AUF WIE VIELE STELLEN BESTIMMEN WIR DEN MITTELWERT?

von JAN BIRNBAUM
übersetzt von Karl Schick

... und deshalb ist der Mittelwert 73 : 192 = 0,380 208 3. Wie oft haben wir so etwas gesehen? Der Taschenrechner verführt praktisch dazu, solche Antworten mit einer oft ganz falschen Genauigkeit zu geben. Wenn die Meßwerte, von denen der Mittelwert berechnet wurde, auf 7 Dezimalstellen genau wäre, sowäre das Ergebnis nicht zu beanstanden; aber das ist nicht sehr wahrscheinlich.

Die Frage erhebt sich jedoch: Auf wie viele Dezimalstellen soll der Mittelwert angegeben werden? Um das zu beantworten, wollen wir ein typisches Beispiel betrachten.

Tabelle 1. Dicke von Schuhleder (in mm): Häufigkeitstabelle

Wert	1,25	1,30	1,32	1,34	1,35	1,37	1,38	1,40	1,41
Häufigkeit	3	3	4	2	5	3	3	5	6
Wert	1,4[illegible]	1,43	1,44	1,45	1,46	1,47	1,48	1,49	1,50
Häufigkeit	10	8	6	6	6	6	7	8	7
Wert	1,5[illegible]	1,52	1,53	1,54	1,55	1,56	1,57	1,58	1,59
Häufigkeit	6	9	8	10	8	9	4	9	8
Wert	1,6[illegible]	1,61	1,62	1,63	1,64	1,65	1,66	1,67	1,68
Häufigkeit	7	6	6	9	6	5	5	6	5
Wert	1,6[illegible]	1,70	1,72	1,73	1,75	1,78	1,83		
Häufigkeit	4	3	2	2	2	2	1		

Tabelle 1 gibt die Dicke, gemessen in mm, kleinerer Proben von Schuhleder an, nachdem das Leder gespannt worden ist. Diese Meßwerte können auf zwei Dezimalstellen genau angegeben werden. Auf wie viele Stellen ist der Mittelwert genau?

Zunächst müssen wir genau wissen, was auf "zwei Dezimalstellen genau" bedeutet. Jeder der drei Werte, die mit 1,25 angegeben sind, liegt irgendwo zwischen 1,245 und 1,255; jeder der drei Werte mit 1,30 liegt irgendwo zwischen 1,295 und 1,305 usw. Allgemein kann man sagen, daß jeder einzelne Meßwert, der mit x bezeichnet wird und auf zwei Dezimalstellen genau angegeben ist, wirklich zwischen x - 0,005 und x + 0,005 liegt.

Nun können wir zwei Grenzfälle betrachten: Den ersten erhalten wir, wenn jeder Wert, der mit x gemessen worden ist, in Wirklichkeit den Wert x - 0,005 hat. In diesem Fall ist der wahre Mittelwert der Meßwerte

$$\frac{1}{n}\sum(x - 0{,}005) = \frac{1}{n}\sum x - \frac{1}{n}\sum 0{,}005 = \bar{x} - 0{,}005.$$

(In dieser Rechnung bedeutet $\sum$ die Summe von allen Meßwerten, n ist die Anzahl der Meßwerte und $\bar{x}$ ist der Mittelwert, der sich aus Tabelle 1 ergibt). Den zweiten Grenzfall erhalten wird, wenn jeder Wert, der mit x gemessen wurde, in Wirklichkeit den Wert x + 0,005 hat. In diesem Fall ist der wahre Mittelwert der Meßwerte $\bar{x}$ + 0,005.

Es ist klar, daß der wahre Mittelwert irgendwo zwischen diesen beiden Grenzfällen liegt, also zwischen $\bar{x}$ - 0,005 und $\bar{x}$ + 0,005. Danach scheint der berechnete Mittelwert auf zwei Stellen genau zu sein.

Wir wollen jedoch in dieses Problem noch etwas tiefer eindringen. Die Grenzfälle $\bar{x}$ - 0,005 und x + 0,005 sind sicherlich unwahrscheinlicher als andere Werte. Zu erwarten ist, daß einige Meßwerte größer und andere kleiner als die angegebenen Werte sind. Daher werden sich, wenn wir sie alle addieren, um den Mittelwert zu berechnen, die positiven und negativen Abweichungen teilweise ausgleichen und das Ergebnis wird näher bei dem berechneten Mittelwert als bei dem Wert x $\pm$ 0,005

liegen. Um das jedoch quantitativ zu analysieren, müssen wir ein Wahrscheinlichkeitsmodell für diese Situation angeben und verwenden.

Wir wollen annehmen, daßein Meßwert, der mit x bezeichnet wird, in Wahrheit mit gleicher Wahrscheinlichkeit zwischen x - 0,005 und x + 0,005 liegt. Das scheint in diesem Fall nicht eine unvernünftige Annahme zu sein, obwohl die zugrundeliegende Verteilung für die Daten wahrscheinlich nicht eine Gleichverteilung ist. Der Bereich der Werte erstreckt sich über 0,60 und dies ist viel mehr als die Spannweite der durch einen einzelnen Wert abgedeckt wird (0,01). Wenn diese Annahme jedoch unvernünftig sein sollte, dann ist anzunehmen, daß der berechnete Mittelwert sich in Wirklichkeit als genauer und nicht als weniger genau erweisen wird als derjenige, zu dem wir gelangen. Darüber wird mehr am Ende des Artikels gesagt werden.

Nun haben wir in Tabelle 1 insgesamt 240 Abweichungen von den zugrundeliegenden wahren Werten, wobei jede Abweichung zwischen - 0,005 und + 0,005 liegt. Nach unserer Voraussetzung folgt, daß wir ein Zufallsergebnis von 240 Werten einer gleichmäßigen stetigen Verteilung haben zwischen - 0,005 und + 0,005. Der berechnete Mittelwert x ist der wahre Mittelwert der Meßwerte vermehrt um den Mittelwert dieser 240 Abweichungen. Leider kennen wir jedoch den letzten Mittelwert nicht; aber nach unserer angenommenen Wahrscheinlichkeitsverteilung können wir ein Intervall angeben, das ihn wahrscheinlich einschließt. Das machen wir folgendermaßen.

Wir wissen, daß für eine genügend große Stichprobe der Größe n einer beliebigen Verteilung mit dem Mittelwert μ und der Standardabweichung σ der Mittelwert dieser Verteilung angenähert normal verteilt ist und zwar mit dem Mittelwert μ und der Standardabweichung

$\sigma/\sqrt{n}$. Das trifft sicherlich in unserem Fall zu, bei dem wir eine stetige Gleichverteilung haben, aus der eine Stichprobe im Umfang 240 entnommen wurde. Daher haben wir, wenn wir den Mittelwert und die Standardabweichung der stetigen Gleichverteilung herausfinden können, mit guter Annäherung die Verteilung des Mittelwertes der Abweichungen.

Nun ist für eine stetige Gleichverteilung mit einer Spannweite b - a der Mittelwert (a + b)/2 und die Standardabweichung $(b - a)/\sqrt{12}$. In unserem Fall ist a = - 0,005 und b = 0,005 und daher ist der Mittelwert 0 und die Standardabweichung $0{,}01/\sqrt{2880}$.

In 99 % aller Fälle liegt eine standarsierte Zufallsvariable (z.B. eine solche mit dem Mittelwert 0 und der Standardabweichung 1) zwischen -2,58 und + 2,58. Daher liegt für unser Beispiel in 99 % aller Fälle der Mittelwert der Stichprobe zwischen $-2{,}58\cdot 0{,}1/\sqrt{2880}$ und $+2{,}58\cdot 0{,}1/\sqrt{2880}$, also zwischen $\pm$ 0,000481.

Daher können wir schließen, daß mit einer hohen Wahrscheinlichkeit der berechnete Mittelwert innerhalb von + 0,005 des wahren Mittelwerts liegt. Nun ist man versucht zu schließen, daß der berechnete Mittelwert auf drei Dezimalstellen genau ist. Das ist leider jedoch nicht so, wie wir aus den Meßwerten der Tabelle 1 erkennen können. Der berechnete Mittelwert ist 367,16/240 = 1,52983 und daher liegt nach unserer Überlegung der wahre Mittelwert zwischen 1,5293 und 1,5303. Der erste Wert ist 1,529 bei drei Dezimalstellen und der zweite 1,530. Trotzdem ist das Ergebnis deutlich besser als die ursprüngliche Genauigkeit von zwei Dezimalstellen. Am besten gibt man das Ergebnis als Intervall an: So liegt für Tabelle 1 der Mittelwert im Intervall [1,5293; 1,5303]. Es wird selten möglich sein, dieses Intervall durch eine Dezimalzahl

mit einer vorgegebenen Stellenzahl zusammenzufassen, weil, wie in unserem Beispiel, gewöhnlich eine gewisse Zweideutigkeit vorliegt.

Wir können unsere Argumentation folgendermaßen verallgemeinern. Es soll angenommen werden, daß ein Meßwert auf r Dezimalstellen genau ist und der berechnete Mittelwert $\bar{x}$ ist. Dann wird der wahre Mittelwert im Intervall $[\bar{x} - 5\cdot 10^{-(r+2)}, \bar{x} + 5\cdot 10^{-(r+2)}]$ liegen, wenn der Stichprobenumfang 200 oder größer ist; ist der Stichprobenumfang 20 000 oder mehr, wird der wahre Mittelwert im Intervall $[\bar{x} - 5\cdot 10^{-(r+3)}, \bar{x} + 5\ 10^{-(r+3)}]$ liegen usw.

Theoretisch folgt daraus, daß wir mit genügend vielen Meßwerten den Mittelwert bis zu jeder benötigten Genauigkeit bestimmen können, unabhängig vom Grad der Genauigkeit der ursprünglichen Meßwerte (wobei selbstverständlich eine richtige Berechnung vorausgesetzt wird). In der Praxis würden wir jedoch eine unzumutbare Vergrößerung unserer Stichprobe benötigen, selbst dann, wenn nur eine relativ kleine Verbesserung der Genauigkeit erreicht werden soll.

Abschließend sollte ich darauf hinweisen, daß die Ergebnisse, die ich hier angegeben habe, wahrscheinlich etwas zurückhaltend sind. Im allgemeinen wird die Annahme einer stetigen Gleichverteilung für die Abweichungen eine größere Standardabweichung liefern als diejenige, die man tatsächlich erhält; denn die Verteilung ist wahrscheinlich schief nach rechts oder nach links vom Null-Abweichungs-Punkt aus. Aus diesem Grund möchte ich anregen, daß eine minimale Stichprobengröße von 50 eher als eine solche von 200 eine vernünftige Faustregel ist.

Originaltitel in 'Teaching Statistics', Vol. 5 (1983) Nr. 3
To How Many Decimal Places Do We Mean?