International Journal of Modern Mathematics () (),

Maximizing a combinatorial expression arising from crowd estimation

Donald M. Davis

Received October 30, 2009; Revised May 14, 2010

Abstract

We determine, within 1, the value of N for which $\sum_i \binom{s_1}{i}\binom{s_2}{N}\binom{s_1}{N-i}\binom{N}{i}$ achieves its maximum value. Here s_1 and s_2 are fixed integers. This problem arises in studying the most likely value of $|A \cup B \cup C|$ if A and C are disjoint sets of cardinality s_1, and $|B| = s_2$. Attempting to remove the 1 unit of indeterminacy leads to interesting conjectures about a family of rational functions.

Keywords: Binomial coefficients, size of union.

2000 Mathematics Subject Classification: 05A10,60C05.

1 Introduction

The question considered here arises from problems involving estimating sizes of crowds. You count sizes of certain subsets and want to estimate the size of the union.

The case considered here involves three sets A, B, and C with the property that $A \cap C = \emptyset$, but B may intersect the other sets. Suppose also that $|A| = s_1$, $|B| = s_2$, and $|C| = s_3$. What is the most likely value for $|A \cup B \cup C|$? This question was suggested to the author by Fred Cohen, along with the mathematical model which we now present.

The assumption being made is that any choice of i people in $A \cap B$ and j other people in $B \cap C$ is equally likely, for all i and j. For example, it is equally likely that $|A \cap B| = 1$ with that person a specified person of A and a specified person of B or that $|A \cap B| = 3$ with those people a specified subset of A and a specified subset of B. This can be formulated in the following way.

Suppose $\widehat{A}$, $\widehat{B}$, and $\widehat{C}$ are disjoint sets with $|\widehat{A}| = s_1$, $|\widehat{B}| = s_2$, and $|\widehat{C}| = s_3$. The sample space consists of all 4-tuples

$$(A_1, B_1, B_2, C_2) \subset (\widehat{A}, \widehat{B}, \widehat{B}, \widehat{C})$$

such that $|A_1| = |B_1|$, $|B_2| = |C_2|$, and $B_1 \cap B_2 = \emptyset$. Thus A_1 and B_1 correspond to $A \cap B$ in the earlier formulation, and we have

$$|A \cup B \cup C| = s_1 + s_2 + s_3 - |B_1| - |B_2|.$$

We assume that each element of the sample space is equally likely. Let $E_{i,j}$ denote the event that $|B_1| = i$ and $|B_2| = j$. Then

$$|E_{i,j}| = \tbinom{s_1}{i}\tbinom{s_2}{i}\tbinom{s_2-i}{j}\tbinom{s_3}{j}.$$

If E_N is the event that $|B_1| + |B_2| = N$, i.e., that $|A \cup B \cup C| = s_1 + s_2 + s_3 - N$, then

$$|E_N| = \sum_{i+j=N} |E_{i,j}|.$$

Hence the most likely value of $|A \cup B \cup C|$ is $s_1 + s_2 + s_3 - N$, where N maximizes

$$\sum_i \tbinom{s_1}{i}\tbinom{s_2}{i}\tbinom{s_2-i}{N-i}\tbinom{s_3}{N-i} = \sum_i \tbinom{s_1}{i}\tbinom{s_2}{N}\tbinom{s_3}{N-i}\tbinom{N}{i}.$$

We focus attention primarily on the case in which $s_1 = s_3$. In this case we obtain in Corollary 1.1 a simple formula for the maximizing N within 1, and in 1.2 a much-less-tractable formula which removes the indeterminacy. In Section 3, we attempt to obtain a more useful approximation to 1.2, and in doing so we notice fascinating patterns in a family of rational functions, but can only conjecture that these patterns persist. See Table 3.3 and Conjecture 3.2. In Section 4, we consider the general case when s_1 and s_3 need not be equal. Our results there are somewhat similar, but not so complete.

Our main theorem is

Theorem 1.1. *Let $f_{s_1,s_2}(N) := \binom{s_2}{N}\sum\limits_i \binom{s_1}{i}\binom{s_1}{N-i}\binom{N}{i}$ for integer values of N. For each s_1 and s_2, there is an integer, which we denote by $g(s_1, s_2)$, such that $f_{s_1,s_2}(N)$ is an increasing function of N for $N \le g(s_1, s_2)$, and a decreasing function of N for $N \ge g(s_1, s_2)$. Moreover,*

$$g(s_1, s_2) = \left[2s_1 + s_2 + \tfrac{3}{2} - \sqrt{4s_1^2 + 4s_1 + (s_2 + \tfrac{1}{2})^2}\right] + \delta \tag{1.1}$$

with $\delta = 0$ or 1.

Corollary 1.1. *The maximum value of $\sum\limits_i \binom{s_1}{i}\binom{s_2}{N}\binom{s_1}{N-i}\binom{N}{i}$ occurs when*

$$N = \left[2s_1 + s_2 + \tfrac{3}{2} - \sqrt{4s_1^2 + 4s_1 + (s_2 + \tfrac{1}{2})^2}\right] + \delta$$

with $\delta = 0$ or 1. The most likely value of $|A \cup B \cup C|$ in the situation discussed above, with $s_3 = s_1$, is $\left[\sqrt{4s_1^2 + 4s_1 + (s_2 + \frac{1}{2})^2} - \frac{3}{2}\right] - \delta$.

It is conceivable that f_{s_1,s_2} might achieve equal maxima at both N and $N + 1$. In such a case, we accept either as an allowable value of $g(s_1, s_2)$.

It was pointed out by the referee that the unimodality of $f_{s_1,s_2}(N)$ follows from a result of Walkup ([1]) which states that a binomial convolution of log-concave functions is log-concave. Since the sequence of binomial coefficients is log-concave, and log-concave implies unimodal, the unimodality of our f follows. But finding the mode is much more delicate.

To illustrate the efficacy of our formula, we consider the typical case $s_1 = 15$, and tabulate in Table 1.1 the actual values of $g(15, s_2)$ for all s_2, and in Table 1.2 the five values of s_2 for which $\delta = 1$ in (1.1). Note how in these five cases the expression whose integer part appears in (1.1) falls slightly short of the required value.

The following proposition generalizes the beginning and end of Table 1.1. Theorem 1.1 is true with $\delta = 0$ in these cases. Note also that the case $d = 0$ of part (b) of Proposition 1.1 shows that if B is much larger than A and C, then the most likely occurrence is that both A and C are contained in B.

Proposition 1.1.

a. If $s_2 \leq \frac{1}{2}(\sqrt{8s_1+9}-1)$*, then* $g(s_1, s_2) = s_2$.

b. For all s_1, s_2*, we have* $g(s_1, s_2) \leq 2s_1$*. For* $0 \leq d \leq 4$*,*

$$g(s_1, s_2) \geq 2s_1 - d \text{ iff } s_2 \geq \tfrac{2}{d+1}(s_1^2 + s_1) - \tfrac{d+2}{2}. \tag{1.2}$$

Table 1.1: Values of $g(15, s_2)$

s_2	$g(15, s_2)$
$[1, 6]$	s_2
$[7, 10]$	$s_2 - 1$
$[11, 12]$	$s_2 - 2$
$[13, 15]$	$s_2 - 3$
$[16, 17]$	$s_2 - 4$
18	13
$[19, 21]$	14
$[22, 23]$	15
$[24, 26]$	16
$[27, 29]$	17
$[30, 33]$	18
$[34, 37]$	19
$[38, 42]$	20
$[43, 48]$	21
$[49, 55]$	22
$[56, 64]$	23
$[65, 76]$	24
$[77, 92]$	25
$[93, 117]$	26
$[118, 157]$	27
$[158, 238]$	28
$[239, 478]$	29
≥ 479	30

Table 1.2: Cases in which equality does not hold in (1.1) when $s_1 = 15$ and $\delta = 0$

s_2	$g(15, s_2)$	$2s_1 + s_2 + \frac{3}{2} - \sqrt{4s_1^2 + 4s_1 + (s_2 + \frac{1}{2})^2}$
6	6	5.84
10	9	8.78
15	12	11.85
17	13	12.91
19	14	13.89

Next we introduce the polynomials involved in the proof. We will usually replace s_1 by x, both because it will occur as a variable in polynomials, and so that we can use the notation $x_i = x(x-1)\cdots(x-i+1)$. For a nonnegative integer d, define a polynomial

$P_d(x)$ of degree $2d$ by

$$P_d(x) = \sum_{i=0}^{d} \frac{(x_i)^2(x_{d-i})^2}{i!(d-i)!}.$$

We will prove the following key result in Section 2.

Lemma 1.1. *When $P_{d+1}(x)$ is divided by $P_d(x)$, the quotient is $\frac{2}{d+1}x^2 - \frac{2d}{d+1}x + \frac{d}{2}$. Let $R_d(x)$ denote the remainder. If f is as in 1.1, then, if x and d are integers,*

$$f_{x,s_2}(2x-d) \geq f_{x,s_2}(2x-d-1) \text{ iff } s_2 \geq \tfrac{2}{d+1}(x^2+x) - \tfrac{d+2}{2} + \tfrac{R_d(x)}{P_d(x)}. \tag{1.3}$$

In Section 2, we will also prove the following result, the proof of which is less straightforward than that of Lemma 1.1.

Lemma 1.2. *For $d \geq 5$ and $x > d/2$, $-0.5 < R_d(x)/P_d(x) < 0$.*

Now we can prove the main theorem.

Proof of Theorem 1.1. By Lemma 1.2, increasing d by 1 changes $R_d(x)/P_d(x)$ by at most 1/2, and clearly it decreases $\frac{2}{d+1}(x^2+x) - \frac{d+2}{2}$ by more than 1/2. Thus, for $x > d/2$, the RHS of (1.3) is a decreasing function of d, at least for integer values of d. This, with (1.3), implies the unimodality part of the theorem, with maximum of $f_{s_1,s_2}(N)$ occurring for $N = 2s_1 - d$ for the smallest integer d such that the RHS of (1.3) is satisfied.

We will show that Lemma 1.2 also implies that

$$[2s_1 - d_1] \leq g(s_1,s_2) \leq [2s_1 - d_1] + 1, \tag{1.4}$$

where d_1 satisfies

$$s_2 = \tfrac{2}{d_1+1}(s_1^2+s_1) - \tfrac{d_1+2}{2}.$$

This value is $d_1 = -s_2 - \frac{3}{2} + \sqrt{4s_1^2 + 4s_1 + (s_2+\frac{1}{2})^2}$, yielding (1.1).

To prove (1.4), write $d_1 = d_2 - t$, with $0 \leq t < 1$, and d_2 an integer. Thus $[2s_1 - d_1] = 2s_1 - d_2$. The RHS of (1.3) is satisfied using d_2 since, using 1.2 at the last step,

$$s_2 = \tfrac{2}{d_1+1}(s_1^2+s_1) - \tfrac{d_1+2}{2} \geq \tfrac{2}{d_2+1}(s_1^2+s_1) - \tfrac{d_2+2}{2} \geq \tfrac{2}{d_2+1}(s_1^2+s_1) - \tfrac{d_2+2}{2} + \tfrac{R_{d_2}(s_1)}{P_{d_2}(s_1)}.$$

Therefore $g(s_1,s_2) \geq 2s_1 - d_2$. We will now show that the RHS of (1.3) is not satisfied using $d = d_2 - 2$, which implies $g(s_1,s_2) \leq 2s_1 - d_2 + 1$, hence completing the proof.

To see this, let $h(d) = s_2 - \left(\frac{2}{d+1}(s_1^2+s_1) - \frac{d+2}{2}\right)$. Then $h(d_1) = 0$ and $d_1 - (d_2 - 2) > 1$, hence $h(d_2-2) < -\frac{1}{2}$. Therefore when $d = d_2 - 2$, using Lemma 1.2 again, $s_2 - \left(\frac{2}{d+1}(s_1^2+s_1) - \frac{d+2}{2} + \frac{R_d(s_1)}{P_d(s_1)}\right) < 0$, as desired. □

In terms of R_d/P_d, we give in 1.2 a precise result about whether $\delta = 0$ or 1 in Theorem 1.1. The usefulness of this is limited by the complicated nature of R_d/P_d. In Section 3, we discuss a very strong conjecture regarding R_d/P_d, which, if proved, would make Theorem 1.2 more useful. See Theorem 3.1. The evidence for this conjecture leads to remarkable conjectural patterns among some rational functions. See Table 3.3 and Conjecture 3.2.

Theorem 1.2. *Let*

$$d_0=\left[\sqrt{4s_1^2+4s_1+(s_2+\tfrac{1}{2})^2}-s_2-\tfrac{3}{2}\right].$$

Then (1.1) is true with $\delta=1$ *iff*

$$0<\tfrac{2}{d_0+1}(s_1^2+s_1)-\tfrac{d_0+2}{2}-s_2\le-\tfrac{R_{d_0}(s_1)}{P_{d_0}(s_1)}. \tag{1.5}$$

Proof. Let $d_1=\sqrt{4s_1^2+4s_1+(s_2+\frac{1}{2})^2}-s_2-\frac{3}{2}$. Then $s_2=\frac{2}{d_1+1}(s_1^2+s_1)-\frac{d_1+2}{2}$.

If d_1 is an integer, then $d_0=d_1$ and the (>0)-condition in (1.5) is not satisfied, and the RHS of (1.3) is not satisfied when $d=d_1-1$ by an argument similar to the proof of 1.1. Hence $g(s_1,s_2)=2s_1-d_1$, verifying Theorem 1.2 in this case.

If d_1 is not an integer, then (1.1) is true with $\delta=1$ iff the RHS of (1.3) is satisfied using d_0 (since $2s_1-d_0=[2s_1-d_1]+1$), but the RHS of (1.3) is exactly the $\le$-part of (1.5). Note that the (>0)-part of (1.5) is certainly satisfied in this case, since the middle expression in (1.5) equals 0 using d_1, and is a strictly decreasing function of d. $\square$

2 Combinatorial proofs

In this section we prove Lemma 1.1, Proposition 1.1, and Lemma 1.2.

Proof of Lemma 1.1. Cancelling common factors in the binomial coefficients involving s_2, we find that the LHS of (1.3) is equivalent to

$$\begin{aligned}&(s_2-2x+d+1)\sum\frac{(x!)^2(2x-d)!}{(i!)^2((2x-d-i)!)^2(x-i)!(d+i-x)!}\\ \ge\quad&(2x-d)\sum\frac{(x!)^2(2x-d-1)!}{(i!)^2((2x-d-1-i)!)^2(x-i)!(d+1+i-x)!}.\end{aligned}$$

Cancelling $(2x-d)!$ and letting $j=x-i$, we obtain the equivalent condition

$$\begin{aligned}&(s_2-2x+d+1)\sum\frac{(x!)^2}{(x-j)!^2(x+j-d)!^2j!(d-j)!}\\ \ge\quad&\sum\frac{(x!)^2}{(x-j)!^2(x+j-d-1)!^2j!(d+1-j)!}.\end{aligned}$$

Multiplying both sides by $(x!)^2$, the inequality becomes

$$(s_2-2x+d+1)\sum\frac{(x_j)^2(x_{d-j})^2}{j!(d-j)!}\ge\sum\frac{(x_j)^2(x_{d+1-j})^2}{j!(d+1-j)!},$$

and this readily yields

$$f_{x,s_2}(2x-d)\ge f_{x,s_2}(2x-d-1)\text{ iff }s_2\ge\frac{P_{d+1}(x)}{P_d(x)}+2x-d-1. \tag{2.1}$$

Our next claim is that the leading terms of $P_d(x)$ are given by

$$P_d(x) = \frac{2^d}{d!}x^{2d} - \frac{2^{d-1}}{(d-2)!}x^{2d-1} + \frac{2^{d-3}}{3(d-2)!}(3d^2-5d+4)x^{2d-2} + \text{lower}. \qquad (2.2)$$

The proof of (2.2) makes frequent use of

$$\sum \tfrac{1}{i!(d-i)!} = \tfrac{2^d}{d!}, \qquad (2.3)$$

which is true since each side is the coefficient of x^d in $e^x \cdot e^x = e^{2x}$. Then (2.3) is exactly the coefficient of x^{2d} in $P_d(x)$. The coefficient of x^{2d-1} is

$$-\sum_{i=0}^{d} \frac{i(i-1)+(d-i)(d-i-1)}{i!(d-i)!} = -2\sum_{i=0}^{d} \tfrac{1}{(i-2)!(d-i)!} = -\tfrac{2^{d-1}}{(d-2)!}.$$

We have used symmetry in the first step. Note that the $i(i-1)$ comes as $\sum\limits_{j=0}^{i-1} 2j$.

The next coefficient of $P_d(x)$ is obtained similarly, but involves much more work. The coefficient of x^{2d-2} in $(x_i)^2(x_{d-i})^2$ is

$$\sum_{j=0}^{i-1} j^2 + \sum_{j=0}^{d-i-1} j^2 + 4\sum_{j=0}^{i-1} j \sum_{j=0}^{d-i-1} j + \sum_{1\le j_1<j_2<i} 4j_1j_2 + \sum_{1\le j_1<j_2<d-i} 4j_1j_2.$$

Noting that

$$2\sum_{1\le j_1<j_2<i} j_1j_2 = \left(\sum_{j=1}^{i-1} j\right)^2 - \sum_{j=1}^{i-1} j^2,$$

we obtain

$$\begin{aligned} & -\sum_{j=0}^{i-1} j^2 - \sum_{j=0}^{d-i-1} j^2 + (i-1)i(d-i-1)(d-i) + 2\big(\tfrac{(i-1)i}{2}\big)^2 + 2\big(\tfrac{(d-i-1)(d-i)}{2}\big)^2 \\ = \; & -\tfrac{(i-1)i(2i-1)}{6} - \tfrac{(d-i-1)(d-i)(2d-2i-1)}{6} + (i-1)i(d-i-1)(d-i) \\ & + \tfrac{i^2(i-1)^2}{2} + \tfrac{(d-i)^2(d-i-1)^2}{2}. \end{aligned}$$

Thus, using symmetry, the coefficient of x^{2d-2} in $\sum\limits_{i=0}^{d} \dfrac{(x_i)^2(x_{d-i})^2}{i!(d-i)!}$ is

$$\sum \tfrac{i(i-1)}{(i-2)!(d-i)!} + \sum \tfrac{1}{(i-2)!(d-i-2)!} - \tfrac{1}{3}\sum \tfrac{2i-1}{(i-2)!(d-i)!}.$$

Next note that

$$\begin{aligned}
\sum \tfrac{i(i-1)}{(i-2)!(d-i)!} &= \sum \tfrac{(j+1)(j+2)}{j!(d-2-j)!}\\
&= 2\frac{2^{d-2}}{(d-2)!}+3\frac{2^{d-3}}{(d-3)!}+\sum\frac{(j-1)+1}{(j-1)!(d-2-j)!}\\
&= \frac{2^{d-1}}{(d-2)!}+3\frac{2^{d-3}}{(d-3)!}+\frac{2^{d-4}}{(d-4)!}+\frac{2^{d-3}}{(d-3)!}\\
&= \frac{2^{d-4}}{(d-2)!}(8+8(d-2)+(d-2)(d-3))\\
&= \frac{2^{d-4}}{(d-2)!}(d^2+3d-2).
\end{aligned}$$

Also,

$$\sum \tfrac{2i-1}{(i-2)!(d-i)!}=2\sum\tfrac{1}{(i-3)!(d-i)!}+3\sum\tfrac{1}{(i-2)!(d-i)!}=\tfrac{2^{d-2}}{(d-2)!}(d-2+3).$$

Hence the desired coefficient equals

$$\begin{aligned}
&\frac{2^{d-4}}{(d-2)!}(d^2+3d-2)+\frac{2^{d-4}}{(d-4)!}-\frac{2^{d-2}}{3(d-2)!}(d+1)\\
=\ &\frac{2^{d-3}}{3(d-2)!}(3d^2-5d+4),
\end{aligned}$$

as asserted in (2.2).

Next we claim that

$$\frac{P_{d+1}(x)}{P_d(x)}=\tfrac{2}{d+1}x^2-\tfrac{2d}{d+1}x+\tfrac{d}{2}+\tfrac{R_d(x)}{P_d(x)}, \tag{2.4}$$

with $\deg(R_d(x))<2d$, the claim of the first part of Lemma 1.1. This can be discovered by division of polynomials, using (2.2), but it is simpler just to verify that

$$\begin{aligned}
&\left(\tfrac{2^d}{d!}x^{2d}-\tfrac{2^{d-1}}{(d-2)!}x^{2d-1}+\tfrac{2^{d-3}}{3(d-2)!}(3d^2-5d+4)x^{2d-2}\right)\left(\tfrac{2}{d+1}x^2-\tfrac{2d}{d+1}x+\tfrac{d}{2}\right)\\
=\ &\tfrac{2^{d+1}}{(d+1)!}x^{2d+2}-\tfrac{2^d}{(d-1)!}x^{2d+1}+\tfrac{2^{d-2}}{3(d-1)!}(3(d+1)^2-5(d+1)+4)x^{2d}+\text{lower}.
\end{aligned}$$

Combining (2.1) and (2.4), we obtain (1.3). □

Proof of Proposition 1.1. (a). One can easily show that the bracketed expression in (1.1) is always less than s_2+1, and equals s_2 if and only if s_2 satisfies the hypothesis of 1.1(a). Noting that $g(s_1,s_2)$ is $\ge$ the bracketed expression of (1.1) by Theorem 1.1, and is $\le s_2$ since $f_{s_1,s_2}(s_2+1)=0$, part (a) follows.

(b). The first statement is true since $f_{s_1,s_2}(2s_1+1)=0$.

After cancelling common factors in the numerator and denominator, one can compute that

$$\frac{-R_d(x)}{P_d(x)} = \begin{cases} 0 & d = 0, 1 \\ (2x-1)/(3(2x^2-2x+1)) & d = 2 \\ 3(x-1)/(4(x^2-x+1)) & d = 3 \\ 6(2x^3-9x^2+3x-9)/(5(2x^4-8x^3+14x^2-5x+3)) & d = 4 \end{cases}$$

Thus (1.2) follows immediately from (1.3) if $d = 0$ or 1.

Note that (1.3) is only meaningful if $2x > d$. If $d = 2$, then $x \geq 2$. For such x, $-0.2 \leq R_2(x)/P_2(x) < 0$. Since, for integer x, $\frac{2}{3}(x^2+x) - 2$ is either an integer or an integer plus 1/3, an integer s_2 satisfies

$$s_2 \geq \tfrac{2}{3}(x^2+x) - 2 + \tfrac{R_2(x)}{P_2(x)} \text{ iff } s_2 \geq \tfrac{2}{3}(x^2+x) - 2,$$

and so (1.2) follows from (1.3).

A similar argument works for $d = 3$ and 4. For $d = 3$, we have that $\frac{2}{d+1}(x^2+x) - \frac{d+2}{2}$ is an integer plus 1/2, and for $x \geq 2$, $-0.25 \leq R_3(x)/P_3(x) < 0$. If $d = 4$ and $x > 3$, then $\frac{2}{d+1}(x^2+x) - \frac{d+2}{2}$ is an integer or an integer plus t with $t \geq 0.2$, while $-0.12 \leq R_4(x)/P_4(x) < 0$. If $d = 4$ and $x = 3$, (1.3) says $s_2 \geq 2.8 - 0.55$, while the hypothesis says $s_2 \geq 2.8$. These are, of course, equivalent. $\square$

Proof of Lemma 1.2. We begin by removing common factors in $P_d(x)$ and $P_{d+1}(x)$. Since parity of d plays a role, we let $d = 2b + \epsilon$ with $\epsilon \in \{0, 1\}$. When considering $R_{2b+\epsilon}(x)$, we let, for $\delta \in \{0, 1\}$,

$$\widetilde{P}_{2b+\epsilon+\delta}(x) := \frac{P_{2b+\epsilon+\delta}(x)}{\prod_{i=0}^{b-1+\epsilon}(x-i)^2}.$$

Note that $\widetilde{P}_{2b+\epsilon+1}(x)/\widetilde{P}_{2b+\epsilon}(x)$ has the same quotient as $P_{2b+\epsilon+1}(x)/P_{2b+\epsilon}(x)$, while its remainder $\widetilde{R}_{2b+\epsilon}(x)$ satisfies $\widetilde{R}_{2b+\epsilon}(x) = R_{2b+\epsilon}(x)/\prod_{i=0}^{b-1+\epsilon}(x-i)^2$, and hence $\widetilde{R}_{2b+\epsilon}(x)/\widetilde{P}_{2b+\epsilon}(x) = R_{2b+\epsilon}(x)/P_{2b+\epsilon}(x)$.

To prove the lemma, we will prove

1. $\widetilde{P}_{2b+\epsilon}(x+b) > 0$ for $x > 0$,

2. $\widetilde{R}_{2b+\epsilon}(x+b) < 0$ for $x > 0$, and

3. $\widetilde{R}_{2b+\epsilon}(x+b) + \frac{1}{2}\widetilde{P}_{2b+\epsilon}(x+b) > 0$ for $x > 0$.

We have, with $c_{i,b}=1$ unless $i=b$, while $c_{b,b}=\frac{1}{2}$,

$$\begin{aligned}&\tfrac{(2b+\epsilon+\delta)!}{2}\widetilde{P}_{2b+\epsilon+\delta}(x+b)\\
=&\begin{cases}\sum\limits_{i=0}^{b} c_{i,b}\binom{2b}{i}\prod\limits_{j=1}^{b-i}(x-j+1)^2\prod\limits_{j=b-i+1}^{b}(x+j)^2 & \epsilon+\delta=0\\
\sum\limits_{i=0}^{b}\binom{2b+1}{i}\prod\limits_{j=0}^{b-i}(x-j)^2\prod\limits_{j=b-i+1}^{b}(x+j)^2 & \epsilon+\delta=1\\
\sum\limits_{i=0}^{b+1} c_{i,b+1}\binom{2b+2}{i}\prod\limits_{j=0}^{b-i}(x-j-1)^2\prod\limits_{j=b-i+1}^{b}(x+j)^2 & \epsilon+\delta=2.\end{cases}\end{aligned}$$

Part (1) is true since $\widetilde{P}_{2b+\epsilon}(x+b)$ is a sum of nonnegative terms including the term $\prod\limits_{j=1-\epsilon}^{b}(x+j)^2$, which is positive for $x>0$.

Next we consider (2) with $\epsilon=1$. We compute $q_{2b+1}(x+b)=\frac{1}{b+1}(x^2-x)+\frac{1}{2}$. We have

$$\begin{aligned}\tfrac{(2b+2)!}{2}\widetilde{R}_{2b+1}(x+b) &= \tfrac{(2b+2)!}{2}(\widetilde{P}_{2b+2}(x+b)-q_{2b+1}(x+b)\widetilde{P}_{2b+1}(x+b))\\
&= \sum_{i=0}^{b}\left(\prod_{j=1}^{b-i}(x-j)^2\prod_{j=b-i+1}^{b}(x+j)^2\right)F_i,\end{aligned}$$

where

$$\begin{aligned}F_i &= \tbinom{2b+2}{i}(x-b+i-1)^2-(2b+2)\tbinom{2b+1}{i}\left(\tfrac{1}{b+1}(x^2-x)+\tfrac{1}{2}\right)+\tfrac{1}{2}\tbinom{2b+2}{b+1}\delta_{i,b}x^2\\
&= \left(\tbinom{2b+1}{i-1}-\tbinom{2b+1}{i}+\delta_{i,b}\tbinom{2b+1}{b}\right)x^2-2\left(\tbinom{2b+1}{i-1}+(b-i)\tbinom{2b+2}{i}\right)x\\
&\quad+\tbinom{2b+2}{i}(b+1-i)^2-(b+1)\tbinom{2b+1}{i}\\
&= \begin{cases}\binom{2b+1}{i-1}(x-(b+1-i))^2-\binom{2b+1}{i}(x+b-i)^2 & \\ \quad+\binom{2b+1}{i}(2(b-i)^2+b-2i) & i<b\\ \binom{2b+1}{b-1}(x-1)^2-b\binom{2b+1}{b} & i=b.\end{cases}\end{aligned}$$

Here $\delta_{i,b}$ is the Kronecker delta.

The first term of (the last form of) F_i and second term of F_{i-1}, when multiplied by the appropriate double products, cancel. Thus we obtain

$$\begin{aligned}&\tfrac{(2b+2)!}{2}\widetilde{R}_{2b+1}(x+b) \qquad (2.5)\\
=&\sum_{i=0}^{b-1}\left(\prod_{j=1}^{b-i}(x-j)^2\prod_{j=b-i+1}^{b}(x+j)^2\right)\tbinom{2b+1}{i}(2(b-i)^2+b-2i)\\
&-\left(\prod_{j=1}^{b}(x+j)^2\right)b\tbinom{2b+1}{b}.\end{aligned}$$

One can easily prove that

$$\sum_{i=0}^{b-1} \binom{2b+1}{i}(2(b-i)^2+b-2i) - b\binom{2b+1}{b} = 0.$$

Also $2(b-i)^2+b-2i>0$ iff $i<b+\frac{1}{2}-\frac{1}{2}\sqrt{2b+1}$. Note also that the double products are increasing with i for $x>0$. Thus our expression for $\frac{(2b+2)!}{2}\widetilde{R}_{2b+1}(x+b)$ is of the form $\sum\limits_{i=0}^{b}\alpha_i\beta_i$ with $0\le\alpha_1\le\cdots<\alpha_b$, $\sum\beta_i=0$, and $\beta_i>0$ iff $i<i_0$. Such a sum is negative.

The proof for (2) when $\epsilon=0$ is extremely similar. We have $q_{2b}(x+b)=\frac{2}{2b+1}x^2+\frac{b}{2b+1}$. Then

$$\tfrac{(2b+1)!}{2}\widetilde{R}_{2b}(x+b) = \sum_{i=0}^{b}\prod_{j=1}^{b-i}(x-j+1)^2\prod_{j=b-i+1}^{b}(x+j)^2\cdot F_i,$$

where now

$$\begin{aligned} F_i &= \binom{2b+1}{i}(x-b+i)^2 - c_{i,b}\binom{2b}{i}(2x^2+b) \\ &= \begin{cases} \binom{2b}{i-1}(x-(b-i))^2 - \binom{2b}{i}(x+b-i)^2 + \binom{2b}{i}(2(b-i)^2-b) & i<b \\ \binom{2b}{b-1}x^2 - b\binom{2b-1}{b-1} & i=b. \end{cases} \end{aligned}$$

The rest of the argument follows exactly the same steps as in the last paragraph of the above proof of the case $\epsilon=1$, using $\sum_{i=0}^{b-1}\binom{2b}{i}(2(b-i)^2-b)-b\binom{2b-1}{b-1}=0$.

The proof of (3) is essentially the same, except that we are subtracting 1/2 from $q_{2b+\epsilon}(x+b)$. The effect, when $\epsilon=1$, is to add $(b+1)\binom{2b+1}{i}$ to F_i. The replacement for (2.5) is

$$\begin{aligned} &\tfrac{(2b+2)!}{2}(\widetilde{R}_{2b+1}(x+b)+\tfrac{1}{2}\widetilde{P}_{2b+1}(x+b)) \\ =\ &\sum_{i=0}^{b-1}\left(\prod_{j=1}^{b-i}(x-j)^2\prod_{j=b-i+1}^{b}(x+j)^2\right)\binom{2b+1}{i}(2(b-i)^2+2b-2i+1) \\ &+\left(\prod_{j=1}^{b}(x+j)^2\right)\binom{2b+1}{b} \end{aligned}$$

which is clearly positive for $x>0$. The proof when $\epsilon=0$ is similar. □

3 Conjectures about $R_d(x)/P_d(x)$

In Theorem 1.2, we determined the precise value of our focal function $g(s_1,s_2)$ in terms of $R_d(s_1)/P_d(s_1)$. In order to make this result useful, we need better information about the family of functions $R_d(x)/P_d(x)$. In this section, we present several conjectures about this family of functions, one of which is supported by remarkable patterns. See Table 3.3 and Conjecture 3.2. We also discuss their implications.

We now state the simplest of these conjectures.

Conjecture 3.1. If $d \ge 5$ and $x \ge \frac{1}{2}(d+\sqrt{d+2})$, then $-\frac{R_d(x)}{P_d(x)} \ge \frac{.995(d-2)}{2x+d-2}$.

The implication of this conjecture is given by the following theorem.

Theorem 3.1. *Assume Conjecture 3.1 and* $s_2 > \frac{1}{2}(\sqrt{8s_1+9}-1)$. *Let*

$$d_0 = \left[\sqrt{4s_1^2+4s_1+(s_2+\tfrac{1}{2})^2}-s_2-\tfrac{3}{2}\right] \ge 5.$$

Then (1.1) is true with $\delta = 1$ *if*

$$0 < \tfrac{2}{d_0+1}(s_1^2+s_1) - \tfrac{d_0+2}{2} - s_2 \le \tfrac{0.995(d_0-2)}{2s_1+d_0-2}. \tag{3.1}$$

For $s_1 \le 38$, the only cases in which (1.1) is true with $\delta = 1$ which are missed by this theorem are $(s_1, s_2) = (6,4)$, $(18,56)$, $(36,16)$, and $(38,155)$. The significance of the 0.995 in 3.1 is that it is, to three decimal places, the largest number for which the inequality appears to be true.

Proof of Theorem 3.1. Let $d = \sqrt{4s_1^2+4s_1+(s_2+\frac{1}{2})^2}-s_2-\frac{3}{2}$. The theorem is vacuously true if $d_0 = d$ because $\frac{2}{d+1}(s_1^2+s_1)-\frac{d+2}{2}-s_2 = 0$. So we assume d is not an integer. Then $2s_1-d_0 = [2s_1-d]+1$, and so, using (1.3), the assertion that (1.1) is true with $\delta = 1$ can be stated as

$$s_2 \ge \tfrac{2}{d_0+1}(s_1^2+s_1) - \tfrac{d_0+2}{2} + \tfrac{R_{d_0}(s_1)}{P_{d_0}(s_1)}.$$

This will follow from (3.1) and our assumption of Conjecture 3.1 once we know that $s_1 \ge \frac{1}{2}(d_0+\sqrt{d_0+2})$.

Since $\sqrt{4s_1^2+4s_1+(s_2+\frac{1}{2})^2}-s_2-\frac{3}{2}$ is a decreasing function of s_2, and $d_0 < d$, it suffices to prove $s_1 \ge \frac{1}{2}(d+\sqrt{d+2})$ if

$$\begin{aligned} d &= \sqrt{4s_1^2+4s_1+(\tfrac{1}{2}\sqrt{8s_1+9})^2}-\tfrac{1}{2}\sqrt{8s_1+9}-1 \\ &= 2s_1+\tfrac{1}{2}-\sqrt{2s_1+\tfrac{9}{4}}. \end{aligned}$$

Solving the latter equation for s_1 yields exactly $s_1 = \frac{1}{2}(d+\sqrt{d+2})$. □

Extensive `Maple` calculation led the author to expect that, for $d \ge 5$ and $x > d/2$,

$$\frac{R_d(x)}{P_d(x)} \approx \frac{-(d-2)}{2x+d-2}. \tag{3.2}$$

To understand how good is the approximation (3.2), we consider the ratio of the two sides, using the reduced versions $\widetilde{R}$ and $\widetilde{P}$. As this will be close to 1, we study

$$\begin{aligned} Q_d(x) &:= 1 - \frac{\widetilde{R}_d(x)/\widetilde{P}_d(x)}{-(d-2)/(2x+d-2)} \\ &= \frac{(d-2)\widetilde{P}_d(x) + (\widetilde{P}_{d+1}(x) - q_d(x)\widetilde{P}_d(x))(2x+d-2)}{(d-2)\widetilde{P}_d(x)}, \end{aligned} \tag{3.3}$$

where $q_d(x) = \frac{2}{d+1}x^2 - \frac{2d}{d+1}x + \frac{d}{2}$ is the quotient in (2.4). We would like to prove that $Q_d(x) \approx 0$ in some sense, when $x > d/2$.

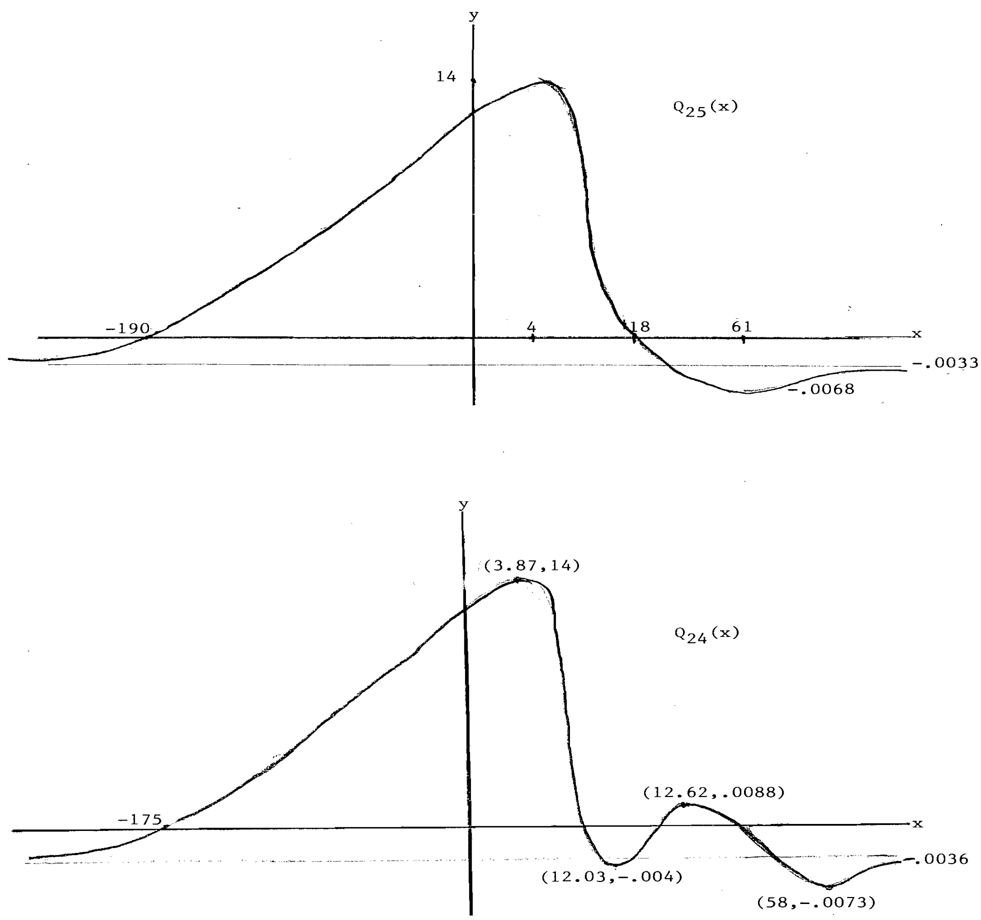
Similarly to the methods in deriving (2.2), we can show that

$$\lim_{x\to\pm\infty} Q_d(x) = \tfrac{-2}{(d+1)(d-2)}. \tag{3.4}$$

To see this, note that the numerator and denominator of (3.3) are both polynomials of degree $2[\frac{d}{2}]$. The desired limit in (3.4) is the ratio of their leading coefficients. We omit the details in this computation.

We begin by considering $Q_{25}(x)$. It is a ratio of two polynomials of degree 24. `Maple` computes that the derivative of $Q_{25}(x)$ is 0 only at $x \approx 4.0409$ and 60.50336. Moreover, `Maple` plots the graph of $Q_{25}(x)$, which turns out to look something like the rough sketch in the top half of Figure 3.1. This sketch is not at all to scale. We are particularly interested in the values for $x \geq 13$. As x increases from 13 to 60.5, $Q_{25}(x)$ decreases from 0.01672 to -0.006867.

Figure 3.1.



The amazing observation is that, for all odd d, $Q_d(x)$ apparently has a form very similar to that in the top half of Figure 3.1, with only one local maximum and one local minimum, which are the absolute maximum and minimum. Note that $Q_d(x)$ is a ratio of two polynomials of degree $d-1$, yet it apparently has this simple form for all odd d.

In Table 3.3, we tabulate for odd d satisfying $5 \le d \le 61$, the values, x_{max} and x_{min}, of x where the derivative $Q'_d(x)$ equals 0, the values of $Q_d(x)$ at these points, which will be absolute maximum and minimum values, and the limiting value Q_{lim} of $Q(x)$ as $x \to \pm\infty$.

Table 3.3: Max, min, and lim of $Q_d(x)$ when d is odd

d	x_{max}	$Q_d(x_{\text{max}})$	x_{min}	$Q_d(x_{\text{min}})$	Q_{lim}
5	0.6874697648	3.069437405	10.48133802	−0.1360199330	−0.1111111111
7	1.044428415	4.216907295	15.48679761	−0.06528852351	−0.05000000000
9	1.371314912	5.374255987	20.49141428	−0.03966514187	−0.02857142857
11	1.706180162	6.540442715	25.49465519	−0.02724322575	−0.01851851852
13	2.040065994	7.712042960	30.49698936	−0.02018382313	−0.01298701299
15	2.373641870	8.886960397	35.49873537	−0.01574295269	−0.009615384615
17	2.707177063	10.06413875	40.50008573	−0.01274397486	−0.007407407407
19	3.040658416	11.24291343	45.50115931	−0.01060966757	−0.005882352941
21	3.374105921	12.42285407	50.50203244	−0.009028116216	−0.004784688995
23	3.707530371	13.60364584	55.50275603	−0.007817967242	−0.003968253968
25	4.040938032	14.78513590	60.50336525	−0.006867535404	−0.003344481605
27	4.374333149	15.96714209	65.50388509	−0.006104769071	−0.002857142857
29	4.707718646	17.14958715	70.50433380	−0.005481366771	−0.002469135802
31	5.041096602	18.33236558	75.50472499	−0.004963889096	−0.002155172414
33	5.374468538	19.51542724	80.50506904	−0.004528540629	−0.001897533207
35	5.707835592	20.69872494	85.50537395	−0.004157982647	−0.001683501683
37	6.041198635	21.88222143	90.50564603	−0.003839317503	−0.001503759398
39	6.374558339	23.06588687	95.50589030	−0.003562775224	−0.001351351351
41	6.707915237	24.24969671	100.5061108	−0.003320834926	−0.001221001221
43	7.041269754	25.43363246	105.5063109	−0.003107623257	−0.001108647450
45	7.374622231	26.61767652	110.5064932	−0.002918493910	−0.00101112234
47	7.707972951	27.80181575	115.5066600	−0.002749728194	−0.000925925926
49	8.041322143	28.98603869	120.5068132	−0.002598318198	−0.00085106383
51	8.374670000	30.17033570	125.5069545	−0.002461807379	−0.00078492936
53	8.708016684	31.35469854	130.5070851	−0.002338171717	−0.00072621641
55	9.041362332	32.53912015	135.5072062	−0.002225729996	−0.00067385445
57	9.374707059	33.72359445	140.5073188	−0.002123075281	−0.00062695925
59	9.708050964	34.90811616	145.5074238	−0.002029022010	−0.00058479532
61	10.041394134	36.09268071	150.5075219	−0.001942564749	−0.00054674686

When d is even, the functions Q_d fall into almost the same pattern, except that they have an additional wiggle between $\frac{d}{2}$ and $\frac{d}{2}+1$. For example, a schematic graph of $Q_{24}(x)$ is given in the bottom half of Figure 3.1. Note that the graph is drawn wildly out of scale. Similarly to Q_{25}, it has an absolute maximum at $x = 3.87$ and an absolute minimum at $x = 58.003$. But instead of decreasing steadily between these, it has an additional single local minimum and local maximum which occur between $x = 12$ and 13. `Maple` calculations strongly suggest that for all even d, the graph of Q_d will have a form similar to that in the bottom half of Figure 3.1, and that the positions and values of the maxima and minima will have patterns extremely similar to those for odd d in Table 3.3. We will not pursue those here, as we prefer to concentrate on the simpler situation when d is odd.

The reader will immediately be struck by the pattern in Table 3.3, which seems especially striking for $x_{\text{min}}(d)$. We have extended these calculations through $d = 151$, using 80 digits of accuracy in `Maple`. Then for $k = 3, \ldots, 10$, we have found the real numbers $c_0, \ldots, c_k$ which satisfy

$$x_{\text{min}}(d) = \tfrac{5}{2}d - 2 + \sum_{i=0}^{k} \frac{c_i}{(d-1)^i}, \qquad d = 51,\, 61, \ldots, 51 + 10k. \tag{3.5}$$

Each c_i seems to stabilize as k increases in (3.5). Moreover, using the formula (3.5) for $x_{\text{min}}(d)$ derived using just a few values of d gives agreement with computed values of $x_{\text{min}}(d)$ for all odd values of d to an increasing number of decimal places as k increases. In addition, we have, with $k = 10$, c_0 equals, to 19 decimal places, $.0104166666666666666 \approx 1/96$.

Conjecture 3.2. For odd $d \geq 5$, there are numbers x_{max} and x_{min} such that $Q_d(x)$ is decreasing for $x_{\text{max}} \leq x \leq x_{\text{min}}$, and increasing elsewhere. There are real numbers c_i for $i \geq 1$ such that

$$x_{\text{min}}(d) = \tfrac{5}{2}d - 2 + \tfrac{1}{96} + \sum_{i=1}^{\infty} \frac{c_i}{(d-1)^i}.$$

The initial digits of $c_1, \ldots, c_6$ are $-.176504629629629$, $.16562740498$, $.20004439$, $.291872$, $.3215$, and $.28$.

The 3-digit repetend in c_1 leads one to guess that $c_1 = -305/12^3$ and that the c_i are all rational numbers.

We have performed similar analyses for x_{max} and $Q_d(x_{\text{min}})$. The initial terms are apparently $x_{\text{max}} = \frac{d}{6} - \frac{1}{8} - \frac{1}{64(d-1)}$ and $Q_d(x_{\text{min}}) = -\frac{1}{12(d-1)}$. The series for $Q_d(x_{\text{min}})$ seems to converge more slowly than the others.

We wish to emphasize that we cannot prove that $Q_d(x)$ for odd d has a unique maximum and minimum. This is all based on `Maple` calculations obtained by setting its derivative equal to 0, where Q_d is a ratio of two polynomials of degree $d-1$.

Table 3.4: Evidence for Conjecture 3.1

d	$Q_d(\frac{1}{2}(d+\sqrt{d+2}))$
5	$-.08632297$
6	$-.0563567189$
7	$-.031250000$
8	$-.021426047$
9	$-.012806353$
10	$-.008467636$
11	$-.004769088$
12	$-.002577860$
13	$-.000764689$
14	0.0004262575
15	0.001391319
16	0.002066212
17	0.002604694
18	0.002994356
19	0.003300396
20	0.00352431
21	0.00369610
22	0.00382002
23	0.00391095
24	0.00397283
25	0.00401357
26	0.00403620
27	0.004045174
28	0.004042664
29	0.0040313148
30	0.004012644
31	0.003988285

Conjecture 3.1 is equivalent to saying that $Q_d(x) \le .005$ for $x \ge \frac{1}{2}(d+\sqrt{d+2})$. This latter statement would follow from Conjecture 3.2 expanded to include a formula for $Q_d(x_{\min})$ and to include even values of d, together with a proof that $Q_d(\frac{1}{2}(d+\sqrt{d+2})) \le .005$. Some justification for this conjecture is given by Table 3.4, which also shows why we use .005.

There is one value of $Q_d(x)$, occurring just before the crucial range $x > d/2$, for which the value of $Q_d(x)$ is easily determined. This is given in the following result, whose easy proof we omit.

Proposition 3.1. *If d is odd, then $Q_d(\frac{d-1}{2}) = 1$. If d is even, then $Q_d(\frac{d}{2}) = \frac{-2}{(d+1)(d-2)}$.*

4 The general case (s_1 and s_3 not necessarily equal)

In this section, we present our analysis of the general case, which is similar to, but not nearly so thoroughly developed as, the case $s_1 = s_3$ considered in the preceding sections.

For arbitrary s_1, s_2, and s_3, now let

$$f_{s_1,s_2,s_3}(N) := \binom{s_2}{N}\sum_i \binom{s_1}{i}\binom{s_3}{N-i}\binom{N}{i}.$$

The argument of Walkup mentioned in the introduction implies that this f_{s_1,s_2,s_3} is a unimodal function of N. We can find the value of N at which f achieves a maximum by an analysis extremely similar to that employed in the case $s_1 = s_3$.

The formula for f is symmetric in s_1 and s_3. We write $s_1 = x$ and $s_3 = x+\Delta$, $\Delta \geq 0$. Let $d = s_1 + s_3 - N$, and

$$P_{\Delta,d}(x) = \sum j!(d-j)!\binom{x+\Delta}{j}^2\binom{x}{d-j}^2 = \sum \frac{((x+\Delta)_j)^2(x_{d-j})^2}{j!(d-j)!}.$$

Generalizing (2.4), which is the case $\Delta = 0$, we have

$$\frac{P_{\Delta,d+1}}{P_{\Delta,d}} = \frac{2x^2 - 2(d-\Delta)x + \frac{1}{2}d(d+1-2\Delta) + \Delta^2}{d+1} + \frac{R_{\Delta,d}(x)}{P_{\Delta,d}(x)}. \tag{4.1}$$

The easy generalization of (1.3) is

$$\begin{aligned} f_{x,s_2,x+\Delta}(2x+\Delta-d) &\geq f_{x,s_2,x+\Delta}(2x+\Delta-d-1) \\ &\text{iff} \\ s_2 &\geq \tfrac{2x^2+2(\Delta+1)x+\Delta^2+\Delta}{d+1} - \tfrac{d+2}{2} + \tfrac{R_{\Delta,d}(x)}{P_{\Delta,d}(x)}, \end{aligned} \tag{4.2}$$

where $R_{\Delta,d}(x)$ is the remainder in (4.1). If we assume this remainder is negligible, then imposing equality in (4.2) and recalling $s_1 = x$ and $s_3 = x + \Delta$ yields

$$d = \sqrt{2s_1^2 + 2s_1 + 2s_3^2 + 2s_3 + (s_2 + \tfrac{1}{2})^2} - s_2 - \tfrac{3}{2}$$

and

$$N = s_1 + s_2 + s_3 + \tfrac{3}{2} - \sqrt{2s_1^2 + 2s_1 + 2s_3^2 + 2s_3 + (s_2 + \tfrac{1}{2})^2}, \tag{4.3}$$

as nice a generalization of 1.2 and (1.1) as one could possibly desire. This yields

$$\sqrt{2s_1^2 + 2s_1 + 2s_3^2 + 2s_3 + (s_2 + \tfrac{1}{2})^2} - \tfrac{3}{2}$$

as the most likely number of elements in the union, assuming remainder terms are negligible. More analysis of the remainder terms is required.

We have seen that when $s_1 = s_3$, the remainder terms can apparently only affect the value of N by 1. In Table 4.5 we present data when $s_3 = s_1 + 8$, indicating rather good agreement. Here "actual N" is where the maximum actually occurs.

Table 4.5: Comparison of actual N and formula N

s_1	s_3	s_2	actual N	(4.3)
4	12	4	4	2.20621862
4	12	5	4	2.94878520
4	12	6	5	3.64427035
4	12	7	6	4.29480265
4	12	8	6	4.90267008
4	12	9	7	5.47025916
4	12	10	7	6.00000000
4	12	11	7	6.49431892
4	12	12	8	6.95559936
4	12	13	8	7.38615134
4	12	14	9	7.78818860
4	12	15	9	8.16381295
4	12	16	9	8.51500450
4	12	17	9	8.84361678
4	12	18	10	9.15137575
4	12	19	10	9.43988174
4	12	20	10	9.71061354
12	20	4	4	3.26186337
12	20	5	5	4.11613751
12	20	6	6	4.94207761
12	20	7	6	5.74010932
12	20	8	7	6.51071593
12	20	9	8	7.25443290
12	20	10	9	7.97184215
12	20	11	9	8.66356603
12	20	12	10	9.33026127
12	20	13	11	9.97261301
12	20	14	11	10.59132893
12	20	15	12	11.18713359
12	20	16	12	11.76076312
12	20	17	13	12.31296031
12	20	18	13	12.84446999
12	20	19	14	13.35603495
12	20	20	14	13.84839221

References

[1] D. W. Walkup, *Pólya sequences, binomial convolution and the union of random sets*, J. Appl. Probability **13** (1976), 76–85.

Donald M. Davis: Department of Mathematics, Lehigh University, Bethlehem, PA 18015, USA

E-mail address: dmd1@lehigh.edu