Ratio Math. 3
(1992), 35 - 48

CHECKING THE RANDOMNESS OF PSEUDORANDOM SEQUENCES FROM A BAYESIAN VIEWPOINT(*)

Giuseppe Di Biase(**) **and Antonio Maturo**(*)
Università "G. D'Annunzio" - Chieti

Sommario. Viene affrontato il problema della verifica di casualità di successioni di numeri pseudocasuali, soprattutto da un punto di vista bayesiano. Nel lavoro vengono proposti dei criteri operativi di carattere convenzionale adatti per la verifica della bontà dei generatori ed alcune analisi bayesiane. In particolare viene esaminata, da un punto di vista teorico, quella sostitutiva del test classico delle frequenze, la cui applicazione numerica è stata effettuata dagli stessi autori in [5].

Abstract. After illustrating the classic and the bayesian approaches to the analysis of randomness of pseudorandom sequences, the bayesian viewpoint is further dealt with, since up to now, it has not been exhaustively analysed. A bayesian analysis wich can replace the frequency test is also presentated. The numerical application and its results can be found in [5].

(*) Work carried out within the Research Project MURST 40% entitled "Mathematical Statistics". National coordinator L. Piccinato.

(*) Dipartimento di Scienze, Storia dell'Architettura e Restauro - Viale Pindaro, 42 - 65127 Pescara

(**) Facoltà di Farmacia - Via degli Agostiniani - Chieti

1. STATISTICAL ANALYSIS OF PSEUDORANDOM GENERATORS

Let N be the set of natural numbers, and X a continuous random variable uniformly distributed over (0,1).

Let $\{X_i\}_{i\in N}$ be a sequence of independent random variables having the same distribution as X. If x_i (for each $i\in N$) is a realization of X_i, the sequence $\{x_i\}_{i\in N}$ is called a **sequence of random numbers**.

For each $m\in N$, $X^{(m)}$ is a standard discrete random variable with parameter m uniformly distributed, i.e. with values in {0, 1/m, 2/m, ..., (m-1)/m}, each with probability 1/m.

$X^{(m)}$ has the characteristic function

$$H_m(u) = (1/m)(e^{iu}-1)/(e^{iu/m}-1).$$

Therefore it results that:

$$\lim_{m\to+\infty} H_m(u) = (e^{iu}-1)/iu,$$

where the value of the limit defines the characteristic function of the random variable X, so that the sequence $\{X^{(m)}\}_{m\in N}$ converges in distribution to X.

Since computers employ only rational numbers with a limited number of decimal-coded digits, the sequence $\{X_i\}_{i\in N}$ can be replaced by a sequence $\{X_i^{(m)}\}_{i\in N}$ of independent random variables having the same distribution as $X^{(m)}$, for sufficiently large m. Instead of the sequence $\{x_i\}_{i\in N}$, it is then considered the sequence $\{x_i^{(m)}\}_{i\in N}$ where $x_i^{(m)}$ corresponds to a realization of $X_i^{(m)}$, called **sequence of random numbers** too.

In practice, given a value of m, the sequence of values $x_i^{(m)}$ taken by $X_i^{(m)}$, $i\in N$, is replaced by a sequence $\{z_i\}_{i\in N}$ of values obtained with the computer applying the recursive formula:

$$y_{i+1} = f(\alpha_1, \alpha_2, \ldots, \alpha_h; y_i), \quad i\in N, \tag{1.1}$$

where f is a function of the variable y_i, defined from $I_m = \{0, 1, 2, \ldots, m-1\}$ to I_m, and depending on the parameters $\alpha_1, \alpha_2, \ldots, \alpha_h$. Therefore, if the starting number $y_1 \in I_m$ is fixed, a sequence $\{y_i\}_{i\in N}$ of elements of set I_m is obtained.

Subsequently, assuming $z_i = y_i/m$, a sequence whose space is {0, 1/m, ..., (m-1)/m} is obtained.

This procedure may seem completely illogical. In fact, according to it, they are not considered random numbers, but numbers obtained out of a formula in a deterministic way. These are called **pseudorandom numbers.**

The procedure can be giustified by the subjective definition of pseudorandom numbers (see i.e. Bruno Baldessari (1987), [1], Alfredo Rizzi (1989), [11]) according to which the sequence $\{z_i^{(m)}\}_{i\in N}$ is said **pseudorandom** if:

(1) a user who knows the generator (1.1) can univocally generate the sequence;

(2) a user who does not know the generator, can assume that the sequence $\{z_i\}_{i\in N}$ is a realization of a sequence $\{X_i^{(m)}\}_{i\in N}$ of standard discrete random variables with parameter m uniformly distribuited.

In fact, if the function f and its parameters are given in the correct way, the evaluation of y_{i+1} starting from y_i is the result of a "randomness effect" similar to the one resulting from mixing cards. As a matter of fact, even in this case, the card position is deterministic and its randomness depends on the fact that the operator does not know it.

In the formula (1.1), to obtain sequences concording with Baldessari's subjective definition, it is necessary that the operator who generates the sequence knows well its matematical and statistical properties, so that it may seem random to any observer.

We obtain this by assuming f egual to a function as simple as possible. For example the formula

$$y_{i+1} = (ay_i+b) \bmod m, \quad i\in N, \quad a,b\in I_m \tag{1.2}$$

is largely used.

The mathematical and statistical properties of the sequence $S=\{y_i\}_{i\in N}$, obtained by generator (1.2), vary considerably according to the variation of the two parameters a e b, of the modulo m and of the starting point y_1 (see A. Rizzi (1977), [10], Cera and Maturo (1983), [2] and [3], Maturo (1989), [7], Maturo and Piscione (1990), [8]).

Let then consider a sequence $\{y_i\}_{i\in N}$ whose space I_m is obtained from any formula of the (1.1) type, and let $z_i=y_i/m$.

A classical approach largely used in order to prove the randomness of the sequence is the following.

Once we have fixed $n\in N$, we ask if the null hypothesis H_0 can be accepted: z_i are determinations of independent random variables Z_i having continuous uniform distribution or discrete normalized uniform distribution with parameter m of the random variable $Z=X^{(m)}$.

For this purpose check the finite sequence $S=\{z_i\}_{i\in\{1,2,\ldots,n\}}$ of lenght n by means of a set of statistical tests.

Some of the most common tests are (see A. Rizzi (1977) e (1989), [10] e [11], Knuth (1969), [6], Maturo (1989), [8], Cera e Maturo (1983) e (1990), [2] e [4], Tausworthe (1965), [14]): Test on the mean; on the numbers variance; on the numbers frequence; on the digits frequence; on the number pairs frequence;

on the digit pairs frequence; on the numbers correlation; on the digits correlation; on the complete sequences; the runs test; the gaps test.

The non-parametric checking methods, used in the aforementioned studies are based on the following steps:

a) consider suitable statistics $D_k=D_k(Z_1, Z_2, \ldots, Z_n)$, $k=1, 2, \ldots, r$, assuming nonnegative real values that, **in certain situations regarded as «ideal»**, assume the value zero. The D_k usually have asintotic or chi-square distribution or they are equal to absolute value of a symmetric distribution with respect to zero and tabulated (i.e. Student's t-distribution or standardized normal one).

b) For each k a "small" number $\alpha_k \in (0,1)$ is fixed, (usualy $\alpha_k=0.01$ or $\alpha_k=0.001$ or a value between the two), called level of significance, and consequently a real number β_k such that prob $(D_k \geq \beta_k)=\alpha_k$ is determined. The pair $T_k=(D_k,\beta_k)$ is called **statistical test at level of significance α_k.**

c) Let d_k the value of D_k for $Z_i=z_i$, with $i=1, 2, \ldots, n$. If $d_k \geq \beta_k$ H_0 is rejected with respect to the test T_k, if this not, H_0 is accepted.

d) The sequences for which the hypothesis H_0 is accepted, with respect to each of the T_k tests, are considered "acceptable" with respect to the set of tests.

The principle according to which the hypothesis H_0 is rejected, if $d_k \geq \beta_k$, is based on the following points:

(1) **probabilistic evaluation**: the event $E_k=(D_k \geq \beta_k)/H_0$ has very small probability α_k of occurring. Since he has occurred, that leads to consider the hypothesis H_0 false;

(2) **psychological evaluation**: if $d_k \geq \beta_k$ then the value of D_k is very far from the, "ideal value", zero.

Some authors (see i.e. R. Scozzafava (1989), [12], A. Maturo and A. Piscione (1990), [8]) have made the following observations:

a) the confusion between E_k and the event $A_k=H_0/(D_k \geq \beta_k)$. According to (1), since the probability of E_k is little, the probability of A_k is little too.

Such conclusions lead to paradoxical consequences.

In fact anyone can see that, whatever the finite sequence $\underline{z}$ of lenght n is, if the z_i are numbers with q decimal digits, being prob $(z_i/H_0)=10^{-q}$ and then it result prob $(\underline{z}/H_0)=10^{-nq}$.

Since, for each $\alpha_k>0$, there is a positive integer n such that $10^{-nq}<\alpha_k$, for a fixed α_k it is possible to determine a positive integer m such that, for each $n \geq m$, the event $\underline{z}/H_0$ has probability lower than α_k.

Therefore, following the same procedure described in Par. 1 we are led to reject H_0 whatever $\underline{z}$ would be, as long as $n \geq m$.

b) The disturbing element put in by (2). The psychological point of view lead us "to reject the tail" of distribution of D_k, since in it there are values wich are very far from the ideal value of D_k, "0". On the other hand, from a rigorously probabilistic point of view, if a is a nonnegative real number minor than β_k, and b is a real number such that the event: $E_{ak}=(a \leq D_k \leq b)/H_0$, has a probability α_k, the event E_{ak} may play the same role as the event E_k in the accepting-rejecting game. For example, assuming a=0, the test would lead us to reject the "first part" of the distribution.

c) The lack of consideration of alternative hypotheses to H_0. If we assume that the event $A_{0k}=H_0/(D_k \geq \beta_k)$ has very little probability, that is not a sufficient reason for rejecting H_0. In fact it may happen that there are many alternative hypotheses H_i, with i=0, 1, 2, ..., m, and that all the events $A_{ik}=H_i/(D_k \geq \beta_k)$ have very little probability. In this case, then, if $d_k \geq \beta_k$, should all the hypotheses be rejected?

At this point the replacement of the classical checking methods by a Bayesian one, is proposed, since the latter gives the following advantages:

a) apart from the hypotheses H_0, it takes into consideration a set of alternative hypotheses to be compared with H_0;

b) it only gives probabilistic outcomes, without recurring to procedures which do not belong to the probability theory.

Therefore the problem is how to find Bayesian analysis which can replace classical models. However, we cannot think of replacing them with the Bayesian analysis based on numbers frequency alone, in fact this may lead to erroneus results.

Using such procedures it may appear that the banal sequences, wich are not random, such as $\{z_i=(i \bmod m)/m\}_{i\in N}$, look better than others. These drawbacks, however, are not sufficient for refusing the Bayesian approach.

2. BAYESIAN ANALYSIS

In order to check the randomness of pseudorandom sequences from a bayesian viewpoint, it is suitable to replace each classical test with a bayesian one. For example, a sample randomness check can be done by replacing each of the eleven quoted tests with a bayesian analysis.

Therefore, the following general procedure is proposed.

From the pseudorandom numbers finite sequence $\underline{z}=\{z_1, z_2, \ldots, z_n\}$, obtained as shown in Par. 1, another one of the kind:

$$T=\{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_{n-r}\}$$

comes out, where r is a suitable positive integer, r<<n, with the $\underline{u}_i$ vectors of R^k, $k \geq 1$, functions of the vector $(z_1, z_2, \ldots, z_n) \in R^n$ and realizations of the k-dimensional random variables $\underline{U}_i$ which are equidistributed and independents.

The $\underline{U}_i$ depend on certain parameters, and for their suitables values, they have the same distribution as those requested by the classical test replaced.

After all suppose that:

(1) the vectors $\underline{u}_i$ are realizations of random variables $\underline{U}_i^{(\alpha_1,\alpha_2,\dots,\alpha_h)}$ which are equidistributed, independent and functions of the parameters $(\alpha_1, \alpha_2, \dots, \alpha_h)$;

(2) in ideal conditions of randomness, the parameters $\alpha_1, \alpha_2, \dots, \alpha_h$ take the values $\mu_1, \mu_2, \dots, \mu_h$.

Let A be the set of the values taken by the parameter $\underline{\alpha}=(\alpha_1, \alpha_2, \dots, \alpha_h)$. Suppose that on a suitable subalgebra of $\mathscr{P}(A)$ to which belong all the elementary events $\{a\}$, with $a \in A$, a comparative probability has been assigned by a relation of "not more possible" (see i.e. Scozzafava (1984), [13], and Maturo (1989), [9]).

Suppose, moreover, that, being $\mathscr{F}$ the collection of finite subsets of A, a positive real function f, defined in A, and called pseudodensity, has been assigned.

The function f is strictly compatible with the comparative probability definition, being such that, for each pair of events of $\mathscr{F}$, $E_1=\{a_1, a_2, \dots, a_n\}$, $E_2=\{b_1, b_2, \dots, b_m\}$, $a_i \in A$, $b_j \in A$, $i=1, 2, \dots, n$, $j=1, 2, \dots, m$, it follows:

$$\sum_i^n f(a_i) \le \sum_j^m f(b_j) \Leftrightarrow E_1 \text{ is not more possible than } E_2. \qquad (2.1)$$

Let us consider two possibilities:

(a) A is a countable set, $A=\{a_1, a_2, \dots, a_n, \dots\}$;

(b) A is a measurable R^h-set with non-zero measure, and $\int_A f(\underline{\alpha})\, d(\underline{\alpha})$ exists, being either finite or not.

If (a) is true compute the sum of the series $\sum_{i=1}^{\infty} f(a_i)$.

If this sum is finite and equals to a real number M, then being: $p(a_i)=f(a_i)/M$, the function: $p: a_i \in A \rightarrow p(a_i)$ is a probability.

If (b) is true and $\int_A f(\underline{\alpha})\, d(\underline{\alpha})$ is a finite real number M, then: $p(\underline{\alpha})=f(\underline{\alpha})/M$, is probability density function (pdf).

In the following the simbol $\int_A f(\underline{\alpha})\, d(\underline{\alpha})$ will indicate also: $\sum_{i=1}^{\infty} f(a_i)$, and the expression "probability density function" to name also the probabilities in the (a) case.

In many papers R. Scozzafava (see i.e. [13]), has shown that the Bayes theorem, valid for prior distributions given by probabilities or probability density functions can be generalized to the case in which the prior distribution is simply given by a pseudodensity.

In the case at hand, if $f(\underline{\alpha})$, $\underline{\alpha}=(\alpha_1, \alpha_2, \dots, \alpha_h) \in A$, is the prior pseudodensity, and $p(\underline{u}/\underline{\alpha})$ is the pseudodensity of $\underline{U} = (U_1, U_2, \dots, U_n)$ conditional on $\underline{\alpha}$, the generalized Bayes theorem takes up the form:

$$g(\underline{\alpha}/\underline{u}) = k\ f(\underline{\alpha})\ p(\underline{u}/\underline{\alpha}) \tag{2.2}$$

where k is a function of $\underline{u}$ and $g(\underline{\alpha}/\underline{u})$ is the posterior pseudodensity.

If $g(\underline{\alpha}/\underline{u})$ is a pdf it then follows:

$$k = \frac{1}{\int_A g(\underline{\alpha}/\underline{u})\, d\underline{\alpha}}.$$

By applying this formula (2.2), the problem is to estimate if $g(\underline{\alpha}/\underline{u})$ "aims to help" the value $\underline{\mu}=(\mu_1, \mu_2, \ldots, \mu_h)$.

For this tree conventional operative criteria can be followed:

1. Criterion of the distance of the absolute maximum of g from the ideal point $\underline{\mu}$

The extreme absolute point $\underline{\beta}$ of the function $g(\underline{\alpha}/\underline{u})$ can be identified, under the hypothesis of existence and uniqueness, and then the distance between $\underline{\beta}$ and $\underline{\mu}$ can be evaluated. According to the Bayesian analysis, the sequences with a minor distance are to be preferred.

The procedure then requires:

a) the identification the collection of the lines passing through $\underline{\mu}$:

$$\begin{cases} x_1 = \mu_1 + \tau_1 t \\ x_2 = \mu_2 + \tau_2 t \\ \ldots\ldots\ldots & , t \in R \\ \ldots\ldots\ldots \\ x_h = \mu_h + \tau_h t \end{cases}$$

with $\tau_1, \tau_2, \ldots, \tau_h$ satisfying the condition:

$$\tau_1^2 + \tau_2^2 + \ldots + \tau_h^2 = 1.$$

This means that for t=0 we obtain the point $\underline{\mu}$ and that $|t|$ is equal to the distance of the point $\underline{x}=(x_1, x_2, \ldots, x_h)$ from $\underline{\mu}$.

b) The search of the extreme point of g on a suitable number of such lines by solving problems with one variable functions and by comparing the outcomes obtained. If g is a regular function, h directions pairwise orthogonals can be sufficient.

c) Being $\underline{\mu}^{(1)}$ the extreme point of the constriction of g at the union of the considered lines (if there are many extreme points, the one with the smaller $|t|$, will be considered), assume a positive number "small" σ and consider the distance $\delta^{(1)}$ between $\underline{\mu}^{(1)}$ and $\underline{\mu}$.

If $\delta^{(1)}<\sigma$ the generator is considered "good" with respect to the adopted criterion, but, in order to "accept" it, further analysis must be carried out, where further factors of randomness are taken into consideration.

2. Criterion of probability in the intervals or in the circles with center $\underline{\mu}$

Let $g(\underline{\alpha}/\underline{u})$ be a probability density function.

Fixed either a suitable interval or circle I, with center $\underline{\mu}$, calculate the probability that $\underline{\alpha}$ belongs to I. According to the bayesian analysis the sequences with a major probability are preferred.

3. Criterion of the confidence interval or confidence circle

Fixed a level of probability θ and calculated the extreme point $\underline{\beta}$ of the function $g(\underline{\alpha}/\underline{u})$ we determine an interval or circle I_θ with center $\underline{\beta}$ such that the probability that $\underline{\alpha}$ belongs to I_θ is equal to θ.

It has to be admitted that the sequence $\underline{u}$ is pseudorandom, according to the proceeding analysis and at confidence level θ, if and only if $\theta \in I_\theta$.

Let us consider some of such analyses:

1. Bayesian analysis of the frequences; let's $\underline{u}_i=z_i$, for each $i \in \{1, 2, \dots, n\}$;

2. Bayesian analysis of the pairs; for each $i \in \{1, 2, \dots, n-1\}$ let's $\underline{u}_i=[z_i, z_{i+1}]'$. It results: $r=1$;

3. Bayesian analysis of the k-ple; it is a generalization of the previous analyses. Let's $\underline{u}_i=[z_i, z_{i+1}, \dots, z_{i+k-1}]'$, for each $i \in \{1, 2, \dots, n-k+1\}$. It results: $r=k-1$;

4. Bayesian analysis of the runs; consider m sequences $\underline{z}^j = \{z^j_1, \dots, z^j_n\}$ and suppose that for each i and j $z^j_i \neq z^j_{i+1}$ (it is in pratice always true). Assume:

$$\underline{u}_{ij} = \begin{cases} 0 \text{ for } i>1 \text{ and } z^j_i < z^j_{i+1} < z^j_{i+2} \text{ or } z^j_i > z^j_{i+1} > z^j_{i+2} \\ 1 \text{ otherwise} \end{cases}$$

for each $i \in (1, 2, \dots, n-2\}$ and for each $j \in (1, 2, \dots, m\}$.

For each sequence $\underline{z}^j$ the number (called runs number)

$$v_j = \sum_{i=1}^{n} u_{ji}$$

can be considered. This numbers can be considered as independents.

The ideal distribution of the v_j (see i.e. A. Rizzi (1977), [10]) is the gaussian one, with mean (2n-1)/3 and variance (16n-29)/90.

The parameters on wich the inference procedure is made are the mean θ and the variance σ^2.

Assuming $\underline{\alpha}=(\theta,\sigma)$, $\underline{v} = (v_1, v_2, \ldots, v_m)$, the relation (2.2) is replaced by:

$$g(\underline{\alpha}/\underline{v}) = k\ f(\underline{\alpha})\ p(\underline{v}/\underline{\alpha}).$$

The Bayesian analysis, illustrated in this work, equivalent to the frequency classical test, can be described in the following way:

(1) let us consider the hypothesis H_0 as an element of a collection of hypotheses $(H_\alpha)_{\alpha\in A}$, where A is a set of indexes, such that for a certain value μ of $A{:}H_\mu = H_0$ then results. Then, in a subjective way and following suitable criteria, let us assign, to each hypothesis H_α, the value $f(\alpha)$, $\alpha\in A$, of a probability density function defined in A or, more generally, of a pseudodensity.

(2) Starting from a given sequence $\underline{z}=\{z_1, z_2, \ldots, z_n\}$ of pseudorandom numbers varying between [0,1), calculate the likelihoods $p(\underline{z}/H_\alpha)$, $\alpha\in A$;

(3) using the Bayes formula

$$g(H_\alpha/\underline{z}) = k\ f\ (H_\alpha)\ p(\underline{z}/H_\alpha), \tag{2.3}$$

with k function only of $\underline{z}$, calculate the posterior probability density function (or more generally, the pseudodensity);

(4) by suitable conventional criteria, for example **1**, **2** or **3**, consider if the hypothesis $\alpha=\mu$ is preferred.

Starting from the H_0 hypothesis of Z_i independence and equidistribution, it follows that each sequence S' obtained as a permutation of $\underline{z}$, originates the same likelihoods and, therefore, the same g values.

This implies that the Bayes theory **gives information only about the frequences** of the z_i values taken by the Z_i, and therefore **it can replace only the frequency classical test.**

However, such Bayesian analysis is just one particolar case of those we have considered in the $\underline{u}_i=z_i$ hypotheses.

Assuming other hypoteses for the $\underline{u}_i$, we obtain further Bayesian analysis which can replace further classical tests.

Now an example of Bayesian analysis based on frequences will be illustrated.

3. AN EXAMPLE OF BAYES ANALYSIS AS FREQUENCY TEST

Assume each random variable Z_i, (i=1, ..., n), varying in the interval [0,1), and distribuited as a Beta one, with parameters: a,b>0. The probability density function is then:

$$p(z;a,b)=\begin{cases}[1/B(a,b)]z^{a-1}(1-z)^{b-1}, \text{ for } z\in(0,1)\\ 0 \text{ for } z\notin(0,1)\end{cases} \tag{3.1}$$

where $B(a,b)=\int_0^1 z^{a-1}(1-z)dz$

The choice can be justified accordingly to the following points:

(i) the Beta distribution varies in the interval (0,1), like any element of the pseudorandom number sequence;

(ii) assuming a determinated set of values for the parameters (a=b=1), the Beta law is trasformed in the continuous uniform distribution, as it happens following null hypothesis.

Since the a and b parameters of the Beta distribution are unknown, they can be thought as random numbers on which the statistical inference is drawn, with the aim of obtaining a posterior probability density function.

In the inference procedure, the examined parameters are therefore collected in a random vector $\underline{\theta}$=(a,b).

As a prior distribution for each of the two parameters, the Gamma law with two parameters τ, μ>0, can be assumed whose pdf is expressed by:

$$h(x;\tau,\mu)=\begin{cases}[\tau/\Gamma(\mu)]e^{-\tau x}(\tau x)^{\mu-1}, \text{ for } x>0\\ 0 \text{ for } x<0\end{cases} \tag{3.2}$$

If the two components of $\underline{\theta}$ can be considered as independent, the probability density function of the random vector $\underline{\theta}$=(a,b) is the following one:

$$f(a,b;\tau,\mu)=[\tau^{\mu}/\Gamma(\mu)]e^{-\tau a}a^{\mu-1}[\tau^{\mu}/\Gamma(\mu)]e^{-\tau b}b^{\mu-1}, \qquad (3.3)$$

with $(a,b)\in(0,+\infty)\text{x}(0,+\infty)$.

The partial derivatives of the function f are continuous in the set: $A=(0,+\infty)\text{x}(0,+\infty)$; then the extreme points correspond to the ones that make nul these derivatives. It follows therefore:

$$\frac{\partial f}{\partial a}=0 \text{ for } a=\frac{\mu-1}{\tau}, \qquad \frac{\partial f}{\partial b}=0 \text{ for } b=\frac{\mu-1}{\tau} \qquad (3.4)$$

Then, only one extreme point of f function exists, this being the point: $P=((\mu-1)/\tau,(\mu-1)/\tau))$.

It is therefore:

$$\frac{\partial^2 f}{\partial a^2}=\left[(\tau^{\mu}/\Gamma(\mu))\right]^2 b^{\mu-1}e^{-\tau b}e^{-\tau a}a^{\mu-3}\left[-2\mu(\mu-1)+\mu^2a^2+(\mu-1)(\mu-2)\right]$$

$$\frac{\partial^2 f}{\partial b^2}=\left[(\tau^{\mu}/\Gamma(\mu))\right]^2 a^{\mu-1}e^{-\tau a}e^{-\tau b}b^{\mu-3}\left[-2\mu(\mu-1)+\mu^2b^2+(\mu-1)(\mu-2)\right]$$

$$\frac{\partial^2 f}{\partial a\,\partial b}=\left[(\tau^{\mu}/\Gamma(\mu))\right]^2[(\mu-1)a^{\mu-2}e^{-\tau a}-\tau e^{-\tau a}a^{\mu-1}]b^{\mu-1}e^{-\tau b}[(\mu-1)b-\tau]$$

Assuming that the partial derivatives are equal to zero if valuated in (1,1), it results that:

$$\tau=\mu-1 \qquad (3.5)$$

A direct consequence of this hypothesis is that:

$$\frac{\partial^2 f}{\partial a^2}<0, \quad \frac{\partial^2 f}{\partial b^2}<0, \quad \frac{\partial^2 f}{\partial a\,\partial b}=0$$

and therefore the Hessian matrix:

$$H=\begin{bmatrix} \dfrac{\partial^2 f}{\partial a^2} & \dfrac{\partial^2 f}{\partial a\,\partial b} \\ \dfrac{\partial^2 f}{\partial a\,\partial b} & \dfrac{\partial^2 f}{\partial b^2} \end{bmatrix}$$

is negative definite. Therefore P is a maximum for f, and this hypothesis corresponds to substantiate the null hypothesis: a=1, b=1.

Denoting by $p(\underline{z}/(a,b))$ the pdf of $\underline{Z}=(Z_1, Z_2, \ldots, Z_n)$, conditional on $\underline{\theta}=(a,b)$, according to the Bayes theorem, it follows therefore:

$$g((a,b)/\underline{z}) = k\, f(a,b;\, \tau,\mu)\, p(\underline{z}/(a,b)) \qquad (3.6)$$

where, if $g((a,b)/\underline{z})$ is a pdf,

$$k = \frac{1}{\iint_A g((a,b)/\underline{z})}\, da\, db. \qquad (3.7)$$

According to the independence condition:

$$p(\underline{z}/(a,b)) = \prod_{i=1}^{n} p(z_i/(a,b)) =$$
$$= [\Gamma(a+b)/\Gamma(a)\Gamma(b)]^n (z_1 z_2 \ldots z_n)^{a-1}[(1-z_1)(1-z_2)\ldots(1-z_n)]^{b-1}.$$

Therefore, considering the relation (3.5), it follows that:

$$g((a,b)/\underline{z}) = k[\tau^{\mu}/\Gamma(\mu)]^2 (ab)^{\tau} e^{-\tau(a+b)} *$$
$$*[\Gamma(a+b)/\Gamma(a)\Gamma(b)]^n (z_1 \ldots z_n)^{a-1}[(1-z_1)\ldots(1-z_n)]^{b-1}. \qquad (3.8)$$

Let assume that $\underline{\theta}$ has a uniform distribution. This can be derived interpreting the relation (3.3) as a pseudodensity.

Therefore, by assuming two arbitrary points, belonging to A, (a_1, b_1) and (a_2, b_2), calculate the limit:

$$\lim_{\tau\to 0} \frac{f(a_1, b_1; \tau, \mu)}{f(a_2, b_2; \tau, \mu)}.$$

It results equal to:

$$\lim_{\tau\to 0} \frac{[\tau^{\mu}/\Gamma(\mu)]e^{-\tau a} a_1^{\mu-1}[\tau^{\mu}/\Gamma(\mu)]e^{-\tau b} b_1^{\mu-1}}{[\tau^{\mu}/\Gamma(\mu)]e^{-\tau a} a_2^{\mu-1}[\tau^{\mu}/\Gamma(\mu)]e^{-\tau b} b_2^{\mu-1}} =$$
$$= \left[\frac{(a_1 b_1)}{(a_2 b_2)}\right]^{\mu-1}. \qquad (3.9)$$

The limit of the pseudodensity expressed by (3.3), is, a part from a costant:

$$f(a,b;\ 0,\mu) = (ab)^{\mu-1}. \tag{3.10}$$

According to the relation (3.5), being τ=0, it then results: μ=1. The limit (3.9) then equals 1, and the pseudodensity (3.10) is the uniform one.

It is easy to derive that the one expressed by the (3.10) is not a pdf. Infact, being μ>0, it follows that:

$$\iint_A (ab)^{\mu-1}\, da\, db = +\infty.$$

Nonetheless, the Bayes Theorem (3.6) is valid altogether, also if the condition (3.7) is not satisfied.

In the particular case: τ=0, assuming f(a,b; τ,μ) as the pseudodensity, expressed by the relation (3.10) it results that:

$$f(a,b;\ 0,\mu) = (ab)^{\mu-1}$$

and the (3.8) is equal to:

$$\begin{aligned} g((a,b)/\underline{z}) &= k\,(ab)^{\mu-1}[\Gamma(a+b)/\Gamma(a)\Gamma(b)]^n * \\ &*(z_1z_2\ldots z_n)^{a-1}[(1-z_1)(1-z_2)\ldots(1-z_n)]^{b-1}. \end{aligned} \tag{3.11}$$

If the (3.5) is valid, the (3.11) then reduces to:

$$\begin{aligned} g((a,b)/\underline{z}) &= k\,[\Gamma(a+b)/\Gamma(a)\Gamma(b)]^n * \\ &*(z_1\ldots z_n)^{a-1}[(1-z_1)\ldots(1-z_n)]^{b-1}. \end{aligned} \tag{3.12}$$

The numerical application and its results of this procedure based on the criterion of the distance of the absolute maximum of g from the ideal point $\underline{\mu}$ can be found in G. Di Biase and A. Maturo (1990), [5].

REFERENCES

1. B. Baldessari (1987). *Aspetti probabilistici della crittografia*, Atti del I Simposio Nazionale su "Stato e Prospettive della Ricerca Crittografica in Italia", pp. 9-21.
2. N. Cera e A. Maturo (1983). *Confronto fra alcuni generatori di numeri pseudocasuali*, Facoltà di Architettura, Pescara.

3. N. Cera e A. Maturò (1983). *Generazione di numeri pseudocasuali per mezzo di relazioni di ricorrenza in campi di Galois*, Facoltà di Architettura, Pescara.
4. N. Cera e A. Maturo (1990). *Analisi della bontà di alcuni generatori di numeri pseudocasuali per la cifratura dei messaggi e la simulazione*, Ratio Mathematica, 2, pp. 83-86.
5. G. Di Biase e A. Maturo (1990). *Su un'analisi bayesiana per la verifica di casualità di successioni pseudorandom: software applicativo ed interpretazione dei risultati.* Atti del Convegno "Classificazione e analisi dei dati, Metodi, Software, Applicazioni", 11-12 ott. 1990, Pescara, pp. 115-129.
6. D.E. Knuth (1969). *The art of Computer programming vol. 2*, Addison Wesley, London.
7. A. Maturo (1989). *Numeri pseudocasuali*, Libreria dell'Università, Pescara.
8. A. Maturo e A. Piscione (1990). *Probabilità e statistica con il calcolatore: problematiche di carattere logico ed operativo,* Metron, vol. XLVIII, n. 14, pp. 509-532.
9. A. Maturo (1989). Probabilità finitamente additive a valori nel campo delle serie bilatere, Rendiconti di Matematica VII, 9, pp. 67-85.
10. A. Rizzi (1977). *Generazione di distribuzioni statistiche mediante un elaboratore elettronico*, Istituto di Statistica e Ricerca sociale C. Gini, Roma.
11. A. Rizzi (1989). *Verifiche di Pseudo-Casualità in Crittografia*, Atti del 2º Simposio Nazionale su "Stato e Prospettive della Ricerca Crittografica in Italia", pp. 3-21.
12. R. Scozzafava (1989). *La probabilità soggettiva e le sue applicazioni*, Masson, Milano.
13. R. Scozzafava (1984). *A Survey of Some Common Misunderstandings Concerning the Role and Meaning of Finitely Additive Probabilities in Statistical Inference*, Statistica, 44, pp. 21-45.
14. R.C. Tausworthe (1965). *Random numbers generated by linear recurrence modulo two*, Mathematics of Computation, 19.