Zeitschr. f. math. Logik und Grundlagen d. Math.
Bd. 26, S. 385–394 (1980)

EXACT EXPRESSIONS FOR SOME RANDOMNESS TESTS

by PETER GÁCS in Rochester, New York (U.S.A.)

0. Introduction

The notion of randomness which v. MISES attempted to formalise first has received a long time only moderate attention in contemporary probability theory. It was hard to find a convincing distinction between random and nonrandom elements of a probability space and, for most probabilists, it has been up to now not clear how much would one be happier finding it. Though for a statistician nothing seems to be more interesting than the question about randomness. Given an element ω of the event space Ω as the outcome of an experiment, and a distribution P he wants to find out how justified it is to suppose that the underlying distribution to the experiment was P; i.e. that ω is *random* w.r.t. P. However, his model is slightly different because in the typical cases he has an access to a *large number of independently repeated experiments* $P^n = P \times P \times \ldots \times P$ and what he wishes to decide on the basis of $\omega = (\omega_1, \ldots, \omega_n)$ is only the question about P, the product structure taken for granted. The decisions can then be made on the basis of central limit theorems, and it is, roughly said, the investigation of the conditions of such decisions to which most of mathematical statistics is devoted.

There are some highly interesting statistical situations where the product-space framework is not applicable: e.g. prediction problems or testing of pseudo-random sequences.

After its revival in the sixties by the work of KOLMOGOROV and MARTIN-LÖF (continued by LEVIN, CHAITIN, SCHNORR) the modern theory of randomness approaches now to a satisfiable form and its solutions to these problems are of convincing simplicity and generality.

Unfortunately, to understand them one has to learn some computability theory, and if later one tries to apply them one notes with some disappointment the large gap between theoretical and practical computability.

The present paper does not bridge this gap, either. It gives some more exact relations between complexity and randomness and one can only hope that when the theory using general computability will be more perfect then the chances to find its practical extension increase.

In Section 1 we give the necessary definitions, in Section 2 some known results on MARTIN-LÖF's tests. In Section 3 apriori probability and its known relation to randomness is described. Section 4 is devoted to various definitions of complexity and their estimates. In Section 5 we give some exact expressions for MARTIN-LÖF's test in terms of different types of complexities. The possibility of such expressions shows once again the technical flexibility of the complexity apparatus. Finally, in Section 6, we follow up the connections of LEVIN's uniform tests to the previous theory and introduce a somewhat modified uniform test and a simple one having the conservation property.

1. Basic definitions

Notations. All logarithms are to the base 2. N is the set of natural numbers, $\mathsf{N}_k = \{0, 1, \ldots, k-1\}$. Q is the set of rational numbers, R the set of real numbers, $\mathsf{R}_+ = \mathsf{R} \cap (0, \infty)$, $\overline{\mathsf{R}} = \mathsf{R} \cup \{\infty\}$, $\overline{\mathsf{R}}_+ = \mathsf{R}_+ \cup \{\infty\}$. $\mathsf{N}_2^* = \bigcup \mathsf{N}_2^n \cup \{\Lambda\}$, where Λ is the so-called empty word. We fix a recursive one-to-one correspondence $\varkappa: \mathsf{N}_2^* \to \mathsf{N}$ with $\varkappa(x) \geqq l(x)$, where $l(x)$ is the length of the word x. $\Omega = \mathsf{N}_2^\infty =$ the set of infinite binary sequences. For $x, y \in \mathsf{N}_2^*$, xy is the concatenation of x and y, $x \subset y$ iff $\exists z.\, xz = y$. For $\omega \in \mathsf{N}_2^\infty$, $\omega = \omega_1\omega_2 \ldots (\omega_i \in \mathsf{N}_2)$ and $\omega^n = \omega_1 \ldots \omega_n$.

For two nonnegative functions f, g let $f \lesssim g$ mean that a $c > 0$ exists with $cf \leqq g$. $f \approx g$ iff $f \lesssim g$ and $g \lesssim f$, $f \leqq\!\!\!\lesssim g$ iff $\exp(f) \lesssim \exp(g)$, $f \asymp g$ iff $f \leqq\!\!\!\lesssim g$ and $g \leqq\!\!\!\lesssim f$.

We consider N_2^∞ with its usual topology determined by the basis $\{x\mathsf{N}_2^\infty \mid x \in \mathsf{N}_2^*\}$, R with the (usual) topology determined by $\{(r_1, r_2) \mid r_1, r_2 \in \mathsf{Q}, r_1 < r_2\}$.

Let $\mathscr{M}$ be the set of all probability measures over Ω, with its (weak) topology determined by the subbasis of sets of the form $\{\mu \in \mathscr{M} \mid r_1 < \mu(x) < r_2\}$. Here for $x \in \mathsf{N}_2^*$, $\mu(x) = \mu(x\mathsf{N}_2^\infty)$, and $r_1, r_2 \in \mathsf{Q}$.

Sometimes N, N_2^* will also be considered as discrete topological spaces (essentially equivalent by the correspondence $\varkappa$). In any of these (locally compact) topological spaces X a basis $\{U_i \mid i \in \mathsf{N}\}$ is given *with a fixed enumeration*. New spaces can be constructed by products.

Definition. An element of the space X will be called *computable* if $\{i \mid x \in U_i\}$ is enumerable. An open set $G \subset X$ is called *constructively open* if $\{i \mid \bar{U}_i \subset G\}$ is enumerable. $F \subset X$ is called *constructively closed* if F^c, the complement of F, is constructively open. A function $f: X \to \mathsf{R}$ is called *semicomputable* (from below) if $\{(r, x) \mid r \in \mathsf{R}, x \in X, f(x) > r\}$ is constructively open. f is called *computable* if f and $-f$ are semicomputable. Especially, a measure is called computable iff it is computable as an element of $\mathscr{M}$. It is easy to see that this is equivalent to the condition that $\mu: \mathsf{N}_2^* \to \mathsf{R}$ be computable as a function.

2. Martin-Löf's tests

MARTIN-LÖF declared for nonrandom those elements ω of Ω which belong to some constructively definable set of measure 0. Since rather different approaches to randomness lead to an equivalent definition, there is a wide agreement that MARTIN-LÖF's random elements are convenient to be considered as "the" random ones. Every element of positive probability (in discrete spaces typically every element) is random, and one can only speak of their *degree of nonrandomness* (deficiency of randomness). Different "tests" assign different degrees of nonrandomness. It is hard to vote for one test as the most natural but most proposed tests are asymptotically equal to MARTIN-LÖF's (in the case of computable measures).

Definition ([1]). Let μ be a computable measure. A *Martin-Löf-test* (ML-test) is a semicomputable function $d: \mathsf{N}_2^\infty \to \overline{\mathsf{R}}$ with $\forall k.\mu\{\omega \mid d(\omega) > k\} < 2^{-k}$.

Theorem 2.1 (MARTIN-LÖF). *Among the ML-tests there is a maximal one in the sense of the ordering* $\leqq\!\!\!\lesssim$.

Definition. We fix a maximal ML-test once for all for each μ and call it $d_M(\omega \mid \mu)$, a *universal ML-test*. A sequence ω will be called *random* iff $d_M(\omega \mid \mu) < \infty$.

The notion of a random sequence is more invariant than ML-tests. One can imagine rather different reasonable ways of measuring nonrandomnesses present in sequences but the question which sequences are random at all will be answered equivalently by most of them. We generalize now the notion of a test.

Definition. Given a computable measure μ, a semicomputable function $d: \Omega \to \overline{\mathsf{R}}$ is a *test* if $\lim_m \mu(\omega \mid d(\omega) > m) = 0$ recursively, i.e. a recursive function $m(k)$ exists with $\mu(\omega \mid d(\omega) > m(k)) < 2^{-k}$. For a test d we have

$$\{\omega \mid d(\omega) = \infty\} \subseteqq \{\omega \mid d_M(\omega) = \infty\}. \tag{2.1}$$

Definition. In case of equality in (2.1) the test is called *universal*. (Every universal test d giving in the following is *asymptotically equal* to d_M, i.e. $\lim_{d \to \infty} d_M(\omega)/d(\omega) = 1$ holds.)

3. Apriori probability

Apriori probability will not be itself a measure, rather the lower bound of the set of measures satisfying a r.e. set of simple conditions, a semimeasure. Our definition of a semimeasure is here more elementary than the final one given in [11] and used in Section 6.

Definition ([2]). A *semimeasure* over Ω is a function $\varphi: \mathsf{N}_2^* \to \mathsf{R}_+$ with

$$\varphi(x) \geqq \varphi(x\,0) + \varphi(x\,1),\ \varphi(\Lambda) \leqq 1.$$

The set of semimeasures will be denoted by $\mathscr{S}$. $\mathscr{S}$ can be given the same topology as $\mathscr{M}$ thus $\mathscr{M}$ being a subspace of $\mathscr{S}$. It is easy to prove that for any semimeasure φ, $\varphi(x) = \inf\{\mu(x) \mid \mu \geqq \varphi, \mu \in \mathscr{M}\}$, and that the lower bound of any set of measures is a semimeasure.

A semimeasure is called *semicomputable* if it is semicomputable as a function from N_2^* to R^+. It is not hard to see that φ is semicomputable iff the set $\{\mu \in \mathscr{M} \mid \mu \geqq \varphi\}$ is constructively closed in $\mathscr{M}$. Semimeasures over the discrete spaces $\mathsf{N}, \mathsf{N}_2^*$ will also be considered. (A semimeasure over N is given by a function $\varphi: \mathsf{N} \to \mathsf{R}^+$; φ must satisfy the condition $\Sigma_x \varphi(x) \leqq 1$.) Semimeasures over N_2^* correspond to those over N by $\varkappa$.

Theorem 3.1 (Levin [2]). *In the set of all semicomputable semimeasures there is a maximal element with respect to the relation* $\lesssim$.

Let us call a fixed maximal semimeasure $M = M_\Omega$ the *apriori probability* over Ω. The apriori probabilities over $\mathsf{N}, \mathsf{N}_2^*$ will be denoted by $M_{\mathsf{N}}, M_{\mathsf{N}_2^*}$ (writing freely $M_{\mathsf{N}}(x)$ for $M_{\mathsf{N}}(\varkappa(x))$; we do not write out the subscript if no misunderstanding may arise). Levin's following theorem establishes the role of apriori probability in determining randomness.

Theorem 3.2 (Levin [3], see also Schnorr [4] for a special case of a related result of [3]). $d_S(\omega \mid \mu) = \sup_n \log M_\Omega(\omega^n) - \log \mu(\omega^n)$ *is a universal test for any computable measure* μ.

Remarks 1. $\log (M_\Omega(\omega^n)/\mu(\omega^n))$ is bounded from below for every fixed computable μ, so ω is random w.r. to μ iff $\mu(\omega^n) \approx M_\Omega(\omega^n)$.

2. d_S is defined also for noncomputable μ and for all $\mu \geqq M_\Omega$ one has $d_S(\omega \mid \mu) \leqq 0$. In other words, if we measure nonrandomness by d_S, we find that "all sequences are random with respect to the apriori probability". This statement was proved to be true also for any reasonable "uniform tests" (see Section 6).

4. Complexity

The numbers

$$H_{\mathsf{N}}(x) = -\log M_{\mathsf{N}}(x), \quad H_{\Omega}(x) = -\log M_{\Omega}(x)$$

can by many reasons be considered as a measure of the complexity of the finite sequence x.

Definition ([5]). *Kolmogorov's complexity* of the sequence x given y with respect to a partial recursive function $A: \mathsf{N}_2^* \times \mathsf{N}_2^* \to \mathsf{N}_2^*$ is defined as

$$K_A(x \mid y) = \min\{l(p) \mid A(p, y) = x\}.$$

Put $K_A(x) = K_A(x \mid \Lambda)$.

Kolmogorov proved that there is a p.r. function U with $K_U \leqq K_A$ for any other p.r. A. Fixing such a U, we define K as K_U. $K(x \mid y)$ is called *Kolmogorov's complexity of x given y*. A slight modification in KOLMOGOROV's definition of complexity proved very useful in the applications.

Definition ([6–8, 12]). A set $E \subset \mathsf{N}_2^*$ is said to be *prefix-free* if

$$\forall x, y.\ x, y \in E \to x \not\leqq y.$$

If we confine us to the set of p.r. functions $\mathscr{F} = \{A \mid \forall y.\{p \mid A(p, y) \text{ is defined}\} \text{ is prefix-free}\}$ then in this class a function T can be found with $K_T \leqq K_A$ for all $A \in \mathscr{F}$.

Definition. $K_T(x \mid y)$ is called *the complexity of x given y*.

Notation. We will sometimes use T' instead of T defined by

$$T'(p, x) = y \quad \text{iff} \quad \exists q \leqq p.\ T(q, x) = y. \tag{4.0}$$

Fixing such a T we have

Theorem 4.1 (LEVIN [6]). $H_{\mathsf{N}}(x) \asymp K_T(x)$.

We define here $M_{\Omega}(x \mid y)$ only for y as an element of the discrete space N_2^*. For this we take a maximal (w.r. to $\lesssim$) one among the *conditional semicomputable semimeasures*, i.e. semicomputable functions $\varphi: \mathsf{N}_2^* \times \mathsf{N}_2^* \to \mathsf{R}^+$, where $\varphi(x \mid y)$ is for each y a semimeasure over Ω.

The three kinds of complexity defined before are numerically close to each other. Indeed, it is easy to see (and well-known) that

Theorem 4.2.

(a) $K \leqq H_{\mathsf{N}} \leqq K + 2 \cdot \log K$, (b) $H_{\Omega} \leqq H_{\mathsf{N}}$,

(c) *for any prefix-free r.e. set* $E \subseteq \mathsf{N}_2^*$, $\exists c.\ \forall x \in E, y \in \mathsf{N}_2^*.\ H_{\mathsf{N}}(x \mid y) \leqq H_{\Omega}(x \mid y) + c$.

By $\varkappa$, K and H_{N} are defined on N as well as on N_2^* and their order of magnitude can be estimated. There is no nontrivial p.r. lower estimate to them (see [2]). As to the upper estimates, the least monotone ones can be found. For any function $f: \mathsf{N} \to \mathsf{N}$ its least monotone upper bound is $f^*(n) = \max_{k \leqq n} f(k)$.

Theorem 4.3 ([5, 7, 9]).

$$K^*(n) \asymp \log n, \qquad H_{\mathsf{N}}(n) \asymp \log n + H([\log n]).$$

Note that H is not computable, only semicomputable from above. It gives rise to a number of somewhat weaker computable upper estimates like

$$\log n + \log\log n + 2\log\log\log n.$$

Theorem 4.4 (LEVIN). $K(x) \asymp \min\{i \mid H_{\mathsf{N}}(x \mid i) \leqq i\}$, $\quad K(x) \asymp H_{\mathsf{N}}(x \mid K(x))$.

Remark. $K(x \mid y)$ is similarly expressible.

Proof. $H_{\mathsf{N}}(x \mid K(x)) \leqq K(x)$ is obvious from the definition of these quantities. Having an i with $H_{\mathsf{N}}(x \mid i) \leqq i$ we have $K(x) \leqq i$. To show this, let us define the function:

$$A(p) = T''(p, l(p)) \quad \text{(see (4.0))}.$$

For any i, p, x with $l(p) \leqq i$, $T''(p, i) = x$ we have $A(p\,0^{i-l(p)}) = x$. Hence

$$K(x) \leqq K_A(x) \leqq i.$$

5. Tests expressed by complexities

Though by Theorem 3.2 we have a most satisfiable characterization of randomness by the behavior of the apriori probability (whose logarithm is a sort of complexity), two questions are of some technical interest:

a) to express MARTIN-LÖF's test by some complexity,

b) to see how the other complexities are suitable to express randomness.

For our characterisation of MARTIN-LÖF's test we introduce an auxiliary complexity.

Definition. $\bar{H}(x; k) = \min\{i \mid H(x \mid k - i) \leqq i\}$.

This definition has sense for both H_Ω and H_{N}. $\bar{H}(x; k \mid y)$ can be defined similarly, with y everywhere in the condition. Then we have by Theorem 4.4:

$$\bar{H}_{\mathsf{N}}(x; k \mid k) \asymp K(x \mid k). \tag{5.1}$$

Thus $\bar{H}_{\mathsf{N}}$ can be considered as a generalization of K.

Remarks.

1. $\bar{H}_\Omega \leqq \bar{H}_{\mathsf{N}}$. This follows easily from the relation $H_\Omega(x \mid k - i) \leqq H_{\mathsf{N}}(x \mid k - i)$.
2. $\bar{H}$ is obviously semicomputable from above.
3. Similarly to (4.1) we have

$$\bar{H}(x; k) \asymp H(x \mid k - \bar{H}(x; k)). \tag{5.2}$$

As we have seen in Theorem 3.2 the testing of the randomness of ω w.r. to μ naturally involves a comparison of $-\log \mu(\omega^n)$ with some complexity of ω^n.

Notation. $l_\mu(\omega^n) = [-\log \mu(\omega^n)]$.

Theorem 5.1. *For a fixed computable measure* μ,

$$d_M(\omega \mid \mu) \asymp \sup_n l_\mu(\omega^n) - \bar{H}_\Omega(\omega^n; l_\mu(\omega^n)) \asymp \sup_n l_\mu(\omega^n) - \bar{H}_{\mathsf{N}}(\omega^n; l_\mu(\omega^n)).$$

Proof. Let us denote the expressions in Theorem 5.1 by d_{ML} and d_{MC} respectively. By the remark following the definition of $\bar{H}$ we have $d_{MC} \leqq d_{ML}$. We have to show $d_{ML} \leqq d_M$ and $d_M \leqq d_{MC}$.

$d_{ML} \leqq d_M$ will be proved if we show that d_{ML} is semicomputable (this is clear from the definition) and that

$$\mu\{\omega \mid d_{ML}(\omega \mid \mu) > m\} \lesssim 2^{-m}.$$

Now by (5.2) we have with a $c > 0$

$$\begin{aligned}\{\omega \mid d_{ML}(\omega \mid \mu) = m\} &\subseteq \{\omega \mid \exists n.\, l_\mu(\omega^n) - H_\Omega(\omega^n \mid m) \geqq m - c\} \\ &= \{\omega \mid \exists n.\, M_\Omega(\omega^n \mid m)/\mu(\omega^n) \geqq 2^{m-c}\}.\end{aligned}$$

The following simple lemma holds for any semimeasure φ and measure μ.

Lemma 5.1. $\mu\{\omega \mid \exists n.\, \varphi(\omega^n)/\mu(\omega^n) > \alpha\} < \alpha^{-1}$.

Proof. Put $n(\omega)$ the first n (if it exists) with $\varphi(\omega^n)/\mu(\omega^n) > \alpha$. The set of finite sequences

$$A = \{x \mid \exists \omega.\, x = \omega^{n(\omega)}\}$$

has the prefix property, hence $\Sigma_{x\in A}\varphi(x) \leqq 1$ by the definition of semimeasures. On the other hand, for $x \in A$ we have $\mu(x) \leqq \varphi(x).\, \alpha^{-1}$, hence

$$\mu\{\omega \mid \exists n.\, \varphi(\omega^n)/\mu(\omega^n) > \alpha\} = \Sigma_{x\in A}\mu(x) \leqq \alpha^{-1}.$$

Applying the lemma to the semimeasures $M_\Omega(\omega^n \mid m)$ we get

$$\mu\{\omega \mid d_{ML}(\omega \mid \mu) = m\} \leqq 2^{-m+c}.$$

Now we prove $d_M \leqq d_{MC}$. By the semicomputability of d_M there is a recursive sequence $\{(m_t, x_t)\}_{t\in\mathbf{N}}$ of elements of $\mathbf{N} \times \mathbf{N}_2^*$ with

$$\{(m_t, x_t) \mid t \in \mathbf{N}\} = \{(m, x) \mid \forall \omega \in x\mathbf{N}_2^\infty.\, d_M(\omega \mid \mu) > m\}.$$

Let $t(m, \omega)$ be the first $t \in \mathbf{N}$ (if it exists) with $m \leqq m_t$ and $\omega \in x_t\mathbf{N}_2^\infty$. The set $U = \{(m, x) \mid \exists \omega.\, x = x_{t(m,\omega)}\}$ is easily seen to be recursively enumerable, $U_m = \{x \mid (m, x) \in U\}$ is a prefix set of finite sequences, $U_m\mathbf{N}_2^\infty = \{\omega \mid d_{ML}(\omega \mid \mu) > m\}$. Hence $\Sigma_{x\in U_m}\mu(x) \leqq 2^{-m}$. But then

$$\varphi(x \mid m) = \begin{cases} \mu(x) \cdot 2^m \text{ for } (m, x) \in U \\ 0 \text{ otherwise} \end{cases}$$

is a "conditional semicomputable semimeasure" over $\mathbf{N}_2^*$, and we have $M_{\mathbf{N}}(x \mid m) \gtrsim \varphi(x, m)$.

Hence for all ω with $d_M(\omega \mid \mu) > m$ there exist an i.e. $n(= l(x_{t(m,\omega)}))$ such that

$$H_{\mathbf{N}}(\omega^n \mid m) \leqq l_\mu(\omega^n) - m; \qquad \bar{H}_{\mathbf{N}}(\omega^n; l_\mu(\omega^n)) \leqq l_\mu(\omega^n) - m$$

(take $i = l_\mu(\omega^n) - m$ in the definition of $\bar{H}$). □

We can deduce from this theorem the first known exact relation between tests and complexity established in [10]. Denote by π_n the equidistribution over $\mathbf{N}_2^n$. A universal ML-test $d_M(x \mid \pi_n)$ can be defined as a greatest (w.r. to $\leqq$) semicomputable (in (x, n)) function $d(x \mid \pi_n)$ with $\forall n, k.\, \pi_n\{x \mid d(x \mid \pi_n) \geqq k\} \leqq 2^{-k}$.

Corollary 5.1 (see [10]). $d_M(x \mid \pi_n) \asymp n - K(x \mid n)$.

Proof. The proof of Theorem 5.1 is analogous (even simpler) over the space $\mathbf{N}_2^*$ instead of Ω, and even if we let n as parameter everywhere in the equations. We thus have for any $x \in \mathbf{N}_2^n$

$$d_M(x \mid \pi_n) \asymp -\log \pi_n(x) - \bar{H}_{\mathbf{N}}(\omega^n; [-\log \pi_n(x)] \mid n)$$
$$= n - \bar{H}_{\mathbf{N}}(x; n \mid n) \asymp n - K(x \mid n),$$

by (5.1). □

From the first part of Theorem 5.1 one can reprove Theorem 3.2. We have by definition $d_S(\omega \mid \mu) \asymp \sup_n - \log \mu(\omega^n) - H_\Omega(\omega^n)$.

Corollary 5.2. *d_S is asymptotically equal to d_M.*

Proof. Let H be $H_{\mathbf{N}}$ or H_Ω. We use the obvious inequality

$$H(x) \leqq H(x \mid j) + H_{\mathbf{N}}(j) \leqq H(x \mid j) + 2 \cdot \log_2 j.$$

Now put $k = [-\log \mu(\omega^n)]$, $\Delta = k - \tilde{H}(\omega^n; k)$. We have by (5.2)

$$k - H(\omega^n) \leqq k - H(\omega^n \mid \Delta) \prec k - \tilde{H}(\omega^n; k),$$

hence $d_S \leqq d_M$. But $k - \tilde{H}(\omega^n; k) \prec k - H(\omega^n \mid \Delta) \leqq k - H(\omega^n) + 2 \cdot \log_2 \Delta$ hence $\Delta - 2 \cdot \log_2 \Delta \leqq k - H(\omega^n)$, $d_M - 2 \cdot \log_2 d_M \leqq d_S \leqq d_M$, i.e. the tests d_S and d_M are asymptotically equal. □

Since the proof does not make any difference between H_Ω and H_{N}, we have also proved the following

Corollary 5.3. $d_C(\omega \mid \mu) := \sup_n l\mu(\omega^n) - H_{\mathsf{N}}(\omega^n)$ *is a universal test, asymptotically equal to* $d_M(\omega \mid \mu)$.

In the special case of the equidistribution this test coincides with the one proposed by CHAITIN [7] and proved to be universal by SCHNORR.

The test $d_C(\omega \mid \mu)$ has some other meaningful characterisations.

Put $t_C(\omega \mid \mu) = 2^{d_C(\omega \mid \mu)} = \sup_n M_{\mathsf{N}}(\omega^n)/\mu(\omega^n)$.

Theorem 5.2. *For any fixed computable measure* μ, $t_C(\omega \mid \mu)$ *is* $\lesssim$*-maximal among the semicomputable functions* $t(\omega)$ *with property* $\int t(\omega)\, \mu(d\omega) \leqq 1$.

Remark. We must fix μ, because the constants in $\lesssim$ are depending on it.

Proof. We have $\sup_n M_{\mathsf{N}}(\omega^n)/\mu(\omega^n) \leqq \Sigma_n M_{\mathsf{N}}(\omega^n)/\mu(\omega^n) = \Sigma_{x \in \mathsf{N}_2^*} M_{\mathsf{N}}(x) \cdot \dfrac{g_x(\omega)}{\int g_x(\omega)\, \mu(d\omega)}$

where

$$g_x(\omega) = \begin{cases} 1 & \text{for } x \subseteq \omega \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$\int d_C(\omega \mid \mu)\, \mu(d\omega) \leqq \Sigma_x M_{\mathsf{N}}(x) \leqq 1.$$

It remains to show that for any semicomputable function $t(\omega)$ with $\int t(\omega)\, \mu(d\omega) \leqq 1$ we have $t(\omega) \lesssim t_C(\omega \mid \mu)$. For $t(\omega)$ one always has recursive sequences x_i, k_i with

$$\sup_i 2^{k_i} g_{x_i} \approx \Sigma_i\, 2^{k_i} g_{x_i} \approx t.$$

Hence $2^{k_i}\mu(x_i) \lesssim M_{\mathsf{N}}(i)$, $\quad 2^{k_i} \lesssim M_{\mathsf{N}}(i)/\mu(x_i)$ because of $\Sigma_i\, 2^{k_i}\mu(x_i) \leqq 1$. We have

$$t(\omega) \lesssim \sup_i \frac{M_{\mathsf{N}}(i)}{\mu(x_i)} \cdot g_{x_i}(\omega) \lesssim \sup_x \frac{M_{\mathsf{N}}(x)}{\mu(x)} \cdot g_x(\omega) = t_C(\omega \mid \mu). \quad \square$$

We can use for the expression of the main term in MARTIN-LÖF's tests KOLMOGOROV's complexity too, as shown by the following theorem.

Theorem 5.3. *Put* $\Delta(\omega^n \mid \mu) = l_\mu(\omega^n) - K(\omega^n \mid n, l_\mu(\omega^n))$. *Then*

$$d_M(\omega \mid \mu) \prec \sup_n \Delta(\omega^n \mid \mu) - \tilde{H}(n, l_\mu(\omega^n); \Delta(\omega^n \mid \mu)).$$

As in the Corollary of Theorem 5.1, we get a simpler universal test espressed by K if we do not require it being a ML-test.

Corollary 5.4. $d_K(\omega \mid \mu) := \sup_n \Delta(\omega^n \mid \mu) - H(n, l_\mu(\omega^n))$ *is a universal test, asymptotically equal to* d_{ML}.

Especially for the Lebesgue measure λ we have the test:

$$d_K(\omega \mid \lambda) = \sup_n (n - K(\omega^n \mid n) - H(n)).$$

Remark. This corollary uses LEVIN's remark in 1972 on the authors unpublished work.

Proof of Theorem 5.3. Let us first prove that the expression d_{MK} in the theorem defines a test. Its semicomputability is clear from the definition. We have to prove $\mu\{\omega \mid d_{MK}(\omega \mid \mu) > m\} \lesssim 2^{-m}$. Like in the proof of Theorem 5.1 we have with some c:

$$\{\omega \mid d_{MK}(\omega \mid \mu) = m\} \subseteqq \{\omega \mid \exists n.\, \Delta(\omega^n \mid \mu) - H(n, l_\mu(\omega^n)) \mid m) \geqq m - c\}$$
$$\subseteqq \bigcup_{n,k} \{\omega \mid l_\mu(\omega^n) = k, k - K(\omega^n \mid n, k) - H(n, k \mid m) \geqq m - c\}.$$

Now

$$\mu\{\omega \mid l_\mu(\omega^n) = k, K(\omega^n \mid n, k) \leqq k - m - H(n, k \mid m) + c\}$$
$$\lesssim 2^{-k} \#\{x \in \mathsf{N}_2^n \mid K(x \mid n, k) \leqq k - m - H(n, k \mid m) + c\} \lesssim 2^{-m+c} M(n, k \mid m)$$

by well-known properties of KOLMOGOROV's complexity. Hence

$$\mu\{\omega \mid d_{MK}(\omega \mid \mu) = m\} \lesssim \Sigma_{n,k} 2^{-m+c} M(n, k \mid m) \lesssim 2^{-m+c}.$$

This proves $d_{MK} \leqq d_M$. The proof of $d_M \leqq d_{MK}$ is based on the following estimate which readily follows from the definition of KOLMOGOROV's complexity. For any $x \in \mathsf{N}_2^*$, $\omega \in x\mathsf{N}_2^\infty$, $n \geqq l(x)$

$$K(\omega^n \mid n, x, l_\mu(\omega^n)) \leqq l_\mu(\omega^n) + \log \mu(x). \tag{5.3}$$

Let us make use now of the enumerability of the set U defined in the proof of Theorem 5.1. We define the conditional semicomputable semimeasure φ over N^2 by

$$\varphi(n, k \mid m) = \begin{cases} 2^{m-1}\mu(\varkappa^{-1}(n)) & \text{if } (m, \varkappa^{-1}(n)) \in U \\ & \text{and } 2^{-k-1} < \mu(\varkappa^{-1}(n)) < 2^{-k+1} \\ 0 & \text{otherwise.} \end{cases}$$

If $d_M(\omega \mid \mu) > m$ then for some $x \in U_m$ with $n = \varkappa(x)$, $k = l_\mu(\omega^n)$ we have by (5.3)

$$K(\omega^n \mid n, k) \leqq k + \log \varphi(n, k \mid m) - m \leqq k - H(n, k \mid m) - m,$$
$$H(n, k \mid m) \leqq \Delta(\omega^n \mid \mu) - m.$$

$\bar{H}(n, k; \Delta(\omega^n \mid \mu)) \leqq \Delta(\omega^n \mid \mu) - m$ follows now by the trick used at the end of the proof of Theorem 5.1. □

6. Uniform tests

In the previous sections we spoke about tests $d(\omega \mid \mu)$ depending both on ω and μ, but μ was always a fixed computable measure and d was required to have certain properties only with respect to ω. The restriction to computable measures would seem to many probabilists injustified since e.g. having the outcome of an experiment it would not be natural to look for a distribution fitting it only among the computable ones. LEVIN defined a general *uniform test* $d_L(\omega \mid \mu)$ in [11] which is a universal test for each fixed computable measure μ. Let us note that any of the expressions of tests by complexities given in the previous section could be taken as the definition of a uniform test. They are all semicomputable in (ω, μ), have some normedness property for each fixed μ, and are in some sense equivalent as long as only computable measures are concerned. But this is not obvious for noncomputable μ, and, in fact, LEVIN's test $d_L(\omega \mid \mu)$ discovers more nonrandom sequences than, say, $d_S(\omega \mid \mu)$. We do not want to repeat all the definitions from [11] necessary to a self-contained definition of the notion of a uniform test (*L-test*). Note that in [11] a *semimeasure* is defined as a concave functional on $C(\Omega)$ and in this section we adopt this definition, differing from that of Section 2. A semimeasure in the old sense is a restriction of a semimeasure in the new sense.

LEVIN's maximal (in the sense of $\leqq$) uniform test d_L as defined in [11] has the properties desirable of any test (we state them in terms of $t_L = \exp d_L$):

(i) $t(\omega \mid \mu)$ *is semicomputable in* (ω, μ); (ii) $\int t(\omega \mid \mu)\, \mu(d\omega) \leqq 1$ *for all* μ.

These properties imply that for any computable distribution μ there exists a constant $c_\mu < \infty$ with $d_L(\omega \mid \mu) \leqq d_C(\omega \mid \mu) + c_\mu$. Here d_C is the test defined in Corollary 3 of Theorem 5.1. On the other hand, it is easy to show that the test

$$d_C(\varphi \mid \psi) = \sup_{x \in \mathbf{N}_2^*} M_{\mathbf{N}}(x) \frac{\varphi(g_x)}{\psi(g_x)}$$

for semimeasures φ, ψ ($0/0 = 0$ by definition) is a uniform L-test. Hence one has

Theorem 6.1. *For any computable measure* μ *there exists a constant* c_μ *with*

$$|d_L(\omega \mid \mu) - d_C(\omega \mid \mu)| \leqq c_\mu.$$

Now I introduce a somewhat modified universal uniform test. Its definition is more explicit and semimeasures do not enter into it. Then I prove that this test also has the properties stated for LEVIN's tests in [11].

Definition. The function $t(\omega \mid \mu)$ is a *P-test* if (i), (ii) are satisfied and further (iii), (iv) holds.

(iii) *For any* $c > 0, \omega$, $\quad \mu c > \nu \Rightarrow t(\omega \mid \mu) < ct(\omega \mid \nu)$.

In other words, $1/t(\omega \mid \mu)$ can be extended to the set of all finite (unnormed) measures to a homogenuous, monotone function of μ.

(iv) $1/t(\omega \mid \mu)$ *is concave.*

Remark. All L-tests are P-tests when restricted to (ω, μ). The concavity requirement seems to be unmotivated but a consequence of it is acceptable as a general feature of tests: $t(\omega \mid \mu) \leqq c$, $t(\omega \mid \nu) \leqq c$ implies $t(\omega \mid \frac{1}{2}\mu + \frac{1}{2}\nu) \leqq c$.

For a P-test we define $t(\varphi \mid \mu) = \varphi(t(\cdot \mid \mu))$, $t(\varphi \mid \psi) = \sup_{\mu \geqq \psi} t(\varphi \mid \mu)$. ($\varphi, \psi$ are semimeasures, μ a measure.) Note that the integral $\varphi(f)$ of a lower semicontinuous function f by a semimeasure φ has a natural definition. Then we have

Theorem 6.2. *There exists a P-test* $t_P(\omega \mid \mu)$ *maximal with respect to* $\lesssim$.

The proof of this theorem does not differ from the proof of any other theorem of this type, so it can readily be omitted.

The relation between t_L and t_P is clear for measures: one has $t_L(\mu \mid \nu) \leqq t_P(\mu \mid \nu)$. The apriori probability M_Ω will retain its remarkable property for t_P:

Theorem 6.3. $t_P(\varphi \mid M_\Omega)$ *is bounded from above by a universal constant* $c < \infty$.

This theorem is proved analogously to Theorem 2 of [11].

Finally, we show that t_P also has the conservation property. A stochastic operator is a monotone linear operator $A: C(\Omega) \to C(\Omega)$ with $A(1) = 1$. Any stochastic operator, defined originally only on continuous functions, can easily be extended to all upper semicontinuous functions $f: \Omega \to \overline{\mathbf{R}}_+$.

Theorem 6.4. *For any computable* (see [11]) *stochastic operator* A *one has*

$$t_P(\varphi A \mid \psi A) \lesssim t_P(\varphi \mid \psi).$$

Proof. Let us define for any ω and measure μ

$$t^A(\omega \mid \mu) = (A t_P(\cdot \mid \mu A))\,(\omega).$$

It is not very hard to see that t_A is a P-test. (To prove the concavity requires some computation.) Hence we have $t_A \lesssim t_P$. Applying φ to this inequality we obtain

$$t_P(\varphi A \mid \mu A) = t^A(\varphi \mid \mu) \lesssim t_P(\varphi \mid \mu).$$

Now we are finished if we can show that $t_P(\varphi A \mid \psi A) = t^A(\varphi \mid \psi)$. We have

$$t_P(\varphi A \mid \psi A) = \sup_{\mu \geqq \psi A} t_P(\varphi A \mid \mu), \qquad t^A(\varphi \mid \psi) = \sup_{\mu \geqq \psi} t_P(\varphi A \mid \mu A).$$

By (iii) it is enough to show therefore that $\mu \geqq \psi A$ implies the existence of a $\nu \geqq \psi$ with $\mu \geqq \nu A$. Put $\nu g = \mu f$ for all g of the form Af. We have then $\mu = \nu A$. The definition of ν is unambiguous because $Ah = 0 \Rightarrow \mu h \geqq \psi A h = 0$. Hence $\mu \geqq 0$ on $KerA$, which implies $\mu = 0$ on $KerA$, i.e. μf depends only on Af. ν is positive (because of $\nu \geqq \psi$) and linear on ImA, $\nu(1) = 1$. Hence it is also continuous and it can be extended continuously to a measure on Ω while retaining the properties $\nu \geqq \psi$, $\mu = \nu A$. □

Remark. $t_L(\varphi \mid \psi)$ also has the property that $\dfrac{1}{t_L(\varphi \mid \psi)}$ is concave in ψ for any semimeasure φ. As to t_P, we can only assert that $\dfrac{1}{t_P(\mu \mid \psi)}$ is concave for any measure μ.

Finally we give a P-test (not necessarily maximal) which also has the conservation property and can be defined explicitly by a formula: Let $\{f_i\}_{i \in \mathbf{N}}$ be a recursive enumeration of all *positive* continuous functions $f_i \colon \Omega \to \mathbf{Q}$ assuming only a finite set of values. Let $f_i > 2^{-i}$. Then the test is defined as

$$t_0(\omega \mid \mu) = \Sigma_i M_{\mathbf{N}}(i) \cdot \frac{f_i(\omega)}{\mu(f_i)}.$$

The proof of the conservation property is straightforward.

Acknowledgement. I want here to express my thanks to Professor C. P. SCHNORR for inspiring discussions on the subject, and to L. A. LEVIN for having made accessible many of his unpublished ideas.

References

[1] MARTIN-LÖF, PER, The definition of random sequences. Information and Control **6** (1966), 602–619.
[2] ZVONKIN, A. K., and L. A. LEVIN, The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. Russ. Math. Surv. **25** (1970), 83.
[3] LEVIN, L. A., On the notion of a random sequence. Soviet Math. Dokl. **14** (1973), 1413.
[4] SCHNORR, C. P., Process complexity and effective random tests. J. Comput. Syst. Sci. **7** (1973), 376.
[5] KOLMOGOROV, A. N., Three approaches to the quantitative definition of information. Prob. Info. Transmission **1** (1965), 1.
[6] LEVIN, L. A., Some theorems on the algorithmic approach to probability theory and information theory. Ph. D. Thesis 1971.
[7] CHAITIN, G. J., A theory of program-size formally identical to information theory. J. A. C. M. **22** (1975), 329.
[8] CHAITIN, G. J., Algorithmic information theory. IBM J. Res. Dev. (July 1977), 350–359.
[9] SCHNORR, C. P., Unpublished manuscript.
[10] MARTIN-LÖF, PER, Algorithmen und zufällige Folgen. Lecture Notes, University of Erlangen 1966.
[11] LEVIN, L. A., Uniform tests of randomness. Soviet Math. Doklady **17** (1976), 337.
[12] GÁCS, P., On the symmetry of algorithmic information. Soviet Math. Dokl. **15** (1974), 1477–1480; Corrections ibid No. 6, V.

(Eingegangen am 15. Oktober 1978)