How to Define Average Class Size (and Deviations from the Average Class Size) in a Way Which Is Most Adequate for Teaching Effectiveness

Olga Kosheleva[1] and Vladik Kreinovich[2]
[1]Department of Teacher Education
[2]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract

When students select a university, one of the important parameters is the average class size. This average is usually estimated as an arithmetic average of all the class sizes. However, it has been recently shown that to more adequately describe students' perception of a class size, it makes more sense to average not over classes, but over all students – which leads to a different characteristics of the average class size. In this paper, we analyze which characteristic is most adequate from the viewpoint of efficient learning. Somewhat surprisingly, it turns out that the arithmetic average *is* the most adequate way to describe the average student's gain due to a smaller class size. However, if we want to describe the effect of *deviations* from the average class size on the teaching effectiveness, then, instead of the standard deviation of the class size, a more complex characteristic is most appropriate.

1 Introduction

Average class size is an important characteristic of a university. The fewer students in a class, the more individual attention can a student get, and thus, the better the student learns. Thus, for students, expected class size is an important criterion for selecting a university.

How average class size is usually estimated. Usually, the average class size is estimated by taking the arithmetic average of all the class sizes $s_1, \ldots, s_c$, i.e., as

$$E \stackrel{\text{def}}{=} \frac{s_1 + \ldots + s_c}{c} = \frac{1}{c} \cdot \sum_{i=1}^{c} s_i. \tag{1}$$

Problem with the usual definition of an average class size, and a more adequate definition. Recent papers [1, 2, 3] show that the student's perception of an average class size does not always coincide with the above quantity E.

Indeed, students gauge an average class size by averaging the sizes of the classes in which they are enrolled. At any given moment of time, we have s_1 students enrolled in a class of size s_1, s_2 students enrolled in a class of size s_2, etc. So, to find the average class size from the student perspective, we must add all these numbers and divide by the total number of students. The resulting estimate is:

$$E_s = \frac{s_1 + \ldots + s_1 \, (s_1 \text{ times}) + \ldots + s_c + \ldots + s_c \, (s_c \text{ times})}{s_1 + \ldots + s_c}.$$

By combining the terms equal to each s_i, we get an equivalent expression

$$E_s = \frac{s_1^2 + \ldots + s_c^2}{s_1 + \ldots + s_c},$$

which can be described as the ratio $\frac{M_2}{E}$, where the second moment M_2 is defined as

$$M_2 \stackrel{\text{def}}{=} \frac{1}{c} \cdot \sum_{i=1}^{c} s_i^2.$$

As usual in statistics, we can represent M_2 as $M_2 = E^2 + V$, where the variance V is defined as

$$V \stackrel{\text{def}}{=} \frac{1}{c} \cdot \sum_{i=1}^{c} (s_i - E)^2. \tag{2}$$

Thus, the student-based estimate of the size of a class can be described as

$$E_s = \frac{M_2}{E} = E + \frac{V}{E}. \tag{3}$$

From this formula, we can see that since $V \geq 0$, this student-based average is always not smaller – and often larger – than the usual estimate (1).

This explains why, in the student's opinion, the official estimate (1) of the class size is usually an underestimation.

What we do in this paper. The main objective of a student is to get a better education. From this viewpoint, what is the best estimate of the average class size that the student should use when selecting a university? This is a problem that we will solve in this paper.

2 Utility-Based Definition of Average Class Size

How class size affects education. A student's gain from the class consists of two parts:

First, there is knowledge that this student gains during lectures and classes, by reading the materials posted on the web (and, more generally, by activities that require individualized interactions – such as a student's personal interaction with the instructor, or the interaction of the instruction with a small study group containing this student). The amount a of this knowledge does not depend on the class size.

Also, there is knowledge that comes from an individualized contact with the constructor:

- from a personal contact with an instructor, or
- from the time that the instructor spends with the small study group containing this student.

Let us assume that the instructor spends time t_p on personal one-on-one interactions with students and time t_g on contacts with student study groups, and let g be an average size of a study group. Then, on average, the time that a student gets for personal contacts with the instructor is equal to $\frac{t_p}{s_i}$. Thus, the resulting knowledge is also inversely proportional to s_i, i.e., equal to $\frac{b_p}{s_i}$, for some constant b_p.

Similarly, the number of study groups of size g formed by s_i students is, on average, equal to s_i/g. So, on average, the instructor spends time $\frac{t_g}{s_i/g}$ with the group containing the given student. This time is also inversely proportional to s_i. Hence, the resulting amount of knowledge is also inversely proportional to s_i, i.e., equal to $\frac{b_g}{s_i}$, for some constant b_g.

So, the amount of knowledge gained via such individualized instructions is equal to $\frac{b_p}{s_i}+\frac{b_g}{s_i}=\frac{b}{s_i}$ for $b \stackrel{\text{def}}{=} b_p+b_g$. The overall student gain (= utility) from studying in a class of size s_i is thus equal to $a+\frac{b}{s_i}$.

Resulting average gain. Now that we have s_1 students with gain $a+\frac{b}{s_1}$, s_2 students with gain $a+\frac{b}{s_2}$, $\ldots$, s_c students with gain $a+\frac{b}{s_c}$, we can compute the average gain u as

$$u=\frac{s_1\cdot\left(a+\frac{b}{s_1}\right)+\ldots+s_c\cdot\left(a+\frac{b}{s_c}\right)}{s_1+\ldots+s_c}.$$

Multiplying each term in the numerator, we conclude that

$$u = \frac{a \cdot s_1 + b + \ldots + a \cdot s_c + b}{s_1 + \ldots + s_k} = \frac{a \cdot \sum\limits_{i=1}^{c} s_i + b \cdot c}{\sum\limits_{i=1}^{c} s_i} = a + b \cdot \frac{c}{\sum\limits_{i=1}^{c} s_i},$$

or, in terms of the average (1):

$$u = a + \frac{b}{E}. \tag{4}$$

Conclusion. The average effect of class size on students is inversely proportional to the average class size E – as it is computed usually, by formula (1). This is a somewhat unexpected result since, as we have mentioned, the average class class as perceived by students is different from E.

3 How to Gauge Deviations from the Average Class Size

Gauging deviations is important. Different classes have different sizes. So, for a student, it it important to know not only the *average* class size (or, alternatively, the average gain), but also the *deviations* from the average class size.

Variance and standard deviation as natural measures of deviation. In statistics, deviation from the average is usually gauged by the variance V or by its square root – standard deviation $\sigma = \sqrt{V}$; see, e.g., [4]

Variance of the class size: traditional approach. In the traditional approach, deviation is described by the variance (2).

Variance: student-based approach. In the student-based approach, the average is equal to $E + \dfrac{V}{E}$. So, for all s_i students from the i-th class, the square of the difference is equal to $\left(s_i - \left(E + \dfrac{V}{E}\right)\right)^2$. Thus, the mean value of this square is equal to

$$V_s = \frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} s_i \cdot \left(s_i - \left(E + \frac{V}{E}\right)\right)^2.$$

By using the fact that $s_i = (s_i - E) + E$, we can represent the expression V_s as the sum of two terms $V_s = V_1 + V_2$, where

$$V_1 = \frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} (s_i - E) \cdot \left(s_i - \left(E + \frac{V}{E}\right)\right)^2$$

and

$$V_2 = \frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} E \cdot \left(s_i - \left(E + \frac{V}{E}\right)\right)^2.$$

In the expression for V_1, we can explicitly separate $s_i - E$ in the squared term, thus getting

$$V_1 = \frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} (s_i - E) \cdot \left((s_i - E) - \frac{V}{E}\right)^2.$$

By explicitly describing the square of the difference, we get

$$\frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} (s_i - E) \cdot \left((s_i - E)^2 - 2 \cdot (s_i - E) \cdot \frac{V}{E} + \frac{V^2}{E^2}\right) =$$

$$\frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} \left((s_i - E)^3 - 2 \cdot (s_i - E)^2 \cdot \frac{V}{E} + (s_i - E) \cdot \frac{V^2}{E^2}\right).$$

Here, the average of $s_i - E$ is 0, the average of $(s_i - E)^2$ is V; so, by defining the skewness

$$S \stackrel{\text{def}}{=} \frac{1}{c} \cdot \sum_{i=1}^{c} (s_i - E)^3, \tag{5}$$

we conclude that

$$V_1 = \frac{S - 2V \cdot \dfrac{V}{E}}{E} = \frac{S}{E} - 2\frac{V^2}{E^2}.$$

In the expression for V_2, we move the common factor E outside the sum, getting

$$V_2 = E \cdot \frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} \left(s_i - \left(E + \frac{V}{E}\right)\right)^2.$$

By explicitly performing the squaring, we conclude that

$$V_2 = E \cdot \frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} \left(s_i^2 - 2 \cdot s_i \cdot \left(E + \frac{V}{E}\right) + \left(E + \frac{V}{E}\right)^2\right).$$

The average of s_i^2 is equal to $M_2 = V + E^2$, the average of s_i is equal to E, so we get

$$V_2 = E \cdot \frac{1}{E} \cdot \left(V + E^2 - 2E \cdot \left(E + \frac{V}{E} \right) + \left(E + \frac{V}{E} \right)^2 \right) =$$

$$\frac{1}{E} \cdot \left(V + E^2 - 2E^2 - 2V + E^2 + 2V + \frac{V^2}{E^2} \right).$$

The terms E and $\frac{1}{E}$ cancel other, and so do the terms $2V$ and $-2V$, and the terms E^2, $-2E^2$; so, we get

$$V_2 = V + \frac{V^2}{E^2}.$$

By combining the formulas for V_1 and V_2, we get the following formula for the student-based variance $V_s = V_1 + V_2$:

$$V_s = V + \frac{S}{E} - \frac{V^2}{E^2}. \qquad (6)$$

Variance: utility approach. As we have shown, the utility of each of s_i students enrolled in the i-the class is equal to $a + \frac{b}{s_i}$, and the average utility is equal to $a + \frac{b}{E}$. Thus, for each of these students, the square of the difference between the actual and average utility is equal to

$$b^2 \cdot \left(\frac{1}{s_i} - \frac{1}{E} \right)^2 = b^2 \cdot \left(\frac{1}{s_i^2} - 2 \cdot \frac{1}{s_i} \cdot \frac{1}{E} + \frac{1}{E^2} \right).$$

So, the corresponding variance is equal to

$$V_u = b^2 \cdot \frac{1}{\sum\limits_{i=1}^{c} n_i} \cdot \sum_{i=1}^{c} s_i \cdot \left(\frac{1}{s_i^2} - 2 \cdot \frac{1}{s_i} \cdot \frac{1}{E} + \frac{1}{E^2} \right),$$

i.e., to

$$V_u = b^2 \cdot \frac{1}{\sum\limits_{i=1}^{c} s_i} \cdot \sum_{i=1}^{c} \left(\frac{1}{s_i} - 2 \cdot \frac{1}{E} + \frac{s_i}{E^2} \right).$$

We know that the average of the values s_i is E. Let us denote

$$M_{-1} \stackrel{\text{def}}{=} \frac{1}{c} \cdot \sum_{i=1}^{c} \frac{1}{s_i}. \qquad (7)$$

Then, the above formula takes the form

$$V_u = b^2 \cdot \frac{1}{E} \cdot \left(M_{-1} - 2 \cdot \frac{1}{E} + \frac{E}{E^2} \right),$$

i.e.,

$$V_u = b^2 \cdot \left(\frac{M_{-1}}{E} - \frac{1}{E^2}\right). \tag{8}$$

Comment. In the above formula, we used the expression M_{-1}, moment of order -1. This moment is closely related to an alternative way of describing the mean of several numbers $s_1, \ldots, s_c$, namely, to the *harmonic mean*

$$\frac{c}{\dfrac{1}{s_1} + \ldots + \dfrac{1}{s_c}}.$$

How to meaningfully compare the utility-based variance with other variances. The first two variances are in terms of number of students, while the utility variance is in terms of its inverse. How can we compare the utility-based variance with other variances?

A variance V means that instead of the exact value E, we have $E+k\cdot\sigma$, where $\sigma = \sqrt{V}$ and k is small number – for normal distributions, with certainty 95%, its absolute value is smaller than 2. The utility variance means that instead of the original value of the utility $u = a + \dfrac{b}{E}$, we have $u + k \cdot \sigma_u$, where $\sigma_u = \sqrt{V_u}$. To meaningfully compare this change with other variances, it is desirable to come up with a value $\sigma_e = \sqrt{V_e}$ for which the corresponding change from E to $E + k \cdot \sigma_e$ will lead to exactly this change from u to $u + k \cdot \sigma_u$.

The change in E changes the original value of $u = a + \dfrac{b}{E}$ to the new value

$$u' = a + \frac{b}{E + k \cdot \sigma_e} = a + \frac{b}{E \cdot \left(1 + k \cdot \dfrac{\sigma_e}{E}\right)}.$$

When the deviation σ_e is small, we can ignore terms which are quadratic and higher order in σ_e and conclude that

$$\frac{1}{1 + k \cdot \dfrac{\sigma_e}{E}} \approx 1 - k \cdot \frac{\sigma_e}{E}.$$

Thus, we get the following expression:

$$u' \approx \frac{b}{E} \cdot \left(1 - k \cdot \frac{\sigma_e}{E}\right) = a + \frac{b}{E} - k \cdot \sigma_e \cdot \frac{b}{E^2}.$$

So, the difference $u' - u$ between this expression and the average utility $u = a + \dfrac{b}{E}$ has the form $k \cdot \sigma_u$, where we denoted $\sigma_u = \sigma_e \cdot \dfrac{b}{E^2}$. Since we know σ_u, we can therefore compute the equivalent standard deviation as $\sigma_e = \sigma_u \cdot \dfrac{E^2}{b}$, and the equivalent variance as $V_e = \sigma_e^2 = V_u \cdot \dfrac{E^4}{b^2}$. Substituting the above expression for V_u, we get

$$V_e = E^3 \cdot M_{-1} - E^2. \quad (9)$$

4 Conclusions

Our objective is to describe the effect of class size on *teaching effectiveness*. It turns out that to gauge average effectiveness, it is sufficient to know the arithmetic average of class sizes. This fact is somewhat unexpected: while this arithmetic average is the mostly used characteristic of the average class size, it is not the most adequate in describing *student perception* of class sizes.

Once we know how the *average* class size affects teaching effectiveness, a natural next question is how *deviations* from the average class size affect teaching effectiveness. At first glance, the above conclusion seems to imply that standard deviation of the class size would be the most adequate characteristic of the effect of deviations on the teaching effectiveness. Again, somewhat unexpectedly, it turns out not be the case: the most adequate characteristic is a more complex expression (9) – that uses both the arithmetic mean $E = \dfrac{s_1 + \ldots + s_c}{c}$ of class sizes *and* their harmonic mean $\dfrac{c}{\dfrac{1}{s_1} + \ldots + \dfrac{1}{s_c}}$.

Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825 (Cyber-ShARE Center of Excellence) and DUE-0926721, by Grant 1 T36 GM078000-01 from the National Institutes of Health, and by a grant on F-transforms from the Office of Naval Research.

The authors are thankful to Dr. Lawrence M. Lesser and to all other participants of the 2012 International Sun Conference on Teaching and Learning for valuable discussions.

References

[1] L. Lesser, "Sizing up class size: a deepr classroom investigation of central tendency", *Mathematics Teacher*, 2009, Vol. 103, No. 5, pp. 376–380.

[2] L. Lesser, "Sizing up class size: additional insights", *Mathematics Teacher*, 2010, Vol. 104, No. 2, pp. 86–87.

[3] L. M. Lesser and K. Kephart, "Setting the Tone: A Discursive Case Study of Problem-Based Inquiry Learning to Start a Graduate Statistics Course for In-Service Teachers", *Journal of Statistics Education*, 2011, Vol. 19, No. 3.

[4] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2004.