ON THE RELIABILITY OF RANDOM NUMBER GENERATORS

Yadolah Dodge and Giuseppe Melfi
Université de Neuchâtel
Groupe de Statistique
CP 805 CH-2002 Neuchâtel
Switzerland
Yadolah.Dodge@unine.ch
Giuseppe.Melfi@unine.ch

KEYWORDS

Simulation, Random and pseudorandom number generators, Normal numbers, π, log 2, Champernowne constants, Tests of randomness, Autocorrelation function.

ABSTRACT

In this paper we discuss major problems related to reliability of random number generators used for simulation studies. We propose the decimals of π as the most reliable random number generator as compared to certain real normal numbers as well as all families of pseudorandom number generators. A new property that a random number generator must have is introduced. Two applications are discussed.

INTRODUCTION

Every simulation study always uses random number generators in order to construct sufficiently large samples. So, in amount, random number generators are a central tool, and in despite of their importance, this topic is not sufficiently studied in modern research. As Coveyou remarked (Coveyou 1969), "random number generation is too important to be left to chance".

The history of random number generators begin with the age of computer. In fact, apart from manually generating random numbers, e.g. by a dice throwing, all random number generators use an algorithm and are called "pseudo-random number generators". So it cannot generate "truly random numbers" because of their deterministic nature.

The aim of this work is to answer to some fundamental questions such as for example: what is really a random number generator? How many properties it must satisfy? Are random number generators sufficiently reliable? What is the state-of-the-art of ideal random number generators? Are normal numbers, as a new class whose normality has been recently proved (Bailey and Crandall 2001), or Champernowne constants (Champernowne 1933) suitable for generating good random numbers?

PSEUDO RANDOM NUMBER GENERATORS

Since the '50's, when the early electronic devices appear to be suitable to be applied to generate random sequences of numbers, many techniques have been proposed. The first method of a certain importance was the middle-square method (Von Neumann 1951). The idea was to take a number n_0 composed by a certain number d of digits (in a given base); the square n_0^2 has twice the number of digits of n_0, so the d central digits of n_0^2 provide a sequence of digits apparenty random. A relation between n_0 and the new sequence of d digits is not trivial, and the iteration of this process provided for a certain time a first satisfactory example of "random number generator". Obviously this method is unable to provide "truly random random numbers" for at least two reasons. The first is that the method is deterministic, and once the hidden algorithm revealed, the sequence is completely previsible. The second reason is that this method only gives periodic sequences.

Nowdays, the most sure method for generating random sequence of numbers is to mixt different pseudorandom number generators, with a special attention in keeping secret the procedure! When a procedure is revealed, a test of randomness can be provided in order that the random number generator fails it.

NORMAL NUMBERS AND IDEAL RANDOM NUMBER GENERATORS

The most current accepted idea of truly random number generator, is provided by normal numbers. A normal number on base b is a real number whose expansion in base b contains each digits strings $d_1 d_2 d_3 \dots d_k$ with the expected frequency $1/b^k$. A normal number *tout court* is a normal number on any base. The definition of normal number corresponds to the definition of ∞-distributed sequences, whose properties have been extensively studied (Knuth 1981). In particular all real numbers, with the exception of a set of measure zero, are normal numbers.

For this reason, the number π is often considered as an ideal random number generator (Dodge 1996). A proof of its normality is not yet available. Recent studies on π (Bailey et al. 1997) have shown some remarkable identities for π as the following:

$$\pi = \sum_{i=0}^{\infty} \frac{1}{16^i}\left(\frac{4}{8i+1} - \frac{2}{8i+4} - \frac{1}{8i+5} - \frac{1}{8i+6}\right).$$

The above identity is of the same kind of the following one for log 2:

$$\log 2 = \sum_{i=1}^{\infty}\frac{1}{i2^i}$$

Bailey and Crandall proved that the number, whose addends are a subset of the terms of the preceding series

$$\alpha_{2,3} = \sum_{i=3,3^2,3^3,\ldots}\frac{1}{i2^i}$$

is normal (Bailey and Crandall 2002). So a proof of the normality of π is perhaps not so far as it was until some years ago. Other relations between π, other real numbers and normality can be found in (Bailey and Crandall 2001). The digits of π satistfy all possible test of randomness (Dodge and Rousson, 1996; Murier and Rousson, 1998). The authors (Dodge and Melfi, to appear) recently investigated 1'000'000 coefficients of the continued fraction of π and have shown that the distribution of these coefficients is in accord with a Khinchin random variable, as for almost all real numbers.

The digits of π have been used as test of efficiency for computers. Modern computation of π is based on algorithms introduced by Brent and Salamin, and developped by Borwein and Borwein (Brent 1976; Salamin 1976; Borwein and Borwein 1984). A new computer architecture is judged more efficient when it calculates the digits of π more rapidly than another. For this reason more than 200 billions digits are known today, so an extensive use of its digits as table of random numbers, e.g. on a CD-ROM or DVD-ROM support is proposed.

RELIABILITY OF A PSEUDO RANDOM NUMBER GENERATOR

Knuth provides a collection of properties that a sequence of numbers must satisfy in order to be classified as "truly random" (Knuth 1981). In this section we propose a property *P* that a sequence must satisfy in addition to those of Knuth. We will show that a real number may be normal, but not satisfying this property of randomness.

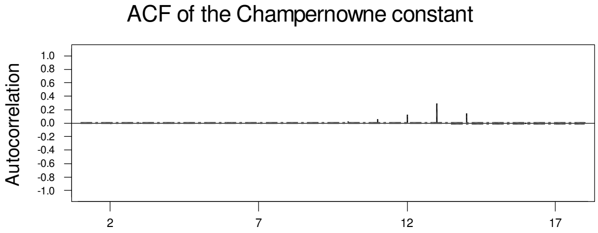
Definition. *We will say that an infinite sequence of digits satisfy the property P if, whenever we cut the sequence at a finite term, digits does not show significant autocorrelations.*

As a significant autocorrelation, a standard choice is an autocorrelation whose *t*-statistic exceeds 2. So a good pseudo-random number generator must provide sequences satisfying property *P*. As is usual in such a kind of definition, almost all real numbers have digits satisfying property *P*, and we conjecture that π also has a digit expansion satisfying property *P*.

On the other hand, the Champernowne constant is an exemple of a normal number, whose digits does not satisfy the property *P*. The Champernowne constant (Champernowne 1933) on base 2 is defined as

$$0.11011100101110...$$

i.e. by concatenating the sequence of natural integers in base 2. Analogously one can define the Champernowne constant for any base *b*. It is always a normal number, but if we consider a finite subsequence of digits, for example, the concatenation of binary expression of numbers 1,2,…, 10000, a strong autocorrelation appear at lags 10,11,12,13,14=$\log_2 10000$, with values of *t*-statistics respectively 2.22, 8.02, 20.32, 41.53, 101.94, 46.67 (See Figure 1). Note that autocorrelation vanishes when the whole infinite sequence of digits is considered. This phenomenon is not surprising, because the digits of a Champernowne constant are locally almost periodic. This is also due to the artificial nature of the definition.

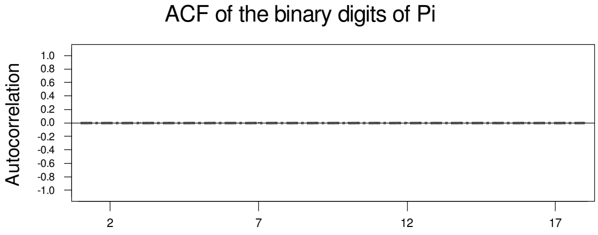


Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.00	-1.31	1.71	8	0.00	0.75	32.13	15	-0.00	-1.21	1.6E+04
2	-0.01	-5.20	28.76	9	0.01	2.22	37.06	16	-0.01	-3.31	1.6E+04
3	-0.00	-0.14	28.78	10	0.02	8.02	101.42	17	-0.00	-0.10	1.6E+04
4	-0.00	-0.79	29.41	11	0.06	20.32	515.01	18	0.01	2.51	1.6E+04
5	-0.00	-1.37	31.29	12	0.12	41.53	2254.22				
6	-0.00	-0.52	31.56	13	0.30	101.94	1.3E+04				
7	-0.00	-0.06	31.56	14	0.15	46.67	1.6E+04				

Figure 1. Autocorrelation function of the first 123'632 digits of the binary Champernowne constant.

One can easily argue that a "random" number in interval (0,1) is of such a nature with probability 0. In particular, we conjecture that π satisfy property *P*, and therefore that the digits of π are "truly random".

In Figure 2 one can see an analysis of the autocorrelation function, similar to that of Figure 1, applied to binary digits of π.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.00	0.55	0.30	8	0.01	1.15	3.70	15	0.00	0.48	11.64
2	0.00	0.08	0.30	9	-0.00	-0.47	3.92	16	-0.01	-1.67	14.42
3	-0.00	-0.25	0.37	10	0.01	0.95	4.82	17	0.01	1.47	16.60
4	0.01	1.23	1.87	11	0.01	1.26	6.42	18	-0.01	-1.42	18.63
5	-0.00	-0.26	1.94	12	-0.00	-0.32	6.52				
6	0.00	0.25	2.00	13	0.01	1.78	9.71				
7	0.00	0.61	2.38	14	0.01	1.30	11.41				

Figure 2. Autocorrelation function of the first 32'770 binary digits of π.

As is has been conjectured, there are no significant values of the autocorrelation function.

REFERENCES

Bailey, D.H. 1988. “The computation of π to 29'360'000 decimal digits using Borweins' quartically convergent algorithm”, Math. Comput. 50, 283-296.

Bailey, D.H., Borwein, P. and Plouffe, S. 1997. “On the rapid computation of various polylogarithmic constants”, Math. Comp. 66, 903-913.

Bailey, D.H. and Crandall, R.E. 2001. "On the Random Character of Fundamental Constant Expansions", Experimental Mathematics 10, no. 2 (June 2001), 175-190

Bailey, D.H. and Crandall, R.E. 2002. “Random Generators and Normal Numbers”, to appear. See also http://www.nersc.gov/~dhbailey/dhbpapers/bcnormal.ps

Borwein, J.M. and Borwein, P.B. 1984. “The arithmetic-geometric mean and fast computation of elementary functions”, SIAM Review 26, 351-365.

Borwein, J.M. and Borwein, P.B. 1987. *Pi and the AGM*, John Wiley, New York.

Brent, R.P. 1976. “Fast multiple-precision evaluation of elementary functions”, J. ACM 23, 242-251.

Champernowne, D.G. 1933. “The Construction of Decimals Normal in the Scale of Ten.” J. London Math. Soc. 8, 254-260.

Coveyou, R.R. 1969. “Random number generation is too important to be left to chance”, Appl. Math., 3, 70-111.

Dodge, Y. 1996. “A natural random number generator”, International Statistical Review 64, 329-344.

Dodge, Y. and Melfi, G., “Rare events in the continued fraction of π ” to appear.

Dodge, Y. and Rousson, V. 1996. “Does π satisfy all statistical tests?”, technical report, University of Neuchâtel.

Knuth, D. 1981. “The art of computer programming, vol.2: Seminumerical algorithms”, Addison-Wesley, Reading, MA.

Murier, T. and Rousson, V. 1998. “On the randomness of the decimal of π ”, Student 2, No. 3, 237-246.

von Neumann, J. 1951. “Various techniques used in connection with random digits”, In the Monte Carlo Method (ed. A.S. Householder et al., 36-38). Nat. Bur. Standards Appl. Math. Ser. no. 12.

Salamin, E. 1976. “Computation of π using arithmetic-geometric mean”, Math. Comp. 30, 565-570.

YADOLAH DODGE obtained his master in applied Statistics from the Utah State University in 1970 and a Ph.D in Statistics with minor in Biometry from the Oregon State University in 1973. He is currently professor of Statistics and Operations Research at the University of Neuchâtel in Switzerland.

GIUSEPPE MELFI was born in Uznach, Switzerland, in 1967. He studied mathematics at the Scuola Normale and at the University of Pisa, Italy. In 1997 he obtained his PhD degree. After three years spent in Lausanne, he moved in Neuchâtel, where at present is Research Assistant.