CSE 312

Foundations of Computing II

Lecture 13: Poisson Distribution

Announcements

- Midterm info is posted
  - Q&A session next Tuesday 4pm on Zoom
  - Practice midterm + other practice materials posted this Wednesday
- Office hour updates
  - Prof. Tessaro: Office hour today already happened. No Friday OH.

Zoo of Random Variables

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

$$P(X = k) = \frac{1}{b - a + 1}$$

$$E[X] = \frac{a + b}{2}$$

$$\text{Var}(X) = \frac{(b-a)(b-a+2)}{12}$$

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

$$P(X = 1) = p, P(X = 0) = 1 - p$$

$$E[X] = p$$

$$\text{Var}(X) = p(1 - p)$$

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

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

$$E[X] = np$$

$$\text{Var}(X) = np(1 - p)$$

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

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

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

$$\text{Var}(X) = \frac{1 - p}{p^2}$$

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

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

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

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

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

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

$$E[X] = n\frac{K}{N}$$

$$\text{Var}(X) = n\frac{K(N-K)(N-n)}{N^2(N-1)}$$

# of coin flips until r heads

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 – Modelling car arrivals at an intersection

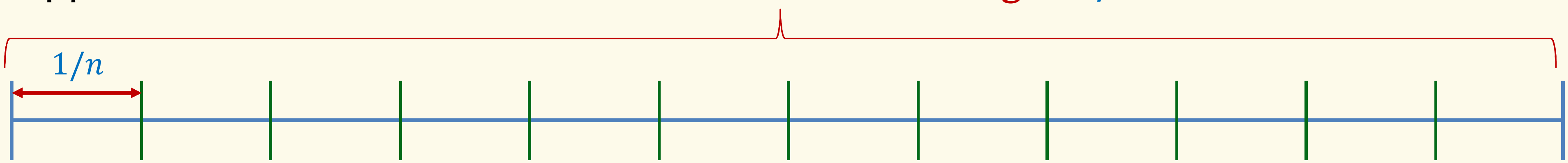
X = # of cars passing through a light in 1 hour

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

X = # cars passing through a light in 1 hour. $\mathbb{E}[X] = 3$ λ lambda

Assume: Occurrence of events on disjoint time intervals is independent

Approximation idea: Divide hour into n intervals of length $1/n$



This gives us n independent intervals

Assume at most one car per interval

$p =$ probability car arrives in an interval

What should p be?

pollev.com/paulbeame028

A. $3/n$

B. $3n$

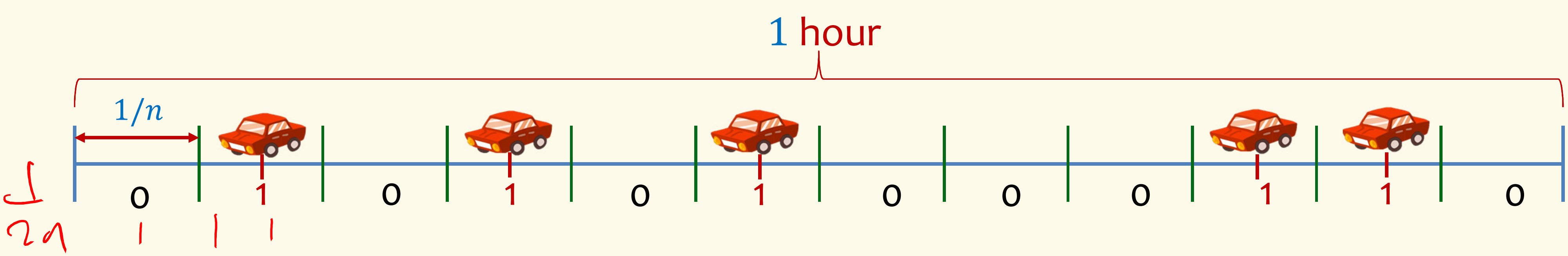
C. 3

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.   Disjoint time intervals are independent.

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



Discrete version: n intervals, each of length $1/n$.

In each interval, there is a car with probability $p = \lambda/n$ (assume ≤ 1 car can pass by)

Each interval is Bernoulli: $X_i = 1$ if car in i^{th} interval (0 otherwise). $P(X_i = 1) = \lambda/n$

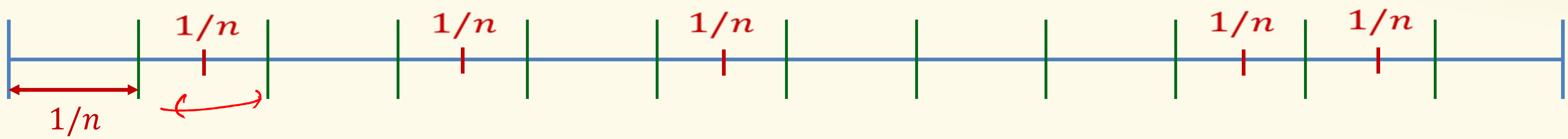
$X = \sum_{i=1}^{n} X_i$ $X \sim \text{Bin}(n, p)$ $P(X = i) = \binom{n}{i}\left(\frac{\lambda}{n}\right)^i\left(1 - \frac{\lambda}{n}\right)^{n-i}$

what kind is X?

indeed! $\mathbb{E}[X] = pn = \lambda$

Don't like discretization

X is binomial $P(X = i) = \binom{n}{i}\left(\frac{\lambda}{n}\right)^{i}\left(1-\frac{\lambda}{n}\right)^{n-i}$



We want now $n \to \infty$

Fix i, $n \to \infty$

$n(n-1)\cdots(n-i+1)$

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

$\frac{n(n-1)\cdots(n-i+1)}{n^i}$

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

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

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 $X \sim \mathrm{Poi}(\lambda)$) and has distribution (PMF):

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

fish in an hour

radioactive decay

frequency of earthquakes etc

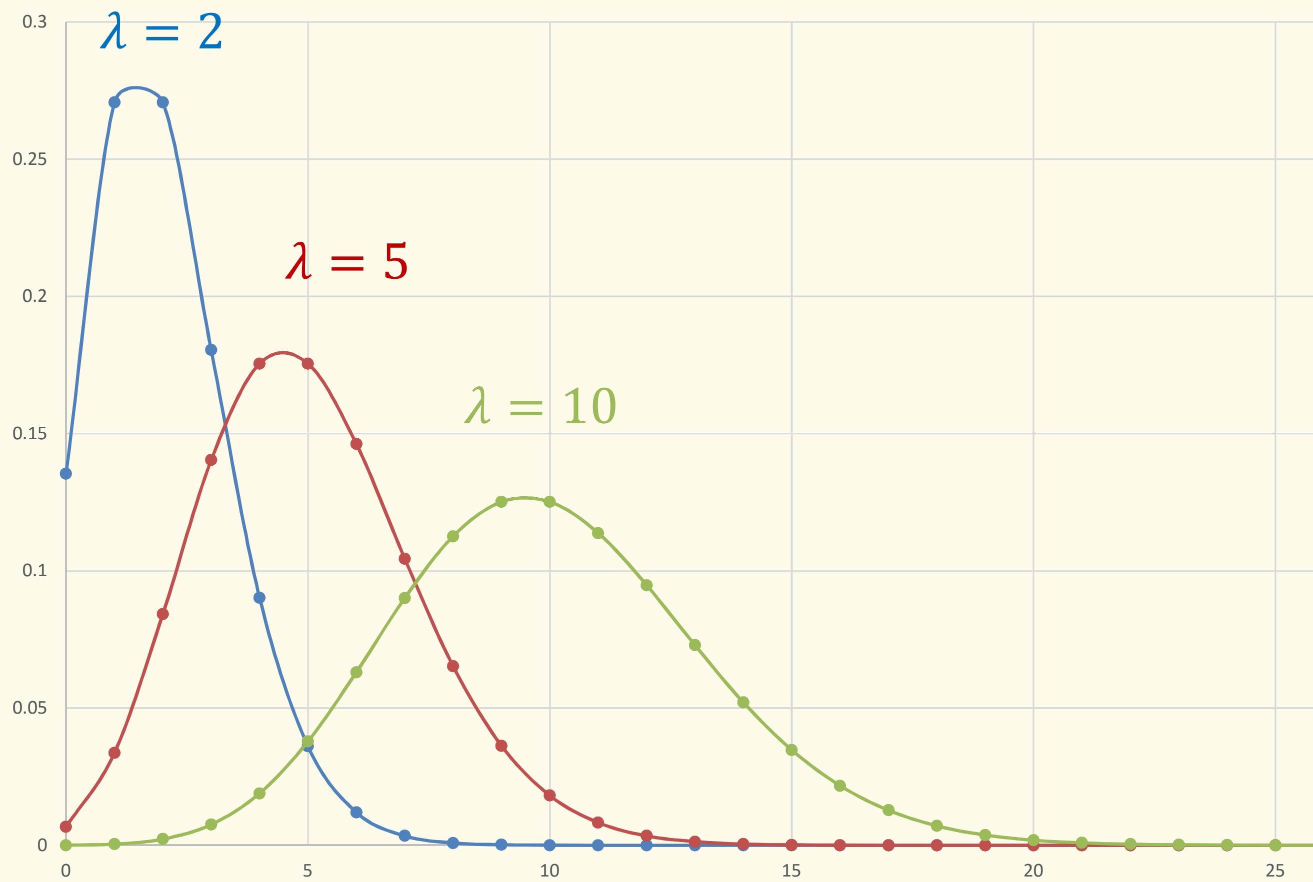
Several examples of "Poisson processes":

- # of cars passing through a traffic light <u>in 1 hour</u>
- # of requests to web servers <u>in an hour</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

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



Validity of Distribution

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

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

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

Fact (Taylor series expansion):

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

Expectation

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

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

$$\mathbb{E}[X] = \lambda$$

$i=0$ $\quad e^{-\lambda}\frac{\lambda^i}{i!} \cdot i = 0$

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

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

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

= 1 (see prior slides!)

Variance

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

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

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

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

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

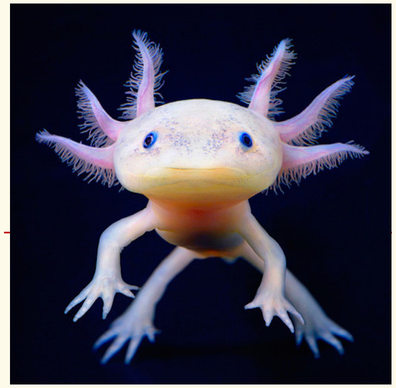
Similar to the previous proof
Verify offline.

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

Poisson Random Variables

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

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



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 $p < 0.1$)

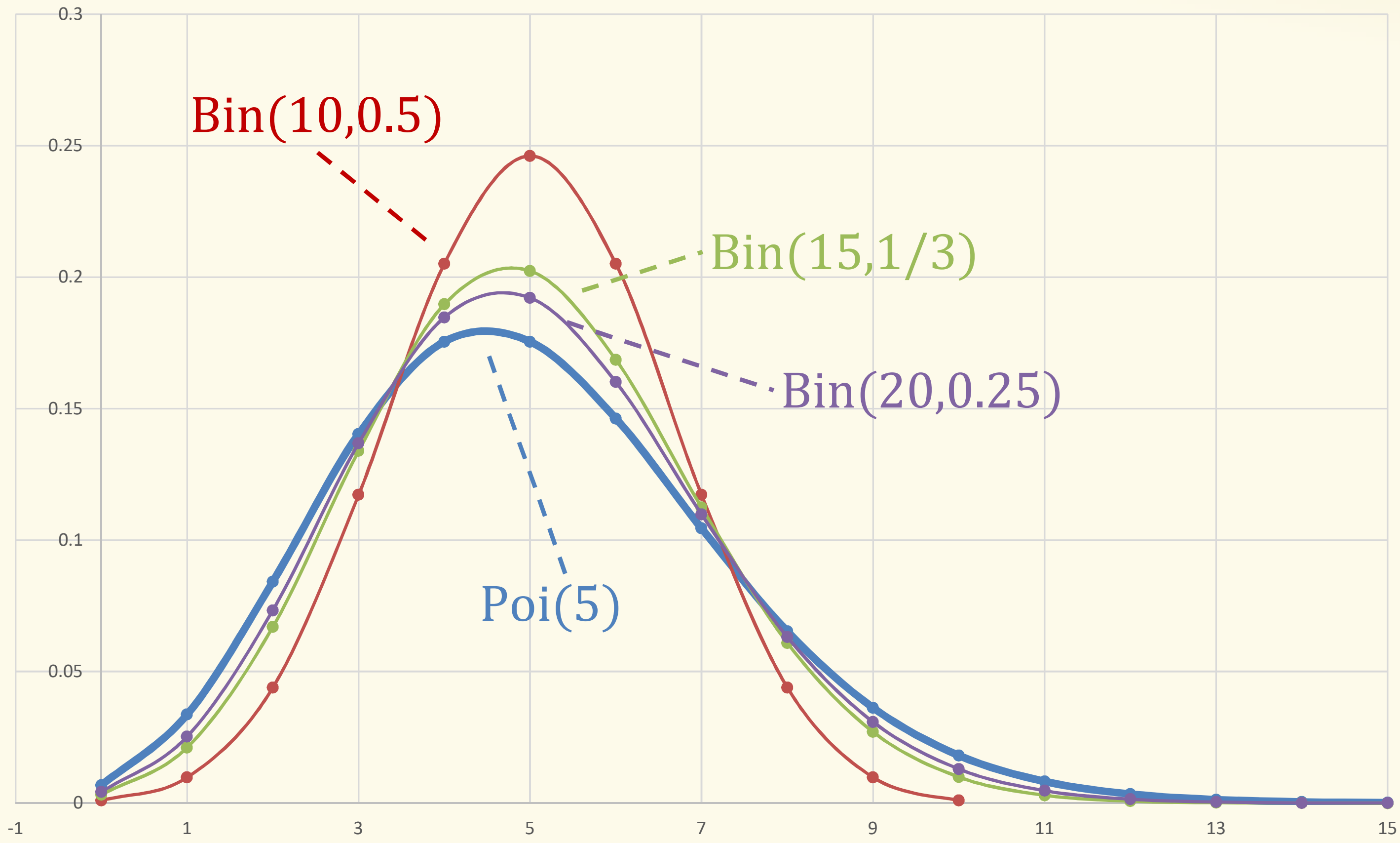
Formally, Binomial approaches Poisson in the limit as $n \to \infty$ (equivalently, $p \to 0$) while holding $np = \lambda$

Probability Mass Function – Convergence of Binomials

$\lambda = 5$

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

$n = 10,15,20$



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

From Binomial to Poisson

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

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

$E[X] = np$

$\text{Var}(X) = np(1-p)$

$n \to \infty$

$np = \lambda$

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

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

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

$E[X] = \lambda$

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

Example -- Approximate Binomial Using Poisson

Consider sending bit string over a network

- Send bit string of length $n = 104$ 10^4
- Probability of (independent) bit corruption is $p = 10^{-6}$

What is probability that message arrives uncorrupted?

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

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

10^4

Using $Y \sim \text{Bin}(104, 10^{-6})$

$$P(Y = 0) \approx 0.990049829$$

Sum of Independent Poisson RVs

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

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

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

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

Sum of Independent Poisson RVs

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

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

$P(Z = z) = ?$

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

pollev.com/paulbeame028

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

Proof

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

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

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

$\sum_{j=0}^{k} \binom{z}{j} \lambda_1^j \lambda_2^{z-j}$

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

$\binom{z}{j}$

Binomial Theorem

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

Poisson Random Variables

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

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

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

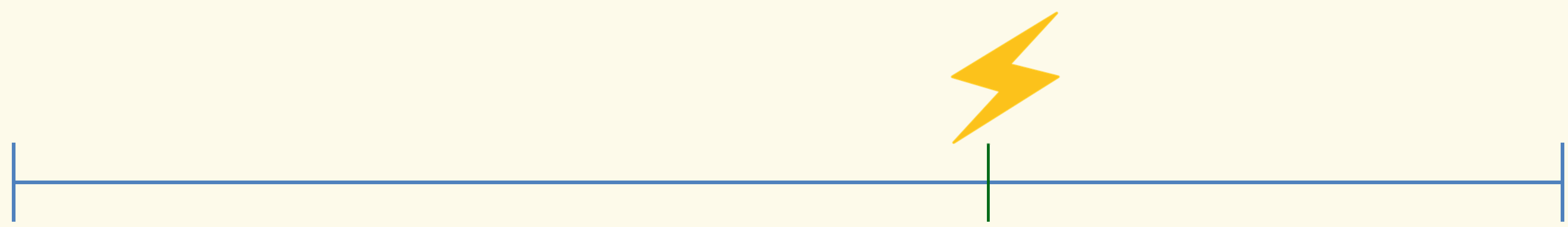
- Continuous Random Variables
- Probability Density Function
- Cumulative Density Function

Often we want to model experiments where the outcome is not 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.

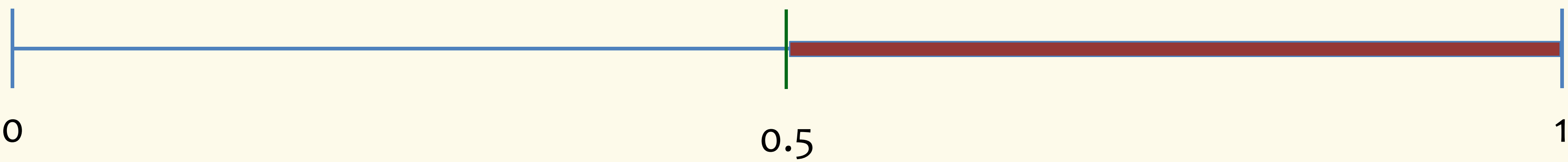


The outcome space is not discrete

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

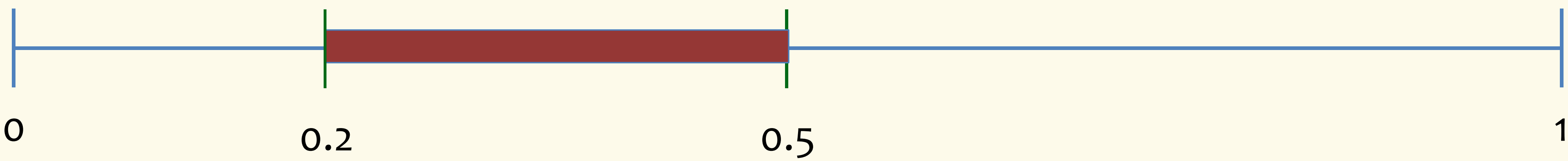
$$P(T \geq 0.5) = ½$$



Lightning strikes a pole within a one-minute time frame

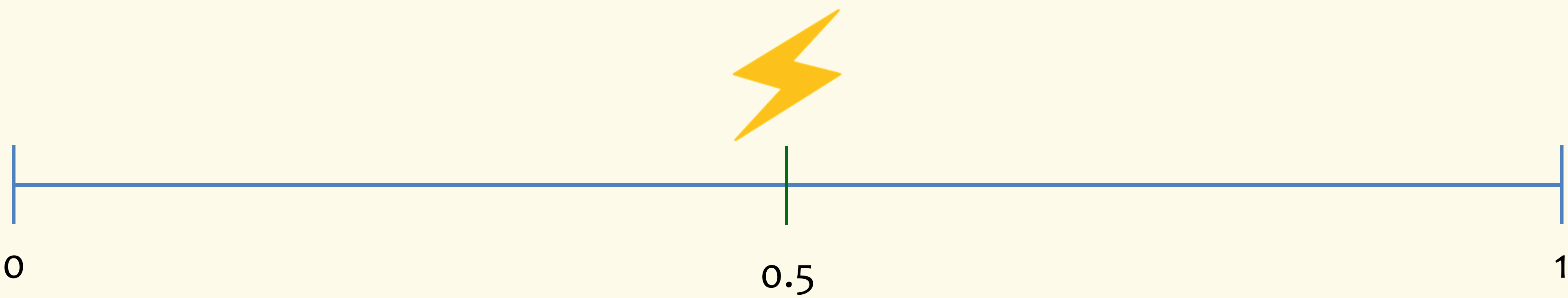
- T = time of lightning strike
- Every point in time within [0,1] is equally likely

$$P(0.2 \leq T \leq 0.5) = 0.5 - 0.2 = 0.3$$



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



$$P(T = 0.5) = 0$$

Bottom line

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