Part I. **(30 points)** It is reported by tourism officials of Paris that the stairs to the top of the Eiffel Tower have 704 steps (=merdiven basamağı) in total. A group of 52 tourists arriving on a tour bus from Istanbul climbs up the stairs to the top of the Eiffel Tower and each tourist counts the number of steps. The average number of steps as counted by the tourists is 715, and the standard deviation of their counts is 49 steps.

1. (8 points) Do the data indicate that the tourists coming from Istanbul in general can count the number of steps more or less correctly on the average?

$$H_o: \mu = 704$$
$$H_a: \mu \neq 704$$

$$z = \frac{715 - 704}{49/\sqrt{52}} = 1.61$$

$$P\text{-value} \approx 100\% - 89\% = 11\%$$

Since P-value > 5% , we do not reject H_o

The tourists from Istanbul can count correctly on the average.

2. (12 points) Another independent group of 16 tourists coming from Tokyo counts the steps as 693 on the average with a standard deviation of 37 steps. Is there a significant difference between the mean counts of people from Istanbul and Tokyo?

$$H_o: \mu_1 = \mu_2$$
$$H_a: \mu_1 \neq \mu_2$$

$$d.f. = 16 + 52 - 2 = 66, \text{ large.}$$

$\Rightarrow$ Use z-table.

$$SE_{diff} = \sqrt{\frac{49^2}{52} + \frac{37^2}{16}} \cong 11.48$$

$$z = \frac{715 - 693}{11.48} = 1.91 \Rightarrow P\text{-value} = (100 - 94.26)\%$$
$$= 5.74\%$$

P-value > 5% $\Rightarrow$ Do not reject H_o.

There is no significant difference between the mean counts of people from Istanbul and Tokyo.
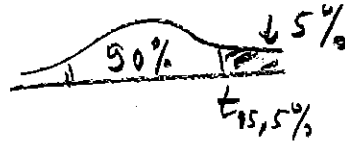
Part I. (continued)

3. (7 points) Construct a 90% confidence interval for the mean number of steps as counted by tourists from Tokyo in general.

$$693 \mp 1.75 (9.25)$$

$$SE = \frac{37}{4} = 9.25 \qquad \Rightarrow [676.8, 709.2]$$

$$t_{15,5\%} = 1.75$$

Since d.f. $= 16-1 = 15$

4. (3 points) Does the interval you constructed in Question 3 contain

    i. the number of steps as reported by tourism officials of Paris?

    Yes; it contains 704.

    ii. the average number of steps that could be counted by all possible tourists from Tokyo?

    We do not know for sure. We are 90% confident that the confidence interval contains this overall average.

Part II. **(20 points)** A medical psychology professor has designed a "TV-cycle" which forces young children to pedal (=pedal çevirmek) for watching TV. That is, the child has to ride a stationary (=sabit) bicycle in front of the TV to keep the TV working and watch it.

A random sample of 5 overweight children is selected. They are provided with only a TV-cycle, and no ordinary TV. Here are the total minutes that each child has watched TV within one week:

$$41, 32, 56, 79, 68$$

1. (6 points) Find the standard deviation of the sample.

Hint: Sum (observation – average)² for all observations, then divide by "n-1", then take the square root.

$$\text{Mean} = \frac{41+32+56+79+68}{5} = 55.2 \text{ min.s}$$

$$SD = \sqrt{\frac{(41-55.2)^2 + (32-55.2)^2 + \cdots + (68-55.2)^2}{4}} \cong 19.2$$

If you cannot solve part 1, then use SD=15 in Question 2 (this is not the correct SD, but OK to use).

2. (12 points) Can the professor conclude that "the children who can access a TV only by a TV-cycle watch on the average about 1 hour per week"?
Perform a test of significance to test this claim.

$$H_0 : \mu = 60 \text{ min.} \qquad n = 5 \quad \text{small sample}$$

$$H_a : \mu \neq 60 \text{ min} \qquad SE = \frac{SD}{\sqrt{n}} = \frac{19.2}{\sqrt{5}} \cong 8.6$$

$$t = \frac{55.2 - 60}{8.6} \cong -0.56$$

$$d.f. = 5-1 = 4 \qquad \text{t-table} \Rightarrow \text{P-value} > \underbrace{25\% \times 2}_{50\%}$$

Large P-value! So, we cannot reject H_0,
The professor can conclude that the children watch TV about 1 hr/week.

3. (2 points) What is your assumption about the distribution of the population of TV-cycle hours of all possible children, for you to conduct the test in Question 1?

We assume this distribution is Normal, so that we can use t-test for a small sample.

Part III. (20 points) It is claimed that the new generation gets bored more quickly than the older generation, possibly due to all the electronic stimulants such as computers, cell phones, and so on.

In particular, the rates of staying in their first job are compared between the younger and older generations. Out of 450 subjects randomly selected from the older generation, 49 of them said they stayed less than a year in their first job. On the other hand, in a sample of 360 subjects from the younger generation, 61 stayed less than a year in their first job.

1. (15 points) Is there sufficient evidence that the younger generation gets bored from their first job at a higher rate than the older generation?

$$H_0 : \ p_1 = p_2$$
$$H_a : \ p_1 < p_2$$

$$\hat{p}_1 = \frac{49}{450} \cong 0.11 \qquad \hat{p}_2 \cong 0.17$$

$$SE_{diff} = \sqrt{\frac{(0.11)(0.89)}{450} + \frac{(0.17)(0.83)}{360}}$$

$$\cong 0.025$$

$$z = \frac{0.11 - 0.17}{0.025} = -2.4$$

$$P\text{-value} = \frac{100 - 98.36}{2} \% = 0.82 \%$$

P-value < 5% ⟹ Reject H_0,

There is highly significant evidence that younger generation has a higher rate of staying shortly in their first job (interpreted as getting bored in this job).

2. (2 points) Is the 95% confidence interval for the difference (of the rates mentioned for the younger and older generations) wider or narrower than the 90% confidence for the same difference? Why?

(Do not construct the intervals, but explain in view of the formula for a confidence interval for the difference).

Since we use a bigger z for 95% CI, it is wider than the 90% CI. $z = 1.96$

$z = 1.65$

3. (3 points) If 200 different 90% confidence intervals are constructed from 200 different samples, all independent from each other, how many of those intervals do you expect to cover the true difference between the younger and older generation?

Each one covers with 30% chance.

⟹ 200 × 90% = 180 of them.

Part IV. **(20 points)** The interviewers working for a polling company noted that "only 1 out of 4 voters" agreed for an interview about their opinion before the last referandum, based on a random sample of size 384. This ratio is well known to be 1 out of 2 for the past polls based on a long history of elections/referandums. 0.50

1. (14 points) Is there sufficient evidence to conclude that the voters in general were more reluctant (=isteksiz) to provide an interview in the last referandum compared with the past elections/referandums? Test at $\alpha = 1\%$.

$$H_0 : p = 0.50$$
$$H_a : p < 0.50$$

$$\hat{p} = \frac{1}{4} = 0.25$$

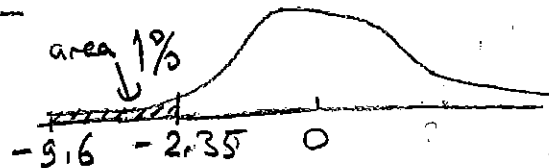$$SE = \sqrt{\frac{(0.50)(0.50)}{384}} = 0.026$$

$$z = \frac{0.25 - 0.50}{0.026}$$

$$= -9.6 \qquad \Rightarrow \quad P\text{-value} \approx 0$$

Clearly, P-value $< 1\%$, as $z_{critical} = -2.35$

Reject H_0 ,

∴ The voters were more reluctant to give an interview in the last referandum.

2. (6 points) Construct a 99% confidence interval for the percentage of voters in the population who were willing to provide an interview before the last referandum.

$$0.25 \mp 2.60 \ (0.022)$$

$$99\% \rightarrow z = 2.60$$

$$\Rightarrow [0.193 \ , \ 0.307]$$

$$SE = \sqrt{\frac{(0.25)(0.75)}{384}} = 0.022$$

$$\Rightarrow [19.3\%, \ 30.7\%]$$

Part V. **(15 points)** In two rival (=rakip) universities, the students are placed mostly from the top 20% of all students who take the university entrance exam. The following table shows the categorization of the students according to their universities and their rankings in the exam. Each cell in the table shows how many students are in that category.

University	Top 5%	Between top 5% and 15%	Between top 15% and 20%	Between top 20% and more
A	1320	3068	984	42
B	1704	5129	1973	98

1. **(5 points)** What is the probability that a randomly selected student from those who entered University A or B, is between top 5% and 15% in ranking?

$$\frac{3068 + 5129}{1320 + 3068 + 984 + 42 + 1704 + 5129 + 1973 + 98}$$

$$= \frac{8197}{14318} \cong 0.57 = 57\%$$

2. **(10 points)** Find the conditional distribution of the rankings of the students in University A.

Total of 5414 std.s in University A.

Top 5%	Top 5-15%	Top 15-20%	Top 20% or more
24%	57%	18%	0.8% ≈ 1%

Prob. D+r:

$\frac{1320}{5414}$ $\frac{3068}{5414}$ $\frac{984}{5414}$ $\frac{42}{5414}$