Small and Large Value Probabilities and Related Limit Theorems

NSF/CBMS Conference UAHuntsville 2012

Part I: Notation and Some Simple Observations

Part II: Gaussian i.i.d. Samples and Examples

Part III: Small Value Probabilities and Limit Theorems

Part IV: Partial Sum Process Results for Supercritical Branching

Part I: Simple Observations and Some Notation.

Notation. (i) B is a real linear space, $\mathcal{B}$ a σ-field of subsets of B, and $q(\cdot)$ a semi-norm on B such that for all $x \in B$ and $r \geq 0$

$$x + rU \in \mathcal{B},$$

where $U = \{x \in B : q(x) < 1\}$.

(ii) If $A \subseteq B$ and $x \in B$ we define the q-distance from A to x to be

$$d_q(x, A) = \inf_{a \in A} q(x - a).$$

Remark. Typically B is a real separable Banach space with norm $q(\cdot)$ and $\mathcal{B}$ the Borel subsets of B, or $B = D[0, T]$ where $D[0, T]$ denotes usual cadlag space of functions with

$$q(x) = \sup_{t \in [0,T]} |x(t)|,$$

and $\mathcal{B}$ the σ-field generated by the mappings $\{x(t) : t \in [0, T]\}$.

Large Values Observation à la Strassen. Let $X, X_1, X_2, \cdots$ be identically distributed B-valued random vectors such that for some $\alpha > 0, c \equiv c_{q,X} > 0$ and $r \to \infty$

$$\log P(q(X) \geq r) \sim -cr^{\alpha}. \tag{1}$$

Then, for every $\epsilon > 0$, (1) implies

$$P(X_n \in (\frac{\log n}{c})^{\frac{1}{\alpha}}(1+\epsilon)(U \textit{ eventually}) = 1, \tag{2}$$

and if the random vectors are independent (1) also implies

$$P(X_n \in (\frac{\log n}{c})^{\frac{1}{\alpha}}(1-\epsilon)U^c \textit{ i.o.}) = 1, \tag{3}$$

Remarks.

(i) Check details using the Borel-Cantelli Lemma. Exponential tails in (2) are crucial for cutoff.

(ii) (2) implies an upper bound on the "rate of growth" in the sense that

$$P(\limsup_{n\to\infty} d_q((\frac{c}{\log n})^{\frac{1}{\alpha}} X_n, U) = 0) = P(\limsup_{n\to\infty} (\frac{c}{\log n})^{\frac{1}{\alpha}} q(X_n) \leq 1) = 1.$$

(iii) (2) and (3) combine to give this rate, i.e.

$$P(\limsup_{n\to\infty} (\frac{c}{\log n})^{\frac{1}{\alpha}} q(X_n) = 1) = 1. \tag{4}$$

Remarks Continued.
(iv) Too Simple? Obviously, and here are some reasons.

(a) Among your favorite stochastic processes, when do you know (1)?

(b) The i.i.d. sample model is quite restrictive since one is frequently interested in the limiting behavior for scaled samples of a fixed process.

(c) If X is a stochastic process with continuous sample paths, say on $[0, T]$, and q is the sup-norm on $C[0, T]$, then (4) determines the rate of growth for the largest absolute values of a typical path from the sample, but what about other properties of the path?

Small Values Observation à la Chung-Wichura. Let $X, X_1, X_2, \cdots$ be identically distributed B-valued random vectors such that for for some $\beta > 0$, $d \equiv d_{q,X} > 0$ and $r \downarrow 0$,

$$\log P(q(X) < r) \sim -dr^{-\beta}. \tag{5}$$

Then, for every $\epsilon > 0$, (5) implies

$$P(X_n \in (\frac{d}{\log n})^{\frac{1}{\beta}}(1-\epsilon)U^c \textit{ eventually}) = 1, \tag{6}$$

and if the random vectors are independent (5) also implies

$$P(X_n \in (\frac{d}{\log n})^{\frac{1}{\beta}}(1+\epsilon)U \textit{ i.o.}) = 1. \tag{7}$$

Remarks.
(i) Check details using the Borel-Cantelli Lemma. Exponential tails in (5) are crucial for cutoff.

(ii) (6) implies a lower bound on the "rate of escape from zero" in the sense that

$$P(\lim_{n\to\infty} d_q((\frac{\log n}{d})^{\frac{1}{\beta}} X_n, U^c) = 0) = P(\liminf_{n\to\infty}(\frac{\log n}{d})^{\frac{1}{\beta}} q(X_n) \geq 1) = 1.$$

(iii) (6) and (7) combine to give this rate, i.e.

$$P(\liminf_{n\to\infty}(\frac{\log n}{d})^{\frac{1}{\beta}} q(X_n) = 1) = 1. \tag{8}$$

Remarks Continued.
(iv) Too Simple? Obviously! Why?

(a) Among your favorite stochastic processes, when do you know (5)? Here things are even harder!

(b) The i.i.d. sample model is quite restrictive since one is frequently interested in the limiting behavior for scaled samples of a fixed process.

(c) If X is a stochastic process with continuous sample paths, say on $[0, T]$, and q is the sup-norm on $C[0, T]$, then (8) determines the rate of escape from the zero function for the largest absolute values of a typical path from the sample, but what about other properties of the path?

Part II: Gaussian i.i.d. Samples and Examples.

Notation.

B is a real separable Banach space with norm $q(\cdot)$ and topological dual space B^*.

X is a centered B-valued Gaussian random vector with $\mu = \mathcal{L}(X)$.

$H_\mu \subseteq B$ is the Hilbert space such μ is determined by considering the pair (E, H_μ) as an abstract Wiener space.

H_μ is the completion of the range of the map $S : B^* \to B$ given by the Bochner integral

$$Sf = \int_B xf(x)d\mu(x), f \in B^*,$$

and the completion is in the inner product norm

$$\langle Sf, Sg\rangle = \int_B f(x)g(x)d\mu(x), f, g \in B^*.$$

Notation Continued.

K is the unit ball of H_μ. For $\epsilon > 0$, $K^\epsilon = K + \epsilon U$.

If X is given by standard Brownian motion on $B = C[0, T]$, then

$$K = \{x \in C[0, T] : x(t) = \int_0^t g(s)ds, t \in [0, T], \int_0^T g^2(s)ds \leq 1\}.$$

$\{\alpha_k : k \geq 1\}$ is a sequence in B^*, orthonormal in $L_2(\mu)$, such that $\{S\alpha_k : k \geq 1\}$ is a CONS in H_μ, and define for $d \geq 1$ the linear operators taking $B \to B$

$$\Pi_d(x) = \sum_{k=1}^{d} \alpha_(x) S\alpha_k \text{ and } Q_d(x) = x - \Pi_d(x).$$

Rates of Clustering Theorem-[GK]. Let $X, X_1, X_2, \cdots$ be identically distributed B-valued centered Gaussian random vectors, and assume $\{d_n\}$ is a sequence of integers such that

$$d_n \geq \inf\{m \geq 1 : E[q(Q_m(X))]/m \leq (\Gamma L_2 n)/(2Ln)^{\frac{1}{2}}\},$$

where $\Gamma = \sup_{x \in K} q(x)$. If $\epsilon_n = (\gamma d_n L_2 n)/Ln$ and $\gamma > 3\Gamma$, then

$$P(X_n/(2Ln)^{\frac{1}{2}} \in K^{\epsilon_n} \textit{ eventually}) = 1.$$

Some Corollaries and Comments.

Corollary 1. Let X be given by standard Brownian motion on $B = C[0, T]$ with $q(\cdot)$ the sup-norm on B, and assume

$$\epsilon_n = \gamma(L_2 n/Ln)^{\frac{2}{3}}$$

Then, for $X, X_1, X_2, \cdots$ identically distributed and $\gamma > 0$ sufficiently large

$$P(X_n/(2Ln)^{\frac{1}{2}} \in K^{\epsilon_n} \textit{ eventually}) = 1.$$

Comments. (i) If $V = \{x \in C[0,T] : \int_0^T x^2(s)ds \leq 1\}$, then for Brownian motion samples as in Corollary 1 and $\gamma > 0$ sufficiently large

$$P(X_n/(2Ln)^{\frac{1}{2}} \in K + \epsilon_n V \textit{ eventually}) = 1,$$

where $\epsilon_n = \gamma(L_2 n)^{\frac{1}{3}}/(Ln)^{\frac{2}{3}}$. Here ϵ_n is smaller since V is larger than U, and our choice of d_n depends on small ball probabilities and the approximation properties of the operators Π_d in the different norms.

(ii) **Scaled Samples of Brownian Paths,[GK].** If $\{W(t) : t \geq 0\}$ is sample continuous Brownian motion,

$$X_n(t) = W(nt)/n^{\frac{1}{2}}, 0 \leq t \leq T, n \geq 1,$$

and

$$\epsilon_n = \gamma(L_3 n)/L_2 n)^{\frac{2}{3}},$$

then for $\gamma > 0$ sufficiently large

$$P(X_n/(2L_2 n)^{\frac{1}{2}} \in K^{\epsilon_n} eventually) = 1.$$

(iii) The results in (ii) are close to being optimal in that [KG] proved for all $\gamma, \theta > 0$ and $\epsilon_n = \gamma/(L_2 n)^{\frac{2}{3}+\theta}$

$$P(X_n/(2L_2 n)^{\frac{1}{2}} \in K^{\epsilon_n} i.o.) = 0.$$

(iv)**Other examples.** Similar results hold for Brownian sheets, fractional Brownian motions, and other Gaussian processes. If $\epsilon_n = \epsilon$, Strassen's ground breaking result for BM and partial sum processes of i.i.d. random variables initiated this type of investigation. Bolthausen-1978 studied rates for the Brownian motion case using special properties of Brownian motion, as Borel (1975), used below, was not a universal part of probability at that stage of the game.

(v) **A Universal Rate and a First Approach to Proofs.** The Rates of Clustering Theorem always holds for $\epsilon_n = \gamma/(Ln)^{\frac{1}{2}}$ and $\gamma > 0$ sufficiently large. Actually $\gamma > 0$ is sufficient, but the argument we give here is only an instructive beginning. The main tool is a result of C. Borell, which also makes a link to small deviation probabilities, and implies

$$P(X/r \in K + aU) \geq \Phi(r + \alpha),$$

where Φ is the $N(0, 1)$ c.d.f. and $\Phi(\alpha) = P(X \in raU)$. Hence for $r = r_n = (2Ln)^{\frac{1}{2}},\ a = \epsilon_n = \gamma/(Ln)^{\frac{1}{2}}$ we have

$$P(X \in (2Ln)^{\frac{1}{2}} K + \sqrt{2}\gamma U) \geq \Phi((2Ln)^{\frac{1}{2}} + \alpha),$$

where $\Phi(\alpha) = P(\sqrt{2}\gamma U)$.

Comment (v) Continued. Taking $\gamma > 0$ sufficiently large that $P(\sqrt{2}\gamma U) > \frac{1}{2}$ we have $\alpha > 0$ and hence

$$p_n \equiv P(X \notin (2Ln)^{\frac{1}{2}} K + \sqrt{2}\gamma U) \leq 1 - \Phi((2Ln)^{\frac{1}{2}} + \alpha),$$

which implies

$$p_n \leq C(Ln)^{-\frac{1}{2}} \exp\{-(\sqrt{2Ln} + \alpha)^2/2\}.$$

Hence $\sum_{n \geq 1} p_n < \infty$ for $\alpha > 0$, and the Borel-Cantelli lemma complete the proof.

(vi) **Another Universal Result, [GK] .** Let B be a separable Banach space. If $X, X_1, X_2, \cdots$ are i.i.d. centered B-valued Gaussian random vectors,

$$G_n = \{X_1/(2Ln)^{\frac{1}{2}}, \cdots, X_n/(2Ln)^{\frac{1}{2}}\},$$

and $\epsilon_n = \gamma/Ln^{\frac{1}{2}}$ for $\gamma > 0$, then

$$P(G_n \subseteq K^{\epsilon_n} eventually) = 1,$$

and

$$P(K \subseteq G_n^{\epsilon_n} eventually) = 1.$$

(vii) If $x \in K$, then determining constants b_n such that

$$P(0 < \liminf_{n \to \infty} b_n d_q(x, \frac{X_n}{(2Ln)^{\frac{1}{2}}}) < \infty) = 1$$

has been studied in a series of papers by [C], [deA], [KG], and [KLT]. These rates again depend on the small ball probabilities of X, and there are different rates for points with $||x||_\mu = 1$. The points $\{Sf : f \in B^*, ||Sf||_\mu = 1\}$ are approached slowest.

Part III: Small Value Probabilities and Limit Theorems

Let $\{X(t) : t \geq 0\}$ be a stochastic process with cadlag paths in $D[0, \infty),\ X(0) = 0$, and for $t \geq 0, n \geq 1$, define

$$\eta_n(t) = M(nt)/(d_\alpha n/L_2 n)^{\frac{1}{\alpha}},$$

where

$$M(t) = \sup_{0 \leq s \leq t} |X(s)|.$$

The parameter α is typically related to the scaling parameter of the process $\{X(t) : t \geq 0\}$, and

$$d_\alpha = \lim_{\epsilon \to 0^+} \epsilon^\alpha \log P(\sup_{0 \leq s \leq 1} |X(s)| \leq \epsilon).$$

Let $\mathcal{M}$ denote the functions $f : [0, \infty) \to [0, \infty]$ such that $f(0) = 0$, f is right continuous on $(0, \infty)$, non-decreasing, and such that $\lim_{t \to \infty} f(t) = \infty$. Also define

$$\mathcal{K}_\alpha = \{f \in \mathcal{M} : \int_0^\infty f^{-\alpha}(s) ds \leq 1\}.$$

The topology on $\mathcal{M}$ is that of weak convergence, i.e. pointwise convergence at all continuity points of the limit function. This topology is metrizable and separable, and if $\{f_n\}$ is a sequence of points in $\mathcal{M}$, then $C(\{f_n\})$ denotes the cluster set of $\{f_n\}$, i.e. all possible subsequential limits of $\{f_n\}$ in the weak topology. If $A \subseteq \mathcal{M}$ we write $\{f_n\} \rightsquigarrow A$ if $\{f_n\}$ is relatively compact in $\mathcal{M}$ and $C(\{f_n\}) = A$.

Then, in [CKL] and [KL] we have for $X = \{X(t) : t \geq 0\}$:

(i) If X is a symmetric stable process with stationary independent increments of index $\alpha \in (0, 2]$, then

$$P(\{\eta_n\} \rightsquigarrow \mathcal{K}_\alpha) = 1.$$

(ii) If X is a fractional Brownian motion process with parameter $\gamma \in (0, 1)$, then for $\alpha = 1/\gamma$

$$P(\{\eta_n\} \rightsquigarrow \mathcal{K}_\alpha) = 1.$$

(iii) Similar results hold for Levy's stochastic area process [KL], partial sum processes built from i.i.d. random variables [R], and also for multigenerational samples of a super critical Galton-Watson branching process [KV]. In this last case the situation is a triangular array, and at each stage n the process is built from the n^{th} generation, i.e. we do not scale a single generation.

(iv) In all of these results we see that with probability one

$$\liminf_{n\to\infty} \eta_n(1) = \liminf_{n\to\infty} \sup_{0\le t\le 1} M(nt)/(d_\alpha n/L_2 n)^{\frac{1}{\alpha}} = 1.$$

Why? Suppose $\liminf_{n\to\infty} \eta_n(1) \le d < 1$ and $\eta_n \rightsquigarrow \mathcal{K}_\alpha$ on $\Omega_0, P(\Omega_0) > 0$. Then, for $\omega \in \Omega_0$ there are random subsequences $\{n_k\}$ such that $\lim_{k\to\infty} \eta_{n_k}(1) = \liminf_{n\to\infty} \eta_n(1) \le d < 1$, and selecting possibly an additional subsequence n_{k_r} we also have $\eta_{n_{k_r}}$ converges to an $f \in \mathcal{K}_\alpha$. Since the η_n and f are increasing, right continuous functions we have from above that $f(t) \le d < 1$ except possibly for countably many $t \in (0, 1)$ where f is discontinuous. Thus

$$\int_0^\infty f^{-\alpha}(s)ds \ge \int_0^1 d^{-\alpha} ds > 1,$$

which is a contradiction to $P(\Omega_0) > 0$. Hence with probability one

$$\liminf_{n\to\infty} \eta_n(1) \ge 1.$$

(iv continued) To see

$$\liminf_{n\to\infty} \eta_n(1) = 1,$$

we take $d > 1$ and define $f(0) = 0, f(t) = d, 0 < t < 1 + \delta$, $f(t) = \infty, t \geq 1 + \delta$. Then, for $\delta > 0$ sufficiently small $f \in \mathcal{K}_\alpha$, and f is continuous at $t = 1$. Hence, with probability one

$$\liminf_{n\to\infty} \eta_n(1) \leq d,$$

and since $d > 1$ was arbitrary our assertion follows.

(v) Note that the small value observations made earlier are applied, not to X itself here, but to the increasing paths formed from X to define the process M.

Something about the Proofs.

To minimize notation we restrict ourselves to X being standard Brownian motion, but the proofs follows similar lines in other situations. However, the details vary, and often require adjustments that are not always immediate.

The proof follows from three facts.

(I) $P(C(\{\eta_n\}) \subseteq \mathcal{K}_\alpha) = 1$.

(II) $P(\{\eta_n\} \text{ is relatively compact in } \mathcal{M}) = 1$.

(III) $P(\mathcal{K}_\alpha \subseteq C(\{\eta_n\})) = 1$.

Let $m \geq 1, \tau > 0$, and assume $0 < t_1 < t_2 < \cdots < t_m$. Then, except for a slight variation when f jumps to infinity at some finite point, a typical neighborhood of $f \in \mathcal{M}$ is of the form

$$N_{f,\tau} = \{g \in \mathcal{M} : f(t_j) - \tau < g(t_j) < f(t_j) + \tau, 1 \leq j \leq m\}.$$

The necessary probability estimate in order to apply Borel-Cantelli arguments in this setting are as follows.

Proposition. Let $\{X(t) : t \geq 0\}$ be a standard Brownian motion. Fix sequences $\{t_i\}_{i=0}^m, \{a_i\}_{i=1}^m$, and $\{b_i\}_{i=1}^m$, where $0 = t_0 < t_1 < \cdots < t_m$, $a_i < b_i, 1 \leq i \leq m$, and $b_1 \leq b_2, \leq \cdots \leq b_m$. Then,

$$\limsup_{\epsilon \to 0^+} \epsilon^2 \log P(a_i\epsilon \leq M(t_i) \leq b_i\epsilon, 1 \leq i \leq m) \leq -\frac{\pi^2}{8} \sum_{i=1}^{m} (t_i - t_{i-1})/b_i^2.$$

In addition, if we assume $a_1 < b_1 \leq a_2 < b_2 \leq \cdots, \leq a_m < b_m$, then

$$\liminf_{\epsilon \to 0^+} \epsilon^2 \log P(a_i\epsilon \leq M(t_i) \leq b_i\epsilon, 1 \leq i \leq m) \geq -\frac{\pi^2}{8} \sum_{i=1}^{m} (t_i - t_{i-1})/b_i^2,$$

where $\frac{\pi^2}{8}$ is the small ball probability constantfor Brownian motion.

Since the sums in the previous proposition are Riemann sums for the function f which is the center of the neighborhood $N_{f,\tau}$, taking $\epsilon > 0$ small and refining the partition by increasing m we see that the limit set should be

$$\{f \in \mathcal{M} : \int_0^\infty f^{-2}(t)dt \leq 1\},$$

and this is the set $\mathcal{K}_2$ of our theorem for the Brownian motion process. Of course, there are many details that have been ignored in this quick picture.

Part IV: Partial Sum Process Results for Supercritical Branching.

Let $\{\xi_{n,j}, j \geq 1, n \geq 1\}$ denote a double array of non-negative integer valued i.i.d. random variables defined on the probability space $(\Omega, \mathcal{F}, P)$, and having probability distribution $\{p_j : j \geq 0\}$, i.e. $P(\xi_{1,1} = k) = p_k$. Then, $\{Z_n : n \geq 0\}$ denotes the Galton-Watson process initiated by a single ancestor $Z_0 \equiv 1$. It is iteratively defined on $(\Omega, \mathcal{F}, P)$ for $n \geq 1$ by

$$Z_n = \sum_{j=1}^{Z_{n-1}} \xi_{n,j}.$$

Let $m = E(Z_1) \in (1, \infty)$. This is the supercritical case, and the probability that the process becomes extinct, namely q, is less than one. The complement of the set $\cup_{n=1}^{\infty}\{Z_n = 0\}$ is the so called survival set, and is denoted by S, $P(S) = 1 - q$, and $Z_n \to \infty$ a.s. on S. Also, $q = 0$ if and only if $p_0 = 0$, and we assume the offspring variance $\sigma^2 = Var(Z_1) \in (0, \infty)$.

On the set $\{Z_{n-1} > 0\}$ define $X_{n,Z_{n-1}}(0) = 0$, and for $0 < t \leq 1$ set

$$X_{n,Z_{n-1}}(t) = \frac{1}{\sigma\sqrt{Z_{n-1}}}\{\sum_{j=1}^{\lfloor tZ_{n-1}\rfloor}(\xi_{n,j} - m) + c_{n,Z_{n-1}}(t)\},$$

where $c_{n,Z_{n-1}}(t) = (tZ_{n-1} - \lfloor tZ_{n-1}\rfloor))(\xi_{n,\lfloor tZ_{n-1}\rfloor+1} - m)$.

On $\{Z_{n-1} = 0\}$ we define $X_{n,Z_{n-1}}(t) = 0$ for $0 \leq t \leq 1$. Hence $X_{n,Z_{n-1}}(\cdot)$ denotes an element of the space of continuous functions on [0,1] that vanish at zero with sup-norm given by $q(\cdot)$.

The analogue of Strassen's result for these processes is the following:

Theorem.[KV] Assume $E(Z_1^2(L(Z_1))^r) < \infty$ for some $r > 4$, and that K denotes the limit set for the Strassen type LIL for Brownian motion. Then

$$P(\lim_{n\to\infty} d_q(\frac{X_{n,Z_{n-1}}}{(2Ln)^{\frac{1}{2}}}, K) = 0) = 1.$$

In addition, if S denotes the survival set of the process, then

$$P(C(\{\frac{X_{n,Z_{n-1}}}{(2Ln)^{\frac{1}{2}}}\}) = K|S) = 1.$$

The maximal process and limit set used in connection to our Chung-Wichura law of the logarithm in this setting are

$$M_{n,Z_{n-1}}(t) = \sup_{0\le s\le t} |X_{n,Z_{n-1}}(s)|, \ \ 0 \le t \le 1,$$

$$\mathcal{K}_2 = \{f \in \mathcal{M} : \int_0^1 f^{-2}(s)ds \le 1\},$$

and we recall $d_2 = \frac{\pi^2}{8}$.

The Chung-Wichura law in this setting is:

Theorem. Assume $E(Z_1^2(L(Z_1))^r) < \infty$ for some $r > 4$, and that S denote the survival set of the process. Then,

$$P(\{\sqrt{\frac{Ln}{d_2}} M_{n,Z_{n-1}}(\cdot) \rightsquigarrow \mathcal{K}_2 | S) = 1.$$

Remarks.

(1) Complete analogues of the above type were proved for $r(n)$ generations of the branching chain, $1 \leq r(n) \leq n, r(n) \to \infty$, and the limit sets are given by

$$K_\infty = \{(f_1, f_2, \cdots) \in (K \times K \times \cdots) : \sum_{k \geq 1} \int_0^1 (f_k'(s))^2 ds \leq 1\},$$

and

$$\mathcal{K}_\infty = \{(h_1, h_2, \cdots) \in (\mathcal{K}_2 \times \mathcal{K}_2 \times \cdots) : \sum_{k=1}^{\infty} \int_0^1 h_k^{-2}(s) ds \leq 1\},$$

which is somewhat surprising as these are the limits one would expect if the successive generations were totally independent.

(2) The fact that $r > 4$ in the moment assumption in these theorems results from the use of standard estimates for the Prokhorov distance in the classical invariance theorem. That these estimates are essentially best possible can be seen from work of Borovkov and also Sahanenko. Thus an attempt at reducing $r > 4$ to , say $r > 1$, would seem to require a substantially different approach than what we use here. In particular, in the setting of functional limit theorems of high dimension, the difficulties imposed when working with partial sums from successive generations of a branching process make many typical LIL arguments along subsequences unavailable.

(3) Using these methods one can prove analogues of these results for triangular arrays of independent random variables under a variety of conditions. For example, such results hold as long as the row lengths have length $n^{8+\delta}$, the random variables are identically distributed with three moments, and the rows are independent, but there are other conditions that suffice as well. The additional assumption that the rows of the triangular array have some form of independence is necessary to show that the cluster set formed is all of the relevant limit set. In the supercritical branching process model no additional assumptions need be made, and although the rows are not independent, there is enough asymptotic independence when combined with the conditional Borel-Cantelli lemma to allow a proof.

References.

Bolthausen, E. (1978). On the speed of convergence in Strassen's law of the iterated logarithm, *Ann. Probab.*, 6, 668-672.

Borell, C. (1975). The Brunn-Minkowski inequality in Gauss space, *Invent. Math.*, 30, 207-216.

Borovkov, A.A. (1973). On the rate of convergence for the invariance principle, *Th. Probab. Appl.*, 18, 207-225.

Chung, K.L. (1948). On the maximum partial sums of sequences of independent random variables, *Trans. Amer. Math. Soc.*, 2, 205-233.

[CKL] Chen, X, Kuelbs, J. and Li, W.V. (2000). A functional LIL for symmetric stable processes, *Ann. Probab.*, 28, 258-276.

[C] Csáki, E. (1980). A relation between Chung's and Strassen's Laws of the iterated logarithm, *Z. Wahrsch. Verw. Gebiete*, 54, 287-301.

[deA] de Acosta, A. (1983). Small deviations in the functional central limit theorem with applications to functional laws of the iterated logarithm, *Ann. Probab.*, 11, 78-101.

[GK-1] Goodman, V. and Kuelbs, J. (1988). Rates of convergence for increments of Brownian motion, *J. Theor. Probab.*, 1, 27-63.

[GK-2]] Goodman, V. and Kuelbs, J. (1991). Rates of clustering for some Gaussian self-similar processes, *Probab. Th. Rel. Fields*, 88, 47-75.

[KG] Grill, K. (1987). On the rate of convergence in Strassen's law of the iterated logarithm, *Probab.Th. Rel. Fields*, 74, 583-589.

[KL1] Kuelbs, J. and Li, W.V. (2002). A functional LIL and some weighted occupation measure results for fractional Brownian motion, *J. Theor. Probab.*, 15, 1007-1030.

[KL-2] Kuelbs, J. and Li, W.V. (2005). A functional LIL for stochastic integrals and the Lévy area process, *J. Theor. Probab.*, 18, 261-290.

[KLT] Kuelbs, J., Li, W.V., and Talagrand, M. (1994). Lim inf results for Gaussian samples and Chung's functional LIL, *Ann. Probab.*, 22, 1879-1903.

[KV] Kuelbs, J. and Vidyashankar, A.N. (2009). Functional limit laws of Strassen and Wichura type for multiple generations of branching processes, *IMS Collections 5, High Dimensional Probability V: The Luminy Volume*, 131-152.

[R] Rushton, J. (2007). A Functional LIL for d-Dimensional Stable Processes; Invariance for Lévy- and Other Weakly Convergent Processes, *J. Theor. Probab.*, 20, 397-427.

[S] Sahanenko, A.I. (1974). Estimates of the rates of convergence in the invariance principle, *Dokl. Akad. Nauk. SSSR*, 1752-55.

[W] Wichura, M. (1973). A functional form of Chung's law of the iterated logarithm for maximum absolute partial sums, unpublished manuscript.