Gewichtete Mittel im Spiegel[1]

RUMA FALK UND AVITAL LAVIE LANN, JERUSALEM

[1] Original: „Weighted Means Through the Looking Glass?" in Teaching Statistics 35 (2013) 2, S. 103–106
Übersetzung: JOACHIM ENGEL, LUDWIGSBURG

Zusammenfassung: *Jedes gewichtete Mittel von zwei Werten hat ein Gegenstück, das im gleichen Abstand zum arithmetischen Mittel liegt. Man erhält dieses Gegenstück durch Rollentausch der beiden Gewichte oder durch Bildung reziproker Gewichte. Diese elementare Beziehung eignet sich für Einführungen in die Statistik.*

1 Einleitung

Das arithmetische Mittel ist unter allen Mittelwerten die am meisten genutzte Größe und ist sowohl Laien wie auch erfahrenen Zuhörern bestens vertraut. In alltäglicher Sprache bedeutet Durchschnitt das Resultat, das man erhält, wenn man alle Beträge addiert und die Summe dann durch die Anzahl der Summanden teilt. Das Mittel ist als „der durchschnittliche Wert einer Menge von Größen" definiert. Trotz vieler Publikationen, die verschiedene Mittel, ihre Größenordnung und ihre Beziehungen zueinander behandeln (z. B. Beckenbach und Bellman 1961; Ercolano 1973; Eves 2003; Hoehn 1984; Hoehn und Niven 1985; Nelsen 1993 S. 55–57) gibt es reichlich empirische Belege dafür, dass das arithmetische Mittel einer Menge von Werten die vorherrschende Lösung ist, die Schülerinnen und Schüler wählen, wenn nach einem Durchschnitt verlangt wird. Und oft denken sie auch in Situationen, die ungleiche Gewichtungen verlangen, dass gleiche Gewichte den gemittelten Werten beigegeben werden sollen.

Von einer Vielfalt von Durchschnitten schauen wir in diesem Aufsatz nur auf gewichtete Mittelwerte. Wir halten es für wichtig, dass Lernende sich von Beginn an eine breitere Vorstellung von Durchschnitten aneignen, von denen das arithmetische Mittel nur ein wichtiges Beispiel darstellt.

2 Zwei spezielle gewichtete Mittel

Ein gewichtetes Mittel W von n positiven Zahlen x_1, $x_2, \ldots, x_n$ ist definiert als

$$W = \frac{w_1 x_1 + w_2 x_2 + \ldots + w_n x_n}{w_1 + \ldots + w_n}$$

mit positiven Gewichten $w_1, w_2, \ldots, w_n$.

Lahn und Falk (2005, 2006) konzentrierten sich auf zwei spezielle gewichtete Mittel: das selbstgewichtete Mittel SW, bei welchem jeder Wert mit sich selbst gewichtet wird, und das harmonische Mittel H, bei dem jeder Wert mit seinem reziproken (oder inversen) Wert gewichtet ist:

$$SW = \frac{x_1^2 + x_2^2 + \ldots + x_n^2}{x_1 + \ldots + x_n}$$

$$H = \frac{\frac{1}{x_1}x_1 + \frac{1}{x_2}x_2 + \ldots + \frac{1}{x_n}x_n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

Stochastik in der Schule 34 (2014) 1, S. 22–25

Da das Gewichten eines Wertes mit sich selbst den Einfluss großer Zahlen verstärkt und das Gewichten mit dem Reziproken den Einfluss kleiner Zahlen vergrößert, ergibt sich für diese Gewichtsschemata $SW \geq A \geq H$, wobei A das arithmetische Mittel von n Zahlen bezeichnet.

Andere Bezeichnungen in der Literatur beziehen sich direkt auf die Reziprozitätseigenschaft für Gewichte: SW wird bei Hoehn (1984) *antiharmonisches Mittel* und bei Eves (2003) und Nelson (1993) *kontraharmonisches Mittel* genannt.

Die Formel für das harmonische Mittel von nur zwei Werten a und b reduziert sich zu

$$H = \frac{2ab}{a+b}$$

In diesem Fall haben SW und H den gleichen Abstand zu A, wie sich algebraisch leicht verifizieren lässt. Dies wird im Folgenden als Spezialfall einer allgemeinen Eigenschaft von Paaren gewichteter Mittel hergeleitet werden. Man beachte, dass nur für n = 2 und nicht allgemein für jede Anzahl von Werten die Differenzen $SW - A$ und $A - H$ gleich sind. Zum Beispiel gilt für die Mittel der drei Zahlen 1, 2, 6 dass $A = 3{,}0$; $SW = 4{,}6$ und $H = 1{,}8$ und somit $4{,}6 - 3{,}0 \neq 3{,}0 - 1{,}8$. Im Folgenden beschränken wir uns nur noch auf gewichtete Mittel von nur zwei Zahlen.

3 Gewichtete Mittel von zwei Werten kommen als Zwillingspaare

Es sei $b > a > 0$. Jeder Punkt im offenen Intervall (a, b) ist ein gewichtetes Mittel $pa + qb$ von a und b mit einem Paar normalisierter Gewichte p und q, die den Bedingungen $0 < p, q < 1$ und $p + q = 1$ genügen. Wir nehmen an, dass $p \neq q$ (für $p = q = 0{,}5$ erhält man das arithmetische Mittel A und die Sätze, die gewichtete Mittel betreffen, gelten trivialerweise auch für A). Die gewichteten Mittel SW und H nehmen mit normalisierten Gewichten die Form an

$$SW = \frac{a}{a+b}a + \frac{b}{a+b}b,$$

$$H = \frac{b}{a+b}a + \frac{a}{a+b}b$$

und wir sehen, dass man diese zwei gewichteten Mittel gerade dadurch erhält, dass man die Gewichte von a und b austauscht.

Interessanterweise hat jeder Punkte im Intervall von a bis b ein ähnliches Gegenstück. Wir listen vier mathematische Sätze bezüglich dieser Paare auf. Der Klarheit willen wird jeder Satz separat bewiesen, obwohl das nicht der sparsamste Weg ist.

Satz 1: Ersetzt man jedes Gewicht von a und b bei einer gewichteten Mittelwertbildung durch seinen reziproken Wert, so erhält man ein anderes gewichtetes Mittel, bei dem die Gewichte von a und b gerade vertauscht sind.

Beweis: Es sei $W_1 = pa + qb$. Invertieren wir jedes Gewicht, so erhält man

$$W_2 = \frac{\frac{1}{p}a + \frac{1}{q}b}{\frac{1}{p} + \frac{1}{q}},$$

aber

$$\frac{\frac{1}{p}}{\frac{1}{p} + \frac{1}{q}} = \frac{pq}{p(p+q)} = \frac{pq}{p} = q;$$

Ähnlich zeigt man

$$\frac{\frac{1}{q}}{\frac{1}{p} + \frac{1}{q}} = p,$$

und somit $W_2 = qa + pb$.

Satz 2: Ein Austauschen der Gewichte zwischen a und b bei einem gewichteten Mittel resultiert in einem anderen gewichteten Mittel, bei dem die Gewichte invertiert sind.

Beweis: Es sei $W_1 = pa + qb$. Ein Austauschen der Gewichte führt zu $W_2 = qa + pb$. Nun ist

$$q = \frac{pq}{p} = \frac{pq}{p(p+q)} = \frac{\frac{1}{p}}{\frac{p+q}{pq}} = \frac{\frac{1}{p}}{\frac{1}{p} + \frac{1}{q}};$$

Ganz analog ergibt sich

$$p = \frac{\frac{1}{q}}{\frac{1}{p} + \frac{1}{q}},$$

so dass

$$W_2 = \frac{\frac{1}{p}a + \frac{1}{q}b}{\frac{1}{p} + \frac{1}{q}}.$$

Satz 3: Austauschen der Gewichte eines gewichteten Mittels von a und b resultiert in einem anderen gewichteten Mittel, das denselben Abstand zu A hat.

Beweis: Es sei $W_1 = pa + qb$. Ein Austauschen der Gewichte führt zu $W_2 = qa + pb$. Ohne Einschränkung der Allgemeinheit können wir annehmen, dass $p > q$, so dass $W_2 > W_1$. Es sei $p = 0{,}5 + d$ und $q = 0{,}5 - d$, wobei $0 < d < 0{,}5$. Dann folgt

$$\begin{aligned} W_2 - A &= (0{,}5 - d)a + (0{,}5 + d)b - 0{,}5a - 0{,}5b \\ &= d(b-a), \end{aligned}$$

und

$$A - W_1 = 0{,}5a + 0{,}5b - (0{,}5 + d)a - (0{,}5 - d)b$$
$$= d(b - a) = W_2 - A\,.$$

Satz 4: Wenn zwei gewichtete Mittel die gleiche Entfernung zu A haben, dann sind ihre Gewichte vergetauscht.

Beweis: Es sei $W_1 = p_1 a + (1 - p_1)b$ und $W_2 = p_2 a + (1 - p_2)b$, wobei $0 < p_1,\ p_2 < 1$. Ohne Einschränkung der Allgemeinheit nehmen wir an, dass $W_2 > W_1$. Vorausgesetzt $W_2 - A = A - W_1$, erhält man

$$p_2 a + (1 - p_2)b - 0{,}5a - 0{,}5b$$
$$= 0{,}5a + 0{,}5b - p_1 a - (1 - p_1)b\,.$$

Dies reduziert sich zu

$$a(p_1 + p_2) + b(1 - p_1 + 1 - p_2) = a + b\,.$$

Daher ist $p_1 + p_2 = 1$, oder $p_2 = 1 - p_1$ und $1 - p_2 = p_1$, so dass $W_2 = (1 - p_1)a + p_1 b$, d. h. die Gewichte von a und b in W_1 wurden gerade vertauscht.

Wir wiederholen – bei einem gewichteten Mittel sind die Gewichte von a und b genau dann und nur dann vertauscht, wenn jedes Gewicht invertiert ist, und genau dann wenn die beiden sich ergebenden gewichteten Mittel symmetrisch um das arithmetische Mittel liegen. Insbesondere haben nicht nur SW und H gleichen Abstand zu A, sondern ebenso ihre jeweiligen verallgemeinerten Versionen

$$\frac{a^n a + b^n b}{a^n + b^n} \text{ und } \frac{b^n a + a^n b}{a^n + b^n},$$

die als verallgemeinerte antiharmonische und verallgemeinerte harmonische Mittel bezeichnet werden (Hoehn 1984). Wenn n immer mehr wächst, dann driften diese verallgemeinerten Mittel immer weiter auseinander hin zu den Endpunkten a und b.

4 Ein Beispiel mit Geschwindigkeiten

Betrachten wir die aus dem Leben gegriffene Frage nach der durchschnittlichen Geschwindigkeit von Autos auf der Straße. Alles hängt von den genauen Details ab, wie die Frage zu verstehen ist (Falk et al. 2005).

Zuerst nehmen wir an, dass viele Fahrzeuge mit konstanter Geschwindigkeit auf einer langen Straße entlang fahren, die eine Hälfte der Autos mit einer Geschwindigkeit von 100 km/h, die andere Hälfte der Autos fährt mit 60 km/h. Beide Arten sind homogen entlang der Straße verstreut. Ein Radargerät, platziert an einem festen Punkt am Straßenrand, erfasst die Geschwindigkeit von allen Fahrzeugen, die im Verlauf einer Stunde verbeikommen. Was wird der Mittelwert dieser aufgezeichneten Messungen sein (Lann und Falk 2005)? Schnellere Autos passieren öfter das Radargerät, und zwar proportional zu der Distanz, die sie während der vorgegebenen Zeit von 1 Stunde passieren, bzw. proportional zu ihrer Geschwindigkeit. Konsequenterweise wird jede Geschwindigkeit mit ihrer eigenen Größe gewichtet und die mittlere aufgezeichnete Geschwindigkeit ist $SW = (100^2 + 60^2)/(100 + 60) = 85$ km/h, was das arithmetische Mittel (80 km/h) der Geschwindigkeiten der Gesamtmenge aller Autos übersteigt.

Als zweites nehmen wir an, ein Auto fahre von A nach B mit einer Geschwindigkeit von 100 km/h und zurück von B nach A mit 60 km/h. Um seine mittlere Geschwindigkeit über die ganze Fahrt hinweg zu ermitteln, muss man die gesamte Fahrtstrecke durch die gesamte Fahrtzeit dividieren. Aber die Strecke ergibt sich als Produkt aus Geschwindigkeit und Zeit, so dass die Geschwindigkeit in jeder Richtung zu gewichten ist mit der Zeit, die in dieser Richtung gefahren wurde. Je langsamer die Geschwindigkeit, um so länger dauert es, dieselbe Distanz zu fahren. Daher ist das Gewicht jeder Geschwindigkeit in reziproker Beziehung zu seiner Größe und die durchschnittliche Geschwindigkeit ist das harmonische Mittel der zwei Geschwindigkeiten: $H = 2 \cdot 100 \cdot 60/(100 + 60) = 75$ km/h, was weniger ist als das arithmetische Mittel von 100 und 60.

Diese beiden Mittel – äquidistant zu $A = 80$ – haben vertauschte normalisierte Gewichte: $SW = 0{,}375 \cdot 60 + 0{,}625 \cdot 100$ und $H = 0{,}625 \cdot 60 + 0{,}375 \cdot 100$.

In beiden Fällen ist die erste intuitive Antwort von Lernenden auf die Frage nach der durchschnittlichen Geschwindigkeit das arithmetische Mittel A (Falk und Lann 2008). Es ist daher ratsam, Lernende so früh wie möglich im Statistikkurs für die Möglichkeit unterschiedlicher Gewichte zu sensibilisieren, die von der zugrunde liegenden Situation herbeigeführt wird. In der Radar-Geschichte verleiht die gleiche Zeitdauer, in der schnelle und langsame Fahrzeuge erfasst werden, den schnelleren Autos ein zusätzliches Gewicht (weil sie eine größere Strecke abdecken), während bei der Hin- und Rückfahrt zwischen A und B dieselbe Strecke mit unterschiedlicher Geschwindigkeit durchfahren wird, was den langsameren Autos ein größeres Gewicht verleiht (weil es mehr Zeit braucht).

Die Behauptung, dass die Interpretation eines Datensatzes und die Bestimmung von Häufigkeiten (oder

Gewichten) Kenntnis darüber verlangt, wie diese Information erlangt wurde, kann bis auf Fisher (1934) zurück geführt werden. Er bedauerte die ‚ganze Geringschätzung der Bedingungen der Datenerhebung' auf dem Gebiet der Humangenetik. Falk und Lann (2008) fanden heraus, dass eine vergleichbare Missachtung von Faktoren, die Anlass für ungleiche Gewichtungen geben, das Urteil von Lernenden in einem weiten Spektrum von Kontexten charakterisiert.

5 Vernunft als Schlüsselfaktor

Es ist leicht einzusehen, dass sich das gewichtete Mittel normalisierter Gewichte nicht ändert, wenn beide Gewichte p und q mit einer positiven Konstanten c multipliziert werden. Das Verhältnis p/q bleibt dabei erhalten und das gewichtete Mittel $(cpa + cqb)/(cp + cq)$ ändert sich nicht. Das bedeutet, dass der Wert des gewichteten Mittels von a und b einzig durch das Verhältnis zwischen den Gewichten bestimmt ist. In Anbetracht dieser Tatsache wird die Behauptung, dass der Austausch der Gewichte zwischen a und b gleichbedeutend dazu ist, die Gewichte zu invertieren, augenscheinlich: wenn p und q die Rollen tauschen, dann ändert sich ihr Verhältnis von p/q zu q/p, und wenn p und q durch ihre jeweiligen reziproken Zahlen ersetzt werden, dann ändert sich das Verhältnis auf die gleiche Art.

6 Schlussfolgerung

Obwohl scheinbar trivial, so müssen wir bekennen, dass wir uns trotz der Vertrautheit mit der Beziehung zwischen SW und H bisher noch nicht der Allgemeinheit der Aussagen bewusst waren, dass jedes gewichtete Mittel von zwei Zahlen einen Zwillingspartner hat, im gleichen Abstand zu A, bei dem die Gewichte gerade die Rollen ausgetauscht haben und die korrespondierenden Gewichte die reziproken Werte zueinander sind.

Es kann sowohl förderlich wie lehrreich sein, aus verschiedenen Blickrichtungen die Aufmerksamkeit von Lernenden auf solche Paare von symmetrisch gewichteten Mitteln zu lenken. Dies kann ihre konzeptionellen Vorstellungen von einem Mittel erweitern und gleichzeitig die Fixierung auf das arithmetische Mittel und die habituelle Zuweisung von gleichen Gewichten zu allen Summanden unterminieren. Der Vorgang des Auswählens, wie die unterschiedlichen Werte bei einer Mittelung zu gewichten sind, erfordert eine sorgfältige Analyse der zugrunde liegende Fragestellung und einer Prüfung, wie die Daten erhoben wurden. Dies möge dann zu einer Routine führen, die die Methoden der Erhebung (Fisher 1934) als einen ersten Schritt beim Lösen statistischer Probleme in den Blick nimmt.

Danksagung

Wir danken Raphael Falk für das Interesse an unserer Fragestellung und für die Mithilfe in allen Phasen dieser Arbeit. Jean-Francois Caullier löste unser erneutes Interesse an gewichteten Mitteln durch seine Fragen aus.

Literatur

Beckenbach, E.; Bellman, R. (1961): An Introduction to Inequalities. Washington, DC: The Mathematical Association of America.

Ercolano, J. L. (1973): Remarks on the neglected mean. Mathematics Teacher, 66 (3), 253–55.

Eves, H. (2003): Means appearing in geometric figures. *Mathematics Magazine*, 76(4), 292– 94.

Falk, R.; Lann, A. (2008): The allure of equality: Uniformity in probabilistic and statistical judgment. *Cognitive Psychology*, 57(4), 293– 334.

Falk, R.; Lann, A.; Zamir, S. (2005): Average speed bumps: Four perspectives on averaging speeds. *Chance*, 18(1), 25–32.

Fisher, R. A. (1934): The effect of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6, 13–25.

Hoehn, L. (1984): A geometrical interpretation of the weighted mean. *College Mathematics Journal*, 15(2), 135–39.

Hoehn, L.; Niven, I. (1985): Averages on the move. *Mathematics Magazine*, 58(3), 151–56.

Lann, A.; Falk, R. (2002): An average with unimaginative weights: When the weights equal the values. In: A. D. Cockburn and E. Nardi (eds.) Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education, Vol. 1, p.288. Norwich: University of East Anglia.

Lann, A.; Falk, R. (2005): A closer look at a relatively neglected mean. *Teaching Statistics*, 27(3), 76–80.

Lann, A.; Falk, R. (2006). Tell me the method, I'll give you the mean. *The American Statistician*, 60(4), 322–27.

Nelsen, R. B. (1993). Proofs Without Words: Exercises in Visual Thinking. Washington, DC: The Mathematical Association of America.

Anschrift der Verfasser

Ruma Falk and Avital Lavie Lann
Hebrew University
Jerusalem
Israel
rfalk@cc.huji.ac.il