Lagrange's Identity and Congressional Apportionment

Tommy Wright

Abstract. This note makes use of Lagrange's Identity to provide a bridge between an insightful motivation and an elementary derivation of the *method of equal proportions*. The method of equal proportions is the current method for apportioning the 435 seats in the U.S. House of Representatives among the 50 states, following each decennial census.

1. INTRODUCTION. Following each decennial census of the United States, the *method of equal proportions* has been used since the 1940 Census to reallocate or reapportion the 435 seats in the U.S. House of Representatives among the states, according to their latest population counts. For a scholarly, technical, and historical presentation of previous methods that have been used, see [**3**]. Surprisingly, the computational details of the method of equal proportions are unknown to many, though the details are quite simple. Moreover, the rationale for the method of equal proportions seems a bit mysterious. This note makes use of Lagrange's Identity to provide a bridge between an insightful motivation and an elementary derivation of the method of equal proportions.

Based on early work by [**6**], the method of equal proportions (Working Rule) is due to Huntington [**7, 8**]. Huntington's proof of optimality [**8**] is indirect and lacks motivation for many. In collaboration with Huntington, Owens [**9**] takes a statistical approach, using weighted least squares as motivation for deriving the method of equal proportions (also see [**12**]). Wright and Cobb [**11**] provide an elementary exposition. While others have sought to highlight and compensate for the limitations of the method of equal proportions (e.g., [**1, 4**], and [**5**]), this is not the objective of this note. We use Lagrange's Identity to further clarify why the method of equal proportions helps to meet the following requirement as expressed in Article 1, Section 2, Clause 3 of the United States Constitution: "Representatives ... shall be apportioned among the several states ... according to their respective numbers ... The number of Representatives shall not exceed one for every thirty thousand, but each state shall have at least one Representative ..."

2. DERIVATION USING LAGRANGE'S IDENTITY.

2.1. Notation and the Method of Equal Proportions. Following the presentation in [**12**], let H_i be the number of seats to be assigned to state i with population P_i for $i = 1, 2, \ldots, 50$. Also let $P = \sum_i P_i$ be the total population of the states and $H = \sum_i H_i$ be the total number of seats in the U.S. House of Representatives. For all states i and j where $i \neq j$, "the state value P_i/H_i should be as nearly equal as possible to the state value P_j/H_j" was taken by Huntington [**7**] as a fundamental principle. Huntington's definition of "inequality" was *relative difference*, i.e.,

$$\frac{\frac{P_i}{H_i} - \frac{P_j}{H_j}}{\min\left(\frac{P_i}{H_i}, \frac{P_j}{H_j}\right)}. \tag{1}$$

http://dx.doi.org/10.4169/amer.math.monthly.121.06.523
MSC: Primary 90C10

His precise formulation of the problem is as follows.

Postulate I. *For two states, i and j, the relative difference between P_i/H_i and P_j/H_j should be as small as possible.*

Postulate II. *In a satisfactory apportionment, there should be no pair of states capable of being "improved" by a transfer of representatives within that pair—the word "improvement" being understood in the sense implied by Postulate I.*

These postulates lead to the following theorem.

Theorem. *For any given values of $P_1, P_2, P_3, \ldots, P_{50}$ and H, there will always be one and only one "satisfactory" apportionment in the sense defined by Postulates I and II.*

Formally and in practice, the method of equal proportions (or *Huntington's Method*) for satisfactory apportionment can be stated in terms of the following 3-step working rule.

Working Rule (*Method of Equal Proportions*)

Step 1: First, assign one seat (or Representative) to each state.

Step 2: Compute the array of *priority values*, where each row corresponds to one of the states (for simplicity, we assume that the states are ordered so that $P_1 \geq P_2 \geq P_3 \geq \cdots \geq P_{50}$):

$$\begin{array}{cccc} \frac{P_1}{\sqrt{1\cdot 2}} & \frac{P_1}{\sqrt{2\cdot 3}} & \frac{P_1}{\sqrt{3\cdot 4}} & \cdots, \\ & \vdots & & \\ \frac{P_i}{\sqrt{1\cdot 2}} & \frac{P_i}{\sqrt{2\cdot 3}} & \frac{P_i}{\sqrt{3\cdot 4}} & \cdots, \\ & \vdots & & \\ \frac{P_{50}}{\sqrt{1\cdot 2}} & \frac{P_{50}}{\sqrt{2\cdot 3}} & \frac{P_{50}}{\sqrt{3\cdot 4}} & \cdots. \end{array}$$

Step 3: Pick the $H - 50$ largest priority values from the array in Step 2, along with the associated states. Each state gets an additional seat each time one of its priority values is among the $H - 50$ largest values.

The largest priority values will be those in which the numerator is large (states with larger populations) and the denominator is small (that is, $\sqrt{1\cdot 2}, \sqrt{2\cdot 3}, \sqrt{3\cdot 4}, \ldots$).

In practice, the computation of priority values for any state may stop with the last result greater than $\frac{P}{H'}$, where P is the total population of all the states, and H' is a number larger than the largest value of H likely to be required [**7**].

2.2. An Insightful Motivation. Motivated by Huntington's formulation of the problem (Postulates I and II), for fixed P_i and P_j and the constraint $\sum_{i=1}^{50} H_i = H$, we want to determine H_i so that

$$\frac{P_i}{H_i} = \frac{P_j}{H_j} \tag{2}$$

for all i and j, such that $i < j$.

For state i, the ratio $\frac{P_i}{H_i}$ is the number of persons represented by a Representative; and (2) expresses the desire to have the number of persons represented by a Representative from state i be equal to the number of persons represented by a Representative from state j for all i and j for $i < j$.

Equivalently, for all i and j such that $i < j$, (2) can be written as

$$\frac{P_i}{\sqrt{H_i}}\sqrt{H_j} - \frac{P_j}{\sqrt{H_j}}\sqrt{H_i} = 0. \tag{3}$$

For the given values of P_i and the constraint that $\sum_{i=1}^{50} H_i = H$, equality is rarely possible, so we seek to determine H_i so that the differences on the left-hand side of (3) are as near to 0 as possible. Hence, mathematical ease leads us to determine H_i subject to the constraint $\sum_{i=1}^{50} H_i = H$ that minimizes the sum of squares

$$\sum_{1 \le i < j \le 50} \left(\frac{P_i}{\sqrt{H_i}}\sqrt{H_j} - \frac{P_j}{\sqrt{H_j}}\sqrt{H_i} \right)^2. \tag{4}$$

2.3. Lagrange's Identity. For any real numbers $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$, Lagrange's Identity (see, for example, [**2**, p. 27] and [**10**])

$$\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) - \left(\sum_{i=1}^{n} a_i b_i\right)^2 = \sum_{1 \le i < j \le n} \left(a_i b_j - a_j b_i\right)^2 \tag{5}$$

follows easily from mathematical induction.

Taking $a_i = \frac{P_i}{\sqrt{H_i}}$, $b_i = \sqrt{H_i}$, and $n = 50$, (5) becomes

$$\left(\sum_{i=1}^{50} \frac{P_i^2}{H_i}\right)\left(\sum_{i=1}^{50} H_i\right) - \left(\sum_{i=1}^{50} P_i\right)^2 = \sum_{1 \le i < j \le 50} \left(\frac{P_i}{\sqrt{H_i}}\sqrt{H_j} - \frac{P_j}{\sqrt{H_j}}\sqrt{H_i} \right)^2. \tag{6}$$

So minimizing (4) is now equivalent to minimizing the left-hand side of (6). But because $(\sum_{i=1}^{50} P_i)^2$ is fixed following a specific census and $\sum_{i=1}^{50} H_i = H$ where H is fixed, minimizing the left-hand side of (6) is equivalent to minimizing

$$\sum_{i=1}^{50} \frac{P_i^2}{H_i}. \tag{7}$$

From (7), we can obtain the Working Rule following the derivation in Section 2 of [**12**], which is highlighted in the remainder of this note for convenience to the reader.

Note that

$$\frac{1}{H_i} = 1 - \frac{1}{1 \cdot 2} - \frac{1}{2 \cdot 3} - \frac{1}{3 \cdot 4} - \cdots - \frac{1}{(H_i - 1) \cdot (H_i)}, \tag{8}$$

and the result in (7) becomes

$$\sum_{i=1}^{50} P_i^2 \left(\frac{1}{H_i}\right) = \sum_{i=1}^{50} P_i^2 - \sum_{i=1}^{50} \left(\frac{P_i^2}{1 \cdot 2} + \frac{P_i^2}{2 \cdot 3} + \cdots + \frac{P_i^2}{(H_i - 1)\,(H_i)}\right).$$

Because $\sum_{i=1}^{50} P_i^2$ is fixed, we determine $H_1, H_2, \ldots, H_{50}$ subject to $\sum_{i=1}^{50} H_i = H$ to maximize

$$\sum_{i=1}^{50} \left(\frac{P_i^2}{1 \cdot 2} + \frac{P_i^2}{2 \cdot 3} + \cdots + \frac{P_i^2}{(H_i - 1)\,(H_i)}\right). \tag{9}$$

Now (9) is maximized by picking the largest terms in the overall sum, subject to the constraint.

Because H is fixed and each state must have at least one Representative, we see that (9) will be maximized, our contraint will be satisfied, and each state will have at least one Representative if

(i) first each state is assigned one seat, and

(ii) then each state receives an additional seat each time it has a term in (9) that appears among the $H - 50$ largest terms in (9).

Because P_i and H_i are all positive, selecting the $H - 50$ largest terms among

$$\frac{P_i^2}{1 \cdot 2}, \frac{P_i^2}{2 \cdot 3}, \ldots, \frac{P_i^2}{(H_i - 1)(H_i)}$$

for $i = 1, 2, \ldots, 50$ is the same as selecting the $H - 50$ largest terms among

$$\frac{P_i}{\sqrt{1 \cdot 2}}, \frac{P_i}{\sqrt{2 \cdot 3}}, \ldots, \frac{P_i}{\sqrt{(H_i - 1)(H_i)}} \tag{10}$$

for $i = 1, 2, \ldots, 50$.

Note that for (8) and (9) to be valid, we must have $H_i \geq 2$. Because each state gets at least one seat, we will only need to consider states to see if they get additional seats, and in these cases $H_i \geq 2$. Thus, we have derived the *method of equal proportions*.

3. CONCLUDING REMARK. As was noted by one of the referees, the appropriate criterion to use for selection of a reasonable method of apportionment is a topic of considerable mathematical interest and controversy. Seemingly equivalent criteria can lead to quite different methods. For example, exchanging P and H in equation (2) leads to an equally reasonable desire to have

$$\frac{H_i}{P_i} = \frac{H_j}{P_j} \tag{2'}$$

for all $i < j$.

Following through with the analogous details of (3), (4), (5), and (6) leads to choosing H_i to minimize

$$\sum_{i=1}^{50} \frac{H_i^2}{P_i} \tag{7'}$$

subject to $\sum_i H_i = H$. However, as is shown in [**9**, p. 963–964] and noted in [**7**, p. 868], this leads to the *method of major fractions* advocated by Willcox, whose derivation is given as follows from (7′). For any positive integer H_i, we have

$$H_i^2 = 1 + 3 + 5 + \cdots + (2H_i - 1).$$

Hence, (7′) becomes

$$\begin{aligned} \sum_{i=1}^{50} \frac{H_i^2}{P_i} &= \sum_{i=1}^{50} \left(\frac{1}{P_i} + \frac{3}{P_i} + \frac{5}{P_i} + \cdots + \frac{2H_i - 1}{P_i} \right) \\ &= 2 \sum_{i=1}^{50} \left(\frac{.5}{P_i} + \frac{1.5}{P_i} + \frac{2.5}{P_i} + \cdots + \frac{H_i - .5}{P_i} \right). \end{aligned}$$

Thus, from an array as follows (where for convenience and without loss of generality we continue to assume $P_1 \geq P_2 \geq \cdots \geq P_{50}$),

$$\begin{array}{cccc} \frac{.5}{P_1} & \frac{1.5}{P_1} & \frac{2.5}{P_1} & \cdots \\ \frac{.5}{P_2} & \frac{1.5}{P_2} & \frac{2.5}{P_2} & \cdots \\ & \vdots & & \\ \frac{.5}{P_{50}} & \frac{1.5}{P_{50}} & \frac{2.5}{P_{50}} & \cdots \end{array}$$

we pick the entry in the first position of each row, because each state gets one seat initially, and then we pick the $H - 50 = 385$ smallest entries in columns 2, 3, ... of this array and use as H_i the total number of entries selected from row i.

Alternatively, from the array below,

$$\begin{array}{cccc} \frac{P_1}{.5} & \frac{P_1}{1.5} & \frac{P_1}{2.5} & \cdots \\ \frac{P_2}{.5} & \frac{P_2}{1.5} & \frac{P_2}{2.5} & \cdots \\ & \vdots & & \\ \frac{P_{50}}{.5} & \frac{P_{50}}{1.5} & \frac{P_{50}}{2.5} & \cdots \end{array}$$

we pick the entry in the first position of each row, because each state gets one seat initially, and then we pick the $H - 50 = 385$ largest entries in columns 2, 3, ... of this array and use as H_i the total number of entries selected from row i.

Furthermore, [**3**, pp. 103–104] shows that *Webster's Method* minimizes the function in (7′) where Webster's method is given as follows.

Webster's Method. Find a divisor D_W such that

$$\sum_i \left[\frac{P_i}{D_W} \right] = H,$$

where $[x]$ is the nearest integer to x, and assign

$$H_i = \left[\frac{P_i}{D_W} \right]$$

seats to state i.

Disclaimer. This note is published to inform interested parties of research and to encourage discussion. The views expressed are those of the author and not necessarily those of the U.S. Bureau of the Census.

REFERENCES

1. R. A. Agnew, Optimal Congressional apportionment, *Amer. Math. Monthly* **115** (2008) 297–303.
2. T. M. Apostol, *Mathematical Analysis*. Second edition. Addison-Wesley, Reading, MA, 1973.
3. M. L. Balinski, H. P. Young, *Fair Representation: Meeting the Ideal of One Man, One Vote*. Second edition. Brookings Institution Press, Washington, DC, 2001.
4. M. Balinski, What is just? *Amer. Math. Monthly* **112** (2005) 502–511.
5. G. Grimmett, Stochastic apportionment, *Amer. Math. Monthly* **111** (2004) 299–307.
6. J. A. Hill, Letter addressed to William C. Houston, Chairman of the House Committee on Census, in House, *Apportionment of Representatives*, House Report 12, 62nd Congress, 1st Session (April 25, 1911).
7. E. V. Huntington, A new method of apportionment of representatives, *Q. Pub. Amer. Stat. Assoc.* **17** (1921) 859–870.
8. E. V. Huntington, The apportionment of representatives in Congress, *Trans. Amer. Math. Soc.* **30** (1928) 85–110.
9. F. W. Owens, On the apportionment of representatives, *Q. Pub. Amer. Stat. Assoc.* **17** (1921) 958–968.
10. T. Wright, Lagrange's identity reveals correlation coefficient and straight-line connection, *The American Statistician* **46** (1992) 106–107.
11. T. Wright, G. Cobb, Counting and Apportionment: Foundations of America's Democracy, in *Statistics: A Guide to the Unknown*. Fourth edition. Edited by R. Peck, G. Casella, G. Cobb, R. Hoerl, D. Nolan, R. Starbuck, and H. Stern. Duxbury Press, Pacific Grove, CA, 2005. 35–67.
12. T. Wright, The equivalence of Neyman optimum allocation for sampling and equal proportions for apportioning the U.S. House of Representatives, *The American Statistician* **66** no. 4 (2012) 217—224.

TOMMY WRIGHT received his Ph.D. in statistics from The Ohio State University in 1977. He is Chief of the Center for Statistical Research and Methodology, U.S. Bureau of the Census. He also serves as adjunct faculty in mathematics and statistics at Georgetown University.
U.S. Bureau of the Census, Washington, DC 20233
tommy.wright@census.gov