Università degli studi di Genova

DIBRIS

Department of Computer Science and Technology,
Bioengineering, Robotics and System Engineering

Research Track 2

Third Assigment

Statistical Analysis on the First Assignment

Author:

Colomba Ilaria

Student ID:

s4829201

Professor:

Carmine Tommaso Recchiuto

May 30, 2024

Contents

1 Introduction **2**

2 Testing of Hypothesis **2**

3 Set up of the experiment **2**

4 Analysis of the data and computations **3**

4.1 Data visualization . . . 3

4.2 Data computation . . . 4

5 Conclusions **5**

1 Introduction

In the third assignment of the course Research Track 2 it is asked to conduct a statistical analysis of the first assignment from the course Research Track 1. The aim of the statistical analysis it to compare the implementation of the code with the ones done by a colleague and to determine which implementation works better with different dispositions of tokens the arena. The randomness and variability of the experiments is set modifying the file "$two\_colours\_assignment\_arena.py$".

The aim of this first assignment is to make a simulator for the robotics. The simulator is able to simulate a robot in a bidimensional environment, to make it move around the arena and interact with golden token which are the objects collocated around the arena. The robot is able to look around the space where it is moving and to make decisions in order to grab any token initially collocated around the environment and to make any token grabbed and released in a specific part of the arena.
In order to elaborate the data, some different tools are used like $MATLAB$.

2 Testing of Hypothesis

Testing of Hypothesis is a statistical method used to make inferences or conclusions about a population parameter based on sample data. It involves formulating two competing hypotheses and using sample data to determine which hypothesis is supported by the evidence.
The two hypothesis are the following this analysis:

- the null hypothesis (H0): it assumes that there are not significant differences between the two implementations.
- the alternative hypothesis (H1): it might demonstrate that the two algorithms works differently and so that one implementation can be better than the other one

Starting the analysis, the assumption is that the null hypothesis is true and that the alternative hypothesis is false. In order to verify that, data is collected from a sample and the distances between the sample estimate and the hypothesised population parameter are measured. To do that the two-sample T-Test is conducted. It is ha statistical method which enables to compare the means between two dataset to determine if they are significantly different or not.

When this statistical test is used, it is important to compute a value called "p-value" that is the probability if getting a test statistic which is as extreme or more extreme that one observed, assuming the null hypothesis true. Depending on the p-value the null hypothesis is rejected or accepted: if the p-value is smaller that a significant level chosen before (for example α = 0.05) the alternative hypothesis is accepted and the null ones is rejected, if the p-value is bigger than α, the null hypothesis is accepted and the alternative ones is rejected.

3 Set up of the experiment

The primary objective of the experiment was to conduct a statistical analysis focused on measuring the average time required ti complete the task by the robot. By adjusting the radius of the golden tokens position in the arena, it is possible to introduce a randomness and variability of the experiment. The experiment is simulated 30 times each setup, accordingly to minimum number of experiment which are necessary to conduct a reliable statistical analysis using parametric techniques, such as calculating means and standard deviations. This criterion is often associated with the central limit theorem, which states that under certain conditions, the distribution of means from a large number of samples drawn from a population will follow a normal distribution, regardless of the population's distribution itself.

Running a sufficient number of simulations allows us to evaluate the impact of random fluctuations on the overall outcomes. The data obtained from these simulations forms a basis for statistical analysis. The average time for task completion is computed for each simulation set. Moreover, the variability in results is examined to assess the statistical significance of any observed differences. By performing multiple simulations with different random seed parameters, the statistical analysis captures overall performance trends and accounts for the effects of randomness. This methodology enhances the reliability of the analysis, enabling meaningful conclusions about the robot's efficiency in the arena.

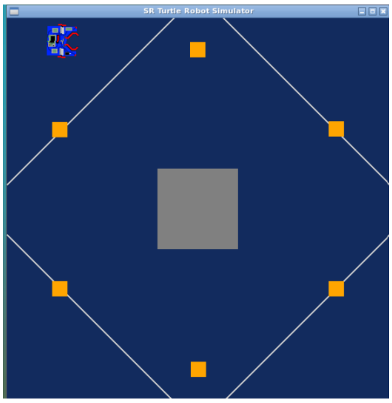

Figure 1: Representation of the arena environment.

4 Analysis of the data and computations

4.1 Data visualization

Experiment	Execution time Robot 1	Execution time Robot 2
1	124.25	98.17
2	126.76	100.14
3	176.43	142.67
4	136.41	111.67
4	130.87	108.67
5	168.38	141.79
6	141.81	108.53
7	128.83	112.19
8	148.27	121.18
9	157.81	114.67
10	166.30	117.19
11	124.26	104.18
12	138.28	112.15
13	131.22	109.26
14	120.79	102.67
15	135.31	107.19
16	148.28	109.87
17	121.78	104.13
18	122.75	102.68
19	117.25	97.67
20	129.28	104.68
21	158.77	131.17
22	134.72	111.96
23	141.30	129.16
24	127.26	107.18
25	116.26	96.17
26	157.76	129.74
27	132.31	113.25
28	160.74	131.69
29	136.85	115.45
30	126.78	100.60

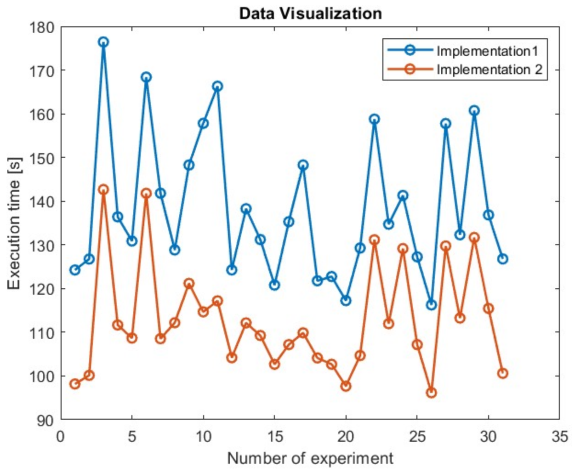

Figure 2: Graphic representation of the execution time of the two implementation.

4.2 Data computation

To do the computation the following algorithm is followed:

1. **Compute Means**:
   - Calculate the means of each group:

$$\text{mean\_group1} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i = \frac{1}{30}\sum_{i=1}^{30} x_i = 138.32$$

$$\text{mean\_group2} = \frac{1}{n_2}\sum_{i=1}^{n_2} x_i = \frac{1}{30}\sum_{i=1}^{30} x_i = 112.82$$

2. **Compute Standard Deviations**:
   - Find the standard deviations of each group to provide an indication of the proximity of the entire dataset to the mean value. The formula used to do that is the following:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

   The results obtained are the following:

$$\text{std\_group1} = 16.24; \text{std\_group2} = 12.45$$

3. **Decide the kind of test to apply**:
   - It is now feasible to perform a statistical analysis using the two-sample T-test to determine the statistical significance of the observed differences. Conducting this analysis is essential for reaching more accurate conclusions about the quality of the two implementations.

4. **Compute Pooled Standard Deviation**:
   - Combine variances of both groups using the formula:

$$\text{pooled\_std} = \sqrt{\frac{(n_1 - 1) \times \text{std\_group1}^2 + (n_2 - 1) \times \text{std\_group2}^2}{n_1 + n_2 - 2}} = 14.4742$$

5. **Compute T-value**:

- Calculate the t-value for a two-sample t-test using the formula:

$$\text{t\_value} = \frac{\text{mean\_group1} - \text{mean\_group2}}{\text{pooled\_std} \times \sqrt{\frac{1}{n1} + \frac{1}{n2}}} = 6.9347$$

6. **Compute the p-value**:
   - An important value in the T-test is the p-value. The p-value helps you assess the strength of evidence against the null hypothesis and make decisions about the hypotheses you're testing.

5 Conclusions

Looking at the data and at the obtained results, it is possible to conclude that there are 29 degrees of freedom (df = 30-1).

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Figure 3: T-table.

Considering the $alpha$ = 0.05 so for a significant level of 5% and with 29 degrees of freedom, comparing the results obtained with the t-table represented in the previous figure, it is possible to conclude that the t-value computed is bigger than the t-vale of the table so: 6.9347 $>$ 2.045. Because of that it is possible to conclude that the null hypothesis $H0$ is rejected and the alternative hypothesis is accepted. That fact suggests that there are important differences between the two implementations: one of the two algorithms works better than the other.
In order to say what of the two implementations works better, it is possible to apply on the data also the one-tail T-test where if the alternative hypothesis is accepted, it allows to understand what of the two algorithms is more performing than the other.