Jakimovski Methods and Almost-Sure Convergence

N.H. Bingham and U. Stadtmüller

1. Introduction

The classical summability methods of Borel (B) and Euler (E(λ), $\lambda > 0$) play an important role in many areas of mathematics. For instance, in summability theory they are perhaps the most important methods other than the Cesàro (C_α) and Abel (A) methods, and two chapters of the classic book of Hardy (1949) are devoted to them. In probability, the distinction between methods of Cesàro-Abel and Euler-Borel type may be seen from the following two laws of large numbers, the first of which extends Kolmogorov's strong law.

THEOREM I. (Lai 1974) *For* $X, X_1, X_2, \ldots$ *independent and identically distributed, the following are equivalent:*
(i) $E|X| < \infty$ *and* $EX = \mu$,
(ii) $X_n \to \mu$ *a.s.* $(n \to \infty)$ (C_α) *for some (all)* $\alpha \geq 1$,
(iii) $X_n \to \mu$ *a.s.* $(n \to \infty)$ (A).

THEOREM II. (Chow 1973) *For* $X, X_1, X_2, \ldots$ *independent and identically distributed, the following are equivalent:*
(i) $E|X|^2 < \infty$ *and* $EX = \mu$,
(ii) $X_n \to \mu$ *a.s.* $(n \to \infty)$ $(E(\lambda))$ *for some (all)* $\lambda > 0$,
(iii) $X_n \to \mu$ *a.s.* $(n \to \infty)$ (B).

Other applications in probability arise through the technique of 'Poissonization', in accordance with Kac's dictum: if you can't solve the problem exactly, then randomise (Kesten 1986, p. 1109; cf. Kac 1949, Hammersley 1950 (pp. 219–224), 1972 (§§7,8), Hammersley et al. 1975, Pollard 1984, p. 117). There are also applications along these lines to combinatorial optimisation (Steele et al. 1987, §3; Steele 1989, §3).

Often the properties of the methods are governed by the fact that their weights — the Poisson and binomial distributions — being convolutions, obey the central limit theorem. Consequently, many such properties extend to matrix methods $A = (a_{nk})$, whose weights are also given by convolutions:

$$a_{nk} = P(S_n = k), \tag{1.1}$$

for (S_n) a random walk (see e.g. Bingham 1981, 1984). There, $S_n = \sum_1^n X_k$ is a sum of independent X_k, identically distributed (and $\mathbb{Z}$-valued). Another important case is that of X_k Bernoulli (0, 1-valued) but not necessarily identically distributed:

$$P(X_n = 1) = p_n, \quad P(X_n = 0) = q_n := 1 - p_n.$$

Writing $p_n = 1/(1 + d_n)$, $(d_n \geq 0)$, this leads to the method $A = (a_{nk})$ defined by

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

the Jakimovski method $[F, d_n]$ (Jakimovski 1959; Zeller and Beekmann 1970 (Ergänzungen, §70)). The motivating examples are:

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

(ii) $d_n = (n-1)/\lambda$, the *Karamata-Stirling method* $\mathrm{KS}(\lambda)$,

(Karamata 1935). Here

$$a_{nk} = \lambda^k S_{nk}/(\lambda)_n,$$

with $(\lambda)_n := \lambda(\lambda+1)\dots(\lambda+n-1)$ and (S_{nk}) the Stirling numbers of the first kind. The Bernoulli representation (1.1) enables both local and global central limit theory to be applied; see Bender (1973) for a perspicuous treatment. In particular, unimodality of Stirling numbers and other weights follows from this; for background see e.g. Hammersley 1951, 1952, 1972 (§§18, 19), Erdős 1953, Harper 1967, Lieb 1968, Bingham 1988.

Our aim here is to extend to Jakimovski methods the law of large numbers (Theorem II), and the corresponding analogue of the law of the iterated logarithm (Lai 1974). This complements the work of Bingham (1988), which gives a similar extension to the basic Tauberian theorem ('O-K-Satz'), due in the Euler case to Knopp in 1923 and in the Borel case to Schmidt in 1925 (Hardy 1949, Theorems 156, 241, 128). For further background on almost-sure convergence behaviour and summability methods, see e.g. Stout 1974 (Chap. 4), Bingham and Goldie 1988.

2. Results

THEOREM 1. *For $X, X_0, X_1, \dots$ independent and identically distributed random variables, and (d_n) as above, the following are equivalent:*

(i) $\mathrm{var}\, X < \infty, \ EX = m$,

(ii) $X_n \to m$ *a.s.* $(E(\lambda)$ *or* $B)$,

(iii) $X_n \to m$ *a.s.* $(KS(\lambda))$,

(iv) $\quad X_n \to m$ a.s. $[F, d_n]$.

In what follows, we restrict the generality slightly. We assume further that $[F, d_n]$ satisfies

$$p_n \to 0 \quad (\text{or } d_n \to \infty).$$

This ensures that $\sigma_n \asymp \sqrt{\mu_n}$ can be strengthened to

$$\sigma_n \sim \sqrt{\mu_n}.$$

The Euler case ($p_n = \lambda/(1+\lambda)$, $d_n = 1/\lambda$) is thereby excluded, but can be handled separately. These two cases together (p_n constant and $p_n \to 0$) cover the cases of main interest (though the result below and its proof may be extended to cover the case $\sigma_n \sim c\sqrt{\mu_n}$, for constant c). In (i) below, 'log' in the denominator means '$\max(1, \log_+)$'.

In Theorem 2, which gives the rates of convergence in Theorem 1, the Karamata-Stirling methods diverge from those of Euler and Borel, and one obtains an iterated logarithm, as in the classical case but unlike the Euler-Borel case (Lai 1974).

THEOREM 2. *The following are equivalent:*

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

(ii) $$\limsup_{x\to\infty} \frac{(4\pi x)^{1/4}}{\log^{1/2} x} \left| \sum_0^\infty e^{-x} \frac{x^k}{k!} X_k \right| = \sigma \quad \text{a.s.},$$

(iii) $$\limsup_{n\to\infty} \frac{(4\pi n)^{1/4}}{\log^{1/2} n} \left| \sum_0^n \binom{n}{k} \lambda^k X_k/(1+\lambda)^n \right| = \sigma(1+\lambda)^{1/4} \quad \text{a.s.},$$

(iv) $$\limsup_{n\to\infty} \frac{(4\pi\lambda \log n)^{1/4}}{\log\log^{1/2} n} \left| \sum_0^n a_{nk} X_k \right| = \sigma \quad \text{a.s.},$$

where $A = (a_{nk})$ is the matrix of the Kamarata-Stirling method $KS(\lambda)$,

(v) $$\limsup_{n\to\infty} \frac{(4\pi\mu_n)^{1/4}}{\log^{1/2} \mu_n} \left| \sum_0^n a_{nk} X_k \right| = \sigma \quad \text{a.s.}$$

where $A = (a_{nk})$ is the matrix of $[F, d_n]$ with $d_n \to \infty$.

Here the equivalence of (i) with (ii) ('LIL for the Borel method') and (iii) ('LIL for the Euler method') is Lai's result, and is included here for comparison. The constant $(1+\lambda)^{1/4}$ in (iii) is $a^{1/4}$, where a is the mean-variance ratio of the Euler method; see Bingham (1984) for a detailed discussion of this parameter and its significance. When $d_n \to \infty$, $\sigma_n \sim \sqrt{\mu_n}$, and $a = 1$.

Our proof of Theorem 2 will involve a non-uniform local limit theorem for the sums S_n in the Bernoulli representation $a_{nk} = P(S_n = k)$. Write

$H_3(x) := x^3 - 3x$ for the third Hermite polynomial, $\kappa_{3,n} := \mu^n_{3,0}$ for the third cumulant (third central moment) of S_n:

$$\kappa_{3,n} := \sum_1^n E[(\xi_j - p_j)^3] = \sum_1^n (p_j - 3p_j^2 + 2p_j^3).$$

Thus $\kappa_{3,n} \sim \sum_1^n p_j = \mu_n,\ (n \to \infty)$, when $p_n \to 0$.

Theorem 3. *For S_n the Bernoulli sum above, $a_{nk} = P(S_n = k)$,*

$$\sup_{k\in\mathbb{Z}} \left(1 + \left|\frac{k-\mu_n}{\sigma_n}\right|^3\right) \times \left|\sigma_n a_{nk} - \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{k-\mu_n}{\sigma_n}\right)^2\right\}\left(1 + H_3\left(\frac{k-\mu_n}{\sigma_n}\right)\frac{\kappa_{3,n}}{3!\sigma_n^3}\right)\right| = o(1/\sigma_n) \quad \textit{as } n \to \infty.$$

This result is closely related to Petrov's non-uniform local limit theorem. The 'uniform' part (taking the '1' term) is the Bernoulli case with $k = 3$ of Theorem 12 of Petrov (1975, VII.3), except that Petrov's condition

$$(*) \qquad \liminf_{n\to\infty} \sigma_n^2/n > 0$$

is violated when $d_n \to \infty$, as in Theorem 2 (iv), (v), since $\sigma_n^2 = \sum_1^n d_j/(1+d_j)^2$. However, to compensate for this, we know the characteristic function of our Bernoulli sum explicitly, and this enables us to handle the error terms in the Fourier analysis of Petrov's method successfully. The 'non-uniform' part (taking the '$|(k-\mu_n)/\sigma_n|^3$' term) is similarly related to Theorem 16 of Petrov (1975, VII.3), except that he has general identical distributions and we have Bernoulli non-identical distributions.

Theorem 3 involves the first term of an expansion of Edgeworth type ($k = 3$ in Petrov's notation). Extensions to Edgeworth expansions of arbitrary length (general k) are also possible, and can be proved by Petrov's method, adapted to our Bernoulli case as in the proof of Theorem 3 below. We shall return to this in Section 4.

3. Proofs

Proof of Theorem 1: We follow the argument of the proof of Theorem 1 of Bingham and Maejima (1985) — BM for short — indicating differences when these arise.

That (i) implies (ii) is Chow's result. Now if $d_n \geq \delta > 0$ for all large n, as assumed, $E(1/\delta) \subset [F, d_n]$ by a result of Meir (1963), Zeller and Beekmann (1970, Ergänzungen, §70); thus (ii) implies (iii) and (iv).

Conversely, the implication from (ii) to (i) is in BM. If (iii) or (iv) holds and $A = (a_{nk})$ denotes the relevant matrix method,

$$\sum a_{nk} X_k \to m \quad \text{a.s.}$$

Write X_k^s for the symmetrisation of X_k (difference of two independent copies of X_k):

$$\sum a_{nk} X_k^s \to 0 \quad \text{a.s.}$$

Split the sum into the sums over $k \leq \mu_n$ and $k > \mu_n$: Y_n and Z_n say. As in BM, $Y_n \to 0$ a.s. Split off the last term of Y_n: arguing as there,

$$a_{n,[\mu_n]} X^s_{[\mu_n]} \to 0 \quad \text{a.s.}$$

But (cf. Bingham 1988)

$$a_{n,[\mu_n]} \sim \frac{1}{\sigma_n \sqrt{2\pi}} \asymp \frac{1}{\sqrt{\mu_n}.\sqrt{2\pi}}$$

and hence

$$X^s_{[\mu_n]} / \sqrt{[\mu_n]} \to 0 \quad \text{a.s.} \quad (n \to \infty).$$

Write N for $[\mu_n]$:

$$X_N^s / \sqrt{N} \to 0 \quad \text{a.s.} \quad (N \to \infty).$$

From this, we obtain (i) as in BM. ■

Proof of Theorem 2: The argument follows that of Theorem 2 of BM with Petrov's non-uniform local limit theorem replaced by Theorem 3.

First, note that by a Borel-Cantelli argument, our moment condition in (i) is equivalent to

$$X_n = o(n^{1/4} \log^{1/2} n) \quad \text{a.s.}$$

We have, writing $\phi(x) := e^{-x^2/2}/\sqrt{2\pi}$,

$$\begin{aligned} &\sum a_{nk} X_k - \sum \phi\left(\frac{k-\mu_n}{\sigma_n}\right) X_k \\ &= \sum \phi\left(\frac{k-\mu_n}{\sigma_n}\right) H_3\left(\frac{k-\mu_n}{\sigma_n}\right) \frac{\kappa_{3,n}}{3!\sigma_n^3} X_k + \sigma_n^{-2} \sum \frac{o(1) X_k}{\left(1 + \left|\frac{k-\mu_n}{\sigma_n}\right|^3\right)}, \end{aligned}$$

the $o(1)$ being uniform in k. Call the two terms on the right the *Edgeworth term* and the *error term*. With probability one, we may replace X_k by $o(k^{1/4} \log^{1/2} k)$ in each. We may then estimate each by the methods of

BM, obtaining $o(\mu_n^{1/4}\log^{1/2}\mu_n)$ (a.s.) in each case. This enables us to reduce (v) (which contains (iv)) to

$$\text{(v}'\text{)} \qquad \limsup_n \frac{(4\pi\mu_n)^{1/4}}{\log^{1/2}\mu_n}\left|\sum \phi\left(\frac{k-\mu_n}{\sigma_n}\right)X_k\right| = \sigma \quad \text{a.s.}$$

This is substantially contained in the paper of Lai (1974), where he uses the result ('LIL for the Valiron method') to prove his results for the Borel and Euler methods (see particularly (16) and between (26) and (27)). Two new complications arise: (a) our mean $\mu_n \to \infty$ is not integer-valued, and (b) our variance $\sigma_n^2 \to \infty$ satisfies $\sigma_n^2 \sim \mu_n$ rather than $\sigma_n^2 = \mu_n$. However, our a.s. bound $X_k = o(k^{1/4}\log^{1/2}k)$ is exactly what is required to reduce our sums to Lai's, to the required accuracy $o(\mu_n^{-1/4}\log^{1/2}\mu_n)$. It suffices to show that

$$\text{(a}'\text{)} \quad \limsup_{\lambda\to\infty}\left\{\frac{\lambda^{1/4}}{\log^{1/2}\lambda}\sum_0^\infty o(k^{1/4}\log^{1/2}k)\times \right.$$
$$\left.\left|\frac{1}{\sqrt{2\pi\lambda}}\exp\left\{-\frac{(k-\lambda)^2}{2\lambda}\right\} - \frac{1}{\sqrt{2\pi[\lambda]}}\exp\left\{-\frac{(k-[\lambda])^2}{2[\lambda]}\right\}\right|\right\} = 0,$$

$$\text{(b}'\text{)} \quad \limsup_{\lambda\to\infty}\left\{\frac{\lambda^{1/4}}{\log^{1/2}\lambda}\sum_0^\infty o(k^{1/4}\log^{1/2}k)\times \right.$$
$$\left.\left|\frac{1}{\sqrt{2\pi\lambda}}\exp\left\{-(1+o(1))\frac{(k-\lambda)^2}{2\lambda}\right\} - \frac{1}{\sqrt{2\pi\lambda}}\exp\left\{-\frac{(k-\lambda)^2}{2\lambda}\right\}\right|\right\} = 0.$$

For (a′), note that if

$$f(\lambda) := \frac{1}{\sqrt{2\pi\lambda}}\exp\left\{-\frac{(k-\lambda)^2}{2\lambda}\right\}$$

then

$$f'(\lambda) = \frac{f(\lambda)}{\lambda}\left\{-\frac{1}{2} + (k-\lambda) + \frac{(k-\lambda)^2}{\lambda}\right\}.$$

Replace the difference $f(\lambda) - f([\lambda])$ by $(\lambda - [\lambda])f'(\lambda_k)$, where $[\lambda] \le \lambda_k \le \lambda$, which may be estimated by

$$\lambda^{-1}f(\lambda)\left\{\frac{1}{2} + |k-\lambda| + \frac{(k-\lambda)^2}{2\lambda}\right\}.$$

The first term is negligible with respect to $f(\lambda)$. For the second, we have to show

$$\frac{\lambda^{1/4}}{\log^{1/2}\lambda}\sum_0^\infty o(k^{1/4}\log^{1/2}k)\frac{|k-\lambda|}{\lambda}\frac{1}{\sqrt{2\pi\lambda}}\exp\left\{-\frac{(k-\lambda)^2}{2\lambda}\right\} \to 0 \quad \text{as } \lambda\to\infty,$$

or

$$\frac{1}{\lambda^{1/4}\log^{1/2}\lambda}\int_0^\infty o(y^{1/4}\log^{1/2}y)\frac{|y-\lambda|}{\sqrt{\lambda}}\frac{1}{\sqrt{2\pi\lambda}}\exp\left\{-\frac{(y-\lambda)^2}{2\lambda}\right\}dy \to 0 \quad \text{as } \lambda\to\infty.$$

Write $(y-\lambda)/\sqrt{\lambda}=t$: thus

$$y^{1/4}=\lambda^{1/4}(1+t/\sqrt{\lambda})^{1/4},\quad \log^{1/2}y=\log^{1/2}\lambda\left(1+\frac{\log(1+t/\sqrt{\lambda})}{\log\lambda}\right)^{1/2}.$$

It remains to consider

$$\int o\left((1+t/\sqrt{\lambda})^{1/4}\left\{1+\frac{\log(1+t/\sqrt{\lambda})}{\log\lambda}\right\}^{1/2}\right)|t|e^{-t^2/2}dt,$$

which tends to 0 as $\lambda\to\infty$, as required. The remaining $((k-\lambda)^2/\lambda)$ term is handled in the same way. Finally, (b′) follows similarly. (A similar analysis is given by Hardy and Littlewood 1916, Thm. 3.4 and Proof of Lemma 2.13.)

In the converse direction, that (iv) or (v) imply (i), follows as in the implication from (ii), (iii) to (i) (Lai 1974, p. 260; BM, p. 389). ∎

Proof of Theorem 3: We consider separately the '1' and '$|(k-\mu_n)/\sigma_n|^3$' terms; call the two parts A and B. Write $x_{k,n}$ for $(k-\mu_n)/\sigma_n$, $\phi_n,\phi_{n,0}$ for the characteristic functions of S_n, S_n-ES_n, c_n for $\kappa_{3,n}/(3!\sigma_n^3)\sim 1/(3!\sigma_n)$.

A:
$$a_{nk}=P(S_n=k)=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-itk}\phi_n(t)\,dt,$$

while for constant c

$$\phi(x)\{1+cH_3(x)\}=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-t^2/2}\{1+c(it)^3\}e^{-itx}\,dx.$$

So
$$2\pi\sigma_n a_{nk}=\int_{-\pi\sigma_n}^{\pi\sigma_n}\exp\{-itx_{k,n}\}\phi_{n,0}(t/\sigma_n)\,dt,$$

$$2\pi\sigma_n a_{nk} - \sqrt{2\pi}\exp\{-x_{k,n}^2\}\{1 + H_3(x_{k,n})c_n\}$$
$$= \int_{-\pi\sigma_n}^{\pi\sigma} \exp\{-itx_{k,n}\}\left(\phi_{n,0}(t/\sigma_n) - e^{-t^2/2}\{1 + (it)^3 c_n\}\right) dt$$
$$+ \int_{|t|\geq\pi\sigma_n} \exp\{-itx_{k,n}\}e^{-t^2/2}\{\dots\}\ dt,$$

$$|\dots| \leq \int_{-\pi\sigma_n}^{\pi\sigma_n} |\dots|\ dt + \int_{|t|\geq\pi\sigma_n} |\dots|\ dt = \mathrm{I} + \mathrm{II},\ \text{say}.$$

Expanding $\phi_{n,0}$ as far as the third cumulant, we find that for $|t| = o(\sigma_n)$ (actually $|t| = o(\sigma_n^{1/6})$ is all we need)

$$\phi_{n,0}(t/\sigma_n) = \exp\left\{-\frac{1}{2}t^2 + (it)^3 c_n + O(t^4\mu_n/\sigma_n^4)\right\}.$$

Now we choose $\epsilon_n \to 0$, and decompose I as the sum of integrals over $|t| \leq \epsilon_n\sigma_n^{1/6}$, $\epsilon_n\sigma_n^{1/6} \leq |t| \leq \sigma_n/4$ and $\sigma_n/4 \leq |t| \leq \pi\sigma_n$:

$$\mathrm{I} = \mathrm{I}_a + \mathrm{I}_b + \mathrm{I}_c, \quad \text{say}.$$

In I_a, $|t| = o(\sigma_n^{1/6})$, and the integrand may be checked to be $e^{-t^2/2}o(1/\sigma_n)$. Hence $\mathrm{I}_a = o(1/\sigma_n)$. For I_b, use Lemma 12 of Petrov (1975, p. 179) on the first term. The integrand is exponentially small in σ_n, hence ('normal tails') so is the integral when $\epsilon_n \to 0$ sufficiently slowly; similarly for the second term. For I_c, the $\{\cdots\}$ term is handled as with I_b. The other term is

$$\mathrm{I}_d \leq \sigma_n \int_{1/4\leq t\leq\pi} |\phi_n(t)|dt + \sigma_n \int_{1/4\leq t\leq\pi} \exp\{-\sigma_n^2 t^2\}(1 + |t|^3\sigma_n^2)dt.$$

By direct estimation,

$$\log|\phi_n(t)| \leq -\sum_1^n p_j(1-p_j)(1-\cos t) = -\sigma_n^2(1-\cos t) \leq -c\sigma_n^2$$

in the range of integration, for some $c > 0$, so the first term is exponentially small; clearly, so is the second. Thus $\mathrm{I} = o(1/\sigma_n)$.

For II, the '1' term in $\dots$ is exponentially small as above, while the 't^3' term is $o(1/\sigma_n)$ as $c_n \sim 1/(3!\sigma_n)$.

B: $$x_{k,n}^3 2\pi\sigma_n a_{nk} = x_{k,n}^3 \int_{-\pi\sigma_n}^{\pi\sigma_n} \exp\{-itx_{k,n}\}\phi_{n,0}(t/\sigma_n)\ dt.$$

Integrating by parts three times, the right is

$$i \int_{-\pi\sigma_n}^{\pi\sigma_n} \exp\{-itx_{k,n}\} D^3 \phi_{n,0}(t/\sigma_n)\ dt.$$

Also

$$\begin{aligned}
&\sqrt{2\pi} x_{k,n}^3 \exp\{-x_{k,n}^2\}(1 + H_3(x_{k,n})c_n) \\
&\qquad = x_{k,n}^3 \int_{-\infty}^{\infty} e^{-t^2/2}(1 + (it)^3 c_n) \exp\{-itx_{k,n}\}\ dt \\
&\qquad = i \int_{-\infty}^{\infty} \exp\{-itx_{k,n}\} D^3[e^{-t^2/2}(1 + (it)^3 c_n)]\ dt,
\end{aligned}$$

integrating by parts three times again.

Subtract, and estimate the difference as a sum of integrals over the interval $[-\pi\sigma_n, \pi\sigma_n]$ and its complement, I and II say, as before. Write (cf. Petrov 1975, p. 209)

$$g_n(t) := \log \phi_n(t/\sigma_n) - \frac{it\mu_n}{\sigma_n} + \frac{1}{2}t^2 - (it)^3 c_n.$$

Then

$$\begin{aligned}
\phi_{n,0}(t/\sigma_n) &= e^{-\frac{1}{2}t^2} \exp\{(it)^3 c_n\} \exp\{g_n(t)\} \\
&= e^{-\frac{1}{2}t^2} \big(1 + (it)^3 c_n + R_n(t)\big) \exp\{g_n(t)\}, \quad \text{say.}
\end{aligned}$$

Because we know $\phi_{n,0}$ explicitly, we can calculate the first three derivatives of $g_n, \exp\{g_n\}$ (and R_n) explicitly. We can then estimate I (splitting it up as before) and II, along the lines above. All remainders are power series, so may be differentiated term-wise. The exponential estimates obtained above are at worst multiplied by polynomials. The extra detail, which is tedious, is omitted. ∎

4. Remarks

1. In BM, an alternative proof of the LIL is given, using a 'weighted l^1 version' of the local limit theorem, due to Bikyalis and Jasjunas (1967). We raise here the question of obtaining a non-identically distributed version of this result, which would provide an alternative proof of Theorem 2.
2. In the special case

$$\sum_1^{\infty} \frac{1}{(1 + d_n)^2} < \infty$$

(which covers the Karamata-Stirling methods), a quite different proof of Theorem 2 may be given, using Poisson instead of normal approximation to reduce to Lai's result for the Borel case. We use Theorem 2 of Barbour (1987) with $l = p = 1$. In (3.15), W is the Bernoulli sum S_n, so (with $h(n) := X_n = o(n^{1/4} \log^{1/2} n)$ a.s.) $Eh(W)$ is the sum $\sum a_{nk} X_k$ to be approximated. In (2.7) with $l = 1$, $\int h \, dQ_1$ is the corresponding 'discrete Borel mean'

$$\sum_0^\infty e^{-\mu_n} \frac{\mu_n^k}{k!} X_k.$$

The error term (in view of Remark 3, p. 765) is $\nu_1/\sqrt{\lambda}$, where

$$\lambda = \mu_n = \sum_1^n \frac{1}{(1+d_j)},$$

$$\nu_1 = \sum_1^n \frac{1}{(1+d_j)^2} \quad (= \mu_n - \sigma_n^2).$$

By assumption, $\nu_1 = O(1)$, so this is $O(1/\sqrt{\lambda})$. Barbour's theorem tells us that the Jakimovski and discrete-Borel means differ by an amount of order $o\big((\lambda^{1/4} \log^{1/2} \lambda)/\sqrt{\lambda}\big) = o(\lambda^{-1/4} \log^{1/2} \lambda)$ (cf. BM, p. 389), which reduces Theorem 2 to the discrete-Borel case. We then use Lai's result for the Borel case, or rather its proof (Lai 1974, p. 258), with $M := [\lambda]$ replaced by $M := [\mu_n]$.

3. Central limit theorems have been given in this context by Embrechts and Maejima (1984), complementing our results on LLN and LIL. Note that their condition (6.1) holds —

$$\sum_k a_{nk}^2 \sim \frac{1}{\sqrt{2}} \sup_k a_{nk} \quad (n \to \infty),$$

which simplifies their Theorems 2 and 3. To see the above, write $\phi_{n,k}$ for $\phi(x_{k,n})$. Then

$$\sum a_{nk}^2 - \sum \phi_{n,k}^2 \le (\sup_k a_{nk} + \sup_k \phi_{n,k}) \sum |a_{nk} - \phi_{n,k}|.$$

The sum is $o(1)$ (Bingham 1988, Proposition, (iii)), while (*loc. cit.*, (ii)) each of the suprema has order $(\sigma_n \sqrt{2\pi})^{-1}$, so the right hand side is $o(1/\sigma_n)$. But

$$\sum \phi_{n,k}^2 \sim \frac{1}{2\sqrt{\pi \sigma_n}}$$

(Hardy 1949, Thm. 140).

4. Closely linked with the E(λ), B and KS(λ) methods considered here is

the Riesz mean $\mathrm{R}(e^{\sqrt{n}}, 1)$ (or 'moving average $M(\sqrt{n})$'; see Bingham and Goldie (1988). Here a functional (Strassen) version of the LIL is available; see de Acosta and Kuelbs (1983), Chan, Csörgő and Révész (1978).

For other LIL results for weighted means, see e.g. Bingham (1986, §15).

5. The Petrov condition (*), whose failure here necessitated our Theorem 3, guarantees that normal rather than Poisson approximation is appropriate. When it fails, as for $\mathrm{KS}(\lambda)$, we may use Poisson approximation as in Remark 2, and Lai's result. This hinges (Lai 1974, p. 258) on large-deviation results approximating Poisson to normal (Hardy 1949, p. 200). This suggests a direct use of large-deviation approximations to normality. Such results are known (Petrov 1975, p. 219, (2.5)), but again only under (*). Accordingly, we raise the question of obtaining large-deviation theorems (and non-uniform local limit theorems for general rather than Bernoulli distributions) when (*) is violated.

6. In Bingham (1984) results are obtained reducing convergence under a 'random-walk method' (a_{nk}) to Valiron convergence for sequences (s_n) of polynomial growth. Here one uses Petrov's non-uniform local limit theorem, with the number of Edgeworth terms retained depending on the degree of polynomial growth. The same method applies here, using the extension of Theorem 3 to general Edgeworth expansions mentioned in Section 2. Thus when

$$s_n = O(n^r) \quad \text{for some } r,$$

Theorem 1 there extends to give the equivalence of

$$\sum a_{nk} s_k \to s$$

and

$$\sum \phi_{n,k} s_k \to s.$$

References

Acosta, A. de and Kuelbs, J. (1983). Limit theorems for moving averages. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 64, 67–123.

Barbour, A.D. (1987). Asymptotic expansions in the Poisson limit theorem. *Annals of Probability* 15, 748–766.

Bender, E.A. (1973). Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory A* 15, 91–111.

Bikyalis, A. and Jasjunas, G. (1967). Limit theorems in the metric of the space L_1, l_1 (in Russian). *Litovskii Matematicheskii Sbornik* 7, 195–218.

Bingham, N.H. (1981). Tauberian theorems and the central limit theorem. *Annals of Probability* 9, 221–231.

——— (1984). On Valiron and circle convergence. *Mathematische Zeitschrift* 186, 273–286.

——— (1986). Variants on the law of the iterated logarithm. *Bulletin of the London Mathematical Society* 18, 433–467.

——— (1988). Tauberian theorems for Jakimovski and Kamarata-Stirling methods. *Mathematika* 35, 216–224.

Bingham, N.H. and Goldie, C.M. (1988). Riesz means and self-neglecting functions. *Mathematische Zeitschrift* 199, 443–454.

Bingham, N.H. and Maejima, M. (1985). Summability methods and almost-sure convergence. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 68, 383–392.

Chan, A.H.C., Csörgő, M., and Révész, P. (1978). Strassen type limit points for moving averages of a Wiener process. *Canadian Journal of Statistics* 6, 57–75.

Chow, Y.-S. (1973). Delayed sums and Borel summability of independent and identically distributed random variables. *Bulletin of the Institute of Mathematics, Academia Sinica,* 1, 207–220.

Embrechts, P. and Maejima, M. (1984). The central limit theorem for summability methods of i.i.d. random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 68, 191–204.

Erdős, P. (1953). On a conjecture of Hammersley. *Journal of the London Mathematical Society* 28, 232–236.

Hammersley, J.M. (1950). On estimating restricted parameters (with discussion). *Journal of the Royal Statistical Society*, Series B, 12, 192–240.

——— (1951). The sums of products of the natural numbers. *Proceedings of the London Mathematical Society* 1, 435–452.

——— (1952). Tauberian theory for the asymptotic forms of statistical frequency functions. *Proceedings of the Cambridge Philosophical Society* 48, 592–599.

——— (1972). A few seedlings of research. *Proceedings of the Sixth Berkeley Symposium for Mathematical Statistics and Probability*, Volume 1, 345–394, University of California Press.

Hammersley, J.M., Lewis, J.W.E., and Rowlinson, J.S. (1975). Relationships between the multinomial and Poisson models of stochastic processes, and between the canonical and grand canonical ensembles in statistical mechanics, with illustrations and Monte-Carlo methods for the penetrable sphere model of liquid-vapour equilibrium. *Sankhyā A* 37, 457–491.

Hardy, G.H. (1949). *Divergent Series.* Clarendon Press, Oxford.

Hardy, G.H. and Littlewood, J.E. (1916). Theorems concerning the summability of series by Borel's exponential method. *Rendiconti Circolo Matematico de Palermo* 41, 36–53. *Collected Works of G.H. Hardy*, Volume VI, 609–628, Clarendon Press, Oxford, 1974.

Harper, L.H. (1967). Stirling behaviour is asymptotically normal. *Annals of Mathematical Statistics* 38, 410–414.

Jakimovski, A. (1959). A generalisation of the Lototsky method. *Michigan Mathematical Journal* 6, 277–296.

Kac, M. (1949). On deviations between theoretical and empirical distributions. *Proceedings of the National Academy of Sciences, U.S.A.* 35, 252–257. In *Mark Kac: Probability, Number Theory and Statistical Physics, Selected Papers*, ed. K. Bacławski and M.D. Donsker, MIT Press, Cambridge, Massachusetts, 1979, 281-286.

Karamata, J. (1935). Théorèmes sur la sommabilité exponentielle et d'autres sommabilités s'y rattachant. *Mathematica (Cluj)* 9, 164–178.

Kesten, H. (1986). The influence of Mark Kac on probability theory. *Annals of Probability* 14, 1103–1128.

Lai, T. -L. (1974). Summability methods for independent, identically distributed random variables. *Proceedings of the American Mathematical Society* 45, 253–261.

Lieb, E.H. (1968). Convexity properties and a generating function for Stirling numbers. *Journal of Combinatorial Theory* 5, 203–206.

Meir, A. (1963). On the $[F, d_n]$-transformation of A. Jakimovski. *Bulletin of the Research Council of Israel* (F: Mathematics) 10, 165–187.

Petrov, V.V. (1975). *Sums of Independent Random Variables.* Springer-Verlag, Berlin.

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer-Verlag, Berlin.

Steele, J.M. (1989). Probabilistic and worst case analyses of classical problems of combinatorial optimization in Euclidean space. *Mathematics of Operations Research*, to appear.

Steele, J.M., Shepp, L.A., and Eddy, W.F. (1987). On the number of leaves of a Euclidean minimal spanning tree. *Journal of Applied Probability* 24, 809–826.

Stout, W.F. (1974). *Almost Sure Convergence.* Academic Press, New York.

Zeller, K. and Beekmann, W. (1970). *Theorie der Limitierungsverfahren.* Springer-Verlag, Berlin.

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

Universität Ulm
Abteilung Mathematik-III
Oberer Eselsberg
7900 Ulm.