MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
STATISTISCHE AFDELING

Report S 266

On probability distributions arising from points on a graph

by

A.R. Bloemena

January 1960

The Mathematical Centre at Amsterdam, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government trough the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

Contents

1. Introduction

Given a set of n points, numbered $1,\ldots,n$, and a $n \times n$ matrix M, with elements m_{ij}, satisfying

(1.1) $m_{ij} = m_{ji}$ $(i \neq j)$,

(1.2) $m_{ii} = 0$,

(1.3) for each i $m_{ij} \neq 0$ for at least one j, and

(1.4) $0 \leqq m_{ij} < \infty$.

The set of points and the matrix M can be interpreted as a finite multigraph (cf. C. BERGE (1958), D. KOENIG (1936)), where the number of joins between point i and j is equal to m_{ij}. If $m_{ij} = 0$, this meansthat there is no join between i and j. Assumption (1.2) states that there are no loops. Assumption (1.3) implies that no point is isolated.

From the n points two samples are taken. We shall consider two cases.

Case I "non free sampling": from the points $1,\ldots,n$ r_1 and r_2 points are chosen at random without replacement $(r_1+r_2 \leqq n)$. The r_1 points will be denoted as black (B) points, the r_2 points as white (W) ones, while finally the $n-r_1-r_2$ remaining points are the red (R) ones.

Case II "free sampling": n independent trials are performed, each trial resulting in the event B with probability p_1, in the event W with probability p_2, and in the event R with probability $1-p_1-p_2$. Point number i is alotted the colour indicated by the outcome of the i-th trial.

Consider the random variables $\underline{x}_{ij}$ and $\underline{y}_{ij}$ $(i,j = 1,\ldots,n)$, defined by

$$\underline{x}_{ii} = 0 \quad \text{spr } 0,$$
$$\underline{y}_{ii} = 0 \quad \text{spr } 0,$$

and for $i \neq j$

$$\underline{x}_{ij} = \begin{cases} 1 & \text{if point i and j are both black} \\ 0 & \text{if not.} \end{cases}$$

$$\underline{y}_{ij} = \begin{cases} 1 & \text{if point i is black and j is white, or point i is white and j is black,} \\ 0 & \text{if not.} \end{cases}$$

$$(1.5) \qquad \underline{x} = \sum_{ij} m_{ij} \ \underline{x}_{ij},$$

$$(1.6) \qquad \underline{y} = \sum_{ij} m_{ij} \ \underline{y}_{ij}.$$

We shall also consider a more general situation. Let be given a set of random variables $\underline{z}_{ij}$, where $\underline{z}_{ii} = 0$ spr o,

while for $i \neq j$ $\underline{z}_{ij}$ is either 0 or 1. Define

$$(1.7) \qquad \underline{z} = \sum_{ij} m_{ij} \ \underline{z}_{ij}.$$

In the following we shall give results on the stochastic properties of $\underline{x}$, $\underline{y}$ and $\underline{z}$. The proofs of these results will be given in a forthcoming thesis.

2. Previous work on the subject

P.A.P. MORAN (1948) considers a „statistical map", equivalent to our graph for $m_{ij} = 0$ or 1, where the points are chosen by „free" and „non free" sampling. He gives for both cases the first and second moments of the number of black-black joins (thus for $\underline{x}$) and the third and fourth moment for the case of free sampling. He proves the asymptotic normality of $\underline{x}$ and $\underline{y}$ (free sampling for a rectangular twodimensional lattice, where there are joins between neighbouring points in the direction of both axis (cf. also P.A.P. MORAN (1947)).

There exists a large number of papers on the subject by P.V. KRISHNA IYER (1948-1953), most of them in an extremely-hard-to-get journal, viz. the Journal of The Indian Society for Agricultural Statistics. As far as we are aware, KRISHNA IYER only deals with rectangular lattices, where neighbouring points are joined in the direction of both axis, but also diagonal joins are considered in a number of his papers. The results of KRISHNA IYER are mostly on the first four moments or cumulants, and statements about asymptotic normality.

A report by BLOEMENA and van EEDEN contains a number of exact results for rectangular lattices (non free sampling). The present report is an outgrowth of this last paper, which arose from a study of the distribution of a statistic, obtained in a psychological test.

Some older papers on the subject are by H. TODD (1940) and D.J. FINNEY (1947).

3. Some graphtheoretical notions

Consider a set S of points and a subset U of the set of all joins between these points. The combination (S,U) is usually called a graph. For a detailed treatment of theory of graphs, we refer to D. KOENIG (1936) and C. BERGE (1958).

For our purpose we use the word "graph" to denote a set of k oriented joins, labelled $J_1, \ldots, J_k$, between ℓ ($2 \leq \ell \leq 2k$) points, such that no points are isolated (are not connected to at least one other point), and loops do not occur. Multiple joins are admitted.

A point to which join J_i is connected will be called the second point of J_i if the orientation of the join is towards the point; if not, it will be called the first point of J_i.

To each graph there corresponds a symmetrical 2k x 2k matrix A, consisting of k^2 2 x 2 block-matrices A_{ij} ($i,j = 1,\ldots,k$), with elements

for $\mu, \lambda = 1,2$

$$a_{i\mu, i\lambda} = 0,$$

and for $i \neq j$

$$a_{i\mu, j\lambda} = \begin{cases} 1 \text{ if the } \mu\text{-th point of } J_i \text{ coincides with the } \lambda\text{-th point of } J_j, \\ 0 \text{ if not} \end{cases}$$

All graphs having the same matrix A are considered to be equivalent. E.g. both

have as matrix

$$\left(\begin{array}{cc|cc|cc} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

and are therefore equivalent.

The k x k matrix with elements

$$b_{ij} = \sum_{\mu=1}^{2} \sum_{\lambda=1}^{2} a_{i\mu, j\lambda}$$

will be called the <u>configurationmatrix</u>.

Consider two graphs G_1 and G_2, each based on ℓ points and k (labelled and oriented) joins. If G_1 and G_2 are not identical, but a permutationmatrix P exists such that for the configuration-matrices B_1 and B_2 the relation

$$B = P\, B_2\, P'$$

holds, we shall say that G_1 and G_2 have the same <u>configuration</u>.

A graph $G = (S,U)$ is called connected if from every point $i \in S$ one can reach any other point of S by travelling along the joins of the set U, neglecting the orientation of the joins. A graph which is not connected, can be decomposed in a number of connected components. This decomposition is unique (cf. D.KÖNIG, 1936, p.15). A configuration-matrix of a not connected graph (if necessary after premultiplication with a permutationmatrix P, and postmultiplication with P') is a logical sum of the configuration-matrices of each of the connected components.

A connected graph with k joins has at most k + 1 points. It has at least two points For l satisfying

$$2 \leq l \leq k+1$$

finitely many, say q_{kl} , distinct configurations exists corresponding to connected graphs based on k joins and l points.
Let $C^{(\alpha)}_{k,l}$ be the α-th one ($\alpha = 1, \ldots, q_k$).
The configuration of a graph having h connected components $(1 \leq h \leq [\frac{l}{2}])$ can now be indicated symbolically by

$$\sum C^{(\alpha_i)}_{k_i, l_i}$$

if the i – th connected component has a configuration $C^{(\alpha_i)}_{k_i, l_i}$.
If among the connected components g_j have the same configuratio

$$C^{(\alpha_j)}_{k_j, l_j} \qquad \text{we may also write} \qquad \sum g_j\, C^{(\alpha_j)}_{k_j, l_j}$$

as the symbool for the configuration of the graph.

By means of the operator $\mathcal{N}(\quad)$, operating on the symbol of a configuration we indicate the number of distinct graphs, having this configuration. It can be proved that if $\sum_i k_i g_i = k$,

$$(3.1)\quad \mathcal{N}\left(\sum_{i=1}^{s} g_i\, C^{(\alpha_i)}_{k_i, l_i}\right) = k! \prod_{i}^{s} \frac{1}{g_i!} \left\{ \frac{\mathcal{N}(C^{(\alpha_i)}_{k_i, l_i})}{k_i!} \right\}^{g_i}.$$

The calculation of $\mathcal{N}(C_{k_i,\ell_i}^{(\alpha_i)})$ proceeds by means of recurrence relations.

4. A ~~general~~ expression for the moments of $\underline{z}$

In order to calculate the k-th moment of $\underline{z}$, we have to consider products like

(4.1) $$m_{i_{1,1}\,i_{1,2}}\,m_{i_{2,1}\,i_{2,2}},\ldots,m_{i_{k,1}\,i_{k,2}}$$

and

(4.2) $$\underline{z}_{i_{1,1}\,i_{1,2}}\,\underline{z}_{i_{2,1}\,i_{2,2}},\ldots,\underline{z}_{i_{k,1}\,i_{k,2}}\quad,$$

where $i_{1,1}\;,\ldots,i_{k,2}$ are, say, l different integers from the range $1,\ldots,n$. To each such products there corresponds a graph. Let each of the subscripts of (4.2) correspond to a point of the graph. If two or more subcripts are equal, they correspond to a same point, thus the graph has l different points in all. Let the first subscript, i_{j1}, of $\underline{z}_{i_{j,1}\,i_{j,2}}$ corresponds to the first point of a join, and the second subscript, i_{j2}, to the second point of the same join. We thus obtain a graph with k oriented joins and l points, no one point being isolated. We assume that always $i_{ji} \neq i_{j2}$ $(j=1,\ldots,k)$, thus no loops arise. Let the graph corresponding to (4.2) have a configuration

$$\sum_{i=1}^{k} C_{k_i,\ell_i}^{(\alpha_i)}\;,$$

then the following assumption on the simultaneous distribution of the $\underline{z}_{ij}$ $(i \neq j)$ is introduced:

Assumption A1

For each $k = 1,\ldots$ the expectation of (4.2) does not depend on the actual value of $i_{1,1},\ldots,\, i_{k,2}$, but only on the configuration $\sum_{i=1}^{h} C_{k_i,\ell_i}^{(\alpha_i)}$.

We therefore introduce the following notation for the expectation of (4.2):

(4.3) $$E_{\left[\sum_{i=1}^{h} C_{k_i,\ell_i}^{(\alpha_i)}\right]}\,\underline{z}^{(1)}\cdots\underline{z}^{(k)}.$$

To indicate the sum over a product of k coefficients $m_{i_{j,1}, i_{j,2}}$ $(i_{j,1} \neq i_{j,2})$, where exactly l out of the 2k subscripts are different, and in such a way, that the graph corresponding to this product has configuration $\sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)}$ we write

$$\sum \lfloor i_1 \neq \cdots \neq i_j,\ i_{j+1}, \ldots, i_l;\ \sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)} \rfloor\ m^{(1)} \cdots m^{(k)},$$ ~~we write~~

if the condition on the subscripts is that in the summation $i_1, \ldots, i_j$ have to be different. Summation over $i_1, \ldots, i_l$ extends from $1, \ldots, n$.

Now one can derive

$$(4.4)\quad E\underline{z}^k = \sum_{l=2}^{2k} \sum_{h=1}^{[\frac{l}{2}]} \sum{}'' \lfloor \sum_{i=1}^{h} k_i = k,\ \sum_{i=1}^{h} l_i = l,\ \alpha_i \rfloor\ \mathcal{N}\Big(\sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)}\Big)\ .$$

$$.\ E \lfloor \sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)} \rfloor\ \underline{z}^{(1)} \cdots \underline{z}^{(k)}\ .\ \sum \lfloor i_1 \neq \cdots \neq i_l;\ \sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)} \rfloor\ m^{(1)} \cdots m^{(k)},$$

where $\sum'' \lfloor \sum_{i=1}^{h} k_i = k,\ \sum_{i=1}^{h} l_i = l,\ \alpha_i \rfloor$ means summation over all configurations with $\sum k_i = k$ and $\sum l_i = l$.

5. The moments of $\underline{x}$

a) non free sampling

Here assumption A 1 is satisfied, as

$$(5.1)\quad E \lfloor \sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)} \rfloor\ \underline{x}^{(1)} \cdots \underline{x}^{(k)} = \frac{\binom{r}{l}}{\binom{n}{l}},$$

where $\sum l_i = l$.

(We omit the subscript on r_1 and p_1, as no danger of confusion arises in the sections on $\underline{x}$.)

From (4.4) we have e.g. after some simplifications

$$E\underline{x} = \frac{r(r-1)}{n(n-1)} \sum_{i,j} m_{ij},$$

$$\sigma^2 = E\underline{x}^2 - (E\underline{x})^2 = 4\,\frac{r(r-1)\cdot(r-2)\cdot(n-r)}{n(n-1)\cdot(n-2)\cdot(n-3)} \sum_i \Big(\sum_j m_{ij} - \frac{\sum_{ij} m_{ij}}{n}\Big)^2$$

$$+ \frac{2r(r-1)\cdot(n-r)\cdot(n-r-1)}{n^2(n-1)^2(n-2)(n-3)} \Big\{ n(n-1)\sum_{ij} m_{ij}^2 - \Big(\sum_{ij} m_{ij}\Big)^2 \Big\}.$$

If $\sum_j m_{ij}$ does not depend on i, the first term of σ^2 is equal to zero. The third reduced moment and the fourth unreduced moment have been calculated as well.

b) <u>free sampling</u>

(5.2) $$E\Big[\sum_{i=1}^{h} C^{(\alpha_i)}_{k_i,\ell_i}\Big] \underline{x}^{(1)} \cdots \underline{x}^{k} = p^{\ell},$$

so e.g.

$$E\,\underline{x} = p^2 \sum_{ij} m_{ij},$$

$$\sigma^2 = 2\,p^2\,(1-p)\Big\{(1-p)\sum_{ij} m_{ij} + 2\,p \sum_{ijk} m_{ij}\, m_{ik}\Big).$$

6. The moments of $\underline{y}$.

Assumption A_1 is satisfied. In order to calculate

$$E\Big[\sum_{i=1}^{h} C^{(\alpha_i)}_{k_i,\ell_i}\Big]\, \underline{y}^{(1)} \cdots \underline{y}^{(k)}$$

we first take a point P_i of the i-th connected component (i=1,...,h) as a reference point. Colour P_i white, next all points connected by a join to P_i are coloured black, then all points connected to these black points are coloured white. If in repeating this procedure one arrives at a point which has already been given one colour, but should be coloured by the just-mentioned rule in the other colour as well, then we conclude that the i-th connected component is <u>not bichromatic</u>.

If no such situation arises one arrives at a stage, where all points have been allotted a colour, viz τ_i points are white and $l_i - \tau_i$ black, we then say that the i-th connected component is <u>bichromatic</u>.

Define

$$(6.1)\qquad \mathcal{B}\left(\sum_i C_{k_i,\ell_i}^{(\alpha_i)}\right) = \begin{cases} 1 & \text{if all connected components of } \sum_i C_{k_i,\ell_i}^{(\alpha_i)} \text{ are bichromatic} \\ 0 & \text{if not.} \end{cases}$$

a) non-free sampling

$$(6.2)\quad E\Big[\sum_{i=1}^{h} C_{k_i,\ell_i}^{(\alpha_i)}\Big]\, \underline{y}^{(1)}\cdots\underline{y}^{(k)} = \mathcal{B}\left(\sum C_{k_i,\ell_i}^{(\alpha_i)}\right) \frac{(n-\ell)!}{n!} \sum_{i=1}^{h} \sum_{\rho_i=0}^{1} \cdot$$

$$\frac{r_1!\; r_2!}{\left(r_1-\sum_i(1-\rho_i)\tau_i-\sum_i\rho_i(\ell_i-\tau_i)\right)!\left(r_2-\sum_i\rho_i\tau_i-\sum_i(1-\rho_i)(\ell_i-\tau_i)\right)!}\,.$$

E.g.

$$E\underline{y} = \frac{2r_1r_2}{n(n-1)} \sum_{ij} m_{ij},$$

$$\sigma^2 = \sum_i\Big(\sum_j m_{ij} - \frac{\sum_{ij} m_{ij}}{n}\Big)^2 \cdot \left\{\frac{r_1r_2(r_2-1)}{n(n-1)(n-2)} + \frac{r_2r_1(r_1-1)}{n(n-1)(n-2)} - \frac{4r_1r_2(r_1-1)(r_2-1)}{n(n-1)(n-2)(n-3)}\right\} +$$

$$+ 4\sum_{ij} m_{ij}^2 \left\{\frac{r_1r_2}{n(n-1)} - \frac{r_1r_2(r_2-1)}{n(n-1)(n-2)} - \frac{r_2r_1(r_1-1)}{n(n-1)(n-2)} + \frac{2r_1r_2(r_1-1)(r_2-1)}{n(n-1)(n-2)(n-3)}\right\} +$$

$$+ 4\Big(\sum_{i,j} m_{ij}\Big)^2 \frac{r_1r_2\,n(r_1+r_2+3) - r_1r_2(2r_1r_2+r_1+r_2+2)}{n^2(n-1)^2(n-2)(n-3)}\,.$$

b) free sampling

$$(6.3)\quad E\Big[\sum_{i=1}^{h} C_{k_i,\ell_i}^{(\alpha_i)}\Big]\, \underline{y}^{(1)}\cdots\underline{y}^{(k)} = \mathcal{B}\left(\sum_i C_{k_i,\ell_i}^{(\alpha_i)}\right) \prod_i^{h}\left(p_1^{\tau_i} p_2^{\ell_i-\tau_i} + p_1^{\ell_i-\tau_i} p_2^{\tau_i}\right).$$

E.g.

$$E\underline{y} = 2p_1p_2 \cdot \sum_{ij} m_{ij},$$

$$\sigma^2 = 4p_1p_2(p_1+p_2-4p_1p_2)\sum_{ijk} m_{ij}m_{ik} + 4p_1p_2(1-p_1-p_2+2p_1p_2)\sum_{ij} m_{ij}\,.$$

7. Tendency towards the normal distribution

The following theorem can be proved.

Theorem 7.1

If

$$(7.1)\qquad \left.\begin{array}{l} \sum_j m_{ij}^a = m_a, \\ m_a < \infty, \\ m_a \text{ independent of } n, \end{array}\right\} \text{ for all } a = 1, \dots,$$

and if r and n tend to infinity such that

$$\lim \frac{r}{n} = o, \text{ with } 0 < \delta < 1,$$

then in the non-free sampling case the distribution of

$$\frac{\underline{x} - E\underline{x}}{\sigma}$$

tends to the standard normal one. $E\underline{x}$ and σ^2 have been given in section

If (7.1) is satisfied and n tends to infinity and p tends to a limit p^* $(0 < p^* < 1)$ then in the case of free sampling the distribution of

$$\frac{\underline{x} - E\underline{x}}{\sigma}$$

tends to the standard normal one. $E\underline{x}$ and σ^2 have been given in section 5b.

If (7.1) is satisfied, and if r_1, r_2 and n tend to infinity such that

$$\left.\begin{aligned} \lim \frac{r_1}{n} &= \delta_1 \\ \lim \frac{r_2}{n} &= \delta_2 \end{aligned}\right\} \qquad \begin{aligned} 0 < \delta_1,\ 0 < \delta_2, \\ \delta_1 + \delta_2 < 1, \end{aligned}$$

then the distribution of

$$\frac{\underline{y} - E\underline{y}}{\sigma}$$

tends to the standard normal one. $E\underline{y}$ and σ^2 have been given in section 6a.

If (7.1) is satisfied, and if n tends to infinity and p_1 and p_2 tend to finit limits p_1^*, and p_2^* $(0 < p_1^*;\ 0 < p_2^*,\ p_1^* + p_2^* < 1)$ then the distribution of

$$\frac{\underline{y} - E\underline{y}}{\sigma}$$

tends to the standard normal one. $E\underline{y}$ and σ^2 have been given in section 6b.

8. Tendency towards the compound POISSON distribution

A theorem on the tendency towards the compound POISSON distribution has been proved for $\underline{z}$. The asymptotic behaviour of $\underline{x}$ and $\underline{y}$ for both the free and the non-free sampling case can be considered as a special case of this theorem. First we introduce some assumptions.

Assumption A1

(section 4) is considered to be satisfied.

Assumption A2

Consider an event $\mathcal{Z}$: $\underline{z}^{(1)}=1 \cap \ldots \cap \underline{z}^{(k)}=1$, where the subscripts of the $\underline{z}$'s correspond to $\sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)}$.

Consider the event $\mathcal{Z} \cap \underline{z}_{ij}=1$, where the configuration is now $\sum_{i=2}^{h} C_{k_i,l_i}^{(\alpha_i)} + C_{k_1+n_1,l_1+1}^{(\beta)}$.

Then if $P[\mathcal{Z}]>0$

$$P[\underline{z}_{ij}=1 \mid \mathcal{Z}] \leqq P[\underline{z}_{ij}=1 \mid \underline{z}_{i1}=1] .$$

This assumption is supposed to be satisfied for k= 1,...

Assumption A3

If in a configuration $\sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)}$, with $\sum_{i=1}^{h} k_i=k$, $\sum_{i=1}^{h} l_i=l$, a point is made to coincide with another one to which it is not connected by a join, thus giving rise to a configuration $\sum_{i=1}^{h'} C_{k_i',l_i'}^{(\alpha_i')}$, with $\sum k_i = k$, $\sum l_i = l-1$.

then if

$$E\left[\sum_{i=1}^{h'} C_{k_i',l_i'}^{(\alpha_i')}\right] \underline{z}^{(1)} \cdots \underline{z}^{(k')} > 0$$

also

$$E\left[\sum_{i=1}^{h} C_{k_i,l_i}^{(\alpha_i)}\right] \underline{z}^{(1)} \cdots \underline{z}^{(k)} > 0 .$$

Assumption A4

If for a configuration $C_{k,l}^{(\alpha)}$

$$E[C_{k,l}^{(\alpha)}]\, \underline{z}^{(1)} \cdots \underline{z}^{(k)}=0$$

then one can find joins such that if they are removed from $C_{k,l}^{(\alpha)}$, a configuration $C_{k',l}^{(\beta)}$ remains with

$$E[C_{k',l}^{(\beta)}]\, \underline{z}^{(1)} \cdots \underline{z}^{(k')} > 0$$

We now state

Theorem 8.1

Let $\underline{z}$ be defined by (1.7), where m_{ij} (i,j=1,...n) satisfies (1.1) through (1.4) and let assumptions $A_1,\ldots,A_4$ be satisfied.

If n tends to infinity and if for every i, j, μ, λ, with $i \neq j$, μ, λ and $\mu \neq \lambda$

$$\lim P\left[\underline{z}_{ij}=1 \mid \underline{z}_{\mu\lambda} =1\right] =0,$$

such that

(8.1) $$\lim E\, \underline{z} = 2\lambda, \quad 0<\lambda<\infty,$$

and

(8.2) $$\lim E\left[C^{(\alpha)}_{k,k+1} \right] \underline{z}^{(1)} \cdots \underline{z}^{(k)} \sum\left[i_1, \ldots, i_{k+1}; C^{(\alpha)}_{k,k+1} \right] m \cdots m = 0$$

for all k= 2,... , and all α,

(8.3)
$$\lim E\left[\sum_{i=1}^{h} C^{(\alpha_i)}_{k_i,l_i} \right] \underline{z}^{(1)} \cdots \underline{z}^{(k)} = \lim \prod_{i=1}^{h} E\left[C^{(\alpha_i)}_{k_i,l_i} \right] \underline{z}^{(k_1+\cdots+k_{i-1}+1)} \cdots \underline{z}^{(k_1+\cdots+k_i)},$$

(where $\sum_i k_i = k$, $\sum_i l_i = l$) for all $h= 2,\ldots,\left[\frac{l}{2}\right]$, $l=3,\ldots,2k$ and all $k= 2\ldots$,

(8.4)
$$\lim \frac{\sum_{ij} m_{ij}^{h}}{\sum_{ij} m_{ij}} = m_h^*, \quad m_h^* < \infty, \quad h = 1, \cdots \quad (m_1^* = 1),$$

and finally,

(8.5) $m_{ij} < M$ for all i and j, where M does not depend on n, then

(8.6) $$\sum_{k=1}^{\infty} \frac{X^k}{k!} \lim E\left(\tfrac{1}{2}\underline{z}\right)^k = \exp\left\{ \lambda \sum_{i=1}^{\infty} m_i^* \frac{X^i}{i!} \right\} - 1,$$

which is the moment-generating function of the component POISSON distribution.

If m_{ij} is either 0 or 1, and if assumptions A_1, A_2, A_3, A_4, are satisfied, together with (8.1), (8.2), and (8.3), then $\frac{1}{2}\underline{z}$ has asymptotically a POISSON distribution with parameter λ.

Assumption (8.2) is satisfied e.g. when $\sum_j m_{ij} = m_1$, independent of i, and

(8.2') $$\lim \frac{E\left[C^{(\alpha)}_{k,k+1} \right] \underline{z}^{(1)} \cdots \underline{z}^{(k)}}{E\left[C^{(1)}_{1,2} \right] \underline{z}^{(1)}} m_1^{k} = 0 \quad \text{for all } k \text{ and } \alpha.$$

E.g when m_1 is a constant independent of n then (8.2') and therefore (8.2) is satisfied.

First special case of Theorem (8.1)

non free sampling

If r and n tend to infinity such that

$$\lim \frac{r}{n} = 0,$$

(8.7) $\lim E\underline{x} = 2\lambda,\ 0<\lambda<\infty,$

(8.8) $\lim \left(\frac{r}{n}\right)^{k+1} \sum_{[i_1,\dots,i_{k+1};C_{k,k+1}^{(\alpha)}]} m^{(1)}\dots m^{(k)} = 0$

for each $k=2,\dots,$ and all α,

if moreover

free sampling

If n tends to infinity and p tends to zero such that

$$\lim E\,\underline{x} = 2\lambda,\quad 0<\lambda<\infty,$$

$$\lim p^{k+1} \sum_{[i_1,\dots,i_{k+1};C_{k,k+1}^{(\alpha)}]} m^{(1)}\dots m^{(k)} = 0$$

for all $k=2,\dots$ and all α,

if moreover

$$\lim \frac{\sum_{ij} m_{ij}^{h}}{\sum_{ij} m_{ij}} = m_h^{*},\quad m_h^{*}<\infty,\quad h=1,\dots\ (m_1^{*}=1),$$

and finally,

$$m_{ij} < M \text{ for all } i \text{ and } j,$$
$$M \text{ does not depend on } n,$$

then

$$\sum_{k=1}^{\infty} \frac{Z^k}{k!} \lim E\,(\tfrac{1}{2}\,\underline{x})^k = \exp\left\{\lambda \sum_{i=1}^{\infty} m_i^{*} \frac{Z^i}{i!}\right\} - 1.$$

If $m_{ij} = 0$ or 1 and if assumptions (8.7) and (8.8) are satisfied $\frac{1}{2}\,\underline{x}$ has a POISSON-distrubition with parameter λ.
Assumption (8.8) is satisfied e.g. when $\sum_j m_{ij} = m_1$ independent of i and $\lim \frac{r}{n} m_1 = 0$

Second special case of Theorem (8.2)

non free sampling

If r_1, r_2 and n tend to infinity such that

$$\lim \frac{r_1}{n} = 0$$

$$\lim \frac{r_2}{n} = 0,$$

(8.9) $\lim E\underline{y} = 2\lambda,\ 0<\lambda<\infty,$

free sampling

If n tends to infinity and p_1 and p_2 tend to zero, such that

$$\lim E\,\underline{y} = 2\lambda,\quad 0<\lambda<\infty,$$

and:

(8.10) $\lim E_{[C_{k,k+1}^{(\alpha)}]}\,\underline{y}^{(1)}\dots\underline{y}^{(k)} \sum_{[i_1,\dots,i_{k+1};C_{k,k+1}^{(\alpha)}]} m^{(1)}\dots m^{(k)} = 0$

for all $k=2,\dots$ and all ,

and if

$$\lim \frac{\sum_{i,j} m_{ij}^{h}}{\sum_{i,j} m_{ij}} = m_h^*, \quad m_h^* < \infty, \quad h = 2, \cdots \quad (m_1^* = 1),$$

and finally,

$m_{ij} < M$, for all i and j , where M is independent of n,

then

$$\sum_{k=1}^{\infty} \frac{Z^k}{k!} \lim E\left(\tfrac{1}{2} \underline{y}\right)^k = \exp\left\{ \lambda \sum_{i=1}^{\infty} m_i^* \frac{Z^i}{i!} \right\} - 1.$$

If assumptions (8.9) and (8.10) are satisfied, and if m_{ij} is either 0 or 1 then $\frac{1}{2}\underline{y}$ has asymptotically a POISSON-distribution with parameter λ.

Assumption (8.10) is satisfied e.g. when $\sum_j m_{ij} = m_1$, independent of i and

$$\lim p_1 m_1 = 0,$$

$$\lim p_2 m_1 = 0.$$

C. BERGE (1958): Théorie des graphes et ses applications, Dunod, Paris.

A.R. BLOEMENA and CONSTANCE van EEDEN, (1959): On probability distributions arising from points on a lattice. Report S 257, Statistics Dept. Mathematical Centr

D.J. FINNEY (1947): The signifance of associations in a square point lattice, Journ. Roy.Stat.Soc. Suppl.(9),99

D. KOENIG (1936): Theorie der endlichen und unendlichen Graphen, Leipzig.

P.V. KRISHNA IYER (1947): Random association of points on a lattice, Nature, 160, 714.

P.V. KRISHNA IYER (1948): The theory of probability distributions of points on a line, JISAS 1, 173-195

P.V. KRISHNA IYER (1949): The first and second moments of some probability distributions arising from points on a lattice and their applications, Biometrika, 36, 135-141.

P.V. KRISHNA IYER (1950a) The theory of probability distributions of points on a lattice. Ann.Math.Stat. 21, 198-217 & 22, 310.

P.V. KRISHNA IYER (1950b,1951,1952) Further constributions to the theory of probability distributions of points on a line. JISAS, 2, 141-160; 3, 80-93; 4, 50-71.

P.A.P. MORAN (1947): Random associations on a lattice Proc.Cambr.Phil.Soc., 43, 321-328

P.A.P. MORAN (1948): The interpretation of Statistical maps. Journ.Roy.Stat.Soc. B 10, 243-251.

H. TODD (1940): Note on the random association in a square point lattice, Jour.Roy. Stat.Soc. Suppl. 7, 78.

JISAS = Journ. of the Indian Society for Agricultural Statistics.