Mariano Méndez-Suárez

DESCRIPTIVE AND INFERENTIAL STATISTICS

PROBLEMS AND SOLUTIONS

LIBROS PROFESIONALES DE EMPRESA

ESIC Editorial

Descriptive and Inferential Statistics

Problems and Solutions

Madrid, 2025

Mariano Méndez-Suárez

Descriptive and Inferential Statistics

Problems and Solutions

ESIC Editorial

April, 2025

Descriptive and Inferential Statistics: Problems and Solutions
Mariano Méndez-Suárez

Avda. de Valdenigriales, s/n
28223 Pozuelo de Alarcón (Madrid)
Tel.: 91 452 41 00
www.esic.edu/editorial
@EsicEditorial

ISBN: 978-84-1192-155-8
Legal Deposit: M-5076-2025

Cover: Zita Moreno Puig
Design: Santiago Díez Escribano
Printed by: Gráficas Dehon

A publication by

ESIC Editorial

Printed in Spain

This book has been printed with ecological ink and sustainable paper.

To Pilar

Contents

1 Introduction to Descriptive Statistics

1.1. Key Concepts and Techniques for Frequency Tables. | 1.2. Exercises on Descriptive Statistics. | 1.3. Proposed Exercises.

Descriptive statistics are used to summarize and better understand large data sets by providing key metrics such as mean, median, mode, variance, and standard deviation. All of this information is obtained through the use of frequency tables, because they allow you to organize and visualize the raw data into categories or intervals that show the frequency with which each value or range of values occurs. Frequency tables, which allow you to systematically organize and summarize data, are a fundamental tool in data analysis. Frequency tables provide a clear visual representation of how often each value occurs by categorizing and counting the occurrences of certain values within a data set. This approach makes it easier to identify patterns, trends, and anomalies. It helps analysts and decision makers gain insight from raw data.

Frequency tables are used in a variety of fields to understand the distribution of various variables. These tables help quantify the frequency of certain events or responses, whether assessing customer preferences, product performance, or engagement levels. This allows you to make decisions based on evidence instead of guesswork.

In addition, frequency tables are essential for the segmentation of data, the identification of target audiences, and the measurement of the impact of different strategies. Their ability to highlight the most common and least common occurrences makes them an invaluable tool for the refinement of approaches and the optimization of results.

1.1. Key Concepts and Techniques for Frequency Tables

Frequency Table Overview

A **frequency table** is a statistical tool used to represent the distribution of data by grouping values into intervals and calculating various frequencies. Key components include:

- *Absolute Frequency:* The number of times a particular value or range of values occurs.
- *Relative Frequency:* The proportion of occurrences relative to the total number of data points.
- *Cumulative Frequency:* The total of frequencies up to the current point.
- *Cumulative Relative Frequency:* The accumulation of relative frequencies up to the current point.

Key Calculations

- *Total Frequency:* Sum of all the absolute frequencies.
- *Relative Frequency:* Calculated by dividing the frequency of each category by the total number of observations.

$$\text{Relative Frequency} = \frac{f_i}{N}$$

  Where:

  – f_i = Frequency of a category
  – N = Total number of observations

- *Cumulative Absolute Frequency:* The running total of absolute frequencies up to a certain class interval.
- *Cumulative Relative Frequency:* Accumulation of relative frequencies up to the current class interval, providing insight into the percentage of data that falls within a range.

$$\text{Cumulative Relative Frequency} = \sum \frac{f_i}{N}$$

Example Step-by-Step Frequency Table Calculation

Let's consider the following data set representing exam scores from 16 students:

5, 7, 8, 7, 6, 9, 5, 6, 7, 8, 9, 5, 6, 9, 10, 7, 7

- *Step 1: Organize the data.* Sort the data in ascending order:

5, 5, 5, 6, 6, 6, 7, 7, 7, 7, 7, 8, 8, 9, 9, 10

- *Step 2: Count the frequencies.* Now count how many times each value appears:
  – The value 5 appears 3 times.
  – The value 6 appears 3 times.
  – The value 7 appears 5 times.
  – The value 8 appears 2 times.
  – The value 9 appears 2 times.
  – The value 10 appears 1 time.
- *Step 3: Create the frequency table.*

Value	Absolute Frequency
5	3
6	3
7	5
8	2
9	2
10	1

- *Step 4: Calculate relative frequency.* To calculate the relative frequency, divide the absolute frequency of each value by the total number of data points (16):

$$\text{Relative Frequency} = \frac{\text{Absolute Frequency}}{16}$$

Value	Absolute Frequency	Relative Frequency
5	3	0.1875
6	3	0.1875
7	5	0.3125
8	2	0.1250
9	2	0.1250
10	1	0.0625

- *Step 5: Calculate cumulative frequencies.* Finally, calculate the cumulative absolute and cumulative relative frequencies:

Value	Absolute Frequency	Relative Frequency	Cumulative Absolute Frequency	Cumulative Relative Frequency
5	3	0.1875	3	0.1875
6	3	0.1875	6	0.3750
7	5	0.3125	11	0.6875
8	2	0.1250	13	0.8125
9	2	0.1250	15	0.9375
10	1	0.0625	16	1.0000

Example 1: Statistical Analysis of Family Units in a Housing Development

In a housing development, a study has been conducted on the number of members that make up the family unit, yielding the following results:

Data:

- Number of People: 2, 3, 4, 5, 6, 7
- Number of Families: 3, 10, 16, 15, 9, 7

Example 2: Statistical Analysis of Family Units in a Neighborhood

In a neighborhood, a study has been conducted to determine the number of members that make up the family unit, and the results are presented in the adjacent table:

Data:

- Number of People: 2, 3, 4, 5, 6, 7
- Number of Families: 28, 63, 99, 112, 42, 0

Explanation of Variables

- *Number of People:* This is the variable in the study because it represents the number of members in each family unit. A variable is something that can change, and in this case, the number of people per family can vary.

- *Number of Families:* This represents the frequency or how often a family with a certain number of members is observed in the study. Frequency indicates the count of occurrences for each value of the variable.

Why Order Matters

The data is initially unordered in terms of processing, but it is important to organize it to understand how families are distributed across different family sizes. Ordering the data makes it easier to calculate cumulative frequencies and relative frequencies, which are essential for understanding the distribution.

Step-by-Step Solution

Example 1 (Housing Development)

Ordered Data:

Number of People	Number of Families
2	3
3	10
4	16
5	15
6	9
7	7

- *Step 1: Total Frequency.* The total frequency is the sum of all the families:

$$N = 3 + 10 + 16 + 15 + 9 + 7 = 60$$

- *Step 2: Relative Frequency.* The relative frequency for each category is calculated by dividing the number of families by the total frequency.

$$f_i = \frac{n_i}{N}$$

For example, for "Number of People = 2":

$$f_1 = \frac{3}{60} = 0.05$$

Repeating for all categories:

Number of People	Number of Families	Relative Frequency
2	3	0.05
3	10	0.16
4	16	0.26
5	15	0.25
6	9	0.15
7	7	0.11

- *Step 3: Cumulative Absolute Frequency.* The cumulative absolute frequency is calculated by adding each frequency to the sum of the previous frequencies:

Number of People	Cumulative Absolute Frequency
2	3
3	3 + 10 = 13
4	13 + 16 = 29
5	29 + 15 = 44
6	44 + 9 = 53
7	53 + 7 = 60

- *Step 4: Cumulative Relative Frequency.* The cumulative relative frequency is calculated by adding each relative frequency to the sum of the previous relative frequencies:

Number of People	Cumulative Relative Frequency
2	0.05
3	0.05 + 0.167 = 0.217
4	0.217 + 0.267 = 0.484
5	0.484 + 0.25 = 0.734
6	0.734 + 0.15 = 0.884
7	0.884 + 0.117 = 1.000

Example 2 (Neighborhood)

Ordered Data:

Number of People	Number of Families
2	28
3	63
4	99
5	112
6	42
7	0

- *Step 1: Total Frequency.* The total frequency is the sum of all the families:

$$N = 28 + 63 + 99 + 112 + 42 + 0 = 344$$

- *Step 2: Relative Frequency.* The relative frequency for each category is calculated by dividing the number of families by the total frequency. For example, for "Number of People = 2":

$$f_1 = \frac{28}{344} \approx 0.0814$$

Repeating for all categories:

Number of People	Number of Families	Relative Frequency
2	28	0.0814
3	63	0.1831
4	99	0.2872
5	112	0.3256
6	42	0.1221
7	0	0.0000

- *Step 3: Cumulative Absolute Frequency.* The cumulative absolute frequency is calculated as follows:

Number of People	Cumulative Absolute Frequency
2	28
3	28 + 63 = 91
4	91 + 99 = 190
5	190 + 112 = 302
6	302 + 42 = 344
7	344

- *Step 4: Cumulative Relative Frequency.* The cumulative relative frequency is calculated as follows:

Number of People	Cumulative Relative Frequency
2	0.0814
3	0.0814 + 0.1831 = 0.2645
4	0.2645 + 0.2872 = 0.5517
5	0.5517 + 0.3256 = 0.8773
6	0.8773 + 0.1221 = 0.9994
7	1.0000

Comparison of Results

- In the housing development, the total number of families was 60, while in the neighborhood, it was 344.
- The distribution of families is more concentrated around 4 and 5 people in both cases, but the neighborhood had a larger number of families overall.
- The neighborhood shows a higher frequency for larger families (5 people) compared to the housing development, where 4 people families are more common.

1.2. Exercises on Descriptive Statistics

Exercise 1: Daily Visits to a Website

Analyze the distribution of daily visits to a website using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Visits per Day	n_i (Absolute Frequency)
0-5	35
5-10	50
10-15	60
15-20	45
20-25	40

Step-by-Step Calculations

To fully understand the distribution of daily visits to a marketing website, we need to calculate the following:

1. *Total Frequency.* The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 35 + 50 + 60 + 45 + 40 = 230$$

2. *Relative Frequency.* The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

   For example, for the category "0-5":

$$f_1 = \frac{35}{230} \approx 0.1522$$

3. *Cumulative Absolute Frequency.* The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:
   - For "0-5": 35
   - For "5-10": 35 + 50 = 85
   - For "10-15": 85 + 60 = 145
   - For "15-20": 145 + 45 = 190
   - For "20-25": 190 + 40 = 230

4. *Cumulative Relative Frequency.* The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:
   - For "0-5": 0.1522
   - For "5-10": 0.1522 + 0.2174 = 0.3696
   - For "10-15": 0.3696 + 0.2609 = 0.6305
   - For "15-20": 0.6305 + 0.1957 = 0.8261
   - For "20-25": 0.8261 + 0.1739 = 1.0000

Final Calculations

Visits per Day	n_i (Absolute Frequency)	f_i (Relative Frequency)	N_i (Cumulative Absolute Frequency)	F_i (Cumulative Relative Frequency)
0-5	35	0.1522	35	0.1522
5-10	50	0.2174	85	0.3696
10-15	60	0.2609	145	0.6305
15-20	45	0.1957	190	0.8261
20-25	40	0.1739	230	1.0000

Exercise 2: Number of Products Sold per Day (Marketing)

Analyze the distribution of the number of products sold per day using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Products Sold per Day	n_i (Absolute Frequency)
0-10	40
10-20	55
20-30	65
30-40	50
40-50	45

Step-by-Step Calculations

To fully understand the distribution of products sold per day, we need to calculate the following:

1. *Total Frequency.* The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 40 + 55 + 65 + 50 + 45 = 255$$

2. *Relative Frequency.* The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

   For example, for the category "0-10":

$$f_1 = \frac{40}{255} \approx 0.1569$$

3. *Cumulative Absolute Frequency.* The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:
   - For "0-10": 40
   - For "10-20": 40 + 55 = 95
   - For "20-30": 95 + 65 = 160
   - For "30-40": 160 + 50 = 210
   - For "40-50": 210 + 45 = 255

4. *Cumulative Relative Frequency.* The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:

- For “0-10”: 0.1569
- For “10-20”: 0.1569 + 0.2157 = 0.3725
- For “20-30”: 0.3725 + 0.2549 = 0.6275
- For “30-40”: 0.6275 + 0.1961 = 0.8235
- For “40-50”: 0.8235 + 0.1765 = 1.0000

Final Calculations

Products Sold per Day	n_i (Absolute Frequency)	f_i (Relative Frequency)	N_i (Cumulative Absolute Frequency)	F_i (Cumulative Relative Frequency)
0-10	40	0.1569	40	0.1569
10-20	55	0.2157	95	0.3725
20-30	65	0.2549	160	0.6275
30-40	50	0.1961	210	0.8235
40-50	45	0.1765	255	1.0000

Exercise 3: Customer Reviews per Day (Marketing)

Analyze the distribution of customer reviews per day using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Reviews per Day	n_i (Absolute Frequency)
0-2	50
2-4	60
4-6	66
6-8	54
8-10	30

Step-by-Step Calculations

To fully understand the distribution of customer reviews per day, we need to calculate the following:

1. *Total Frequency:* The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 50 + 60 + 66 + 54 + 30 = 260$$

2. *Relative Frequency:* The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

For example, for the category “0-2”:

$$f_1 = \frac{50}{260} \approx 0.1923$$

3. *Cumulative Absolute Frequency:* The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:
   - For "0-2": 50
   - For "2-4": 50 + 60 = 110
   - For "4-6": 110 + 66 = 176
   - For "6-8": 176 + 54 = 230
   - For "8-10": 230 + 30 = 260
4. *Cumulative Relative Frequency:* The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:
   - For "0-2": 0.1923
   - For "2-4": 0.1923 + 0.2308 = 0.4231
   - For "4-6": 0.4231 + 0.2538 = 0.6769
   - For "6-8": 0.6769 + 0.2077 = 0.8846
   - For "8-10": 0.8846 + 0.1154 = 1.0000

Final Calculations

Reviews per Day	n_i (Absolute Frequency)	f_i (Relative Frequency)	N_i (Cumulative Absolute Frequency)	F_i (Cumulative Relative Frequency)
0-2	50	0.1923	50	0.1923
2-4	60	0.2308	110	0.4231
4-6	66	0.2538	176	0.6769
6-8	54	0.2077	230	0.8846
8-10	30	0.1154	260	1.0000

Exercise 4: Daily Transactions in a Business

Analyze the distribution of daily transactions in a business using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Transactions per Day	n_i (Absolute Frequency)
0-20	40
20-40	60
40-60	55
60-80	50
80-100	45

Step-by-Step Calculations

To fully understand the distribution of daily transactions in a business, we need to calculate the following:

1. *Total Frequency:* The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 40 + 60 + 55 + 50 + 45 = 250$$

2. *Relative Frequency:* The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

   For example, for the category "0-20":

$$f_1 = \frac{40}{250} = 0.1600$$

3. *Cumulative Absolute Frequency:* The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:
   - For "0-20": 40
   - For "20-40": 40 + 60 = 100
   - For "40-60": 100 + 55 = 155
   - For "60-80": 155 + 50 = 205
   - For "80-100": 205 + 45 = 250

4. *Cumulative Relative Frequency*: The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:
   - For "0-20": 0.1600
   - For "20-40": 0.1600 + 0.2400 = 0.4000
   - For "40-60": 0.4000 + 0.2200 = 0.6200
   - For "60-80": 0.6200 + 0.2000 = 0.8200
   - For "80-100": 0.8200 + 0.1800 = 1.0000

Final Calculations

Transactions per Day	n_i (Absolute Frequency)	f_i (Relative Frequency)	N_i (Cumulative Absolute Frequency)	F_i (Cumulative Relative Frequency)
0-20	40	0.1600	40	0.1600
20-40	60	0.2400	100	0.4000
40-60	55	0.2200	155	0.6200
60-80	50	0.2000	205	0.8200
80-100	45	0.1800	250	1.0000

Exercise 5: Units Produced per Day (Business)

Analyze the distribution of units produced per day using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Units Produced per Day	n_i (Absolute Frequency)
0-10	35
10-20	50
20-30	60
30-40	55
40-50	40

Step-by-Step Calculations

To fully understand the distribution of units produced per day, we need to calculate the following

1. *Total Frequency:* The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 35 + 50 + 60 + 55 + 40 = 240$$

2. *Relative Frequency:* The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

   For example, for the category "0-10":

$$f_1 = \frac{35}{240} \approx 0.1458$$

3. *Cumulative Absolute Frequency:* The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:
   - For "0-10": 35
   - For "10-20": 35 + 50 = 85
   - For "20-30": 85 + 60 = 145
   - For "30-40": 145 + 55 = 200
   - For "40-50": 200 + 40 = 240
4. *Cumulative Relative Frequency:* The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:

- For “0-10”: 0.1458
- For “10-20”: 0.1458 + 0.2083 = 0.3542
- For “20-30”: 0.3542 + 0.2500 = 0.6042
- For “30-40”: 0.6042 + 0.2292 = 0.8333
- For “40-50”: 0.8333 + 0.1667 = 1.0000

Final Calculations

Units Produced per Day	n_i (Absolute Frequency)	f_i (Relative Frequency)	N_i (Cumulative Absolute Frequency)	F_i (Cumulative Relative Frequency)
0-10	35	0.1458	35	0.1458
10-20	50	0.2083	85	0.3542
20-30	60	0.2500	145	0.6042
30-40	55	0.2292	200	0.8333
40-50	40	0.1667	240	1.0000

1.3. Proposed Exercises

Exercise 1: Daily Visits to an E-commerce Website

Analyze the distribution of daily visits to an e-commerce website using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Visits per Day	Frequency
0-10	45
10-20	55
20-30	50
30-40	40
40-50	35
50-60	25

Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.

Exercise 2: Number of Products Sold per Week (Marketing)

Analyze the distribution of the number of products sold per week using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Products Sold per Week	Frequency
0-50	30
50-100	45
100-150	55
150-200	40
200-250	35
250-300	25

Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.

Exercise 3: Customer Support Tickets per Day (Digital Business)

Analyze the distribution of customer support tickets received per day using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Tickets per Day	Frequency
0-10	40
10-20	50
20-30	55
30-40	45
40-50	35
50-60	30

Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.

Exercise 4: Monthly Subscriptions to a Digital Service

Analyze the distribution of monthly subscriptions to a digital service using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Subscriptions per Month	Frequency
0-50	35
50-100	45
100-150	50
150-200	40
200-250	30
250-300	20

Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.

Exercise 5: Daily Transactions in an Online Store

Analyze the distribution of daily transactions in an online store using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

Data Table

Transactions per Day	Frequency
0-20	35
20-40	50
40-60	55
60-80	45
80-100	40
100-120	30

Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.