Central Limit Theorems for Certain Infinite Urn Schemes

SAMUEL KARLIN*

1. Introduction. We consider an urn scheme where N balls are thrown independently at a fixed *infinite* array of cells with probability p_k of hitting the k^{th} cell. Without loss of generality we may assume the cells are ordered in such a way that $p_k \geqq p_{k+1}$ for $k = 1, 2, \cdots$ and of course $p_k > 0$ for all k. Let

$$(1) \qquad X_{N,k} = \text{number of balls in the } k^{\text{th}} \text{ cell after } N \text{ tosses.}$$

If the balls are thrown at the event times of a Poisson process $\{N(t); t \in [0, \infty)\}$ with parameter 1, we let

$$X_{N(t),k} = \text{number of balls in the } k^{\text{th}} \text{ cell at time } t.$$

(Hereafter, we reserve the notation $N(t)$ to represent a Poisson process.) The stochastic process $\{X_{N(t),k}\,; t \geqq 0\}$ for $k = 1, 2, \cdots$ is composed from mutually independent homogeneous Poisson processes with parameters p_k, $k = 1, 2, \cdots$, respectively. In particular, we have

$$\Pr\{X_{N(t),k} = r\} = e^{-tp_k}\frac{(tp_k)^r}{r!}.$$

In sharp contrast, the set of random variables $\{X_{N,k}\}$ with N fixed and k varying are not mutally independent.

The purpose of this paper is to prove central limit theorems and determine the asymptotic behavior of the moments for several special functionals of the processes $\{X_{N(t)\ k}\}_{k=1}^{\infty}$ and $\{X_{N,k}\}_{k=1}^{\infty}$. Some explicit functionals to be considered are the random variables $Z_{N,r}$ (the number of cells containing exactly r balls after N tosses) and

$$(2) \qquad Z_N^* = \sum_{r=1}^{\infty} Z_{N,r}\,,$$

the number of occupied cells. (We shall use the notation $Z_{N,r}^*$ to denote the number of cells containing at least r balls.)

* Research supported in part under contracts NO014-67-A-0112-0015 and NSF GP 4784 at Stanford University, Stanford, California.

Journal of Mathematics and Mechanics, Vol. 17, No. 4 (1967).

An alternative formulation leading to consideration of the quantity Z_N^* is the following. Let $\xi_1, \xi_2, \cdots$ be the independent random variables, each having the same distribution $\Pr\{\xi_n = k\} = p_k$, $k = 1, 2, \cdots$. Then Z_N^* is the number of distinct values assumed by the sequence $\{\xi_1, \xi_2, \cdots, \xi_N\}$.

Of some interest is the random variable

$$U_N^* = \sum_{k=1}^{\infty} U(X_{N,k}), \tag{3}$$

where

$$U(X_{N,k}) = \begin{cases} 1 & \text{if} \quad X_{N,k} \equiv 1 \pmod 2, \\ 0 & \text{otherwise}. \end{cases}$$

Thus U_N^* is the number of cells containing an odd number of balls. A novel interpretation for U_N^* occurs in Spitzer [10] in the following context. Consider an infinite sequence of light bulbs and an infinite sequence of numbers $\{p_k\}$, $p_k > 0$, $\sum_{k=1}^{\infty} p_k = 1$. At time 0 all the light bulbs are "Off". At each unit time ($t = 1, 2, 3, \cdots$) thereafter one of the light bulbs is selected (the k^{th} one with probability p_k) and its switch is turned. Thus it goes on if it was off, and off if it was on. The number of light bulbs on at time N is precisely the random variable U_N^*.

Another random variable of relevance to statistical applications is

$$C_N^* = \sum_{r=1}^{\infty} C_{N,r}, \tag{4}$$

where

$$C_{N,r} = \begin{cases} Np_r & \text{if} \quad X_{N,r} = 0, \\ 0 & \text{if} \quad X_{N,r} > 0, \end{cases}$$

which can be interpreted modulo a factor N, as the percentage of cells not covered resulting from N tosses. (A cell is said to be covered provided it contains at least one ball.)

To underscore concretely the scope of this paper, we state two of the principal theorems.

Theorem 4. *Let $\{p_k\}$ be such that $p_k \geqq p_{k+1}$, $k = 1, 2, \cdots$ and assume that $\alpha(x) = \max\{j \mid p_j \geqq 1/x\} = x^\gamma L(x)$, $0 < \gamma \leqq 1$ where $L(x)$ is a slowly varying function for $x \to \infty$. Then*

$$[Z_N^* - E(Z_N^*)]/(B_N)^{1/2}$$

converges in law to the normal distribution (with mean zero, variance 1) where

$$B_N = \begin{cases} \Gamma(1-\gamma)(2^\gamma - 1)N^\gamma L(N), & 0 < \gamma < 1, \\ NL^*(N), & \gamma = 1, \end{cases} \qquad L^*(N) = \int_0^\infty \frac{e^{-1/y}}{y} L(Ny)\, dy$$

and $E(Z_N^*)$ denotes as usual the expectation of Z_N^*.

Theorem 6. *Let p_k satisfy the conditions of Theorem 4. Then*

$$\frac{U_N^* - E(U_N^*)}{\frac{1}{2}(M(4N))^{1/2}}$$

converges in law to a standard normal distribution where

$$M(N) = \begin{cases} \Gamma(1-\gamma)N^\gamma L(N) & \text{for } 0 < \gamma < 1, \\ NL^*(N) & \text{for } \gamma = 1. \end{cases}$$

The proofs of Theorems 4 and 6 are given in section 4. The variable Z_N^* is simpler to deal with than U_N^* by virtue of the property that Z_N^* is monotone in N whereas U_N^* is clearly not a monotone function.

A multivariate central limit theorem for the vector random variable $\{Z_{N,r_1}, Z_{N,r_2}, \cdots, Z_{N,r_\nu}\}$, $N \to \infty$, $1 \leqq r_1 < r_2 < \cdots < r_\nu$, is described in section 4. A central limit theorem for the random variable C_N^* is also indicated.

The postulate $\alpha(x) = x^\gamma L(x)$ as $x \to \infty$, $0 < \gamma \leqq 1$ is virtually essential in order that Z_N^* be asymptotically normal (see paragraph III of section 6). In this connection note that if $\overline{\lim}_{k\to\infty}(p_{k+1}/p_k) < 1$ we find that $\alpha(x)$ is slowly varying (Lemma 1) but Var Z_N^* is bounded. This fact precludes the possibility that Z_N^*, properly normalized, obeys a central limit law.

In Theorem 1′ of section 2 the asymptotic growth behavior of $E(Z_N^*)$ is determined. We have

$$E(Z_N^*) \sim \begin{cases} \Gamma(1-\gamma)N^\gamma L(N) & \text{for } 0 \leqq \gamma < 1, \\ NL^*(N) & \text{for } \gamma = 1 \end{cases}$$

(see Theorem 1′, section 2).

If $\{p_k\}_{k=1}^\infty$ exhibits erratic behavior, *i.e.*, if $\alpha(x)$ is not a regular function in the sense of Karamata, then $E(Z_N^*)$ increases irregularly and Var Z_N^* could even oscillate wildly. The dependence of $E(Z_N^*)$ and Var Z_N^* on the nature of the parameters is fully elaborated in section 3.

The organization of the analysis proceeds as follows. In sections 2 and 3 we develop the asymptotic growth properties of the mean and variance for the random variables $Z_{N(t)}^*$, $Z_{N(t),r}$, $U_{N(t)}^*$, $C_{N(t)}^*$, Z_N^*, $Z_{N,r}$, U_N^* and C_N^*. Because of the independence property of the variables $\{X_{N(t),k}\}_{k=1}^\infty$, it is then a simple matter to ascertain asymptotic central limit theorems for various functionals of these variables. In section 4, central limit theorems are established for the variables Z_N^*, $Z_{N,r}$, U_N^* and C_N^* by exploiting the relationship between the variables Z_N^*, $Z_{N,r}$, U_N^* and C_N^* and $Z_{N(t)}^*$, $Z_{N(t),r}$, $U_{N(t)}^*$, $C_{N(t)}^*$, respectively, obtained by conditioning on the values of $N(t)$.

In section 5 a strong law of large numbers is validated for the variable $Z_N^*/E(Z_N^*)$ involving no restrictions on the probability sequence $\{p_k\}$. Corresponding strong laws for the variables $Z_{N,r}/E(Z_{N,r})$, $U_N^*/E(U_N^*)$ and $C_N^*/E(C_N^*)$ are confirmed under the conditions of Theorem 4.

Applications of the preceding developments to more general processes of tossing balls into cells are discussed in section 6. Furthermore, for some special cases where the hypotheses of Theorem 4 are not fulfilled, the nature of the limit law for the variable $Z_N^* - E(Z_N^*)$ is discerned.

We close the introduction by citing related literature bearing on the content of this paper. The classical urn problem has the following structure. N balls are tossed into n urns at random such that each ball has probability $1/n$ of falling in any specified urn. Under certain relationships of N and n it is proved that the sequence of vector random variables $\{\eta_{r_1}(N, n), \eta_{r_2}(N, n), \cdots, \eta_{r_r}(N, n)\}$ under proper normalization possesses a limiting multi-dimensional normal distribution as N and $n \to \infty$ where $\eta_{r_i}(N, n)$ denotes the number of urns containing exactly r_i balls (Sevastanov and Cistyakov [8]). Cistyakov [3] also investigated the situation where the probability of a ball falling into the i^{th} urn is

$$a_i = \frac{1}{n} g\left(\frac{i}{n}\right) - \frac{1}{2n^2} g'\left(\frac{i}{n}\right) + O\left(\frac{1}{n^3}\right), \qquad i = 1, 2, \cdots, n,$$

$a_i \geqq 0$, $\sum_{i=1}^n a_i = 1$, and $g(x) \geqq 0$ is some differentiable function defined on $[0, 1]$ such that $\int_0^1 g(x)\, dx = 1$. In this set up he proves for $n/N \to \alpha$ $(0 < \alpha < 1)$ and $N \to \infty$ that the random variable $\eta_0(n, N)$, the number of empty boxes, is asymptotically normally distributed. The arguments rely heavily on saddle point methods. For other investigations pertaining to the classical urn problem we refer to Rényi [7], Bekessey [2] and references therein.

The first discussion of a model of multinomial trials for an infinite array of urns seems to be due to Bahadur [1]. He investigated solely the growth properties of $E(Z_N^*)$ as $N \to \infty$ and secured bounds for Var Z_N^* adequate to demonstrate the fact that $Z_N^*/E(Z_N^*) \to 1$ in probability. The question of almost sure convergence left open by Bahadur is settled affirmatively in this paper.

Spitzer [10, p. 93] studied the recurrence properties of the event $A_N = \{w | U_N^* = 0\}$ exploiting the theory of random walks on a group. Darling [4] discussed some aspects of the same problem using direct arguments. Darling further claims that by refining the analysis set forth in [4] and exploiting the theory of stable processes a proof of the central limit theorem for the variable Z_N^* (under restrictions more special than those of this paper) can be carried out.

I am indebted to Mr. B. Singer for helpful discussions on the contents of this paper.

2. Some preliminaries. The growth properties of moments of Z_N^* and $Z_{N(t)}^*$ are sensitive functions of the probabilities $\{p_k\}$. The asymptotic behavior of the function

$$\alpha(x) = \sum_{\{k; 1/p_k \leqq x\}} 1 = \max\left\{j \mid p_j \geqq \frac{1}{x}\right\} \tag{5}$$

(recall that $p_r \geqq p_{r+1}$, $r = 1, 2, \cdots$ by stipulation) plays a key role in all subsequent analyses. It will be usual to postulate that $\alpha(x)$ is a regular function in the sense of Karamata, *i.e.*, $\alpha(x)$ obeys the asymptotic relation

$$\alpha(x) = x^{\gamma} L(x), \tag{6}$$

for some γ, $0 \leqq \gamma \leqq 1$ where $L(x)$ is a slowly varying function meaning that $L(cx)/L(x) \to 1$ as $x \to \infty$ for each fixed $c > 0$. We will always assume, without loss of generality, that $L(x)$ is continuous and finite at $x = 0$.

It is a formidable if not impossible task to determine the growth behavior of $E(Z_N^*)$ and $\mathrm{Var}(Z_N^*)$ without some condition of the kind expressed in (6). Although the postulate (6) has wide scope and applicability it does not universally prevail. The reader can readily supply examples of $\{p_k\}$ for which (6) fails. We will highlight several concrete cases for which (6) holds after proving the following lemma.

Lemma 1. *If $\{p_k\}$ is such that $P(\xi) = \sum_{k=1}^{\infty} p_k \xi^k$ has a radius of convergence exceeding* 1, *then $\alpha(x)$ defined in* (5) *is a slowly varying function of x.*

Proof. The hypothesis assures us that for some $\rho > 1$, we have

$$p_l \leqq 1/\rho^l \quad \text{for all} \quad l \geqq L(\rho). \tag{7}$$

Now for $c > 1$, consider

$$\alpha(cx) = \sum_{\{k;1/p_k \leqq cx\}} 1 = \alpha(x) + \sum_{\{k;1/cx \leqq p_k < 1/x\}} 1.$$

Let $\gamma(c) = l$ be determined as the least integer satisfying $1/\rho^l < 1/c$. The inequality (7) implies that

$$\sum_{\{k;1/cx \leqq p_k < 1/x\}} 1 \leqq \gamma(c) + L$$

which is a fixed upper bound independent of x. Since $\alpha(x) \uparrow \infty$ as $x \to \infty$ we deduce that

$$\frac{\alpha(cx)}{\alpha(x)} \to 1.$$

An entirely analogous argument proves the same limit relation when $0 < c < 1$. The proof of Lemma 1 is complete.

Remark 1. Imposing the slightly more stringent hypothesis $\overline{\lim}_{k \to \infty} p_{k+1}/p_k = \rho < 1$ (and therefore the radius of convergence of $P(\xi) = \sum_{k=1}^{\infty} p_k \xi^k$ is not smaller than $1/\rho$) we can characterize $\alpha(x)$ by a more familiar formula involving the distribution function $F(u) = \sum_{\{k:k<u\}} p_k$. To this end, we define $A(x)$ by the inequalities

$$1 - F(A(x)) \leqq \frac{1}{x} \leqq 1 - F(A(x)-).$$

Then $A(x) \sim \alpha(x)$ as $x \to \infty$.

Proof. Since $xp_{\alpha(x)} \geqq 1$ by definition, manifestly $\alpha(x) \leqq A(x) + 1$ and hence

$\underline{\lim}_{x\to\infty} A(x)/\alpha(x) \geqq 1$. Now, owing to the condition $\overline{\lim}_{k\to\infty} p_{k+1}/p_k < 1$ we deduce easily that $1/p_k \sum_{r=k}^{\infty} p_r \leqq C$ where C is a bound independent of k. It follows that $\sum_{r=\alpha(Cx)+1}^{\infty} p_r \leqq p_{\alpha(Cx)+1}\, C \leqq 1/x$ and therefore $A(x) \leqq \alpha(Cx) + 1$. But $\alpha(x)$ is slowly varying which implies

$$\overline{\lim_{x\to\infty}} \frac{A(x)}{\alpha(x)} \leqq \overline{\lim_{x\to\infty}} \frac{A(x)}{\alpha(Cx)} \frac{\alpha(Cx)}{\alpha(x)} \leqq 1.$$

The proof is complete.

Example 1. Let $p_k = (1 - \lambda)\, \lambda^{k-1}$, $0 < \lambda < 1$, $k = 1, 2, \cdots$. Then

$$\alpha(x) \sim \frac{\log x}{\log 1/\lambda}, \qquad x \to \infty .$$

Example 2. Suppose that $p_k = e^{-\lambda}\lambda^{k-1}/(k-1)!$, $k = 1, 2, \cdots$. Then,

$$\alpha(x) \sim \frac{\log x}{\log(\log x)}.$$

Example 3. The hypothesis of Lemma 1 is certainly not a necessary condition in order that $\alpha(x)$ be slowly varying. For example, if $p_k = C2^{-k^{\beta}}$ $(0 < \beta < 1)$, where C is a normalizing constant, then the radius of convergence of $P(\xi)$ is 1 but $\alpha(x) \sim (\log x/\log 2)^{1/\beta}$ is slowly varying.

Example 4. If $p_k \sim Ck^{-\beta}$ $(\beta > 1)$ as $k \to \infty$ then clearly $\alpha(x) \sim C^{1/\beta}\, x^{1/\beta}$.

Example 5. If $p_k = b/(k + 1)(\log (k + 1))^{\beta+1}$, where b is a normalizing constant and $\beta > 0$, then $\alpha(x) \sim x/b(\log x)^{1+\beta}$.

Under the conditions of Lemma 1 we can deduce by parallel arguments that the function $\alpha((1 + c)x) - \alpha(x)$, $x > 0$, is uniformly bounded for each $c > 0$. By way of contrast, we have

Lemma 2. *Suppose $p_{k+1}/p_k \to 1$ as $k \to \infty$ (therefore the radius of convergence of $P(\xi) = \sum_{k=1}^{\infty} p_k\, \xi^k$ is one). Then for $c > 0$*

$$\alpha((1 + c)x) - \alpha(x) \to \infty \quad \textit{as} \quad x \to \infty .$$

Proof. For any $\epsilon > 0$ there exists a $K(\epsilon)$ such that $1 \geqq p_{k+1}/p_k \geqq 1 - \epsilon$ for all $k \geqq K(\epsilon)$. Let l be determined as the least integer satisfying $1/(1 - \epsilon)^l < 1/(1 + c)$. An argument paraphrasing that of Lemma 1 shows that

$$\underline{\lim}_{x\to\infty} [\alpha((1 + c)x) - \alpha(x)] \geqq \frac{\log \dfrac{1}{1+c}}{\log \dfrac{1}{1-\epsilon}}$$

and since $\epsilon > 0$ is arbitrary, the right hand side can be made arbitrarily large.

We will also need the elementary facts.

Lemma 3. $\alpha(x)/x \to 0$ as $x \to \infty$ and $\int_1^{\infty}(\alpha(x)/x^2)\, dx \leqq 1$.

Proof. For any prescribed $\epsilon > 0$, choose K large enough and fixed such that $\sum_{k \geq K} p_k \leqq \epsilon$. Then for x sufficiently large $\alpha(x) = K + \sum 1$ where the sum is extended over the set $\{i : xp_i \geqq 1, i \geqq K\}$ and therefore $\alpha(x) \leqq K + \epsilon x$. It follows that $\alpha(x) = o(x)$ as $x \to \infty$. The second statement follows from the definition of $\alpha(x)$.

3. Asymptotic growth of moments. The analysis of the growth behavior of the first moments of $Z^*_{N(t)}$ and Z^*_N is presented first.

Clearly

$$M(t) = E(Z^*_{N(t)}) = \sum_{k=1}^{\infty} (1 - e^{-tp_k}). \tag{8}$$

In view of the definition of $\alpha(x)$ we may write

$$M(t) = \int_0^{\infty} [1 - e^{-t/x}]\, d\alpha(x). \tag{9}$$

Integration by parts and a change of variable yields

$$M(t) = \int_0^{\infty} \frac{t}{x^2} e^{-t/x} \alpha(x)\, dx = \int_0^{\infty} \frac{1}{y^2} e^{-1/y} \alpha(ty)\, dy.$$

Throughout this section, unless stated explicitly to the contrary, we postulate that $\alpha(x)$ is a function of regular growth in the sense of Karamata, *i.e.*, $\lim_{t\to\infty} (\alpha(ty)/\alpha(t)) = y^{\gamma}$ for each $y > 0$ or, equivalently, we have the representation

$$\alpha(x) = x^{\gamma} L(x) \quad \text{for some } \gamma \quad (0 \leqq \gamma \leqq 1), \tag{10}$$

where $L(x)$ is nonnegative and slowly varying as $x \to \infty$.

A standard Abelian argument using (10) produces the result

$$M(t) \sim \alpha(t) \int_0^{\infty} \frac{y^{\gamma}}{y^2} e^{-1/y}\, dy \qquad 0 \leqq \gamma < 1, \qquad t \to \infty \tag{11a}$$

and

$$M(t) = tL^*(t) \qquad \gamma = 1 \tag{11b}$$

where

$$L^*(t) = \int_0^{\infty} \frac{e^{-1/y}}{y} L(ty)\, dy. \tag{12}$$

Observe since the integral $\int_0^{\infty} (e^{-1/y}/y)\, dy$ diverges that

$$\frac{L^*(t)}{L(t)} \to 0 \quad \text{as} \quad t \to \infty. \tag{13}$$

Lemma 4. *When $\alpha(t) = tL(t)$ where $L(t)$ is slowly varying as $t \to \infty$ then $L^*(t)$ is slowly varying as $t \to \infty$.*

Proof. Notice that

$$(14)\qquad tL^*(t) = \int_0^\infty \frac{e^{-1/y}}{y}\, tL(ty)\, dy = \int_0^\infty \frac{e^{-1/y}}{y^2}\, \alpha(ty)\, dy \to \infty \quad \text{as} \quad t \to \infty$$

because $\alpha(t) \to \infty$ as $t \to \infty$.

We also know that $L(y)$ is bounded for $0 \leqq y < \infty$ since $L(y) \to 0$ as $y \to \infty$. Using this fact in conjunction with (14) we infer that for any fixed positive K

$$(15)\qquad \frac{\int_0^{K/t} (e^{-1/y}/y) L(ty)\, dy}{L^*(t)} \leqq \frac{C}{tL^*(t)} \to 0 \quad \text{as} \quad t \to \infty$$

where C is an appropriate positive constant.

Let $c > 0$ be fixed and $\epsilon > 0$ prescribed. Now determine K so large that

$$\left| \frac{L(cz)}{L(z)} - 1 \right| < \epsilon \quad \text{for} \quad z > K.$$

Consider

$$\left| \frac{L^*(ct)}{L^*(t)} - 1 \right| \leqq \frac{\int_0^{K/t} (e^{-1/y}/y)[L(cty) + L(ty)]\, dy}{L^*(t)} + \frac{\int_{K/t}^\infty (e^{-1/y}/y) \left| \frac{L(cty)}{L(ty)} - 1 \right| L(ty)\, dy}{L^*(t)}.$$

By the choice of K and referring to (15) we may conclude that

$$(16)\qquad \overline{\lim_{t\to\infty}} \left| \frac{L^*(ct)}{L^*(t)} - 1 \right| \leqq \epsilon.$$

Since ϵ can be specified arbitrarily small the assertion that $L^*(t)$ is slowly varying is established.

To sum up, we have proved

Theorem 1. *Let $\alpha(x) = x^\gamma L(x)$, $0 \leqq \gamma \leqq 1$ where $L(x)$ is a slowly varying function. (In the case $\gamma = 1$, necessarily $L(x) \to 0$ as $x \to \infty$; see Lemma 3.) Then $M(t) = E(Z^*_{N(t)})$ satisfies*

$$(17)\qquad \begin{aligned} &M(t) \sim \Gamma(1-\gamma) t^\gamma L(t), \qquad 0 \leqq \gamma < 1, \qquad t \to \infty \\ &M(t) \sim tL^*(t), \qquad \gamma = 1 \end{aligned}$$

where

$$(18)\qquad L^*(t) = \int_0^\infty \frac{e^{-1/y}}{y} L(ty)\, dy.$$

Remark 2. It may be of some interest to contrast concretely the growth behavior of $L^*(t)$ with that of $L(t)$. Specifically, if $L(t) \sim 1/(\log t)^\rho$, $\rho > 1$ (*cf.* Example 5 of section 2) then with the assistance of an Abelian theorem for Stieltjes transforms it is not difficult to prove

$$L^*(t) \sim \frac{1}{(\rho - 1)(\log t)^{\rho-1}}, \qquad t \to \infty.$$

In the same vein we find that if

$$L(t) \sim \frac{1}{(\log t)(\log \log t)^\rho}, \qquad \rho > 1$$

then

$$L^*(t) \sim \frac{1}{(\rho - 1)(\log \log t)^{\rho-1}}, \qquad t \to \infty.$$

We next study the growth behavior of $E(Z_N^*) = M_N = \sum_{k=1}^\infty [1 - (1 - p_k)^N]$. For this purpose note the easily established fact (using only the convergence $\sum p_k < \infty$) that

$$\sum_{k=1}^\infty \{(1 - p_k)^N - e^{-p_k N}\} \tag{19}$$

tends to zero as $N \to \infty$. In view of Theorem 1 we obtain immediately

Theorem 1'. *Let $\alpha(x)$ satisfy the conditions of Theorem 1. Then as $N \to \infty$*

$$M_N \sim \begin{cases} \Gamma(1 - \gamma)N^\gamma L(N), & 0 \leqq \gamma < 1 \\ NL^*(N), & \gamma = 1. \end{cases} \tag{20}$$

The same methods establish the following asymptotic results:

(i)

$$E(U_{N(t)}^*) = \frac{1}{2} \sum_{k=1}^\infty [1 - e^{-2tp_k}] \sim \begin{cases} \Gamma(1 - \gamma)2^{\gamma-1}t^\gamma L(t), & 0 \leqq \gamma < 1 \\ 2tL^*(t), & \gamma = 1 \end{cases} \tag{21}$$

where $U_{N(t)}^*$ is the analog of U_N^* (see (3)) with a Poisson arrival pattern of balls tossed into the cells.

(ii) Assuming the balls are tossed at the event times of a Poisson process $N(t)$ with parameter 1, let $M(t; r)$ be the expected number of cells containing exactly r balls ($r \geqq 1$) at time t. Thus

$$M(t; r) = E(Z_{N(t),r}) = \sum_{k=1}^\infty E[\varphi_r(X_{N(t),k})] = \sum_{k=1}^\infty e^{-tp_k} \frac{(tp_k)^r}{r!},$$

where $\varphi_r(\xi) = 1$ for $\xi = r$ and 0 otherwise. Now, we have

$$M(t;r) = \frac{1}{r!}\int_0^\infty e^{-t/x}\,\frac{t^r}{x^r}\,d\alpha(x). \tag{22}$$

Adapting the procedure of (11) we easily deduce the asymptotic formula

$$M(t;r) \sim \begin{cases} \gamma\,\dfrac{\Gamma(r-\gamma)}{\Gamma(r+1)}\,t^\gamma L(t), & 0<\gamma<1,\ r \geqq 1 \text{ or } \gamma = 1,\ r \geqq 2,\ t \to \infty \\ tL^*(t), & \gamma = 1,\ r = 1,\ t\to\infty. \end{cases} \tag{23}$$

The right hand formula of (23) is obviously meaningless when $\gamma = 0$. In this situation the behavior of $M(t;\,1)$ may be quite erratic. For example if $\overline{\lim}_{k\to\infty}\, p_{k+1}/p_k < 1$ holds, in which case $\gamma = 0$ by Lemma 1, the discussion in Remark 3 below can be adapted to demonstrate that $M(t;\,1)$ is bounded; but this function could oscillate irregularly as $t \to \infty$.

(iii) Finally consider the random variable $C^*_{N(t)}$ (see (4)). Clearly

$$C^*(t) = E(C^*_{N(t)}) = t\sum_{k=1}^\infty p_k e^{-p_k t} = M(t;\,1).$$

The asymptotic growth laws as $N \to \infty$ of the mean of the random variables U^*_N, $Z_{N,r}$, C^*_N are derived in a completely analogous manner to that of Theorem 1′. We will not state them explicitly except for the case of U^*_N.

Notice that

$$E(U(X_{N,k})) = \sum_{r=0}^\infty \binom{N}{2r+1} p_k^{2r+1}(1-p_k)^{N-2r-1} = \tfrac{1}{2}[1-(1-2p_k)^N].$$

Therefore

$$E(U^*_N) = \sum_{k=1}^\infty E(U(X_{N,k})) = \frac{1}{2}\sum_{k=1}^\infty [1-(1-2p_k)^N] \tag{24}$$

and this expression behaves asymptotically (see (19)) like

$$\frac{1}{2}\sum_{k=1}^\infty [1-e^{-2Np_k}].$$

Asymptotic growth of variance. In the remainder of this section we investigate the asymptotic behavior of the variance of the random variables $Z^*_{N(t)}$, $U^*_{N(t)}$, $Z_{N(t),r}$, $C^*_{N(t)}$ for $t\to\infty$. We focus attention first with complete details on the random variable $Z^*_{N(t)} = \sum_{k=1}^\infty \varphi(X_{N(t),k})$ where $\varphi(X_{N(t),k}) = 1$ if the k^{th} cell is occupied at time t and 0 otherwise. At this point it is convenient to exploit the property that $\{X_{N(t),k}\}_{k=1}^\infty$ and therefore $\{\varphi(X_{N(t),k})\}_{k=1}^\infty$ constitute sequences of independent random variables. (This special property intimately and decisively depends on the fact that balls are thrown singly at the event times of a Poisson process. Thus especially in the case of a fixed number N of tosses the sequence of random variables $\{X_{N,k}\}_{k=1}^\infty$ is *not* independent.) Because of the independence property we obtain

(25) $$V(t) = \operatorname{Var}[Z^*_{N(t)}] = \sum_{k=1}^{\infty} \operatorname{Var}(Z_{N(t),k}) = \sum_{k=1}^{\infty} (e^{-p_k t} - e^{-2p_k t}) = M(2t) - M(t)$$

(see (8)). It follows immediately from (17) that

(26) $$V(t) \sim \begin{cases} \Gamma(1-\gamma)(2^\gamma - 1)L(t)t^\gamma, & 0 < \gamma < 1, \quad t \to \infty, \\ tL^*(t), & \gamma = 1, \quad t \to \infty. \end{cases}$$

(Notice here the essential requirement that the parameter γ be positive in contrast to (17) where we also permit $\gamma = 0$.) When $\gamma = 0$ by completely analogous considerations we obtain the formula

(27) $$V(t) \sim L(2t) - L(t), \qquad \gamma = 0$$

provided $L(2t) - L(t) \to \infty$ as $t \to \infty$. Consulting Lemma 2 we see that (27) certainly prevails when $\alpha(x)$ is a slowly varying function of x and the radius of convergence of $\sum_{k=1}^{\infty} p_k \xi^k$ is equal to 1.

The growth behavior of $V(t)$ in the case where $\gamma = 0$ can generally be quite irregular. It is illuminating to inquire into the nature of $V(t)$ when $\alpha(x) = L(x)$ is slowly varying in some special circumstances. The following three remarks and Theorem 2 below pertain to this question.

Remark 3. If $\overline{\lim}_{k\to\infty} p_{k+1}/p_k < 1$ then $V(t)$ is bounded.

For the proof, consider $t \to \infty$ and determine k_t such that $1/p_{k_t} \leqq t < 1/p_{k_t+1}$. Also let K_0 be fixed such that $p_{k+1}/p_k < c < 1$ for all $k \geqq K_0$. We estimate $V(t)$ by breaking the sum into three parts;

$$V(t) = \sum_{r=1}^{K_0} [e^{-p_r t} - e^{-2p_r t}] + \sum_{K_0+1}^{k_t} [e^{-p_r t} - e^{-2p_r t}] + \sum_{k_t+1}^{\infty} [\quad] = I_1 + I_2 + I_3 .$$

The first sum consists of a fixed number of terms and obviously goes to zero as $t \to \infty$. The second sum is estimated above by

$$2 \sum_{r=K_0+1}^{k_t} \exp[-p_r/p_{k_t}] \leqq 2 \sum_{r=K_0+1}^{k_t} \exp\left[-\left(\frac{1}{c}\right)^{k_t - r}\right] \leqq 2 \sum_{r=0}^{\infty} \exp\left[-\left(\frac{1}{c}\right)^{r}\right]$$

which is bounded, since $1/c > 1$. The last sum is bounded by

$$2 \sum_{r=k_t+1}^{\infty} (1 - \exp(-2p_r/p_{k_t+1}))$$
$$\leqq 4 \sum_{r=k_t+1}^{\infty} \frac{p_r}{p_{k_t+1}} \leqq 4 \sum_{r=k_t+1}^{\infty} c^{r-k_t-1} < 4 \sum_{r=0}^{\infty} c^r < \infty .$$

Remark 4. Under the conditions of Remark 3 above, it is not necessarily true that $\lim_{t\to\infty} V(t) = v_0$ exists. For example, if $p_k = C2^{-2^k}$, $k = 1, 2, \cdots$ (C is a normalizing constant) then paraphrasing the analysis above we can show that for $t_l = 2^{2^l}$, $l = 1, 2, 3, \cdots$, $V(t_l) \geqq \alpha > 0$ but for $\tau_l = 2^{2^l + l}$, $V(\tau_l) \to 0$ as $l \to \infty$. Observe that

$$\overline{\lim_{k\to\infty}} \frac{p_{k+1}}{p_k} = \overline{\lim_{k\to\infty}} 2^{-2^k} = 0.$$

The conclusion arrived at in Remark 3 for the example at hand affirms that $V(t)$ is uniformly bounded.

We will now exhibit an example where $V(t)$ oscillates unboundedly. Specifically, let

$$p_k = C2^{-2^r}, \quad \frac{r(r-1)}{2} + 1 \leqq k \leqq \frac{r(r+1)}{2}, \quad r = 1, 2, \cdots, \tag{28}$$

(C is a normalizing constant).

Direct examination reveals that for $t_l = 2^{2^l}$, $V(t_l) \geqq l(e^{-1} - e^{-2})$ which tends to ∞ as $l \to \infty$. On the other hand $V(\tau_l) \to 0$, $l \to \infty$ for $\tau_l = 2^{2^l+l}$. Of course by virtue of Lemma 1, $\alpha(x)$ defined by (5) corresponding to the sequence (28) is a slowly varying function.

There is a large class of probability sequences $\{p_k\}_1^\infty$ for which $\lim_{t\to\infty} V(t)$ exists. The following theorem indicates one such class.

Theorem 2. *Let $\{p_k\}_1^\infty$ be such that $\overline{\lim}_{k\to\infty} p_{k+1}/p_k < 1$. Suppose $\alpha(x)$ satisfies*

$$\lim_{x\to\infty} \frac{1}{x} \int_0^x [\alpha(2\xi) - \alpha(\xi)]\, d\xi = \gamma_0 \tag{29}$$

with $0 < \gamma_0 < \infty$. Then $\lim_{t\to\infty} V(t)$ exists and equals γ_0.

Proof. From the definition of $\alpha(x)$ we can write

$$V(t) = \int_0^\infty [e^{-t/x} - e^{-2t/x}]\, d\alpha(x). \tag{30}$$

Integration by parts yields

$$V(t) = \int_0^\infty \left[\frac{2t}{x^2} e^{-2t/x} - \frac{t}{x^2} e^{-t/x}\right] \alpha(x)\, dx = \int_0^\infty \frac{t}{x^2} e^{-t/x}[\alpha(2x) - \alpha(x)]\, dx. \tag{31}$$

Another integration by parts gives the formula

$$V(t) = \int_0^\infty \left(\frac{2t}{x^3} - \frac{t^2}{x^4}\right) e^{-t/x} \left(\int_0^x [\alpha(2\xi) - \alpha(\xi)]\, d\xi\right) dx \tag{32}$$

and a change of variable produces

$$V(t) = \int_0^\infty e^{-1/y} \left(\frac{2}{y^2} - \frac{1}{y^3}\right) \frac{1}{yt} \left(\int_0^{yt} [\alpha(2\xi) - \alpha(\xi)]\, d\xi\right) dy. \tag{33}$$

By virtue of (29) we infer (by dominated convergence since $\alpha(2\xi) - \alpha(\xi)$ is bounded; see the remark preceding Lemma 2) that

$$\lim_{t\to\infty} V(t) = \gamma_0 . \tag{34}$$

An extension of this theorem to deal with higher moments of $Z^*_{N(t)}$ will be indicated in Section 6.

Example 6. An example of $\{p_k\}$ fulfilling the hypothesis of Theorem 2 is

$$p_r = (1 - \rho)\rho^{r-1}, \qquad r = 1, 2, 3, \cdots, \qquad 0 < \rho < 1. \tag{35}$$

It is evident that $\alpha(x) = [\log_{1/\rho} x(1 - \rho)]$ where the bracket symbol $[h]$ signifies the greatest integer not exceeding h and logarithm with a subscript means that the logarithm is evaluated to that base.

It is a simple matter to verify here that

$$\lim_{x \to \infty} \frac{1}{x} \int_0^x \{\alpha(2u) - \alpha(u)\}\, du = \log_{1/\rho} 2 = \gamma_0 .$$

Thus for the case of (35), Theorem 2 tells us that Var $(Z^*_{N(t)})$ converges to $\log_{1/\rho} 2$ as $t \to \infty$.

We record next the outcome of the calculations on the variance of the variables $U^*_{N(t)}$, $Z_{N(t),r}$, $C^*_{N(t)}$.

(a)

$$\operatorname{Var}[U^*_{N(t)}] = \tfrac{1}{4}M(4t) \tag{36}$$

where $M(t)$ is defined in (8).

(b)

$$\operatorname{Var}[Z_{N(t),r}] \sim \begin{cases} \dfrac{\gamma}{\Gamma(r+1)}\, t^\gamma L(t)\left[\Gamma(r-\gamma) - \dfrac{2^\gamma}{2^{2r}} \dfrac{\Gamma(2r-\gamma)}{\Gamma(r+1)}\right], \\ \qquad 0 < \gamma < 1, \quad r \geqq 1 \quad \text{or} \quad \gamma = 1, \quad r \geqq 2, \\ tL^*(t), \qquad \gamma = 1, \quad r = 1 \quad \text{for} \quad t \to \infty. \end{cases} \tag{37}$$

(c)

$$\operatorname{Var} C^*_{N(t)} \sim \gamma\Gamma(2-\gamma)[1 - 2^{\gamma-2}]t^\gamma L(t), \qquad 0 < \gamma \leq 1 \quad \text{for} \quad t \to \infty. \tag{38}$$

It is not difficult to determine the growth behavior of the variance of the variables Z^*_N, $Z_{N,r}$, C^*_N and U^*_N. We briefly illustrate the method in the case of Z^*_N and U^*_N. A direct computation gives

$$\begin{aligned} \operatorname{Var} Z^*_N &= E\left[\sum_{k=1}^\infty \varphi(X_{N,k})\right]^2 - \left[\sum_{k=1}^\infty E(\varphi(X_{N,k}))\right]^2 \\ &= \sum_{k=1}^\infty [(1-p_k)^N - (1-p_k)^{2N}] \\ &\qquad + \sum_{1 \leq k \neq l \leq \infty} [(1 - p_k - p_l)^N - (1-p_k)^N(1-p_l)^N] \\ &= I_1 + I_2 \end{aligned} \tag{39}$$

where

$$\varphi(u) = \begin{cases} 1, & u > 0, \\ 0, & u = 0. \end{cases}$$

In view of (19) we know that

$$\sum_{k=1}^{\infty} [(1 - p_k)^N - (1 - p_k)^{2N}] - \sum_{k=1}^{\infty} [e^{-p_k N} - e^{-2p_k N}]$$

tends to zero as $N \to \infty$. Now executing obvious estimates on the terms of I_2 yields

$$\begin{aligned} |I_2| &\leqq \sum_{1 \leqq k \neq l < \infty} p_k p_l N(1 - p_k)^N (1 - p_l)^N \\ &\leqq \max_{1 \leqq k \leqq K} [N(1 - p_k)^N] + \left(\sum_{k=K+1}^{\infty} p_k\right) \sum_{l=1}^{\infty} (1 - \exp[-Np_l]). \end{aligned} \tag{40}$$

The inequality (40) implies $|I_2| = o(M_N)$ as $N \to \infty$ where $M_N = E(Z_N^*)$. Now assume $\alpha(x) = x^\gamma L(x)$, $0 < \gamma \leqq 1$ where $L(x)$ is slowly varying. In view of (17) we find that

$$\begin{aligned} \operatorname{Var} Z_N^* &\sim \sum_{k=1}^{\infty} [e^{-p_k N} - e^{-2p_k N}] \\ &\sim \begin{cases} \Gamma(1 - \gamma)(2^\gamma - 1)N^\gamma L(N), & 0 < \gamma < 1, \\ NL^*(N), & \gamma = 1. \end{cases} \end{aligned} \tag{41}$$

The determination of the asymptotic growth behavior of $\operatorname{Var} Z_{N,r}$, $\operatorname{Var} C_N^*$ and $\operatorname{Var} U_N^*$ proceed analogously. The explicit formula for $\operatorname{Var} U_N^*$ is needed later. We have

$$\begin{aligned} \operatorname{Var} [U_N^*] &= E\left(\sum_{k=1}^{\infty} U(X_{N,k})\right)^2 - [E(U_N^*)]^2 \\ &= \tfrac{1}{4} \sum_{k=1}^{\infty} (1 - (1 - 2p_k)^{2N}) \\ &\qquad + \tfrac{1}{4} \sum_{i \neq j} [(1 - 2p_i - 2p_j)^N - (1 - 2p_i)^N (1 - 2p_j)^N]. \end{aligned} \tag{42}$$

The last calculation is based on the identity

$$E(U(X_{N,k})U(X_{N,l})) = \tfrac{1}{4}\{1 - (1 - 2p_k)^N - (1 - 2p_l)^N + (1 - 2p_k - 2p_l)^N\},$$
$$1 \leq k \neq l < \infty$$

(*cf.* (24)). An analysis paralleling that of (40) reveals that the second sum is of smaller order of magnitude than $\frac{1}{4} \sum_{k=1}^{\infty} \{1 - (1 - 2p_k)^{2N}\}$.

We conclude this section by citing the result of additional calculations on higher moments under the condition $\alpha(x) = x^\gamma L(x)$, $0 < \gamma \leqq 1$ where $L(x)$ is slowly varying. A straightforward induction establishes

(43) $$E\{Z^*_{N(t)} - M(t)\}^{2m} \sim d_{\gamma,m}[M(t)]^m \qquad (m \text{ is a positive integer})$$

where $d_{\gamma,r}$ denotes an appropriate positive constant and $M(t)$ is defined in (8). Similar asymptotic relations obtain for the higher central moments of the variables $C^*_{N(t)}$, $Z_{N(t),r}$ and $U^*_{N(t)}$.

4. Asymptotic normality. The objective of this section is to establish asymptotic normality for the random variables Z^*_N , $Z_{N,r}$, U^*_N and C^*_N , $N \to \infty$ under the conditions of Theorem 4. The *modus operandi* runs as follows. By standard techniques in the spirit of the Linderberg conditions, the asymptotic normality of the variables $Z^*_{N(t)}$, $Z_{N(t),r}$, $U^*_{N(t)}$, and $C^*_{N(t)}$ is easily ascertained. Then we infer the limit behavior of Z^*_N , $Z_{N,r}$, U^*_N and C^*_N from that of $Z^*_{N(t)}$, $Z_{N(t),r}$, $U^*_{N(t)}$ and $C^*_{N(t)}$, respectively, by suitably conditioning on the values of $N(t)$, and exploiting certain smoothness properties with respect to N inherent to the random variables in question. The detailed analysis is presented first in the case of Z^*_N .

Asymptotic normality of $Z^*_{N(t)}$ and Z^*_N . Observe the representation

(44) $$Z^*_{N(t)} = \sum_{k=1}^{\infty} \varphi(X_{N(t),k})$$

where

$$\varphi(u) = \begin{cases} 1, & u > 0, \\ 0, & u = 0. \end{cases}$$

Since $\{X_{N(t),k}\}_{k=1}^{\infty}$ constitute a sequence of independent Poisson variables with parameters $\{tp_k\}_{k=1}^{\infty}$, respectively, it is clear that for each t

$$Y_{N(t),k} = \varphi(X_{N(t),k}) \qquad k = 1, 2, \cdots$$

are mutually independent binomial variables and in particular $Y_{N(t),k}$ are uniformly bounded. A simple variant of the method used in connection with the Linderberg type criteria for triangular arrays establishes, *mutatis mutandis,* that

$$\frac{Z^*_{N(t)} - E(Z^*_{N(t)})}{(\operatorname{Var}(Z^*_{N(t)}))^{1/2}}$$

converges in law to the standard normal distribution function provided $\operatorname{Var} Z^*_{N(t)} \to \infty$ as $t \to \infty$.

We summarize the above discussion in a more general setting as follows.

Theorem 3. *The variables*

$$\frac{Z^*_{N(t)} - E(Z^*_{N(t)})}{(\operatorname{Var} Z^*_{N(t)})^{1/2}}, \qquad \frac{Z_{N(t),r} - E(Z_{N(t),r})}{(\operatorname{Var} Z_{N(t),r})^{1/2}}, \qquad \frac{U^*_{N(t)} - E(U^*_{N(t)})}{(\operatorname{Var} U^*_{N(t)})^{1/2}}$$

and

$$\frac{C^*_{N(t)} - E(C^*_{N(t)})}{(\mathrm{Var}\, C^*_{N(t)})^{1/2}}$$

converge in law to the standard normal distribution provided $\mathrm{Var}(Z^*_{N(t)})$, $\mathrm{Var}(Z_{N(t),r})$, $\mathrm{Var}(U^*_{N(t)})$ *and* $\mathrm{Var}(C^*_{N(t)})$, *respectively, tend to* ∞ *as* $t \to \infty$.

The formula (36) in conjuction with the fact that $M(t) \to \infty$ as $t \to \infty$ shows that Var $U^*_{N(t)}$ always increases to ∞ as $t \to \infty$ and therefore $U^*_{N(t)}$, properly normalized, is asymptotically normally distributed. This property does not persist without some restrictions on the $\{p_k\}$ in the case of the quantities $\mathrm{Var}(Z^*_{N(t)})$, $\mathrm{Var}(Z_{N(t),r})$ and $\mathrm{Var}(C^*_{N(t)})$; Remarks 3 and 4 of section 3 are pertinent here.

We are now prepared to prove Theorem 4 cited in the introduction.

Theorem 4. *Let* $\{p_k\}_{k=1}^\infty$ *be such that* $\alpha(x) = x^\gamma L(x)$, $0 < \gamma \leqq 1$ *where* $L(x)$ *is slowly varying (see* (6)*). Then*

$$\frac{Z^*_N - E(Z^*_N)}{(B_N)^{1/2}} \xrightarrow{Law} \textit{Normal (mean 0, var 1)}$$

where

$$B_N = \begin{cases} \Gamma(1-\gamma)(2^\gamma - 1)N^\gamma L(N), & 0 < \gamma < 1, \\ NL^*(N), & \gamma = 1, \end{cases}$$

and $L^*(N)$ *is defined in* (18).

To ease the exposition, the proof of Theorem 4 commences with three preliminary lemmas.

Lemma 5. *Suppose the conditions of Theorem 4 hold. Let* $N' = [N + c\,N^{1/2}]$. ($[h]$ *designates the largest integer not exceeding h.) Then for any* $c_0 > 0$

$$\sup_{|c| \leq c_0} \frac{|E(Z^*_N) - E(Z^*_{N'})|}{(B_N)^{1/2}} \to 0 \quad \text{as} \quad N \to \infty \tag{45}$$

Proof. For definiteness take $c > 0$ and consider first the case of $0 < \gamma < 1$. We will prove the inequality

$$|E(Z^*_N) - E(Z^*_{N'})| \leqq KN^{\gamma - 1/2}L(N) \tag{46}$$

for some constant K depending on c_0 only. In fact since $N' > N$, we have

$$0 \leqq E(Z^*_{N'}) - E(Z^*_N) \leqq \sum_{k=1}^\infty (1-p_k)^N[1 - (1-p_k)^{c'N^{1/2}}]$$

for $c' = c + 1$. By elementary estimates we get

$$\sum_{k=k_0}^\infty (1-p_k)^N[1-(1-p_k)^{c'N^{1/2}}] \leqq \sum_{k=k_0}^\infty e^{-p_kN}(1 - \exp(-c''N^{1/2}p_k))$$

$$\leqq c'' \frac{1}{N^{1/2}} \sum_{k=1}^\infty e^{-p_kN}(Np_k),$$

where k_0 is determined such that $p_k < \frac{1}{2}$ when $k \geqq k_0$. Now the asymptotic growth of $\sum_{k=1}^{\infty} e^{-p_k N}(Np_k)$ was evaluated in (23) to be of the order $\gamma\, \Gamma(1-\gamma)$ $N^\gamma\, L(N)$, and therefore (46) is confirmed. Further scrutiny of the above argument reveals that the constant K in (46) can be chosen independent of c provided c varies in a bounded set. When $\gamma = 1$, the right hand side in (46) has to be replaced by $KN^{1/2}L^*(N)$.

The conclusion of (45) is established since $N^{(\gamma-1)/2}\, L(N) \to 0$ in the case $\gamma < 1$ and when $\gamma = 1$ because of the property $(L^*(N))^{1/2} \to 0$.

Lemma 6. *Let the hypotheses and notation of Lemma 5 prevail. For any prescribed $\epsilon > 0$ we have*

$$\Pr\left\{\frac{|Z_{N'}^* - Z_N^*|}{(B_N)^{1/2}} > \epsilon\right\} \to 0 \quad \text{as} \quad N \to \infty \tag{47}$$

uniformly for c restricted to a bounded domain.

Proof. Applying Tchebycheff's inequality for nonnegative random variables we obtain

$$\Pr\left\{\frac{|Z_{N'}^* - Z_N^*|}{(B_N)^{1/2}} > \epsilon\right\} \leqq \frac{|E(Z_{N'}^*) - E(Z_N^*)|}{\epsilon (B_N)^{1/2}}$$

and the right hand side tends to zero in view of the result of Lemma 5.

Lemma 7. *Let $N(t)$ be a Poisson process with parameter 1. Under the same hypotheses as Lemma 5*

$$E(Z_{N(t)}^*) - E(Z_{[t]}^*) \to 0 \quad \text{as} \quad t \to \infty \tag{48}$$

where $[t]$ denotes, as usual, the greatest integer not exceeding t.

Proof. We obtain

$$\begin{aligned} |E(Z_{N(t)}^*) - E(Z_{[t]}^*)| &\leqq \sum_{k=1}^{\infty} |e^{-p_k t} - (1-p_k)^{[t]}| \leqq C' \sum_{k=1}^{\infty} e^{-p_k t} p_k^{\,2} t \\ &\leqq \frac{C'}{t} \sum_{k=1}^{\infty} e^{-p_k t} p_k^{\,2} t^2 \leqq C'' \begin{cases} t^{\gamma-1} L(t), & 0 < \gamma < 1, \\ L^*(t), & \gamma = 1 \end{cases} \end{aligned}$$

and thereby (48) follows.

We are now in possession of the ingredients required to complete the proof of Theorem 4.

Proof of Theorem 4. We start with the identity

$$\Pr\left\{\frac{Z_{N(t)}^* - E(Z_{[t]}^*)}{(B_{[t]})^{1/2}} \leqq x\right\} = \sum_{k=0}^{\infty} \Pr\left\{\frac{Z_k^* - E(Z_{[t]}^*)}{(B_{[t]})^{1/2}} \leqq x\right\} \frac{t^k e^{-t}}{k!} \tag{49}$$

obtained by conditioning on the values of $N(t)$. With ϵ given and t specified sufficiently large we may determine $C_1(\epsilon)$ large enough fixed and independent of t so that

$$\sum_{k=t-C_1(\epsilon)t^{1/2}}^{t+C_1(\epsilon)t^{1/2}} \frac{t^k e^{-t}}{k!} \geqq 1 - \epsilon.$$

Lemma 5 asserts for $t > t_0(\epsilon, \delta)$ and all k satisfying $|k - t| \leqq C_1 t^{1/2}$ the inequality $\Pr\{|Z^*_{[t]} - Z^*_k| > \delta (B_{[t]})^{1/2}\} \leqq \epsilon$ where δ is preassigned arbitrarily small but fixed.

It follows for $t > t_0(\epsilon, \delta)$ that

$$\Pr\left\{\frac{Z^*_k - E(Z^*_{[t]})}{(B_{[t]})^{1/2}} \leqq x\right\} \leqq \Pr\left\{\frac{Z^*_{[t]} - E(Z^*_{[t]})}{(B_{[t]})^{1/2}} \leqq x + \frac{Z^*_{[t]} - Z^*_k}{(B_{[t]})^{1/2}}\right\}$$
$$\leqq F_{[t]}(x + \delta) + \epsilon,$$

where by definition

$$F_{[t]}(\xi) = \Pr\{Z^*_{[t]} - E(Z^*_{[t]}) \leqq \xi(B_{[t]})^{1/2}\}.$$

The above estimates applied in (49) yield

$$\Pr\{Z^*_{N(t)} - E(Z^*_{[t]}) \leqq x(B_{[t]})^{1/2}\} \leqq F_{[t]}(x + \delta) + \epsilon \tag{50}$$

provided $t \geqq t_0(\varepsilon, \delta)$. The asymptotic normality of $Z^*_{N(t)}$ as enunciated in Theorem 4 coupled with the result of Lemma 7 and the asymptotic relation $\operatorname{Var} Z^*_{N(t)} \sim B_{[t]}$ implies

$$\lim_{t\to\infty} \Pr\{Z^*_{N(t)} - E(Z^*_{[t]}) \leqq x(B_{[t]})^{1/2}\} = \Phi(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{1/2}} e^{-\xi^2/2}\, d\xi.$$

Proceeding to a limit ($t \to \infty$) and taking account of (50) produces the inequality

$$\Phi(x) \leqq \varliminf_{t\to\infty} F_{[t]}(x + \delta), \qquad -\infty < x < \infty \tag{51}$$

for any $\delta > 0$. In a similar way we deduce

$$\Phi(x) \geqq \varlimsup_{t\to\infty} F_{[t]}(x - \delta), \qquad -\infty < x < \infty. \tag{52}$$

A standard argument based on the relations (51) and (52) proves

$$\lim_{N\to\infty} F_N(x) = \lim_{t\to\infty} F_{[t]}(x) = \Phi(x).$$

The demonstration of Theorem 4 is complete.

In the case of the family of random variables $\{Z_{N,r}\}_{r=1}^{\infty}$ we can establish a multivariate central limit theorem by suitably adapting the method of the proof of Theorem 4. The estimates paralleling Lemma 5–7 for the random variables $Z_{N,r}$ are derived with the help of the relations $Z_{N,r} = Z^*_{N,r} - Z^*_{N,r-1}$ and exploitation of the basic fact that $Z^*_{N,r}$ is monotone in N for fixed r.

The precise theorem is as follows.

Theorem 5. *Suppose* $\alpha(x) = x^\gamma L(x)$, $0 < \gamma < 1$. *Let* $r_1 < r_2 < \cdots < r_\nu$ *be* ν *prescribed positive integers. Then the vector random variable*

$$\frac{Z_{N,r_1} - E(Z_{N,r_1})}{(N^\gamma L(N))^{1/2}}, \quad \frac{Z_{N,r_2} - E(Z_{N,r_2})}{(N^\gamma L(N))^{1/2}}, \cdots, \frac{Z_{N,r_\nu} - E(Z_{N,r_\nu})}{(N^\gamma L(N))^{1/2}}$$

has an asymptotic $(N \to \infty)$ *multivariate normal distribution with zero mean vector and convariance matrix*

$$\sigma_{r_i,r_j} = -\frac{\gamma\Gamma(r_i + r_j - \gamma)}{r_i\,!\,r_j\,!} 2^{\gamma - r_i - r_j}, \qquad i \neq j,$$

$$\sigma_{r_i,r_i} = \frac{\gamma}{\Gamma(r_i + 1)}\left[\Gamma(r_i - \gamma) - 2^{-2r_i+\gamma}\frac{\Gamma(2r_i - \gamma)}{\Gamma(r_i + 1)}\right], \qquad i = 1, 2, \cdots.$$

The corresponding version of Theorem 5 when $\gamma = 1$ is somewhat surprising, as there is a sharp difference in growth behavior for the random variable $Z_{N,1}$ compared with the random variables $Z_{N,r}$, $r > 1$. Thus we have

Theorem 5′. *Suppose* $\alpha(x)$ *obeys the asymptotic relation* $\alpha(x) = xL(x)$, $x \to \infty$, *where* $L(x)$ *is slowly varying and tends to zero as* $x \to \infty$ *(see Lemma 3). Then as* $N \to \infty$

$$\frac{Z_{N,1} - E(Z_{N,1})}{(NL^*(N))^{1/2}} \xrightarrow{Law} Normal\ (0, 1),$$

($L^*(N)$ *is defined in* (18)) *while for* $r \geqq 2$

$$\frac{Z_{N,r} - E(Z_{N,r})}{(b_r NL(N))^{1/2}} \xrightarrow{Law} Normal\ (0, 1),$$

where

$$b_r = \frac{\Gamma(r - 1)}{\Gamma(r + 1)} - 2^{1-2r}\frac{\Gamma(2r - 1)}{[\Gamma(r + 1)]^2}.$$

Notice that the asymptotic normalizing variance for $Z_{N,1}$ is of a larger order of magnitude than that of $Z_{N,r}$, $r > 2$ (since $L^*(N)/L(N) \to \infty$ as $N \to \infty$; see (13)).

The analysis in Lemma 6 exploited decisively the inherent monotonicity of the random variable Z_N^* as a function of N. This property is definitely lacking for the variable U_N^*. Nevertheless, we have

Theorem 6. *Let* $\{p_k\}_1^\infty$ *satisfy the conditions of Theorem 4. We have*

$$\frac{U_N^* - E(U_N^*)}{\frac{1}{2}(M(4N))^{1/2}} \xrightarrow{Law} Normal\ (mean\ 0, var\ 1)$$

where $M(t)$ *is defined in* (8).

The burden of the proof is to establish the analogs of Lemmas 5–7. We will need the elementary fact

$$M(N) \leqq M(4N) \leqq 4M(N). \tag{53}$$

Lemma 5'. *Let the hypotheses of Theorem 6 prevail. Let* $N' = [N + cN^{1/2}]$. *Then for any* $c_0 > 0$

$$\sup_{|c|<c_0} \frac{|E(U_N^*) - E(U_{N'}^*)|^2}{M(N)} \to 0 \quad \text{as} \quad N \to \infty. \tag{54}$$

Proof. For definiteness, we examine the circumstance $N' > N$; the alternative case is handled by similar means. Consulting (24) and using the trivial inequalities $1 - x^N \leqq (1 - x)N$, $0 < x < 1$, and $ye^{-y} \leqq 1 - e^{-y}$, $y > 0$, we obtain

$$\frac{|E(U_N^*) - E(U_{N'}^*)|^2}{M(N)} = \frac{1}{4} \frac{\left(\sum_{k=1}^{\infty} (1 - 2p_k)^N \{1 - (1 - 2p_k)^{cN^{1/2}}\}\right)^2}{M(N)}$$

$$\leqq \frac{c^2}{4N} \frac{\left(\sum_{k=1}^{\infty} e^{-2p_kN} 2p_kN\right)^2}{M(N)} \leqq c^2 \frac{M(N)}{N}.$$

But $M(N)/N \to 0$ as $N \to \infty$ and the assertion of (54) is thereby confirmed.

In the proof of Lemma 6 we employed Tchebycheff's inequality for positive random variables. In the present context we invoke the Tchebycheff inequality involving second moments. Specifically, we need to estimate

$$P\left\{\frac{|U_N^* - N_{N'}^*|}{(M(4N))^{1/2}} > \epsilon\right\} \leqq \frac{E((U_N^* - U_{N'}^*)^2)}{\epsilon^2 M(4N)}.$$

Lemma 6'. *Adhering to the notation and hypotheses of Lemma 5' we have*

$$\frac{E((U_N^* - U_{N'}^*)^2)}{M(N)} \to 0 \quad \text{as} \quad N \to \infty. \tag{55}$$

Proof. For the calculation of the expectation in (55) it is convenient to express U_N^* in the form

$$U_N^* = \sum_{k=1}^{\infty} U(X_{N,k}) \qquad (cf.\ (3))$$

where

$$U(X_{N,k}) = \begin{cases} 1 & \text{if} \quad X_{N,k} \equiv 1 \pmod 2 \\ 0 & \text{otherwise.} \end{cases}$$

Writing $U_{N,k}$ to abbreviate $U(X_{N,k})$, we have

$$E(U_{N'}^* - U_N^*)^2 = \sum_{k=1}^{\infty} E\{U_{N,k} - U_{N',k}\}^2$$

$$+ \sum_{1 \leq k \neq l < \infty} E\{(U_{N',k} - U_{N,k})(U_{N',l} - U_{N,l})\}.$$

For definiteness and ease of exposition take $N' = N + cN^{1/2}$ where $c > 0$ and $cN^{1/2}$ an integer. Observe that

$$\begin{aligned}
E(U_{N',k} - U_{N,k})^2 &= \Pr\{[U_{N',k} - U_{N,k}]^2 = 1\} \\
&= \Pr\{U_{N,k} = 0, U_{N',k} = 1\} + \Pr\{U_{N,k} = 1, U_{N',k} = 0\} \\
&= \Pr\{U_{N,k} = 0\} \Pr\{U_{cN^{1/2},k} = 1\} + \Pr\{U_{N,k} = 1\} \Pr\{U_{cN^{1/2},k} = 1\} \\
&= \tfrac{1}{2}(1 + (1 - 2p_k)^N)\tfrac{1}{2}(1 - (1 - 2p_k)^{cN^{1/2}}) \\
&\qquad + \tfrac{1}{2}(1 - (1 - 2p_k)^N)\tfrac{1}{2}(1 - (1 - 2p_k)^{cN^{1/2}}) \\
&= \tfrac{1}{2}[1 - (1 - 2p_k)^{cN^{1/2}}].
\end{aligned}$$

In a similar manner for $k \neq l$ by conditioning on the values of $U_{N,k}$ and $U_{N,l}$ we obtain

$$\begin{aligned}
E\{(U_{N',k} - U_{N,k})(U_{N',l} - U_{N,l})\} &= \Pr\{U_{N,k} = U_{N,l} = 1, U_{N',k} = U_{N',l} = 0\} \\
&\quad + \Pr\{U_{N,k} = U_{N,l} = 0, U_{N',k} = U_{N',l} = 1\} \\
&\quad - \Pr\{U_{N,k} = U_{N',l} = 1, U_{N',k} = U_{N,l} = 0\} \\
&\quad - \Pr\{U_{N,k} = U_{N',l} = 0, U_{N',k} = U_{N,l} = 1\} \\
&= I_1 + I_2 - I_3 - I_4 .
\end{aligned} \tag{56}$$

A straightforward calculation with the aid of Bayes rule gives

$$\begin{aligned}
I_1 &= (\tfrac{1}{4})^2\{1 - (1 - 2p_k)^N - (1 - 2p_l)^N + (1 - 2p_k - 2p_l)^N\} \\
&\qquad \times \{1 - (1 - 2p_k)^{cN^{1/2}} - (1 - 2p_l)^{cN^{1/2}} + (1 - 2p_k - 2p_l)^{cN^{1/2}}\} \\
I_2 &= (\tfrac{1}{4})^2\{1 + (1 - 2p_k)^N + (1 - 2p_l)^N + (1 - 2p_k - 2p_l)^N\} \\
&\qquad \times \{1 - (1 - 2p_k)^{cN^{1/2}} - (1 - 2p_l)^{cN^{1/2}} + (1 - 2p_k - 2p_l)^{cN^{1/2}}\}
\end{aligned}$$

and corresponding expressions for I_3 and I_4. Combining and simplifying we get

$$\begin{aligned}
E[(U_{N'}^* - U_N^*)^2] &= \frac{1}{4} \sum_{k=1}^{\infty} [1 - (1 - 2p_k)]^{cN^{1/2}} \\
&\quad + \sum_{1 \le k \ne l < \infty} \tfrac{1}{4}(1 - 2p_k - 2p_l)^N \\
&\quad \times [1 - (1 - 2p_l)^{cN^{1/2}} - (1 - 2p_k)^{cN^{1/2}} + (1 - 2p_l - 2p_k)^{cN^{1/2}}].
\end{aligned} \tag{57}$$

The second sum of (57) can be estimated above by

$$\Gamma_N = C\left(\sum_{k=1}^{\infty} e^{-p_k N}(1 - e^{-cN^{1/2}p_k})\right)^2$$

where C is an appropriate constant independent of N. In the course of the proof of Lemma 5′ we established the order relation

$$\Gamma_N = o(M(N)). \tag{58}$$

Furthermore, note that

$$\frac{1}{2}\sum_{k=1}^{\infty}[1-(1-2p_k)^{cN^{1/2}}] \leqq \frac{M(2cN^{1/2})}{2} \sim C'N^{\gamma/2}L(N^{1/2}) = o(M(N)). \tag{59}$$

By virtue of (58) and (59) the assertion of (55) follows. The proof of Lemma 6′ is complete.

Lemma 7′. *$E(U^*_{N(t)} - U^*_{[t]}) \to 0$ as $t \to \infty$.*

Proof. Identical to that of Lemma 7.

With Lemmas 5′–7′ in hand, the proof of Theorem 6 is accomplished paralleling the proof of Theorem 4.

By tedious considerations via the method of moments we can establish asymptotic normality for U^*_N with no restrictions on $\{p_k\}$.

The growth behavior of the random variable C^*_N can be handled by techniques paraphrasing those used in the proof of Theorem 6. The asymptotic relation

$$\frac{\left(\sum_{k=1}^{\infty} Np_k(1-p_k)^N(1-(1-p_k)^{N^{1/2}})\right)^2}{\sum_{k=1}^{\infty} Np_k(1-p_k)^N} \to 0 \quad \text{as} \quad N \to \infty$$

is used. We state the result and omit the detailed proof.

Theorem 7. *Let $\{p_k\}_{k=1}^{\infty}$ satisfy the conditions of Theorem 4. Then for $a_\gamma = \gamma\Gamma(2-\gamma)(1-2^{\gamma-2})$ we have*

$$\frac{C^*_N - E(C^*_N)}{(a_\gamma)^{1/2}N^{\gamma/2}(L(N))^{1/2}} \xrightarrow{Law} Normal\ (mean\ 0, var\ 1).$$

5. Some strong laws. The principal objective of this section is to establish a strong law for the variables Z^*_N. We will also discuss almost sure convergence of the variables $Z_{N,r}/E(Z_{N,r})$, $C^*_N/E(C^*_N)$, and $U^*_N/E(U^*_N)$.

We begin by examining the validity of the weak law for the variable Z^*_N. Inspection of (39) reveals that

$$\operatorname{Var} Z^*_N \leqq \sum_{k=1}^{\infty}[(1-p_k)^N - (1-p_k)^{2N}] \leqq M_N$$

and therefore

$$E\left(\left(\frac{Z^*_N}{M_N} - 1\right)^2\right) \leqq \frac{1}{M_N} \tag{60}$$

which trivally implies

(61) $$\frac{Z_N^*}{M_N} \to 1 \qquad \text{in probability}$$

because $M_N \to \infty$ as $N \to \infty$.

The result of (61) was pointed out by Bahadur [1] using essentially the above argument. A strengthening to probability one convergence carrying no limitations on the sequence $\{p_k\}$ is the substance of the next theorem.

Theorem 8. *Let $M_N = E(Z_N^*)$. Then as $N \to \infty$, $Z_N^*/M_N \to 1$ w.p. 1 (with probability 1).*

Proof. It is convenient to extend M_N by linear interpolation to M_t defined for $t \geqq 0$ with $M_0 = 0$. Obviously, M_t is strictly increasing piecewise linear and tends to $+\infty$ as $t \to \infty$. Determine t_n such that $M_{t_n} = n^2$.

Because of (60) and since M_t is monotonic we have

(62) $$E\left(\frac{Z_{[t_n]+1}^*}{M_{[t_n]+1}} - 1\right)^2 \leqq \frac{1}{M_{[t_n]+1}} \leqq \frac{1}{M_{t_n}} = \frac{1}{n^2}.$$

It follows that

$$\sum_{n=1}^{\infty} E\left(\frac{Z_{[t_n]+1}^*}{M_{[t_n]+1}} - 1\right)^2 < \infty$$

from which we may conclude that

$$\frac{Z_{[t_n]+1}^*}{M_{[t_n]+1}} \to 1 \qquad \text{w.p. 1.}$$

But $|Z_{m+1}^* - Z_m^*| \leqq 1$ for all integer m and consequently

(63) $$\frac{Z_{[t_n]}^*}{M_{t_n}} \to 1 \qquad \text{w.p. 1.}$$

For each positive integer m we specify n depending on m satisfying

$$[t_n] \leqq t_n \leqq m < t_{n+1} \leqq [t_{n+1}] + 1.$$

Since Z_N^* is monotonic in N it follows that

(64) $$Z_{[t_n]}^* \leqq Z_m^* \leqq Z_{[t_{n+1}]+1}^*$$

and therefore

(65) $$\frac{Z_{[t_n]}^*}{M_{t_{n+1}}} \leqq \frac{Z_m^*}{M_m} \leqq \frac{Z_{[t_{n+1}]+1}^*}{M_{t_n}}.$$

These relations and the trivial fact that $M_{t_{n+1}+1}/M_{t_n} \to 1$ for $n \to \infty$ clearly imply on account of (63) that $Z_m^*/M_m \to 1$ as $m \to \infty$ w.p. 1.

Remark 6. The key ingredient underlying the above analysis is the monotonicity property of Z_N^* . Obviously the random variable $Z_{N,r}^*$, the number of occupied cells containing ***at least*** r balls after N tosses, is also monotonic in N. The identical proof as above establishes

$$\frac{Z^*_{N,r}}{E(Z^*_{N,r})} \to 1 \quad \text{w.p. 1.,} \quad N \to \infty. \tag{66}$$

It does not follow that

$$\frac{Z_{N,r}}{E(Z_{N,r})} = \frac{Z^*_{N,r} - Z^*_{N,r+1}}{E(Z^*_{N,r}) - E(Z^*_{N,r+1})} \tag{67}$$

necessarily converges w.p. 1 and counterexamples can readily be constructed.

In the circumstance that $\alpha(x) = x^\gamma L(x)$, $0 < \gamma \leqq 1$ ($L(x)$ slowly varying), then a standard argument taking account of the asymptotic formula for $E(Z_{N,r})$ leads *via* the representation (67) to the conclusion $Z_{N,r}/E(Z_{N,r}) \to 1$ w.p. 1. More generally, we obtain

Theorem 9. *Let $\{p_k\}_{k=1}^\infty$ be such that $\alpha(x) = x^\gamma L(x)$, $0 < \gamma \leqq 1$ where $L(x)$ is slowly varying. Then $W_N/E(W_N) \to 1$, w.p. 1 where W_N is any one of the variables $Z_{N,r}$ (r fixed), U_N^** or *C_N^**.

Proof. A direct proof invokes the Borel Cantelli lemma with the aid of the estimate

$$E\left(\left(\frac{W_N}{E(W_N)} - 1\right)^{2m}\right) \leqq \frac{c_m}{N^{\gamma m}[L(N)]^m} \tag{68}$$

where c_m is an appropriate constant and m is a positive integer determined large enough so that $\gamma m > 1$. The validation of (68) is accomplished with some labor by induction on m. We omit the details.

With regard to the random variable U_N^* a direct computation yields

$$\frac{\text{var}\,(U_N^*)}{[E(U_N^*)]^2} \leq \frac{C}{E(U_N^*)}$$

for some positive constant C independent of N. It follows that (*cf.* (61))

$$\frac{U_N^*}{E(U_N^*)} \to 1$$

in probability as $N \to \infty$. It is reasonable to expect almost sure convergence entailing no restriction on $\{p_k\}$. This assertion is correct; we sketch its proof. First extend $\tilde{M}_N = E(U_N^*)$ by linear interpolation to $\tilde{M}_t$ for $t > 0$. Next, determine t_n such that $\tilde{M}_{t_n} = n^2$. Now we infer exactly as in the proof of Theorem 8 that

$$\frac{U^*_{[t_n]}}{\tilde{M}(t_n)} \to 1 \quad \text{w.p. 1.} \tag{69}$$

Observe from the calculation of Lemma 5′ and the facts that $\tilde{M}(t_{n+1})/\tilde{M}(t_n) \to 1$ and $\tilde{M}(t)$ is increasing that for any m satisfying $t_n \leqq m \leqq t_{n+1}$

$$E\left(\frac{U_m^* - U_{[t_n]}^*}{\tilde{M}(t_n)}\right)^2 \leqq \frac{C}{n^2}$$

for an appropriate absolute constant C. Select for each n an integer m satisfying $t_n \leqq m_n < t_{n+1}$. The above estimate implies

$$\sum_{n=1}^{\infty} E\left(\frac{U_{m_n}^* - U_{[t_n]}^*}{\tilde{M}_{t_n}}\right)^2 < \infty$$

from which we deduce

(70) $$\frac{U_{m_n}^*}{\tilde{M}_{t_n}} - \frac{U_{[t_n]}^*}{\tilde{M}_{t_n}} \to 0 \qquad \text{w.p. 1.}$$

Comparing (69) and (70) we conclude that

$$\frac{U_{m_n}^*}{\tilde{M}_{t_n}} \to 1 \qquad \text{w.p. 1.}$$

It follows that

$$\frac{U_N^*}{E(U_N^*)} = \frac{U_N^*}{\tilde{M}_N} \to 1 \qquad \text{w.p. 1}$$

as claimed.

We conclude this section with a statement of the strong law for the variable $Z_{n(t)}^*$ where $n(t)$, $t \geqq 0$ is a suitable integer valued process.

Theorem 10. *Let $n(t)$ be a positive integer valued stochastic process obeying the strong law $n(t)/t \to 1$ w.p. 1 as $t \to \infty$. Furthermore, suppose $E(n(t)) = t$. Then $Z_{n(t)}^*/M_t \to 1$ w.p. 1. (For the definition of M_t, see the proof of Theorem* 8.)

Proof. Manifestly, by virtue of Theorem 8, we deduce that $Z_{n(t)}^*/M_{n(t)} \to 1$ w.p. 1 as $t \to \infty$ since $n(t) \to \infty$ w.p. 1. Now by virtue of the fact that $M_t - \sum_{i=1}^{\infty} (1 - e^{-p_i t}) \to 0$ as $t \to \infty$ (see (19)), an elementary analysis proves that

$$\overline{\lim_{t\to\infty}} \left(\frac{M_{[t(1+\epsilon)]}}{M_{[t(1-\epsilon)]}}\right) \leqq \frac{1+\epsilon}{1-\epsilon}.$$

With the aid of this inequality it is routinely confirmed that $M_{n(t)}/M_t \to 1$ w.p. 1 as $t \to \infty$. The proof of Theorem 10 is complete.

A corresponding version of Theorem 9 with N replaced by a process $n(t)$ is available.

6. Other limit theorems and applications. Let the conditions of Theorem 4 be in force throughout this section.

I. Let $n(t)$, $t \geqq 0$ (t may be a discrete or continuous variable), denote a stochastic process whose state space is the nonnegative integers. We further postulate that

(71) $$\frac{n(t) - \lambda t}{\sigma t^{1/2}} \text{ converges in law as } t \to \infty \text{ to } \Psi(\cdot),$$

where Ψ is a nondegenerate proper distribution.

Consider the family of random variables $Z^*_{n(t)}$ = the number of occupied cells after $n(t)$ tosses. We claim that

(72) $$\frac{Z^*_{n(t)} - E(Z^*_{[\lambda t]})}{(B_{[\lambda t]})^{1/2}} \xrightarrow{\text{Law}} \text{Normal } (0, 1).$$

Indeed, conditioning on the values of $n(t)$ we obtain

(73) $$\Pr\left\{\frac{Z^*_{n(t)} - E(Z^*_{[\lambda t]})}{(B_{[\lambda t]})^{1/2}} \leq x\right\} = \sum_{k=0}^{\infty} \Pr\{Z^*_k - E(Z^*_{[\lambda t]}) \leq x(B_{[\lambda t]})^{1/2}\} \Pr\{n(t) = k\}.$$

The main contribution to the sum occurs for those values of k satisfying

(74) $$\lambda t - ct^{1/2} \leq k \leq \lambda t + ct^{1/2}$$

for c sufficiently large. With the aid of Lemmas 5 and 6 we infer that for each k satisfying (75)

(75) $$\lim_{t\to\infty} \Pr\{Z^*_k - E(Z^*_{[\lambda t]}) \leq x(B_{[\lambda t]})^{1/2}\} = \Phi(x),$$

where $\Phi(x)$ denotes the standard normal distribution function. A familiar argument using (73) and the result of (75) validates the assertion of (72).

Under the same conditions on $n(t)$, asymptotic normality can be established for $Z_{n(t),r}$, $C^*_{n(t)}$ and $U^*_{n(t)}$ as well.

II. A simple application of the result of Paragraph I is the following. For each integer t, we let $n(t)$ denote the number of tosses required until the first cell contains t balls. Trivially

$$\frac{n(t) - \dfrac{t}{p_1}}{\sigma t^{1/2}} \xrightarrow{\text{Law}} \text{Normal } (0, 1), \qquad \sigma^2 = \frac{1 - p_1}{p_1^2}.$$

Clearly $Z^*_{n(t)}$ is the number of occupied cells at the time when the first cell collects its t^{th} ball. It follows that

$$\frac{Z^*_{n(t)} - E(Z^*_{[\lambda t]})}{(B_{[t]})^{1/2}} \text{ is asymptotically normal,}$$

where $\lambda = (1 - p_1)/p_1$.

As a second application, suppose balls are tossed at the event times of a renewal process with interarrival distribution $F(\xi)$ possessing mean μ and variance σ^2. Let $n(t)$ represent the renewal function which is the number of tosses in time t. Then it is well known that as $t \to \infty$

$$\frac{n(t) - \frac{t}{\mu}}{\mu^{-3/2}\sigma t^{1/2}} \xrightarrow{\text{Law}} \text{standard normal.}$$

The result of paragraph I implies that $Z^*_{n(t)}$ is again asymptotically normal.

Let at each unit of time a random number K^* of balls be tossed at the urns. Let the probability generating function of K^* be

$$f(s) = \sum_{k=1}^{\infty} a_k s^k, \qquad a_k \geqq 0, \qquad f(1) = 1$$

where $f'(1) = \mu$ and $f''(1) + f'(1) - [f'(1)]^2 = \sigma^2 < \infty$. We designate by N^* the random variable representing the number of balls tossed after N units of time. Then

$$\frac{Z^*_{N^*} - E(Z^*_{[\mu N]})}{(B_N)^{1/2}} \xrightarrow{Law} \text{Normal (Mean 0, Var } b)$$

where b is a suitable positive constant depending on f.

Asymptotic normality also obtains for the more general case where the pattern of arrival of balls forms a compound renewal process.

III. We concentrate on the family of variables $Z^*_{n(t)}$ where $N(t)$ is a Poisson process with parameter 1. In Theorem 2 we exposed a number of cases for which $V(t) = \text{Var } Z^*_{N(t)}$ tends to a finite limit as $t \to \infty$.

Extending the method of proof Theorem 2 we can prove

Theorem 11. *Let $\{p_k\}_{k=1}^{\infty}$ be such that $\overline{\lim}_{k\to\infty} p_{k+1}/p_k < 1$. Suppose $\alpha(x)$ (see (3)) satisfies*

(76) $$\lim_{x\to\infty} \frac{1}{x}\int_0^x [\alpha(r\xi) - \alpha((r-1)\xi)]\, d\xi = \delta_r > 0$$

for each integer $r = 2, 3, 4 \cdots$. Then

(77) $$\lim_{t\to\infty} E([Z^*_{N(t)} - E(Z^*_{N(t)})]^m)$$

exists for $m = 1, 2, 3, \cdots$.

Since all the moments of $Z^*_{N(t)} - E(Z^*_{N(t)})$ converge ($t \to \infty$), it is tempting to conclude on the basis of (77) that $Z^*_{N(t)} - E(Z^*_{N(t)})$ converges in law. A rigorous validation requires verification that the limiting moments determine a unique distribution function. Lacking further specific knowledge of the constants $\{\delta_r\}$ makes it impossible to ascertain uniqueness for the relevant moment problem.

However, for the special circumstance

(78) $$p_k = (1-\rho)\rho^{k-1}, \qquad k = 1, 2, 3, \cdots \qquad (0 < \rho < 1)$$

we can deduce that

$$Z^*_{N(t)} - E(Z^*_{N(t)})$$

converges in law to a proper random variable (label it W).

In this case

$$\delta_r = \log_{1/p}\left(\frac{r}{r-1}\right), \qquad r = 2, 3, \cdots .$$

The Carleman criteria for determinateness of the moment problem (see [9, p. 19]) can be checked using the estimate

$$\lim_{t\to\infty} \sum_{k=1}^{\infty} E[\varphi(X_{N(t),k}) - E(\varphi(X_{N(t),k})]^m \leqq (m-1) \log_{1/p}\left(\frac{m}{m-1}\right),$$
$$m = 2, 3, \cdots ,$$

where $\varphi(u) = 1$ for $u > 0$ and $\varphi(u) = 0$ otherwise.

The distribution function of W is difficult to identify. However, in principle we can calculate all its moments. Thus, for example, the first three moments are

$$E(W) = 0, \qquad E(W^2) = \log_{1/p} 2, \qquad E(W^3) = \log_{1/p} \tfrac{9}{8}.$$

Under the conditions of Theorem 11 we can also show that the moments $E(Z_{N(t),r})^m$ ($r > 0$ and fixed), converge as $t \to \infty$ for each $m = 1, 2, 3, \cdots$. Furthermore, in the case of (78), $Z_{N(t),r}$ actually converges in law.

IV. We close the paper with a statement of some unresolved problems.

(i) The limit properties of Z^*_N, $N \to \infty$, when $\alpha(x)$ is slowly varying is in general unsettled. In the special case where $\alpha(x)$ is slowly varying and $\alpha(2x) - \alpha(x) \to \infty$ then it is possible to prove that Z^*_N, appropriately normalized, is asymptotically normal. When $\overline{\lim}_{k\to\infty} p_{k+1}/p_k < 1$ then $\sup_N \operatorname{Var}(Z^*_N) < \infty$. It would be of interest to discern the behavior of the variable $Z^*_N - E(Z^*_N)$ as $N \to \infty$ (*cf.* paragraph III of this section).

(ii) According to the setup of paragraph I of this section, consider the case where

$$\frac{n(t) - At^a}{Bt^b} \xrightarrow{\text{Law}} \Psi(\cdot)$$

for some a and b positive and constants A and B. (A could be zero.) What is now the character of the random variable $Z^*_{n(t)}$ for t large?

References

[1] Bahadur, R. R., On the number of distinct values in a large sample from an infinite discrete distribution, *Proceedings of the National Institute of Sciences of India*, **26A**, Supp. II (1960) 67–75.

[2] Békéssey, A., On classical occupancy problems I and II, *Magyar Tud. Akad. Mat. Kutato Int. Közl*, **8** (1963) 59–71 and **9** (1964) 133–141.

[3] Chistyakov, V. P., On the calculation of the power of the test of empty boxes, *Theory of Probability and its Applications*, **9** (1964).

[4] Darling, D. A., Some limit theorems associated with multinomial trials, *Proceedings of Fifth Berkeley Symposium* (forthcoming).

[5] Doetsch, G., *Theorie und Anwendung der Laplace-Transformation*, Dover Publications, New York, 1943.

[6] Hardy, G. H., *Divergent Series*, Clarendon Press, Oxford, 1949.

[7] Rényi, A., Three new proofs and a generalization of a theorem of Irving Weiss, *Magyar. Tud. Akad. Mat. Kutato. Int. Közl*, **7** (1962) 203–204.

[8] Sevastyanov, B. A. & Christyakov, V. P. Asymptotic normality in the classical ball problem, *Theory of Probability and its Applications*, **9** (1964) 198–211.

[9] Shohat, J. A. & Tamarkin, J. D., *The Problem of Moments*, American Math. Society Publications, New York, 1943.

[10] Spitzer, F., *Principles of Random Walk*, D. Van Nostrand Co., Inc., New York, 1964.

[11] Widder, D. V., *The Laplace Transform*, Princeton University Press, Princeton, N. J., 1946.

Stanford University
Date Communicated: February 20, 1967