FURTHER CONTRIBUTIONS TO THE THEORY OF PROBABILITY DISTRIBUTIONS OF POINTS ON A LINE—I

By P. V. Krishna Iyer

Indian Council of Agricultural Research, New Delhi

1. Introduction

A number of distributions arising from m points, each possessing one of k characters or colours, arranged at random on a line has been discussed by the author in a previous communication (Iyer, 1948). The distributions considered covered the cases of both free and non-free sampling. In free sampling, the character of each point is determined on the null hypothesis independently of the character of the other points. In non-free sampling the number of points possessing the different characters is fixed beforehand.

Owing to the limitations of the methods used before (Iyer, 1948), it was not possible to discuss fully the limiting forms of the distributions when k, the number of characters which the points could take, was greater than two. In this paper, we shall discuss somewhat fully the limiting forms of these distributions by obtaining the difference equations connecting the moment-generating functions (M.G.F.) of the distributions for free sampling. For the sake of convenience, as in the previous paper, we may describe the characters as colours, say black, white, red, etc. The probabilities of the points taking the colours black, white, red, etc., are $p_1, p_2, p_3, \ldots p_k$ subject to the condition that their sum is unity.

2. Black-Black Joins

The difference equation connecting the M.G.F.'s of this distribution was obtained by Iyer (1948) by actually writing down the general expression for the r-th factorial moment and substituting in the M.G.F. We shall derive this equation by a new approach which is simpler and which can be used to obtain similar equations for other distributions of a complicated nature.

It has been established (Iyer, 1949) that the r-th factorial moment is $r!$ multiplied by the expectation for r black-black joins. Now r black-black joins can be obtained from m points in the following independent ways: (i) all the r joins belong to the last $(m-1)$ points, (ii) $(r-1)$ joins belong to the last $(m-2)$ points while one join comes from the first two points, (iii) $(r-2)$ joins belong to the last $(m-3)$ points, while the last three points give two joins, etc. This means, the expectation of r black-black joins from m points is given by

$$\mathcal{E}(r, m) = \mathcal{E}(r, m-1) + \sum_{s=1}^{r} p^{s+1}\, \mathcal{E}(r-s, m-s-1), \qquad (2.1)$$

where $\mathcal{E}(r-s, m-s-1)$ represents the expectation for $(r-s)$ joins from $(m-s-1)$ points. Since

$$\mathcal{E}(r, m) = \frac{\mu'_{[r,m]}}{r\,!}, \qquad (2.2)$$

(2.1) reduces to

$$\frac{\mu'_{[r,m]}}{r\,!} = \frac{\mu'_{[r,m-1]}}{r\,!} + \sum_{s=1}^{r} \frac{p^{s+1}\,\mu'_{[r-s,\, m-s-1]}}{(r-s)\,!}. \qquad (2.3)$$

We now note that

$$M_m = 1 + \frac{\theta\mu'_{[1,m]}}{1!} + \frac{\theta^2\mu'_{[2,m]}}{2!} + \ldots \frac{\theta^r\mu'_{[r,m]}}{r!} \ldots, \qquad (2.4)$$

where $\theta = (e^t - 1)$ and M_m is the M.G.F. of the distribution for m points.

Substituting from (2.3) for $\mu'_{[r,m]}$ in (2.4), it reduces to

$$\begin{aligned} M_m &= M_{m-1} + \theta p^2 M_{m-2} + \theta^2 p^3 M_{m-3} + \ldots \theta^{m-1} p^m M_0 \\ &= M_{m-1} + p^2\theta\, \frac{E^{m-1} - (p\theta)^{m-1}}{E - p\theta}\, M_0, \end{aligned} \qquad (2.5)$$

where $E^r M_0 = M_r$. Operating both sides of (2.5) by $(E - p\theta)$, we get

$$(E - p\theta)\,(M_m - m_{m-1}) = p^2\theta\,\{E^{m-1} - (p\theta)^{m-1}\}\, M_0, \qquad (2.6)$$

which reduces to

$$M_{m+1} - (1 + p\theta)\, M_m + p\,(1-p)\,\theta\, M_{m-1} = 0, \qquad (2.7)$$

since $M_0 = 0$.

3. Black-White Joins

Wishart and Hirschfeld (1936) have obtained the difference equation of this distribution for the case of two colours. The methods used by them are rather complicated. We shall obtain the general expression for the distribution of black-white joins for points of k colours.

As in the case of black-black joins, r black-white joins can be selected from m points in the following ways which are all independent of one another. They are (i) all the r joins belong to the last $(m-1)$ points, (ii) $(r-1)$ joins belong to the last $(m-2)$ points while the first two points give one black-white join, (iii) $(r-2)$ joins come from the last $(m-3)$ points with the first three points giving two black-white joins, etc. Thus

$$\mathscr{E}(r, m) = \mathscr{E}(r, m-1) + 2\Sigma p_1^s p_2^s\, \mathscr{E}(r-2s+1, m-2s) + \Sigma p_1^s p_2^s (p_1+p_2)\, \mathscr{E}(r-2s, m-2s-1). \quad (3.1)$$

Therefore

$$\frac{\mu'_{[r,m]}}{r!} = \frac{\mu'_{[r,m-1]}}{r!} + 2\Sigma \frac{p_1^s p_2^s \mu'_{[r-2s+1,\, m-2s]}}{(r-2s+1)!} + \Sigma \frac{p_1^s p_2^s (p_1+p_2)\mu'_{[r-2s,\, m-2s-1]}}{(r-2s)!}, \quad (3.2)$$

where s can take values from 1 to $\frac{r}{2}$ or $\frac{r-1}{2}$ according as r is even or odd, and p_1 and p_2 are the probabilities of the points being black or white.

Setting this value of $\mu'_{[r,m]}$ in M_m we find that

$$M_m = M_{m-1} + 2\Sigma p_1^s p_2^s \theta^{2s-1} M_{m-2s} + \Sigma p_1^s p_2^s (p_1+p_2)\,\theta^{2s} M_{m-2s-1} \quad (3.3)$$

This can be written as

$$M_m - M_{m-1} = 2p_1p_2\theta\,[E^{m-2} + p_1p_2\theta^2 E^{m-4} + (p_1p_2\theta^2)^2 E^{m-6} + \ldots]\, M_0 + p_1p_2(p_1+p_2)\,\theta^2\,[E^{m-3} + p_1p_2\theta^2 E^{m-5} + (p_1p_2\theta^2)^2 E^{m-7}\ldots]\, M_0$$

$$= \left[\frac{2p_1p_2\theta E^m + p_1p_2(p_1+p_2)\,\theta^2 E^{m-1}}{E^2 - p_1p_2\theta^2}\right] M_0 \quad (3.4)$$

Operating both sides of (3.4) by $(E^2 - p_1p_2\theta^2)$, the equation becomes

$$(E^2 - p_1p_2\theta^2)(M_m - M_{m-1}) = 2p_1p_2\theta\, M_m + p_1p_2(p_1+p_2)\,\theta^2 M_{m-1},$$

i.e., $M_{m+2} - M_{m+1} = 2p_1p_2\theta M_m + p_1p_2\theta^2\{(p_1+p_2-1)M_{m-1} + M_m\}$. (3.5)

When the points can take only two colours, $p_1 + p_2 = 1$ and (3.5) reduces to

$$M_{m+2} - M_{m+1} - p_1 p_2 \theta (2 + \theta) M_m = 0. \tag{3.6}$$

This equation corresponds to the one given by Wishart and Hirschfeld (1936).

Equation (3.6) can also be obtained by the following procedure: r joins from the m points can be considered to be made up of (i) r joins from the last $(m-1)$ points, (ii) $(r-1)$ joins from the last $(m-2)$ points with one more join from the first two points, and (iii) $(r-2)$ joins from the last $(m-2)$ points with two joins from the first three points. The expectations for items (i) and (ii) are $\mathscr{E}(r, m-1)$ and $2p_1p_2\mathscr{E}(r-1, m-2)$ respectively. Item (iii) occurs only in the cases (a) and (b) shown below:

(r − 2) *joins from* (m − 2) *points*

× × × . . . .
$p_1\, p_2\, p_1$ (a)

× × × . . . .
$p_2\, p_1\, p_2$ (b)

$(r-2)$ *joins from* $(m-2)$ *points*

(p_1 and p_2 represent the probabilities of the points being black and white). It will be noted that p_1p_2 is common in both (a) and (b). Since $(p_1 + p_2) = 1$, the expectation for (a) and (b) together is $p_1p_2\mathscr{E}(r-2, m-2)$. Therefore,

$$\mathscr{E}(r, m) = \mathscr{E}(r, m-1) + 2p_1p_2\, \mathscr{E}(r-1, m-2) + p_1p_2\, \mathscr{E}(r-2, m-2),$$

i.e.,

$$\frac{\mu'_{[r,m]}}{r!} = \frac{\mu'_{[r,m-1]}}{r!} + 2p_1p_2 \frac{\mu'_{[r-1,\, m-2]}}{(r-1)!} + p_1p_2 \frac{\mu'_{[r-2,\, m-2]}}{(r-2)!}. \tag{3.8}$$

Substituting the above value of $\mu'_{[r,m]}$ in M_m, it can be seen that this reduces to

$$M_m - M_{m-1} - p_1p_2\theta (2 + \theta) M_{m-2} = 0. \tag{3.9}$$

4. Joins between Points of Different Colours

The author (1948) obtained the first and the second moments for the distribution of the total number of joins between points of different colours by working out the non-free distributions for (i) two black and $(m-2)$ white points, (ii) one black, one white and $(m-2)$ red points, and (iii) one black, one white, one red and $(m-3)$ green points. We shall obtain the first four cumulants by finding the expectations for one, two, three and four joins.

The probability for a join between two adjacent points belonging to different colours is

$$2\,\Sigma\, p_r\, p_s, \qquad r > s. \tag{4.1}$$

For m points on a line there are $(m-1)$ ways of obtaining a join between two adjacent points. Therefore the first moment is

$$\mu_1' = 2\,(m-1)\,\Sigma p_r p_s. \tag{4.2}$$

The second factorial moment is twice the expectation for two joins. Two joins between points of different colours can be formed from (i) three consecutive points and (ii) four points partitioned into two groups each having two adjacent points. The probability for (i) can be seen to be

$$(\,\Sigma\, p_r^2 p_s + 6\,\Sigma\, p_r\, p_s\, p_t). \tag{4.3}$$

The number of ways of selecting three consecutive points from m points is $(m-2)$. Therefore, the expectation for (i) is

$$(m-2)\,(\Sigma\, p_r^2 p_s + 6\,\Sigma\, p_r\, p_s\, p_t). \tag{4.4}$$

The chance of obtaining a join between points of different colours is $2\,\Sigma p_r\, p_s$. Therefore, the probability of obtaining two independent joins is $4\,(\,\Sigma p_r\, p_s)^2$. Now two independent joins can be formed from m points in $\binom{m-2}{2}$ ways. Therefore, the expectation for item (ii) is

$$\binom{m-2}{2}\,4\,(\,\Sigma p_r\, p_s)^2. \tag{4.5}$$

Hence

$$\mu'_{[2]} = 2\,(m-2)\,(\,\Sigma p_r^2 p_s + 6\,\Sigma p_r p_s p_t) + 4\,(m-2)^{(2)}\,(\,\Sigma p_r p_s)^2, \tag{4.6}$$

aud μ_2 reduces to

$$2\,(m-1)\,(1^2) + 2\,(m-2)\,(21) + 12\,(m-2)\,(1^3) - 4\,(3m-5)\,(1^2)^2, \tag{4.7}$$

where (21) and (1^t) stand for the symmetric functions $\Sigma p_r^2 p_s$ and $\Sigma p_1\, p_2 \ldots p_t$ in $p_1, p_2, \ldots p_k$ respectively.

The third factorial moment is 3! times the expectation for three joins. Three joins can be obtained from (i) four consecutive joins, (ii) five points partitioned into two parts, and (iii) six points divided into three groups of two each. The probabilities for (i), (ii) and (iii) are

$$(2\,\Sigma p_r^2 p_s^2 + 6\,\Sigma p_r^2 p_s p_t + 24\,\Sigma p_r\, p_s p_t p_u), \tag{4.8}$$

$$(2\,\Sigma p_r p_s)\,(\Sigma p_r^2 p_s + 6\,\Sigma p_r p_s p_t), \tag{4.9}$$

and

$$(2\,\Sigma p_r p_s)^3 \tag{4.10}$$

respectively. Note that the probabilities for (ii) and (iii) are calculated by using the results (4.1) and (4.3).

The sum of the expectations for (i), (ii) and (iii) gives

$$\begin{aligned}\frac{\mu'_{[3]}}{3!} &= (m-3)\,\{2\,(2^2) + 6\,(21^2) + 24\,(1^4)\}\\ &\quad + \binom{m-3}{2}\{2\,(1^2)\}\,\{(21) + 6\,(1^3)\}\\ &\quad + \binom{m-3}{3}\{2\,(1^2)\}^3.\end{aligned} \tag{4.11}$$

It can be shown now that

$$\begin{aligned}\kappa_3 &= (m-1)\,a + 6\,(m-2)\,b + 6\,(m-3)\,c + 12\,(5-2m)\,ab\\ &\quad + 3\,(5-3m)\,a^2 + 4\,(5m-11)\,a^3,\end{aligned} \tag{4.12}$$

where $a = 2\,(1^2)$, $b = \{(21) + 6\,(1^3)\}$, $c = \{2\,(2^2) + 6\,(21^2) + 24\,(1^4)\}$.

For the fourth factorial moment, we determine the expectation for obtaining four joins. Four joins can be formed from (i) five consecutive points, (ii) six points partitioned into two groups of two and four and two sets of three adjacent points, (iii) seven points divided into three sets of two, two and three adjacent points and (iv) eight points, divided into four groups of two points each. The probabilities for (i), (ii), (iii) and (iv) are

$$\{(32) + 2\,(31^2) + 36\,(21^3) + 12\,(2^21) + 5!\,(1^4)\},$$

$$\{2\,(1^2)\}\,\{2\,(2^2) + 6\,(21^2) + 24\,(1^4)\},\ \{(21) + 6\,(1^3)\}^2,$$

$\{2\,(1^2)\}^2\,\{(21) + 6\,(1^3)\}$ and $\{2\,(1^2)\}^4$ respectively.

Note that the expectations for (ii), (iii) and (iv) are obtained by using (4.1), (4.3) and (4.8). From these it can be seen that

$$\frac{\mu'_{[4]}}{4!} = (m-4)\,d + 2\binom{m-4}{2}ac + \binom{m-4}{2}b^2 + 3\binom{m-4}{3}a^2b + \binom{m-4}{4}a^4, \qquad (4.13)$$

where d is the probability of having four joins from five adjacent points. The fourth cumulant can now be worked out and it reduces to

$$\begin{aligned}&(m-1)\,a + 14\,(m-2)\,b + 36\,(m-3)\,c + 24\,(m-4)\,d\\ &-7\,(3m-5)\,a^2 - 12\,(5m-16)\,b^2 - 72\,(2m-5)\,ab\\ &-24\,(5m-17)\,ac + 24\,(5m-11)\,a^3 + 24\,(15m-44)\,a^2b\\ &-2\,(105m-279)\,a^4. \end{aligned} \qquad (4.14)$$

B. V. Sukhatme also has obtained the first four cumulants for this distribution from the first principles. His paper will be published shortly.

We shall obtain now the difference equations satisfied by the moment-generating-function. For the sake of simplicity let us confine our discussion in the first instance to the case of three colours.

It has been pointed out before that r joins can be considered to have been obtained in the following independent ways: (i) all the r joins belong to the last $(m-1)$ points, (ii) $(r-1)$ joins belong to the last $(m-2)$ points while the first two points give one join, and (iii) the last $(m-2)$ points contain $(r-2)$ joins while the first three give two joins. The expectations for items (i) and (ii) are $\mathcal{E}(r, m-1)$ and $2\,(p_1p_2 + p_2p_3 + p_3p_1)\,\mathcal{E}(r-1, m-2)$ respectively. The expectation for (iii) can be evaluated by considering the different ways of obtaining two joins from the first three points. There are twelve different ways of having two joins from the first three joints. They are indicated below:

(1) $\overset{\times}{p_1}\ \overset{\times}{p_2}\ \overset{\times}{p_1}$	(4) $\overset{\times}{p_1}\ \overset{\times}{p_3}\ \overset{\times}{p_1}$	(7) $\overset{\times}{p_2}\ \overset{\times}{p_3}\ \overset{\times}{p_1}$
(2) $\overset{\times}{p_2}\ \overset{\times}{p_1}\ \overset{\times}{p_2}$ (a)	(5) $\overset{\times}{p_1}\ \overset{\times}{p_3}\ \overset{\times}{p_2}$ (b)	(8) $\overset{\times}{p_2}\ \overset{\times}{p_3}\ \overset{\times}{p_2}$ (c)
(3) $\overset{\times}{p_1}\ \overset{\times}{p_2}\ \overset{\times}{p_3}$	(6) $\overset{\times}{p_3}\ \overset{\times}{p_1}\ \overset{\times}{p_3}$	(9) $\overset{\times}{p_3}\ \overset{\times}{p_2}\ \overset{\times}{p_3}$
(10) $\overset{\times}{p_2}\ \overset{\times}{p_1}\ \overset{\times}{p_3}$	(11) $\overset{\times}{p_3}\ \overset{\times}{p_1}\ \overset{\times}{p_2}$	(12) $\overset{\times}{p_3}\ \overset{\times}{p_2}\ \overset{\times}{p_1}$

In the above figure, p_1, p_2 and p_3 stand for black, white and red points. It will be observed that p_1p_2, p_1p_3 and p_2p_3 are common in the sets (*a*), (*b*) and (*c*) respectively. Further the last points in the sets (*a*), (*b*) and (*c*) are p_1, p_2 and p_3. Since $p_1 + p_2 + p_3 = 1$, the expectation for *r* joins from the arrangements of the first three points shown in the configurations 1 to 9 above is

$$(p_1p_2 + p_2p_3 + p_3p_1)\ \mathscr{E}(r - 2, m - 2).$$

We have now to consider the expectation for *r* joins when the first three points are arranged as in (10), (11) and (12), excluding all the cases included in (i) and (ii), and also in (*a*), (*b*) and (*c*). It can easily be seen that such cases arise only when the fourth point of the arrangements beginning with (10), (11) and (12) gives three joins. The different ways of obtaining three joins from four points based on the arrangements in (10), (11) and (12) are as shown below:

$$\left.\begin{matrix}(10.1)\ \overset{\times}{p_2}\,\overset{\times}{p_1}\,\overset{\times}{p_3}\,\overset{\times}{p_1}\\ (10.2)\ \overset{\times}{p_2}\,\overset{\times}{p_1}\,\overset{\times}{p_3}\,\overset{\times}{p_2}\end{matrix}\right\}\quad \left.\begin{matrix}(11.1)\ \overset{\times}{p_3}\,\overset{\times}{p_1}\,\overset{\times}{p_2}\,\overset{\times}{p_1}\\ (11.2)\ \overset{\times}{p_3}\,\overset{\times}{p_1}\,\overset{\times}{p_2}\,\overset{\times}{p_3}\end{matrix}\right\}\quad \left.\begin{matrix}(12.1)\ \overset{\times}{p_3}\,\overset{\times}{p_2}\,\overset{\times}{p_1}\,\overset{\times}{p_2}\\ (12.2)\ \overset{\times}{p_3}\,\overset{\times}{p_2}\,\overset{\times}{p_1}\,\overset{\times}{p_3}\end{matrix}\right\}$$

The first three points of the above arrangements contain points of all colours, p_1, p_2 and p_3. Considering the fourth point, p_1, p_2 and p_3 occur each twice. Hence the expectation of *r* joins from configurations beginning with (10), (11) and (12) is

$$2p_1p_2p_3\ \mathscr{E}(r - 3, m - 3).$$

Thus we obtain

$$\frac{\mu'_{[r,m]}}{r!} = \frac{\mu'_{[r,m-1]}}{r!} + 2\,\Sigma p_rp_s\frac{\mu'_{[r-1,m-1]}}{(r-1)!} + \Sigma p_rp_s\frac{\mu'_{[r-2,m-2]}}{(r-2)!} + 2p_1p_2p_3\frac{\mu'_{[r-3,m-3]}}{(r-3)!} \qquad (4.14)$$

Substituting this value in M_m, we get

$$M_m - M_{m-1} = 2\theta M_{m-2}\,\Sigma p_rp_s + \theta^2 M_{m-2}\,\Sigma p_r\,p_s + 2\theta^3 p_1p_2p_3M_{m-3}. \quad (4.15)$$

The same argument will give similar difference equations for four or more colours also. Before writing down the general expression for *k* colours, in order to give a clearer idea of the arguments used above, the equation for four colours is also obtained in detail.

Let the points take any one of four colours, say black, white, red and green with probabilities p_1, p_2, p_3 and p_4. We shall express the expectation for r joins from m points in terms of the distributions arising from $(m-1)$, $(m-2)$ and lower number of points. Now r joins can be obtained from m points in the following ways which are all different from one another: (i) all the r joins belong to the last $(m-1)$ points, (ii) $(r-1)$ joins belong to the last $(m-2)$ points, while the first two points give one join, and (iii) the first three points give two joins and the last $(m-2)$ points contain $(r-2)$ joins. The expectations for (i) and (ii) are $\mathscr{E}(r, m-1)$ and $2\Sigma p_r p_s\, \mathscr{E}(r-1, m-2)$ respectively. To obtain the expectation of (iii), which is rather complicated, we proceed as follows: There are 36 arrangements which give two joins from the first three points and $(r-2)$ joins from the last $(m-2)$ points. They may be split up in the manner indicated below:

(1) $\overset{\times}{p_1}\,\overset{\times}{p_2}\,\overset{\times}{p_1}$	(A)	(5) $\overset{\times}{p_1}\,\overset{\times}{p_3}\,\overset{\times}{p_1}$	(B)	(9) $\overset{\times}{p_1}\,\overset{\times}{p_4}\,\overset{\times}{p_1}$	(C)	(13) $\overset{\times}{p_2}\,\overset{\times}{p_3}\,\overset{\times}{p_1}$	(D)
(2) $\overset{\times}{p_2}\,\overset{\times}{p_1}\,\overset{\times}{p_2}$		(6) $\overset{\times}{p_1}\,\overset{\times}{p_3}\,\overset{\times}{p_2}$		(10) $\overset{\times}{p_1}\,\overset{\times}{p_4}\,\overset{\times}{p_2}$		(14) $\overset{\times}{p_2}\,\overset{\times}{p_3}\,\overset{\times}{p_2}$	
(3) $\overset{\times}{p_1}\,\overset{\times}{p_2}\,\overset{\times}{p_3}$		(7) $\overset{\times}{p_3}\,\overset{\times}{p_1}\,\overset{\times}{p_3}$		(11) $\overset{\times}{p_1}\,\overset{\times}{p_4}\,\overset{\times}{p_3}$		(15) $\overset{\times}{p_3}\,\overset{\times}{p_2}\,\overset{\times}{p_3}$	
(4) $\overset{\times}{p_1}\,\overset{\times}{p_2}\,\overset{\times}{p_4}$		(8) $\overset{\times}{p_1}\,\overset{\times}{p_3}\,\overset{\times}{p_4}$		(12) $\overset{\times}{p_4}\,\overset{\times}{p_1}\,\overset{\times}{p_4}$		(16) $\overset{\times}{p_2}\,\overset{\times}{p_3}\,\overset{\times}{p_4}$	
(17) $\overset{\times}{p_2}\,\overset{\times}{p_4}\,\overset{\times}{p_1}$	(E)	(21) $\overset{\times}{p_3}\,\overset{\times}{p_4}\,\overset{\times}{p_1}$	(F)	(25) $\overset{\times}{p_2}\,\overset{\times}{p_1}\,\overset{\times}{p_3}$		(29) $\overset{\times}{p_4}\,\overset{\times}{p_1}\,\overset{\times}{p_2}$	
(18) $\overset{\times}{p_2}\,\overset{\times}{p_4}\,\overset{\times}{p_2}$		(22) $\overset{\times}{p_3}\,\overset{\times}{p_4}\,\overset{\times}{p_2}$		(26) $\overset{\times}{p_2}\,\overset{\times}{p_1}\,\overset{\times}{p_4}$		(30) $\overset{\times}{p_4}\,\overset{\times}{p_1}\,\overset{\times}{p_3}$	
(19) $\overset{\times}{p_2}\,\overset{\times}{p_4}\,\overset{\times}{p_3}$		(23) $\overset{\times}{p_3}\,\overset{\times}{p_4}\,\overset{\times}{p_3}$		(27) $\overset{\times}{p_3}\,\overset{\times}{p_1}\,\overset{\times}{p_2}$		(31) $\overset{\times}{p_3}\,\overset{\times}{p_2}\,\overset{\times}{p_1}$	
(20) $\overset{\times}{p_4}\,\overset{\times}{p_2}\,\overset{\times}{p_4}$		(24) $\overset{\times}{p_4}\,\overset{\times}{p_3}\,\overset{\times}{p_4}$		(28) $\overset{\times}{p_3}\,\overset{\times}{p_1}\,\overset{\times}{p_4}$		(32) $\overset{\times}{p_3}\,\overset{\times}{p_2}\,\overset{\times}{p_4}$	
(33) $\overset{\times}{p_4}\,\overset{\times}{p_2}\,\overset{\times}{p_3}$		(34) $\overset{\times}{p_4}\,\overset{\times}{p_2}\,\overset{\times}{p_1}$		(35) $\overset{\times}{p_4}\,\overset{\times}{p_3}\,\overset{\times}{p_1}$		(36) $\overset{\times}{p_4}\,\overset{\times}{p_3}\,\overset{\times}{p_2}$	

The arrangements A, B, C, D, E and F have p_1p_2, p_1p_3, p_1p_4, p_2p_3, p_2p_4 and p_3p_4 respectively as the first two points, also p_1, p_2, p_3 and p_4 occur in each of the sets. Therefore the expectation for r joins from the configurations shown in A, B, C, D, E and F is $(\Sigma p_r p_s)\, \mathscr{E}(r-2, m-2)$. We have now to consider the remaining twelve (25–36) arrangements. Each of them can give three joins with a fourth point adjacent to them in three ways. There are thus another 36 arrangements in which the first four points contain three joins while the remaining $(m-3)$ points give $(r-3)$ joins. We group these arrangements as follows:

$$\underbrace{\begin{array}{cc} \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_2\,p_1\,p_3\,p_1 & p_3\,p_1\,p_2\,p_1 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_2\,p_1\,p_3\,p_2 & p_3\,p_2\,p_1\,p_2 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_3\,p_1\,p_2\,p_3 & p_3\,p_2\,p_1\,p_3 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_2\,p_1\,p_3\,p_4 & p_3\,p_1\,p_2\,p_4 \end{array}}_{\text{I}} \qquad \underbrace{\begin{array}{cc} \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_2\,p_1\,p_4\,p_1 & p_4\,p_1\,p_2\,p_1 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_2\,p_1\,p_4\,p_2 & p_4\,p_2\,p_1\,p_2 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_2\,p_1\,p_4\,p_3 & p_4\,p_1\,p_2\,p_3 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_4\,p_1\,p_2\,p_4 & p_4\,p_1\,p_2\,p_4 \end{array}}_{\text{II}}$$

$$\underbrace{\begin{array}{cc} \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_3\,p_1\,p_4\,p_1 & p_4\,p_1\,p_3\,p_1 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_3\,p_1\,p_4\,p_2 & p_4\,p_1\,p_3\,p_2 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_3\,p_1\,p_4\,p_3 & p_4\,p_3\,p_1\,p_3 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_4\,p_1\,p_3\,p_4 & p_4\,p_3\,p_1\,p_4 \end{array}}_{\text{III}} \qquad \underbrace{\begin{array}{cc} \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_4\,p_2\,p_3\,p_1 & p_3\,p_2\,p_4\,p_1 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_4\,p_2\,p_3\,p_2 & p_3\,p_2\,p_4\,p_2 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_3\,p_2\,p_4\,p_3 & p_4\,p_3\,p_2\,p_3 \\ \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_4\,p_2\,p_3\,p_4 & p_4\,p_3\,p_2\,p_4 \end{array}}_{\text{IV}}$$

$$\underbrace{\begin{array}{cccc} \times\;\times\;\times\;\times & \times\;\times\;\times\;\times & \times\;\times\;\times\;\times & \times\;\times\;\times\;\times \\ p_3\,p_2\,p_1\,p_4 & p_4\,p_2\,p_1\,p_3 & p_4\,p_3\,p_1\,p_2 & p_4\,p_3\,p_2\,p_1 \end{array}}_{\text{V}}$$

It is obvious that the expectation for r joins for the arrangements commencing with those indicated in I, II, III and IV is $2\,(p_1p_2p_3 + p_1p_2p_4 + p_1p_3p_4 + p_2p_3p_4)\ \mathscr{E}\,(r-3,\, m-3)$. We are now left with the configurations shown in V. Each of them can give four joins in three ways with a fifth point adjoining them. These configurations are given below:

$$\left.\begin{array}{c} \times\;\times\;\times\;\times\;\times \\ p_3\,p_2\,p_1\,p_4\,p_1 \\ \times\;\times\;\times\;\times\;\times \\ p_3\,p_2\,p_1\,p_4\,p_2 \\ \times\;\times\;\times\;\times\;\times \\ p_3\,p_2\,p_1\,p_4\,p_3 \\ \times\;\times\;\times\;\times\;\times \\ p_4\,p_3\,p_2\,p_1\,p_4 \end{array}\right\}\text{X} \qquad \left.\begin{array}{c} \times\;\times\;\times\;\times\;\times \\ p_4\,p_2\,p_1\,p_3\,p_1 \\ \times\;\times\;\times\;\times\;\times \\ p_4\,p_2\,p_1\,p_3\,p_2 \\ \times\;\times\;\times\;\times\;\times \\ p_4\,p_3\,p_2\,p_1\,p_3 \\ \times\;\times\;\times\;\times\;\times \\ p_4\,p_2\,p_1\,p_3\,p_4 \end{array}\right\}\text{Y} \qquad \left.\begin{array}{c} \times\;\times\;\times\;\times\;\times \\ p_4\,p_3\,p_1\,p_2\,p_1 \\ \times\;\times\;\times\;\times\;\times \\ p_4\,p_3\,p_2\,p_1\,p_2 \\ \times\;\times\;\times\;\times\;\times \\ p_4\,p_3\,p_1\,p_2\,p_3 \\ \times\;\times\;\times\;\times\;\times \\ p_4\,p_3\,p_1\,p_2\,p_4 \end{array}\right\}\text{Z}$$

The expectation of r joins for the arrangements commencing with those shown in (X), (Y) and (Z) is $3p_1p_2p_3p_4\ \mathcal{E}(r-4, m-4)$. From the above discussion, it is obvious that

$$\begin{aligned}\mathcal{E}(r, m) = \mathcal{E}(r, m-1) + 2\mathcal{E}(r-1, m-2)\,\Sigma p_r\, p_s\, + \\ + \mathcal{E}(r-2, m-2)\,\Sigma p_r\, p_s + 2\mathcal{E}(r-3, m-3) \\ \Sigma p_r\, p_s\, p_t + 3\mathcal{E}(r-4, m-4)\, p_1p_2p_3p_4\end{aligned} \tag{4.16}$$

We get therefore that

$$\frac{\mu'_{[r,m]}}{r!} = \frac{\mu'_{[r,\,m-1]}}{r!} + 2\,\Sigma p_r p_s \frac{\mu'_{[r-1,\,m-1]}}{(r-1)!} +$$

$$\Sigma p_r\, p_s \frac{\mu'_{[r-2,\,m-2]}}{(r-2)!} + 2\,\Sigma p_r p_s p_t \frac{\mu'_{[r-3,\,m-3]}}{(r-3)!} + 3p_1p_2p_3p_4 \frac{\mu'_{[r-4,\,m-4]}}{(r-4)!} \tag{4.17}$$

Substituting this value of the r-th factorial moment in M_m, we find that

$$\begin{aligned}M_m = M_{m-1} + 2\theta M_{m-2}\,\Sigma p_r p_s + \theta^2 M_{m-2}\,\Sigma p_r\, p_s + 2\theta^3 M_{m-3}\,\Sigma p_r\, p_s\, p_t \\ + 3\theta^4\, p_1p_2p_3p_4 M_{m-4}.\end{aligned} \tag{4.18}$$

By proceeding in this manner, it can be shown that the difference equation for k colours reduces to the form

$$M_m - M_{m-1} = 2\theta a_2 M_{m-2} + \sum_{r=2}^{k} (r-1)\,\theta^r\, a_r\, M_{m-r}, \tag{4.19}$$

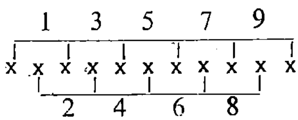
where a_r stands for (1^r), the monomial symmetric function of degree r in $p_1, p_2, \ldots p_k$.

5. Triplets, Quadruplets, etc.

We may define triplets, quadruplets . . . . s-plets as sets of three, four . . . . s adjacent points of a specified colour. The factorial moments for the distribution of the s-plet both for free and non-free sampling have been obtained by B. V. Sukhatme (1950). The difference equations satisfied by the M.G.F.'s of these distributions are derived in this section.

Triplets. r triplets from m points can be considered to be composed of (i) r triplets from $(m-1)$ points, (ii) one triplet from the first three points, *plus* $(r-1)$ triplets from the remaining $(m-3)$

points, (iii) two triplets from the first four points with $(r-2)$ of them from the remaining $(m-4)$ points, (iv) three or two triplets taken in such a way as to include the first five points and $(r-3)$ or $(r-2)$ triplets from the last $(m-5)$ points, (v) four or three triplets covering the first six points and $(r-4)$ or $(r-3)$ triplets from the remaining $(m-6)$ points and so on. When we consider the first $(s+2)$ points, it is possible to have $s, (s-1), (s-2), \ldots \left(\frac{s}{2}+1\right)$ or $\left(\frac{s+1}{2}\right)$ triplets according as s is even or odd, covering all the $(s+2)$ points. The number of ways of having $(s-k)$ triplets from $(s+2)$ points is $\binom{s-1-k}{k}$. This statement is illustrated below. Let there be eleven points on a line.



The maximum number of triplets that can be obtained from eleven points is nine. Suppose we want to find the number of ways of having six triplets, which connect all the eleven points. This can be achieved by removing three of the seven 2–8 triplets in such a manner that the six triplets left should have all the eleven points distributed among them. This can be done by removing three triplets from 2–8 satisfying the condition that no two of them are adjoining ones. The number of ways of removing three triplets subject to this restriction is $\binom{5}{3}$. Thus $\binom{5}{3}$ is the number of ways of having six triplets connecting the eleven points.

Hence we find that

$$\frac{\mu'_{[r,m]}}{r!} = \mathcal{E}(r, m) = \mathcal{E}(r, m-1) + \sum_{s=1}^{m-2} p^{s+2} \times \Sigma \binom{s-1-k}{k} \times \mathcal{E}(r-s+k, m-s-2), \tag{5.1}$$

where the second summation extends from $k=0$ to the greatest integer less than $(s-1)\frac{1}{2}$. Substituting the above value of $\mu'_{[r,m]}$ in the moment-generating-function for the distribution of triplets, it reduces to

شماره دفتر اساس ف ش

کلاسمان پرونده ۳۱۷

تاریخ صدور ۴۲/۱۰/۱۴

ستاد بزرگ ارتشتاران

اداره نظام وظیفه عمومی

برگ معافیت از خدمت زیر پرچم

اسم پرویز شهرت [illegible] پدر محمد علی شماره شناسنامه ۷ محل صدور شناسنامه [illegible]

تاریخ تولد ۱۲۹۷ [illegible] محل اقامت طهران طبق تصویبنامه هیئت دولت درزمان صلح

از خدمت زیر پرچم معاف میگردد.

درموقع تغییر محل اقامت باید حوزه مربوطه را آگاه سازد

رئیس دفتر احضار حوزه

رئیس حوزه [illegible]

گروه ۲ سال [illegible] ۴۶

$$M_m - M_{m-1} = \sum_{s=1}^{m-2} p^{s+2} \left\{ \Sigma \binom{s-1-k}{k} \theta^{s-k} \right\} M_{m-s-2}$$

$$= p^3\theta\, [E^{m-3} + p\theta\,(E^{m-4} + pE^{m-5}) + p^2\theta^2\,(E^{m-5}$$

$$+ 2pE^{m-6} + p^2E^{m-7}) + p^3\theta^3\,(E^{m-6} + 3pE^{m-7} + 3p^2E^{m-8}$$

$$+ p^3E^{m-9}) + \ldots p^s\theta^s \left\{ E^{m-s-3} + \binom{s}{1} pE^{m-s-4} \right.$$

$$\left. + \binom{s}{2} p^2E^{m-s-5} \ldots \right\} \ldots]\, M_0$$

$$= p^3\theta\, [E^{m-3} + p\theta\,(1 + pE^{-1})\,E^{m-4} + p^2\theta^2\,(1 + pE^{-1})^2\,E^{m-5}$$

$$+ p^3\theta^3\,(1 + pE^{-1})^3\,E^{m-6} + \ldots p^s\theta^s\,(1 + pE^{-1})^s\,E^{m-s-3}$$

$$+ \ldots p^{m-3}\theta^{m-3}\,(1 + pE^{-1})^{m-3}]\, M_0$$

$$= p^3\theta\, \frac{E^{m-2} - \{p\theta\,(1 + pE^{-1})\}^{m-2}}{E - p\theta\,(1 + pE^{-1})}\, M_0 \qquad (5.2)$$

Operating both besides by $E - p\theta\,(1 + pE^{-1})$ the above equation reduces to

$$M_{m+1} - (1 + p\theta)\, M_m + pq\theta M_{m-1} + p^2q\theta M_{m-2} = 0. \qquad (5.3)$$

Quadruplets. A similar method will give the difference equation satisfied by the M.G.F.'s of the distribution of quadruplets and s-plets. But the derivation becomes more and more complicated. For the sake of clarifying the principles used in this analysis, the difference equation for the quadruplets also is derived in detail.

As in triplets, we shall set out the different ways of obtaining r quadruplets from m points. Now r quadruplets can be considered to be made up of (i) r quadruplets from the last $(m-1)$ points, (ii) one quadruplet from the first four points and $(r-1)$ of them from the last $(m-4)$ points, (iii) $(r-2)$ quadruplets from the last $(m-5)$ points and two from the first five points, (iv) $(r-3)$ or $(r-2)$ quadruplets from the last $(m-6)$ points with three and two respectively from the first six points, (v) $(r-4)$ or $(r-3)$ or $(r-2)$ quadruplets from the last $(m-7)$ points, with four, three or two respectively from the first seven points and so on.

The number of ways of obtaining $(s-k)$ quadruplets from $(s+3)$ [illegible]s covering all the $(s+3)$ points is

$$\sum_{t=0}^{k} \binom{s-k-1}{k-t} \binom{k-t}{t} \qquad (5.4)$$

The condition to be satisfied here is that no three of the quadruplets removed can be adjoining ones. It will now follow that

$$\frac{\mu'_{[r,m]}}{r!} = \frac{\mu'_{[r,\,m-1]}}{r\,!} + \sum_{s=0}^{m-3} p^{s+3} \left[\Sigma \left\{ \sum_{t=0}^{k} \binom{s-k-1}{k-t} \binom{k-t}{t} \right\} \times \frac{\mu'_{[r-s+k,\; m-s-3]}}{(r-s+k)!} \right] \tag{5.5}$$

where k extends from 0 to the greatest integer less than $\frac{2(s-1)}{3}$. Using this value it can be seen that

$$M_m - M_{m-1} = p^4\theta\,[E^{m-4} + p\theta\,(1 + pE^{-1} + p^2E^{-2})\,E^{m-5}$$

$$+ p^2\theta^2\,(1 + pE^{-1} + p^2E^{-2})^2\,E^{m-6} + \ldots]\,M_0$$

$$= \frac{p^4\theta\,[E^{m-3} - \{p\theta\,(1 + pE^{-1} + p^2E^{-2})\}^{m-3}]}{E - p\theta\,(1 + pE^{-1} + p^2E^{-2})}\,M_0$$

Operating both sides by $E - p\theta\,(1 + pE^{-1} + p^2E^{-2})$, we get

$$M_{m+1} - (1 + p\theta)\,M_m + pq\theta M_{m-1} + p^2q\theta M_{m-2} + p^3q\theta M_{m-3} = 0. \tag{5.6}$$

It can be seen now that the difference equation for an s-plet, *i.e.*, a set containing s points of a specified character is

$$M_{m+1} - M_m\,(1 + p\theta) + pq\theta \sum_{r=1}^{s-1} p^{r-1}\,M_{m-r} = 0. \tag{5.7}$$

6. Black Runs of Length l

The first and the second moments for free sampling have been given by Mood (1940). The second moment given by Mood is not correct. The correct moment is

$$2\,[3p^{2l}q^2 + 3\,(m - 2l - 2)\,p^{2l}q^3 + (m - 2l - 2)\,(m - 2l - 3)\,p^{2l}q^4]$$

$$+ 2p^lq + (m - l - 1)\,p^lq^2 - [2p^lq + (m - l - 1)\,p^lq^2]^2.$$

The higher moments of this distribution for both free and non-free sampling have been obtained by B. V. Sukhatme (1950). We shall derive the difference equation satisfied by the M.G.F.'s of this distribution.

Following the arguments used in the previous sections, can be formed from m points in the following ways:

(i) All the runs are distributed in the last $(m-1)$ points, (ii) $(r-1)$ runs belong to the last $(m-r-2)$ points and one run preceded and followed by points of different colours belong to $(r+2)$ points in the beginning [runs formed from $(r+1)$ points are not to be considered because such configurations are included in the $(r-1)$ runs considered from $(m-r-2)$ points], (iii) $(r-2)$ runs are distributed in the $(m-2r-3)$ points while two runs are obtained from the first $(2r+3)$ points and so on. From this consideration, it will follow that

$$M_m - M_{m-1} = \Sigma\, p^{ls} q^{s+1}\, \theta^s\, M_{m-ls-s-1}, \tag{6.1}$$

where s extends from 1 to the lowest integer greater than s/l. Using the operator E, (6.1) reduces to

$$M_{m+l+1} - M_{m+l} - p^l q\theta M_m + p^{l+1} q\theta M_{m-1} = 0. \tag{6.2}$$

7. Runs of Length l or More

Mood (1940) has derived the first and the second moments of these distributions. Mosteller (1941) has obtained the probability of getting runs of length l or more for two kinds of elements. The method developed by the author (1949) is used here to obtain the factorial moments and the M.G.F. of this distribution.

The first moment is the expectation for a run of length l or more. A black-run of length l or more can be formed from (i) the first l black points and (ii) l black points followed or preceded by a white point. There is only one way of obtaining (i), while (ii) can be had in $(m-l)$ ways. Therefore, the expectation, μ_1' for a black-run of length l or more is

$$\mu'_{[1,m]} = p + (m-l)\, p^l q. \tag{7.1}$$

The expectation for two runs of length l or more is $\frac{1}{2}\mu'_{[2,m]}$. Two black-runs of length l or more can be obtained in the following ways: (i) from $2l$ black points with a white point in the centre and (ii) $2l$ black points with two white points as shown below:

$$\overbrace{\cdots\cdots\cdots}^{l}\ \mathrm{x}\ \overbrace{\cdots\cdots\cdots}^{l}\ \text{ and }\ \overbrace{\cdots\cdots\cdots}^{l}\ \mathrm{x};\ \mathrm{x}\ \overbrace{\cdots\cdots\cdots}^{l}$$

Dots and crosses represent black and white points respectively. It can be easily seen that the expectations for (i) and (ii) are

$$(m-2l)\, p^{2l} q \text{ and } \binom{m-2l}{2} p^{2l} q^2 \text{ respectively.}$$

Therefore

$$\mu'_{[2,m]} = 2\,(m - 2l)\,p^{2l}q + (m - 2l)^{(2)}\,p^{2l}q^2. \tag{7.2}$$

The above expression differs from that given by Mood (1940). The third factorial moment is 3! times the expectation for three runs of length l or more. Three such runs can be obtained only in the following two ways shown below:

$\overbrace{\cdot\cdot\cdot\cdot\cdot}^{l}\times\overbrace{\cdot\cdot\cdot\cdot\cdot}^{l};\times\overbrace{\cdot\cdot\cdot\cdot\cdot}^{l}$ (1) $3l$ black points and two white points

$\underbrace{\cdot\cdot\cdot\cdot\cdot}_{l}\times;\times\underbrace{\cdot\cdot\cdot\cdot\cdot}_{l};\times\underbrace{\cdot\cdot\cdot\cdot\cdot}_{l}$ (2) $3l$ black points and three white points.

The expectations for (1) and (2) are

$$\binom{m-3l}{2}p^{3l}q^2 \text{ and } \binom{m-3l}{3}p^{3l}q^3 \text{ respectively.}$$

Hence

$$\mu'_{[3,m]} = 3\,(m - 3l)^{(2)}\,p^{3l}q^2 + (m - 3l)^{(3)}\,p^{3l}q^3 \tag{7.3}$$

Similarly the 4th factorial moment works out to

$$\mu'_{[4,m]} = 4\,(m - 4l)^{(3)}\,p^{4l}q^3 + (m - 4l)^{(4)}\,p^{4l}q^4 \tag{7.4}$$

The r-th factorial moment is

$$\mu'_{[r,m]} = r\,(m - rl)^{(r-1)}\,p^{rl}q^{r-1} + (m - rl)^{(r)}\,p^{rl}q^r \tag{7.5}$$

The difference equation satisfied by the M.G.F.'s is

$$M_m - M_{m-1} - p^l q\theta M_{m-l-1} = 0 \tag{7.6}$$

It is possible to have r runs of length l in m points as follows: (i) all the r runs belong to the last $(m - 1)$ points and (ii) $(r - 1)$ runs belong to the last $(m - l - 1)$ points while one run preceded or succeeded by a white point comes from the first $(l + 1)$ points. Thus

$$\mathscr{E}\,(r, m) = \mathscr{E}\,(r, m - 1) + p^l q \mathscr{E}\,(r - 1, m - l - 1),$$

i.e.,

$$\frac{\mu'_{[r,m]}}{r!} = \frac{\mu'_{[r,\,m-1]}}{r!} + \frac{p^l q\,\mu'_{[r-1,\,m-l-1]}}{(r - 1)!} \tag{7.7}$$

This will be evident from (7.5) also.

When $l = 1$, the above distribution reduces to that of the number of black runs. This distribution has already been discussed by the author (1948) by a different approach. The factorial moments reduce to the following:

$$\mu'_{[1,\, m]} = (m-1)\, pq + p, \tag{7.8}$$

$$\mu'_{[2,\, m]} = 2\,(m-2)\, p^2 q + (m-2)^{(2)}\, p^2 q^2, \tag{7.9}$$

$$\mu'_{[3,\, m]} = 3\,(m-3)^{(2)}\, p^3 q^2 + (m-3)^{(3)}\, p^3 q^3, \tag{7.10}$$

$$\mu'_{[4,\, m]} = 4\,(m-4)^{(3)}\, p^4 q^3 + (m-4)^{(4)}\, p^4 q^4 . \tag{7.11}$$

The cumulants obtained from the above moments agree with those given by the author (1948). The moments for non-free sampling can be obtained from those for free sampling by substituting

$$\frac{n_1^{(r)}\, n_2^{(s)}}{m^{(r+s)}} \text{ for } p^r q^s.$$

It may also be noted that the M.G.F. difference equation for the distribution of the number of black runs is

$$M_m - M_{m-1} - pq\theta M_{m-2} = 0. \tag{7.12}$$

8. Total Number of Black-Black and Black-White Joins for Two Colours

The factorial moments and the difference equations satisfied by the M.G.F.'s of the distributions for black-black and black-white joins have been obtained in earlier sections. It may sometimes be desirable to develop some statistical tests on the basis of both these distributions together. The first and the second factorial moments and the difference equations for the distribution of the total number of black-black and black-white joins are given below:

$$\mu'_{[1,\, m]} = (m-1)\, p\, (1+q) \tag{8.1}$$

$$\mu'_{[2,\, m]} = 2\,(m-2)\, p\,(1+pq) + (m-2)^{(2)}\, [p\,(1+q)]^2 \tag{8.2}$$

$$M_m - (1+p\theta)\, M_{m-1} - pq\theta\,(1+\theta)\, M_{m-2} = 0. \tag{8.3}$$

The difference equation has been obtained by the method described in the latter part of Section 2.

9. Total Number of Black and White Runs of Length s

It has not been possible to obtain the M.G.F. difference equation for this distribution. However, the first two factorial moments are given here.

$$\mu'_{[1,\, m]} = 2pq\,(p^{s-1} + q^{s-1}) + (m - s - 1)\, p^2q^2\,(p^{s-2} + q^{s-2}) \quad (9.1)$$

$$\mu'_{[2,\, m]} = 6p^2q^2\,(p^{2s-2} + q^{2s-2}) + 6\,(m - 2s - 2)\, p^3q^3\,(p^{2s-3} + q^{2s-3})$$
$$+ (m - 2s - 2)^{(2)}\, p^4q^4\,(p^{2s-4} + q^{2s-4}) + 2\,\{2p^sq^s + 4\,(m - 2s - 1)$$
$$p^{s+1}q^{s+1} + (m - 2s - 2)^{(2)}\, p^{s+2}q^{s+2}\}. \quad (9.2)$$

10. Total Number of Black and White Runs of Length s or More

The M.G.F. difference equation of this distribution also has not yet been obtained. The author hopes to obtain them soon. The first and the second factorial moments reduce to the following expressions:

$$\mu'_{[1,\, m]} = (p^s + q^s) + (m - s)\, pq\,(p^{s-1} + q^{s-1}) \quad (10.1)$$

$$\mu'_{[2,\, m]} = 2\,(m - 2s)\, pq\,(p^{2s-1} + q^{2s-1}) + (m - 2s)^{(2)}\, p^2q^2$$
$$(p^{2s-2} + q^{2s-2}) + 4\,(m - 2s + 1)\, p^sq^s$$
$$+ 2\,(m - 2s)^{(2)}\, p^{s+1}\, q^{s+1} \quad (10.2)$$

Substituting $s = 1$, (10.1) and (10.2) reduce to

$$\mu'_{[1,\, m]} = 2\,(m - 1)\, pq + 1, \quad (10.3)$$

and

$$\mu'_{[2,\, m]} = 2\,(m - 2)\, pq + 2\,(m - 2)^{(2)}\, p^2q^2$$
$$+ 4\,(m - 1)\, pq + 2\,(m - 2)^{(2)}\, p^2q^2$$
$$= 2\,(3m - 4)\, pq + 4\,(m - 2)^{(2)}\, p^2q^2 \quad (10.4)$$

From this μ_2 reduces to the same expression as the one given by the author in a previous communication (1948).

11. Discussion on the Limiting Forms of the Distributions

The difference equations satisfied by the M.G.F.'s of the various distributions take the form

$$M_m - M_{m+1} + \theta \sum_{s=1}^{k} f_s(\theta)\, M_{m-s} = 0, \quad (11.1)$$

where $f_s(\theta)$ is a function in $\theta = e^t - 1$. The solution of this equation is

$$M_m = \sum_{1}^{k} c_r\, a_r^{\,m}, \quad (11.2)$$

where c_r's are constants and α_r's are the roots of the equation

$$x^k - x^{k-1} - \theta \sum_{s=1}^{k} f_s(\theta)\, x^{k-s} = 0. \tag{11.3}$$

The above equation has all the roots, excepting one, zero when $\theta = 0$. Let this non-zero root be α_1. Now the r-th cumulant, κ_r, is

$$\left[\left(\frac{d}{dt}\right)^r \log M_m\right]_{t=0}$$

$$= \left[m\left(\frac{d}{dt}\right)^r \log \alpha_1 + \left(\frac{d}{dt}\right)^r \log\left\{1 + \left(\frac{c_2}{c_1}\right)\left(\frac{\alpha_2}{\alpha_1}\right)^m + \left(\frac{c_3}{c_1}\right)\left(\frac{\alpha_3}{\alpha_1}\right)^m\right.\right.$$

$$\left.\left. + \ldots\right\} + \left(\frac{d}{dt}\right)^r \log c_1\right]. \tag{11.4}$$

Since $\alpha_2, \alpha_3, \alpha_4 \ldots$ are zero when $\theta = 0$, κ_r is a linear expression in m when $r \leqslant m$. Thus all the cumulants of the distributions are linear expressions in m and therefore the γ's tend to the limit zero when m tends to infinity. Hence the distributions tend to the normal form when m tends to infinity. It may, however, be observed that when the probability of the points taking the colours black, white, etc., is very small, the distributions like those of black-black, black-white joins, etc., will tend to the Poisson form. Sukhatme, B. V. (1950) will discuss in detail the values of m and p for which these distributions can be assumed to be normal for all practical purposes.

12. Summary

By using certain special methods developed by the author, the moments and the difference equations satisfied by the M.G.F.'s of a number of distributions that arise from m points possessing one of k characters or colours arranged at random on a line have been discussed in this paper. The distributions considered are (i) the number of joins between adjacent points of the same or different colours, (ii) total number of runs of a given length or more for one or two specified characters or colours. It has been shown that all the distributions, excepting in cases where the probability of the points assuming a specified colour is very small, tend to the normal form when m tends to infinity.

REFERENCES

Krishna Iyer, P. V. .. "The theory of probability distributions of points on a line," *J. Ind. Soc. Agric. Stat.*, 1948, **1**, 173.

——————— .. "Calculation of factorial moments of certain probability distributions," *Nature*, 1949, **164**, 282.

Mood, A. M. .. "The distribution theory of runs," *Ann. Math. Stat.*, 1943, **11**, 367.

Mosteller, F. .. "Note on an application of runs to quality Control," *ibid.*, 1941, **12**, 228.

Sukhatme, B. V. .. "Contributions to the theory of probability distributions of points on a line," 1950 (to be published).

Wishart, J. and Hirschfeld, H. O. .. "A theorem concerning the distribution of joins between line segments," *J. Math. Soc.* London, 1936, **11**, 227.