

X.

Über die Bestimmung der Präzision einer Beobachtungsmethode nach der Methode der kleinsten Quadrate.

[Vierteljahrsschrift der naturforschenden Gesellschaft in Zürich, 1860, S. 76—83.]

In seiner ersten Begründung der Methode der kleinsten Quadrate ging Gauß (Theoria motus corp. coel.) von der Voraussetzung aus, daß der wahrscheinlichste Wert einer beliebig oft auf dieselbe Weise direkt gemessenen Größe das arithmetische Mittel aus den durch diese Messungen erhaltenen Werten ist, und kam auf diese Weise zu dem Ausdruck

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

für die Wahrscheinlichkeit, daß ein Beobachtungsfehler seinem Werte nach in dem unendlich kleinen Intervall zwischen t und $t + dt$ liegt; in diesem Ausdruck bedeutet h eine positive Konstante, welche für verschiedene Beobachtungsmethoden im allgemeinen auch verschiedene Werte hat, und zwar leuchtet ein, daß eine Beobachtungsmethode desto zuverlässiger ist, je größer der Wert der ihr zugehörigen Konstante h ist; denn die Wahrscheinlichkeit

$$\frac{h}{\sqrt{\pi}} \int_{-a}^{+a} e^{-h^2 t^2} dt = \frac{1}{\sqrt{\pi}} \int_{-ha}^{+ha} e^{-u^2} du$$

dafür, daß ein Fehler seinem absoluten Werte nach die positive Größe a nicht überschreitet, ist desto größer, je größer h ist. Aus diesem Grunde hat Gauß die Größe h die Präzision der Beobachtungsmethode genannt; in einer späteren Abhandlung (Zeitschrift für Astronomie usw. von Lindenau und Bohnenberger, Bd. I, 1816) hat er ferner gezeigt, wie man den wahrscheinlichsten Wert der Präzision einer Beobachtungsmethode bestimmen kann, wenn eine Reihe wirklich gemachter Beobachtungsfehler bekannt ist. Es wird für das Folgende nützlich sein, hier den von Gauß zu diesem Zwecke eingeschlagenen Weg wieder in Erinnerung zu bringen, welcher auf dem Satze über die Wahrscheinlichkeit a posteriori beruht.

Ist h die wahre Präzision der Beobachtungsmethode, so ist die Wahrscheinlichkeit, daß bei m aufeinanderfolgenden Beobachtungen die Fehler

t_1, t_2, \dots, t_m
gemacht werden, gleich

$$\alpha = \left(\frac{h}{\sqrt{\pi}} \right)^m e^{-h^2 S} dt_1 dt_2 \dots dt_m,$$

worin zur Abkürzung

$$S = t_1^2 + t_2^2 + \dots + t_m^2$$

gesetzt ist. A priori, d. h. ehe irgend eine Messung vorgenommen ist, haben wir keinen Grund, der Präzision einer uns unbekanntem Beobachtungsmethode einen Wert h eher beizulegen als einen anderen; folglich ist a posteriori, d. h. nachdem wirklich die Beobachtungsfehler t_1, t_2, \dots, t_m gemacht sind, die Wahrscheinlichkeit der Hypothese, daß h der wahre Wert der Präzision ist, proportional dem α , also proportional dem Ausdruck

$$h^m e^{-h^2 S},$$

welcher für

$$\frac{1}{2h^3} = \frac{S}{m}, \text{ also } h = \sqrt{\frac{m}{2S}}$$

ein Maximum wird; es ist also dies der wahrscheinlichste Wert der Präzision der Beobachtungsmethode.

In allen wirklichen Fällen liegt aber die Sache ganz anders. Die Objekte der Beobachtungen sind lineare Funktionen

$$v_1, v_2, \dots, v_m$$

von gewissen unbekanntem Größen $x, y, z \dots$, deren Anzahl n höchstens gleich der Anzahl m der Beobachtungen und deren Wertbestimmung gerade der Zweck dieser Beobachtungen ist. Sind nun

$$k_1, k_2, \dots, k_m$$

die durch die Beobachtungen gelieferten Werte von v_1, v_2, \dots, v_m , so bestimmt die aus dem obigen Wahrscheinlichkeitsgesetz eines beliebigen Fehlers t gefolgerte Methode der kleinsten Quadrate die Werte der Unbekanntem $x, y, z \dots$ durch die Forderung, daß die Quadratsumme

$$(k_1 - v_1)^2 + (k_2 - v_2)^2 + \dots + (k_m - v_m)^2 = \Omega$$

ein Minimum werden soll. Wären nun diese wirklich die wahren Werte der Unbekanntem, so wären die entsprechenden Werte der Differenzen

$$k_1 - v_1, k_2 - v_2, \dots, k_m - v_m$$

auch die wahren Beobachtungsfehler, und man könnte versucht sein, den wahrscheinlichsten Wert der Präzision h nach der früheren Regel zu bestimmen, indem man statt S nur das Minimum Ω_0 der Funktion Ω zu substituieren brauchte, so daß also

$$\sqrt{\frac{m}{2 \Omega_0}}$$

als wahrscheinlichster Wert von h anzusehen wäre. Daß diese Formel aber nicht richtig sein kann, bemerkt man am deutlichsten in dem Falle, wo $n = m$ ist; dann können nämlich die gemachten Beobachtungen sämtlich durch ein und dasselbe Wertsystem x, y, z, \dots befriedigt werden, Ω_0 ist $= 0$, und man würde $h = \infty$, also das Resultat erhalten, daß die Beobachtungsmethode höchstwahrscheinlich absolut genau ist, während doch erst dann ein Urteil über die Präzision gestattet ist, wenn ein Überschuß von Beobachtungen vorliegt.

In einer späteren Abhandlung (Theoria combinationis etc. art. 39) in welcher das Prinzip des arithmetischen Mittels und damit zugleich das obige Wahrscheinlichkeitsgesetz eines Fehlers t ganz verlassen ist, hat Gauß für eine ähnliche Frage (die nach dem wahrscheinlichsten Werte des sogenannten mittleren Fehlers) die richtige Antwort gegeben, welche, auf die frühere Darstellungsweise übertragen, den Ausdruck

$$\sqrt{\frac{m-n}{2 \Omega_0}}$$

als wahrscheinlichsten Wert der Präzision h liefert, so daß also das Minimum Ω_0 als eine Summe von nur $(m - n)$ Fehlerquadraten zu behandeln ist. Man sieht, daß diese Formel in dem Falle $n = m$ unter die ganz unbestimmte Form $\frac{0}{0}$ tritt, und in der Tat ist in diesem Falle gar kein Schluß auf die Präzision gestattet.

Es erscheint nun wünschenswert, einen Beweis dieses Satzes auch aus dem obigen Wahrscheinlichkeitsgesetz abzuleiten, da dies meines Wissens in befriedigender Weise noch nicht geschehen ist*). Dazu führt folgender einfache Weg.

*) So z. B. geht Wittstein (Anhang zu der Übersetzung von Naviers Differentialrechnung) von dem unrichtigen Satze aus, daß, wenn h die wahre Präzision ist, der wahrscheinlichste Wert eines Fehlerquadrates $= \frac{1}{2h^2}$, statt 0 ist.

In der Hypothese B , daß h, x, y, z, \dots die wahren Werte der Präzision, der ersten, zweiten, dritten usw. Unbekannten sind, ist die Wahrscheinlichkeit, daß für die Funktionen

$$v_1, v_2, \dots v_m$$

die Werte

$$k_1, k_2, \dots k_m$$

durch Beobachtung geliefert, daß also die Beobachtungsfehler

$$k_1 - v_1, k_2 - v_2, \dots k_m - v_m$$

gemacht werden, proportional dem Ausdruck

$$h^m e^{-h^2 \Omega};$$

da nun alle denkbaren Hypothesen B a priori gleich wahrscheinlich sind, so ist a posteriori, d. h. nachdem wirklich die Werte $k_1, k_2, \dots k_m$ beobachtet sind, die Wahrscheinlichkeit der Hypothese B proportional demselben Ausdruck; dieselbe ist daher

$$= C h^m e^{-h^2 \Omega} dh dx dy dz \dots,$$

worin

$$\frac{1}{C} = \int_0^{\infty} dh \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz \dots h^m e^{-h^2 \Omega}.$$

Fragt man nun nach dem wahrscheinlichsten Wertsystem von h, x, y, z, \dots , so würde man untersuchen müssen, für welche Werte h, x, y, z, \dots der Ausdruck

$$h^m e^{-h^2 \Omega}$$

ein Maximum wird. Allein wir fragen nach dem wahrscheinlichsten Werte der Präzision allein; wir haben daher zunächst den Ausdruck der Wahrscheinlichkeit herzustellen, daß der Wert der Präzision zwischen h und $h + dh$ liegt. Diesen erhält man aus dem Vorhergehenden durch Integration über alle reellen Werte von x, y, z, \dots . Es ist aber nach bekannten Sätzen

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz \dots e^{-h^2 \Omega} = K \frac{1}{h^n} e^{-h^2 \Omega},$$

worin K von h unabhängig ist; folglich ist das aus den gemachten Beobachtungen resultierende Wahrscheinlichkeitsgesetz für die Präzision von der Form

$$H \cdot h^{m-n} e^{-h^2 \Omega} dh,$$

worin

$$\frac{1}{H} = \int_0^{\infty} h^{m-n} e^{-h^2 \Omega_0} dh$$

ist. Vergleicht man diese Form mit der früheren

$$H' h^m e^{-h^2 S} dh, \quad \text{wo} \quad \frac{1}{H'} = \int_0^{\infty} h^m e^{-h^2 S} dh,$$

welche sich ergab, wenn m wahre Beobachtungsfehler vorlagen, deren Quadratsumme $= S$ war, so findet man in der Tat vollständige Übereinstimmung, wenn man das Minimum Ω_0 der Summe von m Fehlerquadraten wie eine Summe von $m - n$ wirklichen Fehlerquadraten ansieht. Der wahrscheinlichste Wert zu der Präzision ist daher wirklich

$$= \sqrt{\frac{m - n}{2 \Omega_0}}.$$

Hiermit ist der eigentliche Gegenstand dieser Mitteilung beendigt; zum Schluß mag noch folgende Bemerkung gemacht werden. Wir haben als wahrscheinlichsten Wert h einen anderen gefunden, als denjenigen, welcher dem h in dem wahrscheinlichsten System von Werten h, x, y, z, \dots zukommt. Man könnte nun befürchten, daß auch die Bestimmung der wahrscheinlichsten Werte von x, y, z, \dots , wenn sie nach demselben Prinzip ausgeführt, wenn also für jede einzelne Unbekannte besonders der wahrscheinlichste Wert aufgesucht würde, von der durch die Methode der kleinsten Quadrate geforderten Regel abweichen könnte. Allein man überzeugt sich leicht, daß diese Befürchtung ungegründet ist, und daß das System der wahrscheinlichsten Werte von x, y, z, \dots übereinstimmt mit dem wahrscheinlichsten Wertsystem dieser Unbekannten.

Das letztere ist offenbar dasjenige, für welches die Quadratsumme Ω ein Minimum wird, und darin besteht ja gerade der Hauptsatz der Methode der kleinsten Quadrate; die entsprechenden Werte der n Unbekannten x, y, z, \dots findet man bekanntlich dadurch, daß man, was immer möglich ist, die Funktion Ω auf die Form

$$\Omega = Y^2 + Z^2 + \dots + X^2 + \Omega_0$$

bringt, worin Y eine lineare Funktion aller n Unbekannten ist, die dadurch bestimmt wird, daß $\Omega - Y^2$ unabhängig von y wird; ähnlich

ist Z eine lineare Funktion der übrigen $(n - 1)$ Unbekannten, und dadurch bestimmt, daß $\Omega - Y^2 - Z^2$ unabhängig von y, z wird, usf., so daß endlich X eine lineare Funktion von der n^{ten} Unbekannten x allein ist. Die Werte, welche Ω zu einem Minimum machen, sind diejenigen, welche die n Gleichungen

$$X = 0, \dots Z = 0, Y = 0$$

befriedigen, und das letzte Glied Ω_0 in dieser Form stellt offenbar den Minimumwert von Ω dar.

Fragt man nun aber nach dem wahrscheinlichsten Wert der Unbekannten x allein, so hat man zunächst den Ausdruck der Wahrscheinlichkeit abzuleiten, daß der Wert dieser Unbekannten zwischen den Grenzen x und $x + dx$ enthalten ist. Diesen erhält man durch Integration des obigen Wertes

$$C h^m e^{-h^2 \Omega} dh dx dy dz \dots$$

in bezug auf alle zulässigen Werte der Unbekannten h, y, z, \dots . Bringt man die Summe Ω auf die oben erwähnte Form, so gibt die sukzessive Integration in bezug auf die $(n - 1)$ Unbekannten y, z, \dots ein Resultat

$$C' h^{m-n+1} e^{-h^2 (X^2 + \Omega_0)} dh dx,$$

worin C' unabhängig von h und x ist; integriert man endlich noch in bezug auf h , so erhält man für die gesuchte Wahrscheinlichkeit den Ausdruck

$$\frac{c dx}{(X^2 + \Omega_0)^{\frac{m-n+2}{2}}},$$

worin

$$\frac{1}{c} = \int_{-\infty}^{+\infty} \frac{dx}{(X^2 + \Omega_0)^{\frac{m-n+2}{2}}};$$

und hieraus folgt, daß derjenige Wert von x , für welchen $X = 0$ wird, unter allen der wahrscheinlichste ist. Dieser Wert stimmt daher wirklich mit dem durch die Methode der kleinsten Quadrate erhaltenen überein.