

Chapter IV

Measures of Dispersion (Exercises)

Part I:

1.

The following table shows the annual rainfall (mm) X and its relative frequencies in a certain region.

Interval	Frequency
[0, 200[0.155
[200, 300[0.356
[300, 400[0.090
[400, 500[0.066
[500, 600[0.058
[600, 800[0.116
[800, 1200[0.159

1. Determine the distribution function $F(x)$ and find and interpret $F(700) - F(250)$
2. Estimate and interpret the lower and the upper quartiles of the above distribution.
3. Plot the underlying distribution as a histogram and illustrate the results of 1. and 2. geometrically.

2.

The following table shows the aircraft waiting times and their frequencies at a certain airport. Aircraft waiting time X , in minutes, for a departing aircraft is defined to be the difference between scheduled departure time and actual departure time:

Interval	Frequency
[0, 10[32
[10, 20[40
[20, 45[25
[45, 90[13
[90, 180[15

1. Determine the distribution function $F(x)$ and find and interpret $F(23)$.
2. Estimate and interpret the average waiting time and the standard deviation of the above distribution.

3. Sketch the underlying distribution as a histogram and interpret the results of 2. geometrically.

3.

Students are surveyed as to their number of close friends (individuals to whom they would not be afraid to confide a deep secret). The frequency distribution of replies is given below:

Number of Friends	Frequency
0	2
1	3
2	5
3	5
4	4
5	1

- How many students were surveyed?
- Compute the relative frequencies and construct a histogram.
- What percentage of the students had 2 or fewer friends?
- Find the median, mean and standard deviation.

4.

The following table shows the frequency distribution of the height of 200 participants of a statistics lecture:

Height [m] from... under...	Frequency
1.40 1.54	12
1.54 1.68	38
1.68 1.78	50
1.78 1.88	58
1.88 1.98	30
1.98 2.08	12

- Name and characterise the variable.
- Calculate the relative frequencies and plot them as a histogram.
- Find the percentage of the participants being smaller than 1.75 m
- Find the percentage of the participants being at least 1.60 m but less than 1.90 m tall.
- Find the average height of the participants as arithmetic mean. Check the validity of your result by determining the coefficient of variation.

5.

Spot prices per barrel of crude oil reached their highest levels in history during June and July of 2008. The following data give the spot prices (in dollars) of a barrel of crude oil for 14 business days from June 30, 2008, through July 18, 2008:

139.96	141.06	143.74	145.31	141.38	136.06	135.88
141.47	144.96	145.16	138.68	134.63	129.43	128.94

1. Find the mean for these data.
2. Construct a frequency distribution table for these data using a class width of 3.00 and the lower boundary of the first class equal to 128.00
3. Find the mean of the grouped data.
Compare the means from parts 1. and 3. If the two means are not equal, explain why they differ

Part II (SPSS):

1.

Data File: GSS2000R.sav (Variable: *prestg80*)

1. How many cases are available for the analysis? How many are missing?
2. Find the mean and the standard deviation of the distribution of occupational prestige score.
3. Find the median and the interquartile range
4. Is the distribution skewed to the right or to the left?
5. Which measure of center and spread should be reported for this variable?

(Hint:

There are two *rules of thumb* here:

Rule 1:

When the value of skew statistic is 2 times the value of the skewness standard error, the median is preferred.

Rule 2:

When the skewness is smaller than -1.0 or larger than +1.0 the distribution is badly skewed and the median is a better measure of central tendency.)

Use Rule 2 in our problem.

2.

Data File: *survey_sample.sav* (Variable: *tvhours*)

Calculate and interpret the following measures of central tendency and variability:

Mean, Median, Lower and Upper Quartile, Minimum, Maximum
Standard Deviation.

3.

Data File: *GSS2000R.sav* (Variable: *agekdborn*)

1. Identify the direction of the skewing in the distribution of the variable *agekdborn*.
2. Which measure of center or spread should be reported for this variable?
3. Verify the following statement: “Survey respondents who were male were older when first child was born than survey respondents who were female.”
4. Compare the interquartile ranges of the two groups.