PETER J. GRABNER

Block distribution in random strings


<http://www.numdam.org/item?id=AIF_1993__43_2_539_0>
BLOCK-DISTRIBUTION IN RANDOM STRINGS

by Peter J. GRABNER

1. Introduction.

We investigate some properties of infinite sequences of independent random variables, which take the values 0 and 1 with probabilities $p$ and $q$ respectively (Bernoulli's scheme). It is one of the basic results of probability theory that the limit relation

$$
\lim_{N \to \infty} \frac{\# \{ 1 \leq n \leq N - k : x_n x_{n+1} \ldots x_{n+k} = a_1 \ldots a_k \}}{N} = \mu_k(A)
$$

holds in probability for all blocks $A = a_1 \ldots a_k$ of a given constant length $k$ ($\mu_k(A)$ is the $k$-fold product measure generated by $\mu(\{0\}) = p$ and $\mu(\{1\}) = q$). This result can also be naturally imbedded into ergodic theory: consider the infinite product space $X = \{0,1\}^\mathbb{N}$ equipped with the infinite product measure $\mu_\infty$ generated by $\mu$. Then the shift operator $S$ (Bernoulli shift) on $X$ defined by $S(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$ is an ergodic transformation on $X$ (cf. e.g. [Wa]) and the above relation is a consequence of Birkhoff's ergodic theorem.

It is now natural to ask how fast (depending on $N$) $k$ could grow such that this relation persists. In order to answer this question we introduce a
special notion of discrepancy (cf. [HI], [KN])

\[
D^k_N(x_1, \ldots, x_N) = \max_{A \in \{0,1\}^k} \frac{1}{\sqrt{p^k \mu_k(A)}} \left| \frac{\# \{1 \leq n \leq N-k : x_nx_{n+1} \ldots x_{n+k} = a_1 \ldots a_k \}}{N} \right| - \mu_k(A).
\]

The following calculations will show that this is a proper measure for the distribution behaviour of the sequence \(x_1, x_2, \ldots\). Note that this definition agrees with the definition in [FKT] for \(p = q = \frac{1}{2}\).

**Definition.** — A sequence \(x_1, x_2, \ldots\) is called \(k(N)\)-distributed with respect to \(\mu\) if

\[
\lim_{N \to \infty} D^k_N(x_1, \ldots, x_N) = 0.
\]

Our Theorem will show under which conditions almost all sequences are \(k(N)\)-distributed. Without loss of generality assume that \(p \leq q\). The notation \(\log_p n\) is the logarithm to base \(\frac{1}{p}\) : \(\log_p n = \log \frac{1}{p} n\).

**Theorem.** — Let \(k(N)\) be a non-decreasing sequence of positive integers. Then the following \(0\-1\)-law holds

\[
\mu_\infty \left( \lim_{N \to \infty} D^k_N(x_1, \ldots, x_N) = 0 \right) = \begin{cases} 
1 & \text{if } \log_p n - \log_p n - k(n) \to \infty \\
0 & \text{otherwise}.
\end{cases}
\]

It clearly follows from Kolmogoroff’s \(0\-1\)-law or the fact that the set

\[
\left\{ \lim_{N \to \infty} D^k_N(x_1, \ldots, x_N) = 0 \right\}
\]

is invariant under the (ergodic) shift \(S\), that the only possible values for the above probability are 0 and 1. The proof of this theorem will use bivariate correlation polynomials, which are a generalization of Guibas’ and Odlyzko’s correlation polynomials in one variable (cf. [GO]). Using these polynomials we are able to compute the probability generating functions of the events we are interested in.

2. Generating Functions.

Throughout this section let \(A = a_1a_2 \ldots a_k\) be a 0-1-string of length \(k\). We are interested in the cardinalities of the following subsets of the set
$S_{r,s}$ of strings containing $r$ digits 0 and $s$ digits 1:

\begin{equation}
S_{r,s}(r, s) = \# \{ B \in S_{r,s} : B \text{ contains } A \text{ only at the end} \}
\end{equation}

\begin{equation}
A_{r,s}(r, s) = \# \{ B \in S_{r,s} : B \text{ contains } A \text{ only at the beginning and at the end} \}
\end{equation}

\begin{equation}
h_A(r, s) = \# \{ B \in S_{r,s} : B \text{ does not contain } A \}.
\end{equation}

In order to compute the generating functions of these quantities we introduce the bivariate autocorrelation polynomial $[AA](z, w)$:

\begin{align*}
[AA](z, w) &= \left\{ \begin{array}{ll}
1 & \text{if } a_1 a_2 \ldots a_k = a_{r+s+1} a_{r+s+2} \ldots a_k \text{ and the string } a_1 a_2 \ldots a_{r+s} \text{ contains } r \text{ digits 0 and } s \text{ digits 1 } \\
0 & \text{otherwise,}
\end{array} \right.
\end{align*}

where $[z^r w^s]P(z, w)$ as usual denotes the coefficient of $z^r w^s$ in $P(z, w)$. We are now ready to formulate

**Proposition 1.** — The generating functions of the combinatorial expressions (2.1) are given by

\begin{align*}
F_A(z, w) &= \sum_{r,s=0}^{\infty} f_A(r, s) z^r w^s = \frac{z^{0(A)} w^{1(A)}}{z^{0(A)} w^{1(A)} + (1 - z - w)[AA](z, w)} \\
G_A(z, w) &= z^{0(A)} w^{1(A)} + \frac{(z + w - 1) z^{0(A)} w^{1(A)}}{z^{0(A)} w^{1(A)} + (1 - z - w)[AA](z, w)} \\
H_A(z, w) &= \frac{[AA](z, w)}{z^{0(A)} w^{1(A)} + (1 - z - w)[AA](z, w)},
\end{align*}

where $0(A)$ and $1(A)$ denote the number of 0's and 1's in $A$ respectively.

The proof of this proposition is analogous to the proof of the corresponding results for ordinary generating functions (cf. [GO]).

**Remark 1.** — Obviously these results can be generalized to any finite alphabet.

As in [FKT] we use these functions to compute the probability generating function (p.g.f.) of all strings containing the substring $A$ exactly $r$ times:

\begin{align*}
\Phi_A^{(r)}(z) &= \frac{z^{-kr}}{\mu_k(A)} F_A(pz, qz)^2 G_A(pz, qz)^{r-1} \text{ for } r \geq 1 \\
\Phi_A^{(0)}(z) &= H_A(pz, qz).
\end{align*}

Inserting the results of Proposition 1 and setting

\begin{equation}
P(z) = \frac{1}{\mu_k(A)} [AA](pz, qz)
\end{equation}
yields
\[
\Phi_A^{(r)}(z) = \frac{z^k}{\mu_k(A)} \frac{((1 - z)(P(z) - \frac{1}{\mu_k(A)}) + z^k)^{r-1}}{((1 - z)P(z) + z^k)^{r+1}}
\]
\[
\Phi_A^{(0)}(z) = \frac{P(z)}{(1 - z)P(z) + z^k}.
\]

3. Proof of the Theorem.

We split the proof into two parts; first we show that almost all sequences are \( k(N) \) distributed if \( \lambda p n - \lambda p \lambda p n - k(n) \to \infty \). Using our p.g.f. results we can write
\[
\mu_\infty (\#\{0 \leq n \leq N - k : x_{n+1} \ldots x_{n+k} = a_1 \ldots a_k \} = r) = p_A^{(r)}(N) = [z^N]\Phi_A^{(r)}(z) = \frac{1}{2\pi i} \oint_C \Phi_A^{(r)}(z) \frac{dz}{z^{N+1}}.
\]
In order to be able to estimate the integral we need information on the the zeros of the polynomial \((1 - z)P(z) + z^k\).

**Lemma 1.** — The zero of smallest modulus \( z_0 \) of \((1 - z)P(z) + z^k\)
is real and positive and satisfies the estimate
\[
z_0 > 1 + C\mu_k(A)
\]
for a positive constant \( C \) only depending on \( p \).

**Proof.** — As \( F_A(pz, qz) \) is a p.g.f. and \((1 - z)P(z) + z^k\) is the denominator of this rational function the zero of smallest modulus has to be positive and \( \geq 1 \). Investigation of the derivative shows the existence of the constant \( C \).

Let now
\[
k(n) = \lambda p n - \lambda p \lambda p n - \lambda p \psi(n),
\]
where \( \psi(n) \to \infty \). We need estimates for the probability that the number of occurrences \( Z_N(A) \) of a block \( A \) deviates too far from the mean value:
\[
L_N(\delta_A) = \mu_\infty (Z_N(A) < N\mu_k(A)(1 - \delta_A)) \quad \text{and} \quad
U_N(\delta_A) = \mu_\infty (Z_N(A) > N\mu_k(A)(1 + \delta_A)).
\]
These probabilities are sums of the $p_A^{(r)}(N)$ defined in (3.1):

$$L_N(\delta_A) = \sum_{r < N \mu_k(A)(1-\delta_A)} p_A^{(r)}(N) \quad \text{and}$$

$$U_N(\delta_A) = \sum_{r > N \mu_k(A)(1+\delta_A)} p_A^{(r)}(N).$$

(3.3)

We will use the integral representation (3.1) to estimate these quantities.

For convenience we now introduce some notations

$$Q(z) = (1-z)P(z) + z^k$$

(3.4)

$$a(z) = \frac{z^k}{Q(z)^2}, \quad b(z) = 1 + \frac{z - 1}{\mu_k(A)Q(z)}.$$  

This gives

$$\Phi_A^{(r)}(z) = \frac{1}{\mu_k(A)} a(z)b(z)^{r-1}$$

for $r \geq 1$. Observe further that

$$a(1 \pm \varepsilon) = 1 + O \left( \frac{\varepsilon}{\mu_k(A)} \right)$$

$$b(1 \pm \varepsilon) = 1 \pm \frac{\varepsilon}{\mu_k(A)} + O \left( \frac{\varepsilon^2}{\mu_k(A)^2} \right)$$

$$b^j(1 \pm \varepsilon) = \exp \left( \pm \frac{\varepsilon j}{\mu_k(A)} + O \left( \frac{\varepsilon^2 j}{\mu_k(A)^2} \right) \right)$$

$$\left( 1 \pm \varepsilon \right)^{-n} = \exp \left( \mp n \varepsilon + O(\varepsilon^2) \right).$$

(3.5)

We can now write

$$U_N(\delta_A) = \frac{1}{2\pi i} \oint_C \frac{1}{\mu_k(A)} a(z) \frac{b^j(z)}{1-b(z)} \frac{dz}{z^{N+1}},$$

where $j = [N\mu_k(A)(1 + \delta_A)]$. As all the power series involved have positive coefficients and because of Lemma 1 we can estimate

$$U_N(\delta_A) \leq \frac{1}{\mu_k(A)} a(1-\varepsilon) \frac{b^j(1-\varepsilon)}{1-b(1-\varepsilon)} (1-\varepsilon)^{-N}$$

for every positive $\varepsilon < C\mu_k(A)$. Using (3.5) yields

$$U_N(\delta_A) \leq \frac{1 + O \left( \frac{\varepsilon}{\mu_k(A)} \right)}{1 + O \left( \frac{\varepsilon}{\mu_k(A)} \right)} \exp \left( \left( N - \frac{j}{\mu_k(A)} \right) \varepsilon + O \left( \frac{\varepsilon^2 j}{\mu_k(A)^2} \right) + O(\varepsilon^2) \right).$$

(3.6)

Inserting $\varepsilon = \left( \frac{\mu_k(A) \ln N}{N} \right)^{\frac{1}{2}}$ into the above inequality yields

$$U_N(\delta_A) \leq \exp(-\delta_A (N \mu_k(A) \ln N)^{\frac{1}{2}} + C_1 \ln N).$$
In the same way we treat the lower tail. Let now \( j = [N\mu_k(A)(1-\delta_A)] \).

Thus we obtain

\[
L_N(\delta_A) = \frac{1}{2\pi i} \int_C \left( \frac{P(z)}{Q(z)} + \frac{a(z)}{\mu_k(A)} \frac{b^j(z) - 1}{b(z) - 1} \right) \frac{dz}{z^{N+1}}.
\]

We can now estimate

\[
L_N(\delta_A) \leq \frac{P(1+\epsilon)}{Q(1+\epsilon)}(1+\epsilon)^{-N} + \frac{1}{\mu_k(A)} b^j(1+\epsilon)a(1+\epsilon)(1+\epsilon)^{-N}.
\]

Using the same value for \( \epsilon \) as above yields

\[
(3.7) \quad L_N(\delta_A) \leq \exp\left(-\delta_A (N\mu_k(A) \log N)^{\frac{1}{2}} + C_2 \log N \right).
\]

Combining this with (3.6) yields

\[
(3.8) \quad \mu_\infty\left( \left| \frac{Z_N(A)}{N} - \mu_k(A) \right| > \delta_A \mu_k(A) \right) \\
\leq \exp(-\delta_A (N\mu_k(A) \log N)^{\frac{1}{2}} + C_3 \log N).
\]

Let now \( \delta_A = \delta \left( \frac{p^k}{\mu_k(A)} \right)^{\frac{1}{2}} \) and observe that \( p^k = \frac{\log N}{N} \psi(N) \).

Therefore we have

\[
(3.9) \quad \mu_\infty(D_N^{k(N)}(\omega) > \delta) \leq 2^{k(N)} \exp(-\delta \psi(N)^{\frac{1}{2}} \log N + C_3 \log N) \\
\leq \exp(-\delta \psi(N)^{\frac{1}{2}} \log N + C' \log N).
\]

We now choose \( \delta \) as a function of \( N \)

\[
\delta = \psi(N)^{-\frac{1}{4}}
\]

and observe that

\[
\sum_{N=1}^{\infty} \exp(-\psi(N)^{\frac{1}{2}} \log N + C' \log N) < \infty.
\]

Thus by the Borel-Cantelli lemma (cf. [Fe]), we obtain the first part of our Theorem.

We now have to prove that almost no series are \( k(N) \)-distributed if \( \log n - \log \log n - k(n) \not\to \infty \) (we confine ourselves to the case \( p < \frac{1}{2} \), because the case \( p = \frac{1}{2} \) has been treated by Grill [Gr]). We introduce a set \( \mathcal{A} \) of strings of length \( k \), which have only trivial autocorrelation and do not overlap each other:

\[
\mathcal{A} = \left\{ \underbrace{0 \ldots 0}_{l} \underbrace{A}_{l+d(k)-2} \underbrace{1 \ldots 1}_{l} \right\},
\]
where $l = \left\{ \frac{k}{3} \right\} + 1$ and $d(k) = k \mod 3$. We need the p.g.f. $\varphi(z)$ of all strings not containing an element of $A$. This function satisfies the equations

\[
\varphi(z) + \varphi_{A_1}(z) + \ldots + \varphi_{A_m}(z) = z\varphi(z) + 1
\]

\[
\varphi_{A_1}(z) = z^k \mu_k(A_1) \varphi(z)
\]

\[
\varphi_{A_m}(z) = z^k \mu_k(A_m) \varphi(z),
\]

where $A_1, \ldots, A_m$ are the elements of $A$ and $\varphi_{A_l}(z)(l = 1, \ldots, m)$ is the p.g.f. of the blocks ending with $A_l$ but containing no further occurrence of any element of $A$. Solving these equations yields

\[
(3.10) \quad \varphi(z) = \frac{1}{1 - z + \mu_k(A)z^k}.
\]

Note that the simplicity of these equations comes from the trivial overlap structure of the elements of $A$.

Because of this simple overlap structure it is easy to see that

\[
(3.11) \quad \phi_{j_1 \ldots j_m}(z) = \frac{(j_1 + \ldots + j_m)!}{j_1! \ldots j_m!} \mu_k(A_1)^{j_1} \ldots \mu_k(A_m)^{j_m} z^{k(j_1 + \ldots + j_m)} \varphi(z)^{j_1 + \ldots + j_m + 1}
\]

is the p.g.f. of all blocks containing $A_l$ exactly $j_l$ times ($l = 1 \ldots m$). As in the first part of the proof we use

\[
(3.12) \quad MN(\delta) = \mu_\infty \left( |Z_N(A_l) - N\mu_k(A_l)| \leq N\mu_k(A_l)\delta_{A_l}, l = 1 \ldots m \right)
\]

\[
= \frac{1}{2\pi i} \oint_C \sum_{|z| \leq N\mu_k(A_l)\delta_{A_l}} \varphi_{j_1 \ldots j_m}(z) \frac{dz}{z^{N+1}},
\]

where $\delta_{A_l} = \delta \left( \frac{p^k}{\mu_k(A_l)} \right)^{\frac{1}{2}}$.

We want to treat (3.12) exactly like the corresponding expressions in the first part of the proof. For this purpose we need information on the zeros of the polynomial $1 - z + \mu_k(A)z^k$.

**Lemma 2.** — The zero of smallest modulus $z_0$ of $1 - z + \mu_k(A)z^k$ is real and satisfies

\[
z_0 > 1 + \mu_k(A).
\]

**Proof.** — The proof of the first statement is as in the proof of Lemma 1. For the proof of the inequality insert $z = 1 + \mu_k(A)$ into the polynomial.
Observe now that

$$\frac{1}{2\pi i} \oint_C \varphi(z)^{J+1} \frac{dz}{z^{N-kJ+1}} \leq \varphi(1+\varepsilon)^{J+1}(1+\varepsilon)^{kJ-N}$$

for $\varepsilon \leq \mu_k(A)$. Inserting $\varepsilon = \mu_k(A) - \frac{J}{N}$ and performing similar calculations as in the first part of the proof yields

(3.13)

$$\frac{1}{2\pi i} \oint_C \varphi(z)^{J+1} \frac{dz}{z^{N-kJ+1}} \leq \frac{1}{\mu_k(A)^{J+1}} \exp \left( - \frac{(J-N\mu_k(A))^2}{2N\mu_k(A)} + O(k\mu_k(A)^2N) \right).$$

Let now $n = N\mu_k(A)$, $J = j_1 + \cdots + j_m$ and $p_i = \frac{\mu_k(A_i)}{\mu_k(A)}$ and insert (3.13) into (3.12) to obtain

(3.14)

$$M_N(\delta) \leq \frac{1}{\mu_k(A)} \sum_{|j_1 - n|, \ldots, \leq n} \frac{J!}{j_1! \cdots j_m!} \prod_{l=1}^m p_l^{j_l} \exp \left( - \frac{(J-n)^2}{2n} + O(k\mu_k(A)n) \right),$$

where $\sum_{l=1}^m p_l = 1$. Thus we have arrived at an expression that we can treat by the normal approximation of the multinomial distribution.

Assume that $N$ runs through a subsequence of $\mathbb{N}$ such that

$$\lim_{N \to \infty} \max \{ \log \log N - k(N) \} = \limsup_{N \to \infty} (\log \log N - k(N)) = \log C < \infty.$$

It will suffice to prove our theorem for the case that $\limsup_{N \to \infty} \log \log N - k(N) > -\infty$, such that $0 < C < \infty$. Observe now that $N\mu_k(A_i)\delta_{A_i} = \delta \sqrt{CN\mu_k(A_i)} \log N$ and use Stirling’s formula

$$\left( \frac{n}{e} \right)^n \sqrt{2\pi n} \leq n! \leq \frac{11}{10} \left( \frac{n}{e} \right)^n \sqrt{2\pi n}$$

and
to obtain
\[(3.15)\]
\[
M_N(\delta) \leq \frac{11 \exp(O(k\mu_k(A)n))}{10 \mu_k(A)\sqrt{2\pi}} \sum_{|j_i - np_i| \leq \delta \sqrt{CN\mu_k(A)}} \sqrt{J} \frac{J^j}{j_1 \cdots j_m} \prod_{i=1}^m p_i^{j_i} \exp\left(-\frac{(J-n)^2}{2n}\right)
\]
\[= \frac{11}{10} \frac{\exp(O(k\mu_k(A)n))}{n^{m-12} \mu_k(A) (2\pi)^{m-1} \sqrt{p_1 \cdots p_m}} \sum_{|x| \leq \delta A_i} \frac{\sqrt{1+n}}{(1+x_1) \cdots (1+x_m)} \]
\[
\times \exp\left(-\frac{n}{2} \sum_{i=1}^m p_i x_i^2 + O(n\eta^3) + O\left(n \sum_{i=1}^m p_i x_i^2\right)\right),
\]
where \(j_i = np_i(1 + x_i)\) and \(J = n(1 + \eta)\). The terms in the last exponential come from \((1 + x)^{1+x} = \exp\left(x + \frac{x^2}{2} + O(x^3)\right)\) for \(x \to 0\) and the observation that \(j_1 + \cdots + j_m = J\) transforms to \(\sum p_i x_i = \eta\). In the following we will use \(p < \frac{1}{2}\) which yields \(\delta A_i \to 0\) for our choice of \(A\) (in the case \(p = \frac{1}{2}\) we have \(\delta A_i = \delta\) and the following arguments cannot be used).

Inserting the definition of \(\delta A_i\) yields the estimate
\[(3.16)\]
\[
\left|n \sum_{i=1}^m p_i x_i^3\right| \leq \frac{(\log N)^{\frac{3}{2}}}{\sqrt{n}} \sum_{i=1}^m \frac{1}{\sqrt{p_i}}
\]
\[
= O(N^{-\frac{3}{2} + \frac{1}{3} \log q + \frac{1}{3} \log (\sqrt{\frac{3}{p}} + \frac{1}{2} q)} (\log N)^{\frac{3}{2} - \frac{3}{8} \log q - \frac{3}{2} \log (\sqrt{\frac{3}{p}} + \frac{1}{2} q)}
\]
and a similar estimate holds for \(n\eta^3\). Using an exponential estimate yields
\[
\frac{1}{\sqrt{1+x_1} \cdots (1+x_m)}
\]
\[
= \exp(O(N^{\frac{3}{2} + \frac{1}{3} \log q + \frac{1}{3} \log (\sqrt{\frac{3}{p}} + \frac{1}{2} q)} (\log N)^{\frac{3}{2} - \frac{3}{8} \log q - \frac{3}{2} \log (\sqrt{\frac{3}{p}} + \frac{1}{2} q)})
\].

Inserting these inequalities into (3.15) and setting \(\alpha = \frac{1}{6} \log q + \frac{1}{3} \log (\sqrt{\frac{3}{p}} + \frac{1}{2} q)\) yields
\[
M_N(\delta) \leq \exp(O(N^{-\frac{3}{2} + \alpha (\log N)^{\frac{3}{2} - \alpha}})) n^{\frac{m+1}{2}} \sqrt{p_1 \cdots p_m}
\]
\[\times \sum_{|x| \leq \delta A_i} \exp\left(-\frac{n}{2} \sum_{i=1}^m p_i x_i^2\right) \left(\frac{1}{(np_1) \cdots (np_m)}\right).
\]
The sum in the last line can be interpreted as a lower Riemann sum for the integral
\[ \int_{|x| \leq \delta_A} \exp \left( -\frac{n}{2} \sum_{l=1}^{m} p_l x_l^2 \right) dx_1 \cdots dx_m \]
using the lattice
\[ \left\{ (x_1, \ldots, x_m) \mid x_l = \frac{j_l}{np_l} - 1, \quad |x_l| \leq \delta_A, \quad l = 1, \ldots, m \right\}. \]
Thus we obtain
\[(3.17)\]
\[ M_N(\delta) \leq \exp \left( O(N^{-\frac{1}{3} + \alpha} (\log N)^{\frac{3}{4} - \alpha}) + O(\log N) \right) (\Phi(\delta \sqrt{C \log N})^m, \]
where
\[ \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-t^2/2} dt \sim 1 - \sqrt{\frac{2}{\pi}} \frac{1}{x} e^{-x^2} \]
for \( x \to \infty \). Therefore we can estimate
\[(3.18)\]
\[ M_N(\delta) \leq \exp \left( -\frac{\sqrt{2}}{\delta \sqrt{C \log N} e^{\frac{1}{6} \delta^2 C \log N}} + O(N^{-\frac{1}{3} + \alpha} (\log N)^{\frac{3}{4} - \alpha}) + O(\log N) \right). \]
Observe now
\[(3.19)\]
\[ m \asymp \left( \frac{N}{\log N} \right)^{\frac{1}{3} \log p} \]
and
\[ \mu_k(A) \asymp \left( \frac{\log N}{N} \right)^{-\frac{1}{3} \log p} \]
\[ n \asymp N^{1 + \frac{1}{3} \log p} (\log N)^{-\frac{1}{3} \log p}. \]
Inserting these estimates into (3.18) yields
\[ M_N(\delta) \leq \exp \left( -D \frac{N^{\frac{1}{3} \log p} - \frac{1}{6} \delta^2 C \log e}{\delta (\log N)^{\frac{1}{3} + \frac{1}{3} \log p}} + O(N^{-\frac{1}{3} + \alpha} (\log N)^{\frac{3}{4} - \alpha}) + O(\log N) \right), \]
where \( D > 0 \) is a constant implied by (3.19). The right hand side tends to 0 for sufficiently small \( \delta > 0 \), because \( \frac{1}{3} \log p > -\frac{1}{3} + \alpha \) holds for \( p < \frac{1}{2} \).

Note that
\[ \mu_\infty(D^k_N(\omega) < \delta) \leq M_N(\delta). \]
Thus the proof is complete.
Remark 2. — Modifying (1.1) one can also investigate discrepancies

\[ D_{N}^{k,\phi}(\omega) \]

\[ \max_{A \in \{0, 1\}^k} \sqrt{\frac{\phi(k)}{\mu_k(A)}} \left( \frac{\#\{1 \leq n \leq N - k : x_n x_{n+1} \cdots x_{n+k} = A\}}{N} - \mu_k(A) \right), \]

where \( \phi \) is a monotonically increasing function. Then the same calculations as above yield

\[ \mu_\infty \left( \lim_{N \to \infty} D_{N}^{k(N),\phi}(\omega) = 0 \right) = \begin{cases} 1 & \text{if } \lim_{N \to \infty} \frac{N\phi(k(N))}{\log N} = \infty \\ 0 & \text{otherwise.} \end{cases} \]

This answers a question posed by Flajolet, Kirschenhofer and Tichy [FKT], Remark 2.

**BIBLIOGRAPHY**


Manuscrit reçu le 13 avril 1992,

Peter J. GRABNER,
Institut für Mathematik
Technische Universität
Steyrergasse 30
8010 Graz (Austria).