# FOUNDATION YEAR 2025-2026

## COURSE - DATA ANALYSIS

Lesson 8

Niccolò Maffezzoli, Ca' Foscari University of Venice

February 5, 2026

# Remember: Poisson, Binomial, Gaussian formulas

- $P(k,p,n) = p^k(1-p)^{n-k}\frac{n!}{k!(n-k)!}$ *Binomial*
  - Fixed number of independent trials and each trial can only have one of two possible outcomes (usually success or failure). $\mu = np$, $\sigma = \sqrt{np(1-p)}$
- $P(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$ *Poisson*
  - The Poisson distribution is used to model the number of times an event occurs in a fixed interval of time or space, when the events occur independently and at a constant average rate $\lambda$; $\sigma = \sqrt{\lambda}$.
- $P(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ *Gaussian*

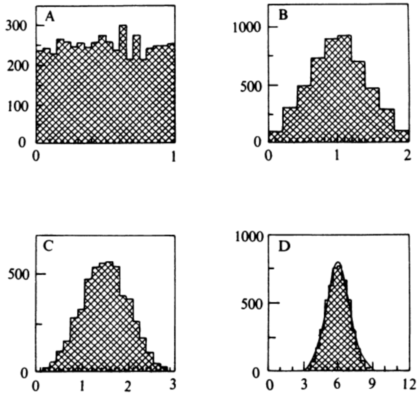
# Central Limit Theorem (CLT) 1/2

If you take the sum X of N independent variables, $x_i$ where i = 1,2,3,... ,N, each taken from a distribution of mean $\mu_i$ and variance $\sigma_i^2$, the distribution for X (the sum):

- $\mu_X = \sum \mu_i$
- $\sigma_X^2 = \sum \sigma_i^2$
- becomes Gaussian as $N \to \infty$ (as the sample size gets bigger)

# Central Limit Theorem (CLT) 2/2

- A. This shows a histogram of 5000 numbers taken at random from a uniform distribution between 0 and 1. It has mean 1/2 and variance 1/12, and is essentially flat -certainly not at all Gaussian.
- B. Here are another 5000 numbers, each the sum of a pair of random numbers of the type shown in A, i.e. $X = x_1 + x_2$. The distribution is triangular, peaking at 1.0.
- C. If three uniformly distributed random numbers are added together, the peak is at 1.5, and the form becomes curved.
- D. When the number of uniformly distributed random numbers increases to 12 at a time, the histogram appears to the eye just like a true Gaussian. The Gaussian of mean 6.0 and $\sigma^2 = 12 \cdot 1/12 = 1.0$ is also shown, and the agreement is seen to be very good indeed.



# Error Propagation

In general measurements are affected by uncertainties, $\sigma$. When different variables are accounted for in a formula (e.g. x or y), uncertainties are propagated ($\sigma_x$, $\sigma_y$). In general, if $z = z(x, y, z, w, ...)$:

$$\sigma_z = \sqrt{\sum_{x_i} \left(\frac{\partial f}{\partial x_i}\right)^2 \cdot \sigma_{x_i}^2} \tag{1}$$

We will just see the 5 simplest cases:

- $z = x + y$
- $z = x - y$
- $z = x \cdot y$
- $z = \frac{x}{y}$
- $z = A \cdot x$

## Addition: z = x + y or Subtraction: z = x - y

If $z = x + y$, or $z = x - y$ and both x and y are affected by uncertainty, i.e. you have $\sigma_x$ and $\sigma_y$, what is $\sigma_z$ ?

$$\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2} \tag{2}$$

## Example

A plane moves at $v_{p1} = (800 \pm 60)$ km/h. At some point it finds head wind that slows it down. The speed of the wind is $v_w = (100 \pm 20)$ km/h. What is the final speed of the plane $v_{p2}$?

- $v_{p2} = v_{p1} - v_w = 800 - 100 = 700$ km/h.
- $\sigma_{vp2} = \sqrt{\sigma_{vp1}^2 + \sigma_{vw}^2} = \sqrt{60^2 + 20^2} = \sqrt{3600 + 400} = \sqrt{4000} = 63.2$ km/h.
- The final speed of the plane is $v_{p2} = (700 \pm 60)$ km/h $= (7 \pm 0.6) \cdot 10^2$ km/h. (remember one digit on the uncertainty!)

## Multiplication: $z = x \cdot y$ or Division: $z = x/y$

If $z = x \cdot y$, or $z = \frac{x}{y}$ and both x and y are affected by uncertainty, i.e. you have $\sigma_x$ and $\sigma_y$, what is $\sigma_z$ ?

$$\frac{\sigma_z}{|z|} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2} \tag{3}$$

## Multiplication with a constant: $z = A \cdot x$

If $z = A \cdot x$, and only x is affected by uncertainty $\sigma_x$, what is $\sigma_z$ ?

$$\sigma_z = |A| \cdot \sigma_x \tag{4}$$

## Example

If the speed of a projectile is given as $v = (200 \pm 10)$ m/s, calculate the distance (and uncertainty) of the projectile in 6 seconds.

$$s = v \cdot t = 200 \cdot 6 = 1200 \ m; \ \sigma_s = t\sigma_v = 6 \cdot 10 = 60 \ m$$

The distance travelled in 6 seconds is $s = (1200 \pm 60)m = (120 \pm 6) \cdot 10 \ m$

# Exercise 1

Two positions are measured on a track:

$$x = 125.40 \pm 0.10 \ m$$

$$y = 124.80 \pm 0.10 \ m$$

Find the distance $z = x - y$ and its uncertainty.

# Exercise 1

Two positions are measured on a track:

$$x = 125.40 \pm 0.10 \ m$$

$$y = 124.80 \pm 0.10 \ m$$

Find the distance $z = x - y$ and its uncertainty.
**SOLUTION**

$$z = 125.40 - 124.80 = 0.60 \ m$$
$$\sigma_z = \sqrt{0.1^2 + 0.1^2} = 0.14 \ m$$
$$\rightarrow z = 0.60 \pm 0.14 \ m$$

Subtracting close numbers gives large relative uncertainty.

# Exercise 2

A metal plate has:

$$x = 30.0 \pm 0.1 \ cm$$

$$y = 20.0 \pm 0.5 \ cm$$

Compute the area and its uncertainty. What variable dominates the uncertainty ?

# Exercise 2

A metal plate has:

$$x = 30.0 \pm 0.1 \ cm$$

$$y = 20.0 \pm 0.5 \ cm$$

Compute the area and its uncertainty. What variable dominates the uncertainty ?
**SOLUTION**

$$A = xy = 30.0 \cdot 20.0 = 600 \ cm^2$$

$$\frac{\sigma_A}{A} = \sqrt{(\frac{0.1}{30.0})^2 + (\frac{0.5}{20.0})^2} \quad (5)$$

$$= \sqrt{0.0033^2 + 0.025^2} = 0.0252 \quad (6)$$

$$\sigma_A = 600 \cdot 0.0252 \simeq 15$$

$$\rightarrow A = 600 \pm 15 \ cm^2$$

The dominant uncertainty comes from y, not x.

# Exercise 3

A student measures the electrical power dissipated by a device using the formula

$$P = V \cdot I,$$

where $V$ is the voltage and $I$ is the current.
The measured values are:

$$V = (12.0 \pm 0.2)\,\text{V}, \qquad I = (1.50 \pm 0.05)\,\text{A}.$$

1. Calculate the value of the power $P$.
2. Using error propagation, calculate the uncertainty $\sigma_P$.
3. Write the final result in the form $P \pm \sigma_P$.

# Exercise 3

A student measures the electrical power dissipated by a device using the formula

$$P = V \cdot I,$$

where $V$ is the voltage and $I$ is the current.
The measured values are:

$$V = (12.0 \pm 0.2)\,\mathrm{V}, \qquad I = (1.50 \pm 0.05)\,\mathrm{A}.$$

1. Calculate the value of the power $P$.
2. Using error propagation, calculate the uncertainty $\sigma_P$.
3. Write the final result in the form $P \pm \sigma_P$.

**SOLUTION**

The power is $P = VI = (12.0)(1.50) = 18.0\,\mathrm{W}$.
For a product, the relative uncertainty is

$$\left(\frac{\sigma_P}{P}\right)^2 = \left(\frac{\sigma_V}{V}\right)^2 + \left(\frac{\sigma_I}{I}\right)^2.$$

Substituting the values:

$$\left(\frac{\sigma_P}{18.0}\right)^2 = \left(\frac{0.2}{12.0}\right)^2 + \left(\frac{0.05}{1.50}\right)^2 = (0.0167)^2 + (0.0333)^2.$$

$$\left(\frac{\sigma_P}{18.0}\right)^2 = 0.00139 \quad \Rightarrow \quad \frac{\sigma_P}{18.0} = 0.037 \Rightarrow \sigma_P = (18.0)(0.037) \approx 0.67\,\mathrm{W}$$

$$\boxed{P = (18.0 \pm 0.7)\,\mathrm{W}}$$