Statistics 1: Sections 3 and 5
CSUS, Spring 2010
Practice Problems for Exam 1

Note: These practice problems are NOT a comprehensive list of the material covered for Exam 1. You are responsible for all the sections we covered in Chapter 1.1 – 4.5 of the text. These problems are meant to be a sample of the type of problem you might see on the exam. I have not timed this exam. It may differ slightly in length from the 50 minute exam you will take.

One 8.5”x11” sheet of handwritten notes allowed. Calculators allowed. The actual exam will have space for your work on the exam paper itself.

1) For the following dataset: 10,6,2,7,100
   a) Calculate the mean and median.
      \( \bar{x} = \frac{125}{5} = 25 \) median: order data 2, 6, 7, 10, 100. Thus, median = 7
   b) Which is a better measure of central tendency, the mean or median? Why?
      The median is a better measure of center since the data contain an outlier, 100, which distorts the mean.
   c) Calculate the standard deviation using the formula
      \[ s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \]
      
      \[ \begin{array}{|c|c|c|}
          \hline
          x_i & x_i - \bar{x} & (x_i - \bar{x})^2 \\
          \hline
          10 & 10-25 = -15 & 225 \\
          6 & 6-25 = -19 & 361 \\
          2 & 2-25 = -23 & 529 \\
          7 & 7-25 = -18 & 324 \\
          100 & 100-2 = 75 & 5625 \\
          \hline
          \sum (x_i - \bar{x})^2 &= 7064 \\
          \hline
        \end{array} \]
      \[ s = \sqrt{\frac{7064}{5-1}} = 42.02 \]
   d) If you added a fixed number, k, to each value in the data set, would the standard deviation change? If so, how?
      No, the standard deviation will not change. Adding the same value to each data point will not change how spread out the data are. Since the standard deviation measures the spread of the data, it will not change when k is added to each data point.
   e) A portion of the ordered exam scores for a class of 100 students is shown below. Will removing the lowest score affect the standard deviation? If so, how?
      Yes, removing the lowest score, 6, will decrease the standard deviation. Since the lowest score is an outlier, it’s squared deviation from the mean, \((x_i - \bar{x})^2\), will be large. This will cause the standard deviation to be large. When the lowest score is removed, this very large deviation from the mean will no longer be included in the standard deviation calculation.
   f) Use Chebychev’s Rule to determine the value of k such that the interval \( \bar{x} \pm k \cdot s \) contains at least 96% of the data.
      According to Chebychev’s Rule, the proportion of the data in \( \bar{x} \pm k \cdot s \) is at least \( 1 - \frac{1}{k^2} \)
      So, we need to solve \( 1 - \frac{1}{k^2} = 0.96 \) which gives \( k=5 \).
2) Use the histogram to answer the questions below.

![Length of 200 Human Pregnancies](image)

a) The value 240 represents the upper boundary of the left-most rectangle in the histogram. What is the upper boundary of the second rectangle from the left?

250

b) Use the histogram to estimate the relative frequency of the category having upper bound 260.

About 38/200

c) Use the histogram to estimate the total number of pregnancies lasting 260 days or less.

About 38 + 9 + 6 = 53

d) The mean and standard deviation for this data are 269 days and 14.2 days, respectively. Calculate mean +/- two standard deviations for this data.

269 ± 2(14.2) gives (240.6, 297.4)

e) Use the histogram to approximate the percent of data falling in the interval calculated in part (d).

About 195/200 = 97.5%

f) Which would give you the best estimate of the percent of data falling in the interval in part (d) -- Chebychev’s Rule or the Empirical Rule? Give reasons for your choice then use the chosen rule to estimate the percent of data in the interval from part (d).

The Empirical Rule would give the best estimate since the data are bell-shaped. Empirical Rule says approximately 95% of the data will fall in the interval in part (d).
3) (20 points) A random sample of voters in Clark County were classified by political party and gender. The results are shown below. What is the probability a randomly selected voter is

<table>
<thead>
<tr>
<th></th>
<th>Democrat</th>
<th>Republican</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>27</td>
<td>36</td>
<td>7</td>
</tr>
<tr>
<td>Female</td>
<td>18</td>
<td>10</td>
<td>2</td>
</tr>
</tbody>
</table>

a) A democrat?
   45/100
b) Not a democrat?
   55/100
c) Female and democrat?
   18/100
d) A democrat or a female?
   
   \[
   \frac{27+18+10+2}{100} = \frac{57}{100}
   \]
e) Are the events voter is a democrat and voter is a republican mutually exclusive? Why or why not?
   Yes, these events are mutually exclusive since no person can be both a democrat and a republican at the same time. A person is only allowed to register in one political party at any given time.

4) The following data give exam scores for 28 students in a CSUS mathematics course, arranged from lowest to highest.

47  49  49  53
57  61  63  66
66  72  73  75
75  76  79  81
82  85  85  85
85  88  88  89
90  95  95  100

a) Find the first quartile, median and 85th percentile of these scores.
   \[
   \text{Position of } Q_1 = 0.25 \times (28+1) = 7.25 \text{ so } Q_1 = \frac{63+66}{2} = 64.5
   \]
   \[
   \text{Position of median} = 0.5 \times (29) = 14.5 \text{ so median} = \frac{76+79}{2} = 77.5
   \]
   \[
   \text{Position of 85th percentile} = 0.85 \times (29) = 24.65, \text{ so 85th percentile} = \frac{89+90}{2} = 89.5
   \]
b) Construct a boxplot for these data.
5) There are three people on a committee: Al, Bob and Carl. One person is selected to bring donuts for the next committee meeting and a person is selected to take minutes at the meeting. The same person may be assigned both tasks. Assume these selections are made at random.

b) List the outcomes in the sample space, S. Let the outcome AB mean Al brings donuts and Bob takes minutes. Then S = {AA, AB, AC, BA, BB, BC, CA, CB, CC}

c) List the outcomes in the event Bob is assigned at least one task. {AB, BA, BB, BC, CB}

d) What is the probability Bob is assigned at least one task? 5/9

e) What is the probability Bob brings donuts and Carl takes minutes? 1/9

f) What is the probability the same person is assigned both tasks? 3/9

g) What is the probability Bob is assigned no task? 4/9

6) Identify the following variables as qualitative or quantitative. Classify the quantitative variables as discrete or continuous.

a) Marital status - qualitative

b) Time until an alkaline battery wears out – quantitative continuous

c) Brand of an alkaline battery - qualitative

d) Whether or not a subject has disease X - qualitative

e) Number of votes a political candidate receives – quantitative discrete

f) The pounds of sugar consumed by a person in a week – quantitative continuous