THE SUBSTITUTIVE MEAN AND CERTAIN SUBCLASSES OF THIS GENERAL MEAN

By Edward L. Dodd

1. Introduction. No general agreement has been reached, so far as I know, as to what constitutes a mean. A *necessary condition* which appears to meet with general approval is that a single-valued mean of a set of numbers all equal to a constant c should itself be equal to c. However, there appears to be some valid objection against imposing any *other* proposed condition as *necessary*.

Of course, intermediacy is a condition that suggests itself at once. Indeed, in certain mean value theorems in general analysis—such as the First Theorem of the Mean for integral calculus, which I mention in Section 3—intermediacy is the main feature.

However, O. Chisini [1] insisted that intermediacy or internality is not the chief characteristic of a statistical mean. Rather, a mean is a number to take the place, by substitution, of each of a set of numbers in general different. Such a mean may well be called a *representative* or *substitutive* mean.

Chisini defined m to be a mean of $x_1, x_2, \cdots, x_n$, relative to a function F, provided that

$$F(m, m, \cdots, m) = F(x_1, x_2, \cdots, x_n). \tag{1.1}$$

If, for example,

$$F(x_1, x_2, \cdots, x_n) = \Sigma x_i^2 = \Sigma m^2 = nm^2, \tag{1.2}$$

the mean m thus obtained is the root-mean-square

$$m = \pm [(1/n)\Sigma x_i^2]^{1/2}. \tag{1.3}$$

The choice of F, Chisini noted, depended upon the use to be made of the mean.

Suppose now that $f(x_1, x_2, \cdots, x_n)$ is such a function that one value of

$$f(x, x, \cdots, x) = x. \tag{1.4}$$

And suppose that this f is taken as a particular F for (1.1) to determine a mean m *implicitly*; thus

$$f(m, m, \cdots, m) = f(x_1, x_2, \cdots, x_n). \tag{1.5}$$

Then, from (1.5) and (1.4) it follows that one value of

$$f(x_1, x_2, \cdots, x_n) = m. \tag{1.6}$$

And thus f determines the mean m both *explicitly* and *implicitly*.

It should be noted that the $F = \Sigma x_i^2$ in (1.2) is *not* itself a mean of the x_i.

If, in (1.2), we take $x_1 = -2$, $x_2 = 1$, $x_3 = 1$, then the double-valued mean $m = \pm 2^{1/2}$ results. Now $-2^{1/2}$ is *internal*; e.i. $-2 < -2^{1/2} < 1$; but $2^{1/2}$ is *external*, for $2^{1/2} > 1 > -2$. But since here $\Sigma x_i = 0$, it follows also that the *standard deviation* of -2, 1, 1, is the *external mean* $2^{1/2}$. Chisini [1], indeed, used the root mean square to show the possibility of external means. External means have been noted by other writers, [2–7].

It is noteworthy that a number of writers [8–12] have used the condition (1.4) (in general, with f single-valued) as one of a set of axioms to characterize particular means. Sometimes, this has appeared in weaker form as $f(1, 1, \cdots, 1) = 1$.

This paper will be concerned primarily with the mean of a finite number n, of variates, $x_1, x_2, \cdots, x_n$. Possible generalizations will be mentioned briefly in Section 8.

In the conception of the substitutive mean, m, as I have been using it for some time, emphasis is laid upon the *explicit* form for m; and provision is made for *multiple* values.

Definition of the Substitutive Mean. *Let $f(x_1, x_2, \cdots, x_n)$ be a function of n variables, $x_1, x_2, \cdots, x_n$ defined at least for one set of equal values, $x_i = k$. If c is any number such that $f(c, c, \cdots, c)$ is defined, let one value of*

$$f(c, c, \cdots, c) = c. \tag{1.7}$$

Then $f(x_1, x_2, \cdots, x_n)$ will be said to be a substitutive mean of $x_1, x_2, \cdots, x_n$.

If an original formulation of a problem does not assign to a function a value when the variables are all equal, it is sometimes possible to assign such values by continuity considerations, such as are commonly used in the "evaluation" of indeterminate forms. This will be discussed in Section 6.

In the following, when the word mean is used, it will designate the substitutive mean as defined above.

2. Classification of Means already made.

Some general classes of means have already been distinguished. One important basis for a classification of means is the kind of data to be used. The data may be only qualitatively distinguishable. Then numbers may be assigned to qualities. For dealing in a very general way with all kinds of data, C. Gini and L. Galvani [13], and G. Pietra [14], distinguished between data in rectilineal series, in cyclical series, and in unconnected series. These three classes are associated respectively with the straight line, the circle, and a regular polyhedron (in three dimensions, the regular tetrahedron, and in n dimensions, a polyhedron with $n + 1$ vertices each at the same distance from each of the other n vertices).

For one definition of the arithmetic mean of a cyclical series, Gini uses the center of gravity principle; and this mean is computed with the aid of sines and cosines. By mechanical means, such an arithmetic mean of dates—for example, of dates of weddings—as days of a year can be found. On the rim of a wheel delicately suspended and marked off for the 365 days or 366 days of a year, let small weights proportional to the number of weddings on a day be placed in the spaces assigned to the individual days. Then when the wheel comes to rest, the arithmetic mean of the dates will be found at the lowest point of the rim. In the special case where the center of gravity of the system is at the center of the circle, the mean is indeterminate, or we may say that every day is a mean day.

Also, for cyclical series the arithmetic mean and the median are defined by other methods, using such principles as minimizing the sum of the squares of deviations or the sum of the absolute deviations.

The properties of means may be made the basis of a classification, either those properties which have been evolved by writers [8–12], [15–18] who have characterized specific means by sets of axioms, or those properties which seem of special importance in making distinctions. Two such properties will now be mentioned.

Gini [19] recognizes two large classes of means: "*A*) medie ferme, B) medie lasche," the latter (loose) class including the median and mode for which values do not depend upon all the data. To describe this latter mean m of arguments x_i, we might write $\partial m/\partial x_i = 0$ as applying to several if not most of the arguments over wide ranges instead of at isolated points.

Subclasses of A or firm means as given by Gini will be discussed in Section 4.

Another rather large classification distinguishes between simple means and their weighted forms. In a case often encountered, where the weights are whole numbers indicating frequencies of occurrence this distinction is of little significance. In the more general case, however, where weights may give ratings of the efficiency of measuring instruments or the weights may be negative [6, 20], more direct attention needs to be paid the weighted forms.

To supplement classifications already proposed, I am indicating in the next section a descent from the substitutive mean, the most general of all means, down through two classes of means less general, which I am calling the summational mean and the quasi-arithmetic mean, to the more specific mean known as the associative mean, studied in particular by M. Nagumo, [21] A. Kolmogoroff, [22] and B. de Finetti, [2].

The foregoing subclasses of the general or substitutive mean are based primarily on structure, the way the mean is formed.

3. The Summational Mean, Quasi-Arithmetic Mean, and Associative Mean. The summational mean, now to be defined, is a generalization of the weighted arithmetic mean.

$$W = \frac{c_1x_1 + c_2x_2 + \cdots + c_nx_n}{c_1 + c_2 + \cdots + c_n}, \qquad \Sigma c_i \neq 0. \tag{3.1}$$

It is to be noted that although W is not a symmetric function of x_i, W is a symmetric function of $c_i x_i$. In the generalization Q, the following features of W are retained:

1. Certain weights c_i being given, Q is a symmetric function of $c_i x_i$.
2. This Q may be determined from sums of n *terms*, each term involving *one and only one* x_i.

DEFINITION. *Let Σ denote a summation for $i = 1, 2, \cdots, n$. Suppose that*

$$F\{y, \Sigma f_1(c_i x_i, y), \Sigma f_2(c_i x_i, y), \cdots, \Sigma f_k(c_i x_i, y)\} = 0 \tag{3.2}$$

has a solution, $y = Q$ which is a substitutive mean of $x_1, x_2, \cdots, x_n$. Then Q will be called a summational mean of $x_1, x_2, \cdots, x_n$, relative to the functions f_1, $f_2, \cdots f_k$, and F.

Sometimes it is possible to express Q as

$$Q = G\{\Sigma g_1(c_i x_i), \Sigma g_2(c_i x_i), \cdots, \Sigma g_k(c_i x_i)\}. \tag{3.3}$$

Among summational means, those of most frequent use involve in a special way but one summation. Thus with $\psi(x)$ a function, which would usually be taken as continuous, this m satisfies

$$\psi(m)\Sigma c_i = \Sigma c_i \psi(x_i). \tag{3.4}$$

But this, with $c_i > 0$, is just an algebraic analogue or prologue to the First Theorem of the Mean for integral calculus—the c_i to be replaced by a positive integrable function. Without further specification, this mean m may have an uncountably infinite number of values. But if it be required that $\psi(x)$ be a continuous increasing function, and that $c_i > 0$, then m is unique.

In a series of papers, C. E. Bonferroni [20], [23–27] used means such as m in (3.4) for statistical and actuarial problems. And, as he had in mind [28] distinctly the notion of substitution, he was in a sense a forerunner of Chisini. E. L. Dodd [29] made use of a mean m defined with the aid of n continuous increasing functions $\psi_i(x)$, thus:

$$\Sigma c_i \psi_i(m) = \Sigma c_i \psi_i(x_i), \qquad c_i > 0. \tag{3.5}$$

If $g_i(x) = c_i \psi_i(x)$, this can be written

$$\Sigma g_i(m) = \Sigma g_i(x_i). \tag{3.6}$$

In one paper, C. E. Bonferroni [20], as already noted, used weights which might be either positive or negative.

Some such mean as m in (3.4) has been used by a number of writers. Here $\psi(m)$ is a weighted arithmetic mean of $\psi(x_i)$; and thus it is natural to call m a quasi-arithmetic mean of x_i.

DEFINITION. *Let $\Sigma c_i \neq 0$. If m is a solution of*

$$\psi(m)\Sigma c_i = \Sigma c_i \psi(x_i), \tag{3.4}$$

then m will be called a quasi-arithmetic mean of x_i, with weights c_i, and relative to the function $\psi(x)$.

Sufficient conditions for the existence of this mean m are: (1) That $\psi(x)$ be continuous in the interval I, finite or infinite, in which the observations x_i lie; (2) That either $c_i > 0$ for each i, or that $\psi(x)$ take on all real values, as x runs through I.

It will be helpful to picture geometrically the double transformation or mirroring represented by (3.4). Points x_i on the horizontal axis are carried vertically to the curve $y = \psi(x)$ and then reflected horizontally to the y axis. For the points y_i, on the y axis thus obtained the arithmetic mean $\bar{y}$ or "center of gravity" is obtained. Then $\bar{y}$ is carried horizontally to the curve and reflected vertically to the x-axis. The abscissas m of points on the x-axis thus obtained are means of the given x_i, relative to this $\psi(x)$.

It may happen (Dodd [3 p. 746]) that the curve $y = \psi(x)$ contains horizontal segments, as in the curve for temperature y of ice-water-steam which has absorbed a quantity x of heat. In this case the mean m may be an "interval," an uncountable set of real numbers. Indeterminateness over an interval is a well known feature of the *median* of an *even* number of variates. In fact, a paper of D. Jackson [30] was for the purpose of indicating one method of selecting a single value from this interval of indeterminateness, as a median.

It may be noted that a mean of n variables becomes, when $n = 1$, a function of a single variable; and thus it appears possible to implant in a mean of n variables almost any peculiarity found in a function of one variable.

A special case of the quasi-arithmetic mean is the associative mean m which under some general conditions has been shown [2, 21, 22] to satisfy

$$n\psi(m) = \Sigma\psi(x_i), \qquad i = 1, 2, \cdots, n; \tag{3.7}$$

where $\psi(x)$ is a continuous increasing function.

If $f_n(x_1, x_2, \cdots, x_n)$ is an associative mean, then by definition, $f_n(x_1, x_2, \cdots, x_n)$ is unaltered when any k of the n variates are each replaced by the mean f_k of that set.

4. The Gini means as summational. Having distinguished firm means from loose means, Gini [19] noted that in the former class, a variate might appear as a base, as an exponent, or both as base and exponent. In general, these variates are to be positive. Gini then listed ten means of a decidedly broad character, some of them generalizing the combinatorial means treated by A. Durand [31] and O. Dunkel [32]. See also G. Pietra [37].

These ten means involve only the four simple arithmetic operations and root extraction. For many purposes they are best expressed in the form given by the author. However, to show that these means are summational, logarithms will be used to reduce products to sums.

Let

$$
(4.1)\quad
\begin{aligned}
S^p &= \Sigma x_i^p \qquad\qquad i = 1, 2, \cdots, n;\\
{}_nC_c &= n!/c!(n-c)!, \text{ a binomial coefficient;}\\
P_c &\text{ be any one of the } {}_nC_c \text{ products of } c \text{ different elements taken from } x_1, x_2, \cdots, x_n;\\
P_c^p &= (P_c)^p, \text{ the } p^{\text{th}} \text{ power of } P_c;\\
Z_c &= \Sigma P_c, \text{ the sum of all the } {}_nC_c \text{ products } P_c;\\
Z_c^p &= \Sigma P_c^p.
\end{aligned}
$$

In the expressions which follow, it is assumed that the denominators are not zero.

The ten means, as defined in Gini's Equations I, II, $\cdots$, X, will be designated here by $m_1, m_2, \cdots, m_{10}$; and their logarithms, with base arbitrary, will now be given.

$$
(4.2)\quad
\begin{aligned}
\log m_1 &= (\log S^p - \log n)/p\\
\log m_2 &= (\log Z_c - \log {}_nC_c)/c\\
\log m_3 &= (\log Z_c^p - \log {}_nC_c)/cp\\
\log m_4 &= (\log S^p - \log S^q)/(p-q)\\
\log m_5 &= \Sigma x_i^p \log x_i/S^p\\
\log m_6 &= (\log Z_c - \log Z_d - \log {}_nC_c + \log {}_nC_d)/(c-d)\\
\log m_7 &= (\log Z_c^p - \log Z_d^p - \log {}_nC_c + \log {}_nC_d)/(c-d)p\\
\log m_8 &= (\log Z_c^p - \log Z_c^q)/c(p-q)\\
\log m_9 &= \Sigma P_c^q \log P_c/cZ_c^q\\
\log m_{10} &= (\log Z_c^p - \log Z_d^q - \log {}_nC_c + \log {}_nC_d)/(cp-dq).
\end{aligned}
$$

As noted by the author, the foregoing include some well known special means. Thus, m_1 is the power mean, which for $p = 1, 2, -1$, becomes respectively the arithmetic mean, the root mean square, and the harmonic mean. If $p \to 0$, then the limit of m_3 and of m_7 is the geometric mean. If $p = 0, 1, 2$, and $q = p - 1$, then m_4 is respectively the harmonic, the arithmetic, and the contra-harmonic mean.

For each of the ten means, Gini gives an appropriate name. Those involving binomial coefficients are combinatorial, a mean like the contra-harmonic with denominator other than a constant is biplanar, the more simple means monoplanar.

When in the following, I show that certain combinatorial expressions may be replaced by sums, it is not implied that this replacement would simplify computation.

To prove that $m_1, m_2, \cdots, m_{10}$ are all summational means, it may be noted that n, p, q, c, d, ${}_nC_c$, and ${}_nC_d$ are constants. Moreover, S^p is the symmetric sum of the pth powers of x_i, thus with *only one* x_i in each term, and $i = 1, 2, \cdots, n$. And, since Z_c, Z_c^p, Z_d, and Z_d^p are symmetric polynomials in the x_i, they may be expressed as polynomials in S^1, S^2, $\cdots$, by a well known theorem of algebra. Hence among the ten means, the only one that requires special attention is the ninth mean, m_9.

To show that m_9 is a summational mean, we need only examine the numerator of the right member. Let this numerator be N.

$$N = \Sigma P_c^q \log P_c. \tag{4.3}$$

Then

$$qN = (x_1^q x_2^q \cdots x_c^q)(\log x_1^q + \cdots + \log x_c^q) + \cdots. \tag{4.4}$$

Thus, if we set $y_i = x_i^q$, we may write

$$qN = (y_1 y_2 \cdots y_c)(\log y_1 + \cdots + \log y_c) + \cdots. \tag{4.5}$$

The coefficient of $\log y_1$ in this right member is the sum of all products of c different factors which include y_1.

Now, let Y_r be the sum of the products of r different factors taken from $y_1, y_2, \cdots, y_n$; and let T_r be the sum of the products of r different factors taken from $y_2, y_3, \cdots, y_n$. Then it is evident that

$$Y_r = T_r + y_1 T_{r-1}; \qquad T_r = Y_r - y_1 T_{r-1}. \tag{4.6}$$

If, now, we set $Y_0 = 1$, it follows that

$$T_{c-1} = Y_{c-1} - y_1 Y_{c-2} + y_1^2 Y_{c-3} - \cdots + (-1)^{c-1} y_1^{c-1} Y_0. \tag{4.7}$$

Hence, in qN, the coefficient of $\log y_1$ is

$$y_1 T_{c-1} = y_1 Y_{c-1} - y_1^2 Y_{c-2} + \cdots + (-1)^{c-1} y_1^c Y_0. \tag{4.8}$$

Thus in qN, the terms containing $\log y_1$ are

$$Y_{c-1} y_1 \log y_1 - Y_{c-2} y_1^2 \log y_1 + \cdots + (-1) y_1^c \log y_1. \tag{4.9}$$

Now let

$$U_r = \Sigma y_i^r \log y_i, \qquad i = 1, 2, \cdots, n. \tag{4.10}$$

Then,

$$qN = Y_{c-1} U_1 - Y_{c-2} U_2 + \cdots + (-1)^{c-1} Y_0 U_c. \tag{4.11}$$

Thus, qN is here constructed from sums of n terms with but a *single* y_i in any term.

Likewise, with y_i replaced by x_i^q, a term contains but a single x_i.

5. Transformations. A function $f(x_1, x_2, \cdots, x_n)$ is not in general a mean of its arguments x_i. However, it is often possible to make a substitution $x_i = \phi(y_i)$ so that

$$f[\phi(y_1), \phi(y_2), \cdots, \phi(y_n)] = g(y_1, y_2, \cdots, y_n), \tag{5.1}$$

is a mean of its arguments y_i.

The required substitution is sometimes obvious, as in the case of the estimate s of scale

$$s = [(1/n)\Sigma(x_i - m)^2]^{1/2} = [(1/n)\Sigma y_i^2]^{1/2}. \tag{5.2}$$

Here s is a mean of y_i, although it is not a mean of x_i.

DEFINITION. *Let $y = \psi(x)$, in general multiple valued, be defined in an interval I, finite or infinite, the values of y lying in an interval J. Suppose that for each y in J, there is at least one x in I such that $\psi(x) = y$. Let any such x be designated by $\phi(y)$. Then $\phi(y)$ will be called the inverse of $\psi(x)$. It follows that one value of*

$$\psi[\phi(y)] = y. \tag{5.3}$$

THEOREM. *Let*

$$z = f(x_1, x_2, \cdots, x_n), \tag{5.4}$$

in general multiple valued, be defined when each x_i is in some interval I, finite or infinite. With x in I, set

$$\psi(x) = f(x, x, \cdots, x); \tag{5.5}$$

and suppose that $y = \psi(x)$ has an inverse, $x = \phi(y)$ defined in J. Let $x_i = \phi(y_i)$ be substituted into f to form the function

$$w = f[\phi(y_1), \phi(y_2), \cdots, \phi(y_n)] = g(y_1, y_2, \cdots, y_n). \tag{5.6}$$

Then w is a mean of y_i, defined when y_i is in J. It is thus a mean of $\psi(x_i)$, where x_i is in I.

If further, $\psi(x)$ is a continuous increasing function of x, then for a given set of x_i, the values of z and w are identical. The same is true for a given set of n values y_i.

PROOF. If each $y_i = c$, a number in J, then

$$f[\phi(y_1), \cdots, \phi(y_n)] = f[\phi(c), \cdots, \phi(c)] = \psi[\phi(c)]. \tag{5.7}$$

And one value of $\psi[\phi(c)]$ is c, from the definition of the inverse function $\phi(y)$. Moreover, if a number c' is taken in I, then $\psi(c')$ is some number in J, which we may call c; and the argument above is applicable. Finally, if $\psi(x)$ is continuous and increasing, then a number x_i in I is associated with one and only one y_i in J; and vice versa. Thus w and z become identical.

In the foregoing, we started with f which is not a mean of its arguments x_i, and obtained g which is a mean of y_i. Something like the reverse of this is possible. The last member of (5.2) is a mean of y_i. It was obtained by treating m as a constant, with respect to x_i. If, however, m is an estimate for location and is taken as $(1/n)\Sigma x_i$, and this is substituted into (5.2) then

$$s = \{[(n-1)/n]\Sigma x_i^2 - (2/n)\Sigma x_i x_j\}^{1/2}, \qquad i < j. \tag{5.8}$$

This s is now not a mean of x_i; for if x equal any constant c, then $s = 0$. Furthermore, there exists no single valued continuous increasing function $x = \phi(y)$ such that if $x_i = \phi(y_i)$ is substituted into (5.8), s will be a mean of the y_i. Thus the elimination of m from (5.2) interferes with the status of s as a mean of the x_i.

6. Indeterminate Forms that arise in testing for Means. Sometimes a function f is substantially continuous. But the investigation leading to the function fails to assign to the function a value for certain values of the argument x, or arguments, $x_1, x_2, \cdots, x_n$. However, values are often assignable which will make the function continuous. This is the usual occurrence when, in curve fitting, parameters are estimated. In general, the measurements are assumed to be not all alike. However, when a general function such as $\Sigma x_i/n$ for location is obtained, we do not hesitate to assign to this function the value c when each $x_i = c$, to make the function continuous.

As another illustration of "indeterminate forms," consider the Jackson [30] median, M, of four numbers $x_1 \leqq x_2 < x_3 \leqq x_4$, viz.,

$$M = (x_4x_3 - x_2x_1)/(x_4 + x_3 - x_2 - x_1). \tag{6.1}$$

A direct substitution of $x = c$, renders M indeterminate. But if $x_i \to c$, indeed, if merely $x_2 \to c$, and $x_3 \to c$, so also does M.

In a recent paper, R. Cisbani [33] generalizes means suggested by Dunkel [32] and L. Galvani [34] by setting up

$$y_j(x) = \left[n^{-1} \sum_{i=1}^{n} (a^j + ih)^{-x/j}\right]^{-1/x}, \qquad j \neq 0, \qquad x \neq 0; \tag{6.2}$$

and letting $n \to \infty$. There results an integral with the value

$$\bar{y}_j(x) = \left[\frac{b^{x+j} - a^{x+j}}{(x/j + 1)(b^j - a^j)}\right]^{1/x}, \tag{6.3}$$

for the case, $x \neq j$. This mean set up as a mean of an infinite number of variates turns out to be also a mean of the two numbers a and b,—which for $b = a$ becomes indeterminate. But as b approaches a, so also does $\bar{y}_j(x)$ approach a. This is also true for the special cases $x = -j$, etc.

In testing to see if a function m of x_i is a mean of these numbers, a difficulty sometimes arises, because a substitution of $x_i = c$ and $m = c$ into the equation which implicitly defines m will put zeros into denominators. An aid in such testing will now be formulated as a theorem, although the ideas involved are not essentially new.

THEOREM. *Let $f(x)$ be a continuous increasing function of x defined for each real x. Let*

$$f(0) = 0. \tag{6.4}$$

Given n real distinct numbers

$$x_1 < x_2 < \cdots < x_{n-1} < x_n , \tag{6.5}$$

n positive numbers, k_i, and a real number C.
Set

$$F(x) = \frac{k_1}{f(x_1 - x)} + \cdots + \frac{k_n}{f(x_n - x)} - C. \tag{6.6}$$

Then $F(x) = 0$ has $n - 1$ real roots m_j, such that

$$x_1 < m_1 < x_2 < m_2 < \cdots < m_{n-1} < x_n ; \tag{6.7}$$

also, a root less than x_1, provided

$$\Sigma k_i/f(+\infty) < C; \tag{6.8}$$

or a root greater than x_n, provided

$$\Sigma k_i/f(-\infty) > C. \tag{6.9}$$

PROOF. Since $f(x)$ is a continuous increasing function of x, so also is $k_i/f(x_i - x)$, except for the single value, $x = x_i$. So also, then, is $F(x)$, except when $x = x_1$ or x_2 or $\cdots$ or x_n. But

$$F(x_i + 0) = -\infty ; F(x_{i+1} - 0) = +\infty. \tag{6.10}$$

Hence, between x_i and x_{i+1}, there exists a root m_i, of $F(x) = 0$.
Moreover, since

$$F(-\infty) = [\Sigma k_i/f(+\infty)] - C; F(x_1 - 0) = +\infty ; \tag{6.11}$$

it follows that there is a root less than x_1, provided (6.8) is satisfied. Likewise, there is a root greater than x_n if (6.9) is satisfied.

The use of this theorem in testing for means is simple. Keeping the x_i distinct, the equation $F(x) = 0$ determines $(n - 1)$ numbers, m_j, such that if $x_i \to c$, so also do these $m_j \to c$. Employing continuity to define m_j when each $x_i = c$, we may say that each m_j is a mean of x_i ; $j = 1, 2, \cdots (n - 1)$; $i = 1, 2, \cdots n$, when the conditions of this theorem are satisfied. If $F(x) = 0$ has still another root, m, this m will not in general be a mean of x_i.

7. Summational Means arising in the Estimation of Parameters of Frequency Distributions. In curve fitting, the estimation of parameters leads in general to summational means. If the method of moments is used, the first step is to find the moments by summation. I have already considered estimates for location and scale by this method [7], and by the R. A. Fisher method of maximum likelihood [4]. A further study of the results of the likelihood method will now be made.

By this method, products which first appear are reduced to sums by logarithms, and the means found are, in general, summational. Some idea of the forms of these means can be obtained by examining a rather general form of frequency function which includes the Pearson Type I, and involves parameters with estimates $p > 0$ and $q > 0$, in addition to the location m and scale a. Let the observations be $x_1, x_2, \cdots, x_n$; let

$$t_i = (x_i - m)/a; \qquad 0 \leqq t \leqq 1; \qquad a > 0; \tag{7.1}$$

$$y = \frac{1}{a}\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} t^{p-1}(1-t)^{q-1} \tag{7.2}$$

The likelihood L is obtained by multiplying together the n factors obtained by substituting $t = t_1, t_2, \cdots, t_n$.
Then

$$\begin{aligned}\log L = -n \log a + n \log \Gamma(p+q) - n \log \Gamma(p) - n \log \Gamma(q) \\ + (p-1)\sum_1^n \log t_i + (q-1)\sum_1^n \log(1-t_i).\end{aligned} \tag{7.3}$$

From $\partial L/\partial m = 0$, there is obtained

$$P\Sigma \frac{1}{x_i - m} + Q\Sigma \frac{1}{x_i - m - a} = 0; \qquad P = p - 1, \qquad Q = q - 1. \tag{7.4}$$

Suppose $P \neq 0$ and $Q \neq 0$; and as a first case, suppose $P + Q \neq 0$. If each x_i is replaced by x, the above equation leads to $m = x - (Pa)/(P + Q)$.
Then m is a summational mean of

$$x_i' = x_i - (Pa)/(P+Q) \qquad i = 1, 2, \cdots, n; \tag{7.5}$$

as seen by applying the Theorem in Section 5.

Likewise, a is a summational mean of

$$x_i'' = (x_i - m)(P+Q)/P. \tag{7.6}$$

If $P \neq 0$, $Q \neq 0$; but $P + Q = 0$, then (7.4) becomes

$$\Sigma \frac{1}{x_i - m - a} = \Sigma \frac{1}{x_i - m}. \tag{7.7}$$

Now set $y_i = x_i - m$, $C = \Sigma 1/y_i$; and write (7.7) as

$$F(a) = \Sigma \frac{1}{y_i - a} - C = 0. \tag{7.8}$$

This has the form given in (6.6) with x replaced by a, $k_i = 1$, $f(a) = a$. If then $y_1 < y_2 < \cdots < y_n$, there exist $(n - 1)$ solutions a_j of $F(a) = 0$ between y_1

and y_n. And thus keeping the y_i distinct, if $y_i \to c$, so also do the $a_j \to c$. These a_j are then means of y_1, and thus, means of $x_i \div m$.

In the more general case where $P + Q \neq 0$, it is seen also that Q is a summational mean of

$$(7.9) \qquad P\left[\frac{a}{x_i - m} - 1\right].$$

From $\partial L/\partial a = 0$, quite analogous results are obtained. The special case now, however, is given by $P + Q + 1 = 0 = p + q - 1$. And, with the continuity interpretation, a is a mean of $x_i - m$; and moreover, m is a mean of $x_i - a$.

Using now the digamma function

$$(7.10) \qquad F(u) = \frac{d}{du} \log \Gamma(u),$$

set

$$(7.11) \qquad D(p) = F(p + q) - F(p).$$

The condition $\partial L/\partial p = 0$, then leads to

$$(7.12) \qquad D(p) = (1/n)\Sigma(-\log t_i), \qquad 0 < t_i \leqq 1.$$

Now, with $q > 0$, $D(\infty) = 0$, $D(-1 + 0) = \infty$; and $D(p)$ is a continuous decreasing function of p, when $p > -1$. Then, since $-\log t_i > 0$, there is a unique $p > -1$ to satisfy (6.12).

To be useful, here, p should be > 0. But, at all events, the p thus found is a mean of $D^{-1}(-\log t_i)$, where D^{-1} is inverse to D.

The digamma function (7.10) appears also in estimating the parameters for the Pearson Type III.

$$(7.13) \qquad y = \frac{1}{a}\frac{1}{\Gamma(p+1)} e^{-t} t^p, \qquad t = (x - m)/a, \qquad p > -1.$$

By setting $\partial L/\partial p = 0$, it is found that m is the *arithmetic* mean of $x_i - ae^{F(p+1)}$; a is the arithmetic mean of $(x_i - m)e^{-F(p+1)}$; while p is a summational mean of $F^{-1}\{\log (x_i - m)/a\} - 1$, where F^{-1} is the inverse of F. From $\partial L/\partial m = 0$, it is found that m is a summational mean of $x_i - pa$; a is the *harmonic* mean of $(x_i - m)/p$; and p is the harmonic mean of $(x_i - m)/a$. Finally, from $\partial L/\partial a = 0$, there is obtained

$$(7.14) \qquad (1/n)\Sigma x_i = m + a(p + 1),$$

which makes m, a and p each an *arithmetic* mean of a simple function of the observations x_i, when the other two estimates are taken as constants.

Comparison of (5.2) with (5.8) has shown that after complete elimination, estimates may cease to be means. However, it may be noted that s is more frequently exhibited in the form (5.2) where it is a mean than in the form (5.8) where it is not.

8. Generalizations. The extension of results from the discrete or discontinuous case where a mean m depends upon only a finite number of elements to the continuous case is fairly immediate, with integration taking the place of summation, and a distribution or frequency function taking the place of discrete weights, c_i. Stieltjes and Lebesque integrals may be used as well as Riemannian. Such a generalization of the Chisini mean was given by de Finetti [2].

The summational mean, which I have defined as involving possibly several summations, may be generalized likewise.

In terms of set functions, sometimes called functionelles, I gave [35] the following general definition of a mean with a point set H in mind as a distribution function.

DEFINITION. *Let E and H be sets of numbers. Such a number t may be a real number or a vector number $t = (t_1, t_2, \cdots, t_k)$.*

Let E_t be the result of replacing each number of E by a single number t.

Then the mean m of numbers in E, relative to the set H, and to a function f, is given by $m = f(E, H)$; provided that the function f has been so constructed that for each t in E, $f(E_t, H) = t$, or at least one value of this f is t. It is to be understood above that when E is changed to E_t, the set H remains unaltered.

This retains the chief feature of $f(t, t, \cdots, t) = t$ in explicit form or of $f(t, t, \cdots, t) = f(t_1, t_2, \cdots, t_n)$ in implicit form, where t is a mean of $t_1, t_2, \cdots, t_n$.

I used [36] a somewhat less general definition to discuss regression coefficients. All such means may well be called *substitutive* or *representative.*

REFERENCES

[1] O. CHISINI, "Sul concetto di media." *Periodico di Mat.*, Ser. 4, Vol. 9 (1929) pp. 106–116.

[2] BRUNO DE FINETTI, "Sul concetto di media." *Giornale dell' Inst. Italiano degli Attuari*, Vol. 2 (1931) pp. 369-396. See p. 378.

[3] E. L. DODD, "The complete independence of certain properties of means." *Annals of Math.*, Vol. 35 (1934) pp. 740–747.

[4] E. L. DODD, "Internal and external means arising from the scaling of frequency functions." *Annals of Math. Stat.*, Vol. 8 (1937) pp. 12-20.

[5] E. L. DODD, "Regression coefficients as means of certain ratios." *Am. Math. Monthly*, Vol. 44 (1937) pp. 306-308.

[6] E. L. DODD, "Index numbers and regression coefficients as means, internal and external." *Report of Third Annual Research Conference on Economics and Statistics.* Cowles Commission for Research in Economics, Colorado Springs, Colo., pp. 13, 14.

[7] E. L. DODD, "Interior and exterior means obtained by the method of moments." *Annals of Math. Stat.*, Vol. 9 (1938) pp. 153–157.

[8] O. SUTO, "Law of the arithmetic mean." *Tôhoku Math. Jour.*, Vol. 6 (1914–5) pp. 79-81.

[9] E. V. HUNTINGTON, "Sets of independent postulates for the arithmetic mean, the geometric mean, the harmonic mean, and the root-mean-square." *Trans. Am. Math. Soc.*, Vol. 29 (1927) pp. 1–22.

[10] S. NARUMI, "Note on the law of the arithmetic mean." *Tôhoku Math. Jour.*, Vol. 30 (1929) pp. 19–21.

[11] L. TEODORIU, "Sur la définition axiomatique de la moyenne." *Mathematicae*, Vol. 5 (1931) pp. 27–31.

[12] I. NAKAHARA, "Axioms for the weighted means." *Tôhoku Math. Jour.*, Vol. 41 (1936), pp. 424–434.

[13] C. GINI, AND L. GALVANI, "Di talune estensioni dei concetti di media ai caratteri qualitativa." *Metron*, Vol. 8 (1929) pp. 3–209.

[14] G. PIETRA, "The theory of statistical relations with special reference to cyclical series." *Metron*, Vol. 4 (1925) pp. 383–557.

[15] G. SCHIAPERELLI, "Come si possa giustificare l' uso della media arithmetica nel calcolo dei risultato d' osservazione." *Rendicon to Regio Inst. Lombardo*, Ser. 2, Vol. 40 (1907) pp. 752–764.

[16] U. BROGGI, "Sur le principe de la moyenne arithmétique. *L'Enseignement*, Vol. 11 (1909) pp. 14–17.

[17] R. SCHIMMACK, "Der Satz vom arithmetischen Mittel in axiomatischer Begrundung." *Math. Annalen*, Vol. 68 (1910) pp. 125–132.

[18] S. MATSUMURA, "Uber die Axiomatik von Mittelbildungen." *Tôhoku Math. Jour.*, Vol. 36 (1933) pp. 260–262.

[19] C. GINI, "Di una formula comprensiva delle medie." *Metron*, Vol. 13 (1938) pp. 3–22.

[20] C. E. BONFERRONI, "Sulle medie dedotte da funzioni concave." *Giornale di Mat. Finanziaria*, Vol. 9 (1927) pp. 13–24.

[21] M. NAGUMO, "Uber dine Klasse der Mittelwerte." *Japanese Jour. of Math.*, Vol. 7 (1930) pp. 71–79.

[22] A. KOLMOGOROFF, "Sur la notion de la moyenne." *Atti Regio Accademia Lincei*, Ser. 6, Vol. 12 (1930) pp. 388–391.

[23] C. E. BONFERRONI, "Scadenza media e leggi di capitalizzazione." *Giornale di Mat. Finanziaria*, Vol. 3 (1921) pp. 1–25.

[24] C. E. BONFERRONI, "Sulla scadenza media en senso lato." *Giornale di Mat. Finanziaria*, Vol. 4 (1922) pp. 160–181.

[25] C. E. BONFERRONI, "La media esponenziale in matematica finanziaria." *Regio Instituto Superiore di Scienze Economiche e Commerciali-Bari*, 1923–24.

[26] C. E. BONFERRONI, "Della vita matematica come media esponenziale," and Note. Ibid., 1924–25.

[27] C. E. BONFERRONI, "Sulle medie di potenze." *Giornale degli Economisti e Revista di Statistica*, Vol. 66 (1926) pp. 292–296.

[28] C. E. BONFERRONI, "A proposito di espressioni generali per le medie." *Giornale degli Economisti e Revista di Statistica*, Vol. 77 (1937) pp. 354–357.

[29] E. L. DODD, "Functions of measurements under general laws of error." *Skandinavisk Aktuarietidskrift*, Vol. 5 (1922) pp. 133–158.

[30] DUNHAM JACKSON, "Note on the median of a set of numbers." *Bull. Am. Math. Soc.*, Vol. 27 (1921) pp. 160–164.

[31] A. DURAND, "Sur un théoreme relatif a des moyennes." *Bull. des Sciences Math.*, Vol. 26 (1902) pp. 181–183.

[32] O. DUNKEL, "Generalized geometric means and algebraic equations." *Annals of Math.*, Vol. 11 (1909–10) pp. 21–32.

[33] R. CISBANI, "Contributi alla teoria della medie." *Metron*, Vol. 13 (1938) pp. 23–59.

[34] L. GALVANI, "Dei limite a cui tendono alcune medie." *Boll. Un. Mat. Italiano*, 1927.

[35] E. L. DODD, "The chief characteristic of statistical means." Cowles Commission for Research in Economics. *Colorado College Publication*, General Series No. 208 (1936) pp. 89–92.

[36] E. L. DODD, "Certain coefficients of regression or trend associated with largest likelihood." *Actualités Scientifiques et Industrielles*, No. 740 (1938) pp. 5–14.

[37] G. PIETRA, "Intorno alle medie combinatorie", *Supplemento statistico, etc.*, Ser. 2, Vol. 4 (1938). This reference is cited in "I Contributi Italiani al Progresso della Statistica," *Societa Italiana per il Progresso delle Scienze*, September 1939 on page 16, as dealing with the expression of power and combinatorial means, monoplane and biplane, as sums of power.

UNIVERSITY OF TEXAS,
AUSTIN, TEXAS.