A THEOREM ON FACTORIAL MOMENTS AND ITS APPLICATIONS

By P. V. Krishna Iyer

Defense Science Laboratory, New Delhi

0. Summary. The theorem that the sth factorial moment for the sum of N events is $s!$ times the sum of the expectations for any s of the events occurring simultaneously has been proved by induction. The applications of this result in obtaining easily the moments of a number of distributions arising from a sequence of observations belonging to two continuous populations and other cases have been demonstrated.

1. Introduction. A number of distributions arising from a sequence of n observations belonging to a binomial population have been considered by the writer [3], [4] in some of his earlier publications. The moments of these distributions were obtained by using the theorem that the sth factorial moment is equal to $s!$ times the expectation for s of the characters considered in the distribution. Thus for a sequence of observations consisting of A's and B's with the probabilities p and q respectively, the sth factorial moment for the distribution of the total number of AB and BA joins between successive observations is the expectation for s joins like AB and BA in the sequence. It can be seen that there are s different ways of obtaining s joins. They are:

(1) From $(s + 1)$ consecutive observations.

(2) From two sets of l_1 and l_2 consecutive observations such that $l_1 + l_2 - 2$ is equal to s.

(3) From three sets of l_1, l_2 and l_3 consecutive observations such that

$$l_1 + l_2 + l_3 - 3$$

is equal to s.

(4) From k sets of $l_1, l_2, \cdots, l_k$ consecutive observations subject to the condition

$$\sum_1^k l_r - k = s,$$

where k takes values 1 to s.

The sum of the expectations for 1, 2, 3, $\cdots$, s is equal to $1/s!$ (the sth factorial moment for the distribution of the total number of AB and BA joins of the sequence).

The theorem as it stands appears to be applicable only for the distributions arising from a binomial sequence consisting of A's and B's with fixed probabilities p and q. We shall show in this paper that this result can be applied for distributions arising from two samples belonging to populations with cumu-

Received December 19, 1956; revised May 29, 1957.

lative distribution functions F and G. Before discussing this aspect, we shall first give a rigorous proof of the theorem and then show how it can be applied for the case of continuous distributions. The use of the result for distributions arising from Markoff chain is also illustrated.

2. Statement of theorem and proof.

THEOREM. *The* sth *factorial moment about the origin of any statistic X which is the sum of N events, dependent or independent, is equal to* $s!$ *times the sum of the expectations for any s of the events occurring together.*

PROOF. Let the events be denoted by $x_1, x_2, \cdots, x_N$. As in the case of binomial distribution, assume that the x's take value 1 if the event occurs and zero otherwise. Define

$$X = \sum_1^N x_r,$$

$$E(X) = E(\sum x_r) = \sum E(x_r) \tag{1}$$

= the sum of the expectations of the different events

= the sum of the probabilities for the events to occur.

$$E(X^2) = E(\sum x_r)^2 = E(\sum x_r^2) + 2E(\sum x_r x_s), \qquad s > r.$$

Now

$$E(\sum x_r^2) = E(X);$$

hence

$$E(X^2) = E(X) + 2E(\sum x_r x_s)$$

or

$$E\{X(X-1)\} = 2 \sum E(x_r x_s) \tag{2}$$

= 2 (sum of the expectations for any two of the events)

= 2 (sum of the probabilities for any two of the events to occur together).

$$E(X^3) = E(\sum x_r)^3 = E(\sum x_r^3) + 3E\sum (x_r^2 x_s) + 3E(\sum x_r x_s^2) + 6E(\sum x_r x_s x_t), \qquad t > s > r,$$

since

$$E(\sum x_r^3) = E(X),$$

$$E(x_r^2 x_s) = E(x_r x_s),$$

$$E(x_r x_s^2) = E(x_r x_s),$$

$$E(X^3) = E(X) + 6E(\sum x_r x_s) + 3!E(\sum x_r x_s x_t).$$

Substituting the value of $E(\sum x_r x_s)$ from (2), we get

$$E(X^3) = E(X) + 3E\{X(X-1)\} + 3!E(\sum x_r x_s x_t)$$

or

$$(3) \quad E\{X(X-1)(X-2)\} = 3!\sum E(x_r x_s x_t)$$

$= 3!$(sum of the expectations for any three of the events)

$= 3!$(sum of the probabilities for any three of the events to occur together).

Thus the theorem holds good for $s = 1$ to 3.

It may be noted that the results given above hold good even without taking the expectation of both sides because the x's take values 1 or 0 only.

We shall now establish the general relation by induction. For this we show that if

$$(4) \qquad X^{[s]} = s!\,(\sum x_{t_1} x_{t_2} \cdots x_{t_s})$$

holds good for any value of s, it is true for $(s + 1)$ also.

Multiplying both sides of (4) by X we get

$$\begin{aligned}[X^{[s]}X] &= s!\,(\sum x_{t_1} x_{t_2} \cdots x_{t_s})(\sum x_r) \\ &= (s+1)!\,(\sum x_{t_1} x_{t_2} \cdots x_{t_s} x_{t_{s+1}}) \\ &\quad + s!\,[\sum x_{t_1}^2 x_{t_2} \cdots x_{t_s} \\ &\quad + \sum x_{t_1} x_{t_2}^2 \cdots x_{t_s} + \cdots + \sum x_{t_1} x_{t_2} \cdots x_{t_s}^2] \\ &= (s+1)!\sum x_{t_1} x_{t_2} \cdots x_{t_{s+1}}) \\ &\quad + s!\,(\sum x_{t_1} x_{t_2} \cdots x_{t_s}).\end{aligned}$$

Substituting for $\sum x_{t_1} x_{t_2} \cdots x_{t_s}$ from (4), (5) reduces to

$$(6) \qquad X^{[s]}(X - s) = X^{[s+1]} = (s+1)!\,(\sum x_{t_1} x_{t_2} \cdots x_{t_{s+1}}).$$

Taking the expectation of both sides

$$(7) \qquad E\{X^{[s+1]}\} = (s+1)!\sum E(x_{t_1} x_{t_2} \cdots x_{t_{s+1}}).$$

Hence the theorem.

3. Applications. We shall now examine how the above result can be applied for obtaining easily the moments of a number of distributions including those arising from a simple Markoff chain. Some of the distributions considered here have been discussed by Wald and Wolfowitz [7], Mood [6], Mann & Whitney [5], and others.

(1) *Binomial distribution.* It is obvious that the rth factorial moment for the distribution of x, the number of successes out of n trials is given by

$$(8) \qquad \frac{\mu'_{[r]}}{r!} = \binom{n}{r} p^r,$$

where p is the probability for a success.

(2) *Hypergeometric distribution.* This can be deduced from the above by substituting

$$p^r = \frac{(N - M)^{[r]}}{N^{[r]}},$$

where N and M have the usual significance. This follows from the fact that the probability p for the 1*st*, 2*nd*, 3*rd*, $\cdots$ successes are

$$\frac{N - M}{N}, \quad \frac{N - M - 1}{N - 1}, \quad \frac{N - M - 2}{N - 2}, \cdots$$

(3) *Distribution of the number of AB joins between successive observations of a binomial sequence.* We first note that r AB joins can be formed from only r sets of two consecutive observations each and therefore

$$(9) \qquad \frac{\mu'_{[r]}}{r!} = \binom{n - r}{r} p^r q^r.$$

This can be seen from the fact that the probabilities for an AB join is pq and that there are $\binom{n - r}{r}$ ways of obtaining them from n observations in a sequence.

(4) *Distribution of AB joins for binomial sequence of n_1A's and n_2B's.* As in the case of hypergeometric series, we substitute

$$p^r q^s = \frac{n_1^{[r]} n_2^{[s]}}{(n_1 + n_2)^{[r+s]}}$$

in the results given in (3) above. Thus

$$(10) \qquad \mu'_{[r]} = \frac{(n_1 + n_2 - r)^{[r]} n_1^{[r]} n_2^{[r]}}{(n_1 + n_2)^{[2r]}}.$$

(5) *Distribution of AB and BA joins between consecutive observations of a binomial sequence.* Taking for simplicity the third factorial moment, we note that three joins can be obtained from (i) four consecutive observations $ABAB$ or $BABA$, (ii) two sets, one of two and the other of three consecutive observations like $AB - ABA$; $BA - ABA$; $AB - BAB$ and $BA - BAB$, (iii) three sets of each of two consecutive observations AB or BA. The sum of the expectations for the above three ways of obtaining three joins is

$$(11) \qquad \frac{\mu'_{[3]}}{3!} = 2(n - 3)p^2q^2 + 8\binom{n - 3}{2} p^2q^2 + \binom{n - 3}{3} 8p^3q^3.$$

(6) *A sequence formed by pooling two samples A and B belonging to F.* Let two samples A and B of sizes n_1 and n_2 be drawn from a population where cumulative

distribution function is $F(x)$, $F(x)$ being continuous and x taking values from $-\infty$ to $+\infty$. By pooling together A and B and arranging them in ascending or descending order we obtain a sequence of A's and B's as in (4) considered above. Hence the moments of any distribution arising from this sequence can be obtained from the corresponding ones for the binomial sequence by substituting

$$p^r q^s = \frac{n_1^{[r]} n_2^{[s]}}{(n_1 + n_2)^{[r+s]}}.$$

(7) *Same as* (6), $F \neq G$. The calculation of the moments for some of the distributions discussed above is more complicated when $F \neq G$. We shall show below how the present theorem enables us to obtain the moments of these distributions also.

(a) *Number of observations of sample A to the left of the* rth *value of the combined ordered sequence of A*'s *and B*'s.

$\frac{\mu'_{[s]}}{s!}$ = [Number of ways of selecting A's from $(r - 1)$ observations]
$\times$ [Probability that s out of the $(r - 1)$ values to the left of the rth observation belong to A]

Assuming $n_1F(\alpha) + n_2G(\alpha) = r$, the probability that amongst the $(r - 1)$ values to the left of the rth observation there are s A's is

$$\frac{n_1F(\alpha)}{n_1 F(\alpha) + n_2 G(\alpha)} \frac{(n_1 - 1)F(\alpha)}{(n_1 - 1)F(\alpha) + n_2 G(\alpha)} \frac{(n_1 - 2)F(\alpha)}{(n_1 - 2)F(\alpha) + n_2 G(\alpha)} \cdots s \text{ terms.}$$

Using the relation between $F(\alpha)$, $G(\alpha)$ and r we get

$$\text{(12)} \qquad \frac{\mu'_{[s]}}{s!} = \binom{r-1}{s} \frac{n_1^{[s]}[F(\alpha)]^s}{r[r - F(\alpha)][r - 2F(\alpha)] \cdots [r - (s - 1)F(\alpha)]}.$$

(b) *Number of AB and BA joins between successive observations.* As the higher moments are complicated we shall be content to obtain the second moment.

$\frac{\mu'_{[2]}}{2!}$ = the sum of the expectations for two joins from (i) three consecutive observations and (ii) two sets each of two consecutive observations.

Expectation for two joins from three consecutive observations x_1, x_2, and x_3 is given by

$$\text{(13)} \qquad \begin{aligned} &n_1(n_1 - 1)n_2 \int_{-\infty}^{+\infty} \int_{-\infty}^{x_3} \int_{-\infty}^{x_2} \{1 - F(x_3) + F(x_1)\}^{n_1-2} \\ &\qquad \cdot \{1 - G(x_3) + G(x_1)\}^{n_2-1} \, dF(x_1) \, dG(x_2) \, dF(x_3) \\ &\qquad + n_1 n_2(n_2 - 1) \int_{-\infty}^{+\infty} \int_{-\infty}^{x_3} \int_{-\infty}^{x_2} \{1 - G(x_3) + G(x_1)\}^{n_2-2} \\ &\qquad \cdot \{1 - F(x_3) + F(x_1)\}^{n_1-1} \, dG(x_1) \, dF(x_2) \, dG(x_3), \qquad x_1 < x_2 < x_3. \end{aligned}$$

Expectation for four joins from two sets of two consecutive observations each is equal to

$$
\begin{aligned}
&n_1^{[2]}n_2^{[2]}\Bigg[\int_{-\infty}^{+\infty}\int_{-\infty}^{x_4}\int_{-\infty}^{x_3}\int_{-\infty}^{x_2} A\,dF(x_1)\,dG(x_2)\,dF(x_3)\,dG(x_4) \\
&\qquad + \int_{-\infty}^{+\infty}\int_{-\infty}^{x_4}\int_{-\infty}^{x_3}\int_{-\infty}^{x_2} A\,dF(x_1)\,dG(x_2)\,dG(x_3)\,dF(x_4) \\
(14)\qquad &\qquad + \int_{-\infty}^{+\infty}\int_{-\infty}^{x_4}\int_{-\infty}^{x_3}\int_{-\infty}^{x_2} A\,dG(x_1)\,dF(x_2)\,dF(x_3)\,dG(x_4) \\
&\qquad + \int_{-\infty}^{+\infty}\int_{-\infty}^{x_4}\int_{-\infty}^{x_3}\int_{-\infty}^{x_2} A\,dG(x_1)\,dF(x_2)\,dG(x_3)\,dF(x_4),
\end{aligned}
$$

where

$$
\begin{aligned}
A = \{1 - F(x_2) + F(x_1) - F(x_4) + F(x_3)\}^{n_1-2} \\
\times \{1 - G(x_2) + G(x_1) - G(x_4) + G(x_3)\}^{n_2-2}, \\
x_1 < x_2 < x_3 < x_4.
\end{aligned}
$$

From the above it follows that

$$
(15)\qquad \frac{\mu'_{[2]}}{2!} = (13) + (14).
$$

When $F = G$, this reduces to the expression known.

(8) *Mann and Whitney's T-statistic.* In this case the expression for the second factorial moment reduces to the simple form

$$
\begin{aligned}
\mu'_{[2]} &= 2n_1n_2\int_{-\infty}^{+\infty}\int_{-\infty}^{x_3}\int_{-\infty}^{x_2}[(n_1-1)f(x_1)f(x_2)g(x_3) + (n_2-1)f(x_1)g(x_2)g(x_3)] \\
(16)\qquad &\qquad \cdot\, dx_1\,dx_2\,dx_3 + n_1^{(2)}n_2^{(2)}\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{x_2} f(x_1)g(x_2)\,dx_1\,dx_2\right]^2,
\end{aligned}
$$

where $f(x)$ and $g(x)$ are the density functions for F and G.

(9) *AB joins between successive observations for a simple Markoff chain.* Let the matrix of probabilities for a simple Markoff chain be

$$
\begin{bmatrix} p_1 & p_2 \\ q_1 & q_2 \end{bmatrix}.
$$

Taking the probability that the first observation is A or B as P and Q respectively, the probabilities $P_r(A)$ and $Q_r(B)$ that the rth observation is A or B are given by

$$
\begin{aligned}
(17)\qquad P_r(A) &= \frac{p_2}{1-\delta} + \frac{Pq_1 - Qp_2}{1-\delta}\,\delta^{r-1}, \\
Q_r(B) &= 1 - P_r(A),
\end{aligned}
$$

where

$$\delta = p_1 - p_2 \quad \text{and} \quad p_1 > p_2 .$$

When the first observation is B, the conditional probability for the rth observation to be A reduces to

$$P_r(A \mid 1B) = \frac{p_2}{1 - \delta}(1 - \delta^{r-1}). \tag{18}$$

This is the same as given by Bartlett:

In the case of the Markoff chain, unlike the previous cases discussed earlier, the probability of an AB join depends on the position of A in the sequence, and the expectation for two AB joins is given by

$$\begin{aligned} &q_1^2[P_1(A)\{P_3(A \mid 2B) + P_4(A \mid 2B) + P_5(A \mid 2B) + \cdots + P_{n-1}(A \mid 2B)\} \\ &+ P_2(A)\{P_4(A \mid 3B) + P_5(A \mid 3B) + P_6(A \mid 3B) + \cdots + P_{n-1}(A \mid 3B)\} \\ &+ P_3(A)\{P_5(A \mid 4B) + P_6(A \mid 4B) + \cdots + P_{n-1}(A \mid 4B)\} \\ &\qquad \cdots \\ &+ P_{n-3}(A)\{P_{n-1}(A \mid n - 2, B)\}], \end{aligned} \tag{19}$$

where $P_r(A \mid B)$ is the conditional probability that the rth observation is A, given that the sth observation is B when $r > s$. Summing up the above series after substituting for P's from (16) and (17), we get

$$\begin{aligned} \frac{\mu'_{[2]}}{2!} = \Bigg[&\frac{(n-2)(n-3)\alpha}{2} - \frac{\alpha\delta}{1-\delta}\left\{(n-3) - \frac{\delta(1-\delta^{n-3})}{1-\delta}\right\} \\ &- \frac{\beta\delta}{1-\delta}\left\{\frac{1-\delta^{n-3}}{1-\delta} - (n-3)\delta^{n-3}\right\} \\ &+ \beta\left\{\frac{n-3}{1-\delta} - \frac{\delta(1-\delta^{n-3})}{(1-\delta)^2}\right\}\Bigg]\frac{p_2 q_1^2}{1-\delta}, \end{aligned} \tag{20}$$

where

$$\alpha = \frac{p_2}{1-\delta} \quad \text{and} \quad \beta = \frac{Pq_1 - Qp_2}{1-\delta}.$$

It may be added that the result given in this paper can be used for deriving the moments of many other distributions of similar kind.

REFERENCES

1. M. S. Bartlett, *An Introduction to Stochastic Processes*, Cambridge University Press, London, 1955, p. 29.
2. W. J. Dixon, "A criterion for testing the hypothesis that two samples are from the same population", *Ann. Math. Stat.*, Vol. 11 (1940), p. 199.
3. P. V. Krishna Iyer, "The theory of probability distributions of points on a line", *J. Indian. Soc. Agri. Stat.*, Vol. 1 (1948), p. 173.

4. P. V. Krishna Iyer, "The first and second moments of some probability distribution arising from points on a lattice and their applications", *Biometrika*, Vol. 36 (1949), p. 135.
5. H. B. Mann and D. R. Whitney, "On a test whether one of two random variables is stochastically larger than the other," *Ann. Math. Stat.*, Vol. 18 (1947), p. 50.
6. A. M. Mood, "On the asymptotic efficiency of certain nonparametric two sample tests", *Ann. Math. Stat.*, Vol. 25 (1954), p. 514.
7. A. Wald and J. Wolfowitz, "On a test whether two samples are from the same population," *Ann. Math. Stat.*, Vol. 11 (1940), p. 147.