

# 1 Confidence Interval

## 1.1 The Confidence Level (95%) and What it Means

The “95%” in a 95% confidence interval refers to the confidence level, which is the likelihood that the confidence interval contains the true population parameter if we repeated the sampling process many times.

- Imagine you could take hundreds or thousands of new samples from the same population, and for each sample, you calculate a new confidence interval. 95% of these intervals would contain the true population parameter (like the true mean height).
- However, 5% of those intervals would not contain the true population parameter because of random sample variation.

## 1.2 What the 95% Does Not Mean

Incorrect interpretation: “There is a 95% chance that the true population mean is within this particular interval.”

- This is incorrect because the true population parameter is fixed, and the confidence interval is just one possible range based on the sample. The interval either contains the true value or it doesn’t; there is no probability attached to the specific interval after it is calculated.

Correct interpretation: “If we repeated the experiment many times, 95% of the calculated intervals would contain the true mean.”

- Here, we’re talking about a long-run probability: the procedure of generating confidence intervals produces correct results 95% of the time.