ON ZIPF’S LAW

MICHAEL WOODROOFE AND
BRUCE HILL, University of Michigan

Abstract
A Zipf’s law is a probability distribution on the positive integers which
decays algebraically. Such laws describe (approximately) a large class of
phenomena. We formulate a model for such phenomena and, in terms of
our model, give necessary and sufficient conditions for a Zipf’s law to hold.

CLASSICAL OCCUPANCY PROBLEM; LAWS OF LARGE NUMBERS; WEAK CONVERGENCE;
REGULARLY VARYING FUNCTIONS

1. Introduction

By a Zipf’s law with exponent $a$, $a > 0$, we mean a probability distribution
$p_1, p_2, \cdots$ on the positive integers for which $p_s \sim c/s^{1+a}$ as $s \to \infty$. Such laws
describe (approximately) a wide variety of phenomena. In particular, the following
application was observed by Willis (1922). Consider $N$ biological species which
are classified into $M$ (non-empty) genera, and for $s \geq 1$, let $G(s)$ be the number
of genera which contain exactly $s$ species; then, in many cases, the relative frequen-
cies $G(s)/M$ are nearly proportional to $1/s^{1+a}$. Data which support this statement
are also considered by Hill (1970) and by Hill and Woodrooife (1975). Zipf (1949)
gave extensive empirical evidence that Zipf’s laws describe (approximately)
many other phenomena, including such diverse ones as the frequency of word
usage, the distribution of city sizes and the distribution of incomes. Theoretical
models which predict Zipf’s law have been considered by Yule (1924), Hill (1970)
and (1974), Hill and Woodrooife (1975), and Simon (1955).

Hill (1970) introduced the following model which, in the context of biological
species and genera, may be described as follows. Suppose that $N$ species are
allocated to $M$ (non-empty) genera according to the Bose-Einstein form of the
classical occupancy problem. Then $G(s)$ is determined as a random variable.
Suppose also that $M$ is a random variable whose distribution depends on $N$ and
that the distribution function $F_N$ of $\theta = M/N$ converges weakly to a distribution
function $F$ for which $F(0) = 0$ as $N \to \infty$. It was then shown that

$$
\lim E\{M^{-1}G(s)\} = \int_0^1 t(1-t)^{s-1}dF(t)
$$

Received in revised form 14 October 1974.
as $N \to \infty$ for every $s \geq 1$. It was also observed that if $F$ is the Beta distribution with parameters $\alpha > 0$ and $\beta > 0$, then the right side of (1) is

$$
 p_s = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha + 1)} \frac{\Gamma(\beta + s - 1)}{\Gamma(\alpha + \beta + s)} \sim \frac{x^\alpha \Gamma(\alpha + 1)}{\Gamma(\beta)} \left(\frac{1}{s}\right)^{1+s}.
$$

It was then shown by Hill and Woodrooffe (1975) that a more refined model will yield the convergence of the actual proportions $G(s)/M$ to a limit which may differ from the right side of (1) and that the limit may be a Zipf’s law for appropriate choices of the parameters of the model.

In this paper we also consider the more refined model, and we shall give necessary and sufficient conditions for the convergence of $G(s)/M$ to a limit $p_s$ for every $s \geq 1$. We shall then investigate the class of possible limits $p_s$, $s \geq 1$. Under some restrictions on the distribution of the number of genera, we show the following: if there are probabilities $p_s$, $s \geq 1$, for which $G(s)/M$ converges to $p_s$ for every $s \geq 1$, and if $M/N$ is small, then

$$
(2) \quad p_s = \frac{\alpha^2 \Gamma(\alpha) \Gamma(s)}{\Gamma(\alpha + s + 1)} \sim \alpha^2 \Gamma(\alpha) \left(\frac{1}{s}\right)^{1+s}.
$$

That is, the only possible limits for $G(s)/M$, $s \geq 1$, are Zipf’s laws.

In Section 2 we describe our model precisely and give some background material. In Section 3 we give necessary and sufficient conditions for the convergence of $G(s)/M$ to a limit for every $s \geq 1$; and in Section 4 we study the class of possible limits and derive Equation (2).

2. The model

For ease of exposition we shall describe our model in terms of biological species and genera. Our model involves a two-fold classification of species into families of species and, within families, into genera. We suppose that there are $k$ families of species with $N_{ki}$ species and $M_{ki}$ genera in the $i$th family, and we require $1 \leq M_{ki} \leq N_{ki}$, for $i = 1, \ldots, k$. We regard $k$ and $N_{k1}, \ldots, N_{kk}$ as parameters which will eventually tend to infinity, and we suppose that $M_{k1}, \ldots, M_{kk}$ are independent random variables. We further suppose that there is a fixed sequence of distribution functions $H_1, H_2, \ldots$ for which $H_n(n) - H_n(1) = 1$ for $n \geq 1$ and the distribution function of $M_{ki}$ is $H_{N_{ki}}$ for $i = 1, \ldots, k$ and $k \geq 1$. That is we suppose that the distribution of the number of genera in a particular family depends only on the number of species in that family.

Next, we suppose that given the vector $M_k = (M_{k1}, \ldots, M_{kk})$, the $N_{ki}$ species in the $i$th family are allocated to the $M_{ki}$ genera in the $i$th family according to the
Bose-Einstein form of the classical occupancy problem for \( i = 1, \cdots, k \) and that the allocations of species into genera within different families are independent. That is, if \( L_{ij} \) denotes the number of species in the \( j \)th genus of the \( i \)th family, then we require that all \( \binom{N_{ki} - 1}{M_{ki} - 1} \) possible values of \( L_i = (L_{i1}, \cdots, L_{iM_{ki}}) \) are equally likely and that the vectors \( L_{i1}, \cdots, L_i \) are independent, given the vector \( M_k \).

Henceforth, \( P_k \) will denote a probability measure which makes the assumptions of the previous two paragraphs valid, and \( E_k \) will denote expectation with respect to \( P_k \). The dependence of \( P_k \) and \( E_k \) on the vector \( N_k = (N_{k1}, \cdots, N_{kk}) \) will be suppressed in the notation.

Let \( G_k(s) \) denote the number of genera in the \( i \)th family which contain exactly \( s \) species, \( i = 1, \cdots, k \) and \( s \geq 1 \), and let \( v_s(m, n) \) and \( \sigma_s^2(m, n) \) denote the conditional mean and variance of \( G_k(s) \) given \( M_{ki} = m \) when \( N_{ki} = n \). Then, as is shown by Hill (1970),

\[
\begin{align*}
v_s(m, n) &= m \binom{n-s-1}{m-2} \binom{n-1}{m-1}, \\
\sigma_s^2(m, n) &= m(m-1) \binom{n-2s-1}{m-3} \binom{n-1}{m-1} - v_s(m, n)^2.
\end{align*}
\]

**Lemma 1.** Let \( \theta = m/n \). Then there are constants \( c = c_s \), depending only on \( s \), for which

\[
|v_s(m, n) - m\theta(1 - \theta)^{s-1}| \leq c\theta \quad \text{and} \quad \sigma_s^2(m, n) \leq c^2m\theta
\]

for \( 1 \leq m \leq n \) and \( s \geq 1 \).

The lemma may be established by a straightforward, if tedious, induction argument. Let

\[
\mu(n) = \int_0^n xH_s(x)
\]

be the expectation of \( M_{ki} \) when \( N_{ki} = n \) and let

\[
\mu_k = \mu(N_{k1}) + \cdots + \mu(N_{kk})
\]

be the expectation of \( M_k = M_{k1} + \cdots + M_{kk} \). We shall impose some conditions on the behavior of \( \mu(n) \) and \( \mu_k \).

1. \( A_1: \mu(n) \uparrow \infty \) as \( n \uparrow \infty \).
2. \( A_2: \) as \( k \to \infty, N_{k1}, \cdots, N_{kk} \) vary in such a manner that \( k = o(\mu_k) \).

That is, we require the average number of genera per family to be large. We shall also need a condition which will allow us to apply the law of large numbers.

1. \( A_3: \) as \( k \to \infty, N_{k1}, \cdots, N_{kk} \) vary in such a manner that
\[ \sum_{i=1}^{k} \int_{(M_{k_i} \geq \varepsilon \mu_k)} M_{k_i} dP_k = o(\mu_k) \]

for every \( \varepsilon > 0 \).

It is not difficult to see that \( A_3 \) will be satisfied if \( A_1 \) and \( A_2 \) are satisfied and \( M_{k_i} / \mu(n) \) is uniformly integrable with respect to \( H_n, n \geq 1 \). This, however, is not the case of primary concern to us. We shall now discuss some alternative conditions which imply \( A_3 \) when \( A_2 \) is strengthened to

\( A_2' \): Let \( N_{k_1}' \leq N_{k_2}' \leq \cdots \leq N_{k_k}' \) denote the ordered values of \( N_{k_1}, \ldots, N_{k_k} \) and for \( 0 < a < 1 \), let \( j_a \) denote the greatest integer which is less than or equal to \( ka \). Then \( N_{k_k}' \to \infty \) as \( k \to \infty \) and for some values of \( a \) and \( A, 0 < a < 1, A > 0 \), we have \( N_{k_k}' \leq AN_{j_a} \) for all \( k \geq 1 \).

It is easily verified that \( A_1 \) and \( A_2' \) imply \( A_2 \) (see (3) below).

**Lemma 2.** Let Conditions \( A_1 \) and \( A_2' \) be satisfied and let \( N_k = N_{k_1} + \cdots + N_{k_k} \). Then \( A_3 \) will be satisfied if either

(i) \( N_k = O(k^2) \) as \( k \to \infty \); or

(ii) \( \mu(n) = n^\alpha L(n) \) where \( \alpha > 0 \) and \( L \) varies slowly at \( \infty \), and \( N_{k_k}'^{-\alpha} = o\left(k^{-\alpha} L(N_{k_k}'^{-1})\right) \) as \( k \to \infty \).

**Proof.** It will suffice to show that \( N_{k_k}' = \max\{N_{k_1}, \ldots, N_{k_k}\} = o(\mu_k) \) under either set of hypotheses. We have

\[ \mu_k \geq k(1-a)\mu(N_{k_k}') \]

(3)

and

\[ N_{k_k}' \leq AN_{k_k}/(1-a). \]

Since \( \mu(N_{k_k}') \to \infty \) as \( k \to \infty \), it is clear that \( N_{k_k}' = o(\mu_k) \) in case (i). In case (ii) we have

\[ \mu_k \geq k(1-a)\mu(N_k/Ak) \sim (1-a)A^{-\alpha} k^{1-\alpha} N_k L(k^{-1} N_k), \]

from which it follows that \( N_{k_k}' = o(\mu_k) \) in case (ii) too.

3. **Convergence**

We shall need the following notation. Let \( M_k = M_{k_1} + \cdots + M_{k_k} \),

\[ G_k(s) = \sum_{i=1}^{k} G_{k_i}(s), \]

and

\[ \hat{G}_k(s) = \sum_{i=1}^{k} M_{k_i} \theta_{k_i}(1 - \theta_{k_i})^{s-1}, \]
where $\theta_{si} = M_{si}/N_{si}$, $i = 1, \cdots, k$. Thus, $M_k$ and $G_k(s)$ are the total number of genera and the total number of genera which contain exactly $s$ species, respectively. In this section we shall derive a necessary and sufficient condition for the convergence in probability of $M_k^{-1} G_k(s)$ to a limit $p_s$ for every $s \geq 1$, when Conditions $A_1$, $A_2$, and $A_3$ are satisfied.

**Lemma 3.** $E_k\{ |G_k(s) - \hat{G}_k(s)|^2\} \leq c_k^2(h^2 + \mu_k)$ for all $k \geq 1$ and $N_k$ where $c_k$ is as in Lemma 1.

**Proof.** Since $G_{k1}(s), \cdots, G_{kk}(s)$ are conditionally independent given $M_k = (M_{k1}, \cdots, M_{kk})$, $E_k\{ |G_k(s) - \hat{G}_k(s)|^2\}$ is simply

$$E_k\{ \operatorname{Var}[G_k(s) | M_k] \} + E_k\{ E[G_k(s) - \hat{G}_k(s) | M_k]^2 \}$$

$$\leq E_k\{ \sum_{i=1}^k \sigma_i^2(M_{ki}, N_{ki}) \} + c_k^2 h^2 \leq c_k^2(h^2 + \mu_k),$$

as asserted.

Let

$$v(s, n) = \int x \left( \frac{x}{n} \right) \left( 1 - \left( \frac{x}{n} \right)^{s-1} \right) dH(x)$$

$$v_k(s) = v(s, N_{k1}) + \cdots + v(s, N_{kk})$$

for $s \geq 1$ and $k \geq 1$, and observe that $v_k(s) = E_k\{ \hat{G}_k(s) \}$.

**Lemma 4.** Let Conditions $A_1$, $A_2$, and $A_3$ be satisfied. Then $\mu_k^{-1} M_k \to 1$ and $\mu_k^{-1}[G_k(s) - v_k(s)] \to 0$

in probability as $k \to \infty$ for every $s \geq 1$.

**Proof.** We shall first show that $\lim_k \mu_k^{-1} E[M_k - \mu_k] = 0$. Indeed, if $\varepsilon > 0$ is given and $M_{ki}^* = M_{ki}$ if $M_{ki} \leq \varepsilon^2 \mu_k$ and $M_{ki}^* = 0$ if $M_{ki} > \varepsilon^2 \mu_k$, then

$$Q_k = \sum_{i=1}^k E_k[M_{ki}^* - M_{ki}] = o(\mu_k)$$

by $A_3$. Moreover, since $M_{k1}, \cdots, M_{kk}$ are independent,

$$E((M_k^* - \mu_k)^2) \leq \sum_{i=1}^k E[(M_{ki}^*)^2] + Q_k^2 \leq \varepsilon^2 \mu_k^2 + Q_k^2,$$

where $M_k^* = M_{k1}^* + \cdots + M_{kk}^*$. The first assertion of the lemma now follows easily and the second may be established by a similar argument.

**Lemma 5.** Let Assumptions $A_1$ and $A_2$ be satisfied and let $\alpha_k = \mu(N_k)/\mu_k$ for $i = 1, \cdots, k$. If $\beta(n), n \geq 1$, is a sequence of real numbers for which $\lim \beta(n) = \beta$ as $n \to \infty$, then
\[
\lim \sum_{i=1}^{k} \alpha_i \beta(N_{\alpha_i}) = \beta
\]
as \(k \to \infty\).

**Proof.** We may suppose that \(\beta = 0\). Let \(\epsilon > 0\) be given and let \(n_0\) be so large that \(|\beta(n)| \leq \epsilon\) for \(n \leq n_0\). Then

\[
\left| \sum_{i=1}^{k} \alpha_i \beta(N_{\alpha_i}) \right| = \left\{ \sum_{N_{\alpha_i} < n_0} + \sum_{N_{\alpha_i} \geq n_0} \right\} \alpha_i \beta(N_{\alpha_i}) \leq bk\mu(n_0)\mu_k^{-1} + \epsilon
\]

where \(b = \max \{ |\beta(n)| : n \leq n_0 \}\). The lemma follows.

Let

\[
F_a(t) = n \mu(n)^{-1} \left[ 1 - H_a(nt) \right]
\]

for \(0 < t \leq 1\) and \(n \geq 1\). Observe that each \(F_a\) is a right continuous, non-increasing function on \((0, 1]\) and that \(F_a(t) \leq 1/t\) for \(0 < t \leq 1\) and \(n \geq 1\). Thus, the sequence \(F_n, n \geq 1\), is relatively compact with respect to the topology of weak convergence (Feller (1966), Chapter 8). Moreover, it is easily seen that if \(F_n\) converges weakly to a limit \(F\) as \(n \to \infty\), and if \(g\) is any continuous function on \([0, 1]\), then

\[
\lim \int_{0}^{1} g(t)t^2 F_a(dt) = \int_{0}^{1} g(t)t^2 F(dt),
\]

where \(F[\cdot]\) denotes the measure induced on the Borel sets of \([0, 1]\) by \(F\), i.e.,

\[
F([a, b]) = F(a) - F(b) \quad \text{for} \quad 0 \leq a \leq b \leq 1.
\]

**Theorem 1.** Let Assumption \(A_1\) be satisfied and let \(k \to \infty\) in such a manner that \(A_2\) and \(A_3\) are satisfied. Then there are non-negative numbers \(p_s, s \geq 1\), for which \(\lim M_{-1}^s G_s(s) = p_s\) in probability as \(k \to \infty\) for all \(s \geq 1\) and all sequences \(N_k\) which satisfy \(A_2\) and \(A_3\) if, and only if, \(F_n\) converges weakly to a right continuous, non-increasing limit \(F\) as \(n \to \infty\). In this case

\[
p_s = \int_{0}^{1} t^2 (1 - t)^{s-1} F(dt)
\]

for \(s \geq 1\), and \(p_s, s \geq 1\), are probabilities if, and only if, \(\int_{0}^{1} t F(dt) = 1\).

**Proof.** It follows easily from Lemmas 3 and 4 that \(M_{-1}^s G_s(s)\) will converge in probability to a limit \(p_s\) as \(k \to \infty\) in the manner indicated if, and only if, \(\mu_{-1}^s \nu(s)\) converges to \(p_s\). Suppose first that \(F_n\) converges weakly to a limit \(F\) as \(n \to \infty\) and let \(p_s(n) = \nu(s, n) / \mu(n), n \geq 1\). Then
\[ p_s(n) = \int_0^1 t^n (1 - t)^{s-1} F_s(dt) \]

for \( n \geq 1 \), so that \( p_s(n) \) converges to \( p_\alpha \), defined by (4), as \( n \to \infty \) for every \( s \geq 1 \). Therefore, by Lemma 5,

\[ \mu_k^{-1} v_k(s) = \sum_{i=1}^k a_k \pi_i(n_k) \]

converges to \( p_\alpha \) as \( k \to \infty \) (in the manner indicated).

Conversely, suppose that the sequence \( F_n, n \geq 1 \), had two distinct limit points, \( F \) and \( F' \) say, and let us consider the case that \( N_{k+1} = \cdots = N_{k+l} = \lfloor k \rfloor \), the greatest integer which is less than or equal to \( \sqrt{k} \). Then as \( k \to \infty \), \( A_2 \) and \( A_3 \) are satisfied by Lemma 2. By letting \( k \to \infty \) along subsequences for which \( F_n \) approaches \( F \) and \( F' \), we find that both \( p_\alpha \) and

\[ p'_s = \int_0^1 t^n (1 - t)^{s-1} F'(dt) \]

are limit points of \( \mu_k^{-1} v_k(s), k \geq 1 \). In order for \( \mu_k^{-1} v_k(s) \) to converge for every \( s \geq 1 \), we must therefore have \( p_\alpha = p'_s \) for every \( s \geq 1 \); but this implies that \( F = F' \) and consequently that \( F_n \) converges weakly to \( F \) as \( n \to \infty \).

We shall now consider some examples.

**Example 1.** Suppose that \( A_3 \) is satisfied and let \( K_n(t) = H_n(nt), 0 < t \leq 1, n \geq 1 \). If \( K_n \) converges weakly to a distribution function \( K \) for which \( K(0) = 1 \), then \( \mu(n) \sim \alpha n \) as \( n \to \infty \), where

\[ \mu = \int_0^1 t K(dt). \]

It follows easily that \( M_{\alpha}/\mu(n) \) is uniformly integrable with respect to \( H_n, n \geq 1 \), so that \( A_3 \) implies \( A_3 \). It now follows easily from Theorem 1 that if \( k \to \infty \) in such a manner that \( A_2 \) is satisfied, then \( M_k^{-1} G_s(s) \) converges in probability to

\[ p_s = \mu^{-1} \int_0^1 t^n (1 - t)^{s-1} K(dt) \]

for every \( s \geq 1 \).

Since any distribution function \( K \) on \([0, 1]\) may be obtained as the limit of \( K_n \) by an appropriate choice of \( H_n, n \geq 1 \), the class of possible limits described in Example 1 is quite large. This class includes Zipf’s laws with exponents \( \alpha > 1 \), but it does not include any with exponent \( \alpha, 0 < \alpha \leq 1 \). In fact, the expectation of the distribution defined by (5) is \( \mu^{-1} \), which is finite. In order to obtain Zipf’s laws with exponent \( \alpha, 0 < \alpha \leq 1 \), we must therefore consider the case that \( K_n \) becomes degenerate at \( 0 \) as \( n \to \infty \).
Example 2. Suppose that the distribution functions $H_n, n \geq 1,$ are of the form

$$H_n(j) - H_n(j -) = c_n j^{1+\beta}$$

for $j = 1, \cdots, n,$ where $0 < \beta < 1$ and $c_n$ is a normalizing constant. Then $\mu(n) \sim cn^{\alpha}/\alpha$ as $n \to \infty,$ where $\alpha = 1 - \beta$ and $c = \lim c_n,$ so that $A_3$ will be satisfied if $A_1$ and $A_2$ are satisfied and $N_k = O(k^\gamma)$ as $k \to \infty$ with $\gamma = (1 + \beta)/\beta.$ Moreover, it follows easily from (6) that $F_n$ converges weakly to $F,$ where

$$F(t) = \alpha \beta^{-1} \Gamma(\Gamma^{-\beta} - 1)$$

for $0 < t \leq 1.$ Thus, if $A_1$ is satisfied and if $k \to \infty$ in such a manner $A_2$ holds and $N_k = o(k^\gamma),$ then $M_k^{-1} G_k(s)$ converges in probability to

$$p_s = \frac{\alpha^2 \Gamma(\alpha) \Gamma(s)}{\Gamma(\alpha + s + 1)}$$

for every $s \geq 1.$ Of course, the right side of (7) is a Zipf’s law with parameter $\alpha.$

In the next section we show that Example 2 is more general than might appear.

4. Possible limits

In this section we consider the case that the distributions $H_n, n \geq 1,$ are of the form

$$H_n(x) = \frac{H(x)}{H(n)}, \quad x \leq n,$$

where $H$ is a fixed, non-decreasing function on $[0, \infty)$ for which $H(1-) = 0.$ We shall suppose throughout this section that $\mu(n) \to \infty$ with $\mu(n) = o(n)$ as $n \to \infty,$ and we shall describe the class of possible limits of

$$F_n(t) = n \mu(n)^{-1} \left[ 1 - \frac{H(nt)}{H(n)} \right], \quad 0 < t \leq 1, \quad n \geq 1.$$

We shall use the theory of regularly varying functions as described, for example, by Feller (1966), pp. 268–272.

Theorem 2. Let $H$ be a fixed non-decreasing function on $[0, \infty)$ for which $H(1-) = 0$ and let $F_n, n \geq 1,$ be defined by (9). Suppose also that $\mu(n) \to \infty$ with $\mu(n) = o(n)$ as $n \to \infty.$ Then $F_n$ converges weakly to a non-degenerate, right continuous, non-increasing limit $F$ as $n \to \infty$ if, and only if,

$$J(x) = \int_0^x yH(dy)$$

varies regularly with exponent $\alpha, 0 < \alpha \leq 1,$ as $x \to \infty.$ In this case
(10) \[ F(1) = 0 \text{ and } F'(t) = -at^{\alpha-2} \quad 0 < t \leq 1. \]

**Proof.** Let \( K_n(t) = J(nt)/J(n) \), \( 0 < t \leq 1 \), \( n \geq 1 \). Then

\[
K_n(t) = 1 - \int_{t}^{1} uF_n(du)
\]

for \( 0 < t \leq 1 \) and \( n \geq 1 \). Suppose first that \( J \) varies regularly with exponent \( \alpha \), \( 0 < \alpha \leq 1 \) and let \( F \) be any weak limit point of the sequence \( F_n \), \( n \geq 1 \). Then \( K_n(t) \to t^\alpha \) as \( n \to \infty \), so that

\[
1 - \int_{t}^{1} uF(du) = t^\alpha
\]

for \( 0 < t < 1 \). Here we have used the right continuity of \( F \). Since (10) and (11) are equivalent, and since \( F \) was an arbitrary weak limit point of the sequence \( F_n \), \( n \geq 1 \), the 'if' assertion in Theorem 2 follows.

Conversely, suppose that \( F_n \) converges weakly to a non-degenerate, right continuous, non-increasing limit \( F \). Then \( K_n(t) \) converges to

\[
K(t) = 1 - \int_{t}^{1} uF(du)
\]

whenever \( t \) is a continuity point of \( F \). Moreover, \( K \) is positive on some interval since \( F \) is non-degenerate. Since \( J \) is monotone, this also implies that \( J(n+1)/J(n) \to 1 \) as \( n \to \infty \). Therefore, by Lemma 2 of Feller (1966), p. 270, \( J \) varies regularly with some exponent \( \alpha, -\infty < \alpha < \infty \).

It remains to show that \( 0 < \alpha \leq 1 \). That \( \alpha \geq 0 \) follows from the fact that \( J \) is non-decreasing, and that \( \alpha \neq 0 \) follows from the fact that \( F \) is non-degenerate.

Finally, we have

\[
H(2n) - H(n) \leq n^{-1} \mu(2n)H(2n) = o\{H(2n)\}
\]

by Markov’s inequality and the assumption that \( \mu(n) = o(n) \). It now follows easily that \( H(n) = o(n^\epsilon) \) for all \( \epsilon > 0 \). Therefore, \( J(n) = H(n)\mu(n) = o(n^{\epsilon+\epsilon}) \) for all \( \epsilon > 0 \), so that \( \alpha \leq 1 \).

If we now combine Theorems 1 and 2, we obtain

**Theorem 3.** Let \( H \) be given by (8), where \( H \) is a fixed non-decreasing function on \([0, \infty)\), and suppose that \( \mu(n) \to \infty \) with \( \mu(n) = o(n) \) as \( n \to \infty \). Finally, let \( k \to \infty \) in such a manner that \( A_2 \) and \( A_3 \) are satisfied. Then there are probabilities \( p_s, s \geq 1 \), for which

\[
limit M_k^{-1} G_s(s) = p_s
\]
in probability for all $s \geq 1$ if, and only if, $J$ varies regularly with exponent $\alpha$, $0 < \alpha \leq 1$, in which case

$$p_s = \frac{s^2 \Gamma(s) \Gamma(\alpha)}{\Gamma(\alpha + s + 1)}$$

for $s \geq 1$.

References


Willis, J. C. (1922) *Age and Area*. Cambridge University Press.
