University of Groningen

Generation of a random partition of a finite set by an urn model

Stam, A.J.

Published in:
Journal of Combinatorial Theory. Series A

DOI:
10.1016/0097-3165(83)90009-2

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
1983

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Stam, A. J. (1983). Generation of a random partition of a finite set by an urn model. *Journal of Combinatorial Theory. Series A*, *35*(2), 231-240. https://doi.org/10.1016/0097-3165(83)90009-2

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 04-06-2022

JOURNAL OF COMBINATORIAL THEORY, Series A 35, 231–240 (1983)

Generation of a Random Partition of a Finite Set by an Urn Model

A. J. STAM

Mathematisch Instituut Rijksuniversiteit, Postbus 800, 9700 AV Groningen, Netherlands

Communicated by the Managing Editors

Received March 19, 1982

A random partition of $\mathbb{N}_n = \{1,..., n\}$ may be generated by putting the elements of $\mathbb{N}_n$ at random into a stochastic number of cells. This representations is used to prove asymptotic results about the random partition for $n \to \infty$.

1. INTRODUCTION AND NOTATIONS

A partition of $\mathbb{N}_n = \{1,..., n\}$ is an unordered family of disjoint proper subset of $\mathbb{N}_n$ with union $\mathbb{N}_n$. A generic element of the class Π_n of partitions of $\mathbb{N}_n$ will be denoted by π. We have $|\Pi_n| = T_n$, where T_n is the nth Bell or exponential number given by Dobinski's formula

$$T_n = e^{-1} \sum_{k=1}^{\infty} k^n/k!. \tag{1.1}$$

The number of partitions of $\mathbb{N}_n$ consisting of k subsets is $S(n, k)$, a Stirling number of the second kind. See Comtet [2]. We define $S(n, k) = 0$ for $k > n$ and for $k \leqslant 0$, $n \geqslant 1$. The asymptotic behaviour as $n \to \infty$ of T_n, originally derived by Moser and Wyman [9], and the combinatorics of Π_n were studied in [3, 4, 12–15]. We will need the relation

$$T_{n+1}/T_n = n\rho_n^{-1} + \tfrac{1}{2}\rho_n(1 + \rho_n)^{-2} + O(n^{-1}\rho_n), \tag{1.2}$$

where $\rho_n = \rho(n)$ and ρ is the inverse of the function xe^x on $[0, \infty)$, so that $\rho_n \sim \log n$ as $n \to \infty$. Manipulation of the ρ_n is simplified by the relation

$$\rho'(t) = t^{-1}\rho(t)/(1 + \rho(t)). \tag{1.3}$$

The combinatorics of Π_n may be described in terms of probability theory by "choosing an element of Π_n at random," i.e., defining a probability

0097-3165/83 $3.00

Copyright © 1983 by Academic Press, Inc.
All rights of reproduction in any form reserved.

measure P_n on Π_n, or rather on the power set of Π_n, such that $P_n(\{\pi\}) = T_n^{-1}$, $\pi \in \Pi_n$. The number of partitions having property A then is $T_n P_n(A)$, where $P_n(A)$ is the probability that the random partition has property A. The heuristic value of this description lies in the exploitation of concepts and theorems of probability theory, such as independence and the central limit theorem, as in [1, 4]. More may be done by embedding (Π_n, P_n) into a larger probability space $(\Omega, \mathscr{S}, P)$, where P is a probability measure on the σ-algebra $\mathscr{S}$ of subsets of Ω, and π is a stochastic variable on Ω. The most obvious method would be the urn model with a fixed number v of urns. Put the elements of $\mathbb{N}_n$ at random into the urns, i.e., place them independently with probability v^{-1} for each of the urns. The nonempty urns then define a stochastic partition of $\mathbb{N}_n$. This, however, is not a random partition: the resulting probability distribution on Π_n is not the uniform distribution P_n, see [5, 7].

We may obtain P_n by taking a stochastic number U of urns with

$$P(U=u) = q_{nu} = T_n^{-1} e^{-1} u^n/u!, \qquad u = 1, 2, \ldots. \tag{1.4}$$

Note that (1.4) defines a probability distribution by (1.1). Under the condition $U = u$ the elements of $\mathbb{N}_n$ are put at random into the u urns. The contents of the nonempty urns then form a stochastic partition of $\mathbb{N}_n$. We have

THEOREM 1. *This partition is a random partition of $\mathbb{N}_n$, i.e., the probability that the urn procedure generates a prescribed $\pi \in \Pi_n$ is T_n^{-1}.*

Proof. If π consists of k subsets, the conditional probability $P(\pi \mid u)$ of obtaining π given $U = u$ is $u^{(k)} u^{-n}$ for $u \geqslant k$ and zero for $u < k$, where

$$m^{(0)} = 1, \qquad m^{(i)} = m(m-1) \cdots (m-i+1), \qquad m, i \in \mathbb{N}. \tag{1.5}$$

So with (1.4) and (1.1),

$$P(\pi) = \sum_{u=k}^{\infty} q_{nu} u^{(k)} u^{-n} = T_n^{-1} \sum_{u=k}^{\infty} e^{-1}/(u-k)! = T_n^{-1}.$$

A formal definition of the above probability measure P may be given on a countable product space Ω, see Neveu [10, Chap. III].

Theorem 1 implies that any theorem about P_n on Π_n may be derived by considering the random partition to be generated by the urn process. This method, with a different urn model, was applied by Kolchin [6] to the cycles of a random permutation of $\mathbb{N}_n$.

The urn model will be applied to the following problems. In the random partition π let A_i be the element of π that contains i and let

$$W_i = |A_i| - 1, \qquad i = 1, \ldots, n. \tag{1.6}$$

Since a partition defines an equivalence relation on $\mathbb{N}_n$, the stochastic variable W_i also may be defined as the number of elements of $\mathbb{N}_n$ that are equivalent to i. In Section 2 we will study the properties of the probability distribution (1.4) and show that the W_i are asymptotically independent with Poisson distribution.

Let $\pi = \{B_1, ..., B_k\} \in \Pi_n$. Then, if $r \leqslant n$, the nonempty sets $B_j \cap \mathbb{N}_r$ form a partition of $\mathbb{N}_r$, called the relative partition of $\mathbb{N}_r$ induced by π, and denoted $\pi \cap \mathbb{N}_r$. If π is random, $\pi \cap \mathbb{N}_r$ in general is not, i.e., its probability distribution is not P_r. In Section 3 we prove that asymptotically as $n \to \infty$ and $r/n = O(\rho_n^{-1})$, the relative partition may be considered to be generated by an urn model with a nonstochastic number of urns.

These results will be formulated in terms of L_1, the absolute variation distance between probability measures. Let m_1, m_2 be finite measures on $B \subset \mathbb{Z}^k$ with

$$m_i(A) = \sum_{j \in A} p_i(j), \qquad A \subset B, \quad i = 1, 2, \tag{1.7}$$

where, as later on, we write j for $(j_1, ..., j_k)$. The L_1 distance between m_1 and m_2 is

$$\|m_1 - m_2\| = \sum_{j \in B} |p_1(j) - p_2(j)|. \tag{1.8}$$

For ψ: $B \to D \subset \mathbb{Z}^h$ we define the measures $m_i \psi^{-1}$ on D by $m_i \psi^{-1}(G) = m_i(\psi^{-1}(G))$. If m_i is a probability measure, $m_i \psi^{-1}$ is the probability distribution of ψ under m_i. We have the following lemma, easily proved:

LEMMA 1.

$$\|m_1 \psi^{-1} - m_2 \psi^{-1}\| \leqslant \|m_1 - m_2\|,$$

with equality if $p_1(j)/p_2(j) = f \circ \psi(j)$ *for* $p_1(j) + p_2(j) > 0$, *where we define* $a/0 = \infty$, $a > 0$.

2. THE URN MODEL AND THE W_i

For the urn model described in Section 1 we have

THEOREM 2. *The random partition of* $\mathbb{N}_n$ *generated by the urn model and the number* M_0 *of empty urns are independent. The distribution of* M_0 *is Poisson with parameter* 1.

Proof. In order that $M_0 = m$ and that a prescribed partition π consisting of k sets is generated, we must have $U = m + k$. So

$$\begin{aligned} P(\pi, M_0 = m) &= P(U = m + k)\, P(\pi, M_0 = m \mid U = m + k) \\ &= T_n^{-1} e^{-1} (m + k)^n\, (m + k)^{(k)}\, (m + k)^{-n} / (m + k)! \\ &= T_n^{-1} e^{-1} / m! = P(\pi)\, e^{-1} / m!. \end{aligned}$$

For the generating function and the moments of U we find from (1.4) and (1.1)

$$Ez^U = \exp(z - 1)\, T_n^{-1} \sum_{u=1}^{\infty} z^u e^{-z} u^n / u!, \tag{2.1}$$

$$EU^h = T_{n+h} / T_n. \tag{2.2}$$

From (2.2) and (1.2) with (1.3)

$$a_n = EU = n\rho_n^{-1} + \tfrac{1}{2}\rho_n (1 + \rho_n)^{-2} + O(n^{-1}\rho_n), \tag{2.3}$$

$$\sigma_n^2 = \sigma^2(U) = n\rho_n^{-1}(1 + \rho_n)^{-1} + O(1). \tag{2.4}$$

Remark. Let X be the number of subsets of $\mathbb{N}_n$ in the random partition. Since X and M_0 are independent by Theorem 2 and $U = X + M_0$, relation (2.1) corresponds to the convolution property of generating functions and we may use (2.2) to find lower order moments of X.

Let W_i be defined by (1.6). By conditioning with respect to U and the urn into which i lands, we see that the distribution of W_i is weighted binomially. For $k = 0, \ldots, n - 1$,

$$P_n(W_i = k) = \sum_{n=1}^{\infty} q_{nu} \binom{n-1}{k} u^{-k} (1 - u^{-1})^{n-1-k}. \tag{2.5}$$

From (2.5) with (1.1)–(1.3)

$$\begin{aligned} b_n = EW_i &= (n - 1)\, T_{n-1} / T_n \\ &= \rho_n \{1 - (2n)^{-1} - (2n)^{-1}\, (1 + \rho_n)^{-2} + O(n^{-2}\rho_n^2)\}, \end{aligned} \tag{2.6}$$

$$\begin{aligned} \sigma^2(W_i) &= (n - 1)^2\, (T_{n-2} T_n^{-1} - T_{n-1}^2 T_n^{-2}) \\ &\quad + (n - 1)\, T_n^{-1} (T_{n-1} - T_{n-2}) = b_n (1 + b_{n-1} - b_n) \\ &= \rho_n \{1 - (2n)^{-1}\, (3\rho_n^2 + 4\rho_n + 2)(1 + \rho_n)^{-2} + O(n^{-2}\rho_n^2)\}. \end{aligned} \tag{2.7}$$

Now let $Q = Q_{nh}$ be the joint probability distribution of $W_1, \ldots, W_h$ and $R = R_{nh}$ the product of h Poisson probability distributions, each with parameter ρ_n. Then we have

THEOREM 3. *With* L_1 *distance defined by* (1.8) *and* a_n, σ_n *by* (2.3) *and* (2.4),

$$\|Q-R\| \leqslant hT_n^{-1}T_{n-1}\{h+2+2[\rho_n^2\sigma_{n-1}^2+(\rho_n a_{n-1}-n+h)^2]^{1/2}\}. \quad (2.8)$$

Remark. If $h=o(n^{1/2}\rho_n^{-1})$ as $n\to\infty$, we have by (1.2), (2.3), (2.4)

$$\|Q-R\| \sim 2n^{-1/2}\rho_n h \to 0, \qquad n\to\infty. \quad (2.9)$$

From the central limit theorem for the Poisson distribution it follows that $W_1,..., W_k$ are asymptotically normal and independent.

Proof. We write $W=(W_1,..., W_h)$ and $j=(j_1,..., j_h)$. Let D be the event that 1,..., h are put into different urns. We have

$$\begin{aligned} P(D) &= \sum_{u=h}^{\infty} q_{nu} u^{(h)} u^{-h} \\ &\geqslant \sum_{u=1}^{\infty} q_{nu}\{1-\tfrac{1}{2}h(h-1)\,u^{-1}\} = 1-\tfrac{1}{2}h(h-1)\,T_{n-1}T_n^{-1}. \quad (2.10)\end{aligned}$$

By conditioning with respect to the urns into which 1,..., h are put, we have

$$P(D, W=j) = \sum_{u=h}^{\infty} q_{nu} u^{(h)} u^{-h} a(n-h, u, j), \quad (2.11)$$

where the $a(n-h, u, j)$ are probabilities of a multinomial distribution $A = A(n-h, u)$ defined by

$$a(n-h, u, j) = \frac{(n-h)!}{j_1!\cdots j_h!\,(n-h-k)!}\, u^{-k}(1-hu^{-1})^{n-h-k}, \quad (2.12)$$

for $0\leqslant k=j_1+\cdots+j_h\leqslant n-h$ and zero elsewhere. Then, with (2.10) and (2.11)

$$\begin{aligned} \|Q-R\| &= \sum_j |P(D, W=j)+P(D^c, W=j)-R(j)| \\ &\leqslant \sum_j |P(D, W=j)-P(D)\,R(j)| + 2P(D^c) \\ &\leqslant 2P(D^c) + \sum_{u=h}^{\infty} q_{nu} u^{(h)} u^{-h} \|A(n-h, u)-R\|. \quad (2.13)\end{aligned}$$

Let $B=B(n-h, u)$ be the product of h Poisson distributions with parameter $(n-h)\,u^{-1}$, i.e., with the same first moments as the multinomial distribution $A(n-h, u)$. We have

$$\|A-R\| \leqslant \|A-B\| + \|B-R\|. \quad (2.14)$$

It is easily seen that $A(n-h,u)$ and $B(n-h,u)$ satisfy the equality condition of Lemma 1 with $\psi(j)=j_1+\cdots+j_h$, and that $A\psi^{-1}$ is the binomial distribution with parameters $n-h$ and hu^{-1}, whereas $B\psi^{-1}$ is the Poisson distribution with parameter $(n-h)\,hu^{-1}$. It was proved in [8, 11] that $\|A\psi^{-1}-B\psi^{-1}\|\leqslant 3hu^{-1}$. So

$$\|A(n-h,u)-B(n-h,u)\|\leqslant 3hu^{-1}. \tag{2.15}$$

Writing $f_i(j)=\alpha_i^k(j_1!\cdots j_h!)^{-1}$ with $k=j_1+\cdots+j_h$, $i=1,2$, we have

$$\begin{aligned}
&\sum_j |f_1(j)\exp(-h\alpha_1)-f_2(j)\exp(-h\alpha_2)|\\
&\quad\leqslant \sum_j f_1(j)\,|\exp(-h\alpha_1)-\exp(-h\alpha_2)|+\sum_j |f_1(j)-f_2(j)|\exp(-h\alpha_2)\\
&\quad= 2-2\exp(-h\,|\alpha_1-\alpha_2|)\\
&\quad\leqslant 2h\,|\alpha_1-\alpha_2|.
\end{aligned}$$

This gives, by definition of the Poisson distributions R and $B(n-h,u)$,

$$\|R-B\|\leqslant 2h\,|\rho_n-(n-h)\,u^{-1}|. \tag{2.16}$$

From (2.13)–(2.16), with (1.1) and (1.4),

$$\|Q-R\|\leqslant 2P(D^c)+3hT_n^{-1}T_{n-1}+C_n, \tag{2.17}$$

$$\begin{aligned}
C_n&=2h\sum_{u=h}^{\infty} q_{nu}u^{(h)}u^{-h}\,|\rho_n-(n-h)\,u^{-1}|\\
&\leqslant 2h\rho_n T_{n-1}T_n^{-1}\sum_{u=1}^{\infty} q_{n-1,u}\,|u-\rho_n^{-1}(n-h)|\\
&\leqslant 2h\rho T_{n-1}T_n^{-1}\left\{\sum_{u=1}^{\infty} q_{n-1,u}(u-\rho_n^{-1}(n-h))^2\right\}^{1/2}\\
&=2h\rho_n T_{n-1}T_n^{-1}\{\sigma_{n-1}^2+(a_{n-1}-\rho_n^{-1}(n-h))^2\}^{1/2},
\end{aligned} \tag{2.18}$$

with a_n and σ_n given by (2.3) and (2.4). Inequality (2.8) now follows from (2.7), (2.10), and (2.18).

3. Relative Partitions

We need some results about the urn model with a fixed number v of urns, see [5, 7]. When r objects are put at random into v urns we have for $Y=Y_{rv}$, the number of occupied urns,

$$p(r, v, k) = P(Y = k) = S(r, k)\, v^{(k)} v^{-r}, \tag{3.1}$$

$$EY = v - v(1 - v^{-1})^r, \tag{3.2}$$

$$\sigma^2(Y) = v(1 - v^{-1})^r + v(v-1)(1 - 2v^{-1})^r - v^2(1 - v^{-1})^{2r}. \tag{3.3}$$

If $r \to \infty$, $v \to \infty$ so that $r/v \to \alpha \in (0, \infty)$, we have

$$EY = v\theta(r/v) + O(1), \tag{3.4}$$

$$\theta(x) = 1 - e^{-x}, \tag{3.5}$$

$$\sigma^2(Y) = v\eta(r/v) + O(1), \tag{3.6}$$

$$\eta(x) = e^{-2x}(e^x - 1 - x). \tag{3.7}$$

As defined in Section 1, a random partition of $\mathbb{N}_n$ induces a relative partition of $\mathbb{N}_r$ that has a nonuniform probability distribution $G = G_{nr}$ on Π_r. We may consider G_{nr} to be derived from the probability space (Π_n, P_n), but also from the probability space (Ω, P) described under (1.4). It follows from (3.1) and (1.4) that the probability of k subsets in the relative partition is

$$\sum_{u=k}^{\infty} q_{nu} S(r, k)\, u^{(k)} u^{-r} = S(r, k)\, T_n^{-1} \sum_{v=0}^{\infty} e^{-1}(v + k)^{n-r}/v!. \tag{3.8}$$

We will compare G_{nr} with the probability measure H_{nr} on Π_r that corresponds to an urn model with a fixed number of urns. Put the elements of $\mathbb{N}_r$ at random into c_n urns, where

$$c_n = [a_n], \qquad a_n = EU, \tag{3.9}$$

see (1.4) and (2.3). The contents of the nonempty urns form a partition of $\mathbb{N}_r$ with probability distribution H_{nr}. The idea is that if $n \to \infty$, and $r/n \to 0$ sufficiently fast, the ratio U/a_n is sufficiently close to 1. We have

THEOREM 4. *If $n \to \infty$ and r varies with n so that $r = O(a_n)$,*

$$\lim_{n \to \infty} \|G_{nr} - H_{nr}\| = 0. \tag{3.10}$$

Remark. This means that asymptotic results for relative partitions of a random partition of $\mathbb{N}_n$ may be derived from the urn model with a nonstochastic number of urns, e.g., Poisson and normal limits for distribution (3.8) as given in [5, 7].

Proof. First we take $r = \lambda c_n$, where λ is a fixed positive integer. For a given partition $\pi \in \Pi_r$ consisting of k subsets

$$H_{nr}(\pi) = c_n^{(k)} c_n^{-r},$$

$$G_{nr}(\pi) = \sum_{u=1}^{\infty} q_{nu} u^{(k)} u^{-r}.$$

Note that $u^{(k)} = 0$, $u \leqslant k$. From Lemma 1, (3.1) and (3.8),

$$\|H_{nr} - G_{nr}\| = \sum_{k=1}^{r} \left| p(r, c_n, k) - \sum_{u=1}^{\infty} q_{nu} p(r, u, k) \right|$$

$$\leqslant \sum_{u=1}^{\infty} q_{nu} V(r, n, u), \tag{3.11}$$

$$V = V(r, n, u) = \sum_{k=1}^{r} |p(r, c_n, k) - p(r, u, k)|. \tag{3.12}$$

Put

$$D_n = \sum_u{}' q_{nu} V(r, n, u), \qquad E_n = \sum_u{}'' q_{nu} V(r, n, u), \tag{3.13}$$

where $\sum'$ and $\sum''$ mean summation over $|u - c_n| \geqslant K_1 \sigma_n$ and $|u - c_n| < K_1 \sigma_n$, respectively, with $\sigma_n = \sigma(U)$ given by (2.4). Since $V \leqslant 2$, we have by (3.9) and Chebyshev's inequality

$$D_n \leqslant 2(K_1 - 1)^{-2}. \tag{3.14}$$

For $|u - c_n| < K_1 \sigma_n$ we write

$$V(r, n, u) = \sum_{i=1}^{3} V_i(r, n, u), \tag{3.15}$$

$$V_1 = \sum_k |p(r, c_n, k) - f(r, n, k)|, \tag{3.16}$$

$$V_2 = \sum_k |f(r, n, k) - g(r, u, k)|, \tag{3.17}$$

$$V_3 = \sum_k |g(r, u, k) - p(r, u, k)|, \tag{3.18}$$

where we have, with $\beta = r/u$ and θ, η given by (3.5) and (3.7),

$$f = (2\pi c_n \eta(\lambda))^{-1/2} \exp\{-(k - c_n \theta(\lambda))^2 / 2c_n \eta(\lambda)\}, \tag{3.19}$$

$$g = (2\pi u \eta(\beta))^{1/2} \exp\{-(k - u\theta(\beta))^2 / 2u\eta(\beta)\}. \tag{3.20}$$

We have

$$V_1 \leqslant \Sigma^{(1)} \,|p(r, c_n, k) - f(r, n, k)| + \Sigma^{(2)}\{p(r, c_n, k) + f(r, n, k)\},$$

where $\Sigma^{(1)}$ sums over $|k - c_n\theta(\lambda)| \leqslant K_2 c_n^{1/2}$ and $\Sigma^{(2)}$ over $|k - c_n\theta(\lambda)| > K_2 c_n^{1/2}$. To the first term we apply the local central limit theorem [7, Chap. II.2, Theorem 2] for $p(r, c_n, k)$. Note that it applies to the number of empty urns. It shows that the first term is $O(c_n^{-1/2})$. By Chebyshev's inequality, (3.1), and (3.4)–(3.7),

$$\Sigma^{(2)} p(r, c_n, k) \leqslant \gamma_1 K_2^{-2} \eta(\lambda).$$

By comparison with the integral over the corresponding normal density we find

$$\Sigma^{(2)} f(r, c_n, k) \leqslant 2 - 2\Phi(\eta^{-1/2}(\lambda)(K_2 - c_n^{-1/2})),$$

where Φ is the standard normal distribution function. So

$$V_1 \leqslant \gamma_2(K_2)\, c_n^{-1/2} + \gamma_1 K_2^{-2} \eta(\lambda) + 2 - 2\Phi(\gamma_3 K_2). \tag{3.21}$$

In (3.18) we split the domain of summation into $|k - u\theta(\beta)| \leqslant K_2 u^{1/2}$ and $|k - u\theta(\beta)| > K_2 u^{-1/2}$, with $\beta = r/u$. Since $|u - c_n| < K_1 c_n$, it follows from (3.9), (2.3), (2.4) that $0 < \gamma_4(K_1) \leqslant \beta \leqslant \gamma_5(K_1) < \infty$. We apply the local central limit theorem cited above to $p(r, u, k)$ and, since this theorem holds uniformly with respect to β if $0 < a \leqslant \beta \leqslant b < \infty$, we find in the same way as in (3.21)

$$V_3 \leqslant \gamma_6(K_1, K_2)\, u^{-1/2} + \gamma_7 K_2^{-2} \eta(\beta) + 2 - 2\Phi(\eta^{-1/2}(\beta)(K_2 - u^{-1/2})).$$

Again using the fact that $|u - c_n| < K_1 \sigma_n$ and $r = \lambda c_n$ so that $\eta(\beta)$ is bounded, we see

$$V_3 \leqslant \gamma_7(K_1, K_2)\, c_n^{-1/2} + \gamma_8 K_2^{-2} + 2 - 2\Phi(\gamma_9 K_2). \tag{3.22}$$

Comparing (3.17) with the L_1 distance between the corresponding normal densities and noting that $|u - c_n| < K_1 \sigma_n$, we may show that

$$V_2 \leqslant \gamma_{10}(K_1)\, \varepsilon_n, \qquad \varepsilon_n \to 0, \quad n \to \infty. \tag{3.23}$$

From (3.11), (3.13)–(3.15), (3.21)–(3.23), noting that the bounds in (3.21)–(3.23) do not depend on u, we see by first choosing K_1 and K_2 sufficiently large, that $\|G_{nr} - H_{nr}\| < \varepsilon$ for $n \geqslant n(\varepsilon)$.

Finally let us only assume $r = O(a_n)$. By (3.9) and (2.3) there is an integer λ such that $r \leqslant R = \lambda c_n \leqslant n$ for $n \geqslant n_1$. Define the map $\psi\colon \Pi_R \to \Pi_r$ so that $\psi(\pi) = \pi \cap \mathbb{N}_r$. By Lemma 1

$$\|G_{nR}\psi^{-1} - H_{nR}\psi^{-1}\| \leqslant \|G_{nR} - H_{nR}\|. \tag{3.24}$$

But $G_{nR} = P_n \chi^{-1}$, where $\chi\colon \Pi_n \to \Pi_R$ is defined by $\chi(\pi^1) = \pi^1 \cap \mathbb{N}_R$. Since $\psi \circ \chi(\pi^1) = \pi^1 \cap \mathbb{N}_r$, we have $G_{nr} = G_{nR}\psi^{-1}$. Since H_{nR} is the probability distribution of the partion of $\mathbb{N}_R$ generated by throwing the elements of $\mathbb{N}_R$ at random into c_n urns, $H_{nR}\psi^{-1}$ is the distribution of the partition of $\mathbb{N}_r$ generated by putting the elements of $\mathbb{N}_r$ at random into these urns, i.e., $H_{nR}\psi^{-1} = H_{nr}$. Relation (3.10) now follows from (3.24) and the first part of the proof.

REFERENCES

1. E. R. CANFIELD, Application of the Berry–Esséen inequality, *J. Combin. Theory Ser. A* **28** (1980), 17–25.
2. L. COMTET, "Analyse Combinatoire," Tome II, Presses Univ. France, Paris, 1970.
3. J. HAIGH, Random equivalence relations, *J. Combin. Theory Ser. A* **13** (1972), 287–295.
4. L. H. HARPER, Stirling behaviour is asymptotically normal, *Ann. Math. Statist.* **38** (1967), 410–414.
5. N. L. JOHNSON AND S. KOTZ, "Urn Models and Their Application," Wiley, New York/London, 1977.
6. V. F. KOLCHIN, A problem of the allocation of particles in cells and cycles of random permutations, *SIAM Theory Probab. Appl.* **16** (1971), 74–90.
7. V. F. KOLCHIN, B. A. SEVAST'YANOV, AND V. P. CHISTYAKOV, "Random Allocations," Winston, Washington, D.C., 1978.
8. L. LE CAM, On the distribution of sums of independent random variables, *in* "Bernoulli, Bayes, Laplace Anniversary Volume," pp. 179–202, Springer, Berlin, 1965.
9. L. MOSER AND M. WYMAN, An asymptotic formula for the Bell numbers, *Trans. Roy. Soc. Canad.* (3) **49** (1955), 49–54.
10. J. NEVEU, "Mathematical Foundations of the Calculus of Probability," Holden–Day, San Francisco, 1965.
11. YU. V. PROHOROV, Asymptotic behaviour of the binomial distribution, *Uspehi Mat. Nauk.* **8** (1953), 135.
12. V. N. SACHKOV, Random partitions of sets, *SIAM Theory Probab. Appl.* **19** (1974), 184–190.
13. V. N. SACHKOV, "Combinatorial Methods in Discrete Mathematics," Nauka, Moscow, 1977. [Russian]
14. V. N. SACHKOV, "Probabilistic Methods in Combinatorial Analysis," Nauka, Moscow, 1978. [Russian]
15. G. SZEKERES AND F. E. BINET, Notes on Borel fields over finite sets, *Ann. Math. Statist.* **28** (1957), 494–498.

Printed in Belgium