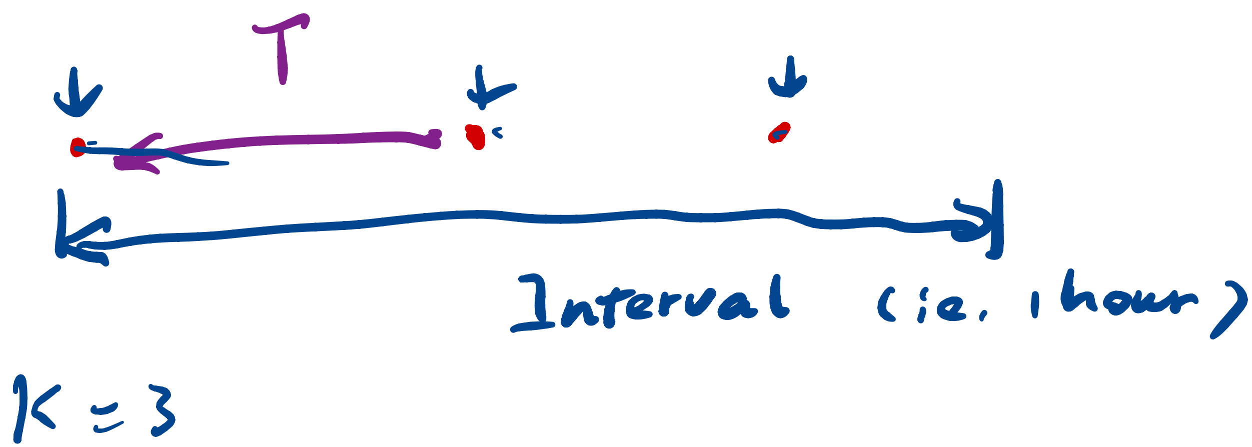
$$E[X]=\frac{1}{p}$$

$$var[X]=\frac{1-p}{p^2}$$

Continous Random Variable

Objectives

- **Poisson** distribution
- Continuous random variable; uniform distribution
- **Exponential** distribution



Motivation for Poisson Distri.

COVID incidences in a time interval,

and many other real world applications.

Motivation for a model called Poisson Distribution

- What's the probability of the **number of incoming customers (k)** in an hour?

- It's widely applicable in physics and engineering both for modeling of time and space.

Simeon D. Poisson Credit: wikipedia
(1781-1840)

DeGroot
Pg 287-288

Poisson Distribution

- A discrete random variable X is called **Poisson** with intensity **λ** (λ>0) if

$$\sum_{k=0}^{\infty} P(X=k) \stackrel{?}{=}$$

$$P(X=k)=\frac{e^{-\lambda}\lambda^{k}}{k!}$$

for integer $k \geq 0$

Simeon D. Poisson
(1781-1840)

λ is the **average rate** *of the event's occurrence*

Poisson Distribution

- **Poisson** distribution is a valid pdf for

$$\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{\lambda}$$

$$\sum_{k=0}^{\infty} P(X=k) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = 1$$

$$e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

Simeon D. Poisson
(1781-1840)

$$P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

for integer $k \geq 0$

λ is the **average rate** *of the event's occurrence*

Poisson Distribution

- **Poisson** distribution is a valid pdf for

$$\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{\lambda} \Rightarrow \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = 1$$

Simeon D. Poisson
(1781-1840)

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

for integer $k \geq 0$

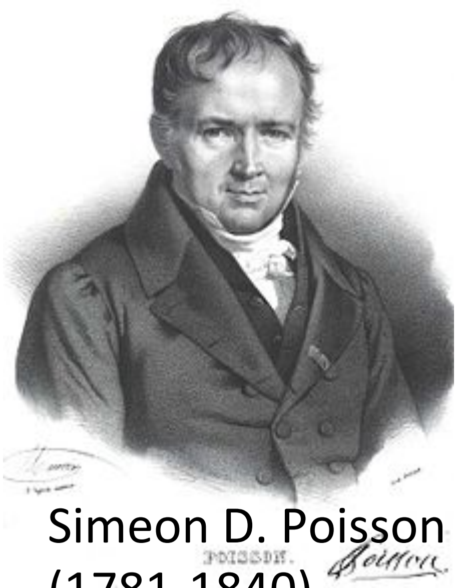
λ *is the* **average rate** *of the event's occurrence*

Expectations of Poisson Distribution

- The expected value and the variance are wonderfully the same! That is **λ**

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

for integer $k \geq 0$



Simeon D. Poisson
(1781-1840)

$$\boxed{\begin{aligned} E[X] &= \lambda \\ var[X] &= \lambda \end{aligned}}$$

$$E[X] = \sum_x x P(x)$$

$$= \sum_k k P(X=k)$$

$$= \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda} \lambda^k}{k!}$$

$$= \sum_{k=1}^{\infty} \cancel{\frac{k}{k}} \frac{e^{-\lambda} \lambda^{k-1} \cdot \lambda}{(k-1)!}$$

$$= \sum_{k=1}^{\infty} \lambda \frac{e^{-\lambda} \lambda^{k-1}}{(k-1)!}$$

$$= \lambda$$

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

$$E[X^2] = \sum x^2 p(x)$$

$$= \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda} \lambda^k}{k!}$$

$$= \sum_{k=2}^{\infty} \frac{k^2}{k(k-1)} \frac{e^{-\lambda} \lambda^{k-2} \cdot \lambda^2}{(k-2)!} + \sum_{k=0}^{1} \frac{k^2 e^{-\lambda} \lambda^k}{k!}$$

$$= \lambda + \lambda^2$$

$$1 + \frac{1}{k-1}$$

$$\mathrm{var}[X] = E[X^2] - E[X]^2 = \lambda$$

Examples of Poisson Distribution

- How many calls does a call center get in an hour?
- How many mutations occur per 100k nucleotides in an DNA strand?
- How many **independent** incidents occur in an interval?

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

for integer $k \geq 0$

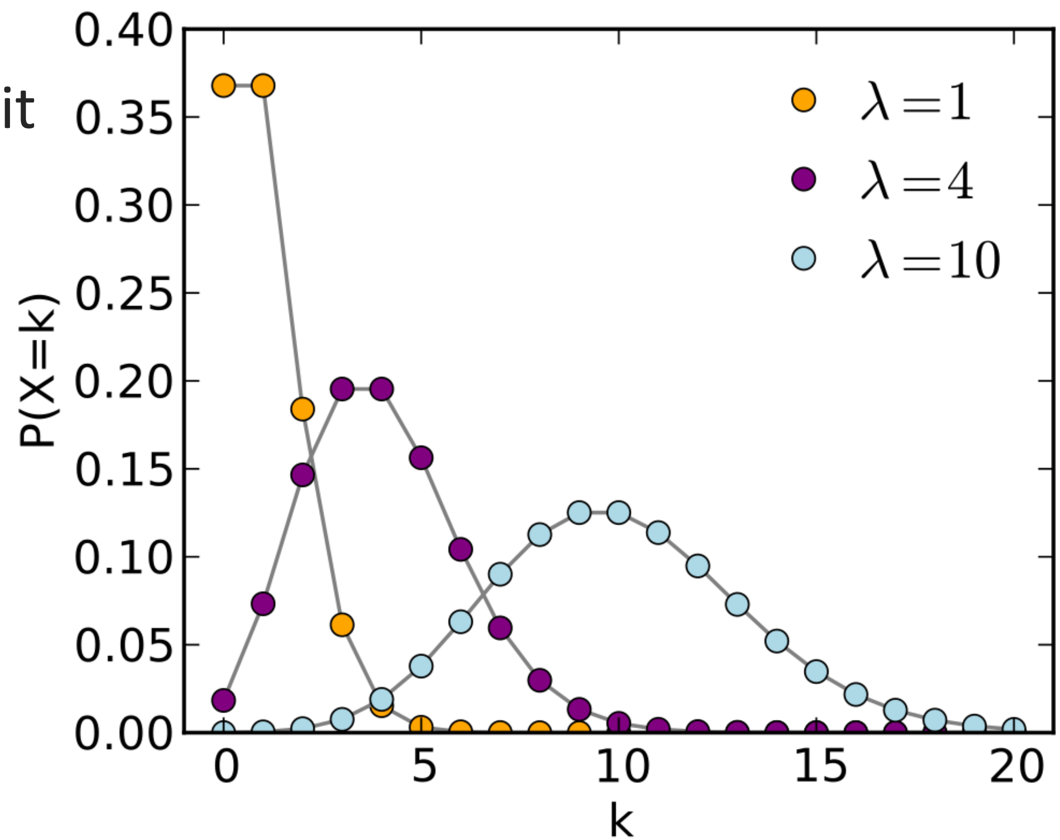
Poisson Distribution: call center

- If a call center receives 10 calls per hour on average, what is the probability that it receives 15 calls in a given hour?
- What is **λ** here? $\lambda = 10$
- What is P(k=15)?

$$P(X=k) = \frac{e^{-10}\, 10^{k}}{k!}$$

$$P(X=15) = \frac{e^{-10} \cdot 10^{15}}{15!}$$

The lines are only to show the trend.



Credit: wikipedia

Q. Poisson Distribution: call center

If a call center receives 4 calls per hour on average.

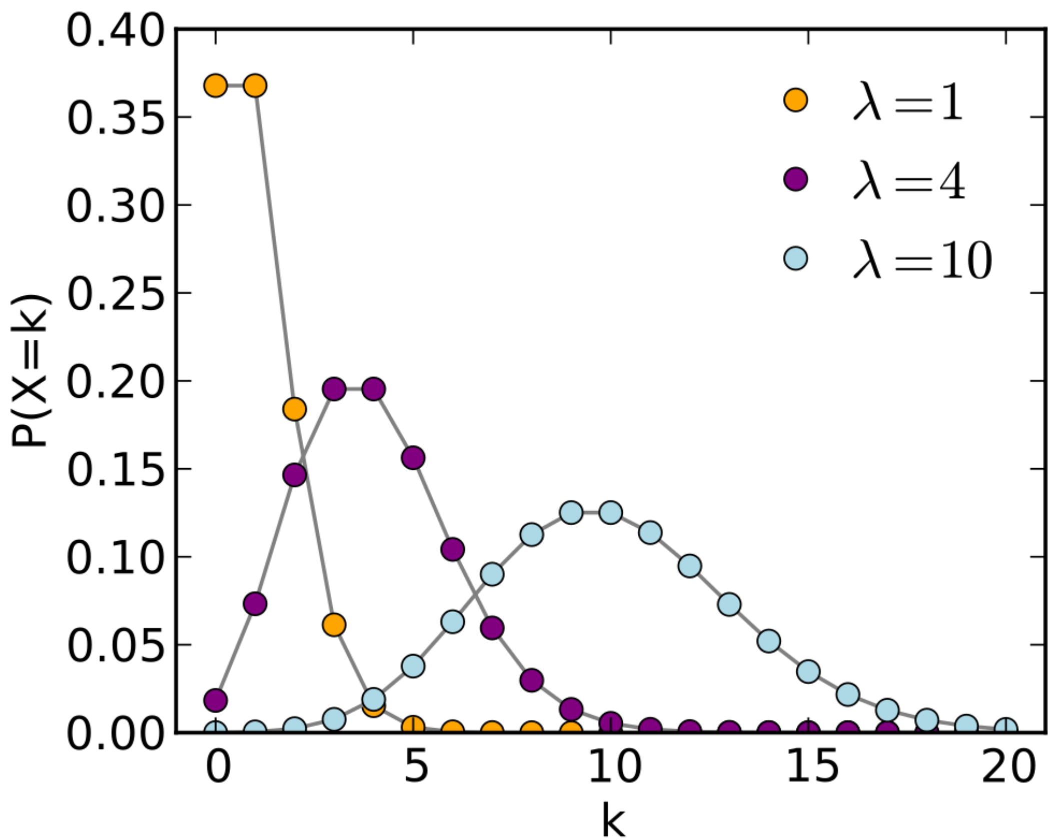
What is intensity **λ** here for an hour?

A. 1

B. 4

C. 8

$\lambda = E[X]$



Credit: wikipedia

Q. Poisson Distribution: call center

If a call center receives 4 calls per hour on average.

$\lambda = 4$

What is probability the center receives 0 calls in an hour?

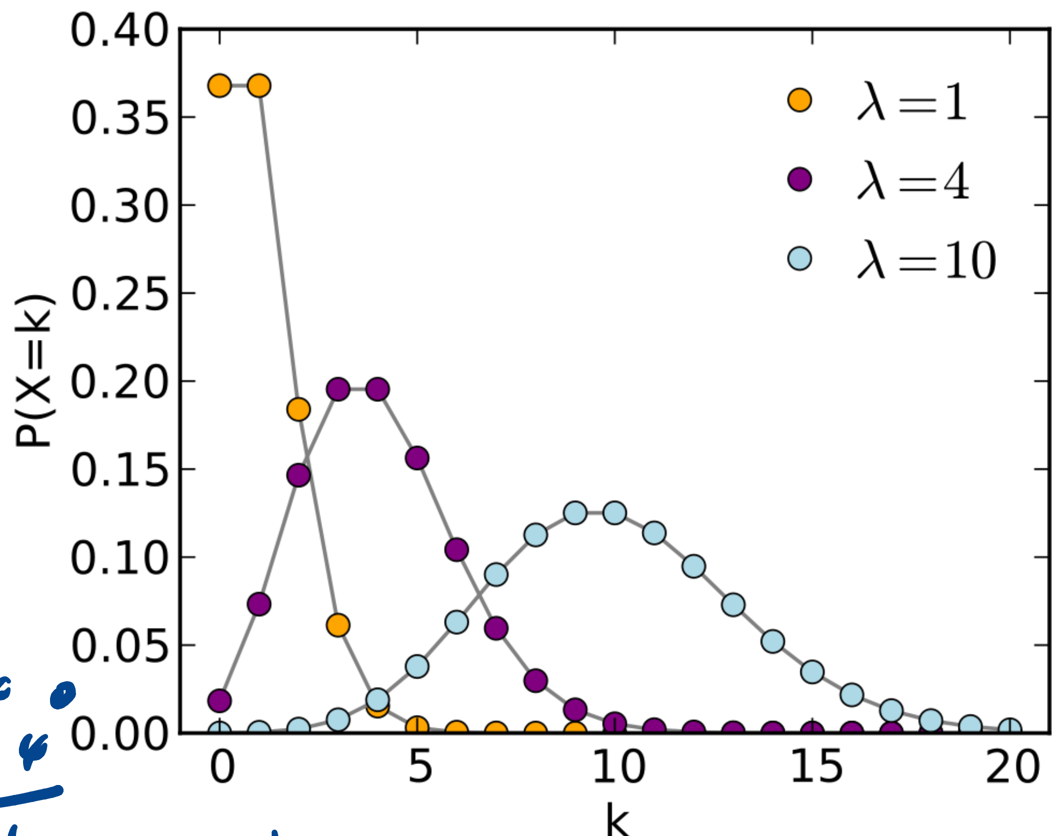
A. $\boxed{}$ e^{-4}

B. 0.5

C. 0.05

$$\frac{e^{-\lambda}\lambda^{k}}{k!} = \frac{e^{-4} \cdot 4^{0}}{0!}$$

$0! = 1$



Credit: wikipedia

Q. Poisson Distribution: call center

- Given a call center receives 10 calls per hour on average, what is the intensity **λ** of the distribution for calls in **Two** hours?

$\lambda_2 = 20$

X_1 2hr

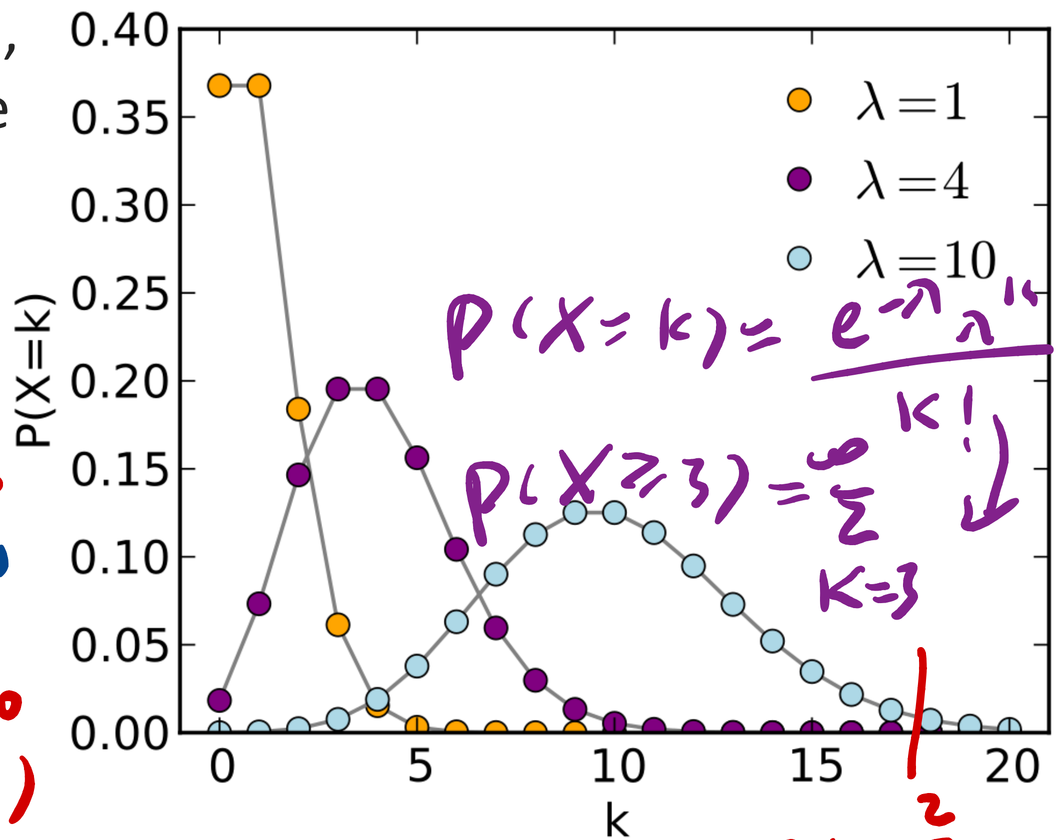
$\lambda_1 = 10$ X_3

X_2 $\lambda_2 = 20$

$X_2 = X_1 + X_3$

$E[X_2] = E[X_1] + E[X_3]$

$= \lambda_1 + \lambda_3 = 2\lambda_1$



$P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$

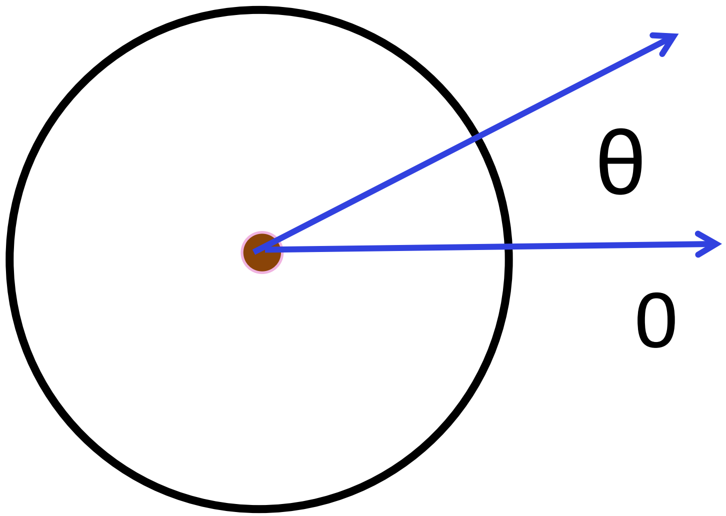
$P(X \geq 3) = \sum_{k=3}^{\infty}$

$= 1 - \sum_{0}^{2} \cdots$

Credit: wikipedia

Example of a continuous random variable

- The spinner



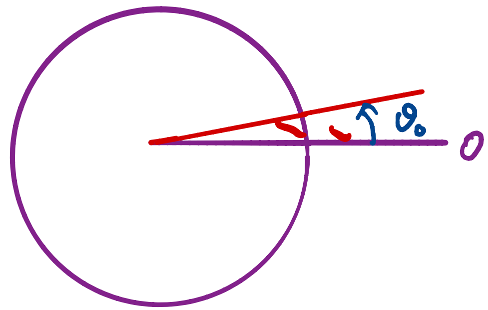
$$\theta \in (0, 2\pi]$$

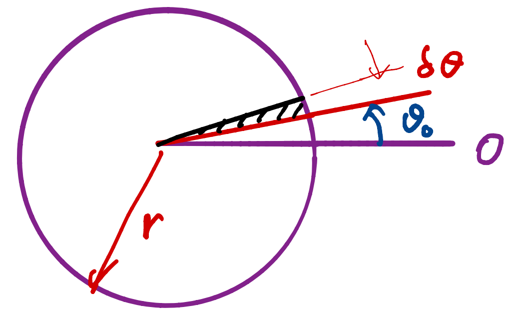
$\cancel{P(\theta = \theta_0)}$

- The sample space for all outcomes is not countable

Spinner example

$\cancel{P(\theta = \theta_0)}$





$$P(\theta_0 < \theta < \theta_0 + \delta\theta) = \frac{\frac{1}{2}\delta\theta \cdot r^2}{\pi r^2}$$

$$\lim_{\delta\theta \to 0} \frac{P(\theta_0 < \theta < \theta_0 + \delta\theta)}{\delta\theta} = \frac{1}{2\pi}$$

Probability density function (pdf)

- For a continuous random variable X, the probability that X=x is essentially zero for all (or most) x, so we can't define $P(X = x)$
- Instead, we define the **probability density function** (pdf) over an infinitesimally small interval dx, $p(x)dx = P(X \in [x, x + dx])$

$$\lim_{dx \to 0} \frac{P(X \in (x, x+dx))}{dx} = p(x)$$

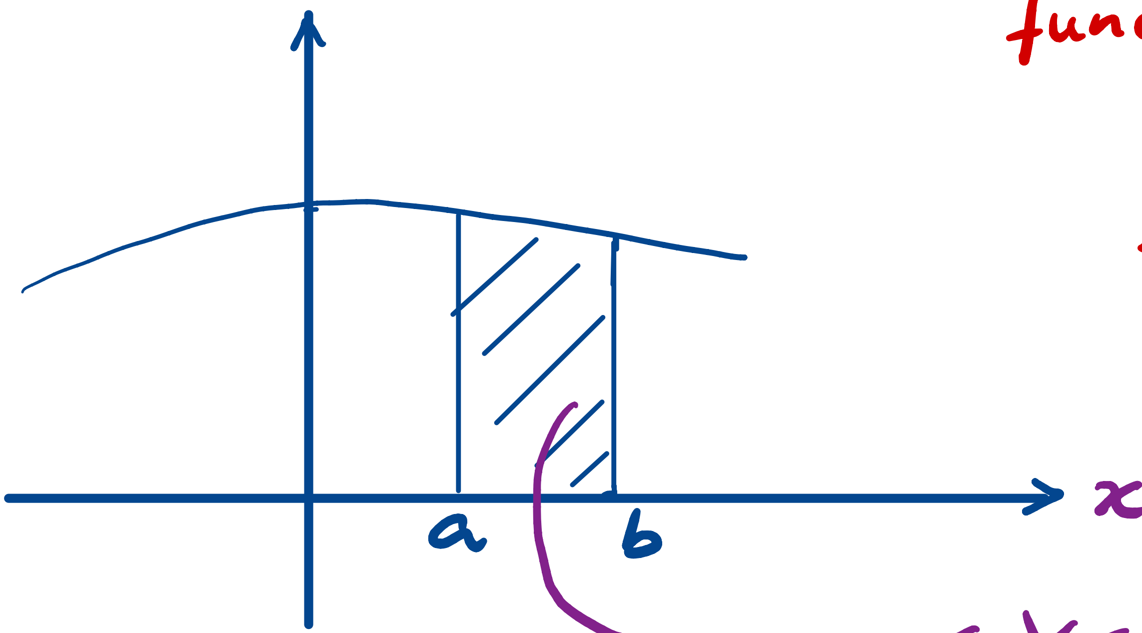
- For $a < b$

$$P(a < X < b) = \int_a^b p(x)dx = P(X \in [a, b])$$

Probability of continuous RV

$p(x) \rightarrow$ Probability density function

$$\int_{-\infty}^{\infty} p(x)\,dx = 1$$



$$P(a \le X \le b) = \int_a^b p(x)\,dx$$

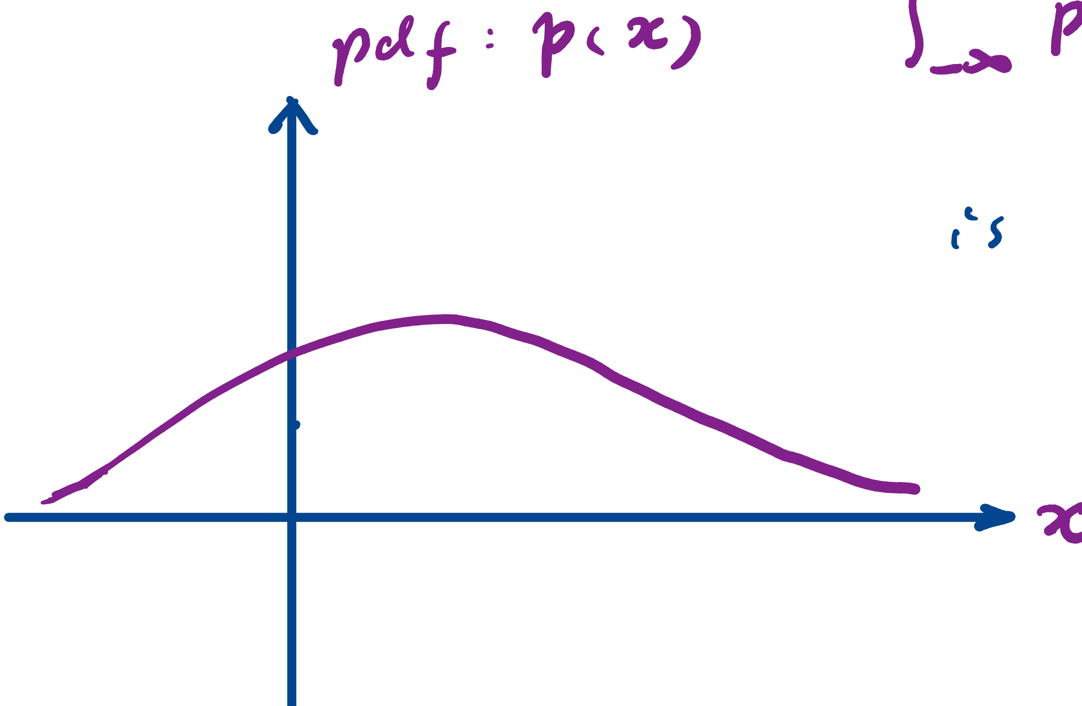
Properties of the probability density function

- $p(x)$ **resembles** the probability function of discrete random variables in that
  - $p(x) \geq 0$ for all x
  - The probability of X taking all possible values is 1.

$P(\Omega) = 1$

$$\int_{-\infty}^{\infty} p(x)dx = 1$$

Area under the pdf curve



$$\int_{-\infty}^{\infty} p(x) = 1$$

is $p(x) = \sin x$ OK?

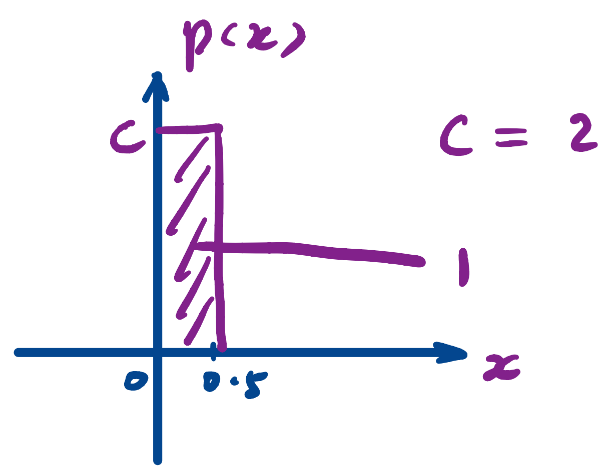
$$p(x) = \begin{cases} c x^3 & x \in (a, b) \\ 0 & \text{ow} \end{cases}$$

Properties of the probability density function

- $p(x)$ **differs** from the probability distribution function for a discrete random variable in that
  - $p(x)$ is not the probability that $X = x$
  - $p(x)$ can exceed 1

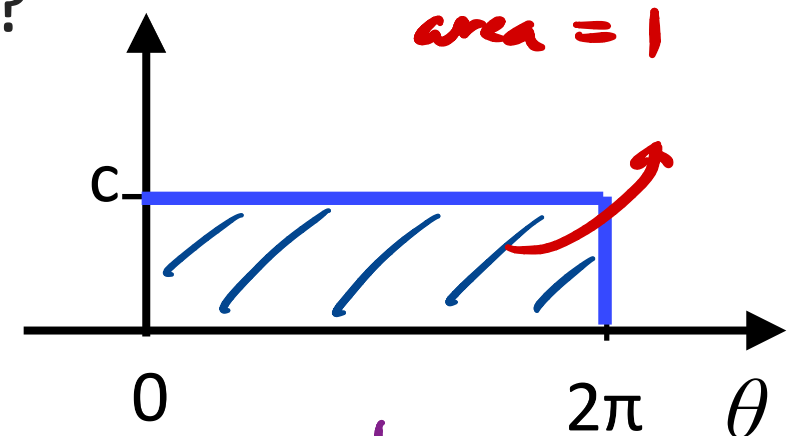
$c \cdot 0.5 = 1$

$\rightarrow c = 2$



Probability density function: spinner

- Suppose the spinner has equal chance stopping at any position. What's the pdf of the angle θ of the spin position?



$$p(\theta) = \begin{cases} c & if \;\; \theta \in (0, 2\pi] \\ 0 & otherwise \end{cases}$$

- For this function to be a pdf,

$c = \frac{1}{2\pi}$

Then $$\int_{-\infty}^{\infty} p(\theta) d\theta = 1$$

$\int_0^{2\pi} c \, d\theta = 1$

$c = \frac{1}{2\pi}$

Probability density function: spinner

- What the probability that the spin angle θ is within $[\frac{\pi}{12}, \frac{\pi}{7}]$?

$$p(\theta) = \begin{cases} \frac{1}{2\pi} & \theta \in [0, 2\pi) \\ 0 & ow \end{cases}$$

$$p(\theta \in [\frac{\pi}{12}, \frac{\pi}{7}])$$

$$= \int_{\frac{\pi}{12}}^{\frac{\pi}{7}} p(\theta)\, d\theta$$

$$= \int_{\frac{\pi}{12}}^{\frac{\pi}{7}} \frac{1}{2\pi}\, d\theta = ?$$

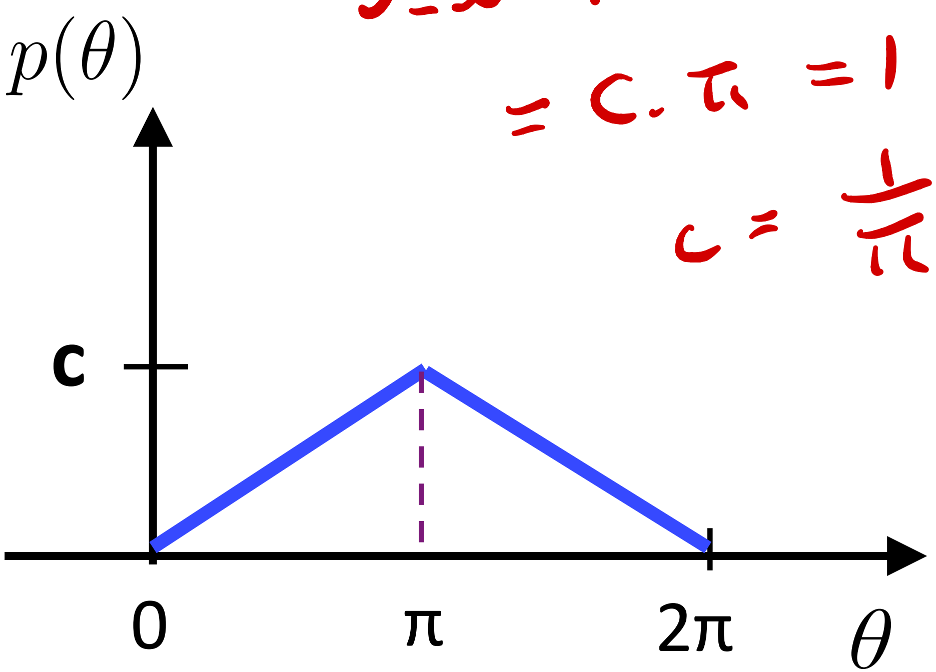
Q: Probability density function: spinner

- What is the constant **c** given the spin angle θ has the following pdf?

$$\int_{-\infty}^{\infty} p(\theta)\, d\theta$$

$$= c \cdot \pi = 1$$

$$c = \frac{1}{\pi}$$



A. 1
B. 1/π
C. 2/π
D. 4/π
E. 1/2π

Expectation of continuous variables

- Expected value of a continuous random variable X

weight

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

- Expected value of function of continuous random variable $Y = f(X)$

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

Probability density function: spinner

- Given the probability density of the spin angle θ

$$p(\theta) = \begin{cases} \frac{1}{2\pi} & if\ \theta \in (0, 2\pi] \\ 0 & otherwise \end{cases}$$

- The expected value of spin angle is

$$E[\theta] = \int_{-\infty}^{\infty} \theta p(\theta) d\theta = \int_0^{2\pi} \theta \cdot \frac{1}{2\pi} d\theta$$

$$= \frac{1}{2\pi} \cdot \frac{\theta^2}{2} \Big|_0^{2\pi} = \pi$$

Properties of expectation of continuous random variables

- The linearity of expected value is true for continuous random variables.

$$\sum \longrightarrow \int$$

$E[X+Y]$
$= E[X]+E[Y]$

$E[kX+b]$
$= kE[X]+b$

- And the other properties that we derived for variance and covariance also hold for continuous random variable

$var[X] = E[X^2] - E^2[X]$

do it on your own

Q.

- Suppose a continuous variable has pdf

$$p(x) = \begin{cases} 2(1-x) & x \in [0,1] \\ 0 & otherwise \end{cases}$$

What is E[X]?

A. 1/2

B. 1/3

C. 1/4

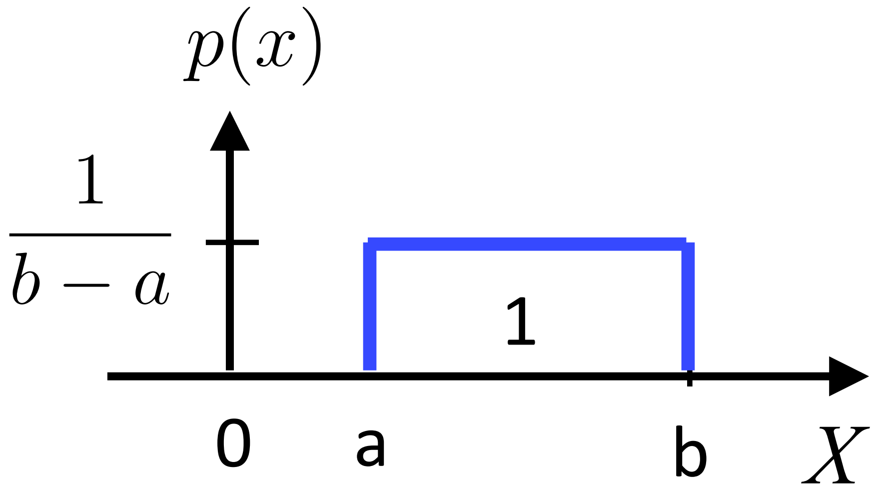
D. 1

E. 2/3

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

Continuous uniform distribution

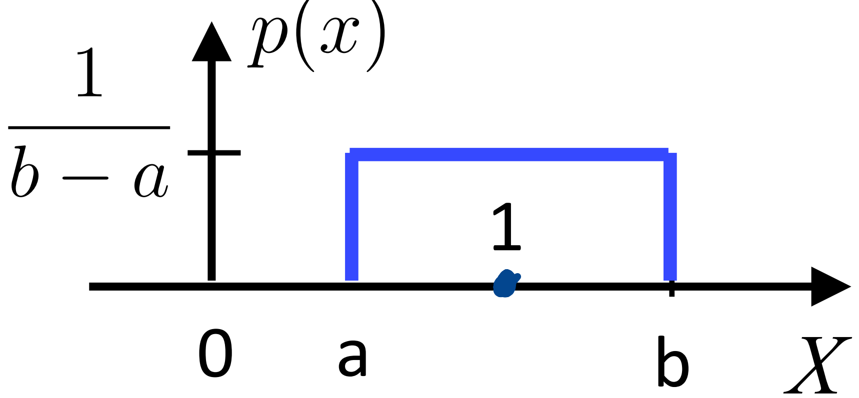
- A continuous random variable X is uniform if



$$p(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{ow} \end{cases}$$

Continuous uniform distribution

- A continuous random variable X is uniform if



$$p(x) = \begin{cases} \frac{1}{b-a} & for \ x \in [a, b] \\ 0 & otherwise \end{cases}$$

$$E[X] = \frac{a+b}{2} \quad \& \quad var[X] = \frac{(b-a)^2}{12}$$

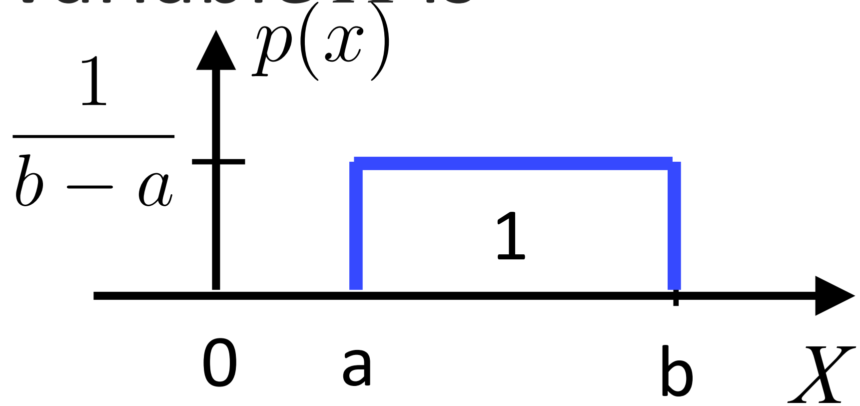
$$var[X] = E[X^2] - (E[X])^2$$

$$= \int_a^b \frac{1}{b-a} \cdot x^2 \, dx - (E[X])^2 = \frac{1}{b-a} \int_a^b x^2 \, dx - (E[X])^2$$

$$= \frac{1}{b-a} \frac{x^3}{3} \Big|_a^b - (E[X])^2$$

Continuous uniform distribution

- A continuous random variable X is uniform if



$$p(x) = \begin{cases} \frac{1}{b-a} & for\ x \in [a,b] \\ 0 & otherwise \end{cases}$$

$$E[X] = \frac{a+b}{2} \quad \& \quad var[X] = \frac{(b-a)^2}{12}$$

- Examples: 1) A dart's position thrown on the target 2) Often associated with **random sampling**

Cumulative distribution of continuous uniform distribution

- Cumulative distribution function (CDF)

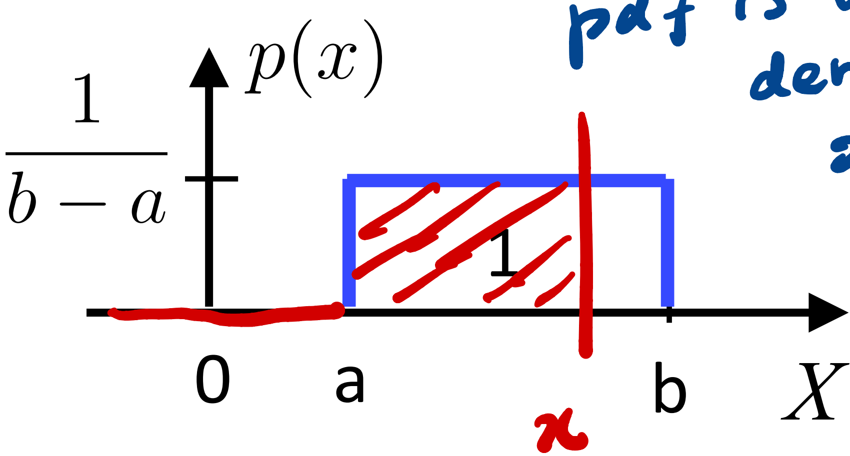
$$P(X \leq x) = \int_{-\infty}^{x} p(x)dx$$

of a uniform random variable X is:

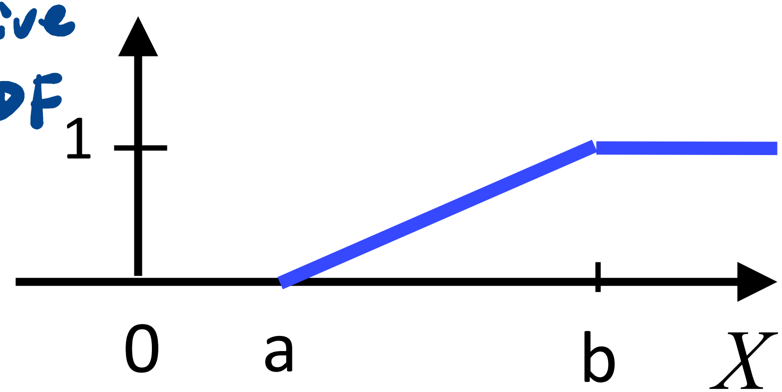
Discrete RV

vs. $P(X \leq x) = \sum_{-\infty}^{x} P(X=x)$

pdf is the derivative of CDF



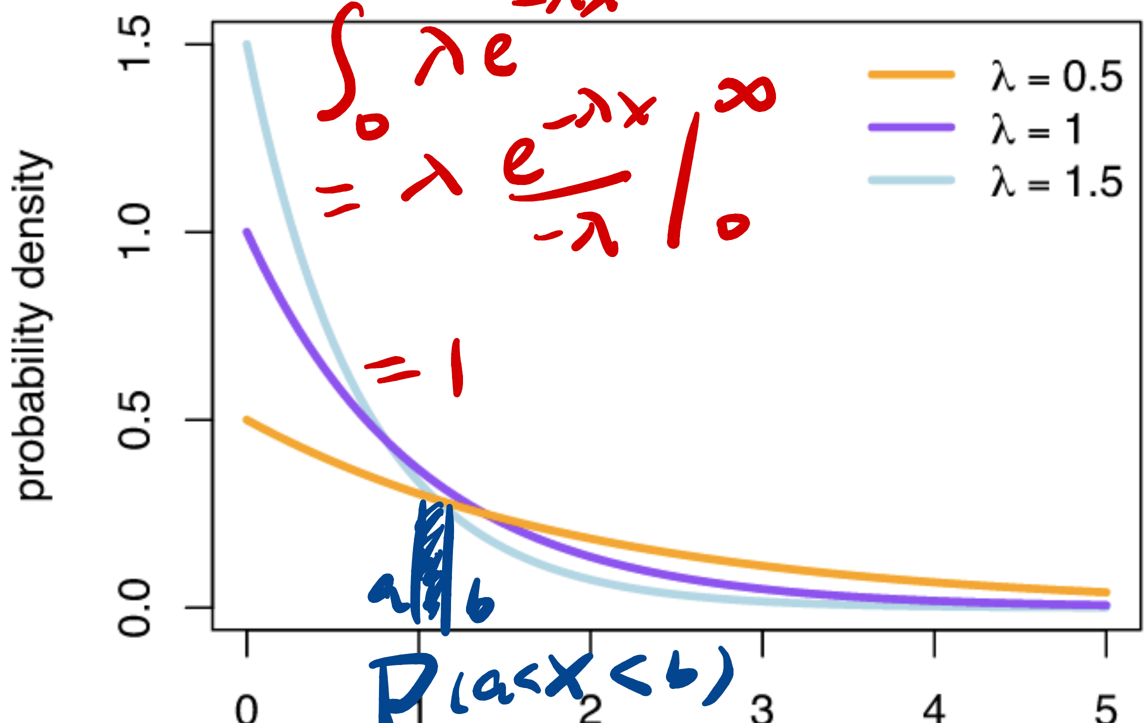
CDF



Exponential distribution

- Common Model for waiting time
- Associated with the Poisson distribution with the same **λ**

$$p(x) = \lambda e^{-\lambda x} \quad for\ x \geq 0$$



$$\int_0^\infty \lambda e^{-\lambda x} = \lambda \frac{e^{-\lambda x}}{-\lambda} \Big|_0^\infty = 1$$

a b

$P(a < X < b)$

Credit: wikipedia

Additional References

- Charles M. Grinstead and J. Laurie Snell "Introduction to Probability”
- Morris H. Degroot and Mark J. Schervish "Probability and Statistics”

Qs for discrete distributions

Q.

- A store staff mixed their fuji  and gala  apples and they were individually wrapped, so they are indistinguishable. Given there are 70% of fuji, if I want to know what is the probability I get 7 fuji in 20 apples? What is the distribution I should use?

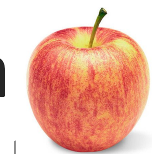
A. Bernoulli B. Binomial C. Geometric

D. Poisson E. Uniform

Q.

- A store staff mixed their fuji  and gala  apples and they were individually wrapped, so they are indistinguishable. Given there are 70% of fuji, if I want to know what is the probability I get 7 fuji in 20 apples? What is the distribution I should use? **What is the probability**?

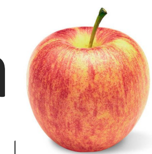
Q.

- A store staff mixed their fuji  and gala  apples and they were individually wrapped, so they are indistinguishable. Given there are 70% of fuji, if I want to know the probability of picking the first gala on the 7th time (I can put back after each pick). What is the distribution I should use?

A. Bernoulli B. Binomial C. Geometric

D. Poisson E. Uniform

Q.

- A store staff mixed their fuji  and gala  apples and they were individually wrapped, so they are indistinguishable. Given there are 70% of fuji, if I want to know the probability of picking the first gala on the 7th time (I can put back after one pick). **What's the probability**?

Q.

- A store staff mixed their fuji  and gala  apples and they were individually wrapped, so they are indistinguishable. Given there are 70% of fuji, **what's the average times of picking to get the first gala?**

See you next time

See You!