Small counts in the infinite occupancy scheme

A. D. Barbour and A. V. Gnedin
University of Zürich and Utrecht University

Abstract

The paper is concerned with the classical occupancy scheme with infinitely many boxes, in which \( n \) balls are thrown independently into boxes 1, 2, \ldots, with probability \( p_j \) of hitting the box \( j \), where \( p_1 \ge p_2 \ge \ldots > 0 \) and \( \sum_{j=1}^{\infty} p_j = 1 \). We establish joint normal approximation as \( n \to \infty \) for the numbers of boxes containing \( r_1, r_2, \ldots, r_m \) balls, standardized in the natural way, assuming only that the variances of these counts all tend to infinity. The proof of this approximation is based on a de-Poissonization lemma. We then review sufficient conditions for the variances to tend to infinity. Typically, the normal approximation does not mean convergence. We show that the convergence of the full vector of \( r \)-counts only holds under a condition of regular variation, thus giving a complete characterization of possible limit correlation structures.

Key words: occupancy problem, normal approximation, poissonization, regular variation.

AMS 2000 Subject Classification: Primary 60F05, 60C05.

Submitted to EJP on September 25, 2008, final version accepted January 14, 2009.
1 Introduction

In the classical occupancy scheme with infinitely many boxes, balls are thrown independently into boxes 1, 2, ..., with probability $p_j$ of hitting the box $j$, where $p_1 \geq p_2 \geq \ldots > 0$ and $\sum_{j=1}^{\infty} p_j = 1$. The most studied quantity is the number of boxes $K_n$ occupied by at least one out of the first $n$ balls thrown. It is known that for large $n$ the law of $K_n$ is asymptotically normal, provided that $\text{Var}[K_n] \to \infty$; see [6, 7] for references and a survey of this and related results. In this paper, we investigate the behaviour of the quantities $X_{n,r}$, the numbers of boxes hit by exactly $r$ out of the $n$ balls, $r \geq 1$.

Under a condition of regular variation, a multivariate CLT for the $X_{n,r}$’s was proved by Karlin [8]. Mikhailov [12] also studied the $X_{n,r}$’s, but in a situation where the $p_j$’s vary with $n$. In this paper, we establish joint normal approximation as $n \to \infty$ for the variables $X_{n,r}, \ldots, X_{n,r_m}$, centred and normalized, assuming only that $\lim_{n \to \infty} \text{Var} X_{n,r_i} = \infty$ for each $i$. We also give examples to show that this condition is not enough to ensure convergence, since the correlation matrices need not converge as $n \to \infty$. The asymptotic behaviour of the moments of the $X_{n,r}$ is thus of key importance, and we discuss this under a number of simplifying assumptions.

The behaviour of these moments, as also of those of $K_n = \sum_{r=1}^{\infty} X_{n,r}$, depends on the way in which the frequencies $p_j$ decay to 0. In the case of power-like decay, $p_j \sim cq^{-1/\alpha}$ with $0 < \alpha < 1$, it is known that, for each fixed $k$, the moments $\mathbb{E}X_{n,r}^k$ have the same order of growth with $n$ for every $r$, and this is the same order of growth as that of $\mathbb{E}K_n^k$; moreover, the limit distributions of $K_n$ and of $X_n := (X_{n,1}, X_{n,2}, \ldots)$ are normal [6, 8]. In contrast, for a sequence of geometric frequencies $p_j = cq^j$ ($0 < q < 1$), there is no way to scale the $X_{n,r}$’s to obtain a nontrivial limit distribution [10], and the moments of $K_n$ have oscillatory asymptotics. In a more general setting such that the $p_j$’s have exponential decay, the oscillatory behaviour of $\text{Var}[K_n]$ is typical [3]. The spectrum of interesting possibilities is, however, much wider: for instance, frequencies $p_j \sim ce^{-j\beta}$, with $0 < \beta < 1$, exhibit a decay intermediate between power and exponential.

Karlin’s [8] multivariate CLT for $X_n$ applies when the index of regular variation is in the range $0 < \alpha < 1$. We complement this by the analysis of the cases $\alpha = 0$ and $\alpha = 1$, showing that for each $\alpha \in [0,1]$ there is exactly one possible normal limit. Finally, we prove that these one-parameter normal laws are the only possible limits of naturally scaled and centred $X_n$. Specifically, we show that a regular variation condition holds if $\text{Var}X_{n,r} \to \infty$ for all $r$ and if all the correlations $\{\text{Corr}(X_{n,r}, X_{n,s}), r,s \geq 1\}$ converge.

2 Poissonization

As in much previous work, we shall rely on a closely related occupancy scheme, in which the balls are thrown into the boxes at the times of a unit Poisson process. The advantage of this model is that, for every $t > 0$, the processes $(N_j(t), t \geq 0)$, counting the numbers of balls in boxes $j = 1, 2, \ldots$, are independent. Let $Y_r(t)$ be the number of boxes occupied by exactly $r$ balls at time $t$. In view of the representation

$$Y_r(t) = \sum_{j=1}^{\infty} 1[N_j(t) = r] \tag{2.1}$$
Lemma 2.1. For any \( m \in \mathbb{N} \) satisfying \( m \leq \frac{1}{2} np_k \), we have

\[
d_{TV}(\mathcal{L}(X_{n,1}, \ldots, X_{n,m}), \mathcal{L}(Y_1(n), \ldots, Y_m(n))) \leq \pi_k + 2k e^{-np_k/10},
\]

where \( \pi_j := \sum_{i=j+1}^{\infty} p_i \).

**Proof.** We begin by noting that, in parallel to (2.1),

\[
X_{n,r} := \sum_{j=1}^{\infty} 1[M_{n,j} = r],
\]

where \( M_{n,j} \) represents the number of balls out of the first \( n \) thrown that fall into box \( j \). Our proof uses lower truncation of the sums (2.1) and (2.3) that define \( Y_r(t) \) and \( X_{n,r} \).

Since \( M_{n,j} \sim \text{Binomial}(n, p_j) \), it follows from the Chernoff inequalities [8] that, if \( m \leq \frac{1}{2} np_k \), then for \( j \leq k \)

\[
\mathbb{P}[M_{n,j} \leq m] \leq \mathbb{P}[M_{n,j} \leq \frac{1}{2} np_j] \leq \exp\{-np_j/10\} \leq \exp\{-np_k/10\},
\]

since the \( p_j \) are decreasing, and \( m \leq \frac{1}{2} np_k \); and the same bound holds also for \( N_j(n) \sim \text{Poisson}(np_j) \).

Hence, defining

\[
X_{n,k,r} := \sum_{j=k+1}^{\infty} 1[M_{n,j} = r], \quad Y_{k,r}(t) := \sum_{j=k+1}^{\infty} 1[N_j(t) = r],
\]

it follows that

\[
d_{TV}(\mathcal{L}(X_{n,1}, \ldots, X_{n,m}), \mathcal{L}(X_{n,k,1}, \ldots, X_{n,k,m})) \leq ke^{-np_k/10}, \quad (2.4)
\]

\[
d_{TV}(\mathcal{L}(Y_1(t), \ldots, Y_m(t)), \mathcal{L}(Y_{k,1}(t), \ldots, Y_{k,m}(t))) \leq ke^{-np_k/10}. \quad (2.5)
\]

But now, from an inequality of Le Cam [4] and Michel [11], we have

\[
d_{TV}(\mathcal{L}(N_j(n), j \geq k+1), \mathcal{L}(M_{n,j}, j \geq k+1)) \leq \pi_k, \quad (2.6)
\]

and the \( X_{n,k,r} \) are functions of \( \{M_{n,j}, j \geq k+1\} \), the \( Y_{k,r}(n) \) of \( \{N_j(n), j \geq k+1\} \). The lemma now follows from (2.4), (2.5) and (2.6).

**Proposition 2.2.** Let \( k(n) \) be any sequence satisfying

\[
k(n) \to \infty \quad \text{and} \quad k(n)e^{-np_k(n)/10} \to 0.
\]

Then, for any sequence \( m(n) \) satisfying \( m(n) \leq \frac{1}{2} np_{k(n)} \) for each \( n \), it follows that

\[
d_{TV}(\mathcal{L}(X_{n,1}, \ldots, X_{n,m(n)}), \mathcal{L}(Y_1(n), \ldots, Y_{m(n)}(n))) \to 0. \quad (2.7)
\]
Proof. Since \( m(n) \leq \frac{1}{2} np_{k(n)} \) for each \( n \), it follows that Lemma \[2.1\] can be applied for each \( n \). Since \( k(n) \to \infty \), it follows that \( \pi_{k(n)} \to 0 \), so that the first element in its bound converges to zero; the second converges to zero also, by assumption. ∎

Remark. Such sequences \( k(n) \) always exist. For instance, one can take

\[
k(n) = \max\{k: \ 20 \log k/p_k \leq n\}.
\]

For this choice, it is immediate that \( k(n) \to \infty \), and that \( np_{k(n)} \geq 20 \log k(n) \to \infty \), entailing also that \( k(n)e^{-np_{k(n)}/10} \leq 1/k(n) \to 0 \). Hence there are always sequences \( m(n) \to \infty \) for which (2.7) is satisfied.

Hence, in particular, any approximation to the distribution of a finite subset of the components of \( Y(n) = (Y_1(n), Y_2(n), \ldots) \) (suitably scaled) remains valid for the corresponding components of \( X_n \), at the cost of introducing an extra, asymptotically negligible, error in total variation of at most

\[
\pi_{k(n)} + 2k(n)e^{-np_{k(n)}/10},
\]

(2.8)

where \( k(n) \) is any sequence satisfying the conditions of Proposition \[2.2\].

3 Normal approximation

As noted above, the distribution of \( Y_r(t) \) is asymptotically normal as \( t \to \infty \) whenever \( \text{Var} Y_r(t) \to \infty \). Here, we consider the joint normal approximation of any finite set of counts \( Y_{r_1}(t), \ldots, Y_{r_m}(t) \) such that \( r_i \geq 1 \) and \( \lim_{t \to \infty} \text{Var} Y_{r_i}(t) = \infty \) for each \( 1 \leq i \leq m \). We measure the closeness of two probability measures \( P \) and \( Q \) on \( \mathbb{R}^m \) in terms of differences between the probabilities assigned to arbitrary convex sets:

\[
d_c(P, Q) := \sup_{A \in \mathcal{C}} |P(A) - Q(A)|,
\]

where ‘\( \mathcal{C} \)’ denotes the class of convex subsets of \( \mathbb{R}^m \). Let

\[
\Phi_r(t) := \mathbb{E}Y_r(t), \quad V_r(t) := \text{Var} Y_r(t), \quad C_{rs}(t) := \text{Cov}(Y_r(t), Y_s(t))
\]

denote the moments of the \( Y_r(t) \), and let

\[
\Sigma_{rs}(t) := C_{rs}(t)/\sqrt{V_r(t)V_s(t)} = \text{Cov}(Y_r(t), Y_s(t))
\]

denote the covariance matrix of the standardized random variables \( Y'_r(t) \) as in (2.2).

Now the random vector \( (Y'_{r_1}(t), \ldots, Y'_{r_m}(t)) \) is a sum of independent mean zero random vectors \( (Y'_{l_{r_1}}(t), \ldots, Y'_{l_{r_m}}(t)) \), \( l \geq 1 \), where \( Y'_{l_{r_i}}(t) := (1[N_l(t) = r] - p_{l,r_i}(t))/\sqrt{V_r(t)} \), and

\[
p_{l,r}(t) := \mathbb{P}[N_l(t) = r] = e^{-tp_l(t)p_l(t)}/r!.
\]

(3.1)

A theorem of Bentkus [1, Thm. 1.1] then shows that

\[
d_c(\mathcal{L}(Y'_{r_1}(t), \ldots, Y'_{r_m}(t)), \text{MVN}_m(0, \Sigma(t))) \leq Cm^{1/4} \beta_t,
\]

368
for an absolute constant $C$, where

$$
\beta_t := \sum_{l \geq 1} \beta_{t,l} \quad \text{and} \quad \beta_{t,j} := \mathbb{E}[\Sigma_R^{-1/2}(t)(Y'_{t,r_1}(t), \ldots, Y'_{t,r_m}(t))^T]_j^3,
$$

and $\Sigma_R(t)$ denotes the $m \times m$ matrix with elements $\{\Sigma_{rs}(t), r, s \in R := \{r_1, \ldots, r_m\}$). Applying this result, we obtain the following theorem.

**Theorem 3.1.** If $\lim_{t \to \infty} V_{r_1}(t) = \infty$ for each $1 \leq i \leq m$, where $1 \leq r_1 < \ldots < r_m$, then, as $t$ and $n$ tend to $\infty$,

$$
d_c(\mathcal{L}(X', \ldots, X', 0, \Sigma_R(t))), \text{MVN}_m(0, \Sigma_R(t))) = O\left(1/\min_{1 \leq i \leq m} \sqrt{V_{r_1}(t)}\right) \to 0;
$$

$$
d_c(\mathcal{L}(X', \ldots, X', 0, \Sigma_R(t))), \text{MVN}_m(0, \Sigma_R(n))) = O\left(\pi_k(n) + 2k(n)e^{-npk(n)/10} + \left\{1/\min_{1 \leq i \leq m} \sqrt{V_{r_1}(n)}\right\}\right) \to 0,
$$

where $k(n)$ is any sequence chosen as for Proposition 2.2 and satisfying $\max_{1 \leq i \leq m} r_j \leq \frac{1}{2} npk(n)$ for each $n$. If, in addition, $\Sigma_R(t) \to \Sigma_R$ as $t \to \infty$, for some fixed $\Sigma_R$, then

$$(Y'_{r_1}(t), \ldots, Y'_{r_n}(t)) \to_d \text{MVN}_m(0, \Sigma_R) \quad \text{and} \quad (X'_{n,r_1}, \ldots, X'_{n,r_m}) \to_d \text{MVN}_m(0, \Sigma_R).$$

**Proof.** All that we need to do is to control the quantity $\beta_t$. This in turn involves bounding the smallest eigenvalue of $\Sigma_R(t)$ away from 0. Now, direct calculation shows that, for any column vector $\alpha \in \mathbb{R}^m$,

$$
a^T \Sigma_R(t) a = \text{Var}\left(\sum_{j=1}^m a_j Y'_{r_j}(t)\right) = \sum_{l \geq 1} \text{Var}\left(\sum_{j=1}^m a_j Y'_{r_j}(t)\right).
$$

Using the definition of $Y'_{r_j}(t)$, this gives

$$
a^T \Sigma_R(t) a = \sum_{l \geq 1} \left\{p^{l,R}(t)E^{l,R,t}(U^2) - \left[p^{l,R}(t)\right]^2 \left\{E^{l,R,t}(U^2)\right\}\right\},
$$

where $p^{l,R}(t) := \sum_{r \in R} p_{l,r}(t)$ and, under the measure $P^{l,R,t}$, $U$ takes the value $a_j/\sqrt{V_{r_j}(t)}$ with probability $p_{l,r_j}(t)/p^{l,R}(t)$, $1 \leq j \leq m$. This in turn implies that

$$
a^T \Sigma_R(t) a \geq \sum_{l \geq 1} \sum_{j=1}^m p^{l,R}(t)(1 - p^{l,R}(t))E^{l,R,t}(U^2),
$$

and since

$$
E^{l,R,t}(U^2) = \sum_{j=1}^m \frac{p_{l,r_j}(t) a_j^2}{V_{r_j}(t)},
$$

it follows that

$$
a^T \Sigma_R(t) a \geq \sum_{l \geq 1} \sum_{j=1}^m \frac{p_{l,r_j}(t) a_j^2}{V_{r_j}(t)} \geq \min_{l \geq 1} (1 - p^{l,R}(t)) a^T a,
$$

369
since $V_r(t) \leq \sum_{l \geq 1} p_{l,r}(t)$. However, for each $l$, $p_{l,R}(t) \leq 1 - p_{l,0}(t) - \sum_{j > r_m} p_{l,j}(t)$, and $(p_{l,r}(t), r \geq 1)$ are just the Poisson probabilities (3.1). Hence $1 - p_{l,R}(t) \geq e^{-1}$ if $tp_l \leq 1$, and $1 - p_{l,R}(t) \geq q(r_m) := \text{Poisson}(1)[(r_m + 1, \infty)]$ if $tp_l > 1$, implying that

$$\min_{l \geq 1} (1 - p_{l,R}(t)) \geq c_R := \min\{e^{-1}, q(m)\} > 0,$$

for all $t$. It thus follows that $a^T \Sigma_R(t)a \geq c_R a^T a$ for all $a \in \mathbb{R}^m$.

It is now immediate that, for any $x \in \mathbb{R}^m$, $|\Sigma_R^{-1/2}(t)x| \leq c_R^{-1/2}|x|$, and hence, since $|Y_{l,r}(t)| \leq 1/\sqrt{V_r(t)}$ a.s., we have

$$|\Sigma_R^{-1/2}(t)Y_{l,r_1}(t), \ldots, Y_{l,r_m}(t)|^2 \leq c_R^{-3/2} \frac{\sum_{j=1}^m |Y_{l,r_j}(t)|^2 \sqrt{m}}{\min_{1 \leq i \leq m} \sqrt{V_r(t)}};$$

taking expectations and adding over $l \geq 1$ gives $\beta_l \leq (m/c_R)^{3/2} / \min_{1 \leq i \leq m} \sqrt{V_r(t)}$, proving the first statement of the theorem. The second follows in view of (2.8).

Thus multivariate normal approximation is always good if the variances of the (unstandardized) components $Y_r(t)$ are large. However, convergence typically does not take place: see a series of examples in Proposition 4.4 below.

### 4 Moments

For normal approximation, in view of Theorem 3.1 we are particularly interested in conditions under which $V_r(t) \to \infty$.

For the moments we have the formulas

$$\Phi_r(t) = \sum_{j=1}^\infty p_{j,r}(t), \quad (4.1)$$

$$V_r(t) = \sum_{j=1}^\infty p_{j,r}(t) \left(1 - p_{j,r}(t)\right) = \Phi_r(t) - 2^{-2r} \binom{2r}{r} \Phi_{2r}(2t), \quad (4.2)$$

$$C_{rs}(t) = -2^{-r-s} \binom{r+s}{r} \Phi_{r+s}(2t), \quad r \neq s, \quad (4.3)$$

where, as above, $p_{j,r}(t) = e^{-tp_j} (tp_j)^r / r!$.

From (4.1) and (4.2) we obtain

$$\Phi_r(t) > V_r(t) > k_r \Phi_r(t),$$

with $k_r > 0$, as is seen by evaluating $\max_{x \geq 0} e^{-x} x^r / r!$ and using the inequalities

$$1 \geq 1 - \frac{e^{-tp_j}(tp_j)^r}{r!} \geq 1 - \frac{e^{-r} r^r}{r!} > 0.$$
It follows that

\[ V_r(t) \to \infty \iff \Phi_r(t) \to \infty; \]

hence, as long as only the convergence to infinity of \( V_r(t) \) is concerned, we can deal with the simpler quantity \( \Phi_r(t) \). This facilitates the proof of the following theorem, showing how the asymptotic behaviour of \( V_r(t) \) for different values of \( r \) is structured.

**Theorem 4.1.** The asymptotic behaviour of the quantities \( V_r(t) \) as \( t \to \infty \) follows one of the following four regimes:

1. \( \lim_{t\to\infty} V_r(t) = \infty \) for all \( r \geq 1 \);
2. \( \limsup_{t\to\infty} V_r(t) = \infty \) for all \( r \geq 1 \), and there exists an \( r_0 \geq 1 \) such that \( \liminf_{t\to\infty} V_r(t) = \infty \) for all \( 1 \leq r \leq r_0 \), and \( \liminf_{t\to\infty} V_r(t) < \infty \) for all \( r > r_0 \);
3. \( \limsup_{t\to\infty} V_r(t) = \infty \) and \( \liminf_{t\to\infty} V_r(t) < \infty \) for all \( r \geq 1 \);
4. \( \sup_t V_r(t) < \infty \) for all \( r \geq 1 \).

**Proof.** Replacing \( V_r \) with \( \Phi_r \) for the argument, the formula (4.1) yields

\[ \Phi_r(t) = \sum_{j \geq 1} e^{-tp_j} \frac{(tp_j)^r}{r!}; \quad \Phi_s(t/2) = \sum_{j \geq 1} e^{-tp_j/2} \frac{(tp_j/2)^s}{s!}. \]

For \( s < r \), the ratio of the individual terms is given by

\[ \frac{e^{-tp_j/2}(tp_j/2)^s}{s!} \geq \min_{y > 0} [e^{y/2}y^{-(r-s)}] \frac{r!}{s!2^s} = \left( \frac{e}{r-s} \right)^{r-s} \frac{r!}{s!2^r}. \]

Hence, for all \( s < r \),

\[ \Phi_s(t/2) \geq \Phi_r(t) \left( \frac{e}{r-s} \right)^{r-s} \frac{r!}{s!2^r}. \quad (4.4) \]

It now follows that if, for some \( r \), \( \lim_{t\to\infty} V_r(t) = \infty \), then \( \lim_{t\to\infty} V_s(t) = \infty \) for all \( 1 \leq s \leq r \) also; and that, if \( \sup_t V_r(t) < \infty \) for some \( r \), then \( \sup_t V_s(t) < \infty \) for all \( s > r \). Hence, to complete the proof, we just need to show that, if \( \sup_t V_r(t) < \infty \) for some \( r \geq 1 \), then \( \sup_t V_s(t) < \infty \).

For this last part, write \( \Phi_r(t) = L_r(t) + R_r(t) \), where

\[ L_r(t) := \sum_{j: tp_j \geq 1} e^{-tp_j}(tp_j)^r/r!; \quad R_r(t) := \sum_{j: tp_j < 1} e^{-tp_j}(tp_j)^r/r!. \quad (4.5) \]

Suppose that \( \sup_t \Phi_r(t) = K < \infty \). Then, for every \( t > 0 \),

\[ L_1(t) \leq r!L_r(t) \leq r!\Phi_r(t) \leq Kr!. \quad (4.6) \]

It thus remains to bound \( R_1(t) \), which in turn can be reduced to finding a bound for

\[ S(t) := \sum_{j: tp_j < 1} tp_j. \]

371
Let \( a_0 \geq a_1 \geq \ldots \geq 0 \) be any decreasing sequence such that \( a_j/a_{j+h} \geq 2 \) holds for some \( h \geq 1 \) and all \( j \geq 1 \). Then \( a_{j+m} \leq a_m \cdot 2^{-i} \) for every \( i \geq 0 \) and \( 0 \leq m < h \). Splitting the \( a_j \)'s into \( h \) subsequences that are dominated by the geometric series, we thus have

\[
\sum_{j \geq 0} a_j \leq \sum_{m=0}^{h} 2a_m \leq 2a_0h.
\]

Now if, for some \( h \geq 1 \), the frequencies \( p_j \) satisfy

\[
p_j/p_{j+h} \geq 2 \quad \text{for all } j \geq 1,
\]

then applying the above result to the sequence \( a_j = tp_{j+\min\{i:tp_i < 1\}} \) for any \( t \) yields the bound \( R_1(t) < S(t) < 2h \), since \( a_0 < 1 \).

On the other hand, if \( p_j/p_{j+h} < 2 \) for some \( j \) and \( h \), then it follows from \( p_j \geq p_{j+1} \geq \ldots \geq p_{j+h} > p_j/2 \) that

\[
L_1(2/p_j) > \sum_{k=j}^{j+h} e^{-2p_k/p_j} \frac{2p_k}{p_j} > e^{-2(h+1)}.
\]

Thus, for any \( h \) such that \( e^{-2(h+1)} > Kr! \), we see that (4.7) must hold, since otherwise (4.6) would be violated for \( t = 2/p_j \). Hence it follows that \( R_1(t) < S(t) < 2e^2Kr! \), and the final part of the lemma is proved. \( \square \)

In particular, in Theorem 3.1 the quantity \( \min_{1 \leq i \leq m} \sqrt{V_r(t)} \) can thus be replaced in the error estimates by \( \Phi_{r_m}(2t) \).

We now turn to finding conditions sufficient for distinguishing the asymptotic behaviour of the \( V_r(t) \). To do so, introduce the measures

\[
v_r(dx) = \sum_{j=1}^{\infty} p_j^r \delta_{p_j}(dx).
\]

Two special cases are \( v_0 \), a counting measure, and \( v_1 \), the probability distribution of a size-biased pick from the \( p_j \)'s. For \( r > 0 \) write (4.1) as

\[
\Phi_r(t) = \frac{t^r}{r!} \int_0^\infty e^{-tx} x^r v_0(dx) = \frac{t^r}{r!} \int_0^\infty e^{-tx} v_r(dx) = \frac{t^{r+1}}{r!} \int_0^\infty e^{-tx} v_r[0,x] dx.
\]

Comparing with standard gamma integrals, it is then immediate that

\[
\liminf_{x \to 0} \frac{v_r[0,x]}{x^r} \leq \liminf_{t \to \infty} \Phi_r(t) \leq \limsup_{t \to \infty} \Phi_r(t) \leq \limsup_{x \to 0} \frac{v_r[0,x]}{x^r}.
\]

This, together with Theorem 4.1 enables us to conclude the following conditions for the convergence to infinity of \( \Phi_r(t) \), and hence equivalently of \( V_r(t) \), expressed in terms of the accessible quantities

\[
\rho_{j,r} := \frac{1}{p_j} \sum_{i=j+1}^{\infty} p_i^r.
\]
Lemma 4.2.

(a) \( \sup_{t \geq 0} \Phi_t(t) < \infty \) for all \( s \geq 1 \) if and only if, for some (and then for all) \( r \geq 1 \), \( \sup_j \rho_{j,r} < \infty \).

(b) If, for some \( r \geq 1 \), \( \lim_{j \to \infty} \rho_{j,r} = \infty \), then \( \lim_{t \to \infty} \Phi_t(t) = \infty \) for all \( 1 \leq s \leq r \).

Proof. If \( p_{j+1} \leq x < p_j \) then

\[
\rho_{j,r} = \frac{\nu_r[0,p_{j+1}]}{p_j} = \frac{\nu_r[0,x]}{p_j} < \frac{\nu_r[0,x]}{x^r} \leq \frac{\nu_r[0,p_{j+1}]}{p_{j+1}} = 1 + \rho_{j+1,r}.
\]

Hence (4.9) can be replaced by the inequalities

\[
\liminf_{j \to \infty} \rho_{j,r} \leq \liminf_{t \to \infty} \Phi_t(t) \leq \limsup_{t \to \infty} \Phi_t(t) \leq 1 + \limsup_{j \to \infty} \rho_{j,r}.
\]

Part (b) of the lemma now follows directly from Theorem 4.1.

For part (a), much as for the last part of the proof of Theorem 4.1 define

\[
h(j) := \max\{l \geq 0 : p_{j+l}/p_j \geq 1/2\}; \quad h^* := \sup_j h(j).
\]

Then it is immediate that

\[
2^{-r} h(j) \leq \rho_{j,r} \leq h^* \sum_{l \geq 1} 2^{-(l-1)} = 2h^*,
\]

so that \( h^* \leq \infty \) if and only if \( \sup_j \rho_{j,r} < \infty \) for some, and then for all, \( r \geq 1 \). We now conclude the proof by showing that \( \sup_{t \geq 0} \Phi_t(t) < \infty \) for all \( s \geq 1 \) if and only if \( h^* \leq \infty \). Defining \( L_r(t) \) and \( R_r(t) \) as in (4.5), we observe that, if \( h^* < \infty \), then

\[
R_r(t) \leq h^* \sum_{l \geq 1} 2^{-r(l-1)} \leq 2h^* \quad \text{and} \quad L_r(t) \leq h^* \sum_{l \geq 1} e^{-2^{l-1} 2^{l/r}} r!,
\]

so that \( \Phi_t(t) = L_r(t) + R_r(t) < \infty \) for all \( r \geq 1 \). On the other hand,

\[
L_r(1/p_j + h(j)) \geq e^{-2h(j)}/r!,
\]

implying that, if \( h^* < \infty \), then \( \limsup_{t \to \infty} \Phi_t(t) = \infty \) for all \( r \geq 1 \). \( \square \)

The familiar ratio test yields simpler sufficient conditions. Thus \( \sup_t \Phi_t(t) < \infty \) for all \( r \geq 1 \) if

\[
\limsup_{j \to \infty} p_{j+1}/p_j < 1,
\]

while \( \lim_{t \to \infty} \Phi_t(t) = \infty \) for all \( r \geq 1 \) if

\[
\lim_{j \to \infty} p_{j+1}/p_j = 1.
\]

For instance, for \( p_j = cq^j \), the geometric distribution with \( 0 < q < 1 \), we have \( p_{j+1}/p_j = q \); hence \( \sup_r \Phi_t(t) < \infty \) for all \( r \), and normal approximation is not adequate for any \( r \). This illustrates possibility 4 in Theorem 4.1. For the Poisson distribution \( p_j = c\lambda^j/j! \), we even have \( p_{j+1}/p_j \to 0 \), and so normal approximation is no good here, either.

Continuing this line, we obtain a further set of conditions.

373
Lemma 4.3. (a) Suppose for some $0 < \lambda < 1$

$$\liminf_{j \to \infty} \frac{p_{j+h}}{p_j} > \lambda \tag{4.11}$$

for every $h \geq 1$. Then $\Phi_r(t) \to \infty$ as $t \to \infty$ for all $r \geq 1$.

(b) The condition $\limsup_{t \to \infty} \Phi_r(t) < \infty$ holds for some (hence for all) $r \geq 1$ if and only if there exists $h \geq 1$ such that

$$\limsup_{j \to \infty} \frac{p_{j+h}}{p_j} \leq \frac{1}{2}. \tag{4.12}$$

Proof. For part (a), assume that $v_0(\lambda x, x) = \# \{ j : \lambda x < p_j < x \} \to \infty$ as $x \to 0$. Then also

$$\Phi_r(1/x) \geq \sum_{\{ j : \lambda x < p_j < x \}} e^{-p_j/x} (p_j/x)^r / r! \geq v_0(\lambda x, x) \min_{\{ y : \lambda < y < 1 \}} [e^{-y} y^r / r!] \to \infty.$$ 

As $x$ decreases, the piecewise-constant function $v_0(\lambda x, x)$ may have downward jumps only at the values $x \in \{ p_j \}$, hence the assumption is equivalent to $v_0(\lambda p_j, p_j) \to \infty$ (as $j \to \infty$), which in turn is readily translated into (4.11).

For part (b), the same estimate with any $0 < \lambda < 1/2$ shows that the condition (4.12) is necessary. In the other direction, suppose that $p_{j+h}/p_j < 3/4$ for all $j \geq 1$. Split $(p_j, j \geq 1)$ into $h$ subsequences $(p_{j+i h}, i \geq 0)$, with $0 \leq s \leq h - 1$. Each of the subsequences has the property that the ratio of any two consecutive elements is at most $3/4$. Hence, as above, the sum of the terms $e^{-p_j}(t p_j)^r / r!$ along a subsequence yields a uniformly bounded contribution to $\Phi_r$.

Examples of irregular behaviour of moments may be constructed by breaking the sequence $(p_j, j \geq 1)$ into finite blocks of sizes $m_1, m_2, \ldots$, and setting the $p_j$'s within the $i$'th block all equal to some $q_i$. We use the notation $V(t) := \text{Var} \left( \sum_{r \geq 1} Y_r(t) \right)$ to denote the variance of the number of occupied boxes.

Example 1. [8, p. 384]. Take $m_i = i$ and $q_i = c 2^{-i}$, with $c$ a normalizing factor to achieve $\sum_i p_j = 1$. Then both $V(t)$ and $\Phi_1(t)$ oscillate between 0 and $\infty$, approaching the extremes arbitrarily closely. This illustrates possibility 3 in Theorem 4.1.

Example 2. As in [3, Example 4.4], take $m_i = 2^i$, $q_i = c 2^{-2^{i+1}}$. Then $\Phi_1(t) \to \infty$, but $\Phi_2(t)$ oscillates between 0 and $\infty$ as $t$ varies; thus $Y_1(t)$ is asymptotically normal, but $Y_2(t)$ is not, and the ratios $p_{j+1}/p_j$ have accumulation points at 0 and 1. This illustrates possibility 2 in Theorem 4.1.

We now extend this example, showing among other things that one can have any value for $r_0$ in behaviour 2 in Theorem 4.1.

Proposition 4.4. Fix $0 < \beta < 1$ and $\alpha > 0$, and take the blocks construction with $m_i = [2^{(1-\beta)^i}]$, $q_i = c m_i^{-(1+\alpha)}$, where $c$ is the appropriate normalizing constant. Then we have

---

In fact, the Poisson sampling model makes sense for arbitrary $p_j$'s, and the enumeration of small counts makes sense if $\sum_i p_j < \infty$.  

374
(i) \( \limsup_{t \to \infty} V_r(t) = \infty \) for all \( r \geq 1 \);

(ii) \( \lim_{t \to \infty} V_r(t) = \infty \) if and only if \( r \beta(1 + \alpha) \leq 1 \);

(iii) \( \lim_{j \to \infty} \rho_{r,j} = \infty \) if and only if \( r \beta(1 + \alpha) < 1 \);

(iv) The quantities \( \Sigma_{r,s}(t) \) do not converge for any \( r \neq s \).

**Proof.** Once again, we work with \( \Phi_r \) instead of \( V_r \), now writing

\[
\Phi_r(t) = \sum_{i \geq 1} m_i e^{-tq_i}(tq_i)^r/r!.
\]  

(4.13)

For part (i), it is enough to consider the subsequence \( t_i := 1/q_i, l \geq 1 \).

For part (ii), split \( \mathbb{R}_+ \) into intervals \( J_l := [q_l^{-1}, q_{l+1}^{-1}), l \geq 1 \); we show that \( \lim_{t \to \infty} \inf_{t \in J_l} \Phi_r(t) = \infty \) if \( r \beta(1 + \alpha) \leq 1 \), and exhibit a subsequence \( (t_l', l \geq 1) \) with \( t_l' \in J_l \) such that \( \lim_{t \to \infty} \Phi_r(t_l') = 0 \) if \( r \beta(1 + \alpha) > 1 \). Indeed, for \( t \in J_l \), taking just the term with \( i = l + 1 \) in (4.13), we obtain

\[
m_{l+1} \exp\{-\phi q_{l+1}/q_l\}(\phi q_{l+1}/q_l)^r/r! \approx m_{l+1} \phi^r \left( \frac{m_{l+1}^{(1-\beta)(1+\alpha)}}{m_{l+1}^{(1+\alpha)}} \right)^r = \phi^r m_{l+1}^{1-r\beta(1+\alpha)},
\]

where we write \( t = \phi/q_l \) with \( 1 \leq \phi \leq q_l/q_{l+1} \sim m_{l+1}^{\beta(1+\alpha)} \), and use the fact that \( \phi q_{l+1}/q_l \leq 1 \) in this range. The notation \( a \sim b \) means that \( \epsilon < a(t)/b(t) < 1/\epsilon \) for some \( \epsilon > 0 \) and all sufficiently large \( t \). For \( r \beta(1 + \alpha) < 1 \), it follows that \( \inf_{t \in J_l} \Phi_r(t) \sim m_{l+1}^{1-r\beta(1+\alpha)} \to \infty \) as \( l \to \infty \).

For \( r \beta(1 + \alpha) = 1 \), take also the term with \( i = l \) in (4.13), giving a combined contribution of at least

\[
m_l e^{-\phi} \frac{\phi^r}{r!} + K \phi^r,
\]

for some \( K > 0 \). It is easily checked that the minimum value of this sum for \( \phi > 1 \) goes to \( \infty \) with \( l \), hence, once again, \( \lim_{t \to \infty} \inf_{t \in J_l} \Phi_r(t) = \infty \).

For \( r \beta(1 + \alpha) > 1 \), these two terms contribute an amount of order

\[
\phi^r \{ m_{l+1}^{(1-\beta)} e^{-\phi} + m_{l+1}^{1-r\beta(1+\alpha)} \},
\]  

(4.14)

to (4.13), which is small as \( l \to \infty \), for example, for \( \phi = 2 \log m_{l+1} \). The sum of the terms in (4.13) for \( i \geq l + 2 \) is of order

\[
\sum_{i \geq l+2} m_i \left( \frac{\phi q_i}{q_l} \right)^r \sim \phi^r \sum_{i \geq l+2} m_i \{ m_i^{(1-\beta)(1-\gamma)} \}^{r(1+\alpha)} = \phi^r O(m_{l+1}^{1-r\beta(1+\alpha)-\eta}),
\]

where \( \eta > 0 \), and hence asymptotically smaller than the second element of (4.14). The sum of the terms in (4.13) for \( i \leq l - 1 \) is of order at most

\[
\left\{ \sum_{i=1}^{l-1} m_i \right\} \exp\{-\phi/q_i\} \left( \frac{\phi q_{l-1}}{q_l} \right)^r,
\]

largest for \( \phi = 1 \) for all \( l \) large enough, when it is of order

\[
m_{l-1}^{1+r\beta(1+\alpha)/(1-\beta)} \exp\{-m_{l-1}^{\beta(1+\alpha)/(1-\beta)} \},
\]

375
asymptotically small as \( l \to \infty \). Hence, for \( t'_i = 2q_i^{-1} \log m_{i+1} \), it follows that \( \lim_{t \to \infty} \Phi_r(t'_i) = 0 \), and therefore that \( \Phi_r(t) \) does not converge to infinity as \( t \to \infty \).

For part (iii), writing \( M_i := \sum_{t=1}^i m_t \), we have

\[
\rho_{r,j} \geq q_i^{-r} \sum_{l \geq i+1} m_l q_l^r \quad \text{whenever} \quad M_{i-1} < j \leq M_i,
\]

with equality for \( j = M_i \). Now

\[
\sum_{l \geq i+1} m_l q_l^r = m_{i+1}^{1-r(1+\alpha)},
\]

and

\[
q_i^{-r} = m_i^{r(1+\alpha)} \sim m_{i+1}^{r(1-\beta)(1+\alpha)}.
\]

Hence \( \rho_{r,M_l} \leq m_{i+1}^{1-r\beta(1+\alpha)} \) is bounded for \( r\beta(1+\alpha) \geq 1 \), and \( \rho_{r,j} \to \infty \) as \( j \to \infty \) if \( r\beta(1+\alpha) < 1 \).

For part (iv), we note that, for \( t = \phi/q_i \), the quantity

\[
\Sigma_{rs}(t) = -2^{-r-s} \binom{r+s}{r} \frac{\Phi_{r+s}(2t)}{\sqrt{V_r(t)V_s(t)}}, \quad r \neq s,
\]

behaves asymptotically, as \( l \) becomes large, in the same way as for the Poisson occupancy scheme with a single block of \( m_l \) boxes with equal frequencies \( q_i \). Computing the limit,

\[
\lim_{l \to \infty} \Sigma_{rs}(\phi/q_i) = -\frac{1}{\sqrt{r!s!}} \frac{e^{-\Phi_{r+s}^{(r+s)/2}}}{\sqrt{\{1 - e^{-\Phi/r!2^r}\}\{1 - e^{-\Phi/s!2^s}\}}},
\]

where \( m_l \) cancels because of the additivity of the moments. As \( \phi \) varies, this limit value varies too, and hence, for \( r \neq s \), the quantities \( \Sigma_{rs}(t) \) do not converge as \( t \to \infty \). \( \square \)

It follows from parts (ii) and (iii) of Proposition 4.4 that the implication in part (b) of Lemma 4.2 cannot be reversed, and from part (iv) that the correlations between different components of \( Y(t) \) need not converge, even when their variances tend to infinity. Hence the approximation in Theorem 3.1 does not necessarily imply convergence. Yet another kind of pathology appears when \( Y_I(t) \) is asymptotically independent of \( (Y_r(t), r > 1) \), as in the following example.

**Example 3.** Suppose that the frequencies in the block construction satisfy \( q_i = 1/i! \), \( m_i = (i-2)! \) (with \( i \geq 2 \)). Since \( q_i^{-r} \sum_{k=i+1}^\infty m_k q_k^r \to \infty \) for each \( r \), we have \( \lim_{t \to \infty} \rho_{j,r} = \infty \), and hence all the variances \( V_r(t) \) go to \( \infty \) by Lemma 4.2(b). On the other hand, \( m_i q_i / \sum_{k=i+1}^\infty m_k q_k \to 0 \), and it follows that

\[
\frac{\Phi_{1+s}(2t)}{\Phi_1(t)} = \frac{2^{s+1} \sum_{k=i+1} m_k q_k e^{-tq_i} \{e^{-tq_i} t^s q_i^s\}}{(s+1)! \sum_{k=i+1} m_k q_k e^{-tq_i}} \to 0
\]

as \( t \to \infty \). Since \( \Phi_{1+s}(2t)/\Phi_1(t) \) is bounded above by (4.4), we conclude that \( \Sigma_{1,s}(t) \to 0 \) for \( s \geq 2 \). It follows that every pair \( (Y_1(t), Y_2(t)), s \geq 2 \), converges in distribution to the standard bivariate normal distribution with independent components. Because the variances go to \( \infty \), Theorem 3.1 guarantees increasing quality of the normal approximation for any finite collection of components \( Y_r(t) \). However, the full vector \( (Y_r, r = 1, 2, \ldots) \) does not converge: see more on this example in Sections 5 and 6.
Part (ii) of Proposition 4.4 also demonstrates that \( \liminf_{j \to \infty} p_{j+1}/p_j = 0 \) does not exclude that \( \Phi_r(t) \to \infty \), hence the condition (4.11) in Lemma 4.3 is not necessary. Finally, by [3, Eqn. 3.1], we have
\[
\frac{1}{2} \Phi_1(2t) < V(t) < \Phi_1(t),
\]
meaning that \( \Phi_1(t) \) is always of the same order as the variance of the number of occupied boxes \( V(t) \). The examples above show that this need not be the case for \( \Phi_r(t) \), when \( r \geq 2 \).

5 Regular variation

We now henceforth assume that \( \Phi_r(t) \to \infty \) for all \( r \geq 1 \). The CLT for each component of \( Y(t) \) then holds, as observed above, and normal approximation becomes progressively more accurate for the joint distribution of any finite collection of components. A joint normal limit for any collection of the standardized components also holds, provided that the corresponding covariances converge. From (4.3) we have
\[
\text{Cov}(Y_r'(t), Y_s'(t)) = \Sigma_{rs}(t) = c(r,s) \frac{\Phi_{r+s}(2t)}{\sqrt{V_r(t)V_s(t)}}, \quad r \neq s.
\]
(5.1)
The RHS converges to a nonzero limit for each pair \( r,s \) if, for each \( r, \Phi_r \approx f \in R_\alpha \), where \( R_\alpha \) denotes the class of functions regularly varying at \( \infty \) with index \( \alpha \), and where, here and subsequently, we write \( a \approx b \) if \( a(t)/b(t) \to c \) as \( t \to \infty \) with \( 0 < c < \infty \). If \( \Phi_r \in R_\alpha \), then the index belongs to the range \( 0 \leq \alpha \leq 1 \), because \( \Phi_r(t) \) cannot converge to 0, and because \( \Phi_r(t)/t \to 0 \).

The results in the next section show that, if the covariances converge for a sufficiently large set of pairs \( r,s \), then this is in fact the only possibility. More formally, we say that then regular variation holds in the occupancy problem, meaning that, for some \( 0 \leq \alpha \leq 1 \) and some rate function \( f \in R_\alpha \),
\[
\Phi_r \approx f \quad \text{for all } r \geq 2.
\]
(5.2)
This setting of regular variation extends the original approach by Karlin [8] in the special case \( \alpha = 0 \), and, moreover, it covers all possible limiting covariance structures (Theorem 6.4).

Observe that the functions \( t^{-r} \Phi_r \) satisfy
\[
\frac{d^r}{dt^r} \{ t^{-1} \Phi_1(t) \} = (-1)^r r! \{ t^{-r} \Phi_r(t) \},
\]
(5.3)
thus, in particular, they are completely monotone. This taken together with the standard properties of regularly varying functions [2] implies that, if \( \Phi_r \in R_\alpha \) for some \( 0 \leq \alpha < 1 \) and \( r \geq 1 \), then the same is true for all \( r \geq 1 \), and we can choose the rate function \( f = \Phi_1 \). The case \( \alpha = 1 \) is special. If \( \Phi_r \in R_1 \) for some \( r \geq 2 \), then all \( \Phi_r \) for \( r \geq 2 \) are of the same order of growth and \( \Phi_1 \in R_1 \), but \( \Phi_1 \gg \Phi_2 \) (this motivates the choice \( r \geq 2 \) in (5.2)).

A necessary condition for (5.2) is \( \lim_{j \to \infty} p_{j+1}/p_j = 1 \), as follows from the next lemma.

**Lemma 5.1.** If \( \liminf_{j \to \infty} p_{j+1}/p_j < 1 \) then \( \Phi_r \) is not regularly varying for \( r \geq 2 \), and \( \Phi_1 \) is not regularly varying with index \( \alpha < 1 \).
Proof. We have

\[ t^{-2} \Phi_2(t) = \sum_{j=1}^{\infty} e^{-t p_j \beta} = \int_0^1 e^{-tx} v_2(dx) \]

with \( v_2[0, x] := \sum_{j=1}^{\infty} p_j^2 1[p_j \leq x] \). Suppose \( t^{-2} \Phi_t \in R_{-\beta} \), then \( 1 \leq \beta \leq 2 \) and, by Karamata’s Tauberian theorem, also \( v_2[0, t^{-1}] \in R_{-\beta} \). Because \( \beta \neq 0 \), the latter implies that \( v_2[at^{-1}, bt^{-1}] \in R_{-\beta} \), i.e. that

\[ v_2[at^{-1}, bt^{-1}] \sim (b^\beta - a^\beta) \ell(t) t^{-\beta}, \quad t \to \infty \quad (5.4) \]

for any positive \( a < b \). However, the assumption of the lemma allows to choose \( a < b < 1 \) such that \( v_2[ap_j, bp_j] = 0 \) for infinitely many \( j = j_k \), so \( (5.4) \) fails for \( t = 1/p_{j_k} \to \infty \). The contradiction shows that \( t^{-2} \Phi_2(t) \) cannot be regularly varying. The assertions regarding \( r \neq 2 \) can be derived in the same way. \( \square \)

The example below shows that \( \Phi_r \) may be regularly varying for \( r = 1 \) alone.

Example 3 (continued). Let \( g(t) = v_1[0, t^{-1}] = \sum_{j=1}^{\infty} p_j 1[p_j \leq t^{-1}] \). We have the general estimates

\[ t^{-1} \Phi_1(t) \geq e^{-1} g(t) \]

and, for \( a > 1 \) and any \( \epsilon > 0 \),

\[
t^{-1} \Phi_1(t) - (at)^{-1} \Phi_1(at) \\
\leq \epsilon g(at/\epsilon) + \{g(t/\log\{1/\epsilon g(t)\}) - g(at/\epsilon)\} + \sum_{j=1}^{\infty} p^j e^{-t p_j} 1[p_j > t^{-1} \log\{1/\epsilon g(t)\}] \\
\leq 2 \epsilon g(t) + \{g(t/\log\{1/\epsilon g(t)\}) - g(at/\epsilon)\}.
\]

Applying these to the block construction with \( q_i = 1/i! \) and \( m_i = (i - 2)! \), we observe that \( g(t) \asymp I(t^{-1}) \) and that \( g(t/\log\{1/\epsilon g(t)\}) - g(at/\epsilon) \) involves at most two \( q_i \), each of the corresponding terms being of the order of \( I(t)^{-2} \), where \( I(t) := \min\{i : i! \geq t\} \). It follows that \( t^{-1} \Phi_1(t) \in R_0 \), whence \( \Phi_1 \in R_1 \) and \( \Phi_1 \gg \Phi_r \) for \( r \geq 2 \). However, \( q_{i+1}/q_i \to 0 \), therefore Lemma \[5.1\] implies that \( \Phi_r \notin R_1 \) for \( r \geq 2 \).

The proper case of regular variation with index \( 0 < \alpha < 1 \) can be characterized by Karlin’s condition \[8\] Equation 5]

\[ v_0[0, x] := \#\{j : p_j \geq x\} \sim \ell(1/x)x^{-\alpha}, \quad x \downarrow 0, \quad (5.5) \]

where and henceforth the symbol \( \ell \) stands for a function of slow variation at \( \infty \). Other equivalent conditions are (see \[3\])

\[
\Phi(t) := \int_0^1 (1 - e^{-tx}) v_0(dx) \sim \Gamma(1 - \alpha) t^\alpha \ell(t), \\
v_r[0, x] \sim \frac{\alpha}{r - \alpha} x^{r-\alpha} \ell(1/x) \text{ for some } r \geq 1, \\
\Phi_r(t) \sim \frac{\alpha r!}{\Gamma(r - \alpha)} t^\alpha \ell(t) \text{ for some } r \geq 1, \\
p_j \sim \ell^*(j) t^{-1/\alpha},
\]

\[378\]
where \( \ell^*(y) = 1/\{\ell^1/y^{1/2}\} \), and \( # \) denotes the de Bruijn conjugate of a slowly varying function \([2]\). Note that \( V_r(t) \) then has the same order of growth, in view of \((4.2)\), yielding behaviour as in possibility 1 of Theorem 4.1.

We describe next all possible forms of multivariate convergence under the regular variation.

**The proper case** \( 0 < \alpha < 1 \). The joint CLT for
\[
\frac{Y_r(t) - \Phi_r(t)}{\sqrt{t^\alpha \ell(t)}}, \quad r = 1, 2, \ldots
\]
in \( \mathbb{R}^\infty \) holds with the limiting covariance matrix \( S \) computed from \((4.3)\) as
\[
S_{rs} = -\frac{a\Gamma(r + s - \alpha)}{r! s! \ell^{2s-\alpha}}, \quad r \neq s
\]
\[
S_{rr} = \frac{\alpha}{r!} \left( \Gamma(r - \alpha) - \Gamma(2r - \alpha) - \frac{r! \ell^{2r-\alpha}}{\ell^{2r-\alpha}} \right).
\]

This agrees with Karlin \([8\) Theorem 5\).

**The case of 'rapid' regular variation with** \( \alpha = 1 \). If \((5.5)\) holds with \( \alpha = 1 \) then \( \ell(t) \) must approach 0 as \( t \to \infty \) sufficiently fast to have \( \sum p_j < \infty \). In this situation we have \( \Phi_r(t) \sim (r^2 - r)^{-1} \ell(t) t \) for \( r > 1 \) but \( \Phi_1(t) \sim \ell_1(t) t \) with some \( \ell_1 \gg \ell \). In fact, \( X_n,1 \sim K_n \) as \( n \to \infty \) almost surely. Because the scaling of \( Y_1(t) \) is faster than that for other \( Y_r(t) \)’s, it follows from \((4.2)\) and \((4.3)\) that \( \Sigma_1 \ell(t) \to 0 \) for all \( r \geq 2 \), so that the CLT holds with \( Y_1(t) \) asymptotically independent of \( (Y_r(t), r \geq 2) \). The limiting covariance matrix of
\[
\frac{Y_r(t) - \Phi_r(t)}{\sqrt{t \ell(t)}}, \quad r \geq 2
\]
is obtained by setting \( \alpha = 1 \) in the above formulas \((5.6), (5.7)\) for \( S \). Our multivariate result extends in this case the marginal convergence that was stated in \([8\) Thm 5'\].

**The case of slow variation** \( \alpha = 0 \). Karlin’s condition \((5.5)\) with \( \alpha = 0 \) is too weak to control the \( \Phi_r(t) \)’s. We impose therefore a slightly stronger condition
\[
\nu_1[0,x] := \sum_{\{j:p_j \leq x\}} p_j \sim x \ell_1(1/x), \quad (5.8)
\]
which is equivalent to \( \Phi_r \in R_0 \) for any (and hence for all) \( r \geq 1 \). To illustrate the difference, note that in the geometric case, with \( p_j = (1 - q)^{q^{j-1}} \), \( 0 < q < 1 \), we have \( \ell(1/x) \sim \log q(1/x) \), whereas \( \nu_1[0,x] = \log [\log (1-x)] \) is not regularly varying, since \( \nu_1[0,x]/x \) jumps infinitely often from \( (1 - q)^{-1} \) to \( q(1 - q)^{-1} \) as \( x \to 0 \). The geometric case can be contrasted to the one with frequencies \( p_j = ce^{-j\beta} \), \( 0 < \beta < 1 \), for which we have \( \ell(1/x) \sim c |\log x|^{1/\beta} \) and \( \nu_1[0,x]/x \sim c |\log x|^{1/\beta - 1} \).

---

2 One example is \( p_j = c/j!\{\log (j + 1)^{\beta+1} \}, \beta > 0 \), in which case \( \ell(t) \sim 1/c(\log t)^{\beta+1} \).

3 Mikhailov \([12]\) indicated yet other situation where the \( X_n,1 \)’s for \( r > 1 \) all behave similarly, but their behaviour is distinct from that of \( X_n,1 \).
By \([6, \text{Prop. 15}]\), the general connection between \(\ell_1\) in (5.8) and \(\ell\) in (5.5) is

\[
\ell(1/x) = \int_x^1 u^{-1}\ell_1(1/u)\,du, \quad 0 < x < 1.
\]

Adopting (5.8) we have \(\nu_r[0,x] \sim r^{-1}x^r\ell_1(1/x)\), \(r \geq 1\), and the situation is then very similar to that in the proper case: we have \(\Phi_r(t) \sim r^{-1}\ell_1(t)\) and

\[
\frac{Y_r(t) - \Phi_r(t)}{\sqrt{\ell_1(t)}}, \quad r \geq 1
\]

jointly converge in law to a multivariate Gaussian limit with covariance matrix \(S\) given by

\[
S_{rs} = -\frac{1}{(r+s)2^{r+s}} \binom{r+s}{r}, \quad r \neq s, \quad (5.9)
\]

\[
S_{rr} = \frac{1}{r - \frac{1}{r^{2r+1}}} \left(\frac{2r}{r}\right). \quad (5.10)
\]

This applies, for instance, to the frequencies \(p_j \sim c e^{-j\beta}\) \((0 < \beta < 1)\). This case of slow variation seems not to have been considered before.

### 6 Convergence of the covariances

We will show in this section that regular variation is essential for the multivariate convergence of the whole standardized vector of counts, so that all possible limit covariance structures are those characterized in the previous section. Our starting point is the following lemma, which asserts that the regular variation is forced by the convergence of the ratios of \(\Phi_r\)'s.

**Lemma 6.1.** Suppose for some \(r \geq 1\)

\[
\lim_{t \to \infty} \Phi_{r+1}(t)/\Phi_r(t) = c. \quad (6.1)
\]

Then \((r - 1)/(r + 1) \leq c \leq r/(r + 1)\) and \(\Phi_r \in R_{\alpha}\) with \(\alpha := r - c(r + 1)\). Moreover, we then always have

\[
\lim_{t \to \infty} \frac{\Phi_s(t)}{\Phi_r(t)} = \frac{r!\Gamma(s - \alpha)}{s!\Gamma(r - \alpha)} \quad (6.2)
\]

and \(\Phi_s \in R_{\alpha}\) for all \(s \geq 1\), unless \(\alpha = 1\). If (6.1) holds with \(r > 1\) and \(c = (r - 1)/(r + 1)\), then \(\Phi_s \in R_1\) for \(s \geq 2\), and (6.2) is still true (in particular, \(\Phi_1 \gg \Phi_2\)).

**Proof.** A monotone density result which dates back to von Mises and Lamperti \([9]\) says that the convergence \(tg'(t)/g(t) \to \beta\) implies \(g \in R_{\beta}\) (this holds for arbitrary \(\beta\), including \(\pm \infty\)). This result applied to \(g(t) = t^{-r}\Phi_r(t)\) yields the regular variation \(\Phi_r \in R_{\alpha}\), with some \(0 \leq \alpha \leq 1\). The rest follows from (5.3), monotonicity and the general behaviour of the regularly varying functions under integration and differentiation \([2]\).

To apply the lemma, we need to pass from the convergence of covariances (5.1) to the convergence of a ratio as in (6.1). To this end, it is useful to exclude zero limits.
Lemma 6.2. If \( \lim \sup_t \Phi_s(t) = \infty \) for any \( s \geq 1 \), then no correlation \( \Sigma_{r,r'}(t) \) with \( 2 \leq r < r' \) can converge to zero.

Proof. (i) Let \( m_j := \# \{ l : 2^{-(j+1)} < p_l \leq 2^{-j} \} \). Then, if \( m^* := \sup_j m_j < \infty \), it follows that, for \( 2^j \leq t < 2^{j+1} \),

\[
\sum_{k \geq 0} \sum_{\{l: 2^{-(k+1)} < p_l \leq 2^{-k} \}} (tp_l)^s e^{-tp_l} \leq \sum_{k \geq 0} m_k 2^{(j+1-k)\lambda} \exp\{-2^{j-k-1}\}
\]

\[
\leq m^* \left( \sum_{k \geq j+1} 2^{(j+1-k)} + 2^j \sum_{k=0}^j 2^{(j-k)} \exp\{-2^{j-k-1}\} \right)
\]

\[
\leq m^* \left( 2 + 2^j \sum_{l \geq 0} 2^l \exp\{-2^{l-1}\} \right) = m^* c_j < \infty,
\]

uniformly in \( j \), which contradicts \( \lim \sup_t \Phi_s(t) = \infty \). Hence \( \sup_j m_j = \infty \).

(ii) Given any \( j_0 \), there exists some \( j \geq j_0 \) such that

\[
m_k \leq m_j, \quad 0 \leq k \leq j; \quad m_k \leq 3^{k-j} m_j, \quad k \geq j. \tag{6.3}\]

To see this, first take \( j_1 \geq j_0 \) such that \( m_{j_1} = \max \{ m_k, 0 \leq k \leq j_1 \} \), as can always be done, since \( \sup_j m_j = \infty \). Then let \( j_2 := \max \{ k \geq j_1 : m_k \geq 3^{k-j_1} m_{j_1} \} \); this is finite, since \( 1 \geq \sum_{l \geq 1} p_l \geq m_{j_1} 2^{-(j_1+1)} \) for each \( j \geq 0 \). Finally, take \( j_3 = \arg \max_{j_1 \leq j \leq j_2} m_j \); then \( j_3 \) satisfies the requirements of (6.3).

(iii) Now suppose that \( j \) satisfies (6.3). Then, much as in part (i), for any \( r \geq 2 \),

\[
r! \Phi_r(2^j) \leq \sum_{k \geq 0} m_k 2^{(j-k)r} \exp\{-2^{j-k-1}\}
\]

\[
\leq \left( \sum_{k \geq j+1} m_j 3^{k-j} 2^{(j-k)} + m_j \sum_{k=0}^j 2^{(j-k)} \exp\{-2^{j-k-1}\} \right)
\]

\[
\leq m_j \left( 3 + \sum_{l \geq 0} 2^l \exp\{-2^{l-1}\} \right) = c'_r m_j,
\]

with \( c'_r < \infty \), whereas also, just from the indices \( l \) with \( 2^{-(j+1)} < p_l \leq 2^{-j} \), we have

\[
r! \Phi_r(2^{j+1}) \geq m_j e^{-2}.
\]

This implies that

\[
\Phi_{r+r'}(2t)/\sqrt{\Phi_r(t)\Phi_r(t)} \geq \frac{e^{-2}/\{r + r'\}!}{\sqrt{c'_r c'_r/r!r'!}} > 0
\]

for \( t = 2^j \), whenever \( j \) satisfies the requirements of (6.3), and there are infinitely many such. Hence the correlations \( \Sigma_{r,r'}(t) \) with \( r' > r \geq 2 \) cannot converge to zero. 

\[\square\]
Note that the correlations $\Sigma_{1,s}(t), s > 1$, converge to zero in the case of regular variation with index $\alpha = 1$. Example 3 illustrates that $\Sigma_{1,s}(t)$ may also converge to zero when regular variation in the sense of (5.2) does not hold.

**Lemma 6.3.** If $g$ is continuous and positive, and $g(2t)/\sqrt{g(t)} \to k$ as $t \to \infty$, with $0 < k < \infty$, then $g(t) \to k^2$.

**Proof.** Given $\epsilon > 0$, let $t_\epsilon$ be such that $g(2t) \leq k\sqrt{(1 + \epsilon)g(t)}$ for all $t \geq t_\epsilon$. Let $K_\epsilon := \sup_{t \in J_\epsilon} g(t)$, where $J_\epsilon := [t_\epsilon, 2t_\epsilon]$. Then, for all $t \in J_\epsilon$ and all $n \geq 0$, we have

$$g(2^n t) \leq \{k^2(1 + \epsilon)\}^{1-2^{-n}} \{g(t)\}^{2^{-n}} \leq k^2(1 + \epsilon)k^{-2^n}.$$

Thus $\limsup_t g(t) \leq k^2$. A similar argument shows that $\liminf_t g(t) \geq k^2$, proving the lemma. □

**Theorem 6.4.** Suppose the correlations $\Sigma_{r,s}(t)$ converge, as $t \to \infty$, for $r, s$ satisfying $2 \leq r < s$ and $r + s \leq 12$. Then the following is true:

(i) (5.2) holds with some $0 \leq \alpha \leq 1$,

(ii) the correlations $\Sigma_{r,s}(t)$ converge for all $r, s$,

(iii) $(Y'_1(t), r = 1, 2, \ldots)$ converges weakly to one of the multivariate normal laws described in Section 5 for the cases $0 < \alpha < 1$, $\alpha = 1$ and $\alpha = 0$, with the obvious adjustment of (5.6), (5.7) and (5.9), (5.10) in order to have standard normal marginals,

(iv) the same multivariate normal limit holds for the normalized and centred $X_n$.

**Proof.** For short, write $V_j = V_j(t), f_j = \Phi_j(t)$ and $F_j = \Phi_j(2t)$. By Lemma 6.2, the $\Sigma_{r,s}(t)$ converge to nonzero limits, whence, for $r, s$ in the required range,

$$\frac{F_{r+s}}{\sqrt{V_r V_s}} \approx \frac{F_{r+s}}{\sqrt{V_{r+1} V_{s-1}}}$$

and hence $V_r V_s \approx V_{r+1} V_{s-1}$. From this, $V_5 \approx V_3 V_4/V_2, V_6 \approx V_3 V_5/V_2 \approx V_3/V_2, V_6 \approx V_8/V_2, V_7 \approx V_5/V_2, V_8 \approx V_7/V_2$, and substituting in $V_2 V_6 \approx V_5 V_5$ we get $V_4/V_2 \approx (V_3/V_2)^2$. Continuing in this way yields

$$\frac{V_j}{V_2} \approx \left(\frac{V_3}{V_2}\right)^{j-2} \text{ for } 2 \leq j \leq 10.$$  \hspace{1cm} (6.4)

From this and $F_j^2 \approx V_2 V_{j-2}$, we obtain

$$\frac{F_j}{V_2} \approx \left(\frac{V_3}{V_2}\right)^{j/2-2} \text{ for } 5 \leq j \leq 12.$$  \hspace{1cm} (6.5)

Substituting (6.4) and (6.5) in $f_j = V_j + c_j F_{2j}$ (recall (4.2)) yields

$$\frac{f_j}{V_2} \approx \left(\frac{V_3}{V_2}\right)^{j/2-2} \text{ for } 3 \leq j \leq 6.$$  \hspace{1cm} (6.6)
This offers two ways of expressing $F_j$ for $j = 5, 6$: using (6.5) or (6.6), but with the argument $2t$ for the latter. The first gives

$$F_5 \approx V_2(t) \left( \frac{V_3(t)}{V_2(t)} \right)^{1/2}, \quad F_6 \approx V_2(t) \left( \frac{V_3(t)}{V_2(t)} \right),$$

and the second gives

$$F_5 \approx V_2(2t) \left( \frac{V_3(2t)}{V_2(2t)} \right)^3, \quad F_6 \approx V_2(2t) \left( \frac{V_3(2t)}{V_2(2t)} \right)^4.$$

It follows that

$$\frac{F_6}{F_5} \approx \left( \frac{V_3(t)}{V_2(t)} \right)^{1/2} \approx \frac{V_3(2t)}{V_2(2t)}.$$

Applying Lemma [6.3] to $g(t) = V_3(t)/V_2(t)$ shows that this must converge, hence from (6.6) the ratio $\Phi_4(t)/\Phi_3(t)$ must converge too. Parts (i), (ii), (iii) of the theorem now follow from Lemma [6.1] and part (iv) follows by de-Poissonization. □

Combining Theorem 6.4 and Lemma 5.1 we arrive at a very simple test for the convergence, which is easy to check in the examples of Section 4:

**Corollary 6.5.** The condition $\lim_{j \to \infty} p_{j+1}/p_j = 1$ is necessary for the convergence of the (normalized and centred) $X_n$ to a multivariate normal law.

It should be stressed that the condition is by no means sufficient. For instance, the frequencies $p_j = c\{2 + \sin(\log j)\}/j^2$ satisfy $p_{j+1}/p_j \to 1$ but do not have the property of regular variation due to the oscillating sine factor. Thus in this case $X_n$ has no distributional limit.

**Acknowledgement** The authors would like to thank Adrian Röllin and Bero Roos for helpful discussions.

**References**


