Assignment: statistical analysis

For Research Track II Course

Martina Germani

Universty of Genoa, Master's Degree in Robotics Engineering

September, 2022

Abstract

In this assignment I have performed a statistical analysis of the first assignment of Research Track I course, considering two different implementations (mine and the professor's solution) and testing if there are differences in the average distance from the obstacles, which are represented by golden boxes.
The statistical test I chose in order to test my hypotheses is a two-tailed t-test, through which I succeeded in rejecting the null hypotheses proving my alternative hypotheses with a level of significance of 5%.

1. Introduction

For this assignment I performed a statistical analysis of the first assignment of Research Track I course, where there is a robot which moves in a circuit in the counter-clockwise direction, avoiding to touch the golden boxes, but grabbing silver box when the robot is close to it and moving it behind itself.
The hypothesis that I wanted to test is that the average distance that the robot maintains from the obstacles is different in the case of two different algorithms: one implemented by me and the second one implemented by the professor.
For testing my hypothesis, I randomly collected 65 samples during the execution of each implementation, where these samples represent the distance between the robot and the golden boxes.
Since the two groups are different, I performed a two-tailed test, and in particular, the statistical test I chose to perform is the t-test, which allows to judge the significance of difference between the means of two samples.
By applying this statistical test I succeeded in rejecting the null hypothesis which considers equal the two average distances, proving that there are differences between the two algorithms.

2. Hypothesis

The first thing that it's necessary to do when you want to perform a statistical test is to choose the hypothesis that you have to test.
This choice is very important and the hypothesis should be formulated in a correct way in order to make it testable.
My goal is to prove that one implementation performs better than the other one in the given circuit and as performance evaluator I chose the distance from the obstacles, that, in this particular environment, are represented by golden boxes.
So, the formulation of the *alternative hypothesis* H_1 that I wanted to test is:

"***The average distance between the robot and the obstacles is different in the two implementations***".

Whereas the *null hypothesis* H_0 that I wanted to reject is:

"***The average distance between the robot and the obstacles is the same in the two implementations***".

3. Experiment

Given the testable hypothesis defined above, for testing it I decided to randomly collect data in one cycle of the robot around the environment.

In particular, the data collected represent the current distance between the robot and the golden boxes while the robot moves in the arena.
In order to have a normal distribution of the samples around the mean value, I randomly collected 65 samples for each group, and in the below figures, (Figure 1) and (Figure 2), I reported these values: group1 contains the value collected by my

```
Group1:  [1.03077641 1.01630528 1.02864927 1.0002867  1.0028642  0.9383
 0.95271614 0.87527333 1.10007654 1.06394999 0.56135869 0.63741024
 1.65442863 0.87623902 0.81665268 0.79113552 0.83944266 0.95043329
 0.53689342 0.7031073  0.623758   0.81949776 0.90277569 0.98816402
 0.53137181 0.60678808 0.77040704 0.54853243 0.66592495 0.7883248
 0.91074483 0.98444376 0.93901204 0.68766017 0.60742169 0.58248529
 0.62207485 0.86419309 0.75499538 1.06211524 0.90600874 0.67674854
 0.67674822 0.63101931 0.83207105 0.95031825 0.81678809 1.06377629
 1.07067925 0.57332692 0.57332692 0.62404095 0.68653356 0.71983926
 0.83035345 0.93981779 1.00717393 0.93933372 0.94372681 1.19934228
 0.7043706  0.71780783 0.90961878 0.72154682 0.72530548]
```

Figure 1: Samples collected by my implementation

```
Group2:  [2.42641462 2.5944931  2.46025033 0.83674479 1.6245664  2.18
 1.90927588 1.8415528  1.77404324 1.50260527 1.49273225 1.41911968
 1.50079165 0.49618727 1.40349742 0.4980495  0.98768823 0.98768823
 1.3887748  2.35772987 2.50189908 2.1651346  2.04251881 1.41648902
 0.82214372 2.7145153  1.64175937 1.08602355 0.85658796 1.38015245
 1.96314291 1.70681023 3.14626727 1.36551779 2.00907714 2.35572905
 2.17250368 0.49735553 0.49731108 1.35604241 1.26411263 1.39981495
 2.15055675 1.93115635 1.87833681 0.49998962 2.39687499 1.68141498
 1.82787901 2.047784   2.27433409 2.50535203 1.47265926 2.16077339
 1.90577916 1.69511216 0.49819657 1.46089714 2.32040501 1.61136555
 1.53247526 1.47265926 2.16077339 2.08912943 2.37625638]
```

Figure 2: Samples collected by professor's implementation

implementation, whereas group2 contains the values collected by professor's implementation.

Below in (Figure 3) and (Figure 4) I reported the samples distribution of each group, where, as described by the *Central Theorem Limit*, you can see that the sampling distribution tends quite closer to the normal distribution since the number of samples is enough large, and the two groups have an equal variance because they were independently extracted.

Moreover in these figures you can also see the computed mean for each group:

- $\mu_1 = 0.831949$
- $\mu_2 = 1.692334$

Therefore, already at first sight, you can observe that the means of the two groups are significantly different, but I proceeded anyway by carrying out a statistical test.

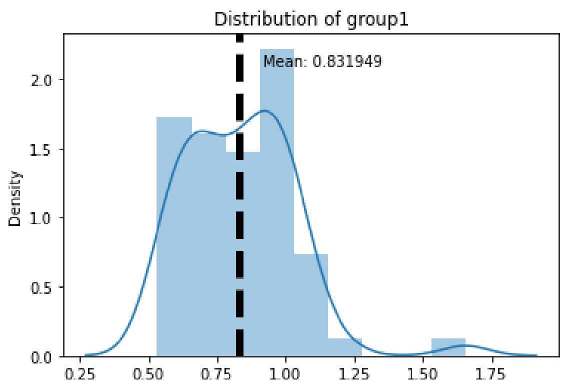

Figure 3: Samples distribution of my implementation

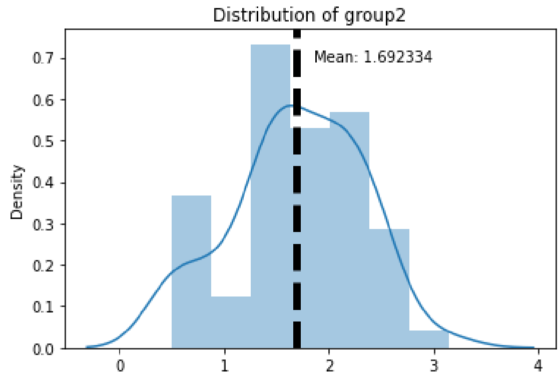

Figure 4: Samples distribution of professor's implementation

4. Statistical test

First of all, since I wanted to test if the two groups are different, I decided to perform a two-tailed test. The statistical test that I chose to perform is the **t-test**, which is used to compare the means of two groups.
The t-test is a parametric method that can be used when the samples satisfy the hypothesis of normal distribution, equal variance and independence.
In particular, there are two types of t-tests:

1. *independent t-test*, known as *Student's t-test*, which can be used when the two groups under comparison are independent of each other,
2. *paired t-test*, which can be used when the two groups under comparison are dependent on each other.

So, I performed an independent t-test with the following hypothesis:

- H_0: $\mu_1 = \mu_2$ null hypothesis,
- H_1: $\mu_1 \neq \mu_2$ alternative hypothesis.

and by choosing $\alpha = 0.05$.

Below I report the Python code used:

```
def independent_ttest(data1, data2, alpha):
    mean1, mean2 = mean(data1), mean(data2)
    se1, se2 = sem(data1), sem(data2)
    sed = sqrt(se1**2.0+se2**2.0)
    t_stat = (mean1-mean2)/sed
    df = len(data1)+len(data2)-2
    cv = t.ppf(1.0-alpha, df)
    p = (1.0 - t.cdf(abs(t_stat), df))*2.0
```

```
    return t_stat, df, cv, p

t_stat, df, cv, p = independent_ttest
    (my_implementation,
    ref_implementation, 0.05)

if abs(t_stat)<=cv:
    print("Accept the null hypothesis
    that the means are equal")
else:
    print("Reject the null hypothesis
    that the means are equal")


if p>0.05:
    print("Accept the null hypothesis
    that the means are equal")
else:
    print("Reject the null hypothesis
    that the means are equal")
```

Where, first of all it's necessary to compute the means and the standard errors of the mean of the two groups, followed by the standard error deviation. At this point, I computed the t statistic using the following formula:

$t = \frac{(mean1 - mean2)}{\sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}}}$

Then it's necessary to compute the inverse cumulative distribution function
Alternatively you can use the pre-existing function:

```
t_stat, p = ttest_ind(data1, data2)

if p> alpha:
    print("Fail to reject H0")
else:
    print("Reject H0")
```

which takes as input the arrays containing the samples and it returns as output the values of t and p.

5. Result

The result of the t-test is reported in the (Table 1).

Table 1: Table with results of the t-test.

$tstat$	$pvalue$
10.67	$2.01*10^{-19}$

So, if we consider the t-test table, attached in the references, for a number of degree of freedom $df = 128$, and with a level of significance of 5%, the $t_{table} = 1.984$, so, since $t = 10.67$, I can reject the null hypothesis H_0.

This is also proved by the functions reported above, which reject H_0 and that reported a p value equal to $p = 2.01*10^{-19}$ which represents the level of uncertainty with which we can assume that the two sets of data belong to two different distributions.
Moreover, the computed t value is very big, and this means that there is a big difference between the mean of the two groups: this was already evident as already announced in the section 3 of this report.
In conclusion, with this test I proved that the average distance between the robot and the obstacles is different in the two implementations, and so one implementation is better than the other one.