

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

Agenda for Monday, Aug 26, 2024

Chapter 2 Displaying Data	2
Group activity	7

Reminders

- ! Canvas Homeworks due } wed HW1 HW2 by 11:59 pm
- First Group Quiz on Wednesday, August 28th in class
con 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 graphs
- Pie charts
- Time series graphs
- Scatter Plots
- Histograms
- Stem and Leaf Plots
- Pareto charts
- Frequency Distributions

Section 2.1 Frequency Distributions

Frequency distributions are used to display both numerical and categorical variables

Frequency distributions are tables with two columns:

Data values	Frequency
-------------	-----------

can be quantitative or qualitative

number of times each data value appears in the data set

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, 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
Maybe	9
Yes	23

Adding up all the frequencies gives us the sample size.

$$23 + 9 = 32 \text{ student}$$
$$n = 32 \text{ (} n \text{ is the sample size)}$$

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

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

$$\text{relative frequency} = \frac{\text{frequency}}{\text{sample size}}$$

adds up to 100%

Summarizing and Displaying a Quantitative Variable with a Frequency Distribution

Example 2. How many siblings do you have?

Data: 2, 1, 2, 1, 2, 3, 3, 1, 0, 2, 1, 1, 4, 1,
3, 2, 2, 2, 2, 0, 2, 1, 1, 4

1. Ungrouped Frequency Distribution - used when data value range is small.

range = highest value - lowest value

• Frequency Distribution

$n = 24$

Data values	Frequency	Relative frequency
0	2	$2/24 = 0.0833 = 8.3\%$
1	7	$7/24 = 0.2916 = 29.2\%$
2	10	$10/24 = 41.7\%$
3	3	$3/24 = 12.5\%$
4	2	$2/24 = 8.3\%$

• Relative Frequency Distribution

- Cumulative relative frequency distribution

Data values	Relative frequency	Cumulative relative frequency
0	8.3 %	8.3 %
1	29.2 %	$8.3 + 29.2 = 37.5 \%$
2	41.7 %	$37.5 + 41.7 = 79.2 \%$
3	12.5 %	$79.2 + 12.5 = 91.7 \%$
4	8.3 %	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.

5 classes

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

class 1: scores between
: 63-69
:

Lower Limits - lowest value of each class

63, 70, 77, 84, 91

Upper Limits - highest value of each class
69, 76, 83, 90, 97

⑦ **Class Width** - the difference between any two consecutive lower limits

ex: $70 - 63 = 7$

or $\frac{91 - 63}{5} = 5.4$

$= \frac{97 - 63}{5} = 7.4$

$= 6.8$
round up to 7

class width = $\frac{\text{range}}{\text{\# of classes}}$ OR (you round up)

Class Midpoints - halfway point of the class

class 1 midpoint: $\frac{63 + 69}{2} = \frac{132}{2} = 66$

class 3 midpoint: $\frac{77 + 83}{2} = 80$

Class Boundaries

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 - 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 55 43 40
56 36

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

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

next class

2. Find **Range**

$$\text{Range} = \text{highest value} - \text{lowest value} = 58 - 15 = 43$$

3. Determine number of classes wanted

5 classes (it's a smaller data set)

4. Find Class Width

$$\text{is } \frac{\text{range}}{\# \text{ of classes}} = \frac{43}{5} = 8.6 \text{ round up to } 9$$

5. Determine first lower limit

15 (to correspond to smallest data value)

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 $20 = n$