J. H. B. KEMPERMAN

Probability methods in the theory of distributions modulo one

Compositio Mathematica, tome 16 (1964), p. 106-137
<http://www.numdam.org/item?id=CM_1964__16__106_0>

Probability methods in the theory of distributions modulo one*

by

J. H. B. Kemperman

1. Introduction

One of the purposes of the present paper is to demonstrate that many of the methods and results of probability theory play an important role in the theory of distributions modulo 1, as is also apparent from the work of Cassels [1], Cigler [3], Hlawka [10], Kac [11], Kesten [13], [14], Stapleton [19] and many others.

The following sections 2, 3 and 5 are entirely expository. The results in the sections 6, 7 and 8 are new, and also some of the results in the sections 4 and 9.

2. Random variables

Many problems in the theory of distributions modulo 1 involve the asymptotic behavior of the sum

$$R_n = \sum_{k=1}^{n} g(x_k); \tag{2.1}$$

here, $g(x)$ is a given real and bounded Borel measurable function satisfying

$$g(x+1) = g(x), \quad \int_0^1 g(x)dx = 0.$$

Quite often, $x_k = x_k(\theta)$ $(k = 1, 2, \ldots)$ depends on a parameter $0 \leqq \theta \leqq 1$ and one is interested in statements on $\{R_n(\theta)\}$ holding at least for almost all θ.

More generally, consider a measure space

$$(\Omega, \mathscr{B}, \mu), \tag{2.2}$$

($\mathscr{B}$ a σ-field of subsets of Ω, μ a nonnegative measure on $\mathscr{B}$), which

*) Nijenrode lecture.

Some of this work was supported by the National Science Foundation under G–24470.

is a probability space in the sense that $\mu(\Omega) = 1$; example: $\Omega = [0, 1]$ with $\mathscr{B}$ as the σ-field of all Borel subsets and μ as the Lebesgue measure.

Suppose further that $x_k = x_k(\theta)$ is a given real and measurable function on Ω. Here, a function $x(\theta)$ from Ω to a given topological space (say, the reals) is said to be measurable if for each open subset U of this space one has $\{\theta\colon x(\theta) \in U\} \in \mathscr{B}$. Relative to the probability space (2.2), (which is kept fixed in most problems), such a measurable function is also called a *random variable.*

If $x(\theta)$ is real-valued then its so-called distribution function is defined as

$$F(z) = \mu\{\theta\colon x(\theta) \leqq z\}, \qquad -\infty < z < +\infty.$$

In particular, one may be interested in the asymptotic behavior of the distribution function

$$F_n(z) = \mu\{\theta\colon R_n(\theta) \leqq z\}$$

of the random variable $R_n = R_n(\theta)$, when n is large.

For instance, if $\Omega = [0, 1]$ as above and

$$x_k(\theta) = a^k\theta, \qquad (k = 1, 2, \ldots), \tag{2.3}$$

with $a \geqq 2$ as a fixed integer, then [11] one has for a large class of functions g that, for each $-\infty < z < +\infty$,

$$\lim_{n\to\infty} F_n(\sigma z\sqrt{n}) = \Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-u^2/2}\,du; \tag{2.4}$$

here, σ denotes a positive constant depending on g.

Or, take Ω as the unit square of points $\theta = (\theta', \theta'')$ together with the σ-field of Borel sets and the Lebesgue measure μ. Let further

$$x_k(\theta) = k\theta' + \theta''.$$

Then, as was shown by Kesten [13], [14], one has for $g(x) = x - [x] - \frac{1}{2}$, and also when $g(x)+c$ $(0 \leqq x < 1)$ is equal to the characteristic function of an interval, that

$$\lim_{n\to\infty} F_n(\sigma z \log n) = \frac{1}{\pi}\int_{-\infty}^{z}\frac{du}{1+u^2},$$

with σ as a positive constant.

3. Independent random variables

Relative to a fixed probability space (2.2), a collection $\{y_k, k \in I\}$ of random variables (sometimes called a stochastic process) is said to be a collection of *independent* random variables if

$$(3.1)\quad \mu\{\theta : y_{k_1}(\theta) \in U_1, \ldots, y_{k_r}(\theta) \in U_r\} = \prod_{\nu=1}^{r} \mu\{\theta : y_{k_\nu} \in U_\nu\},$$

for each choice of the finitely many distinct $k_\nu \in I$ and each choice of the open sets U_ν. In the special case, where each $y_k(\theta)$ assumes at most denumerably many values, this is equivalent to

$$(3.2)\quad \mu\{\theta: y_{k_1}(\theta) = a_1, \ldots, y_{k_r}(\theta) = a_r\} = \prod_{\nu=1}^{r} \mu\{\theta: y_{k_\nu} = a_\nu\}.$$

Now suppose that $\{y_k = y_k(\theta), k = 1, 2, \ldots\}$ is a *uniformly bounded* sequence of real-valued and *independent* random variables. Further, put

$$m_k = \int y_k(\theta)\mu(d\theta), \quad \sigma_k^2 = \int (y_k(\theta)-m_k)^2\mu(d\theta),$$

and

$$s_n = (\sigma_1^2 + \ldots + \sigma_n^2)^{1/2}.$$

Then one has the following important result, compare [8] and [16].

THEOREM 3.1. *Assume in addition that* $s_n \to \infty$. *Then*

$$(3.3)\qquad \lim_{n\to\infty} \mu\{\theta: \sum_{k=1}^{n} (y_k(\theta)-m_k) \leqq z s_n\} = \Phi(z),$$

for each fixed real value z.

Let further $\psi(t)$ *be any non-decreasing positive function. Then, for almost* $[\mu]$ *all* $\theta \in \Omega$, *the inequality*

$$(3.4)\qquad \sum_{k=1}^{n} (y_k(\theta)-m_k) > \sqrt{s_n^2\, \psi(s_n^2)}$$

holds for only finitely many or for infinitely many n, *depending on whether the integral*

$$(3.5)\qquad \int_1^\infty e^{-\frac{1}{2}\psi(t)} \sqrt{\psi(t)}\, dt/t$$

converges or diverges.

Usually, one takes $\psi(t)$ as one of the functions

$\psi(t) = (2+\delta)\log\log t,$

$\psi(t) = 2\log\log t + (3+\delta)\log\log\log t,$

$\psi(t) = 2\log_2 t + 3\log_3 t + 2\log_4 t + \ldots + 2\log_{r-1} t + (2+\delta)\log_r t,$

$(r \geqq 4)$; for these functions, (3.5) converges or diverges according to whether $\delta > 0$ or $\delta \leqq 0$.

Theorem 3.1 applies for instance to the example (2.3). Thus, let $\Omega = [0, 1]$ with Lebesgue measure μ, and let $a \geqq 2$ be a fixed integer. Let $u_k(\theta)$, $k = 1, 2, \ldots$, denote the sequence of random variables defined (for almost $[\mu]$ all θ) by

$$\theta = \sum_{k=1}^{\infty} u_k(\theta) a^{-k}, \qquad (0 \leqq \theta < 1), \tag{3.6}$$

and

$$u_k(\theta) \in \{0, 1, \ldots, a-1\}.$$

Using the criterion (3.2), it is easily seen that the $u_k(\theta)$ are *independent* random variables such that, for all $k = 1, 2, \ldots$,

$$\mu\{\theta \colon u_k(\theta) = j\} = a^{-1}, \qquad j = 0, 1, \ldots, a-1. \tag{3.7}$$

Now, consider

$$R_n(\theta) = \sum_{k=1}^{n} g(x_k(\theta)),$$

with

$$x_k(\theta) = a^{k-1}\theta, \qquad g(x) = x-[x]-\tfrac{1}{2} = ((x))-\tfrac{1}{2}. \tag{3.8}$$

By (3.6),

$$((a^k\theta)) = \sum_{j=k+1}^{\infty} u_j(\theta) a^{k-j} = \sum_{j=k+1}^{n} u_j(\theta) a^{k-j} + a^{k-n}((a^n\theta)),$$

$(k < n)$, hence,

$$R_n(\theta) = \sum_{k=1}^{n} y_k(\theta) + \varepsilon_n(\theta), \tag{3.9}$$

where

$$y_k(\theta) = \frac{1-a^{-k}}{a-1}\left(u_k(\theta) - \frac{a-1}{2}\right), \quad \varepsilon_n(\theta) = \frac{1-a^{-n}}{a-1}\{((a^n\theta))-\tfrac{1}{2}\}.$$

Note that $|\varepsilon_j(\theta)| \leqq \frac{1}{2}(a-1)^{-1}$. Clearly, the sequence $\{y_k(\theta)\}$ satisfies the conditions of Theorem 3.1 with $m_k = 0$ and

$$\sigma_k^2 = \frac{1}{a} \sum_{\nu=0}^{a-1} \left\{\frac{1-a^{-k}}{a-1}\left(\nu - \frac{a-1}{2}\right)\right\}^2 = \frac{1}{12}\frac{a+1}{a-1}(1-a^{-k})^2.$$

Thus,

$$s_n^2 = \sum_{k=1}^{n} \sigma_k^2 = \frac{1}{12}\frac{a+1}{a-1}\left(n - \frac{2a-1}{a^2-1} + O(a^{-n})\right),$$

hence,

$$s_n = \sqrt{\frac{a+1}{a-1}\frac{n}{12}} + O(n^{-1/2}),\ \log\log s_n^2 = \log\log n + O(1/\log n).$$

It follows by (3.3) that, for each fixed real number z,

$$\lim_{n\to\infty} \mu\{\theta\colon R_n(\theta) \leqq z\sqrt{n}\} = \Phi(z\sqrt{12(a-1)/(a+1)}).$$

This is a special case of (2.4). Moreover, by (3.4) and (3.9), we have for almost all θ that

$$R_n(\theta)^2 \left[\frac{a+1}{a-1}\frac{n}{12}\right]^{-1} > 2\log_2 n + 3\log_3 n + 2\log_4 n + \ldots + 2\log_{r-1} n$$
$$+ (2+\delta)\log_r n + 0(1/\log n),$$

($r \geqq 4$), for only finitely many n or for infinitely many n depending on whether $\delta > 0$ or $\delta = 0$. In particular,

$$\overline{\lim_{n\to\infty}}\, R_n(\theta)/\sqrt{n\log\log n} = \sqrt{\frac{1}{6}\frac{a+1}{a-1}},$$

for almost all $0 \leqq \theta \leqq 1$. Similarly, (replacing y_k by $-y_k$),

$$\underline{\lim_{n\to\infty}}\, R_n(\theta)/\sqrt{n\log\log n} = -\sqrt{\frac{1}{6}\frac{a+1}{a-1}}.$$

4. Infinite product spaces

Let G denote a fixed compact Hausdorff space. When considering G as a measurable space, we shall *always* mean the pair $(G, \mathscr{B})$, where $\mathscr{B}$ denotes the σ-field of all Borel subsets of G, that is, the smallest σ-field containing all open subsets of G. Similarly, by a measure ν on G we shall always mean a finite (usually, nonnegative) measure on $\mathscr{B}$ which, moreover, is regular; (for ν nonnegative this means that to each $B \in \mathscr{B}$ and each $\varepsilon > 0$ there corresponds a closed subset F of B such that $\nu(F) > \nu(B)-\varepsilon$.)

Let G_k $(k = 0, 1, \ldots)$ be a copy of G and consider the infinite product

$$G^\infty = G_0 \times G_1 \times \ldots$$

which is again a compact space. Each point $\theta \in G^\infty$ may be regarded as an infinite sequence

$$\theta = (\theta_0, \theta_1, \theta_2, \ldots)$$

of points θ_k in G; θ_k will be called the k-th coordinate of $\theta \in G^\infty$. Equivalently,

$$\theta = (x_0(\theta), x_1(\theta), x_2(\theta), \ldots), \tag{4.1}$$

where $x_k(\theta)$ denotes the measurable function on G^∞ defined by

$$x_k(\theta) = \theta_k, \qquad k = 0, 1, \ldots$$

Next, let ν_∞ be any probability measure on G^∞, thus, (G^∞, ν_∞) is a probability space. Relative to this probability space, each coordinate function $x_k(\theta)$ is a random variable taking values in G and with

$$\mu_k(U) = \nu_\infty\{\theta \colon \theta_k \in U\}, \qquad U \subset G,$$

as its so-called probability distribution.

Now consider the special case that ν_∞ is a direct product, that is,

$$\nu_\infty = \mu_0 \times \mu_1 \times \mu_2 \times \ldots, \tag{4.2}$$

(where μ_k is regarded as a probability measure on the copy G_k of G). Then the random variables $x_k(\theta)$ $(k = 0, 1, 2, \ldots)$ are *independent* (and conversely). Hence, so are the *real-valued* random variables $y_k = f(x_k(\theta)) = f(\theta_k)$ $(k = 0, 1, \ldots)$, where $f(x)$ is a given real-valued and Borel measurable function on G.

Let us assume that $f(x)$ is also *bounded*, thus, one can apply to the sequence $\{y_k\}$ the assertions of Theorem 3.1, provided that $s_n \to \infty$; (if s_n is bounded then the series $\sum(y_k(\theta)-m_k)$ converges for almost all θ, see [16] p. 236).

For convenience, let us consider the still more special case that ν_∞ is defined by (4.2) with $\mu_k = \mu_1$ for all k; (the random variables $x_k(\theta)$ are then said to be equidistributed). One obtains by (3.4) that, for almost all $[\nu_\infty]$ sequences (4.1), (the exceptional set depending on f), one has

$$\frac{1}{n}\sum_{k=1}^{n} f(x_k) = \int_G f(x)\mu_1(dx) + O\left(\sqrt{\frac{1}{n}\log\log n}\right). \tag{4.3}$$

Here, the *remainder cannot be improved*, except for the trivial case that $f(x)$ is equal to a constant c for almost $[\mu_1]$ all $x \in G$. In particular, we have for almost all $[\nu_\infty]$ sequences (4.1) that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(x_k) = \int_G f(x)\mu_1(dx). \tag{4.4}$$

This much holds for any real and Borel measurable function f on G, such that the integral in (4.4) exists, namely, by the so-called strong law of large numbers, see [16] p. 239.

From now on, let us assume that the topology of the compact space G has a countable base, (in other words, G is a metricspace).

It then follows from (4.4) that *almost all* $[\nu_\infty]$ *sequences* (4.1) *have the asymptotic distribution* μ_1, (hence, at least one does), in the sense that (4.4) holds for each $f \in C(G)$.

Here, $C(G)$ denotes the collection of all complex-valued continuous functions on G. We shall also regard $C(G)$ as a Banach space with norm

$$||f|| = \sup_{x \varepsilon G} |f(x)|.$$

Because G is metric there exists a *denumerable* collection $\{f_i\}$ of real-valued $f_i \in C(G)$ such that the finite linear combinations of the f_i form a dense subset of $C(G)$. If (4.4) holds for each f_i in such a collection it automatically holds for each $f \in C(G)$. This proves the above statement in italics.

The result (4.3) is more or less known; sometimes the remainder can even be shown to be uniform with respect to a class of measurable functions f, see Cassels [1].

Let us now demonstrate an analogous result for averages of the type

$$g_n(\theta) = \sum_{k=0}^{\infty} a_{nk} f(\theta_k, \theta_{k+1}, \ldots, \theta_{k+p-1}). \tag{4.5}$$

Here, the a_{nk} ($n, k = 0, 1, \ldots$) are given real numbers. Further, $f(\xi_1, \ldots, \xi_p)$ is a *given* real-valued and measurable function on the p-fold direct product $G \times \ldots \times G$ ($p \geqq 1$ fixed), such that

$$|f(\xi_1, \ldots, \xi_p)| \leqq 1 \tag{4.6}$$

and

$$\int f(\xi_1, \ldots, \xi_p) \mu_1(d\xi_1) \ldots \mu_1(d\xi_p) = 0. \tag{4.7}$$

The following result is new. Here, we take again the probability measure ν_∞ on G^∞ of the form (4.2) with $\mu_k = \mu_1$ for all k.

THEOREM 4.1. *Let* $\{\varepsilon_n\}$ *be any sequence of positive numbers such that*

$$\sum_{n=0}^{\infty} e^{-\varepsilon_n / t_n} < \infty, \tag{4.8}$$

where

$$t_n = \sum_{k=0}^{\infty} a_{nk}^2. \tag{4.9}$$

Then we have for almost all $[\nu_\infty]$ *points* $\theta \in G^\infty$ *that, for* n *sufficiently large,*

$$|g_n(\theta)| \leqq \sqrt{2p\varepsilon_n}. \tag{4.10}$$

The case $p = 1$ of Theorem 4.1 sharpens a result of Hlawka [10] p. 233. If (a_{nk}) is as in (4.3), that is $a_{nk} = n^{-1}$ if $1 \leqq k \leqq n$ and zero otherwise, then $t_n = 1/n$ and (4.8) holds with

$$\sqrt{\varepsilon_n} = \sqrt{\frac{2}{n} \log n}\,.$$

The resulting assertion (4.10) is slightly weaker than the optimal result (4.3).

For the proof of Theorem 4.1 (and also in section 6) we shall need the following auxiliary result.

LEMMA 4.2. *Let* $\{Z_k, k = 0, 1, \ldots\}$ *be a sequence of independent real-valued random variables such that* $|Z_k| \leqq 1$ *for all* k. *Let further* $\{c_k\}$ *be a sequence of real constants such that*

$$t = s^2 = \sum_{k=0}^{\infty} c_k^2 < \infty, \qquad (s \geqq 0). \tag{4.11}$$

Then

$$S = \sum_{k=0}^{\infty} c_k(Z_k - m_k), \qquad (m_k = E(Z_k)), \tag{4.12}$$

satisfies

$$P(S > \delta s) \leqq e^{-\delta^2/2}, \quad P(S < -\delta s) \leqq e^{-\delta^2/2}, \tag{4.13}$$

for each number $\delta > 0$.

As is the case for most results in the theory of probability, Lemma 4.2 holds with respect to any underlying probability space $(\Omega, \mathscr{B}, P)$, (in particular with respect to (G^∞, ν_∞)).

By $EZ = E(Z)$ we mean (here and in the future) the integral

$$E(Z) = \int_\Omega Z(\theta)P(d\theta),$$

whenever the right hand integral is (absolutely) convergent. Further, it is known ([15] p. 236) that (4.12) converges for almost $[P]$ all $\theta \in \Omega$ whenever the Z_k are independent and satisfy

$$\sum c_k^2\, E(Z_k - E(Z_k))^2 < \infty.$$

The latter is implied by (4.11) and $|Z_k| \leqq 1$. Finally,

$$P(S > \delta s) = P\{\theta \colon S(\theta) > \delta s\}.$$

Proof of lemma 4.2. Put

$$\varphi_k(u) = \log E(e^{uZ_k}). \qquad (u \text{ real}).$$

Then $\varphi_k(0) = 0$, $\varphi_k'(0) = m_k$ and

$$\varphi_k''(u) \leqq E(Z_k^2 e^{uZ_k})/E(e^{uZ_k}) \leqq 1,$$

by $|Z_k| \leqq 1$. Therefore,

$$\varphi_k(u) \leqq m_k u + u^2/2.$$

Hence, letting S' denote the truncation $\sum_{k=0}^{M}$ of (4.12),

$$\begin{aligned} E(e^{uS'}) &= \prod_{k=0}^{M} E\exp(c_k u(Z_k - m_k)) \\ &= \exp\sum_{k=0}^{M} (\varphi_k(c_k u) - m_k c_k u) \leqq \exp\sum_{k=0}^{\infty} (c_k u)^2/2 = e^{s^2u^2/2}; \end{aligned}$$

(in the first step we used the assumed independence of the Z_k). On the other hand,

$$\begin{aligned} E(e^{uS'}) &\geqq P(S' > \delta s)e^{\delta s u} \quad \text{if } u > 0, \\ &\geqq P(S' < -\delta s)e^{-\delta s u} \text{ if } u < 0. \end{aligned}$$

Taking $u = \pm\delta/s$, one obtains (4.13), (first with S replaced by its truncation $S'(\theta)$, afterwards, using Egorov, say, for $S(\theta)$ itself).

COROLLARY 4.3. Let $Z_0, Z_1, \ldots$ be a sequence of real-valued random variables, such that there exists a partition of $\{0, 1, \ldots\}$ into p disjoint sets D_j with the property that for each $j = 1, \ldots, p$ the random variables $\{Z_k, k \in D_j\}$ are independent. Then $|Z_k| \leqq 1$ $(k = 0, 1, \ldots)$, (4.11) and (4.12) together imply that, for all $\eta > 0$,

$$(4.14) \qquad P(|S| > \eta) \leqq 2p\exp(-\eta^2/(2pt)).$$

Proof. Let $s_j = \sqrt{t_j}$ and S_j correspond to the subsequence $\{Z_k, k \in D_j\}$. Then, by (4.13), one has for all points θ outside a set of P-measure $\leqq p(2e^{-\delta^2/2})$ that

$$|S_j(\theta)| \leqq \delta s_j \; (j = 1, \ldots, p),$$

hence,

$$|S(\theta)| \leqq \delta\sum_{j=1}^{p}\sqrt{t_j} \leqq \delta\sqrt{pt} = \eta, \qquad (\text{say}).$$

Proof of theorem 4.1. Apply the above corollary with (G^∞, ν_∞) as the underlying probability space. Then $\{x_k(\theta) = \theta_k\}$ is a sequence of independent random variables taking values in G. Let further

$$Z_k = Z_k(\theta) = f(\theta_k, \theta_{k+1}, \ldots, \theta_{k+p-1}).$$

Clearly, $\{Z_{hp+j-1}, h = 0, 1, \ldots\}$ is a sequence of independent random variables, for each fixed $j = 1, \ldots, p$. Further, by (4.6) and (4.7), we have $|Z_k| \leqq 1$ and $E(Z_k) = 0$. It follows by (4.5), (4.9) and (4.14) that

$$\nu_\infty\{\theta: |g_n(\theta)| > \sqrt{2p\varepsilon_n}\} \leqq 2pe^{-\varepsilon_n/t_n}.$$

Thus, (4.8) implies the desired result (4.10).

5. The complete asymptotic distribution of a sequence

In this and the following sections,

$$A = (a_{nk}, n, k = 0, 1, \ldots)$$

will denote a *fixed* real matrix such that

(5.1) $$a_{nk} \geqq 0, \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} = 1$$

and

(5.2) $$\lim_{n\to\infty} \sum_{k=-1}^{\infty} |a_{n,k} - a_{n,k+1}| = 0, \qquad (a_{n,-1} = 0).$$

Note that (5.2) implies that

(5.3) $$\lim_{n\to\infty} a_{nk} = 0 \textit{ uniformly in } k.$$

Thus, A is a regular summation matrix such that $u_k \geqq 0$ and A-lim $u_k = u$ imply $u \geqq 0$. Moreover, if $\{u_k\}$ is bounded then

$$A\text{-}\lim(u_{k+1} - u_k) = 0.$$

Let further G denote a fixed second countable compact Hausdorff space, say, the reals modulo 1. To each bounded linear functional $\mu(f)$ on the Banach space $C(G)$, (that is, $\mu(f)$ is linear and complex-valued such that $|\mu(f)| \leqq c\|f\|$ for some constant c), there corresponds a *unique* regular Borel measure μ on G such that

$$\mu(f) = \int_G f(x)\mu(dx) \text{ for each } f \in C(G).$$

Conversely, this formula associates to each such measure μ a bounded linear functional $\mu(f)$ on $C(G)$. In view of this, a bounded linear functional $\mu(f)$ on $C(G)$ will also be called a measure on G. This measure is real and nonnegative if $\mu(f) \geqq 0$ for each (real and) nonnegative $f \in C(G)$. If, moreover, $\mu(1) = 1$ then $\mu(f)$ is called a probability measure on G.

Let

$$x_{n,k} \in G, \qquad n, k = 0, 1, \ldots, \tag{5.4}$$

be a given double sequence $\{x_{nk}\}$ of points in G; (in many applica-rions, $x_{nk} = x_k$ is independent of n in which case $\{x_{nk} = x_k\}$ is called a simple sequence). A measure μ_1 on G will be called a limiting measure of this double sequence if, for some sequence $0 \leqq n_0 < n_1 < n_2 < \ldots$ of integers n_j, one has

$$\lim_{\substack{n=n_j \\ j\to\infty}} \sum_{k=0}^{\infty} a_{nk} f(x_{nk}) = \int_G f(x)\mu_1(dx) = \mu_1(f) \tag{5.5}$$

for all $f \in C(G)$. Note that, by (5.1), μ_1 must be a probability measure on G. By the remarks following (4.4) (and a diagonal procedure) there exists at least one limiting measure.

The collection of all limiting measures of $\{x_{nk}\}$ will be denoted as

$$V_1 = V_1\{x_{nk}\}.$$

If it consists of a single measure μ_1 then μ_1 is called the asymptotic *A-distribution* of the double sequence $\{x_{nk}\}$.

We shall also be interested in the joint distribution of successive elements $x_{n,k}, x_{n,k+1}, \ldots, x_{n,k+r-1}$, that is, in the asymptotic A-distribution of the sequence of points

$$x_{nk}^{(r)} = (x_{n,k}, x_{n,k+1}, \ldots, x_{n,k+r-1}) \tag{5.6}$$

in the r-fold direct product

$$G^r = G_0 \times \ldots \times G_{r-1}.$$

(Here, G_i denotes a copy of G.) Thus, consider the measure

$$\mu_{r,n}(f) = \sum_{k=0}^{\infty} a_{nk} f(x_{nk}^{(r)}), \qquad f \in C(G^r), \tag{5.7}$$

on G^r having a mass $a_{nk} \geqq 0$ at the point $x_{nk}^{(r)}$, and let $V_r = V_r\{x_{nk}\} = V_1\{x_{nk}^{(r)}\}$ denote the non-empty collection of limit points of the sequence of measures $\{\mu_{r,n}, n = 0, 1, \ldots\}$. That is, a probability measure μ_r on G^r belongs to $V_r\{x_{nk}\}$ if and only if there exists an increasing sequence of positive integers n_j such that

$$\lim_{j\to\infty}\sum_{k=0}^{\infty} a_{n_j,k}\, f(x_{n_j,k},\ldots,x_{n_j,k+r-1}) = \int_{G^r} f\,d\mu_r,$$

for each function $f = f(\theta_0,\ldots,\theta_{r-1})$ in $C(G^r)$. If $V_r\{x_{nk}\}$ consists of a single probability measure μ_r on G^r, we call μ_r the *r-dimensional A-distribution* of the double sequence $\{x_{nk}\}$.

Finally, consider the infinite product space

(5.8) $$G^\infty = G_0\times G_1\times G_2\times\ldots$$

and the measures

(5.9) $$\mu_{\infty,n}(f) = \sum_{k=0}^{\infty} a_{nk}\, f(x_{n,k}^{(\infty)}), \qquad f\in C(G^\infty),$$

on G^∞. Here,

(5.10) $$x_{n,k}^{(\infty)} = (x_{n,k}, x_{n,k+1},\ldots,x_{n,k+r-1}, x_{n,k+r},\ldots)$$

denotes the point in G^∞ whose r-th coordinate is equal to $x_{n,k+r}$.

Let again

$$V_\infty = V_\infty\{x_{nk}\} = V_1\{x_{n,k}^{(\infty)}\}$$

denote the non-empty collection of limit points of the sequence $\{\mu_{\infty,n}\}$. If it consists of a single probability measure μ_∞ on G^∞ we call μ_∞ the *complete A-distribution* of the double sequence $\{x_{nk}\}$ of points x_{nk} in G.

In any case, if $1 \leqq s < r \leqq \infty$ then V_s is precisely the set of projections (marginals) of the measures $\mu_r\in V_r$ on G^r onto the component G^s of $G^r = G^s\times G_s\times G_{s+1}\times\ldots\times G_{r-1}$.

By T we shall denote the *shift transformation*

$$T(\theta_0,\theta_1,\theta_2,\ldots) = (\theta_1,\theta_2,\ldots)$$

in G^∞. In other words, if $\theta\in G^\infty$ and $T\theta = \theta'$ then the r-th coordinate θ'_r of θ' is equal to the $(r+1)$-th coordinate θ_{r+1} of θ, $(r = 0, 1, 2, \ldots)$.

By (5.10), we can write (5.9) as

(5.11) $$\mu_{\infty,n}(f) = \sum_{k=0}^{\infty} a_{nk} f(T^k x_n^{(\infty)}), \qquad f\in C(G^\infty),$$

where

$$x_n^{(\infty)} = (x_{n0}, x_{n1}, x_{n2},\ldots)\in G^\infty.$$

Similarly, (5.7) can be written as

(5.12) $$\mu_{r,n}(f) = \sum_{k=0}^{\infty} a_{nk} f(T^k x_n^{(\infty)}) \text{ with } f\in C(G^r),$$

provided we identify the function $f = f(\theta_0, \ldots, \theta_{r-1})$ on G^r with the function $f(\theta) = f(\theta_0, \ldots, \theta_{r-1})$ on G^∞ which is independent of the coordinates $\theta_r, \theta_{r+1}, \ldots$ of θ.

For $f = f(\theta) = f(\theta_0, \theta_1, \ldots)$ as any function on G^∞, let us define

$$(Tf)(\theta) = f(T\theta) = f(\theta_1, \theta_2, \ldots).$$

Because T is an open and continuous map of G^∞ onto G^∞, we have $Tf \in C(G^\infty)$ if and only if $f \in C(G^\infty)$. Further, $Tf \geqq 0$ if and only if $f \geqq 0$. Finally, by (5.2) and (5.11),

$$\lim_{n\to\infty} [\mu_{\infty,n}(Tf) - \mu_{\infty,n}(f)] = 0 \qquad \text{for each } f \in C(G^\infty). \tag{5.13}$$

It follows that each limit point μ_∞ of $\{\mu_{\infty,n}\}$, that is, each probability measure $\mu_\infty \in V_\infty\{x_{nk}\}$ satisfies

$$\mu_\infty(Tf) = \mu_\infty(f) \tag{5.14}$$

for all $f \in C(G^\infty)$; interpreting $\mu_{\infty,n}$ and μ_∞ as an integral, (5.13) and (5.14) even hold for any bounded and Borel measurable function $f(\theta)$ on G^∞.

A measure μ_∞ on G^∞ satisfying (5.14) will be called an *invariant measure*. The collection of all invariant probability measures on G^∞ will be denoted as I_∞. We thus have proved that (5.1) and (5.2) imply

$$V_\infty\{x_{n,k}\} \subset I_\infty. \tag{5.15}$$

This result has important consequences.

But let us first take up the question whether to each $\nu_\infty \in I_\infty$ there corresponds a double sequence $\{x_{nk}\}$ such that

$$\nu_\infty \in V_\infty\{x_{nk}\}. \tag{5.16}$$

At this point, if desired, the reader could also turn to section 7.

6. Sequences having a preassigned complete distribution

In this section, we shall only require that $A = (a_{nk})$ is a *real* matrix satisfying

$$\sum_{k=0}^{\infty} |a_{nk}| \leqq M, \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} = 1 \tag{6.1}$$

and (5.2), ($M \geqq 1$ denoting a fixed constant); we shall also write $a_{n,k} = a_n(k)$.

Note that, by (5.3) and (6.1), $t_n \to 0$ where

$$t_n = \sum_{k=0}^{\infty} a_{nk}^2, \qquad (n = 0, 1, \ldots). \tag{6.2}$$

Hence, there exists a sequence $0 \leqq n_0 < n_1 < n_2 < \ldots$ of integers such that

$$\sum_{j=0}^{\infty} \exp(-\varepsilon/t_{n_j}) < \infty \quad \text{for each} \quad \varepsilon > 0.$$

Therefore, the following result implies that the question (5.16) has always a positive answer.

THEOREM 6.1. *Suppose that*

$$\sum_{n=0}^{\infty} e^{-\varepsilon/t_n} < \infty \quad \textit{for each} \quad \varepsilon > 0. \tag{6.3}$$

Let $\nu_\infty \in I_\infty$ be arbitrarily given. Then there exists a sequence $\{\theta_k\}$ of points in G such that the simple sequence $\{x_{nk} = \theta_k\}$ has a (unique) complete A-distribution equal to ν_∞.

If A and A' are two real and regular summation matrices, they are said to be consistent for bounded sequences if a bounded sequence $\{u_n\}$ is A-summable to u whenever it is A'-summable to u and conversely. Clearly, Theorem 6.1 applied to two such matrices yields two equivalent conclusions. For instance, the ordinary Cesaro summation method $A = (C, 1)$ is known to be consistent for bounded sequences with many other summation methods, compare [2].

Assume that (6.3) holds. We may also assume that

$$a_{nk} = 0 \text{ for } k > \lambda(n), \text{ where } \lambda(n) \leqq \lambda(n+1), \; \lambda(n) \to \infty. \tag{6.4}$$

For, if (6.4) does not hold, choose $\lambda(n)$ such that

$$\sum_{k=\lambda(n)}^{\infty} |a_{nk}| \leqq 2^{-n}, \qquad \lambda(n) \geqq \lambda(n-1),$$

$(n = 0, 1, \ldots)$. Then the summation method

$$\begin{aligned} a'_{nk} &= a_{nk} \text{ if } k \leqq \lambda(n), \\ &= 0 \text{ if } k > \lambda(n), \end{aligned}$$

is consistent with A for bounded sequences. Moreover, (5.2) and (6.3) imply the corresponding relations for (a'_{nk}).

Now, Let M_r $(r = 1, 2, \ldots)$ be a given sequence of positive integers (to be chosen in a suitable manner), and put

$$N_r = M_1 + 2M_2 + \ldots + rM_r, \qquad (N_0 = 1), \tag{6.5}$$

thus, $N_r \to \infty$. Let further

(6.6) $I_r(h) = \{k: k = N_{r-1}+(h-1)r+j-1;\, j = 1, \ldots, r\}$,

where $r = 1, 2, \ldots$ and $h = 1, \ldots, M_r$. By (6.5), the $I_r(h)$ define a *partition* of the nonnegative integers into disjoint intervals.

LEMMA 6.2. *For $p = 1, 2, \ldots$, let J_p denote the set of nonnegative integers k such that the p integers $k, k+1, \ldots, k+p-1$ all belong to one and the same interval $I_r(h)$. Then, for each $p = 1, 2, \ldots$,*

$$\lim_{n\to\infty} \sum_{\substack{k=0\\ k\notin J_p}}^{\infty} |a_{nk}| = 0. \tag{6.7}$$

Proof. If $p = 1$ then J_p contains all nonnegative integers. Thus, let $p \geqq 2$ be fixed. By (6.6),

$$\sum_{\substack{k=0\\ k\notin J_p}}^{\infty} |a_{nk}| = \sum_{k=0}^{N_{p-1}-1} |a_{nk}| + \sum_{r=p}^{\infty} \sum_{h=1}^{M_r} \sum_{q=1}^{p-1} |a_n(N_{r-1}+hr-q)|.$$

By (5.3), it suffices to show that $\lim_{n\to\infty} d_{nq} = 0$ for each fixed $q = 1, 2, \ldots$, where

$$d_{nq} = \sum_{r=q+1}^{\infty} \sum_{h=1}^{M_r} |a_n(N_{r-1}+hr-q)|.$$

From (5.2), (5.3), (6.1) and (6.5), one easily sees that

$$\lim_{n\to\infty} (d_{n,q}-d_{n,q+1}) = 0$$

for each $q \geqq 1$, while

$$\sum_{q=1}^{\infty} d_{nq} \leqq \sum_{k=0}^{\infty} |a_{nk}| \leqq M.$$

Thus, for q and Q as fixed positive integers,

$$0 \leqq Qd_{nq} \leqq \sum_{s=1}^{Q} d_{ns}+o(1) \leqq M+o(1),$$

showing that $d_{nq} \to 0$ as $n \to \infty$.

LEMMA 6.3. *The sequence $\{M_r\}$ of positive integers can be chosen in such a way that, for some sequence $\{\varepsilon_n\}$ of positive numbers converging to zero, one has*

$$\sum_{n=0}^{\infty} e^{-\varepsilon_n/u_n} < \infty. \tag{6.8}$$

Here,

$$u_n = \sum_{r=1}^{\infty} r \sum_{k=N_{r-1}}^{N_r-1} a_{nk}^2. \tag{6.9}$$

Proof. By (6.3), there exist integers $0 = h_1 < h_2 < \ldots$ such that

$$\sum_{n=h_j}^{\infty} e^{-j^{-2}/t_n} < 2^{-j} \qquad \text{if } j \geqq 2. \tag{6.10}$$

Define $\varepsilon_n = j^{-1}$ for $h_j \leqq n < h_{j+1}$, thus, $\varepsilon_n \to 0$. Further, given the M_s with $1 \leqq s < r$, choose M_r so large that the N_r defined by (6.5) satisfies

$$N_r > \lambda(h_{r+1}), \qquad (r = 1, 2, \ldots). \tag{6.11}$$

Clearly, (6.10) implies (6.8) provided one has, for each fixed $j \geqq 1$, that $h_j \leqq n < h_{j+1}$ implies $u_n \leqq jt_n$; (for, then $\varepsilon_n/u_n \geqq j^{-2}/t_n$ by the definition of ε_n). Hence, by (6.2), (6.4) and (6.9), it suffices to show that

$$n < h_{j+1}, r \geqq j+1, k \geqq N_{r-1}$$

together imply $k > \lambda(n)$, (thus, $a_{nk} = 0$). Indeed, in this case one has

$$k \geqq N_{r-1} \geqq N_j > \lambda(h_{j+1}) \geqq \lambda(n),$$

by (6.11) and $\lambda(m) \leqq \lambda(m+1)$.

Remark. The above proof shows that $\{\varepsilon_n\}$ can be chosen as any nullsequence such that $\sum \exp(-\varepsilon_n/(\varphi_n t_n)) < \infty$ for *some* sequence $\{\varphi_n\}$ tending to infinity; (for the ordinary $(C, 1)$-summation method this means $\varepsilon_n \to 0$, $n\varepsilon_n/\log n \to \infty$). For, choosing

$$N_r > \max_{[\varphi_n] \leqq r} \lambda(n),$$

one has $u_n \leqq \varphi_n t_n$; (if $r > \varphi(n)$ and $k \geqq N_{r-1}$ then $k > \lambda(n)$ thus $a_{nk} = 0$).

Proof of theorem 6.1. Let $\nu_\infty \in I_\infty$ be an arbitrary but fixed invariant probability measure on

$$G^\infty = G_0 \times G_1 \times G_2 \times \ldots, \tag{6.12}$$

G_k denoting a copy of G. Let ν_r denote its projection onto the component $G^r = G_0 \times G_1 \times \ldots \times G_{r-1}$ of G^∞. Because ν_∞ is invariant under the translation T, we have, for each $1 \leqq p < r \leqq \infty$ and $0 \leqq j \leqq r-p$, that the projection of the measure ν_r on G^r onto the component $G_j \times G_{j+1} \times \ldots \times G_{j+p-1}$ of G^r is precisely equal to ν_p in the sense that

$$(6.13) \qquad \nu_r(f(\theta_j, \theta_{j+1}, \ldots, \theta_{j+p-1})) = \nu_p(f(\theta_0, \ldots, \theta_{p-1})),$$

for every bounded and measurable function $f(\theta_0, \ldots, \theta_{p-1})$ on G^p.

Now, choose $\{M_r\}$ as any fixed sequence of positive integers which has the property mentioned in Lemma 6.3. For $r = 1, 2, \ldots$ and $h = 1, \ldots, M_r$, let us consider the r-fold direct product

$$(6.14) \qquad G_h^r = \prod_{k \in I_r(h)} G_k = G_{N_{r-1}+(h-1)r} \times \ldots \times G_{N_{r-1}+hr-1}.$$

In view of (6.5), the space (6.12) may be regarded as the direct product

$$(6.15) \quad G^\infty = G_1^1 \times \ldots \times G_{M_1}^1 \times \ldots \times G_1^r \times \ldots \times G_{M_r}^r \times \ldots$$

We now define a probability measure σ_∞ on G^∞ obtained by assigning to each component G_h^r the measure ν_r; afterwards, we take the direct product of these measures. In other words, the measure space $(G^\infty, \sigma_\infty)$ is defined as the direct product

$$(6.16) \qquad (G^\infty, \sigma_\infty) = \prod_{r=1}^{\infty} \prod_{h=1}^{M_r} (G_h^r, \nu_r)$$

of the measure spaces (G_h^r, ν_r).

We assert that, *for almost all* $[\sigma_\infty]$ points $\theta = (\theta_0, \theta_1, \theta_2, \ldots)$ in G^∞, the corresponding sequence $\{\theta_k\}$ of points in G has a unique complete A-distribution which is equal to the given ν_∞. Thus, there is at least one such sequence, proving Theorem 6.1.

We must show that, for almost all $[\sigma_\infty]$ points $\theta \in G^\infty$,

$$(6.17) \qquad \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} f(\theta_k, \theta_{k+1}, \ldots) = \nu_\infty(f)$$

holds for each $f \in C(G^\infty)$. In fact, we may restrict f to a denumerable collection $\{f_i\}$ spanning a dense subset of $C(G^\infty)$; (such a collection exists because with G also G^∞ is second countable). The f_i can moreover be chosen as functions depending only on a finite number of coordinates θ_k. After all, from the definition of the product topology, we have for each continuous function $f = f(\theta_0, \theta_1, \ldots)$ on G^∞ that

$$\lim_{k\to\infty} f(\theta_0, \theta_1, \ldots, \theta_k, x_0, x_0, x_0, \ldots) = f(\theta)$$

holds *uniformly* in θ; ($x_0 \in G$ fixed).

Thus, let p be a fixed positive integer and let $f(\theta_0, \ldots, \theta_{p-1})$ be an arbitrary but fixed bounded and measurable real-valued function on the p-fold direct product $G \times \ldots \times G$. It suffices to prove that, for almost all $[\sigma_\infty]$ points $\theta = (\theta_0, \theta_1, \ldots)$, one has

(6.18) $$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} f(\theta_k, \theta_{k+1}, \ldots, \theta_{k+p-1}) = \nu_p(f);$$

(compare (6.13)). By (6.1), this certainly holds when f is a constant function, thus, without loss of generality we may assume that

(6.19) $$\nu_p(f) = 0, \qquad |f(\theta_0, \ldots, \theta_{p-1})| \leqq 1.$$

By Lemma 6.2, the summation in (6.18) may be restricted to $k \in J_p$. Thus, it remains to show that, for almost $[\sigma_\infty]$ all θ, one has

(6.20) $$\lim_{n\to\infty} S_n(\theta) = 0.$$

Here,

(6.21) $$S_n(\theta) = \sum_{r=p}^{\infty} \sum_{h=1}^{M_r} Z_{r,h}^{(n)}(\theta),$$

where,

(6.22) $$Z_{r,h}^{(n)}(\theta) = \sum_{q=0}^{r-p} a_n(N_{r,h}+q) f(\theta_{N_{r,h}+q}, \ldots, \theta_{N_{r,h}+q+p-1})$$

with $N_{r,h} = N_{r-1}+(h-1)r$; (compare (6.5), (6.15) and the definition of J_p). Now notice that $Z_{r,h}^{(n)}(\theta)$ depends only on the coordinates θ_k with $k \in I_r(h)$. Recall that the intervals $I_r(h)$ are disjoint. Hence, by (6.14) and the definition (6.16) of σ_∞, the measurable functions $Z_{r,h}^{(n)}$ $(r = p, p+1, \ldots; h = 1, \ldots, M_r)$ are *independent* real-valued random variables *relative* to the probability space $(G^\infty, \sigma_\infty)$, (compare (3.1)). Moreover, by (6.13), (6.16), (6.19) and (6.22),

$$E(Z_{r,h}^{(n)}) = \int_{G^\infty} Z_{r,h}^{(n)}(\theta)\sigma_\infty(d\theta) = \nu_p(f) = 0,$$

for all $r \geqq p$, $1 \leqq h \leqq M_r$. Finally, by (6.19), (6.22) and Cauchy's inequality,

$$|Z_{r,h}^{(n)}|^2 \leqq r \sum_{q=0}^{r-1} a_n(N_{r-1}+(h-1)r+q)^2 = c_{r,h}^{(n)}, \qquad \text{say.}$$

By (6.5) and (6.9),

$$\sum_{r=p}^{\infty} \sum_{h=1}^{M_r} c_{r,h}^{(n)} \leqq u_n.$$

It follows by (6.21) and Lemma 4.2 (with $\{c_k Z_k\}$ replaced by $\{Z_{r,h}^{(n)}\}$) that

$$\sigma_\infty\{\theta\colon |S_n(\theta)| > \delta\sqrt{u_n}\} \leqq 2e^{-\delta^2/2},$$

for all $\delta > 0$. Choosing $\delta = \sqrt{2\varepsilon_n/u_n}$, we obtain from (6.8) that

$$\sum_{n=0}^{\infty} \sigma_\infty\{\theta\colon |S_n(\theta)| > \sqrt{2\varepsilon_n}\} < \infty.$$

This in turn implies that, for almost all $[\sigma_\infty]$ points $\theta \in G^\infty$, there exists an integer $n_0(\theta)$ such that

$$|S_n(\theta)| \leqq \sqrt{2\varepsilon_n} \text{ for } n \geqq n_0(\theta).$$

But $\{\varepsilon_n\}$ converges to zero, thus, (6.20) holds for almost all $[\sigma_\infty]$ points θ.

REMARK. At least for $A = (C, 1)$, the special case of Theorem 6.1, where ν_∞ is ergodic with respect to T, is also an immediate consequence of the individual ergodic theorem.

On the other hand, suppose that the invariant probability measure ν_∞ is not ergodic with respect to T. Then there exists a measurable subset W of G^∞ such that $0 < \nu_\infty(W) < 1$ while $\theta \in W$ if and only if $T\theta \in W$. If $f(\theta) = 1$ or 0, depending on whether θ does or does not belong to W, then the left hand side of (6.17) is always equal to $f(\theta)$, thus, (6.17) holds for no $\theta \in G^\infty$ whatsoever.

What we have shown is that (6.17) is true for many points θ, more precisely, for almost all $[\sigma_\infty]$ points $\theta \in G^\infty$, whenever $f(\theta)$ is a fixed bounded and measurable function on G^∞ which *depends on only finitely many coordinates* θ_k, thus, also when $f(\theta)$ is the uniform limit of a sequence of such functions, in particular, if $f \in C(G^\infty)$.

7. A general moment problem

Here, the assumptions and notations are those of section 5. By $L(G)$ we shall denote the *real* linear vector space consisting of all the real-valued and continuous functions f on G. We further put

$$L(G^\infty) = L^\infty.$$

Be given any subset

$$\{f_i, i \in D_0\} \tag{7.1}$$

of L^∞, (no restrictions on the cardinality of the index set D_0). Let further

$$\{\rho_i, i \in D_0\}$$

be any real-valued function on D_0. Finally, let $\{x_{nk}\}$ be a given double sequence of points in G satisfying

(7.2) $$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} f_i(x_{n,k}, x_{n,k+1}, \ldots) = \rho_i \text{ for each } i \in D_0.$$

In view of (5.1), we may and will assume that for some distinguished element $0 \in D_0$, say, we have

$$f_0(\theta) \equiv 1, \ \rho_0 = 1.$$

In many applications, the condition (7.2) will arise as follows. Let Q be a given index set. For each $q \in Q$, let Γ_q be a given compact metric space and $g_q = g_q(\theta)$ a given *continuous* mapping of G^∞ into Γ_q; (for example, if G is an additively written compact group, let $\Gamma_q = G$ and $g_q(\theta) = \theta_q - \theta_0$, $q = 1, 2, \ldots$). Further, put

$$y_{nk}^{(q)} = g_q(T^k x_n^{(\infty)}) = g_q(x_{n,k}, x_{n,k+1}, \ldots)$$

and *assume* that, for each $q \in Q$, the double sequence $\{y_{nk}^{(q)}\}$ of points in Γ_q has a *given* probability measure ν_q an Γ_q as its (one-dimensional) A-distribution, that is,

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} f_q(g_q(T^k x_n^{(\infty)}) = \nu_q(f_q),$$

for each $q \in Q$ and each $f_q \in L(\Gamma_q)$.

Returning to the general condition (7.2), let $g \in L^\infty$ be a further real and continuous function on G^∞. What then can be said about the set of accumulation points of the sequence

$$\{\mu_{\infty,n}(g) = \sum_{k=0}^{\infty} a_{nk} g(x_{n,k}, x_{n,k+1}, \ldots)\},$$

$(n = 0, 1, \ldots)$? More generally, if

$$\{g_j, j \in D_1\} \subset L^\infty$$

is a given subset of $L(G^\infty)$, find all the real-valued functions

$$\{\sigma_j, j \in D_1\}$$

on D_1 such that there exists a double sequence $\{x_{nk}\}$ of points in G satisfying (7.2) and, further, for some $\mu_\infty \in V_\infty\{x_{nk}\}$,

(7.3) $$\mu_\infty(g_j) = \sigma_j \text{ for all } j \in D_1.$$

Note that a double sequence $\{x_{nk}\}$ satisfies (7.2) if and only if

(7.4) $$\mu_\infty(f_i) = \rho_i \text{ for all } i \in D_0,$$

and all $\mu_\infty \in V_\infty\{x_{nk}\}$.

By (5.16), each $\mu_\infty \in V_\infty\{x_{nk}\}$ belongs to the collection I_∞ of invariant probability measures on G^∞. Thus, if $\{x_{nk}\}$ satisfies (7.2) then, for a given function σ_j on D_1, (7.3) can happen for some $\mu_\infty \in V_\infty\{x_{nk}\}$ *only if* there exists at least one $\mu_\infty \in I_\infty$ satisfying both (7.3) and (7.4). Note that this (necessary) condition is completely *independent of the original summation method* $A = (a_{nk})$, (which is assumed to satisfy (5.1) and (5.2)).

In most applications, for instance, if $A = (C, 1)$, one has

$$(7.5) \qquad \sum_{n=0}^{\infty} \exp(-\varepsilon/\sum_{k=0}^{\infty} a_{nk}^2) < \infty \text{ for each } \varepsilon > 0.$$

In this case, the above condition is *also sufficient.* For, suppose that the real function σ_j on D_1 is such that there exists a $\mu_\infty \in I_\infty$ satisfying both (7.3) and (7.4); let this μ_∞ be fixed. By (7.5) and Theorem 6.1, there exists a *simple* sequence $\{x_{nk} = x_k\}$ of points x_k in G such that $V_\infty\{x_k\}$ is the one-element set $\{\mu_\infty\}$. In particular, by (7.3) and (7.4),

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} f_i(x_k, x_{k+1}, \ldots) = \rho_i \text{ for all } i \in D_0,$$

and

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} g_j(x_k, x_{k+1}, \ldots) = \sigma_j \text{ for all } j \in D_1.$$

In this way, we arrive at the following:

Fundamental moment problem. Given $\{f_i \in L^\infty, i \in D_0\}$, $\{g_j \in L^\infty, j \in D_1\}$ and the real numbers ρ_i, $i \in D_0$, what are the necessary and sufficient conditions on the set of real numbers $\{\sigma_j, j \in D_1\}$ in order that there exists at least one $\mu_\infty \in I_\infty$ satisfying both (7.3) and (7.4)?

A complete (though not always useful) answer to this problem is given by the following Theorem 7.1, whose proof is based on the Hahn-Banach theorem.

Let K denote the cone in L^∞ consisting of all $f \in L^\infty$ satisfying $f(\theta) \geqq 0$ for all $\theta \in G^\infty$. Let further K' denote the cone consisting of all $f \in L^\infty$ which in at least one way can be written as

$$f = f_1 + h - Th \quad \text{with} \quad f_1 \in K \text{ and } h \in L^\infty.$$

We shall write

$$(7.6) \qquad f \prec g \text{ if and only if } g - f \in K'.$$

This defines a partial ordering in L^∞, ($f \prec f$ while $f \prec g$, $g \prec h$

imply $f \prec h$), which is invariant under addition and under multiplication by a nonnegative real number; (one can show that $f \prec 0 \prec f$ if and only if f is of the form $f = h - Th$).

Next, for each $g \in L^\infty$, put

$$Q(g) = \inf\{\sum_{i \in D_0} \alpha_i \rho_i \colon g \prec \sum_{i \in D_0} \alpha_i f_i\}; \tag{7.7}$$

(each sum finite, that is, all but finitely many α_i equal to zero). Note that, by $f_0(\theta) = 1$, $\rho_0 = 1$,

$$-\infty \leqq Q(g) \leqq \sup_{\theta \in G^\infty} g(\theta) < \infty.$$

THEOREM 7.1. *A necessary and sufficient condition for the existence of a $\mu_\infty \in I_\infty$ satisfying (7.4) is that $Q(0) = 0$.*

A necessary and sufficient condition for the existence of a $\mu_\infty \in I_\infty$ satisfying both (7.3) and (7.4) is that

$$\sum_{j \in D_1} \beta_j \sigma_j \leqq Q(\sum_{j \in D_1} \beta_j g_j) \tag{7.8}$$

for each choice of the real numbers β_j, $j \in D_1$, all but finitely many β_j equal to zero.

In particular, given $g \in L^\infty$ and σ real, there exists a $\mu_\infty \in I_\infty$ with $\mu_\infty(g) = \sigma$ if and only if

$$p(g) \leqq \sigma \leqq q(g), \tag{7.9}$$

where

$$p(g) = \sup\{\alpha \colon \alpha \prec g\}, \quad q(g) = \inf\{\beta \colon g \prec \beta\},$$

(α and β ranging over the constant functions); $\sigma = q(g)$ is even attained by an ergodic $\mu_\infty \in I_\infty$.

The proof of Theorem 7.1 and certain related results will be given elsewhere. In applying Theorem 7.1, one of the main difficulties lies in the difficulty of computing the quantity $Q(g)$, more precisely, in the question whether a *given* function $g \in L^\infty$ satisfies $p(g) \geqq 0$. The easiest case is that where $g(\theta)$ depends on only finitely many coordinates $\theta_0, \theta_1, \ldots, \theta_{r-1}$. For, then $p(g) \geqq 0$ can be shown to hold if and only if

$$\sum_{k=1}^{n} g(\theta_k, \theta_{k+1}, \ldots, \theta_{k+r-1}) \geqq 0$$

whenever $\theta_{n+k} = \theta_k$ for $k = 1, \ldots, r-1$.

As an illustration, take $G = \{0, 1, 2, 3\}$ as the group of integers modulo 4. Using the latter criterion, one finds that there exists a $\mu_\infty \in I_\infty$ satisfying

$$\mu_\infty(\cos\frac{\pi}{2}(\theta_1-\theta_0)) = 0,\ \mu_\infty(\cos\frac{\pi}{2}\theta_0) = \sigma',\ \mu_\infty(\sin\frac{\pi}{2}\theta_0) = \sigma'',$$

if and only if $|\sigma''| \leqq \frac{1}{2}$ and $|\sigma'| + |\sigma''| \leqq 1$.

As to (7.9), one usually has $p(g) < q(g)$. It can be shown that $p(g) = q(g)$ implies that

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} g(T^k\theta) = p(g)$$

holds for *all* $\theta \in G^\infty$ in a *uniform* fashion. This result is related to the work of Lorentz [17].

8. Strictly stationary stochastic processes

As we have seen in section 7, many problems on the asymptotic distribution of a double sequence $\{x_{nk}\}$ are *equivalent* to a problem of the following form. Given that $\mu_\infty \in I_\infty$ satisfies

$$\mu_\infty(f) = \rho(f) \text{ for } f \in F \subset C(G^\infty), \tag{8.1}$$

(F and the complex-valued function $\rho(f)$ on F given), what then can be said about the values $\mu_\infty(f)$ when $f \in C(G^\infty)$, $f \notin F$?

This problem may in turn be formulated as a problem in the theory of probability, namely, by taking (G^∞, μ_∞) as the underlying probability space. Then each measurable function $X = X(\theta)$ on G^∞ may be regarded as a random variable. If X is complex-valued, one would usually write $\mu_\infty(X(\theta))$ as $E(X)$. Note that the component functions

$$X_k = X_k(\theta) = \theta_k, \qquad k = 0, 1, \ldots, \tag{8.2}$$

are random variables taking values in the compact space G. Because μ_∞ is invariant under the translation T, the so-called joint probability distribution

$$\mu_\infty\{\theta\colon (X_k(\theta), X_{k+1}(\theta), \ldots, X_{k+r-1}(\theta)) \in B\} = \mu_r(B), \quad \text{say,}$$

of $X_k, X_{k+1}, \ldots, X_{k+r-1}$ is the same for all $k \geqq 0$; (here, B denotes any measurable subset of the r-fold direct product $G \times \ldots \times G$). For this reason, the sequence of random variables $\{X_k\}$ is said to be a (strictly) *stationary stochastic process*, (see Doob [7], chapters 10 and 11, Loève [16], chapters 9 and 10). Clearly, any moment problem of the type (8.1) is equivalent to a problem for the stationary stochastic process $\{X_k\}$.

As an illustration, let $U(x)$ denote a fixed function from G into the ring Γ of all $s \times s$ complex-valued matrices, (s fixed, $\Gamma =$ complex numbers if $s = 1$), such that each element $u_{ij}(x)$ of $U(x)$ is a bounded and measurable function on G. Then,

$$(8.3) \qquad Y_k = U(X_k) = U(\theta_k), \qquad k = 0, 1, \ldots,$$

defines a stationary process of random variables taking values in Γ. Its so-called correlation function is defined by

$$(8.4) \qquad R(h) = E(Y_{k+h}Y_k^*) = \int_{G^\infty} U(\theta_{k+h})U(\theta_k)^* \, \mu_\infty(d\theta),$$

$h = \ldots, -1, 0, 1, \ldots$; $(R(h) \in \Gamma$ is independent of $k \geqq \max(0, -h))$. Here, if V is a matrix then V^* denotes its adjoint ($=$ transposed conjugate of V). In particular,

$$(8.5) \qquad R(-h) = R(h)^*.$$

We assert that the $R(h)$ cannot be small without

$$(8.6) \qquad \Phi = E(Y_k) = \int_{G^\infty} U(\theta_k)\mu_\infty(d\theta)$$

being also small. More precisely, let m and q denote positive integers and consider the nonnegative definite matrix

$$V(\theta) = \left(\sum_{j=0}^{m-1} Y_{jq} - m\Phi\right)\left(\sum_{k=0}^{m-1} Y_{kq}^* - m\Phi^*\right).$$

Integrating $V(\theta)$ over all of G^∞ with respect to the nonnegative measure μ_∞, and using (8.4), one obtains that for each choice of the positive integers m and q

$$(8.7) \qquad 0 \ll \Phi\Phi^* \ll \frac{1}{m^2} \sum_{h=-m+1}^{m-1} (m-1-|h|)R(hq);$$

here, $V \ll W$ denotes the property that the matrix $W-V$ is nonnegative definite. In particular, using (8.5), one has $\Phi = 0$ whenever

$$(8.8) \qquad \lim_{\substack{m=m_j \\ j\to\infty}} \frac{1}{m} \sum_{h=1}^{m} \left(1 - \frac{h}{m}\right) R(hq_j) = 0,$$

for some choice of the sequences of positive integers $\{q_j\}$ and $\{m_j\}$ with m_j tending to infinity.

In certain problems, (where $\mu_\infty \in V_\infty\{x_{nk}\}$), the measure μ_∞ is in a natural way equal to the average of certain measures $\mu_\infty^{(0)}, \ldots, \mu_\infty^{(L-1)}$ in I_∞, $(\mu_\infty^{(j)} \in V_\infty\{x_{n,kL+j}\})$. Applying (8.7) with μ_∞ replaced by

$\mu_\infty^{(j)}$ and then adding over j, one finds that (8.8) implies not only $\Phi = 0$ but even

$$\int_{G^\infty} U(\theta_k)\mu_\infty^{(j)}(d\theta) = 0 \text{ for all } j = 0, 1, \ldots, q-1.$$

In the remainder of this section, let us consider the special case that G is an additively written compact group, (not necessarily commutative). The random variables $X_k \in G$ are said to be *uniformly distributed* if their (common) distribution $\mu_1(B)$ coincides with the normalized Haar measure on G. As is well-known, this is the case if and only if

$$\Phi = \int_G U(\theta_k)\mu_1(d\theta_k) = E(U(X_k)) = 0,$$

for *each* irreducible non-trivial unitary representation $U = U(x)$ of G; (by a unitary representation of G we mean a continuous mapping $U(x)$ from G into some group Γ consisting of all $s \times s$ unitary matrices, such that $U(x-y) = U(x)U(y)^*$).

By (8.3) and (8.4), we have for each such representation that

$$R(h) = E(U(X_{k+h}-X_k)),$$

thus, $R(h) = 0$ if $X_{k+h}-X_k \in G$ is uniformly distributed. Consequently, by (8.7), we have the following result.

THEOREM 8.1. *Let $\{X_k\}$ be a strictly stationary stochastic process of random variables taking values in the additive compact group G. Suppose further that $X_{qh}-X_0 \in G$ is uniformly distributed for all $h = 1, 2, \ldots$, (q a fixed positive integer). Then X_0 itself is uniformly distributed, (hence, also the other X_k).*

The above result characterizes the Haar measure, (for, consider $X'_k = X_k+c$ with $c \in G$ constant). One may regard Theorem 8.1 as the probabilistic counterpart of the following generalization, essentially due to Hlawka [9] and Cigler [4], of the van der Corput [6] difference theorem. Here, $A = (a_{nk})$ is as in section 5.

THEOREM 8.2. *Let $\{x_{nk}\}$ be a given double sequence of points in G. Suppose that, for each $h = 1, 2, \ldots$, the double sequence*

$$\{x_{n,k+qh}-x_{n,k}\} \qquad (q \textit{ fixed}),$$

has the uniform distribution on G as its asymptotic A-distribution. Then so has the original double sequence $\{x_{n,k}\}$.

Proof. Let ν_1 denote the uniform (= Haar) measure on G of mass 1. It is given that

(8.9) $$\lim_{n=\infty} \sum_{k=0} a_{nk} f(x_{n,k+qh} - x_{n,k}) = \nu_1(f)$$

for each $f \in C(G)$ and each $h \geqq 1$. Let $g \in C(G)$ be given. We must prove that the sequence

$$\{\sum_{k=0}^{\infty} a_{nk} g(x_{nk})\}$$

converges to $\nu_1(g)$. Draw first a convergent subsequence and then a further subsequence with indices n_j, say, such that the limit

$$\lim_{\substack{n=n_j \\ j\to\infty}} \sum_{k=0}^{\infty} a_{nk} f(x_{n,k}, x_{n,k+1}, \ldots) = \mu_\infty(f)$$

exists for each function $f(\theta) = f(\theta_0, \theta_1, \ldots)$ in $C(G^\infty)$. We must prove that $\mu_\infty(g(\theta_0)) = \nu_1(g)$.

By (8.9),

(8.10) $\mu_\infty(f(\theta_{qh} - \theta_0)) = \nu_1(f)$ for each $f \in C(G)$, each $h \geqq 1$.

By (5.15), $\mu_\infty \in I_\infty$. Consequently, with (G^∞, μ_∞) as the underlying probability space, the sequence of random variables $X_k(\theta) = \theta_k$ $(k = 0, 1, \ldots)$ satisfies the conditions of Theorem 8.1. It follows that X_0 is uniformly distributed, thus,

$$\mu_\infty(g(\theta)) = E(g(X_0)) = \int_G g(x)\nu_1(dx) = \nu_1(g),$$

proving Theorem 8.2.

In fact, the two theorems may be considered equivalent. In deriving Theorem 8.1 from Theorem 8.2, we start with a μ_∞ satisfying (8.10) and $\mu_\infty \in I_\infty$, and then consider a sequence $\{x_{nk} = x_k\}$ having the complete distribution μ_∞, (say, with respect to $A = (C, 1)$). That such a sequence exists follows from Theorem 6.1.

By (8.7), (compare (8.8)), it is possible to weaken considerably the conditions of Theorem 8.1. In an obvious way, compare the above proof, this allows us to weaken in turn the conditions of Theorem 8.2. A direct proof of an even more general result is given in [12].

9. Normal numbers

Let $a \geqq 2$ denote a fixed integer and take G as the finite set

$$G = \{0, 1, \ldots, a-1\}.$$

Let further K denote the additive group of real numbers modulo 1, (0 and 1 identified). Each number ξ in K has a unique expansion

$$(9.1) \qquad \xi = \sum_{k=0}^{\infty} x_k(\xi)a^{-k-1} \pmod 1 \text{ with } x_k(\xi) \in G,$$

provided we do not allow (as we will) an expansion with $x_k(\xi) = a-1$ for all large k.

The number ξ is said to be normal to the base a if, for each choice of the positive integer r and the elements $y_0, \ldots, y_{r-1}$ in G, one has

$$(9.2) \qquad \lim_{n\to\infty} \frac{1}{n} [\text{no. of } 0 \leqq k \leqq n : (x_k(\xi), \ldots, x_{k+r-1}(\xi)) = (y_0, \ldots, y_{r-1})] = a^{-r}.$$

As usual, let

$$G^\infty = G_0 \times G_1 \times G_2 \times \ldots$$

denote the compact space of points

$$\theta = (\theta_0, \theta_1, \theta_2, \ldots) \text{ with } \theta_k \in G.$$

Then

$$(9.3) \qquad \varphi(\xi) = (x_0(\xi), x_1(\xi), x_2(\xi), \ldots)$$

defines a mapping from K into G^∞. The only points θ not in the range of φ are those for which $\theta_k = a-1$ for all large k. The set of these points $\theta \in G^\infty$ will be denoted as S_a. As is easily seen, $\varphi(\xi)$ has only denumerably many discontinuities, namely precisely those points ξ in K admitting an expansion (9.1) with $x_k(\xi) = 0$ for all large k. The set of these points will be denoted as Q_a.

Let us further define

$$(9.4) \qquad \psi(\theta) = \sum_{k=0}^{\infty} \theta_k a^{-k-1} \pmod 1 \quad \text{for } \theta \in G^\infty,$$

which is a continuous mapping from G^∞ onto K. Clearly,

$$(9.5) \qquad \psi(\varphi(\xi)) = \xi,\ \varphi(\psi(\theta)) = \theta,$$

except that the latter is not true when $\theta \in S_a$. The homomorphic mapping $\xi \to a\xi \pmod 1$ of K unto itself will simply be denoted as $a\xi$. If further T denotes the shift transformation in G^∞ it is clear that

$$(9.6) \qquad \varphi(a^k\xi) = T^k\varphi(\xi), \quad \psi(T^k\theta) = a^k\psi(\theta).$$

Recall that ψ is continuous and that φ has only denumerably many discontinuities. Using (9.5), it is easily seen that the formulae

$$\mu'(f) = \mu(f(\varphi(\xi))), \; f \in C(G^\infty) \tag{9.7}$$

and

$$\mu(g) = \mu'(g(\psi(\theta))), \; g \in C(K) \tag{9.8}$$

establish a 1 : 1 correspondence between the probability measures μ on K and the probability measures μ' on G^∞ satisfying $\mu'(S_a) = 0$. Usually, we shall drop the prime, thus, denoting corresponding measures by the same symbol. In this way there corresponds to the Lebesgue measure on K the measure ν_∞ on G^∞ which is the product of the uniform measures on the components G_k, that is,

$$\nu_\infty\{\theta : \theta_0 = y_0, \ldots, \theta_{r-1} = y_{r-1}\} = a^{-r}.$$

Hence, by (9.2), the number ξ in K is normal if and only if the sequence of points $\{x_k(\xi)\}$ in G has a complete asymptotic distribution equal to ν_∞; (in the sequel we shall take $A = (C, 1)$, though any A satisfying (5.1) and (5.2) would do).

We shall prefer to call the number ξ instead a ν_∞-normal number (to the base a). In the same way, if μ_∞ is any invariant probability measure on G^∞ the number ξ will be said to be μ_∞-normal if the sequence $\{x_k(\xi)\}$ has a complete asymptotic distribution equal to μ_∞. Given $\mu_\infty \in I_\infty$, there always exists such a number ξ. This result due to Pyateckii-Šapiro [18] is in fact an immediate consequence of Theorem 6.1; (in a similar way, introducing an appropriate product space, one arrives at a new proof of a more general result due to Cigler [3] p. 95). In the special case that μ_∞ is ergodic with respect to T one even has, by the individual ergodic theorem, that almost $[\mu_\infty]$ all numbers ξ are μ_∞-normal, (to the base a).

Note that, by (9.6) and (9.8), a measure μ_∞ in I_∞ corresponds to a probability measure on K satisfying

$$\mu_\infty(f(a\xi)) = \mu_\infty(f(\xi)) \text{ for all } f \in C(K).$$

Further, for ξ to be normal relative to $\mu_\infty \in I_\infty$ it is necessary and sufficient that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} f(T^k\varphi(\xi)) = \mu_\infty(f)$$

for all $f \in C(G^\infty)$.

Similarly, the sequence $\{a^k\xi\}$ has (mod 1) the asymptotic distribution μ_∞ if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} g(a^k\xi) = \mu_\infty(g)$$

for all $g \in C(K)$, or equivalently, for all bounded functions g on K continuous at almost all $[\mu_\infty]$ points ξ.

Hence, by (9.6) and the continuity of ψ, if ξ is μ'_∞-normal then $\{a^k\xi\}$ has modulo 1 the asymptotic distribution μ_∞ defined by (9.8). Conversely, by (9.6) and the continuity of φ at each point $\xi \notin Q_a$, if $\{a^k\xi\}$ has modulo 1 the asymptotic distribution μ_∞ such that $\mu_\infty(Q_a) = 0$ then ξ is μ'_∞-normal (μ'_∞ defined by (9.7)). This generalizes a result of Wall [20]. If $\mu_\infty \in I_\infty$ satisfies $\mu_\infty(Q_a) > 0$ it may happen that $\{a^k\xi\}$ has modulo 1 an asymptotic distribution μ_∞ while $\{x_k(\xi)\}$ has no complete distribution at all; (for instance, take long blocks $x_k(\xi) = 0$ followed by long blocks $x_k(\xi) = a-1$).

The following result due to Pyateckii-Šapiro [18] (and a generalization to ergodic μ_∞ due to Cigler [3]) plays a useful role in the theory of normal numbers; see [3] and [5].

Let $\{x_k\}$ be a given sequence of points in G and put

$$q_n(y_0, \ldots, y_{r-1}) = \frac{1}{n} (\text{no. of } 0 \leqq k \leqq n : x_k = y_0, \ldots, x_{k+r-1} = y_{r-1}).$$

Then $\{x_k\}$ has the complete asymptotic distribution ν_∞ (= Lebesgue measure) provided

$$\overline{\lim_{n\to\infty}}\ q_n(y_0, \ldots, y_{r-1}) \leqq ca^{-r} \tag{9.5}$$

for each choice of the positive integer r and the y_j in G.

Here, $c > 1$ denotes a fixed constant. I claim that the same conclusion holds if in (9.5) we replace c by a constant $1 < c_r \leqq \infty$ in such a way that

$$\underline{\lim_{r\to\infty}} \frac{1}{r} \log c_r = 0, \tag{9.6}$$

(say, $c_r = \exp(r^{1-\varepsilon})$ with $0 < \varepsilon < 1$). The condition (9.6) can hardly be weakened; for, consider a sequence $\{x_k\}$ having a product measure different from ν_∞ as its complete distribution.

Proof. Let $\mu_\infty \in V_\infty\{x_k\}$ be fixed; by (5.15), $\mu_\infty \in I_\infty$. We must prove that μ_∞ coincides with ν_∞. Let μ_r denote the projection of μ_∞ onto $G^r = G_0 \times \ldots \times G_{r-1}$, ($G_k$ a copy of G). Similarly, let $\nu_r = \nu_1 \times \ldots \times \nu_1$ denote the projection of ν_∞ onto G^r.

It follows by (9.5), (with c replaced by c_r), that the density

$$f_r = f_r(\theta_0, \ldots, \theta_{r-1}) = d\mu_r/d\nu_r$$

satisfies $f_r \leqq c_r$. Thus, the quantity

$$H_r = \int_{G^r} \log f_r \, d\mu_r$$

satisfies

$$(9.7) \qquad 0 \leqq H_r \leqq \log c_r.$$

That $H_r \geqq 0$ follows easily from the convexity of the function $z \log z$; in fact, $H_r = 0$ if and only if $\mu_r = \nu_r$.

Finally, it is known, compare [15] p. 48, that $\mu_\infty \in I_\infty$ implies

$$(9.8) \qquad H_{r+t} \geqq H_r + H_t, \quad r, t = 1, 2, \ldots$$

This together with (9.6) and (9.7) gives that $H_r = 0$ for all r, thus, $\mu_r = \nu_r$ for all r, thus, $\mu_\infty = \nu_\infty$.

The idea which lies behind the above result is the fact that a $\mu_\infty \in I_\infty$ which is absolutely continuous with respect to an ergodic ν_∞ (for instance the Lebesgue measure) must coincide with ν_∞, compare [3]. After all, the density $f(\theta)$ of μ_∞ relative to ν_∞ is an invariant function and thus, constant for almost $[\nu_\infty]$ all θ.

For the benefit of the reader, we present the following simple proof of (9.8) (which applies equally well to the more general case that G is any measurable space, μ_∞ and ν_∞ invariant probability measures on G^∞, with ν_∞ a product measure).

By $\mu_\infty \in I_\infty$ the two functions

$$g_1(\theta) = f_{r+t}(\theta_0, \theta_1, \ldots, \theta_{r+t-1})$$
$$g_2(\theta) = f_r(\theta_0, \ldots, \theta_{r-1}) f_t(\theta_r, \theta_{r+1}, \ldots, \theta_{r+t-1}),$$

have their ν_∞-integral equal to 1, while

$$H_{r+t} - H_r - H_t = \int_{G^\infty} \log \frac{g_1}{g_2} \, d\mu_\infty$$
$$= \int_{G^\infty} \left(\log \frac{g_1}{g_2}\right) g_1 d\nu_\infty = \int_{G^\infty} \psi\left(\frac{g_1}{g_2}\right) g_2 d\nu_\infty.$$

Here, $\psi(z) = z \log z - z + 1$ ($z \geqq 0$, $\psi(0) = 1$), is clearly nonnegative, thus, (9.8) obtains.

REFERENCES

CASSELS, J. W. S.
[1] An extension of the law of the iterated logarithm, Proc. Cambridge Philos. Soc. 47 (1951) 55–64.

CIGLER, J.
[2] Asymptotische Verteilung reeller Zahlen mod 1, Monatshefte für Math. 64 (1960) 201–225.

CIGLER, J.
[3] Der Individuelle Ergodensatz in der Theorie der Gleichverteilung mod 1, J. reine angew. Math. 205 (1960) 91–100.

CIGLER, J.
[4] Über eine Verallgemeinerung des Hauptsatzes der Theorie der Gleichverteilung, J. reine angew. Math. 210 (1962) 141–147.

CIGLER, J. und HELMBERG, J.
[5] Neuere Entwicklungen der Theorie der Gleichverteilung, Jahresber. Deutschen Math. Ver. 64 (1961) 1–50.

CORPUT, J. G. VAN DER
[6] Diophantische Ungleichungen I, Zur Gleichverteilung Modulo Eins, Acta Math. 56 (1931) 373–456.

DOOB, J.
[7] Stochastic processes, John Wiley and Sons, New York, 1953.

FELLER, W.
[8] The general form of the so-called law of the iterated logarithm, Trans. Amer. Math. Soc. 54 (1943) 373–402.

HLAWKA, E.
[9] Zur formalen Theorie der Gleichverteilung in kompakten Gruppen, Rend. Circ. Mat. Palermo Ser. II, 4 (1955) 1–15.

HLAWKA, E.
[10] Folgen auf kompakten Räumen, Abh. math. Sem. Univ. Hamburg 20 (1956) 223–241.

KAC, M.
[11] Probability methods in some problems of analysis and number theory, Bull. Amer. Math. Soc. 55 (1949) 641–665.

KEMPERMAN, J. H. B.
[12] On the distribution of a sequence in a compact group, this journal.

KESTEN, H.,
[13] Uniform distribution mod 1, Annals Math. 71 (1960) 445–471.

KESTEN, H.
[14] Uniform distribution mod 1, II, Acta Arith. 7 (1961) 355–360.

KHINTCHIN, A. I.
[15] Mathematical foundations of information theory, Dover Publ., New York, 1957.

LOÈVE, M.
[16] Probability Theory, 2nd ed., D. van Nostrand Co., Princeton, 1960.

LORENTZ, G. G.

[17] A contribution to the theory of divergent series, Acta Math. 80 (1948) 167–190.

PYATECKII-ŠAPIRO, I. I.

[18] On the laws of distribution of the fractional parts of an exponential function, Izv. Akad. Nauk SSSR Ser. Mat. 15 (1951) 47–52.

STAPLETON, J. H.

[19] Asymptotic distributions modulo 1 and its extension to abstract spaces, Thesis, Purdue University, 1957.

WALL, D. D.

[20] Normal numbers, Thesis, University of California, Berkeley, 1949.

(Oblatum 29-5-63) University of Rochester.