Università di Genova

Final Report Research Track 2

Author: Federico Malatesta

Professor: Carmine Recchiuto

Course: Research Track 2

Abstract:

This document presents the final report for the Research Track 2 course.
The report contains a statistical comparison between two different implementations of an algorithm developed to move the robot and make it group the tokens present in the environment around a clustering point.
The goal is to demonstrate a difference in the efficiency of the algorithms if any.

Introduction

The aim of the project is to perform a statistical analysis on the first assignment, considering two different implementations (own, and a solution of one of a colleague) and testing which one performs "better", when tokens are randomly placed in the environment.
As performance evaluators it's possible to consider:

- the average time required to finish the task
- the number of success / failures

Hypothesis

The statistical test chosen is the *Paired T-Test* because it allows the comparison between two samples with paired observations. This comes from the fact that the measurement of the time taken to collect the tokens was done using the same environment for both algorithms.
The null hypothesis states that $\mu_d = 0$ denoting with μ_d the average difference between the measurements in the two correlated samples, and this implies that there is no mean difference between the two samples. However, it is expected to be rejected if the average difference observed between the two samples is significantly different from zero.

Experimental setup

The first change made to the colleague's code in order to make it comparable was to set the turning velocity and the linear velocity equal in both codes.
Then as written above the chosen algorithm is the *Paired T-Test* and the number of recorded events is 30.
The operating algorithm of the used statistic test is as follows:

- Calculate the difference ($d_i = y_i - x_i$) between the two observations on each pair, making sure you distinguish between positive and negative differences.
- Calculate the mean difference $\bar{d}$.
- Calculate the standard deviation of the differences, s_d, and use this to calculate the standard error of the mean difference $SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$.
- Calculate the T-statistic, which is given by $T = \frac{\bar{d}}{SE(\bar{d})}$. Under the null hypothesis, this statistic follows a t-distribution with $n - 1$ degrees of freedom.
- Use tables of the t-distribution to compare your value for T to the t_{n-1} distribution. This will give the p-value for the paired t-test.

Regarding the code, to obtain the desired effect, the *token.location* line was modified using the following methodology:

$$angle = \text{random.uniform}(0, 2\pi)$$

This code allows to generate a random angle between 0 and 2π and in order to be sure to use the same configuration of tokens for both the used algorithm it's also necessary to instantiate a seed with: *random.seed* equal for each program, that changes each new iteration.

Results

The result of the 30 measurements carried out is contained in the following table:

Table 1: Execution times

Malatesta	Odetti	Difference
225.525	231.926	6.401
224.011	230.080	6.069
227.048	229.832	2.784
221.968	231.324	9.356
221.476	225.557	4.081
221.962	237.601	15.639
223.010	299.758	76.748
216.976	230.159	13.183
218.470	231.071	12.601
217.989	304.834	86.845
227.542	227.585	0.043
226.486	235.718	9.232
221.982	233.280	11.298
218.964	233.326	14.362
223.995	229.665	5.670
226.017	228.514	2.497
226.504	236.433	9.929
225.529	223.984	1.545
221.011	231.982	10.971
224.994	231.951	6.957
220.026	320.023	99.997
217.986	232.710	14.724
224.029	230.190	6.161
222.523	227.454	4.931
222.466	234.972	12.506
220.463	228.556	8.093
219.964	234.790	14.826
223.984	229.889	5.905
227.055	231.962	4.907
220.511	233.417	12.906

Resulting in $\bar{d} = 16.372$ and $s_d = 24.379$.
Now it's possible to compute $SE(\bar{d})$ using the formula written previously and it's $SE(\bar{d}) = 4.45$.
The next step is to compute $T = \frac{\bar{d}}{SE(\bar{d})} = 3.68$.

Discussion of the results

The correspondent *P-value* that can be extracted from tables is 0.001 and so comparing it with $\alpha = 0.05$ it's possible to conclude that H_0 is rejected and H_1 is accepted.
Therefore my algorithm under the specified condition is better then my colleague's one with a confidence of 99.95%

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										