Classical Limit Theorems and High Entropy MIXMAX random number generator

Hayk Poghosyan¹, Konstantin Savvidy², and George Savvidy³

¹ Yerevan Physics Institute, 2 Alikhanyan Br., 375036, Yerevan, Armenia (E-mail: haykpoghos@gmail.com)
² Demokritos National Research Center, Ag. Paraskevi, Athens, Greece (E-mail: k.savvidis@cern.ch)
³ Demokritos National Research Center, Ag. Paraskevi, Athens, Greece (E-mail: savvidy@inp.demokritos.gr)

Abstract. We investigate the interrelation between the distribution of stochastic fluctuations of independent random variables in probability theory and the distribution of time averages in deterministic Anosov C-systems. On the one hand, in probability theory, our interest dwells on three basic topics: the laws of large numbers, the central limit theorem and the law of the iterated logarithm for sequences of real-valued random variables. On the other hand we have chaotic, uniformly hyperbolic Anosov C-systems defined on tori which have mixing of all orders and nonzero Kolmogorov entropy. These extraordinary ergodic properties of Anosov C-systems ensure that the above classical limit theorems for sums of independent random variables in probability theory are fulfilled by the time averages for the sequences generated by the C-systems. The MIXMAX generator of pseudorandom numbers represents the C-system for which the classical limit theorems are fulfilled.

Keywords: .

1 Introduction

Our intention in this article is to consider the behaviour of deterministic Anosov C-systems in parallel with the classical limit theorems of probability theory, demonstrating that they possess the properties which are inherent to the independent and identically distributed random variables defined in probability theory.

We investigate the interrelation between the distribution of stochastic fluctuations of independent random variables in probability theory and the distribution of time averages in deterministic Anosov C-systems. On the one hand, in probability theory, our interest dwells on three basic topics: the laws of large numbers, the central limit theorem and the law of the iterated logarithm for sequences of real-valued random variables [1–5,7–13]. On the other hand we have chaotic, uniformly hyperbolic Anosov C-systems defined on tori which
have mixing of all orders and nonzero Kolmogorov entropy [14–26]. These extraordinary ergodic properties of Anosov C-systems ensure that the above classical limit theorems for sums of independent random variables in probability theory are fulfilled by the time averages for the sequences generated by the C-systems [49–52]. The MIXMAX generator of pseudorandom numbers represents the homogeneous C-system for which the classical limit theorems are fulfilled [25–27,41,42].

The present paper is organised as follows. In section two we shall overview the classical limit theorems in probability theory. In section three the basic properties of the Anosov C-systems will be defined, their spectral properties and the entropy will be presented. In section four a parallel between the classical limit theorems of probability theory and behaviour of deterministic dynamical C-systems will be derived and the mapping dictionary between the two systems will be established. We shall analyse the law of large numbers, the central limit theorem and law of the iterated logarithm in the case of C-system MIXMAX generator of pseudorandom numbers. The C-system nature of the MIXMAX generator provides well define mathematical background and guaranty the uniformity of generated sequences.

2 Classical Limit Theorems in Probability Theory

Consider an infinite sequence of independent and identically distributed random variables \( \xi_1(x), \xi_2(x), \ldots \) on the interval \( 0 \leq x \leq 1 \) having finite mean values \( M_{\xi_k} = \mu \) and finite variance \( \sigma^2 = M((\xi_k - \mu)^2) \) \( (0 < \sigma^2 < \infty) \) [1]. One of the fundamental questions of interest in probability theory is the limiting behaviour of the sum [1–5,7,13]

\[
S_n = \xi_1 + \xi_2 + \ldots + \xi_n = \sum_{k=1}^{n} \xi_k
\]

as \( n \to \infty \). By the classical central limit theorem the difference between the average \( S_n/n \) and \( \mu \) multiplied by the factor \( \sqrt{n} \) converges in probability to the normal distribution \( \Phi(\frac{z}{\sigma}) \)

\[
P \left\{ \sqrt{n}(\frac{S_n}{n} - \mu) < z \right\} \to \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{z} e^{-\frac{y^2}{2\sigma^2}} \, dy \quad \text{for every } z. \tag{2}
\]

The estimates of the convergence rate in the above central limit theorem were obtained by Lyapunov, Berry, Esseen and others [8–12]. For independent and identically distributed random variables having finite absolute third moments \( \chi^3 = M_{\xi_k^3} < \infty \) it has the form:

\[
\sup_{z} \left| P \left\{ \sqrt{n}(\frac{S_n}{n} - \mu) < z \right\} - \Phi(\frac{z}{\sigma}) \right| \leq \frac{1}{\sqrt{n}} \left( \frac{\chi}{\sigma} \right)^3 . \tag{3}
\]

By the Kolmogorov strong law of large numbers the average \( S_n/n \) converges almost surely to the common mean value \( \mu \) of the random variables \( \xi_k(x) \), that is

\[
P \left\{ \lim_{n \to \infty} \frac{S_n}{n} = \mu \right\} = 1 . \tag{4}
\]
Under the same conditions as in the above theorems Petrov [13] has derived the estimates of the order of growth of the sums $S_n$ (1). The following growth estimates take place

$$
P \left\{ \lim_{n \to \infty} \frac{S_n - \mu n}{n^{1/2} + \epsilon} = 0 \right\} = 1, \tag{5}
$$

$$
P \left\{ \lim_{n \to \infty} \frac{S_n - \mu n}{n^{1/2} (\ln n)^{1/2} + \epsilon} = 0 \right\} = 1, \tag{6}
$$

$$
P \left\{ \lim_{n \to \infty} \frac{S_n - \mu n}{n^{1/2} \ln n (\ln \ln n)^{1/2} + \epsilon} = 0 \right\} = 1, \tag{7}
$$

for arbitrary $\epsilon > 0$. This result means that the random variables $S_n - \mu n$ cannot grow faster than $n^{1/2 + \epsilon}$ or $n^{1/2} (\ln n)^{1/2 + \epsilon}$ or $n^{1/2} (\ln \ln n)^{1/2 + \epsilon}$ and so on. The theorem on the law of the iterated logarithm for a sequence of random variables $\{\xi_k\}$ involve conditions under which the sequence $\lim_{n \to \infty} \sup S_n - \mu n \sqrt{\frac{2}{\ln \ln n}} = \sigma$ converges almost surely. This relation strengthens the estimates provided by the strong law of large numbers (4) and (5). For the independent and identically distributed random variables (1) the following Kolmogorov theorem of the iterated logarithm take place [2–4]

$$
P \left\{ \lim_{n \to \infty} \sup S_n - \mu n \sqrt{\frac{2}{\ln \ln n}} = \sigma \right\} = 1, \tag{8}
$$

$$
P \left\{ \lim_{n \to \infty} \inf S_n - \mu n \sqrt{\frac{2}{\ln \ln n}} = -\sigma \right\} = 1, \tag{9}
$$

that is a maximal possible growth of the sum is $\sigma \sqrt{2n \ln \ln n}$. In order to gain an intuitive understanding of this result it is worth to calculate the probability of large fluctuations of the sum $S_n$ using the central limiting theorem (2). It follows from (2) that for the arbitrary positive number $b$, $z = \sigma b \sqrt{2 \ln \ln n}$ and large $n$ take place the following relation

$$
P \left\{ \sqrt{n} \left( \frac{S_n}{n} - \mu \right) \geq \sigma b \sqrt{2 \ln \ln n} \right\} \to 1 - \frac{1}{\sqrt{2\pi \sigma^2}} \int_{-\infty}^{\sigma b \sqrt{2 \ln \ln n}} e^{-\frac{y^2}{2\sigma^2}} dy, \tag{10}
$$

where the right hand side has the asymptotic $1/2b(\pi \ln \ln n)^{1/2}(\ln n)^{b^2}$ and therefore

$$\frac{1}{(\ln n)^{(1+\delta)b^2}} \leq P \left\{ \frac{S_n(x) - \mu n}{\sqrt{2n \ln \ln n}} \geq b \sigma \right\} \leq \frac{1}{(\ln n)^{b^2}}. \tag{11}
$$

Considering the subsequence $n = q^m$, where $q$ is a fixed integer, one can derive from (8) the celebrated law of the iterated logarithm (6) [2–5,13,7]. The law of the iterated logarithm is a refinement of the law of large numbers (4) and specifies the global behaviour of the asymptotic sequence of the sum $S_n$ since the quantity in the limit in (6) depends not only on single $n$ but the totality of the remainder of the sum. Using the central limiting theorem (2) now for the fluctuations in the interval $(-\epsilon \sqrt{2 \ln \ln n}, +\epsilon \sqrt{2 \ln \ln n})$ one can get that

$$
P \left\{ \left| \frac{S_n(x) - \mu n}{\sqrt{2n \ln \ln n}} \right| < \epsilon \right\} \to \frac{1}{\sqrt{2\pi \sigma^2}} \int_{-\sigma \epsilon \sqrt{2 \ln \ln n}}^{\sigma \epsilon \sqrt{2 \ln \ln n}} e^{-\frac{y^2}{2\sigma^2}} dy \to 1, \tag{12}
$$
meaning that the sum (1) scaled by the factor $\sqrt{2n \ln \ln n}$ is less than any $\epsilon > 0$ with probability approaching one, but will be occasionally visiting points in the interval $(-\sigma, \sigma)$ in accordance with the theorem (6).

Our goal is to compare the asymptotic behaviour of the sum $S_n$ which have been establish in the above limit theorems in probability theory with the asymptotic behaviour of the corresponding quantities defined for deterministic dynamical C-systems and specifically for the C-system which have been implemented into the MIXMAX generator [41,42,25–27].

### 3 Classical Limit Theorems and Deterministic C-systems

\begin{figure}
\centering
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{freq_dist}
\caption{(a) The frequency distribution histogram for the underlying variable $\phi_i(f,n)$ in (23). The dimension of the C-system generator is $N = 17$, the iteration time is $n = 10^4$, the bins are of equal size $\epsilon = 0.01$ and the total number of phase space points $x$ is $I = 10^6$. The function $f(x) = \cos 2\pi(x_1 + \ldots + x_{17})$. The mean value is $\langle \phi \rangle = 0.000470253$ and the standard deviation $\langle \phi^2 \rangle = \sigma_f^2 = 0.499867$.}
\end{subfigure}
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{cdf}
\caption{(b) The p-value of the cumulative distribution function (CDF) for the Kolmogorov-Smirnov test is $p = 0.909337$.}
\end{subfigure}
\end{figure}

Fig. 1. a) The frequency distribution histogram for the underlying variable $\phi_i(f,n)$ in (23). The dimension of the C-system generator is $N = 17$, the iteration time is $n = 10^4$, the bins are of equal size $\epsilon = 0.01$ and the total number of phase space points $x$ is $I = 10^6$. The function $f(x) = \cos 2\pi(x_1 + \ldots + x_{17})$. The mean value is $\langle \phi \rangle = 0.000470253$ and the standard deviation $\langle \phi^2 \rangle = \sigma_f^2 = 0.499867$. b) The p-value of the cumulative distribution function (CDF) for the Kolmogorov-Smirnov test is $p = 0.909337$.

With that aim let us now consider the statistical properties of deterministic dynamical C-systems. The hyperbolic Anosov C-systems defined on a torus have mixing of all orders and nonzero Kolmogorov entropy [14–25,27]. The statistical properties of a C-system defined by the map {∀ $x \in M : x \rightarrow x_n = T^n x$} are characterised by the behaviour of the correlation functions of observables {f(x)} on the phase space $M$

$$D_n(f, g) = \langle f(x)g(T^n x) \rangle - \langle f(x) \rangle \langle g(x) \rangle,$$

where $T$ denotes the Anosov C-system and $\langle \ldots \rangle$ the phase space averages [24,27,45,46,27,45,41,26,42]. These correlation functions decay exponentially, meaning that the observables on the phase space become independent and
uncorrelated exponentially fast \[28–40\]. For the C-systems defined on the \(N\)-dimensional torus \(19\) the upper bound on the exponential decay of the correlation functions is universal and is defined by the value of the system entropy \(h(T)\) \[25\]:

\[
|D_n(f,g)| \leq C e^{-nh(T)\nu},
\]

where \(C = C(f,g)\) and \(\nu = \nu(f,g)\) depend only on the observables and are positive numbers. This result allows to define the decorrelation time \(\tau_0\) for the observable \(f(x)\) as \[25\]

\[
\tau_0 = \frac{1}{h(T)\nu_f},
\]

where \(\nu_f = \nu(f,f)\). The index \(\nu_f\) is increasing linearly \(\nu_f = 2pN\) with the dimension \(N\) of the C-system, where \(p\) is the order of smoothness of the function \(f(x)\) \[25\]. The entropy \(h(T)\) is also increases linearly \(h(T) = \frac{2}{\pi} N\) \[42\], therefore \[25\]

\[
\tau_0 = \frac{\pi}{4pN^2} .
\]

This result justifies the statistical/probabilistic description of the C-systems \[24\] and have important consequences in the form of the law of large numbers and central limit theorem for Anosov C-systems \[49–52\]. The time average of the observable \(f(x)\) on \(\mathcal{M}\)

\[
\bar{f}_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)
\]

behaves as a superposition of quantities which are statistically independent, therefore \[14,49\]

\[
\lim_{n \to \infty} \bar{f}_n(x) = \langle f \rangle
\]

and the fluctuations of the time averages \(14\) from the phase space average \(\langle f \rangle\) multiplied by \(\sqrt{n}\) have at large \(n \to \infty\) the Gaussian distribution \[48–52\]:

\[
\lim_{n \to \infty} m\left\{ x : \sqrt{n}\left( \bar{f}_n(x) - \langle f \rangle \right) < z \right\} = \frac{1}{\sqrt{2\pi\sigma_f^2}} \int_{-\infty}^{z} e^{-\frac{y^2}{2\sigma_f^2}} dy ,
\]

where \(m\) is the invariant measure on the phase space \(\mathcal{M}\) and the value of the standard deviation \(\sigma_f\) is a sum

\[
\sigma_f^2 = \langle f^2(x) \rangle - \langle f(x) \rangle^2 + 2 \sum_{n=1}^{+\infty} [\langle f(T^n x)f(x) \rangle - \langle f(x) \rangle^2].
\]

These results allow to trace a parallel between the classical limit theorems of probability theory and behaviour of deterministic dynamical C-systems. The theorems \((15)\) and \((16)\) which taking place for the deterministic C-systems
are in fine analogy with the theorems (4) and (2) in probability theory. This analogy can be made explicit if one use the dictionary:

\[ \xi_k(x) \iff f(T^k x) \]

\[ S_n(x) \iff \sum_{k=0}^{n-1} f(T^k x) \]

\[ \frac{S_n(x)}{n} \iff \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x). \tag{18} \]

In the next section we shall consider the fine examples of the C-systems defined on the tori. These systems represent a large class of C-systems which can be easily realised on a computer platform in the form of computer algorithms. These algorithms are used to generate pseudorandom numbers of high quality and represent the called MIXMAX pseudorandom number generators \cite{27,41,26,54–56} and in particular the generator passes all statistical U01 tests \cite{53}.

4 MIXMAX C-systems Generator

\[ \text{Fig. 2. The frequency distribution histogram for the variable } \chi(t) \text{ defined in (25). The "time" interval is taken as } [0.2, 2]. \text{ The other parameters are the same as in Fig. 1.} \]

The linear automorphisms of the unit hypercube \( \mathcal{M}^N \) in Euclidean space \( E^N \) with coordinates \( (x_1, ..., x_N) \)\cite{14,27,41,26} is defined as follows:

\[ x_i^{(k+1)} = \sum_{j=1}^{N} T_{ij} x_j^{(k)} \mod 1, \quad k = 0, 1, 2, ... \tag{19} \]

where the components of the vector \( x^{(k)} \) are \( x^{(k)} = (x_1^{(k)}, ..., x_N^{(k)}) \). The phase space \( \mathcal{M}^N \) of the systems (19) can also be considered as the \( N \)-dimensional
torus \([14,27,41,26]\). The operator \(T\) acts on the initial vector \(x^{(0)}\) and produces a phase space trajectory \(x^{(n)} = T^n x^{(0)}\) on a torus. The C-system is defined by the integer matrix \(T\) which has a determinant equal to one \(\text{Det} T = 1\) and has no eigenvalues on the unit circle [14]:

\[
1) \text{Det} T = \lambda_1 \lambda_2 \ldots \lambda_N = 1, \quad 2) |\lambda_i| \neq 1, \quad \forall \ i.
\]

(20)

The measure \(dm = dx_1 \ldots dx_N\) is invariant under the action of \(T\). The conditions (20) guarantee that \(T\) represents Anosov C-system [14] and therefore as such it is a Kolmogorov K-system [15–19] with mixing of all orders and of nonzero entropy. The C-system (19) has a nonzero Kolmogorov entropy \(h(T)\) [14,17,19–21,26]:

\[
h(A) = \sum_{|\lambda_a|>1} \ln |\lambda_a|.
\]

(21)

We shall consider a family of matrix operators \(T\) of dimension \(N\) introduced in [41]. The operators \(T\) fulfill the C-condition (20) and represents a C-system [27,41,42] with entropy:

\[
h(A) = \sum_a \ln |\lambda_a| \approx \frac{2}{\pi} N
\]

(22)

which is increases linearly with the dimension \(N\) of the matrix. Our aim is to study the asymptotic behaviour of the sum \(S_n\) as \(n \to \infty\) for the pseudorandom number generator MIXMAX [41,42] which is defined by the equations (19).

In order to study the asymptotic behaviour in equation (16) as \(n \to \infty\) we shall consider first the following variable

\[
\phi_i(f,n) = \sqrt{n} \left( \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x_i) - \langle f \rangle \right)
\]

(23)

which depends on initial phase space vector \(x_i\) of the \(N\)-dimensional unit hypercube \(M^N\), the function \(f(x)\) and the number of iterations \(n\). In order to calculate the number of vectors \(x_i, i = 1, \ldots, I\) which fulfill the inequality \(\phi_i(f,n) < z\) we shall construct the frequency distribution of the underlying variable \(\phi_i(f,n)\). The bins will be taken of equal size \(\epsilon\). The Fig. 1 represents the distribution function calculated for the MIXMAX generator of size \(N = 17\) and the comparison with the Gaussian distribution

\[
\rho(\phi) = \frac{I \epsilon}{\sqrt{2 \pi \sigma_f^2}} e^{-\frac{\phi^2}{2 \sigma_f^2}}
\]

(24)

shown on the Fig.2 as a solid blue line. We have used the Kolmogorov-Smirnov test to calculate the \(p\)-value. The \(p\)-value of the cumulative distribution function (CDF) for the Kolmogorov-Smirnov test here was \(p = 0.909337\). The null hypothesis that the data is distributed according to the normal distribution is not rejected at the 0.1\% level based on the Kolmogorov-Smirnov test.
Fig. 3. Two histograms of the variable $\Gamma_n = S_n / \sqrt{2n \ln \ln n}$ in (27). The dimension of the C-system generator is $N = 240$ and the iteration time $n = 10^9$. The total number of the initial phase space points $x_i$ is $I = 200$. Here in $S_n$ data we have subtracted the term $\mu n$. The sum $S_n$ grows approximately as $\sigma \sqrt{n}$. The standard deviation for the observable $f(x) = x$ is equal to $\sigma = 1/\sqrt{12}$. The distribution function of the supremum of the $\Gamma_n$ was calculated for the values $n$ in the interval $n \in [m, 10^9]$ for $m = 15$ and $m = 10^5$. As one can see the distribution of $\Gamma_n$ of the supremum values (28) is tightens towards $\sigma = 1/\sqrt{12}$ from below in accordance with the Kolmogorov law of the iterated logarithm (6). On the first histogram, at $m = 15$, there are 115 events smaller and 90 events larger that $\sigma \approx 0.27$. On the second histogram, at $m = 10^5$, there are 18 events smaller and 8 events larger that $\sigma \approx 0.27$.

Introducing a new parameter $t = p/n$, where $p$ is an integer number $p \in \mathbb{Z}$ and the alternative variable

$$\chi_i(t) = t\sqrt{n} \left( \frac{1}{E_h} \sum_{k=0}^{tn-1} f(T^k x_i) - \langle f \rangle \right)$$  \hspace{1cm} (25)

we can find the distribution function for the variable $\chi$. It was proven that the variable $\chi$ is described in accordance with the Wiener-Feynman process [30]

$$\rho(\chi, t) = \frac{I\epsilon}{\sqrt{2\pi \sigma_f^2 t}} e^{-\frac{\chi^2}{2\sigma_f^2 t}}. \hspace{1cm} (26)$$

On Fig. 2 one can see that with the increasing "time" $t$ the distribution evolve as in (26). The useful analogy will be if one consider $I\epsilon$ as the number of "particle" at the initial time of the diffusion and $D = \sigma_f^2$ as the diffusion coefficient. Considering the central limit theorem we were performing iterations for the relatively small values of $n \approx 10^4$. In order to study the large fluctuations of the sum $S_n$ described by the law of the iterated logarithm we
generated sequences of increasing length \( n = 10^7 - 10^9 \) and then constructed the distribution function of the maximum values of the variable

\[
\Gamma_n(x) = \frac{\sum_{k=0}^{n-1} f(T^k x) - \langle f \rangle_n}{\sqrt{2 n \ln \ln n}}
\]  

at the tail of the sequences

\[
\lim \Gamma := \lim_{m \to \infty} (\sup_{n \geq m} \Gamma_n).
\]  

We illustrated the distribution function for the values \( n \) in the interval \( n \in [m, 10^9] \) for \( m = 15 \) and \( m = 10^5 \) on Fig. 3. As one can see the distribution of the supremum values is tightens towards \( \sigma = 1/\sqrt{12} \).

5 Acknowledgement

This work was supported in part by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 644121.

References


48. M. Kac, *On the Distribution of Values of Sums of the Type \( \sum f(2^k t) \)*, Annals of Mathematics, **47** (1946) 33-49

49. V. P. Leonov, *On the central limit theorem for ergodic endomorphisms of the compact commutative groups*, Dokl. Acad. Nauk SSSR, **124** No: 5 (1969) 980-983


55. ROOT, Release 6.04/06 on 2015-10-13

56. GEANT/CLHEP, Release 2.3.1.1 on 2015-11-10.