RATES OF CONVERGENCE FOR SOME FUNCTIONALS IN PROBABILITY\footnote{These results were obtained while the author was being partially supported by the National Science Foundation under grant NSF GP-21063.}

BY STANLEY SAWYER

Yeshiva University

Let \( \{x_1, x_2, \ldots \} \) be a sequence of i.i.d.r.v. with mean zero, variance one, and \( \{1/P(|x_\lambda| \geq \lambda) \leq C \exp(-\alpha \lambda^2) \) for positive \( \alpha, \varepsilon \). Let \( f(t, x) \) (with its first partial derivatives) be of slow growth in \( x \), let \( F_n(x) \) be the distribution function of \((1/n) \sum_{i=1}^{n} f(k/n, s_k/n) \) where \( s_k = x_1 + x_2 + \cdots + x_k \), and let \( F(x) \) be the distribution function of \( \int_0^x f(t, w(t)) \, dt \) where \( |w(t)| \) is Brownian motion. Then \( \sup_{x} |P_n(x) - P(x)| = O((\log n)^{3/2}/n^{1/2}) \) provided \( P(x) \) has a bounded derivative. The proof uses the Skorokhod representation; also, a theorem is proven which would indicate that the Skorokhod representation cannot be used in general to obtain a rate of convergence better than \( O(1/n^4) \). A corresponding result is obtained if (1) is replaced by the existence of a finite \( p \)th moment, \( p \geq 4 \).

1. Introduction. Let \( \{x_1, x_2, \ldots, x_n, \ldots \} \) be a sequence of independent and identically distributed random variables with \( E(x_n) = 0 \), \( E(x_n^2) = 1 \), and

\[
P(|x_\lambda| \geq \lambda) \leq C e^{-\alpha \lambda^2}
\]

for some \( \alpha > 0 \), \( \varepsilon > 0 \). Let \( s_n = x_1 + x_2 + \cdots + x_n \) and let \( \{w(t): 0 \leq t < \infty\} \) be standard Brownian motion. Then (see Section 2)

**Theorem 1.** Let \( f(s, x) \in C(R^n) \) be a function such that \( f \) and its partial derivatives of order one are of slow growth in \( x \); i.e. satisfy inequalities of the form

\[
|Df(s, x)| \leq \Omega (1 + |x|^\varepsilon),
\]

and assume that the probability distribution \( P(\sum f(t, w(t)) \, dt \leq \lambda) \) has a bounded density (i.e., bounded derivative in \( \lambda \). Then, for \( \{x_n\} \) as above

\[
\sup \left| P(\frac{1}{n} \sum f\left(\frac{k}{n}, \frac{s_k}{n}\right) \leq \lambda) - P(\sum f(t, w(t)) \, dt \leq \lambda) \right| = O\left(\frac{\log n}{n^\beta}\right)
\]

where \( \beta = \beta(\varepsilon, \alpha) \).

If, in Theorem 1, (1.1) is replaced by

\[
(1.1)' \quad E(|x_n|^p) < \infty
\]

for some \( p \geq 4 \), the arguments in Section 2 also go through, and the difference in (1.2) has the uniform bound

---

Received January 19, 1971.
\begin{equation}
O\left(\frac{(\log n)^\beta}{n^{\frac{a}{2p+a+b}}}\right), \quad \beta = ap/8.
\end{equation}

For example, if the \{x_k\} have a moment generating function and \(f(s, x) = x^a\) (so that \(\epsilon = 1, a = 2\)), we have

\[
\sup_x \left| P \left( \frac{1}{n^a} \sum_1^a s_k \leq \lambda \right) - P \left( \int_0^1 \gamma(t) f(w(t)) \, dt \leq \lambda \right) \right| = O\left( \frac{(\log n)^{a}}{n^a} \right).
\]

If we only assumed \(E(x_k^a) < \infty\), the rate would be \(O(\log n/n^4)\).

The limiting distribution in Theorem 1 is an application of the Invariance Principle ([1]). The density condition in the above example is satisfied, since for any \(\gamma(t) \in L^2(0, 1), \int_0^1 \gamma(t) f(w(t)) \, dt \) is known to have an integrable characteristic function and hence a bounded continuous density (see Section 4).

Skorokhod (1965) considered a similar problem, and proved

\[
E \left[ g \left( \frac{1}{n} \sum_1^a f \left( \frac{k}{n}, \frac{s_k}{n^3} \right) \right) \right] = E[g(\int_0^1 f(t, w(t)) \, dt)] + A/n^4 + O(1/n)
\]

for uniformly bounded random variables \(x_k\), a fixed expression \(A\), and any function \(g(y)\) with a bounded second derivative. Skorokhod's proof, however, is not sufficient to yield (1.2).

The proof of Theorem 1 is based on the Skorokhod representation ([1], [14]) and a martingale inequality of Burkholder. The Skorokhod representation, applied to the variables \(x_k/n^4\), provides random times \(\{\tau_j^{(n)}\}\) such that if

\begin{equation}
x_n(k/n) = w(\sum_1^a \tau_j^{(n)}), \quad x_n(t) = x_n(k/n), \quad k/n \leq t < (k + 1)/n, 0 \leq t \leq 1,
\end{equation}

the variables \(x_n((k/n) - x_n((k - 1)/n); 1 \leq k \leq n\) are independent and have the same distribution as \(x_n/n^4\). Consequently \(x_n(k/n) \equiv s_k/n^4\) and

\[
\frac{1}{n} \sum_1^a f \left( \frac{k}{n}, \frac{s_k}{n^3} \right) \equiv \frac{1}{n} \sum_1^a f \left( \frac{k}{n}, w(\sum_1^a \tau_j^{(n)}) \right).
\]

Theorem 1 is derived from

**Proposition.** For all \(b < \infty\) and \(\beta = \beta(\epsilon, a)\) as before

\begin{equation}
P \left[ \left| \frac{1}{n} \sum_1^a f \left( \frac{k}{n}, w(\sum_1^a \tau_j^{(n)}) \right) - \int_0^1 f(t, w(t)) \, dt \right| \geq \frac{(\log n)^{\beta}}{n^{\beta}} \right] = O(1/n^4).
\end{equation}

In contrast, given a functional which depends essentially on a single time, such a result would be impossible. That is

**Theorem 2.** (See Section 3.) Let \(x_k\) be a sequence of identically distributed independent random variables with \(E(x_k) = 0, E(x_k^2) = 1,\) and \(E(x_k^4) < \infty,\) and define \(x_n(t)\) by (1.4). Then
\[ (1.6) \quad \lim_{n \to \infty} P\left( x_n(1) - w(1) \leq \frac{c}{n^\delta} \right) = \frac{1}{\pi} \int_0^r e^{-u^2} \int_0^{u^2} e^{-v^2} \, dv \, du \]

where \( c \) is a positive constant (\( c = \sigma(\tau_1) \)) satisfying \( 0.75 \, E(x_k)^4 < c < 1.25 \, E(x_k)^4 \).

As a consequence

\[ (1.7) \quad \lim_{n \to \infty} P\left( \max_{0 \leq t \leq 1} |x_n(t) - w(t)| \geq \frac{A}{n^\beta \log n} \right) = 1 \]

for all \( A > 0 \). Hence it appears that, except in special cases such as (1.2), the Skorokhod representation cannot be used to prove rates of convergence better than \( O(n^{-4}) \). In fact, it was this negative result which led us initially to (1.2) in an unsuccessful attempt to find a (true) rate of convergence worse than \( O(n^{-4}) \). Thus it would appear (i.e., one would conjecture) that all rates of convergence coming from the Invariance Principle for decent \( \{x_n\} \) are \( O(n^{-4}) \). (Here "decent" means \( E(|x_n|) < \infty \).)

In view of Theorem 2, the rate of convergence obtained by Rosenkrantz (1969) for the von Mises statistic cannot be improved beyond \( O(n^{-4}) \), at least by his methods. It is not known, however, whether the approach of Sazonov (1969) suffers a similar limitation, although the results obtained are not as strong.

A partial converse of (1.7) has been obtained by Fraser (1971), Section 7. For \( x_k \) satisfying (1.1), Fraser (essentially) shows

\[ P\left( \max_{0 \leq t \leq 1} |x_n(t) - w(t)| \geq \frac{(\log n)^{\beta}}{n^\delta} \right) = O(1/n^\delta) \]

for all finite \( b \), where \( \beta \) is as in Theorem 1 (\( a = 1 \)). In particular, arguing as in Rosenkrantz (1967), one concludes

\[ \sup_{\lambda} |P(\Phi(x(\cdot)) \leq \lambda) - P(\Phi(w(\cdot)) \leq \lambda)| = O\left(\frac{(\log n)^{\beta}}{n^\delta}\right) \]

where \( \Phi(x(\cdot)) \) is any functional defined and uniformly Lipschitz on the continuous functions on \([0, 1]\), which is also such that the probability distribution of \( \Phi(w(\cdot)) \) has a bounded density.

See references [7]-[13] for other uniform rate of convergence theorems arising from the Invariance Principle. A version of Theorem 2 in the form of a law of the iterated logarithm has been given by Kiefer (1969).

2. Proof of Theorem 1.

Proof of Theorem 1 Given Equation (1.5). Set \( \delta = (\log n)^b/n^\delta \). Then, by (1.5), for all \( \lambda \) and \( b < \infty \)
(2.1) \[ P\left( \frac{1}{n} \sum_{k} f\left( x_k, \frac{s_k}{n^2} \right) \leq \lambda \right) \leq P\left( \frac{1}{n} \sum_{k} f\left( x_k, \frac{s_k}{n^2} \right) \leq \lambda + \delta \right) + O(1/n^4) \]

\[ \geq P\left( \frac{1}{n} \sum_{k} f\left( x_k, \frac{s_k}{n^2} \right) \leq \lambda - \delta \right) - O(1/n^4) \]

\[ \left| P\left( \frac{1}{n} \sum_{k} f\left( x_k, \frac{s_k}{n^2} \right) \leq \lambda \right) - P\left( \frac{1}{n} \sum_{k} f\left( x_k, \frac{s_k}{n^2} \right) \leq \lambda - \delta \right) \right| \]

\[ \leq P\left( \lambda - \delta \leq \frac{1}{n} \sum_{k} f\left( x_k, \frac{s_k}{n^2} \right) \leq \lambda + \delta \right) + O(1/n^4). \]

By hypothesis, the distribution \( P\left( \frac{1}{n} \sum_{k} f\left( x_k, \frac{s_k}{n^2} \right) \leq \lambda \right) \) has a density bounded by some constant, say \( L \). Hence the difference in (2.1) has the uniform bound \( 2L\delta + O(1/n^4) \).

We now state a lemma.

**Lemma 1.** Let \( y_1, y_2, \ldots, y_n \) be identically distributed independent random variables with \( E(y_k) = 0 \), and let \( d_1, d_2, \ldots, d_n \) be random variables with \( |d_k| \leq M \). Assume further that each \( d_k \) is \( \mathcal{B}_{k-1} \) measurable, where, for \( 1 \leq k \leq n \), \( \mathcal{B}_k \) is a \( \sigma \)-algebra which is independent of \( y_{k+1}, \ldots, y_n \). (E.g. \( \mathcal{B}(y_1, y_2, \ldots, y_k) \). Then, there exist universal constants \( C_p \) such that

\[ (2.2) \quad E\left( \left| \frac{d_1 y_1 + d_2 y_2 + \cdots + d_n y_n}{n^{1/p}} \right|^p \right) \leq C_p M^p E\left( |y_1|^p \right) \quad 2 \leq p < \infty. \]

**Proof.** The basic step is a martingale inequality of Burkholder (1966) ([2] page 1502),

\[ E(\sum_{k} d_k y_k^p) = C_p E(\sum_{k} d_k^2 y_k^2)^{p/2} \]

for \( 1 < p < \infty \). Thus

\[ E\left( \left| \frac{d_1 y_1 + d_2 y_2 + \cdots + d_n y_n}{n^{1/p}} \right|^p \right) \leq C_p E\left( \left( \frac{d_1^2 y_1^2 + \cdots + d_n^2 y_n^2}{n} \right)^{p/2} \right) \]

\[ \leq C_p M^p E\left( \left( \frac{1}{n} \sum_{k} y_k^2 \right)^{p/2} \right) \]

\[ \leq C_p M^p E\left( \frac{1}{n} \sum_{k} |y_k|^p \right) \]

\[ = C_p M^p E(|y_1|^p) \]

by the identity \( |(1/n) \sum_{k} a_k|^p \leq (1/n) \sum_{k} |a_k|^p \) for real numbers and \( p \geq 1 \).

**Proof of Equation (1.5).** For any \( A > 1 \), define \( f^A(s, x) \in C^0(R^p) \) such that

\[ (2.3) \quad f^A(s, x) = f(s, x), \quad |x| \leq A, \]

\[ |Df^A(s, x)| \leq 2\Omega(1 + A^a) = M \]

uniformly in \( s \) and \( x \), where \( D \) is either the identity operator or a first partial derivative. Thus
(2.4) \[ P\left( \left| \frac{1}{n} n \sum_{k} f\left( \frac{k}{n}, w\left( \sum_{j} \tau_{j}^{(n)} \right) \right) - \int_{0}^{1} f(t, w(t)) \, dt \right| \geq 6 \delta \right) \]
\[ \leq P\left( \left| \frac{1}{n} n \sum_{k} f\left( \frac{k}{n}, w\left( \sum_{j} \tau_{j}^{(n)} \right) \right) - \int_{0}^{1} f(t, w(t)) \, dt \right| \geq 6 \delta \right) \]
\[ + P(\max_{0 \leq t \leq 1} |w(t)| \geq A) + P(\sum_{j} \tau_{j} > 2), \]
where here and in the following we abbreviate \( \tau_{j} = \tau_{j}^{(n)} \) and \( \tau = \tau_{i}^{(1)} \). Thus
\[ P(\max_{0 \leq t \leq 1} |w(t)| \geq A) \leq 2P(\max_{0 \leq t \leq 1} |w(t)| \geq A) = 2 \left( \frac{2}{\pi} \right)^{\frac{1}{4}} \int_{0}^{\infty} e^{-\frac{u^{2}}{4A}} \, du \]
\[ = 2 \left( \frac{2}{\pi} \right)^{\frac{1}{4}} \int_{0}^{\infty} e^{-\frac{u^{2}}{4A}} \, du \leq 2e^{-\frac{A}{4}}. \]
(see [3] page 392) and
\[ P(\max_{0 \leq t \leq 1} |w(t)| \geq A) = P(\max_{0 \leq t \leq 1} |w(t)| \geq A/2) \leq 2e^{-\frac{A}{4}}. \]
Also, the random times \( \{\tau_{j}^{(n)}\} \) are independent and identically distributed for a fixed \( n, \tau_{j}^{(n)} \approx (1/n) \tau \), and \( E(\tau) = E(x_{1}) = 1 \) (see [1]). Thus by Lemma 1
\[ P(\sum_{j} \tau_{j} > 2) = P(n^{-\frac{1}{2}} \sum_{j} (n_{j} - 1) > n^{\frac{1}{2}}) \]
\[ \leq \frac{1}{n^{p/2}} E\left( \left| n^{-\frac{1}{2}} \sum_{j} (n_{j} - 1) \right|^{p} \right) \]
\[ \leq \frac{1}{n^{p/2}} C_{p} E(\left| \tau - 1 \right|^{p}), \quad p \geq 2. \]
Now, for all \( p \geq 1 \), by Sawyer (1967), Section 2
(2.5) \[ E(\tau^{p}) \leq 4p \Gamma(p) E(|x_{1}|^{p}) < \infty. \]
Finally, if \( A = \log n \), the last two terms in (2.4) are both \( O(1/n^{p}) \) for all \( b \).
We apply Taylor's formula to \( f^{b}(s, x) \), and suppress the superscript \( A \) in \( f^{b}(s, x) \) from this point on.
\[ \int_{0}^{1} f(t, w(t)) \, dt \]
\[ = \sum_{j} \left( f\left( \frac{k}{n}, w\left( \sum_{j} \tau_{j} \right) \right) \right) \tau_{k+1} \]
\[ = \sum_{j} \left( f\left( \frac{k}{n}, w\left( \sum_{j} \tau_{j} \right) \right) \right) \tau_{k+1} \]
(2.6) \[ + \sum_{j} \int_{0}^{1} \left[ \frac{\partial f}{\partial s} \left( \frac{k}{n}, w\left( \sum_{j} \tau_{j} \right) \right) \right] \, ds \]
\[ = \sum_{j} \int_{0}^{1} \left[ \frac{\partial f}{\partial s} \left( \frac{k}{n}, w\left( \sum_{j} \tau_{j} \right) \right) \right] \, ds \]
where \( \zeta_{k} = \sum_{j} \tau_{j}^{(n)} \). Thus
\begin{align*}
\frac{1}{n} \sum_{i}^{n} f\left(\frac{k}{n}, w(\sum_{i}^{n} \tau_{i})\right) dt - \frac{1}{n} \sum_{i}^{n} f\left(\frac{k}{n}, w(\sum_{i}^{n} \tau_{i})\right) \\
= \frac{1}{n} \sum_{\tau_{i} \neq 0}^{n} f\left(\frac{k}{n}, w(\sum_{i}^{n} \tau_{i})\right) (\tau_{k+1} - 1) \\
- \int_{0}^{t} f(t, w(t)) dt + (1/n) f(0, 0) - (1/n) f(1, w(\sum_{i}^{n} \tau_{i})) \\
+ \Phi_{6} + \Phi_{7},
\end{align*}

where we let \( \Phi_{1}, \ldots, \Phi_{4} \) be the first four terms in (2.7) and \( \Phi_{5}, \Phi_{6} \) the last two terms in (2.6). We estimate the difference in (2.7) as follows:

\begin{align*}
P\left( \left| \frac{1}{n} \sum_{i}^{n} f\left(\frac{k}{n}, w(\sum_{i}^{n} \tau_{i})\right) dt - \frac{1}{n} \sum_{i}^{n} f\left(\frac{k}{n}, w(\sum_{i}^{n} \tau_{i})\right) \right| > 6\delta \right) \\
\leq \sum_{i}^{n} P(|\Phi_{k}| > \delta) \\
= \frac{1}{(n\delta)^{p}} \sum_{i}^{n} E(|n^{p} \Phi_{k}|^{p})
\end{align*}

for all \( p > 0 \). In the first term we apply Lemma 1 with \( d_{k} = f(k/n, w(\sum_{i}^{n} \tau_{i})) \) and \( y_{k} = n\tau_{k+1} - 1 \).

\begin{align*}
E(|n^{p} \Phi_{k}|^{p}) \leq C_{p} M^{p} E(|\tau - 1|^{p}).
\end{align*}

Now an inspection of the proof of Burkholder's inequality shows that \( C_{p} = O((c_{0} p)^{p}) \) as \( p \to \infty \), while by (2.5) and (1.1)

\begin{align*}
E(\tau^{p}) = O(p^{p}(2p/\epsilon)^{p/2}).
\end{align*}

Hence there exists a constant \( c > 0 \) such that

\begin{align*}
E(|n^{p} \Phi_{k}|^{p}) \leq \Omega_{0} M^{p}(cp)^{p}.
\end{align*}

Now if \( p = \gamma \log n \) for \( \gamma = 1/c \) and \( \delta = (\log n)^{\rho}/n^{\gamma}, \ n \geq 3 \), then by (2.3)

\begin{align*}
(n\delta)^{-p} E(|n^{p} \Phi_{k}|^{p}) \leq \Omega_{0} (4\Omega_{0})^{\gamma \log n} ((\log n)^{\rho} \log n)/((\log n)^{\log n})/((\log n)^{\rho} \log n) \\
= \Omega_{0} n^{\rho \log \log n} \\
= O(1/n^{\rho \log \log n})
\end{align*}

for some \( \epsilon > 0 \) provided \( (\beta - a)/\epsilon - 1 > 0 \); i.e. \( \beta > a + c \). In particular any term satisfying an estimate of the form

\begin{align*}
E(|n^{p} \Phi_{k}|^{p}) \leq \Omega_{0} C_{p} M^{p} E(\tau^{mp}) E(|x|^{p})
\end{align*}

can be estimated similarly, perhaps with a larger \( \beta \).

For the second term in (2.8)

\begin{align*}
||\Phi_{4}|| = \left| \int_{0}^{t} f(t, w(t)) dt \right| \leq M |\sum_{i}^{n} \tau_{i} - 1| \\
|n^{p} \Phi_{4}| \leq (M/n^{b}) |\sum_{i}^{n} (n\tau_{i} - 1)| \\
E(|n^{p} \Phi_{4}|^{p}) \leq C_{p} M^{p} E(|\tau - 1|^{p})
\end{align*}

which is exactly the same as (2.9). The terms \( \Phi_{5} \) and \( \Phi_{6} \) give no trouble,
since after multiplication by \( n^k \) they converge uniformly to zero. Using the inequality \(|s - k/n| \leq |s - \sum_{j=1}^k \tau_j| + |\sum_{j=1}^k \tau_j - k/n|\) in the integral in \( \Phi_0 \) and integrating, we obtain
\[
|\Phi_0| \leq M \sum_{j=0}^{n-1} \left( \frac{1}{2} \tau_{k+1}^2 + \tau_{k+1} |\sum_{j=1}^k \tau_j - k/n| \right)
\]
\[
|n^k \Phi_0| \leq \frac{M}{n^k} \sum_{j=0}^{n-1} \left( n \tau_{k}^2 + \frac{M}{n} \sum_{j=1}^{n} (n \tau_{k})^p \left| \frac{1}{n} \sum_{j=1}^{k-1} (n \tau_{j} - 1) \right|^p \right)
\]
\[
E(|n^k \Phi_0|^p) \leq \frac{(2M)^p}{n^{p/2}} E(\tau^p) + \frac{(2M)^p}{n^k} \sum_{j=0}^{n-1} E \left( (n \tau_{k})^p \left| \frac{1}{n} \sum_{j=1}^{k-1} (n \tau_{j} - 1) \right|^p \right)
\]
\[
\leq \frac{(2M)^p}{n^{p/2}} E(\tau^p) + C_p (2M)^p E(\tau^p) E(|\tau - 1|^p)
\]
by independence and Lemma 1. This is of the form (2.10); for the sixth term:
\[
|n^k \Phi_0| \leq \frac{M}{n} \sum_{j=0}^{n-1} mn^k \sum_{j=1}^{k+1} |w(s) - w(\sum_{j=1}^k \tau_j)| \ ds.
\]
The terms in the series above are independent and identically distributed by construction, and
\[
\sum_{j=1}^{k+1} |w(s) - w(\sum_{j=1}^k \tau_j)| \ ds \equiv \frac{1}{mn^k} \sum_{j=1}^{k+1} |w(s)| \ ds.
\]
This is because \( n \tau_{1}^{(n)} \) bears the same relation to \( n^k \omega(t/n) \) as \( \tau \) does to \( w(t) \) (see [1]) and thus
\[
\sum_{j=1}^{k+1} |w(s)| \ ds = \frac{1}{n} \sum_{j=1}^{n \tau_{1}^{(n)}} \left| w \left( \frac{s}{n} \right) \right| \ ds
\]
\[
= \frac{1}{mn^k} \sum_{j=1}^{n \tau_{1}^{(n)}} \left| n^k w \left( \frac{s}{n} \right) \right| \ ds \equiv \frac{1}{mn^k} \sum_{j=1}^{k+1} |w(s)| \ ds.
\]
Hence for \( p \geq 1 \)
\[
(2.11) \quad E(|n^k \Phi_0|^p) \leq M^p E(\sum_{j=1}^{k+1} |w(s)| \ ds)^p.
\]
Now \( \tau = \inf \{ t; \omega(t) \in [x, G(x)] \} \), where \( x \leq x_k \), \( x \) is independent of \( \{\omega(t)\} \), and \( G(y) \) is a certain function ([1], [14]). Thus
\[
\sum_{j=1}^{k+1} |w(s)| \ ds \leq \tau \left( |x| + |G(x)| \right)
\]
\[
E(\sum_{j=1}^{k+1} |w(s)| \ ds)^p \leq E(\tau^p) E(\left( |x| + |G(x)| \right)^p)^p
\]
Now if \( b \) is sufficiently small so that \( |G(\pm b)| < \infty, b > 0 \),
\[
E(|G(x)|^p) = \int |G(y)|^p P(x \in dy)
\]
\[
\leq |G(b)|^p + |G(-b)|^p + (1/b) \int |G(y)|^p |y| P(x \in dy)
\]
\[
\leq Q_b^p + (1/b) E(|x|^{p+1})
\]
and (2.11) reduces to an estimate of the form (2.10). We have now shown
that the right-hand side of (2.8) is $O(1/n^b)$ for every finite $b$, and the proof grinds to a halt.

**Corollary 1.** \[ \frac{1}{n} \sum_{i=1}^{n} f\left( \frac{k}{n}, w\left( \sum_{j=1}^{s_i} \tau_j \right) \right) = \frac{1}{n} f(t, w(t)) dt + O\left( \frac{(\log n)^{\delta}}{n^b} \right) \text{ a.s.} \]

**Proof.** Use the Borel-Cantelli lemma in (1.5).

**Corollary 2.** For any $g(y) \in C^1(R)$ with $\int |g'(y)| dy < \infty$,

\[ E\left[ g\left( \frac{1}{n} \sum_{i=1}^{n} f\left( \frac{k}{n}, \frac{n s_i}{n^t} \right) \right) \right] = E[g\left( \frac{1}{n} f(t, w(t)) dt \right)] + O\left( \frac{(\log n)^{\delta}}{n^b} \right). \]

**Proof.** For any random variable $Y$,

\[ E(g(Y)) = \int g(y) P(Y \in dy) = \int g'(y) P(Y \leq y) \, dy + g(\infty). \]

Now use (1.2).

**Remark.** To derive (1.3) assuming only that $E(|x_i|^q) < \infty$, $q \geq 4$, we set $\delta = 1/n^b$ and, in (2.8), estimate the sixth term and the first half of the fifth term with $p = q/4$, and the other terms with $p = q/2$. Using $p = (q - 1)/3$ in the sixth term would give the sharper estimate

(1.3') \[ O\left( \frac{(\log n)^{\delta}}{n^{(q-1)/(2p+1)}} \right). \]

**3. Proof of Theorem 2.** By properties of the Skorokhod representation (see [1] page 276 +), $\tau_{j}^{(n)} \equiv n^{-1} \tau$, where $\tau = \tau_{1}^{(1)}$ and $E(\tau) = E(x_{1}^{\delta}) = 1$. Define a constant $c > 0$ by $c^4 = \sigma^2(\tau) = E(\tau - 1)^2$. Since by Sawyer (1967), (2.4),

\[ (\frac{1}{4}) E(x_{1}^{\delta}) \leq \sigma^2(\tau) \leq 2E(x_{1}^{\delta}) \]

we conclude $\frac{3}{4} E(x_{1}^{\delta}) < c < (\frac{3}{4}) E(x_{1}^{\delta})$.

We continue to suppress the superscripts in $\tau_{j}^{(n)}$, and write

(3.1) \[ P(x_n(1) - w(1) \leq \lambda) = P(w(\sum_{i=1}^{n} \tau_j) - w(1) \leq \lambda, \sum_{i=1}^{n} \tau_j \leq 1) \]

\[ + P(w(\sum_{i=1}^{n} \tau_j) - w(1) \leq \lambda, \sum_{i=1}^{n} \tau_j > 1) \]

and handle the (easier) first term first. By construction, $\sum_{i=1}^{n} \tau_j$ is a Markov time, i.e. does not anticipate the future. Hence, by one of the forms of the strong Markov property, and properties of Brownian motion,

(3.2) \[ P(w(\sum_{i=1}^{n} \tau_j) - w(1) \leq \lambda, \sum_{i=1}^{n} \tau_j \leq 1) \]

\[ = \int_{0}^{\lambda} P(w(s) - w(1) \leq \lambda) P(\sum_{i=1}^{n} \tau_j \in s + ds) \]

\[ = \int_{0}^{\lambda} P(w(1 - s) \leq \lambda) P(\sum_{i=1}^{n} \tau_j \in s + ds) \]

\[ = (2\pi)^{-1} \frac{1}{4} \int_{-\infty}^{\lambda} \int_{-\infty}^{\lambda} e^{-u^2} du \, P(\sum_{i=1}^{n} \tau_j \in s + ds). \]

Assume $\lambda < 0$ for definiteness. Viewing the above as a double integral and interchanging the order of integration, we obtain
(3.3)

\[
(2\pi)^{-\frac{3}{2}} \int_{0}^{\infty} P\left( \sum_{i}^{\infty} \tau_{j} \leq 1 - \frac{2^{2}}{u^{2}} \right) e^{-iu^2} du
\]

by the Central Limit Theorem and a second interchanging of order of integration, plus an error term which is small uniformly in \( \lambda \). Note that (3.3) is half of the right-hand side of (1.6). The same expression is also obtained when \( \lambda > 0 \).

For the second term in (3.1), assume \( \sum_{i}^{\infty} \tau_{j} > 1 \) and \( \sum_{i}^{\infty} \tau_{j} \leq 1 < \sum_{i}^{\infty+1} \tau_{j} \). Then

\[w(\sum_{i}^{\infty} \tau_{j}) - w(1) = \sum_{k=1}^{\infty} (w(\sum_{i}^{\infty+1} \tau_{j}) - w(\sum_{i}^{\infty} \tau_{j})) + w(\sum_{i}^{\infty+1} \tau_{j}) - w(1) .\]

Since the \( \{ \sum_{i}^{\infty} \tau_{j} \} \) are consecutive Markov times by construction, the above is a sum of \( n - k \) independent random variables, all but the last having the same distribution as \( x_{j}/n^{2} \). Let \( \mathcal{B}_{1} \) be the \( \sigma \)-algebra generated by \( \{ w(t) : 0 \leq t \leq 1 \} \), and assume that the variables \( x_{j} \) themselves are independent of \( \{ w(t) \} \). The second term in (3.1) then becomes

\[
P(w(\sum_{i}^{\infty} \tau_{j}) - w(1) \leq \lambda, \sum_{i}^{\infty} \tau_{j} \geq 1)
\]

\[
= \sum_{k=1}^{n-1} E\left( \mathbb{1}_{\{ \sum_{i}^{\infty} \tau_{j} \leq 1, \sum_{i}^{\infty+1} \tau_{j} > 1 \}} P\left( \frac{x_{1} + x_{2} + x_{n-k-1}}{n^{2}} + \frac{y_{k}}{n^{2}} \leq \lambda/\mathcal{B}_{1} \right) \right)
\]

\[
= \sum_{k=1}^{n} E\left( \mathbb{1}_{\{ \sum_{i}^{\infty} \tau_{j} \leq 1, \sum_{i}^{\infty+1} \tau_{j} > 1 \}} P\left( \frac{x_{1} + x_{2} + x_{n-k-1}}{n^{2}} + \frac{y_{k}}{n^{2}} \leq \lambda/\mathcal{B}_{1} \right) \right)
\]

where \( y_{k} = n^{2}[w(\sum_{i}^{\infty+1} \tau_{j}) - w(1)] \). Now by construction, where \( x \equiv x_{k+1} \n\)

\[
\tau_{j}^{(n)}(k+1) = \sup \left\{ t : w(t + \sum_{i}^{\infty} \tau_{j}^{(n)}) - w(\sum_{i}^{\infty} \tau_{j}^{(n)}) \in \left\{ \frac{x}{n}, \frac{G(x)}{n} \right\} \right\}
\]

and \( |y_{k}| \leq |x| + |G(x)| \). Hence by (2.12)

\[
E(y^{2}/\mathcal{B}_{1}) \leq 2E(x^{2}) + 2(Q_{x} + (1/b)E|x^{2}|) \leq C .
\]

Now for all \( \varepsilon > 0 \), where \( [x] \) denotes the greatest integer less than or equal to \( x \),

\[
\sum_{i}^{\infty+1} P(\sum_{i}^{\infty} \tau_{j} \leq 1 < \sum_{i}^{\infty+1} \tau_{j})
\]

\[
= P(\sum_{i}^{\infty} \tau_{j} > 1) - P(\sum_{i}^{\infty+1} \tau_{j} > 1)
\]

\[
= P(\sum_{i}^{\infty} (n\tau_{j} - 1) > 0) - P(\sum_{i}^{\infty+1} (n\tau_{j} - 1) > [\varepsilon n^{2}])
\]

\[
\to (2\pi)^{-\frac{3}{2}} \int_{0}^{\infty} \mathbb{1}_{\left\{ x \right\}} e^{-iu^2} du
\]

and similarly
\[ \sum_{k=1}^{n} P\left( \sum_{i=1}^{n-k} \tau_j \leq 1 < \sum_{i=1}^{n-k+1} \tau_j \right) = P\left( \sum_{i=1}^{n-Mn^\frac{1}{2}} \tau_j > 1 \right) \\
= P\left( \sum_{i=1}^{n-Mn^\frac{1}{2}} (n\pi_j - 1) > [Mn^\frac{1}{2}] \right) \\
\to (2\pi)^{-\frac{1}{4}} \int_{\sqrt{n} \lambda / 2 \epsilon}^{\infty} e^{-\frac{1}{2}u^2} \, du. \]

Hence the second sum in (3.4), within an error which is small for large \( n \) uniformly in \( n \) and \( \lambda \), is over the range
\[ \epsilon n^\frac{1}{2} \leq k \leq Mn^\frac{1}{2}. \]

For these \( k \)
\[ P(s_k/n^\frac{1}{2} + y/n^\frac{1}{2} \leq \lambda/\mathcal{B}_1) \leq P(s_k/n^\frac{1}{2} \leq \lambda + \epsilon) + P(|y| > \epsilon n^\frac{1}{2}/\mathcal{B}_1) \]
\[ \geq P(s_k/n^\frac{1}{2} \leq \lambda - \epsilon) - P(|y| > \epsilon n^\frac{1}{2}/\mathcal{B}_1) \]
and by (3.5) and the Central Limit Theorem
\[ |P(s_k/n^\frac{1}{2} + y/n^\frac{1}{2} \leq \lambda/\mathcal{B}_1) - P(s_k/n^\frac{1}{2} \leq \lambda)| \]
\[ \leq P(2\epsilon n^\frac{1}{2} - \epsilon n^\frac{1}{2} \leq s_k/n^\frac{1}{2} \leq 2\epsilon n^\frac{1}{2} + \epsilon n^\frac{1}{2}) + P(|y| > \epsilon n^\frac{1}{2}/\mathcal{B}_1) \]
\[ = O(\epsilon n^\frac{1}{2}) + o(1) + O(1/\epsilon n^2) \]
where \( o(1) \) is uniform in \( \epsilon \) and \( \lambda \). Setting \( \epsilon = 1/n^\frac{1}{2} \), and ignoring errors which are small as \( n \to \infty \) uniformly in \( \lambda \), the expression in (3.4) becomes
\[ \sum_{k=1}^{n} P\left( \sum_{i=1}^{n-k} \tau_j \leq 1 < \sum_{i=1}^{n-k+1} \tau_j \right) P\left( \frac{x_1 + \cdots + x_k}{n^\frac{1}{2}} \leq \lambda \right) \]
\[ = \sum_{k=1}^{n} P\left( \sum_{i=1}^{n-k} \tau_j \leq 1 < \sum_{i=1}^{n-k+1} \tau_j \right) \Phi\left( \frac{n^\frac{1}{2}}{k} \right) \]
\[ = \sum_{k=1}^{n} \left[ P\left( \sum_{i=1}^{n-k+1} \tau_j > 1 \right) - P\left( \sum_{i=1}^{n-k} \tau_j > 1 \right) \right] \Phi\left( \frac{n^\frac{1}{2}}{k} \right) \]
\[ = \sum_{k=1}^{n} \left[ P\left( \sum_{i=1}^{n-k+1} (n\pi_j - 1) > k + 1 \right) - P\left( \sum_{i=1}^{n-k} (n\pi_j - 1) > k \right) \right] \Phi\left( \frac{n^\frac{1}{2}}{k} \right) \]
\[ = \sum_{k=1}^{n} \left[ \Phi\left( \frac{k + 1}{c^2(n - k + 1)^{\frac{1}{2}}} \right) - \Phi\left( \frac{k}{c^2n^{\frac{1}{2}}} \right) \right] \Phi\left( \frac{n^\frac{1}{2}}{k} \right) \]
\[ = \sum_{k=1}^{n} \left[ \Phi\left( \frac{k + 1}{c^2n^{\frac{1}{2}}} \right) - \Phi\left( \frac{k}{c^2n^{\frac{1}{2}}} \right) \right] \Phi\left( \frac{n^\frac{1}{2}}{k} \right) \]
\[ = \frac{1}{2\pi c^2} \int_{\frac{n}{M(n+1)}}^{\infty} e^{\frac{1}{2}v^2} \, dv \]
where \( \Phi(\lambda) \) is the standard normal distribution function, and \( \Phi^c(\lambda) = 1 - \Phi(\lambda) \). Setting \( \lambda = \mu c/n^\frac{1}{2} \) and letting \( n \to \infty \), we obtain
\[ \frac{1}{2\pi c^2} \int_{0}^{\mu} e^{-\frac{1}{2}v^2} \, dv \int_{-\infty}^{\frac{\mu - n^\frac{1}{2}}{2}} e^{-\frac{1}{2}u^2} \, du \]
\[ = \frac{1}{2\pi} \int_{0}^{\infty} e^{-\frac{1}{2}v^2} \int_{-\infty}^{\frac{\mu - n^\frac{1}{2}}{2}} e^{-\frac{1}{2}u^2} \, du \, dv \]
which is the other half of (1.6).

4. An auxiliary result.

**Theorem 3.** Let \( F = \int_0^t \gamma(t)w(t)^2\, dt \) for \( \gamma(t) \in L^1(0, 1) \) and Brownian motion \( w(t) \). Then, the probability distribution \( P(F \leq \lambda) \) has a bounded continuous density (i.e., derivative).

**Proof.** Let \( N_k \) be a sequence of independent standard normal variables, and let \( \{b_k(u): 1 \leq k < \infty\} \) be a complete orthonormal system in \( L^2(0, 1) \). A Brownian motion can then be defined by

\[
\tag{4.1}
w(t) = \sum_{i=1}^{\infty} N_k \int_0^t b_k(u)\, du .
\]

This series converges almost surely for each \( t \), since, by Parseval

\[
\sum_{i=1}^{\infty} (\int_0^t b_k(u)\, du)^2 = \int_0^t \chi_{(0,t)}(u)^2\, du = t < \infty .
\]

Hence \( w(t) \) is a Gaussian process with zero mean. Another application of Parseval’s identity gives

\[
E(w(t)w(s)) = \sum_{i=1}^{\infty} (\int_0^t b_k(u)\, du)(\int_0^s b_k(v)\, dv) = \min\{s, t\}
\]

and \( \{w(t): 0 \leq t \leq 1\} \) is Brownian motion. Consequently

\[
\tag{4.2}
F \cong \int_0^t \gamma(t)w(t)^2\, dt = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} N_k N_j \int_0^t b_k(u)\, du \int_0^s b_j(v)\, dv \, dt = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} N_k N_j \int_0^t b_k(u) b_j(v) \int_{t \wedge s}^{t \vee s} \gamma(t)\, dt \, du \, dv ,
\]

where the interchanging of summation and integration can be justified by the fact that the series (4.1) converges uniformly a.s. (See Walsh (1967).) Now, let \( \{b_k(u): 1 \leq k < \infty\} \) be the complete orthonormal system in \( L^2(0, 1) \) determined by the Fredholm equation

\[
\tag{4.3}
\int_0^t R(u, v)b(v)\, dv = \lambda_k b(u), \quad \int_0^t b(u)^2\, du = 1 ,
\]

\[
R(u, v) = \int_{t \wedge s}^{t \vee s} \gamma(t)\, dt .
\]

Then (4.2) reduces to

\[
\tag{4.4}
F \cong \sum_{i=1}^{\infty} \lambda_k N_k^2 E(e^{itF^p}) = \Pi_{i=1}^{\infty} \frac{1}{(1 + 4s^2 \lambda_k^2)} .
\]

Hence \( E(e^{itF}) = O(1/s^p) \) for all \( p \), and \( P(F \leq \lambda) \) has a density

\[
f(x) = (1/2\pi) \int_{-\infty}^{\infty} \exp(-ix\gamma)(s)\, ds .
\]

If \( \gamma(t) \geq 0 \), the same argument also applies to \( F_i = \int_0^t \gamma(t)(w(t) + a(t))^2\, dt \) for any function \( a(t) \) with \( \int_0^1 \gamma(t)a(t)^2\, dt < \infty \), and thus \( F_i \) also has a bounded density. All of this is a generalization of a classical technique of Kac and Siegert (1947).
REFERENCES