Confidence Intervals

Decision Making Under Uncertainty (Transitional Year) - 2025/2026
ISCTE-IUL

Afonso Moniz Moreira[1,2]

[1] ISCTE-IUL, Department of Quantitative Methods for Economics and Management
[2] CMVM - Securities Market Commission, Market Supervision Department

Warning/Disclaimer

- This set of slides is not, nor does it intend to be, a replacement for the core bibliography of the Decision Making Under Uncertainty course.
- This set of slides is not, nor does it intend to be, a rigorous source for studying the course topics.
- The sole purpose of these slides is to help the instructor guide the classes as colloquially as possible without carrying unnecessary formalism.
- Accordingly, statistical formalism is removed whenever possible to streamline students' first exposure to the topics.

What are they? What are they for?

- A confidence interval, a CI, is a way of performing inference in which we use a point estimator and its standard deviation.

What are they? What are they for?

- A confidence interval, a CI, is a way of performing inference in which we use a point estimator and its standard deviation.
- The point estimator has its standard deviation weighted by a critical value added and subtracted, yielding an interval.

What are they? What are they for?

- A confidence interval, a CI, is a way of performing inference in which we use a point estimator and its standard deviation.
- The point estimator has its standard deviation weighted by a critical value added and subtracted, yielding an interval.
- The computation of this critical value will be detailed later on.

What are they? What are they for?

- A confidence interval, a CI, is a way of performing inference in which we use a point estimator and its standard deviation.
- The point estimator has its standard deviation weighted by a critical value added and subtracted, yielding an interval.
- The computation of this critical value will be detailed later on.
- Thus, when we use a realization of the random sample (i.e., the observed sample), we obtain an interval instead of a single number.

What are they? What are they for?

- A confidence interval, a CI, is a way of performing inference in which we use a point estimator and its standard deviation.
- The point estimator has its standard deviation weighted by a critical value added and subtracted, yielding an interval.
- The computation of this critical value will be detailed later on.
- Thus, when we use a realization of the random sample (i.e., the observed sample), we obtain an interval instead of a single number.
- We immediately grasp the range of values that **may** contain the parameter of interest.

What are they? What are they for?

- As an example, let us consider the sample mean ($\bar{X}$), which is a centered point estimator for the population mean X, and its standard deviation $\sigma_{\bar{X}}$, respectively:

What are they? What are they for?

- As an example, let us consider the sample mean $(\bar{X})$, which is a centered point estimator for the population mean X, and its standard deviation $\sigma_{\bar{X}}$, respectively:

$$\bar{X} = \sum_{i=1}^{N} X_i, \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}, \tag{1}$$

then the confidence interval, in general terms, will look like this:

What are they? What are they for?

- As an example, let us consider the sample mean ($\bar{X}$), which is a centered point estimator for the population mean X, and its standard deviation $\sigma_{\bar{X}}$, respectively:

$$\bar{X} = \sum_{i=1}^{N} X_i, \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}, \tag{1}$$

then the confidence interval, in general terms, will look like this:

$$]CI[_\mu =]\bar{X} - \text{c.v} \times \sigma_{\bar{X}}; \bar{X} + \text{c.v} \times \sigma_{\bar{X}}[\tag{2}$$

What are they? What are they for?

- As an example, let us consider the sample mean ($\bar{X}$), which is a centered point estimator for the population mean X, and its standard deviation $\sigma_{\bar{X}}$, respectively:

$$\bar{X} = \sum_{i=1}^{N} X_i, \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}, \tag{1}$$

then the confidence interval, in general terms, will look like this:

$$]CI[_{\mu} =]\bar{X} - \text{c.v} \times \sigma_{\bar{X}}; \bar{X} + \text{c.v} \times \sigma_{\bar{X}}[\tag{2}$$

- Therefore, a confidence interval is a point estimator with a margin of error given by the standard deviation weighted by the critical value.

What are they? What are they for?

- So how do we obtain the critical value after all?

What are they? What are they for?

- So how do we obtain the critical value after all?
- The critical value is a quantile of the distribution of the point estimator we are using in the inference.

What are they? What are they for?

- So how do we obtain the critical value after all?
- The critical value is a quantile of the distribution of the point estimator we are using in the inference.
- It is determined by the choice of the interval's confidence level, commonly denoted by λ, and vice versa.

What are they? What are they for?

- So how do we obtain the critical value after all?
- The critical value is a quantile of the distribution of the point estimator we are using in the inference.
- It is determined by the choice of the interval's confidence level, commonly denoted by λ, and vice versa.
- The higher the confidence level (λ), the larger the critical value ($c.v^{\lambda}$) because we are considering a greater portion of the estimator's distribution.

What are they? What are they for?

- So how do we obtain the critical value after all?
- The critical value is a quantile of the distribution of the point estimator we are using in the inference.
- It is determined by the choice of the interval's confidence level, commonly denoted by λ, and vice versa.
- The higher the confidence level (λ), the larger the critical value (c.v^{λ}) because we are considering a greater portion of the estimator's distribution.
- The most frequent confidence levels are 90%, 95% and 99%.

What are they? What are they for?

- So how do we obtain the critical value after all?
- The critical value is a quantile of the distribution of the point estimator we are using in the inference.
- It is determined by the choice of the interval's confidence level, commonly denoted by λ, and vice versa.
- The higher the confidence level (λ), the larger the critical value ($c.v^{\lambda}$) because we are considering a greater portion of the estimator's distribution.
- The most frequent confidence levels are 90%, 95% and 99%.
- These give rise to distinct critical values, namely $c.v^{90\%}$, $c.v^{95\%}$ and $c.v^{99\%}$, which differ according to the distribution of the point estimator we are using.

What are they? What are they for?

- However, the quantiles of the estimator's distribution are never given directly by the confidence level λ but as a function of it.

What are they? What are they for?

- However, the quantiles of the estimator's distribution are never given directly by the confidence level λ but as a function of it.
- The estimator's distribution quantiles are therefore given by:

$$\text{c.v}_a^\lambda = Q_{\frac{1-\lambda}{2}} = Q_{\frac{\alpha}{2}} \quad \text{and} \quad \text{c.v}_b^\lambda = Q_{\lambda+\frac{1-\lambda}{2}} = Q_{\lambda+\frac{\alpha}{2}}. \tag{3}$$

What are they? What are they for?

- However, the quantiles of the estimator's distribution are never given directly by the confidence level λ but as a function of it.
- The estimator's distribution quantiles are therefore given by:

$$\text{c.v}_a^\lambda = Q_{\frac{1-\lambda}{2}} = Q_{\frac{\alpha}{2}} \quad \text{and} \quad \text{c.v}_b^\lambda = Q_{\lambda+\frac{1-\lambda}{2}} = Q_{\lambda+\frac{\alpha}{2}}. \tag{3}$$

- If the estimator's distribution is symmetric, then $\text{c.v}_a^\lambda = -\text{c.v}_b^\lambda \iff Q_{\frac{1-\lambda}{2}} = -Q_{\lambda+\frac{1-\lambda}{2}} \iff Q_{\frac{\alpha}{2}} = -Q_{\lambda+\frac{\alpha}{2}}$.

What are they? What are they for?

- However, the quantiles of the estimator's distribution are never given directly by the confidence level λ but as a function of it.
- The estimator's distribution quantiles are therefore given by:

$$\text{c.v}_a^\lambda = Q_{\frac{1-\lambda}{2}} = Q_{\frac{\alpha}{2}} \quad \text{and} \quad \text{c.v}_b^\lambda = Q_{\lambda+\frac{1-\lambda}{2}} = Q_{\lambda+\frac{\alpha}{2}}. \tag{3}$$

- If the estimator's distribution is symmetric, then $\text{c.v}_a^\lambda = -\text{c.v}_b^\lambda \iff Q_{\frac{1-\lambda}{2}} = -Q_{\lambda+\frac{1-\lambda}{2}} \iff Q_{\frac{\alpha}{2}} = -Q_{\lambda+\frac{\alpha}{2}}$.
- The relationship $\alpha = 1 - \lambda$ will be clarified when we discuss hypothesis tests; for now it is simply an equation.

What are they? What are they for?

- For the previous example, we know that the distribution of the sample mean ($\bar{X}$) is Gaussian (i.e., Normal), so if we want a 95% confidence level we must compute the 97.5 quantile of the Gaussian (i.e., there must be 2.5% left in each tail), which for the standard Gaussian is $\text{c.v}^{95\%} = Z_{97.5} = 1.96$.

What are they? What are they for?

- For the previous example, we know that the distribution of the sample mean ($\bar{X}$) is Gaussian (i.e., Normal), so if we want a 95% confidence level we must compute the 97.5 quantile of the Gaussian (i.e., there must be 2.5% left in each tail), which for the standard Gaussian is $\text{c.v}^{95\%} = Z_{97.5} = 1.96$.
- We can therefore specify interval (2) further:

$$]CI_{95\%}[_{\mu} = \left] \underbrace{\sum_{i=1}^{N} \frac{1}{N} X_i}_{\bar{X}} - \underbrace{1.96}_{\text{c.v}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} ; \underbrace{\sum_{i=1}^{N} \frac{1}{N} X_i}_{\bar{X}} + \underbrace{1.96}_{\text{c.v}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} \right[\tag{4}$$

What are they? What are they for?

- For the previous example, we know that the distribution of the sample mean ($\bar{X}$) is Gaussian (i.e., Normal), so if we want a 95% confidence level we must compute the 97.5 quantile of the Gaussian (i.e., there must be 2.5% left in each tail), which for the standard Gaussian is $c.v^{95\%} = Z_{97.5} = 1.96$.
- We can therefore specify interval (2) further:

$$]CI_{95\%}[_{\mu} = \left] \underbrace{\sum_{i=1}^{N} \frac{1}{N} X_i}_{\bar{X}} - \underbrace{1.96}_{c.v^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} ; \underbrace{\sum_{i=1}^{N} \frac{1}{N} X_i}_{\bar{X}} + \underbrace{1.96}_{c.v^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} \right[\tag{4}$$

- How can we interpret this interval?

What are they? What are they for?

- It is easier to understand the interpretation if we write the interval as follows:

What are they? What are they for?

- It is easier to understand the interpretation if we write the interval as follows:

$$\mathbb{P}\left[\underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} - \underbrace{1.96}_{\text{c.v}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} < \mu < \underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} + \underbrace{1.96}_{\text{c.v}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}}\right] = 0.95 \tag{5}$$

What are they? What are they for?

- It is easier to understand the interpretation if we write the interval as follows:

$$\mathbb{P}\left[\underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} - \underbrace{1.96}_{\text{c.v}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} < \mu < \underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} + \underbrace{1.96}_{\text{c.v}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}}\right] = 0.95 \quad (5)$$

- The probability that interval (5) contains the true value of the population mean is 95%.

What are they? What are they for?

- It is easier to understand the interpretation if we write the interval as follows:

$$\mathbb{P}\left[\underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} - \underbrace{1.96}_{\text{c.v}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} < \mu < \underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} + \underbrace{1.96}_{\text{c.v}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}}\right] = 0.95 \tag{5}$$

- The probability that interval (5) contains the true value of the population mean is 95%.
- In other words, if we plug 100 realizations of the random sample into the previous interval, 95 of them will contain the true value of μ, 5 of them will not.

What are they? What are they for?

- However, we **cannot** state which of the 100 concrete intervals calculated contains the true μ, **why?**

What are they? What are they for?

- However, we **cannot** state which of the 100 concrete intervals calculated contains the true μ, **why?**
- Intervals (4) and (5) are called random intervals because they depend on the random variables $(X_1, ..., X_N)$; the sample's realization has not yet replaced them.

What are they? What are they for?

- However, we **cannot** state which of the 100 concrete intervals calculated contains the true μ, **why?**
- Intervals (4) and (5) are called random intervals because they depend on the random variables $(X_1, ..., X_N)$; the sample's realization has not yet replaced them.
- Attention!!! The previous interpretation is for the random CI and not for the concrete one.

What are they? What are they for?

- However, we **cannot** state which of the 100 concrete intervals calculated contains the true μ, **why?**
- Intervals (4) and (5) are called random intervals because they depend on the random variables $(X_1, ..., X_N)$; the sample's realization has not yet replaced them.
- Attention!!! The previous interpretation is for the random CI and not for the concrete one.
- Once we substitute the observed sample it becomes a concrete confidence interval, because it has specific/observed values; the random variables $(X_1, ..., X_N)$ give way to their realizations $(x_1, ..., x_N)$.

What are they? What are they for?

- However, we **cannot** state which of the 100 concrete intervals calculated contains the true μ, **why?**
- Intervals (4) and (5) are called random intervals because they depend on the random variables $(X_1, ..., X_N)$; the sample's realization has not yet replaced them.
- Attention!!! The previous interpretation is for the random CI and not for the concrete one.
- Once we substitute the observed sample it becomes a concrete confidence interval, because it has specific/observed values; the random variables $(X_1, ..., X_N)$ give way to their realizations $(x_1, ..., x_N)$.
- When we are presented with a concrete CI we can only expect it to be one of those that contain the true value of μ.

Real Examples of Using Confidence Intervals I

- Legislative and Presidential Elections:

Real Examples of Using Confidence Intervals I

- Legislative and Presidential Elections:
  - Even though a point estimate is presented for each candidate, it is then stated that these values take into account a confidence interval of $x\%$, usually 95%.

Real Examples of Using Confidence Intervals I

- Legislative and Presidential Elections:
  - Even though a point estimate is presented for each candidate, it is then stated that these values take into account a confidence interval of x%, usually 95%.
  - The midpoint of the interval corresponds to the point estimate. However, an estimator has variance and this must be considered when analyzing the phenomenon. Therefore, a confidence interval accounts for a margin of error that has an underlying confidence level.

Real Examples of Using Confidence Intervals I

- Legislative and Presidential Elections:
  - Even though a point estimate is presented for each candidate, it is then stated that these values take into account a confidence interval of $x\%$, usually 95%.
  - The midpoint of the interval corresponds to the point estimate. However, an estimator has variance and this must be considered when analyzing the phenomenon. Therefore, a confidence interval accounts for a margin of error that has an underlying confidence level.
  - It is precisely with these bands of variation that projections can be made throughout election night.

Real Examples of Using Confidence Intervals II

- Clinical trials for the effectiveness of a new medical treatment to correct myopia:

Real Examples of Using Confidence Intervals II

- Clinical trials for the effectiveness of a new medical treatment to correct myopia:
    - We want to infer the true mean reduction of diopters in the patients who underwent the new treatment.

Real Examples of Using Confidence Intervals II

- Clinical trials for the effectiveness of a new medical treatment to correct myopia:
  - We want to infer the true mean reduction of diopters in the patients who underwent the new treatment.
  - Note that it is different to state that the mean of this reduction is a certain value ξ than to state that the true value lies between ζ and η with a given probability.
  - We are considering a margin of error, which corresponds to accounting for the estimator's variability at a given confidence level.

Steps to Build a Confidence Interval

- 1st Step: Find a point estimator with suitable properties.
- 2nd Step: Set a confidence level λ.

Steps to Build a Confidence Interval

- 1st Step: Find a point estimator with suitable properties.
- 2nd Step: Set a confidence level λ.
- 3rd Step: Know the sample size n.

Steps to Build a Confidence Interval

- 1st Step: Find a point estimator with suitable properties.
- 2nd Step: Set a confidence level λ.
- 3rd Step: Know the sample size n.
- 4th Step: Know the sampling distributions of the random variables (i.e., of the point estimator) used to estimate the parameter.
- These steps culminate in what is called the pivot variable method, where the crux of the matter, as we will see, is the choice of the point estimator (i.e., the pivot variable) for building the interval.

What are the conditions for choosing the pivot variable?

- A successful choice of the pivot variable (i.e., the point estimator) obeys the following points:

What are the conditions for choosing the pivot variable?

- A successful choice of the pivot variable (i.e., the point estimator) obeys the following points:
- 1st - It must contain the parameter to be estimated in its expression.

What are the conditions for choosing the pivot variable?

- A successful choice of the pivot variable (i.e., the point estimator) obeys the following points:
- 1st - It must contain the parameter to be estimated in its expression.
- 2nd - It must not contain any other unknown parameters in its expression.

What are the conditions for choosing the pivot variable?

- A successful choice of the pivot variable (i.e., the point estimator) obeys the following points:
- 1st - It must contain the parameter to be estimated in its expression.
- 2nd - It must not contain any other unknown parameters in its expression.
- 3rd - Its distribution cannot depend on the parameter to be estimated, nor on any other values that are unknown.

What pivot variable options are there?

- This table of pivot variables will be provided in the assessment moments. In this sense, you only need to know what to look for.

What pivot variable options are there?

- This table of pivot variables will be provided in the assessment moments. In this sense, you only need to know what to look for. Tables Link on GitHub
- The choice of pivot variable depends on the problem and on the data available in it.

The Simplest Confidence Interval

- Example of a confidence interval for the mean of a population with known variance.

The Simplest Confidence Interval

- Example of a confidence interval for the mean of a population with known variance.
- 2nd Step - Set a confidence level λ.

The Simplest Confidence Interval

- Example of a confidence interval for the mean of a population with known variance.
- 2nd Step - Set a confidence level λ.
- Let us consider a 95% confidence level.

The Simplest Confidence Interval

- Example of a confidence interval for the mean of a population with known variance.
- 2nd Step - Set a confidence level λ.
- Let us consider a 95% confidence level.
- 3rd Step - Know the sample size n.

The Simplest Confidence Interval

- Example of a confidence interval for the mean of a population with known variance.
- 2nd Step - Set a confidence level λ.
- Let us consider a 95% confidence level.
- 3rd Step - Know the sample size n.
- Let us consider a sample of size $n = 100$.

The Simplest Confidence Interval

- Example of a confidence interval for the mean of a population with known variance.
- 2nd Step - Set a confidence level λ.
- Let us consider a 95% confidence level.
- 3rd Step - Know the sample size n.
- Let us consider a sample of size $n = 100$.
- Now we go to the table to find a pivot variable that meets the following steps:

The Simplest Confidence Interval

- Example of a confidence interval for the mean of a population with known variance.
- 2nd Step - Set a confidence level λ.
- Let us consider a 95% confidence level.
- 3rd Step - Know the sample size n.
- Let us consider a sample of size $n = 100$.
- Now we go to the table to find a pivot variable that meets the following steps:
- 1st Step - Find a point estimator with suitable properties.

The Simplest Confidence Interval

- Example of a confidence interval for the mean of a population with known variance.
- 2nd Step - Set a confidence level λ.
- Let us consider a 95% confidence level.
- 3rd Step - Know the sample size n.
- Let us consider a sample of size $n = 100$.
- Now we go to the table to find a pivot variable that meets the following steps:
- 1st Step - Find a point estimator with suitable properties.
- 4th Step - Know the sampling distributions of the random variables (i.e., of the point estimator) used to estimate the parameter.

The Simplest Confidence Interval

- Therefore, when analyzing the table, the pivot variable we will use is:

The Simplest Confidence Interval

- Therefore, when analyzing the table, the pivot variable we will use is:

$$\bar{X} = \sum_{i=1}^{n} \frac{1}{n} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \tag{6}$$

The Simplest Confidence Interval

- Therefore, when analyzing the table, the pivot variable we will use is:

$$\bar{X} = \sum_{i=1}^{n} \frac{1}{n} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \tag{6}$$

- Consequently, it is also true that:

The Simplest Confidence Interval

- Therefore, when analyzing the table, the pivot variable we will use is:

$$\bar{X} = \sum_{i=1}^{n} \frac{1}{n} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \tag{6}$$

- Consequently, it is also true that:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1). \tag{7}$$

The Simplest Confidence Interval

- Therefore, when analyzing the table, the pivot variable we will use is:

$$\bar{X} = \sum_{i=1}^{n} \frac{1}{n} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \tag{6}$$

- Consequently, it is also true that:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1). \tag{7}$$

- This pivot variable satisfies all the previous criteria.

The Simplest Confidence Interval

- By definition of a symmetric confidence interval, we have:

The Simplest Confidence Interval

- By definition of a symmetric confidence interval, we have:

$$\mathbb{P}\left[\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \text{c.v}^{95\%}\right] = 0.95 \qquad (8)$$

The Simplest Confidence Interval

- By definition of a symmetric confidence interval, we have:

$$\mathbb{P}\left[\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \text{c.v}^{95\%}\right] = 0.95 \tag{8}$$

- Specifying the previous probability is only possible if we know the pivot variable's distribution.

The Simplest Confidence Interval

- By definition of a symmetric confidence interval, we have:

$$\mathbb{P}\left[\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \text{c.v}^{95\%}\right] = 0.95 \tag{8}$$

- Specifying the previous probability is only possible if we know the pivot variable's distribution.
- By manipulating the inequality inside the probability, we obtain:

The Simplest Confidence Interval

- By definition of a symmetric confidence interval, we have:

$$\mathbb{P}\left[\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \text{c.v}^{95\%}\right] = 0.95 \tag{8}$$

- Specifying the previous probability is only possible if we know the pivot variable's distribution.
- By manipulating the inequality inside the probability, we obtain:

$$\mathbb{P}\left[\bar{X} - \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}}\right] = 0.95 \tag{9}$$

The Simplest Confidence Interval

- Another way of writing the **random** interval is:

The Simplest Confidence Interval

- Another way of writing the **random** interval is:

$$]CI_{95\%}[_{\mu} = \left] \bar{X} - \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}} ; \bar{X} + \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}} \right[\qquad (10)$$

The Simplest Confidence Interval

- Another way of writing the **random** interval is:

$$]CI_{95\%}[_{\mu} = \left]\bar{X} - \mathrm{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}}; \bar{X} + \mathrm{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}}\right[\qquad (10)$$

- At this point, the interpretation is that the probability of the previous interval containing the true value of the population mean is 95%.

The Simplest Confidence Interval

- Another way of writing the **random** interval is:

$$]CI_{95\%}[_{\mu} = \left]\bar{X} - \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}}; \bar{X} + \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}}\right[\qquad (10)$$

- At this point, the interpretation is that the probability of the previous interval containing the true value of the population mean is 95%.
- This interval is random because it depends on the random sample $(X_1, ..., X_n)$ contained in the sample mean $\bar{X}$.

The Simplest Confidence Interval

- If we now want to compute a concrete confidence interval, we simply replace it with the realization of the random sample.

The Simplest Confidence Interval

- If we now want to compute a concrete confidence interval, we simply replace it with the realization of the random sample.
- Let us consider that the sample mean is $\bar{x} = 10$ and the population standard deviation is $\sigma = 2$, then the concrete confidence interval is:

$$]CI_{95\%}[^{*}_{\mu} = \left] 10 - 1.96 \times \frac{2}{\sqrt{100}} ; 10 + 1.96 \times \frac{2}{\sqrt{100}} \right[\qquad (11)$$

The Simplest Confidence Interval

- If we now want to compute a concrete confidence interval, we simply replace it with the realization of the random sample.
- Let us consider that the sample mean is $\bar{x} = 10$ and the population standard deviation is $\sigma = 2$, then the concrete confidence interval is:

$$]CI_{95\%}[^*_\mu = \left]10 - 1.96 \times \frac{2}{\sqrt{100}}; 10 + 1.96 \times \frac{2}{\sqrt{100}}\right[\quad (11)$$

- More specifically:

$$]CI_{95\%}[^*_\mu =]10 - 0.392; 10 + 0.392[=]9.608; 10.392[\quad (12)$$

- The only thing we can state is that this interval is one of the 95 that contain the true value of μ.

Exercise Nº1 I

A radio station wants to estimate the average amount of time a family spends per day listening to that station. A random sample of 81 families was collected, and an average daily listening time of 2.4 hours and a standard deviation of 0.7 hours were calculated. In that city, how much time does a family spend on average per day listening to that station? Answer by providing a point estimate and a 90% confidence interval. Interpret the meaning of the values found.

Exercise Nº4 I

When estimating the mean of a normal population by means of a 90% confidence interval, what should be the minimum sample size so that the width of that interval is less than $\frac{\sigma}{9}$, given that σ is known?

Exercise Nº6 I

The QUITEGOOD cookie factory is considering purchasing a new automatic packaging machine that, according to the technicians of the company selling it, guarantees an average of 60 packages per minute. The selling company placed the machine on consignment at QUITEGOOD's facilities. Over 25 one-minute periods, the number of packaged units was counted, yielding the following results:

$$\sum_{i=1}^{25} x_i = 1400 \text{ packages} \qquad \sum_{i=1}^{25} x_i^2 = 81000 \text{ packages}^2$$

(a) Build a 95% confidence interval for the average number of packages wrapped per minute. Should we accept the statement made by the selling company's technicians? Justify.

Exercise Nº6 II

(b) What should the sample size be if we intend to reduce the interval's width by half?

Exercise Nº9 I

The amount of wine per bottle, measured in centiliters, is a random variable with a normal distribution whose true mean and variance are unknown. A random sample of 25 bottles yielded the value 0.0625 centiliters^2 as the estimate for the corrected sample variance.

(a) Estimate the standard deviation of this population using a 90% confidence interval.

Recently there have been several consumer complaints claiming that the information on the bottle labels stating that each bottle contains 75 centiliters of the precious nectar is not accurate. This statement is refuted by the company's quality control manager,

Exercise Nº9 II

who, based on a sample of 25 bottles, produced the following confidence interval:

$$[I_\lambda]^*_\mu =]74,8968;75,1032[$$

(b) What value would you propose as an estimate for the average amount of wine per bottle? Justify your answer appropriately.

(c) What is the margin of error associated with the confidence interval obtained?

(d) Determine the confidence level $\lambda = 1 - \alpha$ used by the quality control manager when constructing that interval.

Exercise Nº9 III

(e) Comment on the truthfulness of the following statement from the consumer rights association when confronted with the confidence interval presented: ≪It is our belief that only half of the bottles marketed respect the lower and upper limits of that interval≫.

Exercise Nº10 I

If during a STOP operation on the EN1 national road, out of 600 cars, 114 had severe electrical system failures, build a 95% confidence interval for the true proportion of cars with severe electrical system failures traveling on that road. Make any assumptions you deem necessary.

Exercise Nº11 I

A company intends to launch a new product in a city of one million inhabitants. In the market study conducted, 1000 people were surveyed and 800 of them stated that they would very unlikely use that new product.

(a) How many inhabitants are expected to use the new product? Justify your answer.

(b) Build a 95% confidence interval for the true proportion of inhabitants who will use the new product. Interpret the result you obtained.

(c) How many people must be added to the initially collected sample if we want the margin of error to be at most 1%?

Exercise Nº17 I

In two Sports Medicine Centers, the intention was to compare the athletes' average weights, so two random samples of 10 athletes each were collected, yielding the following results: means of 77 kg and 68 kg, respectively, with standard deviations of 6 kg and 10 kg. Determine a 90% confidence interval for the difference between the athletes' mean weights in the two centers.

Exercise Nº20 I

As part of a financial institution's structural training plan, two types of learning methods were used to improve English skills. Method A, called conventional, was used with 100 workers and method B, called modern, was used with 120 workers. At the end of the course, the workers took an assessment test, yielding the following results:

	Number of Passes	Number of Fails
Method A	88	12
Method B	105	15

Given these results, the training plan manager states that ≪the modern method proved to be more effective≫. Comment on this statement using a confidence interval with $\lambda = 0.90$.

Exercises on Output Interpretation

- In the following link we have the confidence interval exercises with output interpretation: ▸ Output Exercises Link on GitHub