RT2: assignment 3

Statistical Analysis using t-Student Test: Comparison of Task Completion Time between Two Robots

Andrea Bolla – 4482930

Abstract:

This report presents the results of a statistical analysis conducted to compare the task completion time between Robot 1, where the algorithm was implemented by one my colleague, and Robot 2, working with my algorithm. The analysis aimed to determine if there are statistically significant differences in task efficiency between the two robots. The study involved multiple experiments, where both robots successfully completed the task, except in one instance where Robot 2 failed. The analysis included calculating the t-value, p-value, and confidence interval to assess the significance of the observed differences in task completion time.

Introduction:

The objective of this study was to investigate and compare the task completion time between Robot 1 and Robot 2. The analysis sought to evaluate the relative performance and efficiency of the two robots in completing the given task, changing at each test only the gold boxes position.
In that assignment, we were required, given a settled arena, to look a silver box within the environment and after that put this silver box near the golden box. Within the conclusion we ought to have gotten silver and gold boxes disseminated in sets. My test is based on the time required to accomplish that assignment.

Hypotheses:

In a t-test, two hypotheses are formulated: the null hypothesis (H0) and the alternative hypothesis (H1). These hypotheses allow for the comparison of sample data to determine if there is a significant difference between the means of the two groups under study. Here are the definitions of the two hypotheses:

- Null Hypothesis (H0): This hypothesis states that there is no significant difference between the means of the two groups. In other words, any observed differences between the groups can be attributed to chance or sampling error. H0 is typically formulated as "the means of the two groups are equal" or "the difference between the means is zero."

- Alternative Hypothesis (H1): This hypothesis states that there is a significant difference between the means of the two groups under study. It indicates that the observed differences cannot be attributed to chance or sampling error but are present. H1 can be formulated as "the means of the two groups are different" or "the difference between the means is not zero."

During the analysis, a sample is collected from each of the two groups, and the t-value is calculated based on the sample data. Subsequently, the p-value is calculated, which represents the probability of obtaining a t-value as extreme as or more extreme than the observed value, assuming that H0 is true. If the p-value is less than a predetermined significance threshold (e.g., 0.05), H0 is rejected in favor of H1, indicating a significant difference between the means of the two groups. Otherwise, if the p-value is greater than the significance threshold, H0 cannot be rejected, indicating insufficient evidence to claim a significant difference.

The level of significance is an important concept. It is always a percentage (usually 5%) which should be chosen with great care. In case we take the significance level at 5 per cent, then this implies that H_0 will be rejected when the sampling result (i.e., observed evidence) has less than 5% probability of occurring if H_0 is true. In other words, the 5 per cent level of significance means that researcher is taking a 5 per cent risk of rejecting the null hypothesis when H_0 happens to be true. Thus, the significance level is the maximum value of the probability of rejecting H_0 when it is true and is usually determined in advance before testing the hypothesis.

For this statistical analysis, I formulated a null hypothesis H0 that states there is no significant difference between the means of the two groups, and an significance threshold of 5%.

The degrees of freedom can be calculated using the formula:

DOF = (n1 + n2) - 2

where n1 and n2 are the sample sizes of the two populations. In my case, with 21 samples in each population, the degrees of freedom would be:

DOF = (21 + 21) - 2 = 40

To find the critical value at a 5% significance level with 40 degrees of freedom, I used a t-distribution table. Based on it, the critical value at a 5% significance level and 40 degrees of freedom is approximately ±2.0211. This means that if the calculated t-value falls outside the range of -2.021 to 2.021, I can reject the null hypothesis in favor of the alternative hypothesis.

Procedure:

To verify the hypotheses, I made some experiments by slightly changing the gold boxes position in the arena, by modifying the angular offset parameter. I made 22 experiments increasing the angular offset by 0.1 at each experiment in the range [0*pi, 2*pi] (e.g., 0, 0.1*pi, 0.2*pi …). To have comparable data, 22 simulations were carried out for both robots, each one with a different scenario given from a different angular offset. The time the robot needs to accomplish the tasks is calculated for each simulation and for both robots.

All the data were imported into an Excel file, then loaded on MATLAB to manage and plot the results and to calculate the t-test.

Robot 1: Anna	Robot 2: Andrea	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

Fig.1 Timetable for Robot 1 and 2 at each offset. The time for Robot 2 with offset 1.9*pi in red if a failure

Firstly, I remove the row with offset 1.9*pi due to a failure of Robot 2, then, I computed the mean value and the standard deviation of all the data recorded using the MATLAB's function and then I plotted the results (Fig. 1 & 2), and I benchmarked the two simulation (Fig. 3). In the end, I calculated the t-student test.

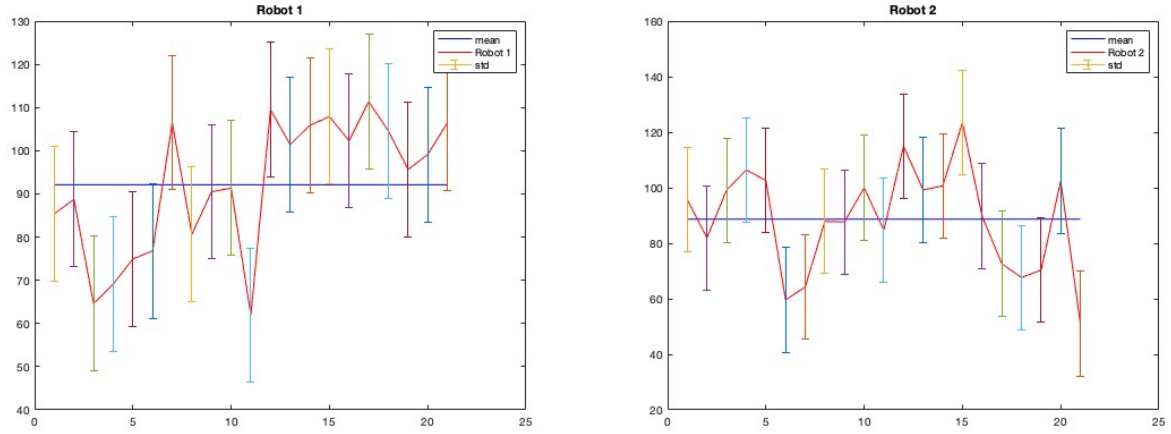

Fig. 2 Plot Robot 1 and Robot 2

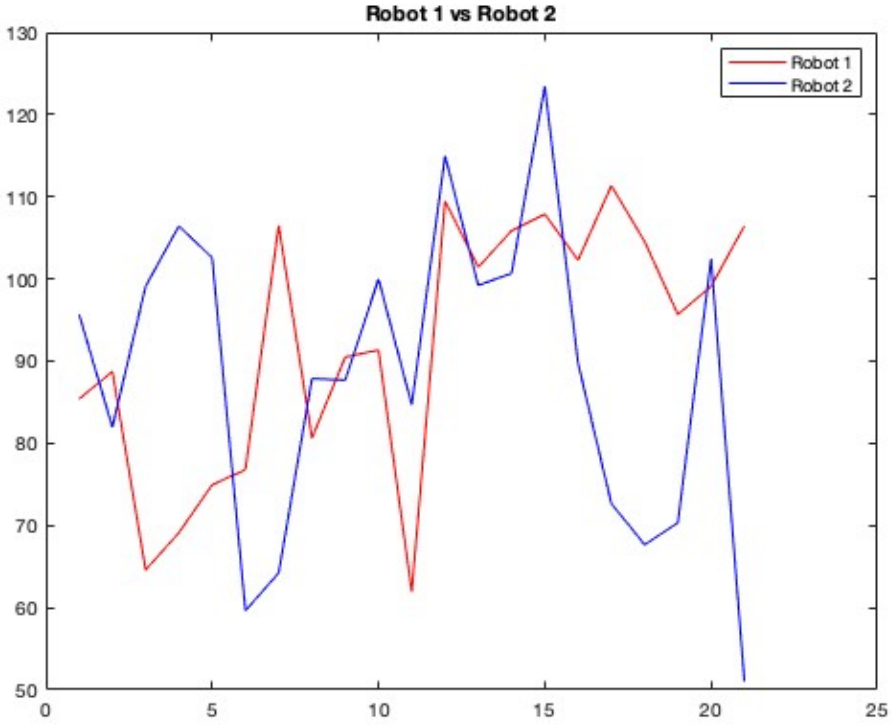

Fig. 3: Benchmark for the two Robots

Results:

The results of the analysis are outlined below:

The *t-value* of 0.62026 indicates a moderate difference in task completion time between Robot 1 and Robot 2.

However, the corresponding *p-value* of 0.54209 suggests that this difference is not statistically significant. The high p-value indicates that the observed variation in task completion time could plausibly occur by chance, and there is no conclusive evidence to support a consistent and significant difference between the two robots.

The *confidence interval* of [-8.1627, 15.0712] further supports these findings. As the interval includes zero, it implies that the true mean difference in task completion time may be zero, indicating no significant disparity in performance between Robot 1 and Robot 2.

Conclusion:

Based on the statistical analysis, there is *insufficient evidence* to conclude that there is a significant difference in task completion time between Robot 1 and Robot 2. Although Robot 2 experienced a failure in one instance, the overall analysis suggests

that this incident does not contribute to a consistent and statistically significant performance difference between the two robots.

It is important to acknowledge the limitations of this study. The sample size is relatively small, and the analysis is based on a single task. Future research with a larger sample size and alternative tasks would provide a more comprehensive understanding of the relative effectiveness of Robot 1 and Robot 2.