# PROBABILITIES OF COMPETING BINOMIAL RANDOM VARIABLES

WENBO V. LI,\* University of Delaware

VLADISLAV V. VYSOTSKY,\*\* Arizona State University, Steklov Mathematical Institute and St. Petersburg State University

#### Abstract

Suppose that both you and your friend toss an unfair coin n times, for which the probability of heads is equal to  $\alpha$ . What is the probability that you obtain at least d more heads than your friend if you make r additional tosses? We obtain asymptotic and monotonicity/convexity properties for this competing probability as a function of n, and demonstrate surprising phase transition phenomenon as the parameters d, r, and  $\alpha$  vary. Our main tools are integral representations based on Fourier analysis.

*Keywords:* Binomial random variable; number of successes; competing random variables; probability of winning; coin tossing; phase transition

2010 Mathematics Subject Classification: Primary 60B99; 60F99; 42A61

#### 1. Introduction

Suppose that both you and your friend toss an unfair coin n times, for which the probability of heads is equal to  $\alpha$ . What is the competing probability that you obtain at least d more heads than your friend if you make r additional tosses?

For a fair coin with  $\alpha = \frac{1}{2}$  and r = d = 1, this is Example 3.33 of [3, p. 118]: 'Adam tosses a fair coin n + 1 times, Andrew tosses the same coin n times. What is the probability that Adam gets more heads than Andrew?'

Two solution methods are offered in [3]. One method uses the symmetry, easily yielding the answer  $\frac{1}{2}$ . The other method is a direct computation, using the identity

$$\sum_{i=0}^{n} \sum_{j=i+1}^{n+1} \binom{n+1}{j} \binom{n}{i} = 2^{2n}; \tag{1}$$

the author noted that 'a combinatorial solution to this problem is neither elegant nor easy to handle'.

The same problem also appeared as Problem 21 of the self-test problems and exercises on page 115 of [7]: 'if A flips n + 1 and B flips n fair coins, show that the probability that A gets more heads than B is  $\frac{1}{2}$ .' This problem appears at the end of the chapter on conditional

Received 28 January 2011; revision received 23 January 2012.

<sup>\*</sup> Postal address: Department of Mathematical Sciences, University of Delaware, 501 Ewing Hall, Newark, DE 19716, USA. Email address: wli@math.udel.edu

Supported in part by NSF grants DMS-0805929, NSFC-6398100, and CAS-2008DP173182.

<sup>\*\*</sup> Postal address: School of Mathematical and Statistical Sciences, Arizona State University, PO Box 871804, Tempe, AZ 85287-1804, USA. Email address: vysotsky@asu.edu

The author began this work at the University of Delaware, and was supported in part by the grant NSh. 4472-2010-1.

probability and independence, along with the hint that one should 'condition on which player has more heads after each has flipped *n* coins'.

What happens if the coin is unfair is not mentioned in either [3] or [7]. In this paper we consider the competing probability of two independent binomials and the associated phase transition behaviors as the parameters vary.

To pose the problem formally, let  $S_k$  and  $S'_n$  be independent binomial random variables, that is,  $S_k = X_1 + \cdots + X_k$  and  $S'_n = X'_1 + \cdots + X'_n$ , where  $\{X_i, X'_i\}_{i \ge 1}$  are independent, identically distributed random variables that equal 1 with probability  $\alpha$  and 0 with probability  $1 - \alpha$ . In the textbooks [3] and [7] the probabilities

$$p_n := P\{S_{n+1} \ge S'_n + 1\}$$

were considered, while we are interested in the more general

$$p_n^{r,d} := P\{S_{n+r} \ge S_n' + d\},\$$

where, of course,  $p_n^{1,1} = p_n$ ; we always assume that  $\alpha \in (0, 1)$  and  $r, d \ge 1$ . Clearly, we can write the combinatorial expression

$$p_n^{r,d} = \sum_{i=0}^n \sum_{j=i+d}^{n+r} P\{S_{n+r} = j\} P\{S_n' = i\}$$

$$= \sum_{i=0}^n \sum_{j=i+d}^{n+r} \binom{n+r}{j} \binom{n}{i} \alpha^{i+j} (1-\alpha)^{2n+r-i-j}$$
(2)

by using independence and binomial probabilities. Unlike (1), (2) is of little use in analyzing finer behaviors as the parameters vary. On the other hand, by the central limit theorem, for any r, d, and  $\alpha$  as above,

$$\lim_{n \to \infty} p_n^{r,d} = \lim_{n \to \infty} P\left\{ \frac{S_{n+r}}{\sqrt{n}} \ge \frac{S_n' + d}{\sqrt{n}} \right\} = P\{N > N'\} = \frac{1}{2},\tag{3}$$

where N and N' are independent normal random variables with mean 0 and variance  $\alpha(1-\alpha)$ . This standard technique allows us to find the limit, but tells us nothing about the mode and exact rate of convergence. The following theorem provides a useful integral representation for  $p_n^{r,d}$  which implies a precise asymptotic.

**Theorem 1.** For any positive integers r and d,

$$p_n^{r,d} = \frac{1}{2} + \frac{1}{\pi} \int_0^1 Q_\alpha^n(x) (1 - x^2)^{-1/2} P_\alpha^{r,d}(x) \, \mathrm{d}x,\tag{4}$$

where

$$Q_{\alpha}(x) := 1 - 4\alpha(1 - \alpha)(1 - x^{2}) \tag{5}$$

and

$$P_{\alpha}^{r,d}(x) := \sum_{j=0}^{r+d} \frac{\mathrm{d}^{2j}}{\mathrm{d}t^{2j}} (t^{j-d} (1 - \alpha + \alpha t)^r) \bigg|_{t=-1} \frac{(2x)^{2j}}{(2j)!}.$$
 (6)

As a consequence,

$$\lim_{n \to \infty} \sqrt{n} \left( p_n^{r,d} - \frac{1}{2} \right) = \frac{1}{4\sqrt{\pi}} \frac{2\alpha r - 2d + 1}{\sqrt{\alpha(1 - \alpha)}}.$$
 (7)

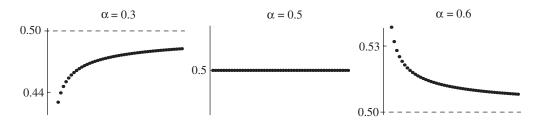


FIGURE 1: The sequence  $p_n$ ,  $0 \le n \le 50$ , for different  $\alpha$ .

The actual degree of  $P_{\alpha}^{r,d}(x)$  is  $2\max(r-d,d-1)$  as the higher coefficients in (6) vanish. In Section 2 we give another convenient formula for  $P_{\alpha}^{r,d}(x)$  in terms of Chebyshev polynomials of the second kind, and then easily show that

$$P_{\alpha}^{r,d}(1) = 2\alpha r - 2d + 1.$$

Note that, for any  $\alpha \in (0, 1)$ , the function  $Q_{\alpha}(x)$  is increasing on [0, 1] with  $Q_{\alpha}(0) \geq 0$  and  $Q_{\alpha}(1) = 1$ . Therefore, as n increases, the main contribution to the integral in (4) comes from a decreasing small neighborhood of 1, and a standard analysis of (4) implies (7). Another advantage of the integral representation (4) stems from the isolation of the variable n from the parameters r and d.

Since the integral representation for  $p_n^{r,d}$  in (4) can be integrated out via the trigonometric substitution  $x = \cos(t/2)$ , it seems possible to check the equivalence of (2) and (4) by pure algebraic manipulations. However, such an approach provides no probabilistic insights into the integral representation and no clues on how we discovered it. We will present a proof, via a combination of one-step-back analysis and Fourier analytic methods, along the same lines as our initial derivation of the representation.

Next we consider the mode of convergence in terms of monotonicity and convexity properties in the simplest setting, r=d=1. Throughout this paper, by 'increasing/decreasing' we mean 'strictly increasing/strictly decreasing'.

**Theorem 2.** The sequence  $p_n = p_n^{1,1} = P\{S_{n+1} \ge S_n' + 1\} = P\{S_{n+1} > S_n'\}, \ n = 1, 2, ...,$  is monotone and convex/concave. More precisely,  $p_n$  is increasing and concave when  $\alpha < \frac{1}{2}$ , decreasing and convex when  $\alpha > \frac{1}{2}$ , and equal to  $\frac{1}{2}$  for all n when  $\alpha = \frac{1}{2}$ .

It is remarkable that the sequence  $p_n^{1,1}$  is monotone starting from its first term. We first discovered this feature when studying some very specific probabilistic properties of the so-called double-sided exponential random variables; see the discussion at the end of the paper.

Observe that, as  $\alpha$  decreases from 1 to 0, the behavior of the sequence  $p_n^{1,1}$  changes instantly at the critical value  $\frac{1}{2}$ ; see Figure 1. This phenomenon is more pronounced in the  $r \geq 2$  and d=1 setting. In order to give the best statement of our results, from this point on we assume that  $n \geq 0$  rather than  $n \geq 1$ , which was natural for the introduction of the problem. We say that a sequence  $a_n, n \geq 0$ , is unimodal with the mode  $N \geq 1$  if  $a_0 < a_1 < \cdots < a_{N-1} \leq a_N > a_{N+1} > a_{N+2} > \cdots$ .

**Theorem 3.** For  $r \ge 2$ , the sequence  $p_n^{r,1}$  is increasing when  $0 < \alpha \le 1/(2r)$ , unimodal when  $1/(2r) < \alpha \le 1/(r+1)$ , and decreasing when  $1/(r+1) < \alpha < 1$ . In addition, the mode  $N_{\alpha,r}$ 

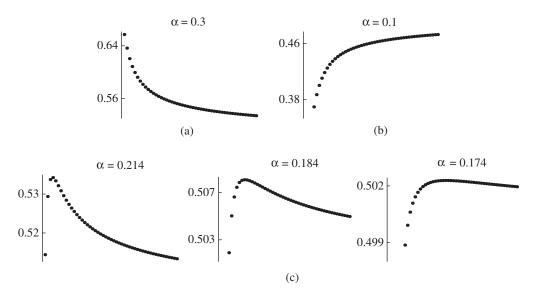


FIGURE 2: The sequence  $p_n^{3,1}$ ,  $0 \le n \le 50$ : monotone modes (a) and (b) and the phase transition (c) between them for decreasing values of  $\alpha$ . In the transitional mode (c) the point of maximum drifts to the right.

satisfies

$$N_{\alpha,r} \sim \frac{r-1}{4r} \left(\alpha - \frac{1}{2r}\right)^{-1} \quad as \ \alpha \downarrow \frac{1}{2r},$$
 (8)

while, for the maximum,

$$p_{N_{\alpha},r} - \frac{1}{2} \sim \frac{4}{3} \sqrt{\frac{r^5}{\pi (2r-1)(r-1)}} \left(\alpha - \frac{1}{2r}\right)^{3/2} \quad as \ \alpha \downarrow \frac{1}{2r}.$$
 (9)

In Figure 2 we show the peculiar 'phase transition' in the behavior of  $p_n^{r,1}$  as  $\alpha$  decreases. For large  $\alpha$ , the sequence is decreasing while, as  $\alpha$  gets smaller, reaching the critical value 1/(r+1),  $p_n^{r,1}$  becomes unimodal and the point of maximum drifts to the right as  $\alpha$  decays to 1/(2r); see Figure 2(c). When  $\alpha$  reaches the critical value 1/(2r),  $p_n^{r,1}$  becomes increasing, which corresponds to the limit case  $N_{\alpha,r} = \infty$ . It is indeed remarkable that in the transitional mode the sequence  $p_n^{r,1}$  is unimodal. Relations (8) and (9) describe the 'speed' of the phase transition near the critical value  $\alpha = 1/(2r)$ .

For the general setting, as parameters  $\alpha$ , d, and r vary, we have the following slightly less precise result which still captures the phase transition phenomenon between the monotone modes of convergence.

**Theorem 4.** (i) If  $1 \le r \le d-1$ , the sequence  $p_n^{r,d}$  is increasing (for any  $0 < \alpha < 1$ ).

- (ii) If  $d \le r \le 2d 2$ , the sequence  $p_n^{r,d}$  is
  - (a) increasing (for all n) when  $0 < \alpha < d/(r+1)$ ;
  - (b) increasing for large n when  $d/(r+1) \le \alpha < (2d-1)/(2r)$ ;
  - (c) decreasing for large n when  $(2d-1)/(2r) \le \alpha < 1$ .

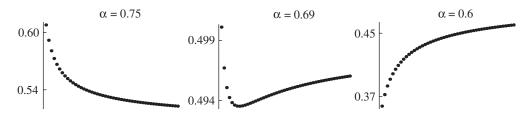


FIGURE 3: Phase transition in the behavior of  $p_n^{5,4}$ ,  $0 \le n \le 50$ , for decreasing values of  $\alpha$ .

- (iii) If r = 2d 1, the sequence  $p_n^{r,d}$  is
  - (a) increasing when  $0 < \alpha < \frac{1}{2}$ ;
  - (b) identically equal to  $\frac{1}{2}$  when  $\alpha = \frac{1}{2}$ ;
  - (c) decreasing when  $\frac{1}{2} < \alpha < 1$ .
- (iv) If  $r \geq 2d$ , the sequence  $p_n^{r,d}$  is
  - (a) increasing for large n when  $0 < \alpha \le (2d 1)/(2r)$ ;
  - (b) decreasing for large n when  $(2d-1)/(2r) < \alpha \le d/(r+1)$ ;
  - (c) decreasing (for all n) when  $d/(r+1) < \alpha < 1$ .

Theorem 4(i) is in some sense degenerate while (ii) uncovers another type of behavior of  $p_n^{r,d}$ ; see Figure 3. However, it is sufficient to concentrate on the case  $r \ge 2d - 1$ , which is already familiar from Theorems 2 and 3, as all the results of (ii) for  $d \le r \le 2d - 2$  follow from (iv) by a certain duality relation given in Section 3.

## 2. One-step-back analysis and Fourier method

It is natural to look at the increments of  $p_n^{r,d}$  in n. We start with the elementary but very useful one-step-back analysis.

**Lemma 1.** For any  $r, d \ge 1$ ,

$$p_{n+1}^{r,d} - p_n^{r,d} = \alpha (1 - \alpha) \sum_{i=0}^r P\{S_r = i\} (q_n^{(i-d+1)} - q_n^{(i-d)}),$$

where the sequence, defined by

$$q_n^{(i)} := P\{S_n' - S_n = i\},\$$

is decreasing in i for  $i \ge 0$  and satisfies  $q_n^{(-i)} = q_n^{(i)}$ .

Proof. We write

$$p_n^{r,d} = P\{S_{n+r} \ge S_n' + d\} = P\{S_r'' \ge S_n' - S_n + d\} = \sum_{i=-n}^n q_n^{(i)} P\{S_r \ge i + d\}$$

with  $S_r'' := S_{n+r} - S_r$ , and, similarly,

$$p_{n+1}^{r,d} = P\{S_r'' \ge S_n' + X_{n+1}' - S_n - X_{n+1} + d\} = \sum_{i=-n}^n q_n^{(i)} P\{S_r \ge X_1' - X_1'' + i + d\}.$$

Hence,

$$\begin{aligned} p_{n+1}^{r,d} - p_n^{r,d} &= \alpha (1 - \alpha) \sum_{i = -n}^n q_n^{(i)} (P\{S_r \ge i + d + 1\} + P\{S_r \ge i + d - 1\} \\ &- 2 P\{S_r \ge i + d\}) \\ &= \alpha (1 - \alpha) \sum_{i = -n}^n q_n^{(i)} (P\{S_r = i + d - 1\} - P\{S_r = i + d\}), \end{aligned}$$

which yields the first statement of the lemma.

The relation  $q_n^{(-i)} = q_n^{(i)}$  is obvious. Let us use induction to show that  $q_n^{(i)} > q_n^{(i+1)}$  for any n and  $0 \le i \le n$ . The initial case n = 1 is trivial. We use the law of total probability to get from n to n + 1:

$$q_{n+1}^{(i)} = \alpha (1 - \alpha)(q_n^{(i-1)} + q_n^{(i+1)}) + (\alpha^2 + (1 - \alpha)^2)q_n^{(i)}.$$

Then

$$q_{n+1}^{(i)} - q_{n+1}^{(i+1)} = \alpha(1-\alpha)(q_n^{(i-1)} - q_n^{(i+2)}) + (\alpha^2 + (1-\alpha)^2 - \alpha(1-\alpha))(q_n^{(i)} - q_n^{(i+1)}),$$

which is positive by the induction hypothesis.

In order to gain more quantitative information, we employ the powerful Fourier method. Recall that, for any integer-valued random variable Z, it holds that

$$P\{Z = 0\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} E e^{itZ} dt.$$
 (10)

Indeed,

$$\int_{-\pi}^{\pi} E e^{itZ} dt = \int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} e^{itk} P\{Z=k\} dt = \sum_{k=-\infty}^{\infty} \int_{-\pi}^{\pi} e^{itk} dt P\{Z=k\} = 2\pi P\{Z=0\}.$$

**Lemma 2.** For any  $r, d \ge 1$ ,

$$p_{n+1}^{r,d} - p_n^{r,d} = -\frac{4\alpha(1-\alpha)}{\pi} \int_0^1 Q_\alpha^n(x) \sqrt{1-x^2} \tilde{P}_\alpha^{r,d}(x) dx,$$

where

$$\tilde{P}_{\alpha}^{r,d}(x) := \sum_{i=d}^{r} P\{S_r = i\} U_{2(i-d)}(x) - \sum_{i=0}^{d-1} P\{S_r = i\} U_{2(d-i-1)}(x)$$
(11)

is defined in terms of Chebyshev polynomials  $U_k(x)$  of the second kind.

Recall that Chebyshev polynomials of the second kind, defined by  $U_k(x) = \sin(k+1)\theta/\sin\theta$  with  $x = \cos\theta$ , are orthogonal on [-1, 1] with weight  $\sqrt{1-x^2}$ . It is remarkable that  $\tilde{P}_{\alpha}^{r,d}(x)$  is explicitly expressed in terms of  $U_k(x)$ , whose properties are well known. In the next section we will show that  $\tilde{P}_{\alpha}^{r,d}(x) = P_{\alpha}^{r,d}(x)$ , and, thus, (11) actually serves as another useful representation of the polynomials defined in Theorem 1.

*Proof of Lemma 2.* With the reminder  $q_n^{(k)} = P\{S_n' - S_n = k\}$ , we use (10) to write

$$q_n^{(k)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{E} \, e^{it(S_n - S_n' - k)} \, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_\alpha^n(t) e^{-ikt} \, dt,$$

where

$$\varphi_{\alpha}(t) := (1 - \alpha + \alpha e^{it})(1 - \alpha + \alpha e^{-it}) = 1 - 4\alpha(1 - \alpha)\sin^2\frac{1}{2}t$$

is the characteristic function of  $S_1 - S'_1$ . By the symmetry of  $\varphi_{\alpha}(t)$ ,

$$q_n^{(k)} = \frac{1}{\pi} \int_0^{\pi} \varphi_\alpha^n(t) \cos kt \, \mathrm{d}t. \tag{12}$$

Combining (12) with Lemma 1, we obtain

$$p_{n+1}^{r,d} - p_n^{r,d} = \frac{\alpha(1-\alpha)}{\pi} \int_0^{\pi} \varphi_{\alpha}^n(t) \sum_{i=0}^r P\{S_r = i\} (\cos(i-d+1)t - \cos(i-d)t) dt.$$
 (13)

As for  $k \geq 0$ , it holds that

$$\cos(k+1)t - \cos kt = -2\sin^2\frac{t}{2} \cdot \frac{\sin(k+t/2)}{\sin(t/2)} = -2\sin^2\frac{t}{2} \cdot U_{2k}\left(\cos\frac{t}{2}\right);$$

thus, (13) becomes

$$p_{n+1}^{r,d} - p_n^{r,d} = -\frac{2\alpha(1-\alpha)}{\pi} \int_0^{\pi} \varphi_{\alpha}^n(t) \sin^2 \frac{t}{2} \cdot \tilde{P}_{\alpha}^{r,d} \left(\cos \frac{t}{2}\right) dt,$$

and we conclude the proof with the change of variable  $x = \cos(t/2)$ .

### 3. Proofs of the theorems

Proof of Theorem 1. First we check that

$$\tilde{P}_{\alpha}^{r,d}(x) = P_{\alpha}^{r,d}(x). \tag{14}$$

Using the representation

$$U_{2k}(x) = \sum_{j=0}^{k} (-1)^j \binom{2k-j}{j} (2x)^{2k-2j} = \sum_{j=0}^{k} (-1)^{k-j} \binom{k+j}{2j} (2x)^{2j}$$

from Section 6.10.7 of [8], we obtain

$$\begin{split} \tilde{P}_{\alpha}^{r,d}(x) &= \sum_{i=d}^{r} \sum_{j=0}^{i-d} P\{S_r = i\} (-1)^{i-j-d} \binom{i+j-d}{2j} (2x)^{2j} \\ &- \sum_{i=0}^{d-1} \sum_{j=0}^{d-i-1} P\{S_r = i\} (-1)^{d-i-j-1} \binom{d-i+j-1}{2j} (2x)^{2j}. \end{split}$$

We use the conventions  $\binom{n}{k} := (1/k!) \prod_{m=n-k+1}^{n} m$  and  $\binom{n}{0} := 1$  to write binomial coefficients with negative n. This allows us to note that

$$\binom{d-i+j-1}{2j} = \binom{i+j-d}{2j}$$

for any  $i, j, d \ge 0$ . The next important observation is that in both double sums the summation in j could be taken from 0 to  $\infty$ . Indeed, any  $j \ge i - d + 1$  does not give a contribution to the first sum as the product  $\prod_{m=i-j-d+1}^{i+j-d} m$ , which corresponds to the binomial coefficient, includes a zero factor because i+j-d>0 and  $i-j-d+1\le 0$ . The same applies to the second double sum for  $j \ge d-i$  as the product  $\prod_{m=d-i-j}^{d-i+j-1} m$  is 0 since d-i+j-1>0 while  $d-i-j\le 0$ . Thus,

$$\tilde{P}_{\alpha}^{r,d}(x) = \sum_{i=0}^{r} \sum_{j=0}^{\infty} P\{S_r = i\}(-1)^{i-j-d} \binom{i+j-d}{2j} (2x)^{2j}$$
$$= \sum_{j=0}^{\infty} \frac{d^{2j}}{dt^{2j}} (E t^{S_r+j-d}) \Big|_{t=-1} \frac{(2x)^{2j}}{(2j)!}.$$

Then (14) follows since the generating function of  $S_r$  is  $\operatorname{E} t^{S_r} = (1 - \alpha + \alpha t)^r$ . It now remains to use Lemma 2 and (11) to write the telescoping sum

$$p_{n+k}^{r,d} - p_n^{r,d} = \sum_{i=0}^{k-1} p_{n+i+1}^{r,d} - p_{n+i}^{r,d} = -\frac{4\alpha(1-\alpha)}{\pi} \int_0^1 \sum_{i=0}^{k-1} Q_\alpha^{n+i}(x) \sqrt{1-x^2} P_\alpha^{r,d}(x) dx.$$

As  $k \to \infty$ ,  $p_{n+k}^{r,d} \to \frac{1}{2}$  by (3), and simplifying the sum of the geometric series with (5), we obtain (4) by the dominated convergence theorem.

For the precise asymptotic relation (7), we return to the convenient trigonometric substitution  $x = \cos(t/2)$  in (4) and obtain

$$p_n^{r,d} - \frac{1}{2} = \frac{1}{2\pi} \int_0^{\pi} \varphi_{\alpha}^n(t) P_{\alpha}^{r,d} \left(\cos \frac{t}{2}\right) dt,$$

with  $\varphi_{\alpha}(t) = Q_{\alpha}(\cos(t/2)) = 1 - 4\alpha(1 - \alpha)\sin^2(t/2)$ . Choose a  $\delta > 0$  such that  $t/4 \le \sin(t/2) \le t/2$  on  $[0, \delta]$ , and observe that  $\varphi_{\alpha}(t)$  is decreasing on  $[0, \pi]$  and  $\varphi_{\alpha}(0) = 1$ . We have

$$\int_{0}^{\pi} \varphi_{\alpha}^{n}(t) P_{\alpha}^{r,d} \left(\cos \frac{t}{2}\right) dt = \int_{0}^{\delta} \varphi_{\alpha}^{n}(t) P_{\alpha}^{r,d} \left(\cos \frac{t}{2}\right) dt + O(\varphi_{\alpha}^{n}(\delta))$$

$$= \frac{1}{\sqrt{n}} \int_{0}^{\delta \sqrt{n}} \varphi_{\alpha}^{n} \left(\frac{s}{\sqrt{n}}\right) P_{\alpha}^{r,d} \left(\cos \frac{s}{2\sqrt{n}}\right) ds + O(\varphi_{\alpha}^{n}(\delta)), \quad (15)$$

where

$$\lim_{n \to \infty} \varphi_{\alpha}^{n} \left( \frac{s}{\sqrt{n}} \right) = \lim_{n \to \infty} \left( 1 - 4\alpha (1 - \alpha) \sin^{2} \frac{s}{2\sqrt{n}} \right)^{n} = e^{-\alpha (1 - \alpha)s^{2}}.$$
 (16)

Now

$$\lim_{n \to \infty} \int_0^\infty \mathbf{1}_{[0,\delta\sqrt{n}]}(s) \varphi_\alpha^n \left(\frac{s}{\sqrt{n}}\right) P_\alpha^{r,d} \left(\cos\frac{s}{2\sqrt{n}}\right) ds = \int_0^\infty e^{-\alpha(1-\alpha)s^2} P_\alpha^{r,d}(1) ds$$
$$= \frac{1}{2} \sqrt{\frac{\pi}{\alpha(1-\alpha)}} P_\alpha^{r,d}(1) \tag{17}$$

by the dominated convergence theorem, with  $e^{-\alpha(1-\alpha)s^2/4} \sup_{0 \le x \le 1} |P_{\alpha}^{r,d}(x)|$  an integrable majorant. Upon combining the above arguments we can complete the proof of Theorem 1 once we find  $P_{\alpha}^{r,d}(1)$ .

It is well known that  $U_k(1) = k + 1$ ; hence, (11) and (14) imply that

$$P_{\alpha}^{r,d}(1) = \sum_{i=d}^{r} P\{S_r = i\}(2i - 2d + 1) - \sum_{i=0}^{d-1} P\{S_r = i\}(2d - 2i - 1)$$

$$= E(2S_r - 2d + 1)$$

$$= 2\alpha r - 2d + 1.$$
(18)

*Proof of Theorem 2.* As  $P_{\alpha}^{1,1}(x) = 2\alpha - 1$ , Lemma 2 yields

$$p_{n+1} - p_n = \frac{4\alpha(1-\alpha)(1-2\alpha)}{\pi} \int_0^1 Q_{\alpha}^n(x) \sqrt{1-x^2} \, \mathrm{d}x.$$

It is readily seen that  $p_n$  is monotone because the integrand is nonnegative; moreover, the integrand is monotone in n for each x, so  $p_{n+1} - p_n$  is monotone, implying the convexity of  $p_n$ .

Note that the monotonicity of  $p_n$  could be obtained directly from Lemma 1, whose proof requires only an elementary one-step-back analysis. The same is true for the convexity of  $p_n$ , but an additional study of the properties of  $q_n^{(i)}$  should be done.

It is also worth mentioning that the asymptotic of  $p_n - \frac{1}{2}$  can be found via a purely probabilistic argument, without the need to use the Fourier method. Indeed, the trivial identities  $1 = 2 P\{S_n > \tilde{S}_n\} + P\{S_n = \tilde{S}_n\}$  and  $P\{S_{n+1} > \tilde{S}_n\} = \alpha P\{S_n = \tilde{S}_n\} + P\{S_n > \tilde{S}_n\}$  imply that

$$p_n = \frac{1}{2} + \frac{2\alpha - 1}{2} P\{S_n = \tilde{S}_n\}.$$

The asymptotic of the probability on the right-hand side is given by the classical local limit theorem.

We now skip to the proof of Theorem 4 as it partly covers Theorem 3.

*Proof of Theorem 4.* For  $1 \le r \le d-1$ , the result immediately follows from Lemma 1. For  $r \ge 2d-1$ , we start with the proofs of cases (iii)(c) and (iv)(c). By Lemma 1 we write

$$p_{n+1}^{r,d} - p_n^{r,d} = \alpha (1 - \alpha) \left( \sum_{i=0}^{d-1} (P\{S_r = i\} - P\{S_r = 2d - i - 1\}) (q_n^{(i-d+1)} - q_n^{(i-d)}) + \sum_{i=2d}^{r} P\{S_r = i\} (q_n^{(i-d+1)} - q_n^{(i-d)}) \right).$$

The second sum is nonpositive. We now show that

$$a_i := \frac{P\{S_r = 2d - i - 1\}}{P\{S_r = i\}} > 1, \qquad 0 \le i \le d - 1,$$

when  $\alpha > d/(r+1)$  to prove that the first sum is negative. This statement generalizes the well-known result that the (last) maximum of binomial coefficients  $P\{S_r = i\}$  occurs at  $i = [(r+1)\alpha]$ .

Consider the ratio

$$\frac{a_{i+1}}{a_i} = \frac{(1-\alpha)^2(i+1)(2d-i-1)}{\alpha^2(r-i)(r-2d+i+2)}$$

and rewrite it in the form

$$\left(\frac{1-\alpha}{\alpha}\right)^2 \frac{d^2 - (d-i-1)^2}{(r-d+1)^2 - (d-i-1)^2} = \left(\frac{1-\alpha}{\alpha}\right)^2 \left(1 - \frac{(r-d+1)^2 - d^2}{(r-d+1)^2 - (d-i-1)^2}\right)$$

to observe that this quantity increases in i for  $0 \le i \le d-1$  when  $r \ge 2d$  and is constant when r = 2d - 1. Now, as  $\alpha > d/(r+1)$ , in both cases we have

$$\frac{a_{i+1}}{a_i} \le \frac{a_d}{a_{d-1}} = \left(\frac{1-\alpha}{\alpha}\right)^2 \frac{d^2}{(r-d+1)^2} < 1,$$

and, thus,  $a_i$  is decreasing. Then

$$a_i \ge a_{d-1} = \frac{P\{S_r = d\}}{P\{S_r = d - 1\}} > 1$$

because  $\alpha > d/(r+1)$ .

Next we present the proofs of cases (iii)(a), (iv)(a), and (iv)(b). First note that  $\alpha < d/(r+1) \le \frac{1}{2}$ . Then, by Lemma 2, the sign of  $p_{n+1}^{r,d} - p_n^{r,d}$  for large n is opposite to the sign of  $\tilde{P}_{\alpha}^{r,d}(1) = P_{\alpha}^{r,d}(1) = 2\alpha r - 2d + 1$ ; see (14) and (18).

The case  $\alpha = (2d-1)/(2r)$  requires more attention. Here  $P_{\alpha}^{r,d}(1) = 0$  and the sign of  $P_{n+1}^{r,d} - P_n^{r,d}$  for large n coincides with that of  $\partial P_{\alpha}^{r,d}(1)/\partial x$ . We use the formulae  $T_n'(x) = nU_{n-1}(x)$  and  $T_n''(1) = (n-1)n^2(n+1)/3$ , where  $T_n(x) = \cos(n\arccos x)$  are Chebyshev polynomials of the first kind, to find that  $U_n'(1) = n(n+1)(n+2)/3$ . Hence, arguing as in the proof of (18), we obtain

$$\frac{\partial}{\partial x} P_{\alpha}^{r,d}(1) = \frac{1}{3} E(2S_r - 2d)(2S_r - 2d + 1)(2S_r - 2d + 2).$$

As the expectation on the right-hand side equals

$$\left. \frac{\partial^3}{\partial t^3} (E t^{2S_r - 2d + 2}) \right|_{t=1} = \left. \frac{\partial^3}{\partial t^3} ((1 - \alpha + \alpha t^2)^r t^{2 - 2d}) \right|_{t=1},$$

we substitute  $\alpha = (2d - 1)/(2r)$  and after some simplifications obtain

$$\frac{\partial}{\partial x} P_{(2d-1)/(2r)}^{r,d}(1) = \frac{2(2d-1)(1+4d^2+3r+2r^2-2d(2+3r))}{3r^2}.$$
 (19)

In order to check that this expression is positive, note that the partial derivative in r of the last factor in the numerator is equal to 4r + 3 - 6d > 0 as  $r \ge 2d$ . Hence, the minimal value of this factor is attained at r = 2d and equals 2d + 1 > 0.

For  $d \le r \le 2d - 1$ , we observe the duality relation

$$p_n^{r,d}(\alpha) = 1 - p_n^{r,r-d+1}(1-\alpha), \tag{20}$$

which follows from

$$p_n^{r,d}(\alpha) = P\{S_{n+r}(\alpha) \ge S_n'(\alpha) + d\} = P\{n + r - S_{n+r}(\alpha) \le n - S_n'(\alpha) + r - d\}$$

by comparing tails. Here we temporarily changed the notation to stress that  $p_n^{r,d}$  is a function of  $\alpha$ .

Now case (iii)(b) follows immediately, while (20) implies that case (ii) is equivalent to case (iv), which was proved above. Indeed, if  $d \le r \le 2d - 2$  then  $r \ge 2d'$ , where d' := r - d + 1, and if  $r \ge 2d$  then  $d' \le r \le 2d' - 2$ . Thus, the function  $(r, d) \mapsto (r, r - d + 1)$  is a bijection between the sets  $\{(r, d) : d \le r \le 2d - 2\}$  and  $\{(r, d) : r \ge 2d\}$  while 1 - d/(r + 1) = d'/(r + 1) and 1 - (2d - 1)/(2r) = (2d' - 1)/(2r).

Proof of Theorem 3. It is well known that the maximum of  $U_k(x)$  on [0, 1] occurs at x = 1. Hence, (11), (14) and (18) imply that  $P_{\alpha}^{r,1}(x) \leq P_{\alpha}^{r,1}(1) = 2\alpha r - 1$  for  $x \in [0, 1]$ , and, by Lemma 2, the sequence  $p_n^{r,1}$  is increasing when  $0 < \alpha \leq 1/(2r)$ . The decrease of  $p_n^{r,1}$  for  $1/(r+1) < \alpha < 1$  is already covered by Theorem 4(iv)(c).

For a proof of unimodality in the transitional zone, it suffices to check that  $P_{\alpha}^{r,1}(x)$  has only one root on [0,1] when  $1/(2r) < \alpha \le 1/(r+1)$ . Indeed, let  $P_{\alpha}^{r,1}(x_0) = 0$  for some  $x_0 \in (0,1)$ , and let  $p_{k+1}^{r,1} - p_k^{r,1} \le 0$  for some  $k \ge 0$ . We claim that, for any n > k, it holds that

$$Q_{\alpha}^{n}(x)\sqrt{1-x^{2}}P_{\alpha}^{r,d}(x) > Q_{\alpha}^{n-k}(x_{0})Q_{\alpha}^{k}(x)\sqrt{1-x^{2}}P_{\alpha}^{r,d}(x), \quad x \in [0,1] \setminus \{x_{0}\}.$$
 (21)

First observe from (6) that  $P_{\alpha}^{r,1}(0)=-(1-2\alpha)^r<0$  while  $P_{\alpha}^{r,1}(1)=2\alpha r-1>0$ , so  $P_{\alpha}^{r,1}(x)<0$  on  $(0,x_0)$  and  $P_{\alpha}^{r,1}(x)>0$  on  $(x_0,1)$ . Then we obtain (21) using the fact that  $Q_{\alpha}(x)$  is positive and increasing in x on [0,1]. Now integrate (21) over [0,1] and apply Lemma 2 to obtain  $p_{n+1}^{r,1}-p_n^{r,1}< Q_{\alpha}^{n-k}(x_0)(p_{k+1}^{r,1}-p_k^{r,1})\leq 0$  for all n>k. Thus, unimodality follows from the uniqueness of the root. Note that, by Lemma 1,

$$p_1^{r,1} - p_0^{r,1} = \alpha(1-\alpha)(P\{S_r = 0\} - P\{S_r = 1\}) \ge 0,$$

ensuring that the mode satisfies  $N_{\alpha,r} \geq 1$ .

Clearly, we prove that  $P_{\alpha}^{r,1}(x)$  has only one root on [0,1] when  $1/(2r) < \alpha < 1/(r+1)$  if we check that  $P_{\alpha}^{r,1}(x)$  is increasing on [0,1] when  $\alpha \le 1/(r+1)$ . We claim the stronger statement: with exception of the constant term, all coefficients of the polynomial  $P_{\alpha}^{r,1}(x)$  are positive when  $\alpha \le 1/(r+1)$ .

By Theorem 1, we show that

$$\frac{\mathrm{d}^{2j}}{\mathrm{d}t^{2j}}(t^{j-1}(1-\alpha+\alpha t)^r)\bigg|_{t=-1}$$

is positive for any  $1 \le j \le r - 1$ . The Leibnitz formula gives

$$\frac{\mathrm{d}^{2j}}{\mathrm{d}t^{2j}}(t^{j-1}(1-\alpha+\alpha t)^r)\bigg|_{t=-1} = \sum_{k=j+1}^{2j \wedge r} (-1)^{k-j-1} a_k,$$

where

$$a_k := \binom{2j}{k} \frac{r! (j-1)!}{(r-k)! (k-j-1)!} \alpha^k (1-2\alpha)^{r-k},$$

and it suffices to prove that the  $a_k$  are decreasing. We have

$$\frac{a_{k+1}}{a_k} = \frac{\alpha}{1 - 2\alpha} \frac{(r - k)(2j - k)}{(k - j)(k + 1)},$$

which is obviously decreasing in k, so

$$\frac{a_{k+1}}{a_k} \le \frac{a_{j+2}}{a_{j+1}} = \frac{\alpha}{1 - 2\alpha} \frac{(r - j - 1)(j - 1)}{j + 2} < \frac{\alpha}{1 - 2\alpha} (r - 2) \le \frac{r - 2}{r - 1} < 1$$

when  $\alpha \leq 1/(r+1)$ . This completes the proof of the unimodality of  $p_n^{r,1}$ .

To prove (8), we use an argument similar to that used to prove (7). By Lemma 2 and the change of variable  $x = \cos(t/2)$ , we have

$$\frac{\pi(p_{n+1}^{r,1} - p_n^{r,1})}{2\alpha(1 - \alpha)} = -\int_0^{\pi} \varphi_{\alpha}^n(t) \sin^2 \frac{t}{2} \cdot P_{\alpha}^{r,d} \left(\cos \frac{t}{2}\right) dt.$$

Representing  $P_{\alpha}^{r,1}(x)$  via its Taylor polynomial at x=1, we have

$$\begin{split} &\frac{\pi(p_{n+1}^{r,1} - p_n^{r,1})}{2\alpha(1 - \alpha)} \\ &= -\int_0^\pi \varphi_\alpha^n(t) \sin^2 \frac{t}{2} P_\alpha^{r,1}(1) \, \mathrm{d}t \\ &+ \int_0^\pi \varphi_\alpha^n(t) \sin^2 \frac{t}{2} \left[ \frac{\partial}{\partial x} P_\alpha^{r,1}(1) \left( 1 - \cos \frac{t}{2} \right) - \frac{1}{2} \frac{\partial^2}{\partial x^2} P_\alpha^{r,1}(\theta(t)) \left( 1 - \cos \frac{t}{2} \right)^2 \right] \mathrm{d}t \\ &=: -I_1(\alpha, n) + I_2(\alpha, n), \end{split}$$

where  $\cos(t/2) < \theta(t) < 1$ .

Now assume that  $n = n(\alpha) \to \infty$  as  $\alpha \downarrow 1/(2r)$ . We claim that

$$I_1(\alpha, n(\alpha)) \sim \frac{\sqrt{\pi}r^3}{2(2r-1)^{3/2}} \frac{2\alpha r - 1}{n(\alpha)^{3/2}}, \quad I_2(\alpha, n(\alpha)) \sim \frac{\sqrt{\pi}(r-1)r^3}{4(2r-1)^{3/2}} \frac{1}{n(\alpha)^{5/2}}, \quad \text{as } \alpha \downarrow \frac{1}{2r}.$$

Indeed, arguing as in (15) and (17) and using (16), which of course holds when  $\alpha \to 1/(2r)$ , we obtain

$$I_1(\alpha, n(\alpha)) \sim \frac{P_{\alpha}^{r,1}(1)}{4n^{3/2}} \int_0^{\delta n^{1/2}} \varphi_{\alpha}^n \left(\frac{s}{\sqrt{n}}\right) s^2 ds \sim \frac{P_{\alpha}^{r,1}(1)}{4n^{3/2}} \int_0^{\infty} \exp\left(-\frac{2r-1}{4r^2}s^2\right) s^2 ds$$

by the dominated convergence theorem. It only remains to compute the integral and use the equation  $P_{\alpha}^{r,1}(1) = 2\alpha r - 1$  to obtain the first part of (22). Similarly,

$$I_2(\alpha, n(\alpha)) \sim \frac{\partial P_{\alpha}^{r,1}(1)/\partial x}{32n^{5/2}} \int_0^{\delta n^{1/2}} \varphi_{\alpha}^n \left(\frac{s}{\sqrt{n}}\right) s^4 ds$$
  
  $\sim \frac{\partial P_{1/(2r)}^{r,1}(1)/\partial x}{32n^{5/2}} \int_0^{\infty} \exp\left(-\frac{2r-1}{4r^2}s^2\right) s^4 ds,$ 

as the term coming from the second derivative of  $P_{\alpha}^{r,1}(x)$  gives a lower order contribution, and we obtain the second relation in (22) by computing the integral and using (19).

Clearly,  $N_{\alpha,r} \to \infty$  as  $\alpha \downarrow 1/(2r)$ , and (22) implies that  $I_k(\alpha, N_{\alpha,r} - 1) \sim I_k(\alpha, N_{\alpha,r})$ , k = 1, 2. By the definition of  $N_{\alpha,r}$ , we have  $I_2(\alpha, N_{\alpha,r} - 1) \ge I_1(\alpha, N_{\alpha,r} - 1)$  and  $I_1(\alpha, N_{\alpha,r}) > I_2(\alpha, N_{\alpha,r})$ ; hence,

$$I_1(\alpha, N_{\alpha,r}) \sim I_2(\alpha, N_{\alpha,r})$$
 as  $\alpha \downarrow \frac{1}{2r}$ .

We now apply (22) again to obtain (8).

For (9), we argue in the same manner as above. Omitting the details, we obtain

$$2\pi \left(p_{N_{\alpha,r}} - \frac{1}{2}\right) \sim \frac{P_{\alpha}^{r,1}(1)}{N_{\alpha,r}^{1/2}} \int_{0}^{\infty} \exp\left(-\frac{2r-1}{4r^2}s^2\right) ds$$
$$-\frac{\partial P_{1/(2r)}^{r,1}(1)/\partial x}{8N_{\alpha,r}^{3/2}} \int_{0}^{\infty} \exp\left(-\frac{2r-1}{4r^2}s^2\right) s^2 ds,$$

and the required relation follows.

### 4. Additional remarks and related questions

Here we mention some issues that deserve more attention than we can provide. Then we explain our initial interest in the problem.

Simulations show that, as in the case  $p_n^{r,1}$  with  $r \ge 2$ , the sequence  $p_n^{r,d}$  with  $r \ge 2d$  is increasing when  $0 < \alpha < (2d-1)/(2r)$  and unimodal in the transitional phase  $(2d-1)/(2r) < \alpha \le d/(r+1)$ . The proofs presented here do not admit a reasonable generalization even for d=2. For example, simulations show that the coefficient of  $P_\alpha^{r,2}(x)$  at x is negative and all the other coefficients are positive. This could neither be easily proved nor used for a proof similar to the one given here for d=1 as  $P_\alpha^{r,2}(x)$  has two roots on [0,1] for some  $\alpha$ . Once unimodality in the transitional phase is established for general d, the asymptotic of the mode  $N_\alpha^{r,d}$  can be found by exactly the same argument as used for (8).

Another open problem is to verify that the sequence  $p_n^{r,d}$  is convex/concave in the monotone modes, as observed by simulations. By the same argument as in the proof of Theorem 1, we easily get convexity (concavity, to be precise) of  $p_n^{r,1}$  when  $0 < \alpha \le 1/(2r)$ . Unfortunately, the same method does not seem to work for  $1/(r+1) < \alpha < 1$ .

Next we should mention that, after a preliminary draft of this paper was finished, we found that a related problem was studied by Addona *et al.* [1]. Their analysis used the multivariate form of Zeilberger's algorithm to find representations and estimates based on a linear combination of two Legendre polynomials. Also, after submitting this paper, Tamas Lengyel and the referee pointed out the paper [6] on approximating point spread distributions. Our Fourier analytic method can also be used to simplify some of the arguments in both [1] and [6], and provide integral representations for general r, d, p, and q; see [2] for details.

Finally, we give our motivations for studying the problem considered in this paper. We say that a random variable Z has a double-sided exponential distribution if it has a density  $\rho(x)$  of the form  $\rho(x) = \alpha a \mathrm{e}^{-ax}$  for x > 0 and  $\rho(x) = (1 - \alpha) b \mathrm{e}^{bx}$  for x < 0, where a, b > 0 and  $\alpha \in (0, 1)$ . Such a specific random variable has very nice properties from the point of view of fluctuation theory; see [4, Chapter VI.8]. Here E Z = 0 means  $b = (1 - \alpha)a/\alpha$ , so any centered double-sided exponential variable could be written in the form

$$Z = \frac{1}{a} \left( XY - \frac{\alpha}{1 - \alpha} (1 - X)Y \right),$$

where X is a Bernoulli random variable that equals 1 and 0 with probability  $\alpha$  and  $1 - \alpha$ , respectively, Y is a standard exponential random variable with density  $e^{-x}$  for x > 0, and X and Y are independent.

Consider independent, identically distributed double-sided exponential random variables  $Z_n$ . Initially, our goal was to show that the sequence  $P\{Z_1 + \cdots + Z_n > 0\} - \frac{1}{2}$  does not change its sign as n increases. This property was verified by simulations, and a rigorous proof was required.

One of the possible approaches is to show that the sequence  $P\{Z_1 + \cdots + Z_n > 0\}$  is monotone since, by the central limit theorem, its limit equals  $\frac{1}{2}$ .

Introducing independent sequences of standard exponential random variables  $Y_n$  and Bernoulli random variables  $X_n$ , we write

$$P\{Z_{1} + \dots + Z_{n} > 0\}$$

$$= \sum_{k=1}^{n} {n \choose k} \alpha^{k} (1 - \alpha)^{n-k} P\{Y_{1} + \dots + Y_{k} > \frac{\alpha}{1 - \alpha} (Y_{k+1} + \dots + Y_{n})\}$$

$$= \sum_{k=1}^{n} {n \choose k} \alpha^{k} (1 - \alpha)^{n-k} P\{Y_{1} + \dots + Y_{k} > \alpha (Y_{1} + \dots + Y_{n})\}$$

$$= \sum_{k=1}^{n-1} {n \choose k} \alpha^{k} (1 - \alpha)^{n-k} P\{U_{k,n-1} > \alpha\} + \alpha^{n},$$

where we have used the well-known (see [5]) representation of uniform order statistics in terms of exponential random variables. Recall the definition. Let  $U_1, \ldots, U_m$  be independent random variables that are uniformly distributed on [0, 1], and arrange them in ascending order. The kth element of this new sequence is called the kth order statistics of  $U_1, \ldots, U_m$ , denoted by  $U_{k,m}$ . By definition,  $U_{1,m} \leq \cdots \leq U_{m,m}$ , but all the inequalities are strict with probability 1.

Note that  $P{U_{k,m} > \alpha} = P{S_m < k}$ ; hence,

$$P\{Z_1 + \dots + Z_n > 0\} = \sum_{k=1}^{n-1} \binom{n}{k} \alpha^k (1 - \alpha)^{n-k} P\{S_{n-1} < k\} + \alpha^n = P\{S'_n > S_{n-1}\} = p_{n-1}.$$

We see that the monotonicity of  $P\{Z_1 + \cdots + Z_n > 0\}$  is exactly the monotonicity of  $p_n$  given by Theorem 2.

## Acknowledgements

The authors would like to thank Tamas Lengyel and the anonymous referee for pointing out the paper [6], and for their insightful remarks and suggestions.

#### References

- [1] ADDONA, V., WAGON, S. AND WILF, H. (2011). How to lose as little as possible. Ars Math. Contemp. 4, 29-62.
- [2] DONG, Z., LI, W. V. AND SONG, C. (2011). Integral representations for binomial sums of chances of winning. In preparation.
- [3] GHAHRAMANI, S. (2005). Fundamentals of Probability, 3rd edn. Prentice Hall.
- [4] FELLER, W. (1966). An Introduction to Probability Theory and Its Applications, Vol. II. John Wiley, New York.
- [5] KARLIN, S. (1968). A First Course in Stochastic Processes. Academic Press, New York.
- [6] LENGYEL, T. (2011). On approximating point spread distributions. J. Statist. Comput. Simul. 81, 1333–1344.
- [7] Ross, S. (2010). A First Course in Probability, 8th edn. Prentice Hall.
- [8] ZWILLINGER, D. (ed.) (2002). CRC Standard Mathematical Tables and Formulae, 31st edn. Chapman and Hall/CRC.