The Annals of Probability
1980, Vol. 8, No. 1, 148–156

THREE LIMIT THEOREMS FOR SCORES BASED ON OCCUPANCY NUMBERS

BY M. P. QUINE
University of Cambridge

Let N balls be distributed independently and at random into n boxes. Let ρ_{nj} denote the number of balls in the jth box. Let $(c_0, c_1, c_2, \cdots)$ be a sequence of real numbers. Three limit theorems are proved for the sum $\Sigma_{j=1}^n c_{\rho_{nj}}$ as N and n tend to infinity in such a way that $N/n \to 0$.

0. Let N balls be distributed independently and at random into n boxes, in such a way that each ball has probability $1/n$ of landing in any given box. Denote by ρ_{nj} the number of balls in the jth box. Let $(c_0, c_1, c_2, \cdots)$ be a sequence of real numbers, and denote by $m = m(c_0, c_1, c_2, \cdots)$ the unique integer $(m \geqslant 2)$ such that

$$c_1 - c_0 = (c_2 - c_0)/2 = \cdots = (c_{m-1} - c_0)/(m-1) \neq (c_m - c_0)/m.$$

We assume that $m < \infty$, for in the case $m = \infty$, the quantity of interest in this paper, $\Sigma c_{\rho_{nj}} - nEc_{\rho_{n1}}$, vanishes. We will show that as N and n tend to infinity in such a way that $N/n \to 0$, convergence in distribution to a normal, Poisson, or degenerate law may occur. Theorems 1 and 2 generalize results of Békéssy (1963). We impose a condition on the sequence (c_i) in terms of

$$d_i = \max(|c_1 - c_0|, |c_2 - c_0|/2, \cdots, |c_i - c_0|/i), \quad i = 1, 2, \cdots.$$

1. Normal convergence. The result of this section generalizes a result of Békéssy (1963), who dealt with the case $c_i = \delta_{ki}$, $k \geqslant 0$. A discussion of this and other special cases may be found in Johnson and Kotz (1977). The analogue of the theorem in the case $N/n \to \alpha$, $0 < \alpha < \infty$, was first proved by Harris and Park (1971) although special cases were known much earlier (see Johnson and Kotz); Quine (1979) contains a discussion of extensions in this case. The case $N/n \to \infty$ has been dealt with (Békéssy (1963)) when $c_i = \delta_{ki}$, but there seem to be no general results available.

THEOREM 1. *Let $n, N \to \infty$ and $N/n \to 0$. If $\Sigma d_i^2 i^{2m+2}/i! < \infty$, and $N^m/n^{m-1} \to \infty$, then*

$$(N^m/n^{m-1})^{-\frac{1}{2}} \Sigma_{j=1}^n (c_{\rho_{nj}} - Ec_{\rho_{nj}}) \to_{\mathfrak{D}} N(0, \sigma^2),$$

where

$$\sigma^2 = (m(c_1 - c_0) - (c_m - c_0))^2/m!.$$

Received April 18, 1978.
AMS 1970 *subject classifications.* Primary 60F05.
Key words and phrases. Limit theorems, normal, Poisson, degenerate, occupancy numbers.

The mean $Ec_{\rho_{nj}}$ may be replaced by Ec_{τ_n}, where τ_n is a Poisson random variable (rv) *with mean N/n.*

PROOF. We consider the boxes in groups of size $k_n = [n/N]$, $[x]$ denoting the largest integer $\leqslant x$. There are $l_n = [n/k_n]$ ($\geqslant N$) groups of this size. Put

$$\beta_{nj} = \Sigma_{i=(j-1)k_n+1}^{jk_n} c_{\rho_{ni}}, \qquad \rho_{nj}^* = \Sigma_{i=(j-1)k_n+1}^{jk_n} \rho_{ni}, \qquad 1 \leqslant j \leqslant l_n;$$

$$\mu_{ni} = E(\beta_{n1}|\rho_{n1}^* = i), \qquad \sigma_{ni}^2 = \mathrm{Var}(\beta_{n1}|\rho_{n1}^* = i), \qquad 0 \leqslant i \leqslant N.$$

Note $(\beta_{n1}, \cdots, \beta_{nl_n})$ are exchangeable rv's, as are $(c_{\rho_{n1}}, \cdots, c_{\rho_{nN}})$.

LEMMA 1. *As $n \to \infty$,*

$$\mu_{ni} = k_n c_0 + i(c_1 - c_0) + \binom{i}{m}(c_m - c_0 - m(c_1 - c_0))/k_n^{m-1} + 0(1/k_n^m),$$

$$\sigma_{ni}^2 = 0 \qquad i < m$$

$$= \binom{i}{m}(c_m - c_0 - m(c_1 - c_0))^2/k_n^{m-1} + 0(1/k_n^m) \qquad i \geqslant m.$$

PROOF. If we can prove the lemma when $c_0 = 0$, then the general case follows trivially by considering $c_i' = c_i - c_0$. Indeed the same remark applies to Theorem 1. So for the rest of this section, we will assume $c_0 = 0$. Then

$$\mu_{ni} = k_n E(c_{\rho_{n1}}|\rho_{n1}^* = i)$$

$$= k_n\left(c_1 \Sigma_{j=0}^{m-1} j\binom{i}{j} k_n^{-j}(1 - 1/k_n)^{i-j} + \Sigma_{j=m}^{i} c_j \binom{i}{j} k_n^{-j}(1 - 1/k_n)^{i-j}\right)$$

$$= k_n c_1 E(\rho_{n1} I(\rho_{n1} < m)|\rho_{n1}^* = i) + c_m\binom{i}{m}\Big/ k_n^{m-1} + 0(1/k_n^m);$$

we take $\binom{i}{m} = 0$ if $i < m$. Since

$$E(\rho_{n1} I(\rho_{n1} < m)|\rho_{n1}^* = i) = i/k_n - E(\rho_{n1} I(\rho_{n1} \geqslant m)|\rho_{n1}^* = i)$$

$$= i/k_n - m\binom{i}{m}\Big/ k_n^m + 0(1/k_n^{m+1}),$$

the first part of the lemma now follows.

Next we note

$$E(c_{\rho_{n1}} c_{\rho_{n2}}|\rho_{n1}^* = i) = c_1^2 E(\rho_{n1}\rho_{n2} I(\rho_{n1}, \rho_{n2} < m)|\rho_{n1}^* = i)$$

$$+2c_1 c_m (i - m)\binom{i}{m}\Big/ k_n^{m+1} + 0(1/k_n^{m+2}),$$

and

$$E(c_{\rho_{n1}}^2|\rho_{n1}^* = i) = c_1^2 E(\rho_{n1}^2 I(\rho_{n1} < m)|\rho_{n1}^* = i) + c_m^2\binom{i}{m}/k_n^m + 0(1/k_n^{m+1}).$$

If $i < m$, then $\beta_{n1} = ic_1$ and σ_{ni}^2 vanishes. If $i \geqslant m$,

$$\sigma_{ni}^2 = k_n\big(E\big(c_{\rho_{n1}}^2|\rho_{n1}^* = i\big) - E\big(c_{\rho_{n1}}c_{\rho_{n2}}|\rho_{n1}^* = i\big)\big)$$
$$+ k_n^2\big(E\big(c_{\rho_{n1}}c_{\rho_{n2}}|\rho_{n1}^* = i\big) - E^2\big(c_{\rho_{n1}}|\rho_{n1}^* = i\big)\big)$$
$$= c_1^2 \operatorname{Var}\big(\Sigma_{j=1}^{k_n}\rho_{nj}I(\rho_{nj} < m)|\rho_{n1}^* = i\big)$$
$$+ k_n^{-m+1}\binom{i}{m}\big(c_m^2 + 2c_1c_m(i-m) - 2c_1c_m i\big) + 0(k_n^{-m})$$

from the above results. Now

$$\operatorname{Var}\big(\Sigma_{j=1}^{k_n}\rho_{nj}I(\rho_{nj} < m)|\rho_{n1}^* = i\big) = \operatorname{Var}\big(\Sigma_{j=1}^{k_n}\rho_{nj}I(\rho_{nj} \geqslant m)|\rho_{n1}^* = i\big)$$

and it is easy to check that

$$E\big(\rho_{n1}^2I(\rho_{n1} \geqslant m)|\rho_{n1}^* = i\big) = m^2k_n^{-m}\binom{i}{m} + 0\big(k_n^{-m-1}\big),$$

$$E\big(\rho_{n1}\rho_{n2}I(\rho_{n1}, \rho_{n2} \geqslant m)|\rho_{n1}^* = i\big) = \frac{m^2 i!}{k_n^{2m}(m!)^2(i-2m)!} + 0\big(1/k_n^{2m+1}\big).$$

Combining these with earlier results gives

$$\operatorname{Var}\big(\Sigma_{j=1}^{k_n}\rho_{nj}I(\rho_{nj} < m)|\rho_{n1}^* = i\big) = m^2k_n^{-m+1}\binom{i}{m} + 0(1/k_n^m),$$

and the second part of the lemma follows.

REMARK. Examination of the proof of this lemma shows that the order term in the expansion of μ_{ni} is dominated by $2d_i i^{(m+1)}/k_n^m$.

We now define a random function

$$X_n(t) = l_n^{-\frac{1}{2}}\Sigma_{j=1}^{[l_nt]}\big(\beta_{nj} - E\beta_{nj}\big), \qquad 0 \leqslant t \leqslant 1.$$

If $(\delta_{i1}^{(n)}, \delta_{i2}^{(n)}, \cdots)$ are independent and identically distributed (i.i.d.) rv's and $P(\delta_{i1}^{(n)} \leqslant x) = P(\beta_{n1} \leqslant x|\rho_{n1}^* = i)$, then $E(\delta_{i1}^{(n)}) = \mu_{ni}$, $\operatorname{Var}(\delta_{i1}^{(n)}) = \sigma_{ni}^2$. For $0 \leqslant t \leqslant 1$, $0 \leqslant i \leqslant N$, write

$$X_n^{(i)}(t) = l_n^{-\frac{1}{2}}\Sigma_{j=1}^{[l_nt]}\big(\delta_{ij}^{(n)} - \mu_{ni}\big),$$
$$\Omega_n^{(i)}(t) = l_n^{-\frac{1}{2}}\Sigma_{j=1}^{[l_nt]}I(\rho_{nj}^* = i);$$

put

$$Y_n(t) = l_n^{-\frac{1}{2}}\big(\Sigma_{i=0}^N\mu_{ni}\Sigma_{j=1}^{[l_nt]}I(\rho_{nj}^* = i) - [l_nt]E\beta_{n1}\big).$$

Then the finite dimensional distributions of X_n coincide with those of

$$\Sigma_{i=0}^N X_n^{(i)} \circ \Omega_n^{(i)} + Y_n.$$

Let $\Rightarrow$ denote weak convergence in (D, d), i.e., in the space of right-continuous functions with left limits, endowed with the Skorokhod topology.

LEMMA 2. *If $i \geqslant m$, then*

$$\sigma_{ni}^{-1}X_n^{(i)} \Rightarrow W^{(i)},$$

where $W^{(i)}$ is a standard Brownian motion; $W^{(i)}$ and $W^{(j)}$ are independent for $i \neq j$.

PROOF. According to, e.g. McLeish (1974), we need only check

$$E\frac{(\delta_{i1}^{(n)}-\mu_{ni})^2}{\sigma_{ni}^2}I\left(\left(\frac{\delta_{i1}^{(n)}-\mu_{ni}}{l_n\sigma_{ni}^2}\right)^2>\varepsilon\right)\to 0$$

for each $\varepsilon > 0$. But $|\delta_{i1}^{(n)}| \leqslant i\Sigma_{j=1}^{i}|c_j|$, so this Lindeberg quantity vanishes for sufficiently large n so long as $l_n\sigma_{ni}^2 \to \infty$, which in view of Lemma 1 is equivalent to $n/k_n^m \to \infty$, i.e., $N^m/n^{m-1} \to \infty$.

LEMMA 3. *Write* $\Omega^{(i)}(t) = e^{-1}t/i!$. *Then* $\Omega_n^{(i)} \Rightarrow \Omega^{(i)}$.

PROOF. Take $0 = t_0 < t_1 < \cdots < t_q = 1$. Then

$$P\left(\max_{0\leqslant j\leqslant q}|\Omega_n^{(i)}(t_j)-\Omega^{(i)}(t_j)|>\varepsilon\right)\leqslant\Sigma_{j=0}^{q}P\left(|\Omega_n^{(i)}(t_j)-\Omega^{(i)}(t_j)|>\varepsilon\right),$$

and an argument based on Chebyshev's inequality shows that each member of this finite sum $\to 0$ (note $\mathrm{Cov}(I(\rho_{n1}^* = i), I(\rho_{n2}^* = i)) \to 0$). Thus the finite-dimensional distributions of $\Omega_n^{(i)}$ converge as required. If $t_1 \leqslant t \leqslant t_2$, then

$$E|R_n^{(i)}(t)-R_n^{(i)}(t_1)|\,|R_n^{(i)}(t_2)-R_n^{(i)}(t)|\leqslant E^{\frac{1}{2}}\left(R_n^{(i)}(t)-R_n^{(i)}(t_1)\right)^2 \times E^{\frac{1}{2}}\left(R_n^{(i)}(t_2)-R_n^{(i)}(t)\right)^2,$$

where $R_n^{(i)} = \Omega_n^{(i)} - \Omega^{(i)}$, and the right-hand side $\to 0$ since, for example,

$$E\left(R_n^{(i)}(t)-R_n^{(i)}(t_1)\right)^2 = l_n^{-2}E\left(\Sigma_{j=[l_nt_1]}^{[l_nt]}I(\rho_{nj}^*=i)-\frac{[l_nt]-[l_nt_1]}{i!e}\right)^2 \to 0.$$

The lemma now follows from, e.g. Billingsley (1968, page 128).

LEMMA 4. *As* $n \to \infty$,

$$k_n^{(m-1)/2}(X_n(1)-Y_n(1))\to_{\mathfrak{D}}N(0,\sigma^2).$$

PROOF. Lemmas 1, 2 and 3, and Theorem 4.2 of Serfozo (1973), imply

$$k_n^{(m-1)/2}\Sigma_{i=0}^{N}X_n^{(i)}\circ\Omega_n^{(i)}\Rightarrow\Sigma\sigma_iW^{(i)}\circ\Omega^{(i)},$$

where

$$\sigma_i^2=\binom{i}{m}(mc_1-c_m)^2$$

(note Serfozo (1973, Section 5) and Billingsley (1968, Section 17)). The lemma follows on taking $t = 1$.

LEMMA 5. *As* $n \to \infty$,

$$k_n^{(m-1)/2}X_n(1)\to_{\mathfrak{D}}N(0,\sigma^2).$$

PROOF. In view of Lemma 4, it suffices to show $k_n^{(m-1)/2}Y_n(1) \to_p 0$, or a fortiori,

(1) $$k_n^{m-1}N^{-1}\mathrm{Var}\left(\Sigma_{j=1}^{l_n}\mu_{n\rho_{nj}^*}\right) \to 0,$$

which we now prove. Write $a_n = Nk_n/n$; note $a_n \leq 1$. Routine analysis shows that if $i \leq N/2$,

$$\left|P(\rho_{n1}^* = i) - \frac{a_n^i e^{-a_n}}{i!}\right| \leq \frac{q(i)}{i!N},$$

where $q(i)$ is a quadratic in i, and for $i > N/2$,

$$P(\rho_{n1}^* = i) \leq a_n^i e^{-a_n}/i!$$

once $N > 4$, in which case, therefore,

$$\left|E\mu_{n\rho_{n1}^*}^2 - \Sigma_{i \leq N/2}\mu_{ni}^2 a_n^i \frac{e^{-a_n}}{i!}(1 + N^{-1}q(i))\right| \leq \Sigma_{i>N/2}\mu_{ni}^2 a_n^i \frac{e^{-a_n}}{i!}.$$

The result

(2) $$E\mu_{n\rho_{n1}^*}^2 = \Sigma\, \mu_{ni}^2 a_n^i e^{-a_n}/i! + 0(1/N)$$

now follows so long as $\Sigma i^2\mu_{ni}^2/i! = 0(1)$. Conditional on $(\rho_{n1}^* = i)$, β_{n1} is of the form $a_1c_1 + \cdots + a_ic_i$, where $\Sigma_{j=1}^i ja_j = i$. Thus

$$|\mu_{ni}| \leq \Sigma_{j=1}^i ja_jd_i = id_i,$$

so that (2) is true if $\Sigma i^4 d_i^2/i! < \infty$.

Similar arguments show that

$$E\mu_{n\rho_{n1}^*}\mu_{n\rho_{n2}^*} = \left(\Sigma\, \mu_{ni}a_n^i e^{-a_n}/i!\right)^2 + 0(1/N)$$

and

$$\mathrm{Cov}(\mu_{n\rho_{n1}^*}, \mu_{n\rho_{n2}^*}) = -N^{-1}\left(\Sigma\, \mu_{ni}a_n^i e^{-a_n}(i - a_n)/i!\right)^2 + 0(1/N^2).$$

Using these two equations together with (2), we obtain

$$\mathrm{Var}\Sigma_{j=1}^{l_n}\mu_{n\rho_{nj}^*} = l_n\left(\Sigma\, \mu_{ni}^2 a_n^i e^{-a_n}/i! - \left(\Sigma\, \mu_{ni}a_n e^{-a_n}/i!\right)^2\right)$$
$$- l_n^2N^{-1}\left(\Sigma\, \mu_{ni}a_n^i(i - a_n)e^{-a_n}/i!\right)^2 + 0(1).$$

According to the remark after Lemma 1, μ_{ni} is of the form

$$\mu_{ni} = ic_1 + k_n^{-m+1}\binom{i}{m}\beta + k_n^{-m}\omega_{ni},$$

with $\beta = c_m - mc_1$, $|\omega_{ni}| \leq 2d_i i^{(m+1)}$. Thus tedious calculations show that, subject to $\Sigma\omega_{ni}^2/i! = 0(1)$,

$$\mathrm{Var}\Sigma_{j=1}^{l_n}\mu_{n\rho_{nj}^*} = 0(N/k_n^m) + 0(1),$$

and since $\Sigma d_i^2 i^{2m+2}/i! < \infty$ by assumption, (1) and the lemma now follow.

LEMMA 6. *As* $n \to \infty$,

$$(N^m/n^{m-1})^{-\frac{1}{2}}\Sigma_{j=k_nl_n+1}^n\left(c_{\rho_{nj}} - Ec_{\rho_{nj}}\right) \to_p 0.$$

PROOF. It is easy to see that

$$\begin{aligned} E|c_{\rho_{n1}}| &\leqslant N|c_1|/n + \Sigma_{i=2}^N |c_i|(N/n)^i/i! \\ &\leqslant N|c_1|/n + (N/n)^2 \Sigma_{i=2}^N i d_i/i! \qquad \text{once } N \leqslant n \\ &= 0(N/n). \end{aligned}$$

Thus

$$E\left|\Sigma_{j=k_n l_n+1}^n \left(c_{\rho_{nj}} - Ec_{\rho_{nj}}\right)\right| < 2k_n E|c_{\rho_{n1}}| = 0(1),$$

and the lemma follows.

The main part of Theorem 1 follows easily from Lemmas 5 and 6. As for the mean, routine calculations based on

$$Ec_{\rho_{n1}} = \Sigma_{j=1}^N c_j \binom{N}{j} n^{-j}(1 - 1/n)^{N-j}$$

show that $Ec_{\rho_{n1}} = Ec_{\tau_n} + 0(1/n)$.

2. Poisson convergence. The result of this section appears, in case $c_i = \delta_{ki}$, in Békéssy (1963); more extensive work on this case has been done by V. F. Kolchin (see Johnson and Kotz (1977) for references).

THEOREM 2. *If $n, N \to \infty$, $N^m/n^{m-1} \to A$, $0 < A < \infty$, and $\Sigma d_i^2 i^{2m+2}/i! < \infty$, then*

$$\{c_m - c_0 - m(c_1 - c_0)\}^{-1}\{\Sigma_{j=1}^n (c_{\rho_{nj}} - c_0) - N(c_1 - c_0)\} \to_{\mathfrak{D}} P(A/m!),$$

where $P(\lambda)$ is a Poisson rv *with mean λ.*

PROOF. The proof follows roughly the lines of that of Theorem 1. We again assume without loss of generality that $c_0 = 0$. We now divide the boxes into groups of size $k_n = [n/BN]$, where $B \geqslant 1$ will be specified subsequently. The quantities l_n, β_{nj}, ρ_{nj}^*, μ_{ni}, σ_{ni}^2 are then defined in terms of this new k_n just as in the previous section. We observe that the statement and proof of Lemma 1 remain unchanged. Furthermore, if $\Omega_n^{(i)}$ is defined in terms of the new k_n, then Lemma 3 continues to hold with

$$\Omega^{(i)}(t) = te^{-1/B}B^{-i}/i!.$$

Write

$$\begin{aligned} X_n(t) &= \Sigma_{j=1}^{[l_n t]}(\beta_{nj} - E\beta_{nj}), \\ X_n^{(i)}(t) &= \Sigma_{j=1}^{[l_n t]}(\delta_{ij}^{(n)} - \mu_{ni}), \\ Y_n(t) &= \Sigma_{j=1}^{[l_n t]}(\mu_{n\rho_{nj}^*} - E\mu_{n\rho_{nj}^*}). \end{aligned}$$

Then, as in the previous section, X_n has the same finite dimensional distributions as $\Sigma_{i=0}^N X_n^{(i)} \circ \Omega_n^{(i)} + Y_n$.

LEMMA 7. *For* $i \geqslant m$,

$$(c_m - mc_1)^{-1} X_n^{(i)} \Rightarrow V^{(i)},$$

where $(V^{(i)}(t) + \lambda_i t,\ 0 \leqslant t \leqslant 1)$ *is a Poisson process*, i.e., *a process with independent increments and with*

$$P\big(V^{(i)}(t) + \lambda_i t = k\big) = (\lambda_i t)^k e^{-\lambda_i t}/k!,$$

$$\lambda_i = AB^m \binom{i}{m}.$$

PROOF. Since $X_n^{(i)}$ is composed of i.i.d. rv's which are asymptotically negligible, it is necessary only to show that $(c_m - mc_1)^{-1} X_n^{(i)}(1)$ converges in distribution to $P\Big(AB^m\binom{i}{m}\Big)$ (see Prohorov (1956, page 197)). According to Brown and Eagleson (1971), sufficient conditions for such convergence are that

$$l_n E(\delta_{i1}^{(n)} - \mu_{ni})^2 / (c_m - mc_1)^2 \rightarrow AB^m \binom{i}{m},$$

and that for each $\varepsilon > 0$,

$$l_n E(\delta_{i1}^{(n)} - \mu_{ni})^2 I\left(\left|\frac{\delta_{i1}^{(n)} - \mu_{ni}}{c_m - mc_1} - 1\right| > \varepsilon\right) \rightarrow 0. \tag{3}$$

The first of these conditions follows directly from Lemma 1. The indicator in (3) is bounded by

$$I(\delta_{i1}^{(n)} = ic_1) + I\left(\left|\frac{\delta_{i1}^{(n)} - \mu_{ni}}{c_m - mc_1} - 1\right| > \varepsilon,\ \delta_{i1}^{(n)} = c_m + (i - m)c_1\right)$$
$$+ I\big(\delta_{i1}^{(n)} \neq ic_1,\ \delta_{i1}^{(n)} \neq c_m + (i - m)c_1\big).$$

The second of these indicators vanishes for all sufficiently large values of n because of the asymptotic behaviour of μ_{ni}. Thus for all large n,

$$E(\delta_{i1}^{(n)} - \mu_{ni})^2 I\left(\left|\frac{\delta_{i1}^{(n)} - \mu_{ni}}{c_m - mc_1} - 1\right| > \varepsilon\right)$$

$$\leqslant 2\binom{i}{m}^2 (c_m - mc_1)^2 / k_n^{2m-2} + 4\big(i\Sigma_{j=1}^{i}|c_j|\big)^2 P(\delta_{i1}^{(n)} \neq ic_1, \neq c_m + (i - m)c_1)$$

and since this last probability is not greater than

$$P\big(\cup_{j=1}^{k_n}(\rho_{nj} > m) | \rho_{n1}^* = i\big) \leqslant k_n P(\rho_{n1} > m | \rho_{n1}^* = i)$$
$$= 0(k_n^{-m}),$$

(3) now follows. Let $\bar{P}(\lambda) = P(\lambda) - \lambda$.

LEMMA 8. *As* $n \rightarrow \infty$,

$$(c_m - mc_1)^{-1} \Sigma_{i=0}^{N} X_n^{(i)} \circ \Omega_n^{(i)}(1) \rightarrow_{\mathfrak{D}} \bar{P}(A/m!).$$

PROOF. c.f. the proof of Lemma 4.

LEMMA 9. *As* $n \to \infty$,

$$\operatorname{Var}\Sigma_{j=1}^{l_n} \mu_{n\rho_{nj}^*} \leqslant c_1^2/B + 0(N/n).$$

PROOF. The result follows easily from the 'tedious calculations' involved in proving Lemma 5.

LEMMA 10. *As* $n \to \infty$,

$$E|\Sigma_{j=k_n l_n+1}^{n}(c_{\rho_{nj}} - Ec_{\rho_{nj}})| \leqslant 2|c_1|/B + 0(N/n).$$

PROOF. cf. the proof of Lemma 6.

We come now to the proof of the first part of Theorem 2. Write

$$S_n = \Sigma_{j=1}^{n}(c_{\rho_{nj}} - Ec_{\rho_{nj}})/(c_m - mc_1).$$

Lemmas 8, 9 and 10, together with the discussion preceding Lemma 7, show that

$$S_n = Z_{n1} + Z_{n2} + Z_{n3},$$

where $Z_{n1} \to_{\mathfrak{D}} \bar{P}(A/m!)$, $EZ_{ni} = 0$, $i = 1, 2, 3$,

$$\operatorname{Var} Z_{n2} \leqslant c_1^2 B^{-1}(c_m - mc_1)^{-2} + 0(N/n),$$

and

$$E|Z_{n3}| \leqslant 2|c_1|B^{-1}|c_m - mc_1|^{-1} + 0(N/n).$$

If $c_1 = 0$, the first part of Theorem 2 follows easily (with $B = 1$, say). If $c_1 \neq 0$, we argue as follows. Given $\varepsilon > 0$, choose B so large that for some integer n_B, $\operatorname{Var} Z_{n2} < \varepsilon^3$ and $E|Z_{n3}| < \varepsilon^2$ for $n > n_B$. Then

$$\begin{aligned} P(S_n \leqslant x) &\leqslant P(S_n \leqslant x, |Z_{n2}| < \varepsilon, |Z_{n3}| < \varepsilon) + P(|Z_{n2}| \geqslant \varepsilon) + P(|Z_{n3}| \geqslant \varepsilon) \\ &\leqslant P(Z_{n1} \leqslant x + 2\varepsilon) + 2\varepsilon \qquad \text{if } n > n_B \\ &\leqslant F(x + 3\varepsilon) + 3\varepsilon \end{aligned}$$

for all sufficiently large n, where F is the distribution function of $\bar{P}(A/m!)$. Similarly,

$$F(x - 3\varepsilon) - \varepsilon \leqslant P(Z_{n1} \leqslant x - 2\varepsilon) \leqslant P(S_n \leqslant x) + 2\varepsilon$$

for all large n, and the convergence in distribution of S_n now follows. Theorem 2 now follows from the result (still assuming $c_0 = 0$)

$$Ec_{\rho_{n1}} = n^{-1}\left(Nc_1 + \frac{A}{m!}(c_m - mc_1)\right) + o(1/n),$$

which follows in turn after some algebra from

$$Ec_{\rho_{n1}} = c_1 E\rho_{n1} I(\rho_{n1} < m) + c_m \binom{N}{m} n^{-m}(1 - 1/n)^{N-m} + o(N/n)^{m+1}.$$

3. Degenerate convergence.

THEOREM 3. *If* $n, N \to \infty$, $N^m/n^{m-1} \to 0$, *then*

$$P(\Sigma_{j=1}^{n} c_{\rho_{nj}} = nc_0 + N(c_1 - c_0)) \to 1.$$

PROOF. Once more, we assume without loss of generality that $c_0 = 0$. Then

$$P\left(\Sigma_{j=1}^{n} c_{\rho_{nj}} \neq Nc_1\right) = P\left(\cup_{j=1}^{n}(\rho_{nj} \geqslant m)\right)$$
$$\leqslant nP(\rho_{n1} \geqslant m).$$

Now

$$\begin{aligned} P(\rho_{n1} < m) &= \Sigma_{j=0}^{m-1}\binom{N}{j} n^{-j}(1 - 1/n)^{N-j} \\ &= (1 - 1/n)^N \Sigma_{j=0}^{m-1}(N/n)^j/j! + o(1/n) \\ &= \Sigma_{i,j=0}^{m-1}\frac{(-1)^i}{i!j!}(N/n)^{i+j} + o(1/n) \\ &= 1 + 0(N/n)^m + o(1/n) \\ &= 1 + o(1/n), \end{aligned}$$

so that

$$P\left(\Sigma_{j=1}^{n} c_{\rho_{nj}} \neq Nc_1\right) = o(1).$$

REFERENCES

BÉKÉSSY, A. (1963). On classical occupancy problems I. *Publ. Math. Inst. Hung. Acad. Sci.* **8** 59–71.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BROWN, B. M. and EAGLESON, G. K. (1971). Martingale convergence to infinitely divisible laws with finite variances. *Trans. Amer. Math. Soc.* **162** 449–453.

HARRIS, B. and PARK, C. J. (1971). The distribution of linear combinations of the sample occupancy numbers. *Indag. Math.* **33** 121–134.

JOHNSON, N. L. and KOTZ, S. (1977). *Urn Models and Their Application*. Wiley, New York.

MCLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probability* **2** 620–628.

PROHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Probability Appl.* **1** 157–214.

QUINE, M. P. (1979). A functional central limit theorem for a generalized occupancy problem. *Stochastic Process Appl.* **9**.

RÉNYI, A. (1962). Three new proofs and a generalisation of a theorem of Irving Weiss. *Publ. Math. Inst. Hung. Acad. Sci.* **7** 203–214.

SERFOZO, R. F. (1973). Weak convergence of superpositions of randomly selected partial sums. *Ann. Probability* **1** 1044–1056.

DEPARTMENT OF MATHEMATICAL STATISTICS
UNIVERSITY OF SYDNEY
SYDNEY, N.S.W. 2006
AUSTRALIA