Mathematical Properties of the Average

Kalid Azad
http://www.princeton.edu/~kazad/

October 31, 2002

Abstract

The average is a simple but powerful concept. This article explains some of its interesting, non-obvious properties (such as being the center-of-gravity).[1]

Contents

1 Introduction and definition

Definition 1.1 *The average is defined as*

$$Average = \frac{1}{n}\sum a_i$$

A simple interpretation: each nummber(a_i) gets to 'vote' for itself; the average is the sum of these 'votes'. Each number gets $\frac{1}{n}$th the total vote, hence the $\frac{1}{n}$. This democratic decision of the average is nice, fair, and also gives us some cool properties.

2 Distance to the average

First, consider how far each number is from the average. Suppose we have three numbers a, b and c, with an average of $\frac{a+b+c}{3}$.

How far is each number from the average? Well, each number (a, b, c) is simply a point on the number line and the distance to the average is a simple subtraction. So, the distance from a to the average is

$a - \frac{a+b+c}{3} = \frac{2a-b-c}{3}$

Likewise, the distance of b and c to the average is $\frac{2b-a-c}{3}$ and $\frac{2c-a-c}{3}$, respectively. These numbers, by themselves, aren't that useful.

However, let's ask a more important question: what is the *total* distance of a, b and c to the average? The answer:

$\frac{2a-b-c}{3} + \frac{2b-a-c}{3} + \frac{2c-a-b}{3} = 0$

Now this is interesting. The sum of the distances to the average (from each point) is zero. Note that we are not talking about the absolute value of distance (such as 'point a is 3 units away from the average'). We are using positive and negative distances (our convetion was $a_i - average$). A positive distance means a point is greater than the average, and a negative distance indacates the point is less than the average.

This example was just for 3 numbers. Of course, it works in general as well. Take n numbers (a_1 to a_n) and call the average a_x $(= \frac{1}{n}\sum^N a_i)$. The sum of the distances to the average is

$$\begin{aligned}\sum^N (a_i - a_x) &= \sum^N a_i - \sum^N a_x \\ &= (n * a_x) - (a_x * n) \\ &= 0\end{aligned}$$

Not a hard proof, eh?

3 Interpretation

This result means that the average is 'midway' between the numbers in an interesting way. Imagine you are sitting at the average, with the smaller numbers to your left and the larger numbers to your right (recall that the smaller numbers have negative distance, and the larger ones have positive distance).

The facts

- Distance from average to smaller numbers is negative
- Distance to larger numbers is positive
- Total distance (from average to larger and smaller numbers) is 0

Conclusion: The sum of distances from the smaller numbers = sum of distance to larger numbers. If they were unequal, the total distance would not be zero. Since the total distance *is* zero, then they must be equal. Another way of putting it: if you have a bunch of good and bad effects, and the net result is zero, then the sum of the good effects = sum of bad effects.

(If that isn't satisfying (and I hope it is!), here is a math explanation.
[2])

Fact to take away: sum of distance above average = sum of distance below average

4 Consequences

Here are a few ways we can use this fact:

- **Changing numbers but keeping the average the same.** If we keep the total distance to the average the same, the average stays the same (if we reduce the smallest number by some

[2] I had hoped the intuitive argument had worked, but I guess not. Let S be the subset of numbers smaller than the average, and L be the subset larger. Let E be the subset of all numbers equal to the average. E, S and L are mutually exclusive (a number cannot be both greater, less than, or equal to the average).

Call the average a_x, and let S_i be an element in S (similarly for L and E). The facts are

- $\sum (S_i - a_x) < 0$, true because each $S_i < a_x$ by definition
- $\sum (L_i - a_x) > 0$, true for similar reasons
- $\sum (E_i - a_x) = 0$, true because each $E_i = a_x$ by definition
- $\sum a_i = E + L + S$, since the sets are mutually exclusive and each number a_i must fall into exactly one category
- $\sum a_i - a_x = 0$, shown earlier

Therefore,

$$\begin{aligned}\sum a_i - \sum a_x &= E + L + S - \sum a_x \\ 0 &= \sum (S_i - a_x) + \sum (L_i - a_x) + \sum (E_i - a_x) \\ 0 &= \sum (S_i - a_x) + \sum (L_i - a_x) \\ -\sum (S_i - a_x) &= \sum (L_i - a_x)\end{aligned}$$

which is the result we sought to prove. Phew! Sorry for the roundabout proof, but you asked for it!

amount, and increase the largest number by the same amount, the average stays the same). You've probably had an intuitive feeling for this in a while: the average of $(9, 10, 11)$ = average of $(7, 10, 13)$. We pull the lowest number down 2 and bring the highest one up 2 to compensate. Also, it shows how you can 'redistriubte' a change in distance. Suppose we have 4 points, and the greatest one goes up by 10. How can we maintain the old average? Well, we can reduce a smaller element by 10, or reduce two smaller elements by a total of 10. Again, as long as the *total* distance stays the same (contribution from all points), we are ok. Example: average of $(4, 6, 8, 10)$ = average of $(4 - 2, 6 - 1, 8, 10 + 3)$. The $+3$ cancels the -2 and -1.

- **Center of gravity!** Suppose we have numerous 1-pound at different positions along a board, and we want to balance the board. At what point can we balance it? In physics terms, each weight applying a torque to the board, equal to weight * distance. Since the weights are the same (1-pound), the torque is simply equal to distance.

  We want the location where all the torques, positive and negative, cancel out, so the board does not spin (it balances!). The answer, my friend, is the average! Remember, the average is the point where the sum of distances (torques) above the average = sum of distances (torques) below! So, the board will balance at the center of gravity: hoo-ha! I had always wanted a proof for this (I suspect you have as well), and now you have one. Discussion of different weights to follow :).

5 Weighted Average

Suppose each element a_i has a corresponding weight w_i. This 'weight' is essentially how important the particular number is. A number with no importance has a weight of 0. To give one number twice the importance of another, give it twice the weight.

The typical example: suppose you took a test (got 90) and a quiz (got an 80). Tests are twice as important as quizes. How would you take the average? One way to do it is to count the test as if we took it twice, and got 90 both times. After all, it is twice as important, and gets twice as many 'votes'. So,

$$average = \frac{90 + 90 + 80}{3}$$

However, let's use the weight method. Give the test a weight of 2, and the quiz a weight of 1. Thus, there are 3 total votes (This is important: the denominator is not the number of elements: it is the sum of all the weights, or votes. Previously, each element got one vote, so denominator = number of votes = number of elements).

$$average = \frac{2 * 90 + 80}{3}$$

In general, the formula is

$$WeightedAverage = \frac{1}{\sum w_i} * \sum a_i w_i$$

This looks worse than it is. Instead of dividing by n, we divide by the sum of the weights. Instead of adding up a_i directly, we multiply each a_i by w_i, then add them up.

Hey, wouldn't it be cool if the sum of the weights was 1? Then we wouldn't have to do any division at the end. This is a commonly used tactic: make each w_i represent the percentage of the total weight. The sum of all the percentages is 1, so no division is required. This is often used in academic settings (i.e., tests count 50%, quizes 30%, and homework 20%).

Another use is when each element is different for some reason or another. In the center of gravity example, each weight did not have to be 1-pound. Instead, have the distance of each weight be a_i and the mass be w_i. The weighted average is the center of gravity; this is how the calculations are actually done (I will put up an article about this someday).

6 Conclusions

The average is a simple concept, but has interesting properties

- Distance above average = distance below. Changes that maintain this relationship keep the old average.
- Weighted averages can reflect differences in 'importance' of elements.
- The average explains why the center-of-gravity is a balance point.