MOMENT RECURRENCE RELATIONS FOR BINOMIAL, POISSON AND HYPERGEOMETRIC FREQUENCY DISTRIBUTIONS[1]

By John Riordan

1. **Introduction.** This paper gives the development of recurrence relations for moments about the origin and mean of binomial, Poisson, and hypergeometric frequency distributions from the basis of the moment arrays defined by H. E. Soper.[2] This procedure has the advantage of expressing the moments in terms of coefficients which are alike for the three distributions and are derivable by a single process, thus providing a degree of formal coordination of the distributions. For both kinds of moments, the coefficients satisfy relatively simple recurrence relations, the use of which leads to recurrence relations for the moments, thus unifying the derivation of these relations for the three distributions. The relations derived in this way for the hypergeometric distribution are apparently new. Apparently new recurrence relations for certain auxiliary coefficients in the expression of the moments about the mean of binomial and Poisson distributions are also given.

This course of development involves repetition of a number of well-known results which is justified, it is hoped, by the unification obtained.[3]

---

[1] Presented to the American Mathematical Society, Sept. 3, 1936.

[2] *Frequency Arrays*, Cambridge, 1922.

[3] The following bibliography is taken from a paper *On the Bernoulli Distribution*, Solomon Kullback, Bull. Am. Math. Soc., **41**, 12, pp. 857–864, (Dec., 1935):

A. Fisher, *The Mathematical Theory of Probabilities*, 2d ed., p. 104 ff.

H. L. Rietz, *Mathematical Statistics*, 1927, p. 26 ff.

V. Mises, *Wahrscheinlichkeitsrechnung*, 1931, pp. 131–133.

Risser and Traynard, *Les Principes de la Statistique Mathématique*, 1933, pp. 39–40 and 320–321.

V. Romanovsky, *Note on the moments of the binomial* $(q + p)^n$ *about its mean*, Biometrika, vol. 15 (1923), pp. 410–412.

A. T. Craig, *Note on the moments of a Bernoulli distribution*, Bull. Am. Math. Soc., vol. 40 (1934), pp. 262–264.

A. R. Crathorne, *Moments de la binomiale par rapport à L'origine*, Comptes Rendus, vol. 198 (1934), p. 1202;

A. A. K. Aygangar, *Note on the recurrence formulae for the moments of the point binomial*, Biometrika, vol. 26 (1934), pp. 262–264.

To this, besides Soper's tract already mentioned, should be added:

Ch. Jordan, *Statistique Mathematique*, Paris, 1927.

K. Pearson, *On Certain Properties of the Hypergeometric Series . . .*, Phil. Mag., **47**, pp. 236–246 (1899).

2. **Moment Arrays.** As developed by Soper, frequency distributions may be exhibited by frequency arrays, in the case of a single variate, in the form:

$$f(A) = \sum_x p_x A^x \tag{2.1}$$

where p_x are the frequencies with which the measures, x, of the character, A, occur in a population.

The substitution $A = e^\alpha$ leads to the moment about the origin array:

$$\begin{aligned} f(e^\alpha) &= \sum_x p_x e^{x\alpha} \\ &= \sum_x p_x \left(1 + x\alpha + \frac{x^2\alpha^2}{2!} + \cdots\right) \\ &= \sum_s m_s \frac{\alpha^s}{s!} \end{aligned} \tag{2.2}$$

where

$$m_s = \sum_x p_x x^s$$

The symbol α is a logical or umbral symbol serving merely to identify the moments in the expansion of the array.

The moment array for moments about the mean is found from the relation:

$$\begin{aligned} \phi(e^\alpha) &= e^{-m_1\alpha} f(e^\alpha) \\ &= \sum_s \mu_s \alpha^s/s! \end{aligned}$$

where m_1 is the first moment about the origin.

The moment arrays for the distributions concerned are as follows:

Binomial $\quad f(e^\alpha) = [1 + p(e^\alpha - 1)]^n = \sum_{x=0}^{n} \binom{n}{x} p^x(e^\alpha - 1)^x$

Poisson $\quad f(e^\alpha) = e^{a(e^\alpha - 1)} = \sum_{x=0}^{\infty} \frac{a^x(e^\alpha - 1)^x}{x!}$

Hypergeometric $\quad f(e^\alpha) = \sum_{x=0}^{\infty} \frac{(l)_x(r)_x}{(n)_x} \frac{(e^\alpha - 1)^x}{x!}$

where the parameters p, n, and a for the binomial and Poisson have the usual significance. The parameters for the hypergeometric distribution, with the substitution $r = s$, follow Soper; Pearson (loc. cit.) uses q, r, n, where $q = l/n$. The notation $(l)_x$ means

$$(l)_x = l(l - 1) \cdots (l - x + 1).$$

It will be seen that, with the usual interpretation of $\binom{n}{x}$ as zero for $x > n$, the three distributions so far as concerns α may be exhibited by a function of the form

$$f(e^\alpha) = \sum_{x=0}^{\infty} A_x(e^\alpha - 1)^x$$

where A_x of course depends on the distribution concerned.

3. **Moments About the Origin.** The moments about the origin can then be defined by the equation:

$$\sum_{s=0}^{\infty} m_s \frac{\alpha^s}{s!} = \sum_{x=0}^{\infty} A_x(e^\alpha - 1)^x \tag{3.1}$$

and

$$\sum_{x=0}^{\infty} A^x(e^\alpha - 1)^x = \sum_{x=0}^{\infty} A_x \sum_{v=0}^{x} (-1)^{x-v} \binom{x}{v} e^{v\alpha}$$
$$= \sum_{s=0}^{\infty} \frac{\alpha^s}{s!} \sum_{x=0}^{s} x!\, A_x S_{x,s}$$

where $S_{x,s}$ is a Stirling number of the second kind, as used by Jordan (loc. cit.) and defined by

$$x!\, S_{x,s} = \sum_{v=0}^{x} (-1)^{x-v} \binom{x}{v} v^s = \Delta^x 0^s,$$

$\Delta^x 0^s$ being in the language of the finite difference calculus, a "difference of nothing" that is $\Delta^x n^s \mid n = 0$.

The internal series terminates at s because $S_{x,s} = 0,\ x > s$, as is readily apparent in the finite difference expression. Further $S_{0,s} = 0,\ s \neq 0;\ S_{0,0} = 1$.

By equating coefficients in equation (3.1), m_s, the sth moment about the origin, is given by

$$m_s = \sum_{x=0}^{s} x!\, A_x\, S_{x,s}. \tag{3.2}$$

The particular forms for the three distributions are as follows:

$$m_s = \sum_{x=0}^{s} (n)_x\, p^x\, S_{x,s} \qquad \textit{Binomial} \tag{3.3}$$

$$m_s = \sum_{x=0}^{s} a^x\, S_{x,s} \qquad \textit{Poisson} \tag{3.4}$$

$$m_s = \sum_{x=0}^{s} \frac{(l)_x (r)_x}{(n)_x} S_{x,s} \qquad \textit{Hypergeometric} \tag{3.5}$$

The Stirling numbers have the following recurrence relation (Jordan loc. cit.):

$$S_{x,s+1} = xS_{x,s} + S_{x-1,s}. \tag{3.6}$$

This relation in conjunction with equations (3.3)–(3.5) leads to moment recurrence relations. The procedure is illustrated for the binomial distribution as follows:

$$\begin{aligned} m_{s+1} &= \sum_{x=0}^{s+1} (n)_x\, p^x\, S_{x,\, s+1} \\ &= \sum_{x=0}^{s+1} (n)_x\, p^x\, (x\, S_{x,\, s} + S_{x-1,\, s}) \\ &= p\, D_p\, m_s + (npm_s - p^2\, D_p\, m_s) \\ &= npm_s + pq\, D_p\, m_s \end{aligned}$$

where $q = 1 - p$.

The steps in the process are expanded as follows:

$$\begin{aligned} \sum_{x=0}^{s+1} (n)_x p^x\, x\, S_{x,\, s} &= \sum_{x=0}^{s} (n)_x\, p^x\, x\, S_{x,\, s} \\ &= \sum_{x=0}^{s} (n)_x\, S_{x,\, s}\, pD_p(p^x) \\ &= pD_p\, m_s \end{aligned}$$

$$\begin{aligned} \sum_{x=0}^{s+1} (n)_x\, p^x\, S_{x-1,\, s} &= \sum_{x=0}^{s+1} (n - x + 1)\, (n)_{x-1}\, p^x\, S_{x-1,\, s} \\ &= n \sum_{x=1}^{s} (n)_x\, p^{x+1}\, S_{x,\, s} - \sum_{x=1}^{s} x(n)_x\, p^{x+1}\, S_{x,\, s} \\ &= np\, m_s - p^2 D_p\, m_s \end{aligned}$$

The results for the three distributions are as follows:

(3.7) $$m_{s+1} = npm_s + pqD_pm_s$$ *Binomial*

(3.8) $$m_{s+1} = am_s + aD_am_s$$ *Poisson*

(3.9) $$m_{s+1} = \frac{lr}{n}\, m_s(l - 1, r - 1, n - 1) - (n + 1)\Delta_n m_s$$ *Hypergeometric*

Here D_p and D_a denote differentiation with respect to p and a, respectively, and Δ_n denotes the difference operation with respect to n. For the hypergeometric distribution the moments are functions of l, r, and n as well as of s; $m_s(l - 1, r - 1, n - 1)$ is the same function of $l - 1$, $r - 1$ and $n - 1$ as $m_s(l, r, n)$ is of l, r, n. Equation (3.9) appears to be new.

For convenience of reference, a short table of the Stirling numbers of the second kind follows:

$s \backslash x$	$S_{x,s}$ 0	1	2	3	4	5
0	1					
1	0	1				
2	0	1	1			
3	0	1	3	1		
4	0	1	7	6	1	
5	0	1	15	25	10	1

4. **Moments About the Mean.** As shown in Section 2 above, moments about the mean may be defined as follows:

$$\sum_{s=0}^{\infty} \mu_s \frac{\alpha^s}{s!} = \sum_{x=0}^{\infty} A_x e^{-m_1\alpha} (e^\alpha - 1)^x \tag{4.1}$$

where m_1 is the first moment about the origin:

$$\begin{aligned} m_1 &= np \quad \textit{Binomial} \\ &= a \quad \textit{Poisson} \\ &= lr/n \ \textit{Hypergeometric} \end{aligned}$$

Now

$$\begin{aligned} \sum_{x=0}^{\infty} A_x e^{-m_1\alpha} (e^\alpha - 1)^x &= \sum_{x=0}^{\infty} A_x \sum_{v=0}^{x} (-1)^{x-v} \binom{x}{v} e^{(v-m_1)\alpha} \\ &= \sum_{s=0}^{\infty} \frac{\alpha^s}{s!} \sum_{x=0}^{s} x!\, A_x\, \sigma_{x,s} \end{aligned}$$

where

$$x!\,\sigma_{x,s} = \sum_{v=0}^{x} (-1)^{x-v} \binom{x}{v} (v - m_1)^s = \Delta^x (-m_1)^s.$$

It will be observed that for $m_1 = 0$, $\sigma_{x,s} = S_{x,s}$. The internal series terminates at s for the same reason as before.

The moments about the mean are then given by:

$$\mu_s = \sum_{x=0}^{s} x!\, A_x\, \sigma_{x,s} \tag{4.2}$$

The particular forms for the three distributions are as follows:

$$\mu_s = \sum_{x=0}^{s} (n)_x\, p_x\, \sigma_{x,s} \qquad \textit{Binomial} \tag{4.3}$$

$$\mu_s = \sum_{x=0}^{s} a^x\, \sigma_{x,s} \qquad \textit{Poisson} \tag{4.4}$$

$$\mu_s = \sum_{x=0}^{s} \frac{(l)_x (r)_x}{(n)_x}\, \sigma_{x,s} \qquad \textit{Hypergeometric.} \tag{4.5}$$

The coefficients $\sigma_{x,s}$ satisfy the following recurrence relation:[4]

(4.6) $$\sigma_{x,s+1} = (x - m_1)\sigma_{x,s} + \sigma_{x-1,s}$$

which in conjunction with equations (4.3)–(4.5) leads to moment recurrence relations as before. The actual derivation is somewhat complicated by the circumstance that $\sigma_{x,s}$ is a function of m_1 and therefore of the frequency parameters, rather than a constant as before. The derivation is illustrated for the binomial distribution as follows:

$$\begin{aligned}
\mu_{s+1} &= \sum_{x=0}^{s+1} (n)_x p^x \sigma_{x,s+1} \\
&= \sum_{x=0}^{s+1} (n)_x p^x [(x - np)\sigma_{x,s} + \sigma_{x-1,s}] \\
&= \sum_{x=0}^{s} (n)_x \sigma_{x,s} pD_p(p^x) - np\mu_s + \sum_{x=0}^{s+1} (n)_x p^x \sigma_{x-1,s} \\
&= pD_p\mu_s + nsp\mu_{s-1} - np\mu_s + np\mu_s - p^2[D_p\mu_s + ns\mu_{s-1}] \\
&= pq\,[ns\mu_{s-1} + D_p\mu_s].
\end{aligned}$$

The steps in the process are expanded as follows:

$$\begin{aligned}
\sum_{x=0}^{s} (n)_x \sigma_{x,s} pD_p(p^x) &= \sum_{x=0}^{s} (n)_x [pD_p(p^x \sigma_{x,s}) - p^x pD_p(\sigma_{x,s})] \\
&= pD_p\mu_s - p \sum_{x=0}^{s} (n)_x p^x(-ns\sigma_{x,s-1}) \\
&= pD_p\mu_s + nsp\mu_{s-1} \\
\sum_{x=0}^{s+1} (n)_x p^x \sigma_{x-1,s} &= \sum_{x=0}^{s+1} (n - x + 1)(n)_{x-1} p^x \sigma_{x-1,s} \\
&= n \sum_{x=0}^{s} (n)_x p^{x+1} \sigma_{x,s} - \sum_{x=0}^{s} x(n)_x p^{x+1} \sigma_{x,s} \\
&= np\,\mu_s - p^2 [D_p\,\mu_s + ns\,\mu_{s-1}].
\end{aligned}$$

The relation $D_p\sigma_{x,s} = -ns\sigma_{x,s-1}$ is obtained from the definition equation of $\sigma_{x,s}$ (with $m_1 = np$).

The resulting recurrence relations for the three distributions are as follows:

(4.7) $$\mu_{s+1} = nspq\,\mu_{s-1} + pq\,D_p\,\mu_s$$ *Binomial*

(4.8) $$\mu_{s+1} = as\mu_{s-1} + a\,D_a\,\mu_s$$ *Poisson*

---

[4] Jordan, loc. cit. or E. C. Molina, *An Expansion for Laplacian Integrals . . .*, Bell System Technical Journal, **11**, p. 571.

$$(4.9)\qquad \mu_{s+1} = (n+1)\left[\mu_s - \sum_{v=0}^{s}\binom{s}{v}K_1^v\,\mu_{s-v}\,(l, r, n+1)\right] \quad \textit{Hypergeometric}$$
$$-\frac{lr}{n}\left[\mu_s - \sum_{v=0}^{s}\binom{s}{v}K_2^v\,\mu_{s-v}(l-1, r-1, n-1)\right]$$

where,

$$K_1 = \frac{-lr}{n(n+1)} = \Delta_n\frac{lr}{n}$$
$$K_2 = \frac{(l-1)(r-1)}{(n-1)} - \frac{lr}{n}.$$

The last of these, which appears to be new, seems to be of formal interest only.

The coefficients $\sigma_{x,s}$ are related to the Stirling numbers by the expression:

$$\sigma_{x,s} = \sum_{v=0}^{s-x}(-1)^v\binom{s}{v}S_{x,s-v}\,m_1^v = \sum_{v=0}^{s-x}a_v m_1^v$$

and consequently can be exhibited with detached coefficients in the form $a_0 + a_1 + a_2 + \cdots + a_{s-x}$. For the binomial and Poisson distributions certain simplifications, to be developed in the section following, in equations (4.3) and (4.4) may be made. For the hypergeometric distribution it appears necessary to use equation (4.5); the following short table of $\sigma_{x,s}$, employing the detached coefficients mentioned above, is given for this purpose:

$s \backslash x$	$\sigma_{x,s}$: 0	1	2	3	4	5
1	0−1	1				
2	0+0+1	1−2	1			
3	0+0+0−1	1−3+3	3−3	1		
4	0+0+0+0+1	1−4+6−4	7−12+6	6−4	1	
5	0+0+0+0+0−1	1−5+10−10+5	15−35+30−10	25−30+10	10−5	1

5. Binomial and Poisson Moments About the Mean—Simplified Formulas.

5.1 Binomial. From examination of the first few moments about the mean, it appears expedient[5] to write the formulas:

$$(5.1.1)\qquad \begin{aligned}\mu_{2s} &= \sum_{x=1}^{s}\alpha_{x,2s}\,(npq)^x\\ \mu_{2s+1} &= (q-p)\sum_{x=1}^{s}\alpha_{x,2s+1}\,(npq)^x\end{aligned}$$

[5] The kind of expression chosen admits of some variety. A recurrence relation for coefficients in the expansion $\mu_s = \sum_{x=1}^{s}\alpha_{x,s}\,p^x$ has been given by E. H. Larguier, *On a Method For Evaluating the Moments of a Bernoulli Distribution,* Bull. Am. Math. Soc., **42**, 1, p. 24 (Abstract 8); I am indebted to Mr. Larguier for the opportunity of examining his results in advance of publication.

When these are substituted into the moment recurrence relation, the coefficients are found to be related as follows:

$$\alpha_{x,2s} = [x + pqD_{pq}]\alpha_{x,2s-1} + (2s - 1)\alpha_{x-1,2s-2}$$
$$-2pq[1 + 2x + 2pqD_{pq}]\alpha_{x,2s-1}$$
$$\alpha_{x,2s+1} = [x + pqD_{pq}]\alpha_{x,2s} + 2s\alpha_{x-1,2s-2}$$

or, in general,

$$(5.1.2)\qquad \begin{aligned}\alpha_{x,s+1} &= [x + pqD_{pq}]\alpha_{x,s} + s\alpha_{x-1,s-1}\\ &\quad - pq[1 - (-1)^s][1 + 2x + 2pqD_{pq}]\alpha_{x,s}\end{aligned}$$

Using detached coefficients of powers of pq as outlined above, these coefficients may be exhibited as follows:

$s \backslash x$	$\alpha_{x,s}$ 1	2	3	4
2	1			
3	1			
4	1 − 6	3		
5	1 − 12	10		
6	1 − 30 + 120	25 − 130	15	
7	1 − 60 + 360	56 − 462	105	
8	1 − 126 + 1680 − 5040	119 − 2156 + 7308	490 − 2380	105
9	1 − 252 + 5040 − 20160	246 − 6948 + 32112	1918 − 13216	1260

It may be noted that the coefficients of the first column in conjunction with equations (5.1.1) give the binomial seminvariants.

Equations (5.1.1) make the coefficients functions of pq only; a slight alteration makes the coefficients functions of n only. Thus:

$$(5.1.3)\qquad \begin{aligned}\mu_{2s} &= \sum_{x=1}^{s} \beta_{x,2s}(pq)^x\\ \mu_{2s+1} &= (q - p)\sum_{x=1}^{s} \beta_{x,2s+1}(pq)^x\end{aligned}$$

and the coefficients are found to satisfy the recurrence relation:

$$(5.1.4)\qquad \beta_{x,s+1} = x\beta_{x,s} + ns\beta_{x-1,s-1} - [1 - (-1)^s](2x - 1)\beta_{x-1,s}.$$

These coefficients may be exhibited by a rearrangement of the table given above as may be seen by comparing equations (5.1.1) and (5.1.3). The first few coefficients are as follows:

$s \backslash x$	$n^{-1}\beta_{x,s}$ 1	2	3
2	1		
3	1		
4	1	$-6+3$	
5	1	$-12+10$	
6	1	$-30+25$	$120-130+15$

5.2 **Poisson.** The Poisson moments about the mean may be expressed as follows:

$$\mu_s = \sum_{x=0}^{[s/2]} \alpha_{x,s}\,\alpha^x \tag{5.2.1}$$

where [ ] represents "integral part of" and

$$\alpha_{x,s+1} = x\alpha_{x,s} + s\alpha_{x-1,s-1}. \tag{5.2.2}$$

The coefficients $\alpha_{x,s}$ are the constant terms in the expressions for the corresponding binomial distribution coefficients in powers of pq.

BELL TELEPHONE LABORATORIES.