SUMMABILITY METHODS AND ALMOST-SURE CONVERGENCE

N.H. Bingham and L.C.G. Rogers

§0. Introduction

This paper explores links between probability theory and summability theory. Such links are to be expected, since a summability method is essentially a (limit of) a weighted average, while the use of weighted averages – be they expectations, sample means, or variants thereof – is ubiquitous in probability and statistics.

The paper falls into two parts. In §1, we present three results (Theorems 1–3) on limits of occupation times (and for comparison, a result of Brosamler, Theorem 4), the theme being the interplay between density properties of sets and limiting properties of occupation times of sets by random processes. In §§2–4, we survey the general area of links between probability and summability, focusing particularly on the i.i.d. case, and comparing the strengths of the integrability conditions on the distribution and the summability method in the a.s. convergence statement.

To make the paper self-contained, we review here the summability methods that appear below. For background, see e.g. Hardy (1949).

Cesàro methods $C_\alpha, \alpha > 0$: $s_n \rightarrow s (C_\alpha)$ means

$$\frac{1}{A_n^\alpha} \sum_{k=0}^{n} A_{n-k}^{\alpha-1} s_k \rightarrow s \quad (n \rightarrow \infty); \quad A_n^\alpha := (\alpha + 1) \ldots (\alpha + n)/n!.$$ 

Abel method $A$: $s_n \rightarrow s (A)$ means

$$(1 - r) \sum_{k=0}^{\infty} s_k r^k \rightarrow s \quad (r \uparrow 1).$$ 

Riesz method $R(\lambda_n, k)$ of order $k$ based on $\lambda_n \uparrow \infty$: $s_n \rightarrow s \quad (R, \lambda_n, k)$ means

$$\frac{k}{x^k} \int_{0}^{x} A_\lambda(t)(x-t)^{k-1} dt \rightarrow s \quad (x \rightarrow \infty)$$

($A_\lambda(x) := s_n$ for $\lambda_n < x \leq \lambda_{n+1}$).

Euler method $E(\lambda), \lambda > 0$: $s_n \rightarrow s (E(\lambda))$ means

$$(1 + \lambda)^{-n} \sum_{k=0}^{n} s_k \binom{n}{k} \lambda^k \rightarrow s \quad (n \rightarrow \infty).$$ 

Borel method $B$: $s_n \rightarrow s (B)$ means

$$e^{-\lambda} \sum_{k=0}^{\infty} s_k \lambda^k/k! \rightarrow s \quad (\lambda \rightarrow \infty).$$ 

§1. Limits of occupation times
The theme of the results of this section is the use of summability methods to link density properties of sets with limit behaviour of occupation times of these sets by random processes.

We begin by considering sets $A \subset \mathbb{R}^+$; the process will be Brownian motion on $\mathbb{R}$ with unit drift,

$$X_t = B_t + t$$

with $B$ standard Brownian motion; the summability method will be the Cesàro method; $|\cdot|$ denotes Lebesgue measure.

**Theorem 1.** A set $A \subset \mathbb{R}^+$ has Cesàro density $c$,

$$\frac{1}{t} |A \cap [0,t]| \to c \quad (t \to \infty),$$

if and only if its occupation time by drifting Brownian motion $X$ satisfies the strong law

$$\frac{1}{t} \int_0^t I(X_u \in A) du \to c \quad (t \to \infty) \quad \text{a.s.}$$

The random-walk analogue of this result is due to Stam (1968) and Meilijson (1973); for extensions see Bingham and Goldie (1982), Högnäs and Mukherjea (1984), Berbee (1987).

**Proof.** Drifting Brownian motion $X$ is a Lévy process with Lévy exponent $\psi(s) = s + \frac{1}{2} s^2$. Its first-passage process $\tau = (\tau_u)_{u>0}$, where

$$\tau_u := \inf\{ t : X_t > u \}, \quad \tau_0 := 0$$

is a subordinator as $X$ is spectrally negative. Its Lévy exponent $\eta(s)$ satisfies $\psi(\eta(s)) = s$ (see e.g. Bingham (1975), §4), so

$$\eta(s) = -1 + (1 + 2s)^{\frac{1}{2}}.$$ 

Thus $E\tau_1 = \eta'(0) = 1$, $E\tau_u = u$, and by the strong law

$$\tau_u/u \to 1 \quad \text{a.s.} \quad (u \to \infty).$$

Write

$$\xi_n := \int_{\tau_{n-1}}^{\tau_n} I_A(X_u) du, \quad \mu_n := E\xi_n, \quad \tilde{\xi}_n := \xi_n - \mu_n.$$ 

Then the $\tilde{\xi}_n$ are independent zero-mean random variables with

$$\text{var} \tilde{\xi}_n \leq E\xi_n^2 \leq E(\tau_n - \tau_{n-1})^2 = E\tau_1^2 < \infty,$$

so the $\tilde{\xi}_n$ are bounded in $L_2$. The martingale

$$M_n := \sum_{j=1}^{n} \tilde{\xi}_j/j$$

is thus bounded in $L_2$, so almost-surely convergent. By Kronecker’s lemma, this gives

$$\frac{1}{n} \sum_{k=1}^{n} \tilde{\xi}_k = \frac{1}{n} \sum_{k=1}^{n} \xi_k - \frac{1}{n} \sum_{k=1}^{n} \mu_k \to 0 \quad \text{a.s.} \quad (n \to \infty)$$

(1)
Write \((L(t, x))\) for the local time of \(X\), jointly continuous in \(t\) and \(x\) by Trotter’s theorem (see Rogers and Williams (1987), 101). Then

\[
\sum_{1}^{n} \mu_k = E \int_{0}^{\tau_n} I_A(X_u) du
= E \int_{0}^{\tau_n} I_A(x)L(\tau_n, x)dx = \int_{0}^{\tau_n} I_A(x)EL(\tau_n, x)dx.
\]

(2)

To compute \(EL(\tau_n, x)\), we use Tanaka’s formula:

\[
X_{\tau_k}^- = -\int_{\tau_k}^{0} I_{\{X_u \leq 0\}} d(B_u + u) + \frac{1}{2}L(\tau_k, 0).
\]

(Rogers and Williams (1987), IV.43.6). When we take expectations, the stochastic integral with respect to \(B\) contributes nothing, since integrability of \(\tau_k\) implies \(L^2\)-boundedness of the stochastic integral \(\int_{0}^{\tau_k} I_{\{X_u \leq 0\}} dB_u\).

Since \(X_{\tau_k}^- = 0\), we deduce that

\[
EL(\tau_k, 0) = 2E \int_{0}^{\tau_k} I(X_u \leq 0) du = 2E \left( \int_{0}^{\infty} - \int_{\tau_k}^{\infty} \right) I(X_u \leq 0) du.
\]

Now the all-time minimum of \(X_t\) is exponentially distributed with parameter 2 (see e.g. Bingham (1975), Prop. 5b applied to \(-X\)):

\[
P(X_t \leq 0 \text{ for some } t | X_0 = k) = e^{-2k}.
\]

Using this and the strong Markov property at time \(\tau_k\),

\[
EL(\tau_k, 0) = 2(1 - e^{-2k})E \int_{0}^{\infty} I(X_u \leq 0) du = c(1 - e^{-2k}), \text{ say.}
\]

Similarly,

\[
EL(\tau_k, x) = c(1 - e^{-2k(x-k)}) \quad (0 < x < k).
\]

The constant \(c = 2 \int_{0}^{\infty} P(X_u \leq 0) du\) is easily evaluated by simple calculus to be 1.

Hence by (2),

\[
\sum_{1}^{n} \mu_k = \int_{0}^{\tau_n} I_A(x) \left( 1 - e^{-2(n-x)} \right) dx,
\]

so

\[
\sum_{1}^{n} \mu_k \leq \int_{0}^{\tau_n} I_A(x) dx \leq \sum_{1}^{n} \mu_k + \int_{0}^{\tau_n} e^{-2(n-x)} dx \leq \sum_{1}^{n} \mu_k + \frac{1}{2}.
\]

In particular,

\[
\frac{1}{n} \sum_{1}^{n} \mu_k - \frac{1}{n} \int_{0}^{\tau_n} I_A(x) dx \rightarrow 0.
\]

(3)

Combining (1) and (3),

\[
\frac{1}{n} \int_{0}^{\tau_n} I_A(X_u) du - \frac{1}{n} \int_{0}^{\tau_n} I_A(x) dx \rightarrow 0 \quad \text{a.s.}
\]
Now $\tau_t/t \to 1$ a.s., and the integrands are bounded. Hence
\[
\frac{1}{t} \int_0^t I_A(X_u)du - \frac{1}{t} \int_0^t I_A(x)dx \to 0 \quad \text{a.s.} \quad (t \to \infty),
\]
which proves the result, and more. In fact, Theorem 1 is of equi-convergence rather than convergence character: the difference above converges though neither term need do so. This is to be expected, in view of the similar nature of the random-walk result (Bingham and Goldie (1982), Theorems 1,2,2').

Use of Trotter’s theorem in a similar context may be found in Kendall and Westcott (1987), Theorem 6.7.

When Theorem 1 applies,
\[
\frac{1}{t} \int_0^t P(X_u \in A)du \to c \quad (t \to \infty);
\]

\[
P(X_t \in A) \to c \quad \text{in the Cesàro sense}.
\]

If we ask instead for pointwise convergence here, we need $A$ to have density $c$ in a sense correspondingly stronger than the Cesàro sense:

THEOREM 2.

(i) \[ P(X_t \in A) \to c \quad (t \to \infty) \]

if and only if

(ii) \[ \frac{1}{u\sqrt{t}}|A \cap [t, t + u\sqrt{t}]| \to c \quad (t \to \infty) \quad \text{for all } u > 0. \]

Proof. Statement (i) is
\[
\frac{1}{(2\pi t)^{\frac{1}{2}}} \int_0^\infty I_A(y) \exp\left(-\frac{1}{4}(t - y)^2/t\right)dy \to c \quad (t \to \infty),
\]
or
\[
I_A(x) \to c \quad (V) \quad (x \to \infty),
\]

where $V$ is the Valiron method of summability (cf. Hardy (1949), §§9.10, 9.16). Statement (ii), of ‘moving-average’ type, is known to be equivalent to
\[
I_A(.) \to c \quad (R(\psi, 1))
\]

where $R(\psi, 1)$ is a Riesz mean of order 1 (cf. Hardy (1949), §§4.16, 5.16); for the equivalence, see Bingham (1981), Bingham and Goldie (1988). But for bounded functions, $V$ and $R(\psi, 1)$ are known to be equivalent (Bingham and Tenenbaum (1986)).

The density condition (ii) is strictly stronger than the Cesàro density condition in Theorem 1; see Bingham (1981), §1. The Riesz and Valiron methods above are closely linked to the Euler and Borel methods; see §3 below, and for background, Bingham (1984a), (1984c).

Somewhat more classical are the corresponding results for standard (driftless) Brownian motion. Recall the arc-sine law – the law on $[0,1]$ with density $1/(\pi x^{\frac{1}{2}}(1-x)^{\frac{1}{2}})$. The next result is the Brownian analogue of results of Davydov and Ibragimov (1971), Davydov (1973), (1974); cf. Bingham and Goldie (1982), Th.B.
THEOREM 3. For $A \subset \mathbb{R}^+$ and $B$ standard Brownian motion, the following are equivalent:

(i) $\frac{1}{t} |A \cap [0, t]| \to c,$

(ii) $P(B_t \in A) \to c,$

(iii) $\frac{1}{t} \int_0^t I(B_u \in A) du$ converges in law.

and then the limit law is that of $c\xi$ where $\xi$ is arc-sine.

The special case $A = \mathbb{R}^+, c = 1$ is Lévy’s arc-sine law (Lévy (1939)). For a modern proof of this classical result, see Williams (1979), III.38.10, or Rogers and Williams (1987) VI.53; further references are Kac (1951), Itô and McKean (1965), p.57, Williams (1969), Takács (1981), Pitman and Yor (1986), Karatzas and Shreve (1988), p.273 and p.422.
COROLLARY. (Lévy’s arc-sine law).

\[ \frac{1}{t} \int_0^t I(B_u \geq 0) du \text{ has the arc-sine law for each } t > 0. \]

Proof. By Brownian scaling, the law is the same for each \( t \), and so coincides with the limit law as \( t \to \infty \), which is arc-sine by the theorem.

Proof of Theorem 3. We now deduce the theorem from the corollary (showing the equivalence of the two results). We present a streamlined proof in the spirit of the proof of Theorem 1. Note that (i) is

\[ (i') \quad \frac{1}{t} \int_0^\infty g_1(x/t) I_A(x) dx \to c \quad (t \to \infty) \]

with \( g_1 := I_{[0,1]} \); while (ii) is

\[ (i'') \quad \frac{1}{t} \int_0^\infty g_2(x/t) I_A(x) dx \to c \quad (t \to \infty) \]

with \( g_2(x) := \exp\{-\frac{1}{2}x^2\}/\sqrt{2\pi} \). Now \( g_1, g_2 \) have Mellin transforms

\[ \hat{g}_1(s) := \int_0^\infty g_1(x) x^s dx = \int_0^1 x^s dx = 1/(1 + is), \]

\[ \hat{g}_2(s) := \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}x^2} x^s dx = \frac{1}{\sqrt{2\pi}} \Gamma(\frac{1}{2} + is)/\sqrt{\pi}. \]

Both are non-zero for all real \( s \), so both \( g_1 \) and \( g_2 \) may be used as Wiener kernels in the Mellin form of Wiener’s Tauberian theorem (Hardy (1949), Th.232) since \( f := I_A \) is bounded.

(i) \( \Rightarrow \) (ii): Use Wiener’s theorem as above (Davydov and Ibragimov (1971)).

(i) \( \Rightarrow \) (iii): The measures \( \mu_t \) defined by

\[ \mu_t(x) := \frac{1}{t} \int_0^t I_A(y) dy \]

converge weakly to \( c \) times Lebesgue measure on \( \mathbb{R}^+ \) as \( t \to \infty \). Also, by Brownian scaling, \( L(t,x) = cL(t/c^2,x/c) \) in law. So

\[ \frac{1}{t} \int_0^t I_A(B_u) du = \frac{1}{t} \int_A dy L(t,y) \]

= \( \frac{1}{t} \int_A dy \sqrt{t} L(1,y/\sqrt{t}) \) in law

= \( \int A(v\sqrt{t}) L(1,v) dv \)

= \( \int L(1,v) du \sqrt{t}(v) \)

\( \to c \int_0^\infty L(1,v) dv \)

(by compact support of \( L(1,v) \))

\[ = c \int_0^1 I_{R^+}(B_u) du \]

\[ = c \xi, \]

with \( \xi \) arc-sine by Lévy’s result.
(iii) ⇒ (i): Taking expectations,
\[ \int EL(1, v) d\mu, \sqrt{t} \rightarrow \frac{1}{2}c. \]

Now
\[ EL(1, v) = \int_0^1 \frac{e^{-\frac{1}{2}v^2/t}}{\sqrt{2\pi}} \sqrt{t} \rightarrow f(v), \text{ say}, \]
where
\[ \int_0^\infty f(v) v^i \, dv = \int_0^1 t^{-\frac{1}{2}} dt \int_0^\infty e^{-\frac{1}{2}v^2/t} v^i \, dv/\sqrt{2\pi} = \int_0^1 t^{-\frac{1}{2}i} dt \int_0^\infty e^{-\frac{1}{2}u^2} u^i \, du/\sqrt{2\pi} = (1 + \frac{i}{2}s)^{-1} \hat{g}(s), \]
which is non-zero for real \( s \) as above. Thus \( f \) is a Wiener kernel, and (i) follows by Wiener’s Tauberian theorem as above.

To obtain strong-law behaviour as in Theorem 1, one needs to coarsen the Cesàro averaging, rather than refining it as in Theorem 2. The appropriate summability method is the logarithmic one (or Riesz mean \( R(\log n, 1) \)). Logarithmic averages were introduced in probability theory by Lévy (1937), 270 (cf. Chung and Erdős (1951), Th.6, Erdős and Hunt (1953), Th.4); the result below may thus be dubbed ‘Lévy’s strong arc-sine law’. For extensions, see Révész’s contribution to this volume.

**THEOREM 4.**
\[ \frac{1}{\log t} \int_1^t I(B_u \geq 0) du/u \rightarrow \frac{1}{2} \text{ a.s. (} t \rightarrow \infty) \]

**First Proof.** Writing \( u = e^v \) and replacing \( t \) by \( e^t \), we have to show
\[ \frac{1}{t} \int_0^t I(B(e^v) \geq 0) dv \rightarrow \frac{1}{2} \text{ a.s. (} t \rightarrow \infty). \]

Now \( Y(t) := e^{-\frac{1}{2}t}B(e^t) \) is an Ornstein-Uhlenbeck process (see e.g. Karlin and Taylor (1981), 380), so we have to show
\[ \frac{1}{t} \int_0^t I_{[0, \infty)}(Y_u) du \rightarrow \frac{1}{2} \text{ a.s. (} t \rightarrow \infty). \]

Now the speed measure \( m \) of an Ornstein-Uhlenbeck process is finite, and so may be scaled to a probability measure \( \pi \), which is Gaussian with mean zero. This follows from the stochastic differential equation for the Ornstein-Uhlenbeck process: see Rogers and Williams (1987), V.5.2(ii), V.52.1-2. The result now follows from the ergodic theorem for diffusions,
\[ \frac{1}{t} \int_0^t f(Y_u) du \rightarrow \int f(x) d\pi(x) \text{ a.s. (} t \rightarrow \infty), \]
with \( f = I_{[0, \infty)} \) and \( \pi \) Gaussian, mean 0 (Rogers and Williams (1987), V.53.5).

Lévy’s strong arc-sine law was rediscovered independently (on an equivalent formulation) by Brosamler (1973), Th.1. Use of the Ornstein-Uhlenbeck process in this context may also be found on Brosamler (1986), 314, (1988), 563-4. We thank Michael Lacey for these observations.

**Second Proof.** This follows from the pathwise central limit theorem, again taking \( f = I_{[0, \infty)} \) and using symmetry of a mean-zero Gaussian measure. See Brosamler (1988), Th.1.6; cf. Schatte (1988), Lacey and Philipp (1989+), Fisher (1990+).

The relationship between the three summability methods used in this section may be expressed by
\[ R(\sqrt{n}, 1) \subset R(n, 1) \subset R(\log n, 1). \]
The general result, comparing $R(\lambda_n, k)$ for different $\lambda_n$ and the same $k$, is the first consistency theorem for Riesz means; see e.g. Chandrasekharan and Minakshisundaram (1952), Ch.1.

§2. Cesàro and Riesz means

We turn now to more traditional links between summability methods and strong laws. Let $X, X_1, X_2, \ldots$ be independent and identically distributed (iid) random variables. The classical Kolmogorov strong law

$$E|X| < \infty \text{ and } EX = \mu \Leftrightarrow \frac{1}{n} \sum_{k=1}^{n} X_k \to \mu \text{ a.s.}$$

may be rephrased as

$$X \in L_1 \text{ and } EX = \mu \Leftrightarrow X_n \to \mu \text{ a.s. } (C),$$

where $C(= C_1)$ is the Cesàro method of summability. There is a Cesàro method $C_\alpha$ for every positive $\alpha$ (Hardy (1949), V-VII); it was shown by Lai (1974a) that $C_\alpha$ may be replaced here by $C_\alpha$ for any $\alpha \geq 1$, or by the Abel method $A$. There are similar versions of the law of the iterated logarithm (Gaposhkin (1965), Lai (1974a)).

For $0 < \alpha < 1$ the situation is different: a.s. $C_\alpha$-convergence is tied to membership of $L_{1/\alpha}$, not to $L_1$: for $p \geq 1$,

$$X \in L_p \text{ and } EX = \mu \Leftrightarrow X_n \to \mu \text{ a.s. } (C_{1/p})$$

(Déniel and Derriennic (1988)).

One may improve the forward implication here (which is the harder and more important) by replacing $C_{1/p}$ by a more stringent summability method. It turns out that such a method is provided by the Riesz mean $R_p := R(\exp \int_{1}^{n} dx/x^{1/p}, 1) : R_p \subset C_{1/p}$. For $p = 1$, $R_1 = C_1$, but the inclusion is strict for $p > 1$; for details see Bingham (1989).

The Riesz formulation also extends to moments more general than powers. For suitable functions $\phi$, Riesz means $R_\phi := R(\exp \int_{1}^{n} dx/\phi(x), 1)$ may be linked similarly with membership of a class of Orlicz type, $L_\phi := \{ X : E\phi^{\lambda}(|X|) < \infty \}$:

$$X \in L_\phi \text{ and } EX = \mu \Leftrightarrow X_n \to \mu \text{ a.s. } (R_\phi).$$

Also, $R_\phi$ may be written as a summability method of moving-average (or ‘delayed-average’) type (Bingham and Goldie (1988); Chow (1973); Lai (1974b)): Riesz convergence here is

$$\frac{1}{u\phi(x)} \sum_{x \leq n < x+u\phi(x)} X_n \to \mu \text{ a.s. } \forall u > 0.$$

This moving-average formulation allows one to use results of LIL type by de Acosta and Kuelbs (1983). These authors also consider the Banach-valued case. Further, they give detailed results for the case of slow growth of $\phi - \phi(x) = c \log x$, or $o(\log x)$ – when strong laws of the above type break down. They are replaced by results of Erdős-Rényi type, where one obtains, instead of the a.s. limit $\mu$ above (‘a.s. invariance principle’), an a.s. limit superior, $\alpha = \alpha(u)$, which as $u$ varies completely determines the law of $X$ (‘a.s. non-invariance principle’). For background on the invariance/non-invariance dichotomy, see Deheuvels and Steinebach (1987).

§3. Euler, Borel and related methods
We recall the classical summability methods of Euler \((E(\lambda), \lambda > 0)\) and Borel \((B)\); see Hardy (1949), VIII, IX. These are closely related; methods of Euler-Borel type are perhaps the most important classical summability methods after those of Cesàro-Abel type. They possess an analogue of the above law of large numbers (Chow (1973)) and law of the iterated logarithm (Lai (1974a)), displayed as the Euler and Borel cases of Theorems 5 and 6 below.

In the proofs of these results for \(E(\lambda)\) and \(B\), the most important feature of the Euler weights \(\left(\frac{n}{k}\right)^\lambda / (1 + \lambda)^n\) and the Borel weights \(e^{-x}x^k/k\) is that they arise (for \(x = n\)) as \(n\)-fold convolutions of the binomial and Poisson distributions respectively, allowing use of the central limit theorem in some form. One may seek to generalise this, and consider weighted sums \(\sum a_{nk}X_k\), where the matrix \(A = (a_{nk})\) is of convolution type:

\[
a_{nk} = P(S_n = k).
\]

Here \(S_n = \sum_1^n \xi_k\), with the \(\xi_n\) independent, \(Z\)-valued random variables. There are two important cases:

\(\text{i}(a)\) \(\xi_n\) identically distributed (with mean \(m\) and variance \(d^2\), say). Then \((S_n)\) is a random walk. \(S_n\) has mean \(nm\) and variance \(nd^2\), and \(A\) is called a summability method of random-walk type (Bingham (1984b), (1984c)).

\(\text{i}(b)\) \(\xi_n\) \(\{0, 1\}\)-valued (Bernoulli): \(P(\xi_n = 1) = p_n\), say, \(P(\xi_n = 0) = q_n := 1 - p_n\). Then the Bernoulli sum \(S_n\) has mean \(\mu_n := \sum_1^n p_k\), variance \(\sigma_n^2 := \sum_1^n p_k(1 - p_k)\). Writing \(p_n = 1/(1 + d_n)(d_n \geq 0)\), one then has

\[
\prod_{j=1}^n \left(\frac{x + d_j}{1 + d_j}\right) = \sum_{k=0}^n a_{nk}x^k.
\]

The method \(A = (a_{nk})\) is the Jakimovski method \([F, d_n]\) (Jakimovski (1959); Zeller and Beekmann (1970); Ergänzungen, §70). The motivating examples are:

\(\text{i}(i)\) \(d_n = 1/\lambda\), the Euler method \(E(\lambda)\) above

\(\text{i}(ii)\) \(d_n = (n - 1)/\lambda\), the Karamata-Stirling method \(KS(\lambda)\) (Karamata (1935); Bingham (1988)).

**THEOREM 5.** The following are equivalent:

\(\text{i}(i)\) \(\text{var} X < \infty, EX = \mu\)

\(\text{i}(ii)\) \(X_n \to \mu \text{ a.s.} (E(\lambda), \text{ or } B)\)

\(\text{i}(iii)\) \(X_n \to \mu \text{ a.s.} (A)\), for \(A\) a random-walk method

\(\text{i}(iv)\) \(X_n \to \mu \text{ a.s.} (KS(\lambda))\), for some (all) \(\lambda > 0\)

\(\text{i}(v)\) \(X_n \to \mu \text{ a.s.} [F, d_n]\), for \(d_n \geq \varepsilon > 0\) for some \(\varepsilon\) and large \(n\).

**THEOREM 6.** The following are equivalent:

\(\text{i}(i)\) \(EX = 0, \text{var} X = \sigma^2, E(|X|^4 / \log^2 |X|) < \infty\)

\(\text{i}(ii)\) \(\lim_{x \to \infty} \frac{4\pi x}{\log x} \int_0^\infty e^{-x}x^k X_k \, dx \to \sigma a \text{ a.s.}\)

\(\text{i}(iii)\) \(\lim_{n \to \infty} \frac{4\pi n}{\log n} \left[\sum_0^n \left(\frac{n}{k}\right)\lambda^k X_k / (1 + \lambda)^n\right] = \sigma(1 + \lambda) \text{ a.s.}\)

\(\text{i}(iv)\) \(\lim_{n \to \infty} \frac{4\pi n}{\log n} \left[\sum a_{nk} X_k \right] = \sigma a \text{ a.s.}\)

where \(A = (a_{nk})\) is a random-walk method with mean-variance ratio \(a := m/d^2\),
\( \limsup_{n \to \infty} \frac{(4\pi \lambda \log n)^{1/4}}{\log \log n} |\sum_0^n a_{nk}X_k| = \sigma \text{ a.s., with } A = KS(\lambda) \)

\( \limsup_{n \to \infty} \frac{(4\pi \mu n)^{1/4}}{\log \log \mu n} |\sum_0^n a_{nk}X_k| = \sigma \text{ a.s.} \)

with \( A = [F, d_n] \) a Jakimovski method with \( d_n \to \infty \).

Here the Euler and Borel parts are due to Chow (1973) and Lai (1974a) respectively; the random-walk parts are in Bingham and Maejima (1985); the Jakimovski and Karamata-Stirling parts are in Bingham and Stadtmüller (1990). The proofs proceed by using normal approximation on the weights \( a_{nk} \), specifically Petrov’s local limit theorem (Petrov (1975), VII.3, Th.16) to reduce to the case

\[
a_{nk} = \frac{1}{\sigma(2\pi n)^{1/2}} \exp\left\{-\frac{1}{2} (k - n\mu)^2/n\sigma^2\right\}
\]

(or analogue in the Bernoulli, non-identically distributed case). This reduces to the Valiron summability method (Bingham (1984c); cf. the proof of Theorem 2), and one argues as in Lai (1974a),(16). We note in passing that Poisson, rather than normal, approximation is also possible (Bingham and Stadtmüller (1990),§4.2). This involves the Chen-Stein method, which has been studied extensively recently; see for instance Stein (1986); Barbour (1987); Arratia, Goldstein and Gordon (1989).

We note the the \( KS(\lambda) \) methods have numerous probabilistic uses, in contexts such as random permutations, records, and greatest convex minorants; for details and references, see Bingham (1988), §3.2. Recent applications include work of Hansen (1987), (1990) on random mappings and the Ewens sampling formula of mathematical genetics.

§4. Complements

1. **Bernstein polynomials.** The classical proof of the Weierstrass approximation theorem (due to S.N. Bernstein in 1912),

\[
f(x) = \lim_{n \to \infty} \sum_0^n f(k/n) \binom{n}{k} x^k (1-x)^{n-k}, \quad f \in C[0,1],
\]

has led to many results linking laws of large numbers with summability methods (here the Euler, but others also); for background see Lorentz (1953); Feller (1971), VII; Goldstein (1975), (1976).

2. **Density estimation.** The Bernstein approximation theorem provides one route into the important subject of density estimation, specifically, estimators of smoothed histogram type. For details and references, see Gawronski (1985).

3. **Non-parametric regression.** Asymptotics of matrix transforms \( \sum a_{nk}X_k \) have applications to non-parametric estimation of regression curves. For details and references, see Stadtmüller (1984); Lai and Weh (1982).

4. **Time series.** Similarly, the a.s. behaviour of sums \( \sum a_{nk}X_k \) has applications to time-series models; see Lai and Weh (1982).
REFERENCES

A. de ACOSTA and J. KUELBS (1983): Limit theorems for moving averages of random vectors. Z. Wahrschein. 64, 67-123.


K.-L. CHUNG and P. ERDŐS (1951): Probability limit theorems assuming only the first moment, I. Memoirs Amer. Math. Soc. 6 (four papers on probability).


