- Academic Year: 2023-2024

- Department: Computer Science and Engineering

- Course: Introduction to Probability and Descriptive Statistics

Solution TD1

Exercise 1

For each of the following studies, specify: the statistical unit, the population, the characteristic under study, and its nature.

1. Study of the validity period of electric lamps.

    - Statistical Unit: An individual electric lamp.
    - Population: All electric lamps under study.
    - Characteristic: Validity period.
    - Nature: Quantitative continuous.

2. Study of worker absenteeism, in days, in a factory during the year 2018.

    - Statistical Unit: An individual worker.
    - Population: All workers in the factory in 2018.
    - Characteristic: Number of absenteeism days.
    - Nature: Quantitative discrete.

3. Distribution of first-year MI students according to the grade obtained on the Baccalaureate diploma.

    - Statistical Unit: An individual first-year MI student.
    - Population: All first-year MI students.
    - Characteristic: Grade on the Baccalaureate diploma.
    - Nature: Qualitative ordinal.

4. Modeling the number of collisions involving two cars at a set of 100 randomly chosen road intersections in a city. Data are collected over a period of one year, and the number of accidents at each intersection is thus measured.

- Statistical Unit: An individual road intersection.
- Population: The 100 randomly chosen road intersections.
- Characteristic: Number of collisions involving two cars.
- Nature: Quantitative discrete.

Solution to Exercise 2

1. Population: The population consists of students in groups 1, 2, and 3 of section 1 of the first year MI.

2. Characteristic and Its Nature: The characteristic studied is the blood type of each student, which is a qualitative nominal variable.

3. Set of Modalities: The modalities include the blood types A, B, AB, and O.

4. Statistical Table Including Relative Frequencies in Percentages: The formula for the Relative Frequency (%) is given by:

$$\text{Relative Frequency (\%)} = \left(\frac{\text{Frequency of Blood Type}}{\text{Total Observations}} \right) \times 100$$

Assuming the calculation of frequencies from the data, the table below represents a hypothetical example of how these might be tabulated along with their relative frequencies.

Table 1: Relative Frequencies of Blood Types

Blood Type	Frequency	Relative Frequency (%)
A	24	40%
B	12	20%
AB	6	10%
O	18	30%

5. Graphical Representations: Suitable graphical representations for this type of characteristic include pie charts and bar graphs.

Exercise 3

A cereal manufacturer conducts a survey to verify if the cereal boxes indeed contain 500 grams as indicated on the packaging. A sample of 1000 boxes produced in one day is checked. The following data is obtained:

Weight Range (g)	Number of Boxes (n_i)	Increasing Cum. Freq.	Decreasing Cum. Freq.
$[490, 496[$	33	33	1000
$[496, 498[$	168	201	967
$[498, 500[$	415	616	799
$[500, 502[$	293	909	384
$[502, 504[$	75	984	91
$[504, 510[$	16	1000	16

1. The population studied is the cereal boxes produced by the manufacturer in one day, with a sample size of 1000 boxes.

2. The characteristic under study is the weight of the cereal boxes, which is quantitative and continuous.

3. The increasing and decreasing cumulative frequencies were calculated based on the provided data. The results are as follows:

    Increasing Cumulative Frequency: [33, 201, 616, 909, 984, 1000]

    Decreasing Cumulative Frequency: [1000, 967, 799, 384, 91, 16]
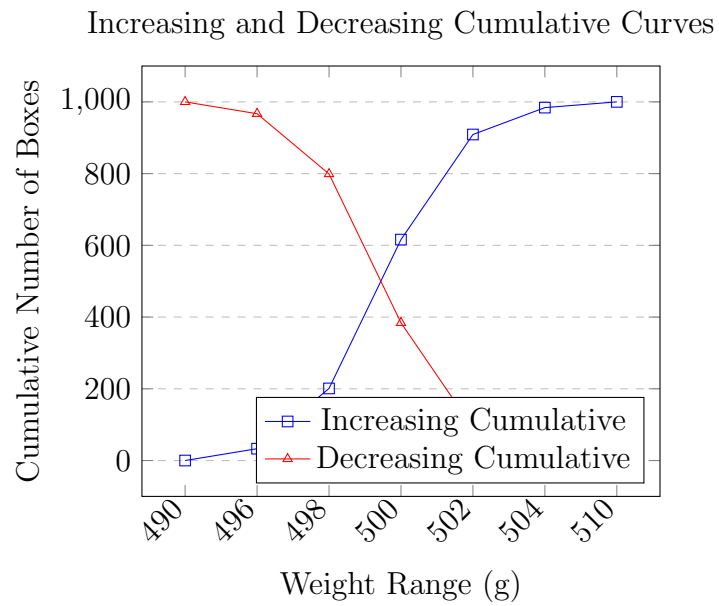
4. **The cumulative curve**,

Figure 1: Comparison of Increasing and Decreasing Cumulative Curves