Geometric and Poisson Distributions

CS70: Discrete Mathematics and Probability Theory

UC Berkeley – Summer 2026

Lecture 21

Ref: Note 18

Today

Two more distributions with important properties...

Geometric distribution

- Fail until you succeed...
- Memoryless distribution
- Application to coupon collecting

Poisson distribution

- Events happen at a given expected rate
- Limiting distribution of increasingly fine-grained Binomial
- Sum of two Poisson is Poisson

Recall: Distribution – Flipping Until Heads

Experiment: Flip a coin until you get a "heads"

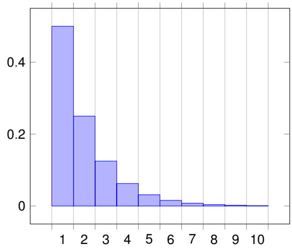
Random variable $X =$ "how many flips were needed to get a heads"

Outcomes:

$\texttt{H}$: $X = 1$
$\texttt{TH}$: $X = 2$
$\texttt{TTH}$: $X = 3$
$\cdots$
$\texttt{TTTTTTH}$: $X = 7$
$\cdots$

Let event $T_i =$ "tails on flip number i"

$\mathbb{P}[X = 1] =$
$\mathbb{P}[X = 2] =$
$\mathbb{P}[X = 3] =$
$\cdots$
$\mathbb{P}[X = k] =$



Continues infinitely to the right...

Common Distributions: Geometric

Experiment: Flip a coin until you get a "heads"

Generalize: Bernoulli trials with success probability p, repeat until success

Random variable $X =$ "how many flips were needed to get a heads"

Outcomes:

$\texttt{H}$: $X = 1$
$\texttt{TH}$: $X = 2$
$\texttt{TTH}$: $X = 3$
$\dots$
$\texttt{TTTTTTH}$: $X = 7$
$\dots$

Distribution of X? $\mathbb{P}[X = 1] = p$ $\mathbb{P}[X = 2] = (1-p)p$ ...

For "$X = i$", have $i - 1$ "failures" followed by one "success"

$\Rightarrow$ Each failure has probability $1 - p$

$\Rightarrow$ Success has probability p

$\Rightarrow$ $\mathbb{P}[X = i] = (1-p)^{i-1}p$

Common Distributions: Geometric

Experiment: Bernoulli trials with success probability p, repeat until success

Random variable $X =$ "number of trials needed"

$\mathbb{P}[X = i] = (1-p)^{i-1}p$

Sanity check: Do probabilities add up to 1?

$$\sum_{i=1}^{\infty}(1-p)^{i-1}p \;=\; p\sum_{i=0}^{\infty}(1-p)^{i} \;=\; p\cdot\frac{1}{p} = 1$$

Uses formula for infinite geometric sum – if you don't remember:

$$\begin{aligned}
S &= (1-p)^0 + (1-p)^1 + (1-p)^2 + \cdots \\
\underline{(1-p)S} &= \underline{\phantom{(1-p)^0 + {}}(1-p)^1 + (1-p)^2 + \cdots} \\
[1-(1-p)]\,S &= (1-p)^0 \qquad\qquad \Longrightarrow \quad pS = 1 \qquad \Longrightarrow \quad S = \frac{1}{p}
\end{aligned}$$

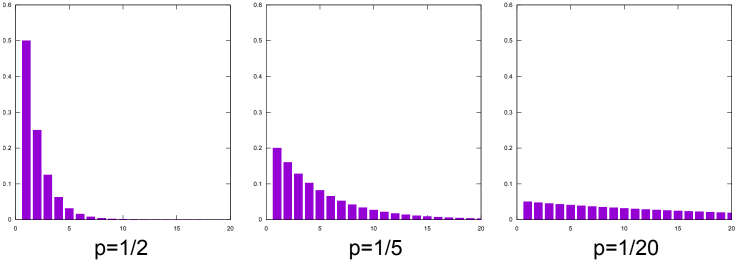
So valid distribution... we say "X has geometric distribution with parameter p"

Notation: $X \sim \text{Geometric}(p)$ or $X \sim \text{Geom}(p)$

Geometric Distribution: Plots

$$X \sim \text{Geom}(p) \qquad \text{so} \qquad \mathbb{P}[X = i] = (1-p)^{i-1} p \quad (i \geq 1)$$

What do distributions "look like"?



Smaller p values (probability of success) lead to flatter distributions.

Interesting application is system operation – "meaning" flipped though

System operates correctly (hopefully many steps) until it fails (once)

Bernoulli "Failure": Operates correctly for one hour (probability $1-p$)

Bernoulli "Success": Fails in this hour (probability p)

Geometric Distribution: Expectation

$$X \sim \text{Geom}(p) \qquad \text{so} \qquad \mathbb{P}[X=i] = (1-p)^{i-1}p \quad (i \geq 1)$$

Expected value?

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} i \cdot \mathbb{P}[X=i] = \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1}p$$

Nobody remembers this formula.... so:

$$\begin{array}{rrcl}
 & \mathbb{E}[X] & = & p + 2(1-p)p + 3(1-p)^2p + 4(1-p)^3p + \cdots \\
- & \underline{(1-p)\mathbb{E}[X]} & = & \underline{\qquad (1-p)p + 2(1-p)^2p + 3(1-p)^3p + \cdots} \\
 & p\mathbb{E}[X] & = & p + \ (1-p)p + \ (1-p)^2p + \ (1-p)^3p + \cdots \\
\Longrightarrow & p\mathbb{E}[X] & = & \sum_{i=1}^{\infty} \mathbb{P}[X=i] = 1.
\end{array}$$

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

With previous system example: Mean Time To Failure (MTTF)

$\Rightarrow$ Common spec for hard drives and other systems

Common Distributions: Variance of Geometric

Experiment: Bernoulli trials with success probability p, repeat until success

Random variable $X =$ "number of trials needed"

$\mathbb{P}[X = i] = (1-p)^{i-1}p$ and $\mathbb{E}[X] = \frac{1}{p}$

$$\begin{array}{rrcl}
& \mathbb{E}[X^2] & = & p + 4p(1-p) + 9p(1-p)^2 + \ldots \\
- & \underline{(1-p)\mathbb{E}[X^2]} & = & \underline{\qquad p(1-p) + 4p(1-p)^2 + \ldots} \\
& p\mathbb{E}[X^2] & = & p + 3p(1-p) + 5p(1-p)^2 + \ldots \\
& & = & 2(p + 2p(1-p) + 3p(1-p)^2 + ..) \quad \mathbb{E}[X]! \\
& & & -(p + p(1-p) + p(1-p)^2 + \ldots) \quad \text{Distribution.} \\
& p \cdot \mathbb{E}[X^2] & = & 2 \cdot \mathbb{E}[X] - 1 \\
& & = & 2 \cdot \frac{1}{p} - 1 = \frac{2-p}{p} \qquad \Longrightarrow \quad \mathbb{E}[X^2] = \frac{2-p}{p^2}
\end{array}$$

$$\Longrightarrow \text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

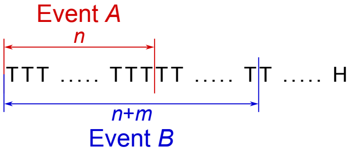
The notes calculate expected value and variance with different techniques...

For expected value: "Tail Sum Formula" *For variance: Calculus trick*

Geometric Distribution: Memoryless Property

"Memoryless": Time-based distribution is the same from any point in time.

$\Rightarrow$ "Gambler's Fallacy": False belief that past losses affect probability



Events:

$A =$ "still going after n tosses" (i.e., "first n tosses are tails")

$B =$ "still going after $n+m$ tosses"

And let $C =$ "first m tosses are tails"

Memoryless means $\mathbb{P}[B \mid A] = \mathbb{P}[C]$ ("m more" same as "m from beginning")

If X is number of tosses to get a heads (success), then

$$\mathbb{P}[X > n+m \mid X > n] = \mathbb{P}[B \mid A] = \mathbb{P}[C] = \mathbb{P}[X > m] .$$

Intuition: Coin has no memory....

Geometric Distribution: Memoryless Proof

Theorem: Let $X \sim \text{Geometric}(p)$. Then, for for all $n, m \geq 0$,

$$\mathbb{P}[X > n+m \,|\, X > n] = Pr[X > m] \ .$$

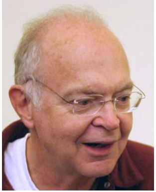
Proof:
First note that $\mathbb{P}[X > k]$ is the probability that the first k trials fail: $(1-p)^k$.

Then

$$\begin{aligned}
\mathbb{P}[X > n+m \,|\, X > n] &= \frac{\mathbb{P}[X > n+m \text{ and } X > n]}{\mathbb{P}[X > n]} \\
&= \frac{\mathbb{P}[X > n+m]}{\mathbb{P}[X > n]} \\
&= \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m \\
&= \mathbb{P}[X > m]. \qquad \square
\end{aligned}$$

Recall the Coupon Collector Problem

Hot trend for 2026: Cards for Turing Award winners in cereal boxes!



Donald Knuth
1974

Richard Karp
1985

Shafi Goldwasser
2012

There are n different cards (unlimited copies), each box has a random person

With random variables, define $X =$ "number of boxes before getting all cards"

Question: What is $\mathbb{P}[X > m]$?

Previous lecture: If $m = n \ln(2n) \approx n(\ln n + 0.693)$, then $\mathbb{P}[X > m] \le \frac{1}{2}$.

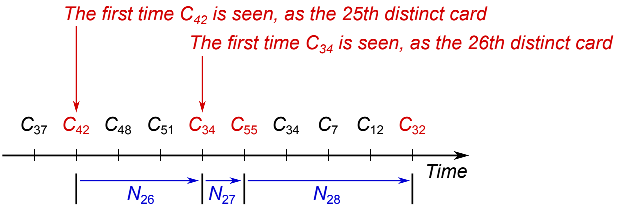
$\Rightarrow$ This is like a median – what about the *mean*? .. or $\mathbb{E}[X]$?

Coupon Collecting: Using Geometric r.v.s

$X =$ "number of boxes bought before getting all cards"

$\Rightarrow$ What is $\mathbb{E}[X]$?

Break this down into time-spans for "receiving a card we haven't seen before"



$N_i =$ "number of boxes after getting $i-1$ different cards to get ith distinct"

$\Rightarrow$ $X = N_1 + N_2 + \cdots + N_n$

N_1 : Always on first step, so $\mathbb{P}[N_1 = 1] = 1$

N_2 : Get first card again with probability $\frac{1}{n}$, new (success!) with probability $\frac{n-1}{n}$

$\Rightarrow$ $N_2 \sim \text{Geometric}\left(\frac{n-1}{n}\right)$

For *ith* card, success with any of the $n-(i-1)$ unseen cards

$\Rightarrow$ $N_i \sim \text{Geometric}(\frac{n-(i-1)}{n})$

Coupon Collecting

$X =$ "number of boxes bought before getting all cards"

$\Rightarrow$ What is $\mathbb{E}[X]$?

Break this down into time-spans for "receiving a card we haven't seen before"

$N_i =$ "number of boxes after getting $i-1$ different cards to get ith distinct"

$\Rightarrow$ $X = N_1 + N_2 + \cdots + N_n$

$\Rightarrow$ $N_i \sim \text{Geometric}(\frac{n-(i-1)}{n})$

From Geometric distribution: $\mathbb{E}[N_i] = \frac{n}{n-i+1}$

$$\mathbb{E}[X] = \sum_{i=1}^{n} \frac{n}{n-i+1} = \sum_{i=1}^{n} \frac{n}{i} = n \sum_{i=1}^{n} \frac{1}{i}$$

$\sum_{i=1}^{n} \frac{1}{i}$ is called the "Harmonic Sum" with notation H_n

$\Rightarrow$ Known: $H_n \approx \ln n + \gamma_E$ where $\gamma_E \approx 0.5772$ is "Euler's constant"

So $\mathbb{E}[X] = nH_n \approx n(\ln n + \gamma_E)$

Compare to previous result: If $m \approx n(\ln n + 0.693)$, then $\mathbb{P}[X > m] \leq \frac{1}{2}$.

Common Distributions: Poisson

Motivation and Derivation

You have been hired as a high-priced consultant by fast-food chain "Twelve Guys" to model customer demand at their new outpost on Zemlox 3.

Question 1: How many customers arrive in an hour?

You know: Average is λ.

Arrivals are independent (assumption)
Rate is constant

What is the distribution?

For staffing in surges, would like to know...

Question 2: $\mathbb{P}[2\lambda \text{ arrivals }]$?

Approach: Cut hour into n intervals of length $1/n$.
$\mathbb{P}[\text{one arrival}] = \lambda/n$ ... $\mathbb{P}[\text{ two arrivals }] = (\lambda/n)^2$ (small when n is large!)
"Arrival" is modeled as Bernoulli trial – count is modeled as Binomial

Common Distributions: Poisson

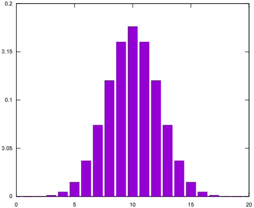
Properties:

$\Rightarrow$ Number of Arrivals with known expected value λ (fixed rate)

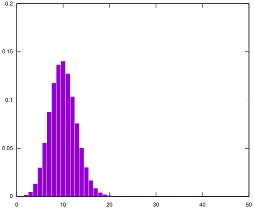
$\Rightarrow$ Arrivals independent and "happens or not" with certain probability

First thought: Binomial? n timeslots $\rightarrow$ Bernoulli success $= \lambda/n$

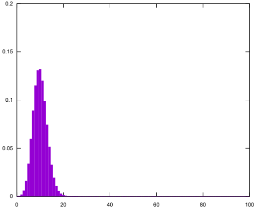
Example: Mean $\lambda = 10$ – Out of 20 time slots? – Out of 50? 100?



$\lambda = 10, \; n = 20$
$X \sim \text{Bin}(20, \frac{1}{2})$
$\mathbb{P}[X = k] = 0$ for $k > 20$



$\lambda = 10, \; n = 50$
$X \sim \text{Bin}(50, \frac{1}{5})$
$\mathbb{P}[X = k] = 0$ for $k > 50$



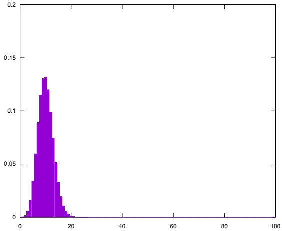
$\lambda = 10, \; n = 100$
$X \sim \text{Bin}(100, \frac{1}{10})$
$\mathbb{P}[X = k] = 0$ for $k > 100$

No reason for fixed upper limit – seems to keep shape, extended to the right

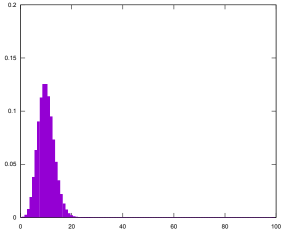
$\Rightarrow$ Is there a valid distribution as $n \rightarrow \infty$? Yes! Poisson distribution

Common Distributions: Poisson

Pictures first... n gets big – really big – in the limit:



$\lambda = 10,\ n = 100$
$X \sim \text{Bin}(100, \frac{1}{10})$
$\mathbb{P}[X = k] = 0$ for $k > 100$



$\lambda = 10,\ n$ "in the limit"
$X \sim \text{Poisson}(10)$
$\mathbb{P}[X = k] > 0$ for all $k \geq 0$

Poisson on the right – non-zero probabilities for $X = k$ for any $k \in \mathbb{N}$

Some sample values with $X \sim \text{Poisson}(10)$:

$$\left.\begin{array}{l} \mathbb{P}[X = 8] \approx 0.1126 \\ \mathbb{P}[X = 10] \approx 0.1251 \\ \mathbb{P}[X = 20] \approx 0.00187 \\ \mathbb{P}[X = 1000] \approx 1.128 \times 10^{-1572} \end{array}\right\} \quad \text{Spoiler: } \mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda}$$

Common Distributions: Poisson – Derivation

Set-up: Random variables $X_n \sim \text{Binomial}(n, \lambda/n)$

Problem: Find $\lim_{n\to\infty} \mathbb{P}[X_n = k]$ for some k

$\Rightarrow$ k is constant – independent of n

$$\begin{aligned}
\mathbb{P}[X_n = k] &= \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{with } p = \lambda/n \\
&= \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k}
\end{aligned}$$

$$\lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} = 1$$

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \lim_{n\to\infty} \left(\left(1 - \frac{\lambda}{n}\right)^{(n-k)/\lambda}\right)^{\lambda} = (e^{-1})^{\lambda} = e^{-\lambda}$$

$$\Longrightarrow \quad \lim_{n\to\infty} \mathbb{P}[X_n = k] = \frac{\lambda^k}{k!} e^{-\lambda}$$

Poisson Distribution: Sanity Check!

$$X \sim \text{Poisson}(\lambda) \qquad \Longrightarrow \qquad \mathbb{P}[X=k] = \frac{\lambda^k}{k!}e^{-\lambda}$$

Sanity check: Does $\sum_{k=0}^{\infty} \mathbb{P}[X=k] = 1$?

$$\sum_{k=0}^{\infty} \mathbb{P}[X=k] = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

The Taylor series expansion of $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$

$$\Longrightarrow \qquad \sum_{k=0}^{\infty} \mathbb{P}[X=k] = e^{-\lambda}\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1 \qquad \checkmark$$

So this is a valid probability distribution.

Poisson Distribution: Expectation

$$X \sim \text{Poisson}(\lambda) \quad \Longrightarrow \quad \mathbb{P}[X = k] = \tfrac{\lambda^k}{k!} e^{-\lambda}$$

What is $\mathbb{E}[X]$?

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\
&= \lambda e^{-\lambda} e^{\lambda} = \lambda
\end{aligned}$$

Recap: We started wanting a distribution with mean λ

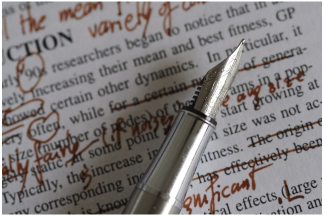
$\Rightarrow$ This has mean λ – excellent!

Poisson Distribution: Example

A proofreader found that writers make errors at an average rate of 2 per page.

Average of 2 per page – or 1 per half page – or 4 per two pages...
 $\Rightarrow$ Constant *rate*, regardless of length
 $\Rightarrow$ Poisson!

$X =$"number of errors on this page"
 $\Rightarrow$ $X \sim \text{Poisson}(2)$



Let's get a feel for this with some values...

$\mathbb{E}[X] = 2$ Of course!

$\mathbb{P}[X = 2] \approx 0.271$ Chance of exactly 2 errors

$\mathbb{P}[X \leq 2] \approx 0.6766$ About 2/3 probability of at *most* 2 errors

$\mathbb{P}[X = 0] \approx 0.135$ Probability of no errors

$\mathbb{P}[X = 4] \approx 0.090$ Less chance of double errors than none

$\mathbb{P}[X \geq 4] \approx 0.143$ But higher chance of two *or more* errors

Common Distributions: Poisson Variance

Let $X \sim \text{Poisson}(\lambda)$, so for $m \geq 0$

$$\mathbb{P}[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}.$$

Variance? Use the formula? Ugh. Done in the notes...

But since we derived Poisson from Binomial first...
$\Rightarrow$ Poisson comes from the limit as $n \to \infty$ of $X_n \sim \text{Binomial}(n, \frac{\lambda}{n})$.

For any n, we have $\mathbb{E}[X_n] = pn = \frac{\lambda}{n} \cdot n = \lambda$.

And also: $\text{Var}(X_n) = p(1-p)n = \frac{\lambda}{n} \cdot \left(1 - \frac{\lambda}{n}\right) n = \lambda \left(1 - \frac{\lambda}{n}\right)$.
$\Rightarrow \lim_{n \to \infty} \text{Var}(X_n) = \lambda$

So for Poisson:

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

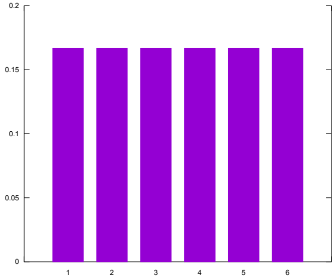
$\text{Var}(X) = \lambda$

Sums of Random Variables: Distribution

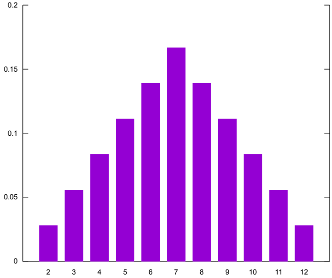
Question: If X and Y have same distribution, does $X + Y$?

To explore: Roll two 6-sided dice, giving two values (RVs) D_1 and D_2

Distribution of D_1 and D_2 is Uniform(6)



Distribution of $D_1 + D_2$



Is $D_1 + D_2$ uniform?

But Poisson is different!

The Sum of Two Poisson r.v.s is Poisson

Theorem: If X and Y are independent random variables with $X \sim \text{Poisson}(\lambda_1)$ and $Y \sim \text{Poisson}(\lambda_2)$, then $Z = X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

Proof: The distributions for X and Y give

$$\mathbb{P}[X = m] = \frac{\lambda_1^m}{m!} e^{-\lambda_1} \quad \text{and} \quad Pr[Y = m] = \frac{\lambda_2^m}{m!} e^{-\lambda_2} \ .$$

Then with $Z = X + Y$, and since X and Y are independent,

$$\mathbb{P}[Z = m] = \sum_{i=0}^{m} \mathbb{P}[X = i, Y = m - i] = \sum_{i=0}^{m} \frac{\lambda_1^i}{i!} e^{-\lambda_1} \frac{\lambda_2^{m-i}}{(m-i)!} e^{-\lambda_2} \ .$$

Rearranging terms and using the binomial theorem,

$$\mathbb{P}[Z = m] = e^{-(\lambda_1+\lambda_2)} \sum_{i=0}^{m} \frac{1}{i!(m-i)!} \lambda_1^i \lambda_2^{m-i} = e^{-(\lambda_1+\lambda_2)} \frac{1}{m!} \sum_{i=0}^{m} \frac{m!}{i!(m-i)!} \lambda_1^i \lambda_2^{m-i}$$

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

So $Z \sim \text{Poisson}(\lambda_1 + \lambda_2)$. $\square$

Summary

Added two more common distributions to our toolbox!

Geometric distribution: $X \sim \text{Geometric}(p)$

- $\mathbb{P}[X = k] = (1-p)^{k-1}p \quad (\text{for } i \geq 1) \quad ; \quad \mathbb{E}[X] = \frac{1}{p} \quad ; \quad \text{Var}(X) = \frac{1-p}{p^2}$
- Memoryless distribution
- Used to analyze expected number of steps in coupon collecting

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

- $\mathbb{P}[X = k] = \frac{k}{k!}e^{-\lambda} \quad (\text{for } i \geq 0) \quad ; \quad \mathbb{E}[X] = \lambda \quad ; \quad \text{Var}(X) = \lambda$
- Events happen at a given expected rate
- Limiting distribution of increasingly fine-grained Binomial
- Sum of two Poisson is Poisson