Problem Set 7

JP

10/8/2021

# Paired-Samples t-test

1. Get the difference between each pair of scores
2. Square each difference score
3. Get the mean difference
4. Get the variance/standard deviation of differences
5. Get the standard error of the mean differences
6. Get the t-obtained value
7. Get the degrees of freedom
8. Get the effect size
9. Find out if the t-obtained value is statistically significant
10. Report the t-test statistic finding. Remember to include the means and standard deviations for the two time points.

*t*(9) = 4.88, *p* < .05

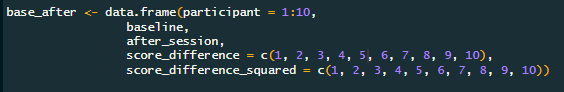
baseline (*M* = 1.63, *SD* = 0.33)

after sessions (*M* = 5.23, *SD* = 2.27)

# Formulas

You are conducting an experiment where you are interested in how many words a one year old toddler can say. Your experimental manipulation is that every child will be given two two-hour long sessions to try and build a stronger vocabulary. You want to know if these lessons are enough for children to have a stronger stronger vocabulary than when they first arrived. You decided to get baseline data by having them say all the words they know and then you asked a second time after the vocabulary sessions. What test are you running and is the difference statistically significant?

**NOTE** once you fill in the values for score\_difference and score\_difference\_squared you will run everything from **base\_after** **<-** to the double parantheses **))** like the picture below. The values in the screenshot are not the correct answers.



set.seed(093021)  
  
baseline <- rnorm(10, mean = 1.5, sd =.5)  
after\_session <- rnorm(10, mean = 4.1, sd = 2)  
  
baseline <- round(baseline, 2)  
after\_session <- round(after\_session, 2)  
  
base\_after <- data.frame(participant = 1:10,  
 baseline,  
 after\_session,  
 score\_difference = c("\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_",  
 "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_"),  
 score\_difference\_squared = c("\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_",  
 "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_"))  
  
base\_after

## participant baseline after\_session score\_difference score\_difference\_squared  
## 1 1 2.05 6.60 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 2 2 1.25 3.82 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 3 3 1.75 6.05 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 4 4 1.11 3.87 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 5 5 1.46 5.79 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 6 6 1.75 1.87 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 7 7 1.61 5.91 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 8 8 1.85 6.71 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 9 9 2.09 2.29 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 10 10 1.35 9.36 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_

base\_after$baseline

## [1] 2.05 1.25 1.75 1.11 1.46 1.75 1.61 1.85 2.09 1.35

base\_after$after\_session

## [1] 6.60 3.82 6.05 3.87 5.79 1.87 5.91 6.71 2.29 9.36

You are conducting an experiment where you are interested in getting smokers to stop smoking can help build lung capacity. You are going to test this by conducting an smoking cessation program for a year and having your sample run a mile. Before your program begins, you have every participant run a mile and get the times for each participant. After the year of the program to stop participants from smoking, you decide to have them run a mile again and track the times. What test are you running and is the difference statistically significant?

set.seed(093021)  
  
smoking <- rnorm(7, mean = 20, sd = 8.1)  
no\_smoking <- rnorm(7, mean = 8, sd = 5.3)  
  
smoking <- round(smoking, 2)  
no\_smoking <- round(no\_smoking, 2)

smoke\_df <- data.frame(participant = 1:7,  
 smoking,  
 no\_smoking,  
 score\_difference = c("\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_",  
 "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_"),  
 score\_difference\_squared = c("\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_",  
 "\_\_\_\_\_\_\_\_\_\_\_\_\_", "\_\_\_\_\_\_\_\_\_\_\_\_\_"))  
  
smoke\_df

## participant smoking no\_smoking score\_difference score\_difference\_squared  
## 1 1 28.97 11.75 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 2 2 15.91 14.22 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 3 3 24.05 6.37 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 4 4 13.70 14.61 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 5 5 19.34 7.27 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 6 6 24.07 13.16 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_  
## 7 7 21.83 7.39 \_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_

smoke\_df$smoking

## [1] 28.97 15.91 24.05 13.70 19.34 24.07 21.83

smoke\_df$no\_smoking

## [1] 11.75 14.22 6.37 14.61 7.27 13.16 7.39