			μ	σ^2
Bernoulli	Bern(p)	$P(X=1)=p$, $P(X=0)=1-p$	p	$p(1-p)$
	Bin(n,p)	$P(X=k)=C_n^k p^k (1-p)^{n-k}$	np	$np(1-p)$
	Geo(p)	$P(X=k)=(1-p)^{k-1}\cdot p$, $k \geq 1$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
	Poi(λ)	$Bin(n,p) \xrightarrow{n\to\infty}$ $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$	λ	λ

$\mathbb{E} = V_{ar}$

$\left(1+\frac{1}{n}\right)^n \to e$ 2.71

e

$U(a,b)$ $f(x)=\frac{1}{b-a}$

$\frac{a+b}{2}$ $\frac{(b-a)^2}{12}$

exp
neb.

$Exp(\lambda)$, $f(x)=\lambda e^{-\lambda x}$, $x \geq 0$, 0 else

$\frac{1}{\lambda}$, $\frac{1}{\lambda^2}$

3B1B

$N(\mu, \sigma^2)$

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = 1$$

🧀 🧀 01 Distribution detective: Which one fits?

Match each scenario to the most appropriate distribution. Justify each choice in one sentence.

- a. The number of typos on a randomly selected page of a 500-page book, if typos occur randomly at an average rate of 0.5 per page.
- b. Whether a randomly selected email is spam (yes/no), given 40% of emails are spam.
- c. The number of heads in 20 coin flips.
- d. The exact time (in minutes) you wait for the next bus, if buses arrive completely randomly at an average rate of 4 per hour.
- e. A randomly chosen real number between 0 and 10.

Ber Bin(1, p)

Binomial

[0, 10]

$P(X=30) =$

= 0.5

100

$n = 50.000$

$p = 0.001$

30 40 50

🧀 🧀 02 Name that distribution

For each scenario, identify the distribution, state its parameter(s), and write the PMF or PDF.

- a. A call center receives calls at an average rate of 8 per hour. Let X be the number of calls received between 2:00 PM and 3:00 PM.
- b. A software update crashes with probability 0.03. An IT department pushes the update to 200 computers independently. Let Y be the number of computers that crash.
- c. A sensor measures temperature continuously, but due to manufacturing imprecision, the true reading is somewhere between 98.5°C and 101.5°C with no value more likely than another. Let T be the measured temperature.
- d. A quality inspector tests light bulbs one by one. Each bulb independently fails inspection with probability 0.15. Let N be the number of bulbs tested until the first failure.
- e. The time between earthquakes in a seismically active region averages 4 months. Let W be the waiting time (in months) until the next earthquake.

🧀 🧀 🧀 03 Mystery distributions: Identify from data

A researcher collects data from three different experiments and computes summary statistics:

Dataset A: $n = 500$ observations, all values are either 0 or 1. Sample mean ≈ 0.23, sample variance ≈ 0.177.

$\mathrm{Bin}(1,p)$

Dataset B: $n = 1000$ observations, values range from 0 to 47. Sample mean ≈ 12.1, sample variance ≈ 11.8.

$E(x) = p = 0.23$

Dataset C: $n = 800$ observations, values are positive reals ranging from 0.001 to 14.2. Sample mean ≈ 2.5, sample variance ≈ 6.3.

E $\quad$ E $\quad$ Var $\quad$ $\mathrm{Var}\,X = (1-p)\cdot p$

For each dataset:

- a. Identify the most likely distribution family.
- b. Estimate the parameter(s) of that distribution from the summary statistics.
- c. For Dataset B, the researcher notices that these are counts of customer complaints per day at a call center. Does this context support your answer? What if instead they were counts of “successes” in 50 independent trials per observation?

12.1

04 The "obvious" Bernoulli that isn't

A weighted die shows 6 with probability $\frac{1}{3}$ and each of 1–5 with probability $\frac{2}{15}$.

- a. Define a Bernoulli random variable X for "rolling a 6." State p and compute $E[X]$ and $\text{Var}[X]$.
- b. Define a different Bernoulli random variable Y for "rolling an even number." Compute $E[Y]$ and $\text{Var}[Y]$.
- c. For which event is the variance larger? Explain intuitively why maximum Bernoulli variance occurs at $p = 0.5$.

05 Memoryless waiting: ~~Geometric~~ intuition

A slot machine pays out with probability $p = 0.05$ on each play.

- a. What is the expected number of plays until the first payout?
- b. You've already played 50 times with no payout. What is the expected *additional* number of plays until you win?
- c. A gambler says: "I'm due for a win soon because I've lost so many times." In 2–3 sentences, explain why this reasoning is flawed.

$P(x=k)=(1-p)^{k-1}\cdot p$

$0{,}05$

$\sum_{k=}^{\infty} kP(x=k)$

$k(1-p)^{k-1}\cdot p$

$E[X]=\frac{1}{p}=20$

$A \to X$, $B \to Y$

$m=50$ $\quad n$

$P(A|B)=\frac{A\cap B}{P(B)}$

$P(x>m+n \mid x>m) = P(x>n) = \frac{P(x>m+n)}{P(x>m)} = \frac{0.95^{m+n}}{0.95^{m}} = 0.95^{n}$

06 Binomial: Quality control decision

A factory produces chips with defect probability $p = 0.02$. A batch of $n = 100$ chips is inspected.

- a. Let X be the number of defective chips. State the distribution of X and compute $E[X]$ and $\mathrm{Var}[X]$.
- b. The batch is rejected if more than 5 chips are defective. Without computing $P[X > 5]$ exactly, explain why $P[X > 5]$ is small.
- c. If p increases to 0.10, recompute $E[X]$. How does this change the rejection decision intuitively?

$\cancel{(1-p)(1-y)(1-p)(1-y)\cdot p}$

$$\frac{P(X=3 \mid X \cancel{-1} \; 2)}{P(\cancel{X-1})}$$

p

$\cancel{1,0}$

1

07 Poisson: Rare events approximation

A website has 10,000 visitors per day. Each visitor independently has a 0.0003 probability of reporting a bug.

- a. Let X be the number of bug reports per day. Which distribution is a good approximation here, and what is the parameter?
- b. What is the probability of receiving at least one bug report?

🧀 🧀 08 Exponential: Memoryless lifetimes

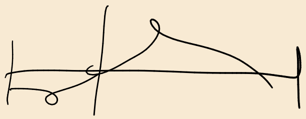
A light bulb's lifetime (in years) follows $\mathrm{Exp}(\lambda = 0.5)$.

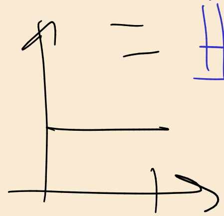
- a. Compute $E[X]$ and the probability that the bulb lasts more than 3 years.
- b. Given that the bulb has already lasted 2 years, what is the probability it lasts at least 1 more year?
- c. Compare with the discrete case: if bulb failure each year is Bernoulli with $p = 0.4$, and $Y \sim \mathrm{Geo}(0.4)$ counts years until failure, compute $P[Y > 3 \mid Y > 2]$ and $P[Y > 1]$. What do you notice?

🧀 🧀 09 Uniform: The broken stick problem

A stick of length 1 is broken at a uniformly random point $X \sim U(0, 1)$.

- a. What is the expected length of the left piece?
- b. Let $Y = X(1 - X)$ be the product of the two piece lengths. Compute $E[Y]$.
- c. What break point x maximizes $Y = x(1 - x)$? Compare this to $E[X]$.





$$\mathbb{E}[X(1-X)] = \mathbb{E}[X - X^2] = \mathbb{E}[X] - \mathbb{E}[X^2]$$

$$\mathbb{E}[X] = \frac{0+1}{2}$$

$$= \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x \frac{1}{1-0}\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2} - 0 = \frac{1}{2}$$

$$\int_0^1 x^2 \cdot 1\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$$

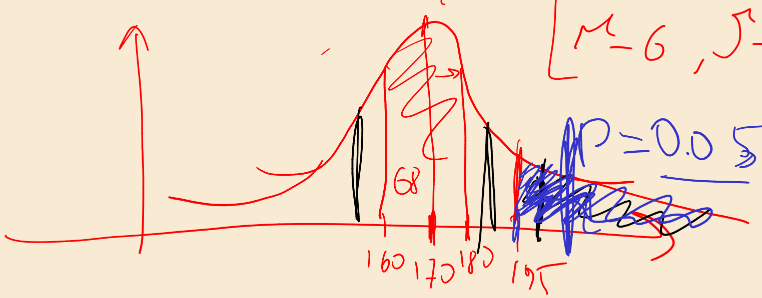
$$\frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

$$\int_0^1 x(1-x)\,1 = \left[\frac{x^2}{2} - \frac{x^3}{3}\right]\Big|_0^1$$

🧀 10 Normal: The 68-95-99.7 rule in action

Human heights in a population follow $N(170, 100)$ (mean 170 cm, variance 100 cm²).

- a. What is σ? What proportion of people are between 160 cm and 180 cm tall?
- b. A person is 2.5 standard deviations above the mean. How tall are they?
- c. Standardize the height $X = 155$ cm. Interpret the z-score: is this person unusually short?



$[\mu-\sigma, \mu+\sigma]$

$P = 0.05$

$F(180) - F(160)$

$\frac{X-\mu}{\sigma}$

150 - 190 ~ 95%

max $X(1-\cancel{X})$

X $\quad 1 \cdot 0 = 0$

$1-X$ $\quad 0.8$

$\mathbb{E}(X(1-X)) \stackrel{?}{=} \frac{1}{4}$

$= \mathbb{E}[X](1-\mathbb{E}(X))$

$\frac{1}{2} \cdot \frac{1}{2}$

🧀 🧀 11 Normal: Standardization and comparison

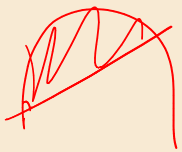
Test A has scores $\sim N(500, 10000)$ (so $\sigma = 100$). Test B has scores $\sim N(50, 100)$ (so $\sigma = 10$).

- a. Alice scores 680 on Test A. Bob scores 72 on Test B. Compute both z-scores.
- b. Who performed better relative to their test population?
- c. Explain why comparing raw scores (680 vs 72) is meaningless without standardization.
- d. You're given a coin that shows heads with unknown probability p. You flip it 100 times and observe 65 heads. If the coin were fair ($p = 0.5$), what are $E[X]$ and $\mathrm{SD}[X]$ for the number of heads? How many standard deviations away from the mean is 65? What can you conclude about whether $p = 0.5$?

$g(x) = (X)(1-X)$

$\mathbb{E}(g(x)) \underset{\le}{\neq} g(\mathbb{E}[X])$

jenson



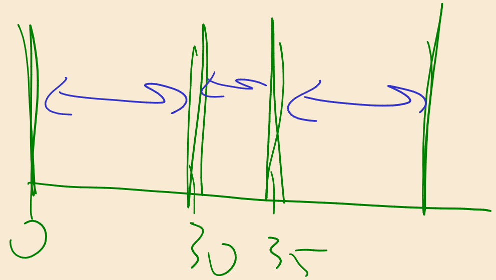
y

$\mathbb{E}[g(x)] \ge g(\mathbb{E}(X))$

🧀 🧀 12 The Poisson-Exponential connection

Customers arrive at a shop according to a Poisson process with rate $\lambda = 4$ per hour.

- a. What distribution does the number of arrivals in 1 hour follow? State its mean and variance.
- b. What distribution does the time between consecutive arrivals follow? State its mean.
- c. If no customer has arrived in the last 15 minutes, what is the probability that the next customer arrives within 10 minutes?



$\text{Pos}(\lambda t)$

$0-1$

$1-2$

$[0,1]$

$\frac{1}{\infty} = 0$

expone

$$P(T > 15+10 \mid T > 15) = \frac{P(T > 15+10)}{P(T > 15)}$$

$$= P(T > 10)$$

$P(T > 10)$

🧀 🧀 🧀 13 The “inspection paradox”

Buses arrive according to a Poisson process with rate $\lambda = 6$ per hour (i.e., one every 10 minutes on average). You arrive at the bus stop at a uniformly random time.

- a. What is the distribution of time between consecutive buses? Compute its expected value.
- b. Intuitively, would you expect your average wait time to be 5 minutes (half the inter-arrival time)?
- c. The “inspection paradox” says you’re more likely to arrive during a *long* gap than a short one. Without computing, explain in 2–3 sentences why your expected wait might actually be *longer*

PDF $\lambda e^{-\lambda t}$

$T \sim$

CDF $\frac{\lambda^k}{k!} e^{-\lambda}$ pmf

$N(t) = \# \; t$

$N(t) \sim Pois(\lambda t)$

PDF = (CDF)'

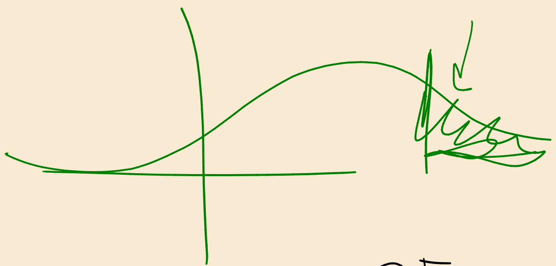
$\{T > t\} = \{$no event $t) =$

$= \{N(t) = 0\}$

$1 + e^{-\lambda t} = +\lambda e^{-\lambda t}$

$P(T > t) = P(N(t) = 0)$

$\frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$



$P(T < t) = 1 - P(T > t) =$

CDF Fun

$= 1 - e^{-\lambda t}$

🧀 🧀 🧀 14 The prosecutor’s fallacy: Conditional thinking

In a city of 1 million people, a crime is committed. DNA evidence matches the suspect with a 1-in-10,000 error rate (i.e., a random person matches with probability 0.0001).

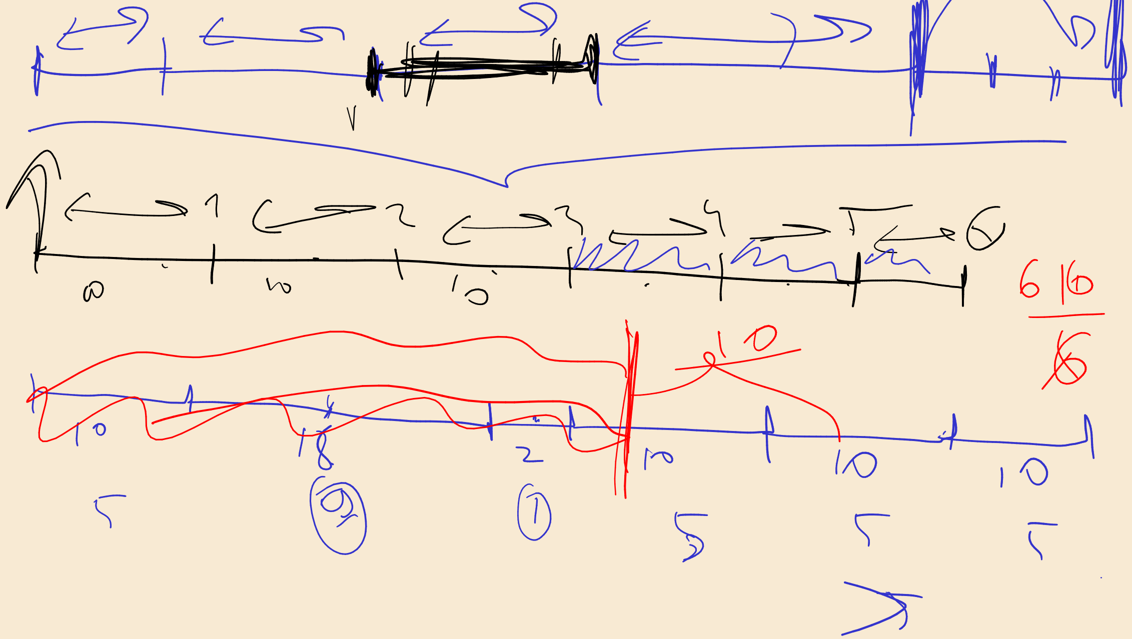
- a. Model the number of matching individuals in the city as a random variable. What distribution is appropriate? What is its expected value?
- b. The prosecutor argues: “The probability of a false match is 0.0001, so the defendant is 99.99% certain to be guilty.” Is this reasoning correct?
- c. If we assume the guilty person is definitely in the city, use Bayes-like reasoning to argue that the suspect’s probability of guilt depends on the expected number of matches.

13 The "inspection paradox"

Buses arrive according to a Poisson process with rate $\lambda = 6$ per hour (i.e., one every 10 minutes on average). You arrive at the bus stop at a uniformly random time.

- a. What is the distribution of time between consecutive buses? Compute its expected value.
- b. Intuitively, would you expect your average wait time to be 5 minutes (half the inter-arrival time)?
- c. The "inspection paradox" says you're more likely to arrive during a *long* gap than a short one. Without computing, explain in 2–3 sentences why your expected wait might actually be *longer* than 5 minutes.

Exp(6) $\quad \frac{1}{6}$



$\frac{6 \cdot 10}{6} = \boxed{10}$

🧀 🧀 🧀 14 The prosecutor's fallacy: Conditional thinking

In a city of 1 million people, a crime is committed. DNA evidence matches the suspect with a 1-in-10,000 error rate (i.e., a random person matches with probability 0.0001).

- a. Model the number of matching individuals in the city as a random variable. What distribution is appropriate? What is its expected value?
- b. The prosecutor argues: "The probability of a false match is 0.0001, so the defendant is 99.99% certain to be guilty." Is this reasoning correct?
- c. If we assume the guilty person is definitely in the city, use Bayes-like reasoning to argue that the suspect's probability of guilt depends on the expected number of matches.

DNA

$\frac{1}{10.000}$

$P(DNA | \ldots) = Bin.(n = 999.999, p = 0.0001)$

$1 - 0.0001$

0.999

$X \sim P(\lambda = n \cdot p) = P(\lambda = 100)$

$E = 1$

100

$P(\ldots | DNA)$

$P(A|B) \neq 1$

$P(B|A)$

$$P(S \ldots | DNA) = \frac{P(DNA|S) \cdot P(S)}{P(DNA)}$$

1, 2, 3, 4, 5

$\frac{1}{1+100}$

99.99%

1%