Statistics Analysis

RT2 Assignment

Samuele Pedrazzi, mat. 4673604

0. Introduction

Data analysis is the final section of the Research Track 2 assignment. On the first assignment of the Research Track 1 course, we were requested to undertake a statistical analysis, selecting some parameters to compare our results to the solution supplied by professor Carmine Recchiuto.

0.0. Tools for the statistical analysis

To do an ideal statistical analysis, I utilized MATLAB 2022a as my primary work environment, where I wrote the code to gather all of the information needed to conduct a legitimate statistical test and generate interesting graphs.

0.1. Environment setup

Instead of making the placement of the silver tokens random, I chose to employ four alternative arena designs with varied numbers of tokens to create different configurations. I built up four arenas in particular:

- **7** tokens (standard arena)
- **8** tokens
- **9** tokens
- **10** tokens

The location of the tokens is always set, which means that I just added another token while maintaining the others in the same place as before.

0.2. Data collection

The data collected were those relating to both projects: the personal one and a given solution.

The following are the most important data gathered:

- **Elapsed time** : measure of time taken with a control loop in the assignment code that the robot takes to complete a lap of the circuit, depending on the environment arena.
- **Distances** : measures of distance from the robot to the walls, characterized from golden tokens.

The distance measurements of the robot from golden tokens were taken at each iteration of the control cycle, taking into account the number of the circuit lap that the robot is traveling on. In addition, I took the time that the robot takes to complete a lap (more precisely the time is initialized when the robot grabs the first silver token and stopped after collecting the last one) for 5 times, so this means that the robot make 5 laps. The distances and times measured were taken for each arena configuration.

To have comparable data, data from each occurrence (collection of 5 laps) has been regarded *legitimate* only if the robot completed the laps without going clockwise and without slamming into the barriers.

1. Results

1.0. Preface

In most cases, the hypothesis is seen to be the most important tool in research. Its major purpose is to generate fresh ideas for experiments and observations.
Simply explained, a hypothesis is a statement that makes an unproven prediction about something. It's a form of ***educated assumption***; in fact, many studies (in robotics and other subjects) are designed specifically to test hypotheses.
Hypothesis testing is a common approach for determining if a sample of data provides enough evidence for a hypothesis to allow for generalization. As a result of hypothesis testing, we may make probability claims regarding ***population parameters*** (s).
The theory may not be proved beyond a reasonable doubt, but it is accepted in practice if it has weathered critical testing.

1.1. Basic concepts

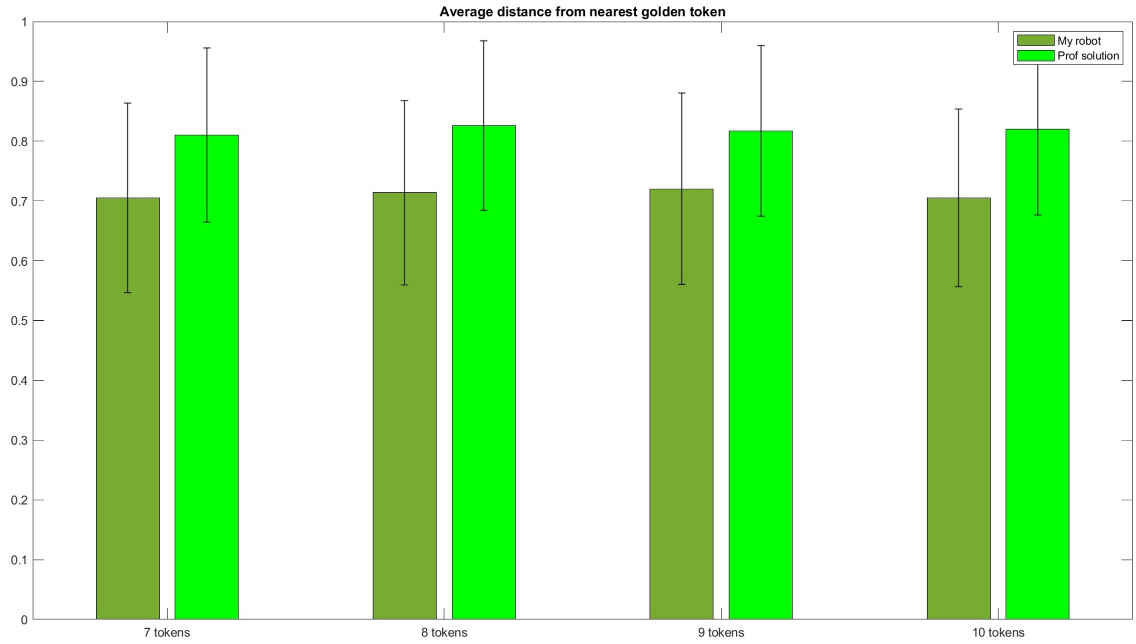
The terms "*null hypothesis*" and "*alternative hypothesis*" are frequently used in the context of statistical analysis. The null hypothesis is what we call the idea that if we compare technique A to method B in terms of superiority and proceed on the premise that both ways are equally good.

In contrast, we may believe that technique A is superior or method B is inferior, in which case we are proposing an *alternative hypothesis*. The *null hypothesis* is denoted by H_0, whereas the alternative hypothesis is denoted by H_a.

We should conclude that something else is true if our sample data do not support this *null hypothesis*. The *alternative hypothesis* is what we come up with after rejecting the *null hypothesis*. In other terms, the *alternative hypothesis* refers to the collection of alternatives to the *null hypothesis*. If we accept H_0, we must reject Ha, and if we reject H_0, we must accept H_a.

1.2. Mean distances

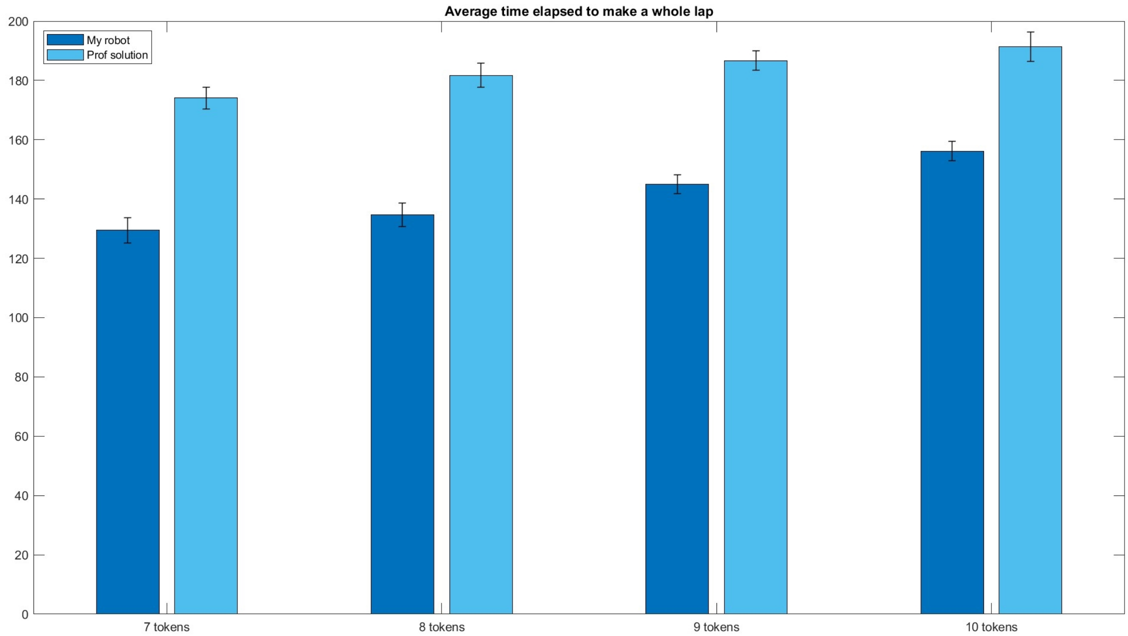
After collecting all the data regarding the distances, all the averages were made for each configuration of the arena circuit. I decide to make a comparison between the averall distance from the walls for each configuration, and plot it using a bar plot.



It is clearly visible that in each configuration the distance from the walls of my robot is on average higher with respect to the given solution, meaning that my robot stay farther from the walls.

1.3. Mean times

After gathering all of the data on the timings, I compute the average time elapsed to make a lap for each configuration of the arena. Also in this case the comparison between my mean times and the ones of the solution is plotted using a bar plot.



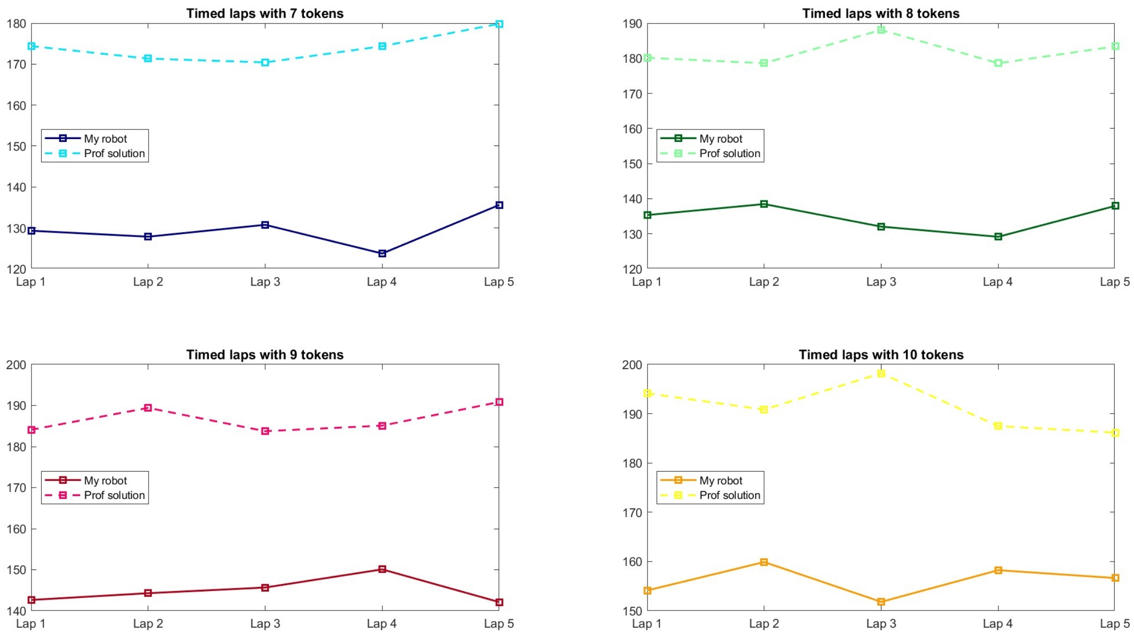
In this case, I discovered that my robot was faster than the robot in the offered solution for each configuration. Furthermore, the robot of the given solution spends roughly the same amount of time doing a full lap, but the time consumed by my robot grows according to the number of tokens.

Differences in implementation of the algorithm to move the robot can be given due to different given speeds or more or less distant localization of a silver token and subsequent approach to it, as well as to the check with respect to the walls, which allows the robot to be more in the middle of the roadway.

1.4. Lap time comparison

As shown in the image below, it is noticeable that, in general, regardless of the quantity of tokens, the robot is quicker in the way given by me. This certainly depends on how the algorithm for moving the robot in the environment was developed, as both my solution and the professor's were run on the same system and under the same conditions.

Regarding the time on a single lap, it can be observed that they are quite constant, because the timings in different laps fluctuate by a few seconds in different configurations of silver tokens, so normally the robot completes a lap in the same time whether it is the first lap or the second lap, and so on.



1.5. T-Test

A t-test (sometimes known as a Student's t-test) is a method for conducting hypothesis testing to evaluate the means of one or two populations. A t-test can be used to determine whether a single group differs from a known value (a one-sample t-test), whether two groups differ from each other (an independent two-sample t-test), or whether there is a significant difference in paired measurements (a paired, or dependent samples t-test), which is what we are looking for.

As previously stated, I chose a *paired sample* t-test since I need to compare the same experiment with two distinct implementations. I also select $\alpha = 0.05$ as the standard value (α is the significance threshold, representing the likelihood of rejecting the null hypothesis when it is true). A significance threshold of 0.05, for example, as in our situation, suggests a 5% probability of concluding that a difference exists when there is no real difference, indicating the likelihood of a type I mistake in statistics).

Regarding the t-test result, H equal to 0 indicates a failure to reject the *null hypothesis* with the specified significance level α, whereas H equal to 1 indicates a rejection of the *null hypothesis* with the given significance level α. The **p-value** is another t-test outcome that measures the level of evidence that a result is not merely a likely chance occurrence (under the assumption that the null hypothesis is correct).

I do three t-tests, one for each arena layout, on the overall average distance from the walls, the average duration, and the average distance from the walls taken in a single lap.

For the first t-test, I compare the mean distance from the walls in each of the four arena configurations and both implementations, thus my data sets are made up of four items.

The result of the first t-test is:

	Value
H	1
p	0.00012

For the second t-test, I compared two data sets, each of which had five elements: the average lap time for each setup.

The result here is:

	Value
H	1
p	1.68e-16

For the last component of the t-test, I chose the mean distances for each lap and configuration, resulting in four t-tests in total. The code compared a data set with dimension 5 in each t-test.

Here’s the results:

	7 tokens	8 tokens	9 tokens	10 tokens
H	1	1	1	1
p	2.19e-04	9.63e-04	0.0033	1.27e-04

3. Conclusions

After a series of tests to ensure that the data was taken in the most accurate manner possible, and attempting to create the most similar conditions for the execution of my algorithm and that of the professor in order to achieve a consistent result, I discovered that in any case, both results are completely non-deterministic, and both methods are efficient in some ways and in others not.

I chose a t-test because it is appropriate for the situation; it examines two sets of data that differ from each other without regard for standard deviation or variance, as opposed to, say, a z-test, which examines the averages of data sets that differ from each other but requires the standard deviation or variance to be provided.