On a functional central limit theorem for random walks conditioned to stay positive

Bolthausen, E

Abstract: Let \{X_k : k \geq 1\} be a sequence of i.i.d.rv with \( E(X_i) = 0 \) and \( E(X_i^2) = \sigma^2 \), \( 0 < \sigma^2 < \infty \). Set \( S_n = X_1 + \cdots + X_n \). Let \( Y_n(t) = \frac{S_k}{\sigma \sqrt{n}} \) for \( t = k/n \) and suitably interpolated elsewhere. This paper gives a generalization of a theorem of Iglehart which states weak convergence of \( Y_n(t) \), conditioned to stay positive, to a suitable limiting process.

DOI: 10.1214/aop/1176996098

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: http://doi.org/10.5167/uzh-23167

Originally published at:
ON A FUNCTIONAL CENTRAL LIMIT THEOREM FOR RANDOM WALKS CONDITIONED TO STAY POSITIVE

BY ERWIN BOLTHAUSEN

Universität Konstanz

Let \(\{X_k: k \geq 1\}\) be a sequence of i.i.d. \(\text{rv}\) with \(E(X_1) = 0\) and \(E(X_1^2) = \sigma^2, 0 < \sigma^2 < \infty\). Set \(S_n = X_1 + \cdots + X_n\). Let \(Y_n(t)\) be \(S_n/\sigma \sqrt{n}\) for \(t = k/n\) and suitably interpolated elsewhere. This paper gives a generalization of a theorem of Iglehart which states weak convergence of \(Y_n(t)\), conditioned to stay positive, to a suitable limiting process.

1. Introduction. Let \(\{X_i\}_{i \in \mathbb{N}}\) be a sequence of i.i.d. \(\text{rv}\) with \(E(X_i) = 0\) and \(E(X_i^2) = \sigma^2\) where \(0 < \sigma^2 < \infty\). Let \(S_k = X_1 + \cdots + X_k\) and \(Y_n(t)\) be the continuous process on \([0, 1]\) for which \(Y_n(k/n) = S_k/\sigma \sqrt{n}\) and which is linearly interpolated elsewhere.

It is well known (see e.g., [2]) that \(Y_n(t)\) converges weakly in \((C[0, 1], \rho)\) to the Brownian motion process, where \(C[0, 1]\) is the set of continuous functions on \([0, 1]\) and \(\rho\) the supremum metric.

Let now \(C^+ = \{f \in C : f(t) \geq 0 \text{ for } t \in [0, 1]\}\). We have \(P(Y_n \in C^+) > 0\) for each \(n\). So the definition of conditional probabilities is elementary. Let \(Y_n^+\) be the \(Y_n\)-process conditioned to stay positive. That is for all Borel-sets \(A \subset C[0, 1]\) we set \(P(Y_n^+ \in A) = P(Y_n \in A | Y_n \in C^+)\). We remark that \(C^+\) is a null set for the measure of the Brownian motion. Iglehart proved [3] weak convergence of the \(Y_n^+\) process to the Brownian meander process \(W^+\) which is defined by

\[
W^+(t) = \frac{1}{(1 - \tau)^{1/4}} W(\tau + (1 - \tau)t), \quad 0 \leq t \leq 1
\]

with \(W\) the Brownian process and \(\tau = \sup\{t \in [0, 1]: W(t) = 0\}\). (Notice that \(\tau < 1\) a.s.)

Iglehart assumed \(E|X|^3 < \infty\) and \(X\) nonlattice or integer valued with span 1. It is shown in this paper that these extra assumptions are superfluous. Iglehart calculates the finite-dimensional distributions and proves tightness. Then he identifies the process with (1.1) for which Belkin [1] calculated the finite dimensional distributions. The proof given here requires no computation. It is based on identifying \(\lim_{n \to \infty} Y_n^+(t) = W(T + t) - W(T) = W^+(t)\) for an appropriate random time \(T\) and uses only the continuous mapping theorem (Theorem 5.1 in [2]).

2. Notations and preliminary lemmas. For \(s \in (0, \infty)\) let \(C^s\) be the set of

Received May 27, 1975; revised November 10, 1975.
AMS 1970 subject classifications. Primary 60F05; Secondary 60J15.
Key words and phrases. Conditioned limit theorem, functional central limit theorem, random walks, weak convergence.

480
continuous functions on $[0, s]$ (or $[0, \infty)$ for $s = \infty$) and $\mathcal{B}^*$ the smallest $\sigma$-algebra such that the mappings $C^* \ni f \to f(t) \in \mathbb{R}$ are measurable.

Let $P^*$ be the measure of the Brownian motion on $(C^*, \mathcal{B}^*)$.

$$T^*: C^* \to \mathbb{R}^+ = [0, \infty] \text{ is the mapping with}$$

$$T^*(f) = \inf \{ t: f(u) \geq f(t) \text{ for } i \leq u \leq t + 1 \leq s \}, \quad (\inf \emptyset = \infty).$$

We set $T = T^\omega$ and $P = P^\omega$ for simplicity.

**Lemma 2.1.** For all $s \in (0, \infty]$ $T^s$ is $\mathcal{B}^s$-measurable.

**Proof.** If $v = s - (u + 1) > 0$ then $\{ T^s \leq u \} = \bigcap_{n \in \mathbb{N}} \{ f \in C^*: \text{there exists a rational } r \leq u + 1/n \text{ with } f(r) < \min_{t \leq t < r} f(t + 1/n) + 1/n \}$, which is easily seen to belong to $\mathcal{B}^s$.

**Lemma 2.2.** $P(T < \infty) = 1$.

**Proof.** Let $A_\varepsilon = \{ f \in C^*: \text{ex. } s \leq 1 - \varepsilon \text{ with } f(s) \leq f(u) \text{ for } s \leq u \leq s + \varepsilon \}$. Now we have $A_\varepsilon \uparrow \{ f \in C^*: \text{f nonincreasing} \}$ as $\varepsilon \downarrow 0$. We infer $P(A_\varepsilon) \uparrow 1$ for $\varepsilon \downarrow 0$. If $\varphi: C^\omega \to C^\omega$ is defined by $\varphi(f)(t) = \varepsilon^{-1} f(\varepsilon t)$ then $\varphi$ is measure preserving (see [5] page 246) and $\varphi(A_\varepsilon) \subset \{ T < \infty \}$ so $P(T < \infty) \geq P(A_\varepsilon)$ for all $\varepsilon > 0$.

**Lemma 2.3.** The following three statements are true for all $s \in (0, \infty]$.

$$P^*(f(T^s) = f(T^s + 1)) = 0;$$

$$P^*(T^s = s - 1) = 0;$$

$$P^*(\text{ex. } u \in (0, 1) \text{ with } f(T^s) = f(T^s + u)) = 0.$$
By (2.2) there is as \( \tau < \delta \) so that
\[
\inf_{T + \tau \leq u \leq T + \tau + 1} f(u) > f(T).
\]
Now (2.4) gives \( \varepsilon = \frac{1}{3} (\inf_{T + \tau \leq u \leq T + \tau + 1} f(u) - f(T)) > 0 \).

If \( \rho(f, f') < \varepsilon \) and \( f' \) is such that \( T(f) \leq f'(f) \leq T(f) + \tau \) and \( f'(f') = \inf_{T \leq u \leq T + \tau} f'(u) \), then \( T(f') \leq f'(f) \leq T(f) + \delta \).

(II) To show the other inequality note that
\[
\lim_{n \to \infty} (\inf \{ T(f') : \rho(f, f') < 1/n \}) = \lambda \leq T(f).
\]

Let \( \{ f_n \}_{n \in \mathbb{N}} \) be a sequence with \( \rho(f, f_n) \leq 1/n \) and \( \lim_{n \to \infty} T(f_n) = \lambda \). Let \( \varepsilon > 0 \).

By the continuity of \( f \) and the uniform convergence of \( f_n \), there exists \( n_0 \) such that for \( n \geq n_0 \) we have:
\[
\inf_{T \leq u \leq T + 1} f(u) \geq \inf_{T \leq u \leq T + 1} f(u) - \varepsilon \geq \inf_{T \leq u \leq T + 1} f(u) - 2\varepsilon \geq f(T(f_n)) - 3\varepsilon \geq f(\lambda) - 4\varepsilon.
\]

So \( \inf_{T \leq u \leq T + 1} f(u) \geq f(\lambda) \) which implies \( T(f) \leq \lambda \) completing the proof of Lemma 2.4.

Let \( u \) be the function in \( C^1 \) which is everywhere equal \(-1\). We define a map \( \Phi_s : C^* \to C^1 \)
\[
\Phi_s(f)(t) = f(T^s(f) + t) \quad \text{for} \quad T^s(f) < \infty
\]
\[
= u \quad \text{for} \quad T^s(f) = \infty.
\]

We write \( \Phi = \Phi_\infty \) for simplicity.

A straightforward conclusion of Lemma 2.4 is

**Lemma 2.5.** For each \( s \in (0, \infty] \) \( \Phi_s \) is continuous \( P^* \) a.s. on \((C^*, \rho)\).

3. Sums of independent random variables conditioned to stay positive. Let \( X_1, X_2, \ldots, \) be i.i.d. rv with \( E(X_1) = 0 \); \( E(X_i^2) = \sigma^2 < \infty \) (\( \sigma^2 > 0 \)) and \( S_k = \sum_{j=1}^k X_j \). \( T_n = \inf \{ k : S_{k+i} \geq S_k \} \) for \( i = 1, \ldots, n \). Clearly \( T_n < \infty \) holds a.s. We set \( Z_k = S_{T_{n+k}} - S_{T_n} \).

**Lemma 3.1.** For each sequence of real numbers \( a_1, \ldots, a_n \)
\[
P(S_k \leq a_k, k = 1, \ldots, n \mid S_k \geq 0, k = 1, \ldots, n) = P(Z_k \leq a_k, k = 1, \ldots, n).
\]

**Proof.** This is an easy consequence of the independence and identical distribution of the \( X_i \):

If \( B_j = \bigcup_{s=0}^{j-1} \{ S_s \leq S_r \} \) for \( s + 1 \leq r \leq \min (j, s + n) \) we have
\[
P(S_{T_{n+k}} - S_{T_n} \leq a_k \text{ for } k = 1, \ldots, n) = \sum_{j=0}^{\infty} P(S_{j+k} - S_j \leq a_k \text{ for } k = 1, \ldots, n \mid T_n = j)P(T_n = j).
\]
\[
\begin{align*}
\sum_{k=1}^{\infty} P(S_{j+k} - S_j \leq a_k \mid S_{j+k} \geq S_j) & \quad \text{for } k = 1, \ldots, n \mid S_{j+k} \geq S_j \\
& \quad \text{for } k = 1, \ldots, n \text{ and } B_i \mid P(T_n = j) \\
& = P(S_k \leq a_k, k = 1, \ldots, n \mid S_k \geq 0, k = 1, \ldots, n) \\
& \quad \text{since } T_n < \infty \text{ a.s.}
\end{align*}
\]

We set \(Y_n(k/n) = (1/n^4)S_k\) for \(k \geq 0\) and \(Y_n(t)\) linearly interpolated.

Let \(Q_n\) be the probability measure defined on \((C^\infty, \mathcal{B}^\infty)\) by this process. Let \(\Pi_s : C^\infty \to C^s\) be the projection map and \(\Phi, C^+\) defined as above. We remark that \(P^s = P\Pi_s^{-1}\).

Let \(Q_n\Pi_1^{-1}(dx \mid C^+\) be the probability measure on \(C^s\) which is defined by
\[Q_n\Pi_1^{-1}(A \mid C^+\) = \frac{Q_n(\Pi_1^{-1}(A \cap C^+))}{Q_n(\Pi_1^{-1}(C^+))}
\]
for \(A \in \mathcal{B}^1\).

**Theorem 3.2.** The probability measures \(Q_n\Pi_1^{-1}(dx \mid C^+)\) converge weakly to \(P\Phi^{-1}\) (on \((C^1, \rho)\)).

**Proof.** We have proved in Lemma 3.1 that
\[(3.2) \quad Q_n\Pi_1^{-1}(dx \mid C^+) = Q_n\Phi^{-1}(dx) \text{ holds.}
\]

Now by Donsker’s theorem (see [2]), \(Q_n\Pi_1^{-1}\) converges weakly to \(P^s\) for \(s < \infty\). With regard to Lemma 2.5 we have for \(s < \infty\)
\[(3.3) \quad Q_n(\Phi_s\Pi_s)^{-1} \to P^s\Phi_s^{-1} \text{ weakly.}
\]
(Theorem 5.1 in [2].)

Let \(A\) be a continuity set in \(\mathcal{B}^1\), that is \(P\Phi^{-1}(\partial A) = 0\). We are going to show that
\[(3.4) \quad \lim_{n \to \infty} Q_n\Phi^{-1}(A) = P\Phi^{-1}(A).
\]
The theorem then follows. (3.4) doesn’t follow directly from (3.3) because we have there the assumption \(s < \infty\). Set
\[D = \{f \in C^1 : \min_{-\frac{1}{2} < t < \frac{1}{2}} f(t) \geq -\frac{1}{2}\}.
\]
Without loss of generality we can assume \(A \subset D\). (If not: replace \(A\) by \(A \cap D\) noticing \(Q_n\Phi^{-1}(D') = P\Phi^{-1}(D') = P\Phi^{-1}(\partial D) = 0\).

Let \(\varepsilon > 0\) be given. According to Lemma 2.2 we have \(P(T < \infty) = 1\). So there exists a real number \(c > 0\) such that \(P(T \leq c - 1) \geq 1 - \varepsilon\).

We choose \(n_0\) such that for \(n \geq n_0\)
\[(3.5) \quad \left|Q_n\Pi_1^{-1}(T^c < \infty) - P^c(T^c < \infty)\right| \leq \varepsilon.
\]
(According to Lemma 2.4 \(T^c < \infty\) is a continuity set with respect to \(P^c\). (3.5) then follows by Donsker’s theorem.)

We infer from (3.5) and the setting of \(c\):
\[(3.6) \quad P(\Phi_s\Pi_s \neq \Phi) \leq \varepsilon,
\]
(3.7) \[ Q_n(\Phi \Pi_t \neq \Phi) \leq 2\varepsilon. \]

(We have \( \{ \Phi \Pi_t = \Phi \} \cap \{ T < \infty \} = \{ T \Pi_t < \infty \} = \{ T \leq c - 1 \}. \)

We choose \( n_1 \leq n_0 \) such that for \( n \geq n_1 \)

(3.8) \[ |Q_n(\Phi \Pi_t)^{-1}(A) - P^o \Phi^{-1}_t(A)| \leq \varepsilon. \]

(The element \( u \) doesn't belong to \( \partial A \) because we assumed \( A \subset D \). It is easily seen that \( (\Phi \Pi_t)^{-1}(\partial A) \subset \Phi^{-1}(\partial A) \), so we infer that \( P(\Phi \Pi_t)^{-1}(\partial A) = P^o \Phi^{-1}_t(\partial A) = 0 \) and the existence of an \( n_1 \), such that (3.8) holds then follows from (3.3).)

For \( n \geq n_1 \) we have:

\[
|Q_n \Phi^{-1}(A) - P\Phi^{-1}(A)| \leq |Q_n \Phi^{-1}(A) - Q_n(\Phi \Pi_t)^{-1}(A)| \\
+ |Q_n(\Phi \Pi_t)^{-1}(A) - P^o \Phi^{-1}_t(A)| \\
+ |P(\Phi \Pi_t)^{-1}(A) - P\Phi^{-1}(A)| \\
\leq Q_n(\Phi \neq \Phi \Pi_t) + \varepsilon + P(\Phi \neq \Phi \Pi_t) \leq 4\varepsilon.
\]

So \( \lim_{n \to \infty} Q_n \Phi^{-1}(A) = P\Phi^{-1}(A) \) which is (3.4) and the proof is complete.

So far we have proved that \( Y^+_n \) converges weakly to \( P\Phi^{-1} \) which is \( W(T + t) - W(T) \) \( 0 \leq t \leq 1 \). It remains to identify \( W(T + \cdot) - W(T) \) with the Brownian meander \( W^+ \). But this clearly follows from Iglehart's result. We give a sketch of a proof using the methods of the present paper: Let \( X_t = \pm 1 \) each with probability \( \frac{1}{2} \). Set \( \mu_n = \inf \{ k \leq n : \text{the sequence } S_n, \ldots, S_n \text{ does not change sign} \} \) and let \( \nu_n = n - \mu_n \) (remark that \( \nu_n \geq 1 \)). We define \( \tilde{Y}_n(t) \) as follows:

\[
\tilde{Y}_n(k/\nu_n) = (1/\nu_n)^i[S_{\nu_n+i}] \quad \text{for} \quad 0 \leq k \leq \nu_n \\text{and linearly interpolated elsewhere}. \]

\( \tilde{Y}_n(\cdot) \) has the same distribution as \( Y^+_n(\cdot) \) where \( \{ Y^+_n \}_{n \in \mathbb{N}} \) and \( \nu_n \) are independent.

Define \( \tau' : C^1 \to [0, 1] \) by \( \tau'(f) = \inf \{ t \in [0, 1] : f(s) \text{ does not change sign for } s \in [t, 1] \} \). Further, define \( \Psi : C^1 \to C^1 \) by \( \Psi(f)(t) = |(1 - \tau')^{-1}f((1 - \tau')t)| \) for \( \tau' \in [0, 1] \), and \( \Psi(f) \) identically zero for \( \tau' = 1 \). We then have \( \tilde{Y}_n = \Psi(Y_n) \), which is identical in law to \( Y^+_n \). Now \( \tau' = \tau = \sup \{ t \in [0, 1] : f(t) = 0 \} \) \( P^1 \)-a.s.

(This can be proved in the same way as the statements of Lemma 2.3). So \( W^+ \) has the same distribution as \( \Psi(W) \). It can be shown by the same methods as in Lemma 2.4 and 2.5 that \( \Psi \) is \( P^1 \)-a.s. continuous on \( C^1 \), \( \rho \). The continuous mapping theorem implies \( \tilde{Y}_n \to W^+ \) and so \( Y^+_n \to W^+ \) in distribution. By Theorem 3.2 \( Y^+_n \to W(T + \cdot) - W(T) \). Clearly \( \nu_n \to \infty \) in distribution. This is sufficient for \( Y^+_n \to W(T + \cdot) - W(T) \) because \( \{ Y^+_n \} \) and \( \nu_n \) are independent. It follows that \( W^+ \) and \( W(T + \cdot) - W(T) \) have the same distribution.

Acknowledgment. I would like to thank L. Rogge and the referee for helping in many ways to improve the original manuscript.

REFERENCES


FACHBEREICH STATISTIK
UNIVERSITÄT KONSTANZ
D-775 KONSTANZ
POSTFACH 7733
GERMANY