Exercise 7: Statistical inference, confidence intervals

1. What from the following is necessary to perform random sampling?

1.1 Normal distribution of the scores in population

1.2 Every individual (measurement unit) in the population has to have the same probability of being chosen in the sample.

1.3 Selection of every individual (measurement unit) has to be fully independent from selection of any other individual.

2. Scores from Wechsler intelligence scales are normally distributed with $\mu = 100$ and $\sigma = 15$. Imagine we testes a random sample of 9 people, computed their M and SD and we repeated the whole procedure 1000 times.

2.1 Estimate SD of the 1000 means.

2.2 Approximately what percentage of these sampling means with N=9 would be higher than 105? And higher than 110?

2.3. Approximately what percentage of these sampling means with N=9 would lie between scores 95 and 105? A between 90 and 110?

2.4 Will these sampling means be normally distributed?

2.5 What is variance of this sampling distribution of means?

2.6 If N=225 (instead of 9), what would be the value of standard error?

2.7 If N=225, what percentage of sampling means would lie between 99 a 101?

2.8 Will the distribution of sampling means be approximately normal, even if the scores distribution in the population wasn't normal?

3. Are the following pairs equivalent?

3.1 standard error of mean AND standard deviation of sampling distribution of mean

3.2 σ^2/n AND standard error of mean

3.3 σ_m^2 AND variance of sampling distribution of mean

3.4 population variance σ^2 AND n times σ_m^2

3.5 mean of the sampling distribution of mean AND σ_m^2

3.6 µ *AND* $((\sum X)/n$

3.7 m *AND* $(\sum x)/n$

3.8 s^2 *AND* $\sum x^2/(n-1)$

4. Answer the following questions:

4.1 If we conduct many studies on different topics and we always construct 95% confidence intervals for the estimated statistics, how many of these intervals will contain the estimated parameter?

4.2 In which of the following cases would increase in sample size cause the biggest constriction of confidence interval?

a) from 5 to 25
b) from 10 to 30
c) from 40 to 60

4.3 Which estimation type informs better about the estimation precision – point estimate or interval estimate?

4.4 The relationship of z and σ_m is equivalent and t and…

a) σ
b) σ^2
c) s
d) s_m

4.5 Under what conditions does it apply that: $\sigma = \sigma_m$?

4.6 What is the relationship between chosen confidence level and confidence interval width?

4.7 Which mathematical theorem says that sampling distribution of mean with growing N approximated normal distribution, no matter what is the variable distribution in population?

4.8 If we know σ, does it apply that m +/- 1,96σ_m makes 95% confidence interval for any N?

4.9 Let's assume a variable with normal distribution and known σ. If we randomly take two samples with N=100 and we compute 68% confidence interval in both samples, will these two intervals be the same?

4.10 How big must be N, so that standard deviation of sampling distribution of mean σ_m was only 10% of standard deviation s of variable X distribution?

5. We conduct a research of whether and how long hospitalizations harm children in development. On of our research questions is, whether there isn't intellect development retardation following the hospitalization. For this purpose, we gave an intelligence test to 30 long-term hospitalized children with results: m_{IQ}=98, s_{IQ}11.

5.1 Compute 95% confidence interval for mean intelligence in the population of long-term hospitalized children (μ_{IQ}).

5.2 Further we found out that the length of hospitalization in days correlates with IQ, r=-0,1. Compute 95% confidence interval for correlation between hospitalization length and IQ (ρ).

6. A researcher is interested in reading efficiency in college students. He measured number of words read per one minute in 6 college students: 200, 240, 300, 410, 450 a 600.

6.1 Compute mean and standard deviation.

6.2 What sampling distribution will you use for confidence interval construction, normal o t-distribution? Why?

6.3 Construct 95% confidence interval for mean.

6.4 Will 99% confidence interval be narrower of wider? Construct the 99% confidence interval for mean as well.

6.5 The researcher thought that 6 people in the sample are too few. He decided to pretend that he included 18 people in the sample and he kept the mean and standard deviation he measured before on 6 people. How will the confidence intervals for mean change? What will change? Construct 95% confidence interval.

7. A researcher is interested in Stroop effect. In every tested person he measured the time in which they name 60 colourful rectangles and the time in which they name 60 words printed with colourful ink, but when these words were also colours names (different than the ink colour). Times for both measurement and their difference can be seen in the following table:

Colourful rectangles	Colourful words	Difference
17	38	21
15	58	43
18	35	17
20	39	19
18	33	15
20	32	12
20	45	25
19	52	33
17	31	14
21	29	8

7.1 Compute 95% confidence intervals for mean difference in times from the table.

7.2 The table shows only a part from the whole data. In the research, 47 people were tested and mean difference was 16.362 seconds and SD was 7.470. Compute 95% confidence intervals for mean difference.

8. Several students suspected bartender that he is not fair to his guests. They ordered 8 beers and measured their volume: 0.51, 0.462, 0.491, 0.466, 0.461, 0.503, 0.495, 0.488 (in litters, note: in Czech republic one beer should have 0.5 litter). Construct 95% confidence intervals for mean.

9. A researcher is interested in intelligence of children in leisure groups. In children who visit painting group he measured the following IQ scores: 128, 117, 139, 122, 101, 152, 120, 127, 136, 108.

9.1 Construct 95% confidence intervals for mean IQ assuming that population standard deviation of IQ is 15. What will this assumption influence?

9.2 Construct 95% confidence intervals for mean IQ, if you don't know the population standard deviation.

12. In 1961 height of 15 randomly chosen 10-year-old boys living in Czechoslovakia was measured. Construct 95% confidence intervals for mean height, if you know that in 1951 standard deviation of height σ = 6.4 cm was measure and it is known that height variability won't change much.

Measured heights:
130, 140, 136, 141, 139, 133, 149, 151, 139, 136, 138, 142, 127, 139, 147

13. Construct 95% confidence interval for correlation coefficient ρ, if you know that correlation r=0.45 between monthly income and sport and recreation expenses was measured in 20 families.

16. If 95% confidence interval includes 0, will 99% confidence interval include 0 as well?