STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

STATISTISCHE AFDELING

Report S 257

On probability distributions arising from points on a lattice

by

Dr Constance van Eeden
and
Ir A.R. Bloemena

September 1959

1. Introduction

In this report the following problem is treated:

Consider a lattice of $N = km$ points arranged in a rectangle of m columns and k rows ($k \leqq m$). From these N points n are chosen at random without replacement. In the following these points will be called black points, the other N-n points are called white points.

The problem to be considered concerns the joins between black points, where two points are said to be joined if they are adjacent in a horizontal, vertical or diagonal way.

In a rectangle the points are not equivalent. For k and $m \geqq 3$ $(m-2)(k-2)$ points have joins with 8 other points, $2(m+k-4)$ points have joins with 5 other points and 4 points have joins with 3 other points. Therefore we suppose the lattice to be wrapped around a torus and the ends joined. Then each point on the torus has joins with 8 other points. For k=1 with $m > 2$ and k=2 with $m > 2$ it is sufficient to wrap the rectangle around a cylinder and to join the smaller ends. Then each point on the cylinder has joins with 2 other points for k=1 and with 5 other points for k=2.

We now consider the random variable $\underline{x}$: the number of black-black joins.

Problems closely related to the above described one were considered by P.V. KRISHNA IYER (1949, 1950) and by P.A.P. MORAN (1947, 1948). KRISHNA IYER considered a rectangular lattice of m columns and k rows, where the ends are not joined. Concerning the choice of the black points he considers two cases:

1. the case where each point may be black or white independently with probability p and 1-p. Then the number of black points is a random variable $\underline{n}$,
2. the case described above, where a fixed number n of points is chosen at random.

Concerning the joins between the points KRISHNA IYER also considers two cases:

a. in his 1949-paper the case described above, where two points are said to be joined if they are adjacent in a horizontal, vertical or diagonal way,

b. in his 1950-paper the case where two points are said to be joined if they are adjacent in a horizontal or vertical way.[1)]

The random variables considered by KRISHNA IYER are the number of black-black joins and the number of black-white joins. In all cases he gives the first and second moment; in some cases also the third and fourth moment. Further he states, without a stringent proof, that all distributions tend to the normal if k and m both tend to infinity. Here he supposes, without clearly mentioning it, that in the case of a fixed number of black points

(1;1) $$0 < \liminf_{N\to\infty} \frac{n}{N} \leq \limsup_{N\to\infty} \frac{n}{N} < 1.$$

The asymptotic distribution of the number of black-black joins being the same for the case of a rectangle and a torus if k and m both tend to infinity, this (incomplete) proof of KRISHNA IYER suggests that the random variable $\underline{x}$ defined above has asymptotically, for k and $m \to \infty$, a normal distribution if (1;1) is satisfied.

P.A.P. MORAN considers an arbitrary lattice with a fixed or a stochastic number of black points. For both cases he gives the first and second moment of the number of black-black joins; for the case of a stochastic number of black points also the third and fourth moment. Further he proves the asymptotic normality for the case of a rectangular lattice with a stochastic $\underline{n}$ and with joins in a horizontal and vertical way.

The abovementioned papers are not the only ones concerning this subject. KRISHNA IYER also published some papers in the "Journal of the Indian Society of Agricultural Statistics". Further Prof. Dr. D. VAN DANTZIG drew our attention to the fact that the problem is related to the "order-disorder" problem in physics. So also in papers concerning this problem results may perhaps be found which are important for our problem.

The results obtained in this paper are

1. The exact distribution of $\underline{x}$ for

----

1) KRISHNA IYER also gives generalizations for more than two dimensions and more than two colours.

a) $k=2$ with $n = 2,3,$ and 4
b) $k \geqq 3$ with $n = 2$ and 3
c) $k = m = n = 4$,

2. The exact probabilities $P[\underline{x}=0]$ and $P[\underline{x}=1]$ for $k = 2$ and 3,
3. The first, second and third moment of $\underline{x}$,
4. The asymptotic distribution of $\underline{x}$ for $N \to \infty$ and $\lim_{N\to\infty} \frac{n}{N} = 0$.

   This asymptotic distribution is a Poisson-distribution.

2. Notation

The columns in the rectangle are numbered from 1 to m; the rows are numbered from 1 to k. The black points are numbered, in the order in which they are chosen, from 1 to n, whereas the pairs of black points are numbered in an arbitrary but fixed way, from 1 to $\binom{n}{2}$.

Unless explicitely stated otherwise

i	takes the values	$1,\ldots,m$,
j	" " "	$1,\ldots,k$,
λ and μ	take " "	$1,\ldots,n$,
ν, κ and γ	" " "	$1,\ldots,\binom{n}{2}$.

If the ν^{th} pair of black points consists of the points λ and μ with $\lambda < \mu$ then the point λ will be called the first and the point μ the second point of the ν^{th} pair.
A black point surrounded by white points only is called an <u>isolated black point</u>. Two adjoining black points are called a <u>joined pair</u> (of black points) irrespective of their other joins. An isolated joined pair of black points is called a <u>pair of twins</u>; a coherently joined set of h black points, which is isolated from all other black points is called a <u>run</u> of h.
All points being equivalent the first black point may be chosen on an arbitrary but fixed place and the distribution of $\underline{x}$ may be found by deriving this distribution under the condition that the first black point has this fixed position. Unless explicitely stated otherwise we derive the distribution of $\underline{x}$ and its moments in this way.

3. The distribution of $\underline{x}$ for k=1

For k=1 the exact distribution of $\underline{x}$ is known for each N and n. Therefore we only mention some wellknown properties of this distribution.

The exact distribution of $\underline{x}$ may e.g. be derived from the distribution of the number of runs of black points. If $\underline{r}$ is this number of runs then (cf. e.g. W.L. STEVENS (1939) and H.A. KUIPERS (1957)) 2)

$$(3;1) \qquad P[\underline{r}=r] = \frac{\binom{n}{r}\binom{N-n-1}{r-1}}{\binom{N-1}{n-1}} .$$

It may easily be seen that

$$(3;2) \qquad \underline{r} + \underline{x} = n$$

and from (3;1) and (3;2) it follows that

$$(3;3) \qquad P[\underline{x}=x\,|\,k=1] = \frac{\binom{n}{x}\binom{N-n-1}{n-x-1}}{\binom{N-1}{n-1}} .$$

Consequently the random variable $\underline{x}$ has, for k=1, a hypergeometric distribution with

$$(3;4) \qquad \begin{cases} \mu \overset{\text{def}}{=\!=} \mathcal{E}(\underline{x}\,|\,k=1) = \dfrac{n(n-1)}{N-1} , \\ \sigma^2 \overset{\text{def}}{=\!=} \sigma^2(\underline{x}\,|\,k=1) = \dfrac{n(n-1)(N-n)(N-n-1)}{(N-1)^2(N-2)} . \end{cases}$$

----

2) The distribution of $\underline{r}$ may also be derived from the distribution of the number of runs of black points in a row of points of length N', where the ends are not joined. If $\underline{r}'$ is this number of runs, then (cf. e.g. W. FELLER (1950), p.59, exercise 15)

$$P[\underline{r}'=r] = \frac{\binom{n-1}{r-1}\binom{N'-n-1}{r}}{\binom{N'}{n}} .$$

Then the distribution of $\underline{r}$ is obtained by substituting N-1 for N' (cf. e.g. H.A. KUIPERS (1957), p.5).

Now let $\underline{x}'$ denote the number of white-white joins. Then

(3;5) $\underline{x}' = N-n-\underline{r} = \underline{x}+N-2n.$

Consequently

(3;6) $$P[\underline{x}'=x|k=1] = \frac{\binom{N-n}{x}\binom{n-1}{N-n-x-1}}{\binom{N-1}{N-n-1}} .$$

We now consider the asymptotic distributions of $\underline{x}$ and $\underline{x}'$ for $N\to\infty$. The random variables $\underline{x}$ and $\underline{x}'$ having hypergeometric distributions we may use the well-known asymptotic properties of this distribution (cf. e.g. C. VAN EEDEN and J.Th. RUNNENBURG (1959)). Then we obtain the following results

1. if $\lim_{N\to\infty} \mathcal{E}(\underline{x}|k=1) = \lim_{N\to\infty} \frac{n^2}{N} = 0$ then

   (3;7) $\lim_{N\to\infty} P[\underline{x}=0|k=1] = 1.$

2. if $\lambda \overset{def}{=} \lim_{N\to\infty} \frac{n^2}{N}$ exists and $0<\lambda<\infty$, then the random variable $\underline{x}$ has asymptotically for $N\to\infty$ a Poisson distribution with parameter λ.

3. if $\lim_{N\to\infty} \sigma^2(\underline{x}) = \infty$ [3], then the random variables $\frac{\underline{x}-\mathcal{E}\underline{x}}{\sigma(\underline{x})}$ and $\frac{\underline{x}'-\mathcal{E}\underline{x}'}{\sigma(\underline{x}')}$ both have asymptotically for $N\to\infty$ a standard normal distribution.

----

3) If this condition is satisfied then one of the following conditions is satisfied

a) $\lim_{N\to\infty} \frac{n}{N} = 0, \quad \lim_{N\to\infty} \frac{n^2}{N} = \infty,$

b) $0 < \liminf_{N\to\infty} \frac{n}{N} \leq \limsup_{N\to\infty} \frac{n}{N} < 1,$

c) $\lim_{N\to\infty} \frac{n}{N} = 1, \quad \lim_{N\to\infty} \frac{(N-n)^2}{N} = \infty.$

4. if $\lambda' \overset{\text{def}}{=} \lim_{N\to\infty} \mathcal{E}(\underline{x}'|k=1) = \lim_{N\to\infty} \frac{(N-n)^2}{N}$ exists and $0<\lambda'<\infty$, then the random variable $\underline{x}'$ has asymptotically for $N\to\infty$ a Poisson distribution with parameter λ'.

5. if $\lim_{N\to\infty} \frac{(N-n)^2}{N} = 0$ then

$$(3.8) \qquad \lim_{N\to\infty} P\left[\underline{x}'=0|k=1\right] = 1.$$

We now consider the situation that x pairs of twins and n-2x isolated black points occur. This situation is denoted by S_x and we have

$$(3;9) \qquad P\left[\underline{x}=x|k=1\right] \geqq P\left[S_x|k=1\right].$$

The equality sign in (3;9) holds if and only if x=0 and x=1. Further we have (cf. W. FELLER (1950), p. 59, exercise 16[4])

$$(3;10) \qquad P\left[S_x|k=1\right] = \frac{\binom{n-x}{n-2x}\binom{N-n}{n-x}}{\binom{N-1}{n}}.$$

From (3;3) and (3;10) then follows

$$(3;11) \qquad \frac{P\left[S_x|k=1\right]}{P\left[\underline{x}=x|k=1\right]} = P\left[S_x|\underline{x}=x;k=1\right] = \prod_{i=1}^{x-1}\left(1-\frac{x}{n-i}\right).$$

Consequently, for each finite x, we have

$$(3;12) \qquad \lim_{N\to\infty} P\left[S_x|\underline{x}=x;k=1\right] = 1.$$

----

4) FELLER considers a row of N' points, where the ends are not joined. Then the probability for the black points to be arranged in r runs of which r_i are of length i $(i=1,\dots,\nu;\ \sum_{i=1}^{\nu} r_i=r,\ \sum_{i=1}^{\nu} ir_i=n)$ is

$$\frac{r!}{\prod_{i=1}^{\nu} r_i!} \quad \frac{\binom{N'-n+1}{r}}{\binom{N'}{n}}.$$

Substituting $\nu=2$, $r_1+r_2 = r = n-x$, $r_1+2r_2 = n$ and $N' = N-1$ (cf. footnote 2) we obtain (3;10).

This means that, if $\frac{n^2}{N}$ has a positive, finite limit, asymptotically only pairs of twins of black points and isolated black points occur. If $\frac{(N-n)^2}{N}$ has a positive, finite limit asymptotically only pairs of twins of white points and isolated white points occur.

4. Some properties of the distribution of $\underline{x}$ for k=2

In this section some properties of the distribution of $\underline{x}$ for k=2 will be derived. The trivial case that m=2[5] is left out of consideration; then $m \geqq 3 (N \geqq 6)$ and each point has joins with 5 other points.

In section 4.1 we derive the exact distribution of $\underline{x}$ for n=2, 3 and 4; section 4.2 contains a general expression for $P[\underline{x}=0|k=2]$ and $P[\underline{x}=1|k=2]$; formulae for $\mathcal{E}(\underline{x}|k=2)$, $\sigma^2(\underline{x}|k=2)$ and $\mathcal{E}\{(\underline{x}-\mathcal{E}\underline{x})^3|k=2\}$ are given in section 4.3.

4.1. The exact distribution of $\underline{x}$ for n=2, 3 and 4

If N=6 each point has joins with each of the 5 other points. Consequently (cf. footnote 5)

(4.1;1) $$P[\underline{x} = \tfrac{1}{2}n(n-1)|k=2,\ N=6] = 1 \qquad 0 \leqq n \leqq 6.$$

If $m > 3$ and n=2 the random variable $\underline{x}$ takes one of the values 0 and 1. If the first black point is chosen N-1 points are available for the second black point; 5 of these N-1 points have joins with the first black point, consequently

(4.1;2) $$\begin{cases} P[\underline{x}=0|k=2,\ n=2] = \dfrac{N-6}{N-1}, \\ P[\underline{x}=1|k=2,\ n=2] = \dfrac{5}{N-1}. \end{cases}$$

We now derive the distribution of $\underline{x}$ for n=3 and we first consider the case that $N \geqq 10$. For the second black point with respect to the first one the following situations may be distinguished: (* denotes a black point, . a white point)

----

5) If m=2 each point in the rectangle has joins with each of the three other points. Consequently

$$P[\underline{x} = \tfrac{1}{2}n(n-1)\ |k=m=2] = 1 \qquad \text{for } 0 \leqq n \leqq 4 .$$

(4.1;3)

			probability
A	1)	. . . . * * .	$\frac{2}{N-1}$
	2)	. . . * . . . * . .	$\frac{2}{N-1}$
B		. . * . . . . * . .	$\frac{1}{N-1}$
C	1)	. . . . * . *	$\frac{2}{N-1}$
	2)	. . . . * . . * . .	$\frac{2}{N-1}$
D		the other situations	$\frac{N-10}{N-1}$.

First situation A.1 is considered. If the third black point is chosen $\underline{x}$ takes one of the values 1,2 and 3 and the number of times $\underline{x}$ takes these values may be counted as follows:

1 2 3 3 2 1

1 2 * * 2 1 ,

where for the possible positions of the third point the values assumed by $\underline{x}$ are indicated. Consequently for k=2, N$\geqq$10 and n=3

$$(4.1;4)\quad \begin{cases} P[\underline{x}=1 \mid A.1] = \frac{N-8}{N-2}, \\ P[\underline{x}=2 \mid A.1] = \frac{4}{N-2}, \\ P[\underline{x}=3 \mid A.1] = \frac{2}{N-2}. \end{cases}$$

For situation A.2 the same values of $\underline{x}$ with the same probabilities occur.

In an analogous way the distributions of $\underline{x}$ under the conditions B, C and D may be obtained. This gives for k=2, N$\geqq$10 and n=3:

	$P[\underline{x}=x\,\vert A]$	$P[\underline{x}=x\,\vert B]$	$P[\underline{x}=x\,\vert C]$	$P[\underline{x}=x\,\vert D]$
0	0	0	$\frac{N-10}{N-2}$	$\frac{N-12}{N-2}$
1	$\frac{N-8}{N-2}$	$\frac{N-6}{N-2}$	$\frac{6}{N-2}$	$\frac{10}{N-2}$
2	$\frac{4}{N-2}$	0	$\frac{2}{N-2}$	0
3	$\frac{2}{N-2}$	$\frac{4}{N-2}$	0	0

From this table it follows that

(4.1;5)

x	$P[\underline{x}=x\,\vert k=2,\ N \geqq 10,\ n=3]$
0	$\frac{(N-10)(N-8)}{(N-1)(N-2)}$
1	$\frac{3(5N-38)}{(N-1)(N-2)}$
2	$\frac{24}{(N-1)(N-2)}$
3	$\frac{12}{(N-1)(N-2)}$

Similarly it may be seen that (4.1;5) also holds for N=8.

In an analogous way the distribution of $\underline{x}$ for n=4 may be obtained:

x	$P[\underline{x}=x \mid k=2, N=8, n=4]$	$P[\underline{x}=x \mid k=2, N \geqq 10, n=4]$
0	0	$\frac{(N-14)(N-12)(N-10)}{(N-1)(N-2)(N-3)}$
1	0	$\frac{6(N-10)(5N-56)}{(N-1)(N-2)(N-3)}$
2	$\frac{1}{35}$	$\frac{3(57N-550)}{(N-1)(N-2)(N-3)}$
3	0	$\frac{48(N-4)}{(N-1)(N-2)(N-3)}$
4	$\frac{24}{35}$	$\frac{96}{(N-1)(N-2)(N-3)}$
5	$\frac{8}{35}$	$\frac{48}{(N-1)(N-2)(N-3)}$
6	$\frac{2}{35}$	$\frac{12}{(N-1)(N-2)(N-3)}$

4.2. A general expression for $P[\underline{x}=0 \mid k=2]$ and $P[\underline{x}=1 \mid k=2]$

A general expression for

$$P[\underline{x}=0 \mid k=2]$$

may be obtained as follows.
The random variable $\underline{x}$ takes the value 0 if and only if the following situation occurs:
each column of the rectangle contains at most one black point and between two columns containing a black point at least one column with two white points occurs.
This situation is denoted by S. The situation S cannot occur if $n > \frac{k}{2} = \frac{N}{4}$; consequently

$$(4.2;1) \qquad P[\underline{x}=0 \mid k=2] = 0 \qquad \text{if} \qquad n > \frac{N}{4} .$$

We now compute $P[S \mid k=2]$ for $n \leqq \frac{N}{4}$. If the first black point is chosen $n-1$ isolated columns must be chosen from $\frac{N}{2} - 1$ columns. This can be done in (cf. (3;3) for $x=0$)

$$\binom{\frac{N}{2}-n-1}{n-1}$$

ways. In each of these n-1 columns a black point may be chosen in two ways. Consequently the number of ways in which n-1 black points may be chosen from N-1 points such that S occurs is

$$2^{n-1}\binom{\frac{N}{2}-n-1}{n-1}.$$

Consequently

$$(4.2;2) \qquad P[\underline{x}=0\,|\,k=2] = 2^{n-1}\frac{\binom{\frac{N}{2}-n-1}{n-1}}{\binom{N-1}{n-1}}$$

for $n \leqq \frac{N}{4}$. If $n > \frac{N}{4}$ then $n-1 > \frac{N}{2}-n-1$, so

$$\binom{\frac{N}{2}-n-1}{n-1} = 0 \qquad \text{for } n > \frac{N}{4},$$

i.e. (4.2;2) also holds for $n > \frac{N}{4}$.

$P[\underline{x}=1\,|\,k=2]$ is not derived under the condition that the first black point has a fixed position. This probability may easier be found as follows.

Let, for $\nu=1,\ldots,\binom{n}{2}$,

$$(4.2;3) \qquad \underline{x}_\nu = \begin{cases} 1 \text{ if the } \nu^{th} \text{ pair of black points is a joined pai} \\ 0 \text{ if the } \nu^{th} \text{ pair of black points is not a joined pair.} \end{cases}$$

Then

$$(4.2;4) \qquad P[\underline{x}=1\,|\,k=2] = \sum_\nu P[\underline{x}=1 \text{ and } \underline{x}_\nu=1].$$

Further, if $\underline{x}_\nu=1$, the following situations may be distinguished

A. The two points of the ν^{th} black pair are in the same column. Then (cf. (4.2;2))

$$(4.2;5) \qquad P[\underline{x}=1\,|\,\underline{x}_\nu=1, A, k=2] = 2^{n-2}\frac{\binom{\frac{N}{2}-(n-1)-1}{n-2}}{\binom{N-2}{n-2}}.$$

B. The two points of the ν^{th} black pair are not in the same column. Then (cf. (4.2;2))

(4.2;6) $$P\left[\underline{x}=1 \mid \underline{x}_\nu=1,\ B,\ k=2\right] = 2^{n-2}\,\frac{\binom{\frac{N}{2}-1-(n-1)-1}{n-2}}{\binom{N-2}{n-2}}\ .$$

Further

(4.2;7) $$P\left[\underline{x}_\nu=1 \text{ and } A \mid k=2\right] = \frac{1}{N-1}\ ,\quad P\left[\underline{x}_\nu=1 \text{ and } B \mid k=2\right] = \frac{4}{N-1}\ ,$$

consequently (cf. (4.2;4))

(4.2;8) $$P\left[\underline{x}=1 \mid k=2\right] = \sum_\nu \Big\{ P\left[\underline{x}_\nu=1 \text{ and } A \mid k=2\right].P\left[\underline{x}=1 \mid \underline{x}_\nu=1, A, k=2\right] +$$
$$+ P\left[\underline{x}_\nu=1 \text{ and } B \mid k=2\right].P\left[\underline{x}=1 \mid \underline{x}_\nu=1, B, k=2\right]\Big\} =$$
$$= \frac{\binom{n}{2}2^{n-2}}{(N-1)\binom{N-2}{n-2}}\left[\binom{\frac{N}{2}-n}{n-2} + 4\binom{\frac{N}{2}-n-1}{n-2}\right]\ .$$

4.3. <u>The mean, variance and third moment of $\underline{x}$.</u>

In this section formulae for $\mathcal{E}(\underline{x} \mid k=2)$, $\sigma^2(\underline{x} \mid k=2)$ and $\mathcal{E}\{(\underline{x}-\mathcal{E}\underline{x})^3 \mid k=2\}$ will be derived.

If $\underline{x}_\nu$ $(\nu=1,\ldots,\binom{n}{2})$ are defined by (4.2;3) then

(4.3;1) $$\underline{x} = \sum_\nu \underline{x}_\nu$$

and (cf. also (4.1;2))

(4.3;2) $$\begin{cases} P\left[\underline{x}_\nu = 1 \mid k=2\right] = \frac{5}{N-1}\ , \\ P\left[\underline{x}_\nu = 0 \mid k=2\right] = \frac{N-6}{N-1}\ . \end{cases}$$

Now we need the following moments

$$\mathcal{E}(\underline{x}_\nu \mid k=2),\quad \mathcal{E}(\underline{x}_\nu^2 \mid k=2) \text{ and } \mathcal{E}(\underline{x}_\nu^3 \mid k=2)$$

$$\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \mid k=2) \text{ and } \mathcal{E}(\underline{x}_\nu^2 x_\kappa \mid k=2) \qquad \text{for } \nu \neq \kappa$$

$$\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \underline{x}_\gamma) \qquad \text{for } \nu \neq \kappa \neq \gamma\ ,\ \nu \neq \gamma.$$

From (4.3;2) it follows that

(4.3;3) $\mathcal{E}(\underline{x}_\nu|k=2) = \mathcal{E}(\underline{x}_\nu^2|k=2) = \mathcal{E}(\underline{x}_\nu^3|k=2) = P[\underline{x}_\nu = 1|k=2] = \frac{5}{N-1}$.

Consequently

(4.3;4) $\mathcal{E}(\underline{x}|k=2) = \sum_\nu \mathcal{E}(\underline{x}_\nu|k=2) = \binom{n}{2} \frac{5}{N-1} = \frac{5n(n-1)}{2(N-1)}$.

Further we have, for $\nu \neq \kappa$,

(4.3;5) $\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa|k=2) = \mathcal{E}(\underline{x}_\nu^2 \underline{x}_\kappa|k=2) = P[\underline{x}_\nu=1 \text{ and } \underline{x}_\kappa=1|k=2]$.

For the ν^{th} and κ^{th} pair of black points the following situations may be distinguished.

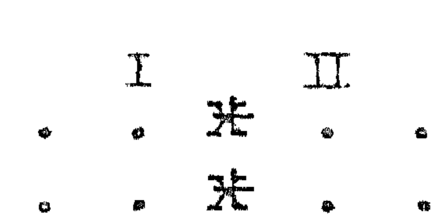
1. the ν^{th} and κ^{th} pair together consist of three black points. The number of such couples of pairs of black points with $\nu \neq \kappa$ is $n(n-1)(n-2)$ and

(4.3;6) $P[\underline{x}_\nu=1 \text{ and } \underline{x}_\kappa=1|k=2] = P[\underline{x}_\nu=1|k=2].P[\underline{x}_\kappa=1|\underline{x}_\nu=1,k=2] =$

$$= \frac{5}{N-1} \cdot \frac{4}{N-2} = \frac{20}{(N-1)(N-2)} .$$

2. the ν^{th} and κ^{th} pair together consist of four black points. The number of such couples of pairs of black points with $\nu \neq \kappa$ is $\frac{1}{4} n(n-1)(n-2)(n-3)$ and we consider the following two situations.

a. the two black points of the ν^{th} pair are in the same column (cf. fig.)



Then the first black point of the κ^{th} pair is

a.1 in one of the columns I or II

a.2 not in one of the columns I or II,

where

(4.3;7) $\begin{cases} P[\underline{x}_\kappa=1|k=2;\ a.1] = \frac{3}{N-3} \\ P[\underline{x}_\kappa=1|k=2\ \ a.2] = \frac{5}{N-3} \end{cases}$

Consequently

(4.3;8) $P[\underline{x}_\kappa=1|k=2,a] = \frac{4}{N-2} \cdot \frac{3}{N-3} + \frac{N-6}{N-2} \cdot \frac{5}{N-3} = \frac{5N-18}{(N-2)(N-3)}$.

b. the two black points of the ν^{th} pair are in adjacent columns

$$\begin{array}{cccccc} & & ^{I} & & ^{II} & \\ . & . & III & IV & . & . \\ . & . & * & * & . & . \end{array} \quad \text{or} \quad \begin{array}{cccccc} & & ^{I} & & ^{II} & \\ . & . & III & * & . & . \\ . & . & * & IV & . & . \end{array}$$

Then the first black point of the κ^{th} pair is

b.1. in one of the columns I or II

b.2. in one of the points III or IV

b.3. in one of the other points,

where

$$(4.3;9) \quad \begin{cases} P[\underline{x}_\kappa=1 \,|\, k=2; b.1] = \frac{4}{N-3} \\ P[\underline{x}_\kappa=1 \,|\, k=2; b.2] = \frac{3}{N-3} \\ P[\underline{x}_\kappa=1 \,|\, k=2; b.3] = \frac{5}{N-3}, \end{cases}$$

consequently

$$(4.3;10) \quad P[\underline{x}_\kappa=1 \,|\, k=2, b] = \frac{4}{N-2} \cdot \frac{4}{N-3} + \frac{2}{N-2} \cdot \frac{3}{N-3} + \frac{(N-8)}{(N-2)} \cdot \frac{5}{N-3} = \frac{5N-18}{(N-2)(N-3)} .$$

So if the ν^{th} and κ^{th} pair together consist of four black points then

$$(4.3;11) \quad P[\underline{x}_\nu=1 \text{ and } \underline{x}_\kappa=1 \,|\, k=2] = \frac{5(5N-18)}{(N-1)(N-2)(N-3)} .$$

Consequently (cf. (4.3;6) and (4.3;11))

$$(4.3.12) \quad \sum_{\nu \neq \kappa}\sum \mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \,|\, k=2) = \sum_{\nu \neq \kappa}\sum \mathcal{E}(\underline{x}_\nu^2 \underline{x}_\kappa \,|\, k=2) =$$

$$= n(n-1)(n-2)\frac{20}{(N-1)(N-2)} + \frac{1}{4}\, n(n-1)(n-2)(n-3)\frac{5(5N-18)}{(N-1)(N-2)(N-3)}$$

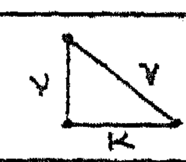
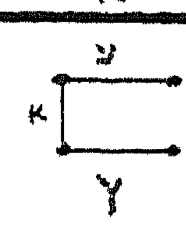
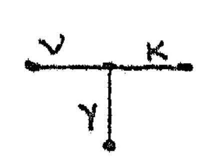
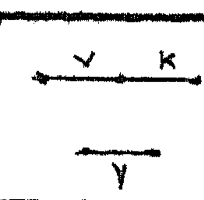
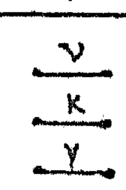
Further (cf.(4.3;1))

$$(4.3;13) \quad \mathcal{E}(\underline{x}^2 \,|\, k=2) = \sum_\nu \mathcal{E}(\underline{x}_\nu^2 \,|\, k=2) + \sum_{\nu \neq \kappa}\sum \mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \,|\, k=2)$$

and from (4.3;4),(4.3;12) and (4.3;13) then follows

$$(4.3;14) \qquad \sigma^2(\underline{x}\,|\,k=2) = \frac{5n(n-1)(N-6)(N-n)(N-n-1)}{2(N-1)^2(N-2)(N-3)} .$$

In an analogous way $\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \underline{x}_\gamma\,|\,k=2)$ for $\nu \neq \kappa \neq \gamma$ and $\nu \neq \gamma$ may be obtained. The ν^{th}, κ^{th} and γ^{th} pair of black points together consist of 3, 4, 5 or 6 points and we find for $N \geqq 8$

number of black points	situation	number of three pairs of black points	$\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \underline{x}_\gamma\,|\,k=2) = P[\underline{x}_\nu = \underline{x}_\kappa = \underline{x}_\gamma = 1\,|\,k=2]$
3		$n(n-1)(n-2)$	$\frac{12}{(N-1)(N-2)}$
4	a) 	$3n(n-1)(n-2)(n-3)$	$\frac{68}{(N-1)(N-2)(N-3)}$
	b) 	$n(n-1)(n-2)(n-3)$	$\frac{60}{(N-1)(N-2)(N-3)}$
5		$\frac{3}{2}n(n-1)(n-2)(n-3)(n-4)$	$\frac{100N-496}{(N-1)(N-2)(N-3)(N-4)}$
6		$\frac{1}{8}n(n-1)(n-2)(n-3)(n-4)(n-5)$	$\frac{125N^2-1350N+3784}{(N-1)(N-2)(N-3)(N-4)(N-5)}$

Further (cf. (4.3;1))

$$(4.3;15) \quad \mathcal{E}(\underline{x}^3\,|\,k=2) = \sum_\nu \mathcal{E}(\underline{x}_\nu^3\,|\,k=2) + 3 \sum_{\nu\neq\kappa}\sum \mathcal{E}\underline{x}_\nu^2 \underline{x}_\kappa + \sum_{\substack{\nu\neq\kappa\neq\gamma \\ \nu\neq\gamma}}\sum\sum \mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \underline{x}_\gamma\,|\,k=2) .$$

Consequently (cf.(4.3;3),(4.3;12) and the table above)

$$(4.3;16) \quad \mathcal{E}(\underline{x}^3\,|\,k=2) = \frac{5n(n-1)}{2(N-1)} + \frac{72n(n-1)(n-2)}{(N-1)(N-2)} +$$

$$+ \frac{1}{4}\,n(n-1)(n-2)(n-3)\,\frac{75N+786}{(N-1)(N-2)(N-3)}$$

$$+ n(n-1)(n-2)(n-3)(n-4)\,\frac{150N-744}{(N-1)(N-2)(N-3)(N-4)} +$$

$$+ \frac{1}{8}\,n(n-1)(n-2)(n-3)(n-4)(n-5)\,\frac{125N^2-1350N+3784}{(N-1)(N-2)(N-3)(N-4)(N-5)} .$$

Further

$$(4.3;17) \qquad \mathcal{E}(\underline{x} - \mathcal{E}\underline{x})^3 = \mathcal{E}\underline{x}^3 - 3\,\mathcal{E}\underline{x}\,\sigma^2(\underline{x}) - (\mathcal{E}\underline{x})^3 .$$

Consequently (cf. (4.3;4), (4.3;14) and (4.3;16))

(4.3;17) $\mathcal{E}\{(\underline{x}-\mathcal{E}\underline{x})^3 \mid k=2\} =$

$$= \begin{cases} 0 \text{ for } N=6 \\ n(n-1)(N-n)(N-n-1)\dfrac{(N-1)(5N^3-133N^2+694N+184)+4n(N-n)(N^2+48N-424)}{2(N-1)^3(N-2)(N-3)(N-4)(N-5)} \text{ for } N \geqq 8. \end{cases}$$

5. Some properties of the distribution of $\underline{x}$ for k=3

5.1. The exact distribution of $\underline{x}$ for n=2, n=3 and N=16 with n=4

If N=9 each point has joins with each of the 8 other points. Consequently (cf. also (4.1;1))

(5.1;1) $P[\underline{x} = \frac{1}{2}n(n-1) \mid k=m=3] = 1$ for $0 \leqq n \leqq 9$.

For n=2 we have (cf. also (4.1;2))

(5.1;2) $$\begin{cases} P[\underline{x}=0 \mid k \geqq 3, n=2] = \dfrac{N-9}{N-1} \\ P[\underline{x}=1 \mid k \geqq 3, n=2] = \dfrac{8}{N-1} \end{cases}.$$

For n=3 the exact distribution of $\underline{x}$ may be obtained in the same way as in section 4.1. In this case we have

x	$P[\underline{x}=x \mid n=3, k=3, N \geqq 12]$	$P[\underline{x}=x \mid n=3, k>3, N \geqq 16]$
0	$\frac{(N-15)(N-12)}{(N-1)(N-2)}$	$\frac{N^2-27N+194}{(N-1)(N-2)}$
1	$\frac{6(4N-45)}{(N-1)(N-2)}$	$\frac{24(N-13)}{(N-1)(N-2)}$
2	$\frac{54}{(N-1)(N-2)}$	$\frac{96}{(N-1)(N-2)}$
3	$\frac{38}{(N-1)(N-2)}$	$\frac{24}{(N-1)(N-2)}$

Further we have for n=k=m=4 (N=16)

x	$P[\underline{x}=x \mid n=k=m=4]$
0	0,0066
1	0,0352
2	0,2198
3	0,3165
4	0,3429
5	0,0703
6	0,0088

5.2. A general expression for $P[\underline{x}=0 \mid k=3]$ and $P[\underline{x}=1 \mid k=3]$

If k=3 a general expression for

$$P[\underline{x}=0 \mid k=3] \text{ and } P[\underline{x}=1 \mid k=3]$$

may be obtained in the same way as in section 4.2.

In this case the random variable $\underline{x}$ is equal to zero if and only if the following situation occurs:
each column contains at most one black point and between two columns containing a black point at least one column with three white points occurs.
In an analogous way as in section 4.2 we then find

$$(5.2;1) \qquad P[\underline{x}=0 \mid k=3] = 3^{n-1} \frac{\binom{\frac{N}{3}-n-1}{n-1}}{\binom{N-1}{n-1}}$$

and

$$(5.2;2) \qquad P[\underline{x}=1 \mid k=3] = \frac{\binom{n}{2}3^{n-2}}{(N-1)\binom{N-2}{n-2}} \left\{ 2\binom{\frac{N}{3}-n}{n-2} + 6\binom{\frac{N}{3}-n-1}{n-2} \right\}.$$

5.3. The mean, variance and third moment of $\underline{x}$ for $k \geq 3$.

Formulae for the moments of $\underline{x}$ may be obtained in the same way as in section 5.2 for k=2. We again consider the random variables $\underline{x}_\nu (\nu=1,\ldots,\binom{n}{2})$ (cf. (4.2;3)). Then

$$(5.3;1)\qquad \begin{cases} P[\underline{x} = 0 \mid k \geq 3] = \frac{N-9}{N-1} \\ P[\underline{x} = 1 \mid k \geq 3] = \frac{8}{N-1} \end{cases} .$$

As in section 4.3 we need the following moments

$$\mathcal{E}(\underline{x}_\nu \mid k \geq 3),\quad \mathcal{E}(\underline{x}_\nu^2 \mid k \geq 3) \text{ and } \mathcal{E}(\underline{x}_\nu^3 \mid k \geq 3)$$

$$\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \mid k \geq 3) \text{ and } \mathcal{E}(\underline{x}_\nu^2 \underline{x}_\kappa \mid k \geq 3) \qquad \text{for } \nu \neq \kappa$$

$$\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \underline{x}_\gamma \mid k \geq 3) \qquad \text{for } \nu \neq \kappa \neq \gamma,\ \nu \neq \gamma .$$

From (5.3;1) it follows that

$$(5.3;2)\quad \mathcal{E}(\underline{x}_\nu \mid k \geq 3) = \mathcal{E}(\underline{x}_\nu^2 \mid k \geq 3) = \mathcal{E}(\underline{x}_\nu^3 \mid k \geq 3) = P[\underline{x}_\nu = 1 \mid k \geq 3] = \frac{8}{N-1} .$$

Consequently

$$((5.3;3)\quad \mathcal{E}(\underline{x} \mid k \geq 3) = \sum_\nu \mathcal{E}(\underline{x}_\nu \mid k \geq 3) = \binom{n}{2} \frac{8}{N-1} = \frac{4n(n-1)}{N-1} .$$

For the ν^{th} and κ^{th} pair of black points ($\nu \neq \kappa$) we again consider the following situations (cf. section 4.3)

1. the ν^{th} and κ^{th} pair together consist of three black points.

Then

$$(5.3;4)\quad \mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \mid k \geq 3) = \mathcal{E}(\underline{x}_\nu^2 \underline{x}_\kappa \mid k \geq 3) =$$

$$= P[\underline{x}_\kappa = 1 \mid k \geq 3] . P[\underline{x}_\nu = 1 \mid \underline{x}_\kappa = 1; k \geq 3] = \frac{8}{N-1} \cdot \frac{7}{N-2} = \frac{56}{(N-1)(N-2)} .$$

2. the ν^{th} and κ^{th} pair together consist of four black points.

Then it may be proved (in the same way as in section 4.3) that

$$(5.3;5)\qquad \mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \mid k \geq 3) = \mathcal{E}(\underline{x}_\nu^2 \underline{x}_\kappa \mid k \geq 3) = \frac{16(4N-15)}{(N-1)(N-2)(N-3)} .$$

Consequently (cf. also (4.3;12))

$$(5.3;6)\ \sum\sum_{\nu \neq \kappa} \mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \mid k \geq 3) = n(n-1)(n-2) \frac{56}{(N-1)(N-2)} +$$

$$+ \frac{1}{4} n(n-1)(n-2)(n-3) \frac{16(4N-15)}{(N-1)(N-2)(N-3)}$$

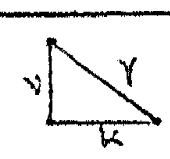
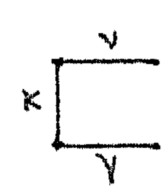
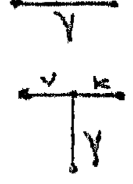
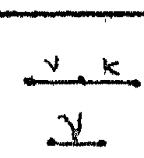
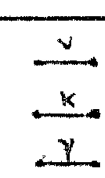
and from

$$(5.3;7)\quad \mathcal{E}(\underline{x}^2 \mid k \geq 3) = \sum_\nu \mathcal{E}(\underline{x}_\nu^2 \mid k \geq 3) + \sum\sum_{\nu \neq \kappa} \mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \mid k \geq 3)$$

then follows

$$(5.3;8) \qquad \sigma^2(\underline{x}\,|\,k \geqq 3) = \frac{4n(n-1)(N-9)(N-n)(N-n-1)}{(N-1)^2(N-2)(N-3)} .$$

In an analogous way $\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \underline{x}_\gamma\,|\,k \geqq 3)$ for $\nu \neq \kappa \neq \gamma$, $\nu \neq \gamma$ may be obtained. The ν^{th}, κ^{th} and γ^{th} pair of black points together consist of 3, 4, 5 or 6 points and we find for $N > 9$ (cf. also section 4.3)

number of black points	situation	number of three pairs of black points	$\mathcal{E}(\underline{x}_\nu \underline{x}_\kappa \underline{x}_\gamma) = P[\underline{x}_\nu = \underline{x}_\kappa = \underline{x}_\gamma = 1]$	
			$k=3,\ m \geqq 4$	$k \geqq 4$
3		$n(n-1)(n-2)$	$\frac{38}{(N-1)(N-2)}$	$\frac{24}{(N-1)(N-2)}$
4	a) 	$3n(n-1)(n-2)(n-3)$	$\frac{354}{(N-1)(N-2)(N-3)}$	$\frac{368}{(N-1)(N-2)(N-3)}$
	b) 	$n(n-1)(n-2)(n-3)$	$\frac{336}{(N-1)(N-2)(N-3)}$	$\frac{336}{(N-1)(N-2)(N-3)}$
5		$\frac{3}{2}n(n-1)(n-2)(n-3)(n-4)$	$\frac{448N-2388}{(N-1)(N-2)(N-3)(N-4)}$	$\frac{448N-2416}{(N-1)(N-2)(N-3)(N-4)}$
6		$\frac{1}{8}n(n-1)(n-2)(n-3)(n-4)(n-5)$	$\frac{512N^2-5760N+17232}{(N-1)(N-2)(N-3)(N-4)(N-5)}$	$\frac{512N^2-5760N+17344}{(N-1)(N-2)(N-3)(N-4)(N-5)}$

From this table it follows that (cf. also (4.3;15))

$$(5.3;9) \qquad \mathcal{E}(\underline{x}^3\,|\,k=3, m \geqq 4) = \frac{4n(n-1)}{N-1} + \frac{206n(n-1)(n-2)}{(N-1)(N-2)} +$$

$$+ n(n-1)(n-2)(n-3)\,\frac{48N+1218}{(N-1)(N-2)(N-3)} +$$

$$+ 6n(n-1)(n-2)(n-3)(n-4)\,\frac{112N-597}{(N-1)(N-2)(N-3)(N-4)} +$$

$$+ \frac{1}{8}\,n(n-1)(n-2)(n-3)(n-4)(n-5)\,\frac{512N^2-5760N+17232}{(N-1)(N-2)(N-3)(N-4)(N-5)}$$

and

$$(5.3;10) \qquad \mathcal{E}(\underline{x}^3\,|\,k \geqq 4) = \frac{4n(n-1)}{N-1} + \frac{192n(n-1)(n-2)}{(N-1)(N-2)} +$$

$$+ n(n-1)(n-2)(n-3)\,\frac{48N+1260}{(N-1)(N-2)(N-3)}$$

$$+ 24n(n-1)(n-2)(n-3)(n-4)\,\frac{28N-151}{(N-1)(N-2)(N-3)(N-4)} +$$

$$+ \frac{1}{8}\,n(n-1)(n-2)(n-3)(n-4)(n-5)\,\frac{512N^2-5760N+17344}{(N-1)(N-2)(N-3)(N-4)(N-5)}$$

From (5.3;9) and (5.3;10) then follows (cf. also (4.3;17))

(5.3;11) $\mathcal{E}\{(\underline{x}-\mathcal{E}\underline{x})^3 \mid k \geqq 3\} =$

$$\begin{cases} = 0 \quad \text{for } N=9 \\ = n(n-1)(N-n)(N-n-1)\dfrac{4(N-1)(N^3+45N^2+374N+54)+2n(N-n)(11N^2+10N-1557)}{(N-1)^3(N-2)(N-3)(N-4)(N-5)} \quad \text{for } k=3, m\geqq 4 \\ = n(n-1)(N-n)(N-n-1)\dfrac{4(N-17)[(N-1)(N^2-21N-4)+2n(N-n)(N+23)]}{(N-1)^3(N-2)(N-3)(N-4)(N-5)} \quad \text{for } k\geqq 4 . \end{cases}$$

6. The asymptotic distribution of $\underline{x}$ for $N \to \infty$

In this section we consider the asymptotic distribution of $\underline{x}$ for $N \to \infty$.

Now let a denote the number of joins of each point, then

(6;1) $\begin{cases} a=2 \text{ for } k=1 \\ a=5 \text{ for } k=2 \\ a=8 \text{ for } k\geqq 3 \end{cases}$.

Let further (cf. section 3) $\underline{x}'$ denote the number of white-white joins, then (cf.(3.5))

(6.2) $$\underline{x}' = \underline{x} + \frac{a}{2}N - an.$$

Further

(6;3) $$\begin{cases} \mathcal{E}\underline{x} = \dfrac{an(n-1)}{2(N-1)} \qquad \mathcal{E}\underline{x}' = \dfrac{a(N-n)(N-n-1)}{2(N-1)} \\ \sigma^2(\underline{x}) = \sigma^2(\underline{x}') = \dfrac{an(n-1)(N-a-1)(N-n)(N-n-1)}{2(N-1)^2(N-2)(N-3)} \end{cases}.$$

Now the following cases may be distinguished (cf. section 3)

1. $\lim_{N\to\infty} \mathcal{E}\underline{x} = 0$

2. $\lambda \overset{\text{def}}{=} \lim_{N\to\infty} \mathcal{E}\underline{x}$ exists and $0<\lambda<\infty$

3. $\lim_{N\to\infty} \sigma^2(\underline{x}) = \infty$[6)]

----

6) If this condition is satisfied, then (cf. footnote 3) one of the following conditions is satisfied

a) $\lim_{N\to\infty} \frac{n}{N} = 0, \ \lim_{N\to\infty} \frac{n^2}{N} = \infty$

b) $0 < \liminf_{N\to\infty} \frac{n}{N} \leqq \limsup_{N\to\infty} \frac{n}{N} < 1$

c) $\lim_{N\to\infty} \frac{n}{N} = 1, \ \lim_{N\to\infty} \frac{(N-n)^2}{N} = \infty$.

4. $\lambda' \stackrel{\text{def}}{=} \lim_{N\to\infty} \mathcal{E}\,\underline{x}'$ exists and $0<\lambda'<\infty$

5. $\lim_{N\to\infty} \mathcal{E}\,\underline{x}' = 0.$

It may easily be seen that (cf. (3;7) and (3;8))

$$(6;4)\qquad \begin{cases} 1.\ \lim_{N\to\infty} P[\underline{x}=0] = 1 & \text{if} \quad \lim_{N\to\infty} \mathcal{E}\,\underline{x} = 0 \\ 2.\ \lim_{N\to\infty} P[\underline{x}'=0] = 1 & \text{if} \quad \lim_{N\to\infty} \mathcal{E}\,\underline{x}' = 0. \end{cases}$$

We now consider the case where $\lim_{N\to\infty} \sigma^2(\underline{x})=\infty$. For k=1 the asymptotic normality of $\frac{\underline{x}-\mathcal{E}\underline{x}}{\sigma(\underline{x})}$ and $\frac{\underline{x}'-\mathcal{E}\underline{x}'}{\sigma(\underline{x}')}$ under this condition follows from the properties of the hypergeometric distribution (cf. section 3). For $k\geq 2$ A.R. BLOEMENA (1960) proved the asymptotic normality of $\frac{\underline{x}-\mathcal{E}\underline{x}}{\sigma(\underline{x})}$ and $\frac{\underline{x}'-\mathcal{E}\underline{x}'}{\sigma(\underline{x}')}$ under the condition (cf. footnote 6)

$$(6;5)\qquad 0 < \liminf_{N\to\infty} \frac{n}{N} \leq \limsup_{N\to\infty} \frac{n}{N} < 1 .$$

This asymptotic normality does not only hold for k <u>and</u> $m\to\infty$ (cf. section 1), but also for finite k and $m\to\infty$. The asymptotic normality for $k\geq 2$ under the condition (cf. footnote 6)

$$(6;6)\qquad \begin{cases} 1.\ \lim_{N\to\infty} \frac{n}{N} = 0 \text{ and } \lim_{N\to\infty} \frac{n^2}{N} = \infty \\ \quad \text{or} \\ 2.\ \lim_{N\to\infty} \frac{n}{N} = 1 \text{ and } \lim_{N\to\infty} \frac{(N-n)^2}{N} = \infty \end{cases}$$

has not yet been proved.
We now consider the case that

$$(6;7)\qquad \lambda = \lim_{N\to\infty} \mathcal{E}\,\underline{x}$$

exists and $0<\lambda<\infty$. Then

$$(6;8)\qquad \lim_{N\to\infty} \frac{n}{N} = 0.$$

It will be proved that in this case the random variable $\underline{x}$ has asymptotically for $N\to\infty$ a Poisson distribution with parameter λ, i.e.

(6;9) $$\lim_{N\to\infty} P[\underline{x}=x] = \frac{e^{-\lambda}\lambda^x}{x!} .$$

For k=1 this follows from the properties of the hypergeometric distribution (cf. section 3). For k=2 and k=3 with x=0 or x=1 (6;9) follows from the exact formulae for this probability (cf. (4.2;2), (4.2;8), (5.2;1) and (5.2;2)). For $P[x=0|k=2]$ e.g. it follows from (4.2;2) that for $n \leq \frac{N}{4}$

(6;10) $$P[\underline{x}=0|k=2] = \prod_{i=1}^{n-1} \frac{N-2n-2i}{N-i} = \prod_{i=1}^{n-1} \left(1 - \frac{2n+i}{N-i}\right) ,$$

consequently (cf. (6;8))

(6;11) $$\lim_{N\to\infty} \ln P[\underline{x}=0|k=2] = \lim_{N\to\infty} \sum_{i=1}^{n-1} \ln\left(1 - \frac{2n+i}{N-i}\right) =$$

$$= -\lim_{N\to\infty} \sum_{i=1}^{n-1} \frac{2n+i}{N} = -\lim_{N\to\infty} \frac{5n^2}{2N} = -\lambda .$$

In an analogous way (6;9) may be proved for k=2 with x=1 and for k=3 with x=0 and x=1.

In the general case (6;9) may be proved as follows.

The random variable $\underline{x}$ takes the value 0 if and only if each black point is an isolated black point. Now suppose i-1 black points ($1 \leq i-1 \leq n-1$) are chosen in such a way that they are all isolated. Then for the i^{th} black point N - (i-1) points are available and at least N - (a+1)(i-1) of these points have no joins with the (i-1) black points. Consequently

(6;12) $$P[\underline{x}=0] \geq \prod_{i=2}^{n} \frac{N - (a+1)(i-1)}{N - (i-1)}$$

and from (6;12) it follows that

(6;13) $$\liminf_{N\to\infty} P[\underline{x}=0] \geq \lim_{N\to\infty} \prod_{i=2}^{n} \left\{1 - \frac{a(i-1)}{N-(i-1)}\right\} = e^{-\lambda} .$$

Now consider the situation that x pairs of twins occur and n-2x isolated black points (cf. section 3). This situation is denoted by S_x and we have

$$(6;14) \qquad P[\underline{x}=x] \geqq P[S_x] ,$$

where the equality sign holds if and only if x=0 and x=1.

Now we have

$$(6;15) \qquad P[\underline{x}_\nu=1 | S_x] = \frac{x}{\binom{n}{2}} .$$

Consequently, for $x > 0$,

$$(6;16) \qquad P[S_x] = \frac{P[\underline{x}_\nu=1]}{P[\underline{x}_\nu=1|S_x]} \cdot P[S_x|\underline{x}_\nu=1] = \frac{P[\underline{x}_\nu=1]}{x} \binom{n}{2} P[S_x|\underline{x}_\nu=1] =$$

$$= \frac{\mathcal{E}\underline{\underline{x}}}{x} P[S_x|\underline{x}_\nu=1] .$$

Further

$$(6;17) \qquad \lim_{N\to\infty} \{P[S_x|\underline{x}_\nu=1] - P[S_{x-1}]\} = 0,$$ [7)]

consequently

$$(6;18) \qquad \lim_{N\to\infty} \left\{ P[S_x] - \frac{\mathcal{E}\underline{\underline{x}}}{x} P[S_{x-1}] \right\} = 0.$$

Now we have (cf. (6;13))

$$(6;19) \qquad \liminf_{N\to\infty} P[\underline{x}=0] = \liminf_{N\to\infty} P[S_o] \geqq e^{-\lambda} .$$

If we suppose that, for a certain value of x,

$$(6;20) \qquad \liminf_{N\to\infty} P[S_x] \geqq \frac{e^{-\lambda}\lambda^x}{x!}$$

then it follows from (6;18) that (6;20) holds for x+1. From (6;19) then follows

$$(6;21) \qquad \liminf_{N\to\infty} P[\underline{x}=x] \geqq \liminf_{N} P[S_x] \geqq \frac{e^{-\lambda}\lambda^x}{x!}$$

and from $\sum_x P[\underline{x}=x] = 1$ then follows

----

7) This will be proved at the end of this section.

(6;22) $\lim\limits_{N\to\infty} P\left[\underline{x}=x\right] = \lim\limits_{N\to\infty} P\left[S_x\right] = \dfrac{e^{-\lambda}\lambda^x}{x!}$.

Consequently $\underline{x}$ possesses asymptotically a Poisson distribution with parameter λ and asymptotically only pairs of twins of black points and isolated black points occur.
An alternative proof of asymptotic Poisson distribution may be found in A.R. BLOEMENA (1960).
Further if $\lambda' \overset{def}{=} \lim\limits_{N\to\infty} \mathcal{E}\underline{x}'$ exists and $0<\lambda'<\infty$, the random variable $\underline{x}'$ has asymptotically a Poisson distribution with parameter λ'. In this case asymptotically only pairs of twins of white points and isolated white points occur.

Proof of (6;17)

Let $\underline{y}_\nu$ denote the number of joins of the ν^{th} pair with the other n-2 black points, then

(6;23) $\mathcal{E}\,(\underline{y}_\nu \mid \underline{x}_\nu=1) < \dfrac{2an}{N}$,

consequently

(6;24) $\lim\limits_{N\to\infty} \mathcal{E}(\underline{y}_\nu \mid \underline{x}_\nu=1) = 0$ and $\lim\limits_{N\to\infty} P\left[\underline{y}_\nu=0 \mid \underline{x}_\nu=1\right] = 1.$

Further

(6;25) $P\left[S_x \mid \underline{x}_\nu=1\right] = P\left[S_x \text{ and } \underline{y}_\nu=0 \mid \underline{x}_\nu=1\right] =$

$= P\left[\underline{y}_\nu=0 \mid \underline{x}_\nu=1\right] \,.\, P\left[S_x \mid \underline{y}_\nu=0, \underline{x}_\nu=1\right]$.

Consequently (cf. (6;24))

(6;26) $\lim\limits_{N\to\infty} \left\{P\left[S_x \mid \underline{x}_\nu=1\right] - P\left[S_x \mid \underline{y}_\nu=0, \underline{x}_\nu=1\right]\right\} = 0.$

Further, if $\underline{y}_\nu=0$ and $\underline{x}_\nu=1$, there are at least N-2a points available for the n-2 black points, obtained by omitting the ν^{th} pair. Consequently, if $\underline{y}_\nu=0$ and $\underline{x}_\nu=1$, the situation S_x is identical with the situation that these n-2 black points chosen from at least N-2a points give the situation S_{x-1}. Therefore

(6;27) $\lim\limits_{N\to\infty} \left\{P\left[S_x \mid \underline{y}_\nu=0,\ \underline{x}_\nu=1\right] - P\left[S_{x-1}\right]\right\} = 0.$

From (6;26) and (6;27) then follows

(6;28) $\lim\limits_{N\to\infty} \left\{P\left[S_x \mid \underline{x}_\nu=1\right] - P\left[S_{x-1}\right]\right\} = 0.$

References

BLOEMENA, A.R. (1960), On probability distributions arising from points on a graph, Report S 266 of the Statistical Department of the Mathematical Centre, Amsterdam.

VAN EEDEN, C. and J.Th. RUNNENBURG (1959), Limiting distributions in a 2x2-table, Report S 254 (M 80a) of the Statistical Department of the Mathematical Centre, Amsterdam.

FELLER, W. (1950), An introduction to probability theory and its applications, John Wiley and Sons, Inc., New York.

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

KRISHNA IYER, P.V. (1950), The theory of probability distributions of points on a lattice, Ann. Math. Stat. 21 198-217 . (Corrections: Ann. Math. Stat. 22 (1951) 310).

KUIPERS, H.A. (1957), Over de verdeling van het aantal runs in reeksen van alternatieven, Report S 219 (Ov 7) of the Statistical Department of the Mathematical Centre, Amsterdam.

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

MORAN, P.A.P. (1948), The interpretation of statistical maps, Journal Royal Stat. Soc. B 10, 243-251.

STEVENS, W.L. (1939), Distribution of groups in a sequence of alternatives, Ann. of Eug. 9, 10-17.