Supplementary Material

Pre-test

English Version of the Previous Knowledge Pre-test and its Scoring Instructions (In blue)

Identification number:

1) The following shows the timing expressed in seconds for a group of 9 students in a 25 meters swimming competition: 12, 14, 16, 14, 18, 16, 13, 14, 18
   ¿What is the mean, the median, and the mode of the timing of this group?
   Mean: 15 → 1 point
   Median: 14 → 1 point
   Mode: 14 → 1 point

2) In the previous distribution, ¿What would the quartile 1 and the quartile 3 be?
   Quartile 1: 13.5 → 1 point
   Quartile 2: 17 → 1 point

3) In a different group of 5 students, the mean for the 25 meters swimming competition is 20 seconds. The incorporation of an additional member to this group changes the group mean to 19 seconds. What is the time of this additional student?
   a. 119 seconds
   b. 19 seconds
   c. 14 seconds → 1 point
   d. 15 seconds
   e. None of the above

4) What is the most appropriate measure of central tendency to address the performance of this group of students? Why?
   a. La median, because it is an ordinal variable.
   b. The mean, because it is an ordinal variable.
   c. The median, because it is an interval variable.
   d. The mean, because it is a ratio variable. → 1 point
   e. None of the above.

1) The heart rate per minute of a group of 20 adults is displayed in the dot diagram below. For example, 3 adults have a rate of 60 beats per minute. Based on this data set, how many individuals from a <u>similar group of 40 adults</u> would be expected to have a heart rate of <u>at least</u> <u>90 beats</u> per minute?

   a. 12 adults → 1 point
   b. 6 adults
   c. 3 adults
   d. 28 adults
   e. None of the above

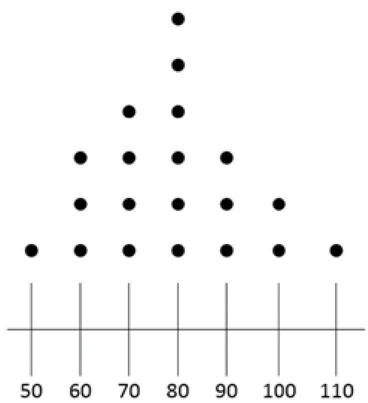


Figure S1: Study Design

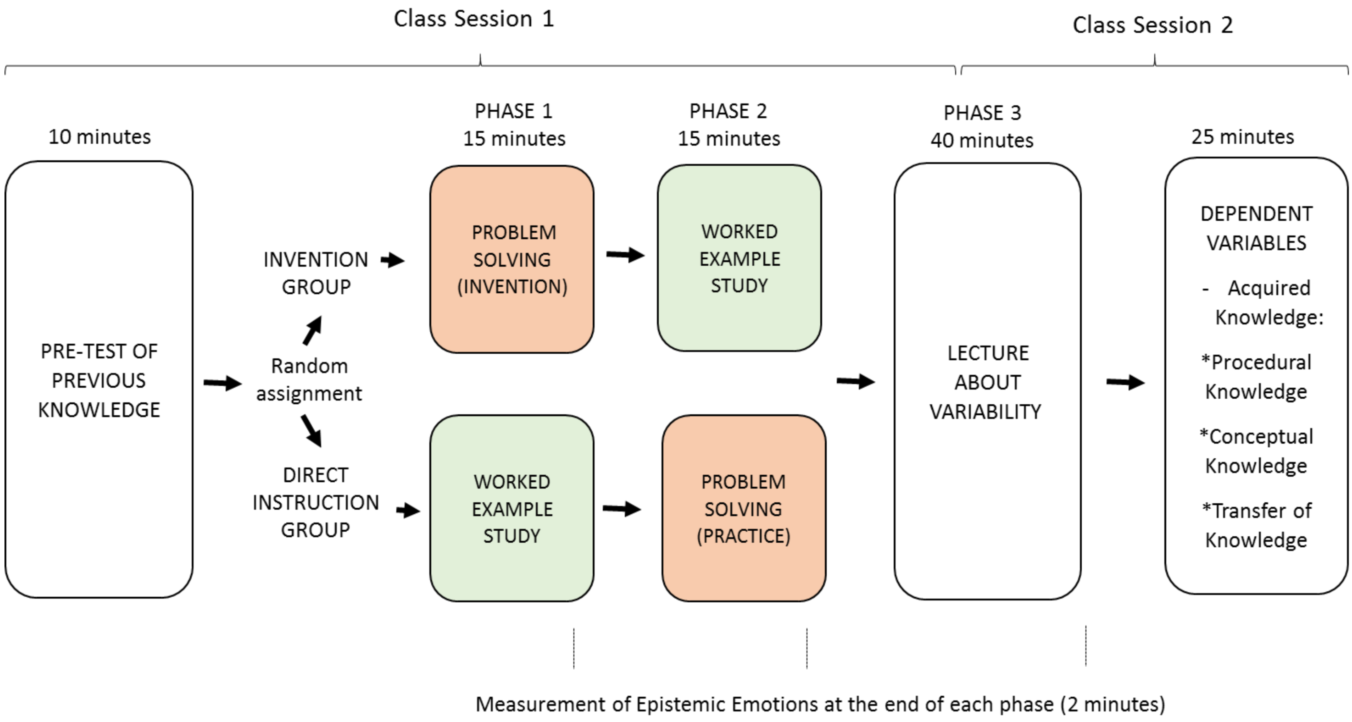
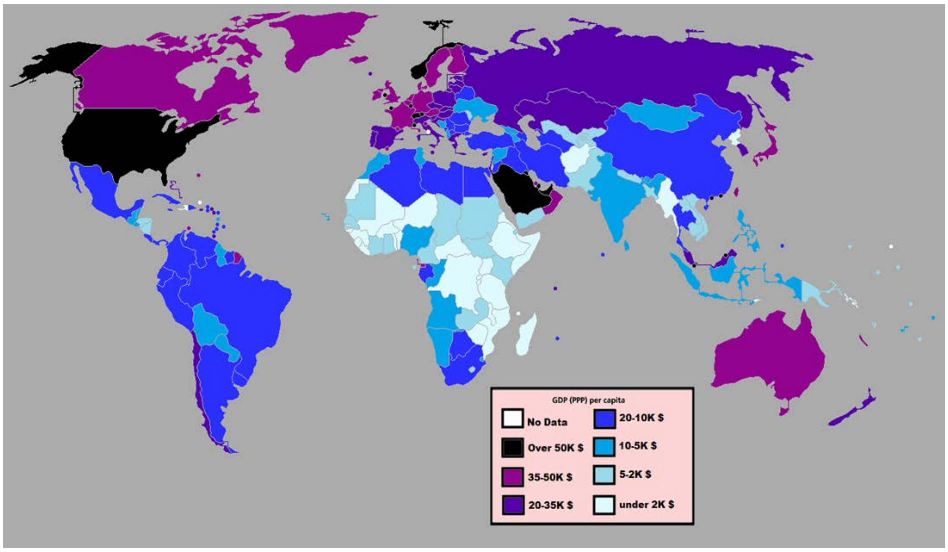


Figure S2. Worked Example

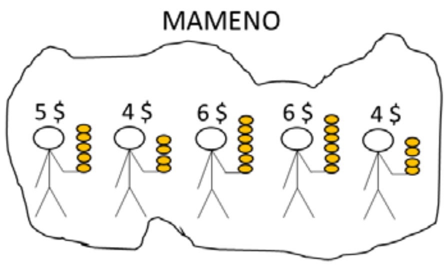


This image shows the average income (GDP per capita) of the countries in 2012. We can see that, for example, Spain and Russia have a similar average income. However, is the average income alone enough to give us an idea of the wealth of the inhabitants?

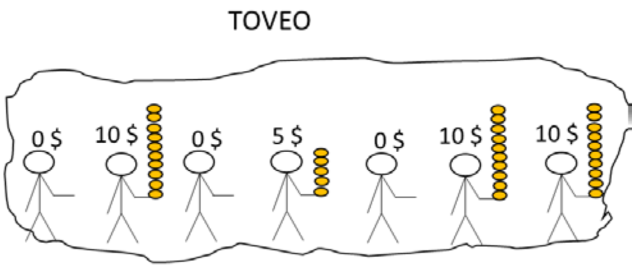
Below you can see four imaginary countries. Although the average income is the same in all of them, it is easy to see that they differ on how strongly inequality affects the wealth of their inhabitants.
How can we measure inequality?

There are multiple solutions that can help us measure inequality and compare the inequality of these four countries. In the following pages we will show you three of the most accepted solutions.

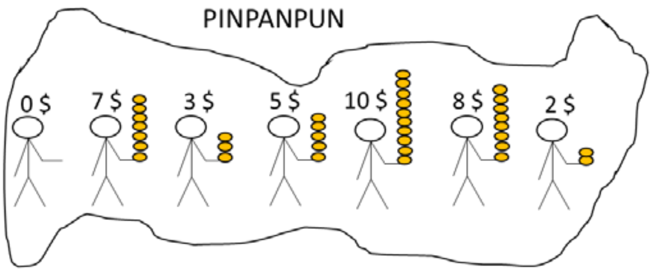
Try to learn their calculation procedures and comprehend their affordances and limitations.



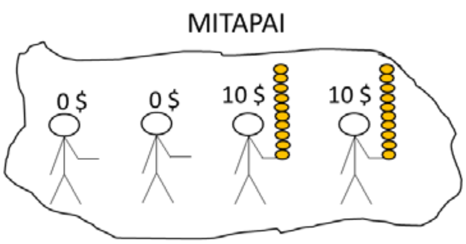
Average Income: (5+ 4 + 6 + 6 + 5) / 5 = **5**



Average Income: (0 + 10 + 0 + 5 + 0 +10 + 10) / 7 = **5**



Average Income: (0 + 7 + 3 + 5 + 10 + 8 + 2) / 7 = **5**



Average Income: (0 + 0 + 10 + 10) / 4 = **5**

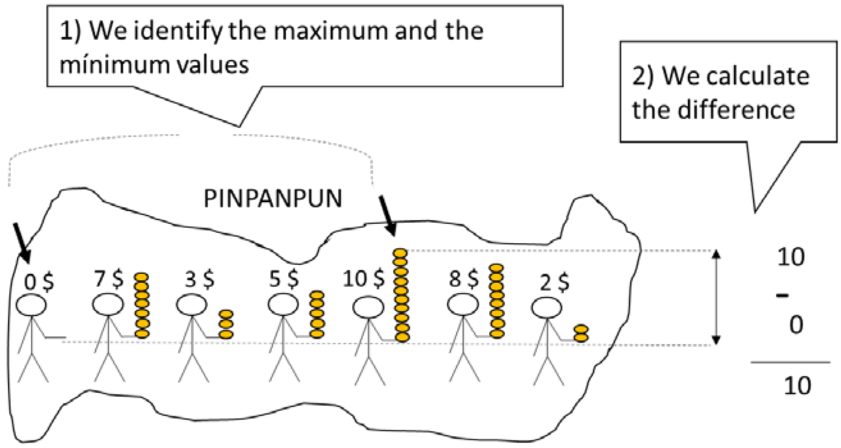
(continue to the next page)

SOLUTION 1: THE RANGE

- It is the difference between the maximum and the minimum value of the distribution.

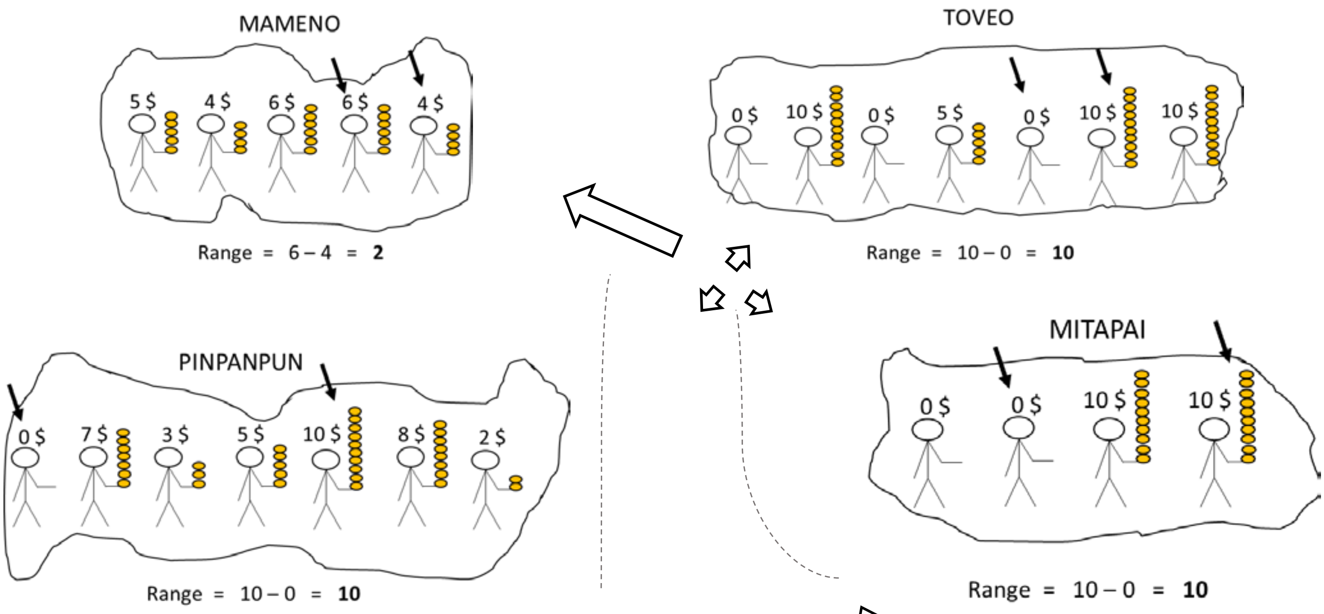
How is it calculated?

$$R = \max_i(x_i) - \min_i(x_i)$$



Range = 10

Solution:



ADVANTAGES OF THE RANGE:

It is easy to be calculated, and in some cases is a good indicator of inequality. For example, the lower range of Mameno certainly corresponds with a lower inequality in this country than in the rest of the countries.

LIMITATIONS:

It only takes into account the two extreme values of the distribution, the richest and the poorest person. Therefore, it is not possible to identify the different degree of inequality between Toveo, Mitapai and Pinpanpun, in which the extreme values are the same.

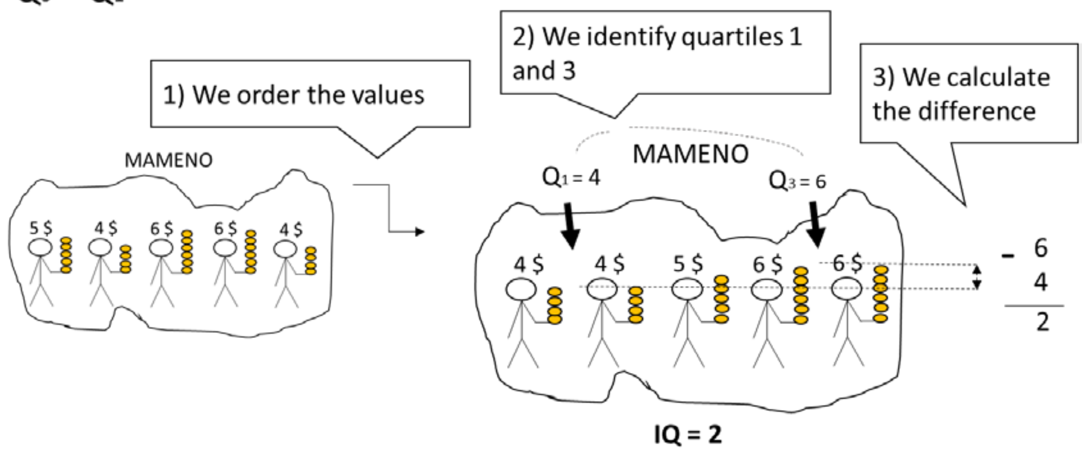
(continue to the next page)

SOLUTION 2: INTERQUARTILE RANGE (IQ)

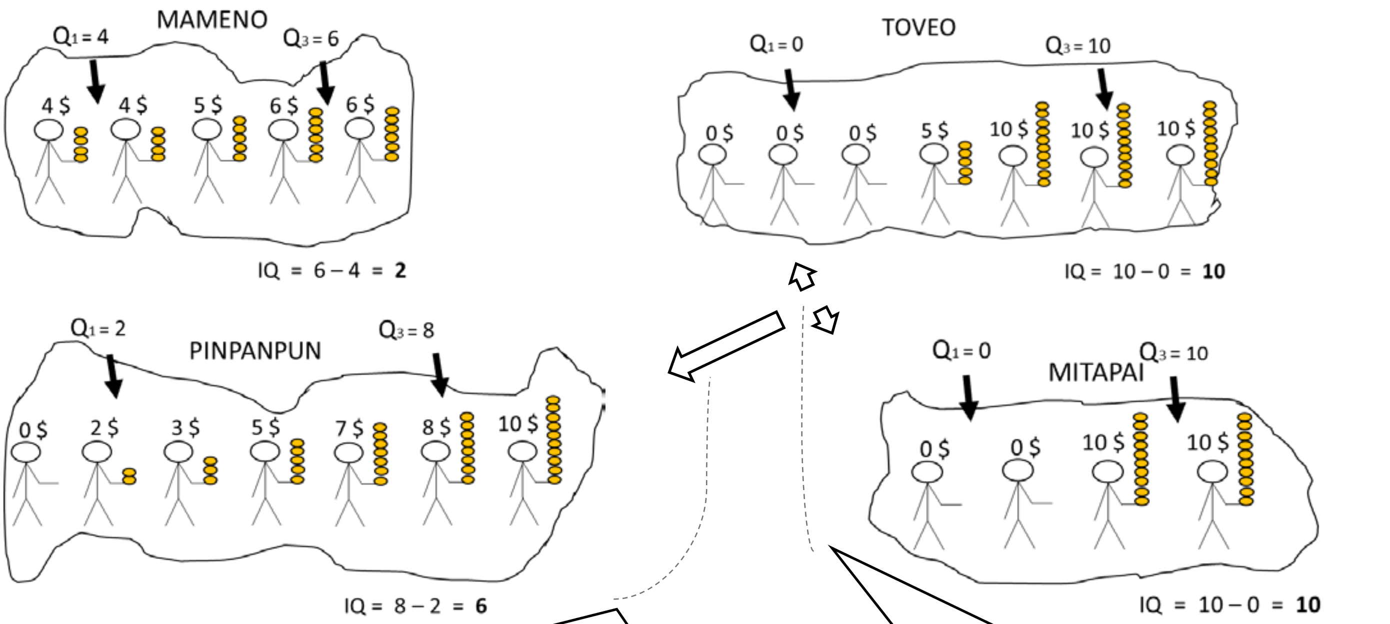
- It is the difference between quartile 3 and quartile 1. In other words, the amplitude of the range that contains the central half of the distribution.

How is it calculated?

$$IQ = Q_3 - Q_1$$



Solution:



ADVANTAGES OF THE INTERQUARTILE RANGE:

It is not distorted by extreme scores because it only considers the central differences.

It is relatively comprehensive. Although it only considers differences in the central half of the distribution, these differences are expected to be representative of differences in the whole distribution. It allows us to see that there is more equality in Pinpanpun (IQ = 6) than in Toveo and Mitapai (IQ =

LIMITATIONS OF THE INTERQUARTILE RANGE:

Its exhaustiveness can be improved because it does not consider the differences across all cases. That is why we cannot identify the different levels of inequality between Toveo and Mitapai, where Q_1 and Q_3 are the same.

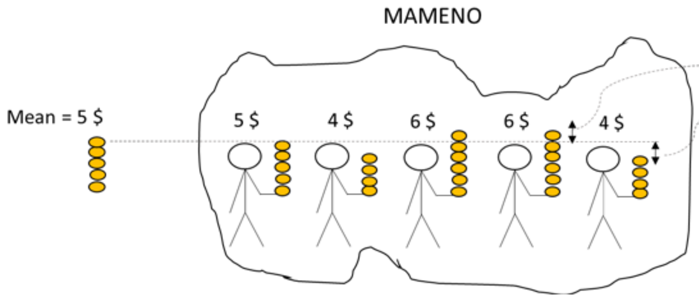
(continue to the next page)

SOLUTION 3: THE STANDARD DEVIATION

- It is an average of the differences from the mean.

How is it calculated?

1) We calculate the differences from the mean $x_i - \bar{x}$



$6 - 5 = \mathbf{1}$

$4 - 5 = \mathbf{(-1)}$

$6 - 5 = \mathbf{1}$

$4 - 5 = \mathbf{(-1)}$

$5 - 5 = \mathbf{0}$

2) We squared them. $(x_i - \bar{x})^2$

This step is necessary to convert the differences in possitive values. Otherwise, when summing them up in the next step the negative values would supress the positive values.
Ej:
1 + (-1) + 1 + (-1) + 0 = 0

$1^2 = \mathbf{1}$

$(-1)^2 = \mathbf{1}$

$1^2 = \mathbf{1}$

$(-1)^2 = \mathbf{1}$

$0^2 = 0$

3) We average the differences: We first sum them up, and then we divide the total by the number of cases. With it we obtain the variance (σ^2)

$$\frac{\Sigma (x_i - \bar{x})^2}{N}$$

$$\frac{1+1+1+1+0}{5} = \frac{4}{5} = \mathbf{0{,}8}$$

$\boldsymbol{\sigma^2 = 0{,}8}$

4) Because before we have squared the differences, we have to invert this process calculating the square root. With it we obtain the standard deviation (σ).

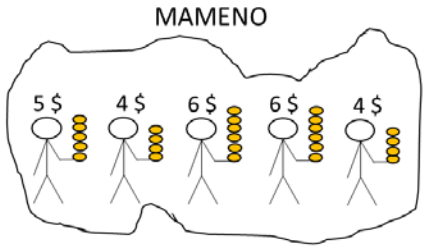
$$\sqrt{\frac{\Sigma (x_i - \bar{x})^2}{N}}$$

$\sqrt{0{,}8} = \mathbf{0{,}89}$

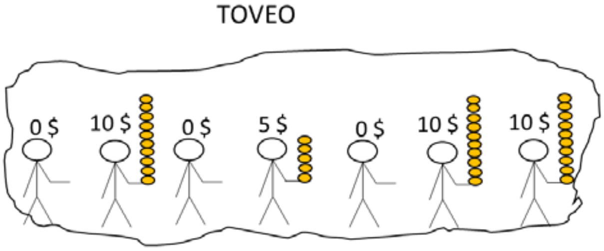
$\boldsymbol{\sigma = 0{,}89}$

(continue to the next page)

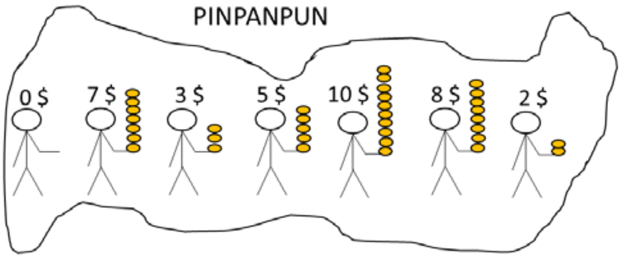
Solution:



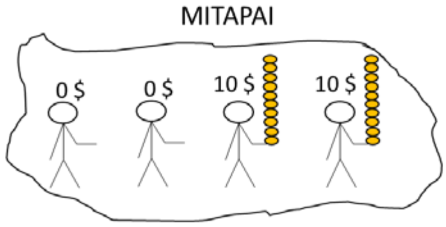
$$\sigma = \sqrt{\frac{(5-5)^2+(4-5)^2+(6-5)^2+(6-5)^2+(4-5)^2}{5}} = \mathbf{0{,}89}$$



$$\sigma = \sqrt{\frac{(0-5)^2+(10-5)^2+(0-5)^2+(5-5)^2+(0-5)^2+(10-5)^2+(10-5)^2}{7}} = \mathbf{4{,}62}$$



$$\sigma = \sqrt{\frac{(0-5)^2+(7-5)^2+(3-5)^2+(5-5)^2+(10-5)^2+(8-5)^2+(2-5)^2}{7}} = \mathbf{3{,}29}$$



$$\sigma = \sqrt{\frac{(0-5)^2+(0-5)^2+(10-5)^2+(10-5)^2}{4}} = \mathbf{5}$$

ADVANTAGES OF THE STANDARD DEVIATION:

It is a very comprehensive indicator, as it considers the differences of all the cases in the distribution. This allow us to identify the different degrees of inequality of the 4 countries, even when they are very subtle, for example when we compare Mitapai and Toveo.

LIMITATIONS OF THE STANDARD DEVIATION:

Although in general it is quite resistant to extreme values, it also considers them in its calculation. This is the reason why, when we find very asymmetric distributions, where there are very extreme values towards one side of the distribution, the standard deviation can be distorted.

It is not always appropriate to be used with ordinal variables. In ordinal variables the differences between values are uncertain, and the standard deviation is based on the calculation of these exact differences.

When we are in any of these situations, we can use the interquartile range.

Only if you have finished studying these contents, in the reverse of the page you can practice developing the procedures for the calculation of the inter-quartile range and standard deviation in Toveo, Pinpanpun, and Mitapai. Do your solutions correspond with the solutions given here?

Lecture Description

The power-point to support the lecture can be seen in https://www.dropbox.com/sh/aa6p3hs8esyf5xa/AACTvpVlEbdEtLVfBIbe9j7aa?dl=0. This file includes animations to stagger the contents, comments with proposed explanations to give to students, and indications about the approximative time allocated for each explanation. Also, a general description of these aspects is provided below:

1. Introductory explanation of the concept of variability, the different types of variability measures, their complementation with central tendency measures, and the use of all of them depending on the different type of variables (slides 1-5, approximate time: 3 minutes).

2. Introductory explanation of the standard deviation, when we use it and why (slides 6-8, approximate time: 4 minutes).

3. The students are given time to calculate the standard deviation for the distributions in the following problem. They can refer back to the explanations in the Worked Example (slide 9, approximate time: 5 minutes).

PROBLEM: The scores obtained by two groups of 5 students in a validated test of knowledge are:

- Group A: 4, 5, 6, 7, 8
- Group B: 4, 6, 6, 6, 8

In which group of students is there more variability? Justify your answer with the calculation of a variability measure

4. Feedback on the calculations and interpretation of results in the previous problem (slides 10-11, approximate time: 2 minutes).
5. Feedback on incorrect or partial calculations of the standard deviation (slides 12-21, approximate time: 7 minutes).
6. Short explanation of the different symbols used to refer to standard deviation depending on whether we refer to samples or populations (slide 22, approximate time: 1 minute).
7. Short explanation of the formula used in SPSS or other inferential statistics programs to calculate the standard deviation: the unbiased standard deviation (slide 23, approximate time: 1 minute).
8. Explanation of the properties of the standard deviation (slides 24-25, approximate time: 2 minutes).
9. Introductory explanation of the interquartile range: when and why we use it (slides 26-30, approximate time: 5 minutes).
10. Students are given time to calculate the interquartile range of the distributions in the following problem. They can refer back to the explanations in the Worked Example (slide 31, approximate time: 3 minutes).

PROBLEM: The scores obtained by 2 groups of 7 students in a exam are:

GRUPO A	GRUPO B
Outstanding (3)	Outstanding (3)
Failed (0)	Failed (0)
Very Good (2)	Outstanding (3)
Honors (4)	Very Good (2)
Very Good (2)	Failed (0)
Aprobado (1)	Failed (0)
Very Good (2)	Outstanding (3)

¿In which group of students there is more variability? Justify your answer with the calculation of a variability measure

11. Feedback on the calculations and interpretation of the interquartile range from the previous problem, including the discussion of the partial solution of the range (slides 32-33, approximate time: 2 minutes).
12. Explanation of the property of resistance to extreme values and the appropriateness of the interquartile range when we deal with ratio or interval variables in asymmetric distributions (slide 34, approximate time: 2 minutes).
13. Introductory Explanation to the Coefficient of Variation: How we calculate it, when and why we use it (slides 35-37, approximate time: 3 minutes).
14. Overview of the contents covered (slide 38, approximate time: 1 minute).

Learning Post-test

English Version of the Learning Post-test and its Scoring Instructions (In blue)

Identification number:

Please answer the following questions:

1) Calculate the standard deviation (SD) of the following set of marks on a test: 3, 6, 5, 4, 7
   a. 1
   b. $\sqrt{2}$ → 1 point
   c. 2
   d. $\sqrt{1}$

2) The owners of two cinemas, A and B, argue that their respective cinema enjoys a more consistent attendance. They collected the daily attendance of their cinemas for 11 random days. The results of their data collection are shown below:

	Cinema A	Cinema B
Mean, M	*72*	*75*
Standard Deviation, SD	*10*	*14*

Which cinema do you think presents a <u>more consistent</u> attendance?

   a. Cinema A → 1 point
   b. Cinema B
   c. Both enjoy equally consistent attendance.
   d. None of the above.

3) Below we can see the data from 2 distributions. In which distribution there is more variability?

	Distribution A	Distribution B
Median	*6*	*8*
Interquartile Range	*3*	*2*

   a. In distribution A → 1 point
   b. In distribution B
   c. Both have the same variability
   d. Not enough information to decide

4) In calculating the standard deviation, why is it important to divide the sum of squared deviations by *n*?

Assign 1 point if mentioned the need to average differences, or made reference to the idea that otherwise the indicator would be contaminated by the number of cases.

Assign 0.5 points if just mentioned a general need for averaging

Assign 0 points if none of the above.

5) As we can see in the following table, the standard deviation is greater in group A than in group B, while the interquartile range is greater in group B than in group A. How can it be explained?

	Group A	Group B
Standard Deviation	5,37	2,77
Interquartile Range	2	4

a) In the Group A there must be more extreme scores than in Group B.
b) In the Group B there must be more extreme scores than in Group A. → 1 point
c) The interquartile range indicates the central tendency, and the standard deviation indicates the variability, and therefore they do not have to coincide.
d) It is imposible, both are measures of variability and they have to provide consistent results.

6) Consider the following six datasets:

a. (1, 5, 6, 10)
b. (4, 4, 4, 4)
c. (101, 102, 103, 104)
d. (7, 8, 9, 10)
e. (1, 2, 9, 10)
f. (1, 2, 3, 4)

6.1) Which dataset has the smallest SD? B → 1 point
6.2) Which dataset has the largest SD? E → 1 point
6.3) Which datasets have the same SD? C, D, and F → 1 point

7) A data set consisting of <u>five</u> numbers has mean, M = 7, and standard deviation, SD = 4. If each of the five numbers is increased by 2, what are the new mean and SD?

a. M = 7, SD = 4
b. M = 9, SD = 4 → 1 point
c. M = 7, SD = 6
d. M = 9, SD = 6

8) An equal number of students competed in the 100m sprint and 100m swim finals. The timings (in seconds) of the champions of the 100m sprint and 100m swim are shown below, as are the average timings and the SDs of the finalists in the two competitions.

	100m sprint	100m swim
Champion	10s	40s
Average of the Finalists, M	12s	45s
SD of the Finalists	1s	3s

Assuming all else being equal, between the two champions, who is the better performer?

a. The sprint champion → 1 point
b. The swim champion
c. Both
d. Not enough information to decide

9) David's scores for Mathematics, Physics and Chemistry in the final examinations are given below. His class's performance for the three subjects is also given below:

	Mathematics	Physics	Chemistry
David's Score	95	90	85
Class Average, M	80	80	80
Class SD	15	5	4

9.1) Relative to his class, in which subject did David perform the best? Physics → 1 point

9.2) Relative to his class, in which subject did David perform the worst? Mathematics → 1 point

The total score of this posttest can vary from 0 to 12 points, and it will constitute the general acquired knowledge measure. This measure can be disaggregated in other learning measures:

- The procedural knowledge score can vary from 0 to 3, and it will be calculated summing up scores in items 1-3.

- The conceptual knowledge score can vary from 0 to 6, and it will be calculated summing up scores in items 4-7.

- The transfer of knowledge score can vary from 0 to 3, and it will be calculated summing up scores in items 8-9.