Homework 05

Christina Lee

10/6/2021

1.

You hypothesize that the average person is smarter than Sarah Palin. You know her IQ is 100. You give an IQ test to 100 randomly selected people, and get a mean of 104 and standard deviation of 22. Please show your work for each question.

a. What is your null hypothesis?

$$H_0 : \mu = 100$$

b. What is your research hypothesis?

$$H_a : \mu > 100$$

c. What is your test statistic?

SE = SD/ sqrt n = 22/sqrt(100)
SE = 2.2

$$Test\ statistic = \frac{\bar{x} - \mu_0}{se}$$

$$\frac{104 - 100}{2.2} = 1.\bar{81}$$

d. Do you prefer a one-tailed or two-tailed test here, and why?

I would prefer doing a one tailed test because we are only interested in whether or not x > 100 (only in one specific direction). Such that, we are only interested if the average person is "smarter" than Sarah Palin. Moreover, the given sample mean, 104 is already slightly greater than 100, which is why I don't think it is needed to conduct a two-tail test to test the other direction, however, it is more stringent and a better practice to do two-tailed tests which is why I will still conduct a two tailed test.

e. What is your a and threshold (t statistic) value or values for your rejection region? (Whatever a you prefer is fine, just be sure to state it and explain why you chose it.)

This is a two tail test, and we have a threshold value of 1.98 and -1.98
alpha: 0.05 (standard and conservative alpha value)

```
qt(0.975,99) # upper threshold
```

```
## [1] 1.984217
```

```
qt(0.025,99) # lower threshold
```

```
## [1] -1.984217
```

f. Can you reject the null under a one-tailed test?

Yes, since p-value 0.04 is < 0.05 we reject the null hypothesis. We can also look at the t statistic value. Since our t statistic value 1.81 > threshold value 1.66 we can also reject the null hypothesis.

```
1-pt(1.81,99)   # p-value under one tailed test.
```

[1] 0.03666486
```
qt(0.95,99)    # threshold value for one tailed test.
```

[1] 1.660391

   g. Can you reject the null under a two-tailed test?

To get the p-value for a two-tailed test, we take our p-value for the one tailed test and multiply it by 2.

$$2 * 0.04 = 0.08$$

P-value for the two tailed test would be $0.08 > 0.05$ therefore, we fail to reject the null hypothesis.

We can also look at the threshold values for a two tailed test, which is 1.98 and -1.98. Since our t statistic $1.81 < 1.98$ we fail to reject the null hypothesis.
```
qt(0.975,99) # upper threshold
```

[1] 1.984217
```
qt(0.025,99) # lower threshold
```

[1] -1.984217

   h. What is your 95% confidence interval?

This is the 95% CI for two tailed test.

$$CI = \bar{x} \pm t * \frac{sd}{\sqrt{n}}$$

$$CI = 104 \pm 1.98 * \frac{22}{\sqrt{100}}$$

$$CI = 104 \pm 4.356$$

$$95\% \, CI = [99.6, 108.4]$$

   i. What is the p-value for your test results?

P-value for a one tailed test with an alpha of 0.05 is 0.04. Again, since 0.04 is < 0.05 we reject the null hypothesis.
```
1-pt(1.81,99)
```

[1] 0.03666486

P-value for a two tailed test with an alpha of 0.05 is 0.08. Again, since 0.08 is > 0.05 we fail to reject the null hypothesis.
```
2*(1-pt(1.81,99))
```

[1] 0.07332972

   2.

You hypothesize that men and women have different skill levels in playing Tetris. To test this, you have 50 men and 50 women play the game in a controlled setting. The mean score of the men is 1124 with a standard deviation of 200 and the mean score for the women is 1245, also with a standard deviation of 200.

   a. Are these scores statistically significantly different? Show your work.

$$H_a : \mu_m \neq \mu_w$$

$$H_o : \mu_m = \mu_w$$

This is our equation for se_diff.

$$se_d iff = \sqrt{se_1^2 + se_2^2}$$

Let's first calculate the se for men and women.

$$men = se_1 = \frac{sd}{\sqrt{n}} = \frac{200}{\sqrt{50}} = 28.3$$

$$women = se_2 = \frac{sd}{\sqrt{n}} = \frac{200}{\sqrt{50}} = 28.3$$

Therefore, our se_diff is equal to 40.

$$se_d iff = \sqrt{28.3^2 + 28.3^2} = 40$$

We can now calculate our t statistic.

$$T statistic = \frac{\bar{x}_w - \bar{x}_m}{se_d iff}$$

$$T statistic = \frac{1245 - 1124}{40} = 3$$

Lastly, we need to calculate the degrees of freedom. Since our n and sd are equal across the two groups, we can simply calculate the df like such:

$$df = 2n - 2$$

$$df = 2 * 50 - 2 = 98$$

```r
qt(0.975,98) # upper threshold for two-tailed test.
```

```
## [1] 1.984467
```

```r
qt(0.025,98) # lower threshold for two-tailed test.
```

```
## [1] -1.984467
```

```r
2 * pt(3,98,lower.tail = FALSE) # P-value of two tailed test. alpha= 0.05
```

```
## [1] 0.003423264
```

1. T statistics = 3. Threshold = 1.98. The t statistic is above the critical threshold, so we reject the null and conclude that the difference in scores is statistically significant.
2. The p-value is 0.0034 < 0.05, so we again reject the null.
3. 95% CI = difference in mean (1245-1124=121) plus or minus the threshold value times the SE_diff : 121 plus or minus 1.98 * 40 = [41.8,200.2]. This does not include 0, so again we reject the null.

b. Do you reject your hypothesis or the null? What do you conclude from this experiment?

We reject the null hypothesis and accept our research hypothesis. This means that the difference in score is indeed statistically significant –men and women do have different skill levels in playing Tetris.

3.

You think drinking the night before an exam might help performance on the exam the next morning. To test this, you select 100 of your closest friends, and randomly get 50 of them drunk the night before the exam, which you denote the treatment group. The next day, the treatment group gets a mean of 78 with a standard deviation of 10 and the control group gets a 75 with a standard deviation of 5.

a. Does the evidence show that drinking helped exam performance?

$$H_a : \mu_t > \mu_c$$

$$H_o : \mu_t = \mu_c$$

This is our equation for se_diff.
$$se_diff = \sqrt{se_1^2 + se_2^2}$$

Let's first calculate the se for treatment and control.

$$treatment = se_1 = \frac{sd}{\sqrt{n}} = \frac{10}{\sqrt{50}} = 1.41$$

$$control = se_2 = \frac{sd}{\sqrt{n}} = \frac{5}{\sqrt{50}} = 0.71$$

Therefore, our se_diff is equal to 1.58.

$$se_diff = \sqrt{1.41^2 + 0.71^2} = 1.58$$

We can now calculate our t statistic.
$$T statistic = \frac{\bar{x}_w - \bar{x}_m}{se_diff}$$
$$T statistic = \frac{78 - 75}{40} = 1.9$$

Lastly, we need to calculate the degrees of freedom. Since our n and sd are not equal across the two groups, we calculate the df using the following equation:

$$df = \frac{se_diff^4}{se_a^4/n_a - 1 + se_b^4/n_b - 1}$$

$$df = \frac{1.58^4}{1.41^4/50 - 1 + 0.71^4/50 - 1} = 73$$

```
qt(0.975,73) # upper threshold for two-tailed test.
```

```
## [1] 1.992997
```

```
qt(0.025,73) # lower threshold for two-tailed test.
```

```
## [1] -1.992997
```

```
2 * pt(1.9,73,lower.tail = FALSE) # P-value of two tailed test. alpha= 0.05
```

```
## [1] 0.0613828
```

1. T statistics = 1.90 Threshold = 1.99. The t statistic is not above the critical threshold, so we fail to reject the null and conclude that drinking does not help improve performance on exams.

2. The p-value is 0.06 > 0.05, so we again fail to reject the null.

3. 95% CI = difference in mean (78-75=3) plus or minus the threshold value times the SE_diff : 3 plus or minus 1.99 * 1.58 = [-0.14,6.14]. This does include 0, so again we fail to reject the null.

4.

Using data of your choosing (or using simulated data), use R to conduct the following tests, and explain the results you get:

a. A standard one-sample hypothesis test.

To conduct this one-sample hypothesis test, I used R's built in data set "co2" which measured the atmospheric concentrations of CO2 (ppm) monthly from 1959 to 1997, and a total of 468 observations were made. Here, I want to see if the mean co2 measured was different from 335 ppm.

```
t.test(co2, alternative="two.sided", mu=335)
```

```
##
##  One Sample t-test
##
## data:  co2
## t = 2.9683, df = 467, p-value = 0.003148
## alternative hypothesis: true mean is not equal to 335
## 95 percent confidence interval:
##  335.6941 338.4130
## sample estimates:
## mean of x
##  337.0535
```

```
qt(0.975,467)
```

```
## [1] 1.965057
```

```
qt(0.025,467)
```

```
## [1] -1.965057
```

Results:
t statistic is 2.97, df is 467, sample mean= 337, threshold hold value is plus or minus 1.97, with a p-value of 0.003 and a 95% confidence interval of [335.7, 338.4].
Since the p-value 0.003 < 0.05 we reject the null.
t statistic 2.97 > 1.97 threshold value so, we reject the null.
95% CI [335.7,338.4] does not include 0, so again we reject the null.

Therefore, the difference in measured co2 from the "co2" data set is indeed significantly different from 335 ppm.

b. A difference-in-means test with independent samples.

I think sleeping 8 hrs before test can help improve test taking abilities the next day. I picked 30 random students and asked 15 of them to sleep 8 hrs before test day (treatment) and accessed their test scores.

```
treatment = c(80,87,85,100,79,80,88,69,79,77,69,89,96,91,82)
control = c(70,66,80,78,80,71,90,82,64,59,60,80,63,93,91)
t.test(treatment,control,conf=0.95, alternative="two.sided")
```

```
##
##  Welch Two Sample t-test
##
```

```
## data:  treatment and control
## t = 2.2302, df = 26.332, p-value = 0.03447
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   0.652096 15.881237
## sample estimates:
## mean of x mean of y
##  83.40000  75.13333
```

```r
qt(0.975,26.3)
```

```
## [1] 2.054389
```

```r
qt(0.025,26.3)
```

```
## [1] -2.054389
```

Results:
t statistic is 2.23, df is 26.3, threshold hold value is plus or minus 2.05 with a p-value of 0.034 and a 95% confidence interval of [0.65, 15.88].
Since the p-value $0.034 < 0.05$ we reject the null.
t statistic $2.23 > 2.05$ threshold value so, we reject the null.
95% CI [0.65,15.88] does not include 0, so again we reject the null.

Therefore, the difference in test scores between our treatment and control group is indeed significantly different.

    c. A difference-in-means test with dependent samples (ie, a paired t-test).

Here I wish to test to see if giving mice treatment x will impact their weight(grams) in 5 months. In this experiment, I will have 20 total mice, which I will first measure their weight before the treatment and then measure the weights of the same 20 mice after the treatment to see if there is a significant difference in their weights.

```r
beforeT = c(119, 143.6, 187.7, 213, 210.4, 132.9, 122.2, 185.5, 205.2, 193.7,119, 123.1, 198.5, 264,
            293.3, 122, 114.6, 176.2, 134.2, 116.8)
afterT = c(232, 201.2, 223.1, 328, 327.9, 343, 303.9, 332.3, 342.2,211, 234, 331, 265.5, 323, 317.5,
           319, 335, 365.2, 321.7, 247.3)
# Create a data frame
mice_weight <- data.frame (group = rep(c("beforeT", "afterT"), each=20),
                           weight = c(beforeT, afterT)
                           )
# Compute paired t-test
t.test(weight ~ group, data= mice_weight, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  weight by group
## t = 8.5724, df = 19, p-value = 5.915e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   95.57256 157.31744
## sample estimates:
## mean of the differences
##                 126.445
```

Results:

6

t statistic is 8.57, df is 19, mean differences is 126.44, with a p-value of 5.915e-8 and a 95% confidence interval of [95.57, 157.32].
Since the p-value 5.915e-8 < 0.05 we reject the null.
95% CI [95.57, 157.32] does not include 0, so again we reject the null.

Therefore, the difference in weight between our treatment and control group is indeed significantly different. Hence, treatment x can significantly impact mice weights.

    d. Manually verify the results in (a) using the mean and sd as calculated by R (ie, you don't have to manually calculate the mean or sd by hand!).

Here, we can see that all results matches with the results in (a).

```
mean(co2) # mean of co2 data set by R
```

```
## [1] 337.0535
```

```
sd(co2) # sd of co2 data set by R
```

```
## [1] 14.96622
```

SE = SD/ sqrt n = 14.97/sqrt(468)
SE = 0.69

$$Test\ statistic = \frac{\bar{x} - \mu_0}{se}$$

$$\frac{337.05 - 335}{0.69} = 2.97$$

```
2*(1-pt(2.97,467)) # p-value by R
```

```
## [1] 0.003131382
```

This is the 95% CI:

$$CI = \bar{x} \pm t * \frac{sd}{\sqrt{n}}$$

$$CI = 337.05 \pm 1.97 * \frac{14.97}{\sqrt{468}}$$

$$CI = 337.05 \pm 1.97 * 0.69$$

$$95\%\ CI = [335.69, 338.41]$$