AGAINST ALL ODDS: INSIDE STATISTICS

Marsha Davis
Eastern Connecticut State University
Mathematics Department
83 Windham Street, Willimantic, CT 06226
davisma@easternct.edu

Abstract

The new online Against All Odds (AAO) series is guaranteed to breathe life into your introductory statistics course. Picking up were the original AAO left off in the 1980s, the new series maintains the same emphasis on "doing" statistics. Video modules take students on location to where people from all walks of life are using statistics in their work, while online materials allow students to practice and review what they have learned. This paper introduces sample materials that include descriptions of a video segment, interactive applets and activities designed to actively engage students, and exercises that involve real-world data and use of technology.

Introduction

Over the years, the number of sections of *Introductory Statistics* has greatly increased. Students enrolled in *Introductory Statistics* have diverse backgrounds, interests, and reasons for taking the course. All too often *Introductory Statistics* is viewed as unpleasant, difficult, and/or boring by students. Given the importance of statistics and statistical reasoning in an increasingly complex and information-rich world, ways must be found to engage students with real-world contexts and activities that support learning the basic elements of statistical thinking and the important concepts that underlie statistical reasoning, particularly concepts that students find difficult. The *Against All Odds: Inside Statistics* (*AAO*) series, funded by Annenberg and the Consortium for Mathematics and Its Applications (COMAP), is designed to do just that.

AAO is an online resource (www.learner.org/courses/againstallodds) that consists of 30 modules on topics from descriptive statistics, probability, and inference. Each module consists of a video, an activity (some of which use interactive applets), written support material, and exercises that involve real-world data and use of technology. These online materials can be used either as the only course materials or as supplements to other course materials. The videos, activities, interactive applets, exercises and support materials provide a rich learning environment that helps statistics come alive. While *AAO* is designed to support a standard general-education-requirement introductory statistics course, with careful selection and a bit of creativity, *AAO* materials can be adapted for use with a wide range of students. For example, some materials from the early modules could be adapted for use with middle school students, while other materials could be adapted for use in a statistics course for mathematics majors, or for teacher professional development (particularly preparing teachers for the statistics and probability content in the *Common Core State Standards for Mathematics*).

Materials

This paper will describe a sample video segment showing statistics in action, present one of the interactive applets and several activities that rely on that applet, and share sample exercises based on real datasets. These materials can be accessed at:

www.learner.org/courses/againstallodds

Two-Way Tables Video: The Happiness Survey

Everywhere we turn, we find surveys. Watch the local news and we get asked to respond to a survey. Buy an item online, and we get asked to complete a questionnaire. Analyzing categorical survey data one question at a time is easy, but it is also not very interesting. The interesting results come from investigating the relationships between the responses to two (or more) questions – in other words, from analyzing data that can be organized into two-way tables. However, college students (even mathematics majors!) struggle with conditional percentages. For example, on a question regarding political affiliation, students often have trouble differentiating between the percentage of women who are Democrats and the percentage of Democrats who are women. Prior to introducing this topic formally, the video described below can help students get a handle on marginal and conditional distributions.

The context for the video is the Happiness Survey that was part of Somerville, Massachusetts' 2011 annual census. The video focuses on two of the survey questions:

- How happy do your feel right now?
- How would you rate the beauty or physical setting of Somerville?

For the video, Happiness ratings are boiled down into three categories: Unhappy, So-So, and Happy. Ratings of Somerville's physical beauty are categorized as Bad, OK, and Good. Participants' responses to these two questions are organized into the two-way table shown in Table 1.

		Physical Beauty			
		Bad	OK	Good	Total
Happiness	Unhappy	90	123	62	275
	So-so	555	972	610	2137
	Happy	541	1426	1406	3373
Total		1186	2521	2078	5785

Table 1. Results from rating happiness and Somerville's physical beauty.

From here students can easily find marginal distributions of Happiness and Physical Beauty. For example, a majority (58%) of Somerville participants responded that they were happy while slightly less than 36% rated Somerville's beauty as good.

The more interesting question, a question that cannot be answered by the marginal percentages, is whether happy people have a more positive view of Somerville's physical beauty than unhappy people. For that we need to compute the conditional distributions of Physical Beauty for each level of the Happiness variable (Table 2).

		Physical Beauty			
		Bad	OK	Good	Total
Happiness	Unhappy	32.73%	44.73%	22.55%	100%
	So-so	25.97%	45.48%	28.54%	100%
	Happy	16.04%	42.28%	41.68%	100%

Table 2. Conditional distribution of Physical Beauty for each Happiness category.

From Table 2, we discover that 42% of happy people rated Somerville's beauty as good, compared to only 23% of unhappy people. Clearly there is a connection between Happiness and Physical Beauty. Upon further inspection, we note that as the level of Happiness goes up, the percentage of Good ratings goes up and the percentage of Bad ratings goes down. The bar chart shown in Figure 1 effectively illustrates these patterns.

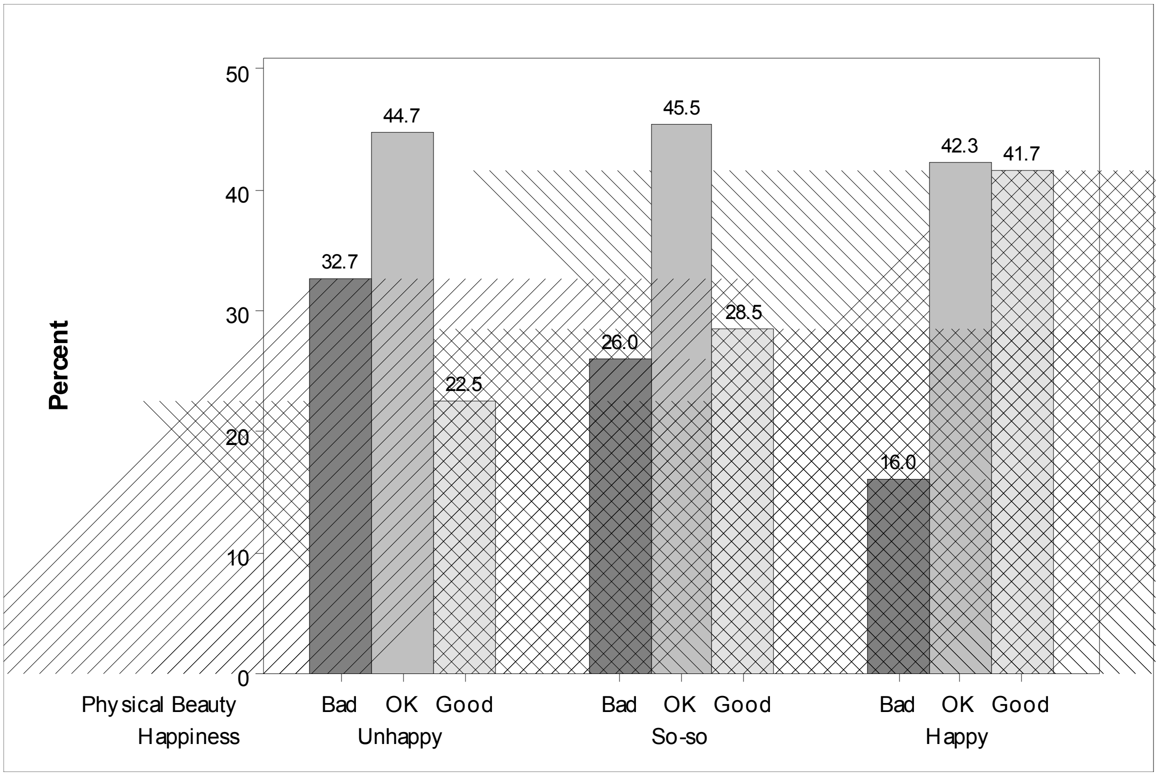

Figure 1. Conditional distribution of Physical Beauty for each level of Happiness.

Now that Somerville has identified a connection between resident's happiness levels and their ratings of the city's physical beauty, officials want to dig more deeply into this connection for the next survey in an effort to improve resident's satisfaction with Somerville.

The activity for Unit 13, *Two-Way Tables*, is directly connected to this video. Students are asked to conduct their own Happiness survey. I have done this on the first day of class for the last two years. In addition to asking students to rate their level of happiness on the first day of class and to rate the beauty of our campus, I have also asked them to give their year (Freshman, Sophomore, Junior, Senior) and gender, which has produced some interesting results. In my introductory statistics course for mathematics majors, it appears that a higher percentage of juniors are happy on the first day of class compared to seniors. Also, as is true of the Somerville survey, happier students are more likely to rate the physical beauty of campus as good compared to unhappy students.

Wafer Thickness Interactive Applet

Over 20 years ago, Cobb (1991) provided a direction for improving the teaching of *Introductory Statistics*:

> Almost any course in statistics can be improved by more emphasis on data and concepts, at the expense of less theory and fewer recipes. To the maximum extent feasible, calculations and graphics should be automated. Any introductory

course should take as its main goal helping students to learn the basics of statistical thinking.

Moore (1990) put understanding variability and appropriate ways to quantify and model variability at the core of his description of statistical thinking. However, students in introductory statistics courses usually focus on describing central tendency and neglect variability (Gould, 2004). Activities based on *AAO*'s Wafer Thickness interactive applet can help address this problem.

The Wafer Thickness interactive simulates measuring the thickness of polished wafers used in the manufacturing of microchips. Figure 2 shows a screen capture of the interactive in action. For this screenshot, the second of a sample of five polished wafers is being measured and the results are posted in real time. The interactive posts the thickness measurement of each wafer and simultaneously creates a histogram of the data as they become available. The interactive applet allows students to compare histograms of up to three samples of data. In addition, the data can be saved in a csv file for export into Excel or statistical software so that students can conduct more sophisticated analyses.

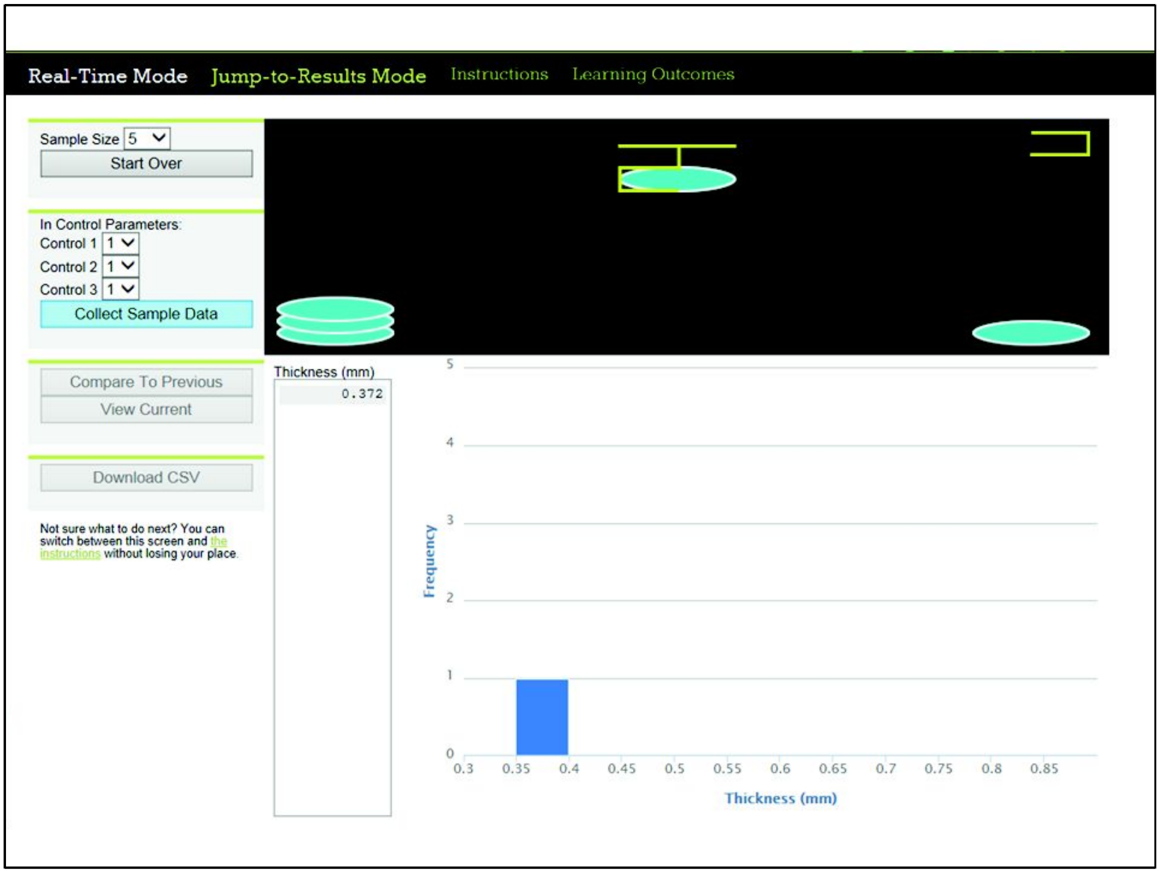

Figure 2. Screen shot of Wafer Thickness interactive applet.

<u>Introductory Activity: Determining the Optimal Control Settings</u>

The first activity that relies on this interactive appears in Unit 3, *Histograms*, which is early in an introductory statistics course. The activity focuses on quality control in the production of the polished wafers. Using the interactive, students can set three controls at three different levels. These controls affect the thickness distribution of the polished wafers. The final task asks students to make a recommendation for the control settings so that the product is consistently close to the target thickness of 0.5 mm. This activity has been class tested in a variety of settings: introductory statistics courses for both math and non-math majors, in-service course for teachers at the graduate level, and with middle school students. Below are some observations on the results:

- Students were totally engaged in the activity. Each group had their own data. The group data varied from group to group.
- Students understood how to make a histogram from watching the construction of histograms in real time as data were collected. This reduced the class time needed to cover the construction of histograms and descriptions of their shape.
- Students observed that in a sample of wafers produced under the same control settings, wafer thickness varied from wafer to wafer.
- Students discovered that the distributions of wafer thicknesses differed from sample to sample even when produced under the same control settings.
- Students were able to observe variation due to the control settings even in the presence of variability within each sample and from sample to sample.
- In order to compare wafer thickness under different control settings, students created informal measures of both center and spread.
- Students discovered from class discussion that different groups came to different conclusions for the final task. It was clear that more needed to be learned before a more consistent recommendation across all groups could be reached. Students bought into the activity and started the semester ready to learn.

Follow-up Activities

In one follow-up activity, students imported the data into a statistics package. Then they created their own graphic displays and computed numeric summaries to add to their descriptions of the data. Figure 3 shows a graphic display of comparative boxplots created by one group.

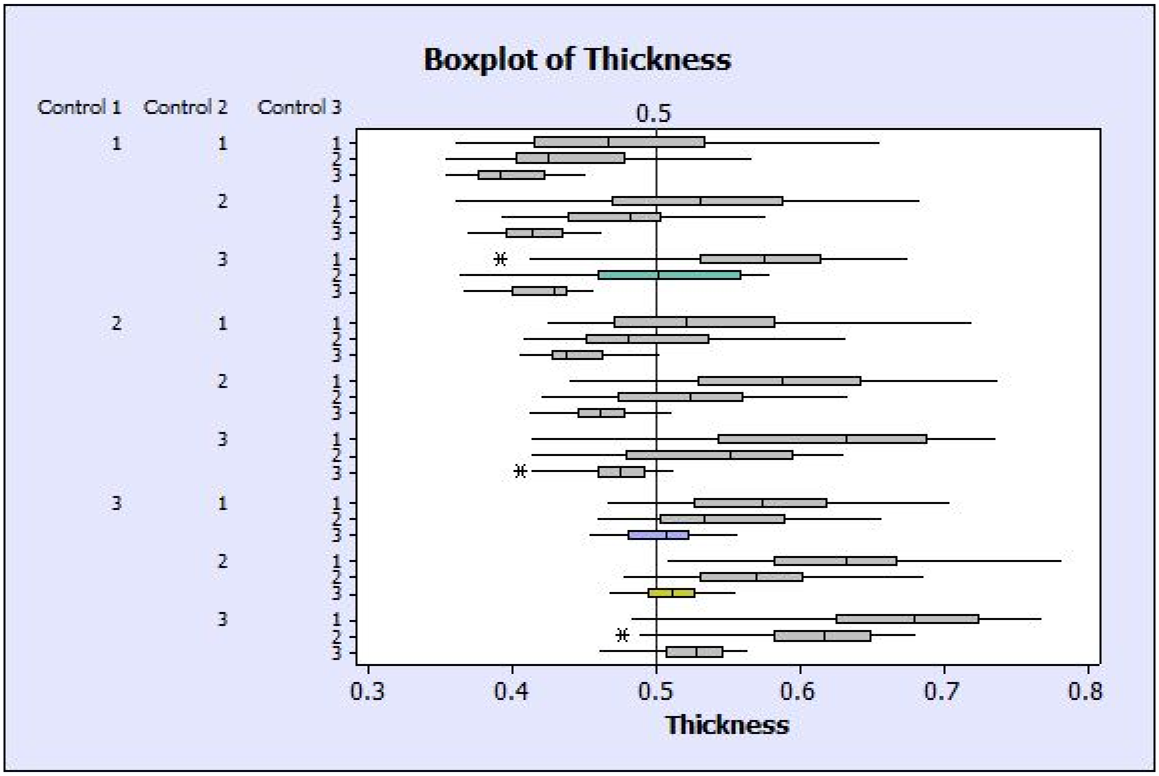

Figure 3. Graphic display of wafer thicknesses under all possible control settings.

Even with this added level of information, there were still some disagreements among groups about the best recommendation. Thus, students discovered that there is not always a "best" solution. In this case, there were several good solutions and the "best" solution depended on whether the group thought it was better to be closest to "on target" or were willing to trade being a little "off target" for some reduced variability.

In the math major course, this activity was extended even further to include a theoretical solution based on the underlying triangular distributions that governed the applet's data generation:

$$f(x \mid a,b,c) = \begin{cases} 0 & x < a \\ \dfrac{2(x-a)}{(b-a)(c-a)} & a \le x \le c \\ \dfrac{2(b-x)}{(b-a)(b-c)} & c < x \le b \\ 0 & x > b \end{cases}$$

- Control 1 changed the settings for *a*: (1) 0.35, (2) 0.40, (3) 0.45.

- Control 2 changed the settings for $(c-a)/(b-a)$: (1) 0.20, (2) 0.50, (3) 0.80.
- Control 3 changed the settings for $(b-a)$: (1) 0.35, (2) 0.25, (3) 0.12.

Students then computed the theoretical mean and standard deviation for the 27 possible triangular distributions and again were faced with the tradeoff of selecting the control settings associated with having the mean closest to the target or trading a little bit of bias for reduced variability.

ANOVA Activity

The Wafer Thickness applet produces great data for a project on one-way ANOVA, a topic typically covered at the end of a semester or in the second semester of a two-semester sequence. ANOVA is particularly difficult for students to grasp conceptually because here we establish a difference in means by analyzing variation. In this activity, the question is whether mean wafer thickness differs depending on control levels for each control separately. The activity is structured as three experiments. In Experiment 1, control 1 is varied from level 1 to 2 to 3 while controls 2 and 3 are set at level 2. In Experiments 2 and 3, controls 2 and 3, respectively, are varied while the other two controls are held fixed at level 2. For each experiment, samples of size 10 are collected in Real-Time mode so that students can watch the data being collected. The data can be exported for analysis with spreadsheet or statistical software. Because the applet generates random data, each student (or group of students) works with different data. However, results are generally similar:

- The mean thicknesses of wafers produced under the three settings of control 1 differ significantly.
- The mean thicknesses of wafers produced under the three settings of control 2 do not differ significantly.
- For control 3, the underlying assumption of equal standard deviations for ANOVA is not satisfied.

After individuals or groups have completed the activity, there is opportunity to share results. Students can observe that the same experiment results in different values for the F-statistic. In most cases, the conclusions turn out to be the same. However, it is possible that one individual/group's conclusions could differ from the majority due to sampling variability.

Exercises/Review Questions

In addition to videos and activities, the online materials include an overview of each topic along with exercises and review questions that give students an opportunity to practice what they have learned. Generally exercises and review questions are based on real data. An example of two review questions from Unit 11, *Fitting a Line to Data,* based on the data in Table 3 follows.

1. Table 3 contains data on mercury concentration in tissue samples from 20 largemouth bass taken from Lake Natoma (California). Only fish of legal/edible size were used in this study.

a. We want to be able to predict mercury concentration from fish length. Which variable is the explanatory variable and which is the response variable?

b. Fit a least-squares line to the data from Table 3. Report its equation. (Round the slope and y-intercept to four decimals.) Also, show a scatterplot of the data along with a graph of the least-squares line.

c. Make a residual plot. Based on your plot, is the least-squares model adequate to describe the overall pattern in these data? Explain.

d. Interpret the slope and y-intercept of the least-squares line in the context of this problem. Do these make sense in the given context? Explain why or why not.

2. Return to the data in Table 3. Use your answer to 1(b) to make the following predictions:

a. Predict the mercury concentration in a largemouth bass that is 430 mm in length. Is this prediction an example of interpolation or extrapolation? Explain.

b. Predict the mercury concentration in a largemouth bass that is 90 mm in length, which is below the legal/edible size. Is this an example of interpolation or extrapolation? Explain.

Total Length (mm)	Mercury Concentration (µg/g wet wt.)
341	0.515
353	0.268
387	0.450
375	0.516
389	0.342
395	0.495
407	0.604
415	0.695
425	0.577
446	0.692

490	0.807
315	0.320
360	0.332
385	0.584
390	0.580
410	0.722
425	0.550
480	0.923
448	0.653
460	0.755

Table 3. Fish length and mercury concentration in fish tissue samples.

Conclusion

The new online *Against All Odds*: *Inside Statistics* series will breathe life into any introductory statistics course. Providing students with videos showing statistics in action and activities in which students can collect and analyze their own data can engage students and support their learning of difficult statistical concepts. The *AAO* written materials include summaries of statistical techniques/concepts as well as exercises/review questions, which provide opportunities for students to review and put into practice what they have learned.

References

Cobb, George W. (1991). "Teaching statistics: more data, less lecturing," *Amstat News*, December 1991, pp. 1, 4.

Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. www.corestandards.org/Math.

Garfunkel, S. (Project Manager) & Davis, M. (Chief content Advisor/Writer), & Villiger, M. (Writer/Producer/Director). (2013). *Against All Odds: Inside Statistics*. Washington, DC: Annenberg Learner. www.learner.org/courses/againstallodds/.

Gould, R. (2004). Variability: One Statistician's View. Statistics Education Research Journal, 3(3), pp. 7 – 16.

Moore, D. (1990). Uncertainty. In L. Steen (Ed.). On the Shoulders of Giants: New Approaches to Numeracy, pp 95 – 138. Washington. DC: National Academy Press.