

**UCLA**

**Department of Statistics, UCLA**

**Title**

Determination of the Accuracy of the Observations, by C. F. Gauss, Translation with preface

**Permalink**

<https://escholarship.org/uc/item/4n21q6bx>

**Author**

Ekström, Joakim

**Publication Date**

2013-11-12

# **Determination of the Accuracy of the Observations**

by Carl Friedrich Gauss.

*Translation by Joakim Ekström, with preface.*

ABSTRACT. *Bestimmung der Genauigkeit der Beobachtungen* is the second of the three major pieces that Gauss wrote on statistical hypothesis generation. It continues the methodological tradition of *Theoria Motus*, producing estimates by maximizing probability density, however absence of the change-of-variables theorem causes technical difficulties that compromise its elegance. In *Theoria Combinationis*, Gauss abandoned the aforementioned method, hence placing *Bestimmung der Genauigkeit* at a crossroads in the evolution of Gauss's statistical hypothesis generation methodology. The present translation is paired with a preface discussing the piece and its historical context.

## Translator's Preface and Discussion

by Joakim Ekström, UCLA Statistics

Carl Friedrich Gauss (1777-1855) published three major pieces on statistical hypothesis generation: *Theoria Motus* (1809), *Bestimmung der Genauigkeit* (1816) and *Theoria Combinationis* (1821) (see Sheynin, 1979). *Theoria Motus* was translated into English by C. H. Davis in 1858, *Theoria Combinationis* was translated into English by G. W. Stewart in 1995, but an English translation of *Bestimmung der Genauigkeit* has, in spite of great efforts, not been found in the literature. Hence the present translation.

*Bestimmung der Genauigkeit der Beobachtungen*, as its complete title reads, is an interesting historical text for many reasons. In it, Gauss uses the statistical hypothesis generation method of *Theoria Motus* for the purpose of estimating the standard deviation, basically, but is challenged by technical difficulties. In *Theoria Combinationis*, Gauss abandoned the aforementioned method in favor of a more rudimentary, and substantially less ambitious, method that considers linear combinations of the observations. As such, *Bestimmung der Genauigkeit* is conceptually and chronologically at a crossroads in the evolution of Gauss's statistical hypothesis generation methodology.

In addition, *Bestimmung der Genauigkeit* contains a first discussion of the estimator property that Fisher (1922) termed asymptotic relative efficiency, the definition of the Gauss error function, and Hald (1999) even argues that the piece contains the first application of the method of maximum likelihood.

This preface aims to discuss and provide historical context to *Bestimmung der Genauigkeit* and Gauss's statistical hypothesis generation methodology.

### 1. HISTORICAL CONTEXT

Fundamentally, the statistical hypothesis generation method of *Theoria Motus* and *Bestimmung der Genauigkeit* is based on the idea of Jakob Bernoulli (1654-1705), as discussed in *Ars Conjectandi* (1713), that empirical observations should be evaluated through the concept of probability. More precisely, the method is based on Bernoulli's fifth axiom: "*Between two, the one that seems more probable should always be chosen.*"

While elegant in principle, Bernoulli's approach has a fundamental weakness, which was first articulated by fellow philosopher-mathematician Gottfried Leibniz (1646-1716) in correspondence with Bernoulli (1703). In many applications, for instance Gauss's field of astronomy, there are commonly infinitely many possibilities that each has probability zero; i.e. there exists continuous, or non-atomic, probability distributions. Therefore, if every possibility has probability zero then choosing the most probable is not a meaningful course of action.

A pragmatic circumvention of this conundrum was proposed by Johann Lambert (1760; 1765), and subsequently by Jakob's nephew Daniel Bernoulli (1778), through a statistical criterion. Utilizing probability density, if one possibility has greater probability density than another, then by the density criterion the former is deemed more probable than the latter. Gauss employed the density criterion

in *Theoria Motus* and *Bestimmung der Genauigkeit*; the statistical criterion continues to be widely used to this day.

The difficulty in applying the density criterion is that relevant probability distributions need to be derived. In *Theoria Motus*, Gauss sought to determine the most probable Kepler orbit given observations of a heavenly body, and applied the Gauss-Pearson decomposition  $x = \mu + u$  where  $x$  is the observation,  $\mu$  the ideal part, i.e. the true position of the heavenly body, and  $u$  the observational error. A hypothesized ideal part  $\nu$  yields the representation  $x = \nu + e$  where  $e$  is the representation residual, and by Gauss's theory of errors the most probable ideal part corresponds to the most probable representation residual under the probability distribution of the observational error (see Ekström, 2012, for a comprehensive discussion). In *Theoria Motus*, under assumptions of statistically independent and normally distributed observational errors, maximal density is obtained by minimizing a sum of squares.

In *Bestimmung der Genauigkeit*, derivations of probability distributions are much more challenging. If the absolute value of the observational error is denoted by  $y$ , and the standard deviation of its probability distribution by  $\sigma$ , then under the normal distribution assumption the quotient  $y/\sigma$  is chi distributed with one degree of freedom. However, deriving the density function of this transformed random variable requires the change-of-variables theorem, which had not yet been fully developed in the early nineteenth century. In the absence of this needed change-of-variables result, Gauss developed a change-of-variables formula, according to which the transformed random variable,  $y/\sigma$ , is chi distributed with two degrees of freedom. Note that in the notation of Gauss the standard deviation satisfies  $\sigma = (\sqrt{2}h)^{-1}$ , where  $h$  is his accuracy measure.

A remark with respect to the density criterion is that, because the probability density function of the chi distribution with one degree of freedom is strictly decreasing, maximizing density yields an extreme value. Specifically, the most probable standard deviation under the density criterion is identically zero, regardless of the observations. This example illustrates the sometimes bizarre consequences of the density criterion. By contrast, the chi distribution with two degrees of freedom, which Gauss's change-of-variables formula yielded, has its mode at one, and as a consequence the estimate of the standard deviation is the intuitive and very sensible square-root of the mean square error.

## 2. GAUSS'S CHANGE-OF-VARIABLES FORMULA

If a random variable  $W$  can be expressed as a transformation  $T$  of a random variable  $U$ ,  $W = T(U)$ , then the probability that  $W$  attains an element of a set  $B$  can be determined through the identity

$$\text{Prob}(W \in B) = \text{Prob}(T(U) \in B) = \text{Prob}(U \in T^{-1}(B)), \quad (1)$$

where  $T^{-1}$  denotes inverse of  $T$ . Consequently, if  $U$  has a probability density function, the probability that  $W$  attains an element of  $B$  can be obtained by integrating the probability density function over the set  $T^{-1}(B)$ . While elegant, this method is in general difficult to apply directly because determining inverse images of sets is often quite laborious.

By the change-of-variables theorem, if the inverse transformation  $T^{-1}$  is differentiable and injective on an open set that contains  $B$ , then

$$\int_{T^{-1}(B)} f d\lambda = \int_B (f \circ T^{-1}) |J_{T^{-1}}| d\lambda,$$

where  $J_{T^{-1}}$  denotes the Jacobian determinant of  $T^{-1}$  and  $\lambda$  denotes the Lebesgue measure (see Rudin, 1987). Consequently, if both  $W$  and  $U$  have probability density functions,  $f_W$  and  $f_U$  respectively, and  $W = T(U)$  where  $T$  is differentiable and injective, then

$$\int_B f_W d\lambda = \int_B (f_U \circ T^{-1}) |J_{T^{-1}}| d\lambda, \quad (2)$$

and thus the identity  $f_W = (f_U \circ T^{-1}) |J_{T^{-1}}|$  is obtained.

Since the Lebesgue measure was constructed at the turn of the twentieth century, the change-of-variables theorem was not available to Gauss who died in 1855. However, simple versions can be derived through the fundamental theorem of calculus, which was well established in Gauss's days. In spite of this, Gauss did not use that result, but constructed a new change-of-variables formula for this particular requirement. In the following, Gauss's change-of-variables formula is discussed in some detail.

In *Theoria Motus*, Gauss argued that if  $P$  and  $Q$  are two probability distributions and one accepts the premise  $\text{Prob}(\mathcal{L}(W) = P) = \text{Prob}(\mathcal{L}(W) = Q) > 0$ , where  $\mathcal{L}(W)$  denotes the probability distribution of  $W$ , then it holds that

$$\frac{\text{Prob}(\mathcal{L}(W) = P | W(\omega) \in E)}{\text{Prob}(\mathcal{L}(W) = Q | W(\omega) \in E)} = \frac{P(E)}{Q(E)},$$

for any subset  $E$  satisfying  $Q(E) > 0$ . This relation can be shown through an application of Bayes' theorem; the issue of constructing topological product spaces of probability distributions and values of random variables is not discussed in the present text. In *Bestimmung der Genauigkeit*, probability densities are treated by method of infinitesimal calculus, allowing Gauss to propose analogously

$$\frac{\text{Dens}(\mathcal{L}(W) = P | W(\omega) \in \{b\})}{\text{Dens}(\mathcal{L}(W) = Q | W(\omega) \in \{b\})} = \frac{\text{Dens}(W(\omega) \in \{b\} | \mathcal{L}(W) = P)}{\text{Dens}(W(\omega) \in \{b\} | \mathcal{L}(W) = Q)},$$

where  $b$  is some value in the range of  $W$ .

More specifically, if  $\{T_\alpha\}_{\alpha \in A}$  is an indexed set of transformations,  $A$  being its index set, then the indexed set  $\{\mathcal{L}(T_\alpha(U))\}_{\alpha \in A}$  is a set of probability distributions, and if the probability density function of  $T_\alpha(U)$  is denoted by  $f_{T_\alpha(U)}$ , then Gauss's density identity can be expressed

$$\text{Dens}(\mathcal{L}(W) = \mathcal{L}(T_\alpha(U)) | W(\omega) \in \{b\}) = \frac{c f_{T_\alpha(U)}(b)}{f_{T_{\alpha_0}(U)}(b)},$$

where  $c$  is a constant and  $\alpha_0 \in A$ . Furthermore, if the identity function is an element of  $\{T_\alpha\}_{\alpha \in A}$  and if the probability distribution  $\mathcal{L}(T_\alpha(U))$  is unique for each  $\alpha \in A$ , then the identity can be written

$$\hat{f}_W(\alpha) = \text{Dens}(W = T_\alpha(U) | W(\omega) = b) = \frac{cf_{T_\alpha(U)}(b)}{f_U(b)}, \quad (3)$$

which is Gauss's change-of-variables formula.

As it is constructed, the function  $\hat{f}_W$  of Equation (3) is defined on the index set,  $A$ , rather than the range of the transformed random variable. However, in the applications of Gauss the index sets are taken so that they equal the ranges of the transformed random variables; thus alleviating this inconsistency. In *Bestimmung der Genauigkeit*, the the index sets are either the non-negative real numbers or the real numbers, matching the ranges of the transformed random variables studied.

### 3. FIRST DERIVATION OF THE ACCURACY MEASURE

In Section 3 of *Bestimmung der Genauigkeit*, Gauss uses his change-of-variables formula to derive the probability density function of his accuracy measure,  $h$ , which is defined by  $\sigma = (\sqrt{2}h)^{-1}$ . Letting  $y$  denote the absolute value of the observational error, it is natural to use the Gauss-Pearson decomposition  $y = \sigma u$ , where  $\mathcal{L}(u) = \chi_1$ , i.e. the chi distribution with 1 degree of freedom. In this part of *Bestimmung der Genauigkeit*, Gauss approaches the observation as a manifested quantity, and the observation is therefore treated as a constant. Solving for  $\sigma$  yields  $\mathcal{L}(\sigma) = \mathcal{L}(yu^{-1})$ , and so the ideal part to be estimated is treated as a random variable.

By the change-of-variables theorem, the probability density function of the random variable  $yu^{-1}$  is  $f_{yu^{-1}}(z) = f_\sigma(z) = cz^{-2}e^{-y^2/2z^2}$ , where  $c$  is a constant and  $z$  the real-valued argument. By applying Gauss's change-of-variables formula, using the known probability density functions  $f_u(z) = c_1e^{-z^2/2}$  and  $f_y(z) = f_{\sigma u}(z) = c_2\sigma^{-1}e^{-z^2/2\sigma^2}$ , one obtains  $\hat{f}_\sigma(z) = \tilde{c}z^{-1}e^{-y^2/2z^2}$ , where  $c_1$ ,  $c_2$  and  $\tilde{c}$  are constants. If Gauss's accuracy measure  $h$  is used instead of the standard deviation,  $\sigma$ , the change-of-variables theorem yields  $f_h(z) = ce^{-z^2y^2}$ , i.e.  $\mathcal{L}(\sqrt{2}hy) = \chi_1$ , while application of Gauss's change-of-variables formula produces  $\hat{f}_h(z) = \tilde{c}ze^{-z^2y^2}$ , i.e.  $\mathcal{L}(\sqrt{2}hy) = \chi_2$ .

It may also be noted that  $\hat{f}_\sigma$  and  $\hat{f}_h$  are inconsistent relative to the change-of-variables theorem in the sense that, letting  $R(h) = (\sqrt{2}h)^{-1} = \sigma$ , it holds that  $\hat{f}_\sigma \neq (\hat{f}_h \circ R^{-1})|_{J_{R^{-1}}}$ ; hence contradicting Equation (2). In fact it holds that  $\hat{f}_\sigma = (\hat{f}_h \circ R^{-1})$ , i.e. Gauss's formula does not account for the Jacobian determinant. This circumstance explains why Gauss's change-of-variables formula did not give rise to corresponding problems in *Theoria Motus*, because with the additive Gauss-Pearson decomposition  $x = \mu + u$  solving for the ideal part  $\mu$  yields  $\mu = x - u$ , which is a transformation that has Jacobian determinant absolute value equal to one. In this case, Equation (1) can easily be applied directly, as well.

Given the derived density function  $\hat{f}_\sigma(z) = \tilde{c}z^{-1}e^{-y^2/2z^2}$ , application of the density criterion yields that the most probable value of  $h$  is  $(\sqrt{2}y)^{-1}$ , and with  $m$  statistically independent observations the most probable value of  $h$  is  $\sqrt{m/2\vec{y}^t\vec{y}}$ , where  $\vec{y}$  and  $\vec{y}^t$  denote the vector  $(y_1, \dots, y_m)$  and its transpose respectively. This corresponds to estimating the variance by the mean square observation, a result

that Gauss expressed comfort with; according to Sheynin (1979) the estimate had been discussed by Laplace in the year 1815.

In Section 4, Gauss derives the probability density function of the observational error of his estimate. If the estimate  $\sqrt{m/2\bar{y}^t\bar{y}}$  is denoted by  $H$ , then this observation of  $h$  is equipped with Gauss-Pearson decomposition  $H = h - \lambda$  where  $h$  is the true accuracy and  $\lambda = h - H$  denotes the observational error. By using  $y = (\sqrt{2}H)^{-1}$ , the function  $\hat{f}_h$  can be written  $\hat{f}_h(z) = cze^{-z^2/2H^2}$ , and by Equation (1)  $\hat{f}_{h-z}(H) = \hat{f}_h(H+z)$ . Gauss's change-of-variables formula, Equation (3), yields

$$\hat{f}_\lambda(z) = \frac{c\hat{f}_{h-z}(H)}{\hat{f}_h(H)} = \frac{\tilde{c}(H+z)e^{-(H+z)^2/2H^2}}{He^{-H^2/2H^2}} = \tilde{c}(1+z/H)e^{-z/H}e^{-z^2/2H^2}.$$

Given  $m$  statistically independent observations  $y_1, \dots, y_m$ ,

$$\hat{f}_\lambda(z) = \tilde{c}^m (1+z/H)^m e^{-mz/H} e^{-mz^2/2H^2},$$

and by using the Maclaurin series for  $\log(1+x)$  one obtains the approximation  $(1+x)^m \approx e^{mx} e^{-mx^2/2}$ , and it follows

$$\hat{f}_\lambda(z) \approx \tilde{c}^m (e^{mz/H} e^{-mz^2/2H^2}) e^{-mz/H} e^{-mz^2/2H^2} = \tilde{c}^m e^{-mz^2/H^2}.$$

Hence Gauss concludes that for large  $m$ ,  $\lambda$  is approximately mean zero normally distributed with variance  $H^2/2m$ , and consequently his accuracy estimate,  $h = H + \lambda$ , is approximately normal,  $\mathcal{L}(h) = N(H, H^2/2m)$ .

#### 4. DETERMINATION OF 50% CONFIDENCE INTERVALS

In Sections 4 and 5, Gauss discusses determination of probable limits of the true value of the accuracy measure, which in modern terminology are referred to as endpoints of 50% confidence intervals.

In general, the preferred method of Gauss is to firstly obtain results of the form  $Q \sim N(q, q^2c^2/m)$ , where  $Q$  is the sample estimate of a true value  $q$ ,  $c$  a constant, and  $m$  the sample size. A 50% prediction interval for  $Q$  is  $q(1 \pm \sqrt{2}\rho c/\sqrt{m})$ , where  $\rho = 0.4769363\dots$  is the number such that the 75% percentile of the standard normal distribution equals  $\sqrt{2}\rho$ . Gauss then states that  $Q(1 \pm \sqrt{2}\rho c/\sqrt{m})$  is a 50% confidence interval for  $q$ , a maneuver which may be termed the prediction-confidence interval substitution.

In a modern argument, the assumed probability distribution of  $Q$  implies that the random variable converges in probability to  $q$  as the sample size goes to infinity, and by an application of Slutsky theorems an asymptotic 50% confidence interval for  $q$  is  $Q(1 \pm \sqrt{2}\rho c/\sqrt{m})$ . Hence, by substitution of the true value for the sample estimate, an approximate confidence interval for the true value is obtained from a prediction interval for the sample estimate. Consequently, this prediction-confidence interval substitution is correct asymptotically. The challenge of this method lies in finding relevant sample estimates that satisfies  $Q \sim N(q, q^2c^2/m)$ .

In Section 4, Gauss uses the estimate  $H = \sqrt{m/2\bar{y}^t\bar{y}}$ , derived in Section 3, to find a 50% confidence interval for  $h$ . In this part of *Bestimmung der Genauigkeit*, the observation is considered a given

constant and the ideal part to be estimated is considered a random variable. Since Gauss obtained  $\mathcal{L}(h) = N(H, H^2/2m)$ , the probable limits for  $h$  are  $H(1 \pm \rho/\sqrt{m})$ .

The twenty-first century standard solution of this problem looks like the following: Since the observations are assumed statistically independent and identically distributed  $N(0, (\sqrt{2}h)^{-2})$ ,  $2H^{-2}$  is identified as a sample mean and is hence, suitably normalized, asymptotically normal by the central limit theorem. Also,  $H$  is asymptotically normal by Cramér's theorem, and  $\sqrt{m}(H - h)$  converges in distribution to  $N(0, h^2/2)$ . Thus a 50% prediction interval for  $H$  is given by  $h(1 \pm \rho/\sqrt{m})$ , and by the prediction-confidence interval substitution an asymptotic 50% confidence interval for  $h$  is given by  $H(1 \pm \rho/\sqrt{m})$ . Even though Gauss uses his imperfect change-of-variables formula, he still reaches the correct asymptotic confidence interval, which is quite remarkable.

At the conclusion of Section 4, Gauss claims that 50% confidence intervals for the quantity  $r = \rho/h$ , i.e. the median absolute error, are both  $R/(1 \pm \rho/\sqrt{m})$  and  $R(1 \pm \rho/\sqrt{m})$ , where  $R = \rho/H$ . By Cramér's theorem, the second confidence interval is asymptotically correct. Although Gauss offers minimal explanation, the first confidence interval is given in the same sentence as the confidence interval for  $h$ , and it is therefore possible that the first confidence interval for  $r$  is obtained through a transformation of the confidence interval for  $h$ . Specifically, if  $g(x) = \rho/x$ , so that  $g(h) = r$ , then, because  $g$  is continuous and strictly decreasing on the positive reals, the image of the interval  $H(1 \pm \rho/\sqrt{m})$  under  $g$  is  $R/(1 \pm \rho/\sqrt{m})$ .

## 5. SECOND DERIVATION OF THE ACCURACY MEASURE

In Section 5, Gauss proceeds along a line of reasoning which from a twenty-first century perspective is more conventional; studying distributions of statistics. By using results of Pierre-Simon Laplace, Gauss claims that sums of powers of the absolute values of the observations,  $S^{(n)} = y_1^n + \dots + y_m^n$  in the present notation, are normally distributed assuming the sample size,  $m$ , is a large number. Under the normal distribution assumption, the expected value and variance of  $S^{(n)}$  are computed, and then construction of 50% prediction intervals for  $S^{(n)}$  is straightforward.

However, the aim is to construct confidence intervals for his accuracy measure, and for this purpose Gauss proposes the transformation  $T_n(z) = r \sqrt[n]{z/mK^{(n)}}$ , where  $K^{(n)}$  denotes the expected value of  $y^n$ . Gauss claims that the most probable value, which equals the expected value under the normal distribution assumption, of  $T_n(S^{(n)})$  is  $r$ , however the equality  $E(\sqrt[n]{S^{(n)}}) = \sqrt[n]{E(S^{(n)})}$  holds only if  $n = 1$  or if the random variable is one-point distributed, as follows by Jensen's inequality. Further, Gauss uses the equality  $\text{Var}(\sqrt[n]{S^{(n)}}) = \sigma^2 \mu^{2(1-n)}/n^2$ , where  $\mu$  and  $\sigma$  denote the mean and standard deviation of  $S^{(n)}$ , which also only holds if  $n = 1$ . But asymptotically, by Cramér's theorem the claims hold for all  $n$ , and through this asymptotical approximation it holds approximately that  $T_n(S^{(n)}) \sim N(r, r^2 A^{(n)}/n^2 m)$ , where  $A^{(n)}$  satisfies  $\text{Var}(S^{(n)}) = mK^{(n)2}A^{(n)}$ .

Thus, for  $n = 1, 2, 3, \dots$ , asymptotic 50% prediction intervals for  $T_n(S^{(n)})$  are given by  $r(1 \pm \sqrt{2}\rho/n\sqrt{A^{(n)}/m})$ , and asymptotic 50% confidence intervals for  $r$  are obtained through the prediction-confidence interval substitution, i.e.  $T_n(S^{(n)})(1 \pm \sqrt{2}\rho/n\sqrt{A^{(n)}/m})$ . For  $n = 2$ , Gauss notes that the 50% confidence interval is identical to that obtained in Section 4.

On more than one occasion, Gauss derives results that fundamentally rest on Cramér's theorem (for reference, see, e.g., Ferguson, 1996). This circumstance raises the question of whether Gauss was aware of this twentieth-century result. In writing, Gauss offers little explanation other than stating that the results hold clearly. Though fundamentally, Cramér's theorem is an application of local linear approximation of differentiable functions, and considering Gauss's works on power series expansions it is not at all inconceivable that Gauss understood this result. At the same time, the fact that he did not elevate the result into a theorem, or at least discuss it in some detail, must be taken as circumstantial evidence that Gauss did not fully understand the entirety of this elegant result. An additional possibility is that Gauss presumed asymptotic normality, and then simply computed the means and variances.

## 6. EFFICIENCY, NON-PARAMETRICS, AND ILLUSTRATION

Section 6 of *Bestimmung der Genauigkeit* discusses a property of the estimates of  $r$  that in modern terminology is referred to as asymptotic relative efficiency. This treatment may well be one of the first discussions of this estimator property, which often is attributed to Fisher (1922). According to the computations of Gauss, the estimate based on the sum of squares has the smallest variance and therefore utilizes the observations most efficiently.

Despite being less efficient, i.e. less accurate at a given sample size, Gauss argues that the most easily computed estimate well can be used; thereby demonstrating a willingness to accept a trade-off of accuracy for ease of computation. It should be noted that the simplest estimate is asymptotically normal through the central limit theorem only, while the other estimates are asymptotically normal through the central limit theorem paired with Cramér's theorem; a circumstance that typically yields a worse asymptotical approximation of the distribution of those latter estimates.

In Section 7, Gauss discusses a 50% confidence interval for  $r$  that is based on the sample median of the absolute values. While no citation is given, the results Gauss uses could well be of Laplace, since Laplace studied the sample median and is referenced in Section 5. The 50% confidence interval given is correct asymptotically. Section 8 provides a numerical example from astronomy, which was the scientific discipline of Gauss's professorship.

## 7. DISCUSSION

*Bestimmung der Genauigkeit* is an interesting historical document for many reasons. In and of itself, it demonstrates that Gauss maintained an interest in statistics and the advancement of science through empirical observation. The text also demonstrates that Gauss sometimes were, to an extent, pragmatic; using approximations when reasonable and arguing that a less efficient estimate in

some instances can be preferred on the ground that it is more easily computed. As a service, Gauss provides the reader with numerical values so to facilitate easier computation. Like in *Theoria Motus*, the derivations rest on an assumption of normally distributed observational errors. In its entirety, *Bestimmung der Genauigkeit* provides a rare unmediated insight into Gauss's statistical hypothesis generation methodology.

In the statistics literature, *Bestimmung der Genauigkeit* has been discussed in the context of the early history of the method of maximum likelihood (Hald, 1999). In Section 3, Gauss derives the most probable value of his accuracy measure in a way that has technical similarities to the twentieth-century likelihood method. However, at close examination it is quite clear that Gauss is using a method that was conventional at his time and also applied in *Theoria Motus*; deriving the probability distribution of the observational error and using the mode as the most probable value, as per the density criterion of Lambert and Bernoulli. The technical similarities between Gauss's derivations and the method of maximum likelihood are a result of Gauss's imperfect change-of-variables formula.

Possibly due to elegance-compromising technical difficulties caused by his imperfect change-of-variables formula, Gauss abandoned the method of *Theoria Motus* and *Bestimmung der Genauigkeit*, i.e. estimation through Bernoulli's fifth axiom and the density criterion. In *Theoria Combinationis*, Gauss used a more rudimentary and substantially less ambitious method based on the premise of restricting estimates to linear combinations of the observations. Given an  $m$ -sized sample  $\vec{x}$  of statistically independent and identically distributed real-valued observations, the linear combinations  $\vec{a}^t \vec{x}$  are unbiased estimates when  $\vec{a}^t(1, \dots, 1) = 1$ . Since the variances of those estimates are proportional to  $\vec{a}^t \vec{a}$ , the minimum variance unbiased linear combination estimator is  $m^{-1}(1, \dots, 1)^t \vec{x}$ , i.e. the arithmetic mean. *Theoria Combinationis* was Gauss's last major piece on statistical hypothesis generation.

#### ACKNOWLEDGEMENTS

This work was supported by the Swedish Council for Working Life and Social Research, project 2010-1406. The author is grateful to Ulrike Grömping for valuable comments and suggestions for improvements of the translation.

#### REFERENCES

- Bernoulli, D. (1778). Diiudicatio maxime probabilis plurium obseruationum discrepantium atque verisimillima inductio inde formanda. *Acta Acad. Scient. Imper. Petrop., pro Anno 1777, Pars prior*, 3–23. English translation by C. G. Allen, 1961.
- Bernoulli, J. (1713). *Ars Conjectandi*. Basel: Thurnisiorum Fratrum. English translation by E. D. Sylla, 2006.
- Ekström, J. (2012). Statistical hypothesis generation: Determining the most probable subset. *UCLA Statistics Preprints*, 641.
- Ferguson, T. S. (1996). *A Course in Large Sample Theory*. New York: Chapman & Hall.

- Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Phil. Trans. R. Soc. Lond. A*, 222, 309–368.
- Gauss, C. F. (1809). *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium*. Hamburg: F. Perthes und I. H. Besser. English translation by C. H. Davis, 1858.
- Gauss, C. F. (1812). Disquisitiones generales circa seriem infinitam: Pars prior. *Commentationes societatis regiae scientiarum Gottingensis recentiores*, 2, 1–46.
- Gauss, C. F. (1816). Bestimmung der Genauigkeit der Beobachtungen. *Zeitschrift für Astronomie und verwandte Wissenschaften*, 1, 185–197.
- Gauss, C. F. (1821). Theoria combinationis observationum erroribus minimis obnoxiae: Pars prior. *Commentationes societatis regiae scientiarum Gottingensis recentiores*, 5, 33–62. English translation by G. W. Stewart, 1995.
- Hald, A. (1999). On the history of maximum likelihood in relation to inverse probability and least squares. *Statist. Sci.*, 14, 214–222.
- Lambert, J. H. (1760). *Photometria, sive de mensura et gradibus luminis, colorum et umbrae*. Augsburg: Typis Christophori Petri Detleffsen. English translation by D. L. DiLaura, 2001.
- Lambert, J. H. (1765). Das Mittel zwischen den Fehlern. In *Beiträge zum Gebrauche der Mathematik und deren Anwendung*, (pp. 296–313). Berlin: Verlage des Buchladens der Realschule.
- Leibniz, G. W. (1703). Letter to Jakob Bernoulli, December 3, 1703. In C. I. Gerhardt (Ed.) *Leibnizens mathematische Schriften, Erste Abtheilung, III*, 1855, (pp. 79–86). Berlin: H. W. Schmidt. English translation by B. Sung, 1966.
- Rudin, W. (1987). *Real and Complex Analysis, third ed.* Singapore: McGraw-Hill.
- Sheynin, O. B. (1979). C. F. Gauss and the theory of errors. *Arch. Hist. Exact Sci.*, 20, 21–72.

**Zeitschrift**  
für  
**Astronomie**  
und  
**verwandte Wissenschaften.**

---

*March and April 1816.*

---

**XII. Determination of the Accuracy of the Observations.**

*by Professor Gauss,*

Knight of the Royal Hanoverian Guelphic Order.

1.

In the justification of the so-called Method of Least Squares, it is assumed that the probability of an observational error  $\Delta$  can be expressed through the formula

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2},$$

where  $\pi$  denotes the half circumference,  $e$  the basis of the hyperbolic logarithm, and  $h$  a constant<sup>1</sup> that, by Section 178 of *Theoria Motus Corporum Coelestium*, can be viewed as the measure of the accuracy of the observations.<sup>2</sup> When using the Method of Least Squares to estimate the most probable value of the quantity that the observations depend on, then knowledge of the constant  $h$  is not needed; even the *ratio* of the accuracy of the estimate to the accuracy the observations is independent of  $h$ . Still, knowledge of the accuracy measure  $h$  is interesting and instructive in and of itself, and I will therefore show how one through the observations may reach such knowledge.

2.

First, I am allowing myself to precede the subject matter with a few explanatory remarks. For convenience, I denote the value of the integral

$$\int \frac{2e^{-t^2} dt}{\sqrt{\pi}},$$

---

<sup>1</sup>In *Theoria Motus*, Gauss argued, through making a parallel with the method of the arithmetic mean, that this should be accepted as the probability of all observational errors. During the nineteenth century, this claim was commonly referred to as the *law of errors*.

<sup>2</sup>In modern notation,  $h^{-1} = \sigma\sqrt{2}$  where  $\sigma$  denotes the standard deviation.

from 0 to  $t$ , by  $\Theta(t)$ .<sup>3</sup> A few separate values provide an understanding of the shape of this function. One has

$$\begin{aligned}
 0.5000000 &= \Theta(0.4769363) = \Theta(\rho), \\
 0.6000000 &= \Theta(0.5951161) = \Theta(1.247790\rho), \\
 0.7000000 &= \Theta(0.7328691) = \Theta(1.536618\rho), \\
 0.8000000 &= \Theta(0.9061939) = \Theta(1.900032\rho), \\
 0.8427008 &= \Theta(1) = \Theta(2.096716\rho), \\
 0.9000000 &= \Theta(1.1630872) = \Theta(2.438664\rho), \\
 0.9900000 &= \Theta(1.8213864) = \Theta(3.818930\rho), \\
 0.9990000 &= \Theta(2.3276754) = \Theta(4.880475\rho), \\
 0.9999000 &= \Theta(2.7510654) = \Theta(5.768204\rho), \\
 1 &= \Theta(\infty).
 \end{aligned}$$

The probability that the error of an observation lies between the limits  $-\Delta$  and  $+\Delta$ , or, disregarding the sign, is not greater than  $\Delta$ , is

$$= \int \frac{he^{-h^2x^2} dx}{\sqrt{\pi}}$$

when one stretches the integral from  $x = -\Delta$  to  $x = +\Delta$ , or two times the same integral when taken from  $x = 0$  to  $x = \Delta$ , thus

$$= \Theta(h\Delta).$$

The probability that the error is not less than  $\frac{\rho}{h}$  is thus  $= \frac{1}{2}$ , or equal to the probability of the contrary; we shall name this quantity the *probable error*, and denote it by  $r$ .<sup>4</sup> By contrast, the probability that the error exceeds  $2.438664r$  is only  $\frac{1}{10}$ ; the probability that the error rises above  $3.818930r$  is only  $\frac{1}{100}$ , and so forth.

3.

We shall now assume that the errors  $\alpha, \beta, \gamma, \delta, \dots$ , have been manifested through  $m$  actual observations, and investigate what can be concluded from these with respect to the values of  $h$  and  $r$ . If one makes two suppositions, in which the true value of  $h$  is either set to  $= H$  or  $= H'$ , then the probabilities of observing the errors  $\alpha, \beta, \gamma, \delta, \dots$  relate as<sup>5</sup>

$$He^{-H^2\alpha^2} He^{-H^2\beta^2} He^{-H^2\gamma^2} \dots \quad \text{to} \quad H'e^{-H'^2\alpha^2} H'e^{-H'^2\beta^2} H'e^{-H'^2\gamma^2} \dots,$$

i.e. as

$$H^m e^{-H^2(\alpha^2+\beta^2+\gamma^2+\dots)} \quad \text{to} \quad H'^m e^{-H'^2(\alpha^2+\beta^2+\gamma^2+\dots)}.$$

<sup>3</sup>In modern terminology,  $\Theta(t)$  is the (Gauss) error function.

<sup>4</sup>In modern terminology,  $r$  is the median absolute value.

<sup>5</sup>Apparently, a statistical independence assumption is also made.

This relationship therefore expresses the probabilities that the true value of  $h$  was  $H$  or  $H'$ , after the realization of these errors (*T. M. C. C.* Section 176); or, the probability of each possible value of  $h$  is proportional to the quantity<sup>6</sup>

$$h^m e^{-h^2(\alpha^2+\beta^2+\gamma^2+\dots)}.$$

The *most probable* value of  $h$  is consequently that which maximizes this quantity, from which one derives the familiar rule

$$= \sqrt{\frac{m}{2(\alpha^2 + \beta^2 + \gamma^2 + \dots)}}.$$

The most probable value of  $r$  is thus

$$= \rho \sqrt{\frac{2(\alpha^2 + \beta^2 + \gamma^2 + \dots)}{m}} = 0.6744897 \sqrt{\frac{\alpha^2 + \beta^2 + \gamma^2 + \dots}{m}}.$$

This result holds generally, whether  $m$  is large or small.

4.

One understands easily that the smaller  $m$  is, the less reliable these determinations of  $h$  and  $r$  are, in terms of accuracy. For this purpose we develop the degree of accuracy one should attach to these determinations, in the case where  $m$  is a large number. For convenience, we denote by  $H$  the previously derived most probable value of  $h$ , i.e.

$$\sqrt{\frac{m}{2(\alpha^2 + \beta^2 + \gamma^2 + \dots)}},$$

and note that the probability that  $H$  is the true value of  $h$  to the probability that  $H + \lambda$  is the true value of  $h$  relate as

$$H^m e^{-\frac{m}{2}} \quad \text{to} \quad (H + \lambda)^m e^{-\frac{m(H+\lambda)^2}{2H^2}},$$

or as<sup>7</sup>

$$1 \quad \text{to} \quad e^{-\frac{\lambda^2 m}{H^2} (1 - \frac{1}{3} \frac{\lambda}{H} + \frac{1}{4} \frac{\lambda^2}{H^2} - \frac{1}{5} \frac{\lambda^3}{H^3} + \dots)}.$$

The second term is only appreciable, relative to the first, when  $\frac{\lambda}{H}$  is small, and therefore we may allow ourselves to use

$$1 \quad \text{to} \quad e^{-\frac{\lambda^2 m}{H^2}},$$

instead of the specified relation. Simply put, the probability that the true value of  $h$  lies between  $H + \lambda$  and  $H + \lambda + d\lambda$  is very close to

$$= K e^{-\frac{\lambda^2 m}{H^2}} d\lambda,$$

where  $K$  is a constant specified so that the integral

$$\int K e^{-\frac{\lambda^2 m}{H^2}} d\lambda,$$

<sup>6</sup>This derived probability distribution is incorrect; see translator's preface.

<sup>7</sup>This step utilizes the Maclaurin series for  $\log(1+x)$ ; see preface.

between appropriate limits with respect to  $\lambda$ , shall = 1. Instead of such limits, it is here permitted to take the limits  $-\infty$  and  $+\infty$ , since the size of  $m$  clearly renders  $e^{-\frac{\lambda^2}{H^2}}$  negligible as soon as  $\frac{\lambda}{H}$  seizes to be a small fraction, whereby

$$K = \frac{1}{H} \sqrt{\frac{m}{\pi}}.$$

Hence the probability that the true value of  $h$  lies between  $H - \lambda$  and  $H + \lambda$  is

$$= \Theta\left(\frac{\lambda}{H} \sqrt{m}\right),$$

thus this probability =  $\frac{1}{2}$  when

$$\frac{\lambda}{H} \sqrt{m} = \rho.$$

The odds are thus one-to-one that the true value of  $h$  lies between  $H(1 - \frac{\rho}{\sqrt{m}})$  and  $H(1 + \frac{\rho}{\sqrt{m}})$ , or that the true value of  $r$  falls between

$$\frac{R}{1 - \frac{\rho}{\sqrt{m}}} \quad \text{and} \quad \frac{R}{1 + \frac{\rho}{\sqrt{m}}},$$

where  $R$  denotes the most probable value of  $r$  that was derived in the preceding Section. One can name these limits the *probable limits of the true values of  $h$  and  $r$* .<sup>8</sup> Clearly, we can here also set the probable limits for the true value of  $r$  as  $R(1 - \frac{\rho}{\sqrt{m}})$  and  $R(1 + \frac{\rho}{\sqrt{m}})$ .

5.

In the preceding investigation, we have been of the view that we consider  $\alpha, \beta, \gamma, \delta, \dots$  as given quantities, and seek the amount of probability that the true values of  $h$  and  $r$  lie between certain known limits. However, one can also look at the matter from another point of view, and under the presumption, that the observational errors are subjected to a certain probability law that, in turn, determines the probability that the expected value of the sum of squares of  $m$  observational errors falls between known limits. This problem, given the condition that  $m$  is a large number, has already been solved by *Laplace*,<sup>9</sup> as has the problem of determining the probability that the sum of  $m$  observational errors itself falls between certain known limits. One can easily generalize this investigation even further; I am here content with stating this result.

Let  $\phi(x)$  denote the probability of the observational error  $x$ , so that  $\int \phi(x) dx = 1$  when one stretches the integral from  $-\infty$  to  $\infty$ . We shall generally denote by  $K^{(n)}$  the value of the integral

$$\int \phi(x) x^n dx,$$

<sup>8</sup>In modern terminology, these limits are the endpoints of a 50% confidence interval.

<sup>9</sup>*Laplace* is presumably Pierre-Simon Laplace (1749-1827), French mathematician and astronomer.

between these limits.<sup>10</sup> Let further  $S^{(n)}$  denote the sum

$$\alpha^n + \beta^n + \gamma^n + \delta^n + \dots$$

where  $\alpha, \beta, \gamma, \delta, \dots$  denote undetermined  $m$  observational errors. The parts in that sum should all be taken as positive, also for odd  $n$ .

Then  $mK^{(n)}$  is the most probable value of  $S^{(n)}$ , and the probability that the true value of  $S^{(n)}$  falls between the limits  $mK^{(n)} - \lambda$  and  $mK^{(n)} + \lambda$  is<sup>11</sup>

$$= \Theta \left( \frac{\lambda}{\sqrt{2m(K^{(2n)} - K^{(n)2})}} \right).$$

Consequently the probable limits of  $S^{(n)}$  are

$$mK^{(n)} - \rho\sqrt{2m(K^{(2n)} - K^{(n)2})}$$

and

$$mK^{(n)} + \rho\sqrt{2m(K^{(2n)} - K^{(n)2})}.$$

This result holds generally, for every law of the observational errors. If we turn to the case where

$$\phi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$$

is taken, then we find

$$K^{(n)} = \frac{\Pi \frac{1}{2}(n-1)}{h^n \sqrt{\pi}},$$

where the characteristic  $\Pi$  is taken in the meaning of *Disquisitiones generales circa seriem infinitam* (Comm. nov. soc. Gotting. T. U., Section 28).<sup>12</sup> Thus,<sup>13</sup>

$$\begin{aligned} K^I &= 1, & K^I &= \frac{1}{h\sqrt{\pi}} \\ K^{II} &= \frac{1}{2h^2}, & K^{III} &= \frac{1}{h^3\sqrt{\pi}} \\ K^{IV} &= \frac{1\cdot 3}{4h^4}, & K^V &= \frac{1\cdot 2}{h^5\sqrt{\pi}} \\ K^{VI} &= \frac{1\cdot 3\cdot 5}{8h^6}, & K^{VII} &= \frac{1\cdot 2\cdot 3}{h^7\sqrt{\pi}}, \quad \text{and so forth.} \end{aligned}$$

Consequently, the most probable value of  $S^{(n)}$  is

$$\frac{m\Pi \frac{1}{2}(n-1)}{h^n \sqrt{\pi}},$$

<sup>10</sup>In modern terminology,  $K^{(n)}$  is the expected value of the  $n$ :th power.

<sup>11</sup>Apparently, this as an application of the central limit theorem. For normal distributions the mode equals the mean, and  $\text{Var}(S^{(n)}) = m(K^{(2n)} - K^{(n)2})$ .

<sup>12</sup>The function  $\Pi$  is defined  $\Pi(\lambda) = \int_0^\infty y^\lambda e^{-y} dy$ , i.e.  $\Pi(\lambda) = \Gamma(\lambda+1) = \lambda\Gamma(\lambda)$ , where  $\Gamma$  denotes the Gamma function. Thus  $\Pi(-1/2) = \sqrt{\pi}$  and  $\Pi(n) = n!$  for  $n \in \mathbb{N}$ .

<sup>13</sup>Apparently, these are the expected values of the powers of the absolute value.

and the probable limits of the true value of  $S^{(n)}$  are

$$\frac{m\Pi_{\frac{1}{2}}^{(n-1)}}{h^n\sqrt{\pi}} \left( 1 - \rho\sqrt{\frac{2}{m} \left( \frac{\Pi(n-\frac{1}{2})\sqrt{\pi}}{(\Pi_{\frac{1}{2}}^{(n-1)})^2} - 1 \right)} \right)$$

and

$$\frac{m\Pi_{\frac{1}{2}}^{(n-1)}}{h^n\sqrt{\pi}} \left( 1 + \rho\sqrt{\frac{2}{m} \left( \frac{\Pi(n-\frac{1}{2})\sqrt{\pi}}{(\Pi_{\frac{1}{2}}^{(n-1)})^2} - 1 \right)} \right).$$

If, as above, one sets  $\frac{\rho}{h} = r$ , so that  $r$  represents the probable observational error, then, clearly, the most probable value of

$$\rho^n \sqrt{\frac{S^{(n)}\sqrt{\pi}}{m\Pi_{\frac{1}{2}}^{(n-1)}}}$$

is  $r$ ,<sup>14</sup> and the probable limits of the value of that quantity are

$$r \left( 1 - \frac{\rho}{n} \sqrt{\frac{2}{m} \left( \frac{\Pi(n-\frac{1}{2})\sqrt{\pi}}{(\Pi_{\frac{1}{2}}^{(n-1)})^2} - 1 \right)} \right)$$

and

$$r \left( 1 + \frac{\rho}{n} \sqrt{\frac{2}{m} \left( \frac{\Pi(n-\frac{1}{2})\sqrt{\pi}}{(\Pi_{\frac{1}{2}}^{(n-1)})^2} - 1 \right)} \right).$$

Hence, the odds are also one-to-one that  $r$  lies between the limits<sup>15</sup>

$$\rho^n \sqrt{\frac{S^{(n)}\sqrt{\pi}}{m\Pi_{\frac{1}{2}}^{(n-1)}}} \left( 1 - \frac{\rho}{n} \sqrt{\frac{2}{m} \left( \frac{\Pi(n-\frac{1}{2})\sqrt{\pi}}{(\Pi_{\frac{1}{2}}^{(n-1)})^2} - 1 \right)} \right)$$

and

$$\rho^n \sqrt{\frac{S^{(n)}\sqrt{\pi}}{m\Pi_{\frac{1}{2}}^{(n-1)}}} \left( 1 + \frac{\rho}{n} \sqrt{\frac{2}{m} \left( \frac{\Pi(n-\frac{1}{2})\sqrt{\pi}}{(\Pi_{\frac{1}{2}}^{(n-1)})^2} - 1 \right)} \right).$$

For  $n = 2$ , these limits are

$$\rho\sqrt{\frac{2S^{II}}{m}} \left( 1 - \frac{\rho}{\sqrt{m}} \right)$$

and

$$\rho\sqrt{\frac{2S^{II}}{m}} \left( 1 + \frac{\rho}{\sqrt{m}} \right),$$

which agree entirely with those found above (Section 4). In general, the limits for even  $n$  are

$$\rho\sqrt{2}^n \sqrt{\frac{S^{(n)}}{m \cdot 1 \cdot 3 \cdot 5 \cdots (n-1)}} \left( 1 - \frac{\rho}{n} \sqrt{\frac{2}{m} \left( \frac{(n+1)(n+3)\cdots(2n-1)}{1 \cdot 3 \cdot 5 \cdots (n-1)} - 1 \right)} \right)$$

and

$$\rho\sqrt{2}^n \sqrt{\frac{S^{(n)}}{m \cdot 1 \cdot 3 \cdot 5 \cdots (n-1)}} \left( 1 + \frac{\rho}{n} \sqrt{\frac{2}{m} \left( \frac{(n+1)(n+3)\cdots(2n-1)}{1 \cdot 3 \cdot 5 \cdots (n-1)} - 1 \right)} \right),$$

<sup>14</sup>This is correct for  $n = 1$ , but only asymptotically so for  $n = 2, 3, 4, \dots$ ; see translator's preface.

<sup>15</sup>This is the prediction-confidence interval substitution; see translator's preface.

and for odd  $n$  they are

$$\rho^n \sqrt{\frac{S^{(n)}\sqrt{\pi}}{m \cdot 1 \cdot 3 \cdot 5 \cdots \frac{1}{2}(n-1)}} \left( 1 - \frac{\rho}{n} \sqrt{\frac{1}{m} \left( \frac{1 \cdot 3 \cdot 5 \cdots (n-1)\pi}{(2 \cdot 4 \cdot 6 \cdots (n-1))^2} - 2 \right)} \right)$$

and<sup>16</sup>

$$\rho^n \sqrt{\frac{S^{(n)}\sqrt{\pi}}{m \cdot 1 \cdot 3 \cdot 5 \cdots \frac{1}{2}(n-1)}} \left( 1 + \frac{\rho}{n} \sqrt{\frac{1}{m} \left( \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)\pi}{(2 \cdot 4 \cdot 6 \cdots (n-1))^2} - 2 \right)} \right).$$

6.

I attach the numerical values for the simplest cases:

Probable limits of  $r$

- I.  $0.8453473 \frac{S^I}{m} \left( 1 \pm \frac{0.5095841}{\sqrt{m}} \right)$
- II.  $0.6744897 \sqrt{\frac{S^{II}}{m}} \left( 1 \pm \frac{0.4769363}{\sqrt{m}} \right)$
- III.  $0.5771897 \sqrt[3]{\frac{S^{III}}{m}} \left( 1 \pm \frac{0.4971987}{\sqrt{m}} \right)$
- IV.  $0.5125017 \sqrt[4]{\frac{S^{IV}}{m}} \left( 1 \pm \frac{0.5507186}{\sqrt{m}} \right)$
- V.  $0.4655532 \sqrt[5]{\frac{S^V}{m}} \left( 1 \pm \frac{0.6355080}{\sqrt{m}} \right)$
- VI.  $0.4294972 \sqrt[6]{\frac{S^{VI}}{m}} \left( 1 \pm \frac{0.7557764}{\sqrt{m}} \right)$

From this one can also see that determination method II is the most advantageous of them all. When using Formula II, one hundred observational errors give a result that is as reliable as<sup>17</sup>

- 114 using I,
- 109 using III,
- 133 using IV,
- 178 using V,
- 251 using VI.

However, Formula I has the advantage of the most convenient computation, and since it is not much less precise than II one may use this formula when one does not have the sum of the squared observational errors readily available, or wishes to know it.

7.

Even more convenient, although considerably less accurate, is the following procedure. One orders all  $m$  observational errors, absolute values taken, by their size and names the middle most, if  $m$  is odd, or the arithmetic mean of the two middle most, if  $m$  is even, by  $M$ .<sup>18</sup> It can be shown, to which

<sup>16</sup>There is possibly a misprint in the last numerator.

<sup>17</sup>In modern terminology, this property is referred to as asymptotic relative efficiency.

<sup>18</sup>In modern terminology, this is the sample median of absolute values.

in this place nothing further can be provided, that given a large number of observations the most probable value of  $M$  is  $r$ , and that the probable limits of  $M$  are<sup>19</sup>

$$r \left( 1 - e^{\rho^2} \sqrt{\frac{\pi}{8m}} \right)$$

and

$$r \left( 1 + e^{\rho^2} \sqrt{\frac{\pi}{8m}} \right);$$

or, that the probable limits of the value of  $r$  are<sup>20</sup>

$$M \left( 1 - e^{\rho^2} \sqrt{\frac{\pi}{8m}} \right)$$

and

$$M \left( 1 + e^{\rho^2} \sqrt{\frac{\pi}{8m}} \right),$$

or in numbers

$$M \left( 1 \pm \frac{0.7520974}{\sqrt{m}} \right).$$

This procedure is thus only slightly more accurate than application of Formula VI, and one must randomly draw 249 observational errors to accomplish the same as with 100 observational errors when using Formula II.

## 8.

The application of any of these methods on the 48 observational errors from *Bode's*<sup>21</sup> annual astronomical book for 1818, Page 234, of the right ascensions of the Polar star by *Bessel*,<sup>22</sup> yielded

$$\begin{aligned} S^I &= 60.''46, \\ S^{II} &= 110.600, \\ S^{III} &= 250.341118. \end{aligned}$$

From this follows the most probable value of  $r$  through

Formula I	:	1.''065,	probable uncertainty	$\pm 0.''078,$
II	:	1.024,		$\pm 0.070,$
III	:	1.001,		$\pm 0.072,$
Section 7	:	1.045,		$\pm 0.113,$

an agreement that could barely have been expected. *Bessel* got 1.'' 067 himself, and seems therefore to have computed according to Formula I.

<sup>19</sup>Although uncredited, this result is possibly of Laplace.

<sup>20</sup>This is the prediction-confidence interval substitution; see translator's preface.

<sup>21</sup>*Bode* is presumably Johann Elbert Bode (1747-1826), German astronomer.

<sup>22</sup>*Bessel* is presumably Friedrich Bessel (1784-1846), German mathematician and astronomer.

\* \* \* \*

At this occasion, I share yet another correction that should be made to *Theoria Motus Corporum Coelestium*, Pages 218 and 219. One reads, namely,

Page 218	Line 3	instead of	$e^{\frac{-h^2 s^2}{\delta'''}}$ ,	$e^{-hh'''s^2}$ ;
	Line 4	instead of	$\sqrt{\frac{1}{\delta'''}}$ ,	$\sqrt{\delta'''}$ ;
Page 219	Line 8	instead of	$\sqrt{A}, \sqrt{B'}, \sqrt{C''}, \sqrt{D'''}$ ,	$\frac{1}{\sqrt{A}}, \frac{1}{\sqrt{B'}}, \frac{1}{\sqrt{C''}}, \frac{1}{\sqrt{D'''}}$ .

These incorrectnesses were caused through the circumstance that a different terminology was used in an earlier stage of the investigation; their exchanges were in the indicated instances neglected. The numerical example is, as one sees, computed according to the corrected terminology, although one computational error is made. The equation on Page 219, Line 5 from bottom, should namely be

$$6633r = 12707 + 2P - 9Q + 123R,$$

and consequently the accuracy of  $r$ , Page 220,

$$= \sqrt{\frac{2211}{21}} = 7.34.$$

These corrections were brought to my attention by Mr. *Nicolai*,<sup>23</sup> to whom I here therefore express my most profound thanks.

UCLA DEPARTMENT OF STATISTICS, 8125 MATHEMATICAL SCIENCES BUILDING, BOX 951554, LOS ANGELES CA, 90095-1554

*E-mail address:* joakim.ekstrom@stat.ucla.edu

---

<sup>23</sup>*Nicolai* is presumably Friedrich Bernhard Gottfried Nicolai (1793-1846), German astronomer.