1 Review of the last lecture

- Instructor’s name: AKI.
- Visit TA/Instructor’s office by Sep 10 to get 2 points.
- Location: Medical Science Center. North-West corner of Univ. Ave and Charter st.
- Ask questions anytime or visit TA/Instructor and get email rebate points! We count your effort.
- A plenty of rebate point will help you if your final total score is less than 1 point of the borders.
- Exams are all multiple choice. No partial points. Time to show your understanding!
- Sample exams will be provided with answers. (last year’s exams).
- All your score will be recorded in learnuw.
- Ch 4 and Ch 12 will be super difficult.
- HWs are for your understanding. You get full points once you submit.

2 Ch1 About Statistics

- 3 concepts (p6): Population, Sample, Unit
  
  eg1) Population= All students in the world. Sample = 30 students in UW Madison. Unit = student.
  
  eg2) Population= All banana in the world. Sample = 10 banana I bought at Copps. Unit = banana.
  
  Def: Population: The set of all records.
  Def: Sample: subset of population.
  Def: Unit: Single entity. This is usually a person or object

(p19) "One goal of statistics is to understand and quantify the variation in the data and, if possible, to identify sources that contribute to this variation."

Why do we need to introduce the concept "sample"? Why don’t we just use population? Because in reality, it is impossible to get all information of population. Statistics will give you some technique to guess population’s characteristics from sample which is partial information of population.

Q) Is Statistics useful all the time? If not, counter examples?

3 Ch2 Organizing Data and Describing Patterns

- (p26): Basic types of data.
  Categorical (Qualitative)
    Nominal (no relation between each letter) eg) Blood type (O, A, B, AB).
    Ordinal (some ordering is possible) eg) Academic grade (A, B, C, D, E, F).
  Measurement (Quantitative)
    Discrete (values in steps) eg) Human height (170.0 cm, 170.5 cm, 171.0 cm, 171.5 cm etc) by digital measurement.
    Continuous (values in interval)
eg) Human height (170 cm, 170.002 cm, 172.2 cm) by analogue measurement.

**Note 1**: As a nature, human height is continuous, but if you change experiment (i.e. use different measurement machine, it can be discrete.)

- For pie chart, see p30
- For dot diagram, see p35, p36.
- Histogram of the Frequency Distribution (Categorical data)

\[
Relative \text{ frequency} = \frac{\text{Frequency in the category}}{\text{Total number of observations}} \quad (1)
\]

eg)

<table>
<thead>
<tr>
<th>Responses</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Oppose</td>
<td>184</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>400</strong></td>
<td><strong>1.00</strong></td>
</tr>
</tbody>
</table>

Frequency Table (Table 1 (p29) modified example)

Histogram of the frequency distribution is just a figure with x-axis = category, y-axis = relative frequency.

**UNEQUAL class interval** case:

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 5)</td>
<td>30</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>[5, 10)</td>
<td>35</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>[10, 15)</td>
<td>35</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>1.00</strong></td>
<td></td>
</tr>
</tbody>
</table>

\[
Height = \frac{\text{Relative frequency}}{\text{Width of interval}} \quad (3)
\]

\[
Relative \text{ frequency} = \frac{\text{Class frequency}}{\text{Total number of observations}} \quad (2)
\]
• Total area=1 (must be).
  In the above example,
  \[(5-0)*0.060 + (10-5)*0.070 + (20-10)*0.035=1.\]

• p39 Fig 9 (a) is NOT histogram (because total area ≠ 1).
  Fig 9 (b) is histogram.

• (p40) Stem-and-leaf display.
  eg) You have data: 31, 34, 35, 40, 42, 42, 67.
  Then stem-and-leaf display is

  \[
  \begin{array}{c}
  3 \\
  4 \\
  5 \\
  6 \\
  \end{array}
  \]