ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 7, Number 4, Fall 1977

SUCCESS PROBABILITIES OF A CONDITIONED RANDOM WALK

W. D. KAIGH

ABSTRACT. Let $X_1, X_2, \cdots$ be a sequence of i.i.d. Bernoulli random variables with $P[X_1 = +1] = p$ and $P[X_1 = -1] = 1 - p$ for some fixed p with $0 < p < 1$ and form the simple random walk with initial state an arbitrary positive integer z by setting $S_0 = z$, $S_n = z + X_1 + \cdots + X_n$, $n \geqq 1$. We consider the chance behavior of the increments X_k conditioned by the time of first entry of the zero state by the corresponding random walk. In the context of the classical problem of gambler's ruin, we determine the conditional probability of a "win" by the gambler at any stage prior to his ruin, given only his initial capital z and the time of ruin.

1. We consider the version of the classical ruin problem in which a gambler with initial capital z competes in a sequence of independent games against an infinitely rich adversary. At the conclusion of each game the gambler's fortune is either increased or decreased by one unit according to the outcome of some random experiment. The gambler has a fixed but arbitrary success probability p of winning at each stage and the sequence of trials continues indefinitely until the cumulative fortune of the gambler diminishes to zero, i.e., until the gambler's initial capital is exhausted and he is "ruined." In probabilistic terminology we assume the individual gains $X_1, X_2, \cdots$ are independent and identically distributed Bernoulli random variables with $P[X_1 = +1] = p$ and $P[X_1 = -1] = 1 - p$ for some fixed p with $0 < p < 1$. The gambler's cumulative fortune after n trials is given by $S_0 = z$, $S_n = z + X_1 + \cdots + X_n$ for $n = 1, 2, \cdots$. The sequence $\{S_n\}$ is said to form a simple random walk with initial state z and absorbing state at the origin.

Historically, problems associated with gambler's ruin such as the probability of ultimate ruin, expected duration of the game, etc. have received much attention in probability theory. A thorough discussion is presented in Chapter XIV of Feller (1968). In this note we consider the following

PROBLEM. Given only the amount of initial capital z and the time of ruin n, determine the (conditional) probability that the kth trial resulted in a "win."

Received by the editors on August 24, 1976.

Copyright © 1977 Rocky Mountain Mathematics Consortium

Again in probabilistic terminology, for integers n and z with $1 \leqq z \leqq n$ we define the random time T_z to be the first index n with $S_n = 0$ and we wish to determine the conditional probability $P[X_k = +1 \mid T_z = n]$ for each $k = 1, 2, \cdots, n$.

We remark that extensive effort concerning the (asymptotic) chance behavior of the random walk $\{S_n\}$ conditioned by the events $\{T_0 > n\}$ and $\{T_0 = n\}$ has been conducted by several authors, e.g., Belkin (1972), Iglehart (1974), Kaigh (1975) and (1976), but here we consider the conditional increments X_k instead of the cumulative sums S_k.

2. As stated previously our goal is to calculate $P[X_k = \pm 1 \mid T_z = n]$ which will be denoted by $p^{\pm}(z; n; k)$.

To provide comparison and insight note that the event $\{T_z = n\}$ implies the event $\{S_n = 0\}$ and consider $P[X_k = \pm 1 \mid S_n = 0]$. An easy symmetry argument shows that since the conditional expectation $E(S_n \mid S_n = 0) = 0$ and since X_k assumes only the two values ± 1, we have

$$-z/n = \mathrm{E}(X_k \mid S_n = 0) = (-1)P[X_k = -1 \mid S_n = 0] + (+1)P[X_k = +1 \mid S_n = 0],$$

and

$$P[X_k = \pm 1 \mid S_n = 0] = (1 \mp z/n)/2 \text{ for } k = 1, 2, \cdots, n.$$

It is instructive to note that the above expression depends on neither p nor k. The lack of dependence on p is guaranteed by the statistical concept of sufficiency and it is clear that conditioning by the event $\{T_z = n\}$ will not alter this phenomenon. This observation is of definite significance because our subsequent probability computations will simplify to application of elementary counting techniques and the classical definition of probability. However, in contrast to the above, it will be seen that an effect of the additional requirement that the entry of the zero state at time n be the first such entry is that $p^{\pm}(z; n; k)$ is not independent of k.

Following these observations we begin now the calculation of the desired conditional probabilities. To facilitate our counting arguments it is convenient to introduce the notion of a path. For integers x and y with $x > 0$ a *path* from the origin $(0, 0)$ to the point (x, y) is a polygonal line whose vertices have abscissas $0, 1, \cdots, x$ and ordinates $s_0, s_1, \cdots, s_x$ satisfying $s_0 = 0$, $s_x = y$, $s_i - s_{i-1} = x_i = \pm 1$ for $i = 1, 2, \cdots, x$. The preceding definition is easily generalized to include paths between arbitrary points (x_1, y_1) and (x_2, y_2) with integer coordinates. A more detailed discussion is presented on page 68 of Feller (1968).

Conditioning by $T_z = n$ involves consideration only of all paths from $(0, z)$ to $(n, 0)$ which avoid the x-axis until time n. We introduce the following notation: Let $u(z; n)$ be the number of paths with $T_z = n$ and for $k = 1, 2, \cdots, n$ let $v^{\pm}(z; n; k)$ be the number of paths with $T_z = n$ and $x_k = \pm 1$.

The immediate consequence of the earlier discussion concerning the lack of dependence of the conditional probabilities $p^{\pm}(z; n; k)$ on the Bernoulli success probability p, is that given $T_z = n$, the $u(z; n)$ paths satisfying this condition are equally likely. As a result we have that

$$p^{\pm}(z; n; k) = v^{\pm}(z; n; k)/u(z; n) \text{ for } 1 \leqq z \leqq n \text{ and } k = 1, 2, \cdots, n. \tag{1}$$

Next we record several identities involving the terms in (1). Clearly,

$$\begin{aligned} u(z; n) &= v^{+}(z; n; k) + v^{-}(z; n; k), k = 1, 2, \cdots, n \text{ and} \\ v^{+}(z; n; n-1) &= v^{+}(z; n; n) = 0 \text{ for } 1 \leqq z \leqq n. \end{aligned} \tag{2}$$

From page 352 of Feller (1968) we obtain

$$u(z; n) = \frac{z}{n}\binom{n}{(n+2)/2} \text{ for } 1 \leqq z \leqq n, \tag{3}$$

where we adopt the convention that the above and all similar expressions which follow vanish if z and n are of opposite parity. Similarly, binomial coefficients of the form $\binom{N}{K}$ are taken to be zero if it is not the case that K and N are integers with $0 \leqq K \leqq N$.

Since each $X_k = \pm 1$, we note that $v^{\pm}(1; 2n+1; k+1) = v^{\pm}(2; 2n; k)$ and $v^{\pm}(z; n+1; k+1) = v^{\pm}(z+1; n; k) + v^{\pm}(z-1; n; k)$ for $2 \leqq z \leqq n$ and $k = 1, 2, \cdots, n$ to obtain recursively

$$\begin{aligned} v^{\pm}(z; n; k) = \sum_{j=0}^{[(z-1)/2]} (-1)^j \binom{z-j-1}{j} v^{\pm}(1; n+z-1 \\ - 2j; k+z-1-2j) \end{aligned} \tag{4}$$

$$\text{for } 1 \leqq z \leqq n \text{ and } k = 1, 2, \cdots, n.$$

Although a closed form for $u(z; n)$ was exhibited in (2), it is of interest that a similar recursive argument will provide the analogous formula

$$u(z; n) = \sum_{j=0}^{[(z-1)/2]} (-1)^j \binom{z-j-1}{j} u(1; n+z-1-2j) \tag{5}$$

$$\text{for } 1 \leqq z \leqq n.$$

The computational significance of (4) and (5) is that for fixed k and n with $1 \leqq k \leqq n$ the functions $u(\cdot; n)$ and $v^{\pm}(\cdot; n; k)$, $1 \leqq z \leqq n$, are determined by their respective values at $z = 1$. As a result in the following section we devote our attention to $u(1; n)$ and $v^{\pm}(1; n; k)$ to ultimately obtain a computational expression for (1).

3. Our immediate objective is to determine $v^{\pm}(1; 2n + 1; k)$. Since we have from (3) that

$$u(1; 2n+1) = \frac{1}{2n+1}\binom{2n+1}{n+1} = \frac{1}{n+1}\binom{2n}{n} \tag{6}$$

it is clear from (1) that this task is equivalent to a determination of $p^{\pm}(1; 2n + 1; k) = P[X_k = \pm 1 \mid T_1 = 2n + 1]$. To compute this we note that $p^{+}(1; 2n + 1; k) + p^{-}(1; 2n + 1; k) = 1$ and obtain

$$\begin{aligned} E(X_k \mid T_1 = 2n+1) &= (+1)\,p^{+}(1; 2n+1; k) \\ &\quad + (-1)\,p^{-}(1; 2n+1; k) \\ &= 2p^{+}(1; 2n+1; k) - 1 \\ &\quad \text{for } k = 1, 2, \cdots, 2n+1. \end{aligned} \tag{7}$$

Employing the linearity property of conditional expectation we have

$$\begin{aligned} E(X_k \mid T_1 = 2n+1) &= E(S_k - S_{k-1} \mid T_1 = 2n+1) \\ &= E(S_k \mid T_1 = 2n+1) \\ &\quad - E(S_{k-1} \mid T_1 = 2n+1) \\ &\quad \text{for } k = 1, 2, \cdots, 2n+1. \end{aligned} \tag{8}$$

A combination of (7) with (8) provides

$$\begin{aligned} p^{+}(1; 2n+1; k) &= [1 + E(S_k \mid T_1 = 2n+1) - E(S_{k-1} \mid T_1 \\ &= 2n+1)]/2 \\ &\text{for } k = 1, 2, \cdots, 2n+1. \end{aligned} \tag{9}$$

From (9) we see that it suffices to determine $E(S_k \mid T_1 = 2n + 1)$ for $k = 0, 1, \cdots, 2n + 1$. We perform this computation in a direct manner first obtaining the conditional probability distributions $P[S_k = x \mid T_1 = 2n + 1]$ and then employing the definition of expectation. As an initial step we introduce further notation. For $k = 0, 1, \cdots, 2n + 1$ let $w(x; 2n + 1; k)$ be the number of paths with $T_1 = 2n + 1$ and $S_k = x$. A brief reflection will show that the possible conditional values for S_k satisfy $S_{2n+1} \equiv 0$ and $1 \leqq S_k \leqq \min(k + 1, 2n - k + 1)$ for $0 \leqq k \leqq 2n$.

An application of the elementary multiplicative counting rule yields

$$w(x; 2n+1; k) = u(x; k+1)\, u(x; 2n+1-k) \quad \text{for } k = 0, 1, \cdots, 2n+1. \tag{10}$$

The conditional probability distribution for S_k is obtained following division of (10) by $u(1; 2n+1)$. An application of (3) then provides

$$\begin{aligned} P[S_k = x \mid T_1 = 2n+1] &= [(n+1)x^2/(k+1)(2n-k+1)] \\ &\quad \binom{k+1}{(k+x+1)/2}\binom{2n-k+1}{(2n-k+x+1)/2}\Big/\binom{2n}{n} \\ &\quad \text{for } k = 0, 1, \cdots, 2n \text{ and } 1 \leqq x \leqq \min(k+1, 2n-k+1); \end{aligned} \tag{11}$$

$$P[S_{2n+1} = 0 \mid T_1 = 2n+1] = 1.$$

To illustrate with a numerical example we take $2n + 1 = 21$. From (6) we compute $u(1; 21) = 16796$ paths from $(0, 1)$ to $(21, 0)$ satisfying $T_1 = 21$. We take $k = 13$ and apply (11) to obtain the possible conditional values for S_{13} as 2, 4, 6, 8 with corresponding conditional probabilities 6006/16796, 8008/16796, 2574/16796, 208/16796, respectively.

Although not required here it is of interest that the transition density of the Brownian excursion stochastic process (see [5]) can be obtained by passage to the limit in a suitable normalization of (11).

To compute $E(S_k \mid T_1 = 2n+1)$ it is necessary to evaluate the series

$$\mathrm{E}(S_k \mid T_1 = 2n+1) = \sum_x x\, P[S_k = x \mid T_1 = 2n+1]. \tag{12}$$

To perform this we consider separately the cases k even and k odd so we must evaluate each of the following:

$$\begin{aligned} \mathrm{E}(S_{2k} \mid T_1 = 2n+1) &= \sum_x (2x+1)\, P[S_{2k} = 2x+1 \mid T_1 = 2n+1] \\ \mathrm{E}(S_{2k+1} \mid T_1 = 2n+1) &= \sum_x (2x+2)\, P[S_{2k+1} = 2x+2 \mid T_1 = 2n+1]. \end{aligned} \tag{13}$$

Examination of (11) and (13) indicates an apparent difficulty for a direct calculation. Instead we employ the identity

$$\begin{aligned} 2x+1 = [&(k+x+1)(n-k+x+1) \\ &- (k-x)(n-k-x)]/(n+1) \end{aligned}$$

to simplify binomial coefficients, and after manipulation the first series in (13) becomes

$$E(S_{2k} \mid T_1 = 2n + 1) = \sum_x (2x + 1)^2 \binom{2k}{k + x} \binom{2n - 2k}{n - k - x} \Big/ \binom{2n}{n}$$
$$- \sum_x (2x + 1)^2 \binom{2k}{k + x + 1} \binom{2n - 2k}{n - k - x - 1} \Big/ \binom{2n}{n}.$$

Using symmetry of the binomial coefficients involved and further manipulation, we obtain

$$E(S_{2k} \mid T_1 = 2n + 1)$$
$$= 4 \sum_{y=0}^{2k} |y - k| \binom{2k}{y} \binom{2n - 2k}{n - y} \Big/ \binom{2n}{n} \tag{14}$$
$$+ \binom{2k}{k} \binom{2n - 2k}{n - k} \Big/ \binom{2n}{n}.$$

The identity

$$(2x + 2) = [(k + x + 2)(n - k - x + 1) - (k - x)(n - k - x - 1)]/(n + 1)$$

and similar treatment provides

$$E(S_{2k+1} \mid T_1 = 2n + 1)$$
$$= 4 \sum_{y=0}^{2k+1} |y - (2k + 1)/2| \binom{2k + 1}{y} \binom{2n - 2k - 1}{n - y} \Big/ \binom{2n}{n}. \tag{15}$$

The series appearing in (14) and (15) have respective closed forms

$$\sum_{y=0}^{2k} |y - k| \binom{2k}{y} \binom{2n - 2k}{n - y} \Big/ \binom{2n}{n}$$
$$= [k(n - k)/n] \binom{2k}{k} \binom{2n - 2k}{n - k} \Big/ \binom{2n}{n}$$

$$\sum_{y=0}^{2k+1} |y - (2k + 1)/2| \binom{2k + 1}{y} \binom{2n - 2k - 1}{n - y} \Big/ \binom{2n}{n} \tag{16}$$
$$= [(k + 1)(n - k)/n] \binom{2k + 1}{k} \binom{2n - 2k - 1}{n - k} \Big/ \binom{2n}{n}.$$

The validity of (16) is easily seen but since the derivations are quite tedious we omit the verification. It is curious that the above series admit a probabilistic interpretation as the expected absolute deviation about the mean of a symmetric hypergeometric random variable.

Substitution of the expressions of (16) into (14) and (15) produces

$$
\begin{aligned}
\mathrm{E}(S_{2k} \mid T_1 = 2n+1) &= 4[k(n-k)/n] \binom{2k}{k}\binom{2n-2k}{n-k}\Big/\binom{2n}{n} \\
\text{(17)} \qquad &\quad + \binom{2k}{k}\binom{2n-2k}{n-k}\Big/\binom{2n}{n} \\
\mathrm{E}(S_{2k+1} \mid T_1 = 2n+1) &= 4[(k+1)(n-k)/n] \\
&\quad \binom{2k+1}{k}\binom{2n-2k-1}{n-k}\Big/\binom{2n}{n} \\
&\quad \text{for } k = 0, 1, \cdots, n.
\end{aligned}
$$

Following substitution of the terms of (17) into (9) and simplification we obtain

$$
\begin{aligned}
p^{\pm}(1; 2n+1; 2k) &= p^{\pm}(1; 2n+1; 2k+1) \\
\text{(18)} \qquad &= 1/2 \pm [(n-2k)/2n] \binom{2k}{k}\binom{2n-2k}{n-k}\Big/\binom{2n}{n} \\
&\quad \text{for } k = 0, 1, \cdots, n.
\end{aligned}
$$

A combination of (1), (6), and (18) provides

$$
\begin{aligned}
v^{\pm}(1; 2n+1; 2k) &= v^{\pm}(1; 2n+1; 2k+1) \\
\text{(19)} \qquad &= u(1; 2n+1)/2 \pm [(n-2k)/2n(n+1)] \\
&\quad \binom{2k}{k}\binom{2n-2k}{n-n} \\
&\quad \text{for } k = 0, 1, \cdots, n.
\end{aligned}
$$

It is interesting to note the symmetry about $2k = n$ which appears in (17), (18), and (19). A rough interpretation is that with the initial state $z = 1$, the random fluctuations of the initial and final portions of the random walk are probabilistically identical.

Because of the relation (4), formula (19) enables us to compute $v^{\pm}(z; n; k)$ which can be used in (1) to calculate finally the conditional probabilities $p^{\pm}(z; n; k)$. Our problem thus is solved and we conclude with a table illustrating the conditional success probabilities for a gambler with values of initial capital 1, 3, 5, 7, 9, 11 ruined after 11 trials.

TABLE 1

k	$v^+(1; 11; k)$	$v^+(3; 11; k)$	$v^+(5; 11; k)$	$v^+(7; 11; k)$	$v^+(9; 11; k)$	$v^+(11; 11; k)$
1	42 (1.000)	48 (0.533)	27 (0.360)	8 (0.229)	1 (0.111)	0 (0.000)
2–3	28 (0.667)	48 (0.533)	27 (0.360)	8 (0.229)	1 (0.111)	0 (0.000)
4–5	23 (0.548)	43 (0.478)	27 (0.360)	8 (0.229)	1 (0.111)	0 (0.000)
6–7	19 (0.452)	37 (0.411)	25 (0.333)	8 (0.229)	1 (0.111)	0 (0.000)
8–9	14 (0.333)	28 (0.311)	20 (0.267)	7 (0.200)	1 (0.111)	0 (0.000)
10–11	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)
	$u(1; 11) = 42$	$u(3; 11) = 90$	$u(5; 11) = 75$	$u(7; 11) = 35$	$u(9; 11) = 9$	$u(11; 11) = 1$

The numbers appearing in parentheses are the corresponding $p^+(z; n; k)$. The $v(z; n; k)$ are calculated from the appropriate formulas and the following chart containing numerical values for $v(1; n; k)$.

$$v^+(1; 11; k) = \sum_{j=0}^{0} (-1)^j \binom{0-j}{j} v^+(1; 11 - 2j; k - 2j)$$

$$v^+(3; 11; k) = \sum_{j=0}^{1} (-1)^j \binom{2-j}{j} v^+(1; 13 - 2j; k + 2 - 2j)$$

$$v^+(5; 11; k) = \sum_{j=0}^{2} (-1)^j \binom{4-j}{j} v^+(1; 15 - 2j; k + 4 - 2j)$$

$$v^+(7; 11; k) = \sum_{j=0}^{3} (-1)^j \binom{6-j}{j} v^+(1; 17 - 2j; k + 6 - 2j)$$

$$v^+(9; 11; k) = \sum_{j=0}^{4} (-1)^j \binom{8-j}{j} v^+(1; 19 - 2j; k + 8 - 2j)$$

$$v^+(11; 11; k) = \sum_{j=0}^{5} (-1)^j \binom{10-j}{j} v^+(1; 21 - 2j; k + 10 - 2j)$$

k	$v^+(1;1;k)$	$v^+(1;3;k)$	$v^+(1;5;k)$	$v^+(1;7;k)$	$v^+(1;9;k)$	$v^+(1;11;k)$	$v^+(1;13;k)$	$v^+(1;15;k)$	$v^+(1;17;k)$	$v^+(1;19;k)$	$v^+(1;21;k)$
1	0	1	2	5	14	42	132	429	1430	4862	16796
	(0.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
2–3		0	1	3	9	28	90	297	1001	3432	11934
		(0.000)	(0.500)	(0.600)	(0.643)	(0.667)	(0.682)	(0.692)	(0.700)	(0.706)	(0.711)
4–5			0	2	7	23	76	255	869	3003	10504
			(0.000)	(0.400)	(0.500)	(0.548)	(0.576)	(0.594)	(0.608)	(0.618)	(0.625)
6–7				0	5	19	66	227	785	2739	9646
				(0.000)	(0.357)	(0.452)	(0.500)	(0.529)	(0.549)	(0.563)	(0.574)
8–9					0	14	56	202	715	2529	8986
					(0.000)	(0.333)	(0.424)	(0.471)	(0.500)	(0.520)	(0.535)
10–11						0	42	174	645	2333	8398
						(0.000)	(0.318)	(0.406)	(0.451)	(0.480)	(0.500)
12–13							0	132	561	2123	7810
							(0.000)	(0.308)	(0.392)	(0.437)	(0.465)
14–15								0	429	1859	7150
								(0.000)	(0.300)	(0.382)	(0.426)
16–17									0	1430	6292
									(0.000)	(0.294)	(0.375)
18–19										0	4862
										(0.000)	(0.289)
20–21											0
											(0.000)

$u(1;1) = 1$ $u(1;3) = 1$ $u(1;5) = 2$ $u(1;7) = 5$ $u(1;9) = 14$ $u(1;11) = 42$ $u(1;13) = 132$
$u(1;15) = 429$ $u(1;17) = 1430$ $u(1;19) = 4862$ $u(1;21) = 16796$

References

1. B. Belkin, *An invariance principle for conditioned random walk attracted to a stable law*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **21** (1972), 45-64.

2. W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed., Wiley, New York (1968).

3. D. L. Iglehart, *Functional central limit theorems for random walks conditioned to stay positive*, Ann. Prob. 2 (1974), 608-619.

4. W. D. Kaigh, *A conditional local limit theorem for recurrent random walk*, Ann. Prob. 3 (1975), 882-887.

5. ——, *An invariance principle for random walk conditioned by a late return to zero*, Ann. Prob. 4 (1976), 115-121.

University of Texas at El Paso, El Paso, Texas 79968