Statistics Package Exercise: Calculating the Standard Deviation

Learning Objective: Relate measures of center and spread to the shape of the distribution, and choose the appropriate measures in different contexts.

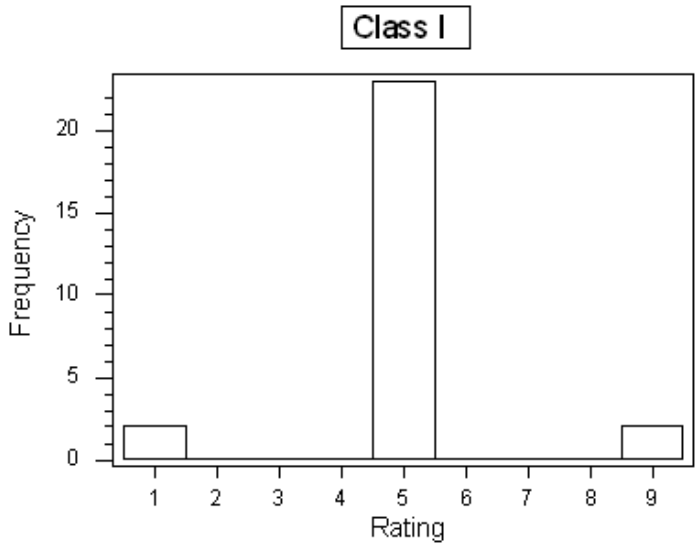
The concept of standard deviation is less intuitive as a measure of spread than the range or the IQR. The following activity is designed to help you develop a better intuition for the standard deviation.

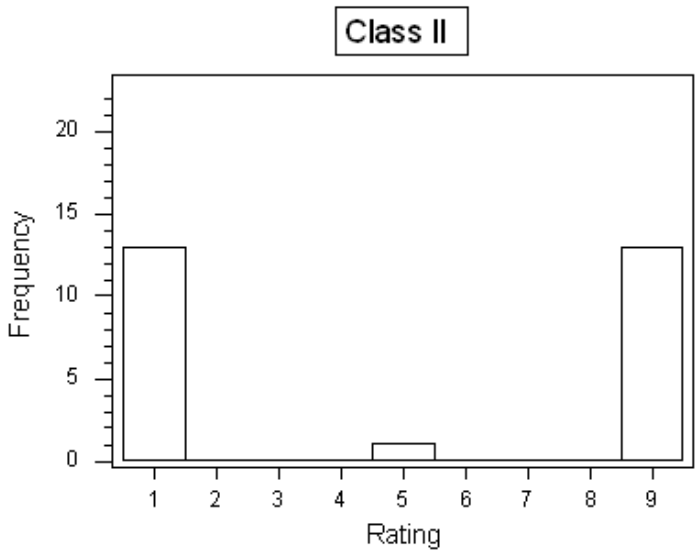
Background

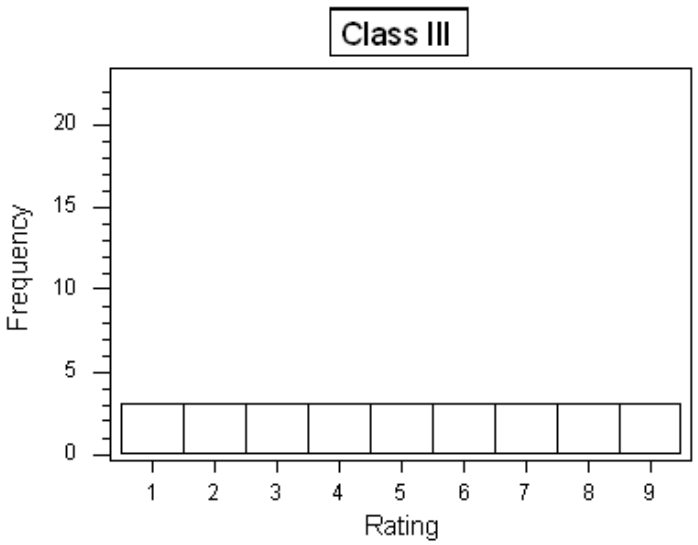
At the end of a statistics course, the 27 students in the class were asked to rate the instructor on a number scale of 1 to 9 (1 being "very poor," and 9 being "best instructor I've ever had"). The following table provides three hypothetical rating data:

Rating	1	2	3	4	5	6	7	8	9
Class I	1	0	0	0	22	0	0	0	1
Class II	12	0	0	0	1	0	0	0	12
Class III	2	2	2	2	2	2	2	2	2

And here are the histograms of the data:







Learn By Doing (1/1 point)

Assume that the average rating in each of the three classes is 5 (which should be visually reasonably clear from the histograms), and recall the interpretation of the SD as a "typical" or "average" distance between the data points and their mean. Judging from the table and the histograms, which class would have the largest standard deviation, and which one would have the smallest standard deviation? Explain your reasoning.

Your Answer:

Revised answer:
LOL I looked at the diagram wrong. I forgot that the x-axis was the rating, and was kinda looking at the frequency by mistake. But if I looked at it correctly, I think I would've arrived at the correct conclusion.

My orig answer:

Our Answer:
In class I, almost all the ratings are 5, which is also the mean. The average distance between the observations and the mean, then, would be very small. In class II most of the observations are far from the mean (at 1 or 9). The average distance between the observations and the mean in this case would be larger. Class III is the case where some of the observations are close to the mean, and some are far, so the average distance between the observations and the mean would be somewhere in between class I and II. This observation would lead me to conclude that the standard deviation would be ranked (from smallest to largest): Class I, Class III, Class II.

Resubmit Reset

Now check your intuition by finding the actual standard deviations of the three rating distributions.

- R StatCrunch TI Calculator Minitab Excel

R Instructions
To open R with the dataset preloaded, right-click here and choose "Save Target As" to download the file to your computer. Then find the downloaded file and double-click it to open it in R.

The data have been loaded into the variable

```
ratings
```

which has three variables (column titles)

```
Class.I
```

,

```
Class.II
```

, and

```
Class.III
```

. Enter the command

```
ratings
```

to see the data.

Now you can calculate the standard deviations for each variable with the command:

```
sapply(ratings, sd)
```

Note: Using R-Notice that R understands that your three columns represent different data sets with different names, and R computes the standard deviations of the three data sets separately. The above code can easily be modified for other commands for descriptive statistics such as

```
mean
```

or

```
median
```

. If the data frame is formatted differently, consider the

```
aggregate()
```

function instead of

```
sapply()
```

.

Learn By Doing (1/1 point)

What are the standard deviations of the three rating distributions? Was your intuition correct?

Your Answer:

My initial intuition before viewing the answers and realizing i was looking at the chart wrong was definitely wrong hahaha. But amending my understanding, yeah, it does sound intuitively correct.

Class I =1.57,
Class II = 4.00,
Class III = 2.63

Our Answer:

Here are the three standard deviations:Class I: 1.6Class II: 4.0Class III: 2.6Note that through this example, we also learn that the number of distinct values represented in a histogram does not necessarily indicate greater variability.

Resubmit Reset