Mathematik für Wirtschaftswissenschaftler

(K. Steffen, Heinrich-Heine-Universität Düsseldorf, WS 2006/07)

1.5 Mittelwerte

Eine erste Information über einen Satz erhobener Daten — und oft auch die einzige darin enthaltene Information, die ausgenutzt wird — liefert die Bildung eines Mittelwerts zu diesen Daten. Auch und gerade bei der Beurteilung ökonomischer Daten spielen Mittelwerte eine wichtige Rolle. Nun gibt es verschiedene Mittelwertbildungen, die auch verschiedene Werte liefern. Daher muss man, um die Aussagekraft von ausgewiesenen Mittelwerten beurteilen zu können, etwas über verschiedene Mittelungsverfahren und über den Größenvergleich der Mittelwerte wissen, welche diese Verfahren auf derselben Datengrundlage berechnen. Darum geht es in diesem Abschnitt. Zunächst geben wir eine sehr allgemeine Definition von Mittelungsverfahren für einen endlichen Datensatz, d.h. für endlich viele gegebene reelle Zahlen.

DEFINITION: Unter einem **Mittelungsverfahren** (einer **Mittelwertbildung**) für n reelle Zahlen aus einem Intervall I in $\mathbb{R}$ verstehen wir eine Vorschrift, die je n Zahlen $x_1, x_2, \ldots, x_n$ aus I eine reelle Zahl $M(x_1, x_2, \ldots, x_n)$ zuordnet und die folgenden Bedingungen genügt:

(i) $\min(x_1, x_2, \ldots, x_n) \leq M(x_1, x_2, \ldots, x_n) \leq \max(x_1, x_2, \ldots, x_n)$;

(ii) (*Monotonie*) Vergrößerung einer Zahl x_i in I bei Festhalten der Werte aller anderen Zahlen x_j, $j \neq i$, verkleinert den Wert von $M(x_1, x_2, \ldots, x_n)$ nicht;

(iii) (*Symmetrie*) Vertauschung der Zahlen x_j untereinander ändert den Wert von $M(x_1, x_2, \ldots, x_n)$ nicht.

Die Zahl $M(x_1, x_2, \ldots, x_n)$ heißt dann der **Mittelwert** von $x_1, x_2, \ldots, x_n$ bezüglich des Mittelungsverfahrens M. ∎

Das ist leicht zu verstehen: Die erste Bedingung besagt einfach, dass $M(x_1, x_2, \ldots, x_n)$ immer zwischen der größten und der kleinsten der zu mittelnden Zahlen liegen muss. Insbesondere liegt $M(x_1, x_2, \ldots, x_n)$ dann auch in dem Intervall I, dem die zu mittelnden Zahlen entnommen wurden. Als Intervalle in den Anwendungen kommen hier übrigens hauptsächlich $I = \mathbb{R}$ oder $I = \mathbb{R}_{>0}$ in Frage. Die zweite Bedingung bedeutet, das der Wert $M(x_1, x_2, \ldots, x_n)$ größer wird, oder jedenfalls nicht kleiner, wenn man eine der zu mittelnden Zahlen vergrößert. (Bei gewissen in der Praxis durchaus verwendeten Mittelungsverfahren braucht der Wert nicht echt größer zu werden.) Die dritte Bedingung schließlich besagt, dass der Wert $M(x_1, x_2, \ldots, x_n)$ von der Reihenfolge der Zahlen $x_1, x_2, \ldots, x_n$ nicht abhängen soll. (Sind einige der Zahlen untereinander gleich, so kommt es aber sehr wohl darauf an, wie oft jede vorkommt; wenn sich diese "Vielfachheiten" ändern, so wird sich im Allgemeinen auch der Wert $M(x_1, x_2, \ldots, x_n)$ ändern.) Diese drei Bedingungen sind unmittelbar einleuchtende Forderungen, die man an jedes vernünftige Mittelungsverfahren stellen wird.

Der Punkt ist, dass wir, um keine praktisch relevanten Mittelungsverfahren mit der Definition auszuschließen, nicht noch *mehr* Bedingungen gefordert haben. Das hat als Konsequenz, dass es im Sinne der Definition sehr viele und sehr unterschiedliche Mittelwertbildungen gibt, auch solche, die einem eher abwegig erscheinen. Z.B. ist die Bildung des Maximums $\max(x_1, x_2, \ldots, x_n)$ zu n gegebenen Zahlen $x_i \in \mathbb{R}$ ein zulässiges Mittelungsverfahren im Sinne der Definition, ebenso auch die Bildung des Minimums $\min(x_1, x_2, \ldots, x_n)$. Die Maximumbildung oder Minimumbildung wird man in der Praxis wohl kaum als Mittelungsverfahren benutzen, es gibt aber, wie wir sehen werden, auch durchaus eindrucksvolle "wissenschaftliche" Mittelwertformeln, welche Mittelwerte liefern, die immer ziemlich nahe beim Maximum der zu mittelnden Zahlen liegen (oder immer nahe beim Minimum).

Außerdem kann man das Ergebnis einer Mittelbildung erheblich dadurch verändern, dass man (mit mehr oder weniger guten Begründungen) Gewichte einführt oder die Daten manipuliert. Gewichte bewirken, dass die einzelnen zu mittelnden Zahlen das Ergebnis unterschiedlich stark beeinflussen. Die Manipulation von Daten erfolgt oft so, dass einzelne aus dem Rahmen fallende Werte zu "Ausreißern" deklariert und dann bei der Mittelbildung schlicht weggelassen werden. Aber auch ohne solche Maßnahmen im Auge zu haben, kann man sagen (wir werden das noch präzisieren): Für jedes gewünschte Ergebnis zwischen Minimum und Maximum der zu mittelnden Zahlen gibt es auch eine "Mittelwertformel", welche dieses Ergebnis liefert.

- *Die Angabe eines Mittelwertes ist daher ohne Informationswert, wenn das dabei verwendete Mittelungsverfahren nicht beschrieben wird (einschließlich vorgenommener Gewichtungen und Gewinnung der Daten). Die Wahl eines bestimmten Mittelungsverfahrens muss durch anwendungsbezogene Überlegungen gerechtfertigt werden.*

Bevor wir zu Beispielen von Mittelungsverfahren kommen, nennen wir noch eine Eigenschaft, die viele wichtige Mittelungsverfahren auf dem Grundintervall $\mathbb{R}$ oder $\mathbb{R}_{>0}$ besitzen. Diese Eigenschaft besagt, dass sich der Mittelwert $M(x_1, x_2, \ldots, x_n)$ zum Beispiel verdoppelt, wenn wir jeden der zu mittelnden Zahlenwerte x_j durch den doppelten Wert $2x_j$ ersetzen; entsprechend auch bei Multiplikation mit anderen Faktoren. Man spricht, wenn eine Mittelwertbildung diese Eigenschaft hat, von der **Homogenität** der Mittelwertbildung, d.h. für alle $s > 0$ und alle $x_1, x_2, \ldots, x_n$ gilt

$$M(sx_1, sx_2, \ldots, sx_n) = s \cdot M(x_1, x_2, \ldots, x_n)\,.$$

Da aber nicht alle in der Praxis relevanten Mittelungsverfahren homogen sind, haben wir diese Eigenschaft nicht zum Bestandteil der allgemeinen Definition von Mittelwertbildungen gemacht.

BEISPIELE und DISKUSSION von Mittelungsverfahren:

Für alle nachfolgend angegebenen Mittelungsverfahren sind die Bedingungen (i), (ii) und (iii) aus der obigen Definition mit Hilfe einer Monotoniediskussion für die entsprechenden Terme leicht nachzuprüfen; darauf gehen wir nicht mehr ein. Die Mittelwertverfahren sind auch alle homogen.

1) Am häufigsten verwendet — aber nicht immer zu Recht — wird das sogenannte **arithmetische Mittel**:

$$AM(x_1, x_2, \ldots, x_n) \quad := \quad \frac{1}{n}\sum_{i=1}^{n} x_j \quad = \quad \frac{x_1 + x_2 + \ldots + x_n}{n}\,.$$

Hier addiert man einfach die zu mittelnden Zahlen x_j und dividiert durch die Anzahl der Summanden. Der arithmetische Mittelwert wird relativ stark durch sog. *Ausreißer* beeinflußt, d.h. durch Werte x_i, die erheblich größer oder kleiner sind als die meisten anderen x_j (und die vielleicht auf fehlerhafter Datenerhebung beruhen). Daher ist arithmetische Mittelbildung oft nicht sachgerecht und sollte nicht blindlings angewendet werden, nur weil dieses Mittel leicht zu berechnen ist. Zum Beispiel ist es nicht sinnvoll, durchschnittliche Studienzeiten durch arithmetische Mittelbildung zu berechnen; denn wenige Fälle von extrem langer Studiendauer beeinflussen den Wert des arithmetischen Mittels erheblich, die Ursachen liegen aber meist im außeruniversitären Bereich, während man mit der mittleren Studiendauer eher einen Parameter im Auge hat, der eine Information über Studienbedingungen und Studienverhalten geben soll. Beispiele für arithmetische Mittelbildung in der Ökonomie sind zahllos: Durchschnittliche Umsätze, Kosten, Gewinne, Preise,... in mehreren Perioden oder für mehrere Produkte bzw. für mehrere Produktionsstätten eines Unternehmens usw.

2) Häufig sind die zu mittelnden Zahlen nicht "gleichberechtigt", sondern mit unterschiedlichem Gewicht zu berücksichtigen. Unter einem Satz von n **Gewichten** verstehen wir immer die Vorgabe von n positiven Zahlen $s_1, s_2, \ldots, s_n$ mit Summe $s_1+s_2+\ldots+s_n = 1$. (Insbesondere sind dann alle Gewichte s_j Zahlen zwischen 0 und 1, wenn $n \geq 2$ ist. Man könnte auch den Wert Null für Gewichte zulassen; das läuft darauf hinaus, dass die Zahlen mit Gewicht Null bei der Mittelbildung einfach weggelassen werden.) Gewichte werden oft als Prozentsätze angegeben, $s_j = p_j\% = \frac{1}{100}p_j$; die Zahlen $p_1, \ldots, p_n$ müssen sich dann natürlich zu 100 addieren. Das mit den Gewichten $s_1, \ldots, s_n$ aus den reellen Zahlen $x_1, x_2, \ldots, x_n$ gebildete **gewichtete arithmetische Mittel** ist

$$AM(x_1, x_2 \ldots, x_n; s_1, s_2, \ldots, s_n) \quad := \quad \sum_{j=1}^{n} s_j x_j \quad = \quad s_1x_1 + s_2x_2 + \ldots + s_nx_n .$$

Bei dieser Mittelbildung multipliziert man also jede der zu mittelnden Zahlen mit ihrem zugehörigen Gewicht und addiert dann die entstandenen Produkte auf. Wählt man alle Gewichte gleich, also $s_1 = s_2 = \ldots = s_n = \frac{1}{n}$, so ergibt sich das gewöhnliche ("ungewichtete") arithmetische Mittel der x_j.

Mitunter sind die Gewichte s_j nicht direkt so vorgegeben, dass sie sich zu 1 summieren, sondern man kennt nur gewisse positive "Referenzgrößen" r_j und möchte die gegebenen Zahlen mit Gewichten s_j mitteln, die in denselben Verhältnissen $s_1 : s_2 : \ldots : s_n$ stehen wie die Referenzgrößen $r_1 : r_2 : \ldots : r_n$ (d.h. die Quotienten $s_i : s_j$ und $r_i : r_j$ sind für alle $i \neq j$ gleich). Dies bedeutet, dass man die Gewichte $s_j := r_j/r$ zu wählen hat, wobei $r := r_1+r_2+\ldots+r_n$ die Summe der Referenzgrößen ist (denn genau bei dieser Wahl der s_j stehen die s_j in denselben Verhältnissen zueinander wie die r_j und haben die Summe 1). Für je n gegebene reelle Zahlen $x_1, x_2, \ldots, x_n$ und positive Zahlen $r_1, r_2, \ldots, r_n$ ist also das **arithmetische Mittel im Verhältnis der Referenzgrößen** $r_j > 0$ gleich

$$\frac{\sum_{j=1}^{n} r_j x_j}{\sum_{j=1}^{n} r_j} \quad = \quad \frac{r_1x_1 + r_2x_2 + \ldots + r_nx_n}{r_1 + r_2 + \ldots + r_n} .$$

Das arithmetische Mittel der Zahlen x_j im Verhältnis $1 : 1 : \ldots : 1$ ist also ihr gewöhnliches ("ungewichtetes") arithmetisches Mittel.

Soll das arithmetische Mittel zu einer endlichen Zahlenfolge gebildet werden, in denen Zahlen mehrfach auftreten können (Z.B. Klausurergebnisse), so kann man stattdessen auch die verschiedenen auftretenden Zahlen arithmetisch im Verhältnis ihrer Fallzahlen mitteln, d.h. im Verhältnis der Häufigkeiten ihres Auftretens in der Folge. Das Ergebnis ist dasselbe, weil die Summe der mit ihren Fallzahlen multiplizierten verschiedenen Zahlen gleich der Summe aller Zahlen der ursprünglichen Zahlenfolge mit "Wiederholungen" ist und die Summe aller Fallzahlen gleich der Anzahl der Glieder der ursprünglichen Folge.

In der Wahrscheinlichkeitstheorie treten gewichtete arithmetische Mittel $s_1x_1+\ldots+s_nx_n$ als *Erwartungswert einer Zufallsgröße* auf, welche die n verschiedenen Werte $x_1,\ldots,x_n$ annehmen kann und keine anderen. Das Gewicht $s_j \geq 0$ ist dabei die Wahrscheinlichkeit dafür, dass der Wert x_j auftritt, und die Summe dieser Wahrscheinlichkeiten ist 1, weil die Zufallsgröße mit Sicherheit einen der Werte x_j annimmt. Der Erwartungswert ist im Allgemeinen kein Wert, den die Zufallsgröße annehmen kann; insofern ist die Benennung etwas irreführend. (Beim Würfeln werden z.B. die Werte $1,\ldots,6$ der Augenzahl jeweils mit Wahrscheinlichkeit $\frac{1}{6}$ angenommen, der Erwartungswert, also hier das arithmetische Mittel der möglichen Augenzahlen, ist aber $\frac{7}{2}$ und somit keine mögliche Augenzahl.) Vielmehr gibt der Erwartungswert einen in diesem Kontext sinnvollen Mittelwert der möglichen Werte an, welche die Zufallsgröße annehmen kann, nämlich das gewichtete arithmetische Mittel mit den entsprechenden Wahrscheinlichkeiten als Gewichten.

3) Ein Beispiel aus der Ökonomie für die Bildung eines gewichteten arithmetischen Mittels ist der **Durchschnittspreis** für eine auf n Märkten abgesetzte Ware: Sind $p_1,\ldots,p_n$ die auf diesen Märkten jeweils erzielten Preise (pro abgesetzter Einheit), so darf man natürlich nicht die Zahlen p_j direkt arithmetisch mitteln, sondern man muss sie im Verhältnis der jeweils abgesetzten Mengen $q_1,\ldots,q_n$ (Einheiten) mitteln, um einen ökonomisch sinnvollen Mittelwert zu erhalten. (Ein hoher erzielter Preis auf einem Markt, auf dem fast nichts abgesetzt wurde, darf den Mittelwert nicht so stark beeinflussen, wie ein niedriger Preis auf einem andern Markt mit großem Umsatz.) Analog werden die **Durchschnittskosten** eines Produktionsfaktors (z.B. eines Rohstoffs), dessen Preis von Monat zu Monat schwankt, als arithmetisches Mittel der in n Monaten zu zahlenden Preise $p_1,\ldots,p_n$ im Verhältnis der jeweils eingekauften Mengen $q_1,\ldots,q_n$ definiert; etwaige Restbestände des eingekauften Produktionsfaktors werden am Jahresende mit diesen Durchschnittskosten bewertet und bilanziert. In beiden Situationen wird hier also das gewichtete arithmetische Mittel $\sum_{j=1}^n q_jp_j / \sum_{j=1}^n q_j$ gebildet.

Indexzahlen sind Verhältnisse von zwei (gewichteten oder ungewichteten) arithmetischen Mitteln ökonomischer Variablen und werden benutzt, um die zeitliche Entwicklung wichtiger ökonomischer Parameter zu beschreiben (Preisindex, Umsatzindex, Aktienindex, ...). Dabei erhebt man die Werte $x_1,\ldots,x_n$ von gewissen positiven ökonomischen Variablen in einer aktuellen "Berichtsperiode" und stellt diese den "Basiswerten" $\mathring{x}_1,\ldots,\mathring{x}_n$ dieser Variablen aus einer früheren "Basisperiode" gegenüber, indem man den Quotienten der gewichteten arithmetischen Mittel der beiden Sätze von Variablenwerten bildet. Dabei muss man natürlich für beide Mittelungen dieselben Gewichte nehmen, sonst ist der Vergleich der Mittelwerte sinnlos. Eine Indexzahl sieht also, wenn die Gewichte im gleichen Verhältnis wie bei den positiven Referenzgrößen $r_1,\ldots,r_n$ gewählt sind, folgendermaßen aus

$$I \quad = \quad \frac{\sum_{j=1}^n r_jx_j \Big/ \sum_{j=1}^n r_j}{\sum_{j=1}^n r_j\mathring{x}_j \Big/ \sum_{j=1}^n r_j} \quad = \quad \frac{\sum_{j=1}^n r_jx_j}{\sum_{j=1}^n r_j\mathring{x}_j}\,.$$

Statt der Zahl I wird oft der Prozentsatz $100 \cdot I\%$ angegeben.

Die Indexzahl I fällt größer als 1 aus, wenn der Mittelwert vom Bezugszeitpunkt bis zum aktuellen Zeitpunkt zugenommen hat, bzw. kleiner als 1, wenn er abgenommen hat. Genauer ist der aktuelle Mittelwert gleich $100 \cdot I\%$ vom Mittelwert aus der Basisperiode, und $I-1$ gibt die relative Zunahme bzw. Abnahme des Mittelwertes von der Basisperiode bis zur Berichtsperiode an (d.h. die Zunahme/Abnahme beträgt $100 \cdot (I-1)\%$ des Mittelwerts aus der Basisperiode).

Betrachten wir z.B. einen "Warenkorb" von n Gütern, mit Marktpreisen p_j (Geldeinheiten pro Mengeneinheit), umgesetzten Mengen q_j (Mengeneinheiten) und Umsätzen $u_j = p_j \cdot q_j$ (Geldeinheiten) im aktuellen Jahr. Für die Bildung eines mittleren Umsatzes ist ungewichtete arithmetische Mittelbildung sinnvoll, d.h. man mittelt die Umsätze der einzelnen Güter im Verhältnis $1:1:\ldots:1$. Dementsprechend ist der

$$\textbf{Umsatzindex:} \qquad \frac{\sum_{j=1}^{n} u_j}{\sum_{j=1}^{n} \mathring{u}_j} = \frac{\sum_{j=1}^{n} p_j q_j}{\sum_{j=1}^{n} \mathring{p}_j \mathring{q}_j}.$$

Bei der Bildung eines mittleren Preises ist es aber sicher nicht sinnvoll, einfach das arithmetische Mittel der Preise der einzelnen Güter zu nehmen; denn eine Verteuerung bei einem wenig verkauften Produkt sollte nicht so ins Gewicht fallen wie eine Verteuerung bei einem viel umgesetzten Produkt. Also wird man die Preise der n Güter im Verhältnis der umgesetzten Mengen mitteln. Damit entfällt auch das Problem bei der Definition der Mengenangaben q_j, wenn die umgesetzte Mengen der Güter des Warenkorbs in verschiedenen Einheiten gemessen werden (Stückzahlen, Gewichtseinheiten,...); denn die Produkte $q_j p_j = u_j$ sind unabhängig von der Wahl der Mengeneinheiten. Man hat nun zwei verschiedene Möglichkeiten der Gewichtung, die sich anbieten: Man kann sich auf die umgesetzten Mengen in der aktuellen Berichtsperiode beziehen (*Index nach Paasche*) oder auf die umgesetzten Mengen in der früheren Basisperiode (*Index nach Laspeyres*). Demnach ist der

$$\textbf{Preisindex:} \qquad \frac{\sum_{j=1}^{n} q_j p_j}{\sum_{j=1}^{n} q_j \mathring{p}_j} \quad (\textit{Paasche}), \qquad \text{bzw.} \qquad \frac{\sum_{j=1}^{n} \mathring{q}_j p_j}{\sum_{j=1}^{n} \mathring{q}_j \mathring{p}_j} \quad (\textit{Laspeyres}).$$

Diese Preisindizes haben wir in Abschnitt 1.1 schon im Zusammenhang mit den Rechenregeln für Summen diskutiert. Beide bestimmen die mittleren Preise so, als hätten sich die umgesetzten Mengen seit der Basisperiode nicht geändert, wobei im ersten Index die aktuellen umgesetzten Mengen, im zweiten die umgesetzten Mengen in der Basisperiode angesetzt werden. Analog hat man auch zwei Möglichkeiten, einen Mengenindex zu definieren, der die Änderung der mittleren umgesetzten Mengen von der Basisperiode bis zur Berichtsperiode beschreibt. Hier ist die einfache arithmetische Mittelung der umgesetzten Mengen ebenfalls nicht sinnvoll, schon wegen der unterschiedlichen Einheiten, in denen die umgesetzten Mengen gemessen werden. (Es macht schließlich einen beträchtlichen zahlenmäßigen Unterschied, ob man z.B. verkaufte Streichhölzer in Stückzahlen oder in Kilogramm angibt.) Das Problem entfällt, wenn man die umgesetzten Mengen im Verhältnis der Preise mittelt, und so erhält man den

$$\textbf{Mengenindex:} \qquad \frac{\sum_{j=1}^{n} p_j q_j}{\sum_{j=1}^{n} p_j \mathring{q}_j} \quad (\textit{Paasche}), \qquad \text{bzw.} \qquad \frac{\sum_{j=1}^{n} \mathring{p}_j q_j}{\sum_{j=1}^{n} \mathring{p}_j \mathring{q}_j} \quad (\textit{Laspeyres})\,.$$

Bei beiden Mengenindizes werden die mittleren Umsätze so ermittelt, als hätten sich die Preise seit der Basisperiode nicht verändert, wobei im ersten Fall die aktuellen, im zweiten Fall die Basispreise angenommen werden. Insofern sind beide Indizes ein Maß für die Entwicklung der durchschnittlich abgesetzten Mengen von der Basis- bis zur Berichtsperiode.

Die Vor- und Nachteile der Index-Definitionen nach Paasche und Laspeyres liegen auf der Hand: Bei der Definition nach Laspeyres kann man die Gewichte aus den Daten der Basisperiode entnehmen, das spart Erhebungskosten und ermöglicht, die Gewichte für mehrere aufeinander folgende Berichtsperioden beizubehalten. So verfährt z.B. das statistische Bundesamt bei der Ermittelung eines Lebenshaltungskostenindex. Allerdings entspricht die Gewichtung nach einigen Perioden dann möglicherweise kaum noch der aktuellen Situation, während die Indizes nach Paasche sich eben immer an der aktuellen Situation orientieren. Deshalb müssen die angenommenen Umsätze für die Güter des "Lebensmittelkorbs" von Zeit zu Zeit aktualisiert werden (wie auch die Zusammensetzung des Korbs überhaupt den eingetretenen Entwicklungen angepasst werden muss).

4) Das geometrische Mittel von n positiven (!) Zahlen $x_1, \ldots, x_n$ ist das multiplikative Analogon des arithmetischen Mittels. Statt die Zahlen x_j zu addieren und durch die Anzahl n zu dividieren, multipliziert man also die x_j und zieht aus dem Produkt die n-te Wurzel (denn weil die Erhebung in die n-te Potenz das multiplikative Analogon der Vervielfachung mit dem Faktor n ist, hat man das Ziehen der n-ten Wurzel als multiplikatives Analogon der Division durch n anzusehen). Somit ist das **geometrische Mittel** definiert durch

$$GM(x_1, x_2, \ldots, x_n) \quad := \quad \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n} \quad = \quad \prod_{j=1}^{n} x_j^{1/n} \qquad\qquad (x_j > 0)\,.$$

Der Name "geometrisches Mittel" kommt von folgender geometrischen Interpretation des geometrischen Mittels $\sqrt{ab}$ von zwei positiven Zahlen: Dieser Mittelwert $\sqrt{ab}$ ist gerade die Seitenlänge eines Quadrats mit demselben Flächeninahlt ab wie ein Rechteck mit Seitenlängen a und b. Die Positivität der Zahlen x_j, die man hier mittelt, ist wichtig. Bei negativen Faktoren kann das Produkt negativ werden und das Wurzelziehen ist dann nicht möglich. Bei nichtnegativen Faktoren ist das Produkt Null und damit das geometrische Mittel gleich dem Minimum der zu mittelnden Zahlen, wenn eine davon den Wert Null hat; deshalb ist es nicht besonders sinnvoll, für die x_j hier auch den Wert Null zuzulassen, und wir tun das auch nicht. Ein wesentlicher Unterschied zwischen dem arithmetischen und dem geometrischen Mittel ist, dass letzteres von "Ausreißern nach oben" viel weniger stark beeinflusst wird als die arithmetische Mittelbildung. Ein extremes Beispiel mag das verdeutlichen: Mittelt man 10 Zahlen, von denen 9 gleich 1 sind und eine gleich 1000, so ist das arithmetische Mittel größer als 100, das geometrische Mittel aber kleiner als 2!

Da Logarithmieren die Multiplikation in Addition verwandelt, kann man die Bildung des geometrischen Mittels der x_j auch so beschreiben: Man schreibt die Zahlen als Potenzen einer Basis $a > 1$, also $x_j = a^{\xi_j}$ mit $\xi_j = \log_a x_j$, und mittelt dann die Exponenten ξ_j arithmetisch:

$$GM(x_1, x_2, \ldots, x_n) = a^{AM(\xi_1, \xi_2, \ldots, \xi_n)} \qquad (x_j = a^{\xi_j} > 0)\,.$$

Hier sieht man nun, wie man zu vorgegebenen positiven Gewichten $s_j > 0$ mit $s_1 + s_2 + \ldots + s_n = 1$ das **gewichtete geometrische Mittel** der positiven Zahlen x_j zu definieren hat: Man bildet einfach das entsprechend gewichtete arithmetische Mittel der Exponenten, was man wegen $a^{s_j \xi_j} = x_j^{s_j}$ auch so schreiben kann:

$$GM(x_1, x_2, \ldots, x_n; s_1, s_2, \ldots, s_n) \quad := \quad \prod_{j=1}^{n} x_j^{s_j} \qquad (x_j > 0)\,.$$

Das geometrische Mittel im Verhältnis der Referenzgrößen $r_j > 0$ mit Summe r ist dann

$$\Big(\prod_{j=1}^{n} x_j^{r_j}\Big)^{\frac{1}{r}} \qquad (x_j > 0, \quad r = r_1 + r_2 + \ldots + r_n)\,.$$

Wählt man alle r_j gleich, also $s_j = \frac{1}{n}$ für $j = 1 \ldots n$, so ergibt sich das gewöhnliche ("ungewichtete") geometrische Mittel.

5) Eine Anwendung in der Ökonomie findet das geometrische Mittel bei der Berechnung des effektiven Zinssatzes für einen Vorgang, bei dem ein Kapital in n Zinsperioden gleicher Dauer $\frac{1}{m}$ (Jahre, $m \in \mathbb{N}$) mit evtl. unterschiedlichen Aufzinsungsfaktoren $q_1, \ldots, q_n$ verzinst wird. Der **mittlere Aufzinsungsfaktor** $q_* > 1$ für eine Periode ist dann derjenige, der bei Anwendung in jeder Zinsperiode nach n Perioden dasselbe Ergebnis liefert. Die Gleichung für q_* ist also $q_*^n \cdot K_0 = q_1 \cdot q_2 \cdot \ldots \cdot q_n \cdot K_0$, und die Lösung ist $q_* = \sqrt[n]{q_1 \cdot q_2 \cdot \ldots \cdot q_n} = GM(q_1, q_2, \ldots, q_n)$. Der äquivalente Jahresbezogene Aufzinsungsfaktor ist dann q_*^m und der dadurch bestimmte Jahreszinsfuß, der sogenannte **effektive Zinssatz** $p_{\text{eff}}\%$ für den Verzinsungsvorgang, ist dann gegeben durch $p_{\text{eff}} = 100 \cdot (q_*^m - 1)\%$. Bei einfacher Verzinsung wäre dagegen das arithmetische Mittel der q_j der mittlere Aufzinsungsfaktor für eine Periode gewesen, und als Effektivzinssatz hätte man das arithmetische Mittel der in den einzelnen Perioden verwendeten Jahreszinsfüße erhalten.

Hat man unterschiedliche Aufzinsungsfaktoren q_j, die evtl. in mehreren Zinsperioden Anwendung finden, so ist der mittlere Aufzinsungsfaktor q_* entsprechend das geometrische Mittel der q_j im Verhältnis der Anzahlen m_j der Perioden, in denen mit Aufzinsungsfaktor q_j verzinst wird. Bei kontinuierlicher Verzinsung zu Zinsfüßen p_j in n Perioden beliebiger Dauer t_j ist der Jahres-bezogene mittlere Aufzinsungsfaktor das geometrische Mittel der Jahres-bezogenen Aufzinsungsfaktoren $e^{p_j/100}$ im Verhältnis der Dauern t_j, also

$$\Big[(e^{p_1/100})^{t_1} \cdot \ldots \cdot (e^{p_n/100})^{t_n}\Big]^{\frac{1}{t}} = e^{(t_1 p_1 + \ldots + t_n p_n)/t \cdot 100} =: e^{\tilde{p}/100} \qquad (t = t_1 + \ldots + t_n)\,,$$

und der entsprechende mittlere Jahreszinsfuß $\tilde{p}\%$ ist, wie man sieht, das arithmetische Mittel der in den einzelnen Perioden verwendeten Zinsfüße im Verhältnis der Länge dieser Perioden. (Der effektive Zinssatz, d.h. der konforme Zinsfuß in einem Vergleichsverfahren mit jährlichem Zinszuschlag, ist hier $p_{\text{eff}} = 100 \cdot (e^{\tilde{p}/100} - 1)$ Prozent, und das ist natürlich größer als $\tilde{p}$ Prozent, weil bei kontinuierlicher Verzinsung ja ein kleinerer Zinsfuß schon dasselbe Jahresergebnis liefert.)

6) Für n gegebene nichtnegative Zahlen $x_1, x_2, \ldots, x_n$ heißt die Quadratwurzel aus dem arithmetischen Mittel ihrer Quadrate das **quadratische Mittel**,

$$QM(x_1, x_2, \ldots, x_n) \quad := \quad \sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}} \quad = \quad \Big(\frac{1}{n}\sum_{j=1}^{n} x_j^2\Big)^{\frac{1}{2}}.$$

Bildet man das arithmetische Mittel der Quadrate zu Gewichten $s_1, s_2, \ldots, s_n > 0$ mit $s_1 + s_2 + \ldots + s_n = 1$, so erhält man das entsprechende **gewichtete quadratische Mittel**:

$$QM(x_1, x_2, \ldots, x_n; s_1, s_2, \ldots, s_n) \quad := \quad \sqrt{s_1 x_1^2 + s_2 x_2^2 + \ldots + s_n x_n^2} \quad = \quad \Big(\sum_{j=1}^{n} s_j x_j^2\Big)^{\frac{1}{2}}.$$

Zwar sind diese Ausdrücke auch definiert, wenn einige der Zahlen x_j negativ sind, aber dann liegen sie nicht mehr unbedingt zwischen dem Minimum und dem Maximum der x_j, können also nicht mehr als Mittelwert aufgefasst werden. (Wenn z.B. alle x_j negativ sind, so ist $QM(x_1, x_2, \ldots, x_n) > 0$ sicher größer als das Maximum.) Daher beschränken wir uns beim quadratischen Mitteln auf nichtnegative Zahlen. Das quadratische Mittel wird von "Ausreißern nach oben" noch mehr beeinflußt als das arithmetische. Bei dem schon betrachteten extremen Beispiel, in dem 10 Zahlen gemittelt werden, von denen eine gleich 1000 ist und 9 gleich 1 sind, ist z.B. das arithmetische Mittel 100.9, das quadratische aber ≈ 316.23.

7) Quadratische Mittel sind von großer Bedeutung für die Wahrscheinlichkeitsrechnung und Statistik sowie für die sog. *Fehlerausgleichsrechnung*. Dabei sind z.B. Werte x_j einer mit zufälligen Schwankungen behafteten Variablen erhoben (etwa Meßgrößen in der Qualitätskontrolle), und man sucht einen vernünftig erscheinenden Wert $\overline{x}$, den man als "*wahren Wert*" der Größe interpretieren kann. Nach einem Vorschlag von Gauß kann man z.B. $\overline{x}$ durch die Bedingung festlegen, dass das quadratische Mittel aller "*Fehler*" $|x_j - \overline{x}|$ möglichst klein wird; dies ist die sog. **Methode der kleinsten Quadrate**. Mit quadratischer Ergänzung können wir $\overline{x}$ dann leicht bestimmen: Aus

$$\frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2 = \frac{1}{n}\sum_{j=1}^{n} x_j^2 - \frac{2\overline{x}}{n}\sum_{j=1}^{n} x_j + \overline{x}^{\,2} = \frac{1}{n}\sum_{j=1}^{n} x_j^2 + \Big(\overline{x} - \frac{1}{n}\sum_{j=1}^{n} x_j\Big)^2 - \Big(\frac{1}{n}\sum_{j=1}^{n} x_j\Big)^2$$

lesen wir nämlich ab, dass $\frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2$ genau dann am kleinsten wird, wenn wir für $\overline{x}$ das arithmetische Mittel der x_j einsetzen. Bei dieser Festsetzung des "wahren Wertes" $\overline{x}$ ist dann das quadratische Mittel der Fehler $|x_j - \overline{x}|$ gleich

$$\sqrt{\frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2} = \sqrt{\frac{1}{n}\sum_{j=1}^{n} x_j^2 \; - \overline{x}^{\,2}} \qquad (\overline{x} := \tfrac{1}{n}\textstyle\sum_{j=1}^{n} x_j).$$

Diese Größe heißt **mittlerer Fehler** oder **Standardabweichung** der x_j von ihrem arithmetischen Mittel. Das einfacher zu handhabende Quadrat des Ausdrucks nennt man die **Varianz** der Werte $x_1, \ldots, x_n$,

$$\frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2 = \frac{1}{n}\sum_{j=1}^{n} x_j^2 \; - \overline{x}^{\,2} \qquad (\overline{x} := \tfrac{1}{n}\textstyle\sum_{j=1}^{n} x_j).$$

Mittlerer Fehler und Varianz sind ein Maß für die **Streuung** der Werte $x_1, \ldots, x_n$ der Zufallsgröße. Eine große Varianz deutet z.B. auf große Qualitätsunterschiede hin, wenn es um Qualitätskontrolle geht, oder auf große Meßfehler. Meist ist weniger die Standardabweichung von Interesse als ihr Verhältnis zur Größe des "wahren Wertes" $\overline{x}$, jedenfalls in Situationen, in denen $\overline{x} > 0$ groß ist gegenüber den Fehlern $|x_j - \overline{x}|$. Dieses Verhältnis heißt die **relative Standardabweichung** oder der **Variationskoeffizient**

$$\frac{1}{\overline{x}} \sqrt{\frac{1}{n} \sum_{j=1}^{n} (x_j - \overline{x})^2} = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \frac{x_j^2}{\overline{x}^2} - 1} \qquad (\overline{x} := \tfrac{1}{n} \textstyle\sum_{j=1}^{n} x_j).$$

Eine etwas allgemeinere Aufgabe besteht darin, aus Erhebungen der Werte $x_1, \ldots, x_n$ und $y_1, \ldots, y_n$ zu zwei Variablen, für die ein linearer Zusammenhang $y = ax + b$ angenommen wird, die unbekannten Koeffizienten a, b möglichst gut zu bestimmen. Man kann dies so sehen, dass in der (x, y)-Ebene eine Gerade gesucht ist, welche möglichst nahe bei den Punkten (x_j, y_j), $j = 1 \ldots n$, verläuft. Nimmt man als Maß für die Abweichung der Geraden von den Punkten wieder das quadratische Mittel der "Fehler" $|y_j - ax_j - b|$, d.h. der Entfernung der Punkte (x_j, y_j) von den vertikal darüber oder darunter liegenden Punkten auf der Geraden, so kann man die Aufgabe ganz ähnlich wie oben mit quadratischer Ergänzung lösen. Sind $\overline{x}$ und $\overline{y}$ wie oben die arithmetischen Mittel der x_j bzw. der y_j, so lesen wir aus

$$\frac{1}{n} \sum_{j=1}^{n} (y_j - ax_j - b)^2 = \sum_{j=1}^{n} \frac{(y_j - ax_j)^2}{n} - 2b(\overline{y} - a\overline{x}) + b^2 = \sum_{j=1}^{n} \frac{(y_j - ax_j)^2}{n} + (b - \overline{y} + a\overline{x})^2 - (\overline{y} - a\overline{x})^2$$

direkt ab, dass $b = \overline{y} - a\overline{x}$ sein muss, wenn der quadratische Fehler minimal sein soll. Und wenn wir diesen Wert für b einsetzen, so zeigt

$$\sum_{j=1}^{n} \Big[(y_j - \overline{y}) - a(x_j - \overline{x}) \Big]^2 = \sum_{j=1}^{n} (y_j - \overline{y})^2 - 2a \sum_{j=1}^{n} (x_j - \overline{x})(y_j - \overline{y}) + a^2 \sum_{j=1}^{n} (x_j - \overline{x})^2$$

$$= \sum_{j=1}^{n} (y_j - \overline{y})^2 + \left(\frac{\sum_{j=1}^{n} (x_j - \overline{x})(y_j - \overline{y})}{\sqrt{\sum_{j=1}^{n} (x_j - \overline{x})^2}} - a \sqrt{\sum_{j=1}^{n} (x_j - \overline{x})^2} \right)^2 - \frac{\left(\sum_{j=1}^{n} (x_j - \overline{x})(y_j - \overline{y}) \right)^2}{\sum_{j=1}^{n} (x_j - \overline{x})^2},$$

dass der quadratische Fehler minimal wird, genau wenn wir die gesuchten Zahlen a und b wie folgt wählen:

$$a = \frac{\sum_{j=1}^{n} (x_j - \overline{x})(y_j - \overline{y})}{\sum_{j=1}^{n} (x_j - \overline{x})^2}, \qquad b = \overline{y} - a\overline{x} = \frac{1}{n} \sum_{j=1}^{n} (y_j - ax_j)$$

Die hierdurch festgelegte Gerade, also der Graph der linearen Funktion $y = ax + b$, heißt die **Regressionsgerade** zu den Werten $x_1, \ldots, x_n$ und $y_1, \ldots, y_n$. Aufgabenstellungen wie die hier behandelte, und allgemeinere ähnlicher Art, zur Auswertung erhobener Daten (Meßwerte) fasst man zusammen unter dem Begriff **Regressionsanalyse**, was ein wichtiges Teilgebiet der Statistik und auch der Ökonometrie ist.

8) Der **Potenzmittelwert zum Exponenten t** von $x_1, x_2, \ldots, x_n \in \mathbb{R}_{>0}$ ist definiert durch

$$PM^t(x_1, x_2, \ldots, x_n) := \Big(\frac{1}{n}\sum_{j=1}^{n} x_j^t\Big)^{1/t} = \Big(\frac{x_1^t + x_2^t + \ldots + x_n^t}{n}\Big)^{1/t},$$

also als die t-te Wurzel aus dem arithmetischen Mittel der t-ten Potenzen der x_j; dabei darf t ein beliebiger reeller Exponent $\neq 0$ sein, und die zu mittelnden Zahlen müssen positiv sein (für $t > 0$ kann man auch den Wert Null noch zulassen). Für $t = 1$ ergibt sich z.B. das arithmetische Mittel, für $t = 2$ das quadratische Mittel, und für $t = -1$ erhält man das Reziproke des arithmetischen Mittels der Kehrwerte der x_j, das sogenannte **harmonische Mittel**

$$HM(x_1, x_2, \ldots, x_n) := \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}.$$

Für $t = 0$ sind die Potenzmittelwerte nicht definiert, wir werden aber sehen, dass es sinnvoll ist, das geometrische Mittel als Potenzmittelwert zum Exponenten Null festzulegen, also PM^0 als GM zu definieren. Die Potenzmittelwerte berücksichtigen "Ausreißer nach oben" um so stärker / weniger, je größer / kleiner der Exponent t ist. In unserem Beispiel der Mittelung von 10 Zahlen, von denen 9 gleich 1 sind und eine gleich 1000 ist, ergeben sich für die Exponenten $t = -2, -1, 0, 1, 2, 3$ die Potenzmittelwerte $1.054092\ldots$, $1.110987\ldots$, $1.995263\ldots$, 100.9, $316.2291\ldots$, $464.1588\ldots$.

Bildet man statt des gewöhnlichen arithmetischen Mittels der Potenzen x_j^t oben das gewichtete arithmetische Mittel zu Gewichten $s_1, s_2, \ldots, s_n > 0$ mit $s_1 + s_2 + \ldots + s_n = 1$, so erhält man die **gewichteten Potenzmittelwerte zum Exponenten t**,

$$PM^t(x_1, x_2, \ldots, x_n; s_1, s_2, \ldots, s_n) := \Big(s_1 x_1^t + s_2 x_2^t + \ldots + s_n x_n^t\Big)^{1/t}.$$

9) Eine Mittelbildung ganz anderer Art ist folgende: Man numeriert die zu mittelnden Zahlen $x_1 \leq x_2 \leq \ldots \leq x_n$ der Größe nach und wählt im Fall einer ungeraden Anzahl n die Zahl $x_{(n+1)/2}$ in der mittleren Position, im Fall gerader Anzahl n einen Wert zwischen den beiden Zahlen $x_{n/2}$ und $x_{1+n/2}$ in der mittleren Position. Den so erhaltenen Mittelwert nennt man einen **Median** der Zahlen $x_1, x_2, \ldots, x_n$. Mediane sind dadurch gekennzeichnet, dass ebenso viele der zu mittelnden Zahlen über dem Median liegen wie unter ihm. Die genaue Festlegung bei einer geraden Anzahl n von Zahlen erfolgt meist als **mittlerer Median**, d.h. als arithmetischer Mittelwert $\frac{1}{2}(x_{n/2} + x_{1+n/2})$ der beiden Zahlen in mittlerer Position bei Anordnung der Größe nach. Wir schreiben

$$Med(x_1, x_2, \ldots, x_n)$$

für den Median der Zahlen x_j bei ungerader Anzahl n bzw. für ihren mittleren Median bei gerader Anzahl n. Andere Möglichkeiten der Festlegung bei geradem n sind der kleinstmögliche Median $x_{n/2}$ oder der größtmögliche Median $x_{1+n/2}$. Da es bei der Bildung eines Medians nur darauf ankommt zu zählen, wieviele der Zahlen x_j unter einem gewissen Wert liegen und wieviele darüber, nicht aber auf die Größe der Zahlen, *sind Mediane völlig unsensibel gegen Ausreißer* — darin liegt ihre praktische Bedeutung. In unserem Beispiel der Mittelung von 9 Zahlen mit Wert 1 und einer mit Wert 1000 ist der Median z.B. gleich 1, und das würde auch so bleiben, wenn wir vier der Zahlen beliebig vergrößern würden.

Wegen ihrer Unempfindlichkeit gegen Ausreißer sind Mediane in vielen Situationen die sachlich angemessenen Mittelwerte, z.B. auch bei der Angabe durchschnittlicher Studiendauern. Wie andererseits Unsinn auch mit Medianen getrieben werden kann, zeigt eine

Statistik des Wissenschaftsrates über Studiendauern (1985). Dort wurde als Mittelwert der um ein halbes Semester erhöhte kleinstmögliche Median verwendet. In einem Diplomstudiengang einer Hochschule ergab sich so eine durchschnittliche Sudiendauer von 11.5 Semestern, womit die Hochschule zu den drei Universitäten mit den kürzesten Studiendauern im betreffenden Fach zählte. Tatsächlich hatten im fraglichen Jahr dort aber nur 4 Studierende den Studiengang abgeschlossen, und zwar zwei nach 11 Semestern und zwei nach 17 Semestern. Schon beim mittleren Median wäre die Hochschule im letzten Drittel, beim genau so gerechtfertigten Median 16.5 gar das Schlusslicht in bezug auf die Studiendauer in diesem Fach gewesen. Soviel zum "Ranking von Hochschulen" auf der Basis unsinniger Mittelwertbildungen. (Zur Ehrenrettung des Wissenschaftsrates sei gesagt, dass er sein Mittelungsverfahren und die zugrunde liegenden Daten genau angegeben hat, so dass es möglich war, die unsinnige Mittelwertangabe überhaupt zu erkennen.)

Die Medianbildung zu Gewichten kommt in der Praxis kaum vor. Wenn man an verschiedene der Größe nach numerierte Zahlen x_j denkt, die mit Fallzahlen $m_j \geq 1$ auftreten, welche sich zur Gesamtzahl m addieren, so zählt man einfach für jedes j die Zahl x_j mit der Vielfachheit m_j. Der Median ist dann die Zahl x_{k+1}, wenn $\sum_{j=1}^{k} m_j < \frac{1}{2}m < \sum_{j=1}^{k+1} m_j$ ist, und eine Zahl zwischen x_k und x_{k+1} (z.B. das arithmetische Mittel dieser beiden Werte), wenn $\sum_{j=1}^{k} m_j = \frac{1}{2}m$ ist. Da die Gewichte in dieser Situation $s_j = m_j/m$ sind, wird man allgemein für Zahlen $x_1 \leq x_2 \leq \ldots \leq x_n$ und gegebene Gewichte $s_j > 0$ mit $s_1+s_2+\ldots+s_n = 1$ den **gewichteten Median** als x_{k+1} definieren, wenn $\sum_{j=1}^{k} s_j < \frac{1}{2} < \sum_{j=1}^{k+1} s_j$ ist und als eine Zahl zwischen x_k und x_{k+1} (z.B. als $\frac{1}{2}x_k + \frac{1}{2}x_{k+1}$ beim mittleren gewichteten Median), wenn $\sum_{j=1}^{k} s_j = \frac{1}{2}$ ist. ■

Nachdem wir nun viele verschiedene (aber in der Praxis gebräuchliche) Mittelwertbildungen eingeführt haben, ist es angebracht, etwas zum Größenvergleich für die verschiedenen Mittelwerte zu sagen. Der folgende in der Mathematik bewiesene Lehrsatz enthält alle Vergleichsaussagen, die man in dieser Hinsicht generell machen kann.

SATZ (über den Vergleich verschiedener Mittelwerte):

Für positive reelle Zahlen $x_1, \ldots, x_n > 0$ gilt:

(i) *die* **Ungleichung zwischen dem harmonischen, geometrischen und arithmetischen Mittel:**

$$HM(x_1, \ldots, x_n) \leq GM(x_1, \ldots, x_n) \leq AM(x_1, \ldots, x_n)$$

mit Gleichheit in einer der beiden Ungleichungen nur wenn $x_1 = \ldots = x_n$;

(ii) *die* **Ungleichung zwischen Potenzmittelwerten:**

$$PM^r(x_1, \ldots, x_n) \leq PM^t(x_1, \ldots, x_n) \qquad \text{für} \quad r < t \quad \text{in} \ \mathbb{R}$$

mit Gleichheit nur wenn $x_1 = \ldots = x_n$;

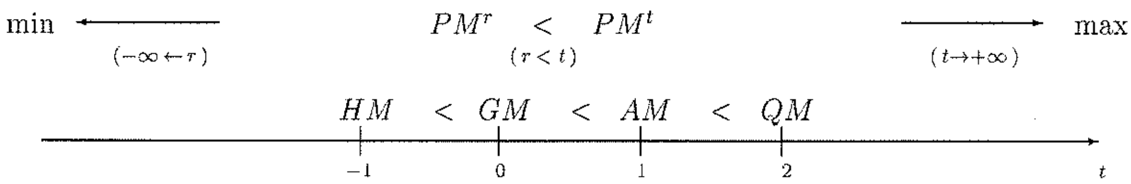
(iii) *der* **Wertebereich der Potenzmittelwerte** *$PM^t(x_1, \ldots, x_n)$ überdeckt das ganze Intervall zwischen dem Minimum der x_j und dem Maximum der x_j wenn man den Exponenten t von $-\infty$ bis $+\infty$ laufen läßt.*

Dabei ist $PM^0(x_1, \ldots, x_n)$ als geometrisches Mittel $GM(x_1, \ldots, x_n)$ definiert. Die Aussagen (i) - (iii) gelten auch, wenn man alle Mittelwerte zu denselben vorgegebenen Gewichten $s_1, \ldots, s_n > 0$ mit $s_1 + \ldots + s_n = 1$ bildet. ■

Die Aussage (iii) erklärt, was wir zu Beginn dieses Abschnitts gesagt haben: Zu jeder gewünschten Zahl zwischen dem Minimum und dem Maximum der zu mittelnden Zahlen gibt es eine Mittelwertformel, welche genau die gewünschte Zahl als Ergebnis liefert, nämlich ein Potenzmittelwert $(x_1^t + \ldots + x_n^t)^{1/t}$ mit passendem reellen Exponenten t. Natürlich würde die Wahl eines extremen Exponenten wie $t = 100$ auffallen, wenn man manipulieren möchte. Aber schon die Auswahl aus den Exponenten -1, 0, 1 und 2, also aus harmonischem, geometrischen, arithmetischem und quadratischen Mittel, die alle gebräuchlich sind, kann das Ergebnis der Mittelung erheblich in die gewünschte Richtung beeinflussen. In Abwandelung eines Ausspruchs, der von Churchill kolportiert wird ("Traue keiner Statistik, die du nicht selbst gefälscht hast"), könnte man daher sagen:

> "*Traue keiner Mittelwertangabe, die du nicht selbst (durch Auswahl einer geeigneten Mittelwertformel) manipuliert hast.*"

Das folgende Diagramm stellt die Aussagen des Satzes graphisch dar für den Fall, dass nicht alle der zu mittelnden positiven Zahlen gleich sind (andernfalls sind auch alle Mittel gleich):



Für die Ungleichung "$GM \leq AM$" gibt es eine einleuchtende ökonomische Begründung: Für Aufzinsungsfaktoren $q_1, \ldots, q_n$ in n aufeinander folgenden Zinsjahren ist nämlich $AM(q_1, \ldots, q_n)$ der mittlere jährliche Aufzinsungsfaktor bei einfacher Verzinsung (Zinsen werden entnommen) und $GM(q_1, \ldots, q_n)$ bei Zinsesverzinsung (vgl. Bsp. 5) oben). Die Ungleichung "$GM \leq AM$" besagt hier also einfach, dass bei Zinsesverzinsung in n Jahren derselbe Zinsgewinn mit einem niedrigeren konstanten Zinsfuß erreicht wird als bei einfacher Verzinsung.

Die Ungleichung klärt aber aber auch weniger evidente Sachverhalte. Wird z.B. in 12 Monaten mit unterschiedlichen Zinsfüßen $p_1\%, \ldots, p_{12}\%$ (p.a.) einfach verzinst, so ist $AM(p_1, \ldots, p_{12})$ der entsprechende Jahres-Zinsfuß. Betrachtet man aber Zinsesverzinsung und setzt für den j-ten Monat den Monats-bezogenen Aufzinsungsfaktor $\sqrt[12]{q_j}$ an, der bei Anwendung in allen 12 Monaten den Aufzinsungsfaktor $q_j = 1 + \frac{p_j}{100}$ für das Jahr bringt, so ist $\sqrt[12]{q_j}$ natürlich kleiner als der Monats-bezogene Aufzinsungsfaktor bei einfacher Verzinsung $1 + \frac{1}{12}\frac{p_j}{100}$. (Das zeigt auch die Bewrnoulli–Ungleichung.) Daher ist nun für den Gesamtzeiteaum von einem Jahr nicht klar, ob die Zinsesverzinsung mit den Monats-bezogenen Aufzinsungsfaktoren $\sqrt[12]{q_j}$ günstiger ist oder die einfache Verzinsung mit den Monats-bezogenen Aufzinsungsfaktoren $1 + \frac{1}{12}\frac{p_j}{100}$. Die "$GM \leq AM$" Ungleichung beantwortet das: $\sqrt[12]{q_1} \cdot \ldots \cdot \sqrt[12]{q_{12}} \leq \frac{1}{12}(q_1 + \ldots + q_{12}) = 1 + \frac{1}{100}\frac{p_1+\ldots+p_n}{12}$, also ist Letzteres günstiger (und zwar echt günstiger, wenn die Zinsfüße $p_1, \ldots, p_{12}$ nicht alle gleich sind). Das ist ökonomisch nicht ohne Weiteres klar, weil wir dabei zwar größere monatliche Aufzinsungsfaktoren haben, aber keinen Zinsverzinsungseffekt.

Auch wenn es für das Verständnis der Mittelwertungleichungen nicht nötig ist, wollen wir hier zum Schluss für Interessierte den Beweis noch skizzieren; denn die Ungleichungen sind ja keineswegs evident.

Für den *Beweis von (i)* schreiben wir $x_j = \mathrm{e}^{y_j}$ und $\overline{x} := \sum_{j=1}^n s_j x_j$, $\overline{y} := \sum_{j=1}^n s_j y_j$. Mit der fundamentalen Ungleichung für die Exponentialfunktion aus 1.4 erhalten wir dann $\mathrm{e}^{y_j - \overline{y}} \geq 1 + y_j - \overline{y}$, und Multiplikation der Ungleichungen mit s_j sowie Addition gibt $\overline{x}/\mathrm{e}^{\overline{y}} \geq 1$ mit Gleichheit nur, wenn $y_j = \overline{y}$ ist für alle j, also auch $x_j = \overline{x}$. Das ist die Ungleichung "$GM \leq AM$", und "$HM \leq GM$" folgt durch Übergang zu Kehrwerten.

Für den *Beweis von (ii)* im Fall $r = 1 < t$ kann man $\sum_{j=1}^n s_j x_j^t = 1$ annehmen (sonst jedes x_j durch die t-te Wurzel aus dieser Summe dividieren). Aus der "$GM \leq AM$" Ungleichung für zwei Zahlen erhalten wir $s_j x_j = s_j^{1-1/t}(s_j x_j^t)^{1/t} \leq (1 - \frac{1}{t})s_j + \frac{1}{t} s_j x_j^t$ und damit $\sum_{j=1}^n s_j x_j \leq (1 - \frac{1}{t}) + \frac{1}{t} \cdot 1 = 1$. Das ist die Ungleichung $PM^1 \leq PM^t$ in diesem Fall. Indem man darin x_j durch x_j^r ersetzt und t durch $\frac{t}{r}$, folgt $PM^r \leq PM^t$ für $0 < r < t$. Für $r < t < 0$ erhält man dieselbe Ungleichung durch Übergang zu Kehrwerten. Für $r < 0 < t$ ist schließlich noch $PM^t = (\sum_{j=1}^n s_j x_j^t)^{1/t} \geq (\prod_{j=1}^n x_j^{s_j t})^{1/t} = GM = PM^0 = (\prod_{j=1}^n x_j^{s_j r})^{1/r} \geq (\sum_{j=1}^n s_j x_j^r)^{1/r} = PM^r$ gemäß (i). Inspektion der Argumente zeigt auch, dass Gleichheit $PM^r = PM^t$ für $r < t$ nur eintritt, wenn alle x_j gleich sind.

Zum *Beweis von (iii)* nehmen wir an, dass etwa x_n die größte der Zahlen x_j ist und bemerken $x_n \geq PM^t \geq s_n^{1/t} x_n$. Für hinreichend große t liegt $s_n^{1/t} = \mathrm{e}^{(1/t)\ln s_n}$ beliebig nahe bei 1 und daher PM^t beliebig nahe beim Maximum x_n . Auf der anderen Seite liegen für hinreichend kleine $t > 0$ die Zahlen $x_j^t = \mathrm{e}^{t \ln x_j}$ beliebig nahe bei 1, und mit der fundamentalen Ungleichung aus 1.4 für die Logarithmusfunktion folgt daher $\ln PM^t = \frac{1}{t}\ln(\sum_{j=1}^n s_j x_j^t) \leq \frac{1}{t}(\sum_{j=1}^n s_j x_j^t - 1) = \frac{1}{t}\sum_{j=1}^n s_j(x_j^t - 1) = \frac{1}{t}\sum_{j=1}^n x_j^t s_j(1 - x_j^{-t}) \leq \frac{1}{t}\sum_{j=1}^n s_j x_j^t(-\ln x_j^{-t}) = \sum_{j=1}^n s_j x_j^t \ln x_j \approx \sum_{j=1}^n s_j \ln x_j = \ln GM$, wobei der im vorletzten Schritt gemachte Fehler beliebig klein ist, wenn $t > 0$ klein genug. Mit Exponentiieren ergibt sich somit, dass PM^t für hinreichend kleine $t > 0$ beliebig wenig größer als GM ist. Damit nimmt $PM^t = (\sum_{j=1}^n s_j x_j^t)^{1/t}$ Werte beliebig nahe beim Maximum der x_j und beliebig nahe beim geometrischen Mittel $GM = PM^0$ an, wenn t im Intervall $]0, \infty[$ variiert. Weil PM^t auch stetig von t abhängt (d.h. die Änderung von PM^t ist kleiner als eine beliebig vorgegebene positive Größe, wenn man $t > 0$ hinreichend wenig verändert — das kann man ähnlich, aber einfacher überlegen), folgt aus dem Zwischenwertsatz der Analysis (siehe 2.1), dass PM^t alle Werte zwischen $GM = PM^0$ und dem Maximum der x_j annimmt, wenn t von 0 bis ∞ läuft. Durch Übergang zu den Reziproken der Mittel für die reziproken Zahlen $1/x_j$ folgt dann, dass PM^t auch alle Werte zwischen dem Minimum und $GM = PM^0$ annimmt, wenn t von $-\infty$ nach 0 läuft. Damit sind alle Aussagen des Satzes bewiesen.

Wie gesagt, es ist für das Verständnis der Mittelwertungleichungen nicht nötig, den gerade skizzierten Beweis nachzuvollziehen – das kann man den Mathematikern überlassen. Aber man sollte sich mit dem Größenvergleich zwischen den verschiedenen in der Ökonomie gebräuchlichen Arten von Mittelwerten auskennen und, wenn ein Argument auf Mittelwerte von Daten gestützt wird, immer fragen (ggf. auch sich selbst), ob die gewählte Art der Mittelbildung sachlich angemessen ist, oder ob möglicherweise eine weniger angemessene Mittelwertbildung verwendet wurde mit der Absicht, den berechneten Mittelwert in eine genehme Richtung größer oder kleiner ausfallen zu lassen.