The Annals of Probability
1982, Vol. 10, No. 4, 927–934

LIMIT THEOREMS FOR SOME RANDOM VARIABLES ASSOCIATED WITH URN MODELS

BY L. FLATTO

Bell Laboratories

Balls are successively thrown, independently and uniformly, in n given urns. Let $N_{n,m}$ be the number of throws required to obtain at least m balls in each urn. Let $N'_{n,m,r}$ be the number of urns containing exactly r balls upon completion of the $N_{n,m}$th throw, $r \geq m$. We prove that, given $N_{n,m} = [n \log n + (m-1)n \log\log n + nx]$, $N'_{n,m,r} \sim e^{-x}(\log n)^{r-m+1}/r!$ as $n \to \infty$ in probability. From this, we derive the following limit law for the joint distribution of $N_{n,1}, \cdots, N_{n,m}$: $\lim_{n\to\infty} P(N_{n,i} \leq n \log n + (i-1)n \log\log n + nx_i;\ 1 \leq i \leq m) = \prod_{i=1}^{m} \exp(-(1/(i-1)!)e^{-x_i})$. This result generalizes earlier work of Erdös and Renyi who obtained the limit law for $N_{n,m}$ as $n \to \infty$.

1. Introduction. Consider the following classical urn problem. Balls are successively thrown, independently and uniformly, in n given urns labeled $1, \cdots, n$. Let $N_{n,m}$, $1 \leq n, m < \infty$, be the number of throws required to obtain at least m balls in each urn, in which case the urns are said to be covered m times. Erdös and Renyi [1] have proven the following limit law.

THEOREM 1. *Let* $N_{n,m} = n \log n + (m-1)n \log\log n + nX_{n,m}$. *Then* $\lim_{n\to\infty} P(X_{n,m} \leq x) = \exp(-(1/(m-1)!)e^{-x})$.

Using this result, it is possible to derive the asymptotic behavior of the expectation $E(N_{n,m})$.

THEOREM 2. $E(N_{n,m}) = n \log n + (m-1)n \log\log n + C_m n + o(n)$ *as* $n \to \infty$, *where* $C_m = \gamma - \log(m-1)!$, γ *being Euler's constant.*

This result had previously been obtained by Newman and Shepp [5], except for the value of C_m.

The above theorems reveal the rather surprising feature that, up to first order terms, it takes $n \log n$ throws for the first cover, each subsequent cover requiring only an additional $n \log\log n$ throws. To obtain a better understanding of this phenomenon, we derive limit theorems for $N'_{n,m}$, $N''_{n,m}$ conditioned on $N_{n,m}$, where $N'_{n,m}$, $N''_{n,m}$ are respectively defined to be the number of urns containing precisely m balls upon completion of the $N_{n,m}$th throw, and the number of throws past the $N_{n,m}$th one required to obtain at least one more ball in each of $N'_{n,m}$ urns. Thus $N_{n,m+1} = N_{n,m} + N''_{n,m}$ for all n, m.

The following two limit theorems will be proven in Sections 2 and 3. We indicate here how they imply $N''_{n,m} \sim n \log\log n$ in probability as $n \to \infty$, $1 \leq m < \infty$, and the specific form of the limit law of Theorem 1.

THEOREM 3. *Let* $N(n, m, x) = [n \log n + (m-1)n \log\log n + nx]$, $[x]$ *denoting the largest integer* $\leq x$. *Let* $N'_{n,m,r}$, $r \geq m$, *equal the number of urns containing precisely* r *balls upon completion of the* $N_{n,m}$th *throw (In particular* $N'_{n,m,m} = N'_{n,m}$). *For each* $\varepsilon > 0$,

Received November 1980; revised July 1981.
AMS 1970 *subject classification.* Primary 60F99; secondary 60C05.
Key words and phrases. Urn model, limit laws, asymptotic independence.

$$\lim_{n\to\infty} P\left(\left| \frac{N'_{n,m,r}}{\frac{e^{-x}(\log n)^{r-m+1}}{r!}} - 1 \right| > \varepsilon \,\middle|\, N_{n,m} = N(n, m, x) \right) = 0$$

uniformly on every finite interval $a \le x \le b$.

THEOREM 4. *Let* N_n^k *be the number of throws necessary to obtain at least one ball in each of the urns* $1, \cdots, k$ $(k \le n)$, *the balls being thrown independently and uniformly into the urns* $1, \cdots, n$. *Then*

$$\lim_{k\to\infty} P(N_n^k \le n \log k + ny) = \exp(-e^{-y}).$$

Theorem 3 states that, given $N_{n,m} = N(n, m, x)$, $N'_{n,m,r} \sim e^{-x}(\log n)^{r-m+1}/r!$ in probability. This result can be derived heuristically. Let $N = N(n, m, x)$ balls be thrown into n urns. The probability of hitting a specific urn is $1/n$. As $n \to \infty$, the number of balls in a given urn becomes Poisson distributed with parameter $\lambda = N/n$. Hence given $N_{n,m} = N(n, m, x)$, the number of urns with r balls should be $\sim(\lambda^r/r!)e^{-\lambda}n \sim e^{-x}(\log n)^{r-m+1}/r!$. Since $N''_{n,m} = N_n^k$ with $k = N'_{n,m}$, we conclude from Theorems 3 and 4 that $N''_{n,m} \sim n \log \log n$ in probability. Theorems 3 and 4 explain the limit law for $N_{n,m}$. For they readily imply

$$\begin{aligned} &\lim_{n\to\infty} P(N_{n,m+1} \le n \log n + mn \log \log n + n\{y - \log m!\} \mid N_{n,m} = N(n, m, x)) \\ (1.1)\qquad &= \lim_{n\to\infty} P\left(N''_{n,m} \le n \log\left(\frac{e^{-x} \log n}{m!}\right) + ny \,\middle|\, N_{n,m} = N(n, m, x) \right) = \exp(-e^{-y}). \end{aligned}$$

We may drop the condition $N_{n,m} = N(n, m, x)$ in the first line of (1.1). Substituting y for $y - \log m!$, we get Theorem 1.

In Section 3, we derive from Theorems 3 and 4 limit laws (unconditioned) for $N'_{n,m}$, $N''_{n,m}$. We also show that the random variables $X_{n,m}$, $1 \le m < \infty$, are asymptotically independent as $n \to \infty$.

We conclude with another explanation of Theorem 1, though only heuristic. Let $m > 1$. Since $N'_{n,1,m-1}$ will be much larger than $N'_{n,1,r}$, $1 \le r < m - 1$, $N_{n,m}$ should roughly equal $N_{n,1} + N''_{n,1,m-1}$, where $N''_{n,1,m-1}$ is the number of throws past the $N_{n,1}$th one required to obtain at least one additional ball in each of the $N'_{n,1,m-1}$ urns. Theorem 3 then yields

$$\begin{aligned} &\lim_{n\to\infty} P(N_{n,m} \le n \log n + (m-1)n \log \log n + n\{y - \log(m-1)!\} \mid \\ (1.2)\qquad &N_{n,1} = [n \log n + nx]) = \lim_{n\to\infty} P\left(N''_{n,1,m-1} \le n \log\left(\frac{e^{-x}(\log n)^{m-1}}{(m-1)!}\right) \right. \\ &\qquad\qquad \left. + ny \mid N_{n,1} = [n \log n + nx] \right) = \exp(-e^{-y}). \end{aligned}$$

We may drop the condition $N_{n,1} = [n \log n + nx]$ in the first line of (1.2). Substituting y for $y - \log(m-1)!$, we get Theorem 1.

Whether the above argument can be made rigorous is left here as an open problem.

The author thanks L. A. Shepp for several helpful discussions concerning the contents of this paper.

2. Proof of Theorem 3. We derive asymptotic formulas for the expectation and variance of $N'_{n,m,r}$ conditioned on $N_{n,m}$. Chebychev's inequality then proves Theorem 3. Let

$$X_i = \begin{cases} 1, & \text{if urn } i \text{ contains exactly } r \text{ balls upon completion of } N_{n,m}\text{th throw.} \\ 0, & \text{otherwise.} \end{cases}$$

Thus $N'_{n,m,r} = X_1 + \cdots + X_n$. Let $P_{n,m,N} = P(N_{n,m} = N)$. Then

$$(2.1)\qquad E(X_1 \cdots X_\ell \mid N_{n,m} = N) = \frac{1}{P_{n,m,N}} (Q_{n,m,\ell,r,N} + R_{n,m,\ell,r,N})$$

where

$$Q_{n,m,\ell,r,N} = P\begin{pmatrix} N_{n,m} = N;\text{ urns } 1, \cdots, \ell \text{ contain exactly } r \text{ balls upon completion of} \\ N\text{th throw; } N\text{th ball falls in one of the urns } 1, \cdots, \ell \end{pmatrix}$$

$$R_{n,m,\ell,r,N} = P\begin{pmatrix} N_{n,m} = N;\text{ urns } 1, \cdots, \ell \text{ contain exactly } r \text{ balls upon completion of} \\ N\text{th throw; } N\text{th ball falls in one of the urns } \ell + 1, \cdots, n \end{pmatrix}$$

We first derive exact and asymptotic formula for $E(X_1 \cdots X_\ell \mid N_{n,m} = N)$.

LEMMA A.

$$\text{i)} \qquad P_{n,m,N} = \frac{1}{n^{m-1}}\binom{N-1}{m-1}\left(1 - \frac{1}{n}\right)^{N-m} P(N_{n-1,m} \le N - m)$$

$$\text{ii)} \qquad Q_{n,m,\ell,m,N} = \ell \frac{(N-1)!}{(m!)^{\ell-1}(m-1)!(N-\ell m)!}\frac{1}{n^{\ell m}}\left(1 - \frac{\ell}{n}\right)^{N-\ell m} P(N_{n-\ell,m} \le N - \ell m)$$

$$(2.2) \qquad Q_{n,m,\ell,r,N} = 0, \qquad r > m$$

$$\text{iii)} \qquad R_{n,m,\ell,r,N} = (n-\ell)\frac{(N-1)!}{(r!)^{\ell}(m-1)!(N-r\ell-m)!}\frac{1}{n^{r\ell+m}}$$

$$\cdot\left(1 - \frac{\ell+1}{n}\right)^{N-r\ell-m} P(N_{n-\ell-1,m} \le N - r\ell - m).$$

PROOF. i) Let the Nth ball fall in urn i, $1 \le i \le n$, the remaining $m - 1$ balls in urn i being thrown in at the specified times $1 \le N_1 \le N_2 \le \cdots \le N_{m-1} \le N - 1$. The probability of this subevent of $(N_m = N)$ is $(1/n^m)(1 - 1/n)^{N-m}P(N_{n-1,m} \le N - m)$. Since there are $n\binom{N-1}{m-1}$ such subevents, we get the desired formula for $P_{n,m,N}$.

ii) If $N_{n,m} = N$ and the Nth ball falls into urn i, then urn i contains m balls upon completion of the Nth throw. Hence $Q_{n,m,\ell,r,N} = 0$ for $r > m$. To derive the formula for Q in case $r = m$, we argue as follows. Let the Nth ball fall in urn i, $1 \le i \le \ell$. Of the first $N - 1$ balls, urn i gets $m - 1$ balls, each of the urns $1, \cdots, i-1, i+1, \cdots \ell$ gets m balls, the remaining $N - \ell m$ balls falling into urns $\ell + 1, \cdots, n$ and covering these m times. The subevent of $(N_{n,m} = N)$ obtained by specifying i and the times at which the balls are thrown into $1, \cdots, \ell$ has probability

$$\frac{1}{n^{\ell m}}\left(l - \frac{\ell}{n}\right)^{N-\ell m} P(N_{n-\ell,m} \le N - \ell m). \quad \text{There are } \ell \cdot \frac{(N-1)!}{(m!)^{\ell-1}(m-1)!(N-\ell m)!}$$

such subevents, so that we obtain the desired formula.

iii) The reasoning is identical with that of ii) and is omitted. To obtain an asymptotic formula for $E(X_1 \cdots X_\ell \mid N_{n,m} = N)$ we require an estimate on the rate of approach of $P(X_{n,m} \le x)$ to its limit $\exp(-1/(m-1)!e^{-x})$.

LEMMA B. $P(X_{n,m} \le x) = \exp\left(-\frac{e^{-x}}{(m-1)!}\right) + O\left(\frac{\log\log n}{\log n}\right).$

REMARK 1. Lemma B is a special case of a result of Kaplan [See the remark in 3, page 216]. The proof presented here is more elementary.

REMARK 2. The above estimate is understood to hold uniformly on any finite interval $-a \le x \le a$. I.e. for any $a > 0$,

$$\left| P(X_{n,m} \le x) - \exp\left(\frac{-e^{-x}}{(m-1)!}\right) \right| \le C(m, a)\frac{\log\log n}{\log n},$$

$2 \le n < \infty$ and $-a \le x \le a$. $C(m, a)$ is a positive constant depending on m, a but independent of n. This uniform interpretation of the 0-sign is assumed throughout this section. (In formulas (2.8), (2.9) the estimates are uniform in the two variables x, j)

REMARK 3. For $m = 1$, the above error estimate can be improved to $O(\log n/n)$. We leave out the derivation as the estimate of Theorem 2.2 more than suffices for our purpose.

PROOF. Assume, without loss of generality, that for given n, x runs only through values for which $N(n, m, x)$ is an integer. Let $|x| \le a$, and choose n sufficiently large so that $N(n, m, x) > 0$ for $|x| \le a$. Let $N = N(n, m, x)$ and

$P_1 = P$(upon completion of Nth throw, some urn has $m - 1$ balls)

$P_2 = P$(upon completion of Nth throw, some urn has less than $m - 1$ balls)

The event $(X_{n,m} > x)$ means that upon completion of Nth throw, some urn has less than m balls. Hence

$$|P(X_{n,m} > x) - P_1| \le P_2 \le n \sum_{j=0}^{m-2} \binom{N}{j} \frac{1}{n^j}\left(1 - \frac{1}{n}\right)^{N-j}. \tag{2.3}$$

Since $\binom{N}{j} \frac{1}{n^j}\left(1 - \frac{1}{n}\right)^{N} \le \left(\frac{N}{n}\right)^j e^{\frac{-jN}{n}} = O\left(\frac{1}{(\log n)^{m-1-j}}\right)$, we get

$$P(X_{n,m} \le x) = 1 - P_1 + O\left(\frac{1}{\log n}\right). \tag{2.4}$$

By inclusion-exclusion

$$1 - P_1 = \sum_{j=0}^{n} (-1)^j \pi_{n,j} \tag{2.5}$$

where $\pi_{n,j} = \binom{n}{j} \cdot P$(in first N throws, each of the urns $1, \cdots, j$ contains exactly $m - 1$ balls). Thus

$$\pi_{n,j} = \binom{n}{j} \cdot \frac{N!}{((m-1)!)^j (N - j(m-1))!} \frac{1}{n^{j(m-1)}} \left(1 - \frac{j}{n}\right)^{N-j(m-1)}. \tag{2.6}$$

Let $\ell = [\log n/\log \log n]$, $n \ge 3$ and write

$$\begin{aligned} 1 - P_1 - \exp\left\{-\frac{1}{(m-1)!} e^{-x}\right\} &= \left(1 - P_1 - \sum_{j=0}^{\ell-1} (-1)^j \pi_{n,j}\right) \\ &+ \sum_{j=0}^{\ell-1} (-1)^j \left(\pi_{n,j} \frac{-e^{-jx}}{j!((m-1)!)^j}\right) - \sum_{j=\ell}^{\infty} \frac{e^{-jx}}{j!((m-1)!)^j} = A_1 + A_2 + A_3. \end{aligned} \tag{2.7}$$

Taking logarithms of both sides of (2.6), a straightforward though somewhat tedious calculation yields

$$\pi_{n,j} = \frac{e^{-jx}}{j!((m-1)!)^j}\left\{1 + O\left(j \frac{\log \log n}{\log n}\right)\right\}, j \le \ell. \tag{2.8}$$

We conclude from (2.8) and Stirling's formula,

$$\pi_{n,\ell} = O\left(\frac{e^{a\ell}}{\ell!}\right) = O\left(\frac{1}{\sqrt{n}}\right), \quad \text{where} \quad |x| \le a. \tag{2.9}$$

From (2.9) and the inclusion-exclusion inequalities

$$\begin{aligned} 1 - P_1 &\le \sum_{j=0}^{k} (-1)^j \pi_{n,j}, \quad k \quad \text{even} \\ 1 - P_1 &\ge \sum_{j=0}^{k} (-1)^j \pi_{n,j}, \quad k \quad \text{odd} \end{aligned} \tag{2.10}$$

we get

$$|A_1| \le \pi_{n,\ell} = O\left(\frac{1}{\sqrt{n}}\right). \tag{2.11}$$

Similarly,

$$|A_3| = O\left(\frac{e^{a\ell}}{\ell!}\right) = O\left(\frac{1}{\sqrt{n}}\right). \tag{2.12}$$

Inserting (2.8) into the series for A_2, we get

$$|A_2| = O\left(\frac{\log\log n}{\log n}\right). \tag{2.13}$$

Lemma B follows from (2.7) and the estimates (2.11) − (2.13).
We let $N = N(n, m, x)$, $s = r - m$, $A = (N_{n,m} = N)$.

LEMMA C.

$$E(X_1 \cdots X_\ell | A) = \left(\frac{e^{-x}}{r!}\frac{(\log n)^{s+1}}{n}\right)^\ell + O\left(\frac{(\log n)^{(s+1)\ell-1}\log\log n}{n^\ell}\right).$$

PROOF. From (2.2), we have

$$\begin{aligned}\frac{Q_{n,m,\ell,m,N}}{P_N} = \frac{\ell}{(m!)^{\ell-1}}\left\{\frac{(N-m)!}{(N-\ell m)!}\,\frac{1}{n^{(\ell-1)m+1}}\,\frac{\left(1-\frac{\ell}{n}\right)^{N-\ell m}}{\left(1-\frac{1}{n}\right)^{N-m}}\right\}.\\ \frac{P(N_{m,m-\ell} \le N - \ell m)}{P(N_{m,n-1} \le N - m)}.\end{aligned} \tag{2.14}$$

Let B be the quantity inside the parentheses. A computation yields

$$B = \frac{(e^{-x}\log n)^{\ell-1}}{n^\ell}\left(1 + O\left(\frac{\log\log n}{\log n}\right)\right). \tag{2.15}$$

It is readily checked that

$$N(n, m, x) - \ell m = N(n-\ell, m, x') \quad \text{where} \quad x' - x = O\left(\frac{\log n}{n}\right). \tag{2.16}$$

It follows from (2.16) and Lemma B that

$$P(N_{m,n-\ell} \le N - \ell m) = \exp(-e^{-x})\left(1 + O\left(\frac{\log\log n}{\log n}\right)\right). \tag{2.17}$$

We conclude from (2.15), (2.17) that

$$\frac{Q_{n,m,\ell,r,N}}{P_N} = \frac{\ell\left(\frac{e^{-x}\log n}{m!}\right)^{\ell-1}}{n^\ell}\left(1 + O\left(\frac{\log\log n}{\log n}\right)\right). \tag{2.18}$$

Similarly, we derive

$$\frac{R_{n,m,\ell,r,N}}{P_N} = \left(\frac{e^{-x}(\log n)^{s+1}}{r!\,n}\right)^\ell\left(1 + O\left(\frac{\log\log n}{\log n}\right)\right). \tag{2.19}$$

Lemma C follows from (2.1), (2.18), 2.19).

LEMMA D. $\operatorname{Cov}(X_1X_2 \mid A) = O\left(\frac{(\log n)^{2s+1}\cdot\log\log n}{n^2}\right)$.

PROOF. Since X_1, X_2 are identically distributed, $\operatorname{Cov}(X_1X_2 \mid A) = E(X_1X_2 \mid A) - E^2(X_1 \mid A)$. Lemma D follows from Lemma C, upon setting $\ell = 1, 2$.

LEMMA E. $E(N'_{n,m} \mid A) = \frac{e^{-x}}{m!}(\log n)^{s+1} + O(\log\log n)$

$$\sigma^2(N'_{n,m} \mid A) = O((\log n)^{2s+1}\cdot\log\log n).$$

PROOF. As the X_i's are exchangeable random variables, we have $E(N'_{n,m} \mid A) = nE(X_1 \mid A)$

$$\begin{aligned}\sigma^2(N'_{n,m} \mid A) &= n\sigma^2(X_1) + n(n-1)\operatorname{Cov}(X_1X_2 \mid A) \le n\,E(X_1 \mid A)\\ &\quad + n(n-1)\operatorname{Cov}(X_1X_2 \mid A).\end{aligned} \tag{2.22}$$

Lemma E follows from Lemmas C and D.

Finally, Theorem 3 is a direct consequence of Lemma E and an application of Chebychev's inequality.

3. Limit Laws. We derive in this section limit laws for the random variables $N'_{n,m}$, $N''_{n,m}$, and the sequence $\{X_{n,1}, X_{n,2}, \cdots, X_{n,m}\}$ as $n \to \infty$. These limits laws are derived from Theorem 3 and from the following.

THEOREM 4. *Let $1 \le k \le n$, $-a \le x \le a$, where $a > 0$. Let N_n^k be the number of throws necessary to obtain at least one ball in each of the urns $1, \cdots, k$, the balls being thrown independently and uniformly into the urns $1, \cdots, n$. Then* $\lim_{k\to\infty} | P(N_n^k \le n\log k + nx) - \exp(-e^{-x}) | = 0$ *uniformly in n and x.*

PROOF. Let $N = n\log k + nx$. Assume $k \ge k_0$ where $\log k_0 > a$. Thus $N > 0$. Without loss of generality, we may assume that for given n and k, x runs only through those values for which N is an integer.

Let ℓ be a positive integer. For $k > k_0$, write

$$\begin{aligned}P(N_n^k \le N) - \exp(-e^{-x}) &= \left(P(N_n^k \le N) - \textstyle\sum_{j=0}^{\ell} (-1)^j \binom{k}{j}\left(1 - \frac{j}{n}\right)^N\right)\\ &\quad + \textstyle\sum_{j=0}^{\ell} (-1)^j\left(\binom{k}{j}\left(1-\frac{j}{n}\right)^N - \frac{e^{-jx}}{j!}\right) - \sum_{j=\ell+1}^{\infty}\frac{(-1)^j}{j!}e^{-jx} = A_1 + A_2 + A_3.\end{aligned} \tag{3.1}$$

By the inclusion-exclusion inequalities

$$|A_1| \le \binom{k}{\ell+1}\left(1 - \frac{\ell+1}{n}\right)^N \le \frac{k^{\ell+1}}{(\ell+1)!}e^{-\frac{\ell+1}{n}N} = \frac{e^{-x}(\ell+1)}{(\ell+1)}. \tag{3.2}$$

A computation shows that for given j,

$$\begin{aligned}&\log\left\{k(k-1)\cdots(k-j+1)\left(1-\frac{j}{n}\right)^N\right\}\\ &\qquad = -jx + O\left(\frac{\log k}{k}\right) \quad \text{uniformly for} \quad |x| \le a,\ n \ge k.\end{aligned} \tag{3.3}$$

Hence, for fixed j,

$$\lim_{k\to\infty}\binom{k}{j}\left(1-\frac{j}{n}\right)^N = \frac{e^{-jx}}{j!} \quad \text{uniformly for} \quad |x| \le a,\ n \ge k. \tag{3.4}$$

Let $\varepsilon > 0$. Choose ℓ so that $\sum_{j=\ell+1}^{\infty} e^{aj}/j! \leq \varepsilon/3$. Then $|A_3| \leq \varepsilon/3$ and, by (3.2), $|A_1| \leq \varepsilon/3$ for $|x| \leq a$, $n \geq k$. By (3.4) $\exists\, K_\varepsilon > k_0$, $\ell \ni |A_2| < \varepsilon/3$ *for* $|x| \leq a$, $n \geq k > K_\varepsilon$. Hence $|P(N_n^k \leq N) - \exp\{-e^{-x}\}| < \varepsilon$ for $|x| \leq a$, $n \geq k > K_\varepsilon$, thereby proving Theorem 4.

THEOREM 5.

$$\lim_{n\to\infty} P\left(\frac{N'_{n,m}}{\frac{1}{m}\log n} \geq x\right) = e^{-x},\ x > 0$$

PROOF. Let $Z_{n,m} = m!\, e^{X_{n,m}} N'_{n,m}/\log n$. By Theorem 1, $\forall\, \varepsilon > 0\ \exists\, a_\varepsilon > 0 \ni P(|X_{n,m}| > a_\varepsilon) < \varepsilon$ for $1 \leq n < \infty$. Let $x_{n1}, x_{n2}, \cdots$ denumerate all possible values of $X_{n,m}$. Then, for $\delta > 0$,

$$P(|Z_{n,m} - 1| > \delta) \leq P(|X_{n,m}| > a_\varepsilon) + \sum_{|x_{nk}| \leq a_\varepsilon} P(|Z_{n,m} - 1| > \delta \mid X_{n,m} = x_{nk}). \tag{3.5}$$

Let $n \to \infty$. We conclude from Theorem 3 and 3.5) that $\lim\sup_{n\to\infty} P(|Z_{n,m} - 1| > \delta) \leq \varepsilon$. Since $\varepsilon > 0$ is arbitrary, this means $\lim_{n\to\infty} P(|Z_{n,m} - 1| > \delta) = 0$. I.e., for given m, $Z_{n,m}$ converges to 1 in probability as $n \to \infty$. Since $N'_{n,m}/((1/m)\log n) = Z_{n,m}\, e^{-X_{n,m}}/(m-1)!$ and $Z_{n,m} \to 1$ in probability, we have

$$\begin{aligned} \lim_{n\to\infty} P\left(\frac{N'_{n,m}}{\frac{1}{m}\log n} \geq x\right) &= \lim_{n\to\infty} P\left(\frac{e^{-X_{n,m}}}{(m-1)!} \geq x\right) \\ &= \lim_{n\to\infty} P(X_{n,m} \leq -\log(m-1)!\, x) = e^{-x}. \end{aligned} \tag{3.6}$$

THEOREM 6. *Let* $1 \leq m < \infty$ and $x_1, \cdots, x_m$ arbitrary real numbers. $\lim_{n\to\infty} P(X_{n,1} \leq x_1, \cdots, X_{n,m} \leq x_m) = \prod_{\iota=1}^{m} \lim_{n\to\infty} P(X_{n,\iota} \leq x_i)$. I.e., $X_{n,1}, \cdots, X_{n,m}$ are asymptotically independent as $n \to \infty$.

PROOF. We assume the results holds for m and show that it then holds for $m + 1$. Let $X_n = (X_{n,1}, \cdots, X_{n,m})$, $x = (x_1, \cdots, x_m)$. By $X_n \leq x_i$, we mean $X_{n,\iota} \leq x_\iota$, $1 \leq 1 \leq m$.

Let $\varepsilon, \delta > 0$. Define $Z_{n,m}$ as above. We have shown in the proof of Theorem 5 that $\exists\, N_{\iota,\delta} \ni P(|Z_{n,m} - 1| > \delta) < \varepsilon$ whenever $n > N_{\iota,\delta}$. Hence for all $x_1, \cdots, x_{m+1}$

$$\begin{aligned} &P(X_n \leq x, X_{n,m+1} \leq x_{m+1}) - P(X_n \leq x, X_{n,m+1} \leq x_{m+1}, 1 - \delta \\ &\qquad \leq Z_{n,m} \leq 1 + \delta)| < \varepsilon \quad \text{for} \quad n > N_{\varepsilon,\delta}. \end{aligned} \tag{3.7}$$

Let $\alpha = (\alpha_1, \cdots, \alpha_m)$ and β denote respectively the values attained by X_n, $Z_{n,m}$. Then

$$\begin{aligned} &P(X_n \leq x, X_{n,m+1} \leq x_{m+1}, 1 - \delta \leq Z_{n,m} \leq 1 + \delta) \\ &\quad = \sum_{\alpha \leq x, |\beta - 1| < \iota} P(X_{n,m+1} \leq x_{m+1} \mid X_n = \alpha, Z_{n,m} = \beta) \cdot P(X_n = \alpha, Z_{n,m} = \beta). \end{aligned} \tag{3.8}$$

Since $N''_{n,m} = N_{n,m+1} - N_{n,m} = n \log\log n + n(X_{n,m+1} - X_{n,m})$, we have

$$\begin{aligned} &P(X_{n,m+1} \leq x_{m+1} \mid X_n = \alpha, Z_{n,m} = \beta) \\ &\quad = P(N''_{n,m} \leq n \log\log n + n(x_{m+1} - \alpha_m) \mid X_n = \alpha, N'_{n,m} = k) \end{aligned} \tag{3.9}$$

where $k = k(\alpha, \beta, n) = \dfrac{\beta}{m!} e^{-\alpha} \log n$.

$N''_{n,m}$ is the number of throws required to place at least one ball in each of the $N'_{n,m}$ urns. It follows that the right side of (3.9) $= P(N_n^k \leq n \log\log n + n(x_{m+1} - \alpha_m)) = P(N_n^k \leq n \log k + n(x_{m+1} + \log(m!/\beta)))$, N_n^k being defined as in Theorem 4. Inserting into (3.8), we get

$$P(X_n \le x, X_{n,m+1} \le x_{m+1}, 1 - \delta \le Z_{n,m} \le 1 + \delta) \tag{3.10}$$
$$= \textstyle\sum_{\alpha \le x, |\beta - 1| \le \varepsilon} P\left(N_n^k \le n \log k + n\left(x_{m+1} + \log \frac{m!}{\beta}\right)\right) \cdot P(X_n = \alpha, Z_{n,m} = \beta).$$

Write

$$P\left(N_n^k = n \log k + n\left(x_{m+1} + \log \frac{m!}{\beta}\right)\right)$$
$$= \left\{P\left(N_n^k = n \log k + n\left(x_{m+1} + \log \frac{m!}{\beta}\right)\right) - \exp\left(\frac{-\beta e^{-x_{m+1}}}{m!}\right)\right\} \tag{3.11}$$
$$+ \left\{\exp\left(\frac{-e^{-\beta x_{m+1}}}{m!}\right) - \exp\left(\frac{-e^{-x_{m+1}}}{m!}\right)\right\} + \exp\left(\frac{-e^{-x_{m+1}}}{m!}\right)$$

and split $\sum$ accordingly into $\sum_1 + \sum_2 + \sum_3$. Thus

$$\textstyle\sum_3 = \exp\left(\frac{-e^{-x_{m+1}}}{m!}\right) \cdot P(X_n \le x, 1 - \delta \le Z_{n,m} \le 1 + \delta).$$

For given $\varepsilon > 0$, choose $\delta = \delta_\varepsilon < ½\ 3\ |\exp(-e^{-\beta x_{m+1}}/m!) - \exp(-e^{-x_{m+1}}/m!)| < \varepsilon$ whenever $|\beta - 1| < \delta_\varepsilon$. $\mathrm{Min}_{\alpha,\beta} k(\alpha, \beta, n) \ge 1/2m!\, e^{-x_m} \log n \to \infty$ as $n \to \infty$. We may therefore apply Theorem 4 to get $\lim_{n\to\infty} \sum_1 = 0$. It follows from (3.10), (3.11) that

$$\limsup_{n\to\infty} P(X_n \le x, X_{n,m+1} \le x_{m+1}, 1 - \delta_\varepsilon \le Z_{n,m} \le 1 + \delta_\varepsilon) \tag{3.12}$$
$$- \exp\left(\frac{-e^{x_{m+1}}}{m!}\right) P(X_n \le x, 1 - \delta_\varepsilon \le Z_{n,m} \le 1 + \delta_\varepsilon) | \le \varepsilon.$$

(3.7) and (3.12) yield

$$\limsup_{n\to\infty} \left| P(X_n \le x, X_{n,m+1} \le x_{m+1}) - \exp\left(\frac{-e^{-x_{m+1}}}{m!}\right) P(X_n \le x) \right| \le 3\varepsilon. \tag{3.13}$$

Since $\varepsilon > 0$ is arbitrary, we conclude from (3.13) the desired result

$$\lim_{n\to\infty} P(X_{n1} \le x_1, \cdots, X_{n,m+1} \le x_{m+1}) = \textstyle\prod_{i=1}^{m+1} \exp\left(\frac{-e^{-x_i}}{(i-1)!}\right).$$

COROLLARY. $\lim_{n\to\infty} P(N''_{n,m} \le n \log \log n + nx) = F_m(x)$ where $F_m(x) = \exp\{(1/(m - 1)!)e^{-x}\} * (1 - \exp\{(1/m!)e^{-x}\})$.

PROOF. Let $N_{n,m} = n \log \log n + n\, Y_{n,m}$. By Theorem 3.3 $(X_{n,m}, X_{n,m+1}) \to_{\mathscr{L}} (X_1, X_2)$ as $n \to \infty$, X_1 and X_2 being independent random variables whose respective distribution functions are $\exp\{(1/(m - 1)!)e^{-x}\}$, $\exp\{(1/m!)e^{-x}\}$. Hence $Y_{n,m} = X_{n,m+1} - X_{n,m} \to_{\mathscr{L}} X_1 - X_2$ as $n \to \infty$. Since $P(X_1 - X_2 \le x) = F_m(x)$, we get $\lim_{n\to\infty} P(N''_{n,m} \le n \log \log n + nx) = \lim_{n\to\infty} P(Y_{n,m} \le x) = F_m(x)$.

REFERENCES

[1] ERDÖS, P. and RENYI, A. (1961). On a classical problem in probability theory. *Magyar Tud. Akad. Kutato Int. Kozl.* **6** 215–219.

[2] JOHNSON, N. L. and KOTZ (1977). *Urn Models and Their Applications.* Wiley, New York.

[3] KAPLAN, N. (1977). A generalization of a result of Erdös and Renyi. *J. Appl. Probability* **14** 212–216.

[4] KOLCHIN, V. F., SEVAST'YANOV-CHISTYAKOV (1978). *Random Allocations.* (translation). Wiley, New York.

[5] NEWMAN, D. J. and SHEPP, L. (1960). The double dixie cup problem. *Amer. Math. Monthly* **67** 58–61.

BELL LABORATORIES
MURRAY HILL, NEW JERSEY 07974