ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 7, Number 4, Fall 1977

ON THE PROBABILITY THAT AN INTEGER CHOSEN ACCORDING TO THE BINOMIAL DISTRIBUTION BE k-FREE

J. E. NYMANN AND W. J. LEAHEY

Introduction. Let s and t be integers chosen from among the first $n + 1$ non-negative integers according to a binomial distribution with parameter p, $0 < p < 1$. Consider the probability that s and t be relatively prime. In [1] we showed that this probability tends to $6/\pi^2$, independent of p, as $n \to \infty$. Suppose now we choose a single integer s from the first $n + 1$ non-negative integers according to a binomial distribution and ask what is the probability that s be square-free. In this paper we show that the techniques of [1] can also be used to show that this probability is $6/\pi^2$ in the limit. In fact we show something more, viz., that the probability that s be k-free, k any integer greater than 1, is $1/\zeta(k)$ where ζ denotes the Riemann zeta-function. (s is k-free if and only if s is not divisible by the k-th power of any prime.) In section 1 we deal with the case $k > 2$ and in section 2, with the case $k = 2$.

1. Let n be a non-negative integer and denote by N_n the set of integers $0, 1, 2, \cdots, n$. Let P_n be a probability distribution on N_n and let Q_k denote the set of non-negative k-free integers. Set $Q_k(n) = Q_k \cap N_n$. For any positive integer d, let $A_n(d) = \{j \in N_n : j \equiv 0 \pmod{d}\}$. We then have the following.

LEMMA 1. *Let P_n be any probability measure on N_n. Then for $n > 1$,*

$$P_n(Q_k(n)) = \sum_{1 \leqq d \leqq n^{1/k}} \mu(d)\{P_n(A_n(d^k)) - P_n(\{0\})\}.$$

PROOF. Let $p_1 < p_2 < \cdots < p_s$ be the primes less than or equal to $n^{1/k}$. Then, if $\tilde{Q}_k(n)$ denotes the complement of $Q_k(n)$ in N_n, we have

$$\tilde{Q}_k(n) = \bigcup_{i=1}^{s} A_n(p_i^k).$$

Therefore

Received by the editors on October 29, 1975, and in revised form on April 26, 1976.

$$P_n(Q_k(n)) = 1 - P_n(\tilde{Q}_k(n)) = 1 - P_n\left(\bigcup_{i=1}^{s} A_n(p_i{}^k)\right)$$

$$= 1 - \sum_{r=1}^{s} \sum_{(i_1,i_2,\cdots,i_r)} (-1)^{r-1} P_n(A_n(p_{i_1}^k)$$

$$\cap A_n(p_{i_2}^k) \cap \cdots \cap A_n(p_{i_r}^k)),$$

where the inner sum is taken over all r-tuples $(i_1, i_2, \cdots, i_r)$ such that $1 \leqq i_1 < i_2 < \cdots < i_r \leqq s$. Now it is clear that if $(d_1, d_2) = 1$, then $A_n(d_1) \cap A_n(d_2) = A_n(d_1 d_2)$. Hence this last expression can be rewritten as

$$1 + \sum_{r=1}^{s} \sum_{(i_1,i_2,\cdots,i_r)} (-1)^r P_n(A_n((p_{i_1} p_{i_2} \cdots p_{i_r})^k)).$$

Now if $(p_{i_1} p_{i_2} \cdots p_{i_r})^k > n$, $A_n((p_{i_1} p_{i_2} \cdots p_{i_r})^k) = \{0\}$. Hence this last expression is the same as

$$\sum_{1 \leqq d \leqq n^{1/k}} \mu(d) P_n(A_n(d^k)) + \sum_{r=1}^{s} \sum_{p_{i_1} p_{i_2} \cdots p_{i_r} > n^{1/k}} \mu(p_{i_1} p_{i_2} \cdots p_{i_r}) P_n(\{0\}).$$

Since

$$\sum_{d | p_1 p_2 \cdots p_s} \mu(d) = 0,$$

$$\sum_{r=1}^{s} \sum_{p_{i_1} p_{i_2} \cdots p_{i_r} > n^{1/k}} \mu(p_{i_1} p_{i_2} \cdots p_{i_r}) = - \sum_{1 \leqq d \leqq n^{1/k}} \mu(d).$$

This observation completes the proof of the lemma.

If P_n is the uniform distribution on N_n ($P_n(j) = (n+1)^{-1}$ for all $j \in N_n$) then it is easy to check that $|P_n(A_n(d)) - d^{-1}| < n^{-1}$ uniformly in d. Using this estimate along with Lemma 1 and the fact that $\sum \mu(d)\, d^{-k} \to 1/\zeta(k)$, it is not difficult to prove

$$\lim_{n \to \infty} P_n(Q_k(n)) = 1/\zeta(k)$$

for all $k \geqq 2$.

From now on P_n will always be taken to be a binomial distribution relative to some fixed parameter p with $0 < p < 1$. Thus $P_n(j) = \binom{n}{j} p^j (1-p)^{n-j}$. For $1 \leqq d \leqq n$ define $\epsilon_n(d)$ by

$$\epsilon_n(d) = P_n(A_n(d)) - d^{-1} = \sum_{j \equiv 0(d)} \binom{n}{j} p^j(1-p)^{n-j} - d^{-1}.$$

LEMMA 2. *$|\epsilon_n(d)| \ll n^{-1/2}$ uniformly in d.*

PROOF. See [1].

THEOREM 3. *If P_n is a binomial distribution, then $\lim_{n\to\infty} P_n(Q_k(n)) = 1/\zeta(k)$ for all $k \geqq 3$.*

PROOF. By Lemma 1 we have

$$\begin{aligned} P_n(Q_k(n)) &= \sum_{1 \leq d \leq n^{1/k}} \mu(d)\,\{P_n(A_n(d^k)) - P_n(\{0\})\} \\ &= \sum_{1 \leq d \leq n^{1/k}} \mu(d)\,\{d^{-k} + \epsilon_n(d^k) - (1-p)^n\} \\ &= \sum_{1 \leq d \leq n^{1/k}} \mu(d)d^{-k} + \sum_{1 \leq d \leq n^{1/k}} \mu(d)\epsilon_n(d^k) \\ &\quad - (1-p)^n \sum_{1 \leq d \leq n^{1/k}} \mu(d). \end{aligned}$$

The first sum tends to $1/\zeta(k)$ while the last sum goes to zero as $n \to \infty$. For the middle sum we have by Lemma 2

$$\left| \sum_{1 \leq d \leq n^{1/k}} \mu(d)\epsilon_n(d^k) \right| \ll n^{1/k} n^{-1/2}.$$

Thus for $k > 2$ this term goes to zero which proves the theorem.

2. In this section we show that $\lim_{n\to\infty} P_n(Q_2(n)) = 6/\pi^2 (=1/\zeta(2))$ where P_n is a binomial distribution. As in the proof of Theorem 3 it is sufficient to show that

$$\lim_{n\to\infty} \sum_{1 \leq d^2 \leq n} |\epsilon_n(d^2)| = 0.$$

We need the following lemmas. For proofs of the first two we refer to [1].

LEMMA 4.

$$\sum_{|k-pn| > pn^{3/4}} \binom{n}{k} p^k (1-p)^{n-k} \ll n^{-1}.$$

LEMMA 5. *If* $d > p(n + n^{3/4})$, *then* $|\epsilon_n(d)| \ll d^{-1}$ *uniformly in* d.

LEMMA 6. *Let* K_n *be the number of integers* d *which satisfy* $pn^{3/4} \leqq d^2 \leqq p(n - n^{3/4})$ *and which have the property that for some integer* k, kd^2 *is in the interval* $(p(n - n^{3/4}), p(n + n^{3/4}))$. *Then* $K_n \ll n^{3/8}$.

PROOF. Let $u = pn$, $v = pn^{3/4}$ and let $s = [(u + v)/v]$. Suppose $kd^2 \in (u - v, u + v)$. Then we must have $2 \leqq k \leqq s$. For each such k we ask how many possible d's are there such that $kd^2 \in (u - v, u + v)$. Such d's must lie in the interval

$$(((u - v)/k)^{1/2}, ((u + v)/k)^{1/2}).$$

Hence there are not more than $z_k = ((u + v)^{1/2} - (u - v)^{1/2})k^{-1/2} + 1$ of them. Now it is easy to verify that $(u + v)^{1/2} - (u - v)^{1/2} \leqq (2v^2/u)^{1/2}$. Therefore

$$K_n = \sum_{k=2}^{s} z_k < (2v^2/u)^{1/2} \sum_{k=2}^{s} k^{-1/2} + s - 1 \leqq v(2s/u)^{1/2} + s - 1 \ll n^{3/8}.$$

We now state and prove our main result as

THEOREM 7. *Let* P_n *be a binomial distribution. Then*

$$\lim_{n \to \infty} P_n(Q_2(n)) = 6/\pi^2.$$

PROOF. As stated at the beginning of this section we need to show

$$\lim_{n \to \infty} \sum_{1 \leqq d^2 \leqq n} |\epsilon_n(d^2)| = 0. \tag{1}$$

Let $n_1 = pn^{3/4}$, $n_2 = p(n - n^{3/4})$ and $n_3 = p(n + n^{3/4})$. The sum in (1) can then be written as

$$\sum_{1 \leqq d^2 \leqq n} = \sum_{1 \leqq d^2 \leqq n_1} + \sum_{n_1 < d^2 \leqq n_2} + \sum_{n_2 < d^2 \leqq n_3} + \sum_{n_3 < d^2 \leqq n}. \tag{2}$$

(We assume that n is large enough so that $n_1 < n_2$ and $n_3 < n$.) We will examine each of these sums separately. By Lemma 2

$$\sum_{1 \leqq d^2 \leqq n_1} |\epsilon_n(d^2)| \ll n_1^{1/2} n^{-1/2} \leqq n^{3/8} n^{-1/2} = n^{-1/4}$$

and hence the first term on the right-hand side of (2) goes to zero as $n \to \infty$. A similar argument works for the third sum on the right-hand side of (2).

By Lemma 5 $|\epsilon_n(d^2)| \ll d^{-2}$ for $d^2 > p(n + n^{3/4})$. Hence the fourth sum

$$\sum_{n_3 < d^2 \leqq n} |\epsilon_n(d^2)| \ll \sum_{n_3 < d^2 \leqq n} d^{-2} < \sum_{d=[n_3{}^{1/2}]}^{\infty} d^{-2}$$

and hence goes to zero because it is less than the tail of a convergent series.

The second sum on the right-hand side of (2) is somewhat more difficult to deal with. We break it into two parts

$$\sum_{n_1 < d^2 \leqq n_2} = \sum_{n_1 < d^2 \leqq n_2}{}' + \sum_{n_1 < d^2 \leqq n_2}{}'' \tag{3}$$

where the summation with the prime on it is taken over those d^2 which have the property that for some integer k, kd^2 is in the interval (n_2, n_3) and the double primed summation is taken over the remaining d^2. By Lemmas 2 and 6 we have

$$\sum_{n_1 < d^2 \leqq n_2}{}' |\epsilon_n(d^2)| \ll n^{3/8} n^{-1/2} = n^{-1/8}.$$

Hence the single primed sum goes to zero as $n \to \infty$. We now examine the double primed sum. Recall that

$$\epsilon_n(d^2) = \sum_{k \equiv 0(d^2)} \binom{n}{k} p^k (1-p)^{n-k} - d^{-2}.$$

For the d^2 under consideration we have by Lemma 4

$$\sum_{k \equiv 0(d^2)} \binom{n}{k} p^k(1-p)^{n-k} = \sum_{\substack{k \equiv 0(d^2) \\ |k-pn| > pn^{3/4}}} \binom{n}{k} p^k (1-p)^{n-k}$$

$$\leqq \sum_{|k-pn| > pn^{3/4}} \binom{n}{k} p^n (1-p)^{n-k} \ll n^{-1}.$$

Hence for those d^2, $|\epsilon_n(d^2)| \ll d^{-2}$. Thus for the double primed sum

$$\sum_{n_1 < d^2 \leqq n_2}{}'' |\epsilon_n(d^2)| \ll \sum_{n_1 < d^2 \leqq n_2} d^{-2} < \sum_{d=[n_1{}^{1/2}]}^{\infty} d^{-2}$$

and hence goes to zero as $n \to \infty$. This completes the proof of Theorem 7.

REFERENCE

1. J. E. Nymann and W. J. Leahey, *On the probability that integers chosen according to the binomial distribution are relatively prime,* Acta Arithmetica **31** (1976), 205-211.

THE UNIVERSITY OF TEXAS AT EL PASO, EL PASO, TEXAS 79968