CONVERGENCE OF SEQUENCES OF REGULAR FUNCTIONALS OF EMPIRICAL DISTRIBUTIONS TO PROCESSES OF BROWNIAN MOTION

By

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 726

January 1971

CONVERGENCE OF SEQUENCES OF REGULAR FUNCTIONALS OF EMPIRICAL DISTRIBUTIONS TO PROCESSES OF BROWNIAN MOTION[1]

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

For partial cumulative sums of independent and identically distributed random variables (iidrv) having a positive (finite) variance, weak convergence to Brownian motion processes has been established by Donsker (1951, 1952). The result is extended here to von Mises' (1947) differentiable statistical functions and Hoeffding's (1948) U-statistics. Few applications are sketched.

1. Introduction. Let $\omega = \{X_1, X_2, \ldots\}$ be a sequence of iidrv with each X_i having a distribution function (d.f.) $F(x)$, $x \varepsilon R^p$, the $p(\geq 1)$-dimensional Euclidean space. Let $g(X_1, \ldots, X_m)$, symmetric in its arguments, be a Borel-measurable kernel of degree $m(\geq 1)$, and consider the regular functional

$$\theta(F) = \int \cdots \int_{R^{pm}} g(x_1, \ldots, x_m) dF(x_1) \cdots dF(x_m), \tag{1.1}$$

defined over $F \varepsilon \mathcal{F} = \{F: \theta(F) < \infty\}$. For a sample $(X_1, \ldots, X_n)$, consider the empirical d.f.

$$F_n(x, \omega) = n^{-1} \sum_{i=1}^{n} c(x - X_i), \quad x \varepsilon R^p, \tag{1.2}$$

[1]Work supported by the Army Research Center, Durham.

where for $u = (u_1,\ldots,u_p)$, $c(u) = 1$ if $u_i \geq 0$, $1 \leq i \leq p$, and $c(u) = 0$, otherwise. Then, the corresponding regular functional of $F_n(x,\omega)$ is defined by

$$(1.3) \qquad \theta(F_n,\omega) = \int\cdots\int_{R^{pm}} g(x_1,\ldots,x_m)dF_n(x_1,\omega)\cdots dF_n(x_m,\omega).$$

Under certain regularity conditions, von Mises (1947) derived the asymptotic normality of $n^{\frac{1}{2}}[\theta(F_n,\omega)-\theta(F)]$; a comparatively simpler proof was subsequently devised by Hoeffding (1948), who primarily considered the unbiased estimators (U-statistics)

$$(1.4) \qquad U_n(\omega) = n^{-[m]} \sum_{P_{n,m}} g(X_{i_1},\ldots,X_{i_m}), \quad n \geq m,$$

(where $P_{n,m} = \{1 \leq i_1 \neq \cdots \neq i_m \leq n\}$ and $n^{-[m]} = \{n\cdots(n-m+1)\}^{-1}$), and showed that under essentially the same conditions, $n^{\frac{1}{2}}[U_n(\omega)-\theta(F)]$ is asymptotically normal and $n^{\frac{1}{2}}[\theta(F_n,\omega)-U_n(\omega)] \overset{P}{\to} 0$, as $n\to\infty$.

In particular when m=1, $\theta(F_n,\omega) = U_n(\omega) = n^{-1}\sum_{i=1}^n g(X_i) = \bar{g}_n(\omega)$, say. If we assume that $E\{[g(X)-\theta(F)]^2\} = \sigma^2$ is positive (finite), and let

$$(1.5) \qquad Y_n(t,\omega) = \{[nt]\bar{g}_{[nt]}(\omega) + (nt-[nt])g(X_{[nt]+1})-nt\,\theta(F)\}/\{\sigma n^{\frac{1}{2}}\},$$

where [s] denotes the integral part of $s(\geq 0)$ and $t \in I = \{t: 0 \leq t \leq 1\}$, then by the Donsker theorem [cf. Billingsley (1968, p. 68)], as $n\to\infty$

$$(1.6) \qquad Y_n(\omega) = [Y_n(t,\omega), t \in I] \overset{\mathcal{D}}{\to} W = [W_t, t \in I],$$

where $W_t, t\geq 0$, is a standard Brownian motion, and $\overset{\mathcal{D}}{\to}$ stands for convergence in distribution.

In the present paper, the above result is extended to general $\theta(F_n,\omega)$ and $U_n(\omega)$, and some possible applications are briefly sketched. Throughout the paper, we assume that $m\geq 2$.

2. The main theorem. For every $h(0\leq h\leq m)$, let

$$g_h(x_1,\ldots,x_h) = E\{g(x_1,\ldots,x_h,X_{h+1},\ldots,X_m)\},\ g_0 = \theta(F); \tag{2.1}$$

$$\zeta_h(F) = E\{[g_h(X_1,\ldots,X_h)-\theta(F)]^2\}(\leq\zeta_m(F)),\ \zeta_0(F) = 0. \tag{2.2}$$

Our basic assumptions are (I,IIa) or (I,IIb), where (I) $\theta(F)$ is stationary of order 0 i.e., $\zeta_1(F)>0$, (IIa) $\zeta_m(F)<\infty$, and (IIb)

$$\max_{1\leq i_1\leq\cdots\leq i_m\leq m}\{g^2(X_{i_1},\ldots,X_{i_m})\} = \zeta^*(F)<\infty. \tag{2.3}$$

[Note that (IIb) $\Rightarrow$ (IIa).] Let then, for every $n\geq 1$,

$$\begin{aligned} Y_n(t,\omega) &= \{m[n\zeta_1(F)]^{\frac{1}{2}}\}^{-1}\{([nt]+1)(nt-[nt])\theta(F_{[nt]+1},\omega) \\ &\quad + [nt]([nt]+1-nt)\theta(F_{[nt]},\omega)-nt\theta(F)\},\ t\ \varepsilon\ I. \end{aligned} \tag{2.4}$$

Also, we define $Y_n^*(t,\omega)$ as in (2.4), with $\theta(F_k,\omega)$ being replaced by $U_k(\omega)$ for $m\leq k\leq n$, and $\theta(F_k,\omega)$ by 0 for $k<m$. Let then

$$Y_n(\omega) = [Y_n(t,\omega),t\ \varepsilon\ I],\ Y_n^*(\omega) = [Y_n^*(t,\omega),\ t\ \varepsilon\ I]; \tag{2.5}$$

$$\rho(Y_n(\omega),Y_n^*(\omega)) = \sup_{t\varepsilon I}|Y_n(t,\omega)-Y_n^*(t,\omega)|. \tag{2.6}$$

Then, we have the following theorem:

Theorem 2.1. Under (I) and (IIa), as $n\to\infty$,

$$Y_n^*(\omega) \overset{\mathcal{D}}{\to} W, \tag{2.7}$$

while under (I) and (IIb), as $n\to\infty$,

$$Y_n(\omega) \overset{\mathcal{D}}{\to} W, \text{ and } \rho(Y_n(\omega), Y_n^*(\omega)) \overset{P}{\to} 0. \tag{2.8}$$

The proof of the theorem is postponed to section 4.

3. Some results on $\theta(F_n,\omega)$ and $U_n(\omega)$. For every $h(1\le h\le m)$, define

$$V_{n,h}(\omega) = \int_{R^{ph}}\cdots\int g_h(x_1,\ldots,x_h)\Pi_{j=1}^h d[F_n(x_j,\omega)-F(x_j)], \tag{3.1}$$

so that

$$nV_{n,1}(\omega) = \sum_{i=1}^n [g_1(X_i)-\theta(F)] = S_n(\omega), \quad \text{say.} \tag{3.2}$$

Now, writing $dF_n(x_j,\omega) = dF(x_j)+d[F_n(x_j,\omega)-F(x_j)]$, $i\le j\le m$, we obtain from (1.1), (1.3), (3.1) and some simplifications that

$$[\theta(F_n,\omega)-\theta(F)-mV_{n,1}(\omega)] = \sum_{h=2}^m \binom{m}{h} V_{n,h}(\omega), \quad n\ge 1. \tag{3.3}$$

Let then $\mathcal{B}_n$ be the σ-field generated by $\{X_i, i\le n\}$, and let

$$V_{n,2}^*(\omega) = n^2 V_{n,2}(\omega), \quad \theta_2^*(F) = \int g_2(x,x)dF(x). \tag{3.4}$$

Lemma 3.1. $\{V^*_{n,2}(\omega)+n[\theta(F)-\theta^*_2(F)], \mathcal{B}_n\}$ forms a martingale sequence.

Proof. By (1.2), (3.1) and (3.4),

$$V^*_{n,2}(\omega) = V^*_{n-1,2}(\omega) + 2\sum_{i=1}^{n-1} \int_{R^{2p}}\cdots\int g_2(x_1,x_2)d[c(x_1-X_i)-F(x_1)]d[c(x_2-X_n)-F(x_2)]$$

(3.5)

$$+ \int_{R^{2p}}\cdots\int g_2(x_1,x_2)d[c(x_1-X_n)-F(x_1)]d[c(x_2-X_n)-F(x_2)].$$

Also, for all $i\leq n-1$,

(3.6) $$E\{d[c(x_1-X_i)-F(x_1)]d[c(x_2-X_n)-F(x_2)] \mid \mathcal{B}_n\} = 0,$$

$$E\{d[c(x_1-X_n)-F(x_1)]d[c(x_2-X_n)-F(x_2)] \mid \mathcal{B}_n\}$$

(3.7)

$$= (\delta_{x_1x_2})dF(x_1)-dF(x_1)dF(x_2),$$

where $\delta_{x_1x_2}$ is 1 or 0 according as $x_1=x_2$ or not. Hence,

(3.8) $$E\{V^*_{n,2}(\omega) \mid \mathcal{B}_{n-1}\} = V^*_{n-1,2}(\omega) + \theta^*_2(F)-\theta(F), \quad n\geq 2,$$

and the lemma follows.

Now, for all $r_j\geq 1$, $j=1,\ldots,\ell(\geq 1)$, $\sum_{j=1}^{\ell} r_j=2s$, $s\geq 1$,

$$|E(\{d[c(x_1-X_1)-F(x_1)]\}^{r_1} \ldots \{d[c(x_\ell-X_\ell)-F(x_\ell)]\}^{r_\ell})|$$

(3.9) $$= 0, \text{ if at least one of } r_1,\ldots,r_\ell = 1,$$

$$\leq dF(x_1)\ldots dF(x_\ell), \text{ otherwise;}$$

and hence, the maximum ℓ for which (3.9) is different from zero is s, where $r_1=\ldots=r_s=2$.

Lemma 3.2. Under (2.3), for every $\varepsilon>0$,

$$\lim_{n\to\infty} P\{\omega:\ \max_{1\le k\le n} k|V_{k,2}(\omega)|>\varepsilon n^{\frac{1}{2}}\} = 0. \tag{3.10}$$

Proof. For every $\varepsilon>0$, there exists an $n_0(\varepsilon)$, such that

$$|\theta(F)-\theta_2^*(F)|<\tfrac{1}{2}\varepsilon n^{\frac{1}{2}} \text{ for all } n\ge n_0(\varepsilon). \tag{3.11}$$

Hence, for $n\ge n_0(\varepsilon)$, by definition in (3.4),

$$\begin{aligned} &P\{\omega:\ \max_{1\le k\le n} k|V_{k,2}(\omega)|>\varepsilon n^{\frac{1}{2}}\} \\ &\le P\{\omega:\ \max_{1\le k\le n} k^{-1}|V_{k,2}^*(\omega)+k[\theta(F)-\theta^*(F)]|>\tfrac{1}{2}\varepsilon n^{\frac{1}{2}}\}. \end{aligned} \tag{3.12}$$

Now, under (2.3), by (3.1) and (3.9),

$$\begin{aligned} E\{V_{k,2}^*(\omega) + k[\theta(F)-\theta_2^*(F)]\}^2 &\le 2\{E[V_{k,2}^*(\omega)]^2 + k^2[\theta(F)-\theta_2^*(F)]^2\} \\ &\le C_1k^2\zeta^*(F), \text{ where } C_1(<\infty) \text{ does not depend on } k. \end{aligned} \tag{3.13}$$

Hence, by theorem 1 of Chow (1960), (i.e., the semi-martingale extension of the Hájek-Rènyi inequality), Lemma 3.1, (3.12) and (3.13), it follows that for $n\ge n_0(\varepsilon)$,

$$P\{\omega:\ \max_{1\le k\le n} k|V_{k,2}(\omega)|>\varepsilon n^{\frac{1}{2}}\}$$

(3.14)
$$\le (4C_1\zeta^*(F)/n\varepsilon^2)(1+\textstyle\sum_{j=1}^{n-1}\{j^{-2}-(j+1)^{-2}\}j^2)$$

$$\le (8C_1\zeta^*(F)[\log n]/n\varepsilon^2) \to 0 \text{ as } n\to\infty. \quad \text{Q.E.D.}$$

<u>Lemma 3.3</u>. <u>If</u> $m\ge 3$ <u>and</u> (2.3) <u>holds, then for every</u> $\varepsilon>0$,

(3.15)
$$\lim_{n\to\infty} P\{\omega:\ \max_{1\le k\le n} k|\textstyle\sum_{h=3}^{m}\binom{m}{h}V_{K,h}(\omega)|>\varepsilon n^{\frac{1}{2}}\} = 0.$$

<u>Proof</u>. By (2.3), (3.1) and (3.9), for all $k\ge 1$,

(3.16)
$$E\{[\textstyle\sum_{h=1}^{3}\binom{m}{h}V_{K,h}(\omega)]^2\} \le C_2k^{-3}\zeta^*(F),\ C_2<\infty.$$

Hence, by the Bonferroni and the Markov inequality,

$$P\{\omega:\ \max_{1\le k\le n} k|\textstyle\sum_{h=3}^{m}\binom{m}{h}V_{k,h}(\omega)|>\varepsilon n^{\frac{1}{2}}\}$$

(3.17)
$$\le \textstyle\sum_{k=1}^{n} P\{\omega:\ k|\sum_{h=3}^{m}\binom{m}{h}V_{k,h}(\omega)|>\varepsilon n^{\frac{1}{2}}\}$$

$$\le \textstyle\sum_{k=1}^{n} \{C_2\zeta^*(F)/n\varepsilon^2\}k^{-1} \le C_2(\log n)\zeta^*(F)/n\varepsilon^2,$$

which converges to 0 as $n\to\infty$. Q.E.D.

From (3.2), (3.3), (3.10) and (3.15), we have for every $\varepsilon'>0$,

(3.18)
$$\lim_{n\to\infty} P\{\omega:\ \max_{1\le k\le n} k|\theta(F_k,\omega)-\theta(F)-mV_{k,1}(\omega)|>\varepsilon' n^{\frac{1}{2}}\} = 0.$$

Now, by (1.4), for all $n \geq m$,

$$U_n(\omega) = n^{-[m]} \sum_{P_{n,m}} \int_{R^{pm}} \cdots \int g(x_1, \ldots, x_m) \Pi_{j=1}^{m} \, d[c(x_j - X_{i_j})]$$

(3.19)

$$= \theta(F) + \sum_{h=1}^{m} \binom{m}{h} U_{n,h}(\omega),$$

where $U_{n,1}(\omega) = V_{n,1}(\omega)$ is given by (3.2), and for $2 \leq h \leq m$,

$$U_{n,h}(\omega) = n^{-[h]} \sum_{P_{n,h}} \int_{R^{kh}} \cdots \int g(x_1, \ldots, x_h) \Pi_{j=1}^{h} \, d[c(x_j - X_{i_j}) - F(x_j)]. \tag{3.20}$$

Lemma 3.4. $\{n^{[h]} U_{n,h}(\omega), \mathcal{B}_n\}$ _forms a martingale sequence for ever_ $1 \leq h \leq m$.

The proof is similar to Lemma 3.1 and may be found in Hoeffding (1961).

Lemma 3.5. _If_ $\zeta_m(F) < \infty$, _then for every_ $h (2 \leq h \leq m)$ _and_ $\varepsilon > 0$,

$$\lim_{n\to\infty} P\{\omega: \max_{1 \leq k \leq n} k |U_{k,h}(\omega)| > \varepsilon n^{\frac{1}{2}}\} = 0. \tag{3.21}$$

By virtue of Lemma 3.4 and the Chow-Hájek-Rènyi inequality, the proof follows exactly as in Lemma 3.2 after noting that $E\{U_{k,h}^2(\omega)\} = O(k^{-h})$, $2 \leq h \leq m$. By (3.19) and Lemma 3.5, we obtain that for every $\varepsilon' > 0$,

$$\lim_{n\to\infty} P\{\omega: \max_{1 \leq k \leq n} k |U_k(\omega) - \theta(F) - mU_{k,1}(\omega)| > \varepsilon' n^{\frac{1}{2}}\} = 0. \tag{3.22}$$

Now from (3.1) and (3.20), for $n \geq m$,

$$V^*_{n,2}(\omega)-n^{[2]}U_{n,2}(\omega) = \sum_{i=1}^n \int\int_{R^{2p}} g_2(x_1,x_2)d[c(x_1-X_i)-F(x_1)]d[c(x_2-X_i)-F(x_2)]$$

(3.23)

$$= Z^*_n(\omega), \text{ say.}$$

Then, by the same technique as in Lemmas 3.1 and 3.2, we have the following.

Lemma 3.6. $\{Z^*_n(\omega)+n[\theta(F)-\theta^*_2(F)], \mathcal{B}_n\}$ forms a martingale sequence, and hence, under (2.3), for every $\varepsilon>0$,

(3.24) $$\lim_{n\to\infty} P\{\omega: \max_{1\le k\le n} k^{-1}|Z^*_k(\omega)|>\varepsilon n^{\frac{1}{2}}\} = 0.$$

Finally, from (3.18), (3.21), Lemma 3.3 and 3.6, we obtain the following.

Lemma 3.7. Under (2.3), for every $\varepsilon>0$,

(3.25) $$\lim_{n\to\infty} P\{\omega: \max_{m\le k\le n} k|\theta(F_k,\omega)-U_k(\omega)|>\varepsilon n^{\frac{1}{2}}\} = 0.$$

4. The proof of the main theorem. For $t\varepsilon I$, define

(4.1) $$Y^o_n(t,\omega) = \{S_{[nt]}(\omega)+(nt-[nt])[g_1(X_{[nt]+1})-\theta(F)]\}/\{n\zeta_1(F)\}^{\frac{1}{2}},$$

where $S_k(\omega)$ is defined by (3.2). Then, by the Donsker theorem [cf. Billingsley (1968, p. 68)], as $n\to\infty$

(4.2) $$Y^o_n(\omega) = [Y^o_n(\omega,t), t\varepsilon I] \overset{\mathcal{D}}{\to} W = [W_t, t\varepsilon I].$$

Now, by (2.5), (3.2), (3.18) and (4.1), under (2.3),

$$\rho(Y_n(\omega),Y_n^o(\omega)) = \sup_{t\varepsilon I}|Y_n(t,\omega)-Y_n^o(t,\omega)|$$

(4.3)

$$= \max_{1\leq k\leq n} k|\theta(F_k,\omega)-\theta(F)-mV_{k,1}(\omega)|/\{m[n\zeta_1(F)]^{\frac{1}{2}}\} \overset{P}{\to} 0, \text{ as } n\to\infty.$$

Hence, by (4.2) and (4.3), under (2.3),

(4.4)
$$Y_n(\omega) \overset{P}{\to} Y_n^o(\omega) \overset{\mathfrak{D}}{\to} W.$$

Similarly, by (3.19) and (4.1), as $n\to\infty$,

(4.5)
$$Y_n^*(\omega) \overset{P}{\to} Y_n^o(\omega) \overset{\mathfrak{D}}{\to} W.$$

Further, by definition, $U_k(\omega) = 0$ for $k<m$, and hence,

$$\rho(Y_n(\omega),Y_n^*(\omega)) \leq \max_{1\leq k<m} k|\theta(F_k,\omega)|/\{m[n\zeta_1(F)]^{\frac{1}{2}}\}$$

(4.6)

$$+ \max_{m\leq k\leq n} k|\theta(F_k,\omega)-U_k(\omega)|/\{m[n\zeta_1(F)]^{\frac{1}{2}}\},$$

where by the Chebyshev and the Bonferroni inequalities, for every $\varepsilon>0$,

$$P\{\omega: \max_{1\leq k<m} k|\theta(F_k,\omega)|/\{m[n\zeta_1(F)]\}^{\frac{1}{2}} >\varepsilon\}$$

(4.7)
$$\leq \sum_{k=1}^{m-1} P\{\omega: |\theta(F_k,\omega)|>\varepsilon[n\zeta_1(F)]^{\frac{1}{2}}\}$$

$$\leq [(m-1)\zeta^*(F)]/[n\varepsilon^2\zeta_1(F)]\to 0 \text{ as } n\to\infty, \text{ [by (2.3)]}$$

and hence, by Lemma 3.7 and (4.6), $\rho(Y_n,Y_n^*) \overset{P}{\to} 0$ as $n\to\infty$. Q.E.D.

5. Some applications.

(I) Asymptotic normality for random sample sizes. For every index variable $r(\geq 1)$, let N_r be an integer valued (non-negative) random variable and n_r is a real (positive) number, such that

$$(5.1) \qquad \lim_{r\to\infty} n_r=\infty \text{ and } \lim_{r\to\infty} (N_r/n_r)=1, \text{ in probability.}$$

This implies that for every $\delta>0$,

$$(5.2) \qquad \lim_{r\to\infty} P\{|n_r^{-1}N_r-1|>\delta\} = 0.$$

Now, by the tightness property [cf. Billingsley (1968, p. 54)] of W_t, $t\geq 0$, for every $\varepsilon>0$ and $\eta>0$, there exists a $\delta>0$, such that

$$(5.3) \qquad P\{\sup_{|t-s|<\delta}|W_t-W_s|>\varepsilon|t,s\varepsilon I\}<\eta.$$

Hence, by theorem 2.1, (5.2) and (5.3), we have as $r\to\infty$,

$$N_r^{\frac{1}{2}}[\theta(F_{N_r},\omega)-\theta(F)] \overset{p}{\sim} n_r^{\frac{1}{2}}[\theta(F_{N_r},\omega)-\theta(F)]$$

$$\overset{p}{\sim} n_r^{\frac{1}{2}}[\theta(F_{n_r},\omega)-\theta(F)],$$

and the asymptotic normality readily follows.

A similar result for $U_{N_r}(\omega)$ has been obtained by Sproule (1969) by an indirect method involving some elaborate analysis. Our proof follows from theorem 2.1, (5.2) and (5.3).

(II) Limiting distributions of $\sup_{t\in I} Y_n(t,\omega)$ etc. By virtue of theorem 2.1 and well-known results on Brownian motion processes [viz., Billingsley (1968, p. 79)], we obtain that for all $\lambda>0$,

(5.5) $$\lim_{n\to\infty} P\{\sup_{t\in I} Y_n(t,\omega)>\lambda\} = \lim_{n\to\infty} P\{\sup_{t\in I} Y_n^*(t,\omega)>\lambda\} = \sqrt{\frac{2}{\pi}} \int_\lambda^\infty e^{-\frac{1}{2}t^2} dt,$$

(5.6) $$\lim_{n\to\infty} P\{\omega: \sup_{t\in I} |Y_n(t,\omega)|>\lambda\} = \lim_{n\to\infty} P\{\omega: \sup_{t\in I} |Y_n^*(t,\omega)|>\lambda\}$$
$$= 1 - \sum_{k=-\infty}^{\infty} (-1)^k [\Phi((2k+1)\lambda) - \Phi((2k-1)\lambda)],$$

where $\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^x \exp(-\frac{1}{2}t^2)dt$, $-\infty<x<\infty$.

(III) An application to rank statistics. Consider the kernel

(5.7) $$\phi(x_1,x_2) = 1, 0 \text{ or } -1 \text{ according as } x_1+x_2 \text{ is } >, = \text{ or } < 0,$$

and assume that the X_i have a continuous d.f. $F(x)$, $x\varepsilon R^1$. Then

(5.8) $$\theta(F) = \int_{-\infty}^\infty [1-2F(-x)]dF(x) \quad (=0 \text{ when } F \text{ is symmetric about } 0).$$

The corresponding $\theta(F_n,\omega)$ is given by

(5.9) $$n\theta(F_n,\omega) = 2n^{-1} \sum_{i=1}^n R_{ni} \operatorname{sgn}(X_i) = 2 W_n, \text{ say},$$

where R_{ni} = Rank of $|X_i|$ among $|X_1|,\ldots,|X_n|$, sgn u=1, 0 or -1 according as u is >, = or < 0, and W_n is the classical Wilcoxon signed rank statistic. Also, here

$$\zeta_1(F) = \int_{-\infty}^{\infty} [1-2F(-x)]^2 dF(x) - \theta^2(F), \tag{5.10}$$

which equals to $^1/_3$ when F is symmetric about 0. Thus, if we let

$$Y_n(t,\omega) = \{(nt-[nt])W_{[nt]+1} + ([nt]+1-nt)W_{[nt]} - \tfrac{1}{2}nt\theta(F)\}/[n\zeta_1(F)]^{\frac{1}{2}}. \tag{5.11}$$

$0 \leq t \leq 1$, we have from theorem 2.1 that if $\zeta_1(F)>0$,

$$Y_n(\omega) \overset{\mathcal{D}}{\to} W, \tag{5.12}$$

and thus the sequence $\{W_k - \tfrac{1}{2}k\theta(F);\ k \geq 1\}$ though not linear in the basic random variables is attracted by the Brownian motion process.

REFERENCES

[1] BILLINGSLEY, P. (1968). Convergence of Probability Measures. John Wiley & Sons, New York.

[2] CHOW, Y. S. (1960). A martingale inequality and the law of large numbers. Proc. Amer. Math. Soc. 11, 107-111.

[3] DONSKER, M. (1951). An invariance principle for certain probability limit theorems. Mem. Amer. Math. Soc. 6, 1-12.

[4] DONSKER, M. (1952). Justification and extension of Doob's heuristic approach to the Kolmogoroll-Smirnov theorems. Ann. Math. Statist. 23, 277-281.

[5] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 19, 293-325.

[6] HOEFFDING, W. (1961). The strong law of large numbers for U-statistics. Inst. Statist. Univ. N. Carolina Mimeo Rep. No. 302.

[7] MISES, R. V. (1947). On the asymptotic distribution of differentiable statistical functions. Ann. Math. Statist. 18, 309-348.

[8] SPROULE, R. N. (1969). A sequential fixed-width confidence interval for the mean of a U-statistic. Doctoral Dissertation, Univ. N. Carolina, Chapel Hill.