UNIVERSITÀ DEGLI STUDI DI GENOVA

DIBRIS

DEPARTMENT OF COMPUTER SCIENCE AND TECHNOLOGY,
BIOENGINEERING, ROBOTICS AND SYSTEM ENGINEERING

RESEARCH TRACK II

Third Assignment

Fundaments of Statistics

Author:

Moriconi Michele

Student ID:

s4861803

Professor:

Carmine Recchiuto

June 1, 2023

1 Assignment description

Perform a statistical analysis on the first assignment, considering two different implementations (yours, and a solution of one of your colleagues) and testing which one performs better, when silver and golden tokens are randomly placed in the environment.

As performance evaluators you may possibly consider:

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

Possibly, you can also vary the number of boxes to see how the algorithm behaves. You can add obstacles by modifying the file: robot-sim/sr/robot/arenas/two_colours_assignment_arena.py

Write a report composed of:

- Hypotheses made (null hypothesis and alternative hypothesis)
- Description and motivation of the experimental setup (types of experiments, number of repetitions)
- Results
- Discussion of the results with statistical analysis
- Conclusion (is the hypothesis proven?)

2 Introduction

The assignment was carried out in collaboration with my colleague Giovanni Rocca. We compared the algorithms we implemented in the first assignment of the course Research Track I to perform a statistical analysis of the speed of completion in a random environment.

3 Hypotheses

I compared my colleague's algorithm with mine wanting to confirm the superiority of the latter. Usually, the alternative hypothesis is the one that one wishes to prove, while the null hypothesis is what one wishes to reject. Considering that I took the assumption in which the two algorithms are equal as the null hypothesis and the premise that my algorithm is superior as the alternative hypothesis.

4 Experimental setup

As mentioned before in the Introduction section the goal of the experiment is to perform a statistical analysis of the algorithms' speed in a random environment. In the environment in which the simulation takes place are present six golden and six silver markers as shown in Figure 1. I wanted to make the environment random, but also to be able to replicate the experiment for both algorithms. To achieve this I used a function that generates a random number from a seed and then I used that number to set the angular offset of the golden markers in the circle.

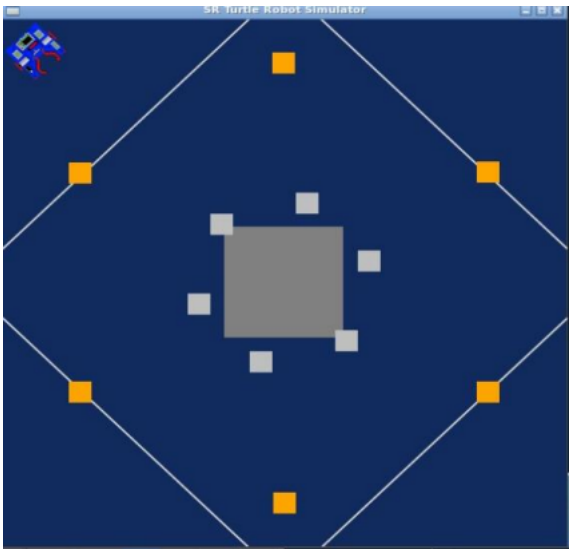

Figure 1: Environment

The statistical test I chose is the Paired T-Test which can be used to compare two different approaches applied to the same scenario. I measured the time elapsed from the start of the simulation to the completion of the task 30 times for each algorithm.

5 Results

The measurements of the experiment are shown in the following table.

Number	Moriconi	Rocca	Seed	Difference
1	113.2568562	99.31476092	10	13.94209528
2	84.685815	92.33062196	11	-7.6448071
3	94.732783	94.38754201	12	0.34524107
4	92.821157	151.74904418	13	-58.9278872
5	84.209522	91.35502696	14	-7.145504952
6	82.673911	98.55895519	15	-15.88504434
7	82.176113	104.0243602	16	-21.84824729
8	85.170658	155.7155468	17	-70.54488873
9	88.190233	85.33834004	18	2.851892948
10	90.208793	83.81672406	19	6.392068863
11	105.762834	137.593051	20	-31.83021712
12	74.680162	90.78722215	21	-16.10706019
13	78.673735	95.52155805	22	-16.8478229
14	103.792728	143.052345	23	-39.25961709
15	89.696134	100.8769982	24	-11.18086433
16	86.680724	105.44752	25	-18.76679587
17	90.673854	92.78944707	26	-2.115592957
18	94.730558	98.7431879	27	-4.012629747
19	82.203786	88.69668794	28	-6.492902041
20	75.666856	83.683676	29	-8.016819954
21	80.158795	80.27108097	30	-0.112285852
22	89.190372	101.2363181	31	-12.04594612
23	91.191099	91.69763303	32	-0.5065341
24	80.172738	84.69903183	33	-4.526293755
25	77.666526	82.18091893	34	-4.514392853
26	76.642797	80.68608212	35	-4.043285131
27	89.407257	91.69872308	36	-2.291466236
28	105.632253	93.2082839	37	12.42396903
29	92.8466649	95.2108779	38	-2.36421299
30	78.3020508	82.1743731	39	-3.87232232

In the first column is shown the repetition's number, in the second one the times of completion of my algorithm, in the third the times of completion of my colleague's algorithm, in the fourth the seed used to change the environment, and in the last one the difference between my times and my colleague's.

I also calculated the following values:

- The mean of the differences: -11,16493913;
- The standard deviation: 18,4548198;
- The standard error of the mean difference: 3,369373699;
- The absolute value of the t-statistic: 3,313654148.

6 Discussion of the results

I computed the standard error of the mean difference using the formula: $SE(\overline{d}) = \frac{S_d}{\sqrt{n}}$, where $\overline{d}$ is the mean of the differences, S_d is the standard deviation, and n is the number of experiments.

I calculated the t-statistic with the formula: $T = \frac{\overline{d}}{SE(\overline{d})}$.

Using the tables of the t-distribution, shown in Figure 2, I compared my value for T to the t_{n-1} distribution. This will give the p- value for the paired t-test.

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 2: T-Table

To extrapolate the p-value from the t-table I had to find the higher number that is less than T in the 29^{th} row. I used that row because it is the one corresponding to the number of experiments minus one. I used the value relative to the one-tailed distribution because I had a clear ”direction” in mind. By using a one-tailed test, I focused the statistical power of the test on detecting whether the mean difference is significantly greater (or less) than zero in the ”direction” specified by my alternative hypothesis.

7 Conclusion

In conclusion the p-value that I extracted from the t-table is 0.005. So I accepted the alternative hypothesis, that stated that my algorithm is better than my colleague's, with a confidence of 99.5% and I rejected the null hypothesis with the .5% error probability.

Contents