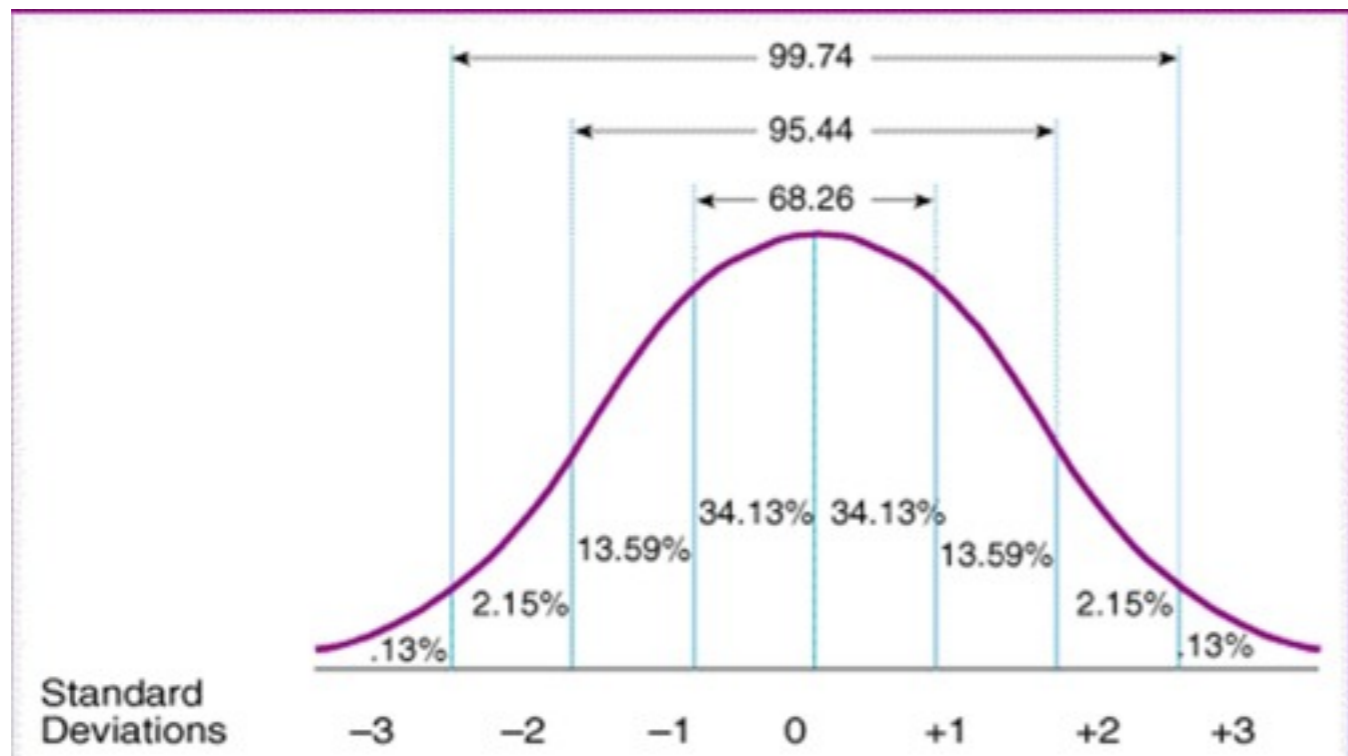


Inferential statistics - Confidence intervals: definition



Ask Mish



IN a NORMAL DISTRIBUTION approximative:

- 68% of the data are within one SD (-1; +1)
- 95% are within 2 SD (-2;+2)
- 99.7% are within 3 SD (-3;+3)
- 0.3% are beyond 3 SD.

CONFIDENCE INTERVALS:

- If you have a **N=sample size** -> you take and measure **n outcomes** -> you can calculate the **mean \bar{X}_0** from these n outcomes. However the result is a distribution of outcomes around this mean. **Confidence intervals** is about how far from the mean you want to go to feel confident with the result.
- It is generally accepted that a 95% interval around the mean (meaning 2 SD above and 2 SD below the mean) would give a good estimation of the sample. This means that from 100 outcomes, based on the 95% CI formula you will be able to recognize as good 95 outcomes. You will make mistakes in 5 outcomes which you will recognize as good when in reality they are not.

