SOME RESULTS ON RESTRICTED OCCUPANCY THEORY[1]

By John E. Freund and Arthur N. Pozner

Virginia Polytechnic Institute

1. Introduction. If k judges rate a product on a discrete scale, say $0, 1, 2, \cdots$, or m, it is not only important to know the average rating assigned to the product, but it is also important to know the consistency of the ratings. Whereas the average rating is r/k, where r is the total number of points assigned to the product by the k judges, it seems reasonable to use the variance of the ratings, i.e.,

$$s^2 = \sum_{i=0}^{m} \left(i - \frac{r}{k}\right)^2 \left(\frac{v_i}{k}\right), \tag{1}$$

as a measure of consistency. Here v_i stands for the number of judges who gave the product the rating i.

In order to test the consistency of such ratings, it will be necessary to find a suitable mathematical model which will assign low probabilities to cases in which the ratings are either very inconsistent or overly consistent. It is felt that the appropriate model is provided by that of restricted occupancy theory, in which we consider as equally likely all possible distributions of r indistinguishable objects among k cells with at most m objects per cell. With this model we shall then test whether it is reasonable to suppose that the r points given by the k judges are randomly distributed among the k judges.

Related problems dealing with the probability that x cells contain more than q objects when r objects are distributed among k cells, and the probaility that x cells contain m objects when r objects are distributed among k cells with at most m objects per cell were investigated by Baticle, [1], [2], and [3], with reference to applications to casualty insurance, merchandizing, and transportation.

2. A restricted occupancy distribution. Restricted occupancy theory deals with the distribution of r objects among k cells if a maximum of m objects is permitted per cell, the cells are distinguishable, and empty cells are permitted. In this paper we shall investigate the distribution of the variables v_i $(i = 0, 1, 2, \cdots, m)$, standing for the number of cells occupied by i objects, respectively, if r indistinguishable objects are distributed among k cells with a maximum of m objects per cell.

Since r and k are assumed to be fixed, it should be noted that the variables v_i are subject to the two linear restrictions

$$\sum_{i=0}^{m} v_i = k \qquad \text{and} \qquad \sum_{i=0}^{m} i \cdot v_i = r, \tag{2}$$

and any $m - 1$ of the v_i will thus determine the remaining two.

Received December 30, 1954, revised November 16, 1955.

[1] Research sponsored by the Office of Ordnance, U.S. Army, Contract No. DA-36-034-ORD-1527RD.