L25: Flavors of T

April 6, 2019

Flavors of T

Flavors of T

In the guest lecture Dr. Riddell introduced the T test - which relaxes some of the assumptions we needed for the Z test. Both the Z test and the T test we have learned so far have compared a sample to a hypothesized mean.

- ▶ The tests we have looked at so far are one sample tests, meaning we have one variable of interest, and we take one sample. We're interested in knowing how the one sample differs from some null hypothesis ($H_0 : \mu = 68mm$)

Today and next lecture we will look at additional types of T test:

- ▶ Comparison of the means of two samples
- ▶ Comparison of two means from paired samples

Comparing means from two populations

Next up: comparing two population means

In many types of public health studies, rather than comparing one sample to a hypothesize value, we are interested in comparing two samples from two different populations or two groups within a population. We want to shed light on the question: Do these two samples come from two groups with different underlying means? Or,

$H_O : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 \neq 0$ (two-sided alternative)

Comparing two population means

We like to compare two means! In public health, we like to:

▶ run randomized controlled trials where we compare a treated subgroup to a placebo group. Do their mean health outcomes differ?
▶ conduct observation studies where we have exposed and unexposed individuals. Do their mean health outcomes differ?

This is covered in chapter 18 in your book. Note that we are skipping around a bit.

Comparing two population means, graphically

- ▶ Make two histograms, one for each sample
- ▶ Compare their shapes, centers (means or medians) and spreads (standard deviations)
- ▶ Could instead make two box plots and compare their medians and IQRs

Example with heights from two different populations

Comparing means from two populations
Two sample T test
CI for two sample t-test

Are the heights of US and Dutch-born men different?

```
## # A tibble: 2 x 4
##   country sample_mean sample_sd length
##   <fct>         <dbl>     <dbl>  <int>
## 1 Dutch          184.      6.75    100
## 2 USA            175.      7.16    100
```

Example with heights from two different populations

L25: Flavors of T

Comparing means from two populations
Two sample T test
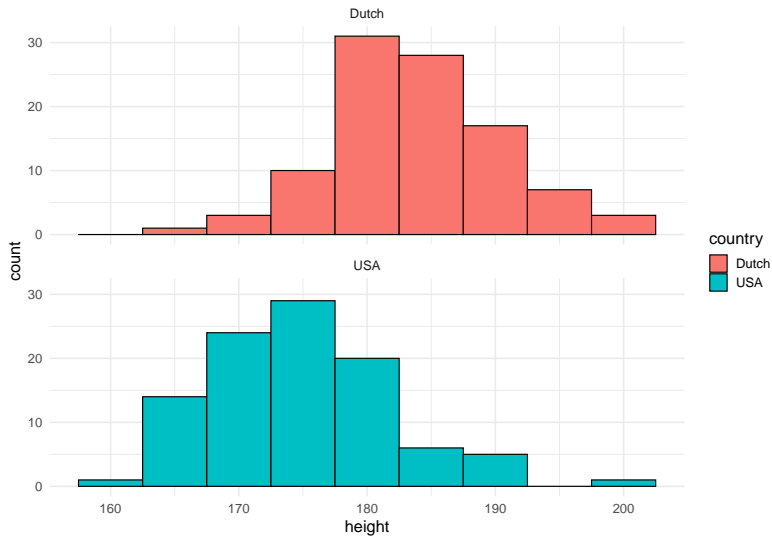CI for two sample t-test

Plotting two distributions:

ggplot(height_data, aes(x = height)) + geom_histogram(aes(fill = country), binwidth = 5, col = "black") + theme_minimal(base_size = 15) + facet_wrap(~country, nrow = 2)

Example with heights from two different populations
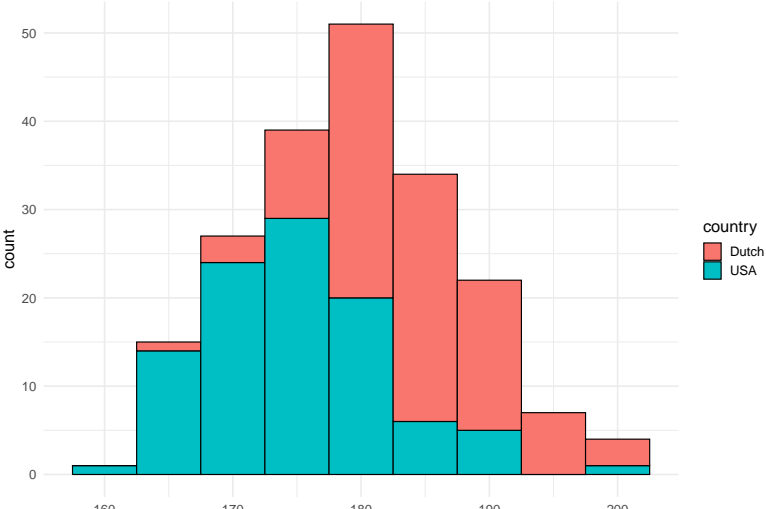
Example with heights from two different populations

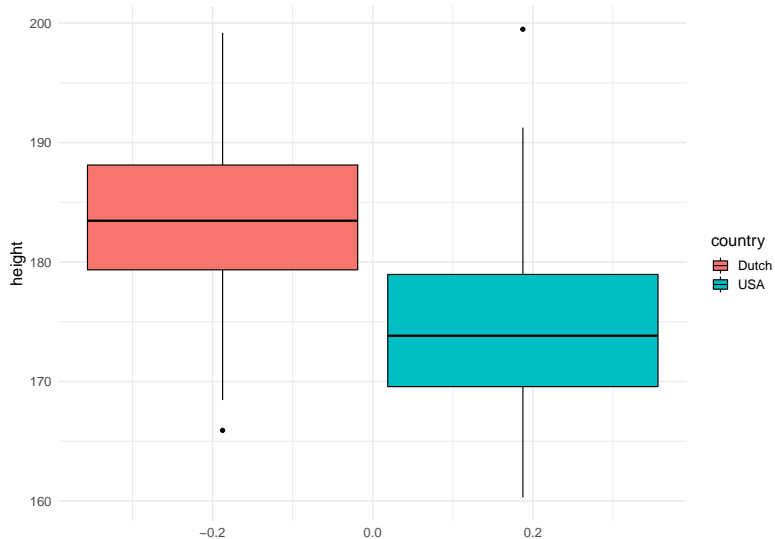If we take out the last line of that code, the histograms will be plotted on one grid. Notice what happens to the columns. . .

Example with heights from two different populations

Revisiting the code for box plots :

ggplot(height_data, aes(y = height)) + geom_boxplot(aes(fill = country), col = "black") + theme_minimal(base_size = 15)

Example with heights from two different populations

Two sample T test

Independent or Non-Independent samples?

An important requirement for using a basic t-test as a tool for this comparison, is that the observations from the two groups are independent from one another. If there is a relationship between the two groups we will will use a different type of t-test, the paired t-test which we will introduce next lecture.

When we compare these two samples, we also assume that both populations were sampled in the same (random) way and that the measurement of the variable we are comparing was done in the same way.

Conditions for inference comparing two means

- ▶ Both populations are Normally distributed. The means and standard deviations of the populations are unknown. In practice, it is enough that the distributions have similar shapes and that the data have no strong outliers.

Notation

Notation for the population parameters

Population	Variable	Population mean	Population SD
1	x_1	μ_1	σ_1
2	x_2	μ_2	σ_2

Notation for the sample statistics

Population	Sample size	Sample mean	Sample SD
1	n_1	$\bar{x}_1$	s_1
2	n_2	$\bar{x}_2$	s_2

To perform inference about the difference between $\mu_1 - \mu_2$ between the means of the two populations, we start from the difference $\bar{x}_1 - \bar{x}_2$ between the means of the two samples.

Two-sample t procedures

▶ With one-sample procedures, we had one $\bar{x}$ and we would draw the sampling distribution for $\bar{x}$. It was centered at μ with standard error $\sigma/\sqrt{n}$

▶ With two-samples, we have two sample averages $\bar{x}_1$ and $\bar{x}_2$.

▶ What do we do?

Example with heights from two different populations

Are the heights of US and Dutch-born men different?

$\bar{x}_{USA} = 174.7042$ and $\bar{x}_D = 183.6690$ so the difference is $\bar{x}_D - \bar{x}_{USA} = 8.9648$

What would happen if we took another set of two samples of USA and Dutch-born men? We would expect these sample means to change. We could draw an approximate sampling distribution for the difference between these two means.

The sampling distribution of $\bar{x}_1 - \bar{x}_2$

The distribution of the difference between two independent random variables has a mean equal to the difference of their respective means and a variance equal to the sum of their respective variances. That is:

The mean of the sampling distribution for $\bar{x}_1 - \bar{x}_2$ is $\mu_1 - \mu_2$.

The standard deviation of the sampling distribution is: $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Our *estimate* of the standard deviation of the sampling distribution is:
$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Recall the one sample t-test!

$$\frac{\bar{x} - \mu}{s/\sqrt{n}}$$

We need to generalize this by replacing each piece in the t-test by the calculations on the previous slide:

The two-sample t-test is therefore:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

The two-sample t statistics has approximately a t distribution. The approximation is accurate when both sample sizes are greater than or equal to 5.

Degrees of freedom for the two-sample t-test...

is bananas.

$$df = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{1}{n_1-1}(\frac{s_1^2}{n_1})^2 + \frac{1}{n_2-1}(\frac{s_2^2}{n_2})^2}$$

Often this is approximated by assuming that the degrees of freedom is $= n_1 - 1$ or $n_2 - 1$ whichever is smaller, or by making an assumption that the variance in the two samples is equal and approximating the df with $n_1 + n_2 - 2$

We will NOT calculate df by hand - we will use R, or we will tell you to use one of the approximation methods.

Hypothesis testing when you have two samples

$H_0 : \mu_1 - \mu_2 = 0$, obtain the two-sample t-test statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where the test p-value is the probability, when H_0 is true, of getting a test statistics t at least as extreme in the direction of H_a as that obtained, and is computed as the corresponding area under the t distribution with the appropriate degrees of freedom.

Example, continued

Let R do the work for you:

Comparing means from two populations
Two sample T test
CI for two sample t-test

```
t.test(height_data_wide %>% pull(usa),
       height_data_wide %>% pull(dutch),
       alternative = "two.sided")
```

```
##
##  Welch Two Sample t-test
##
## data:  height_data_wide %>% pull(usa) and height_data_wide %>% pull(dutch)
## t = -9.1103, df = 197.35, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -10.905418  -7.024282
## sample estimates:
## mean of x mean of y
##  174.7042  183.6690
```

Example, continued

Note that t.test gives you both the t-test results (t-statistic (called "t" in the output), df, and p-value), as well as the 95% CI. We got both because we performed a two-sided test.

Example: Zika vaccine 2017 article

L25: Flavors of T

Comparing means from two populations
Two sample T test
CI for two sample t-test

ORIGINAL ARTICLE

Safety and Immunogenicity of an Anti–Zika Virus DNA Vaccine — Preliminary Report

Pablo Tebas, M.D., Christine C. Roberts, Ph.D., Kar Muthumani, Ph.D., Emma L. Reuschel, Ph.D., Sagar B. Kudchodkar, Ph.D., Faraz I. Zaidi, M.S., Scott White, M.D., Amir S. Khan, Ph.D., Trina Racine, Ph.D., Hyeree Choi, B.S., Jean Boyer, Ph.D., Young K. Park, J.D., et al.
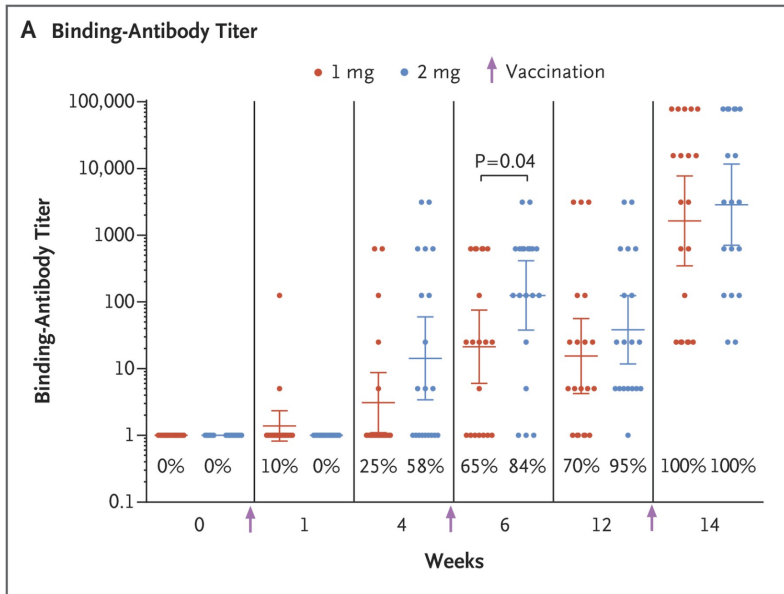
Example: Zika vaccine

L25: Flavors of T

Comparing means from two populations
Two sample T test
CI for two sample t-test

STATISTICAL ANALYSIS The antibody-binding response that was assessed on ELISA is reported as the proportion of participants in whom an antibody response developed at a given time point and as the geometric mean titer (both with 95% confidence intervals). We used Fisher's exact test to determine positive response rates and Student's t-test to compare the magnitude of the log-transformed antibody response between the two dose groups and within individuals as the change from baseline.

Example: Zika vaccine

A Binding-Antibody Titer

• 1 mg   • 2 mg   ↑ Vaccination

P=0.04

Binding-Antibody Titer

100,000
10,000
1000
100
10
1
0.1

0% 0% 10% 0% 25% 58% 65% 84% 70% 95% 100% 100%

0   1   4   6   12   14
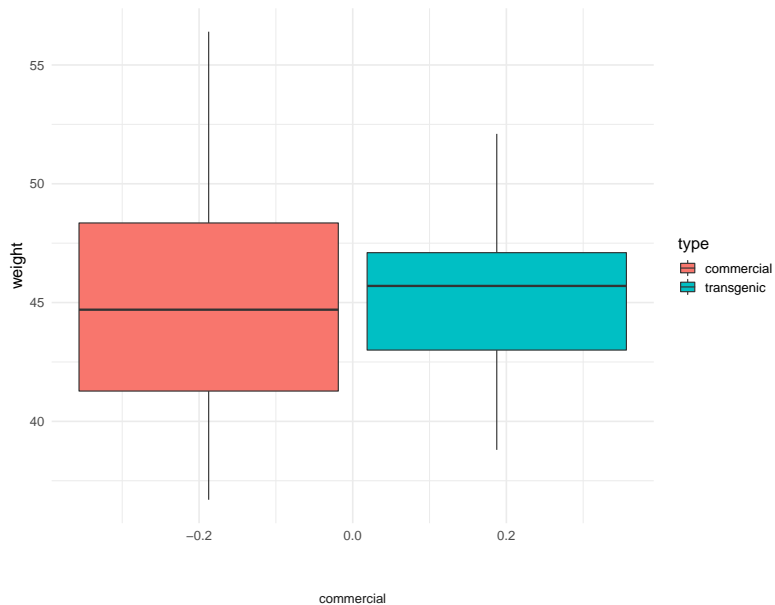
Weeks

Example: Transgenic chickens

Infection of chickens with the avian flu is a threat to both poultry production and human health. A research team created transgenic chickens resistant to avian flu infection. Could the modification affect the chicken in other ways? The researchers compared the hatching weights (in grams) of 45 transgenic chickens and 54 independently selected commercial chickens of the same breed.

Example: Transgenic chickens

commercial

Estimate the size of the difference between the two means

Comparing means from two
populations
Two sample T test
CI for two sample t-test

```
means <- chicken_data %>%
  group_by(type) %>%
  summarise(mean_weight = mean(weight))

diff_means <- means[1, 2] - means[2, 2]
diff_means
```

```
##    mean_weight
## 1  -0.1533333
```

The estimated mean difference is -0.153 grams.

Estimate the standard error

Comparing means from two populations
Two sample T test
CI for two sample t-test

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

```
chicken_stats <- chicken_data %>%
  group_by(type) %>%
  summarise(mean_weight = mean(weight),
            sd_weight = sd(weight),
            n = length(weight))
```

Use the output to calculate the SE:

$$SE = \sqrt{\frac{4.568872^2}{54} + \frac{3.320836^2}{45}} = 0.7947528$$

Calculate the t-statistic

Comparing means from two populations
Two sample T test
CI for two sample t-test

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$t = \frac{(44.98889 - 45.14222) - (0)}{0.7947528} = -0.1929279$$

What is the chance of observing the t-statistic -0.193 on the t-distribution with the appropriate degrees of freedom?

To answer this, we would need to calculate the degrees of freedom using that crazy formula. We won't do this. Instead, we will ask R to do the test for us (and verify that our calculated t-statistic matches R's test)

L25: Flavors of T

Comparing means from two
populations
Two sample T test
CI for two sample t-test

t.test in R

Pay attention to the arguments specified by `t.test`. The first argument is the weight data for the commercial chickens and the second argument is the weight data for the transgenic chickens.

```
t.test(commercial_weight, transgenic_weight, alternative = "two.sided")
```

```
##
##  Welch Two Sample t-test
##
## data:  commercial_weight and transgenic_weight
## t = -0.19293, df = 95.344, p-value = 0.8474
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.731044  1.424377
## sample estimates:
## mean of x mean of y
##  44.98889  45.14222
```

Robustness of the two-sample t procedures

▶ These procedures are more robust than the one-sample t procedures, especially if the data are skewed.

▶ When the sizes of the two samples are equal and the two populations being compared have similar shapes, the t procedures will work well for sample sizes as small as $n_1 = n_2 = 5$.

▶ When the two populations have different shapes, larger samples are needed (e.g., one skewed left and the other skewed right).

CI for two sample t-test

Confidence intervals for the two-sample t-test

For a one sample t-test the CI looked like this:

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

When we have two samples it will look like this:

$$(\bar{x_1} - \bar{x_2}) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where t^* is the critical value with area C between $-t^*$ and t^* under the t density curve with the appropriate degrees of freedom.

R Recap

We have done two types of t-test so far.

A one sample t- test will take the form:

t.test(x = x variable, alternative = greater, less or two.sided, mu = null hypothsis value)

A two sample t-test will take the form:

t.test(first sample data, second sample data, alternative = greater, less or two.sided)

Parting Humor

L25: Flavors of T

Comparing means from two populations
Two sample T test
CI for two sample t-test