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 $\sum_{i} \binom{s_1}{i} \binom{s_2}{N-i} \binom{s_3}{N} \binom{s_4}{i}$ 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 $\hat{A}$, $\hat{B}$, and $\hat{C}$ are disjoint sets with $|\hat{A}| = s_1$, $|\hat{B}| = s_2$, and $|\hat{C}| = s_3$. The sample space consists of all 4-tuples $(A_1, B_1, B_2, C_2) \subset (\hat{A}, \hat{B}, \hat{B}, \hat{C})$ 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

$$|E_{i,j}| = \binom{n_1}{i} \binom{n_2}{j} \binom{n_3}{s_1 - i} \binom{n_3}{s_2}.$$ 

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

$$\sum_i \binom{n_1}{i} \binom{n_2}{s_1 - i} \binom{n_3}{s_2} \binom{n_3}{N - i}.$$ 

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 $f_{s_1,s_2}(N) := \binom{N}{s_1} \sum_i \binom{n_1}{i} \binom{n_2}{s_1 - i} \binom{n_3}{N - i}$ 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 $N \leq g(s_1, s_2)$, and a decreasing function of $N$ for $N \geq g(s_1, s_2)$. Moreover,

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

(1.1)

with $\delta = 0$ or 1.

**Corollary 1.1.** The maximum value of $\sum_i \binom{n_1}{i} \binom{n_2}{s_1 - i} \binom{n_3}{s_2} \binom{n_3}{N - i}$ occurs when

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

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

$$\sqrt{4s_1^2 + 4s_1 + \left(s_2 + \frac{1}{2}\right)^2} - \frac{3}{2} - \delta.$$ 

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 $g(s_1, s_2)$. 


Maximizing a combinatorial expression

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.**

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

\[ b. \text{ For all } s_1, s_2, \text{ we have } g(s_1,s_2) \leq 2s_1. \text{ For } 0 \leq d \leq 4, \]

\[ g(s_1,s_2) \geq 2s_1 - d \text{ iff } s_2 \geq \frac{2}{\pi} (s_1^2 + s_1) - \frac{d+2}{2^2}. \]  

(1.2)
Table 1.1: Values of $g(15, s_2)$

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

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

<table>
<thead>
<tr>
<th>$s_2$</th>
<th>$g(15, s_2)$</th>
<th>$2s_1 + s_2 + \frac{3}{2} - \sqrt{4s_1^2 + 4s_1 + (s_2 + \frac{1}{2})^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6</td>
<td>5.84</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>8.78</td>
</tr>
<tr>
<td>15</td>
<td>12</td>
<td>11.85</td>
</tr>
<tr>
<td>17</td>
<td>13</td>
<td>12.91</td>
</tr>
<tr>
<td>19</td>
<td>14</td>
<td>13.89</td>
</tr>
</tbody>
</table>

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
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,

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

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

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

where $d_1$ satisfies

$$s_2 = \frac{2}{d_1+1} (s_1^2 + s_1) - \frac{d_1+2}{2}. \quad (1.5)$$

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

$$s_2 = \frac{2}{d_2+1} (s_2^2 + s_1) - \frac{d_2+2}{2} \geq \frac{2}{d_2+1} (s_1^2 + s_1) - \frac{d_2+2}{2} \geq \frac{2}{d_2+1} (s_1^2 + s_1) - \frac{d_2+2}{2} + \frac{R_d(s_1)}{P_d(s_1)}. \quad (1.6)$$

Therefore $g(s_1, s_2) \geq 2s_1 - d_2$. We will now show that the RHS of (1.1) 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 $h(d) = s_2 - \left( \frac{2}{d_2+1} (s_1^2 + s_1) - \frac{d_2+2}{2} \right)$. 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,

$$s_2 - \left( \frac{2}{d_2+1} (s_1^2 + s_1) - \frac{d_2+2}{2} + \frac{R_d(s_1)}{P_d(s_1)} \right) < 0, \quad \text{as desired.} \quad \square$$

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
\[ d_0 = \left[ \sqrt{4s_1^2 + 4s_1 + (s_2 + \frac{1}{2})^2} - s_2 - \frac{3}{2} \right]. \]
Then (1.1) is true with \( \delta = 1 \) iff
\[ 0 < \frac{2}{d_{i+1}}(s_1^2 + s_1) - \frac{d_i + 2}{2} - s_2 \leq -\frac{R_{d_0}(s_1)}{R_{s_0}(s_2)}. \] (1.5)

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

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 \( \leq \)-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
\[
(s_2 - 2x + d + 1) \sum_{i=0}^{(x!)} (x!)(2x-d)! \frac{(x!)(2x-d)!}{(d+i-x)!} 
\geq (2x-d) \sum_{i=0}^{(x!)} (x!)(2x-d-1)! \frac{(x!)(2x-d-1)!}{(d+i-x)!}.
\]
Cancelling \((2x-d)!\) and letting \( j = x-i \), we obtain the equivalent condition
\[
(s_2 - 2x + d + 1) \sum_{j=0}^{(x!)} (x - j)^2 (x + j - d)^2 j! (d - j)! 
\geq \sum_{j=0}^{(x!)} (x - j)^2 (x + j - d - 1)^2 j! (d - j)!.
\]
Multiplying both sides by \((x!)^2\), the inequality becomes
\[
(s_2 - 2x + d + 1) \sum_{j=0}^{(x!)} \frac{(x-j)^2 (x+j-d)^2}{j! (d-j)!} \geq \sum_{j=0}^{(x!)} \frac{(x-j)^2 (x+j-1-d)^2}{j! (d+1-j)!},
\]
and this readily yields
\[
f_{x,s_2}(2x-d) \geq f_{x,s_2}(2x-d-1) \text{ iff } s_2 \geq \frac{P_{d+1}(x)}{P_d(x)} + 2x - d - 1. \] (2.1)
Our next claim is that the leading terms of $P_d(x)$ are given by

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

The proof of (2.2) makes frequent use of

$$\sum_i \frac{1}{i(d-i)!} = \frac{x^d}{d!}, \quad (2.3)$$

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

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

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

The next coefficient of $P_d(x)$ is obtained similarly, but involves much more work. The coefficient of $x^{2d-2}$ in $(x_1)^2(x_{d-1})^2$ is

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

Noting that

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

we obtain

$$-\sum_{j=0}^{i-1} j^2 - \sum_{j=0}^{d-i-1} j^2 + (i-1)i(d-i-1)(d-i) + 2 \left( \frac{(i-1)i}{2} \right)^2 + 2 \left( \frac{(d-i-1)(d-i)}{2} \right)^2$$

$$= -\left( \frac{(i-1)i(2i-1)}{6} \right) - \left( \frac{(d-i-1)(d-i)(2d-2i-1)}{6} \right) + (i-1)i(d-i-1)(d-i)$$

$$+ \frac{(i-1)^2}{2} + \frac{(d-i)^2(d-i-1)^2}{2}.$$

Thus, using symmetry, the coefficient of $x^{2d-2}$ in $\sum_{i=0}^d \frac{1}{i!d!(d-i)!}$ is

$$\sum_{i=0}^d \frac{\binom{i}{2}(i-1)}{i!d!(d-i)!} + \sum_{i=0}^d \frac{1}{i!d!(d-i-1)!} - \frac{1}{3} \sum_{i=0}^d \frac{2i-1}{i!d!(d-i)!}.$$
Next note that
\[
\sum_{i=2}^{n} \frac{(i-1)(i+2)}{(d-2)! (d-3)!} = \sum_{i=1}^{d-3} \frac{(j+1)(j+2)}{(d-2-j)!} \\
= 2 \frac{2^{d-2}}{(d-2)!} + 3 \frac{2^{d-3}}{(d-3)!} + \sum_{j=1}^{d-4} \frac{(j-1)(d-2-j)!}{(d-2-j)!} \\
= \frac{2^{d-4}}{(d-2)!} + 3 \frac{2^{d-3}}{(d-3)!} + \frac{(d-3)!}{(d-4)!} \\
= \frac{2^{d-4}}{(d-2)!} (8 + 8(d-2) + (d-2)(d-3)) \\
= \frac{2^{d-4}}{(d-2)!} (d^2 + 2d - 2).
\]

Also,
\[
\sum_{i=2}^{n} \frac{2i-1}{(i-2)(i-3)!} = 2 \sum_{i=3}^{d-3} \frac{1}{(i-3)(i-2)!} + 3 \sum_{i=2}^{d-4} \frac{1}{(i-2)(i-3)!} = \frac{2^{d-2}}{(d-2)!} (d - 2 + 3).
\]

Hence the desired coefficient equals
\[
\frac{2^{d-4}}{(d-2)!} (d^2 + 3d - 2) + \frac{2^{d-4}}{(d-3)!} \frac{2^{d-2}}{3(d-2)!} (d + 1) \\
= \frac{2^{d-3}}{3(d-2)!} (3d^2 - 5d + 4),
\]
as asserted in (2.2).

Next we claim that
\[
P_{d+1}(x) = P_d(x) \frac{2^d x^{2d} - \frac{2^d}{d+1} x^{2d+1} + \frac{2^d}{d(d-1)} (3d^2 - 5d + 4) x^{2d+2}}{P_d(x)}, \tag{2.4}
\]
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
\[
\left( \frac{2^d}{d+1} x^{2d} - \frac{2^d}{d(d-1)} x^{2d+1} + \frac{2^d}{3(d-2)!} (3d^2 - 5d + 4) x^{2d+2} \right) \left( \frac{2^d}{d+1} x^{2d} - \frac{2^d}{d(d-1)} x^{2d+1} + \frac{2^d}{3(d-2)!} (3d^2 - 5d + 4) x^{2d+2} \right) = \frac{2^{d+1}}{(d+1)!} x^{2d+2} - \frac{2^d}{(d-1)!} x^{2d+1} + \frac{2^d}{3(d-1)!} (3(d+1)^2 - 5(d+1) + 4) x^{2d} + \text{lower}.
\]

Combining (2.1) and (2.4), we obtain (1.3). \(\square\)

**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 \(\geq\) the bracketed expression of (1.1) by Theorem 1.1, and is \(\leq s_2\) 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

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

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

\[ s_2 \geq \frac{2}{3}(x^2 + x) - 2 + \frac{R_2(x)}{P_2(x)} \iff s_2 \geq \frac{2}{3}(x^2 + x) - 2, \]

and so (1.2) follows from (1.3).

A similar argument works for \( d = 3 \) and \( 4 \). For \( d = 3 \), we have that \( \frac{2}{3} (x^2 + x) - \frac{d+2}{2} \) 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 \( \frac{2}{3} (x^2 + x) - \frac{d+2}{2} \) 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.

**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 \( \delta \in \{0, 1\} \),

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

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

To prove the lemma, we will prove

1. \( \tilde{P}_{2b+\epsilon}(x+b) > 0 \) for \( x > 0 \),
2. \( \tilde{R}_{2b+\epsilon}(x+b) < 0 \) for \( x > 0 \), and
3. \( \tilde{R}_{2b+\epsilon}(x+b) + \frac{1}{2} \tilde{P}_{2b+\epsilon}(x+b) > 0 \) for \( x > 0 \).
We have, with \(c_{i,b} = 1\) unless \(i = b\), while \(c_{b,b} = \frac{1}{2}\),

\[
\frac{(2b+\epsilon+\delta)!}{2} \tilde{P}_{2b+\epsilon+\delta}(x+b) = \sum_{i=0}^{b} c_{i,b}(2b) \prod_{j=1}^{b-i} (x-j+1)^2 \prod_{j=b-i+1}^{b} (x+j)^2 \epsilon + \delta = 0
\]

\[
= \sum_{i=0}^{b} (2b+1) \prod_{j=0}^{b-i} (x-j)^2 \prod_{j=b-i+1}^{b} (x+j)^2 \epsilon + \delta = 1
\]

\[
= \sum_{i=0}^{b} c_{i,b+1}(2b+2) \prod_{j=0}^{b-i} (x-j-1)^2 \prod_{j=b-i+1}^{b} (x+j)^2 \epsilon + \delta = 2.
\]

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

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

\[
\frac{(2b+2)!}{2} \tilde{R}_{2b+1}(x+b) = \frac{(2b+2)!}{2}(\tilde{P}_{2b+2}(x+b) - q_{2b+1}(x+b)\tilde{P}_{2b+1}(x+b))
\]

\[
= \sum_{i=0}^{b} \left( \prod_{j=1}^{b-i} (x-j)^2 \prod_{j=b-i+1}^{b} (x+j)^2 \right) F_{i},
\]

where

\[
F_{i} = \binom{2b+2}{i}(x-b+i-1)^2 - (2b+2) \binom{2b+1}{i} \left( \frac{1}{x_{2b}^{2}}(x^2 - x) + \frac{1}{2} \right) + \frac{1}{2} \binom{2b+2}{i} \delta_{i,b} x^2
\]

\[
= \left( \binom{2b+1}{i-1} - \binom{2b+1}{i} + \delta_{i,b} \binom{2b+1}{i} \right) x^2 - 2 \left( \binom{2b+1}{i-1} + (b-i) \binom{2b+1}{i} \right) x
\]

\[
+ \binom{2b+2}{i} (b + 1 - i)^2 - (b + 1) \binom{2b+1}{i}
\]

\[
\begin{cases} 
\binom{2b+1}{i} (x-(b+1-i))^2 - \binom{2b+1}{i} (x+b-i)^2 & \text{if } i < b \\
\binom{2b+1}{b-i} (x-1)^2 - b \binom{2b+1}{b+1} & \text{if } i = b.
\end{cases}
\]

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

\[
\frac{(2b+2)!}{2} \tilde{R}_{2b+1}(x+b) = \sum_{i=0}^{b} \left( \prod_{j=1}^{b-i} (x-j)^2 \prod_{j=b-i+1}^{b} (x+j)^2 \right) \binom{2b+1}{i} (2(b-i)^2 - b - 2i)
\]

\[
- \left( \prod_{j=1}^{b} (x+j)^2 \right) b \binom{2b+1}{b+1}.
\]
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 \(\frac{(2b+2)!}{2}\overline{R}_{2b+1}(x + b)\) is of the form \(\sum \alpha_i \beta_i\) with \(0 \leq \alpha_1 \leq \cdots < \alpha_b\), \(\sum \beta_i = 0\), 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
\[ \frac{(2b+1)!}{2}\overline{R}_{2b}(x + b) = \sum_{i=0}^{b} \prod_{j=1}^{b-i} (x - j + 1)^2 \prod_{j=b-i+1}^{b} (x + j)^2 : F_i, \]

where now
\[ F_i = \begin{cases} \binom{2b+1}{i}(x - b + i)^2 - c_{i,b}\binom{2b}{i}(2x^2 + b) & \text{if } i < b \\ \left(\binom{2b}{i-1}x^2 - b\binom{2b-1}{b-1}\right) & \text{if } i = b. \end{cases} \]

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 \(q_{2b+1}(x+b)\). The effect, when \(\epsilon = 1\), is to add \((b+1)\binom{2b+1}{b}\) to \(F_i\). The replacement for (2.5) is
\[ \frac{(2b+2)!}{2}\overline{R}_{2b+1}(x + b) + \frac{1}{2} \overline{R}_{2b+1}(x + b) \]
\[ = \sum_{i=0}^{b-1} \left( \prod_{j=1}^{b-i} (x - j)^2 \prod_{j=b-i+1}^{b} (x + j)^2 \right) \binom{2b+1}{i}(2(b-i)^2 + 2b - 2i + 1) \]
\[ + \left( \prod_{j=1}^{b} (x + j)^2 \right) \binom{2b+1}{b} \]

which is clearly positive for \(x > 0\). The proof when \(\epsilon = 0\) is similar. \(\square\)

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 \( d \geq 5 \) and \( x \geq \frac{1}{2} (d + \sqrt{d + 2}) \), then \(- \frac{R_d(x)}{P_d(x)} \geq \frac{0.995(d - 2)}{2x + d - 2}\).

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

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

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

\[
0 < \frac{2}{d_0 + 1} (s_1^2 + s_1) - \frac{d_0 + 2}{2} - s_2 \leq \frac{0.995(d_0 - 2)}{2s_1 + d_0 - 2}.
\] (3.1)

For \( s_1 \leq 38 \), 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 \( \frac{2}{d_0}(s_1^2 + s_1) - \frac{d_0 + 2}{2} - s_2 = 0 \). 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

\[
s_2 \geq \frac{2}{d_0 + 1} (s_1^2 + s_1) - \frac{d_0 + 2}{2} + \frac{R_{d_0}(s_1)}{P_{s_0}(s_1)}.
\]

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

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 \( s_1 \geq \frac{1}{2} (d + \sqrt{d + 2}) \) if

\[
da = \sqrt{4s_1^2 + 4s_1 + (\frac{1}{2} \sqrt{8s_1 + 9})^2} - \frac{1}{2} \sqrt{8s_1 + 9} - 1
\]

\[
= 2s_1 + \frac{1}{2} - \sqrt{2s_1 + \frac{9}{4}}.
\]

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

Extensive Maple calculation led the author to expect that, for \( d \geq 5 \) and \( x > d/2 \),

\[
\frac{R_d(x)}{P_d(x)} \approx \frac{-(d - 2)}{2x + d - 2}.
\] (3.2)
To understand how good is the approximation (3.2), we consider the ratio of the two sides, using the reduced versions $\tilde{R}$ and $\tilde{P}$. As this will be close to 1, we study

$$Q_d(x) := 1 - \frac{\tilde{R}_d(x)/\tilde{P}_d(x)}{-(d - 2)/(2x + d - 2)} = \frac{(d - 2)\tilde{P}_d(x) + (\tilde{P}_{d+1}(x) - q_d(x)\tilde{P}_d(x))(2x + d - 2)}{(d - 2)\tilde{P}_d(x)}, \quad (3.3)$$

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

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

To see this, note that the numerator and denominator of (3.3) are both polynomials of degree $2\lceil\frac{d}{2}\rceil$. 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 $-0.00687$. 
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 $5 \leq d \leq 61$, the values, $x_{\text{max}}$ and $x_{\text{min}}$, 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 $Q_{\text{lim}}$ of $Q(x)$ as $x \to \pm \infty$. 

Figure 3.1.
Maximizing a combinatorial expression

<table>
<thead>
<tr>
<th>d</th>
<th>$x_{\text{max}}$</th>
<th>$Q_d(x_{\text{max}})$</th>
<th>$x_{\text{min}}$</th>
<th>$Q_d(x_{\text{min}})$</th>
<th>$Q_{\lim}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.6874697648</td>
<td>3.069437405</td>
<td>10.48133802</td>
<td>-0.1360199330</td>
<td>-0.111111111</td>
</tr>
<tr>
<td>7</td>
<td>1.044428415</td>
<td>4.21607295</td>
<td>15.48679761</td>
<td>-0.06528852351</td>
<td>-0.050000000</td>
</tr>
<tr>
<td>9</td>
<td>1.371341912</td>
<td>5.374255987</td>
<td>20.49141428</td>
<td>-0.03966514187</td>
<td>-0.02857142857</td>
</tr>
<tr>
<td>11</td>
<td>1.706180162</td>
<td>6.540442715</td>
<td>25.49465519</td>
<td>-0.02724232575</td>
<td>-0.01851851852</td>
</tr>
<tr>
<td>13</td>
<td>2.040065994</td>
<td>7.712042960</td>
<td>30.49698936</td>
<td>-0.02018382313</td>
<td>-0.01298701299</td>
</tr>
<tr>
<td>15</td>
<td>2.373641870</td>
<td>8.886960397</td>
<td>35.49873537</td>
<td>-0.01574295269</td>
<td>-0.009615384615</td>
</tr>
<tr>
<td>17</td>
<td>2.707177063</td>
<td>10.06413875</td>
<td>40.50008573</td>
<td>-0.01274397486</td>
<td>-0.007407407407</td>
</tr>
<tr>
<td>19</td>
<td>3.040658416</td>
<td>11.24291343</td>
<td>45.50115931</td>
<td>-0.01060966757</td>
<td>-0.005882532941</td>
</tr>
<tr>
<td>21</td>
<td>3.374105921</td>
<td>12.42285407</td>
<td>50.50203244</td>
<td>-0.009028116216</td>
<td>-0.004784688995</td>
</tr>
<tr>
<td>23</td>
<td>3.707530371</td>
<td>13.60364584</td>
<td>55.50275603</td>
<td>-0.00781796724</td>
<td>-0.003968253968</td>
</tr>
<tr>
<td>25</td>
<td>4.040938032</td>
<td>14.78513590</td>
<td>60.50336525</td>
<td>-0.006867535404</td>
<td>-0.003344481605</td>
</tr>
<tr>
<td>27</td>
<td>4.374333149</td>
<td>15.96714209</td>
<td>65.50388509</td>
<td>-0.006104769071</td>
<td>-0.002857142857</td>
</tr>
<tr>
<td>29</td>
<td>4.707718646</td>
<td>17.14958715</td>
<td>70.50433380</td>
<td>-0.00548136677</td>
<td>-0.002469135802</td>
</tr>
<tr>
<td>31</td>
<td>5.041096602</td>
<td>18.33236558</td>
<td>75.50472499</td>
<td>-0.004963889096</td>
<td>-0.002155712414</td>
</tr>
<tr>
<td>33</td>
<td>5.37468538</td>
<td>19.51542724</td>
<td>80.50506904</td>
<td>-0.004528540629</td>
<td>-0.001897533207</td>
</tr>
<tr>
<td>35</td>
<td>5.707835952</td>
<td>20.69872494</td>
<td>85.50537395</td>
<td>-0.004157982647</td>
<td>-0.001683501683</td>
</tr>
<tr>
<td>37</td>
<td>6.041198635</td>
<td>21.8822143</td>
<td>90.50563460</td>
<td>-0.003839317503</td>
<td>-0.001503759398</td>
</tr>
<tr>
<td>39</td>
<td>6.374583339</td>
<td>23.06588687</td>
<td>95.50589030</td>
<td>-0.003562775224</td>
<td>-0.001351351351</td>
</tr>
<tr>
<td>41</td>
<td>6.707915237</td>
<td>24.24969671</td>
<td>100.5061108</td>
<td>-0.00320834926</td>
<td>-0.001221001221</td>
</tr>
<tr>
<td>43</td>
<td>7.041269754</td>
<td>25.43363246</td>
<td>105.5063109</td>
<td>-0.003107623257</td>
<td>-0.001108647450</td>
</tr>
<tr>
<td>45</td>
<td>7.374622231</td>
<td>26.61767652</td>
<td>110.5064932</td>
<td>-0.002919849930</td>
<td>-0.00101112234</td>
</tr>
<tr>
<td>47</td>
<td>7.707972951</td>
<td>27.80181575</td>
<td>115.5066600</td>
<td>-0.002749728194</td>
<td>-0.000925925926</td>
</tr>
<tr>
<td>49</td>
<td>8.041322143</td>
<td>28.98603869</td>
<td>120.5068132</td>
<td>-0.002598313198</td>
<td>-0.00085106383</td>
</tr>
<tr>
<td>51</td>
<td>8.374670000</td>
<td>30.17033570</td>
<td>125.5069545</td>
<td>-0.002461807379</td>
<td>-0.00078492936</td>
</tr>
<tr>
<td>53</td>
<td>8.708016684</td>
<td>31.35469854</td>
<td>130.5070851</td>
<td>-0.00233817177</td>
<td>-0.00072621641</td>
</tr>
<tr>
<td>55</td>
<td>9.041362332</td>
<td>32.53912015</td>
<td>135.5072062</td>
<td>-0.002225729996</td>
<td>-0.00067385445</td>
</tr>
<tr>
<td>57</td>
<td>9.374707059</td>
<td>33.72359445</td>
<td>140.5073188</td>
<td>-0.002123075281</td>
<td>-0.00062695925</td>
</tr>
<tr>
<td>59</td>
<td>9.708050964</td>
<td>34.90811616</td>
<td>145.5074238</td>
<td>-0.002092022010</td>
<td>-0.00058479532</td>
</tr>
<tr>
<td>61</td>
<td>10.041394134</td>
<td>36.09268071</td>
<td>150.5075219</td>
<td>-0.001942564749</td>
<td>-0.0005467486</td>
</tr>
</tbody>
</table>

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

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

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 \( x_{\text{max}} \) and \( x_{\text{min}} \) 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

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

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 \( x_{\text{max}} \) 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 \( d - 1 \).
Conjecture 3.1 is equivalent to saying that $Q_d(x) \leq .005$ for $x \geq \frac{1}{2}(d + \sqrt{d + 2})$. 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 $Q_d(\frac{1}{2}(d + \sqrt{d + 2})) \leq .005$. 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 \( Q_d \left( \frac{d-1}{2} \right) = 1 \). If \( d \) is even, then \( Q_d \left( \frac{d}{2} \right) = \frac{-2}{(d+1)(d-2)} \).

4 The general case \((s_1 \text{ and } s_3 \text{ not necessarily equal})\)

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

\[
 f_{s_1,s_2,s_3}(N) := \left( \frac{s_1}{N} \right) \sum_i \left( \frac{s_2}{s_3} \right) \left( \begin{array}{c} N \vspace{1pt} \end{array} \right) \left( \begin{array}{c} i \vspace{1pt} \end{array} \right).
\]

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

\[
P_{\Delta,d}(x) = \sum j!(d-j)! \left( \begin{array}{c} x+\Delta \vspace{1pt} \end{array} \right) \left( \begin{array}{c} j \vspace{1pt} \end{array} \right)^2 = \sum \left( \frac{(x+\Delta)}{j!(d-j)!} \right)^2.
\]

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

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

The easy generalization of (1.3) is

\[
f_{x,s_2,x+\Delta}(2x + \Delta - d) \geq f_{x,s_2,x+\Delta}(2x + \Delta - d - 1)
\]

iff

\[
s_2 \geq \frac{2x^2 + 2(p+1)x + p^2 + 2\Delta + \frac{d}{2} + \frac{R_{\Delta,d}(x)}{P_{\Delta,d}(x)}}{2d+2}, \tag{4.2}
\]

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

\[
d = \sqrt{2s_1^2 + 2s_1 + 2s_2 + 2s_3 + (s_2 + \frac{1}{2})^2} - s_2 - \frac{3}{2}
\]

and

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

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

\[
\sqrt{2s_1^2 + 2s_1 + 2s_2^2 + 2s_3 + (s_2 + \frac{1}{2})^2} - \frac{3}{2}
\]

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 $N$

<table>
<thead>
<tr>
<th>$s_1$</th>
<th>$s_3$</th>
<th>$s_2$</th>
<th>actual $N$</th>
<th>(4.3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>12</td>
<td>4</td>
<td>4.20621862</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>5</td>
<td>2.94878520</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>6</td>
<td>3.64427035</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>7</td>
<td>4.29482065</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>8</td>
<td>4.90267008</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>9</td>
<td>5.47025916</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>10</td>
<td>6.00000000</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>11</td>
<td>6.49431892</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>12</td>
<td>6.9555936</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>13</td>
<td>7.38615134</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>14</td>
<td>7.78818860</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>15</td>
<td>8.16381295</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>16</td>
<td>8.51500450</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>17</td>
<td>8.84361678</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>18</td>
<td>9.15137575</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>19</td>
<td>9.43988174</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>20</td>
<td>9.71061354</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>4</td>
<td>4.26186337</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>5</td>
<td>4.11613751</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>6</td>
<td>4.94207761</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>7</td>
<td>5.74010932</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>8</td>
<td>6.51071593</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>9</td>
<td>7.25443290</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>10</td>
<td>7.97184215</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>11</td>
<td>8.66356603</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>12</td>
<td>9.33026127</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>13</td>
<td>9.97261301</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>14</td>
<td>10.59132893</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>15</td>
<td>11.18713359</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>16</td>
<td>11.76076312</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>17</td>
<td>12.31296031</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>18</td>
<td>12.84446999</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>19</td>
<td>13.35603495</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>20</td>
<td>13.84839221</td>
<td></td>
</tr>
</tbody>
</table>
References


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

\textit{E-mail address:} dmd1@lehigh.edu