FACTORIAL MOMENTS AND CUMULANTS OF DISTRIBUTIONS ARISING IN MARKOFF CHAINS

By P. V. Krishna Iyer

Indian Council of Agricultural Research, New Delhi

Introduction

The writer (1950) has considered a number of distributions arising from points of k characters distributed at random on a lattice. These distributions can be treated on the basis of simple or complex Markoff chains. Their cumulants have been calculated by obtaining the factorial moments. Though this method has been found to be satisfactory, the reduction involved in the calculation of the cumulants is so heavy that it has become necessary to examine the possibility of simplifying this procedure. The cumulants of most of the distributions are linear functions of the number of points on the lattice and it appears that it should be possible to develop a simple method for obtaining the cumulants. Moran (1948) has made some progress in this direction. But he has dealt only with the corrected moments and as they are not linear in the number of points on the lattice, the methods are not simple. The purpose of this paper is to give (i) a rigorous proof for the author's result for calculating the factorial moments and (ii) a new method for obtaining the cumulants directly for distributions arising from Markoff chains.

2. Factorial Moments

Theorem.—The r-th factorial moment for the distribution of the total number of some defined character arising in Markoff chains on a line or on a lattice is $r!$ (the expectation for r of the characters).

Proof.—Let the variate (*i.e.*, the defined character) assume the value *one* or *zero* according as it occurs or does not occur. Suppose that the character can occur in N ways and let $x_1, x_2, \ldots x_N$ be the variables for these N ways. Assume

$$X = \sum_{1}^{N} x_r.$$

Then

$$E(X) = E(\overset{N}{\underset{1}{\Sigma}} x_r) = NE(x_r)$$

$$= \text{expectation for the total number of characters} \tag{2.1}$$

$$E(X^2) = E(\overset{N}{\underset{1}{\Sigma}} x_r)^2 = E(\overset{N}{\underset{1}{\Sigma}} x_r^{\,2}) + 2E\,\Sigma\,(x_r x_s)$$

$$= E(X) + N(N-1)(Ex_r x_s)$$

Transposing we get

$$EX(X-1) = 2!\,(\text{expectation for two of the characters}) \tag{2.2}$$

$$E(X^3) = E(\overset{N}{\underset{1}{\Sigma}} x_r)^3 = E(\Sigma x_r^{\,3}) + 3E(\Sigma x_r^{\,2} x_s) + 3E(\Sigma x_r x_s^{\,2})$$

$$+ 3!\,E(\Sigma\, x_r x_s x_t)$$

$$= E(X) + 6E(\Sigma x_r x_s) + 3!\,E(\Sigma x_r x_s x_t)$$

$$= E(X) + 6E\,X(X-1) + 3!\,E(\Sigma x_r x_s x_t)$$

Transposing we obtain

$$EX(X-1)(X-2) = 3!\,(\text{expectation for three characters}) \tag{2.3}$$

Similarly it can be shown that

$$E(X^r) = E\left(\overset{N}{\underset{1}{\Sigma}} x_s\right)^r$$

$$= E(\Sigma x_s^{\,r}) + \frac{r!}{u!\,v!}\,E(\Sigma x_l^{\,u} x_m^{\,v}) + \frac{r!}{u!\,v!\,w!}\,E(\Sigma x_l^{\,u} x_m^{\,v} x_n^{\,w}).$$

$$\ldots + r!\,E(\Sigma x_l x_m \ldots r \text{ terms})$$

The above expression reduces to

$$E(X^r) = \Delta o^r E(X) + \frac{\Delta^2 o^r}{2!}\,E(X^{(2)}) + \frac{\Delta^3 o^r}{3!}\,E(X^{(3)}) \ldots +$$

$$\frac{\Delta^{r-1} o^r}{(r-1)!}\,E(X^{(r-1)}) + r!\,E(\Sigma x_l x_m \ldots r \text{ terms}) \tag{2.4}$$

Using the identity that

$$x^r = x^{(1)}\,\Delta o^r + \frac{x^{(2)}\Delta^2 o^r}{2!} \ldots \frac{x^{(r)}}{r!}\,\Delta^r o^r$$

it can be easily seen that (2.3) reduces to

$$E(X^{(r)}) = r!\,(\text{expectation for } r \text{ of the characters}) \tag{2.5}$$

3. Calculation of Cumulants

It can be shown that κ_r for the distribution of the total number of black-black joins for n points on a line is given by

$$\kappa_r' = r! \sum_{t=0}^{r} (n-t) \sum \frac{\kappa'_{s_1 s_2 s_3 \ldots s_t}}{s_1!\, s_2! \ldots s_t!}$$

where $s_1 + s_2 \ldots s_t = r$ and $\kappa'_{s_1 s_2 \ldots s_t}$ stand for the joint product cumulant of

$$(x_1 - p^2)^{s_1} (x_2 - p^2)^{s_2} \ldots (x_t - p^2)^{s_t}.$$

This result can be extended for the distributions on a lattice and also for that of runs, etc.

We shall first establish this result for the linear case. Let there be n points on a line. The number of ways in which a black-black or a black-white (including white-black) join can be obtained is $(n-1)$. Let the variables $x_1, x_2, \ldots x_{n-1}$ take the values one or zero according as a join occurs or does not occur. Now

$$\begin{aligned}\kappa_4 = \kappa \left(\sum_1^n z_r\right)^4 &= \kappa\,(\Sigma z_r^4) + \frac{4!}{3!\,1!}\,\kappa\,\{\Sigma\,(z_r^3 z_s + z_r z_s^3)\} \\ &+ \frac{4!}{2!\,2!}\,\kappa\,(\Sigma z_r^2 z_s^2) + \frac{4!}{2!\,1!\,1!}\,\kappa\,\{\Sigma\,(z_r^2 z_s z_t \\ &+ z_r z_s^2 z_t + z_r z_s z_t^2)\} + 4!\,\kappa\,(\Sigma z_r z_s z_t z_u), \qquad (3.1)\end{aligned}$$

where $z_r = (x_r - a_r)$ and a_r being the expectation of x_r. It can be easily seen that

$$\kappa\,(z_r^{p_1} z_s^{p_2} z_t^{p_3} p_u^{p_4}) = 0,$$

where $r > s+1 > t+1 > u+1$. Hence (3.1) reduces to

$$\begin{aligned}\kappa_4 &= (n-1)\,\kappa\,(z^4) + \frac{4!}{3!\,1!}\,(n-2)\,\{\kappa\,(z_1^3 z_2) + \kappa\,(z_1 z_2^3)\} \\ &+ \frac{4!}{2!\,2!}\,(n-2)\,\kappa\,(z_1^2 z_2^2) + \\ &\quad \frac{4!}{2!\,1!\,1!}\,(n-3)\,\kappa\,\{(z_1^2 z_2 z_3) + \kappa\,(z_1 z_2^2 z_3) + \kappa\,(z_1 z_2 z_3^2)\} \\ &+ 4!\,(n-4)\,\kappa\,(z_1 z_2 z_3 z_4) \qquad (3.2) \\ &= (n-1)\,\kappa_4' + \frac{4!}{3!\,1!}\,(n-2)\,\{\kappa'_{31} + \kappa'_{13}\} + \frac{4!}{2!\,2!}\,\kappa'_{22} \\ &+ \frac{4!}{2!\,1!\,1!}\,(n-3)\,\{\kappa'_{211} + \kappa'_{121} + \kappa'_{112}\} \\ &+ 4!\,(n-4)\,\kappa'_{1111}, \qquad (3.3)\end{aligned}$$

where κ' stands for the appropriate cumulants indicated by the suffixes.

Similarly it can be shown that

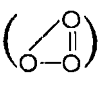
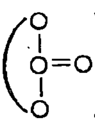
$$\kappa_r = r!\, \Sigma\, (n-t)\, \Sigma\, \frac{\kappa'_{s_1 s_2 \ldots s_t}}{s_1!\, s_2' \ldots s_t!} \tag{3.4}$$

The above result can be extended for the distributions arising from a lattice of points. Representing in terms of Moran's figures (1948)

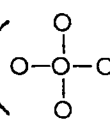
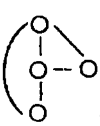
$$\kappa_4 = \alpha_1 \kappa'_4 (\text{O}\equiv\text{O}) + \frac{4!}{3!\,1!}\alpha_2 \{\kappa'_{31} (\text{O}\equiv\text{O}-\text{O}) + \kappa'_{13} (\text{O}-\text{O}\equiv\text{O})\}$$

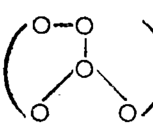
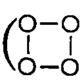
$$+ \frac{4!}{2!\,2!}\alpha_3 \kappa'_{22} (\text{O}=\text{O}=\text{O}) + \frac{4!}{2!\,1!\,1!}\alpha_4 \{\kappa'_{211} (\text{O}=\text{O}-\text{O}-\text{O})$$

$$+ \kappa'_{121} (\text{O}-\text{O}=\text{O}-\text{O}) + \kappa'_{112} (\text{O}-\text{O}-\text{O}=\text{O})\}$$

$+ \frac{4!}{2!\,1!\,1!} 3\alpha_5 \kappa'_{211}$ () $+ \frac{4!}{2!\,1!\,1!} 3\alpha_6 \kappa'_{211}$ ()

$$+ 4!\alpha_7 \kappa'_{1111} (\text{O}-\text{O}-\text{O}-\text{O}-\text{O})$$

$+ 4!\alpha_8 \kappa'_{1111}$ () $+ 4!\alpha_9 \kappa'_{1111}$ ()

$+ 4!\alpha_{10} \kappa'_{1111}$ () $+ 4!\alpha_{11} \kappa'_{1111}$ () (3.5)

where $\alpha_1, \alpha_2, \ldots \alpha_{11}$ are the number of ways of obtaining the given or similar configurations on the lattice.

4. Applications

Using the method developed in Section 3 it has been shown that the fourth cumulants for the distribution of black-black, black-white (including white-black also) and the total number of joins between points of varying characters can be obtained comparatively easily for points lying on a line and on a lattice.

(i) *Fourth cumulants of distributions for points on a line.*—For this we have to evaluate κ'_4, κ'_{31}, κ'_{13}, κ'_{22}, κ'_{211}, κ'_{121}, κ'_{112} and κ'_{1111}.

$$\kappa_4' = E(x_1 - a_2)^4 - 3\{E(x_1 - a_2)^2\}^2$$

$$= a_2 - 7a_2^2 + 12a_2^3 - 6a_2^4, \tag{4.1}$$

where a_2 is the probability for a black-black, black-white or a join between points of varying characters.

$$\kappa'_{13} = \kappa'_{31} = E(x_1 - a_2)^3 (x_2 - a_2) - 3\{E(x_1 - a_2)^2\}\{E(x_1 - a_2)(x_2 - a_2)\}$$

$$= a_3 - 6a_3a_2 + 6a_3a_2^2 - a_2^2 + 6a_2^3 - 6a_2^4 \tag{4.2}$$

where a_3 is the probability for two joins from three adjacent points.

$$\kappa'_{22} = E(x_1 - a_2)^2 (x_2 - a_2)^2 - \{E(x_1 - a_2)^2\}\{E(x_2 - a_2)^2\} - 2\{E(x_1 - a_2)(x_2 - a_2)\}^2$$

$$= a_3 - a_2^2 - 4a_2a_3 - 2a_3^2 + 4a_2^3 + 8a_2^2a_3 - 6a_2^4 \tag{4.3}$$

$$\kappa'_{112} = \kappa'_{211} = E(x_1 - a_2)^2 (x_2 - a_2)(x_3 - a_2) - \{E(x_1 - a_2)^2\}\{E(x_1 - a_2)(x_2 - a_2)\}$$

$$= a_4 - 2a_2a_3 - 2a_2a_4 + 4a_2^2a_3 + a_2^3 - 2a_2^4 \tag{4.4}$$

where a_4 represents the probability for three joins from four points.

$$\kappa'_{121} = E(x_1 - a_2)(x_2 - a_2)^2 (x_3 - a_2) - 2\{E(x_1 - a_2)(x_2 - a_2)\}^2$$

$$= a_4 - 2a_2a_3 + a_2^3 - 2a_2a_4 - 2a_3^2 + 8a_2^2a_3 - 4a_2^4 \tag{4.5}$$

$$\kappa'_{1111} = E(x_1 - a_2)(x_2 - a_2)(x_3 - a_2)(x_4 - a_2) - \{E(x_1 - a_2)(x_2 - a_2)\}^2$$

$$= a_5 - 2a_2a_4 - a_3^2 + 3a_2^2a_3 - a_2^4 \tag{4.6}$$

where a_5 is the probability for four joins from five points.

The values of a_2, a_3, a_4 and a_5 for black-black, black-white (including white-black) and for any kind of join is shown in Table I below:

Using (3.3) κ_4 can now be written easily. For black-black joins it reduces to

$$\begin{aligned}\kappa_4 = {} & (n-1)(p^2 - 7p^4 + 12p^6 - 6p^8) + 8(n-2)(p^3 - p^4 - 6p^5 \\ & + 6p^6 + 6p^7 - 6p^8) + 6(n-2)(p^3 - p^4 - 4p^5 + 2p^6 \\ & + 8p^7 - 6p^8) + 24(n-3)(p^4 - 2p^5 - p^6 + 4p^7 - 2p^8) \\ & + 12(n-3)(p^4 - 2p^5 - 3p^6 + 8p^7 - 4p^8) \\ & + 24(n-4)(p^5 - 3p^6 + 3p^7 - p^8)\end{aligned} \tag{4.7}$$

TABLE I

Values of a's *for different distributions*

Probabilities of	B-B joins	B-W and W-B joins	Any join between points of different characters
One join from two points a_2 ..	p^2	$2p_1p_2$	$2\Sigma p_rp_s$
Two joins from three points a_3 ..	p^3	$p_1p_2(p_1+p_2)$	$\Sigma(p_r^2p_s+p_rp_s^2)+6\Sigma p_rp_sp_t$
Three joins from four points a_4 ..	p^4	$2p_1^2p_2^2$	$2\Sigma p_r^2p_s^2+6\Sigma(p_r^2p_sp_t+p_rp_s^2p_t+p_rp_sp_t^2)$ $+24\Sigma p_rp_sp_tp_u$
Four joins from five points a_1 ..	p^5	$p_1^2p_2^2(p_1+p_2)$	$\Sigma p_r^2p_s^2(p_r+p_s)+2\Sigma p_rp_sp_t(p_r^2+p_s^2+p_t^2)$ $+36\Sigma p_rp_sp_tp_u(p_r+p_s+p_t+p_u)$ $+12\Sigma p_rp_sp_t(p_rp_s+p_sp_t+p_tp_r)$ $+120\Sigma p_rp_sp_tp_up_v$

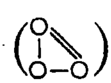
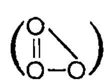
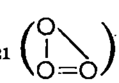
Similarly for white-black and black-white joins

$$
\begin{aligned}
\kappa_4 = {} & (n-1)\{2p_1p_2 - 28p_1^2p_2^2 + 96p_1^3p_2^3 - 96p_1^4p_2^4\} \\
& + 8(n-2)\{p_1p_2(p_1+p_2) - 12p_1^2p_2^2(p_1+p_2) \\
& + 24p_1^3p_2^3(p_1+p_2) - 4p_1^2p_2^2 + 48p_1^3p_2^3 - 96p_1^4p_2^4\} \\
& + 6(n-2)\{p_1p_2(p_1+p_2) - 4p_1^2p_2^2 - 8p_1^2p_2^2(p_1+p_2) \\
& - 2p_1^2p_2^2(p_1+p_2)^2 + 32p_1^3p_2^3 + 64p_1^4p_2^4(p_1+p_2) \\
& - 96p_1^4p_2^4\} + 24(n-3)\{2p_1^2p_2^2 - 4p_1^2p_2^2(p_1+p_2) \\
& - 8p_1^3p_2^3 + 16p_1^3p_2^3(p_1+p_2) + 8p_1^3p_2^3 - 32p_1^4p_2^4\} \\
& + 12(n-3)\{2p_1^2p_2^2 - 4p_1^2p_2^2(p_1+p_2) + 8p_1^3p_2^3 \\
& - 8p_1^3p_2^3 - 2p_1^2p_2^2(p_1+p_2)^2 + 32p_1^3p_2^3(p_1+p_2) \\
& - 64p_1^4p_2^4\} + 24(n-4)\{p_1^2p_2^2(p_1+p_2) - 8p_1^3p_2^3 \\
& - p_1^2p_2^2(p_1+p_2)^2 + 12p_1^3p_2^3(p_1+p_2) - 16p_1^4p_2^4\} \qquad (4.8)
\end{aligned}
$$

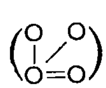
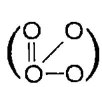
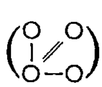
κ_4 for the total number of joins between points of different characters can be written in a similar manner.

(ii) *Fourth cumulants of distributions for points on a lattice.*—The probabilities for one, two, three and four joins from two, three, four and five points for the various possible configurations are shown in Table II.

Now to obtain the fourth cumulants of the distributions we have to evaluate the κ''s for the different configurations shown in Table II. κ'_4, κ_{31}', κ_{13}' and κ_{22}' are the same as given in (4.1), (4.2) and (4.3).

κ_{112}'  $= \kappa'_{121}$  $= \kappa'_{121}$  $=$

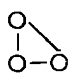
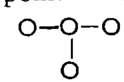
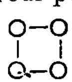
$$
\begin{aligned}
& E(x_1 - a_2)^2(x_2 - a_2)(x_3 - a_2) - 2\mu_{11}^2 - \mu_2\mu_{11} \\
& = A_3 - 5A_3a_2 + 10A_3a_2^2 + 2a_2^3 - 6a_2^4 - 2A_3^2 \qquad (4.9)
\end{aligned}
$$

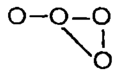
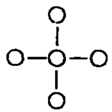
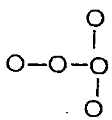
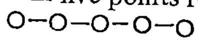
κ'_{211}  $= \kappa'_{112}$  $= \kappa'_{121}$ 

$$
\begin{aligned}
& = A_4 - 3a_2A_3 - 2a_2A_4 + 10a_2^2a_3 + 2a_2^3 \\
& - 2a_3^2 - 6a_2^4 \qquad (4.10)
\end{aligned}
$$

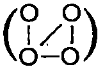
TABLE II

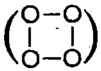
Probabilities for different configurations

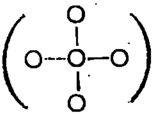
Configurations	Probability for joins of			No. of configurations on the lattice
	B-B	B-W and W-B	Different colours	
One join from two points O—O a_2	p^2	$2p_1p_2$	$2\ (1^2)$	α_1
Two joins from three points O—O—O a_3	p^3	$p_1p_2\,(p_1+p_2)$	$(21)+6\ (1^3)$	$\alpha_2=\alpha_3$
Three joins from three points  A_3	p^3	0	$6\ (1^3)$	α_5
Three joins from four points meeting at a common point  A_4	p^4	$p_1p_2\,(p_1^2+p_2^2)$	$(31)+6\ (21^2)+24\ (1^4)$	α_6
Three joins from four points forming a chain O—O—O—O a_4	p^4	$2p_1^2p_2^2$	$2\ (2^2)+6\ (21^2)+24\ (1^4)$	α_4
Four joins from four points forming a ring  B_4	p^4	$2p_1^2p_2^2$	$2\ (2^2)+4\ (21^2)+24\ (1^4)$	α_{11}

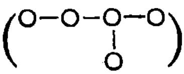
Four joins from four points three of which form a Δ  B_4'	p^4	0	$4\,(21^2) + 24\,(1^5)$	α_9
Four joins from five points meeting at a common point  C_5	p^5	$p_1 p_2\,(p_1^3 + p_2^3)$	$(41) + 8\,(31^2) + 6\,(2^2 1) + 36\,(21^3) + 120\,(1^5)$	α_8
Four joins from five points such that three meet at a point  C_5'	p^5	$p_1^2 p_2^2\,(p_1 + p_2)$	$(32) + 4\,(31^2) + 36\,(21^3) + 10\,(2^2 1) + 120\,(1^5)$	α_{10}
Four joins from five points forming a chain  a_5	p^5	$p_1^2 p_2^2\,(p_1 + p_2)$	$(32) + 2\,(31^2) + 36\,(21^3) + 12\,(2^2 1) + 120\,(1^5)$	α_7

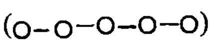
The κ''s arising from three joins forming a chain has already been given in (4.4) and (4.5).

κ'_{1111}  $= B'_4 - 2a_2a_4 - a_2A_4 - a_2A_3 + 9a_3a_2^2 - 4a_2^4 - 2a_3^2$ (4.11)

κ'_{1111}  $= B_4 - 4a_2B_4 + 8a_2^2a_3 - 3a_3^4 - 2a_3^2$ (4.12)

κ'_{1111}  $= B_5 - 4a_2A_4 + 12a_2^2a_3 - 6a_2^4 - 3a_3^2$ (4.13)

κ'_{1111}  $= C_5' - 2a_2a_4 - a_2A_4 + 5a_2^2a_3 - 2a_2^4 - a_3^2$ (4.14)

κ'_{1111}  = same as given in (4.6).

Substituting the values of a, A, B, C, etc., in (3.5) from Table II the κ_4's can be obtained. For black-black joins it reduces to

$$\begin{aligned}\kappa_4 = {} & \alpha_1 (p^2 - 7p^4 + 12p^6 - 6p^8) + 8\alpha_2 (p^3 - p^4 - 6p^5 + 6p^6 \\ & + 6p^7 - 6p^8) + 6\alpha_3 (p^3 - p^4 - 4p^5 + 2p^6 + 8p^7 - 6p^8) \\ & + 24\alpha_4 (p^4 - 2p^5 - p^6 + 4p^7 - 2p^8) + 12\alpha_4 (p^4 - 2p^5 \\ & - 3p^6 + 8p^7 - 4p^8) + 36\alpha_5 (p^3 - 5p^5 + 10p^7 - 6p^8) \\ & + 36\alpha_6 (p^4 - 3p^5 - 2p^6 + 10p^7 - 6p^8) + 24\alpha_7 (p^5 - 3p^6 \\ & + 3p^7 - p^8) + 24\alpha_8 (p^5 - 7p^6 + 12p^7 - 6p^8) + 24\alpha_9 \\ & (p^4 - p^5 - 5p^6 + 9p^7 - 4p^8) + 24\alpha_{10} (p^5 - 4p^6 + 5p^7 \\ & - 2p^8) + 24\alpha_{11} (p^4 - 6p^6 + 8p^7 - 3p^8), \end{aligned} \tag{4.15}$$

where the values of the α's obtained by B. V. Sukhatme (1951) are as follows:

$$\begin{aligned} \alpha_1 &= (4b - 3a + 2),\ \alpha_2 = \alpha_3 = (28b - 26a + 44), \\ \alpha_4 &= (184b - 327a + 554),\ \alpha_5 = (4b - 4a + 4), \\ \alpha_6 &= (56b - 92a + 148),\ \alpha_7 = (1168b - 2614a + 5580) \\ \alpha_8 &= (70b - 130a + 240),\ \alpha_9 = (72b - 108a + 156), \\ \alpha_{10} &= (1032b - 2166a + 4416),\ \alpha_{11} = (12b - 17a + 23) \end{aligned}$$

in which $b = mn$ and $a = m + n$, m and n being the number of points on the sides of the lattice.

Similarly by substituting the values of a, A, B and C given in Table II the 4th cumulant for the distribution of B-w and w-B joins and also for the total number of joins between points of different colours and also for the total number of joins between points of different colours can be written.

5. Summary

A rigorous proof for the author's result for calculating the factorial moments of distributions arising in Markoff chains has been given. A new simple method for obtaining the cumulants of such distributions has also been developed.

References

Krishna Iyer, P. V. .. "The theory of probability distribution of points on a lattice," *Ann. Math. Stat.*, 1950, **21**, 198.

———— .. "Further contributions to the theory of probability distributions of points on a line," *J. Ind. Soc. Agri. Stat.*, 1950 & 51, **2**, **3** & **4**, 141, 80 and 50.

Moran, P. A. P. .. "The interpretation of statistical maps," *J. Roy. Stat. Soc.*, 1948, **B 10**, 243.

Sukhatme, B. V. .. "On certain probability distributions leading to non-parametric tests and their applications," Unpublished Thesis, 1951.

ERRATA

(Vol. IV, No. 1, p. 61)

Eighth row of the table—*Read* " 18 " *in place of* " 17 ".

3rd and 4th lines from the bottom—*Read* " 18 " *in place of* " 17 ".