TEACHING INTRODUCTORY STATISTICS WITH IN-CLASS DATA COLLECTION

Dr. Heidi Hulsizer, Hampden-Sydney College
GAISE Recommendations

1: Emphasize statistical literacy and develop statistical thinking.

2: Use real data.

3: Stress conceptual understanding, rather than mere knowledge of procedures.
GAISE Recommendations

4: Foster active learning in the classroom.

5: Use technology for developing concepts and analyzing data.

6: Use assessments to improve and evaluate student learning.
Naked, Realistic, & Real Data

- Naked Data – Made-up data with no context.
- Example: Find the correlation coefficient

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
Naked, Realistic, & Real Data

- Realistic – Context added to Made-up data
- Example: Simpson’s Paradox Lesson*

\[
\begin{array}{|c|c|c|}
\hline
 & \text{Helicopter} & \text{Ambulance} \\
\hline
\text{Victim died} & 64 & 260 \\
\hline
\text{Victim survived} & 136 & 840 \\
\hline
\text{Total} & 200 & 1100 \\
\hline
\end{array}
\]
## Naked, Realistic, & Real Data

<table>
<thead>
<tr>
<th>Serious</th>
<th>Helicopter</th>
<th>Ambulance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Victim died</td>
<td>48</td>
<td>60</td>
</tr>
<tr>
<td>Victim survived</td>
<td>52</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Less serious</th>
<th>Helicopter</th>
<th>Ambulance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Victim died</td>
<td>16</td>
<td>200</td>
</tr>
<tr>
<td>Victim survived</td>
<td>84</td>
<td>800</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>1000</td>
</tr>
</tbody>
</table>
Naked, Realistic, & Real Data

- Real – Data from a real situation that is of interest to students.
- Example:

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>8</th>
<th>3</th>
<th>6</th>
<th>5</th>
<th>2</th>
<th>2</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Age Boys First Ride an ATV
Examples for Regression

- Grade on Exam vs. Study Time

<table>
<thead>
<tr>
<th>Time Studying (minutes)</th>
<th>Grade (out of 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>91</td>
</tr>
<tr>
<td>90</td>
<td>65</td>
</tr>
<tr>
<td>120</td>
<td>68</td>
</tr>
<tr>
<td>10</td>
<td>92</td>
</tr>
<tr>
<td>60</td>
<td>51</td>
</tr>
<tr>
<td>180</td>
<td>82</td>
</tr>
<tr>
<td>20</td>
<td>83</td>
</tr>
<tr>
<td>45</td>
<td>81</td>
</tr>
</tbody>
</table>
Examples for Regression

- Cumulative GPA vs. Course Grade
- Age vs. Number of Parking Tickets/Car Accidents
- Number of Cars Owned vs. Number of Cars Wrecked
Examples for Hypothesis Testing

- Rolling dice
  - Also good for Law of Large Numbers and Central Limit Theorem

- Poker chips out of a bag
Examples for Z-Test, T-Test

- How many children do you want to have? ($\mu = 2.5$)
- How many bicycles have you owned in your lifetime? ($\mu = 3$, $\sigma = 1.3$)
- How many pets do you own?
- What is the ideal age to get married? ($\mu = 27$ women, $\mu = 29$ men$^2$) *skewed

* http://en.wikipedia.org/wiki/Age_at_first_marriage
Examples for 2-Sample T-Test

- At what age did you first ride an ATV? Rural vs. Urban

<table>
<thead>
<tr>
<th>Age Boys First Ride an ATV</th>
<th>Rural</th>
<th>2</th>
<th>3</th>
<th>8</th>
<th>3</th>
<th>6</th>
<th>5</th>
<th>2</th>
<th>2</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>2</td>
<td>3</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>Urban</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Any Z/T – Test example divided by gender.
Examples for Proportions or Binomial Distributions

- Do you support vaccinations for all children?
- Would you report cheating to a professor?
- Do you live on-campus? ($p = ?$ your campus)
- Do you have a car? ($p = .95^1$)

^1http://photos.state.gov/libraries/cambodia/30486/Publications/everyone_in_america_own_a_car.pdf
Examples for Two-Way Tables

- How often do you use Facebook?
  - More than once a day, Once a day, Several times a week, Once a week, Less than once a week

- Where were you raised?
  - Rural, Urban, Suburban

- Have you ever broken a bone? Or, how many bones have you broken?

- Grade level?
  - Freshmen, Sophomore, Junior, Senior
How to collect data:

- Go around the room and ask
- Slips of paper
- With other work
- Student demonstration
- Beginning of the semester survey (could be online)

Websites
- [http://www.worldometers.info/](http://www.worldometers.info/)
- [http://lib.stat.cmu.edu/DASL/](http://lib.stat.cmu.edu/DASL/)
Warnings

- Anonymous/Confidential
- Nice numbers are nice sometimes (Simpson’s paradox)
- Be prepared for wild p-values
- Be prepared for outliers and be able to compare results without them
Technology

- TI Calculators
- Excel
- Online Resources/Applets
References
