Simple Random Walk Statistics. Part I: Discrete Time Results

Walter Katzenbeisser, Wolfgang Panny

Institut für Statistik
Wirtschaftsuniversität Wien

Forschungsberichte

Bericht 1
1990

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

SIMPLE RANDOM WALK STATISTICS
PART I: DISCRETE TIME RESULTS

W. KATZENBEISSER and W. PANNY

University of Economics, Vienna

In a famous paper Dwass [1967] proposed a method to deal with rank order statistics, which constitutes a unifying framework to derive various distributional results. In the present paper an alternative method is presented, which allows to extend Dwass's results in several ways, viz. arbitrary endpoints, horizontal steps, and arbitrary probabilities for the three step types. Regarding these extensions the pertaining rank order statistics are extended as well to simple random walk statistics. This method has proved appropriate to generalize all results given by Dwass. Moreover, these discrete time results can be taken as a starting point to derive the corresponding results for randomized random walks by means of a limiting process.

Introduction. In the literature on rank order statistics two generic approaches for determining distributional results are generally distinguished, namely Gnedenko's method and Dwass's method (cf. Mohanty [1979]). Gnedenko's method is essentially of combinatorial nature and involves a random walk consisting of *nonindependent* steps in the sense that the number of +1's and −1's is given in advance, i.e. the endpoint is fixed (in most cases $(2n, 0)$); the main tools are various path counting techniques. In contrast, the basis of the Dwass method is the consideration of a transient random walk ($p < 1/2$) with *independent* steps. On that basis one often conveniently obtains an expression $h(p)$ for the probability of the corresponding event in the unrestricted random walk, where $h(p)$ reflects the dependency of this probability on p. Technically, the distribution of the interesting rank order statistic can be obtained by expanding $h(p)/(1-2p)$ in powers of pq. For details of both methods the reader is refered to Šidak [1973], Mohanty [1979], Aneja and Kanwar Sen [1972], and Dwass [1967] of course.

In this paper we directly approach the p.g.f. for the interesting rank order statistic where *nonindependent* steps in the underlying random walk are considered in the sense that the difference ℓ between the number of +1's and −1's is fixed, i.e. for all endpoints (n, ℓ), $n = 0, 1, \ldots$ simultaneously, ℓ fixed. It turns out that the p.g.f.'s for all statistics considered in Dwass [1967] can be derived from a single basic p.g.f., say $\Psi_{h,m,\ell}(y)$. Once the interesting p.g.f. has been obtained — essentially by convoluting appropriate versions of $\Psi_{h,m,\ell}(y)$, which corresponds to the concatenation of respective segments in the underlying paths — one only has to extract the coefficients in the resulting p.g.f., which is more or less a routine task. Our method allows to generalize Dwass's results in several ways. First, we consider arbitrary endpoints (n, ℓ), permitting to deal with rank order statistics for unequal sample sizes also. Secondly, we introduce horizontal steps as third step type. Moreover, we allow arbitrary probabilities α, β, γ for the three step types. Regarding these extensions the pertaining rank order statistics are extended as well to "simple random walk statistics", where the term simple random walk is used in the sense of Cox and Miller [1965]. In part II it will be shown how the pertaining results can be translated to the corresponding results for randomized random walks by means of a limiting process.

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

$$\mathbf{P}(\mathbf{X}_k = 1) = \alpha, \quad \mathbf{P}(\mathbf{X}_k = 0) = \beta, \quad \mathbf{P}(\mathbf{X}_k = -1) = \gamma,$$

where $\alpha + \beta + \gamma = 1$. Consider the random walk

$$\mathbf{S}_k = \mathbf{S}_0 + \sum_{j=1}^{k} \mathbf{X}_j, \quad k = 1, 2, \ldots, n \quad \text{with} \quad \mathbf{S}_n = \ell,$$

i.e. a simple random walk in the sense of Cox and Miller [1965] starting at $\mathbf{S}_0$ and leading to ℓ after n steps. Confining to $\mathbf{S}_0 = 0$ actually constitutes no restriction at all. So this assumption will be made in the sequel if not explicitly stated otherwise. For $\alpha = \gamma = 1/2$ and $\mathbf{S}_0 = 0$ a sample path of $\mathbf{S}_{n_1+n_2} = n_1 - n_2$, $\mathbf{S}_0 = 0$ corresponds to the graph of the rank order indicators, where the corresponding sample sizes are n_1 and n_2, respectively. In Dwass [1967] a number of rank order statistics are compiled and their probability distributions are given for the special case $n_1 = n_2$ which is equivalent to the condition $\ell = 0$.

All results contained in this paper are based on the probability generating function $\Psi_{h,m,\ell}(y)$,

$$\Psi_{h,m,\ell}(y) = \sum_{n \geq 0} p(h, m, \ell, n) y^n, \quad h, m \geq 0.$$

$p(h, m, \ell, n)$ gives the probability that a particle obeying a random walk with absorbing barriers at $-m$ and h reaches the state ℓ when it started from state 0, i.e.

$$p(h, m, \ell, n) = \mathbf{P}(-m \leq \mathbf{S}_1 \leq h, \ldots, -m \leq \mathbf{S}_{n-1} \leq h, -m \leq \mathbf{S}_n = \ell \leq h \,|\, \mathbf{S}_0 = 0).$$

The definition of $p(h, m, \ell, n)$ shows that touching the barriers is admissible and the particle is absorbed only if it crosses them. Following Barton and Mallows [1965] this type of absorption could be termed weak sense absorption. It has been shown (cf. Panny [1984], Katzenbeisser and Panny [1984]) that

$$\Psi_{h,m,\ell}(y) = \frac{\rho^{\frac{|\ell|+\ell}{2}}}{\gamma} v^{|\ell|} \frac{\alpha v^2 + \beta v + \gamma}{1 - \rho v^2} \frac{\left[1 - (\rho v^2)^{m+1-\frac{|\ell|-\ell}{2}}\right]\left[1 - (\rho v^2)^{h+1-\frac{|\ell|+\ell}{2}}\right]}{1 - (\rho v^2)^{h+m+2}}, \quad (1)$$

where $\rho=\alpha/\gamma$ and the substitution $y=g(v)=v/(\alpha v^2+\beta v+\gamma)$ has been used. This substitution is crucial for our approach because it considerably simplifies the original generating function in terms of y. The generating function also comprises the one-sided cases, viz.

$$\Psi_{\infty,m,\ell}(y)=\frac{\rho^{\frac{|\ell|+\ell}{2}}}{\gamma}v^{|\ell|}(\alpha v^2+\beta v+\gamma)\frac{1-(\rho v^2)^{m+1-\frac{|\ell|-\ell}{2}}}{1-\rho v^2}, \tag{2}$$

$$\Psi_{h,\infty,\ell}(y)=\frac{\rho^{\frac{|\ell|+\ell}{2}}}{\gamma}v^{|\ell|}(\alpha v^2+\beta v+\gamma)\frac{1-(\rho v^2)^{h+1-\frac{|\ell|+\ell}{2}}}{1-\rho v^2}, \tag{3}$$

and the unrestricted case

$$\Psi_{\infty,\infty,\ell}(y)=\frac{\rho^{\frac{|\ell|+\ell}{2}}}{\gamma}v^{|\ell|}\frac{\alpha v^2+\beta v+\gamma}{1-\rho v^2}. \tag{4}$$

Explicit expressions for the probabilities $p(h,m,\ell,n)$ can be derived by an application of Cauchy's integral formula,

$$p(h,m,\ell,n)=\frac{1}{2\pi i}\oint\frac{\Psi_{h,m,\ell}(y)}{y^{n+1}}dy. \tag{5}$$

Since $y=v/\gamma+O(v^2)$ as $v\to 0$, and $g(v)$ is analytic in a sufficiently small neighborhood of 0, (5) remains valid also after substituting $y=g(v)$. Hence

$$p(h,m,\ell,n)=\frac{1}{2\pi i}\oint\frac{\Psi_{h,m,\ell}(g(v))}{g^{n+1}(v)}g'(v)dv,$$

where

$$\frac{g'(v)}{g^{n+1}(v)}=\gamma\frac{1-\rho v^2}{v^{n+1}}(\alpha v^2+\beta v+\gamma)^{n-1}.$$

Consequently

$$p(h,m,\ell,n)=\sum_{j=0,\pm1,\ldots}\rho^{-jD}\left[\binom{n;\alpha,\beta,\gamma}{n+\ell+2jD}-\rho^{h+1}\binom{n;\alpha,\beta,\gamma}{n+\ell-2(h+1)+2jD}\right], \tag{1'}$$

$$p(\infty,m,\ell,n)=\binom{n;\alpha,\beta,\gamma}{n+\ell}-\rho^{-(m+1)}\binom{n;\alpha,\beta,\gamma}{n+\ell+2(m+1)}, \tag{2'}$$

$$p(h,\infty,\ell,n)=\binom{n;\alpha,\beta,\gamma}{n+\ell}-\rho^{h+1}\binom{n;\alpha,\beta,\gamma}{n+\ell-2(h+1)}, \tag{3'}$$

$$p(\infty,\infty,\ell,n)=\binom{n;\alpha,\beta,\gamma}{n+\ell}, \tag{4'}$$

where $D=h+m+2$ and where generalized trinomial coefficients (GTC's) $\binom{n;\alpha,\beta,\gamma}{k}$ are used. They have the generating function $(\alpha v^2+\beta v+\gamma)^n$, i.e.

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

and $[v^k]P(v)$ denotes the coefficient of v^k in $P(v)$. GTC's are quasi-symmetric, i.e.

$$\binom{n;\alpha,\beta,\gamma}{n-k} = \rho^{-k}\binom{n;\alpha,\beta,\gamma}{n+k}$$

and comprise binomial coefficients as a special case, viz.

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

They are connected to 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:

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

An integral representation is

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

which can easily be verified by the residue theorem.

List of random variables. The following list contains all rank order statistics considered by Dwass. Regarding the possibility of horizontal steps the pertaining definitions had to be adapted, bearing in mind also their suitability for randomized random walks. However, in the absence of horizontal steps (i.e. $\beta = 0$) they are equivalent to Dwass's definitions. Fig.1 illustrates the definitions in terms of a random walk diagram.

I. N_n, *the number of visits to zero.* A *visit to zero* occurs if $\mathrm{S}_k = \mathrm{S}_{k+1} = \mathrm{S}_{k+2} = \ldots = \mathrm{S}_{k+m} = 0$, and $\mathrm{S}_{k-1} \neq 0$, $\mathrm{S}_{k+m+1} \neq 0$ for $0 \leq k \leq k+m \leq n$. If $\mathrm{S}_0 = 0$ then a *visit to zero* begins at the origin by definition. Correspondingly, if $\mathrm{S}_n = 0$ then a *visit to zero* terminates at the end-point by definition. This may be summarized by saying that if there should be one or more consecutive horizontal steps coinciding with the x-axis (i.e. $m > 0$), this counts only as a single *visit to zero.*

II. $\mathrm{N}_n^+, \mathrm{N}_n^-$, *the number of positive and negative sojourns.* A *positive (negative) sojourn* occurs, if $\mathrm{S}_k, \mathrm{S}_{k+1}, \mathrm{S}_{k+2}, \ldots, \mathrm{S}_{k+m} > 0\,(< 0)$, and $\mathrm{S}_{k-1} = 0$, $\mathrm{S}_{k+m+1} = 0$ for $0 < k \leq k+m < n$.

III. $\mathrm{N}_n(r)$, *the number of visits to* r. A *visit to* r occurs if $\mathrm{S}_k = \mathrm{S}_{k+1} = \mathrm{S}_{k+2} = \ldots = \mathrm{S}_{k+m} = r$, and $\mathrm{S}_{k-1} \neq r$, $\mathrm{S}_{k+m+1} \neq r$ for $0 \leq k \leq k+m \leq n$. If $\mathrm{S}_0 = r$

then a *visit to r* begins at the origin by definition. Correspondingly, if $\mathbf{S}_n = r$ then a *visit to r* terminates at the end-point by definition. This may be summarized by saying that if there should be one or more consecutive horizontal steps coinciding with the line $y = r$ (i.e. $m > 0$), this counts only as a single *visit to r*.

IV. $\mathbf{N}_n^+(r)$, *the number of positive sojourns with respect to r*. A *positive sojourn with respect to r* occurs, if $\mathbf{S}_k, \mathbf{S}_{k+1}, \mathbf{S}_{k+2}, \ldots, \mathbf{S}_{k+m} > r$ and $\mathbf{S}_{k-1} = r$, $\mathbf{S}_{k+m+1} = r$ for $0 < k \leq k + m < n$.

V. $\mathbf{N}_n^*(r)$, *the number of crossings of r*. A *crossing of r* occurs if $\mathbf{S}_{k-1} = r \mp 1$, $\mathbf{S}_k = \mathbf{S}_{k+1} = \mathbf{S}_{k+2} = \ldots = \mathbf{S}_{k+m} = r$, $\mathbf{S}_{k+m+1} = r \pm 1$, $0 < k \leq k + m < n$. Again, if there should be one or more consecutive horizontal steps coinciding with the line $y = r$ (i.e. $m > 0$), this counts only as a single *crossing of r*.

VI. $\mathbf{D}_n^+, \mathbf{D}_n^-$, *the one-sided maximum deviations.* $\mathbf{D}_n^+ = \max_{0 \leq k \leq n} \mathbf{S}_k$. Analogously, $\mathbf{D}_n^- = \min_{0 \leq k \leq n} \mathbf{S}_k$.

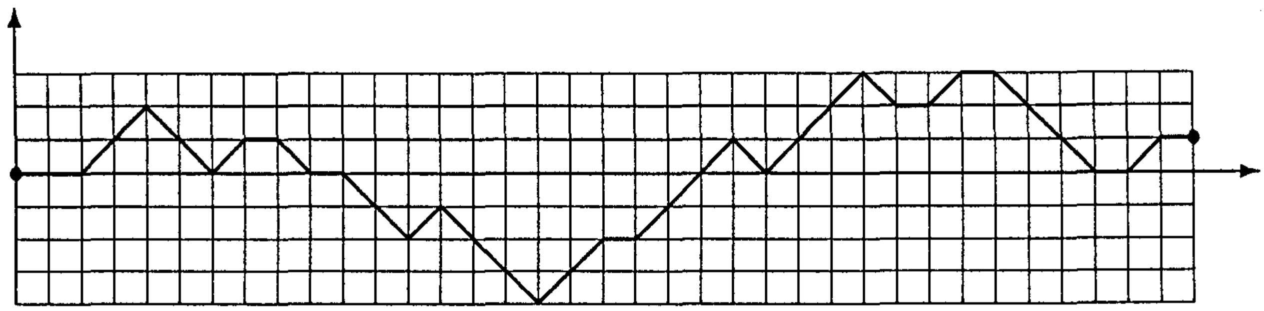
VII. $\mathbf{D}_n$, *the two-sided maximum deviation.* $\mathbf{D}_n = \max_{0 \leq k \leq n} |\mathbf{S}_k|$.

VIII. $\mathbf{Q}_n$, *the number of times where the maximum is achieved.* The maximum is achieved, if $\mathbf{S}_k = \mathbf{S}_{k+1} = \mathbf{S}_{k+2} = \ldots = \mathbf{S}_{k+m} = \mathbf{D}_n^+$ and $\mathbf{S}_{k-1}, \mathbf{S}_{k+m+1} < \mathbf{D}_n^+$, $0 \leq k \leq k + m \leq n$. By definition, if $\mathbf{S}_0 = \mathbf{D}_n^+$ the path starts with a maximum; accordingly, if $\mathbf{S}_n = \mathbf{D}_n^+$ the path ends with a maximum. As before, if there should be one or more consecutive horizontal steps coinciding with the line $y = \mathbf{D}_n^+$ (i.e. $m > 0$), this counts only as a single maximum.

IX. $\mathbf{Q}_{n,k}$, *the position of the k-th return to zero.* $\mathbf{Q}_{n,k}$ is the first index of the k-th *return to zero*. The *visit to zero* that automatically occurs at the origin is not taken into account, i.e. $\mathbf{Q}_{n,1} \geq 2$.

X. $\mathbf{R}_n^+$, *the position where the maximum is first achieved.* $\mathbf{R}_n^+ = \min\{k | \mathbf{S}_k = \mathbf{D}_n^+\}$.

XI. $\mathbf{L}_n$, *the number of steps strictly above the x-axis.*



I: $\mathbf{N}_n = 6$; II: $\mathbf{N}_n^+ = 4$, $\mathbf{N}_n^- = 1$; III: $\mathbf{N}_n(1) = 7$, $\mathbf{N}_n(2) = 4$;
IV: $\mathbf{N}_n^+(1) = 2$, $\mathbf{N}_n^+(2) = 2$; V: $\mathbf{N}_n^*(1) = 4$, $\mathbf{N}_n^*(2) = 2$;
VI: $\mathbf{D}_n^+ = 3$, $\mathbf{D}_n^- = 4$; VII: $\mathbf{D}_n = 4$; VIII: $\mathbf{Q}_n = 2$;
IX: $\mathbf{Q}_{n,1} = 6$, $\mathbf{Q}_{n,2} = 9$, $\mathbf{Q}_{n,3} = 21$, $\mathbf{Q}_{n,4} = 23$, $\mathbf{Q}_{n,5} = 33$;
X: $\mathbf{R}_n^+ = 26$; XI: $\mathbf{L}_n = 21$.

Fig. 1: *Random walk diagram, values of random variables*

Distribution results. In the following formulae the condition $S_0 = 0$ and the conjunct $S_n = \ell$ are omitted for convenience. That is we write e.g. $P(N_n > k)$ as a shorthand notation for $P(N_n > k, S_n = \ell \mid S_0 = 0)$. For each random variable we give the distribution for the general case and its specialization to the binomial case $\alpha = \gamma = 1/2$. The binomial results cover Dwass's results, which can be checked by specializing on $\ell = 0$, substituting $2n$ for n and dividing by the probability of the conditioning event, viz. $P(S_{2n} = 0 \mid S_0 = 0) = 2^{-2n}\binom{2n}{n}$.

I. For all $k \geq 0$ we have

$$P(N_n > k) = 2^k \rho^{k+\frac{|\ell|+\ell}{2}} \sum_{j\geq 0} \rho^j \binom{-k}{j} \binom{n;\alpha,\beta,\gamma}{n-|\ell|-2k-2j}.$$

For the special case $\alpha = \gamma = 1/2, \beta = 0$ one obtains

$$P(N_n > k) = 2^{k-n} \binom{n-k}{\frac{n-|\ell|}{2} - k}.$$

II. For all $k \geq 0$ we have

$$P(N_n^+ \geq k) = P(N_n^- \geq k) = \rho^{k+\frac{|\ell|+\ell}{2}} \binom{n;\alpha,\beta,\gamma}{n-|\ell|-2k}.$$

For the binomial case $\alpha = \gamma = 1/2, \beta = 0$ one obtains

$$P(N_n^+ \geq k) = P(N_n^- \geq k) = 2^{-n} \binom{n}{\frac{n-|\ell|}{2} - k}.$$

III. For all $k \geq 0$ and $r \geq 0$ we have

$$P(N_n(r) > k) = 2^k \rho^{-k+\frac{(\ell-r)-|\ell-r|}{2}} \sum_{j\geq 0} \rho^{-j} \binom{-k}{j} \binom{n;\alpha,\beta,\gamma}{n+r+|\ell-r|+2k+2j},$$

$$P(N_n(r) = 0) = \begin{cases} \binom{n;\alpha,\beta,\gamma}{n+\ell} - \rho^r \binom{n;\alpha,\beta,\gamma}{n+\ell-2r} & \text{for } \ell \leq r, \\ 0 & \text{for } \ell \geq r. \end{cases}$$

Observe that the distribution of $N_n(r)$ does not depend on r if $\ell \geq r$. Specializing on the binomial case we get for all $k \geq 0$

$$P(N_n(r) > k) = 2^{k-n} \binom{n-k}{\frac{n-r-|\ell-r|}{2} - k},$$

$$P(N_n(r) = 0) = \begin{cases} 2^{-n} \left[\binom{n}{\frac{n+\ell}{2}} - \binom{n}{\frac{n+\ell}{2} - r} \right] & \text{for } \ell \leq r, \\ 0 & \text{for } \ell \geq r. \end{cases}$$

IV. For all $k > 0$ and $r \geq 0$ we have

$$\mathbf{P}(\mathbf{N}_n^+(r) \geq k) = \rho^{-k+\frac{(\ell-r)-|\ell-r|}{2}} \binom{n;\alpha,\beta,\gamma}{n+r+|\ell-r|+2k},$$

$$\mathbf{P}(\mathbf{N}_n^+(r) = 0) = \begin{cases} \binom{n;\alpha,\beta,\gamma}{n+\ell} - \rho^{r+1}\binom{n;\alpha,\beta,\gamma}{n+\ell-2r-2} & \text{for } \ell \leq r, \\ \binom{n;\alpha,\beta,\gamma}{n+\ell} - \rho^{-1}\binom{n;\alpha,\beta,\gamma}{n+\ell+2} & \text{for } \ell \geq r. \end{cases}$$

Here again the distribution of $\mathbf{N}_n^+(r)$ does not depend on r if $\ell \geq r$. For the binomial case this entails for all $k > 0$ and $r \geq 0$

$$\mathbf{P}(\mathbf{N}_n^+(r) \geq k) = 2^{-n}\binom{n}{\frac{n-r-|\ell-r|}{2} - k},$$

$$\mathbf{P}(\mathbf{N}_n^+(r) = 0) = \begin{cases} 2^{-n}\left[\binom{n}{\frac{n+\ell}{2}} - \binom{n}{\frac{n+\ell}{2} - r - 1}\right] & \text{for } \ell \leq r, \\ 2^{-n}\left[\binom{n}{\frac{n+\ell}{2}} - \binom{n}{\frac{n+\ell}{2}+1}\right] & \text{for } \ell \geq r. \end{cases}$$

The subcase $\ell = 0$ has been delt with in Mihalevic [1952].

V. For $r > 0$ we have

$$\mathbf{P}(\mathbf{N}_n^*(r) \geq k) = \begin{cases} \rho^{\ell+k-1}\binom{n;\alpha,\beta,\gamma}{n-\ell-2k+2} & \ell > r,\ k = 1,3,5,\ldots \\ \rho^{r+k-1}\binom{n;\alpha,\beta,\gamma}{n+\ell-2k-2r+2} & \ell < r,\ k = 2,4,6,\ldots \\ \rho^{r+k}\binom{n;\alpha,\beta,\gamma}{n-2k-r} & \ell = r,\ k = 1,2,3,\ldots \end{cases}$$

$$\mathbf{P}(\mathbf{N}_n^*(r) = 0) = \begin{cases} \binom{n;\alpha,\beta,\gamma}{n+\ell} - \rho^{r+1}\binom{n;\alpha,\beta,\gamma}{n+\ell-2r-2} & \ell \leq r, \\ 0 & \ell > r. \end{cases}$$

Similar as before, the distribution of $\mathbf{N}_n^*(r)$ does not depend on r for $\ell > r > 0$. For $r = 0$ one obtains

$$\mathbf{P}(\mathbf{N}_n^*(0) \geq k) = \begin{cases} \rho^{k+\frac{\ell+|\ell|}{2}}\binom{n;\alpha,\beta,\gamma}{n-2k-|\ell|} & \ell \neq 0,\ k = 0,1,2,\ldots \\ 2\rho^{k+1}\sum_{j\geq 0}(-1)^j\rho^j\binom{n;\alpha,\beta,\gamma}{n-2k-2j-2} & \ell = 0,\ k = 1,2,3,\ldots \end{cases}$$

$$\mathbf{P}(\mathbf{N}_n^*(0) = 0) = 2\left[\binom{n;\alpha,\beta,\gamma}{n} - \rho\binom{n;\alpha,\beta,\gamma}{n-2}\right] - \beta^n \qquad \ell = k = 0.$$

For the binomial case $\alpha = \gamma = 1/2$ we have for $r > 0$

$$\mathbf{P}(\mathbf{N}_n^*(r) \geq k) = \begin{cases} 2^{-n}\binom{n}{\frac{n-\ell}{2} - k + 1} & \ell > r,\ k = 1,3,5,\ldots \\ 2^{-n}\binom{n}{\frac{n+\ell}{2} - k - r + 1} & \ell < r,\ k = 2,4,6,\ldots \\ 2^{-n}\binom{n}{\frac{n-r}{2} - k} & \ell = r,\ k = 1,2,3,\ldots \end{cases}$$

$$\mathbf{P}(\mathbf{N}_n^*(r) = 0) = \begin{cases} 2^{-n}\left[\binom{n}{\frac{n+\ell}{2}} - \binom{n}{\frac{n+\ell}{2} - r - 1}\right] & \ell \leq r, \\ 0 & \ell > r. \end{cases}$$

For $r = 0$ and all $k \geq 0$ one obtains

$$\mathbf{P}(\mathbf{N}_n^*(0) \geq k) = \begin{cases} 2^{-n}\binom{n}{\frac{n+|\ell|}{2} + k} & \ell \neq 0, \\ 2^{-n+1}\binom{n-1}{\frac{n}{2} - k - 1} & \ell = 0, \end{cases}$$

which is equivalent to Kanwar Sen's result [1965]. The subcase $\ell = 0, r \geq 0$ is treated in Csáki and Vincze [1961]. Incidentally, the corresponding formula of Dwass seems to be erroneous.

VI(a). For all $k \geq 0$ we have

$$\mathbf{P}(\mathbf{D}_n^+ \geq k) = \rho^k \binom{n;\alpha,\beta,\gamma}{n+\ell-2k}, \quad \mathbf{P}(\mathbf{D}_n^- \geq k) = \rho^{-k}\binom{n;\alpha,\beta,\gamma}{n+\ell+2k},$$

which is equivalent to (14) of Katzenbeisser and Panny [1984]. For the binomial case this becomes

$$\mathbf{P}(\mathbf{D}_n^+ \geq k) = 2^{-n}\binom{n}{\frac{n+\ell}{2} - k}, \quad \mathbf{P}(\mathbf{D}_n^- \geq k) = 2^{-n}\binom{n}{\frac{n+\ell}{2} + k}.$$

This is equivalent to Lemma 1 in Feller [1968, p.89]. For $\ell = 0$ this is the well-known result due to Gnedenko and Korolyuk [1951].

VI(b). For all $k, r \geq 0$ the joint distribution of both maximal deviations can be given by

$$\begin{aligned} \mathbf{P}(\mathbf{D}_n^+ < k, \mathbf{D}_n^- < r) = &\sum_{j=0,\pm1,\ldots} \rho^{-j(k+r)}\binom{n;\alpha,\beta,\gamma}{n+\ell+2j(k+r)} \\ &- \rho^k \sum_{j=0,\pm1,\ldots} \rho^{-j(k+r)}\binom{n;\alpha,\beta,\gamma}{n+\ell-2k+2j(k+r)}. \end{aligned}$$

This is Theorem 1 of Katzenbeisser and Panny [1984]. For the binomial case we can write

$$\mathbf{P}(\mathbf{D}_n^+ < k, \mathbf{D}_n^- < r) = 2^{-n} \sum_{j=0,\pm1,\ldots} \binom{n}{\frac{n+\ell}{2} + j(k+r)} - 2^{-n} \sum_{j=0,\pm1,\ldots} \binom{n}{\frac{n+\ell}{2} - k + j(k+r)},$$

which comprises the result of Gnedenko and Rvaceva [1952].

VII. For all $k \geq 0$ we have

$$\mathbf{P}(\mathbf{D}_n < k) = \sum_{j=0,\pm1,\ldots} \rho^{-2jk} \binom{n;\alpha,\beta,\gamma}{n+\ell+4jk} - \rho^k \sum_{j=0,\pm1,\ldots} \rho^{-2jk} \binom{n;\alpha,\beta,\gamma}{n+\ell-2k+4jk},$$

which is equivalent to (13) of Katzenbeisser and Panny [1984]. For the binomial case we get

$$\mathbf{P}(\mathbf{D}_n < k) = 2^{-n} \sum_{j=0,\pm1,\ldots} \binom{n}{\frac{n+\ell}{2} + 2jk} - 2^{-n} \sum_{j=0,\pm1,\ldots} \binom{n}{\frac{n+\ell}{2} - k + 2jk},$$

which covers the result due to Gnedenko and Korolyuk [1951].

VIII(a). For all $k, r \geq 0$ we have

$$\mathbf{P}(\mathbf{D}_n^+ \geq k, \mathbf{Q}_n > r) = \rho^{k+r} \sum_{j\geq 0} \rho^j \binom{-r}{j} \binom{n;\alpha,\beta,\gamma}{n+\ell-2k-2r-2j},$$

which entails for the binomial case

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

The last formula is equivalent to (4.9) in Mohanty [1979, p.93].

VIII(b). For all $r > 0$ we have

$$\mathbf{P}(\mathbf{Q}_n = r) = \rho^{r-1+\frac{|\ell|+\ell}{2}} \sum_{j\geq 0} \rho^j \binom{-r}{j} \binom{n;\alpha,\beta,\gamma}{n-|\ell|-2r-2j+2}.$$

For the binomial case Vandermonde's convolution formula yields

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

IX(a). For all i, k, r subject to $2 \leq 2k \leq i \leq n - |\ell|, 0 \leq 2r \leq n - i - |\ell|$, i.e. for all possible cases, we have

$$\mathbf{P}(\mathbf{Q}_{n,k} = i,\ \mathbf{N}_n = k + r + 1) = \alpha 2^{k+r} \rho^{k+r+1+\frac{|\ell|+\ell}{2}}$$
$$\times \sum_{j \geq 0} \rho^j \binom{-k}{j} \left[\frac{1}{\rho} \binom{i-1; \alpha, \beta, \gamma}{i - 2k - 2j} - \binom{i-1; \alpha, \beta, \gamma}{i - 2k - 2j + 2} \right]$$
$$\times \sum_{j \geq 0} \rho^j \binom{-r-1}{j} \left[\frac{1}{\rho} \binom{n-i; \alpha, \beta, \gamma}{n-i-|\ell|-2r-2j} - \binom{n-i; \alpha, \beta, \gamma}{n-i-|\ell|-2r-2j-2} \right].$$

For the binomial case this simplifies to the following: for all i, k, r subject to $2 \leq 2k \leq i \leq n - |\ell|, 0 \leq 2r \leq n - i - |\ell|, n - i > 0$ i.e. for all possible cases excluding only the trivial case $n - i = 0$ we have

$$\mathbf{P}(\mathbf{Q}_{n,k} = i,\ \mathbf{N}_n = k + r + 1) = 2^{-n+k+r} \frac{k}{i-k} \binom{i-k}{\frac{i}{2}} \frac{r + |\ell|}{n - i - r} \binom{n-i-r}{\frac{n-i+|\ell|}{2}},$$

which checks with the right hand side of Dwass's result. On the left hand side of Dwass's result the visit to zero that automatically occurs at the origin has been overlooked. The natural formulation of the event reads: the k-th *return to zero* occurs at i and there are $k + r$ *returns to zero* altogether. Using $\mathbf{N}_n$, which counts the *visits to zero* in the formal definition of the event, makes it necessary to consider the *visit to zero* at the origin, too.

IX(b). For all i, k subject to $2 \leq 2k \leq i \leq n - |\ell|, 0 \leq n - i - |\ell|$, i.e. for all possible cases, we have

$$\mathbf{P}(\mathbf{Q}_{n,k} = i,\ \mathbf{N}_n > k) = \alpha\, 2^k \rho^k \binom{n-i; \alpha, \beta, \gamma}{n-i+\ell}$$
$$\times \sum_{j \geq 0} \rho^j \binom{-k}{j} \left[\frac{1}{\rho} \binom{i-1; \alpha, \beta, \gamma}{i - 2k - 2j} - \binom{i-1; \alpha, \beta, \gamma}{i - 2k - 2j + 2} \right].$$

Specializing on the binomial case we get

$$\mathbf{P}(\mathbf{Q}_{n,k} = i,\ \mathbf{N}_n > k) = 2^{-n+k} \frac{k}{i-k} \binom{i-k}{\frac{i}{2} - k} \binom{n-i}{\frac{n-i+\ell}{2}},$$

where again all possible cases $(2 \leq 2k \leq i \leq n - |\ell|, 0 \leq n - i - |\ell|)$ are covered.

X(a). For all k, m satisfying $0 < k \leq m \leq n + \ell - k \leq n$, which excludes only the trivial case $k = m = 0$ we have

$$\mathbf{P}(\mathbf{D}_n^+ = k,\ \mathbf{R}_n^+ = m) = \alpha \left[\binom{m-1; \alpha, \beta, \gamma}{m + k - 2} - \rho^{-1} \binom{m-1; \alpha, \beta, \gamma}{m + k} \right]$$
$$\times \left[\binom{n-m; \alpha, \beta, \gamma}{n - m + \ell - k} - \rho \binom{n-m; \alpha, \beta, \gamma}{n - m + \ell - k - 2} \right].$$

For the binomial case one obtains for the same range of k, m-values

$$\mathbf{P}(\mathbf{D}_n^+ = k,\, \mathbf{R}_n^+ = m) = 2^{-n} \frac{k(k+1-\ell)}{m(n-m+1)} \binom{m}{\frac{m+k}{2}} \binom{n-m+1}{\frac{n-m-k+\ell}{2}}.$$

For $\ell = 0$ this specializes to Vincze's result [1957].

X(b). Let us first consider the non-trivial case $k > 0$, $n > 0$. For all possible values of k, viz. $\max\{2, 2\ell\} \le k \le n+\ell$, we have

$$\mathbf{P}(\mathbf{D}_n^+ + \mathbf{R}_n^+ = k) = \alpha \sum_{j \ge 0} \left[\binom{k-j-2; \alpha, \beta, \gamma}{k-2} - \rho^{-1} \binom{k-j-2; \alpha, \beta, \gamma}{k} \right] \times \left[\binom{n-k+j+1; \alpha, \beta, \gamma}{n-k+\ell} - \rho \binom{n-k+j+1; \alpha, \beta, \gamma}{n-k+\ell-2} \right].$$

For the trivial case $k = 0$, $n \ge 0$ we obtain

$$\mathbf{P}(\mathbf{D}_n^+ + \mathbf{R}_n^+ = 0) = \begin{cases} \binom{n; \alpha, \beta, \gamma}{n+\ell} - \rho \binom{n; \alpha, \beta, \gamma}{n+\ell-2} & \ell \le 0, \\ 0 & \ell > 0. \end{cases}$$

Let us now turn to the binomial case. We first consider the non-trivial case $k > 0$, $n > 0$. The following formula is valid for all possible values of k, viz. $\max\{2, 2\ell\} \le k \le n+\ell$, more explicitly: $k = 2, 4, \ldots, n+\ell$, $n+\ell$ even, of course.

$$\mathbf{P}(\mathbf{D}_n^+ + \mathbf{R}_n^+ = k) = \frac{2^{-n+2}}{k(n-k+\ell)} \sum_{j \ge 1} j(j+1-\ell) \binom{k-j-1}{\frac{k}{2}-1} \binom{n-k+j}{\frac{n-k+\ell}{2}-1},$$

For $\ell = 0$ we have the following closed form expression

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

which shows that $\mathbf{D}_n^+ + \mathbf{R}_n^+$ is uniformly distributed over $k = 0, 2, \ldots, n$ for that case. For the trivial case $k = 0$, $n \ge 0$ we obtain

$$\mathbf{P}(\mathbf{D}_n^+ + \mathbf{R}_n^+ = 0) = \begin{cases} 2^{-n+1} \frac{|\ell|+1}{n+|\ell|+2} \binom{n}{\frac{n+|\ell|}{2}} & \ell \le 0, \\ 0 & \ell > 0. \end{cases}$$

XI. In general the range of k is $\frac{\ell+|\ell|}{2} \le k \le n + \frac{\ell-|\ell|}{2}$. Let us first consider the non-trivial case $k > 0$, $n > 0$. For $\ell \le 0$ we get

$$\mathbf{P}(\mathbf{L}_n = k) = \alpha \sum_{j \ge 0} \left[\binom{k-j-2; \alpha, \beta, \gamma}{k-2} - \rho^{-1} \binom{k-j-2; \alpha, \beta, \gamma}{k} \right] \times \left[\binom{n-k+j+1; \alpha, \beta, \gamma}{n-k+\ell} - \rho \binom{n-k+j+1; \alpha, \beta, \gamma}{n-k+\ell-2} \right].$$

For $\ell > 0$ one obtains

$$\mathbf{P}(\mathbf{L}_n = k) = \alpha \sum_{j \geq 0} \left[\binom{k-j-1; \alpha, \beta, \gamma}{k+\ell-2} - \rho^{-1} \binom{k-j-1; \alpha, \beta, \gamma}{k+\ell} \right] \times \left[\binom{n-k+j; \alpha, \beta, \gamma}{n-k} - \rho \binom{n-k+j; \alpha, \beta, \gamma}{n-k-2} \right].$$

For the trivial case $k = 0$, $n \geq 0$ we obtain

$$\mathbf{P}(\mathbf{L}_n = 0) = \begin{cases} \binom{n; \alpha, \beta, \gamma}{n+\ell} - \rho \binom{n; \alpha, \beta, \gamma}{n+\ell-2} & \ell \leq 0, \\ 0 & \ell > 0. \end{cases}$$

The last two expressions show that the random variables $\mathbf{D}_n^+ + \mathbf{R}_n^+$ and $\mathbf{L}_n$ have the same distribution for $\ell \leq 0$. This remains true for the binomial case, of course. Accordingly, for the case $\ell \leq 0$, $n > 0$, $k = 2, 4, \ldots, n+\ell$, $n+\ell$ even, the following formula applies

$$\mathbf{P}(\mathbf{L}_n = k) = \frac{2^{-n+2}}{k(n-k+\ell)} \sum_{j \geq 1} j(j+1-\ell) \binom{k-j-1}{\frac{k}{2}-1} \binom{n-k+j}{\frac{n-k+\ell}{2}-1},$$

which becomes

$$\mathbf{P}(\mathbf{L}_n = k) = 2^{-n} \frac{1}{\frac{n}{2}+1} \binom{n}{\frac{n}{2}},$$

for $\ell = 0$, $k = 0, 2, \ldots, n$ (cf. Chung and Feller [1949], Feller [1968, p.94f.]). For $\ell > 0$, $n > 0$, $k = \ell, \ell+2, \ldots, n$; $n+\ell$ even, we have

$$\mathbf{P}(\mathbf{L}_n = k) = \frac{2^{-n+2}}{(n-k)(k+\ell)} \sum_{j \geq 0} (j+\ell)(j+1) \binom{k-j-1}{\frac{k+\ell}{2}-1} \binom{n-k+j}{\frac{n-k}{2}-1},$$

For the trivial case $k = 0$, $n \geq 0$ we obtain as before

$$\mathbf{P}(\mathbf{L}_n = 0) = \begin{cases} 2^{-n+1} \frac{|\ell|+1}{n+|\ell|+2} \binom{n}{\frac{n+|\ell|}{2}} & \ell \leq 0, \\ 0 & \ell > 0. \end{cases}$$

Proofs of distribution results.

I. Fig. 2 shows a sample path for the event $\mathbf{N}_n = 3$.

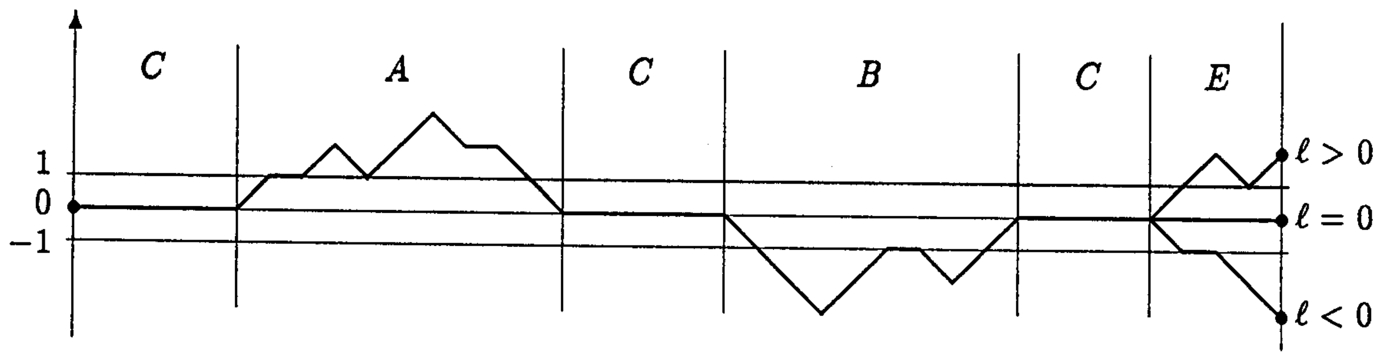


Fig. 2: *Visits to zero* **(I)**

A path corresponding to the event $\mathbf{N}_n = k$ can symbolically be characterized by $C[(A+B)C]^{k-1}E$, $k = 1,2,\ldots$, where the segments $A, B, C,$ and E are illustrated in Fig.2. The p.g.f.'s for the segments are:

$$\Upsilon_C(y) = \Psi_{0,0,0}(y),$$

$$\Upsilon_A(y) = \alpha\gamma y^2 \Psi_{\infty,0,0}(y) = \frac{\alpha\gamma v^2}{(\alpha v^2 + \beta v + \gamma)^2} \Psi_{\infty,0,0}(y),$$

$$\Upsilon_B(y) = \alpha\gamma y^2 \Psi_{0,\infty,0}(y) = \frac{\alpha\gamma v^2}{(\alpha v^2 + \beta v + \gamma)^2} \Psi_{0,\infty,0}(y) = \Upsilon_A(y),$$

$$\Upsilon_E(y) = \begin{cases} \alpha y \Psi_{\infty,0,\ell-1}(y) = (\rho v)^\ell & \ell > 0, \\ 1 & \ell = 0, \\ \gamma y \Psi_{0,\infty,\ell+1}(y) = v^{-\ell} & \ell < 0. \end{cases}$$

Hence, the p.g.f. $\phi_I(k,\ell;y)$ for $\mathbf{P}(\mathbf{N}_n = k)$ reads

$$\begin{aligned} \phi_I(k,\ell;y) &= 2^{k-1}\Upsilon_A^{k-1}(y)\Upsilon_C^k(y)\Upsilon_E(y) \\ &= 2^{k-1}\frac{\rho^{\frac{|\ell|+\ell}{2}}}{\gamma}\frac{v^{|\ell|}}{\rho v^2}(\alpha v^2 + \beta v + \gamma)\Big(\frac{\rho v^2}{1+\rho v^2}\Big)^k. \end{aligned}$$

Consequently, the p.g.f. $\Phi_I(k,\ell;y)$ for the tail probabilities $\mathbf{P}(\mathbf{N}_n > k, k = 0,1,\ldots)$ is

$$\Phi_I(k,\ell;y) = 2^k \frac{\rho^{\frac{|\ell|+\ell}{2}}}{\gamma}\frac{v^{|\ell|}}{1-\rho v^2}(\alpha v^2 + \beta v + \gamma)\Big(\frac{\rho v^2}{1+\rho v^2}\Big)^k.$$

Picking out the coefficient of y^n by means of Cauchy's integral formula where we have to take into account that

$$\frac{dy}{y^{n+1}} = \gamma \frac{1-\rho v^2}{v^{n+1}}(\alpha v^2 + \beta v + \gamma)^{n-1} dv$$

yields the given result.

II. Fig.3 shows a sample path for the event $\mathbf{N}_n^+ = 2$.

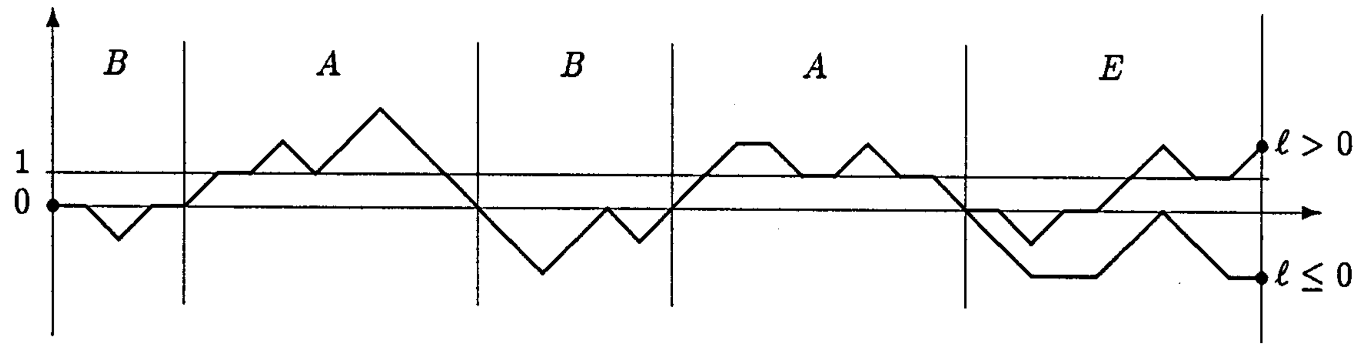


Fig. 3: *Positive sojourns* **(II)**

A path corresponding to the event $\mathbf{N}_n = k$ can symbolically be written as $(BA)^k E$, $k = 0, 1, \ldots$, where the segments A, B, and E are illustrated in Fig.3. The p.g.f.'s for the segments are:

$$\Upsilon_A(y) = \alpha\gamma y^2 \Psi_{\infty,0,0}(y),$$
$$\Upsilon_B(y) = \Psi_{0,\infty,0}(y),$$
$$\Upsilon_E(y) = \begin{cases} \Psi_{0,\infty,0}(y)\, \alpha y\, \Psi_{\infty,0,\ell-1}(y) & \ell > 0, \\ \Psi_{0,\infty,\ell}(y) & \ell \leq 0. \end{cases}$$

Hence, the p.g.f. for $\mathbf{P}(\mathbf{N}_n^+ = k)$ is

$$\phi_{II}(k,\ell;y) = [\Upsilon_A(y)\Upsilon_B(y)]^k \Upsilon_E(y) = \frac{\rho^{\frac{|\ell|+\ell}{2}}}{\gamma} v^{|\ell|} (\rho v^2)^k (\alpha v^2 + \beta v + \gamma),$$

and the p.g.f. for $\mathbf{P}(\mathbf{N}_n^+ \geq k)$ is

$$\Phi_{II}(k,\ell;y) = \frac{\rho^{\frac{|\ell|+\ell}{2}}}{\gamma} v^{|\ell|} \frac{(\rho v^2)^k}{1-\rho v^2} (\alpha v^2 + \beta v + \gamma).$$

Again, the coefficient of y^n can be determined by means of Cauchy's integral formula and the result follows. Proceeding correspondingly for $\mathbf{N}_n^-$ leads to the same generating functions.

III. This generalizes the *visits to zero* dealt with in I. It is easily seen from the decomposition given in Fig.4 that we only have to shift the line $y = 0$ into the line $y = r$ for the paths considered in I and prefix them by those paths (segment S) which have their first visit to r in their connecting point. Hence, the p.g.f. $\Phi_{III}(k,\ell,r;y)$ for $\mathbf{P}(\mathbf{N}_n(r) > k)$ is $\Psi_{r-1,\infty,r-1}(y)\, \alpha y\, \Phi_I(k, \ell - r; y)$, which results in

$$\Phi_{III}(k,\ell,r;y) = 2^k \frac{\rho^{r+\frac{|\ell-r|+(\ell-r)}{2}}}{\gamma} \frac{v^{r+|\ell-r|}}{1-\rho v^2} (\alpha v^2 + \beta v + \gamma) \Big(\frac{\rho v^2}{1+\rho v^2}\Big)^k,$$

and the coefficient of y^n can be extracted by the same method as before. The result for the trivial case $k = 0$ follows from a straight forward application of (2′).

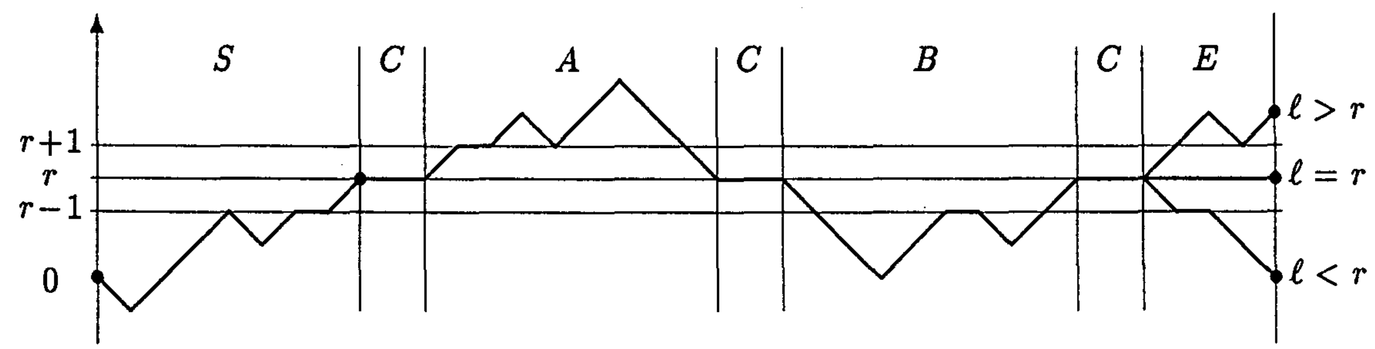


Fig.4: *Visits to r* **(III)**

IV. This generalizes II in the same way as III generalizes I. To obtain the p.g.f. we may proceed as under III, i.e. choosing an appropriate prefix for the paths considered in II, which leads to the p.g.f. $\Psi_{r-1,\infty,r-1}(y)\,\alpha y\,\Phi_{II}(k,\ell-r;y)$ for $\mathbf{P}(\mathrm{N}_n^+(r)\geq k)$, viz.

$$\Phi_{IV}(k,\ell,r;y)=\frac{\rho^{r+\frac{|\ell-r|+(\ell-r)}{2}}}{\gamma}v^{r+|\ell-r|}\frac{(\rho v^2)^k}{1-\rho v^2}(\alpha v^2+\beta v+\gamma).$$

For $k=0$ we have to distinguish the cases $\ell\geq r$ and $\ell<r$. For the first case the general formula applies, because r is visited at least once and the connecting point actually exists. The second case is equivalent to the event $\mathrm{N}_n(r+1)=0$.

V. Fig.5 shows an appropriately decomposed sample path for the case $r>0$, $k>0$.

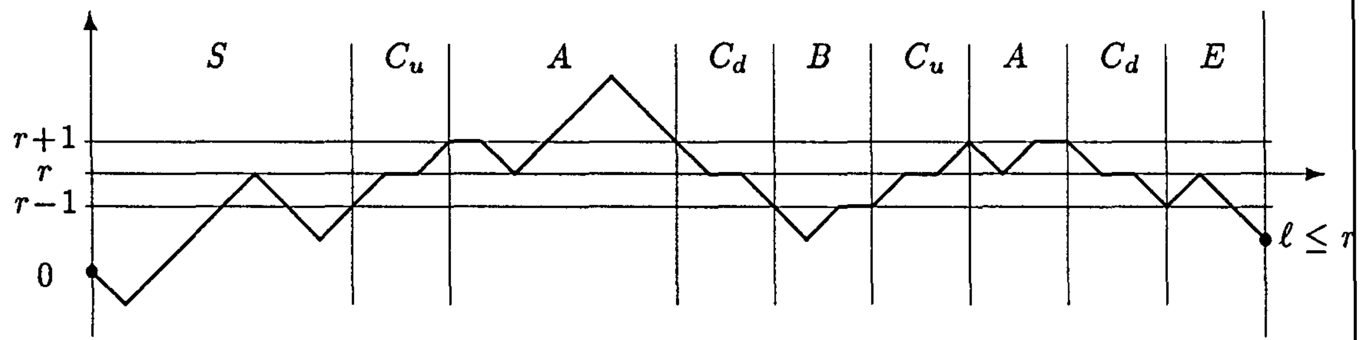


Fig. 5: *Number of crossings of r* **(V)**

A path corresponding to the event $\mathrm{N}_n^*(r)=k$ can symbolically be represented as $SC_u(AC_dBC_u)^{\lfloor (k-1)/2\rfloor}(AC_d)^{(k-1)\bmod 2}E$. The p.g.f.'s for the segments are:

$$\begin{aligned}
\Upsilon_S(y)&=\Psi_{r,\infty,r-1}(y),\\
\Upsilon_{C_u}(y)&=(\alpha y)^2\Psi_{0,0,0}(y),\\
\Upsilon_A(y)&=\Psi_{\infty,1,0}(y),\\
\Upsilon_{C_d}(y)&=(\gamma y)^2\Psi_{0,0,0}(y),\\
\Upsilon_B(y)&=\Psi_{1,\infty,0}(y),\\
\Upsilon_E(y)&=\begin{cases}\Psi_{\infty,1,\ell-r-1}(y) & k=1,3,\ldots,\\ \Psi_{1,\infty,\ell-r+1}(y) & k=2,4,\ldots.\end{cases}
\end{aligned}$$

Hence, the p.g.f. for $\mathbf{P}(\mathrm{N}_n^*(r)=k)$, $r>0,k>0$ is

$$\phi_V(k,\ell,r;y)=\Upsilon_S(y)[\Upsilon_{C_u}(y)]^{\lceil\frac{k}{2}\rceil}[\Upsilon_A(y)\Upsilon_{C_d}(y)]^{\lfloor\frac{k}{2}\rfloor}[\Upsilon_B(y)]^{\lceil\frac{k}{2}\rceil-1}\Upsilon_E(y).$$

Summation over k in the resulting p.g.f. and determining the corresponding coefficient yields the given tail probabilities. Of course, if $\ell\neq r$ then $k=1,3,5,\ldots$ implies $\ell>r$, and $k=0,2,4,\ldots$ implies $\ell<r$. $\mathbf{P}(\mathrm{N}_n^*(r)=0)$ follows from an application of (3′).

Above p.g.f. applies for $r=0$ too, but it only covers the cases $\ell>0$, $k=1,3,\ldots$, and $\ell<0$, $k=2,4,\ldots$. For the cases $\ell<0$, $k=1,3,\ldots$, and $\ell>0$,

$k = 2, 4, \ldots$, we have to consider also paths which have a down-crossing as their first crossing, symbolically $S'C_d(BC_uAC_d)^{\lfloor (k-1)/2 \rfloor}(BC_u)^{(k-1) \bmod 2}E'$, where

$$\Upsilon'_S(y) = \Psi_{\infty,0,1}(y),$$

$$\Upsilon'_E(y) = \begin{cases} \Psi_{1,\infty,\ell+1}(y) & k = 1, 3, \ldots, \\ \Psi_{\infty,1,\ell-1}(y) & k = 2, 4, \ldots. \end{cases}$$

Combining the Υ's results in the following p.g.f. for $P(\mathrm{N}_n^*(0) = k), k > 0, \ell \neq 0$,

$$\phi(k, \ell, 0; y) = \frac{1}{\gamma} (\rho v^2)^k \rho^{\frac{|\ell|+\ell}{2}} v^{|\ell|} (\alpha v^2 + \beta v + \gamma). \tag{6}$$

For the trivial case $r = 0$, $k = 0$, $\ell \neq r$ we have the p.g.f.

$$\begin{cases} \Psi_{\infty,0,\ell}(y) & \ell > 0, \\ \Psi_{0,\infty,\ell}(y) & \ell < 0, \end{cases}$$

which is also comprised in (6). Summation over k and extracting the coefficients in the usual way leads to the given result.

It only remains to consider the case $r = 0$, $\ell = 0$. Let us first deal with the subcase $k > 0$. The corresponding p.g.f. can be obtained from (6) in the following way

$$\begin{aligned} \phi(k, 0, 0; y) &= \phi(k, 1, 0; y)\, \gamma y\, \Psi_{0,0,0}(y) + \phi(k, -1, 0; y)\, \alpha y\, \Psi_{0,0,0}(y) \\ &= 2 \frac{(\rho v^2)^{k+1}}{\gamma} \frac{\alpha v^2 + \beta v + \gamma}{1 + \rho v^2}, \end{aligned}$$

and the result follows. The trivial subcase $k = 0$ is a consequence of $(2')$ and $(3')$, viz. $p(\infty, 0, 0, n) + p(0, \infty, 0, n) - \beta^n$, which gives the probability that a path leads to the point $(n, 0)$ without crossing the axis $y = 0$.

VI. The results are direct consequences of formulae $(3')$, $(2')$, and $(1')$.

VII. This follows from $(1')$ by substituting $k - 1$ for m and h.

VIII. Let us first consider the case $k > 0, \ell < k$. Fig.6 shows an appropriately decomposed sample path (with $r = 3$).

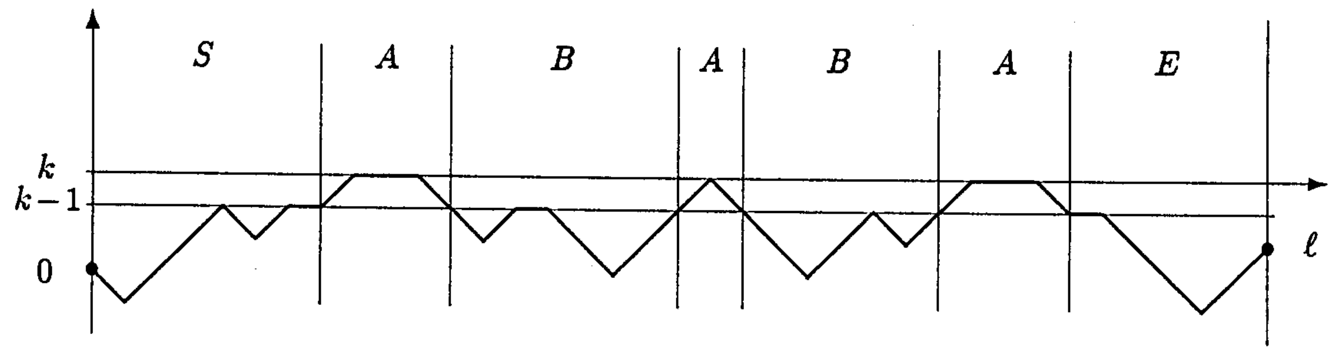


Fig. 6: *Number of times where the maximum is achieved* **(VIII)**

A path corresponding to the event $\mathbf{D}_n^+ = k$, $\mathbf{Q}_n = r$ can symbolically be written as $SA(BA)^{r-1}E$. The p.g.f.'s for the individual segments are

$$\begin{aligned}\Upsilon_S(y) &= \Psi_{k-1,\infty,k-1}(y),\\ \Upsilon_A(y) &= \alpha\gamma y^2\, \Psi_{0,0,0}(y),\\ \Upsilon_B(y) &= \Psi_{0,\infty,0}(y),\\ \Upsilon_E(y) &= \Psi_{0,\infty,\ell-k+1}(y).\end{aligned}$$

Hence, the p.g.f. for $\mathbf{P}(\mathbf{D}_n^+ = k,\ \mathbf{Q}_n = r)$, $r > 0$, $k > 0$, $\ell < k$ is

$$\Upsilon_S(y)\Upsilon_A^r(y)\Upsilon_B^{r-1}(y)\Upsilon_E(y),$$

which entails

$$\phi_{VIII}(k,\ell,r;y) = \frac{1}{\gamma}\Big(\frac{\rho v^2}{1+\rho v^2}\Big)^r(\alpha v^2+\beta v+\gamma)\frac{\big(\rho v^2\big)^{k-1}}{v^\ell}.$$

A separate investigation of the remaining cases shows that the last formula applies for all cases, i.e. for $r > 0$, $k \geq 0$, $\ell \leq k$. Summation over k and r and extraction of the corresponding coefficient yields the result VIII(a). Accordingly, summation over all possible values of k ($k \geq \max\{0,\ell\}$) for fixed r in the p.g.f. leads to VIII(b).

IX(a). The probability under consideration can be expressed as $[\gamma\mathbf{P}(\mathbf{N}_{i-1}=k-1,\mathbf{S}_{i-1}=1)+\alpha\mathbf{P}(\mathbf{N}_{i-1}=k-1,\mathbf{S}_{i-1}=-1)]\ \mathbf{P}(\mathbf{N}_{n-i}=r+1,\mathbf{S}_{n-i}=\ell)$. The occurring probabilities are known from I.

IX(b). Obviously, this probability can be written as $[\gamma\mathbf{P}(\mathbf{N}_{i-1} = k-1,\mathbf{S}_{i-1} = 1)$ $+\alpha\mathbf{P}(\mathbf{N}_{i-1}=k-1,\mathbf{S}_{i-1}=-1)]\,\mathbf{P}(\mathbf{S}_{n-i}=\ell)$, and the involved probabilities are known from I and (4′).

X(a). We have $\mathbf{P}(\mathbf{D}_n^+ = k,\ \mathbf{R}_n^+ = m) = p(k-1,\infty,k-1,m-1)\,\alpha\, p(0,\infty,\ell-k,n-m)$ and the result follows from an application of (3′).

X(b). For the non-trivial case one only has to sum over all complementary values of $\mathbf{D}_n^+$ and $\mathbf{R}_n^+$, i.e. $\mathbf{D}_n^+ + \mathbf{R}_n^+ = k$, and the result follows after some manipulations. For the trivial case $k = 0, \ell \leq 0$ the probability is $p(0,\infty,\ell,n)$.

The derivation of the closed form expression for the binomial case and $\ell = 0$ is a little intricate. Here we may adopt the following path-combinatorial argument: Spezializing the general formula for $\mathbf{P}(\mathbf{D}_n^+ + \mathbf{R}_n^+ = k)$ to the case at hand (viz.$\alpha = \gamma = 1/2$, $\ell = 0$), we see that the resulting expression equals

$$\frac{1}{2}\sum_{j\geq 0} p(\infty,0,j,k-j-2)\,p(\infty,0,j+1,n-k+j+1).$$

By reverting the paths contributing to $p(\infty,0,j+1,n-k+j+1)$ and shifting their end-points into $(n,0)$ we see that

$$p(\infty,0,j+1,n-k+j+1) = \mathbf{P}(\mathbf{S}_n = 0, \mathbf{S}_\kappa \geq 0 \text{ for } k-j-1 \leq \kappa \leq n \mid \mathbf{S}_{k-j-1} = j+1).$$

On the other hand we obviously have

$$p(\infty,0,j,k-j-2) = \mathbf{P}(\mathbf{S}_{k-j-2}=j, \mathbf{S}_\kappa \geq 0 \text{ for } 0 \leq \kappa \leq k-j-2 \,|\, \mathbf{S}_0 = 0),$$

and

$$\mathbf{P}(\mathbf{S}_{k-j-1}=j+1 \,|\, \mathbf{S}_{k-j-2}=j) = \frac{1}{2}.$$

Hence, a single term in the above sum gives the probability that a non-negative path starting at the origin leads to the point $(n,0)$ and has an edge from $(k-j-2,j)$ to $(k-j-1,j+1)$, $j \geq 0$. Fig. 7 illustrates the situation.

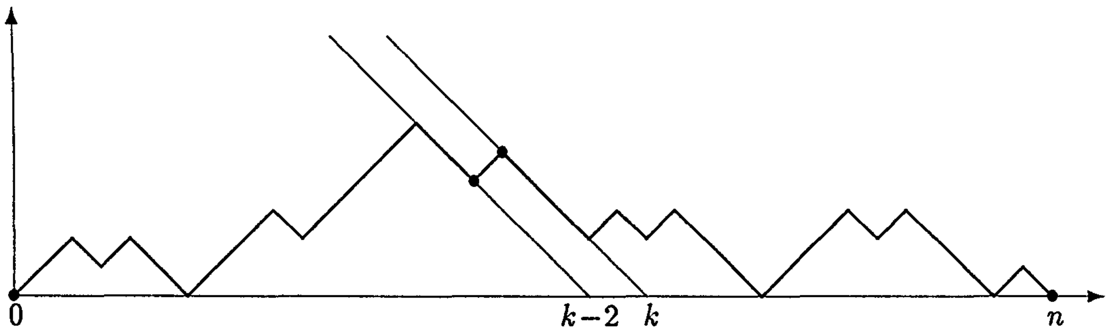


Fig. 7

Since every non-negative (binomial) path bridges the stripe between the lines $y = k-2-x$ and $y = k-x$ by exactly one such edge the whole sum represents the probability $p(\infty,0,0,n)$ and the result follows. Unfortunately, this construction works only for the binomial case and even there only if $\ell = 0$.

XI. Let us first consider the non-trivial case $k > 0$. We have to distinguish between the subcases $\ell \leq 0$ and $\ell > 0$. For $\ell \leq 0$ we may refer to Fig. 3. A path corresponding to the event $\mathbf{L}_n = m$, $\mathbf{N}_n^+ = k$ can symbolically be written as $(BA)^k E$. Hence

$$\mathbf{P}(\mathbf{L}_n = m) = \sum_{k \geq 1} [y^m]\{\Upsilon_A^k(y)\}\, [y^{n-m}]\{\Upsilon_B^k(y)\Upsilon_E(y)\},$$

and the result follows by the usual procedure. For $\ell > 0$ the path can be characterized by $B(AB)^k E'$, where E' represents the (strictly positive) segment from the last visit to zero to the end-point. The pertaining p.g.f. reads

$$\Upsilon_{E'}(y) = \alpha y\, \Psi_{\infty,0,\ell-1}(y),$$

which entails

$$\mathbf{P}(\mathbf{L}_n = m) = \sum_{k \geq 1} [y^m]\{\Upsilon_A^k(y)\Upsilon_{E'}(y)\}\, [y^{n-m}]\{\Upsilon_B^{k+1}(y)\},$$

and the result follows. As to the trivial case $k = 0$, $n \geq 0$ we have $\mathbf{P}(\mathbf{L}_n = 0) = \mathbf{P}(N_n(1) = 0)$, and the result follows from III.

References.

ANEJA, K. G. and KANWAR SEN [1972] Random walk and distributions of rank order statistics, *SIAM J. Appl. Math.* **23**, 276–287.

BARTON, D. E. and MALLOWS, C. L. [1965] Some aspects of the random sequence, *Ann. Math. Statist.* **36**, 236–260.

CHUNG, K. L. and FELLER, W. [1949] Fluctuations in coin tossing, *Proc. Natl. Acad. Sci. USA.* **35**, 605–608.

COX, D. R. and MILLER, H. D. [1968] *The Theory of Stochastic Processes*, Methuen, London.

CSÁKI, E. and VINCZE, I. [1961] On some problems connected with the Galton test, *Publ. Math. Inst. Hungarian Acad. Sci. Ser. A*, **6**, 97–109.

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

FELLER, W. [1968] *An Introduction to Probability Theory and its Applications*, 3rd. ed., Wiley, New York.

GNEDENKO, B. V. and KOROLYUK, V. S. [1951] On the maximum discrepancy between two empirical distributions (Russian), *Dokl. Akad. Nauk SSSR.* **80**, 525–528 (*Sel. Transl. Math. Statist. Prob.* 1961, 13–16).

GNEDENKO, B. V. and RVACEVA, E. L. [1952] On a problem of the comparison of two empirical distributions (Russian), *Dokl. Akad. Nauk SSSR.* **82**, 513–516 (*Sel. Transl. Math. Statist. Prob.* 1961, 69–72).

KANWAR SEN [1965] On some combinatorial relations concerning the symmetric random walk, *Publ. Math. Inst. Hungarian Acad. Sci.* **9**, 335–357.

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

MIHALEVIC, V. S. [1952] On the mutual disposition of two empirical distribution functions (Russian), *Dokl. Akad. Nauk SSSR.* **85**, 485–488 (*Sel. Transl. Math. Statist. Prob.* 1961, 63–68).

MOHANTY, S. G. [1979] *Lattice Path Counting and Applications*, Academic Press, New York.

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

ŠIDAK, Z. [1973] Applications of random walks in non-parametric statistics, *Bull. Inst. Internat. Statist., Proc. 39. Session* **45**, Book 3, 34–43.

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

DEPARTMENT OF STATISTICS
UNIVERSITY OF ECONOMICS
AUGASSE 2–6
A-1090 VIENNA, AUSTRIA

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF ECONOMICS
AUGASSE 2–6
A-1090 VIENNA, AUSTRIA