Rounding up integers : A surprising occurrence of π

John Coffey, Cheshire, UK.

January 2014

Question
Show that the limit of this process as $N \to \infty$ is π. Choose a positive integer N and round it up to the next multiple of $N-1$ ($= d_1$, say). Call this p_1. Round p_1 up to the next multiple of $N-2 = d_2$ and call this p_2. Continue decreasing the number d_k by 1 until it has reduced to $d_{N-2} = 2$, and continue rounding up to the next multiple of d_k (except when d_k divides p_{k-1}, in which case set $p_k = p_{k-1}$). A sequence of integers p_k is thereby produced. Call the final one F.
Show that $N^2/F \to \pi$.

I found this stated in a book on mathematical constants by Steven Finch. My friend Mr Jim Pearson of Salford has also looked at this problem independently of myself and noted that if the divisors d_k are reduced by 2 rather than by 1, then N^2/F seems to converge to 2π. More generally, reducing d_k in steps of S causes $N^2/F \to S\pi$. Remarkable!

Answer
I give below my own proof and exploration of the structure of the sequence of rounded-up numbers k for finite values of N. In parallel and independently, the problem has been solved elegantly and more succinctly by Jim Pearson. I present the essence of his answer too.

1 Structure of the rounding-up sequence

It is often helpful to start with numerical examples so I have written a computer program to simulate the rounding-up process. Figure 1 plots the discrepancies $N^2/F - \pi$ against N for all N in the range 10 to 150. Clearly the trend is for N^2/F to get closer to π as N increases, though there is a systematic underestimate. For much higher N the error continues to decrease. So the proposition in the Question looks essentially correct, though perhaps not exact.

Table 1 shows the sequence being generated from the seed $N = 126$ and finishing with $F = d_{124} = 5014$. $N^2/F = 3 \cdot 166$, a modest approximation to π. The columns are as follows: k is an index, d_k and m_k are the two factors which give p_k, the latest value in the sequence, and σ_k notes the amount by which the multiplier m_k is being increased from the row above.

Note the following features as we look down the table:

- The sequence p_k initially consists of a several subsequences in each of which m_k increases with k by a constant value of σ_k. Initially $\sigma_k = 1$. This is because the ratio $N-(k+1) : N-k$ is

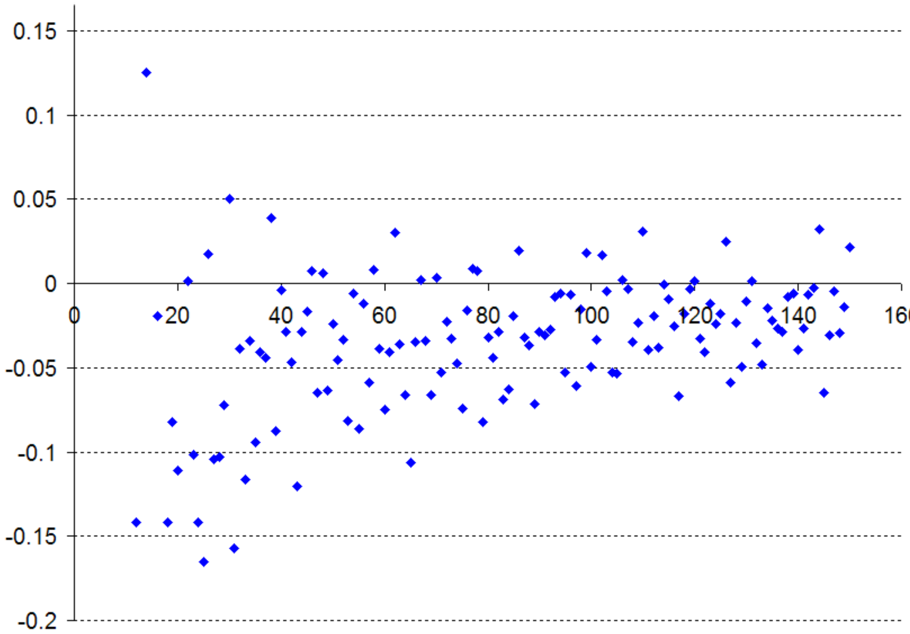

Figure 1: Error in estimating π from N^2/F for all N, $10 \le N \le 150$.

just less than 1 so $p_k/(N-k-1)$ is just greater than m_k. Then the addition of one extra unit of $N-k-1$ takes p_{k+1} well to the right of p_k on the number line. For small k, p_{k+1} is almost twice p_k.

- As k increases, $N-k$ decreases to the point where 2 extra units of $N-k-1$ are needed to take p_{k+1} to the right of p_k.
- Further down the sequence, 3 are required, then 4, *etc.*. Here p_{k+1} is less far to the right of p_k than when k is smaller.
- The subsequences cease to be distinguishable once σ increases from one row to the next, and by increasing greater amounts.
- At the end of the process, at the bottom of Table 1, either $p_k = p_{k-1}$ or $p_k = p_{k-1} + 1$. Here the multiplier m_k increases greatly from $k-1$ to k, inversely as d_k falls from 4 to 3 to 2, to make their product $p_k = d_k m_k$ the final constant, F.

Figure 2 plots the convergence and picks out the first two subsequences. I will divide the plotted points in Figure 2 into three regions as indicated:

1. Region 1 in which subsequences indexed by σ can be clearly distinguished. There must be at least two successive values of k for which σ has the same value.
2. A transition region, starting about 2/3 way across the graph, where the arcs of subsequences can no longer be recognised.
3. The tail of the curve as the final rounded-up value F is reached, at the very right of the figure.

k	$d_k = N - k$	m_k	p_k	σ_k
0	126	1	126	
1	125	2	250	1
2	124	3	372	1
3	123	4	492	1
...				1
62	64	63	4032	1
63	63	64	4032	1
64	62	66	4092	2
65	61	68	4148	2
...				2
79	47	96	4512	2
80	46	99	4554	3
81	45	102	4590	3
...				3
87	39	120	4680	3
88	38	124	4712	4
...				4
92	34	140	4760	4
93	33	145	4785	5
...				5
95	31	155	4805	5
96	30	161	4830	6
97	29	167	4843	6
98	28	173	4844	6
99	27	180	4860	7
100	26	187	4862	7
101	25	195	4875	8
102	24	204	4896	9
103	23	213	4899	9
104	22	223	4906	10
105	21	234	4914	11
106	20	246	4920	12
107	19	259	4921	13
108	18	274	4932	15
109	17	291	4947	17
110	16	310	4960	19
111	15	331	4965	21
112	14	355	4970	24
113	13	383	4979	28
114	12	415	4980	32
115	11	453	4983	38
116	10	499	4990	46
117	9	555	4995	56
118	8	625	5000	70
119	7	715	5005	90
120	6	835	5010	120
121	5	1002	5010	167
122	4	1253	5012	251
123	3	1671	5013	418
124	2	2507	5014	836

Table 1: The calculation for $N = 126$. Dots indicate lines omitted because they follow the same pattern in σ_k.

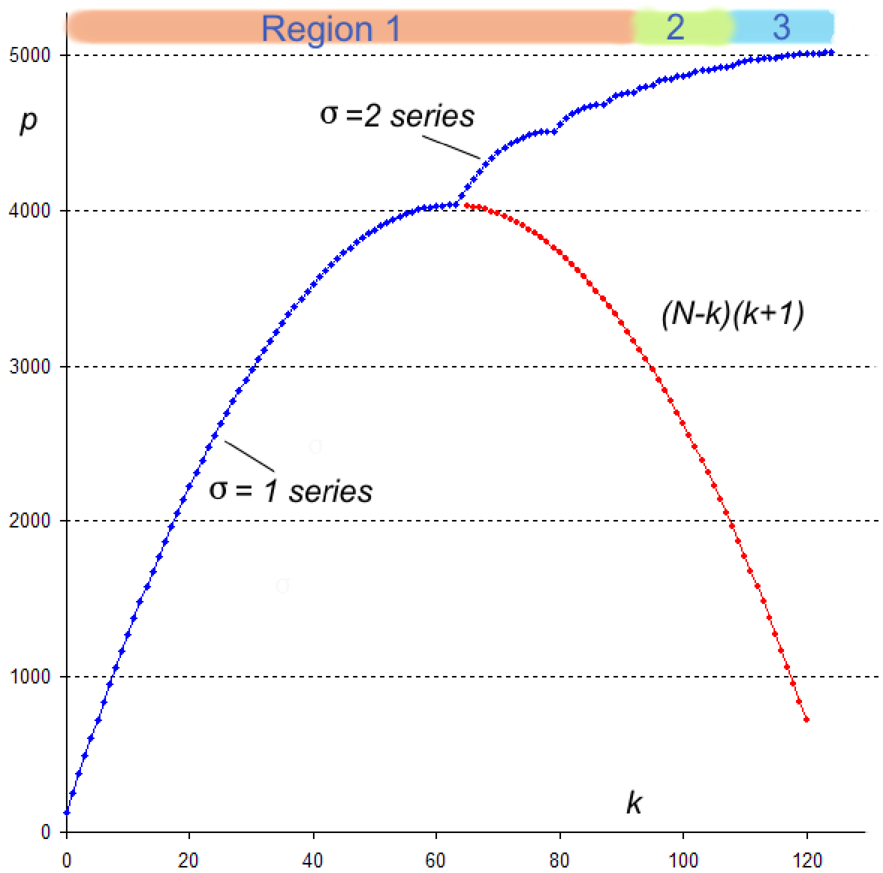

Figure 2: The rounded up number p_k versus k, where $N-k$ is the current divisor, for $N=126$.

2 Region 1 : σ subsequences

In trying to understand the first region it seems worth determining the values of k, d_k, m_k and p_k at the junctions between the subsequences where σ_k jumps in value. I will call these junction values K_σ, D_σ and M_σ respectively. The table below shows the case of $\sigma=1$.

k	$d_k=N-k$	m_k	p_k
0	N	1	N
1	$N-1$	2	$2(N-1)$
2	$N-2$	3	$3(N-2)$
...			...
$k-1$	$N-(k-1)$	k	$k(N-k+1)$
k	$N-k$	$k+1$	$(k+1)(N-k)$

This subsequence can continue so long as $p_k \ge p_{k-1}$ or

$$(k+1)(N-k) \;\ge\; k(N-k+1), \qquad \text{so} \quad k \le \frac{N}{2}.$$

Let $K_1 = N/2$ denote this limiting value of k for the $\sigma=1$ sequence. The largest value of p_k is

$$p_{K_1} \;=\; \frac{N}{2}\left(\frac{N}{2}+1\right).$$

Write $d_{N/2} = D_1$ and $m_{N/2} = M_1$. For $N = 126$, $k = N/2 = 63$ does identify the last row in Table 1 at which $\sigma = 1$, and the value there is $D_1 \times M_1 = 63 \times 64$. Notice how d and m have swapped in values from row $k = 62$ above. The curve of red dots in Figure 2 shows how, if the $\sigma = 1$ series were to continue, it would start reducing in a parabolic arc. Instead, the rounding-up process passes over to $\sigma = 2$.

We will be looking for a pattern in the values at the end of each subsequence, so next consider $\sigma = 2$ which starts with $k = K_1 + 1 = N/2 + 1$.

k	$d_k = N - k$	m_k	p_k
$K_1 + 1$	$N - (K_1 + 1)$	$M_1 + 2$	$\frac{1}{4}(N-2)(N+6)$
$K_1 + 2$	$N - (K_1 + 2)$	$M_1 + 4$	$\frac{1}{4}(N-4)(N+10)$
$K_1 + 3$	$N - (K_1 + 3)$	$M_1 + 6$	$\frac{1}{4}(N-6)(N+14)$
$\dots$			$\dots$
$K_1 + h - 1$	$N - (K_1 + h - 1)$	$M_1 + 2h - 2$	$\frac{1}{4}(N-2h+2)(N+4h-2)$
$K_1 + h$	$N - (K_1 + h)$	$M_1 + 2h$	$\frac{1}{4}(N-2h)(N+4h+2)$

This subsequence can continue so long as

$$(N-2h)(N+4h+2) \;\geq\; (N-2h+2)(N+4h-2), \qquad \text{so} \quad h \leq \frac{N+2}{8},$$

making $k \leq (5N+2)/8$. The largest value in this subsequence, attained at $k = K_2 = (5N+2)/8$, is

$$D_2 \times M_2 \;\equiv\; \tfrac{1}{8}(3N-2) \times \tfrac{3}{4}(N+2) \;=\; \tfrac{3}{32}(3N-2)(N+2)\,.$$

For $N = 126$, $k = K_2 = (5N+2)/8 = 79$ as the last row in Table 1 for which $\sigma = 2$. Here $d = D_2 = 47$ and $m = M_2 = 96$.

Appendix 1 gives details of the calculation for σ from 3 to 7. Table 2 collects the values of k and the two factors D_σ and M_σ of the rounded-up number at the end of each subsequence. These values are exact in the following limited sense: they are correct provided the values of D_σ, M_σ and hence their product are all integers. Once the length of subsequence h from this formula is a fraction, the formulae have broken down and no longer correctly model the discrete nature of the rounding-up process.

Subject to the above proviso, for a general value of σ the length h_σ of the subsequence is given by the condition

$$(D_{\sigma-1} - h)(M_{\sigma-1} + \sigma h) \;=\; (D_{\sigma-1} - h + 1)(M_{\sigma-1} + \sigma(h-1)),$$

$$i.e. \quad h_\sigma \;=\; \tfrac{1}{2}\left(D_{\sigma-1} - \frac{M_{\sigma-1}}{\sigma} + 1\right). \tag{1a}$$

Then

$$K_\sigma = K_{\sigma-1} + h_\sigma, \qquad D_\sigma = D_{\sigma-1} - h_\sigma \qquad M_\sigma = M_{\sigma-1} + \sigma h_\sigma\,. \tag{1b}$$

The initial conditions are $D_0 = N$, $M_0 = 0$ from which $h_1 = (N+1)/2$. The whole sequence in σ can be built from this.

σ	h	K_σ	D_σ	M_σ
1	$\frac{1}{2}(N+2)$	$\frac{1}{2}N$	$\frac{1}{2}N$	$\frac{1}{2}(N+2)$
2	$\frac{1}{8}(N+2)$	$\frac{1}{8}(5N+2)$	$\frac{1}{8}(3N-2)$	$\frac{3}{4}(N+2)$
3	$\frac{1}{16}(N+2)$	$\frac{1}{16}(11N+6)$	$\frac{1}{16}(5N-6)$	$\frac{15}{16}(N+2)$
4	$\frac{5}{128}(N+2)$	$\frac{1}{128}(93N+58)$	$\frac{1}{128}(35N-58)$	$\frac{35}{32}(N+2)$
5	$\frac{7}{256}(N+2)$	$\frac{1}{256}(193N+130)$	$\frac{1}{256}(63N-130)$	$\frac{315}{256}(N+2)$
6	$\frac{21}{1024}(N+2)$	$\frac{1}{1024}(793N+562)$	$\frac{1}{1024}(231N-562)$	$\frac{693}{512}(N+2)$
7	$\frac{33}{2048}(N+2)$	$\frac{1}{2048}(1619N+1190)$	$\frac{1}{2048}(429N-1190)$	$\frac{3003}{2048}(N+2)$

Table 2: Summary of exact integer values for the subsequences $\sigma = 1$ to 7.

2.1 Mr Jim Pearson's solution

The question posed has also been answered, independently and more succinctly, by my friend Jim Pearson of Salford. Without numerical exploration he derived the key equation Eq 1. He considered the limiting condition of N and σ very large, and so ignored the constant terms in Table 2. Straight way he generated a table of limiting values of the two factors D_σ and M_σ of the rounded-up integer p_σ, retaining only terms in N. To his table he added columns showing the ratios of successive values $D_\sigma/D_{\sigma-1}$ and $M_\sigma/M_{\sigma-1}$.

I reproduce some of his results in Table 3. The second column gives the factor D_σ, and the fourth column gives M_σ in the limit $N \to \infty$ where the constant terms in Table 2 can be ignored. The rounded-up number at the end of the $\sigma = 1$ sequence is approximately $p_1 = N^2/4$. Hence at $\sigma = 2$ this has been scaled by the product of ratios to $N^2/4 \times \frac{3}{4}.\frac{3}{2}$ and so on. Thus

$$p_\sigma = \frac{N^2}{4}\,\frac{3.3}{2.4}\,\frac{5.5}{4.6}\,\frac{7.7}{6.8}\,\ldots \quad \text{so} \quad \frac{N^2}{p_\sigma} \to 4\,\frac{2.4}{3^2}\,\frac{4.6}{5^2}\,\frac{6.8}{7^2}\ldots \tag{2}$$

This is very similar to the famous Wallis's product for $\pi/2$, written down in 1655:

$$\frac{\pi}{2} = \frac{2.2}{1.3}\,\frac{4.4}{3.5}\,\frac{6.6}{5.7}\,\frac{8.8}{7.9}\ldots \tag{3}$$

σ	D_σ	$D_\sigma/D_{\sigma-1}$	M_σ	$M_\sigma/M_{\sigma-1}$
1	$1N/2$		$1N/2$	
2	$3N/8$	$3/4$	$3N/4$	$3/2$
3	$5N/16$	$5/6$	$15N/16$	$5/4$
4	$35N/128$	$7/8$	$35/32N$	$7/6$
5	$63N/256$	$9/10$	$315/256N$	$9/8$
6	$231N/1024$	$11/12$	$693N/512$	$11/10$

Table 3: Ratios of successive values of the two factors D_σ and M_σ in the limit of $N \to \infty$.

In fact Eq 2 is twice Wallis's formula, though with the fractions paired the alternative way. Without being a formal proof, this argument clearly answers the Question. The table shows clearly how the blocks of paired fractions in Wallis's formula arise from these two ratios of successive subsequences. Jim's succinct presentation turns the rounding-up procedure into an attractive practical illustration of the structure of Wallis's formula.

2.2 Exact values in Region 1 for N finite

I now return to my own investigation of the sequence of rounded-up numbers. With the help of Sloane's on-line encyclopaedia of integer sequences I am led to consider the Maclaurin expansion of $1-\sqrt{(1-x)}$:

$$\frac{x}{2}+\frac{x^2}{8}+\frac{x^3}{16}+\frac{5\,x^4}{128}+\frac{7\,x^5}{256}+\frac{21\,x^6}{1024}+\frac{33\,x^7}{2048}+\frac{429\,x^8}{32768}+\frac{715\,x^9}{65536}+\frac{2431\,x^{10}}{262144}+\cdots.$$

So this is a generating function for the coefficients of $N+2$ in h_σ, describing how the span of k from 1 to $N-2$ is divided into subsequences of ever decreasing length. The coefficient of x^σ is

$$\frac{(2\sigma-2)!}{2^{2\,\sigma-1}\,(\sigma-1)!\,\sigma!}. \tag{4}$$

If $x\to 1$ the expansion probably breaks down, but in a formal sense it converges to 1. This would imply that the sum of all h_σ is $(N+2)$ as $N\to\infty$.

Eq 4 also shows that the denominators should be considered as advancing in odd powers of σ. I will therefore rewrite Table 2 as Table 4. h' is the numerator of the coefficient of $N+2$ in h and μ_σ is the numerator of the coefficient of $N+2$ in M_σ. K_σ and D_σ both have a constant term. Calling this Q_σ, I will write them explicitly as $K_\sigma=(\kappa_\sigma N+Q_\sigma)/2^{2\sigma-1}$, $D_\sigma=(\delta_\sigma N-Q_\sigma)/2^{2\sigma-1}$. Each number j in the second column denotes a power of 2 such that 2^j is the denominator of all coefficients in the row. As an example, the bottom row means that the $\sigma=7$ subsequence has duration $h_7=132(N+2)/2^{13}$

σ	power	h'	κ_σ	δ_σ	Q_σ	μ_σ
1	1	1	1	1	0	1
2	3	1	5	3	2	6
3	5	2	22	10	12	30
4	7	5	93	35	58	140
5	9	14	386	126	260	630
6	11	42	1586	462	1124	2772
7	13	132	6476	1716	4760	12012

Table 4: Numerators of the coefficients of N in h, K_σ, D_σ and M_σ, and the numerator of the constant terms Q_σ in Table 2. The Power column denotes the power of 2 in each denominator.

and ends when $k = K_7 = (6476N + 4760)/2^{13}$. The value of D_7 is $(1716N - 4760)/2^{13}$ and the other factor M_7 is $12012(N + 2)/2^{13}$.

For obtaining explicit expressions for these coefficients, Sloane's table of sequences is again invaluable. I find that the numerators of all the coefficients of N are related to factorial and binomial coefficients $C(n, r) \equiv {}^nC_r$:

- The numerators of h_σ are Catalan numbers:

$$h'_\sigma = \frac{1}{\sigma} C(2\sigma - 2, \sigma - 1) = \frac{(2\sigma - 2)!}{\sigma!(\sigma - 1)!} \tag{5}$$

.

- K_σ has numerator $\kappa = 2^{2\sigma - 1} - C(2\sigma - 1, \sigma)$.
- D_σ has numerator $\delta = C(2\sigma - 1, \sigma) = (2\sigma - 1) \times$ numerator of h_σ.
- M_σ has numerator $\mu = \dfrac{(2\sigma - 1)!}{[(\sigma - 1)!]^2} = \sigma\,\delta_\sigma = \sigma(2\sigma - 1) \times$ numerator of h_σ.
- All have denominator $2^{2\sigma - 1}$.

The sequence of constant terms Q is not listed in Sloane's encyclopaedia. I have found that

$$Q_\sigma = 4Q_{\sigma - 1} + 2{h'}_\sigma\,. \tag{6}$$

Thus $Q_7 = 4 \times 1124 + 2 \times 132$. As explicit formula for the Q_σ is obtained in the next section.

To summarise, at the end of the σ^{th} subsequence, the rounded-up number p_σ is the product of two factors D_σ and M_σ and is given by

$$p_\sigma = \frac{\sigma\delta_\sigma}{2^{4\sigma - 2}} (\delta_\sigma N - Q_\sigma)(N + 2)\,, \qquad \delta_\sigma = \frac{(2\sigma - 1)!}{\sigma!(\sigma - 1)!} \tag{7}$$

In those cases where D_σ and M_σ are both integers, this is exact; otherwise it is a fair approximation. A few numerical trials suggest that the true rounded-up number will exceed p_σ, and this error probably increases towards the end of the rounding-up process.

The Question poses that $\lim_{n\to\infty} N^2/F = \pi$, equivalent to $\lim_{\sigma\to\infty} N^2/(D_\sigma M_\sigma) = \pi$. In what follows I am assuming that these Region 1 formulae continue to hold to the end of the rounding-up process, though in a later section I show that they fail. Within this assumption, and neglecting the terms which do not contain N, we have

$$\frac{N^2}{D_\sigma\, M_\sigma} \approx \frac{2^{4\sigma - 2}\,[(\sigma - 1)!]^2}{C(2\sigma - 1, \sigma)\,(2\sigma - 1)!} = \frac{2^{4\sigma - 2}\,\sigma\,[(\sigma - 1)!]^4}{[(2\sigma - 1)!]^2}\,. \tag{8}$$

Some numerical values show slow convergence to π, but from above – not below as in Figure 1.

σ	p
10	3.221089
20	3.181105
40	3.161289
60	3.15471
100	3.149456
200	3.145522
500	3.143164
1000	3.142378
30000	3.141619

The expression can also be written in terms of the double factorial function, defined by

$$(2n)!! \;=\; 2.4.6.8....(2n-2).2n \;=\; 2^n\, n! \text{ for } x \text{ even and by}$$

$$(2n-1)!! \;=\; 3.5.7.9.....(2n-1) \;=\; \tfrac{(2n)!}{2^n\, n!} \text{ for } x \text{ odd. Thus}$$

$$\frac{N^2}{p_\sigma} \;=\; \frac{2^{2\sigma}\,(\sigma-1)!\,\sigma!}{[(2\sigma-1)!!]^2} \;=\; \frac{(2\sigma)!!^2}{\sigma\,(2\sigma-1)!!^2} \;=\; \frac{2\sigma+1}{\sigma}\,\frac{(2\sigma)!!^2}{(2\sigma-1)!!\,(2\sigma+1)!!}.$$

$$=\; \frac{2\sigma+1}{\sigma}\,\frac{(2.4.6.8.10.12....\,.2\sigma)^2}{(3.5.7.9.11.13.\;...\;.2\sigma-1)(3.5.7.9.11.13.\;...\;.2\sigma+1)}\;. \tag{9}$$

This is very similar to Wallis's product for $\pi/2$ already given as Eq 3:

$$\frac{\pi}{2} \;=\; \frac{2.2}{1.3}\,\frac{4.4}{3.5}\,\frac{6.6}{5.7}\,\frac{8.8}{7.9}...$$

which shows the same component fractions at which an evaluation may be terminated. The factor $(2\sigma+1)/\sigma$ tends to 2 as $\sigma \to \infty$, in which limit Eq 8 tends to twice Wallis's formula. Bear in mind that σ is the index for the subsequences, as in Table 1, and increases indefinitely. Therefore the Region 1 formulae of Eq 8 does predict that

$$\frac{N^2}{p_\sigma} \;\to\; \pi \text{ as } \sigma \to \infty\;.$$

However, I again state that this does not in itself mean that the actual rounding-up process with integers converges to the same limit.

2.3 Proof of Region 1 formulae and evaluation of constants Q

It remains is necessary to prove that the relations at Eq 5 *et seq* are self consistent and do satisfy Eq 1a, b. Use induction on σ. Because the expressions become complicated when the exact formulae are used, I will treat this in two stages. First ignore the constant terms and the $\frac{1}{2}$ in Eq 1a and deal only with the terms involving N. In this limit $N \to \infty$

$$h_{\sigma+1} \;\approx\; \tfrac{1}{2}\left(D_\sigma - \frac{M_\sigma}{\sigma+1}\right) \;=\; \tfrac{1}{2}(2\sigma-1)h_\sigma\left(1-\frac{\sigma}{\sigma+1}\right) \;=\; \frac{(2\sigma-1)}{2(\sigma+1)}\,h_\sigma$$

$$=\; \frac{(2\sigma-1)}{2(\sigma+1)}\,\frac{(2\sigma-2)!\,N}{2^{2\sigma-1}\,\sigma!\,(\sigma-1)!} \;=\; \frac{(2\sigma-1)!\,N}{2^{2\sigma}\,(\sigma+1)!\,(\sigma-1)!}$$

$$=\; \frac{(2\sigma-1)!\,N}{2^{2\sigma}\,(\sigma+1)!\,(\sigma-1)!}\,\frac{2\sigma}{2\sigma} \;=\; \frac{(2\sigma)!\,N}{2^{2\sigma+1}\,(\sigma+1)!\,\sigma!} \tag{10}$$

This is indeed the expression for $h_{\sigma+1}$. From this

$$D_{\sigma+1} \equiv D_\sigma - h_{\sigma+1} \;=\; \frac{C(2\sigma-1,\sigma)\,N}{2^{2\sigma-1}} \;-\; \frac{(2\sigma)!\,N}{2^{2\sigma+1}\,(\sigma+1)!\,\sigma!}$$

$$=\; \frac{(2\sigma-1)!\,N}{2^{2\sigma-1}\,\sigma!\,(\sigma-1)!}\,\frac{2^2\,\sigma(\sigma+1)}{2^2\,\sigma(\sigma+1)} \;-\; \frac{(2\sigma)!\,N}{2^{2\sigma+1}\,(\sigma+1)!\,\sigma!}$$

$$=\; \frac{(2\sigma)!\,N}{2^{2\sigma+1}\,(\sigma+1)!\,\sigma!}[2\sigma+2-1] \;=\; \frac{(2\sigma+1)!\,N}{2^{2\sigma+1}\,(\sigma+1)!\,\sigma!}$$

which corresponds with substituting $\sigma \to \sigma + 1$ in the expression for D_σ. Finally for the factor M

$$M_{\sigma+1} \equiv M_\sigma + (\sigma+1) h_{\sigma+1} \;=\; \frac{(2\sigma-1)!\, N}{2^{2\sigma-1}\,(\sigma-1)!^2}\,\frac{(2\sigma)^2}{2^2\,\sigma^2} \;+\; \frac{(\sigma+1)(2\sigma)!\, N}{2^{2\sigma+1}\,(\sigma+1)!\,\sigma!}$$

$$= \frac{N}{2^{2\sigma+1}\,\sigma!^2}\left[(2\sigma)^2\,(2\sigma-1)! + (2\sigma)!\right] \;=\; \frac{N}{2^{2\sigma+1}\,\sigma!^2}\,(2\sigma)!(2\sigma+1) \;=\; \frac{N}{2^{2\sigma+1}\,\sigma!^2}\,(2\sigma+1)!$$

which is also correct and establishes proof as $N \to \infty$.

The case of finite N is more complicated because of the $N+2$ and the constant coefficients Q. Use the notation in the bullet points by Eq 5 and write $2^{2\sigma-1} = \tau$. Then Eq 1 takes the form

$$\frac{h'_{\sigma+1}}{4\tau}(N+2) \;=\; \frac{\delta_\sigma}{2\tau}N \;-\; \frac{Q_\sigma}{2\tau} - \frac{\mu_\sigma}{2\tau(\sigma+1)}(N+2) \;+ \tfrac{1}{2}$$

Separate out the terms in N to find

$$\tfrac{1}{2}h'_{\sigma+1} \;=\; \delta_\sigma \;-\; \frac{\sigma}{\sigma+1}\,\delta_{\sigma+1}$$

where $\mu_\sigma = \sigma\delta_\sigma$ has been used (see near Eq 5). This reduces to

$$h'_{\sigma+1} \;=\; \frac{2}{\sigma+1}\,\delta_{\sigma+1} \;=\; \frac{(2\sigma)!}{(\sigma+1)!\,\sigma!}$$

consistent with the analysis above for $N \to \infty$. That means that the constant terms must satisfy

$$h'_{\sigma+1} \;=\; -Q_\sigma \;-\; \frac{2\sigma}{\sigma+1}\,\delta_\sigma \;+\; \tau$$

$$Q_\sigma \;=\; \tau - \frac{(2\sigma)!}{\sigma!^2} \;=\; 2^{2\sigma-1} \;-\; C(2\sigma,\sigma). \tag{11}$$

This is a satisfying closed expression for the constants in Tables 2 and 4 and shows the strong similarity to $\kappa_\sigma = 2^{2\sigma-1} \;-\; C(2\sigma-1,\sigma)$: see function definitions next to Eq 5. It is straightforward to verify that the recursion relation Eq 6 follows from Eqs 8 and 9, as follows

$$4Q_\sigma + 2h'_{\sigma+1} \;=\; 2^{2\sigma+1} \;-\; \frac{(2\sigma)!}{(\sigma+1)!^2}\left[4(\sigma+1)^2 - 2(\sigma+1)\right]$$

and the bracket [...] evaluates to $4\sigma^2 + 6\sigma + 2 = (2\sigma+2)(2\sigma+1)$, making the expression into $Q_{\sigma+1}$. The Q values are twice those in the Sloane on-line database, sequence A008549.

2.4 Summary of Region 1

After all this algebra let me pull together the formulae characterising Region 1. The rounding-up process starts with integer N and each step is indexed by k. The rounded-up integer p is the product of two factors, $d_k \times m_k$, the first decreasing by 1 for each increment in k. This Region is characterised by consisting of a number of obvious subsequences in each of which m_k increases by a constant amount σ as k increments by 1. The values of d_k and m_k at the end of each subsequence are labelled D_σ and M_σ. The binomial coefficient $C(2\sigma,\sigma) = (2\sigma)!/\sigma!^2$ features in the values of both factors. Call this $\mathbb{C}_\sigma$.

$$D_\sigma \;=\; \frac{(\delta_\sigma N - Q_\sigma)}{2^{2\sigma-1}}, \qquad M_\sigma \;=\; \frac{\sigma\delta_\sigma}{2^{2\sigma-1}}(N+2)\,,$$

$$\delta_\sigma \;=\; \tfrac{1}{2}\mathbb{C}_\sigma, \qquad Q_\sigma \;=\; 2^{2\sigma-1} - \mathbb{C}_\sigma\,. \tag{12a}$$

Hence

$$D_\sigma = \frac{\mathbb{C}_\sigma}{2^{2\sigma}}(N+2) - 1 \ , \qquad M_\sigma = \frac{\sigma\mathbb{C}_\sigma}{2^{2\sigma}}(N+2) \, . \tag{12b}$$

The number of rounding-up steps in the σ^{th} subsequence is

$$h_\sigma = \frac{\mathbb{C}_\sigma}{2^{2\sigma}(2\sigma-1)}(N+2) \, . \tag{12c}$$

For suitably low values of σ these formula give integers which correctly follow the rounding-up algorithm.

The formulae for Region 1 of the rounding-up process cease to be meaningful once the sequence length h_σ has reduced to less than 2 and certainly by 1. The breakdown of the Eq 12 marks the beginning of Region 2. Before examining this, however, I want to jump forwards to the end of the process in Region 3, then come back to the transitional Region 2.

3 Region 3: close of rounding-up procedure

As the rounding-up procedure comes to an end, one factor, d_k, steps down .... 5, 4, 3, 2 while the other factor m_k increases roughly as .... $F/5$, $F/4$, $F/3$, $F/2$ such that their product $p_k = d_k m_k$ come close to its final value F. Thus the d_k are known, but without stepping through the whole rounding-up procedure we cannot know how the sequence of m_k for any given N will close. However, it should be possible to say something about how *on average* the procedure will close. This section considers the probability of the sequence p_k reaching F from some assumed value a few steps back in the process.

To illustrate the argument, take the case of $d = 6$, $m = m_6$, say, so $p_6 = 6m_6$[1]. The next step is to divide by 5, as which the remainder could be 0, 1, 2, 3 or 4, requiring respectively that 0, 4, 3, 2, 1 be added to p_6 to form p_5. On average, with m_6 a typical, virtually random number, each of these additions will occur with equal probability. The mean amount added is $(0+1+2+3+4)/5 = 2$. Assuming that $p_5 \approx p_6 + 2$ is also a typical number, we divide by 4 and add 0, 3, 2, or 1. The average added is 3/2. Then divide by 3 and add on average 1, and finally by 2 and add on average 1/2. The cumulated increment from p_6 is $2 + \frac{3}{2} + 1 + \frac{1}{2} = 5$.

To generalise, suppose we pick up the procedure at $d_k = \tau$. This is divided by $\tau - 1$ and the average amount added in rounding-up is the arithmetic mean of 0, 1, 2, 3, ... $\tau - 2$, which is $\frac{1}{2}(\tau - 2)$. The accumulated addition to the end of the process is

$$\frac{(\tau-2)}{2} + \frac{(\tau-3)}{2} + \frac{(\tau-4)}{2} + \ldots + \frac{3}{2} + \frac{2}{2} + \frac{1}{2}\,, \qquad \tau - 2 \text{ terms.}$$

This sums to

$$\tfrac{1}{4}(\tau-1)(\tau-2) \tag{13}$$

and shows that on average $F = p_{final}$ is approached by parabolic growth. For example, the last 110 steps in the process for large N would on average contribute about 3000 to F. The probability of obtaining a particular cumulative increment is examined in Appendix 2.

It is natural now to ask whether τ can be taken to be the limiting value of d_k for Region 1 so that Region 3 can be extended backwards through the transitional Region 2 and stitched on to the end of Region 1. The next section addresses this join.

[1]Note that in this notation the subscripts denote the matching value of d; 6 does not mean the 6th rounded-up number from the start of the whole procedure.

4 Region 2: transition

		Exact					Formulae			
k	D	M	σ	p	h	K	D	M	σ	p
15	15	16	1	240	16	15	15	16	1	240
16	14	18	2	252						
17	13	20	2	260						
18	12	22	2	264						
19	11	24	2	264	4	19	11	24	2	264
20	10	27	3	270						
21	9	30	3	270	2	21	9	30	3	270
22	8	34	4	272	1.3	22.3	7.8	35.0	4	271.3
23	7	39	5	273	0.9	23.1	6.9	39.4	5	270.7
					0.7	23.8	6.2	43.3	6	269.3
24	6	46	7	276	0.5	24.3	5.7	46.9	7	267.6
					0.4	24.7	5.3	50.3	8	265.7
					0.3	25.1	4.9	53.4	9	263.6
25	5	56	10	280	0.3	25.4	4.6	56.4	10	261.5
					0.3	25.6	4.4	59.2	11	259.4
					0.2	25.8	4.2	61.9	12	257.3
					0.2	26.0	4.0	64.5	13	255.3
26	4	70	14	280	0.2	26.2	3.8	67.0	14	253.2

Table 5: Half way through the rounding-up calculation for $N = 30$, showing the breakdown of the Region 1 formulae. Comparison of the actual rounding up process (left) with the formulae for Region 1 (right).

Consider Table 5 which shows the central part of the rounding-up calculation for $N = 30$ and the end of Region 1. The left panel of five columns is equivalent to Table 1 and gives the correct integer values. Notice how the subsequence $\sigma = 4$ occupies only 1 row, and how below that an increasing number of σ values are omitted. The 6-column panel on the right are values from the factorial/binomial formulae collected together in Eq 12. They know no restriction on σ. Notice that when the predicted subsequence length h is less than 2, fractional values appear. Furthermore, the product p_σ reaches a maximum at this point and starts to decrease, whereas the integer process generates a strictly increasing sequence.

The σ value at which Region 1 ends is larger when N is large. I have examined the formulae of Eq 12 numerically up to $\sigma = 2400$. Some illustrative values at which $p_{\sigma+1} = p_\sigma$ are listed in Table 6. These marks the end of Region 1.

The asymptotic values of functions of σ as $\sigma \to \infty$ can be obtained using Stirling's factorial approximation or the asymptotic expansion of the gamma function:

$$\ln(n!) \sim n\ln(n) - n \qquad \text{or, more accurately,} \qquad n! \sim n^n e^{-n}\sqrt{2\pi n}\left(1 + \frac{1}{12n} + \frac{1}{288n^2} + \dots\right). \quad (14)$$

Using the second version $\mathbb{C}_\sigma \equiv C(2\sigma, \sigma)$ approaches

$$\mathbb{C}_\sigma \sim \frac{2^{2\sigma}}{\sqrt{\pi\sigma}}\left(1 - \frac{1}{8\sigma} + \frac{1}{128\sigma^2} - \dots\right), \qquad \sigma \to \infty. \quad (15)$$

N	σ	N	σ
30	3.5	1000	42.3
60	6.0	1200	47.9
126	10.2	1500	55.7
200	14.1	1700	60.6
400	22.7	2000	67.6
600	29.9	2200	72.1
800	36.3	2400	76.3

Table 6: Numerically interpolated values of σ at which the product p_σ reaches a maximum.

Substitute this into Eq 12b :

$$D_\sigma \sim \frac{(N+2)}{\sqrt{\pi\sigma}}\left(1-\frac{1}{8\sigma} + \dots\right) - 1, \qquad M_\sigma \sim \frac{\sigma(N+2)}{\sqrt{\pi\sigma}}\left(1-\frac{1}{8\sigma} + \dots\right).$$

$$p_\sigma \sim \frac{(N+2)^2}{\pi}\left(1-\frac{1}{8\sigma} + \dots\right)^2 - \sqrt{\frac{\sigma}{\pi}}\,(N+2)\left(1-\frac{1}{8\sigma} + \dots\right). \qquad (16)$$

Notice how the σ have cancelled in the leading term here and so, if the formula were actually to hold to the end of the rounding-up process, we would have $(N+2)^2/F \to \pi$ rather than the $N^2/F \to \pi$ posed in the Question. The π comes from the asymptotic expansion of the gamma function.

We have some discretion as to where Region 1 is judged to end. Table 6 gave the condition that $p_\sigma = p_{\sigma+1}$, but we might choose $h = \theta,\ 1 \le \theta < 2$. From Eq 12c

$$h_\sigma \sim \frac{N+2}{2\sqrt{\pi}\,\sigma^{3/2}} \qquad \text{so} \qquad \sigma|_{last} \approx \left(\frac{N+2}{2\sqrt{\pi}\,\theta}\right)^{2/3}.$$

Numerically I have fitted a function of this form to the values in Table 6 and obtain close fit with

$$\sigma|_{last} \approx 0\cdot425(N+2)^{2/3}.$$

This gives $\theta = 1\cdot02$, an encouraging result. $\sigma|_{last}$ can be substituted into D_σ at Eq 16 to find the value of the factor d_k at which Region 1 ends:

$$d_k|_{Region1} \approx 0\cdot865N'^{2/3}\left(1-\frac{5}{17\,N'^{2/3}}+\dots\right) - 1 \approx 0\cdot865\,N'^{2/3} - 1, \qquad N' \equiv N+2. \qquad (17)$$

From this, some $(N,\ d)$ values at the end of Region 1 are : $(30, 7\cdot5)$, $(126, 21)$, $(200, 28\cdot5)$, $(800, 73\cdot5)$, $(1000, 85)$, $(2000, 136)$. In the remainder of the process d_k reduces in steps of 1 down to 2, so Eq 17 measures the number of steps left before the rounding-up process reaches its final value, F. The fraction of the rounding-up process occupied by this is about $0\cdot865/N'^{1/3}$ which decreases slowly with N, being about 9% for $N = 1000$.

We are finally able to match Region 1 to Region 2, by equating the d_k in Eq 17 with the τ in Eq 13. Then, taking $\tau = 0\cdot865N'^{2/3} - 1$, the Region 2+ Region 3 tail of the process would add about

$$0\cdot187N'^{4/3} - N'^{2/3}. \qquad (18)$$

For $N = 126$ used in Table 1, Region 1 was shown in §3 to end at about $\sigma = 10$, $d_k = 21$. Eq 18 predicts that 95 will be added to 4914, giving $F = 5009$. The true value is 5014. Bear in mind, of course, that Eq 18 is an *average* or typical addition in Region 3; we cannot know the actual addition without actually carrying out the integer rounding-up.

This Region 1 + Region 3 analysis of the rounding-up process predicts that the limiting value on average will be

$$F \approx \frac{N'^2}{\pi}\left(1-\frac{1}{8\sigma}+\ldots\right)^2 - \sqrt{\frac{\sigma}{\pi}}N'\left(1-\frac{1}{8\sigma}+\ldots\right) + 0\cdot 187N'^{4/3} - N'^{2/3}\,..$$

$$\approx \frac{N'^2}{\pi}\left(1-\frac{10}{17N'^{2/3}}\right) - 0\cdot 368\,N'^{4/3}\left(1-\frac{5}{17N'^{2/3}}\right) + 0\cdot 187N'^{4/3} - N'^{2/3}\,.$$

$$F \approx \frac{N'^2}{\pi} - 0\cdot 368\,N'^{4/3} - 0\cdot 88\,N'^{2/3}\,, \qquad N' = N+2 \tag{19}$$

So does this work in practice? My impression from comparison with a few numerical simulations on the exact rounding-up process is that it underestimates F. I therefore have undertaken a more extensive numerical comparison to see if the same type of formula might apply, though with different coefficients. The best fit for 305 values of N from 10 to 4400, including all between 10 and 150 used for Figure 1, occurs with

$$F \approx \frac{(N+2)^2}{\pi} - 0\cdot 05(N+2)^{4/3} - 3(N+2)^{2/3}\,,$$

a significantly lower contribution from $N'^{4/3}$. Having said that, just as good fit is obtained with

$$F \approx \frac{(N+2)^2}{\pi} - (N+2)\,,$$

and the best fit of all is with the simple

$$F \approx \frac{(N+0\cdot 36)^2}{\pi}\,. \tag{20}$$

So, despite the involved arguments about Region 1 not continuing to the end of the process, and Region 3 adding a little, I find very lttle numerical evidence to support a term in $N^{4/3}$. It does seem clear, however, that a function involving $(N+\beta)^2/2F$, $0\cdot 3 < \beta \le 2$, is a better approximation to π than the simple N^2/F posed in the Question.

John Coffey, January 2014

5 Appendix 1 : Details of calculations for $\sigma = 3$ to 7

For $\sigma = 3$, redefining the auxiliary index h, the pattern is

k	$d_k = N - k$	m_k	p_k
K_2+1	$N-(K_2+1)$	M_2+3	$\frac{3}{32}(3N-10)(N+6)$
K_2+2	$N-(K_2+2)$	M_2+6	$\frac{3}{32}(3N-18)(N+10)$
$\ldots$			$\ldots$
K_2+h-1	$N-(K_2+h-1)$	M_2+3h-3	$\frac{3}{32}(3N-8h+6)(N+4h-2)$
K_2+h	$N-(K_2+h)$	M_2+3h	$\frac{3}{32}(3N-8h-2)(N+4h+2)$

This subsequence will continue provided

$$(3N - 8h - 2)(N + 4h + 2) \ \ge \ (3N - 8h + 6)(N + 4h - 2), \qquad i.e. \quad h \le \frac{N+2}{16}.$$

The largest value in this subsequence is attained at $k = K_3 = (11N + 6)/16$ and is

$$D_3 \times M_3 \ \equiv \ \tfrac{1}{16}(5N - 6) \times \tfrac{15}{16}(N + 2) \ = \ \tfrac{15}{256}(5N - 6)(N + 2).$$

For $N = 126$ this evaluates to 4680, in agreement with Table 1. Also the end-of-subsequence values $K_3 = 87$, $M_3 = M_2 + (N + 2)/16 = 15(N + 2)/16 = 120$ are correctly given.

Case $\sigma = 4$.

k	$d_k = N - k$	m_k	p_k
$K_3 + 1$	$N - (K_3 + 1)$	$M_3 + 4$	$\frac{1}{256}(5N - 22)(15N + 94)$
$K_3 + 2$	$N - (K_3 + 2)$	$M_3 + 8$	$\frac{1}{256}(5N - 38)(15N + 158)$
$\dots$			$\dots$
$K_3 + h - 1$	$N - (K_3 + h - 1)$	$M_3 + 4h - 4$	$\frac{1}{256}(5N - 16h + 10)(15N + 64h - 34)$
$K_3 + h$	$N - (K_3 + h)$	$M_3 + 4h$	$\frac{1}{256}(5N - 16h - 6)(15N + 64h + 30)$

This ends when

$$(5N - 16h - 6)(15N + 64h + 30) \ = \ (5N - 16h + 10)(15N + 64h - 34), \qquad i.e. \ \ h = \frac{5(N+2)}{128}.$$

The largest value of p_k in this subsequence, attained at $k = K_4 = (93N + 58)/128$, is

$$D_4 \times M_4 \ \equiv \ \frac{1}{128}(35N - 58) \times \frac{35}{32}(N + 2) \ = \ \frac{35}{4096}(35N - 58)(N + 2).$$

For $N = 126$ this evaluates to 4760, in agreement with Table 1. Also the end-of-subsequence values $K_4 = 92$, $D_4 = 34$, $M_4 = 140$ are correctly given.

For $\sigma = 5$ the subsequence length is determined by

$$(35N - 128h - 58)(7N + 32h + 14) \ \ge \ (35N - 128h + 70)(7N + 32h - 18), \qquad i.e. \ \ h \le \frac{7(N+2)}{256}.$$

The largest value of the rounded-up integer in this subsequence occurs at $k = K_5 = (193N + 130)/256$ and is

$$D_5 \times M_5 \ \equiv \ \frac{1}{256}(63N - 130) \times \frac{315}{256}(N + 2) \ = \ \frac{315}{2^{16}}(63N - 130)(N + 2).$$

For $N = 126$ these give fractional values, but I have checked them against the calculation for $N = 254$ and they do give correctly $K_5 = 192$, $D_5 = 62$, $M_5 = 315$ and product $p_{K_5} = 19530$.

For $\sigma = 6$ the critical values are

$$h_6 = \frac{21}{1024}(N + 2), \quad K_6 = \frac{1}{1024}(793N + 562), \quad D_6 = \frac{1}{1024}(231N - 562), \quad M_6 = \frac{693}{512}(N + 2).$$

For $\sigma = 7$ they are

$$h_7 = \frac{33}{2048}(N + 2), \quad K_7 = \frac{1}{2048}(1619N + 1190), \quad D_7 = \frac{1}{2048}(429N - 1190), \quad M_7 = \frac{3003}{2048}(N + 2).$$

6 Appendix 2 : Probability distribution of rounded-up numbers in Region 3

Let us pick up the example of $d = 6$, $m = m_6$, say, so $p_6 = 6m_6$, and recall that in this notation the subscripts denote the matching value of d. The next step is to divide by 5, so 0, 4, 3, 2, 1 would need to be added to form p_5, all with equal probability. This is represented in the line in Table 7 labelled 'Start: divide by 5' where the line of five 1s count respectively the number of 0, 1, 2, 3, or 4 added. The panel below labelled 'Div by 4' shows the adding of further numbers 0, 1, 2, or 3 at the division by 4. 'Sum' denotes the sum of rows; the values here mean respectively that there is only on way that 0 could be added (0+0), 2 ways to add 1 (0+1, 1+0), three to add 2 (2+0, 1+1, 0+2), and so on. This probability distribution is carried forwards to the division by 3 and that in turn to division by 2, at which the rounding-up process stops. The bottom row, which sums to 120 = 5!, counts the ways in which values a could occur through rounding-up from p_6 to p_2. The most likely value added is 5 and this has probability $22/120 = 18 \cdot 3\%$. There is a 52% ([20+22+20]/120) chance that the amount added will be 4, 5 or 6.

	a =	0	1	2	3	4	5	6	7	8	9	10
Start: divide by 5	add	1	1	1	1	1						
Div by 4	add 0	1	1	1	1	1						
	1		1	1	1	1	1					
	2			1	1	1	1	1				
	3				1	1	1	1	1			
	sum	1	2	3	4	4	3	2	1			
Div by 3	add 0	1	2	3	4	4	3	2	1			
	1		1	2	3	4	4	3	2	1		
	2			1	2	3	4	4	3	2	1	
	sum	1	3	6	9	11	11	9	6	3	1	
Div by 2	add 0	1	3	6	9	11	11	9	6	3	1	
	1		1	3	6	9	11	11	9	6	3	1
	sum	1	4	9	15	20	22	20	15	9	4	1

Table 7: Development of cumulate probability of adding a to a starting number p_6 through the rounding-up process when the first divisor is 5.

Table 8 shows the equivalent calculation starting at p_7 and dividing first by 6. One is equally likely to add 7 or 8, so the mean added is $7 \cdot 5$, in agreement with Eq 13 with $\tau = 7$. The row sum is 720 = 6! and the probability of adding 7 or 8 is 14%.

I have carried out some numerical simulation to check this process starting each time with a random number and counting the number of times the various integers are added on. These give confidence that this model is correct. For completeness, Table 9 gives the probability weightings for adding various numbers, for starting positions p_3 to p_8. In each row the sum is $t!$ where t is the first divisor used. The probability of adding a at the end of the rounding-up process is then the row value divided by $t!$. The integers seem to be called Mahonian numbers and appear in the Sloane on-line encyclopaedia as A008302.

	$a =$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Div 6	add	1	1	1	1	1	1										
Div 5	add 0	1	1	1	1	1	1										
	1		1	1	1	1	1	1									
	2			1	1	1	1	1	1								
	3				1	1	1	1	1	1							
	4					1	1	1	1	1	1						
	sum	1	2	3	4	5	5	4	3	2	1						
Div 4	add 0	1	2	3	4	5	5	4	3	2	1						
	1		1	2	3	4	5	5	4	3	2	1					
	2			1	2	3	4	5	5	4	3	2	1				
	3				1	2	3	4	5	5	4	3	2	1			
	sum	1	3	6	10	14	17	18	17	14	10	6	3	1			
Div 3	add 0	1	3	6	10	14	17	18	17	14	10	6	3	1			
	1		1	3	6	10	14	17	18	17	14	10	6	3	1		
	2			1	3	6	10	14	17	18	17	14	10	6	3	1	
	sum	1	4	10	19	30	41	49	52	49	41	30	19	10	4	1	
Div 2	add 0	1	4	10	19	30	41	49	52	49	41	30	19	10	4	1	
	1		1	4	10	19	30	41	49	52	49	41	30	19	10	4	1
	sum	1	5	14	29	49	71	90	101	101	90	71	49	29	14	5	1

Table 8: Equivalent table for rounding up from p_7.

Start position	First divide by	Add $a =$ 0	1	2	3	4	5	6	7	8	9	10	11	12	
3	2	1	1												
4	3	1	2	2	1										
5	4	1	3	5	6	5	3	1							
6	5	1	4	9	15	20	22	20	15	9	4	1			
7	6	1	5	14	29	49	71	90	101	101	90	71	49	29	...
8	7	1	6	20	49	98	169	259	359	455	531	573	573	531	...

Table 9: The number of ways that a can be added to a starting integer in rounding-up, for start positions from p_2 to p_8. Rows extend right so as to be symmetric.