What does the mean mean?

Nicholas N. Watier
Claude Lamontagne
Sylvain Chartier
University of Ottawa

Key Words: Arithmetic Mean; Constructivism; Least Squares; Analytic Geometry.

Abstract

The arithmetic mean is a fundamental statistical concept. Unfortunately, social science students rarely develop an intuitive understanding of the mean and rely on the formula to describe or define it. According to constructivist pedagogy, educators that have access to a variety of conceptualizations of a particular concept are better equipped to teach that concept in a meaningful way. With this in mind, this article outlines five conceptualizations of the arithmetic mean and discusses how each conceptualization can be presented in the classroom. Educators can use these conceptualizations in order to foster insight into the mean.

1. Introduction

Social science students are often apprehensive (Sciutto 1995) and anxious (e.g. Schacht and Stewart 1990; Zeidner 1991) towards statistics courses. Unfortunately, negative attitudes towards statistics are difficult to change and are directly related to academic performance (Garfield and Ben-Zvi 2007). Moreover, a poor understanding of basic statistical concepts persists after graduation (Groth and Bergner 2006) and negative attitudes are maintained in the workplace (Huntley, Schneider and Aronson 2000). Considering that competence in quantitative methods is a requisite for many careers, it is of considerable importance that students possess a basic understanding of statistical concepts. Needless to say, there is a demand for teaching methods and strategies that aim to reduce trepidation and present statistics in a manner that is both meaningful and enjoyable to students.

Constructivist approaches to teaching statistics have exhibited promise in this regard. For example, active learning exercises (Bartsch 2006; Dolinksy 2001; Helman and Horswill 2003), relating course content to students' background knowledge (Dillbeck 1983), and encouraging students to focus on concepts rather than computation (Hastings 1982) can improve attitudes and performance in introductory statistics courses. According to constructivist pedagogy, the role of educators is to provide opportunities for students to build upon their own understanding of a particular concept (von Glasersfeld 1989). The opportunities that educators can present to students are directly related to the variety of ways in which students might conceptualize a particular topic. As a result, a task of educators is to:

> … construct a hypothetical model of the particular conceptual worlds of the students they are facing. One can hope to induce changes in their ways of thinking only if one has some inkling as to the domains of experience, the concepts, and the conceptual relations the student possess at the moment. (von Glasersfeld 1996, p.7)

Thus, according to constructivist pedagogy, statistics educators should develop hypotheses about their students' current level of understanding and anticipate areas of the subject material that might be difficult to comprehend.

The arithmetic mean is a fundamental statistical concept and is considered to be a core topic in introductory statistics courses (Landrum 2005). However, an intuitive understanding of the mean often escapes students (e.g. Russell and Mokros 1996; Watson 2006). For example, students are capable of defining the mean algorithmically, but they are unable to provide any conceptual insight (i.e. it is the value that minimizes the sum of squared deviations; the value of the mean might not be a member of the data set) (Matthews and Clarke 2007). Like other fundamental statistical concepts (e.g. the general linear model, see Chartier and Faulkner 2008), the mean can be conceptualized in a variety of ways. Consequently, it would be beneficial for students if statistics educators: (a) have at their disposal a variety of possible conceptualizations of the mean; and (b) develop hypotheses regarding how a student might conceptualize the mean. This would aid educators in developing an eclectic range of opportunities to present to students so that students can construct their own understanding of the mean rather than parrot the formula. The purpose of this article is to outline five conceptualizations of the mean and discuss how each conceptualization can be presented in the classroom.

As a preliminary, it should be noted that the conceptualizations are presented in order of mathematical profundity. This ordering scheme follows what we consider to be an appropriate progression of insight into the mean for a social science student. Initially, the mean is conceptualized somewhat informally, relying less on a mathematical description and more on a figurative description. As the student progresses to more advanced courses in statistics, the mean is conceptualized in more formal, mathematical terms. For each conceptualization, we discuss the appropriate audience (e.g. introductory course, upper-year, or graduate level) and the recommended mathematics knowledge that would facilitate understanding.

2. The Socialist Conceptualization

The socialist conceptualization is intended for students in an introductory statistics course in the social sciences. It requires an understanding of the basic arithmetic operators, with which social science students are generally comfortable. This conceptualization also makes use of anthropomorphism; human qualities are assigned to the mean in order to foster meaning.
The mean can be described as a "socialist" measure of central tendency, that is - the mean is the value that each participant receives if the sum is divided equally among all members of the group (Gravetter and Wallnau 2002). Students are well aware that the formula for the mean involves summing the scores in the distribution and dividing the sum by the number of scores in the distribution:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} . (1)$$

However, they often wonder why summing the scores and dividing the sum by n produces a meaningful measure of central tendency. The socialist conceptualization provides an answer: dividing the sum by n is equivalent to determining how the sum can be equally proportioned to all of the scores in the distribution. From this point of view, the mean is an intuitively appealing measure of central tendency because it is the single value that is common to all scores.
The etymology of mean and sum illustrates this point well. Sum originates from the Latin noun summa, which refers to the whole or the total of the parts (Partridge 1977). Thus, the sum of a distribution represents its whole. Dividing the sum by n is equivalent to partitioning the whole of the distribution equally to all of the scores. It is not surprising, then, to find that the etymology of mean contains a "shared by all" connotation (Onions 1966).

3. The Fulcrum Conceptualization

The second conceptualization of the mean makes use of an engineering analogy and is also geared towards an introductory level statistics course. The mean can be described as the fulcrum that is unique to each distribution (Weinberg and Schumaker 1962). As Figure 1 illustrates, if the scores in a distribution are represented on a number line, then the value of the mean is the location on the number line where the sum of negative deviations from that location is equal to the sum of positive deviations from that location. Thus, the mean is equivalent to the balancing point of deviations in a distribution.
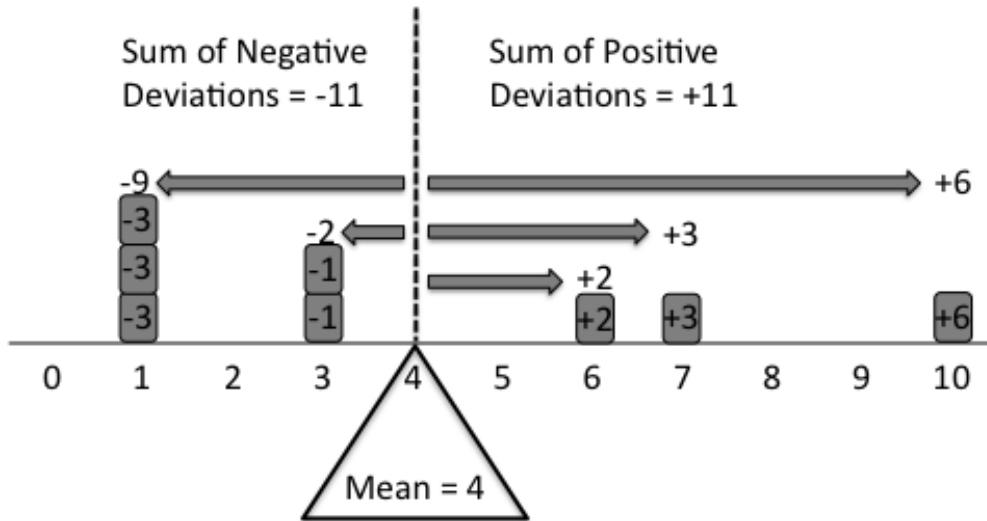
Figure 1. The mean conceptualized as a fulcrum. The distribution {1,1,1,3,3,6,7,10} is represented on a number line. Each grey box represents the existence of a score in the distribution. The number inside each box represents the deviation from the mean. The sum of negative deviations from the mean is equal to the sum of positive deviations from the mean. Thus, the mean acts as a balancing point in the distribution.

Presenting the mean as a fulcrum allows students to visualize the location of the mean relative to all of the scores in the distribution. Consequently, several interesting properties of the mean can be discovered that might otherwise go unnoticed. For example, the fact that the sum of deviations from the mean will always be equal to zero is explicitly presented in the fulcrum conceptualization. Additionally, students may notice that the value of the mean is restricted to the range of scores in the distribution. Hence, they could realize that it is impossible for the value of the mean to be greater than the maximum score or less than the minimum score. Finally, students could discover that the location of the mean is not necessarily the middle of the distribution, which would be useful for helping to distinguish between the mean and the median.

4. Presenting the Socialist and Fulcrum Conceptualizations in the Classroom

Taken together, the socialist and fulcrum conceptualizations capture several aspects of the meaning of the mean. The socialist conceptualization explicitly addresses the rationale behind the formula, whereas the fulcrum conceptualization addresses some of the consequences of using the mean as a representative value of the distribution (i.e. the sum of deviations from mean are equal to zero). Furthermore, the socialist and fulcrum conceptualizations require minimal recourse to mathematics, and as a result, are likely to appeal to students in introductory statistics courses that suffer from mathematics anxiety.

Appendix A describes a sample activity that could be used to introduce the socialist and fulcrum conceptualizations in the classroom. Classroom discussions are well suited for fostering insight into statistical concepts (Garfield and Everson 2009). Consequently, the activity is in the form of a class discussion that centers on a grade comparison problem. Throughout the discussion students are encouraged to offer tentative solutions to the problem. The educator is responsible

for precluding certain solutions and encouraging others. Ultimately, the aim of the discussion is to guide students to arrive at the socialist and fulcrum conceptualizations on their own. Suggestions for doing so are in Appendix A.

5. The Algebraic form of the Least Squares Conceptualization

The least squares criterion can be used to present the mean as a measure of central tendency. The algebraic and geometric forms of the least squares criterion are appropriate for students in upper-year statistics courses who have experience with mean deviations or with the fulcrum conceptualization. The algebraic form is particularly useful for providing a formal criterion as to why the mean is used as a measure of central tendency. If educators decide to present the proof of the least squares criterion, they should be aware that students are expected to know how to expand and differentiate a single variable quadratic polynomial and how to analyze a quadratic function in Cartesian space.

Formally, the goal of least squares is to find the value of c that minimizes the quadratic function $SS(c)$:

$$SS(c) = \sum_{i=1}^{n}(x_i - c)^2 . (2)$$

Legendre, who was the first to present the criterion (Stigler 1986), noted that it could be used to determine a measure of central tendency:

> We see, therefore, that the method of least squares reveals, in a manner of speaking, the center around which the results of observations arrange themselves, so that the deviations from that center are as small as possible. (Legendre 1805, p.75, as cited in Stigler 1986, p.14)

A suitable measure of central tendency should produce a value that is representative of the distribution. Legendre's insight suggests that representative can be formally defined as the value that minimizes the sum of squared deviations. Not surprisingly, the mean is the value that satisfies the least squares criterion.

Returning to Equation 2, the proof that the mean is the value of c that minimizes $SS(c)$ can be presented to students using a combination of algebra, calculus, and geometry. If $SS(c)$ is expanded and rearranged, then $SS(c)$ is equal to

$$nc^2 - 2c\sum_{i=1}^{n}X_i + \sum_{i=1}^{n}X_i^2 . (3)$$

Note that Equation 3 has the general form of the quadratic formula,

$$SS(c) = Ax^2 + Bx + C , \text{ where } A = n, B = -2\sum_{i=1}^{n}X_i, C = \sum_{i=1}^{n}X_i^2, x = c. (4)$$

As Figure 2 illustrates, a quadratic function takes the form of a parabola when graphed in Cartesian space. A parabola will open upward if the A coefficient of the quadratic function is positive. The vertex of a parabola that opens upward is located where the quadratic function reaches its minimum value. In the case of SS(c), the location of the vertex is the value that minimizes the sum of squared deviations. If the quadratic function is expressed in the general form, as demonstrated in Equation 4, then the coordinate of the vertex can be found by calculating the derivative of the function at equilibrium (i.e. where the slope of the parabola equals zero) and solving for x:

$$\frac{d(\text{SS}(c))}{dx} = 2Ax + B = 0$$

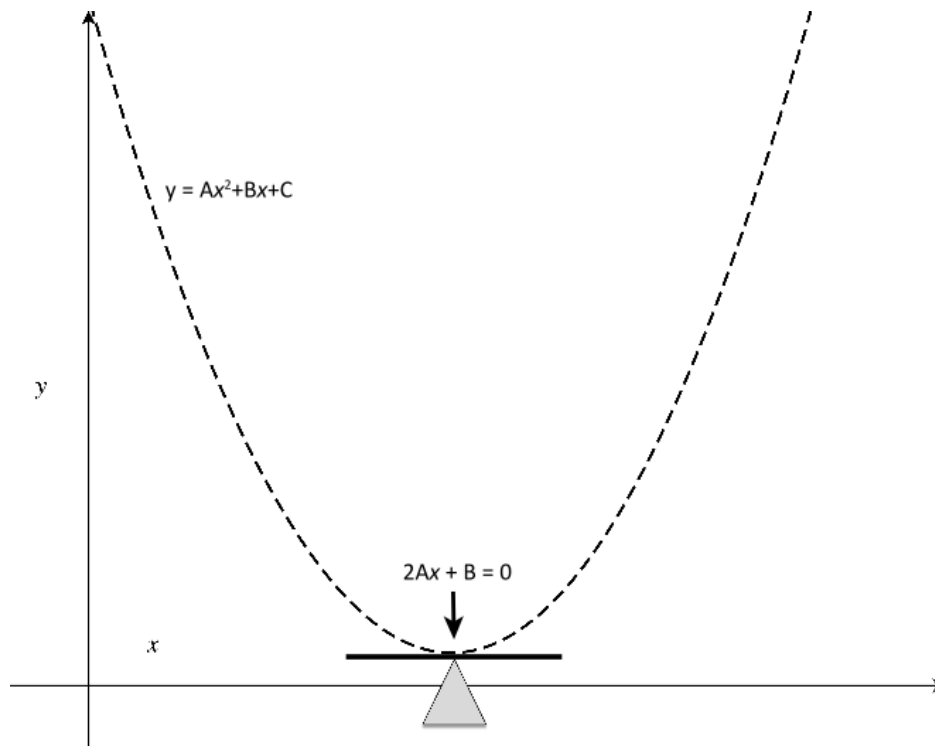$$\Rightarrow x = \frac{-B}{2A} \qquad . \ (5)$$

Figure 2. A graph of the quadratic function $y = Ax^2 + Bx + C$. The function is expressed in the general form of the quadratic formula. The A coefficient of the function is positive, which results in a parabola that opens upward. The vertex of the parabola is located where the derivative of the function is at equilibrium.

Substituting the values of A, B, and x from Equation 4 into Equation 5, the value of c that minimizes the function SS(c) is then

$$\frac{-(-2\sum_{i=1}^{n} X_i)}{2n} = \bar{x}. \ (6)$$

Simplifying Equation 6 reveals that the mean (Equation 1) minimizes the sum of squared deviations. Thus, the mean is the unique value in a distribution that satisfies the least squares criterion.

There are advantages and disadvantages of presenting the least squares criterion. On the one hand, it introduces the method of least squares, which is useful for understanding other fundamental statistical concepts (e.g. variance, correlation, regression, normalization). On the other hand, it may deter students from attempting to understand the mean beyond the formula because the proof relies on recourse into geometry and calculus. While most university students in social science programs will have completed algebra and calculus courses at a secondary level, some could be discouraged by the mathematics and revert to memorizing the formula rather than learning the rationale behind the formula, which would be counter-productive to understanding. Thus, if educators decide to present the proof, they should do so alongside a review of the necessary mathematics.

6. The Geometric form of the Least Squares Conceptualization

Alternatively, the least squares criterion can be presented geometrically. From this point of view, the mean is equivalent to the coordinates of a new origin in n-dimensional space that minimizes the length of the hypotenuse formed from combining vectors of independent observations. The benefit of presenting the geometry of the least squares criterion is that it provides visual representations of the mean and the method of least squares. The geometric form also illustrates how the Pythagorean Theorem, which is often referenced in popular culture (e.g. The Wizard of Oz, The Simpsons), can be used in the context of statistics.

The rationale for this conceptualization is illustrated in Figure 3 with a distribution that contains two observations. As shown in Figure 3, a distribution with two independent observations can be represented in 2D space as orthogonal vectors originating from the standard origin. The two vectors can be joined to create a right triangle. The Pythagorean Theorem can be used to determine the length of the hypotenuse of the right triangle:

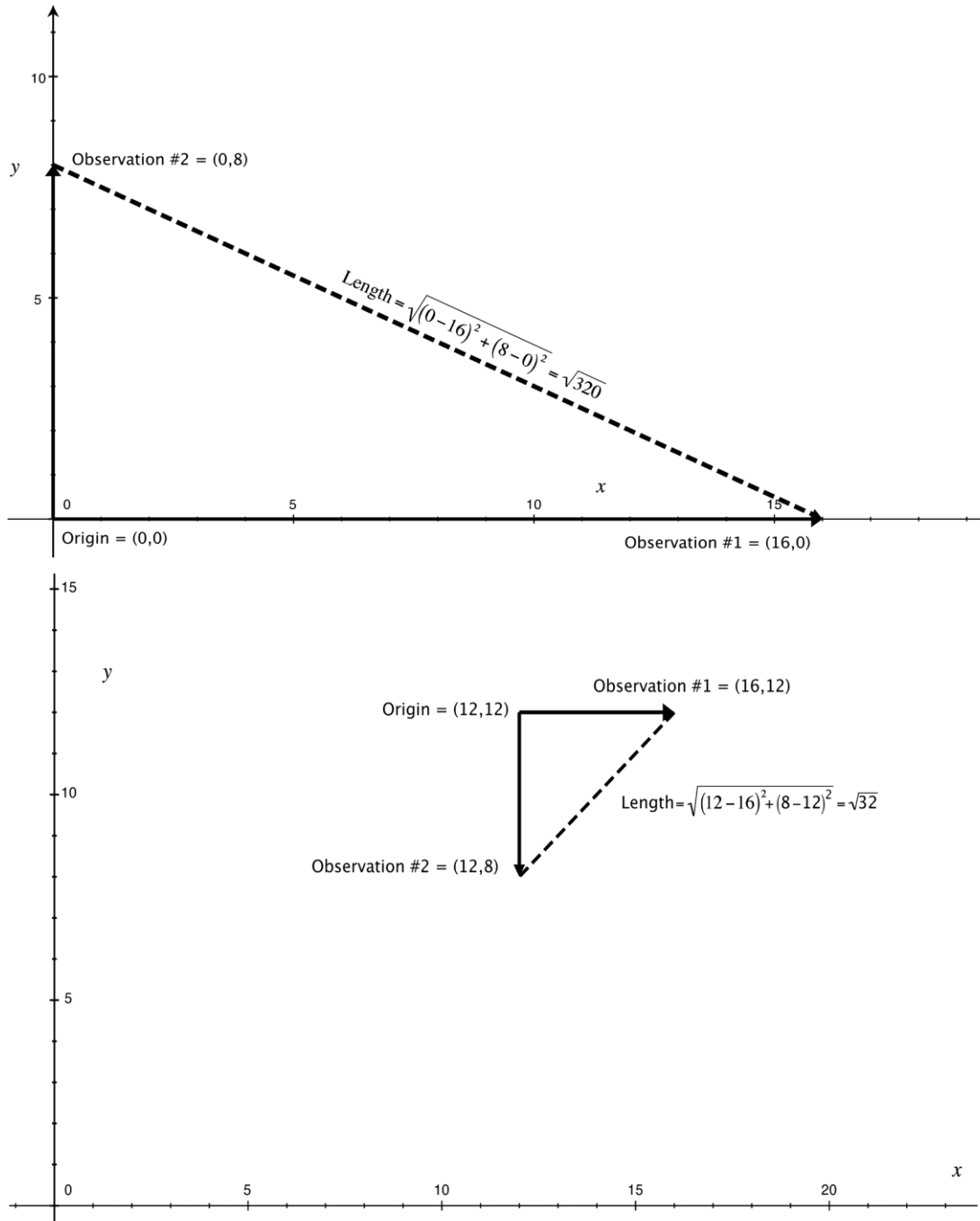$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \ . \ (7)$$

Figure 3. A geometric representation of the least squares criterion. A distribution of two independent observations {8,16} can be represented as orthogonal vectors. Joining the two vectors produces a right triangle. The length of the hypotenuse of the triangle is the square root of the sum of squared deviations of the observations from the origin. The top graph demonstrates the length of the hypotenuse (dashed line) when the origin is set at (0,0). The bottom graph demonstrates the length of the hypotenuse (dashed line) when the origin is set at the value of the mean (12,12). If the mean is used as the origin, the length of the hypotenuse (i.e. sum of squared deviations) is at a minimum.

Note that the squared length of the hypotenuse is equal to the sum of the squared deviations of the observations from the standard origin. The selection of (0,0) as the origin is arbitrary. According to the least squares criterion, a measure of central tendency should minimize the sum of squared deviations. Geometrically, the least squares criterion amounts to selecting a new set of coordinates for the origin that minimize the length of a hypotenuse formed from combining vectors of independent observations. As the bottom panel of Figure 3 shows, the ideal coordinates for the origin will be equal to the value of the mean. Thus, the mean can be understood as the optimal location of the origin for minimizing the length of a hypotenuse formed from combining vectors of independent observations. Although this example is in 2D space, the rationale generalizes to a distribution of size n in n-dimensional space using the generalized Pythagorean Theorem (Lin and Lin 1990). Note that this is constrained to situations where the origin is in the form of $(a, a, ..a)$.

7. Presenting the Least Squares Conceptualizations in the Classroom

The fulcrum conceptualization provides an ideal starting point for presenting the least squares criterion. Recall that the fulcrum conceptualization illustrates that the sum of deviations from the mean is equal to zero. Importantly, the fulcrum conceptualization brings attention to the importance of using deviations from a point in the distribution as a means of determining a suitable measure of central tendency. Once students realize the relevance of calculating deviations, they can be asked to develop their own mathematical criterion for a measure of central tendency and express it algebraically. If they can provide an algebraic expression for their criterion, then they can be asked to use their criterion to derive a formula for a measure of central tendency. If they cannot develop their own criterion, they can be encouraged to speculate about the mathematical properties of an ideal measure of central tendency. After students have provided some speculative responses, they can investigate whether minimizing the sum of squared deviations is a desirable criterion, and if so, how they can determine the value that minimizes the sum of squared deviations.

To facilitate insight into the algebraic and geometric forms of the least squares criterion, we developed an interactive computer exercise for students. The exercise can be downloaded at: http://www.amstat.org/publications/jse/v19n2/LeastSquaresDemonstration.zip. Unzip the file named LeastSquaresDemonstration.nbp.

To run the file requires the Wolfram Mathematica[®] Player to be installed on the computer. The software can be downloaded for free at: http://www.wolfram.com/cdf-player/. Follow the instructions below:

- Go to http://www.wolfram.com/cdf-player/.

- Enter an e-mail address in the response box.

- Click on the Start Download button. After you have installed the software on your computer then you should be able to open the file LeastSquaresDemonstration.nbp. (Note – You must Enable Dymanics to run the demonstration.)

The exercise (see Appendix B) depicts, for a distribution of two data, how the sum of squared deviations varies as a function of the value of the mean, and similarly, how the length of a hypotenuse varies as a function of the value of the origin. When the appropriate value of the mean is selected, or when the value of the mean is used as the origin, students can discover that the sum of squared deviations or the length of the hypotenuse is at a minimum, respectively. We recommend that students complete the exercise only after the least squares conceptualizations have been discussed in the classroom. A worksheet for the exercise can be found in Appendix B.

8. The Vector Conceptualization

Analytic geometry can be used to illustrate how fundamental concepts such as central tendency, variability, degrees of freedom, correlation, and regression are based on a small number of underlying principles that arise naturally from manipulating vectors in Euclidian *n*-dimensional space. Although some educators consider geometric representations to be more insightful for novices compared with algebraic representations (Margolis 1979; Bryant 1984; Saville and Wood 1986; Bring 1996), based on our experiences, we feel that the analytic geometry approach is not suitable for the majority of social science students at the undergraduate level because they become overwhelmed with the mathematics. Thus, the final conceptualization is intended for advanced undergraduate or graduate level students who are interested in learning how analytic geometry can be used to express statistical concepts.

Within the context of analytic geometry, the mean can be conceptualized as a projection of an *n*-dimensional observation vector onto a one-dimensional subspace that contains vectors with a constant at each component (Wickens 1995). The vector in the subspace that has the shortest distance to the observation vector will have the value of the mean at each component.

In the geometric approach, a distribution of scores is represented in participant space rather than in variable space. In participant space, each axis represents a participant rather than a variable. Variables are represented as vectors. The components or coordinates of a vector are analogous to the scores on a variable. For example, the distribution {2,3,6,4,4} is represented in participant space as a 5-dimensional vector with 5 components, each component representing a participant's score on a variable. In contrast, the distribution in variable space would be represented as 5 points on a single dimension (i.e. 5 points on a line).

The vector that represents scores in a distribution is called an observation vector and is denoted by **x**. The number of scores dictates the dimensionality of **x**. If **x** contains *n* components, then **x** lies in *n*-dimensional participant space. Similarly, if every participant has the same score of 1, the vector will be specially denoted by **1**. An ideal measure of central tendency should produce the same value for every participant in the distribution and it should produce a value that is closest to all of the scores in the distribution. With this in mind, discovering a measure of central tendency using analytic geometry amounts to finding a value (i.e. $\bar{x}$) that when multiplied by **1** produces the shortest distance between **x** and $\bar{x}\mathbf{1}$.

The distance between **x** and $\bar{x}\mathbf{1}$ is represented by **e**:

$$\mathbf{e} = \mathbf{x} - \bar{x}\mathbf{1} \quad (8)$$

As [Figure 4](#) illustrates, the shortest distance between **x** and $\bar{x}\mathbf{1}$ occurs when **e** and **1** are perpendicular, that is – when **e** is at a right angle from **1** and forms a right triangle with **x**.
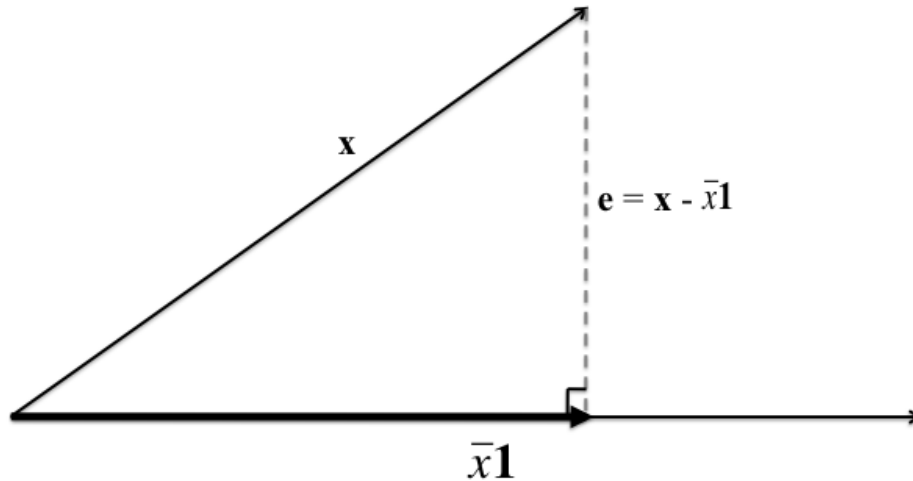

Figure 4. An *n*-dimensional observation vector, **x**, is projected onto a one-dimensional subspace. The vector in the subspace that is closest to **x** is $\bar{x}\mathbf{1}$, which is a scalar multiple of the identity vector **1**. The difference between **x** and $\bar{x}\mathbf{1}$ is represented by **e.** The value of the scalar can be determined by exploiting the orthogonally of **e** and $\bar{x}\mathbf{1}$, that is - by evaluating the dot product of **e** and $\bar{x}\mathbf{1}$, and solving for $\bar{x}$.

Because **e** and **1** are perpendicular, their dot product is equal to zero, as described by the following equation:

$$\mathbf{e}^{\mathrm{T}}\mathbf{1}=0. \text{ (9)}$$

Note that $^{\mathbf{T}}$ represents the transpose operator. The value of $\bar{x}$ that minimizes the distance between **x** and $\bar{x}\mathbf{1}$ can be determined by substituting [Equation 8](#) into [Equation 9](#) and solving for $\bar{x}$, which results in the formula for the mean:

$$(\mathbf{x} - \bar{x}\mathbf{1})^{\mathrm{T}}\,\mathbf{1} = 0$$
$$(\mathbf{x}^{\mathrm{T}}\mathbf{1} - \bar{x}\mathbf{1}^{\mathrm{T}}\mathbf{1}) = 0$$
$$\sum_{i=1}^{n} x_i - \bar{x}n = 0 \qquad (10)$$
$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

We envision this conceptualization to be used in a graduate level course that explores the use of analytic geometry in statistics. By presenting the mean as the coordinates of a vector in a subspace, students are provided with a concrete example of how vector decomposition and projection can be used to express statistical concepts. Thus, the vector conceptualization of the mean can be used as a primer for presenting higher-level statistical analysis within the context of

analytic geometry. Indeed, students can apply the same rationale for the proof of the vector conceptualization of the mean to derive the formulae for regression coefficients.

9. Conclusion

Obviously, certain conceptualizations will be more appealing to a student in an introductory course than others. The socialist conceptualization provides an intuitive explanation for the formula, and the fulcrum conceptualization explicitly reveals several mathematical properties of the mean. Moreover, both conceptualizations provide a foundation for learning the higher-level conceptualizations, and at the same time do not require an extensive recourse into mathematics. For these reasons, the socialist and fulcrum conceptualizations are likely to be the most accessible to a novice. The algebraic and geometric forms of the least squares criterion are suitable for students in upper-year courses who would like a formal proof as to why the mean is used as a measure of central tendency or who would like to know how the Pythagorean Theorem can be applied in a statistical context. Furthermore, these conceptualizations can be used as a primer when teaching the method of least squares. Finally, presenting the mean in the context of analytic geometry would be useful for introducing analytic geometry to advanced undergraduate or graduate level students, especially in multivariate statistics courses where linear algebra is ubiquitous.

Links between conceptualizations can be highlighted in the classroom. For example, the socialist conceptualization presents the idea of distributing a sum of scores equally to all members of a group. The fulcrum conceptualization extends this idea by seeking to determine a point in the distribution that allows for the sum of deviations from that point to be equally distributed or balanced among all members of a group. The utility of calculating deviations from a point in a distribution in order to determine a method of central tendency can be used to bridge the gap between the fulcrum and least squares conceptualizations. Finally, the proof of the vector conceptualizations captures aspects of the socialist and fulcrum conceptualizations. In the former case, the vector conceptualization seeks to determine an equivalency between a vector that represents scores from participants (i.e. $\mathbf{x}$) and a vector that represents a score that is common to all the participants (i.e. $\bar{x}\mathbf{1}$). In the latter case, if $\mathbf{e}$ is squared and differentiated with respect to $\bar{x}$, it would result in the same solution described in [Equation 5](#).

In summary, the mean can be characterized as a socialist, a fulcrum, a solution to the least squares criterion, a set of ideal coordinates of the origin in n-dimensional space, and the coordinates of a vector in a one-dimensional subspace of an n-dimensional participant space. Statistics educators can use these conceptualizations of the mean to foster a deeper level of understanding in the classroom and encourage students to internalize the semantics of the mean in addition to the formula.

Appendix A

The purpose of this exercise is to place students in a situation where they propose the socialist and fulcrum conceptualizations of the mean. The exercise is a structured classroom discussion, where students are encouraged to express speculative solutions to a problem. Below, we describe the problem, summarize a discussion with three participants that attempted to solve the problem, and provide suggestions for educators who wish to use a similar strategy for presenting the socialist and fulcrum conceptualizations in the classroom. The participants were undergraduate social science students who had completed an introductory statistics course.

The Problem

Grades are a topic of interest to students. Consequently, a grade comparison problem is well suited for fostering insight into the mean. The problem is as follows:

> Imagine that your friend recently received his final grade from his introductory statistics course. He received a 55 out of 100, and is quite discouraged. In fact, he comes to the conclusion that he simply is not good at statistics. Being the kind and statistically literate friend that you are, you want to change his mind. Is there any other important information that your friend should know about before he makes his conclusion?

Discussion with Participants

All of the participants initially responded to the problem by wanting to know the grades of the other students in the course. We provided the participants with a distribution of nine scores, which included their friend's score. The distribution did not contain any outliers. At this point, a single strategy emerged from all of the participants: order the scores in the distribution from highest to lowest and mark where their friend's score was in relation to the others. Note that this would be an ideal strategy if outliers were present in the distribution.

The participants were asked why they were comparing their friend's score to every other score. They responded by noting that this strategy allows them to compare their friend with the majority of the class, and if their friend scored higher than the majority, then their friend should not feel bad about his mark. We then noted that, although it is an acceptable strategy, it would be quite cumbersome if the class size were 100 or 250 students. We then asked the participants if they could discover a method that requires only one comparison. The participants responded by comparing their friend's score with the highest or lowest in the distribution, and that this would allow them to determine how their friend performed relative to the poorest or greatest student.

We pointed out to the participants that the lowest or highest score in the distribution does not reflect the overall class performance, which the participants agreed was of interest. They proposed that the midpoint of the distribution provides a measure of overall class performance. According to the participants, the midpoint can be considered representative of the class because it equally separates the highest scores from the lowest scores in the distribution and allows for a single comparison to be calculated.

We intervened here by noting that although the midpoint satisfies our criteria, it does so at the cost of ignoring the other scores in the distribution, that is – the value of the midpoint does not change if the extreme scores were removed from the distribution or if the cardinality of a score changes as long as its rank remains the same. We encouraged the participants to discover an alternative method of finding a single value that reflects the overall class performance. All of the participants then proposed to calculate the mean, but they could not explain why the equation for the mean results in a single value that reflects the overall class performance. We directed the participants' attention to the terms in the equation of the mean, and asked them to speculate on why they are calculating the sum of the scores in the distribution. One participant responded that the sum "transcends the data", and that it "means something more than itself". We asked the participant to elaborate on his idea, but he could only repeat his statement. The other participants noted that the sum combines all of the scores in the distribution into a single value, but this value is much larger that most of the scores in the distribution. One participant then commented that it might be beneficial to take into account the number of scores in the distribution when calculating the sum. It was at this point where we described the socialist conceptualization of the mean.

Afterwards, we asked the participants to revisit their initial "score-by-score" comparison strategy, but using the mean as the anchor of comparison instead of their friend's score. The participants calculated the mean deviation for every score in the distribution and ranked them in order of magnitude. We then asked the participants if the mean deviations offer any useful information. One participant commented that mean deviations could be used to determine how every other student performed relative to the class as a whole. Another participant noted that there would always be positive and negative deviations from the mean. It was at this point where we described the fulcrum conceptualization and ended the discussion.

Psychologically, the participants' responses to the problem were very interesting. Participants initially addressed the problem by adopting a "local" comparison strategy. They compared their friend's score with every other score in the distribution in order to obtain a relative estimate of their friend's competence. However, once they realize that their "score-by-score" comparison strategy becomes too cumbersome with a large sample size, they turn their attention to the lowest score or highest score in the distribution. At this point, their strategy shifts to more "global" comparisons, such that they begin to search for a single score that is representative of the entire distribution. After noting that the extremes of the distribution are not representative of overall class performance, they use the midpoint to make their comparison. The midpoint is more global compared with the highest or lowest score, but it fails to incorporate all of the scores in the distribution. To address this problem, the participants proposed the mean, which is a more "global" comparison compared with the median.

Suggestions for Educators

A classroom discussion that centers on a problem is ideal for placing students in a situation where they discover the socialist and fulcrum conceptualization of the mean on their own. The discussion must be structured so that in attempting to solve the problem, students propose: 1) that they need to determine a representative value of the distribution; 2) that they need to take into account all of the scores in the distribution and that summation serves this purpose; and 3) that they must also account for the number of scores in a distribution. Once these proposals have

been voiced in the classroom and students agree upon a satisfactory solution to the problem, the educator can elaborate on how these proposals relate to the socialist and fulcrum conceptualizations.

Appendix B

The purpose of this exercise is to illustrate the geometric and algebraic forms of the least squares conceptualization of the mean.

Instructions

For this exercise, you will need to download and install the Wolfram Mathematica® Player, which is freely available at: http://www.wolfram.com/cdf-player/. Afterwards, download the Mathematica® notebook file that contains the exercise, which is located at: http://www.amstat.org/publications/jse/v19n2/LeastSquaresDemonstration.zip. Unzip the notebook file named LeastSquaresDemonstration.nbp. Once you open the notebook file with the Wolfram Mathematica® Player, you should be presented with two plots, as depicted in Figure 1. (Note – You must Enable Dymanics to run the demonstration.)
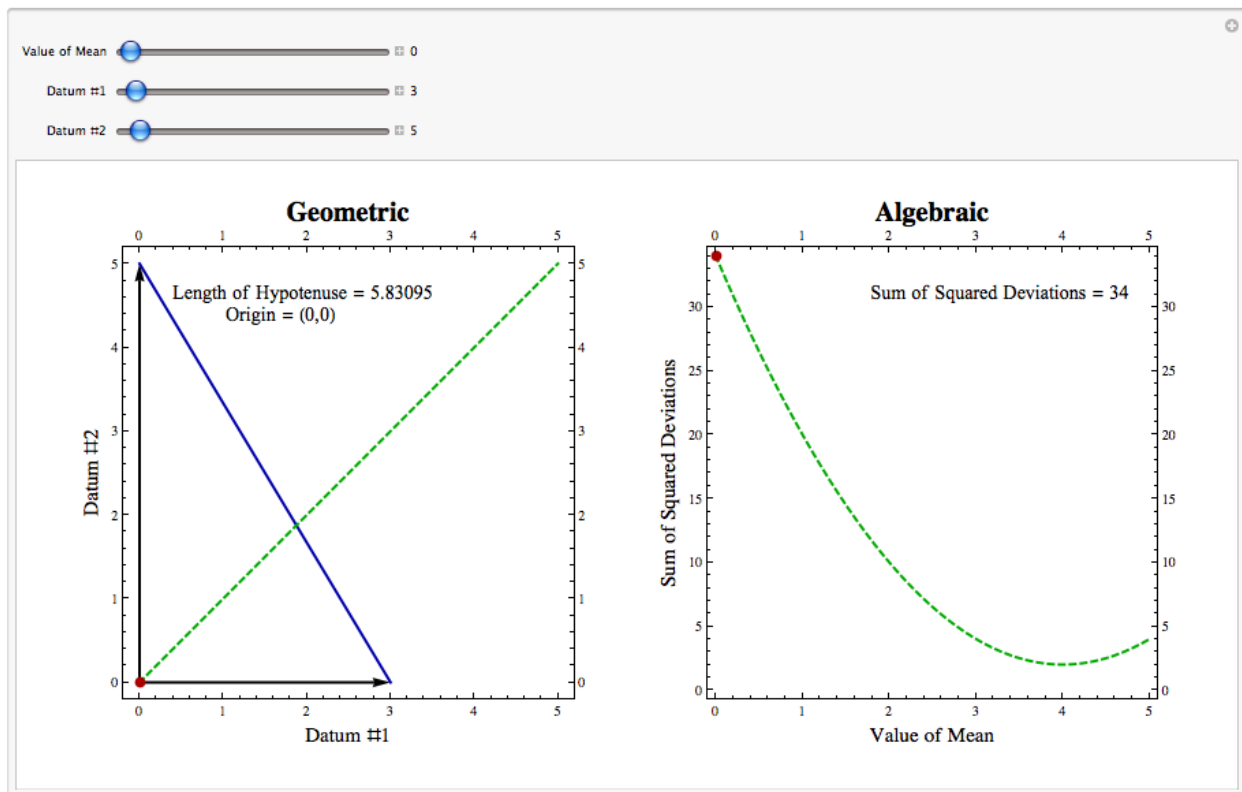

Figure 1. The Least Squares Conceptualizations of the Mean.

The plot on the left illustrates the geometric form of the least squares conceptualization of the mean. It contains two vectors (black arrows). Each vector represents a datum. You can change the value of each datum by moving the appropriate sliders. The default values for the data are 3 and 5. The blue line is the hypotenuse that is formed when the two vectors are combined to make a right-angle triangle. The red dot represents the location of the origin. The default value is (0,0). The dashed green line indicates the permissible locations of the origin. You can select a new location of the origin by moving the slider entitled "Value of Mean".

The plot on the right illustrates the algebraic form of the least squares conceptualization of the mean. The green curve represents how the sum of squared deviations varies as a function of the value of the mean. The red dot represents a value of the mean. The default value is 0. You can change the value of the mean by moving the slider entitled "Value of Mean".

Questions

1. Say you have the following distribution of data: {3,5}.

    a. What is the length of the hypotenuse when the origin is:

        i. (0,0)

        ii. (1,1)

        iii. (3,3)

        iv. (5,5)

        v. (4,4)

    b. Describe what happens to the length of the hypotenuse when the location of the origin changes. What happens to the length of the hypotenuse when the location of the origin is equal to the appropriate value of the mean for this distribution?

    c. Aside from using the value of the mean as the location for the origin, is there any other location for the origin that can produce a shorter hypotenuse?

    d. Describe what happens to the sum of squared deviations when the value of the mean changes. When the appropriate value of the mean is selected, what happens to the sum of squared deviations?

    e. Is there any other value that can produce a smaller sum of squared deviations?

2. Create your own distribution of data. You can select any number from 1-100 for Datum #1 and Datum #2.

    a. Calculate the mean for your distribution.

    b. When the appropriate value of the mean for your distribution is used as the origin, what happens to the length of the hypotenuse?

    c. Describe what happens to the sum of squared deviations as the value of the mean changes. When the appropriate value of the mean is selected, what happens to the sum of squared deviations?

3. Can you think of a distribution where the sum of squared deviations or the length of the hypotenuse will be equal to zero?

4. What can you conclude in regards to the relationship between the mean and the sum of squared deviations or the length of the hypotenuse?

Acknowledgement

References

Bartsch, R.A. (2006), "Improving Attitudes Toward Statistics in the first Class," *Teaching of Psychology*, 33(3) 197-198.

Bring, J. (1996), "A Geometric Approach to Compare Variables in a Regression Model," *The American Statistician*, 50(1), 57-62.

Bryant, P. (1984), "Geometry, Statistics, Probability: Variations on a Common Theme," *The American Statistician*, 38(1), 38-48.

Chartier, S. and Faulkner, A. (2008), "General Linear Models: An Integrated Approach to Statistics," *Tutorials in Quantitative Methods for Psychology*, 4(2), 65-78.

Dillbeck, M.C. (1983), "Teaching Statistics in Terms of the Knower," *Teaching of Psychology* 10(1) 18-20.

Dolinksy, B. (2001), "An Active Learning Approach to Teaching Statistics," *Teaching of Psychology* 28(1), 55-56.

Garfield, J. and Ben-Zvi, D. (2007), "How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics," *International Statistical Review*, 75(3), 372-396.

Garfield, J. and Everson, M. (2009), "Preparing Teachers of Statistics: A Graduate Course for Future Teachers," *Journal of Statistics Education* [Online] 17(2). www.amstat.org/publications/jse/v17n2/garfield.html

Gravetter, F.J. and Wallnau, L.B. (2002), *Essentials of statistics for the behavioural sciences*, 4th edition, Pacific Grove: Wadsworth.

Groth, R.E. and Bergner, J.A. (2006), "Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode," *Mathematics Thinking and Learning*, 8, 37–63.

Hastings, M.W. (1982), "Statistics: Challenge for Students and the Professor," *Teaching of Psychology*, 9(4) 221-222.

Helman, S. and Horswill, M.S. (2003), "Does the introduction of non-traditional teaching techniques improve psychology undergraduates' performance in statistics?," *Psychology Learning and Teaching* 2(1) 12-16.

Huntley, D., Schneider, L., and Aronson, H. (2000), "Clinical Interns' Perception of Psychology and Their Place Within it: The Decade of the 1990s," *The Clinical Psychologist*, 53(4), 3-11.

Landrum, R.E. (2005), "Core Terms in Undergraduate Statistics," *Teaching of Psychology*, 32(4) 249-251.

Lin, S. and Lin, Y. (1990), "The n-Dimensional Pythagorean Theorem," *Linear and Multilinear Algebra* 26, 9-13.

Margolis, M.S. (1979), "Perpendicular Projections and Elementary Statistics," *The American Statistician*, 33(3) 131-135.

Mathews, D. and Clark, J.M. (2007), "Successful Students' Conceptions of Mean, Standard Deviation, and The Central Limit Theorem," Unpublished Paper. Retrieved June 2 2010, from http://www1.hollins.edu/faculty/clarkjm/stats1.pdf.

Onions, C. (1966), *The Oxford dictionary of English etymology*, Oxford: Clarendon.

Partridge, E. (1977), *Origins: A Short Etymological Dictionary of Modern English*, New York: Macmillian Publishing Co., Inc.

Russell, S.J. and Mokros, J. (1996), "What do Children Understand about Average?" *Teaching Children Mathematics* 2(6), 360-364.

Saville, D.J. and Wood, G.R. (1986), "A Method for Teaching Statistics Using N-Dimensional Geometry," *The American Statistician*, 40(3) 205-214.

Schacht, S. and Stewart, B.J. (1990), "What's Funny about Statistics? A Technique for Reducing Student Anxiety," *Teaching Sociology* 18(1), 52-56.

Sciutto, M.J. (1995), "Student-centered methods for decreasing anxiety and increasing interest level in undergraduate statistics courses," *Journal of Instructional Psychology* 22(3) 277-280.

Stigler, S.M. (1986), *The History of Statistics: The Measurement of Uncertainty before 1900*, Harvard University Press: Cambridge, MA.

von Glasersfeld, E. (1989), "Constructivism in Education," In T. Husen and T. N. Postlethwaite, (Eds.) *The International Encyclopedia of Education*, Supplement Vol.1. New York: Pergamon Press 162–163.

von Glasersfeld, E. (1996), "Introduction: Aspects of constructivism," In C. Fosnot (Ed.), *Constructivism: Theory, perspectives, and practice*, New York: Teachers College Press, 3-7.

Watson, J.M. (2006), *Statistical Literacy at School: Growth and Goals*, Mahwah, NJ: Lawrence Erlbaum.

Weinberg, G.H. and Schumaker, J.A. (1962), *Statistics: An intuitive approach*, California: Wadsworth.

Wickens, T.D. (1995), *The Geometry of Multivariate Statistics*, Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Zeidner, M. (1991), "Statistics and mathematics anxiety in social science students—some interesting parallels," *British Journal of Educational Psychology*, 61, 319-328.

Nicholas N. Watier
University of Ottawa
136 Jean Jacques Lussier
Ottawa, Ontario, Canada
K1N 6N5
Email: nwati078@uottawa.ca
Phone: (613) 562-5800 x7150

Claude Lamontagne
University of Ottawa
136 Jean Jacques Lussier
Ottawa, Ontario, Canada
K1N 6N5
Email: Claude.Lamontagne@uOttawa.ca
Phone: (613) 562-5800 x4300

Sylvain Chartier
University of Ottawa
136 Jean Jacques Lussier
Ottawa, Ontario, Canada
K1N 6N5
Email: sylvain.chartier@uOttawa.ca
Phone: (613) 562-5800 x4419

Volume 19 (2011) | Archive | Index | Data Archive | Resources | Editorial Board | Guidelines for Authors | Guidelines for Data Contributors | Guidelines for Readers/Data Users | Home Page | Contact JSE | ASA Publications

20