Report Assignment 3

Statistics Analysis

Università di Genova

Course: Research Track 2

Professor: Carmine Recchiuto

Student: Veronica Gavagna s5487110

s.y. 2022/23

Abstract:

This study compares the failure and success rates of two algorithms (Algorithm 1 and Algorithm 2) under the same initial conditions. The chi-square test was used to assess if there is a significant difference in their performance.
Data on failures and successes were collected for a sample of 80 tests, and a contingency table was constructed. The null hypothesis assumed no significant difference, while the alternative hypothesis suggested a difference.
The second algorithm consistently failed when more than 6 boxes were involved, so two analyses were conducted, one with the full sample and one with a reduced sample excluding over 7-box cases. Expected frequencies were calculated assuming equal rates.
The chi-square statistic was computed, and the degrees of freedom were 1. The null hypothesis was rejected in the full sample test ($p < 0.001$), but not in the reduced sample ($p < 0.05$).

Introduction:

This study presents a chi-square analysis conducted to compare the failure and success rates of two algorithms, Algorithm 1 (Veronica Gavagna) and Algorithm 2 (Simone Borelli) when provided with the same initial conditions with a random number of boxes between 2 and 10, and a constant internal radius of 1 and external radius of 2.
The chi-square test was employed to determine if there exists a statistically significant difference in the performance of the two algorithms.

Data were collected on the number of failures and successes for each algorithm under identical initial conditions. A contingency table on an 80-size sample test was constructed to organize the observed frequencies, with Algorithm 1 and Algorithm 2 represented in the rows, and failures and successes in the columns.

The hypotheses were defined as follows:
the null hypothesis (H0) stated that there is no significant difference in the failure and success rates between Algorithm 1 and Algorithm 1, while the alternative hypothesis (Ha) suggested that there is a significant difference.

Methodology:

1. Data Collection:
   The number of successes and failures were recorded for both Algorithm 1 and Algorithm 2 under identical initial conditions. The radius of the internal circle was set to 1 and the radius of the external circle was set to 2. The number of boxes was chosen randomly between 2 and 10.
   A contingency table was constructed to organize the observed frequencies, with two rows representing the two algorithms and two columns representing the number of failures and successes.

2. Hypotheses:
   - Null Hypothesis (H0): There is no significant difference in the failure and success rates between Algorithm 1 and Algorithm 2.
   - Alternative Hypothesis (Ha): There is a significant difference in the failure and success rates between Algorithm 1 and Algorithm 2.

3. Calculation of Expected Frequencies:
   Under the assumption that both algorithms have the same failure and success rates, the expected frequencies were calculated for each cell of the contingency table.
   The expected frequency for the two categories (success and failure) was determined using the formula:
   - Expected success: $\frac{Success\ Alg1+Success\ Alg2}{2}$
   - Expected failure: $\frac{Failure\ Alg1 + Failure\ Alg2}{2}$

4. Chi-square Test:
   The chi-square statistic was calculated using the formula:

$$chi - square = \sum_{1}^{2} \frac{(Succ.Obs_i - Succ.Exp_i)^2}{Succ.Exp_i} + \frac{(Fail.Obs_i - Fail.Exp_i)^2}{Fail.Exp_i}$$

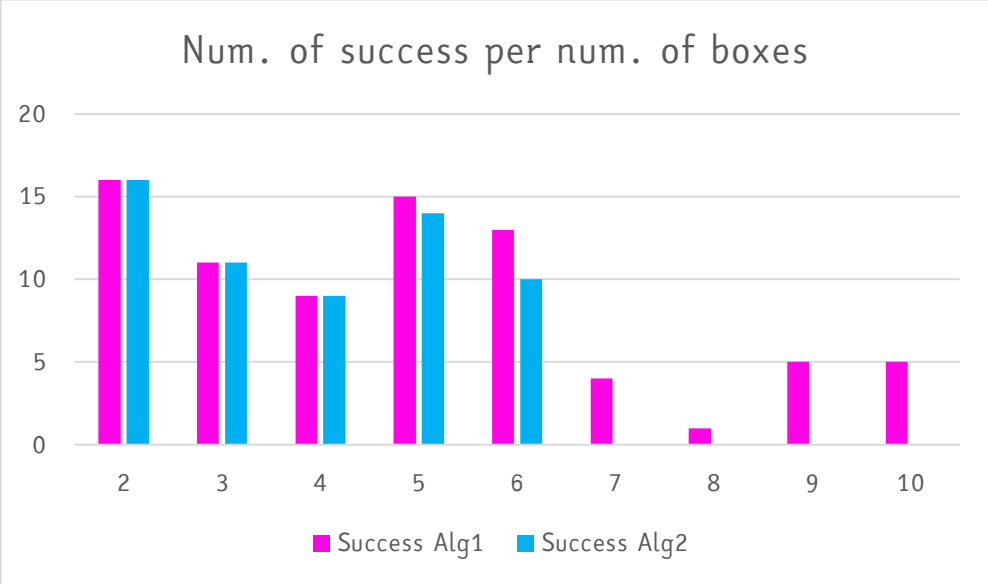
   The calculated chi-square value was compared to the critical chi-square value corresponding to the desired significance level and degrees of freedom.

	P										
DF	0.995	0.975	0.2	0.1	0.05	0.025	0.02	0.01	0.005	0.002	0.001
1	.0004	.00016	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.55	10.828

If the calculated chi-square value was greater than the critical chi-square value, the null hypothesis is rejected, indicating a significant difference in the failure and success rates between the algorithms. Otherwise, the null hypothesis is not rejected.

Results:

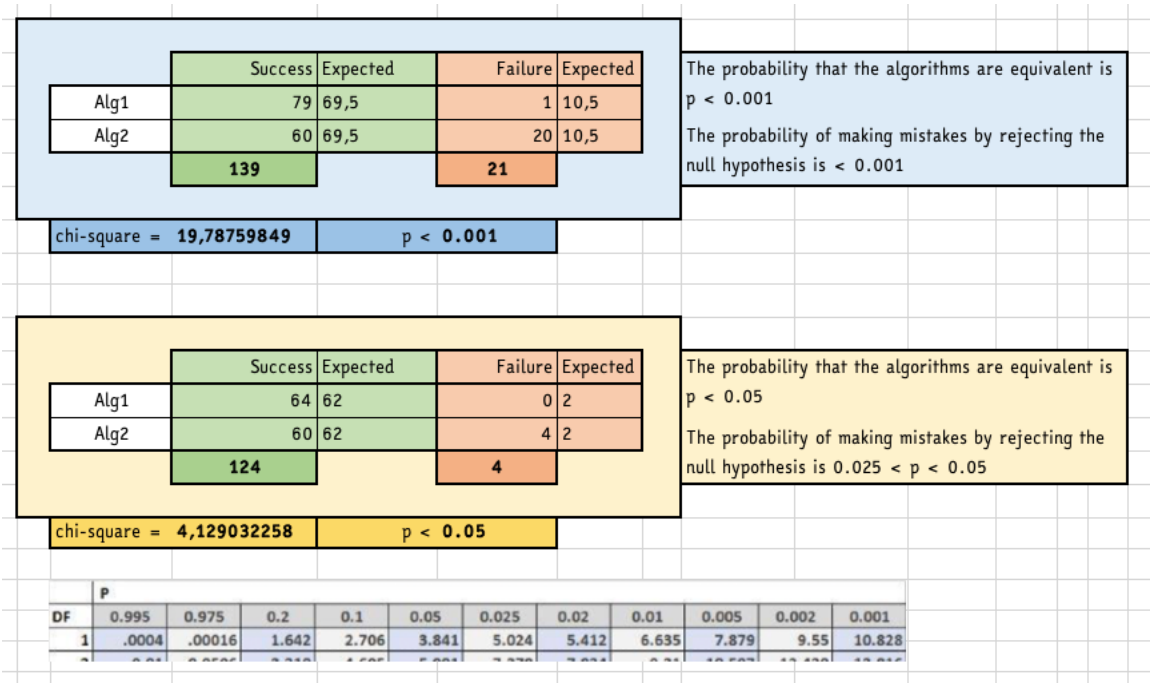
The results of the chi-square analysis provided insights into the performance of Algorithm 1 and Algorithm 2. If the null hypothesis was rejected, it indicated a significant difference in the failure and success rates between the algorithms, suggesting that one algorithm outperformed the other. Conversely, if the null hypothesis was not rejected, it implied that there was insufficient evidence to suggest a significant difference in performance.



Analyzing the data collected, it is possible to notice that the second algorithm fails every time over the 7 boxes. Because of that, two analysis was made: one with the full-size sample and one with a 64-size sample test containing data corresponding to a maximum of 6 boxes.

The percentage of success and failure of the two algorithms are reported in the following picture: in blue with the full-size sample and in yellow with the reduced-size sample.

Expected frequencies were calculated assuming both algorithms have the same failure and success rates. The chi-square statistic was computed using the observed and expected frequencies. Degrees of freedom were determined based on the number of rows and columns in the contingency table (DF = 2 – 1 = 1).



	Success	Expected	Failure	Expected
Alg1	79	69,5	1	10,5
Alg2	60	69,5	20	10,5
	139		21	

chi-square = 19,78759849 p < 0.001

The probability that the algorithms are equivalent is p < 0.001

The probability of making mistakes by rejecting the null hypothesis is < 0.001

	Success	Expected	Failure	Expected
Alg1	64	62	0	2
Alg2	60	62	4	2
	124		4	

chi-square = 4,129032258 p < 0.05

The probability that the algorithms are equivalent is p < 0.05

The probability of making mistakes by rejecting the null hypothesis is 0.025 < p < 0.05

	P										
DF	0.995	0.975	0.2	0.1	0.05	0.025	0.02	0.01	0.005	0.002	0.001
1	.0004	.00016	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.55	10.828

Conclusion:

The goal of the chi-square analysis in this study is to compare the failure and success rates of two algorithms, Algorithm 1 and Algorithm 2, given the same initial conditions. The chi-square test is employed to determine if there exists a statistically significant difference in the performance of the two algorithms.

By conducting the chi-square analysis, the study aims to assess whether Algorithm 1 and Algorithm 2 exhibit significantly different patterns of failures and successes. This analysis provides a statistical approach to evaluate the relative strengths and weaknesses of the algorithms under consideration.

The results of the chi-square analysis will help determine if there is evidence to support the hypothesis that Algorithm 1 and Algorithm 2 differ significantly in their failure and success rates. The findings will contribute to understanding the comparative performance of the algorithms and can guide decision-making in selecting the most suitable algorithm for a given task or application.

Overall, the goal of the chi-square analysis is to provide a quantitative assessment of the performance differences between Algorithm 1 and Algorithm 2, offering insights into their relative effectiveness under the same initial conditions.

In conclusion, this study utilized a chi-square analysis to compare the failure and success rates of two algorithms, Algorithm 1 and Algorithm 2, when subjected to the same initial conditions. The data collected on failures and successes provided insights into the relative strengths and weaknesses of the algorithms.

The analysis revealed that Algorithm 2 consistently failed when more than 7 boxes were involved, leading to two separate analyses: one with the full sample of 80 tests and another with a reduced sample excluding the over 7-box cases. The null hypothesis assumed no significant difference between the algorithms, while the alternative hypothesis suggested a difference.

By comparing the calculated chi-square value with the critical chi-square value corresponding to the desired significance level and degrees of freedom, the null

hypothesis was rejected in the first sample test because the probability of committing an error by rejecting it less than 0.001, while in the second sample test, the probability of committing an error by rejecting it is less than 0.05.

It is important to note that the chi-square test assumes certain conditions, such as the independence of observations and an adequate sample size. These assumptions should be verified to ensure the validity of the results.

Based on the findings of this study, further analysis and investigation can be conducted to identify the factors that contribute to the observed differences in the failure and success rates between Algorithm 1 and Algorithm 2. This may involve examining additional variables, expanding the sample size, or conducting experiments with different initial conditions such (as the position of the boxes) to gain a more comprehensive understanding of their relative strengths and weaknesses.