# Normal Distribution

Complete Reference Guide · statisticsfundamentals.com

FREE DOWNLOAD

**DEFINITION — OPTIMISED FOR QUICK REFERENCE**

Normal distribution (Gaussian distribution / bell curve) is a continuous probability distribution symmetric about its mean $\mu$, where data clusters near the centre and becomes rarer toward the extremes. Defined by mean $\mu$ (location) and standard deviation $\sigma$ (spread). Total area under the curve = 1. Mean = Median = Mode.

## Three-Format Definition

**Plain English**

Most values cluster in the middle; fewer values appear at the extremes. Plot them and you get a symmetric bell-shaped curve.

**Statistical**

A symmetric, unimodal continuous probability distribution where Mean = Median = Mode, with tails extending infinitely in both directions.

**Mathematical**

$f(x; \mu, \sigma) = (1 / \sigma\sqrt{(2\pi)}) \cdot \exp(-(x-\mu)^2 / 2\sigma^2)$ for $x \in (-\infty, +\infty)$

## Normal Distribution Bell Curve

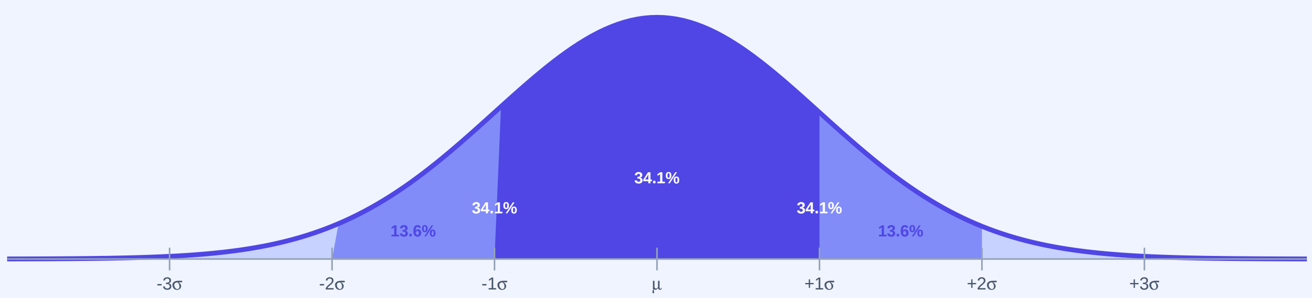


Figure 1: Bell curve showing 68-95-99.7% regions. Shading: 1σ (dark), 2σ (medium), 3σ (light).

# Normal Distribution Formula & Properties

FREE DOWNLOAD

PROBABILITY DENSITY FUNCTION (PDF)

$$f(x) = (1 / \sigma\sqrt{(2\pi)}) \times e^{-(x-\mu)^2 / 2\sigma^2}$$

$\mu$ = mean · $\sigma$ = standard deviation · $e \approx 2.71828$ · $\pi \approx 3.14159$

## Formula Symbol Breakdown

| Symbol | Name | Description |
|---|---|---|
| **x** | Observed value | The data point being evaluated |
| **$\mu$ (mu)** | Mean | Centre of the distribution; peak of the bell curve |
| **$\sigma$ (sigma)** | Std deviation | Controls the width/spread of the curve |
| **$\sigma^2$** | Variance | Square of $\sigma$; appears in the exponent |
| **e** | Euler's number | $\approx$2.71828; base of the natural logarithm |
| **$\pi$** | Pi | $\approx$3.14159; part of the normalisation constant |
| **$1/\sigma\sqrt{(2\pi)}$** | Normalisation | Ensures total area under curve = 1 |

## Z-Score Formula (Standardisation)

$$Z = (X - \mu) / \sigma$$

X = observed value · $\mu$ = mean · $\sigma$ = standard deviation

CUMULATIVE DISTRIBUTION FUNCTION (CDF)

$$F(x) = P(X \leq x) = \Phi( (x-\mu) / \sigma )$$

## Key Properties of Normal Distribution

| Property | Value |
|---|---|
| **Skewness** | 0 (perfectly symmetric) |
| **Kurtosis (excess)** | 0 (total kurtosis = 3, mesokurtic) |
| **Mean = Median = Mode** | All equal $\mu$ exactly |
| **Support** | $(-\infty, +\infty)$ — infinite |
| **Total probability** | 1.0 (area under curve) |
| **Moment generating fn** | $M(t) = \exp(\mu t + \sigma^2 t^2/2)$ |

# 68-95-99.7 Empirical Rule & Z-Score Example

FREE DOWNLOAD

## The 68-95-99.7 Empirical Rule

| Range | % of Data | Z-Score Band | IQ Example ($\mu$=100, $\sigma$=15) | Exact % |
|---|---|---|---|---|
| $\mu \pm 1\sigma$ | ~68% | -1 to +1 | 85 to 115 | 68.27% |
| $\mu \pm 2\sigma$ | ~95% | -2 to +2 | 70 to 130 | 95.45% |
| $\mu \pm 3\sigma$ | ~99.7% | -3 to +3 | 55 to 145 | 99.73% |
| $\mu \pm 4\sigma$ | ~99.99% | -4 to +4 | 40 to 160 | 99.994% |
| $\mu \pm 6\sigma$ | Six Sigma | -6 to +6 | Virtually all | 99.9999998% |

## How to Use the Z-Table — Step by Step

1. **Identify your values**
   Note: observed value X, population mean $\mu$, standard deviation $\sigma$
2. **Calculate the Z-score**
   Apply $Z = (X - \mu) / \sigma$ to convert to the standard normal scale
3. **Look up the Z-table**
   Find the cumulative probability $P(Z < z)$ in the standard normal table
4. **Interpret the result**
   Table value = proportion below X. For $P(X > \text{value})$ subtract from 1.

WORKED EXAMPLE — IQ SCORES

**What % of people have IQ above 115? ($\mu = 100$, $\sigma = 15$)**

Step 1: $Z = (115 - 100) / 15 = 15/15 = 1.00$

Step 2: Look up Z = 1.00 in Z-table $\rightarrow$ $P(Z < 1.00) = 0.8413$

Step 3: $P(X > 115) = 1 - 0.8413 = 0.1587$

**Answer: 15.87% of people score above 115 on an IQ test.**

# Real-World Examples & Distribution Compari

FREE DOWNLOAD

## Real-World Examples of Normal Distribution

| Domain | Key Statistics & Details |
|---|---|
| **IQ Scores** | μ = 100, σ = 15 (Wechsler scale). About 68% score between 85–115. Only 2.3% above 130. |
| **Human Height** | US adult males: μ ≈ 70 in (178 cm), σ ≈ 3 in. Approximately normal. |
| **Birth Weight** | Full-term US newborns: μ = 3,400 g, σ = 500 g. 95% weigh 2,400–4,400 g. |
| **Standardised Tests** | SAT: μ = 1,010, σ ≈ 210. GRE Quantitative: μ = 150, σ = 8.9. |
| **Manufacturing (Six Sigma)** | ±6σ = 3.4 defects per million opportunities. Machine tolerances are normal. |
| **Measurement Error** | Repeated lab measurements produce normally distributed errors (basis of least squares). |

## Normal vs. Other Distributions

| Feature | Normal | T-Distribution | Log-Normal | Binomial (approx) |
|---|---|---|---|---|
| **Shape** | Bell / symmetric | Bell / heavier tails | Right-skewed | Bell (large n) |
| **Support** | (–∞, +∞) | (–∞, +∞) | (0, +∞) | {0,1,...,n} |
| **Mean** | μ | 0 (standard) | $\exp(\mu+\sigma^2/2)$ | np |
| **When to use** | Large n; known σ | Small n; unknown σ | Skewed positive data | np≥5 & n(1-p)≥5 |
| **Parameters** | μ, σ | Degrees of freedom | μ, σ of log(X) | n, p |

# Testing Normality, Misconceptions & Tools

FREE DOWNLOAD

## Statistical Tests for Normality

| Test | Best For | Null Hypothesis |
|---|---|---|
| **Shapiro-Wilk** | n < 2,000 | Data is normally distributed |
| **Kolmogorov-Smirnov** | Large samples | Data follows specified distribution |
| **Anderson-Darling** | General; tail-sensitive | Data is normally distributed |
| **D'Agostino-Pearson** | Skewness + kurtosis | Skewness=0 and Kurtosis=3 |
| **Jarque-Bera** | Large n; econometrics | Skewness=0 and Kurtosis=3 |

## 5 Common Misconceptions

■ **MYTH: "All data is normally distributed"**

✓ REALITY: Many real datasets are skewed (income, web traffic). Always verify.

■ **MYTH: "Normal means good or usual"**

✓ REALITY: "Normal" is mathematical — not a value judgement. Disease data can be normal.

■ **MYTH: "The tails reach zero"**

✓ REALITY: Tails extend to ±∞ and only approach zero asymptotically — never touch.

■ **MYTH: "Large samples are always normal"**

✓ REALITY: CLT applies to sample MEANS, not raw data. Skewed data stays skewed.

■ **MYTH: "Bell-shaped = normal distribution"**

✓ REALITY: t-distribution and Cauchy are also bell-shaped but are NOT normal.

## Quick Tool Reference

| Platform | Function | Purpose |
|---|---|---|
| Excel | `=NORM.DIST(x, mean, sd, TRUE)` | CDF: P(X ≤ x) |
| Excel | `=NORM.INV(prob, mean, sd)` | Inverse CDF |
| Python | `scipy.stats.norm.cdf(x, loc=μ, scale=σ)` | CDF probability |
| Python | `scipy.stats.norm.ppf(p, loc=μ, scale=σ)` | Inverse CDF (quantile) |
| R | `pnorm(q, mean=μ, sd=σ)` | CDF probability |
| R | `qnorm(p, mean=μ, sd=σ)` | Inverse CDF (quantile) |