# A RECONSIDERATION OF SHEPPARD'S CORRECTIONS

By W. T. Lewis[1]

In computing the moments of a frequency distribution it is customary to find first what are known as the raw moments. These are obtained on the assumption that all the material of each class interval is concentrated at the middle point of the interval. It introduces what is called a grouping error because in fact the material does not all lie at the middle point. To compensate for this error W. F. Sheppard[2] derived a set of corrections. The hypothesis underlying his method is that the distribution may be regarded as similar to one to which the Euler-MacLaurin summation formula without its end terms may be applied. He presupposed such a curve, found its true moments, and then the raw moments that would be obtained if its area were concentrated at several equidistant abscissae. The relationship between these raw moments and the true moments of the curve furnished him with the corrections required for that distribution. If now our observed distribution may be supposed to be sufficiently like that one, we may use his corrections also on the observed data. One may note four points of criticism.

(1) The given distribution may not be similar to the one suggested, in the sense that it would be close to such a curve if the intervals of grouping were made very small; or at all events the purpose of finding the moments may be in part to decide whether or not it would become such a curve, and so one would not like to assume that to be true at the outset. A special case of importance in which this last is true occurs when one is finding the moments of a sample in order to determine whether it may have been drawn from a presupposed universe. It is inexact to use raw moments but it is illogical to use corrections that have been proved only for the universe being tested.

(2) Sheppard's argument does not make use of the one certain fact that is given in the hypothesis, viz: that the partial area of the given distribution over each class interval is exactly as stated. In fact, if, following the argument of some authors, the given curve be assumed to be exponential, it obviously cannot have partial areas everywhere exactly equal to the several given frequencies, for in particular its partial area is not zero beyond the given range.

(3) It is common to find distributions which do not have high contact at the ends of the range and for them Sheppard's corrections certainly fail. To obviate this criticism new corrections have been derived by Pairman and Pear-

[1] With the assistance of Burton H. Camp.

[2] The true values are given on page 220 of "Mathematical Part of Elementary Statistics, by Camp, D. C. Heath and Company, 1931.

son for the so-called abrupt cases. These new corrections are adequate to care for the abrupt cases but involve so much computation that it is a fair question whether it would not be simpler, first to distribute the given material over each interval by a smoothing process, and then to find without corrections the moments of the smoothed distribution.

(4) Even if one admits Sheppard's method in general, waiving the dubious question as to whether it is proper to start with an assumed curve instead of starting with the given distribution, it is doubtful whether there are any curves which have exactly the properties required. The high contact hypothesis may be put in different language as follows: using the notation of the Handbook[3] page 92, let $f(x)$ be the curve and $x_t$ be the middle point of the slice. It is assumed that

$$\sum_t c x_t^r f^{(i)}(x_t) = \int_{-\infty}^{\infty} x^r f^{(i)}(x)\, dx; \qquad i = 0, 1, \cdots; \qquad r = 0, 1, \cdots;$$

$c$ being the class interval. This means that if the moments of the curve be found by using *mid-ordinates times class interval*, instead of *areas*, one will obtain exactly the true moments of the curve, and that this will remain true for all the curves which are derivatives of this curve. This property is certainly not true of the normal curve; but it is almost true when $r$ and the class interval are both small, and it is probably due to this fact that Sheppard's corrections seem to be good in practice.

Moreover, this high contact hypothesis cannot be true for any function over a limited range if the function is developable in Taylor's series about one end of the range. For the only function which has the required properties is identically zero, since the function and all its derivatives are required to vanish at that end of the range.

The primary purpose of this paper, therefore is to derive corrections similar to Sheppard's with a different set of assumptions. The results may be used as an approximate substitute for both Sheppard's and Pairman's. That is, they will apply approximately to both extreme cases and to the intermediate cases; on the whole they give better results than Sheppard's and are not so difficult to administer as Pairman's.

The argument runs as follows. When a distribution is given merely by class intervals, there is no way of knowing exactly what the distribution would have been had the class intervals been smaller; we do not know that we have a sample from an exponential curve, and even if we did we would not know that this sample would lie close to the exponential in form. We shall, however, try to draw a graduating curve in such a manner that (a) its partial area over each class interval will equal the frequency of the given distribution over that interval; and (b) its form within each class interval will be such that it will pass smoothly into the adjacent portions to the right and left. A good way to do this is by a

[3] H. L. Rietz, "Handbook of Math. Stat." Houghton Mifflin Co. (1924).

freehand graph, frankly recognizing that there are many forms that will do equally well. To obtain a numerical result it is necessary to use the equation of some curve. Again frankly recognizing that there are many types which will do equally well we choose the simplest to handle:

$$y = a + bt + ct^2.$$

Let the relative frequency distribution be defined by $f(i)$, $-m \leqq i \leqq n$, $m, n, i$ being integers. To satisfy (a) we have the equation

$$\int_{i-\frac{1}{2}}^{i+\frac{1}{2}} y \, dt = f(i).$$

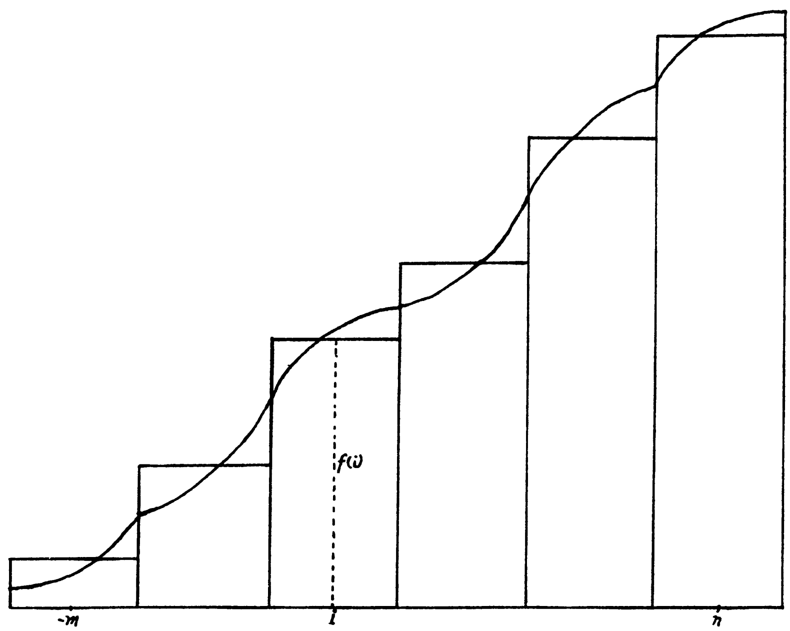


FIG. 1

To satisfy (b) we shall let

$$y = \tfrac{1}{2}[f(i) + f(i + 1)] \text{ if } t = i + \tfrac{1}{2}.$$

The latter will hold for all values of $i$ from $-m$ to $n - 1$ inclusive, but the end intervals require special treatment. Here in order to satisfy as well as possible both the high contact and the abrupt cases, we wish to let the material be distributed according to the way the curve is behaving over the two nearest intervals on the right (at $n$) or left (at $-m$) rather than by the addition of zero frequencies beyond the given limits. To do this we let the slope of the parabolas be zero at the extremes:

$$\frac{dy}{dt} = 0 \qquad \text{at } t = -m - \tfrac{1}{2} \text{ and } t = n + \tfrac{1}{2}.$$

Then, if for example the frequencies are increasing as one nears the right end interval, the curve will rise over the right end interval; if they are decreasing, it will fall. These three conditions are sufficient to determine a continuous curve of the sort indicated in the figure. The exact moments of the curve may be found by integration and expressed in terms of the raw moments. The details are tedious and of an elementary nature and will be given only for the mean value $\bar{\nu}_1$.

To determine the coefficients of the parabola $y = a + bt + ct^2$ for the rectangle at $t = i$ we may write the following three equations; the first complying with the requirement that the area under the parabola from $t = i - \frac{1}{2}$ to $t = i + \frac{1}{2}$ equals the area of the rectangle at $t = i$, the second and third giving the ordinates at $i - \frac{1}{2}$ and $i + \frac{1}{2}$ respectively:

$$f(i) = \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} (a + bt + ct^2)dt\,,$$

$$\frac{f(i) + f(i+1)}{2} = a + b\,(i + \tfrac{1}{2}) + c\,(i + \tfrac{1}{2})^2\,,$$

$$\frac{f(i) + f(i-1)}{2} = a + b\,(i - \tfrac{1}{2}) + c\,(i - \tfrac{1}{2})^2\,.$$

Solving these three simultaneous equations we get for $a$, $b$, and $c$:

$$a = (\tfrac{5}{4} - 3i^2)\,f(i) + \left(\frac{3i^2}{2} - \frac{i}{2} - \frac{1}{8}\right)f(i+1) + \left(\frac{3i^2}{2} + \frac{i}{2} - \frac{1}{8}\right)f(i-1)\,,$$

$$b = 6if(i) + (\tfrac{1}{2} - 3i)f(i+1) - (\tfrac{1}{2} + 3i)f(i-1)\,,$$

$$c = -\,3f(i) + \tfrac{3}{2}f(i+1) + \tfrac{3}{2}f(i-1)\,,$$

and these hold for $-m + 1 \leqq i \leqq n - 1$.

For the parabola $y = a_1 + b_1t + c_1t^2$ over the first rectangle, i.e., where $i = -m$, we get the equations:

$$f(-m) = \int_{-m-\frac{1}{2}}^{-m+\frac{1}{2}} (a_1 + b_1\,t + c_1\,t^2)dt\,,$$

$$\frac{f(-m) + f(-m+1)}{2} = a_1 + b_1\,(-m + \tfrac{1}{2}) + c_1\,(-m + \tfrac{1}{2})^2\,,$$

$$b_1 + 2\,c_1\,(-m - \tfrac{1}{2}) = 0\,,$$

and their solutions:

$$a_1 = \tfrac{3}{4}\,(m^2 + m - \tfrac{1}{12})\,f(-m+1) - \tfrac{3}{4}\,(m^2 + m - \tfrac{17}{12})\,f(-m)\,,$$

$$b_1 = \tfrac{3}{4}\,(2m + 1)\,f(-m+1) - \tfrac{3}{4}\,(2m+1)\,f(-m)\,,$$

$$c_1 = \tfrac{3}{4}f(-m+1) - \tfrac{3}{4}f(-m)\,.$$

Similarly for the parabola $y = a_n + b_n t + c_n t^2$ through the last rectangle at $i = n$ we get

$$f(n) = \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} (a_n + b_n t + c_n t^2) dt \,,$$

$$\frac{f(n) + f(n-1)}{2} = a_n + b_n (n - \tfrac{1}{2}) + c_n (n - \tfrac{1}{2})^2 \,,$$

$$b_n + 2\, c_n n + c_n = 0 \,,$$

and for the constants

$$a_n = \tfrac{3}{4} (n^2 + n - \tfrac{1}{12}) f(n-1) - \tfrac{3}{4} (n^2 + n - \tfrac{17}{12}) f(n) \,,$$

$$b_n = -\tfrac{3}{4} (1 + 2n) f(n-1) + \tfrac{3}{4} (1 + 2n) f(n) \,,$$

$$c_n = \tfrac{3}{4} f(n-1) - \tfrac{3}{4} f(n) \,.$$

Having obtained the constants for the graduating curve we will determine the moments of this curve in terms of those of the given frequency distribution.

*Notation:* Let the class interval be $c = 1$; let $\nu_s = \sum_{i=-m}^{n} i^s f(i)$ be the uncorrected $s^{\text{th}}$ moment of the given frequency distribution about the given origin; let $\mu_s = \sum_{i=-m}^{n} (i - \nu_1)^s f(i)$ be the uncorrected $s^{\text{th}}$ moment of the given frequency distribution about its uncorrected mean; let $\bar{\nu}_s$ be the corrected value of the $s^{\text{th}}$ moment about the given origin; and let $\bar{\mu}_s$ be the corrected value of the $s^{\text{th}}$ moment about the corrected mean. Thus $\nu_s$ and $\mu_s$ apply to the rectangles, and $\bar{\nu}_s$ and $\bar{\mu}_s$ apply to the curves as follows:

$$\bar{\nu}_s = \sum_{i=-m+1}^{n-1} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} t^s (a + bt + ct^2) dt + \int_{-m-\frac{1}{2}}^{-m+\frac{1}{2}} t^s (a_1 + b_1 t + c_1 t^2) dt + \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} t^s (a_n + b_n t + c_n t^2) dt \,,$$

$$\bar{\mu}_s = \sum_{i=-m+1}^{n-1} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} (t - \bar{\nu}_1)^s (a + bt + ct^2) dt + \int_{-m-\frac{1}{2}}^{-m+\frac{1}{2}} (t - \bar{\nu}_1)^s (a_1 + b_1 t + c_1 t^2) dt + \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} (t - \bar{\nu}_1)^s (a_n + b_n t + c_n t^2) dt \,.$$

Using these symbols we have for the first moment about the given origin:

$$\bar{\nu}_1 = \sum_{i=-m+1}^{n-1} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} t (a + bt + ct^2) dt + \int_{-m-\frac{1}{2}}^{-m+\frac{1}{2}} t (a_1 + b_1 t + c_1 t^2) dt + \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} t (a_n + b_n t + c_n t^2) dt$$

$$= \sum_{-m+1}^{n-1} \left[ai + b\left(i^2 + \frac{1}{12}\right) + c\left(i^3 + \frac{i}{4}\right)\right]$$

$$+ \left[- a_1 m + b_1\left(m^2 + \frac{1}{12}\right) - c_1\left(m^3 + \frac{m}{4}\right)\right]$$

$$+ \left[a_n n + b_n\left(n^2 + \frac{1}{12}\right) + c_n\left(n^3 + \frac{n}{4}\right)\right].$$

Substituting the values for the constants this becomes

$$\bar{\nu}_1 = \sum_{-m+1}^{n-1} \left\{i\left[\left(\frac{5}{4} - 3i^2\right)f(i) + \left(\frac{3i^2}{2} - \frac{i}{2} - \frac{1}{8}\right)f(i+1) + \left(\frac{3i^2}{2} + \frac{i}{2} - \frac{1}{8}\right)f(i-1)\right]\right.$$

$$+ (i^2 + \tfrac{1}{12})\,[6if(i) + (\tfrac{1}{2} - 3i)\,f(i+1) - (\tfrac{1}{2} + 3i)\,f(i-1)]$$

$$\left.+ \left(i^3 + \frac{i}{4}\right)[- 3f(i) + \tfrac{3}{2}f(i+1) + \tfrac{3}{2}f(i-1)]\right\}$$

$$+ \{- m\,[\tfrac{3}{4}\,(m^2 + m - \tfrac{1}{12})\,f(- m + 1) - \tfrac{3}{4}\,(m^2 + m - \tfrac{17}{12})\,f(- m)]$$

$$+ (m^2 + \tfrac{1}{12})\,[\tfrac{3}{4}\,(2m + 1)\,f(- m + 1) - \tfrac{3}{4}\,(2m + 1)\,f(- m)]$$

$$- \left(m^3 + \frac{m}{4}\right)[\tfrac{3}{4}f(- m + 1) - \tfrac{3}{4}f(- m)]\}$$

$$+ \{n[\tfrac{3}{4}\,(n^2 + n - \tfrac{1}{12})\,f(n-1) - \tfrac{3}{4}\,(n^2 + n - \tfrac{17}{12})\,f(n)]$$

$$+ \left(n^2 + \frac{1}{12}\right)\left[-\frac{3}{4}\,(1 + 2n)\,f(n-1) + \frac{3}{4}\,(1 + 2n)\,f(n)\right]$$

$$+ \left(n^3 + \frac{n}{4}\right)\left[\frac{3}{4}f(n-1) - \frac{3}{4}f(n)\right]\Bigg\}.$$

$$\bar{\nu}_1 = \sum_{-m+1}^{n-1}\left[if(i) + \frac{1}{24}f(i+1) - \frac{1}{24}f(i-1)\right] + \frac{1}{16}f(- m + 1)$$

$$- \left(m + \frac{1}{16}\right)f(- m) - \frac{1}{16}f(n-1) + \left(n + \frac{1}{16}\right)f(n).$$

$$\sum_{-m+1}^{n-1} if(i) = \sum_{-m}^{n} if(i) - (- m)\,f(- m) - nf(n) = \nu_1 + mf(- m) - nf(n).$$

$$\frac{1}{24}\sum_{i=-m+1}^{n-1} f(i+1) = \frac{1}{24}\sum_{k=-m+2}^{n} f(k) = \frac{1}{24}\sum_{-m}^{n} f(k) - \frac{1}{24}f(- m + 1) - \frac{1}{24}f(- m)$$

$$= \frac{1}{24} - \frac{1}{24}f(- m + 1) - \frac{1}{24}f(- m).$$

$$\frac{1}{24}\sum_{i=-m+1}^{n-1} f(i-1) = \frac{1}{24}\sum_{j=-m}^{n-2} f(j) = \frac{1}{24}\sum_{-m}^{n} f(j) - \frac{1}{24}f(n-1) - \frac{1}{24}f(n)$$

$$= \frac{1}{24} - \frac{1}{24}f(n-1) - \frac{1}{24}f(n)\,.$$

$$\bar{\nu}_1 = \nu_1 + mf(-m) - nf(n) + \frac{1}{24} - \frac{1}{24}f(-m+1) - \frac{1}{24}f(-m) - \frac{1}{24}$$

$$+ \frac{1}{24}f(n-1) + \frac{1}{24}f(n) + \frac{1}{16}f(-m+1)$$

$$- \left(m + \frac{1}{16}\right)f(-m) - \frac{1}{16}f(n-1) + \left(n + \frac{1}{16}\right)f(n)\,.$$

$$\bar{\nu}_1 = \nu_1 - \frac{5}{48}f(-m) + \frac{5}{48}f(n) + \frac{1}{48}f(-m+1) - \frac{1}{48}f(n-1)\,.$$

Using this same notation and method for the higher moments we get

$$\bar{\mu}_2 = \nu_2 - \frac{1}{12} - \bar{\nu}_1^2 + \left(\frac{5m}{24} + \frac{7}{80}\right)f(-m) + \left(\frac{5n}{24} + \frac{7}{80}\right)f(n)$$

$$+ \left(\frac{-m}{24} - \frac{1}{240}\right)f(-m+1) + \left(\frac{-n}{24} - \frac{1}{240}\right)f(n-1)\,.$$

$$\bar{\mu}_3 = \nu_3 - 3\bar{\nu}_1\bar{\mu}_2 - \frac{\bar{\nu}_1}{4} - \bar{\nu}_1^3 + f(-m)\left[\frac{-5}{16}m^2 - \frac{21}{80}m - \frac{17}{120}\right]$$

$$+ f(n)\left[\frac{5}{16}n^2 + \frac{21}{80}n + \frac{17}{120}\right] + f(-m+1)\left[\frac{m^2}{16} + \frac{m}{80} + \frac{1}{120}\right]$$

$$+ f(n-1)\left[\frac{-n^2}{16} - \frac{n}{80} - \frac{1}{120}\right].$$

$$\bar{\mu}_4 = \nu_4 - 4\bar{\mu}_3\bar{\nu}_1 - 6\bar{\mu}_2\bar{\nu}_1^2 - \bar{\nu}_1^4 - \frac{\bar{\mu}_2}{2} - \frac{\bar{\nu}_1^2}{2} - \frac{17}{64}$$

$$+ f(-m)\left[\frac{5m^3}{12} + \frac{21m^2}{40} + \frac{17m}{30} + \frac{313}{1680}\right] + f(n)\left[\frac{5n^3}{12} + \frac{21n^2}{40} + \frac{17n}{30} + \frac{313}{1680}\right]$$

$$+ f(-m+1)\left[\frac{-m^3}{12} - \frac{m^2}{40} - \frac{m}{30} - \frac{1}{336}\right] + f(n-1)\left[\frac{-n^3}{12} - \frac{n^2}{40} - \frac{n}{30} - \frac{1}{336}\right].$$

## SPECIAL CASES

The above formulae are rather long and in practice the special cases below will frequently be preferred.

(*a*) We may usually take the origin at or very near the middle of the range so that $m = n$, at least approximately.

If $m = n$:

$$\bar{\nu}_1 = \nu_1 - \frac{5}{48}f(-m) + \frac{5}{48}f(n) + \frac{1}{48}f(-m+1) - \frac{1}{48}f(n-1).$$

$$\bar{\mu}_2 = \nu_2 - \frac{1}{12} - \bar{\nu}_1^2 + \left(\frac{5m}{24} + \frac{7}{80}\right)[f(-m) + f(n)] + \left(\frac{-m}{24} - \frac{1}{240}\right)[f(-m+1) + f(n-1)].$$

$$\bar{\mu}_3 = \nu_3 - 3\bar{\nu}_1\bar{\mu}_2 - \frac{\bar{\nu}_1}{4} - \bar{\nu}_1^3 + \left[\frac{5m^2}{16} + \frac{21m}{80} + \frac{17}{120}\right][f(n) - f(-m)] + \left[\frac{m^2}{16} + \frac{m}{80} + \frac{1}{120}\right][f(-m+1) - f(n-1)].$$

$$\bar{\mu}_4 = \nu_4 - 4\bar{\mu}_3\bar{\nu}_1 - 6\bar{\mu}_2\bar{\nu}_1^2 - \bar{\nu}_1^4 - \frac{\bar{\mu}_2}{2} - \frac{\bar{\nu}_1^2}{2} - \frac{17}{64} + \left[\frac{-m^3}{12} - \frac{m^2}{40} - \frac{m}{30} - \frac{1}{336}\right][f(-m+1) + f(n-1)] + \left[\frac{5m^3}{12} + \frac{21m^2}{40} + \frac{17m}{30} + \frac{313}{1680}\right][f(-m) + f(n)].$$

(*b*) Except in the abrupt cases the end frequencies and the difference between those next to the ends will be so small (relative to unity) that they will have a negligible effect on the corrections. If $m = n$ as in (*a*), and if also

$$f(-m) = f(n) = 0 \text{ and } f(-m+1) - f(n-1) = 0:$$

$$\bar{\nu}_1 = \nu_1.$$

$$\bar{\mu}_2 = \nu_2 - \bar{\nu}_1^2 - \frac{1}{12} + f(-m+1)\left[\frac{-m}{12} - \frac{1}{120}\right].$$

$$\bar{\mu}_3 = \nu_3 - 3\bar{\nu}_1\bar{\mu}_2 - \frac{\bar{\nu}_1}{4} - \bar{\nu}_1^3.$$

$$\bar{\mu}_4 = \nu_4 - 4\bar{\mu}_3\bar{\nu}_1 - 6\bar{\mu}_2\bar{\nu}_1^2 - \bar{\nu}_1^4 - \frac{\bar{\mu}_2}{2} - \frac{\bar{\nu}_1^2}{2} - \frac{17}{64} + f(-m+1)\left[\frac{-m^3}{6} - \frac{m^2}{20} - \frac{m}{15} - \frac{1}{168}\right].$$

These formulae have been written in the form which makes the computing simple. The following makes a comparison with Sheppard's corrections easy.

$$\bar{\nu}_1 = \nu_1 .$$

$$\bar{\mu}_2 = \mu_2 - \frac{1}{12} + f(-m+1)\left[\frac{-m}{12} - \frac{1}{120}\right].$$

$$\bar{\mu}_3 = \mu_3 + \nu_1\left(\frac{m}{4} + \frac{1}{40}\right)f(-m+1).$$

$$\bar{\mu}_4 = \mu_4 - \frac{\mu_2}{2} - \frac{43}{192} + f(-m+1)\left[\frac{-m^3}{6} - \frac{m^2}{20} - \frac{m}{40} - \frac{1}{560} - \frac{m\nu_1^2}{2} - \frac{\nu_1^2}{20}\right].$$

The following special case is also useful in comparing my formulae with Sheppard's.

(*c*) Let $f(-m) = \frac{1}{5} f(-m+1)$ and $f(n) = \frac{1}{5} f(n-1)$. This produces a graduating curve which is exactly tangent to the $t$-axis at the ends of the range and is everywhere continuous—though it does not have continuous derivatives at certain isolated points. It is, however, a curve which to the eye cannot be distinguished from the type assumed in the Euler-MacLaurin theorem, which lies at the base of Sheppard's formulae. My corrections become:

$$\bar{\nu}_1 = \nu_1 ,$$

$$\bar{\mu}_2 = \mu_2 - \frac{1}{12} + \frac{1}{15}[f(-m) + f(n)],$$

$$\bar{\mu}_3 = \mu_3 - \frac{\nu_1}{5}[f(-m) + f(n)] + \left[\frac{-m}{5} - \frac{1}{10}\right]f(-m) + \left[\frac{n}{5} + \frac{1}{10}\right]f(n),$$

$$\bar{\mu}_4 = \mu_4 - \frac{\mu_2}{2} - \frac{43}{192} + \frac{2}{5}\left[(\nu_1 + m)^2 + \nu_1 + m + \frac{29}{84}\right]f(-m)$$
$$+ \frac{2}{5}\left[(\nu_1 - n)^2 - \nu_1 + n + \frac{29}{84}\right]f(n).$$

Sheppard's are:

$$\bar{\nu}_1 = \nu_1 ,$$

$$\bar{\mu}_2 = \mu_2 - \frac{1}{12},$$

$$\bar{\mu}_3 = \mu_3 ,$$

$$\bar{\mu}_4 = \mu_4 - \frac{\mu_2}{2} + \frac{7}{240}.$$

Let us compare my results with Sheppard's in the very special case in which $f(-m) = f(n) = 1/7$, $f(0) = 5/7$, $m = n = 1$. The odd moments vanish. My corrections for $\mu_2$ and $\mu_4$ are

$$\bar{\mu}_2 = 0.2214, \qquad \bar{\mu}_4 = 0.1870.$$

Sheppard's are

$$\bar{\mu}_2 = 0.2024, \qquad \bar{\mu}_4 = 0.1720.$$

The numerical difference between the $\bar{\mu}_2$'s is 0.0190, and the numerical difference between the $\bar{\mu}_4$'s is 0.0150.

This example shows that Sheppard's corrections are not valid to the precision to which they are usually given if they are to be used for the purpose of correcting raw moments. The last term in the fourth moment correction, 7/240, might equally well be, for example, $-43/192$ as in my special case. This will become more evident to the reader if he will draw the curve indicated in this example. To the eye it will appear exactly like the kind specified in the Euler-MacLaurin theorem; for example, much like the normal curve. Now suppose one adopted for the moment the point of view (which I have criticized earlier) of starting with the curve used in this example, breaking it up into three partial areas and then finding the relation between the true and the raw moments. The partial areas found would be exactly those used in this example and this method would give us Sheppard's corrections, but they would not be exactly correct, for in this instance my formulae give exactly the relationship between the true and the raw moments. The difference is due to the fact that in this instance the assumptions permitting the use of the Euler-MacLaurin theorem in abbreviated form are not justified for this curve. But there is no way of telling at the outset, if one has given initially only the partial areas, whether precisely this curve or another which to the eye would appear very much like it is truly the curve which will graduate the same material when subjected to a finer classification.