MEASUREMENT OF DIVERSITY: MULTIPLE CELL CONTENTS

F. N. DAVID
UNIVERSITY OF CALIFORNIA, RIVERSIDE

1. Introduction

In a previous paper [1], we discussed the distribution of a score, representing a diversity index, when a number of different colored balls are randomly dropped into $M(=mn)$ identical compartments, one ball only permitted per compartment. We now generalize and at the same time simplify the previous setup by allowing more than one ball per box.

2. Notation

A box of $m \times n = M$ compartments is supposed. There will be $(m - 1)(n - 1)$ crossover points each of which will be surrounded by four compartments. Denote these crossover points by (ij), $i = 1, 2, \cdots, m - 1$, $j = 1, 2, \cdots, n - 1$. Let there be K_1 balls in s colors with k_ℓ the number of balls of the ℓth color and

$$\sum_{\ell=1}^{s} k_\ell = K_1. \tag{2.1}$$

These K_1 balls are supposedly dropped randomly into the M compartments, with no limitation on the individual compartment capacity. Consider the (ij)th crossover point. Let T_{ij} be the total number of balls in the four compartments surrounding (ij). Let $t_{ij\ell}$ be the total number of balls of the ℓth color in the same four compartments so that

$$T_{ij} = \sum_{\ell=1}^{s} t_{ij\ell}. \tag{2.2}$$

The number of joins between balls of like and unlike colors will be, omitting the factor of one half,

$$T_{ij}^{(2)} = \sum_{\ell=1}^{s} t_{ij\ell}^{(2)} + \sum_{\ell \neq h} t_{ij\ell} t_{ijh}, \tag{2.3}$$

for the four boxes. Summed for all values of i and j, we have

$$\sum_{i=1}^{m-1} \sum_{j=1}^{n-1} T_{ij}^{(2)} = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \sum_{\ell=1}^{s} t_{ij\ell}^{(2)} + \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \sum_{\ell \neq h} t_{ij\ell} t_{ijh}. \tag{2.4}$$

This investigation was partially supported by USPHS Research Grant No. GM-10525-0 8 National Institutes of Health, Public Health Service.