

Measurement and Summary Statistics Practice

Part A. Types of Variables

Complete the following table.

	Example 1	Example 2	Example 3	Example 4
Variable:	Support for Democracy	Interest in Politics	Presidential Vote Choice in 2016 (US)	Party Identification
Conceptual Definition of Variable:				Party identification is the political party with which someone has a strong psychological attachment; it is not necessarily the political party with which someone is registered.
Operational Definition of Variable:		Amount of time spent reading political news, discussing politics, and/or engaging in political activities.	The candidate for whom someone reports voting on a survey.	
Values of Variable:				
Type of Variable: Nominal, Ordinal, or Interval	Ordinal			

Part B. Summary Statistics

Below are the results from the top five finishers of the 2017 Nathan's Hot Dog Eating Contest in which contestants try to eat as many hot dogs and buns as they can in 10 minutes. (Men's results)

Name	Number of hot dogs and buns eaten
Chestnut	72
Cincotti	60
Stonie	48
Breeden	38.5
Esper	44.5

1. What is the median number of hot dogs consumed?
2. What is the mean number of hot dogs consumed?
3. What is the variance for the number of hot dogs consumed?
4. What is the standard deviation for the number of hot dogs consumed?
5. What kind of variable is number of hot dogs consumed?
6. What are the values of the variable "number of hot dogs consumed" that we observe?

Additional Resources

Types of Variables (Nominal, Ordinal, Interval)

- *Note, there's yet another type of variable (ratio) that we did not discuss in class, but you'll see this in the examples that follow. Do not worry about ratio variables.
- Definitions, visualizations, examples
 - <https://stats.idre.ucla.edu/other/mult-pkg/whatstat/what-is-the-difference-between-categorical-ordinal-and-interval-variables/>
 - <https://www.graphpad.com/support/faq/what-is-the-difference-between-ordinal-interval-and-ratio-variables-why-should-i-care/>
 - <http://www.statisticshowto.com/nominal-ordinal-interval-ratio/>
- Videos
 - <https://www.youtube.com/watch?v=5OrOhxoWlis> (nominal)
 - <https://www.youtube.com/watch?v=hZxnfnt5v8> (good review of earlier concepts too. I strongly recommend viewing this one)
 - <https://www.youtube.com/watch?v=8iSbvun0tMc> (nominal)
 - <https://www.youtube.com/watch?v=U8JyBN9AkmY> (ordinal)
 - <https://www.youtube.com/watch?v=A5zlhbmBghI> (broader overview with an example)

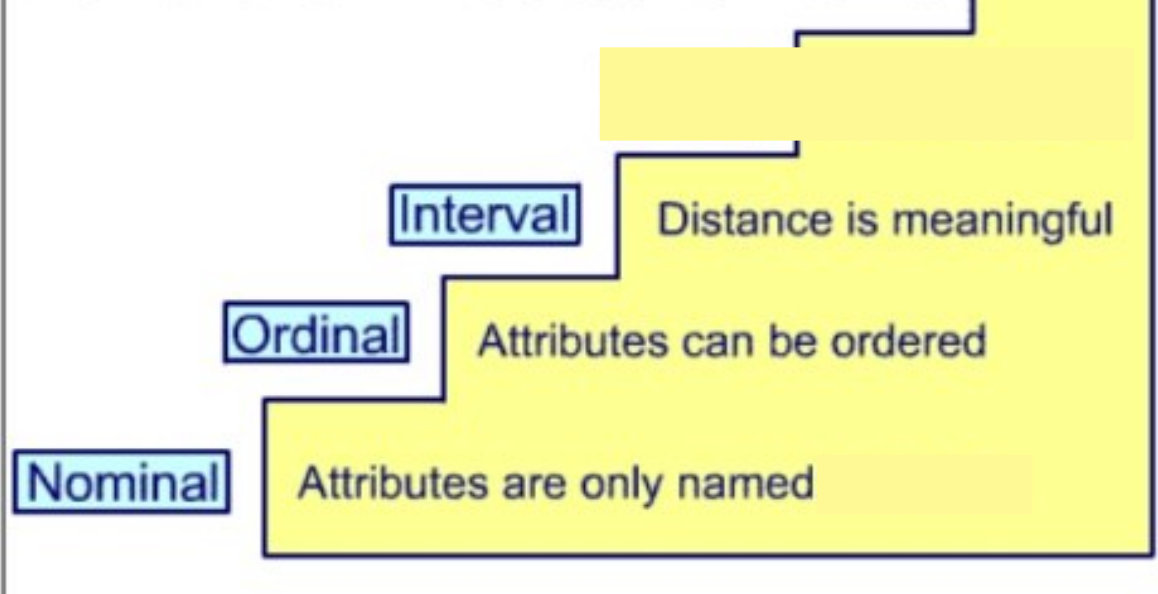
Conceptual vs. Operational Definitions

- Definitions, visualizations, examples
 - <http://www.sfu.ca/personal/archives/richards/Zen/Pages/Chap3.htm>
 - <http://willsull.net/la597/resources/Jan--24/5-Bernard.pdf>
- Videos
 - <https://www.youtube.com/watch?v=HNpMfAtiC5o>
 - https://www.youtube.com/watch?v=6E2_Bqncdo4
 - <https://www.youtube.com/watch?v=DhjHy0gmL38>

Summary Statistics

- Central Tendency Practice Problems with Answers and Explanations
 - https://www.khanacademy.org/math/probability/data-distributions-a1/summarizing-center-distributions/e/mean_median_and_mode
 - <http://faculty.webster.edu/woolfm/measct.html>
 - <https://study.com/academy/lesson/measures-of-central-tendency-definitions-practice.html>
- Dispersion
 - https://www.khanacademy.org/math/ap-statistics/summarizing-quantitative-data-ap/measuring-spread-quantitative/e/standard_deviation (don't worry about the difference between sample and population standard deviation)
 - <https://www.khanacademy.org/math/statistics-probability/summarizing-quantitative-data/variance-standard-deviation-sample/e/variance>
 - <https://www.ixl.com/math/algebra-1/variance-and-standard-deviation>
 - <https://onlinecourses.science.psu.edu/stat414/node/66>
 - <http://www.mathsisfun.com/data/standard-deviation.html>
 - <https://www.youtube.com/watch?v=ztfaPxCqGiE>
 - https://www.youtube.com/watch?v=qqOyy_NjflU

Levels of Measurement



Measurement and Summary Statistics Practice – Key

Part A. Types of Variables

Complete the following table.

	Example 1	Example 2	Example 3	Example 4
Variable:	Support for Democracy	Interest in Politics	Presidential Vote Choice in 2016 (US)	Party Identification
Conceptual Definition of Variable:	A society that supports democracy is one that emphasizes tolerance, trust, political activism, and post-materialist values, not just one that pays lip service to the ideals of democracy. (Ronald Inglehart)	Someone who is interested in politics values learning and talking about government, policy, politicians, campaigns, elections, and other political topics; s/he does not avoid learning about or participating in politics.	Vote choice is the candidate for whom a person votes for public office in an election; it is not the candidate for whom someone does not vote; nor is it a candidate for whom a person simply supports in a campaign, but does not vote in the election	Party identification is the political party with which someone has a strong psychological attachment; it is not necessarily the political party with which someone is registered.
Operational Definition of Variable:	Level of agreement with the following statement: “Is it good to have a strong leader who does not have to bother with parliament and elections?”	Amount of time spent reading political news, discussing politics, and/or engaging in political activities.	The candidate for whom someone reports voting on a survey.	The political party selected in response to this question: “Do you consider yourself to be a Republican, a Democrat, an Independent, or something else?”
Values of Variable:	Strongly disagree, disagree, neither agree nor disagree, agree, strongly agree	0 hours per week, 0.5 hours per week, 1 hour per week, 1.5 hours per week, 2.25 hours per week, ...	Donald Trump, Hillary Clinton, Other	Republican, Democrat, Independent, Something Else
Type of Variable: Nominal, Ordinal, or Interval	Ordinal	Interval	Nominal	Nominal

Part B. Summary Statistics

Below are the results from the top five finishers of the 2017 Nathan's Hot Dog Eating Contest in which contestants try to eat as many hot dogs and buns as they can in 10 minutes. (Men's results)

Name	Number of hot dogs and buns eaten
Chestnut	72
Cincotti	60
Stonie	48
Breeden	38.5
Esper	44.5

1. What is the median number of hot dogs consumed?
 - a. Begin by putting all of the observations in order from least to greatest: 38.5, 44.5, 48, 60, 72
 - b. Find the value of the observation in the middle: 48 (our third observation is in the middle since we have five total observations. The value of our third observation is 48).
2. What is the mean number of hot dogs consumed?

Recall the formula for mean:

$$\text{Mean} = \bar{X} = \frac{X_1 + X_2 + \dots + X_N}{N} = \frac{\sum X_i}{N}$$

Plug and Chug: X_1 corresponds to the first observation in our dataset (72), X_2 corresponds to the second observation in our dataset (60), and so on. Thus, we are going to add all of the values we observe and divide that total by the number of observations we have (5):

$$\frac{72 + 60 + 48 + 38.5 + 44.5}{5} = \frac{263}{5} = 52.6$$

3. What is the variance for the number of hot dogs consumed?

Recall the formula for variance:

$$\text{Variance} = \frac{\sum (X_i - \bar{X})^2}{N - 1}$$

To calculate variance, then, we're going to take the value of each observation (X_i) and subtract the mean from it. Then, we will square this value. We'll do this for each observation and then add all of these up. Then, we'll divide by the number of observations we have in our data (minus 1). I think that the easiest way to calculate variance is by constructing a table to organize your calculations as I'll show below.

X_i	Value Observed	Mean	$X_i - \text{Mean}$	$(X_i - \text{Mean})^2$
X_1	72	52.6	$72 - 52.6 = 19.4$	$19.4^2 = 376.4$
X_2	60	52.6	$60 - 52.6 = 7.4$	$7.4^2 = 54.8$
X_3	48	52.6	$48 - 52.6 = -4.6$	$(-4.6)^2 = 21.2$
X_4	38.5	52.6	$38.5 - 52.6 = -14.1$	$(-14.1)^2 = 198.8$
X_5	44.5	52.6	$44.5 - 52.6 = -8.1$	$(-8.1)^2 = 65.6$

Now, add up the calculations in the rightmost column. This gives you the numerator to the variance formula:

$$376.4 + 54.8 + 21.2 + 198.8 + 65.6 = 716.8$$

Finally, divide by the number of observations, five in our case, minus one (N-1):

$$\frac{716.8}{5-1} = \frac{716.8}{4} = 179.2$$

4. What is the standard deviation for the number of hot dogs consumed?

Recall that the standard deviation is the square root of the variance. Thus:

$$\sqrt{179.2} = 13.4$$

5. What kind of variable is number of hot dogs consumed?

Interval / Continuous

6. What are the values of the variable “number of hot dogs consumed” that we observe?

We observe: 72, 60, 48, 38.5, and 44.5. All possible values could include anything from 0 onward (though, the record is 72).