Math 142: Monday, Aug 26th, Day 3 Notes

Agenda for Monday, Aug 26, 2024

Reminders

- ! Canvas Homeworks due } today & Wed by 11:59p
- First Group Quiz on Wednesday, August 28th in class on chapter 1

Recall:

Chapter 2 Displaying Data

Chapter 2 Overview

- We can summarize and describe data values in a variety of ways.
- You'll probably recognize these displays and summaries.
- This chapter is a fast review of these concepts so we all agree on terms, notation, and methods.

Bar charts
Pie charts
Frequency Distributions - today
Histograms
Stem and Leaf Plots
scatter Plots
etc...

Section 2.1 Frequency Distributions

Frequency distributions are used to display both numerical and categorical variables

A frequency distribution is a table with two columns

Data values	Frequency

→ how many times does the data value appear in the data set

↙ can be qualitative or quantative

Summarizing and Displaying a Categorical Variable with a Frequency Distribution

Categorical variables are easy to summarize in a frequency table that lists the number of cases in each category along with its name.

Example 1. The students in my Mth 142 classes were polled as to whether they plan to transfer to a university or not. Construct a frequency distribution and a relative frequency distribution for the responses below.

Data: Maybe, Maybe, Maybe, Yes, Maybe, Yes, Yes, Yes, Yes, Yes, Yes, Yes, Yes Yes, Yes, Maybe, Maybe, Yes, Maybe, Yes, Yes, Yes, Yes, Yes, Yes, Yes, Yes, Yes, Maybe, Yes, Yes, Maybe, Yes.

1. Frequency Distribution (or Frequency Table) for the data above.

Data Values	Frequency
Yes	23
Maybe	9

add up all frequencies and get the sample size

$n = 32$

2. Relative Frequency Distribution (or Relative Frequency Table) for the data above.

Data values	Relative frequency
Maybe	$23/32 = 0.71875 = 71.9\%$
Yes	$9/32 = 0.28125 = 28.1\%$

add up to 100%

relative frequency $=$ frequency / sample size

Summarizing and Displaying a Quantitative Variable with a Frequency Distribution

Example 2. How many siblings do you have?

Data: 5, 1, 2, 5, 2, 3, 3, 4, 15, 1, 1, 2, 1, 3, 1, 1, 1

1. Ungrouped Frequency Distribution

- used when data value range is small (or we have few possible data values)

range = highest value − lowest value

- Frequency Distribution

Data values	Frequency	Relative frequency	Cumulativ rel. freq
1	7	7/17 = 41.2%	
2	4	4/17 = 23.5%	
3	3	3/17 = 17.6%	
5	2	2/17 = 11.8%	
15	1	1/17 = 5.9%	

- Relative Frequency Distribution

n = 17

- Cumulative relative frequency distribution

Data values	Frequency	Relative frequency	cumulative rel-freq
1	7	7/17 = 41.2%	41.2%
2	4	4/17 = 23.5%	41.2 + 23.5 = 64.7%
3	3	3/17 = 17.6%	64.7 + 17.6 = 82.3%
5	2	2/17 = 11.8%	82.3 + 11.8 = 94.1%
15	1	1/17 = 5.9%	94.1 + 5.9 = 100%

When data values range is large it is not practical to list each value individually. We group the data into classes.

2. Grouped Frequency Distribution

Example 3. Consider the following grouped frequency distribution.

Test scores	f
63 - 69	3
70 - 76	7
77 - 83	9
84 - 90	5
91 - 97	1

5 classes

Class 1 : 63-69

Lower Limits - lowest value of each class

ex: what are the lower limits:

63, 70, 77, 84, 91

Upper Limits - highest value of each class

ex: 69, 76, 83, 90, 97

Class Width - difference between any two consecutive lower limits (or upper)

ex: 76 − 69 = 7

ex: $\frac{97-63}{5} = 6.8$ → up 7

(round up)

OR class width = $\frac{\text{range}}{\#\text{ of classes}}$

Class Midpoints

- are the halfway point in each class
- is the average of the lower and upper limit of the class

ex: midpoint for class 3?

$\text{midpoint} = \frac{77+83}{2} = 80$

for class 5 is 94

Class Boundaries

next class

Test scores	Boundaries	f
63 - 69		3
70 - 76		7
77 - 83		9
84 - 90		5
91 - 97		1

Guidelines for classes:

- There should be between 5-20 classes
- The classes must be mutually exclusive (means no overlap) no data value can be in more than one class
- The classes must be exhaustive – every data value must be in a class
- The classes must be continuous (no gaps)
- The classes must be equal width.

Group Activity 1

Get in groups of 3 or 4 and work on this example.

Example 4. Twenty randomly selected students were asked to estimate the number of hours that they had spent studying in the past week (in and out of class). The responses are recorded below.

15 58 37 42 20 27 36 57 29
42 51 28 46 29 58 58 43 40
58 36

Create a grouped frequency distribution and calculate boundaries for each class by following the steps below.

1. First order the data in ascending order. Use your calculator.

15, 20, 27, 28, 29, 29, 36, 36, 37, 40
42, 42, 43, 46, 51, 55, 56, 57, 58, 58

2. Find Range

range = highest − lowest = 58 − 15 = 43

3. Determine number of classes wanted

5 classes since it's a smaller data set

4. Find Class Width

$$\frac{\text{range}}{\#\text{ of classes}} = \frac{43}{5} = 8.6$$

round to 9 class width

5. Determine first lower limit

15

6. Use class width to find all the other lower limits

15, 15+9=24,

7. Use the lower limits to find upper limits

~~8. Find the class boundaries~~

Classes	~~Boundaries~~	frequency
15-23		2
24-32		4
33-41		4
42-50		4
51-59		6

add up to 20=n