

## THE PROBABILITY THAT THE NUMBER OF POINTS ON A COMPLETE INTERSECTION IS SQUAREFREE

ERIC SCHMIDT

**ABSTRACT.** We consider the asymptotic probability that integers chosen according to a binomial distribution will have certain properties: (i) that such an integer is not divisible by the  $k$ th power of a prime, (ii) that any  $k$  of  $s$  chosen integers are relatively prime and (iii) that a chosen integer is prime. We also prove an analog of the Dirichlet divisor problem for the binomial distribution. We show how these results yield corresponding facts concerning the number of points on a smooth complete intersection over a finite field.

**1. Introduction.** Let  $\alpha \in (0, 1)$ , and let  $n$  be a nonnegative integer. The probability that  $n$  Bernoulli trials with chance of success  $\alpha$  yield precisely  $t$  successes is given by the binomial distribution. In [7], Nymann and Leahey show that the probability that  $k$  integers chosen according to the binomial distribution are relatively prime approaches  $1/\zeta(k)$  as  $n \rightarrow \infty$  with  $\alpha$  fixed, where  $\zeta$  is the Riemann zeta function. The same authors, in [8], show likewise that an integer chosen according to the binomial distribution is  $k$ -free, that is, not divisible by the  $k$ th power of a prime, with probability  $1/\zeta(k)$  as  $n \rightarrow \infty$ . These results are analogous to classical results yielding the same limiting values when the integers are chosen from a uniform distribution. We prove further statements about choosing integers according to the binomial distribution. We generalize the result on  $k$ -free integers to obtain Theorem 3.1. We consider the analog of the Dirichlet divisor problem for the binomial distribution, obtaining Theorem 4.2. The result on relatively prime integers is generalized to consider choosing  $s$  integers such that any  $k$  of them are relatively prime, yielding an analog of the

---

2010 AMS *Mathematics subject classification.* Primary 11N37, Secondary 11G25.

*Keywords and phrases.* Binomial distribution, asymptotics, squarefree, divisor, complete intersection.

Received by the editors on November 13, 2015, and in revised form on June 14, 2016.

result of [6]. Finally, we prove that the probability of choosing a prime from the binomial distribution becomes arbitrarily small.

We now briefly describe our motivation in considering such questions. In [4], Bucur, et al., gave a probabilistic model for the number of points on a smooth projective plane curve of degree  $d$  over a finite field of order  $q$ . They showed, roughly, that the number of points approximately follows a binomial distribution and that the total error becomes arbitrarily small if  $d$  and  $q$  vary in an appropriate fashion. Bucur and Kedlaya generalized this result [5] showing that, in a similar manner, the number of points on a smooth, projective, complete intersection of hypersurfaces of specified degrees also follows a binomial model. Hence, our results can be used to make statements concerning the probability that randomly chosen smooth plane curves (more generally, complete intersections) possess the previously mentioned properties. For example, the probability that a smooth plane curve over  $\mathbf{F}_q$  of degree  $d$  has a squarefree number of points approaches  $1/\zeta(2)$  as  $q \rightarrow \infty$  and  $d \gg q$ . (More general and precise statements are given later.)

**2. Notation and preliminaries.** We employ the standard notation  $f(x) \sim g(x)$ , defined  $f(x)/g(x) \rightarrow 1$  as  $x \rightarrow \infty$ .

Throughout, we consider a sequence  $\{\alpha_n\}_{n \geq 0}$  with  $\alpha_n \in (0, 1)$ . We will write  $\beta_n = 1 - \alpha_n$ . We define the binomial probability mass function

$$B_{\alpha_n, n}(t) = \binom{n}{t} \alpha_n^t (1 - \alpha_n)^{n-t},$$

for nonnegative integers  $t$  and  $B_{\alpha_n, n}(t) = 0$  for negative  $t$ . Thus, the parameter  $\alpha$  varies depending upon  $n$ ; this is essential to our applications in Section 7. In order to simplify notation, we will set  $B_n(t) = B_{\alpha_n, n}(t)$  and

$$B_n(S) = \sum_{t \in S} B_n(t).$$

The binomial distribution has mean  $a_n = \alpha_n n$  and standard deviation  $\sigma_n = (n\alpha_n\beta_n)^{1/2}$ . It will be convenient to approximate the binomial distribution with a normal distribution, with probability density function  $N_n(x) = (2\pi n\alpha_n\beta_n)^{-1/2} e^{-(1/2)u_n^2(x)}$ , where  $u_n(x) = (x - a_n)/\sigma_n$ . The relationship between the binomial distribution and the normal approximation is given by the following.

**Theorem 2.1.** ([9]). *There is a constant  $\lambda_2 \in \mathbf{R}$  such that*

$$\sum_{j \in \mathbf{Z}} |B_n(j) - N_n(j)| = |\beta_n - \alpha_n| \sigma_n^{-1} \lambda_2 + O(\sigma_n^{-2}).$$

*Proof.* This is [9, Theorem 3], which also explicitly determines the value of  $\lambda_2$ .  $\square$

We will use the following generalization of a result of Nymann and Leahey [7, Lemma 3]. Intuitively, it states that the probability of a number chosen from a binomial distribution divisible by  $d$  is approximately  $1/d$ .

**Lemma 2.2.** *Let*

$$\varepsilon_n(d) = \left( \sum_{j=0}^{\infty} B_n(jd) \right) - d^{-1}.$$

*Then,  $|\varepsilon_n(d)| = O((\alpha_n \beta_n n)^{-1/2})$  uniformly in  $d$  as  $n \rightarrow \infty$ .*

*Proof.* Nymann and Leahey show, in the proof of [7, Lemma 3], that  $|\varepsilon_n(d)| \leq 3B_n(\lfloor (n+1)\alpha_n \rfloor)$ . The result follows from Theorem 2.1.  $\square$

We will call a function on the nonnegative integers *negligible* if it is  $O((\alpha_n n)^{-c})$  for all  $c \in \mathbf{R}$ . Some facts concerning negligible functions are listed next.

**Lemma 2.3.** *Let  $g(n)$  be bounded by a polynomial, and fix  $\varepsilon > 0$ . The following functions are negligible:*

- (i)  $N_n(f(n))$ , if  $\beta_n n \rightarrow \infty$  and  $|f(n) - \alpha_n n| \geq (\alpha_n n)^{1/2+\varepsilon}$  for  $n \gg 0$ ,
- (ii)  $N_n(t)$ , for a fixed  $t \in \mathbf{R}$ , if  $\alpha_n n \rightarrow \infty$ ,
- (iii)  $N_n(n)g(n)$ , if  $\beta_n n^{1-\varepsilon} \rightarrow \infty$ ,
- (iv)  $f(n)g(\alpha_n n)$ , if  $f$  is negligible,
- (v)  $f(n)g(n)$ , if  $f$  is negligible and  $\alpha_n n^{1-\varepsilon} \rightarrow \infty$ .

*Proof.*

- (i) The second hypothesis shows that

$$N_n(f(n)) \leq \sigma^{-1} \exp(-(\alpha_n n)^\varepsilon / (2\beta_n))$$

for  $n \gg 0$ . We can now verify the result if  $\beta_n \rightarrow 0$ , or if  $\beta_n$  is bounded away from 0, from which the general result follows.

(ii) For large fixed  $n$ , the function  $\alpha_n \mapsto N_n(t)$  decreases near 1. Thus, we may assume that  $\beta_n$  is bounded away from 0, and the result follows from (i).

(iii) For large fixed  $n$ , the function  $\alpha_n \mapsto N_n(n)$  increases near 0. Thus, we may assume that  $\alpha_n$  is bounded away from 0, and the result follows readily.

(iv) Obvious.

(v) This follows since  $n = o((\alpha_n n)^{1/\varepsilon})$ . □

**3. Integers that are  $k$ -free.** For  $k \geq 2$ , a positive integer is said to be  $k$ -free if it is not divisible by the  $k$ th power of a prime. Nymann and Leahey determined [8] the probability that an integer chosen according to the binomial distribution is  $k$ -free, assuming that  $\alpha$  is constant. Here, we prove the same by a different method, while allowing  $\alpha$  to vary.

**Theorem 3.1.** *Suppose that  $\alpha_n n \rightarrow \infty$ , and  $\beta_n n^{1-2/k}$  is bounded away from 0. As  $n \rightarrow \infty$ , the probability of an integer chosen according to the binomial distribution being  $k$ -free tends to  $1/\zeta(k)$ .*

*Proof.* Let  $S_k(x)$  be the set of all  $k$ -free integers at most  $x$  and  $f_k(x) = \#S_k(x)$ . Using Theorem 2.1, we have

$$B_n(S_k(n)) = \sum_{j \in S_k(n)} N_n(j) + O((\alpha_n \beta_n n)^{-1/2}).$$

Now,  $f_k(t) \sim t/\zeta(k)$  according to a classical result. Accordingly, we apply summation by parts to obtain

$$\sum_{j \in S_k(n)} N_n(j) = N_n(n)f_k(n) - \int_1^n N'_n(t) \left( \frac{t}{\zeta(k)} + R_k(t) \right) dt,$$

where  $R_k(t)$  is a remainder term to be described. The term  $N_n(n)f_k(n)$  is negligible (Lemma 2.3 (iii)). Moreover,

$$- \int_1^n N'_n(t) \frac{t}{\zeta(k)} dt = \frac{1}{\zeta(k)} \left( - (N_n(n)n - N_n(1)) + \int_1^n N_n(t) dt \right).$$

Here,  $N_n(n)n$  and  $N_n(1)$  are negligible. In order to show that

$$(3.1) \quad \lim_{n \rightarrow \infty} \int_1^n N_n(t) dt = 1,$$

perform a change of variables  $t = \sigma_n v + a_n$  to obtain

$$\frac{1}{\sqrt{2\pi}} \int_{(1-a_n)/\sigma_n}^{(n-a_n)/\sigma_n} e^{-(1/2)v^2} dv.$$

Since the restrictions on  $\alpha_n$  and  $\beta_n$  show that the limits of integration approach  $-\infty$  and  $\infty$ , respectively, as  $n \rightarrow \infty$ , (3.1) is established.

Now, we consider the remainder term. We have  $R_k(t) = O(g_k(t))$ , where  $g_k(t) = t^{1/k} \exp(-Ak^{-3/2} \log^{1/2} t)$  for some  $A$ , [10, page 213]. Therefore,

$$\begin{aligned} \left| \int_1^n N'_n(t) R_k(t) dt \right| &\ll \int_1^n |N'_n(t)| g_k(t) dt \\ &= \int_1^{a_n} N'_n(t) g_k(t) dt - \int_{a_n}^n N'_n(t) g_k(t) dt. \end{aligned}$$

Integrating by parts and ignoring negligible terms, we obtain

$$2N_n(a_n)g_k(a_n) - \int_1^n N_n(t)g'_k(t) dt.$$

The term on the left is  $O((\alpha_n \beta_n n)^{-1/2} g_k(a_n))$ , which is  $o(1)$  using  $g_k(t) = o(t^{1/k})$  together with the restrictions on  $\alpha_n$  and  $\beta_n$ . The integral on the right may be replaced with

$$\int_{\ell_n}^n N_n(t)g'_k(t) dt,$$

where  $\ell_n = a_n - a_n^{3/4}$ . (The error accrued from this replacement is negligible since the integrand is negligible on the excluded interval, and the width of the excluded interval is less than  $a_n$ .) Since  $g'_k(t) = O(t^{-1/2})$ , this latter integral is bounded above by a constant times

$$\ell_n^{-1/2} \int_{\ell_n}^n N_n(t) dt,$$

which goes to 0 since  $\ell_n$  increases without bound.  $\square$

**4. The number-of-divisors function.** The result here is not a probability calculation, but it is of a similar flavor, which will be used in Section 5. It concerns the function  $\tau(n)$ , the number of divisors of  $n$ . For a function  $f$ , we write  $f_{(1)}(x) = f(x+1) - f(x)$  for the first differences of  $f$ . Similarly,  $f_{(2)}$  denotes the first differences of  $f_{(1)}$ . In this section, we will denote  $N_n$  simply by  $N$ . We will write

$$T(x) = \sum_{1 \leq j \leq x} \tau(j).$$

Dirichlet showed that

$$T(x) = x \log x + (2\gamma - 1)x + \Delta(x),$$

where  $\Delta(x) = O(\sqrt{x})$ , and  $\gamma$  is Euler's constant [11]. We establish an analog of this for the binomial distribution. Our argument here is similar to, but more complex than, that of the previous section. First, we need the following.

**Lemma 4.1.** *If  $\alpha_n n \rightarrow \infty$  and  $\beta_n n^{1-\varepsilon} \rightarrow \infty$ , then, for  $t \in \mathbf{Z}$  and  $c \in \mathbf{R}$ ,*

$$\sum_{j=t}^n N(j) = 1 + O((\alpha_n n)^{-c}).$$

This may be proved using the Euler-Maclaurin summation formula by an argument similar to that of [3, pages 43, 44].

We now turn to the main claim of this section.

**Theorem 4.2.** *If  $\alpha_n n^{1-\varepsilon} \rightarrow \infty$  and  $\beta_n n^{1-\varepsilon} \rightarrow \infty$ , then*

$$\sum_{j=1}^n B_n(j) \tau(j) = \log(\alpha_n n) + 2\gamma + O((\alpha_n n)^{-1/4} \beta_n^{-1}).$$

*Proof.* From Theorem 2.1 and the estimate  $\tau(x) = o(x^\eta)$  for any  $\eta > 0$  [2, page 296], we have

$$\sum_{j=1}^n B_n(j) \tau(j) = \sum_{j=1}^n N(j) \tau(j) + o((\alpha_n^{-1/2} \beta_n^{-1/2} n^{-1/2+\eta})).$$

The restrictions on  $\alpha_n$  and  $\beta_n$  show that the  $o$ -term is  $O((\alpha_n\beta_n n)^{-1/4})$ . Summation by parts yields

$$\sum_{j=1}^n N(j)\tau(j) = N(n)T(n) - \sum_{j=1}^{n-1} N_{(1)}(j)T(j).$$

The term  $N(n)T(n)$  is negligible; thus, we are left to describe the asymptotics of

$$(4.1) \quad - \sum_{j=1}^{n-1} N_{(1)}(j)T(j).$$

Now, we write  $T(j)$  as  $j \log j + (2\gamma - 1)j + \Delta(j)$ . Each of the terms yields a sum to examine. First, we have

$$(4.2) \quad - \sum_{j=1}^{n-1} N_{(1)}(j)j \log j = -N(n)n \log n + \sum_{j=1}^{n-1} N(j+1)(j \log j)_{(1)}.$$

Similar to the reasoning for the previous term,  $-N(n)n \log n$  is negligible. As for the sum on the right, the mean value theorem implies that  $\log(j+1) \leq (j \log j)_{(1)} \leq \log(j+1) + 1$ . Therefore, a lower bound for this sum is

$$(4.3) \quad \sum_{j=2}^n N(j)(\log(j-1) + 1).$$

Writing  $\log(j-1) + 1$  as  $(1 + \log a_n) + (\log(j-1) - \log a_n)$ , the sum breaks down into

$$\sum_{j=2}^n N(j)(1 + \log a_n),$$

which, from Lemma 4.1, is  $1 + \log a_n$  with negligible error, and

$$\sum_{j=2}^n N(j)(\log(j-1) - \log a_n).$$

With negligible error, we contract the range of summation to  $(a_n - a_n^c, a_n + a_n^c)$  for some  $c$  in  $(1/2, 3/4)$ . (*Proof of negligibility:*  $N(j)$  is negligibly small on the excluded intervals, the logarithmic term is  $O(n)$  and the range excluded has length less than  $n$ .) From the mean value theorem,  $|\log(j-1) - \log(a_n)| \leq |(j-1) - a_n|/a_n$ . For  $j$  in the new

range of summation,  $|(j - 1) - a_n|/a_n \leq a_n^{c-1} + 1/a_n$ . Hence, the sum is  $o((\alpha_n n)^{-1/4})$ . Therefore, (4.3) converges to  $1 + \log a_n$  with error  $o((\alpha_n n)^{-1/4})$ . A similar argument shows that the upper bound

$$\sum_{j=2}^n N(j)(1 + \log j)$$

also converges to  $1 + \log a_n$  with the same error bound. Hence, (4.2) itself does.

The second sum from (4.1) to examine is

$$-\sum_{j=1}^{n-1} N_{(1)}(j)(2\gamma - 1)j = (2\gamma - 1)\left(-N(n)n + \sum_{j=1}^n N(j)\right).$$

The term  $-N(n)n$  is negligible, and thus, from Lemma 4.1, we obtain  $2\gamma - 1$  in the limit.

Thus far, we have found the main term  $\log(\alpha_n n) + 2\gamma$ . In analyzing (4.1), it remains to examine the sum

$$(4.4) \quad -\sum_{j=1}^{n-1} N_{(1)}(j)\Delta(j).$$

We show that (4.4) is  $O((\alpha_n n)^{-1/4}\beta_n^{-1})$ . In order to do this, we write it as

$$-N_{(1)}(n - 1)\Delta_2(n - 1) + \sum_{j=1}^{n-2} N_{(2)}(j)\Delta_2(j),$$

where

$$\Delta_2(j) = \sum_{i=1}^j \Delta(i).$$

As usual,  $-N_{(1)}(n - 1)\Delta_2(n - 1)$  is negligible. For the sum on the right, we use a result from [11]:

$$\Delta_2(j) = \frac{1}{2}j \log j + \left(\gamma - \frac{1}{4}\right)j + O(j^{3/4}).$$

Again, after substituting for  $\Delta_2(j)$ , we divide into cases. First, letting

$f(x) = x \log x$ , we have

$$\sum_{j=1}^{n-2} N_{(2)}(j)f(j) = N_{(1)}(n-1)f(n-1) - \sum_{j=1}^{n-2} N_{(1)}(j+1)f_{(1)}(j).$$

Here,  $N_{(1)}(n-1)f(n-1)$  is negligible, and the expression which remains is

$$(4.5) \quad -N(n)f_{(1)}(n-1) + N(2)f(2) + \sum_{j=1}^{n-2} N(j+1)f_{(2)}(j).$$

Ignoring the negligible terms, we consider the sum that remains. From the mean value theorem,

$$f_{(2)}(j) \leq \sup_{x \in [j, j+1]} f'_{(1)}(x) \leq \sup_{y \in [j, j+2]} f''(y) = 1/j.$$

(There is no ambiguity since  $(f_{(1)})' = (f')_{(1)}$ .) Hence, by restricting the range of summation, we can prove that (4.5) is  $O(1/(\alpha_n n))$ . Next,

$$\sum_{j=1}^{n-2} N_{(2)}(j)j$$

is negligible, as  $j_{(2)} = 0$ . Finally, we must consider

$$\sum_{j=1}^{n-2} N_{(2)}(j) O(j^{3/4}).$$

This is bounded above by a constant times

$$(4.6) \quad \sum_{j=1}^{n-2} |N_{(2)}(j)|j^{3/4}.$$

Now, the mean value theorem implies that  $\sup |N_{(1)}(j)| \leq \sup |N'(j)| = O(1/(\alpha_n n))$ . Moreover,  $N''$  has exactly two zeros. Therefore, if we define  $S = \{j \in \mathbf{Z} \mid \text{sgn } N_{(2)}(j) \neq \text{sgn } N_{(2)}(j+1)\}$ , then the size of  $S$  is bounded above by a constant independent of  $n$ . Now, (4.6) is bounded above by the sum of the usual negligible terms and of

$$2 \sum_{j \in S} |N_{(1)}(j+1)|j^{3/4} + \sum_{j=1}^{n-2} |N_{(1)}(j+1)|(j^{3/4})_{(1)}.$$

The left sum is  $O((\alpha_n n)^{-1/4} \beta_n^{-1})$ . Since  $(j^{3/4})_{(1)} \leq j^{-1/4}$ , the sum on the right is, besides negligible terms, bounded above by

$$2 \sum_{j \in U} N(j+1)j^{-1/4} + \sum_{j=1}^{n-2} N(j+1)(j^{-1/4})_{(1)},$$

where  $U = \{j \in \mathbf{Z} \mid \text{sgn } N_{(1)}(j) \neq \text{sgn } N_{(1)}(j+1)\}$ . Since  $N'$  has a single root at  $a_n$ , the sum on the left is  $O((\alpha_n n)^{-3/4} \beta_n^{-1/2})$ . Restricting the range of summation, we see that the sum on the right is  $O((\alpha_n n)^{-5/4})$ . This completes the proof.  $\square$

**5. Integers  $k$ -wise relatively prime.** In [12], Tóth determined the probability that an  $s$ -tuple integer is pairwise coprime. Hu [6] generalized this to the situation in which any  $k$  of the chosen integers are coprime. Despite the generality of this result, we show here that his argument can be made to apply to a binomial distribution instead of a uniform distribution. In this section, we closely follow [6], making changes when appropriate to our problem.

An  $s$ -tuple of integers is defined to be  $k$ -wise relatively prime if any  $k$  of them are relatively prime, and to be  $k$ -wise relatively prime to an integer  $u$  if any  $k$  of them, together with  $u$ , are relatively prime. (These conditions hold vacuously if  $k > s$ . This observation allows us to dispense with the multiple cases considered by Hu.) Hu found that the probability of  $s$  integers being  $k$ -wise relatively prime when chosen according to the uniform distribution is

$$A_{s,k} = \prod_p \left( 1 - \sum_{m=k}^s B_{1/p,n}(m) \right).$$

The main result of this section (Corollary 5.5) is that  $A_{s,k}$  is also the probability that  $s$  integers are  $k$ -wise relatively prime when chosen according to a binomial distribution. Note that the case  $s = k$  reproduces the results of [7].

We use notation similar to Hu's, namely, for a tuple  $\mathbf{u} = (u_1, \dots, u_{k-1})$ , let  $S_{s,k}^{(\mathbf{u})}(n)$  denote the set of  $s$ -tuples of integers  $(a_1, \dots, a_s)$  in  $[1, n]$  that are  $k$ -wise relatively prime and  $i$ -wise relatively prime to  $u_i$  for  $1 \leq i \leq k-1$ . Define  $Q_{s,k}^{(\mathbf{u})}(n) = B_n^s(S_{s,k}^{(\mathbf{u})}(n))$ . For integers  $a, b > 0$ , Hu defines  $(a, b]$  to be the product, over primes  $p$  dividing  $a$ , of the

largest power of  $p$  dividing  $b$ . Set  $[b, a) = (a, b]$ . Define, for any positive integer  $j$ ,

$$j * \mathbf{u} = \left( \frac{u_1(j, u_2)}{(j, u_1)}, \dots, \frac{u_{k-2}(j, u_{k-1})}{(j, u_{k-1})}, \frac{j u_{k-2}}{(\prod_{i=2}^{k-1} [j, u_i])(j, u_{k-1})} \right).$$

(Here  $(x, y)$  denotes  $\gcd(x, y)$ .) Importantly, if  $\mathbf{u}$  is a pairwise coprime tuple of positive integers, then so is  $j * \mathbf{u}$ .

**Lemma 5.1.** *For  $\mathbf{u}$  pairwise coprime,*

$$Q_{s+1,k}^{(\mathbf{u})}(n) = \sum_{\substack{j=1 \\ (j, u_1)=1}}^n B_n(j) Q_{s,k}^{(j*\mathbf{u})}(n).$$

*Proof.* Hu [6, page 1265] observed that, for an  $(s + 1)$ -tuple  $(\mathbf{s}, a_{s+1})$  of integers in  $[1, n]$ , we have  $(\mathbf{s}, a_{s+1}) \in S_{s+1,k}^{(\mathbf{u})}(n)$  if and only if

- (i)  $\mathbf{s}$  is  $k$ -wise relatively prime;
- (ii) for  $i \in [1, k - 2]$ , we have that  $\mathbf{s}$  is  $i$ -wise relatively prime to  $u_i$  and to  $(a_{s+1}, u_{i+1})$ ;
- (iii)  $\mathbf{s}$  is  $(k - 1)$ -wise relatively prime to  $u_{k-1}$  and to  $a_{s+1}$ ; and
- (iv)  $(a_{s+1}, u_1) = 1$ .

This justifies the equalities

$$Q_{s+1,k}^{(\mathbf{u})}(n) = \sum_{\substack{a_{s+1}=1 \\ (a_{s+1}, u_1)=1}}^n B_n(a_{s+1}) Q_{s,k}^{(a_{s+1}*\mathbf{u})}(n) = \sum_{\substack{j=1 \\ (j, u_1)=1}}^n B_n(j) Q_{s,k}^{(j*\mathbf{u})}(n),$$

where  $j * \mathbf{u} = (u_1(j, u_2), \dots, u_{k-2}(j, u_{k-1}), j u_{k-1})$ .

In order to complete the proof, we need only show that  $S_{s,k}^{(j*\mathbf{u})}(n) = S_{s,k}^{(j*\mathbf{u})}(n)$ . The argument is contained in Hu [6, pages 1265, 1266]. (He only claims that the sets have the same cardinality, but his argument shows that they are the same set.) □

**Lemma 5.2.** *For integers  $u, m \geq 1$  with  $(m, u) = 1$ , we have*

$$\sum_{\substack{a=1 \\ (a,u)=1 \\ m|a}}^n B_n(a) = \frac{\varphi(u)}{mu} + O((\alpha_n \beta_n n)^{-1/2} \theta(u)),$$

where  $\varphi$  is Euler's totient function and  $\theta(u)$  is the number of squarefree divisors of  $u$ .

*Proof.* The desired sum equals

$$\sum_{\substack{a=1 \\ m|a}}^n B_n(a) \sum_{d|(a,u)} \mu(d) = \sum_{\substack{a=1 \\ m|a}}^n B_n(a) \sum_{\substack{d|a \\ d|u}} \mu(d) = \sum_{d|u} \mu(d) \sum_{\substack{j \geq 1 \\ md|j}} B_n(j).$$

Applying Lemma 2.2, this is

$$\sum_{d|u} \mu(d) \left( \frac{1}{md} + O((\alpha_n \beta_n n)^{-1/2}) \right) = \frac{1}{m} \sum_{d|u} \frac{\mu(d)}{d} + O((\alpha_n \beta_n n)^{-1/2} \theta(u)).$$

Since

$$\sum_{d|u} \mu(d)/d = \varphi(u)/u,$$

we are done. □

**Lemma 5.3.** *Define*

$$f_{s,k,i}(u_i) = \prod_{p|u_i} \left( 1 - \frac{\sum_{m=i}^{k-1} \binom{s}{m} (p-1)^{k-1-m}}{\sum_{m=0}^{k-1} \binom{s}{m} (p-1)^{k-1-m}} \right)$$

and

$$g_{s,i}(d) = d^i \prod_{p|d} \sum_{m=0}^i \binom{s}{m} \left( 1 - \frac{1}{p} \right)^{i-m} \frac{1}{p^m}.$$

Then, we have

$$\frac{f_{s,k,i}(u_i)}{f_{s,k,i+1}(u_i)} = \sum_{d|u_i} \frac{\mu(d) \binom{s}{i} \omega(d)}{g_{s,i}(d)}, \quad i = 1, \dots, k-2,$$

and

$$f_{s,k,k-1}(u_{k-1}) = \sum_{d|u_{k-1}} \frac{\mu(d) \binom{s}{k-1}^{\omega(d)}}{g_{s,k-1}(d)},$$

where  $\omega(n)$  is the number of distinct prime factors of  $n$ . (When evaluating the above expressions we put  $0^0 = 1$ .)

*Proof.* See [6, Lemma 4]. □

**Theorem 5.4.** *Let  $\delta(s, k)$  be the maximum value of  $\binom{s-1}{i}$  for  $1 \leq i \leq k-1$ . Suppose that  $\alpha_n n^{1-\varepsilon} \rightarrow \infty$  and  $\beta_n n^{1/4} \rightarrow \infty$ . For  $s \geq 1$  and  $k \geq 2$ , uniformly in  $u_i$  with the  $u_i$  pairwise coprime, we have*

$$Q_{s,k}^{(\mathbf{u})}(n) = A_{s,k} \prod_{i=1}^{k-1} f_{s,k,i}(u_i) + O((\alpha_n \beta_n n)^{-1/2} \theta(u_1) \log^{\delta(s,k)} n).$$

*Proof.* Our proof parallels that of [6, Theorem 1]. We proceed by induction on  $s$ . For  $s = 1$ , Lemma 5.2 shows that

$$Q_{s,k}^{(\mathbf{u})}(n) = \frac{\varphi(u_1)}{u_1} + O((\alpha_n \beta_n n)^{-1/2} \theta(n)),$$

from which the result follows since  $A_{1,k} = 1$ ,  $f_{1,k,1}(u_1) = \varphi(u_1)/u_1$  and  $f_{1,k,i}(u_1) = 1$  for  $i > 1$ .

Next, we will prove the result for  $s + 1$ , assuming it for  $s$ . We obtain, using Lemma 5.1,

$$\begin{aligned} Q_{s+1,k}^{(\mathbf{u})}(n) &= \sum_{\substack{j=1 \\ (j,u_1)=1}}^n B_n(j) Q_{s,k}^{(j^* \mathbf{u})}(n) = \sum_{\substack{j=1 \\ (j,u_1)=1}}^n B_n(j) A_{s,k} \\ &\times \prod_{i=1}^{k-2} f_{s,k,i} \left( \frac{u_1(j, u_2)}{(j, u_1)} \right) f_{s,k,k-1} \left( \frac{j u_{k-1}}{(\prod_{i=2}^{k-1} [j, u_i]) (j, u_{k-1})} \right) \\ &\quad + O(B_n(j) (\alpha_n \beta_n n)^{-1/2} \theta(u_1(j, u_2)) \log^{\delta(s,k)} n) \\ (*) &= A_{s,k} \prod_{i=1}^{k-1} f_{s,k,i}(u_i) \sum_{\substack{j=1 \\ (j,u_1)=1}}^n B_n(j) \end{aligned}$$

$$\begin{aligned} &\times \prod_{i=1}^{k-2} \frac{f_{s,k,i}((j, u_{i+1}))}{f_{s,k,i+1}((j, u_{i+1}))} f_{s,k,k-1} \left( \frac{j}{\prod_{i=2}^{k-1} [j, u_i]} \right) \\ &+ O \left( (\alpha_n \beta_n n)^{-1/2} \theta(u_1) \log^{\delta(s,k)} n \sum_{j=1}^n B_n(j) \theta(j) \right). \end{aligned}$$

Using Theorem 4.2, we have

$$\sum_{j=1}^n B_n(j) \theta(j) \leq \sum_{j=1}^n B_n(j) \tau(j) = O(\log n).$$

We also have

$$\begin{aligned} &\sum_{\substack{j=1 \\ (j, u_1)=1}}^n B_n(j) \prod_{i=1}^{k-2} \frac{f_{s,k,i}((j, u_{i+1}))}{f_{s,k,i+1}((j, u_{i+1}))} f_{s,k,k-1} \left( \frac{j}{\prod_{i=2}^{k-1} [j, u_i]} \right) \\ &= \sum_{\substack{j=1 \\ (j, u_1)=1}}^n B_n(j) \prod_{i=1}^{k-2} \sum_{d_i | (j, u_{i+1})} \frac{\mu(d_i) \binom{s}{i}^{\omega(d_i)}}{g_{s,i}(d_i)} \\ &\quad \times \sum_{d_{k-1} | j / (\prod_{i=2}^{k-1} [j, u_i])} \frac{\mu(d_{k-1}) \binom{s}{k-1}^{\omega(d_{k-1})}}{g_{s,k-1}(d_{k-1})} \\ &= \sum_{\substack{d_1 \cdots d_{k-1} e = j \leq n \\ d_i | (j, u_{i+1}), 1 \leq i \leq k-2 \\ d_{k-1} | j / ([j, u_2] \cdots [j, u_{k-1}]) \\ (j, u_1)=1}} B_n(j) \prod_{i=1}^{k-1} \frac{\mu(d_i) \binom{s}{i}^{\omega(d_i)}}{g_{s,i}(d_i)} \\ &= \sum_{\substack{d_1 \cdots d_{k-1} \leq n \\ d_i | u_{i+1}, 1 \leq i \leq k-2 \\ (d_{k-1}, u_i)=1, 1 \leq i \leq k-1}} \sum_{\substack{e \leq n / (d_1 \cdots d_{k-1}) \\ (e, u_1)=1}} B_n(d_1 \cdots d_{k-1} e) \prod_{i=1}^{k-1} \frac{\mu(d_i) \binom{s}{i}^{\omega(d_i)}}{g_{s,i}(d_i)}. \end{aligned}$$

Using Lemma 5.2, we have

$$\sum_{\substack{j=1 \\ (j, u_1)=1}}^n B_n(j) \prod_{i=1}^{k-2} \frac{f_{s,k,i}((j, u_{i+1}))}{f_{s,k,i+1}((j, u_{i+1}))} f_{s,k,k-1} \left( \frac{j}{\prod_{i=2}^{k-1} [j, u_i]} \right)$$

$$\begin{aligned}
 &= \sum_{\substack{d_1 \cdots d_{k-1} \leq n \\ d_i | u_{i+1}, 1 \leq i \leq k-2 \\ (d_{k-1}, u_i) = 1, 1 \leq i \leq k-1}} \prod_{i=1}^{k-1} \frac{\mu(d_i) \binom{s}{i}^{\omega(d_i)}}{g_{s,i}(d_i)} \\
 &\times \left( \frac{\varphi(u_1)}{u_1 d_1 \cdots d_{k-1}} + O((\alpha_n \beta_n n)^{-1/2} \theta(u_1)) \right) \\
 &= \frac{\varphi(u_1)}{u_1} \sum_{\substack{d_1 \cdots d_{k-1} \leq n \\ d_i | u_{i+1}, 1 \leq i \leq k-2 \\ (d_{k-1}, u_i) = 1, 1 \leq i \leq k-1}} \prod_{i=1}^{k-1} \frac{\mu(d_i) \binom{s}{i}^{\omega(d_i)}}{d_i g_{s,i}(d_i)} \\
 &+ O\left( (\alpha_n \beta_n n)^{-1/2} \theta(u_1) \sum_{d \leq n} \frac{\delta(s+1, k)^{\omega(d)}}{d} \right),
 \end{aligned}$$

since  $g_{s,i}(d_i) \geq d_i$ .

If we define

$$h_{s,i}(p) = \left( 1 - \frac{\binom{s}{i-1}}{p \sum_{m=0}^{i-1} \binom{s}{m} (p-1)^{i-1-m}} \right),$$

we have

$$\begin{aligned}
 &\frac{\varphi(u_1)}{u_1} \sum_{\substack{d_i | u_{i+1}, 1 \leq i \leq k-2 \\ (d_{k-1}, u_i) = 1, 1 \leq i \leq k-1}} \prod_{i=1}^{k-1} \frac{\mu(d_i) \binom{s}{i}^{\omega(d_i)}}{d_i g_{s,i}(d_i)} \\
 &= \frac{\varphi(u_1)}{u_1} \prod_{i=2}^{k-1} \prod_{p | u_i} h_{s,i}(p) \prod_{p \nmid u_1 \cdots u_{k-1}} h_{s,k}(p) \\
 &= \prod_{i=1}^{k-1} \prod_{p | u_i} h_{s,i}(p) (h_{s,k}(p))^{-1} \prod_p h_{s,k}(p),
 \end{aligned}$$

together with the error terms

$$O\left( \sum_{d > n} \frac{\delta(s+1, k)^{\omega(d)}}{d^2} \right) = O\left( \sum_{d > n} \frac{\tau_{\delta(s+1, k)}(d)}{d^2} \right) = O(n^{-1} \log^{\delta(s+1, k)-1} n),$$

from [12, Lemma 3 (b)], and

$$\begin{aligned} & O\left((\alpha_n\beta_n n)^{-1/2}\theta(u_1)\sum_{d\leq n}\frac{\delta(s+1,k)\omega(d)}{d}\right) \\ &= O\left((\alpha_n\beta_n n)^{-1/2}\theta(u_1)\sum_{d\leq n}\frac{\tau_{\delta(s+1,k)}(d)}{d}\right) \\ &= O((\alpha_n\beta_n n)^{-1/2}\theta(u_1)\log^{\delta(s+1,k)}n), \end{aligned}$$

from [12, Lemma 3 (a)]. We substitute into (\*) to obtain

$$\begin{aligned} Q_{s+1,k}^{(\mathbf{u})}(n) &= A_{s,k}\prod_p h_{s,k}(p)\prod_{i=1}^{k-1} f_{s,k,i}(u_i)\prod_{p|u_i} h_{s,i}(p)(h_{s,k}(p))^{-1} \\ &\quad + O(n^{-1}\log^{\delta(s+1,k)-1}n) \\ &\quad + O((\alpha_n\beta_n n)^{-1/2}\theta(u_1)\log^{\delta(s+1,k)}n) \\ &\quad + O((\alpha_n\beta_n n)^{-1/2}\theta(u_1)\log^{\delta(s,k)+1}n) \\ &= A_{s+1,k}\prod_{i=1}^{k-1} f_{s+1,k,i}(u_i) + O((\alpha_n\beta_n n)^{-1/2}\theta(u_1)\log^{\delta(s+1,k)}n). \end{aligned}$$

This establishes the claim for  $s + 1$ . □

**Corollary 5.5.** *If  $\alpha_n n^{1-\varepsilon} \rightarrow \infty$  and  $\beta_n n^{1/4} \rightarrow \infty$ , then the probability that  $s$  integers chosen according to the binomial distribution are  $k$ -wise relatively prime approaches  $A_{s,k}$ , as  $n \rightarrow \infty$ .*

**6. Prime numbers.** Let  $\Pi$  be the set of all prime numbers. Here, we seek information on the behavior of  $B_n(\Pi)$ , as  $n \rightarrow \infty$ . We can show that  $B_n(\Pi) \rightarrow 0$ . If we use the prime number theorem, we can deduce a bit more, as follows.

**Theorem 6.1.** *If  $\alpha_n n \rightarrow \infty$  and  $\beta_n n^{1-\varepsilon} \rightarrow \infty$ , then*

$$\limsup_{n \rightarrow \infty} B_n(\Pi) \log \log(\alpha_n n) \leq 1.$$

*Proof.* (The germ of this proof is found in [1, pages 101–103].) Let  $x_n = (\alpha_n n)^{1/2}$ . For any  $j$ , denote the primorial, the product of all

primes at most  $j$ , by  $j\#$ . Let  $p_m$  denote the  $m$ th prime. For any  $n$ , let  $h(n)$  be such that  $y_n = p_{h(n)}\#$  is the least primorial not less than  $x_n$ . Now, for  $m \gg 0$ , we have  $p_m\# > p_{m+1}^2$  [10, page 246]. Therefore, for  $n \gg 0$ , we have  $p_{h(n)} < x_n^{1/2}$ ; thus,  $y_n < x_n^{3/2}$ . Write

$$B_n(\Pi) = \sum_{2 \leq p \leq y_n} B_n(p) + \sum_{p > y_n} B_n(p).$$

Using Theorem 2.1, we obtain

$$\sum_{2 \leq p \leq y} B_n(p) \leq \sum_{j=2}^y B_n(j) = \sum_{j=2}^y N_n(j) + O((\alpha_n \beta_n n)^{-1/2}).$$

The sum on the right is bounded by  $y_n N_n(y_n)$ . Since  $y_n < a_n^{3/4}$ , we see that  $y_n N_n(y_n)$  is negligible.

Now, for primes  $p > y_n$ , we have  $(p, y_n) = 1$ . Hence, from Lemma 5.2,

$$\sum_{p > y_n} B_n(p) \leq \sum_{\substack{j=1 \\ (j, y_n)=1}}^n B_n(j) = \frac{\varphi(y_n)}{y_n} + O((\alpha_n \beta_n n)^{-1/2} \theta(y_n)).$$

We know that  $\theta(y_n) = o(y_n^\eta)$ , for all  $\eta > 0$ , so the  $O$ -term is  $o(1)$ . According to one version of the prime number theorem [2, page 79],  $p_{h(n)} \sim \log y_n$ ; thus,

$$\begin{aligned} \frac{\varphi(y_n)}{y_n} &= \prod_{i=1}^{h(n)} \left(1 - \frac{1}{p_i}\right) \leq \left(\sum_{j=1}^{p_{h(n)}} \frac{1}{j}\right)^{-1} \sim \frac{1}{\log p_{h(n)}} \\ &\sim \frac{1}{\log \log y_n} < \frac{1}{\log \log x_n} \sim \frac{1}{\log \log(\alpha_n n)}. \end{aligned} \quad \square$$

**7. Applications.** Here, we show how the results obtained in the previous sections may be applied to complete intersections over finite fields. Consider an integer  $m \geq 1$ . Let  $S_d$  be the set of homogeneous polynomials  $F(X_0, \dots, X_m)$  of degree  $d$  over  $\mathbf{F}_q$ . For  $F \in S_d$ , let  $H_F$  denote the hypersurface  $F(X_0, \dots, X_m) = 0$ . For a prime power  $q$ , we let

$$\nu_q = \#\mathbf{P}^m(\mathbf{F}_q) = 1 + q + q^2 + \dots + q^m.$$

We also denote by  $p$  the characteristic of  $\mathbf{F}_q$ . The key theorem is the following.

**Theorem 7.1.** *Let  $1 \leq j \leq m$  be an integer, and consider tuples  $\mathbf{d} = (d_1, \dots, d_j)$  of positive integers such that  $(m+1)\nu_q < d_1 \leq \dots \leq d_j$ . For any*

$$\mathbf{f} = (f_1, \dots, f_j) \in S_{d_1} \times \dots \times S_{d_j},$$

let

$$H_{\mathbf{f}} = H_{f_1} \cap \dots \cap H_{f_j}.$$

Consider some  $R \subseteq \mathbf{P}^m(\mathbf{F}_q)$  of size  $t$ . Suppose  $q, d_1, \dots, d_j$  vary such that  $d_1 \rightarrow \infty$  and  $d_j = o((q^{d_1/\max(m+1,p)})^{1/m})$ . Then, the probability that a smooth  $H_{\mathbf{f}}$  of dimension  $m - j$  contains the points of  $R$  but no other point of  $\mathbf{P}^m(\mathbf{F}_q)$  is

$$\left( \frac{q^{-j}L(q, m, j)}{1 - q^{-j} + q^{-j}L(q, m, j)} \right)^t \left( \frac{1 - q^{-j}}{1 - q^{-j} + q^{-j}L(q, m, j)} \right)^{\nu_q - t} + O((d_1 - (m + 1)\nu_q + 1)^{-(2j-1)/m} + d_j^m q^{-d_1/\max(m+1,p)}),$$

where

$$L(q, m, j) = \prod_{i=1}^{j-1} (1 - q^{-(m-i)}).$$

*Proof.* This follows from [5, Theorem 1.2 and Corollary 1.3], cf., the discussion preceding Section 3 on pages 551, 552. Note that there is a small error in the statement of Corollary 1.3. In order to ensure smoothness at all points of the complete intersection, we must take  $z$  to be at least  $(m + 1)g + h$  when applying Theorem 1.2. Thus,  $g$  should be replaced with  $(m + 1)g$  in the error term. □

Theorem 7.1 states, in essence, that the distribution of the number of points on a smooth complete intersection over a finite field  $\mathbf{F}_q$  is approximately a binomial distribution with parameter

$$\alpha_q = (q^{-j}L(q, m, j))/(1 - q^{-j} + q^{-j}L(q, m, j)).$$

When applying this theorem, restrictions need to be made on the behavior of the degrees of the hypersurface sections relative to the order  $q$  of the field such that the error term goes to 0.

**Theorem 7.2.** Fix  $m \geq 2$  and  $1 \leq j \leq m - 1$ . Suppose that  $\{q_i\}_{i \geq 1}$  is a sequence of prime powers increasing to infinity, and suppose the integers  $d_{i,1} \leq \dots \leq d_{i,j}$  go to infinity in such a way that

$$d_{i,1}/2^{m\nu_{q_i}/(2j-1)} \longrightarrow \infty$$

and

$$d_{i,j} = o((2^{-\nu_{q_i}} q_i^{d_{i,1}/\max(m+1,p)})^{1/m}).$$

For a fixed  $k \geq 2$ , the probability that a smooth complete intersection (formed by intersecting hypersurfaces of degrees  $d_{i,1}, \dots, d_{i,j}$ ) has the number of points  $k$ -free is, in the limit,  $1/\zeta(k)$ .

*Proof.* For any  $n \geq q_1$ , let the integer  $r(n)$  be maximal such that  $\nu_{q_{r(n)}}$  does not exceed  $n$ . Let

$$\alpha_n = (q_{r(n)}^{-j} L(q_{r(n)}, m, j)) / (1 - q_{r(n)}^{-j} + q_{r(n)}^{-j} L(q_{r(n)}, m, j)).$$

We have  $\alpha_n = \Omega(n^{-j/m})$ , and thus, Theorem 3.1 tells us that, as  $n \rightarrow \infty$ , the probability of integers chosen according to the binomial distribution being  $k$ -free approaches  $1/\zeta(k)$ . From Theorem 7.1, this is also true of the number of points on a smooth complete intersection since our hypotheses ensure that the error term in the theorem (multiplied by  $2^{\nu_{q_{r(i)}}$ , the maximum number of choices for  $R$  in the theorem) goes to 0. □

**Theorem 7.3.** Fix  $m \geq 2$ ,  $s \geq 2$ ,  $k \geq 2$  and  $1 \leq j \leq m - 1$ . Suppose that  $\{q_i\}_{i \geq 1}$  is a sequence of prime powers increasing to infinity, and suppose that the integers  $d_{i,1} \leq \dots \leq d_{i,j}$  go to infinity in such a way that

$$d_{i,1}/2^{m\nu_{q_i}/(2j-1)} \longrightarrow \infty$$

and

$$d_{i,j} = o((2^{-\nu_{q_i}} q_i^{d_{i,1}/\max(m+1,p)})^{1/m}).$$

The probability that  $s$  smooth complete intersections  $H_1, \dots, H_s$  (formed by intersecting hypersurfaces of degrees  $d_{i,1}, \dots, d_{i,j}$ ) have the numbers of points on  $H_1, \dots, H_s$  to be  $k$ -wise relatively prime is, in the limit, the number  $A_{s,k}$  defined in Section 5.

*Proof.* Similar to the previous theorem. □

We may also prove a similar theorem based on Theorem 6.1. (We do not give the full statement.) The restrictions on  $q_i, d_{i,1}, \dots, d_{i,j}$  in these theorems heavily depend upon the error term in [5, Theorem 1.2]. If a better error term were found, this could imply a relaxation on these restrictions.

**Acknowledgments.** This paper is based on a portion of my dissertation, written under the direction of Prof. Jeff Achter. I would also like to thank him for helping to prepare this paper for submission.

#### REFERENCES

1. George E. Andrews, *Number theory*, Dover, New York, 1994.
2. Tom M. Apostol, *Introduction to analytic number theory*, Springer-Verlag, New York, 1976.
3. N.G. de Bruijn, *Asymptotic methods in analysis*, North-Holland Publishing Co., Amsterdam, 1961.
4. Alina Bucur, Chantal David, Brooke Feigon and Matilde Lalín, *Fluctuations in the number of points on smooth plane curves over finite fields*, J. Num. Th. **130** (2010), 2528–2541.
5. Alina Bucur and Kiran S. Kedlaya, *The probability that a complete intersection is smooth*, J. Th. Nomb. Bordeaux **24** (2012), 541–556.
6. Jerry Hu, *The probability that random positive integers are  $k$ -wise relatively prime*, Int. J. Num. Th. **9** (2013), 1263–1271.
7. J.E. Nymann and W.J. Leahey, *On the probability that integers chosen according to the binomial distribution are relatively prime*, Acta Arith. **31** (1976), 205–211.
8. ———, *On the probability that an integer chosen according to the binomial distribution be  $k$ -free*, Rocky Mountain J. Math. **7** (1977), 769–774.
9. Yu.V. Prohorov, *Asymptotic behavior of the binomial distribution*, Select. Transl. Math. Stat. Prob. **1** (1961), 87–95.
10. J. Sándor, B. Crstici and Dragoslav S. Mitrinović, *Handbook of number theory I*, Kluwer Academic, Dordrecht, 2006.
11. Sanford Segal, *A note on the average order of number-theoretic error terms*, Duke Math. J. **32** (1965), 279–284.
12. László Tóth, *The probability that  $k$  positive integers are pairwise relatively prime*, Fibonacci Quart. **1** (2002), 13–18.

22433 NE MARKETPLACE DRIVE, #N1098, REDMOND, WA 98053

**Email address:** [eric41293@comcast.net](mailto:eric41293@comcast.net)