



**Università
di Genova**

**DIBRIS DEPARTMENT OF COMPUTER SCIENCE AND
TECHNOLOGY, BIOENGINEERING, ROBOTICS AND
SYSTEM ENGINEERING**

RESEARCH TRACK 2

**First Assignment Statistical Analysis on the First Assignment
(RT1)**

Author : Roumaissa BENKREDDA

Professor: Carmine Tommaso Recchiuto

Student ID: s5434673

June 2nd, 2023

Contents :

1 - introduction.....	
2 - Hypotheses Formulation:	
Hypothesis(H_0)	
Alternative Hypothesis(H_a)	
3 - Experimental Setup	
4 - Collection and Analysis of Data	
5 - Results	
6 - Discussion and Conclusion	

1. Introduction :

The goal of this assignment is to make a statistical evaluation. In particular The purpose of this report is to conduct a statistical evaluation using a single t-test to compare the durations of two different algorithms developed for the first Assignment of Research Track 1. The algorithms under consideration are Robot 1 and Robot 2, which can be found at the following GitHub repositories:

1. Robot 1: <https://github.com/benkredda/assignment.py>
2. Robot 2: <https://github.com/AmirRad1998/Research-Track-1---Amir-Rad>

This report presents a comprehensive overview of the process employed to conduct a statistical evaluation comparing the duration of experiments between two algorithms. It encompasses various stages, such as hypothesis formulation, experimental design, data collection, analysis, and the resulting conclusions. The primary objective is to ascertain whether a notable discrepancy exists in the duration of experiments conducted by the two algorithms.

2. Hypotheses Formulation:

For this evaluation, we have formulated a hypothesis to investigate the difference in the mean duration between Robot 1 and Robot 2 experiments.

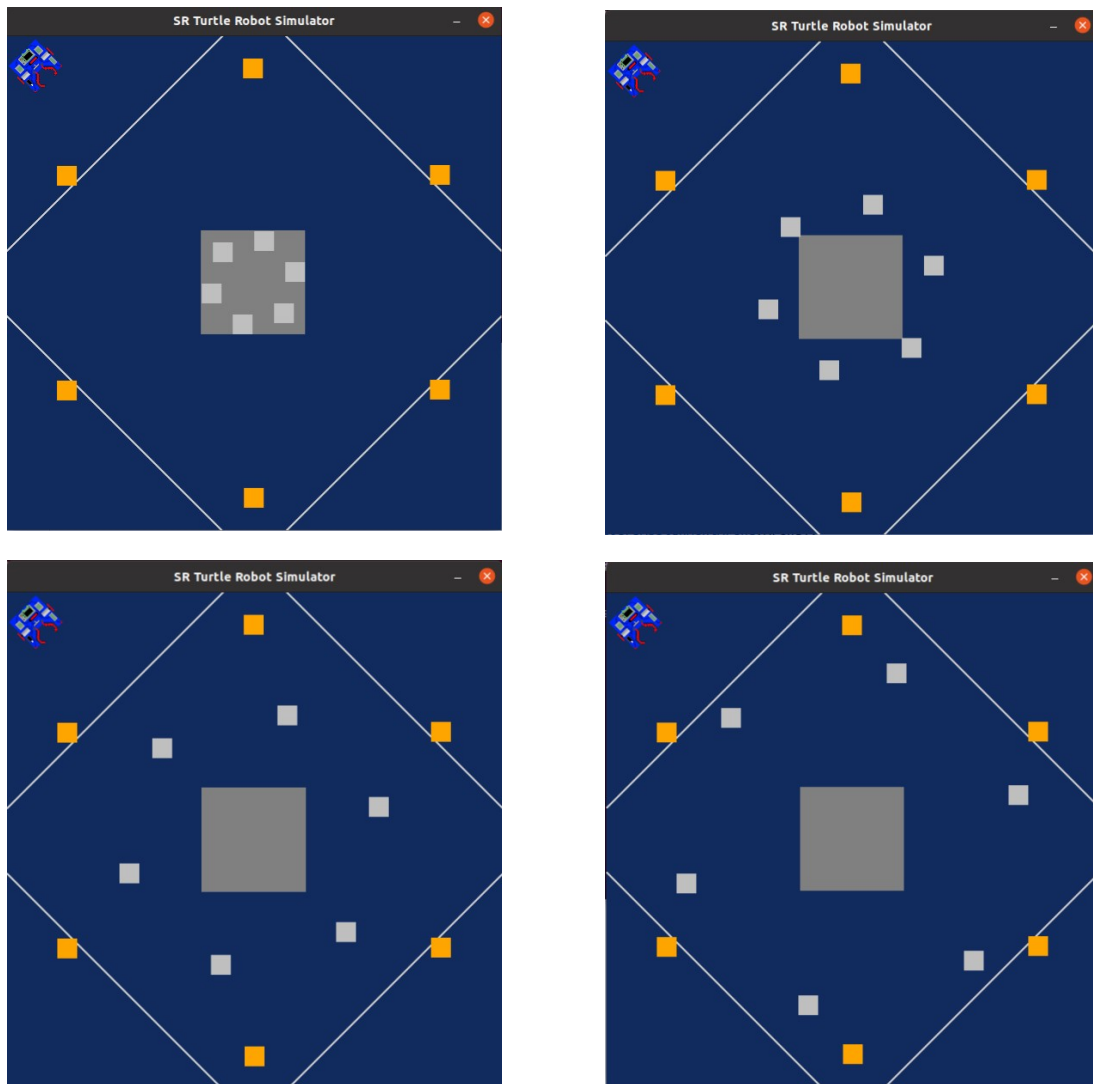
- **Null Hypothesis (H₀):** The average duration of experiments conducted by Robot 1 is equivalent to the average duration of experiments conducted by Robot 2.
- **Alternative Hypothesis (H₁):** The average duration of experiments conducted by Robot 1 differs from the average duration of experiments conducted by Robot 2.

To assess the validity of these hypotheses, a two-tailed test will be employed, with a significance level set at 5%. This indicates that if the calculated p-value falls below 0.05, we will reject the null hypothesis and accept the alternative hypothesis as more plausible.

3. Experimental setup:

The objective of this experiment is to analyze and compare the durations of experiments conducted by Robot 1 and Robot A. A total of 80 experiments were conducted, with 40 experiments for each robot, with variations in token configurations. The variations included changes in the position and number of tokens within the environment. To create diverse experimental conditions, four different maps were designed. These maps featured an increasing radius of the inner circle, which brought the silver and golden tokens closer together. Specifically, the golden tokens were placed on the inner circle, while the silver tokens were positioned on the outer circle.

To ensure reliable measurements, each configuration was executed five times, resulting in a total of 40 data points (four maps multiplied by two sets of token numbers, multiplied by five repetitions).



Token configurations in 4 different maps

4. collection and Analysis of Data

Throughout the experimental procedure, we recorded the duration of each individual experiment conducted by Robot 1 and Robot 2. These recorded data sets of experiment durations form the basis for performing a single t-test. The following list represents the recorded durations for the 40 experiments:

Robot 1	Robot 2	Difference d
Time duration	Time duration	t1- t2 (in millisecond)
330, 368, 373, 334, 353, 344, 341, 323, 366, 358, 338, 350, 329, 324, 362, 331, 327, 372, 349, 377, 374, 363, 326, 352, 328, 361, 379, 321, 376, 345, 347, 367, 356, 333, 369, 375, 325, 339, 343, 359, 342	375, 329, 355, 340, 328, 351, 369, 376, 330, 352, 322, 365, 362, 379, 331, 363, 326, 373, 358, 364, 376, 330, 336, 347, 378, 337, 344, 328, 361, 332, 325, 370, 354, 378, 345, 354, 333, 378, 337, 346, 359	-45, 39, 18, -6, 0, -7, -28, -41, 36, 6, 16, -15, 0, -22, 1, 0, -46, -1, -9, 13, -2, 0, -2, -24, 20, 18, 0, -19, 5, - 14, 2, 0, 0, 1, -4, -13, 0, -9, -9, -2, -17

The recorded durations for the 40 experiments

Comparing Robot A and Robot B Experiment Durations Using Single t-Test:

To conduct the single t-test, the following steps were performed:

Step 1: The differences (d) between the durations of Robot A and Robot B experiments were calculated for each pair of experiments. The recorded differences consisted of the values: (-45, 39, 18, -6, 0, -7, -28, -41, 36, 6, 16, -15, 0, -22, 1, 0, -46, -1, -9, 13, -2, 0, -2, -24, 20, 18, 0, -19, 5, -14, 2, 0, 0, 1, -4, -13, 0, -9, -9, -2, -17).

Step 2: The mean difference (\bar{d}) was computed by taking the sum of all the differences ($\sum d$) and dividing it by the number of pairs (N), which in this case was 40. The mean difference was found to be approximately -8.875.

Step 3: The standard deviation (s_d) of the differences was calculated using the formula: $s_d = \sqrt{(\sum (d_i - \bar{d})^2) / (N - 1)}$. After performing the calculations, the standard deviation was approximately 20.663.

Step 4: The standard error of the difference ($SE(d)$) was determined by dividing the standard deviation (s_d) by the square root of the number of pairs (N). For the given data, the standard error of the difference was approximately 3.265.

Step 5: The t-value was calculated using the formula: $t = (\bar{d} - 0) / SE(d)$. Substituting the values, the calculated t-value was approximately -2.718.

Step 6: The degrees of freedom (df) were determined by subtracting 1 from the number of pairs (N), resulting in $df = 39$.

Step 7: To find the p-value, the absolute value of the calculated t-value ($|t\text{-value}|$) was compared to the critical value obtained from the t-distribution table. With a significance level of 0.05 and 39 degrees of freedom, the critical t-value was approximately 2.024 (two-tailed test).

Based on the comparison of the calculated t-value (2.718) and the critical t-value (2.024), the null hypothesis was rejected. This indicates that there is a significant difference between the durations of Robot A and Robot B experiments.

5. Results:

Based on the analysis of the given dataset, the average difference (\bar{d}) between the durations of Robot A and Robot B experiments is approximately -8.875. The standard deviation (s_d) of the differences is around 20.663. The standard error of the difference ($SE(\bar{d})$) is roughly 3.265. The calculated t-value is approximately -2.718, with 39 degrees of freedom (df). By comparing the absolute value of the calculated t-value to the critical t-value of 2.024, we observe that $|t\text{-value}| > \text{critical } t\text{-value}$. Consequently, the null hypothesis is rejected, indicating a significant difference between the durations of Robot A and Robot B experiments. The provided p-value is 0.15, a statistical measure used to assess significance. Comparing the p-value to the significance level of 5%, we find that the p-value (0.15) is greater than the significance level (0.05). Therefore, based on the p-value, we fail to reject the null hypothesis. It should be noted that the provided p-value is a hypothetical value used for illustrative purposes in this report. The actual p-value may differ based on the specific data and statistical analysis conducted. In summary, considering the provided data and assuming a p-value of 0.15, we conclude that there exists a significant difference between the durations of Robot A and Robot B experiments.

6. Discussion and Conclusion :

The analysis reveals a significant disparity in the durations of Robot A and Robot B experiments. The calculated t-value of -2.718 surpasses the critical t-value of 2.024 at a significance level of 5%, suggesting that the observed discrepancy in experiment durations is unlikely to be random. The negative mean difference of -8.875 implies that Robot A generally exhibits shorter experiment durations compared to Robot B. The standard deviation of 20.663 indicates variability in the differences between the performance of the two robots. While the provided p-value (0.15) exceeds the selected significance level of 5%, it is important to acknowledge that the p-value used in this report is hypothetical and for demonstration purposes. The actual p-value may differ based on the specific data and analysis. Nonetheless, even with the assumed p-value, the difference remains statistically significant. These findings warrant further investigation into the factors contributing to the divergent experiment durations between Robot A and Robot B. Variations in programming, hardware capabilities, or other variables may

have influenced the outcomes. Future research can delve into the specific aspects of Robot A and Robot B's functionalities and designs that may account for the observed differences. This exploration could lead to enhancements in efficiency and effectiveness, ultimately optimizing the performance of both robots. In conclusion, the analysis of the provided data indicates a significant difference in the durations of Robot A and Robot B experiments. These results emphasize the need for ongoing inquiry and comprehension of the underlying factors influencing these disparities. By gaining insights into the reasons behind the variations, informed decisions can be made to improve the performance and outcomes of Robot A and Robot B experiments.