Please write your name and student ID below, and circle your section. With your signature, you certify that you have not observed poor or dishonest conduct on the part of your classmates. You also certify that you have not been a party to poor or dishonest conduct, and that the work on this exam is solely your own.

Name: ___________________________ Student ID: ___________________________

Signature: ___________________________ Date: ___________________________

Section: 102 (3pm-4pm) 101 (4pm-5pm)

Answer the questions in the spaces provided. There are questions on the front and back of each page. This midterm covers the material from Lectures 1 through 13, and Homeworks 1 through 6. Show your work, including labeling quantities (such as z-scores). The clearer that your work is, the easier it is to award partial or full credit. If you do not show your work, you will not receive credit. You are welcome to leave your answers as fractions. If you use decimals, please round all answers to two significant figures, and hold your rounding until the final calculation.

<table>
<thead>
<tr>
<th>Question</th>
<th>Points</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>
1. In June 2013 the United States House of Representatives was voting on a farm bill. The pie chart below was featured in a Washington Post article about the proposed law and the ensuing vote (the color has been changed to black and white). The original article can be viewed at [http://goo.gl/N6pBd](http://goo.gl/N6pBd).

<table>
<thead>
<tr>
<th>Segment</th>
<th>Percentage</th>
<th>Cost (2014-2023)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food stamps and Nutrition</td>
<td>79.2%</td>
<td>$743.9 billion</td>
</tr>
<tr>
<td>Crop insurance</td>
<td>9.9%</td>
<td>$93 billion</td>
</tr>
<tr>
<td>Commodity programs</td>
<td>6.0%</td>
<td>$40.1 billion</td>
</tr>
<tr>
<td>Conservation</td>
<td></td>
<td>$56.7 billion</td>
</tr>
<tr>
<td>Everything else</td>
<td></td>
<td>$5.8 billion</td>
</tr>
</tbody>
</table>

(a) (2 points) Please give *two distinct reasons* why this pie chart is not an ideal data display.

(b) (2 points) Please construct a better display of this data, including filling in any missing values (if needed).
2. (Hypothetical) *Happy Runners* magazine is interested in creating a training program to entice the public to subscribe to their magazine. They contact all of their subscribers via email asking them to follow the training program attached to the email. The training program consists of running sets of intervals 3 times per week for 8 weeks. They are asked to record the time it takes them to run one mile both before and after completing the training program. About 40% of the runners respond, submitting “before” and “after” times. On average, there is a 45 second improvement in the mile time.

(a) (3 points) Describe the method of comparison, and explain how/if it is being used in this case.

(b) (3 points) *Happy Runners* magazine begins publishing ads claiming that “a study has shown that following our training program will improve your mile time by 45 seconds. Subscribe today, and receive this program for free”. Some friends of yours want to subscribe, as they are interested in improving their times. Do you think they should do this? Give at least two reasons.
(c) (3 points) Give at least two examples of possible confounders that could affect this study.

(d) (3 points) Give at least two ways in which this study could be improved.
3. A manufacturing process produces dumbbells used in exercise studios. It is known that the distributions of the actual weights of each type of dumbbell is normal.

(a) (4 points) The dumbbells labeled as 5 pounds have an average weight of 5 pounds, with a spread of 0.1 pounds. A batch of 300,000 dumbbells is manufactured. Find the number of dumbbells that weigh between 5.175 and 5.22 pounds. (Hint: Begin by drawing the normal curve and shading the area desired.)

(b) (3 points) The dumbbells labeled as 10 pounds have an average weight of 10 pounds. 10% of these dumbbells weigh 9.74 pounds or less. What is the sd for the dumbbells labeled as 10 pounds? Be sure to include units.
4. On the planet Kerbin food is a scarce, expensive resource. The scatterplot below shows the relationship between weight (in pounds) and yearly salary (in dollars), for all 5,000 kerbans, with a correlation of 0.56.

\[
\text{Weight vs. Salary} \\
\text{r = 0.56}
\]

(a) (2 points) The government on Kerbin is transitioning from the currency dollars to the currency yen. Each yen will be worth 17 dollars. With the information given, can you calculate the new correlation between weight (in pounds) and salary (in yen)? If yes, explain and do so. If no, why not?
(b) (2 points) Kerbin is a caste oriented society, where members of the Moho clan are eligible only for low-wage menial jobs, members of the Eve clan are eligible only for mid-wage, service oriented jobs, and members of the Duna clan are eligible only for high-wage, business and technology jobs. When looking at only the members of the Eve clan, the correlation between weight and salary will be (circle one, and explain your answer):
(Hint: try grouping the scatterplot above by clan)

A. Higher than 0.56  B. Lower than 0.56  C. About 0.56  D. Cannot tell

(c) (2 points) Two workers at the statistical bureau on Kerbin were tasked with analyzing the data on weight vs. salary. They were asked to each complete the task of computing the correlation between the two variables. Bill looked at the data collected on all 5,000 of his fellow kerbans, and calculated the correlation of 0.56. Jebediah, on the other hand, looked at the data grouped into clan membership (Moho, Eve, and Duna). He computed group averages for salary and weight for each of these clans, then computed the correlation of those group averages. Jebediah’s correlation will be (circle one):
(Hint: try sketching Jebediah’s scatterplot)

A. Higher than 0.56  B. Lower than 0.56  C. About 0.56  D. Cannot tell

Please explain your choice, and whether you prefer Bill’s method or Jebediah’s method.
5. (Hypothetical) Measurements of the amount of daily rainfall in two towns (Redley and Berking) were measured for 60 days. A scatter plot is shown below.

Redley Rain: mean = 3.25 in  sd = 1.2 in
Berking Rain: mean = 3.65 in  sd = 1.5 in
 correlation: r = 0.60

(a) (2 points) Predict the daily rainfall in Berking when Redley receives 1.85 inches of rain. Include units.

(b) (2 points) The prediction in (a) is likely to be off by how much? Include units.
(c) (4 points) Among all days when Redley receives 1.85 inches of rain, approximately what percent of days feature more than 0.8 inches of rain in Berking?

(d) (3 points) Fill in the blank: About _____ % of the days in Berking have rainfall between 3.05 inches and 5.9 inches.

(e) (3 points) Predict the daily rainfall in Redley when Berking receives 2.6 inches of rain. Does this match your answer from part (a)? Does this concern you? Why or why not?
6. (4 points) An apartment manager has only 12 units for rent. All of the available units are on either the first or second floor. Some apartments are corner units, some are not. 5 available units are on the top floor. 4 available units are corner units. 2 available units are corner units and are also on the bottom floor. She has 3 applicants who want to rent from her and they all want corner units on the bottom floor. She decides that the best way to assign units is to do so totally at random. What is the chance that two applicants receive their preference (corner units on the bottom floor), while one applicant comes close to their preference (either a corner unit OR a bottom floor unit)?

7. A fair die is rolled 7 times.
   (a) (2 points) What is the chance that in those 7 rolls, we see 3 “1”s?

   (b) (3 points) What is the chance that not all 7 rolls are “5”s or “6”s?
8. A box has 30 strips of paper in it. Each strip has a number written on it in a color. There are 3 colors (red, blue, yellow) and the numbers 0 through 9. That is, there are the numbers 0 through 9 in red, in yellow, and in blue. Each strip is folded in half, so that the color and number combination cannot be seen.

(a) (2 points) 3 strips are drawn from the box without replacement. If only the third strip is unfolded, what is the chance that it is a yellow 6?

(b) (3 points) 2 strips are drawn from the box with replacement. What is the chance that the first strip is blue and the second strip is either a green or a 7?

(c) (3 points) 4 strips are drawn from the box with replacement. What is the chance that at most 3 of them have an even number? (Assume that 0 is an even number).