Walter Katzenbeisser and Wolfgang Panny

Some further Results on the Height of Lattice Path

Paper

Some further Results on the Height of Lattice Path

Walter Katzenbeisser, Wolfgang Panny

Institut für Statistik
Wirtschaftsuniversität Wien

Forschungsberichte

Bericht 3
1990

http://statmath.wu-wien.ac.at/

SOME FURTHER RESULTS ON THE HEIGHT OF LATTICE PATHS

W. KATZENBEISSER and W. PANNY

Department of Statistics
University of Economics
Augasse 2–6
A–1090 Vienna
Austria

Abstract. This paper deals with the joint and conditional distributions concerning the maximum of random walk paths and the number of times this maximum is achieved. This joint distribution was studied first by Dwass [1967]. Based on his result, the correlation and some conditional moments are derived. The main contributions are however asymptotic expansions concerning the conditional distribution and conditional moments.

Keywords. Lattice paths, simple random walks, rank order statistics, asymptotic expansions.

AMS Subject Classification. 60J15, 41A60, 62G30.

Proposed running title. On the height of lattice paths.

SOME FURTHER RESULTS ON THE HEIGHT OF LATTICE PATHS

W. KATZENBEISSER and W. PANNY

University of Economics, Vienna

Introduction. Let $\mathbf{X}_k$, $k = 1, 2, \ldots$, be independent and identically distributed random variables with

$$\mathbf{P}(\mathbf{X}_k = 1) = \mathbf{P}(\mathbf{X}_k = -1) = 1/2.$$

Consider the simple random walk

$$\mathbf{S}_m = \sum_{k=1}^{m} \mathbf{X}_k,\ m = 1, 2, \ldots, 2n \quad \text{with} \quad \mathbf{S}_0 = 0 \quad \text{and} \quad \mathbf{S}_{2n} = 0\,,$$

i.e. a simple random walk starting at 0 and leading to 0 after $2n$ steps. There is a natural bijection between such random walks and lattice paths which helps to better visualize the situation and allows to use lattice paths combinatorics. A number of random variables can be associated to this random walk. Perhaps the most prominent amongst these is $\mathbf{D}_n^+$, the maximal (one-sided) height,

$$\mathbf{D}_n^+ = \max_{0 \le m \le 2n} \mathbf{S}_m$$

which essentially constitutes the test statistic of the one-sided Kolmogorov-Smirnov two-sample test with equal sample sizes. There is an extremely large literature on the Kolmogorov-Smirnov test; we only want to quote Kemperman [1959], Durbin [1973], Katzenbeisser and Panny [1984a], Panny [1984]. Alternative notions of height are discussed in Panny and Prodinger [1985]. Another random variable closely related to $\mathbf{D}_n^+$ is $\mathbf{Q}_n$,

$$\mathbf{Q}_n = [\text{number of times where the random walk reaches its maximum}].$$

The distribution of $\mathbf{Q}_n$ and the joint distribution of $\mathbf{D}_n^+$ and $\mathbf{Q}_n$ are treated (among other things) in Dwass [1967]. Some related results are included in Vincze [1957], Vincze [1959] and Mohanty [1968].

In this paper we consider the conditional distribution of $\mathbf{D}_n^+$ given $\mathbf{Q}_n$ and the conditional moments, taking Dwass's results as a starting point. Exact and asymptotic results will be derived, our main concern, however, is the analysis of the asymptotic behavior. As main tools for the asymptotic analysis we rely on the so-called generalized trinomial coefficients and the Γ-function method.

EXACT RESULTS

Probability distribution. From Dwass [1967, p.1047] it is well known that

$$\mathbf{P}(\mathbf{D}_n^+ \geq k,\, \mathbf{Q}_n \geq r) = \frac{\binom{2n-r+1}{n-k-r+1}}{\binom{2n}{n}}, \tag{1}$$

where $k = 0,1,2,\ldots,\ r = 1,2,\ldots,\ k+r \leq n+1$, and

$$\mathbf{P}(\mathbf{Q}_n \geq r) = \frac{\binom{2n-r+1}{n-r+1}}{\binom{2n}{n}}. \tag{2}$$

From (1) one easily obtains

$$\mathbf{P}(\mathbf{D}_n^+ = k,\, \mathbf{Q}_n = r) = \frac{2k+r-1}{2n-r+1}\frac{\binom{2n-r+1}{n+k}}{\binom{2n}{n}}. \tag{3}$$

Another representation for (3) reads

$$\mathbf{P}(\mathbf{D}_n^+ = k,\, \mathbf{Q}_n = r) = \frac{\binom{2n-r}{n+k-1} - \binom{2n-r}{n+k}}{\binom{2n}{n}}. \tag{4}$$

From (2) we get

$$\mathbf{P}(\mathbf{Q}_n = r) = \frac{\binom{2n-r}{n-1}}{\binom{2n}{n}}. \tag{5}$$

Let $p_{k|r}$ denote the conditional distribution of $\mathbf{D}_n^+$, given $\mathbf{Q}_n$, i.e.,

$$p_{k|r} = \mathbf{P}(\mathbf{D}_n^+ = k \mid \mathbf{Q}_n = r).$$

From (3) and (5) the following expression for $p_{k|r}$ can be derived:

$$p_{k|r} = \frac{2k+r-1}{2n-r+1}\frac{\binom{2n-r+1}{n+k}}{\binom{2n-r}{n-1}} = \frac{2k+r-1}{n}\frac{\binom{2n-r+1}{n+k}}{\binom{2n-r+1}{n}}. \tag{6}$$

Using (4) instead of (3) we get the alternative representation:

$$p_{k|r} = \frac{\binom{2n-r}{n+k-1} - \binom{2n-r}{n+k}}{\binom{2n-r}{n-1}}.$$

Moments. The s-th conditional moment can be expressed by

$$\mathbf{E}(\mathbf{D}_n^{+^s} \mid \mathbf{Q}_n = r) = \sum_{k\geq 0} k^s p_{k|r} = \frac{1}{n\binom{2n-\rho}{n}} \sum_{k\geq 0} (2k^{s+1} + \rho k^s)\binom{2n-\rho}{n+k}, \tag{7}$$

where (6) has been used and $\rho = r - 1$. Unfortunately, there exists no closed form for the sum (7) (cf. Graham, Knuth, Patashnik [1989, p.165]). However, the range of summation can considerably be reduced (namely from $0 \le k \le n-\rho$ to $0 \le j \le s+1$) if we use hypergeometric functions. We first observe that

$$\sum_{k\ge 0} [k]_s \binom{2n-\rho}{n+k} = \frac{(2n-\rho)!}{n!\,(n-\rho)!} \sum_{j\ge 0} \frac{(-n+\rho)_{s+j}(1)_{s+j}}{(n+1)_{s+j}} \frac{(-1)^{s+j}}{j!}$$
$$= s! \binom{2n-\rho}{n+s} {}_2F_1(-n+\rho+s, s+1; n+s+1; -1),$$

where $(k)_j = k(k+1)\cdots(k+j-1)$ and $[k]_j = k(k-1)\cdots(k-j+1)$. Converting from powers to factorials by means of Stirling numbers of the second kind $\left\{ {n \atop k} \right\}$, viz.

$$k^s = \sum_{j=0}^{s} \left\{ {s \atop j} \right\} [k]_j ,$$

we have

$$\sum_{k\ge 0} k^s \binom{2n-\rho}{n+k} = \sum_{j=0}^{s} \left\{ {s \atop j} \right\} j! \binom{2n-\rho}{n+j} {}_2F_1(-n+\rho+j, j+1; n+j+1; -1) .$$

Putting all together we obtain the following expression for $\mathbf{E}(\mathbf{D}_n^{+^s} \mid \mathbf{Q}_n = r)$:

$$\frac{1}{n\binom{2n-\rho}{n}} \sum_{j=0}^{s+1} \left(2\left\{ {s+1 \atop j} \right\} + \rho \left\{ {s \atop j} \right\} \right) j! \binom{2n-\rho}{n+j} {}_2F_1(-n+\rho+j, j+1; n+j+1; -1) .$$

Unfortunately, the last expression is not too useful from a numerical point of view since the hypergeometric functions must be computed in general (which essentially corresponds to the summation over all k in formula (7)). Only for some particular arguments special values are known, e.g. for $s = 0$, $\rho = 0, 1$ (cf. Abramowitz, Stegun [1972, p.557]).

The following results on the expectation and variance of $\mathbf{D}_n^+$ and $\mathbf{Q}_n$ and on their covariance are not directly used in the sequel. However, they may perhaps be helpful to view the problems from a wider perspective. It is well known (cf. Katzenbeisser, Panny [1984a, p.169]) that

$$\mathbf{E}(\mathbf{D}_n^+) = \frac{1}{2}\left(\frac{2^{2n}}{\binom{2n}{n}} - 1 \right), \qquad \mathbf{var}(\mathbf{D}_n^+) = \frac{1}{4}\left(4n + 1 - \frac{2^{4n}}{\binom{2n}{n}^2} \right).$$

Using (5) a straightforward computation yields

$$\mathbf{E}(\mathbf{Q}_n) = \frac{2n+1}{n+1}, \qquad \mathbf{var}(\mathbf{Q}_n) = \frac{n^2\,(2n+1)}{(n+1)^2\,(n+2)} .$$

From (4) we obtain after some manipulations the following closed form expression:

$$\mathbf{E}(\mathbf{D}_n^+\mathbf{Q}_n) = \frac{2^{2n}}{\binom{2n}{n}} - \frac{2n+1}{n+1}.$$

Hence

$$\mathbf{cov}(\mathbf{D}_n^+, \mathbf{Q}_n) = \frac{1}{2(n+1)}\left(\frac{2^{2n}}{\binom{2n}{n}} - 2n - 1\right)$$

and

$$\mathbf{corr}(\mathbf{D}_n^+, \mathbf{Q}_n) = -\frac{1}{n}\sqrt{\frac{n+2}{2n+1}}\,\frac{(2n+1)\binom{2n}{n}2^{-2n} - 1}{\sqrt{(4n+1)\binom{2n}{n}^2 2^{-4n} - 1}},$$

which shows that $\mathbf{D}_n^+$ and $\mathbf{Q}_n$ are asymptotically uncorrelated.

ASYMPTOTIC ANALYSIS

The exact formulae presented in the preceeding section may become rather tedious to compute for large n. Moreover, it is not too easy to get an impression of their behavior for various arguments. This is especially true for the conditional moments. The asymptotic approximations to be derived in the present section may be helpful to avoid these problems; they are easy to compute and reflect their dependency on their arguments in elementary terms. Moreover, it will be seen that they provide a numerical accuracy suitable for most practical purposes. Finally we would like to mention that the techniques applied in this section are not restricted to the present context but might prove useful for similar problems.

Probability distribution. Let us first introduce the so called generalized trinomial coefficients (see Katzenbeisser and Panny [1984b], Panny [1984]) defined by

$$\binom{n;\alpha,\beta,\gamma}{k} = [v^k](\alpha v^2 + \beta v + \gamma)^n, \tag{8}$$

where $[v^k]Q(v)$ denotes the coefficient of v^k in the power series $Q(v)$. They are connected to the ordinary trinomial coefficients by the relation

$$\binom{n;\alpha,\beta,\gamma}{n+k} = \sum_{\substack{a,b,c\geq 0\\ a+b+c=n\\ a-c=k}} \binom{n}{a,b,c}\alpha^a\beta^b\gamma^c,$$

which entails the following representation as a hypergeometric function:

$$\binom{n;\alpha,\beta,\gamma}{n+k} = \alpha^k\beta^{n-k}\binom{n}{k}\,{}_2F_1\!\left(-\frac{n-k}{2}, -\frac{n-k-1}{2}; k+1; \frac{4\alpha\gamma}{\beta^2}\right).$$

An integral representation is

$$\binom{n;\alpha,\beta,\gamma}{n+k} = \frac{(\alpha/\gamma)^{k/2}}{\pi}\int_0^\pi \cos kt\,(\beta + 2\sqrt{\alpha\gamma}\cos t)^n\,dt.$$

Clearly, the GTC's are related to ordinary binomial coefficients by

$$\binom{n;1/2,0,1/2}{2m} = \binom{n}{m}2^{-n}.$$

Theorem 3 of Panny [1984] gives asymptotic approximations to the GTC's for the symmetric case $\alpha = \gamma$. A weaker version, suitable for our purposes reads:

Lemma 1. *Let $N = \alpha n$, $d = 1$ if $\alpha < 1/2$, and $d = 2$ if $\alpha = 1/2$. Furthermore, let $n \equiv k(2)$ for the purely binomial case $\alpha = 1/2$. Then, the asymptotic behavior of the GTC's is*

$$\binom{n;\alpha,\beta,\alpha}{n+k} = \frac{d}{2\sqrt{N\pi}}\Big[1 + \frac{1-6\alpha}{192N}\big(\frac{k^4}{N^2} - 12\frac{k^2}{N} + 12\big)\Big]e^{-k^2/4N} + O(n^{-5/2})$$

as $n \to \infty$, the O-term holding uniformly for all k.

Lemma 1 can be obtained by an application of Laplace's method for integrals to the integral representation given above. For details the reader is refered to Panny [1984, p.15].

Using GTC's (6) can be expressed as:

$$p_{k|r} = \frac{2k+r-1}{n}\frac{\binom{2n-r+1;1/2,0,1/2}{(2n-r+1)+(2k+r-1)}}{\binom{2n-r+1;1/2,0,1/2}{(2n-r+1)+(r-1)}}. \tag{9}$$

An application of Lemma 1 to (9) furnishes the following

Theorem 1. *Let $r = O(n^{\frac{1+\epsilon}{2}})$, $\epsilon > 0$. Then the asymptotic behavior of $p_{k|r}$ can be described by:*

$$\begin{aligned} p_{k|r} &= \frac{2k+r-1}{n}\exp\Big\{-\frac{2k(k+r-1)}{2n-r+1}\Big\}(1+O(n^{-1+\epsilon})) \\ &= \frac{2k+r-1}{n}\Big(\exp\Big\{-\frac{2k(k+r-1)}{2n-r+1}\Big\} + O(n^{-1+\epsilon})\Big), \end{aligned} \tag{10}$$

as $n \to \infty$.

Theorem 1 follows from a straight forward application of Lemma 1. Perhaps it is helpful to consider the O-term in more detail. It is not hard to see that the binomial coefficients in (6) (and, of course, the GTC's in (9) also) have the property that the coefficient in the numerator is always less than the coefficient in the denominator except for $k = 0$, where they are equal. If we restrict the range of r to $r = O(n^{\frac{1+\epsilon}{2}})$ the quotient of the GTC's becomes

$$\frac{e^{-x} + O(n^{-1})}{(1+O(n^{-1+\epsilon}))e^{-y}} \tag{11}$$

by an application of Lemma 1, where $x = (2k+r-1)^2/(2(2n-r+1))$ and $y = (r-1)^2/(2(2n-r+1))$. Consequently, the relative error of the denominator is at most $1+O(n^{-1+\epsilon})$. After division the essential part of the numerator becomes

$$(1+O(n^{-1+\epsilon}))e^{-x}e^{y}$$

which equals

$$e^{-x+y}+O(n^{-1+\epsilon}), \tag{12}$$

since $x \geq y$. For the error term in the numerator it suffices to observe that it can not increase after division because the quotient (11) is at most equal to 1. Consequently, the resulting error E is at most $O(n^{-1})$ which is covered by the O-term in (12). It should perhaps be mentioned that for very large values of k it may happen that E dominates the whole expression (12), producing a considerable *relative* error. However, for such values of k expression (12) is exponentially small. Accordingly the *absolute* error is of the same order for that case.

Tables 1 and 2 show exact and approximate values for $p_{k|r}$, where for each k the first (second) row corresponds to exact (approximate) values. According to Theorem 1 the range of r is restricted to $r \leq n^{0.7}$, which corresponds to a relatively large value of ϵ. It can be seen from these tables that even for small values of n (and large values of r) the approximation shows a good accuracy.

Table 1 $\quad \mathbf{P}(D_n^+ = k \mid \mathbf{Q}_n = r), \quad n=10$

$k\backslash r$	1	2	3	4	5
0	0.000	0.100	0.200	0.300	0.400
	0.000	0.100	0.200	0.300	0.400
1	0.182	0.245	0.291	0.318	0.327
	0.181	0.243	0.286	0.312	0.321
2	0.273	0.273	0.255	0.223	0.182
	0.268	0.266	0.247	0.216	0.179
3	0.252	0.206	0.157	0.110	0.070
	0.244	0.198	0.151	0.108	0.072
4	0.168	0.113	0.070	0.038	0.018
	0.162	0.110	0.070	0.041	0.022
5	0.084	0.046	0.022	0.009	0.003
	0.082	0.047	0.025	0.012	0.005
6	0.031	0.014	0.005	0.001	0.000
	0.033	0.016	0.007	0.003	0.001
7	0.009	0.003	0.001	0.000	
	0.010	0.004	0.001	0.000	
8	0.002	0.001	0.000		
	0.003	0.001	0.000		
9	0.000	0.000			
	0.001	0.000			
10	0.000				
	0.000				

Table 2 $\mathbf{P}(D_n^+ = k \mid \mathbf{Q}_n = r), \quad n = 50$

$k \backslash r$	1	2	3	4	5	10	15
0	0.000 0.000	0.020 0.020	0.040 0.040	0.060 0.060	0.080 0.080	0.180 0.180	0.280 0.280
1	0.039 0.039	0.058 0.058	0.075 0.075	0.092 0.092	0.108 0.108	0.177 0.177	0.226 0.226
2	0.074 0.074	0.089 0.089	0.102 0.102	0.114 0.114	0.125 0.125	0.161 0.160	0.171 0.171
3	0.100 0.100	0.110 0.110	0.118 0.118	0.125 0.124	0.130 0.129	0.137 0.136	0.122 0.122
4	0.117 0.116	0.121 0.120	0.123 0.123	0.124 0.124	0.124 0.123	0.109 0.108	0.082 0.082
5	0.122 0.121	0.121 0.120	0.118 0.117	0.115 0.114	0.110 0.110	0.082 0.082	0.052 0.053
6	0.117 0.117	0.112 0.111	0.106 0.105	0.099 0.099	0.092 0.092	0.058 0.058	0.031 0.032
7	0.106 0.105	0.097 0.097	0.089 0.088	0.081 0.080	0.073 0.072	0.039 0.039	0.018 0.018
8	0.090 0.089	0.080 0.079	0.071 0.070	0.062 0.062	0.054 0.054	0.025 0.025	0.009 0.010
9	0.072 0.071	0.062 0.062	0.053 0.053	0.046 0.045	0.039 0.038	0.015 0.015	0.005 0.005
10	0.054 0.054	0.046 0.046	0.038 0.038	0.032 0.032	0.026 0.026	0.009 0.009	0.002 0.003
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
15	0.007 0.007	0.005 0.005	0.003 0.004	0.002 0.003	0.002 0.002	0.000 0.000	0.000 0.000

For $k = O(n^{\frac{1+\epsilon}{2}})$, r fixed, the following version of (10) can be derived:

$$\begin{aligned} p_{k|r} &= \frac{1}{N+\rho/2}\Big(2k + \rho - \frac{2\rho}{N}k^2 + \frac{\rho^2}{N^2}k^3 - \frac{\rho^2}{N}k\Big)e^{-\frac{k^2}{N}}\left(1 + O(N^{-1})\right) \\ &= \frac{1}{N}\Big(2k + \rho - \frac{2\rho}{N}k^2 + \frac{\rho^2}{N^2}k^3 - \frac{\rho^2}{N}k - \frac{\rho}{N}k\Big)e^{-\frac{k^2}{N}}\left(1 + O(N^{-1})\right), \end{aligned} \tag{13}$$

as $N \to \infty$, where $\rho = r - 1$ and $N = n - \rho/2$. This version can be derived by writing the exponential function in (10) as a product, viz. $\exp\{-k^2/N\}\exp\{-k\rho/N\}$, and replacing the second factor by its Maclaurin expansion. Version (13) will prove useful for the asymptotics of the conditional moments. Notice that the restriction on k and r is not unreasonable because for other values the joint tail probability (1) gets exponentially small.

Table 2 $\mathbf{P}(D_n^+ = k \,|\, \mathbf{Q}_n = r), \quad n = 50$

$k \backslash r$	1	2	3	4	5	10	15
0	0.000	0.020	0.040	0.060	0.080	0.180	0.280
	0.000	0.020	0.040	0.060	0.080	0.180	0.280
1	0.039	0.058	0.075	0.092	0.108	0.177	0.226
	0.039	0.058	0.075	0.092	0.108	0.177	0.226
2	0.074	0.089	0.102	0.114	0.125	0.161	0.171
	0.074	0.089	0.102	0.114	0.125	0.160	0.171
3	0.100	0.110	0.118	0.125	0.130	0.137	0.122
	0.100	0.110	0.118	0.124	0.129	0.136	0.122
4	0.117	0.121	0.123	0.124	0.124	0.109	0.082
	0.116	0.120	0.123	0.124	0.123	0.108	0.082
5	0.122	0.121	0.118	0.115	0.110	0.082	0.052
	0.121	0.120	0.117	0.114	0.110	0.082	0.053
6	0.117	0.112	0.106	0.099	0.092	0.058	0.031
	0.117	0.111	0.105	0.099	0.092	0.058	0.032
7	0.106	0.097	0.089	0.081	0.073	0.039	0.018
	0.105	0.097	0.088	0.080	0.072	0.039	0.018
8	0.090	0.080	0.071	0.062	0.054	0.025	0.009
	0.089	0.079	0.070	0.062	0.054	0.025	0.010
9	0.072	0.062	0.053	0.046	0.039	0.015	0.005
	0.071	0.062	0.053	0.045	0.038	0.015	0.005
10	0.054	0.046	0.038	0.032	0.026	0.009	0.002
	0.054	0.046	0.038	0.032	0.026	0.009	0.003
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
15	0.007	0.005	0.003	0.002	0.002	0.000	0.000
	0.007	0.005	0.004	0.003	0.002	0.000	0.000

For $k = O(n^{\frac{1+\epsilon}{2}})$, r fixed, the following version of (10) can be derived:

$$\begin{aligned} p_{k|r} &= \frac{1}{N+\rho/2}\Big(2k+\rho-\frac{2\rho}{N}k^2+\frac{\rho^2}{N^2}k^3-\frac{\rho^2}{N}k\Big)e^{-\frac{k^2}{N}}\big(1+O(N^{-1})\big) \\ &= \frac{1}{N}\Big(2k+\rho-\frac{2\rho}{N}k^2+\frac{\rho^2}{N^2}k^3-\frac{\rho^2}{N}k-\frac{\rho}{N}k\Big)e^{-\frac{k^2}{N}}\big(1+O(N^{-1})\big), \end{aligned} \tag{13}$$

as $N \to \infty$, where $\rho = r-1$ and $N = n-\rho/2$. This version can be derived by writing the exponential function in (10) as a product, viz. $\exp\{-k^2/N\}\exp\{-k\rho/N\}$, and replacing the second factor by its Maclaurin expansion. Version (13) will prove useful for the asymptotics of the conditional moments. Notice that the restriction on k and r is not unreasonable because for other values the joint tail probability (1) gets exponentially small.

Moments. Let us introduce the $f_s(N)$-functions, defined by

$$f_s(N) = \sum_{k\geq 1} k^s e^{-\frac{k^2}{N}}.$$

Applying them to (13), the s-th conditional moment of $\mathbf{D}_n^+$, given $\mathbf{Q}_n$ can be expressed as:

$$\frac{1}{N+\rho/2}\Big(2f_{s+1} + \rho\big(f_s - \frac{2}{N}f_{s+2}\big)\Big)\left(1+O(N^{-1})\right),$$

where the arguments of the $f_s(N)$-functions have been omitted for convenience. Using the $f_s(N)$-functions we have extended the range of summation for k beyond the range of validity of (13). But it is not hard to see that this produces only an exponentially small error, which is covered by the given O-term.

Asymptotic equivalents for the $f_s(N)$-functions can conveniently be derived by means of the so called Γ-function method (cf. Knuth [1973, p.131]). This has been done in Panny [1984, p.72]. They could also be obtained by an application of Euler's summation formula. The following Lemma gives the resulting asymptotic expansions for the $f_s(N)$-functions.

Lemma 2. *Let $s = 0, 1, 2, \ldots$. Then we have*

$$f_s(N) = \frac{1}{2}\Gamma\Big(\frac{s+1}{2}\Big)N^{\frac{s+1}{2}} + \sum_{k=0}^{m-1}\frac{(-1)^k}{k!}\zeta(-2k-s)N^{-k} + O(N^{-m}),$$

for all $m > 0$ as $N \to \infty$.

Hence the asymptotic behavior of the conditional moments of $\mathbf{D}_n^+$ can be given by:

$$\mathbf{E}(\mathbf{D}_n^{+\,s} \mid \mathbf{Q}_n = r) = \frac{1}{N+\rho/2}\Big(\Gamma\big(\frac{s+2}{2}\big)N^{\frac{s+2}{2}} - \frac{\rho s}{2}\Gamma\big(\frac{s+1}{2}\big)N^{\frac{s+1}{2}}\Big)\left(1+O(N^{-1})\right). \quad (14)$$

Using stronger versions of (13) this method allows to get further terms of the asymptotic series (14). In principle it could be extended to achieve an O-term as small as we please. Considering the next term in (14) and reverting back to the actual quantity n one obtains the following result, summarized in

Theorem 2. *Let $\rho = r - 1$. Then we have*

$$\begin{aligned}\mathbf{E}(\mathbf{D}_n^{+\,s} \mid \mathbf{Q}_n = r) = {} & \Gamma\big(\frac{s+2}{2}\big)n^{\frac{s}{2}} \\ & - \frac{\rho s}{2}\Gamma\big(\frac{s+1}{2}\big)n^{\frac{s-1}{2}} \\ & - \frac{(s+4)(s+2) - 6(1-\rho+\rho^2)(s+2) + 12\rho^2}{24}\Gamma\big(\frac{s+2}{2}\big)n^{\frac{s-2}{2}} \\ & + O(n^{\frac{s-3}{2}}) \quad \text{as } n \to \infty.\end{aligned}$$

Specializing on $s = 1, 2$ we have

$$\mathbf{E}(\mathbf{D}_n^+ \,|\, \mathbf{Q}_n = r) = \frac{\sqrt{\pi}}{2} n^{1/2} - \frac{\rho}{2} + \frac{2\rho^2 - 6\rho + 1}{16} \sqrt{\pi} n^{-1/2} + O(n^{-1}),$$
$$\mathbf{E}(\mathbf{D}_n^{+2} \,|\, \mathbf{Q}_n = r) = n - \frac{\rho\sqrt{\pi}}{2} n^{1/2} + \frac{\rho^2 - 2\rho}{2} + O(n^{-1/2}),$$
$$\mathbf{var}(\mathbf{D}_n^+ \,|\, \mathbf{Q}_n = r) = \frac{4-\pi}{4} n + \frac{\rho^2 - 4\rho}{4} - \frac{2\rho^2 - 6\rho + 1}{16} \pi + O(n^{-1/2}).$$

In the following tables exact and approximate values for the conditional expectation and variance of $\mathbf{D}_n^+$, given $\mathbf{Q}_n$ are shown. For each value of r, the first (second) row gives the exact (approximate) values.

Table 3 $\mathbf{E}(\mathbf{D}_n^+ \,|\, \mathbf{Q}_n = r)$

$r \backslash n$	10	20	50	100	500	1000
1	2.838	3.988	6.282	8.873	19.822	28.028
	2.838	3.988	6.282	8.873	19.822	28.028
2	2.196	3.388	5.719	8.329	19.302	27.514
	2.197	3.389	5.720	8.329	19.302	27.514
3	1.696	2.888	5.219	7.829	18.802	27.014
	1.697	2.889	5.220	7.829	18.802	27.014
4	1.302	2.469	4.774	7.369	18.321	26.528
	1.338	2.488	4.782	7.373	18.322	26.528
5	0.988	2.114	4.375	6.945	17.858	26.055
	1.118	2.186	4.408	6.962	17.861	26.056

Table 4 $\mathbf{var}(\mathbf{D}_n^+ \,|\, \mathbf{Q}_n = r)$

$r \backslash n$	10	20	50	100	500	1000
1	1.95	4.09	10.53	21.26	107.10	214.41
	1.95	4.10	10.53	21.26	107.10	214.41
2	1.98	4.13	10.57	21.30	107.14	214.44
	1.99	4.13	10.57	21.30	107.14	214.44
3	1.73	3.88	10.32	21.05	106.89	214.19
	1.74	3.88	10.32	21.05	106.89	214.19
4	1.40	3.50	9.89	20.59	106.39	213.68
	1.20	3.35	9.78	20.51	106.35	213.66
5	1.07	3.08	9.36	19.99	105.68	212.94
	0.38	2.53	8.96	19.69	105.53	212.83

References

ABRAMOWITZ, M., STEGUN, I. [1972] *Handbook of Mathematical Functions*, 10th printing, National Bureau of Standards.

DURBIN, J. [1973] *Distribution Theory for Tests Based on the Sample Distribution Function*, Regional Conference Series in Applied Mathematics 9, SIAM, Philadelphia/Pennsylvania.

DWASS, M. [1967] Simple random walk and rank order statistics, *Ann. Math. Statist.* **38**, 1042–1053.

GRAHAM, R. L., KNUTH, D. E., PATASHNIK, O. [1989] *Concrete Mathematics*, Addison-Wesley, Reading/Mass.

KATZENBEISSER, W., PANNY, W. [1984a] The moments of the teststatistics $D_{n,n}$ and $D^+_{n,n}$ of the Kolmogorov-Smirnov test and their asymptotic behavior, in: *Proc. of the 4th Pannonian Symp. on Math. Stat.*, W. Grossmann, G. Pflug, I. Vincze, W. Wertz, eds., 167–183; Akadémiai Kiadó, Budapest.

KATZENBEISSER, W., PANNY, W. [1984b] Asymptotic results on the maximal deviation of simple random walks, *Stochastic Processes Appl.* **18**, 263–275.

KEMPERMAN, J. H. B. [1959] Asymptotic expansions for the Smirnov test and for the range of cumulative sums, *Ann. Math. Statist.* **30**, 448–462.

KNUTH, D. E. [1973] *The Art of Computer Programming*, Vol.3, Addison-Wesley, Reading/Mass.

MOHANTY, S. G. [1968] On some generalizations of a restricted random walk, *Studia Scientiarum Mathematicarum Hungarica* **3**, 225–241.

PANNY, W. [1984] *The Maximal Deviation of Lattice Paths*, Athenäum/Hain/Hanstein, Königstein/Ts.

PANNY, W., PRODINGER, H. [1985] The expected height of paths for several notions of height, *Studia Scientiarum Mathematicarum Hungarica* **20**, 119–132.

VINCZE, I. [1957] Einige zweidimensionale Verteilungs- und Grenzverteilungssätze in der Theorie der geordneten Stichproben, *Publ. Math. Inst. Hungarian Acad. Sci* **2**, 183–203.

VINCZE, I. [1959] On some joint distributions and some joint limiting distributions in the theory of order statistics, II; *Publ. Math. Inst. Hungarian Acad. Sci* **4**, 29–41.

Institut für Statistik
Wirtschaftsuniversität Wien

Augasse 2–6
A–1090 Vienna
Austria