ON THE NUMBER OF TIMES AN INTEGER OCCURS AS A BINOMIAL COEFFICIENT

H. L. ABBOTT, P. ERDÖS, AND D. HANSON

Let $N(t)$ denote the number of times the integer $t > 1$ occurs as a binomial coefficient; that is, $N(t)$ is the number of solutions of $t = \binom{n}{r}$ in integers n and r. We have $N(2) = 1$, $N(3) = N(4) = N(5) = 2$, $N(6) = 3$, etc. In a recent note in the research problems section of the MONTHLY, D. Singmaster [1] proved that

$$N(t) = O(\log t). \tag{1}$$

He conjectured that $N(t) = O(1)$ but pointed out that this conjecture, if it is in fact true, is perhaps very deep. In [1] and [5], Singmaster points out that $N(t) = 6$ for the following values of $t \leqq 2^{48}$; $t = 120$, 210, 1540, 7140, 11628 and 24310. It has been shown by Singmaster [5] and D. Lind [6] that $N(t) \geqq 6$ infinitely often. Singmaster has verified that the only value of $t \leqq 2^{48}$ for which $N(t) \geqq 8$ is $t = 3003$, for which $N(t) = 8$.

In this note we obtain some additional information about the behavior of $N(t)$. In Theorem 1 we prove that the average and normal order of $N(t)$ is 2; in fact, we prove somewhat more than this, namely, the number of integers t, $1 < t \leqq x$, for which $N(t) > 2$ is $O(\sqrt{x})$. (See [4] p. 263 and p. 356, for the definitions of average and normal order.) In Theorem 2 we give an upper bound for $N(t)$ in terms of the number of distinct prime factors of t. Our main result is Theorem 3, in which we show that (1) can be improved to $N(t) = O(\log t/\log\log t)$. Finally, in Theorem 4, we consider the related problem of determining the number of representations of an integer as a product of consecutive integers.

THEOREM 1. *The average and normal order of $N(t) = 2$.*

Proof. For integral x, let n be defined by $\binom{2n-2}{n-1} < x \leqq \binom{2n}{n}$ so that $n = O(\log x)$. We have

$$\sum_{1<t\leqq x} N(t) = 2 \sum_{\substack{1<\binom{m}{r}\leqq x \\ 2r\leqq m}} 1 - \sum_{1<\binom{2k}{k}\leqq x} 1$$

$$(2)\qquad \begin{aligned} &= 2\Big\{ \sum_{1<\binom{m}{1}\leq x} 1 + \sum_{1<\binom{m}{2}\leq x} 1 + \sum_{\substack{1<\binom{m}{r}\leq x \\ 3\leq r\leq m/2}} 1\Big\} - \sum_{1<\binom{2k}{k}\leq x} 1 \\ &= 2x + 2\sqrt{2}x^{1/2} + O(x^{1/3}n) \\ &= 2x + 2\sqrt{2}x^{1/2} + O(x^{1/3}\log x). \end{aligned}$$

It follows that the average order of $N(t)$ is 2.

Let $f(x)$ be the number of integers t, $1 < t \leqq x$, such that $N(t) = 2$ and $g(x)$ the number such that $N(t) > 2$, so that $f(x) + g(x) = x - 2$. We have

$$(3)\qquad \begin{aligned} \sum_{1<t\leqq x} N(t) &\geqq 2f(x) + 3g(x) + 1 \\ &= 2(x - 2 - g(x)) + 3g(x) + 1 \\ &= 2x + 2g(x) - 3. \end{aligned}$$

It follows from (2) and (3) that $g(x) = O(x^{1/2})$ and this implies that the normal order of $N(t)$ is 2.

THEOREM 2. *Let $w(t)$ denote the number of distinct prime factors of the integer $t > 1$. For all t satisfying $w(t) < \log t/\log\log t$ we have*

$$(4)\qquad N(t) < \frac{2w(t)\log t}{\log t - w(t)\log\log t}.$$

Proof. The theorem can be verified directly for $t \leqq 20$. In what follows we therefore assume $t \geqq 21$. Let $k = k(t)$ be the largest integer for which $t = \binom{n}{k}$ for some $n \geqq 2k$. Then clearly

$$(5)\qquad N(t) \leqq 2k.$$

By an easy induction argument we have, for $k \geqq 4$, $t = \binom{n}{k} \geqq \binom{2k}{k} \geqq e^k$. Since we are assuming $t \geqq 21 > e^3$, the inequality $t \geqq e^k$ holds for all $k \geqq 1$. Equivalently,

$$(6)\qquad k \leqq \log t \text{ and } \log k \leqq \log\log t.$$

Let P^α be the highest power of the prime P which divides t. Then, according to the well-known theorem of Legendre,

$$\alpha = \sum_{i=1}^{[\log_P n]} \left\{ \left[\frac{n}{P^i}\right] - \left[\frac{n-k}{P^i}\right] - \left[\frac{k}{P^i}\right] \right\}.$$

Each term in the sum on the right is either 0 or 1. The number of non-zero terms is therefore α and we must have

(7) $$P^{\alpha} \leqq n.$$

From $t = \binom{n}{k}$ and the inequality $\binom{n}{k} \geqq \left(\frac{n}{k}\right)^{k}$, we obtain

(8) $$n \leqq kt^{1/k}$$

and from (7) and (8) it follows that

$$t = \Pi P^{\alpha} \leqq n^{w(t)} \leqq k^{w(t)} t^{w(t)/k}.$$

If we take logarithms and substitute from the second inequality in (6) we get, after some manipulations,

$$k \leqq \frac{w(t) \log t}{\log t - w(t) \log\log t},$$

and this, together with (5), yields (4). This completes the proof of Theorem 2.

We come now to our main result.

THEOREM 3. $N(t) = O(\log t/\log\log t)$.

Proof. We shall need to make use of the following deep result of A. E. Ingham [2] on the distribution of the primes: If $\alpha \geqq 5/8$, there is a prime between x and $x + x^{\alpha}$ for all sufficiently large x.

For a given integer t, let $S = \{n : t = \binom{n}{k}$ for some $k \leqq n/2\}$. Write $S = S_1 \cup S_2$ where $S_1 = \{n : n \in S, n > (\log t)^{6/5}\}$ and $S_2 = \{n : n \in S, n \leqq (\log t)^{6/5}\}$. We first estimate the size of S_1. Let $n \in S_1$ and let $t = \binom{n}{k}$. We have at our disposal the following inequalities:

(9) $$t = \binom{n}{k} \geqq \left(\frac{n}{k}\right)^{k}$$

(10) $$t \geqq e^{k} \text{ (see the proof of Theorem 2)}$$

(11) $$n > (\log t)^{6/5}.$$

Thus

$$k \leqq \frac{\log t}{\log n/k} \leqq \frac{\log t}{\log (n/\log t)} \leqq \frac{\log t}{\log (\log t)^{1/5}}$$
$$= O\left(\frac{\log t}{\log\log t}\right),$$

where we have used, successively, (9), (10) and (11). It follows that

$$|S_1| = O(\log t/\log\log t).$$

Next we must estimate the size of S_2. Let N be the largest number in S_2 and let $t = \binom{N}{K}$. We have the inequalities

$$N \leqq (\log t)^{6/5} \text{ and } t \leqq N^K$$

from which we get $N \leqq (K \log N)^{6/5}$. This in turn implies, for N sufficiently large,

$$N \leqq K^{8/5} < K^{8/5} + K,$$

and it is easy to see that this last inequality implies

$$(N - K) + (N - K)^{5/8} \leqq N.$$

We are now in a position to apply the theorem of Ingham. By this theorem, there is a largest prime P satisfying $K \leqq N - K < P \leqq N$. It follows that P divides t and hence that $n \geqq P$ for all $n \in S_2$. Hence all of the numbers in S_2 lie between P and N. The number of numbers in S_2 is thus

$$|S_2| \leqq N - P \leqq P^{5/8} \leqq N^{5/8} \leqq (\log t)^{3/4} = O(\log t/\log\log t),$$

where, in obtaining the second inequality, we again appeal to Ingham's result. This completes the proof of Theorem 3.

We remark that if one makes use of the unproved conjecture of Cramér [3] asserting that there is a prime between x and $x + (\log x)^2$ for all sufficiently large x, then our argument gives $N(t) = O((\log t)^{2/3+\varepsilon})$. The proof is basically the same as before, except that one puts $S_1 = \{n: n \in S, \log n > (\log t)^{1/3-\varepsilon}\}$. We omit the rather laborious details of the argument.

We conclude with a brief discussion of a somewhat related problem. Let $G(t)$ denote the number of representations of the positive integer t as a product of consecutive integers; that is, $G(t)$ is the number of solutions of $t = (n+1)(n+2)\cdots(n+l)$ in integers n and l. For any such solution we have $t \geqq l!$ and consequently we get $G(t) = O(\log t/\log\log t)$. For this problem, however, we can get a substantially stronger result.

THEOREM 4. $G(t) = O(\sqrt{\log t})$.

Proof. Let $S = \{l: t = (n+1)(n+2)\cdots(n+l) \text{ for some } n\}$. Let L_0 be the largest number in S and let

$$S_1 = \{l: l \in S, L_0 - C(\log t)^{1/2} < l \leqq L_0\} \text{ and } S_2 = \{l: l \in S, l \leqq L_0 - C(\log t)^{1/2}\}.$$

C is a constant. It is clear that $|S_1| \leqq C(\log t)^{1/2}$. It remains to estimate the size of $|S_2|$. Let 2^α be the highest power of 2 which divides t. Then, for some constant C_1,

$$\alpha \geqq \sum_{j=1}^{\infty} \left[\frac{L_0}{2^j}\right] \geqq L_0 - C_1 \log L_0. \tag{12}$$

Let L be the largest number in S_2 and let $t = (N+1)(N+2)\cdots(N+L)$. Let 2^β be the highest power of 2 which divides one of $(N+1),(N+2),\cdots,(N+L)$, say $N+k$. Then

$$\alpha = \beta + \sum_{j=1}^{\infty}\left[\frac{L-k}{2^j}\right] + \sum_{j=1}^{\infty}\left[\frac{k-1}{2^j}\right]. \tag{13}$$

In fact (13) follows from the observation that the first sum on the right is the exponent to which 2 divides the product $(N+k+1)(N+k+2)\cdots(N+L)$, while the second sum is the exponent to which 2 divides the product $(N+1)(N+2)\cdots(N+k-1)$. It follows from (13) that

$$\alpha \leqq \beta + \sum_{j=1}^{\infty}\left[\frac{L}{2^j}\right] \leqq \beta + L. \tag{14}$$

Thus,

$$\begin{aligned}\beta &\geqq \alpha - L \\ &\geqq (L_0 - C_1\log L_0) - (L_0 - C(\log t)^{1/2}) \\ &\geqq C(\log t)^{1/2} - C_1\log L_0 \\ &\geqq C_2(\log t)^{1/2},\end{aligned} \tag{15}$$

where we have used (14), (12), the definition of S_2 and the estimate $L_0 = O(\log t)$. We need two further inequalities; the first of which is obvious. These are

$$(N+1)^L \leqq t \tag{16}$$

and, for t sufficiently large,

$$N+1 \geqq 2^{\beta-1}. \tag{17}$$

To obtain (17) we simply have to notice that $N+L \geqq N+k \geqq 2^\beta$, so that $N+1 \geqq 2^\beta - (L-1)$ and (17) now follows from (15) and the fact that $L = O(\log t)$.

It now follows from (15), (16) and (17) that $L \leqq C_3\,(\log t)^{1/2}$, where C_3 is a positive constant depending on C_2, and hence on C. This completes the proof of Theorem 4.

We remark that by choosing $C = (1+\varepsilon)(\log 2)^{-1/2}$, our argument yields $G(t) < (2+\varepsilon)(\log t/\log 2)^{1/2}$ for every $\varepsilon > 0$, provided $t \geqq t_0(\varepsilon)$.

References

1. D. Singmaster, How often does an integer occur as a binomial coefficient, this MONTHLY, 78 (1971) 385–386.

2. A. E. Ingham, On the difference between consecutive primes, Quart. J. Math. Oxford, 8 (1937) 255–266.

3. H. Cramér, On the order of magnitude of the difference between consecutive prime numbers, Acta. Arith., 2 (1936) 23–46.

4. G. H. Hardy and E. M. Wright, Introduction to the Theory of Numbers, Oxford University Press, New York. 1960.

5. D. Singmaster, Repeated binomial coefficients and Fibonacci Numbers, (to appear in Fibonacci Quarterly).

6. Douglas Lind, The quadratic field $Q(\sqrt{5})$ and a certain diophantine equation, Fibonacci Quarterly, 6 (1968) 86–94.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ALBERTA, EDMONTON, ALBERTA, CANADA.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALGARY, CALGARY, ALBERTA, CANADA.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SASKATCHEWAN, REGINA, SASKATCHEWAN, CANADA.