Bobo Numbers, Bobbers, and Bears—Experiences in Undergraduate Research

Steve Kifowit
Prairie State College
skifowit@prairiestate.edu

November 17, 2016

Undergraduate research experiences are well known for their effects on student learning and engagement. The presenter has mentored community college students whose research has won awards and even been published. In this presentation, you will learn some things about Bobo numbers, falling ladders and floating bobbers, former Chicago Bears head coach Lovie Smith, and student research, in general.

1 What is undergraduate research?

The concept of undergraduate research is simple enough—it is research done by undergraduates. However, students (and faculty) in courses at the freshman and sophomore level may have trouble distinguishing between class projects and real research. Experts warn that schoolwork should not be called "research" unless it contains all those qualities that we have come to expect of research [4].

Undergraduate research...

- must involve original contributions,
- must involve public dissemination of results,
- need not make a great impact, and
- is not valued for its research productivity.

2 The value of undergraduate research

Undergraduate research has been identified as a practice of high value for students, faculty, and institutions. In the first two years, many college students have not settled into their major courses, but nonetheless, research experiences at their level can be deeply meaningful and can provide significant contributions to learning. The following benefits of undergraduate research are well documented (see [4], [10], and the references therein).

- Values for the student:
    1. Increases engagement and improves learning
    2. Leads to higher order thinking, application, and integration
    3. Provides skills that facilitate life-long learning
    4. Improves confidence and positively affects attitude
    5. Increases chances of continuing/completing education
    6. Adds to resume and employment potential
- Values for faculty:
    1. Informs teaching and is good pedagogy
    2. Brings fresh perspectives to the classroom
    3. Provides variety and excitement
    4. May provide greater sense of purpose
- Values for the institution:
    1. Creates long-term connections between students and faculty/campus
    2. Affects college reputation
    3. Can be useful recruiting tool

3 Problems with undergraduate research

At any level, real research consumes a great deal of time and can be fraught with difficulties. Research at the freshman and sophomore levels has some additional, unique challenges.

- Students lack necessary skills (prerequisites, proof, sophistication, tools, etc.)
- Students lack time and motivation
- Faculty lack time and motivation
- Institutional support falls short
- Doable problems can be difficult to pose
- Students tend to be naive and think too big

4 Finding research problems

It can be difficult to find interesting research problems that are original, yet elementary enough for beginning students. When beginning students pose their own research questions, the questions often lack clarity and focus. For faculty wishing to pursue student research, but not knowing where to start, here are some ideas:

- Many journals focusing on undergraduate mathematics publish articles that pose questions.
- Many journals have problem sections with problems ripe for generalization.
- Exercises in textbooks can lead to deeper questions.
- Problems in textbooks can lead to Mythbusters-type experiments.
- Sequences are great sources for projects. In fact, there are many conjectures and open questions in the Online Encyclopedia of Integer Sequences.
- Statistics problems are plentiful.

5 Some samples of undergraduate research

In 2010, Prairie State College (PSC) began to participate in a regional STEM competition. The competition motivated PSC faculty to pursue an undergraduate research agenda and provided a forum for the dissemination of the results of that research. Some of that research has won awards and been published. Four such projects are described below.

5.1 Lovie Smith and separating hyperplanes

At the end of 2012, Chicago Bears head coach Lovie Smith was fired. In the Chicagoland area, Lovie Smith was still the talk of the town well into the spring 2013 semester. A student wishing to do a research project at that time had the following idea: *"I want to prove that Lovie Smith should not have been fired."* This student was hoping to use some sort of statistical analysis to provide the required proof, but after several hours of thought and conversation, we decided there was little new insight we could provide by using statistics. Rather, we decided to search for features that might "separate" winning and losing coaches.

The goal of the project was an unbiased, objective assessment of the conference-winning potential of former Chicago Bears head coach Lovie Smith. Since Lovie Smith is a play-off winning coach, he was only compared to other modern play-off coaches (i.e. NFL head coaches whose teams have made it to the play-offs at least once since 2000). Data were collected on all 52 NFL head coaches who in 2013 had made it to the play-offs since 2000. Seven features of coaches were identified. Quantitative data measuring these features were determined for all 52 coaches. *Excluding Lovie Smith*, the 51 remaining coaches were separated into two groups:

Group 1 (Winning Coaches): Coaches were labeled as winning coaches if their team won its conference at some point since 2000. 24 coaches are in this group.

Group 2 (Losing Coaches): Coaches were labeled as losing coaches if their team never won its conference title since 2000. 27 coaches are in this group.

In order to rate Lovie Smith, we attempted to find a hyperplane in 7-dimensional space that separates Group 1 coaches from Group 2 coaches. A determination of Lovie Smith's location relative to the hyperplane would classify him, relative to his peers, as a winning coach or a losing coach.

A plane in n-dimensional space (i.e a hyperplane) is the set of all points $(x_1, x_2, \ldots, x_n)$ satisfying an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

Two sets of points in n-dimensional space, A and B, are called *linearly separable* if there exists a hyperplane that separates the points, that is,

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n > b, \text{ for all points } (x_1, x_2, \ldots, x_n) \text{ in } A$$

and

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n < b, \text{ for all points } (x_1, x_2, \ldots, x_n) \text{ in } B.$$

For separable data, a separating hyperplane is found by solving a certain linear programming problem. The method was originally proposed in 1991 by Bennett and Mangasarian [1]. If data are not linearly separable, the method yields a "best" separating hyperplane, which almost separates the data in a well-defined sense.

Attempting to separate the 24 winning (Group 1) coaches from the 27 losing (Group 2) coaches, we sought a "best" separating hyperplane in 7-dimensional space:

$$\alpha_1x_1 + \alpha_2x_2 + \alpha_3x_3 + \alpha_4x_4 + \alpha_5x_5 + \alpha_6x_6 + \alpha_7x_7 = \beta.$$

The hyperplane is determined by solving the following linear programming problem for the variables α_1, α_2, $\ldots$, α_7, β, y_1, y_2, $\ldots$, y_{24}, z_1, z_2, $\ldots$, z_{27}.

Minimize

$$F = 0\alpha_1 + 0\alpha_2 + \cdots + 0\alpha_7 + 0\beta + \frac{1}{24}\left[y_1 + y_2 + \cdots + y_{24}\right] + \frac{1}{27}\left[z_1 + z_2 + \cdots + z_{27}\right]$$

subject to

$$\begin{array}{ccccccccccccccc}
-w_{1,1}\alpha_1 & - & w_{1,2}\alpha_2 & - & \cdots & - & w_{1,7}\alpha_7 & + & \beta & - & y_1 & & & \leq & -1 \\
-w_{2,1}\alpha_1 & - & w_{2,2}\alpha_2 & - & \cdots & - & w_{2,7}\alpha_7 & + & \beta & & - & y_2 & & \leq & -1 \\
\vdots & & & & \vdots & & & & \vdots & & & \ddots & & & \vdots \\
-w_{24,1}\alpha_1 & - & w_{24,2}\alpha_2 & - & \cdots & - & w_{24,7}\alpha_7 & + & \beta & & & - & y_{24} & \leq & -1 \\
\ell_{1,1}\alpha_1 & + & \ell_{1,2}\alpha_2 & + & \cdots & + & \ell_{1,7}\alpha_7 & - & \beta & & & - & z_1 & \leq & -1 \\
\vdots & & & & \vdots & & & & \vdots & & & & \ddots & & \vdots \\
\ell_{27,1}\alpha_1 & + & \ell_{27,2}\alpha_2 & + & \cdots & + & \ell_{27,7}\alpha_7 & - & \beta & & & & - \; z_{27} & \leq & -1
\end{array}$$

and

$$y_i \geq 0 \quad i = 1, 2, \ldots 24$$

$$z_j \geq 0 \quad j = 1, 2, \ldots 27$$

where $W = [w_{ij}]$ and $L = [\ell_{ij}]$ are 24×7 and 27×7 matrices, respectively, whose rows are points in 7-dimensional space. The elements of W describe the features of the winning (Group 1) coaches, while the elements of L describe the features of the losing (Group 2) coaches.

The data (i.e the coaches) are linearly separable if and only if the minimum value of F is zero, in which case each $y_i = 0$ and each $z_j = 0$. See [1] or [3] for details of the method.

We used Scilab's `linpro` function to solve the linear programming problem, obtaining the following hyperplane:

$$36.594x_1 + 15.227x_2 + 0.0907x_3 - 0.124x_4 - 0.423x_5 + 0.369x_6 - 0.0839x_7 = 33.812.$$

The minimum value of the F was found to be 0.277. Because this value is nonzero, the coaches are not linearly separable using the 7 features we identified. However, the hyperplane above represents a "best" separating hyperplane according to the criteria described in [1] and [3].

Each coach is classified by substituting his features into the left-hand-side of the hyperplane equation. If a number greater than 33.812 is obtained, the coach is on the winning side of the hyperplane. If a number less than 33.812 is obtained, the coach is on the losing side of the hyperplane. The hyperplane above correctly classifies all but two of the 51 coaches. Dan Reeves is incorrectly classified as a losing coach, and Rex Ryan is incorrectly classified as a winning coach.

A coach's signed distance from the hyperplane can be found by using his features to compute

$$D = 36.594x_1 + 15.227x_2 + 0.0907x_3 - 0.124x_4 - 0.423x_5 + 0.369x_6 - 0.0839x_7 - 33.812.$$

A positive D classifies the coach as a winning coach, while a negative D classifies the coach as a losing coach.

Recall that Lovie Smith's data were not used in determining the best separating hyperplane. Our hyperplane was the result of the analysis of the "training data" for 51 coaches. When Lovie Smith's data are input into the distance equation, his distance from the hyperplane is

$$D = -0.987.$$

This value puts Lovie Smith on the losing side of the hyperplane. In addition, if we rank all 52 coaches according to distance from the hyperplane, Lovie Smith's distance ranks him 26th out of 52. Overall, coaches' distances range from 13.738 for Jim Harbaugh to -18.994 for Eric Mangini.

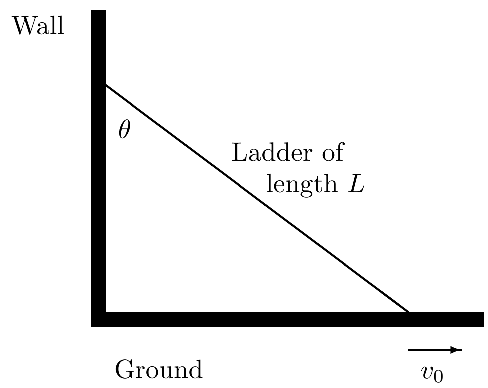
Based on Lovie Smith's classification relative to our hyperplane, we conclude:

- Relative to modern play-off coaches, Lovie Smith has low conference winning potential.
- If conference winning is desired, Lovie Smith's firing is justifiable.

5.2 The falling ladder problem

The falling ladder problem has the top of a ladder of length L sliding down a vertical wall as its base moves away from the wall at a constant speed, v_0 (see Figure 1). Most often the problem asks to find the speed of the top of the ladder as it falls, and paradoxically, the speed approaches infinity. To resolve the paradox, one must recognize that a real ladder will eventually come away from the wall as it falls. Once it does so, the ladder must be treated as a physical pendulum.

Figure 1: The falling ladder



Several very nice articles have been written on the falling ladder paradox [6, 7, 8, 11]. My students noticed, however, that there are absolutely no references to any actual experiments. Wishing to mimic the Mythbusters, we set out to do some experiments. Along the way, we discovered that much remained unsaid about the problem.

It is well established [11] that related rates apply until the ladder has fallen to the point where the angle θ reaches the critical angle

$$\theta_c = \cos^{-1}\left(\sqrt[3]{\frac{2v_0^2}{3gL}}\right),$$

which gives the critical height

$$y_c = L\cos\theta_c = \sqrt[3]{\frac{2L^2v_0^2}{3g}}.$$

Once $\theta > \theta_c$, the ladder must be modeled as a pendulum. Perhaps our most interesting discovery was that a closed-form formula can easily be derived for the speed of the top of the ladder, even after related rates no longer apply:

$$\frac{dy}{dt} = -L\sin\theta\,\sqrt{A - B\cos\theta}, \quad \theta > \theta_c, \tag{1}$$

where

$$A = \frac{3}{4}\left(\frac{12v_0g}{L^2}\right)^{2/3} \qquad \text{and} \qquad B = \frac{3g}{L}.$$

We were very surprised to find that this formula does not appear in any falling ladder discussion that we have encountered. We were also surprised to find that some modern calculus textbooks seem to ignore (or not notice) the falling ladder paradox.

Our real-world experiments proved to be much more difficult than expected. We were a bit naive in hoping to simply measure the speed of a falling ladder. Eventually we constructed a device consisting of a simulated ladder, an inertial beam, a pivot mechanism, and a disengagement indicator from which we obtained consistent results. Though we obtained good results, we found that the theoretical predictions always underestimated our observed values of y_c. Air resistance provides a reasonable explanation for the discrepancies, and it should probably not be ignored (depending on your size and style of ladder).

Over the last 30 years, the falling ladder paradox has generated a great deal of interest, but there are other related-rates paradoxes that have attracted little, or no, attention. We mention only the *boat/bobber problem.* In this problem, a boat or a fisherman's bobber is pulled in at a constant rate toward a dock above the water level. Students are asked to find the speed of the boat/bobber as it approaches the dock. As in the falling ladder problem, the related-rates approach predicts speeds tending to infinity. However, as any fisherman can attest, the floating object will come out of the water and swing as a shrinking pendulum at some point as it approaches the dock. (Incidentally, we believe that the boat in [9, page 155, exercise 28] is too close to the dock for related rates to apply.)

5.3 Bobo numbers

In a Classroom Capsule published in the College Mathematics Journal, E. Ray Bobo [2] studied the sequence $\{a(n)\}_{n=2}^{\infty}$, where $a(n)$ is the least positive integer such that

$$\frac{1}{n} + \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{a(n)} > 1.$$

Bobo showed that $a(n)$ takes one of three values: $\lfloor ne \rfloor - 2$, $\lfloor ne \rfloor - 1$, or $\lfloor ne \rfloor$, where $\lfloor \cdot \rfloor$ is the integer floor function and e is the base of the natural logarithm. He also took a special interest in the set of n's for which $a(n) = \lfloor ne \rfloor$:

$$\begin{aligned} \mathcal{B} &= \{n : a(n) = \lfloor ne \rfloor\} \\ &= \{4, 11, 18, 25, 32, 36, 43, 50, 57, 64, 71, 75, 82, 89, 96, 103, 114, \\ &\quad 121, 128, 135, 142, 146, 153, 160, 167, 174, 185, 192, 199, \ldots\}. \end{aligned}$$

While Bobo did not rule out the possibility that $a(n) = \lfloor ne \rfloor - 2$, his numerical experiments (through $n = 2115$) provided no examples. Furthermore, Bobo observed that the gaps between consecutive elements of $\mathcal{B}$ formed a pattern of 7's, 4's, and 11's. At the end of his article he posed several questions, two of which were:

A: Does the pattern in the gaps of $\mathcal{B}$ persist, or does chaos eventually take over?

B: Is $a(n) = \lfloor ne \rfloor - 2$ impossible?

In Spring 2011, one of my students chose the research project of answering Bobo's questions. In addition to showing that $a(n) = \lfloor ne \rfloor - 2$ is indeed impossible, we obtained a number of significant results concerning the set $\mathcal{B}$. We first named the elements of $\mathcal{B}$ the Bobo numbers in honor of E. Ray Bobo. Then we began to chip away at Question A, attempting to unravel the pattern in the Bobo numbers. Here are some of our results. Details can be found in [5]. In all that follows, the notation $\{\cdot\}$ is used to denote the fractional part of the real number.

Proposition 1 *For $x \geq 1$, let $g(x) = xe - 1 - x\exp\left(1 - \frac{1}{2x} - \frac{1}{2xe-2}\right)$. If n is a Bobo number, then $\{ne\} > g(n)$. Furthermore, g is an increasing function with* $\lim_{x\to\infty} g(x) = \frac{1}{2}(e-1)$.

Proposition 2 *Suppose n is a Bobo number and k is a positive integer. If $3e - 8 \leq \{ke\} \leq 9 - 3e$, then $(n + k)$ is not a Bobo number. This directly rules out gaps of size 1,2,3,5,6,8,9, or 10 between consecutive Bobo numbers.*

Proposition 3 *Suppose n is an integer and $n \geq 2$. If $\{ne\} \geq \frac{1}{2}(e - 1)$, then n is a Bobo number.*

Definition 1 *A Bobo number, n, is said to be* exceptional *if $\{ne\} < \frac{1}{2}(e - 1)$. A Bobo number that is not exceptional is called* unexceptional.

Proposition 4 *For any positive integer n, there exists an integer k in $\{1, 2, 3, \ldots, 11\}$ such that $(n + k)$ is a Bobo number. As a consequence, the gaps between consecutive Bobo numbers cannot exceed 11.*

Proposition 5 *Let n be a positive integer and let $A_n = \{1, 2, 3 \ldots, n\}$. Referring only to unexceptional Bobo numbers, as $n \to \infty$, the proportion of those in A_n approaches $\frac{1}{2}(3 - e) \approx 0.14086$. Furthermore,*

1. *the proportion of those preceding 4-gaps approaches $\frac{1}{2}(7e - 19) \approx 0.013986$,*
2. *the proportion of those preceding 7-gaps approaches $\frac{1}{2}(41 - 15e) \approx 0.11289$, and*
3. *the proportion of those preceding 11-gaps approaches $\frac{1}{2}(7e - 19) \approx 0.013986$.*

5.4 Exceptional Bobo numbers

Refer to Definition 1 above and note that the Bobo numbers come in two varieties: exceptional and unexceptional. The exceptional Bobo numbers are rare! Within the first 6 billion positive integers, there are more than 845 million Bobo numbers, only four of which are exceptional: 36, 9045, 5195512, and 5311399545. By the end of the Bobo number research project described above and the publication of [5], only the first two were known. The next two were found more recently by doing exhaustive computer searches. In Spring 2016, two of my students took on the task of finding and describing the exceptional Bobo numbers.

Here are some new results:

- If n is an exceptional Bobo number, then
$$\frac{e-1}{2} - \frac{1}{2ne - 10 + 3e} < \{ne\} < \frac{e-1}{2}.$$
This result improves on Proposition 1 above, and it sets the stage for using continued fractions.
- Exceptional Bobo numbers always follow a gap of size 4.
- There are infinitely many exceptional Bobo numbers, and they can all be obtained from the denominators of the odd-indexed convergents of the continued fraction for $(e - 1)/2$.
- There are two types of exceptional Bobo numbers. Type-1 EBNs are those obtained directly from the continued fraction convergents. Their computation requires only integer arithmetic. Type-2 EBNs are computed from Type-1 EBNs. They must be confirmed as exceptional using high-precision floating-point arithmetic.

Here are some details to illustrate our results...

The continued fraction expansion for $\frac{e-1}{2}$ is given by

$$\frac{e-1}{2} = [0; 1, 6, 10, 14, 18, \ldots] = \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{1}{10 + \cfrac{1}{14 + \cfrac{1}{18 + \cdots}}}}}.$$

The convergents with indices 3,5,7, 9, and 11 are, respectively,

$$\frac{61}{71}, \quad \frac{15541}{18089}, \quad \frac{8927353}{10391023}, \quad \frac{9126481321}{10622799089}, \quad \frac{14586253530421}{16977719590391}.$$

The first five EBNs are of Type-1 and follow directly from the denominators by adding 1 and dividing by 2. They are: 36, 9045, 5195512, 5311399545, 8488859795196. The sixth EBN is of Type-2. It is

$$3 \times 8488859795196 - 1 = 25466579385587.$$

All Type-1 EBNs follow the pattern described above. The Type-2 EBNs have the form

$$mn - \frac{m-1}{2},$$

where n is a Type-1 EBN and $m \geq 3$ is an odd integer. Most numbers of this form, however, are not exceptional Bobo numbers. We are in the process of writing up our results, but in the meantime, we would be happy to share the details.

References

[1] K. P. Bennett and O. L. Mangasarian, *Robust linear programming discrimination of two linearly inseparable sets*, Optimization Methods and Software, 1 (1992), pp. 23–34.

[2] E. R. Bobo, *A sequence related to the harmonic series*, College Math. J., 26 (1995), pp. 308–310.

[3] R. A. Bosch and J. A. Smith, *Separating hyperplanes and the authorship of the disputed federalist papers*, American Mathematical Monthly, 105 (1998), pp. 601–608.

[4] D. F. Brakke, M. L. Crowe, and K. Karukstis, *Perspective: Reasons deans and provosts (and presidents) should value, support, and encourage undergraduate research*, CUR Quarterly, 30 (2009), pp. 10–14.

[5] D. T. Clancy and S. J. Kifowit, *A closer look at Bobo's sequence*, College Math. J., 45 (2014), pp. 199–206.

[6] W. P. Fox and J. B. Fox, *Modeling the sliding/falling ladder paradox*, PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies, 13 (2003), pp. 277–286.

[7] M. Freeman and P. Palffy-Muhoray, *On mathematical and physical ladders*, American Journal of Physics, 53 (1985), pp. 276–277.

[8] S. Kapranidis and R. Koo, *Variations on the sliding ladder problem*, The College Mathematics Journal, 39 (2008), pp. 374–379.

[9] R. Larson and B. H. Edwards, *Calculus*, Brooks/Cole, Cengage Learning, 9th ed., 2010.

[10] D. Lopatto, *Science in Solution: The Impact of Undergraduate Reasearch on Student Learning*, Research Corporation for Science Advancement, 2009.

[11] P. Scholten and A. Simoson, *The falling ladder paradox*, The College Mathematics Journal, 27 (1996), pp. 49–54.