The Annals of Mathematical Statistics
1972, Vol. 43, No. 1 273-284

RATES OF CONVERGENCE FOR SOME FUNCTIONALS IN PROBABILITY[1]

By Stanley Sawyer

Yeshiva University

Let $\{x_1, x_2, \cdots\}$ be a sequence of i.i.d.r.v. with mean zero, variance one, and (1) $P(|x_k| \geqq \lambda) \leqq C \exp(-\alpha\lambda^\varepsilon)$ for positive α, ε. Let $f(t, x)$ (with its first partial derivatives) be of slow growth in x, let $F_n(x)$ be the distribution function of $(1/n) \sum_1^n f(k/n, s_k/n^{\frac{1}{2}})$ where $s_k = x_1 + x_2 + \cdots + x_k$, and let $F(x)$ be the distribution function of $\int_0^1 f(t, w(t))\, dt$ where $\{w(t)\}$ is Brownian motion. Then $\sup_x |F_n(x) - F(x)| = O((\log n)^\beta/n^{\frac{1}{4}})$ provided $F(x)$ 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 $O(1/n^{\frac{1}{4}})$. A corresponding result is obtained if (1) is replaced by the existence of a finite pth moment, $p \geqq 4$.

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

$$(1.1) \qquad P(|x_k| \geqq \lambda) \leqq Ce^{-\alpha\lambda^\varepsilon}$$

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

THEOREM 1. *Let $f(s, x) \in C^1(R^2)$ 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*

$$|Df(s, x)| \leqq \Omega\,(1 + |x|^a)\,,$$

and assume that the probability distribution $P(\int_0^1 f(t, w(t))\, dt \leqq \lambda)$ has a bounded density (i.e., bounded derivative in λ). Then, for $\{x_k\}$ as above

$$(1.2) \quad \sup_\lambda \left| P\left(\frac{1}{n} \sum_1^n f\left(\frac{k}{n}, \frac{s_k}{n^{\frac{1}{2}}}\right) \leqq \lambda\right) - P\left(\int_0^1 f(t, w(t))\, dt \leqq \lambda\right) \right| = O\left(\frac{(\log n)^\beta}{n^{\frac{1}{4}}}\right)$$

where $\beta = \beta(\varepsilon, a)$.

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

$$(1.1)' \qquad E(|x_k|^p) < \infty$$

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

Received January 19, 1971.

[1] These results were obtained while the author was being partially supported by the National Science Foundation under grant NSF GP-21063.

$$O\left(\frac{(\log n)^\beta}{n^{p/(2p+8)}}\right), \qquad \beta = ap/8. \tag{1.3}$$

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

$$\sup_\lambda \left|P\left(\frac{1}{n^2}\sum_1^n s_k^2 \leqq \lambda\right) - P\left(\int_0^1 w(t)^2\,dt \leqq \lambda\right)\right| = O\left(\frac{(\log n)^{10}}{n^{\frac{1}{4}}}\right).$$

If we only assumed $E(x_k^4) < \infty$, the rate would be $O(\log n/n^{\frac{1}{8}})$.

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 $\gamma(t) \in L^1(0, 1)$, $\int_0^1 \gamma(t)w(t)^2\,dt$ 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

$$E\left[g\left(\frac{1}{n}\sum_1^n f\left(\frac{k}{n}, \frac{s_k}{n^{\frac{1}{2}}}\right)\right)\right] = E\left[g\left(\int_0^1 f(t, w(t))\,dt\right)\right] + A/n^{\frac{1}{2}} + O(1/n)$$

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 $\{x_k/n^{\frac{1}{2}}\}$, provides random times $\{\tau_j^{(n)}\}$ such that if

$$\begin{aligned} x_n(k/n) &= w\left(\sum_1^k \tau_j^{(n)}\right), && \text{and} \\ x_n(t) &= x_n(k/n), && k/n \leqq t < (k+1)/n,\ 0 \leqq t \leqq 1, \end{aligned} \tag{1.4}$$

the variables $\{x_n(k/n) - x_n((k-1)/n)\colon 1 \leqq k \leqq n\}$ are independent and have the same distribution as $x_k/n^{\frac{1}{2}}$. Consequently $x_n(k/n) \cong s_k/n^{\frac{1}{2}}$ and

$$\frac{1}{n}\sum_1^n f\left(\frac{k}{n}, \frac{s_k}{n^{\frac{1}{2}}}\right) \cong \frac{1}{n}\sum_1^n f\left(\frac{k}{n}, w\left(\sum_1^k \tau_j^{(n)}\right)\right).$$

Theorem 1 is derived from

PROPOSITION. *For all $b < \infty$ and $\beta = \beta(\varepsilon, a)$ as before*

$$P\left[\left|\frac{1}{n}\sum_1^n f\left(\frac{k}{n}, w\left(\sum_1^k \tau_j^{(n)}\right)\right) - \int_0^1 f(t, w(t))\,dt\right| \geqq \frac{(\log n)^\beta}{n^{\frac{1}{4}}}\right] = O(1/n^b). \tag{1.5}$$

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) $$\lim_{n\to\infty} P\left(x_n(1) - w(1) \leqq \frac{\lambda c}{n^{\frac{1}{4}}}\right) = \frac{1}{\pi} \int_0^\infty e^{-\frac{1}{2}u^2} \int_{-\infty}^{\lambda/u^{\frac{1}{2}}} e^{-\frac{1}{2}v^2}\, dv\, du$$

where c is a positive constant ($c^2 = \sigma(\tau_1^{(1)})$) *satisfying* $0.75\, E(x_k^4)^{\frac{1}{4}} < c < 1.25\, E(x_k^4)^{\frac{1}{4}}$.

As a consequence

(1.7) $$\lim_{n\to\infty} P\left(\max_{0\leqq t\leqq 1} |x_n(t) - w(t)| \geqq \frac{A}{n^{\frac{1}{4}} \log n}\right) = 1$$

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 $O(n^{-\frac{1}{4}})$. 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 $O(n^{-\frac{1}{4}})$. Thus it would appear (i.e. one would conjecture) that all rates of convergence coming from the Invariance Principle for decent $\{x_k\}$ are $O(n^{-\frac{1}{4}})$. (Here "decent" means $E(|x_k|^6) < \infty$.)

In view of Theorem 2, the rate of convergence obtained by Rosenkrantz (1969) for the von Mises statistic cannot be improved beyond $O(n^{-\frac{1}{4}})$, 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

$$P\left(\max_{0\leqq t\leqq 1} |x_n(t) - w(t)| \geqq \frac{(\log n)^\beta}{n^{\frac{1}{4}}}\right) = O(1/n^b)$$

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

$$\sup_\lambda |P(\Phi(x_n(\cdot)) \leqq \lambda) - P(\Phi(w(\cdot)) \leqq \lambda)| = O\left(\frac{(\log n)^\beta}{n^{\frac{1}{4}}}\right)$$

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 $\delta = (\log n)^\beta / n^{\frac{1}{4}}$. Then, by (1.5), for all λ and $b < \infty$

$$(2.1)\quad P\Big(\frac{1}{n}\sum_1^n f\Big(\frac{k}{n}, \frac{s_k}{n^{\frac12}}\Big) \leqq \lambda\Big) \leqq P(\textstyle\int_0^1 f(t, w(t))\,dt \leqq \lambda + \delta) + O(1/n^b)$$

$$\geqq P(\textstyle\int_0^1 f(t, w(t))\,dt \leqq \lambda - \delta) - O(1/n^b)$$

$$\Big|P\Big(\frac{1}{n}\sum_1^n f\Big(\frac{k}{n}, \frac{s_k}{n^{\frac12}}\Big) \leqq \lambda\Big) - P(\textstyle\int_0^1 f(t, w(t))\,dt \leqq \lambda)\Big|$$

$$\leqq P(\lambda - \delta \leqq \textstyle\int_0^1 f(t, w(t))\,dt \leqq \lambda + \delta) + O(1/n^b)\,.$$

By hypothesis, the distribution $P(\int_0^1 f(t, w(t))\,dt \leqq \lambda)$ has a density bounded by some constant, say L. Hence the difference in (2.1) has the uniform bound $2L\delta + O(1/n^b)$.

We now state a lemma.

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

$$(2.2)\qquad E\Big(\Big|\frac{d_1y_1 + d_2y_2 + \cdots + d_ny_n}{n^{\frac12}}\Big|^p\Big) \leqq C_pM^pE(|y_1|^p) \quad 2 \leqq p < \infty.$$

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

$$E(|\textstyle\sum_1^n d_ky_k|^p) \leqq C_pE((\sum_1^n d_k^2y_k^2)^{p/2})$$

for $1 < p < \infty$. Thus

$$E\Big(\Big|\frac{d_1y_1 + d_2y_2 + \cdots + d_ny_n}{n^{\frac12}}\Big|^p\Big) \leqq C_pE\Big(\Big(\frac{d_1^2y_1^2 + \cdots + d_n^2y_n^2}{n}\Big)^{p/2}\Big)$$

$$\leqq C_pM^pE\Big(\Big(\frac{1}{n}\sum_1^n y_k^2\Big)^{p/2}\Big)$$

$$\leqq C_rM^pE\Big(\frac{1}{n}\sum_1^n |y_k|^p\Big)$$

$$= C_pM^pE(|y_1|^p)$$

by the identity $|(1/n)\sum_1^n a_k|^p \leqq (1/n)\sum_1^n |a_k|^p$ for real numbers and $p \geqq 1$.

PROOF OF EQUATION (1.5). For any $A > 1$, define $f^A(s, x) \in C^1(R^2)$ such that

$$(2.3)\qquad f^A(s, x) = f(s, x)\,, \qquad |x| \leqq A,$$
$$|Df^A(s, x)| \leqq 2\Omega(1 + A^a) = M$$

uniformly in s and x, where D is either the identity operator or a first partial derivative. Thus

$$(2.4)\quad P\Big(\Big|\frac{1}{n}\sum_1^n f\Big(\frac{k}{n}, w(\textstyle\sum_1^k \tau_j^{(n)})\Big) - \int_0^1 f(t, w(t))\,dt\Big| \geqq 6\delta\Big)\Big)$$

$$\leqq P\Big(\Big|\frac{1}{n}\sum_1^n f^A\Big(\frac{k}{n}, w(\textstyle\sum_1^k \tau_j^{(n)})\Big) - \int_0^1 f^A(t, w))\,dt\Big| \geqq 6\delta\Big)$$

$$+ P(\max_{0\leqq t\leqq 2}|w(t)| \geqq A) + P(\textstyle\sum_1^n \tau_j > 2)\,,$$

where here and in the following we abbreviate $\tau_j = \tau_j^{(n)}$ and $\tau = \tau_1^{(1)}$. Thus

$$P(\max_{0\leqq t\leqq 1}|w(t)| \geqq A) \leqq 2P(\max_{0\leqq t\leqq 1} w(t) \geqq A) = 2\Big(\frac{2}{\pi}\Big)^{\frac12}\int_A^\infty e^{-\frac12 u^2}\,du$$

$$= 2\Big(\frac{2}{\pi}\Big)^{\frac12}\int_0^\infty e^{-\frac12(u+A)^2}\,du \leqq 2e^{-\frac12 A^2}$$

(see [3] page 392) and

$$P(\max_{0\leqq t\leqq 2}|w(t)| \geqq A) = P(\max_{0\leqq t\leqq 1}|w(t)| \geqq A/2^{\frac12}) \leqq 2e^{-\frac14 A^2}\,.$$

Also, the random times $\{\tau_j^{(n)}\}$ are independent and identically distributed for a fixed $n, \tau_j^{(n)} \cong (1/n)\tau$, and $E(\tau) = E(x_1^2) = 1$ (see [1]). Thus by Lemma 1

$$P(\textstyle\sum_1^n \tau_j > 2) = P(n^{-\frac12}\sum_1^n (n\tau_j - 1) > n^{\frac12})$$

$$\leqq \frac{1}{n^{p/2}} E(|n^{-\frac12}\textstyle\sum_1^n (n\tau_j - 1)|^p)$$

$$\leqq \frac{1}{n^{p/2}} C_p E(|\tau - 1|^p)\,, \qquad p \geqq 2.$$

Now, for all $p \geqq 1$, by Sawyer (1967), Section 2

$$(2.5)\qquad E(\tau^p) \leqq 4p\Gamma(p)E(|x_k|^{2p}) < \infty\,.$$

Finally, if $A = \log n$, the last two terms in (2.4) are both $O(1/n^b)$ for all b.

We apply Taylor's formula to $f^A(s, x)$, and suppress the superscript A in $f^A(s, x)$ from this point on.

$$\int_0^{\sum_1^n \tau_j} f(t, w(t))\,dt$$

$$= \textstyle\sum_0^{n-1} \int_{\zeta_k}^{\zeta_{k+1}} f(t, w(t))\,dt$$

$$(2.6)\quad = \sum_0^{n-1} f\Big(\frac{k}{n}, w(\textstyle\sum_1^k \tau_j)\Big)\tau_{k+1}$$

$$+ \sum_0^{n-1}\int_{\zeta_k}^{\zeta_{k+1}} \frac{\partial f}{\partial s}\Big(\frac{k}{n} + \bar\theta_{kn}(s), w(\textstyle\sum_1^k \tau_j) + \theta_{kn}(s)\Big)\Big(s - \frac{k}{n}\Big)\,ds$$

$$+ \sum_0^{n-1}\int_{\zeta_k}^{\zeta_{k+1}} \frac{\partial f}{\partial x}\Big(\frac{k}{n} + \bar\theta_{kn}(s), w(\textstyle\sum_1^k \tau_j) + \theta_{kn}(s)\Big)$$

$$\times [w(s) - w(\textstyle\sum_1^k \tau_j)]\,ds\,,$$

where $\zeta_k = \sum_1^k \tau_j^{(n)}$. Thus

$$\int_0^1 f(t, w(t))\,dt - \frac{1}{n}\sum_1^n f\Big(\frac{k}{n}, w(\textstyle\sum_1^k \tau_j)\Big)$$

$$(2.7)\qquad = \frac{1}{n}\sum_0^{n-1} f\Big(\frac{k}{n}, w(\textstyle\sum_1^k \tau_j)\Big)(n\tau_{k+1} - 1)$$
$$- \int_1^{\sum_1^n \tau_j} f(t, w(t))\,dt + (1/n)f(0,0) - (1/n)f(1, w(\textstyle\sum_1^n \tau_j))$$
$$+ \Phi_5 + \Phi_6\,,$$

where we let $\Phi_1, \cdots, \Phi_4$ be the first four terms in (2.7) and Φ_5, Φ_6 the last two terms in (2.6). We estimate the difference in (2.7) as follows:

$$P\Big(\Big|\int_0^1 f(t, w(t))\,dt - \frac{1}{n}\sum_1^n f\Big(\frac{k}{n}, w(\textstyle\sum_1^k \tau_j)\Big)\Big| > 6\delta\Big)$$
$$(2.8)\qquad \leqq \sum_1^6 P(|\Phi_k| > \delta)$$
$$= \frac{1}{(n^{\frac12}\delta)^p}\sum_1^6 E(|n^{\frac12}\Phi_k|^p)$$

for all $p > 0$. In the first term we apply Lemma 1 with $d_k = f(k/n, w(\sum_1^k \tau_j))$ and $y_k = n\tau_{k+1} - 1$.

$$(2.9)\qquad E(|n^{\frac12}\Phi_1|^p) \leqq C_p M^p E(|\tau - 1|^p)\,.$$

Now an inspection of the proof of Burkholder's inequality shows that $C_p = O((c_0 p)^{2p})$ as $p \to \infty$, while by (2.5) and (1.1)

$$E(\tau^p) = O(p^p(2p/\varepsilon)^{2p/\varepsilon})\,.$$

Hence there exists a constant $c > 0$ such that

$$E(|n^{\frac12}\Phi_1|^p) \leqq \Omega_0 M^p (cp)^{cp}\,.$$

Now if $p = \gamma \log n$ for $\gamma = 1/c$ and $\delta = (\log n)^\beta/n^{\frac12}$, $n \geqq 3$, then by (2.3)

$$(n^{\frac12}\delta)^{-p} E(|n^{\frac12}\Phi_1|^p) \leqq \Omega_0(4\Omega)^{\gamma\log n}((\log n)^{a\gamma\log n})((\log n)^{\log n})/(\log n)^{\beta\gamma\log n}$$
$$= \Omega_0 n^{\gamma\log 4\Omega}/n^{[\gamma(\beta-a)-1]\log\log n}$$
$$= O(1/n^{\varepsilon\log\log n})$$

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

$$(2.10)\qquad E(|n^{\frac12}\Phi_k|^p) \leqq \Omega^p C_p M^p E(\tau^{mp})^j E(|x|^{lp})$$

can be estimated similarly, perhaps with a larger β.

For the second term in (2.8)

$$|\Phi_2| = \Big|\int_1^{\sum_1^n \tau_j} f(t, w(t))\,dt\Big| \leqq M\,|\textstyle\sum_1^n \tau_j - 1|$$
$$|n^{\frac12}\Phi_2| \leqq (M/n^{\frac12})\,|\textstyle\sum_1^n (n\tau_j - 1)|$$
$$E(|n^{\frac12}\Phi_2|^p) \leqq C_p M^p E(|\tau - 1|^p)$$

which is exactly the same as (2.9). The terms Φ_3 and Φ_4 give no trouble, since after multiplication by $n^{\frac{1}{2}}$ they converge uniformly to zero. Using the inequality $|s - k/n| \leqq |s - \sum_1^k \tau_j| + |\sum_1^k \tau_j - k/n|$ in the integral in Φ_5 and integrating, we obtain

$$|\Phi_5| \leqq M \sum_0^{n-1} (\tfrac{1}{2}\tau_{k+1}^2 + \tau_{k+1} |\textstyle\sum_1^k \tau_j - k/n|)$$

$$|n^{\frac{1}{2}}\Phi_5| \leqq \frac{M}{n^{\frac{1}{2}}}\frac{1}{n} \sum_1^n (n\tau_k)^2 + \frac{M}{n}\sum_1^n (n\tau_k)\left|\frac{1}{n^{\frac{1}{2}}}\sum_1^{k-1}(n\tau_j - 1)\right|$$

$$E(|n^{\frac{1}{2}}\Phi_5|^p) \leqq \frac{(2M)^p}{n^{p/2}} E(\tau^{2p}) + \frac{(2M)^p}{n^{\frac{1}{2}}}\sum_1^n E\left((n\tau_k)^p\left|\frac{1}{n}\sum_1^{k-1}(n\tau_j - 1)\right|^p\right)$$

$$\leqq \frac{(2M)^p}{n^{p/2}} E(\tau^{2p}) + C_p(2M)^p E(\tau^p)E(|\tau - 1|^p)$$

by independence and Lemma 1. This is of the form (2.10); for the sixth term:

$$|n^{\frac{1}{2}}\Phi_6| \leqq \frac{M}{n}\sum_0^{n-1} nn^{\frac{1}{2}} \int_{\xi_k}^{\xi_{k+1}} |w(s) - w(\textstyle\sum_1^k \tau_j)|\, ds\,.$$

The terms in the series above are independent and identically distributed by construction, and

$$\int_{\sum^k \tau_j}^{\sum^{k+1}\tau_j} |w(s) - w(\textstyle\sum_1^k \tau_j)|\, ds \cong \frac{1}{nn^{\frac{1}{2}}}\int_0^\tau |w(s)|\, ds\,.$$

This is because $n\tau_1^{(n)}$ bears the same relation to $n^{\frac{1}{2}}w(t/n)$ as τ does to $w(t)$ (see [1]) and thus

$$\int_0^{\tau_1^{(n)}} |w(s)|\, ds = \frac{1}{n}\int_0^{n\tau_1^{(n)}} \left|w\left(\frac{s}{n}\right)\right| ds$$

$$= \frac{1}{nn^{\frac{1}{2}}}\int_0^{n\tau_1^{(n)}} \left|n^{\frac{1}{2}}w\left(\frac{s}{n}\right)\right| ds \cong \frac{1}{nn^{\frac{1}{2}}}\int_0^\tau |w(s)|\, ds\,.$$

Hence for $p \geqq 1$

$$E(|n^{\frac{1}{2}}\Phi_6|^p) \leqq M^p E((\textstyle\int_0^\tau |w(s)|\, ds)^p)\,. \tag{2.11}$$

Now $\tau = \inf\{t;\, w(t) \in \{x, G(x)\}\}$, where $x \cong x_k$, x is independent of $\{w(t)\}$, and $G(y)$ is a certain function ([1], [14]). Thus

$$\int_0^\tau |w(s)|\, ds \leqq \tau(|x| + |G(x)|)$$

$$E((\textstyle\int_0^\tau |w(s)|\, ds)^p) \leqq E[\tau^{2p}]^{\frac{1}{2}}[E((|x| + |G(x)|)^{2p})]^{\frac{1}{2}}$$

Now if b is sufficiently small so that $|G(\pm b)| < \infty$, $b > 0$,

$$\begin{aligned} E(|G(x)|^p) &= \textstyle\int |G(y)|^p P(x \in dy) \\ &\leqq |G(b)|^p + |G(-b)|^p + (1/b)\textstyle\int |G(y)|^p\, |y|\, P(x \in dy) \\ &\leqq Q_b{}^p + (1/b)\, E(|x|^{p+1}) \end{aligned} \tag{2.12}$$

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. $\frac{1}{n}\sum_1^n f\left(\frac{k}{n}, w(\sum_1^k \tau_j)\right) = \int_0^1 f(t, w(t))\,dt + O\left(\frac{(\log n)^\beta}{n^{\frac{1}{4}}}\right)$ a.s.

PROOF. Use the Borel-Cantelli lemma in (1.5).

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

$$E\left[g\left(\frac{1}{n}\sum_1^n f\left(\frac{k}{n}, \frac{s_k}{n^{\frac{1}{2}}}\right)\right)\right] = E[g(\textstyle\int_0^1 f(t, w(t))\,dt)] + O\left(\frac{(\log n)^\beta}{n^{\frac{1}{4}}}\right).$$

PROOF. For any random variable Y,

$$E(g(Y)) = \int g(y)P(Y \in dy) = -\int g'(y)P(Y \leqq y)\,dy + g(\infty).$$

Now use (1.2).

REMARK. To derive (1.3) assuming only that $E(|x_k|^q) < \infty$, $q \geqq 4$, 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

$$O\left(\frac{(\log n)^\beta}{n^{(p-1)/(2p+4)}}\right). \tag{1.3$'$}$$

3. **Proof of Theorem** 2. By properties of the Skorokhod representation (see [1] page 276 +), $\tau_j^{(n)} \cong n^{-1}\tau$, where $\tau = \tau_1^{(1)}$ and $E(\tau) = E(x_1^2) = 1$. Define a constant $c > 0$ by $c^4 = \sigma^2(\tau) = E((\tau - 1)^2)$. Since by Sawyer (1967), (2.4),

$$(\tfrac{1}{3})E(x_1^4) \leqq \sigma^2(\tau) \leqq 2E(x_1^4)$$

we conclude $(\frac{3}{4})E(x_1^4)^{\frac{1}{4}} < c < (\frac{5}{4})E(x_1^4)^{\frac{1}{4}}$.

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

$$\begin{aligned} P(x_n(1) - w(1) \leqq \lambda) &= P(w(\textstyle\sum_1^n \tau_j) - w(1) \leqq \lambda, \sum_1^n \tau_j \leqq 1) \\ &\quad + P(w(\textstyle\sum_1^n \tau_j) - w(1) \leqq \lambda, \sum_1^n \tau_j > 1) \end{aligned} \tag{3.1}$$

and handle the (easier) first term first. By construction, $\sum_1^n \tau_j$ 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,

$$\begin{aligned} &P(w(\textstyle\sum_1^n \tau_j) - w(1) \leqq \lambda, \sum_1^n \tau_j \leqq 1) \\ &\quad = \textstyle\int_0^1 P(w(s) - w(1) \leqq \lambda)P(\sum_1^n \tau_j \in s + ds) \\ &\quad = \textstyle\int_0^1 P(w(1-s) \leqq \lambda)P(\sum_1^n \tau_j \in s + ds) \\ &\quad = (2\pi)^{-\frac{1}{2}} \textstyle\int_0^1 \int_{-\infty}^{\lambda(1-s)^{-\frac{1}{2}}} e^{-\frac{1}{2}u^2}\,du\, P(\sum_1^n \tau_j \in s + ds). \end{aligned} \tag{3.2}$$

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

$$
\begin{aligned}
&(2\pi)^{-\frac{1}{2}} \int_0^\infty P\Big(\textstyle\sum_1^n \tau_j \leqq 1 - \dfrac{\lambda^2}{u^2}\Big) e^{-\frac{1}{2}u^2}\, du \\
&\quad = (2\pi)^{-\frac{1}{2}} \int_0^\infty P\Big(n^{-\frac{1}{2}} \textstyle\sum_1^n (n\tau_j - 1) \leqq -\dfrac{\lambda^2 n^{\frac{1}{2}}}{u^2}\Big) e^{-\frac{1}{2}u^2}\, du \qquad (3.3)\\
&\quad = (2\pi)^{-\frac{1}{2}} \int_0^\infty \int_{-\infty}^{-(\lambda^2 n^{\frac{1}{2}}/c^2u^2)} e^{-\frac{1}{2}v^2}\, dv\, e^{-\frac{1}{2}u^2}\, du \\
&\quad = (2\pi)^{-\frac{1}{2}} \int_0^\infty e^{-\frac{1}{2}v^2} \int_{-\infty}^{\lambda n^{\frac{1}{4}}/cv^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\, du\, dv
\end{aligned}
$$

by the Central Limit Theorem and a second interchanging of order of integration, plus an error term which is small uniformly in λ. 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 $\sum_1^n \tau_j > 1$ and $\sum_1^k \tau_j \leqq 1 < \sum_1^{k+1} \tau_j$. Then

$$
w(\textstyle\sum_1^n \tau_j) - w(1) = \sum_{k+1}^{n-1} (w(\sum_1^{l+1} \tau_j) - w(\sum_1^l \tau_j)) + w(\sum_1^{k+1} \tau_j) - w(1)\,.
$$

Since the $\{\sum_1^l \tau_j\}$ 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 $x_k/n^{\frac{1}{2}}$. Let $\mathscr{B}_1$ be the σ-algebra generated by $\{w(t)\colon 0 \leqq t \leqq 1\}$, and assume that the variables x_k themselves are independent of $\{w(t)\}$. The second term in (3.1) then becomes

$$
\begin{aligned}
&P(w(\textstyle\sum_1^n \tau_j) - w(1) \leqq \lambda,\ \sum_1^n \tau_j \geqq 1) \\
&\quad = \textstyle\sum_0^{n-1} E\Big(\chi_{[\sum_1^k \tau_j \leqq 1 < \sum_1^{k+1} \tau_j]}\, P\Big(\dfrac{x_1 + x_2 + x_{n-k-1}}{n^{\frac{1}{2}}} + \dfrac{y_k}{n^{\frac{1}{2}}} \leqq \lambda/\mathscr{B}_1\Big)\Big) \qquad (3.4)\\
&\quad = \textstyle\sum_1^n E\Big(\chi_{[\sum_1^{n-k} \tau_j \leqq 1 < \sum_1^{n-k+1} \tau_j]}\, P\Big(\dfrac{x_1 + x_2 + x_{k-1}}{n^{\frac{1}{2}}} + \dfrac{y_{n-k}}{n^{\frac{1}{2}}} \leqq \lambda/\mathscr{B}_1\Big)\Big)
\end{aligned}
$$

where $y_k = n^{\frac{1}{2}}[w(\sum_1^{k+1} \tau_j) - w(1)]$. Now by construction, where $x \cong x_{k+1}$

$$
\tau_{k+1}^{(n)} = \sup \Big\{t\colon w(t + \textstyle\sum_1^k \tau_j^{(n)}) - w(\sum_1^k \tau_j^{(n)}) \in \Big\{\dfrac{x}{n^{\frac{1}{2}}}, \dfrac{G(x)}{n^{\frac{1}{2}}}\Big\}\Big\}
$$

and $|y_k| \leqq |x| + |G(x)|$. Hence by (2.12)

$$
E(y^2/\mathscr{B}_1) \leqq 2E(x^2) + 2(Q_b^{\,2} + (1/b)E(|x|^3)) \leqq C\,. \qquad (3.5)
$$

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

$$
\begin{aligned}
&\textstyle\sum_1^{[\varepsilon n^{\frac{1}{2}}]} P(\sum_1^{n-k} \tau_j \leqq 1 < \sum_1^{n-k+1} \tau_j) \\
&\quad = P(\textstyle\sum_1^{n-[\varepsilon n^{\frac{1}{2}}]} \tau_j \leqq 1 < \sum_1^n \tau_j) \\
&\quad = P(\textstyle\sum_1^n \tau_j > 1) - P(\sum_1^{n-[\varepsilon n^{\frac{1}{2}}]} \tau_j > 1) \\
&\quad = P(\textstyle\sum_1^n (n\tau_j - 1) > 0) - P(\sum_1^{n-[\varepsilon n^{\frac{1}{2}}]} (n\tau_j - 1) > [\varepsilon n^{\frac{1}{2}}]) \\
&\quad \to (2\pi)^{-\frac{1}{2}} \int_0^{\varepsilon/c^2} e^{-\frac{1}{2}u^2}\, du
\end{aligned}
$$

and similarly

$$\begin{aligned}\textstyle\sum_{[Mn^{\frac12}]}^{n} P(\sum_1^{n-k}\tau_j \leqq 1 < \sum_1^{n-k+1}\tau_j) &= P(\textstyle\sum_1^{n-[Mn^{\frac12}]}\tau_j > 1)\\ &= P(\textstyle\sum_1^{n-[Mn^{\frac12}]}(n\tau_j - 1) > [Mn^{\frac12}])\\ &\to (2\pi)^{-\frac12}\textstyle\int_{M/c^2}^{\infty} e^{-\frac12 u^2 du}\,.\end{aligned}$$

Hence the second sum in (3.4), within an error which is small for large n uniformly in n and λ, is over the range

$$\varepsilon n^{\frac12} \leqq k \leqq Mn^{\frac12}\,.$$

For these k

$$\begin{aligned}P(s_k/n^{\frac12} + y/n^{\frac12} \leqq \lambda/\mathscr{B}_1) &\leqq P(s_k/n^{\frac12} \leqq \lambda + \varepsilon) + P(|y| > \varepsilon n^{\frac12}/\mathscr{B}_1)\\ &\geqq P(s_k/n^{\frac12} \leqq \lambda - \varepsilon) - P(|y| > \varepsilon n^{\frac12}/\mathscr{B}_1)\end{aligned}$$

and by (3.5) and the Central Limit Theorem

$$\begin{aligned}&|P(s_k/n^{\frac12} + y/n^{\frac12} \leqq \lambda/\mathscr{B}_1) - P(s_k/n^{\frac12} \leqq \lambda)|\\ &\quad\leqq P(\lambda n^{\frac14} - \varepsilon n^{\frac14} \leqq s_k/n^{\frac14} \leqq \lambda n^{\frac14} + \varepsilon n^{\frac14}) + P(|y| > \varepsilon n^{\frac12}/\mathscr{B}_1)\\ &\quad= O(\varepsilon n^{\frac14}) + \sigma(1) + O(1/n\varepsilon^2)\end{aligned}$$

where $\sigma(1)$ is uniform in ε and λ. Setting $\varepsilon = 1/n^{3/8}$, and ignoring errors which are small as $n \to \infty$ uniformly in λ, the expression in (3.4) becomes

$$\begin{aligned}&\textstyle\sum_1^n P(\sum_1^{n-k}\tau_j \leqq 1 < \sum_1^{n-k+1}\tau_j)P\left(\dfrac{x_1 + \cdots + x_k}{n^{\frac12}} \leqq \lambda\right)\\ &= \textstyle\sum_1^n P(\sum_1^{n-k}\tau_j \leqq 1 < \sum_1^{n-k+1}\tau_j)\Phi\left(\lambda\left(\dfrac{n}{k}\right)^{\frac12}\right)\\ &= \textstyle\sum_1^n [P(\sum_1^{n-k+1}\tau_j > 1) - P(\sum_1^{n-k}\tau_j > 1)]\Phi\left(\lambda\left(\dfrac{n}{k}\right)^{\frac12}\right)\\ &= \textstyle\sum_1^n [P(\sum_1^{n-k+1}(n\tau_j - 1) > k + 1) - P(\sum_1^{n-k}(n\tau_j - 1) > k)]\Phi\left(\lambda\left(\dfrac{n}{k}\right)^{\frac12}\right)\\ &= \textstyle\sum_1^n \left(\Phi^c\left(\dfrac{k+1}{c^2(n-k+1)^{\frac12}}\right) - \Phi^c\left(\dfrac{k}{c^2(n-k)^{\frac12}}\right)\right)\Phi\left(\lambda\left(\dfrac{n}{k}\right)^{\frac12}\right)\\ &= \textstyle\sum_2^n \left(\Phi\left(\dfrac{k+1}{c^2 n^{\frac12}}\right) - \Phi\left(\dfrac{k}{c^2 n^{\frac12}}\right)\right)\Phi\left(\lambda\left(\dfrac{n}{k}\right)^{\frac12}\right)\\ &= \dfrac{1}{2\pi c^2 n^{\frac12}}\textstyle\sum_1^{M(n)^{\frac12}} e^{k^2/2c^4 n}\int_{-\infty}^{\lambda(n/k)^{\frac12}} e^{-\frac12 u^2}\,du\end{aligned}$$

where $\Phi(\lambda)$ is the standard normal distribution function, and $\Phi^c(\lambda) = 1 - \Phi(\lambda)$. Setting $\lambda = \mu c/n^{\frac14}$ and letting $n \to \infty$, we obtain

$$\begin{aligned}&\frac{1}{2\pi c^2}\int_0^M e^{-\frac12 v^2/c^4}\int_{-\infty}^{\mu c v^{-\frac12}} e^{-\frac12 u^2}\,du\,dv\\ &\qquad= \frac{1}{2\pi}\int_0^\infty e^{-\frac12 u^2}\int_{-\infty}^{\mu v^{-\frac12}} e^{-\frac12 v^2}\,dv\,du\end{aligned}$$

which is the other half of (1.6).

4. An auxiliary result.

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

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

$$w(t) = \sum_1^\infty N_k \int_0^t b_k(u)\,du\,. \tag{4.1}$$

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

$$\sum_1^\infty \left(\int_0^t b_k(u)\,du\right)^2 = \int_0^1 \chi_{(0,t)}(u)^2\,du = t < \infty\,.$$

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

$$E(w(t)w(s)) = \sum_1^\infty \left(\int_0^t b_k(u)\,du\right)\left(\int_0^s b_k(v)\,dv\right) = \min\{s, t\}$$

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

$$\begin{aligned} F \cong \int_0^1 \gamma(t)w(t)^2\,dt &= \sum_1^\infty \sum_1^\infty N_k N_j \int_0^1 \gamma(t) \int_0^t b_k(u)\,du \int_0^t b_j(v)\,dv\,dt \\ &= \sum_1^\infty \sum_1^\infty N_k N_j \int_0^1 \int_0^1 b_k(u)b_j(v) \int_{\max\{u,v\}}^1 \gamma(t)\,dt\,du\,dv\,, \end{aligned} \tag{4.2}$$

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 $\{b_k(u): 1 \leqq k < \infty\}$ be the complete orthonormal system in $L^2(0, 1)$ determined by the Fredholm equation

$$\begin{gathered} \int_0^1 R(u, v)b(v)\,dv = \lambda_k b(u)\,, \qquad \int_0^1 b(u)^2\,du = 1\,, \\ R(u, v) = \int_{\max\{u,v\}}^1 \gamma(t)\,dt\,. \end{gathered} \tag{4.3}$$

Then (4.2) reduces to

$$\begin{gathered} F \cong \sum_1^\infty \lambda_k N_k^2 \\ E(e^{sF}) = \prod_1^\infty E(e^{s\lambda_k N_k^2}) = \prod_1^\infty (1 - 2s\lambda_k)^{-\frac{1}{2}} \\ |E(e^{isF})| = \left(\prod_1^\infty \frac{1}{(1 + 4s^2\lambda_k^2)}\right)^{\frac{1}{4}}\,. \end{gathered}$$

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

$$f(x) = (1/2\pi) \int_{-\infty}^{+\infty} \exp(-ixs)g(s)\,ds\,.$$

If $\gamma(t) \geqq 0$, the same argument also applies to $F_1 = \int_0^1 \gamma(t)(w(t) + a(t))^2\,dt$ for any function $a(t)$ with $\int_0^1 \gamma(t)a(t)^2\,dt < \infty$, and thus F_1 also has a bounded density. All of this is a generalization of a classical technique of Kac and Siegert (1947).

REFERENCES

[1] BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading. Chapter 13.

[2] BURKHOLDER, D.L. (1966). Martingale transforms. *Ann. Math. Statist.* **37** 1494–1504.

[3] DOOB, J.L. (1953). *Stochastic Processes*. Wiley, New York.

[4] FRASER, DAVID F. (1971). Rates of convergence for a probabilistic solution to the Dirichlet problem. Ph. D. Dissertation, Brown University.

[5] KAC, M. and SIEGERT, A.J. (1947). An explicit representation for stationary Gaussian processes. *Ann. Math. Statist.* **18** 434–442.

[6] KIEFER, J. (1969). On the deviations in the Skorokhod-Strassen approximation scheme. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* **13** 321–332.

[7] NAGAEV, S. (1969). An estimate of the rate of convergence of the distribution of the maximum of sums of independent random variables. *Siberian Math. J.* **10** 443–458. (English translation.)

[8] NEVZOROV, V.B. and PETROV, V.V. (1969). On the distribution of the maximum cumulative sum of independent random variables. *Theŏr. Probability Appl.* **14** 682–687. (English translation.)

[9] ROSENKRANTZ, W. (1967). On rates of convergence for the invariance principle. *Trans. Amer. Math. Soc.* **129** 542–552.

[10] ROSENKRANTZ, W. (1969). A rate of convergence for the von Mises statistic. *Trans. Amer. Math. Soc.* **139** 329–337.

[11] SAWYER, S. (1967). A uniform rate of convergence for the maximum absolute value of partial sums in probability. *Comm. Pure Appl. Math.* **20** 647–658.

[12] SAWYER, S. (1968). Uniform limit theorems for the maximum cumulative sum in probability. *Trans. Amer. Math. Soc.* **132** 363–367.

[13] SAZONOV, V.V. (1969). An improvement of a convergence-rate estimate. *Theor. Probability Appl.* **14** 640–651. (English translation.)

[14] SKOROKHOD, A. (1965). *Studies in the Theory of Random Processes*. Addison-Wesley, Reading. Chapter 7. (English translation.)

[15] WALSH, J. (1967). A note on the uniform convergence of stochastic processes. *Proc. Amer. Math. Soc.* **18** 129–132.