

INSH5301 Intro Computational Statistics

Ali Banijamali

02/10/2020

1. You teacher has a friend, Dr. Huang, he thinks his maximum fastball velocity is faster than average adult. His max velocity is 52 miles per hour (True story). So you teacher measured 100 randomly selected adult, and obtain a mean velocity of 48 miles per hour, and standard deviation of 5 miles per hour (This part is fiction). Please answer the following questions, measure in miles per hour (mph) and show your work for each question.

1.a. Formalize the null hypothesis.

ANS.

We are only interested in values of $\mu < 52$, so I will only use one-tailed test.

Based on that, our Null hypothesis is:

$$H_0 : \mu = 52$$

1.b. What's the alternative hypothesis?

ANS.

Our Alternative (Research) hypothesis is: $H_a : \mu < 52$

1.c. What is test statistic can you use to test the hypothesis?

ANS.

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

$$se = \frac{s}{\sqrt{n}} = \frac{5}{10} = 0.5$$

$$\text{Test Statistics} = \frac{48 - 52}{0.5} = -8$$

1.d. Do you prefer a one-tailed or two-tailed test in this case? Explain why.

ANS.

A two-tailed test is more conservative and it would have a narrower rejection region (2 regions), but I think since we are only interested to see if the average is below Dr. Huang's record, one-tailed would be enough. We are interested only in one tail of the distribution. If we were to use a two-tailed test, our research hypothesis had to be in the following form:

$$H_a : \mu \neq 52$$

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

ANS.

If we want to define the rejection region, we must assume an alpha and do it based on that. Let's assume alpha 0.05 and a one-tailed test. Since our research hypothesis is less than null, the rejection region would be:

```
qt(0.05, 100-1) # 100-1 is degree of freedom
```

```
## [1] -1.660391
```

Alternatively, we could calculate the p-value for our test. Since the T-statistic is negative and we are using one-tailed, our p-value is:

```
pt(-8, 100-1) # 100-1 is degree of freedom
```

```
## [1] 1.200152e-12
```

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

ANS.

Yes, according to 1.e., both by p-value and the rejection region, we can reject the null hypothesis.

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

ANS.

Let's test based on a two tailed. For two tailed, assuming a p-value of 0.05, the rejection region would be:

```
# Two tailed: 0.05/2=0.025
```

```
# Test Statistic Larger than:
```

```
qt(0.975, 100-1) # 100-1 is degree of freedom
```

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

```
# and T-statistic Smaller than:
```

```
qt(0.025, 100-1)
```

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

Since $-8 < -1.984217$, the null hypothesis is still rejected.

1.h. What is the 95% confidence interval of the test?

ANS.

The 0.95 confidence interval is:

$$P(\bar{x} - 1.96se \leq \mu \leq \bar{x} + 1.96se)$$

$$\bar{x} = 48$$

$$se = \frac{s}{\sqrt{n}} = \frac{5}{10} = 0.5$$

Therefore,

$$CI_{0.95} = [48 - 1.984 \times 0.5, 48 + 1.984 \times 0.5] = [47.008, 48.992]$$

It can be seen that the null mean is outside the confidence interval.

1.i. What is the p-value of the test results?

ANS.

The p-value for one-tailed was calculated in 1.e., Let's calculate it for a two-tailed test here:

Since our t-statistic is negative, the one-tailed p-value is calculated using `pt(-8, 100-1)`. The two-tailed p-value is twice the p-value of one-tailed:

```
2*pt(-8, 100-1) # 100-1 is degree of freedom
```

```
## [1] 2.400304e-12
```

2. Your friend thinks that men and women have different skill levels in playing Tetris - for whatever reason -. To test this, you have 50 men and 50 women play the game in a controlled setting. The mean score of the men is 1,124 with a standard deviation of 200 and the mean score for the women is 1,245, also with a standard deviation of 200.

2.a. Test if the male average is statistically greater than the female average.

ANS.

The 3 parts of this question are very similar to a great extent, I think for part a and b, we should use one-tailed test and look at one side of the distribution (positive side), but for part c, we should do a two-tailed test. Let's start by defining the hypothesis: $H_0 : \mu_{men} = \mu_{women}$

Our Alternative (Research) hypothesis is: $H_a : \mu_{men} > \mu_{women}$

$$\text{Test Statistic} = \frac{\bar{x}_{men} - \bar{x}_{women}}{se_{diff}}$$

$$se_{diff} = \sqrt{se_1^2 + se_2^2}$$

$$se_{men} = \frac{s_{men}}{\sqrt{n_{men}}} = \frac{200}{\sqrt{50}} = 28.28427$$

$$se_{women} = \frac{s_{women}}{\sqrt{n_{women}}} = \frac{200}{\sqrt{50}} = 28.28427$$

$$se_{diff} = 40$$

$$\text{Test Statistic} = \frac{1124 - 1245}{40} = -3.025$$

Now let's calculate the degree of freedom. Since n and s are equal in both samples:

$$df = 2n - 2 = 2(50) - 2 = 98$$

Since our research hypothesis is greater than null, assuming a p-value of 0.95, the rejection region would be:

```
qt(0.95, 98)
```

```
## [1] 1.660551
```

Test statistic (-3.025) is less than 1.660551. Therefore, we will reject the research theory. I don't understand the point of doing this test in the first place?! The average of men is already less than women, how can we prove that men have a higher average?

2.b. Test if the female average is statistically greater than the male average.

ANS.

To a great extent the same as part a. with some differences in the details:

$$H_0 : \mu_{women} = \mu_{men}$$

Our Alternative (Research) hypothesis is: $H_a : \mu_{women} > \mu_{men}$

$$\text{Test Statistic} = \frac{\bar{x}_{women} - \bar{x}_{men}}{se_{diff}}$$

$$se_{diff} = \sqrt{se_1^2 + se_2^2}$$

$$se_{men} = \frac{s_{men}}{\sqrt{n_{men}}} = \frac{200}{\sqrt{50}} = 28.28427$$

$$se_{women} = \frac{s_{women}}{\sqrt{n_{women}}} = \frac{200}{\sqrt{50}} = 28.28427$$

$$se_{diff} = 40$$

$$\text{Test Statistic} = \frac{1245 - 1124}{40} = 3.025$$

Now let's calculate the degree of freedom. Since n and s are equal in both samples:

$$df = 2n - 2 = 2(50) - 2 = 98$$

Since our research hypothesis is greater than null, assuming a p-value of 0.95, the rejection region would be:

```
qt(0.05, 98)
```

```
## [1] -1.660551
```

Now here, the 3.025 is greater than 1.660551. It is far away in the right (positive) tail. So we can reject the null hypothesis and accept the research hypothesis.

2.c. Test if the two averages are different.

ANS.

The test samples are two independent groups. First, Let's define the Null and Alternative Hypothesis:

$$H_0 : \mu_{women} = \mu_{men}$$

$$H_a : \mu_{women} \neq \mu_{men}$$

In this case, our Test Statistic is calculated as following:

$$\text{Test Statistic} = \frac{\bar{x}_{women} - \bar{x}_{men}}{se_{diff}}$$

$$se_{diff} = \sqrt{se_1^2 + se_2^2}$$

$$se_1 = \frac{s_1}{\sqrt{n_1}} = \frac{200}{\sqrt{50}} = 28.28427$$

$$se_2 = \frac{s_2}{\sqrt{n_2}} = \frac{200}{\sqrt{50}} = 28.28427$$

$$se_{diff} = 40$$

$$\text{Test Statistic} = \frac{1245 - 1124}{40} = 3.025 \quad \# \text{ Here it doesn't matter which one is } se_1 \text{ and which one is } se_2$$

Now let's calculate the degree of freedom. Since n and s are equal in both samples:

$$df = 2n - 2 = 2(50) - 2 = 98$$

Now we can find our rejection region assuming an $\alpha = 0.05$:

```
# Two tailed: 0.05/2=0.025
# Test Statistic Larger than:
qt(0.975, 98) # 98 is degree of freedom
```

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

```
# and T-statistic Smaller than:
qt(0.025, 98)
```

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

Since the T-statistic (3.025) is larger than 1.984467, the alternate hypothesis is accepted. Meaning that the averages are different.

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.

Does the evidence show that drinking helped exam performance?

ANS.

This question is very similar to 2.c. Only the degree of freedom is different. I will use the same two-tailed test and a p-value of 0.05:

The test samples are two independent groups. First, Let's define the Null and Alternative Hypothesis:

$$H_a : \mu_{drink} > \mu_{no drink}$$

$$H_0 : \mu_{drink} = \mu_{no drink}$$

In this case, our Test Statistic is calculated as following:

$$\text{Test Statistic} = \frac{\bar{x}_{drink} - \bar{x}_{no drink}}{se_{diff}}$$

$$se_{diff} = \sqrt{se_{drink}^2 + se_{no drink}^2}$$

$$se_{drink} = \frac{s_{drink}}{\sqrt{n_{drink}}} = \frac{10}{\sqrt{50}} = 1.414214$$

$$se_{no drink} = \frac{s_{no drink}}{\sqrt{n_{no drink}}} = \frac{5}{\sqrt{50}} = 0.7071068$$

$se_{diff} = 1.581139$
 Test Statistic = $\frac{78-75}{1.581139} = 1.897366$

Now let's calculate the degree of freedom. Since s is different in both samples:

$df = \frac{se_{diff}^4}{\frac{se_a^4}{n_a-1} + \frac{se_b^4}{n_b-1}} = \frac{1.581139^4}{\frac{1.414214^4}{50-1} + \frac{0.7071068^4}{50-1}} = 72.05877$ Now we can find our rejection region assuming an $\alpha = 0.05$:

```
# Two tailed: 0.05/2=0.025
# Test Statistic Larger than:
qt(0.975, 72.05877) # 98 is degree of freedom
```

```
## [1] 1.993436
```

```
# and T-statistic Smaller than:
qt(0.025, 72.05877)
```

```
## [1] -1.993436
```

Since the T-statistic (1.897366) is smaller than 1.984467, the alternate hypothesis is rejected. Meaning that drinking doesn't affect the exam results. However, if we had used the one-tailed test, the research hypothesis could be acceptable (1.897366 > 1.666276)

4. Using data of your choosing use R to conduct the following tests, and explain (i.e. interpret) the results you get. You can download your own data from the internet - in such case provide a link to it - or data from R's built-in datasets as long as you are not using the class examples for this week:

4.a A standard one-sample hypothesis test.

```
air <- airquality
t.test(air$Temp, alternative="two.sided", mu=50)

##
## One Sample t-test
##
## data: air$Temp
## t = 36.437, df = 152, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 50
## 95 percent confidence interval:
## 76.37051 79.39420
## sample estimates:
## mean of x
## 77.88235
```

Here we tested whether the average of the Temperature is different from 50 degrees ($H_0 : \mu_0 = 50$). The alternative (research) hypothesis is proved here. The p-value is less than 2.2e-16, much less than 0.05 or even 0.01. Also the null mean is outside the 95 percent confidence interval.

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

ANS. Here, I use the same dataset from the previous part. I will check to see if the average temperature is different in July compared to September:

```
july <- air[air$Month==7, 'Temp']
september <- air[air$Month==9, 'Temp']
# We are testing whether the diff in means is 0 (means are not different) in a two-sample test
t.test(july, september, alternative="two.sided", mu=0)
```

```
##
## Welch Two Sample t-test
##
## data:  july and september
## t = 4.0927, df = 43.127, p-value = 0.0001834
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  3.552674 10.453778
## sample estimates:
## mean of x mean of y
## 83.90323 76.90000
```

According to the results, means are not the same, they are different and the p-value for it is 0.0001834, much smaller than 0.05. Also the null mean is outside the 95 percent confidence interval.

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

ANS. Since it is hard to find such data, I randomly made two groups of numbers. Let's say group.1 are the results of test.1 on 10 people and group.2 are the results of test.2 on the same group. Since these numbers are completely random, I expect to see no significant difference between them

```
set.seed(1)
group.1 <- sample(1:10,10,replace=T)
group.2 <- sample(1:10,10,replace=T)
# We are testing whether the diff in means is 0 (means are not different) in a two-sample test
t.test(group.1, group.2, alternative="two.sided", mu=0, paired=T)
```

```
##
## Paired t-test
##
## data:  group.1 and group.2
## t = -2.7222, df = 9, p-value = 0.02352
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -6.2254334 -0.5745666
## sample estimates:
## mean of the differences
## -3.4
```

Here we can not reject the null hypothesis. the p-value is 0.9405 and the mean is inside the 95 percent confidence interval. Therefore our research hypothesis is rejected and there is no difference between two groups. We expected this result.

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

```
# mean:
air.mean <- mean(air$Temp)
air.mean
```

```
## [1] 77.88235
```

```
# standard deviation:
air.sd <- sd(air$Temp)
air.sd
```

```
## [1] 9.46527
```

```

# n (count):
air.n <- length(air$Temp)

# standard error:
air.se <- air.sd/sqrt(air.n)
air.se

## [1] 0.7652217

# test score:
air.ts <- (air.mean-50)/air.se # 50 is the mu
air.ts

## [1] 36.43696

# region according to two-tailed and p-value = 0.95:
# Test Statistic Larger than:
qt(0.975, air.n-1)

## [1] 1.975694

# and T-statistic Smaller than:
qt(0.025, air.n-1)

## [1] -1.975694

```

our t-statistic (36.43696) is far away from 1.975694. Therefore the null hypothesis is rejected.