Humanistic Mathematics Network Journal

Issue 17 Article 13

5-1-1998

Real Data, Real Math, All Classes, No Kidding

Martin Vern Bonsangue
California State University, Fullerton

Real Data, Real Math, All Classes, No Kidding

Martin Vern Bonsangue
Department of Mathematics
California State University
Fullerton, CA 92634
mbonsangue@fullerton.edu

Unlike most university professors, I took the scenic route. For three years I taught mathematics and choral music at a junior high school in East LA; for six years, at the high school next door to it; and for six more years, at the community college a few miles down the freeway from both of them. So, when I work with colleagues who teach at the elementary, middle school, or high school level, I really see us all as just that: *colleagues*. I know lots of people that would be fabulous at my university doing my job, and, frankly, from time to time I wouldn't mind being back at Sierra Vista Junior High.

WHEN AM I EVER GONNA USE THIS STUFF?

So when I get the opportunity to do mathematics with colleagues around the country as part of The National Faculty Institutes and Academic Sessions, I usually try to look at myself from the other end of the pencil. What would I want if someone were going to help me in my mathematics class? What would I not want? Why?

It's clear to me what I would *not* want: to be given the same old stuff in the same old way. I wouldn't even want the same old stuff in a new way (Uri Treisman once remarked that a good pedagogy will not fix a bad curriculum). What I would want is an experience that will cause me to think differently about what I do and why I do it. That means that I have had a mathematical experience that has, I hope, taken my breath away, and, at least, has caused me (as Arsenio Hall used to say) to go "hmmmm." Consequently, that's what I try to provide for my colleagues during these Institutes: a mathematical moment that will affect their professional lives, not just because the dynamic of the shared experience, but *because of the mathematics itself*.

So what is my greatest need as a teacher? To have a good answer to the shop-worn question,

When am I ever gonna use this stuff?

I am embarassed to admit that until recently I basically ignored this question. I gave shop-worn answers like, "Math teaches you how to think, so it doesn't matter," or "Well, we can solve word problems with math," neither of which I believed in. The truth is, I didn't know. After more than two decades of studying mathematics pretty seriously, I didn't have practical applications of percents, algebra, and calculus other than the examples in the books, most of which were either contrived or trivialized, like a 23-minute sitcom that states, develops, and resolves a major crisis. I'd like to think that I taught well, and that most of my students did well, and that some of them enjoyed me and even the math itself. But I still lacked real application of the subject.

In the past, the problem was always access: how do you get hold of real data? Where do you go to look in the library for AIDS reports, carbon dioxide levels, mortality rates, or earthquake incidences? How current is printed information? And who has the time? But with the internet, all of these things are literally a moment away. And, with private companies making internet access increasingly common for educators and non-educators alike, "surfing the net" is no longer done only by computer geeks. Now almost any geek can do it.

So the search for using real data sets in my own classes has launched a somewhat second career for me. This semester I am teaching a course entitled "Calculus for Biologists," a one-semester course built around real data sets that utilize the powerful ideas of calculus, that is, how quantities change in relation to one another, in contexts that are relevant for scientists because of their reality. After only a few weeks it is clear to me how the mechanics of the language mathematics, such as algebraic manipulations, serve primarily to enhance one's understanding of and ability to describe the dynamics of a physical phenomenon. And, if the algebra is minimized, most of the ideas would be accessible to persons with considerably less mathematics background. Indeed, exposure to these phenomena would create for the student a context in which the algebraic structure could take on physical charactersitics, as was described by the Greeks so long ago, but since lost in 20th century textbooks.

What follows is an account of two such lessons that I did with two different groups of middle school teachers. The first lesson, *Earthquakes!*, was done with a group of math/science teachers in Long Beach in southern California. The second lesson, *Math from the Crypt*, was done with a group of teachers from the Mississippi Delta region. Each uses technology in a meaningful way, but is not technology-dependent. After the data are in hand, then the mathematics really begins.

EARTHQUAKES!

Most of us who live on the west coast have experienced an earthquake (some of my out-of-state colleagues say that they would *never* live in California for that reason). Indeed, southern California natives (such as myself) have lived through some memorable shakers: the Magic Mountain earthquake of 1971 (magnitude 7.0); the Whittier Narrows ("Shake and Bake") earthquake of 1986 (magnitude 6.8); and, more recently, the Northridge earthquake in January of 1994 (magnitude 6.9). Although most of us in southern California like to think we're pretty savvy about earthquakes (we know the lingo - Richter scale, epicenter, aftershock), my observation is that we actually harbor many false ideas about earthquakes. For example, how are earthquakes caused? Are they triggered by hot weather? Is "The Big One" likely to happen? Although these questions are geological in nature and require some understanding of the earth's formation, some mathematical observations about earthquake frequencies (how often) and magnitudes (how big) can provide insight. But first, try the Earthquake "Quiz" below. Most of the Californians got fewer than two out of the five questions correct.

What Do You Believe About West Coast Earthquakes?

1. Given that a "felt" earthquake (4.0 or higher) has occurred on the west coast, the chance that it is a "severe" earthquake (6.0 or higher) is about:
a) 50 % b) 25 % c) 10 % d) 5 %
e) 2 % f) less than 1 %

2. Generally speaking, earthquakes with deeper epicenters (10 km or more) will tend to be more/less/about the same in magnitude as those nearer the surface.

3. Southern California tends to get more/less/about the same number of "felt" earthquakes (4.0 or higher) than does northern California.

4. Earthquakes in southern California tend to be more/less/about the same in magnitude as those in northern California.

5. Small earthquakes (between 2.5 and 4.0) constitute about what percentage of all west coast earthquakes?
a) more than 99 % b) about 90 % c) about 75 %
d) about 50 % e) about 25 % f) about 10 %
g) less than 5 %

To explore answers to these (and other) questions, we turned to the internet, hunting earthquake data. While there are several good geological sites that post recent data, we found the Earthquake Laboratory at the University of Washington to be very current and easy to use. The Internet address is:
http://www.iris.washington.edu/FORMS/event.search.form.html
We downloaded earthquake data over a six-year period for earthquakes whose epicenters were in the latitude and longitude range for the west coast (from Baja to Washington). The download yielded 90 (electronic) pages worth of data, a sample of 782 earthquakes of magnitude 4.0 or greater! We used a spreadsheet program to generate the descriptive statistics for this sample (Table 1).

Table 1 yields answers to some of the "quiz" questions almost immediately. For example, the typical earthquake in northern California has a magnitude of about 4.5, same as that in southern California. However, there were about twice as many "felt" earthquakes in southern California (213 compared to 123), while northern California quakes tended to have much deeper epicenters (11.7 km compared to 6.1 km). While these answers trigger more questions that are geologic in nature (e.g., why are northern CA quakes deeper?), they do help bring one's beliefs about earthquake behavior into line with reality. Perhaps the most interesting question centers around the relative frequency of big vs. small earthquakes (Earthquake Quiz, questions 1 and 5). A graph of the number of earth-

Table 1: Summary of Earthquake Data

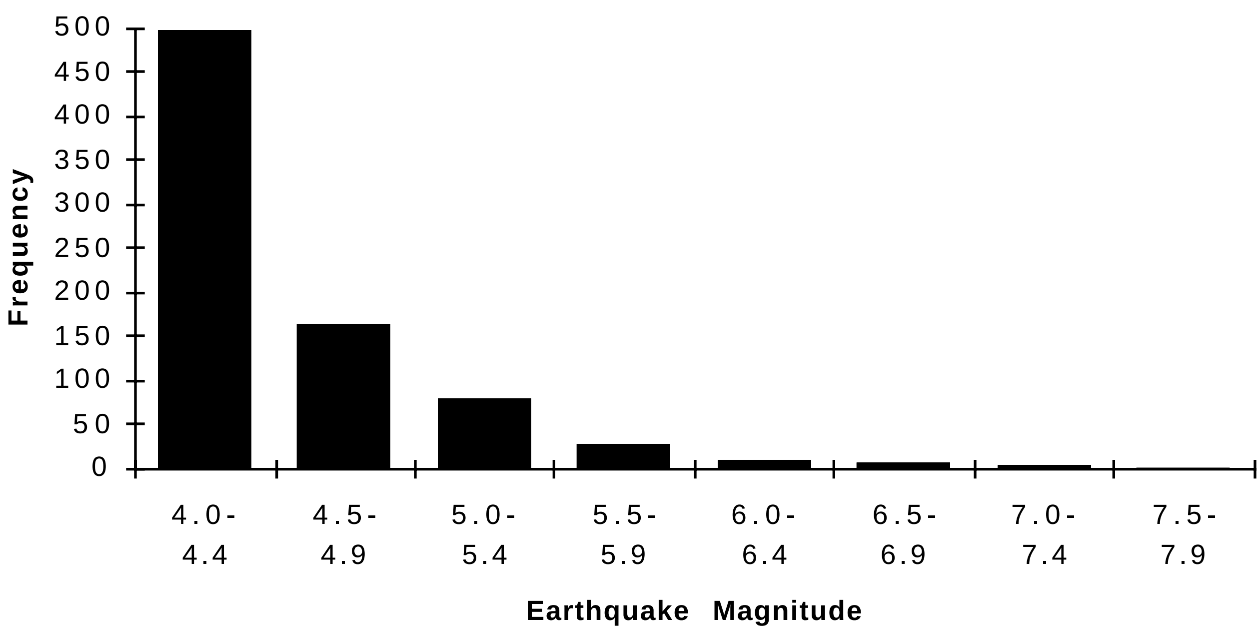
Earthquake Data by Magnitude

Line Nos.	Magnitude	n	Mean	St. Dev.	Med.	Mode	Avg. Depth (km)
2 to 2	7.5 - 7.9	1	7.6	0	7.6	7.6	1
3 to 5	7.0 - 7.4	3	7.06	0.06	7.1	7.1	13
6 to 11	6.5 - 6.9	6	6.73	0.12	6.75	6.8	14.3
12 to 20	6.0 - 6.4	9	6.16	0.13	6.2	6.3	10.6
21 to 46	5.5 - 5.9	26	5.6	0.13	5.6	5.5	8.1
47 to 124	5.0 - 5.4	79	5.18	0.14	5.2	5	8.4
125 to 286	4.5 - 4.9	162	4.67	0.15	4.6	4.5	8
287 to 783	4.0 - 4.4	497	4.15	0.14	4.1	4	7.6
2 to 783	Total Data	782	4.47	0.54	4.3	4	7.9

Earthquake Data by Region

Line Nos.	Location	n	Mean	St. Dev.	Med.	Mode	Avg. Depth (km)
1 to 17	Baja	17	4.3	0.31	4.2	4	10.3
18 to 230	Southern Cal.	213	4.49	0.55	4.3	4	6.1
231 to 331	Central Cal.	101	4.41	0.45	4.2	4	6.6
332 to 454	Cal./Nev. Border	123	4.48	0.51	4.3	4	4.8
455 to 582	Northern Cal.	128	4.49	0.67	4.2	4.2	11.7
583 to 729	Oregon	147	4.44	0.52	4	4	9.5
729 to 782	Wash./Vanc. Is.	53	4.58	0.51	4.1	4.1	10.4
1 to 782	TOTAL DATA	782	4.47	0.54	4.3	4	7.9

Distribution of "Felt" West Coast Earthquakes
January 1, 1990 through July 11, 1996 (N = 782)



quakes of different magnitude ranges (7.5-7.9, 7.0-7.4, etc.) shows an intuitive yet stunning example of exponential decline.

The data pictured in the chart above suggest that the total number T of earthquakes of magnitude M is related in an exponentially decreasing fashion. In the language of algebra,

$$T = Ab^M,$$

for some constant A and base b. Although earthquake magnitudes are measured in base ten, we found that using the base 7.1 gives a better "fit" to the data and therefore might be a better model for west coast earthquake frequencies. The data clearly show that big earthquakes (6.0 or larger) represent a fraction (18/782, or about 2.3 %) of all "felt" earthquakes over the past six years. Thus, while "The Big One" is possible, it is unlikely, and certainly does not merit the hysteria portrayed by the media every time an earthquake occurs.

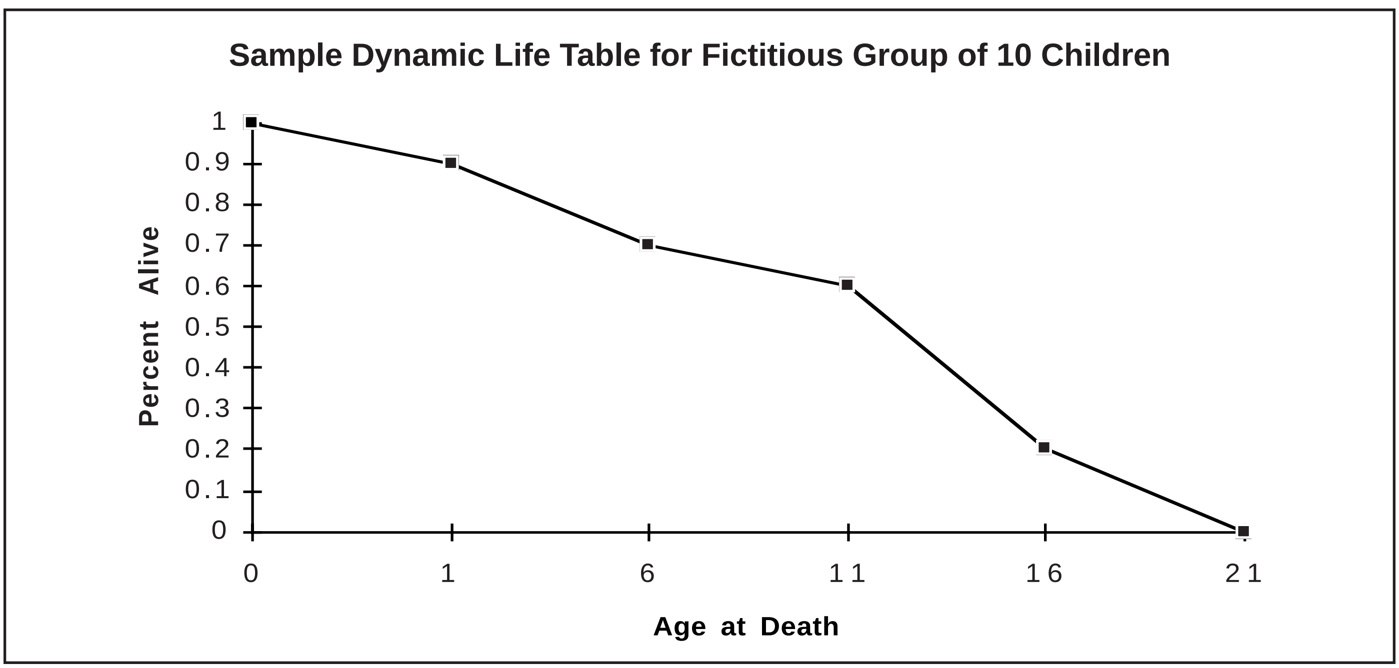
The earthquake data are a scientifically documented, accessible source of information that gives insight into understanding earthquake issues. In exploring this data, we dealt with many aspects of arithmetic and advanced algebra as well as incorporating technology in an integral way. The mathematics was powerful because it occurred in the context of a situation that was tied to the teachers' individual and collective experience.

MATH FROM THE CRYPT: INVESTIGATING THE PAST AT ST. PETER'S CEMETERY IN OXFORD, MISSISSIPPI

Description of Project

Demography is the study of the age structure and growth rate of populations. The *life table* is one way of summarizing key demographic variables, including age-specific mortality, survivorship, and expectation of further life. Once these data are compiled, we can use them to investigate demographic patterns and processes, such as differences in the survival rate or life expectancy of different groups of organisms.

The simplest way to construct a life table is to follow a group (or *cohort*) of organisms from birth, recording the age at which each individual dies, until all individuals of the original cohort have died. The result of this approach is termed a *dynamic life table*. However, cohort data are difficult and time-consuming to obtain, because the table cannot be completed until the entire cohort has died - which could take decades, in the cases of elephants or seabirds, or even centuries, as for trees such as bristlecone pines (which may live 2,500 years). Consequently, ecologists often construct life tables using other types of information. The approach we used was to gather data on the age of death of a sample of individuals, and to use these data to estimate mortality rates and to calculate other vital statistics. This approach yields a static life table, with entries that are age-specific, even though the sample is a composite, made up of individuals who started life at different times.



In this project we focused on human demography, in part because of our obvious interest in understanding the patterns and causes of death among people, but also because the data are readily available - thanks to our cultural tradition of memorializing our deceased relatives, and information about their lives, on gravestones and tombs. During our visit to the cemetery participants recorded information from gravestones. Data were separated by century of birth, sex, race, or other variables, depending on the question(s) on which teams wanted to focus. Questions which teams were addressing included the following:

•What is the general shape of the survivorship curve for your various datasets?

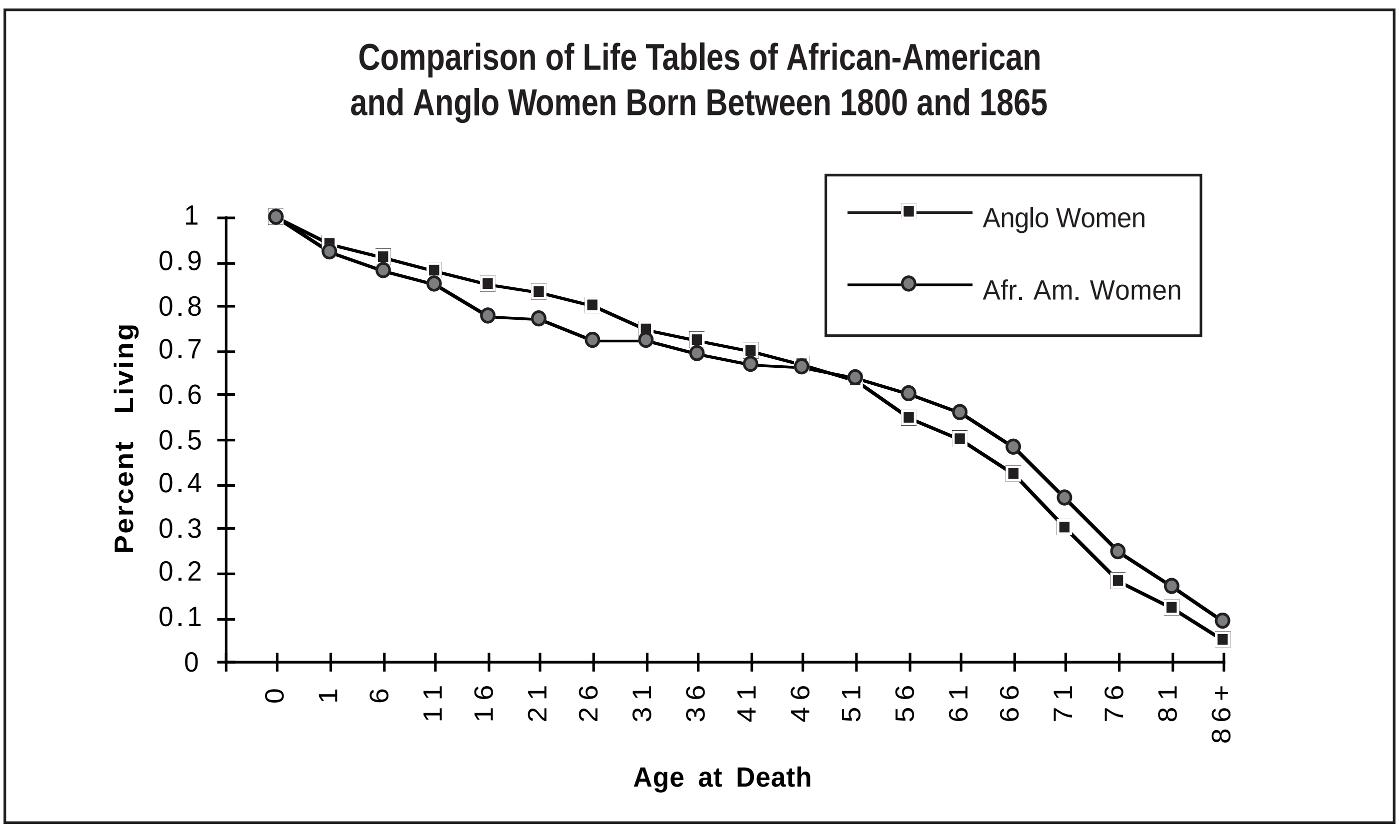
•Is there any major differences between the survivorship curves or life expectancy for people born before 1800 vs. those born in the 1800s vs. those born in this century? What biological (including medical) changes might account for any differences?

•Do the survivorship curves or life expectancies for men differ from those for women? Are the differences, if any, consistent from one century to the next? What biological factors might account for any differences?

•Do any of the datasets show marked differences compared to the recent life table for the United States population? What are the most obvious differences, and how might you explain them?

•The answers to the preceding questions might be erroneous if our data did not accurately represent the demography of people in any of the time periods. What types of biases, if any, can you envision and how might they skew the results (as well as affect our responses to the questions)?

The technique of creating life tables is a straightforward application of cumulative percentages. After recording birth and death dates, sex (inferred), and ethnicity (when known) for a sample, a simple tally is taken. Using the standard interval groupings of five years (after the first year of life), a small sample of ten children deceased before age 21 might show this: one person died at birth, two made it past their first birthday but not to 6, one person lived past six but died before age 11, four lived past eleven but died before 16, and two lived past sixteen but died before age 21. Thus, at age one, 9 people (90 %) were still alive; at age six, 70 % were alive; at age eleven, 60 % were alive; at age sixteen, 20 % were alive, and by age 21, 0 %



were alive. Thus, a graph showing survivorship rates for each age interval using this fictitious sample would be represented as shown on the previous page.

Real Data: An African Legacy of Strength

Since St. Peter's cemetery dates back to the 18th century, there were a number of comparisons that teams could make, such as comparing life tables by sex, by century, and even by race (it turned out that Section 5 of the nine-section cemetery was a pre-civil war slave section). Although some of the markers were unreadable or even absent, many of those could be identified by taking a rubbing.

Independent samples showed two findings that were consistent. First, both African-American and Anglo women outlived their male counterparts during both the 19th and 20th centuries, a result that is consistent with current lifetables for all races. Second, and perhaps less intuitive, African-American women outlived Anglo women across during both centuries, including during the age of American slavery. The graph below shows data collected by a team of African-American teachers illustrating the life tables of black women and white women born between 1800-1865. Note that the African-American group shows a marked decrease during the late teens and early twenties (probably attributed to childbirth issues), but shows a strong survivorship after age 50 and on into old age. The result depicted in this graph was corroborated by two other groups using independent samples, indicating the validity of this finding, namely, that African-American women showed stronger survivorship than their Anglo counterparts despite their status as slaves. The group whose data are shown here presented their results entirely on the computer, using overhead graphics. The group was impressive not only in its use of technology, but in its understanding of what it had found. The work of these teachers commanded the respect and admiration of all of their colleagues in the institute.

SUMMARY

Using real data as a catalyst to explore mathematics has made a lot of sense to me. Both student and teacher are partners in trying to figure out what the data mean, and which, if any, mathematical models might be useful tools to make predictions. Of course, the models are far from perfect; indeed, part of the problem that scientists face is to decide which type of equation is appropriate, and over what interval is it valid. Students will disagree on solutions as well, causing a certain level of *angst* for both them and the teacher in regards to grading.

Nonetheless, I have tried to incorporate such problems into my calculus class this semester. It would be untrue to say that there have not been drawbacks. First, it takes a lot more preparation time for me to find the data sets and incorporate them in a useful and appropriate way. Second, I have less control over what the students actually learn from these types of problems, since there is often no clear answer (or even question, for that matter). And third, it takes class time away from other, more traditional activities, such as my lecturing on textbook material.

Both student and teacher are partners in trying to figure out what the data mean, and which, if any, mathematical models might be useful tools to make predictions.

Paired with each of these concerns, though, is a benefit. First, *I* am more engaged in *thinking* about the calculus than I have ever been in the past. I have been especially struck by the importance of viewing a function as continuous, in which case the rules for derivatives and integrals apply, or discrete, so that average rather than instantaneous rates of change make sense. Second, it seems that my students have done more thinking about calculus on their own, based on their written projects, than have students in past classes, based on less thoughtful responses to original application questions. And third, scores on mechanics-based exams involving derivatives and integrals have been at least as high as those from past years, even though I have spent less class time lecturing on and going over these processes. Perhaps the greatest benefit, though, is that (hopefully) most of the students in this course will have a pretty good answer the next time someone asks them, "When am I ever gonna use this stuff?"