On The Number of Times where a Simple Random Walk reaches a Nonnegative Height

Katzenbeisser, Walter; Panny, Wolfgang

On The Number of Times where a simple Random Walk reaches a nonnegative Height

Walter Katzenbeisser, Wolfgang Panny

Institut für Statistik
Wirtschaftsuniversität Wien

Forschungsberichte

Bericht 59
September 1998

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

ON THE NUMBER OF TIMES WHERE A SIMPLE RANDOM WALK REACHES A NONNEGATIVE HEIGHT

W. KATZENBEISSER and W. PANNY

University of Economics, Vienna

The purpose of this note is to generalize the distribution of the local time of a purely binomial random walk for simple random walks allowing for three directions with different probabilities.

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

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

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

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

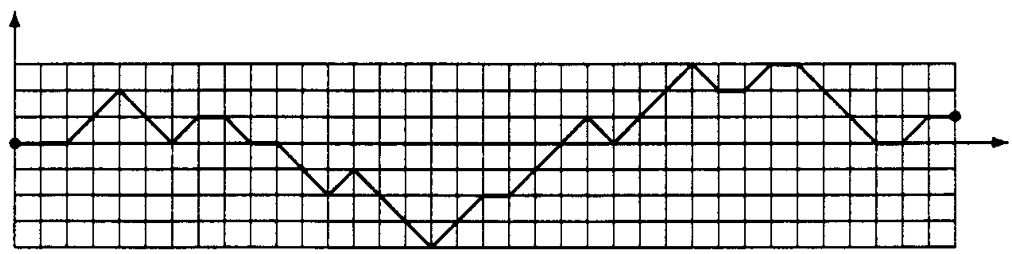
i.e. a simple random walk in the sense of Cox and Miller (1965) starting at $\mathbf{S}_0$ and leading to ℓ after n steps. Confining to $\mathbf{S}_0 = 0$ actually constitutes no restriction at all. So this assumption will be made in the sequel.

The purpose of this paper is to discuss distributional properties of the random variable

$$\mathbf{N}_n(r) = [\text{the number of visits to height } r]\,,$$

with $r \geq 0$; of course, for $r = 0$, $\mathbf{N}_n(0)$ counts the number of times where a simple random walk visits the origin.

A *visit to height* r occurs if $\mathbf{S}_k = \mathbf{S}_{k+1} = \mathbf{S}_{k+2} = \ldots = \mathbf{S}_{k+m} = r$, and $\mathbf{S}_{k-1} \neq r$, $\mathbf{S}_{k+m+1} \neq r$ for $0 \leq k \leq k+m \leq n$. 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 height* r. If $\mathbf{S}_0 = r$ then a *visit to height* r begins at the origin by definition. Correspondingly, if $\mathbf{S}_n = r$ then a *visit to height* r terminates at the end-point by definition. The following figure shows a sample path with $\mathbf{N}_n(0) = 6$, $\mathbf{N}_n(1) = 7$, $\mathbf{N}_n(2) = 4$, $\mathbf{N}_n(3) = 2$, and $\mathbf{S}_n = 1$:



In a recent paper Katzenbeisser and Panny (1996) derived the joint probability distribution $\mathbf{P}(\mathbf{N}_n(r) > k, \mathbf{S}_n = \ell)$, with ℓ *arbitrary* but fixed: For all $k \geq 0$ and $r \geq 0$ we have

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

where $\rho = \alpha/\gamma$ and generalized trinomial coefficients (GTC) are used. They have generating function $(\alpha v^2+\beta v+\gamma)^n$, i.e.

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

where $[v^k]P(v)$ denotes the coefficient of v^k in $P(v)$; a probabilistic interpretation of the GTC's is given by

$$\mathbf{P}(\mathbf{S}_n = \ell) = \binom{n;\alpha,\beta,\gamma}{n+\ell} \quad (2)$$

for all admissible ℓ, i.e. the probability that an unrestricted simple random walk reaches the state ℓ after n steps and $\ell \in \{-n, -n+1, \dots, n\}$. Further properties of the GTC's can be found in Panny(1984) or Böhm (1993).

For the *symmetric random walk* with $\alpha = \gamma$ and therefore $\rho = 1$ we get from (1)

$$\mathbf{P}(\mathbf{N}_n(r) > k, \mathbf{S}_n = \ell) = 2^k \sum_{j\geq 0} \binom{-k}{j} \binom{n;\alpha,\beta,\alpha}{n-r-|\ell-r|-2k-2j}, \quad (3)$$

which further can be specialized for the *purely binomial random walk* with $\alpha = \gamma = 1/2$ and therefore $\beta = 0$:

$$\mathbf{P}(\mathbf{N}_n(r) > k, \mathbf{S}_n = \ell) = 2^{k-n} \binom{n-k}{\frac{n-r-|\ell-r|}{2} - k}. \quad (4)$$

This result follows from the fact that the GTC's and the ordinary binomial coefficients are related by

$$\binom{n;1/2,0,1/2}{2m} = \binom{n}{m} 2^{-n}, \quad (5)$$

cf. Panny (1984) and an application of Vandermonde's convolution formula. Obviously, by means of formulae (1) and (2) the conditional distribution $\mathbf{P}(\mathbf{N}_n(r) > k | \mathbf{S}_n = \ell)$ can be derived; specializations for $r = 0, \ell = 0$ are discussed in Katzenbeisser and Panny (1984); moreover, substituting $n \to 2n$ we obtain for the *purely binomial random walk* the well known result due to Dwass (1967):

$$\mathbf{P}(\mathbf{N}_{2n}(0) > k | \mathbf{S}_{2n} = 0) = 2^k \frac{\binom{2n-k}{n-k}}{\binom{2n}{n}}. \quad (6)$$

$\mathbf{N}_{2n}(0)|\mathbf{S}_{2n}=0$ counts the number of times where a purely binomial random walk terminating at zero after n steps returns to the origin; the moments of $\mathbf{N}_{2n}(0)>k|\mathbf{S}_{2n}=0$ are discussed in Katzenbeisser and Panny (1986) and Kemp (1987). Statistical properties of an alternative test to the Kolmogorov-Smirnov two sample test with equal sample sizes, based on $\mathbf{N}_{2n}(0)|\mathbf{S}_{2n}=0$ are discussed in Katzenbeisser and Hackl (1986).

The basic tool in deriving formula (1) is a generating function $\Phi(k,\ell,r;y)$ for the joint probabilities $\mathbf{P}(\mathbf{N}_n(r)>k,\mathbf{S}_n=\ell)$ given by

$$\begin{aligned}\Phi(k,\ell,r;y)&=\sum_{n\geq 0}\mathbf{P}(\mathbf{N}_n(r)>k,\mathbf{S}_n=\ell)y^n=\\&=2^k\frac{\rho^{r+\frac{|\ell-r|+(\ell-r)}{2}}}{\gamma}\frac{v^{r+|\ell-r|}}{1-\rho v^2}(\alpha v^2+\beta v+\gamma)\left(\frac{\rho v^2}{1+\rho v^2}\right)^k,\end{aligned}\tag{7}$$

where the substitution $y=g(v)=v/(\alpha v^2+\beta v+\gamma)$ has been used, which considerably simplifies the calculations. An explicit expression for $\mathbf{P}(\mathbf{N}_n(r)>k,\mathbf{S}_n=\ell)$ can be found by an application of Cauchy's integral formula

$$\mathbf{P}(\mathbf{N}_n(r)>k,\mathbf{S}_n=\ell)=\frac{1}{2\pi i}\oint\frac{\Phi(k,\ell,r;y)}{y^{n+1}}dy\,,$$

which yields after the change of variables $y=g(v)$

$$\mathbf{P}(\mathbf{N}_n(r)>k,\mathbf{S}_n=\ell)=\frac{1}{2\pi i}\oint\frac{\Phi(k,\ell,r;g(v))}{g^{n+1}(v)}g'(v)dv\,,$$

with

$$\frac{g'(v)}{g^{n+1}}=\gamma\frac{1-\rho v^2}{v^{n+1}}(\alpha v^2+\beta v+\gamma)^{n-1}\,,\tag{8}$$

for further technical details cf. Katzenbeisser and Panny (1996, p.324).

This generating function can also be used to derive distributional properties of $\mathbf{N}_n(r)$ irrespective, where the random walk terminates after n steps, which will be done in the sequel of this paper.

2. The distribution of $\mathbf{N}_n(r)$. Of course, the interesting marginal probabilities $\mathbf{P}(\mathbf{N}_n(r)>k)$ can be derived by summation over all admissible values for ℓ in the joint distribution (1) or, equivalently, in the generating function (7):

$$\begin{aligned}\sum_{\ell}\Phi(k,\ell,r;y)&=\sum_{\ell}\sum_{n\geq 0}\mathbf{P}(\mathbf{N}_n(r)>k,\mathbf{S}_n=\ell)y^n=\\&=\sum_{n\geq 0}\mathbf{P}(\mathbf{N}_n(r)>k)y^n:=\Phi(k,r;y)\,,\end{aligned}$$

and $\mathbf{P}(\mathbf{N}_n(r)>k)$ is given as $[y^n]\,\Phi(k,r;y)$. Therefore, to obtain $\Phi(k,r;y)$ we have to sum over all admissible values of ℓ, i.e. $-n\leq\ell\leq n$ in the corresponding generating function

(7). However, a careful investigation of the summation shows that the contribution of all terms with $\ell > n$ and $\ell < -n$ to this sum are zero;thus we can take

$$\sum_{\ell} \rho^{\frac{|\ell-r|+(\ell-r)}{2}} v^{|\ell-r|} = \frac{1}{1-\rho v} + \frac{v}{1-v}.$$

Therefore we get for the generating function $\Phi(k,r;y)$

$$\Phi(k,r;y) = 2^k \rho^r \frac{1}{\gamma} \frac{1}{1-\rho v^2} (\alpha v^2 + \beta v + \gamma) \Big(\frac{\rho v^2}{1+\rho v^2} \Big)^k v^r \left[\frac{1}{1-\rho v} + \frac{v}{1-v} \right], \tag{9}$$

and we only have to extract $[y^n]\,\Phi(k,r;y)$, where the substitution (8), i.e. $y = g(v) = v/(\alpha v^2 + \beta v + \gamma)$ has to be taken into account. Technically, $\mathbf{P}(\mathbf{N}_n(r) > k)$ can be obtained as

$$[v^n] \left\{ 2^k \rho^{k+r} v^{2k+r} (\alpha v^2 + \beta v + \gamma)^n \frac{1}{(1+\rho v^2)^k} \left[\frac{1}{1-\rho v} + \frac{v}{1-v} \right] \right\},$$

and is given in the following

Theorem 1. For all $r \geq 0$ and $k \geq 0$

$$\begin{aligned} \mathbf{P}(\mathbf{N}_n(r) > k) =& 2^k \rho^{k+r} \sum_{i \geq 0} \rho^i \binom{-k}{i} \times \\ & \sum_{j \geq 0} \left[\rho^j \binom{n;\alpha,\beta,\gamma}{n-2k-r-2i-j} + \binom{n;\alpha,\beta,\gamma}{n-2k-r-1-2i-j} \right]. \end{aligned} \tag{10}$$

For some special cases we get immediately: Let $r = k = 0$ then we find

$$\mathbf{P}(\mathbf{N}_n(0) > 0) = \sum_{j \geq 0} \left[\rho^j \binom{n;\alpha,\beta,\gamma}{n-j} + \binom{n;\alpha,\beta,\gamma}{n-1-j} \right],$$

and because of the quasi-symmetry property of the GTC's (Panny (1986)), i.e.

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

we obtain

$$\begin{aligned} \mathbf{P}(\mathbf{N}_n(0) > 0) =& \sum_{j \geq 0} \left[\binom{n;\alpha,\beta,\gamma}{n+j} + \binom{n;\alpha,\beta,\gamma}{n-1-j} \right] = \\ =& \mathbf{P}(\mathbf{S}_n \geq 0) + \mathbf{P}(\mathbf{S}_n \leq -1) = 1. \end{aligned}$$

Moreover, for arbitrary $r > 0$ and $k = 0$ we get

$$\mathbf{P}(\mathbf{N}_n(r) > 0) = \mathbf{P}(\mathbf{S}_n \geq r) + \rho^r \mathbf{P}(\mathbf{S}_n \leq -r-1).$$

Finally, for the simple special case $r = n$ and $k = 0$ we find from (10)

$$\mathbf{P}(\mathbf{N}_n(n) > 0) = \rho^n \binom{n;\alpha,\beta,\gamma}{0} = \alpha^n ,$$

which is the probability that a simple random walk moves n consecutive steps upwards.

Formula (10) specializes for the *symmetric random* walk to

$$\mathbf{P}(\mathbf{N}_n(r) > k) = 2^k \sum_{i\geq 0} \binom{-k}{i} \sum_{j\geq 0} \left[\binom{n;\alpha,\beta,\alpha}{n-2k-r-2i-j} + \binom{n;\alpha,\beta,\alpha}{n-2k-r-1-2i-j} \right] . \tag{11}$$

Moreover, for the *purely binomial random walk* we find

$$\mathbf{P}(\mathbf{N}_n(r) > k) = 2^{k-n} \sum_{j\geq 0} \binom{n-k}{\lfloor \frac{n-r-j}{2} \rfloor - k} , \tag{12}$$

which follows from formulae (11), (5), and repeated applications of Vandermonde's convolution formula. From (12) we get for example for all $k \geq 0$

$$\mathbf{P}(\mathbf{N}_n(r) = k+1) = \begin{cases} 2^{k-n} \binom{n-k}{\frac{n+r}{2}} , & n-r \quad \text{even} \\ 2^{k-n+1} \binom{n-1-k}{\frac{n+r-1}{2}} , & n-r \quad \text{odd} , \end{cases}$$

which specializes for $r = 0$:

$$\mathbf{P}(\mathbf{N}_n(0) = k+1) = 2^{k-2\lfloor \frac{n}{2} \rfloor} \binom{2\lfloor \frac{n}{2} \rfloor - k}{\lfloor \frac{n}{2} \rfloor} ,$$

showing that

$$\mathbf{P}(\mathbf{N}_n(0) = k+1) = \mathbf{P}(\xi(0,n) = k) ,$$

where $\xi(x,n) = \#[k : 0 < k \leq n, \mathbf{S}_k = x]$, is the *local time* of the *purely binomial random walk*, cf. Revesz (1990, p.95, Theorem 9.3).

3. The Moments of $\mathbf{N}_n(r)$. In principle, the generating function $\Phi(k,r;y)$ can also be used to obtain the moments of the random variable $\mathbf{N}_n(r)$. Multiplication of $\Phi(k,r;y)$ by $(k+1)^s - k^s$ and summation over all admissible $k \geq 0$ in (9) leads to a generating function

$$\begin{aligned} \Phi(r;s;y) &= \sum_{k\geq 0} [(k+1)^s - k^s] \Phi(k,r;y) = \sum_{n\geq 0} \sum_{k\geq 0} [(k+1)^s - k^s] \mathbf{P}(\mathbf{N}_n(r) > k) y^n = \\ &= \sum_{n\geq 0} \mathbf{E}(\mathbf{N}_n^s(r)) y^n , \end{aligned}$$

and the s-th moment of $\mathbf{N}_n(r)$ is given as $[y^n]\Phi(r;s;y)$.

Specializing on $s = 1$ we get the generating function for the expectation of $\mathbf{N}_n(r)$:

$$\Phi(r;1;y) = \sum_{k\geq 0} \Phi(k,r;y) = \sum_{n\geq 0}\sum_{k\geq 0} \mathbf{P}(\mathbf{N}_n(r) > k) y^n = \sum_{n\geq 0} \mathbf{E}(\mathbf{N}_n(r)) y^n .$$

Again, a careful investigation of the summation necessary shows that we can take

$$\sum_{k\geq 0} \left(\frac{2\rho v^2}{1+\rho v^2}\right)^k = \frac{1+\rho v^2}{1-\rho v^2}$$

and we get, after taking the substitution (8) into account:

$$\mathbf{E}(\mathbf{N}_n(r)) = [v^n] \left\{ \rho^r v^r (\alpha v^2 + \beta v + \gamma)^n (1+\rho v^2) \frac{1}{1-\rho v} \frac{1}{1-v} \right\} .$$

Extracting this coefficient yields the expression for the expectation of $\mathbf{E}(\mathbf{N}_n(r))$:

Theorem 2. The expectation of $\mathbf{N}_n(r)$ is given by

$$\begin{aligned} \mathbf{E}(\mathbf{N}_n(r)) &= \rho^r \sum_{i\geq 0} \rho^i \sum_{j\geq 0} \left[\binom{n;\alpha,\beta,\gamma}{n-r-i-j} + \rho \binom{n;\alpha,\beta,\gamma}{n-r-2-i-j} \right] = \\ &= \rho^r \binom{n;\alpha,\beta,\gamma}{n-r} + \frac{\rho^r(1+\rho)}{1-\rho} \sum_{j\geq 1} (1-\rho^j) \binom{n;\alpha,\beta,\gamma}{n-r-j} . \end{aligned} \tag{13}$$

Using the quasi-symmetry property of the GTC's, we have for the special case $r = 0$

$$\begin{aligned} \mathbf{E}(\mathbf{N}_n(0)) &= \binom{n;\alpha,\beta,\gamma}{n} + \frac{1+\rho}{1-\rho} \sum_{j\geq 1} \left[\binom{n;\alpha,\beta,\gamma}{n-j} - \binom{n;\alpha,\beta,\gamma}{n+j} \right] = \\ &= \mathbf{P}(\mathbf{S}_n = 0) + \frac{1+\rho}{1-\rho} \left[\mathbf{P}(\mathbf{S}_n \leq -1) - \mathbf{P}(\mathbf{S}_n \geq 1) \right] . \end{aligned}$$

Of course, equation (13) can be specialized for the symmetric- as well as for the purely binomial random walk.

For the *symmetric random walk* we get from (13)

$$\mathbf{E}(\mathbf{N}_n(r)) = \binom{n;\alpha,\beta,\alpha}{n-r} + 2 \sum_{j\geq 1} j \binom{n;\alpha,\beta,\alpha}{n-r-j} . \tag{14}$$

For the special case $r = n$ we obtain immediately from (13) and (14): $\mathbf{E}(\mathbf{N}_n(n)) = \alpha^n$.

For the *purely binomial random walk* formula (14) specializes :

$$\mathbf{E}(\mathbf{N}_n(r)) = \begin{cases} 2^{-n} \left[\binom{n}{\frac{n-r}{2}} + 4 \sum_{j\geq 1} j \binom{n}{\frac{n-r}{2}-j} \right] , n-r \quad \text{even} \\ 2^{-n} \sum_{j\geq 0} (4j+2) \binom{n}{\frac{n-r-1}{2}-j} \ , n-r \quad \text{odd} \end{cases} \tag{15}$$

which yields for $r = n$: $\mathbf{E}(\mathbf{N}_n(n)) = (\frac{1}{2})^n$.

Moreover for the special case $r = 0$ we can further simplify formula (15). Because of $\sum_{j=0}^{\alpha} j\binom{2\alpha}{\alpha-j} = \alpha\binom{2\alpha-1}{\alpha}$ for all integers α, and $\sum_{j=0}^{\alpha} j\binom{2\alpha+1}{\alpha-j} = (2\alpha+1)\binom{2\alpha-1}{\alpha} - 2^{2\alpha-1}$ for $\alpha = 1, 2, \ldots$, cf. Riordan (1968, p.34) we can rewrite (15):

$$\mathbf{E}(\mathbf{N}_n(0)) = \begin{cases} (1+n)2^{-n}\binom{n}{\frac{n}{2}} & , n \text{ even}, \\ n2^{-n+2}\binom{n-2}{\frac{n-1}{2}} & , n \text{ odd}, n \geq 3. \end{cases}$$

REFERENCES

BÖHM, W. [1993] *Markovian Queuing Systems in Discrete Time*, Mathematical Systems in Economics, Vol.137, Hain, Frankfurt/Main.

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

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

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

KATZENBEISSER, W. and PANNY, W. [1986] A note on the higher moments of the random variable T associated with the number of returns of a simple random walk, *Adv.Appl.Prob.*, **18**, 279-282

KATZENBEISSER, W. and HACKL, P. [1986] An alternative to the Kolmogorov-Smirnov two-sample test, *Comm.Statist.-Theor.Meth.*, **15**(4), 1163-1177

KATZENBEISSER, W. and PANNY, W. [1996] Simple random walk statistics. Part I: Discrete time results, *J.Appl.Prob.*, **33**, 311-330.

KEMP, A. W. [1987] The moments of the random variable for the number of returns of a simple random walk. *Adv.Appl.Prob.*, **19**, 505-507.

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.

REVESZ, P. [1990] *Random Walk in Random and Non-Random Environments*, World Scientific, Singapore.

RIORDAN, J. [1968] *Combinatorical Identities*, J.Wiley&Sons, New York.