Statistical Analysis of the First Assignment in RT1

Miriam Anna Ruggero

June 1, 2023

1 Introduction

This report presents a statistical analysis of the execution of a program with two different implementations. The primary objective is to compare the performance of these implementations using the T-Test statistical method.
The program under consideration is a simple simulation involving the placement of golden and silver tokens. These tokens are arranged on two concentric circles, with the silver tokens positioned on the inner circle and the golden tokens on the outer circle. The silver tokens are equally spaced along the circumference of the inner circle, and the same applies to the golden tokens on the outer circle. In addition, the simulation involves a robot that moves and is tasked with picking up the silver tokens from the inner circle and placing them near the corresponding golden tokens on the outer circle, effectively pairing them together. The execution time of the program's different implementations will be analyzed to determine which one exhibits better performance in terms of efficiency and speed.

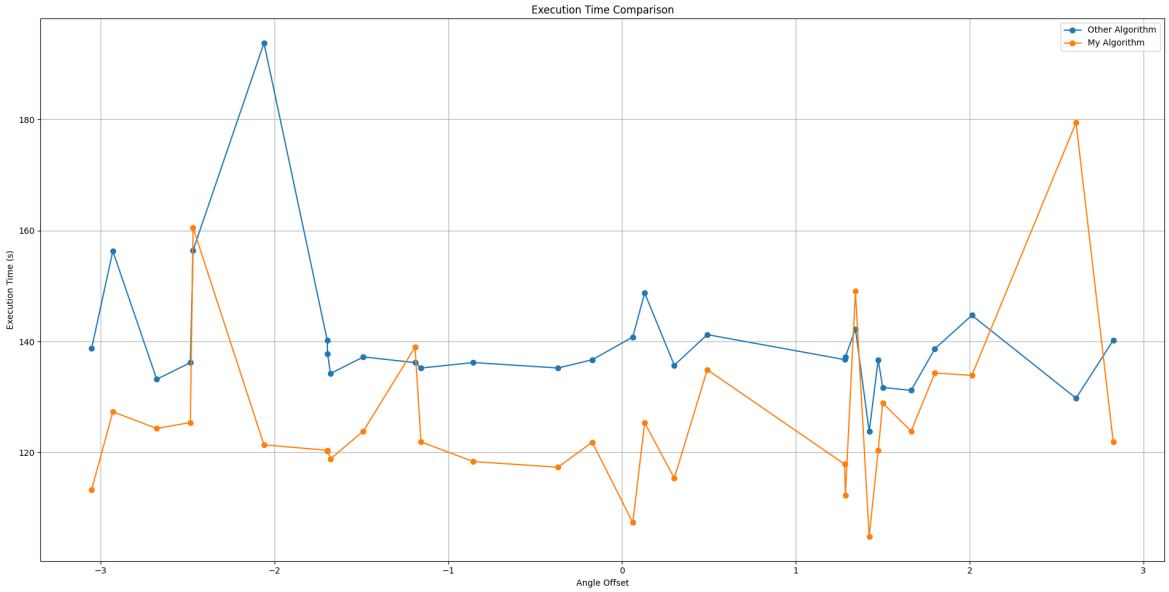

Figure 1: Time comparison

2 Data and Analysis

2.1 Data collection and description

The analysis was conducted on the execution speeds of an algorithm in 30 different scenarios. In order to modify the arrangement of golden tokens, 30 random values were collected ranging from -pi to pi. These values were used to replace the "angle-offset" parameter, thus affecting the relative positioning of the tokens. Subsequently, the algorithm was executed 30 times, and the total execution times of the program were recorded.Then, the same experiment was repeated with another algorithm, and the same data was collected. The data is described in the following table:

	angle offset	my implementation	other implementation
1	-2,060424597	121,3933911	193,788177967
2	1,285208181	112,321095	137,229959011
3	2,012904269	133,918118	144,727675915
4	-2,677467967	124,3382721	133,202096939
5	1,800200872	134,3297551	138,712718010
6	1,280954395	117,894413	136,746514082
7	-1,191339204	138,9849582	136,207404137
8	-1,696353823	120,406168	140,225404978
9	-2,484088381	125,4198809	136,207803011
10	1,474160518	120,3659189	136,723016977
11	-2,93018648	127,3605928	156,299800873
12	-1,156747467	121,911978	135,228929996
13	-0,171594251	121,8825049	136,721837044
14	0,130268327	125,3756559	148,747511148
15	-0,36826719	117,3565621	135,236062050
16	-0,856482791	118,3809879	136,221502066
17	-1,4887443	123,85625	137,229980946
18	1,663733295	123,853457	131,202381849
19	0,490334709	134,9258821	141,253770113
20	-3,052077598	113,340066	138,805794954
21	-1,694369179	120,3308229	137,817510128
22	0,061899363	107,3766999	140,823760033
23	1,422226214	104,902539	123,773239136
24	2,612489442	179,451473	129,802911997
25	0,300823182	115,380271	135,733092070
26	2,826926283	121,890105	140,250714064
27	-2,468290453	160,520401	156,445524931
28	1,342997203	149,050493	142,277531862
29	-1,677858795	118,863039	134,241894960
30	1,5	128,913923	131,734001160

Figure 2: Collected data

Here is a preliminary analysis of the provided data:

- Mean of the "My Implementation" column: 126.143
- Mean of the "Other Implementation" column: 140.398
- Standard deviation of the "My Implementation" column: 14.889
- Standard deviation of the "Other Implementation" column: 11.957
- Minimum value of the "My Implementation" column: 104.903
- Minimum value of the "Other Implementation" column: 123.773
- Maximum value of the "My Implementation" column: 179.451
- Maximum value of the "Other Implementation" column: 193.788

Significant differences are observed between the two implementations in terms of execution times. The mean of the "My Implementation" column is lower compared to the mean of the "Other Implementation" column, indicating better overall performance of the algorithm. The execution times for both implementations appear to follow different distributions. The "My Implementation" column has a lower mean and a higher standard deviation compared to the "Other Implementation" column. This suggests that the execution times for the "My Implementation" vary more widely, potentially indicating different performance characteristics or factors influencing the results.

2.2 Statistical Test

To compare and analyze this type of obtained results, the T-test is the most appropriate statistical test. The T-test is specifically designed to assess the significance of differences between two groups or conditions. It allows for a rigorous evaluation of the means or averages obtained from the data and helps determine if there is a significant distinction between the groups being compared. By utilizing the T-test, we can confidently evaluate the statistical significance of the observed differences and draw reliable conclusions from the analysis. The two-sample t-test assumes that the data in each group follow a normal distribution and that the variances of the two groups are approximately equal. The test calculates a t-value, which represents the difference between the means of the two groups relative to the variability within each group. The t-value is obtained by dividing the difference between the means of the two groups by the standard error of the difference. The standard error is a measure of the variability within each group and takes into account the sample sizes and variances. The t-value is then compared to a critical value from the t-distribution based on the degrees of freedom and the chosen significance level. If the calculated t-value exceeds the critical value, it suggests that there is a significant difference between the means of the two groups.

2.3 Execution of T-test

The null hypothesis (H_0) represents the assumption that there is no significant difference or no significant effect between the variables or groups under study. It assumes that there are no real relationships or effects in the underlying population.On the other hand, the alternative hypothesis (H_1) represents the hypothesis that you want to support or demonstrate. In contrast to the null hypothesis, the alternative hypothesis states that there is a significant difference or effect between the variables or groups being examined. It describes the hypothesis that is accepted if the data provide sufficient evidence to reject the null hypothesis.

Hypothesis:

- H_0: $\mu_1 = \mu_2$ **null hypothesis**: there's no significant difference between time execution
- H_1: $\mu_1 \neq \mu_2$ **alternative hyphothesis**: there's a significant difference between time execution

Data:

- $x_1 = 126.143$ mean value of 'my implementation'
- $x_2 = 140.398$ mean value of 'other implementation'
- $s_1 = 14.889$ standard deviation of 'my implementation'
- $s_2 = 11.957$ standard deviation of 'other implementation'
- $N_1 = 30$ number of samples of 'my implementation'
- $N_2 = 30$ number of samples of 'other implementation'
- dof = N_1 + N_2 - 2 = 58 degrees of freedom of the 2 samples

Method:

The first step is to compute the Pooled variance.The Pooled variance is the weighted average of the variances, using as weights N-1:

$$\sigma^2_{pooled} = \frac{(N_1-1)s_1^2+(N_2-1)s_2^2}{dof} = 182.348$$

Which leads to the pooled, estimated SE of the sampling distribution of the difference of means:

$$\sigma_{x_1-x_2} = \sqrt{\sigma^2_{pooled}(\frac{1}{N_1} + \frac{1}{N_2})} = 3.486$$

Now it's possible to compute the t-value:

$$t_{x_1-x_2} = \frac{x_1-x_2}{\sigma_{x_1-x_2}} = \text{-}4.008$$

Finally we can compare the t-value obtained with the critical value retrieved from the t-distribution table(Figure 3),based on degrees of freedom and significance level α= 0.05. Based on these numerical findings, we can confidently reject the null hypothesis, H_0, as $|t_{x_1-x_2}| < t_{crit} = 2.0017$, indicating a significant deviation from the expected value. Moreover, the calculated value of $t_{x_1-x_2}$ is notably lower than the standard value $t_{crit} = 2.0017$ obtained from the t-two-tailed distribution with 58 degrees of freedom and significance level α= 0.05. This provides ample evidence to support the acceptance of the alternative hypothesis H_1.

3 Conclusion

Based on a significance level of 5% α= 0.05 , the t-test results provide significant statistical evidence to support the alternative hypothesis that there is a significant difference in time between the two algorithms.

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%

Figure 3: T-table