Case-based Social Statistics I  
CSSS 321  
Professor: Mark S. Handcock

Solutions to Homework 2  
Due Friday, January 21, 2005

Problems to be handed in:
1) Exercise 2.68 from page 53 of Chapter 2: “Turning Data Into Information” of MOS.

Solution(5): This will be different for each student. The calculation is:

\[ z = \frac{\text{height} - \text{mean}}{s}. \]

Use the mean and standard deviation relevant to your gender. For example, my height is 5’8” (or 68 inches). I’m female, so my \( z \)-score is

\[ z = \frac{68 - 65}{2.5} = 1.2. \]

Note that the \( z \)-score will be negative if the height is less than the mean. Notice also that if the height equals the mean, the result is \( z = 0 \).

2) Exercise 2.83 from page 55 of Chapter 2: “Turning Data Into Information” of MOS.

Solution(15): This will be different for each student. The following example looks at data on the number of tracks on 24 CDs.

\begin{align*}
15 & 5 \ 12 \ 9 \ 15 \ 15 \ 14 \ 20 \ 19 \ 12 \ 19 \ 19 \\
12 & 4 \ 12 \ 18 \ 9 \ 9 \ 5 \ 10 \ 13 \ 14 \ 16 \ 12
\end{align*}

(a) To find the five number summary, first we put the data in order from smallest to largest.

\begin{align*}
4 & 5 \ 9 \ 9 \ 9 \ 10 \ 12 \ 12 \ 12 \ 12 \ 12 \ 13 \\
14 & 14 \ 15 \ 15 \ 15 \ 16 \ 18 \ 19 \ 19 \ 19 \ 19 \ 20
\end{align*}

Since there are an even number of values (24), the median will be the average of the two middle values. So median is \( \frac{13 + 14}{2} = 13.5 \). Then lower quartile is the middle of the values less than the median. Again we have an
even number (12), so we average the middle two. So the lower quartile is \( \frac{10 + 12}{2} = 11 \). Finding the median of the values greater than the median gives an upper quartile of \( \frac{15 + 16}{2} = 15.5 \). The extreme values of the minimum and the maximum, or 4 and 21. So my five number summary looks like:

<table>
<thead>
<tr>
<th>Number of CD tracks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Quartiles</td>
</tr>
<tr>
<td>Extremes</td>
</tr>
</tbody>
</table>

(b) Below is a boxplot of the data.

(c) Below is a dotplot of the data.
(d) The shape of the data is pretty symmetric. It looks like the points at 4 and 5 could be outliers. (e) The mean of my data is 13.25. That is pretty close to the median of 13.5, which is as expected, since the data looks fairly symmetric.

3) Exercise 2.90 from page 55 of Chapter 2: “Turning Data Into Information” of MOS.

Solution(5): Consider the categories given on page 17. In this case we have one categorical variable (martial status) and one quantitative variable (ounces of coffee consumed per day), so the possible questions to consider are 4a and 4b. The natural question is the first (Question 4a): Is the average amount of coffee consumed per day the same for married people as it is for single people? The second question does not apply as martial status is not ordered.

Although we were not asked, the answer to the question is unclear logically and up to empirical exploration (that is, what does the data say?)

4) Submit electronically exercises 22, 64, and 66 from Unit A5 of CyberStats.

A5 Ex. 22. What is the median height, approximately? How tall is the tallest person? About what percentage of these athletes are taller than six feet? Are these people generally taller than the average male?

Solution(4): The median height is about 74 inches. The tallest is about 84 inches. We can determine what percentage of these athletes are taller than six feet (72 inches) by counting them in the window created by clicking the “Data” button. 102 out of 132 are taller than 72 inches. This is 77%. Alternatively, the boxplot indicates that the lower quantile is at 72 so about 75% are above 72” (because 72” is Q1). Yes, they are taller than average as the average male is about 5’10” or 70 inches

A5 Ex. 64. Estimate the median.

Solution(1): From the boxplot, it is about 31.
A5 Ex. 66. Is there skewness? If so, which way?

Solution(2): Yes, there is skewness towards the higher temperatures.

Submit electronically exercises 21 and 22 from Unit A6 of CyberStats.

A6 Ex. 21. Find the two values Means ±s. How many of the temperatures are between those two numbers? What percentage of the temperatures is between those two numbers? Is it reasonably close to 2/3?

Solution(5): The mean and standard deviation are 65.34 and 7.98, respectively, so Mean ±s is from 57.36 to 73.32. 258 of the values are between those two numbers, which is 70.49%. It’s close to 2/3.

A6 Ex. 22. Find the two values Mean ±2 × s. How many of the temperatures are between those two numbers? What percentage of the temperatures is between those two numbers? Is it reasonably close to 95%?

Solution(5): Mean ±2 × s is from 49.38 to 81.30. 348 of the values are between those two numbers, which is 95.08%. It’s very close to 95%.

6) In this question we will analyze the “Study: ‘Cutting TV, video game use can make kids less aggressive’”.

The article is a Associated Press report of a study about the effects of TV and video game use on violence among children. The article is at:

http://www.canoe.ca/CNEWSFeatures0101/15_kids-ap.html

You should read the article and answer the questions posed below.

The original study being reported on is available on-line from the UW library. Search for the the title “Archives of Pediatrics and Adolescent Medicine”. It will give you a link to the online version where you find the article in the January 2001 edition.

Here are the questions:

a) Answer each of the 6 questions on Lecture 1-4 with respect to the article

Solution(7):

1. What are the claims made in the article?
   The major claim of the article is that the effects of televised violence in kids are reversible. It is claimed that there exists a relationship between media exposure and children behavior and it is modifiable.

2. Are the claims supported in the article? If so, how?
   The claims are somewhat supported in the article. The results of a study suggesting that kids with less TV viewing are less aggressive compared to kids without intervention are reported. Also, two experts
(one associated with the study, and one independent of the study) are quoted.

3. Is there data in the article used to support the claims?
   Yes, the article reports data from the study mentioned above, and this data is used to support the claims.

4. Is there data in the references for the claims?
   Yes, the journal article in the *Archives of Pediatrics and Adolescent Medicine* contains the data from the original study.

5. Can you think of alternative explanations for the data in the article? Is there data in the article to support these alternatives?
   One possible alternative is that the kids in the treatment group know that they are watched and therefore change their behaviors during the period of the study. The data in the article showed 25% decrease in the peer reports of aggression and fewer observed incidents on the playground for the treatment group, which might indicate that kids behaved better when being watched.

6. Does the data presented support the claims?
   Somewhat. The data quoted in the article does seem to show that aggression was less in the treatment group than in the control group.

**Solution (7):**

b) Briefly describe each of the components of the study given in Lecture 1-5 with respect to the study.

- Component 1: The individuals or objects studied and how they were selected
  Subjects were 225 3rd and 4th graders from 2 comparable public elementary schools. Students who had parental consent were selected to participate. One school was randomly selected to be the treatment school. 60% of the participating children in the treatment school were randomly selected for playground observation.

- Component 2: The exact nature of the measurements made and the questions asked
  The researchers measured changes in verbal and physical aggression by observing the playground behavior of about 50 children from each
school. The researchers also recorded the percentage change of peer reports of aggression between 2 schools. Additionally, the researchers determined how many TVs each participant had access to and how much time they spent watching TV.

- Component 3: The setting or context in which the measurement were made
  All measurements were done in the natural school setting. The two schools were chosen to be comparable. Of course all students for the treatment group were in one school and all students for the control group were in a different school. Thus differences between the school in terms of student backgrounds, and within school programs could effect the results (e.g. One might have had football practice before testing, while the other had music lessons).

- Component 4: The extraneous differences between groups being compared
  Remember that extraneous differences are differences between the control and treatments group besides the treatment itself. For example, the results might be misleading if kids from the treatment group have some other activities or classes (i.e. ethic class) at the same time of the study that might have discouraged them from acting violent.

  The control group should have received some type of lessons during the same time frame, but on subjects unrelated to the the violence treatment. Also they should have some other activity scheduled for the times that the treatment groups was restricted from TV viewing. This isolates the effects of TV viewing and the content of the lessons from the format of the study.

- Component 5: The magnitude of any claimed effects of differences
  Is the claimed reduction of the number of incidents and peer reports a huge difference compared to the variation of the number of these incidents and reports? In the journal article, the authors call the effects for perr-nominated agression and observed verbal aggression small to medium, and medium effects, respectively.

- Component 6: The source of the research and the funding
  The source of the research was the Standford University School of Medicine. The study was funded in part by grants from the American
Heart Association and the National Heart, Lung and Blood Institute. There might be potential biases if the source or funding of the research was from groups against violence on TV and video games, but that seems unlikely in this case.

• Component 7: The researchers who had contact with the participants There might be potential changes of behavior among kids on the playgrounds if they knew that they were observed by the researchers.

c) Briefly answer these questions about the study itself.

i) What were the two treatments in the study?

Solution (1): The two treatments are 18 lessons on reducing the use of television, videotapes and video games over six months and the limitation on the hours of TV viewing (the treatment) and not getting the lessons/restrictions (the control).

ii) Who were the subjects? How many subjects were there?

Solution (2): The subjects are 3rd and 4th grade children at two comparable public elementary schools in San Jose, CA. From one school, 120 children received no intervention served as the control group. At the other school, 105 children served as treatment group.

iii) What pre-treatment characteristics were recorded?

Solution (1): Peer reports of aggression, playground observation of aggression, and average weekly hours of TV, video watching and game playing were recorded before the treatment.

iv) What were the main outcome measurements?

Solution (1): The main outcome measurements are changes in verbal and physical aggression observed on the playground and number of peer reports of aggression between the two groups.

d) Describe the method of treatment assignment. Why do you think they assigned the treatments randomly?

Solution (1): The treatment school was selected randomly from the two candidate schools. For playground observation, the students were selected randomly from all participants. Randomly selected subjects are to ensure to obtain a sample that is representative to the population that we are interested in.

e) Was the treatment assignment “blind” to the subjects? The researcher (i.e. “double blind”)
Solution(2): The treatment is not “blind” to the subjects as they took 18 lessons regarding reduction of TV viewing. The treatment is not “blind” to the researcher as they trained the teachers who conducted the 18 lessons.

Hence the students knew they were in a study, and probably knew that they were being observed for violence and TV. It is somewhat likely that this modified their behavior. Of course, this is hard to avoid in a study of this type. On way to ameliorate the effects is to “treat” the control group in a similar form as the treatment group, but do not give them the substance of the treatment, that is, information about the violence and the content of TV versus other forms of entertainment. Such changes would make the results much stronger.

However, the studies goals and hypothesis were not made clear to the students, parents and teachers thereby reducing the likelihood of changing behavior toward (or away) from the researchers goals.

Hence they were blinded to this aspect of the study.

f) Given the information in the report, do you think the article title, “Cutting TV, video game use can make kids less aggressive” is a reasonable claim?

Solution(2): I think the statement from the title of the article is too strong provided the study was based on only 2 schools and the researchers did not assess what kids watched. A less strong statement can be that “Cutting TV,...could make kids less aggressive”.

g) If you could add one paragraph to the newspaper article to give more information from the study, what would you say?

Solution(2): I would add more details about the size, variation, and significance of the observed reduction of aggression in the treatment group.

Overall this is quite a good article in that it reports the nuances of the study (e.g. small number of schools), and lets outside researchers have a say also. The underlying study is quite professional as well - these forms of studies that involve observations of interventions are difficult to perform and design.

Note, however, that the reporters (and researchers) desire to confirm the “obvious” and “natural” notion that video games and TV influence violence permeates the writing. Indeed the selection of the question and the design of the study itself are somewhat biased toward finding results. Part of this is the underlying business of science.
The researchers have probably been funded to do the small study reported in the article. At the end they are beginning to report on a larger study on about 900 students at 12 schools over a longer time span. This study would have been much more expensive. Do you think they would have the same chance of funding if the results of the study reported in the article had shown little or mixed results?

Feel free to ask me questions in class about the material, or on the discussion board. You can discuss it amongst yourselves, but please submit separate work.