

Frequency Distributions

Chapter 2

Class Outline – 7/08/08

- Review of basic terms
- Frequency Distributions – Chap. 2
- Workshop #1

Chapter 2 – Frequency

- Need way to organize data

Subject	X	Y	Z
1	4	8	2
2	3	7	2
3	7	6	2
4	9	5	6
5	1	1	6
6	5	2	8
7	0	3	5
8	2	6	7
9	6	5	9
10	3	0	9
11	8	9	3
12	4	1	1
13	9	5	4
14	1	3	0

Frequency Distribution Tables

- Frequency Distribution:
 - Organized presentation of the number of individuals located in each category on the scale of measurement
 - Provides a picture of how the scales are *distributed*
 - Lets us see our data
- Two Components:
 - Set of categories that make up the measurement scale
 - In some order (i.e. highest to lowest)
 - All possible values of X , not just actual scores
 - The frequency (number of individuals) in each category

Frequency Distribution Tables

- Let's create one!
- Raw quiz scores ($N = 15$)
 - 8, 9, 5, 7, 3
 - 10, 6, 8, 5, 9
 - 9, 6, 4, 10, 7
- How well did most students do?
- Quick check ($\sum f = N$)

Calculations from Frequency Distribution Tables

- Always remember that a frequency distribution is not the same thing as the raw data
- To compute $\sum X$:
 - Take the data out of the frequency table
 - Same thing for $\sum X^2$
 - Alternatively, can use all info presented in table
 - X, f, Xf
 - Let's practice
 - How calculate $\sum X^2$?

Other Info in Tables

- Can include columns for *proportion* and *percentage*
- Proportion $p = f/N$
 - fraction, but more often decimals
- Percentage $\% = p(100) = f/N (100)$
 - whole numbers
 - cumulative percentage
- Let's add this info to our table

Grouped Frequency Distribution Tables

- What do you do when you have many values of X ?
 - Example = exam scores (max 100 points)
 - Table will be huge!
 - Purpose is to simplify data
 - So... divide scores into groups
 - Ex. 90's, 80's, 70's, etc.

Grouped Tables

- Rules:
 - 1: have about 10 class intervals (groups)
 - More is too cumbersome, less is oversimplified
 - 2: width of intervals should be simple number
 - 2, 5, 10, 20...
 - 3: bottom score in each interval should be a multiple of the width
 - Makes it easier to understand table
 - 4: all intervals should be same width
 - Cover range of scores completely with no gaps, so every score falls within an interval

Determining Interval Width

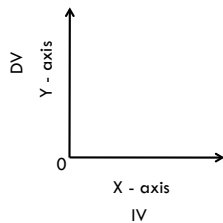
- First need to determine range of scores
 - ▣ Range = highest - lowest
- Next determine how many intervals needed at different widths
 - ▣ range/ width = # intervals needed (round up to whole number)
- Example, min = 5, max = 55
 - ▣ Range = ?
 - ▣ Width = 2, intervals = ?
 - ▣ Width = 5, intervals = ?

Determining Interval Width

- Statistical Approach
 - ▣ Exam scores (min = 42, max = 89)
 - ▣ Width = 5, # intervals = 10 (round up from 9.4)
- Conceptual Approach
 - ▣ Exam scores (min = 42, max = 89), but (min possible = 0, max possible = 100)
 - ▣ May want to look at frequency of letter grades
 - ▣ Width = 10, # intervals = 10
 - ▣ 90's, 80's, 70's represent frequency A's, B's, C's, etc.
 - ▣ Can also use with nominal data (ex. gender)

Frequency Distribution Graphs

- Graphing review:



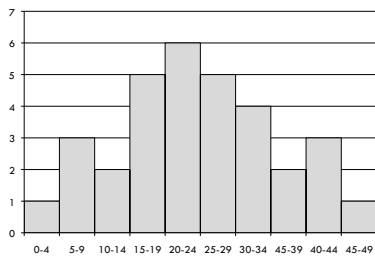
Histograms

- Used with interval or ratio data
- List numerical scores along X-axis
- Above each X value, draw a bar that corresponds to frequency of that value
- Width of each bar extends to *real limits* (no gaps between bars)

X	f
5	3
4	4
3	2
2	1
1	2

Let's try one:

Grouped Frequency Histogram



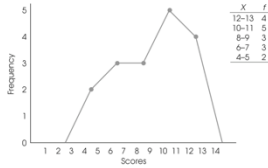
X	f
0-4	1
5-9	3
10-14	2
15-19	5
20-24	6
25-29	5
30-34	4
35-39	2
40-44	3
45-49	1

Polygons

- Used with interval or ratio data
- List numerical scores along X-axis
- Above each X value, place a dot corresponding to the frequency of that value
- Draw a straight line connecting each dot
 - line should touch x-axis at one interval below and above max interval
- Let's try one!
 - (use same frequency table as with histogram)

Grouped Frequency Polygon

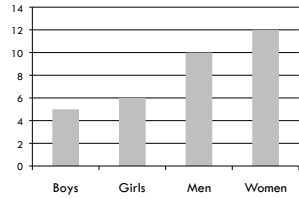
- List all values of X along X-axis (even though have grouped them)
- Place the dot in the center of each interval



- Alternately, can list intervals along X-axis, as we did with histograms

Bar Graphs

- Used with nominal or ordinal data
- Same as a histogram, but with spaces between bars
 - Represents distinct categories
 - Ordinal data can't assume equal interval

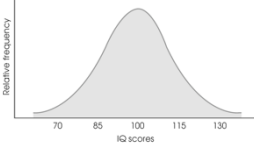


Population Distributions

- Exact Frequencies
 - Graph just as before with histogram, polygon, or bar graph
 - Ex. population = our class, exact frequency of grades is known
- Relative Frequencies
 - Exact frequencies aren't known because population is large
 - Ex. students in stats classes at Rutgers over past 4 years
 - Representative sample

Graphing Relative Frequencies

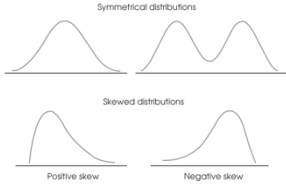
- Use smooth curve
 - Represents absence of exact scores



- Example of a normal distribution

Relative Frequency Distributions

- Central tendency
- Variability
- Shape of Distribution
 - Symmetrical
 - Bimodal
 - Positively Skewed
 - Negatively Skewed



Workshop #1

Break into groups of 3
You need your calculator!
You may use your book
