**Point Estimators and Confidence Intervals**

**Learning Objectives**:

Computation and interpretation of confidence interval for:

* mean
* the difference between two means
* proportion
* the difference between two proportion

Recapitulation (Additional Requests).

**Problem**: A study has been conducted to compare the effect of a drug (Nebivolol) in decreasing arterial blood pressure compared with an older drug (Propranolol). Systolic and diastolic pressures were measured for each subject at the inclusion in the study and after 3 months of treatment. Furthermore, information related with the presence (1) or absence (0) of two adverse effects (*dry mouth* and *nausea or vomiting*) were collected. The collected data are in the file named **Nebiv.xlsx**.

**Requests**

1. Create a new Microsoft Excel document named ***ConfidenceIntervals.xlsx*** and save it in the **Lab09** folder.
2. Copy the collected data from ***Nebiv.xlsx*** to ***ConfidenceIntervals.xlsx****.* Save the file.
3. Assess the normal distribution of variable named *Baseline SBP*. Do not impose any classes of frequency! Put the results in a new sheet named Histogram. Write if the data looks to follow a normal distribution (symmetry around the mean).
4. Under the assumption of normal distribution compute descriptive statistics for *3months SBP* separately for subjects who received Nebivolol and for those who received Propranolol. Put the results in a new sheet named *Final SBP*.

**One quantitative variable: Point estimator and associated confidence interval**

1. Copy the whole table with data in a new sheet named *Point estimators*.
2. Sort the whole table with the data according with treatment. Using the predefine function (COUNTIF) count how many subjects received Nebivolol.
3. For subjects who received Nebivolol compute using predefine functions the arithmetic mean (AVERAGE) and standard deviation (STDEV) for 3 months systolic blood pressure (*3 months SBP*).
4. Specify the value of the point estimator for 3 months systolic blood pressure.
5. The 95% confidence interval will be calculated to withdraw the sampling variability on the point estimator for the 3 months systolic blood pressure. Create first a table with the following structure:

|  |  |
| --- | --- |
|  | Value |
| Mean (m) | =AVERAGE(array) |
| Standard deviation (s) | =STDEV(array) |
| Significance level (α) | 0.05 |
| Sample size (n) | =countif(array, "nebivolol") |
| tdf | =T.INV.2T(α, n-1) |
| 95%CI lower bound |  |
| 95%CI upper bound |  |

Apply the following formula to calculate the lower and upper bound of 95% confidence interval associated to the mean:



Using CONCATENATE function, write in a single cell the calculated 95%CI as type (95%CI lower bound; 95%CI upper bound). In the next cell to the right of the concatenated confidence interval write its interpretation.

Interpretation by example: We are 95% confident that the population treated with Nebivolol has on average SBP from 145 to 147.

**Two quantitative variables: Point estimator and associated confidence interval**

We are interested in the efficacy of the Nebivolol (is the mean SBP of subjects treated with Nebivolol different by the mean of SBP of subjects treated with Propranolol?)

1. To answer the above question several calculation must be done. Please create a table like the one bellow:

|  |  |  |
| --- | --- | --- |
|  | Nebivolol | Propranolul |
| Mean (m) |  |  |
| Standard deviation (s) |  |  |
| Sample size (n) |  |  |

Fill the above table with values calculated using predefined functions.

1. Compute the 95% confidence interval for the difference between means.

|  |  |
| --- | --- |
|  | Value |
| Mean difference (m1-m2) |  |
| Significance level (α) | 0.05 |
| Standard error for the difference between means (SE) | 8.23 |
| tdf | 1.96 |
| 95%CI lower bound |  |
| 95%CI upper bound |  |

The formula of 95% confidence interval for the difference between two means:

, where



Using CONCATENATE function, write in a single cell the calculated 95%CI as type (95%CI lower bound; 95%CI upper bound). In the next cell to the right of the concatenated confidence interval write its interpretation.

Interpretation by example: Since the 95% confidence interval for the difference between means we can say that the means are not statistically different.

**One qualitative variable: Point estimator and associated confidence interval**

1. Create a graphical representation to illustrate the apparition of *nausea or vomiting* according to the received treatment. Use pivot table for this request.
2. Which is the chance to have *nausea or vomiting* as adverse effect for subjects treated with nebivolol? Create a table like the one bellow and fill it with the requested information:

|  |  |
| --- | --- |
| Probability to have *nausea or vomiting* (f) |  |
| Significance level (α) | 0.05 |
| Standard error (SE) |  |
| Zα | 1.96 |
| 95%CI lower bound |  |
| 95%CI upper bound |  |

Formulas to be apply:

 where 

1. Using CONCATENATE function, write in a single cell the calculated 95%CI as type (95%CI lower bound; 95%CI upper bound) and use number with 2 decimals. In the next cell to the right of the concatenated confidence interval write its interpretation.

**Two qualitative variables: Point estimator and associated confidence interval**

Is the probability of dry mouth on subjects who received nebivolol significantly different by the probability of dry mouth on subjects who received propranolol?

1. Create a graphical representation to illustrate the relation between treatment and apparition of *dry mouth*.
2. Compute the 95% confidence interval associated to the different by filling the table bellow

|  |  |
| --- | --- |
| Probability to have *dry mouth* on subjects who received Nebivolol (f1) |  |
| Probability to have *dry mouth* on subjects who received Propranolol (f2) |  |
| f1-f2 |  |
| Sample size for Nebivolol (n1) |  |
| Sample size for Propranolol (n2) |  |
| Standard error (SE) |  |
| Zα | 1.96 |
| 95%CI lower bound |  |
| 95%CI upper bound |  |

Use the following formulas:

, where



Using CONCATENATE function, write in a single cell the calculated 95%CI as type (95%CI lower bound; 95%CI upper bound) and use number with two decimals. In the next cell to the right of the concatenated confidence interval write its interpretation.

**Additional Requests**

1. Create a new column named TreatmentEffic. Display for each subject if the treatment was efficient using the following criterion:

A treatment is considered efficient if the subjects have the SBP at three months less than 130 mmHG

1. Create a new variable named MAP baseline (mean arterial pressure). Calculate for each subject the value of MAP using the following formula:

MAP baseline = (2\*baseline DBP+ baseline SBP)/3

1. Create a graphical representation to investigate the relation between age and baseline SBP.
2. Create a graphical representation to see the distribution of treatment according with gender.
3. Compute the coefficient of variation for male patient for the following variables: Age (years), Baseline SBP (mmHG), 3months SBP (mmHG), 3months DBP (mmHG).
4. Compute the following descriptive statistics for 3months SBP (mmHG), 3months DBP (mmHG) separately for subjects treated with Nebivolol and those treated with Propranolol: mean, median, mode, quartile 1, quartile 3, standard error.
5. Which is the probability that a subject who receive Nebivolol to have dry mouth?