Elementary Statistics: Math 080

Jordan Hanson

July 10, 2020

Whittier College Department of Physics and Astronomy

Unit 0 Outline

1. Topics from Chapter 1: 1.1, 1.2, 1.3
    - What is a statistic?
    - Probability examples
    - Data and sampling
2. Topics from Chapter 2: 2.1 - 2.4, 2.5 - 2.8
    - Data visualization
    - Location of the data in numerical space
3. Topics from Chapter 3: 3.1, 3.2, 3.3
    - Two rules of probability

Topics from Chapter 2

Stemplots

Useful for numbers like *grades*. Most significant digit is the category.

Stem	Leaves
0	
1	
2	
3	
4	[3.0]
5	[6.0]
6	[7.0, 9.0]
7	[8.0, 0.0, 8.0, 1.0, 2.0, 5.0, 7.0]
8	[8.0, 3.0, 4.0, 6.0, 2.0, 1.0, 2.0, 1.0]
9	[8.0, 7.0, 1.0, 4.0]

Table 1: A *stemplot* of a grade distribution.

Stemplots

Procedure:

1. Identify the approximate order of magnitude of the sample.
2. Within that order of magnitude, create ≈ 10 *stems*, corresponding to the base-10 digits.
3. For each data point, call the non-most significant digits the *leaves* and drop the leaves in the category with the matching leaf.

Professor example: What is the stemplot of

```
[11, 22, 33, 44, 55, 66]
```

Stemplots

Procedure:

1. Identify the approximate order of magnitude of the sample.
2. Within that order of magnitude, create ≈ 10 *stems*, corresponding to the base-10 digits.
3. For each data point, call the non-most significant digits the *leaves* and drop the leaves in the category with the matching leaf.

Let's create a stemplot of:

1. Our ages in MATH080
2. My age and the rest of my department

(Stemplots lead in to the topic of histograms)

Histograms

Histograms are a tool for measuring *probability distributions*. The inputs are the data points and the corresponding relative frequencies, or plain frequencies.

How many textbooks or books did you purchase for school last year? (Type in the chat).

1. Determine the bins, or *binning*
2. For each data point, drop it into the appropriate bin
3. Each time a measurement is dropped into a bin, the *count* increases by 1.
4. If a histogram displays plain frequencies, it is called *un-normalized.*
5. If a histogram displays relative frequencies, it is called *normalized.*

Histograms

1. Histogram of books, by hand
2. Repeat with Excel/Calc

Practice with the FREQUENCY function in Calc/Excel:

```
=FREQUENCY(A1:A99; B1:B11)
```

Then press **control+shift+enter** to execute on arrays of data and bins. To *normalize*, input the relative frequencies, or divide frequecies by *N*. Assume the data is in C column:

```
=C1/N ...
```

Conclusion

Unit 0 Outline

1. Topics from Chapter 1: 1.1, 1.2, 1.3
    - What is a statistic?
    - Probability examples
    - Data and sampling
2. Topics from Chapter 2: 2.1 - 2.4, 2.5 - 2.8
    - Data visualization
    - Location of the data in numerical space
3. Topics from Chapter 3: 3.1, 3.2, 3.3
    - Two rules of probability