Journal of Mathematical Sciences: Advances and Applications
Volume 11, Number 2, 2011, Pages 69-77

SOME WONDERFUL STATISTICAL PROPERTIES OF PI-NUMBER DECIMAL DIGITS

MORTEZA KHODABIN

Department of Mathematics
Islamic Azad University
Karaj Branch
Karaj
Iran
e-mail: m-khodabin@kiau.ac.ir

Abstract

In this paper, we have considered an elementary statistical inference about the first 40960 digits of $\pi - 3$. We can see that, in the wonderful, in apogee of irregularity, they have intuitive regularity. Based on this study and according to number of decimal digits used for this analysis, we can guess that probably, this intuitive regularity can be established for all Pi-number decimal digits.

1. Introduction

π (sometimes written pi) is a mathematical constant, whose value is the ratio of any circle's circumference to its diameter in Euclidean space, this is the same value as the ratio of a circle's area to the square of its radius [1, 3, 8, 9]. It is approximately equal to 3.14159265 in the usual decimal notation. Many formulae from mathematics, science, and engineering involve π, which makes it one of the most important

2010 Mathematics Subject Classification: 60J10, 62G10, 94A17.

Keywords and phrases: Pi-number, uniform distribution, run tests, chi-square test, Shannon entropy.

Received September 23, 2011

mathematical constants. π is an irrational number, which means that its value cannot be expressed exactly as a fraction m / n, where m and n are integers [1, 24]. Consequently, its decimal representation never ends or repeats. It is also a transcendental number, which implies, among other things, that no finite sequence of algebraic operations on integers (powers, roots, sums, etc.) can be equal to its value; proving this was a late achievement in mathematical history and a significant result of 19th century German mathematics. Throughout the history of mathematics, there has been much effort to determine π more accurately and to understand its nature; fascination with the number has even carried over into non-mathematical culture. Probably, because of the simplicity of its definition, the concept of π has become entrenched in popular culture to a degree far greater than almost any other mathematical construct. It is, perhaps, the most common ground between mathematicians and non-mathematicians. Reports on the latest, most-precise calculation of π (and related stunts) are common news items [2, 4, 5, 7, 12, 13]. The decimal representation of π truncated to 50 decimal places is: π = 3.14159265358979323846264338327950288419716939937510... . Various online web sites provide π to many more digits [20, 21, 22, 23, 26]. While the decimal representation of π has been computed to more than a trillion (1012) digits [18], elementary applications, such as estimating the circumference of a circle, will rarely require more than a dozen decimal places. For example, the decimal representation of π truncated to 11 decimal places is good enough to estimate the circumference of any circle that fits inside the earth with an error of less than one millimeter, and the decimal representation of π truncated to 39 decimal places is sufficient to estimate the circumference of any circle that fits in the observable universe with precision comparable to the radius of a hydrogen atom. Because π is an irrational number, its decimal representation does not repeat, and therefore does not terminate [10, 14, 17]. This sequence of non-repeating digits has fascinated mathematicians and laymen alike, and much effort over the last few centuries has been put into computing ever more of these digits and investigating π's properties. Despite much analytical work, and supercomputer calculations that have determined over 1 trillion digits of the decimal representation of π, no simple base-10 pattern in the digits has ever been found. Digits of the decimal representation of π are available on many web pages, and there is software for calculating the decimal representation of π to billions of digits on any personal computer [2, 11, 20, 21, 23, 26]. The current record for the decimal expansion of π, if verified, stands at 5 trillion digits. This statistical survey stands at 40960 digits. It is not known if is normal [6], although the first 30 million digits are very uniformly distributed [4].

This paper is organized as follows:

In Section 2, the frequency table of decimal digits for the first 40960 digits of $\pi - 3$ is introduced. In Section 3, the randomness test of these digits is presented. In Section 4, the goodness of fit test (versus uniform distribution) is done. The entropy of digits distribution is obtained in Section 5. In Section 6, the one step transition frequency matrix and one step transition probability matrix are shown. Finally, the conclusion is drawn in Section 7.

2. Frequency Table

The following frequency distribution of decimal digits is found for the first 40960 digits of $\pi - 3$. It shows no statistically significant departure from a uniform distribution.

Table 1. Frequency table

Digits	Frequency	Percent
0	4086	10
1	4160	10.2
2	3971	9.7
3	4067	9.9
4	4102	10
5	4117	10.1
6	4120	10.1
7	4085	10
8	4148	10.1
9	4104	10
Total	40960	100

Namely, distribution of all digits is uniform.

3. Randomness Test

The runs test procedure tests whether the order of occurrence of values of a variable is random. A run is a sequence of like observations. A sample with too many or too few runs suggests that the sample is not random. For more details, see [15, 19]. The following table summarizes results of run test:

Table 2. Runs test

	Digits
Test value (median)	5
Cases $<$ test value	20386
Cases $\geq$ test value	20574
Total cases	40960
Number of runs	20582
Z	1.002
Asymp. sig. (2-tailed)	0.316

So by sig (0.316), we do not reject the hypothesis of randomness, i.e., the all digits $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ distributed randomly in the first 40960 digits of $\pi - 3$.

4. Goodness of Fit Test

The chi-square test procedure tabulates a variable into categories and computes a chi-square statistic [15, 19]. This goodness-of-fit test compares the observed and expected frequencies in each category to test that all categories contain the same proportion of values or test that each category contains a user-specified proportion of values. The chi-square test table for our data is shown in the following table:

Table 3. Results for goodness of fit test

Digits	Observed N	Expected N	Residual
0	4086	4096	– 10
1	4160	4096	64
2	3971	4096	– 125
3	4067	4096	– 29
4	4102	4096	6
5	4117	4096	21
6	4120	4096	24
7	4085	4096	– 11
8	4148	4096	52
9	4104	4096	8
Total	40960	40960	

Chi-square test statistic is 6.007 with $\text{df} = 9$ and asymp. sig = 0.739. Therefore, we cannot reject the hypothesis of uniform distribution of digits.

5. Entropy of Digits Distribution

The entropy of a random variable is defined in terms of its probability distribution and can be shown to be a good measure of randomness or uncertainty. Let X be a random variable with probability mass function

$$p(x) = \Pr(X = x), \quad x \in \mathbb{A}.$$

Definition 5.1. The Shannon's entropy [25, 27] $H(X)$ of random variable X is defined by

$$H(X) = -\sum_{x \in \mathbb{A}} p(x) \log(p(x)). \tag{5.1}$$

So, the entropy of Pi-number decimal digits distribution is given by

$$H(X) \simeq -0.1 \log_2 0.1 - 0.1 \log_2 0.1 - \ldots - 0.1 \log_2 0.1 = 3.3219, \tag{5.2}$$

where is the maximum entropy between all discrete distributions, i.e., the diffusion of decimal digits of Pi-number have the most disorder.

6. One Step Transition Matrices

Now, look at the following table that show number of one step transition from one digit to another digit only in one step.

Table 4. One step transition frequency matrix

States	0	1	2	3	4	5	6	7	8	9	Total
0	400	428	390	406	400	402	400	425	397	438	4086
1	428	431	383	413	434	397	428	387	434	425	4160
2	383	419	394	411	366	413	422	402	390	371	3971
3	392	412	398	377	432	429	379	408	398	442	4067
4	400	419	403	411	365	413	427	432	416	416	4102
5	410	384	408	422	430	412	418	427	415	391	4117
6	418	410	422	402	419	422	398	404	397	428	4120
7	422	444	390	421	378	376	428	404	435	387	4085
8	405	404	399	402	432	442	406	420	433	405	4148
9	428	409	384	402	446	411	414	376	433	401	4104
Total	4086	4160	3971	4067	4102	4117	4120	4085	4148	4104	40960

Look at marginal sums. With deferent component, the corresponding sum is equal. (No wonder!)

The following table shows one step ratio frequency matrix between digits, obtained from above table.

Table 5. One step transition probability matrix

States	0	1	2	3	4	5	6	7	8	9	Total
0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
4	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
5	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
6	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
7	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
8	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
9	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
Total	1	1	1	1	1	1	1	1	1	1	10

We see that, the probability of transmission from specific digit to next digit in one step, is equal to 0.1, i.e., the transmission between all digits in one step is completely random. Or from specific digit, the next digit is selected from $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ uniformly. This endorsement the runs test for randomness.

7. Conclusion

In this article, we investigated some statistical properties of the first 40960 digits of $\pi - 3$. Based on these results, we saw that all digits $\{0, 1, 2, \ldots, 9\}$ distributed uniformly in 40960 decimal digits. Also, by the one step transition probability matrix, the probability of transmission from specific digit to next digit in one step, obtained 0.1, i.e., the transmission between all digits in one step is completely random and we can conclude that since, these digits do not obey from any predictable statistical model, the prediction (with high confidence) of next digit in decimal digits of Pi-number is impossible.

References

[1] V. Adamchik and S. Wagon, A simple formula for π, Amer. Math. Monthly 104 (1997), 852-855.

[2] D. Anderson, The Pi-Search Page.

http://www.angio.net/pi/piquery

[3] J. Arndt and C. Haenel, Pi-Unleashed, 2nd Edition, Springer-Verlag, Berlin, 2001.

[4] D. H. Bailey, The computation of decimal digit using Borwein's quartically convergent algorithm, Math. Comput. 50 (1988), 283-296.

[5] D. H. Bailey, P. B. Borwein and S. Plouffe, On the rapid computation of various polylogarithmic constants, Math. Comput. 66 (1997), 903-913.

[6] D. H. Bailey and R. E. Crandall, On the random character of fundamental constant expansions, Exper. Math. 10 (2001), 175-190.

http://www.nersc.gov/ dhbailey/dhbpapers/baicran.pdf

[7] D. H. Bailey, J. M. Borwein, N. J. Calkin, R. Girgensohn, D. R. Luke and V. H. Moll, Experimental Mathematics in Action, Wellesley, MA: A K Peters, 2007.

[8] Petr Beckmann, A History of Pi, Barnes and Noble Publishing, ISBN 0880294183, 1989.

[9] Ari. Ben-Menahem, Historical Encyclopedia of Natural and Mathematical Sciences: Jones was first to use? for the ratio (perimeter/diameter) of a circle, in 1706, 2009.

[10] Lennart Berggren, Jonathan M. Borwein and Peter B. Borwein (eds.), Pi: A Source Book, Springer, 1999 (2nd Ed.), ISBN 978-0-387-98946-4.

[11] J. M. Borwein, P. B. Borwein and D. H. Bailey, Ramanujan, modular equations, and approximations to Pi, or how to compute one billion digits of Pi, Amer. Math. Monthly 96 (1989), 201-219.

[12] J. M. Borwein and D. H. Bailey, Mathematics by Experiment: Plausible Reasoning in the 21st Century, Wellesley, MA: A K Peters, 2003.

[13] J. M. Borwein, Talking About Pi.

http://www.cecm.sfu.ca/personal/jborwein/picover.html

[14] C. K. Caldwell and H. Dubner, Primes in Pi, J. Recr. Math. 29 (1998), 282-289.

[15] G. W. Corder and D. I. Foreman, Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach, Wiley, New Jersey, 2009.

[16] T. Cover and J. A. Thomas, Elements of Information Theory, Wiley, New York, 2006.

[17] W. W. Gibbs, A digital slice of Pi, The new way to do pure math: Experimentally, Sci. Amer. 288 (2003), 23-24.

[18] X. Gourdon and P. Sebah, PiFast: The Fastest Program to Compute Pi.

http://numbers.computation.free.fr/Constants/PiProgram/pifast.html

[19] K. Gouri, R. Bhattacharyya and A. Johnson, Statistical Concepts and Methods, John Wiley and Sons, 1977.

[20] http://www.pi314.net/eng/ramanujan.php

[21] http://numbers.computation.free.fr/Constants/Pi/pi.html

[22] Y. Kanada, New World Record of Pi: 51.5 Billion Decimal Digits

http://www.cecm.sfu.ca/personal/jborwein/Kanada50b.html

[23] Pi Calculated to Record Number of Digits, bbc.co.uk. 2010-01-06. [Retrieved 2010-01-06].

http://news.bbc.co.uk/1/hi/technology/8442255.stm

[24] Alfred S. Posamentier and Ingmar Lehmann, Pi: A Biography of the World's Most Mysterious Number, Prometheus Books, 2004, ISBN 978-1-59102-200-8.

[25] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (379-423) (1984), 623-656.

[26] H. J. Smith, Computing Pi.

http://www.geocities.com/hjsmithh/Pi.html

[27] S. Wagon, Is normal? Math. Intel. 7 (1985), 65-67.