CONFIDENCE INTERVALS

The objective is the construction of confidence intervals for population parameters. When the underlying distribution of a statistic is approximately normal [1], the confidence interval will be of the form

Point estimate $\pm$ margin of error

The question becomes: how do we build the margin of error, i.e. how to identify the appropriate test statistics.

$\longrightarrow$ When σ is known

If the population standard deviation σ is known, a (100-α)% confidence interval for μ is given by:

$$\bar{\mathrm{X}} \pm z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$$

$\bar{\mathrm{X}}$: sample mean; n: sample size; σ: population standard deviation
z: z value from the z-test statistics; α: significance level
$z_{\alpha/2} = 1.96$ for a 95% confidence interval

Use cases:

- estimation of a population proportion p (how close is the sample proportion $\hat{p}$ to the true value of the parameter)
- estimation of a population mean in the unlikely case that σ is known

$\longrightarrow$ When σ is unknown

If the population standard deviation σ is not known, a (100-α)% confidence interval for μ is given by:

$$\bar{\mathrm{X}} \pm t_{\alpha/2} \times \frac{s}{\sqrt{n}}$$

We use s the unbiased sample standard deviation to estimate σ and therefore need to implement the t-procedure instead of the z procedure.

$\bar{\mathrm{X}}$: sample mean; n: sample size; s: unbiased sample standard deviation
t: t value from the t distribution with $n-1$ degrees of freedom
α: significance level

[1] One of the major justification for supposing a normal distribution is the central limit theorem.