American Open Journal of Statistics, **2011, 1, 128-138**
doi:10.4236/ojs.2011.12015 Published Online July 2011 (http://www.SciRP.org/journal/ojs)

Fibonacci Congruences and Applications

René Blacher
Laboratory LJK, Université Joseph Fourier, Grenoble, France
E-mail: *rene.blacher@aliceadsl.fr*

Received May 24, 2011; *revised June* 14, 2011; *accepted June* 20, 2011

Abstract

When we study a congruence $T(x) \equiv ax$ modulo m as pseudo random number generator, there are several means of ensuring the independence of two successive numbers. In this report, we show that the dependence depends on the continued fraction expansion of m/a. We deduce that the congruences such that m and a are two successive elements of Fibonacci sequences are those having the weakest dependence. We will use this result to obtain truly random number sequences x_n. For that purpose, we will use non-deterministic sequences y_n. They are transformed using Fibonacci congruences and we will get by this way sequences x_n. These sequences x_n admit the IID model for correct model.

Keywords: Fibonacci Sequence, Linear Congruence, Random Numbers, Dependence, Correct Models

1. Introduction

In this paper, we present a new method using Fibonacci sequences to obtain real IID sequences x_n of random numbers[1]. To have random number two methods exists : 1) use of pseudo-random generators (for example the linear congruence), 2) use of random noise (for example Rap music).

But, up to now *no completely reliable solution had been proposed* ([1]-[3]). To set straight this situation, Marsaglia has created a Cd-Rom of random numbers by using sequences of numbers provided by Rap music. But, he has not proved that the sequence obtained is really random.

However, by using Fibonacci congruence, there exists simple means of obtaining random sequences whose the quality is sure (cf [4]): one uses the same method as Marsaglia, but one transforms the obtained sequence by Fibonacci congruences. Then, one obtains sequence of real $x_n$such that the IID model is a correct model of x_n.

1.1. Fibonacci Congruence

Linear congruences $T(x) \equiv ax$ mod (m) are often used as pseudo-random generators. In this case, we try to choose a and m so that successive pseudorandom numbers behave as independent. Of course, we can only ensure that it is the case of p successive numbers p where $2^p \le m$. To choice a and m, one can use the spectral test or the results of Dieter (cf [5]) which allow to choose the best "a".

Unfortunately, the conditions which ensure the independence of three successive numbers are not those which ensure the best independence of two successive numbers,for example.

Indeed, in this paper, we will study the conditions which ensure the independence of only two successive numbers and we will see with astonishment that this is the Fibonacci congruence which provides the best empirical independence.

We shall study the set $E_2 = \left\{\ell, \overline{T(\ell)} \mid \ell \in \{0,1,\cdots,m-1\}\right\}$ when $\overline{z} \equiv z$ modulo m and $0 \le \overline{z} < m$ if $z \in \mathbb{Z}$. We will understand that this dependence depends on the continued fraction $\frac{m}{a}$, *i.e.* it depends on sequences r_n and h_n defined in the following way.

Notations 1.1 *Let* $r_0 = m$, $r_1 = a$. *One denotes by* r_n *the sequence defined by* $r_n = h_{n+1}r_{n+1} + r_{n+2}$ *the Euclidean division of* r_n *by* r_{n+1} *when* $r_{n+1} \ne 0$. *Moreover, one denotes by d the smallest integer such as* $r_{d+1} = 0$. *One sets* $r_{d+2} = 0$.

One sets $k_0 = 0$, $k_1 = 1$ and $k_{n+2} = h_{n+1}k_{n+1} + k_n$ if $n+1 \le d$.

Then, dependence depends on the h_i's: more they are small, more the dependence is weak.

Theorem 1 *Let* $(x_0, y_0) \in E_2$. *Let* $R^0 = \{[x_0, x_0 + k_n] \otimes [y_0, y_0 + r_{n-2}[\}$ *and let* $R_0 = \overline{R^0}$,

[1] By abuse of language, we will call "IID sequence" (independent identically distribution) the sequence of random numbers.

be the rectangle R^0 *modulo m. Then*

If n is even,

$E_2 \cap R_0 = \left\{\left(\overline{x_0 + k_{n-1}\ell}\,,\overline{y_0 + r_{n-1}\ell}\right) \middle| \ell = 0,1,2,\cdots,h_{n-1}\right\}$.

Moreover the points $\left(\overline{x_0 + k_{n-1}\ell}\,,\overline{y_0 + r_{n-1}\ell}\right)$ *are lined up modulo m.*

If n is odd,

$E_2 \cap R_0$

$= \left\{\left(\overline{x_0 + k_{n-2} + k_{n-1}\ell}\,,\overline{y_0 + r_{n-2} - r_{n-1}\ell}\right) \middle| \ell = 0,1,2,\cdots,h_{n-1}\right\}$.

Moreover, the points $\left(\overline{x_0 + k_{n-2} + k_{n-1}\ell}\,,\overline{y_0 + r_{n-2} - r_{n-1}\ell}\right)$ *are lined up modulo m.*

Of course, in general, it is only on the border that R_0, the rectangle modulo *m*, satisfies $R_0 \neq R^0$. If not, R_0 is a normal rectangle.

For example if $x_0 = y_0 = 0$, this theorem means that the rectangle $[0,k_n/2]\otimes[r_{n-2}/2,r_{n-2}[$ does not contain points of E_2 if *n* is even: $E_2 \cap \{[0,k_n/2]\otimes[r_{n-2}/2,r_{n-2}[\} = \varnothing$. If h_{n-1} is large, that will mean that an important rectangle of $\mathbb{R}^2$ is empty of points of E_2: that will mark a breakdown of independence.

As $h_i \geq 1$, the congruence which defines the best independence of E_2 will satisfy $h_i = 1$ and $h_d = 2$. In this case **we call it congruence of Fibonacci**. Indeed, there exists n_0 such that $a = fi_{n_0}$ and $m = fi_{n_0+1}$ where fi_n is the sequence of Fibonnacci: $fi_1 = fi_2 = 1$, $fi_{n+2} = fi_{n+1} + fi_n$. As a matter of fact the sequence r_n is the sequence of Fibonacci except for the last terms (i.e. except for $fi_1 = fi_2 = 1$). It is also the case for sequence k_n.

Remark 1.1 *In fact, when we use sequence* h_n, *we use Euclidean Algorithm. Now, Dieter has also used this algorithm to compute the dependence of* $\left(T^n(x_0),T^{n+1}(x_0)\right)$ *when* $m = 2^e$. *But he has not understood the part of the* h_i*'s in this dependence.*

1.2. Application: Building of Random Sequence

Unfortunately, congruences of Fibonacci cannot be used in order to directly generate good pseudo random sequences because $T^2 = \pm Id$ where Id is the identity (cf page 141 of [6]). Indeed in this case, the pseudo random sequence $x_n = \overline{T}^n(x_0)$ checks $x_{n+2} = \pm x_n$. However, one can use congruence of Fibonacci in order to build IID sequences by transforming some random noise y_n.

Definition 1.2 *Let* $q \in \mathbb{N}^*$. *Let T be the congruence of Fibonacci modulo m. We define the function of Fibonacci* T_q *by* $T_q = Pr_q\hat{T}$ *where*

1) $\hat{T}(x) = \overline{T}(mx)/m$,

2) $Pr_q(z) = \overline{\overline{0,b_1b_2\cdots b_q}}$ when $z = \overline{\overline{0,b_1b_2\cdots}}$ is the binary writing of *z*.

We choose $y_n \in \{0/m,1/m,\cdots,(m-1)/m\}$, $n = 1, 2,\cdots,N$. Then, y_n admits for correct model a sequence of random variables Y_n defined on a probability space $(\Omega,\mathcal{A},P)$. Then, we will impose to y_n that the conditional probabilities of Y_n admit densities with Lipschitz coefficient bounded by K_0 not too large.

In fact, since Y_n has discrete value, we can always assume that Y_n has a continuous density.

Notations 1.3 *We denote by* μ_m *be the uniform measure defined on* $\{0/m,1/m,\cdots,(m-1)/m\}$ *by* $\mu_m(k/m) = 1/m$ *for all* $k \in \{0,1,\cdots,m-1\}$.

For all permutation ϕ *of* $\{1,2,\cdots,N\}$, *for all* $n \in \{1,2,\cdots,N\}$, *we denote by* $f_{n,\phi}\left(.\middle|y_1',y_2',\cdots\right)$ *the conditional density with respect to* μ_m of $Y_{\phi(n)} = y$ given $Y_{\phi(n-1)} = y_1', Y_{\phi(n-2)} = y_2',\cdots$.

Since Y_n is discrete, we can also assume that $f_{n,\phi}\left(.\middle|y_1',y_2',\cdots\right)$ has a finite Lipschitz coefficient.

Notations 1.4 *We denote by* K_0 *a constant such that, for all permutation* ϕ *of* $\{1,2,\cdots,N\}$, *for all* $n \in \{1,2,\cdots,N\}$,
$\left|f_{n,\phi}\left(y\middle|y_1',y_2',\cdots\right) - f_{n,\phi}\left(y'\middle|y_1',y_2',\cdots\right)\right| \leq K_0|y - y'|$. *In order to simplify the proofs we suppose* $K_0 > 1$.

Now, we shall prove easily that the conditional probabilities of $T_q(Y_n)$ check

$$P\left\{T_q\left(Y_{\phi(n)}\right) = x_0 \middle| Y_{\phi(n-1)} = y_1', Y_{\phi(n-2)} = y_2',\cdots,Y_{\phi(N-1)} = y_{N-1}'\right\}$$

$$= 1/m\left[1 + O(1)K_0\,2^q/m\right]$$

Then we shall choose *m* and *q* such that $|\varepsilon| \leq K_0\,2^q/m$ is small enough. We shall deduce that, for all Borel set $Bo \subset \left\{0/2^q,1/2^q,\cdots,\left(2^q - 1\right)/2^q\right\}^N$,
$P\{(X_1,\cdots,X_N) \in Bo\} = L(Bo)\left[1 + O(1)N\varepsilon\right]$ where L is the measure corresponding to the Borel measure in the case of discrete space : $L\left(\left\{k_1/2^q,\cdots,k_N/2^N\right\}\right) = 1/2^{Nq}$.

Then $x_n = T_q(y_n)$ cannot be differentiated from an IID sequence. Indeed, it is wellknown that, for a sample x_n, there is many models correct : in particular, if x_n is extracted of an IID sequence, models such that $P\{(X_1,\cdots,X_N) \in Bo\} = L(Bo)[1 + \varepsilon]$ are correct if ε is small enough with respect to N. Reciprocally, if the sequence of random variables X_n checks $P\{(X_1,\cdots,X_N) \in Bo\} = L(Bo)\left[1 + \varepsilon_{Bo}\right]$, the model IID is also a correct model for the sequence X_n.

Thus one will be able to admit that the IID model is a correct model for the sequences x_n. As a matter of fact, one will be even able to admit that there exists another correct model $Y_n^{\theta_0}$ of y_n such that $T_q\left(Y_n^{\theta_0}\right)$ is exactly the IID sequence.

Now there exists noises y_n such that K_0 is not too large. For example these sequences can be built by using texts. In this case we can prove the result : in order that x_n is IID, it suffices that y_n admits a correct model such that K_0 is not too large. However, it is a condition which can be imposed easily by transforming some noises. The advantage compared with the CD-Rom of Marsaglia is that this result is proved. Of course, we tested such sequences.

So finally we can indeed build sequences x_n admitting for correct model the IID model by using Fibonacci congruences. This means that, a priori, these sequences x_n behave as random sequences. It is always possible that they do not satisfy certain tests. But it will be a very weak probability as we know it is the case for samples of sequences of IID random variables.

We point out that a first version of these results are in [4]. Moreover, all these results and the proofs are detailed in [7].

Note that to use the congruence of Fibonacci method is completely different from the method using Fibonacci sequence with $X_{n+1} = X_n + X_{n-1}$ modulo m, which is moreover a bad generator : cf page 27 of [1].

2. Dependence Induced by Linear Congruences

In this section, we study the set $E_2 = \left\{\ell, T(\ell) \middle| \ell \in \{0,1,\cdots,m-1\}\right\}$ when T is a congruence $T(x) \equiv ax$ modulo m.

2.1. Notations

We recall that we define sequences r_n and h_n by the following way: we set $r_0 = m$, $r_1 = a$ and $r_n = h_{n+1} r_{n+1} + r_{n+2}$, the Euclidean division of r_n by r_{n+1} when $r_{n+1} \neq 0$. One denotes by d the smallest integer such as $r_{d+1} = 0$. One sets $r_{d+2} = 0$. Moreover, $k_0 = 0$, $k_1 = 1$ and $k_{n+2} = h_{n+1} k_{n+1} + k_n$ if $n+1 \le d$.

Then $$\frac{m}{a} = h_1 + \cfrac{1}{h_2 + \cfrac{1}{h_3 + \cfrac{1}{h_4 + \cdots}}}.$$

Therefore, $h_n \ge 1$ for all $n = 1, 2, \cdots, d$ and $r_{d-1} = h_d r_d + r_{d+1} = h_d r_d + 0 = h_d r_d$. The full sequence r_n is thus the sequence $r_0 = m$, $r_1 = a$, $\cdots$, $r_{d+1} = 0$, $r_{d+2} = 0$. Then, if T is a Fibonacci conguence, r_n is the Fibonacci sequence fi_n, except for the last terms.

Remark that if $h_n = 1$ for $n = 1,2,\cdots,d-1$, k_n is also the Fibonacci sequence for $n = 1,2,\cdots$,d. Indeed by definition, $k_0 = 0$, $k_1 = 1$ and $k_{n+2} = h_{n+1} k_{n+1} + k_n$ if $n+1 \le d$.

2.2. Theorems

Now, in order to prove the theorem 1, it is enough to prove the following theorem.

Theorem 2 *Let* $n \in \{2,3,\cdots,d\}$. *Then*

If n is even,
$$E_2 \cap \{[0,k_n[\otimes[0,r_{n-2}[\} = \{(k_{n-1}\ell, r_{n-1}\ell) | \ell = 0,1,2,\cdots,h_{n-1}\}.$$
Moreover the points $(k_{n-1}\ell, r_{n-1}\ell)$ *are lined up.*

If n is odd,
$$E_2 \cap \{]0,k_n] \otimes]0,r_{n-2}]\} = \{(k_{n-2} + k_{n-1}\ell, r_{n-2} - r_{n-1}\ell) | \ell = 0,1,2,\cdots,h_{n-1}\}.$$

Moreover, the points $(k_{n-2} + k_{n-1}\ell, r_{n-2} - r_{n-1}\ell)$ *are lined up.*

Then, if there exists h_i large, there is a breakdown of independence. For example if n = 2, it is a wellknown result. Indeed, $m = r_0$, $r_1 = a$, $k_1 = 1$ and $k_2 = h_1 = \lfloor m/a \rfloor$ where $\lfloor x \rfloor$ means the integer part of x. Thus, the rectangle $Rect_2 = [0, m/(2a)] \otimes [m/2, m[$ will not contain any point of E_2. However, this rectangle has its surface equal to $m^2/(4a)$. Thus if "*a*" is not sufficiently large, i.e if h_1 is too large, there is breakdown of independence.

We confirm by graphs the previous conclusion. We suppose m = 21. If a = 13, we have a Fibonacci congruence: cf **Figure 1**. If one chooses a = 10, $sup(h_i) = 20$: cf **Figure 2** . If one chooses a = 5, $sup(h_i) = 5$: cf **Figure 3**.

Then, in order to avoid any dependence, it is necessary that $sup(h_i)$ is small.

2.3. Distribution of T([c,c'[) When *T* Is a Fibonacci Congruence

We assume that T is a Fibonacci congruence. Let $I = [c,c'[\cap \{0,1,\cdots,m-1\}$ where $c,c' \in \{0,1,\cdots,m-1\}$. We are interested by $\overline{T}^{-1}(I)$ or $\overline{T}(I)$ because $T^2 = \pm Id$. Since $\overline{T}(I)$ behaves as independent of *I*, normally, we should find that $\overline{T}(I)$ and, therefore $\overline{T}^{-1}(I)$, is well distributed in $\{0,1,\cdots,m-1\}$. As a matter of fact it is indeed the case.

Indeed, let k^n, $n = 1, 2, \cdots, c' - c$, be a permutation of $\{c, c+1, \cdots, c'-1\}$ such that $\overline{T}^{-1}(k^1) < \overline{T}^{-1}(k^2) < \overline{T}^{-1}(k^3) < \cdots < \overline{T}^{-1}(k^{c'-c})$. Then, for all numerical simulations which we executed, one has always obtained

$$\left| \overline{T}^{-1}(k^r)/m - r/N(I) \right| \le \varphi(m)/N(I)$$

where $\varphi(m) \ll Log(m)$. In fact, it seems that $\varphi(m)$ is of the order of Log(Log(m)). Moreover, $Max_{r=0,1,\cdots,N(I)-1} \left| N(I) T^{-1}(k^r)/m - r \right|$ seems maximum when I is large enough : $c' - c > m/2$.

For example, in **Figures 4**, **5** and **6**, we have the graphs $N(I) T^{-1}(k^r)/m - r$, $r = 0, 1, \cdots, N(I) - 1$ for

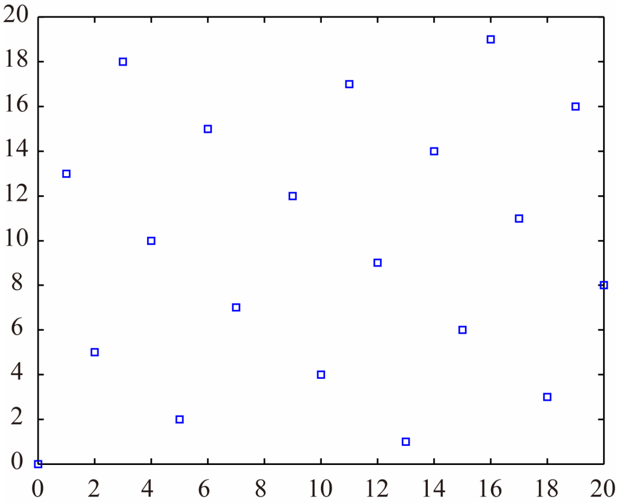

Figure 1. Fibonacci congruence.

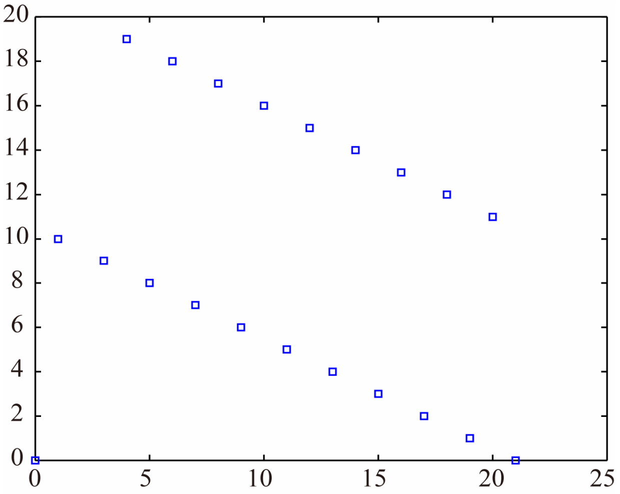

Figure 2. sup(hi) = 20.

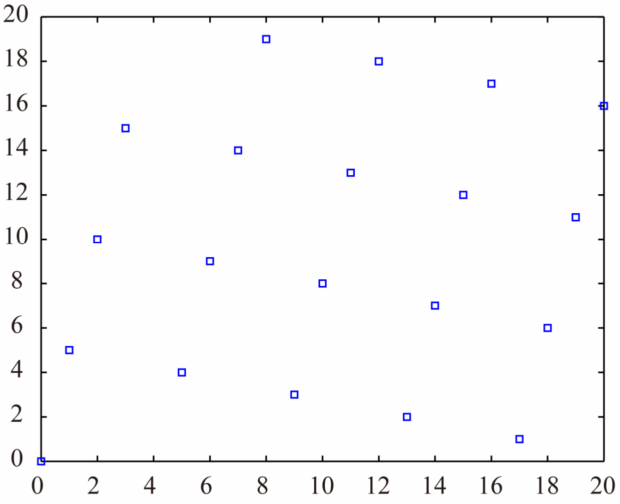

Figure 3. sup(hi) = 5.

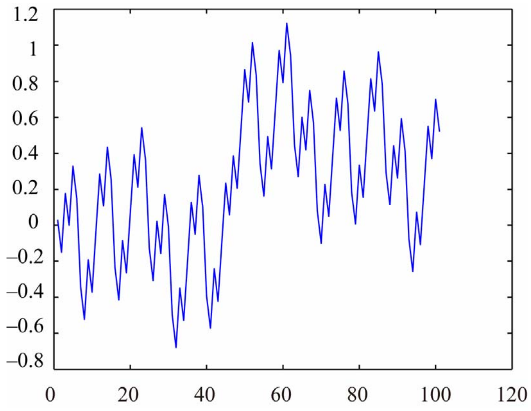

Figure 4. *a* = 1346269, *m* = 2178309.

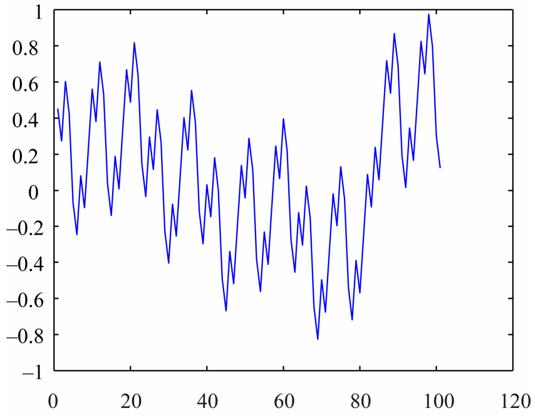

Figure 5. *a* = 121393, *m* = 196418.

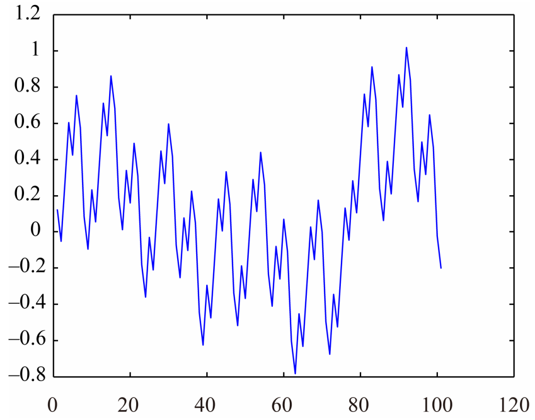

Figure 6. *a* = 10946, *m* = 17711.

various Fibonacci congruences when $c' - c = 100$.

3. Proof of Theorem 2

In this section, the congruences are conguences modulo m. Now the first lemma is obvious.

Lemma 3.1 *For* $n = 3,4,\cdots,d+1$, $k_{n+1} > k_n > k_{n-1}$. *Moreover* $k_{n+2} = h_{n+1}k_{n+1} + k_n$ *is the Euclidean division* *of* k_{n+2} *by* k_{n+1}.

Now, we prove the following results.

Lemma 3.2 *Let n* = 0, 1, 2 ,$\cdots$, *d. If n is even,* $\overline{k_n a} = m - r_n$. *If n is odd,* $\overline{k_n a} = r_n$.

Proof. We prove this lemma by recurrence. For *n* = 0, $\overline{k_n a} = \overline{0} = 0 = m - m = m - r_0$. For *n* = 1, $\overline{k_n a} = \overline{a} = a = r_1$.

We suppose that it is true for *n*.

One supposes *n* even. Then,
$k_{n+1}a \equiv ah_n k_n + ak_{n-1} \equiv -h_n r_n + r_{n-1} = r_{n+1}$.

One supposes *n* odd. Then,
$k_{n+1}a \equiv ah_n k_n + ak_{n-1} \equiv h_n r_n - r_{n-1} = -r_{n+1} \equiv m - r_{n+1}$.

Therefore, $\overline{k_{n+1}a} = m - r_{n+1}$.

Lemma 3.3 *Let n* = 2,3,$\cdots$,*d + 1. Let* $t \in \{1,2,\cdots,k_n - 1\}$. *If* $n \geq 2$ *is even,* $r_{n-1} \leq \overline{at} < m - r_n$. *If* $n \geq 3$ *is odd,* $m - r_{n-1} \geq \overline{at} > r_n$.

Moreover, if $n \geq 2$ is even, $\overline{k_n a} = m - r_n$. If $n \geq 3$ is odd, $\overline{k_n a} = r_n$.

Proof. The second assertion is lemma 3.2. Now, we prove the first assertion by recurrence.

One supposes n = 2. Then, $m = r_0 = h_1 r_1 + r_2 = h_1 a + r_2$. Moreover, $k_2 = h_1$. If $1 \leq t < h_1 = k_2$, $r_1 = a \leq at < h_1 a = m - r_2$.

If $h_1 = k_2 = 1$, $\{1,2,\cdots,k_2 - 1\} = \varnothing$. In this case, we study $t \in \{1,2,\cdots,k_3 - 1\}$ where $k_3 = h_2 k_2 + k_1 = h_2 + 1$. Then, $1 \leq t \leq h_2$. Then, $at \equiv tak_2 \equiv -tr_2$.

Moreover,

$$\begin{aligned} &m - r_2 \geq m - tr_2 \geq m - h_2 r_2 = r_0 - h_2 r_2 \\ &= r_0 - (r_1 - r_3) = r_3 + (r_0 - r_1) > r_3 \end{aligned}.$$

Therefore, because $at \equiv m - tr_2$, $\overline{at} = m - tr_2$.

Therefore, $m - r_2 \geq \overline{at} > r_3$.

One supposes that the first assertion is true for *n* where $2 \leq n \leq d$.

Let $0 < t' < k_{n+1}$. Let $t' = fk_n + e$ be the Euclidean division of *t'* by k_n: $e < k_n$.

Then, $f \leq h_n$. If not,
$t' \geq (h_n + 1)k_n + e \geq h_n k_n + k_{n-1} = k_{n+1}$.

One supposes n even.

In this case, $r_{n-1} \leq \overline{at} < m - r_n$ for $t \in \{1,2,\cdots,k_n - 1\}$.

Moreover,
$at' \equiv fak_n + ae \equiv f(m - r_n) + ae \equiv -fr_n + ae$.

First, one supposes $e = 0$. Then, $f \geq 1$.

Moreover, because $n \geq 2$,

$$\begin{aligned} &m - r_n \geq m - fr_n \geq m - h_n r_n = m - (r_{n-1} - r_{n+1}) \\ &= r_0 - r_{n-1} + r_{n+1} \geq r_0 - r_1 + r_{n+1} > r_{n+1} \end{aligned}.$$

Therefore, because $at' \equiv -fr_n$, $\overline{at'} = m - fr_n$.

Therefore, $m - r_n \geq \overline{at'} > r_{n+1}$.

Now, one supposes $f < h_n$ and $e > 0$.

By recurrence,

$$\begin{aligned} &m - r_n \geq \overline{ae} \geq \overline{ae} - fr_n \geq r_{n-1} - fr_n \geq r_{n-1} - (h_n - 1)r_n \\ &= r_n + r_{n+1} > r_{n+1} \end{aligned}.$$

Therefore, because $at' \equiv -fr_n + ae$, $\overline{at'} = \overline{ae} - fr_n$.

Therefore, $m - r_n \geq \overline{at'} > r_{n+1}$.

One supposes $f = h_n$, $e \neq k_{n-1}$ and $e > 0$.

If $e \neq k_{n-1}$, $\overline{ae} \neq \overline{k_{n-1}a}$. Indeed, if not, $\overline{a(e - k_{n-1})} = 0$. For example, if $e - k_{n-1} > 0$, $k_n > e - k_{n-1} > 0$. Then, because our recurence, $\overline{a(e - k_{n-1})} > r_{n-1} > 0$: it is impossible.

Now, if $n = 2$, $\overline{k_{n-1}a} = \overline{k_1 a} = \overline{a} = r_1 = r_{n-1}$.

Moreover, if $n > 2$, $n \geq 4$. Then, by recurence $\overline{k_{n-1}a} = r_{n-1}$.

Then, if $e \neq k_{n-1}$, $\overline{ae} \neq \overline{k_{n-1}a} = r_{n-1}$. Then, $\overline{ae} > r_{n-1}$.

Moreover,
$m - r_n \geq \overline{ae} \geq \overline{ae} - fr_n > r_{n-1} - fr_n \geq r_{n-1} - h_n r_n = r_{n+1}$.

Therefore, because $at' \equiv -fr_n + ae$, $\overline{at'} = \overline{ae} - fr_n$.

Therefore, $m - r_n \geq \overline{at'} > r_{n+1}$.

One supposes $f = h_n$ and $e = k_{n-1}$. Then, $t' = h_n k_n + k_{n-1} = k_{n+1}$. It is opposite to the assumption.

Then, in all the cases, for $t' \in \{1,2,\cdots,k_{n+1} - 1\}$, $m - r_n \geq \overline{at'} > r_{n+1}$. Therefore, the lemma is true for *n* + 1 if *n* is even. Then, it is also true for *n* + 1 = 3.

One supposes *n* odd with $n \geq 3$. One proves the recurrence by the same way as if *n* is even (cf [7]). Then the lemma is true for n+1.

Lemma 3.4 *The following inequalities holds*: $k_{d+1} \leq m$.

Proof. If $t \in \{1,2,\cdots,k_{d+1} - 1\}$, by lemma 3.3, $r_d \leq \overline{at} < m - r_{d+1}$ or $m - r_d \geq \overline{at} > r_{d+1}$, *i.e.* $r_d \leq \overline{at} < m$ or $m - r_d \geq \overline{at} > 0$ where $r_d > 0$. Then, $0 < \overline{at} < m$ or $m > \overline{at} > 0$.

Then, if $k_{d+1} > m$, there exists $t_0 \in \{1,2,\cdots,k_{d+1} - 1\}$ such that $t_0 = m$, *i.e.* $\overline{at_0} = \overline{am} = 0$. It is impossible.

Lemma 3.5 *Let* $t, t' \in \{1,2,\cdots,k_{d+1} - 1\}$ *such that* $\overline{at} = \overline{at'}$. *Then, t=t'.*

Proof. Suppose $t > t'$. Then, $a(t - t') \equiv 0$ and $\overline{a(t - t')} = 0$. Then, by lemma 3.3, $r_d \leq \overline{a(t - t')} < m - r_{d+1}$ or $m - r_d \geq \overline{a(t - t')} > r_{d+1} = 0$ where $r_d > 0$. Then, $0 < \overline{a(t - t')}$. It is a contradiction.

Lemma 3.6 *Let n* = 1, 2, $\cdots$,*d. Let* $H_n = h_1 k_1 + h_2 k_2 + h_3 k_3 + \cdots + h_n k_n$. *Then,* $H_n = k_{n+1} + k_n - 1$

The proof is basic.

Lemma 3.7 *Let n* = 1, 2, 3, ..., *d* – 1 . *Let* $L_n = \{t \mid t = 0,1,2,\cdots,H_n\}$. *Then, for all* $n \geq 1$, $L_{n+1} = \{t = l + gk_{n+1} \mid l \in L_n, g \leq h_{n+1}\}$.

Proof. Let $l \in L_n$, $l \le H_n$. Let $g \le h_{n+1}$. Therefore, if $t = l + gk_{n+1}$, $t \le H_n + h_{n+1}k_{n+1} = H_{n+1}$. Therefore, $\left\{ t = l + gk_{n+1} \middle| l \in L_n, g \le h_{n+1} \right\} \subset L_{n+1}$.

Reciprocally, let $t \in L_{n+1}$ and let $t = fk_{n+1} + e$, $e < k_{n+1}$ be the Euclidean division of t by k_{n+1}.

We know that $H_n = k_{n+1} + k_n - 1 \ge k_{n+1}$. Therefore, $e \le H_n$. Therefore, $e \in L_n$.

Therefore, if $f \le h_{n+1}$,
$t = fk_{n+1} + e \in \left\{ t = l + gk_{n+1} \middle| l \in L_n, g \le h_{n+1} \right\}$.

Moreover, if $f > h_{n+1} + 1$,

$$t = fk_{n+1} + e \ge \left(h_{n+1} + 2\right)k_{n+1} + e \ge h_{n+1}k_{n+1} + 2k_{n+1}$$
$$= H_{n+1} - H_n + 2k_{n+1} = H_{n+1} - k_{n+1} - k_n + 1 + 2k_{n+1} \;.$$
$$= H_{n+1} + k_{n+1} - k_n + 1 \ge H_{n+1} + 1$$

Therefore, $t \notin L_{n+1}$.

Then, suppose $f = h_{n+1} + 1$. Then,

$$t = fk_{n+1} + e = \left(h_{n+1} + 1\right)k_{n+1} + e$$
$$= h_{n+1}k_{n+1} + k_{n+1} + e = H_{n+1} - H_n + k_{n+1} + e \qquad .$$
$$= H_{n+1} - k_{n+1} - k_n + 1 + k_{n+1} + e = H_{n+1} - k_n + 1 + e$$

Because $t \in L_{n+1}$ and $t = H_{n+1} - k_n + 1 + e$, $e + 1 - k_n \le 0$. Therefore, $e \le k_n - 1$.

Therefore, $t = fk_{n+1} + e = h_{n+1}k_{n+1} + k_{n+1} + e$,
where $k_{n+1} + e \le k_{n+1} + k_n - 1 = H_n$

Therefore, $t = h_{n+1}k_{n+1} + e'$ where $e' \le H_n$.

Therefore, $t \in \left\{ t = l + gk_{n+1} \middle| l \in L_n, g \le h_{n+1} \right\}$.

Therefore, $L_{n+1} \subset \left\{ t = l + gk_{n+1} \middle| l \in L_n, g \le h_{n+1} \right\}$.

Therefore, $L_{n+1} = \left\{ t = l + gk_{n+1} \middle| l \in L_n, g \le h_{n+1} \right\}$.

Lemma 3.8 *Let* $F_n = \left\{ \overline{at} \middle| t = 0,1,2,\cdots,H_n \right\}$.

Let $E_n = \left\{ \overline{at} + km \middle| t = 0,1,2,\cdots,H_n, k \in Z \right\}$. *We set* $E_n = \left\{ o_s^n \middle| s \in Z \right\}$ where $o_0^n = 0$ *and* $o_{s+1}^n > o_s^n$ *for all* $s \in Z$.

Then, for all $s \in Z$, $o_{s+1}^n - o_s^n = r_n$ *or* $o_{s+1}^n - o_s^n = r_{n+1}$.

Proof We prove this lemma by recurrence.

Suppose n = 1. Then, $r_1 = a$, $H_1 = h_1k_1 = k_2 = h_1$. Therefore,

$$F_1 = \left\{ \overline{at} \middle| t = 0,1,2,\cdots,h_1 \right\} = \left\{0, a, 2a, \cdots, h_1a\right\}$$
$$= \left\{0, r_1, 2r_1, \cdots, h_1r_1 = m - r_2\right\} \qquad .$$

Therefore, the lemma is true for *n* = 1.

Suppose that the lemma is true for n.

Then, $E_{n+1} = \left\{ \overline{at} + km \middle| t=0,1,2,\cdots,H_{n+1}, k \in Z \right\}$, where

$H_{n+1} = h_1k_1 + h_2k_2 + h_3k_3 + \cdots + h_{n+1}k_{n+1} = H_n + h_{n+1}k_{n+1}$.

Because $t \in \left\{0,1,2,\cdots,H_{n+1}\right\}$, $t \in L_{n+1}$. By lemma 3.7, *If* $t \in L_{n+1}$, *t=l+gk*$_{n+1}$ where $g \le h_{n+1}$. By lemma 3.2,

$$\overline{at} \equiv \overline{a\left(l + gk_{n+1}\right)} \equiv \overline{al} + \left(-1\right)^{n+2} gr_{n+1} \equiv \overline{al} + \left(-1\right)^n gr_{n+1} \quad .$$

Therefore,

$$E_{n+1} = \left\{ \overline{at} + km \middle| t \in L_{n+1}, k \in Z \right\}$$
$$= \left\{ \overline{at} + km \middle| t = l + gk_{n+1}, l \in L_n, g \le h_{n+1}, k \in Z \right\}$$
$$= \left\{ \overline{al} + \left(-1\right)^n gr_{n+1} + km \middle| l \in L_n, g \le h_{n+1}, k \in Z \right\}$$
$$= \left\{ f + \left(-1\right)^n gr_{n+1} + km \middle| f \in F_n, g \le h_{n+1}, k \in Z \right\}$$
$$= \left\{ o_s^n + \left(-1\right)^n gr_{n+1} + km \middle| s \in Z, g \le h_{n+1}, k \in Z \right\}$$

Suppose that n is even.

Then, $o_s^n + \left(-1\right)^n gr_{n+1} = o_s^n + gr_{n+1} \le o_s^n + r_n - r_{n+2}$ because $gr_{n+1} \le h_{n+1}r_{n+1} = r_n - r_{n+2}$.

Use the recurrence. Suppose $o_{s+1}^n - o_s^n = r_n$. Then, $o_s^n + \left(-1\right)^n gr_{n+1} \le o_s^n + r_n - r_{n+2} = o_{s+1}^n - r_{n+2}$.

Therefore,

$$\left\{ o_t^{n+1} \middle| o_s^n \le o_t^{n+1} < o_{s+1}^n \right\}$$
$$= \left\{ o_s^n < o_s^n + r_{n+1} < \cdots < o_s^n + h_{n+1}r_{n+1} < o_{s+1}^n \right\} \qquad .$$

Therefore, $o_{t+1}^{n+1} - o_t^{n+1} = r_{n+1}$ or r_{n+2} if $o_s^n \le o_t^{n+1} < o_{t+1}^{n+1} \le o_{s+1}^n$.

Suppose $o_{s+1}^n - o_s^n = r_{n+1}$. Then, *s* is fixed .

Let $T = \min\left\{ t = 0,1,\cdots, \middle| o_{s+t+1}^n - o_{s+t}^n = r_n \right\}$. Therefore, $o_{s+T+1}^n - o_{s+T}^n = r_n$.

Let $O = \cup_{t=0}^T \left\{ o_{s+t}^n + gr_{n+1} \middle| 0 \le g \le h_{n+1} \right\}$.

Then,
$O = \left\{ o_s^n, o_{s+1}^n, \cdots, o_{s+T-1}^n \right\} \cup \left\{ o_{s+T}^n + gr_{n+1} \middle| 0 \le g \le h_{n+1} \right\}$.

Therefore, $O = \left\{ o'_s, o'_{s+1}, \cdots, o'_{s+K} \right\}$ where $o'_{s'+1} - o'_{s'} = r_{n+1}$. Moreover, $o_{s+T+1}^n - o'_{s+K} = r_n - h_{n+1}r_{n+1} = r_{n+2}$.

Therefore, if $o_{t'}^{n+1}$ and $o_{t'+1}^{n+1} \in \left\{ o_t^{n+1} \middle| o_s^n \le o_t^{n+1} \le o_{s+T+1}^n \right\}$, $o_{t'+1}^{n+1} - o_{t'}^{n+1} = r_{n+1}$ or r_{n+2}.

Suppose that n is odd. One proves this result by the same way as when n is even (cf [7]).

Proof 3.9 *Now one proves theorem* 2.

Suppose that n is even.

Then, $\overline{k_{n-1}a} = r_{n-1}$, $2\overline{k_{n-1}a} = 2r_{n-1}$, ...
$h_{n-1}\overline{k_{n-1}a} = h_{n-1}r_{n-1} = r_n - r_{n-2}$.

Now, $\overline{ak_{n-1}\ell} = \overline{\ell r_{n-1}} = \ell r_{n-1}$ for $\ell = 0,1,2,\cdots,h_{n-1}$.

$$\left\{ \left(k_{n-1}\ell, r_{n-1}\ell\right) \middle| \ell = 0,1,2,\cdots,h_{n-1} \right\}$$
$$= \left\{ \left(k_{n-1}\ell, \overline{ak_{n-1}\ell}\right) \middle| \ell = 0,1,2,\cdots,h_{n-1} \right\} \subset E_2$$

Therefore,

Moreover, $r_{n-2} = h_{n-1}r_{n-1} + r_n$. On the other hand, by lemma 3.8 , all the points of $E_2 = \left(t, \overline{at}\right)$, $t \le H_{n-1}$, have ordinates distant of r_n or r_{n-1}.

Therefore, if there is other points of $E_2 \cap \left\{ \left[0, H_{n-1}\right] \otimes \left[0, r_{n-2}\right[\right\}$ that the points $\left\{ \left(k_{n-1}\ell, r_{n-1}\ell\right) \middle| \ell = 0,1,2,\cdots,h_{n-1} \right\}$, there exists $\ell_0 \in \left\{1,2,\cdots,h_{n-1}\right\}$ and

$(x_1,y_1)\in E_2\cap\{[0,H_{n-1}]\otimes[0,r_{n-2}]\}$ such that $r_{n-1}\ell_0-y_1=r_n$.

Because $H_{n-1}=k_n+k_{n-1}-1<k_{n+1}\le k_{d+1}$, by lemma 3.5, there exists an only $t\in\{1,\cdots,H_{n-1}\}$, such that $\overline{at}=y_1$: $t=x_1$. Because $y_1\neq 0$, there exists an only $t\in\{0,1,\cdots,H_{n-1}\}$, such that $\overline{at}=y_1$.

Now, $r_{n-1}\ell_0-y_1=\overline{a\ell_0k_{n-1}}-\overline{at}=r_n=\overline{-ak_n}$. Then, $\overline{a\ell_0k_{n-1}}-\overline{-ak_n}=\overline{at}$. Then, $\overline{a(\ell_0k_{n-1}+k_n)}=\overline{at}$.

Because $r_{d-1}=h_dr_d$ with $r_{d-1}>r_d$, $h_d\ge 2$. Moreover, $d\ge n\ge 2$. Then, $d-1>0$. Then, $k_{d-1}>0$.

Then, by lemma 3.4, $2k_n-k_{n-2}\le 2k_d<2k_d+k_{d-1}\le h_dk_d+k_{d-1}=k_{d+1}\le m$. Then, $0<\ell_0k_{n-1}+k_n<k_{d+1}$.

Then, by lemma 3.4,

$$0<k_{n-1}+k_n\le \ell_0k_{n-1}+k_n\le h_{n-1}k_{n-1}+k_n\le k_n-k_{n-2}+k_n$$
$$=2k_n-k_{n-2}\le 2k_d<2k_d+k_{d-1}\le h_dk_d+k_{d-1}=k_{d+1}\le m.$$

Then, $0<\ell_0k_{n-1}+k_n<k_{d+1}$.

Now $0<t\le H_{n-1}=k_n+k_{n-1}-1<k_d+k_{d-1}\le k_{d+1}$. Moreover, $0<\ell_0k_{n-1}+k_n<k_{d+1}$.

Then, because $\overline{a(\ell_0k_{n-1}+k_n)}=\overline{at}$, by lemma 3.5, $t=\ell_0k_{n-1}+k_n$.

Then, $t=\ell_0k_{n-1}+k_n\ge k_{n-1}+k_n>H_{n-1}$. It is a contradiction.

Therefore, there is not other points of $E_2\cap\{[0,H_{n-1}]\otimes[0,r_{n-2}[\}$ that $\{(k_{n-1}\ell,r_{n-1}\ell)|\ell=0,1,2,\cdots,h_{n-1}\}$.

Therefore, there is not other points of $E_2\cap\{[0,k_n[\otimes[0,r_{n-2}[\}$ that the points $\{(k_{n-1}\ell,r_{n-1}\ell)|\ell=0,1,2,\cdots,h_{n-1}\}$.

Therefore,

$E_2\cap\{[0,k_n[\otimes[0,r_{n-2}[\}=\{(k_{n-1}\ell,r_{n-1}\ell)|\ell=0,1,2,\cdots,h_{n-1}\}$.

According to what precedes,

$$\left\{\left(k_{n-1}\ell,\overline{ak_{n-1}\ell}\right)|\ell=0,1,2,\cdots,h_{n-1}\right\}$$
$$=\left\{\left(k_{n-1}\ell,r_{n-1}\ell\right)|\ell=0,1,2,\cdots,h_{n-1}\right\}$$

is located on the straight line $y=(r_{n-1}/k_{n-1})x$ if n is even.

Suppose that n is odd. One proves this result by the same way as when n is even (cf [7]).

4. Models Equivalent with a Margin of ε

4.1. Correct Models

In general terms, one can always suppose that y_n is the realization of a sequence of random variables Y_n defined on a probability space (Ω,A,P): $y_n=Y_n(\omega)$ where $\omega\in\Omega$ and where Y_n is a correct model of y_n. As a matter of fact, there exist an infinity of correct models of y_n. It is thus necessary to be placed in the set of all the possible random variables.

Notations 4.1 *One considers the sequences of random variables Y_n^θ, n = $1,\cdots,N$, defined on the probabilities spaces (Ω,A,P_θ), $\theta\in\Theta$:*

$(Y_1^\theta,Y_2^\theta,\cdots,Y_N^\theta):\Omega\to\{0/m,1/m,\cdots,(m-1)/m\}^N$. *One assumes that $Y_n^\theta=Y_n$ for all $\theta\in\Theta$.*

For example, one can assume that $\Omega=\{0/m,1/m,\cdots,(m-1)/m\}^N$ and $(Y_1,\cdots,Y_N)=(Id,\cdots,Id)$.

It thus raises the question to define what is a correct model. Indeed, if a model Y_n^θ is not correct, it is however possible that $y_n=Y_n^\theta(\omega)$. Now, in the case where the model Y_n^θ is IID, to define a correct model is a generalization of the problem of the definition of an IID sequence. Then, it is a very complex problem (cf [1]).

However, generally, one feels well that correct models exist. In fact, it is a traditional assumption in science. In weather for example, the researchers seek a correct model, which implies its existence (if not, why to try to make forecasts?).

One could thus admit that like a conjecture or a postulate without defining exactly what is a correct model. Cependant, a more detailed study is in [7].

4.2. Models Equivalent

4.2.1. The Problem

Let $Y_n^{\theta_2}$ and $Y_n^{\theta_1}$ be two sequences of random variables such that, for all Borel set Bo,

$$P\left\{\left(Y_1^{\theta_2},\cdots,Y_N^{\theta_2}\right)\in Bo\right\}=P\left\{\left(Y_1^{\theta_1},\cdots,Y_N^{\theta_1}\right)\in Bo\right\}\left[1+Ob(1)\varepsilon\right],$$

where Ob(.) means the classical O(.) with the additional condition $|Ob(1)|\le 1$. One supposes that $Y_n^{\theta_1}$ is a correct model of the sequence y_n, n = 1, 2, $\cdots$,N. One wants to prove that $Y_n^{\theta_2}$ is also a correct model of y_n if ε is small enough.

4.2.2. Example

Let us suppose that we have a really IID sequence of random variables X_n^ε with uniform distribution on [0,1/2] and [1/2,1] and with a probability such as $P\{X_n^\varepsilon\in[1/2,1]\}=0,500[1+\varepsilon]$ where $\varepsilon=0,001$. Then, this sequence has not the uniform distribution on [0,1]. However, if we have a sample with size 10, we will absoluetely not understand that X_n^ε has not the uniform distribution on [0,1]. It is wellknown that one need samples with size larger than N = 1000 minimum in order to test this difference.

For example, one cannot test significantly H_0: " X_n^θ has the uniform distribution" against $H_1(\varepsilon)$:

“ $P\left\{X_n^\theta \in Bo\right\} = L(Bo)\left[1+Ob(1)\varepsilon\right]$ ” if $\sqrt{N}\ \varepsilon \le 1/10$. Indeed, if $\sqrt{N}\ \varepsilon = 1/10$ and b = 2, the probability of obtaining $\dfrac{\sum_{n=1}^{N}\left[1_{[1/2,1]}\left(X_n^\theta\right)-1/2\right]}{\sqrt{N/4}} \ge 2$ is about 0.0466 under $H_1(\varepsilon)$ and about 0.0455 under H_0 : *i.e.* the probability of rejecting the assumption IID, H_0, under $H_1(\varepsilon)$ is not much bigger than that of rejecting H_0 if X_n^θ is really IID (cf also section 4.3 of [8]).

Then it is no possible to differentiate the IID model and models such that $P\left\{\left(X_1^\theta,\cdots,X_N^\theta\right)\in Bo\right\} = L(Bo)\left[1+Ob(1)\varepsilon\right]$.

4.2.3. Border of Correct Models

Now there is a problem : for example, use a realization x_n of the IID model, and let $X_n^{\theta_1}$ be a model checking $P\left\{\left(X_1^{\theta_1},\cdots,X_N^{\theta_1}\right)\in Bo\right\} = L(Bo)\left[1+Ob(1)\varepsilon^1\right]$ where ε^1 is small enough but not very small. Let $X_n^{\theta_2}$ be a model such that

$$P\left\{\left(X_1^{\theta_2},\cdots,X_N^{\theta_2}\right)\in Bo\right\}$$
$$= P\left\{\left(X_1^{\theta_1},\cdots,X_N^{\theta_1}\right)\in Bo\right\}\left[1+Ob(1)\varepsilon\right]^{\cdot}$$

Then, $X_n^{\theta_2}$ can be not a correct model: it is enough that ε^1 is in extreme cases of the possible values of the ε ’s such that $P\left\{\left(X_1^{\theta_2},\cdots,X_N^{\theta_2}\right)\in Bo\right\} = L(Bo)\left[1+Ob(1)\varepsilon\right]$, $sup_{Bo}\left(Ob(1)\right)=1$, imply that X_n^θ is a correct model.

Then, there are models more correct than others. These are models $Y_n^{\theta_0}$ such that, if $P\left\{\left(Y_1^\theta,\cdots,Y_N^\theta\right)\right.$, Y_n^θ is also a correct model where ε is small enough, but not too small. For example, $\varepsilon = 1/10$, 1/100 or at worst $\varepsilon = 1/N$ if need be (cf section 5-7 of [7]).

It seems clear that such models exist. For example it is assumed that this is the case when x_n is sample of an IID sequence X_n^ε and that it is a good realization of X_n^ε .

On the other hand, we know that it should exist estimates of models (these estimates are easier to calculate in some cases as texts). Then, we can choose as model $Y_n^{\theta_1}$, the model provided by these estimates : close models will also correct models.

All these points are detailed in [7].

4.3. Exact IID Model

Then, generally, if Y_n^θ is a correct model such that $T_q\left(Y_n^\theta\right)$ cannot be differentiated with the IID model, one will be able to choose another correct model $Y_n^{\theta_0}$ close to Y_n^θ and such that $T_q\left(Y_n^{\theta_0}\right)$ is exactly the IID model. Indeed one proves easily the following proposition (cf proposition 5-1 of [8]).

Proposition 4.1 *One assumes that m is large enough. Let* $Y_n^{\theta_c}$ *be a correct model of the sequence* y_n *. One assumes that there exists* $\varepsilon_Y > 0$ *such that if* Y_n^θ *is a model satisfying, for all Borel set Bo,*

$$P\left\{\left(X_1^\theta,\cdots,X_N^\theta\right)\in Bo\right\} = P\left\{\left(X_1^{\theta_c},\cdots,X_N^{\theta_c}\right)\in Bo\right\}\left[1+Ob(1)\varepsilon_Y\right],$$

then Y_n^θ *is a correct model of* y_n *.*

One assumes also that, for all $\left(k_1,\cdots,k_N\right)$,

$$P\left\{\left\{T_q\left(Y_1^{\theta_c}\right)=k_1/2^q\right\}\cap\cdots\cap\left\{T_q\left(Y_N^{\theta_c}\right)=k_N/2^q\right\}\right\}$$
$$=\frac{1}{2^{qN}}\left[1+\varepsilon_{k_1,\cdots,k_N}(q)\right]$$

where $sup_{k_1,\cdots,k_N}\left|\varepsilon_{k_1,\cdots,k_N}(q)\right| = \varepsilon_X(q)$. *One assumes that* $\varepsilon_X(q)$ *is increasing, that* $\varepsilon_X(1) \ll \varepsilon_Y$ *and that there exists* $q_1 \in N^*$ *such that* $\varepsilon_X(q_1)$ *is small enough.*

Then, there exists $q_0 \in N^*$ *and a correct model* $Y_n^{\theta_0}$ *of the sequence* $\left\{y_n\right\}_{n=1,\cdots,N}$ *such that, for all* $\left(k_1,\cdots,k_N\right)$,

$$P\left\{\left\{T_{q_0}\left(Y_1^{\theta_c}\right)=k_1/2^q\right\}\cap\cdots\cap\left\{T_{q_0}\left(Y_N^{\theta_c}\right)=k_N/2^q\right\}\right\}=\frac{1}{2^{q_0 N}}$$

5. Approximation Theorem

5.1. Theorem

In this section, we assume that T is a Fibonacci congruence and we use Fibonacci function T_q *in order to build IID sequences.*

Theorem 3 *We keep the notations* 1.3 *and* 1.4 *and notations of section* 2.3. *Let* $\gamma(m)=\left[2+\varphi(m)\right]$. *We assume* $\gamma(m)NK_0\,2^q/m \approx 0$ *and* $m/K_0 \gg 1$. *Then, for all Borel set Bo,*

$$P\left\{\left(X_1,\cdots,X_N\right)\in Bo\right\} = L(Bo)\left[1+\frac{\gamma(m)Ob'(1)NK_0}{m/2^q}\right].$$

where $\left|Ob'(1)\right|$ *is increased by a number close to* 1.

If K_0 *is not too large, there is no difficulty to choose m and q in such a way that* $\varepsilon \le \gamma(m)2^q N K_0/m$ *is small enough. Therefore,*

$$P\left\{\left(X_1,\cdots,X_N\right)\in Bo\right\} = L(Bo)\left[1+Ob(1)\varepsilon\right].$$

5.2. Proof

Because, by Section 2.3, *the points of* $\overline{T}^{-1}(mI)$ *are well distributed in* $\left\{0,1,\cdots,m-1\right\}$, *it is easy to prove that the sum of points of* $h_N'\left(\overline{T}^{-1}(mI)\right)$ *will be close* $card\left(mI\cap\left\{0,1,\cdots,m-1\right\}\right)/m$ (*e.g. cf* **Figure** 7). *Then, we have the following lemma* (*cf also proposition* 6-1 of [7]).

Lemma 5.1 *Let* h_N *be the probability density function of* $Y \in \left\{0/m,1/m,\cdots,(m-1)/m\right\}$, *with respect to* μ_m :

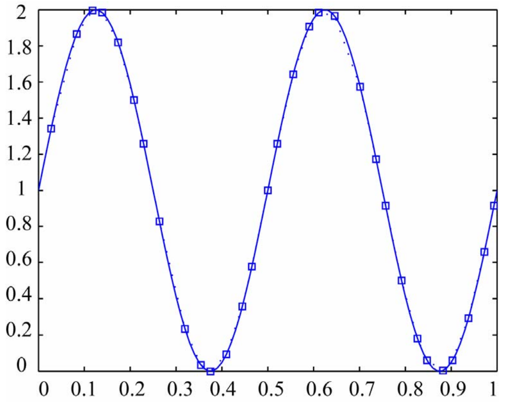

Figure 7. Points of $\boldsymbol{h'_N\left(\overline{T}^{-1}(mI_k)\right)}$ **when** $\boldsymbol{h'_N(t)=\sin(4\pi t)+1}$**.**

$\int_0^1 h_N(u)\,\mu_m(du)=1$. *Let* $h'_N=(1/c_0)\,h_N$ *such that* $\int_0^1 h'_N(u)\,\mathrm{d}u=1$.

Let $K'_0\in R_+$ *such that* $\left|h_N(r)-h_N(r')\right|\le K'_0\left|r'-r\right|$ *and* $\left|h'_N(r)-h'_N(r')\right|\le K'_0\left|r'-r\right|$ when $r,r'\in[0,1]$. *One supposes* $2^q/m\approx 0$, *and* $m/K'_0\gg 1$.

Then, *the following equality holds*:

$$P\left\{\overline{T}(mY)/m\in I_k\right\}=L(I_k)\left[1+\frac{\gamma(m)Ob'(1)K'_0}{m/2^q}\right],$$

where $I_k=\left[k/2^q,(k+1)/2^q\right[$, $L(I_k)=1/2^q$.

Then, **one can prove theorem 3**. Indeed, by applying lemma 5.1 when Y has for distribution the distribution of the conditional probability of $Y_{\phi(n)}$ given $Y_{\phi(n-1)}=y'_1,Y_{\phi(n-2)}=y'_2,\cdots$, we have

$$P\left\{\overline{T}(mY)/m\in I_k\right\}$$
$$=P\left\{X_{\phi(n)}\in I_k\left|Y_{\phi(n-1)}=y'_1,Y_{\phi(n-2)}=y'_2,\cdots\right.\right\}$$
$$=L(I_k)\left[1+\frac{\gamma(m)Ob'(1)K_0}{m/2^q}\right].$$

Now, one proves easily that

$$P\left\{X_{\phi(n)}\in I_k\left|X_{\phi(n-1)}=x'_1,X_{\phi(n-2)}=x'_2,\cdots\right.\right\}$$
$$=\sum_{y'_{s_1}\in T_q^{-1}(x'_1)}\cdots\sum_{y'_{s_{N-1}}\in T_q^{-1}(x'_{N-1})}\eta_{y'_{s_1},\cdots,y'_{s_{N-1}}}$$
$$\cdot P\left\{X_{\phi(n)}\in I_k\left|Y_{\phi(n-1)}=y'_{s_1},Y_{\phi(n-2)}=y'_{s_2},\cdots\right.\right\},$$

where

$$\eta_{y'_{s_1},\cdots,y'_{s_{N-1}}}=\frac{P\left\{\left\{Y_{\phi(n-1)}=y'_{s_1}\right\}\cap\cdots\cap\left\{Y_{\phi(n-N+1)}=y'_{s_{N-1}}\right\}\right\}}{\sum_{y'_{s_1}\in T_q^{-1}(x'_1)}\cdots\sum_{y'_{s_{N-1}}\in T_q^{-1}\{x'_{N-1}\}}P\left\{\left\{Y_{\phi(n-1)}=y'_{s_1}\right\}\cap\cdots\cap\left\{Y_{\phi(n-N+1)}=y'_{s_{N-1}}\right\}\right\}}$$

Then, $\sum_{y'_{s_1}\in T_q^{-1}(x'_1)}\cdots\sum_{y'_{s_{N-1}}\in T_q^{-1}(x'_{N-1})}\eta_{y'_{s_1},\cdots,y'_{s_{N-1}}}=1$.

We deduce that

$$P\left\{X_{\phi(n)}\in I_k\left|X_{\phi(n-1)}=x'_1,X_{\phi(n-2)}=x'_2,\cdots\right.\right\}$$
$$=L(I_k)\left[1+\frac{\gamma(m)Ob'(1)K_0}{m/2^q}\right].$$

Then, one proves by basic methods (cf proposition 6.2 of [7]) that, for all $I_{k_1}\otimes\cdots\otimes I_{k_N}$,

$$P\left\{(X_1,\cdots,X_N)\in I_{k_1}\otimes\cdots\otimes I_{k_N}\right\}$$
$$=\prod_{s=1}^{N}\left(L\left(I_{k_s}\right)\left[1+Ob(1)\varepsilon\right]\right),$$

where $|\varepsilon|\le\dfrac{\gamma(m)\left|Ob'(1)\right|K_0}{m/2^q}$. Because $\gamma(m)NK_0\,2^q/m\approx 0$, we deduce that

$$P\left\{\left(X_1^\theta,\cdots,X_N^\theta\right)\in I_{k_1}\otimes\cdots\otimes I_{k_N}\right\}$$
$$=\frac{1}{2^{Nq}}\left[1+\frac{\gamma(m)Ob'(1)NK_0}{m/2^q}\right].$$

Now, we deduce that, for all Borel set $Bo\subset\left\{0/2^q,,\cdots,\left(2^q-1\right)/2^q\right\}^N$,

$$P\left\{(X_1,\cdots,X_N)\in Bo\right\}=L(Bo)\left[1+\frac{\gamma(m)Ob'(1)NK_0}{m/2^q}\right].$$

6. Choice of Random Noises

6.1. Use of Texts

Now, we suppose that we use sequences y'_n and $y''_n\in\left\{0/m,1/m,\cdots,(m-1)/m\right\}$, n = 1,2,⋯,N, obtained from independent texts. In order to reduce K_0 we add modulo m a text and a text written backward:

$y_n=\left[\overline{y'_n+y''_{N-n+1}+rand_0(n)}\right]/m$ where $rand_0(n)$ is pseudo-random sequences which have good empirical independence assumptions for p successive pseudo random numbers when $p\ge 3$. In an obvious way, the texts are realizations of sequences of random variables : for example, one can take as model, the set of the possible texts provided with the uniform probability

As a matter of fact, we add $rand_0(n)$ to have sequences y_n which have a good randomness (cf [9], or chapter 3 of [6]).

In particular, a priori, "$P\{Y_n = y\}$ is not too different from 1/m" is a reasonable assumption. Moreover, (y_n, y_{n+1}, y_{n+2}) has a empirical distribution close to independence and texts behaves as Q-dependent sequences (cf [6]). Then, for all permutation ϕ, a three-dimensional model $\left(Y_{\phi(n)}, Y_{\phi(n+1)}, Y_{\phi(n+2)}\right)$ with a continuous density and a Lipschitz coefficient not too big will be a good model. By the same way,

$P\left\{Y_{\phi(n)} = y \middle| Y_{\phi(n-1)} = y'_1, Y_{\phi(n-2)} = y'_2, \cdots\right\}$ is not too different of $P\{Y_n = y\}$ which is not too different from 1/m (cf [7]). In this case, one can prove that generally K_0 is small cf [7].

On the other hand, to increase K_0 a good way is to use the Central Limit theorem. In fact one can combine the two methods : cf [7].

6.2. Example

Now it may be necessary to do some transformations to get the $my'_n \in \{0, 1, \cdots, m-1\}$ in the case where the letters and symbols are provided by sequences a(*j*), $j = 1, 2, \cdots, N_3$, $a(j) \in \{0, 1, \cdots, 255\}$, $j = 1, 2, \cdots, N_3$.

One sets $N_0 = \lfloor N_3 / r_1 \rfloor$. We choose two consecutive elements a and m of the Fibonacci sequence : m can be chosen with respect to N_0. Then, we choose r_1 such that $a < 32^{r_1} \le m$.

1) We set $c(j) = \overline{a(j)}$ mod $\kappa = 32$

2) We set $d(n) = \sum_{r=1}^{r_1} c(r_1(n-1)+r)\kappa^{r-1}$ for $j = 1, 2, \cdots, N_0$.

3) We set $y'_n = \left\lfloor d(n) m / \kappa^{r_1} \right\rfloor / m$ for $j = 1, 2, \cdots, N_0$.

By using this technique, we have created real IIID sequences x_n. We have used a sequence *c*(*j*) with $N_3 = 20{,}000{,}000$. This sequence was obtained from dictionary, encyclopedia, and Bible. We choose $r_1 = 20$, $a < 4 * 10^{36} \le m$, $q = 70$. Then $N_0 = 10^6$. Then, we have estimated $K_0 = 0.01$. In order to avoid any error we have choose $K_0 = 10^4$ in the building of x_n.

We have tested the sequence x_n. We have used the classical Diehard tests (cf [1] [2]) and the higher order correlation coefficients (cf [10]). Results are in accordance with what we waited: the hypothesis "randomness" is accepted by all these tests (cf [7]).

One can can download this sequence in [11]

7. Conclusion: Building of Random Sequence

By theorem 3 one can find models correct Y_n^θ such that $P\left\{\left(X_1^\theta, \cdots, X_N^\theta\right) \in Bo\right\} = L(Bo)\left[1 + Ob(1)\varepsilon\right]$ where ε is small if K_0 is not too large. Now it is possible to build such sequences concretely, for example by using texts studied in section 6. In this case, coefficient K_0 depend on the choice of m, *i.e.* of r_1. But, K_0 increases very little when r_1 increases. Even, in some cases, it seems that it decreases. Then, at most $2^q / m$ decreases much more quickly than K_0 increases.

So by taking m large enough and by choosing well q, we found ε small enough in a way that there exists correct models which checks the conditions of proposition 4.1. Then, there exists m sufficiently large and q sufficiently small and a correct model $Y_n^{\theta_0} \in \{0/m, 1/m, \cdots, (m-1)/m\}$ such that $T_q\left(Y_n^{\theta_0}\right)$ is the IID model.

Then, this result show that **one can build sequences x_n such that the model IID is a correct model of x_n.**

That means that x_n behaves like any IID sample: a priori, x_n can check not the properties which one expects from a IID sample like certain tests, but that occurs only with a probability equal to that of any IID sample.

By this method, we therefore have a mean to value the technique used by Marsaglia to create its CD-ROM. We arrive in fact *to prove mathematically* that the sequence obtained can be regarded a priori as random, what Marsaglia did not.

Remark 7.1 *One might wonder if the sequence built adding text and pseudo-random sequences is not an IID sequence. It is a similar hypothesis which Marsaglia does when he built its CD-Rom. This also corresponds to results of* [9]. *But in fact, nothing is proved.*

It is maybe possible to prove it but that seems complicated. Finally, it is much easier to apply the functions T_q: in this case, it requires only that K_0 is not too big. It's an hypothesis much simpler to be verified and it does not require many efforts in some cases. That is why we choose to build IID sequences using this technique.

8. References

[1] D. E. Knuth, "The Art of Computer Programming," 3rd Edition, Addison-Wesley, Boston, 1998.

[2] J. Gentle, "Random Number Generation and Monte Carlo Method," Springer 13, 1984, pp. 61-81.

[3] A. Menezes, P. Van Oorschot, S. Vanstone, "Handbook of Applied Cryptography," CRC Press, London, 1996. doi:10.1201/9781439821916

[4] R. Blacher, "Solution Complete Au Probleme des Nombres Aleatoires," 2004. http://www.agro-montpellier.fr/sfds/CD/textes/blacher1.pdf

[5] U. Dieter, "Statistical Interdependence of Pseudo Random numbers Generated by the Linear Congruential

Method," In: S. K. Zaremba, Ed., *Applications of Number Theory to Numerical Analysis*, Academic Press, New York, 1972, pp. 287-317.

[6] R. Blacher, "A Perfect Random Number Generator," Rapport de Recherche LJK Universite de Grenoble, 2009. http://hal.archives-ouvertes.fr/hal-00426555/fr/

[7] R. Blacher, "Fibonacci Congruences and Applications," Rapport de Recherche LJK Universite de Grenoble, 2011. http://hal.archives-ouvertes.fr/hal-00587108/fr/.

[8] R. Blacher, "Correct models," Rapport de Recherche LJK Universite de Grenoble, 2010. http://hal.archives-ouvertes.fr/hal-00521529/fr/

[9] L. Y. Deng and E. O. George, "Some Characterizations of the Uniform Distribution with Applications to Random Number Generation," *Annals of the Institute of Statistical Mathematics*, Vol. 44, No. 2, 1992, pp. 379-385. doi:10.1007/BF00058647

[10] R. Blacher, "Higher Order Correlation Coefficients," *Statistics*, Vol. 25, No. 1, 1993, pp. 1-15. doi:10.1080/02331889308802427

[11] R. Blacher, File of random Number, 2009. http://www-ljk.imag.fr/membres/Rene.Blacher/GEAL/node3.html.