

The proof consists in integrating $f(x)F(y)dxdy$ over the area bounded by the two lines $x + y = s$ and $x + y = s + \Delta s$, as shown in Figure 3.

THEOREM 4. *If x obeys a law $\int_{-\infty}^{\infty} f(x)dx = 1$, and y obeys a law $\int_{-\infty}^{\infty} F(y)dy = 1$, then the difference, $w = x - y$, will obey the law $\int_{-\infty}^{\infty} R(w)dw = 1$, where $R(w) = \int_{-\infty}^{\infty} f(w + y) F(y) dy$.*

The proof consists in integrating $f(x)F(y)dxdy$ over the area bounded by the two lines $x - y = w$ and $x - y = w + \Delta w$, as shown in Figure 4.

HARVARD UNIVERSITY.

MOMENTS ABOUT THE ARITHMETIC MEAN OF A HYPERGEOMETRIC FREQUENCY DISTRIBUTION

BY HAROLD D. LARSEN

In a recent paper¹ Kirkman has developed a method of continuation for obtaining the moments of a binomial distribution. Although other investigators² have found various methods which are perhaps superior from the standpoint of elegance and compactness, Kirkman's method is of some importance inasmuch as it is adaptable to use in a course in elementary statistics. With this thought in mind, we shall extend Kirkman's method to obtain the moments of the hypergeometric distribution of Table I.³

TABLE I

Variate v	Relative Frequency P_v
0	${}_n C_0 \alpha^{(0)} \beta^{(n)} / N^{(n)}$
1	${}_n C_1 \alpha^{(1)} \beta^{(n-1)} / N^{(n)}$
2	${}_n C_2 \alpha^{(2)} \beta^{(n-2)} / N^{(n)}$
.	.
.	.
n	${}_n C_n \alpha^{(n)} \beta^{(0)} / N^{(n)}$

¹ W. J. Kirkman, "Moments About the Arithmetic Mean of a Binomial Frequency Distribution," *Ann. Math. Statist.*, vol. vi, no. 2, June, 1935, pp. 96-101.

² For example, J. Riordan, "Moment Recurrence Relations for Binomial, Poisson and Hypergeometric Frequency Distributions," *Ann. Math. Statist.*, vol. viii, no. 2, June, 1937, pp. 103-111.

³ For the Poisson distribution, this method degenerates into the application of a well-known recursion formula.



The hypergeometric distribution above can be conceived as being generated in the following manner. From an urn containing N balls, $\alpha = Np$ white and $\beta = Nq$ black, n balls are drawn without replacements. The probability that exactly v of the balls are white is

$$P_v = {}_n C_v \alpha^{(v)} \beta^{(n-v)} / N^{(n)},$$

where

$$\alpha^{(v)} = \alpha(\alpha - 1)(\alpha - 2) \cdots (\alpha - v + 1),$$

$$\alpha^{(0)} = 1, \text{ etc.}$$

It may be noted in passing that the hypergeometric distribution reduces to a binomial distribution when $n = 1$, or $N = \infty$.

For the distribution of Table I, let m_k denote the k th moment about the origin, and let μ_k denote the k th moment about the arithmetic mean. Then by definition

$$m_k = \sum_{v=0}^n v^k P_v,$$

and

$$\mu_k = \sum_{v=0}^n (v - m_1)^k P_v.$$

It is apparent that these moments are functions of the parameters α, β, n and N . In particular,

$$m_k = F(\alpha, \beta, n, N).$$

We shall have need of the hypergeometric distribution of Table II. For the latter distribution, let ν_k denote the k th moment about the origin; i. e.,

$$\nu_k = \sum_{v=0}^{n-1} v^k P'_v.$$

TABLE II

v	P'_v
0	${}_{n-1}C_0(\alpha - 1)^{(0)}\beta^{(n-1)}/(N - 1)^{(n-1)}$
1	${}_{n-1}C_1(\alpha - 1)^{(1)}\beta^{(n-2)}/(N - 1)^{(n-1)}$
2	${}_{n-1}C_2(\alpha - 1)^{(2)}\beta^{(n-3)}/(N - 1)^{(n-1)}$
⋮	⋮ ⋮ ⋮ ⋮ ⋮
⋮	⋮ ⋮ ⋮ ⋮ ⋮
$n - 1$	${}_{n-1}C_{n-1}(\alpha - 1)^{(n-1)}\beta^{(0)}/(N - 1)^{(n-1)}$

Comparing Table I with Table II, we see at once that

$$(1) \quad \nu_k = F(\alpha - 1, \beta, n - 1, N - 1).$$

In other words, ν_k is equal to the expression obtained from m_k upon replacing α , n , and N respectively by $\alpha - 1$, $n - 1$, and $N - 1$.

Now consider

$$\begin{aligned} m_k &= \sum_{v=0}^n v^k P_v \\ &= \sum_{v=1}^n v^k P_v. \end{aligned}$$

Replacing v by $v + 1$, we have

$$\begin{aligned} m_k &= \sum_{v=0}^{n-1} (v+1)^k \cdot P_{v+1} \\ &= \frac{n\alpha}{N} \sum_{v=0}^{n-1} (v+1)^{k-1} \frac{(n-1)!}{v!(n-v-1)!} \frac{(\alpha-1)^{(v)} \beta^{(n-v-1)}}{(N-1)^{(n-1)}} \\ &= \frac{n\alpha}{N} \sum_{v=0}^{n-1} (v+1)^{k-1} P'_v, \end{aligned}$$

whence, expanding the binomial and summing term by term,

$$(2) \quad m_k = \frac{n\alpha}{N} \{ \nu_{k-1} + {}_{k-1}C_1 \nu_{k-2} + {}_{k-1}C_2 \nu_{k-3} + \cdots + 1 \}.$$

By repeated use of (1) and (2), we can obtain quite readily the moments about the origin for the distribution of Table I. It follows by definition that

$$m_0 = \sum_{v=0}^n P_v = 1,$$

and, similarly,

$$\nu_0 = \sum_{v=0}^{n-1} P'_v = 1.$$

Setting $k = 1$ in (2), we have

$$m_1 = \frac{n\alpha}{N} \cdot \nu_0 = n\alpha/N.$$

Setting $k = 2$, and then using (1), we obtain

$$\begin{aligned} m_2 &= \frac{n\alpha}{N} \{ \nu_1 + \nu_0 \} \\ &= \frac{n\alpha}{N} \left\{ \frac{(n-1)(\alpha-1)}{N-1} + 1 \right\} \\ &= \frac{n^{(2)} \alpha^{(2)}}{N^{(2)}} + \frac{n\alpha}{N}. \end{aligned}$$

In a similar manner,

$$\begin{aligned} m_3 &= \frac{n\alpha}{N} \{ \nu_2 + 2\nu_1 + \nu_0 \} \\ &= \frac{n\alpha}{N} \left\{ \frac{(n-1)^{(2)}(\alpha-1)^{(2)}}{(N-1)^{(2)}} + 3 \frac{(n-1)(\alpha-1)}{N-1} + 1 \right\} \\ &= \frac{n^{(3)}\alpha^{(3)}}{N^{(3)}} + 3 \frac{n^{(2)}\alpha^{(2)}}{N^{(2)}} + \frac{n\alpha}{N}. \end{aligned}$$

The coefficients are seen to follow the same law as for the binomial distribution. As a matter of fact, if we replace $\alpha^{(r)}/N^{(r)}$ by p^r in the above m 's, we obtain precisely the corresponding formulae for the binomial distribution. The coefficients for some of the higher moments are

$$m_4 = \{1, 6, 7, 1\}$$

$$m_5 = \{1, 10, 25, 15, 1\}$$

$$m_6 = \{1, 15, 65, 90, 31, 1\}.$$

The moments about the arithmetic mean can now be determined from the foregoing m 's by means of the semi-recursion formula

$$(3) \quad \mu_k = m_k - {}_k C_{1\mu_{k-1}} m_1 - {}_k C_{2\mu_{k-2}} m_1^2 - \dots$$

I have tried several formulae for this purpose, but it seems impossible to avoid a great deal of tedious reduction. Since the reduction in any case only involves algebraic manipulation, the details will be omitted. The formulae for the first few moments follow:

$$\mu_0 = 1$$

$$\mu_1 = 0$$

$$\mu_2 = npq \frac{N-n}{N-1}$$

$$\mu_3 = npq(q-p) \frac{(N-n)(N-2n)}{(N-1)(N-2)}.$$

If the higher moments are required in a practical problem, it appears to be the best course to first calculate the values of the m 's, and then use (3).