RT2 Statistics report

Paul PHAM DANG, Mamoru OTA

June 2, 2025

1 Abstract of the Experiment

This experiment compares two obstacle avoidance strategies based on the planning method Bug 0. For the first strategy, when an obstacle is encountered, the robot turns to the left, while in the second method, it turns to the right. We use a task where a mobile robot must reach a goal while navigating around obstacles. Performance is evaluated based on the distance traveled and the success rate, with the aim of statistically determining which strategy is more effective.

2 Hypotheses Made

In this experiment, it is assumed that there is no essential difference in performance between the two obstacle avoidance algorithms: left avoidance and right avoidance. The objective is to statistically examine this assumption. The hypotheses are formulated as follows:

Travel Distance

- Null hypothesis (H_0): There is no significant difference in the mean distance traveled between left and right avoidance.
$$\mu_{\text{left}} = \mu_{\text{right}}$$
- Alternative hypothesis (H_1): There is a significant difference in the mean distance traveled between left and right avoidance.
$$\mu_{\text{left}} \neq \mu_{\text{right}}$$

The paired t-test is used for this analysis.

Success Rate

- Null hypothesis (H_0): There is no significant difference in the success rate between left and right avoidance.
$$P_{\text{success,left}} = P_{\text{success,right}}$$
- Alternative hypothesis (H_1): There is a significant difference in the success rate between left and right avoidance.
$$P_{\text{success,left}} \neq P_{\text{success,right}}$$

The chi-square test is used for this analysis.

3 Description and Motivation of the Experimental Setup

In this experiment, we used a simulation environment built on ROS Noetic and Gazebo 11. Multiple fixed obstacles were placed in a square environment with coordinates from -10 to 10 along x and y-axis. The robot has to move to a specified goal point by avoiding the obstacles. We generated a list of random points with the library *random* of Python with a uniform distribution between $[-9; 9]$ for the 30 goal points. When the robot reached the current goal, the next target point was sent in sequence. If the robot got stuck due to an obstacle or other reason, it would retry from the starting point (0, 1), and if it still could not reach the goal, it was judged to have failed. In addition, if it took too long to reach the goal (more than 3 minutes), it was also considered to have failed.

We performed 30 trials for each method and designed the experiment so that we could compare the performance of each avoidance method by applying two different algorithms, left avoidance and right avoidance, to the same target point list. Performance was evaluated using the distance traveled to reach the goal and the goal-reaching success rate.

Distance traveled serves as a reliable metric for evaluating the robot's trajectory efficiency. Unlike time-based measurements, it is less affected by variations in processing speed or frame rate, which can differ across execution environments. Therefore, distance was chosen as the primary evaluation criterion. To calculate it, we computed the Euclidean distance between consecutive points received at each callback of the /odometry topic subscriber. The total distance traveled is then obtained by summing these individual distances. The collected distance traveled data showed a shape that was roughly close to a normal distribution using histograms, so a t-test was selected as the method to test whether there was a significant difference in the average distance of each method. In addition, by applying both left and right avoidance to each goal point, the two measured values to be compared were obtained under the same conditions and can be considered paired data. Therefore, instead of a regular t-test (comparison of two independent groups), we adopted a "paired t-test," which performs analysis based on the difference between observed values. In the paired t-test, the difference $d_i = x_i - y_i$ between each pair is calculated, and the t-value is calculated using the following formula:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \tag{1}$$

Here, $\bar{d}$ is the average of the differences, s_d is the standard deviation of the differences, and n is the number of pairs. This makes it possible to eliminate the influence of environmental factors such as the goal location and obstacle placement and more accurately detect effects due to differences in the avoidance algorithms themselves.

Furthermore, the t-value obtained by the t-test is used to calculate the p-value (probability of significance) by comparing it with a t-distribution with degrees of freedom df $= n - 1$ to determine whether the difference is statistically significant. This p-value represents the probability that a bias greater than the observed difference would occur by chance when the null hypothesis (there is no difference between the average distances of both methods) is correct. The standard for judgment is generally interpreted as follows:

- If $|t|$ is large (2 or more) and the corresponding p-value is less than 0.05, it is considered that there is a statistically significant difference, and the null hypothesis is rejected.
- If $|t|$ is small and the p-value is 0.05 or more, it is determined that "there is no difference," and the null hypothesis is not rejected.

On the other hand, the success rate was treated as categorical data (success/failure) indicating whether the robot reached the goal or not. The number of successes/failures for each method was cross-tabulated, and the chi-square test was applied to verify whether there was a statistical bias in their frequency distribution. In the chi-square test, the difference between the observed value O_{ij} and the expected value E_{ij} is squared, divided by the expected value, and summed for all cells to obtain the chi-square statistic using the following formula:

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{2}$$

The chi-square test is a standard method for evaluating whether the difference between the observed value and the expected value for two or more categories is due to chance and is suitable for comparing the frequency of success/failure, as in this case. In this way, by selecting the statistical test most appropriate for the data structure and distribution characteristics of each data type (distance traveled and success rate), we are able to increase the reliability and validity of the analysis results.

Furthermore, to determine the significance of the test results, it is necessary to consider the p-value and degrees of freedom in addition to the chi-squared statistic obtained. The degrees of freedom depend on the structure of the cross-tabulation table, such as df $=$ (rows $-$ 1) $\times$ (columns $-$ 1). Based on the obtained chi-squared value and the degrees of freedom, the p-value is calculated according to the chi-squared distribution. The standard for judgment is generally interpreted as follows:

- If the p-value is less than 0.05, it is considered that there is a statistically significant difference, and the null hypothesis is rejected.
- If the p-value is 0.05 or more, it is determined that "there is no difference," and the null hypothesis is not rejected.

4 Results

Travel Distance

Table 1 shows the average and standard deviation of travel distance for each avoidance method.

Table 1: Average and standard deviation of travel distance

Method	Average	Standard Deviation
Left	13.85	12.52
Right	12.84	9.18

To assess the assumption of normality required for the paired t-test, we examined the distribution of distance differences between the two methods. Figure 1 and Figure 2 provide visualizations of the difference distribution (left - right) and standardized Z-scores respectively.

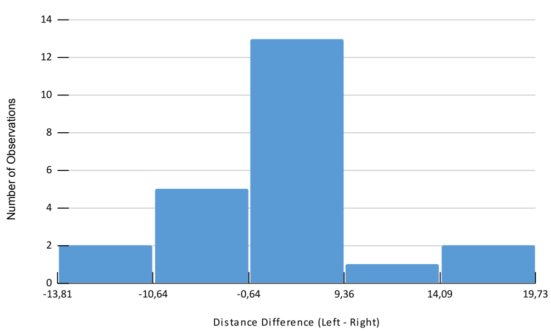

Figure 1: Histogram of distance differences

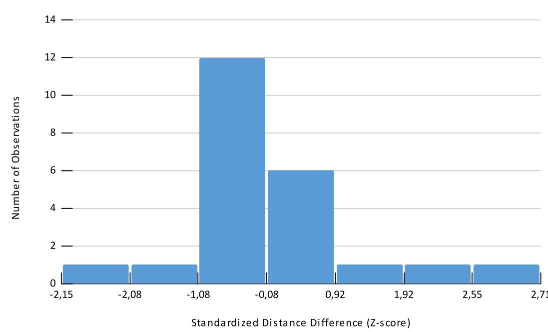

Figure 2: Histogram of standardized Z-scores

For the statistical comparison of travel distances, we conducted a paired t-test using only the data from trials where the robot started and reached the same target for both the left and right avoidance strategies, ensuring a fair comparison under identical conditions. The resulting t-statistic was $t = 0.70$, with $df = 22$ degrees of freedom, and the corresponding p-value was $p = 0.49$.

Success Rate

Table 2 shows the observed frequency of success and failure for each avoidance strategy. Table 3 shows the expected frequencies under the null hypothesis of no association between strategy and success rate.

Table 2: Observed frequency

Method	Success	Failure	Total
Left	28	2	30
Right	26	4	30
Total	54	6	60

Table 3: Expected frequency

Method	Success	Failure	Total
Left	27	3	30
Right	27	3	30
Total	54	6	60

We conducted a chi-square test based on the 2×2 contingency table above. The resulting chi-square statistic was $\chi^2 = 0.74$ with $df = 1$ degree of freedom, and the corresponding p-value was $p = 0.39$.

5 Discussion of the Results

Travel Distance

The visual inspection of the distance differences (Figure 1) and the standardized Z-scores (Figure 2) suggest that the distribution of differences is approximately normal. This observation supports the validity of using a paired t-test for comparing the travel distances between the two methods.

The result of the paired t-test yielded a t-statistic of $t = 0.70$ and a p-value of $p = 0.49$. The relatively small t-statistic indicates that the difference in average travel distances between the two methods is minor compared to the variability within the data. Since the p-value is considerably greater than the commonly used significance level of $\alpha = 0.05$, we fail to reject the null hypothesis. This implies that there is no statistically significant difference in the mean travel distance between the left and right avoidance strategies.

Success Rate

For the success rate, a chi-square test was performed using the observed and expected frequencies. The test resulted in a chi-square statistic of $\chi^2 = 0.74$ with 1 degree of freedom, and a p-value of $p = 0.39$.

Similar to the result for travel distance, the p-value is greater than the significance level of $\alpha = 0.05$, and therefore, we also fail to reject the null hypothesis in this case. This suggests that there is no statistically significant difference in the success rates between the two avoidance strategies.

Further Considerations

The two avoidance strategies may be fundamentally similar in terms of their underlying logic and performance characteristics, especially under the specific experimental conditions employed in this study. As a result, such variations may not manifest as significant performance differences, leading to overlapping distributions in the observed outcomes.

Future research could better evaluate each avoidance strategy by investigating a wider range of environmental scenarios and task complexities. Introducing more complex obstacle configurations, for example, could reveal performance differences that are currently undetectable. Furthermore, comparative analysis with a wider variety of avoidance algorithms could provide deeper insights into the design principles that contribute to effective robot navigation in real-world environments.

6 Conclusion

In this experiment, we statistically evaluated the performance of two obstacle avoidance strategies—left and right—based on two key metrics: travel distance and success rate. Using a paired t-test and a chi-square test, respectively, we found no statistically significant differences between the two methods for either metric. The t-test for travel distance yielded a p-value of 0.49, and the chi-square test for success rate resulted in a p-value of 0.39, both exceeding the conventional significance threshold of 0.05.

These findings suggest that, under the specific simulated conditions of the experiment, both avoidance strategies exhibit comparable effectiveness. The absence of significant differences may reflect inherent similarities in their logic or the limitations of the experimental setup in distinguishing finer performance variations.

However, it is important to emphasize that failing to reject the null hypothesis does not constitute statistical proof of equivalence. As such, our findings should be interpreted as tentative support for the possibility that there is no meaningful performance difference between the two strategies under the tested conditions. If we wanted to prove that there is no difference statistically, we need infinite data, which is impossible. So we should prove it with a mathematical proof instead.

Future research should consider more challenging scenarios, such as complex obstacle configurations, to more clearly elucidate the nuances in the effectiveness of strategies. Increasing the range of algorithms compared will also enable us to identify the characteristics that best contribute to robust and efficient navigation.

Appendix: Raw Data

The following tables summarize the raw results of the 30 trials conducted for each obstacle avoidance strategy. Each row represents a trial number, traveled distance and the final status. The status falls into one of four categories:

- REACHED: The robot successfully reached the new goal from the last goal position.
- REACHED after blocked: The robot is blocked with an obstacle when it started from the last goal position, but managed to reach the goal from the initial position (0,1).
- REACHED from start: The robot started from the initial position (0,1), and then reached the goal.
- FAILED: The robot failed to reach the goal due to timeout or repeated blocked.

Table 4: Trial Results for Left

No.	Distance	Status
1	13.5	REACHED from start
2	5.69	REACHED
3	32.45	REACHED
4	17.64	REACHED after blocked
5	14.03	REACHED
6	20.02	REACHED after blocked
7	12.49	REACHED
8	31.56	REACHED
9	3.19	REACHED
10	–	FAILED
11	2.85	REACHED from start
12	76.88	REACHED
13	8.67	REACHED after blocked
14	5.85	REACHED
15	–	FAILED
16	7.42	REACHED from start
17	6.05	REACHED
18	39.36	REACHED
19	15.88	REACHED after blocked
20	5.52	REACHED
21	39.35	REACHED
22	17.57	REACHED
23	11.8	REACHED
24	2.29	REACHED
25	3.9	REACHED after blocked
26	15.15	REACHED
27	7.55	REACHED
28	6.41	REACHED
29	41.34	REACHED
30	1.54	REACHED

Table 5: Trial Results for Right

No.	Distance	Status
1	15.31	REACHED from start
2	6.36	REACHED
3	3.81	REACHED after blocked
4	17.99	REACHED after blocked
5	25.65	REACHED
6	14.34	REACHED after blocked
7	12.68	REACHED
8	30.22	REACHED
9	4.17	REACHED
10	–	FAILED
11	2.86	REACHED from start
12	–	FAILED
13	21.62	REACHED from start
14	4.95	REACHED
15	–	FAILED
16	21.23	REACHED from start
17	5.29	REACHED
18	29.59	REACHED
19	13.48	REACHED after blocked
20	6.22	REACHED
21	24.78	REACHED
22	–	FAILED
23	14.36	REACHED from start
24	1.67	REACHED
25	3.86	REACHED after blocked
26	16.69	REACHED
27	8.13	REACHED
28	6.37	REACHED
29	21.61	REACHED
30	1.92	REACHED