The Normal Distribution

Antonio Aguirre

University of California, Santa Cruz
Winter 2025

Introduction

The **Normal Distribution** is one of the most fundamental probability distributions, widely used in statistics, natural sciences, and engineering. It models continuous data where values cluster around a central mean, following a symmetric bell-shaped curve.

Probability Density Function (PDF)

A random variable X follows a normal distribution with mean μ and variance σ^2 if its probability density function (PDF) is:

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

Key properties:

- **Symmetric** around μ.
- **Unimodal**, meaning it has a single peak at μ.
- The **spread** of the distribution is determined by σ^2.
- The **standard normal distribution** is a special case with $\mu = 0$ and $\sigma^2 = 1$.

—

1. Properties and Operations for Normal Probabilities

Symmetry

For any normal variable $X \sim \mathcal{N}(\mu, \sigma^2)$:

$$X \leq \mu - a = X \geq \mu + a.$$

Complement Rule

$$X \geq x = 1 - X \leq x.$$

Probability Between Two Values

$$a \leq X \leq b = X \leq b - X \leq a.$$

Standardization

A normal variable $X \sim \mathcal{N}(\mu, \sigma^2)$ can be transformed into a standard normal variable $Z \sim \mathcal{N}(0, 1)$ using:

$$Z = \frac{X - \mu}{\sigma}.$$

—

2. Linear Transformations of a Normal Distribution

If $X \sim \mathcal{N}(\mu, \sigma^2)$, and we define a new variable:

$$Y = aX + b,$$

where $a \neq 0$, then Y follows:

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

Key Result: Linear Transformations

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b$ is also normally distributed with:

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

—

3. Linear Combination of Independent Normals

If $X_1, X_2, \ldots, X_n$ are independent normal variables, where $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, and we define:

$$Y = \sum_{i=1}^{n} a_i X_i,$$

then Y follows:

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

—

4. Sample Mean Distribution

If $X_1, X_2, \ldots, X_n$ are i.i.d. random variables:

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

then the **sample mean** is defined as:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Since the X_i's are normal, the sample mean $\bar{X}_n$ follows:

$$\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

—

5. Determining Minimum Sample Size

We determine the minimum sample size n such that:

$$P(|\bar{X}_n - \mu| \leq E) \geq p.$$

Solution Steps

1. Rewrite the event:

$$P(|\bar{X}_n - \mu| \leq E) = 2\Phi(z) - 1,$$

where:

$$z = \frac{E\sqrt{n}}{\sigma}.$$

2. Solve for n:

$$n = \left(\frac{z_p \sigma}{E}\right)^2.$$

Key Formula: Minimum Sample Size

$$n = \left(\frac{z_p \sigma}{E}\right)^2, \quad \text{where } z_p = \Phi^{-1}\left(\frac{p+1}{2}\right).$$