CLASSROOM CAPSULES

Average Perceived Class Size and Average Perceived Population Density

Clifford H. Wagner (w44@psu.edu), The Pennsylvania State University at Harrisburg, Middletown, PA 17057 doi:10.4169/193113409X458741

It is claimed that mean class size, as perceived by students, is always greater than a school's advertised average class size (see [**1**], [**2**], [**3**]). We give here a new, quick proof, go on to consider the perceptions of students taking multiple classes, and generalize to average population densities within a state or region. As a bonus, we invoke two famous inequalities.

Perceived class size. Consider a school with n classes of size $x_1, x_2, \ldots, x_n$. The advertised average class size for the school is the simple arithmetic mean

$$\bar{x} = \frac{\sum x_i}{n}.$$

But from the perspective of individual students, particularly if each student enrolls in exactly one class, the average perceived class size is more accurately represented by the self-weighted mean

$$\bar{x}_{sw} = \frac{\sum x_i x_i}{\sum x_i} = \frac{\sum x_i^2}{\sum x_i},$$

as explained by Hemenway [**1**], who quantifies the difference between $\bar{x}_{sw}$ and $\bar{x}$, Lann and Falk [**2**], who provide additional examples and references, and Schwenk [**4**]. The simple mean assigns a weight of 1 to each x_i, while the self-weighted mean assigns a weight of x_i to each x_i; that is each class is weighted by the number of attendees.

That the self-weighted mean must be at least as large as the arithmetic mean is stated formally as:

Theorem 1. *If $\{x_i\}$ is a positive sequence of length n, then*

$$\bar{x}_{sw} = \frac{\sum x_i^2}{\sum x_i} \geq \frac{\sum x_i}{n} = \bar{x},$$

with equality if and only if the sequence is constant.

This result follows easily by invoking Chebyshev's Summation Inequality [**3**, pp. 36–37], the first of our famous inequalities. This states that when two nondecreasing sequences have the same length, the mean of the product sequence is at least as large as the product of the two individual sequence means.

What if a student is enrolled in several classes? An obvious answer to this question, one ignored by the cited authors, is for each student to simply average their own several class sizes. Surprisingly, if we then average these perceptions of the various students, the result bears no relation to either $\bar{x}$ or $\bar{x}_{sw}$, as shown by the following two examples.

First, consider the simple but interesting case of a school with five students and five classes: $\{A, B, C\}$, $\{A, B, D\}$, $\{A, C, D\}$, $\{B, C, D\}$, $\{A, E\}$. The mean class size is

$$\bar{x} = \frac{3+3+3+3+2}{5} = \frac{14}{5} = 2.8,$$

and the self-weighted mean is

$$\bar{x}_{sw} = \frac{9+9+9+9+4}{3+3+3+3+2} = \frac{40}{14} \approx 2.857.$$

The five students, A, B, C, D, and E, have individual perceived average class sizes of 2.75, 3, 3, 3, and 2, respectively, and the mean of the individual perceived average class sizes is

$$\frac{2.75+3+3+3+2}{5} = 2.75,$$

a value that is even smaller than $\bar{x}$.

Now consider five students and two classes: $\{A, B, C\}$, $\{A, B, C, D, E\}$. Here the average of the students' average perceived class sizes is

$$\frac{4+4+4+5+5}{5} = 4.4,$$

and is larger than both $\bar{x} = 4$ and $\bar{x}_{sw} = 4.25$.

Perceived population density. Here is another example comparing overall versus individual perceptions. It is a bit more advanced because it is based upon averages of rates. Consider a state whose two counties have the characteristics given in Table 1 (land area is in square miles; population density is in persons per square mile).

The overall mean population density is

$$\bar{d}_{yw} = \frac{\sum x_i}{\sum y_i} = \frac{\sum \frac{x_i}{y_i} y_i}{\sum y_i} = 1{,}195.4,$$

Table 1. County Population Densities

County, i	*Population*	*Land Area*	*Population Density*
1	$x_1 = 120{,}044$	$y_1 = 1{,}234.85$	$x_1/y_1 = 97.2$
2	$x_2 = 1{,}517{,}550$	$y_2 = 135.09$	$x_2/y_2 = 11{,}233.6$
Overall	$\sum x_i = 1{,}637{,}594$	$\sum y_i = 1{,}369.94$	$\bar{d}_{yw} = 1{,}195.4$

which is also a weighted mean of the county population densities weighted by their land areas (the subscript yw means y-weighted, that is, weighted by the sequence $\{y_i\}$). But the mean perceived population density, using population values as weights, is

$$\bar{d}_{xw} = \frac{\sum \frac{x_i}{y_i} x_i}{\sum x_i} = 10{,}417.3.$$

This larger mean is surely much closer to the perception of nearly all residents.

In fact, the mean perceived population density $\bar{d}_{xw}$ is greater than or equal to the overall mean population density $\bar{d}_{yw}$ for every such data set. And of course this principle applies to regions and subregions other than states and counties.

Theorem 2. *If $\{x_i\}$ and $\{y_i\}$ are positive sequences of length n, then*

$$\bar{d}_{xw} = \frac{\sum \frac{x_i}{y_i} x_i}{\sum x_i} \geq \frac{\sum x_i}{\sum y_i} = \bar{d}_{yw},$$

with equality if and only if the ratios $\frac{x_i}{y_i}$ are all equal.

Proof. This theorem is an easy corollary of our second famous inequality, Cauchy's Inequality [**3**, pp. 30–31]. This states that if $\{a_i\}$ and $\{b_i\}$ are sequences of length n, then

$$\sum a_i^2 \sum b_i^2 \geq \left(\sum a_i b_i\right)^2,$$

with equality if and only if the sequences are proportional. If we let $a_i = x_i/\sqrt{y_i}$ and $b_i = \sqrt{y_i}$ for each i, then Cauchy asserts that

$$\sum \frac{x_i^2}{y_i} \sum y_i \geq \sum x_i \sum x_i,$$

which, upon rearrangement, gives the desired result. ■

Table 1 is real data, describing the *two* largest counties in Pennsylvania: Lycoming county has the largest land area, and Philadelphia county has the largest population. If we consider all 67 counties of Pennsylvania [**5**], the mean densities are $\bar{d}_{yw} = 274.0$, and $\bar{d}_{xw} = 2054.6$. And as expected, the mean perceived population density is larger.

So what's the difference between $\bar{d}_{xw}$ and $\bar{d}_{yw}$? By taking variances into account, we can sharpen the previous inequalities into equations that quantify the differences between overall means and perceived means. In the case of class sizes, Hemenway [**1**] showed that

$$\bar{x}_{sw} - \bar{x} = \sigma^2/\bar{x},$$

where σ^2 is the variance of the data set $\{x_i\}$. Since variance is always nonnegative, Theorem 1 is a corollary of this equation, which in turn is an easy corollary of the well-known fact that

$$n\sigma^2 = \sum (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2.$$

Similarly, using $\{y_i\}$ for weights, the weighted mean population density is

$$\bar{d}_{yw} = \frac{\sum \frac{x_i}{y_i} y_i}{\sum y_i} = \frac{\sum x_i}{\sum y_i},$$

and the corresponding weighted variance is

$$\sigma_{yw}^2 = \frac{\sum \left(\frac{x_i}{y_i} - \bar{d}_{yw}\right)^2 y_i}{\sum y_i} = \frac{\sum \frac{x_i{}^2}{y_i}}{\sum y_i} - \bar{d}_{yw}^2,$$

a quantity that is clearly nonnegative.

Dividing the last identity by $\bar{d}_{yw}$, we have this unifying theorem:

Theorem 3. *If $\{x_i\}$ and $\{y_i\}$ are positive sequences of length n, then*

$$\bar{d}_{xw} - \bar{d}_{yw} = \sigma_{yw}^2 / \bar{d}_{yw}.$$

Thus, the difference between $\bar{d}_{xw}$ and $\bar{d}_{yw}$ is directly proportional to the variability of the county densities, and inversely proportional to $\bar{d}_{yw}$. Also note that Theorem 2 is a corollary of Theorem 3. Moreover, letting $y_i = 1$ for each i produces Hemenway's identity, and so Theorem 1 also is a corollary of Theorem 3.

Conclusion. Choosing the most useful descriptors of a data set requires careful consideration, and if it also leads one to learn classical results by two pioneers in the theory of inequalities, so much the better.

Acknowledgment. The author thanks Professor Vedula N. Murty for his helpful suggestions and advice on this topic.

References

1. D. Hemenway, Why your classes are larger than "average," *Math. Mag.* **55** (1982) 162–164.
2. A. Lann and R. Falk, A closer look at a relatively neglected mean, *Teaching Stat.* **27** (2005) 76–80.
3. D. S. Mitrinović, *Analytic Inequalities*, Springer-Verlag, New York, 1970.
4. A. Schwenk, Distortion of average class size: The Lake Wobegon effect, this JOURNAL **37** (2006) 293–296.
5. U.S. Census Bureau, *Census 2000, Summary File 1 (SF 1)*, generated by the author using American Factfinder; available at `http://factfinder.census.gov/` (24 August 2006).

Differentiating the Arctangent Directly

Eric Key (ericskey@uwm.edu), University of Wisconsin-Milwaukee, Milwaukee, WI 53201-0413
doi:10.4169/193113409X458750

We will show by a direct argument that the inverse tangent function is differentiable at all values in its domain.