Nathan Cooper
STAT 100
HW 5
7/24/17

Problem 1:

a) They are testing if gas mileage is greater than 28 mpg, so it is one-tailed.

b) Null hypothesis is the mean gas mileage remains 28 mpg, Alternative Hypothesis is the mean gas mileage increases > 28 mpg.

c) The type 1 error sis that the mean gas mileage remains 28 mpg, but we measure a mean gas mileage greater than 28 mpg. The CEO then wastes the money to install 100,000 units on the new cars. Also, failing to meet advertised improved gas mileage could result in a class action law suit.

d) The type 2 error is that the mean gas mileage is greater than 28 mpg, but they measure no significant increase above 28 mpg. Then the CEO does not buy the 100,000 units. No money is wasted, but it is a missed investment.

Problem 2:

a) H_o is that the spindle is 5 mm, H_a is that the spindle is not 5mm.

b) H_o is that the mean income is \$42,500/yr. H_a is that mean income is greater than \$42,500/yr.

c) H_o is that the mean commute distance remains 15 miles. H_a is that mean commute distance is greater than 15 miles.

Problem 3:

μ = \$250.00, n = 100, x_bar = 234.85 , s = 95.23.
H_o is the mean donation is \$250.00
H_a is the mean donation is not \$250.00
This is a two tailed test.
$t_{stat} = |X\text{-bar} – \mu)| / (s/\sqrt{n}) > 1.96$ to reject Ho.
$t_{stat} = |\ 234.85-250|/(95.23/(100)\text{^}0.5) =$ **1.59**
1.59 is not greater than 1.96 so we cannot reject the null hypothesis. There is not a statistically significant difference between our sample mean of \$234.85 and the claimed population mean of \$250.00. We can claim that the mean donation is \$250.00.

Problem 4:

```
> prop.test(170,500,p=.4)

        1-sample proportions test with continuity correction

data:  170 out of 500, null probability 0.4
X-squared = 7.2521, df = 1, p-value = 0.007082
alternative hypothesis: true p is not equal to 0.4
95 percent confidence interval:
 0.2988758 0.3836225
sample estimates:
   p
0.34
```

p = 0.40 is not within our confidence interval. We must reject the null hypothesis. The candidate's approval rating is not 40%.

Problem 5:

x_bar = 5.23, n = 64, s = 0.24, μ = 5.5

```
>pt((5.23-5.5)/(0.24/64^0.5), df=63)
[1] 3.237498e-13
```

We reject the null the population mean is 5.5 oz and accept the alternative that population mean is less than 5.5 oz.

Problem 6:

μ = 50000, n =30, x_bar = 46500, s = 9800

```
> tsum.test(mean.x = 46500, n.x = 30,s.x=9800,mu=50000,alternative =
'less')

        One-sample t-Test

data:  Summarized x
t = -1.9562, df = 29, p-value = 0.03007
alternative hypothesis: true mean is less than 50000
95 percent confidence interval:
       NA 49540.12
sample estimates:
mean of x
    46500
```

We must reject the null hypothesis that population mean is 50,000 miles an accept the alternative hypothesis that population mean is less than 50,000 miles.

Problem 7:

μ = 14, n =40, x_bar = 18, s = 6

```
> tsum.test(mean.x = 18, n.x = 40,s.x=6,mu=14,alternative =
'two.sided')

        One-sample t-Test

data:  Summarized x
t = 4.2164, df = 39, p-value = 0.0001426
alternative hypothesis: true mean is not equal to 14
95 percent confidence interval:
 16.08111 19.91889
sample estimates:
mean of x
       18
```

We must reject the null hypothesis that the mean number of cigarettes smoked is 14, and accept the alternative hypothesis that the mean number of cigarettes smoke is not 14.

Problem 8:

p = 0.60, n = 64, x= 34

```
> prop.test(34,64,0.60, alternative = 'two.sided')

        1-sample proportions test with continuity
        correction

data:  34 out of 64, null probability 0.6
X-squared = 0.99023, df = 1, p-value = 0.3197
alternative hypothesis: true p is not equal to 0.6
95 percent confidence interval:
 0.4032547 0.6554866
sample estimates:
      p
0.53125
```

We have insufficient evidence to reject the null hypothesis that true p is 0.6. We can claim p = 0.6.

Problem 9:

p = 0.13, n = 76, x = 2

```
> prop.test(2,76,0.13, alternative = 'less')

        1-sample proportions test with continuity
        correction

data:  2 out of 76, null probability 0.13
X-squared = 6.3363, df = 1, p-value = 0.005915
alternative hypothesis: true p is less than 0.13
95 percent confidence interval:
 0.00000000 0.08571316
sample estimates:
         p
0.02631579
```

I disagree with Newsweek. The contingent survey has sufficient evidence to reject the null and accept the alternative hypothesis that the percent of Americans who claim to have interacted with angles is less than 0.13.

Problem 10:

p = 0.585, n = 1000, x = 622

```
> prop.test(622,1000,0.585, alternative = 'greater')

        1-sample proportions test with continuity
        correction

data:  622 out of 1000, null probability 0.585
X-squared = 5.4876, df = 1, p-value = 0.009576
alternative hypothesis: true p is greater than 0.585
95 percent confidence interval:
 0.5959762 1.0000000
sample estimates:
    p
0.622
```

He is correct. There is sufficient evidence for him to reject the null hypothesis and accept the alternative hypothesis that p is greater than 0.585.

Problem 11:

x_bar = 83.6, s_x = 4.3, n_x = 36, y_bar = 79.2, s_y = 3.8, n_y = 36

```
> tsum.test(mean.x = 83.6,s.x = 4.3,n.x = 36, mean.y = 79.2,s.y =
3.8,n.y = 36)

        Welch Modified Two-Sample t-Test

data:  Summarized x and y
t = 4.6005, df = 68.957, p-value = 1.859e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 2.491991 6.308009
sample estimates:
mean of x mean of y
     83.6      79.2
```

The 95% confidence interval does not span 0. Therefore, we must reject the null hypothesis and accept the alternative that the difference of the means is not 0.

Problem 12:

a) yes_1 = 32, n_1 =52, yes_2 = 163, n_2 = 209

```
> prop.test(c(32,163),c(52,209))

        2-sample test for equality of proportions with
        continuity correction

data:  c(32, 163) out of c(52, 209)
X-squared = 5.1265, df = 1, p-value = 0.02356
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.320193854 -0.008845527
sample estimates:
   prop 1    prop 2
0.6153846 0.7799043
```

b) The 95% confidence interval does not span zero. We must reject the null hypothesis and accept the alternative hypothesis that the proportion of teen mothers to twenty-something mothers is different.

Problem 13:

yes_1 = 44, n_1 =80, yes_2 = 41, n_2 = 90 yes_2 was rounded to 41 from the 40.5 calculated from 0.45 reported in the problem text.

```
> prop.test(c(44,41),c(80,90))

        2-sample test for equality of proportions with
        continuity correction

data:  c(44, 41) out of c(80, 90)
X-squared = 1.1569, df = 1, p-value = 0.2821
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.06726415  0.25615304
sample estimates:
   prop 1    prop 2
0.5500000 0.4555556
```

We have insufficient evidence to reject the null hypothesis that the two proportions are equal.

Problem 14:

```
before <- c(6,8,10,9,5,12,9,7)

after <- c(10,12,9,12,8,13,8,10)

> t.test(before,after, alternative = "greater",paired=TRUE)

        Paired t-test

data:  before and after
t = -2.7325, df = 7, p-value = 0.9854
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -3.38669      Inf
sample estimates:
mean of the differences
                     -2
```

We have insufficient evidence to reject the null hypothesis that the mean productivity increased.

Problem 15:

```
> stocks <- c(0.07,0.58,0.24,0.04,0.07)
> bonds <- c(48,323,79,16,63)
> chisq.test(x=bonds,p=stocks)

        Chi-squared test for given probabilities

data:  bonds
X-squared = 41.692, df = 4, p-value = 1.932e-08
```

We must reject the null hypothesis that mutual fund investors' attitudes toward corporate bonds are the same as their attitude toward stocks and accept the alternative that their attitudes are different.

Problem 16:

```
> eastern_ed <- c(1105,31,229,485)
> western_ed <- c(574,15,186,344)
> table_list <- c(eastern_ed,western_ed)
> table_list
[1] 1105   31  229  485  574   15  186  344
> wsj <- matrix(nrow = 4, ncol = 2,table_list)
> rownames(wsj)=c("Full Time", "Part Time", "Self-
Employed","Unemployed")
> colnames(wsj)=c("Eastern Ed.", "Western Ed.")
> wsj
              Eastern Ed. Western Ed.
Full Time            1105         574
Part Time              31          15
Self-Employed         229         186
Unemployed            485         344
> chisq.test(wsj)

        Pearson's Chi-squared test

data:  wsj
X-squared = 23.373, df = 3, p-value = 3.376e-05
```

We must reject the null hypothesis that the employment status of the two groups is independent of region and accept the alternative hypothesis that the employment status depends on region.

Problem 17:

```
>mydata=read.csv("http://people.fas.harvard.edu/~mparzen/stat100/movi
es.csv")

> table_list <- c(mydata$length)
> run_time <- table_list
> rating <- as.factor(c(rep(1,6),rep(2,6),rep(3,6),rep(4,6)))
> rating <- as.factor(rating)
> summary(aov(run_time~rating))
            Df Sum Sq Mean Sq F value  Pr(>F)
rating       3   3632  1210.7   6.656 0.00269 **
Residuals   20   3638   181.9
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Yes, at the 5% significance level we can conclude that that at least one group has a different mean run time as compared to the other groups.