

Recall from lecture that confidence intervals come out of Neyman-Pearson ideas. However, instead of making a choice between two possible parameter values, we estimate an entire interval of values within which we suspect the true parameter value lies. Thus, we can use our confidence interval estimates to do two-sided hypothesis tests in the NHST framework.

First, let's re-create the `setosa.petal.length` dataset:

```
> library(dplyr)
> setosa.petal.length <- iris %>% filter(Species == "setosa") %>%
  select(Petal.Length)
```

Recall that we used this dataset to test the null hypothesis that the population mean petal length was 1.3 cm. Now we will test this null hypothesis using a confidence interval instead.

```
> t.test(setosa.petal.length$Petal.Length, conf.level = 0.95) # technically
the confidence level is 0.95 by default
```

Question #1 Paste the output from R. Write a sentence interpreting the 95% confidence interval for the population mean petal length of *setosa* irises.

One Sample t-test

```
data: setosa.petal.length$Petal.Length
t = 59.528, df = 49, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 1.412645 1.511355
sample estimates:
mean of x
 1.462
```

There is 95% confidence that the range of values for μ is between (1.412645, 1.511355)

Question #2 Based on our confidence interval, does the population mean petal length of *setosa* irises appear to be greater than 1.3 cm or less than 1.3 cm? Or can you not tell? Explain your answer.

The population mean of the petal length appears to be bigger because the interval is from (1.412645, 1.511355). The value of the mean needs to be inside said confidence interval.

Question #3 Based on our confidence interval, can we reject $H_0: \mu = 1.3$ in favor of the two-sided NHST alternative $H_a: \mu \neq 1.3$ at the 5% significance level? Why or why not?

We can reject our null hypothesis because the p-value is so low, meaning we have a high t-statistic. Thus implying we have at least one of the means not equal.

Now, let's look at the difference of population mean sepal lengths between *setosa* and *versicolor*. We'll need the original *iris* dataset for this one. Since our *setosa* and *versicolor* samples are unrelated, this is a two-sample *t* confidence interval.

```
> iris_sv <- iris %>% filter(Species %in% c("setosa", "versicolor"))
> t.test(Sepal.Length ~ Species, data = iris_sv, conf.level = 0.95)
```

Question #4 Paste the output table from R. Write a sentence interpreting the 95% confidence interval for the difference of population mean sepal lengths between *setosa* and *versicolor*.

Welch Two Sample t-test

```
data: Sepal.Length by Species
t = -10.521, df = 86.538, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-1.1057074 -0.7542926
sample estimates:
mean in group setosa mean in group versicolor
      5.006           5.936
```

On average, the *versicolor* plants are larger than the *setosa* plants.

Question #5 Based on our confidence interval, which species has longer petals? Or can you not tell? Explain your answer.

On average, the *versicolor* plants are larger than the *setosa* plants. The mean for *setosa* is 5.006 and 5.936 for *versicolor*, further corroborating the statement above.

Question #6 Based on our confidence interval, can we reject $H_0: \mu_1 - \mu_2 = 0$ in favor of the two-sided NHST alternative $H_a: \mu_1 - \mu_2 \neq 0$ at the 5% significance level? Why or why not?

You can reject the null hypothesis because 0 is not in the confidence interval.

Finally, we will do a *t* confidence interval for the population mean of paired differences (matched pairs confidence interval). We used the *straight_jeans2* dataset for this one and investigated the mean of the paired differences in maximum back pocket width (*maxWidthBackMens* and *maxWidthBackWomens*).

There's actually a way to do a paired test without getting the difference first; we just almost always look at a histogram of the differences first. Just to get the practice, we'll use a 99% CI.

```
> t.test(x = straight_jeans2$maxWidthBackMens, y =
straight_jeans2$maxWidthBackWomens, paired = TRUE, conf.level = 0.99)
```

or (if you like to do the subtraction first and then just run the one-sample test) you could copy the code from Lab 19 and include the confidence level argument.

```
> jeans_diff <- straight_jeans2 %>% mutate(maxWidthBackDiff =  
maxWidthBackMens - maxWidthBackWomens)  
> t.test(jeans_diff$maxWidthBackDiff, conf.level = 0.99)
```

Question #7 Paste the output from R. Write a sentence interpreting the 99% confidence interval for the population mean of the paired differences in maximum back pocket width.

Paired t-test

data: straight_jeans2\$maxWidthBackMens and straight_jeans2\$maxWidthBackWomens
t = 1.2936, df = 19, p-value = 0.2113
alternative hypothesis: true difference in means is not equal to 0
99 percent confidence interval:
-0.4361529 1.1561529
sample estimates:
mean of the differences
0.36

We are 95% confident that the men's jeans are between 0.4361529 cm narrower and 1.1561529 wider than women's jeans of the same brand.

Question #8 Based on our confidence interval, whose jeans have wider back pockets – men or women? Or can you not tell? Explain your answer.

Since our confidence interval includes the value of 0, we cannot accurately determine if either gender will have wider/narrower pockets.

Question #9 Based on our confidence interval, can we reject $H_0: \mu_d = 0$ in favor of the two-sided NHST alternative $H_a: \mu_d \neq 0$ at the 1% significance level? Why or why not?

We cannot reject the null hypothesis because our confidence interval contains 0.

Question #10 Based on your answers in this lab, what information can you get from a confidence interval that you cannot get from a hypothesis test?

A confidence interval, as the name implies, an interval of values in which the population mean can be at a 95% confidence level. A hypothesis test locks us into using one population mean and requires us to make a fail to/reject based on these criteria.