A Statistical Analysis of Execution Times for Two Robot Programs

Anna Possamai

5621738

Introduction

This statistical analysis focuses on the efficiency of two programs used for the movement of a robot in an environment with obstacles. The first program was implemented by me, while the second by my colleague Andrea Bolla.

The main objective is to assess whether there are significant differences in the timing of implementation between the two programs to determine which is more efficient in the specific context.

The execution times of both programs were collected and analyzed, using a data set that includes 22 measurements for each program. Note how in one measure, the second algorithm failed.

Through the use of statistical techniques, such as the t-test and the confidence interval, it is possible to assess the importance of differences in execution times and to draw conclusions based on the results obtained.

Hypothesis

The null hypothesis (H0) states that there are no significant differences in the average execution time between the two programs used for the movement of the robot. In other words, the average execution time of the first program is equal to the average execution time of the second program.

The alternative hypothesis (H1) suggests that there are significant differences in the average execution time between the two programs, namely that the average execution time of the first program is different from the average execution time of the second.

Methodology of statistical analysis:

The statistical analysis was carried out using the t-test to compare the averages of the execution times of the two programs used for the movement of the robot. The t-test was chosen as it is particularly suitable for comparing the averages of two groups when a relatively small sample of data is available, as in our case, where 22 observations were collected for each program. In addition, the t-test is based on the assumption of data normality, which is reasonable to assume for the time distributions of program execution, considering that the performance of programs may approximate a normal distribution.

First of all, the t distribution was calculated.
Then, the p-value, which represents the probability of obtaining a result at least as extreme as the one observed, was calculated assuming that the null hypothesis (H0) is true.
By evaluating the p-value with respect to the level of significance chosen (in our case, 5%), it is possible to determine if there are statistically significant differences between the averages of the execution times of the two programs: if the p-value obtained is lower than the level of significance, then H0 is rejected in favor of H1, indicating a significant difference between the averages of the two programs.
However, if p is higher than the level of significance, H0 cannot be rejected, indicating a lack of significant difference between the two programs.

The degrees of freedom (DOF) represent the number of independent values that can vary within an analysis. In the t analysis, the degrees of freedom are calculated using the following formula:

DOF = (n1 + n2) - 2

where n1 and n2 represent the number of observations in the two groups.
In our case, n1 = n2 = 21 (21 experiments were performed for each robot).

Therefore, the degrees of freedom are calculated as follows:

DOF = (21 + 21) - 2 = 40

In addition, the 95% confidence interval has been calculated for the difference between the average execution times of the two programs. This interval provides an estimated interval within which the true difference in execution time averages is expected to fall, with a probability of 95%.

Results:

A series of experiments was run in for each program in which I made small changes to the position of the gold boxes in the arena by changing their angular offset parameter. I ran a total of 22 experiments, increasing the angular offset by 0.1 per experiment in the range [0*π, 2*π]. For each simulation, I recorded the time it took each robot to complete its assigned tasks in an Excel file.

I then analyzed the data and processed it by using MATLAB. By analyzing the collected times before making the t test, I removed the times related to the offset of 1.9*π because the second program had failed.

I also used MATLAB to run the t-test, which allowed me to evaluate the statistical significance of any observed differences between the two robots. In particular, the mean and standard deviation of all recorded data were calculated, and finally the t-test values were calculated, using the MATLAB functions.

Times for Anna's program	Times for Andrea's program	Angular offset
85.39	95.67	0*pi
88.73	81.96	0,1*pi
64.59	99.11	0,2*pi
69.12	106.44	0,3*pi
74.95	102.56	0,4*pi
76.79	59.62	0,5*pi
106.5	64.23	0,6*pi
80.63	87.87	0,7*pi
90.49	87.69	0,8*pi
91.35	99.99	0,9*pi
61.93	84.67	1*pi
109.45	114.96	1,1*pi
101.42	99.23	1,2*pi
105.87	100.67	1,3*pi
107.89	123.45	1,4*pi
102.28	89.74	1,5*pi
111.35	72.67	1,6*pi
104.55	67.65	1,7*pi
95.67	70.33	1,8*pi
81.96	360.3	1,9*pi
99.11	102.44	1,95*pi
106.44	51.01	2*pi

Timetable for the two programs with the respective offsets. The time for Robot 2 with offset 1.9*pi is red to indicate a failure.

The results of the T-Test calculated on MATLAB are the following:

- T Value: 0.62026
- P-value: 0.54209
- Confidence interval: [-8.1627, 15.0712]

The results of the statistical analysis of the execution times of the two programs indicate that there are no significant differences between the average execution times. The calculated value t, which measures the difference between the averages of the execution times of the two programs, is 0.62026. A value t close to zero suggests that the averages of the execution times of the two programs are very similar.

The p-value, which represents the probability of obtaining a result at least as extreme as the one observed, considering that the null hypothesis is true, is 0.54209. This value indicates that there is no statistical evidence to reject the null hypothesis, which supports the equality of the averages of the times of execution of the two programs. In other words, there is not enough evidence to say that one average is significantly different from the other.

The 95% confidence interval for the difference between the average execution times is [-8.1627, 15.0712]. This interval represents an estimated interval within which the true difference between the averages of execution time is expected to fall with a probability of 95%. Since the interval contains zero, there is insufficient evidence to claim that there is a significant difference between the averages of the execution times of the two programs.

Conclusions:

Based on the statistical analysis carried out, we can conclude that there are no significant differences in terms of execution time between the two programs considered. Both programs have a similar performance in terms of efficiency, regardless of changing the angle offsets for golden obstacles.

It is important to consider some limitations of this analysis. First, the sample of data used may not be representative of the entire population of program execution times. In addition, the analysis is based on the assumption of normality of data and independence of observations, which may not be fully satisfied. Further studies with a larger sample and a more rigorous experimental design may be necessary to confirm these results.