ARS MATHEMATICA CONTEMPORANEA

Also available at http://amc-journal.eu
ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 12 (2017) 329–336

Integral representations for binomial sums of chances of winning

Zhao Dong *

Key Laboratory of Random Complex Structures and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, P. R. China

Wenbo V. Li

Department of Mathematical Sciences, University of Delaware, Newark, DE 19716

Chunwei Song †

School of Mathematical Sciences, & LMAM, Peking University, Beijing 100871, P. R. China

Received 22 March 2013, accepted 16 December 2016, published online 30 January 2017

Abstract

Addona, Wagon and Wilf (*Ars Math. Contemp.* 4 (1) (2011), 29-62) examined a problem about the winning chances in tossing unbalanced coins. Here we present some integral representations associated with such winning probabilities in a more general setting via using certain Fourier transform method. When our newly introduced parameters (r, d) are set to be $(0, 1)$, one of our results reduces to the main formula in the above reference.

Keywords: Binomial sum, integral representation, probabilistic analysis, unbalanced coin.

Math. Subj. Class.: 11B65, 33D45, 42A16, 60C05, 91A60

1 Introduction: How to make it fair and fun?

One of the final two papers published by Herbert S. Wilf (1932-2012) is a joint work with Addona and Wagon entitled "How to lose as little as possible", which investigates an intriguing problem of disadvantaged player Alice competing with Bob [1]:

*The first author was supported in part by NSFC Grants #11271356, #11371041 and #11431014.
†The third author was supported in part by the 973 Program project #2013CB834201.

E-mail addresses: dzhao@amt.ac.cn (Zhao Dong), wli@math.udel.edu (Wenbo V. Li), csong@math.pku.edu.cn (Chunwei Song)

Suppose Alice has a coin with probability of heads equal to q $(0 < q < 1)$, Bob has a different coin with probability of heads equal to p $(0 < p < 1)$, and that $q < p$. They toss their coins independently n times each. The rule says that Alice wins if and only if she gets strictly more heads than Bob does. Clearly, in the above setting Alice's odds of winning are

$$\mathbb{P}(S_n > T_n) = \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} \sum_{k=j+1}^{n} \binom{n}{k} q^k (1-q)^{n-k}, \tag{1.1}$$

where the random variable S_n (resp. T_n) stands for the number of heads that Alice gets (resp. Bob gets) after n tosses.

For convenience, let

$$f(n) = f(n, p, q) = \mathbb{P}(S_n > T_n). \tag{1.2}$$

In search of the choice of n that maximizes Alice's chances of winning, it is shown in [1] that $f(n)$ is essentially unimodal, and sharp bounds on the turning point $N(q, p)$ are given. Their analysis uses the multivariate form of Zeilbergers algorithm [2]. In particular, one of the main results of [1] that provides a key role in the proof of unimodality and in the derivation of the turning point is the following:

Theorem 1.1. *With $f(n)$ defined above,*

$$\frac{f(n+1) - f(n)}{((1-p)(1-q))^{n+1}} = (y + \frac{1}{2}(1+xy))\phi_n(xy) - \frac{1}{2}\phi_{n+1}(xy), \tag{1.3}$$

where $x = p/(1-p)$, $y = q/(1-q)$, $\phi_n(z) = \sum_{j=0}^{n} \binom{n}{j}^2 z^j = (1-z)^n P_n(\frac{1+z}{1-z})$, and $P_n(u)$ is the classical Legendre polynomial:

$$P_n(u) = \frac{1}{\pi} \int_0^{\pi} (u + \sqrt{u^2 - 1}\cos t)^n dt.$$

(Note that $|u| = |\frac{1+xy}{1-xy}| > 1$ for $xy \in (0,1) \cup (1, +\infty)$; the case $xy = 1$ yields $p = 1 - q$ and (1.3) may be verified directly from (1.1) without using the Legendre polynomial.)

Explicitly, the numerator of the left hand side of (1.3), which is the essential part, may be expressed via

$$f(n+1) - f(n) = \frac{1}{\pi} \int_0^{\pi} \psi^n(t)(q - pq - \sqrt{pq(1-p)(1-q)}\cos t)dt, \tag{1.4}$$

where $\psi(t) = 1 - p - q + 2pq + 2\sqrt{pq(1-p)(1-q)}\cos t$. In fact, the above expression (1.4) is found by first using the multivariate form of Zeilberger's algorithm and then proved mathematically (with ease once the formula is found). Also, it follows from (1.4) that the probability for Alice to win with n tosses is

$$\mathbb{P}(S_n > T_n) = \frac{1}{\pi} \int_0^{\pi} \frac{1 - \psi^n(t)}{1 - \psi(t)} (q - pq - \sqrt{pq(1-p)(1-q)}\cos t)dt.$$

To be fair and fun, here in this paper we consider a more general setting. Since Alice has a weaker coin, why should not she toss it for r more times than Bob does? And if that becomes the fact, maybe we should investigate the chances for Alice to get at least d ($d \geq 1$) more heads than Bob does.

Now formally, let $S_n \sim$ Bin (n, q) and $T_m \sim$ Bin (m, p). Suppose Alice tosses her coin $n+r$ ($r \geq 0$) times and Bob tosses his n times, and that Alice wins iff she gets at least d ($d \geq 1$) more heads than Bob does. Under the new rule, the probability for Alice to win is

$$\mathbb{P}(S_{n+r} \geq T_n + d) = \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} \sum_{k \geq j+d}^{n+r} \binom{n+r}{k} q^k (1-q)^{n+r-k}.$$

For convenience, we let

$$f_{r,d}(n) = f_{r,d}(n, p, q) = \mathbb{P}(S_{n+r} \geq T_n + d). \tag{1.5}$$

Apparently, the function $f(n)$ in this setting is $f_{0,1}(n)$.

In order to study the turning point, investigations of the difference $f_{r,d}(n+1) - f_{r,d}(n)$ are needed. In Section 2 we introduce probabilistic preliminaries with Fourier analysis blended. In Section 3 we provide several representations of the difference function based on certain trigonometric integrals.

Throughout this work we adopt the commonly used convention for the generalized binomial coefficients: for $\alpha \in \mathbb{R}$, $j \in \mathbb{Z}$,

$$\binom{\alpha}{j} = \begin{cases} \frac{\alpha(\alpha-1)\cdots(\alpha-j+1)}{j!}, & \text{if } j \geq 1; \\ 1, & \text{if } j = 0; \\ 0, & \text{if } j < 0. \end{cases}$$

2 Probabilistic analysis: Lens of Fourier method

To attack on $f_{r,d}(n+1) - f_{r,d}(n)$, we adopt the Fourier analysis approach used in [4], where the special case $p = q$ has been studied.

The following known fact [3, p. 95] will be useful in our studies. Let Z be an integer-valued random variable. It holds that

$$\mathbb{P}(Z = k) = \frac{1}{2\pi} \int_0^{2\pi} \varphi_Z(t) e^{-ikt} dt, \tag{2.1}$$

where $\varphi_Z(t)$ is the characteristic function of Z.

Lemma 2.1. *For any $r, d, n \in \mathbb{N}$,*

$$f_{r,d}(n+1) - f_{r,d}(n) = \sum_{j=0}^{r} \big(q(1-p)\mathbb{P}(T_n - S_n - k = j + 1 - d) - p(1-q) \\ \mathbb{P}(T_n - S_n - k = j - d) \big) \mathbb{P}(S_r = j).$$

Proof. For convenience, let $g(n, k) := \mathbb{P}(T_n - S_n - k = 0)$. Note that

$$f_{r,d}(n) = \mathbb{P}(S_{n+r} \geq T_n + d) = \mathbb{P}(S'_r \geq T_n - S_n + d) = \sum_{k=-n}^{n} g(n, k) \mathbb{P}(S'_r \geq k + d).$$

Here in this proof technically S'_r is independent of S_n and has the same distribution as S_r. Similarly,

$$\begin{aligned} f_{r,d}(n+1) &= \mathbb{P}(S'_r \geq T_n - S_n + Y_1 - X_1 + d) \\ &= \sum_{k=-n}^{n} g(n,k)\mathbb{P}(S'_r \geq Y_1 - X_1 + k + d). \end{aligned}$$

Comparing the two formulae above, we arrive at

$$\begin{aligned} &f_{r,d}(n+1) - f_{r,d}(n) \\ &= \sum_{k=-n}^{n} g(n,k)(p(1-q)\mathbb{P}(S'_r \geq k+d+1) + q(1-p)\mathbb{P}(S'_r \geq k+d-1) \\ &+ (2pq - p - q)\mathbb{P}(S'_r \geq k+d)) \\ &= \sum_{k=-n}^{n} g(n,k)(q(1-p)\mathbb{P}(S'_r = k+d-1) - p(1-q)\mathbb{P}(S'_r = k+d)) \\ &= \sum_{j=0}^{r} (q(1-p)g(n, j+1-d) - p(1-q)g(n, j-d))\mathbb{P}(S'_r = j). \end{aligned}$$

□

Corollary 2.2. *For any* $r, d, n \in \mathbb{N}$, $f_{r,d}(n+1) - f_{r,d}(n) =$

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi^n(t) \sum_{j=0}^{r} (\mathbb{P}(S_r = j))(q(1-p)e^{-i(j+1-d)t} - p(1-q)e^{-i(j-d)t}))dt,$$

where

$$\begin{aligned} \varphi(t) = \varphi(t,p,q) &:= Ee^{it(Y_1 - X_1)} \\ &= (1 - p + pe^{it})(1 - q + qe^{-it}) \\ &= 1 - p - q + 2pq + (p + q - 2pq)\cos t + i(p-q)\sin t \end{aligned}$$

is the characteristic function of $Y_1 - X_1$ *with that* $Y_1 \sim$ *Bin* $(1,p)$ *and* $X_1 \sim$ *Bin* $(1,q)$.

Proof. This is an immediate consequence of (2.1) and Lemma 2.1. □

Corollary 2.3. *More specifically, for* $r = 0$, $d = 1$, *and* $n \in \mathbb{N}$,

$$f_{0,1}(n+1) - f_{0,1}(n) = \frac{1}{2\pi}\int_0^{2\pi} \varphi^n(t)(q(1-p) - p(1-q)e^{it})dt. \tag{2.2}$$

In the case $r = 0$ and $d = 1$, the formula of Corollary 2.2 reduces to the formula by Addona et al [1] as will be shown in Example 3.3.

3 Integral representations

The difference $f_{r,d}(n+1)-f_{r,d}(n)$ may be evaluated directly. In fact it depends on certain trigonometric integrals. Before we proceed the following fact is needed.

Lemma 3.1. *For all nonnegative integers a, b, c, let $J(a,b,c)=\int_0^{2\pi}\cos^a t\sin^b t\cos(ct)dt$ and $K(a,b,c)=\int_0^{2\pi}\cos^a t\sin^b t\sin(ct)dt$. Then*

$$J(a,b,c)=2\pi\sum_s\frac{(-1)^{b/2+s}}{2^{a+b+1}}\binom{b}{s}\Big[\binom{a}{(a+b-c)/2-s}+\binom{a}{(a+b+c)/2-s}\Big]$$
$$K(a,b,c)=2\pi\sum_s\frac{(-1)^{(b-1)/2+s}}{2^{a+b+1}}\binom{b}{s}\Big[\binom{a}{(a+b-c)/2-s}-\binom{a}{(a+b+c)/2-s}\Big],$$

where for convenience we assume that $(-1)^u=0$ if $u\notin\mathbb{Z}$.

Proof. Note that $J(a,b,c)=0$ whenever b is odd, and that

$$\int_0^{2\pi}e^{itm}dt=\begin{cases}2\pi, & \text{if } m=0;\\ 0, & \text{otherwise.}\end{cases}$$

Hence we have

$$\begin{aligned}J(a,b,c)&=\int_0^{2\pi}\cos^a t\sin^b t\cos(ct)dt\\&=\int_0^{2\pi}(\frac{e^{it}+e^{-it}}{2})^a(\frac{e^{it}-e^{-it}}{2i})^b(\frac{e^{ict}+e^{-ict}}{2})dt\\&=\sum_{l,s}\frac{(-1)^{b-s}}{2^{a+b+1}i^b}\int_0^{2\pi}\binom{a}{l}e^{itl-it(a-l)}\binom{b}{s}e^{its-it(b-s)}(e^{ict}+e^{-ict})dt\\&=\sum_{l,s}\frac{(-1)^{b/2-s}\binom{a}{l}\binom{b}{s}}{2^{a+b+1}}(\int_0^{2\pi}e^{it(2l-a+2s-b+c)}dt+\int_0^{2\pi}e^{it(2l-a+2s-b-c)}dt)\\&=2\pi\sum_s\frac{(-1)^{b/2+s}}{2^{a+b+1}}\binom{b}{s}\left[\binom{a}{(a+b-c)/2-s}+\binom{a}{(a+b+c)/2-s}\right].\end{aligned}$$

A similar calculation yields the result for $K(a,b,c)$. □

Now we rewrite Corollary 2.2 in a more explicit form:

Theorem 3.2. *For any $r, d, n\in\mathbb{N}$,*

$$\begin{aligned}&f_{r,d}(n+1)-f_{r,d}(n)\\&=\frac{1}{2\pi}\int_0^{2\pi}\varphi^n(t)\sum_{j=0}^{r}[\binom{r}{j}q^j(1-q)^{r-j}(q(1-p)\cos(j+1-d)t-p(1-q)\cos(j-d)t\\&-q(1-p)i\sin(j+1-d)t+p(1-q)i\sin(j-d)t]dt,\end{aligned}$$

where

$$\varphi(t)=1-p-q+2pq+(p+q-2pq)\cos t+i(p-q)\sin t.$$

Note that by Theorem 3.2, $f_{r,d}(n+1) - f_{r,d}(n)$ reduces to

$$\frac{1}{2\pi}\sum_{j=0}^{r}\binom{r}{j}q^j(1-q)^{r-j}(I_1+I_2+I_3+I_4)$$

where the integrals I_k's ($1 \le k \le 4$) are defined as,

$$\begin{aligned}
I_1 &= \int_0^{2\pi} \varphi^n(t)q(1-p)\cos(j+1-d)tdt\\
I_2 &= -\int_0^{2\pi} \varphi^n(t)p(1-q)\cos(j-d)tdt\\
I_3 &= -\int_0^{2\pi} \varphi^n(t)q(1-p)i\sin(j+1-d)tdt\\
I_4 &= \int_0^{2\pi} \varphi^n(t)p(1-q)i\sin(j-d)t)dt.
\end{aligned}$$

All results reduce to such integrals. We shall represent $f_{r,d}(n+1) - f_{r,d}(n)$ by the "basis" $\int_0^{2\pi}\cos^{2k} tdt$, $k \ge 0$. To do this, we partition the real parts of $\varphi^n(t)(q(1-p)\cos(j+1-d)t$, etc, according to the "order" $2k$, i.e., those of the form constant times $\cos^a t\sin^b t\cos(ct)$ or $\cos^a t\sin^b t\sin(ct)$ where $2k-1 \le a+b \le 2k$. The notation $[\cos^{2k} t]_I f$ means that in the expansion of f, we group all "homogeneous" terms (of form $\cos^a t\sin^b t\cos(ct)$ or $\cos^a t\sin^b t\sin(ct)$ where $2k-1 \le a+b \le 2k$) and take the ratio of the integrals $\frac{\sum_{a,b}\int_0^{2\pi}\cos^a t\sin^b t\cos(ct)dt}{\int_0^{2\pi}\cos^{2k} tdt}$ or $\frac{\sum_{a,b}\int_0^{2\pi}\cos^a t\sin^b t\sin(ct)dt}{\int_0^{2\pi}\cos^{2k} tdt}$ as the corresponding coefficient. For example, based on Lemma 3.1, we have

$$\begin{aligned}
&[\cos^{2k} t]_I(\cos^{2k-1-2m} t\sin^{2m} t\cos(ct))\\
=&\frac{(2k)!!}{(2k-1)!!}\sum_s\frac{(-1)^{m+s}}{2^{2k}}\binom{2m}{s}(\binom{2k-1-2m}{(2k-1-c)/2-s}+\binom{2k-1-2m}{(2k-1+c)/2-s}),\\
&[\cos^{2k} t]_I(\cos^{2k-2m} t\sin^{2m} t\cos(ct))\\
=&\frac{(2k)!!}{(2k-1)!!}\sum_s\frac{(-1)^{m+s}}{2^{2k+1}}\binom{2m}{s}(\binom{2k-2m}{(2k-c)/2-s}+\binom{2k-2m}{(2k+c)/2-s}).
\end{aligned}$$

Similarly we obtain the formulas for $[\cos^{2k} t]_I(\cos^{2k-2m} t\sin^{2m-1} t\sin(ct))$ and $[\cos^{2k} t]_I(\cos^{2k+1-2m} t\sin^{2m-1} t\sin(ct))$. Consequently, $[\cos^{2k} t]_I\{I_1\}$, etc, may be found and the following theorem follows. To keep the cleanness we omit the proof details.

Theorem 3.3. *For $r, d \in \mathbb{N}$,*

$$f_{r,d}(n+1) - f_{r,d}(n) = \frac{1}{2\pi}\sum_j\binom{r}{j}q^j(1-q)^{r-j}\sum_k a_{n,k}(j,d)\int_0^{2\pi}\cos^{2k}(t)dt,$$

where (i) $a_{n,k}(j,d) =$

$$\frac{n!}{(2k)!(n+1-2k)!}\frac{(2k)!!\binom{2k}{k+(j-d+1)/2}}{(2k-1)!!}(1-p-q+2pq)^{n-2k}(q(1-p)p(1-q))^k$$
$$(q^{-1}(1-p)^{-1}p(1-q))^{\frac{j-d+1}{2}}\{(k+(j-d+1)/2)p+(n+1-k+(j-d+1)/2)q$$
$$-(n+2+j-d)pq-(k+(j-d+1)/2)\},\ \textit{if } j-d \textit{ is odd;}$$

and (ii) $a_{n,k}(j,d) =$

$$\frac{n!}{(2k)!(n+1-2k)!}\frac{(2k)!!\binom{2k}{k+(j-d)/2}}{(2k-1)!!}(1-p-q+2pq)^{n-2k}(q(1-p)p(1-q))^k$$
$$(q^{-1}(1-p)^{-1}p(1-q))^{\frac{j-d}{2}}\{(-n-1+k+(j-d)/2)p$$
$$-(k-(j-d)/2)q+(n+1-j+d)pq+k\},\ \textit{if } j-d \textit{ is even.}$$

We conclude with three examples.

Example 3.1.

$$f_{1,1}(n+1)-f_{1,1}(n) = \sum_k a_{n,k}(1,1)\frac{1}{2\pi}\int_0^{2\pi}\cos^{2k}(t)dt,$$

where

$$a_{n,k}(1,1) = \frac{n!}{(2k)!(n+1-2k)!}(1-p-q+2pq)^{n-2k}2^{2k}(1-q)$$
$$(pq(1-p)(1-q))^k((-n-1+k)p-kq+(n+1)pq+k).$$

Example 3.2. For $m \in \mathbb{N}$,

$$(i)\ f_{0,2m}(n+1)-f_{0,2m}(n) = \sum_k a_{n,k}(0,2m)\frac{1}{2\pi}\int_0^{2\pi}\cos^{2k}(t)dt,$$

where $a_{n,k}(0,2m) :=$

$$\frac{n!}{(2k)!(n+1-2k)!}(1-p-q+2pq)^{n-2k}\frac{(2k)!!\binom{2k}{k+m}}{(2k-1)!!}(pq(1-p)(1-q))^k$$
$$(q(1-p)p^{-1}(1-q)^{-1})^{-m}((-n-1+k-m)p-(k-m)q+(n+1+2m)pq+k).$$

$$(ii)\ f_{0,2m+1}(n+1)-f_{0,2m+1}(n) = \sum_k a_{n,k}(0,2m+1)\frac{1}{2\pi}\int_0^{2\pi}\cos^{2k}(t)dt,$$

where $a_{n,k}(0,2m+1) :=$

$$\frac{n!}{(2k)!(n+1-2k)!}(1-p-q+2pq)^{n-2k}\frac{(2k)!!\binom{2k}{k+m}}{(2k-1)!!}(pq(1-p)(1-q))^k$$
$$(q(1-p)p^{-1}(1-q)^{-1})^{-m}((k-m)p+(n+1-k-m)q+$$
$$(n+1-2m)pq+m-k).$$

Theorem 3.3 actually provides a perspective to generalize the Legendre type representations discussed in [1].

Finally we exhibit the equivalence of Corollary 2.3 and (1.4).

Example 3.3.

$$f_{0,1}(n+1)-f_{0,1}(n)=\frac{1}{\pi}\int_0^{\pi}\psi^n(t)(q-pq-\sqrt{pq(1-p)(1-q)}\cos t)dt,$$

where $\psi(t)=1-p-q+2pq+2\sqrt{pq(1-p)(1-q)}\cos t$.

Proof. In fact, specializing Theorem 3.3,

$$\begin{aligned}f_{0,1}(n+1)-f_{0,1}(n)&=\frac{1}{2\pi}\sum_{k=0}^{\lfloor (n+1)/2\rfloor}a_{n,k}(0,1)\int_0^{2\pi}\cos^{2k}t dt\\&=\frac{1}{\pi}\int_0^{\pi}\sum_{k=0}^{\lfloor (n+1)/2\rfloor}a_{n,k}(0,1)\cos^{2k}t dt\\&=\frac{1}{\pi}\int_0^{\pi}\psi^n(t)(q-pq-\sqrt{pq(1-p)(1-q)}\cos t)dt.\end{aligned}$$

□

Thus we have rediscovered (1.4).

Acknowledgement

One of our coauthors, Professor Wenbo V. Li passed away on January 26, 2013 [5]. Wenbo, dear friend and passionate mathematician, is to be memorized by many.

The authors wish to thank an anonymous referee for many helpful comments and suggestions.

References

[1] V. Addona, S. Wagon and H. Wilf, How to lose as little as possible, Ars Math. Contemp. **4** (2011), 29–62.

[2] M. Apagodu and D. Zeilberger, Multi-variable Zeilberger and Almkvist-Zeilberger algorithms and the sharpening of Wilf-Zeilberger theory, *Adv. in Appl. Math.* **37** (2006), 139–152.

[3] R. Durrett, *Probability: Theory and Examples*, 3rd edition, Brooks/Cole–Thomson Learning, Belmont, CA, 2005.

[4] W. V. Li and V. V. Vysotsky, Probabilities of competing binomial random variables, *J. Appl. Probab.* **49** (2012), 731–744.

[5] Univeristy of Delaware website, In Memoriam: Long-time UD math professor Wenbo V. Li dies, http://www1.udel.edu/udaily/2013/jan/Li012813.html.