Evaluating the Randomness of Digit Sequences via Statistical Tests

Sanjana Prasanth
Redondo Union High School, Redondo Beach, CA

Abstract
This study applies three statistical tests, the chi-squared test, the permutation test, and the gap test, to assess whether the first 15 million digits of the number π (pi) exhibit randomness. Groups of one through five digits are examined.

Keywords: random digits, number π (pi), chi-squared test, permutation test, gap test

1 Introduction

Many programs rely on random numbers to generate unique and varied outputs. However, no program can produce truly random numbers; even pseudo-random generators, whose outputs appear random, follow deterministic rules and eventually repeat. As a result, many "random" numbers are, in fact, predictable, an issue that can be problematic depending on the application.

It is therefore important to test numerical sequences for randomness and ensure that no systematic bias is present. A truly random sequence should be both unpredictable and have all values occur with equal probability. Because randomness implies equal likelihood across outcomes, a random set should follow a uniform distribution. Pairs of values should also follow a uniform distribution to ensure that the same values do not consistently occur next to each other. Ideally, a random set of numbers would be uniform for single values, pairs, triplets, or any grouping of digits.

A decimal number is termed **normal** if every sequence of g consecutive digits appears with probability 10^{-g} [1]. For example, the one-digit values 0–9 should each occur with probability $10^{-1} = 0.1$, and the two-digit values 00–99 should each occur with probability $10^{-2} = 0.01$. This means that the digits of a normal number follow a discrete uniform distribution and can be tested for uniformity. By this definition, the digits of a normal number would exhibit the same patterns as a set of truly random numbers.

An example of a seemingly normal number is π, which has been tested extensively using numerous methods, all suggesting that its digits occur randomly [2,3]. Because π is an infinite, non-repeating number that cannot be fully tested, Bailey et al. (2012) concluded that it is impossible to definitively claim its digits are random. However, they found that the more digits are analyzed, the stronger the evidence supporting the randomness of π's digits becomes.

In this study, we apply several statistical tests of randomness to evaluate whether they agree that π can be considered normal, essentially testing its digits for randomness up to groups of five digits. Common theoretical approaches for testing randomness include the chi-squared, gap, and permutation tests; therefore, we focused on examining whether these methods yield results consistent with prior findings [4]. The number π is an ideal candidate for such analysis due to the vast number of available digits and the absence of any known pattern, although these tests can be applied to any numerical sequence. For this analysis, we used the first 15 million digits of π obtained from `calculat.io`.

2 Goodness-of-Fit Tests for Uniformity: Theory

2.1 Hypotheses

Suppose we are given a string of $n \times k$ digits, and we group them into a sample of size n consisting of k-digit numbers $X_1, \ldots, X_n$. We would like to test whether these k-digit numbers are equally likely to appear in the sample [5]. Thus, the null hypothesis is given as

$$H_0 : P(X = i) = \frac{1}{k}, \quad i = 1, 2, \ldots, k,$$

and the alternative hypothesis is

$$H_1 : P(X = i) \neq \frac{1}{k} \quad \text{for some } i.$$

For example,

- For single-digit numbers (0–9), $k = 10$, so each digit is expected with probability 1/10.
- For double-digit numbers (00–99), $k = 100$, so each number is expected with probability 1/100.
- For triple-digit numbers (000–999), $k = 1000$, so each number is expected with probability 1/1000.
- And so forth for more digits.

2.2 Chi-Squared Test

Let O_i denote the observed frequency of outcome i ($i = 1, \ldots, k$). Under H_0, the expected frequency is

$$E_i = n \cdot \frac{1}{k} = \frac{n}{k}.$$

The chi-squared test statistic is

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}.$$

Under the null hypothesis,

$$\chi^2 \sim \chi^2_{k-1},$$

where $k-1$ is the number of degrees of freedom (since $\sum_i O_i = n$ provides one constraint). Hence, the p-value is computed as

$$p\text{-value} = P\big(\chi^2_{k-1} > \chi^2\big).$$

2.3 Gap Test

The gap test is used to check for randomness by examining the gaps between occurrences of a specified set of digits. Let $S \subset \{X_1, \ldots, X_k\}$ be a chosen subset of "success" numbers. A gap is defined as the length of a sequence between successive occurrences of a number from S. We would like to test the null hypothesis that the numbers are independent and uniformly distributed. Under H_0, the gaps between occurrences of numbers in S follow a geometric distribution with success probability $p = |S|/k$. Let G_i denote the length of the i-th gap. Then under H_0, the probability mass function is

$$P(G_i = x) = (1 - p)^x p, \quad x = 0, 1, 2, \ldots,$$

where $p = |S|/k$ is the probability that a digit belongs to S.

To test the fit, we collect the observed frequencies of gaps of length $0, 1, 2, \ldots$ up to some maximum gap m. Let O_i be the observed number of gaps of length i, and $E_i = N \cdot (1-p)^i p$ the expected number under H_0, where N is the total number of gaps observed. The chi-squared test statistic is then

$$\chi^2 = \sum_{i=0}^{m} \frac{(O_i - E_i)^2}{E_i}.$$

Under H_0, $\chi^2 \sim \chi^2_m$, assuming expected frequencies E_i are at least 5. Otherwise, categories are pooled.

2.4 Permutation Test

A permutation test is a test used to check whether the sequence $X_1, \ldots, X_n$ exhibits randomness, based on the assumption that any ordering of the elements is equally likely under the null hypothesis. The hypotheses are:

- H_0: $X_1, \ldots, X_n$ are independently and identically distributed and the order is random.
- H_1: $X_1, \ldots, X_n$ exhibit a non-random ordering or pattern.

Choose a test statistic $T(X_1, \ldots, X_n)$ that measures a property of interest in the sequence, for example:

- Number of ascending runs or descending runs.
- Maximum deviation from expected frequency in a block of digits.
- Sum of differences between consecutive digits.

Let t_{obs} be the value of the statistic computed from the original sequence. Now carry out the following permutation procedure:

1. Randomly permute the sequence of digits to generate a new sequence under H_0.

2. Compute the test statistic T for this permuted sequence.

3. Repeat steps 1–2 many times (e.g., $B = 10,000$ permutations) to generate the null distribution of T.

The p-value is estimated as the proportion of permutations where the test statistic is at least as extreme as the observed value:

$$p\text{-value} = \frac{\#\{T_{\text{perm}} \geq t_{\text{obs}}\}}{B}.$$

If p-value is small, the observed ordering is unlikely under the null hypothesis of randomness.

3 Applications and Results

To test the digits of π for uniformity, the first 15 million digits were partitioned into sets of 15 million one-digit numbers, 7.5 million two-digit numbers, 5 million three-digit numbers, 3.75 million four-digit numbers, and 3 million five-digit numbers. We apply the chi-squared test to determine whether these groupings follow a discrete uniform distribution. The resulting p-values are summarized in the table below.

	One-digit	Two-digit	Three-digit	Four-digit	Five-digit
p-value	0.9068	0.8724	0.8915	0.9185	0.8279

As all the p-values exceed 0.05, we accept the hypothesis that the digits of *pi* are random up to groups of five. Next, we apply the permutation test and get the following p-values:

	One-digit	Two-digit	Three-digit	Four-digit	Five-digit
p-value	0.7482	0.8756	0.8904	0.9160	0.8362

For the permutation tests, the full set of 15 million digits of π was not used, as the computational time required would have been excessive. Instead, each test was conducted using a sample size such that each possible value would theoretically occur about 30 times.

For one-digit values, this corresponded to 300 digits, since there are ten possible outcomes. Two-digit values were tested using 6,000 digits, three-digit values with 90,000 digits, four-digit values with 1,200,000 digits, and five-digit values with all 15,000,000 digits. Because these sample sets are smaller, the resulting p-values are expected to be slightly less precise than those from the chi-squared tests; however, they lead to similar conclusions regarding the uniformity of π since all of them are larger than 0.05.

Finally, the gap test was applied to assess uniformity. The gap test evaluates each value individually, beginning with one-digit numbers. For one-digit values, it tests the randomness of the digits 1–9, and for two-digit values, it tests the randomness of the numbers 1–99. To assess uniformity, the proportion of values with p-values greater than 0.05 (thus, accepted as random) was recorded. Due to the extensive computational time required to test all four- and five-digit numbers, the analysis was limited to three-digit values. The table below presents the proportions of p-values larger than 0.05.

	One-digit	Two-digit	Three-digit
proportions	0.8888	0.7878	0.8688

Once again, each p-value exceeds 0.05, indicating that π can be considered random up to at least three digits according to the gap test.

4 Conclusion and Discussion

The goodness-of-fit tests considered in this study, the chi-squared test, the permutation test, and the gap test, demonstrate that the first 15 million digits of π follow a uniform distribution, at least up to groups of five digits. This finding supports the conclusion that π is a normal number, consistent with numerous previous studies. As noted earlier, the properties of the digits of a normal number closely resemble those of a truly random set of numbers. Therefore, the tests employed here to demonstrate the uniformity of π's digits and its normality could also be applied to evaluate whether any set of numbers can be considered random. While this study examined groups of two to five digits, similar analyses could be performed on pairs, triplets, quadruplets, and quintuplets in an arbitrary set of random numbers.

4.1 Future Research Direction

Here, π was chosen to demonstrate these goodness-of-fit uniformity tests because it has an abundance of digits that have been shown to behave randomly. However, the same tests could be applied to evaluate the normality of other irrational numbers by examining the randomness of their digits. For example, they could be used to analyze the digits of e, the golden ratio, or $\sqrt{2}$. Although it is impossible to have a complete dataset of π's digits, as it is infinitely long, these tests could potentially be applied to even more digits of π to reach stronger conclusions regarding its normality. Future research could also explore additional goodness-of-fit tests for assessing randomness. For instance, the Kolmogorov–Smirnov test, equidistribution or frequency test, and serial test are all theoretically applicable for evaluating the uniformity of pseudo-random numbers [6-8].

Supplemental Materials

The complete R code used in this study and the dataset containing the first 15 million digits of π are available in the GitHub repository at

`https://github.com/sanjanaprasanth09/Randomness-Testing.git.`

Readers are encouraged to access the repository to review and reproduce the analyses.

Acknowledgments

I would like to give thanks especially to Dr. Olga Korosteleva, Professor of Statistics at California State University, Long Beach, for guiding me and providing mentorship through this project, as well as giving me the inspiration to continue searching for interesting avenues of research. I'm also grateful to the UCLA ORMC program for helping me to develop my mathematical skills and knowledge, especially to Dr. Oleg Gleizer for introducing this research opportunity to me. In addition, I'd like to express gratitude to my parents for constantly encouraging and assisting me through this process. Without them, I would not have the passion or skills to explore my interests and complete this research.

References

[1] Sierpinski, W. (1988). *Elementary Theory of Numbers*, Second English Edition (edited by A. Schinzel), Elsevier Science Ltd.

[2] Sevcik, C. (2009). *Fractal analysis of Pi normality.* arXiv preprint arXiv:0906.1278.

[3] Bailey, D. H. et al. (2012). Normality and the Digits of π, *Experimental Mathematics*, 21(4): 375-384.

[4] Kennedy, T. (2016). Chapter 3: Pseudo-random Number Generators. In *Monte Carlo Methods – A Special Topics Course.* University of Arizona. Available at: https://math.arizona.edu/∼tgk/

[5] Knuth, D. (1997). *The Art of Computer Programming, Volume 2, Seminumerical Algorithms*, Addison-Wesley Professional, 3rd edition.

[6] Bikos, A., Nastou, P. E., Petroudis, G., and Y. C. Stamatiou. (2023). Random number generators: Principles and applications, *Cryptography*, 7(4), 54.

[7] Hussain, I. and A. Khalique. (2019). Random Number Generators and Their Applications: A Review, *International Journal of Research in Electronics & Computer Engineering*, 7(2): 1777-1781.

[8] L'Ecuyer, P. (1994). Uniform Random Number Generation, *Annals of Operations Research*, 53: 77-120.