

# Wald's Identity for the Fair Coin-Tossing Games and Some Applications

Ying-Chao Hung\*

Department of Statistics National Chengchi University No. 64, Sec. 2, ZhiNan Rd., Wenshan District Taipei 11605, Taiwan (R.O.C)

**Abstract:** We provide a simple proof of Wald's second identity for a class of problems that can be formulated as a fair coin-tossing game. The identity provides a useful technique for computing the expected stopping time and allows us to explore the second-order moment of the so-called heads-minus-tails distribution. We also introduce some interesting applications related to this simple identity.

**Keywords:** Wald's identity, fair coin-tossing games, heads-minus-tails distribution, stopping time.

## 1. INTRODUCTION

Let  $\{X_j\}_{j \in \mathbb{Z}_+}$  be a sequence of i.i.d. random variables that represent the consecutive outcomes of tossing a fair coin, i.e.,  $P(X_j = 1) = P(X_j = -1) = \frac{1}{2}$ , where  $X_j = 1$  represents the outcome of a head and  $X_j = -1$  the outcome of a tail. Let  $N$  be the stopping time of any fair coin-tossing game with respect to the filtration  $F_n = \sigma\{X_1, \dots, X_n\}$  and consider the partial sum process  $S_n = X_1 + \dots + X_n$ ,  $S_0 = 0$ . Therefore,  $S_N$  represents the partial sum at the stopping time  $N$ , which is the difference between the number of heads and the number of tails (also known as the "heads-minus-tails") when the game stops. Note that such a fair coin-tossing game has many applications. For example, in the gambler's ruin problem  $S_N$  denotes the player's fortune earned when the game stops; in the symmetric random walk  $S_N$  denotes the location when it stops; just to name a few.

Conventional studies for the partial sum processes have focused on their moments by utilizing the results from the central limit theorem [1, 2] or through direct calculations based on the heads-minus-tails distribution [3, 4]. For the stopped processes  $S_N$ , researchers are mostly interested in finding the value of  $E(N)$ . The computation of  $E(N)$  has been largely dependent on the strong theoretical results from the fundamental Wald's equation and the martingale stopping theorem [1, 2, 5]. However, these theoretical results may not directly apply to examples such as fair coin-tossing games or symmetric random walks (think about solving  $E(N)$  by using Wald's equation but  $E(S_N) = E(X_j) = 0$ ). To overcome this problem, the Wald's second identity turns out to be quite useful.

The Wald's second identity states that, let  $\{X_j\}_{j \in \mathbb{Z}_+}$  be a sequence of i.i.d. random variables with  $E(X_1) = 0$  and  $E(X_1^2) < \infty$ , then for a stopping time  $N$  with  $E(N) < \infty$  we have  $E(S_N^2) = E(X_1^2)E(N)$  [2, 6-9]. Note that for the fair coin-tossing game, this identity simply reduces to  $\text{Var}(S_N) = E(N)$ . In this letter we provide another proof of Wald's second identity for the fair coin-tossing games. The identity can serve as a useful technique for computing the expected stopping time. Further, it allows us to explore the second-order moment of the heads-minus-tails distribution, which was not fully discussed in literature. In Section 2 we introduce the identity and sketch a simple proof of it. In Section 3 we introduce some interesting applications related to this identity.

## 2. ANOTHER PROOF OF WALD'S SECOND IDENTITY FOR THE FAIR COIN-TOSSING GAMES

Based on the formulation of fair coin-tossing games introduced in Section 1, the reduced form of Wald's second identity is shown as follows.

**Theorem 1** For any stopping time  $N$  with respect to the filtration  $F_n = \sigma\{X_1, \dots, X_n\}$  and such that  $E(N) < \infty$ , we have  $\text{Var}(S_N) = E(N)$ .

**Proof:** Since  $E(N) < \infty$ , the fundamental Wald's equation gives that  $E(S_N) = E(\sum_{i=1}^N X_i) = E(N)E(X_1) = 0$ . To prove that  $\text{Var}(S_N) = E(N)$ , it then suffices to show  $E(S_N^2) = E(N)$ . Let  $Z_n = S_n^2 - n$ , since

$$\begin{aligned} E(Z_{n+1} | X_1, \dots, X_n) &= E(S_{n+1}^2 - (n+1) | X_1, \dots, X_n) \\ &= E((S_n + X_{n+1})^2 - (n+1) | X_1, \dots, X_n) \\ &= S_n^2 - n + 2S_n E(X_{n+1} | X_1, \dots, X_n) \\ &= Z_n + 0 = Z_n \end{aligned}$$

\*Address correspondence to this author at the Department of Statistics National Chengchi University No. 64, Sec. 2, ZhiNan Rd., Wenshan District Taipei 11605, Taiwan (R.O.C); Tel: 886-2-29387115; Fax: 886-2-29398024; E-mail: hungy@nccu.edu.tw

$\{Z_n\}_{n \geq 1} = \{S_n^2 - n\}_{n \geq 1}$  is a martingale with respect to the filtration  $F_n$ . A similar argument shows that  $\{S_{N \wedge n}^2 - (N \wedge n)\}_{n \geq 1}$  is also a martingale with respect to  $F_n$ . Hence, we have that

$$E(S_{N \wedge n}^2 - (N \wedge n)) = E(S_{N \wedge 1}^2 - (N \wedge 1)) = E(S_1^2 - 1) = 0$$

which implies  $E(S_{N \wedge n}^2) = E(N \wedge n)$  for all  $n = 1, 2, \dots$ . Since  $S_{N \wedge n}^2 \geq 0$ , by Fatou's Lemma we have that

$$E(\liminf_{n \rightarrow \infty} S_{N \wedge n}^2) \leq \liminf_{n \rightarrow \infty} E(S_{N \wedge n}^2) = \liminf_{n \rightarrow \infty} E(N \wedge n) \quad (1)$$

The reverse of Fatou's lemma gives that

$$E(\limsup_{n \rightarrow \infty} S_{N \wedge n}^2) \geq \limsup_{n \rightarrow \infty} E(S_{N \wedge n}^2) = \limsup_{n \rightarrow \infty} E(N \wedge n) \quad (2)$$

Since  $(N \wedge n)$  is increasing with  $n$ ,  $\limsup_{n \rightarrow \infty} (N \wedge n) = \liminf_{n \rightarrow \infty} (N \wedge n) = N$ , and thus

$$E(\limsup_{n \rightarrow \infty} S_{N \wedge n}^2) = E(\liminf_{n \rightarrow \infty} S_{N \wedge n}^2) = E(N) \quad (3)$$

Since we also know that  $\liminf_{n \rightarrow \infty} S_{N \wedge n}^2 = \limsup_{n \rightarrow \infty} S_{N \wedge n}^2 = S_N^2$ , by Eq. (1)-(3) we conclude that

$$E(\liminf_{n \rightarrow \infty} S_{N \wedge n}^2) = E(\limsup_{n \rightarrow \infty} S_{N \wedge n}^2) = E(S_N^2) = E(N)$$

**Remark 1.** The readers can refer to [2, 6-8] for alternative proofs of the Wald's second identity.

**Remark 2.** The following example shows that Theorem 1 is not correct when  $E(N) = \infty$ . Consider tossing a fair coin and stop when  $S_n = 1$ , it is obvious that  $Var(S_N) = 0$ . However, a little algebra shows that  $E(N) = \infty$ , which simply violates the identity shown in Theorem 1.

### 3. APPLICATIONS

In this section we introduce some interesting applications by utilizing the identity shown in Theorem 1.

#### APPLICATION 1 (THE OCCURRENCE OF STRINGS)

Let us consider the sample space  $\Omega = \{-1, 1\}$ , the  $n$ -tuple  $\Omega_n = \Omega \times \dots \times \Omega$  for each integer  $n \geq 1$ , and  $\Omega^* = \bigcup_{n \geq 1} \Omega_n$ . Any element  $A \in \Omega^*$  is called a string. Further, if  $A \in \Omega_n$ , then  $A$  is called a string of length  $n$ . For example, suppose we toss a fair coin and stop until the particular outcome/pattern  $(H, H, T, T, H)$  appears. Then the corresponding string is  $A = (1, 1, -1, -1, 1)$ , which is of length 5.

We next introduce a general framework for finding the solution of the "expected occurrence time" of strings. Let  $\{X_j\}_{j \geq 1}$  be a sequence of i.i.d. random variables such that

$$P(X_1 = 1) = P(X_1 = -1) = \frac{1}{2}. \text{ The string } A = (a_1, a_2, \dots, a_m) \text{ is a sub-string of the string } B = (b_1, b_2, \dots, b_n) \text{ if } a_i = b_i, a_2 = b_{i+1}, \dots, a_m = b_{i+m-1} \text{ for some } 1 \leq i \leq n - m + 1. \text{ For each string } A \in \Omega_n, \text{ let } T_A \text{ be the first occurrence time of the string } A \text{ (with}$$

respect to the stochastic process  $\{X_j\}_{j \geq 1}$ ) which is defined as  $T_A = \inf \{m : (X_{m-n+1}, X_{m-n+2}, \dots, X_m) = A\}$ , or  $T_A = \infty$  if no such  $m$  exists. For any given  $k$  strings  $A_1, A_2, \dots, A_k \in \Omega^*$  such that  $A_i$  is not a sub-string of  $A_j$  and  $A_j$  is not a sub-string of  $A_i$  for any  $1 \leq i < j \leq k$ , let  $N_k = \min\{T_{A_1}, T_{A_2}, \dots, T_{A_k}\}$ , i.e.,  $N_k$  is the first occurrence time of one of the strings  $A_1, A_2, \dots, A_k$ .

For any string  $S = (s_1, s_2, \dots, s_k)$  in  $\Omega^*$ , define  $P(S) = \prod_{i=1}^k P(X_j = s_i)$ . For any two strings  $A$  and  $B$  of the same length, define also the operation  $A \circ B$  as

$$A \circ B = \sum_{S \in \lambda(A)} \prod_{\pi(B)} P(S)^{-1}$$

where  $\lambda(A)$  and  $\pi(B)$  are the suffix and prefix of  $A$  and  $B$ , respectively (see [10] for details). The following lemma, which was introduced in [10-12], is essential for solving  $E(N_k)$ .

**Lemma 1** Let  $A_1, \dots, A_k$  be  $k$  distinct strings in  $\Omega^*$  ( $2 \leq k < 2^n$ ) and  $N_k = \min\{T_{A_1}, \dots, T_{A_k}\}$ . For each  $i = 1, 2, \dots, k$ , let  $P_k = P\{T_{A_1}, \dots, N_k\}$ . Then we have the following system of  $k+1$  linear equations (with  $k+1$  variables):

$$\begin{pmatrix} 0 & 1 & & \dots & & 1 \\ \vdots & & & & & \\ \vdots & & (A_i \circ A_j - A_j \circ A_i)_{i+1, j+1} & & & \\ \vdots & & & & & \\ 1 & & & & & \end{pmatrix} \begin{pmatrix} E(N_k) \\ p_1 \\ \vdots \\ p_k \end{pmatrix} = \begin{pmatrix} 1 \\ A_1 \circ A_1 \\ \vdots \\ A_k \circ A_k \end{pmatrix}$$

Note that given any  $k$  distinct strings  $A_1, \dots, A_k$ ,  $E(N_k)$  can be simply solved by the linear equations in Lemma 1. Based on the result of Theorem 1, we then obtain the variance of  $S_{N_k}$ .

**Remark 3.** Lemma 1 also applies to a general setting that  $\Omega = \{w_1, w_2, \dots, w_m\}$  [10-13]. The current setting  $\Omega = \{-1, 1\}$  for the coin-tossing games is just a special case of it.

**Example 1.** Consider two strings  $A_1 = (H, H, T) = (1, 1, -1)$  and  $A_2 = (H, T, H) = (1, -1, 1)$ . Let  $N_2$  be the occurrence time until one of the strings  $A_1$  and  $A_2$  appears. By Lemma 1 we have that

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 6 \\ 1 & 6 & 0 \end{pmatrix} \begin{pmatrix} E(N_2) \\ p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ 10 \end{pmatrix}$$

Solving the above linear equations we have that  $E(N_2) = 6$ , which also gives that  $Var(S_{N_2}) = 6$  by Theorem 1.

**Example 2.** Consider three strings  $A_1 = (H, H, Y) = (1, 1, -1)$ ,  $A_2 = (H, T, H) = (1, -1, 1)$ , and  $A_3 = (H, T, T, H) = (1, -1, -1, 1)$ . Let  $N_3$  be the occurrence time until one of the strings  $A_1, A_2$ , and  $A_3$  appears. By Lemma 1 we have that

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 6 & 6 \\ 1 & 6 & 0 & 8 \\ 1 & 14 & 16 & 0 \end{pmatrix} \begin{pmatrix} E(N_3) \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ 10 \\ 18 \end{pmatrix}$$

Solving the above linear equations we have that  $E(N_3) = 38/1$ , which also gives that  $Var(S_{N_3}) = 38/7$  by Theorem 1.

**Example 3.** Toss a fair coin and let  $N$  be the number of tosses until either  $m$  consecutive heads are observed or  $n$

consecutive tails are observed. Let  $A_1 = (1, 1, \dots, 1)$  (of length  $m$ ) and  $A_2 = (-1, -1, \dots, -1)$  (of length  $n$ ). Solving the linear equations in Lemma 1, we obtain that  $E(N_2) = (2^{m+1} - 2)(2^{n+1} - 2)/(2^{m+1} + 2^{n+1} - 4)$ . By Theorem 1 we then have that  $\text{Var}(S_{N_2}) = E(N_2) = (2^{m+1} - 2)(2^{n+1} - 2)/(2^{m+1} + 2^{n+1} - 4)$ .

**APPLICATION 2 (A TWO-PLAYER RUIN PROBLEM)**

Consider a coin-tossing game with two players (Player I and Player II) where each player has a 50% chance of winning for each toss. For each toss, Player I wins if a head appears, while Player II wins if a tail appears. After each toss of the coin, the loser must transfer one penny to the winner. Suppose Player I has  $b$  pennies and Player II has  $a$  pennies at the beginning of the game, and the game stops when one of the players goes broke. Let  $N$  be the number of tosses when the game stops. The question is, what is the expectation of  $N$ ?

This two-player ruin problem can be formulated as follows. It is clear that  $N = \inf \{n : S_n = a \text{ or } S_n = -b\}$ . First, we have that  $E(S_N) = 0$  by the fundamental Wald's equation.

Second, it is clear that  $P(S_N = a) = \frac{b}{a+b}$  and  $P(S_N = -b) = \frac{a}{a+b}$ .

Thus,  $\text{Var}(S_N) = (-b)^2 \left(\frac{a}{a+b}\right) + a^2 \left(\frac{b}{a+b}\right) = ab = E(N)$ , which is the direct result of Theorem 1.

**Remark 4.** Application 2 can be also formulated as a well-known symmetric random walk, where one starts at zero and stops when he reaches  $a$  or  $-b$ ,  $a$  and  $b$  are positive integers. In this case,  $S_n$  represents the location after  $n$  steps, while  $S_N = a$  or  $-b$ . In general, a direct computation of  $E(N)$  requires a more intricate procedure. The readers can refer to [6, 8] for more details.

**APPLICATION 3 (THE NEGATIVE BINOMIAL RANDOM VARIABLE)**

Let  $N$  be the total number of tosses needed to obtain  $n$  heads, then  $N$  is a negative binomial random variable with parameters  $n$  and  $p=1/2$ . Let  $Y$  be the number of heads minus the number of tails in  $N$  tosses (i.e.  $Y = n - (N-n) = 2n - N$ ). By Theorem 1,  $\text{Var}(Y) = E(N) = 2n$ , which agrees with the fact that  $\text{Var}(Y) = \text{Var}(2n - N) = \text{Var}(N) = 2n$  when  $p = 1/2$ .

**APPLICATION 4 (THE BEST-OF- (2n - 1) SERIES)**

The best-of- $(2n-1)$  format refers to a head-to-head competition between two teams (or players) in which one team must win  $n$  games (i.e. the majority of the games) to win the series. For example, the most common playoff format in the major North American sports (such as NBA, MLB, NHL, etc) is best-of-seven. This format indicates that each team must win four out of seven games in order to win a playoff series. Suppose now Team I has a 50% chance of winning Team II in a single game and the best-of- $(2n-1)$  series is considered. Let  $N$  be the number of games two teams play until one of the teams wins  $n$  games. Therefore,

$$E(N) = \sum_{k=n}^{2n-1} kP(N = k) = \sum_{k=n}^{2n-1} 2k \binom{k-1}{n-1} 2^{-k} = 2n \sum_{k=n}^{2n-1} \binom{k}{n} 2^{-k}$$

Since it can be shown that  $\sum_{k=n}^{2n} \binom{k}{n} 2^{-k} = 1$  (see later for the proof), by Theorem 1 we have that

$$\text{Var}(S_N) = E(N) = 2n \left\{ 1 - \binom{2n}{n} 2^{-2n} \right\}$$

which is the variance of the difference between the number of games the winner wins and the number of games the loser wins.

Now we show how to prove  $\sum_{k=n}^{2n} \binom{k}{n} 2^{-k} = 1$ . Since

$$\sum_{k=n}^{2n} \binom{k}{n} 2^{-k} = \sum_{k=n}^{2n} \binom{k}{k-n} 2^{-k} = 2^{-n} \sum_{j=0}^n \binom{n+j}{j} 2^{-j}$$

it suffices to show  $\sum_{j=0}^n \binom{n+j}{j} 2^{-j} = 2^n$ . We prove this by mathematical induction. Let  $F(n) = \sum_{j=0}^n \binom{n+j}{j} 2^{-j}$ , first, it is

clear that  $F(2) = 1 + 3 \cdot 2^{-1} + 6 \cdot 2^{-2} = 2^2$ . Suppose it is true that  $F(n) = 2^n$ , then we need to show that  $F(n+1) = 2^{n+1}$ . To see this, note that

$$\begin{aligned} F(n+1) &= \sum_{j=0}^{n+1} \binom{n+1+j}{j} 2^{-j} = \sum_{j=0}^{n+1} \left(1 + \frac{j}{n+1}\right) \binom{n+j}{j} 2^{-j} \\ &= \sum_{j=0}^{n+1} \binom{n+j}{j} 2^{-j} + \frac{1}{(n+1)} \sum_{j=0}^{n+1} j \binom{n+j}{j} 2^{-j} \\ &= 2^n + \binom{2n+1}{n+1} 2^{-(n+1)} + \frac{1}{(n+1)} \sum_{j=0}^{n+1} j \binom{n+j}{j} 2^{-j} \end{aligned} \tag{4}$$

We next examine the summation term in the right hand side of Eq. (4). Note that

$$\begin{aligned} \sum_{j=0}^{n+1} j \binom{n+j}{j} 2^{-j} &= \sum_{j=1}^{n+1} j \binom{n+j}{j} 2^{-j} = \sum_{j=1}^n (n+j) \binom{n+j-1}{j-1} 2^{-j} \\ &= \sum_{i=0}^n (n+1+i) \binom{n+i}{i} 2^{-(i+1)} \\ &= (n+1) \sum_{i=0}^n \binom{n+i}{i} 2^{-(i+1)} + \sum_{i=0}^n i \binom{n+i}{i} 2^{-(i+1)} \end{aligned} \tag{5}$$

Let  $T(n) = \sum_{i=0}^n i \binom{n+i}{i} 2^{-i}$ , Eq. (5) then reduces to

$$\sum_{j=0}^{n+1} j \binom{n+j}{j} 2^{-j} = \left(\frac{n+1}{2}\right) F(n) + \frac{1}{2} T(n) = \left(\frac{n+1}{2}\right) 2^n + \frac{1}{2} T(n) \tag{6}$$

Since the term in the left hand side of Eq. (6) can be written as  $T(n) + (n+1) \binom{2n+1}{n+1} 2^{-(n+1)}$ , Eq. (6) can be further written as

$$T(n) + (n+1) \binom{2n+1}{n+1} 2^{-(n+1)} = \left(\frac{n+1}{2}\right) 2^n + \frac{1}{2} T(n)$$

based on which we obtain that

$$T(n) = (n+1) 2^n - (n+1) \binom{2n+1}{n+1} 2^{-n} \tag{7}$$

Taking the results of Eq. (7) and Eq. (6) into Eq. (4), we then have that

$$\begin{aligned} F(n+1) &= 2^n + \binom{2n+1}{n+1} 2^{-(n+1)} + \frac{1}{(n+1)} \left\{ \left(\frac{n+1}{2}\right) 2^n + \frac{1}{2} T(n) \right\} \\ &= 2^n + \binom{2n+1}{n+1} 2^{-(n+1)} + 2^{n-1} + 2^{n-1} - \binom{2n+1}{n+1} 2^{-(n+1)} \\ &= 2^{n+1} \end{aligned}$$

The proof is now complete.

**Remark 5.** The readers can refer to an alternative proof of  $E(N) = 2n \left\{ 1 - \binom{2n}{n} 2^{-2n} \right\}$  given in [14]. Based on the result, it is easy to show that  $E(N)$  (or  $\text{Var}(S_N)$ ) has the following properties: (i) It is increasing with  $n$ ; and (ii)  $E(N)/2n$  (or  $\text{Var}(S_N)/2n$ ) increases to 1 as  $n$  increases to infinity.

#### CONFLICT OF INTEREST

None declared.

#### ACKNOWLEDGEMENT

None declared.

#### REFERENCES

- [1] P. Billingsley, “*Probability and Measure*”, 3rd Ed. John Wiley & Sons, New York 1995.
- [2] R. Durrett, “*Probability: Theory and Examples*”, 3rd ed. Belmont CA: Thomson Brooks/Cole, 2005.
- [3] M. Abramowitz and I. A. Stegun, Eds., “*Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*”, 9th printing. New York: Dover, 1972.
- [4] B. D. Hughes, “*Random Walks and Random Environments*, Vol. 1: *Random Walks*”. New York: Oxford University Press, 1995.
- [5] S. M. Ross, “*Stochastic Processes*”, 2nd ed. New York-Chichester-Brisbane-Toronto-Singapore: John Wiley & Sons, 1996.
- [6] Y. S. Chow, H. Robbins, and D. Siegmund, “*Great Expectations: The Theory of Optimal Stopping*”, Boston: Houghton Mifflin, 1971.
- [7] Y. S. Chow and H. Teicher, “*Probability Theory: Independence, Interchangeability*”, Martingales, 3rd Ed. New York: Springer-Verlag, 2003.
- [8] A. N. Shiryaev, “*Probability*”, Berlin-Heidelberg-New York-Tokyo: Springer-Verlag, 1984.
- [9] A. Wald, “*Sequential Analysis*”, New York: John Wiley & Sons, 1947.
- [10] Y. C. Hung, R. Chen, A. Zame, and M. R. Chen, “A note on the first occurrence of strings”, *Electron. J. Comb.*, vol. 17, # N5, 2010.
- [11] R. Chen, A. Zame, and B. Rosenberg, “On the first occurrence of strings”, *Electron. J. Combin.*, vol. 16, pp. R29, 2009.
- [12] H.V. Gerber and S. Y. R. Li, “The occurrence of sequence patterns in repeated experiments and hitting times in a Markov chain”, *Stoch. Process. Appl.*, vol. 11, pp. 101-108, 1981.
- [13] R. Chen and A. Zame, “On the fair coin-tossing games”, *J. Multivariate Anal.*, vol. 9, pp. 150-157, 1979.
- [14] T. Lengyel, “A Combinatorial Identity and the World Series”, *SIAM Rev.*, vol. 35, pp. 294-297, 1993.

Received: December 20, 2011

Revised: February 16, 2012

Accepted: February 19, 2012

© Ying-Chao Hung; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.