Statistics Pre-course
Probability - Part 2

Ilaria Mozzetta

Tor Vergata University of Rome

September 2026

Main probability distributions

- Often you do not have to derive the distribution of a random variable on your own.
- You can choose from a catalogue of known random variables whose functions are known and deeply investigated.
- You then select the one that is the most adequate to the phenomenon under analysis.

Known discrete random variables

- Uniform
- Bernoulli
- Binomial
- Poisson
- Geometric
- Hypergeometric
- ...

Discrete Uniform Distribution

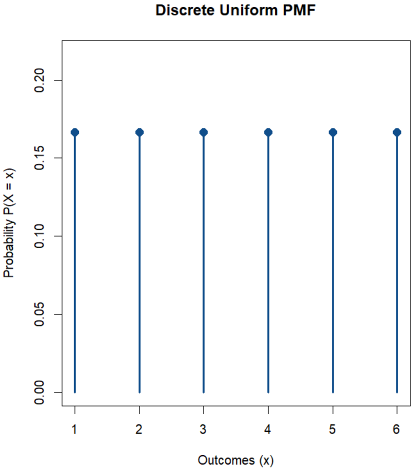
- This is the simplest distribution for whole numbers.
- Every possible outcome from 1 to n has the exact same chance of happening.

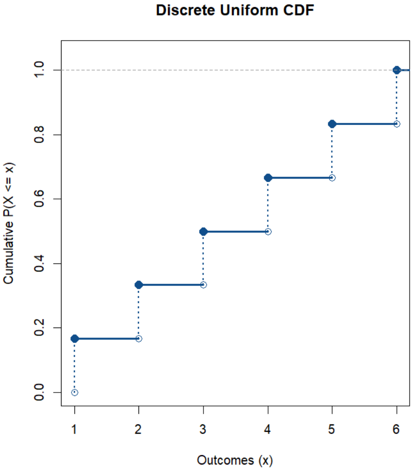
$$\mathbb{P}(X = x) = \frac{1}{n} \quad \text{for} \quad x = 1, 2, \ldots, n$$

- $E(X) = \frac{n+1}{2}$
- $Var(X) = \frac{n^2-1}{12}$

  Example: Rolling a fair 6-sided die. Every number from 1 to 6 has a $\frac{1}{6}$ chance of coming up.

Visualization





Bernoulli

- Assume that a random experiment has two possible outcomes (typically addressed as success or failure).
- The random variable X representing the result of the experiment can take either 0 or 1 as values. We have that:

  Probability of success $\mathbb{P}(X = 1) = p$

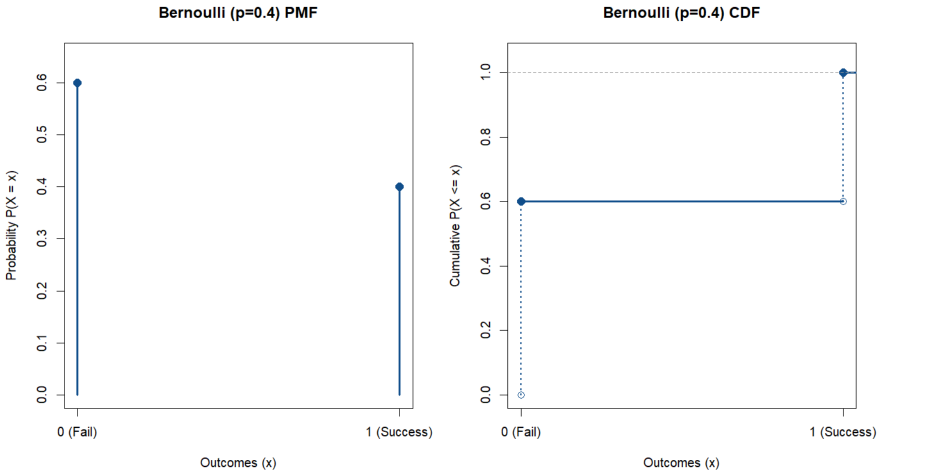
  Probability of failure $\mathbb{P}(X = 0) = 1 - p$
- The random variable X is distributed as a Bernoulli:

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

$$\mathbb{P}(X = x) = \begin{cases} 0 & \text{with probability } 1 - p \\ 1 & \text{with probability } p \end{cases}$$

Example: result of an exam (pass or fail)

Visualization



$E(X)$ AND $V(X)$

If:

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

Then:

- **Expected Value:**

$$\begin{aligned} E[X] &= \sum x \cdot P(X = x) \\ &= (1 \cdot p) + (0 \cdot (1 - p)) \\ &= p \end{aligned}$$

- **Variance:**

$$\begin{aligned} E[X^2] &= (1^2 \cdot p) + (0^2 \cdot (1 - p)) = p \\ Var(X) &= E[X^2] - (E[X])^2 \\ &= p - p^2 \\ &= p(1 - p) \end{aligned}$$

Example

Let X be the random variable representing the price behaviour of a Microsoft's stock.

- $X = 1$ if the price goes up
- $X = 0$ if the price goes down (assuming it cannot stay fixed)

The price can go up with probability $3/5$. Then X follows a Bernoulli distribution with parameter $p = 3/5$.

$$X \sim \text{Bernoulli}(3/5)$$

$$X = \begin{cases} 0 & \text{with probability } \frac{2}{5} \\ 1 & \text{with probability } \frac{3}{5} \end{cases}$$

Binomial Distribution

- What if we repeat that same Bernoulli event multiple times and count how many total Successes we get?
- This is called the **Binomial Distribution**.
- To use it, we only need to know two numbers:
  - n: the total number of times we try.
  - p: the chance of success on each individual try.

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

$$\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad \text{for n} = 0, 1, 2, \ldots, n$$

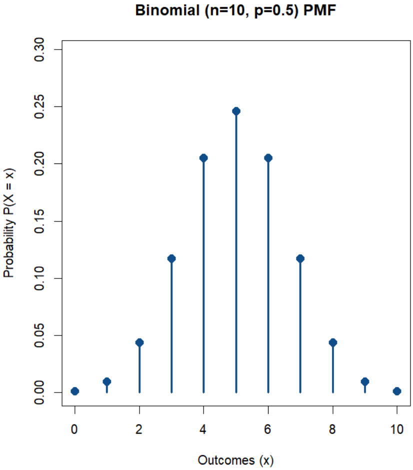
Can we always use the Binomial?

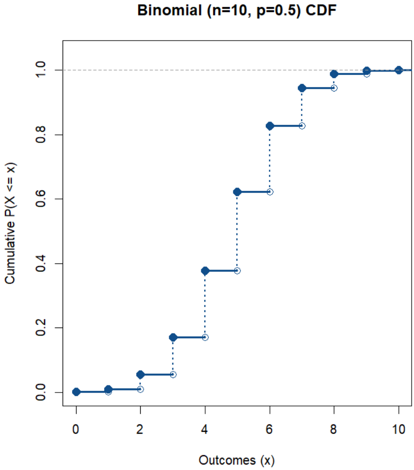
If:

- Each of the n trials has only two possible outcomes. The outcome we are interested in is called success and the other failure
- Each trial has the same probability of success. If the probability of a success is p then the probability of a failure is $1 - p$
- The n trials are independent. The result of one does not affect the results of other trials.

Then X follows a Binomial distribution with parameters n and p.

Visualization





Expected Value

$$\mathbb{E}(X) = \sum_{x \in \mathcal{X}} x p_x = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} \frac{n(n-1)!}{(x-1)!(n-x)!} p^x (1-p)^{n-x}$$

$$= np \sum_{z=0}^{s} \frac{s!}{z!(s-z)!} p^z (1-p)^{s-z} = np$$

Towards the variance

$$
\begin{aligned}
\mathbb{E}[X(X-1)] = \sum_{x \in \mathcal{X}} x(x-1)p_x &= \sum_{x=0}^{n} x(x-1)\binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} x(x-1)\frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} \\
&= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x} \\
&= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!((n-2)-(x-2))!} p^{x-2} (1-p)^{(n-2)-(x-2)} \\
&= n(n-1)p^2 \sum_{z=0}^{s} \frac{s!}{z!(s-z)!} p^z (1-p)^{s-z} = n(n-1)p^2
\end{aligned}
$$

Variance

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

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

$$\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = n(n-1)p^2 + np - (np)^2$$

$$= np[np - p + 1 - np] = np(1-p)$$

Consider $X \sim \text{Binomial}(n, p)$ as the sum of n independent Bernoulli random variables Y_i.
Remember that if $Y_1, \ldots, Y_n \sim \text{Bernoulli}(p)$ then $\mathbb{E}[Y_i] = p$ and $\mathbb{V}[Y_i] = p(1-p)$ $\forall i$, which is enough to prove

$$\mathbb{E}(X) = \mathbb{E}\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} \mathbb{E}(Y_i) = np$$

Moreover, since $Y_1, \ldots, Y_n$ are independent we have that

$$\mathbb{V}(X) = \mathbb{V}\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} \mathbb{V}(Y_i) = np(1-p)$$

Example

A quality control audit finds that 20% of microchips produced on a line are defective (meaning 80% are functional). What is the probability that out of 8 randomly selected microchips, exactly 5 are functional?

A basketball player makes her free throws 75% of the time ($p = 0.75$). If she takes 5 shots, what is the probability of:

- Exactly 3 successful shots?
- At least 1 successful shot?

Example

1. Microchip Problem ($X \sim \text{Binomial}(n = 8, p = 0.80)$)

$$\mathbb{P}(X = 5) = \binom{8}{5}(0.80)^5(0.20)^{8-5}$$

$$\mathbb{P}(X = 5) = 56 \cdot (0.32768) \cdot (0.008) \approx 0.1468 \quad (14.68\%)$$

2. Free Throw Problem ($Y \sim \text{Binomial}(n = 5, p = 0.75)$)

- **Exactly 3 hits:**

$$\mathbb{P}(Y = 3) = \binom{5}{3}(0.75)^3(0.25)^2 = 10 \cdot (0.4219) \cdot (0.0625) \approx 0.2637 \quad (26.37\%)$$

- **At least 1 hit:**

$$\mathbb{P}(Y \geq 1) = 1 - \mathbb{P}(Y = 0) = 1 - \binom{5}{0}(0.75)^0(0.25)^5$$

$$\mathbb{P}(Y \geq 1) = 1 - (0.25)^5 = 1 - 0.00098 = 0.99902 \quad (99.90\%)$$

Exercise

An airline knows from historical data that 10% of passengers who book a flight do not show up ($p = 0.10$).

If 12 passengers book tickets for a flight, let X be the number of no-shows ($X \sim \text{Binomial}(12, 0.10)$).

Questions

1. What is the probability that **exactly 2** passengers do not show up?
2. What is the probability that **at most 1** passenger does not show up ($\mathbb{P}(X \leq 1)$)?
3. What are the expected number of no-shows $\mathbb{E}(X)$ and the variance $\mathbb{V}(X)$?

Geometric distribution

The Geometric Distribution gives the distribution of the number X of Bernoulli trials needed to get one success.

If the probability of success in each trial is p, then the probability of observing a success on the x-th trial (after $x-1$ failures) is

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

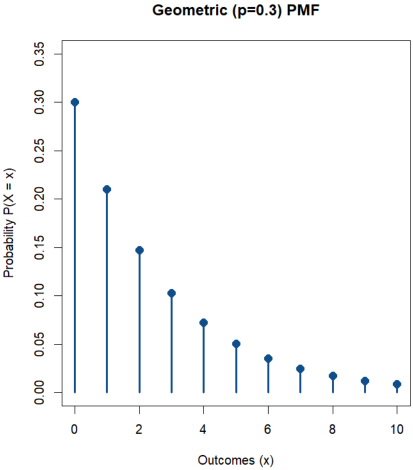
$$p_X(x) = \mathbb{P}(X = x) = (1-p)^{x-1}p$$
$$F_X(x) = 1 - (1-p)^x$$

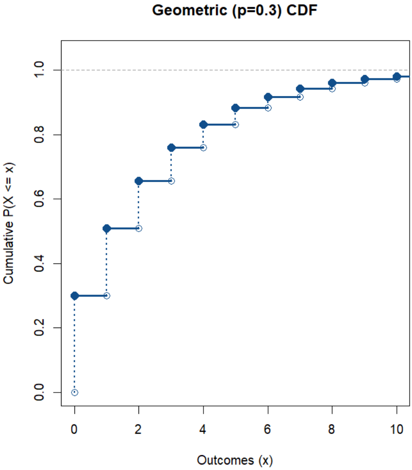
Geometric

The Geometric distribution is a suitable model for a random variable X if:

- X is the result of an experiment which requires a sequence of independent trials
- There are only two possible outcomes for each trial (success or failure)
- Each trial has the same probability of success p

Visualization





EXPECTED VALUE

$$\mathbb{E}(X) = \sum_{x \in \mathcal{X}} x p_X(x) = \sum_{x=0}^{\infty} x(1-p)^{x-1} p$$

$$= p \sum_{x=1}^{\infty} x(1-p)^{x-1} = p \sum_{x=1}^{\infty} -\frac{d}{dp}(1-p)^x = -p \frac{d}{dp} \sum_{x=1}^{\infty} (1-p)^x$$

Using the rule for geometric series we get

$$\mathbb{E}(X) = -p \frac{d}{dp} \frac{1}{p} = -p \left(-\frac{1}{p^2}\right) = \frac{1}{p}$$

Variance

$$\mathbb{E}(X^2) = \sum_{x=0}^{\infty} x^2 p_X(x) = \sum_{x=0}^{\infty} x^2 (1-p)^{x-1} p = p \sum_{x=0}^{\infty} x^2 (1-p)^{x-1}$$

Substitute $q = (1-p)$ and solve the infinite series

$$\mathbb{E}(X^2) = p \sum_{x=0}^{\infty} x^2 q^{x-1} = p \frac{1+q}{(1-q)^3} = p \frac{2-p}{p^3} = \frac{2-p}{p^2}$$

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

$$\mathbb{V}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Example

A software tester finds that a rare system glitch occurs in automated runs with a probability of $p = 0.05$ (5%) per run.

Let X be the number of test runs needed to detect the **first glitch** ($X \sim \text{Geometric}(0.05)$).

Questions

1. What is the probability that the first glitch is detected on the **4th test run**?
2. How many test runs do you expect to execute before detecting the first glitch?

Example

1. First glitch on the 4th run ($\mathbb{P}(X = 4)$)

$$\mathbb{P}(X = x) = (1 - p)^{x-1} p$$

$$\mathbb{P}(X = 4) = (1 - 0.05)^{4-1}(0.05) = (0.95)^3(0.05)$$

$$\mathbb{P}(X = 4) = (0.8574) \cdot 0.05 \approx 0.0429 \quad (4.29\%)$$

2. Expected number of runs ($\mathbb{E}(X)$)

$$\mathbb{E}(X) = \frac{1}{p} = \frac{1}{0.05} = 20 \text{ test runs}$$

Interpretation

You expect to wait through 19 successful runs before finding the glitch on average on the 20th run.

Exercise

An IT help desk analyst knows that a caller has a password reset issue with probability $p = 0.20$ (20%).

Questions

1. What is the probability that the first password reset request occurs on the **3rd call**?
2. What is the probability that it takes **at most 2 calls** to encounter the first password reset request?
3. Compute the expected number of calls $\mathbb{E}(X)$ and the standard deviation σ_X.

Poisson

The Poisson distribution is known as the distribution of rare events. It is typically used to model counts, i.e. the number of events in a given interval of time (or space).

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

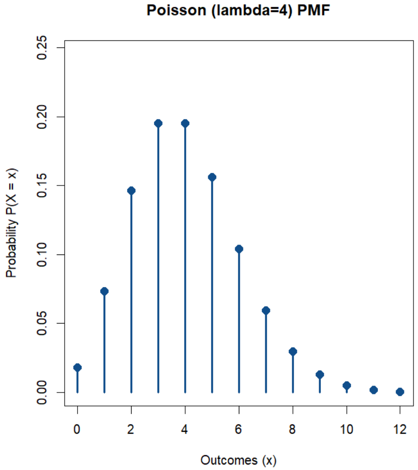
$$p_X(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x \in \{0, 1, 2, \dots\}$$

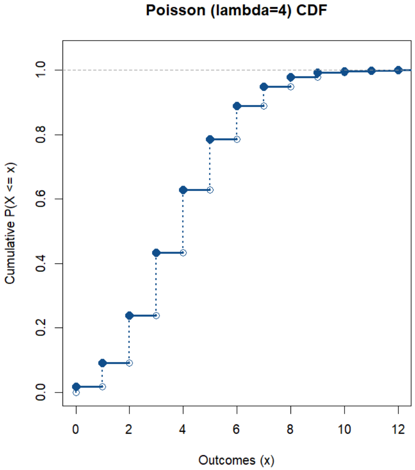
$$F_X(x) = \sum_{k \leq x} p_X(k)$$

Examples:

- number of clients calling a call center
- number of defects per square meter of manufactured goods
- number of earthquakes in a year

Visualization





Expected Value

Recall: $e^{\alpha} = \sum_{s=0}^{\infty} \frac{\alpha^s}{s!}$

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} x p_x = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} = \sum_{x=1}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-1)!}$$

$$= e^{-\lambda} \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = e^{-\lambda} \lambda \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} = \lambda$$

Variance

$$\mathbb{E}[X(X-1)] = \sum_{x \in \mathcal{X}} x(x-1)p_x = \sum_{x=0}^{\infty} x(x-1)\frac{\lambda^x e^{-\lambda}}{x!} = \sum_{x=2}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-2)!}$$

$$= e^{-\lambda}\lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = e^{-\lambda}\lambda^2 \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} = \lambda^2$$

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

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

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

Example

A bank branch receives an average of $\lambda = 3$ customer arrivals per 10-minute window during peak hours.

Let X be the number of arrivals in a 10-minute period ($X \sim \text{Poisson}(3)$).

Questions

1. What is the probability of receiving **exactly 2** arrivals in a 10-minute window?
2. What is the probability of receiving **no arrivals** ($X = 0$)?

Example

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

1. Exactly 2 arrivals ($\mathbb{P}(X = 2)$)

$$\mathbb{P}(X = 2) = \frac{3^2 e^{-3}}{2!} = \frac{9 \cdot 0.0498}{2} \approx 0.2240 \quad (22.40\%)$$

2. No arrivals ($\mathbb{P}(X = 0)$)

$$\mathbb{P}(X = 0) = \frac{3^0 e^{-3}}{0!} = \frac{1 \cdot 0.0498}{1} \approx 0.0498 \quad (4.98\%)$$

Exercise

A cloud web server experiences an average of $\lambda = 2$ critical error alerts per day.

Questions

1. What is the probability of receiving **exactly 4** alerts in a given day?
2. What is the probability of receiving **at least 1** alert in a day?
3. What are the mean $\mathbb{E}(X)$ and variance $\mathbb{V}(X)$ for this process?

Known Continuous Random Variables

- Uniform
- Exponential
- Normal
- Student's t
- Gamma
- Beta
- ...

From Discrete to Continuous Uniform:

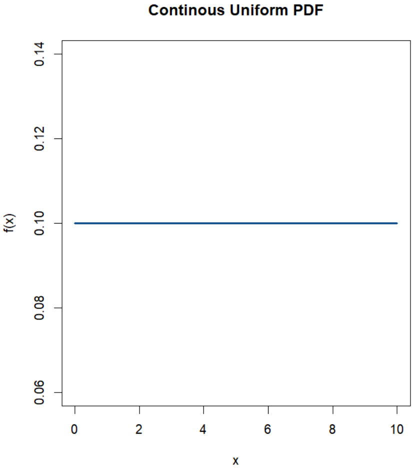
- **Discrete Uniform:** X takes finite values $\{x_1, \dots, x_n\}$ with equal point probability $P(X = x_i) = \frac{1}{n}$.
- **What Changes in Continuous?** X takes values in a continuous interval $[a, b]$.
  - Individual point probabilities are zero: $P(X = x) = 0$.
  - The probability density $f_X(x)$ is constant across $[a, b]$.

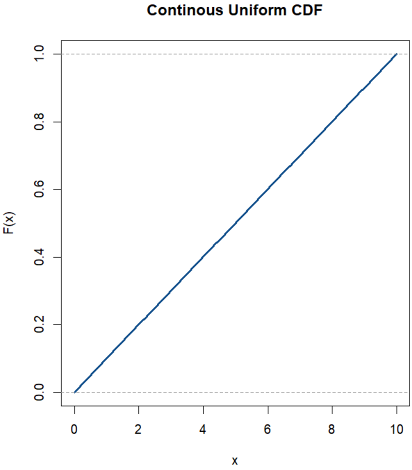
$$X \sim \text{Unif}(a, b)$$

$$f_X(x) = \frac{1}{b-a}, \quad F_X(x) = \frac{x-a}{b-a}$$

Example: The arrival time of a bus between the moment you reach the stop and midnight.

Visualization





Continuous Uniform: Expected Value & Exercise

Expected Value Derivation:

$$\mathbb{E}[X] = \int_a^b x f_X(x)\,dx = \int_a^b x \frac{1}{b-a}\,dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$
$$= \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right) = \frac{(b-a)(a+b)}{2(b-a)} = \frac{a+b}{2}$$

Exercise 1: Prove that $\mathbb{V}[X] = \frac{(b-a)^2}{12}$.

Exercise 2

The current (in mA) measured in a piece of copper wire follows $X \sim \text{Unif}(0, 48)$.

- Write down the formula for the PDF $f_X(x)$ and CDF $F_X(x)$.
- Calculate the mean $\mathbb{E}[X]$ and variance $\mathbb{V}[X]$.

Exponential Distribution

A random variable X follows an Exponential distribution with parameter $\lambda > 0$, denoted $X \sim \text{Exp}(\lambda)$, if its PDF is:

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

- The intuition behind an Exponential random variable is that the larger the value, the less likely it is.
- Typically used to model waiting time until a specific event occurs.
- The rate parameter λ represents the average number of events per unit time.

Examples:

- Time until an earthquake strikes.
- Amount spent by a customer during a supermarket visit.

Exponential: PDF and CDF

For $X \sim \text{Exp}(\lambda)$ with rate $\lambda > 0$:

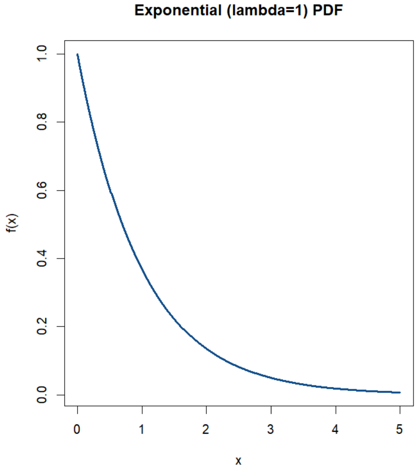
- **Probability Density Function (PDF):**

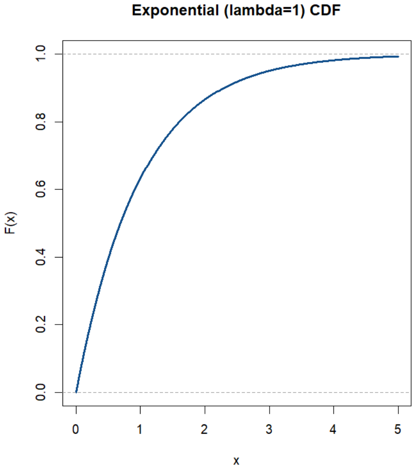
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

- **Cumulative Distribution Function (CDF):**

$$F_X(x) = \int_0^x \lambda e^{-\lambda t}\, dt = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Visualization





Exponential: Expected Value

$$\mathbb{E}[X] = \int_0^\infty x f_X(x)\, dx = \int_0^\infty x(\lambda e^{-\lambda x})\, dx = \lambda \int_0^\infty x e^{-\lambda x}\, dx$$

Integration by parts ($\int u\, dv = uv - \int v\, du$):

$$\text{Let } u = x \implies du = dx, \quad dv = e^{-\lambda x} dx \implies v = -\frac{1}{\lambda} e^{-\lambda x}$$

$$\begin{aligned}
\mathbb{E}[X] &= \lambda \left(\left[-\frac{x e^{-\lambda x}}{\lambda} \right]_0^\infty - \int_0^\infty -\frac{1}{\lambda} e^{-\lambda x}\, dx \right) \\
&= \lambda \left(0 + \frac{1}{\lambda} \left[-\frac{e^{-\lambda x}}{\lambda} \right]_0^\infty \right) \\
&= \lambda \cdot \frac{1}{\lambda^2} = \frac{1}{\lambda}
\end{aligned}$$

Exercises

Exercise 1:

Find the variance $\mathbb{V}[X] = \frac{1}{\lambda^2}$ by computing $\mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

Exercise 2: Job Arrival

If jobs arrive every 15 seconds on average ($\lambda = 4$ per minute), what is the probability of waiting 30 seconds (0.5 min) or less?

Exercise 3: Statistics Study Time

Study time X for researchers follows $X \sim \text{Exp}(\lambda)$ with mean $\mathbb{E}[X] = 15$ minutes.

1. State the parameter λ and write the PDF $f_X(x)$.
2. Find $P(60 \leq X \leq 120)$, the probability of studying between 1 and 2 hours.

Normal Distribution

The Normal or Gaussian is the queen of all random variables.

- Represents many natural, physical, and economic phenomena.
- Approximates other distributions via the Central Limit Theorem.
- Fundamental to sampling theory and statistical inference.

A traditional parametrization: $\mathbb{E}[X] = \mu$, $\mathbb{V}[X] = \sigma^2$

Normal: PDF and CDF

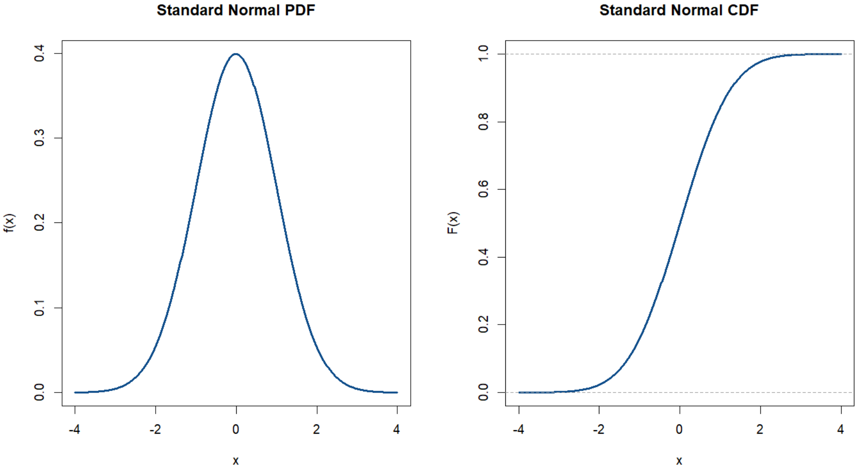
For $X \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu \in \mathbb{R}$ and $\sigma > 0$:

- **Probability Density Function (PDF):**

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad x \in \mathbb{R}$$

- **Cumulative Distribution Function (CDF):**

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

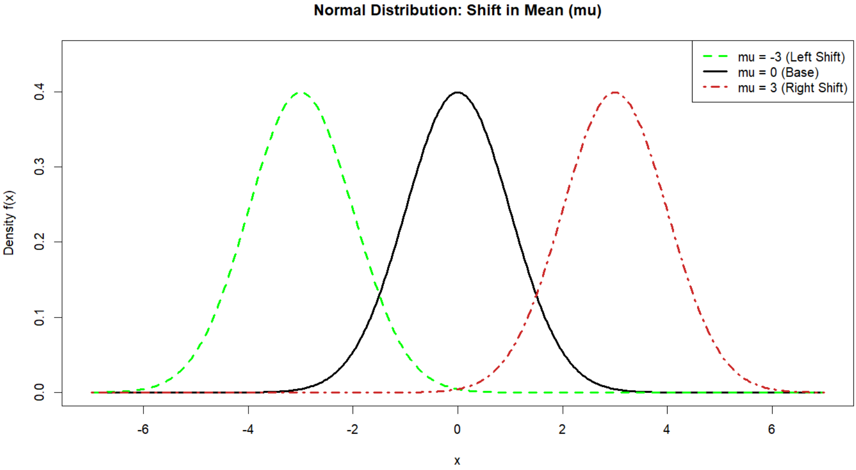

Standard Normal PDF
f(x)
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4
x
Standard Normal CDF
F(x)
0.0
0.2
0.4
0.6
0.8
1.0
-4
-2
0
2
4
x

Normal: Effect of Varying μ

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

- The mean μ acts as the **location parameter**.
- Changing μ shifts the bell curve horizontally along the x-axis.
- The shape, peak height, and width of the curve remain unchanged.
- The highest point (mode/median/mean) always lies at $x = \mu$.

Visualization

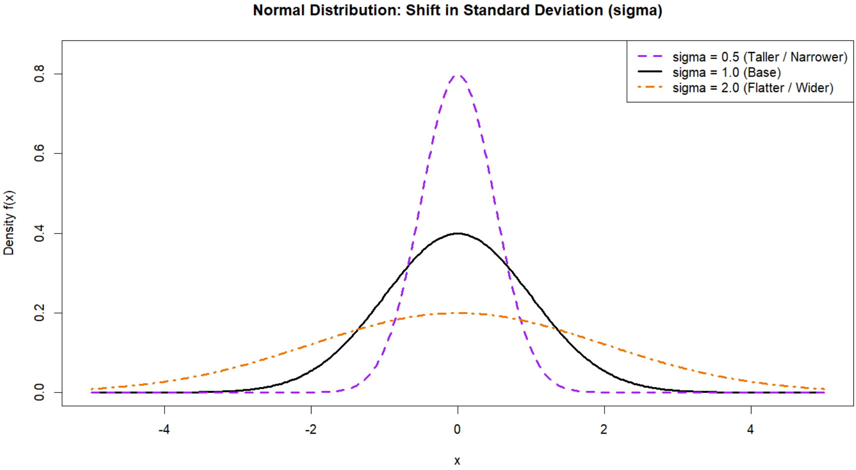


Normal: Effect of Varying σ

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

- The standard deviation σ acts as the **scale parameter**.
- Controls the spread and concentration of probability mass:
  - **Smaller σ:** Concentrated mass, taller and narrower peak.
  - **Larger σ:** Dispersed mass, flatter and wider bell curve.
- The total area under the density curve remains 1.

Visualization



1. Linear Transformation: If $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = aX + b$, then

$$Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

2. Linear Combination of Independent Normals: If $X_1, \ldots, X_n$ are independent with $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, then

$$Y = \sum_{i=1}^{n} a_i X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

Standard Normal Distribution

When $\mu = 0$ and $\sigma^2 = 1$, $Z \sim \mathcal{N}(0, 1)$ is called the **Standard Normal Distribution**.

Standardization: Any $X \sim \mathcal{N}(\mu, \sigma^2)$ can be transformed into a Standard Normal Z:

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

Expected value and Variance:

$$\mathbb{E}[Z] = \mathbb{E}\left[\frac{X - \mu}{\sigma}\right] = \frac{\mathbb{E}[X] - \mu}{\sigma} = \frac{\mu - \mu}{\sigma} = 0$$

$$\mathbb{V}[Z] = \mathbb{V}\left[\frac{X - \mu}{\sigma}\right] = \frac{\mathbb{V}[X]}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1$$

Standard Normal CDF $\Phi(z)$ & Tables

The cumulative distribution function of $Z \sim \mathcal{N}(0,1)$ is denoted by $\Phi(z)$:

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

Symmetry & Cumulative Properties

- $P(Z > z) = 1 - \Phi(z)$
- $\Phi(-z) = 1 - \Phi(z)$
- $P(a \leq Z \leq b) = \Phi(b) - \Phi(a)$

Values for $\Phi(z)$ are looked up in Standard Normal (Z) probability tables.

Z TABLES

Standard Normal Probabilities

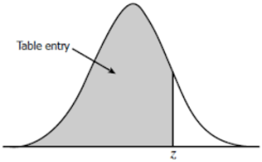


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177

Reading the Standard Normal (Z) Table

We want to find the cumulative probability $\Phi(z) = P(Z \leq z)$.

- **Row Header:** First digit and first decimal place of z (e.g., 1.9).
- **Column Header:** Second decimal place of z (e.g., 0.06).
- **Intersection:** The probability value $\Phi(1.96) = 0.9750$.

Useful Symmetry Rules for Z-Tables

- $P(Z > z) = 1 - \Phi(z)$
- $P(Z \leq -z) = 1 - \Phi(z) \quad \Longrightarrow \quad \Phi(-1.96) = 1 - 0.9750 = 0.0250$
- $P(a \leq Z \leq b) = \Phi(b) - \Phi(a)$

Example: Normal Distribution

A beverage filling machine fills bottles with an average volume of $\mu = 500$ mL and standard deviation $\sigma = 10$ mL ($X \sim \mathcal{N}(500, 100)$).

Questions

1. What is the probability that a randomly chosen bottle has **less than 485 mL**?
2. What is the probability that a bottle contains **between 490 mL and 515 mL**?

Solution: Normal Distribution Example

1. **Less than 485 mL ($\mathbb{P}(X < 485)$)**

$$Z = \frac{X - \mu}{\sigma} = \frac{485 - 500}{10} = -1.50$$

$$\mathbb{P}(X < 485) = \mathbb{P}(Z < -1.50) = \Phi(-1.50) \approx 0.0668 \quad (6.68\%)$$

2. **Between 490 mL and 515 mL ($\mathbb{P}(490 < X < 515)$)**

$$Z_1 = \frac{490 - 500}{10} = -1.00 \quad \text{and} \quad Z_2 = \frac{515 - 500}{10} = +1.50$$

$$\mathbb{P}(-1.00 < Z < 1.50) = \Phi(1.50) - \Phi(-1.00)$$

$$\mathbb{P}(-1.00 < Z < 1.50) = 0.9332 - 0.1587 = 0.7745 \quad (77.45\%)$$

Exercises

- The time (in minutes), X, that is needed to solve this Statistics exercise is normally distributed with mean 5 and standard deviation 10. When I solved it at home, it took me 6.2 minutes. What is the probability that a random PhD student faster than me?
- X is a normally distributed random variable with mean 30 and standard deviation 4.
  - Find $P(X < 40)$
  - Find $P(X > 21)$
  - Find $P(30 < X < 35)$
- Entry to a certain University is determined by a national test. The scores on this test are normally distributed with a mean of 500 and a standard deviation of 100. Tom wants to be admitted to this university and he knows that he must score better than at least 70% of the students who took the test. Tom takes the test and scores 585. Will he be admitted to this university?