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 $C_\alpha, \alpha > 0$: $\quad s_n \to s\ (C_\alpha)$ means

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

Abel method A : $\quad s_n \to s\ (A)$ means

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

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

$$\frac{k}{x^k}\int_0^x A_\lambda(t)(x-t)^{k-1}dt \to s \quad (x \to \infty)$$

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

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

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

Borel method B : $\quad s_n \to s\ (B)$ means

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

§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 $A \subset R^+$; the process will be Brownian motion on 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 $A \subset R^+$ 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

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

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

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

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

$$M_n := \sum_1^n \tilde{\xi}_j/j$$

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

$$\frac{1}{n}\sum_1^n \tilde{\xi}_k = \frac{1}{n}\sum_1^n \xi_k - \frac{1}{n}\sum_1^n \mu_k \to 0 \quad \text{a.s.} \tag{1}$$

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

$$\begin{aligned}\sum_1^n \mu_k &= E\int_0^{\tau_n} I_A(X_u)du \\ &= E\int_0^n I_A(x)L(\tau_n,x)dx \\ &= \int_0^n I_A(x)EL(\tau_n,x)dx.\end{aligned} \tag{2}$$

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

$$X^-_{\tau_k} = -\int_0^{\tau_k} I_{\{X_u\le 0\}}d(B_u+u) + \tfrac{1}{2}L(\tau_k,0).$$

(Rogers and Williams (1987), IV.43.6). When we take expectations, the stochastic integral with respect to B contributes nothing, since integrability of τ_k implies L^2-boundedness of the stochastic integral $\int_0^{t\wedge\tau_k} I_{\{X_u\le 0\}}dB_u$. Since $X^-_{\tau_k}=0$, we deduce that

$$\begin{aligned}EL(\tau_k,0) &= 2E\int_0^{\tau_k} I(X_u\le 0)du \\ &= 2E\left(\int_0^\infty - \int_{\tau_k}^\infty\right) I(X_u\le 0)du.\end{aligned}$$

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

$$P(X_t\le 0 \text{ for some } t\,|X_0=k)=e^{-2k}.$$

Using this and the strong Markov property at time τ_k,

$$\begin{aligned}EL(\tau_k,0) &= 2(1-e^{-2k})E\int_0^\infty I(X_u\le 0)du \\ &= c(1-e^{-2k}),\ \text{say}.\end{aligned}$$

Similarly,

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

The constant $c=2\int_0^\infty P(X_u\le 0)du$ is easily evaluated by simple calculus to be 1.

Hence by (2),

$$\sum_1^n \mu_k = \int_0^n I_A(x)\ \ (1-e^{-2(n-x)})dx,$$

so

$$\sum_1^n \mu_k \le \int_0^n I_A(x)dx \le \sum_1^n \mu_k + \int_0^n e^{-2(n-x)}dx \le \sum_1^n \mu_k + \tfrac{1}{2}.$$

In particular,

$$\frac{1}{n}\sum_1^n \mu_k - \frac{1}{n}\int_0^n I_A(x)dx \to 0. \tag{3}$$

Combining (1) and (3),

$$\frac{1}{n}\int_0^{\tau_n} I_A(X_u)du - \frac{1}{n}\int_0^n I_A(x) \to 0 \quad \text{a.s.}$$

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) $$P(X_t \in A) \to c \qquad (t \to \infty)$$

if and only if

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

Proof. Statement (i) is

$$\frac{1}{(2\pi t)^{\frac{1}{2}}}\int_0^\infty I_A(y)\exp\{-\tfrac{1}{2}(t-y)^2/t\}dy \to c \quad (t \to \infty),$$

or

$$I_A(x) \to c \qquad (V) \qquad (x \to \infty),$$

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

$$I_A(.) \to c \quad (R(e^{\sqrt{n}}, 1))$$

where $R(e^{\sqrt{n}}, 1)$ 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 $R(e^{\sqrt{n}}, 1)$ 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 $A \subset R^+$ 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)
$$\frac{1}{t}\int_0^t I(B_u \in A)du \quad \textit{converges in law.}$$

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

The special case $A = R^+, c = 1$ 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).

$\frac{1}{t}\int_0^t I(B_u \geq 0)du$ *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

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

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

$$\text{(ii}'\text{)} \qquad \frac{1}{t}\int_0^\infty g_2(x/t)I_A(x)dx \to c \quad (t \to \infty)$$

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

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

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

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 μ_t defined by

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

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

$$\begin{aligned}
\frac{1}{t}\int_0^t I_A(B_u)du &= \frac{1}{t}\int_A dyL(t,y) \\
&= \frac{1}{t}\int_A dy.\sqrt{t}L(1, y/\sqrt{t}) \quad \text{in law} \\
&= \int I_A(v\sqrt{t})L(1,v)dv \\
&= \int L(1,v)d\mu_{\sqrt{t}}(v) \\
&\to c\int_0^\infty L(1,v)dv
\end{aligned}$$

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

$$\begin{aligned}
&= c\int_0^1 I_{R^+}(B_u)du \\
&= c\xi,
\end{aligned}$$

with ξ arc-sine by Lévy's result.

(iii) $\Rightarrow$ (i): Taking expectations,

$$\int EL(1,v)d\mu_{\sqrt{t}}(v) \to \tfrac{1}{2}c.$$

Now

$$EL(1,v) = \int_0^1 \frac{e^{-\frac{1}{2}v^2/t}}{\sqrt{2\pi}} \frac{dt}{\sqrt{t}} = f(v), \text{ say,}$$

where

$$\begin{aligned}\int_0^\infty f(v)v^{is}dv &= \int_0^1 t^{-\frac{1}{2}}dt \int_0^\infty e^{-\frac{1}{2}v^2/t}v^{is}dv/\sqrt{2\pi} \\ &= \int_0^1 t^{\frac{1}{2}is}dt \int_0^\infty e^{-\frac{1}{2}u^2}u^{is}du/\sqrt{2\pi} \\ &= (1+\tfrac{1}{2}is)^{-1}\hat{g}_2(s),\end{aligned}$$

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 $R(\log n, 1)$; Hardy (1949), Th.37). 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. $\frac{1}{\log t}\int_1^t I(B_u \geq 0)du/u \to \frac{1}{2}$ a.s. $(t \to \infty)$.

First Proof. Writing $u = e^v$ and replacing t by e^t, we have to show

$$\frac{1}{t}\int_0^t I(B(e^v) \geq 0)dv \to \tfrac{1}{2} \quad \text{a.s.} \quad (t \to \infty).$$

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

$$\frac{1}{t}\int_0^t I_{[0,\infty)}(Y_v)dv \to \tfrac{1}{2} \quad \text{a.s.} \quad (t \to \infty).$$

Now the speed measure m of an Ornstein-Uhlenbeck process is finite, and so may be scaled to a probability measure π, 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,

$$\frac{1}{t}\int_0^t f(Y_u)du \to \int f(x)d\pi(x) \quad \text{a.s.} \quad (t \to \infty),$$

with $f = I_{[0,\infty)}$ and π 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

$$R(e^{\sqrt{n}}, 1) \subset R(n,1) \subset R(\log n, 1).$$

The general result, comparing $R(\lambda_n, k)$ for different λ_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

$$E|X| < \infty \text{ and } EX = \mu \Leftrightarrow \frac{1}{n}\sum_1^n X_k \to \mu \text{ a.s.}$$

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_α for every positive α (Hardy (1949), V-VII); it was shown by Lai (1974a) that C may be replaced here by C_α 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_α-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 ϕ, 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, $L_\phi := \{X : E\overleftarrow{\phi}(|X|) < \infty\}$:

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

Also, R_ϕ 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 μ 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 $\binom{n}{k}\lambda^k/(1+\lambda)^n$ 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 ξ_n independent, Z-valued random variables. There are two important cases:

ı(a) ξ_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)).

ı(b) ξ_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

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

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

ı(i) $d_n = 1/\lambda$, the Euler method $E(\lambda)$ above

ı(ii) $d_n = (n-1)/\lambda$, the Karamata-Stirling method $KS(\lambda)$ (Karamata (1935); Bingham (1988)).

THEOREM 5. *The following are equivalent*:

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

ı(ii) $X_n \to \mu$ a.s. ($E(\lambda)$, or B)

ı(iii) $X_n \to \mu$ a.s. (A), *for* A *a random-walk method*

ı(iv) $X_n \to \mu$ a.s. $(KS(\lambda))$, *for some (all)* $\lambda > 0$

ı(v) $X_n \to \mu$ a.s. $[F, d_n]$, *for* $d_n \geq \varepsilon > 0$ *for some* ε *and large* n.

THEOREM 6. *The following are equivalent*:

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

ı(ii) $\limsup_{x\to\infty} \frac{(4\pi x)^{\frac14}}{\log^{\frac12} x}\left|\sum_0^\infty e^{-x}\frac{x^k}{k!}X_k\right| = \sigma$ a.s.

ı(iii) $\limsup_{n\to\infty} \frac{(4\pi n)^{\frac14}}{\log^{\frac12} n}\left|\sum_0^n \binom{n}{k}\lambda^k X_k/(1+\lambda)^n\right| = \sigma(1+\lambda)^{\frac14}$ a.s.

ı(iv) $\limsup_{n\to\infty} \frac{(4\pi n)^{\frac14}}{\log^{\frac12} n}\left|\sum a_{nk}X_k\right| = \sigma a^{\frac14}$ a.s.

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

ı(vi) $\limsup_{n\to\infty} \frac{(4\pi\mu_n)^{\frac{1}{4}}}{\log^{\frac{1}{2}} \mu_n} |\sum_0^n a_{nk}X_k| = \sigma$ 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

$$a_{nk} = \frac{1}{\sigma(2\pi n)^{\frac{1}{2}}} \exp\{-\tfrac{1}{2}(k-n\mu)^2/n\sigma^2\}$$

(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.

R. ARRATIA, L. GOLDSTEIN and L. GORDON (1989): Two moments suffice for Poisson approximation: the Chen-Stein method. *Ann. Probab.* **17**, 9-25.

A.D. BARBOUR (1987): Asymptotic expansions in the Poisson limit theorem. *Ann. Probab.* **15**, 748-766.

H.C.P. BERBEE (1987): Convergence rates in the strong law for bounded mixing sequences. *Probab. Th. Rel. Fields* **74**, 255-276.

N.H. BINGHAM (1975): Fluctuation theory in continuous time. *Adv. Appl. Probab.* **7**, 705-766.

N.H. BINGHAM (1981): Tauberian theorems and the central limit theorem. *Ann. Probab.* **9**, 221-231.

N.H. BINGHAM (1984a): On Euler and Borel summability. *J. London Math. Soc.* (2) **29**, 141-146.

N.H. BINGHAM (1984b): Tauberian theorems for summability methods of random-walk type. *J. London Math. Soc.* (2) **30**, 281-287.

N.H. BINGHAM (1984c): On Valiron and circle convergence. *Math. Z.* **186**, 273-286.

N.H. BINGHAM (1988): Tauberian theorems for Jakimovski and Karamata-Stirling methods. *Mathematika* **35**, 216-224.

N.H. BINGHAM (1989): Moving averages. *Almost Everywhere Convergence I* (ed. G.A. Edgar and L. Sucheston), Academic Press, 131-144.

N.H. BINGHAM and C.M. GOLDIE (1982): Probabilistic and deterministic averaging. *Trans. Amer. Math. Soc.* **269**, 453-480.

N.H. BINGHAM and C.M. GOLDIE (1988): Riesz means and self-neglecting functions. *Math. Z.* **199**, 443-454.

N.H. BINGHAM and M. MAEJIMA (1985): Summability methods and almost-sure convergence. *Z. Wahrschein.* **68**, 383-392.

N.H. BINGHAM and U. STADTMÜLLER (1990): Jakimovski methods and almost-sure convergence. *Disorder in Physical Systems* (*J.M. Hamersley Festschrift*, ed. G.R. Grimmett and D.J.A. Welsh), Oxford University Press, 5-18.

N.H. BINGHAM and G. TENENBAUM (1986): Riesz and Valiron means and fractional moments. *Math. Proc. Cambridge Phil. Soc.* **99**, 143-149.

G.A. BROSAMLER (1973): The asymptotic behaviour of certain additive functionals of Brownian motion. *Invent. Math.* **20**, 87-96.

G.A. BROSAMLER (1986): Brownian occupation times on compact manifolds. *Seminar on Stochastic Processes 1985*, 290-322, Birkhäuser.

G.A. BROSAMLER (1988): An almost-everywhere central limit theorem. *Math. Proc. Cambridge Phil. Soc.* **104**, 561-574.

K. CHANDRASEKHARAN and S. MINAKSHISUNDARAM (1952): *Typical Means*, Oxford University Press.

Y.-S. CHOW (1973): Delayed sums and Borel summability of independent, identically distributed random variables. *Bull. Inst. Math. Acad. Sinica* **1**, 207-220.

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).

Yu.A. DAVYDOV (1973): Limit theorems for functionals of processes with independent increments. *Th. Probab. Appl.* **18**, 431-441.

Yu.A. DAVYDOV (1974): Sur une classe de fonctionelles des processus stables et marches aléatoires. *Ann. Inst. H. Poincaré (B)* **10**, 1-29.

Yu.A. DAVYDOV and I.A. IBRAGIMOV (1971): On asymptotic behaviour of some functionals of processes with independent increments. *Th. Probab. Appl.* **16**, 162-167.

P. DEHEUVELS and J. STEINEBACH (1987): Exact convergence rates in strong approximation laws for large increments of partial sums. *Probab. Th. Rel. Fields* **76**, 369-393.

Y. DÉNIEL and Y. DERRIENNIC (1988): Sur la convergence presque sûre, au sens de Cesàro d'ordre α, $0 < \alpha < 1$, des variables aléatoires indépendantes et identiquement distribuées. *Probab. Th. Rel. Fields* **79**, 629-636.

P. ERDŐS and G.A. HUNT (1953): Changes of signs of sums of random variables. *Pacific J. Math.* **3**, 678-687.

W. FELLER (1971): *An Introduction to Probability Theory and its Applications, Vol.II*, 2nd ed., Wiley.

A. FISHER (1990+): A pathwise central limit theorem for random walks. *Ann. Probab.*

V.F. GAPOSHKIN (1965): The law of the iterated logarithm for Cesàro's and Abel's methods of summation. *Th. Probab. Appl.* **10**, 411-420.

W. GAWRONSKI (1985): Strong laws for density estimators of Bernstein type. *Period. Math. Hungar.* **16**, 23-43.

J.A. GOLDSTEIN (1975): Some applications of the law of large numbers. *Bol. Soc. Bras. Mat.* **6**, 25-38.

J.A. GOLDSTEIN (1976): Semigroup-theoretic proofs of the central limit theorem and other limit theorems of analysis. *Semigroup Forum* **12**, 189-206.

J.C. HANSEN (1987): A functional central limit theorem for random mappings. *Ann. Probab.* **17**, 317-332.

J.C. HANSEN (1990): A functional central limit theorem for Ewens' sampling formula. *J. Appl. Probab.* **27**, 28-43.

G.H. HARDY (1949): *Divergent Series.* Oxford University Press.

G. HÖGNÄS and A. MUKHERJEA (1984): On the limit of the average of values of a function at random points. *Lecture Notes in Math.* **1064**, Springer, 204-218 (Probab. Groups VII).

K. ITÔ and H.P. McKEAN (1965): *Diffusion Processes and Their Sample Paths.* Springer, Berlin.

A. JAKIMOVSKI (1959): A generalisation of the Lototsky method. *Michigan Math. J.* **6**, 277-296.

M. KAC (1951): On some connections between probability theory and differential and integral equations. *Proc. 2nd Berkeley Symp. Math. Statists. Prob.*, University of California Press, Berkeley, 189-215.

J. KARAMATA (1935): Théoremès sur la sommabilité exponentielle et d'autres sommabilités s'y rattachant. *Mathematica (Cluj)* **9**, 164-178.

I. KARATZAS and S.E. SHREVE (1988): *Brownian Motion and Stochastic Calculus.* Springer, Berlin.

S. KARLIN and H.L. TAYLOR (1981): *A Second Course in Stochastic Processes.* Academic Press.

W.S. KENDALL and M. WESTCOTT (1987): One-dimensional classical scattering processes and the diffusion limit. *Adv. Appl. Probab.* **19**, 81-105.

M. LACEY and W. PHILIPP (1990): A note on the almost-sure central limit theorem. *Statist. Probab. Letters* **9**, 201-205.

T.-L. LAI (1974a): Summability methods for independent, identically distributed random variables. *Proc. Amer. Math. Soc.* **45**, 253-261.

T.-L. LAI (1974b): Limit theorems for delayed sums. *Ann. Probab.* **2**, 432-440.

T.-L. LAI and C.Z. WEI (1982): A law of the iterated logarithm for double arrays of independent random variables, with applications to regression and time-series models. *Ann. Probab.* **10**, 320-335.

P. LÉVY (1937): *Théorie de l'Addition des Variables Aléatoires.* Gauthier-Villars, Paris.

P. LÉVY (1939): Sur certaines processus stochastiques homogènes. *Compositio Math.* **7**, 283-339.

G.G. LORENTZ (1953): *Bernstein Polynomials.* University of Toronto Press.

I. MEILIJSON (1973): The average of the values of a function at random points. *Israel J. Math.* **15**, 193-203.

V.V. PETROV (1975): *Sums of Independent Random Variables.* Springer.

J.W. PITMAN and M. YOR (1986): Asymptotic properties of planar Brownian motion. *Ann. Prob.* **14**, 733-778.

L.C.G. ROGERS and D. WILLIAMS (1987): *Diffusions, Markov Processes and Martingales, Volume 2: Itô Calculus.* Wiley.

P. SCHATTE (1988): On strong versions of the central limit theorem. *Math. Nachrich- ten* **137**, 249-256.

U. STADTMÜLLER (1984): A note on the law of the iterated logarithm for weighted sums of random variables. *Ann. Statist.* **12**, 35-44.

A.J. STAM (1968): Laws of large numbers for functionals of random walks with positive drift. *Compositio Math.* **19**, 299-333.

C. STEIN (1986): Approximate computation of expectations. *Inst. Math. Statist. Lecture Notes* **7**.

L. TAKACS (1981): The arc-sine law of P. Lévy. *Contributions to Probability* (*Lukács Festschrift*, ed. Gani and Rohatgi), Academic Press, New York, 49-63.

D. WILLIAMS (1969): Markov properties of Brownian local time. *Bull. Amer. Math. Soc.* **75**, 1035-1036.

D. WILLIAMS (1979): *Diffusions, Markov Processes and Martingales, Volume 1: Foundations.* Wiley.

K. ZELLER and W. BEEKMANN (1970): *Theorie der Limitierungsverfahren.* Springer.

Department of Mathematics
Royal Holloway and Bedford New College
Egham Hill, Egham
Surrey TW20 0EX
England

School of Mathematical Sciences
Queen Mary & Westfield College
Mile End Road
London E1 4NS
England