ON THE DISTRIBUTION OF ERRORS IN N^{th} TABULAR DIFFERENCES[1]

By Arnold N. Lowan and Jack Laderman

In the construction of mathematical tables, a frequent method of checking the computed values of the tabulated function is to apply a differencing test. This test consists of computing the tabular differences of some suitable order n and comparing them with the theoretical values of the differences computed to a higher degree of accuracy by an independent method. Whenever the absolute deviation of a tabular difference $\bar{\Delta}^n$ from the corresponding theoretical difference Δ^n exceeds some predetermined upper bound, the entries giving rise to the difference in question are investigated. Thus, in the computation of the functions $Si(x)$, $Ci(x)$, $Ei(x)$ and $Ei(-x)$, it was found desirable to check the final manuscript by comparing the tabular second differences with the values of the second differences computed to a higher degree of accuracy by an independent method.[2]

A study of the distribution of errors suggested the following problem: If we assume a rectangular distribution of the errors in the entries of a mathematical table, what is the distribution of errors in the n^{th} tabular differences?

For the sake of mathematical simplicity, it will be convenient to idealize the problem as follows: Consider $n + 1$ randon numbers $x_0, x_1, x_2, \cdots x_n$, drawn from any rectangular distribution. When arranged in a definite order, these $n + 1$ values give rise to an n^{th} difference Δ^n. If these $n + 1$ numbers are rounded to k decimal places, the new approximate values $\bar{x}_0, \bar{x}_1, \cdots \bar{x}_n$, will give rise to another n^{th} difference $\bar{\Delta}^n$.

We shall investigate the distribution of the error $\Delta^n - \bar{\Delta}^n$.

The explicit expression for $\Delta^n - \bar{\Delta}^n$ is given by:

$$\Delta^n - \bar{\Delta}^n = C_0^n E_n - C_1^n E_{n-1} + C_2^n E_{n-2} - \cdots + (-1)^n C_n^n E_0$$
$$= w_0 + w_1 + w_2 + \cdots + w_n (\text{say})$$

where $E_i = x_i - \bar{x}_i$ and C_r^n are the binomial coefficients.

---

[1] The results reported in this paper were obtained in the course of work done by the Project for the Computation of Mathematical Tables. O. P. No. 765-97-3-10. Work Projects Administration, N. Y. C., operated under the sponsorship of Dr. Lyman J. Briggs, Director of the National Bureau of Standards. The authors wish to express their appreciation to the W. P. A. and to the Sponsor of this Project for permission to publish these results.

[2] The above functions were computed for $x = 0(0.0001)2.0000$ to 9 places of decimals and $x = 2(0.001)10.000$ to 9 decimals or significant figures. For the independent method of computation of the second differences, see article by A. N. Lowan in the *Bulletin of the American Mathematical Society*, August, 1939.

The distribution of any one of the E's is

$$f(E)\,dE = \begin{cases} 10^k\,dE & \text{if} \quad -\tfrac{1}{2}10^{-k} \le E \le \tfrac{1}{2}10^{-k} \\ 0 & \text{elsewhere.} \end{cases}$$

The subsequent developments are based on the fundamental theorem which states that the characteristic function of the distribution of the sum of any number of random variables is the product of the characteristic functions of the distributions of the individual variables.[3] The characteristic function, $g(t)$, of $f(x)$ is defined as follows:

(1) $$g(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\,dx.$$

As is well known, the inversion of (1) is given by:

(2) $$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx} g(t)\,dt.$$

It can be easily seen that the distribution of w_i is:

$$f(w_i)\,dw_i = \begin{cases} \dfrac{10^k}{C_i^n}\,dw_i, & \text{if} \quad -\dfrac{10^{-k}C_i^n}{2} \le w_i \le \dfrac{10^{-k}C_i^n}{2} \\ 0 & \text{elsewhere.} \end{cases}$$

and its characteristic function is:

$$g_i(t) = \frac{\sin \tfrac{1}{2}(10^{-k}C_i^n t)}{\tfrac{1}{2}(10^{-k}C_i^n t)}.$$

On the basis of the theorem, above mentioned, the characteristic function of the distribution of $\Delta^n - \bar{\Delta}^n = y$ (say) is:

(3) $$G(t) = \prod_{i=0}^{n} \frac{\sin \tfrac{1}{2}(10^{-k}C_i^n t)}{\tfrac{1}{2}(10^{-k}C_i^n t)}.$$

The desired frequency function, $F(y)$ is given by the inversion of

$$G(t) = \int_{-\infty}^{\infty} e^{ity} F(y)\,dy.$$

From (2) and (3), we get:

$$F(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ity} \prod_{i=0}^{n} \frac{\sin \tfrac{1}{2}(10^{-k}C_i^n t)}{\tfrac{1}{2}(10^{-k}C_i^n t)} \cdot dt$$

---

[3] See, for instance, Harald Cramér, *Random Variables and Probability Distributions*, Cambridge Tracts in Mathematics and Mathematical Physics No. 36 (1937) p. 36.

which may be written as:

$$F(y) = \frac{10^{k(n+1)}}{\pi \prod_{i=0}^{n} C_i^n} \int_{-\infty}^{\infty} \cos (2\,ty) \cdot \prod_{i=0}^{n} \sin (10^{-k} C_i^n t) \cdot \frac{dt}{t^{n+1}}. \tag{4}$$

The problem now reduces to the evaluation of the integral in (4). In the evaluation of this integral, it is convenient to consider even and odd values of n separately.

Case 1. When n is even

$$\prod_{i=0}^{n} \sin (10^{-k} C_i^n t) = \frac{(-1)^{\frac{1}{2}n}}{2^n} [P_{n+1} - P_n + P_{n-1} - \cdots + (-1)^{\frac{1}{2}n} P_{\frac{1}{2}n+1}]$$

where P_{n+1-r} denotes the sum of the sines of $n + 1 - r$ of the angles taken positively and the remaining r taken negatively, the negative angles being taken in every combination.[4] Thus $\cos (2\,ty) \prod_{i=0}^{n} \sin (10^{-k} C_i^n t)$ can be expressed as the sum of products of a sine by a cosine. By employing the identity $\sin A \cos B = \frac{1}{2}\{\sin (A + B) + \sin (A - B)\}$, each term can be written as the sum of two sines. Hence the integral under consideration can be written as the sum of integrals of the form

$$\int_{-\infty}^{\infty} \frac{\sin at}{t^{n+1}}\, dt.$$

Integrating by parts n times in succession, we obtain:

$$\int_{-\infty}^{\infty} \frac{\sin at}{t^{n+1}}\, dt = \frac{(-1)^{\frac{1}{2}n} a^n}{n!} \int_{-\infty}^{\infty} \frac{\sin at}{t}\, dt \tag{5}$$

But

$$\int_{-\infty}^{\infty} \frac{\sin at}{t}\, dt = \begin{cases} \pi & \text{for } a > 0 \\ -\pi & \text{for } a < 0. \end{cases}$$

(6) Therefore

$$\int_{-\infty}^{\infty} \frac{\sin at}{t^{n+1}}\, dt = \begin{cases} \dfrac{(-1)^{\frac{1}{2}n} a^n \pi}{n!} & \text{for } a > 0 \\[2ex] \dfrac{(-1)^{\frac{1}{2}n+1} a^n \pi}{n!} & \text{for } a < 0. \end{cases}$$

By use of (6) the integral in (4) can be readily evaluated.

Case 2. When n is odd.

$$\prod_{i=0}^{n} \sin 10^{-k} C_i^n t = \frac{(-1)^{\frac{1}{2}(n+1)}}{2^n} [Q_{n+1} - Q_n + Q_{n-1} - \cdots + (-1)^{\frac{1}{2}(n+1)} \cdot \tfrac{1}{2} Q_{\frac{1}{2}(n+1)}]$$

where Q_{n+1-r} denotes the sum of cosines of the sum of $n + 1 - r$ of the angles taken positively and the remaining r taken negatively, the negative angles being

[4] See E. W. Hobson, *Plane Trigonometry*, Seventh Edition (1928) pp. 50-51.

taken in every combination. As in Case I, the integral in (4) can be expressed as the sum of integrals of the form:

$$\int_{-\infty}^{\infty} \frac{\cos at}{t^{n+1}}\, dt.$$

Integrating by parts, we obtain:

$$\int_{-\infty}^{\infty} \frac{\cos at}{t^{n+1}}\, dt = -\frac{a}{n} \int_{-\infty}^{\infty} \frac{\sin at}{t^{n}}\, dt.$$

The second member of this equation has been treated in Case I. It follows that:

$$(7) \qquad \int_{-\infty}^{\infty} \frac{\cos at}{t^{n+1}}\, dt = \begin{cases} \dfrac{(-1)^{\frac{1}{2}(n+1)} a^n \pi}{n!} & \text{for } a > 0 \\[2ex] \dfrac{(-1)^{\frac{1}{2}(n-1)} a^n \pi}{n!} & \text{for } a < 0. \end{cases}$$

By means of the integrals, (6) and (7), $F(y)$ can be obtained for any n. The results for $n = 1$, 2, and 3 are given below:

$n = 1$

$$F(y) = \begin{cases} 10^{2k} y + 10^k & \text{for } -10^{-k} < y \le 0 \\ -10^{2k} y + 10^k & \text{for } 0 \le y < 10^{-k} \\ 0 & \text{elsewhere} \end{cases}$$

$n = 2.$

$$F(y) = \begin{cases} \dfrac{10^{3k}}{4} y^2 + 10^{2k} y + 10^k & \text{for } -2\cdot 10^{-k} < y \le -10^{-k} \\[2ex] -\dfrac{10^{3k}}{4} y^2 + \dfrac{10^k}{2} & \text{for } -10^{-k} \le y \le 10^{-k} \\[2ex] \dfrac{10^{3k}}{4} y^2 - 10^{2k} y + 10^k & \text{for } 10^{-k} \le y < 2\cdot 10^{-k} \\[2ex] 0 & \text{elsewhere} \end{cases}$$

$n = 3.$

$$F(y) = \begin{cases} \dfrac{10^{4k}}{54} y^3 + \dfrac{2\cdot 10^{3k}}{9} y^2 + \dfrac{8\cdot 10^{2k}}{9} y + \dfrac{32\cdot 10^k}{27} & \\ & \text{for } -4\cdot 10^{-k} < y \le -3\cdot 10^{-k} \\[2ex] -\dfrac{10^{4k}}{54} y^3 - \dfrac{10^{3k}}{9} y^2 - \dfrac{10^{2k}}{9} y + \dfrac{5\cdot 10^k}{27} & \text{for } -3\cdot 10^{-k} \le y \le -2\cdot 10^{-k} \\[2ex] \dfrac{10^{2k}}{9} y + \dfrac{10^k}{3} & \text{for } -2\cdot 10^{-k} \le y \le -10^{-k} \end{cases}$$

$n = 3$.—*cont.*

$$F(y) = \text{(cont.)} \begin{cases} -\frac{10^{4k}}{27} y^3 - \frac{10^{3k}}{9} y^2 + \frac{8 \cdot 10^k}{27} & \text{for} \quad -10^{-k} \le y \le 0 \\ \frac{10^{4k}}{27} y^3 - \frac{10^{3k}}{9} y^2 + \frac{8 \cdot 10^k}{27} & \text{for} \quad 0 \le y \le 10^{-k} \\ -\frac{10^{2k}}{9} y + \frac{10^k}{3} & \text{for} \quad 10^{-k} \le y \le 2 \cdot 10^{-k} \\ \frac{10^{4k}}{54} y^3 - \frac{10^{3k}}{9} y^2 + \frac{10^{2k}}{9} y + \frac{5 \cdot 10^k}{27} & \text{for} \quad 2 \cdot 10^{-k} \le y \le 3 \cdot 10^{-k} \\ -\frac{10^{4k}}{54} y^3 + \frac{2 \cdot 10^{3k}}{9} y^2 - \frac{8 \cdot 10^{2k}}{9} y + \frac{32 \cdot 10^k}{27} & \text{for} \quad 3 \cdot 10^{-k} \le y < 4 \cdot 10^{-k} \\ 0 & \text{elsewhere.} \end{cases}$$

In general, $F(y)$ is an even continuous function, vanishing for $| y | \ge 2^{n-1} 10^{-k}$ and defined by different n^{th} degree polynomials in different intervals.

The above frequency functions were derived on the assumption that the x's are random numbers drawn from a rectangular distribution. However, the results may be applied to the entries of a mathematical table provided the rounding errors are horizontally distributed and the difference under consideration is of such an order that the digits in the decimal place corresponding to the last place given in the table are also horizontally distributed. These conditions are frequently satisfied. Since data on the errors in the second differences of a table of $Ci(x) = \int_{\infty}^{x} \frac{\cos x}{x} dx$ given to 9 decimals was available, a study was made of a sample of 1000 of these errors. The theoretical and observed frequencies for this sample are given in the following table:

Error	−2 to −1.5	−1.5 to −1.0	−1.0 to −.5	−.5 to 0	0 to .5	.5 to 1.0	1.0 to 1.5	1.5 to 2	
Theoretical Frequency	10.4	72.9	177.1	239.6	239 6	177.1	72.9	10.4	1000.0
Observed Frequency	9	68	161	272	243	174	63	10	1000

By applying Pearson's χ^2-test, it is found that the observed frequencies show no significant deviations from the theoretical frequencies.

NEW YORK