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 \( Y_n(t) \) be \( S_k/\sigma n^{1/2} \) 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: [https://doi.org/10.1214/aop/1176996098](https://doi.org/10.1214/aop/1176996098)

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: [https://doi.org/10.5167/uzh-23167](https://doi.org/10.5167/uzh-23167)

Originally published at:
DOI: [https://doi.org/10.1214/aop/1176996098](https://doi.org/10.1214/aop/1176996098)
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.\,rv with \( E(X_k) = 0 \) and \( E(X_k^2) = \sigma^2 \), \( 0 < \sigma^2 < \infty \). Set \( S_n = X_1 + \cdots + X_n \). Let \( Y_n(t) \) be \( S_n/\sigma n^\alpha \) 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.\,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 \( Y_n(k/n) = S_k/\sigma n^\alpha \) 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

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

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

Iglehart assumed \( E|X_i|^\alpha < \infty \) and \( X_i \) non-lattice 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 $\mathcal{B}^s$ the smallest $\sigma$-algebra such that the mappings $C^* \ni f \mapsto f(t) \in \mathbb{R}$ are measurable.

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

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

$$(2.1) \quad T^*(f) = \inf\{t: f(u) \geq f(t) \quad \text{for} \quad t \leq u \leq t + 1 \leq s\}, \quad (\inf \emptyset = \infty) .$$

We set $T = T^\infty$ and $P = P^\infty$ 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 $\{T^s \leq u\} = \bigcap_{n \in \mathbb{N}} \{f \in C^*: \exists \text{ there exists a rational } r \leq u + 1/n \text{ with } f(r) < \min_{1 \leq i \leq n} f(r + i/n) + 1/n\}$, which is easily seen to belong to $\mathcal{B}^s$.

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

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

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

$$(2.2) \quad P^s(f(T^s) = f(T^s + 1)) = 0 ;$$

$$(2.3) \quad P^s(T^s = s - 1) = 0 ;$$

$$(2.4) \quad P^s(\text{ex. u} \in (0, 1) \text{ with } f(T^s) = f(T^s + u)) = 0 .$$

**Proof.** We set $m(t) = \min_{0 \leq s \leq t} W(t)$. $D(t) = W(t) - m(t)$ has the same finite-dimensional distributions as $|W(t)|$ (see [5] page 193). Observe now that $T^s = \inf\{t \leq s - 1: m(t) = m(t + 1)\}$. Now $T^s = s - 1$ implies $D(s - 1) = 0$ which has $P$ measure 0. This proves (2.3).

Let $U = \{\text{ex. } u < v < w \text{ with } m(u) = m(v) = m(w) \text{ and } D(u) = D(v) = D(w) = 0\}$. Then $U \subset \bigcup_{r,s \in \mathbb{Q}} \{\min_{0 \leq s \leq r} W(t) = \min_{0 \leq s \leq r+1} W(t)\}$ and the last has $P$ measure 0. This proves (2.4).

It suffices to prove (2.2) for $s = \infty$. With probability one, the hitting time process $\{T_x: x \geq 0\}$ $(T_x = \inf\{t: W(t) = -x\})$ has no jumps of length one. This follows from its Lévy decomposition (see Section 1.7 of [4]). Together with $P(U) = 0$ this yields (2.2).

**Lemma 2.4.** For each $s \in (0, \infty)$ $T^s$ is a continuous $P^s$ a.e. on $(C^s, \rho)$.

**Proof.** By (2.3) it suffices to consider the case $s = \infty$. Let $f$ be such that $T(f) < \infty$ and $f$ does not belong to the null sets defined in (2.2)—(2.4).

(I) We first prove that for all $\delta > 0$ there exists an $\varepsilon > 0$ with

$$ T(f') \leq T(f) + \delta \quad \text{when} \quad \rho(f, f') < \varepsilon . $$

By (2.2) there is as $\tau < \delta$ so that
\[ \inf_{T+1 \leq u \leq T+1+\tau} f(u) > f(T). \]

Now (2.4) gives $\varepsilon = \frac{1}{4} (\inf_{T+1 \leq u \leq T+1+\tau} f(u) - f(T)) > 0$.

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

(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 $\{f_n\}_{n \in N}$ 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:
\[
\inf_{T\leq u \leq T+\lambda+1} f(u) \geq \inf_{T\leq f(f_n) \leq T+\lambda+1} f(u) - \varepsilon \\
\geq \inf_{T\leq f(f_n) \leq T+\lambda+1} f(u) - 2\varepsilon \\
\geq f_n(T(f_n)) - 2\varepsilon \geq f(T(f_n)) - 3\varepsilon \geq f(\lambda) - 4\varepsilon.
\]

So $\inf_{T\leq u \leq T+\lambda+1} f(u) \geq f(\lambda)$ 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 $\Phi_s: C^s \to C^1$
\[ \Phi_s(f)(t) = f(Ts(f) + t) \quad \text{for} \quad Ts(f) < \infty \]
\[ = u \quad \text{for} \quad Ts(f) = \infty. \]

We write $\Phi = \Phi_s$ for simplicity.

A straightforward conclusion of Lemma 2.4 is

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

3. Sums of independent random variables conditioned to stay positive. Let $X_1, X_2, \ldots$, be i.i.d. rv with $E(X_i) = 0; E(X_i^2) = \sigma^2 < \infty$ ($\sigma^2 > 0$) and $S_k = \sum_{j=1}^k X_j$. $T_n = \inf\{k: S_{k+i} \geq S_k \text{ for } i = 1, \ldots, n\}$. 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 $a_1, \ldots, a_n$
\[ (3.1) \quad P(S_k \leq a_k, k = 1, \ldots, n | S_k \geq 0, k = 1, \ldots, n) \]
\[ = P(Z_k \leq a_k, k = 1, \ldots, n). \]

**Proof.** This is an easy consequence of the independence and identical distribution of the $X_i$.

If $B_j = \bigcup_{i=0}^{j-1} \{S_i \leq S_r \text{ for } s + 1 \leq r \leq \min(j, s + n)\}$ we have
\[ P(S_{T_n+k} - S_{T_n} \leq a_k \text{ for } k = 1, \ldots, n) \]
\[ = \sum_{j=0}^{\infty} P(S_{j+k} - S_j \leq a_k \text{ for } k = 1, \ldots, n | T_n = j)P(T_n = j) \]
\[ = \sum_{j=0}^{\infty} P(S_{j+k} = S_j \leq a_k \text{ for } k = 1, \ldots, n | S_{j+k} \geq S_j \text{ for } k = 1, \ldots, n \text{ and } B_t^s) P(T_n = j) \]
\[ = P(S_k \leq a_k, k = 1, \ldots, n | S_k \geq 0, k = 1, \ldots, n) \]
\[ \text{since } T_n < \infty \text{ a.s.} \]

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

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

Let \( Q_n \Pi_s^{-1}(dx | C^+) \) be the probability measure on \( C^c \) which is defined by
\[ Q_n \Pi_s^{-1}(A | C^+) = Q_n(\Pi_s^{-1}(A \cap C^+))/Q_n(\Pi_s^{-1}(C^+)) \]
for \( A \in B^1 \).

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

**Proof.** We have proved in Lemma 3.1 that
\[ (3.2) \quad Q_n \Pi_s^{-1}(dx | C^+) = Q_n \Phi^{-1}(dx) \quad \text{holds.} \]
Now by Donsker's theorem (see [2]), \( Q_n \Pi_s^{-1} \) converges weakly to \( P^s \) for \( s < \infty \).

With regard to Lemma 2.5 we have for \( s < \infty \)
\[ (3.3) \quad Q_n(\Phi_s \Pi_s)^{-1} \to P^s \Phi_s^{-1} \quad \text{weakly.} \]
(Theorem 5.1 in [2].)

Let \( A \) be a continuity set in \( B^1 \), 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
\[ D = \{ f \in C^1 : \min_{x \in I_{1/2}} f(t) \geq -1/2 \} \]
Without loss of generality we can assume \( A \subset D \). (If not: replace \( A \) by \( A \cap D \) noticing \( Q_n \Phi^{-1}(D^c) = P \Phi^{-1}(D^c) = P \Phi^{-1}(\partial D) = 0 \).

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 \( n \geq n_0 \)
\[ (3.5) \quad |Q_n \Pi_s^{-1}(T^e < \infty) - P^s(T^e < \infty)| \leq \varepsilon. \]
(According to Lemma 2.4 \( T^e < \infty \) is a continuity set with respect to \( P^e \). (3.5) then follows by Donsker's theorem.)

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

(We have \( \{\Phi, \Pi e = \Phi\} \cap \{T < \infty\} = \{T^e \Pi e < \infty\} = \{T \leq c - 1\}\).)

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

\[ |Q_n(\Phi e \Pi e)^{-1}(A) - P^o \Phi e^{-1}(A)| \leq \varepsilon. \]

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

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

\[
\begin{align*}
|Q_n \Phi^{-1}(A) - P\Phi^{-1}(A)| & \leq |Q_n \Phi^{-1}(A) - Q_n(\Phi e \Pi e)^{-1}(A)| \\
& \quad + |Q_n(\Phi e \Pi e)^{-1}(A) - P^o \Phi e^{-1}(A)| \\
& \quad + |P(\Phi e \Pi e)^{-1}(A) - P\Phi^{-1}(A)| \\
& \leq Q_n(\Phi \neq \Phi e \Pi e) + \varepsilon + P(\Phi \neq \Phi e \Pi e) \leq 4\varepsilon.
\end{align*}
\]

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 \( Y^+ \) 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 \( X = \pm 1 \) each with probability \( \frac{1}{2} \). Set \( \mu_n = \inf\{k \leq n: \text{the sequence } S_n, \ldots, S_k \text{ does not change sign}\} \) and let \( \nu_n = n - \mu_n \) (remark that \( \nu_n \geq 1 \)). We define \( \bar{Y}_n(t) \) as follows:

\[
\bar{Y}_n(k/n) = (1/n) |S_{k/n}| \quad \text{for } 0 \leq k \leq n \text{ and linearly interpolated elsewhere.}
\]

\( \bar{Y}_n(\cdot) \) has the same distribution as \( Y^+_{\nu_n}(\cdot) \) where \( \{Y^+_{\nu_n}\}_{\nu_n} \) and \( \nu_n \) are independent. Define \( \tau^e: C^1 \to [0, 1] \) by \( \tau^e(f) = \inf\{t \in [0, 1]: f(s) \text{ does not change sign for } s \in [t, 1] \} \). Further, define \( \Psi: C^1 \to C^1 \) by \( \Psi(f)(t) = |(1 - \tau^e - \tau f(\tau^e + (1 - \tau^e) t)| \) for \( \tau^e \in [0, 1] \), and \( \Psi(f) \) identically zero for \( \tau^e = 1 \). We then have \( \bar{Y}_n = \Psi(Y_n) \), which is identical in law to \( Y^+ \). Now \( \tau^e = \tau = \sup\{t \in [0, 1]: f(t) = 0\} \) \( P^1 \)-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 \( \bar{Y}_n \to W^+ \) and so \( Y^+_{\nu_n} \to W^+ \) in distribution. By Theorem 3.2 \( Y^+_{\nu_n} \to W(T +, \cdot) - W(T) \). Clearly \( \nu_n \to \infty \) in distribution. This is sufficient for \( Y^+_{\nu_n} \to W(T +, \cdot) - W(T) \) because \( \{Y^+_{\nu_n}\} \) 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