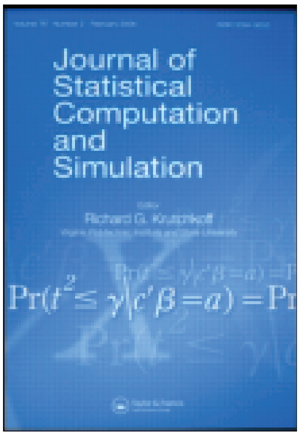

Journal of Statistical Computation and Simulation
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gscs20

On approximating point spread distributions

Tamás Lengyel [a]

[a] Department of Mathematics, Occidental College, 1600 Campus Road, Los Angeles, CA, 90041, USA

Version of record first published: 24 Feb 2011.

To cite this article: Tamás Lengyel (2011): On approximating point spread distributions, Journal of Statistical Computation and Simulation, 81:10, 1333-1344

To link to this article: http://dx.doi.org/10.1080/00949655.2010.484901

Journal of Statistical Computation and Simulation
Vol. 81, No. 10, October 2011, 1333–1344

On approximating point spread distributions

Tamás Lengyel*

Department of Mathematics, Occidental College, 1600 Campus Road, Los Angeles, CA 90041, USA

(*Received 19 January 2010; final version received 8 April 2010*)

We discuss some properties of the point spread distribution, defined as the distribution of the difference of two independent binomial random variables with the same parameter n including exact and approximate probabilities and related optimization issues. We use various approximation techniques for different distributions, special functions, and analytic, combinatorial and symbolic methods, such as multi-summation techniques. We prove that in case of unequal success rates, if these rates change with their difference kept fixed and small, and n is appropriately bounded, then the point spread distribution only slightly changes for small point differences. We also prove that for equal success rates p, the probability of a tie is minimized if $p = 1/2$. Numerical examples are included for the case with $n = 12$.

Keywords: Skellam distribution; approximating distributions; asymptotic enumeration; special functions; multi-summation

AMS Subject Classification: 62E17; 05A16; 33F10

1. Introduction

Questions regarding the point spread distribution in certain sports present a rich variety of interesting problems. In high-scoring sports, e.g. basketball, the underlying scoring distributions can be modelled by independent binomial random variables, while in low-scoring sports, e.g. baseball, hockey, and soccer, independent Poisson random variables might be used. Point spreads are often used to set fairly equal winning odds in terms of scoring differences in matches between players or teams of widely different strengths. In fact, bookmakers set a point spread to even the game for betting purposes [1]. On the other hand, there are some infamous cases involving point-shaving to beat the projected point spread through the bribing of players from the favoured team by gamblers.

In the simplified model, let X and Y represent the number of points scored by two teams or players in n 'games', respectively. We assume that X and Y are independent, binomially distributed random variables and introduce the probability $f_d(n, p, \varepsilon)$ of a point spread $X - Y$ which is at least as large as $d \geq 0$ points in n games with respective success rates p and $p - \varepsilon$. Section 2 is devoted to the discussion of the symmetry of the point spread distribution $f_d(n, p, \varepsilon)$ (cf. Theorem 2.1). We study the probability $f_0(n, p, \varepsilon) - f_1(n, p, \varepsilon)$ of a tie and that of an exact point spread d, $f_d(n, p, \varepsilon) - f_{d+1}(n, p, \varepsilon)$, in Section 3, and briefly describe a random-walk-based approach for

*Email: lengyel@oxy.edu

ISSN 0094-9655 print/ISSN 1563-5163 online

http://dx.doi.org/10.1080/00949655.2010.484901
http://www.tandfonline.com

the calculation of the distribution in Section 4. The point spread distribution is approximated in Sections 5 and 6 by a normal distribution in Theorem 5.1 and by a Skellam distribution if the success rates are close to one. We extend some of the calculations to negative values of d in Section 5. In Section 7, we ask some optimization questions such as: for a fixed $\varepsilon \geq 0$, what value of p will maximize $f_d(n, p, \varepsilon)$, a question settled in Theorem 7.1 for the case with $f_0(n, p, 0)$. The last section is concerned with the distribution of the absolute point spread. The main results are summarized in Theorems 2.1, 5.1, and 7.1.

For purely illustrative purposes, we consider an example which touches upon some aspects of the point spread distribution.

Example Players 1 and 2 are two basketball players. Player 1 makes 65% of his free-throws, while Player 2 is even better and makes 75%. They have a contest in which they each shoot 12 free-throws. You can assume that the scoring luck is totally independent of the previous or current scoring experience of either player. We are interested in the following probabilities.

(A) What is the probability that this free-throw contest will end in a tie?
(B) What is the probability that Player 2 will win?
(C) What are the chances that, say, Player 2 scores two more than Player 1?
(D) How about at least two more?

The answers can be calculated easily with most statistical software, and perhaps in the most economical way and conveniently in S-PLUS, at least in terms of the frugality of the necessary code. In fact, the answers are one-liners in S-PLUS: for (A) we get 0.1533 by `sum(dbinom(0:12,12,0.65)* dbinom(0:12,12,0.75))`, for (B) 0.6249 is obtained by `sum(dbinom(1:12,12,0.75)* pbinom(0:11,12,0.65))`, for (C) `sum(dbinom(2:12,12,0.75)* dbinom(0:10,12,0.65))` results in 0.1676, and for (D) `sum(dbinom(2:12,12,0.75)*pbinom(0:10,12,0.65))` yields 0.4482, with four significant digits. Other values related to the point spread distribution can be derived in a similar fashion.

We focus on questions (B) and (D), in particular. After calculating these values, we tested similar settings within a given range but with the same difference in shooting success rates and found, somewhat surprisingly, that the answers obtained changed only slightly in (B) and (D). In other words, if there are some external circumstances that are equally influencing both players to play better or worse, then for small differences, the point spread distribution will change only to a very small degree. We apply different approximation and summation techniques, a random walk approach, and discuss related optimization problems to help explain this observation. Calculations for $n = 12$ are included to illustrate numerical aspects and features of the different approaches.

2. Symmetry

Now we prove an interesting symmetry property.

Theorem 2.1 *Let $f_d(n, p, \varepsilon)$ denote the probability that, in n trials, the 'stronger' player with success rate p accumulates at least $d \geq 0$ points more than the other (equally strong or 'weaker') player with success rate $p - \varepsilon$ with $\varepsilon \geq 0$. The function $f_d(n, p, \varepsilon)$, $\varepsilon \leq p \leq 1$, is symmetric about $(1 + \varepsilon)/2$ for every $n \geq 1$.*

Proof We observe that

$$f_d(n, p, \varepsilon) = \sum_{k=d}^{n} \binom{n}{k} p^k (1-p)^{n-k} \sum_{j=0}^{k-d} \binom{n}{j} (p-\varepsilon)^j (1-p+\varepsilon)^{n-j}. \quad (1)$$

Let us assume that $\varepsilon \le p \le 1$. The claim is that $f_d(n, p, \varepsilon) = f_d(n, 1-p+\varepsilon, \varepsilon)$. By Equation (1), we have that

$$f_d(n, 1-p+\varepsilon, \varepsilon) = \sum_{k=d}^{n} \binom{n}{k} (1-p+\varepsilon)^k (p-\varepsilon)^{n-k} \sum_{j=0}^{k-d} \binom{n}{j} (1-p)^j p^{n-j}.$$

We can view the second summation as the probability that the player with success rate p loses at most $k-d$ games, while $\binom{n}{k}(1-p+\varepsilon)^k(p-\varepsilon)^{n-k}$ is the probability that the player with success rate $p-\varepsilon$ loses exactly k games. Thus, the combined expression gives the probability of winning by at least d games by the former player. ■

Note that this property guarantees that it suffices to deal with $f_d(n, p, \varepsilon)$ with $p \ge (1+\varepsilon)/2$. Sometimes, we use the short notation f_d instead of $f_d(n, p, \varepsilon)$.

Remark 1 The definition of f_d can be extended to every real p by Equation (1), and obviously, the extended function remains symmetric. This can be proved by taking the Taylor expansion of the polynomial f_d about $(1+\varepsilon)/2$ which has only terms with even exponents. Similarly, the statement holds true for any entire function (whose Taylor series converges to the function itself everywhere).

Remark 2 Note that the degree of polynomial f_d in p is $2n$. Moreover, Equation (1) shows that $p^d(1-p+\varepsilon)^d$ is always a factor of f_d independent of n. By focusing on the highest powers of p (including sign) in Equation (1), it can be proved that the leading coefficient of $f_d(n, p, \varepsilon)$, as a polynomial in p, is $(-1)^d c_{n,d}$ with

$$c_{n,d} = \binom{2n-1}{n-d}, \quad n \ge d,$$

independent of ε. It follows, for instance, that $c_{n,0} = c_{n,1} = \binom{2n}{n}/2$.

For $d = 0$, by more detailed calculations, we can get that

$$f_0(n, p, \varepsilon) = (1+\varepsilon)^n - n(1+\varepsilon)^{n-1}(1+n\varepsilon)p + \frac{n}{4}(1+\varepsilon)^{n-2}(\varepsilon^2 n^3 + 6\varepsilon n^2 - \varepsilon^2 n - 2\varepsilon n + 6n - 2)p^2 + \cdots + \binom{2n-1}{n} p^{2n}.$$

However, as we will see, it is more important to have the expansion of f_d about $p = (1+\varepsilon)/2$ rather than about $p = 0$.

3. The probability of a tie and exact point spread

If $p = 1$ then the probability of a tie is $(1-\varepsilon)^n$. In general, we find that the probability of a tie is

$$T = f_0(n, p, \varepsilon) - f_1(n, p, \varepsilon) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \binom{n}{k} (p-\varepsilon)^k (1-p+\varepsilon)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k}^2 (p(p-\varepsilon))^k ((1-p)(1-p+\varepsilon))^{n-k}$$

$$= ((1-p)(1-p+\varepsilon))^n \sum_{k=0}^{n} \binom{n}{k}^2 \left(\frac{p(p-\varepsilon)}{(1-p)(1-p+\varepsilon)}\right)^k,$$

and thus,

$$T = \begin{cases} ((1-p)(1-p+\varepsilon))^n \left(1 - \frac{p(p-\varepsilon)}{(1-p)(1-p+\varepsilon)}\right)^n & \\ \quad \times P_n\left(\frac{1+p(p-\varepsilon)/(1-p)(1-p+\varepsilon)}{1-p(p-\varepsilon)/(1-p)(1-p+\varepsilon)}\right), & \text{if } \frac{1+\varepsilon}{2} < p < 1, \\ \left(\frac{1-\varepsilon^2}{4}\right)^n \binom{2n}{n}, & \text{if } p = \frac{1+\varepsilon}{2}, \end{cases} \tag{2}$$

with the nth Legendre polynomial $P_n(x)$ (cf. [2]). If $x < -1$ then $P_n(x)$ can be approximated by

$$P_n(x) \sim \frac{(-1)^n}{\sqrt{2\pi n}(x^2-1)^{1/4}} (-x + \sqrt{x^2-1})^{n+1/2} \tag{3}$$

as $n \to \infty$. For instance, the answer 0.1533 to question (A) can be approximated by 0.1542 this way since $p < 1$ guarantees that the argument of P_n in (3) is less than -1. With the notation

$$\Delta = p - \frac{1+\varepsilon}{2}, \tag{4}$$

we get the approximation of the answer T

$$T \approx \left(\Delta - \frac{1}{2}\right)^{2n} \left(1 - \frac{(\Delta+1/2)^2}{(\Delta-1/2)^2}\right)^n P_n\left(-\frac{1}{4\Delta} - \Delta\right)$$

for a small ε if $\Delta > 0$. Of course, we can also use the approximation (3) for the last factor. For example, with $n = 12$, $p = 0.75$, and $\varepsilon = 0.01$, we get that the exact value is 0.1859 while the above approximation with Equation (3) results in 0.1857. Also note that by Equation (2), we have that

$$f_0\left(n, \frac{1+\varepsilon}{2}, \varepsilon\right) - f_1\left(n, \frac{1+\varepsilon}{2}, \varepsilon\right) \sim \frac{(1-\varepsilon^2)^n}{\sqrt{n\pi}}$$

for $\Delta = 0$ and $n \to \infty$.

We can generalize the above approach for any difference $d \geq 0$. By using the hypergeometric function ${}_2F_1$ (cf. [2]), we obtain the following theorem.

THEOREM 3.1 *The probability $f_d(n, p, \varepsilon) - f_{d+1}(n, p, \varepsilon)$ of an exact point spread of d is*

$$((1-p)\,(1-p+\varepsilon))^n \left(\frac{p}{1-p}\right)^d \binom{n}{d} {}_2F_1\left(d-n, -n, 1+d; \frac{p\,(p-\varepsilon)}{(1-p)\,(1-p+\varepsilon)}\right)$$

if $(1+\varepsilon)/2 \leq p < 1$.

For instance, Theorem 3.1 gives the answer $f_0(12, 0.75, 0.10) - f_1(12, 0.75, 0.10) = 0.1533$ to question (A). Similarly, $f_2(12, 0.75, 0.10) - f_3(12, 0.75, 0.10) = 0.1676$ answers question (C).

4. Random walk approach

Assume that in the shooting competition the players shoot in an alternating fashion. Let X_i be 1 or 0 if the stronger player scores or misses his free throw in the ith trial. We define Y_i similarly. Then $X = \sum_{k=1}^{n} X_i$ and $Y = \sum_{k=1}^{n} Y_i$. We can model the problem with a random walk which has step sizes $X_i - Y_i, i = 1, 2, \ldots, n$: the walk moves one step to the right or left if only the stronger or the other player scores, respectively, and stays in place if either both or none of the players score. With $P = p(1-q)$, $Q = q(1-p)$, and $q = p - \varepsilon$, we have that

$$P(X - Y = k) = [z^k]\left(Pz + (1-P-Q) + \frac{Q}{z}\right)^n = [z^{n+k}](Pz^2 + (1-P-Q)z + Q)^n$$

where $[z^k]g(z)$ stands for the coefficient of the term z^k in the (Laurent) power series expansion of $g(z)$ about 0. By adding the coefficients for $k = d, d+1, \ldots, n$, we can obtain $f_d(n, p, \varepsilon)$. Formally, in terms of a Laurent series with an essential singularity at 0, we can write that

$$f_d(n, p, \varepsilon) = P^n\,[z^{n+d}]\left(1 + \frac{1}{z} + \frac{1}{z^2} + \cdots\right)\left(z^2 + \frac{1-P-Q}{P}z + \frac{Q}{P}\right)^n.$$

We note, however, that this approach results in a double sum for f_d and thus, offers little or no improvement over Equation (1) unless some computer algebra system is easily available for extracting coefficients. In some cases though, it might be helpful. For example, coefficient extraction is easy when $p = q = 1/2$ and thus, $\varepsilon = 0$, $P = Q = 1/4$, and identity (15) follows immediately. In another example, we can get a good approximation for the probability of a tie $T = P(X - Y = 0)$ (cf. Equation (2)) by using standard facts regarding random walks. In fact, [3, Proposition VII. 10] suggests

$$T \sim \frac{\left(1 - (\sqrt{P} - \sqrt{Q})^2\right)^{n+1/2}}{2(PQ)^{1/4}\sqrt{\pi n}}$$

as $n \to \infty$ by using the asymptotic number of 'bridges' (and its generalization for finding the probability that a random walk of length n is a bridge from altitude 0 to 0), hence providing an approximation for T which looks simpler than the one given in Section 3. For example, in this way we get the approximate answer 0.1542 to (A). Note the difference in the rate of decrease of the approximated T as n grows: it is exponential if $p \neq q$ and of order $1/\sqrt{n}$ if $p = q$. In fact, T decreases as $n \to \infty$, as can be shown in the way Lemma 5 of [4] is proved according to Wagon [5]. The author conjectures that for $n \geq 5$, $p = (1+\varepsilon)/2$ switches from being the location of the minimum of T to that of the maximum as ε grows. Of course, the symmetry of T, i.e. $f_0(n, p, \varepsilon) - f_1(n, p, \varepsilon)$, in p about $(1+\varepsilon)/2$ follows by Theorem 2.1, and this yields that this particular location gives either a (local) minimum or maximum.

5. Approximation by normal

With some reasonable bounds on the number of trials n and success rates p and $q = p - \varepsilon$ with $\varepsilon > 0$ for Players 2 and 1, respectively, we first approximate $f_d(n, p, \varepsilon)$ for $d \geq 1$ and then extend the range of the approximation for $d \leq 0$. As we observed above, we can restrict our attention to the range $p \geq (1+\varepsilon)/2$.

Let X and Y represent the number of successful shots made by Players 2 and 1, respectively. Clearly, X and Y are independent, binomially distributed random variables, $X \sim \mathcal{N}[\mu =$

$np, \sigma = \sqrt{np(1-p)}]$ and $Y \sim \mathcal{N}[\mu = nq, \sigma = \sqrt{nq(1-q)}]$ approximately, even for small values of n, as long as $9p/(1-p) \le n$, i.e. $p \le n/(n+9)$, although this condition can be relaxed in practice. Therefore, the distribution of $X - Y$ is approximately $\mathcal{N}[\mu = n(p-q) = n\varepsilon, \sigma = \sqrt{n(p(1-p)+q(1-q))}]$.

To answer (A)–(D) we need the probabilities $P(X-Y=0)$, $P(X-Y \ge 1)$, $P(X-Y=2)$, and $P(X-Y \ge 2)$, respectively. For example, $f_1(n,p,\varepsilon) = P(X-Y \ge 1)$ can be approximated by

$$1 - \Phi\left(\frac{0.5 - n\varepsilon}{\sqrt{n(p(1-p)+q(1-q))}}\right) \tag{5}$$

using the so-called continuity correction. We can use a simple and fairly accurate approximation for the normal distribution function

$$\Phi(x) \approx 0.5 + 0.1x(4.4 - x), \quad \text{for } 0 \le x \le 2.2.$$

In fact, it is good to two decimal places.

We consider $\sqrt{n(p(1-p)+q(1-q))}$ from Equation (5). We first take $p(1-p)+q(1-q)$ and rewrite it in terms of the fixed $\varepsilon = p - q$ and then express it as a function of $\Delta = p - (1+\varepsilon)/2$. We get that

$$\sqrt{n(p(1-p)+q(1-q))} = \sqrt{\frac{n}{2}\left(1 - 4\left(p - \frac{1+\varepsilon}{2}\right)^2 - \varepsilon^2\right)}$$
$$\approx \sqrt{\frac{n}{2}}\left(1 - 2\left(p - \frac{1+\varepsilon}{2}\right)^2 - \frac{1}{2}\varepsilon^2\right)$$

if ε and Δ are small, more precisely, if $\sqrt{n}(\Delta^4 + \varepsilon^4)$ is small.

We can use a quadratic or finer approximation of the argument of Φ about $p = (1+\varepsilon)/2$ in Equation (5). We proceed with the quadratic approximation. Let us assume that $2\varepsilon\Delta^4\sqrt{2n}$ is small. This assumption will guarantee that the above approximation introduces only negligible errors when we replace x by x_1, x_2, and x_d below. We set

$$x_1 = \left(0.5\sqrt{\frac{2}{n}} - \varepsilon\sqrt{2n}\right)\left(1 + 2\left(p - \frac{1+\varepsilon}{2}\right)^2 + \frac{\varepsilon^2}{2}\right) \le 0,$$

for $0.5/\varepsilon \le n \le 2/\varepsilon^2$ and use

$$1 - \Phi(x) \approx 0.5 + 0.1(-x)(4.4 + x) \tag{6}$$

with $x = x_1$. This provides us with a quadratic approximation in x_1. Note that to answer (D), we need a slight modification of x_1. We use approximation (6) with setting x to

$$x_2 = \left(1.5\sqrt{\frac{2}{n}} - \varepsilon\sqrt{2n}\right)\left(1 + 2\left(p - \frac{1+\varepsilon}{2}\right)^2 + \frac{\varepsilon^2}{2}\right) \le 0$$

for $1.5/\varepsilon \le n \le 2/\varepsilon^2$. In general, for a difference of d points we set

$$x_d = \left((d - 0.5)\sqrt{\frac{2}{n}} - \varepsilon\sqrt{2n}\right)\left(1 + 2\left(p - \frac{1+\varepsilon}{2}\right)^2 + \frac{\varepsilon^2}{2}\right) \le 0 \tag{7}$$

for $(d - 0.5)/\varepsilon \le n \le 2/\varepsilon^2$.

We use the above quadratic approximation (6) to $f_d(n, p, \varepsilon)$ about $(1+\varepsilon)/2$. For example, for $d = 1$ we get that

$$\text{the constant coeff:}\quad \left(0.5 - \frac{0.311127}{\sqrt{n}} - \frac{0.05}{n}\right) + \varepsilon\left(0.622254\sqrt{n} + 0.2\right) - \varepsilon^2\left(0.2n + \frac{0.155563}{\sqrt{n}} + \frac{0.05}{n}\right) + O\left(\varepsilon^3\right), \tag{8}$$

$$\text{the coeff of } \left(p - \frac{1+\varepsilon}{2}\right)^2 :\quad -\left(\frac{0.622254}{\sqrt{n}} + \frac{0.2}{n}\right) + \varepsilon\left(1.24451\sqrt{n} + 0.8\right) - \varepsilon^2\left(0.8n + \frac{0.1}{n}\right) + O\left(\varepsilon^3\right). \tag{9}$$

Note that if $d \le 0$ then $x_d < 0$ in Equation (7), and we need only the condition that $n \le 2/\varepsilon^2$. In general, we have the following theorem.

THEOREM 5.1 *For an arbitrary integer d, we get the approximation* $f_d(n, p, \varepsilon) \approx c_0 + c_2(p - (1+\varepsilon)/2)^2$ *with*

$$c_0 = \left(0.5 - \frac{0.311127(2d-1)}{\sqrt{n}} - \frac{0.05(4d(d-1)+1)}{n}\right) + \varepsilon(0.622254\sqrt{n} + 0.2(2d-1)) - \varepsilon^2\left(0.2n + \frac{0.155563(2d-1)}{\sqrt{n}} + \frac{0.05(4d(d-1)+1)}{n}\right) + O(\varepsilon^3), \tag{10}$$

$$c_2 = -\left(\frac{0.622254(2d-1)}{\sqrt{n}} + \frac{0.2(4d(d-1)+1)}{n}\right) + \varepsilon(1.24451\sqrt{n} + 0.8(2d-1)) - \varepsilon^2\left(0.8n + \frac{(2d-1)^2}{10n}\right) + O(\varepsilon^3), \tag{11}$$

provided that $(d - 0.5)/\varepsilon \le n \le 2/\varepsilon^2$ *and* $2\varepsilon\Delta^4\sqrt{2n}$ *is small with* $\Delta = p - (1+\varepsilon)/2$.

Of course, if p is close to $(1+\varepsilon)/2$ and ε is small then the constant term $0.5 - 0.311127(2d-1)/\sqrt{n} - 0.05(4d(d-1)+1)/n$ suffices in order to get a good approximation of f_d.

Note that using a higher degree approximation in Equation (5) will not change the shape of the approximating function (7) since the Maclaurin series of $1/\sqrt{a-y}$, $a > 0$, in y has only positive coefficients. Thus, it will not help in finding a better match for f_d.

The last approximation (11) explains the changes in f_d, mentioned in Section 1, if we change p but keep n, ε, and d fixed.

Tables 1 and 2 show the exact probabilities and their approximations via Equations (10) and (11) after dropping the terms with $O(\varepsilon^3)$. Note that in some cases the condition $(d - 0.5)/\varepsilon \le n$ is violated and thus, the approximation fails to reach an acceptable accuracy.

6. Poisson approximation

If $p - \varepsilon$ is close to 1 then the approximation (5) by normal distribution does not work except for large n. In this case, one might try to use approximation by the Poisson distribution for $n - X$ and $n - Y$, respectively. For instance, if $p - \varepsilon \ge 0.90$ and $n(1 - p + \varepsilon) \le 10$ then

Table 1. The values of $f_1(12, p, \varepsilon)$ with $p = 0.55, 0.60, \ldots, 0.80$ and $\varepsilon = 0.01, 0.05, 0.10$.

	$\varepsilon = 0.01$		$\varepsilon = 0.05$		$\varepsilon = 0.10$	
p	Exact	Approximation	Exact	Approximation	Exact	Approximation
0.55	0.4386	0.4290	0.5176	0.5177	0.6151	0.6171
0.60	0.4376	0.4280	0.5178	0.5179	0.6156	0.6176
0.65	0.4359	0.4263	0.5180	0.5182	0.6173	0.6193
0.70	0.4332	0.4238	0.5185	0.5188	0.6203	0.6220
0.75	0.4291	0.4206	0.5190	0.5195	0.6249	0.6258
0.80	0.4228	0.4166	0.5198	0.5203	0.6317	0.6308

Table 2. The values of $f_2(12, p, \varepsilon)$ with $p = 0.55, 0.60, \ldots, 0.80$ and $\varepsilon = 0.01, 0.05, 0.10$.

	$\varepsilon = 0.01$		$\varepsilon = 0.05$		$\varepsilon = 0.10$	
p	Exact	Approximation	Exact	Approximation	Exact	Approximation
0.55	0.2864	0.2191	0.3579	0.3242	0.4539	0.4429
0.60	0.2835	0.2147	0.3563	0.3223	0.4536	0.4426
0.65	0.2782	0.2073	0.3531	0.3185	0.4527	0.4417
0.70	0.2699	0.1967	0.3479	0.3128	0.4509	0.4403
0.75	0.2578	0.1830	0.3401	0.3053	0.4482	0.4383
0.80	0.2400	0.1662	0.3283	0.2959	0.4439	0.4357

$n - X \sim$ Poisson $[n(1-p)]$ and $n - Y \sim$ Poisson $[n(1-p+\varepsilon)]$, approximately. Therefore, the distribution of $X - Y$ can be approximated by the distribution of the difference of two independent Poisson random variables $n - Y$ and $n - X$. The difference of two Poisson random variables follows a Skellam distribution. We have that

$$f_d(n, p, \varepsilon) \approx P(X - Y \geq d) = P((n - Y) - (n - X) \geq d)$$
$$= \sum_{k \geq d} e^{-(\mu_1+\mu_2)} \left(\frac{\mu_1}{\mu_2}\right)^{k/2} I_k(2\sqrt{\mu_1\mu_2})$$

with $\mu_1 = n(1 - p + \varepsilon)$ and $\mu_2 = n(1 - p)$, and $I_k(x)$ being the modified Bessel function of the first kind [2]. For example, we get that $f_2(12, 0.95, 0.05) = 0.2247$ while the above approximation yields 0.2283.

7. Optimization

One might wonder what the largest possible probabilities in (B) and (D) are. Note that using the probabilistic context, $f_d(n, p, \varepsilon)$ increases as $n \to \infty$ and $\varepsilon > 0$, $d \geq 0$, and $p, \varepsilon \leq p \leq 1$ are kept fixed. Clearly,

$$f_0(n, p, 0) = 1 - f_1(n, p, 0), \tag{12}$$

and its value, $(1 + T)/2$, can be determined by identity (2), possibly using the approximation (3) for T. By Remark 2, $f_d(n, p, \varepsilon)$ goes to ∞ for d even and to $-\infty$ for d odd as $|p|$ grows indefinitely.

It is more interesting to look for

$$\max_{\varepsilon \leq p \leq 1} f_d(n, p, \varepsilon)$$

for different values of d with n and ε kept fixed. Of course, by symmetry, we can, as we will do, focus on the range $(1 + \varepsilon)/2 \leq p \leq 1$.

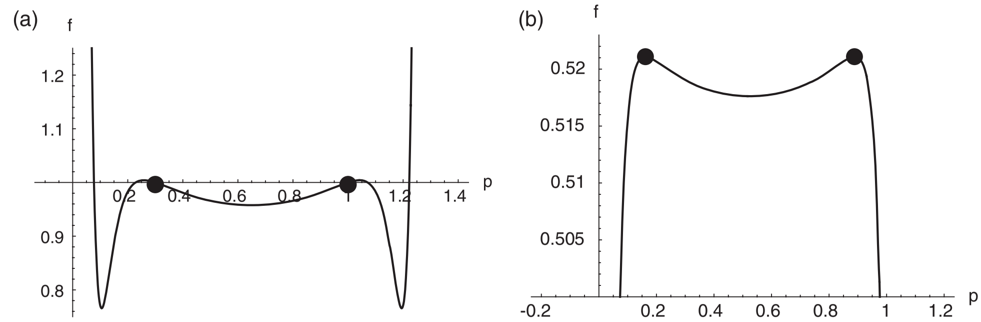

Figure 1. The largest probabilities in two examples with the maximum locations emphasized.

We observe that the shape of f_d about $p = (1+\varepsilon)/2$ changes from concave up to concave down as d increases. For $\varepsilon = 0$, the shape of f_0 is that of a distorted parabola opening up which is fairly flat about its vertex for all $n \geq 1$. The shape becomes concave down at $p = 1/2$ for $d \geq 1$. For other values of ε this change at $p = (1+\varepsilon)/2$ happens at a higher value of d depending on n and ε, too. For example, with $\varepsilon = 0.3$, $(\partial^2/\partial p^2) f_d(12, 0.65, 0.3)$, $(\partial^2/\partial p^2) f_d(13, 0.65, 0.3)$, and $(\partial^2/\partial p^2) f_d(16, 0.65, 0.3)$ become negative at $d = 4$, 5, and 6, respectively.

As a consequence, typically, the maximum occurs at $p = 1$ when d is small (cf. Figure 1(a), $f_0(12, p, 0.3)$). However, when it is not the case, the approximation methods of Sections 5 and 6 can hardly help. In fact, for $n = 12$, $\varepsilon = 0.05$, and $d = 1$ the optimum is found around $p = 0.8883$, and the probability appears to be sharply decreasing as p increases from this value on (cf. Figure 1(b), $f_1(12, p, 0.05)$). On the other hand, as d grows, it appears that the maximum occurs at $p = (1+\varepsilon)/2$.

We prove only the following theorem.

THEOREM 7.1 *The polynomial $f_0(n, p, 0)$ is concave up, and thus, it takes its minimum at $p = 1/2$. Its maximum is taken at* 0 *and* 1.

Proof By the notation (4), we get that

$$f_d(n,p,\varepsilon) = \left(\left(\frac{1-\varepsilon}{2} - \Delta\right)\left(\frac{1+\varepsilon}{2} - \Delta\right)\right)^n \sum_{k=d}^{n} \binom{n}{k} \left(\frac{\Delta + (1+\varepsilon)/2}{(1-\varepsilon)/2 - \Delta}\right)^k \times \sum_{j=0}^{k-d} \binom{n}{j} \left(\frac{\Delta + (1-\varepsilon)/2}{(1+\varepsilon)/2 - \Delta}\right)^j .$$

With $\varepsilon = 0$, this simplifies to

$$\frac{(1-2\Delta)^{2n}}{2^{2n}} \sum_{k=d}^{n} \binom{n}{k} \left(\frac{2\Delta+1}{1-2\Delta}\right)^k \sum_{j=0}^{k-d} \binom{n}{j} \left(\frac{2\Delta+1}{1-2\Delta}\right)^j , \tag{13}$$

which can be expanded as a function of Δ^2

$$f_d(n,p,0) = \sum_{k=0}^{n} c_{2k}(n,d)\Delta^{2k}$$

according to Theorem 2.1. In general, determining $c_{2k}(n,d)$ requires the evaluation of a triple sum according to Formula (13). On the other hand, we can easily derive that the coefficient of the term $\Delta^2 = (p - 1/2)^2$ is

$$c_2(n, d) = \frac{1}{2^{2n-2}} \sum_{k=d}^{n} \sum_{j=0}^{k-d} \binom{n}{k}\binom{n}{j}(2(k + j - n)^2 - n).$$

For instance, $c_2(n, n) = -n/2^{2n-2} < 0$ follows immediately. With some calculations and simplifications, we obtain the recurrence relation

$$c_2(n, d + 1) = c_2(n, d) - \frac{1}{2^{2n-2}} \sum_{k=d}^{n} \binom{n}{k}\binom{n}{k-d}(2(2k - d - n)^2 - n), \quad \text{for } d \geq 0,$$

with the initial condition

$$c_2(n, 0) = \frac{1}{2^{2n-2}}\binom{2n-2}{n-1} \sim \frac{1}{\sqrt{\pi(n-1)}}, \tag{14}$$

which already guarantees the convexity of $f_0(n, p, 0)$ at $p = 1/2$. We note that Equation (14) follows from the observation (12) which implies that $c_2(n, 1) = -c_2(n, 0)$. In a similar fashion, by Equations (2) and (12), $c_0(n, 1) = 1 - c_0(n, 0)$ implies that

$$f_0\left(n, \frac{1}{2}, 0\right) = c_0(n, 0) = \frac{1}{2}\left(1 + \frac{\binom{2n}{n}}{2^{2n}}\right). \tag{15}$$

With considerably more calculations, we obtain all coefficients

$$c_{2k}(n, 0) = \frac{1}{2^{2n-2k}}\binom{2k-1}{k}\binom{2n-2k}{n-k}, \quad k = 1, 2, \ldots, n, \tag{16}$$

which proves the convexity of $f_0(n, p, 0)$ everywhere. The minimum of $f_0(n, p, 0)$ is taken at $p = 1/2$, while the locations of the maximum are 0 and 1 since the function is symmetric about $p = 1/2$ by Theorem 2.1.

Note that Equation (16) can be verified by finding that

$$(2(n+2)N - (2n+1))(2(n+2)N^2 - (2n+3)(1+4\Delta^2)N + 8(n+1)\Delta^2)$$

is an annihilating operator (cf. [6,7]), using Zeilberger's algorithm [6] by calling the `Zb` function of the `Zb` or the certificate finding `FindRecurrence` function of the `MultiSum` [7] Mathematica package, for

$$\sum_{k=1}^{n} \frac{1}{2^{2n-2k}}\binom{2k-1}{k}\binom{2n-2k}{n-k}\Delta^{2k} = \frac{\binom{2n}{n}}{2^{2n+1}}\left(-1 + {}_2F_1\left(\frac{1}{2}, -n, \frac{1}{2} - n, (2\Delta)^2\right)\right), \tag{17}$$

with N being the forward shift operator with respect to n. It also annihilates the double sum $f_0(n, p, 0) - c_0(n, 0)$ which can be numerically verified for particular values of n. (Note that the left factor of the annihilator changes to $N - 1$ if we include the constant term, indicating that the right factor annihilates the sums up to an inhomogeneous part which is free of n; in fact, it is $(1 - 4x^2)/2$.) The symbolic verification can be effectively done by the latest version of the Mathematica package `Sigma` as it was pointed out by Schneider [8]. The proof is complete after comparing the initial values of Equation (17) and the double sum $f_0(n, p, 0) - c_0(n, 0)$ based on Formula (13). ■

We note that for $d = 0$, Formula (13) simplifies to

$$f_0(n, p, 0) = \frac{(1-2\Delta)^{2n}}{2^{2n}} \sum_{k=0}^{n} \sum_{j=0}^{n} \binom{n}{k} \binom{n}{k+j} \left(\frac{2\Delta+1}{1-2\Delta}\right)^{2k+j}.$$

We add that Equation (12) also implies that $c_{2k}(n, 1) = -c_{2k}(n, 0)$, $k \geq 1$, and thus, $f_1(n, p, 0)$ becomes concave down at $p = 1/2$.

Note that for $d = n$ it is straightforward to prove the following theorem.

Theorem 7.2 *The roots of* $(\partial/\partial p) f_n(n, p, \varepsilon)$ *are* 0, $(1+\varepsilon)/2$, *and* $1+\varepsilon$ *with multiplicity* $n-1$, 1, *and* $n-1$, *respectively. Thus,*

$$\max_{\varepsilon \leq p \leq 1} f_n(n, p, \varepsilon) = f_n\left(n, \frac{1+\varepsilon}{2}, \varepsilon\right).$$

We note that very recently the problem of finding

$$\max_n (1 - f_0(n, p, \varepsilon))$$

appeared in [9,4] with the pairs $p = 0.51, \varepsilon = 0.01$, and $p = 0.101, \varepsilon = 0.001$. Here $1 - f_0(n, p, \varepsilon)$ corresponds to the probability that the weaker player scores more than the stronger one. The constant term approximation (10) of Theorem 5.1 and numerical evidence suggest that the optimum value n and the corresponding probability $1 - f_0(n, p, \varepsilon)$ must be close to $1/(2\varepsilon)$ and $0.5 - 0.88\sqrt{\varepsilon}$, respectively, for any sufficiently small $\varepsilon > 0$ and $|\Delta|$. We observe that only ε seems to matter as long as $|\Delta|$ is small, in agreement with our findings regarding (B) and (D). The results can be easily generalized for $1 - f_d(n, p, \varepsilon)$ with the difference $d \leq 0$, i.e. the probability that the weaker player scores at least $-d + 1 \geq 1$ points more than the stronger one, since in this case the approximation (7) works even for small values of n. From the point of view of the weaker player, with conveniently switching the sign of d, we get the maximum probability that the weaker player wins by more than $d \geq 0$ points and the corresponding n are close to $0.5 - 0.88\sqrt{\varepsilon(1+2d)}$ and $(1+2d)/(2\varepsilon)$, respectively, for any sufficiently small $\varepsilon > 0$ and $|\Delta|$.

On another note, Wagon [5] conjectures that $f_0(n, p, \varepsilon)$, $\varepsilon \leq p \leq 1$, takes its minimum at $(1+\varepsilon)/2$. In general, for $f_d(n, p, \varepsilon)$ with $d \geq 1$, the author believes that $p = (1+\varepsilon)/2$ switches from the location of the maximum to that of the minimum as ε grows, leaving the original maximization problem open.

8. The absolute spread difference for $p = q = 1/2$

We can consider the absolute spread difference between X and Y. Here we deal with the special case $p = q = 1/2$ which implies $\varepsilon = 0$. Since $X - Y \sim \mathcal{N}[\mu = 0, \sigma = \sqrt{n/2}]$ approximately, we get that $|X - Y|$ is approximately of half-normal distribution, i.e. the distribution of the absolute value of a normally distributed random variable centred at zero with $\sigma = \sqrt{n/2}$. Determining the moments of $|X - Y|$ raises some interesting questions involving double summations.

Theorem 8.1 *For the* r*th raw moment of* $|X - Y|$ *we get that*

$$m_r = \sum_{k=0}^{n} \sum_{j=0}^{n} \frac{|k-j|^r}{2^{2n}} \binom{n}{k} \binom{n}{j} = 2 \sum_{k=0}^{n} \sum_{j=0}^{k} \frac{(k-j)^r}{2^{2n}} \binom{n}{k} \binom{n}{j} \qquad (18)$$

with the particular values

$$m_0 = 1 - \frac{\binom{2n}{n}}{2^{2n}}, \quad m_1 = \frac{n}{2^{2n}}\binom{2n}{n}, \quad m_2 = \frac{n}{2}, \quad m_3 = \frac{n^2}{2^{2n}}\binom{2n}{n},$$
$$m_4 = \frac{n(3n-1)}{4}, \quad m_5 = \frac{n^2(2n-1)}{2^{2n}}\binom{2n}{n}, \quad \text{and} \quad m_6 = \frac{n(15n^2-15n+4)}{8}.$$

Of course, $E(|X-Y|^2) = \text{var}(X-Y) = n/2$. We also observe that $m_0 = 1 - P(X = Y) = 1 - \binom{2n}{n}/2^{2n}$ which provides us with an alternative proof of Equation (15).

The proof of Theorem 8.1 can be accomplished by the use of the Mathematica package `MultiSum` (cf. [7]) or `Sigma` (cf. [10]). Further values of m_r have been determined by Schneider [8]. Note that finding moments with even indices might be easier since the summation variables have standard bounds if we use the first form of m_r in Equation (18) while odd indices require the second form. The closed form for m_1 was originally suggested by John Essam and derived in [7].

We note that Theorem 8.1 is in agreement with the moments of the corresponding half-normal distribution in an asymptotic sense.

Acknowledgement

The author thanks a referee, Carsten Schneider, Gregory P. Tollisen, and Stan Wagon for helpful comments that improved the presentation of this paper.

References

[1] L.L. Colquitt, N.H. Godwin, and S. Swidler, *Betting on long shots in NCAA basketball games and implications for skew loving behavior*, Finance Res. Lett. 1 (2004), pp. 119–126.
[2] L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht, 1974.
[3] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.
[4] V. Addona, S. Wagon, and H. Wilf, *How to lose as little as possible*, pp. 1–26, `arXiv:1002.1763v1`. Available at http://arxiv.org/abs/1002.1763, 9 February 2010.
[5] S. Wagon, personal communication, 2010.
[6] M. Petkovsek, H.S. Wilf, and D. Zeilberger, *A=B*, AK Peters, Ltd., Wellesley, MA, 1996.
[7] K. Wegschaider, *Computer generated proofs of binomial multi-sum identities*, Master's thesis, RISC, J. Kepler University, 1997. Available at `http://www.risc.uni-linz.ac.at/publications/download/risc_2245/diplom.ps`
[8] C. Schneider, personal communication, 2009–2010.
[9] Problem of the week: Flip some coins, Macalester College 2009, Problem #1128. Available at http://mathforum.org/wagon/fall09/p1128.html
[10] C. Schneider, *Symbolic summation assists combinatorics*, Sem. Lothar. Combin. 56 (2007), pp. 1–36.