University of Genoa

Robotics Engineering

RESEARCH TRACK 2

Report assignment3 about statistical analysis

Professore:

Carmine recchiuto

Studente:

Massimo Carlini, s4678445

Session: 2022-2023

Sommario

ABSTRACT

This report aims to explain the methodology for conducting a statistical analysis based on the performance of two programs we implemented during the first assignment of the Research Track1 course. The programs considered are mine and my colleague Matteo Forni. These programs are designed to make the robot perform a simple action, which is to match all the silver tokens in the arena with the concentrically generated golden tokens located outside the silver ones. The experiment we need to carry out is to verify the performance by varying the radius at which the tokens are generated.

EXPERIMENT DESCRIPTION

The experiment involves measuring the times it takes for the two algorithms to complete their task while varying the angle at which the tokens are generated. In my case, I decided to perform 50 simulations to have more data for statistical analysis. The tokens are generated within a range of [-pi, pi], allowing for different scenarios and increasing the complexity of execution. To effectively evaluate the performance of the two programs, we need to calculate the average time taken for each simulation.

HYPOTHESIS

Before analyzing the available data, it is necessary to formulate the "null hypothesis" and an "alternative hypothesis." Only after analyzing the data can we determine which of the two hypotheses is confirmed or refuted.

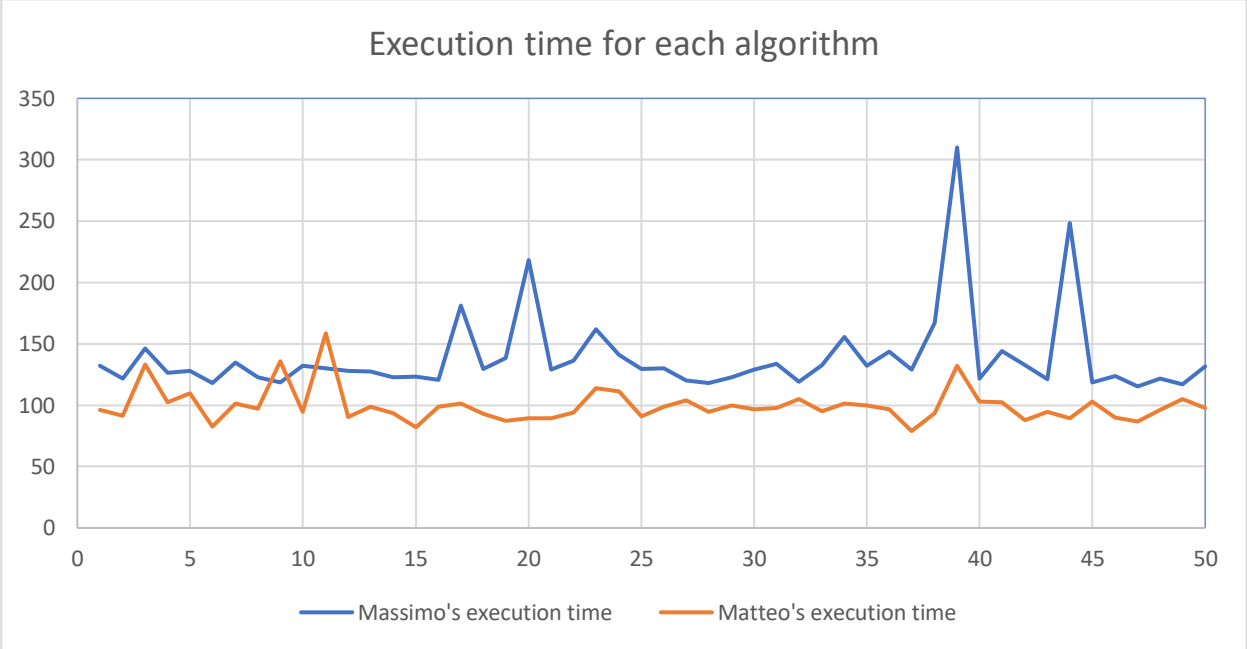
The null hypothesis assumes that the two programs do not differ significantly in terms of performance.

The alternative hypothesis, on the other hand, assumes that the two programs are statistically significantly different, with one algorithm having better performance than the other.

Once the hypotheses are established, we can proceed with the actual analysis of the data, considering the null hypothesis to be true. The analysis will be based on the T-Test, a statistical test used to compare the means of two samples and determine if there are significant differences between them.

ANALYSIS

Based on two datasets with 50 simulations each, we have enough data to create a graph that shows the differences between the two algorithms.



After obtaining the execution times, I calculated the mean times for both the algorithm.

$\mu_{Massimo}$ = 138.7633 μ_{Matteo} = 99.0575

After computing the mean time, I also calculated the standard deviation, which indicates how close the dataset is to the mean. In other words, the smaller the standard deviation, the more precise the data.

$\sigma_{Massimo}$= 34.6121 σ_{Matteo} = 14.2680

As can be seen, the two standard deviations of the algorithms are not close. However, a thorough analysis is needed to verify if there is a substantial difference. Comparing only the mean values is not sufficient.

As mentioned earlier, I decided to perform a T-Test to determine if there are significant differences between the two datasets under consideration.

To proceed with the considerations, I need to calculate:

- σ^2_{pooled}= 700.7890
- $\sigma_{\bar{x}_1-\bar{x}_2}$= 5.2423
- $t_{\bar{x}_1-\bar{x}_2}$ = 7.4504

CONCLUSIONS

Based on the conducted analysis, we can draw the following conclusions regarding the performance of the two algorithms under consideration.

After performing a T-Test to compare the mean execution times of the two algorithms, we obtained a t-value of 7.4504. We then compared this value with the critical value from the reference table for a 5% confidence level, which is 2.009 for our specific case.

Since the obtained t-value (7.4504) is much larger than the critical reference value (2.009), we can confirm that the two algorithms are statistically significantly different. Therefore, we can reject the null hypothesis that states there are no significant differences between the performance of the two algorithms.

This conclusion leads us to support the alternative hypothesis, which suggests that one of the algorithms performs better than the other.

It is important to note that the probability of the observed differences being due to chance is less than 5%, providing a solid basis for asserting the significant difference in the performance of the two algorithms.