Regents Park Publishers

Data Analytics

T1LM3

Excel Tutorials

Content

Axioms

Axioms

An **Axiom** is a mathematical statement that is **assumed** to be true.

- a self-evident truth that requires no proof
- a universally accepted principle or rule
- a proposition that is assumed without proof for the sake of studying the consequences that follow from it
- Formulas that we use are based on axioms

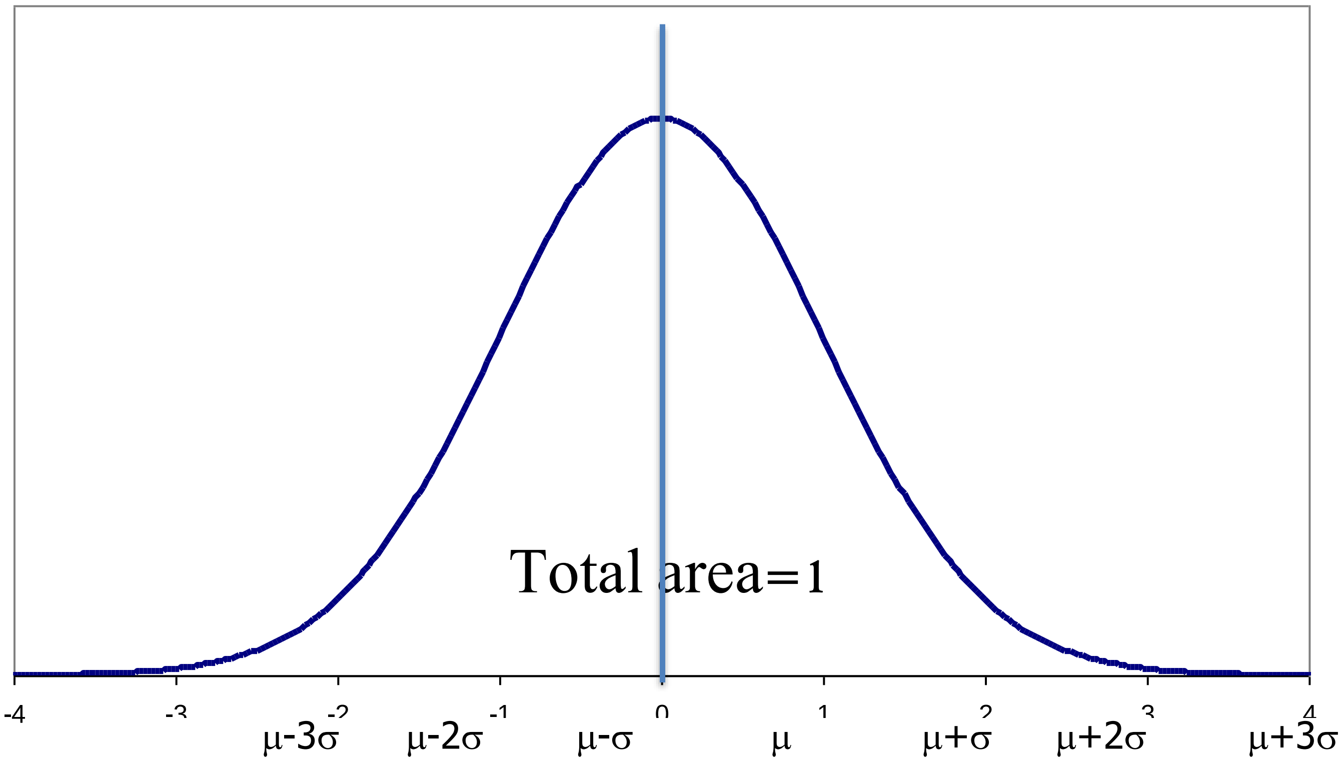
Normal Probability Distribution

Introduction

- The Normal (or Gaussian) distribution is probably the most used (and abused) distribution in statistics.

- Normal random variables are **continuous** (they can take any value on the real line) so the Normal distribution is an example of a **continuous probability distribution.**

The Normal Distribution



The graph of the normal distribution is called the normal (or bell) curve.

The Normal Distribution

- The mean, median, and mode are the same.
- The normal curve is symmetric about its mean.
- The total area under the normal curve is one.
- The normal curve approaches, but never touches, the x-axis.

Normal Distributions and Probability

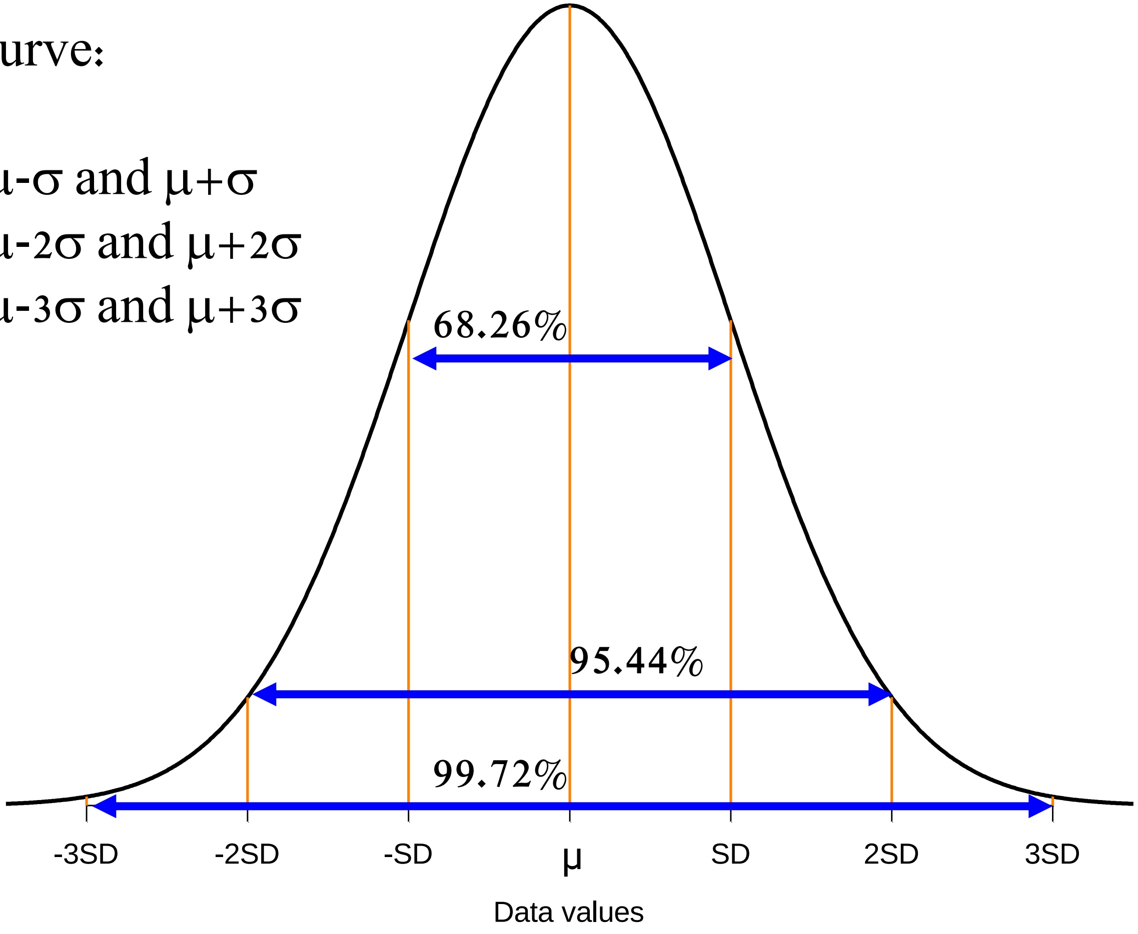
- We do not talk about the probability $P[X=x]$ for continuous random variables. Rather we talk about the probability that the random variable falls in an interval, i.e. $P[x_1 \le X \le x_2]$.

- $P[x_1 \le X \le x_2]$ can be determined by finding the area under the normal curve between x_1 and x_2.

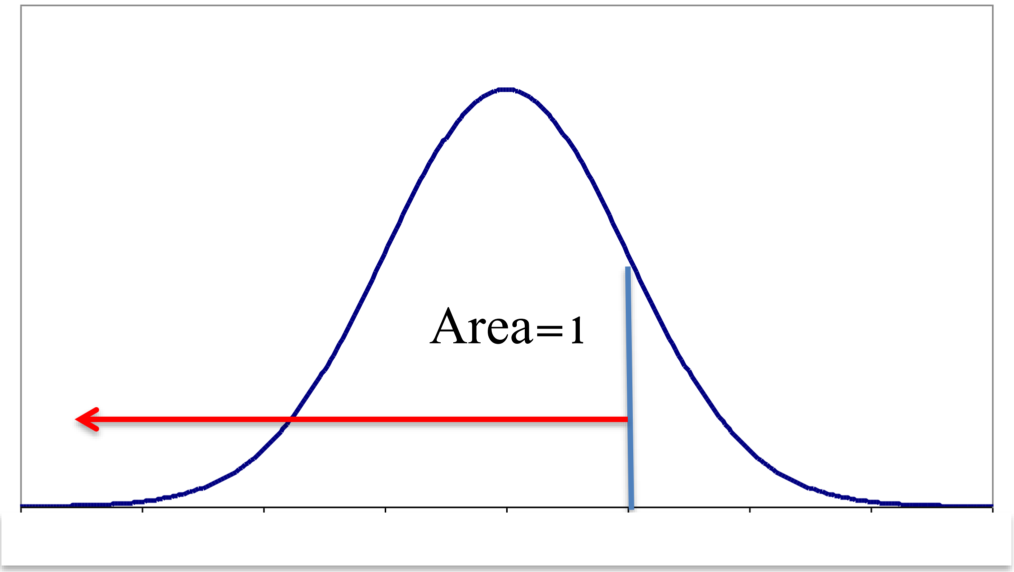
Normal Distributions and Probability

The areas under the curve:

68.26% lies between μ-σ and μ+σ
95.44% lies between μ-2σ and μ+2σ
99.72% lies between μ-3σ and μ+3σ

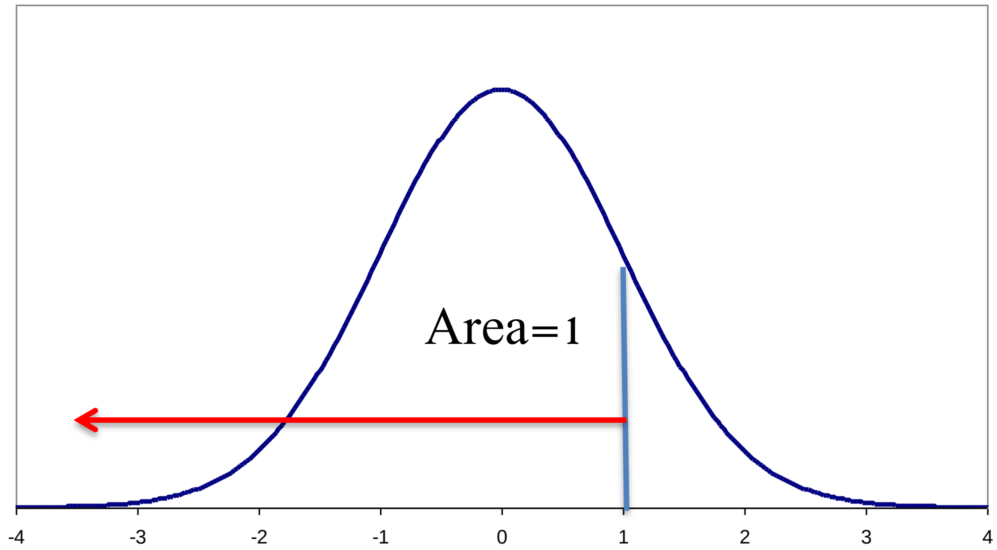


Excel Function



The area under the normal distribution
from x to -∞ can be computed using the EXCEL function
NORMDIST

Excel Function



The area under the **standard** normal distribution from x to $-\infty$ can be computed using the EXCEL function **NORMSDIST**

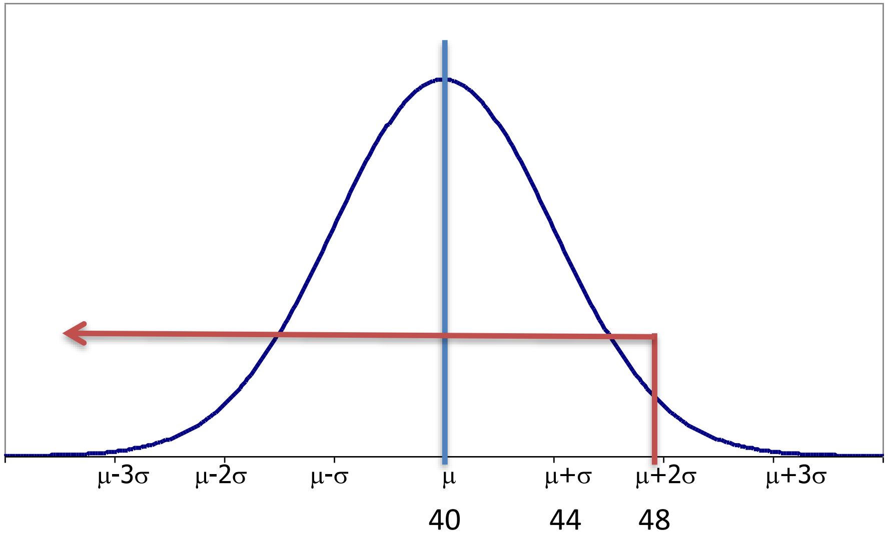
Examples 1-3

Example 1

- The mean length of a fish is 40cm and the standard deviation is 4 cm. What is the probability that the length of a randomly selected fish is less than 48cm?
- 48cm is two standard deviations above the mean so the area to the left of 48cm is 0.5+0.475 = **0.9772**

or

NORMDIST (48,40,4,1)= 0.9772



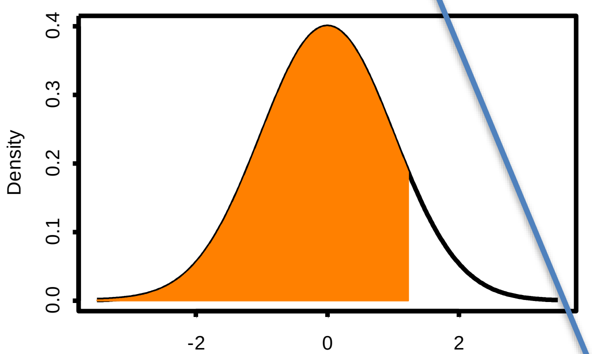
Example-2

- Find the area under the normal distribution curve between -0.12 and 1.23.
  - We use our previous approach:
    - $P[-0.12 \leq X \leq 1.23] = P[X \leq 1.23] - P[X \leq -0.12]$
  - In EXCEL:
    - NORMDIST(1.23,0,1,1)-NORMDIST(-0.12,0,1,1)

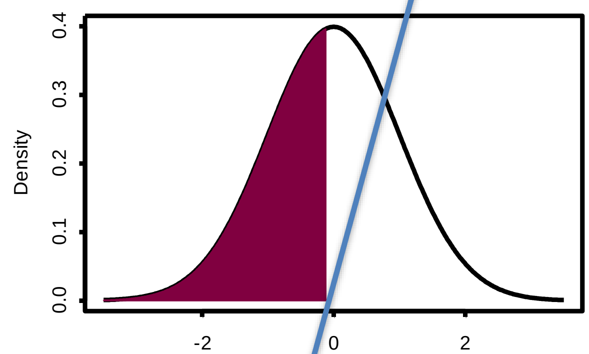
NORMDIST(x,μ,σ,1)

Example-3

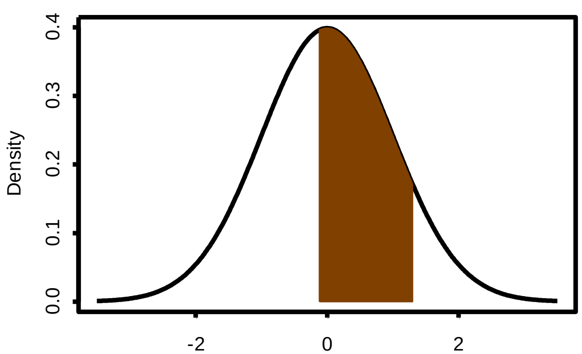
$P[X \le 1.23] =$ **0.8907**



$P[X \le -0.12] =$ **0.4522**



$P[X \le 1.23] - P[X \le -0.12] =$



0.8907-0.4522 = 0.4384

Standardized Normal Probability Distribution

The Standard Normal Distribution

The normal distribution with a mean of 0 and a standard deviation of 1 is called the

Standard Normal Distribution

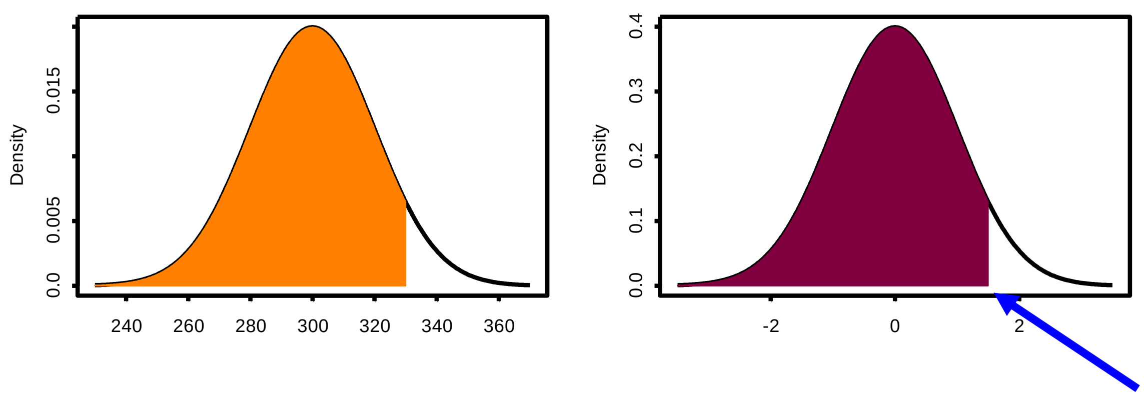
- The standard normal distribution and the z-score:

$$z = \frac{\text{value - mean}}{\text{standard deviation}} = \frac{x - \mu}{\sigma}$$

Standardization

- We can transform any normal distribution into a standard normal distribution by subtracting the mean and dividing by the standard deviation.

$\mu=300;\ \sigma=20;\ X=330$ $\mu=0;\ \sigma=1;\ Z=1.5$



Area=0.933 in both cases

$Z = (X-300)/20$

Standardization

- To find the probability that $X \leq Y$ if X is normally distributed with mean μ and standard deviation σ.
  - Compute the z-score: $z = (y - \mu)/\sigma$
  - Calculate the area under the normal curve between -∞ and z
  - We could calculate this area directly using the EXCEL function:

Standardize

Examples 1-3

Examples 1-3

- The average swimming speed of a fish population is 2 m.s. (standard deviation 0.5). You select a fish at random. What is the probability that:
  - Its swimming speed is less than 1 m.s.
  - Its swimming speed is greater than 2.5 m.s.
  - Its swimming speed is between 2 and 3 m.s.

Example -1

- The average swimming speed of a fish population is 2 m.s. (standard deviation 0.5).
  You select a fish at random.
  What is the probability that:

  - It's swimming speed is less than 1 m.s.
    $= P(z < (1-2)/.5) = P(z < -2) =$ **0.0228**

Examples-2

- The average swimming speed of a fish population is 2 m.s. (standard deviation 0.5). You select a fish at random. What is the probability that:
  - Is swimming speed is greater than 2.5 m.s.

$= P(z > (2.5-2)/.5) = P(z > 1) = 1 - P(z \leq 1)$

= 0.159

Example-3

- The average swimming speed of a fish population is 2 m.s. (standard deviation 0.5). You select a fish at random. What is the probability that:
  - Its swimming speed is between 2 and 3 m.s.

  - $= P((2-2)/.5 \leq z \leq (3-2)/0.5)$
    $= P(0 \leq z \leq 2) = P(z \leq 2) - P(z \leq 0)$

= 0.477

Regents Park Publishers

T1LM3

End