The Annals of Probability
1984, Vol. 12, No. 2, 335–360

THE BOUNDED LAW OF THE ITERATED LOGARITHM FOR THE WEIGHTED EMPIRICAL DISTRIBUTION PROCESS IN THE NON-I.I.D. CASE

BY MICHAEL B. MARCUS AND JOEL ZINN

Texas A & M University

Using a simple symmetrization procedure, an upper bound is obtained for the probability distribution of various kinds of weighted empirical distribution processes where the underlying real valued random variables are not identically distributed. These probability bounds are used to obtain bounded laws of the iterated logarithm for empirical processes with different kinds of weighting. They are also used to obtain a one sided version of Daniel's theorem in the non-i.i.d. case.

0. Introduction. Using a simple symmetrization procedure for Banach space valued random variables, we obtain a number of strong limit theorems for the empirical distribution process. New proofs of some classical results, such as the exponential bound of Dvoretzky, Kiefer and Wolfowitz [3] are obtained as well as some new results that are valid when the independent random variables that occur in the empirical process are not identically distributed. In [1], Bretagnolle extends the Dvoretzky, Kiefer, Wolfowitz result to the non-i.i.d. case by showing that the identically distributed case is extremal and therefore gives an upper bound for an analogous statement involving independent random variables which are not identically distributed. We approach the problem of independent non-identically distributed random variables directly and hence we can obtain some results which have no analogues in the i.i.d. case.

In what follows $\{X_k\}_{k=1}^{\infty}$ will denote a sequence of independent, non-negative, real valued random variables which are not necessarily identically distributed, $\{\eta_k\}_{k=1}^{\infty}$ will denote a sequence of random variables with finite expectation such that the pairs $\{(\eta_k, X_k)\}_{k=1}^{\infty}$ are independent in k and $\{\varepsilon_k\}$ will denote a Rademacher sequence (i.e. a sequence of independent, symmetric random variables taking the values ± 1) independent of $\{(\eta_k, X_k)\}$. Also, for a real valued function $g(t)$, $t \geq 0$, let $\| g \| := \sup_{t \geq 0} | g(t) |$ and let $Lx = \max(1, \log x)$ and $L_2 x = L(Lx)$. We will now state some of the main results of this paper.

THEOREM 0.1. *Let* $\{c_n\}_{n=1}^{\infty}$ *be a sequence of real numbers and* $a_n := (\sum_{k=1}^{n} c_k^2 L_2(\sum_{k=1}^{n} c_k^2))^{1/2}$. *Let*

$$\mathscr{U}_n(t) = \sum_{k=1}^{n} c_k(I_{[X_k \geq t]} - P(X_k \geq t)). \tag{0.1}$$

Received November 1982.
AMS 1980 *subject classifications*. Primary, 60B12, 60F15; secondary, 62F25, 62F12.
Key words and phrases. Empirical distribution process, law of the iterated logarithm, Daniel's theorem.

Then

$$\limsup_{n\to\infty} \frac{\| \mathscr{U}_n \|}{a_n} < \infty \quad \text{a.s.} \tag{0.2}$$

In the case when $\{X_k\}$ are i.i.d., the compact version of this result, under some mild regularity conditions on the $\{c_n\}$, is due to Vanderzanden [12].

THEOREM 0.2. *Let $\psi(t)$, $t \geq 0$ be a non-negative, non-decreasing function and let $\{b_n\}$ be an increasing sequence of positive real numbers with $\lim_{n\to\infty} b_n = \infty$. Assume that*

$$\limsup_{n\to\infty} \frac{| \sum_{k=1}^n \varepsilon_k \psi(X_k) |}{b_n} < \lambda \quad \text{a.s.} \tag{0.3}$$

for some $\lambda > 0$. Furthermore, assume that

$$\limsup_{n\to\infty} \sup_{\psi(t) > 2\lambda b_n} b_n^{-1} \psi(t) \textstyle\sum_{k=1}^n P(X_k \geq t) < \beta \tag{0.4}$$

for some $\beta > 0$. Then

$$\limsup_{n\to\infty} \frac{\| \psi(t) \sum_{k=1}^n (I_{[X_k \geq t]} - P(X_k \geq t)) \|}{b_n} < 1120\lambda + 20\beta \quad \text{a.s.} \tag{0.5}$$

Theorems 0.1 and 0.2 are actually special cases of Theorem 0.3.

THEOREM 0.3. *Let $\psi(t)$ be as in Theorem 0.2. Let*

$$Z_k(t) = \psi(t)(\eta_k I_{[X_k \geq t]} - E\eta_k I_{[X_k \geq t]}), \quad S_n = \textstyle\sum_{k=1}^n \varepsilon_k \eta_k \psi(X_k)$$

and

$$\mathscr{S}_n(t) = \textstyle\sum_{k=1}^n Z_k(t).$$

Let $\{b_n\}$ be an increasing sequence of positive numbers with $\lim_{n\to\infty} b_n = \infty$ and assume that

$$\limsup_{n\to\infty} \frac{| S_n |}{b_n} < \lambda \quad \text{a.s.} \tag{0.6}$$

for some $\lambda > 0$. Furthermore, assume that for some $\beta < \infty$

$$\limsup_{n\to\infty} \sup_{1 \leq j \leq n} \sup_{t > 0} b_n^{-1} \psi(t) \,|\, \textstyle\sum_{k=1}^j E\eta_k I_{[X_k \geq t]} I_{[|\eta_k \psi(X_k)| > 2\lambda b_n]} \,|\, < \beta/2. \tag{0.7}$$

Then

$$\limsup_{n\to\infty} \frac{\| \mathscr{S}_n \|}{b_n} < 1120\lambda + 20\beta \quad \text{a.s.} \tag{0.8}$$

Essentially Theorem 0.2 and 0.3 state that whenever the partial sums of the symmetrization of the random variables $\{\eta_k \psi(X_k)\}_{k=1}^\infty$ satisfies a bounded law of the iterated logarithm with respect to $\{b_n\}$ then so does the weighted empirical process formed with $\{\eta_k\}$, $\{X_k\}$ and ψ. As is well known, there are many theorems on the bounded law of the iterated logarithm for real valued random variables, see e.g. Theorems 7.5.1 [2] and Theorem 1 [11]. Each of these yields a result about an appropriate weighted empirical process. Of course we also require that (0.4) or (0.7) holds but, as we point out in Remark 3.1, they are, in some sense, necessary conditions for (0.5) or (0.8). In applying Theorem 7.3.1. [2] and Theorem 1 [11] to our Theorem 0.2 or 0.3 it is very easy to see that (0.5) or respectively, (0.7) is satisfied. We shall go into greater detail on this point in Section 3. When the $\{X_k\}$ are i.i.d., Theorem 0.2 is subsumed by a result of James [6]. We shall discuss this further in Remark 4.4.

In Section 1 we obtain bounds on the probability distribution of weighted empirical processes. Lemma 1.1 is a generalization of the Dvoretzky, Kiefer, Wolfowitz result. Lemma 1.7 is our main result. It is used to prove Theorem 0.3 in Section 3. Further applications of Lemma 1.7 to Daniel's Theorem in the non-i.i.d. case and to generalize some inequalities of Mason [9] are given in Section 4. In Section 2 we slightly extend a useful result of Nagaev and Volodin [10] which simplifies the proving of strong laws. A necessary condition for the bounded law of the iterated logarithm for weighted empirical processes is also given in Section 2. Section 3 is devoted to proofs of Theorems 0.1, 0.2 and 0.3 and to some examples of when Theorem 0.2 and 0.3 hold. Section 4 is concerned with applications of Lemma 1.7. In Section 5 we give a new proof of an upper bound for Daniel's Theorem in the non-i.i.d. case which is somewhat sharper than some recent results of van Zuijlen [13], [14], [15].

Our approach to many of the problems considered in this paper was used earlier in [8] in the study of weak ℓ^p norms for sequences of independent random variables. It was in [8] that the relationship between probability estimates for these norms and Daniel's Theorem was first recognized. We would like to thank Gilles Pisier for helpful discussions about these topics.

1. Some inequalities for the probability distribution of $\| \mathscr{U}_n \|$ and $\| \mathscr{S}_n \|$. The main result of this section is Lemma 1.7, an inequality for the probability distribution of $\mathscr{S}_n$ (defined in Theorem 0.3). This is also valid for $\mathscr{U}_n$ (see (0.1)) since $\mathscr{U}_n$ is a special case of $\mathscr{S}_n$. Nevertheless we begin this Section by obtaining an exponential inequality for $\| \mathscr{U}_n \|$ itself. This is because our proof in this case is quite elementary and could even be used as a simple proof of the exponential inequality of Dvoretzky, Kiefer and Wolfowitz (although we don't obtain the best constant).

LEMMA 1.1. *For $\mathscr{U}_n(t)$, $t \geq 0$ as defined in Theorem 0.1 and for all $\lambda \geq 0$*

$$P\left(\frac{\| \mathscr{U}_n \|}{(\sum_{k=1}^n c_k^2)^{1/2}} > \lambda\right) \leq \exp\left(\frac{-\lambda^2}{4}\right)\left[1 + 2\sqrt{2\pi}\lambda \exp\left(\frac{\lambda^2}{8}\right)\right] \leq [1 + 2\sqrt{2\pi}\lambda]\exp\left(\frac{-\lambda^2}{8}\right). \tag{1.1}$$

Proof. We use a symmetrization technique that has been used before by various authors (see e.g. Lemma 2 [7]). Let Z be a random variable with values in some complete separable linear metric space B, let $|||\ |||$ be a measurable norm or pseudo-norm on B and let Z' be an independent copy of Z. Let E_Z ($E_{Z'}$) denote expectation with respect to the measure induced by Z (Z') and assume that φ, $\varphi(0) \geq 0$, is a non-decreasing convex function. Then by Jensen's inequality and the fact that $E \| Z \| \geq \| EZ \|$ we get

$$\begin{aligned} E\varphi(|||Z - Z'|||) &= E_Z E_{Z'}\varphi(|||Z - Z'|||) \\ &\geq E_Z\varphi(|||Z - EZ'|||) = E\varphi(|||Z - EZ|||). \end{aligned} \tag{1.2}$$

Now let $\{Y_k\}$ be a sequence of independent random variables with values in B and let $\{Y_k'\}$ be an independent copy of $\{Y_k\}$. Set $Z = \sum_{k=1}^n Y_k$ and $Z' = \sum_{k=1}^n Y_k'$. By (1.2)

$$\begin{aligned} E\varphi(|||\textstyle\sum_{k=1}^n (Y_k - EY_k)|||) &\leq E\varphi(|||\textstyle\sum_{k=1}^n (Y_k - Y_k')|||) \\ &= E\varphi(|||\textstyle\sum_{k=1}^n \varepsilon_k(Y_k - Y_k')|||) \\ &\leq E\varphi(|||\textstyle\sum_{k=1}^n \varepsilon_k Y_k||| + |||\textstyle\sum_{k=1}^n \varepsilon_k Y_k'|||) \end{aligned} \tag{1.3}$$

where $\{\varepsilon_k\}$ is a Rademacher sequence independent of $\{Y_k\}$ and $\{Y_k'\}$.

To obtain (1.1) it is clearly enough to assume that $\sum_{k=1}^n c_k^2 = 1$. We use (1.3) with $Y_k = c_k I_{[X_k \geq t]}$, $Y_k' = c_k I_{[X_k' \geq t]}$ (where $\{X_k'\}$ is an independent copy of $\{X_k\}$) and $\varphi(x) = e^{\beta x}$ and get

$$E \exp\{\beta \| \mathscr{U}_n \|\} \leq E \exp\{\beta \| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{[X_k \geq t]} \| + \beta \| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{[X_k' \geq t]} \|\}. \tag{1.4}$$

Let us be more explicit about the probability space in (1.4). We can assume that $\{X_k\}$, $\{X_k'\}$ and $\{\varepsilon_k\}$ are defined, respectively on the probability spaces $(\Omega_X, \mathscr{F}_X, P_X)$, $(\Omega_{X'}, \mathscr{F}_{X'}, P_{X'})$ and $(\Omega_\varepsilon, \mathscr{F}_\varepsilon, P_\varepsilon)$ with corresponding expectation operators E_X, $E_{X'}$ and E_ε. The random variables in (1.4) can be defined on the product probability space $(\Omega_X \times \Omega_{X'} \times \Omega_\varepsilon, \mathscr{F}_X \times \mathscr{F}_{X'} \times \mathscr{F}_\varepsilon, P_X \times P_{X'} \times P_\varepsilon)$. By Fubini's Theorem we can write the right side of (1.4) as

$$\begin{aligned} &E_X E_{X'} E_\varepsilon \exp\{\beta \| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{[X_k \geq t]} \| + \beta \| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{[X'_k \geq t]} \|\} \\ &\leq E_X[E_\varepsilon \exp\{2\beta \| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{[X_k \geq t]} \|\}]^{1/2} E_{X'}[E_\varepsilon \exp\{2\beta \| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{[X_k' \geq t]} \|\}]^{1/2} \\ &= E_X E_\varepsilon \exp\{2\beta \| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{X_k \geq t]} \|\}, \end{aligned} \tag{1.5}$$

where we use the Schwarz inquality twice and the fact that $\{X_k'\}$ is an independent copy of $\{X_k\}$. Let us now fix $\omega \in \Omega_X$ and consider

$$E_\varepsilon \exp\{2\beta \| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{[X_k(\omega) \geq t]} \|\}. \tag{1.6}$$

Let $\{X_{\pi(k,\omega)}\}_{k=1,\cdots,n}$ be a non-increasing rearrangement of $\{X_k(\omega)\}_{k=1,\cdots,n}$. Then, clearly,

$$\| \textstyle\sum_{k=1}^n \varepsilon_k c_k I_{[X_k(\omega) \geq t]} \| = \sup_{1 \leq j \leq n} | \textstyle\sum_{k=1}^j \varepsilon_{\pi(k,\omega)} c_{\pi(k,\omega)} |. \tag{1.7}$$

Since $\{\varepsilon_k\}$ and $\{X_k\}$ are independent, $\{\varepsilon_{\pi(k,\omega)}\}$ is a Rademacher sequence, i.e., a sequence of independent symmetric random variables taking on the values ± 1.

Thus by Lévy's inequality, for $u \geq 0$,

$$(1.8)\quad \begin{aligned} &P_\varepsilon[\sup_{1\leq j\leq n} | \textstyle\sum_{k=1}^{j} \varepsilon_{\pi(k,\omega)} c_{\pi(k,\omega)} | > u] \\ &\qquad \leq 2P_\varepsilon[| \textstyle\sum_{k=1}^{n} \varepsilon_{\pi(k,\omega)} c_{\pi(k,\omega)} | > u] \leq 4 \exp(-u^2/2), \end{aligned}$$

where, at the last step, we use the well-known sub-Gaussian inequality (see e.g. Lemma 5.2, Chapter 2 [5]), and the fact that $\sum_{k=1}^{n} c_k^2 = 1$. It follows from (1.7) and (1.8) that (1.6)

$$(1.9)\qquad \leq 1 + \exp(2\beta^2) \int_0^\infty 8\beta \exp\left(-\frac{(x-2\beta)^2}{2}\right) dx = G(\beta).$$

By Chebyschev's inequality

$$P(\| \mathscr{U}_n \| > \lambda) \leq e^{-\beta\lambda} E \exp\{\beta \| \mathscr{U}_n \| \}$$

and by (1.4), (1.5), (1.6) and (1.9) this is

$$\leq e^{-\beta\lambda} G(\beta).$$

Letting $\beta = \lambda/4$ we get (1.1).

REMARK 1.2. When $c_k = 1$; $k = 1, \cdots, n$ and the $\{X_k\}$ are i.i.d., (1.2) gives the exponential bound of Dvoretzky, Kiefer and Wolfowitz [3] and when the $\{X_k\}$ are not identically distributed we get the extension due to Bretagnolle [1]. However, we do not get the best constant in the exponent, i.e., the upper bound in [1] and [3] is $C \exp(-2\lambda^2)$ for some constant C. Nevertheless the proof of Lemma 1.1 should still be of interest because it is completely trivial using nothing more than the Jensen and Schwarz inequalities. Also (1.7) shows exactly why these exponential bounds are not a function of the specific sequence $\{X_k\}$. As far as we know, for arbitrary $\{c_k\}$, Lemma 1.1 is a new result.

In order to obtain a bound for the probability distribution of $\mathscr{S}_n$ (defined in Theorem 0.3) we use a relation which appears in the middle of the proof of Lemma 3.4 [8]. For completeness we will include the relation here as Corollary 1.4. It follows immediately from the following lemma which is itself a simple consequence of summation by parts.

LEMMA 1.3 *Let $\{a_k\}$ and $\{b_k\}$ be two sequences of non-negative numbers with $b_k \geq a_k$ and let $\{\alpha_k\}_{k=1}^{n}$ be a sequence of real numbers*

(i) *If a_k/b_k is non-increasing, then*

$$\sup_{1\leq j\leq n} | \textstyle\sum_{k=1}^{j} \alpha_k a_k | \leq \sup_{1\leq j\leq n} | \sum_{k=1}^{j} \alpha_k b_k |.$$

(ii) *If a_k/b_k is non-decreasing, then*

$$\sup_{1\leq j\leq n} | \textstyle\sum_{k=1}^{j} \alpha_k a_k | \leq 2 \sup_{1\leq j\leq n} | \sum_{k=1}^{j} \alpha_k b_k |.$$

(iii) *Let $\{\theta_k\}$ be a sequence of independent symmetric real valued random variables. Then for all $\lambda \geq 0$,*

$$P(| \textstyle\sum_{k=1}^{n} \theta_k a_k | \geq \lambda) \leq 2P(| \sum_{k=1}^{n} \theta_k b_k | \geq \lambda).$$

PROOF. Let $T_j = \sum_{k=1}^{j} \alpha_k b_k$, $T_0 = 0$. Then

$$\sum_{k=1}^{j} \alpha_k a_k = \sum_{k=1}^{j} \frac{a_k(T_k - T_{k-1})}{b_k} = \sum_{k=1}^{j} \frac{a_k T_k}{b_k} - \sum_{k=1}^{j-1} \frac{a_{k+1}T_k}{b_{k+1}}$$

$$= \frac{a_j T_j}{b_j} + \sum_{k=1}^{j-1} T_k\left(\frac{a_k}{b_k} - \frac{a_{k+1}}{b_{k+1}}\right).$$

In (i), $(a_k/b_k) - (a_{k+1}/b_{k+1}) \geq 0$. Hence

$$|\sum_{k=1}^{j} \alpha_k a_k| \leq \sup_{1\leq \ell \leq j} |T_\ell| \left[\frac{a_j}{b_j} + \sum_{k=1}^{j-1}\left(\frac{a_k}{b_k} - \frac{a_{k+1}}{b_{k+1}}\right)\right]$$

$$= \frac{a_1}{b_1} \sup_{1\leq \ell \leq j} |T_\ell| \leq \sup_{1\leq \ell \leq n} |T_\ell|.$$

In (ii), $(a_k/b_k) - (a_{k+1}/b_{k+1}) \leq 0$. Therefore

$$|\sum_{k=1}^{j} \alpha_k a_k| \leq \sup_{1\leq \ell \leq j} |T_\ell| \left[\frac{a_j}{b_j} + \sum_{k=1}^{j-1}\left(\frac{a_{k+1}}{b_{k+1}} - \frac{a_k}{b_k}\right)\right]$$

$$= \sup_{1\leq \ell \leq j} |T_\ell| \left[\frac{a_j}{b_j} + \left(\frac{a_j}{b_j} - \frac{a_1}{b_1}\right)\right] \leq 2 \sup_{1\leq \ell \leq n} |T_\ell|.$$

For (iii) choose a permutation $\{\pi(k)\}_{k=1}^{n}$ on $[1, \cdots, n]$ such that $a_{\pi(k)}/b_{\pi(k)}$ is non-increasing. Then, by (i)

$$|\sum_{k=1}^{n} \theta_{\pi(k)} a_{\pi(k)}| \leq \sup_{1\leq j\leq n} |\sum_{k=1}^{j} \theta_{\pi(k)} b_{\pi(k)}|$$

and consequently

$$P(|\sum_{k=1}^{n} \theta_k a_k| \geq \lambda) = P(|\sum_{k=1}^{n} \theta_{\pi(k)} a_{\pi(k)}| \geq \lambda)$$

$$\leq P(\sup_{1\leq j\leq n} |\sum_{k=1}^{j} \theta_{\pi(k)} b_{\pi(k)}| \geq \lambda).$$

By Lévy's inequality, this is

$$\leq 2P(|\sum_{k=1}^{n} \theta_{\pi(k)} b_{\pi(k)}| \geq \lambda) \leq 2P(|\sum_{k=1}^{n} \theta_k b_k| \geq \lambda).$$

COROLLARY 1.4. *Let $\{Y_k\}_{k=1}^{n}$ be independent, non-negative random variables defined on some probability space $(\Omega, \mathscr{F}, P)$, $\omega \in \Omega$. Let $\pi(k) = \pi(k, \omega)$, $k = 1, \cdots, n$ be a random permutation on $[1, \cdots, n]$ so that $\{Y_{\pi(k)}\}_{k=1}^{n}$ is a non-increasing rearrangement of $\{Y_k\}$. Let $\{\alpha_k\}_{k=1}^{n}$ be real numbers. Then*

$$\sup_{1\leq j\leq n} |Y_{\pi(j)} \sum_{k=1}^{j} \alpha_{\pi(k)}| \leq 2 \sup_{1\leq j\leq n} |\sum_{k=1}^{j} \alpha_{\pi(k)} Y_{\pi(k)}|. \tag{1.10}$$

PROOF. Clearly for $k \leq j$, $Y_{\pi(k)} \geq Y_{\pi(j)}$ and $Y_{\pi(j)}/Y_{\pi(k)}$ is non-decreasing. Therefore by Lemma 1.3 (ii)

$$Y_{\pi(j)} |\sum_{k=1}^{j} \alpha_{\pi(k)}| \leq 2 |\sum_{k=1}^{j} \alpha_{\pi(k)} Y_{\pi(k)}| \leq 2 \sup_{1\leq j\leq n} |\sum_{k=1}^{j} \alpha_{\pi(k)} Y_{\pi(k)}|.$$

The inequality in (1.10) follows immediately.

We now obtain an upper bound for the probability distribution of $\mathscr{S}_n$.

LEMMA 1.5. *Let* $\{X_k\}$, $\{\eta_k\}$, S_n, $\mathscr{S}_n$ *and* $\psi(t)$ *be as in Theorem* 0.3. *Assume that* $64\delta^{-2}\sum_{k=1}^n E\eta_k^2\psi^2(X_k) < 1$. *Then for* λ, $\delta > 0$

$$P[\sup_{1\le j\le n}\|\mathscr{S}_j\| > \lambda+\delta] \le \frac{8P[|S_n| > (\lambda/4)]}{1-64\delta^{-2}\sum_{k=1}^n E\eta_k^2\psi^2(X_k)}. \tag{1.11}$$

PROOF. Let Z, Z' and $|||\ |||$ be as in Lemma 1.1. Then

$$P(|||Z||| > \lambda+\delta)P(|||Z||| < \delta) \le P(|||Z-Z'||| > \lambda). \tag{1.12}$$

Therefore, taking $Z = \{\mathscr{S}_j\}_{j=1}^n$ and $|||Z||| = \sup_{1\le j\le n}\|\mathscr{S}_j\|$ in (1.12) we get

$$P(\sup_{1\le j\le n}\|\mathscr{S}_j\| > \lambda+\delta) \le \frac{P(\sup_{1\le j\le n}\sup_{t>0}|\psi(t)\sum_{k=1}^j(\eta_k I_{[X_k\ge t]}-\eta_k' I_{[X_k'\ge t]})| > \lambda)}{1-P(\sup_{1\le j\le n}\|\mathscr{S}_j\| > \delta)}. \tag{1.13}$$

The term in the numerator on the right in (1.13) is clearly equal to

$$\begin{aligned} &P(\sup_{1\le j\le n}\sup_{t>0}|\psi(t)\sum_{k=1}^j \varepsilon_k(\eta_k I_{[X_k\ge t]}-\eta_k' I_{[X_k'\ge t]})| > \lambda) \\ &\qquad \le 2P(\sup_{1\le j\le n}\sup_{t>0}|\psi(t)\sum_{k=1}^j \varepsilon_k\eta_k I_{[X_k\ge t]}| > \lambda/2) \end{aligned} \tag{1.14}$$

which by Lévy's inequality

$$\le 4P(\sup_{t>0}|\psi(t)\sum_{k=1}^n \varepsilon_k\eta_k I_{[X_k\ge t]}| > \lambda/2). \tag{1.15}$$

Note that in the above we take $\{\eta_k', X_k'\}$ to be independent copies of $\{\eta_k, X_k\}$ just as we did in the proof of Lemma 1.1.

Let $\{X_{\pi(k)}\}_{k=1}^n$ be a non-increasing rearrangement of $\{X_k\}_{k=1}^n$. Then since ψ is non-decreasing

$$\begin{aligned} \sup_{t>0}|\psi(t)\sum_{k=1}^n \varepsilon_k\eta_k I_{[X_k\ge t]}| &\le \sup_{1\le j\le n}|\psi(X_{\pi(j)})\sum_{k=1}^j \varepsilon_{\pi(k)}\eta_{\pi(k)}| \\ &\le 2\sup_{1\le j\le n}|\sum_{k=1}^j \varepsilon_{\pi(k)}\eta_{\pi(k)}\psi(X_{\pi(k)})| \end{aligned} \tag{1.16}$$

by (1.10). By (1.16), (1.15) is

$$\le 4P(\sup_{1\le j\le n}|\sum_{k=1}^j \varepsilon_{\pi(k)}\eta_{\pi(k)}\psi(X_{\pi(k)})| > \lambda/4),$$

which by Lévy's inequality

$$\le 8P(|\sum_{k=1}^n \varepsilon_{\pi(k)}\eta_{\pi(k)}\psi(X_{\pi(k)})| > \lambda/4) = 8P(|\sum_{k=1}^n \varepsilon_k\eta_k\psi(X_k)| > \lambda/4).$$

We next consider the denominator in (1.11). By Chebyshev's inequality

$$P(\sup_{1\le j\le n}\|\mathscr{S}_j\| > \delta) \le \delta^{-2}E(\sup_{1\le j\le n}\|\mathscr{S}_j\|^2). \tag{1.17}$$

By (1.3) with $\varphi(x) = x^2$ we get

$$\begin{aligned} E_\varepsilon(\sup_{1\le j\le n}\|\mathscr{S}_j\|^2) &\le 4E_\varepsilon(\sup_{1\le j\le n}\sup_{t>0}|\psi(t)\sum_{k=1}^j \varepsilon_k\eta_k I_{[X_k\ge t]}|^2) \\ &\le 8E_\varepsilon(\sup_{t>0}|\psi(t)\sum_{k=1}^n \varepsilon_k\eta_k I_{[X_k\ge t]}|^2) \end{aligned} \tag{1.18}$$

where, at the last step, we use Lévy's inequality. By (1.16) the last term in (1.18)

$$\le 32E_\varepsilon(\sup_{1\le j\le n}|\sum_{k=1}^j \varepsilon_{\pi(k)}\eta_{\pi(k)}\psi(X_{\pi(k)})|^2)$$

and finally, using Lévy's inequality yet again, this is

$$\le 64E_\varepsilon(|\textstyle\sum_{k=1}^n \varepsilon_k\eta_k\psi(X_k)|^2).$$

Taking expectation with respect to (η_k, X_k) and using (1.17) we get

$$P(\sup_{1\le j\le n}\|\mathscr{S}_j\| > \delta) \le 64\delta^{-2}E(|\textstyle\sum_{k=1}^n \varepsilon_k\eta_k\psi(X_k)|^2)$$

and this completes the proof of the lemma.

REMARK 1.6(a). If we take $\psi(t) = 1$ and $\eta_k = c_k$, $1 \le k \le n$ for c_k as in Lemma 1.1 then Lemma 1.5 gives, in analogy with (1.1),

$$P\left(\frac{\|\mathscr{U}_n\|}{(\sum_{k=1}^n c_k^2)^{1/2}} > \lambda\right) \le C\exp(-\alpha\lambda^2)$$

for constants C and α, although α will be quite a bit smaller than ⅛.

REMARK 1.6(b). Lemma 1.5 holds with $\sup_{j\le n}\|\mathscr{S}_j\|$ replaced by

$$\|\psi(\cdot)\textstyle\sum_{k=1}^\infty[\eta_k I_{[X_k\ge\cdot]} - E\eta_k I_{[X_k\ge\cdot]}]\|$$

and n replaced by ∞ wherever else it occurs in the statement of Lemma 1.5. This is because all the sums involved in Lemma 1.5 are Cauchy.

The next lemma is useful when $E\eta_k^2\psi^2(X_k)$ does not exist.

LEMMA 1.7. *Following the notation of Lemma* 1.5 *let* λ, δ, $\beta > 0$ *and let* γ *be such that*

$$64\delta^{-2}\textstyle\sum_{k=1}^n E\eta_k^2\psi^2(X_k)I_{[|\eta_k\psi(X_k)|\le\gamma]} < 1. \tag{1.19}$$

Furthermore, assume that

$$\sup_{1\le j\le n}\sup_{t>0}\psi(t)|\textstyle\sum_{k=1}^j E\eta_k I_{[X_k\ge t]}I_{[|\eta_k\psi(X_k)|>\gamma]}| \le \beta. \tag{1.20}$$

Then

$$\begin{aligned}&P[\sup_{1\le j\le n}\|\mathscr{S}_j\| > \lambda+\delta+2\beta]\\ &\le \frac{8P[|\sum_{k=1}^n \varepsilon_k\eta_k\psi(X_k)I_{[|\eta_k\psi(X_k)|\le\gamma]}| > \lambda/4]}{1 - 64\delta^{-2}\sum_{k=1}^n E\eta_k^2\psi^2(X_k)I_{[|\eta_k\psi(X_k)|\le\gamma]}} + P[\sup_{1\le k\le n}|\eta_k\psi(X_k)| > \gamma].\end{aligned} \tag{1.21}$$

PROOF. Let

$$\eta_k' = \eta_k I_{[|\eta_k\psi(X_k)|\le\gamma]},\quad \eta_k'' = \eta_k I_{[|\eta_k\psi(X_k)|>\gamma]}.$$

By the triangle inequality

$$\begin{aligned}&P[\sup_{1\le j\le n}\|\mathscr{S}_j\| > \lambda+\delta+2\beta]\\ &\le P[\sup_{1\le j\le n}\|\psi(t)\textstyle\sum_{k=1}^j(\eta_k' I_{[X_k\ge t]} - E\eta_k' I_{[X_k\ge t]})\| > \lambda+\delta]\\ &\quad + P[\sup_{1\le j\le n}\|\psi(t)\textstyle\sum_{k=1}^j(\eta_k'' I_{[X_k\ge t]} - E\eta_k'' I_{[X_k\ge t]})\| > 2\beta].\end{aligned} \tag{1.22}$$

We apply Lemma 1.5 to the first term on the right in (1.22) and see that it is

$$\leq \frac{8P[\,|\sum_{k=1}^{n} \varepsilon_k \eta_k \psi(X_k) I_{[|\eta_k\psi(X_k)|\leq\gamma]}\,| > (\lambda/4)]}{1 - 64\delta^{-2} \sum_{k=1}^{n} E\eta_k^2 \psi^2(X_k) I_{[|\eta_k\psi(X_k)|\leq\gamma]}}.$$

Now since

$$\sup_{1\leq j\leq n} \sup_{t>0} \psi(t)\, |\sum_{k=1}^{j} E\eta_k'' I_{[X_k\geq t]}| \leq \beta$$

by hypothesis, the second term in (1.22) is bounded by $P[\sup_k |\eta_k''| > 0]$ and this is the last term in (1.21).

REMARK 1.8. If $\sup_k |\eta_k \psi(X_k)| \leq \gamma$ then (1.21) is the same as (1.11). If not, we must deal with the condition (1.20). In some sense this is a necessary condition (see Section 2). In the special case that $\sup_k |\eta_k| \leq 1$ we can replace (1.20) by

$$\sup_{\psi(t)>\gamma} \psi(t) \sum_{k=1}^{n} P[X_k \geq t] \leq \beta. \tag{1.23}$$

To see this note that in this case (1.20)

$$\leq \max\{\sup_{\psi(t)\leq\gamma} \psi(t) \sum_{k=1}^{n} EI_{[X_k\geq t]} I_{[\psi(X_k)>\gamma]},\ \sup_{\psi(t)>\gamma} \psi(t) \sum_{k=1}^{n} P[X_k \geq t]\}. \tag{1.24}$$

Let $t_0 = \sup\{t: \psi(t) \leq \gamma\}$. If $\psi(t_0) > \gamma$ then

$$\begin{aligned} &\sup_{\psi(t)\leq\gamma} \psi(t) \sum_{k=1}^{n} EI_{[X_k\geq t]} I_{[\psi(X_k)>\gamma]} \\ &\qquad \leq \gamma \sum_{k=1}^{n} P[X_k \geq t_0] \leq \sup_{\psi(t)>\gamma} \psi(t) \sum_{k=1}^{n} P[X_k \geq t]. \end{aligned} \tag{1.25}$$

If $\psi(t_0) \leq \gamma$ then

$$\begin{aligned} &\sup_{\psi(t)\leq\gamma} \psi(t) \sum_{k=1}^{n} EI_{[X_k\geq t]} I_{[\psi(X_k)>\gamma]} \\ &\qquad \leq \gamma \sum_{k=1}^{n} P[X_k > t_0] \leq \sup_{\psi(t)>\gamma} \psi(t) \sum_{k=1}^{n} P[X_k \geq t]. \end{aligned} \tag{1.26}$$

Using (1.24), (1.25) and (1.26) we see that (1.23) implies (1.20) when $\sup_k |\eta_k| \leq 1$.

We now consider the case in which $\eta_k = 1$ for all k and give a version of Lemma 1.7 involving infinite sums.

LEMMA 1.9. *Following the notation of Lemma 1.5 let $\eta_k = 1$ for all k and assume that $\sum_{k=1}^{\infty} P(X_k \geq t) < \infty$ for all $t > 0$. Let $\lambda, \delta, \beta > 0$ and let γ be such that*

$$64\delta^{-2} \sum_{k=1}^{\infty} E\psi^2(X_k) I_{[\psi(X_k)\leq\gamma]} < 1 \tag{1.27}$$

and

$$\sup_{\psi(t)>\gamma} \psi(t) \sum_{k=1}^{\infty} P(X_k \geq t) \leq \beta. \tag{1.28}$$

Furthermore, assume that for every $\varepsilon > 0$ there exists an $n(\varepsilon)$ such that for $n, m \geq n(\varepsilon)$

$$\sup_{\psi(t)>\gamma} \psi(t) \sum_{k=n}^{m} P(X_k \geq t) < \varepsilon. \tag{1.29}$$

Then

$$P[\sup_{t>0} | \psi(t) \textstyle\sum_{k=1}^{\infty} (I_{[X_k \geq t]} - P(X_k \geq t)) | > \lambda + \delta + 2\beta]$$

$$\leq \frac{8P[| \sum_{k=1}^{\infty} \varepsilon_k \psi(X_k) I_{[|\psi(X_k)| \leq \gamma]} | > (\lambda/4)]}{1 - 64\delta^{-2} \sum_{k=1}^{\infty} E\psi^2(X_k) I_{[|\psi(X_k)| \leq \gamma]}} + P[\sup_{1 \leq k < \infty} \psi(X_k) \geq \gamma]. \tag{1.30}$$

PROOF. The lemma is a corollary of Lemma 1.7 with (1.20) replaced by (1.23). For some $\varepsilon > 0$ let $\lambda = \delta = \beta = \varepsilon/4$ in Lemma 1.7. By (1.27) we can find an $n_1(\varepsilon)$ so that for $n, m \geq n_1(\varepsilon)$

$$\frac{1024}{\varepsilon^2} \textstyle\sum_{k=n}^{m} E\psi^2(X_k) I_{[\psi(X_k) \leq \gamma]} \leq \varepsilon.$$

It follows by Chebyshev's inequality that for $n, m \geq n_1(\varepsilon)$

$$P(| \textstyle\sum_{k=n}^{m} \varepsilon_k \psi(X_k) I_{[\psi(X_k) \leq \gamma]} | > (\varepsilon/16)) \leq (\varepsilon/4).$$

Of course by (1.29) we can find an $n_2(\varepsilon)$ such that for $n, m \geq n_2(\varepsilon)$

$$\sup_{\psi(t) > \gamma} \psi(t) \textstyle\sum_{k=n}^{m} P(X_k \geq t) \leq (\varepsilon/8).$$

Now note that there is nothing to prove unless $P[\sup_{1 \leq k < \infty} \psi(X_k) > \gamma] < 1$ and in this case we have

$$\textstyle\sum_{k=1}^{\infty} P(\psi(X_k) > \gamma) < \infty.$$

Therefore we can find an $n_3(\varepsilon)$ such that for $n, m \geq n_3(\varepsilon)$

$$P[\sup_{n \leq k \leq m} \psi(X_k) > \gamma] \leq \textstyle\sum_{k=n}^{m} P(\psi(X_k) > \gamma) \leq \varepsilon.$$

Using all these in Lemma 1.7 with (1.20) replaced by (1.23) we have that for n, $m \geq \sup_{1 \leq i \leq 3} n_i(\varepsilon)$,

$$P(\sup_{t>0} | \psi(t) \textstyle\sum_{k=n}^{m} (I_{[X_k \geq t]} - P(X_k \geq t)) | > \varepsilon) \leq 3\varepsilon/(1 - \varepsilon). \tag{1.31}$$

Therefore, since by (1.31) $\| \mathscr{S}_n \|$ is Cauchy in probability, we can use Lemma 1.7 with (1.20) replaced by (1.23) to get (1.30).

2. Necessary conditions for the law of the iterated logarithm. We begin with a construction and lemma of Nagaev and Volodin [10] which greatly simplifies the problem of obtaining iterated log laws. Given an increasing sequence $\{b_n\}$ fix $c > 1$ and consider the intervals $(0, c]$, $(c, c^2]$, $(c, c^3]$, $\cdots$. From these discard all intervals for which $\{b_n\} \cap (c^s, c^{s+1}] = \phi$ and relabel the remaining intervals $(c^{s_r}, c^{s_r+1}]$, $r = 1, 2, \cdots$ in such a way that $s_r < s_{r+1}$. (We will assume that $b_1 > c$ in order to avoid problems of notation.) Let $n_r = \sup\{n{:}b_n \in (c^{s_r}, c^{s_r+1}]\}$ and consider the subsequence $\{b_{n_r}\}$. It follows from the definition of $\{b_{n_r}\}$ that for $j < r$

$$\frac{b_{n_j}}{b_{n_r}} < c^{j-r+1}. \tag{2.1}$$

Given such a sequence $\{b_n\}$ and $c > 1$ we will refer to the corresponding subsequence of the integers $\{n_r\}$ as a N-V subsequence.

Our Lemma 2.1 was proved by Nagaev and Volodin on the real line. We present it in the setting of linear measurable spaces; however, except for one point at which we use the Ottavani-Skorohod inequality instead of Lévy's inequality our proof is identical to the proof of Nagaev and Volodin. $(B, \mathscr{B}, \|\cdot\|)$ is a linear measurable space, where B denotes a real vector space, $\mathscr{B}$ a σ-algebra of subsets of B and $\|\cdot\|$ a semi-norm on B if (i) addition, scalar multiplication and $\|\cdot\|$ are $\mathscr{B}$ measurable operations on B, (ii) there exists a subset F of the $\mathscr{B}$ measurable linear functionals on B such that

$$\|x\| = \sup_{f\in\mathscr{F}} |f(x)|, \quad x \in B$$

(property (ii) is not used in this paper). Note that $D[0, \infty)$ with the supremum norm is a linear measurable space. In this case $\mathscr{B}$ is the σ-field generated by evaluations.

LEMMA 2.1. *$(B, \mathscr{B}, \|\cdot\|)$ is a linear measurable space. Let $\{Y_k\}$ be a sequence of independent random variables with values in B, $S_n = \sum_{k=1}^n Y_k$ and $\{b_n\}$, $b_1 > 1$, an increasing sequence of positive numbers with $\lim_{n\to\infty} b_n = \infty$. Then, if*

$$\text{(2.2)} \qquad \limsup_{n\to\infty} \frac{\|S_n\|}{b_n} < a \quad \text{a.s.}$$

we have

$$\text{(2.3)} \qquad \lim_{n\to\infty} P(\|S_n\| > ab_n) = 0$$

and, for the N-V *subsequence $\{n_r\}$ based on $\{b_n\}$ and $c > 1$*

$$\text{(2.4)} \qquad \sum_{r=1}^\infty P(\|S_{n_r} - S_{n_{r-1}}\| > 2ab_{n_r}) < \infty.$$

Conversely, if (2.3) *and* (2.4) *are satisfied then*

$$\text{(2.5)} \qquad \limsup_{n\to\infty} \frac{\|S_n\|}{b_n} \le \left(4c + \frac{2}{c-1}\right)a \quad \text{a.s.}$$

Furthermore,

$$\text{(2.6)} \qquad \limsup_{n\to\infty} \frac{\|S_n\|}{b_n} = 0 \quad \text{a.s.}$$

if and only if (2.3) *and* (2.4) *hold for all $a > 0$.*

PROOF. That (2.2) implies (2.3) and (2.4) is completely elementary. Fix $c > 1$ and consider the corresponding N-V subsequence $\{n_r\}$. Since b_n is increasing

$$\frac{\|S_{n_r} - S_{n_{r-1}}\|}{b_{n_r}} \le \frac{\|S_{n_r}\|}{b_{n_r}} + \frac{\|S_{n_{r-1}}\|}{b_{n_{r-1}}}.$$

Hence

$$\text{(2.7)} \qquad \limsup_{r\to\infty} \frac{\|S_{n_r} - S_{n_{r-1}}\|}{b_{n_r}} \le 2 \limsup_{n\to\infty} \frac{\|S_n\|}{b_n} < 2a \quad \text{a.s.}$$

Since $\{(\|S_{n_r} - S_{n_{r-1}}\|)/b_{n_r}\}_{r=1}^\infty$ are independent random variables we see from (2.7) and the Borel-Cantelli Lemma that (2.4) is satisfied. Also, since

$$P\left(\frac{\| S_n \|}{b_n} > a\right) \leq P\left(\sup_{m \geq n} \frac{\| S_m \|}{b_m} > a\right)$$

we see from (2.2) that (2.3) holds.

We now show that (2.3) and (2.4) imply (2.2). Let

$$T_j = \frac{S_{n_j} - S_{n_{j-1}}}{b_{n_j}}.$$

By (2.4) there exists a random integer $j_0 < \infty$ a.s. such that $\| T_j \| \leq 2a$ for $j \geq j_0$. Therefore on a set of measure one

$$\frac{\| S_{n_r} \|}{b_{n_r}} \leq \frac{\| S_{n_{j_0-1}} \|}{b_{n_r}} + \sum_{j=j_0}^{r} \frac{b_{n_j}}{b_{n_r}} \| T_j \| \leq \frac{\| S_{n_{j_0-1}} \|}{b_{n_r}} + \frac{2a}{c} \sum_{j=j_0}^{r} (c^{-1})^{r-j}$$

by (2.1). It follows that

$$\limsup_{r \to \infty} \frac{\| S_{n_r} \|}{b_{n_r}} \leq \frac{2a}{c-1} \quad \text{a.s.} \tag{2.8}$$

We now use a norm version of the Ottaviani-Skorohod lemma. Let

$$U_r := \max_{n_{r-1} < n \leq n_r} \frac{\| S_n - S_{n_{r-1}} \|}{b_{n_r}}.$$

We have

$$P(U_r > 4a) \leq \frac{P(\| S_{n_r} - S_{n_{r-1}} \| > 2ab_{n_r})}{1 - \max_{n_{r-1} < n \leq n_r} P(\| S_{n_r} - S_n \| > 2ab_{n_r})}.$$

Clearly for $n_{r-1} < n \leq n_r$

$$P(\| S_{n_r} - S_n \| > 2ab_{n_r}) \leq P(\| S_{n_r} \| > ab_{n_r}) + P(\| S_n \| > ab_n).$$

Therefore, by (2.3)

$$\lim_{r \to \infty} \max_{n_{r-1} < n \leq n_r} P(\| S_{n_r} - S_n \| > 2ab_{n_r}) = 0$$

so we can find an r_0 such that for $r \geq r_0$

$$P(U_r > 4a) \leq 2P(\| S_{n_r} - S_{n_{r-1}} \| > 2ab_{n_r}).$$

It now follows from (2.4) that

$$\sum_r P(U_r > 4a) < \infty.$$

Therefore, for $n_{r-1} < n \leq n_r$ and r sufficiently large

$$\frac{\| S_n \|}{b_n} \leq \frac{b_{n_r}}{b_n} \frac{\| S_n - S_{n_{r-1}} \|}{b_{n_r}} + \frac{b_{n_{r-1}}}{b_n} \frac{\| S_{n_{r-1}} \|}{b_{n_{r-1}}}$$

(2.9)

$$\leq cU_r + \frac{\| S_{n_{r-1}} \|}{b_{n_{r-1}}} \leq \left(4c + \frac{2}{c-1}\right)a \quad \text{a.s.,}$$

where we use (2.8) and the fact that $\{b_{n_r}\}$ is an N-V subsequence.

We also see from (2.9) that if (2.3) and (2.4) hold for all $a > 0$ then we get (2.6), whereas (2.6) implies that (2.3) and (2.4) hold for all $a > 0$ since this is contained in the fact that (2.2) implies (2.3) and (2.4).

In the next theorem we obtain a necessary condition for the law of the iterated logarithm in the case $\eta_k = 1$ for all $k \geq 1$, i.e. in the setting of Theorem 0.2. This shows that in some sense (0.4) is a necessary condition for (0.5) (and similarly, (0.7) is a necessary condition for (0.8) in Theorem 0.3). Note that

$$\sup_{t>0} | \psi(t) \textstyle\sum_{k=1}^{n} (I_{[X_k \geq t]} - P(X_k \geq t)) |$$

(see (0.5)) is identically zero if each of the X_k is equal to a constant with probability 1. This degenerate case is ruled out by (2.10) below.

THEOREM 2.2. *Using the notation of Theorem* 0.2 *assume that*

(2.10) $$\limsup_{n\to\infty} \sup_{t>0} P[X_n = t] \leq 1 - \delta$$

for some $\delta > 0$ *and that*

(2.11) $$\limsup_{n\to\infty} b_n^{-1} \| \psi(t) \textstyle\sum_{k=1}^{n} (I_{[X_k \geq t]} - P(X_k \geq t)) \| \leq \lambda \quad \text{a.s.}$$

Then, it follows that

(2.12) $$\limsup_{n\to\infty} \frac{\psi(X_n)}{b_n} \leq \frac{4\lambda}{\delta} \quad \text{a.s.}$$

and for any $\rho > \delta^{-1}$

(2.13) $$\limsup_{n\to\infty} \sup_{\psi(t) > 4\lambda\rho b_n} b_n^{-1} \psi(t) \textstyle\sum_{k=1}^{n} P(X_k \geq t) \leq \lambda.$$

PROOF. Let

$$\mathscr{T}_n(t) = \psi(t) \textstyle\sum_{k=1}^{n} (I_{[X_k \geq t]} - P(X_k \geq t))$$

and $\mathscr{W}_n(t) = \mathscr{T}_n(t) - \mathscr{T}_{n-1}(t)$. Then, since

$$b_n^{-1} \| \mathscr{W}_n \| \leq b_n^{-1} \| \mathscr{T}_n \| + b_{n-1}^{-1} \| \mathscr{T}_{n-1} \|,$$

it follows from (2.11) that

(2.14) $$\limsup_{n\to\infty} b_n^{-1} \| \mathscr{W}_n \| \leq 2\lambda \quad \text{a.s.}$$

Note that

$$\begin{aligned}\| \mathscr{W}_n \| &= \sup_{t\le X_n} | \mathscr{W}_n(t) | \vee \sup_{t>X_n} | \mathscr{W}_n(t) | \\ &\ge \psi(X_n)[(1 - P(X_n \ge t)) \vee P(X_n > t)] \ge (\delta\psi(X_n)/2).\end{aligned} \tag{2.15}$$

Therefore, (2.12) follows from (2.14) and (2.15). We now obtain (2.13). By (2.12) for any $\rho > \delta^{-1}$, for a.a. ω there exists $n_0 = n_0(\omega)$ such that

$$\frac{\psi(X_n)}{b_n} \le 4\lambda\rho \quad \text{for all} \quad n \ge n_0(\omega).$$

Now choose $n_1 = n_1(\omega) \ge n_0(\omega)$ such that

$$\max_{k\le n_0} \frac{\psi(X_k)}{b_{n_1}} \le 4\lambda\rho.$$

Then, for $n \ge n_1$

$$\begin{aligned}\max_{k\le n} \frac{\psi(X_k)}{b_n} &= \max\left\{\max_{k\le n_0} \frac{\psi(X_k)}{b_n}, \max_{n_0<k\le n} \frac{\psi(X_k)}{b_n}\right\} \\ &\le \max\left\{\max_{k\le n_0} \frac{\psi(X_k)}{b_{n_1}}, \max_{n_0<k\le n} \frac{\psi(X_k)}{b_k}\right\} \le 4\lambda\rho.\end{aligned}$$

It follows that for $n \ge n_1$,

$$\begin{aligned}&\sup_{\psi(t)>4\lambda\rho b_n} b_n^{-1}\psi(t) \left| \textstyle\sum_{k=1}^n (I_{[X_k\ge t]} - P(X_k \ge t)) \right| \\ &\qquad = \sup_{\psi(t)>4\lambda\rho b_n} b_n^{-1}\psi(t) \textstyle\sum_{k=1}^n P(X_k \ge t),\end{aligned}$$

since for $n \ge n_1$ and $k \le n$, $\psi(X_k) \le 4\lambda\rho b_n$ and so

$$\sup_{\psi(t)>4\lambda\rho b_n} \textstyle\sum_{k=1}^n I_{[X_k\ge t]} = 0.$$

Using (2.11) and (2.16) we get (2.13).

3. Proofs of Theorems 0.1, 0.2, and 0.3.

Theorems 0.1 and 0.2 follow from Theorem 0.3. Nevertheless we will begin with a proof of Theorem 0.1 based on Lemma 1.1. Then we will obtain Theorem 0.3 as a consequence of Lemma 1.7 and remark on why it implies the other two theorems.

PROOF OF THEOREM 0.1. If $\sum_{k=1}^\infty c_k^2 < \infty$, then it follows immediately from Lemma 1.5 that $\lim\sup_{n\to\infty} \| \mathscr{U}_n \| < \infty$ a.s. Therefore assume that $\lim_{n\to\infty} \sum_{k=1}^n c_k^2 = \infty$. Let $c = e$, and for $\{a_n\}$ as defined in Theorem 0.1, construct the associated N-V subsequence $\{a_{n_r}\}$.

By Lemma 2.1, (0.2) will hold if there exists a $\lambda > 0$ such that

$$\textstyle\sum_{r=1}^\infty P(\| \mathscr{U}_{n_r} - \mathscr{U}_{n_{r-1}} \| > 2\lambda a_{n_r}) < \infty \tag{3.1}$$

and

$$\lim_{n\to\infty} P(\| \mathscr{U}_n \| > \lambda a_n) = 0. \tag{3.2}$$

By Lemma 1.1 with $\mathscr{U}_n$ replaced by $\mathscr{U}_{n_r} - \mathscr{U}_{n_{r-1}}$ we have

$$P(\| \mathscr{U}_{n_r} - \mathscr{U}_{n_{r-1}} \| > 2\lambda a_n) \leq [1 + \sqrt{2\pi}4\lambda(L_2(\textstyle\sum_{k=1}^{n_r} c_k^2))^{1/2}]\exp\left\{-\frac{\lambda^2}{2} L_2(\textstyle\sum_{k=1}^{n_r} c_k^2)\right\}. \tag{3.3}$$

Recall that

$$n_r = \sup\{n: ((\textstyle\sum_{k=1}^{n} c_k^2)L_2(\sum_{k=1}^{n} c_k^2))^{1/2} \in (e^{s_r}, e^{s_r+1}]\}.$$

This implies that

$$2s_r \leq L(\textstyle\sum_{k=1}^{n_r} c_k^2) + L_3(\sum_{k=1}^{n_r} c_k^2) \leq 2L(\sum_{k=1}^{n_r} c_k^2)$$

and consequently, since $s_r \geq r$, that

$$Lr \leq Ls_r \leq L_2(\textstyle\sum_{k=1}^{n_r} c_k^2).$$

Therefore since $x \exp(-x^2/2)$ decreases for $x \geq 1$, the right side of (3.3) is, for $\lambda > 1$,

$$\leq [1 + 4\sqrt{2\pi}\lambda](Lr)^{1/2}\exp\left\{-\frac{\lambda^2 Lr}{2}\right\}$$

which is summable in r for $\lambda > \sqrt{2}$. Thus we get (3.1) for $\lambda > \sqrt{2}$.

Using Lemma 1.1 again we see that

$$P(\| \mathscr{U}_n \| > \lambda a_n) \leq [1 + 2\sqrt{2\pi}\lambda]\exp\left\{-\frac{\lambda^2}{8} L_2(\textstyle\sum_{k=1}^{n} c_k^2)\right\}$$

so (3.2) holds for all $\lambda > 0$. This completes the proof of Theorem 0.1.

PROOF OF THEOREM 0.3. Given $\{b_n\}$ let $c = 2$ and form the N-V subsequence $\{b_{n_r}\}$. By Lemma 2.1, in order to obtain (0.8) we need only show that

$$\textstyle\sum_{r=1}^{\infty} P(\| \mathscr{S}_{n_r} - \mathscr{S}_{n_{r-1}} \| > 2(112\lambda + 2\beta)b_{n_r}) < \infty \tag{3.4}$$

and

$$\lim_{n\to\infty} P(\| \mathscr{S}_n \| > (112\lambda + 2\beta)b_n) = 0. \tag{3.5}$$

For ease of notation set $I_r = \{n: n_{r-1} < n \leq n_r\}$. Also, note that by symmetry

$$\textstyle\sum_{k\in I_r} \varepsilon_k\eta_k\psi(X_k) = \sum_{k\in I_r} \varepsilon_k\eta_k\psi(X_k)I_{[|\eta_k\psi(X_k)|\leq 2\lambda b_{n_r}]} + \sum_{k\in I_r} \varepsilon_k\eta_k\psi(X_k)I_{[|\eta_k\psi(X_k)|>2\lambda b_{n_r}]}$$

and

$$\textstyle\sum_{k\in I_r} \varepsilon_k\eta_k\psi(X_k)I_{[|\eta_k\psi(X_k)|\leq 2\lambda b_{n_r}]} - \sum_{k\in I_r} \varepsilon_k\eta_k\psi(X_k)I_{[|\eta_k\psi(X_k)|>2\lambda b_{n_r}]}$$

have the same distribution. This implies that

$$P(|\textstyle\sum_{k\in I_r} \varepsilon_k\eta_k\psi(X_k)I_{[|\eta_k\psi(X_k)|\leq 2\lambda b_{n_r}]}| > 2\lambda b_{n_r}) \leq 2P(|\sum_{k\in I_r} \varepsilon_k\eta_k\psi(X_k)| > 2\lambda b_{n_r}). \tag{3.6}$$

The main step in this proof is to use Lemma 1.7 applied to $\mathcal{S}_{n_r} - \mathcal{S}_{n_{r-1}}$ with λ replaced by $8\lambda b_{n_r}$, δ by $108\lambda b_{n_r}$, β by βb_{n_r} and with $\gamma = 2\lambda b_{n_r}$. In order to satisfy (1.20) in this case we must have

$$\sup_{1\le j\le n_r-n_{r-1}} \sup_{t>0} \psi(t) \left| \sum_{k=n_{r-1}+1}^{n_{r-1}+j} E\eta_k I_{[X_k\ge t]} I_{[|\eta_k\psi(X_k)|>2\lambda b_{n_r}]} \right| \le \beta b_{n_r}. \tag{3.7}$$

This is satisfied for r sufficiently large by (0.7) and the triangle inequality. Thus, for r sufficiently large

$$\begin{aligned} &P(\| \mathcal{S}_{n_r} - \mathcal{S}_{n_{r-1}} \| > (116\lambda + 2\beta) b_{n_r}) \\ &\quad \le \frac{32P(|\sum_{k\in I_r} \varepsilon_k \eta_k \psi(X_k)| > 2\lambda b_{n_r})}{1 - 64(108\lambda b_{n_r})^{-2} \sum_{k\in I_r} E\eta_k^2 \psi^2(X_k) I_{[|\eta_k\psi(X_k)|\le 2\lambda b_{n_r}]}} \\ &\qquad + P(\sup_{k\in I_r} |\eta_k \psi(X_k)| > 2\lambda b_{n_r}). \end{aligned} \tag{3.8}$$

We now show that (3.8) is summable in r. First we use a Lemma of Hoffmann-Jørgensen to estimate the sum in the denominator of the first term to the right of the inequality sign, (Theorem 3.1, [4]). We get

$$\begin{aligned} &\sum_{k\in I_r} E\eta_k^2 \psi^2(X_k) I_{[|\eta_k\psi(X_k)|\le 2\lambda b_{n_r}]} \\ &\quad = E\left| \sum_{k\in I_r} \varepsilon_k \eta_k \psi(X_k) I_{[|\eta_k\psi(X_k)|\le 2\lambda b_{n_r}]} \right|^2 \\ &\quad \le 18[t_0^2 + E \sup_{k\in I_r} |\varepsilon_k \eta_k \psi(X_k) I_{[|\eta_k\psi(X_k)|\le 2\lambda b_{n_r}]}|^2] \end{aligned} \tag{3.9}$$

where

$$t_0 = \inf\{t > 0: P(|\textstyle\sum_{k\in I_r} \varepsilon_k \eta_k \psi(X_k) I_{[|\eta_k\psi(X_k)|\le 2\lambda b_{n_r}]}| > t) \le 1/72\}. \tag{3.10}$$

By (3.6) and Lévy's inequality

$$P(|\textstyle\sum_{k\in I_r} \varepsilon_k \eta_k \psi(X_k) I_{[|\eta_k\psi(X_k)|\le 2\lambda b_{n_r}]}| > t) \le 4P(|\sum_{k=1}^{n_r} \varepsilon_k \eta_k \psi(X_k)| > t). \tag{3.11}$$

Using (3.11) in (3.10) along with (0.6) and (2.3) we see that there exists an integer r_0 such that for $r \ge r_0$, $t_0 \le \lambda b_{n_r}$. Therefore, (3.9)

$$\le 18[(\lambda b_{n_r})^2 + (2\lambda b_{n_r})^2]$$

and the denominator in the first term on the right of the inequality sign in (3.8)

$$\ge 1 - 64(108\lambda b_{n_r})^{-2} 18 \cdot 5(\lambda b_{n_r})^2 > 1/2.$$

Therefore, for $r \ge r_0$

$$\begin{aligned} &P(\| \mathcal{S}_{n_r} - \mathcal{S}_{n_{r-1}} \| > (116\lambda + 2\beta) b_{n_r}) \\ &\quad \le 64P(|\textstyle\sum_{k\in I_r} \varepsilon_k \eta_k \psi(X_k)| > 2\lambda b_{n_r}) + \sum_{k\in I_r} P(|\eta_k \psi(X_k)| > 2\lambda b_{n_r}). \end{aligned} \tag{3.12}$$

The first term on the right in (3.12) is summable by (0.6) and Lemma 2.1 applied to S_n. The second term on the right in (3.12) is summable by (0.6), the Borel-Cantelli Lemma and Lévy's inequality.

By exactly the same estimate we used in (3.8) for $\mathcal{S}_{n_r} - \mathcal{S}_{n_{r-1}}$ and the preceding argument we have, for all n sufficiently large

$$P(\| \mathscr{S}_n \| > (112\lambda + 2\beta)b_n)$$
$$\leq 128P(| \sum_{k=1}^n \varepsilon_k \eta_k \psi(X_k) | > 2\lambda b_n) + P(\sup_{k \leq n} | \eta_k \psi(X_k) | > 2\lambda b_n)$$

and this goes to zero as $n \to \infty$ by (0.6) and Lemma 2.1.

Thus we have established (3.4) and (3.5). This completes the proof of the theorem.

PROOF OF THEOREM 0.2. Use Theorem 0.3 with $\eta_k = 1$ for all k along with (1.23). This gives Theorem 0.2 except that we get $\beta/2$ instead of β on the right side of (0.4). To get β note that because the terms in (3.7) are positive we don't need to use the triangle inequality. (I.e. (3.7) follows from (0.4) in this case.)

PROOF OF THEOREM 0.1 AS A COROLLARY OF THEOREM 0.3. Let $\eta_k = c_k$, $k = 1, \cdots$ and let $\psi(s) = 1$ for all $s > 0$. If we take $\lambda > 1/2$ and $b_n = a_n$ (for a_n as in Theorem 0.1) we see that (0.7) is identically zero. Thus to obtain Theorem 0.1 we need only show that

$$\limsup_{n \to \infty} \frac{| \sum_{k=1}^n \varepsilon_k c_k |}{(\sum_{k=1}^n c_k^2 L_2 (\sum_{k=1}^n c_k^2))^{1/2}} < \lambda \quad \text{a.s.} \tag{3.13}$$

for some $\lambda > 0$. This, of course, is a very simple form of the law of the iterated logarithm for real valued random variables. Since we do not know of any reference for the fact that (3.13) holds without any conditions on the $\{c_k\}$ we will give a proof of (3.13). By the subgaussian inequality (Lemma 5.2, Chapter 2 [5]), for integers p and q,

$$P\left(\frac{| \sum_{k=p}^q \varepsilon_k c_k |}{(\sum_{k=p}^q c_k^2)^{1/2}} > u\right) \leq 2 \exp\left(\frac{-u^2}{2}\right).$$

This inequality enables us to use Lemma 2.1 to obtain (3.13) exactly as in the first proof of Theorem 0.1 given at the beginning of this section. The only difference is that we consider $\sum_{k=1}^n \varepsilon_k c_k$ instead of $\mathscr{U}_n$.

REMARK 3.1. It follows from Theorem 2.2 that some condition like (0.4) is necessary to (0.5) to hold. Therefore we can not entirely dispense with (0.4). (The same argument applies to (0.7).) It would be tidy if (0.3) implied (0.4), but it is easy to see that it does not. Suppose (0.3) holds and that $\lim_{t \to \infty} \psi(t) = \infty$ (otherwise (0.4) is satisfied). Let X_k be defined on Ω_k and let $A_k \subset \Omega_k$ be such that $\text{Prob}(A_k) = p_k$ where $\sum_{k=1}^\infty p_k < \infty$. Define $\{X_k\}$ on $\Omega = \otimes_{k=1}^\infty \Omega_k$. Define

$$X_k' = \begin{cases} X_k & \text{on} \quad \Omega_k - A_k \\ t_k & \text{on} \quad A_k \end{cases}$$

where t_k is some number for which

$$\psi(t_k) \geq \frac{k \lambda b_k}{p_k}.$$

Then, clearly, $\{X'_k\}$ also satisfies (0.3) however for any integer N, if $n > N$

$$\sup_{\psi(t)>N\lambda b_n} b_n^{-1}\psi(t) \sum_{k=1}^n P[X_k \geq t] \geq \psi(t_n)P[X_n \geq t_n] \geq n\lambda.$$

Therefore, (0.3) can not imply any condition like (0.4).

We end this section by considering some well known sufficient conditions for the law of the iterated logarithm for real valued random variables. Using Theorems 0.2 and 0.3 we immediately get results for the corresponding weighted empirical processes since, as we shall show, in these cases it is completely elementary that (0.4) or respectively (0.7) is satisfied.

In the examples below we let $Y_k = \varepsilon_k\psi(X_k)$, $S_n = \sum_{k=1}^n Y_k$, $s_n^2 = \sum_{k=1}^n EY_k^2$ and $b_n^2 = 2s_n^2 L_2 s_n^2$.

EXAMPLE 1. In Theorem 1, [11], Teicher extends the classical Kolmogorov LIL. One of the hypotheses of this theorem is: for all $\varepsilon > 0$,

$$\frac{1}{s_n^2} \sum_{j=1}^n \int_{|x|>\varepsilon s_j (L_2 s_j^2)^{-1/2}} x^2 F_j\,(dx) \to 0 \quad \text{as} \quad n \to \infty. \tag{3.14}$$

We now show that (0.4) follows from (3.14). We have

$$\begin{aligned}\sup_{\psi(t)>2\lambda b_n} b_n^{-1}\psi(t) \sum_{k=1}^n P(X_k \geq t) &\leq \sup_{\psi(t)>2\lambda b_n} b_n^{-1}\psi(t) \sum_{k=1}^n P(\psi(X_k) \geq \psi(t)) \\ &\leq \sup_{\psi(t)>2\lambda b_n} \frac{b_n^{-1}\psi(t) \sum_{k=1}^n EY_k^2 I_{[|Y_k|\geq\psi(t)]}}{\psi^2(t)} \\ &\leq \frac{\sum_{k=1}^n EY_k^2 I_{[|Y_k|\geq 2\lambda b_n]}}{2\lambda b_n^2},\end{aligned}$$

which goes to zero by (3.14), since

$$b_n \geq s_n$$

and

$$b_n \geq \varepsilon s_j/\sqrt{L_2 s_j^2}, \quad j \leq n.$$

EXAMPLE 2. In Theorem 7.5.1 [2], Chung also gives a LIL. One hypothesis is: for some $0 < \varepsilon < 1$ and $A < \infty$

$$\frac{\sum_{j=1}^n E|Y_j|^3}{(\sum_{j=1}^n EY_j^2)^{3/2}} \leq \frac{A}{(Ls_n)^{1+\varepsilon}}. \tag{3.15}$$

Proceeding in a manner similar to Example 1, we have

$$\begin{aligned}&\sup_{\psi(t)>2\lambda b_n} b_n^{-1}\psi(t) \sum_{k=1}^n P(|Y_k| \geq \psi(t)) \\ &\qquad \leq \sup_{\psi(t)>2\lambda b_n} b_n^{-1}\psi(t) \frac{\sum_{k=1}^n E|Y_k|^3}{\psi^3(t)} \leq \frac{\sum_{k=1}^n E|Y_k|^3}{4\lambda^2 b_n^3},\end{aligned}$$

which goes to zero as n goes to infinity by (3.15).

4. Applications of Lemma 1.7. We need the following well known result in what follows, which we include for the sake of completeness.

LEMMA 4.0. *Let $\psi(t)$, $t \geq 0$ be a non-negative, non-decreasing, left continuous function. Let $a = \inf\{t:\psi(t) > 0\}$. Then there exists a sequence of functions $\{\psi_m(t)\}$ which are continuous with $\psi_m(t) = 0$, $t \in [0, a]$ and strictly increasing for $t > a$ such that $\psi_m(t)$ increases in m and $\lim_{m\to\infty}\psi_m(t) = \psi(t)$ for all $t > 0$.*

PROOF. We define $\psi^{-1}(u) = \inf\{s \geq 0:\psi(s) > u\}$. Note that for $t > 0$ and $u \geq 0$, $\psi^{-1}(u) < t$ if and only if $u < \psi(t)$. This shows that for $t > 0$, $\lambda\{u \geq 0:\psi^{-1}(u) < t\} = \psi(t)$, where λ is Lebesgue measure. Let $\eta \geq 0$ be any non-negative random variable with respect to $([0, \infty), Q)$ which has a distribution function which is mutually absolutely continuous with respect to Lebesgue measure on $[0, \infty)$. We consider (ψ^{-1}, η) as defined on the product space $([0, \infty) \times [0, \infty), \lambda \times Q)$ and for each integer $m > 1$ define

$$\psi_m(t) = \lambda \times Q\{\psi^{-1} + (\eta/m) < t\}, \quad t \geq 0.$$

One can check that $\{\psi_m\}$ satisfies the conditions stated in the lemma.

Our main result, Theorem 0.3, followed immediately from Lemma 1.7. There are a number of other interesting results which follow, more or less immediately, from this lemma. We first consider a one-sided version of Daniel's Theorem in the independent (but not necessarily identically distributed) case. This has been studied by van Zuijlen [13], [14], [15] and independently in [8].

THEOREM 4.1. *Let $\{X_n\}$ be independent non-negative, real-valued* r.v.'s. *Then*

$$\sup_{c>0} cP\left(\sup_{t>0} \frac{\sum_{k=1}^n I_{[X_k\geq t]}}{\sum_{k=1}^n P(X_k \geq t)} > c\right) \leq 1030. \tag{4.1}$$

PROOF. Let $\psi(t) = (\sum_{k=1}^n P(X_k \geq t))^{-1}$, and note that ψ is non-decreasing and left continuous. We now show that for this function ψ we have

$$\sum_{k=1}^n sP(\psi(X_k) \geq s) \leq 1. \tag{4.2}$$

To see this let $\{\psi_m(t)\}$ be a sequence of continuous functions, strictly increasing in t, such that for each t, $\psi_m(t)$ increases in m, $\psi_m(0) = 0$, and $\lim_{m\to\infty}\psi_m(t) = \psi(t)$, $t > 0$. Such functions exist by Lemma 4.0. By the definition of $\psi(t)$ we have

$$\sup_{t>0}\psi_m(t) \sum_{k=1}^n P(X_k \geq t) \leq 1$$

from which it follows that

$$\sup_{s>0} s \sum_{k=1}^n P(\psi_m(X_k) \geq s) \leq 1. \tag{4.3}$$

The statement in (4.2) now follows by monotonicity.

Now by (4.2) we have

$$\sum_{k=1}^{n} E\psi^2(X_k)I_{[\psi(X_k)\le\gamma]} \le 2\int_0^{\gamma} \sum_{k=1}^{n} sP(\psi(X_k) \ge s)\, ds \le 2\gamma, \tag{4.4}$$

and furthermore

$$\begin{aligned} &P(|\textstyle\sum_{k=1}^{n} \varepsilon_k\psi(X_k)I_{[\psi(X_k)\le\gamma]}| \ge (\gamma/4)) \\ &\quad\le 16\gamma^{-2}\textstyle\sum_{k=1}^{n} E\psi^2(X_k)I(\psi(X_k)\le\gamma) \le (32/\gamma). \end{aligned} \tag{4.5}$$

Next we note that (in the notation of Theorem 0.3)

$$\mathscr{S}_n(t) = \frac{\sum_{k=1}^{n} I_{[X_k\ge t]}}{\sum_{k=1}^{n} P(X_k \ge t)} - 1. \tag{4.6}$$

We now apply Lemma 1.7 with $\gamma = \lambda = (c/2) = 3$, $\delta = (c/2)$, $\beta = 1$ and $\eta_k \equiv 1$. First assume $c \ge 1024$. Then (1.19) and (1.20) easily hold by (4.2) and the definition of ψ. Hence, by (1.21), (4.2), (4.5) and (4.6)

$$\begin{aligned} &P\left(\sup_{t>0} \frac{\sum_{k=1}^{n} I_{[X_k\ge t]}}{\sum_{k=1}^{n} P(X_k \ge t)} > c\right) \\ &\quad\le P(\|\mathscr{S}_n\| > c-1) = P(\|\mathscr{S}_n\| > \lambda+\delta+2\beta) \\ &\quad\le \frac{(8)(32)/\lambda}{1-(64/\delta^2)\,2\lambda} + \frac{1}{\lambda} \le \frac{1030}{c}. \end{aligned} \tag{4.6a}$$

For $c \le 1024$

$$P\left(\sup_{t>0} \frac{\sum_{k=1}^{n} I_{[X_k\ge t]}}{\sum_{k=1}^{n} P(X_k \ge t)} > c\right) \le 1 \le \frac{1030}{c}. \tag{4.6b}$$

Combining (4.6a) and (4.6b) we get (4.1).

REMARK 4.2. Our approach to Theorem 4.1 is essentially that of Corollary 3.5 [8]. In [15] van Zuijlen obtains (4.1) and with a better constant. Our purpose in presenting (4.1) here is to show that the inequality in Lemma 1.7 is sharp enough to give these results (except for the constant). In the i.i.d. case the constant is 1 and it seems reasonable to expect that the constant is 1 in the non-i.i.d. case as well. In Section 5 we will give another approach to this generalization of Daniel's theorem and obtain a bound that is a little better than those of [14] and [15] and discuss further our conjecture about the constant being equal to one.

In the next Lemma we obtain a bound for the weighted empirical process for random variables with tail distribution given by the Pareto distribution.

LEMMA 4.2. *Let $\theta(p) = \theta$ be a random variable satisfying (for some $0 < p < \infty$)*

$$P[\theta \geq \lambda] = 1 \wedge \lambda^{-p}, \tag{4.7}$$

and let $\{\theta_k\}$ be i.i.d. *copies of θ. Let $\{a_k\}$ be non-negative real numbers such that $\sum_{k=1}^{\infty} a_k^p < \infty$. Then*

$$\begin{aligned}(\sup_{\lambda>0}\lambda^{p/q}P[\sup_{t>0}t^q \mid \textstyle\sum_{k=1}^{\infty} (I_{[a_k\theta_k \geq t]} - P[a_k\theta_k \geq t]) \mid > \lambda])^{1/p} \\ \leq (257c_{p,q})^{1/p}(\textstyle\sum_{k=1}^{\infty} a_k^p)^{1/p}\end{aligned} \tag{4.8}$$

where $p/2 < q \leq p$ and $c_{p,q} = 3^{p/q}(p/(2q-p))$.

PROOF. We use Lemma 1.9 with $\psi(t) = t^q$, $\eta_k \equiv 1$, $\gamma = \lambda = \delta$ and $X_k = a_k\theta_k$. Without loss of generality we may assume that the $\{a_k\}$ are non-increasing, and since (4.8) is homogeneous in $\{a_k\}$ we can assume $\sum_{k=1}^{\infty} a_k^p = 1$. Note that

$$\sup_{t^q>1} t^q \sum_{k=1}^{\infty} P(a_k\theta_k \geq t) \leq \sup_{t>1} \frac{\sum_{k=1}^{\infty} a_k^p}{t^{p-q}} \leq 1.$$

Hence (1.28) holds with $\beta = 1$ and $\lambda = \gamma > 1$ and clearly (1.29) also holds. Furthermore

$$P(\sup_k \psi(a_k\theta_k) > \lambda) \leq \lambda^{-p/q}. \tag{4.9}$$

Now consider

$$\begin{aligned}\sum_{k=1}^{\infty} E(a_k\theta_k)^{2q} I_{[a_k\theta_k \leq \lambda^{1/q}]} &= p \sum_{k=1}^{\infty} a_k^{2q} \int_1^{\lambda^{1/q}a_k^{-1}} u^{2q-p-1}\, du \\ &\leq \left(\frac{p}{2q-p}\right)\lambda^{2-(p/q)}.\end{aligned} \tag{4.10}$$

Therefore, for $\lambda > (128p/(2q-p))^{q/p}$, the left side of (1.27) is less than ½. Hence by Lemma 1.9, (4.9), (4.10) and Chebyshev's inequality we get

$$\begin{aligned}P(\sup_{t>0} t^q \mid \textstyle\sum_{k=1}^{\infty} (I_{[a_k\theta_k \geq t]} - P(a_k\theta_k \geq t)) \mid > 2\lambda + 2) \\ \leq \left(16 \cdot 16 \frac{p}{2q-p} + 1\right)\lambda^{-p/q}.\end{aligned}$$

Also since $3\lambda \geq 2\lambda + 2$, letting $u = 3\lambda$ we get for $u > 3(128p/(2q-p))^{q/p}$

$$P(\sup_{t>0} t^q \mid \textstyle\sum_{k=1}^{\infty} (I_{[a_k\theta_k \geq t]} - P(a_k\theta_k \geq t)) \mid > u) \leq \left(3^{p/q} \cdot \frac{257p}{2q-p}\right)u^{-p/q}. \tag{4.11}$$

Of course (4.11) is also valid for $u \leq 3(128p/(2q-p))^{q/p}$ since a probability is less than or equal to 1. Thus we get (4.8).

If we take $a_1 = \cdots = a_n = n^{-1/p}$; $a_{n+1} = \cdots = 0$ in (4.8) we get probability estimates for i.i.d. sequences which may be useful in constructing confidence intervals.

COROLLARY 4.3. *Let $\theta(p) = \theta$ and $\{\theta_k\}$ be as in Lemma 4.2. Then for $p/2 < q \le p$ and all $n \ge 1$,*

$$\left(\sup_{\lambda>0}\lambda^{p/q}P\left[\sup_{t>0}t^q \frac{1}{n^{q/p}} \mid \textstyle\sum_{k=1}^n (I_{[\theta_k \ge t]} - P(\theta_k \ge t)) \mid > \lambda\right]\right)^{1/p} \le (257c_{p,q})^{1/p} \tag{4.12}$$

or, equivalently

$$\left(\sup_{\lambda>0}\lambda^{p/q}P\left(\sup_{t>0}t^q \frac{1}{n} \mid \textstyle\sum_{k=1}^n (I_{[\theta_k \ge t]} - P[\theta_k \ge t]) \mid > \lambda\right)\right)^{1/p} \le \frac{(257c_{p,q})^{1/p}}{n^{(1/q-1/p)}}, \tag{4.13}$$

where $c_{p,q}$ is given in Lemma 4.2. In particular, if we let U be a uniformly distributed random variable on [0, 1] and let $\{U_k\}$ be i.i.d. copies of U, then for $\frac{1}{2} < q \le 1$ and all $n \ge 1$,

$$\sup_{\lambda>0}\lambda^{1/q}P\left[\sup_{0\le s\le 1}\left(\frac{1}{s}\right)^q \frac{1}{n} \mid \textstyle\sum_{k=1}^n (I_{[U_k \le s]} - s) \mid > \lambda\right] \le \frac{257c_{1,q}}{n^{1/q-1}}, \tag{4.14}$$

where $c_{1,q}$ is given in Lemma 4.2.

PROOF. Inequalities (4.12) and (4.13) follow immediately from (4.8). To obtain (4.14) note that when $p = 1$, $\theta(1) = \theta = 1/U$. Using this in (4.13) and setting $t = 1/s$ we get (4.14).

REMARK 4.4. If $q = 1$, (4.14) is a special case of Daniel's Theorem (with a larger constant). However for $\frac{1}{2} < q < 1$ the rate of decay of the probability in (4.14) is faster than λ^{-1}, which is the rate of decay in Daniel's Theorem. Thus (4.14) can be viewed as a generalization of Daniel's Theorem in this case and, by a change of variables, in general. It is not exactly a new result, it is given in Mason's Theorem 1 [9], although Mason's result is a little more restrictive than (4.14) in that the supremum is not taken over all $0 < s \le 1$. To obtain Mason's result from (4.14) let $q = 1 - \nu$ and $\lambda = an^{-\nu}$. Then for $0 \le \nu < \frac{1}{2}$ we get by (4.14)

$$P[\sup_{0<s\le 1}(1/s)^{1-\nu}n^{\nu}\{(1/n) \mid \textstyle\sum_{k=1}^n (I_{[U_k \le s]} - s) \mid\} > a] \le C_\nu a^{-1/(1-\nu)}.$$

All these observations demonstrate that Lemma 1.7 is interesting even in the i.i.d. case. It is also interesting to see what Theorem 0.2 gives for i.i.d. uniformly distributed random variables. In Theorem 0.2 let $X_k = 1/U_k$ for $\{U_k\}$ as given in Corollary 4.3 and let $b_n = (nL_2n)^{1/2}$. Then (0.4) holds for $\lambda > \sqrt{2}$ if $E\psi^2(1/U_1) < \infty$ or, equivalently, if

$$\int_0^1 \psi^2\left(\frac{1}{s}\right) ds < \infty. \tag{4.15}$$

If (4.15) holds then, since $\psi(1/s)$ is non-increasing as s increases from 0,

$\lim_{s\to 0} s\psi^2(1/s) = 0$. Thus for $s \le s(\varepsilon)$ sufficiently small

$$s\psi(1/s) < \varepsilon[\psi(1/s)]^{-1}. \tag{4.16}$$

Letting $t = \frac{1}{2}$, condition (0.4) of Theorem 0.2 is satisfied if

$$\limsup_{n\to\infty} \sup_{\psi(1/s)>2b_n} b_n^{-1}\psi(1/s)ns = 0$$

which, clearly, is implied by (4.16). Thus we have that if (4.5) holds then

$$\limsup_{n\to\infty} \sup_{0\le s\le 1} \frac{\psi(1/s) \,|\, \sum_{k=1}^n (I_{[U_k\le s]} - s) \,|}{(nL_2n)^{1/2}} \le 1120\sqrt{2} \quad \text{a.s.} \tag{4.17}$$

This is not as good as James' [6] result applied to the bounded LIL, but it is pretty close. (Actually we have a slightly weaker hypothesis on the smoothness of ψ.) Of course, it is well known that the bounded LIL for the weighted empirical process can hold even if the independent real valued random variables do not satisfy the LIL.

5. More on Daniel's Theorem. If $\{X_k\}$ are i.i.d., non-negative, real-valued random variables, then Daniel's theorem gives

$$P\left(\sup_{t>0} \frac{\sum_{k=1}^n I_{[X_k\ge t]}}{\sum_{k=1}^n P(X_k \ge t)} > \lambda\right) \le \frac{1}{\lambda} \tag{5.1}$$

(where we define $0/0 = 0$). This is easy to prove since

$$M_t = \frac{I_{[X_k\ge t]}}{P(X_k \ge t)}$$

is a martingale. We would like to obtain (5.1) even when the $\{X_k\}$ are not identically distributed. Considering Theorem 4.1 and the references following it, the problem comes down to getting the constant 1 on the right in (4.1). Van Zuijlen [14] has already shown that for all $\varepsilon > 0$, for c sufficiently large the constant in (4.1) can be taken to be $1 + \varepsilon$.

Here we obtain an upper bound for the left side of (5.1) when $\{X_k\}$ are not necessarily identically distributed that gives somewhat sharper bounds than those obtained by van Zuijlen. Before doing this, however, let us remark that for $X_k = a_k\theta_k$, for $\{\theta_k\}$ as in Lemma 4.2 we do get (5.1). This result is given in Remark 3.9 (iii) of [8].

The main result of this section is the following inequality which is a generalization of Theorem 3.3 [8].

THEOREM 5.1. *Let $\{X_k\}$ and $\psi(t)$, $t \ge 0$ be as in Theorem 0.2. Then*

$$P\left(\sup_{t>0}\psi(t) \textstyle\sum_{k=1}^\infty I_{[X_k\ge t]} \ge \lambda\right) \le \sum_{k=1}^\infty \frac{1}{k!}\left[\frac{k}{\lambda}\sup_{t>0}\psi(t) \sum_{k=1}^\infty P(X_k \ge t)\right]^k. \tag{5.2}$$

PROOF. There is nothing to prove unless $\sum_{k=1}^\infty P(X_k \ge t) < \infty$ for all $t > 0$.

Therefore we shall assume that this is the case. Let $X_1^* \ge X_2^* \ge \cdots$ denote a non-increasing rearrangement of $\{X_k\}$. We have

$$\begin{aligned}(5.3)\qquad & P(\sup_{t>0}\psi(t)\ \textstyle\sum_{k=1}^\infty I_{[X_k\ge t]} \ge \lambda) \\ &= P(\sup_k k\psi(X_k^*) \ge \lambda) \le \textstyle\sum_{k=1}^\infty P(\psi(X_k^*) \ge \lambda/k) \\ &= \textstyle\sum_{k=1}^\infty P(\text{at least } k \text{ of the events } \{\psi(X_i) \ge (\lambda/k)\}_{i=1}^\infty \text{ occur}).\end{aligned}$$

By Bonferoni's inequality this is

$$(5.4)\qquad \le \sum_{k=1}^\infty \frac{[\sum_{k=1}^\infty P(\psi(X_i) \ge \lambda/k)]^k}{k!}.$$

Also, clearly

$$(5.5)\qquad (\lambda/k)\ \textstyle\sum_{i=1}^\infty P(\psi(X_i) \ge \lambda/k) \le \sup_{t>0} t\ \textstyle\sum_{i=1}^\infty P(\psi(X_i) \ge t).$$

Therefore by (5.3), (5.4) and (5.5)

$$\begin{aligned}(5.6)\qquad & P(\sup_{t>0}\psi(t)\ \textstyle\sum_{k=1}^\infty I_{[X_k\ge t]} \ge \lambda) \\ &\le \textstyle\sum_{k=1}^\infty (1/k!)(k/\lambda)^k[\sup_{t>0} t\ \textstyle\sum_{j=1}^\infty P(\psi(X_j) \ge t)]^k.\end{aligned}$$

Let $a = \inf\{t:\psi(t) > 0\}$ and let $\{\psi_m(t)\}$ be the sequence of functions described in Lemma 4.0. Clearly (5.6) holds for each ψ_m and

$$\begin{aligned}\sup_{t>0} t(\textstyle\sum_{k=1}^\infty P(\psi_m(X_k) \ge t)) &= \sup_{\psi_m(t)>0}\psi_m(t)\ \textstyle\sum_{k=1}^\infty P(\psi_m(X_k) \ge \psi_m(t)) \\ &= \sup_{t>a}\psi_m(t)\ \textstyle\sum_{k=1}^\infty P(\psi_m(X_k) \ge \psi_m(t)) \\ &= \sup_{t>a}\psi_m(t)\ \textstyle\sum_{k=1}^\infty P(X_k \ge t) \\ &\le \sup_{t>0}\psi(t)\ \textstyle\sum_{k=1}^\infty P(X_k \ge t).\end{aligned}$$

Thus we have

$$\begin{aligned}(5.7)\qquad & P(\sup_{t>0}\psi_m(t)\ \textstyle\sum_{k=1}^\infty I_{[X_k\ge t]} \ge \lambda) \\ &\le \textstyle\sum_{k=1}^\infty (1/k!)(k/\lambda)^k[\sup_{t>0}\psi(t)\ \textstyle\sum_{j=1}^\infty P(X_j \ge t)]^k\end{aligned}$$

and (5.2) follows by monotonicity of ψ_m in m and the continuity (when finite) of the right hand side of (5.2).

COROLLARY 5.2. *Assume that* $\sum_{k=1}^\infty P(X_k \ge t) < \infty$ *for all* $t > 0$. *Then*

$$(5.8)\qquad P\left(\sup_{t>0} \frac{\sum_{k=1}^\infty I_{[X_k\ge t]}}{\sum_{k=1}^\infty P(X_k \ge t)} \ge \lambda\right) \le 1 \wedge \sum_{k=1}^\infty \frac{1}{k!}\left(\frac{k}{\lambda}\right)^k.$$

PROOF. This follows from Theorem 5.1 by taking $\psi(t) = (\sum_{k=1}^\infty P(X_k \ge t))^{-1}$.

REMARK 5.3. The statement in (5.8) is as close as we can come to (5.1). The sum on the right in (5.8) converges for $\lambda > e$, but of course, the probability must be less than or equal to 1. It is easy to see that (5.8) gives

$$\lim_{\lambda\to\infty}\lambda P\left(\sup_{t>0}\frac{\sum_{k=1}^{\infty} I_{[X_k\ge t]}}{\sum_{k=1}^{\infty} P(X_k\ge t)}\ge\lambda\right)=1,$$

a result which was obtained by van Zuijlen in [14].

Van Zuijlen has given two bounds for the left side of (5.8). In [14] he shows that it can be taken to be

$$\frac{1}{\lambda}+\frac{7\pi^2}{3}\frac{1}{\lambda^2(1-(2/\lambda))^4},\quad \lambda>2 \tag{5.9}$$

and in [15] that it can be taken to be

$$\frac{2\pi^2}{3\lambda(1-(1/\lambda))^4},\quad \lambda>1. \tag{5.10}$$

Of course these bounds are meaningless unless they are less than 1 and they are both greater than 1 for $\lambda\le 2e$. We will show that the bound in (5.8) is smaller than the bounds in (5.9) and (5.10) for $\lambda>2e$. We have for $\lambda>2e$

$$\frac{1}{\lambda}+\sum_{k=2}^{\infty}\frac{1}{k!}\left(\frac{k}{\lambda}\right)^k\le\frac{1}{\lambda}+\left(\frac{e}{\lambda}\right)^2\sum_{k=2}^{\infty}\left(\frac{e}{\lambda}\right)^{k-2}\le\frac{1}{\lambda}+2\left(\frac{e}{\lambda}\right)^2$$

which is less than both (5.9) and (5.10) if $\lambda>2e$.

REFERENCES

[1] BRETAGNOLLE, J. (1980). Statistique de Kolmogorov-Smirnov pour un enchantillon non equirepart. *Colloques Internationaux du Centre National de la Recherche Scientifique* **307** 39–44.

[2] CHUNG, K. L. (1974). *A Course in Probability Theory.* Academic, New York.

[3] DVORETZKY, A., KIEFER, J. and WOLFOWITZ (1956). Asymptotic minimax character of the sample distribution. *Ann. Math. Statist.* **27** 642–669.

[4] HOFFMANN-JØRGENSEN, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **52** 159–186.

[5] JAIN, N. C. and MARCUS, M. B. (1978). Continuity of sub-gaussian processes. *Adv. in Probab.* **4**. Dekker, New York.

[6] JAMES, B. R. (1975). A functional law of the iterated logarithm for weighted empirical distributions. *Ann. Probab.* **3** 767–772.

[7] MARCUS, M. B. (1981). Weak convergence of the empirical characteristic function. *Ann. Probab.* **9** 194–201.

[8] MARCUS, M. B. and PISIER, G. (1982). Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes. *Acta Mathematica*, to appear.

[9] MASON, D. M. (1981). Bounds for weighted empirical distribution functions. *Ann. Probab.* **9** 881–884.

[10] NAGAEV, S. V. and VOLODIN, N. A. (1975). On the strong law of large numbers. *Theor. Probab. Appl.* **20** 626–631.

[11] TEICHER, H. (1974). On the law of the iterated logarithm. *Ann. Probab.* **2** 714–728.

[12] VANDERZANDEN, A. J. (1980). Some results for the weighted empirical process concerning the law of the iterated logarithm and weak convergence. Thesis, Michigan State University.

[13] ZUIJLEN, M. C. A. VAN (1976). Some properties of the empirical distribution function in the non-i.i.d. case. *Ann. Statist.* **4** 406–408.

[14] ZUIJLEN, M. C. A. VAN (1978). Properties of the empirical distribution function for independent non-identically distributed random variables. *Ann. Probab.* **6** 250–266.
[15] ZUIJLEN, M. C. A. VAN (1982). Properties of the empirical distribution function for independent non-identically distributed random vectors. *Ann. Probab.* **10** 108–123.

DEPARTMENT OF MATHEMATICS
TEXAS A & M UNIVERSITY
COLLEGE STATION, TEXAS 77843