Journal of the Korean Statistical Society (2001), **30**: *3, pp 529–539*

Some Distribution Results on Random Walk with Unspecified Terminus

Jagdish Saran[1] and Sarita Bansal[2]

ABSTRACT

This paper deals with the distributions of certain characteristics related to a symmetric random walk of n steps ending at an unspecified position, thus generalizing and extending the earlier work.

Keywords: Fluctuations of partial sums $\Sigma \pm 1$ when total is unknown; Random walk with unspecified terminus; Wave; Section; Trail; Tail; Sojourn; Crossing; Upcrossing; Run of returns; Run of positive returns; Run of crossings; Run of upcrossings; Generating function; Unconditional ballot problem.

1. INTRODUCTION

Consider a symmetric random walk of n steps, n being odd or even, ending at an unspecified position. This random walk, for even n, was considered by Chung and Feller(1949), Feller (1957),Csáki and Vincze(1963) and Jain(1966). Again, Aneja(1975, Ch. V), Csáki (1961), Engelberg(1964), Sen(1964,1969), Kaul(1982, Ch. V), Saran and Sen(1981) and Kaul and Sen(1983) treated independently the symmetric random walk of n steps ending at an unspecified position. In this paper we consider the above mentioned symmetric random walk of n steps, n being odd or even, ending at an unspecified position and obtain the distributions of certain characteristics of this random walk, thus generalizing and extending the earlier work.

Consider a sequence of n independent random variables $\{W_i\}$ having a common distribution $P(W_i = +1) = P(W_i = -1) = \frac{1}{2}$ which generates the sequence of partial sums $\{S_j\}$ where $S_0 = 0, S_j = \sum_{i=1}^{j} W_i$ in such a way that the final sum S_n of all the n variate values is unknown. Obviously there are 2^n possible arrays of $(+1)'s$ and $(-1)'s$ each with probability 2^{-n}. The sequence of random

[1]Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi-110007, INDIA, *Email: dustats@del3.vsnl.net.in*

[2]Department of Statistics, PGDAV College, Nehru Nagar, New Delhi-110065, INDIA

variables $\{S_j\}$ defines a random walk on a one - dimensional lattice which can be represented graphically by a polygonal line whose j^{th} side has slope ± 1 and whose j^{th} vertex has the ordinate S_j. It may be noted that i and S_i are always in parity.

It may be pointed out that the results obtained in this paper also have an interpretation in terms of the "unconditional" ballot problem considered by Engelberg(1964) and saran and Sen(1981) formulated below: Suppose that in a ballot with two candidates A and B, the total number of votes equals m and that all 2^m possible arrangements of counting are equally probable (here the number of votes scored by individual candidates are not known). In this context a random walk path ending at an unspecified position can be interpreted as a vote sequence.

For convenience of presentation, we introduce the following symbols:

I. $T_r^+(T_r^-)$ = a positive (negative) trail
= the segment of the path $\{S_j\}$ starting from the x-axis, say at the i^{th} step, and ending at an unspecified position after the n-th step such that $S_i = 0, S_j \geq 0 (S_j \leq 0)$, $j = i+1, \ldots, n$.

II. $T^+(T^-)$ = a positive (negative) tail
= the segment of the path $\{S_j\}$starting from the x-axis at the i^{th} step and ending at an unspecified position after the n^{th} step without touching the x-axis in between such that $S_i = 0$ and $S_j > 0 (S_j < 0), j = i+1, \ldots, n$. (It is obvious that a $T^+(T^-)$ is a $T_r^+(T_r^-)$ but the converse is not true).

III. $T_{re}^+(T_{re}^-)$ = a $T_r^+(T_r^-)$ of even length.

IV. $T_{r0}^+(T_{r0}^-)$ = a $T_r^+(T_r^-)$ of odd length.

V. $T_e^+(T_e^-)$ = a $T^+(T^-)$ of even length.

VI. $T_0^+(T_0^-)$ = a $T^+(T^-)$ of odd length.

VII. $N_n(u)$ = the number of sojourns at height u.
= the number of returns at height u.
= the number of indices i for which $S_i = u, u = 0, 1, 2, \ldots$.

VIII. $N_n^+(u), N_n^-(u)$ = the number of positive and negative sojourns or returns at height u, respectively. Let $0 < i_1 < i_2 < \ldots$ be the indices for which $S_j = u$. Then the sojourn from i_{k-1} to i_k is called positive or negative according as $S_j > u$ or $S_j < u$ for $i_{k-1} < j < i_k$.

IX. $N_n^*(u)$ = the number of crossings of height u.
= the number of indices i for which $S_i = u$ and $(u - S_{i-1})(u - S_{i+1}) < 0, u = 0, 1, 2, \ldots$.

X. $N_n^{+*}(u)$ = the number of upcrossings of height u.
= the number of indices i for which $S_i = u + 1$ and $S_{i-1} = u, u = 0, 1, 2, \ldots$.

XI. $R_n(u)$ = the number of runs of returns of height u of type VII whose number is $N_n(u)$.
= the number of sequences of (consecutive) returns of height u with indices increasing by 2, i.e., the sequence $(i_k, i_{k+1}, \ldots, i_l)$ of return indices of height u will be said to form a run of returns at height u if

(i) $i_j - i_{j-1} = 2, \;\; j = k+1, k+2, \ldots, l$,

(ii) $i_k > i_{k-1} + 2$, and

(iii) $i_{l+1} > i_l + 2$.

XII. $R_n^+(u)$ = the number of runs of positive returns of height u of type VIII whose number is $N_n^+(u)$.
= the definition XI with 'positive return' in place of 'return'.

XIII. $R_n^*(u)$ = the number of runs of crossings of height u of type IX whose number is $N_n^*(u)$.
= the definition XI with 'crossing' in place of 'return'.

XIV. $R_n^{+*}(u)$ = the number of runs of upcrossings of height u of type X whose number is $N_n^*(u)$.
= the definition XI with 'upcrossing' in place of 'return'.

A 'wave' (V) is defined as the segment between two consecutive returns to the origin or x-axis; the segment between the origin and the first return point is also regarded as a wave. A wave V lying above(below) the x-axis is called a positive (negative) wave and is denoted by the symbol $V^+(V^-)$. A 'Section' (W) is defined as the path segment between two consecutive crossings of the x-axis. A section W lying above (below) the x – axis is called a positive (negative) section and is denoted by the symbol $W^+(W^-)$.

2. SOME RELEVANT GENERATING FUNCTIONS (GF's)

The following expansion due to Dwass(1967, Appendix (16)) will be found useful in determining the GF's considered in this section:

$$\frac{\left[\frac{1}{2}\left\{1-(1-4t)^{\frac{1}{2}}\right\}\right]^{k-1}}{(1-4t)^{\frac{1}{2}}}=\sum_{n=k-1}^{\infty}\binom{2n-k+1}{n-k+1}t^n. \tag{2.1}$$

Some of the results we quote below appear in Feller (1968), Kaul(1982, Ch. V) and Kaul and Sen(1983) and the rest are easily derivable from elementary considerations. The following list covers what is needed in the sequel.

(i) Let $G(T_e^+)$ and $G(T_e^-)$ denote, respectively, the GF of a positive tail of even length (say, $2k$) and that of a negative tail of even length(say, $2k$). Then

$$G(T_e^+)=G(T_e^-)=\begin{cases} s^2/2\phi(s)(1-\phi(s)), & k\geq 0\\ \phi(s)/2(1-\phi(s)), & k\geq 1,\end{cases}$$

where $\phi(s)=1-(1-s^2)^{1/2}$.

Proof: The number of paths of $2k(k\geq 0)$ steps relevant to a positive tail $T_e^+$of even length considered above is the same as the number of paths with origin shifted to $(1,1)$ such that $S_j\geq 0, j\geq 1$ which by Sen(1969) is $\binom{2k-1}{k-1}$.

Hence

$$\begin{aligned} G(T_e^+) &= \sum_{k=0}^{\infty}\binom{2k-1}{k-1}t^{2k}\\ &= \sum_{k=0}^{\infty}\binom{2k-1}{k-1}\left(\frac{s}{2}\right)^{2k}, \text{by substituting } t=\tfrac{s}{2}\\ &= \left(\frac{s}{2}\right)^2\sum_{k=0}^{\infty}\binom{2k-1}{k-1}\left(\frac{s}{2}\right)^{2(k-1)}\\ &= \left(\frac{s}{2}\right)^2\sum_{m=-1}^{\infty}\binom{2m+1}{m}\left(\frac{s}{2}\right)^{2m}\end{aligned}$$

$$= \frac{2(s/2)^2}{\phi(s)(1-\phi(s))}, \text{by (2.1)}$$
$$= \frac{s^2}{2\phi(s)(1-\phi(s))}.$$

Similary when $k \geq 1$, the GF of the positive tail T_e^+ of even length is given by

$$\begin{aligned} G(T_e^+) &= \sum_{k=1}^{\infty} \binom{2k-1}{k-1} s^{2k} \\ &= \sum_{k=1}^{\infty} \binom{2k-1}{k-1} s^{2k} - 1 \\ &= \frac{s^2}{2\phi(s)(1-\phi(s))} - 1 \quad \text{(from above)} \\ &= \frac{\phi(s)}{2(1-\phi(s))}. \end{aligned}$$

The result that $G(T_e^+) = G(T_e^-)$ follows from symmetry.

(ii) the GF's $G(T_0^+)$ and $G(T_0^-)$ of positive and negative tails of odd length $2k+1 (k \geq 0)$ are

$$G(T_0^+) = G(T_0^-) = \frac{s}{2(1-\phi(s))}.$$

(iii) the GF's $G(T_{re}^+)$ and $G(T_{re}^-)$ of positive and negative trails of even length 2k are

$$G(T_{re}^+) = G(T_{re}^-) = \begin{cases} \frac{1}{(1-\phi(s))}, & k \geq 0 \\ \frac{\phi(s)}{(1-\phi(s))}, & k \geq 1. \end{cases}$$

(iv) the GF's $G(T_{r0}^+)$ and $G(T_{r0}^-)$ of positive and negative trails of odd length $2k+1 (k \geq 0)$ are

$$G(T_{r0}^+) = G(T_{r0}^-) = \frac{\phi(s)}{s(1-\phi(s))}.$$

(v) the GF of the time to reach height k for the first time is $\left\{\frac{\phi(s)}{s}\right\}^k, k = 1, 2, \ldots$.
(vi) the GF of a wave V (i.e., V^+ or V^-) is $\phi(s)$.
(vii) the GF of a positive (negative) wave $V^+(V^-)$ of length 2 steps is $\frac{s^2}{4}$.

(viii) the GF of a positive (negative) wave $V^+(V^-)$ of length > 2 steps is $\frac{\phi^2(s)}{4}$.
(ix) The GF of a section $W(i.e., W^+$ or $W^-)$ is $\frac{\phi^2(s)}{s^2}$.
(x) the GF of a positive (negative) section $W^+(W^-)$ of length 2 steps is $\frac{s^2}{4}$.
(xi) The GF of a positive(negative) section $W^+(W^-)$ of length > 2 steps is $\frac{\phi^2(s)}{s^2} - \frac{s^2}{4}$.

3. DISTRIBUTIONS OF $N_m(a)$ AND $R_m(a)$

Theorem 3.1. *When m and $a(> 0)$ are in parity, we have*

$$P(N_m(a) = r, R_m(a) = k) = 2^{r-m-1} \binom{r-2}{k-1} \binom{m-2r+3}{\frac{m-a}{2} - k - r + 2} \quad (3.1)$$

and

$$P(N_m(a) = r) = 2^{r-m-1} \binom{m-r+1}{\frac{m+a}{2}}. \quad (3.2)$$

Proof: We shall first prove (3.1) and (3.2) for the case when m and a are both odd. Let $m = 2n+1$. Consider a random walk $\{S_j; S_0 = 0, S_j = \sum_{i=1}^{j} W_i, j \geq 1\}$ denoted by OMDE from O(0,0) to $E(2n+1, 2c\pm1), c = 0, \pm1, \pm2, \ldots, \pm n$, where M and D are the first and the last return points of height a, respectively. Then the path $\{S_j\}$ is divided into three segments, viz. from O to M, M to D and D to E. The first segment from O to M is a first passage to height a with GF $(\frac{\phi(s)}{s})^a$, by (v) of Section 2. The last segment from D to E involves $2t(t \geq 0)$ steps if the ordinate of the point E, viz. 2c+1, is ≥ 0 and is such that it does not return to the horizontal MD or it involves $2t(t \geq 1)$ steps if the ordinate of the point E is < 0 and is such that it does not return to the horizontal MD. Thus the segment DE is a positive tail of even length $2t(t \geq 0)$ or negative tail of even length $2t(t \geq 1)$ with GF's $\frac{s^2}{2\phi(s)(1-\phi(s))}$ and $\frac{\phi(s)}{2(1-\phi(s))}$, respectively, by (i) of Section 2. The path segment MD consists of r+1 waves. The waves comprising a run of returns are of length 2 each and a wave between any two consecutive runs of returns will be of length > 2. But the wave in the beginning of the path segment MD may be of any length. Thus the path segment MD consists of the following:
(a) one wave in the beginning with GF $\phi(s)$,
(b) $(r-k-1)$ waves (V) each of length 2 and each with GF $\frac{s^2}{2}$, by (vii) of Section 2,
(c) $(k-1)$ waves (V) each of length > 2 and each with GF $\frac{\phi^2(s)}{2}$, by (viii) of Section 2.

The $(r-k-1)$ waves(V) each of length 2 are to be combined with $(k-1)$ waves (V) each of length >2, so as to form k runs of total $r-1$ returns which is possible in $\begin{pmatrix} r-k-1+k-1 \\ r-k-1 \end{pmatrix} = \begin{pmatrix} r-2 \\ k-1 \end{pmatrix}$ ways. Therefore, the generating function(GF) of the path entailing r returns of height a and k runs of returns of height a is given by

$$G(s; N_{2n+1}(a)=r, R_{2n+1}(a)=k) = \sum_n P(N_{2n+1}(a)=r, R_{2n+1}(a)=k)s^{2n+1}$$

$$= \begin{pmatrix} r-2 \\ k-1 \end{pmatrix} 2^{r-1} \left(\frac{2}{s}\right)^{a+2k-2r+2} \left(\frac{\phi(s)}{2}\right)^{a+2k-1} \frac{1}{1-\phi(s)} . \tag{3.3}$$

Summation of (3.3) over k from 1 to $r-1$ yields

$$G(s; N_{2n+1}(a)=r) = \left(\frac{\phi(s)}{s}\right)^a \phi^{r-1}(s) \frac{1}{1-\phi(s)} . \tag{3.4}$$

The coefficients of s^{2n+1} in the expansions of (3.3) and (3.4) obtained with the help of (2.1) lead to (3.1) and (3.2), respectively, for the case when m and a are both odd. Likewise (3.1) and (3.2) can be established when both m and a are even.

Theorem 3.2. *When m and $a(>0)$ are not in parity, we have*

$$P(N_m(a)=r, R_m(a)=k) = 2^{r-m} \begin{pmatrix} r-2 \\ k-1 \end{pmatrix} \begin{pmatrix} m-2r+2 \\ \frac{m-a+3}{2}-k-r \end{pmatrix} \tag{3.5}$$

and

$$P(N_m(a)=r) = 2^{r-m} \begin{pmatrix} m-r \\ \frac{m+a-1}{2} \end{pmatrix} . \tag{3.6}$$

Proof: Let us consider the case when m is even and a is odd. Let $m=2n$ and let OMDE be a random walk $\{S_j\}$ as envisaged on the left hand side of (3.5), where M and D denote, respectively, the first and the last return points of height a . We observe that the GF's of the segments OM and MD are of the same form as those of the corresponding segments in Theorem 3.1. But the segment DE of the path is either a T_0^+ or a T_0^- each with GF $\frac{s}{2(1-\phi(s))}$, by (ii) of Section 2. Arguing in a similar manner as in Theorem 3.1, it can be shown that

$$G(s; N_{2n}(a)=r, R_{2n}(a)=k)$$
$$= \begin{pmatrix} r-2 \\ k-1 \end{pmatrix} 2^r \left(\frac{2}{s}\right)^{a-2r+2k+1} \left(\frac{\phi(s)}{2}\right)^{a+2k-1} \frac{1}{1-\phi(s)} . \tag{3.7}$$

Summing (3.7) over k from 1 to $r-1$ gives

$$G(s;N_{2n}(a)=r)=2^r\left(\frac{2}{s}\right)^{a-1}\left(\frac{\phi(s)}{2}\right)^{a+r-1}\frac{1}{1-\phi(s)}. \tag{3.8}$$

The coefficients of s^{2n} in the expansions of (3.7) and (3.8) yield, respectively, (3.5) and (3.6) for the case when m is even and a is odd. The other case when m is odd and a is even follows similarly.

In a similar manner, the following result can easily be established.

Theorem 3.3. *For $a=0$,*

$$P(N_m(0)=r,R_m(0)=k)$$
$$=\begin{cases} 2^{r-m+1}\begin{pmatrix} r-1 \\ k-1 \end{pmatrix}\begin{pmatrix} m-2r \\ \frac{m+1}{2}-k-r \end{pmatrix}, & \text{for } m \text{ odd} \\ 2^{r-m}\begin{pmatrix} r-1 \\ k-1 \end{pmatrix}\begin{pmatrix} m-2r+1 \\ \frac{m}{2}-r-k+1 \end{pmatrix}, & \text{for } m \text{ even} \end{cases} \tag{3.9}$$

and

$$P(N_m(0)=r)=\begin{cases} 2^{r-m+1}\begin{pmatrix} m-r-1 \\ \frac{m-1}{2} \end{pmatrix}, & \text{for } m \text{ odd} \\ 2^{r-m}\begin{pmatrix} m-r \\ \frac{m}{2} \end{pmatrix}, & \text{for } m \text{ even.} \end{cases} \tag{3.10}$$

It may be noted that the results (3.2), (3.6) and (3.10) are in agreement, respectively, with the results (1), (8) and (15) of Saran and Sen(1981). They also verify, respectively, the results [(17),(60)], [(32), (46)] and [(32),(60)] of Kaul (1982, Ch. V).

The distributions of the vectors $(N_m^*(a),R_m^*(a))$, $(N_m^{+*}(a),R_m^{+*}(a))$ and $(N_m^+(a), R_m^+(a))$ can easily be obtained by using similar arguments as used above and are quoted in the next section.

4. OTHER DISTRIBUTION RESULTS

(i) When m and $a(>0)$ are in parity, we have

$$P(N_m^*(a)=r,R_m^*(a)=k)=P(N_m^{+*}(a)=r,R_m^{+*}(a)=k)$$

$$= 2^{-m+1}\binom{r-1}{k-1}\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^{k-i-1}\binom{m+2i-2r+2}{\frac{m-a}{2}-r}, \quad (4.1)$$

$$P(N_m^*(a)=r) = P(N_m^{+*}(a)=r) = 2^{-m+1}\binom{m}{\frac{m+a}{2}+r}, \quad (4.2)$$

$$P(N_m^+(a)=r, R_m^+(a)=k) = 2^{-m}\binom{r-1}{k-1}\sum_{i=0}^{k-1}\binom{k-1}{i}$$

$$\cdot\left[\binom{m+i-2r+2}{\frac{m-a}{2}-k-r+1} + \binom{m+i-2r+1}{\frac{m-a}{2}-k-r}\right], \quad (4.3)$$

and

$$P(N_m^+(a)=r) = 2^{-m}\sum_{i=0}^{r-1}\binom{r-1}{i}\left[\binom{m+1-r}{\frac{m+a}{2}+i+1} + \binom{m-r}{\frac{m+a}{2}+i+1}\right]. \quad (4.4)$$

It may be noted that the first part of (4.2) verifies results (6) and (27) of Kaul and Sen(1983).

(ii) When m and $a(>0)$ are not in parity, we have

$$P(N_m^*(a)=r, R_m^*(a)=k) = P(N_m^{+*}(a)=r, R_m^{+*}(a)=k)$$

$$= 2^{-m}\binom{r-1}{k-1}\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^{k-i-1}\binom{m+2i-2r+3}{\frac{m-a+1}{2}-r}, \quad (4.5)$$

$$P(N_m^*(a)=r) = P(N_m^{+*}(a)=r) = 2^{-m}\binom{m+1}{\frac{m+1+a}{2}+r}, \quad (4.6)$$

$$P(N_m^+(a)=r, R_m^+(a)=k) = 2^{-m}\binom{r-1}{k-1}\sum_{i=0}^{k}\binom{k}{i}\binom{m+i-2r+1}{\frac{m-a+1}{2}-k-r}, \quad (4.7)$$

and

$$P(N_m^+(a)=r) = 2^{-m}\sum_{i=0}^{r-1}\binom{r-1}{i}\left[\binom{m-r}{\frac{m+a+1}{2}+i} + \binom{m-r+1}{\frac{m+a+3}{2}+i}\right]. \quad (4.8)$$

It may be noted that the first part of (4.6) is in agreement with results (14) and (23) of Kaul and Sen(1983).

(iii) When $a = 0$

$$P(N_m^*(0) = r, R_m^*(0) = k) = \begin{cases} 2^{-m} \binom{r-1}{k-1} \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{k-i-1} \binom{m+2i-2r+3}{\frac{m}{2}-r+1}, & \text{for } m \text{ odd} \\ 2^{-m+1} \binom{r-1}{k-1} \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{k-i-1} \binom{m+2i-2r+2}{\frac{m+1}{2}-r}, & \text{for } m \text{ even} \end{cases} \tag{4.9}$$

and

$$P(N_m^*(0) = r) = \begin{cases} 2^{-m+1} \binom{m}{\frac{m+1}{2}+r}, & \text{for } m \text{ odd} \\ 2^{-m+2} \binom{m-1}{\frac{m}{2}+r}, & \text{for } m \text{ even} \end{cases} \tag{4.10}$$

For even m the second part of (4.1), the second part of (4.2), (4.3) and (4.4), and for odd m the second part of (4.5), the second part of (4.6), (4.7) and (4.8) hold good for $a = 0$ too.

It may be noted that the result (4.10) verifies the results (18) and (31) of Kaul and Sen(1983).

References

Aneja, K.G. (1975). Random Walk and Rank Order Statistics, Ph.D. Thesis, University of Delhi, India.

Chung, K.L. and Feller, W. (1949). "Fluctuations in Coin tossing", *Proceedings of the National Academy of Science*, U.S.A., Vol 35, 605-608.

Csáki, E.(1961). "On the Number of Intersectioins in the One Dimensional Random Walk", *Publication of the Mathematical Institute of the Hungarian Academy of Sciences.*, Vol.. 6, 281-286.

Csáki, E and Vincze, I.(1961). "On Some Combinatorial Relations Concerning the Symmetric Random Walk",*Acta Scientiarum mathematicarum*, Vol. 24, 231-235.

Dwass, M.(1967). " Simple Random Walk and Rank Order Statistics", *Annals of Mathematical Statistics,* Vol. 38, 1042-1053.

Engelberg, O. (1964). "Exactand Limiting Distributions of the Number of Lead Positions in *"unconditional"* Ballot Problems, *Journal of Applied Probability*, Vol. 1, 168-172.

Feller, W. (1957). "The Number of Zeros and of Changes of Sign in a Symmetric Random Walk", *L' Enseignement Math.*, Vol. 3, 229-235.

Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, Vol. I, 3rd edition, New York, Wiley.

Jain, G.C. (1966). Random Walk and Their Orientation with Reference to Special Points, Ph.D. Thesis, University of Delhi, India.

Kaul, C.L. (1982). Contributions in Rank Order Statistics, Ph.D. Thesis, University of Delhi, India.

Kaul, C.L. and Sen, K. (1983). "Fluctuations of Partial Sums $\Sigma \pm 1$ when Total is Unknown", *Journal of the Indian Statistical Association*, Vol. 21, 83-90.

Saran, J. and Sen, K. (1981). "Some Distribution Results on Unconditional Ballot Problems", *Mathematische Operationsforschung und Statistik, Series Statistics*, Vol. 12, No. 4, 613-619.

Sen, K.(1964). " On Some Combinatorial Relations Concerning the Symmetric Random Walk", *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, Vol.. 9, 335-357.

Sen, K.(1969). " Paths of an Odd Number of Steps with Final Position Unspecified", *Journal of the Indian Statistical Association*, Vol. 7, 107-135.