THE ACCURACY OF LEAST SQUARES SOLUTIONS.
PART II: THE STANDARD DEVIATION OF THE ERRORS OF LINEAR EQUATIONS OF CONDITION

By T. E. Sterne
Harvard College Observatory
Communicated October 24, 1934

Often we know the results of \( n \) observations, \( c_1', \ldots, c_n' \), such that \( c_i' \) may be supposed to have been drawn at random from a normal universe of unknown standard deviation \( \sigma \) (the same for all the observations), and with mean \( a_i \xi_a + b_i \xi_b + d_i \xi_d + \ldots + m_i \xi_m \). Here the coefficients \( a_i \ldots m_i \) are supposed to be known, while the \( m \) quantities \( \xi_a \ldots \xi_m \) are unknown. We suppose that \( n > m \), and the question that immediately arises is (A): If we know only the values of \( c_i' \) and of \( a_i, b_i, \) etc., what are the values of the \( \xi ' s \)? The answer to question (A) has been given\(^1\) in a previous paper (referred to henceforth as Part I) in these PROCEEDINGS. The next question, perhaps, in order of importance is (B): If we know only the values of \( c_i' \) and of \( a_i, b_i, \) etc., what is the value of \( \sigma \)?

The classical answer to the problem is well known; it is furnished by the method of least squares. In the notation of Part I,

\[
\sigma = s',
\]

and this equation has been frequently employed in the practical evaluation of the mean or standard error of the equations of condition of least squares. That equation (10) is not correct has long been known, the resulting defect in the ordinary theory of least squares having been pointed out by Simon Newcomb,\(^2\) who, however, took no steps to correct the theory. While the answer to question (B) is not essential to the answer to question (A), as is apparent from the treatment given in Part I, it may nevertheless be of interest; and it is in any case desirable to present it for the sake of completeness.

Equation (8) of Part I is

\[
p(s, a)dsda = \text{const.} \times s^{n-m-1} e^{-\frac{(s-m)^2 + a^2|d/d\alpha|}{2\sigma^2}} dsda
\]
giving, in the notation of that paper, the distribution of samples of \( s' \) and \( a' \). Integrate over all \( a \). Then

\[
p(s)ds = \text{const.}^{IV} \times s^{n-m-1}e^{-\frac{(n-m)s^2}{2\sigma^2}} ds
\]

(8')

is the distribution of \( s' \) in the samples. Let \( y = \sigma/s \). Then the distribution of \( y \) in the samples is

\[
p(y)dy = B \times y^{m-n-1}e^{-\frac{n-m}{2\sigma^2}} dy
\]

(11)

in which the constant \( B \) is defined by

\[
B^{-1} = \frac{1}{2} \left( \frac{2}{n-m} \right)^{\frac{n-m}{2}} \Gamma \left( \frac{n-m}{2} \right).
\]

Equation (11) gives the chance that \( \sigma \) lies between \( ys' \) and \( (y + dy)s' \); it is the answer to question \( (B) \), and all the properties of the solution follow from it.

Equation (11), and the definition of \( B \), show that under the hypotheses of the first paragraph, \( p(y) \) does not depend upon \( s' \). The mode of \( y \) is \( [(n - m)/(n - m + 1)]^{1/2} \); the curve is skew. The mean, \( \bar{y} \), is infinite when \( n \leq m + 1 \), and when \( n > m + 1 \),

\[
\bar{y} = \left( \frac{n - m}{2} \right)^{1/2} \Gamma \left( \frac{n - m - 1}{2} \right)/ \Gamma \left( \frac{n - m}{2} \right).
\]

(12)

The mean value of \( y^2 \) is infinite when \( n \leq m + 2 \), and when \( n > m + 2 \),

\[
\bar{y}^2 = (n - m)/(n - m - 2).
\]

(13)

Hence the standard deviation of \( y \) about \( \bar{y} \), when \( n > m + 2 \), is

\[
\sigma_y = \left[ \frac{n - m}{n - m - 2} - \frac{n - m}{2} \Gamma^2 \left( \frac{n - m - 1}{2} \right)/ \Gamma^2 \left( \frac{n - m}{2} \right) \right]^{1/2}.
\]

(14)

The mean value of \( \sigma \), on the data of the first paragraph, is \( s'\bar{y} \), and the standard error of this value is \( s'\sigma_y \).

It is possible to obtain equation (9) of Part I from equation (8) and the results of the present paper. If \( \sigma \) is known, then the distribution of \( a' \), obtained by integrating (8), is

\[
p(a)da = \frac{1}{\sigma} \text{const.}^{V} e^{-\frac{a^2|\Delta/\Delta\alpha|}{2\sigma^2}} da,
\]

where \( \text{const.}^{V} \) does not depend upon \( \sigma \). If \( s' \) is known, the distribution of \( \sigma \) is, by (11),

\[
p(\sigma) d\sigma = Bs'^{n-m}\sigma^{m-n-1} e^{-\frac{(n-m)s'^2}{2\sigma^2}} d\sigma,
\]
By the law of mutually exclusive events, the distribution of $a'$ when $s'$ is known can be found by integrating the present $p(a)p(s)\,ds\,d\sigma$ over all $\sigma$. Equation (9) is then in fact obtained, showing the consistency of the results given in these papers. Equation (9) is the answer to question (A); and while the answer to question (A) has thus been shown to follow from the answer to question (B), the more direct derivation of equation (9) by the method of Part I is to be preferred.

As in Part I, no use has been made here of inverse probabilities. While the writer has been able to derive the results given in Part I by inverse probabilities, the inverse arguments are considerably inferior to the direct arguments of Parts I and II.

---

CHROMOSOME NUMBERS IN CERTAIN RICCIACEAE

By Margaret B. Siler

Department of Botany, University of Wisconsin

Communicated October 22, 1934

The Ricciaceae have been but little studied cytologically. A preliminary examination of the chromosomes of Riccia sorocarpa and R. Austini showed that further study of this genus might be interesting. Hence, a more extensive examination of various Riccias and of Ricciocarpus natans was undertaken. Table 1 shows the chromosome numbers thus far determined. All counts were made from equatorial plates.

Plants of Ricciocarpus natans from three sources give the same count, 9 chromosomes in the gametophyte (Fig. 1). These chromosomes are all small, but one is very minute, apparently nearly spherical. Garber and Lewis reported 4 as the gametophytic number for R. natans. The small size of the chromosomes and the large proportion of indistinct plates are disadvantageous, but the occasional plate in which the chromosomes are spread out shows 9 very clearly.

The genus Riccia is divided into two groups, Ricciella and Euriccia, on the basis of air-chamber structure. The latter group is much the larger. In the former, counts have been obtained for three species.

In a young antheridium of Riccia Sullivantii were found many plates favorable for counting. These invariably contain 7 large chromosomes; many show in addition an eighth very small chromosome (Fig. 2). All the chromosomes (except the smallest) are longer and thicker than any of those of Ricciocarpus natans.