Stochastic Processes and their Applications 24 (1987) 109-131
North-Holland

A CENTRAL LIMIT THEOREM FOR $D(A)$-VALUED PROCESSES

Richard F. BASS and Ronald PYKE
Department of Mathematics, University of Washington, Seattle, WA 98195, USA

Received 11 November 1986

Let $D(A)$ be the space of set-indexed functions that are outer continuous with inner limits, a generalization of $D[0, 1]$. This paper proves a central limit theorem for triangular arrays of independent $D(A)$ valued random variables. The limit processes are not restricted to be Gaussian, but can be quite general infinitely divisible processes. Applications of the theorem include construction of set-indexed Lévy processes and a unified central limit theorem for partial sum processes and generalized empirical processes. Results obtained are new even for the $D[0, 1]$ case.

AMS 1980 Subject Classifications: Primary 60F05; Secondary 60F17, 60B12.

D-space * empirical processes * partial sum processes * Skorokhod topology * Lévy processes * central limit theorem * set-indexed functions * subpoissonian

1. Introduction

The purpose of this paper is to state and prove a central limit theorem (CLT) for sums of independent D-valued random processes, where the limiting process is neither constrained to be Gaussian nor continuous. To set the context more precisely, let $X_{n1}, X_{n2}, \ldots, X_{nm_n}$ be, for each integer $n \geq 1$, a finite set of independent random quantities. If the X_{nj}'s take values in a space in which an addition operation $+$ is defined, let $S_n = X_{n1} + \cdots + X_{nm_n}$ denote their sum. In the important classical case in which the X_{nj}'s are real valued, the solution to the central limit problem was known by the end of the 1930's; cf. Gnedenko and Kolmogorov (1954). When the X_{nj}'s take values in more general linear spaces, the problem is far from being fully resolved.

There are a large number of results concerning the CLT when the X_{nj} take values in a Banach space. These include, among many others, the CLT for iid random variables in $C(S)$ for compact metric spaces S by Jain and Marcus (1975) and the Banach space CLT of Pisier (1975); cf. Araujo and Giné (1980).

When we turn from Banach spaces to the spaces of right continuous functions with left limits, the D spaces introduced in 1956 (cf. Prohorov (1956) and Skorokhod (1956)), the literature is much sparser. Although D-spaces under sup norms are

Prepared under National Science Foundation Grants MCS-83-00581 and MCS-82-02861.

Banach spaces, this is not the case when Skorokhod-like topologies are used. Following Prohorov (1956), results that focus on CLT's for $D[0, 1]$ and its generalizations, all with Gaussian limits, include the result of Fisz (1959) for sums of iid processes with independent increments, the extension to $D[0, 1]^d$ by Neuhaus (1971) and to general $D(T)$ by Straf (1971), the CLT's for $D[0, 1]^d$ by Bickel and Wichura ((1971), Theorem 5 for partial-sums and Theorem 6 for empiricals) and the general CLT for $D[0, 1]$ by Hahn (1978) for the case of Gaussian limiting processes. Extensions of D-spaces to non-compact index sets include that of $D[0, \infty)$, cf. Lindvall (1973).

The central limit problem is of course much broader than those aspects of it that involve Gaussian limits. As in the classical case, which always pertains when one considers finite-dimensional distributions, there is the full scope of infinitely divisible or Lévy processes, and for this the special structure of the D-spaces is essential. A recent example of this for $D[0, 1]$ is the CLT for stochastic integrals by Giné and Marcus (1983). For the general space $\boldsymbol{D}(\boldsymbol{A})$, the present authors (Bass and Pyke (1985)) obtained CLT's for partial-sum processes in the domains of normal attraction of stable processes.

In this paper we derive a CLT for quite general sums of independent $\boldsymbol{D}(\boldsymbol{A})$-valued random set functions where the limits may also be discontinuous. We consider the case where each X_{nj} is a random set function defined on a large family $\boldsymbol{A}$ of Borel subsets of the unit cube $I^d = [0, 1]^d$, $d \geq 1$ and taking values in $\boldsymbol{D}(\boldsymbol{A})$, the space of set functions having "inner limits and outer continuity"; see (2.1) below. This space contains the set of continuous set functions $\boldsymbol{C}(\boldsymbol{A})$ and is a generalization of it in the same way that $D[0, 1]$ generalizes $C[0, 1]$. Triangular arrays $\{X_{nj}: 1 \leq j \leq n, n \geq 1\}$ having independence within rows are considered. The choice of $m_n = n$ summands in the n-th row leads to no loss of generality; notice that we do not assume that the summands are infinitesimal but rather assume directly that the finite dimensional distributions converge.

The necessary definitions, notation and preliminary properties are set out in Section 2. In Section 3, the concept of a subpoissonian r.v. is introduced and a Bernstein-like bound is derived for the tail probabilities of such r.v.'s. This concept is key to the paper since we assume that the stratified components of the summands are subpoissonian.

The processes under consideration are allowed to have purely atomic discontinuities. The proof of the CLT is divided into two parts. In Section 4 we obtain the necessary tightness result for the processes from which all atoms whose absolute mass exceeds a specified level have been removed. We refer to this as the "small" atom case. The processes made up of the remaining "large" atoms are handled in Section 5. In the former, the sup norm and uniform topology is adequate, while for the latter particular compact subsets of purely atomic set functions must be used. The main result, Theorem 6.1, is given in Section 6. In this section we also give applications to Lévy processes, partial-sum processes and generalized empirical processes. Example 6.1 shows how our theorem lead to a construction of set-indexed

Lévy processes, giving a new proof of the results of Adler and Feigin (1984) and Bass and Pyke (1984). Example 6.2 applies our theorem to give a central limit theorem for set-indexed partial sum processes where the summands are in the normal domain of attraction of a stable law of index α, $\alpha \in (1, 2)$; this gives a new proof of the results of Bass and Pyke (1985), Section 5. In this context, notice that the classical one-dimensional partial-sum process can be expressed as a normalized sum of set functions $X_j \delta_{j/n}$ in contrast to the empirical process defined in terms of δ_{X_i}, where δ_x is point mass at x. This illustrates how the two types of processes are unified once one considers triangular arrays. The two types of processes can be classified according as to whether the atoms of the set functions have random masses at fixed locations (partial-sums) or fixed masses at random locations (empirical). This then suggests the case of random masses at random locations, such as $V_j \delta_{X_j}$ for example. When the mass V_j is in the domain of attraction of a stable law, this leads to an interesting application of the CLT, Example 6.3, in which the limit is non-Gaussian.

We have primarily considered the discontinuous elements of $D(\mathbf{A})$. In Section 7, we show how existing results on CLT's for empirical processes on $C(\mathbf{A})$-valued processes can be combined with our main theorem.

The CLT of Section 6 considers processes that take values in the subfamily $D_0(\mathbf{A})$ of $D(\mathbf{A})$ which is the closure of the linear span of all continuous or purely atomic members of $D(\mathbf{A})$. In Section 8, we introduce the larger subfamily $D_{SM}(\mathbf{A})$ which is the closure of the linear span of all members of $D(\mathbf{A})$ that are either continuous or the restriction of a signed measure. Two examples of families of compact subsets of $D_{SM}(\mathbf{A})$ are given which should be suitable for many applications.

In the case where $\mathbf{A}$ is the class of intervals $[0, t]$, $0 \le t \le 1$, $D(\mathbf{A})$ becomes $D[0, 1]$. The CLT of this paper gives new results even for this case.

2. Preliminaries

Given a Borel set $A \subset I^d = [0, 1]^d$, let A^0 be the interior of A with respect to the relative topology on I^d, and let $A^\delta = \{t \in I^d$: the Euclidean distance of t to A is $< \delta\}$ be the open δ-neighborhood around A. Define the Hausdorff metric by

$$d_H(A, B) = \inf\{\varepsilon : A \subset B^\varepsilon \text{ and } B \subset A^\varepsilon\}.$$

We will assume $\mathbf{A}$ is a collection of closed subsets of I^d satisfying the following:

Hypothesis A. (i) $\mathbf{A}$ is closed with respect to d_H.

(ii) for each $\delta > 0$, there is a finite subset $\mathbf{A}_\delta$ of $\mathbf{A}$ such that whenever $A \in \mathbf{A}$ there exists $B \in \mathbf{A}_\delta$ with $A \subset B^0 \subset B \subset A^\delta$.

Hypothesis A implies that $\mathbf{A}$ is totally bounded with inclusion with respect to d_H and that $\mathbf{A}$ is compact.

We now proceed to define $\boldsymbol{D}(\boldsymbol{A})$, $\boldsymbol{D}_0(\boldsymbol{A})$, and d_0. See Bass and Pyke (1985) for further details.

Definition 2.1. A function $x: \boldsymbol{A} \to \mathbb{R}$ is outer continuous with inner limits if

(i) $A, A_n \in \boldsymbol{A}$, $A \subset A_n$, $d_H(A, A_n) \to 0$ implies $x(A_n) \to x(A)$.

(ii) $A, A_n \in \boldsymbol{A}$, $A_n \subset A^0$, $d_H(A, A_n) \to 0$ implies $\lim_{n\to\infty} x(A_n)$ exists.

Let

$$\boldsymbol{D}(\boldsymbol{A}) = \{x: \boldsymbol{A} \to \mathbb{R}: x \text{ is outer continuous with inner limits}\}. \tag{2.1}$$

Let d_G be the Hausdorff metric on $\boldsymbol{A} \times \mathbb{R}$. Define the graph function $G: \boldsymbol{D}(\boldsymbol{A}) \to \boldsymbol{A} \times \mathbb{R}$ by letting $G(x)$ be the closure of $\{(A, x(A)): A \in \boldsymbol{A}\}$ with respect to d_G. We define the pseudometric d_D on $\boldsymbol{D}(\boldsymbol{A})$ by

$$d_D(x, y) = d_G(G(x), G(y)).$$

Let us say that x is *purely atomic* if there exist finitely many reals $a_1, \dots, a_n$ (the size of the atoms) and $t_1, \dots, t_n \in I^d$ (the locations of the atoms) such that

$$x(A) = \sum_{t_i \in A} a_i \quad \text{for all } A \in \boldsymbol{A}.$$

For such x, let

$$\operatorname{gap}(x) = \inf_{i \neq j} |t_i - t_j|, \qquad \operatorname{Variation}(x) = \sum_i |a_i|.$$

Let

$$\boldsymbol{D}_0(\boldsymbol{A}) = \{x: \boldsymbol{A} \to \mathbb{R}: \text{there exist purely atomic } j_m, \text{ continuous } c_m \text{ on } \boldsymbol{A} \text{ such that } \|x - (j_m + c_m)\|_{\boldsymbol{A}} \to 0 \text{ as } m \to \infty\}.$$

Here $\|\cdot\|_{\boldsymbol{A}}$ is the sup norm for functions on $\boldsymbol{A}$.

Two of the main results of Bass and Pyke (1985) are

Proposition 2.1. *Suppose η and R are reals, and $h(\cdot)$ and $N(\cdot)$ are real functions. Let $F_{PA}(h, N, \eta, R)$ be the set of purely atomic x such that*

(i) $\operatorname{gap}(x) \geq \eta$,

(ii) Variation $(x) \leq R$,

(iii) *for each δ, there exist sets $A_1, \dots, A_{N(\delta)} \in \boldsymbol{A}$ so that*

  (a) *every point of $G(x)$ is within δ (with respect to d_G) of some $(A_i, x(A_i))$, $i = 1, \dots, N(\delta)$, and*

  (b) *$A_i^{h(\delta)} \setminus A_i$ contains no atoms of x.*

Then $F_{PA}(h, N, \eta, R)$ is compact relative to d_D.

Proposition 2.2. *Suppose $\Delta_m \downarrow 0$, η_m, R_m, M_m are reals, and h_m, N_m, ω_m are functions on the positive reals with $\omega_m(r) \to 0$ as $r \to 0$. Let $F_{D_0}(h_m, N_m, \eta_m, R_m, \omega_m, M_m, \Delta_m)$ be the set of $x \in D_0(A)$ such that for each m there exist $J_m(x) \in F_{PA}(h_m, N_m, \eta_m, R_m)$ and continuous $C_m(x)$ with*

(i) (a) $\|C_m(x)|_A \leq M_m$,

(b) $\sup_{A,B \in A, d_H(A,B)<r} |C_m(x)(B) - C_m(x)(A)| \leq \omega_m(r)$ *for all r,*

(ii) $\|x - (J_m(x) + C_m(x))\|_A \leq \Delta_m$.

Then $F_{D_0}(h_m, N_m, \eta_m, R_m, \omega_m, M_m, \Delta_m)$ is compact relative to d_D.

For $\delta > 0$ and ν a measure, let $A(\delta, \nu)$ be the smallest (in cardinality) collection of subsets of I^d such that whenever $B \in A$, there exist $A, A^+ \in A(\delta, \nu)$ with $A \subset B \subset A^+$ and $\nu(A^+ \setminus A) < \delta$. Let $H(\delta, \nu)$, the log entropy, be the logarithm of the cardinality of $A(\delta, \nu)$. In the special case when ν is Lebesgue measure, we will denote the log entropy by $H_L(\delta)$.

We now turn to the random elements to which our central limit theorem pertains. We consider a triangular array $\{X_{nj}(\cdot): n = 1, 2, \ldots, j = 1, \ldots, n\}$ of elements of $D_0(A)$. With little or no loss of generality, we may assume A contains all singletons $\{t\}$, so that there is no ambiguity as to what is meant by an atom of X_{nj}. For $A \in A$, $J \subset \mathbb{R}$, let $Y_{nj}(A, J)$ be the sum of those atoms of X_{nj} whose location is in A and whose size is in J. Thus $Y_{nj}(\cdot J)$ represents a stratification of $X_{nj}(\cdot)$. Let $N_{nj}(A, J)$ be the number of atoms involved in this definition of $Y_{nj}(A, J)$. We center Y_{nj} by letting

$$X_{nj}(A, J) = Y_{nj}(A, J) - EY_{nj}(A, J \cap [-1, 1]),$$

so that if $J \subset [-1, 1]$, then $EX_{nj}(A, J) = 0$. Let

$$S_n(A, J) = \sum_{j=1}^{n} X_{nj}(A, J), \qquad S_n(A) = \sum_{j=1}^{n} X_{nj}(A),$$

$$N_n(A, J) = \sum_{j=1}^{n} N_{nj}(A, J).$$

We will interpret $X_{nj}(\cdot, J)$ as a stratification of $X_{nj}(\cdot)$, and similarly for S_n.

The assumptions on the triangular array are contained in the following:

Hypothesis B. (i) For each n, the $X_{nj}(\cdot)$, $j = 1, \ldots, n$, are independent,

(ii) the finite dimensional distributions of $S_n(\cdot)$ converge,

(iii) for each n, j, $X_{nj}(\cdot) = X_{nj}(\cdot, \mathbb{R} \setminus \{0\})$ on A,

(iv) for each n,

$$P(X_{nj_1} \text{ and } X_{nj_2} \textit{ have an atom in the same location for some } 1 \leq j_1 < j_2 \leq n) = 0,$$

(v) for each y, $\{E \sum_{j=1}^{n} Y_{nj}(\cdot, [-y, y]), n = 1, 2, \cdots\}$ is an equicontinuous family of continuous functions on A.

Conditions B(i), (ii) are the critical ones. Condition B(iii) says two things; first that the $X_{nj}(\cdot)$ are centered based on the atoms of size ≤ 1 in absolute value, and second, that $X_{nj}(\cdot)$ has no nontrivial continuous component. For (ii) to hold, the X_{nj}'s have to be centered in some fashion anyway, while the case where the X_{nj} have a continuous component is considered in Section 7. If (iv) is not satisfied, it can be circumvented using the method of Section 7; (v) will be trivially satisfied in most cases.

3. Subpoissonian random variables

By analogy to the word subgaussian, we introduce the term *subpoissonian*. (Further motivation is provided by the authors' surnames.)

Definition 3.1. A r.v. X is subpoissonian with parameters (θ, b) if, for all s,

$$E\,e^{sX} \leq \exp(\theta(e^{bs}+e^{-bs}-2)).$$

Obviously, if X_i, $i=1,\dots,n$, are independent and subpoissonian with parameters (θ_i, b), then $\sum_{i=1}^n X_i$ will be subpoissonian with parameters $(\sum_{i=1}^n \theta_i, b)$.

The symmetrization of a r.v. X is given by $X-X'$ where X' is an independent copy of X. Clearly, if X is subpoissonian with parameters (θ, b), then $X-X'$ is subpoissonian with parameters $(2\theta, b)$.

The following lemma provides a criterion for a r.v. to be subpoissonian. We use the convention that $\sum_{i=1}^0=0$.

Lemma 3.1. *Suppose Y_i, $i=1,2,\dots$, are independent symmetric r.v.'s that are bounded by b in absolute value. Suppose N is a nonnegative integer-valued r.v. that is independent of the Y_i's and is stochastically smaller than a Poisson (θ) r.v. Let $X=\sum_{i=1}^N Y_i$. Then X is subpoissonian with parameters $(\theta/2, b)$.*

Proof. First of all, because of symmetry we can write Y_i as $|Y_i|\varepsilon_i$, where ε_i is independent of Y_i and is $+1$ or -1, each with probability $\frac{1}{2}$. Since e^x+e^{-x} is increasing in x for $x>0$,

$$\varphi_i(s) := E\,e^{sY_i} = E\,e^{s\varepsilon_i|Y_i|} = E(e^{s|Y_i|}+e^{-s|Y_i|})/2 \leq (e^{sb}+e^{-sb})/2. \tag{3.1}$$

By enlarging the probability space if necessary, we may assume that there exists a Poisson (θ) r.v. W that is independent of the Y_i's such that $N\leq W$. Since $(e^x+e^{-x})/2\geq 1$ for all x,

$$E\,e^{sX} = E\prod_{i=1}^{N}\varphi_i(s) \leq E[(e^{bs}+e^{-bs})/2]^N \leq E[e^{bs}+e^{-bs})/2]^W. \tag{3.2}$$

By a straightforward calculation, the last expression is bounded by

$$\exp\left(\frac{\theta}{2}(e^{bs}+e^{-bs}-2)\right). \qquad \square$$

A corollary of the above lemma that will be useful in applications is

Corollary 3.2. *Suppose in Lemma* 3.1 *that* N *is a Binomial* (n, p) *r.v. Then* X *is subpoissonian* $(np/2, b)$.

Proof. Letting $r = \ln((e^{bs} + e^{-bs})/2) \geq 0$, by (3.2) we have

$$\begin{aligned} E\, e^{sX} \leq E\, e^{rN} &= ((1-p) + p\, e^r)^n = \exp(n \ln(1 + p(e^r - 1))) \\ &\leq \exp(np(e^r - 1)) \\ &= \exp\left(\frac{np}{2}(e^{bs} + e^{-bs} - 2)\right). \quad \square \end{aligned}$$

The importance of the subpoissonian concept is due to the following Bernstein-like estimate that applies to subpoissonian r.v.'s.

Proposition 3.3. *If* X *is subpoissonian with parameters* (θ, b), *then*

$$P(X \geq \lambda) \leq \exp\left(-\frac{\lambda^2}{4\theta b + 2\lambda b/3}\right).$$

Proof (cf. Bass and Pyke (1984)). If we let $s = \lambda/(2\theta b^2 + b\lambda/3)$, then

$$bs/3 < 1, \quad (1 - bs/3)^{-1} = (2\theta b + \lambda/3)/2\theta b,$$

and

$$\begin{aligned} e^{bs} + e^{-bs} - 2 &= b^2 s^2 \left(1 + \frac{b^2 s^2}{4 \cdot 3} + \cdots\right) \\ &\leq b^2 s^2 (1 - bs/3)^{-1} = \frac{bs^2}{2\theta}(2\theta b + \lambda/3). \end{aligned}$$

Then, by Chebychev,

$$\begin{aligned} P(X \geq \lambda) &\leq e^{-s\lambda} E\, e^{sX} \leq \exp(\theta(e^{bs} + e^{-bs} - 2) - s\lambda) \\ &\leq \exp\left(\frac{bs^2}{2}(2\theta b + \lambda/3) - s\lambda\right) \\ &= \exp\left(\frac{-\lambda^2}{4\theta b + 2\lambda b/3}\right) \quad \square \end{aligned}$$

4. Small atoms

In this section we obtain a bound on $\|S_n(\cdot, [-a, a])\|_A$. For each $n \geq 1, j = 1, \ldots, n$, $y \in (0, 1)$, $\nu_{nj}(\cdot, y)$ will be a measure on I^d. Set $\nu_n(A, y) = \sum_{j=1}^n \nu_{nj}(A, y)$. We assume

Hypothesis C. For any $0 < y < z \leq 1$ and $J \subset [-z, z] \backslash [-y, y]$,

(i) the symmetrization of $X_{nj}(A, J)$ is subpoissonian with parameters $(\nu_{nj}(A, y), z)$.

(ii) if $s \geq 0$, then $E \exp(sN_{nj}(A, J)) \leq \exp(\nu_{nj}(A, y)(e^s - 1))$,

(iii) $\sup_n y^2 \nu_n(I^d, y) = o(|\ln y|^{-(2+\tau)})$ for some $\tau > 0$ as $y \downarrow 0$,

(iv) $\zeta := 1 + \sup_{n,y,x} H(x, \nu_n(\cdot, y))/H(2x, \nu_n(\cdot, y)) < \infty$,

(v) $$\lim_{\gamma \to 0} \sup \sup_n \int_0^\gamma \int_{G_n^{-1}(2\zeta^{1/2}, y/2)/2}^{\nu_n(I^d, y/2) \vee 1} G_n(x, y)\, dx\, dy = 0,$$

where

$$G_n(x, y) = [H(x, \nu_n(\cdot, y))/x]^{1/2}$$

and G_n^{-1} refers to the inverse with respect to x with y fixed, and

(vi) $$\lim_{y \downarrow 0} \sup \sup_{A \in \mathcal{A}} \sup_n \sum_{j=1}^n \operatorname{Var}(X_{nj}(A, [-y, y])) = 0.$$

Conditions (i), (ii), and (v) are the important ones; the others are needed for technical purposes. If $N_{nj}(A, J)$ is stochastically smaller than a Poisson $(\nu_{nj}(A, y))$ r.v., then (ii) will be satisfied. The purpose of (iii) is to eliminate the case where the atoms of the X_{nj} are too concentrated about 0; (iv) says that the log-entropy H is regularly varying, uniformly in y. Condition (v) gives a bound on the size of H, and (vi) is a condition that will be easily satisfied in most applications.

For any $J \subset R$, let $\alpha_J = \sup\{|y|: y \in J\}$, and define

$$S_n^+(A, J) = \alpha_J N_n(A, J) + \alpha_J E N_n(A, J). \tag{4.1}$$

It should be clear that for fixed A and A^+,

$$\sup_{A \subset B \subset A^+} |S_n(B, J) - S_n(A, J)| \leq S_n^+(A^+ \backslash A, J). \tag{4.2}$$

Suppose Hypothesis C(ii) holds. Then if $J \subset [-y, y]$,

$$E \exp(sN_n(A, J)) \leq \prod_{j=1}^n \exp(\nu_{nj}(A, y)(e^s - 1)) = \exp(\nu_n(A, y)(e^s - 1)). \tag{4.3}$$

Since $N_n(A, J)$ is nonnegative and has a moment generating function, then

$$E N_n(A, J) = \lim_{s \to 0} s^{-1}(E \exp(sN_n(A, J)) - 1) \leq \nu_n(A, y). \tag{4.4}$$

Using the proof of Bass and Pyke (1984), Lemma 2.2, we have

$$P(N_n(A, J) > w) \leq e^{-w},$$

provided $w \geq e^2 \nu(A, y)$. Hence, if we let $w = 9\lambda/10\alpha_J$, we have

$$P(S_n^+(A, J) > \lambda) \leq P(N_n(A, J) > 9\lambda/10\alpha_J) \leq \exp(-\lambda/2\alpha_J), \tag{4.5}$$

provided

$$\lambda/\alpha_J \geq 10\nu_n(A, y). \tag{4.6}$$

We now prove

Proposition 4.1. *Suppose Hypothesis C holds. For $\lambda > 0$, $\varepsilon > 0$, there exists $\alpha \in (0, 1)$ such that if $a < \alpha$,*

$$\sup_n P(\|S_n(\cdot, [-a, a]\|_{\mathcal{A}} > \lambda) < \varepsilon.$$

Proof. Since $EX_{nj}(A, [-y, y]) = 0$ for $y < 1$, using C(vi) gives

$$P(|S_n(A, [-a, a])| > \lambda/2) \leq 4 \operatorname{Var} S_n(A, [-a, a])/\lambda^2 < \tfrac{1}{2}$$

for a sufficiently small. Then by the symmetrization lemma of Pollard (1984, p. 14),

$$P(\|S_n(\cdot, [-a, a])\|_{\mathcal{A}} > \lambda) \leq 2P(\|(S_n - S_n')(\cdot, [-a, a])\|_{\mathcal{A}} > \lambda/2),$$

where S_n' is an independent copy of S_n. It thus suffices to prove the proposition with the X_{nj}'s symmetric (doubling ν_n and halving λ, as necessary).
Let $\beta = \frac{1}{2}$. We will choose numbers δ_{k0}, λ_k, λ_{ki}, s_k and K later so that

$$\sum_{i=1}^{s_k} \lambda_{ki} \leq \lambda_k/4 \tag{A}$$

and

$$\sum_{k=K}^{\infty} \lambda_k \leq \lambda. \tag{B}$$

Let $\delta_{ki} = \delta_{k0}\beta^i$. Once K is chosen, we will let $\alpha = \beta^K$; fix $a < \beta^K$, let $a_k = \beta^k \wedge \beta^K$, and let

$$J_k = [-a_k, -a_{k+1}) \cup (a_{k+1}, a_k].$$

Let us temporarily abbreviate $H(\delta, \nu_n(\cdot, a_{k+1}))$ by $H_{nk}(\delta)$ and $A(\delta, \nu_n(\cdot, a_{k+1}))$ by $A_{nk}(\delta)$. Fix n.

If $B \in \mathcal{A}$, we can find sets $A_i, A_i^+ \subset A_{nk}(\delta_{ki})$ such that $A_i \subset B \subset A_i^+$ and $\nu_n(A_i^+ \setminus A_i, a_{k+1}) < \delta_{ki}$. Writing

$$S_n(B, J_k) = S_n(A_0, J_k) + \sum_{i=1}^{s_k} [S_n(A_i, J_k) - S_n(A_{i-1}, J_k)]$$
$$+ [S_n(B, J_k) - S_n(A_{s_k}, J_k)]$$

and using the facts that

$$\nu_n(A_i \Delta A_{i-1}, a_{k+1}) \leq \nu_n(A_i \Delta B, a_{k+1}) + \nu_n(A_{i-1} \Delta B, a_{k+1})$$
$$\leq 2\delta_{k,i-1} \leq 4\delta_{ki}, \tag{4.7}$$

$$|S_n(A_i, J_k) - S_n(A_{i-1}, J_k)| \leq |S_n(A_i \backslash A_{i-1}, J_k)| + |S_n(A_{i-1} \backslash A_i, J_k)|, \tag{4.8}$$

and

$$|S_n(B, J_k) - S_n(A_{s_k}, J_k)| \leq S_n^+(A_{s_k}^+ \backslash A_{s_k}, J_k), \tag{4.9}$$

we see that the only way $\|S_n(\cdot, J_k)\|_{\mathcal{A}}$ can be larger than λ_k is if either

(i) for some $A_0 \in \mathcal{A}_{nk}(\delta_{k0})$, we have $|S_n(A_0, J_k)| > \lambda_k/4$;

(ii) for some $i = 1, \ldots, s_k$, some $A_i \in \mathcal{A}_{nk}(\delta_{ki})$, $A_{i-1} \in \mathcal{A}_{nk}(\delta_{k,i-1})$ with $\nu_n(A_i \Delta A_{i-1}, a_{k+1}) < 4\delta_{ki}$, we have $|S_n(A_i \backslash A_{i-1}, J_k)| > \lambda_{ki}$ or $|S_n(A_{i-1} \backslash A_i, J_k)| > \lambda_{ki}$; or

(iii) for some $A_{s_k}, A_{s_k}^+ \in \mathcal{A}_{nk}(\delta_{ks_k})$ with $\nu_n(A_{s_k}^+ \backslash A_{s_k}, a_{k+1}) < \delta_{ks_k}$, we have $S_n^+(A_{s_k}^+ \backslash A_{s_k}, J_k) > \lambda_k/4$.

Since the cardinality of $\mathcal{A}_{nk}(\delta_{k,i-1}) \times \mathcal{A}_{nk}(\delta_{ki})$ is bounded by $\exp(2H_{nk}(\delta_{ki}))$, we then have

$$P(\|S_n(\cdot, J_k)\|_{\mathcal{A}} > \lambda_k) \leq p_k + \sum_{i=1}^{s_k} q_{ki} + r_k, \tag{4.10}$$

where

$$p_k = \exp(H_{nk}(\delta_{k0})) \sup_{A_0 \in \mathcal{A}_{nk}(\delta_0)} P(|S_n(A_0, J_k)| > \lambda_k/4); \tag{4.11}$$

$$q_{ki} = 2\exp(2H_{nk}(\delta_{ki})) \sup P(|S_n(B \backslash A, J_k)| > \lambda_{ki}), \tag{4.12}$$

the sup being over pairs A, B with one in $\mathcal{A}_{nk}(\delta_{ki})$, the other in $\mathcal{A}_{nk}(\delta_{k,i-1})$, and $\nu_n(A \Delta B, a_{k+1}) < 4\delta_{ki}$,

and

$$r_k = \exp(2H_{nk}(\delta_{ks_k})) \sup P(S_n^+(B \backslash A, J_k) > \lambda_k/4), \tag{4.13}$$

the sup being over pairs A, B in $\mathcal{A}_{nk}(\delta_{ks_k})$ with $\nu_n(B \backslash A, a_{k+1}) < \delta_{ks_k}$.

Since

$$\|S_n(\cdot, [-a, a])\|_{\mathcal{A}} \leq \sum_{k=K}^{\infty} \|S_n(\cdot, J_k)\|_{\mathcal{A}},$$

and hence

$$P(\|S_n(\cdot, [-a, a])\|_{\mathcal{A}} > \lambda) \leq \sum_{k=K}^{\infty} P(\|S_n(\cdot, J_k)\|_{\mathcal{A}} > \lambda_k)$$
$$\leq \sum_{k=K}^{\infty} \left(p_k + \sum_{i=1}^{s_k} q_{ki} + r_k \right),$$

we will require

$$\sum_{k=K}^{\infty} p_k < \varepsilon/3, \tag{C}$$

$$\sum_{k=K}^{\infty} \sum_{i=1}^{s_k} q_{ki} < \varepsilon/3, \tag{D}$$

and

$$\sum_{k=K}^{\infty} r_k < \varepsilon/3, \tag{E}$$

with K chosen independently of n.

We now proceed to estimate p_k, q_{ki}, and r_k. By Hypothesis C(iii),

$$\sup_n a_k^2 \nu_n(I^d, a_{k+1}) \leq k^{-(2+\tau)} \tag{4.14}$$

for some $\tau > 0$ independent of n and for all k sufficiently large. We will thus require

$$\lambda_{ki} \geq \beta^{1/2} k^{-(1+\tau/3)} \tag{F}$$

By (4.11) and Proposition 3.3,

$$\begin{aligned} p_k &\leq 2 \exp(H_{nk}(\delta_{k0})) \sup_{A_0, \cdot A_{nk}(\delta_0)} \exp\left(\frac{-(\lambda_k/4)^2}{4a_k^2 \nu_n(A_0, a_{k+1}) + 2a_k(\lambda_k/4)/3}\right) \\ &\leq 2 \exp(H_{nk}(\delta_{k0})) \left[\exp\left(\frac{-\lambda_k^2}{128 a_k^2 \nu_n(I^d, a_{k+1})}\right) + \exp\left(\frac{-\lambda_k}{6a_k}\right)\right]. \end{aligned}$$

Provided

$$H_{nk}(\delta_{k0}) \leq \lambda_k / 12 a_k \tag{G}$$

and

$$H_{nk}(\delta_{k0}) \leq \lambda_k^2 / 256 a_k^2 \nu_n(I^d, a_{k+1}), \tag{H}$$

we get

$$p_k \leq 2 \exp\left(\frac{-\lambda_k^2}{256 a_k^2 \nu_n(I^d, a_{k+1})}\right) + 2 \exp\left(\frac{-\lambda_k}{12 a_k}\right).$$

By the definition of a_k, (4.14), and (F), we see that (C) is satisfied provided K is taken sufficiently large.

Next we look at q_{ki}. By (4.12) and Proposition 3.3,

$$q_{ki} \leq 4 \exp(2H_{nk}(\delta_{ki})) \exp\left(\frac{-\lambda_{ki}^2}{16 a_k^2 \delta_{ki} + 2a_k \lambda_{ki}/3}\right).$$

Recall from the definition of ζ in Hypothesis C(iv) that $\zeta \geq 1$. Provided

$$\lambda_{ki} \leq c_1 a_k \delta_{ki}, \quad i = 1, \ldots, s_k, \; c_1 = 80\zeta, \tag{I}$$

and

$$H_{nk}(\delta_{ki}) \le \frac{\lambda_{ki}^2}{c_2 a_k^2 \delta_{ki}}, \quad i=1,\ldots,s_k, \ c_2 = 200\zeta, \tag{J}$$

then

$$q_{ki} \le 4\exp(2H_{nk}(\delta_{ki}))\exp\left(\frac{-\lambda_{ki}^2}{(16+c_1)a_k^2\delta_{ki}}\right) \le 4\exp(-\lambda_{ki}^2/c_2 a_k^2 \delta_{ki}).$$

Provided

$$a_k^2 \delta_{k0} k^{2+\tau} \le c_3, \quad \text{where } c_3 \text{ is independent of } n, \tag{K}$$

then, using (F),

$$\begin{aligned}\sum_{k=K}^{\infty}\sum_{i=1}^{s_k} q_{ki} &\le 4\sum_{k=K}^{\infty}\sum_{i=1}^{\infty}\exp(-(c_2 a_k^2 \delta_{k0}\beta^{i/2}k^{2+2\tau/3})^{-1}) \\ &\le 4\sum_{k=K}^{\infty}\sum_{i=1}^{\infty}\exp\left(\frac{-i(\beta^{-1/2}-1)}{c_2 a_k^2 \delta_{k0} k^{2+2\tau/3}}\right) \\ &\le 4\sum_{k=K}^{\infty}\sum_{i=1}^{\infty}\exp\left(\frac{-i(\beta^{-1/2}-1)k^{\tau/3}}{c_2 c_3}\right),\end{aligned} \tag{4.15}$$

and (D) will be satisfied if K is taken sufficiently large.

Finally, by (4.5) and (4.13),

$$r_k \le \exp(2H_{nk}(\delta_{ks_k}))\exp(-\lambda_k/8a_k)$$

provided

$$\lambda_k/4a_k \ge 10\delta_{ks_k}. \tag{L}$$

Provided

$$H_{nk}(\delta_{ks_k}) \le \lambda_k/32a_k, \tag{M}$$

we get

$$r_k \le \exp(-\lambda_k/16a_k),$$

and using (F), (E) is satisfied if K is large.

Now define

$$\begin{aligned}&\lambda_{ki} = \max(\beta^{i/2}k^{-(1+\tau/3)}, c_4 H_{nk}^{1/2}(\delta_{ki})a_k\delta_{ki}^{1/2}), \quad c_4 = 40\zeta^{1/2}, \\ &s_k = \inf\{i: \lambda_{ki} \ge 40a_k\delta_{ki}\}, \qquad \lambda_k = \max\left(4\sum_{i=1}^{s_k}\lambda_{ki}, k^{-2}\right),\end{aligned} \tag{4.16}$$

and

$$\delta_{k0} = \max(\nu_n(I^d, a_{k+1}), 1)$$

(These quantities depend on n through H_{nk} and ν_n.) Note that for fixed k, $40a_k\delta_{ki} = 40a_k\delta_{k0}\beta^i$ will be less than $\beta^{i/2}k^{-(1+\tau/3)} \le \lambda_{ki}$ for i large, and so $0 \le s_k < \infty$.

From these definitions, we immediately have that (A) and (F) hold. Since $c_4^2 \geq c_2$, (J) follows, and since $\lambda_k \geq \lambda_{ks_k}$, (L) follows. We have (K) from the definition of δ_{k0} and Hypothesis C.

Clearly $\{\phi, I^d\}$ will serve for $A_{nk}(\nu_n(I^d, a_{k+1}))$, and so $H_{nk}(\delta_{k0}) = \ln 2$. Hence (G) and (H) follow from (F), (K), and the definition of δ_{k0}. Since $H_{nk}(x)$ is decreasing, $H_{nk}(\delta_{ks_k}) \leq H_{nk}(\delta_{k0})$, and (M) follows just as we showed (G).

We now consider (I). If $s_k = 0$, (I) is unnecessary. So suppose $s_k > 0$. Then

$$\lambda_{ks_k} \leq \left(\frac{H_{nk}(\delta_{ks_k})}{H_{nk}(\delta_{k,s_k-1})}\right)^{1/2} \lambda_{k,s_k-1} \leq 40\zeta a_k \delta_{k,s_k-1} = 80\zeta a_k \delta_{ks_k}, \tag{4.17}$$

and (I) is satisfied. (The first inequality of (4.17) is trivially satisfied if $\lambda_{ks_k} = \beta^{s_k/2} k^{-(1+\tau/3)}$, since $H_{nk}(x)$ is decreasing.)

It remains to show (B), and since $\sum_k \sum_i \beta^{i/2} k^{-(1+\tau/3)} < \infty$, it suffices to show

$$S_1 = \sum_{k=K}^{\infty} H_{nk}^{1/2}(\delta_{k0}) a_k \delta_{k0}^{1/2}$$

and

$$S_2 = \sum_{\{k:\, k \geq K, s_k > 0\}} \sum_{i=0}^{s_k} H_{nk}^{1/2}(\delta_{ki}) a_k \delta_{ki}^{1/2}$$

can both be made small by taking K large, uniformly in n. We can make S_1 small by using (K) and the fact that $H_{nk}(\delta_{k0}) = \ln 2$. To handle S_2, first note by (I) that if $s_k > 0$,

$$c_4 H_{nk}^{1/2}(\delta_{ks_k}) a_k \delta_{ks_k}^{1/2} \leq \lambda_{ks_k} \leq 80\zeta a_k \delta_{ks_k}.$$

Hence $H_{nk}^{1/2}(\delta_{ks_k}) / \delta_{ks_k}^{1/2} \leq 80\zeta / c_4 = 2\zeta^{1/2}$, and hence $\delta_{ks_k} \geq G_n^{-1}(2\zeta^{1/2}, a_k)$. Then

$$\begin{aligned} S_2 &\leq \sum_{\{k:\, k \geq K, s_k > 0\}} \sum_{i=0}^{s_k} G_n(\delta_{ki}, a_k) \delta_{ki} a_k \leq 2 \sum_{\{k:\, k \geq K, s_k > 0\}} \int_{\delta_{k,s_k+1}}^{\delta_{k0}} G_n(x, a_k)\, \mathrm{d}x\, a_k \\ &\leq 2 \sum_{k=K}^{\infty} \int_{G_n^{-1}(2\zeta^{1/2}, a_k)/2}^{\nu_n(I^d, a_{k+1}) \vee 1} G_n(x, a_k)\, \mathrm{d}x\, a_k \\ &\leq 4 \sup_n \int_0^{\beta_{K+1}} \int_{G_n^{-1}(2\zeta^{1/2}, y/2)/2}^{\nu_n(I^d, y/2) \vee 1} G_n(x, y)\, \mathrm{d}x\, \mathrm{d}y. \end{aligned}$$

That S_2 can be made small by taking K large follows from Hypothesis (C)(v). □

5. Large atoms

In this section we consider the "large" atoms, those above a fixed given value. We assume the following:

Hypothesis D. For each $m = 1, 2, \ldots$, there exists a measure μ_m on $(I^d)^m \times \mathbb{R}$ such that

(i) for each $y > 0$, $\mu_m((I^d)^m \times [-y, y]^c) < \infty$, and

(ii) $\sup_n \sum_{j=1}^{n} P(X_{nj}(\cdot, [-y, y]^c)$ has at least one atom located in each of $A_1, A_2, \ldots, A_m)$

$$\leq \mu_m(A_1 \times A_2 \times \cdots \times A_m \times [-y, y]^c).$$

We first show that Hypothesis D implies the corresponding fact about S_n.

Proposition 5.1. Under Hypothesis D, there exists a measure σ_m on the Borel sets of $(I^d)^m \times \mathbb{R}$ for each $m = 1, 2, \ldots$ such that

(i) for each $y > 0$, $\sigma_m((I^d)^m \times [-y, y]^c) < \infty$ and

(ii) $\sup_n P(S_n(\cdot, [-y, y]^c)$ has at least one atom located in each of $A_1, \ldots, A_m)$

$$\leq \sigma_m(A_1 \times \cdots \times A_m \times [-y, y]^c).$$

Proof. For notational convenience, we prove the proposition for $m = 3$; the generalization to other m should then be obvious. We will fix y and let $J = [-y, y]^c$.

For $S_n(\cdot, J)$ to have at least one atom in each of A_1, A_2, and A_3, either

(i) some $X_{nj}(\cdot, J)$ has an atom in each of them;

(ii) some $X_{nj_1}(\cdot, J)$ has an atom in two of them, while some other $X_{nj_2}(\cdot, J)$ has an atom in the remaining set, or

(iii) for j_1, j_2, j_3 distinct, $X_{nj_i}(\cdot, J)$ has an atom in A_i, for $i = 1, 2, 3$.

The probability of (i) is bounded by

$$\sum_{j=1}^{n} P(X_{nj}(\cdot, J) \text{ has an atom in each of } A_1, A_2, A_3)$$
$$\leq \mu_3(A_1 \times A_2 \times A_3 \times J).$$

The probability of (ii) is less than the sum of three terms of the form

$$\sum_{j_1, j_2} P(X_{nj_1}(\cdot, J) \text{ has an atom in } A_1 \text{ and } A_2) P(X_{nj_2}(\cdot, J) \text{ has an atom in } A_3)$$
$$\leq \mu_2(A_1 \times A_2 \times J)\mu_1(A_3 \times J).$$

Similarly the probability of (iii) is bounded by

$$\sum_{j_1, j_2, j_3} \prod_{i=1}^{3} P(X_{nj_i}(\cdot, J) \text{ has an atom in } A_i)$$
$$\leq \mu_1(A_1 \times J)\mu_1(A_2 \times J)\mu_1(A_3 \times J).$$

If we define σ_3 so that

$$\begin{aligned}\sigma_3(A_1 \times A_2 \times A_3 \times \cdot) = \mu_3(A_1 \times A_2 \times A_3 \times \cdot) &+ \mu_2(A_1 \times A_2 \times \cdot)\mu_1(A_3 \times \cdot) \\ &+ \mu_2(A_1 \times A_3 \times \cdot)\mu_1(A_2 \times \cdot) \\ &+ \mu_2(A_2 \times A_3 \times \cdot)\mu_1(A_1 \times \cdot) \\ &+ \mu_1(A_1 \times \cdot)\mu_1(A_2 \times \cdot)\mu_1(A_3 \times \cdot),\end{aligned}$$

it is clear that σ_3 satisfies the proposition for $m = 3$. $\square$

With the help of Proposition 5.1, we can show that the atoms of S_n of size $\geq y$ in absolute value will be in a compact set with high probability. Fix y, and define

T_n by

$$T_n(A) = \sum_{j=1}^{n} Y_{nj}(A, [-y, y]^c). \tag{5.1}$$

Proposition 5.2. *There exist η, R, $N(\cdot)$, and $h(\cdot)$ such that for all n,*

$$P(T_n \notin F_{PA}(h, N, \eta, R)) < 4\varepsilon.$$

Proof. Fix n. Since $\sigma_m((I^d)^m \times [-y, y]^c)$ is finite for each m, $\sigma_1(I^d \times [-a, a]^c) \to 0$ as $a \to \infty$. So, for a sufficiently large,

$$P(T_n \text{ has an atom of size } > a \text{ in absolute value}) < \varepsilon. \tag{5.2}$$

Define

$$D_\eta = \{(s, t) \in (I^d)^2 : 0 < |s - t| < \eta\}.$$

As $\eta \to 0$, $D_\eta \downarrow \phi$, and hence $\sigma_2(D_\eta \times [-y, y]^c) \to 0$. Thus for η sufficiently small,

$$P(\text{gap}(T_n) < \eta) < \varepsilon. \tag{5.3}$$

If $\text{gap}(T_n) > \eta$, then there are at most $c_1\eta^{-d}$ atoms for a suitable constant c_1, and this fact and (5.2) show that

$$P(\text{Variation}(T_n) > R) < 2\varepsilon$$

if R is sufficiently large.

Now for $\delta = k^{-1}$, we want to show that we can find $N(\delta)$ and $h(\delta)$ such that the probability that there is not a δ-net satisfying Proposition (2.1)(iii) is $< \varepsilon/2^k$. Let $\boldsymbol{B}_q = \bigcup_{l \le q} \boldsymbol{A}_{1/l}$, where $\boldsymbol{A}_{1/l}$ is defined by Hypothesis A(ii). For $m = 1, \ldots, c_1\eta^{-d}$, let

$$W_{qm} = \{(t_1, \ldots, t_m) \in (I^d)^m : \text{for some purely atomic } x \text{ with one atom located at each of } t_1, \ldots, t_m, \{(A, x(A))\}_{A \in \boldsymbol{B}_q} \text{ is not a } \delta\text{-net for the graph } x\}.$$

By the proof of Theorem 4.4 of Bass and Pyke (1985), for each m, $W_{qm} \downarrow \phi$ as $q \to \infty$, and so $\sigma_m(W_{qm} \times [-y, y]^c) \to 0$. Restricting ourselves to the set $(\text{gap}(T_n) > \eta)$, where T_n has at most $c_1\eta^{-d}$ atoms, we then see that the probability that $\{(A, T_n(A))\}_{A \in \boldsymbol{B}_q}$ is not a δ-net for the graph of $T_n(A)$ can be made less than $\varepsilon/2^k$, uniformly in n, by taking q large enough. Define $N(\delta) = \#(\boldsymbol{B}_q)$.

Let $B_1, \ldots, B_{N(\delta)}$ be an enumeration of $\boldsymbol{B}_q$.

$$\begin{aligned} &P(T_n \text{ has an atom in } B_k^{h(\delta)} \setminus B_k \text{ for some } B_k \in \boldsymbol{B}_q) \\ &\quad \le N(\delta) \sup_{B_k \in \boldsymbol{B}_q} P(T_n \text{ has an atom in } B_k^{h(\delta)} \setminus B_k) \\ &\quad \le N(\delta) \sup_{B_k \in \boldsymbol{B}_q} \sigma_1(B_k^{h(\delta)} \setminus B_k \times [-y, y]^c) < \varepsilon/2^k \end{aligned}$$

if $h(\delta)$ is taken sufficiently small.

If $(k+1)^{-1} \le \delta < k^{-1}$, let $N(\delta) = N((k+1)^{-1})$, $h(\delta) = ((k+1)^{-1})$. With these choices of η, R, N, and h, we have proved the proposition. $\square$

6. Main theorem and examples

In this section we state and prove our main theorem and give some examples.

Theorem 6.1. *Suppose A and $\{X_{nj}\}$ satisfy Hypotheses* A–D. *Then $S_n \xrightarrow{w}$ to a $D_0(A)$-valued r.v., where $\xrightarrow{w}$ is with respect to d_D.*

Proof. By Hypotheses B(ii), the finite dimensional distributions converge, and so it suffices to prove tightness. Let $\varepsilon > 0$. Let $\Delta_m = m^{-1}$, and choose $y_m < 1$ sufficiently small so that

$$P(\|S_n(\cdot, [-y_m, y_m])\|_A \geq \Delta_m) < \varepsilon/2^{m+1}. \tag{6.1}$$

This is possible by Proposition 4.1.

Let

$$J_m(S_n) = \sum_{j=1}^{n} Y_{nj}(\cdot, [-y_m, y_m]^c),$$

$$C_m(S_n) = -E \sum_{j=1}^{n} Y_{nj}(\cdot, [-1, 1]\backslash[-y_m, y_m]),$$

so that $J_m(S_n)$ is purely atomic, and by Hypothesis $B(v)$, $C_m(S_n)$ is continuous. Note

$$S_n(\cdot) - (J_m(S_n) + C_m(S_n)) = S_n(\cdot, [-y_m, y_m]).$$

By Proposition 5.1, we can choose η_m small, R_m large, and suitable functions $N_m(\delta)$, $h_m(\delta)$ so that

$$P(J_m(S_n) \notin F_{PA}(h_m, N_m, \eta_m, R_m)) < \varepsilon/2^{m+1}.$$

By Hypothesis $B(v)$, for each n, $C_m(S_n)$, which is deterministic, is an equicontinuous family of functions; moreover, $C_m(S_n)$ is uniformly bounded since $C_m(S_n)(\phi) = 0$.

We therefore can find, using Proposition 2.2, a compact subset F of $D_0(A)$ such that $P(S_n \notin F) < \varepsilon$. This proves tightness, and by Theorem of 4.3 of Bass and Pyke (1985), we have weak convergence. □

Remark. In the case $d = 1$ and $A = ([0, t]: t \in [0, 1])$, where we write $x(t)$ for $x[0, t]$, our topology on $D(A)$ is very close to Skorokhod's M_2-topology. However, it is easy to see (cf. Billingsley (1968, p. 116)) that the F_{D_0} sets of Proposition 2.2 are compact with respect to the more common J_1-topology as well, and so we have weak convergence on $D[0, 1]$ in the usual sense, also. It is perhaps worth remarking that in this case $H_L(x) = O(\ln(1/x))$, a very small log-entropy.

Example 6.1. Set-indexed Lévy processes (cf. Bass and Pyke (1984), Adler and Feigin (1984); see the first reference for any details omitted in the discussion below). Let ρ be a measure on $\mathbb{R}\backslash\{0\}$ with $\int(x^2 \wedge 1)\rho(dx) < \infty$. Let $\rho_n = \rho|_{[-1,1]\backslash[-n^{-1}, n^{-1}]}$. Since

ρ_n is a finite measure, there is no difficulty constructing a mean 0 Levý process, X_{n1}, which has Levý measure ρ_n. For $2 \leq j \leq n$, let $X_{nj} \equiv 0$.

If $J = [-b, b] \setminus [-a, a]$, the moment generating function of the symmetrization of $X_{n1}(A, J)$ is

$$\exp(|A| \int_a^b (e^{sx} + e^{-sx} - 2)[\rho_n(dx) + \rho_n(-dx)])$$
$$\leq \exp(|A| \rho_n(J)(e^{sb} + e^{-sb} - 2)),$$

where $|\cdot|$ is Lebesgue measure, since $e^x + e^{-x}$ is increasing for $x > 0$. If we take $\nu_{n1}(A, y) = |A| \rho([-y, y]^c)$, we see that the symmetrization of $X_{n1}(A, J)$ will be subpoissonian with parameters $(\nu_{n1}(A, a), b)$.

Define μ_m by

$$\mu_m(A_1 \times \cdots \times A_m \times [-y, y]^c) = \prod_{i=1}^m \rho([-y, y]^c)|A_i|.$$

If the A_i's are disjoint,

$$P(X_{n1} \text{ has at least one atom located in each of } A_1, \ldots, A_m)$$
$$= \prod_{i=1}^m P(X_{n1} \text{ has at least one atom located in } A_i)$$
$$= \prod_{i=1}^m (1 - \exp(-\rho([y, y]^c)|A_i|)$$
$$\leq \mu_m(A_i \times \cdots \times A_m \times [-y, y]^c).$$

By linearity, we also have Hypothesis D where the A_i's are not disjoint.

Hypotheses C(iii)-(v) translate to restrictions on ρ and A; the other hypotheses can be routinely checked. By Theorem 6.1, $S_n = X_{n1}$ converges weakly, say to Z. The process Z will be a mean 0 Lévy process indexed by A with Lévy measure $\rho|_{[-1,1]}$. It is then easy to construct a Lévy process with ρ as its Lévy measure.

Suppose now that ρ is the Lévy measure of a stable process of index $\alpha \in (1, 2)$, and that H_L, the log entropy with respect to Lebesgue measure is of the form $K\delta^{-r}$. In what follows, $c_1, c_2, \ldots$ are constants whose exact values are of no importance. We note $\rho([-y, y]^c) = c_1 y^{-\alpha}$, and so $\nu_n(A, y) \leq c_1 y^{-\alpha}|A|$. In order that $\nu_n(A \triangle B, y) < \delta$, we need $|A \triangle B| \leq \delta / c_1 y^{-\alpha}$. Hence $H(\delta, \nu_n(\cdot, y)) \leq H_L(\delta / c_1 y^{-\alpha})$. A simple calculation gives

$$G(x, y) \leq c_2 x^{-(r+1)/2} y^{-\alpha r/2}.$$

Hypothesis C(iv) is satisfied with $\zeta \leq 1 + 2^{1+r}$, while C(iii) is obvious.

Another calculation shows $G^{-1}(c_3, y) \geq (c_4 y^{\alpha r/2})^{-2/(r+1)}$ and

$$\int_{G^{-1}(2\zeta^{1/2}, y/2)/2} G(x, y)\, dx \leq c_5 y^{-\alpha r/2} (y^{-\alpha r/(r+1)})^{(1-r)/2}$$

for y small. The double integral in Hypothesis C(v) will then be finite if

$$\frac{-\alpha r}{2} + \left(\frac{-\alpha r}{r+1}\right)\left(\frac{1-r}{2}\right) > -1,$$

or if $r<(\alpha-1)^{-1}$, the same exponent of metric entropy as that obtained in our previous paper, Bass and Pyke (1984).

Example 6.2. *Partial sum processes* (cf. Bass and Pyke (1985), Section 5). Let $V_j, j\in\{1,2\cdots\}^d$, be mean 0 iid r.v.'s in the domain of normal attraction of a stable law of index $\alpha\in(1,2)$. Let $C_{nj}=n^{-1}(j-1,j]$, and let U_{nj} be independent r.v.'s, independent of the V_j's, with U_{nj} uniform on C_{nj}. Define

$$Z_n(A)=n^{-d/\alpha}\sum_{j\in\{1,\ldots,n\}^d} V_j 1_{A\cap C_{nj}}(U_{nj}).$$

Suppose $H_L(\delta)\leq K\delta^{-r}$ for $r<(\alpha-1)^{-1}$. The weak convergence of the Z_n sequence to a stable Lévy process, obtained in the above reference, is also a consequence of Theorem 6.1. Let us restrict ourselves to the subsequence n^d and index the X's by X_{n^dj}, where

$$X_{n^dj}(A)=n^{-d/\alpha}V_j 1_{A\cap C_{nj}}(U_{nj}).$$

Let us first consider the symmetrization of X_{n^dj}. Suppose $J\subseteq[-z,z]\backslash[-y,y]$ with $0<y<z\leq 1$. Let ε_j be iid r.v.'s that are independent of the V_j and U_{nj}'s and $P(\varepsilon_j=1)=P(\varepsilon_j=-1)=\frac{1}{2}$. It is easy to see that the symmetrization of $X_{n^dj}(A,J)$ has the same distribution as $\sum_{i=1}^N \tilde{V}_i$, where the $\tilde{V}_i$ are iid, independent of N,

$$P(\tilde{V}_i\in B)=P(n^{-d/\alpha}\varepsilon_j V_j\in B\cap J\,|\,n^{-d/\alpha}\varepsilon_j V_j\in J),$$

and N is distributed as a Binomial $(2,(|A\cap C_{nj}|/|C_{nj}|)P(n^{-d/\alpha}\varepsilon_j V_j\in J))$. Clearly, $|\tilde{V}_i|$ is bounded by z, and so by Corollary 3.2, $X_{n^dj}(A,J)-X'_{n^dj}(A,J)$, the symmetrization of $X_{n^dj}(A,J)$, is subpoissonian with parameters

$$((|A\cap C_{nj}|/|C_{nj}|)P(n^{-d/\alpha}\varepsilon_j V_j\in J),\,z).$$

Since V_j is in the domain of attraction of a stable law of index α and $|C_{nj}|=n^{-d}$,

$$n^d P(n^{-d/\alpha}\varepsilon_j V_j\in J)\leq n^d P(|\varepsilon_j V_j|\geq y n^{d/\alpha})\leq c_6 y^{-\alpha}.$$

We may thus take $\nu_{nj}(A,y)$ in Hypothesis C(i) to be $c_6|A\cap C_{nj}|y^{-\alpha}$, and hence $\nu_n(A,y)=c_6|A|y^{-\alpha}$.

Since the number of atoms of $X_{nj}(A,J)$ is

$$\text{Binomial}(1,(|A\cap C_{nj}|/|C_{nj}|)P(n^{-d/\alpha}V_j\in J))$$

and the moment generating function of a Binomial is dominated by that of a Poisson (cf. proof of Corollary 3.2), we see that the above choice of ν_{nj} will satisfy Hypothesis C(ii) also.

Next,

$$\sum_{j\in\{1,\ldots,n\}^d} P(X_{n^dj}(\cdot,[-y,y]^c)\text{ has an atom in }A)$$
$$\leq\sum_j(|A\cap C_{nj}|/|C_{nj}|)P(|V_j|\geq n^{d/\alpha}y)\leq c_6|A|y^{-\alpha}.$$

Since each X_{nj} has at most one atom, Hypothesis D will be satisfied if we take $\mu_m = 0$ for $m \geq 2$, $\mu_1(A \times J) = c_6(\alpha+1)^{-1}|A| \int_J y^{-\alpha-1}\, dy$. The calculations of Example 6.1 above show that the remainder of Hypothesis C is satisfied; see Bass and Pyke (1985) for a proof that the finite dimensional distributions converge. We may then apply Theorem 6.1 to conclude that Z_n converges weakly in $D_0(A)$.

Example 6.3. *Random masses, random locations (Generalized empirical measures).* Suppose one chooses n locations at random and then takes an observation at each location. As $n \to \infty$, one would expect a central limit theorem to hold. Perhaps the simplest model where the limit is not a continuous process is the following.

Let U_j be iid uniform r.v.'s on $(0, 1]^d$, and let V_j be mean 0 iid r.v.'s in the domain of normal attraction of a stable law of index α, $\alpha \in (1, 2)$, independent of the U_j's. Let

$$X_{nj}(A) = n^{-1/\alpha} V_j 1_A(U_j).$$

Using the methods and calculations of Examples 6.1 and 6.2, one can see that Hypotheses C and D are satisfied, provided $H_L(\delta) \leq K\delta^{-r}$, $r < (\alpha-1)^{-1}$. Hypothesis B is routine, and we see that $S_n \xrightarrow{w}$ to a stable Lévy process indexed by A.

Remark. A much more interesting case in Example 6.3 is where the distribution of V_j depends on the value of U_j. This case fits easily into the framework of Theorem 6.1, and with suitable conditions one can obtain a central limit theorem that will apply. Although these processes can be expressed in terms of ordinary empirical processes, the novelty here is the non-Gaussian limit.

Similar remarks apply to Examples 6.1 and 6.2, also. It is not hard to modify Example 6.1 to get processes with independent but not stationary increments. One could modify Example 6.2 to get a central limit theorem for partial sums of non-identically distributed r.v.'s.

7. Continuous components

We have not considered the case where the X_{nj}'s have a nondeterministic continuous component nor the case where the limit law of the S_n sequence has a continuous component (as, for example, in empirical processes converging to a tied-down Brownian process). The primary reason we have not done so is that there is already a large literature concerning continuous limits (see Section 1), and any of these results can be combined with our Theorem 6.1 as follows.

Suppose first of all that Hypothesis A holds. Suppose each X_{nj} can be written as $X_{nj}^D + X_{nj}^C$. Let

$$S_n^C = \sum_{j=1}^n X_{nj}^C \quad \text{and} \quad S_n^D = \sum_{j=1}^n X_{nj}^D.$$

Suppose the joint finite dimensional distributions of (S_n^C, S_n^D) converge. Suppose the X_{nj}^D satisfy Hypotheses B, C, and D. Finally, suppose $S_n^C \xrightarrow{L} S^C \in C(A)$, the continuous functions on A. Here the meaning of $\xrightarrow{L}$ varies according to context. If the X_{nj}^C are themselves in $C(A)$, then $\xrightarrow{L}$ means weak convergence with respect to $\|\cdot\|_A$. It follows that for given ε, there exists a compact subset K_1 of $D_0(A)$ with $\sup_n P(S_n^C \notin K_1) < \varepsilon$: Take $J_m \equiv 0$ and $C_m(S_n^C) = S_n^C$.

If S_n^C is the empirical process, Dudley and Philipp (1983) interpret $\xrightarrow{L}$ in our context as:

There exist continuous Gaussian processes T_n such that

$$\|S_n^C - T_n\|_A \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

For this case, let $\Delta_m = m^{-1}$. Given ε, there exists n_m such that for $n \geq n_m$,

$$P(\|S_n^C - T_n\|_A \geq \Delta_m) < \varepsilon/3.$$

Let

$$C_m(S_n^C) = T_n \text{ and } J_m(S_n^C) = 0 \text{ if } n \geq n_m,$$

and

$$C_m(S_n^C) = -ES_n^C,\ J_m(S_n^C) = S_n^C + ES_n^C \text{ if } n < n_m.$$

By taking N_m and R_m sufficiently large and h_m and η_m sufficiently small, it is not hard, using Proposition 2.2, to see that in this case as well, there exists a compact subset K_1 of $D_0(A)$ with $\sup_n P(S_n^C \notin K_1) < \varepsilon$.

Similar arguments show that for each of the references mentioned in Section 1, S_n^C is in a compact subset K_1 of $D_0(A)$ with large probability. By the proof of Theorem 6.1, we get that, given ε, there exists a compact subset K_2 of $D_0(A)$ with $P(S_n^D \notin K_2) < \varepsilon$. Then, considered as elements of $D_0(A) \times D_0(A)$, we have that (S_n^C, S_n^D) converges weakly, say to (S^C, S^D), where the metric here is $d_D \times d_D$. By Skorokhod's representation theorem (cf. remarks following Theorem 4.3 of Bass and Pyke (1985)), we can find another probability space and processes $\hat{S}_n^C, \hat{S}_n^D, \hat{S}^C, \hat{S}^D$, equal in law to S_n^C, S_n^D, S^C, S^D, respectively, such that

$$\hat{S}_n^C \xrightarrow{d_D} \hat{S}^C, \text{ a.s., and } \hat{S}_n^D \xrightarrow{d_D} \hat{S}^D, \text{ a.s.}$$

Since $\hat{S}^C \in C(A)$ by assumption, then $\hat{S}_n^C \xrightarrow{\|\cdot\|_A} \hat{S}^C$ by Theorem 3.5 of Bass and Pyke (1985). By Lemma 3.3 of that same paper,

$$\hat{S}_n^C + \hat{S}_n^D \xrightarrow{d_D} \hat{S}^C + \hat{S}^D, \quad \text{a.s.}$$

The weak convergence of S_n with respect to d_D follows immediately.

8. Signed measures

Until now, we have considered sums of purely atomic and continuous processes and their limits only. But there are many elements of $\boldsymbol{D}(\boldsymbol{A})$ other than these, e.g., measures concentrated on the surfaces of convex sets. In this section we introduce a set $\boldsymbol{D}_{SM}(\boldsymbol{A})$ intermediate between $\boldsymbol{D}_0$ and $\boldsymbol{D}$. While we do not know if $\boldsymbol{D}_{SM}$ is all of $\boldsymbol{D}$, it contains virtually every example likely to come up in practice.

Let us say that $x:\boldsymbol{A}\to\mathbb{R}$ is a signed measure if x is the restriction to $\boldsymbol{A}$ of a signed measure, also denoted x, on $(I^d, \boldsymbol{B}(I^d))$ of finite total variation. Define

$$\boldsymbol{D}_{SM}(\boldsymbol{A}) = \{x\colon \boldsymbol{A}\to\mathbb{R}\colon \textit{there exist signed measures } j_m \textit{ and continuous functions } c_m \textit{ such that } \|x-(j_m+c_m)\|_{\boldsymbol{A}}\to 0 \textit{ as } m\to 0\}.$$

Clearly $\boldsymbol{D}_0\subset \boldsymbol{D}_{SM}\subset \boldsymbol{D}$.

We will give a condition sufficient for a subset of $\boldsymbol{D}_{SM}(\boldsymbol{A})$ to be compact with respect to d_D. With this criterion, it should not be hard to formulate and prove central limit theorems using the techniques introduced in this paper. We leave such formulations to the reader with particular applications in mind.

First, we need a lemma. Recall that the topology on I^d is the relative one.

Lemma 8.1. *If x is a signed measure, Hypothesis A holds, $A\in\boldsymbol{A}$, and $\delta, \varepsilon>0$, then there exists $B\in\boldsymbol{A}$ with $A\subset B^0\subset A^\delta$ such that $|x(B)-x(A)|<\varepsilon$ and $x(\partial B)=0$.*

Proof. Since in the relative topology $\partial I^d=\emptyset$, we may without loss of generality assume $A\subsetneq I^d$. Suppose x is the restriction of a signed measure, also denoted x, with total variation measure $|x|$. If necessary, make δ smaller so that $|x|(A^\delta\setminus A)<\varepsilon$ and $A^\delta\subsetneq I^d$. By Hypothesis A, there exists $C_{1/2}\in\boldsymbol{A}$ such that $A\subset C^0_{1/2}\subset C_{1/2}\subset A^{\delta/2}$. Pick $b_{1/4}$ so that $A^{b_{1/4}}\subset C^0_{1/2}$, and then pick $C_{1/4}\in\boldsymbol{A}$ so that $A\subset C^0_{1/4}\subset C_{1/4}\subset A^{b_{1/4}}$. Pick $b_{3/4}$ so that $C_{1/2}^{b_{3/4}}\subset A^{\delta/2}$, and then pick $C_{3/4}\in\boldsymbol{A}$ so that $C_{1/2}\subset C^0_{3/4}\subset C_{3/4}\subset C_{1/2}^{b_{3/4}}$. Continue in this way, choosing $C_{1/8}, C_{3/8}$, etc. For $r\in(0,1)$ not a dyadic rational, let $C_r=\bigcap_{s>r,\, s \text{ dyadic}} C_s$. Since $\boldsymbol{A}$ is closed with respect to d_H, we have $C_r\in\boldsymbol{A}$. Since every pair of reals is separated by a dyadic rational, the C_r's are all disjoint, and for all $r\in(0,1)$, $A\subset C^0_r\subset C_r\subset A^\delta$. Since there are uncountably many C_r's, we can find one, call it C_{r_0}, such that $|x|(\partial C_{r_0})=0$. Letting $B=C_{r_0}$ completes the proof. □

If x is a signed measure, define

$$\text{Variation}\,(x) = \text{total variation of } x \text{ with respect to } (I^d, \boldsymbol{B}(I^d)).$$

Since $\boldsymbol{A}$ is compact, it is clear that given a sequence of signed measures x_m with $\sup_m$ Variation(x_m) finite, there is a subsequence $x_{m'}$ and a signed measure x_0 such that $x_{m'}(A)\to x_0(A)$ whenever $x_0(\partial A)=0$. With this observation, the proof of the following theorems are very similar to the proofs of Proposition 3.2 and Theorem 3.4 of Bass and Pyke (1985) and are left to the reader.

Theorem 8.2. *Let N be an integer-valued function, h a real-valued function and R a real number. Let $F_{SM}(h, N, R)$ be the set of all signed measures x such that*

(i) Variation $(x) \leq R$,

(ii) *for each δ, there exist $A_1, \ldots, A_{N(\delta)} \in \mathbf{A}$, possibly depending on x, such that*

(a) *every element of $G(x)$ is within δ of $\{(A_i, x(A_i))\}_{i=1}^{N(\delta)}$ with respect to d_G, and*

(b) $\sup_{A_i \subset B \subset A_i^{h(\delta)}} |x(B) - x(A)| < \delta$, $i = 1, \ldots, N(\delta)$.

Then $F_{SM}(h, N, R)$ is compact relative to d_D.

Theorem 8.3. *Suppose for each $m \geq 1$ that R_m and M_m are reals, h_m and N_m are functions as in Theorem 8.2, ω_m is an increasing function with $\omega_m(r) \to 0$ as $r \to 0$ and $\Delta_m \downarrow 0$. Let $F_D(h, N, R, M, \omega, \Delta)$ be the set of $x \in D_{SM}(A)$ such that for each m there exists a signed measure $J_m(x)$ and $C_m(x) \in C(A)$ with*

(i) $J_m(x) \in F_{SM}(h_m, N_m, R_m)$,

(ii) (a) $\|C_m(x)\|_A \leq M_m$,

(b) $\sup\{|C_m(x)(A) - C_m(x)(B)|: A, B \in \mathbf{A}, d_H(A, B) \leq r\} \leq \omega_m(r)$, and

(iii) $\|x - (J_m(x) + C_m(x)\|_A \leq \Delta_m$.

Then $F(h, N, R, M, \omega, \Delta)$ is compact relative to d_D.

References

R.J. Adler and P. Feigin, On the cadlaguity of random measures, Ann. Probab. 12 (1984) 615-630.

A. Araujo and E. Giné, The Central Limit Theorem for Real and Banach Valued Random Variables (Wiley, New York, 1980).

R.F. Bass and R. Pyke, The existence of set-indexed Lévy processes, Z. Wahrsch. Verw. Geb. 66 (1984) 157-172.

R. Bass and R. Pyke, The space D(A) and weak convergence for set-indexed processes, Ann. Probab. 13 (1985) 860-884.

P.J. Bickel and M.J. Wichura, Convergence criteria for multiparameter stochastic processes and some applications, Ann. Math. Stat. 42 (1971) 1656-1670.

P. Billingsley, Covergence of Probability Measures (Wiley, New York, 1968).

R.M. Dudley and W. Philipp, Invariance principles for sums of Banach space valued random elements and empirical processes, Z. Wahrsch. Verw. Geb. 62 (1983) 509-552.

M. Fisz, A central limit theorem for stochastic processes with independent increments, Studia Math. 18 (1959) 223-227.

E. Giné and M.B. Marcus, The central limit theorem for stochastic integrals with respect to Lévy processes, Ann. Probab. 11 (1983) 58-77.

B.V. Gnedenko and A.N. Kolmogorov, Limit Distributions for Sums of Independent Random Variables (Addison-Wesley, Cambridge, Mass, 1954).

M.G. Hahn, Central limit theorems for D[0, 1], Z. Wahrsch. Verw. Geb. 44 (1978) 89-101.

N.C. Jain and M.B. Marcus, Central limit theorems for $C(S)$-valued random variables, J. Funct. Anal. 19 (1975) 216-231.

T. Lindvall, Weak convergence of probability measures and random functions on the function space $D(0, \infty)$, J. Appl. Probab. 10 (1973) 109-121.

G. Neuhaus, On weak convergence of stochastic processes with multidimensional time parameter, Ann. Math. Stat. 42 (1971) 1285-1295.

G. Pisier, Le théorème limite central et la loi du logarithme iteré dans les éspaces de Banach, Séminaire Maurey-Schwartz, 1975-76. Exposés III et IV (1975).

D. Pollard, Convergence of Stochastic Processes (Springer, New York, 1984).

Yu. V. Prohorov, Convergence of random processes and limit theorems in probability theory, Theory Probab. Appl. 1 (1956) 117-238.

A.V. Skorokhod, Limit theorems for stochastic processes, Theory Probab. Appl. 1 (1956) 289-319.

M.L. Straf, Weak convergence of stochastic processes with several parameters. Proc. Sixth Berkeley Symp. Math. Statist. Probab. 2 (Univ. of California Press, 1972) 187-221.