An Elementary Proof of the Hitting Time Theorem

Remco van der Hofstad and Michael Keane

In this note, we give an elementary proof of the random walk hitting time theorem, which states that, for a left-continuous random walk on $\mathbb{Z}$ starting at a nonnegative integer k, the conditional probability that the walk hits the origin for the first time at time n, given that it does hit zero at time n, is equal to k/n. Here, a walk is called *left-continuous* when its steps are bounded from below by -1.

We start by introducing some notation. Let $\mathbb{P}_k$ denote the law of a random walk starting at $k \geq 0$, let $\{Y_i\}_{i=1}^{\infty}$ be the independent and identically distributed (i.i.d.) steps of the random walk, let $S_n = k + Y_1 + \cdots + Y_n$ be the position of the random walk starting at k after n steps, and let

$$T_0 = \inf\{n : S_n = 0\} \tag{1}$$

denote the walk's first hitting time of the origin. Clearly, $T_0 = 0$ when the walker starts at the origin. Then, the hitting time theorem is the following result:

Theorem 1 (Hitting time theorem). *For a random walk starting at $k \geq 1$ with i.i.d. steps $\{Y_i\}_{i=1}^{\infty}$ satisfying $Y_i \geq -1$ almost surely, the distribution of T_0 under $\mathbb{P}_k$ is given by*

$$\mathbb{P}_k(T_0 = n) = \frac{k}{n}\mathbb{P}_k(S_n = 0). \tag{2}$$

Theorem 1 gives the remarkable conclusion that, conditionally on the event $\{S_n = 0\}$, and regardless of the precise distribution of the steps of the walk $\{Y_i\}_{i=1}^{\infty}$, as long as $Y_i \geq -1$ a.s., the probability of the walk being at 0 *for the first time* is equal to k/n. Theorem 1 has particular importance in the context of branching processes, where, for $k = 1$, the law of T_0 is the same as the total progeny of a branching process with offspring distribution equal to the law of $Y_i + 1$ (see [**5**] or [**10**, Problem 12, p. 234] for this connection, and [**1**, Section 10.4] for a modern application in random graph theory).

The first proofs of Theorem 1 in the special case $k = 1$ can be found in [**9**]. The extension to $k \geq 2$ is in [**7**], or in [**5**] using a result in [**4**]. Most of these proofs make unnecessary use of generating functions, in particular, the Lagrange inversion formula, which the simple proof below does not employ. See also [**6**, pp. 165–167] for a more recent version of the generating function proof. In [**12**], various proofs of the hitting time theorem are given, including a combinatorial proof making use of a relation in [**3**]. A proof for random walks making only steps of size ± 1 using the reflection principle can for example be found in [**6**, p. 79].

The hitting time theorem is equivalent to the ballot theorem, which has a long history dating back to Whitford in 1878 and, independently, Bertrand in 1887 (see [**2, 8, 11**] for overviews of the history and literature). The version of the ballot theorem in [**8**] states that, for a random walk $\{S_n\}_{n=0}^{\infty}$ starting at 0, with exchangeable, nonnegative steps, the probability that $S_m < m$ for all $m = 1, \ldots, n$, conditionally on $S_n = k$, equals $(n - k)/n$. Here, a random vector is called *exchangeable* when any permutation of the vector has the same distribution. The proof of [**8**] borrows from queueing theory methodology, and is related to, yet slightly different from, our proof below. Proofs of the ballot theorem often use induction, but are generally of a more combinatorial nature than the proof given below, which is entirely probabilistic.

To prove the ballot theorem from the hitting time theorem, suppose that $\{S_m\}_{m=0}^n$ is a random walk starting at 0 with exchangeable nonnegative steps, and suppose that $n > k \geq 0$. For $m \leq n$, let

$$\tilde{S}_m = n - k - m + S_n - S_{n-m}. \tag{3}$$

Then $\tilde{S}_0 = n - k$ and $\tilde{S}_m - \tilde{S}_{m-1} = S_{n-m+1} - S_{n-m} - 1$, so $\{\tilde{S}_m\}_{m=0}^n$ is a left-continuous random walk with exchangeable steps starting at $n - k$. Also, $\tilde{S}_n = 0$ if and only if $S_n = k$, and if $S_n = k$ then $\tilde{S}_m = (n - m) - S_{n-m}$, so $\tilde{S}_m > 0$ if and only if $S_{n-m} < n - m$. The conclusion of the ballot theorem now follows from the hitting time theorem (once the hitting time theorem is extended to exchangeable steps; see the proof of Theorem 2 below). The proof of the hitting time theorem from the ballot theorem retraces the above steps, and is left to the reader.

Proof. We prove (2) for all $k \geq 0$ by induction on $n \geq 1$. When $n = 1$, both sides are equal to 0 when $k > 1$ and $k = 0$, and are equal to $\mathbb{P}(Y_1 = -1)$ when $k = 1$. This initializes the induction.

To advance the induction, we take $n \geq 2$, and note that both sides are equal to 0 when $k = 0$. Thus, we may assume that $k \geq 1$. We condition on the first step to obtain

$$\mathbb{P}_k(T_0 = n) = \sum_{s=-1}^{\infty} \mathbb{P}_k(T_0 = n \mid Y_1 = s)\mathbb{P}(Y_1 = s). \tag{4}$$

For $m = 0, \dots, n-1$, let $S_m^* = S_{m+1}$. By the random walk Markov property, when $S_0 = k$ and conditionally on $Y_1 = s$, the random walk path $\{S_m^*\}_{m=0}^{n-1}$ has the same distribution as the random walk path $\{S_m\}_{m=0}^{n-1}$ started at $S_0 = k + s$, so that

$$\mathbb{P}_k(T_0 = n \mid Y_1 = s) = \mathbb{P}_{k+s}(T_0 = n - 1) = \frac{k+s}{n-1}\mathbb{P}_{k+s}(S_{n-1} = 0), \tag{5}$$

where in the last equality we have used the induction hypothesis, which is allowed since $n - 1 \geq 1$, and $k \geq 1$ and $s \geq -1$, so that $k + s \geq 0$. This leads to

$$\mathbb{P}_k(T_0 = n) = \sum_{s=-1}^{\infty} \frac{k+s}{n-1}\mathbb{P}_{k+s}(S_{n-1} = 0)\mathbb{P}(Y_1 = s). \tag{6}$$

We undo the law of total probability, noting that

$$\begin{aligned}
&\sum_{s=-1}^{\infty} (k+s)\mathbb{P}_{k+s}(S_{n-1} = 0)\mathbb{P}(Y_1 = s) \qquad (7)\\
&\quad = \sum_{s=-1}^{\infty} (k+s)\mathbb{P}_k(S_n = 0 \mid Y_1 = s)\mathbb{P}(Y_1 = s)\\
&\quad = \sum_{s=-1}^{\infty} (k+s)\mathbb{P}_k(Y_1 = s \mid S_n = 0)\mathbb{P}_k(S_n = 0)\\
&\quad = \mathbb{P}_k(S_n = 0)\big(k + \mathbb{E}_k[Y_1 \mid S_n = 0]\big),
\end{aligned}$$

where $\mathbb{E}_k[Y_1 \mid S_n = 0]$ is the conditional expectation of Y_1 given that $S_n = 0$ occurs. We next note that the conditional expectation $\mathbb{E}_k[Y_i \mid S_n = 0]$ is independent of i, so that

$$\begin{aligned}\mathbb{E}_k[Y_1 \mid S_n = 0] &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_k[Y_i \mid S_n = 0] \\ &= \frac{1}{n}\mathbb{E}_k\left[\sum_{i=1}^{n} Y_i \mid S_n = 0\right] = -\frac{k}{n},\end{aligned} \tag{8}$$

since $\sum_{i=1}^{n} Y_i = S_n - k = -k$ when $S_n = 0$. Therefore, we arrive at

$$\mathbb{P}_k(T_0 = n) = \frac{1}{n-1}\left[k - \frac{k}{n}\right]\mathbb{P}_k(S_n = 0) = \frac{k}{n}\mathbb{P}_k(S_n = 0). \tag{9}$$

This advances the induction, and completes the proof of Theorem 1. ■

We close this paper by deriving a slightly stronger result than Theorem 1. We first introduce some notation. We write $\vec{m} = (m_{-1}, m_0, m_1, \dots)$, where m_i are nonnegative integers for all $i \geq -1$. Further, for any $\vec{m}$ and $n \geq 1$, let $A_{\vec{m}}(n)$ denote the event that $S_n = 0$, and the random walk $\{S_i\}_{i=0}^{\infty}$ has made m_{-1} steps of size -1, m_0 steps of size 0, m_1 steps of size 1, etc. In other words, we let

$$A_{\vec{m}}(n) = \{S_n = 0\} \cap \left\{\#\{i \in \{1, \dots, n\} : Y_i = s\} = m_s, \forall s \geq -1\right\}. \tag{10}$$

Naturally, $\sum_{i=-1}^{\infty} m_i = n$, and, since $S_n = 0$, we must have that $\sum_{i=-1}^{\infty} i m_i = -k$. Then we obtain the following extension of Theorem 1:

Theorem 2. *Under the assumptions in Theorem 1, for all $k \geq 1$, $\vec{m}$, and $n \geq 1$,*

$$\mathbb{P}_k(\{T_0 = n\} \cap A_{\vec{m}}(n)) = \frac{k}{n}\mathbb{P}_k(A_{\vec{m}}(n)). \tag{11}$$

The proof of Theorem 2 is a minor modification of that of Theorem 1, again using that, conditionally on $A_{\vec{m}}(n)$, the steps $(Y_1, \dots, Y_n)$ are exchangeable. We omit further details. Theorem 1 can be deduced from Theorem 2 by conditioning on the number of steps of all sizes, using (11) followed by a sum over all possible collections of steps of all sizes for which $\sum_{i=-1}^{\infty} m_i = n$ and $\sum_{i=-1}^{\infty} i m_i = -k$.

As indicated above for the ballot theorem, Theorem 1 also follows when the steps $(Y_1, \dots, Y_n)$ are *exchangeable* instead of i.i.d. Indeed, for exchangeable steps, the law of $(Y_1, \dots, Y_n)$ conditionally on $\sum_{i=1}^{n} Y_i = -k$ is still exchangeable, which implies that the conditional expectation $\mathbb{E}_k[Y_i \mid S_n = 0]$ is independent of i. Thus, in this setting, (8) again follows, which allows for the induction argument to be completed. As such, Theorem 2 follows from this generalization of Theorem 1 to exchangeable steps.

ACKNOWLEDGMENTS. The work of RvdH was supported in part by Netherlands Organisation for Scientific Research (NWO). We thank the anonymous referees for a clarification of the history of the problem, as well as many helpful suggestions which improved the presentation.

REFERENCES

1. N. Alon and J. Spencer, *The Probabilistic Method*, 2nd ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley, New York, 2000.
2. D. E. Barton and C. L. Mallows, Some aspects of the random sequence, *Ann. Math. Statist.* **36** (1965) 236–260.
3. M. Dwass, A fluctuation theorem for cyclic random variables, *Ann. Math. Statist.* **33** (1962) 1450–1454.
4. ———, A theorem about infinitely divisible distributions, *Z. Wahrscheinleikheitsth.* **9** (1968) 206–224.
5. ———, The total progeny in a branching process and a related random walk, *J. Appl. Prob.* **6** (1969) 682–686.
6. G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 3rd ed., Oxford University Press, New York, 2001.
7. J. H. B. Kemperman, *The Passage Problem for a Stationary Markov Chain*, Statistical Research Monographs, vol. I, University of Chicago Press, Chicago, IL, 1961.
8. T. Konstantopoulos, Ballot theorems revisited, *Statist. Probab. Lett.* **24** (1995) 331–338.
9. R. Otter, The multiplicative process, *Ann. Math. Statist.* **20** (1949) 206–224.
10. F. Spitzer, *Principles of Random Walk*, 2nd ed., Springer, New York, 1976.
11. L. Takács, Ballots, queues and random graphs, *J. Appl. Probab.* **26** (1989) 103–112.
12. J. G. Wendel, Left-continuous random walk and the Lagrange expansion, this MONTHLY **82** (1975) 494–499.

Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
rhofstad@win.tue.nl

Department of Mathematics and Computer Science, Wesleyan University Middletown, CT 06459-0128, U.S.A.
mkeane@wesleyan.edu, and
Philips Research Laboratories, Building WO 1.075, 5656 AE Eindhoven, The Netherlands

A Result of Strauss:[1] 41/333 = 3/4

Proof. We have

$$\begin{aligned} 41/333 &= 123/999 \\ &= .123\ 123\ 123\ 123\ldots \\ &= 3/4, \end{aligned}$$

where the last equality is derived from a basic waltz step. ▲

—Submitted by Rick Mabry, Shreveport, LA

[1] Johann Strauss II (the “Waltz King”), October 25, 1825–June 3, 1899. Happy birthday, Johann!