

Theoretical and Methodological Foundations of Visual Computing

(Assignment 12)

Please submit your solution to ILIAS by 2025-01-26, 23:55. Solutions will be discussed in the tutorial on Thursday, 2025-01-30. Please note that source code will only be accepted if it compiles.

Exercise 12.1 (12 points)

You and your friend were playing a simple dice game, where both of you would roll five standard six-sided dice, trying to roll a larger sum than the opponent. After playing many rounds and losing your money, you suspect that they might be cheating using weighted dice. So you have recorded the average dice score (sum divided by 5) of your opponent over the last 15 rounds, giving you the sample X :

$$X = \{3.24, 4.39, 5.24, 3.83, 3.50, \\ 3.75, 4.06, 3.83, 3.54, 3.20, \\ 4.28, 3.65, 3.01, 4.69, 3.32\} \quad (1)$$

From your statistics and probability theory background you know that the population standard deviation is 0.76 and expect the mean to be 3.5, but you suspect that the actual mean is higher. Write an R script to perform a Z-test for a null hypothesis H_0 : “population mean is equal to 3.5” and an alternative hypothesis H_a : “population mean is higher than 3.5”. In this task you are not allowed to use the library functions for performing the Z-test.

- Compute the Z-statistic, and then a corresponding p-value. Instead of using a Z-table or a library function, use simple numerical integration to calculate the p-value for the obtained Z-value yourself. *Hint: you could still use a Z-table to validate your p-value computation.*
- Assuming a significance level of 0.05, do we reject the null-hypothesis?

Exercise 12.2 (12 points)

Your friend did a little experiment on the effectiveness on hand washing methods but needs your help with statistical evaluation. Download the data from ILIAS and do all the statistical analysis in R and submit your code as well.

Here is what she did¹: She tested four different methods – washing with water only, washing with regular soap, washing with antibacterial soap (ABS), and spraying hands with antibacterial spray (AS). Her experiment consisted of one experimental factor, the washing method, at four levels.

She suspected that the number of bacteria on her hands before washing might vary considerably from day to day. To help even out the effects of those changes, she generated random numbers to determine the order of the four treatments. Each morning, she washed her hands according to the treatment randomly chosen. Then she placed her right hand on a sterile media plate designed to encourage bacteria growth. She incubated each plate for 2 days at 36°C, after which she counted the bacteria colonies. She replicated this procedure 8 times for each of the four treatments.

- (a) Why is it advisable to perform an ANOVA instead of multiple t-tests when having more than 2 sample means to compare (think about what you learned in the lecture and also do some internet research)?
- (b) What are the assumptions that need to be met to perform an ANOVA? Explain them in one sentence each.
- (c) Ok, now let's get started with the data. To get an initial overview, use R to calculate mean, median, and SD per hand-washing method and visualize the data e.g. in a box or violin plot.
- (d) Check if the data meets one of the ANOVA assumptions. Which assumption did you test? Does the data meet the assumption? Use R to answer the question and show any plots needed to demonstrate your answer.
- (e) Luckily, your friend already checked the ANOVA assumptions and you are good to go. So, perform an ANOVA in R (either with the built-in functions or a library of your choice). What does the result tell you (write one or two sentences)?
- (f) You now think about post-hoc testing and find the Tukey HSD test for pairwise comparison. How is Tukey different from a normal t-test? Research again on the internet.
- (g) Perform the Tukey HSD test in R and visualize the resulting confidence intervals in a line plot similar to Figure 1
- (h) Your friend seems excited about the results and thinks about publishing them in a

¹Data and story taken from: https://das1.datadescription.com/datafile/hand-washing/?_sfm_methods=Analysis+of+Variance&_sfm_cases=4+59943&sf_paged=2

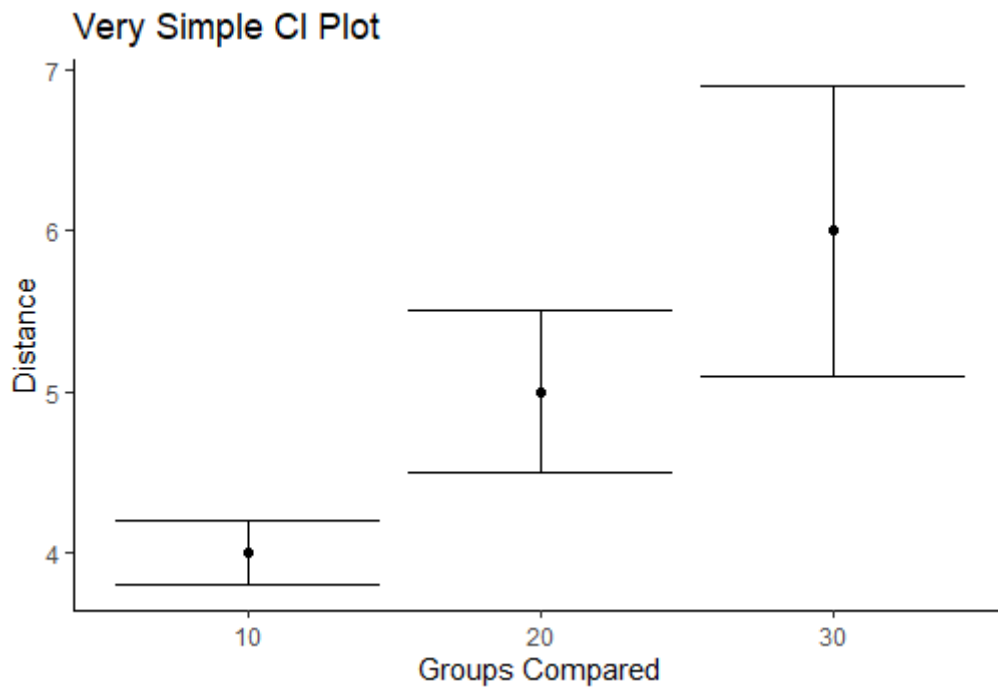


Figure 1: A very simple example of a CI plot, please try to get a prettier one ;). Note, that the points in the middle should be the differences of the mean, while the upper and lower bounds signify the CIs.

scientific journal. You are a bit skeptical of the results because of the experiment design, after having done the previous exercise sheets and listening to the lecture. Give your friend at least 3 pointers on how to improve her experiment. Write a short paragraph.