Math B22
Spring 2014

Exam 1

Love Letter From a Statistician

Dear Valentine-
You're three
standard
deviations above
the mean!

Instructions:

• Please turn off and put away all cell phones and any listening devices.
• Each question is weighted the same, and graded using a 5 point rubric. The exam is worth 100 points.
• Work should be clearly labeled and easy to follow.
• Partial credit will only be given for work showing significant progress toward the correct answer.
• You may use a graphing calculator on this exam.
• Work in pen will not be accepted.
• You may use one 8.5" x 11", or smaller handwritten sheet of notes. Both sides may be used, but photocopies of any kind are not allowed.
• Do not hesitate to ask the instructor if you do not understand what a question is asking, or if you have shown enough work.

Date: __________  Name: __________

Class(circle one):  4:20  6:00
1. The frequency distribution below shows the number of pounds of each snack food eaten during the Super Bowl. (a) Find the percent for each category rounded to the nearest whole percent. (b) Determine the central angles of a pie chart that corresponds to each category.

<table>
<thead>
<tr>
<th>Snack</th>
<th>Pounds (in millions)</th>
<th>Percent</th>
<th>Work for Percent</th>
<th>Central Angle</th>
<th>Work for Central Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potato chips</td>
<td>11.2</td>
<td>37%</td>
<td>$\frac{11.2}{30}$</td>
<td>$133^\circ$</td>
<td>$0.37(360^\circ)$</td>
</tr>
<tr>
<td>Tortilla chips</td>
<td>8.2</td>
<td>27%</td>
<td>$\frac{8.2}{30}$</td>
<td>$97^\circ$</td>
<td>$0.27(360^\circ)$</td>
</tr>
<tr>
<td>Pretzels</td>
<td>4.3</td>
<td>14%</td>
<td>$\frac{4.3}{30}$</td>
<td>$50^\circ$</td>
<td>$0.14(360^\circ)$</td>
</tr>
<tr>
<td>Popcorn</td>
<td>3.8</td>
<td>13%</td>
<td>$\frac{3.8}{30}$</td>
<td>$47^\circ$</td>
<td>$0.13(360^\circ)$</td>
</tr>
<tr>
<td>Nuts</td>
<td>2.5</td>
<td>8%</td>
<td>$\frac{2.5}{30}$</td>
<td>$29^\circ$</td>
<td>$0.08(360^\circ)$</td>
</tr>
</tbody>
</table>

\[ n = 30 \quad \frac{99\%}{(close\ to\ 100\%)} \quad \frac{356^\circ}{(close\ to\ 360^\circ)} \]

2. Males have red blood cell counts with a mean of 5.072 and a standard deviation of 0.395, while females have red blood cell counts with a mean of 4.577 and a standard deviation of 0.382. Who has the higher count relative to the sample from which it came: John (a male) with a count of 4.9 or Jane (a female) with a count of 4.32?

**John**

\[
\hat{z} = \frac{4.91 - 5.072}{0.395} = -0.062 \approx -0.41\]  \((calc\ -0.41013)\)

**Jane**

\[
\hat{z} = \frac{4.32 - 4.577}{0.382} = -0.257 \approx -0.67\]  \((calc\ -0.67277)\)

John has the higher relative count since \(-0.41 > -0.67\).
3. For problems a-d, determine which of the four levels of measurement (nominal, ordinal, interval, ratio) is most appropriate.

   a. The amount of money in a person's wallet.  
      \[\text{ratio}\]
   b. Body temperatures of the patients admitted to an emergency room.  
      \[\text{interval}\]
   c. Types of shoes: flats, heels, pumps, boots, and sneakers.  
      \[\text{nominal}\]
   d. Military titles such as Lieutenant, Captain, and Major.  
      \[\text{ordinal}\]
   e. Based on the information in the previous problems a through d, circle all the ones that also are examples of categorical data.
      \[\text{c and d}\]

4. Identify the type of sampling used at a petroleum refining factory: random, systematic, convenience, stratified, or cluster.

   a. All the barrels of refined oil are labeled with a number. Duplicate labels with the same numbers are placed in a box and thoroughly mixed. Sixty-five labels are drawn from the box, and the oil in those barrels will be inspected for quality control.  
      \[\text{random}\]
   b. The factory foreman stops and checks the weights of the first 3 tanker trucks leaving the compound to make sure the oil pump is filling tanks to full capacity.  
      \[\text{Convenience}\]
   c. The oil is pumped into 7 different holding tanks, and then all the oil in Tank 2 and Tank 5 are randomly selected for refining.  
      \[\text{Cluster}\]
   d. The empty metal barrels that are returned to the factory are sent through a machine one at a time to be cleaned. Every 25th barrel is pulled out of line and is checked for signs of rust.  
      \[\text{Systematic}\]
   e. Employees at the factory work in 7 different departments. Ten employees from each department are selected to participate in a job satisfaction survey.  
      \[\text{Stratified}\]
The data set below represents the amount of garbage (in millions of tons) collected over a period of 19 years.

<table>
<thead>
<tr>
<th>31</th>
<th>40</th>
<th>46</th>
<th>46</th>
<th>52</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>53</td>
<td>58</td>
<td>38</td>
<td>51</td>
</tr>
<tr>
<td>58</td>
<td>53</td>
<td>43</td>
<td>56</td>
<td>57</td>
</tr>
<tr>
<td>54</td>
<td>57</td>
<td>55</td>
<td>(46)</td>
<td></td>
</tr>
</tbody>
</table>

5. For the data set of collected garbage, find the following:

a. mean 
\[ \bar{X} = \frac{49,9 \text{ million tons}}{(49.9)} = 15.2 \% \]  
(calc 49.94736842)

b. standard deviation
\[ S = \sqrt{7.6 \text{ million tons}} (calc 7.488338122) \]

c. mode
mode \(= 46.0 \text{ million tons} \)

d. coefficient of variation.
\[ C.V. = \frac{7.6}{49.9} \times 100 \% = 15.2 \% \]  
(calc 15.23046092)

6. For the data set of collected garbage, find the 5-number summary and construct a boxplot.

<table>
<thead>
<tr>
<th>5-Number Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>min = 31</td>
</tr>
<tr>
<td>Q_1 = P_{25} = 46</td>
</tr>
<tr>
<td>Q_3 = P_{75} = 56</td>
</tr>
<tr>
<td>max = 58</td>
</tr>
</tbody>
</table>

*Data sorted on calculator*

\( h = .25(19) = 4.75 \) \( \rightarrow 5 \text{th position} \)

\( h = .75(19) = 14.25 \) \( \rightarrow 15 \text{th position} \)
The data set below represents the amount of garbage collected (in millions of tons) over a 19 year period. (Show a tally)

\[
\begin{array}{cccccc}
31 & 40 & 46 & 46 & 52 \\
55 & 53 & 58 & 38 & 51 \\
58 & 53 & 43 & 56 & 57 \\
54 & 57 & 55 & 46 & \\
\end{array}
\]

\[\text{min} = 31\]
\[\text{max} = 58\]

7. Construct a frequency distribution of the amount of garbage beginning with a lower class limit of 30 million tons and use a class width of 5 million tons. Also, construct a relative frequency distribution (round the percents to one decimal place).

<table>
<thead>
<tr>
<th>Garbage (millions of tons)</th>
<th>Tally</th>
<th>Freq</th>
<th>Relative Freq.</th>
<th>Work for Rel. Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 - 34</td>
<td>I</td>
<td>1</td>
<td>5.3%</td>
<td>(1/19 = .0526)</td>
</tr>
<tr>
<td>35 - 39</td>
<td>I</td>
<td>1</td>
<td>5.3%</td>
<td>(2/19 = .10526)</td>
</tr>
<tr>
<td>40 - 44</td>
<td>II</td>
<td>2</td>
<td>10.5%</td>
<td>(3/19 = .157895)</td>
</tr>
<tr>
<td>45 - 49</td>
<td>III</td>
<td>3</td>
<td>15.8%</td>
<td>(5/19 = .263158)</td>
</tr>
<tr>
<td>50 - 54</td>
<td>IV</td>
<td>5</td>
<td>26.3%</td>
<td>(7/19 = .368421)</td>
</tr>
<tr>
<td>55 - 59</td>
<td>V</td>
<td>7</td>
<td>36.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

8. a. Use the frequency distribution below to estimate \(\bar{x}\).

\[
\begin{array}{|c|c|c|}
\hline
\text{Number of Minutes} & \text{frequency} \\
\hline
7-18 & 2 \\
19-30 & 5 \\
31-42 & 10 \\
43-54 & 13 \\
55-66 & 10 \\
67-78 & 8 \\
79-90 & 2 \\
\hline
\end{array}
\]

\[
\begin{align*}
X &= \frac{2497}{50} = 49.94 \\
\overline{X} &= 49.9 \text{ minutes}
\end{align*}
\]

b. Based on the above frequency distribution, do the number of minutes spent on the Internet appear to be normally distributed? Explain why or why not.

Yes, the frequencies start low, increase to a max, then drop back to a low. And, there is symmetry.
9. In a recent Pew Poll, 5,000 randomly selected Californian households were surveyed and it was found that 49% of those surveyed said that the use of marijuana should not be made legal. There are 12.5 million households in California.

a. Identify the population.

12.5 million households

b. Identify the sample.

5,000 households

c. Determine whether the sample is likely to be representative of the population and explain why or why not.

Representative since they were randomly selected.

d. Does the 49% in this problem represent a parameter or a statistic? Explain.

49% is a statistic because it describes the sample.

10. What's wrong in each problem below? Give an explanation for each.

a. The graph below was submitted to the local paper with the results of a contest at a middle school that raised donations for the United Way. What's wrong with the graph?

Dollars raised for the United Way

$2,500
$2,250
$2,000
$1,750
$1,500
$1,250
$1,000

8th Grade 7th Grade 8th Grade

The vertical axis does not begin at zero, this makes it look like the 8th graders raised 5 times or more than the 7th graders, it is only about twice as much.

b. A survey included 4,230 responses from Internet users who decided to respond to a question posted by American Online. The survey results showed that 67% said they used credit cards frequently. What's wrong with this survey?

They used a voluntary response sample which causes bias in the data.

c. A pest control company claims that with regular monthly service, you can reduce the number of pests in your home by 150%. What's wrong with that statement?

A 150% reduction is not possible.

100% reduction would eliminate all pests.