Probability in Computing

LECTURE 24

CS 237

Last time

- Uniform Distribution
- Normal Distribution

Today

- **Exponential Distribution**
- **Poisson Process**

Reminders

- HW 11 due tonight

Reading

- P 4.2.2, 11.1.2

Normal Distribution

Tiago decided to fail 60% of the CS237 students. Let $X \sim Normal(\mu_X, \sigma_X^2)$ be distribution of points with $\mu_X = 40$ and $\sigma_X = 5$. How many points you need to pass in this course?

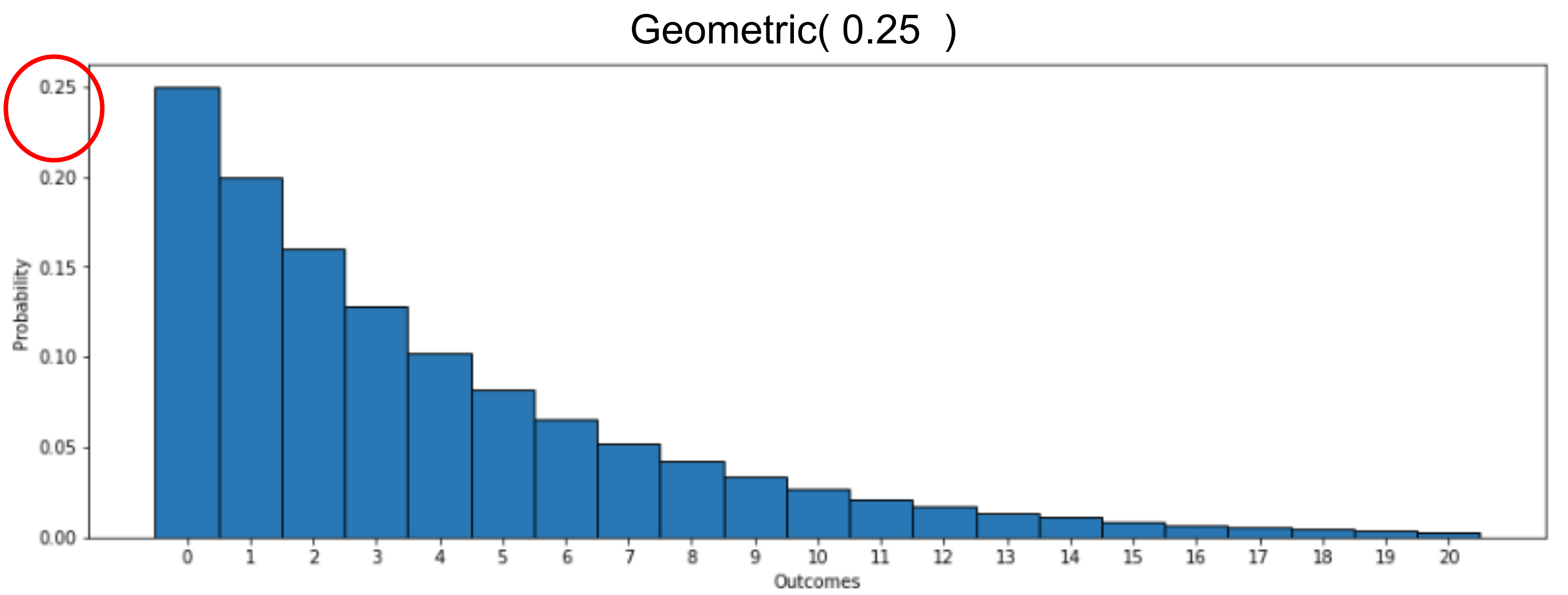
Normal Distribution

Tiago decided to fail 60% of the CS237 students. Let $X \sim Normal(\mu_X, \sigma_X^2)$ be distribution of points with $\mu_X = 40$ and $\sigma_X = 10$. How many points you need to pass in this course?

Comparing Discrete and Continuous Distributions

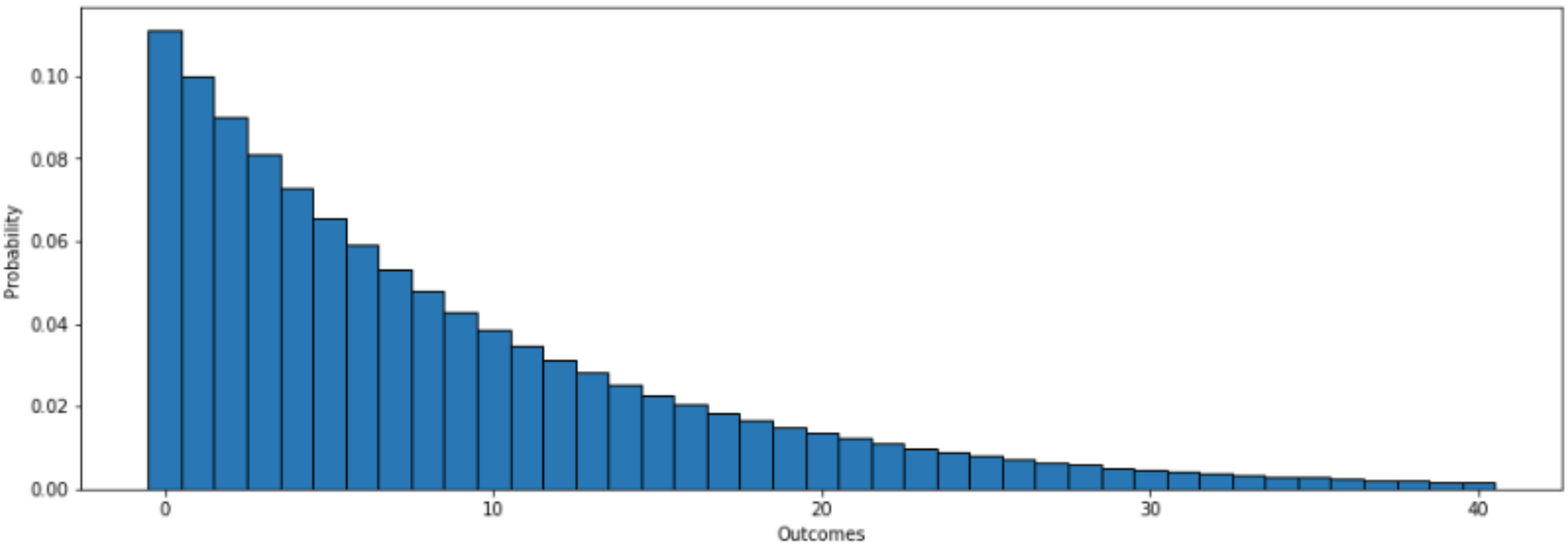
	Discrete		Continuous
Measuring:	Binomial(N,1/2)	⇒	Normal(μ, σ^2)
Waiting:	Geometric(p)	⇒	Exponential(λ)

Exponential Distribution as Limit of Geometric



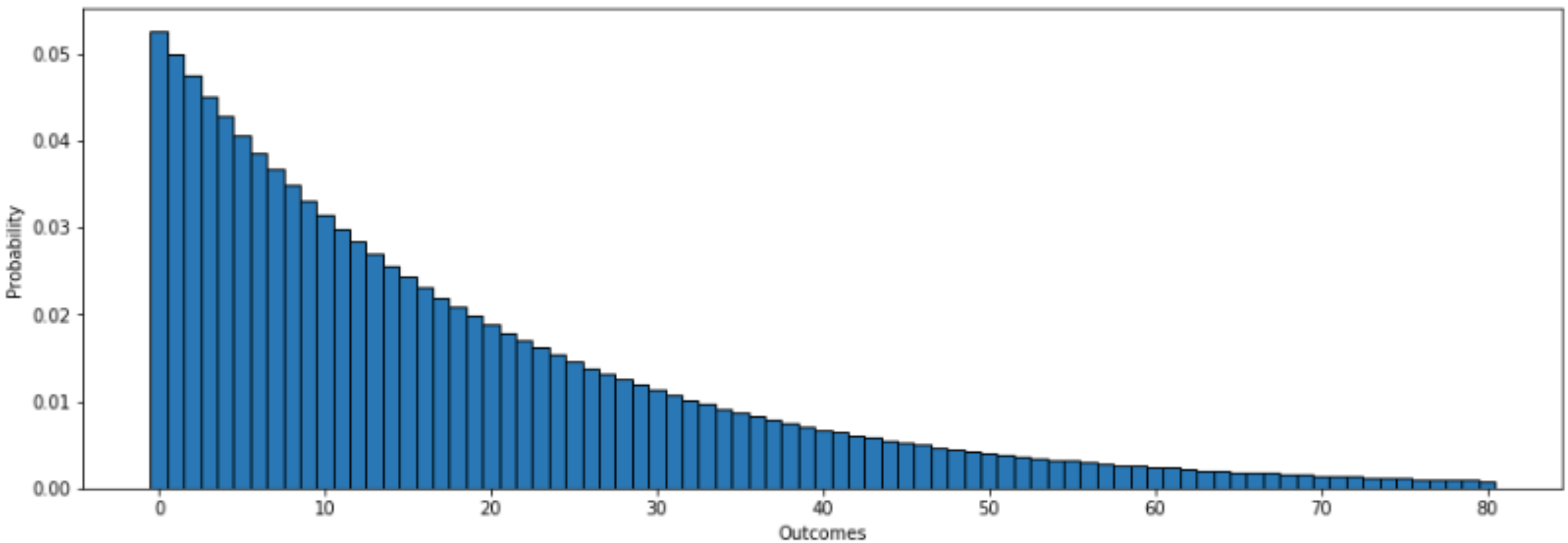
Exponential Distribution as Limit of Geometric

Geometric(0.125)



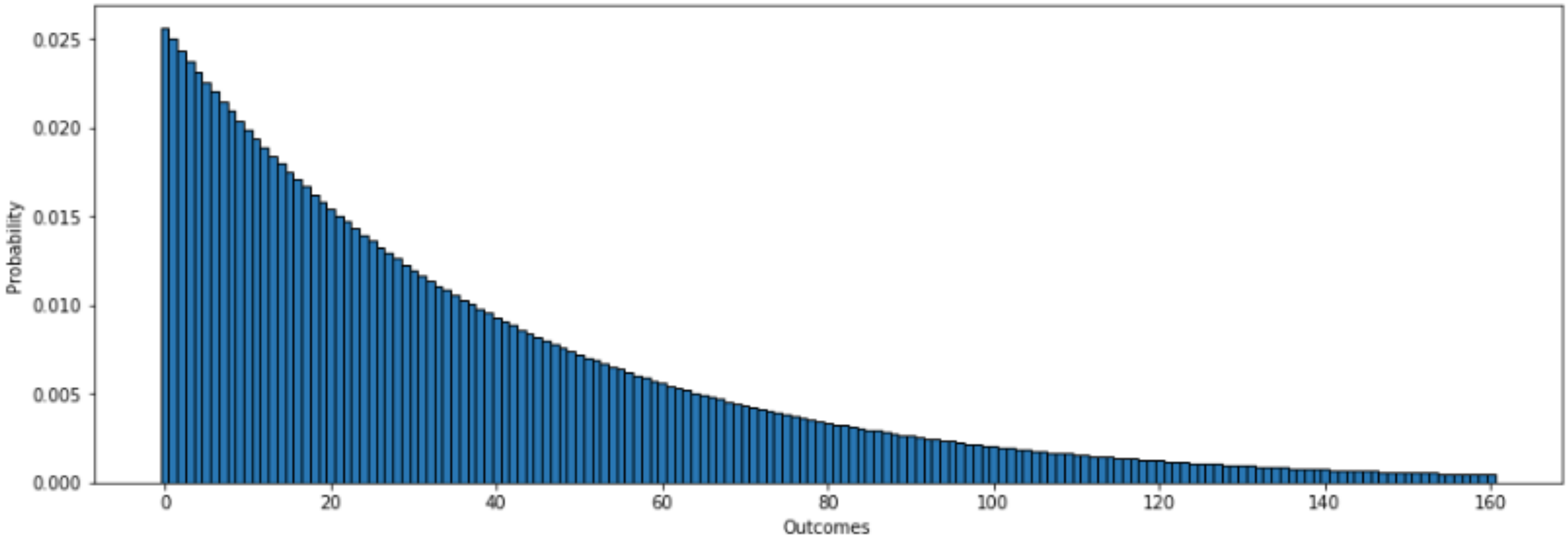
Exponential Distribution as Limit of Geometric

Geometric(0.052)

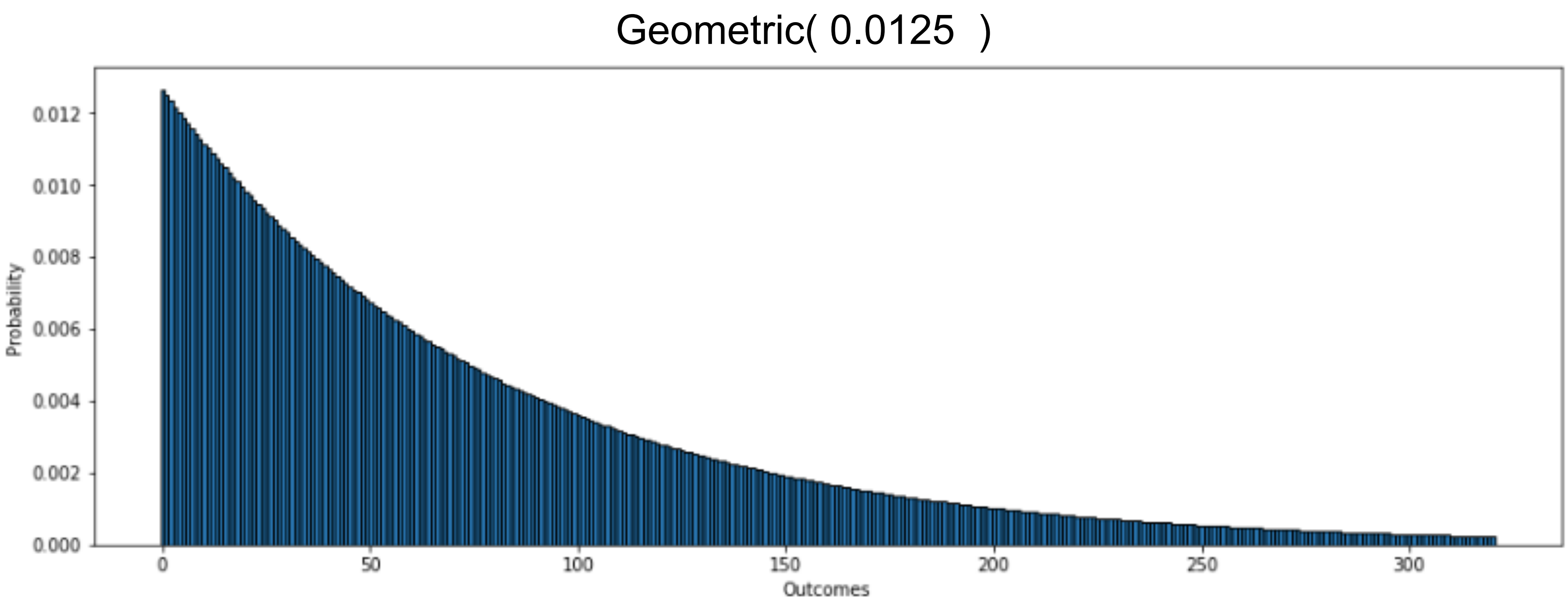


Exponential Distribution as Limit of Geometric

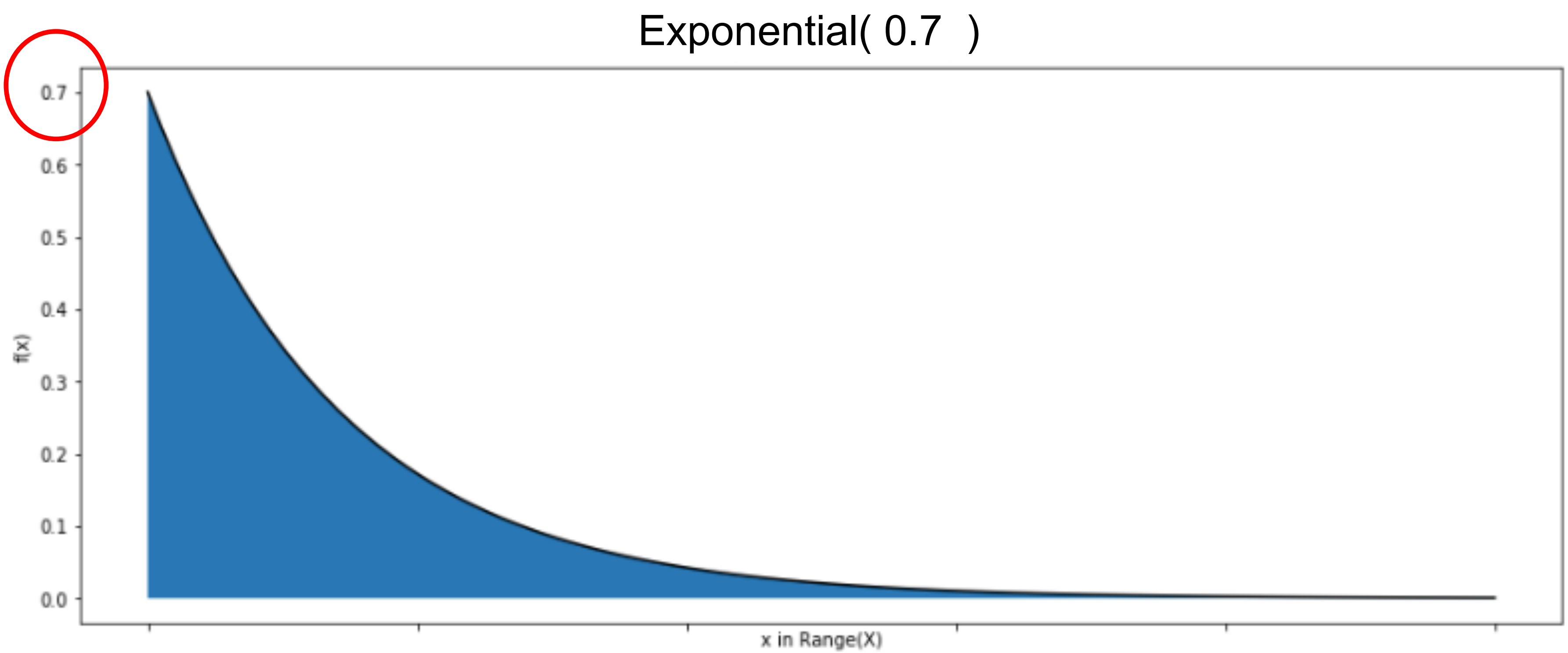
Geometric(0.0252)



Exponential Distribution as Limit of Geometric



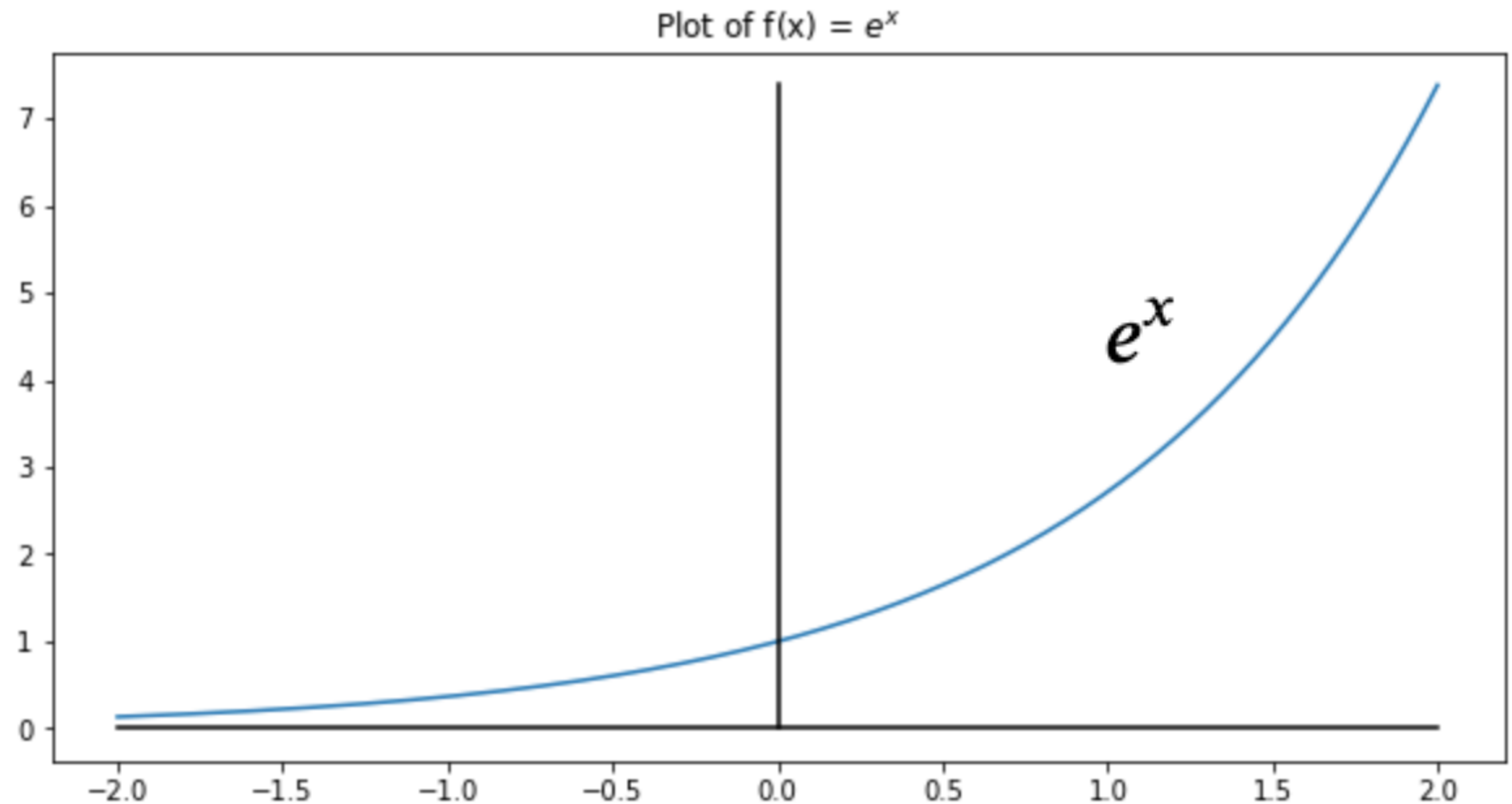
Exponential Distribution as Limit of Geometric



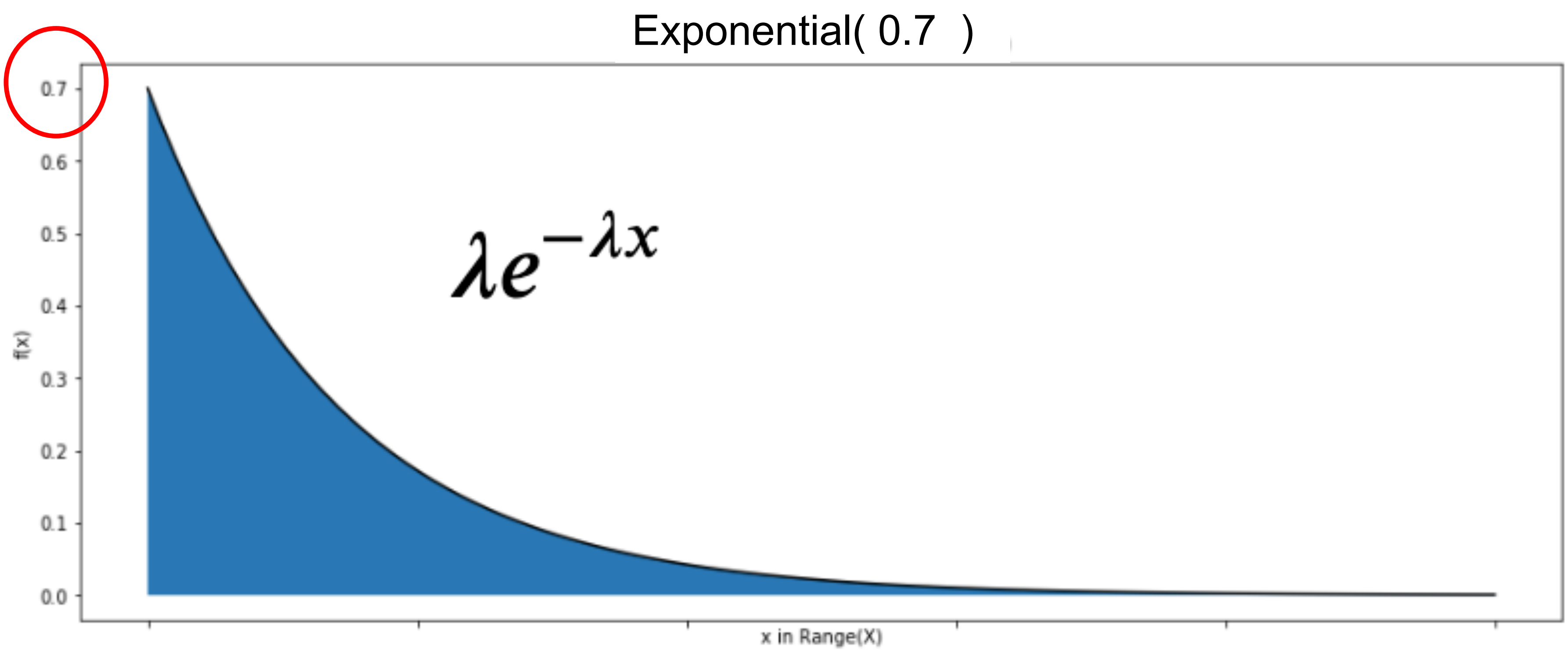
Review: Exponential Function

Since two of the continuous distributions we study (Normal and Exponential) use exponentials, let's think about this a bit....

Here is a graph of the exponential function e^x, where e = 2.71828... (Euler's Constant):



Exponential Distribution as Limit of Geometric



Exponential Distribution

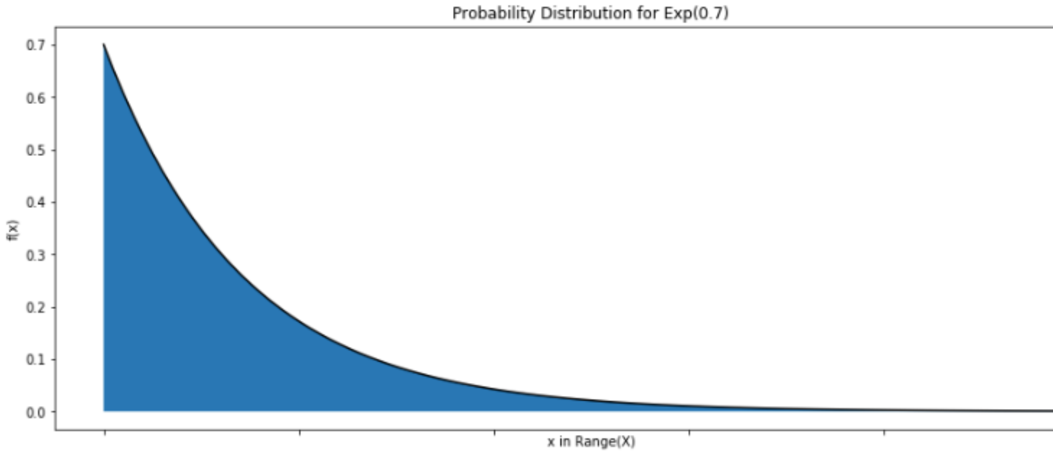
Exponential Distribution, along with the Normal, is one of the most important continuous distributions in probability and statistics.

Formally we say that X is distributed according to the Exponential Distribution with rate parameter $\lambda > 0$, denoted

$$X \sim Exponential(\lambda)$$

if

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t} & if\ t \geq 0 \\ 0 & if\ t < 0 \end{cases}$$



$$F_X(t) = \begin{cases} 1 - e^{-\lambda t} & if\ t \geq 0 \\ 0 & if\ t < 0 \end{cases}$$

and where $\mathbb{E}(X) = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$

Exponential vs Geometric

- Exponential distribution models the time between events, where events occur continuously and independently at a constant rate over time
- Geometric distribution, models the number of trials needed to achieve the first success in a sequence of independent Bernoulli trials

- In other words, it models the waiting time for the next occurrence of an event.
- In other words, it models the number of trials needed to obtain the first "success"

Exponential vs Geometric

- $X \sim Exp(\lambda)$
- $R_X = [0, \infty)$
- $f_X(t) = \lambda e^{-\lambda t}$
- $F_X(t) = 1 - e^{-\lambda t}$
- $\mathbb{E}(X) = \frac{1}{\lambda}$
- $Var(X) = \frac{1}{\lambda^2}$
- $Pr(X > t) = e^{-\lambda t}$
- $Pr(X \leq t) = 1 - e^{-\lambda t}$

- $X \sim Geometric(p)$
- $R_X = [1,2,3, \ldots]$
- $P_X(k) = (1-p)^{k-1} p$
- $F_X(k) = 1 - (1-p)^k$
- $\mathbb{E}(X) = \frac{1}{p}$
- $Var(X) = \frac{1-p}{p^2}$
- $\Pr(X > k) = (1-p)^k$
- $\Pr(X \leq k) = 1 - (1-p)^k$

Exponential Distribution

Examples:

- Time until we observe a shooting star
- Time until a taxi or bus arrives
- Time waiting in line at the post office
- Time until a hard drive breaks down
- Time it takes for the popcorn to finish popping
- Time it takes for your hair to dry

Top Hat question (Join Code: 033357)

Assume that at an intersection the time between accidents is exponential and there are two accidents per day, on average. What is the average time in hours between accidents?

A. 24

B. 12

C. 2

D. ½

E. 30

Top Hat question (Join Code: 033357)

Assume that at an intersection the time between accidents is exponential and there are two accidents per day, on average.

- What is the probability that there will be at least one accident in the next two days? Type your answer with 2 decimal places.

Top Hat question (Join Code: 033357)

Fill in the blank

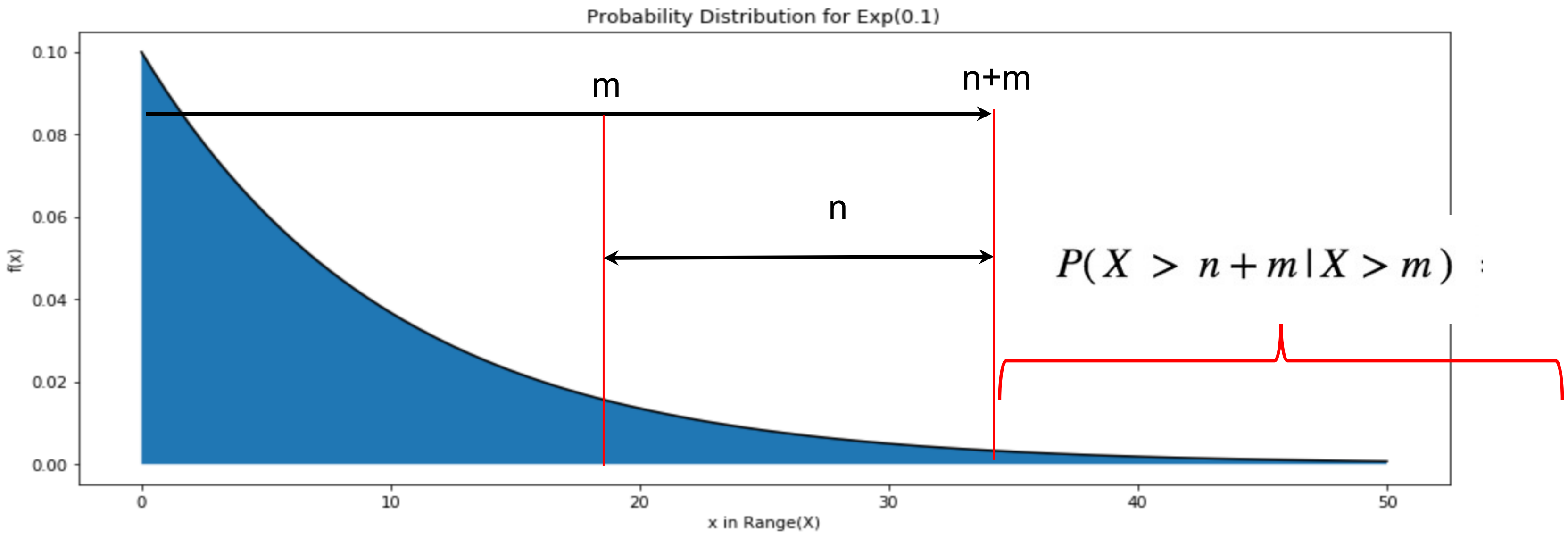
The time (in minutes) Prof Sofya takes per student during OH has an Exponential(0.1) distribution. When you arrive, Prof Sofya is helping another student and there is no one else in front of you. The expected time you will wait is ________ 10 min.

A. greater than

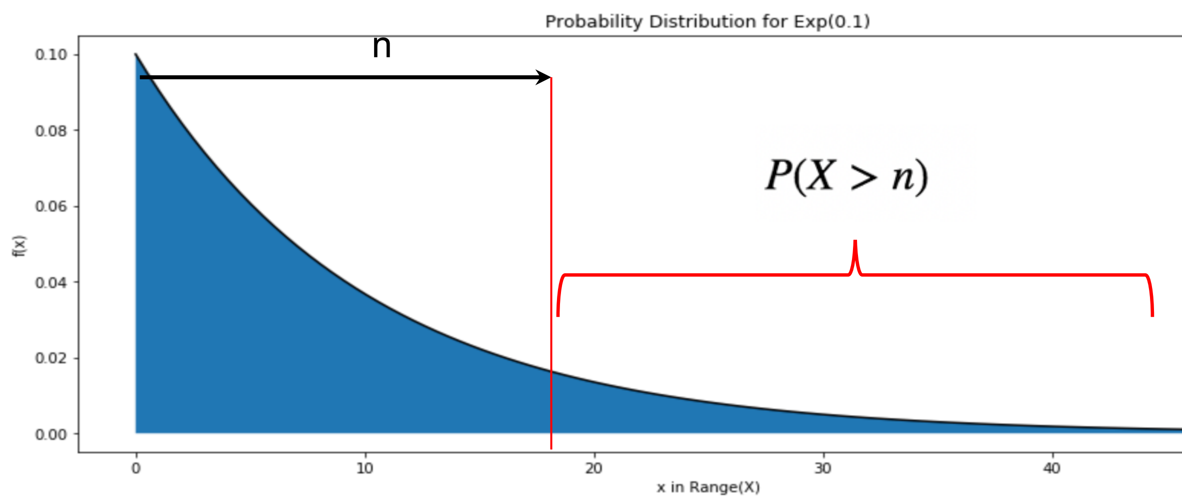
B. equals to

C. smaller than

Exponential Distribution: The Memoryless Property



The exponential, like the geometric, has memoryless property



The Memoryless Property

The time (in hours) Prof Sofya takes per student during OH has an Exponential(6) distribution. You have been waiting for 30 min. What is the probability that you will wait at least 15 min more?

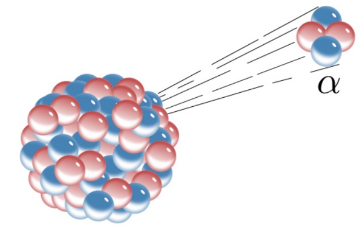
The Memoryless Property

The time (in hours) Prof Sofya takes per student during OH has an Exponential(6) distribution. What is the probability that she will help within 10 to 20 min?

Random Arrivals

Random events over time occur frequently:

- Sneezes in a classroom
- Alpha particles emitted from U 238
- Email arriving in my inbox
- Accidents at an intersection
- Earthquakes, volcanoes, asteroids, ...



What if an asteroid hit the Earth?

BY MARSHALL BRAIN

UP NEXT »

An illustration of an asteroid on its way to Earth. See more space dust images.
PHOTOGRAPHER: ANDREUS AGENCY: DREAMSTIME.COM

An asteroid striking our planet -- it's the stuff of science fiction. Many movies and books have portrayed this possibility ("Deep Impact," "Armageddon," "Lucifer's Hammer," and so on).

An asteroid impact is also the stuff of science fact. There are obvious craters on Earth (and the moon) that show us a long history of large objects hitting the planet. The most famous asteroid ever is the one that hit Earth 65 million years ago. It's thought that this asteroid threw so much moisture and dust in to the atmosphere that it cut off sunlight, lowering temperatures worldwide and causing the extinction of the dinosaurs.

Yellowstone volcano eruption: NASA to SAVE the world from supervolcano erupting

NASA scientists are creating an ambitious plan to prevent an explosion of a Yellowstone volcano that could even end human life by drilling a hole.

THE REALLY BIG ONE

An earthquake will destroy a sizable portion of the coastal Northwest. The question is when.

By Kathryn Schulz

When the 2011 earthquake and tsunami struck Tohoku, Japan, Chris Goldfinger was two hundred miles away, in the city of Kashiwa, at an international meeting on seismology. As the shaking started, everyone in the room began to laugh. Earthquakes are common in Japan—that one was the third of the week—and the participants were, after all, at a seismology conference. Then everyone in the room checked the time.

Seismologists know that how long an earthquake lasts is a decent proxy for its

The next full-margin rupture of the Cascadia subduction zone will spell the worst natural disaster in the history of the continent.

Every year The Federal Highway Administration reports approximately 2.5 Million intersection accidents. Most of these crashes involve left turns.

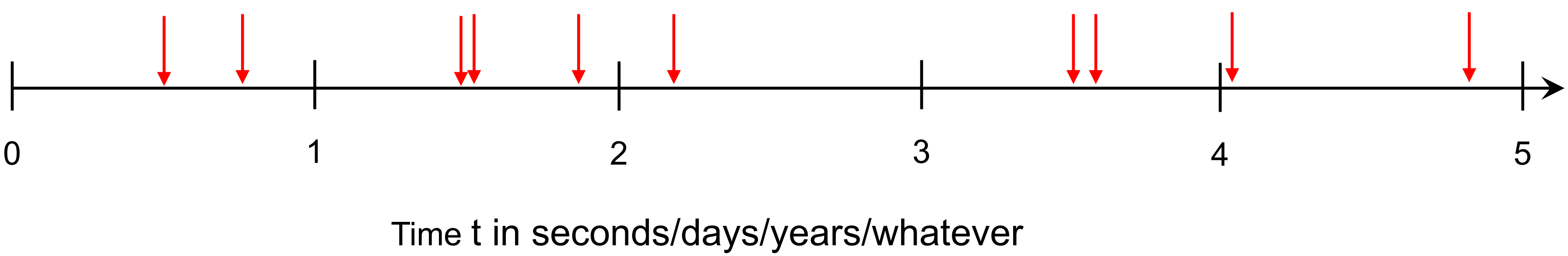
Poisson Process

The Poisson Process concept captures an important way of thinking about events randomly occurring through time (or space)... Two things to remember are

- **Events are discrete** (they happen or they don't – you can think of it as a Bernoulli trial with an outcome of success or failure), but
- **Time and space are continuous**..... the random behavior here is the time of an event.

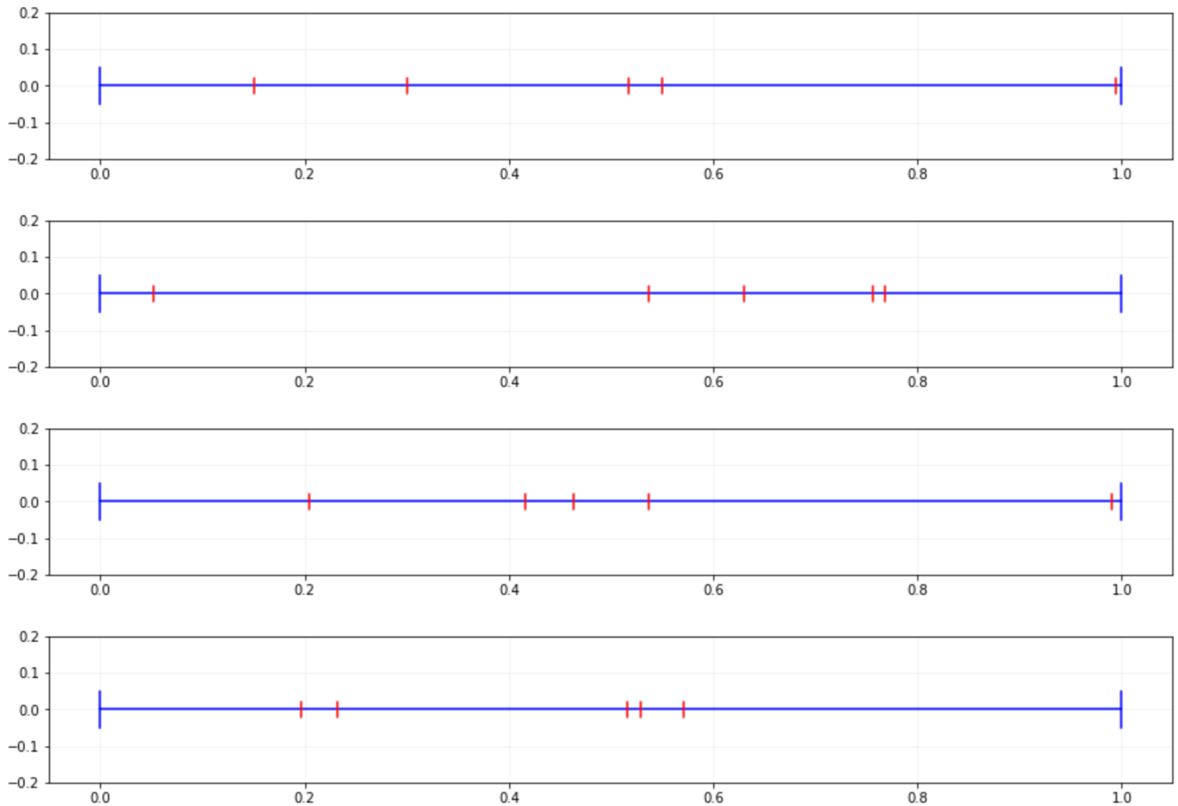
When an event has happened we say it has arrived. You can think of this as a sequence of real numbers giving the **arrival time** of an event:

Arrival Times = { 0.4324... , 0.734 , 1.389... , 1.453... , 2.1546... , ... }



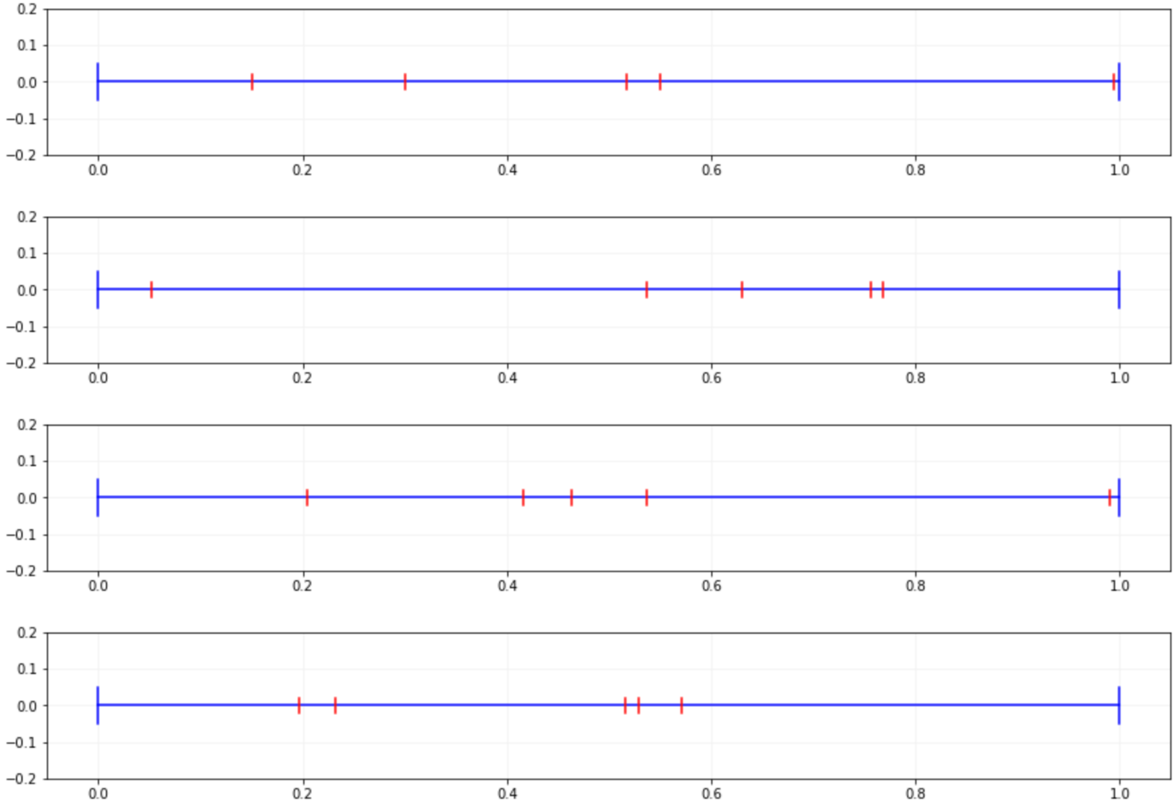
Poisson Process

We can motivate the way a Poisson process is formally defined by considering what happens when we randomly generate arrivals in a unit interval. Suppose each trial of the experiment we generate 5 random numbers in the interval [0..1):



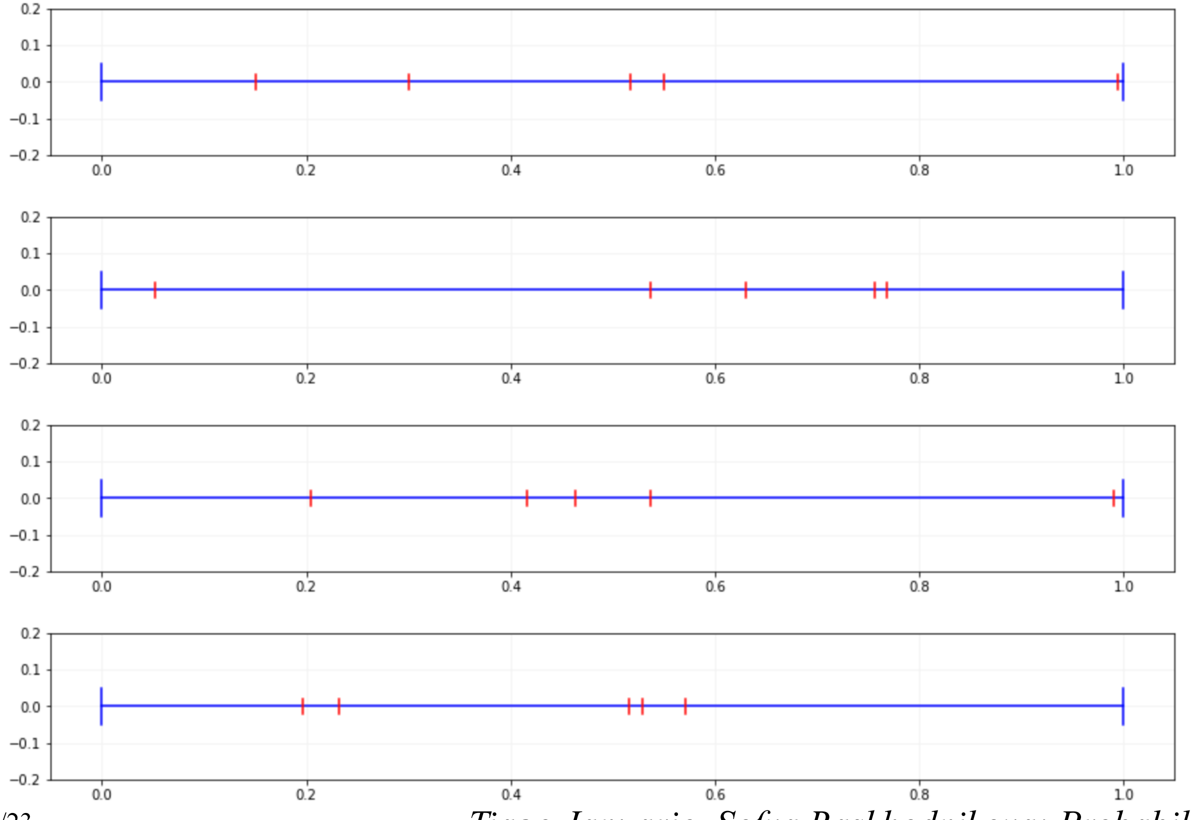
Poisson Process

We know that the probability that a particular arrival occurs in the interval [0.0 .. 0.1) is 1/10; for [0.2 .. 0.5) is 0.3, and for any interval [a..b) it is (b-a).



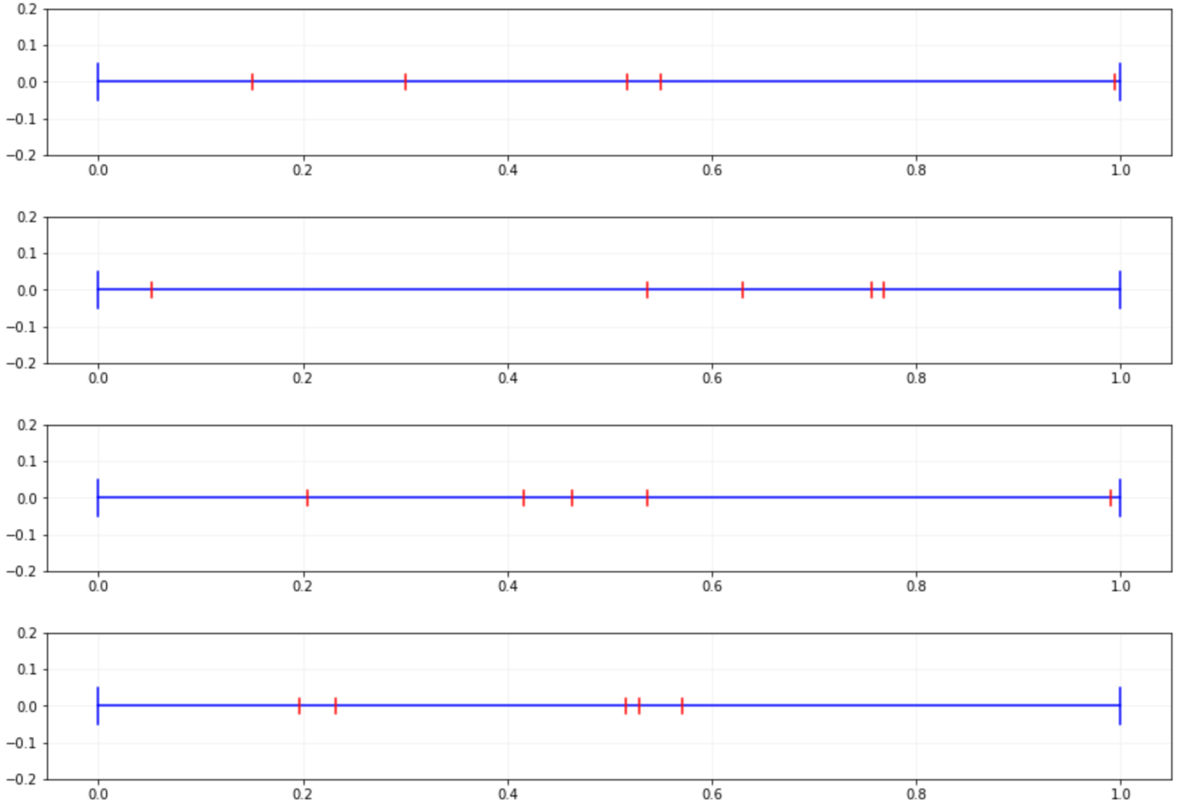
Poisson Process

The probability for any one arrival is equal to the length of the interval.



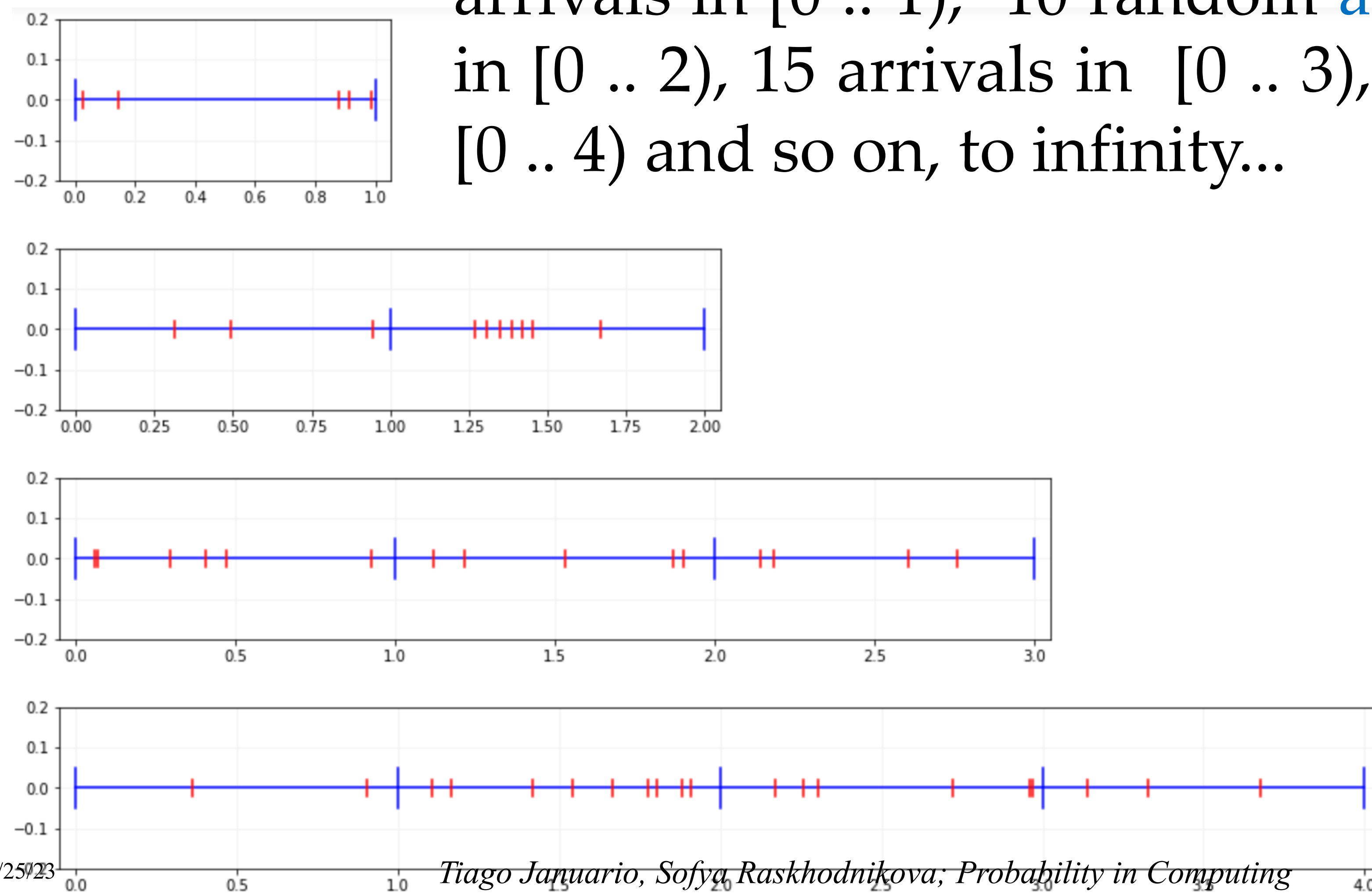
Poisson Process

This is because the arrivals are independent and uniformly distributed in the interval [0..1).



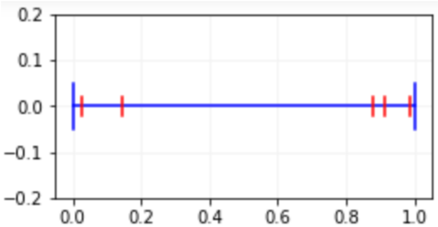
Poisson Process

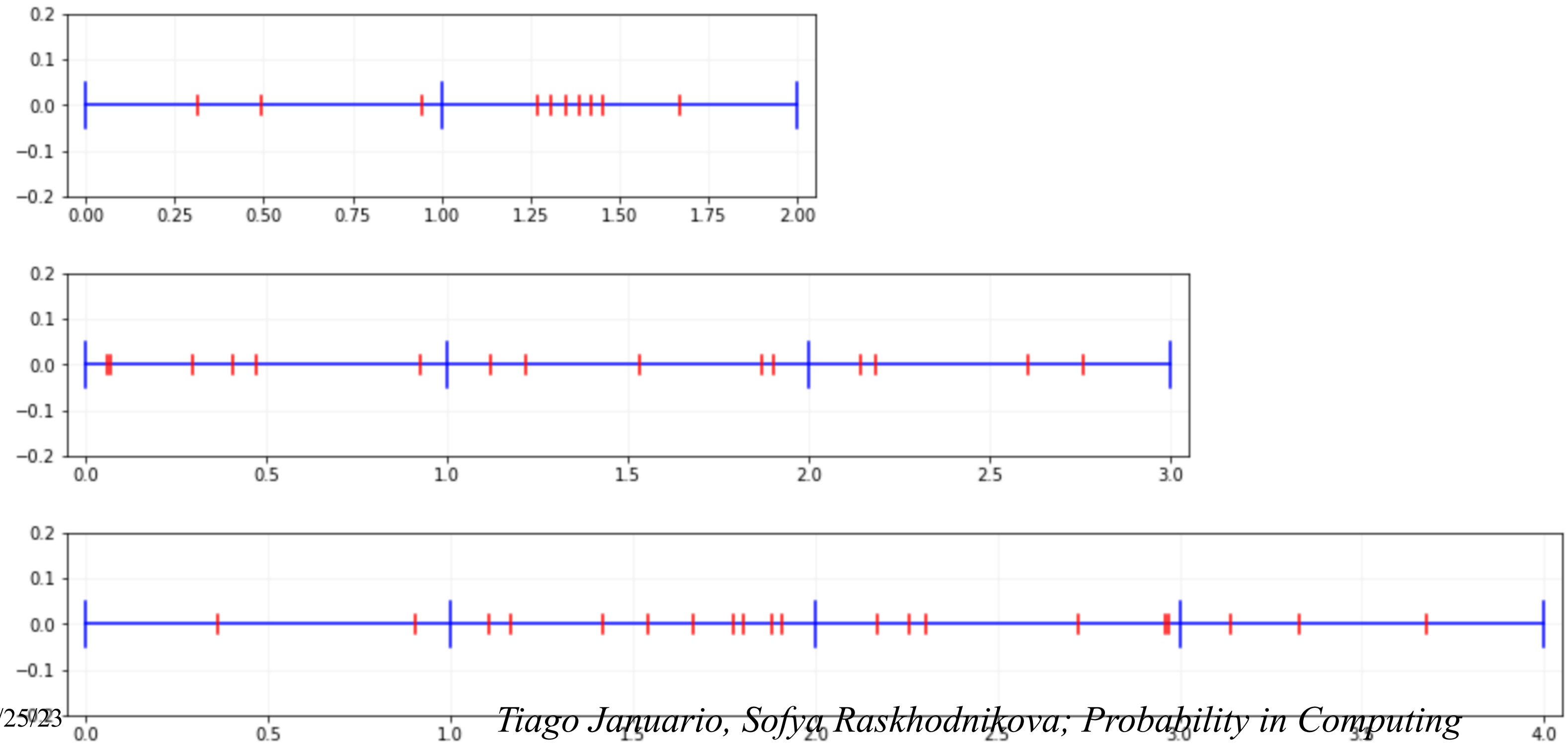
Now suppose we generate 5 random arrivals in [0 .. 1), 10 random arrivals in [0 .. 2), 15 arrivals in [0 .. 3), 20 in [0 .. 4) and so on, to infinity...



Poisson Process

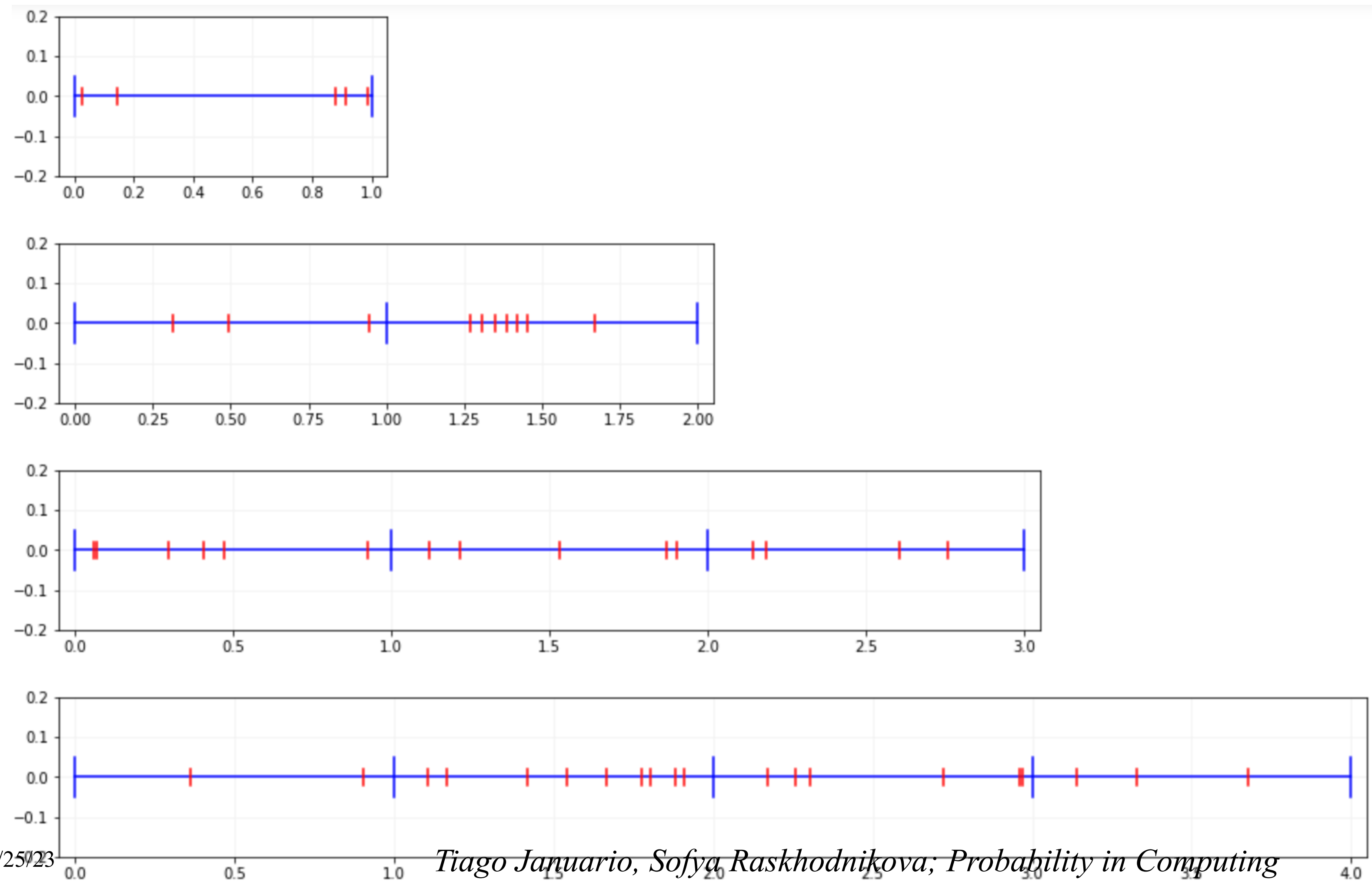
Since the arrivals are independent and distributed uniformly, the mean number of arrivals in each second [0 .. 1), [1 .. 2), [2 .. 3), etc. is still 5.





Poisson Process

RATE = # of arrivals per second is always constant at 5



Poisson Process

Definition: A Poisson Process is a sequence of arrivals over time, where:

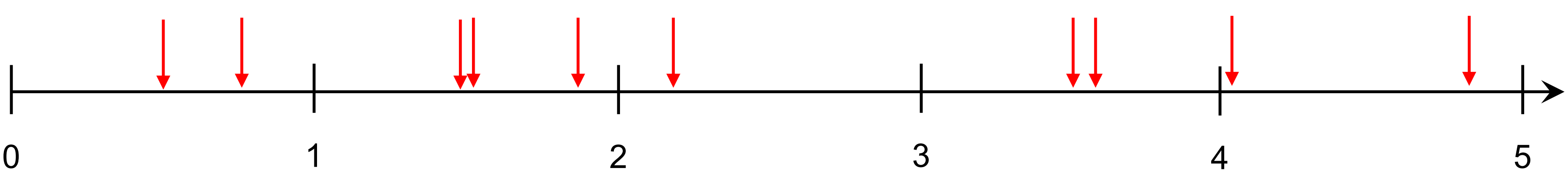
1) The mean arrival rate

$$\lambda = \frac{\text{Mean Number of Arrivals}}{\text{Unit Time}}$$

is a constant over all time, for any unit interval [t .. t+1)

1) The number of arrivals in two non-overlapping intervals is independent; and

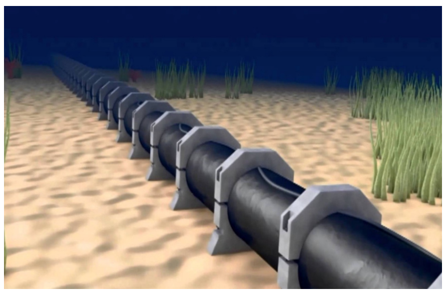
2) The probability of two events occurring at the same time is 0.



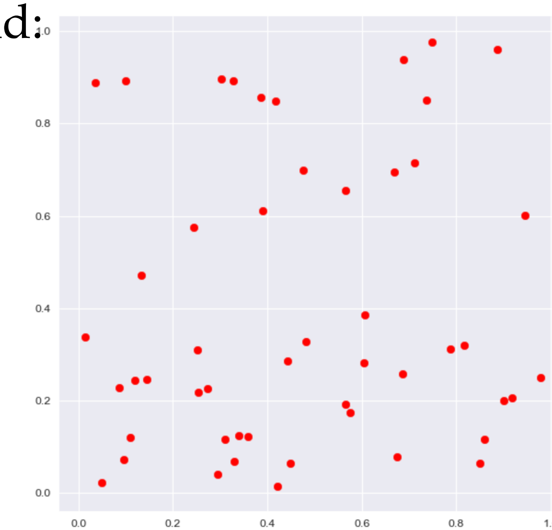
Poisson Process

It is also possible that the continuous dimension is distance in space, in 1 dimension or more than 1. Examples include:

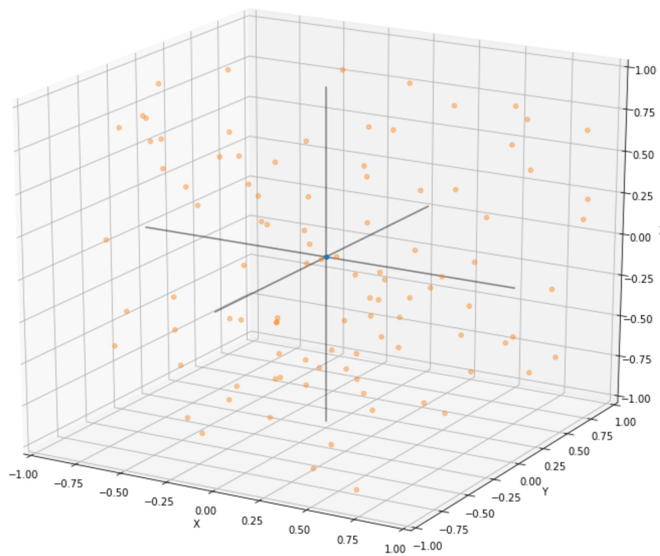
The occurrence of leaks in an undersea pipeline (1D):

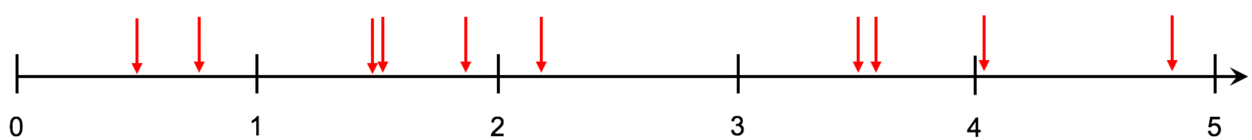


Location of trees in a 1 square mile plot of land:



Location of supernovas in a given cubic gigaparsec volume of space in the last billion years:



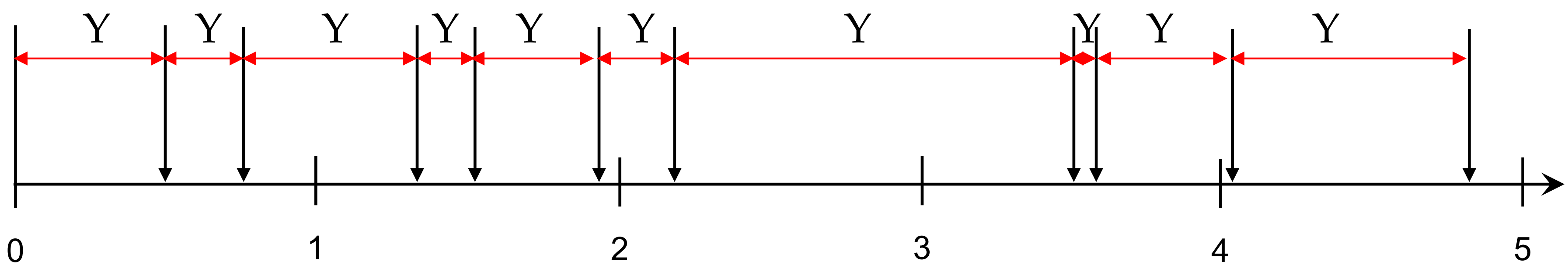


The important point is that events (discrete) occur along 1 or more (continuous) dimensions.

Interarrival Times of a Poisson Process

Suppose we have a Poisson Process, and instead of counting the number of arrivals in each unit interval, we look at the interarrival times, i.e., the amount of time between each arrival.

Intuitively, this is a natural thing to think about: **How long before the next event?**



Let's define the random variable Y = "the arrival time of the first event."

In fact, because the arrivals are independent, at any time t, probabilistically the Poisson process starts all over again (the events don't remember the past!), so in fact

Y is the distribution of the interarrival times, where

$$Y \sim \text{Exponential}(\lambda) \quad \text{where} \quad \lambda = \frac{\text{Mean Number of Arrivals}}{\text{Unit Time}}$$

Poisson Process: Examples

THE REALLY BIG ONE

An earthquake will destroy a sizable portion of the coastal Northwest. The question is when.

By Kathryn Schulz

When the 2011 earthquake and tsunami struck Tohoku, Japan, Chris Goldfinger was two hundred miles away, in the city of Kashiwa, at an international meeting on seismology. As the shaking started, everyone in the room began to laugh. Earthquakes are common in Japan—that one was the third of the week—and the participants were, after all, at a seismology

Poisson Process: Examples

Evidence shows the expected period of megaquakes is once every 243 years. Last one was January 26th, 1700, or 323 years ago. What is probability it will occur in the next ten years?

THE REALLY BIG ONE

An earthquake will destroy a sizable portion of the coastal Northwest. The question is when.

By Kathryn Schulz

When the 2011 earthquake and tsunami struck Tohoku, Japan, Chris Goldfinger was two hundred miles away, in the city of Kashiwa, at an international meeting on seismology. As the shaking started, everyone in the room began to laugh. Earthquakes are common in Japan—that one was the third of the week—and the participants were, after all, at a seismology

Poisson Random Variables

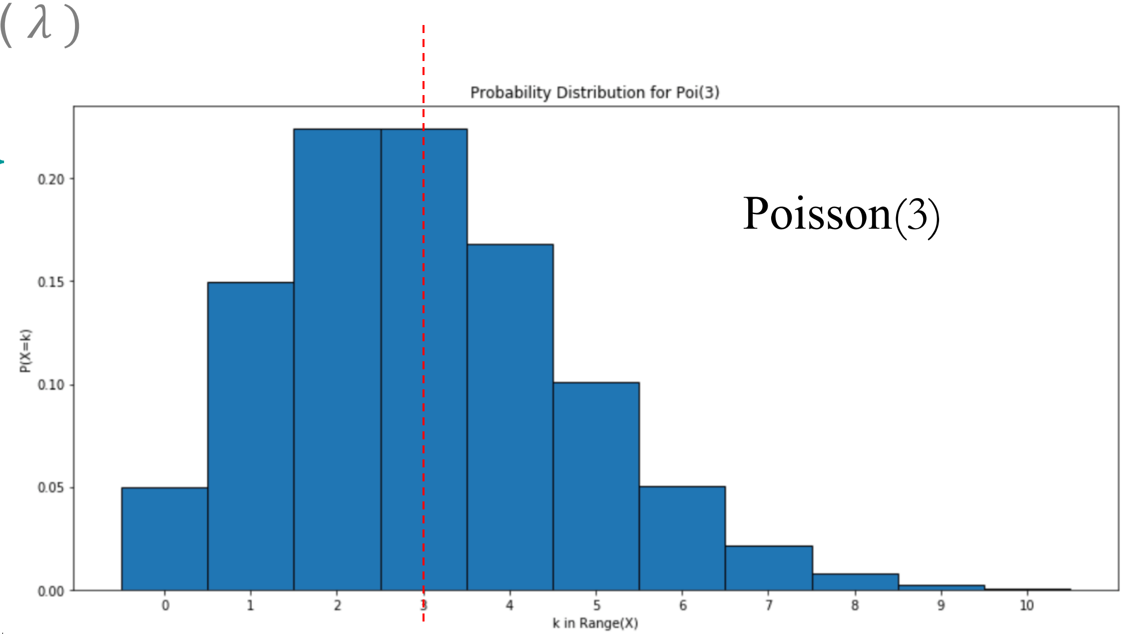
- Consider a Poisson Process
- Fix the unit time interval (e.g., 1 second or 1 year)
- Supposed the mean number of arrivals in a unit interval is λ
- Let X be the number of arrivals in a unit interval. (It has the same distribution for each unit interval.)
- Then we call X a **Poisson Random Variable with rate parameter λ**, denoted $X \sim \text{Poisson}(\lambda)$

$$\text{Range}(X) = \{0,1,2,\ldots\}$$

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

$$\mathbb{E}(X) = \lambda$$

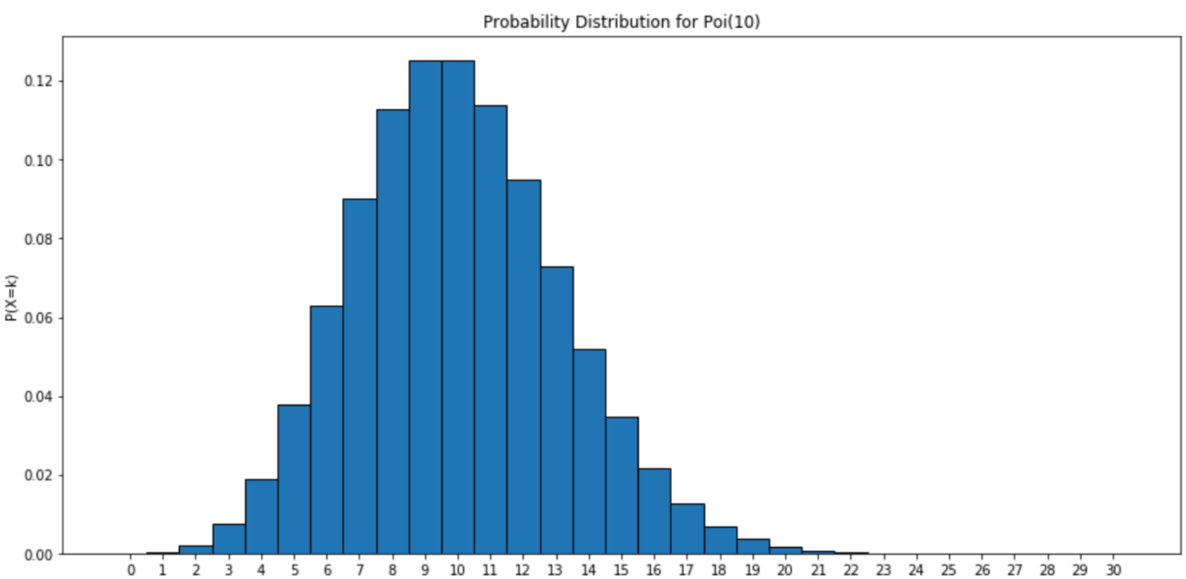
$$\text{Var}(X) = \lambda$$



Poisson Random Variables: Example

- Assume that arrivals of email in my Inbox are a Poisson Process with rate $\lambda = 10$ messages per hour.

What is the probability that I get exactly 10 emails in the next hour?



$$X \sim Poi(10)$$

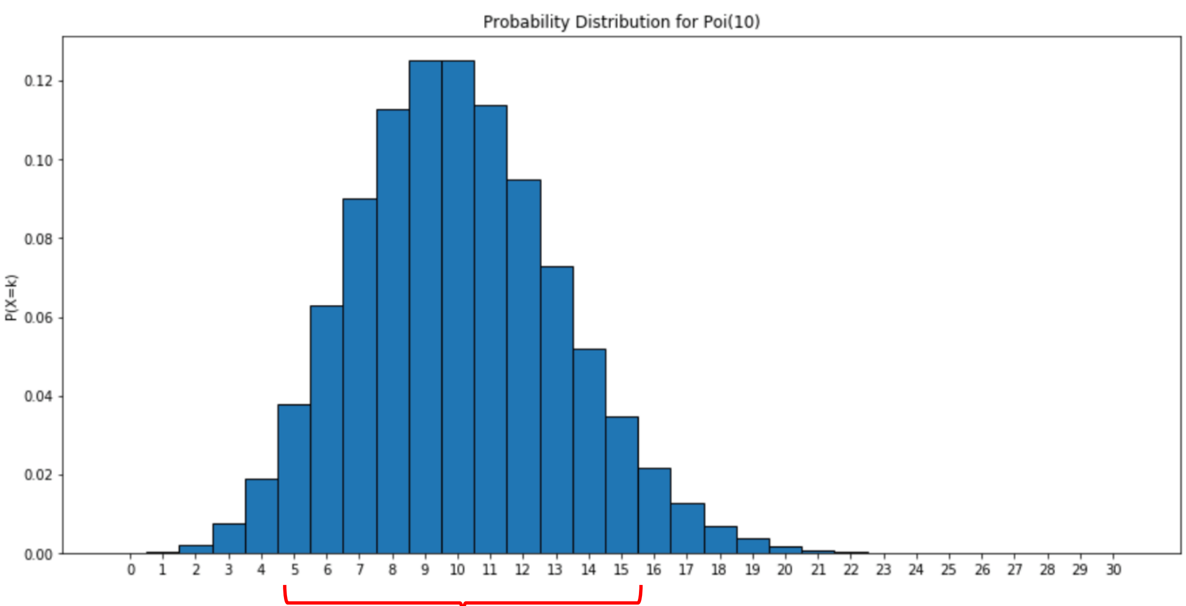
$$R_X = \{0, 1, 2, 3, \ldots\}$$

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

Poisson Random Variables: Example

- Assume that arrivals of email in my Inbox are a Poisson Process with rate λ = 10 messages per hour.

What is the probability that I get between 5 and 15 emails (inclusive) in the next hour?



$P(5 \leq X \leq 15)$

$$X \sim Poi(10)$$

$$R_X = \{0, 1, 2, 3, \ldots\}$$

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