Finding weights to result in a given average

Wagner L. Truppel

April 17, 2015

The problem

Suppose we're given $n > 1$ positive real numbers a_i and a positive real number N such that $N < a_i$, for $1 \leq i \leq n$. The problem is to find n numbers p_i, $0 \leq p_i \leq 1$, such that N is the average of the a_i's with weights given by the p_i's, that is, such that

$$\sum_{i=1}^{n} p_i \, a_i = N \qquad \text{and} \qquad \sum_{i=1}^{n} p_i = 1 \, .$$

Obviously, as stated, the problem admits an infinitely large number of solutions so, instead, let's look for a solution that minimizes the sum of the squares of the p_i's. As it turns out, that solution is unique.

The solution

The solution is most easily obtained by employing the technique of *Lagrange multipliers*. Suppose we construct the function

$$S \equiv S(p_1, p_2, \ldots, p_n \,|\, \lambda_1, \lambda_2) = \sum_{i=1}^{n} p_i^2 - \lambda_1 \, (\sum_{i=1}^{n} p_i - 1) - \lambda_2 \, (\sum_{i=1}^{n} p_i a_i - N)$$

where λ_1 and λ_2 are constants to be determined. Note that if the p_i's satisfy the two conditions of the problem, then S reduces to the sum of the squares of the p_i's.

Now suppose we want to minimize S with respect to all possible choices of the values of the p_i's. A necessary condition is that all the first-order partial derivatives vanish at the location of the extremum or extrema:

$$\frac{\partial S}{\partial p_j} = 2p_j - \lambda_1 - \lambda_2 \, a_j = 0 \, , \qquad (1 \leq j \leq n) \, .$$

These immediately give us the values of the p_i's,

$$p_i = \frac{1}{2}\,(\lambda_1 + \lambda_2\, a_i)\,,$$

in terms of λ_1 and λ_2. In order to find these two constants, we only need to apply the constraint conditions. Summing $2p_j - \lambda_1 - \lambda_2\, a_j = 0$ and $2p_j\, a_j - \lambda_1\, a_j - \lambda_2\, a_j^2 = 0$ over j and making use of the constraints gives us

$$\begin{aligned} n\,\lambda_1 + (\sum_{j=1}^{n} a_j)\,\lambda_2 &= 2 \quad \text{and} \\ (\sum_{j=1}^{n} a_j)\,\lambda_1 + (\sum_{j=1}^{n} a_j^2)\,\lambda_2 &= 2N \end{aligned}$$

which can be written in matrix form as follows

$$\begin{pmatrix} n & \sum_{j=1}^{n} a_j \\ \sum_{j=1}^{n} a_j & \sum_{j=1}^{n} a_j^2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ N \end{pmatrix}.$$

This system of linear equations on λ_1 and λ_2 can be solved using Cramer's rule, resulting in

$$\lambda_1 = 2\,\frac{\sum_{j=1}^{n} a_j^2 - N \sum_{j=1}^{n} a_j}{n \sum_{j=1}^{n} a_j^2 - (\sum_{j=1}^{n} a_j)^2} \qquad \text{and} \qquad \lambda_2 = 2\,\frac{nN - \sum_{j=1}^{n} a_j}{n \sum_{j=1}^{n} a_j^2 - (\sum_{j=1}^{n} a_j)^2}\,.$$

Thus,

$$\boxed{p_i = \frac{\sum_{j=1}^{n} a_j^2 - (N + a_i) \sum_{j=1}^{n} a_j + nNa_i}{n \sum_{j=1}^{n} a_j^2 - (\sum_{j=1}^{n} a_j)^2} \qquad (1 \le i \le n)}$$

minimizes $S = \sum_{i=1}^{n} p_i^2$ subject to the constraints of the original problem. That the value is in fact a minimum, rather than a maximum or inflection point, can be seen from the fact that the second-order partial derivatives of S with respect to the p_i's are positive when the p_i values are positive.

■