The Poisson Distribution

CSE 312 Spring 26

Lecture 13

Discrete Uniform Random Variables (2)

A discrete random variable X **equally likely** to take any (integer) value between integers a and b (inclusive), is uniform.

Notation: $X \sim \mathrm{Unif}(a, b)$

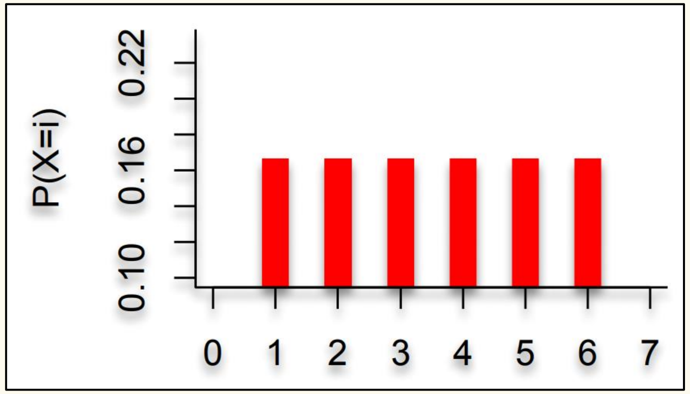
PMF: $\mathrm{P}(X = i) = \frac{1}{b - a + 1}$

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

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

Example: value shown on one roll of a fair die is $\mathrm{Unif}(1,6)$:

- $P(X = i) = 1/6$
- $\mathbb{E}[X] = 7/2$
- $\mathrm{Var}(X) = 35/12$



Bernoulli Random Variables (2)

A random variable X that takes value 1 ("Success") with probability p, and 0 ("Failure") otherwise. X is called a Bernoulli random variable.

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

PMF: $P(X = 1) = p,\ P(X = 0) = 1 - p$

Expectation: $\mathbb{E}[X] = p$ Note: $\mathbb{E}[X^2] = p$

Variance: $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1 - p)$

Examples:

- Coin flip
- Randomly guessing on a MC test question
- A server in a cluster fails
- Whether or not a share of a particular stock pays off or not
- Any indicator r.v.

Binomial Random Variables (3)

A discrete random variable $X = \sum_{i=1}^{n} Y_i$ where each $Y_i \sim \mathrm{Ber}(p)$.

Counts number of successes in n independent trials, each with probability p of success.

X is a Binomial random variable

Notation: $X \sim \mathrm{Bin}(n, p)$

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

Expectation: $\mathbb{E}[X] = np$

Variance: $\mathrm{Var}(X) = np(1-p)$

Examples:

- # of heads in n indep coin flips
- # of 1s in a randomly generated n bit string
- # of servers that fail in a cluster of n computers
- # of bit errors in file written to disk
- # of elements in a bucket of a large hash table
- # of n different stocks that “pay off”

Geometric Random Variables (2)

A discrete random variable X that models the number of independent trials $Y_i \sim \mathrm{Ber}(p)$ up until and including the first success.

X is called a Geometric random variable with parameter p.

Notation: $X \sim \mathrm{Geo}(p)$

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

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

Variance: $\mathrm{Var}(X) = \frac{1-p}{p^2}$

Examples:

- # of coin flips until first head
- # of random guesses on MC questions until you get one right
- # of random guesses at a password until you hit it
- # hash trials until a miner successfully mines a Bitcoin

Today

- Poisson Random Variables

Preview: Poisson

Model: X is # events that occur in an hour

- Events occur at a fixed rate (3 per hour), but at random times.
- The expected number of events in t hours, is $3t$
- Occurrence of events on disjoint time intervals is independent

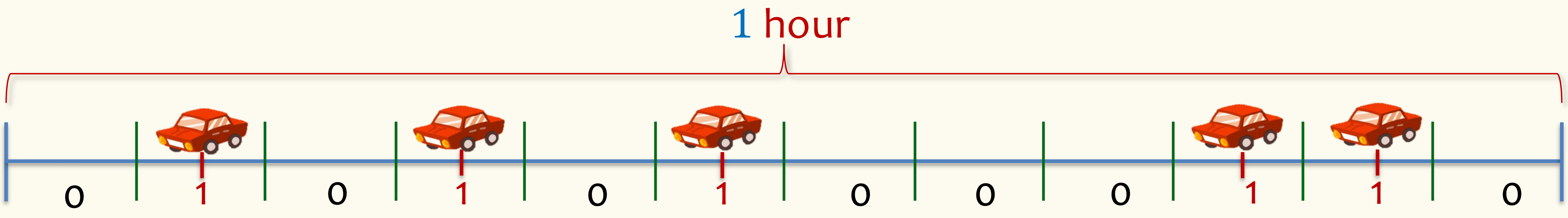
Example – Modelling car arrivals at an intersection

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

How to model?

X = # cars passing through an intersection in 1 hour. Constant rate of arrival.

Disjoint time intervals are independent. Know: $\mathbb{E}[X] = 3$

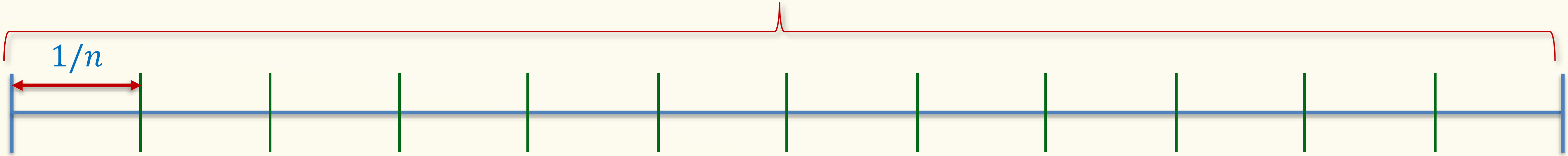


Example – Model the process of cars passing through a light in 1 hour

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

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

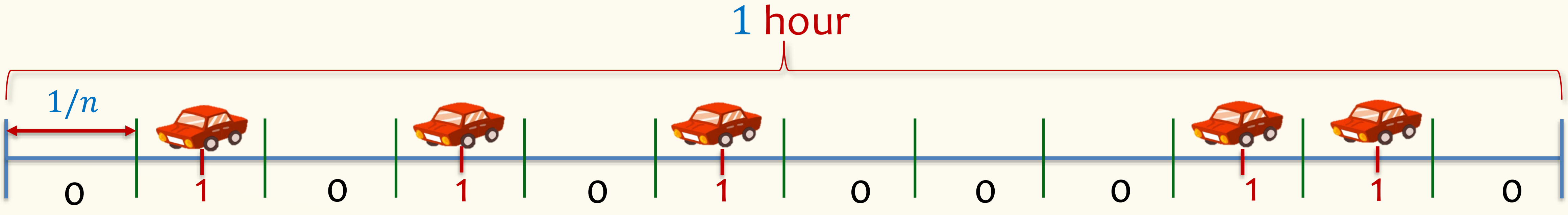
Assume either zero or one car per interval

$p =$ probability car arrives in a single interval

Model as Binomial

X = # cars passing through a light in 1 hour. Disjoint time intervals are independent.

Know: $\mathbb{E}[X] = 3$



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

In each interval, there is a car with probability $p = 3/n$ (assume ≤ 1 car can pass by)

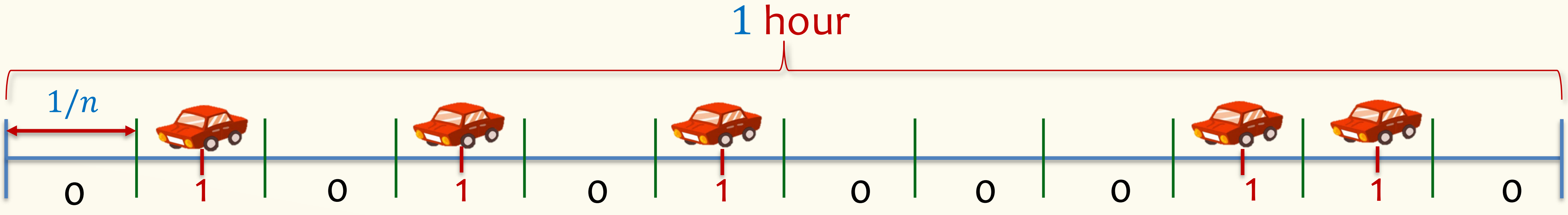
Each interval is Bernoulli: $X_i = 1$ if car in i^{th} interval (0 otherwise). $P(X_i = 1) = 3/n$

$X = \sum_{i=1}^{n} X_i$

More generally

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)

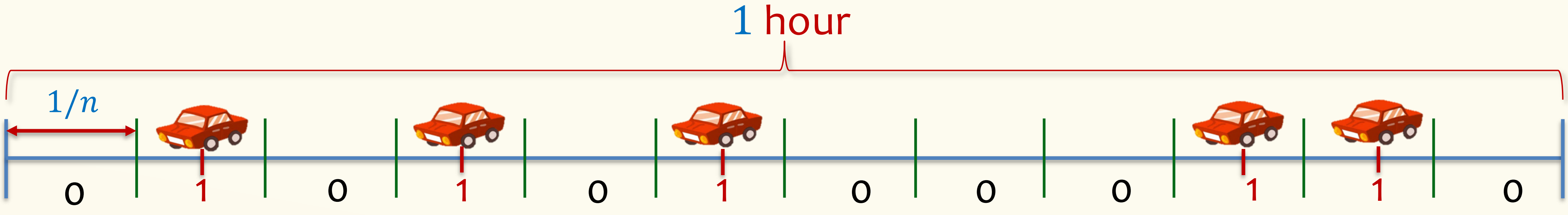
Each interval is Bernoulli: $X_i = 1$ if car in i^{th} interval (0 otherwise). $P(X_i = 1) = \lambda/n$

$X = \sum_{i=1}^{n} X_i$

Model as Binomial

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)

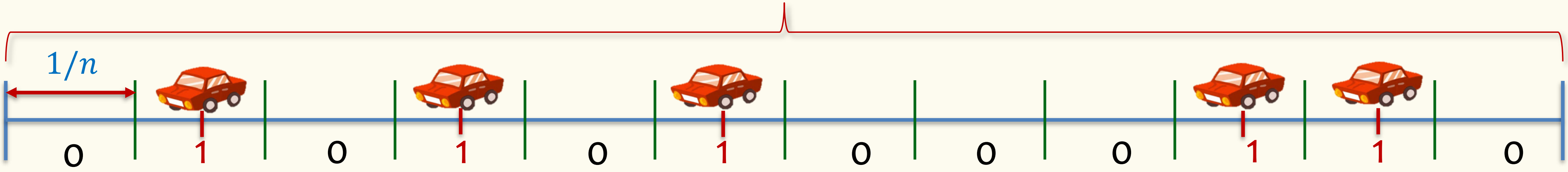
Each interval is Bernoulli: $X_i = 1$ if car in i^{th} interval (0 otherwise). $P(X_i = 1) = \lambda\,/n$

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

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

Don't like discretization

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

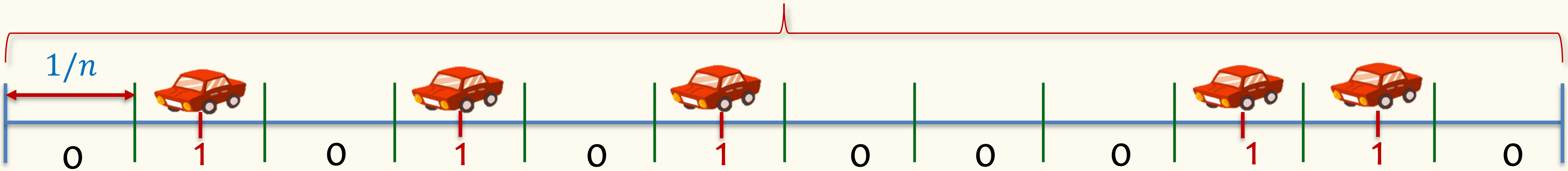


We want now $n \to \infty$

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

Take the limit

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



We want now $n \to \infty$

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

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

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

$$P(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!} \qquad i = 0, 1, 2, \ldots$$

Some examples of "Poisson processes":

- # of cars passing through a traffic light in 1 hour
- # of requests to web servers in an hour
- # of photons hitting a light detector in a given interval
- # of patients arriving to ER within an hour

Assume fixed average rate

Poisson Distribution

Siméon Denis Poisson
1781-1840

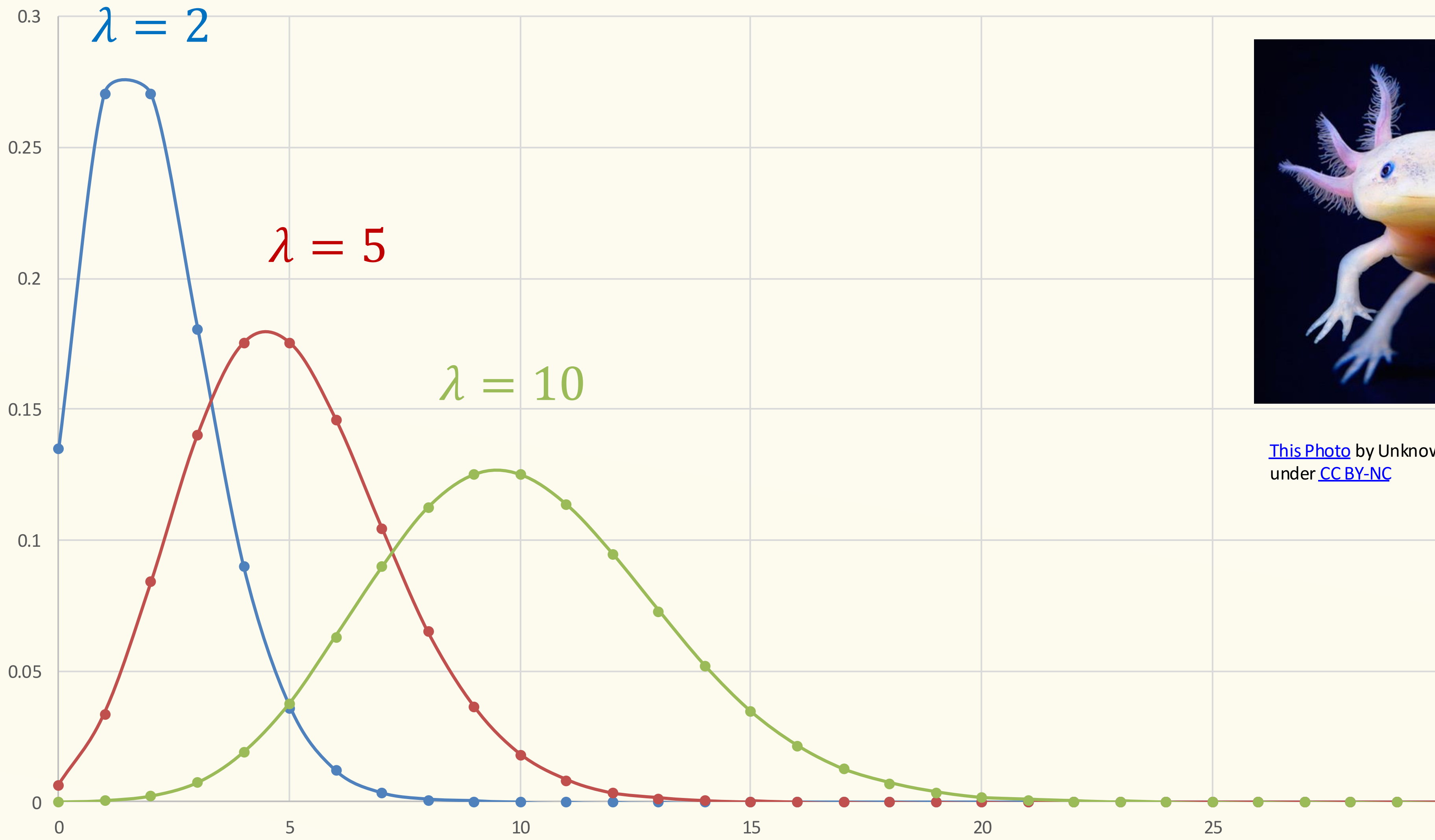
- Suppose "events" happen, independently, at an *average* rate of λ per unit time.
- Let X be the *actual* number of events happening in a given time unit. Then X is a *Poisson* r.v. *with parameter* λ (denoted $X \sim \mathrm{Poi}(\lambda)$) and has distribution (PMF):

$$P(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!} \qquad i = 0, 1, 2, \ldots$$

$$E(X) = \lambda$$
$$\mathrm{Var}(X) = \lambda$$

Probability Mass Function

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



Validity of Distribution

$$P(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!} \qquad i = 0, 1, 2, \ldots$$

Is this a valid probability mass function?
(How do you show that a pmf is valid?)

Calculate sum

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

To show that a pmf is valid, need to check that it takes nonnegative values and that the probabilities sum up to 1.

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

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 $\lambda \geq 0$, then

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

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

Expectation calculation

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

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

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

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

Expectation calculation

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

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

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

Proof.

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

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

= 1 (see prior slides!)

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

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

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

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

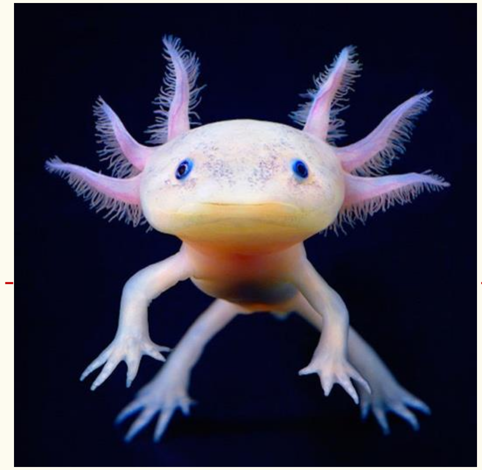
Similar to the previous proof
Verify offline.

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

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 $p < 0.1$)

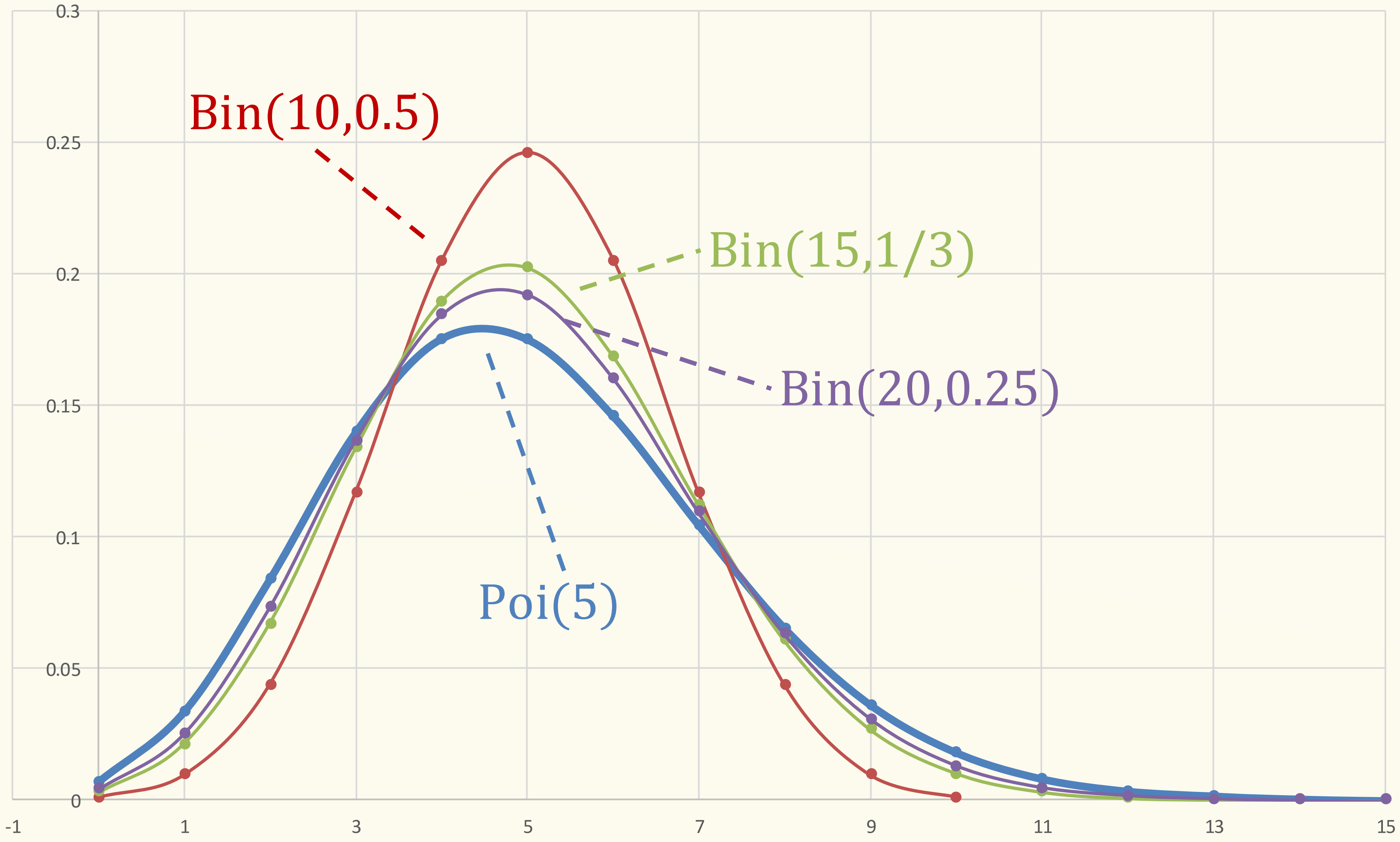
Formally, Binomial approaches Poisson in the limit as $n \to \infty$ (equivalently, $p \to 0$) while holding $np = \lambda$

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

What is probability that message arrives uncorrupted?

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

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

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

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

Sum of Independent Poisson RVs

Let $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$ such that $\lambda = \lambda_1 + \lambda_2$.

Let $Z = X + Y$. What kind of random variable is Z ?

Aka what is the "distribution" of Z ?

Intuition first:

- X is measuring number of (type 1) events that happen in, say, an hour if they happen at an average rate of λ_1 per hour.
- Y is measuring number of (type 2) events that happen in, say, an hour if they happen at an average rate of λ_2 per hour.
- Z is measuring total number of events of both types that happen in, say, an hour, if type 1 and type 2 events occur independently.

Sum of Independent Poisson RVs (2)

Theorem. Let $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$ be independent r.v.s 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 $X_1 \sim \text{Poi}(\lambda_1), \cdots, X_n \sim \text{Poi}(\lambda_n)$ independent 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 (3)

Theorem. Let $X \sim \mathrm{Poi}(\lambda_1)$ and $Y \sim \mathrm{Poi}(\lambda_2)$ 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. $P(Z = z) = \Sigma_{j=0}^{z} P(Y = z - j | X = j)$

https://pollev.com/annakarlin185

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

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

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

Proof

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

Proof

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

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

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

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

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

Binomial Theorem

Poisson Random Variables - Summary

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:

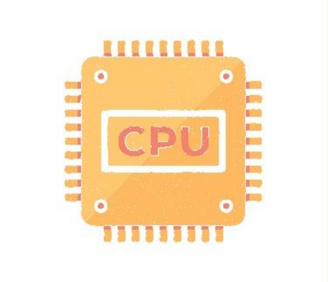
- Events happen at an average rate of λ per time unit
- Number of events happening at a time unit X is distributed according to Poi(λ)

- Poisson approximates Binomial when n is large, p is small, and np is moderate
- Sum of independent Poisson is still a Poisson

Who Fails First?

You have a disk with probability p_1 of failing each day, and a CPU which independently has probability p_2 of failing each day.

$\boldsymbol{p_1}$



$\boldsymbol{p_2}$

Q: What is the probability that the disk fails *before* the CPU?

Law of Total Probability

Law of Total Probability

<u>Theorem</u>: **[Law of Total Probability for Discrete R.V.s]**
Let E be an event. Let Y be a discrete r.v.

$$\boldsymbol{P}\{E\} = \sum_{y} \boldsymbol{P}\{E \cap Y = y\} = \sum_{y} \boldsymbol{P}\{E \mid Y = y\} \cdot \boldsymbol{P}\{Y = y\}$$

For a discrete r.v. X :

$$\boldsymbol{P}\{X = k\} = \sum_{y} \boldsymbol{P}\{X = k \cap Y = y\} = \sum_{y} \boldsymbol{P}\{X = k \mid Y = y\} \cdot \boldsymbol{P}\{Y = y\}$$

Proof: Follows immediately from Law of Total Probability for Events, if we realize that $Y = y$ represents an event and the set of events $Y = y$ over all y form a partition.

Who Fails First?

Disk with prob. p_1 of failing each day, and a CPU with indpt. prob. p_2 of failing each day.

Q: What is the probability that the disk fails *before* the CPU? (Redo using conditioning!)

X_1 = days until disk fails $\sim Geometric(p_1)$

X_2 = days until CPU fails $\sim Geometric(p_2)$

$X_1 \perp X_2$

$\boldsymbol{P}\{X_1 < X_2\} =$

Who Fails First?

Disk with prob. p_1 of failing each day, and a CPU with indpt. prob. p_2 of failing each day.

Q: What is the probability that the disk fails *before* the CPU? (Redo using conditioning!)

X_1 = days until disk fails $\sim Geometric(p_1)$ X_2 = days until CPU fails $\sim Geometric(p_2)$

$X_1 \perp X_2$

$$\boldsymbol{P}\{X_1 < X_2\} = \sum_{k=1}^{\infty} \boldsymbol{P}\{X_1 < X_2 \mid X_1 = k\} \cdot \boldsymbol{P}\{X_1 = k\}$$

$$= \sum_{k=1}^{\infty} \boldsymbol{P}\{k < X_2 \mid X_1 = k\} \cdot \boldsymbol{P}\{X_1 = k\}$$

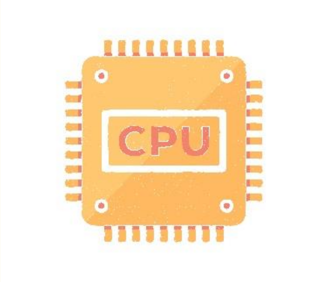
$$= \sum_{k=1}^{\infty} \boldsymbol{P}\{X_2 > k\} \cdot \boldsymbol{P}\{X_1 = k\}$$

$$= \sum_{k=1}^{\infty} (1 - p_2)^k \cdot (1 - p_1)^{k-1} \cdot p_1 = \frac{p_1(1 - p_2)}{1 - (1 - p_2)(1 - p_1)}$$

Who Fails First?

You have a disk with probability p_1 of failing each day, and a CPU which independently has probability p_2 of failing each day.

$\boldsymbol{p_1}$



$\boldsymbol{p_2}$

$$\boldsymbol{P}\{\text{disk fails before CPU fails}\} = \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}$$

But WHY?

Who Fails First?

You have a disk with probability p_1 of failing each day, and a CPU which independently has probability p_2 of failing each day.

$$\boldsymbol{P}\{\text{disk fails before CPU fails}\} = \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}$$

Intuition: Think about flipping 2 coins each day.
There may be many days where both coins are heads.
We only care about the *first day where the coins are not both heads*.

Given that both coins are not heads, what's the probability that coin 1 is H and coin 2 is T?

$\boldsymbol{P}\{\text{coin 1 is H \& coin 2 is T} \mid \text{not both tails}\} =$

Who Fails First?

You have a disk with probability p_1 of failing each day, and a CPU which independently has probability p_2 of failing each day.

$$\boldsymbol{P}\{\text{disk fails before CPU fails}\} = \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}$$

Intuition: Think about flipping 2 coins each day.
There may be many days where both coins are heads.
We only care about the *first day where the coins are not both heads*.

Given that both coins are not heads, what's the probability that coin 1 is H and coin 2 is T?

$$\boldsymbol{P}\{\text{coin 1 is H \& coin 2 is T} \mid \text{not both tails}\} = \frac{\boldsymbol{P}\{\text{coin 1 is H \& coin 2 is T}\}}{\boldsymbol{P}\{\text{not both tails}\}} = \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}$$