Statistical Analysis

Research Track 2 - Second Assignment

University of Genova

In this report, I aim to provide a comprehensive exploration of my personal approach to the statistical analysis of the performance characteristics of two distinct codes related to the first assignment for Research Track 1.

Danial Sabzevari – (5217544)

June 2023

Contents

Introduction:

The second part of the Research Track II Assignment required us to conduct a statistical analysis comparing two sets of code performance:

1. The solution I developed for the first assignment in the Research Track I course.
2. The solution provided by a professor for the same assignment.

Our task was to evaluate which solution performed better in a specific circuit.

In addition, we were instructed to modify the positions of the silver tokens randomly within the circuit.

Data Collection:

I approached this task by considering two different scenarios:

1. Silver tokens were placed in standard positions.
2. Silver tokens were placed randomly.

In both scenarios, I applied the same statistical analysis.

To prevent an unnecessary increase in simulation time, I chose not to introduce additional silver tokens into the circuit. Instead, I concentrated solely on changing their positions.

Data analysis:

I evaluated the following factors as a performance analyst:

1. In terms of silver tokens, the mean time it takes to complete a lap.
2. Silver tokens are randomly and standardly placed in the circuit to make the robot's turn decision method depend on frontal distance from gold tokens.

Methodology:

My first step was to make minor changes to the two scripts to keep their resemblance to the originals. B

Specifically, I configured the scripts to produce two distinct text files:

1. The first file is dedicated to storing data concerning the forward distance from gold tokens.
2. The second file retains information related to the number of laps completed and the corresponding time duration for each lap.

I executed both codes ten times, ensuring the robot completed five laps.

This methodology remains consistent for data collection with both systematically and randomly distributed silver tokens.

Consequently, the following datasets were produced:

1. Ten files that contain data regarding the forward distance from gold tokens when silver tokens are standardly placed.
2. Ten files that record information about the number of laps completed and the corresponding time duration for each lap under conditions of standardly placed silver tokens.
3. Ten files that capture data on the forward distance from gold tokens when silver tokens are placed randomly.
4. Ten files that store data concerning the number of laps finished and the time taken for each lap when silver tokens are randomly positioned.

Diagrams of Distances

After gathering all the necessary data, I calculated the mean distance from gold tokens for each scenario.

Here are the plots I obtained:

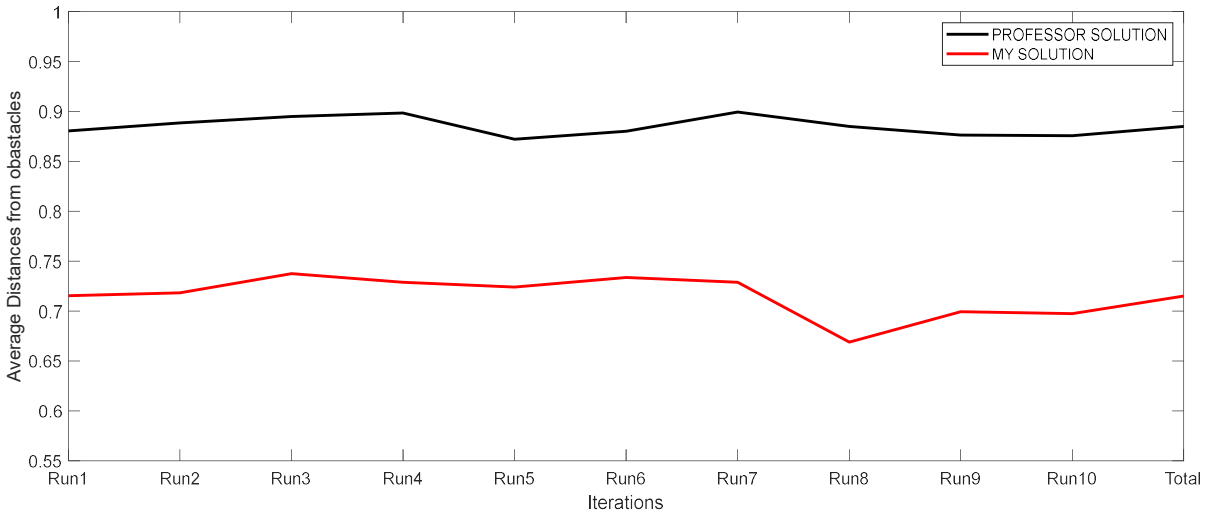

Figure 1: Distance from gold tokens - Randomly

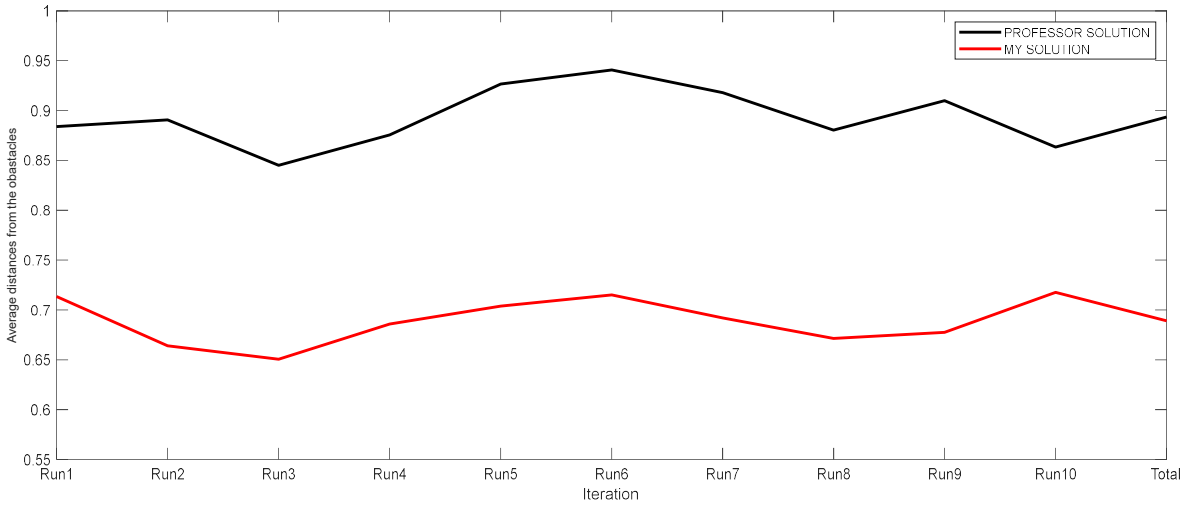

Figure 2: Distance from gold tokens - Standardly

Diagrams of Lap times

Additionally, I calculated the average time utilized by the robot to complete each lap for every code run, resulting in the generation of the following diagrams:

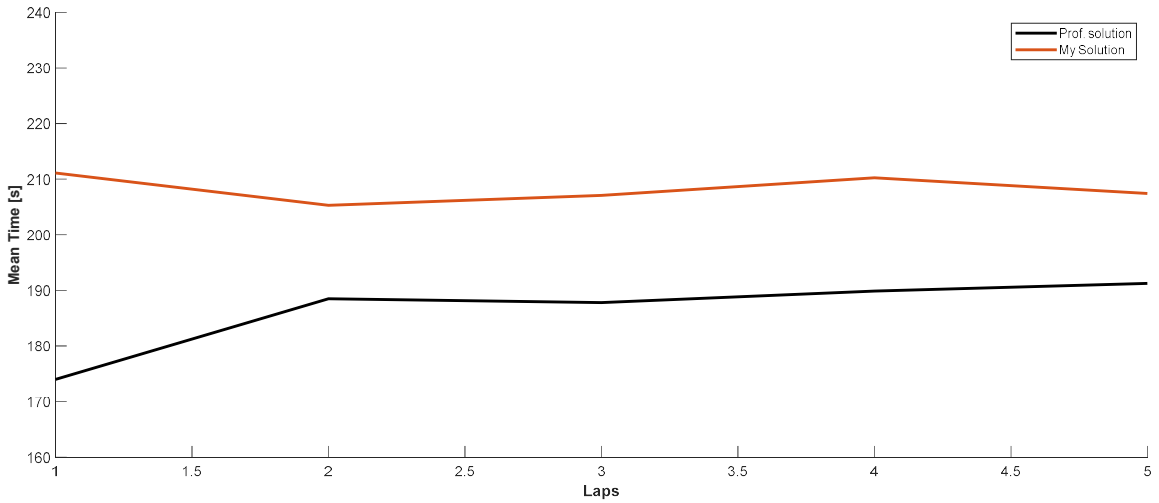

Figure 3: Figure 3: Time taken by the robot to complete each lap - Randomly.

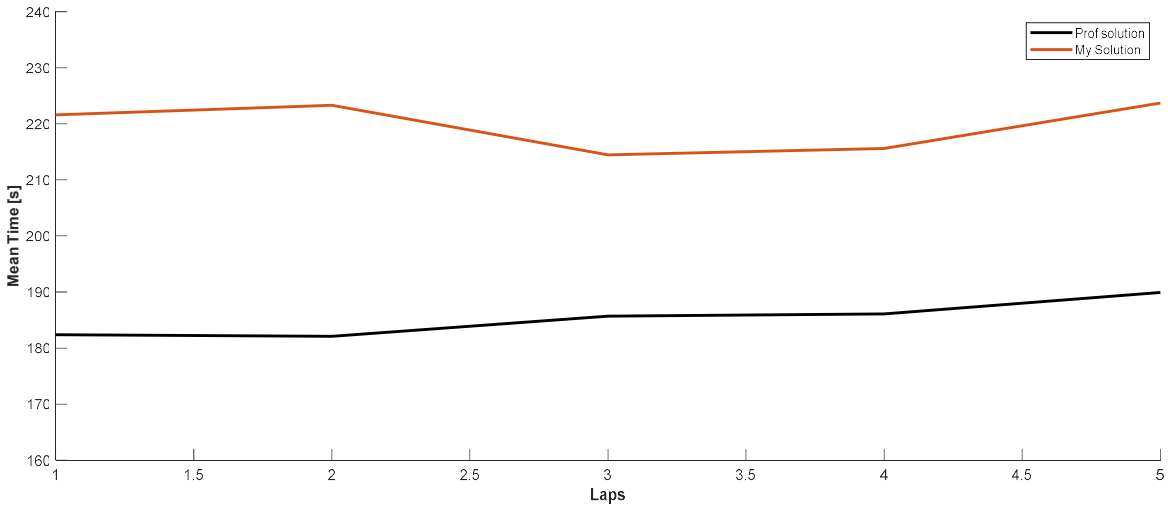

Figure 4: Time taken by the robot to complete each lap - Standardly.

With my data sets, I obtained the following boxplots:

- Boxplot distances from Gold Tokens - Standard.

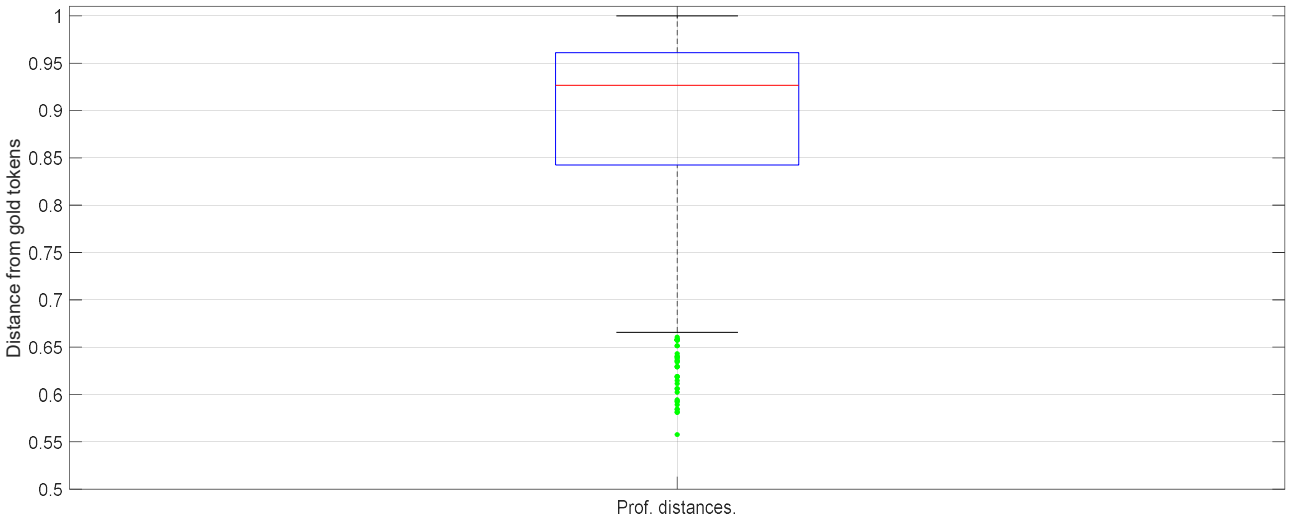

Figure 5: Boxplot distances from Gold Tokens - standard.

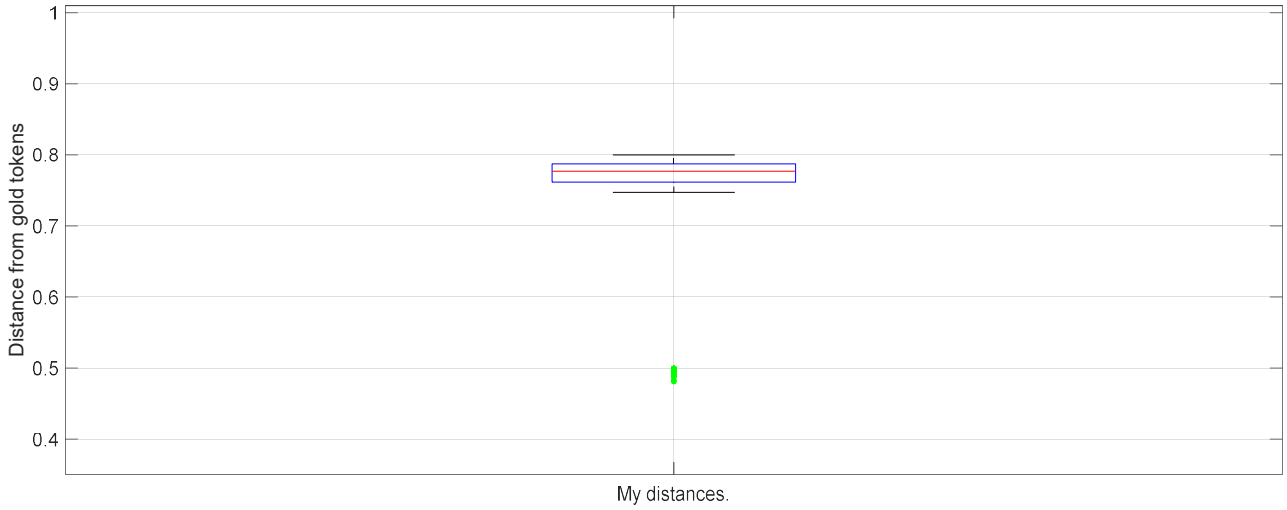

Figure 6: Boxplot distances from Gold Tokens - standard.

- Boxplot distances from Gold Tokens - Randomly.

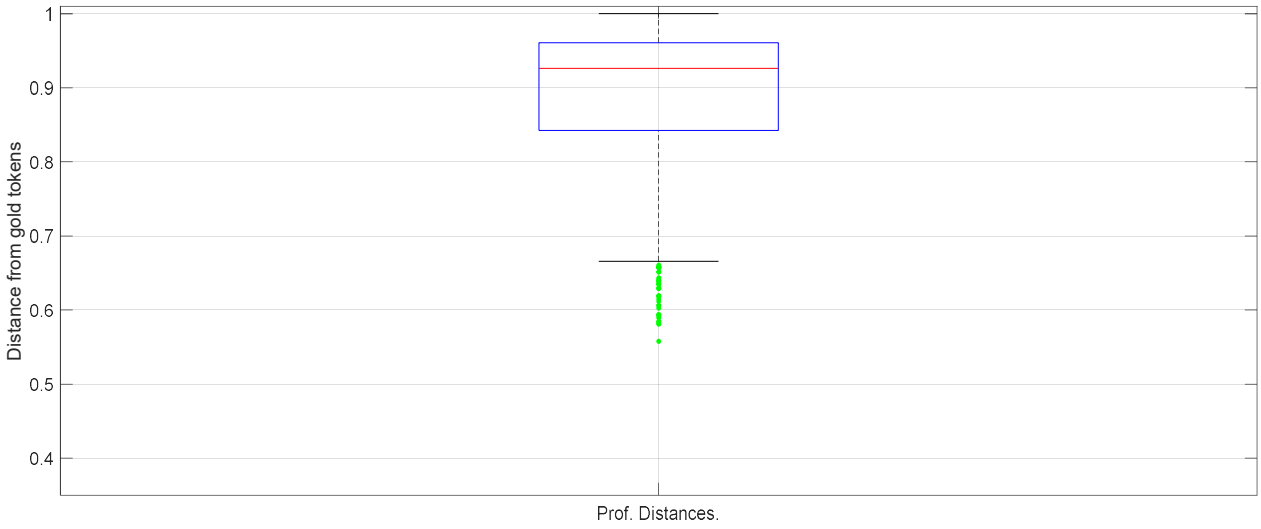

Figure 7: Boxplot distances from Gold Tokens - Randomly.

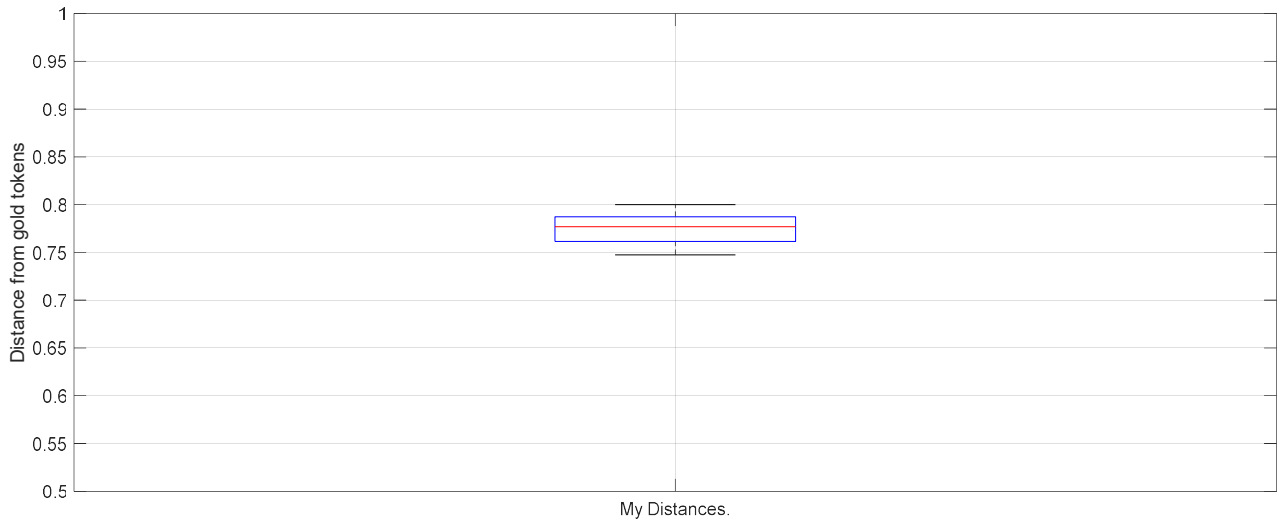

Figure 8: Boxplot distances from Gold Tokens - Randomly.

- A boxplot showing time laps with Silver Tokens **randomly** placed.

Figure 9: Time laps with Silver Tokens randomly placed.

- A boxplot showing time laps with Silver Tokens **standardly** placed.

Figure 10: Time laps with Silver Tokens standardly placed

T-Test:

To determine whether there is a significant difference between the means of my data sets, I used T-Test.
Initially, I performed a Lilliefors test to determine if my data samples were distributed normally.

The following are my tests scenarios.

Hypothesis:

Distances:

Null Hypothesis (H0):

The average distance from obstacles in the professor's solution is equivalent to the average distance from obstacles in my solution.

Alternative Hypothesis (Ha):

The average distance from obstacles in the professor's solution surpasses my solution.

Lap Times:

Null Hypothesis (H0):

The average execution time for the professor's solution matches the average execution time for my solution.

Alternative Hypothesis (Ha):

The professor's solution's mean execution time exceeds my solution's mean execution time.

Conclusion:

Silver Tokens are placed as standard:

> ***Distances:*** The value of h equals 1, leading to the rejection of the null hypothesis.

Consequently, it's safe to assert that, when silver tokens are systematically positioned, professor's solution maintains a larger distance from obstacles than my robot.

> ***Lap Times:*** Since h equals 1, the null hypothesis is rejected at the default 5% significance level.

Consequently, it's evident that when silver tokens are systematically positioned, my solution requires more time to complete a lap than the professor's solution.

Silver Tokens are placed randomly:

> ***Distances:*** Given that h is equal to 1, the null hypothesis is at the default significance level of 5%.

Thus, it can be deduced that even with randomly placed silver tokens, the professor's robot solution maintains a greater distance from obstacles compared to my robot.

> ***Lap Times:*** Given that h is equal to 1, the null hypothesis is at a 5% significance level.

Thus, it can be inferred that my solution takes more time to finish a lap than the professor's solution, even when silver tokens are dispersed randomly.