ON THE DISTRIBUTION OF THE GREATEST COMMON DIVISOR

BY

PERSI DIACONIS and PAUL ERDOS

TECHNICAL REPORT NO. 12
OCTOBER 10, 1977

PREPARED UNDER GRANT
DAAG29-77-G-0031
FOR THE U.S. ARMY RESEARCH OFFICE

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government
Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, UNLESS SO DESIGNATED BY OTHER AUTHORIZED DOCUMENTS.
1. Introduction and Statement of Main Results.

Let M and N be random integers chosen uniformly and independently from \([1, 2, \ldots, x]\). Throughout \((M, N)\) will denote the greatest common divisor and \([M, N]\) the least common multiple. Cesáro (1885) studied the moments of \((M, N)\) and \([M, N]\). Theorems 1 and 2 extend his work by providing explicit error terms. The distribution of \((M, N)\) and \([M, N]\) is given by:

**Theorem 1.**

\[
\begin{align*}
(1) \quad & P_x([M, N] \leq t x^2 \text{ and } (M, N) = k) = \frac{6}{x^2} \frac{1}{k^2} (kt(1-\log kt)) + O_t \left( \frac{\log x}{x} \right) \\
(2) \quad & P_x((M, N) = k) = \frac{6}{x^2} \frac{1}{k^2} + O \left( \frac{\log k}{xk} \right) \\
(3) \quad & P_x([M, N] < t x^2) = 1 + \frac{6}{x^2} \cdot \sum_{j=1}^{1/t} \lfloor t(1-\log jt) - 1 \rfloor + O_t \left( \frac{\log x}{x} \right).
\end{align*}
\]

Where \([x]\) denotes the greatest integer less than or equal to \(x\).

Christopher (1956) gave a weaker form of (2).

(2) easily yields an estimate for the expected value of \((M, N)\):

\[
E_x((M, N)) = \frac{1}{x^2} \sum_{i, j \leq x} (i, j) = \sum_{k \leq x} k P_x((M, N) = k) = \frac{6}{x^2} \log x + O(1).
\]

(2) does not lead to an estimate for higher moments of \((M, N)\). Similarly the form of (3) makes direct computation of moments of \([M, N]\) unwieldy.

Using elementary arguments we will show:
Theorem 2.

(4) \[ E_x[(M,N)] = \frac{6}{\pi^2} \log x + C + O\left(\frac{\log x}{\sqrt{x}}\right) \]

where \( C \) is an explicitly calculated constant.

(5) \[ \text{for } k \geq 2, \quad E_x[(M,N)^k] = \frac{x^{k-1}}{k+1} \left( \frac{\zeta(k)}{\zeta(k+1)} - 1 \right) + O(x^{k-2} \log x) \]

where \( \zeta(z) \) is Riemann's zeta function,

(6) \[ \text{for } k \geq 1, \quad E_x[(M,N)^k] = \frac{x^{2k}}{\zeta(2)(k+1)^2} + O(x^{2k-1} \log x) \]

Section two of this paper contains proofs while section three contains remarks, further references and an application to the statistical problem of reconstructing the sample size given a table of rounded percentages.

2. Proofs of Main Theorems.

Throughout we use the elementary estimate

(2.1) \[ \Phi(x) = \sum_{1 \leq k \leq x} \Phi(k) = \frac{3}{\pi^2} x^2 + R(x) \]

where \( R(x) = O(x \log x) \).

See for example Hardy and Wright (1960) theorem 330. Since

\[ \# \{m,n \leq x: (m,n) = 1\} = 2\Phi(x) + 0(1) \quad \text{and} \quad (m,n) = k \quad \text{if and only if} \quad k|m, \quad k|n \quad \text{and} \quad (m/n, n/k) = 1, \]

we see that \( \# \{m,n \leq x: (m,n) = k\} = 2\Phi\left(\frac{x}{k}\right) + O(1) \).

This proves (2). To prove (1) and (3) we need a preparatory lemma.
Lemma 1. If \( F_x(t) = \#\{m, n \leq x: mn \leq tx^2 \text{ and } (m, n) = 1\} \), then

\[
F_x(t) = \frac{6}{\pi^2} t(1 - \log t)x^2 + O_t(x \log x).
\]

Proof. Consider the number of lattice points in the region \( R_x(t) = \{m, n \leq x: mn \leq tx^2\} \). It is easy to see that there are \( t(1 - \log t)x^2 + O_t(x) = N_x(t) \) such points. Also, the pair \( (m, n) \in R_x(t) \) and \( (m, n) = k \) if and only if \( \left(\frac{m}{k}, \frac{n}{k}\right) \in R_{x/k}(t) \) and \( \left(\frac{m}{k}, \frac{n}{k}\right) = 1 \). Thus

\[
N_x(t) = \sum_{1 \leq d \leq x} F_x/d(t).\]

The standard inversion formula says

\[
F_x(t) = \sum_{1 \leq d \leq x} \mu(d)N_x/d(t) = \frac{6}{\pi^2} t(1 - \log t)x^2 + O_t(x \log x).
\]

Lemma 1 immediately implies that the product of 2 random integers is independent of their greatest common divisor:

Corollary 1.

\[
P_x\{[M N] \leq tx^2 \mid (M, N) = k\} = t(1 - \log t) + O_t k \left(\frac{\log x}{x}\right).
\]

To prove (1) note that

\[
P_x\{[M N] \leq tx^2 \text{ and } (M, N) = k\} = P_x\{[M N] \leq tx^2 \mid (M, N) = k\} \cdot P_x\{ (M, N) = k\}
\]

\[
= P_x\{MN \leq \frac{t}{k} x^2 \mid (M, N) = k\} \cdot P_x\{ (M, N) = k\}.
\]

Use of (2) and Corollary 1 completes the proof of (1). To prove (3) note that
\[
P_X\left(\left[M, N\right] \leq tx^2\right) = P_X\left(\left[M, N\right] > \left[\frac{1}{t}\right]\right) + \sum_{k=1}^{[1/t]} P_X\left(\left[M, N\right] \leq tx^2 \mid \left[M, N\right] = k\right) \cdot P_X\left(\left[M, N\right] = k\right).
\]

Using (2) and Corollary 1 as before completes the proof of Theorem 1.

To prove Theorem 2, write, for \( k \geq 1, \)

\[
\tag{2.2}
\sum_{(m, n) \leq x} (m, n)^k = 2 \sum_{1 \leq m \leq x} \sum_{1 \leq n \leq m} (m, n)^k - \sum_{1 \leq i \leq x} i^k
\]

\[
= 2 \sum_{1 \leq m \leq x} f_k(m) - \frac{x^{k+1}}{k+1} + O(x^k)
\]

where \( f_k(m) = \sum_{d \mid m} \varphi(d)^k n_d) \). Dirichlet's Hyperbole argument (see e.g. Saffari (1970)) yields for any \( t, \)

\[
\tag{2.3}
\sum_{1 \leq m \leq x} f_k(m) = \sum_{1 \leq i \leq t} i^k \frac{x}{I_k} + \sum_{1 \leq i \leq x/t} \varphi(1)I_k \left( \frac{x}{t} \right) - I_k(t) \phi(\frac{x}{t})
\]

where

\[
I_k(t) = \sum_{1 \leq i \leq t} i^k = \frac{t^{k+1}}{k+1} + O(t^k).
\]

When \( k = 1, \) we proceed as follows: Choose \( t = \sqrt{x} \). The first sum on the right side of (2.3) is,

\[
\tag{2.4}
\sum_{1 \leq k \leq \sqrt{x}} \left\{ \frac{3}{\pi^2} \left( \frac{x}{k} \right)^2 + O\left( \frac{x}{k} \log \frac{x}{k} \right) \right\}
\]

\[
= \frac{3}{\pi^2} x^2 \left( \log \sqrt{x} + \gamma + O\left( \frac{1}{\sqrt{x}} \right) \right) + O(x^{3/2} \log x).
\]

\[
\]
The second sum in (2.3) is

\[(2.5) \quad \sum_{1 \leq k \leq \sqrt{x}} \varphi(k)\left(\frac{1}{k} \left(\frac{k}{x}\right)^2 + o\left(\frac{1}{k}\right)\right) = \frac{x^2}{2} \sum_{1 \leq k \leq \sqrt{x}} \frac{\varphi(k)}{k^2} + o(x^{3/2}).\]

Now

\[\sum_{1 \leq k \leq \sqrt{x}} \frac{\varphi(k)}{k^2} = \sum_{1 \leq k \leq \sqrt{x}} \frac{2k+1}{k(k+1)^2} \Phi(k) + \frac{\Phi(\frac{3}{2})}{\sqrt{x}}\]

\[= 2 \sum_{1 \leq k \leq \sqrt{x}} \frac{1}{k(k+1)^2} \left(\frac{3}{2} k^2 + R(k)\right) + \sum_{1 \leq k \leq \sqrt{x}} \frac{\varphi(k)}{k^2 (k+1)^2}\]

\[= \frac{3}{x^2} \log x + d + o\left(\frac{1}{\sqrt{x}}\right),\]

where

\[(2.6) \quad d = \sum_{k=1}^{\infty} \left(\varphi(k) + 2kR(k) - \frac{6}{\pi^2} k(2k+1)/(k(k+1))^2\right) + \frac{6}{\pi^2} (\gamma + \frac{1}{2})\]

and \(\gamma\) is Euler's constant.
Using this in equation (2.5) yields that the second sum in (2.3) is

\[ (2.7) \quad \frac{3x^2}{2\pi^2} \log x + \frac{a}{2} x^2 + O(x^{3/2} \log x). \]

The third term in (2.3) is

\[ (2.8) \quad \frac{1}{2} \frac{3}{\pi^2} x^2 + O(x^{3/2} \log x). \]

Combining (2.8), (2.7) and (2.4) in (2.3) and using this in (2.2) yields:

\[ \sum_{m,n} (m,n) = \frac{6}{\pi^2} x^2 \log x + (d + \frac{6}{\pi^2} (y + \frac{1}{2}) - \frac{1}{2}) x^2 + O(x^{3/2} \log x), \]

where \( d \) is defined in (2.6).

When \( k \geq 2 \), the best choice of \( t \) in (2.3) is \( t = 1 \). A calculation very similar to the case of \( k = 1 \) leads to (3).

We now prove (6). Consider the sum

\[ (2.9) \quad \sum_{i,j \leq x} [i,j]^k = 2 \sum_{i \leq x} \sum_{j \leq i} [i,j]^k + O(x^{k+1}) \]

\[ = 2 \sum_{i \leq x} \sum_{d|i} \sum_{j \leq 1} \left( \frac{i}{d} \right)^k + O(x^{k+1}) \]

\[ = 2 \sum_{i \leq x} \sum_{d|i} \mu(d) \left( \frac{i}{d} \right)^k + O(x^{k+1}) = 2 \sum_{d=1}^x \sum_{j \leq x/d} j^k f_k(j) + O(x^{k+1}). \]

Where
We may derive another expression for $f_k(n)$ by considering the sum

$$
\sum_{i=1}^{n} i^k = \frac{n^{k+1}}{k+1} + R_k(n) = n^k \sum_{d|n} \frac{f_k(d)}{d^k}.
$$

Dividing (2.10) by $n^k$ and inverting yields

$$
\frac{f_k(n)}{n^k} = \frac{1}{k+1} \sum_{d|n} \mu\left(\frac{n}{d}\right) d + \sum_{d|n} \mu\left(\frac{n}{d}\right) \frac{R_k(d)}{d^k}.
$$

or

$$f_k(n) = \frac{n^k}{k+1} \varphi(n) + \sum_{d|n} \mu\left(\frac{n}{d}\right) \left(\frac{n}{d}\right)^k R_k(d) = \frac{n^k \varphi(n)}{k+1} + E(n).
$$

When we substitute this expression for $f_k(j)$ in (2.9) we must evaluate:

$$S_1(y) = \sum_{j \leq y} j^k E(j) = \sum_{j \leq y} j^k \sum_{d|j} \mu\left(\frac{j}{d}\right) \left(\frac{n}{d}\right)^k R_k(d)
$$

$$= \sum_{d \leq x} R_k(d) d^k \sum_{d \leq y/d} \mu(d) i^{2k} = \sum_{d \leq y/d} R_k(d) d^k.
$$

Now $R_k(d)$ is a polynomial in $d$ of degree $k$. Thus,

$$|S_1(y)| \leq \sum_{i \leq y} i^{2k} \left(\frac{y}{i}\right)^{2k+1} = O(y^{2k+1} \log y).$$
We must also evaluate

\[ S_2(y) = \frac{1}{k+1} \sum_{j \leq y} j^k \varphi(j) = \frac{1}{k+1} \left( 2k \sum_{j \leq y} -j^{2k-1} \varphi(j) + o \left( \sum_{j \leq y} j^{2k-2} \varphi(j) \right) \right) \varphi(y) y^{2k} \]

\[ - \frac{6}{\pi^2} \frac{k}{(k+1)} y^{2k+2} + \frac{3}{\pi^2} \frac{1}{(k+1)} y^{2k+2} + o(y^{2k+1} \log y) \]

\[ = \frac{6}{\pi^2 (k+1)} \left( \frac{1}{2} - \frac{k}{2k+2} \right) y^{2k+2} + o(y^{2k+1} \log y) = \frac{3}{\pi^2 (k+1)^2} y^{2k+2} + o(y^{2k+1} \log y) . \]

Substituting in the right side of (2.9) we have

\[ \sum_{i,j \leq x} [i,j]^k = 2 \sum_{d=1}^x d^k \left( S_1 \left( \frac{x}{d} \right) + S_2 \left( \frac{x}{d} \right) \right) + o(x^{k+1}) \]

\[ = \frac{6}{\pi^2 (k+1)^2} x^{2k+2} \sum_{d=1}^x \frac{1}{d^{k+2}} + o(x^{2k+1} \log x) \]

\[ = \frac{\zeta(k+2)}{\zeta(2)} \frac{x^{2k+2}}{(k+1)^2} + o(x^{2k+1} \log x) . \]


1. If \( M_1, M_2, \ldots, M_k \) are random integers chosen uniformly at random then results stated in Christopher (1956) (see also Cohen (1960), Herzog and Stewart (1971), and Neymann (1972)) imply that

\[ P_x \left( (M_1, M_2, \ldots, M_k) = j \right) = \frac{1}{\zeta(k)} \frac{1}{j^k} + o \left( \frac{1}{x^{j-k-1}} \right) k \geq 3. \]

We have not tried to extend theorems 1 and 2 to the \( k \)-dimensional case.
3.1 has an application to a problem in applied statistics. Suppose a population of \( n \) individuals is distributed into \( k \) categories with \( n \) individuals in category \( i \). Often only the proportions \( p_i = \frac{n_i}{n} \) are reported. A method for estimating \( n \) given \( p_i \), \( 1 \leq i \leq k \) is described in Wallis and Roberts (1956), pgs. 184-189. Briefly, let \( m = \min |\sum_{i=1}^{k} p_i b_i| \) where the minimum is taken over all \( k \) tuples \((b_1, b_2, \ldots, b_k)\), with \( b_i \in \{0, \pm 1, \pm 2, \ldots\} \) not all \( b_i \) equal zero. An estimate for \( n \) is \([1/m]\). This method works if the \( p_i \) are reported with enough precision and the \( n_i \) are relatively prime for then the Euclidean algorithm implies there are integers \((b_i)_{i=1}^{k}\) such that \( \sum b_i n_i = 1 \). These \( b_i \) give the minimum \( m = \frac{1}{n} \). If it is reasonable to approximate the \( n_i \) as random integers then (3.1) implies that \( \text{Prob}(n_1 n_2 \cdots n_k = 1) = \frac{1}{\zeta(k)} \) and, as expected, as \( k \) increases this probability goes to 1. For example, \( \frac{1}{\zeta(5)} \approx .964 \), \( \frac{1}{\zeta(7)} \approx .992 \), \( \frac{1}{\zeta(9)} \approx .998 \). This suggests the method has a good chance of working with a small number of categories. Wallace and Roberts (1956) give several examples and further details about practical implementation.

2. The best result we know for \( R(x) \) defined in (2.1) is due to Saltykov (1960). He shows that

\[
R(x) = O(x (\log x)^{2/3} (\log \log x)^{1+\varepsilon}).
\]

Use of this throughout leads to a slight improvement in the bounds of theorems 1 and 2.

3. The functions \((M,N)\) and \([M,N]\) are both multiplicative in the sense of Delange (1969,1970). It would be of interest to derive results similar to Theorems 1 and 2 for more general multiplicative functions.
References


Hardy, G.H. and Wright, E.M. (1960), The Theory of Numbers, Oxford University Press.


Neymann, J.E. (1972), On the Probability that k Positive Integers are Relatively Prime, Jour. Number Th. 4, 469-473.


# On The Distribution of the Greatest Common Divisor

**Authors:** Persi Diaconis and Paul Erdős

**Performing Organization:** Department of Statistics, Stanford University, Stanford, CA 94305

**Contract or Grant Number:** DAAG 29-77-G-0031

**Report Date:** 10/10/77

**Number of Pages:** 10

**Distribution Statement:** Approved for Public Release; Distribution Unlimited.

**Keywords:** Probabilistic number theory, greatest common divisor, least common multiple, multinomial distribution.

**Abstract:** Please see reverse side.
The limiting joint distribution of the least common multiple and the greatest common divisor is determined. The lead term is given for the moments of both marginal distributions. The results are applied to the problem of reconstructing the sample size from a percentage classification.