AN ABSTRACT OF THE THESIS OF

VISHNU BALCHAND JUMANI for the Master of Science
(Name of student) (degree)

in Applied Mathematics presented on May 9, 1969
(Major) (Date)

Title: SHEPPARD'S CORRECTION: MOMENT RELATIONSHIPS AND CORRECTION TERMS

Abstract approved: Redacted for Privacy
(William M. Stone)

This thesis is concerned with the detailed derivation of Sheppard's Correction Formula relating the true to the raw moments by the use of Euler-MacLaurin formula. Special moment relationships have been derived using different special functions, whose application is made to derive correction formula for the semi-invariants. Special investigation is carried out on the error term resulting from the moment relationship mentioned above.

Further, a moment relationship is established for a normal distribution function using a Taylor series approach.

SHEPPARD'S CORRECTION: MOMENT RELATIONSHIPS AND CORRECTION TERMS

by

VISHNU BALCHAND JUMANI

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

June 1969

APPROVED:

Redacted for Privacy

Professor of Department of Mathematics

in charge of major

Redacted for Privacy

Professor & Acting Chairman, Department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 9, 1969

Typed by Charlene Laski for Vishnu Balchand Jumani

ACKNOWLEDGEMENT

I wish to express my appreciation to Dr. William Stone for guiding the development and aiding in the writing of this thesis.

TABLE OF CONTENTS

Chapter I

Introduction

The Sheppard's correction for the calculation of moments of order n of the frequency function $f(x)$, which is assumed to be single-valued and continuous, is based on the fact that the domain of the frequency function is broken up into equal intervals. Let $\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots$ represent an infinite set of points on the x-axis, such that the difference $x_{i+1} - x_i = \omega$ is a constant. Furthermore, let $f(x)$ be so normalized that

$$(1.1) \qquad M_0 = \int_{-\infty}^{\infty} f(x)\, dx = 1.$$

The nth order moment of the frequency function $f(x)$ is defined as

$$(1.2) \qquad M_n = \int_{-\infty}^{\infty} x^n f(x)\, dx,$$

while the calculated or raw moments are defined as

$$(1.3) \qquad \overline{M}_n = \sum_{i=-\infty}^{\infty} x_i^n A_i,$$

where

$$(1.4) \qquad A_i = \int_{x_i - \frac{\omega}{2}}^{x_i + \frac{\omega}{2}} f(x)\, dx = \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i + y)\, dy$$

is the frequency corresponding to the interval $(x_i - \frac{\omega}{2}, x_i + \frac{\omega}{2})$. Another assumption about the frequency $f(x)$ is that it has a high order of contact with the x-axis as x becomes large, such that

$$\lim_{x\to\pm\infty} x^n f^{(2m)}(x) \to 0, \tag{1.5}$$

for all positive integral n and m.

The form of the Euler-MacLaurin identity of interest at this point is that presented in the Cramer (1946) treatise,

$$\sum_{n=n_1}^{n_2} g(a+n\omega) = \int_{n_1}^{n_2} g(a+x\omega)\,dx + \tfrac{1}{2}g(a+n_1\omega) + \tfrac{1}{2}g(a+n_2\omega)$$

$$-\omega \int_{n_1}^{n_2} P_1(x)\, g'(a+x\omega)\,dx. \tag{1.6}$$

If $g(x)$ has continuous derivatives of higher orders the last term can be transformed by repeated partial integration,

$$\sum_{n=n_1}^{n_2} g(a+n\omega) = \int_{n_1}^{n_2} g(a+x\omega)\,dx + \tfrac{1}{2}g(a+n_1\omega) + \tfrac{1}{2}g(a+n_2\omega)$$

$$- \sum_{n=1}^{N} \frac{B_{2n}}{(2n)!}\,\omega^{2n-1}[g^{(2n-1)}(a+n_1\omega) - g^{(2n-1)}(a+n_2\omega)]$$

(1.7)

(1.7) con't

$$+ (-1)^{N+1}\omega^{2N+1} \int_{n_1}^{n_2} P_{2N+1}(x)g^{(2N+1)}(a+x\omega)dx,$$

where the set of even and odd periodic functions is defined by

(1.8) $$P_{2k}(x) = \sum_{n=1}^{\infty} \frac{\cos 2n\pi x}{2^{2n-1}(n\pi)^{2k}}, \quad P_{2k+1}(x) = \sum_{n=1}^{\infty} \frac{\sin 2n\pi x}{2^{2n}(n\pi)^{2k+1}}.$$

Each of these functions has unit period; in particular, $P_1(x)$ is the sawtooth function. The Bernoulli polynomials of order n are given by the identity

(1.9) $$\frac{t^n e^{xt}}{(e^t - 1)^n} = \sum_{k=0}^{\infty} \frac{B_k^{(n)}(x)}{k!} t^k;$$

For $x = 0$ the general set of Bernoulli numbers is defined by

(1.10) $$\frac{t^n}{(e^t - 1)^n} = \sum_{k=0}^{\infty} \frac{B_k^{(n)}}{k!} t^k.$$

That is, $B_0^{(n)} = 1$, $B_1^{(n)} = -\frac{n}{2}$, $B_2^{(n)} = \frac{n(3n-1)}{12}$,
$B_3^{(n)} = -\frac{n^2(n-1)}{8}$, $B_4^{(n)} = \frac{n(15n^3 - 30n^2 + 5n + 2}{240}$,
$B_6^{(n)} = \frac{n(63n^5 - 315n^4 + 315n^3 + 91n^2 - 42n - 16)}{4032}$, ...

The case $n = 1$, written B_{2k} in the identity (1.7), is of most significance: $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$,

$B_6 = \frac{1}{42}$, $B_8 = -\frac{1}{30}$, $B_{10} = \frac{5}{66}$, $B_{12} = -\frac{691}{2730}$, $B_{14} = \frac{7}{6}$,

$B_{16} = -\frac{3617}{510}$, ... , and all odd order numbers, except for B_1, are zero.

The simple Euler-MacLaurin identity of (1.6) may be implemented in the form

$$(1.11) \quad \omega \sum_{i=-\infty}^{\infty} x_i^n f^{(2m)}(x_i) = \int_{-\infty}^{\infty} x^n f^{(2m)}(x)\, dx + R_1 ,$$

The present idea is to discuss the deduction of the Sheppard correction of the raw moments associated with a frequency function of high order contact. The development is based on the condition

$$(1.12) \qquad x^{n+1} \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x+y)\, dy \to 0 \quad \text{as} \quad x \to \pm\infty ,$$

which is not a severe restriction on $f(x)$. Using (1.6) in the form

$$(1.13) \quad \sum_{i=-\infty}^{\infty} x_i^n \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy = \frac{1}{\omega} \int_{-\infty}^{\infty} x^n dx \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x+y)\, dy + R_1$$

it is easy to arrive at a formula giving the raw moments in terms of the true moments

$$(1.14) \qquad \overline{M}_n = \frac{1}{n+1} \sum_{i=0}^{[\frac{n}{2}]} \binom{n+1}{2i+1} \left(\frac{\omega}{2}\right)^{2i} M_{n-2i} ,$$

where [x] is taken to mean "greatest integer $\leq x$".

It is the raw moments which are measurable in the field so it is clearly preferable to invert the infinite set of linear algebraic equations in (1.14). Crámer (1946) writes down the inverse set (Sheppard's Correction Formula)

$$(1.15) \qquad M_n = \sum_{i=0}^{n} \binom{n}{i} (2-2^i) \left(\frac{\omega}{2}\right)^i B_i \overline{M}_{n-i} ,$$

and refers to a paper by Wold (1934) for the procedure involved. A more complete study of the problem was made by Kendall (1938) and Langdon and Ore (1929). Part of this thesis is concerned with the detailed derivation of the several moment relationships. Wold (1934) also derived the correction formula described in the Langdon and Ore (1929) paper for the semi-invariants $\{\lambda_i\}$, which take a prominent place in error analysis. There should be an additive error term in (1.15) but analysis of that is postponed. The study of moment relationships may be made more general by the introduction of an arbitrary function $\phi(x)$, so

$$(1.16) \qquad M[\phi(x)] = \int_{-\infty}^{\infty} \phi(x)\ f(x)\ dx.$$

This thesis is also concerned with the study of the remainder term R_1 appearing in (1.16) and derivation of the moment relationship of a given normal distribution function using a Taylor series expansion.

Chapter II

Moment Relationships

In terms of an arbitrary $\phi(x)$ define a new function $g(x)$ by a functional equation.

$$\frac{1}{\omega}\int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} g(x+y)\,dy = \phi(x). \tag{2.1}$$

If $\phi'(x)$ exists the integral relationship may be changed to a difference relationship,

$$\mathop{\Delta}_{\omega} g(x) = \frac{g(x+\omega) - g(x)}{\omega} = \phi'(x+\tfrac{\omega}{2}). \tag{2.2}$$

So the general moment problem reduces to the evaluation of the integral

$$\begin{aligned} M(\phi) &= \int_{-\infty}^{\infty} \phi(x)\, f(x)\, dy \\ &= \sum_{i=-\infty}^{\infty} g(x_i) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy + R_1 , \end{aligned} \tag{2.3}$$

which follows from the Euler-MacLaurin formula. A second application yields

$$\sum_{i=-\infty}^{\infty} g(x_i) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy + R_1 = \frac{1}{\omega}\int_{-\infty}^{\infty} g(x)\, dx \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x+y) dy, \tag{2.4}$$

or

$$(2.5) \qquad M(\phi) = \int_{-\infty}^{\infty} f(x)\, dx \, \frac{1}{\omega}\int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} g(x+y)\, dy.$$

This relation holds up under the assumption that the condition

$$(2.6) \qquad \lim_{x\to\pm\infty} \left[\int_0^x g(y)\, dy\right]\left[\int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x+y)\, dy\right] = 0$$

is satisfied. The remainder R_1 can be written as

$$R_1 = \frac{1}{\omega}\int_{-\infty}^{\infty} g(x)\, dx \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x+y)\, dy - \sum_{i=-\infty}^{\infty} g(x_i) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy$$

$$(2.7) \qquad = \int_{-\infty}^{\infty} P_1(x)\, F'(x)\, dx,$$

where

$$(2.8) \qquad P_1(x) = \left[\frac{x}{\omega}\right] - \frac{x}{\omega} + \frac{1}{2} = \sum_{i=1}^{\infty} \frac{\sin 2i\pi \frac{x}{\omega}}{\pi i},$$

the sawtooth function. Also,

$$(2.9) \qquad F(x) = g(x) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x+y)\, dy$$

To demonstrate that (2.5) is valid observe that the right member of (2.4) may be integrated by parts, yielding

$$\frac{1}{\omega}\int_{-\infty}^{\infty} g(x)\ dx \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x+y)\ dy = [\int_0^x g(y)\ dy \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x+y)\ dy]\Big|_{-\infty}^{\infty}$$

$$(2.10)\qquad -\int_{-\infty}^{\infty}\int_0^x g(y)\ dy\ [f(x+\tfrac{\omega}{2}) - f(x - \tfrac{\omega}{2})]\ dx$$

$$= \int_{-\infty}^{\infty} f(x)\ dx \int_{x-\frac{\omega}{2}}^{x+\frac{\omega}{2}} g(y)\ dy = \int_{-\infty}^{\infty} f(x)\ dx \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} g(x+y)\ dy\ ,$$

$$= \int_{-\infty}^{\infty} \phi(x)\ f(x)\ dx\ ,$$

where the limit condition of (2.6) has been imposed.

Three special cases of $\phi(x)$ are of interest, the polynomial x^n, the factorial polynomial $x^{(n)} = x(x-\omega)\cdots\cdot[x-(n-1)\omega]$, and the exponential e^{tx}. Consider

$$(2.11)\qquad \phi(x) = x^n, \qquad \phi'(x+\tfrac{\omega}{2}) = n(x+\tfrac{\omega}{2})^{n-1};$$

and (2.2) takes the form

$$(2.12)\qquad \underset{\omega}{\Delta}\ g(x) = n\ \omega^{n-1}(\tfrac{x}{\omega} + \tfrac{1}{2})^{n-1}.$$

Since $\Delta B_n(x) = nx^{n-1}$ (see Milne-Thompson (1933), p. 136) it is clear that

$$(2.13)\qquad g(x) = \omega^n\ B_n(\tfrac{x}{\omega} + \tfrac{1}{2})$$

and the polynomial moment relationship is

$$(2.14)\qquad M(x^n) = \omega^n \sum_{i=-\infty}^{\infty} B_n\left(\frac{x_i}{\omega} + \frac{1}{2}\right) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy + R_1 .$$

The relationship (see Norlund (1924))

$$(2.15)\qquad B_n(x+y) = \sum_{i=0}^{n} \binom{n}{i} B_i(y)\, x^{n-i}$$

is applicable here; the polynomial moments take the form

$$(2.16)\quad M_n = \omega^n \sum_{i=-\infty}^{\infty} \sum_{s=0}^{n} \binom{n}{s} \left(\frac{x_i}{\omega}\right)^{n-s} B_s\left(\frac{1}{2}\right) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy + R_1 .$$

The multiplicative theorem for Bernoulli polynomials of order unity is of interest here. Write the defining equation (1.9), for $n = 1$, in the form

$$(2.17)\qquad \frac{te^{(x+\frac{s}{m})t}}{e^t - 1} = \sum_{k=0}^{\infty} \frac{B_k(x+\frac{s}{m})}{k!}\, t^k ,$$

then sum on the index s:

$$(2.18)\qquad \begin{aligned} \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{s=0}^{m-1} B_k\left(x+\frac{s}{m}\right) &= \frac{te^{xt}}{e^t-1} \sum_{s=0}^{m-1} e^{\frac{st}{m}} = \frac{m\frac{t}{m}\, e^{\frac{mxt}{m}}}{e^{t/m}-1} \\ &= m \sum_{k=0}^{\infty} \frac{B_k(mx)}{k!} \left(\frac{t}{m}\right)^k . \end{aligned}$$

That is,

$$(2.19) \qquad B_k(mx) = m^{k-1} \sum_{s=0}^{m-1} B_k(x+\frac{s}{m}),$$

and, for $x = 0$,

$$(2.20) \qquad B_k + \sum_{s=1}^{m-1} B_k(\frac{s}{m}) = \frac{B_k}{m^{k-1}},$$

or, finally,

$$(2.21) \qquad \sum_{s=1}^{m-1} B_k(\frac{s}{m}) = -(1 - \frac{1}{m^{k-1}}) B_k.$$

For the case at hand take $m = 2$ so (2.16) has the form (1.15) but with an additive error term. The first six relationships, without remainder term, for the polynomial moments in terms of the raw moments are listed by Cramer (1946).

Consider a similar procedure for derivation of the factorial moment relationships. Write

$$(2.22) \qquad \phi(x) = x^{(n)} = \omega^n \frac{x}{\omega}(\frac{x}{\omega} - 1)\cdots\cdots\cdots(\frac{x}{\omega} - n+1).$$

Milne-Thompson (1933), p. 130, establishes that

$$(2.23) \qquad B_n^{(n+1)}(x) = (x-1)(x-2)\cdots\cdots(x-n) = (x-1)^{(n)}$$

so

(2.24) $$\phi(x + \tfrac{\omega}{2}) = \omega^n B_n^{(n+1)}(\tfrac{x}{\omega} + \tfrac{3}{2}).$$

Another well-known property of the generalized Bernoulli polynomials is that

(2.25) $$\frac{d}{dx} B_k^{(n)}(x) = k B_{k-1}^{(n)}(x),$$

so

(2.26) $$\phi'(x + \tfrac{\omega}{2}) = \omega^{n-1} B_{n-1}^{(n+1)}(\tfrac{x}{\omega} + \tfrac{3}{2})$$

and (2.2) has the form

(2.27) $$\underset{\omega}{\Delta} g(x) = n\, \omega^{n-1} B_{n-1}^{(n+1)}(\tfrac{x}{\omega} + \tfrac{3}{2}).$$

The difference relation, corresponding to (2.25), is

(2.28) $$\Delta B_k^{(n)}(x) = k B_{k-1}^{(n-1)}(x),$$

which leads to

(2.29) $$\underset{\omega}{\Delta} g(x) = \omega^n \underset{\omega}{\Delta} B_n^{(n+2)}(\tfrac{x}{\omega} + \tfrac{3}{2})$$

or, finally,

(2.30) $$g(x) = \omega^n B_n^{(n+2)}\left(\frac{x}{\omega} + \frac{3}{2}\right).$$

The factorial moments are now defined by

$$M_{(n)} = \int_{-\infty}^{\infty} \phi(x)\ f(x)\ dx = \sum_{i=-\infty}^{\infty} g(x_i) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\ dy + R_1$$

(2.31)
$$= \omega^n \sum_{i=-\infty}^{\infty} B_n^{(n+2)}\left(\frac{x_i}{\omega} + \frac{3}{2}\right) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\ dy + R_1.$$

Another identity in Milne-Thompson (1933), p. 133, is

(2.32) $$B_k^{(n)}(x+h) = \sum_{s=0}^{k} \binom{k}{s} B_{k-s}^{(n-s)}(h)\ x^{(s)}$$

so

$$B_n^{(n+2)}\left(\frac{x_i}{\omega} + \frac{3}{2}\right) = \sum_{s=0}^{n} \binom{n}{s} B_{n-s}^{(n+2-s)}\left(\frac{3}{2}\right)\left(\frac{x_i}{\omega}\right)^{(s)}$$

$$= \sum_{s=0}^{n} \binom{n}{s} B_s^{(s+2)}\left(\frac{3}{2}\right)\left(\frac{x_i}{\omega}\right)^{(n-s)}.$$

Equation (2.31) may now be written in the form

$$M_{(n)} = \omega^n \sum_{i=-\infty}^{\infty} \sum_{s=0}^{n} \binom{n}{s} B_s^{(s+2)}\left(\frac{3}{2}\right)\left(\frac{x_i}{\omega}\right)^{(n-s)} \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)dy+R_1$$

(2.34)

(2.34) con't

$$= \sum_{s=0}^{n} \omega^s \binom{n}{s} B_s^{(s+2)}(\tfrac{3}{2}) \, \bar{M}_{(n-s)} + R_1 .$$

This is a Wold (1934) result not usually appearing in text-books. Further reduction is possible because of the identity (2.23). The familiar binomial theorem goes as

$$(2.35) \quad (1+t)^{x-1} = 1 + \sum_{n=1}^{\infty} \frac{(x-1)(x-2)\cdots(x-n)}{n!} t^n = \sum_{n=0}^{\infty} \frac{B_n^{(n+1)}}{n!} (x) t^n .$$

Differentiation with respect to x yields

$$\sum_{n=1}^{\infty} \frac{(x-1)\cdots(x-n)}{n!} \left[\frac{1}{x-1} + \frac{1}{x-2} + \cdots\cdots + \frac{1}{x-n}\right] t^n$$

$$(2.36) \qquad = \sum_{n=1}^{\infty} \frac{B_{n-1}^{(n+1)}}{(n-1)!} (x) \; t^n$$

so, at least formally,

$$B_0^{(2)}(\tfrac{3}{2}) = 1$$

$$B_1^{(3)}(\tfrac{3}{2}) = 0,$$

....

and, in general,

$$B_{n-1}^{(n+1)}(\tfrac{3}{2}) = \frac{1}{n}(\tfrac{3}{2}-1)(\tfrac{3}{2}-2)\cdots\cdots(\tfrac{3}{2}-n)[2-2-\tfrac{2}{3}-\tfrac{2}{5}-\cdots\cdots-\tfrac{2}{2n-3}] ,$$

(2.37) $n \geq 3$,

or, finally,

$$B_{n}^{(n+2)}(\tfrac{3}{2}) = (-1)^{n+1}\frac{1\cdot 1\cdot 3\cdots\cdots(2n-1)}{2^n(n+1)}(\tfrac{1}{3}+\tfrac{1}{5}+\cdots\cdots\tfrac{1}{2n-1})$$

(2.38)

$$= (-1)^{n+1}\frac{(2n)!}{2^{2n}(n+1)!}(\tfrac{1}{3}+\tfrac{1}{5}+\cdots\cdots+\tfrac{1}{2n-1}),\ n \geq 2.$$

The first six moments in terms of the raw factorial moments are listed from (2.34) and (2.38):

$$M_{(0)} = M_0 = \overline{M}_{(0)} = 1$$

$$M_{(1)} = \overline{M}_{(1)} ,$$

$$M_{(2)} = \overline{M}_{(2)} - \frac{\omega^2}{12} ,$$

$$M_{(3)} = \overline{M}_{(3)} - \frac{\omega^2}{4}\overline{M}_{(1)} + \frac{\omega^3}{4} ,$$

$$M_{(4)} = \overline{M}_{(4)} - \frac{\omega^2}{2}\overline{M}_{(2)} + \omega^3\,\overline{M}_{(1)} - \frac{71}{80}\omega^4 ,$$

$$M_{(5)} = \overline{M}_{(5)} - \frac{5\omega^2}{6}\,\overline{M}_{(3)} + \frac{5\omega^3}{2}\,\overline{M}_{(2)} - \frac{71\omega^4}{16}\,\overline{M}_{(1)} + \frac{31\omega^5}{8} ,$$

$$M_{(6)} = \overline{M}_{(6)} - \frac{5\omega}{4}\,\overline{M}_{(4)} + 5\omega^3\,\overline{M}_{(3)} - \frac{213\omega^4}{16}\,\overline{M}_{(2)} + \frac{93\omega^5}{4}\,\overline{M}_{(1)}$$

$$- \frac{9129\ ^{6}}{448} ,$$

. . . .

The third choice of function is

$$\phi(x) = e^{tx} , \tag{2.39}$$

and from (2.2)

$$\underset{\omega}{\Delta}\, g(x) = \phi'(x+\tfrac{\omega}{2}) = t\, e^{t(x+\frac{\omega}{2})} \tag{2.40}$$

To take

$$g(x) = \frac{\omega t\, e^{t(x+\frac{\omega}{2})}}{e^{\omega t} - 1} \tag{2.41}$$

is consistent with (2.40). As in the earlier discussions the idea is to develop a general formula which will give the true moments in terms of the raw moments evaluated by measurement in some kind of field operation. Implementation of (1.16) leads to

$$M(\phi) = \sum_{i=-\infty}^{\infty} g(x_i) \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy + R_1$$

$$(2.42) \qquad = \frac{\omega t\, e^{\frac{\omega t}{2}}}{e^{\omega t} - 1} \sum_{i=-\infty}^{\infty} e^{tx_i} \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy + R_1$$

$$= \frac{\omega t\, e^{\frac{\omega t}{2}}}{e^{\omega t} - 1} \overline{M}[e^{tx}] + R_1.$$

The semi-invariants $\{\lambda_n\}$ are defined by

$$(2.43) \qquad M[e^{tx}] = \exp\left(\sum_{n=1}^{\infty} \frac{\lambda_n}{n!} t^n\right) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

and the raw semi-invariants by

$$(2.44) \quad \overline{M}[e^{tx}] = \exp\left(\sum_{n=1}^{\infty} \frac{\overline{\lambda}_n}{n!} t^n\right) = \sum_{i=-\infty}^{\infty} e^{tx_i} \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i+y)\, dy.$$

Let the remainder term be considered negligible so a very simple relationship is

$$M[e^{tx}] = \frac{\omega t \, e^{\frac{\omega t}{2}}}{e^{\omega t} - 1} \bar{M}[e^{tx}]. \tag{2.45}$$

Logarithmic differentiation leads to

$$\begin{aligned} \frac{\frac{d}{dt} M[e^{tx}]}{M[e^{tx}]} - \frac{\frac{d}{dt} \bar{M}[e^{tx}]}{\bar{M}[e^{tx}]} &= \frac{1}{t} + \frac{\omega}{2} - \frac{\omega \, e^{\omega t}}{e^{\omega t}-1} \\ &= \frac{1}{t} - \frac{\omega}{2} - \frac{\omega t}{t(e^{\omega t} - 1)} \end{aligned} \tag{2.46}$$

To implement the indicated differentiation is easy, so (2.46) reduces to

$$\sum_{n=1}^{\infty} (\lambda_n - \bar{\lambda}_n) \frac{t^{n-1}}{(n-1)!} = \frac{1}{t} - \frac{\omega}{2} - \frac{1}{t} \frac{\omega t}{e^{\omega t} - 1} \tag{2.47}$$

The quotient on the right is at once the generating function for the Bernoulli numbers, the case $n = 1$ in (1.10). Recalling that $B_0 = 1$, $B_1 = -\frac{1}{2}$, (2.47) becomes

$$\sum_{n=1}^{\infty} (\lambda_n - \bar{\lambda}_n) \frac{t^{n-1}}{(n-1)!} = -\omega \sum_{n=2}^{\infty} \frac{B_n}{n!} (\omega t)^{n-1}. \tag{2.48}$$

The implication is

$$\lambda_1 = \bar{\lambda}_1 \tag{2.49}$$

and

$$\lambda_n = \bar{\lambda}_n - \frac{B_n \omega^n}{n}, \quad n \geq 2. \tag{2.50}$$

Since the odd order Bernoulli numbers are all zero, exception noted above, the final statement goes as

$$\begin{aligned} \lambda_{2n+1} &= \bar{\lambda}_{2n+1}, \quad n \geq 0, \\ \lambda_{2n} &= \bar{\lambda}_{2n} - \frac{B_{2n} \omega^{2n}}{2n}, \quad n \geq 1. \end{aligned} \tag{2.51}$$

This result is due to Langdon and Ore (1929).

Chapter III

Error terms

In a recent paper Gould and Squire (1967) have exploited the fact that there is a second Euler-Maclaurin summation formula, and that a general form exists such that both the first and second formulas are special cases. Their point is that the remainder term for the one is quite likely to be of opposite sign to that of the other, thus forming a bracket for the true result of some summation or integration. They seemed to be unaware that Hildebrand (1956) had rather extensively discussed the two forms in a readilty accessible textbook, although they did mention a few such names as Jacobi, Darboux, Lindelöf, etc. In a fashion quite similar to the estimation of error for a Taylor series representation of a function Hildebrand writes the error estimate

$$E_N = r \frac{B_{2N+2}\omega^{2N+2}}{(2N+2)!} f^{(2N+2)}(\xi), \quad a < \xi < a+r\omega, \tag{3.1}$$

to go with the first form of the summation formula,

$$\sum_{k=0}^{r} f_k = \frac{1}{\omega}\int_{x_0}^{x_r} f(x)\,dx + \frac{f_0 + f_r}{2} \tag{3.2}$$

$$+ \sum_{n=1}^{N} \frac{B_{2n}}{(2n)!}\,\omega^{2n-1}[f_r^{(2n-1)} - f_0^{(2n-1)}] + E_N ,$$

where $f_k = f(a+k\omega)$, $x_0 = a$, $x_r = a + r\omega$. The second form goes as

$$(3.3) \qquad \sum_{k=0}^{r-1} f_{k+\frac{1}{2}} = \frac{1}{\omega}\int_{x_0}^{x_r} f(x)\,dx$$

$$- \sum_{n=1}^{N} \frac{(1-2^{1-2n})B_{2n}}{(2n)!}\,\omega^{2n-1}[f_r^{(2n-1)} - f_0^{(2n-1)}] + E_N,$$

where

$$(3.4) \qquad E_N = -r\frac{(1-2^{-1-2n})B_{2N+2}}{(2N+2)!}\omega^{2N+2} f^{(2N+2)}(\xi), \quad a<\xi<a+r\omega$$

Note the opposite algebraic signs. In both situations the uncertainty lies in the location of ξ. Knopp (1947) introduces a similar form for the estimate of error for the first formula; by a rather elaborate procedure the uncertainty is introduced by a positive fractional factor,

$$(3.5) \qquad E_N = \theta\frac{B_{2N+2}\,\omega^{2N+2}}{(2N+2)!}[f_r^{(2N+1)} - f_0^{(2N+1)}], \quad 0 < \theta < 1.$$

Both textbooks offer much information and advice on how to safely estimate the error.

In the case at hand it is the Cramer (1946) treatment which is of greatest interest. Construct an identity in terms of the periodic function $P_1(x)$, period unity, of (2.8):

$$\omega\int_k^{k+1} P_1(x)g'(a+\omega x)dx = P_1(x)g(a+\omega x)\Big|_k^{k+1} + \int_k^{k+1} g(a+\omega x)dx$$

(3.6)

$$= -\tfrac{1}{2}g[a+(k+1)\omega] - \tfrac{1}{2}g(a+k\omega) + \int_k^{k+1} g(a+\omega x)dx.$$

Summation from $k = n_1$ to $k = n_2-1$ yields

$$\sum_{k=n_1}^{n_2} g(a+k\omega) = \int_{n_1}^{n_2} g(a+\omega x)dx + \tfrac{1}{2}g(a+n_1\omega) + \tfrac{1}{2}g(a+n_2\omega)$$

(3.7)

$$- \omega\int_{n_1}^{n_2} P_1(x)g'(a+\omega x)dx.$$

If n_1 and n_2 are allowed to become arbitrarily large, left and right, Cramer's (12.2.5) takes the form

(3.8) $$\sum_{-\infty}^{\infty} g(a+k\omega) = \int_{-\infty}^{\infty} g(a+\omega x)dx - \omega\int_{-\infty}^{\infty} P_1(x)g'(a+\omega x)dx,$$

since $g(x)$ must vanish at the extremities of the x-axis.

A special case discussed by Cramer is for $a = 0$, $\omega = 1$, $n_1 = 1$, $n_2 = n$, $g(x) = \frac{1}{x}$.

$$\sum_{k=1}^{n} \frac{1}{k} = \int_1^n \frac{dx}{x} + \frac{1}{2} + \frac{1}{2n} + \int_1^n \frac{P_1(x)}{x^2}\, dx.$$

From the definition of $P_1(x)$ it is clear that

$$0 < \int_n^\infty \frac{P_1(x)}{x^2}\, dx < \frac{1}{n^2}\int_n^\infty P_1(x)\,dx < \frac{1}{8n^2},$$

The $1/8$ is the area of the triangle between an integer k and $k + \frac{1}{2}$. If the Euler constant may be defined as

$$C = \frac{1}{2} + \int_1^\infty \frac{P_1(x)}{x^2}\,dx = 0.5772...$$

then

$$\sum_{k=1}^{n} \frac{1}{k} = \ln n + C + \frac{1}{2n} + \mathcal{O}(\frac{1}{n^2}).$$

Chapter IV

Taylor Series Approach

In this chapter emphasis is placed on the derivation of relationships between the true and the raw moments of a given distribution by means of Taylor series expansions. Formally the definition of (1.4) may be written as

$$A_i = \int_{x_i - \frac{\omega}{2}}^{x_i + \frac{\omega}{2}} f(x)\,dx = \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} f(x_i + y)\,dy$$

(4.1)

$$\sum_{j=0}^{\infty} \frac{f^{(j)}(x_i)}{j!} \int_{-\frac{\omega}{2}}^{\frac{\omega}{2}} y^j\,dy = \omega \sum_{j=0}^{\infty} \frac{(\frac{\omega}{2})^{2j}}{(2j+1)!} f^{(2j)}(x_i) .$$

The raw moments $\overline{M}_n$ of (1.3) may now be written as

$$\overline{M}_n = \sum_{i=-\infty}^{\infty} x_i^n A_i$$

(4.2)

$$= \omega \sum_{j=0}^{\infty} \frac{(\frac{\omega}{2})^{2j}}{(2j+1)!} \sum_{i=-\infty}^{\infty} x_i^n\, f^{(2j)}(x_i) ,$$

$$n = 0,1,2,\ldots \quad .$$

The interchange of order of summation must be justified in each case. The density function $f(x)$ is assumed to satisfy the condition (1.5). The Euler-Maclaurin identity (1.6) leads to

$$\sum_{i=-\infty}^{\infty} x_i^n \, f^{(2j)}(x_i) = \int_{-\infty}^{\infty} x^n \, f^{(2j)}(x)\,dx + R. \tag{4.3}$$

Assume that R is negligible; it is easily established that

$$\int_{-\infty}^{\infty} x^n f^{(2j)}(x)\,dx = \begin{cases} 0, & 2j > n, \\ n(n-1)\ldots(n-2j+1)\int_{-\infty}^{\infty} x^{n-2j} f(x)\,dx, & \end{cases}$$

$$= \begin{cases} 0, & 2j > n, \\ \dfrac{n!}{(n-2j)!}\, M_{n-2j}, & 2j \le n, \quad n = 0,1,2,\ldots\,. \end{cases} \tag{4.4}$$

From (4.2) and (4.4) the set of raw moments in terms of the true moments is

$$\overline{M}_n = \sum_{j=0}^{[\frac{n}{2}]} \left(\frac{\omega}{2}\right)^{2j} \frac{n!}{(2j+1)!\,(n-2j)!}\, M_{n-2j}$$

$$= \frac{1}{n+1} \sum_{j=0}^{[\frac{n}{2}]} \left(\frac{\omega}{2}\right)^{2j} \binom{n+1}{2j+1} M_{n-2j}, \quad n = 0,1,2,\ldots\,, \tag{4.5}$$

the same as (1.4). Note that the inversion of this triangular system of linear equations is not at all an obvious procedure.

The special case of interest is the normal density function,

$$\text{(4.6)} \qquad f(x) = \phi^{(0)}(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}, \quad -\infty < x < \infty.$$

The well-known table of the error function and its first twenty derivatives, published by the Harvard Computation Laboratory (1966), is indicated for choice of notation. That is,

$$\text{(4.7)} \qquad \phi^{(n)}(x) = (-1)^n \phi^{(0)}(x)\ He_n(x),$$

and this set of Hermite polynomials takes the general form

$$\text{(4.8)} \qquad He_n(x) = n! \sum_{k=0}^{[\frac{n}{2}]} \frac{(-1)^k\ x^{n-2k}}{2^k k!(n-2k)!}.$$

Because of the symmetry present only even order polynomials will appear and it is convenient to write

$$He_{2n}(x) = \frac{(-1)^n(2n)!}{2^n} \sum_{k=0}^{n} \frac{(-1)^k(2x^2)^k}{(2k)!(n-k)!} \; .$$

Special cases are $He_0(x) = 1$, $He_2(x) = x^2 - 1$, $He_4(x) = x^4 - 6x^2 + 3$, $He_6(x) = x^6 - 15x^4 + 45x^2 - 15, \ldots\ldots$.

For this normal density function the set of true moments is

$$(4.9) \qquad M_n = \int_{-\infty}^{\infty} x^n \phi^{(0)}(x)\,dx = \begin{cases} 0, & n \text{ odd}, \\ 2\int_0^{\infty} x^n \phi^{(0)}(x)\,dx, & n \text{ even}. \end{cases}$$

It is easily established that

$$(4.10) \qquad M_{2n} = 2\int_0^{\infty} x^{2n}\phi^{(0)}(x)\,dx = \frac{(2n)!}{2^n n!} = (-1)^n He_{2n}(0),$$

$$n = 0,1,2,\ldots \quad .$$

The raw moments depend on the ω interval,

$$(4.11)\quad \begin{aligned}\overline{M}_{2n} &= \sum_{j=0}^{n} \left(\frac{\omega}{2}\right)^{2j} \frac{(2n)!}{(2j+1)!\,(2n-2j)!} M_{2n-2j} \\ &= \frac{(2n)!}{2^n} \sum_{j=0}^{n} \left(\frac{\omega}{2}\right)^{2j} \frac{1}{(2j+1)!\,(n-j)!} \, .\end{aligned}$$

It is instructive to consider the Laplace transform pair:

$$(4.12)\qquad s^{n-\frac{1}{2}} e^{-a\sqrt{s}} \qquad \frac{1}{\sqrt{\pi t}\,(2t)^n}\, e^{-\frac{a^2}{4t}} He_{2n}\left(\frac{a}{\sqrt{2t}}\right) ,$$

$$a,t > 0,\ n = 0,1,2,\ldots \quad .$$

(see Erdelyi, et. al. (1954)). In order to establish the Theta function identity which is of interest the case $a = 0$ may be included. Also, let $n = 0$ be excluded. One transform pair is defined in terms of the sum of residues,

$$(4.13)\qquad s^{n-\frac{1}{2}} \frac{\cosh\sqrt{s}}{\sinh\sqrt{s}} \qquad \sum_{r=1}^{\infty} R_r \, ,$$

where

$$(4.14) \qquad R_r = \lim_{s_r \to -r^2\pi^2} 2s^n e^{st} = 2(-r^2\pi^2)^n e^{-r^2\pi^2 t}.$$

Also,

$$(4.15) \qquad s^{n-\frac{1}{2}} \frac{\cosh\sqrt{s}}{\sinh\sqrt{s}} = s^{n-\frac{1}{2}} \frac{1 + e^{-2\sqrt{s}}}{1 - e^{-2\sqrt{s}}}$$

$$= s^{n-\frac{1}{2}}[1 + 2 \sum_{r=1}^{\infty} e^{-2r\sqrt{s}}],$$

so a second entry in the table is

$$(4.16) \qquad s^{n-\frac{1}{2}} \frac{\cosh\sqrt{s}}{\sinh\sqrt{s}} \quad \frac{1}{\sqrt{\pi t}(2t)^n} [(-1)^n M_{2n}$$

$$+ 2 \sum_{r=1}^{\infty} e^{-r^2/t} He_{2n}(r\sqrt{\tfrac{2}{t}})].$$

For the special $f(x)$ under consideration (4.2) may be written as

$$(4.17)\quad \overline{M}_{2n} = 2\omega \sum_{j=0}^{\infty} \frac{(\frac{\omega}{2})^{2j}}{(2j+1)!} \sum_{i=1}^{\infty} x_i^{2n} \phi^{(0)}(x_i) He_{2j}(x_i)$$

$$= 2\omega \sum_{j=0}^{\infty} \frac{(\frac{\omega}{2})^{2j}}{(2j+1)!} \frac{(-1)^j (2j)!}{2^j} \sum_{k=0}^{j} \frac{(-1)^k 2^k}{(2k)!(j-k)!} \sum_{i=1}^{\infty} x_i^{2n+2k} \phi^{(0)}(x_i).$$

Note that because of symmetry $x_0 = 0$. To use the Theta function identity it is essential that $x_i = i$, $i = 1,2,\cdots$, meaning that the symmetrically spaced intervals are of unit length. It is formally correct to write

$$2(-1)^n \pi^{2n} \sum_{r=1}^{\infty} r^{2n} e^{-r^2\pi^2 t}$$

(4.18)

$$= \frac{1}{\sqrt{\pi t}(2t)^n}[(-1)^n M_{2n} + 2 \sum_{r=1}^{\infty} e^{-r^2/t} He_{2n}(r\sqrt{\tfrac{2}{t}})]$$

and if $2\pi^2 t = 1$ then

$$(4.19)\quad 2 \sum_{r=1}^{\infty} r^{2n} \phi^{(0)}(r) = M_{2n} + 2(-1)^n \sum_{r=1}^{\infty} e^{-2\pi^2 r^2} He_{2n}(2\pi r).$$

Equation (4.17) now has the form

$$\overline{M}_{2n} = \sum_{j=0}^{\infty} \frac{(-1)^j}{2^{3j}(2j+1)} \sum_{k=0}^{j} \frac{2^k}{(2k)!(j-k)!} [M_{2n}$$

(4.20)

$$+ 2(-1)^n \sum_{r=1}^{\infty} e^{-2\pi^2 r^2} He_{2n+2k}(2\pi r)].$$

The double summation before M_n has the value of a certain simple definite integral:

$$\sum_{j=0}^{\infty} \frac{(-1)^j}{2^{3j}(2j+1)} \sum_{k=0}^{j} \frac{2^k}{(2k)!(j-k)!}$$

$$= \sum_{k=0}^{\infty} \frac{2^k}{(2k)!} \sum_{j=k}^{\infty} \frac{(-1)^j}{2^{3j}(2j+1)(j-k)!}$$

(4.21)

$$= \sum_{k=0}^{\infty} \frac{2^k}{(2k)!} \sum_{r=0}^{\infty} \frac{(-1)^{r+k}}{2^{3(r+k)}(2r+2k+1)r!}$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{2^{2k}(2k)!} \int_0^1 x^{2k} e^{-x^2/8} dx$$

$$= \int_0^1 \cos\frac{x}{2} e^{-x^2/8} dx = 0.9222\cdots ,$$

where the decimal approximation has been obtained by the use of Simpson's rule. So, finally,

$$A_0 M_{2n} = \bar{M}_{2n} - 2 \sum_{j=0}^{\infty} \frac{(-1)^{j+n}}{2^{3j}(2j+1)} \sum_{k=0}^{j} \frac{2^k}{(2k)!(j-k)!} \sum_{r=1}^{\infty} e^{-2\pi^2 r^2}$$

$$He_{2n+2k}(2\pi r), \qquad (4.22)$$

where A_0 is the decimal approximation obtained above.

BIBLIOGRAPHY

Cramer, Harald. 1958. Mathematical methods of statistics. Princeton, Princeton University. 575 p.

Erdelyi, Arthur, W. Magnus, F. Oberhettinger and F.G. Tricomi. 1954. Tables of integral transforms. New York, McGraw-Hill. 388 p. (California Institute of Technology. Bateman Manuscript Project. Vol. 1).

Gould, H.W. and W. Squire. 1967. Maclaurin second formula and its generalization. American Mathematical Monthly 74: 44-52.

Hildebrand, F.B. 1956. Introduction to numerical analysis. New York, McGraw-Hill. 511 p.

Kendall, M.G. 1938. The conditions under which Sheppard's correction are valid. Journal of the Royal Statistical Society 101: 592-604.

Knopp, Konrad. 1947. Theory and applications of infinite series. New York, Hafner. 563 p.

Langdon, W.H. and O. Ore. 1930. Semi-invariants and Sheppard's correction. Annals of Mathematics, ser. 2, 31: 230-232.

Milne-Thompson, L.M. 1933. The calculus of finite differences. London, Macmillan. 558 p.

Norlund, Niels Erik. 1924. Vorlesungen über Differinzenrechnung. Berlin, Springer. 551 p.

Richardson, C.H. 1954. An Introduction to the calculus of finite differences. Princeton, New Jersey, Van Nostrand. 142 p.

Todd, John. 1963. Introduction to the constructive theory of functions. New York, Academic. 127 p.

Wilks, Samuel. 1963. Mathematical statistics. New York, Wiley. 644 p.

Wold, Herman. 1933. Sheppard's correction formulae in several variables. Skandinavisk Aktuarietidskrift 16: 248-254.

Wold, Herman. 1934. Sulla correzione di Sheppard. Giornale Italiano delgi Attuari 5: 304.