Part I. Do TV shows with violence worsen memory for commercials (=reklamlar)? To answer this question, researchers divided 324 adults randomly into two groups of 162 each. First group watched a TV program with a violent content rating, and the second group watched a neutral TV program (for general audiences and no violent content). There were 9 commercials in total. After viewing the program, each participant was scored ranging from 0 (no brands recalled) to 9 (all brands recalled). The researchers compared the mean recall scores of the two groups.

1. (5 points) Is this an experiment or an observational study? Explain in at most two sentences.

It is an experiment because the researcher assigns subjects randomly. The subjects themselves do not decide.

2. (5 points) Is it controlled? Explain in at most two sentences.

It is controlled because there is a control group. These are subjects who watch a neutral TV program.

3. (5 points) This study may have been conducted double-blind, although not explicitly stated above. Explain how it can be conducted double-blind, in at most three sentences.

The subjects must not know which group they belong to.
Similarly, the researchers evaluating the results should not know which group a subject belongs to.

4. (5 points) The mean score of the first group was found to be smaller than the mean score of the second group. What general conclusion can you make? State in one sentence.

Since this was a randomized experiment, we can say that TV shows with violence worsen the memory for commercials in general.

5. (5 points) The 95% confidence intervals were formed separately for each group. The confidence interval for the first group was found to be wider than the confidence interval for the second group. If the standard error of the first group is 0.4, what can you say about the standard error of the second group?

a) Smaller than 0.4 b) Approximately equal to 0.4 c) Larger than 0.4

Explain briefly.

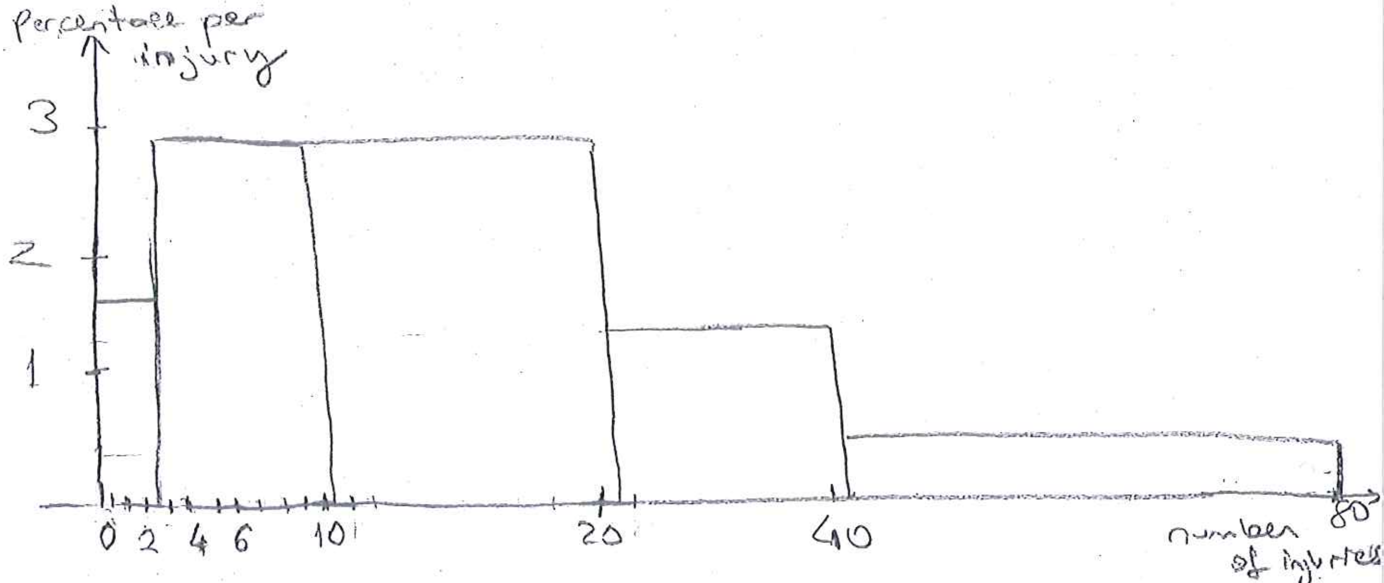
First CI is wider. Formula is $\bar{X} \mp z.SE$, so since z is the same (same confidence level) SE must differ.
Wider CI means larger SE

width of classes

Part II. (25 points) The following is a frequency table for data collected on the number of hand injuries due to using a computer keyboard. Each subject reports the total number of injuries s/he has suffered over a period of a year.

# of injuries	3	8	10	20	40
Class Interval (number of injuries)	0-2	3-10	11-20	21-40	41-80
Frequency	15	69	86	80	50
Rel. Freq	5%	23%	29%	27%	17%
Perc. per (number of injury)	1.67%	2.9%	2.9%	1.3%	0.4%

1. (10 points) Draw the histogram in view of the above table and label both axes. $n = 300$



2. (5 points) Is the median likely to be larger or smaller than the mean? Explain in one sentence by referring to your histogram.

Since the histogram is right-tailed, we expect median to be smaller than the mean.

3. (5 points) In which interval is the 25th percentile? Explain how you found it.

In the interval 3-10, because $5\% + 23\% = 28\% > 25\%$.

4. (5 points) The standard deviation of the data set is found to be 18.7; accordingly, fill in the blanks in the following:

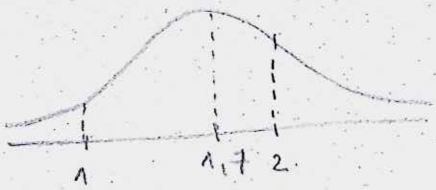
The mean number of injuries in the population can be estimated by the sample mean, give or take, 1.08 injuries.

Hint: Use "sample" and "population" in two of the three blanks above.

$$SE = \frac{18.7}{\sqrt{300}} = 1.08$$

Part III. **(20 points)** A study has been conducted to measure the coping (=baş etmek) skills of caregivers when living or caring for a chronically ill patient. The scores for coping skills are found to be normally distributed with mean 1.7 and standard deviation 0.6.

1. (7 points) Find the percentage of scores between 1 and 2.



$$z_1 = \frac{2-1{,}7}{0{,}6} = 0{,}5$$

$$z_2 = \frac{1-1{,}7}{0{,}6} = -1{,}167$$

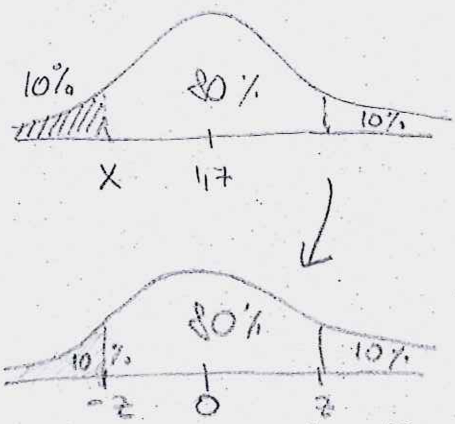
z-score	Area.
0,5	38,29 %
1,167	74,99 %

$$\text{Area} = \frac{1}{2}\left(38{,}29 + 74{,}99\right) = 56{,}64\ \%$$

2. (6 points) The 10^{th} percentile of the score distribution is considered as the cutoff score between the caregivers who can and cannot cope with caring for a chronically ill patient (lower scores indicate coping less). Find this cutoff score.

cutoff score = X.

Area	Z
80%	1,30



$$\frac{X-1{,}7}{0{,}6} = -z_{80} = -1{,}3$$

$$X = -(1{,}30)(0{,}6) + 1{,}7 = 0{,}92$$

3. (5 points) A caregiver with a score of 3 or more is said to cope very well. What is the probability that a randomly selected caregiver can cope very well?

Z	Area
2,167	96,84

$$z = \frac{3-1{,}7}{0{,}6} = 2{,}167$$

Area right of 3 is $\frac{1}{2}(100 - 96{,}84) =$

$= 1{,}58\ \%$

4. (2 points) "The scores of a random sample of caregivers are expected to have an average score of 1.7 and a standard deviation of 0.6, and an approximately bell-shaped histogram." True or False, why (one sentence only)?

True. The expected values for the sample statistics are the parameters from the population like average and SD. Histogram is expected to be similar as well.

Part IV. (20 points) It is well known that income distribution has a long right tail. In 2005, average income in US was reported as \$60,000, standard deviation of income was \$40,000, the median was \$54,000, and the 75th percentile was \$90,000.

1. (5 points) Indicate if the following statements are True or False, and why?

F The standard deviation is a better indicator of variability than the interquartile range.

False, the distribution has a long right tail, therefore interquartile range works better.

F A histogram of many *sample averages*, say from samples of size 200, also has a long right tail.

False. The distribution of sample averages always assumes a bell-shaped curve for many samples.

2. (3 points) What is probability that a randomly selected individual from US in 2005 has an income over \$90,000? What about over \$54,000?

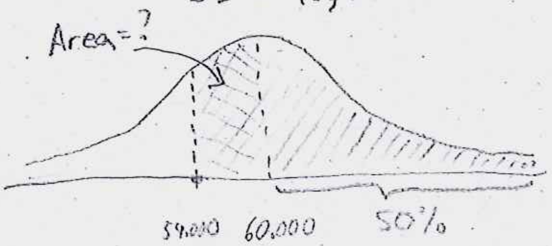
75th percentile was \$90,000 therefore picking a person with greater income has a 25% chance. 54,000 is the median, therefore 50%

3. (7 points) What is the probability that a randomly selected sample of 50 individuals from US in 2005 have an average income over \$54,000?

Z	Area
1,06	70,63

Sample size: N = 50

$SD = 40{,}000 \rightarrow SE = \frac{SD}{\sqrt{50}} \approx 5656{,}8$



$\frac{54{,}000 - 60{,}000}{5656{,}8} =$

$= \frac{-6{,}000}{5656{,}8} = -1.06$

Total Area $= 50\% + \frac{1}{2} \cdot 70.63 = 85,315$

4. (5 points) Construct a 95% confidence interval for the average income level of the population using a recently obtained sample of size 50, which has a mean of \$74,000 and standard deviation of \$45,000.

conf. level → $Z = 2.00$

sample size → $N = 50$

average → $M = 74{,}000$

sample SD → $SD = 45{,}000$

$SE = \frac{SD}{\sqrt{N}} = \frac{45{,}000}{7.071} \approx 6{,}364$

$M + Z \cdot SE = 74000 + 2 \cdot 6364 = 86728$

$M - 2 \cdot SE = 74000 - 2 \cdot 6364 = 61272$

$CI = [61272, 86728]$

Part V. (20 points) In United States (US), it has been found that 80% of all home buyers want a fireplace in their homes.

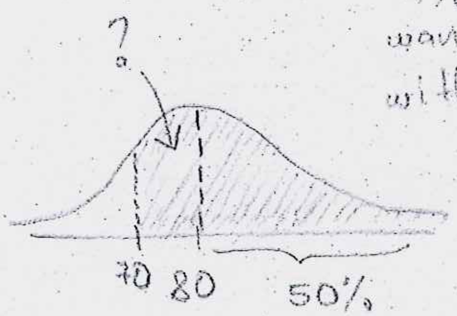
1. (2 points) In a random sample of size 100 from US, what is the expected value of the percentage of home buyers who want a fireplace in their homes?

It is 80%.

2. (7 points) In a random sample of 100 home buyers in US, what is the chance that at least 70 of them want a fireplace?

% of people in a 100-people-sample that want a fireplace is normally distributed with average 80, and

$$SE = \sqrt{\frac{80 \cdot 20}{100}} = \sqrt{\frac{1600}{100}} = 4$$



$$\frac{70-80}{4} = \frac{-10}{4} = -2{,}5$$

z	Area
-2,5	98,76

Area between 70 and 80 is

$\frac{1}{2} \cdot 98{,}76 = 49.38$, therefore total area is $49{,}38 + 50 = 99{,}38\%$

3. (8 points) In Turkey, a recent survey of randomly selected 93 new home buyers indicated that 37 of them wanted a fireplace. Find a 90% confidence interval for the percentage of home buyers in Turkey who want a fireplace in their home.

$p = \frac{37}{93} \approx 0{,}398 \Rightarrow 39{,}8\%$

sample size: N = 93.

$SE = \sqrt{\frac{(39{,}8)(60{,}2)}{93}} \cong 5$

CONF. LEVEL; $z = 1{,}65$

$39{,}8 + (1.65)(5) = \cong 48$

$39{,}8 - (1{,}65)(5) \cong 31.6$

CI = (31.6%, 48%)

4. (3 points) Does the interval you constructed in Question 4 contain the true percentage of home buyers who want a fireplace in Turkey, or not? Explain in one sentence.

There is a 90% chance that the (actual) true percentage is in the conf. interval.