On a Periodic Property of Pseudo-Random Sequences

EVE BOFINGER AND V. J. BOFINGER

North Carolina State College, Raleigh, N.C., and N.S.W. University of Technology, Sydney, Australia

The sequence u_i ($i = 1, 2, \cdots$) formed by taking the principal remainders modulo n of a^i, where n and a are relatively prime positive integers, may be shown to be periodic of period δ, where δ is the smallest positive integer satisfying

$$a^\delta \equiv 1(\text{mod } n).$$

When δ is defined in this way, a is said to belong to the exponent δ modulo n.

It has been suggested by Lehmer [7] that, provided δ is reasonably large, the numbers $u_i n^{-1}$ may be used as uniform variates in the range $(0 \rightarrow 1)$.

In section 1 we shall give a general method for evaluating δ and in section 2 the results of some of the well-known tests for randomness performed on digits generated by this multiplicative congruence method when the multiplier a is chosen to be 3^{19} in order to give a sequence of maximum period for $n = 10^s$.

1. *Evaluation of δ*

Juncosa [4] has discussed the problem of choosing a so that δ is a maximum for $n = 10^s$ and Moshman [8] has calculated δ for $a = 7^{4k+1}$ and $n = 10^s$. Lehmer showed that, when $n = 10^8 + 1$, δ is a maximum if $a = 23$.

The following definitions and theorems I to VI are well known (see, for example, Nagell [9]). Theorems VII, VIII and IX may easily be proved, or references to them may be found in Dickson [2].

In the following discussion, unless otherwise stated, all numbers considered are positive integers.

DEFINITION: Euler's ϕ-function, $\phi(n)$, is defined as the number of positive integers, including 1, less than and relatively prime to n.

THEOREM I: *If n has as distinct prime factors only $p_1, p_2, \cdots, p_r$, then*

$$\phi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)(\cdots\cdots)\left(1 - \frac{1}{p_r}\right).$$

THEOREM II: *If a is relatively prime to n, then*

$$a^{\phi(n)} \equiv 1(\text{mod } n).$$

THEOREM III: *If $a^x \equiv 1(\text{mod } n)$, then x is a multiple of δ.*

COROLLARY: *δ is a divisor of $\phi(n)$.*

DEFINITION: If $\delta = \phi(n)$, then a is called a primitive root of n.

THEOREM IV: *Let n be greater than 1, and a be prime to n. If a belongs to the exponent δ modulo n and if the highest common factor of m and δ equals μ, then a^m belongs to the exponent $\frac{\alpha}{\mu}$ modulo n.*

THEOREM V: *The number n has primitive roots if and only if n can be expressed in one of the forms 2, 4, p^s, $2p^s$, where p is an odd prime.*

THEOREM VI: *If a is a primitive root of the odd prime p, and if $a^{p-1} - 1$ is not divisible by p^2, then a is a primitive root of p^s.*

Let $\delta(n, a)$ denote the exponent modulo n to which a belongs.

THEOREM VII: *If p is an odd prime and a is prime to p and if r is such that p^r is the largest power of p which divides $a^{\delta(p,a)} - 1$, then*

$$\delta(p^s, a) = \begin{cases} p^{s-r}\delta(p, a) & \text{if } s > r, \\ \delta(p, a) & \text{if } s \leqq r. \end{cases}$$

Notice that theorem VI is simply a particular case of theorem VII.

THEOREM VIII: *Let r be greater than 1.*

(i) *If a is congruent to $+1$ modulo 2^r but not 2^{r+1}, then*

$$\delta(2^s, a) = \begin{cases} 2^{s-r} & \text{if } s > r, \\ 1 & \text{if } s \leqq r. \end{cases}$$

(ii) *If a is congruent to -1 modulo 2^r but not 2^{r+1}, then*

$$\delta(2^s, a) = \begin{cases} 2^{s-r} & \text{if } s > r, \\ 1 & \text{if } s = 1, \\ 2 & \text{if } 1 < s \leqq r. \end{cases}$$

THEOREM IX: *If $n = n_1 \cdot n_2$ where n_1 and n_2 are relatively prime, then*

$$\delta(n, a) = \text{l.c.m. of } \delta(n_1, a) \text{ and } \delta(n_2, a).$$

Now if $n = 2^s$ (a convenient choice for a binary machine), theorem VIII gives the period of the sequence immediately.

When $n = 10^s$, δ is a divisor of $4 \cdot 10^{s-1}$, by theorems I and II, and if $s > 1$, 10^s has no primitive roots so that $\delta < 4 \cdot 10^{s-1}$.

Now, by theorem IX, $\delta(10^s, a) = $ l.c.m. of $\delta(2^s, a)$ and $\delta(5^s, a)$, and if 5^r is the largest power of 5 dividing $a^{\delta(5,a)} - 1$, and 2^t is the largest power of 2 dividing $(a \pm 1)$, and $s > r, t$, then

$$\delta(10^s, a) = \text{l.c.m. of } 2^{s-t} \text{ and } 5^{s-r}\delta(5, a).$$

Since $\delta(5, a)$ is a divisor of 4, put $\delta(5, a) = 2^\gamma$ where $\gamma = 0, 1, 2$.

$$\delta(10^s, a) = \begin{cases} 5^{s-r} \cdot 2^{s-t} & \text{if } s - t \geqq \gamma, \\ 5^{s-r} \cdot 2^\gamma & \text{if } s - t < \gamma. \end{cases}$$

If we choose $a = 7$, then $\gamma = 2$ (since $\delta(5, 7) = 4$) and $r = 2$ and $t = 3$ so that

$$\delta(10^s, 7) = \begin{cases} 5 \cdot 10^{s-3} & \text{if } s \geqq 5, \\ 100 & \text{if } s = 4, \end{cases}$$

and by using theorems VII and VIII for $s \leqq r, t$,

$$\delta(10^s, 7) = \begin{cases} 20 & \text{if } s = 3, \\ 4 & \text{if } s = 1, 2. \end{cases}$$

Now by theorem IV,

$$\delta(10^s, 7^5) = \begin{cases} 4 & \text{if } s = 1, 2, 3, \\ 20 & \text{if } s = 4, \\ 10^{s-3} & \text{if } s \geqq 5, \end{cases}$$

which disagrees with Moshman's [8] result. The reason for this is as follows.

Moshman states that if 5^q is the highest power of 5 dividing $(7^4)^{4k+1} - 1$ then q is given by $q = 2 + \left[\frac{k}{6}\right]$, where $[t]$ is the greatest integer less than or equal to t. This can be seen not to hold for the particular case $k = 1$.

The correct expression for q is a particular case of the following theorem.

THEOREM X: *If the highest power of p, an odd prime, dividing $x - 1$ is p^ξ, and the highest dividing y is p^η, where x, y and ξ are positive integers and η is a positive integer or zero, then the highest power of p dividing $x^y - 1$ is $p^{\xi+\eta}$.*

PROOF: Put $x - 1 = \alpha p^\xi$ and $y = \beta p^\eta$, where α, β are not divisible by p. Now $x^y - 1 = (1 + \alpha p^\xi)^{\beta p^\eta} - 1 = \alpha\beta p^{\xi+\eta} +$ terms of higher order in p.

This also holds for $p = 2$ if $\xi > 1$. Thus $q = 2 + \eta$ where 5^η is the highest power of 5 dividing $4k + 1$.

Now the maximum value of $\delta(10^s, a)$ is given by

$$\begin{cases} 5 \cdot 10^{s-2} & \text{if } s \geqq 4, \\ 100 & \text{if } s = 3, \\ 20 & \text{if } s = 2, \\ 4 & \text{if } s = 1. \end{cases}$$

For $\delta(10^s, a)$ to attain these values, a must be chosen so that $\delta(5, a) = 4$, $r = 1$ and $t = 2$. That is, a is a primitive root of 5 and such that $a^4 - 1$ is not divisible by 5^2 and neither $a + 1$ nor $a - 1$ is divisible by 8. A possible value for a is 3.

Notice that the period of individual digits may be found by considering appropriate values of s. For example the maximum period of the least significant digit is 4.

If we are interested only in values of $s \geqq 4$ (that is, we discard the least significant digits because of their shorter periods), we can attain the maximum period of $5 \cdot 10^{s-2}$ by choosing a so that $r = 1$ and $t = 2$ and $\delta(5, a)$ may be 1 or 2. A possible value for a is then 11.

Also by theorem IV, if a is chosen to give the maximum period then a^m, where m is relatively prime to 10, will also give this maximum period.

To start at a different part of the sequence a^i $(i = 1, 2, \cdots)$ we may choose i, in some random manner, and the principal remainder modulo 10^s of a^i may be evaluated on a desk calculator. Or we may generate numbers u_i $(i = 0, 1, 2, \cdots)$ which are the principal remainders modulo 10^s of a^ib where b is randomly chosen to be relatively prime to 10. Depending on the value of b, we obtain numbers from some part of eight sequences, each having a period of $5 \cdot 10^{s-2}$, which exhaust the $4 \cdot 10^{s-1}$ numbers less than and relatively prime to 10^s.

If d_i is the ith digit of the number $a^r \cdot b$, and the period of the number $\sum_{i=1}^{m} d_i 10^{i-1}$ (that is, the number consisting of the first m digits of $a^r \cdot b$) is denoted by δ, then provided $m \geqq \gamma + t$ (where γ and t are as defined in above) the period of the number $\sum_{i=1}^{m+n} d_i 10^{i-1}$ is $10^n \cdot \delta$ when $n \geqq 1$.

Hence each one of the 10^n possible values of $\sum_{i=m+1}^{m+n} d_i 10^{i-1}$ must occur with all δ possible values of $\sum_{i=1}^{m} d_i 10^{i-1}$. This means that when using

$$10^{-(m+n)} \sum_{i=m+1}^{m+n} d_i 10^{i-1}$$

as a random number we know that all possible numbers occur with equal frequencies in a complete period.

2. *Tests for Randomness*

24,000 pseudo-random numbers were generated using a multiplier, a, equal to 3^{19} and taking s equal to 20. For the first 4,000, b was chosen to be 1 and for succeeding groups of 4,000, b was chosen each time to be a random 20-digit number (using the Rand tables [9]). From each group of 1,000 20-digit numbers, 10 sequences of 1,000 digits were obtained, each sequence corresponding to a particular digit position from the 10 most significant.

These sequences were subjected to the frequency, serial and gap tests as described by Kendall and Babington-Smith [5]. The serial test was modified in the following way as suggested by Good [3].

Let n_i be the number of digits equal to i and $\bar{n}_{ij}$ be the number of (ij) sequences in a random cyclic sequence of length N.

Let

$$\psi_1^2 = \sum_{i=0}^{9} \frac{(n_i - N \cdot 10^{-1})^2}{N \cdot 10^{-1}}$$

and

$$\bar{\psi}_2^2 = \sum_{(ij)} \frac{(\bar{n}_{ij} - N \cdot 10^{-2})^2}{N \cdot 10^{-2}}$$

where (ij) runs through its 10^2 possible values.

Now ψ_1^2 has asymptotically a χ_9^2 distribution (a chi-squared distribution with 9 degrees of freedom) and $\bar{\psi}_2^2$ has been considered to have asymptotically a χ_{90}^2 distribution. However, Good [3] has shown that the expected value of $\bar{\psi}_2^2$ is 99 and that it is reasonable to expect that $\nabla\bar{\psi}_2^2 = \bar{\psi}_2^2 - \bar{\psi}_1^2$ has asymptotically a χ_{90}^2 distribution. Billingsley [1] has found the asymptotic distribution of $\bar{\psi}_2^2$ and we have shown (in work to be published) that, in fact, $\nabla\bar{\psi}_2^2$ has asymptotically a χ_{90}^2 distribution. We have used $\nabla\bar{\psi}_2^2$ for the test statistic of the serial test.

To test whether the numbers $u_i 10^{-20}$ might be used as uniform variates, the frequencies with which these numbers fell in the classes $r \cdot 10^{-2}$ to $(r + 1) \cdot 10^{-2}$ where $r = 0, 1, \cdots, 99$ were found and a goodness-of-fit test was carried out for each 1,000 numbers.

The results of these tests were as follows.

Frequency test. Of the 240 values of ψ_1^2, seven were found to be greater than 16.919, the 5 per cent critical value of a χ_9^2 distribution.

Serial test. Of the 240 values of $\nabla\bar{\psi}_2^2$, nine were found to be greater than 113.14, the 5 per cent critical value of a χ_{90}^2 distribution (using the Wilson and Hilferty approximation [6]).

A χ_{11}^2 test, used to examine the goodness of fit of the $\nabla\bar{\psi}_2^2$ values obtained to a χ_{90}^2 distribution, yielded a value of 12.315.

Gap test. The numbers of gaps between zeros of sizes 0 to 99 were found. Some of these 100 classes were pooled so that the expected frequency in each class should be at least 1. The 240 χ_{29}^2 goodness-of-fit tests then yielded nine values greater than the 5 per cent critical value of 42.577.

Uniform test. The 24 χ_{99}^2 goodness-of-fit tests yielded two values greater than the 5 per cent critical value of 123.22.

Acknowledgements

We are grateful to Dr. A. Grandage for suggesting this problem, and to Dr. E. J. Williams for his helpful comments.

REFERENCES

1. P. Billingsley, Asymptotic distributions of two goodness of fit criteria, *Ann. Math. Stat.* **27** (1956), 1123–1129.
2. L. E. Dickson, *History of the Theory of Numbers*, Vol. 1, Chelsea Publishing Company, New York, 1952.
3. I. J. Good, The serial test for sampling numbers and other tests for randomness, *Cambridge Philos. Soc. Proc.* **49** (1953), 276–284.
4. M. L. Juncosa, Random number generation on B.R.L. high-speed computing machines, Ballistic Research Laboratories Report No. 855, 1953, Aberdeen Proving Ground.
5. M. G. Kendall and B. Babington Smith, Randomness and random sampling numbers, *Jour. Royal Stat. Soc.* **101** (1938), 147–166.
6. M. G. Kendall, *The Advanced Theory of Statistics*, Vol. 1, (5th ed , 1952), Hafner Publishing Co., New York, p. 294.
7. D. H. Lehmer, Mathematical methods in large-scale computing units, Symposium on Large-Scale Digital Calculating Machinery 2d, Harvard University, 1949, pp. 141–146.
8. J. Moshman, The generation of pseudo-random numbers on a decimal calculator, *Jour. Assoc. Computing Machinery* **1** (1954), 88–91.
9. T. Nagell, *Introduction to Number Theory*, J. Wiley and Sons, New York, 1951.
10. Rand Corporation, *A Million Random Digits*, The Free Press, Publishers, Glencoe, Illinois.