THE ASYMPTOTICS OF RANDOM SIEVES

G. R. GRIMMETT AND R. R. HALL

Abstract. Let $\mathcal{S}=(s_1, s_2, \dots)$ be a collection of relatively prime integers, and suppose that $\pi(n)=|\mathcal{S}\cap\{1,2,\dots,n\}|$ is a regularly varying function with index α satisfying $0<\alpha<1$. We investigate the "stationary random sieve" generated by $\mathcal{S}$, proving that the number of integers less than k which escape the action of the sieve has a probability mass function with approximate order $k^{-\alpha/2}$ in the limit as $k\to\infty$. This result may be used to deduce certain asymptotic properties of the set of integers which are divisible by no $s\in\mathcal{S}$, in that it gives new information about the usual deterministic (that is, non-random) sieve. This work extends previous results valid when $s_i=p_i^2$, the square of the ith prime.

§1. *Introduction.* Let $\mathcal{S}=(s_1, s_2, \dots)$ be a sequence of relatively prime positive integers satisfying $1<s_1<s_2<\dots$. It is of great interest to understand the statistics of the set U of "$\mathcal{S}$-free" integers, being those which are divisible by no member of $\mathcal{S}$. Of particular interest to number theorists is the case when $s_i=p_i^a$, the ath power of the ith prime, for which U is the set of "a-free numbers".

It was shown in Grimmett (1991) that some asymptotic properties of U may be investigated *via* a certain weak convergence result relating U to the "random sieve generated by $\mathcal{S}$", a random process defined as follows. Let $X_1, X_2, \dots$ be independent random variables, X_i having probability mass function

$$P(X_i=r)=\begin{cases}1/s_i, & \text{if } 1\leqslant r\leqslant s_i,\\ 0, & \text{otherwise,}\end{cases}$$

so that X_i is equally likely to take any value in $\{1,2,\dots,s_i\}$. (Throughout this paper, P stands for probability and E for expectation.) Let $\mathcal{G}=\{g_1, g_2, \dots\}$ be a set of distinct labels, and attach labels to integers in the following way. For each $m\geqslant 1$, we label with g_m each member of the set $\{X_m+ks_m: 0\leqslant k<\infty\}$. The outcome of this process is a random vector $\Gamma=(\Gamma_1, \Gamma_2, \dots)$ of subsets of $\mathcal{G}$, Γ_i being the (random) set of labels of i. It is easy to see that the vectors $(\Gamma_1, \Gamma_2, \dots)$ and $(\Gamma_2, \Gamma_3, \dots)$ have the same distributions, so that Γ is a stationary random process. Of especial interest in this general setup is the set of integers with empty label sets. Writing I_A for the indicator function of the event A, consider the stationary random sequence $I=(I_{\{\Gamma_i=\varnothing\}}: i\geqslant 1)$ of 0's and 1's. It is a simple matter to calculate the mean of the partial sum

$$\Sigma_k=\sum_{i=1}^{k} I_{\{\Gamma_i=\varnothing\}}=|\{i\in\{1,2,\dots,k\}: \Gamma_i=\varnothing\}|$$

[MATHEMATIKA, **38** (1991), 285-302]

thus:

$$E(\Sigma_k) = kP(\Gamma_1 = \emptyset) = k \prod_{s \in \mathscr{S}} \left(1 - \frac{1}{s}\right),$$

which is strictly positive, if, and only if,

$$\sum_{s \in \mathscr{S}} \frac{1}{s} < \infty;$$

we shall assume always that this last sum converges. Of more interest are the higher moments of Σ_k, and particularly its variance. Hall (1982) proved that

$$\operatorname{var}(\Sigma_k) = \sigma^2 \sqrt{k} + O(k^{\varepsilon + (1/3)}) \qquad \text{as} \qquad k \longrightarrow \infty,$$

for some constant σ^2 and all $\varepsilon > 0$, in the special case when $s_i = p_i^2$, the square of the ith prime (see Hall (1989) for bounds on the higher moments of Σ_k). A subsidiary result of this paper will be that $\operatorname{var}(\Sigma_k)$ is (in some suitable sense) approximately k^α under the rather weak assumption that the "distribution function" $\pi(n) = |\mathscr{S} \cap \{1, 2, \ldots, n\}|$ of $\mathscr{S}$ is regularly varying with index α satisfying $0 < \alpha < 1$ (that is, $\pi(cn)/\pi(n) \to c^\alpha$ as $n \to \infty$, for all $c > 0$). It is most striking that $\operatorname{var}(\Sigma_k)$ does not have order k, an observation which reflects the fact that the stationary sequence I has substantial correlations over unbounded distances. That is to say,

$$P(\Gamma_1 = \emptyset, \Gamma_{1+t} = \emptyset) = \prod_{s \in \mathscr{S}} \left(1 - \frac{1}{s}\right) \prod_{\substack{s \in \mathscr{S} \\ s \nmid t}} \left(\frac{1 - (2/s)}{1 - (1/s)}\right)$$

$$\not\longrightarrow P(\Gamma_1 = \emptyset)^2 \qquad \text{as} \qquad t \longrightarrow \infty.$$

It is presumably the case that, possibly under mild extra conditions, $(\Sigma_k - E(\Sigma_k))/\sqrt{\operatorname{var} \Sigma_k}$ converges in distribution as $k \to \infty$, and indeed that the random process $Y_n(t) = \Sigma_{\lfloor kt \rfloor}$ converges weakly as $k \to \infty$, when suitably normalized, to a limit process which is self-similar with index $\frac{1}{2}\alpha$. We are unable to make any substantial progress towards this conjecture. Instead, we extend in this paper the main conclusion of Grimmett (1991). It was shown there that, for the case when $s_i = p_i^2$, the mass function of Σ_k has approximately order $k^{-1/4}$ as $k \to \infty$, in the sense that

$$M_k = \sup_j \{P(\Sigma_k = j)\}$$

satisfies

$$A \leqslant k^{1/4} M_k \leqslant B \sqrt{\log k}, \tag{1.1}$$

for all k and some positive constants A and B. For general $\mathscr{S}$ we have the following result.

THEOREM 1. *Suppose that $\pi(n) = |\mathscr{S} \cap \{1, 2, \ldots, n\}|$ is regularly varying with index α satisfying $0 < \alpha < 1$, and let $\varepsilon > 0$. Then $M_k = \sup_j \{P(\Sigma_k = j)\}$ satisfies*

$$k^{-\varepsilon} \leqslant k^{\alpha/2} M_k \leqslant k^\varepsilon, \tag{1.2}$$

for all large k.

We recall that the function π is said to be *regularly varying* if

$$l(c)=\lim_{n\to\infty}\frac{\pi(cn)}{\pi(n)}$$

exists for all $c>0$; if π is regularly varying then it is easily seen that $l(c)=c^{\alpha}$ for some α, called the *index* of the function. See Feller (1971, p. 275) and Bingham, Goldie and Teugels (1987).

The conclusion (1.2) of Theorem 1 is rather weaker than the special case (1.1) of the squares of the primes; this is of course due to the weakness of the assumption on $\mathscr{S}$ in the general case. Rather more is provable for particular choices of $\mathscr{S}$. In the proof of this result, the full assumption of regular variation is not used; slightly less suffices, but the gain in stating a marginally more general conclusion subject to a very messy hypothesis seems insignificant.

In proving Theorem 1, we shall obtain slightly more than stated there.

THEOREM 2. *Suppose that $\pi(n)=|\mathscr{S}\cap\{1,2,\dots,n\}|$ is regularly varying with index α satisfying $0<\alpha<1$. Then $M_k=\sup_j\{P(\Sigma_k=j)\}$ satisfies*

$$\frac{A}{\sigma_k}\leqslant M_k\leqslant\frac{Bg_\alpha(k)}{\sqrt{\sum_{s>k,s\in\mathscr{S}}k/s}}\qquad\text{for all}\qquad k,\tag{1.3}$$

where A and B are positive constants, $\sigma_k^2=\operatorname{var}(\Sigma_k)$, and $g_\alpha(k)$ is given by

$$g_\alpha(k)=\begin{cases}1, & \text{if } 0<\alpha<\frac{1}{2}\\ \min\left\{1+\frac{1}{k}\left(\sum_{\substack{s>k\\ s\in\mathscr{S}}}\frac{k}{s}\right)^2,\ \sqrt{\log k}\,\{\log\log k\}^{1+\varepsilon}\right\}, & \text{if } \alpha=\frac{1}{2}\\ \sqrt{\log k}, & \text{if } \frac{1}{2}<\alpha<1\end{cases}\tag{1.4}$$

where ε is an arbitrary positive constant.

The principal component of these results is the upper bound for M_k, being an upper bound of order $(\sum_{s>k}k/s)^{-1/2}$; such sums are understood to be over values of s in $\mathscr{S}$. The logarithmic terms in the upper bound are of little importance and may be improved with extra work, especially in the case $\alpha=\frac{1}{2}$. When $\mathscr{S}$ is the set of ath powers of the primes, then $\alpha=1/a$ and

$$\sum_{\substack{s>k\\ s\in\mathscr{S}}}\frac{k}{s}\sim\frac{ak^{1/a}}{(a-1)\log k},$$

so that Theorem 2 implies that

$$\frac{A}{\sigma_k}\leqslant M_k\leqslant B\sqrt{\frac{\log k}{k^{1/a}}},$$

which leads to (1.1) in the case $a=2$.

The above theorems may be applied, *via* the weak convergence theorem of Grimmett (1991), to the usual (non-random) sieve generated by $\mathscr{S}$. For each $m\geqslant 1$, we label with g_m every multiple of s_m. This results in a deterministic sequence $G=(G_1,G_2,\dots)$ of subsets of the label set $\mathscr{G}$, G_i being the set of labels of i. Of principal interest is the set $U(G)=\{m\colon G_m=\varnothing\}$ of "$\mathcal{S}$-free" integers. Let

$$S_k(m)=|U(G)\cap\{m, m+1,\ldots, m+k-1\}|.$$

If $1/s_i$ is summable then

$$p_k(j)=\lim_{n\to\infty}\left\{\frac{1}{n}|\{m\in\{1,2,\ldots,n\}\colon S_k(m)=j\}|\right\}$$

exists and is given explicitly by $p_k(j)=P(\Sigma_k=j)$. Information about the mass function of Σ_k may therefore be translated into information about the number of $\mathcal{S}$-free integers in a typical interval. If π is regularly varying with index α satisfying $\alpha\in(0,1)$, then certainly $1/s_i$ is summable, and it follows from Theorem 1 that, for all $\varepsilon>0$,

$$k^{-\varepsilon}\leqslant k^{\alpha/2}\sup_j\{p_k(j)\}\leqslant k^{\varepsilon}\qquad\text{for all large } k. \tag{1.5}$$

In proving the above theorems, there is a particular complication when $\frac{1}{2}\leqslant\alpha<1$. This difficulty arises from the fact that, if one throws $b=B^{\alpha}$ balls independently at random into B boxes, then the number Y of boxes containing two or more balls has order approximately $b^2/(2B)=\frac{1}{2}B^{2\alpha-1}$, and this tends to infinity as $B\to\infty$, if, and only if, $\alpha>\frac{1}{2}$. We resolve this difficulty by proving a limit theorem for Y as $B\to\infty$, together with an estimate of the rate of convergence therein. Such results may be found in Section 2. Readers interested only in number theory may pass directly to Section 3, which contains the proofs of Theorems 1 and 2.

§2. *The occupancy problem.* We shall use the result of this section in the later proofs of Theorems 1 and 2.

The occupancy problem is one of the most elementary of probability theory. We are provided with B boxes and b balls, and we allocate the balls one by one to the boxes in such a way that each ball is equally likely to go into any of the boxes, and different balls are allocated independently of each other. Let X be the number of non-empty boxes. The asymptotic distribution of X as $b, B\to\infty$, subject to $b=f(B)$ for some given f, has been much studied (see for example Kolchin, Sevastyanov, and Chistyakov (1978) or Chow and Teicher (1978)).

Instead of working with X, we shall work with $Y=b-X$, the number of balls which are allocated to already occupied boxes. It is not difficult to see that $\lambda=E(Y)$ satisfies $\lambda=(1+o(1))b^2/(2B)$ and also $\lambda^{-1}\operatorname{var}(Y)\to 1$ as b, $B\to\infty$ in such a way that $b/B\to 0$. This suggests a Poisson limit theorem which, with an estimate of the rate of convergence, is the subject of the next theorem.

THEOREM 3. *There exist positive constants C_1 and C_2 such that, whenever $1\leqslant 2B\log B\leqslant b^2$ and $b\leqslant C_1B$, then*

$$\left|P(Y=j)-\frac{1}{j!}\lambda^j e^{-\lambda}\right|\leqslant\frac{C_2}{\sqrt{B}\{1+(jB/b^2)\}}, \tag{2.1}$$

for $j=0,1,2,\ldots$.

There is a corresponding result for the total variation distance between Y and the Poisson distribution,

$$d_{TV}(Y,\Pi)=\sup_{A}\{|P(Y\in A)-P(\Pi\in A)|\}, \tag{2.2}$$

where Π is a random variable having the Poisson distribution with mean λ, and the supremum is taken over all subsets A of $\mathbb{Z}=\{0,1,2,\dots\}$. Indeed it is straightforward to deduce from the arguments of Barbour and Holst (1989) that

$$d_{TV}(Y,\Pi)\leqslant 1-\frac{\operatorname{var}(Y)}{E(Y)}, \tag{2.3}$$

and a calculation of an elementary nature then implies that there exists a constant C such that

$$d_{TV}(Y,\Pi)\leqslant C\frac{b}{B}. \tag{2.4}$$

One must work a little harder for the "local" estimate of (2.1), and we shall make use of Theorem II.Q of Barbour, Holst, and Janson (1991) in order to establish this.

The main purpose of Theorem 3 is to establish an error estimate. The condition that $b^2\geqslant 2B\log B$ may be weakened by a closer look at the proofs. More careful asymptotic analysis than that presented here would yield estimates for the constants in (2.1) and (2.4).

Proof of Theorem 3. Suppose that $1\leqslant 2B\log B\leqslant b^2$ and $6\leqslant b\leqslant \frac{1}{2}B$. First we estimate the mean and variance of Y. Clearly

$$E(Y)=b-E(X)=b-B\left\{1-\left(1-\frac{1}{B}\right)^{b}\right\}, \tag{2.5}$$

since each box is empty with probability $(1-B^{-1})^b$. A similar elementary calculation (or see Kolchin *et al.* (1978, p. 5)) shows that

$$\begin{aligned}\operatorname{var}(Y)&=\operatorname{var}(X)\\&=B(B-1)\left(1-\frac{2}{B}\right)^{b}+B\left(1-\frac{1}{B}\right)^{b}-B^2\left(1-\frac{1}{B}\right)^{2b}.\end{aligned} \tag{2.6}$$

It is clear from (2.5) that λ $(=E(Y))$ is given by

$$\lambda=\sum_{i=1}^{b-1}\binom{b}{i+1}\frac{(-1)^{i+1}}{B^i},$$

and hence

$$\left|\lambda-\frac{b^2}{2B}\right|\leqslant\frac{b}{2B}+\frac{b^3}{B^2}\sum_{i=3}^{b}\frac{1}{i!}\leqslant\frac{b^2}{2B}\left(\frac{1}{b}+\frac{b}{B}\right). \tag{2.7}$$

As for the variance, a careful expansion of $\lambda-\operatorname{var}(Y)$ yields

$$\lambda-\operatorname{var}(Y)=B^2\left\{\sum_{i=4}^{2b}\binom{2b}{i}\frac{(-1)^i}{B^i}-\sum_{i=4}^{b}\binom{b}{i}\frac{(-1)^i}{B^i}\right\}+B\sum_{i=3}^{b}\binom{b}{i}\frac{(-2)^i}{B^i}.$$

With some care, one may deduce the existence of a positive constant γ_1 such that

$$\left|\lambda - \operatorname{var}(Y) - \frac{2b^3}{3B^2}\right| \leq \gamma_1 \frac{b^4}{B^3}. \tag{2.8}$$

We denote the balls by $\beta(1), \beta(2), \ldots, \beta(b)$ and the boxes by $\tau(1), \tau(2), \ldots, \tau(B)$. We may assume without loss of generality that $\beta(1)$ is allocated to $\tau(1)$; otherwise we relabel the boxes. The remaining balls are allocated to boxes one by one in a random order. Let $\pi = (\pi_2, \pi_3, \ldots, \pi_b)$ be a random permutation of $(2, 3, \ldots, b)$, and let $\beta'_1 = \beta(1)$, $\beta'_i = \beta(\pi_i)$ for $2 \leq i \leq b$; we place the balls into boxes in the order $\beta'_1, \beta'_2, \ldots, \beta'_b$. Let $Z_2, Z_3, \ldots, Z_b$ be independent random variables each being uniformly distributed on $\{1, 2, \ldots, B\}$; for $i \geq 2$, we place β'_i in the box $\tau(Z_i)$. For each $i \geq 2$, we call $\beta(\pi_i) = \beta'_i$ an *intruder* if it is placed into a box which is already occupied, that is, if $\beta'_j \in \tau(Z_i)$ for some $j < i$. Writing I_i for the indicator function of the event that $\beta(i)$ is an intruder, we have that

$$Y = \sum_{i=2}^{b} I_i,$$

and hence by symmetry

$$E(I_i) = \frac{\lambda}{b-1} \quad \text{for} \quad 2 \leq i \leq b. \tag{2.9}$$

Inherent in the Stein-Chen approach to Poisson convergence is the idea of a "good coupling" of Y to a copy of Y conditioned on the event $\{I_i = 1\}$. To this end, we introduce new random variables $Z'_2, Z'_3, \ldots, Z'_b$ as follows. If $\beta'_i = \beta(\pi_i)$ is an intruder, we define $Z'_i = Z_i$. If on the other hand β'_i is not an intruder, then we let Z'_i be chosen uniformly at random from the set of indices of those boxes which are occupied just prior to the allocation of β'_i.

Fix k such that $2 \leq k \leq b$, and consider the following revised scheme of allocation. Let $l = \pi^{-1}(k)$ be the unique l such that $\pi_l = k$; thus $\beta'_l = \beta(k)$. We allocate β'_i to $\tau(Z_i)$ when $i \neq l$ and to $\tau(Z'_l)$ when $i = l$. Let J_{ik} be the indicator function of the event that, in this new scheme, $\beta(i)$ is allocated to a box which is already occupied, and define

$$W_k = \sum_{i: i \neq k} J_{ik}.$$

A little thought leads to the following observations.

(a) $W_k + 1$ has the same distribution as Y conditioned on the event that $\beta(k)$ is an intruder.

(b) $J_{ik} \leq I_i$ for $i \neq k$.

(c) $J_{ik} = 0$ and $I_i = 1$ where $i \neq k$, if and only if $\beta(k)$ is not an intruder and furthermore $\beta(i)$ is the next ball to be allocated to the box containing $\beta(k)$.

(d) By observation (c), $Y - 1 \leq W_k \leq Y$.

Using (a) and (b), we have from Theorems 2.1 and 2.3 of Barbour and Holst (1989) that

$$d_{TV}(Y, \Pi) \leq \frac{1}{\lambda} \sum_{k=2}^{b} E(I_k) E|Y - W_k| = 1 - \frac{1}{\lambda} \operatorname{var}(Y) \leq \gamma_2 \frac{b}{B}, \tag{2.10}$$

for some $\gamma_2 > 0$, by (2.7) and (2.8), where Π has the Poisson distribution with mean λ. This confirms (2.4) under the assumptions that $1 \leq 2B \log B \leq b^2$ and $6 \leq b \leq \frac{1}{2}B$; it may easily be seen that these assumptions are not essential for (2.4).

Before moving on, we note from (2.9), (2.10), and the fact that the summands in (2.10) are constant for $2 \leq k \leq b$ (the sequence of pairs (I_i, W_i) is of course exchangeable), that, for $2 \leq k \leq b$,

$$E|Y - W_k| = 1 - \frac{1}{\lambda}\operatorname{var}(Y). \tag{2.11}$$

We shall make use of Theorem II.Q of Barbour, Holst, and Janson (1991) in order to obtain the pointwise estimate of the theorem. In order to verify the hypotheses of Theorem II.Q, it suffices to prove that there exist positive constants C_1 ($\leq \frac{1}{2}$) and γ_3 such that, for $2 \leq k \leq b$ and $b/B \leq C_1$, the following four equations are valid:

$$P(|Y - W_k| > 1) = 0, \tag{2.12}$$

$$P(Y \geq 6\lambda) \leq \frac{2}{\sqrt{\lambda}} E|Y - W_k|, \tag{2.13}$$

$$\frac{P(W_k = m-1 \mid Y = m)}{E|Y - W_k|} \leq \gamma_3 \quad \text{for} \quad 1 \leq m \leq 6\lambda, \tag{2.14}$$

$$\gamma_3\left(1 - \frac{1}{\lambda}\operatorname{var}(Y)\right) \leq \frac{1}{4}. \tag{2.15}$$

Equation (2.12) is an immediate consequence of observation (d) above. For (2.13), we use the fact that $P(Y \geq y) \leq P(T \geq y)$ where T has the binomial distribution with parameters b and b/B. This is intuitively clear, since the probability that β_i' is an intruder is no larger than b/B, for all i and whatever the boxes to which $\beta_1', \beta_2', \ldots, \beta_{i-1}'$ have been allocated; in a more rigorous argument, one may show that Y may be coupled to a copy of T in such a way that $Y \leq T$. Now $E(T) = b^2/B$, and $\lambda \leq b^2/B \leq 3\lambda$ by (2.7) so long as $b/B \leq \frac{1}{6}$. Hence, if $b/B \leq \frac{1}{6}$, then

$$P(Y \geq 6\lambda) \leq P(T \geq 2E(T)) \leq e^{-2tb^2/B}\left(1 - \frac{b}{B} + \frac{b}{B}e^t\right)^b \quad \text{for} \quad t > 0,$$

by Markov's inequality (see for example Grimmett and Stirzaker (1982, p. 181)). We set $e^t = 2$ and find that

$$P(Y \geq 6\lambda) \leq (\tfrac{1}{4}e)^{b^2/B}.$$

Therefore, by (2.10) and (2.8),

$$\sqrt{\lambda}\,\frac{P(Y \geq 6\lambda)}{E|Y - W_k|} = \frac{\lambda^{3/2}}{\lambda - \operatorname{var}(Y)}\,P(Y \geq 6\lambda) \leq 2\sqrt{B}\,(\tfrac{1}{4}e)^{b^2/B},$$

if b/B is sufficiently small, say $b/B \leq \gamma_4$ where $0 < \gamma_4 \leq \frac{1}{4}$. Now $b^2 \geq 2B \log B$, and (2.13) follows so long as $b/B \leq \gamma_4$.

For (2.14) we argue as follows. From observations (c) and (d) above, it is the case that $W_k = Y-1$, if, and only if, either (i) $\beta(k)$ is an intruder or (ii) $\beta(k)$ is not an intruder but some later ball is allocated to the same box as $\beta(k)$. Hence the number of values of k for which $W_k = Y-1$ is no greater than twice the number of intruders, and therefore

$$\sum_{k=2}^{b} P(W_k = m-1 \mid Y=m) \leqslant 2E(Y \mid Y=m) = 2m.$$

By symmetry, if $m \leqslant 6\lambda$,

$$P(W_k = m-1 \mid Y=m) \leqslant \frac{2m}{b-1} \leqslant \frac{12\lambda}{b-1},$$

whence

$$\frac{P(W_k = m-1 \mid Y=m)}{E|Y-W_k|} \leqslant \frac{12\lambda^2}{(b-1)\{\lambda - \mathrm{var}(Y)\}}.$$

Using (2.7) and (2.8) again, we find that the last term is smaller than some constant γ_3 so long as b/B is sufficiently small, say $b/B \leqslant \gamma_5$ where $\gamma_5 > 0$.

Having picked γ_3 such that (2.14) holds, we have then that

$$\gamma_3\left(1 - \frac{1}{\lambda}\mathrm{var}(Y)\right) \leqslant \gamma_3\gamma_2\frac{b}{B}$$

by (2.10), and this is smaller than $\frac{1}{4}$, if $b/B \leqslant (4\gamma_3\gamma_2)^{-1}$. We set $C_1 = \min\{\gamma_4, \gamma_5, (4\gamma_3\gamma_2)^{-1}\}$, and this is the first constant in the statement of the theorem.

Suppose that $b/B \leqslant C_1$. The hypotheses of Theorem II.Q of Barbour *et al.* (1991) having been verified (see (2.12)-(2.15)), we deduce from that theorem that there exists a constant γ_6 such that

$$|P(Y=j) - P(\Pi = j)| \leqslant \gamma_6 \frac{\lambda - \mathrm{var}(Y)}{\sqrt{\lambda}(\lambda + j)}, \qquad \text{for} \qquad j \in \mathbb{Z},$$

and the result follows by (2.7) and (2.10).

§3. *Proofs of Theorems* 1 *and* 2. The general plan of the proofs resembles closely that of Grimmett (1991). We suppose that $\mathscr{S} = (s_1, s_2, \dots)$ is a strictly increasing sequence of coprime integers such that $\pi(n) = |\mathscr{S} \cap \{1, 2, \dots, n\}|$ is regularly varying with index $\alpha \in (0, 1)$.

It is an elementary consequence of the assumption of regular variation (see Feller (1971, p. 277)) that

$$\forall \varepsilon > 0, \exists N \quad \text{such that} \quad n^{\alpha-\varepsilon} \leqslant \pi(n) \leqslant n^{\alpha+\varepsilon} \quad \text{for all} \quad n \geqslant N. \tag{3.1}$$

Here is a preliminary lemma. Products and summations over the variable s are understood to be over all $s \in \mathscr{S}$.

LEMMA 1. *Let $\varepsilon > 0$. Then*

$$\sum_s \frac{1}{s^{\alpha-\varepsilon}} = \infty, \qquad \sum_s \frac{1}{s^{\alpha+\varepsilon}} < \infty. \tag{3.2}$$

If $\beta > \alpha$ then

$$n^{\alpha-\beta-\varepsilon} < \sum_{s>n} \frac{1}{s^\beta} < n^{\alpha-\beta+\varepsilon} \quad \text{for all large} \quad n. \tag{3.3}$$

Proof. We have

$$\sum_s \frac{1}{s^\gamma} = \gamma \int_1^\infty \frac{\pi(x)}{x^{\gamma+1}}\, dx,$$

and we substitute $\gamma = \alpha - \varepsilon$, $\gamma = \alpha + \varepsilon$ in turn, in each case applying (3.1) with ε replaced by $\frac{1}{2}\varepsilon$. This proves (3.2). Next

$$\sum_{s>n} \frac{1}{s^\beta} = \beta \int_n^\infty \frac{\pi(x) - \pi(n)}{x^{\beta+1}}\, dx, \tag{3.4}$$

and we insert the upper bound from (3.1), again with ε replaced by $\frac{1}{2}\varepsilon$. For sufficiently large n, the right-hand side of (3.4) does not exceed

$$\beta \int_n^\infty x^{\alpha-\beta-1+\frac{1}{2}\varepsilon}\, dx = \frac{\beta}{\beta - \alpha - \frac{1}{2}\varepsilon}\, n^{\alpha-\beta+\frac{1}{2}\varepsilon},$$

provided $\varepsilon < 2(\beta - \alpha)$. For large n, the right-hand side is smaller than $n^{\alpha-\beta+\varepsilon}$. We also have, from (3.4),

$$\sum_{s>n} \frac{1}{s^\beta} \geq \beta \int_{2n}^\infty \frac{\pi(x) - \pi(\frac{1}{2}x)}{x^{\beta+1}}\, dx \geq \beta(1 - 2^{-\alpha})(1 - \varepsilon) \int_{2n}^\infty \frac{\pi(x)}{x^{\beta+1}}\, dx, \tag{3.5}$$

for large n, since by hypothesis π is regularly varying with index α. We apply (3.1); if n is sufficiently large, this yields

$$\sum_{s>n} \frac{1}{s^\beta} \geq \beta(1 - 2^{-\alpha})(1 - \varepsilon) \int_{2n}^\infty x^{\alpha-\beta-1-\frac{1}{2}\varepsilon}\, dx > n^{\alpha-\beta-\varepsilon}.$$

This proves (3.3).

The lower bound for M_k will follow by Chebyshev's inequality

$$(1 + 4\sigma_k) M_k \geq P(|\Sigma_k - E(\Sigma_k)| < 2\sigma_k) \geq \tfrac{3}{4},$$

where $\sigma_k^2 = \operatorname{var}(\Sigma_k)$. Hence

$$M_k \geq \frac{3/4}{1 + 4\sigma_k},$$

which gives the lower bound of Theorem 2. The corresponding bound of Theorem 1 is a consequence of the following inequality for σ_k.

LEMMA 2. *For* $\varepsilon > 0$, $\sigma_k^2 \leq k^{\alpha+\varepsilon}$ *for all large* k.

Before proving this, we note that it is also the case that $\sigma_k^2 \geq k^{\alpha-\varepsilon}$ for all $\varepsilon > 0$ and all large k. There are at least two ways of proving this, the shorter being to note that it is an immediate consequence of (1.2) and (1.3). A more primitive argument is available, using the forthcoming technique developed to establish the upper bound for M_k; we sketch this at the end of this section.

Proof. We follow Hall (1982) initially. Suppose for convenience that $s_1 > 2$; we shall consider the case $s_1 = 2$ later. As in Hall (1982),

$$\sigma_k^2 = kZ_1 - (kZ_1)^2 + 2Z_2 \sum_{d=1}^{\infty} g(d) \left\{ \frac{k}{d} \left\lfloor \frac{k}{d} \right\rfloor - \frac{1}{2} \left\lfloor \frac{k}{d} \right\rfloor \left(\left\lfloor \frac{k}{d} \right\rfloor + 1 \right) \right\}, \tag{3.6}$$

where

$$Z_i = \prod_s \left(1 - \frac{i}{s}\right) \tag{3.7}$$

and

$$g(d) = \left\{ \begin{array}{lll} 1, & \text{if } d = 1, & \\ \displaystyle\prod_{j=1}^{J} \frac{s_{i_j}}{s_{i_j} - 2}, & \text{if } d = s_{i_1} s_{i_2} \ldots s_{i_J} \quad \text{where} & i_1 < i_2 < \ldots < i_J, \\ 0, & \text{otherwise;} & \end{array} \right\} \tag{3.8}$$

as usual, $\lfloor x \rfloor$ denotes the integer part of x. Now

$$\sum_{d=1}^{\infty} \frac{g(d)}{d^{\beta}} = \prod_s \left(1 + \frac{1}{s^{\beta-1}(s-2)}\right) < \infty \qquad \text{if} \qquad \beta > \alpha \tag{3.9}$$

by (3.2), and

$$\sum_{d=1}^{\infty} \frac{g(d)}{d} = \frac{Z_1}{Z_2}, \qquad \sum_{d=1}^{\infty} \frac{g(d)}{d^2} = \frac{Z_1^2}{Z_2}.$$

Hence

$$\begin{aligned} \sigma_k^2 &= Z_2 \sum_{d=1}^{\infty} g(d) \left\{ \frac{k}{d} - \frac{k^2}{d^2} + 2\frac{k}{d} \left\lfloor \frac{k}{d} \right\rfloor - \left\lfloor \frac{k}{d} \right\rfloor \left(\left\lfloor \frac{k}{d} \right\rfloor + 1 \right) \right\} \\ &= Z_2 \sum_{d=1}^{\infty} g(d) \{k/d\}(1 - \{k/d\}), \end{aligned}$$

where $\{k/d\} = k/d - \lfloor k/d \rfloor$. For $\beta \in (\alpha, 1]$,

$$\sigma_k^2 \leq \tfrac{1}{4} Z_2 k^{\beta} \sum_{d=1}^{k} \frac{g(d)}{d^{\beta}} + Z_2 \sum_{d>k} g(d) \left(\frac{k}{d}\right)^{\beta} \leq Z_2 k^{\beta} \sum_{d=1}^{\infty} \frac{g(d)}{d^{\beta}},$$

and the claim of the lemma follows by (3.9).

Finally suppose that $s_1 = 2$, so that $X_1 \in \{1, 2\}$. Numbers of the form $X_1 + 2r$, for $r \geq 0$, cannot contribute to Σ_k, which may therefore be expressed as

$$\Sigma_k = \begin{cases} \Sigma'(F), & \text{with probability } \tfrac{1}{2}, \\ \Sigma'(\{1, 2, \ldots, k\} \backslash F), & \text{with probability } \tfrac{1}{2}, \end{cases} \tag{3.10}$$

where $F = \{2, 4, 6, \ldots, 2\lfloor \frac{1}{2}k \rfloor\}$ and $\Sigma'(G)$ is the number of members of a set G which escape the action of $s_2, s_3, \ldots$. With a little thought and computation, it may be seen that $\Sigma'(F)$ and $\Sigma'(\{1, 2, \ldots, k\} \backslash F)$ have the same distributions as $\Sigma'_{|F|}$ and $\Sigma'_{k-|F|}$ respectively, where Σ'_r denotes $\Sigma'(\{1, 2, \ldots, r\})$. It follows from (3.10) that

$$\begin{aligned} \sigma_k^2 &= \tfrac{1}{2}(\operatorname{var}(\Sigma'_{|F|}) + \operatorname{var}(\Sigma'_{k-|F|})) + \tfrac{1}{4}(E(\Sigma'_{|F|}) - E(\Sigma'_{k-|F|}))^2 \\ &= \begin{cases} \sigma'^2_{k/2}, & \text{if } k \text{ is even}, \\ \tfrac{1}{2}(\sigma'^2_{(k+1)/2} + \sigma'^2_{(k-1)/2}) + Z_1^2, & \text{if } k \text{ is odd}, \end{cases} \end{aligned}$$

where $\sigma'^2_r = \operatorname{var}(\Sigma'_r)$. The result follows from the previous calculation.

During the calculations which follow, it is useful to remember that, as a consequence of (3.3), $\sum_{s>k} k/s$ is approximately of order k^α.

Turning to the upper bound for M_k, we shall prove the bounds of Theorem 2; the corresponding bound of Theorem 1 will follow by an application of Lemma 1. We divide $\mathcal{S}$ into two parts: $\mathcal{S} = \mathcal{M} \cup \mathcal{N}$ where

$$\mathcal{M} = \{s \in \mathcal{S} : s \leq k\}, \qquad \mathcal{N} = \mathcal{S} \backslash \mathcal{M}.$$

Let T be the set of integers $i \in \{1, 2, \ldots, k\}$ such that $g_m \notin \Gamma_i$ for all $s_m \in \mathcal{M}$. Certainly $|T| \geq \Sigma_k$. Let ζ be such that $0 < \zeta < Z_1$; then

$$P(|T| \leq \zeta k) \leq P(\Sigma_k \leq \zeta k) \leq \frac{\operatorname{var}(\Sigma_k)}{(Z_1 - \zeta)^2 k^2}, \tag{3.11}$$

by Chebyshev's inequality. Hence

$$P(|T| \leq \zeta k) \leq \frac{1 + o(1)}{k} \quad \text{as} \quad k \longrightarrow \infty, \tag{3.12}$$

by Lemma 2. This probability is insignificant compared to the claimed upper bound for M_k, and therefore we need only deal with the case when $\zeta k < t \leq k$.

Assume henceforth that $\zeta k < t \leq k$. Let N be the number of times that some g_j (with $s_j \in \mathcal{N}$) belongs to some Γ_i (with $i \in T$), and let us estimate the mean and variance of N. As in Grimmett (1991),

$$E'(N) = \sum_{s>k} \frac{t}{s}, \qquad \operatorname{var}'(N) = \sum_{s>k} \frac{t}{s} - \sum_{s>k} \frac{t^2}{s^2}, \tag{3.13}$$

where E', var', and later P', denote expectation, variance, and probability, conditional on the event that $|T| = t$. For large k, the (conditional) mean and variance of N have the same order of magnitude, as indicated by the next lemma.

Lemma 3. *There exists $C\in(0,1)$ and an integer K such that*

$$\operatorname{var}^t(N)\geq CE^t(N),$$

for all $0<t\leq k$ and $k\geq K$.

Proof. It suffices to deal with the case $t=k$, since if $t=\nu k$ where $0<\nu\leq 1$ then $E^t(N)=\nu E^k(N)$ and $\operatorname{var}^t(N)\geq \nu \operatorname{var}^k(N)$ by (3.13).

For $c>1$, define

$$h(c)=\frac{1}{c}+\frac{3}{2}\left(1-\frac{1}{c}\right)(2c-c^\alpha-1).$$

It is easily checked that $h(c)\to 1$ and $h'(c)\to -1$ as $c\downarrow 1$; we pick $c>1$ such that $h(c)<1$. With this choice of c, we find K such that

$$\tfrac{1}{2}(c^\alpha+1)\leq\frac{\pi(cn)}{\pi(n)}\leq\tfrac{1}{2}(3c^\alpha-1)\qquad\text{for all}\qquad n\geq K. \tag{3.14}$$

Now

$$\sum_{s>k}\frac{k^2}{s^2}\leq\sum_{k<s\leq ck}\frac{k}{s}+\frac{1}{c}\sum_{s>ck}\frac{k}{s}=\frac{1}{c}\sum_{s>k}\frac{k}{s}+\left(1-\frac{1}{c}\right)\sum_{k<s\leq ck}\frac{k}{s}. \tag{3.15}$$

Furthermore, if $k\geq K$,

$$\begin{aligned}\frac{\sum_{k<s\leq ck}k/s}{\sum_{s>k}k/s}&\leq\frac{\pi(ck)-\pi(k)}{\sum_{m=0}^{\infty}c^{-(m+1)}(\pi(kc^{m+1})-\pi(kc^m))}\\&\leq\frac{\frac{3}{2}(c^\alpha-1)}{\sum_{m=0}^{\infty}c^{-(m+1)}\frac{1}{2}(c^\alpha-1)\{\frac{1}{2}(c^\alpha+1)\}^m}\quad\text{by (3.14)}\\&=\tfrac{3}{2}(2c-c^\alpha-1).\end{aligned}$$

Hence

$$\sum_{s>k}\frac{k^2}{s^2}\leq h(c)\sum_{s>k}\frac{k}{s}$$

by (3.15), and the result follows from (3.13) and the fact that $h(c)<1$.

It is a consequence of the Berry-Esséen bounds that there exist positive constants γ_1 and γ_2 such that

$$P^t(N=i)\leq\gamma_1\frac{E^t(N)}{\operatorname{var}^t(N)^{3/2}}+\frac{\gamma_2}{\sqrt{\operatorname{var}^t(N)}}, \tag{3.16}$$

for all $t\in(\zeta k,k]$ and i, as in equation (3.15) of Grimmett (1990). Now

$$CE^t(N)\leq\operatorname{var}^t(N)\leq E^t(N),\qquad\text{for all}\quad t\quad\text{and all large}\quad k, \tag{3.17}$$

by (3.13) and Lemma 3, implying that there exists a constant γ_3 such that

$$P^t(N=i)\leq\frac{\gamma_3}{\sqrt{E^k(N)}},\qquad\text{for all}\quad i,\quad \zeta k<t\leq k,\quad\text{and all large}\quad k. \tag{3.18}$$

We have by Chebyshev's inequality that

$$\begin{aligned} P^t(|N-E^t(N)|\geqslant\tfrac{1}{2}E^t(N)) &\leqslant 4\frac{\operatorname{var}^t(N)}{E^t(N)^2}\leqslant\frac{4}{E^t(N)} \\ &\leqslant\frac{4}{E^{\zeta k}(N)}, \qquad \text{if } t>\zeta k, \\ &=\frac{4}{\zeta E^k(N)}, \qquad \text{by (3.13).} \qquad (3.19) \end{aligned}$$

In the limit as $k\to\infty$, the last term has smaller order of magnitude than the claimed upper bound of Theorem 2; we may therefore restrict our attention to the case when $|N-E^t(N)|<\frac{1}{2}E^t(N)$.

Let R be the number of distinct integers i lying in T such that $g_j\in\Gamma_i$ for some $s_j\in\mathcal{N}$. Then

$$P^t(R=r)\leqslant\frac{4}{E^t(N)}+\sum_{r\leqslant i<\frac{3}{2}E^t(N)}P^t(R=r\,|\,N=i)P^t(N=i) \qquad (3.20)$$

by (3.19). We may think of R as the number of occupied boxes when N balls are distributed randomly into $|T|$ boxes. Writing $Y=N-R$ for the number of "wasted" balls, that is, balls which are allocated to boxes which are already occupied, we have from (2.7) and (2.8) that $E^t(Y\,|\,N=i)=\lambda_{it}$ where

$$\lambda_{it}=\frac{i^2}{2t}(1+o(1)), \qquad (3.21)$$

and furthermore

$$\operatorname{var}^t(Y\,|\,N=i)=\frac{i^2}{2t}(1+o(1)); \qquad (3.22)$$

here and later, the $o(1)$ terms refer to the limit as $k\to\infty$ and are uniform in $t\in(\zeta k, k]$ and $i\in(\frac{1}{2}E^t(N),\frac{3}{2}E^t(N))$.

At this point we need to treat the cases $\alpha<\frac{1}{2}$, $\alpha=\frac{1}{2}$, and $\alpha>\frac{1}{2}$ separately. Suppose for the moment that $0<\alpha\leqslant\frac{1}{2}$. Using (3.21), (3.22), and (3.13), we find that, if $t\in(\zeta k, k]$ and $i\in(\frac{1}{2}E^t(N),\frac{3}{2}E^t(N))$, it is the case that λ_{it} and $\operatorname{var}^t(Y\,|\,N=i)$ are no larger than

$$\frac{9E^t(N)^2}{8t}(1+o(1))\leqslant\frac{2E^k(N)^2}{\zeta k}, \qquad (3.23)$$

for all large k. Following Grimmett (1991), we obtain by Chebyshev's inequality that

$$\begin{aligned} P^t(R=r\,|\,N=i) &\leqslant P^t(Y\geqslant i-r\,|\,N=i) \\ &\leqslant\frac{\operatorname{var}^t(Y\,|\,N=i)}{(i-r-\lambda_{it})^2} \qquad \text{if} \qquad i-r>\lambda_{it}. \qquad (3.24) \end{aligned}$$

Hence the summation in (3.20) is no larger than

$$\left\{2+\frac{2E^k(N)^2}{\zeta k}+\sum_{r+\mu_k+1\leqslant i<\frac{3}{2}E^t(N)}\frac{\operatorname{var}^t(Y\,|\,N=i)}{(i-r-\mu_k)^2}\right\}\sup_i\{P^t(N=i)\}, \qquad (3.25)$$

where $\mu_k = 2E^k(N)^2/(\zeta k)$. Using (3.18) and (3.23), we see that the last expression is no larger than

$$\left\{2+\frac{6E^k(N)^2}{\zeta k}\right\}\frac{\gamma_3}{\sqrt{E^k(N)}}. \tag{3.26}$$

Therefore, as in (3.10) of Grimmett (1991),

$$\begin{aligned} P(\Sigma_k = j) &\leqslant P(|T| \leqslant \zeta k) + \sum_{t>\zeta k} P^t(R = t-j)P(|T| = t) \\ &\leqslant \frac{1+o(1)}{k} + \sup_{\substack{0\leqslant r\leqslant t \\ \zeta k<t\leqslant k}} \{P^t(R=r)\} \\ &\leqslant \frac{1+o(1)}{k} + \frac{4}{\zeta E^k(N)} + \left\{2+\frac{6E^k(N)^2}{\zeta k}\right\}\frac{\gamma_3}{\sqrt{E^k(N)}} \end{aligned} \tag{3.27}$$

by (3.12), for all large k. If $0<\alpha<\frac{1}{2}$, then $k^{-1}E^k(N)^2 \to 0$ by (3.13) and Lemma 1, thus implying the appropriate part of (1.3).

Suppose next that $\frac{1}{2}<\alpha<1$; the argument is slightly more complicated in this case, and it is here that we shall appeal to Theorem 3. Now $P^t(R=r \mid N=i)$ is the probability that i balls, allocated one by one and independently to t boxes, occupy exactly r boxes. With $Y = N-R$, we have as in (3.20)

$$P^t(R=r) \leqslant \frac{4}{E^t(N)} + \sum_{\frac{1}{2}E^t(N)<i<\frac{3}{2}E^t(N)} P^t(Y=i-r \mid N=i)P^t(N=i) \tag{3.28}$$

(note that the summand is 0 if $i<r$). We may assume as before that $\zeta k < t \leqslant k$. In the notation of Theorem 3, we have from Lemma 1 and (3.13) that

$$(\tfrac{1}{2}E^t(N))^2 \geqslant 2t\log t, \qquad \tfrac{3}{2}E^t(N) \leqslant C_1 t, \qquad \text{for} \qquad \zeta k<t\leqslant k, \tag{3.29}$$

for all large k, and hence

$$\begin{aligned} &P^t(Y=i-r \mid N=i) \\ &\qquad \leqslant P(\Pi_{it}=i-r)+\frac{C_2}{\sqrt{t}}, \qquad \text{for} \qquad \tfrac{1}{2}E^t(N)<i<\tfrac{3}{2}E^t(N), \quad r\in\mathbb{Z}, \end{aligned}$$

by Theorem 3, where Π_{it} is a random variable having the Poisson distribution with mean λ_{it} satisfying (3.21). The summation in (3.28) is therefore no larger than $S_1+S_2+S_3$, where the S_r's are the following sums over the variable i:

$$S_1 = \sum_{\substack{|i-r-\lambda_{it}|\leqslant cE^t(N)\sqrt{t^{-1}\log t} \\ \frac{1}{2}E^t(N)<i<\frac{3}{2}E^t(N)}} P(\Pi_{it}=i-r)P^t(N=i);$$

$$S_2 = \sum_{\substack{|i-r-\lambda_{it}|> cE^t(N)\surd t^{-1}\log t \\ \frac{1}{2}E^t(N)<i<\frac{3}{2}E^t(N)}} P(\Pi_{it}=i-r)P^t(N=i);$$

$$S_3 = \sum_i \frac{C_2}{\sqrt{t}}P^t(N=i) \leqslant \frac{C_2}{\sqrt{\zeta k}}, \qquad \text{for} \qquad \zeta k<t\leqslant k; \tag{3.30}$$

and c is a constant to be picked soon. To bound S_1, we use the facts that

$$P(\Pi_{it}=j) \leq \frac{1}{\sqrt{\lambda_{it}}} \qquad \text{for all} \quad j,$$

and

$$\frac{E^t(N)^2}{9t} \leq \lambda_{it} \leq \frac{2E^t(N)^2}{t} \quad \text{for} \quad \tfrac{1}{2}E^t(N) < i < \tfrac{3}{2}E^t(N) \quad \text{and all large} \quad k, \tag{3.31}$$

by (3.21). Hence, for all large k,

$$S_1 \leq Q \frac{\sqrt{9t}}{E^t(N)} \sup_i \{P^t(N=i)\},$$

where Q is the number of integers i for which $|i-r-\lambda_{it}| \leq cE^t(N)\sqrt{t^{-1}\log t}$. Using (3.21) and a little calculus, one finds that $Q \leq 4cE^t(N)\sqrt{t^{-1}\log t}$ for all large k, uniformly in i and t. Hence

$$S_1 \leq 4c\gamma_3 \sqrt{\frac{\log k}{E^k(N)}}, \tag{3.32}$$

for all large k, where we have used (3.18).

As for S_2,

$$S_2 \leq \sup_{\substack{\frac{1}{2}E^t(N)<i<\frac{3}{2}E^t(N)\\ \zeta k<t\leq k}} P\left(\frac{|\Pi_{it}-\lambda_{it}|}{\sqrt{\lambda_{it}}} > cE^t(N)\sqrt{\frac{\log t}{t\lambda_{it}}}\right).$$

Now Π_{it} is approximately distributed as the sum of λ_{it} independent summands, each having the Poisson distribution with mean 1. For the specified ranges of i and t, it is the case that $cE^t(N)\sqrt{\log t/(t\lambda_{it})} = o(\sqrt{\lambda_{it}})$ as $k \to \infty$. Hence, by large-deviation theory (see for example Feller (1971, p. 553)), there exists $\eta > 0$ such that

$$S_2 \leq \sup_{\substack{\frac{1}{2}E^t(N)<i<\frac{3}{2}E^t(N)\\ \zeta k<t\leq k}} \left\{\exp\left(-\eta c^2 E^t(N)^2 \frac{\log t}{t\lambda_{it}}\right)\right\},$$

which, by (3.31), is no larger than $(\zeta k)^{-\eta c^2/2}$ for all large k. Picking c such that $\eta c^2 > 2\alpha$, say, we find that

$$S_2 \leq (\zeta k)^{-\alpha} \qquad \text{for all large} \quad k, \tag{3.33}$$

which, by Lemma 1, is smaller in order than $\{E^k(N)\}^{-1/2}$.

Combining these inequalities for S_1, S_2 and S_3, we deduce from (3.28) and Lemma 1 that

$$\sup_{\substack{r,t:\\ \zeta k<t\leq k}} \{P^t(R=r)\} \leq \gamma_4 \sqrt{\frac{\log k}{E^k(N)}},$$

for some constant γ_4 and all large k, whence the conclusion of the theorem follows as in (3.27).

There remains only the case $\alpha = \frac{1}{2}$. In the light of (3.27), which is valid when $\alpha = \frac{1}{2}$, it will suffice to show that, for $0 < \varepsilon < 1$, there exists a constant γ_5 such that

$$P(\Sigma_k = j) \leqslant \gamma_5 \frac{\sqrt{\log k}\,\{\log\log k\}^{1+\varepsilon}}{\sqrt{E^k(N)}} \quad \text{for all} \quad \text{large } k. \tag{3.34}$$

Returning to (3.20), we suppose that $\zeta k < t \leqslant k$, and we split the summation into two parts depending on whether $i^2 < 2t \log t$ or $i^2 \geqslant 2t \log t$. The contribution from values of i satisfying $i^2 \geqslant 2t \log t$ may be dealt with in the same manner as was the case $\frac{1}{2} < \alpha < 1$, whence it follows that

$$\sum_{\substack{\frac{1}{2}E^t(N) < i < \frac{3}{2}E^t(N) \\ i^2 \geqslant 2t \log t}} P^t(R = r \mid N = i) P^t(N = i) \leqslant \gamma_6 \sqrt{\frac{\log k}{E^k(N)}}, \tag{3.35}$$

for some constant γ_6 and all large k.

Turning to the remainder of the sum (for $i^2 < 2t \log t$), we have that

$$\begin{aligned} P^t(R = r \mid N = i) &\leqslant P^t(|Y - \lambda_{it}| \geqslant |i - r - \lambda_{it}| \mid N = i) \\ &\leqslant \frac{E^t(g(|Y - \lambda_{it}|) \mid N = i)}{g(|i - r - \lambda_{it}|)} \end{aligned} \tag{3.36}$$

by Markov's inequality (see for example Grimmett and Stirzaker (1982, p. 186)) where $g(x) = x\{\log(1+x)\}^{1+\varepsilon}$. There exists a positive integer S such that, for all $0 < \varepsilon < 1$, $h(x) = g(\sqrt{x})$ is concave on $[S, \infty)$; hence $Z = S + |Y - \lambda_{it}|$ is such that

$$\begin{aligned} E^t(g(|Y - \lambda_{it}|) \mid N = i) &\leqslant E^t(g(Z) \mid N = i) = E^t(h(Z^2) \mid N = i) \\ &\leqslant h(E^t(Z^2 \mid N = i)) = g(\sqrt{E^t(Z^2 \mid N = i)}), \end{aligned} \tag{3.37}$$

by Jensen's inequality (see for example Grimmett and Stirzaker (1982, p. 216)). Now

$$\begin{aligned} E^t(Z^2 \mid N = i) &= S^2 + 2SE^t(|Y - \lambda_{it}| \mid N = i) + \operatorname{var}^t(Y \mid N = i) \\ &\leqslant S^2 + 2S\sqrt{\operatorname{var}^t(Y \mid N = i)} + \operatorname{var}^t(Y \mid N = i) \leqslant \gamma_7 \log k, \end{aligned}$$

for some constant γ_7 and all large k, by (3.22), the Cauchy-Schwarz inequality, and the assumption that $i^2 < 2t \log t$. It follows from (3.36) that

$$\sum_{\substack{r \leqslant i < \frac{3}{2}E^t(N) \\ i^2 < 2t \log t}} P^t(R = r \mid N = i) \leqslant 2g(\sqrt{\gamma_7 \log k}) \left\{ 1 + \sum_{\substack{i: \\ r + \lambda_{it} + 1 \leqslant i < 2t \log t}} \frac{1}{g(|i - r - \lambda_{it}|)} \right\}.$$

With the aid of a change of variables and the fact that $\lambda_{i+1,t} - \lambda_{it} \leqslant i/t$, it may be seen that the last sum is bounded uniformly in k and r, whence

$$\sum_{\substack{r \leqslant i < \frac{3}{2}E^t(N) \\ i^2 < 2t\log t}} P^t(R=r \mid N=i)P^t(N=i)$$

$$\leqslant \gamma_8 \sqrt{\log k}\,\{\log\log k\}^{1+\varepsilon} \sup_i \{P^t(N=i)\},$$

for some constant γ_8 and all large k. We combine this with (3.35), in a way similar to that used in (3.27), to find that

$$P(\Sigma_k = j) \leqslant \frac{1+o(1)}{k} + \gamma_9 \frac{\sqrt{\log k}\,\{\log\log k\}^{1+\varepsilon}}{\sqrt{E^k(N)}},$$

for some constant γ_9 and all large k. This implies (3.34), and the proof is complete.

We presented earlier a proof that $\sigma_k^2 = \operatorname{var}(\Sigma_k)$ satisfies $\sigma_k^2 \geqslant k^{\alpha-\varepsilon}$ for all $\varepsilon > 0$ and all large k; see the remark after the statement of Lemma 2. The same conclusion may be reached without appealing directly to Theorem 1. The argument is roughly as follows. The number $k - \Sigma_k$ of integers "hit" by $\mathscr{S}$ in $\{1, 2, \ldots, k\}$ may be expressed as the sum $M+R$, where $M = |\{1, 2, \ldots, k\} \backslash T|$ is the number of integers hit by members of $\mathscr{M} = \{s \in \mathscr{S}\colon s \leqslant k\}$, and R is the number of integers in T which are hit by members of $\mathscr{N} = \{s \in \mathscr{S}\colon s > k\}$. Therefore $\sigma_k^2 = \operatorname{var}(M+R)$. The random variables M and R are not independent, but nevertheless it may be shown that the "degree of uncertainty" of Σ_k is at least as great as that of R; when correctly phrased, this implies that σ_k^2 cannot be smaller in order than $\operatorname{var}(R)$. Typically, R differs only little from N, which has variance of order equal to its mean, which in turn is around k^α.

Acknowledgement. We thank the Consiglio Nazionale delle Ricerche for support to the first author while visiting the University of Rome II "Tor Vergata" where part of this work was done. We are grateful to Andrew Barbour for his prompt despatch by electronic mail of a draft in TEX of the appropriate part of Barbour *et al.* (1991).

References

A. D. Barbour and L. Holst (1989). Some applications of the Stein–Chen method for proving Poisson convergence. *Advances in Applied Probability*, 21, 74–90.

A. D. Barbour, L. Holst and S. Janson (1991). *Poisson Approximation* (Oxford University Press).

N. H. Bingham, C. M. Goldie and J. Teugels (1987). *Regular Variation* (Cambridge University Press, Cambridge).

Y. S. Chow and H. Teicher (1978). *Probability Theory* (Springer-Verlag, New York).

W. Feller (1971). *An Introduction to Probability Theory and its Applications*, volume 2 (John Wiley and Sons, New York).

G. R. Grimmett (1991). The statistics of sieves and the square-free numbers. *J. London Math. Soc.*, 43, 1–11.

G. R. Grimmett and D. R. Stirzaker (1982). *Probability and Random Processes* (Clarendon Press, Oxford).

R. R. Hall (1982). Squarefree numbers on short intervals. *Mathematika*, 29, 7-17.

R. R. Hall (1989). The distribution of squarefree numbers. *J. für reine u. ange. Math.*, 394, 107-117.

V. F. Kolchin, B. A. Sevastyanov and V. P. Chistyakov (1978). *Random Allocations* (J. Wiley and Sons, New York).

Professor G. R. Grimmett,
School of Mathematics,
University of Bristol,
University Walk,
Bristol. BS8 1TW

Professor R. R. Hall,
Department of Mathematics,
University of York,
Heslington,
York. YO1 5DD

11N25: *NUMBER THEORY; Multiplicative number theory; Distribution of integers with specified multiplicative constraints.*

Received on the 15*th of October*, 1990.