IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 18, Issue 2 Ser. IV (Mar. – Apr. 2022), PP 08-11
www.iosrjournals.org

Averaging of the searchable quantity by three its independent values

Ph.D. Ruslan Yu. Golikov [1]
[1] *(Department of Mathematics / Synergy University, Russia)*

Abstract:
In the course of solving scientific problems, it is not uncommon for the same searchable quantity to have different numerical values and the difference between them to be more than the margin of error in obtaining them. This creates uncertainty in the estimation of its true value. The smallest separate values number of the searchable quantity, at which a deterministic estimate is still possible, is three. The two values of the three that have the greatest difference are mutually exclusive. When the definition is independent, the third, intermediate value plays a clarifying role, indicating the proximity of the searchable quantity to one of the two mutually exclusive values. This paper considers the general solution of equations for the searchable quantity true value and its three separate values. The result is a way of averaging such data, taking into account the clarifying role of the intermediate value. The solution to the problem is presented in algebraic and geometric interpretations, leading to a single solution. As an example, the results of averaging the Hubble constant from its measurement by different research groups are given. Shown a marked discrepancy between the averaging result and classical arithmetic mean of searchable quantity. The averaging method discussed in the paper is applicable for approximate deterministic estimation of experimental and computational data under their true values uncertainty conditions.

Key Word: *deterministic estimate, uncertainty, systematic error, scatter, clarifying, general solution, particular solution, averaging, arithmetic mean, Hubble constant.*

Date of Submission: 05-04-2022 Date of Acceptance: 20-04-2022

I. Introduction

Scientific research in various fields of knowledge relies, in most cases, on direct measurement of physical quantities. Therefore, there is often a natural discrepancy between the individual results obtained and the need to average them as an approximate estimation of the searchable quantity true value. In this case we are discussing a quantity whose values cannot be determined by statistical methods due to the small sample size, i.e. its deterministic estimate, taking into account the contribution of each value obtained. The final result is the numerical value that is closest to the true value of the searchable quantity, taking into account the irreducible variation in its separate values. That is why, when there are separate values of the searchable quantity, each of them has a significant effect on its approximate estimation. In this paper we consider the case of the same quantity three different values, obtained by alternative methods (or different measuring instruments under the same conditions), which corresponds to the minimum acceptable set of initial data. The smaller number of values available makes it impossible to trace their mutual influence. Obviously, two values of the three are mutually exclusive and determine their maximum variation. The third, intermediate value, is clarifying regardless of the order in which it is received, as it is generally closer to one of the other two. Usually, in such cases, researchers apply averaging based on the arithmetic mean. However, this approach to averaging the searchable quantity devalues the clarifying role of its intermediate value, as it sums up equal fractions from each individual value.

II. Algebraic interpretation of the problem

Let the three independently obtained values of the searchable quantity E be E_1, E_2 and E_3. The order in which they are received is irrelevant, and the subscripts are arranged in ascending its numerical values. Each value is associated respectively with an unknown systematic (non-random) error δ_1, δ_2 and δ_3. Each obtained value of the searchable quantity E and the unknown errors of its determination are linked by a linear equations system

$$\begin{cases} E_1 + \delta_1 = E; \\ E_2 + \delta_2 = E; \\ E_3 + \delta_3 = E. \end{cases} \tag{1}$$

By virtue of the right-hand sides equality, let us write the system of equations (1) in the following form

$$\begin{cases} \delta_1 - \delta_2 = E_2 - E_1; \\ \delta_1 - \delta_3 = E_3 - E_1; \\ \delta_2 - \delta_3 = E_3 - E_2. \end{cases} \tag{2}$$

The rank of the expanded matrix of this system is two and since the number of unknowns is three, this system has an infinite number of solutions. Its general solution is

$$\begin{cases} \delta_1 - \delta_2 = E_2 - E_1; \\ \delta_2 - \delta_3 = E_3 - E_2. \end{cases} \tag{3}$$

This structure of the general solution shows that the unknown quantities are related in pairs: δ_1 and δ_2, δ_2 and δ_3. For ease of further analysis, let's write down the general solution (3) as

$$\delta_1;\ \delta_2 = \delta_1 - E_2 + E_1;\ \delta_3 = \delta_2 - E_3 + E_2 = \delta_1 - E_3 + E_1. \tag{4}$$

Since the first and second equations of system (1) are equal, and by problem condition $E_1 < E_2$, the equality can only be fulfilled if $\delta_2 = -\delta_1$:

$$E_1 + \delta_1 = E_2 - \delta_1. \tag{5}$$

From equation (5) it follows that

$$\delta_1 = \frac{E_2 - E_1}{2}. \tag{6}$$

Similarly, from the equality of the second and third equations of system (1) provided $E_2 < E_3$ we obtain

$$\delta_2 = \frac{E_3 - E_2}{2}. \tag{7}$$

Expressions (6) and (7) are two particular solutions of system (3). Each of these, given (4), satisfies the original equations system (1) and gives the same values of the searchable quantity E in it. Thus, the searchable quantity E value for the particular solution (6) will be

$$E = \frac{E_1 + E_2}{2}, \tag{8}$$

and is the arithmetic mean of E_1 and E_2.
For the particular solution (7), the value of the searchable quantity E is

$$E = \frac{E_2 + E_3}{2}, \tag{9}$$

i.e. is the arithmetic mean of E_2 and E_3.
Then the arithmetic mean of the searchable quantity E, taking into account (8) and (9), is

$$E' = \frac{E_1 + 2E_2 + E_3}{4}. \tag{10}$$

On expression (10) base it can be shown that for searchable quantity E arithmetic mean E_{am} of the its independent values (E_1, E_2 and E_3) is only a particular case when the E_2 value is equal to the E_1 and E_3 values arithmetic mean:

$$E' = \frac{E_1 + 2\left(\frac{E_1 + E_3}{2}\right) + E_3}{4} = \frac{E_1 + E_3}{2}; \tag{11}$$

$$E_{am} = \frac{E_1 + \frac{E_1 + E_3}{2} + E_3}{3} = \frac{E_1 + E_3}{2}.$$

III. Averaged estimate of the searchable quantity, using Hubble constant measurements as an example

A number of values for the Hubble constant H_0, determining in cosmology the average velocity of galaxies breaking away in the modern era, have now been obtained using different measurement methods. The value obtained by measuring of the cosmic microwave background radiation left over from the Big Bang using Planck Space Observatory as of 2018 is 67.4 kilometers per second per Megaparsec [(km/s)/Mpc][1]. The value obtained by calculating the distances to galaxies from the luminosity of the Cepheid observed in them using Hubble Space Telescope as of 2019 is 74.03 (km/s)/Mpc[1]. A researchers group led by Wendy L. Freedman from Chicago University has come up with an intermediate result: 69.8 (km/sec)/Mpc after studying the luminosity of the distant Red Giants and Supernovae[1]. The highest Hubble constant value that obtained by taking into account the surface brightness fluctuations of galaxies is 76.5 (km/s)/Mpc[1].

Thus, all these estimates of the Hubble constant are inconsistent with each other. There is as yet no consensus on the reasons for such discrepancies and this allows us to give an average estimate of the individual Hubble constant measurements as a possible guideline. The above estimates of the Hubble constant obtained by the different research groups, as well as the average estimates made using expression (10) for the different measurement datasets, are presented in Table no 1 and Table no 2.

Table no 1: Measured values of Hubble constant and their average estimate (option 1).

Data source	Measured values	Arithmetic mean (E_{am})	Averaging by expression (10) (E')
Measuring of the cosmic microwave background radiation using Planck Space Observatory (E_1)	67.40	71.23	70.88
Wendy Friedman's group studying the luminosity of the distant red giants and supernovae (E_2)	69.80		
Value that obtained by taking into account the surface brightness fluctuations of galaxies (E_3)	76.50		

Table no 2: Measured values of Hubble constant and their average estimate (option 2).

Data source	Measured values	Arithmetic mean (E_{am})	Averaging by expression (10) (E')
Measuring of the cosmic microwave background radiation using Planck Space Observatory (E_1)	67.40	72.64	73.00
Value obtained by calculating the distances to galaxies from the luminosity of the Cepheid observed in them using Hubble Space Telescope (E_2)	74.03		
Value that obtained by taking into account the surface brightness fluctuations of galaxies (E_3)	76.50		

Tables no 1 and no 2 show that for each measurement data set the intermediate value of E_2 is closer to either the lower or the upper limit of the searchable quantity estimate. The averaging results show that for the variants considered, the arithmetic mean Hubble constant values are located substantially closer to the middle of their scatter range, deviating from it by +0.72 and -0.69 (km/s)/Mpc respectively. However, the averages obtained using expression (10) deviate significantly more from the spread range middle: by +1.07 and -1.05 (km/s)/Mpc respectively, as in this case the position of the intermediate value between the range boundaries has a greater influence on the averaging result. It cannot be ruled out that advances in cosmology will allow more accurate measurement methods to be developed in the future and a new Hubble constant estimate will come close to one of the proposed averaged estimates.

IV. Geometric interpretation of the problem

The mean value determined by expression (10) can be represented geometrically as the intersection point of the quadrangle diagonals. Its vertices are common points of linear functions

$$f_i(x) = E_i \pm \delta_i(x), \tag{12}$$

where i=1, 2, 3 and $\delta_i(x)=x$.

The geometric interpretation of the Hubble constant measurements averaged estimate, using the data from Table 1 as an example, is given in Figure 1. The dependencies $\delta_i(x)$ imply a measured values deviation from its true value, and are thus not related to the accuracy of the measurement methods themselves.

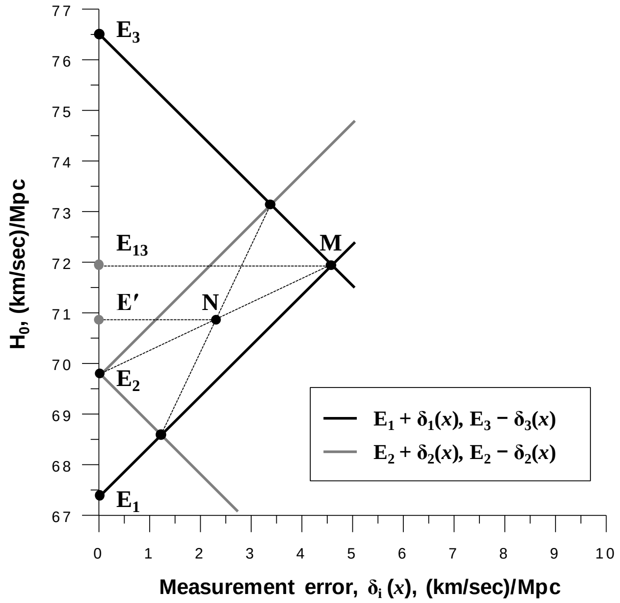

Figure 1 shows that the functions $f_i(x)$ defined by expression (12) have slopes +45° and -45° to the horizontal, hence intersect at right angles and their intersection points form a rectangle. At the same time it follows that the angles ME_1E_3, E_1E_3M and the sides E_1M, ME_3 of right-angled triangle E_1ME_3 are equal. Therefore the height ME_{13} of triangle E_1ME_3 halves its hypotenuse, the line segment E_1E_3, i.e. the point M projection onto vertical axis is E_1 and E_3 arithmetic mean. It is also evident that the height ME_{13} of triangle E_1ME_3 is equal to half of its hypotenuse E_1E_3 and together with the rectangle ME_2 diagonal form a right triangle $ME_{13}E_2$. The point N and its projection on the vertical axis (point E') form the line segment NE', which divides its cathetus E_2E_{13} and hypotenuse ME_2 into line segments, respectively. Since the intersection point N halves the diagonals of the rectangle (hypotenuse ME_2), the cathetus E_2E_{13} of triangle $ME_{13}E_2$ divides into two equal line segments: $E_2E'=E'E_{13}$. This equality leads directly to expression (10), since it is obvious that the E' value is obtained by adding half the length of the line segment E_2E_{13} to the E_2 value:

$$E' = E_2 + \frac{E_{13} - E_2}{2} = \frac{E_1 + 2E_2 + E_3}{4}. \quad (13)$$

V. Conclusion

The limited and mutually contradictory nature of the separate numerical values of the searchable quantity creates natural uncertainty in its true value estimation. Where each result is obtained independently of the others, averaging gives a refined estimate. With a minimum of three separate numerical values, two of them have the largest difference and define the spread limits. The intermediate result is generally closer to one of the bounds, and is thus a clarifying, keeping the estimation of the searchable quantity in question deterministic. This is defined by a general solution to a system of linear algebraic equations, which relates the searchable quantity true value to its three independently obtained numerical values and the errors in obtaining them. The structure of the general solution, which has an infinite number of particular solutions, relates in pairs an intermediate value with two extreme values. Thus, the average estimate can be seen as the arithmetic mean of the two arithmetic means linking the intermediate value to each of the extreme values. In geometric interpretation, this averaging method derives from the common points existence in linear error functions for obtaining individual values of the searchable quantity. Determining the position of the rectangle diagonals intersection point constructed using these functions and their common points gives a solution similar to the algebraic one. As a result, averaging taking into account the intermediate result clarifying role, implies a higher likelihood than the arithmetic mean of the searchable quantity all three numerical values.

References

[1]. Natalie Wolchover. Cosmologists Debate How Fast the Universe Is Expanding // Quanta magazine: August 8, 2019. URL: http://www.quantamagazine.org/cosmologists-debate-how-fast-the-universe-is-expanding-20190808, accessed 06.04.2022.

Ph.D. Ruslan Yu. Golikov. "Averaging of the searchable quantity by three its independent values." *IOSR Journal of Mathematics (IOSR-JM)*, 18(2), (2022): pp. 08-11.