CSE 312

Foundations of Computing II

Lecture 13: The Poisson Distribution

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Anna R. Karlin

Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Alex Tsun, Rachel Lin, Hunter Schafer & myself ☺

Zoo of Discrete RVs! 🦍🐘🦁🐅🦓🐪🦒

$X \sim \text{Unif}(a, b)$

$$P(X = k) = \frac{1}{b - a + 1}$$
$$E[X] = \frac{a + b}{2}$$
$$Var(X) = \frac{(b - a)(b - a + 2)}{12}$$

$X \sim \text{Ber}(p)$

$$P(X = 1) = p, P(X = 0) = 1 - p$$
$$E[X] = p$$
$$Var(X) = p(1 - p)$$

$X \sim \text{Bin}(n, p)$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$
$$E[X] = np$$
$$Var(X) = np(1 - p)$$

$X \sim \text{Geo}(p)$

$$P(X = k) = (1 - p)^{k-1} p$$
$$E[X] = \frac{1}{p}$$
$$Var(X) = \frac{1 - p}{p^2}$$

$X \sim \text{NegBin}(r, p)$

$$P(X = k) = \binom{k - 1}{r - 1} p^r (1 - p)^{k-r}$$
$$E[X] = \frac{r}{p}$$
$$Var(X) = \frac{r(1 - p)}{p^2}$$

$X \sim \text{HypGeo}(N, K, n)$

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$
$$E[X] = n\frac{K}{N}$$
$$Var(X) = n\frac{K(N - K)(N - n)}{N^2(N - 1)}$$

Agenda

- Poisson Distribution
- Approximate Binomial distribution using Poisson distribution

Preview: Poisson

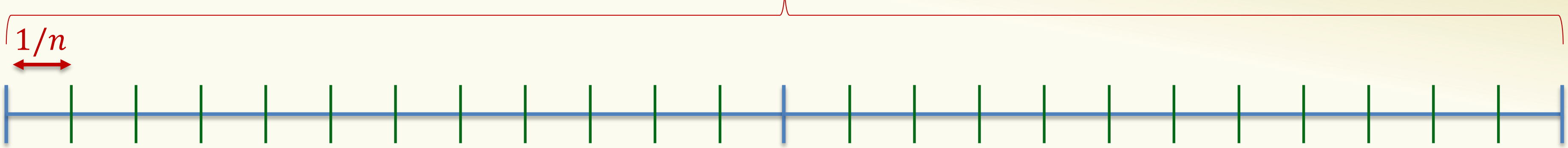
Model: # events that occur in an hour

- Expect to see 3 events per hour (but will be random)
- The expected number of events in t hours, is 3t
- Occurrence of events on disjoint time intervals is independent

Example – Model cars passing through a certain town in 1 hour

X = # cars passing through a certain town in 1 hour

Divide 1 hour into n intervals each of length $1/n$



https://pollev.com/ annakarlin185

What should p be?

Poll:

A. 3/n

B. 3n

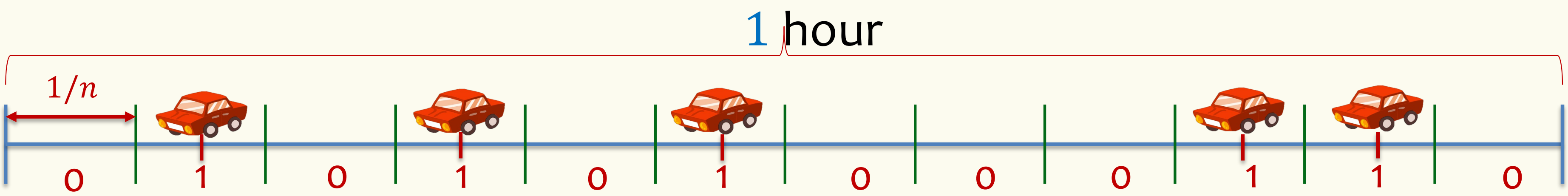
C. 3

D. 3/60

Example – Model the process of cars passing through a light in 1 hour

X = # cars passing through a light in 1 hour

Know: $\mathbb{E}(X) = \lambda$ for some given $\lambda > 0$



Discretize problem: n intervals, each of length $\frac{1}{n}$.

In each interval, a car passes by with probability $\frac{\lambda}{n}$ (assume ≤ 1 car can pass by)

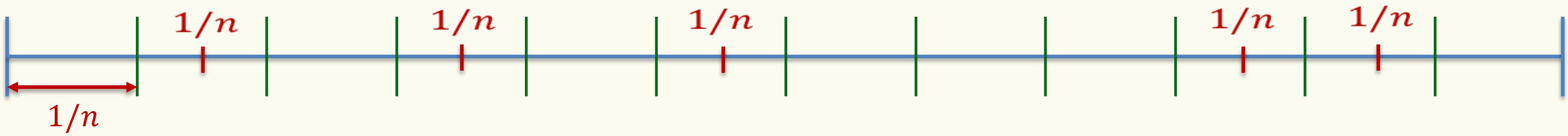
Bernoulli $X_i = 1$ if car in i-th interval (0 otherwise). $\mathbb{P}(X_i = 1) = \frac{\lambda}{n}$

$X = \sum_{i=1}^{n} X_i$ $X \sim$ Binomial(n,p) $\mathbb{P}(X = i) = \binom{n}{i}\left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}$

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

Don’t like discretization

X is <u>Binomial</u> $\mathbb{P}(X = i) = \binom{n}{i}\left(\frac{\lambda}{n}\right)^i\left(1-\frac{\lambda}{n}\right)^{n-i}$



We want now $n \to \infty$

$$\mathbb{P}(X = i) = \binom{n}{i}\left(\frac{\lambda}{n}\right)^i\left(1-\frac{\lambda}{n}\right)^{n-i} = \underbrace{\frac{n!}{(n-i)!\,n^i}}_{\to 1}\frac{\lambda^i}{i!}\underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{\to e^{-\lambda}}\underbrace{\left(1-\frac{\lambda}{n}\right)^{-i}}_{\to 1}$$

$$\to \mathbb{P}(X = i) = e^{-\lambda}\cdot\frac{\lambda^i}{i!}$$

Poisson Distribution

Siméon Denis Poisson
1781-1840

- Suppose "events" happen, independently, at an *average* rate of λ per unit time.
- Let X be the *actual* number of events happening in a given time unit. Then X is a *Poisson* r.v. *with parameter* λ (denoted X ~ Poi(λ)) and has distribution (PMF):

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

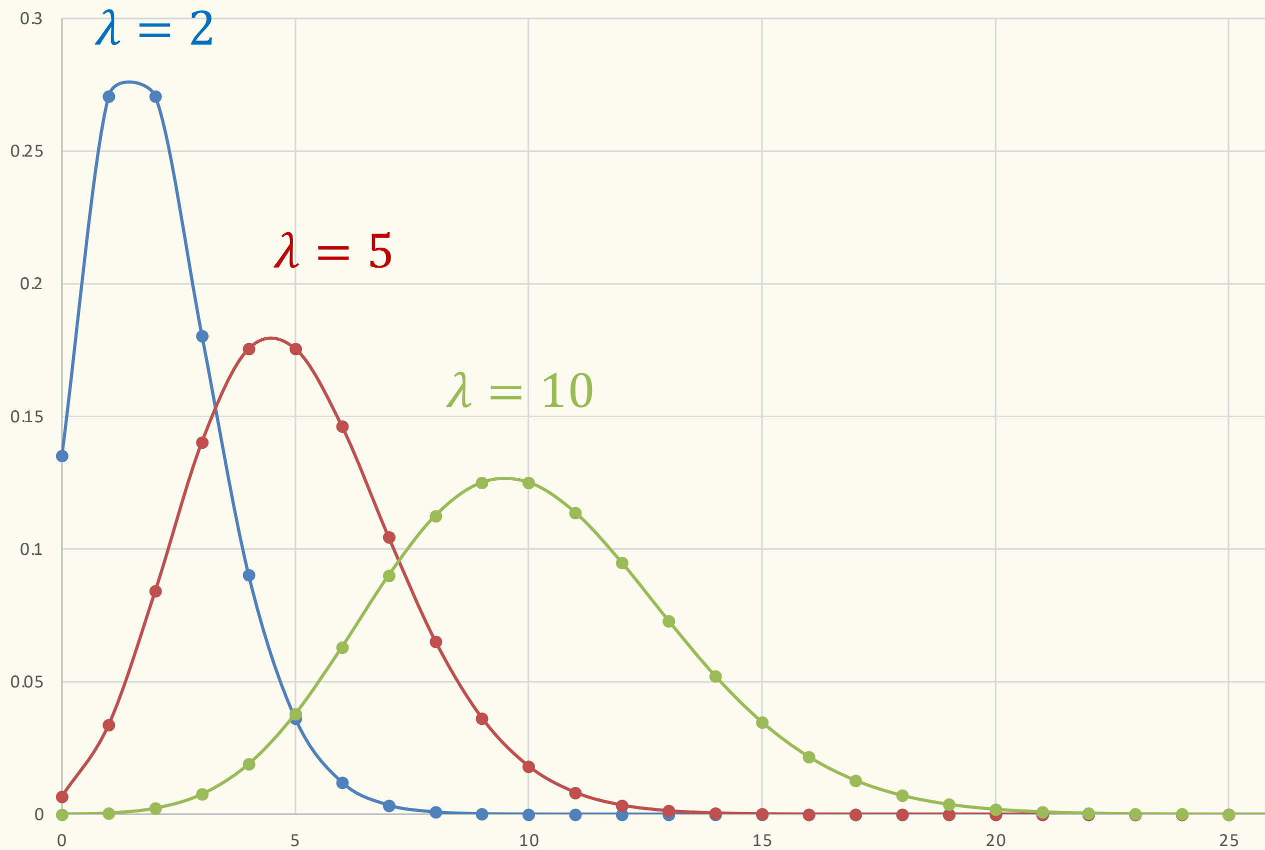
Several examples of "Poisson processes":

- # of cars passing through a certain town <u>in 1 hour</u>
- # of requests to web servers <u>in a minute</u>
- # of photons hitting a light detector <u>in a given interval</u>
- # of patients arriving to ER <u>within an hour</u>

Assume fixed average rate

Probability Mass Function

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$



Validity of Distribution

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

We first want to verify that Poisson probabilities sum up to 1.

$$\sum_{i=0}^{\infty} \mathbb{P}(X = i) =$$

Fact. $\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$

Validity of Distribution

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

We first want to verify that Poisson probabilities sum up to 1.

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

Fact. $\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$

Expectation

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

Theorem. If X is a Poisson RV with parameter λ, then

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

Proof. $\mathbb{E}(X) = \sum_{i=0}^{\infty} i \cdot \mathbb{P}(X = i)$

Expectation

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

Theorem. If X is a Poisson RV with parameter λ, then

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

Proof.

$$\mathbb{E}(X) = \sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} \cdot i = \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!}$$

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

= 1 (see prior slides!)

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

Variance

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

Theorem. If X is a Poisson RV with parameter λ, then $\mathrm{Var}(X) = \lambda$

Proof. $\mathbb{E}(X^2) = \sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} \cdot i^2 \quad = \lambda^2 + \lambda$

$$\Rightarrow \quad \mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Variance

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

Theorem. If X is a Poisson RV with parameter λ, then $\mathrm{Var}(X) = \lambda$

Proof. $\mathbb{E}(X^2) = \sum_{i=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{i!} \cdot i^2 = \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!} i$

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

$$= \lambda \left[\underbrace{\sum_{j=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^j}{j!} \cdot j}_{= \mathbb{E}(X) = \lambda} + \underbrace{\sum_{j=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^j}{j!}}_{= 1} \right] = \lambda^2 + \lambda$$

Similar to the previous proof
Verify offline.

$$\Rightarrow \quad \mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Poisson Random Variables

Definition. A **Poisson random variable** X with parameter $\lambda \geq 0$ is such that for all $i = 0,1,2,3 \ldots,$

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

Poisson approximates Binomial when n is very large, p is very small, and $\lambda = np$ is "moderate" (e.g. $n > 20$ and $p < 0.05$, $n > 100$ and $p < 0.1$)

Formally, Binomial is Poisson in the limit as
$n \to \infty$ (equivalently, $p \to 0$) while holding $np = \lambda$

From Binomial to Poisson

$X \sim \text{Bin}(n, p)$

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

$$E[X] = np$$

$$Var(X) = np(1-p)$$

$$n \to \infty$$
$$np = \lambda$$
$$p = \frac{\lambda}{n} \to 0$$

$X \sim \text{Poisson}(\lambda)$

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

$$E[X] = \lambda$$

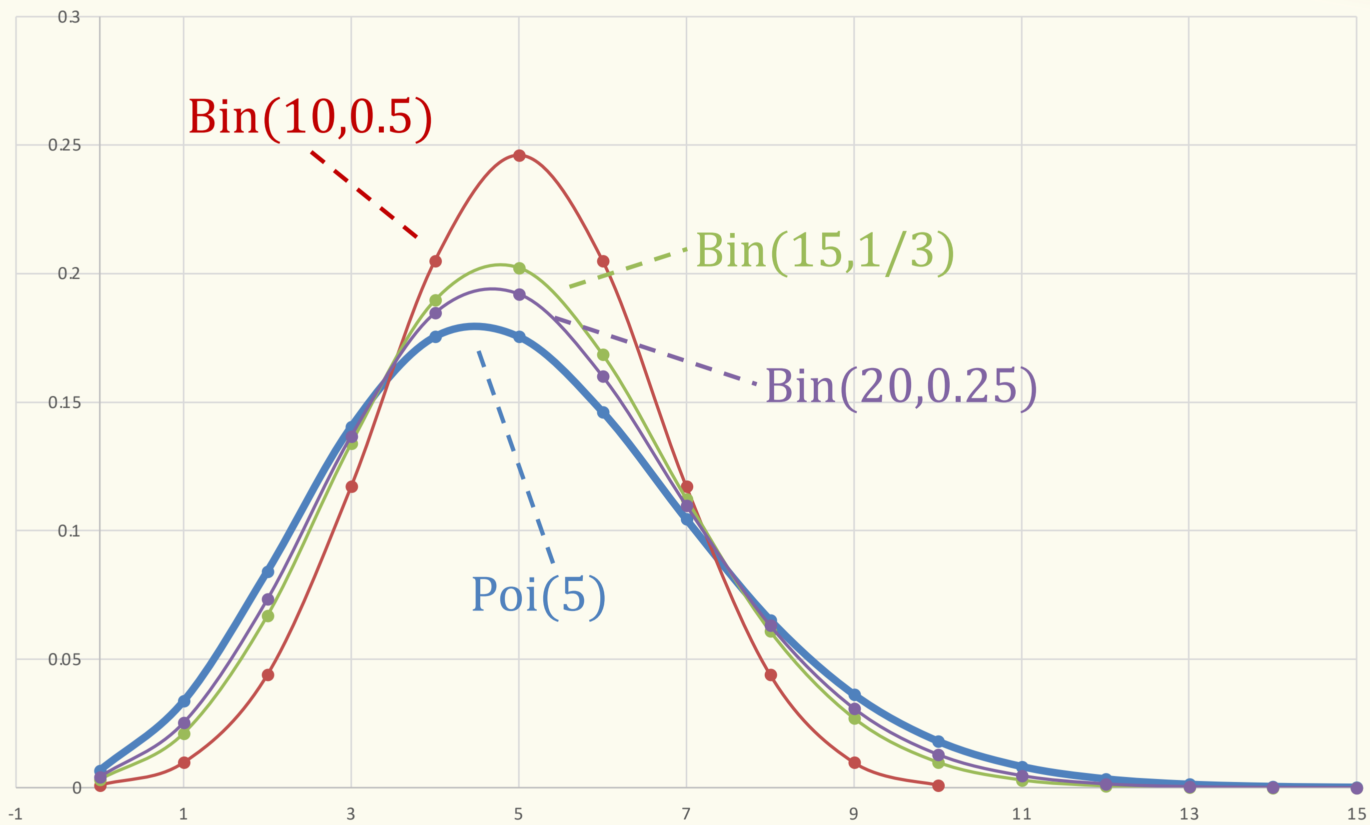
$$Var(X) = \lambda$$

Probability Mass Function – Convergence of Binomials

$\lambda = 5$

$p = \frac{5}{n}$

$n = 10,15,20$



$as\ n \to \infty,\ \ Binomial(n,\ \ p = \ \lambda/n) \to poi(\lambda)$

Example -- Approximate Binomial Using Poisson

Consider sending bit string over a network

- Send bit string of length **n = 10^4**
- Probability of (independent) bit corruption is **p = 10^{-6}**
- What is probability that message arrives uncorrupted?

Using Y ~ Bin(10^4, 10^{-6})

$$\mathbb{P}(Y = 0)$$

Using X ~ Poi(λ = np = $10^4 \bullet 10^{-6}$ = 0.01)

$$\mathbb{P}(X = 0)$$

Example -- Approximate Binomial Using Poisson

Consider sending bit string over a network

- Send bit string of length $n = 10^4$
- Probability of (independent) bit corruption is $p = 10^{-6}$
- What is probability that message arrives uncorrupted?

Using $Y \sim \text{Bin}(10^4, 10^{-6})$

$$\mathbb{P}(Y = 0) \approx 0.990049829$$

Using $X \sim \text{Poi}(\lambda = np = 10^4 \bullet 10^{-6} = 0.01)$

$$\mathbb{P}(X = 0) = e^{-\lambda} \cdot \frac{\lambda^0}{0!} = e^{-0.01} \cdot \frac{0.01^0}{0!} = 0.990049834$$

Sum of Independent Poisson RVs

Theorem. Let $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$ such that $\lambda = \lambda_1 + \lambda_2$.
Let $\mathrm{Z} = (X + Y)$. For all $k = 0,1,2,3 \ldots,$

$$\mathbb{P}(Z = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

More generally, let $X_1 \sim Poi(\lambda_1), \cdots, X_n \sim Poi(\lambda_n)$ such that $\lambda = \Sigma_i \lambda_i$.
Let $\mathrm{Z} = \Sigma_i X_i$

$$\mathbb{P}(Z = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

Sum of Independent Poisson RVs

Theorem. Let $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$ such that $\lambda = \lambda_1 + \lambda_2$.
Let $Z = (X + Y)$. For all $k = 0,1,2,3 \ldots,$

$$\mathbb{P}(Z = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

https://pollev.com/ annakarlin185

$\mathbb{P}(Z = k) = ?$

1. $\mathbb{P}(Z = k) = \Sigma_{j=0}^{k} \mathbb{P}(X = j, Y = k - j)$
2. $\mathbb{P}(Z = k) = \Sigma_{j=0}^{\infty} \mathbb{P}(X = j, Y = k - j)$
3. $\mathbb{P}(Z = k) = \Sigma_{j=0}^{k} \mathbb{P}(Y = k - j | X = j)\mathbb{P}(X = j)$
4. $\mathbb{P}(Z = k) = \Sigma_{j=0}^{k} \mathbb{P}(Y = k - j | X = j)$

Poll:
A. All of them are right
B. The first 3 are right
C. Only 1 is right
D. Don't know

1. $\mathbb{P}(Z = k\ ) = \Sigma_{j=0}^{k}\ \mathbb{P}(X = j, Y = k - j)$
2. $\mathbb{P}(Z = k) = \Sigma_{j=0}^{\infty}\mathbb{P}(X = j, Y = k - j)$
3. $\mathbb{P}(Z = k) = \Sigma_{j=0}^{k}\mathbb{P}(Y = k - j|X = j)\mathbb{P}(X = j)$
4. $\mathbb{P}(Z = k) = \Sigma_{j=0}^{k}\mathbb{P}(Y = k - j|X = j)$

$$\mathbb{P}(Z = k) = \Sigma_{j=0}^{k} \mathbb{P}(X = j, Y = k - j)$$ Law of total probability

$$= \Sigma_{j=0}^{k} \mathbb{P}(X = j)\mathbb{P}(Y = k - j)$$ Independence

$$\mathbb{P}(Z = z) = \Sigma_{j=0}^{k} \mathbb{P}(X = j, Y = z - j)$$

Law of total probability

$$= \Sigma_{j=0}^{k} \mathbb{P}(X = j)\mathbb{P}(Y = z - j) = \Sigma_{j=0}^{k} \; e^{-\lambda_1} \cdot \frac{\lambda_1^j}{j!} \cdot e^{-\lambda_2} \cdot \frac{\lambda_2^{z-j}}{z - j!}$$

Independence

$$= e^{-\lambda} \left(\Sigma_{j=0}^{k} \; \cdot \frac{1}{j! \, z - j!} \cdot \lambda_1^j \lambda_2^{z-j} \right)$$

$$= e^{-\lambda} \left(\Sigma_{j=0}^{k} \; \frac{z!}{j! \, z - j!} \cdot \lambda_1^j \lambda_2^{z-j} \right) \frac{1}{z!}$$

$$= e^{-\lambda} \cdot (\lambda_1 + \lambda_2)^z \cdot \frac{1}{z!} = e^{-\lambda} \cdot \lambda^z \cdot \frac{1}{z!}$$

Binomial Theorem

Poisson Random Variables

Definition. A **Poisson random variable** X with parameter $\lambda \geq 0$ is such that for all $i = 0,1,2,3 \ldots,$

$$\mathbb{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

General principle:

- Events happen at an average rate of λ per time unit
- Number of events happening at a time unit X is distributed according to Poi(λ)
- Poisson approximates Binomial when n is large, p is small, and np is moderate
- Sum of independent Poisson is still a Poisson

Next Time

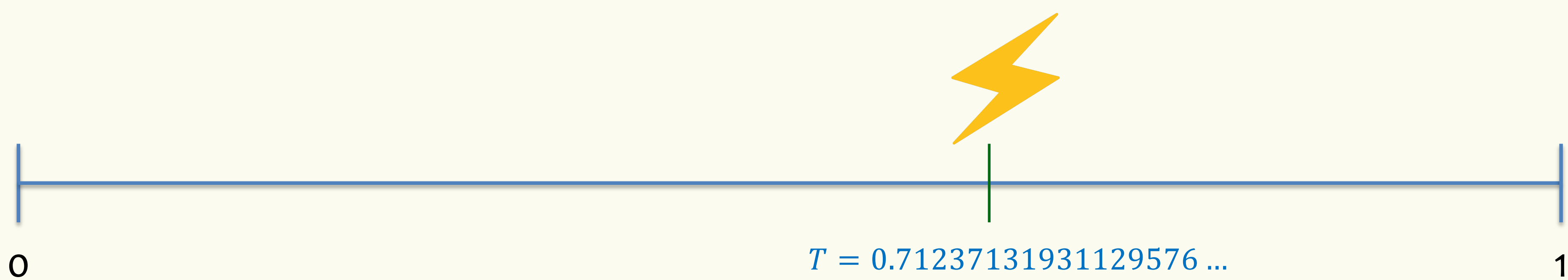
- Continuous Random Variables
- Probability Density Function
- Cumulative Density Function

Often we want to model experiments where the outcome is <u>not</u> discrete.

Example – Lightning Strike

Lightning strikes a pole within a one-minute time frame

- T = time of lightning strike
- Every time within $[0,1]$ is equally likely
  - Time measured with infinitesimal precision.

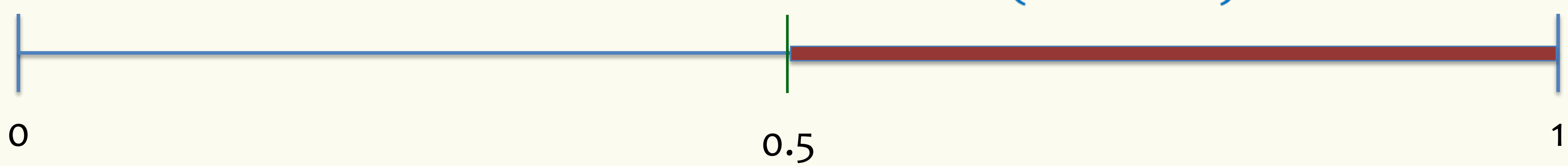


The outcome space is not discrete

Lightning strikes a pole within a one-minute time frame

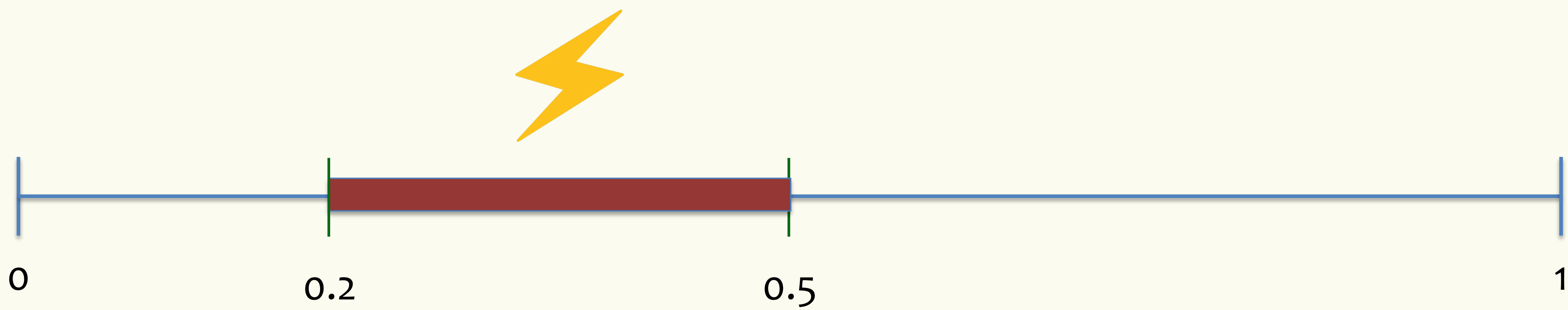
- T = time of lightning strike
- Every point in time within $[0,1]$ is equally likely

$$\mathbb{P}(T \geq 0.5) =$$



Lightning strikes a pole within a one-minute time frame

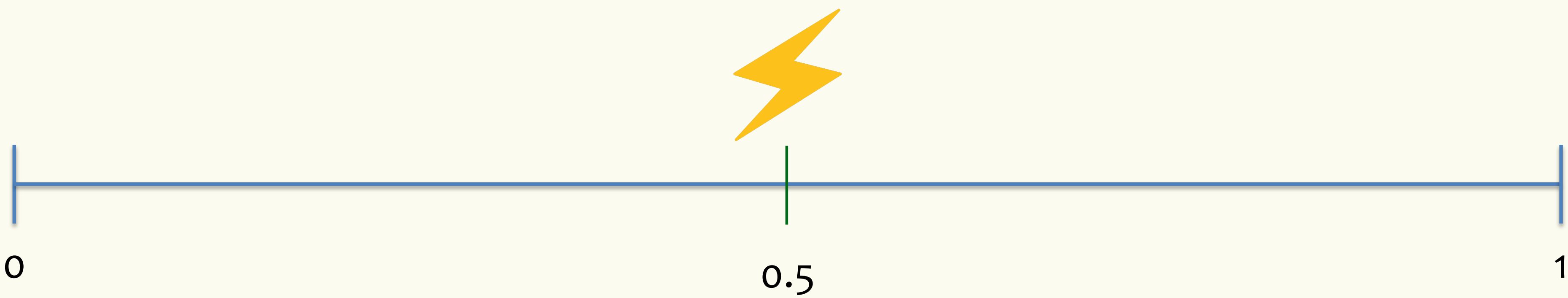
- T = time of lightning strike
- Every point in time within $[0,1]$ is equally likely



$$\mathbb{P}(0.2 \leq T \leq 0.5) =$$

Lightning strikes a pole within a one-minute time frame

- T = time of lightning strike
- Every point in time within $[0,1]$ is equally likely



$$\mathbb{P}(T = 0.5) =$$

Bottom line

- This gives rise to a different type of random variable
- $\mathbb{P}(T = x) = 0$ for all $x \in [0,1]$
- Yet, somehow we want
  - $\mathbb{P}(T \in [0,1]) = 1$
  - $\mathbb{P}(T \in [a, b]) = b - a$
  - …
- How do we model the behavior of T?