INTERIOR AND EXTERIOR MEANS OBTAINED BY THE METHOD OF MOMENTS

By Edward L. Dodd

1. **Introduction—The Substitutive Mean.** A very general mean based upon substitution was proposed by O. Chisini.[1] Briefly stated, this mean M of data $x_1, x_2, \cdots, x_n$, is a number which satisfies some equation of the form

$$G(M, M, \cdots, M) = G(x_1, x_2, \cdots, x_n). \tag{1}$$

If, now,[2]

$$M = F(x_1, x_2, \cdots, x_n) \tag{2}$$

is an *explicit* expression of M, then for each value c which each of the arguments x_i can take on,

$$F(c, c, \cdots, c) = c; \tag{3}$$

or at least one value of this F is c.

Suppose now that $F(x_1, x_2, \cdots, x_n)$ is *any* function of $x_1, x_2, \cdots, x_n$, defined for at least one set of equal arguments c, and such that whenever defined for equal arguments c, at least one value of $F(c, c, \cdots, c) = c$. Such a function, I have called a *substitutive mean*. Various extensions[3] are immediate, such as the use of integration in place of summation. Indeed, point set functions or functionals may be used.[4] Here I shall supplement (3) by a fairly common convention. If $F(c, c, \cdots, c)$ is not originally defined, but as $x_i \to c$ simultaneously, limit $F(x_1, x_2, \cdots, x_n) = c$, in this case, $F(c, c, \cdots, c)$ will be assigned its limiting value c,—thus establishing continuity.

2. **Location and scale as means.** The purpose of this paper is to investigate the nature of the means which arise when the well known *Method of Moments* is used to estimate the values of two important parameters—namely, the location κ and the scale α of a frequency function or distribution. These are taken as associated with the variable x of the distribution thus:

$$t' = (x - \kappa)/\alpha. \tag{4}$$

[1] O. Chisini, "Sul Concetto di media," *Periodico di Matematico*, Series 4, Vol. 9, (1929), pp. 106–116.

[2] E. L. Dodd, "Internal and External Means Arising from the Scaling of Frequency Functions," These *Annals*, Vol. 8, (1937), pp. 18–20.

[3] For an extension of Chisini's results, see: Bruno de Finetti, "Sul Concetto di media," *Giornale dell' Instituto Italiano degli Attuari*, Vol. 2, (1931), pp. 369–396.

[4] E. L. Dodd, "The Chief Characteristic of Statistical Means," Cowles Commission lecture. *Colorado College Publication*, General Series No. 208, (1936), pp. 89–92.

The nature of the distribution is then "specified" by

$$y = \alpha^{-1}\Phi(t'); \tag{5}$$

where Φ may contain other parameters, but in Φ the κ and α appear only in the t' given by (4). For this mode of approach, the reader is referred to R. A. Fisher.[5]

The other parameters which *may* appear in Φ will not be considered in this paper.

The parameters κ and α are in general unknown and unknowable. However, we attempt to get close *estimates* k and a, of κ and α respectively, from a set of observations

$$X_1, X_2, \cdots, X_n. \tag{6}$$

To accomplish this, we have to solve certain equations formed in some way from

$$t = (x - \mathrm{k})/\mathrm{a}, \tag{7}$$

and

$$y = \mathrm{a}^{-1}\Phi(t). \tag{8}$$

These equations (7) and (8) result from (4) and (5) by substituting estimates k and a, respectively, for parameters κ and α.

Now the Method of Moments equates the theoretic moments—those obtained from some such equation as (8) with t replaced by its value in (7)—to the moments obtained from the observation (6).

For the following discussion it will be useful to obtain "*auxiliary*" moments from the $\Phi(t)$ in (8) *before* substitution is made from (7). Such moments, then, *do not depend* at all upon the values ultimately assigned to k and a. It is supposed that

$$\int_{-\infty}^{\infty} \Phi(t)\, dt = 1, \tag{9}$$

so that $\Phi(t)$ gives probability or relative frequency. Here, for finite distributions, $\Phi(t) \equiv 0$ outside the interval of the distribution. We shall assume the existence of the first moment

$$\mu = \int_{-\infty}^{\infty} t\Phi(t)\, dt, \tag{10}$$

and of the variance

$$\sigma^2 = \int_{-\infty}^{\infty} (t - \mu)^2 \Phi(t)\, dt = \int_{-\infty}^{\infty} t^2 \Phi(t)\, dt - \mu^2; \tag{11}$$

and we shall assume that $\sigma > 0$, to eliminate a degenerate case.

[5] R. A. Fisher, "On the Mathematical Foundation of Theoretical Statistics" *Philosophical Transactions of the Royal Society of London*, Series A, Vol. 222, (1921), pp. 309–368.

For the empirical moments of (6), we write

$$\bar{X} = (X_1 + X_2 + \cdots + X_n)/n = \Sigma X_i/n, \tag{12}$$

$$S = \Sigma X_i^2/n = \Sigma(X_i - \bar{X})^2/n + \bar{X}^2. \tag{13}$$

These two moments are, by the Method of Moments, equated respectively to

$$M_1 = \frac{1}{\mathrm{a}} \int_{-\infty}^{\infty} x\Phi[(x - \mathrm{k})/\mathrm{a}]\, dx, \tag{14}$$

$$M_2 = \frac{1}{\mathrm{a}} \int_{-\infty}^{\infty} x^2\Phi[(x - \mathrm{k})/\mathrm{a}]\, dx. \tag{15}$$

But, from (7), (10), and (11) it is easy to see that

$$M_1 = \mathrm{k} + \mathrm{a}\mu, \tag{16}$$

$$M_2 = \mathrm{k}^2 + 2\mathrm{k}\mathrm{a}\mu + \mathrm{a}^2(\sigma^2 + \mu^2); \tag{17}$$

from which it follows that

$$M_2 - M_1^2 = \mathrm{a}^2\sigma^2. \tag{18}$$

Suppose now that τ^2 is the empirical variance,

$$\tau^2 = \Sigma(X_i - \bar{X})^2/n. \tag{19}$$

It follows from (12) and (13), that if $M_1 = \bar{X}$ and $M_2 = S$—as the Method of Moments requires—then

$$\mathrm{a}^2 = \tau^2/\sigma^2 = (\tau/\sigma)^2. \tag{20}$$

And, from (16),

$$\mathrm{k} = \bar{X} - \mathrm{a}\mu. \tag{21}$$

These results may be expressed in the following theorem:

THEOREM I. *The estimated scale* a *in*

$$y = \mathrm{a}^{-1}\Phi(t), \tag{8}$$

where by (7), $t = (x - \mathrm{k})/\mathrm{a}$, *as obtained by the Method of Moments from observations* $X_1, X_2, \cdots, X_n$, *is the root-mean square of* $| X_i - \bar{X} | /\sigma$, *where* $\bar{X}$ *is the arithmetic mean of the* X_i*'s, and* σ^2 *is the "theoretic" variance of* $\Phi(t)$ *itself, as a function of* t*—with no reference to the* k *or* a *in* (7).

Moreover, the estimated location k is a substitutive mean, characterized by (3), and given by

$$\mathrm{k} = \bar{X} - \mathrm{a}\mu = \bar{X} - [\Sigma(X_i - \bar{X})^2/\sigma^2 n]^{\frac{1}{2}}. \tag{21}$$

As regards this final statement, it will be seen that if each $X_i = c$, then $\bar{X} = c$; and hence $\mathrm{k} = c$,—as required by (3). We may say, then, that the right member of (21) obtained as the formal solution of equations which the problem sets up, is a substitutive mean of the elements X_i.

But if each $X_i = c$, then a $= 0$ in (21); and this a may not be used as a scale. However, if any two X_i's are different, a $\neq 0$. And it is evident that as $X_i \to c$ simultaneously, limit k $= c$. If, then, we consider that the right member of (21) is not originally defined for equal values c of the elements X_i, it is to be given its "continuity" value c, in accordance with the common convention already mentioned.

In the special case where the function Φ in (5) chosen to specify the distribution has a first moment μ equal to zero, the estimate k of location given by (21) is seen at once to be the arithmetic mean of the observations $X_1, X_2, \cdots, X_n$.

3. **External means.** In the papers cited, Chisini and DeFinetti gave examples of external means. Indeed, it is not difficult to find means which do not conform to the condition of internality:

$$\text{(22)} \qquad \text{Minimum } (X_i) \leqq \text{Mean } (X_i) \leqq \text{Maximum } (X_i).$$

As a simple illustration, suppose that there are just three measurements $X_1 = 1$, $X_2 = 1$, $X_3 = -2$. The standard deviation $\sqrt{2}$ is greater than each measurement—it is an external mean. In this case also, the estimate of scale mentioned in Theorem I is an external mean of $(X_i - \bar{X})/\sigma$. But, it may be noted that a, the estimate of scale, is an *internal* mean of $| X_i - \bar{X} |/\sigma$.

However, it will be shown now that the estimate k of location may be an external mean, with an *externality not "removable"* by the simple device of using absolute values.

And it may be noted that in the earlier paper cited, I found by the *Method of Maximum Likelihood* estimates of the *scale* a, which were likewise *not removable.*

THEOREM II. *If for the function* $\Phi(t)$ *in* (8), *the second moment is less than twice the square of the second moment, then the estimated location given by*

$$\text{(21)} \qquad \text{k} = X - \text{a}\mu$$

is an external mean of the measurements X_i, *if these are all numerically equal, half of them positive and the other half negative.*

Proof. Let the positive measurements be c, and the negative measurements be $-c$. Then $\bar{X} = 0$; also in (19), $\tau = c$. Hence from (20), a $= c/\sigma$. But by hypothesis, the second moment $\sigma^2 + \mu^2$ of $\Phi(t)$ in (11) is less than $2\mu^2$; and thus $| \mu/\sigma | > 1$. Then, by (21) k $= \bar{X} - \text{a}\mu = (-c/\sigma)\mu$; and hence $| \text{k} | > c$. Either k is greater than every positive measurement c, or it is less than every negative measurement $-c$. In either case, it is an external mean.

COROLLARY. *If in* $\Phi(t)$, *the* t *is subjected to a translation* $t = u + b$, *so that* $\Phi(t) = \Phi(u + b) = \Psi(u)$, *then it is always possible to choose* b *so that the second moment of* $\Psi(u)$ *is less than twice the square of its first moment; and thus if a location* k′ *is obtained from* $\Psi(u)$, *external means may occur. On the other hand, by proper choice of* b, *it is possible to make the first moment zero, so that the location becomes the arithmetic mean* $\bar{X}$ *of the* X_i*'s.*

The first part of this corollary may be seen from (11) which states that Second

Moment $= \mu^2 + \sigma^2$. Translation does not change σ^2, but it can increase μ^2 indefinitely,—making eventually $\mu^2 > \sigma^2$, and thus $\mu^2 + \sigma^2 < 2\mu^2$.

4. **Illustration.** For the Pearson Type III the simplest specification is perhaps with the origin at the start. In this case,

$$\Phi(t) = (p!)^{-1}e^{-t}t^p, \qquad p > -1, \qquad t \geqq 0, \qquad t = (x - \mathrm{k})/\mathrm{a}. \tag{23}$$

Here $p! = \Gamma(1 + p)$. Apart from this numerical factor, $\Phi(t)$ is the integrand of the Gamma function. With $\Phi(t)$ in this form, it is easily seen that the first moment is $(p + 1)$ and the second moment is $(p + 1)(p + 2)$. In the usual[6] case, $p > 0$. Here, then, the second moment is less than twice the square of the first moment. If, then, there are an even number of measurements, all numerically equal, with half the measurements positive and the other half negative, then the estimate k of location as found by the Method of Moments is an *external* mean of the measurements. Such conditions, while sufficient, are by *no means necessary* for externality.

5. **Summary.** Suppose that the specification for a frequency function in x is $\alpha^{-1}\Phi(t')$, where $t' = (x - \kappa)/\alpha$, and that for the unknown scale α and location κ, estimates a and k, respectively, are made by the Method of Moments from a set of n measurements X_i with arithmetic mean $\bar{X}$. Let σ^2 be the variance of $\Phi(t')$. Then the estimate a is the root-mean-square of $| X_i - \bar{X} | /\sigma$, an internal mean. The estimate k of the location is $\bar{X} - \mu\mathrm{a}$, where μ is the first moment of $\Phi(t')$. This is a substitutive mean of the measurements X_i; and it may be external—either greater than Maximum X_i or less than Minimum X_i.

THE UNIVERSITY OF TEXAS

[6] W. Palin Elderton, *Frequency Curves and Correlation*, Second Edition, p. 91.