ON CLASSICAL OCCUPANCY PROBLEMS I.

by

A. BÉKÉSSY

1. Introduction Let there be n "cells" or "urns", and suppose that N "balls" are thrown in the cells independently of each other with the same probability $1/n$. As a result, there will be cells occupied by 0, 1, 2, . . . balls resp. Denoting the number of cells, which contain exactly m balls by $\xi(n, N, m)$, an "occupancy problem" is to determine the probability distribution of the random variables ξ. This well known problem is treated and the probabilites in question are computed by elementary combinatorial methods e. g. in W. FELLER's book [1]. However, the expressions for these probabilities are rather inconvenient, so that various authors worked on determining the corresponding limit distributions. In this respect there are some earlier results due to R. von MISES [3] and to S. M. BERNSTEIN [2]. Recently, I. WEISS [4] has proved that supposing $n \to \infty$, $N \to \infty$, $N/n =$ const., the number of the empty cells, more precisely the *standardized* variables $\xi(n, N, 0)$ tends to be normally distributed. Moreover, F. N. DAVID and D. E. BARTON showed that the same is true more generally for $\xi(n, N, m)$. Their results are summarized in [10]. In a recent paper [5] A. RÉNYI generalized the theorem concerning the normal limit case in an other direction, he proved namely that the condition $N/n =$ const. is not necessary, the sufficient (and at the same time necessary) condition being $D \to \infty$, where

$$D^2 = ne^{-\alpha}[1 - (1 + \alpha)\, e^{-\alpha}],$$

$$\alpha = \frac{N+1}{n},$$

but his paper deals with the special case $m = 0$ only. The purpose of the present paper is to generalize RÉNYI's result to $m \neq 0$, or with other words, to extend the corresponding theorem of DAVID and BARTON.

Theorem. *If both n and N tend to infinity and $\alpha = (N + 1)/n$ is restricted by*

$$n\, \alpha^m e^{-\alpha} \to \infty \tag{1}$$

(since α may be eventually unbounded), whereas $m =$ const., $x =$ const. and $m \geq 2$, then the asymptotic relation

$$\mathbf{P}\left\{\frac{\xi(n, N, m) - E(n, N, m)}{D(n, N, m)} < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \tag{2}$$

holds with

$$E(n, N, m) = n\,\frac{\alpha^m e^{-\alpha}}{m!} \tag{3}$$

and

$$D^2(n, N, m) = n\,\frac{\alpha^m e^{-\alpha}}{m!}\left[1 - \frac{\alpha^m e^{-\alpha}}{m!}\left(1 + \frac{(\alpha - m)^2}{\alpha}\right)\right]. \tag{4}$$

As for $m = 0$ and $m = 1$, the same is true with the only difference that for α not bounded from below the condition

$$n\,\alpha^2 \to \infty \tag{5}$$

is necessary (instead of (1)).

The conditions (1) and (5) are equivalent to $D(n, N, m) \to \infty$.

If $D^2(n, N, m) \to \gamma \neq 0$, then the corresponding random variables are distributed on the limit according to Poisson's law. In order to give a more detailed description, if $n\,\alpha^m\, e^{-\alpha}/m! \to \gamma \neq 0$ because α tends to infinity, then

$$\mathbf{P}\{\xi(n, N, m) = k\} \to \frac{\gamma^k e^{-\gamma}}{k!} \tag{6}$$

with no restriction on m, (see e. g. [1], [3], [8]), but if $n\alpha^m\, e^{-\alpha}/m! \to \gamma$ because α tends to zero, then (6) holds only for $m \geqq 2$, and for the exceptional cases we have

$$\mathbf{P}\{\xi(n, N, 0) - n + N = k\} \to \frac{(\gamma/2)^k e^{-\gamma/2}}{k!} \tag{7}$$

and

$$\mathbf{P}\left\{\frac{N - \xi(n, N, 1)}{2} < x\right\} \to \sum_{k=0}^{[x]} \frac{(\gamma/2)^k e^{-\gamma/2}}{k!} \tag{8}$$

where $\gamma = \lim n\alpha^2$.

Remark. It follows from (8) that the probability of $N - \xi(n, N, 1)$ being an *even* number tends to 1. This somewhat queer phenomenon may be roughly explained as follows. The variable $N - \xi(n, N, 1)$ is the number of balls placed in cells containing more than one ball. In the case $\lim n\alpha^2 < \infty$ there are relatively few balls at all, therefore with probability tending to 1 the small set of cells containing more than one ball consists of cells having exactly two ones.

2. The G-functions of the ξ's. In the course of proving WEISS'theorem, A. RÉNYI has found that a certain generating function of the characteristic functions of $\xi(n, N, 0)$ turns to be a very simple expression. In fact, denoting the characteristic function by $\Phi(n, N, 0, t)$, the G-function

$$G(n, z, 0, t) = \sum_{N=0}^{\infty} \Phi(n, N, 0, t)\,\frac{(nz)^N}{N!}$$

is simply $(e^z + e^{it} - 1)^n$. Similarly, the G-function of the variables $\xi(N, n, m)$ is relatively simple,

$$G(n, z, m, t) = \left[e^z + (e^{it} - 1)\frac{z^m}{m!}\right]^n, \tag{9}$$

where $G(n, z, m, t)$ is defined as

$$\sum_{N=0}^{\infty} \mathbf{E}\,\{e^{it\xi(n,N,m)}\}\frac{(nz)^N}{N!}. \tag{10}$$

In order to prowe (9), the starting point will be the joint distribution of the variables ξ (for various m). Denote $p(n, N, k_0, k_1, \ldots k_m)$ the probability of the event that after having distributed N balls, the number of the cells containing 0, 1, 2, . . ., m balls is $k_0, k_1, k_2, \ldots, k_m$ resp. For the probabilities $p(n, N, k_0, k_1, \ldots, k_m)$ the following recurrence relation holds:

$$\left\{\begin{aligned} p(n, N+1, k_0, k_1, \ldots, k_m) &= p(n, N, k_0, \ldots, k_m)\frac{n - k_0 - \ldots - k_m}{n} + \\ &+ p(n, N, k_0 + 1, k_1 - 1, \ldots, k_m)\frac{k_0 + 1}{n} + \\ &+ p(n, N, k_0, k_1 + 1, k_2 - 1, \ldots, k_m)\frac{k_1 + 1}{n} + \ldots \\ &\ldots + p(n, N, k_0, \ldots, k_{m-1} + 1, k_m - 1)\frac{k_{m-1} + 1}{n} + \\ &+ p(n, N, k_0, \ldots, k_{m-1}, k_m + 1)\frac{k_m + 1}{n}, \end{aligned}\right. \tag{11}$$

expressing the change in the probability of the event characterized by the numbers $k_0, k_1, \ldots, k_m$ after having thrown one more ball. Denoting the characteristic function by $\Phi\,(n, N, t_0, t_1, \ldots, t_m)$ and the G-function

$$\sum_{N=0}^{\infty} \Phi(n, N, t_0, t_1, \ldots, t_m)\frac{(nz)^N}{N!}$$

by $G(n, z, t_0, \ldots, t_m)$, we obtain from (11) the partial differential equation

$$\left\{\begin{aligned} \frac{1}{n}\frac{\partial G}{\partial z} &= G + \frac{1}{n}(e^{i(t_1 - t_0)} - 1)\frac{\partial G}{i\,\partial t_0} + \\ &+ \frac{1}{n}(e^{i(t_2 - t_1)} - 1)\frac{\partial G}{i\,\partial t_1} + \ldots \\ &\ldots + \frac{1}{n}(e^{i(t_m - t_{m-1})} - 1)\frac{\partial G}{i\,\partial t_{m-1}} + \\ &+ \frac{1}{n}(e^{-it_m} - 1)\frac{\partial G}{i\,\partial t_m}, \end{aligned}\right. \tag{12}$$

which may be solved e. g. by the method of characteristics, and the solution satisfying the suitable initial condition $G(n, 0, t_0, \ldots, t_m) = e^{int_0}$ is

$$G(n, z, t_0, t_1, \ldots, t_m) = \left[e^z + \sum_{\mu=0}^{m} \frac{z^\mu}{\mu!} (e^{it_\mu} - 1) \right]^n . \tag{13}$$

From (13) the equation (9) immediately follows.

Probabilities of various events are easily derivable from (9) or (13) by differentiating; e. g. the probability of having k cells, each occupied by m balls is (see [1])

$$\frac{1}{n^N k!} \frac{\partial^{N+k}}{\partial z^N \partial x^k} \left(e^z + (x-1) \frac{z^m}{m!} \right)^n \Bigg|_{\substack{x=0\\z=0}} =$$

$$= \frac{(-1)^k n! \, N!}{n^N k!} \sum_{s=k}^{n} \frac{(-1)^s}{(s-k)! \, (n-s)!} \cdot \frac{(n-s)^{N-ms}}{(m!)^s (N-ms)!}$$

(Put $(a\,!)^{-1}$ equal to zero for $a < 0$).

3. Expectation and variance. By differenciating (9), we have for the expectation $\mathbf{E}\{\xi(n, N, m)\}$ and for the quadratic moment $\mathbf{E}\{\xi^2(n, N, m)\}$

$$\left\{ \mathbf{E}\{\xi(n, N, m)\} = \frac{1}{n^N} \frac{\partial^{N+1}}{i \partial t \, \partial z^N} G(n, z, m, t)_{\substack{t=0\\z=0}} = \frac{n}{n^N} \binom{N}{m} (n-1)^{N-m} , \right. \tag{14}$$

and

$$\left\{ \begin{aligned} &\mathbf{E}\{\xi^2(n, N, m)\} = -\frac{1}{n^N} \frac{\partial^{N+2}}{\partial t^2 \, \partial z^N} G(n, z, m, t)_{\substack{t=0\\z=0}} = \\ &= \frac{n}{n^N} \binom{N}{m} (n-1)^{N-m} + \frac{n(n-1)}{n^N} \binom{N}{m} \binom{N-m}{m} (n-2)^{N-2m} . \end{aligned} \right. \tag{15}$$

Supposing $N \to \infty$, $n \to \infty$ and $N/n^2 \to 0$, it follows from (14) that asymptotically

$$\mathbf{E}\{\xi(n, N, m)\} \sim n \frac{\alpha^m e^{-\alpha}}{m\,!} = E(n, N, m) , \tag{16}$$

(where $\alpha = (N+1)/n$). With more difficulties, — altough the computation is quite elementary — under the same conditions as former one has

$$\begin{aligned} &\mathbf{D}^2\{\xi(n, N, m)\} = \mathbf{E}\{\xi^2(n, N, m)\} - \mathbf{E}\{\xi(n, N, m)\}^2 \sim \\ &\sim \frac{n \alpha^m e^{-\alpha}}{m\,!} \left[1 - \frac{\alpha^m e^{-\alpha}}{m\,!} \left(1 + \frac{(\alpha - m)^2}{\alpha} \right) \right] = D^2(n, N, m) . \end{aligned} \tag{17}$$

The difficulty said above arises because of the fact that $\mathbf{E}\{\xi^2\}$ is asymptotically equal to $\mathbf{E}\{\xi\}^2$ and $\mathbf{D}^2\{\xi\}$ being their difference, one must take account of asymptotic terms of lower order in $\mathbf{E}\{\xi\}$ and $\mathbf{E}\{\xi\}^2$ — too.

The rate $\mathbf{D}^2/\mathbf{E}$ tends to 1 for $\alpha \to 0$, but only if $m \geqq 2$, whereas for $m = 0$

$$\frac{D^2}{E} \sim \frac{\alpha^2}{2},$$

resp. for $m = 0$

$$\frac{D^2}{E} \sim 2\alpha .$$

This peculiarity shows that for $\alpha \to 0$ the asymptotic behaviour of the distribution of $\xi(n, N, 0)$ resp. $\xi(n, N, 1)$ will eventually be different from that of $\xi(n, N, m)$ with $m \geqq 2$.

4. Preliminary remarks concerning the proof of the theorem. It follows from (9) by Cauchy's formula that the characteristic function of $\xi(n, N, m)$ is

$$\Phi(n, N, m, t) = \frac{1}{2\pi i} \cdot \frac{N!}{n^N} \cdot \oint \left[e^z + (e^{it} - 1) \frac{z^m}{m!} \right]^n z^{-N-1} dz , \tag{18}$$

where the path of integration may be taken to be along a circle about the point $z = 0$. Having the characteristic function in the integral form (18), its asymptotic behaviour can be effectively analysed by Riemann's and Debye's method of steepest decents well known in the analysis [9]. On the other hand, according to the continuity theorem, the asymptotics of the characteristic function and that of the distribution correspond to each other. In the present case however, saddle points on the z plane, necessary for employing Debye's method, are not real for real values of the argument t. In order to avoid the inconvenience involved with complex-valued saddle points, we regard in what follows the variable t as pure imaginary, in which case it can be shown that there exists at least one saddle point b on the real positive z axis. As for the validity of the continuity theorem, I. H. CURTISS has shown [7] that it is sufficient to analyse the behaviour of $\psi(x) = \Phi(-ix)$ for real x's in the both sided neighbourhood of $x = 0$. More precisely, in the sense of Curtiss' theorem it is sufficient to show that for x real and constant the asymptotic relation

$$\Psi \left(n, N, m, \frac{x}{D} \right) \cdot \exp \left\{ - \frac{xE}{D} \right\} \sim e^{x^2/2}$$

is valid (since we expect normal limit distribution), where E and D have the values (3) and (4) resp. Hence it is to be proved that

$$\frac{N!}{n^N} \frac{1}{2\pi i} \oint \left(e^z + p \frac{z^m}{m!} \right)^n z^{-N-1} \, dz \sim \exp \left\{ \frac{xE}{D} + \frac{x^2}{2} \right\} \tag{19}$$

(where p denotes $e^{\frac{x}{D}} - 1$ as abbreviation) under the restrictions imposed upon x, n, N and α previously.

The saddle points b, defined as values of z, for which

$$\frac{\partial}{\partial z} \left[\left(e^z + p \frac{z^m}{m!} \right)^n z^{-N-1} \right] = 0 ,$$

are now determined by the equation

$$b = \alpha + (\alpha - m)\, p \frac{b^m e^{-b}}{m!}. \tag{20}$$

There lies on the real positive z axis at least one point b satisfying equation (20). (The later estimations will show that for sufficiently large n only one positive b exists.)

Considering the integral (19), let us take the path of integration through the saddle point b. Putting $z = bw$ we have

$$\Psi\left(n, N, m, \frac{x}{D}\right) = F.J\,, \tag{21}$$

where

$$F = \frac{N!}{n^N} \frac{1}{\sqrt{2\pi N}} \cdot \left(e^b + p\frac{b^m}{m!}\right)^n b^{-N} \tag{22}$$

and

$$J = \frac{\sqrt{N}}{i\sqrt{2\pi}} \oint \left[\frac{e^{b(w-1)} + pw^m \dfrac{b^m e^{-b}}{m!}}{1 + p\dfrac{b^m e^{-b}}{m!}} \right]^n w^{-N-1}\, dw\,. \tag{23}$$

The factor F in (21) is $\left(\text{apart from a factor } \dfrac{N!}{n^N}\dfrac{1}{\sqrt{2\pi N}}\right)$ the value of the integrand in the saddle point b, and as a matter of fact, the asymptotical behaviour of ψ is wholly governed by F, because the other factor J tends to 1 in all cases considered later on.

The proof of the theorem (including also the "Poisson-cases") may be conducted in three steps these beeing the analysis of the behaviour of the saddle point b, the factor F and the integral J respectively.

5. The expansion of b. Let us suppose first $D \to \infty$, then $p = o(1)$, more precisely

$$p = e^{x/D} - 1 = O\left(\frac{e^{\alpha/2}}{\alpha^{m/2}\sqrt{n}}\right) + O\left(\frac{1}{\alpha\sqrt{n}}\right). \tag{24}$$

If α is bounded away both from zero and from infinity, then

$$b = \alpha + O\left(\frac{1}{\sqrt{n}}\right) \tag{25}$$

because of (24), but (25) involves

$$\frac{b^m e^{-b}}{m!} = \frac{\alpha^m e^{-\alpha}}{m!}\left[1 + O\left(\frac{1}{\sqrt{n}}\right)\right]$$

so that from (20)

$$b = \alpha + (\alpha - m)\, p \frac{\alpha^m e^{-\alpha}}{m!}\left[1 + O\left(\frac{1}{\sqrt{n}}\right)\right] \tag{26}$$

follows. After repeated application of (20) and (25) also the more elaborate asymptotic expansions, needed for the subsequent calculations

$$(27)\qquad b = \alpha + (\alpha - m)\, p \frac{\alpha^m e^{-\alpha}}{m!} - \frac{(\alpha - m)^3}{\alpha} p^2 \frac{\alpha^{2m} e^{-2\alpha}}{(m!)^2} + O\left(\frac{1}{n\sqrt{n}}\right)$$

and

$$(28)\qquad p \frac{b^m e^{-b}}{m!} = \frac{b-\alpha}{\alpha - m} = p \frac{\alpha^m e^{-\alpha}}{m!} - \frac{(\alpha - m)^2}{\alpha} p^2 \frac{\alpha^{2m} e^{-2\alpha}}{(m!)^2} + O\left(\frac{1}{n\sqrt{n}}\right)$$

can be deduced.

If $\alpha \to \infty$ then first we have from (20)

$$b > c\alpha > m,$$

for α large with a c positive and tending to 1. It follows therefore that

$$\frac{b^m e^{-b}}{m!} < \frac{(c\,\alpha)^m e^{-c\alpha}}{m!} = O\,(\alpha^m e^{-c\alpha})$$

and

$$p(\alpha - m) \frac{b^m e^{-b}}{m!} = O\left(\frac{e^{-\alpha/(2+\eta)}}{\sqrt{n}}\right)$$

(where $\eta \to 0$), i.e. (25) holds in the present case, too. The expansions (27) and (28) are now deducible as former.

If $\alpha \to 0$ and $m \geqq 2$, then from (20) we have

$$\frac{b}{\alpha} < 1 + m\,|p| \frac{b}{\alpha} \cdot \frac{b^{m-1} e^{-b}}{m!} < 1 + \frac{b}{\alpha} m^2 |p|,$$

thus b/α is bounded from above. With sufficiently large n it follows then $b < C\,\alpha < m$ with a C bounded from above so that we have

$$b^m e^{-b} = O\,(\alpha^m)$$

and

$$p \frac{b^m e^{-b}}{m!} = O\left(\frac{\alpha}{\sqrt{n}}\right),$$

thus

$$(29)\qquad b = \alpha\left[1 + O\left(\frac{1}{\sqrt{n}}\right)\right].$$

With repeated application of (20) and (29) the expansions (27) and (28) are deducible again.

For $\alpha \to 0$ and $m = 1$ we obtain in the same manner as former the asymptotic expansions (27) and (28), but now with the remainder term $O\left(\frac{1}{\alpha^2 n \sqrt{n}}\right)$. Similarly, for $\alpha \to 0$ and $m = 0$ the second of the expansions

$$(30)\qquad b = \alpha + p\,\alpha\, e^{-\alpha} - p^2\, \alpha^2 e^{-2\alpha} + O\left(\frac{1}{n\sqrt{n}}\right)$$

$$(31)\qquad pe^{-b} = \frac{b-\alpha}{\alpha} = pe^{-\alpha} - p^2 \alpha\, e^{-2\alpha} + O\left(\frac{1}{\alpha\, n \sqrt{n}}\right)$$

corresponding to (28), has the remainder term $O\left(\frac{1}{\alpha\, n \sqrt{n}}\right)$ instead of $O\left(\frac{1}{n\sqrt{n}}\right)$.

As $D \to \gamma = \text{const.} \neq 0$, we take $p = e^x - 1 = \text{const.}$ and the asymptotic behaviour of b may be expressed by

$$b = \alpha + (\alpha - m)\frac{p\gamma}{n}(1 + o(1)) \tag{32}$$

if $\alpha \to \infty$ (without restriction on n) or if $\alpha \to 0$ but $m \neq 0,1$. For $\alpha \to 0$ and $m = 0,1$ it is enough to show that

$$b \sim \frac{\alpha}{1+p} \sim \frac{\sqrt{\gamma}}{(1+p)\sqrt{n}}$$

where now $\gamma = \lim_{n\to\infty} n\alpha^2$.

6. The asymptotic behaviour of the factor F. In order to have a convenient form for F, we make use of Stirling's formula and of the equation (20). Thus we obtain

$$\log F = n(\alpha - m)\, p \frac{b^m e^{-b}}{m!} + n \log\left(1 + p\frac{b^m e^{-b}}{m!}\right) - \tag{33}$$

$$- N \log\left(1 + p\frac{\alpha - m}{\alpha} \cdot \frac{b^m e^{-b}}{m!}\right) + o(1)$$

as starting-point. Supposed $D \to \infty$, the terms with logarithmic factors in (33) can be expanded down to $o(1)$, and regard to

$$p = e^{x/D} - 1 = \frac{x}{D} + \frac{x^2}{2D^2} + O\left(\frac{1}{D^3}\right)$$

and to (28), we easily obtain

$$\log F = \frac{xE}{D} + \frac{x^2}{2} + o(1)\,. \tag{34}$$

However, this simple procedure can only be applied if the terms $pb^m e^{-b}/m!$ and $p\dfrac{\alpha - m}{\alpha} \cdot b^m e^{-b}/m!$ have the order $O(1/\sqrt{n})$, that does not hold in the cases $\alpha \to 0$, $m = 0$ and $m = 1$. If $m = 0$, then

$$\log F = n\,\alpha\, pe^{-b} + (n - n\,\alpha)\log(1 + pe^{-b}) + o(1)\,.$$

and in order to avoid the mentioned difficulty, let us put

$$\log(1 + pe^{-b}) = \log(1 + p) + \log\left(1 - \frac{p(1 - e^{-b})}{1 + p}\right),$$

the term $p(1 - e^{-b})$ beeing $O(1/\sqrt{n})$. Similarly, if $m = 1$, then

$$\log F = n(\alpha - 1)\, pb\, e^{-b} + n\log(1 + pb\, e^{-b}) - n\,\alpha \log\left(1 + \frac{\alpha - 1}{\alpha}\, pb\, e^{-b}\right).$$

Considering that from (20)

$$1 + \frac{\alpha - 1}{\alpha} pb\, e^{-b} = \frac{b}{\alpha},$$

and that the quantity b/α may be expressed in the case $m = 1$ more convenient by

$$\frac{b}{\alpha} = \frac{1}{1 - (\alpha - 1)\, pe^{-b}},$$

we may put

$$n\,\alpha \log\left(1 + \frac{\alpha - 1}{\alpha} pb\, e^{-b}\right) =$$

$$= -\, n\,\alpha \log(1 + p) - n\,\alpha \log\left(1 - \frac{p(1 - e^{-b} + \alpha\, e^{-b})}{1 + p}\right).$$

After expanding the logarithmic terms and using (28) the result will be (34) again, nevertheless, the computations are, however elementary, somewhat more cumbersome, because the remainder term in (28) as said above is now considerably greater then $O(n^{-3/2})$.

Supposing $D \to \gamma = \text{const.} \neq 0$ the corresponding results are

$$\log F = \frac{1}{2} \gamma\, p + (n - N) \log(1 + p) + o(1)$$

for $m = 0$, $\alpha \to 0$,

$$\log F = N \log(1 + p) - \frac{1}{2} \gamma \frac{p(p + 2)}{(p + 1)^2} + o(1)$$

for $m = 1$, $\alpha \to 1$ with $\gamma = \lim\limits_{n \to \infty} n\, \alpha^2$, and

$$\log F = \gamma\, p + o(1)$$

with $\gamma = \lim\limits_{n \to \infty} n\, \alpha^m e^{-\alpha}/m\,!$ in all other cases.

7. The asymptotics of the integral J. As to the integral (23), according to Debye's method, one must find and take the steepest descent's path through the critical saddle point. This is, however, in general not necessary, it is sufficient to find such a line as a path of integration, which is convenient to apply Laplace's method (see e. g. [9]). The fact that the integrand will be eventually complex-valued, does not matter, rather the difficulty arises that in the present case the integrand has two parameters, both tending to infinity, their rate α beeing not previously fixed. Applying Laplace's method according to its sense, the integral will be divided into three parts; the first, ,,essential" part J_1, containing the close vicinity of the saddle point b, tends to 1, whereas the other two, "unessential" parts tend to zero. This will be proved as follows.

a) $\alpha \to \infty$, $\lim\inf n\alpha^m e^{-\alpha} > 0$. We take the straight line through the saddle point parallel to the imaginary axis to be the path integration. It is easy to see that the new path is equivalent to the original circle. Putting

$w = 1 + iu$, we have

$$J = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{e^{iub}(1+iu)^{-m} + p\dfrac{b^m e^{-b}}{m!}}{1 + p\dfrac{b^m e^{-b}}{m!}} \right)^n (1+iu)^{n(m-\alpha)} du .$$

Divide the integral into three parts, the "essential" part beeing

$$J_1 = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-n^{-1/5}}^{n^{-1/5}} f(u)\, du ,$$

where $f(u)$, with respect to (20) may be written as

$$\log f(u) = n \log \left(1 + \frac{\alpha - m}{b - m} \left(\frac{e^{ibu}}{(1+iu)^m} - 1 \right) \right) - n(\alpha - m) \log (1 + iu) .$$

Expanding $\log f(u)$ in prowers of u to $O(u^3)$ we obtain

$$\log f(u) = - \frac{(N+1)\, u^2}{2} \left(1 - \frac{m}{\alpha}\right) \left(b - \alpha + \frac{b}{b-m} \right) + O(nb^3 u^3) ,$$

i.e.

$$J_1 = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-n^{-1/5}}^{n^{-1/5}} \exp \left\{ - \frac{N+1}{2} u^2 \left[1 + o(1) + O(\alpha^2 u) \right] \right\} du = 1 + o(1) ,$$

because of $O(\alpha^2 u) = O(\alpha^2 n^{-1/5}) = o(1)$ and $N^{1/2} n^{-1/5} \to \infty$.

As to the "unessential" part

$$J_2 = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{n^{-1/5}}^{\infty} f(u)\, du ,$$

we have the inequality

$$\left| \frac{e^{ibu}(1+iu)^{-m} + p\dfrac{b^m e^{-b}}{m!}}{1 + p\dfrac{b^m e^{-b}}{m!}} \right|^n \leq \left(\frac{1 + |p|\dfrac{b^m e^{-b}}{m!}}{1 - |p|\dfrac{b^m e^{-b}}{m!}} \right)^n = O(e^{c\sqrt{n}}) ,$$

where c is bounded from above, so that

$$|I_2| = O(e^{c\sqrt{n}}) \cdot O\left(\frac{1}{(1 + n^{-2/5})^{\frac{(\alpha - m)n}{2}}} \right) \cdot O(\sqrt{N}) = o(1) ,$$

and similarly

$$I_3 = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\infty}^{-n^{-1/5}} f(u)\, du = o(1) .$$

b) lim sup $\alpha < \infty$. For the path of integration we take the unit circle around the point $w = 0$ on the w-plane. Putting $w = {}^{iu}$, we have from (23)

$$J = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \left[\frac{\exp\{-b(1-e^{iu})\} + pe^{imu} \dfrac{b^m e^{-b}}{m!}}{1 + p\dfrac{b^m e^{-b}}{m!}} \right]^n e^{-iNu}\, du =$$

$$= \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\pi}^{\pi} g(u)\, du ,$$

the essential part being now

$$J_1 = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{-\delta}^{\delta} g(u)\, du ,$$

where $\delta > 0$ is arbitrarily small, but fixed.

For the unessential part

$$I_2 = \frac{\sqrt{N}}{\sqrt{2\pi}} \int_{\delta}^{\pi} g(u)\, du$$

with respect to (20) and (35) the inequality

$$|I_2| = O(\sqrt{N}) \cdot e^{-nb(1-\cos\delta)} \cdot \max_{\delta \leq u \leq \pi} \exp\left\{ n \frac{\alpha - m}{b - m} |p| \frac{b^m}{m!} |e^{imu + b(1-e^{iu})} - e^{-b}| \right\}$$

is easily obtainable. Since the function $C = \exp\{imu + b(1-e^{iu})\} - e^{-b}$ is bounded,

$$|I_2| = O(\sqrt{N}) \cdot O(\exp\{-nb(1 - \cos\delta - o(1)\,|p|\, b^{m-1})\}) \tag{36}$$

follows.

If $p \to 0$, then the final result is

$$|I_2| = O(\sqrt{N}) \cdot O(e^{-N\eta}) = o(1) \tag{37}$$

(where lim inf $\eta > 0$). For $m = 0$ the proof seems at first sight to be incorrect, but considering that for $m = 0$ the function C is bounded by $O(b)$ the term $o(1)$ in (36) can be now replaced by $O(b)$.

The inequality (36) involves (37) also if $p =$ const., $\alpha \to 0$, $n\alpha^m \to$ const. for $m \geq 2$, but the cases $n\alpha^2 \to$ const., $m = 0$ or 1 need special considerations. If $m = 1$, then

$$|I_2| = O(\sqrt{N}) \cdot \max_{\delta \leq u \leq \pi} \left| \frac{\exp\{be^{iu}\} + pbe^{iu}}{e^b + p} \right|^n =$$

$$= O(\sqrt{N}) \cdot \max_{\delta \leq u \leq \pi} |1 + (p+1)\, b(e^{iu} - 1) + O(b^2)|^n .$$

but with regard to $b^2 = O\left(\frac{1}{n}\right)$ we have

$$|I_2| = O(\sqrt{N})\left[1 - 2b(p+1)\left[1 - \cos\delta + O(b)\right]\right]^{n/2},$$

and $p + 1$ being always positive the relation (37) follows. For $m = 0$ the proof of (36) is similar to that for $m = 1$.

There remained to show that the essential part J_1 tends to 1. As former, the logarithm of $\log g(u)$ will be expanded to $O(u^3)$, but care must be taken that, with regard to the possibility of $b \to 0$, the remainder term should have the form $O(bu^3)$, therefore it will be convenient to use a special form of $\log g(u)$, easily obtainable from (20) and (35):

$$\log g(u) = n \log\left\{1 + \frac{a-m}{b-m}\left[\exp\{-b(1-e^{iu})\} - 1 + \right.\right.$$

$$\left.\left. + p\frac{b^m e^{-b}}{m!}(e^{imu} - 1)\right]\right\} - iNu$$

leading to

$$\log g(u) = iu - \frac{bnu^2}{2}\cdot\frac{a-m}{b-m}\left[1 + b + m^2 p\frac{b^{m-1}e^{-b}}{m!} - \frac{a^2}{b}\cdot\frac{b-m}{a-m}\right] + O(bu^3) =$$

$$= iu - \frac{Nu^2}{2}\left[1 + o(1) + O(u)\right],$$

and, having in mind that δ is fixed, but arbitrarily small,

$$J_1 = \frac{\sqrt{N}}{\sqrt{2\pi}}\int_{-\delta}^{\delta} e^{iu - \frac{Nu^2}{2}[1+o(1)+O(u)]}\, du \to 1$$

follows, making the proof of the theorem complete.

(Received January 11, 1963.)

REFERENCES

[1] Feller, W.: *An Introduction to Probability Theory and Its Applications.* Vol. 1. (Second Edition) N. Y. 1957. (Problem 22. page 103).
[2] Bernstein, S. N.: *Teorija Verojatnosztyej.* Moszkva, 1946.
[3] von Mises, R.: "Über Aufteilung und Besetzungswahrscheinlichkeiten". *Revue de la Faculté des Sciences de l'Université d'Istanbul,* N. S. **4** (1939) 145—163.
[4] Weiss, I.: "Limiting distribution in some occupancy problems", *Ann. Math. Stat.,* **29** (1958) 878—884.
[5] Rényi, A.: "Three new proofs and a generalisation of a theorem of Irving Weiss". *MTA Mat. Kut. Int. Közl.* **7** (1962) 203—215.
[6] Békéssy, A.: "Egy elosztási problémára vonatkozó határeloszlástétel új bizonyítása". *MTA III. Osztályának Közleményei* **12** (1962) 329—334.
[7] Curtiss, J. H.: "A note on the theory of moment generating functions". *Ann. Math. Stat.* **13** (1942) 430—433.
[8] Erdős, P.—Rényi, A.: "On a classical problem of porability theory". *MTA Mat. Kut. Int. Közl.* **6** A. (1961) 215—220.
[9] Erdélyi, A.: *Asymptotic expressions.* Dover Publ., 1956.
[10] David, F. N.—Barton, D. E.: *Combinatorial Chance.* London, 1962, Ch. Griffin et Comp.

О КЛАССИЧЕСКИХ ЗАДАЧАХ ЗАПОЛНЕНИЯ ЯЩИКОВ I.

A. BÉKÉSSY

Резюме

Пусть события $a_1, a_2, \ldots, a_n$ являются всеми элементами дискретного и конечного поля событий и пусть их вероятностное распределение равномерно. Справшивается, сколько будет таких событий которые точно m ряда $(m = 0, 1, 2, \ldots)$ состоятся в некотором образце этих событий, состоящего из N элементов. Обозначим число этих событий через $\xi(n, N, m)$. Поставленный вопрос может быть сформирован более наглядно посредством ящиков и шариков таким образом: если распределить N шариков в n ящиках случайным способом, какое будет число ящиков содержащих точно m шариков.

Статья содержит следующую теорему:

если $n \to \infty$ *и* $N \to \infty$ *и* $m =$ *конст. и если, кроме того*

$$D^2 = n \frac{\alpha^m e^{-\alpha}}{m!} \left[1 - \frac{\alpha^m e^{-\alpha}}{m!} \left(1 + \frac{(\alpha - m)^2}{\alpha}\right)\right] \to \infty ,$$

где $\alpha = \dfrac{N+1}{n}$, *тогда предельное распределение стандартизированного случайного переменного* $\xi(n, N, m)$ *является нормальным.* (Сравн. (2), (3), (4)). В качестве дополнения автор приводит в больших чертах также предельный случай $D^2 \to \gamma = \text{konst} \neq 0$.

Теорему, высказанную в статье ранее доказал A. Rényi [5] для частного случая $m = 0$, соответственно доказали ее для общего случая F. N. David и D. E. Barton [10], однако при более сильных ограничениях.