

## MICROSOFT EXCEL & EPI INFO: CONFIDENCE INTERVALS BY EXAMPLE

### Learning Objectives:

- Computation and interpretation of confidence interval for mean. (Exercise 1–4, 6, 7)
- Computation and interpretation of confidence interval for proportion. (Exercise 5 & 8)
- Working with Epi Info for analysis medical data. (Exercise 9)

Create a new Microsoft Excel document named **ConfidenceIntervals.xlsx** and save it in the **Lab09** folder.

### EXERCISE 1

The mean value of the weight calculated on a sample was 76.4 kg. The associated confidence interval for the mean is [53.2; 99.6] kg.

- Rename the first sheet of the **ConfidenceIntervals.xlsx** file as Exercise 1.
- In the Exercise 1 sheet specify the value of punctual estimator of the weight.
- Write in this sheet the interpretation of the confidence interval associated to the mean of weight.

### EXERCISE 2

We are interested to find the confidence interval associated to the mean of cholesterol on a sample of 50 subjects with different diseases. We know that the cholesterol data are normally distributed and we also have the values of the mean in the sample (m) and the standard deviation in the population ( $\sigma$ ).  $\alpha$

- Insert a new sheet named Exercise 2 in the **ConfidenceIntervals.xlsx** file.
- Create in this sheet the following tables:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	a)				b)				c)				d)				e)	
2	m	200			m	200			m	200			m	200			m	180
3	$\sigma$	20			$\sigma$	20			$\sigma$	20			$\sigma$	40			$\sigma$	20
4	n	50			n	50			n	100			n	50			n	50
5	confidence level	0.95			confidence level	0.9			confidence level	0.95			confidence level	0.95			confidence level	0.95
6	$\alpha$	0.05			$\alpha$				$\alpha$				$\alpha$				$\alpha$	
7	Z	1.96			Z				Z				Z				Z	
8																		
9	95%CI-lower bound	194.46			90%CI-lower bound				95%CI-lower bound				95%CI-lower bound				95%CI-lower bound	
10	95%CI-upper bound	205.54			90%CI-upper bound				95%CI-upper bound				95%CI-upper bound				95%CI-upper bound	
11																		
12	CI=	(194.46; 205.54)																

- For each situation presented above:
  - Display the value of significance level ( $\alpha$ ) using the following formula exemplified for the first table (a): =1-B5
  - Display the standard Z using the following formula exemplified for the first table (a): =ABS(ROUND(NORMSINV(B6/2),2))
  - Compute the lower bound of confidence intervals using the following formula exemplified for first table (a): =B2-B7\*B3/SQRT(B4)
  - Compute the upper bound of confidence intervals using the following formula exemplified for first table (a): =B2+B7\*B3/SQRT(B4)
  - Display the confidence limits using the following formula exemplified for the first table (a): =CONCATENATE("(",ROUND(B9,2),";",ROUND(B10,2),")")
  - Do all calculation for tables b), c), d) and e)

By observing the obtained results, please provide an answer to the following questions. Put the answer under the tables.

- 1) A confidence level is (larger/narrow/of the same length) if the confidence level is smaller?
- 2) A confidence level is (larger/narrow/of the same length) if the sample size is bigger?
- 3) A confidence level is (larger/narrow/of the same length) if standard deviation is larger?
- 4) A confidence level is (larger/narrow/of the same length) if confidence level is smaller?
- 5) A confidence level is (larger/narrow/of the same length) if the mean for which is calculated is smaller?

### EXERCISE 3

Under assumption of normal distribution, compute the 95% confidence intervals for mean associated to the weight of a sample of 21 subjects with diabetes, knowing the sample mean ( $m = 75$  kg) and sample standard deviation ( $s = 10$  kg).

Use the following formula:  $\left( m - t_{n-1, 1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}; m + t_{n-1, 1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right)$ , where  $t_{n-1, 1-\alpha/2} = 2.42$

Put the results in a new sheet named *Exercise 3* in the **ConfidenceIntervals.xlsx** file.

### EXERCISE 4

Under assumption of normal distribution, compute the 95% confidence intervals for mean associated to triglycerides of a sample of 100 subjects with hypertension, knowing the sample mean ( $m = 170$  mg/dL) and sample standard deviation ( $s = 16$  mg/dL).

Use the following formula:  $\left( m - Z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}; m + Z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right)$ , where  $Z_{1-\alpha/2} = 1.96$

Put the results in a new sheet named *Exercise 4* in the **ConfidenceIntervals.xlsx** file.

### EXERCISE 5

Compute the 95% confidence interval associated to frequency of pain after vaccination ( $f = 0.25$ ) on a sample of 100 vaccinated children.

Use the following formula:  $\left( f - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{f \cdot (1-f)}{n}}; f + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{f \cdot (1-f)}{n}} \right)$ , where  $Z_{1-\alpha/2} = 1.96$

Put the results in a new sheet named *Exercise 5* in the **ConfidenceIntervals.xlsx** file.

### EXERCISE 6

Save on the **Lab09** folder the file named **OralCancer.xlsx**.

In the sheet named Height compute the 95% confidence intervals for this variable. For the calculation use the following formula:  $\left( m - t_{n-1, 1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} ; m + t_{n-1, 1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right)$ , where  $t_{n-1, 1-\alpha/2}=2.42$ ,  $s$  = the values retrieved by STDEV predefined function, and  $m$  = the value retrieved by AVERAGE predefined function.

## EXERCISE 7

Copy the weight from the Data sheet of **OralCancer.xlsx** file in a new sheet named Exercise 6. Using Data Analysis option, compute for this variable (weight) the 95% confidence interval associated to the mean.

The results must look line in the image bellow (Please do not copy the results from the picture bellow – obtain descriptive statistics for Weight (Greutate) using Data Analysis).

	A	B
1	Greutate	
2		
3	Mean	77.04936
4	Standard Error	1.79286
5	Median	75.60101
6	Mode	#N/A
7	Standard Deviation	13.77122
8	Sample Variance	189.6465
9	Kurtosis	-0.5176
10	Skewness	0.498867
11	Range	55.44145
12	Minimum	55.21503
13	Maximum	110.6565
14	Sum	4545.912
15	Count	59
16	Confidence Level(95.0%)	3.5888
17		
18	lim inf = B3 - B16	73.46056
19	lim sup = B3 + B16	80.63816
20	CI	(73,46; 80,64)

## EXERCISE 8

- Read the Data sheet from the **OralCancer.xlsx** file in EpiInfo.
- Create a new Microsoft Excel file named **ResultsEpiInfo.xlsx** file in the **Lab09** folder.
- Create the frequency table for Gender. Copy the results in the **ResultsEpiInfo.xlsx** and specify the punctual estimator for Female and Male and respectively associated 95% confidence interval for each estimator. Is there any differences between these two estimators (provide the reason able to sustain you answer).
- Create the frequency table for Mouthwash with alcohol. Copy the results in the **ResultsEpiInfo.xlsx** and specify the punctual estimator for Yes and associated 95% confidence interval.
- Specify which from the two requested confidence levels associated to frequencies is larger (is the one for which the frequency is close to 50% or the one for which the frequency is close by extreme values 0% or 100%).