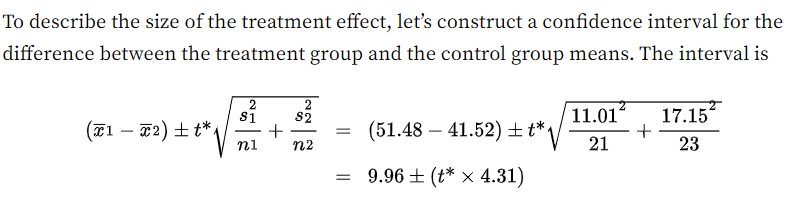
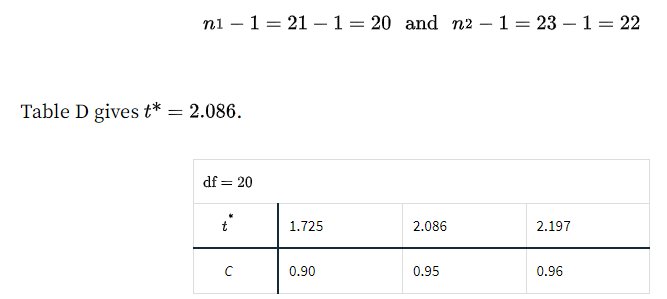
Exercises 2-4 in Part 3\_4 (comparison of two means) are borrowed from the textbook, which contains detailed solutions.

**Exercise 2 (EXAMPLE 7.15 and 7.16 in the book): The exposed group receives directed reading activities, while the control receives none. We will compare the mean of groups in terms of the Degree of Reading Power test score.**

Computing an approximate (depending on how you compute the degrees of freedom – software or the minimum number criteria introduced in the class) 95% confidence interval for the difference in means.



Using the approach in the class, we computer the quantile *t\** for a 95% confidence level as: 

Combining the difference in sample means and the margin of error, we get the final interval output of….

A black numbers and a white background

Description automatically generated

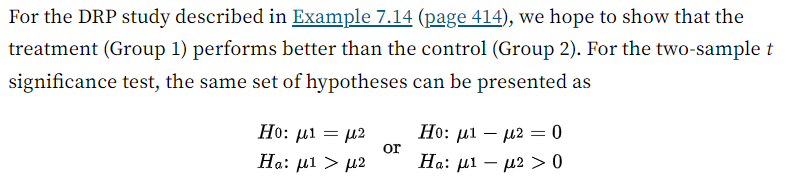
The number will not match the software-generated 95% CI (there will be a difference of ~0.2), as the methods to compute the degrees of freedom (df) will vary. This is because the sampling distribution of the difference of means is only an approximation of the t-distribution, so the *t*-quantile would differ slightly from one method to another.

**Interpretation**

Does the 95% CI include the null value (zero)? If it does not, then we conclude that the classroom that received the directed reading activity sessions has a different mean DRP test score from the group that received none.

**Now, compute the test statistic to see if the null hypothesis should be rejected.**

Start with the t-procedure for two-sample, with unequal population variance.



Is there a difference? (two-sided)

Is there an improvement?

We chose to use a two-sided test (we did not assume one group had a better or worse mean outcome than the other), so we would need to multiply the cumulative probability from the test statistic by a factor of 2 to obtain the p-value.

A math problem with numbers and equations

Description automatically generated

So, the two-sample *t-*test statistic is 2.31, and we can get the corresponding p-value as follows (make sure to multiply for the two-sided test). If using the table, look for the t-distribution table with the df = 20. The p-value will be somewhere between 0.02 and 0.04. For exact computation of p-value using the *pt* function:

P() for the one-sided p-value. For the two-sided probability,

2\*(1- pt(df = 20, q = 2.31)) # R code

**p-value = 0.03168334**

**Interpretation:** The probability of obtaining the test statistic as extreme or more extreme than 2.31 is 0.03, assuming that the null hypothesis is true. Based on the threshold of ¸we will reject the null hypothesis of no difference in the test score between the exposed (had the reading activities training) and the control (no training) groups. In other words, there is evidence that the reading activity sessions make a difference in the DRP test score. Does this match the results from CI at the level of 95% confidence (C = 1-0.05)?

The answer using the table below:

A screenshot of a test

Description automatically generated

**Exercise 3 (EXAMPLE 7.20):**

The equal variance assumption is somewhat questionable, but the textbook proceeded with the pooled t-test. I would do so if the sample sizes were larger than 40, which would give more stable estimates of standard deviations. The two samples look more or less normally distributed (but they should have used the stem/leaf plots rather than the boxplots). Anyhow, following the textbook, let’s use a one-sample pooled test (for a two-sided test, simply multiply the p-value). The book used , my example is based on (99% CI), so it is pretty hard to reject the null. We will also get a quite wide CI (as we are setting a 99% confidence level now, leading to a large margin of error).

The pooled variance and standard deviation are calculated as follows:   
A math problem with numbers and equations

Description automatically generated

If we use the 99% level, then your *t*-quantile will be…

, where the df (noted as k in the slides) will be:

*= 10 + 11 - 2 = 19*

With R studio, t-quantiles for the middle 99% are:

*qt(df = 19, p = 0.01/2)*

*qt(df = 19, p = 0.995)*

The margin of error for 99% CI is 9.23149 based on the calculation below:

2.860935\*7.385\*sqrt(1/10 + 1/11) #R code

So, the 99% CI is

or (-3.96, 14.51)

So, the range includes the difference of zero. Thus, the calcium treatment did not make a difference in terms of systolic blood pressure at the 99% confidence level.

See the results for the 90% confidence level below as used in the textbook. It still did not lead to the evidence to conclude the difference in BP across the exposed and unexposed groups, despite resulting in a much smaller margin of error.

A math equations and numbers

Description automatically generated with medium confidence

**Significance test (EXAMPLE 7.21)**

For the one-sided test,

A math equations and formulas

Description automatically generated with medium confidence

And the alternative for the two-sided test is .

Using the pooled variance from the previous example, the pooled test statistic is

A math equations with numbers and symbols

Description automatically generated

And the computation of the p-value is below. Even with the one-sided test evaluated against (least conservative test), we did not reject the null hypothesis. Multiply the p-value by 2 for the two-sided test and evaluate against . Just like the 99% Confidence Interval, the result here represents the lack of evidence to reject the null hypothesis.

A screenshot of a test

Description automatically generated

**Exercise 4 (Correction: the sample size for the first group is 25, not 9). Example 7-18 from the book.**

A screenshot of a computer

Description automatically generated

A screenshot of a graph

Description automatically generated

A screenshot of a test

Description automatically generated