# Z- score

*The z score is computed by the following formula:*

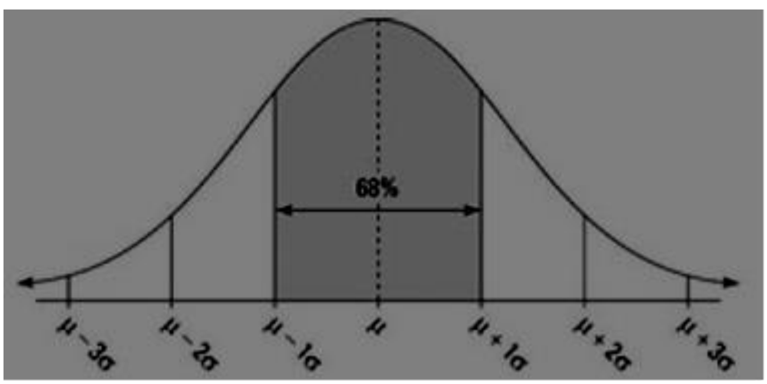
$$z = \frac{X - \mu}{\sigma}$$

*z= standard score*

*= original score    μ= population mean    σ= population standard deviation*

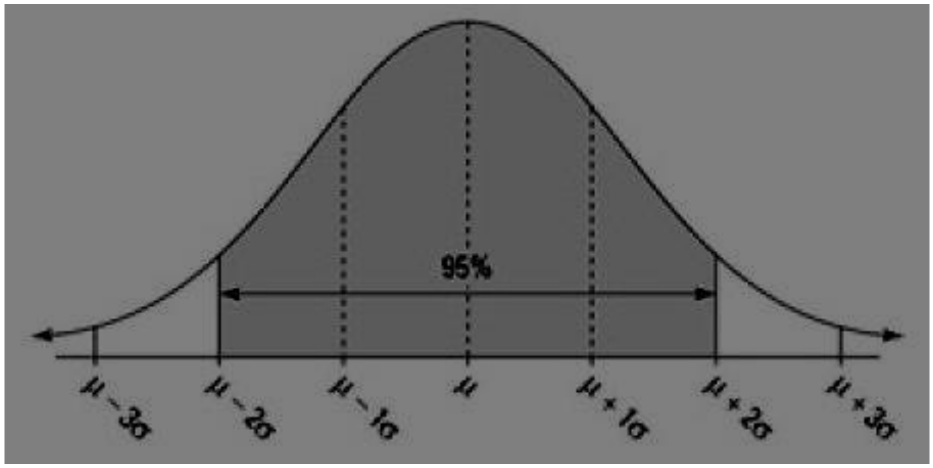
**Z score rule :**

**- One sigma rule: 68% (68.2%) of observations lie between minus & plus one time SD. ( -1Z & +1Z ).**

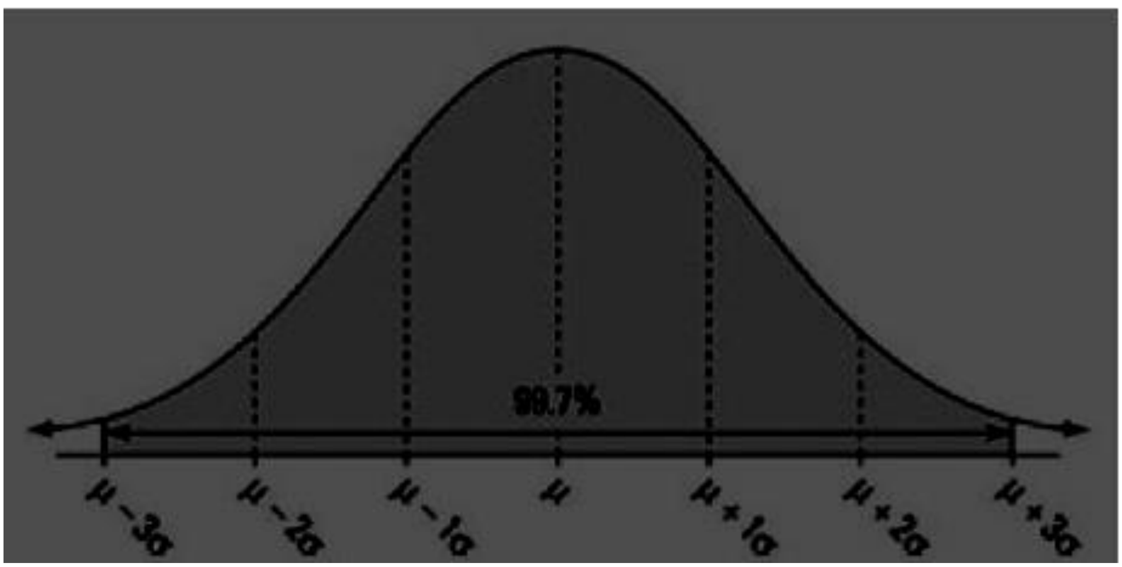





**- Two sigma rule: 95% (95.4%)of observations lie between minus 2 & plus 2 SD units**

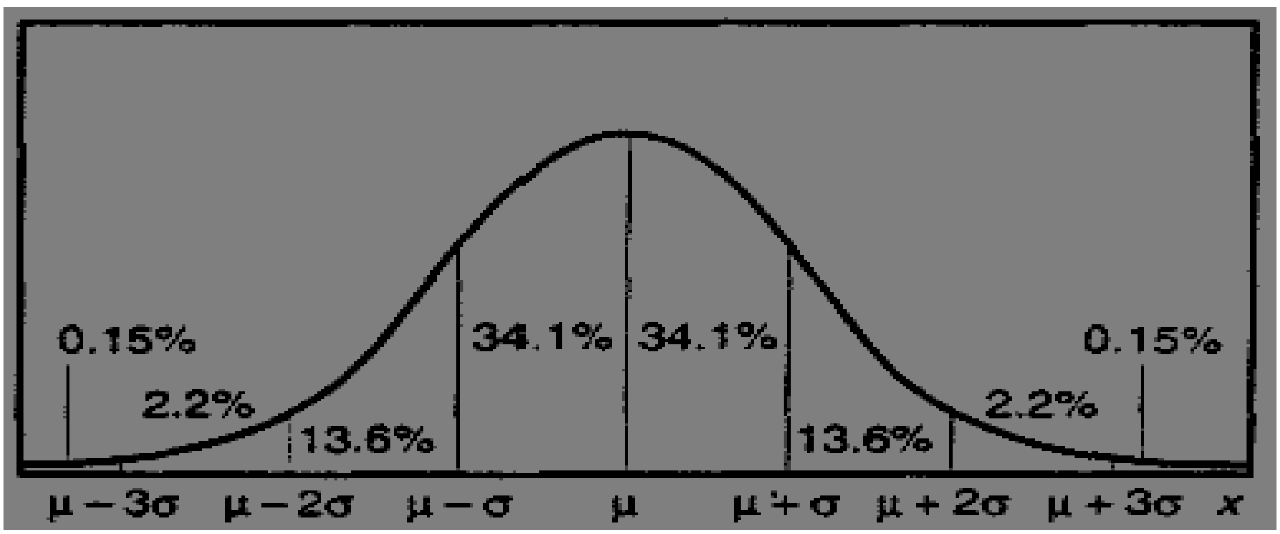



- 99 (99.8%) of observations lie between minus 3 & plus 3 SD units.



**Empirical rules :**

- One sigma rule: 68% (68.2%) of observations lie between minus & plus one time SD. ( -1Z & +1Z ).
- Two sigma rule: 95% (95.4%)of observations lie between minus 2 & plus 2 SD units
- Three sigma rule: 99 (99.8%) of observations lie between minus 3 & plus 3 SD units.




**Z-score: Z score is a measure of the distance that a particular population member is from the population's mean. It is so named because it is frequently used on populations that have a normal distribution, which is also known as a Z distribution.**

**A z-score is also known as a standard score because it is measured in units of standard deviation, which allows observations from different distributions to be compared.**

**In statistics, a z-score (or standard score) is used to compare means from**

**Different normally distributed sets of data. The actual score indicates how many standard deviations an observation is above or below the mean.**

The z- score is useful in research utilizing statistical analysis because it allows for the comparison of observations from different normal distributions. In effect, when items from different data sets are transformed into z-scores, then they may then all be compared. This article will show you how to calculate a z-score (or standard score).

**Example**

- Assuming systolic blood pressure (BP) in normal healthy individuals is normally distributed with **$\mu$ = 120** and **$\sigma$**= 10 mm Hg,
- *What area of the curve is above 130 mm Hg?*
- *What area of the curve is between 100 and 140 mm Hg?*

**Answer**

# z = (130 - 120) / 10 = 1.00, and the area above 1.00 is 0.159. So 15.9% of

normal healthy individuals have a systolic blood pressure above 1 standard deviation (>I30 mm Hg).

# z = (100 - 120) / 10 = -2.00. and = (140 - 120)/10 = 2.00; the area between -2 and +2 is 0.951.

So 95.4% have a systolic blood pressure between -2 and +2 - standard deviations (between 100 and 140 mm Hg).

**Benefit of Sigma rule (Mean± SD) :**

Normal ranges are frequently based on the mean ±2 SDs.

Low SD, the data are all clustered tightly around the mean and the distribution is tall and thin.

Higher SD, the data are more scattered, the distribution is low and wide.

**Calculate 95% Confidence interval.**

**95% Confidence interval:**

The mean derived from the sample is the best available estimate of the population Mean and is referred to as the point estimate.
A mean derived from a sample is unlikely to be a perfect estimate of the population mean.
A range within which we are reasonably confident the true populations mean lies, it is called 95%CI.

- Standard deviation tells us about the variability (spread) in a sample.
- The CI tells us the range in which the true value (the mean if the sample were infinitely large) is likely to be.

- Greater SD gives wider intervals.
- Greater sample sizes give narrower intervals.

**Table of probabilities related to multiples of SDs**

| **Number of SD** | **Probability of observation showing at least as large a deviation from the population mean** |
|---|---|
| 0.674 | 0.50 |
| 1.0 | 0.317 |

| | |
|---|---|
| 1.645 | 0.10 |
| 1.960 | 0.05 |
| 2.0 | 0.046 |
| 2.576 | 0.01 |

**Confidence limits:**

- 68% of observations lie between minus & plus one SD. ( -1Z & +1Z ).
- 95% of observations lie between minus 1.96 & plus 1.96 SD units.
- 99% of observations lie between minus 2.58 & plus 2.58 SD units.
- 99.7% of observations lie between minus 3 & plus 3 SD units.