Polish Development and Research Project within the Smart Growth Operational Programme SGOP 1.1.1.1 of The National Centre for Research and Development: JURAND - National Quantum Random Number Generator (POIR.01.01.01-00-0173/15-00)

Randomness

W. Jacak, J. Jacak, W. Donderowicz, and L. Jacak

CONTENTS

I. CLASSICAL APPROACHES TO RANDOMNESS

The definition of randomness constitutes a not easy task within the probabilistic theories – attempts to formally grasp such an idea, in general, lead to a philosopical problems with interpetations of probabily concept itself. Thus, there exists some general definitions of randomness, with developed formalisms, but yet they are unable to provide a complete and formal way to verify what is randomness itself.

For the simplest possible set of symbols, i.e. a binarny set $\{0,1\}$, any sequnece created form such set can be identified with a real number from an interval $[0,1]$. On can suspect that the problem with randomness definition is correlated with the fact that when the length of such sequence is theoretically assumed to be infinite then their set has the cardinality of a continuum same as the set of all real numbers from an interval $[0,1]$.

A. von Mises's unpredictability

The theory presented by von Mises in 1919 [**? ? ?**], was a proposal of one of the basic principles of probability theory, although its foundaments were not purly mathematical as were defined in terms of physical experiment. It is bases on the idea of a so called collective or random sequence.

Let S be the random experiment, which all possible results constite a finit set $L = \{\alpha_1, \ldots, \alpha_m\}$, called the label or attributes set of this experiment. In this experiment let a finit sample of its results be defined as

$$x = (x_1, \ldots, x_N), x_i \in L, \tag{1}$$

where N is the sample size.

The collective is considered as infinite number of trials of experiment S, written as

$$x = (x_1, \ldots, x_N, \ldots), x_i \in L, \tag{2}$$

for which two assumptions are true:

1. **The statistical stabilization** of the relative frequencies of appearance of each attribute $\alpha \in L$ (i.e. possible result of experiment S) in the above infinite sequence, were frequencie is defined as follows

$$\nu_N(\alpha, x) = \frac{n_N(\alpha, x)}{N}, \tag{3}$$

where $n_N(\alpha, x)$ is the number of appearances of α in first N trials.

The statistical stabilization is defined as approaching of the relative frequency to a certian limit for each α, which is called the probability of the attribute α of the random experiment S:

$$P_x(\alpha) = \lim_{N\to\infty} \nu_N(\alpha, x) . \tag{4}$$

2. **The randomness** of the sequence – which is defined as a statement that the statistical stabilization should be satisfied with respect to a choice of a subsequence from x. But a proper definition of such selection is quite problematic.

   The construction of a subsequence is based on decision whether to keep or to reject n-th element of original sequence – this decision cannot depend on any value α and cannot be based neither on value x_n nor any value of x_i where $i > n$. Thus above selection can be written as

$$f_1, f_2(x_1), f_3(x_1, x_2), \ldots, f_n(x_1, \ldots, x_{n-1}), \ldots, \tag{5}$$

   where f_i returns 0 for rejection and 1 for keeping the i-th element and there must be an infinite number of i indices for which f_i returns 1.

   The randomness condition states that there isn't any strategy which satisfies selection rules and produces a subsequence with diffrent odds than subsequence generated by tossing an ideal coin (heads in i-th toss keeps element x_i).

   Selection of a subsequence is reffered to as a place selection.

For example, in case of $L = \{0, 1\}$ let x be a collective – random sequence. Let n_1 be the first index for which $f_{n_1}(x_1, \ldots, x_{n_1-1}) = 1$, n_2 be the next such index, and so on, then due to randomness condition one can write

$$\exists_{x_{n_i}} \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x_{n_i} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} x_k = P_x(1) . \tag{6}$$

Von Mises never solved the problem of the existance of a random sequence (collective). But one can recall some definitions of collectives with regard to special classes of place selections defined by specific rules.

- **Bernoulli sequences**, which von Mises considered as not proper to fully define a collective, can be considered collectives with very specific place selection rule [] – they are defiend as follows.

  Let $w_k = y_1 \ldots y_k$ is a k-digit word, i.e. string of k integers. For an arbitrary word w_k, the sequence $x \in 2^\omega$ is k-distributed if the probability P that an arbitrary substring $x_n \ldots x_{n+k-1}$ of the sequence x is identicall to word w_k is equal to $\frac{1}{2^k}$, as below

$$P\left(x_n \ldots x_{n+k-1} = y_1 \ldots y_k = \frac{1}{2^k}\right) . \tag{7}$$

  A Bernoulli sequence is a sequence, which is a ∞-distributed, i.e. k-distributed for every positive integer k.

  This definition can be understood as counting the relative frequency of an appearance of an arbitrary word in a analyzed sequence, which should overlap with the probability of that word. But in light of von Mises definition such place selection does not satify the randomness requirements.

- **Mises-Wald collectives** – Wald in his work [] stated a theorem, that for any countable set U of place selections and any probability distribution p on the set L of labels, the set of sequences $X(U; p)$ has the cardinality of the continuum, where

  - U – some family of place selections,
  - $X(U; p) = \{x \in L^\infty : \forall_{\phi\in U} \lim_{N\to\infty} \nu_n(\alpha_j; \phi x) = p_j, j = 1, \ldots, m\}$
  - $\nu_N(a, y), \alpha \in L$ – relative frequency appearance of α label from the set L within the first N elements of sequence $y \in L^\infty$
  - L^∞ – infinit sequence of labels from set L.

  Thus Wald random sequence is based on an arbitraty but denumerable set of place selections.

- **Mises-Church collectives** – Church [] proposed to use selection functions which can be alghoritmized – only the recursive functions ϕ_r are allowed in the definition of von Mises. The set of such place selections is countable []. The result obtained by Wald justify the existance of algorithmically computable selection functions proposed by Church. Here again in light of the von Mises definition such properties of place selection are too restrictive
- **Lambalgen axiomatisation for collectives** – Lambalgen [] proposed an approach which is using so called axiomatisation of the relative independence of sequences. He proposed to determine to keep or to reject an i-th value of binary representation of a sequence $x \in 2^\omega$ upon a decision based on the i-th digit of another sequence $y \in 2^\omega$.

Although the collective definitions seem to complement the von Mises theory, Ville publised a problematic objection [].

Ville stated that, if the label set $L = \{0, 1\}$ and a countable set of place selections $U = \{\phi_n\}$, then

$$\exists_{x\in L^\infty} \begin{cases} \forall_n \lim_{N\to N} \sum_{j=1}^{N} \left(\phi_n x\right)_j = \frac{1}{2}, \\ \forall_N \sum_{j=1}^{N} \left(\phi_n x\right)_j = \frac{1}{2}, \end{cases} \tag{8}$$

From definition, such sequence x is a collective, $x \in X\left(U; \frac{1}{2}\right)$, but it cannot be considered as random.

B. Laplace-Ville and Martin-Löf typicality

As the Villes objection to von Mises theory showed that von Mises-Wald collective do not, in general, satisfy all the randomness requirements, he proposed (which is overlaps with Laplace concept []) an alternative to a collective concept, treated as dual definition of randomness to von Mises approach, based on following idea.

A random sequence should satisfy all properties of probability 1 – those properties are probability laws of the form $\mu\left(\{x \in 2^\omega | A\left(x\right)\}\right) = 1$, where μ is a nomalised measure and A denotes a fromula. Similarly as in case of allowed place selection in von Mises definition, here the problematic statment is "all properties A of probability 1".

Each such property can be interpreted as a test of randomness – thus a random sequence, according to Ville, is a seqnece which passes all possible randomness tests. But such situation is not possible, as one need to choose only a countably many of such properties, or otherwise, in case of uncountable family of sets, their intersection can have probablity lower then 1 or does not have a probability defined at all.

In other words, a random sequence defined in frame of some probabilistic measure should satisfy all probabilistic laws for that measure, or the set of random sequences should be a result of the intersection of all possible properties of probability 1 – but as this interesection is an empty set, thus such condition is never satified. The only way to avoid this is to choose a part of those properties – i.e. some (countable) family of tests. But this leads to ambiguity in described aproach as such choice can be done in many ways, resulting in various tests families. This problem was solved by Martin-Löf in 1970 [] by considering a recursive, thus algorythmic, properties of probabily one, which led to define the recursive algorithimc randomness test theory. But unfortunatelly, despite having the randomness formally defined in point of view of typicality, one arrives at the quite paradoxal situation. It can be proven that in language of algorithmic recursive tests there exists a so called universal algorithmic test which is a complete oracle decideing whether a sequence is truly random or not. But unfortunatelly such a test cannot be constructed in an algorithmically manner and the way it looks remains a mystery.

This results in a situation where in case of specified sequence one cannot check in an algorithmically manner whether the sequence is random or not, although it is known that such algorithmic test exists.

1. Martin-Löf randomness

Formally, Martin-Löf randomness of a real r is defined as follows

$$\forall_i \mu\left(A_i\right) \leq 2^{-i} \Rightarrow \neg\forall_i r \in A_i, \tag{9}$$

where A_i corresponds to a set of an recursively enumerable infinite sequence of sets of intervals.

2. Solovay randomness

Solovay proposed a definition without use of so called convergence regulator or significance levels unlike in the case of Martin-Löf definition.

Solovay randomness of a real r is defined as

$$\sum_i \mu\left(A_i\right) < \infty \Rightarrow \exists_N \forall_{i>N} r \notin A_i. \tag{10}$$

In case of von Mises theory the collectives are considered as more importatnt in comparison with the statistics tests, whereas Ville treats the statistical tests as primary factor of randomness – in Martin-Löf's approach to confirm randomness of a sequence it is required that all computable statisitcal tests are satisfied with probability one. Such situation takes place when considered sequence cannot be algorithmically compressed into a shorter one – this sheds light on a correspondance or rather equivalence between randomness definitions based on statistical arguemnts and those based on alghorithmical complexity.

C. Kolmogorov notion of complexity for randomness

Kolmogorov stated that randomness is directly correlated with complexity, and that complexity must be checked algorithmically [].

Formally, one can define

- $L = \{0,1\}$.
- word is a finit sequence of symbols from L.
- L^* is the set of all words in the alphabet L.
- the algorythmic complexity of a word x with respect to some arbitrary algorithm A is defiend as

  $$K_A\left(x\right) = \min l\left(\pi\right), \tag{11}$$

  where $\{\pi\}$ is the set of all programs which can construct word x using algorithm A, $l\left(\pi\right)$ is length of a program π.
- $\forall_A \exists_{A_0} \exists_{C>0} K_{A_0}\left(x\right) \leq K_A\left(x\right) + C$.
- Kolmogorov complexity $K\left(x_{1:n}\right)$ of initial segments $x_{1:n}$ of a random sequence x asymptotically converges to n:

  $$K\left(x_{1:n}\right) \sim n, \quad n \to \infty, \tag{12}$$
- conditional algorithmic complexity $K\left(x;n\right)$ is the lenght of the program π which generates the word x based on the fact that word x has the length n

Martin-Löf published a theorem [], that for every sequence x on $L = \{0,1\}$ alphabet, below is satified for infinitely many n

$$K\left(x_{1:n};n\right) < n - \log_2 n. \tag{13}$$

This showed that Kolmogorov random sequence, defined as $K\left(x_{1:n}\right) \sim n, \quad n \to \infty$ does not exists.

Kolmogorov randomness as algorithmic complexity suffers from another problem – it is not algorithmically computable, thus form of any optimal algorithm A_0 remains unknown. Nevertheless it is possible to estimamte this complexity, which is some solution of this problem.

D. The randomness by Kolmogorov-Chaitin

Despite above mentioned problems, it is possible to construct a proper randomness definition based on the Kolmogorov idea [].

Definitions:

- An algorithm is a computable (recursive) function $A : L^* \to L^*$ defined on some subset D_A of L^*.

- Kolmogorow complexity of the word $x \in L^*$ with respect to A is the length of the shortest program $\pi \in D_A$ such that
$$A\left(\pi\right) = x : K_A\left(x\right) = l\left(\pi\right). \tag{14}$$
In case when such π does not exists then $K_A\left(x\right) = \infty$.
- Prefix of a word $x = x_1 \ldots x_n$ is denoted as $\hat{x} = x_1 \ldots x_m$ where $m \leq n$.
- A prefix free subset D of L^* is a subset where no word from D is a prefix of any other word from D.

Now let us assume A as an arbitrary algorithm – a computable function, defined on a prefix free domain D. Algorithmic prefix free complexity of a word x from L^* defined on A is the length of the sortests program $\pi \in D$, such that

$$A\left(\pi\right) = x : \tilde{K}_A\left(A\right) = l\left(x\right). \tag{15}$$

Similarly, if such $\pi \in D$ does not exists, then $\tilde{K}_A\left(x\right) = \infty$.

A theorem can be stated that there exists an optimal prefix free algorithm (computable function) A_0, such that

$$\forall_A \exists_{C>0} \quad \tilde{K}_{A_0}\left(x\right) \leq \tilde{K}_A\left(x\right) + C. \tag{16}$$

And finally, Kolmogorov-Chaitin random sequence x from L^∞, $L = \{0, 1\}$ is a sequence which is incompressible – non of its initial segments can be compressed more then for a fixed finit number of bits. This can be written as follows

$$\forall_n \exists_{b>0} \quad \tilde{K}\left(x_{1:n}\right) \geq n - b. \tag{17}$$

The Kolmogorov-Chaitin randomness is equivalent to Martin-L´of randomness [].

Both approaches are based on well developed formalism – in case of Kolmogorov-Chaitin it is complexity-incompressibility, in case of Martin-Löf it is typicality, which is foundation of statisitical pseudo-random sequence testing (e.g. NIST test suite).

Despiete those theoretical foundations both algorithmical approaches do not seem to constitute fundamental condition of randomness. Closer to the essence of the randomness seems to be the idea of randomness as unpredictability, despite of all its drawbacks.

According to Khrennikov and Zeilinger discussion [], it is quite possible that strictly mathematical approach to randomness and formalization of its definition seems to be out of reach, as probably it is insufficient to use only the mathematical tools for the theoretical construction of concept of the randomness itself. It is the physical procedures where the true randomness is hidden – furthermore, not the classical physics itself (as it is formally deterministic) but rather the quantum physics phenomenons, due to theirs fundamental nondeterminism.

II. QUANTUM APPROACH

A. The Copenhagen interpretation of quantum mechanics and von Neumann measuement scheme

The Copenhagen interpretation of quantum mechanis was formulated by N. Bohr and W. Heisenberg in 1927 in Copenhagen, based on an M. Born idea of interpreting the wavefunction in probabilistic manner (e.g. for wavefunction in position representation, its modulus squared represents a probability or probability density of finding a particle in given point or part of space). It is currently referred as a standard interpretation, even though a quite rapid development of the concept of a so called quantum Bayesianism proposed by Fuchs [], which contradicts some essential assumption of Copenhagen interpretation especially in the foundations of measurement process.

In Copenhagen interpretation in the quantum measurement process quantum system randomly selects one of many possible quantum states. This selection is caused by the wavefunction collapse due to the interaction between the observer (some macroscopic external object, characterised by at least Avogadro's number of degrees of freedom – thus classical) and a measured quantum system. In fact this theory states that it is the knowledge of the observer that collapses and not the formal and objective wavefunction.

The scheme of measurement process was the subject of deep studies performed by von Neumann []. As a result von Neumann proposed an "ansatz" – which can be fromulated as an axiom stating that at the moment when measurement occures the wave function of a quantum system, presented in measured observable basis (observable – an hermitian operator corresponding to some physical property of a quantum system, e.g. position, momentum, energy, etc.) collapses, in a truly random manner, to one of the observable base states with non-zero coefficient in initial wavefunction expansion into the measurment base.

For a certain moment in time, one can write

$$\begin{aligned} \psi\left(\boldsymbol{r}\right) &= \sum_{i}^{\dim \mathcal{H}_{\hat{A}}} \boldsymbol{c}_i \phi_i\left(\boldsymbol{r}\right), \\ \hat{A}\phi_i &= \lambda_i \phi_i, \end{aligned} \tag{18}$$

where $\mathcal{H}_{\hat{A}}$ is the Hilbert space spanned by an observable $\hat{A}$ basis $\{\phi_i\}$, and λ_i is an eigenvalue of the operator $\hat{A}$. In such notion the measurement results in choosing, in a random manner, one of ϕ_i state for which c_i is non-zero. In such case the measured quantum system attains randomly choosen state ϕ_i and the λ_i eigenvalue is being mapped in macroscopic number of degrees of freedom of measurement device, which allows to observe it in a classical manner as the measurement result of some physical observable. The modulus squared of coefficients c_i defines only the probabilites of occurences of different ϕ_i states as the result of the measurement. It is assumed that one deals here with the frequentist probabilites which constitute a fundamental problem in light of quantum mechanics formulation, especially a measurement theory. Such assumption imposes the existance of an infinit number of identical copies of the measured system for which the occurances of the measurement resultant states will have the distribution defined by the coefficients c_i. But this stands in contradiction with the properites of quantum measurement.

In quantum mechnics a measurement process is characterized by following fundamental axioms

- quantum measurement is an unrepeatable process,
- quantum measurement is destructive to a part of the quantum information and quantum state,
- quantum measurement is distinguishing the observer.

In other words – as the quantum measurement process irreversibly destroys some part of the quantum information the whole process is unrepeatable. This process distinguish the observer as the one who performed the measurement. Additionally Zureks no-cloning theorem [] ensures a lack of copies of the measured system emphasising the true unrepeatably nature of quantum measurement.

But in light of above the idea of the frequentist probability in quantum measurement cannot be fully justified, as there will never be a situation when quantum system state is not disturbed by the measurement, and thus the measurement on a unknown quantum state can be performed only once (as the this quantum state after the meausrement will be irrevesibly destroyed). It is possible to organize a setup to generate, in a deterministic manner, some state, starting every time form some known initial state – this will lead to

This problem led to a concept of changing the frequentist probability paradigm. Fuchs explains it with use of a comparison to the weather prediction []. When predicting a weather for the next day one deals with a situation which has never taken place before, thus one cannot use the frequentist probability paradigm (as it corresponds to a phenomenon which can be). Instead of that the Bayesian probability paradigms should be considered. The weather prediction must be made upon the knowledge of similar but not identical situations. Thus it can be described with the conditional probability.

Fuchs argues that similar situation takes place for quantum measurement as it is unrepeatable due to its destructive proces type. This approach is called QBism or Quantum Bayesianism [].

Frequentist probability, as an interpretation of probability, defines probability of some event in terms of a limit of its relative frequency when the number of trials tends to the infinity. When the number of trails in infinit then one calls the probability of some event a true probability

$$P\left(x\right) = \lim_{N\to\infty} \frac{n}{N}, \tag{19}$$

where N is the number of all trails, n is the number of trials where the considered event occured.

The difference between QBism and frequentist is that the state of the system can be considered as an objective fact about that system or as a judgement made by an observer on the basis of his prior experience of that system (the QBism view). QBisms questions the soul idea of quantum mechanics – the entanglement.

The existance of non-local quantum correlations – quantum entanglement, requires two assumptions to be fulfilled:

1. No-signalling – there is no communication between the measuring devices.
2. Initial randomness – required for performance of the Bell test.