The Normal Distribution

What is it?

If you take a large sample of measurements from a continuous random variable in a population, a histogram of this data will look bell - shaped.

The probability distribution curve for these data can be approximated using a formula for the normal distribution.

Two pieces of information are required by the formula: the population mean and the population standard deviation.

Notation

μ Population mean

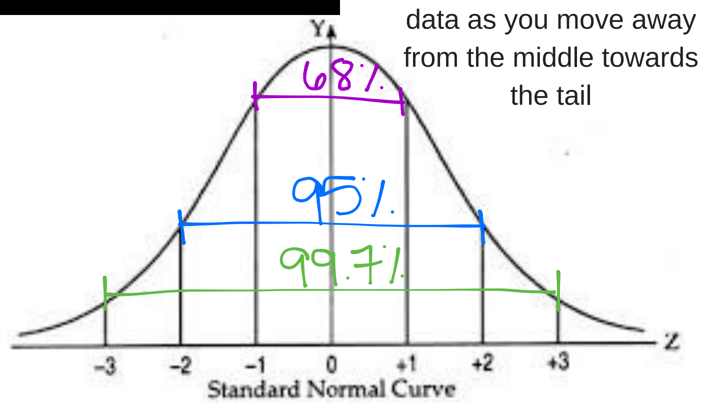
σ Population standard deviation

$\bar{x}$ Sample mean

s Sample standard deviation

Empirical Rule (68-95-99.7 Rule)

Most data is near the middle and there is less data as you move away from the middle towards the tail



Standard Normal Curve

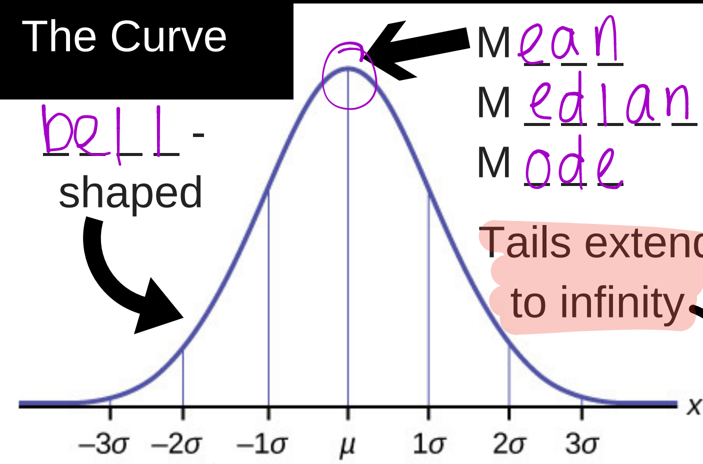
68 % of the data are within 1 standard deviation of the mean

95 % of the data are within 2 standard deviations of the mean

99.7 % of the data are within 3 standard deviations of the mean

The Curve

bell - shaped



Mean
Median
Mode

Tails extend to infinity

Symmetric about the mean

The shape is determined by two parameters

→ μ - where the middle is

→ σ - how wide and tall it is

We often don't know these parameters so we have to estimate them using $\bar{x}$ and s, ideally from a sample > 30

The Normal Distribution

The Area Under the Curve

The area under the curve represents the p robability of all possible outcomes. The total area under the curve is therefore equal to one(100 %)

We can work out the percentage of data that lie within a given range of values using:

1) The Empirical Rule
2) a z-table
3) a statistical software package calculator

Practice

On a standardized exam, the scores are normally distributed with a mean of 170 and a standard deviation of 20. Find the z-score of a person who scored 140 on the exam.

$$z = \frac{140-170}{20} = \frac{-30}{20} = \frac{-3}{2}$$

$z = -1.5$

the score of 140 is 1.5 standard deviations below the mean.

Z-Scores

- A numerical value that tells you how far a value (x) is from the mean. It measures distance in terms of standard deviation.
- For example, a z-score of 2 means that the value (x) is 2 standard deviations above the mean.
- Z-scores are used to calculate these percentages.
- A z-score (z) is calculated using:

$$z = \frac{x-\mu}{\sigma}$$

Real World Examples

Data collected on many natural phenomena, such as height and weight of people, have an approximate normal distribution.

These data are typically influenced by many factors and no single factor overpowers the others.