Journal of Computers, Mechanical and Management
Journal homepage: www.jcmm.co.in
e-ISSN: 3009-075X
AN PUBLISHING

Volume 3 Issue 1

Article Number: 240101

# The Study of the Value of $\pi$ Probability Sampling by Testing Hypothesis and Experimentally

Sanjay Kulkarni*[1] and Sandeep Kulkarni[2]

[1]Department of First Year Engineering, Hope Foundation's Finolex Academy of Management and Technology, Ratnagiri, Maharashtra, India, 415639
[2]Tata Consultancy Services, Pune, Maharashtra, India

**Abstract**

This study evaluated the value of $\pi$ using the Monte Carlo Simulation Method and compared the results with experimental values. The experimental value of $\pi$ was determined by considering a unit circle $|z| = 1$ centered at the origin, inscribed within a square with vertices $(0, 0)$, $(1, 0)$, $(1, 1)$, and $(0, 1)$. Points were randomly generated within the square, where points satisfying $|z| \leq 1$ lay within the circle, and those with $|z| \geq 1$ lay outside the circle but within the square. By selecting large numbers of random pairs and determining their positions relative to the circle, the ratio $\pi = \frac{4n}{N}$ was calculated, where $N$ was the total number of points and $n$ was the number of points within the circle. Larger sample sizes yielded values of $\pi$ closer to the true value. The distribution of Monte Carlo Simulation results, using 20 triplets of random numbers, was examined with non-parametric tests such as Friedman's Test. Ranks were assigned to the 20 random numbers row-wise for each triplet. The null hypothesis, asserting that all triplets had identical effects, was tested and showed significant differences at the 5% level. Additionally, the distribution was tested for goodness of fit using a Chi-Square Test at a 5% significance level. Results indicated that the triplets of random numbers conformed to the expected distribution.

**Keywords:** Random Number; Triplets; Non-Parametric Test; Friedman's Test; Chi-Square Test

## 1 Introduction

The technique of simulation is extensively utilized in the physical sciences and is increasingly becoming a crucial tool for addressing complex problems in managerial decision-making. Scale models of machines are used to simulate plant layouts, and models of aircraft are tested in wind tunnels to determine their aerodynamic characteristics. Simulation, aptly described as a management laboratory, assesses the effects of various alternative policies without disrupting the real system. Techniques such as linear programming, dynamic programming, queuing theory, and network models are insufficient to tackle all significant managerial problems requiring data analysis, each having its limitations. When characteristics such as uncertainty, complexity, dynamic interaction between decisions and subsequent events, and the need to develop detailed procedures with finely divided time intervals combine in one scenario, it becomes too complex for traditional mathematical programming and probabilistic models. Such situations necessitate analysis by alternative quantitative techniques that provide accurate and reliable results. The Monte Carlo method of simulation, developed by mathematicians John von Neumann and Stanislaw Ulam during World War II, was initially used to study neutron travel through various materials. The technique provided an approximate but workable solution to this problem and soon became popular, finding numerous applications in business and industry.

*Corresponding author: sanjay.kulkarni@famt.ac.in
**Received:** 02 September 2023; **Revised:** 16 December 2023; **Accepted:** 16 December 2023; **Published:** 29 February 2023

**DOI:** 10.57159/gadl.jcmm.3.1.240101.

It is now a vital tool in the operations researcher's toolkit [1–7]. In computer science and its applications, new or improved algorithms are compared with existing ones on several datasets to demonstrate superior performance. Let $x_1$ represent the control algorithm, and $x_2, x_3, x_4, \ldots, x_k$ represent the $k-1$ benchmark algorithms. The challenge lies in better judging whether the control algorithm $x_1$ has a significant advantage over other benchmark algorithms on experimental datasets. Due to dataset diversity and various random factors in training and testing, it is rare for the control algorithm to perform better on all datasets. Therefore, meaningful conclusions require statistical hypothesis tests [8]. These tests are categorized into parametric and non-parametric tests [9–11]. Parametric tests assume that the data follows a known probability distribution and make inferences about distribution parameters [10]. Conversely, non-parametric tests typically have no preliminary assumptions about the data distribution, making them applicable in various circumstances. Most non-parametric tests use rankings instead of raw data for hypothesis testing. A transformation procedure is adopted to obtain rankings for control and benchmark algorithms [9, 10]. The choice of statistical tests depends on the specific application, data characteristics, and researcher preferences. Parametric tests often have assumptions regarding data characteristics for comparison. For instance, analysis of variance (ANOVA) requires data to meet conditions such as independence, normality, and homogeneity [10]. When these assumptions are met, parametric tests are more effective [10]. Otherwise, parametric tests can produce biased conclusions. In practical applications, it is rarely possible to verify that algorithm results on different datasets satisfy these assumptions. Therefore, non-parametric tests are commonly considered [9, 12]. When selecting a non-parametric test, it is necessary to distinguish between pairwise and multiple comparisons. Non-parametric tests for pairwise comparisons include the Wilcoxon signed-rank test [13]. For multiple comparisons, non-parametric tests include the Friedman test, multiple sign tests [14], and contrast estimation based on medians [15]. Although non-parametric tests are widely adopted in published papers [16–19], the Friedman test is particularly effective and widely used by many scholars [20, 21]. O'Gorman [22] compared the F-test, Friedman test, and several aligned Friedman tests using Monte Carlo simulation. The Monte Carlo technique employs random numbers and is suitable for problems involving probability where physical experimentation is impractical, and mathematical model formation is impossible. It is a simulation method using a sampling technique. The steps involved in conducting a Monte Carlo simulation are:

- Select the measure of effectiveness of the problem.
- Identify the variables that significantly affect the measure of effectiveness.
- Determine the cumulative probability distribution of each variable selected in step 2.
- Generate a set of random numbers.
- Consider each random number as a decimal value of the cumulative probability distribution.
- Record the value(s) of the variables generated in step 5. Substitute in the formula chosen for the measure of effectiveness and find its simulated value.
- Repeat steps 5 and 6 until the sample is large enough to satisfy the decision maker.

This paper aims to study the value of $\pi$ obtained through the Monte Carlo Simulation method and to compare the results with experimental values. The Monte Carlo Simulation distribution is tested by applying non-parametric hypothesis testing methods, such as Friedman's Test and the Chi-Square Test.

# 2 Methods

## 2.1 Experimental Determination of $\pi$

The coordinate axes $OX$ and $OY$ were drawn. With center $O$, an arc $PR$ of unit radius was drawn, completing the square $OPQR$ as shown in Figure 1. The equation of the circle was established as $x^2 + y^2 = 1$. From the random number table, two numbers were selected, specifically 0.2068 and 0.7295, and were assigned as values to $x$ and $y$ respectively. The point $P_1(0.2068, 0.7295)$ was plotted. If $x^2 + y^2 \leq 1$, then $P_1$ lay inside or on the arc of the circle. Conversely, if $x^2 + y^2 > 1$, $P_1$ lay outside the arc but within the square.

Several pairs of random numbers were continuously selected, and it was determined whether the points represented by these numbers fell within or on the arc, or outside the arc but within the square. Let $N$ represent the total number of points considered, and $n$ represent the number of points that lay in or on the arc. Then, the Equation (1) was established:

$$\frac{n}{N} = \frac{\text{Area enclosed by the arc}}{\text{Area of the square}} \tag{1}$$

As the area enclosed by the arc is $\frac{\pi}{4}$ and the area of the square is 1, the ratio was calculated as given by Equation (2):

$$\frac{n}{N} = \frac{\pi}{4} \tag{2}$$

From this, $\pi$ was calculated using the Equation (3):

$$\pi = \frac{4n}{N} \tag{3}$$

This equation provided the experimental value of $\pi$. The method demonstrated that the larger the sample size $N$, the closer the obtained value was to the true value of $\pi$, effectively illustrating the utility of geometric random sampling for approximating $\pi$, as depicted in Figure 1.

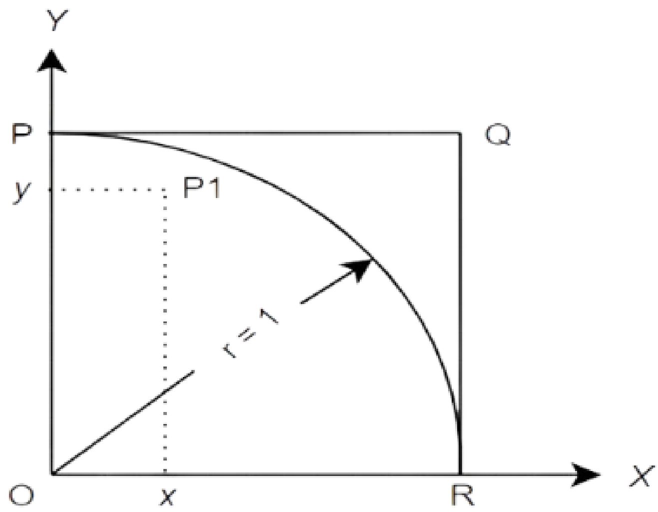


Figure 1: Illustration of points inside or outside the arc but within the square.

## 2.2 Monte Carlo Simulation Approach

Three points were chosen at random on the circumference of a circle using Monte Carlo methods to determine the probability that they lie on the same semicircle. A circle of circumference unity, i.e., of radius $\frac{1}{2\pi}$, was drawn as depicted in Figure 1. A triplet of random numbers (0.48, 0.51 and 0.06) was selected from the random number table. These numbers were plotted as points A, B and C on the circumference in Figure 2, with their positions from point $O$ along the circumference measured anticlockwise. Since the difference between the maximum (0.51) and minimum (0.06) values in this triplet was less than 0.50, the points were determined to lie on the same semicircle. The following general rule was applied to ascertain whether a triplet of random numbers lies on a semicircle:

1. The difference between the maxima and minima needs to be calculated. If this difference is $\leq 0.50$, the triplet is considered to lie on a semicircle.

2. If the difference is $> 0.50$, unity is added to those random numbers in the triplet that were $< 0.50$ and to the minimum random number in the triplet. The new difference between the maxima and minima needs to be then found. If this difference is $\leq 0.50$, the triplet is considered to lie on a semicircle; otherwise, it did not.

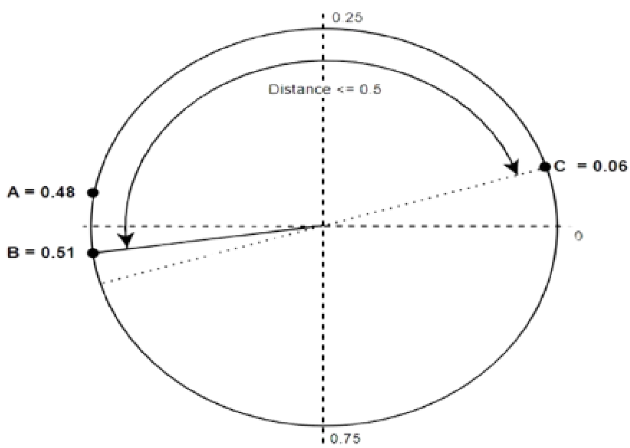


Figure 2: Three points chosen at random on the circumference of a circle using Monte Carlo methods.

### 2.3 Friedman's Test / Two-Way Analysis of Variance by Ranks

Friedman's test is a non-parametric test utilized to identify differences across multiple treatments on the same subjects. Being non-parametric, this test does not assume that the data originates from a specific distribution, such as the normal distribution. This test is effectively an extension of the sign test and is applied when there are more than two treatments, although it reduces to the sign test with only two treatments. This statistic is applicable in two distinct scenarios that may seem different but fundamentally address the same statistical question: either measuring the same quantitative variable at different times or measuring different comparable quantitative variables from the same sample. In both cases, Friedman's test serves to compare the distributions of the variables and thus, it was used in the present work.

### 2.4 Chi Square ($\chi^2$) Test to Test the Goodness of Fit

The $\chi^2$ test, a non-parametric statistic, is utilized to evaluate the degree of correspondence between observed frequencies and those expected under a specified hypothesis. This test is especially appropriate for categorical data as it does not assume a normal distribution. Given these characteristics, the $\chi^2$ test is particularly well-suited for the analysis of triplet data in this study. It was thus used in the present work to provide a methodological framework to assess whether observed variances from expected frequencies were statistically significant, thereby testing the underlying hypotheses of the research.

## 3 Results

### 3.1 Monte Carlo Simulation Outcomes

Table 1 presents the results of the Monte Carlo simulation approach, indicating whether each of 20 triplets lies on a semicircle.

Table 1: The Monte Carlo Simulation Techniques for Triplets of Random Numbers.

| No. | Triplet - 1 | Triplet - 2 | Triplet - 3 | Diff Max & Min | Diff New Max & Min | Triplet on Semi-Circle | Min Diff (x) |
|---|---|---|---|---|---|---|---|
| 1 | 0.21 | 0.11 | 0.71 | 0.60 | 0.90 | | 0.6 |
| 2 | 0.65 | 0.41 | 0.35 | 0.30 | | 0.3 | 0.3 |
| 3 | 0.17 | 0.91 | 0.07 | 0.84 | 0.90 | | 0.84 |
| 4 | 0.34 | 0.12 | 0.43 | 0.31 | | 0.31 | 0.31 |
| 5 | 0.38 | 0.49 | 0.13 | 0.36 | | 0.36 | 0.36 |
| 6 | 0.05 | 0.96 | 0.76 | 0.91 | 0.29 | 0.29 | 0.29 |
| 7 | 0.85 | 0.69 | 0.57 | 0.28 | | 0.28 | 0.28 |
| 8 | 0.63 | 0.41 | 0.03 | 0.60 | 0.62 | | 0.6 |
| 9 | 0.91 | 0.58 | 0.62 | 0.33 | | 0.33 | 0.33 |
| 10 | 0.75 | 0.89 | 0.23 | 0.66 | 0.48 | 0.48 | 0.48 |
| 11 | 0.21 | 0.36 | 0.59 | 0.38 | | 0.38 | 0.38 |
| 12 | 0.39 | 0.19 | 0.21 | 0.20 | | 0.2 | 0.2 |
| 13 | 0.74 | 0.86 | 0.90 | 0.16 | | 0.16 | 0.16 |
| 14 | 0.64 | 0.18 | 0.67 | 0.49 | | 0.49 | 0.49 |
| 15 | 0.20 | 0.72 | 0.34 | 0.52 | 0.86 | | 0.52 |
| 16 | 0.54 | 0.30 | 0.22 | 0.32 | | 0.32 | 0.32 |
| 17 | 0.48 | 0.74 | 0.76 | 0.28 | | 0.28 | 0.28 |
| 18 | 0.02 | 0.07 | 0.64 | 0.62 | 0.95 | | 0.62 |
| 19 | 0.95 | 0.23 | 0.91 | 0.72 | 0.32 | 0.32 | 0.32 |
| 20 | 0.48 | 0.55 | 0.91 | 0.43 | | 0.43 | 0.43 |
| **Total** | | | | | | **15** | **0.4** |

Out of 20 triplets, 15 were found to lie on a semicircle, yielding a calculated probability of $\frac{15}{20} = 0.75$. The experimentally determined value of $\pi$, where $N = 20$ and $n = 15$, was calculated as 3 using Equation 4:

$$\pi = \frac{4n}{N} \tag{4}$$

It was observed that increasing the sample size $N$ would bring the value of $\pi$ closer to its experimental value.

### 3.2 Friedman's Test Results

For the Friedman's test, the ranks of 20 randomly selected numbers were assigned row-wise for each Triplet – 1, Triplet – 2, and Triplet – 3, as illustrated in Table 2.

Table 2: Details of Ranks Assigned Row-wise for Each Triplet

| Sr. No | Triplet 1 | Triplet 2 | Triplet 3 | Rank of Triplet 1 | Rank of Triplet 2 | Rank of Triplet 3 |
|---|---|---|---|---|---|---|
| 1 | 0.21 | 0.11 | 0.71 | 2 | 1 | 3 |
| 2 | 0.65 | 0.41 | 0.35 | 3 | 2 | 1 |
| 3 | 0.17 | 0.91 | 0.07 | 2 | 3 | 1 |
| 4 | 0.34 | 0.12 | 0.43 | 2 | 1 | 3 |
| 5 | 0.38 | 0.49 | 0.13 | 2 | 3 | 1 |
| 6 | 0.05 | 0.96 | 0.76 | 1 | 3 | 2 |
| 7 | 0.85 | 0.69 | 0.57 | 3 | 2 | 1 |
| 8 | 0.63 | 0.41 | 0.03 | 3 | 2 | 1 |
| 9 | 0.91 | 0.58 | 0.62 | 3 | 1 | 2 |
| 10 | 0.75 | 0.89 | 0.23 | 2 | 3 | 1 |
| 11 | 0.21 | 0.36 | 0.59 | 1 | 2 | 3 |
| 12 | 0.39 | 0.19 | 0.21 | 3 | 1 | 2 |
| 13 | 0.74 | 0.86 | 0.90 | 1 | 2 | 3 |
| 14 | 0.64 | 0.18 | 0.67 | 2 | 1 | 3 |
| 15 | 0.20 | 0.72 | 0.34 | 1 | 3 | 2 |
| 16 | 0.54 | 0.30 | 0.22 | 3 | 2 | 1 |
| 17 | 0.48 | 0.74 | 0.76 | 1 | 2 | 3 |
| 18 | 0.02 | 0.07 | 0.64 | 1 | 2 | 3 |
| 19 | 0.95 | 0.23 | 0.91 | 3 | 1 | 2 |
| 20 | 0.48 | 0.55 | 0.91 | 1 | 2 | 3 |
| **Total of Ranks** | | | | **40** | **39** | **41** |

The null hypothesis assumes that all triplets exert identical effects, while the alternative hypothesis suggests that the effects vary among the triplets. Here, $N$ represents the total number of triplets (20), and $k$, the number of conditions, is 3. The total ranks for each column were 40, 39, and 41, respectively. The Friedman test statistic was calculated as 380.16 using Equation (5):

$$FM = \frac{12N}{k(k+1)}\left(\sum R^2 - \frac{k(k+1)^2}{4}\right) \tag{5}$$

With the significance level set at 5% and the degrees of freedom $df = k-1 = 2$, the critical value from the Chi-Square table for 2 degrees of freedom at 5% significance is $FM_{\text{table}} = 5.99$. Given that the computed Friedman test statistic $FM_{\text{Calculated}} = 380.16$ significantly exceeds the critical value $FM_{\text{table}} = 5.99$, the null hypothesis is rejected. This result indicates significant differences in the effects of the triplets.

### 3.3 Chi-Square Goodness-of-Fit Test

The Chi-Square Goodness-of-Fit test was conducted to determine if the distribution of differences between maxima and minima across triplets aligns with a theoretical or expected distribution. The obtained results are shown in Table 3 The hypotheses were formulated as follows:

- Null Hypothesis ($H_0$): The observed distribution matches the expected distribution.
- Alternative Hypothesis ($H_1$): There is a significant difference between the observed and expected distributions.

The expected frequency ($E$) and the observed frequency ($O$) for each category were calculated, leading to the $\chi^2$ statistic using Equation (6).

$$\chi^2 = \sum \frac{(O-E)^2}{E} \tag{6}$$

With $N = 20$ triplets and the calculated $\chi^2$ value of 1.2977, we compare this to the critical value for 19 degrees of freedom at a 5% level of significance, 30.14. Since $1.2977 < 30.14$, we accept the null hypothesis, indicating no significant difference between the observed and expected distributions. This suggests that the triplets follow the expected goodness of fit, underscoring the importance of understanding both the statistical significance of test results and their practical implications in research contexts.

Table 3: Details of Observed and Expected Minimum Difference between Maxima and Minima

| Sr. No | Triplet - 1 | Triplet - 2 | Triplet - 3 | Difference between Maxima and Minima | Difference between New Maxima and New Minima | Minimum Difference between Maxima and Minima (O) | O-E | $(O\text{-}E)^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.21 | 0.11 | 0.71 | 0.60 | 0.90 | 0.6 | 0.2 | 0.04 |
| 2 | 0.65 | 0.41 | 0.35 | 0.30 | | 0.3 | -0.1 | 0.01 |
| 3 | 0.17 | 0.91 | 0.07 | 0.84 | 0.90 | 0.84 | 0.44 | 0.194 |
| 4 | 0.34 | 0.12 | 0.43 | 0.31 | | 0.31 | -0.09 | 0.008 |
| 5 | 0.38 | 0.49 | 0.13 | 0.36 | | 0.36 | -0.04 | 0.002 |
| 6 | 0.05 | 0.96 | 0.76 | 0.91 | 0.29 | 0.29 | -0.11 | 0.0121 |
| 7 | 0.85 | 0.69 | 0.57 | 0.28 | | 0.28 | -0.12 | 0.014 |
| 8 | 0.63 | 0.41 | 0.03 | 0.60 | 0.62 | 0.6 | 0.2 | 0.04 |
| 9 | 0.91 | 0.58 | 0.62 | 0.33 | | 0.33 | -0.07 | 0.005 |
| 10 | 0.75 | 0.89 | 0.23 | 0.66 | 0.48 | 0.48 | 0.08 | 0.006 |
| 11 | 0.21 | 0.36 | 0.59 | 0.38 | | 0.38 | -0.02 | 0.0004 |
| 12 | 0.39 | 0.19 | 0.21 | 0.20 | | 0.2 | -0.2 | 0.04 |
| 13 | 0.74 | 0.86 | 0.9 | 0.16 | | 0.16 | -0.24 | 0.058 |
| 14 | 0.64 | 0.18 | 0.67 | 0.49 | | 0.49 | 0.09 | 0.008 |
| 15 | 0.2 | 0.72 | 0.34 | 0.52 | 0.86 | 0.52 | 0.12 | 0.014 |
| 16 | 0.54 | 0.3 | 0.22 | 0.32 | | 0.32 | -0.08 | 0.006 |
| 17 | 0.48 | 0.74 | 0.76 | 0.28 | | 0.28 | -0.12 | 0.014 |
| 18 | 0.02 | 0.07 | 0.64 | 0.62 | 0.95 | 0.62 | 0.22 | 0.048 |
| 19 | 0.95 | 0.23 | 0.91 | 0.72 | 0.32 | 0.32 | -0.08 | 0.006 |
| 20 | 0.48 | 0.55 | 0.91 | 0.43 | | 0.43 | 0.03 | 0.0001 |
| | | | | | | Total | 0.405 | 0.5256 |

# 4 Discussion

The Monte Carlo Simulation method was employed to estimate the value of $\pi$, utilizing random sampling within a geometric framework. The efficacy of this method is strongly influenced by the sample size $N$; an increase in $N$ leads to an estimated value of $\pi$ that converges more closely to its true value. This convergence is supported by the Law of Large Numbers, which asserts that the average of the results from a large number of trials will approximate the expected value. The Monte Carlo method involves generating pairs of random numbers and determining whether these points lie within a unit circle inscribed in a square. The ratio of points within the circle to the total number of points, multiplied by 4, yields an estimate of $\pi$. This approach is valued for its simplicity and computational efficiency, making it a popular choice for numerical integration and probabilistic simulations. To validate the distribution of the generated random numbers, non-parametric tests such as Friedman's Test and the Chi-Square Test were applied. Friedman's Test, which detects differences in treatments across multiple attempts, indicated that the triplets of random numbers have varying effects. This conclusion was derived by ranking the triplets and calculating the test statistic, which significantly exceeded the critical value at a 5% significance level, suggesting variability among the triplets and highlighting the stochastic nature of random sampling. The Chi-Square Test for goodness of fit assessed the alignment of the observed distribution of random triplets with the expected theoretical distribution. The calculated $\chi^2$ value was significantly lower than the critical value, leading to the acceptance of the null hypothesis. This indicates that the triplets of random numbers closely follow the expected distribution, affirming the reliability of the Monte Carlo Simulation method in generating random samples that conform to theoretical expectations. These findings underscore the importance of employing robust statistical methods to analyze and validate the results of simulations, confirming the effectiveness of the Monte Carlo method and the appropriateness of non-parametric tests in hypothesis testing for random samples.

# 5 Conclusions

This paper investigates the value of $\pi$ obtained through the Monte Carlo Simulation method and compares the results with the experimental value of $\pi$. It also tests the distribution of the Monte Carlo Simulation by applying non-parametric hypothesis testing methods, such as Friedman's Test and the Chi-Square Test. The detailed discussion and analysis provided offer significant insights into the application of these statistical techniques. Key findings include:

- **Accuracy of Monte Carlo Simulation:** The accuracy of the Monte Carlo Simulation improves with larger sample sizes $N$, underscoring the importance of scale in such simulations to achieve closer approximations of mathematical constants.
- **Effect of Random Triplets:** Friedman's Test reveals variability among triplets of random numbers, rejecting the null hypothesis that all triplets have identical effects. This finding is critical in considering the randomness and distribution of data points in simulations.

- **Goodness of Fit:** The Chi-Square Test confirms that the triplets of random numbers adhere well to the theoretical distribution, accepting the null hypothesis of goodness of fit and affirming the consistency of the Monte Carlo method with expected theoretical results.

These findings have broader implications for the fields of computational engineering and data science. By demonstrating the reliability and accuracy of Monte Carlo simulations in estimating $\pi$ and analyzing random distributions, this study provides a robust framework for students, researchers, and data analysts. The application of non-parametric tests like Friedman's Test and the Chi-Square Test provides powerful tools for hypothesis testing, enhancing informed decision-making based on empirical data. The methodologies and results discussed can serve as references for further research and applications in various domains where statistical analysis and simulations are pivotal. Understanding the behavior of random numbers and their distributions is crucial for optimizing algorithms, enhancing data analysis techniques, and improving the accuracy of computational models. In summary, this paper not only confirms the effectiveness of the Monte Carlo Simulation method in approximating $\pi$ but also highlights the importance of statistical hypothesis testing in validating simulation results. These contributions are expected to foster deeper insights and innovations in computational and statistical methods.

## 6 Acknowledgment

The authors are thankful to Management of Hope Foundation's Finolex Academy of Management and Technology, Ratnagiri, for providing excellent computational facilities. Authors also, thankful to unknown referees for their comment and suggestions to substantial improvement of quality of the paper.

## Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding Declaration

This research did not receive any grants from governmental, private, or nonprofit funding bodies.

## Author Contribution

**Sanjay Kulkarni**: Conceptualization, Methodology, Investigation, Visualization, Writing - original draft, review and editing. **Sandeep Kulkarni**: Investigation, Visualization, Resources.

## References

[1] Y. Dodge, *Statistical data analysis and inference.* Elsevier, 2014.

[2] C. Ismay and A. Y. Kim, *Statistical Inference via Data Science.* 2019.

[3] J. S. Milton and J. C. Arnold, *Introduction to Probability and Statistics.* 2007.

[4] R. A. Johnson and C. B. Gupta, *Probability and Statistics for Engineers.* 2007.

[5] A. Chandrasekaran and G. Kavitha, *Probability Statistics Random Processes and Queuing Theory.* 2014.

[6] P. Bruce and A. Bruce, *Practical Statistics for Data Scientists.* 2017.

[7] J. A. Rice, *Mathematical Statistics and Data Analysis.* Thomson Higher Education, 2013.

[8] Z. Liu, F. Blasch, and V. Jhon, "Statistical comparison of image fusion algorithms recommendations," *Inf. Fusion*, 2017.

[9] S. Couch, Z. Kazan, K. Shi, A. Bray, and A. Groce, "Differentially private nonparametric hypothesis testing," in *Proc. Of the ACM Conf. Comput. Commun. Secur.*, 2019.

[10] S. García, A. Fernández, J. Luengo, and F. Herrera, "Advanced non-parametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: experimental analysis of power," *Inf. Sci.*, 2010.

[11] S. D. Pawar and D. T. Shirke, "Nonparametric tests for multivariate multi-sample locations based on data depth," *J. Stat. Comput. Simul.*, 2019.

[12] J. Derrac, S. García, S. Hui, P. N. Suganthan, and F. Herrera, "Analyzing convergence performance of evolutionary algorithms: a statistical approach," *Inf. Sci.*, 2014.

[13] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *J Mach Learn Res.*, vol. 7, no. 1, 2006.

[14] R. G. D. Steel, "A multiple comparison sign test: treatments versus control," *J. Am. Stat. Assoc.*, 1959.

[15] K. Doksum, "Robust procedures for some linear models with one observation per cell," *Ann. Math. Stat.*, 1967.

[16] T. B. Chandra, K. Verma, B. K. Singh, D. Jain, and S. S. Netam, "Coronavirus disease (covid-19) detection in chest x-ray images using majority voting-based classifier ensemble," *Expert Syst. Appl.*, 2021.

[17] M. De Gregorio and M. Giordano, "An experimental evaluation of weightless neural networks for multi-class classification," *Appl. Soft Comput. J.*, 2018.

[18] F. J. Pulgar, F. Charte, A. J. Rivera, and M. J. Del Jesus, "Choosing the proper autoencoder for feature fusion based on data complexity and classifiers: analysis, tips and guidelines," *Inf. Fusion*, 2020.

[19] S. Shi, S. Ding, Z. Zhang, and W. Jia, "Energy-based structural least squares mbsvm for classification," *Appl. Intell.*, 2020.

[20] J. Liu, "Fuzzy support vector machine for imbalanced data with borderline noise," *Fuzzy Sets Syst.*, 2020.

[21] M. Petrović, Z. Miljković, and A. Jokić, "A novel methodology for optimal single mobile robot scheduling using whale optimization algorithm," *Appl. Soft Comput. J.*, 2019.

[22] T. W. O'Gorman, "A comparison of the f-test, friedman's test, and several aligned rank tests for the analysis of randomized complete blocks," *J. Agric. Biol. Environ. Stat.*, 2001.