ON THE EFFECT OF DECIMAL CORRECTIONS ON ERRORS OF OBSERVATION

BY PHILIP HARTMAN AND AUREL WINTNER

The Johns Hopkins University

1. Summary. Let t be the true value of what is being measured and suppose that the error of observation is a symmetric normal distribution of standard deviation σ. The "rounding-off" error due to the reading of measurements to the nearest unit has a distribution and an expected value depending on t and σ. It is shown that, for a fixed $\sigma > 0$, the expected value of the decimal correction, $r(t; \sigma)$, is an analytic function of t which is odd, of period 1, positive for $0 < t < \frac{1}{2}$, and has a convex arch as its graph on $0 \leqq t \leqq \frac{1}{2}$. Furthermore, if $0 < t < \frac{1}{2}$, both $r(t; \sigma)$ and its maximum value, $\operatorname{Max}_t r(t; \sigma)$, are decreasing functions of σ.

2. Introduction. Let X be an error of observation and let $\phi(x)$ denote the density of probability of the distribution of X. In particular,

$$\int_{-\infty}^{+\infty} \phi(x)\, dx = 1, \qquad \text{where } \phi(x) \geqq 0. \tag{1}$$

If t is any fixed number, the density of probability of the distribution of $X + t$ is $\phi(x - t)$.

Besides the "instrumental error of observation", X, there is another error, that of the "rounding-off", which is carried along in the registration of the measurements. It is introduced by the circumstance that, if $\cdots, b, a$ are digits, and if b denotes the last digit considered, then decimal fractions such as $\cdots ba$ and $\cdots ba \cdots$ are registered as $\cdots b$ if $a < 5$ and as $\cdots (b + 1)$ if $a > 5$. Let the unit, in which the measurements are expressed, be so chosen that the first digit neglected becomes the first digit following the decimal point, i.e., that the error of the "rounding-off" is between $\pm\frac{1}{2}$. Then, if t denotes the *true* value of what is being measured, the remark made after (1) shows that *the probability that the error of the decimal corrections be less than* x is given by

$$\sum_{n=-\infty}^{\infty} \int_{n-\frac{1}{2}}^{n-\frac{1}{2}+x} \phi(u - t)\, du,$$

if $|x| \leqq \frac{1}{2}$, whereas this probability is 0 or 1 according as $x < -\frac{1}{2}$ or $x > \frac{1}{2}$. Since the last series can be written in the form

$$\sum_{n=-\infty}^{\infty} \int_{-\frac{1}{2}}^{-\frac{1}{2}+x} \phi(u + n - t)\, du = \int_{-\frac{1}{2}}^{-\frac{1}{2}+x} \sum_{n=-\infty}^{\infty} \phi(u + n - t)\, du, \qquad (\phi \geqq 0), \tag{2}$$

it follows that the density of probability of the error due to the decimal corrections is

$$\sum_{n=-\infty}^{\infty} \phi(x + n - t) \text{ if } |x| < \tfrac{1}{2}, \text{ and } 0 \text{ if } |x| > \tfrac{1}{2}. \tag{3}$$

Consequently, *if* $r = r(t)$ *denotes the expected value of the decimal error induced on the "true" value*, t, *of the observations, then*

$$(4) \qquad r(t) = \int_{|x|<\frac{1}{2}} x \sum_{n=-\infty}^{\infty} \phi(x + n - t)\, dx.$$

Formula (4) is known[1]. It is usually based on its intuitive interpretation which results if, on the one hand, (4) is written in the form

$$(5) \qquad r(t) = \int_{-\infty}^{\infty} s(x)\phi(x - t)\, dx,$$

where

$$(6) \qquad s(x) = x \text{ if } -\tfrac{1}{2} < x < \tfrac{1}{2} \text{ and } s(x) = s(x + 1), \qquad -\infty < x < \infty,$$

and, on the other hand, the periodic function (6) is thought of as representing the *uniform* distribution of the error of "rounding-off" over the *arithmetical continuum* over a period,

$$|x - n| < \tfrac{1}{2}, \qquad (n = 0, \pm 1, \cdots),$$

on the x-axis. Needless to say, the specification of $s(x)$ at the points $x = n + \frac{1}{2}$, which are disregarded in the definition (6), is immaterial, since $s(x)$ occurs in (5) only as an integrable weight-factor, isolated values of which do not influence the integral.

It follows at once from (1), (5) and the continuity (almost everywhere) of (6), that $r(t)$ is continuous.

3. Fourier analysis of $r(t)$. Since the Fourier expansion of the periodic function (6) is

$$(7) \qquad s(x) = -\pi^{-1} \sum_{n=1}^{\infty} (-1)^n n^{-1} \sin 2\pi n x = s(x \pm 1) = \cdots, \qquad (|x| < \tfrac{1}{2}),$$

it follows from (5) that[2]

$$(8) \qquad r(t) = -\pi^{-1} \sum_{n=1}^{\infty} (-1)^n n^{-1} \int_{-\infty}^{\infty} \phi(x) \sin 2\pi n(x + t)\, dx.$$

Hence, if the sine in (8) is expressed in terms of $2\pi nx$ and $2\pi nt$,

$$(9) \qquad \pi r(t) = -\sum_{n=1}^{\infty} (-1)^n n^{-1}(a_n \cos 2\pi nt + b_n \sin 2\pi nt),$$

[1] F. Zernike, "Wahrscheinlichkeitsrechnung und mathematische Statistik, "*Handbuch der Physik*, Vol. 3 (1928), pp. 475–476.

[2] In view of (1), the term-by-term integration leading from (5) to (8) is justified by the fact that the partial sums of the series (7) are uniformly bounded. Correspondingly, the above deduction of (9) and (10) from (4) is equivalent to an application of Poisson's summation formula. In this regard, cf. A. Wintner, "The sum formulae of Euler-Maclaurin and the inversions of Fourier and Möbius," *Am. Jour. of Math.*, Vol. 69 (1947), pp. 685–708, the end of §1 (p. 687) and its application on p. 697.

where

$$(10) \qquad b_n + ia_n = \int_{-\infty}^{\infty} \phi(x) \exp(2\pi inx)\, dx, \qquad (n = 1, 2, \cdots).$$

Let it be assumed that positive and negative errors of observation, when of the same magnitude, are equally probable, i.e., that $\phi(x) = \phi(-x)$. Then (10) shows that a_n becomes 0. Hence, (9) reduces to

$$(11) \qquad r(t) = -\sum_{n=1}^{\infty} (-1)^n (c_n/n) \sin 2\pi nt,$$

where

$$(12) \qquad c_n = \pi^{-1} \int_{-\infty}^{\infty} \phi(x) \cos 2\pi nx\, dx = 2\pi^{-1} \int_0^{\infty}.$$

Clearly, $r(t)$ is an odd function whenever the density $\phi(x)$ is even.

4. The normal case. Suppose that $\phi(x)$ is the density of a symmetric *normal* (Gaussian) distribution. Then, if σ is the positive constant representing the standard deviation of the errors of observation,

$$(13) \qquad \phi(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}x^2/\sigma^2) \qquad (0 < \sigma < \infty).$$

It is clear from (5) and (6) that

$$(14) \qquad r(t) \to s(t) \text{ if } \sigma \to 0 \text{ in } (13).$$

Actually, all that (14) says is a triviality, according to which the total error becomes the decimal error when the measurements become infinitely sharp. In this limiting case, that is, if $r(t) = s(t)$, it is seen from (6) that the graph of the periodic function $r = r(t)$ is piecewise linear, and therefore discontinuous.

If $\sigma = 0$ is replaced by $0 < \sigma < \infty$, the jumps of $r(t)$ at $t = n - \frac{1}{2}$ disappear (cf. the end of §3) and, as will be proved below,

(I) *$r(t)$ is an analytic function which is odd, of period* 1, *and positive for* $0 < t < \frac{1}{2}$ (*hence negative for* $-\frac{1}{2} < t < 0$), *and*

(II) *the graph of $r = r(t)$ over the fundamental interval* $0 \leqq t \leqq \frac{1}{2}$ *is a convex arch, no matter what the value of σ in* (13) *may be.*

Since r now depends both on the "true" value, t, of the observations and the "precision", σ, of the measurements, let r be denoted by $r(t; \sigma)$. It will be shown that

(i) Max $r(t; \sigma)$, *where the* Max *refers to t while σ is fixed, is a decreasing function of σ*, where σ varies on the half-line $0 < \sigma < \infty$; and that, on the same half-line,

(ii) *$r(t; \sigma)$ is a decreasing function of σ at every fixed t contained in the fundamental region* $0 < t < \frac{1}{2}$.

All of this seems to be clear for physical reasons. Actually, it is easy to give examples of distribution laws distinct from (13) for which the above assertions become false.

5. The ϑ_3-function. As is well-known,

$$\int_{-\infty}^{\infty} \exp\left(-\tfrac{1}{2}x^2/\sigma^2\right) \cos ux \, dx = (2\pi\sigma^2)^{\frac{1}{2}} \exp\left(-\tfrac{1}{2}\sigma^2 u^2\right).$$

Hence, the value of the integral (12) is q^{n^2}, if q is an abbreviation for

$$q = \exp(-2\pi^2\sigma^2). \tag{15}$$

Consequently, if $r(t, q)$ is defined, in terms of the above $r(t; \sigma)$, by placing

$$r(t, q) = r(t; \sigma) \text{ in virtue of (15),} \tag{16}$$

then (11) shows that[3]

$$r(t, q) = -\pi^{-1} \sum_{n=1}^{\infty} (-1)^n n^{-1} q^{n^2} \sin 2\pi nt \tag{17}$$

It will be noted that the range, $0 < \sigma < \infty$, of the standard deviation is mapped by (15) on the range

$$0 < q < 1, \tag{18}$$

and that σ decreases or increases according as q increases or decreases.

Let partial differentiations with respect to t and q be denoted by primes and subscripts, respectively:

$$f' = \partial f/\partial t, \qquad f_q = \partial f/\partial q. \tag{19}$$

Thus, from (17),

$$r'(t, q) = -2 \sum_{n=1}^{\infty} (-1)^n q^{n^2} \cos 2\pi nt \tag{20}$$

and, as easily verified from (17),

$$r_q(t, q) = (-4\pi q)^{-1} r''(t, q). \tag{21}$$

Let $\theta(t, q)$ be defined by

$$\theta(t, q) = 1 + 2 \sum_{n=1}^{\infty} q^{n^2} \cos nt \tag{22}$$

(so that $\theta(t, q)$ is, in the main, the elliptic theta-function usually denoted by ϑ_3). It is known that

$$\theta(t, q) > 0 \tag{23}$$

and that[4]

$$\theta'(t, q) < 0 \text{ if } 0 < t < \pi \qquad (\text{hence, } \theta'(t, q) > 0 \text{ if } -\pi < t < 0). \tag{24}$$

The above assertions will be deduced from these facts.

---

[3] Cf. F. Zernike, loc. cit.

[4] For a simple proof, cf. A. Wintner, "On the shape of the angular case of Cauchy's distribution," *Annals of Math. Stat.*, Vol. 18 (1948), pp. 589–593, §6.

6. Proof of (I)–(II) and (i)–(ii). First, it is seen from (17) and (22) that

$$r'(t, q) = 1 - \theta(2\pi t - \pi, q). \tag{25}$$

Hence,

$$r''(t, q) = -2\pi\theta'(2\pi t - \pi, q). \tag{26}$$

If (26) is compared with (24), it is seen that

$$r''(t, q) < 0 \text{ if } 0 < t < \tfrac{1}{2} \qquad (\text{hence, } r''(t, q) > 0 \text{ if } -\tfrac{1}{2} < t < 0). \tag{27}$$

Consequently, (I) and (II) follow, since, in view of (17),

$$r(\pm\tfrac{1}{2}, q) = 0 = r(0, q). \tag{28}$$

Next, (21) and (27) imply that

$$r_q(t, q) > 0 \text{ for } 0 < t < \tfrac{1}{2}. \tag{29}$$

Hence, (ii) follows from the fact that q is a decreasing function of σ.

As to (i), let $t = t(q)$ denote that (unique) t-value on $0 < t < \frac{1}{2}$ at which $r(t, q)$ assumes its maximum value, say r^q; so that

$$r^q = r(t(q), q), \qquad (0 < t(q) < \tfrac{1}{2}). \tag{30}$$

Clearly, $t = t(q)$ is the only t-value on $0 < t < \frac{1}{2}$ for which

$$r'(t, q) = 0. \tag{31}$$

Since $r'(t, q)$ possesses continuous partial derivatives with respect to t and q, and since (27) implies that its partial derivative with respect to t, namely, $r''(t, q)$, does not vanish at $t = t(q)$, it follows that the solution $t = t(q)$ of the equation (31) possesses a continuous derivative. Hence, the function (30) possesses a continuous derivative with respect to q, namely,

$$\frac{dr^q}{dq} = r'(t(q), q)\,\frac{dt(q)}{dq} + r_q(t(q), q). \tag{32}$$

But since $t = t(q)$ is a solution of (31), the identity (32) can be reduced to

$$\frac{dr^q}{dq} = r_q(t(q), q), \qquad (0 < t(q) < \tfrac{1}{2}).$$

Consequently, (i) follows from (29), since q is a decreasing function of σ.