The Annals of Probability
1984, Vol. 12, No. 4, 1205–1212

ORDERED PRIME DIVISORS OF A RANDOM INTEGER

BY STUART P. LLOYD

AT&T Bell Laboratories

Without using the prime number theorem, we obtain the asymptotics of the rth largest prime divisor of a harmonically distributed random positive integer N; harmonic asymptotics are obtained from asymptotics of the zeta distribution via Tauberian methods. (Knuth and Trabb-Pardo need a strong form of the prime number theorem to obtain the distributions when N is uniformly distributed.) A trick brings in Poisson variates, and then we can use the methods developed for the fractional length of the rth longest cycle in a random permutation.

1. Introduction. The asymptotics of the rth largest prime divisor of a random positive integer give rise to a certain limiting distribution F_r, derived by Knuth and Trabb-Pardo in [7]. They observe that their F_r coincides with the limiting distribution of the fractional length of the rth longest cycle in a random permutation, obtained in [9]. They do not explain the coincidence. The present contribution consists of deriving the rth largest prime divisor distribution by a method paralleling that of [9]. No isomorphism of the problems is established; the termwise correspondence is not particularly close. The methods developed here are applicable to other problems involving divisors of a random integer.

2. Ordered prime divisors. We denote by $q_r(n)$ the rth largest prime divisor of a positive integer n. That is, n is the product $n = q_1(n)q_2(n)\cdots$ where $q_r(n)$ is a prime or 1 and is weakly decreasing in r, eventually to 1. It is convenient to have $q_r(n)$ defined for all $r = 1, 2, \cdots$, as we have done. The quantities of direct concern are the $\eta_r(n)$ defined by

$$\begin{aligned} \eta_r(n) &= \frac{\log q_r(n)}{\log n}, \qquad && n > 1, \\ &= \delta_{r1}, && n = 1. \end{aligned} \tag{1}$$

The following distributions of a random positive integer N are familiar. The uniform distribution is some weak limit of the finite approximations

$$P_{u,\nu}\{N = n\} = \begin{cases} 1/\nu, & 1 \le n \le \nu, \\ 0, & n > \nu. \end{cases}$$

The harmonic distribution has finite approximations

$$P_{h,\nu}\{N = n\} = \begin{cases} 1/a_\nu n & 1 \le n \le \nu, \\ 0, & n > \nu, \end{cases}$$

Received July 1982; revised April 1984.

AMS 1980 *subject classifications.* Primary, 60B99; secondary, 10K20.

Key words and phrases. rth largest prime divisor.

where

$$a_\nu = \sum_1^\nu (1/n) = \log \nu + \gamma + O(1/\nu);$$

$\gamma = .577 \cdots$ is Euler's constant.

It is shown in [7] that limiting distributions

$$\lim_{\nu\to\infty} P_{u,\nu}\{\eta_r(N) \le x\} = F_r(x), \quad 0 \le x \le 1/r, \tag{2}$$

exist, and explicit forms for the F_r are given. In [9] the same functions F_r are obtained via their moments. Explicitly, let $E(x)$, $0 < x < \infty$, denote the exponential integral

$$E(x) = \int_x^\infty \frac{e^{-y}}{y}\, dy, \quad 0 < x < \infty,$$

and let $\xi(t)$, $0 < t < \infty$, denote its functional inverse: $E(\xi(t)) = t$, $0 < t < \infty$. Then the moments of F_r are represented as

$$\begin{aligned} G_{r,m} &= \int_0^\infty \frac{[\xi(t)]^m}{m!}\, e^{-t} \frac{t^{r-1}}{(r-1)!}\, dt \\ &= \int_0^{1/r} x^m \, dF_r(x), \quad m \ge 0, \quad r \ge 1. \end{aligned}$$

What we prove here, using the method of moments, is

THEOREM 1. *For the harmomic distribution of N, the asymptotic distribution*

$$\lim_{\nu\to\infty} P_{h,\nu}\{\eta_r(N) \le x\} = F_r(x), \quad 0 \le x \le 1/r,$$

exists, $r = 1, 2, \cdots$.

This is a weaker result than (2), but we obtain it without the prime number theorem, used in [7]. Getting uniform asymptotics from harmonic asymptotics is a delicate matter, requiring control of the error terms or Tauberian conditions; cf. [4]. Indeed, the prime number theorem can be given this setting [8]. (On the other hand, Billingsley does not need the prime number theorem to obtain the asymptotics of the ordered *distinct* prime divisors of uniformly distributed N [3].) The harmonic averages are a poor man's theory of a random positive integer; we let it go at that. Related problems are treated in [1], [5].

3. The zeta distribution. The following family of distributions of N is known by a variety of names; we call it the zeta distribution. Namely, for $s > 1$,

$$P_s\{N = n\} = 1/(\zeta(s)n^s), \quad n = 1, 2, \cdots,$$

where

$$\zeta(s) = \sum_1^\infty 1/n^s, \quad s > 1,$$

is the Riemann zeta function. As $s \downarrow 1$ the unnormalized probabilities $1/n^s$ tend to the unnormalizable harmonic distribution $1/n$. It is well known that Tauberian methods will give harmonic asymptotics in terms of zeta asymptotics at $s = 1$ [4] [6, Volume II, Section XIII.5]. The following is sufficient for our purposes. Assume the Tauberian condition $f(n) \geq 0$, and suppose that $\lim_{s\to 1} E_s\{f(N)\} = A$ with $0 < A < \infty$. Then $\lim_{\nu\to\infty} E_{h,\nu}\{f(N)\} = A$.

Let $p_1 = 2, p_2 = 3, p_3 = 5, \cdots$ be the enumeration of the successive primes. In several places it will simplify the formalism if we define also $p_0 = 1$; this is not regarded as a prime. We will also use the usual convention where p is a variable ranging over the set $\{p_\mu : \mu \geq 1\}$. Let

$$n = \prod_{\mu=1}^{\infty} [p_\mu]^{\alpha_\mu(n)}$$

be the prime factorization of n; for given n, the $\{\alpha_\mu(n)\}$ are defined for all $\mu = 1, 2, \cdots$, and are eventually 0. The Euler factorization

$$\frac{1}{\zeta(s)} = \prod_{\mu=1}^{\infty} \left(1 - \frac{1}{p_\mu^s}\right),$$

together with the prime factorization of N, gives

$$P_s\{\alpha_1(N) = a_1, \alpha_2(N) = a_2, \cdots\} = \prod_{\mu=1}^{\infty} \left[\left(1 - \frac{1}{p_\mu^s}\right)\left(\frac{1}{p_\mu^s}\right)^{a_\mu}\right]$$

for all sequences $(a_1, a_2, \cdots)$ of nonnegative integers eventually 0. But this is to say, the $\{\alpha_\mu(N)\}$ are mutually independent geometrically distributed random variables, the ratio parameter of $\alpha_\mu(N)$ being $1/p_\mu^s$. This independence often makes it possible to evaluate $E_s\{f(N)\}$ explicitly when $f(\cdot)$ is a multiplicative or additive number theoretic function; cf. [2, Chapter 11].

The following device enables us to approximate the geometrically distributed $\{\alpha_\mu(N)\}$ with Poisson variates. If a random nonnegative integer α is geometrically distributed, the generating function of its distribution is

$$E\{x^\alpha\} = \frac{1-\rho}{1-\rho x}, \quad |x| < 1/\rho,$$

where $0 \leq \rho < 1$ is the ratio parameter. The Poisson distribution of mean ρ has generating function $\exp[(x-1)\rho]$, $|x| < \infty$.

LEMMA 1. *The quotient* $[(1-\rho)/(1-\rho x)]/[e^{(x-1)\rho}]$ *is the generating function of a probability distribution:*

$$\frac{(1-\rho)e^{-(x-1)\rho}}{1-\rho x} = \sum_{m=0}^{\infty} \lambda_m(\rho) x^m, \quad |x| < 1/\rho,$$

where

$$\lambda_m(\rho) = (1-\rho)e^{\rho}\rho^m \sigma(m),$$

$$\sigma(m) = \sum_{j=0}^{m} \frac{(-1)^j}{j!}, \quad m = 0, 1, \cdots, \quad 0 \leq \rho < 1. \tag{3}$$

Proof. Explicit calculation. A direct proof of $0 \leq \sigma(m) \leq 1$ is straightforward; alternatively, $1 - \sigma(m)$ is the matching probability of [6, Vol. I, page 101]. □

Lemma 1 is equivalent to the following: the above geometrically distributed α is equidistributed with $\alpha' + \alpha''$ where α' is Poisson with the same parameter as α, α'' has the distribution (3), and α' and α'' are independent. As will appear, the property that makes this useful is $P\{\alpha'' > 0\} = P\{\alpha'' \geq 2\} = O(\rho^2)$, $\rho \to 0$, stemming from $\sigma(1) = 0$.

Let us apply this to each $\alpha_\mu(N)$ of the zeta distribution. We extend the probability space, if necessary, to support independent random positive integers N' and N'' with the following distributions:

$$(4)\qquad \begin{aligned} P_s\{N' = n\} &= \frac{1}{\nu(s)n^s} \prod_{\mu=1}^{\infty} \frac{1}{[\alpha_\mu(n)]!}, \quad n = 1, 2, \cdots, \\ P_s\{N'' = n\} &= \frac{\nu(s)}{\zeta(s)n^s} \prod_{\mu=1}^{\infty} \sigma(\alpha_\mu(n)), \quad n = 1, 2, \cdots, \end{aligned}$$

where

$$\nu(s) = \exp\left[\sum_{\mu=1}^{\infty} \frac{1}{p_\mu^s}\right], \quad s > 1.$$

Then $N = N'N''$ has the zeta distribution, since $\alpha_\mu(N) = \alpha_\mu(N') + \alpha_\mu(N'')$ and $\{\alpha_\mu(N')\}$ are independent Poisson with the appropriate parameters and the independent $\{\alpha_\mu(N'')\}$ have the distribution (3).

The point is that N'' does not diverge as $s \downarrow 1$. The divergence of $N = N'N''$ is carried by the N' factor, and the Poisson $\{\alpha_\mu(N')\}$ are more tractable for our purposes than the geometrical $\{\alpha_\mu(N)\}$. More precisely, the N'' distribution has a limit at $s = 1$ which is a probability distribution. (In fact, this persists for $s > 1/2$, and $E_s\{(N'')^\varepsilon\}$ exists for $s - \varepsilon > 1/2$.)

Lemma 2. *The abscissa of convergence of the Dirichlet series*

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \prod_{\mu=1}^{\infty} \sigma(\alpha_\mu(n)) \left(= \frac{\zeta(s)}{\nu(s)}, \quad \mathscr{R}e(s) > 1,\right)$$

is $s = 1/2$.

Proof. Möbius inversion [2, page 40] of

$$\begin{aligned} \log \zeta(s) &= \sum_p \sum_{m=1}^{\infty} (1/mp^{ms}) \\ &= \sum_{m=1}^{\infty} (1/m) \log \nu(ms), \quad \mathscr{R}e(s) > 1, \end{aligned}$$

gives

$$\log \nu(s) = \sum_{m=1}^{\infty} (\mu(m)/m) \log \zeta(\mathrm{ms}), \quad \mathscr{R}e(s) > 1,$$

absolutely convergent in $\mathscr{R}e(s) > 1$. From this we obtain

$$\frac{\zeta(s)}{v(s)} = \prod_p \left[\frac{\exp(-1/p^s)}{1 - 1/p^s}\right]$$

$$= \prod_{m=2}^{\infty} [\zeta(ms)]^{-\mu(m)/m},$$

and these are convergent in $\mathscr{R}e(s) > ½$, with divergence at $s = ½$. The proof is completed by appeal to the results of [2, Section 11.9]. □

Later we will use

$$v(s) \sim c/(s - 1), \quad s \downarrow 1,$$

where $c \approx .729$ is the constant

$$c = \prod_p \left[e^{1/p}\left(1 - \frac{1}{p}\right)\right] = P_1\{N'' = 1\}.$$

This follows from $\zeta(s) \sim 1/(s - 1)$ and $v(s)/\zeta(s) \to c$.

4. The Poisson process model for N'. With $s > 1$ fixed, we lay off the points $t_1(s) > t_2(s) > \cdots \downarrow 0$ on the positive t-axis, where

$$t_\mu(s) = \sum_{\nu=\mu}^{\infty} (1/p_\nu^s), \quad \mu = 1, 2, \cdots,$$

noting that $t_1(s) = \log v(s) < \infty$. We define the μth interval $I_\mu(s)$ to be $I_\mu(s) = (t_{\mu+1}(s), t_\mu(s)]$, $\mu = 1, 2, \cdots$; the length of $I_\mu(s)$ is $1/p_\mu^s$. We define also $I_0(s) = (t_1(s), \infty)$.

Let a Poisson process take place on the t-axis, with jump rate unity. The number of jumps of the process occurring in $I_\mu(s)$ has the Poisson distribution of mean $1/p_\mu^s$, and the numbers in the various $I_\mu(s)$ are mutually independent. But this is the joint distribution of the $\{\alpha_\mu(N')\}$ for the N' distribution of (4). More completely, define $\Lambda_s(t)$, $0 < t < \infty$, by

$$\Lambda_s(t) = \log p_\mu, \quad t \in I_\mu(s), \quad \mu = 0, 1, \cdots.$$

Let $0 < \tau_1 \leq \tau_2 \leq \cdots$ be the random locations of the successive jumps of the process on $0 < t < \infty$, counting to the right from the origin. The total number of jumps in $(0, t_1(s)]$ is finite w.p.1, so the terms in the series $\sum_{r=1}^{\infty} \Lambda_s(\tau_r)$ are eventually 0 w.p.1. The series thus determines a random positive integer N' according to $\sum_{r=1}^{\infty} \Lambda_s(\tau_r) = \log N'$ w.p.1. The distribution of this N' is the N' distribution of (4).

The usefulness of the method now becomes apparent: the rth term in the series is $\Lambda_s(\tau_r) = \log q_r(N')$, $r = 1, 2, \cdots$. The density of τ_r is $e^{-t}t^{r-1}/(r - 1)!$, $0 < t < \infty$, as is well known. The distribution of $q_r(N')$ is thus

$$P_s\{q_r(N') = p_\mu\} = \int_{I_\mu(s)} e^{-t}\left[\frac{t^{r-1}}{(r - 1)!}\right] dt, \quad \mu = 0, 1, \cdots.$$

The moments of $\log q_r(N')$ are then

$$E_s\{[\log q_r(N')]^m\} = \int_0^\infty [\Lambda_s(t)]^m e^{-t}\left[\frac{t^{r-1}}{(r-1)!}\right] dt, \quad m, r \geq 1.$$

Existence of the moments is assured by comparison with

$$E_s\{(\log N)^m\} = \frac{1}{\zeta(s)} \sum_{n=1}^\infty \frac{(\log n)^m}{n^s}, \quad s > 1.$$

5. Proof of Theorem 1. Let us define $\xi_s(t) = (s-1)\Lambda_s(t)$, $0 < t < \infty$. We will prove $\lim_{s\to 1}\xi_s(t) = \xi(t)$, $0 < t < \infty$, where $\xi(t)$ is the function specified in Section 2. The asymptotics of $\xi(t)$ are [9]

$$\begin{aligned} \xi(t) &\sim e^{-t-\gamma}, & t &\to \infty, \\ &\sim \log(1/t), & t &\to 0. \end{aligned}$$

Let $Y(dx)$ be the discrete measure which assigns measure $1/p_\mu$ to the point $x = \log p_\mu$ for all $\mu = 1, 2, \cdots$; the cumulative is

$$\begin{aligned} Y(x) &= \textstyle\sum_p \{1/p : p \leq e^x\}, & -\infty < x < \infty, \\ &= 0, & -\infty < x < \log 2. \end{aligned}$$

In terms of this,

$$t_\mu(s) = \int_{\log p_\mu -}^\infty e^{-(s-1)x} Y(dx), \quad \mu = 1, 2, \cdots.$$

The asymptotics of $Y(x)$ are as follows [2, Section 4.8]. For a certain constant A,

$$\begin{aligned} Y(x) &= \log x + A + Y_1(x), \quad \log 2 \leq x < \infty, \\ Y_1(x) &= O(1/x), \quad x \to \infty. \end{aligned}$$

(The prime number theorem is equivalent to convergence of $\int^\infty xY_1(dx)$, but we do not need this; cf. [8].) We find

$$\begin{aligned} t_\mu(s) &= \int_{\log p_\mu -}^\infty e^{-(s-1)x}\left[\frac{dx}{x} + Y_1(dx)\right] \\ &= E((s-1)\log p_\mu) - \exp(-(s-1)\log p_\mu)Y_1(\log p_\mu -) \\ &\quad + (s-1)\int_{\log p_\mu}^\infty e^{-(s-1)x} Y_1(x)\, dx. \end{aligned}$$

Since $(s-1)\log p_\mu = \xi_s(t)$, $t \in I_\mu(s)$, this can be rewritten as

$$\begin{aligned} t_\mu(s) &= E(\xi_s(t)) + e^{-\xi_s(t)}O(1/\log p_\mu) \\ &\qquad + (s-1)O(E(\xi_s(t))), \quad I_\mu(s) \ni t. \end{aligned} \tag{5}$$

Now fix $0 < t < \infty$ and let $s \downarrow 1$. The μ for which $I_\mu(s) \ni t$ increases to ∞ as $s \downarrow 1$, and $t_\mu(s) \longrightarrow t$. (The jumps in $\xi_s(t)$ tend to zero:

$$\begin{aligned}(s-1)\log p_{\mu+1} - (s-1)\log p_\mu & \\ &= (s-1)O(1) \quad \text{by [2, Theorem 4.7]} \\ (\text{or } &= (s-1)o(1) \quad \text{by the prime number theorem}).)\end{aligned}$$

From (5) then follows $\lim_{s\to1}\xi_s(t) = \xi(t), \quad 0 < t < \infty$.

We want to show that the moments of $\xi_s(t)$ converge to the moments of $\xi(t)$ relative to the density of τ_r. Since pointwise convergence has just been shown, and since $0 \leq \log q_r(N') \leq \log N$, it will suffice to show that the family $\{[(s-1)\log N]^m : 1 < s \leq s_0\}$ is uniformly integrable with respect to $E_s\{\cdot\}$, $1 < s \leq s_0$. (The corresponding step in [9] is easier.) This is a direct consequence of the following result.

LEMMA 3. *Let $0 < \theta_0 < \infty$ and $1 < s_0 < \infty$ be fixed. Then*

$$P_s\{(s-1)\log N \geq \theta\} \leq Ae^{-\theta}, \quad 1 < s \leq s_0, \quad \theta_0 \leq \theta < \infty,$$

where $1 < A < \infty$ is independent of θ, s.

PROOF. Summation of

$$\frac{1}{n^s} \leq \int_{n-1}^{n} \frac{dx}{x^s}, \quad n > 1,\ s \geq 0,$$

gives

$$\begin{aligned}\textstyle\sum_n \left\{\frac{1}{n^s} : (s-1)\log n \geq \theta\right\} &\leq \frac{1}{s-1}\,[e^{\theta/(s-1)} - 1]^{-(s-1)} \\ &= \frac{e^{-\theta}B(\theta, s)}{s-1}, \quad \theta > 0,\ s > 1,\end{aligned}$$

where

$$B(\theta, s) = [1 - e^{-\theta/(s-1)}]^{-(s-1)}.$$

Since $B(\theta, s)$ is decreasing in θ and increasing in s, we can take $A = B(\theta_0, s_0)$. The inequality $(s-1)\zeta(s) > 1$, $s > 1$, is elementary, and the lemma follows. □

With uniform integrability thus established, we now have

$$\begin{aligned}\lim_{s\downarrow1} & E_s\{[(s-1)\log p_r(N')]^m\} \\ &= \lim_{s\downarrow1} \int_0^\infty [\xi_s(t)]^m e^{-t}\left[\frac{t^{r-1}}{(r-1)!}\right] dt \\ &= \int_0^\infty [\xi(t)]^m e^{-t}\left[\frac{t^{r-1}}{(r-1)!}\right] dt \\ &= m!G_{r,m}, \quad m \geq 0, \quad r \geq 1.\end{aligned}$$

Let us rewrite this in the equivalent form

$$\left(-\frac{d}{ds}\right)^m \sum_{n=2}^{\infty} \frac{[\eta_r(n)]^m}{n^s \prod_\mu [\alpha_\mu(n)]!} \sim \left(-\frac{d}{ds}\right)^m \left\{\frac{cG_{r,m}}{s-1}\right\}, \quad s \downarrow 1,$$

with $\eta_r(n)$ defined in (1). We integrate m times on (s, ∞), which is permissible, and divide by $v(s)$; we obtain the weaker result

$$\lim_{s\downarrow 1} E_s\{[\eta_r(N')]^m\} = G_{r,m}, \quad m \geq 0, \quad r \geq 1.$$

We now argue that this implies

$$\lim_{s\downarrow 1} E_s\{[\eta_r(N)]^m\} = G_{r,m}, \quad m \geq 0, \quad r \geq 1. \tag{6}$$

This follows from the fact that $\eta_r(N') \rightarrow \eta_r(N)$ in distribution: $q_r(N') = q_r(N)$ except on a set of small measure, and $\log N/\log N' \rightarrow 1$ in distribution. We omit the epsilontics. From (6) and the Tauberian theorem,

$$\lim_{\nu\rightarrow\infty} E_{h,\nu}\{[\eta_r(N)]^m\} = G_{r,m}, \quad m \geq 0, \quad r \geq 1,$$

and application of the method of moments [6, Volume II, page 514] completes the proof. □

Properties of the rth shortest cycle in a random permutation were investigated in [9] with the same machinery as for the rth longest cycle. The techniques of the present paper would not be very useful for treating the rth smallest prime divisor of random $N = N'N''$, since the small divisors of N' and N'' remain intertwined even when $N' \rightarrow \infty$.

REFERENCES

[1] ALLADI, K. and ERDŐS, P. (1977). On an additive arithmetic function. *Pacific J. Math.* **71** 275–294.

[2] APOSTOL, T. M. (1976). *Introduction to Analytic Number Theory.* Springer-Verlag, New York.

[3] BILLINGSLEY, P. (1972). On the distribution of large prime divisors. *Period. Math. Hungar.* **2** 283–289.

[4] DIACONIS, P. (1977). Examples for the theory of infinite iteration of summability methods. *Canad. J. Math.* **29** 489–497.

[5] DIACONIS, P. (1980). Average running time of the fast Fourier transform. J. Algorithms **1** 187–208.

[6] FELLER, W. (1970, 1971). An Introduction to Probability Theory and its Applications, I 3rd ed., II 2nd ed. Wiley, New York.

[7] KNUTH, D. E. and TRABB-PARDO, L. (1976/77). Analysis of a simple factorization algorithm. *Theor. Comput. Sci.* **3** 321–348.

[8] NEWMAN, D. J. (1980). Simple analytic proof of the prime number theorem. *Amer. Math. Monthly,* **87** 693–696.

[9] SHEPP, L. A. and LLOYD, S. P. (1966). Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.* **121** 340–357.

AT & T BELL LABORATORIES
WHIPPANY, NEW JERSEY 07981