1st Semester, 2021-2022

3rd Year, Math & CompSci

Mathematical Statistics

Seminar Exercises: Week 4

Exercise 1. User ratings on the website **imdb.com** range from 1 to 10. The following is an (imaginary) dataset of the first 100 ratings for the Netflix series **Squid Game (2021)**:

4	4	9	4	3	8	9	10	3	4
9	8	6	10	10	9	9	8	5	6
1	5	6	10	8	4	10	9	9	9
8	9	10	8	7	7	7	7	2	10
8	10	6	2	7	9	8	6	5	8
6	1	8	5	3	8	4	10	9	8
3	3	10	10	10	10	10	9	7	8
7	7	7	8	7	7	9	8	10	9
7	4	4	10	7	5	2	7	8	8
10	4	6	10	8	8	9	8	6	9

(a) Build the (ungrouped) frequency distribution table for this data (both absolute and relative frequencies);

(b) Compute the arithmetic, geometric and harmonic means of the data;

(c) Find the median, the mode and the range of the data;

(d) Find the quartiles, the interquartile range and the outliers of the data;

(e) Find the moment of order 2, the variance and the coefficient of variation of the data.

Exercise 2. The following data represents the number of days of sick leave taken by each of the 50 employees of a given company over the last 6 weeks:

$$2, 2, 0, 0, 5, 8, 3, 4, 1, 0, 0, 7, 1, 7, 1, 5, 4, 0, 4.0, 1, 8, 9, 7, 0$$
$$1, 7, 2, 5, 5, 4, 3, 3, 0, 0, 2, 5, 1, 3, 0, 1, 0, 2, 4, 5, 0, 5, 7, 5, 1.$$

(a) Construct the grouped frequency distribution table for this dataset;

(b) Plot the data using a line graph;

(c) Plot the data as a histogram and relative frequency polygon.

Exercise 3. The following dataset represents the average net monthly income (in thousands of RON) of 30 Romanian families.

4.5	8.9	4.4	6.5	3.5	7.8	9.4	10.2	23	9.8
7.6	4.9	7.3	19	10	9.4	3.9	9.8	4.5	3.6
12	5.3	16.4	10.5	6.8	4.2	10	9.7	4.9	9.8

(a) Group the data into classes and construct the grouped frequency distribution table using Sturges' rule;

(b) Group the data into classes and construct the grouped frequency distribution table using the formula for the class width;

(c) Find the range, the mean and the median of the data;

(d) Find the quartiles, the interquartile range and the outliers of the data.