Psych 315, Winter 2021, Homework 2 Answer Key

Due Friday, January 15 by 5pm

Name _____ ID _____

Section [AA] (Natalie), [AB] (Natalie), [AC] (Ryan), [AD] (Ryan), [AE] (Kelly), [AE] (Kelly)

Problem 1) Suppose you sample n=8 household incomes from a specific neighborhood and obtain the following values (in thousands of dollars):

X = 18, 6, 31, 5, 9, 22, 8 and 5

a) What is the median household income of this sample?

The scores in order are 5, 5, 6, 8, 9, 18, 22 and 31
The median annual income is 8.5 thousand dollars.

b) Calculate the mean:

$$\bar{X} = \frac{\sum X}{n}$$

$\sum X = 18 + 6 + 31 + 5 + 9 + 22 + 8 + 5 = 104$

The mean is $\frac{104}{8} = 13$ thousand dollars

c) Compare the median and mean and make a comment about the shape of the distribution of incomes for this example. Which is the most appropriate measure of central tendency? (mean median or mode).

If a distribution is positively skewed then the mean will be
greater than the median.
If a distribution is negatively skewed then the mean will be
smaller than the median.
If a distribution is skewed, then the median is most appropriate.

d) Calculate the deviation from the mean for each score:

$(X - \bar{X})$

5, -7, 18, -8, -4, 9, -5 and -8

e) Calculate SS_X, the sum of squared deviations from the mean:

$SS_X = \sum (X - \bar{X})^2$

$SS_X = (18 - 13)^2 + (6 - 13)^2 + (31 - 13)^2 + (5 - 13)^2 + (9 - 13)^2 + (22 - 13)^2 + (8 - 13)^2 + (5 - 13)^2$

$SS_X = 25 + 49 + 324 + 64 + 16 + 81 + 25 + 64 = 648$

f) Calculate the population variance using:

$MS = \frac{\sum (X - \bar{X})^2}{n}$

Round to two decimal points.

$MS = \frac{648}{8} = 81.00$

g) Calculate the population standard deviation by taking the square root of the population variance:

$s = \sqrt{s^2} = \sqrt{\frac{\sum (X - \bar{X})^2}{n}}$

Round to two decimal points.

$s = \sqrt{81.00} = 9.00$

h) Calculate the sum of the squares:

$$\sum X^2$$

324 + 36 + 961 + 25 + 81 + 484 + 64 + 25 = 2000

i) Calculate the square of the sum:

$$(\sum X)^2$$

18 + 6 + 31 + 5 + 9 + 22 + 8 + 5 = 104

$104^2 = 10816$

j) Calculate the the sum of squared deviations SS_X using this formula:

$$SS_X = \sum X^2 - \frac{(\sum X)^2}{n}$$

Check that you got the same answer as for part e above.

$SS_X = 2000 - \frac{10816}{8} = 648$

k) Suppose there was a recession and all of the incomes are cut in half. What is the new mean and the new standard deviation?

The new mean is the old mean divided by 2:

$\frac{13.00}{2} = 6.50$

The new sd is the old sd divided by 2

$\frac{9.00}{2} = 4.50$

l) Suppose there was an economic boom and all of the incomes increased by 10 thousand dollars. What is the new mean and the new standard deviation?

The new mean is the old mean plus 10:
13+10 = 23 thousand dollars

The new sd is the same as the old sd: 9.00 thousand dollars.

Problem 2) Suppose that the preferred outdoor temperature of the population is normally distrubuted with a mean of 68 degrees and a standard deviation of 8 degrees.

a) On the graph below, label the x-axis so that the bell-curve represents this distribution.

| 44 | 52 | 60 | 68 | 76 | 84 | 92 |
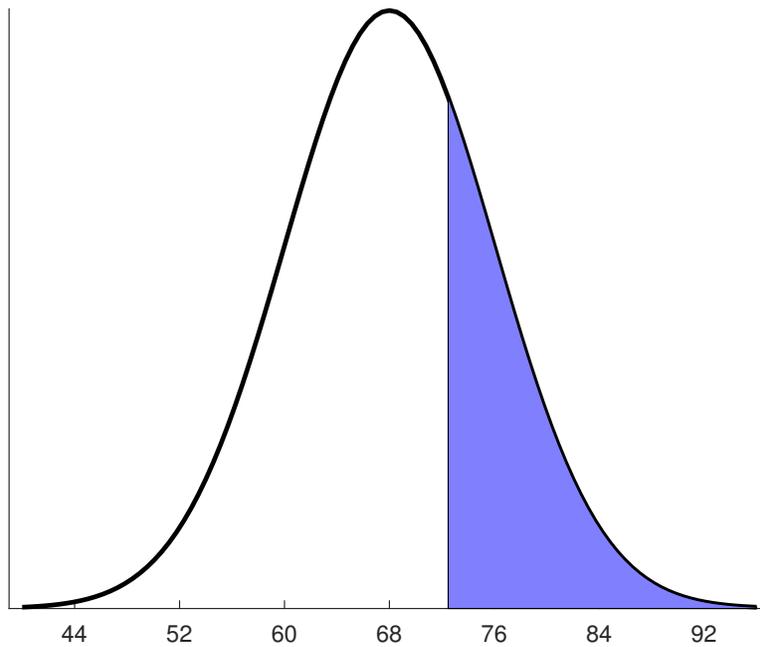
b) Shade the area under the curve for temperatures above 72 degrees.

**c)** Find the z score for 72

$z = \frac{72-68}{8} = 0.5$

d) Use the table to determine what percent of the population prefers a temperature above 72.

Using the table, $Pr(z > 0.50) = 0.3085$

30.85 percent of the population prefers a temperature above 72 degrees.

Problem 3) Assume that a population of men is normally distributed with a mean height of 69 inches and a standard deviation of 3 inches.

a) Draw a smooth bell curve below and label the axes so that it matches the probability distribution of heights.
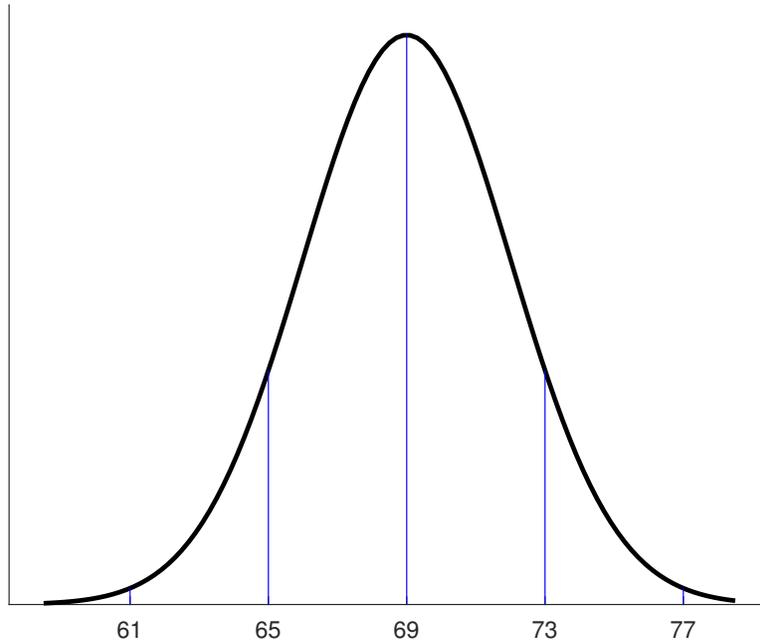
b) Mark the heights of 61, 65, 69, 73 and 77 as vertical lines on this curve.

c) Calculate the z scores for each of these 5 heights.

61 inches: $z = \frac{61-69}{3} = -2.67$

65 inches: $z = \frac{65-69}{3} = -1.33$

69 inches: $z = \frac{69-69}{3} = 0.00$

73 inches: $z = \frac{73-69}{3} = 1.33$

77 inches: $z = \frac{77-69}{3} = 2.67$

c) What proportion of men fall **below** each of these 5 heights?

$Pr(X < 61) = Pr(z < -2.67) = 0.0038$

$Pr(X < 65) = Pr(z < -1.33) = 0.0918$

$Pr(X < 69) = Pr(z < 0.00) = 0.5000$

$Pr(X < 73) = Pr(z < 1.33) = 0.9082$

$Pr(X < 77) = Pr(z < 2.67) = 0.9962$

d) These 5 heights divide the range of heights into 6 sections. Find the proportion of heights that fall within each section. Check that the proportions all add up to 1.

$Pr(X < 61) = Pr(z < -2.67) = 0.0038$

$Pr(61 < X < 65) = Pr(z < -2.67) - Pr(z < -1.33) = 0.0918 - 0.0038 = 0.0880$

$Pr(65 < X < 69) = Pr(z < -1.33) - Pr(z < 0.00) = 0.5000 - 0.0918 = 0.4082$

$Pr(69 < X < 73) = Pr(z < 0.00) - Pr(z < 1.33) = 0.9082 - 0.5000 = 0.4082$

$Pr(73 < X < 77) = Pr(z < 1.33) - Pr(z < 2.67) = 0.9962 - 0.9082 = 0.0880$

$Pr(X > 77) = Pr(z > 2.67) = 1 - Pr(z < 2.67) = 0.0038$

Check: $0.0038 + 0.088 + 0.4082 + 0.4082 + 0.088 + 0.0038 = 1$

e) What is the height for which 75% of the heights fall below? (Round to the nearest tenth of an inch)

Using the table, the z-score for p = 0.7500 is z=0.6745
Converting from z to heights:
X = 69.00 + (0.6745)(3.00) = 71.0 inches

Problem 4) It possible for more than half of the students at UW to have a greater than the average GPA? If so, how could this happen?

This could happen if we assume that 'average' means 'mean'.
If the median is greater than the mean, then more than half of the distribution
is greater than the mean, which can happen if the distribution is
negatively skewed.

Problem 5) Do the following 6 problems using R. For all examples, assume that the population is normally distributed. Provide your answers as a printout.

Very helpful hints can be found in the R files associated with the tutorials:

NormalDistribution.R

The easiest way to submit your answers is to copy the output in the 'Console' window and paste into a document to print. Your printout could look something like this:

```
pnorm(1)
[1] 0.8413447
qnorm(.5)
[1] 0
pnorm(115,100,15)
[1] 0.8413447
```

a) A Google search shows that the work of sponges has a mean of 48 and a standard deviation of 5, find value for which 47.61 percent of the work fall below.

```
qnorm(0.4761,48,5)
[1] 47.70028
```
Answer: 47.70

b) Suppose the determination of neurons has a mean of 35 and a standard deviation of 5.9, find the proportion of the determination below 32.2 and above 41.4

```
pnorm(32.2,35,5.9) + (1- pnorm(41.4,35,5.9))
[1] 0.4565616
```
Answer: 0.4566

c) Suppose the safety of students has a mean of 61 and a standard deviation of 4.5, find the value for which 96.41 percent of the safety falls above.

```
qnorm(1-0.9641,61,4.5)
[1] 52.89827
```
Answer: 52.90

d) Let's assume that the work of elements has a mean of 42 and a standard deviation of 5.2, find values that bracket the middle 63.39 percent of the work.

```
c(qnorm((1-0.6339)/2,42,5.2),qnorm(1-(1-0.6339)/2,42,5.2))
[1] 37.30023 46.69977
```

Answer: between 37.30 and 46.70

e) Let's assume that the violence of teenagers has a mean of 49 and a standard deviation of 3.3, find the proportion of the violence below 49

```
pnorm(49,49,3.3)
[1] 0.5
```
Answer: 0.5000

f) If the snow of apartments has a mean of 1 and a standard deviation of 4.7, find values that bracket the middle 98.64 percent of the snow.

```
c(qnorm((1-0.9864)/2,1,4.7),qnorm(1-(1-0.9864)/2,1,4.7))
[1] -10.59799  12.59799
```
Answer: between -10.60 and 12.60

Problem 6

Use R to calculate some statistics based on your survey. Very helpful hints can be found in the R files associated with the tutorials:

DescriptiveStatistics.R

a) Find the mean of your father's heights:

Big hint: Here's how to calculate the mean of your **mother's** heights:

```
# load in the survey data
survey <-read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W21survey.csv")
# calculate the mean of 'mheight', ignoring missing data:
mean(survey$mheight,na.rm = TRUE)
[1] 64
```
Father's heights can be found in:

```
survey$pheight
  [1] 67 64 70 72 67 70 65 67 69 66 70 NA 64 68 74 71 68 74 68 70 70 69 71 60 70
 [26] 69 67 NA 66 74 70 67 73 68 71 69 70 68 69 70 63 71 NA 70 66 68 72 72 71 66
 [51] 67 75 72 61 69 73 66 68 66 70 72 68 70 72 71 74 68 73 65 64 64 67 65 75 71
 [76] 68 66 69 72 68 70 64 72 69 69 72 65 NA 73 70 74 70 69 72 68 67 69 69 72 75
[101] 69 NA 71 67 75 69 72 75 69 72 65 64 74 70 68 69 62 77 64 65 67 73 70 62 69
[126] 71 70 69 71 NA 74 66 72 72 72 72 71 68 68 NA 70 73 73 70 75 67 68 73 70 65
[151] 72 71
```

```
# answer:
mean(survey$pheight,na.rm = TRUE)
[1] 69.33103
```

b) Find the standard deviation of the heights of the male students.

Big hint: Here's how to calculate the standard deviation of the heights of the female students:

```
# load in the survey data
survey <-read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W21survey.csv")
# find the heights of the female students
```

```
female.height <- survey$height[survey$gender=="Female"]
# calculate the standard deviation
sd(female.height,na.rm = TRUE)
[1] 2.824592
```
The heights of the men can be found with:

```
male.height <- survey$height[survey$gender=="Male"]
```

```
# answer:
sd(survey$height[survey$gender=="Male"],na.rm = TRUE)
[1] 3.152136
```

c) Use R to count the number of students by voting preference

Big hint: Here's how to count where people went to high school:

```
# load in the survey data
survey <-read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W21survey.csv")
# use 'table' to sort the entries in 'school'
school <- table(survey$school)
print(school)
In the US but outside of Washington State
                                       36
                    In Washington State
                                       92
                        Outside of the US
                                       24
```
The voting preferences can be found in:

```
head(survey$vote)
[1] "I never (or can't) vote" "Democrat"
[3] "Democrat"                "I never (or can't) vote"
[5] "I never (or can't) vote" "I never (or can't) vote"
```

```
# answer:
table(survey$vote)
             Democrat I never (or can't) vote              Independent
                  101                      39                        7
                Other               Republican
                    3                        2
```