

Zu der von GAUSS gegebenen Begründung der Methode der kleinsten Quadrate

Von K.LIEBSCHER und D-E.LIEBSCHER, Potsdam-Babelsberg
Die Sterne **53** (1977), 15-21¹.

GAUSS hat sich während seines ganzen Lebens mit astronomischen und geodätischen Problemen befaßt. Die Methode der kleinsten Quadrate, die er dabei zur Ausgleichung der Meß- und Beobachtungsergebnisse benutzte, ist noch heute die Grundlage der gesamten Fehler- und Ausgleichsrechnung, und ihre Prinzipien werden auch heute noch in der Theorie der Schätzverfahren befolgt. GAUSS rechnete mit ihr bereits in seiner Jugendzeit (seit 1794) und empfand die Methode derart als natürlich, daß er sie um der Bekanntgabe der Technik allein nicht veröffentlichte. GAUSS zögerte mit der Veröffentlichung selbst dann noch, als die Methode ihre erste Bewährungsprobe bei der Berechnung der Bahnelemente der Ceres, die PIAZZI 1801 gefunden hatte, und der Wiederauffindung dieses Planetoiden glänzend bestanden hatte. So geschah es, daß während der Abfassung der „Theorie der Bewegung der Himmelskörper“ [2] ein Buch von A.M.LEGENDRE (1752-1838) über die Bestimmung von Kometenbahnen erschien, in dem die Methode der kleinsten Quadrate vorgeführt und erstmalig ihr Name geprägt wurde (1806). LEGENDRE bemühte sich aber nicht um eine tiefergehende wahrscheinlichkeitstheoretische Begründung, sondern führte als Hauptargument die Zweckmäßigkeit und Einfachheit der Methode an. GAUSS hatte bereits längere Zeit vorher die Notwendigkeit der Gründung der Methode auf die Theorie der Wahrscheinlichkeit gesehen. Dies geschah in der oben zitierten Arbeit, die 1809 erschien.

GAUSS hat die Methode der kleinsten Quadrate in zwei weiteren Abhandlungen [3, 4], die 1821 und 1828 erschienen, noch tiefer begründet und ausgebaut. Er schätzte in späteren Jahren seine Methode der kleinsten Quadrate selbst sehr hoch ein. GAUSS, der nicht gern Vorlesungen hielt, kündigte noch am liebsten die „Methode der kleinsten Quadrate“ als Vorlesungsgegenstand an. 1829 übernahm GAUSS diese Methode auch zur Formulierung eines Differentialprinzips der Mechanik², seines „Prinzips des kleinsten Zwanges“ [5]. Dort schrieb er: „Es ist sehr merkwürdig, daß die freien Bewegungen, wenn sie mit den notwendigen Bedingungen nicht bestehen können, von der Natur gerade auf dieselbe Art modificirt werden, wie der rechnende Mathematiker, nach der Methode der kleinsten Quadrate, Erfahrungen ausgleicht, die sich auf unter einander durch notwendige Abhängigkeit verknüpfte Größen beziehen.“

Wir wollen hier zunächst das Grundproblem der Ausgleichsrechnung formulieren, um die von GAUSS gegebenen Begründungen der Methode der kleinsten Quadrate daran anschließend diskutieren zu können. Wir benutzten dabei die in den heutigen Lehrbüchern der Statistik verwendeten Bezeichnungen, um die Konsultation dieser Bücher zu erleichtern.

Wir setzen voraus, ein zu vermessendes Objekt sei gegeben, das unter bestimmten Bedingungen x einen Meßwert $y = f[x, \theta]$ liefern sollte. Mit $\theta = \theta_1, \dots, \theta_m$ bezeichnen wir die zunächst unbekannt Parameter des Systems. Führen wir gleichviel Messungen durch, wie Parameter zu bestimmen sind, so lösen wir diese Aufgabe durch Umkehrung des Gleichungssystems

$$y_i = f[x_i; \theta_1, \dots, \theta_m]; \quad i = 1, \dots, m. \quad (1)$$

Führen wir weitere Messungen durch, dann entdecken wir, daß das erweiterte Gleichungssystem unlösbar wird, da offenbar Fehler in die Meßwerte eingehen. Diese Fehler sind auch in den Messungen zum System (1) enthalten, dort stören sie aber noch nicht die formale Lösbarkeit des Gleichungssystems. Jetzt ist jedoch das Gleichungssystem

$$y_i = f[x_i; \theta] + \varepsilon_i; \quad i = 1, \dots, n; \quad m < n. \quad (2)$$

nur lösbar, wenn wir Fehler $\varepsilon_i \neq 0$ explizit zulassen. Evident gehört — bei gegebenen x_i und y_i — zu jedem Parametersatz $\theta_1, \dots, \theta_m$ ein Satz Fehler $\varepsilon_1, \dots, \varepsilon_m$. Die Aufgabe der Ausgleichsrechnung ist nun, die unbekannt Parameter derart zu bestimmen, daß die Fehler ε so wenig nachteilig als möglich ausfallen.

Wir haben also ein Optimierungsproblem zu lösen. Der prinzipielle Teil eines solchen Problems besteht in der Angabe der Zielfunktion, in unserem Falle eines Bewertungsmaßstabes des n -komponentigen Fehlers

¹Gegenüber der Originalarbeit sind bestimmte Notationen verändert. Die Operatoren E, D² und d sind steil gesetzt, Variablen und Variablenlisten stehen in eckigen Klammern, um Verwechslungen mit Multiplikationen zu vermeiden, und die Exponentialfunktion ist immer als exp[] dargestellt

²Nebenbedingungen schränken die Bewegung eines n -Teilchen-Systems derart ein, daß

$$Z = \sum_{k=1}^n m_k \left(\ddot{\mathbf{r}}_k - \frac{1}{m_k} \mathbf{F}_k \right)^2$$

bei erlaubten Variationen der Beschleunigung ein Minimum annimmt.

ε , deren Extremum zu suchen den technischen Teil des Problems darstellt. Die Methode der kleinsten Quadrate bewertet die Fehler mit der Summe ihrer Quadrate³:

$$S = \sum_{i=1}^n \varepsilon_i^2 \quad (3)$$

Dadurch wird eine wichtige technische Vereinfachung erzielt: Ist die Funktion f in den Parametern θ linear,

$$f[x, \theta] = \sum_{k=1}^m \theta_k g_k[x], \quad (4)$$

so erhalten wir lineare Bestimmungsgleichungen für die Schätzwerte θ der Parameter. Die Zielgröße

$$S = \sum_{i=1}^n \left(y_i - \sum_{k=1}^m \theta_k g_k[x_i] \right)^2 \quad (5)$$

wird minimal an der Stelle $\hat{\theta}$, wenn

$$-\frac{\partial S}{\partial \theta_l} \Big|_{\theta=\hat{\theta}} = \sum_{k=1}^m \hat{\theta}_k \sum_{i=1}^n g_l[x_i] g_k[x_i] - \sum_{i=1}^n y_i g_l[x_i] = 0. \quad (6)$$

Das ist ein System linearer Gleichungen für θ , die unter der Bezeichnung „Normalgleichungen“ bekannt sind.

GAUSS hat die derart sichtbare Zweckmäßigkeit der Wahl (3) von S schon früh erkannt und sie gegen die von P.S.DE LAPLACE (1749-1821) bevorzugte Wahl

$$S = \sum_{i=1}^n |\varepsilon_i| \quad (7)$$

verteidigt. GAUSS sah dabei auch, daß das Kriterium (7) in einfachen Fällen — speziell für das Gleichungssystem $y_i = \theta + \varepsilon_i$ — keine eindeutige Lösung liefert. Darüber hinaus kann (7) unstetig in θ sein. GAUSS war dennoch der Überzeugung, daß der eigentliche Vorteil der Wahl (3) in der wahrscheinlichkeitstheoretischen Begründung liegen muß.

In der Arbeit über die Theorie der Bewegung der Himmelskörper [21] behandelte GAUSS die Fehler als zufällige Größen, denen eine relative Wahrscheinlichkeit zugerechnet werden kann. Wir beschreiben diese durch eine Wahrscheinlichkeitsdichte deren Integral die Wahrscheinlichkeit angeben soll, daß der Fehler zwischen den Integrationsgrenzen liegt:

$$P[\varepsilon_1 < \varepsilon < \varepsilon_2] = \int_{\varepsilon_1}^{\varepsilon_2} \varrho[\varepsilon] d\varepsilon. \quad (8)$$

Zu einer gegebenen Meßreihe M erhält man die gemeinsame Wahrscheinlichkeit der Fehler als Funktion der Parameter, da die Fehler selbst mittels (2) eine Funktion dieser Parameter sind. Dargestellt als Funktion der Parameter heißt die Wahrscheinlichkeit der Fehler „Likelihood-Funktion“ $L[M, \theta]$. GAUSS zeigte nun, daß $L[\theta]$ — die Wahrscheinlichkeit der Fehler — auch als Wahrscheinlichkeit der Parameter θ selbst interpretiert werden kann. Zu diesem Zweck leitete er die heute nach BAYES (gest. 1763) benannte Formel ab. Ist die Wahrscheinlichkeit zweier hypothetischer Parametersätze θ_A und θ_B a priori (d. h. vor der Messung) gleich, so verhalten sich ihre Wahrscheinlichkeiten a posteriori (d. h. nach Kenntnis des Ergebnisses der Messung) wie die Wahrscheinlichkeiten von M unter den beiden Hypothesen. Das Meßergebnis M hat unter den beiden konkurrierenden Hypothesen verschiedene Wahrscheinlichkeiten, die in unserem Fall durch die Wahrscheinlichkeiten der sich ergebenden Fehler definiert sind, d. h. durch die Likelihood-Funktion. Bezeichnen wir mit $P[\theta]$ die Wahrscheinlichkeit a priori des Parametersatzes θ , mit $P[\theta|M]$ die Wahrscheinlichkeit a posteriori, d. h. unter der Bedingung, daß M gemessen worden ist, und mit $P[M|\theta] = L[M, \theta]$ die Wahrscheinlichkeit des Meßergebnisses bei zutreffenden Parametern θ , dann lautet die BAYESSche Formel

$$\frac{P[\theta_A|M]}{P[\theta_B|M]} = \frac{P[\theta_A]P[M|\theta_A]}{P[\theta_B]P[M|\theta_B]}. \quad (9)$$

³Formel (3) wird für unabhängige Fehler gleichen Gewichts verwendet. Sie kann ohne Verlust der im folgenden abgeleiteten Eigenschaft der Erhaltung der Linearität auf untereinander korrelierte Fehler verschiedenen Gewichts verallgemeinert werden. Ausführliche Darstellungen der technischen Seite findet man in [1], [9] und [12].

GAUSS setzte in der zitierten Arbeit die a-priori-Wahrscheinlichkeiten als gleich an und fand, daß die Wahrscheinlichkeit a posteriori proportional der Wahrscheinlichkeit des Meßergebnisses ist. Die Schätzung größter Wahrscheinlichkeit wird somit durch das Maximum der Likelihood-Funktion gegeben⁴.

Sind die einzelnen Fehler unabhängige Größen, dann ist die Likelihood-Funktion das Produkt der Wahrscheinlichkeiten der einzelnen Fehler, in unserem Falle also

$$L[y, \theta] = \prod_{i=1}^n \varrho_i[\varepsilon_i] = \prod_{i=1}^n \varrho_i[y_i - f_i[x_i, \theta]] . \quad (10)$$

GAUSS hat natürlich gesehen, daß die Schätzung nach der größten Wahrscheinlichkeit — die Maximum-Likelihood-Schätzung — und die Schätzung nach der Methode der kleinsten Quadrate — die er ja begründen wollte — nur dann zusammenfallen, wenn L eine Funktion der Summe der Fehlerquadrate ist. Der einzelne Fehler ist dann mit einer Wahrscheinlichkeitsdichte der Form

$$\varrho[\varepsilon] = C_1 \exp[C_2 x^2] ; \quad C_1, C_2 \text{ konstant} \quad (11)$$

verteilt. GAUSS gab deshalb eine Begründung für die Natürlichkeit dieser Verteilung.

GAUSS postulierte zu diesem Zweck, daß das arithmetische Mittel $\bar{y} = \frac{1}{n} \sum_i y_i$ in der Aufgabe $y_i = \mu + \varepsilon_i$ immer Maximum-Likelihood-Schätzung des Erwartungswertes μ ist, unabhängig von der Meßreihe y_1, \dots, y_n . Dann folgt: Aus der Wahrscheinlichkeitsdichte der Einzelmessung

$$\varrho[\varepsilon] = \varrho[y - \mu]$$

ergibt sich die Likelihood-Funktion zu

$$L[y, \mu] = \prod_{i=1}^n \varrho[y_i - \mu] ,$$

und die Bedingungsgleichung für ihr Maximum lautet

$$\frac{1}{L} \frac{\partial L}{\partial \mu} \Big|_{\mu=\hat{\mu}} = - \sum_{i=1}^n \frac{\varrho'[y_i - \hat{\mu}]}{\varrho[y_i - \hat{\mu}]} . \quad (12)$$

Laut Voraussetzung soll $\hat{\mu}$ eine Lösung sein, wenn für y_i die Werte $y_1 = 0$ und $y_2 = \dots = y_n = \frac{n}{n-1} \mu$ eingesetzt werden. Wir nennen $\varphi = \frac{\varrho'}{\varrho}$ und finden

$$-\varphi[-\hat{\mu}] = (n-1) \varphi\left[\frac{\hat{\mu}}{n-1}\right] . \quad (13)$$

φ ist daher eine homogene lineare Funktion

$$\varphi = -k\varepsilon ,$$

und die Wahrscheinlichkeitsdichte ϱ lautet

$$\varrho[\varepsilon] = \sqrt{\frac{k}{2\pi}} \exp\left[-\frac{kx^2}{2}\right] , \quad (14)$$

wobei der konstante Faktor durch die Normierung der Gesamtwahrscheinlichkeit bestimmt wird. Dies ist die GAUSSsche Ableitung der Normalverteilung.

Die Herleitung der Methode der kleinsten Quadrate als Schätzung nach der höchsten Wahrscheinlichkeit hat GAUSS aus eben dem Grunde für unzureichend gehalten, weil sie eine spezielle, wenn auch ausgezeichnete Verteilung der Fehler, die Normalverteilung voraussetzt. GAUSS ist deshalb später in der Arbeit über die Theorie der Beobachtungsfehler [3] anders vorgegangen. Er postulierte dort, daß die Schätzung kleinster Varianz

$$D^2 \hat{\theta} = E(\hat{\theta}[y] - \theta)^2 = \int (\hat{\theta}[y] - \theta)^2 \varrho[y - f[x, \theta]] dy \quad (15)$$

⁴Streng genommen ist die „Wahrscheinlichkeit“ der Parameterwerte nicht auf einer Stufe mit der Wahrscheinlichkeit der Fehler behandelbar, da die Parameter zunächst keine zufälligen, sondern nur unbekannte Größen sind. Weil Formel (9) nur abgeleitet werden kann, wenn θ und M Ereignisse derselben Grundgesamtheit sind, ist die obige Darstellung der Bedeutung der Maximum-Likelihood-Schätzung nur eine Interpretation der Formel (9).

das „am wenigsten nachteilige Spiel“ bedeutet und daß deshalb die Schätzfunktion $\hat{\theta}[y]$ so gewählt werden muß, daß (15) ein Minimum annimmt. GAUSS zeigte nun, daß unabhängig von der Verteilung der Fehler die Summe der Quadrate eine erwartungstreu und konsistente Schätzung der Varianz liefert⁵. Er definierte

$$M^2 = Ey^2 = \int y^2 \varrho[y] dy$$

und

$$N^4 = Ey^4 = \int y^4 \varrho[y] dy$$

und bewies

$$E\left(\frac{\sum_{i=1}^n y^2}{n} - M^2\right)^2 = \frac{N^4 - M^4}{n} \xrightarrow{n \rightarrow \infty} 0. \quad (16)$$

Damit ist die Schätzung nach der Methode der kleinsten Quadrate als Näherung der Schätzung nach der kleinsten Varianz charakterisiert.

GAUSS war sich nach wie vor der Tatsache bewußt, daß er hiermit die Wahl von (3) nur auf die Wahl eines anderen, allerdings allgemeineren Kriteriums (15) zurückgeführt hat, daß aber auch dort eine Willkür bleibt, die, um mit seinen Worten zu schreiben, die Bestimmung des Maßes für „den Nachteil im Spiel“ mit der Natur ausmacht.

Wir wissen heute, daß die Schätzung nach der kleinsten Varianz in dem Fall, den GAUSS behandeln konnte, der Schätzung nach der größten Wahrscheinlichkeit entspricht. GAUSS hat mit (16) das asymptotische Verhalten für große n im Auge. Für diesen Grenzfall gelten folgende Sätze über die Maximum-Likelihood-Schätzungen:

1. Maxima von L stellen konsistente Schätzungen dar.
2. Konsistente Lösungen der Likelihood-Gleichungen $\frac{\partial L}{\partial \theta} = 0$ sind asymptotisch effektiv, d.h., es gibt keine anderen Schätzfunktionen, die asymptotisch eine kleinere Varianz als die Maximum-Likelihood-Schätzungen aufweisen.
3. Konsistente Lösungen der Likelihood-Gleichungen liefern Schätzfunktionen $\hat{\theta}[y]$, die asymptotisch (d.h. für große n) normalverteilt sind.

Damit ist der Kreis geschlossen. Die Schätzung nach der kleinsten Varianz ist in dem Maße auch Schätzung größter Wahrscheinlichkeit, wie durch die große Anzahl der Messungen Normalverteilungen entstehen. Die Normalverteilung spielt eine ausgezeichnete Rolle auch dann, wenn der einzelne Meßfehler nicht normalverteilt ist. Das ist aber nur die eine Seite. Auch der einzelne Meßfehler ist in erstaunlich guter Näherung normalverteilt. Das liegt an weiteren speziellen Eigenschaften der GAUSSschen Normalverteilung, die wir nun noch kurz betrachten wollen.

Ohne uns auf das Problem der Ausgleichsrechnung und der Schätzverfahren zu beziehen, stellen wir die besondere Rolle der Normalverteilung in Hinsicht auf die Entropie und das Grenzverhalten der Verteilung von Summen fest. Wir wollen hier nicht genauer die Begriffe Information und Entropie erläutern⁶ und nur soviel andeuten, daß die Normalverteilung in einem wohldefinierten Sinn diejenige Verteilung ist, die die wenigsten besonderen Strukturen aufweist. Sind Erwartungswert

$$Ey = \int y \varrho[y] dy = \mu \quad (17)$$

und Streuung

$$D^2y = \int (y - \mu)^2 \varrho[y] dy = \sigma^2 \quad (18)$$

gegeben, so nimmt die Entropie (= mittlerer Informationsgehalt)

$$S = \int \varrho[y] \ln \frac{1}{\varrho[y]} dy \quad (19)$$

für die Normalverteilung

$$\varrho[y] = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(y - \mu)^2}{2\sigma^2}\right]$$

⁵Eine Schätzfunktion $\hat{\theta}[y]$ für θ heißt konsistent, wenn sie für große n erwartungstreu ($\lim_{n \rightarrow \infty} E\hat{\theta} = \theta$) und streungsfrei ($\lim_{n \rightarrow \infty} E(\hat{\theta} - \theta)^2 = 0$) ist.

⁶Eine Einführung in den Problemkreis findet man bei Renyi [10], eine ausführliche Darstellung vom Standpunkt der mathematischen Statistik bei KULLBACK [8].

ein Maximum an⁷.

Hinsichtlich der Ausgleichsrechnung ist die Auszeichnung der Normalverteilung als Grenzverteilung von Summen zufälliger Größen von noch augenscheinlicherer Bedeutung. Schon A. DE MOIVRE (1667-1754) stellte fest, daß die in der Binominalverteilung

$$P[X = m] = \binom{n}{m} p^m (1-p)^{n-m} \quad (20)$$

auftretenden Wahrscheinlichkeiten durch die Funktion

$$P[x = m] = C_1 \exp\left[-\frac{(m - np)^2}{2np(1-p)}\right]$$

genähert werden [11]. Er leitete dies speziell für $p = \frac{1}{2}$ ab. LAPLACE verallgemeinerte das Ergebnis von MOIVRE und präziserte es. Der entstehende Grenzwertsatz ist deshalb nach MOIVRE und LAPLACE benannt:

Ist eine Zufallsgröße gegeben, die den Wert 0 mit der Wahrscheinlichkeit p und den Wert 1 mit der Wahrscheinlichkeit $q = 1 - p$ annimmt, dann ist die Wahrscheinlichkeit, daß bei n Versuchen in m Fällen der Wert 0 erscheint, durch (20) gegeben. Die Größe $x = \frac{m - np}{\sqrt{npq}}$ ist dann asymptotisch normalverteilt, d. h. es konvergiert die Wahrscheinlichkeit

$$P[m_1 < m < m_2] \xrightarrow{n \rightarrow \infty} \int_{\frac{m_1 - np}{\sqrt{npq}}}^{\frac{m_2 - np}{\sqrt{npq}}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right] dx . \quad (21)$$

Die Normalverteilung stellt sich hier als Grenzfall der Verteilung einer Summe unabhängiger Zufallsgrößen heraus, deren einzelne Verteilung nichts Gemeinsames mit der Normalverteilung hat außer der Endlichkeit der Streuung und des Erwartungswerts. Die Besonderheiten der Verteilung der einzelnen Summanden werden in der Summe bedeutungslos. Es entsteht eine Verteilung mit den wenigsten Besonderheiten, es entsteht die Normalverteilung.

Die Normalverteilung ergibt sich in einer großen Zahl von Fällen⁸. Man könnte sagen: im Normalfall, denn es müssen viel einschränkendere Bedingungen an die Verteilungen der Teilsommen gestellt werden, wenn andere Grenzverteilungen entstehen sollen. Diese barmacht der Normalverteilung ist der Grund dafür, daß im allgemeinen mit Erfolg die Normalverteilung der Fehler vorausgesetzt und die Methode der kleinsten Quadrate zur Schätzung angewandt werden kann.

Die Fehler können als normalverteilt angesehen werden, wenn sie von sehr vielen verschiedenen Mikroprozessen verursacht werden. Je mehr Faktoren der Fehler enthält, desto sicherer ist er normalverteilt. Hieran sollte man denken, wenn man bei der Betrachtung der Meßergebnisse geneigt ist, sogenannte herausfallende Werte aus dem Protokoll zu streichen, solange man nicht sicher weiß, daß ihnen Versuchsbedingungen zugrunde lagen, die bei den übrigen Meßwerten nicht eingetreten sind. Die einfache Tatsache, daß Meßwerte a posteriori herausfallen, berechtigt nicht zu ihrer Streichung. Selbst wenn man einen Faktor findet, der ihre extreme Stellung begünstigt haben könnte, ist das eher ein Argument für ihr Verbleiben: Schließlich entsteht die Normalverteilung über viele solcher Faktoren. Schon GAUSS hat sich — gestützt auf die wahrscheinlichkeitstheoretische Begründung der Methode der kleinsten Quadrate — gegen das Verfahren der Streichung herausfallender Meßwerte gewandt. In den Göttinger Gelehrten Anzeigen vom 25. 9. 1826 [6] schrieb er über die Grundsätze der Methode im Zusammenhang mit trigonometrischen Messungen: „Allein sowohl die sichere Würdigung, als die vollkommenste Benutzung der Messungen ist nur dann möglich, wenn sie in reiner Authentizität und Vollständigkeit vorliegen, und es wäre daher sehr zu wünschen, dass alle grösseren auf besondere Genauigkeit Anspruch machenden Messungen dieser Art immer mit aller nöthigen Ausführlichkeit bekannt gemacht werden möchten. Nur zu gewöhnlich ist das Gegentheil, wo nur Endresultate für die einzelnen gemessenen Winkel mitgeteilt werden. Wenn solche Endresultate nach richtigen Grundsätzen gebildet werden, indem man durchaus alle einzelnen Beobachtungsreihen, die nicht einen durchaus unstatthaften Fehler gewiss enthalten, dazu concurriren läßt, so ist der Nachtheil freilich lange nicht so gross, als wenn man etwa nur diejenigen Reihen beibehält, die am besten zu den naheliegenden Prüfungsmitteln passen, welche die Summen der Winkel jedes Dreiecks und die Summen der Horizontalwinkel um jeden Punkt herum darbieten. Wo dies

⁷In der Thermodynamik begegnet uns die Normalverteilung in der Form der Maxwellschen Geschwindigkeitsverteilung als Gleichgewichtsverteilung bei gegebener mittlerer Energie. Die Bedingung fester mittlerer Energie entspricht der Bedingung (18), weil die Energie eine quadratische Funktion der Geschwindigkeiten ist.

⁸Eine ausführliche Darstellung der Grenzwertsätze findet man bei GNEDENKO und KOLMOGOROFF [7].

durchaus verwerfliche Verfahren angewandt ist, sei es aus Unbekanntschaft mit den wahren Grundsätzen einer richtigen Theorie, oder aus dem geheimen Wunsche, den Messungen das Ansehen grösserer Genauigkeit zu geben, geht der Maassstab zu einer gerechten Würdigung der Beobachtungen und der aus ihnen abzuleitenden Resultate verloren.“

Literatur

- [1] AHRENS, H., Varianzanalyse, 2. Aufl., Berlin 1968.
- [2] GAUSS, C. F., *Theoria motus corporum coelestium in sectionibus conicis Solem ambientium* (1809), Werke VII, Göttingen 1871.
- [3] GAUSS, C. F., *Theoria combinationis observationum erroribus minimis obnoxiae* (1821), Werke IV. Göttingen, 1880.
- [4] GAUSS, C. F., *Supplementum theoriae combinationis observationum erroribus minimis obnoxiae* (1826), Werke IV, Göttingen 1880.
- [5] GAUSS, C. F., *ber ein neues allgemeines Grundgesetz der Mechanik* (1829), Werke V, Göttingen 1867.
- [6] GAUSS, C. F., Werke IV, Göttingen 1880, Seite 107.
- [7] GNEDENKO, B. W., A. N. KOLOGOROFF, *Grenzverteilungen von Summen unabhängiger Zufallsgrößen*, Übers. a. d. Russ., 2. Aufl., Berlin 1960.
- [8] KULLBACK, S., *Information theory and statistics*, 2. Aufl., New York 1959.
- [9] MOOD, A. M., F. A. GRAYBILL, *Introduction to the theory of statistics*, New York 1963.
- [10] RÉNYI, A., *Wahrscheinlichkeitsrechnung*, 2. Aufl., Berlin 1966,
- [11] TODHUNTER, L., *A history of the mathematical theory of probability*, Cambridge 1865.
- [12] WOLF, H., *Ausgleichsrechnung nach der Methode der kleinsten Quadrate*, Bonn 1968.