JOURNAL OF COMBINATORIAL THEORY, Series A 27, 367–370 (1979)

# Note

## The Number of Dense Arrangements

J. GLAZ

*Department of Mathematics and Statistics, Case Western Reserve University, Cleveland, Ohio 44106*

*Communicated by Gian-Carlo Rota*

Received March 1, 1979

Consider arrangements of $N$ elements of two kinds, $A$ and $B$, such that any segment of length $m$ contains at least $k$ type $A$ elements. We evaluate the number of such arrangements.

## 1. INTRODUCTION AND SUMMARY

The purpose of this paper is to illustrate the use of Karlin–McGregor's Theorem [6], on coincidence probabilities, to an enumeration problem with $N$ elements of two kinds.

Suppose we have $M_1$ type $A$ elements and $M_2$ type $B$ elements, $M_1 + M_2 = N$. Consider arrangements that have at least $k$ $A$ elements in any segment of length $m$, where by a segment of length $m$ we mean a consecutive string of $m$ elements. In the next section we evaluate the number of such arrangements.

Karlin and McGregor's result has been applied to solve various probabilistic problems: $L$-candidate ballot problems—Barton and Mallows [1] and Naus [8]; generalized birthday probabilities—Huntington [3], Naus [8], and Saperstein [9], [10]; clustering probabilities—Huntington and Naus [7], Wallenstein and Naus [11], [12]; multiple coverage probabilities—Glaz and Naus [2].

It is our hope that this method could be applied to other enumeration problems.

## 2. THE MAIN RESULT

Represent each of the $\binom{N}{M_1}$ possible arrangements of the $N$ given elements by a vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)$, where $x_i = 1$ (respectively, $x_i = 0$) if the $i$th element in the arrangement is an $A$ (respectively, $B$). Let $S_{k,m}$ (respectively,



0097-3165/79/060367-04$02.00/0

Copyright © 1979 by Academic Press, Inc.
All rights of reproduction in any form reserved.

$T_{h,m}$) be the set of all arrangements in which any segment of length $m$ contains at least $k$ (respectively, at most $h$) $A$ (respectively, $B$) elements. Then,

$$S_{k,m} = \left\{\mathbf{x};\ \inf_{1\leqslant n\leqslant N-m+1} \sum_{i=n}^{n+m-1} x_i \geqslant k\right\} \tag{1}$$

and

$$T_{h,m} = \left\{\mathbf{x};\ \sup_{1\leqslant n\leqslant N-m+1} \sum_{i=n}^{n+m-1} (1 - x_i) \leqslant h\right\}. \tag{2}$$

LEMMA 1. *For* $1 \leqslant k \leqslant m \leqslant N$ *integers*, $S_{k,m} = T_{m-k,m}$.

*Proof.* Follows from Eqs. (1) and (2).

Let $\partial_{k,m}$ and $\eta_{k,m}$ denote the cardinalities of the sets $S_{k,m}$ and $T_{k,m}$, respectively. It follows from Lemma 1 that $\partial_{k,m} = \eta_{m-k,m}$.

Let $b = N - Lm$, where $L$ is the largest integer in $N/m$. Subdivide the $N$ positions in the arrangements into $L + 1$ parts, $I_i$, $i = 1,..., L + 1$: $I_i = \{(i-1)m + 1, (i-1)m + 2,..., im\}$, for $i = 1, 2,..., L$; and $I_{i+1} = \{Lm + 1, Lm + 2,..., Lm + b = N\}$. Let $n_i$ denote the number of $B$ elements in the $i$th part.

THEOREM 1. *For* $1 \leqslant k \leqslant m$ *and* $M_1 \geqslant kL + \max(0, k + b - m)$ *integers,*

$$\partial_{k,m} = \sum_{P1} \det(c_{ij}), \tag{3}$$

*where for* $1 \leqslant i \leqslant L + 1$

$$\begin{aligned} c_{ij} &= \sum_{t=\beta_{L+1}+1}^{\beta_j} \binom{b}{t - \alpha_i}\binom{m-b}{\beta_j - t}, & j &= 1,..., L \\ &= \binom{b}{\beta_{L+1} - \alpha_i}, & j &= L + 1, \end{aligned} \tag{4}$$

*and*

$$\begin{aligned} \alpha_i &= (L + 1 - i)(m - k + 1) - \sum_{r=i}^{L} n_r, & i &= 1,..., L \\ &= 0, & i &= L + 1, \end{aligned} \tag{5}$$

$$\beta_j = \alpha_j + n_j, \qquad j = 1,..., L + 1. \tag{6}$$

$P1$ *is the set of all partitions of* $M_2$ *into* $L + 1$ *integers* $n_i$ *satisfying* $n_i \leqslant m - k$, *for* $i = 1, 2,..., L$; *and* $n_{L+1} \leqslant \min(b, m - k)$.

*Proof.* Let $n_i$ be the number of $B$ elements in $I_i$ for $i = 1,..., L + 1$ and $\{n_1,..., n_{L+1}\} \in P1$. Following Huntington [3], Naus [8], and Saperstein [10] associate $L + 1$ paths that start at $(0, \alpha_i)$ and end at $(m_i, \beta_i)$ and have $n_i$ moves up and $m_i - n_i$ horizontal moves, where $\alpha_i$ and $\beta_i$ are given by equations (5) and (6), respectively; $m_i = m$ for $i = 1,..., L$ and $m_{L+1} = b$. Then $\eta_{m-k,m}$ is equal to the number of realizations of the $L + 1$ paths defined above, such that none of the paths ever coincide. Saperstein [10] shows that this is equal to the $\det(c_{ij})$, where $c_{ij}$ is given by Eq. (4). It follows from Lemma 1, that for fixed $\{n_1,..., n_{L+1}\} \in P1$, $\partial_{k,m} = \det(c_{ij})$. Summing over all possible $\{n_1,..., n_{L+1}\} \in P1$ we obtain the result given by Eq. (3).

For the case $M/m = L$, $L \geqslant 2$ an integer it follows from Naus [8] and Lemma 1,

THEOREM 2. *For $2 \leqslant k$, $M/m = L$, $L$ an integer, $L \geqslant 2$*

$$\partial_{k,m} = (m!)^L \sum_{P2} \det[1/d_{ij}!\,(m - d_{ij})!],$$

*where*

$$d_{ij} = (j - i)(m - k + 1) - \sum_{t=i}^{j-i} n_t + n_i, \qquad \text{for} \quad 1 \leqslant i < j \leqslant L,$$

$$= (j - i)(m - k + 1) + \sum_{t=j}^{i} n_t, \qquad \text{for} \quad 1 \leqslant j \leqslant i \leqslant L.$$

*$P2$ is the set of all partitions of $M_2$ into $L$ integers $n_i$ satisfying $n_i \leqslant m - k$, $i = 1,..., L$.*

Further simplification is obtained from Naus [8] for the special case of $N/m = L$, $L$ an integer greater than one, and $M_1 > N - 2(m - k)$.

THEOREM 3. *For $k \geqslant 2$, $N/m = L$, $L \geqslant 2$ and $M_1 \geqslant N - 2(m - k + 1)$ integers,*

$$\partial_{k,m} = \binom{N}{M_2} - 2 \sum_{t=m-k+1}^{M_2} \binom{m}{t}\binom{N-m}{M_2 - t} + [L(m - k + 1) - M_2 - 1]\binom{m}{m - k + 1}\binom{N - m}{M_2 - m - k - 1}. \tag{7}$$

*Proof.* It follows from Naus [8, Corollary 2] that for $M_2 < 2h$

$$\eta_{h,m} = \binom{N}{M_2} - 2 \sum_{t=k+1}^{M_-} \binom{m}{t}\binom{N-m}{M_2 - t} + [L(k + 1) - M_2 - 1]\binom{m}{k + 1}\binom{N - m}{M_2 - k - 1}. \tag{8}$$

Using the identity $\partial_{k,m} = \eta_{m-k,m}$ and Eq. (8), we obtain the result given by Eq. (7).

## REFERENCES


1. D. E. BARTON AND C. L. MALLOWS, Some aspects of the random sequence, *Ann. Math. Statist.* **36** (1965), 236–260.
2. J. GLAZ AND J. NAUS, Multiple coverage of the line, *Ann. Probability* **7** (1979), in press.
3. R. J. HUNTINGTON, "Distributions and expectations for clusters in continuous and discrete cases, with Applications," Ph.D. thesis, Rutgers University, 1974.
4. R. J. HUNTINGTON AND J. I. NAUS, A simpler expression for $k$-th nearest neighbour coincidence probabilities, *Ann. Probability* **3** (1975), 894–896.
5. F. K. HWANG, A generalization of the Karlin–McGregor theorem on coincidence probabilities and an application to clustering, *Ann. Probability* **5** (1977), 814–817.
6. S. KARLIN AND G. MCGREGOR, Coincidence probabilities, *Pacific J. Math.* **9** (1959), 1141–1164.
7. J. I. NAUS, The distribution of the size of the maximum cluster of points on a line, *J. Amer. Statist. Assoc.* **60** (1965), 532–538.
8. J. I. NAUS, Probabilities for the generalized birthday problem, *J. Amer. Statist. Assoc.* **69** (1974), 810–815.
9. B. SAPERSTEIN, The generalized birthday problem, *J. Amer. Statist. Assoc.* **67** (1972), 425–428.
10. B. SAPERSTEIN, Note on a clustering problem, *J. Appl. Probability* **12** (1975), 629–632.
11. S. R. WALLENSTEIN AND J. I. NAUS, Probabilities for a $k$-th nearest neighbour problem on the line, *Ann. Probability* **1** (1973), 188–190.
12. S. R. WALLENSTEIN AND J. I. NAUS, Probabilities for the size of largest clusters and smallest intervals, *J. Amer. Statist. Assoc.* **69** (1974), 690–697.