

Mathematics 231

Lecture 2
Liam O'Brien

1

Announcements

- Reading
 - Today M&M 1.1 6-16
 - Next class M&M 1.2 30-44

2

Describing Data

- Variables: Categorical versus quantitative
- Displaying Distributions with graphs
- Bar graphs, pie charts, histograms, stem-and-leaf plots
- Skewness

3

Variables and Distributions

- **Variable:** Any particular characteristic that can take on different “values” for each individual.
- **Examples:** Age, gender, GPA, major
- **Categorical variable:** Places individuals into one of several groups or categories.
- **Quantitative variable:** Takes numerical values for which arithmetic operations are meaningful.

4

Example: Survey Question

- What is your interest in this course?

```
+-----+
| interest |
+-----+
1. |      4 |
2. |      8 |
3. |      5 |
4. |      8 |
5. |      7 |
+-----+
6. |      . |
7. |      7 |
8. |      5 |
9. |      5 |
10. |     6 |
```

5

Distribution of a Variable

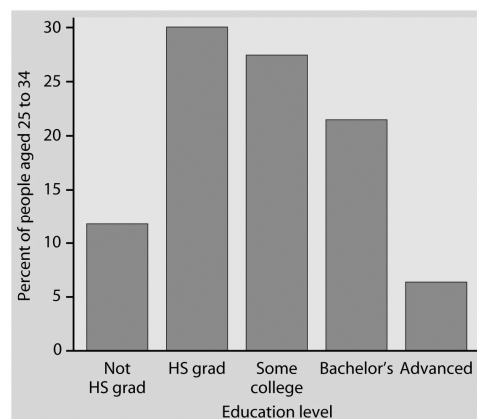
- **Distribution:** Describes what values a variable takes on, and how frequently these values occur.
- The distribution of a variable can be described through both graphics and numerical summaries.
- The three most used characteristics are **shape, center (or location), and spread (or variability)**.
- Shape: A picture is worth a thousand words.

6

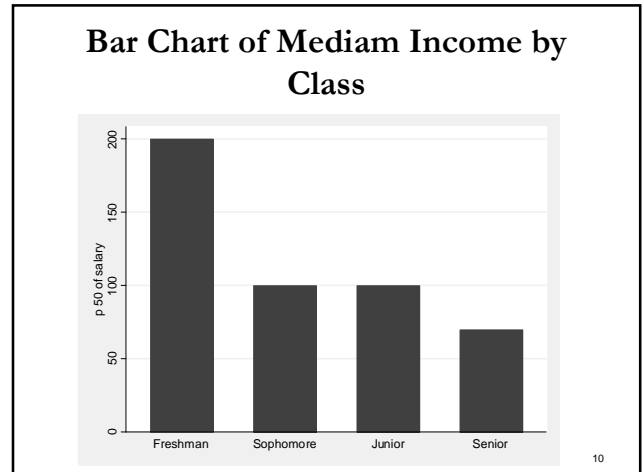
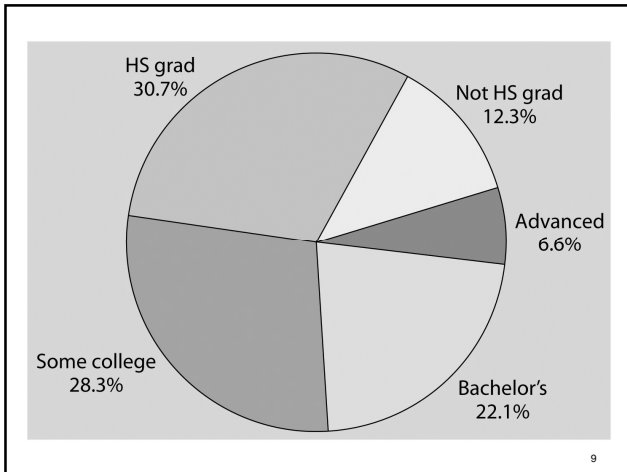
Graphs for Categorical Data

- **Bar graphs:** Display count/percentage of individuals in each category.
- **Pie charts:** Display percentages of individuals in each category as “wedges” in a “pie.”

7



8



Graphs for Quantitative Data

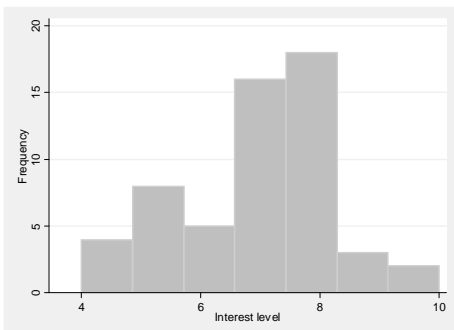
- **Histograms:** Display the distribution of values depicting the (relative) frequencies of observations within intervals (usually of equal size).
- **Boxplots:** Display important summaries of the distribution of values (discuss later).

Histograms

- A histogram does the following:
 1. Breaks the range of values into intervals of equal length.
 2. Displays the count/percentage of observations that fall into each interval (or "bin").
- Note that the number of intervals and interval width are important and can change the way the histogram looks dramatically for the same set of data.

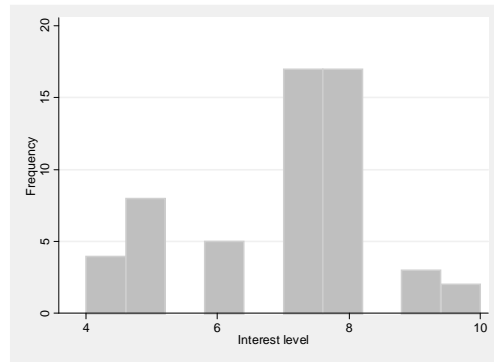
Example: Interest

- What is your interest in this course?



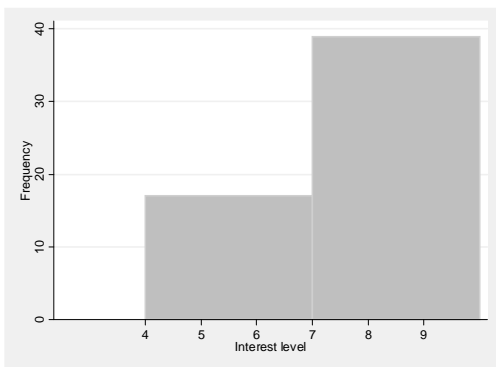
13

- Histogram – interest level from MA231

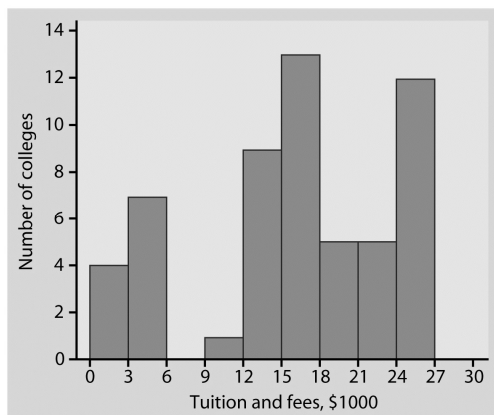


14

- Histogram – interest level from MA231



15

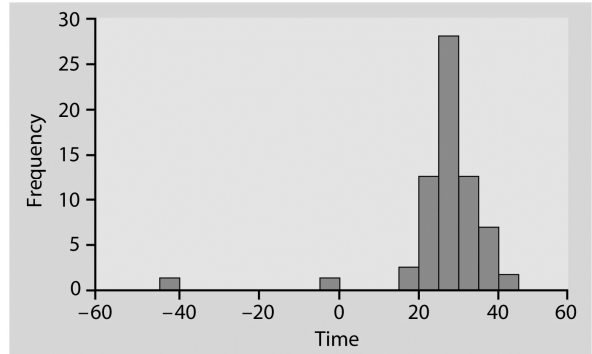


16

TABLE 1.1 Newcomb's measurements of the passage time of light

28	22	36	26	28	28
26	24	32	30	27	24
33	21	36	32	31	25
24	25	28	36	27	32
34	30	25	26	26	25
-44	23	21	30	33	29
27	29	28	22	26	27
16	31	29	36	32	28
40	19	37	23	32	29
-2	24	25	27	24	16
29	20	28	27	39	23

17



18

Shape

- The shape of the distribution can be described by visually inspecting it.
- We generally say it is either symmetric or skewed.
- If it is skewed, it can be skewed right or left.
- We also can state how many peaks it has (one peak = unimodal)

19

Symmetric Distributions

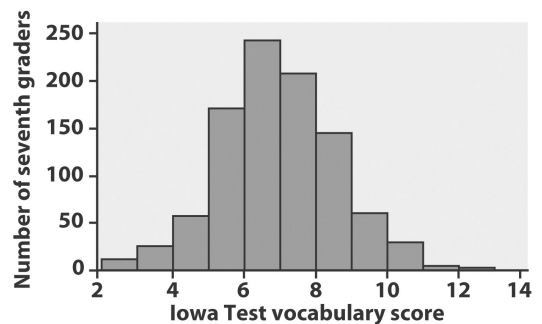


Figure 1-12
Introduction to the Practice of Statistics, Fifth Edition
© 2003 W.H. Freeman and Company

20

Skewed Distribution (right)

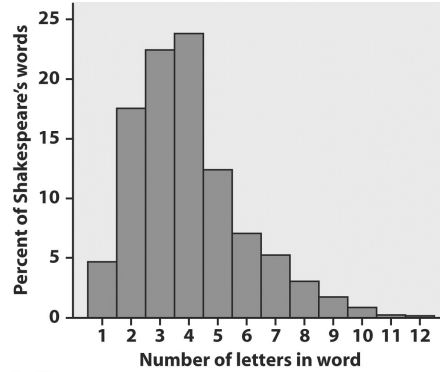


Figure 1-13
Introduction to the Practice of Statistics, Fifth Edition
© 2005 W. H. Freeman and Company

21

Stem-and-Leaf Plots

STEMPLOT

To make a **stemplot**:

1. Separate each observation into a **stem** consisting of all but the final (rightmost) digit and a **leaf**, the final digit. Stems may have as many digits as needed, but each leaf contains only a single digit.
2. Write the stems in a vertical column with the smallest at the top, and draw a vertical line at the right of this column.
3. Write each leaf in the row to the right of its stem, in increasing order out from the stem.

Definition, pg 11
Introduction to the Practice of Statistics, Fifth Edition
© 2005 W. H. Freeman and Company

22

TABLE 1.2

Literacy rates (percent) in Islamic nations

Country	Female percent	Male percent	Country	Female percent	Male percent
Algeria	60	78	Morocco	38	68
Bangladesh	31	50	Saudi Arabia	70	84
Egypt	46	68	Syria	63	89
Iran	71	85	Tajikistan	99	100
Jordan	86	96	Tunisia	63	83
Kazakhstan	99	100	Turkey	78	94
Lebanon	82	95	Uzbekistan	99	100
Libya	71	92	Yemen	29	70
Malaysia	85	92			

Table 1-2
Introduction to the Practice of Statistics, Fifth Edition
© 2005 W. H. Freeman and Company

23

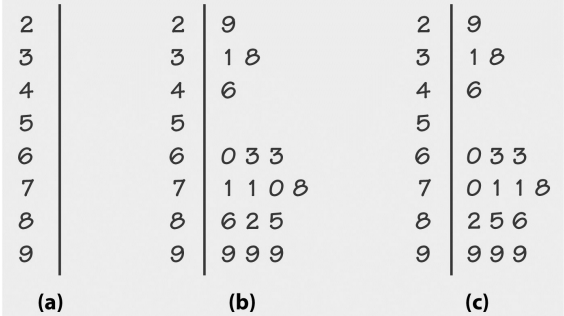


Figure 1-3
Introduction to the Practice of Statistics, Fifth Edition
© 2005 W. H. Freeman and Company

24