The Annals of Applied Probability
1992, Vol. 2, No. 2, 435–459

RANDOM WALK PROCESSES AND THEIR APPLICATIONS IN ORDER STATISTICS

BY LAJOS TAKÁCS

Case Western Reserve University

This paper is concerned with two stochastic processes, namely, a Bernoulli excursion and a tied-down random walk. Three random variables are defined for these processes, each variable representing the area of a random set determined by one of the processes. The aim is to find the distributions and the moments of these random variables and to determine their asymptotic behavior. The results derived for random walks are applied to the theory of order statistics to determine the asymptotic behavior of the moments and the distributions of two statistics which measure the deviation between two empirical distribution functions.

1. Introduction. We shall consider two random walk processes. One is the Bernoulli excursion process $\{\eta_0^+, \eta_1^+, \dots, \eta_{2n}^+\}$, that is, a random walk for which $\eta_{2n}^+ = \eta_0^+ = 0$ and $\eta_i^+ \ge 0$ for $0 \le i \le 2n$. The other is a tied-down random walk $\{\eta_0, \eta_1, \dots, \eta_{2n}\}$ for which $\eta_{2n} = \eta_0 = 0$.

For these processes we define the random variables ω_n, ρ_n and σ_n by

$$2n\omega_n = \sum_{i=1}^{2n} \eta_i^+ \tag{1}$$

for $n \ge 1$ and $\omega_0 = 0$,

$$2n\rho_n = \sum_{i=1}^{2n} (\eta_i + \delta_{2n}) \tag{2}$$

for $n \ge 1$, where

$$\delta_{2n} = -\min(\eta_0, \eta_1, \dots, \eta_{2n}) \tag{3}$$

and $\rho_0 = 0$, and

$$2n\sigma_n = \sum_{i=1}^{2n} |\eta_i| \tag{4}$$

for $n \ge 1$ and $\sigma_0 = 0$.

We shall determine the distributions and the moments of these random variables and their asymptotic behavior as $n \to \infty$.

Afterwards, we apply the results derived for random walks to the theory of order statistics. We assume that $F_n(x)$ and $G_n(x)$ are the empirical distribution functions of two independent samples of size n in the case where the elements of the two samples are independent random variables each having the same

Received April 1990; revised February 1991.
AMS 1980 *subject classifications*. Primary 60J15, 62G30; secondary 60C05.
Key words and phrases. Random walk, order statistics, limit distributions.

continuous distribution function $V(x)$. We define two statistics Θ_n and Δ_n by

$$(5)\qquad \Theta_n/n = \int_{-\infty}^{\infty}[F_n(x) - G_n(x)]\,dV(x) - \min_{-\infty<x<\infty}[F_n(x) - G_n(x)]$$

and

$$(6)\qquad \Delta_n/n = \int_{-\infty}^{\infty}|F_n(x) - G_n(x)|\,dV(x)$$

and determine the asymptotic behavior of the moments and the distributions of these statistics.

2. A Bernoulli excursion. We define a sequence of random variables $\{\eta_0^+, \eta_1^+, \dots, \eta_{2n}^+\}$ in the following way: Let us arrange n white and n black balls in a row in such a way that for every $i = 1, 2, \dots, 2n$, among the first i balls there are at least as many white balls as black balls. The total number of such sequences is given by the nth Catalan number,

$$(7)\qquad C_n = \binom{2n}{n}\frac{1}{n+1}.$$

We have $C_0 = 1$, $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, $C_5 = 42, \dots$.

In 1879, Whitworth [18, 19] showed that the number of ways in which a gains and b losses can be arranged in such a way that the losses are never in excess of the gains is

$$(8)\qquad N(a,b) = \binom{a+b}{a}\frac{a+1-b}{a+1}$$

if $a \geq b$. If $a = b = n$, then (8) reduces to (7).

Let us suppose that all the possible C_n sequences are equally probable and denote by η_i^+ the difference between the number of white balls and the number of black balls among the first i balls in a sequence chosen at random. We have $\eta_{2n}^+ = \eta_0^+ = 0$ and $\eta_i^+ \geq 0$ for $i = 1, 2, \dots, 2n$.

The sequence $\{\eta_0^+, \eta_1^+, \dots, \eta_{2n}^+\}$ is usually called a Bernoulli excursion. We can imagine that a particle performs a random walk on the x-axis. It starts at $x = 0$ and takes $2n$ steps. In the ith step the particle moves either a unit distance to the right or a unit distance to the left according as the ith ball in the row is white or black respectively. At the end of the ith step the position of the particle is $x = \eta_i^+$ for $i = 1, 2, \dots, 2n$.

As an alternative we can assume that the particle starts at time $t = 0$ at the origin and in the time interval $(i-1, i]$, $i = 1, 2, \dots, 2n$, it moves with a unit velocity to the right or to the left according as the ith ball in the row of balls is white or black respectively. Denote by $\eta_n^+(t)$ the position of the particle at time $2nt$ where $0 \leq t \leq 1$. Then $\eta_n^+(i/2n) = \eta_i^+$ for $i = 1, 2, \dots, 2n$.

Let us define a random variable ω_n for $n = 1, 2, \dots$ by

$$(9)\qquad 2n\omega_n = \sum_{i=1}^{2n}\eta_i^+$$

and set $\omega_0 = 0$. By (9),

$$\omega_n = \int_0^1 \eta^+_{[2nt]}\,dt = \int_0^1 \eta_n^+(t)\,dt \tag{10}$$

for $n = 1, 2, \dots$.

The random variable $2n\omega_n$ is a discrete random variable with possible values $n + 2j$, $j = 0, 1, \dots, \binom{n}{2}$. Denote by $f_n(n + 2j)$ the number of sequences $\{\eta_0^+, \eta_1^+, \dots, \eta_{2n}^+\}$ in which $2n\omega_n = n + 2j$. Then we have

$$P\{2n\omega_n = n + 2j\} = f_n(n + 2j)/C_n \tag{11}$$

for $j = 0, 1, \dots, \binom{n}{2}$.

The distribution of $2n\omega_n$ is determined by the generating function

$$\phi_n(z) = \sum_{j=0}^{\binom{n}{2}} f_n(n + 2j) z^j, \tag{12}$$

which can be obtained by the following theorem.

THEOREM 1. *We have*

$$\phi_n(z) = \sum_{i=1}^{n} \phi_{i-1}(z)\phi_{n-i}(z) z^{i-1} \tag{13}$$

for $n = 1, 2, \dots$ *and* $\phi_0(z) = 1$.

PROOF. If $i = 1, 2, \dots, n$ is the smallest positive integer for which $\eta_{2i}^+ = 0$, then in the representation

$$\begin{aligned} \eta_1^+ + \cdots + \eta_{2n}^+ = 2i - 1 + (\eta_1^+ - 1) \\ + \cdots + (\eta_{2i-1}^+ - 1) + \eta_{2i}^+ + \cdots + \eta_{2n}^+, \end{aligned} \tag{14}$$

the sum $(\eta_1^+ - 1) + \cdots + (\eta_{2i-1}^+ - 1)$ has the same distribution as $2(i-1)\omega_{i-1}$ and the sum $\eta_{2i}^+ + \cdots + \eta_{2n}^+$ has the same distribution as $2(n-i)\omega_{n-i}$ and these two random variables are independent. If we use the notation (12), then by (14) we obtain (13), which was to be proved. □

By (13) we obtain that $\phi_1(z) = 1$, $\phi_2(z) = 1 + z$ and $\phi_3(z) = 1 + 2z + z^2 + z^3$. Table 1 contains $f_n(n + 2j)$ for $0 \le j \le \binom{n}{2}$ and $n \le 10$. In Table 1 the rows are wrapped and only j (mod 10) is displayed.

Now let us define

$$F(z, w) = w \sum_{n=0}^{\infty} \phi_n(z)(zw)^n = w \sum_{n=0}^{\infty} \sum_{j=0}^{\binom{n}{2}} f_n(n + 2j) z^{n+j} w^n. \tag{15}$$

TABLE 1
$f_n(n+2j)$

	j									
n	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	1									
2	1	1								
3	1	2	1	1						
4	1	3	3	3	2	1	1			
5	1	4	6	7	7	5	5	3	2	1
	1									
6	1	5	10	14	17	16	16	14	11	9
	7	5	3	2	1	1				
7	1	6	15	25	35	40	43	44	40	37
	32	28	22	18	13	11	7	5	3	2
	1	1								
8	1	7	21	41	65	86	102	115	118	118
	113	106	96	85	73	63	53	42	34	26
	20	15	11	7	5	3	2	1	1	
9	1	8	28	63	112	167	219	268	303	326
	338	338	331	314	293	268	245	215	190	162
	139	116	97	77	63	48	38	28	22	15
	11	7	5	3	2	1	1			
10	1	9	36	92	182	301	434	574	704	813
	901	959	995	1003	990	958	918	862	801	734
	665	598	531	466	405	348	295	249	207	171
	138	111	87	69	52	40	30	22	15	11
	7	5	3	2	1	1				

Since $|\phi_n(z)| \le \phi_n(1) = C_n$ if $|z| \le 1$ and since

$$w \sum_{n=0}^{\infty} C_n w^n = \left[1 - (1-4w)^{1/2}\right]/2 \tag{16}$$

if $|w| \le 1/4$, the series (15) is convergent if $|z| \le 1$ and $|zw| \le 1/4$.

Multiplying (13) by $(zw)^n$ and forming the sum for $n = 1, 2, \dots,$ we obtain

$$[1 - F(z, zw)]F(z, w) = w \tag{17}$$

for $|z| \le 1$ and $|zw| \le 1/4$. The repeated application of (17) leads to the continued fraction

$$F(z,w) = \cfrac{w}{1 - \cfrac{zw}{1 - \cfrac{z^2 w}{1 - \ddots}}}. \tag{18}$$

The continued fraction (18) has been encountered by Ramanujan [10] in the theory of partitions. (See, e.g., Hardy and Wright [6], page 295.)

3. A tied-down random walk. We define a sequence of random variables $\{\eta_0, \eta_1, \dots, \eta_{2n}\}$ in the following way. Let us suppose that a box contains n white and n black balls. We draw all the $2n$ balls one by one without replacement from the box. There are $\binom{2n}{n}$ possible results and they are supposed to be equally probable. Define η_i, $i = 0, 1, \dots, 2n$, as the difference between the number of white balls and the number of black balls among the first i balls drawn ($\eta_{2n} = \eta_0 = 0$).

We can interpret the sequence $\{\eta_0, \eta_1, \dots, \eta_{2n}\}$ as a random walk on the real line. A particle starts at $x = 0$ and in the ith step it moves either a unit distance to the right or a unit distance to the left according as the ith ball drawn is white or black respectively. Altogether n steps are taken in the positive direction and n steps in the negative direction. At the end of the ith step the position of the particle is $x = \eta_i$.

As an alternative we can assume that the particle starts at time $t = 0$ at the origin and in the time interval $(i - 1, i]$, $i = 1, 2, \dots, 2n$, it moves with a unit velocity to the right or to the left according as the ith ball drawn is white or black respectively. Denote by $\eta_n(t)$ the position of the particle at time $2nt$ where $0 \le t \le 1$. Then $\eta_n(i/2n) = \eta_i$ for $i = 1, 2, \dots, 2n$.

Let us define a random variable ρ_n for $n = 1, 2, \dots$ by

$$2n\rho_n = \sum_{i=1}^{2n} (\eta_i + \delta_{2n}), \tag{19}$$

where

$$\delta_{2n} = -\min(\eta_0, \eta_1, \dots, \eta_{2n}), \tag{20}$$

and set $\rho_0 = 0$. By (19),

$$\rho_n = \int_0^1 \eta_{[2nt]}\, dt - \min_{0 \le t \le 1} \eta_{[2nt]} = \int_0^1 \eta_n(t)\, dt - \min_{0 \le t \le 1} \eta_n(t) \tag{21}$$

for $n = 1, 2, \dots$.

The random variable $2n\rho_n$ is a discrete random variable with possible values $n + 2j$, $j = 0, 1, \dots, \binom{n}{2}$. Denote by $h_n(n + 2j)$ the number of random walks $\{\eta_0, \eta_1, \dots, \eta_{2n}\}$ in which $2n\rho_n = n + 2j$. Then we have

$$P\{2n\rho_n = n + 2j\} = h_n(n + 2j) \Big/ \binom{2n}{n} \tag{22}$$

for $j = 0, 1, \dots, \binom{n}{2}$.

For the purpose of finding the distribution of $2n\rho_n$, let us introduce the generating function

$$\psi_n(z) = \sum_{j=0}^{\binom{n}{2}} h_n(n + 2j) z^j. \tag{23}$$

This generating function is determined by the following theorem.

THEOREM 2. *We have* $\psi_0(z) = 1$ *and*

$$\psi_n(z) = 2\sum_{i=1}^{n} i\phi_{i-1}(z)\phi_{n-i}(z)z^{i-1} \tag{24}$$

for $n = 1, 2, \ldots,$ *where* $\phi_n(z)$, $n = 0, 1, 2, \ldots,$ *is determined by the recurrence formula* (13).

PROOF. Let $n \geq 1$. In the random walk $\{\eta_0, \eta_1, \ldots, \eta_{2n}\}$ it may happen that $\delta_{2n} = 0$, that is, $\eta_s \geq 0$ for $0 \leq s \leq 2n$. Then $2n\rho_n$ simply has the same distribution as $2n\omega_n$. If $\delta_{2n} \geq 1$, then let $s = i$ be the first subscript for which $\eta_s = -\delta_{2n}$ and let $s = i + 2(n-1-k)$ be the last subscript for which $\eta_s = -\delta_{2n}$. Then i may be $1, \ldots, 2k+1$ and k may be $0, 1, \ldots, n-1$. Now let us consider a new random walk defined by

$$\{\eta_i + \delta_{2n}, \ldots, \eta_{2n-1}^{+} + \delta_{2n}, \eta_0 + \delta_{2n}, \ldots, \eta_i + \delta_{2n}\}. \tag{25}$$

That is, in the original random walk we transfer the first i steps from the beginning to the end and shift the zero level to $-\delta_{2n}$. For fixed i and k, the random walk (25) has the same stochastic properties as the Bernoulli excursion $\{\eta_0^+, \eta_1^+, \ldots, \eta_{2n}^+\}$ in which $\eta_{2(n-1-k)}^+ = 0$ and $\eta_s^+ > 0$ for $2(n-1-k) < s < 2n$, and $2n\rho_n$ has the same distribution as $\eta_0^+ + \eta_1^+ + \cdots + \eta_{2n}^+$. Under the aforementioned conditions $\eta_0^+ + \cdots + \eta_{2(n-1-k)}^+$ has the same distribution as $2(n-1-k)\omega_{n-1-k}$ and $\eta_{2(n-1-k)}^+ + \cdots + \eta_{2n-1}^+$ has the same distribution as $2k + 1 + 2k\omega_k$, and these two sums are independent random variables. Obviously, $\eta_{2n}^+ = 0$. By the above considerations,

$$\begin{aligned}\psi_n(z) &= \phi_n(z) + \sum_{k=0}^{n-1}\sum_{i=1}^{2k+1} \phi_k(z)\phi_{n-1-k}(z)z^k \\ &= 2\sum_{k=0}^{n-1}(k+1)\phi_k(z)\phi_{n-1-k}(z)z^k\end{aligned} \tag{26}$$

for $n \geq 1$, where $\phi_n(z)$ is defined by (12) and determined by (13). By definition, $\psi_0(z) = 1$. This completes the proof of (24). □

It is worthwhile to point out the significance of formula (24). To find the distribution of $2n\omega_n$, we should determine the generating functions $\phi_0(z), \phi_1(z), \ldots, \phi_n(z)$. If these functions are known, then the distribution of $2n\rho_n$ can immediately be calculated by (24). No extra calculations are needed, although the random variable $2n\rho_n$ is much more complicated than $2n\omega_n$.

By (24) we obtain that $\psi_0(z) = 1$, $\psi_1(z) = 2$, $\psi_2(z) = 2 + 4z$ and $\psi_3(z) = 2 + 6z + 6z^2 + 6z^3$. Table 2 contains $h_n(n + 2j)$ for $0 \leq j \leq \binom{n}{2}$ and $n \leq 10$. In Table 2 the rows are wrapped and only j (mod 10) is displayed.

Let us define

$$\psi(z, w) = \sum_{n=0}^{\infty} \psi_n(z)w^n = \sum_{n=0}^{\infty}\sum_{j=0}^{\binom{n}{2}} h_n(n + 2j)z^j w^n. \tag{27}$$

TABLE 2
$h_n(n + 2j)$

	j									
n	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	2									
2	2	4								
3	2	6	6	6						
4	2	8	12	16	8	8				
5	2	10	20	30	40	40	40	30	20	10
	10									
6	2	12	30	52	78	96	114	120	108	96
	72	60	36	24	12	12				
7	2	14	42	84	140	196	252	308	336	350
	336	308	266	224	168	140	98	70	42	28
	14	14								
8	2	16	56	128	236	368	512	672	816	944
	1032	1072	1064	1008	928	816	720	592	496	384
	304	224	176	112	80	48	32	16	16	
9	2	18	72	186	378	648	978	1368	1782	2202
	2610	2952	3222	3366	3402	3330	3186	2952	2700	2394
	2106	1800	1530	1260	1044	810	648	486	378	270
	198	126	90	54	36	18	18			
10	2	20	90	260	580	1084	1770	2640	3660	4780
	5980	7180	8340	9360	10170	10720	11020	11040	10810	10400
	9790	9120	8320	7520	6680	5900	5080	4380	3680	3100
	2540	2080	1640	1320	1000	780	580	440	300	220
	140	100	60	40	20	20				

Since $|\psi_n(z)| \le \psi_n(1) = \binom{2n}{n}$ for $|z| \le 1$ and since

$$\sum_{n=0}^{\infty} \binom{2n}{n} w^n = (1 - 4w)^{-1/2} \tag{28}$$

for $|w| < 1/4$, the infinite series (27) is convergent for $|z| \le 1$ and $|w| < 1/4$.

If we multiply (24) by w^n and form the sum for $n = 1, 2, \dots$, we obtain

$$\psi(z, w) = 1 + \frac{2w}{1 - F(z, w)} \left(\frac{\partial F(z, w)}{\partial w} \right) \tag{29}$$

for $|z| \le 1$ and $|w| < 1/4$, where $F(z, w)$ is defined by (15) and determined by (17).

We shall consider also a random variable σ_n defined by

$$2n\sigma_n = \sum_{i=1}^{2n} |\eta_i| \tag{30}$$

for $n = 1, 2, \dots$ and $\sigma_0 = 0$. By (30) we have

$$\sigma_n = \int_0^1 |\eta_{[2nt]}| \, dt = \int_0^1 |\eta_n(t)| \, dt \tag{31}$$

for $n = 1, 2, \dots$.

The random variable $2n\sigma_n$ is a discrete random variable with possible values $n+2j$, $j=0,1,\ldots,\binom{n}{2}$. Denote by $q_n(n+2j)$ the number of random walks $\{\eta_0,\eta_1,\ldots,\eta_{2n}\}$ in which $2n\sigma_n = n+2j$. Then we have

$$(32)\qquad P\{2n\sigma_n = n+2j\} = q_n(n+2j)\Big/\binom{2n}{n}$$

for $j=0,1,\ldots,\binom{n}{2}$.

The distribution of $2n\sigma_n$ is determined by the generating function

$$(33)\qquad t_n(z) = \sum_{j=0}^{\binom{n}{2}} q_n(n+2j)z^j,$$

which is given by the following theorem.

THEOREM 3. *We have*

$$(34)\qquad t_n(z) = 2\sum_{i=1}^{n} \phi_{i-1}(z)t_{n-i}(z)z^{i-1}$$

for $n \geq 1$, $t_0(z)=1$ *and* $\phi_n(z)$ *for* $n=0,1,2,\ldots$ *is determined by the recurrence formula* (13).

PROOF. If in the random walk $\{\eta_0,\eta_1,\ldots,\eta_{2n}\}$, $i=1,2,\ldots,n$ is the smallest positive integer for which $\eta_{2i}=0$, then in the representation

$$(35)\qquad 2n\sigma_n = |\eta_1| + \cdots + |\eta_{2i}| + |\eta_{2i+1}| + \cdots + |\eta_{2n}|,$$

the sum $|\eta_1| + \cdots + |\eta_{2i}|$ has the same distribution as $2(i-1)\omega_{i-1} + 2i - 1$, where ω_{i-1} is defined by (1), and $|\eta_{2i+1}| + \cdots + |\eta_{2n}|$ has the same distribution as $2(n-i)\sigma_{n-i}$, and these two random variables are independent. Consequently, (34) holds for $n \geq 1$ where $t_0(z)=1$. In (34) the factor 2 should be included because of the symmetry property of the random walk $\{\eta_0,\eta_1,\ldots,\eta_{2n}\}$. □

By (34) we obtain that $t_0(z)=1$, $t_1(z)=2$, $t_2(z)=4+2z$ and $t_3(z)=8+8z+2z^2+2z^3$. Table 3 contains $q_n(n+2j)$ for $0 \leq j \leq \binom{n}{2}$ and $n \leq 10$. In Table 3 the rows are wrapped and only j (mod 10) is indicated.

If we introduce the generating function

$$(36)\qquad T(z,w) = \sum_{n=0}^{\infty} t_n(z)w^n = \sum_{n=0}^{\infty}\sum_{j=0}^{\binom{n}{2}} q_n(n+2j)z^jw^n,$$

which is convergent if $|z| \leq 1$ and $|w| < 1/4$, then by (15) and (34) we obtain

$$(37)\qquad T(z,w) = 1/[1-2F(z,w)]$$

for $|z| \leq 1$ and $|w| < 1/4$, where $F(z,w)$ is defined by (15) and is determined by (17).

TABLE 3
$q_n(n+2j)$

n	j = 0	1	2	3	4	5	6	7	8	9
1	2									
2	4	2								
3	8	8	2	2						
4	16	24	12	10	4	2	2			
5	32	64	48	40	26	14	14	6	4	2
	2									
6	64	160	160	144	116	74	68	44	30	22
	18	10	6	4	2	2				
7	128	384	480	480	440	328	290	226	164	132
	106	80	56	44	30	26	14	10	6	4
	2	2								
8	256	896	1344	1504	1520	1288	1148	986	772	646
	534	438	340	278	216	184	138	104	80	60
	44	34	22	14	10	6	4	2	2	
9	512	2048	3584	4480	4928	4640	4288	3912	3290	2838
	2434	2074	1722	1444	1196	1022	852	684	570	460
	376	306	242	182	146	108	84	60	48	30
	22	14	10	6	4	2	2			
10	1024	4608	9216	12800	15232	15680	15232	14560	13028	11610
	10308	9048	7850	6768	5818	5046	4384	3694	3176	2692
	2296	1958	1652	1362	1144	938	774	628	516	406
	320	250	194	150	112	84	64	44	30	22
	14	10	6	4	2	2				

4. The asymptotic distribution of ω_n. First, let us consider the moments of $2n\omega_n$. If we form the first r derivatives of (17) with respect to z at $z = 1$, we can determine the generating functions of the first r moments of $2n\omega_n$. Our calculations show that

$$C_n E\{(2n\omega_n + 2n + 1)^r\} = \begin{cases} C_n Q_r(n), & \text{if } r = 0, 2, 4, \ldots, \\ 4^n Q_r(n), & \text{if } r = 1, 3, 5, \ldots, \end{cases} \tag{38}$$

where $Q_r(n)$ is a polynomial of degree $3[r/2]$ in n. In particular, we have $Q_0(n) = Q_1(n) = 1$,

$$3Q_2(n) = 10n^3 + 21n^2 + 14n + 3, \tag{39}$$

$$4Q_3(n) = 15n^3 + 27n^2 + 18n + 4, \tag{40}$$

$$315Q_4(n) = 4420n^6 + 15912n^5 + 24883n^4 + 21564n^3 + 10702n^2 + 2844n + 315 \tag{41}$$

and

$$32Q_5(n) = 565n^6 + 1695n^5 + 2485n^4 + 2105n^3 + 1030n^2 + 280n + 32. \tag{42}$$

Furthermore, we obtain

$$Q_r(n) = M_r(8/\pi)^{r/2}\pi^{[r/2]}n^{3[r/2]} + \cdots \tag{43}$$

for $r = 0, 1, 2, \ldots$. In (43) only the leading term is displayed. The neglected terms have smaller order than the displayed one. In (43),

$$M_r = K_r \frac{4\sqrt{\pi}\, r!}{\Gamma\left(\dfrac{3r-1}{2}\right)2^{r/2}} \tag{44}$$

for $r = 0, 1, 2, \ldots$, where $K_0 = -1/2$, $K_1 = 1/8$ and K_r, $r = 2, 3, \ldots$, can be obtained by the recurrence formula

$$K_r = \frac{(3r-4)}{4}K_{r-1} + \sum_{j=1}^{r-1} K_j K_{r-j}. \tag{45}$$

The asymptotic behavior of the moments of $2n\omega_n$ is given by the following theorem.

THEOREM 4. *We have*

$$\lim_{n\to\infty} E\left\{\left(\omega_n/\sqrt{2n}\right)^r\right\} = M_r \tag{46}$$

for $r = 0, 1, 2, \ldots$, *where* M_r *is determined by* (44) *and* (45).

PROOF. Since

$$C_n 4^{-n} \sim 1/(n^3\pi)^{1/2} \tag{47}$$

as $n \to \infty$, by (38) and (43) we obtain (46). □

THEOREM 5. *There exists a distribution function* $W(x)$ *such that*

$$\lim_{n\to\infty} P\left\{\omega_n/\sqrt{2n} \le x\right\} = W(x) \tag{48}$$

at every continuity point of $W(x)$. *The distribution function* $W(x)$ *is uniquely determined by its moments*

$$\int_0^\infty x^r\, dW(x) = M_r \tag{49}$$

for $r = 0, 1, 2, \ldots$, *where* M_r *is defined by* (44) *and* (45).

PROOF. By (45) we can prove that

$$\lim_{r\to\infty}\left(\frac{4}{3}\right)^r \frac{K_r}{(r-1)!} = \frac{1}{2\pi}. \tag{50}$$

TABLE 4

r	K_r	M_r
0	$-\frac{1}{2}$	1
1	$\frac{1}{8}$	$\sqrt{\frac{\pi}{8}}$
2	$\frac{5}{64}$	$\frac{5}{12}$
3	$\frac{15}{128}$	$\frac{15}{32}\sqrt{\frac{\pi}{8}}$
4	$\frac{1105}{4096}$	$\frac{221}{1008}$
5	$\frac{1695}{2048}$	$\frac{565}{2048}\sqrt{\frac{\pi}{8}}$
6	$\frac{414125}{131072}$	$\frac{82825}{576576}$
7	$\frac{59025}{4096}$	$\frac{19675}{98304}\sqrt{\frac{\pi}{8}}$
8	$\frac{1282031525}{16777216}$	$\frac{256406305}{2234808576}$
9	$\frac{242183775}{524288}$	$\frac{16145585}{92274688}\sqrt{\frac{\pi}{8}}$
10	$\frac{1683480621875}{536870912}$	$\frac{304702375}{2790982656}$

This implies that

$$M_r \sim 3\sqrt{2}\, r\left(\frac{r}{12e}\right)^{r/2} \tag{51}$$

as $r \to \infty$, and

$$\sum_{r=1}^{\infty} \frac{1}{M_r^{1/r}} = \infty. \tag{52}$$

By (46) the sequence $\{M_r\}$ is a moment sequence. Since (52) is satisfied, a theorem of Carleman [2] implies that there exists one and only one distribution function $W(x)$ such that $W(0) = 0$ and (49) holds for $r = 0, 1, 2, \ldots$. By the moment convergence theorem of Fréchet and Shohat [4], we can conclude that (46) implies (48). □

Table 4 contains K_r and M_r for $r \leq 10$.

5. The distribution function $W(x)$. We can easily prove that the finite-dimensional distributions of the process $\{\eta_n^+(t)/\sqrt{2n}\,;\ 0 \leq t \leq 1\}$ converge to the corresponding finite-dimensional distributions of a Markov process $\{\eta^+(t);\ 0 \leq t \leq 1\}$ which is called the Brownian excursion process.

Moreover, the process $\{\eta_n^+(t)/\sqrt{2n}\,;\ 0 \le t \le 1\}$ converges weakly to $\{\eta^+(t);\ 0 \le t \le 1\}$ as $n \to \infty$. For the weak convergence of stochastic processes, we refer to Gikhman and Skorokhod [5], pages 438–495.

In the Brownian excursion process, $\eta^+(t)$ has a density function $f(t,x)$ for $0 < t < 1$. Obviously, $f(t,x) = 0$ for $x \le 0$. If $0 < t < 1$ and $x > 0$, we have

$$f(t,x) = \frac{2x^2}{\sqrt{2\pi t^3(1-t)^3}} e^{-x^2/(2t(1-t))}. \tag{53}$$

If $0 < t < u < 1$, the random variables $\eta^+(t)$ and $\eta^+(u)$ have a joint density function $f(t,x;u,y)$. We have $f(t,x;u,y) = 0$ if $x \le 0$ or $y \le 0$. If $0 < t < u < 1$ and $x > 0$, $y > 0$, we have

$$\begin{aligned} f(t,x;u,y) &= \frac{\sqrt{8\pi}\,xy}{\sqrt{t^3(u-t)(1-u)^3}} \phi\left(\frac{x}{\sqrt{t}}\right)\phi\left(\frac{y}{\sqrt{1-u}}\right) \\ &\quad \times \left[\phi\left(\frac{y-x}{\sqrt{u-t}}\right) - \phi\left(\frac{y+x}{\sqrt{u-t}}\right)\right], \end{aligned} \tag{54}$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{55}$$

is the normal density function. Since $\{\eta^+(t);\ 0 \le t \le 1\}$ is a Markov process, the density functions $f(t,x)$ and $f(t,x;u,y)$ completely determine the finite-dimensional distributions of the process.

Let us define the random variable ω^+ by the integral

$$\omega^+ = \int_0^1 \eta^+(t)\,dt. \tag{56}$$

THEOREM 6. *We have*

$$P\{\omega^+ \le x\} = W(x), \tag{57}$$

where $W(x)$ *is defined by* (49).

PROOF. Since $\{\eta_n^+(t)/\sqrt{2n}\,;\ 0 \le t \le 1\}$ converges weakly to the process $\{\eta^+(t);\ 0 \le t \le 1\}$ and since the integral (56) is a continuous functional on the process $\{\eta^+(t);\ 0 \le t \le 1\}$, (48) implies (57). □

The distribution function $W(x)$ has a density function $W'(x)$ and both $W(x)$ and $W'(x)$ can be calculated explicitly by using the confluent hypergeometric function $U(a,b,x)$. If $0 < a < b$ and $x > 0$, then

$$U(a,b,x) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-tx} t^{a-1} (1+t)^{b-a-1}\,dt, \tag{58}$$

TABLE 5
$\mathrm{Ai}(-a_k) = 0$

k	a_k	k	a_k
1	2.33810741	6	9.02265085
2	4.08794944	7	10.04017434
3	5.52055983	8	11.00852430
4	6.78670809	9	11.93601556
5	7.94413359	10	12.82877675

and, in general, we have

$$U(a-1,b,x) = (a-b+x)U(a,b,x) - xU'(a,b,x). \tag{59}$$

See Slater [14]. We also need the definition of the Airy function

$$\mathrm{Ai}(z) = \frac{1}{\pi}\int_0^\infty \cos\left(\frac{t^3}{3} + tz\right) dt. \tag{60}$$

The function $\mathrm{Ai}(z)$ has zeros only on the negative real axis. We arrange the zeros $z = -a_k$, $k = 1, 2, \dots$, so that $0 < a_1 < a_2 < \cdots$. The derivative $\mathrm{Ai}'(z)$ has zeros also only on the negative real axis, and we arrange them so that $z = -a'_k$, $k = 1, 2, \dots$, and $0 < a'_1 < a'_2 < \cdots$. Tables 5 and 6 contain the first 10 zeros of $\mathrm{Ai}(z)$ and $\mathrm{Ai}'(z)$. The first 50 zeros of $\mathrm{Ai}(z)$ and $\mathrm{Ai}'(z)$ can be found in Miller [9], page 43, for eight decimals. See also Abramowitz and Stegun [1], page 478.

THEOREM 7. *If $x > 0$, we have*

$$W(x) = \frac{\sqrt{6}}{x}\sum_{k=1}^{\infty} e^{-v_k} v_k^{2/3} U(1/6, 4/3, v_k) \tag{61}$$

and

$$W'(x) = \frac{2\sqrt{6}}{x^2}\sum_{k=1}^{\infty} e^{-v_k} v_k^{2/3} U(-5/6, 4/3, v_k), \tag{62}$$

TABLE 6
$\mathrm{Ai}'(-a'_k) = 0$

k	a'_k	k	a'_k
1	1.01879297	6	8.48848673
2	3.24819758	7	9.53544905
3	4.82009921	8	10.52766040
4	6.16330736	9	11.47505663
5	7.37217726	10	12.38478837

TABLE 7

x	$W(x)$	x	$W(x)$
0.05	0.00000000	0.80	0.86427925
0.10	0.00000000	0.85	0.91153523
0.15	0.00000000	0.90	0.94430335
0.20	0.00000000	0.95	0.96610624
0.25	0.00001007	1.00	0.98005322
0.30	0.00071659	1.05	0.98864280
0.35	0.00858774	1.10	0.99374158
0.40	0.04027493	1.15	0.99666130
0.45	0.11029731	1.20	0.99827530
0.50	0.21745116	1.25	0.99913710
0.55	0.34719272	1.30	0.99958179
0.60	0.48159861	1.35	0.99980363
0.65	0.60641841	1.40	0.99991065
0.70	0.71328366	1.45	0.99996060
0.75	0.79906372	1.50	0.99998317

where

(63) $$v_k = 2a_k^3/(27x^2),$$

$U(a,b,x)$ is a confluent hypergeometric function and $z = -a_k$, $k = 1, 2, \dots$, are the roots of Ai$(z) = 0$ *arranged so that* $0 < a_1 < a_2 < \cdots < a_k < \cdots$.

PROOF. From (49) or from the results of Darling [3] and Louchard [8], we can deduce that

(64) $$\int_0^\infty e^{-sx} W(x)\, dx = \sqrt{2\pi} \sum_{k=1}^{\infty} e^{-a_k s^{2/3}/2^{1/3}}$$

for Re$(s) > 0$, where $z = -a_k$, $k = 1, 2, \dots$, are the zeros of the Airy function Ai(z) arranged so that $0 < a_1 < a_2 < \cdots$. We obtain (61) and (62) from (64) by inversion. □

Tables 7 and 8 contain $W(x)$ and $W'(x)$ for $0 < x \leq 1.5$.

6. The asymptotic distribution of ρ_n. Equation (29) makes it possible to determine the moments of $2n\rho_n$. If we form the first r derivatives of (29) with respect to z at $z = 1$, we can determine the generating functions of the first r moments of $2n\rho_n$. Thus we obtain

(65) $$\binom{2n}{n} E\{(2n\rho_n + n)^r\} = \begin{cases} \binom{2n}{n} P_r(n), & \text{if } r = 0, 2, 4, \dots, \\ 4^n P_r(n), & \text{if } r = 1, 3, 5, \dots, \end{cases}$$

TABLE 8

x	$W'(x)$	x	$W'(x)$
0.05	0.00000000	0.80	1.11391248
0.10	0.00000000	0.85	0.78816801
0.15	0.00000000	0.90	0.53444938
0.20	0.00000071	0.95	0.34809189
0.25	0.00113999	1.00	0.21811909
0.30	0.04549111	1.05	0.13165761
0.35	0.33036908	1.10	0.07662545
0.40	0.99108575	1.15	0.04303415
0.45	1.80425970	1.20	0.02333689
0.50	2.42954788	1.25	0.01222620
0.55	2.69798891	1.30	0.00619087
0.60	2.63013465	1.35	0.00303102
0.65	2.33507134	1.40	0.00143530
0.70	1.92967395	1.45	0.00065756
0.75	1.50379914	1.50	0.00029152

where $P_r(n)$ is a polynomial of degree $[3r/2]$ in n. In particular, we have $P_0(n) = 1$, $P_1(n) = n$,

(66) $$3P_2(n) = 10n^3 + 2n^2,$$

(67) $$4P_3(n) = 15n^4 + n^3,$$

(68) $$315P_4(n) = 4420n^6 + 332n^5 + 404n^4 - 116n^3$$

and

(69) $$32P_5(n) = 565n^7 - 165n^6 + 177n^5 - 75n^4 + 10n^3.$$

Furthermore, we obtain

(70) $$P_r(n) = M_r(8/\pi)^{r/2}\pi^{[r/2]}n^{[3r/2]} + \cdots$$

for $r = 0, 1, 2, \ldots$. In (70) only the leading term is displayed. The neglected terms have smaller order than the displayed one. In (70), M_r has the same meaning as in (44).

Similarly to Theorem 4 we have the following result.

THEOREM 8. *We have*

(71) $$\lim_{n\to\infty} E\left\{\left(\rho_n/\sqrt{2n}\right)^r\right\} = M_r$$

for $r = 0, 1, 2, \ldots$, *where* M_r *is determined by* (44) *and* (45).

PROOF. Since

(72) $$\binom{2n}{n}4^{-n} \sim 1/\sqrt{n\pi}$$

as $n \to \infty$, by (65) and (70) we obtain (71). □

THEOREM 9. *We have*

$$\lim_{n\to\infty} P\{\rho_n/\sqrt{2n} \le x\} = W(x) \tag{73}$$

at every continuity point of $W(x)$, *where* $W(x)$ *is the distribution function defined by* (49) *in Theorem* 5.

PROOF. The proof follows along the same lines as the proof of Theorem 5. □

We can easily prove that the finite-dimensional distributions of the process $\{\eta_n(t)/\sqrt{2n}\,;\ 0 \le t \le 1\}$ converge to the corresponding finite-dimensional distributions of a Gaussian process $\{\eta(t);\ 0 \le t \le 1\}$ for which $E\{\eta(t)\} = 0$ for $0 \le t \le 1$ and $E\{\eta(t)\eta(u)\} = t(1-u)$ for $0 \le t \le u \le 1$. The process $\{\eta(t);\ 0 \le t \le 1\}$ is called the Brownian bridge or the tied-down Brownian motion process. Define

$$\rho = \int_0^1 \eta(t)\,dt - \min_{0\le t\le 1} \eta(t). \tag{74}$$

THEOREM 10. *We have*

$$P\{\rho \le x\} = W(x), \tag{75}$$

where $W(x)$ *is defined by* (49).

PROOF. We can prove that the process $\{\eta_n(t)/\sqrt{2n}\,;\ 0 \le t \le 1\}$ converges weakly to the process $\{\eta(t);\ 0 \le t \le 1\}$. Since the integral (74) is a continuous functional on the process $\{\eta(t);\ 0 \le t \le 1\}$, (73) implies (75). □

By Theorems 6 and 10 we can draw the interesting conclusion that

$$P\{\rho \le x\} = P\{\omega^+ \le x\}, \tag{76}$$

that is, ρ and ω^+ have exactly the same distribution. For a direct proof of (76), we refer to Vervaat [16]. The variables ρ and ω^+ are defined for different processes. The random variable ρ is a functional on the Brownian bridge, while ω^+ is a functional on the Brownian excursion.

By (76) the problem of finding the distribution function of ω^+ can be reduced to the problem of finding the distribution function of ρ. But this is not a great advantage because ρ is a complicated functional on the Brownian bridge. However, as Darling [3] showed, ρ can also be expressed as

$$\rho = \max_{0\le t\le 1} \zeta(t), \tag{77}$$

where

$$\zeta(t) = \int_0^1 \eta(u)\,du - \eta(t). \tag{78}$$

In 1961, Watson [17] observed that $\{\zeta(t);\ 0 \le t \le 1\}$ is a Gaussian process for which $E\{\zeta(t)\} = 0$ for $0 \le t \le 1$ and

$$E\{\zeta(t)\zeta(u)\} = r(t-u) \tag{79}$$

for $0 \le t \le 1$ and $0 \le u \le 1$, where

$$r(t) = \frac{1}{2}\left(|t| - \frac{1}{2}\right)^2 - \frac{1}{24} \tag{80}$$

for $|t| \le 1$. By using the representation (77), Darling [3] proved that the Laplace transform of $P\{\rho \le x\}$ is given by (64). In 1984, Louchard [8] proved directly that the Laplace transform of $P\{\omega^+ \le x\}$ is given by (64) and he also calculated the moments of ω^+.

7. The asymptotic distribution of σ_n. We can determine the moments of $2n\sigma_n$ by (37). If we form the first r derivatives of (37) with respect to z at $z = 1$, we can determine the generating functions of the first r moments of $2n\sigma_n$. Thus we obtain

$$\binom{2n}{n} E\{(2n\sigma_n)^r\} = \begin{cases} \binom{2n}{n} T_r(n), & \text{if } r = 0,2,4,\dots, \\ 4^n T_r(n), & \text{if } r = 1,3,5,\dots, \end{cases} \tag{81}$$

where $T_r(n)$ is a polynomial of degree $[3r/2]$ in n. In particular, we have $T_0(n) = 1$, $T_1(n) = n/2$,

$$T_2(n) = \frac{14}{15}n^3 + \frac{1}{5}n^2 - \frac{2}{15}n, \tag{82}$$

$$T_3(n) = \frac{21}{32}n^4 + \frac{1}{16}n^3 - \frac{9}{32}n^2 + \frac{1}{16}n, \tag{83}$$

$$T_4(n) = \frac{76}{45}n^6 + \frac{8}{45}n^5 - \frac{409}{315}n^4 + \frac{88}{315}n^3 + \frac{8}{35}n^2 - \frac{8}{105}n \tag{84}$$

and

$$T_5(n) = \frac{101}{64}n^7 - \frac{7}{16}n^6 - \frac{29}{16}n^5 + \frac{35}{32}n^4 + \frac{39}{64}n^3 - \frac{21}{32}n^2 + \frac{1}{8}n. \tag{85}$$

Furthermore, we obtain

$$T_r(n) = M_r^* \pi^{[r/2]-r/2} 2^{3r/2} n^{[3r/2]} + \cdots \tag{86}$$

for $r = 0,1,2,\dots$. In (86) only the leading term is displayed. The neglected terms have smaller order than the displayed one. In (86),

$$M_r^* = D_r \frac{\sqrt{\pi}\, r!}{\Gamma\left(\frac{3r+1}{2}\right) 2^{r/2}} \tag{87}$$

for $r = 0, 1, 2, \ldots$, where $D_0 = 1$, $D_1 = 1/4$ and

$$D_r = \frac{(3r-2)}{4} D_{r-1} - \frac{1}{2} \sum_{i=1}^{r-1} D_i D_{r-i} \tag{88}$$

for $r \geq 2$.

The asymptotic behavior of the moments of $2n\sigma_n$ is given by the following theorem.

THEOREM 11. *We have*

$$\lim_{n \to \infty} E\left\{\left(\sigma_n/\sqrt{2n}\right)^r\right\} = M_r^* \tag{89}$$

for $r = 0, 1, 2, \ldots$, *where* M_r^* *is determined by* (87) *and* (88).

PROOF. By (72), (81) and (86) we obtain (89). □

THEOREM 12. *There exists a distribution function* $H(x)$ *such that*

$$\lim_{n \to \infty} P\left\{\sigma_n/\sqrt{2n} \leq x\right\} = H(x) \tag{90}$$

at every continuity point of $H(x)$. *The distribution function* $H(x)$ *is uniquely determined by its moments*

$$\int_0^\infty x^r \, dH(x) = M_r^* \tag{91}$$

for $r = 0, 1, 2, \ldots$, *where* M_r^* *is defined by* (87) *and* (88).

PROOF. By (88) we can prove that

$$\lim_{r \to \infty} (4/3)^r D_r/(r-1)! = 1/\pi. \tag{92}$$

This implies that

$$M_r^* \sim \sqrt{2}\left(\frac{r}{12e}\right)^{r/2} \tag{93}$$

as $r \to \infty$, and

$$\sum_{r=1}^{\infty} \frac{1}{(M_r^*)^{1/r}} = \infty. \tag{94}$$

The remaining part of the proof follows along the same lines as the proof of Theorem 5. □

Table 9 contains D_r and M_r^* for $r \leq 10$.

8. The distribution function $H(x)$. We have already mentioned that the process $\{\eta_n(t)/\sqrt{2n}\,;\ 0 \leq t \leq 1\}$ converges weakly to the Brownian bridge

TABLE 9

r	D_r	M_r^*
0	1	1
1	$\frac{1}{4}$	$\frac{1}{4}\sqrt{\frac{\pi}{2}}$
2	$\frac{7}{32}$	$\frac{7}{60}$
3	$\frac{21}{64}$	$\frac{21}{512}\sqrt{\frac{\pi}{2}}$
4	$\frac{1463}{2048}$	$\frac{19}{720}$
5	$\frac{2121}{1024}$	$\frac{101}{8192}\sqrt{\frac{\pi}{2}}$
6	$\frac{495271}{65536}$	$\frac{70753}{7001280}$
7	$\frac{136479}{4096}$	$\frac{45493}{7864320}\sqrt{\frac{\pi}{2}}$
8	$\frac{1445713003}{8388608}$	$\frac{206530429}{36714712320}$
9	$\frac{268122561}{262144}$	$\frac{89374187}{23991418880}\sqrt{\frac{\pi}{2}}$
10	$\frac{1838183317201}{268435456}$	$\frac{1256447927}{305663155200}$

$\{\eta(t); 0 \le t \le 1\}$. Let us define

$$\sigma = \int_0^1 |\eta(t)|\,dt. \tag{95}$$

THEOREM 13. *We have*

$$P\{\sigma \le x\} = H(x), \tag{96}$$

where $H(x)$ is defined by (91).

PROOF. Since σ is a continuous functional on the process $\{\eta(t); 0 \le t \le 1\}$, (90) implies (96). □

In 1982, Shepp [13] proved that

$$\int_0^\infty e^{-zs} E\{e^{-\sigma\sqrt{2}s^{3/2}}\} s^{-1/2}\,ds = -\sqrt{\pi}\,\mathrm{Ai}(z)/\mathrm{Ai}'(z) \tag{97}$$

for $\mathrm{Re}(z) > 0$, where $\mathrm{Ai}(z)$ is the Airy function defined by (60), and $\mathrm{Ai}'(z)$ is the derivative of (60). From (97) Shepp [13] derived a recurrence formula for the determination of the moments of σ and calculated $E\{\sigma^r\}$ for $r \le 5$. He did not

TABLE 10

x	$H(x)$	x	$H(x)$	x	$H(x)$
0.05	0.00000000	0.55	0.93414905	1.05	0.99971169
0.10	0.00056919	0.60	0.95725659	1.10	0.99985542
0.15	0.04345562	0.65	0.97285278	1.15	0.99992950
0.20	0.19556463	0.70	0.98315714	1.20	0.99996658
0.25	0.38791766	0.75	0.98980435	1.25	0.99998460
0.30	0.55745483	0.80	0.99398356	1.30	0.99999311
0.35	0.68822816	0.85	0.99654135	1.35	0.99999700
0.40	0.78389833	0.90	0.99806403	1.40	0.99999873
0.45	0.85228267	0.95	0.99894527	1.45	0.99999948
0.50	0.90052193	1.00	0.99944089	1.50	0.99999979

give details of his proof, but merely indicated that he used asymptotic expansions and let $z \to \infty$. By inverting (97), Rice [11] calculated the density function of σ by using numerical integration. An explicit expression for the distribution function $P\{\sigma \le x\}$ has been given by Johnson and Killeen [7]. By the results of Johnson and Killeen [7], we have the following theorem.

THEOREM 14. *We have*

$$H(x) = \frac{\sqrt{\pi}}{(18)^{1/6}x} \sum_{j=1}^{\infty} e^{-u_j} u_j^{-1/3} \operatorname{Ai}\left((3u_j/2)^{2/3}\right) \tag{98}$$

for $x > 0$, *where*

$$u_j = (a'_j)^3/(27x^2) \tag{99}$$

and $0 < a'_1 < a'_2 < \cdots < a'_j < \cdots$ *are defined such that* $z = -a'_j$, $j = 1, 2, \ldots$, *are the roots of* $\operatorname{Ai}'(z) = 0$.

Table 10 contains $H(x)$ for $0 < x \le 1.5$.

9. Applications in order statistics. There are several known statistics to test the hypothesis that the elements of two independent samples have a common distribution function. Here we consider two such statistics and find the asymptotic behavior of their moments and distributions.

Let $F_n(x)$ and $G_n(x)$ be the empirical distribution functions of two independent samples of size n in the case where the elements of the samples are independent random variables having a common continuous distribution function $V(x)$. Define Θ_n by

$$\Theta_n/n = \int_{-\infty}^{\infty} [F_n(x) - G_n(x)]\, dV(x) - \min_{-\infty<x<\infty} [F_n(x) - G_n(x)] \tag{100}$$

and Δ_n by

$$\Delta_n/n = \int_{-\infty}^{\infty} |F_n(x) - G_n(x)|\, dV(x). \tag{101}$$

Since $V(x)$ is a continuous distribution function, the distributions of Θ_n and Δ_n are independent of $V(x)$, that is, Θ_n and Δ_n are distribution-free statistics. Consequently, to find the distributions of Θ_n and Δ_n, we may assume without loss of generality that $V(x) = x$ for $0 \le x \le 1$, that is, the elements of the samples are independent random variables each having a uniform distribution over the interval $(0, 1)$. In this case (100) and (101) can be expressed in the following equivalent ways:

$$\Theta_n = \sum_{i=1}^{2n} \eta_i \xi_i + \delta_{2n} \tag{102}$$

and

$$\Delta_n = \sum_{i=1}^{2n} |\eta_i| \xi_i, \tag{103}$$

where $\{\eta_0, \eta_1, \dots, \eta_{2n}\}$ is the tied-down random walk defined in Section 3, δ_{2n} is defined by (20) and the random variables $\xi_0, \xi_1, \dots, \xi_{2n}$ are independent of the random walk $\{\eta_0, \eta_1, \dots, \eta_{2n}\}$ and defined in the following way: We choose $2n$ points at random in the interval $(0, 1)$. We assume that the $2n$ points are distributed independently and each point has a uniform distribution over the interval $(0, 1)$. These $2n$ points divide the interval $(0, 1)$ into the $2n + 1$ subintervals. Denote by $\xi_0, \xi_1, \dots, \xi_{2n}$ their lengths.

The random variables $\xi_0, \xi_1, \dots, \xi_{2n}$ are interchangeable and obviously $\xi_0 + \xi_1 + \cdots + \xi_{2n} = 1$. We have

$$E\{\xi_i\} = 1/(2n + 1) \tag{104}$$

and

$$E\{\xi_i^2\} = \frac{2}{(2n + 1)(2n + 2)} \tag{105}$$

for $i = 0, 1, \dots, 2n$. If $0 \le i < j \le 2n$, we have

$$E\{\xi_i \xi_j\} = \frac{1}{(2n + 1)(2n + 2)}. \tag{106}$$

Furthermore, we note that

$$P\{\eta_i = 2k - i\} = \binom{i}{k}\binom{2n - i}{n - k} \Big/ \binom{2n}{n} \tag{107}$$

for $0 \le k \le i \le 2n$ and

$$P\{\delta_{2n} \ge k\} = \binom{2n}{n - k} \Big/ \binom{2n}{n} \tag{108}$$

for $0 \le k \le 2n$.

The problem of finding the moments and the asymptotic distribution of Θ_n was proposed in the mid-1950s by researchers of the Hungarian State Railway. In response, Sarkadi [12] proved that

$$E\{\Theta_n\} = \frac{1}{2}\left[\frac{4^n}{\binom{2n}{n}} - 1\right] \tag{109}$$

and Takács [15] proved that

$$E\{\Theta_n^2\} = \frac{5n}{6} - E\{\Theta_n\}. \tag{110}$$

By the above results we can prove that $\Theta_n/\sqrt{2n}$ has the same limit distribution as $\rho_n/\sqrt{2n}$ as $n \to \infty$.

THEOREM 15. *We have*

$$\lim_{n\to\infty} P\left\{\frac{\Theta_n}{\sqrt{2n}} \le x\right\} = \lim_{n\to\infty} P\left\{\frac{\rho_n}{\sqrt{2n}} \le x\right\} = W(x), \tag{111}$$

where the distribution function $W(x)$ *is defined by* (49) *and is given explicitly by* (61).

PROOF. By (2) and (102) we have

$$\Theta_n - \rho_n = \sum_{i=1}^{2n} \eta_i\left(\xi_i - \frac{1}{2n}\right). \tag{112}$$

By the formulas mentioned above

$$E\{\Theta_n - \rho_n\} = 0 \tag{113}$$

and

$$E\{(\Theta_n - \rho_n)^2\} = 1/12. \tag{114}$$

Accordingly,

$$E\left\{\left(\frac{\Theta_n - \rho_n}{\sqrt{2n}}\right)^2\right\} = \frac{1}{24n} \to 0 \tag{115}$$

as $n \to \infty$. Thus

$$\lim_{n\to\infty}\left(\frac{\Theta_n - \rho_n}{\sqrt{2n}}\right) = 0 \tag{116}$$

in probability. This implies (111). □

The limit theorem (111) suggests, and indeed we can prove, that for every $r = 0, 1, 2, \ldots$ the moments $E\{\Theta_n^r\}$ and $E\{\rho_n^r\}$ show the same asymptotic behavior as $n \to \infty$.

THEOREM 16. *We have*

$$\lim_{n\to\infty} E\left\{\left(\frac{\Theta_n}{\sqrt{2n}}\right)^r\right\} = \lim_{n\to\infty} E\left\{\left(\frac{\rho_n}{\sqrt{2n}}\right)^r\right\} = M_r \tag{117}$$

for $r = 0, 1, 2, \dots$, *where* M_r *is defined by* (44) *and* (45).

PROOF. By making use of (71) and applying the Schwarz inequality, we can prove that

$$\lim_{n\to\infty} \left[E\{\Theta_n^r\} - E\{\rho_n^r\}\right]/n^{r/2} = 0 \tag{118}$$

for $r = 1, 2, \dots$. Hence (117) follows. □

In considering the statistic (101), by (103) we can prove that

$$E\{\Delta_n\} = \frac{n4^n}{2(2n+1)\binom{2n}{n}} \tag{119}$$

and

$$E\{\Delta_n^2\} = \frac{n(7n+3)}{30(n+1)}. \tag{120}$$

Moreover, $\Delta_n/\sqrt{2n}$ has the same limit distribution as $\sigma_n/\sqrt{2n}$ as $n \to \infty$.

THEOREM 17. *We have*

$$\lim_{n\to\infty} P\left\{\frac{\Delta_n}{\sqrt{2n}} \le x\right\} = \lim_{n\to\infty} P\left\{\frac{\sigma_n}{\sqrt{2n}} \le x\right\} = H(x), \tag{121}$$

where the distribution function $H(x)$ *is defined by* (91) *and is given explicitly by* (98).

PROOF. By (4) and (103) we have

$$\Delta_n - \sigma_n = \sum_{i=1}^{2n} |\eta_i|\left(\xi_i - \frac{1}{2n}\right). \tag{122}$$

Thus by (81),

$$E\{\Delta_n - \sigma_n\} = -\frac{4^{n-1}}{\binom{2n}{n}(2n+1)} \sim -\frac{\sqrt{\pi}}{8\sqrt{n}} \tag{123}$$

as $n \to \infty$, and

$$E\{(\Delta_n - \sigma_n)^2\} = \frac{3n^2+9n-2}{60n(n+1)} \to \frac{1}{20} \tag{124}$$

as $n \to \infty$. Accordingly,

$$E\left\{\left(\frac{\Delta_n - \sigma_n}{\sqrt{2n}}\right)^2\right\} = \frac{3n^2+9n-2}{120n^2(n+1)} \to 0 \tag{125}$$

as $n \to \infty$. Thus

$$\lim_{n\to\infty}\left(\frac{\Delta_n - \sigma_n}{\sqrt{2n}}\right) = 0 \tag{126}$$

in probability. This implies (121). □

Moreover, we have the following result.

THEOREM 18. *We have*

$$\lim_{n\to\infty} E\left\{\left(\frac{\Delta_n}{\sqrt{2n}}\right)^r\right\} = \lim_{n\to\infty} E\left\{\left(\frac{\sigma_n}{\sqrt{2n}}\right)^r\right\} = M_r^* \tag{127}$$

for $r = 0, 1, 2, \ldots$, *where* M_r^* *is defined by* (87) *and* (88).

PROOF. We can prove (127) by (89) and (122). □

Finally, we note that if we more generally assume that $F_m(x)$ and $G_n(x)$ are the empirical distribution functions of two independent samples of sizes m and n respectively, and the elements of the two samples are independent random variables each having the same continuous distribution function $V(x)$, then the statistic

$$\Theta_{m,n}^* = \sqrt{\frac{mn}{m+n}}\left\{\int_{-\infty}^{\infty}[F_m(x) - G_n(x)]\,dV(x) - \min_{-\infty<x<\infty}[F_m(x) - G_n(x)]\right\} \tag{128}$$

has the limit distribution function $W(x)$ as $m \to \infty$ and $n \to \infty$, and the statistic

$$\Delta_{m,n}^* = \sqrt{\frac{mn}{m+n}}\int_{-\infty}^{\infty}|F_m(x) - G_n(x)|\,dV(x) \tag{129}$$

has the limit distribution function $H(x)$ as $m \to \infty$ and $n \to \infty$.

I have used the remarkable program Mathematica of Wolfram [20] to calculate the tables in this paper.

REFERENCES

[1] ABRAMOWITZ, M. and STEGUN, I. A. (1965). *Handbook of Mathematical Functions*. Dover, New York.

[2] CARLEMAN, T. (1922). Sur le problème des moments. *C. R. Acad. Sci. Paris* **174** 1680–1682.

[3] DARLING, D. A. (1983). On the supremum of a certain Gaussian process. *Ann. Probab.* **11** 803–806.

[4] FRÉCHET, M. and SHOHAT, J. (1931). A proof of the generalized second-limit theorem in the theory of probability. *Trans. Amer. Soc.* **33** 533–543.

[5] GIKHMAN, I. I. and SKOROKHOD, A. V. (1969). *Introduction to the Theory of Random Processes*. W. B. Saunders, Philadelphia.

[6] HARDY, G. H. and WRIGHT, E. M. (1956). *An Introduction to the Theory of Numbers*, 3rd ed. Oxford Univ. Press.

[7] JOHNSON, B. McK. and KILLEEN, T. (1983). An explicit formula for the c.d.f. of the L_1 norm of the Brownian bridge. *Ann. Probab.* **11** 807–808.

[8] LOUCHARD, G. (1984). The Brownian excursion area: A numerical analysis. *Comput. Math. Appl.* **10** 413–417. [Erratum: (1986) **A12** 375.]

[9] MILLER, J. C. P. (1946). *The Airy Integral*. Cambridge Univ. Press, Cambridge.

[10] RAMANUJAN, S. (1919). Proof of certain identities in combinatorial analysis. *Proc. Cambridge Philos. Soc.* **19** 214–216.

[11] RICE, S. O. (1982). The integral of the absolute value of the pinned Wiener process—calculation of its probability density by numerical integration. *Ann. Probab.* **10** 240–243.

[12] SARKADI, K. (1954). Mozdonyok várakozási idejéröl (On the waiting time of locomotives). *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* **3** 191–194.

[13] SHEPP, L. A. (1982). On the integral of the absolute value of the pinned Wiener process. *Ann. Probab.* **10** 234–239.

[14] SLATER, L. J. (1960). *Confluent Hypergeometric Functions*. Cambridge Univ. Press.

[15] TAKÁCS, L. (1956). Egy közlekedéssel kapcsolatos valószinüségszámitási problémáról (On a probability problem connected with railway traffic). *Publ. Math. Inst. Hungar. Acad. Sci.* **1** 99–107.

[16] VERVAAT, W. (1979). A relation between Brownian bridge and Brownian excursion. *Ann. Probab.* **7** 143–149.

[17] WATSON, G. S. (1974). Goodness-of-fit test on a circle. *Biometrika* **48** 1777–1782.

[18] WHITWORTH, W. A. (1879). Arrangements of m things of one sort and n things of another sort, under certain conditions of priority. *Messenger Math.* **8** 105–114.

[19] WHITWORTH, W. A. (1886). *Choice and Chance* 4th ed. Deighton Bell, Cambridge.

[20] WOLFRAM, S. (1988). *Mathematica. A System for Doing Mathematics by Computer*. Addison-Wesley, Reading, Mass.

DEPARTMENT OF MATHEMATICS AND STATISTICS
CASE WESTERN RESERVE UNIVERSITY
CLEVELAND, OHIO 44106