

Inference: Difference between two means

Stat 145

Standard Error for the Difference between Two Means

$$SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \approx \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Notes

- We won't know σ_1 and σ_2 , so we have to approximate them.
- Approximate by replacing σ_i with s_i (plug-in estimates).
- The approximation is better if each group is approximately normal and if the sample size is larger. When a group has a sample of size 30 or less, we need to be confident that that distribution is unimodal and symmetric. For larger samples, this matters less.
- If both groups are skewed, it is better if they are skewed in the same direction.
- Use t distribution. Degrees of freedom satisfies

$$\min(n_1 - 1, n_2 - 1) \leq \text{df} \leq n_1 - 1 + n_2 - 1 = n_1 + n_2 - 2$$

- When doing the work yourself, use the smaller bound: $\min(n_1 - 1, n_2 - 2)$.
- Software will give a more precise value, which may not be an integer.
- Software values will be closer to the larger bound when the two groups are quite similar (sample sizes are approximately the same and the two standard deviations are approximately the same).
- If we suspect the populations are heavily skewed or unusual in some other way, we prefer other methods (including the bootstrap). But note that the bootstrap also works less well with smaller samples than with larger samples.

Smiles and Leniency

Are punishments more lenient for students who have smiling photos on their fact sheets? How much more?

Here is a data summary.

```
library(Lock5withR) # several data sets used here are from this package
df_stats(Leniency ~ Group, data = Smiles) %>% pander()
```

response	Group	min	Q1	median	Q3	max	mean	sd	n	missing
Leniency	neutral	2	3	4	4.875	8	4.118	1.523	34	0
Leniency	smile	2.5	3.5	4.75	5.875	9	4.912	1.681	34	0

1. Create a picture or pictures of the data before proceeding. What pictures should you make? What are you looking for or at?
2. Use an appropriate method to answer the main questions of the study: Are punishments more lenient for students who have smiling photos on their fact sheets? How much more?

Does tea boost your immune system?

Interferon gamma is a molecule that fights bacteria, viruses, and tumors. In a study to test whether interferon gamma is elevated in tea drinkers, 21 healthy, non-tea-drinkers were randomly assigned to two groups. Eleven of them were asked to drink five or six cups of tea each day and ten were asked to drink that much coffee, but no tea. After two weeks, the amount of interferon gamma in the subjects' blood was measured.

Here is a summary of the data

```
df_stats(InterferonGamma ~ Drink, data = ImmuneTea)
```

```
##           response Drink min   Q1 median   Q3 max  mean   sd  n missing
## 1 InterferonGamma Coffee  0  5.0  15.5 21.0  52 17.70 16.69 10      0
## 2 InterferonGamma   Tea   5 15.5  47.0 53.5  58 34.82 21.08 11      0
```

3. Why do you think one group was asked to drink coffee but not tea?
4. Is this an experiment or an observational study? Why?
5. Before proceeding to make a confidence interval or test a hypothesis, make a picture (or pictures) of your data. What pictures should you make? What are you looking for/at in those pictures?
6. Is there evidence that interferon gamma is elevated in the tea drinkers?

Night lights and weight gain?

A study was conducted to see whether mice who sleep in complete darkness gain more or less weight than mice who are exposed to light at night? That data set includes three light conditions: dark (LD), dim light (DM), and bright light (LL) at night. We don't know (yet) how to deal with three groups, so let's combine the two light groups and compare that combined group to the darkness group.

```
LightatNight2 <- LightatNight %>%
  mutate(some_light = Light != "LD")
df_stats(BMGain ~ some_light, data = LightatNight2)
```

```
##  response some_light min   Q1 median   Q3  max  mean   sd  n missing
## 1  BMGain      FALSE 2.79 4.757  6.33 7.035  8.17 5.926 1.899  8      0
## 2  BMGain      TRUE  3.42 7.430  9.39 10.900 17.40 9.352 3.194 19      0
```

7. This is a pretty small data set, so we need to be quite confident (especially for the smaller group) that the population distribution is approximately normal. Are there any alarming issues in your pictures? (You did make pictures, right?)
8. Does the data show that mice who sleep with light gain more weight than those that sleep in darkness? Answer this using the formula method and using randomization/bootstrap. How do the two results compare?

R can do the whole thing

The function `t.test()` can compute p-values (as long as you are comparing to 0) and confidence intervals for a mean or the difference between two means. Here is an example for the previous data set.

```
t.test(BMGain ~ some_light, data = LightatNight2)
```

```
##  
## Welch Two Sample t-test  
##  
## data: BMGain by some_light  
## t = -3.4, df = 22, p-value = 0.002  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -5.488 -1.362  
## sample estimates:  
## mean in group FALSE mean in group TRUE  
## 5.926 9.352
```

- Just want the p-value or the interval? Here's how to get less output:

```
t.test(BMGain ~ some_light, data = LightatNight2) %>% pval()
```

```
## p.value  
## 0.002341
```

```
t.test(BMGain ~ some_light, data = LightatNight2) %>% confint()
```

```
## mean in group FALSE mean in group TRUE lower upper level  
## 1 5.926 9.352 -5.488 -1.362 0.95
```

- Want a different confidence level? Here's how:

```
t.test(BMGain ~ some_light, data = LightatNight2, conf.level = 0.99) %>% confint()
```

```
## mean in group FALSE mean in group TRUE lower upper level  
## 1 5.926 9.352 -6.231 -0.6195 0.99
```

- Want a one-sided test? Here's how:

```
t.test(BMGain ~ some_light, data = LightatNight2, alternative = "less") # or "greater"
```

```
##  
## Welch Two Sample t-test  
##  
## data: BMGain by some_light  
## t = -3.4, df = 22, p-value = 0.001  
## alternative hypothesis: true difference in means is less than 0  
## 95 percent confidence interval:  
## -Inf -1.717  
## sample estimates:  
## mean in group FALSE mean in group TRUE  
## 5.926 9.352
```

9. Give `t.test()` a try on the other problems where you used the formula method. (The results might not match exactly because `t.test()` might choose a more precise degrees of freedom number.)

The Name of the Game

These methods usually go by the name “two-sample t”. You will often see that label in computer menus and computer output. “Two sample” indicates that we are comparing two groups.