

Introduction to the Statistics

List of Supplementary Exercises

1. In the analysis of the granting of loans, a potentially important variable is the person's income. The manager of a bank collects a database from their account holders and extracts the variable "monthly income (R\$)" for 50 people. Although it is a quantitative variable, they want to perform an analysis through a frequency table. Therefore, it is asked:

- a) Classify the account holders in the income ranges, being: 0-2,000; 2,001-4.000; 4.001- 6.001-8.000; 8.001-10.000 and 10.001-12.000.
- b) Next, elaborate the frequency table for the income ranges above.

The database is in the List of Supplementary Exercises: tab Exercise 1.

2. An analyst of the share market collected the monthly returns of two shares that they intend to recommend to their clients. Calculate the descriptive statistics for the two variables, including the correlation coefficient between returns. The database with the monthly percentage returns is in the List of Complementary Exercises: tab Exercise 2.

3. In a certain game, probability of victory (success) for each new move is $1/6$. If 10 moves are done, which are the following probabilities:

- a) Having a victory in 4 moves.
- b) Having a victory in at least 7 moves.

4. (Source: Fávero and Belfiore, 2017, Cap. 5) Suppose a student answer right three questions for each five tests. Being X the number of attempts up to the 12th right one. Determine the probability of the student needing to do 20 questions to get 12 right.

5. (Source: Fávero and Belfiore, 2017, Cap. 5) Suppose in a certain hospital, 3 people are operated daily with a stomach surgery, following a Poisson distribution. Calculate the probability of 28 people being operated in the next week (7 working days).

6. In the last months, the time was measured between the beginning and end of one of the stages of the production process of a certain product. The average time was calculated in 26.5 minutes and the standard deviation was 4.0 minutes. Knowing that this variable follows a normal distribution, identify the following information:

- a) $P(X > 37)$
- b) $P(X < 20)$
- c) $P(22 < X < 28)$

7. (Source: Fávero and Belfiore, 2017, Cap. 8) A group of 60 readers did an assessment of three novel books and, at the end, they chose one of the three options. Test the null hypothesis that there is no difference in the preference of the readers, at the significance level of 5%.

The data are available in the List of Supplementary Exercises: tab Exercise 7.

8. Data was collected on the amount of rain per day (in millimeters), for 14 days, for two different places. The researcher noticed that place A appears to have greater variability in the rain amount, compared to place B. In this sense, they want to test the hypothesis that the variability in the amount of rain is significantly greater in A than in B. Perform the F-test to test this hypothesis.

The data are available in the List of Supplementary Exercises: tab Exercise 8.

9. (Source: Fávero and Belfiore, 2017, Cap. 7) It is desired to compare the average time for service (min) in 2 hospitals. To do this, a sample was collected with 20 patients in each hospital. Check if there are differences between the average times of waiting in two hospitals. Consider $\alpha = 1\%$.

Additionally, calculate the confidence intervals for the average time for waiting in the two hospitals.

The data are available in the List of Supplementary Exercises: tab Exercise 9.

10. (Source: Fávero and Belfiore, 2017, Cap. 7) A manufacturer of skimmed yogurts states that the amount of calories in each cup is 60 cal. To verify if this information is correct, a random sample with 36 cups is collected, observing that the average amount of calories is 65 cal with a standard deviation of 3.5. Apply the appropriate test and verify if the statement of the manufacturer is true, considering the level of significance of 5%.

11. In the food industry, for the preparation of a certain product, the standard amount of salt is 8g with a standard deviation equal to 1g. In the analysis of a sample of 85 products, an average amount of salt was observed of 8.25g. Evaluate if the amount observed in the sample can be considered statistically major, at the level of significance of 5%.

12. In a sample of 425 measurements, it is observed that the average weight of the products that were measured is 226g, with a standard deviation of 10g. Establish the confidence interval for this average, knowing that the standard deviation observed is exactly its historical value. Consider a level of confidence of 95%.

Reference

Fávero, Luiz Paulo; Belfiore, Patrícia. (2017). Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®. Rio de Janeiro: Elsevier