DER MITTELWERT

Von Dr. S. Škreb.

Die gesamte ältere Klimatologie und ein Teil der Meteorologie fusst auf Mittelwerten. Trotzdem ist über den Mittelwert selbst, über seine Begründung und Bedeutung recht wenig in der gesammten meteorologischen Literatur zu finden, leider ist unter diesem Wenigen auch Falsches. Die mathematische Literatur über den Mittelwert ist bedeutend reicher, doch auch davon ist nur ein kleiner Teil aus der neuesten Zeit in der Meteorologie und Statistik anwendbar. Es ist darum nicht unbegründet den Mittelwert, wie er allgemein gebraucht wird, von allen Seiten zu betrachten um daraus den Wert oder die Wichtigkeit, die ihm faktisch zukommt, zu erkennen.

Die Gewohnheit, dass man bei vielen gleichartigen aber verschiedenen »Dingen« oder »Tatbeständen«[1]) einen Durchschnitt sucht ist wol sehr alt. Die Quelle dieses Vorgehens liegt warscheinlich in der Beschränktheit unseres Gedächtnisses. Wir sind nicht im Stande uns jedes einzelne »Ding« klar zu merken und suchen darum einen »Vertreter« der uns gewissermassen die ganze Menge vorstellen soll. Man könnte dieses Vorgehen oder diesen Vereinfachungstrieb auch auf das Mach'sche Prinzip der Oekonomie des Denkens[2]) zurückführen. Als Beispiel eines (geschätzten) Mittelwertes denke man sich, dass man aus Norddeutschland nach Italien kommt. Man wird dann sagen können, dass im Durchschnitt die Menschen kleiner und die Temperaturen höher sind. Dieser geschätzte Durchschnitt oder Mittelwert entspricht aber in solchen Fällen immer dem häufigsten Wert der Beobachtung.

Prinzipiell sind aber Mittelwert und häufigster Wert zwei verschiedene Begriffe, die auf diese Weise, von Anfang an bis in die neueste Zeit, sehr oft verwechselt werden.

Beim Handel mit vielköpfiger Herde oder mit Stückgut kann es bei der Abschätzung des gesammten Wertes manchmal auch zur Bildung eines Mittelwertes kommen. Man nimmt den mittleren oder durchschnittlichen Wert eines Stückes und erhält durch die Multiplikation mit der Stückzahl den Gesamtwert. Dieser Mittelwert ist ganz anders definiert als der oben erwähnte. Wenn man noch annimmt dass weder der Käufer noch der Händler betrogen werden dürfen, so ist dieser Mittelwert als arithmetisches Mittel rechnerisch bestimmt, ohne Rücksicht auf die Häufigkeit des Vorkommens.

Man sieht aus diesen ungefähren Beispielen, dass der durchschnittliche oder mittlere Wert schon auf primitivster Stufe zwischen dem häufigsten Wert einer Menge und dem rechnerisch als arithmetisches Mittel zu bestimmenden Wert liegen kan. Am einfachsten ist es natürlich wenn beide Werte übereinstimmen (Gauss).

---

[1]) Czuber: Die Entwicklung der Wahrscheinlichkeitstheorie und ihrer Annwendungen. Jahresbericht der deutschen Mathematiker-Vereinigung Bd. VII. 1899.

[2]) Mach: Erkenntniss und Irrtum. Leipzig. 1920.

Es wurde darum, sobald sich irgendwo eine rechnerische Bestimmung eines Mittelwertes als notwendig erwies, das arithmetische Mittel als »der Mittelwert« (schon wegen der einfachsten Ausrechnung) genommen, aber auch stillschweigend angenommen, dass dieser Wert zugleich dem am häufigsten vorkommenden Wert dieser Menge entspricht.

In der Wissenschaft bekam der Mittelwert eine hervorragende Stellung durch die sogenannte »Ausgleichsrechnung nach der Methode der kleinsten Quadrate«[3]). Der grosse deutsche Mathematiker C. F. Gauss nahm das althergebrachte Verfahren des arithmetischen Mittels[4]) an um den wahrscheinlichsten Wert zu erhalten aus einer Reihe wiederholter präziser Messungen einer konstanten physikalischen Grösse (erste Voraussetzung). Wenn man eine konstante Grösse mehrmals mit der grössten Genauigkeit misst, so erhält man Zahlen, die niemals bis in die letzten Dezimalen übereinstimmen.

Die Zahlen widersprechen gewissermassen einander und man muss einen Ausgleich treffen um den wahrscheinlichsten Wert der gemessenen konstanten Grösse aus ihnen zu erhalten. Dass dieser Ausgleich am besten durch die Bildung des arithmetischen Mittels geschieht, ist eine vollständig willkürliche Annahme, die überhaupt nicht bewiesen werden kann aber aus dem vorher gesagten verständlich ist. Es wurde schon eine Menge Scheinbeweise hiezu geliefert,[5]) die es aber kaum plausibel machen können, bei weitem nicht beweisen. Die Differenz zwischen dem Mittelwert und Einzelwert wird hier Fehler genannt und Gauss nahm für die Ausgleichsrechnung weiter an (zweite Vorausetzung), dass die Häufigkeit dieser Fehler bei langen Messungsreihen einer ganz bestimmten mathematischen Funktion entspricht. Wen wir die Grösse des Fehlers mit x seine Häufigkeit mit y bezeichnen so ist nach Gauss $y = ae^{-h^2x^2}$ wobei a und h Konstanten sind. Die Häufigkeit des Fehlers hängt nur von seiner Grösse ab.

Diese Funktion als Kurve dargestellt hat die bekannte vollständig symmetrische Glockenform, mit dem Maximum (für $x = 0$, $y = a$) in der Mitte der Kurve.

Da aber die »Fehler« nur Differenzen gegen den Mittelwert sind, so ist hiedurch auch für die Häufigkeit der Messzahlen selbst dieselbe Funktion vorausgesetzt, wobei der Mittelwert zugleich der häufigste Wert der (genügend langen) Reihe notwendig sein muss, da die Funktion symmetrisch ist. Da weiter die Begriffe der relativen Häufigkeit und der Wahrscheinlichkeit identisch sind so ist der häufigste auch zugleich der warscheinliste Wert.

Wir können jetzt die zwei Voraussetzungen der Ausgleichsrechnung in umgekehrter Reihenfolge logischer verbinden. Da vorausgesetzt wird, dass bei wiederholten genauesten Messungen einer konstanten physikalischen

---

[3]) Czuber E. Enz. d. math. Wiss IDI.
Jordan dr. W. Hdbch der Vermessungskunde I. Bd. Ausgleichsrechnung Stuttgart, 1904.
Helmert: Ausgleichsrechnung Teubner, 1907.

[4]) »Die Regel des arithmetischen Mittels als praktisches Rechnungsverfahren, wurde, bevor Gauss sie zur Grundlage einer wichtigen Theorie erhoben, benutzt, ohne dass man nach einer wissenschaftlichen Begründung geforscht hätte, sie galt als unanfechtbare Eingebung des Verstandes.«
Czuber E.: Theorie der Beobachtungsfehler Leipzig Teubner 1891, Seite 44.

[5]) Czuber E. a. a. O.

Grösse die Häufigkeit einzelner Messzahlen der, oben erwähnten, symmetrischen Gausschen Funktion entspricht, so wird der häufigste also »wahrscheinlichste« Wert dieser Reihe am einfachsten ausgerechnet, wenn man das arithmetische Mittel dieser Messreihe nimmt. Es ist in diesem Falle das arithmetische Mittel nur eine Rechnungsart um den häufigsten also wahrscheinlichsten Wert der Reihe auf die einfachste Art zu erhalten. Es ist darum prinzipiell nicht korrekt den Mittelwert der Ausgleichsrechnung mit den anfangs erwähnten Beispielen wo eine Uebersicht gesucht wird zusammenzustellen.

Die weitere Voraussetzung: dass der häufigste Wert der Messreihe in der Ausgleichsrechnung (als arithmetisches Mittel berechnet) dem »wahren« Werte der gemessenen Grösse am besten entsprechen soll, ist wieder willkürlich, doch ist es hier weniger von Interesse diese Voraussetzung zu untersuchen, da es in der Meteorologie keine »wahren« Mittelwerte geben kann.

Da die »Ausgleichsrechnung« in der Geodäsie und Astronomie wahre Triumphe feierte, wurde das arithmetische Mittel einfach als der »wahre« Wert jeder Messreihe angesehen und die Symmetrie der Häufigkeit der einzelnen Zahlen (oder ihrer Abweichung vom Mittelwert) als notwendig postuliert. Es ist aber aus dem vorher gesagten klar, dass der Mittelwert (selbst in der Ausgleichsrechnung) mit Unrecht auf den Thron der Wahrheit erhoben wurde. Dem Sinne der Ausgleichsrechnung (gleiche Anzahl der positiven und negativen Abweichungen) nach müsste der häufigste Wert der Reihe »wahr« sein, da er faktisch dem wahren Werte am nächsten kommen soll. Dass man hier den häufigsten Wert als arithmetisches Mittel aller Messzahlen (nach der Voraussetzung) erhalten kann, ist eine angenehme Bequemlichkeit, aber keine Naturnotwendigkeit.

Wie der Mittelwert als »wahrer Wert« vorherrscht sieht man klar in der Arbeit von Lamont[6]) der in der Meteorologie verschiedenartige Mittelwerte unterscheiden will. Als »wahre« gelten ihm nur diejenigen, die der Ausgleichsrechnung entsprechen und danach der symmetrichen Häufigkeitskurve folgen. Selbst ein so scharfer und unbefangener Denker, wie Fechner, bedurfte erst eines gewissen moralischen Rucks, um sich zu sagen, dass schlechterdings kein stichhaltiger Grund anzuführen sei, weshalb z. B. die Kurve einer Rekrutentafel symmetrisch sein müsste[7]) und dass ferner eine solche Symmetrie, wenn sie wirklich existierte, geradezu eine im höchsten Grade merkwürdige Erscheinung bilde. Um Beobachtungsreihen, wo jedes einzelne Glied ein Ding für sich, in der Natur bestimmt und nicht nur den »fehlerhaften Mittelwert« darstellt, mathematisch bearbeiten zu können, begründete Fechner die Kollektivmasslehre. Trotz der Arbeiten von Fechner[9]), Czuber [10]), Bruns[8]) u. a. ist noch vielfach in der Meteorologie wie auch in der Statistik die für diese Wissenschaften sinnlose Ausgleichsrechnung siegreich. So erkennt z. B. Žižek (1923!)[11]) wie Lamont (1867) nur Mittelwerte einer symmetrischen Reihe an und sagt sogar in der Statistik: »Die Einzel-

[6]) Lamont: Ueber die Bedeutung arithmetischer Mittelwerte.
[7]) dasselbe gilt für meteorologische Beobachtungsreihen.
[8]) Bruns H.: Wahrscheinlichkeitsrechnung und Kollektivmasslehre Leipzig, Teubner 1906.
[9]) Fechner-Lipps: Kollektivmasslehre. Leipzig, Englemann 1897.
[10]) Czuber: Wahrscheinlichkeitsrechnung.
[11]) Žižek F.: Grundriss der Statistik. München 1923.

werte sind dem Mittelwert gegenüber nur mit zufälligen Fehler behaftete und daher bedeutungslose Abweichung.« Das heisst wol die Natur selbst als bedeutungslos und die errechneten willkürlichen Zahlen als das Wichtigste betrachten! Es mag hier kurz erwähnt werden, dass die Ausgleichsrechnung auch dann möglich wäre, wenn die Häufigkeistkurve der Messungen oder die der »zufälligen Fehler« nicht symmetrisch wäre. Damit wäre nur die rechnerische Unannemlichkeit verbunden, dass man den häufigsten also wahrscheinlichsten (wahren) Wert nicht einfach als arithmetisches Mittel aller Messzahlen finden könnte sondern, dass man zuvor den analytischen Ausdruck für die Häufigkeitskurve finden müsste aus welchem man dann das Maximum als Extremwert nach bekanntem Verfahren finden würde und hiedurch den wahrscheinlichsten (häufigsten) Wert. Es ist ja eben die Bequemlichkeit welche für die symmetrische Häufigkeitskurve der wiederholten Messungen einer konstanten Grösse spricht und keine »tiefliegenden« Gründe.

Die Gaussche symmetrische (und willkürliche) Häufigkeitskurve wird auch Kurve der zufälligen Fehler gennant. Hiedurch scheint es als ob der Begriff der Zufälligkeit überhaupt an die Gaussche Häufigkeitskurve gebunden wäre. So sagt Hann (in Hann-Süring: Lehrbuch der Meteorologie. IV Afl. S. 112), nachdem er die Häufigkeit der Abweichungen vom Mittelwert bei Jahresmitteln des Luftdrucks und der Temperatur (nach Angot) mit der Gausschen Kurve verglichen und übereinstimmend gefunden hat: »Man wird diese Proben für genügend halten um den Satz aufstellen su können, dass auf langjährige Temperatur und Luftdruckmittel die Regeln der Wahrscheinlichkeitsrechnung angewendet werden dürfen«. Da die Regeln der Wahrscheinlichkeitsrechnung nur auf »zufällige Dinge« angewendet werden dürfen, so hat es den Anschein als ob durch den Vergleich mit der Gausschen Kurve der »zufälligen Fehler« die Zufälligkeit der Luftdruck- und Temperaturmittel (oder ihrer Abweichungen) bewiesen wurde. Die Zufälligkeit kann aber durch keine mathematische Formel »bewiesen«, sondern nur durch einen logischen Schluss angenommen werden. Diese Entscheidung ob etwas als zufällig betrachtet werden kann oder nicht, gehört überhaupt nicht in die Mathematik. Ist dieser Schluss aber gefasst und die Zufälligkeit plausibel gemacht, so können die mathematischen Regeln der Wahrscheinlichkeitsrechnung angewendet werden ohne Rücksicht auf die Form der Häufigkeitskurve. Dass der (gar nicht eindeutig zu definierende) Begriff der Zufälligkeit mit der Symmetrie nichts zu schaffen hat, zeigt am schönsten die kolossal entwickelte Lebensversicherung, die auf unsymmetricher Sterbetafel fusst.

Es scheint aber dass an der zitierten Stelle überhaupt nicht Wahrscheinlichkeitsrechnung sondern die Ausgleichsrechnung herangezogen werden sollte, doch auch nicht »mit Recht« wie schon erwähnt wurde.

In der Ausgleichsrechnung müsste (bei direkter Messung) jedes Glied der Messreihe eigentlich dieselbe Zahl ergeben und darum wird angenommen, dass der errechnete häufigste Wert (also der Mittelwert) die zu messende (konstante) Grösse auf mehr Dezimalen ergeben kann als die Einzelmessungen. Diese für die Ausgleichsrechnung auf Grund der symmetrischen Häufigkeitskurve plausible aber unbegründete Annahme wurde auch in die Meteorologie übernommen, wodurch die Übung entstand, dass Mittelwerte

»genauer« angegeben werden sollen als Einzelwerte. Da aber in der Meteorologie dem Mittelwert allgemein überhaupt Nichts in der Natur entspricht, so kann auch die Angabe auf beliebig viel Dezimalen aus diesem Nichts keine »konstante physikalische Grösse« machen für welche eine genauere Angabe einen Sinn hätte.

Stellen wir uns eine meteorologische Beobachtungsreihe vor und bezeichnen sie durch $m_1, m_2, \ldots\ldots m_n$. Der Mittelwert Mn dieser Reiche, als arithmetisches Mittel, wird gewöhntlich so berechnet, dass die Summe aller Zahlen durch ihre Anzahl dividiert wird.

$$M_n = \frac{m_1 + m_2 + \text{-----} + m_n}{n} = \frac{1}{n} \Sigma m$$

Die Entstehung des Mn wird aber klarer dargestellt, wenn man zuvor von jeder Zahl den n-ten Teil nimmt und dann diese n Teile summiert:

$$\frac{m_1}{n} + \frac{m_2}{n} + \text{-----} = \frac{m_n}{n} = M_n$$

Wir sehen, dass Mn von jeder Zahl der Reihe einen Teil hat. Da alle dies Zahlen Messresultate einer bestimmten Grössenart sind so kann auch dem Mittelwert der aus solchen Teilen zusammengesetzt ist kein anderer Charakter zuerkannt werden, als der einer, wenn auch fiktiven, Messzahl oder eines Messresultates.

Der Hauptzweck, zu welchem Mittelwerte in der Meteorologie berechnet werden, ist die Möglichkeit der Vergleichung mit anderen gleichartigen Mittelwerten, da Reihen selbst wegen Unübersichtlichkeit nicht verglichen werden können. Der Mittelwert ist dannach nur ein Vertreter der Reihe. Als solcher müsste er natürlich nicht notwendig als arithmetisches Mittel berechnet werden. Jede beliebige andere Wahlart des Vertreters wäre ebenso »berechtigt« sobald sie allgemein angenommen sein würde, da diese Wahlart nur konventionell bestimmt werden kann. Für das arithmetische Mittel als »Wahlsystem« für den Vertreter der Reihe ist, nach Vorigem, ausschlaggebend die Kommodität und Einfachkeit der Rechnungsart und die aus der Ausgleichsrechnung stammende falsche Überzeugung, dass man so den »wahren« Wert erhält.

Der Mittelwert irgend einer meteorologischen Reihe ist dannach ein zu Vergleichszwecken gewählter Reihenvertreter, gewählt nach dem »Proportionalwahlrecht«, (weil beim arithmetischen Mittel jede Zahl der Reihe um so mehr zu der Grösse ihres Vertreters beiträgt, je grösser sie selbst ist).

Da dem Mittelwert der gleiche Charakter einer (fiktiven) Messzahl zukommt wie allen (faktischen) Messzahlen der Reihe so kann er auch keine grössere (physikalische) Genauigkeit haben, als jede Zahl der Reihe. Das ist ein weiterer Beveiss dazu, dass es absolut keinen Sinn hat den Mittelwert einer meteorologischen Reihe auf mehr Dezimalen anzugeben, als einzelne Glieder der Reihe angegeben sind. Rechnerisch ist es manchmal vorteilhaft, wegen leichterer und schärferer Kontrolle eine Dezimalstelle mehr zu berechnen. Einen physikalischen Sinn hat diese Stelle aber nicht und sollte bei Vergleichen nicht mitgeführt werden.

Wir haben bisher keine Beschränkungen für die Zahlen einer meteorologischen Messreihe angenommen. Wenn wir jetzt die aus der Beobachtung stammende Erfahrung benützen, dass keine meteorologische Zahlenreihe (direkter Beobachtungen) das ganze Gebiet reeller Zahlen von $-\infty$ bis $+\infty$ einnehmen kann, sondern ,dass die Grenzen der in einer Reihe vorkommenden Zahlen allgemein im Endlichen und zwar in ziemlich engem Gebiete liegen, so folgt daraus unbedingt, dass mit wachsender Beobachtungsanzahl jede Zahl der Reihe immer weniger zur Entstehung des Mittelwertes beiträgt. Der Mittelwert wird dann immer mehr unabhängig von weiteren Gliedern der Reihe. Daraus folgt aber n i c h t, dass ein solcher Mittelwert (wenn n sehr gross ist gegenüber jedem m_i) notwendig eine Naturkonstante darstellen muss. In der Annahme der Grenzen der Reihe liegt nicht der Beweis, dass der Mittelwert einer ebensolchen Reihe, die unmittelbar folgt, notwendig mit dem vorhergehenden übereinstimmen muss.

Um auch die Konstanz eines Mittelwertes aus sehr langer Beobachtungsreihe annehmen zu können, ist es, ausser der Annahme der Grenzen der Messzahlen im Endlichen, auch noch notwendig dass die relativen Häufigkeiten einzelner Messzahlen in den erwähnten Grenzen konstant sind.

Daraus folgt, dass man die (vergleichende Mittelwerts-) Klimatologie nur auf der Konstanz der Häufigkeitskurven logisch fundieren kann.

Die Bedeutung des Mittelwertes bei der Häufigkeitskurve soll später erörtert werden.

Dass der Mittelwert allgemein als Vertreter der Reihe angesehen wird, ersieht man aus der allgemein anerkannten Regel, dass nur Mittelwerte der Reihen g l e i c h e r G l i e d e r z a h l zu Vergleichszwecken gebraucht werden. Die Richtigkeit oder Notwendigkeit dieser Regel kann nicht bewiesen werden, sie gilt konvetionell und hat ihre Berechtigung nur bei kürzeren Reihen. Wenn es sich aber um Mittelwerte handelt, die auf Grund konstanter Häufigkeitskurven als Naturkonstanten betrachtet werden sollen, so werden noch andere Verhältnisse von Wichtigkeit, und zwar die sog. Streung der Werte, so dass für die Vergleichung solcher Mittelwerte andere Kriterien gelten, die hier nicht erörtert werden sollen.

Die Zahl, welche eine meteorologische Beobachtungreihe vertreten soll, um diese Reihe mit andern, auf gleiche Art entstandenen, übersichtlich vergleichen zu können, wird konventioneller Weise als arithmetisches Mittel berechnet, (das V e r g l e i c h s m i t t e l ist ein a r i t m e t i s c h e s M i t t e l). Dieser Mittelwert hat, seiner Entstehung nach, den Charakter einer Zahl der Reihe ohne irgend welche weitere Vorrechte, da ja nur an die Stelle einer g r o s s e n A n z a h l f a k t i s c h e r Messzahlen e i n e f i k t i v e Messzahl gesetzt wird. Ohne Anschluss an die Häufigkeitskurve für das betreffende Element kann über den Mittelwert als Naturkonstante nichts ausgesagt werden.

Wir nehmen an, dass uns eine sehr lange Reihe der Messzahlen eines meteorologischen Elementes vorliegt. Es ist schon erwähnt, dass die Grenzen, zwischen welchen diese Zahlen varieren, im Endlichen liegen. Wenn wir jetzt die Häufigkeit, mit welcher einzelne Zahlen in dieser Reihe vorkommen, auszählen und das Resultat graphisch so darstellen dass wir die Grösse der Zahlen als Abscissen, ihre Häufigkeit in dieser Reihe als Ordinaten betrachten und in ein beliebiges Koordinatensystem eintragen, so erhalten wir als graphische Darstellung eine diskrete Punktreihe. Um daraus eine Häufigkeitskurve (Verteilungskurve) zu erhalten, müssen wir zwischen diese eingetragenen Punkte noch unendlich viele Punkte interpolieren. Wir übergehen hiedurch von endlicher Anzahl der Beobachtungszahlen zur unendlichen. Dieser Prozess, der gewöhnlich als nebensächlich betrachtet wird, ist niemals ohne kleine Willkürlichkeiten zu erledigen und stellt den wesentlichen Übergang von Einzelbeobachtungen zum »Klima« (unendliche Beobachtungszahl) dar. Derselbe Prozess vollzieht sich bei der Ausrechnung der formelmässigen analytischen Darstellung dieser Punktfolge. Dabei ist noch die Voraussetzung stillschweigend enthalten, dass die Häufigkeitskurve immer stetig ist und dass sie monoton vom Maximum abfällt ohne Unterbrechungen, Zacken und Sprünge. Gewöhnlich wird wegen leichterer Darstellung noch angenommen, dass nur ein Maximum vorhanden sein soll, und dass es zwischen den Grenzen (aber nicht notwendig symmetrisch) liegt.

Es ist immer möglich die Messzahlenreihe so auszuwählen dass nur ein Maximum vorhanden ist, doch ist es nicht immer möglich zu erreichen, dass das Maximum nicht auf die Grenzwerte fällt. So ergibt z. B. die Häufigkeitskurve der Stundensummen des Niederschlags (nach dem Ombrographen) eine Häufigkeitskurve deren Maximum an der unteren Grenze liegt, weil die kleinen Summen bedeutend häufiger sind als grössere.

Die Methode der analytischen Darstellung der Häufigkeitskurve aus der Messzahlenreihe gab Bruns[12]) an in Anschluss an die erwähnte Gaussche symmetrische Häufigkeitskurve.

Bei dem Verbinden der erwähnten einzelnen diskreten Punkte zu einer Kurve wissen wir aus Erfahrung, dass diese Verbindungslinie (die Häufigkeitskurve) umso glatter verläuft je mehr Daten zur Verfügung stehen. Es muss um eine glatte Kurve zu erhalten eine »gleichmässige Erschöpfung gleichmöglicher Fälle« vorhanden sein.[13])

Da diese Kurven analytisch nur als exponentielle Funktionen dargestellt werden können, so kann keine endliche Punktanzahl angegeben werden durch welche eine solche Kurve sicher bestimmt wäre. Um aber eine »gleichmässige Erschöpfung gleichmöglicher Fälle« zu erreichen, ist jedenfalls eine Zahlenreihe notwendig von sehr grosser Gliederzahl. Um das zu ersehen genügt es an ein einfaches Beispiel aus der Wahrscheinlichkeitsrechnung zu erinnern. Wenn auf den Flächen eines regelmässigen Würfels je eine der Zahlen von 1 bis 6 steht, so ist beim Würfeln zu erwarten, dass jede dieser Zahlen gleich häufig erscheint, weil sie alle gleichmöglich sind. Eine faktische »gleichmässige Erschöpfung« dieser »gleichmöglichen Fälle« wird aber manchmal kaum bei einigen hundert Würfen zu erreichen sein. Dieses einfache Beispiel zeigt uns klar, dass aus kurzen Reihen keine korrekten Häufigkeitskurven zu konstruiren sind. Im Falle des Würfels wäre die Häufigkeitskurve eine Gerade parallel zur Abscissenachse.

Rechnerisch ist es natürlich immer möglich aus einer gegebenen Zahlenfolge etwas auszurechnen, ob das Resultat aber einen physikalischen Sinn hat, bleibt sehr fraglich.

---

[12]) Bruns: Wahrscheinlichkeitsrechnung und Kollektivmasslehre. Leipzig, Teubner 1906.

[13]) Diese Verhältnisse werden auch als »Gesetz der grossen Zahlen« bezeichnet.

Die Untersuchung solcher Häufigkeitskurven verschiedener Elemente ist die Aufgabe der Kollektivmasslehre. Diese Untersuchung kann in zwei Richtungen gehen. R. v. Mises stellt als Grundaufgabe der Kollektivmasslehre[14]): »Man soll das in der Verteilungskurve zum Ausdruck gebrachte Ergebniss wiederholter Beobachtungen durch eine Reihe von Zahlen derart charakterisieren, dass den einzelnen Zahlen eine möglichst anschauliche Bedeutung zukommt.« Bruns[15]) sagt: »das zu erstrebende Ziel ist in der Kollektivmasslehre wie in anderen messenden und zählenden Wissenschaften die Verdichtung jeder Beobachtungsreihe in eine Formel, die alles das ausspricht was die Beobachtung an erkennbarer Gesetzmässigkeit in sich schliesst«.

Durch diese zwei Grundaufgaben der Kollektivmmasslehre ist auch die weitere Entwiklung der Klimatologie gegeben. (Nach Bruns wäre ja die gesamte Meteorologie eine Kollektivmasslehre der Witterungserscheinungen was aber allgemein selbst für die Klimatologie nicht gilt).

Bisher genügte es in der Klimatologie nur die eine charakteristische Zahl der Reihe (nach Mises) zu kennen: den Mittelwert; sobald eine genügende Anzahl der Beobachtungen vorhanden sein wird, ergibt sich die Möglichkeit weitere Zahlen wie den häufigsten Wert, die Streuung usw. zu erkennen und schliesslich die analytische Form der Haufigkeitskurve zu untersuchen um sie mit anderen zu vergleichen.

Wir sind in der Klimatologie eben im Übergangsstadium. Die Mittelwertsklimatologie ist ziemlich für den grössten Teil der Erde bekannt und für weitere charakteristische Zahlen und die Häufigkeitsklimatologie sind unsere Beobachtungsreihen allgemein zu kurz, sie genügen nicht der Hauptforderung: gleichmässige Erschöpfung gleichmöglicher Fälle. Es melden sich aber in der Literatur schon Stimmen von mehreren Seiten, welche auf die Notwendigkeit der Häufigkeitsuntersuchungen in der Klimatologie hinweisen.[16])

Eine vollständige Verkennung dieser Sachlage findet man in Hann-Süring: Lehrbuch der Meteorologie S. 119, wo es heisst: »Statt nach der mittleren Temperatur kann man auch nach jener Temperatur fragen welche innerhalb eines gewissen Zeitraumes am häufigsten auftritt«.

Der häufigste Wert kann niemals »statt« des Mittelwertes errechnet werden, sondern in weiterer und genauerer Untersuchung der klimatischen Verhältnisse. Eine Konkurenz dieser Werte ist sinnlos.

Wie tief aber der Mittelwert als Naturkonstante in der älteren Meteorologie wurzelt ersieht man aus dem Schlussatz (a. a. O. S. 120) der auch die Unbrauchbarkeit der Häufigkeitskurven zeigen soll: »Daraus ergibt sich auch, dass der Scheitelwert nur aus sehr langen Beobachtungsreihen mit einiger Sicherheit berechnet werden kann, und dann gibt es immer noch sekundäre Scheitelwerte die an einem Nachbarorte den Hauptscheitel bilden können, wodurch dann der scheinbare Temperaturunterschied ganz benachbarten Orte leicht ebensoviele Grade erreichen kann als er in Wirklichkeit Zehntelgrade beträgt.«

[14]) R. v. Mises: Jahresbericht der deütsch. Math. Ver.

[15]) Bruns: a. a. O. S. 243.

[16]) Als Pionier dieser Richtung ist in neuster Zeit L. W. Pollak, Prag besonders hervorzuheben der auch die erste grössere Arbeit in dieser Richtung lieferte (Prager Geophysikalische Studien, Prag 1927).

Aus dem Kontext ersieht man, dass »in Wirklichkeit« die Differenz der Mittelwerte bedeutet. Es wären danach Mittelwerte Wirklichkeit und die häufigsten Werte fiktiv. Faktisch ist es umgekehrt.

Wenn wir den Mittelwert als Naturkonstante mit Hilfe der Häufigkeitskurve definieren wollen (unter Voraussetzung der Konstanz der Häufigkeitskurve), so können wir das anschaulich folgendermassen erreichen.

Wenn die Häufigkeitskurve als analytische Funktion f(x) gegeben ist, so ist der Mittelwert definiert durch die Gleichung $M = \frac{S x f(x)\, dx}{S x f(x)\, dx}$. Denkt man sich jetzt die Fläche zwischen der Häufigkeitskurve und der Abscissenachse gleichmässig mit Masse belegt, so ist M die Abscisse des Schwerpunktes der Fläche.[17])

Diese (geometrische) Bedeutung des Mittelwerts ist wol sehr anschaulich, hat aber für die Meteorologie wenig Wichtigkeit. Man sieht danach dass der Mittelwert als Naturkonstante keine bedeutende Rolle spielen kann. Da ist wol der häufigste Wert oder das Maximum der Häufigkeitskurve als Naturkonstante an Wichtigkeit und Anschaulichkeit dem Mittelwert überlegen. Aus dieser Gegenüberstellung ersieht man aber, dass die Sicherheit des Mittelwertes die gleiche ist wie die des häufigsten Wertes (Scheitelwertes), da ja beide direkt von der Häufigkeitskurve abhängen. Es ist nur eine rechnerische Täuschung (unter Mitwirkung der Ausgleichsrechnung) wenn wir, selbst aus langen Messreihen, dem Mittelwert ohne weiteres den Charakter einer Naturkonstante zuschreiben möchten (»normale Mittel«).

Allerdings kann der Mittelwert aus jeder beliebigen Zahlenreihe, blindlings und gedankenlos, auf beliebig viele Dezimalen ausgerechnet werden, ohne das hiebei die eventuelle Inkompatibilität der Zahlen zum Vorschein kommt. Solche Mittel sind das beste Mittel um jedes beliebige Resultat zu erhalten.

Die Häufigkeitsuntersuchungen stellen viel grössere Anforderungen an das logische Denken und die Kritik des Arbeiters. Schon die Auswahl des Kollektivgegenstandes ist nicht immer leicht, wenn man eine klare Häufigkeitskurve, die auch etwas aussagt, erhalten will. Ausserdem ist auch die zu leistende mechanische Rechenarbeit bedeutend grösser. Es ist ein Verdienst von L. W. Pollak, dass er maschinelle Hilfsmittel zu dieser Arbeit gefunden hat.[18])

---

[17]) R. v. Mies: a. a. O.

[18]) L. W. Pollak: Verwendungs statistischer Maschinen in der Klimatologie: Meteor. Ztschft XLIV. 296. 1927, Prager geophysikalische Studien. Seite 22, Die Naturwissenschaften 18. 343. 1930.