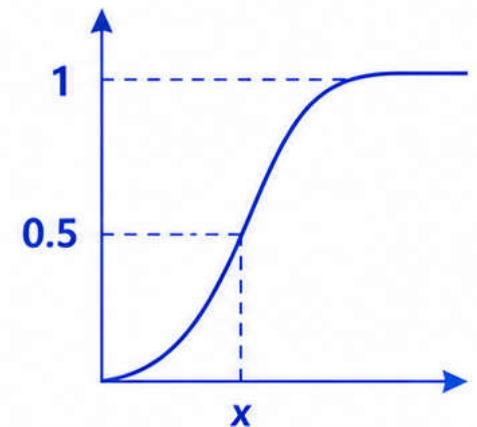
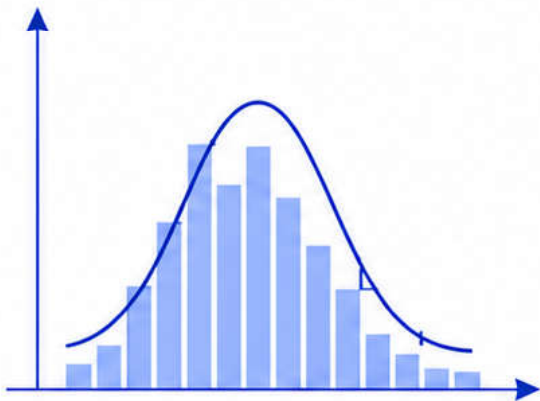


Lecture 6

Probability Distributions I

*Random variables, probability distributions
and their characterization*



PHYSICAL MEASUREMENT METHODS

From the previous lecture

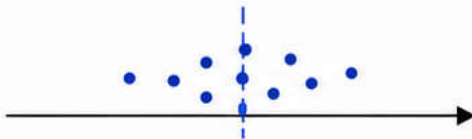


Every measurement has **uncertainty**.

Uncertainty can be **random** (statistical) or **systematic**.

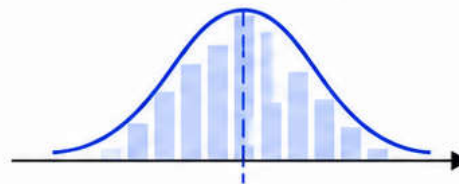
Repeated measurements

Under (nominally) the same conditions we obtain slightly different results.



Distribution of results

The set of results can be described by a **probability distribution**.



Quantitative description

The distribution is characterized by **parameters** (e.g., mean, standard deviation, etc.).

$$\bar{x} \quad \sigma \quad \dots$$



Probability theory provides the language and tools to describe, analyze and interpret the variability of measurement results.

Random variable



A **random variable** is a quantity whose value cannot be predicted exactly before an experiment, but whose possible values and probabilities can be described.

Examples in physical measurements and experiments:



Number of detected photons in a time interval



Number of successful events in repeated trials



Measured voltage of an electrical signal



Lifetime of a particle or a component



Measured temperature or other physical quantity



Probability theory provides the mathematical framework for describing and analyzing **random variables** and their **distributions**.

Why do we need probability distributions?



In physical measurements and experiments, outcomes are **uncertain** and **vary** from one trial to another.

Probability distributions allow us to **describe**, **analyze** and **predict** this variability in a quantitative way.

They help us to:



Quantify **uncertainty** in a systematic and consistent way.



Characterize the behavior of random variables.



Extract **meaningful information** from data.



Compare different systems, methods or models.



Make **predictions** and assess the **reliability** of results.



Probability distributions are the **foundation** of statistical analysis and decision making in physics and engineering.

Two types of random variables



Random variables can be classified into two main types: **discrete** and **continuous**.

Discrete random variable

Can take a finite number or a countably infinite number of **distinct values**.

Examples:

- Number of heads in 10 coin tosses
- Number of detected particles
- Number of defects in a sample
- Number of successful events



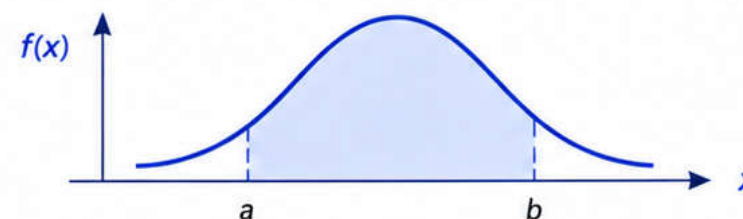
VS.

Continuous random variable

Can take an uncountably infinite number of values within an **interval**.

Examples:

- Length, distance, position
- Time, lifetime
- Temperature
- Voltage, current, intensity



Discrete probability distribution



A discrete probability distribution specifies how a discrete random variable K takes each of its possible values $k_1, k_2, \dots, k_j, \dots$ with certain **probabilities** p_j .

► Definition

If K can take the values

$k_1, k_2, \dots, k_j, \dots$

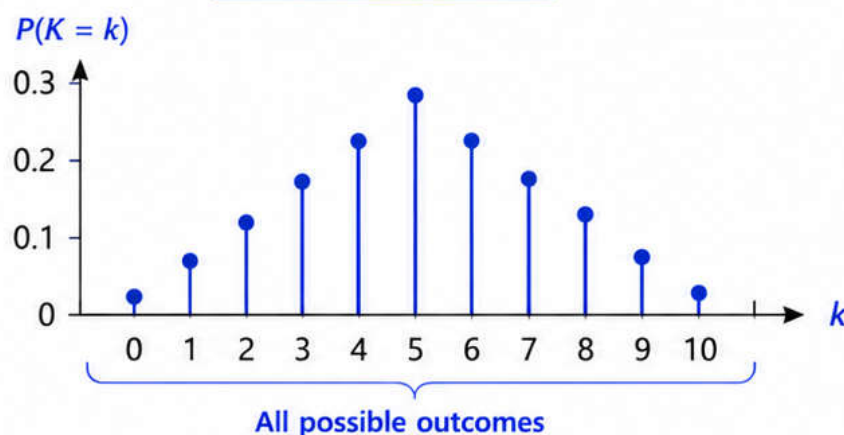
then the probability of each value is assigned:

$$p_j = P(K = k_j)$$

for $j = 1, 2, \dots$

Example: number of heads in 10 coin tosses

$$K = 0, 1, 2, \dots, 10$$



► Normalization

Since the outcomes are mutually exclusive and exhaustive:

$$\sum_{j=1}^{\infty} p_j = 1$$

- $0 \leq p_j \leq 1$ for all j
- Total probability = 1



A discrete probability distribution is often visualized by a **bar chart** or **stem plot**.

Continuous probability distribution



A **continuous probability distribution** describes a continuous random variable X that can take any value in an interval $[a, b]$ (or $(-\infty, \infty)$) via a **probability density function** $f(x)$.

► Definition

A function $f(x)$ is a probability density function (pdf) of X if:

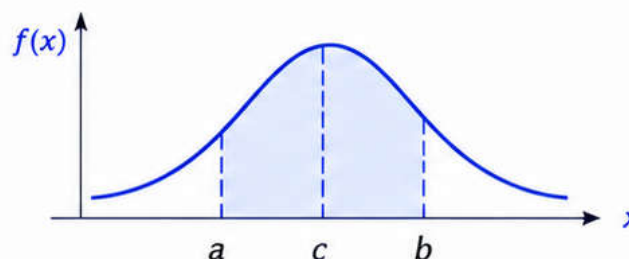
1. $f(x) \geq 0$ for all x
2. $\int_{-\infty}^{\infty} f(x) dx = 1$

For any $a \leq c \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

► Interpretation

$f(x)$ is not a probability, but a **density**. Probabilities are areas under the curve.



$$P(a \leq X \leq b) = \text{Area shaded} = \int_a^b f(x) dx$$

► Properties

- $P(X = x_0) = 0$ for any single point x_0
- $P(X \leq x) = \int_{-\infty}^x f(t) dt$
- $P(X > x) = \int_x^{\infty} f(t) dt = 1 - P(X \leq x)$

Examples

- Measurement errors
- Particle lifetimes
- Time between events
- Voltage, current, temperature, etc.



Continuous distributions are essential for modeling **real-valued measurements** and **natural phenomena**.

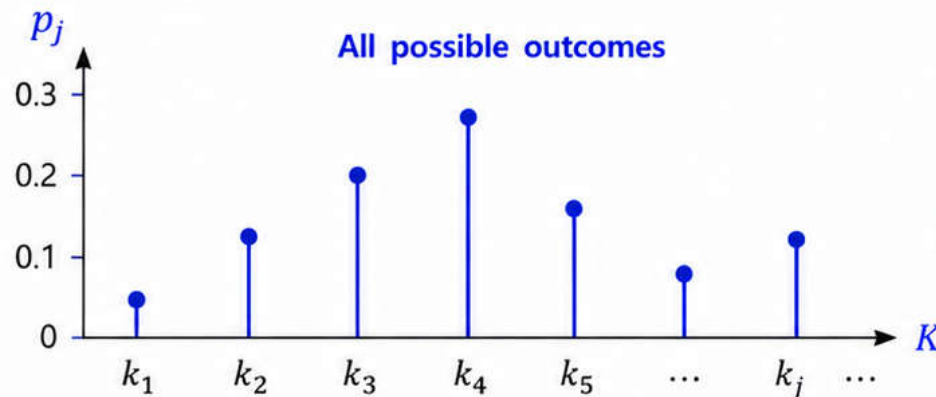
Normalization



For a discrete random variable, the possible outcomes are **mutually exclusive** and **exhaustive**. Therefore, the **total probability** of all possible outcomes must equal 1.

$$\sum_j p_j = 1$$

Interpretation: If we add up the probabilities of all possible values, we account for all possible outcomes—nothing is missing.



Total probability

$$\sum_j p_j = 1$$

(100% of the probability is distributed.)

Simple discrete example: coin toss

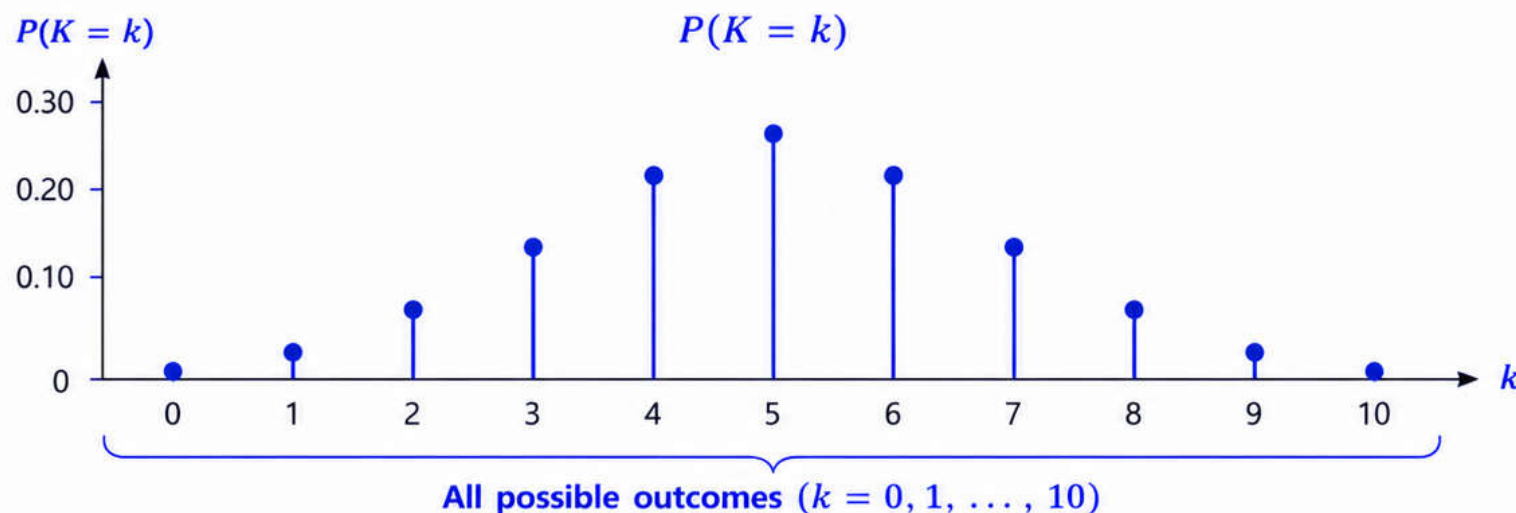


Example: 10 coin tosses

Let K = number of heads in 10 tosses.

Then K can take values $0, 1, \dots, 10$.

- All outcomes are equally likely (fair coin).
- This is a discrete random variable.
- The distribution shown below is the **binomial distribution** with $n = 10$, $p = 0.5$.



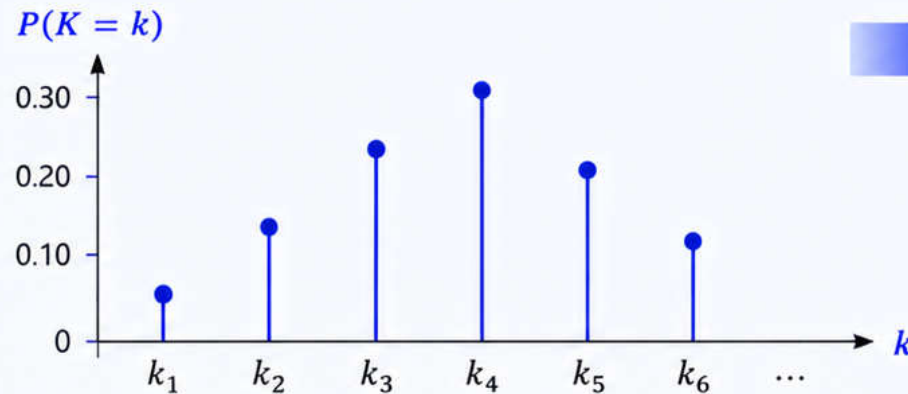
This discrete distribution fully describes the random variable K . It is the foundation for the **binomial distribution** discussed later.

From discrete to continuous

Discrete random variable K

Takes countable values $k_1, k_2, \dots, k_j, \dots$

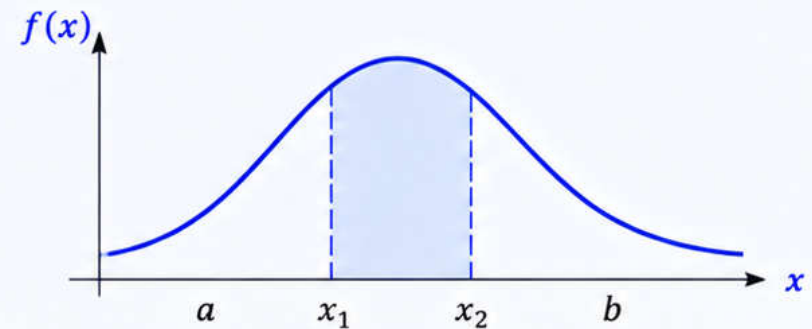
$$P(K = k_j) = p_j$$



Continuous random variable X

Can take any value in an interval (uncountably many).
Individual points become infinitely dense.

Probabilities cannot be assigned
to individual points.



Therefore, **probability must be assigned to intervals** rather than **individual points**.

Cumulative distribution function (CDF)

Definition

$$F(x) = P(X \leq x)$$

Probability that the random variable X does not exceed x .

Three fundamental properties

1 $0 \leq F(x) \leq 1$

A cumulative probability stays between 0 and 1.

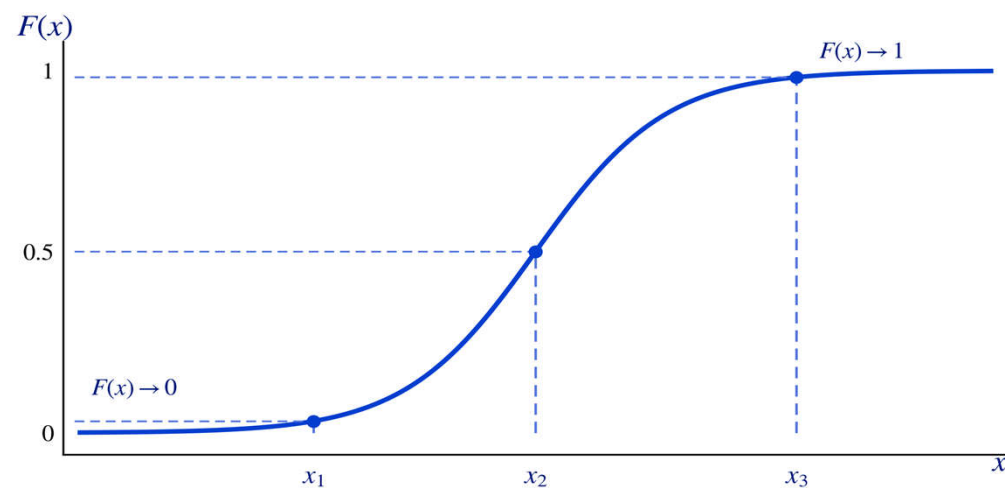
2 $F(-\infty) = 0, \quad F(+\infty) = 1$

Far left: no outcomes counted. Far right: all outcomes counted.

3 $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$

The curve never falls as x increases.

Continuous example: a rising CDF



At x_2 , $F(x_2) = 0.5$: half of the outcomes do not exceed x_2 .

The CDF accumulates probability and encodes the entire distribution in one nondecreasing function.

Probability of an interval of zero width



For a **continuous random variable** X with CDF $F(x)$, the probability of **any interval of zero width** is zero.

$$P(a \leq X \leq a) = 0 \text{ for any real number } a$$

Proof

For any real number a ,

$$P(a \leq X \leq a) = F(a) - F(a^-)$$

Because $F(x)$ is **continuous** at every point a ,

$$F(a^-) = F(a)$$

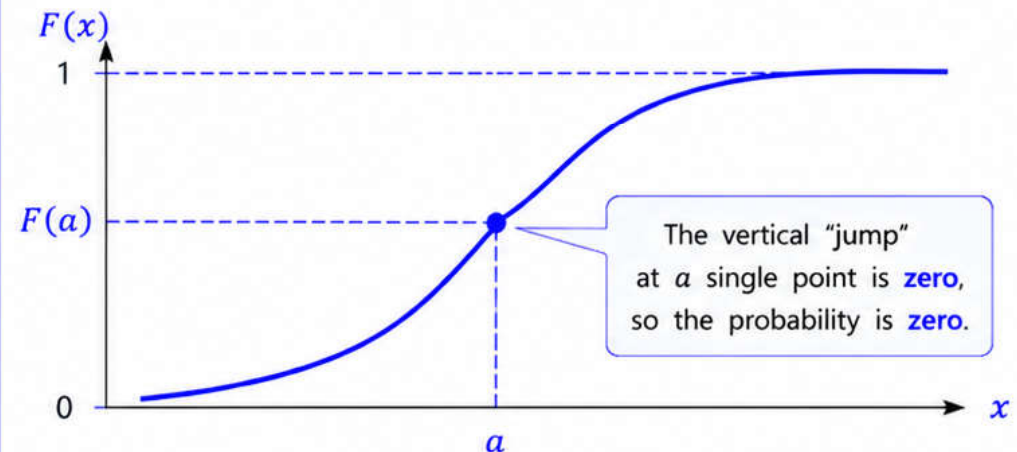


$$P(a \leq X \leq a) = 0$$



Intuitively: an interval with zero width contains no "room" for probability.

Graphical interpretation



For continuous random variables, **probability** is assigned to **intervals**. Intervals with **zero width** (a single point) receive **probability zero**.

Example

$$P(X = 3.7) = 0$$

Probability of an interval of nonzero width



For a **continuous random variable** X with CDF $F(x)$, the probability of an **interval of nonzero width** is **positive**.

$$P(a < X \leq b) = F(b) - F(a) \quad \text{for any } a < b$$

Derivation

For any real numbers $a < b$,

$$\begin{aligned} P(a < X \leq b) &= P(X \leq b) - P(X \leq a) \\ &= F(b) - F(a) \end{aligned}$$

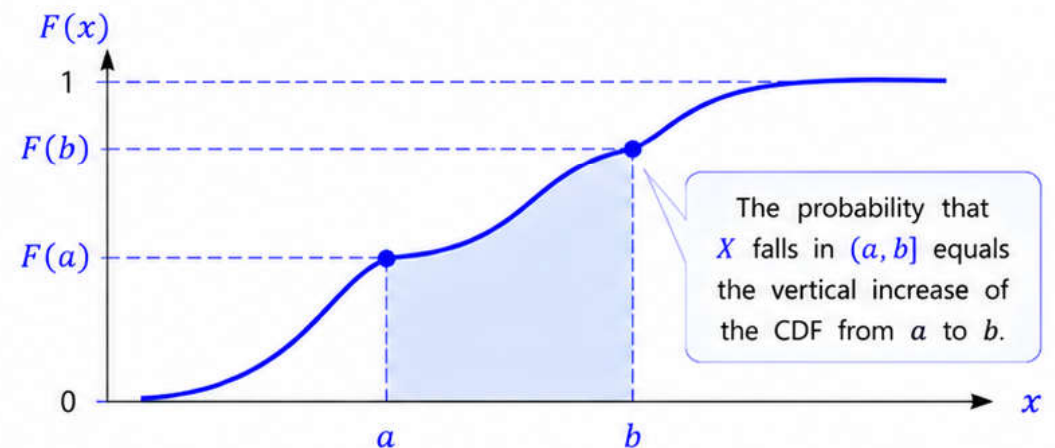


Because $F(x)$ is non-decreasing, $F(b) \geq F(a)$.
 → Therefore $P(a < X \leq b) \geq 0$.

Other equivalent forms

- $P(a \leq X < b) = F(b^-) - F(a)$
 - $P(a \leq X \leq b) = F(b) - F(a)$
 - $P(a < X < b) = F(b^-) - F(a)$
- For a **continuous CDF**,
 $F(b^-) = F(b)$,
 so all definitions coincide.

Graphical interpretation



Area interpretation: $P(a < X \leq b)$ is the "area" under the CDF between a and b .



Intervals with **positive length** ($b - a > 0$) have **positive probability**.

Only intervals with **zero width** ($a = b$) have probability zero.

Probability of intervals: properties



For a **continuous random variable** X with CDF $F(x)$, the probability of **intervals** satisfies important **properties** that follow from the properties of the CDF.

For any real numbers $a < b < c$, all probabilities below are ≥ 0 .

1 Nonnegativity

$$P(a \leq X \leq b) = F(b) - F(a) \geq 0$$

Probability of any interval is never negative.

2 Additivity over adjacent intervals

$$P(a \leq X \leq c) = P(a \leq X \leq b) + P(b \leq X \leq c)$$

The whole interval equals the sum of its parts (no overlap).

3 Additivity over disjoint intervals

$$P(a \leq X \leq b \text{ or } c \leq X \leq d) = P(a \leq X \leq b) + P(c \leq X \leq d)$$

Disjoint intervals add because probability on a single point is zero.

4 Monotonicity

$$\text{If } a_1 \leq a_2 \leq b_2 \leq b_1, \text{ then } P(a_2 \leq X \leq b_2) \leq P(a_1 \leq X \leq b_1)$$

Larger intervals have greater or equal probability.

5 Symmetry by complement

$$P(X \leq a) + P(X > a) = 1 \text{ and } P(X \geq a) + P(X < a) = 1$$

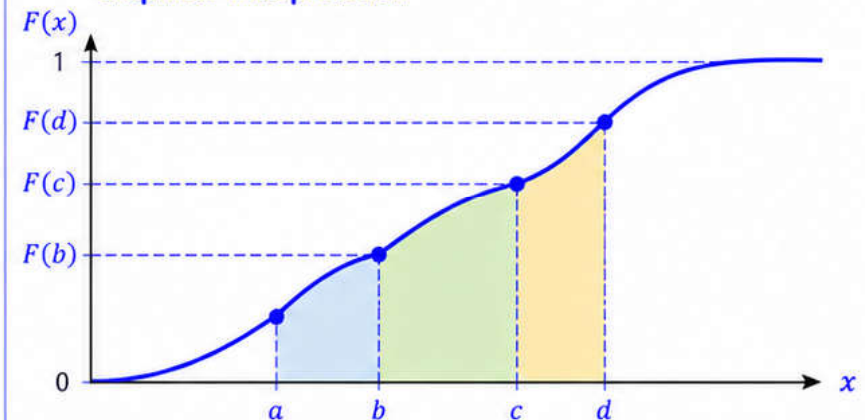
Everything to the left or at a plus everything to the right is 1.

6 Total probability

$$P(-\infty < X < \infty) = 1$$

The entire real line has probability 1.

Graphical interpretation



$$P(a \leq X \leq b) = F(b) - F(a)$$

$$P(b \leq X \leq c) = F(c) - F(b)$$

$$P(c \leq X \leq d) = F(d) - F(c)$$

$$P(a \leq X \leq d) = F(d) - F(a)$$



All interval probability properties follow directly from the fact that the CDF $F(x)$ is **non-decreasing** and **right-continuous**, with $F(-\infty) = 0$ and $F(+\infty) = 1$, and from $P(\text{a single point}) = 0$ for continuous random variables.

Probability of intervals: examples



To compute the probability that a **continuous random variable** X falls in an **interval**, subtract the **CDF value** at the **lower endpoint** from the CDF value at the **upper endpoint**.

$$P(a \leq X \leq b) = F(b) - F(a)$$

for any $a < b$

Examples

1 Finite interval

For $a < b$,

$$P(a \leq X \leq b) = F(b) - F(a)$$

Example: If $F(1.5) = 0.6$ and $F(-0.5) = 0.2$, then

$$P(-0.5 \leq X \leq 1.5) = 0.6 - 0.2 = 0.4$$

2 One-sided intervals

• Left tail: $P(X \leq a) = F(a) - F(-\infty) = F(a)$

Example: If $F(-1) = 0.3$,
 $P(X \leq -1) = 0.3$

• Right tail: $P(X \geq a) = 1 - F(a)$

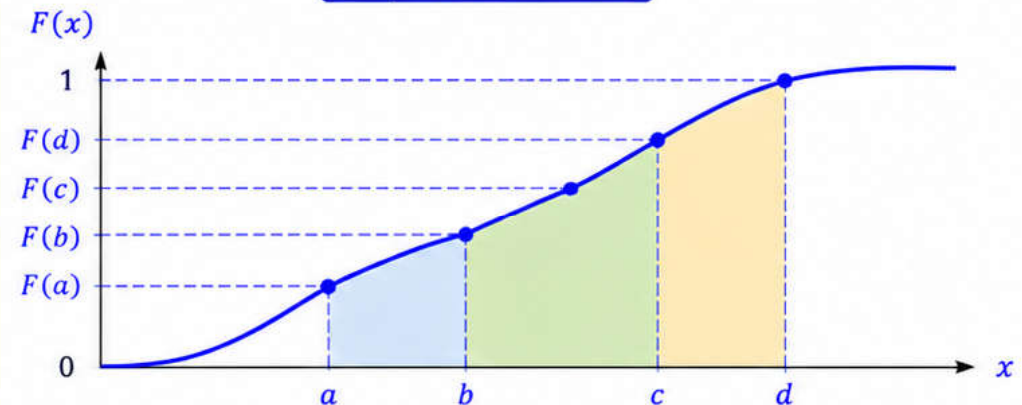
Example: If $F(2) = 0.8$,
 $P(X \geq 2) = 1 - 0.8 = 0.2$

3 Entire real line

$$P(-\infty < X < \infty) = F(\infty) - F(-\infty) = 1 - 0 = 1$$

The total probability is 1.

Graphical illustration



$P(a \leq X \leq b) = F(b) - F(a)$

$P(b \leq X \leq c) = F(c) - F(b)$

$P(c \leq X \leq d) = F(d) - F(c)$

$P(X \leq a) = F(a)$



Summary: For a **continuous random variable**, probabilities are computed using the CDF:

- Use **differences of CDF values** at **endpoints**.
- Include $-\infty$ for **left tails**, ∞ for **right tails**.
- The probability of any interval depends **only on its endpoints**.

Quick checklist

- ✓ Identify the interval.
- ✓ Read the CDF values at the endpoints.
- ✓ Subtract: *upper* – *lower*.

Probability density function

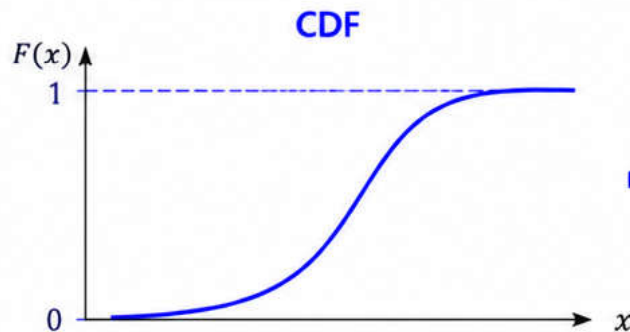


For a **continuous random variable** X with CDF $F(x)$, the **probability density function (PDF)** $f(x)$ is defined as:

$$f(x) = \frac{d}{dx} F(x)$$

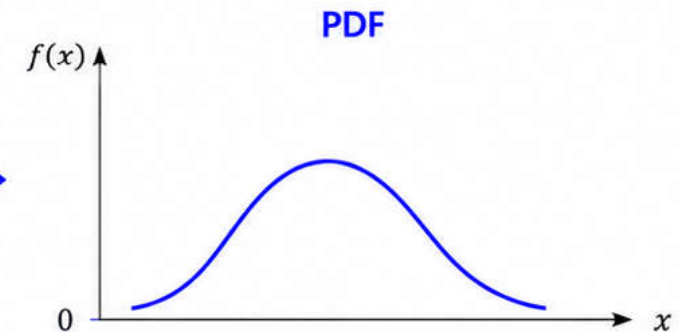
That is, the PDF is the derivative of the CDF.

Visual connection



derivative

$$f(x) = \frac{d}{dx} F(x)$$



The PDF $f(x)$ describes how the probability is distributed along the real line. It is **nonnegative** and its **total area** equals 1.



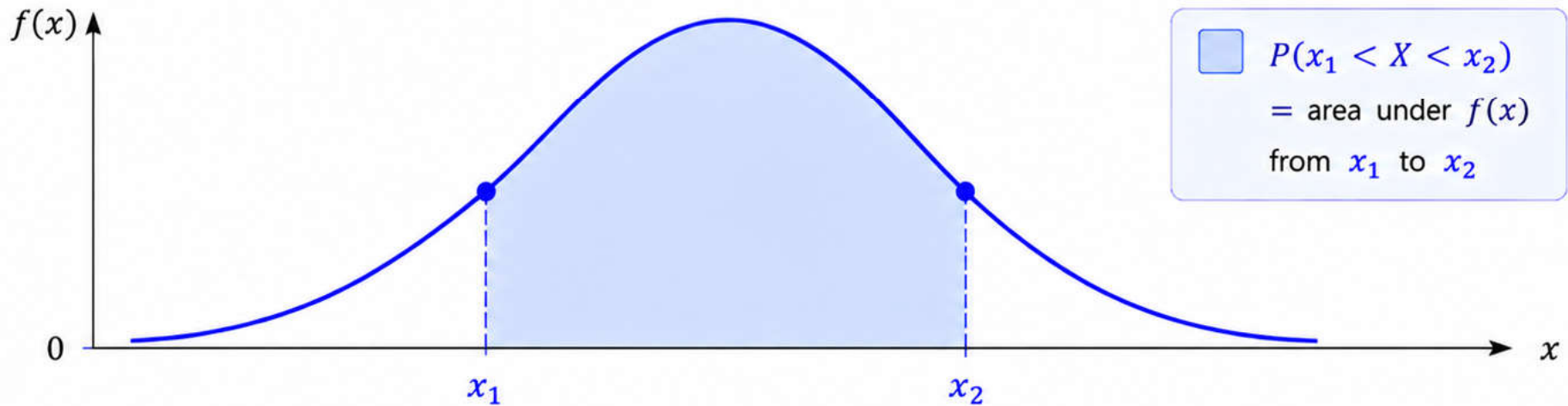
Intuition: The CDF **accumulates** probability, the PDF shows **how fast** it accumulates.

Probability from the PDF



For a **continuous random variable** X with **PDF** $f(x)$, the probability that X lies between x_1 and x_2 is

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x) dx$$



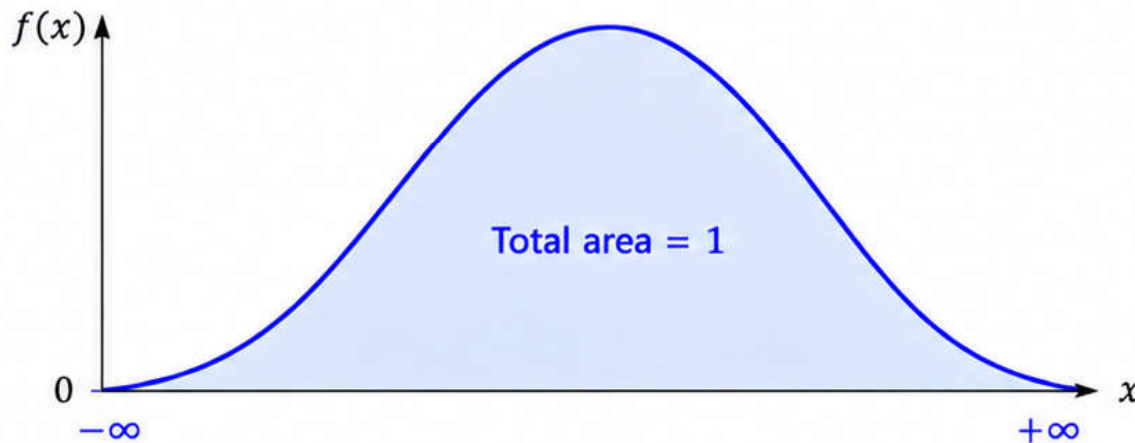
Probability = **area** under the **probability density curve**.

Normalization of the PDF



For any **continuous random variable** X with PDF $f(x)$, the **total area** under the PDF over the entire real line equals 1:

$$\int_{-\infty}^{\infty} f(x) dx = 1$$



Important:

- $f(x)$ itself is **not** a probability.
- A probability is the **area** under $f(x)$ over an interval.
- Values of $f(x)$ can be greater than 1. Only **areas** represent probabilities.



Remember: The **PDF** describes how probability is distributed, but **probability corresponds** to an **area**.



This helps avoid a common student **misunderstanding**.

Mean, median and mode

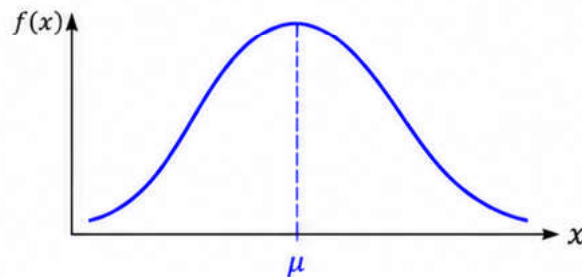


Three important location parameters describe the “center” of a distribution, each in a slightly different way.



1. Mean

Probability-weighted center.

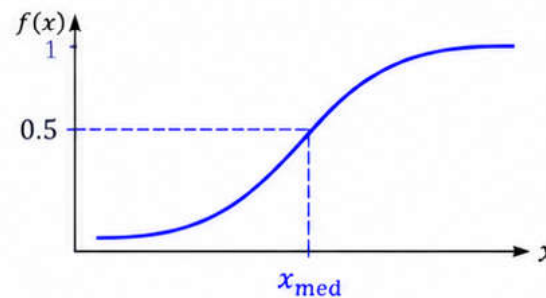


$$\begin{aligned}\mu &= E[X] = \sum_j x_j p_j && \text{(discrete)} \\ &= \int_{-\infty}^{\infty} x f(x) dx && \text{(continuous)}\end{aligned}$$



2. Median

$$P(X < x_{\text{med}}) = P(X > x_{\text{med}}) = \frac{1}{2}$$

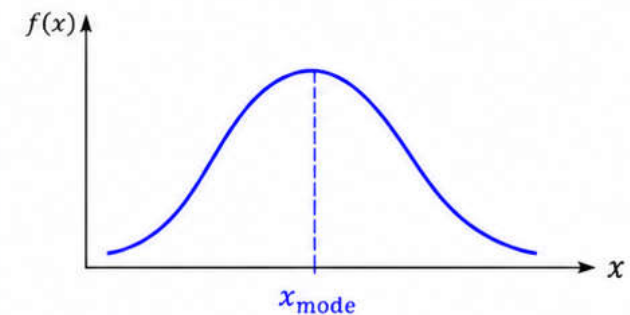


The **median** divides the distribution into two equally probable halves.



3. Mode

Most probable value.
The value where the PDF is maximum.



$$\text{Mode} = \arg \max_x f(x)$$



For symmetric distributions: **mean = median = mode**.

For skewed distributions: **mean $\neq$ median $\neq$ mode**.

26 Mean, Median and Mode: Three Ways to Describe “Typical”

1 / 3

1 MEAN (μ) – the average

The mean is the probability-weighted average.

Discrete: $\mu = \sum_i x_i p_i$

Continuous: $\mu = \int_{-\infty}^{+\infty} x f(x) dx$

Rare large values pull the mean to the right.

2 MEDIAN (x_{median}) – the 50% point

The median divides the distribution into two equal probability regions.

$$P(X \leq x_{\text{median}}) = 0.5$$

50% of the probability is below the median and 50% above it.

3 MODE (x_{mode}) – the most probable value

The mode is the value where the probability (or probability density) is greatest (the highest peak).

$$x_{\text{mode}}: f(x) \text{ is maximum}$$



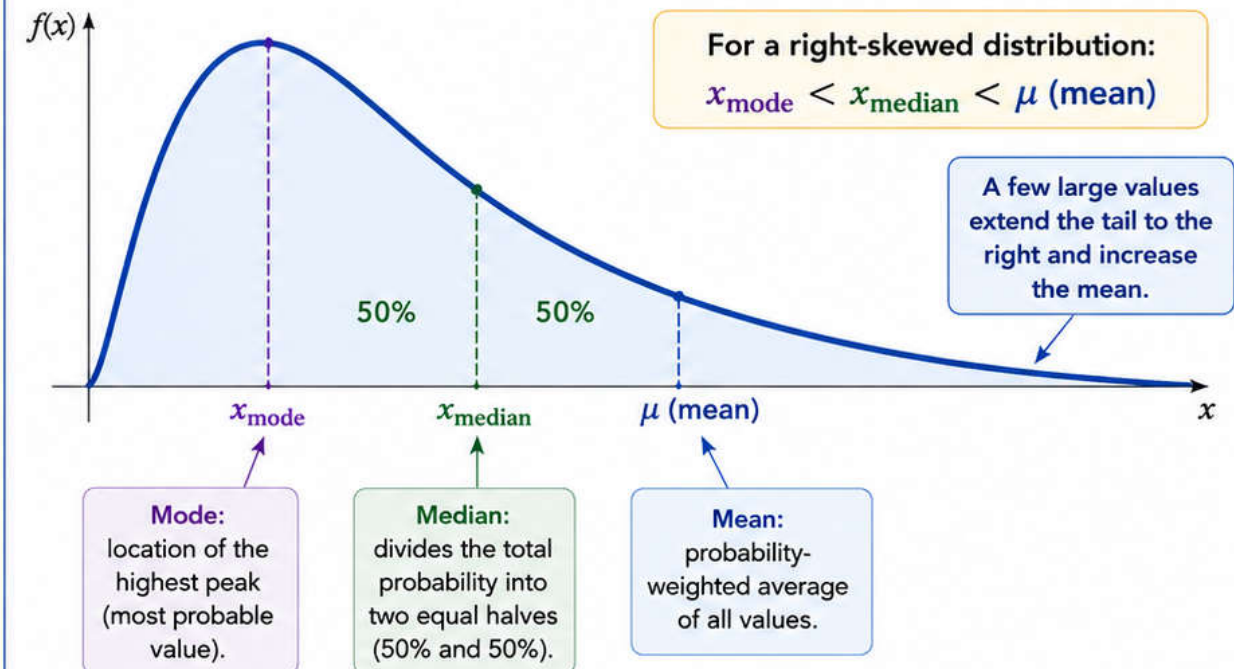
Mean, median and mode answer different questions. There is no universally “best” measure of a typical value.



Why it matters in practice

- The mean is sensitive to extreme values.
- The median is robust and less affected by outliers.
- The mode tells us where the population is most concentrated.
- Which one is most useful depends on the physical problem.

Right-skewed (asymmetric) distribution



Example in particle science

In aerosol measurements we may report:

- Number median diameter ($D_{N,50}$)
- Number mode diameter ($D_{N,\text{mode}}$)
- Number mean diameter ($\bar{D}_N$)

These can be very different when the size distribution is skewed!

Next slide: How this looks for symmetric and asymmetric distributions.



What is plotted?

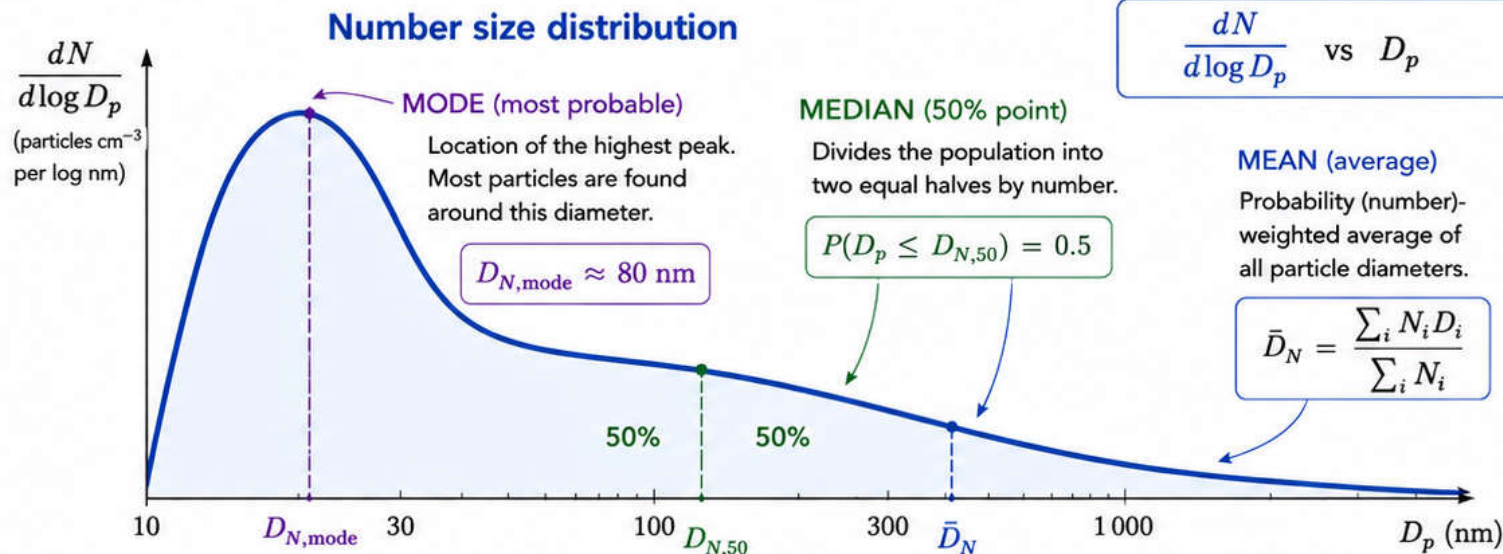
This is a particle number size distribution. It shows how many particles we have in each size range.

$$\frac{dN}{d \log D_p} \quad (\text{particles cm}^{-3} \text{ per log nm})$$

Example aerosol

- Many small particles (~30–100 nm)
- Fewer medium particles (~100–300 nm)
- Very few large particles (>300 nm)

→ The distribution is right-skewed.



Key message: For a skewed distribution, mean > median > mode.

All three are valid, but they represent different aspects of the same aerosol.

Important: These are **number-based quantities**. They describe 50% of the particles by number, not by mass or volume.

Interpretation in practice

Mode ($D_{N,mode}$) → where the particle population is most concentrated by number.

Median ($D_{N,50}$) → half of the particles are smaller, half are larger.

Mean ($\bar{D}_N$) → pulled toward the larger sizes by the long tail.



Always state clearly which weighting is used (number, surface, volume/mass)!

Typical size ranges (example values)

$D_{N,mode}$	~ 50 – 100 nm	Most numerous size range
$D_{N,50}$	~ 100 – 200 nm	Half the particles smaller, half larger
$\bar{D}_N$	~ 200 – 400 nm	Average shifted to larger sizes

Exact values depend on the aerosol source and conditions.

The Same Aerosol Can Have Different "Typical" Sizes

Which particles dominate depends on how we "weight" them.

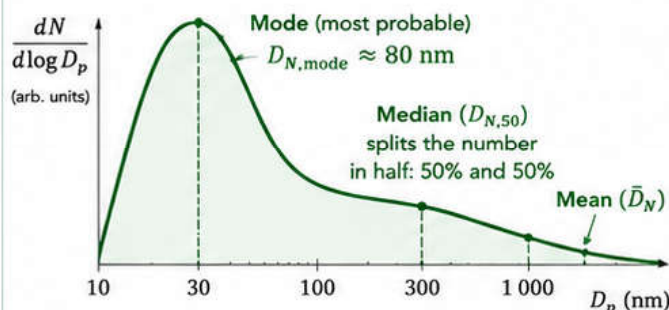


The same particle population → different distributions → different mean, median and mode.

A Number-weighted distribution

$$\frac{dN}{d \log D_p} \quad (\text{particles cm}^{-3} \text{ per log nm})$$

Every particle counts equally (weight = 1).
Small particles dominate because they are far more numerous.



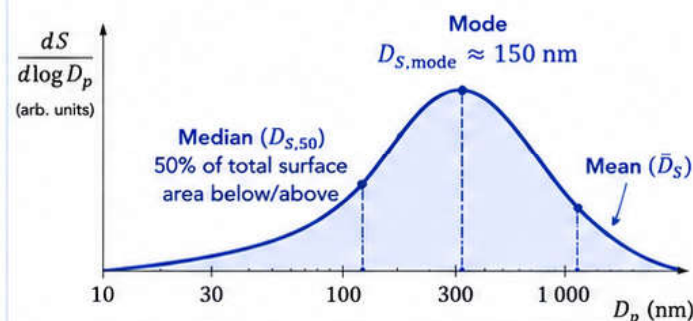
Typical order for right-skewed distributions:

$$D_{N,mode} < D_{N,50} < \bar{D}_N$$

B Surface-area-weighted distribution

$$\frac{dS}{d \log D_p} \propto D_p^2 \frac{dN}{d \log D_p} \quad (\text{surface cm}^2 \text{ cm}^{-3} \text{ per log nm})$$

Each particle is weighted by its surface area ($\propto D_p^2$).
Larger particles contribute much more to total surface area.



Typical order:

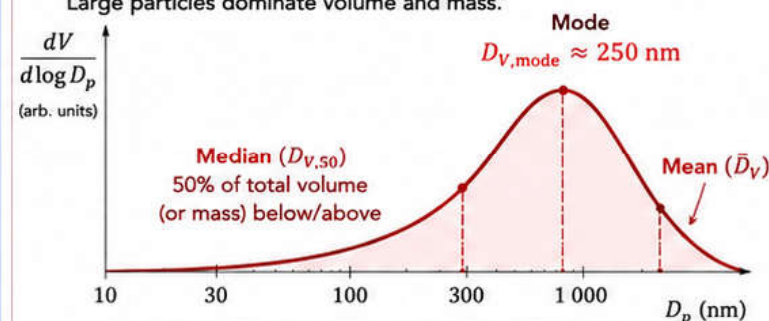
$$D_{S,mode} < D_{S,50} < \bar{D}_S$$

C Volume / Mass-weighted distribution

$$\frac{dV}{d \log D_p} \propto D_p^3 \frac{dN}{d \log D_p} \quad (\text{volume cm}^3 \text{ cm}^{-3} \text{ per log nm})$$

(mass if density ρ is constant: $dM/d \log D_p \propto D_p^3 dN/d \log D_p$)

Each particle is weighted by its volume ($\propto D_p^3$).
Large particles dominate volume and mass.



Typical order:

$$D_{V,mode} < D_{V,50} < \bar{D}_V$$

Number $\xrightarrow{\text{weight by } D_p^2}$ Surface area $\xrightarrow{\text{weight by } D_p^3}$ Volume / Mass



As we weight more strongly by size, the distribution shifts to larger diameters.



Key takeaways

- The same aerosol does not have one unique "characteristic particle size". It depends on how the particle population is weighted.
- Small particles may dominate number concentration, while larger particles can dominate surface area and especially volume or mass.
- Always specify what you are reporting: number, surface, or mass (volume) based!



Why it matters

- Number** → important for counting, collisions, nucleation
- Surface area** → important for chemical reactions, catalysis, condensation
- Volume / Mass** → important for mass loading, radiative effects, deposition, health impacts



Always ask:

- Mean of what?
- Median of what?
- Mode of what?

The answer completely changes the interpretation!



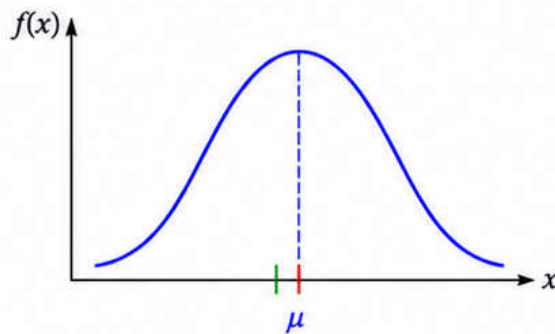
Symmetric vs. asymmetric distributions



Distributions can be **symmetric**, **asymmetric** (skewed), or even **bimodal**.
The **mean**, **median** and **mode** behave differently in each case.

1. Symmetric distribution

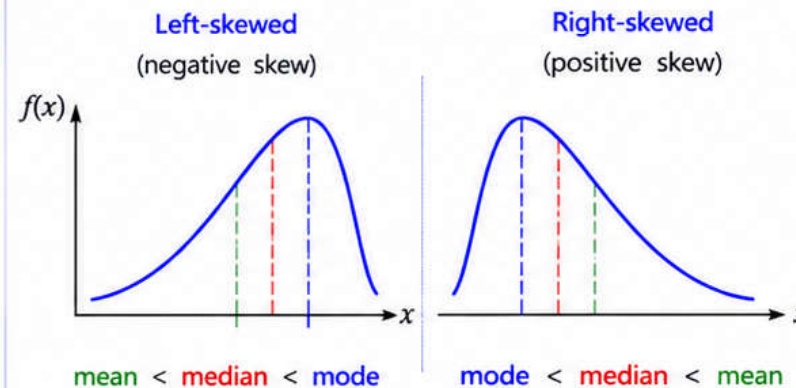
Perfect mirror image around the center.



mean = **median** = **mode**
(all equal)

2. Skewed (asymmetric) distribution

Not symmetric; has a longer tail on one side.

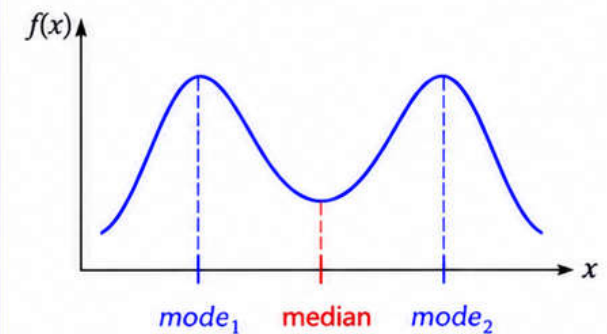


mean < **median** < **mode**

mode < **median** < **mean**

3. Bimodal distribution

Two distinct peaks (two modes).



Two **modes**; **median** lies between them.



Takeaway: The relative positions of **mean**, **median** and **mode** reveal important information about the **shape** and **character** of a distribution.

Variance



The **variance** measures the **average squared deviation** from the mean. It quantifies the **spread** of a distribution.

Definition

$$\sigma^2 = E[(X - \mu)^2]$$

where $\mu = E[X]$ is the mean.

Discrete random variable X

$$\sigma^2 = \sum_j (x_j - \mu)^2 p_j$$

where x_j are the possible values and $p_j = P(X = x_j)$.

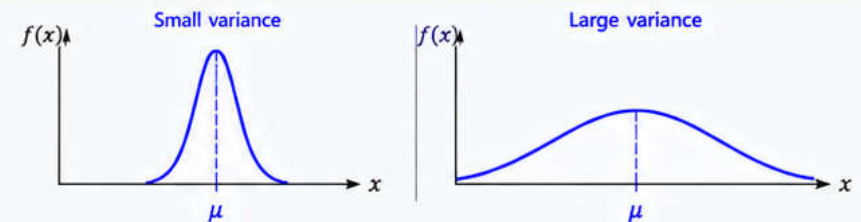
Continuous random variable X

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

where $f(x)$ is the probability density function (PDF).



Note: A larger variance means data are **more spread out**; a smaller variance means data are **more concentrated** around the mean.



Standard deviation

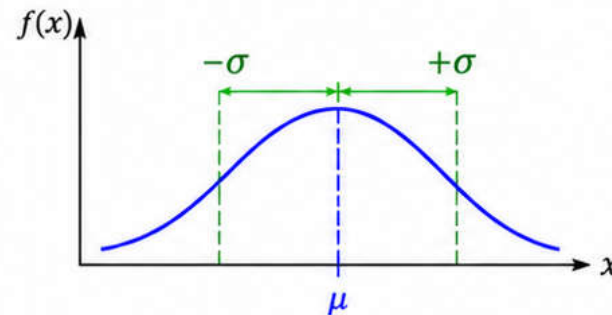


The **standard deviation** is the **square root of the variance**.

$$\sigma = \sqrt{\sigma^2}$$

Key interpretation:

The standard deviation describes the characteristic **spread** around the mean.



For many (bell-shaped) distributions:

$$P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 68\%$$



Important: The standard deviation has the **same dimension** as the variable.

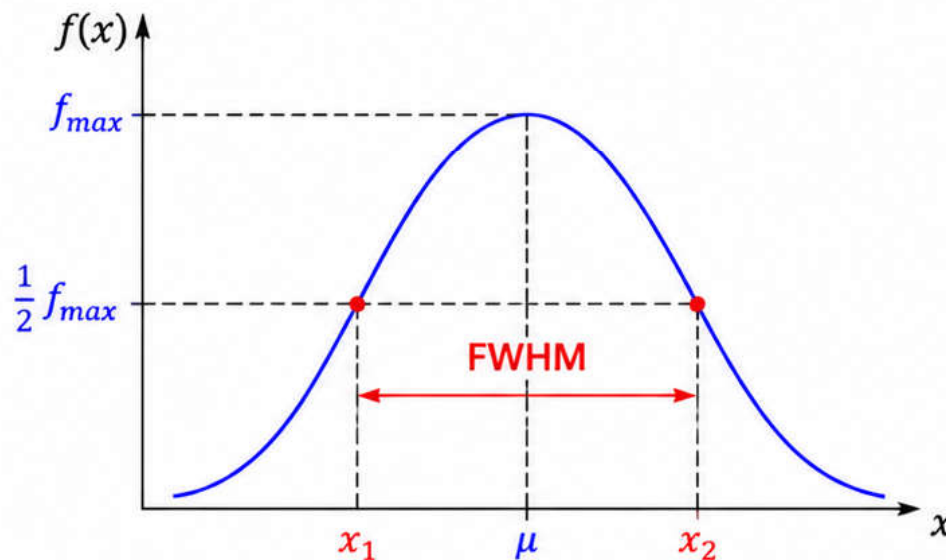
$$[\sigma] = [X]$$

Example: If X is measured in **meters (m)**, then σ is also in **meters (m)**.

Full width at half maximum (FWHM)



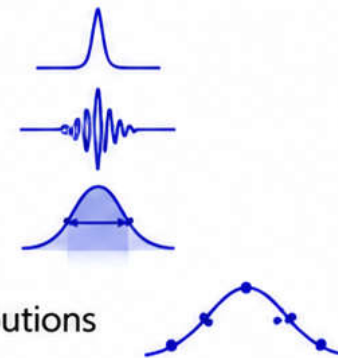
The **full width at half maximum (FWHM)** is the width of a distribution measured at **50%** of its maximum (peak) value.



$$\text{FWHM} = x_2 - x_1$$

Used frequently for:

- spectral lines
- laser pulses
- beam profiles
- experimental distributions

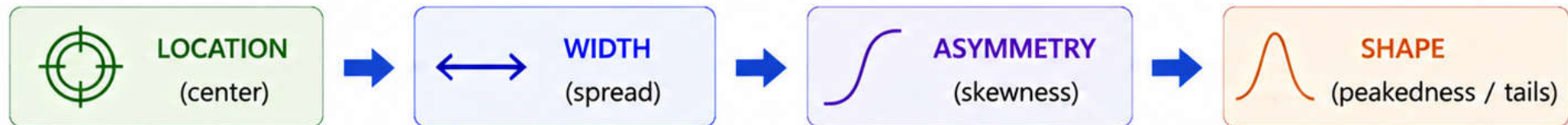


For a Gaussian distribution: $\text{FWHM} = 2\sqrt{2\ln 2} \sigma \approx 2.355 \sigma$

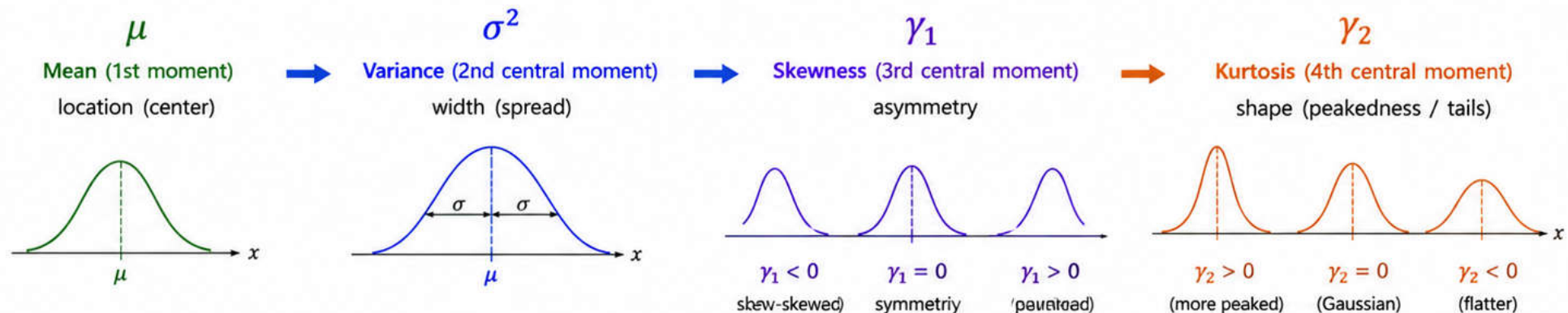
Why introduce moments?



Moments provide a **systematic mathematical description** of distributions. They allow us to characterize a distribution step by step:



Low-order moments describe the main features of a distribution:



Takeaway: Moments summarize increasingly detailed information about a distribution: **location** → **width** → **asymmetry** → **shape**.

Moments



Moments are expected values of **powers** of the random variable. They form the building blocks for characterizing distributions.

The s-th raw moment

$$m_s = E[X^s]$$

Discrete:

$$m_s = \sum_j x_j^s p_j$$

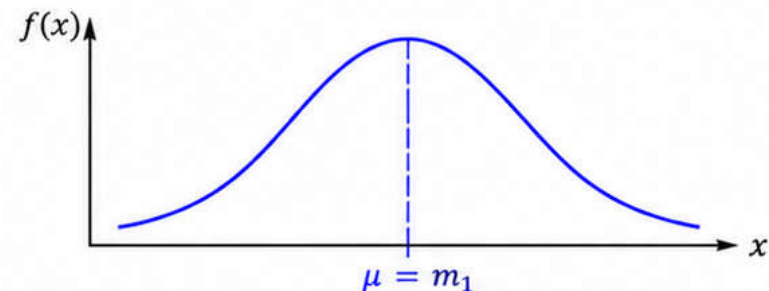
Continuous:

$$m_s = \int_{-\infty}^{\infty} x^s f(x) dx$$

First moment

$$m_1 = E[X] = \mu$$

- The first raw moment is the expected value (**mean**).
- It describes the **location** (center) of the distribution.



Note: Higher-order raw moments ($m_2, m_3, m_4, \dots$) capture increasingly detailed aspects of the distribution.

Moments (continued)

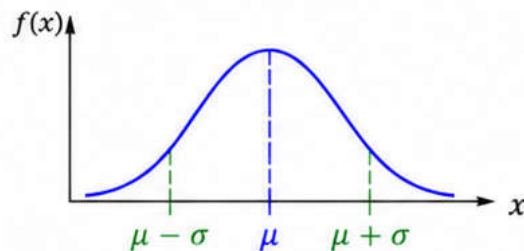


Beyond the first moment, higher moments describe **width**, **asymmetry**, and **shape**. These lead to important **central** (or **cumulants**) that isolate these properties.

Second central moment (variance)

$$\mu_2 = E[(X - \mu)^2] = \sigma^2$$

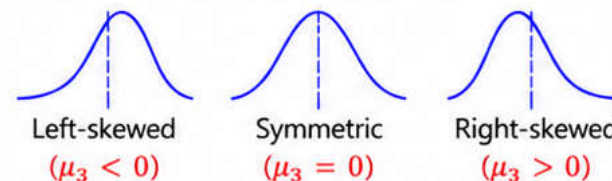
- Measures **spread** (width).
- $\sigma = \sqrt{\mu_2}$ is the standard deviation.



Third central moment (skewness numerator)

$$\mu_3 = E[(X - \mu)^3]$$

- Measures **asymmetry**.
- $\mu_3 > 0$: right-skewed (tail to the right)
- $\mu_3 < 0$: left-skewed (tail to the left)
- $\mu_3 = 0$: symmetric



Fourth central moment (kurtosis numerator)

$$\mu_4 = E[(X - \mu)^4]$$

- Measures **tailedness / peakedness** (shape).
- Compared to normal distribution (kurtosis):

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$$

- $\gamma_2 > 0$: leptokurtic (more peaked, heavy tails)
- $\gamma_2 = 0$: mesokurtic (normal)
- $\gamma_2 < 0$: platykurtic (flatter, light tails)



Note: Central moments are invariant to shifts of origin (depend on $X - \mu$), so they describe **shape** independently of **location**. In practice, we often use **standardized** moments (e.g., skewness $\gamma_1 = \mu_3/\sigma^3$ and kurtosis $\gamma_2 = \mu_4/\sigma^4 - 3$).

Moments and the shape of distributions



Moments summarize key aspects of a distribution. Together, they determine many features of the **shape**.

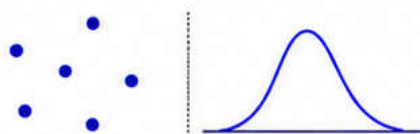
First moment (mean) $\mu = E[X]$	- Determines the location (center) of the distribution. - Changing μ shifts the distribution left or right without changing its shape.	$\mu_1 < \mu_2 < \mu_3$
Second central moment (variance) $\sigma^2 = E[(X - \mu)^2]$	- Measures the spread (width) of the distribution. - Larger $\sigma^2 \rightarrow$ wider distribution. - Smaller $\sigma^2 \rightarrow$ narrower distribution.	$\sigma_1^2 < \sigma_2^2 < \sigma_3^2$
Third central moment (skewness) $\mu_3 = E[(X - \mu)^3]$	- Determines the asymmetry of the distribution. $\mu_3 > 0$: right-skewed (long right tail) $\mu_3 = 0$: symmetric $\mu_3 < 0$: left-skewed (long left tail)	$\mu_3 > 0$ $\mu_3 = 0$ $\mu_3 < 0$
Fourth central moment (kurtosis) $\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$	- Determines the peakedness (tail behavior). $\gamma_2 > 0$: leptokurtic (more peaked, heavy tails) $\gamma_2 = 0$: mesokurtic (normal) $\gamma_2 < 0$: platykurtic (flatter, light tails)	$\gamma_2 > 0$ $\gamma_2 = 0$ $\gamma_2 < 0$



Summary: μ sets the **location**, σ^2 sets the **width**, μ_3 sets the **asymmetry**, and γ_2 sets the **peakedness**.

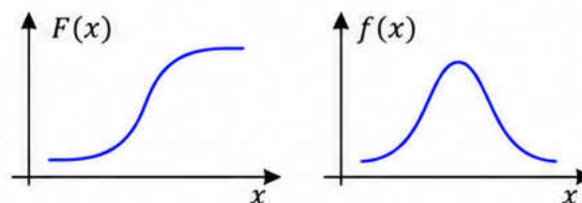
Summary

1 Random variables



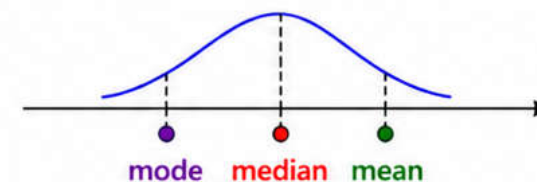
Discrete / continuous

2 Distribution functions



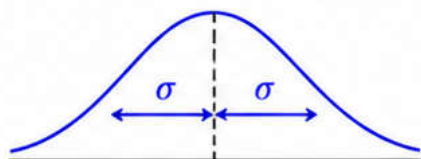
CDF / PDF

3 Location



Mean / median / mode

4 Dispersion



Variance / standard deviation / FWHM

5 Shape



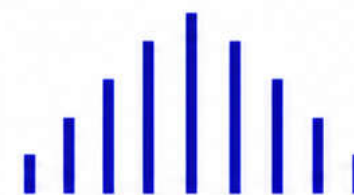
skewness

kurtosis

moments

Moments / skewness / kurtosis

6 Discrete model



Binomial distribution

Take-home messages

- 1 A probability distribution describes both possible values and their **likelihood**.
- 2 Discrete variables assign probabilities to **individual outcomes**; continuous variables assign probabilities to **intervals**.
- 3 The **CDF** and **PDF** provide **complementary** descriptions of continuous distributions.
- 4 Mean, **variance** and **standard deviation** describe location and spread.
- 5 Higher-order **moments** describe **asymmetry** and **shape**.
- 6 The **binomial distribution** describes the number of successes in independent repeated trials.



Next lecture → **Probability Distributions II**