The Annals of Mathematical Statistics
1972, Vol. 43, No. 3, 988-996

A MATRIX OCCUPANCY PROBLEM[1]

BY PATRICK J. EICKER, M. M. SIDDIQUI
AND PAUL W. MIELKE, JR.

Colorado State University

1. Introduction. Suppose that for $i = 1, \cdots, s$, x_i balls, $1 \leqq x_i \leqq T$, are randomly placed in the ith row of an $s \times T$ matrix. It is assumed that: (i) each cell contains at most one ball, and (ii) the balls in any row are distributed independently of the balls in any other row or set of rows. A column of the matrix will be said to have weight j, $0 \leqq j \leqq s$, if exactly j of the s cells in that column are occupied, the other s-j cells being empty. Let C_j, $j = 0, 1, \cdots, s$, denote the number of columns with weight j. The exact (univariate) probability distribution of C_s, the number of full columns, was first given by Mielke and Siddiqui (1965) after being implicitly used by Cowan *et al.* (1963) to describe a temporal attack pattern of three asthmatics. The purpose of the present paper is to study the probability distribution of the basic variables associated with the matrix from which the (multivariate) distribution of the vector $C = (C_0, \cdots, C_s)$ can be derived easily.

We construct an $s \times T$ matrix corresponding to the original matrix replacing "a ball" by a one and "empty" by a zero. Any column of the new matrix is a permutation of j zeros and s-j ones for some $j = 0, 1, \cdots, s$. The entire discussion which follows is in terms of the new matrix. We shall refer to a cell with a zero in it as a 0-cell, and a cell with a one in it as a 1-cell.

Let

$$\mathscr{P} = \{p : p = (a_1, \cdots, a_s), a_j = 0 \quad \text{or} \quad 1, j = 1, \cdots, s\}$$
$$\mathscr{P}_i = \{p \in \mathscr{P} : p = (a_1, \cdots, a_s), \textstyle\sum_j a_j = i\}, \qquad i = 0, 1, \cdots, s.$$

There are 2^s elements in $\mathscr{P}$ and $\binom{s}{k}$ elements in $\mathscr{P}_k$. Note that $\mathscr{P}_0$ and $\mathscr{P}_s$ each have only one element which we denote as p_0 and p_s respectively. For each $p \in \mathscr{P}$, let N_p be the number of columns with structure p, then

$$\eta = \{N_p : p \in \mathscr{P}\}$$

is a set of 2^s random variables. Also,

$$C_j = \textstyle\sum_{p \in \mathscr{P}_j} N_p, \qquad j = 0, 1, \cdots, s,$$

is the number of columns with weight j. For example, if $s = 2$, then

$$\eta = \{N_{00}, N_{01}, N_{10}, N_{11}\} \qquad \text{and} \qquad C_0 = N_{00}, \quad C_1 = N_{10} + N_{01}, \quad C_2 = N_{11}.$$

In Section 2 we obtain the exact distribution of η. In later sections we investigate the asymptotic properties of this distribution under various assumptions on the behavior of x_i's as $T \to \infty$.

Received October 17, 1969; revised November 18, 1971.

[1] This research was supported in part by National Institutes of Health Grant 5 T1 GM 725.