A Baseball Statistics Course

Jim Albert
Bowling Green State University

Journal of Statistics Education Volume 10, Number 2 (2002),
ww2.amstat.org/publications/jse/v10n2/albert.html

Copyright © 2002 by Jim Albert, all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the author and advance notification of the editor.

Key Words: Ability; Measures of batting performance; Situational statistics; Spinner probability model; Sports; Streakiness.

Abstract

An introductory statistics course is described that is entirely taught from a baseball perspective. Topics in data analysis, including methods for one batch, comparison of batches, and relationships, are communicated using current and historical baseball data sets. Probability is introduced by describing and playing tabletop baseball games. Inference is taught by first making the distinction between a player's "ability" and his "performance", and then describing how one can learn about a player's ability based on his season performance. Baseball issues such as the proper interpretation of situational and "streaky" data are used to illustrate statistical inference.

1. An Introductory Statistics Course

Our department offers a one-semester introductory statistics course. This course satisfies the mathematics elective for students majoring in the College of Arts and Sciences and is also required by students in the health college. The general goal of this course is to explain the discipline of statistics and describe in a general way how one draws conclusions from data. The topics of the course include data analysis for one and two variables, elementary probability, and inference for proportions and means.

There are many difficulties and concerns in teaching an introductory statistics course, some of which are listed below:

- It's a required "math" course that few students want to take. Many students are fearful of taking it because they are not comfortable with their mathematical and computational ability.

- Many introductory statistics courses focus on computation and skills instead of the important concepts.
The lecture format in teaching is not conducive to learning statistics.

Students have little interest for the topics and data sets that are discussed in a statistics course.

There is currently a reform movement in the instruction of introductory statistics. Many statistical educators believe that:

- There should be more emphasis on data analysis and less emphasis on topics in probability (Moore 1992).

- There should be less time devoted to lectures and more time spent on active learning by means of directed activities in the classroom, activities in a computer lab, and projects where the students do various parts of a statistical investigation (Hogg 1992).

- There should be more emphasis on concepts and statistical reasoning, and less focus on computation and formulas (Moore 1992).

- The course should be made more relevant to the students by emphasizing connections with everyday life. The Chance course (Snell and Finn 1992) is an excellent illustration of a course that is driven by current events that are reported in the media.

Hogg (1992), summarizing a workshop on statistical education held at Iowa City, discusses several poor characteristics of science and mathematics education. He comments (p. 4) that mathematics and science courses "are not 'fun' because we fail to communicate our enthusiasm and excitement about mathematics and science." Commenting on introductory statistics teaching (p. 6), the workshop participants mention that statisticians "often fail to see any need to convey a sense of excitement."

Many authors discuss the need for statisticians to focus their teaching on the wealth of statistical applications. Willett and Singer (1992, p. 83) state that “learning applied statistics can be made more interesting ... (if we can) ... capitalize on students’ fascination ... for the substantive problems that statistics can address.” These authors describe eight attributes that they believe enhance a data set’s “instructional suitability.” The best data sets:

1. come in raw form,
2. are authentic,
3. include background information,
4. have case-identifying information,
5. are intrinsically interesting,
6. are topical or controversial,
7. offer substantive learning, and
8. lend themselves to a variety of statistical analyses.

Mosteller in Moore (1993, paragraph 34), comments about using data exploration to teach statistics: “I believe that students are very interested in findings from the data and are willing to work hard on it, and so I think data-oriented statistical teaching is a good idea. I have written a book with colleagues on statistics for physicians, and it tries to orient itself toward teaching the course from the point of view of the problems that physicians have - problems of diagnosis, problems of treatment, problems of different dosage levels, problems of tests and the conflicts between tests that are carried out. ... So that course is oriented in a different way from our usual statistics course which tends to teach about statistical topics such as means and variances and regression and analysis of variance. It's more oriented toward the way
the practical people in the field think about the subject matter that they're working with.”

Sowey (1995) talks about the characteristics of a statistics course that makes learning last. He comments on how an instructor can make the student see the “worthwhileness” of the discipline of statistics. The enthusiasm of the teacher and the student’s own discovery of the subject lead to intellectual excitement. Also, the worthwhileness of the discipline can be seen by demonstration of the practical usefulness of statistics. Yilmaz (1996) and Zetterqvist (1997) also discuss how to make the introductory statistics course more effective by linking statistics and real-world situations.

2. Sports and Baseball

Because many college students are interested in sports, either as observers or participants, it seems natural to base a statistics course on data and the associated investigations from various sports. Many students should have backgrounds in the various sports, so they may be better able to understand the statistical concepts, as they are set within the familiar context of sports.

Why did I decide to focus my special statistics course on baseball instead of other sports? First, baseball is the great American game. The game developed in America about 150 years ago, and it is played today using essentially the same rules as in the early days. Second, many students are familiar with the game. Although students may not be familiar with the various baseball statistics, they are familiar with the basic rules of the game and likely have attended some baseball games. Baseball also has a great historical tradition. There are many famous teams and players that one can talk about in a class. Finally, more than any other sport, baseball can be described by the associated statistics.

How is baseball a statistical game? Players (both batters and pitchers) are evaluated by means of their statistics. When a batter comes to bat during a television broadcast, his statistics are flashed on the screen. TV and radio broadcasters routinely use statistics in their discussions. Some of these statistics are announced with the intention of entertaining the audience. Other statistics are used by the broadcasters to make a particular argument regarding the quality or lack-of-quality of a team or a player. More importantly, a player's statistics are used to make decisions about salary, to decide whether to keep or drop a particular player, or to make a trade with another team. Many great players are defined by their associated great statistics. All baseball fans know of Babe Ruth's 60 home runs in 1927, Roger Maris' 61 home runs in 1961, Mark McGwire's 70 home runs in 1998, and Barry Bonds 73 home runs in 2001. Likewise, Bob Gibson is famous for his unusually low 1.12 ERA in 1968, and the "great streak" refers to Joe DiMaggio's 56-game hitting streak in 1941. Baseball has a relatively discrete structure that makes it easy to model probabilistically. A basic event is the result of the confrontation between batter and pitcher, and one can simulate this event by use of dice or spinners.

3. The Baseball Statistics Course

One special section of the introductory statistics course was advertised as a “baseball statistics" course. This section was opened to all students who had an interest in baseball. In the first semester (Fall 2000), 30 students enrolled - 24 were male and 6 were female. Because the material for this course was being developed this academic year, there was no textbook and the course was lecture-driven. Copies of the lecture notes were made available over the class Web site. Homework assignments were given from a special workbook that was written by the instructor. The course grade was determined by three in-class tests and homework assignments.

Every class focused on the analysis of a particular baseball data set and the statistical methods and concepts were discussed in the context of the particular data set. In the next three sections, we outline a
sample of these lectures presented in the three general areas of data analysis, probability, and inference. For each lecture, we focus on the data set and the corresponding questions that would motivate a particular statistical concept or method. (Please contact the author for information about an extensive set of case studies and exercises from baseball that can be used in teaching topics in data analysis, probability, and statistical inference.)

4. Lectures in Data Analysis

"My Tribute to Richie Ashburn" (Data Analysis for One Batch)

This lecture focuses on the baseball data that are found on the back of a usual baseball card - the season hitting or pitching statistics for a particular player. Because the instructor is a Phillies fan, the class looked at the hitting statistics, shown in Table 1, for Richie Ashburn, a member of the Whiz Kids, who was recently inducted in the Hall of Fame.

Table 1. Career Batting Statistics for Richie Ashburn.

| Year | Team | G  | AB  | R  | H  | HR | AVG | SLG | OBP |
|------|------|----|-----|----|----|----|-----|-----|-----|
| 1948 | PHI  | 117| 463 | 78 | 154| 2  | .333| .400| .410|
| 1949 | PHI  | 154| 662 | 84 | 188| 1  | .284| .349| .343|
| 1950 | PHI  | 151| 594 | 84 | 180| 2  | .303| .402| .372|
| 1951 | PHI  | 154| 643 | 92 | 221| 4  | .344| .426| .393|
| 1952 | PHI  | 154| 613 | 93 | 173| 1  | .282| .357| .362|
| 1953 | PHI  | 156| 622 | 110| 205| 2  | .330| .408| .394|
| 1954 | PHI  | 153| 559 | 111| 175| 1  | .313| .376| .442|
| 1955 | PHI  | 140| 533 | 91 | 180| 3  | .338| .448| .449|
| 1956 | PHI  | 154| 628 | 94 | 190| 3  | .303| .384| .385|
| 1957 | PHI  | 156| 626 | 93 | 186| 0  | .297| .364| .392|
| 1958 | PHI  | 152| 615 | 98 | 215| 2  | .350| .441| .441|
| 1959 | PHI  | 153| 564 | 86 | 150| 1  | .266| .307| .362|
| 1960 | CHI  | 151| 547 | 99 | 159| 0  | .291| .338| .416|
| 1961 | CHI  | 109| 307 | 49 | 79 | 0  | .257| .306| .375|
| 1962 | NY   | 135| 389 | 60 | 119| 7  | .306| .393| .426|

In this lecture we focused on a single batting statistic - the on-base percentage (OBP). We graphed the OBP’s for Ashburn using a stemplot and discussed the variability present in this distribution of values. This discussion leads naturally to the concepts of center and spread of a batch. We might next look for a pattern in these OBP values across time. Most athletes mature in ability in the early stages of their career, hit a peak, and then deteriorate in ability towards the end of their career. Can we see this pattern in Ashburn’s OBP values when plotted against time? If we look further at both Ashburn’s OBP and slugging percentages (SLG), we might notice that Ashburn was essentially a singles hitter with relatively little power.

"Barry and Junior" (Comparing Batches)
This lecture compared two of the current great hitters in baseball, Barry Bonds and Ken Griffey, Jr. (Junior). A reasonable measure of batting ability is the OPS, which is equal to the sum of the player’s on-base percentage (OBP) and his slugging percentage (SLG):

\[
\text{OPS} = \text{OBP} + \text{SLG}
\]

(In fact, OPS stands for "On-base percentage Plus Slugging percentage.")

A useful graphical display to compare the season OPS’s for Barry and Junior in side-by-side stemplots as shown in Figure 1.

| BARRY OPS | JUNIOR OPS |
|-----------|-----------|
| 4 | 7 | 4 |
| 7 | 7 |  |
| 2 | 8 | 4 |
| 5 | 8 | 699 |
| 2 | 9 | 23 |
| 7 | 9 | 67 |
| 4300 | 10 | 222 |
| 877 | 10 | 7 |
| 3 | 11 |  |
| 5 | 11 |  |
| | 12 |  |
| | 12 |  |
| | 13 |  |
| 7 | 13 |  |

Figure 1. Side-by-side stemplots of the season OPS’s for Barry Bonds and Ken Griffey Jr. through the 2001 season.

The break point for each stemplot is between the tenth and hundredth places, so that

\[
8 | 699
\]

indicates that Junior had three OPS values .86, .89, and .89. This display indicates that Barry is generally a better hitter than Junior and we can compare medians to describe the difference in hitting. But both players are still active in baseball and Junior, being the younger player, likely will play more baseball seasons. So a fairer comparison might be to plot the OPS for both hitters against age. Figure 2 displays a scatterplot that shows that Junior performed better than Barry for young ages and Barry is doing exceptionally well in his 30’s.
"Great Batting Averages" (Standardization)

In this class, we discussed some great season batting averages in the recent history of baseball: Ted Williams (the last "400" hitter) hit .406 in 1941, Rod Carew hit .388 in 1977, George Brett hit .390 in 1980, and Tony Gwynn hit .394 in 1994. Was Ted Williams’ .406 really the best batting average among the four? Maybe or maybe not. To properly assess greatness, we need to look at each batting average in the context of the entire group of batting averages for that particular season. A standardized score 

$$z = \frac{AVG - \text{mean}}{\text{standard deviation}}$$

is a useful measure of relative standing of a player’s AVG. Here we see that Carew’s .388 corresponded to a z-score of 4.07 and Williams’ .406 average corresponded to a z-score of 3.82. So actually, Carew’s AVG had a higher relative standing and so one could argue that Carew’s accomplishment was more impressive.

“Measures of a Team’s Offensive Performance” (Relationships)

Probably the most-discussed issue among sabermatricians (the people who analyze baseball statistics) is how to evaluate the hitting accomplishments of a player. There are many count statistics that are recorded, such as hits, runs, doubles, and walks. How can we combine these basic statistics to obtain a good measure of batting performance?

The objective of batting is to produce runs and teams, not individuals, produce runs. So to evaluate different batting measures, one needs to look at team data. For the 2000 American League teams, Table 2 shows the runs scored per game (R/G) and four batting measures, the batting average (AVG), the on-base percentage (OBP), the slugging percentage (SLG), and the OPS (OBP + SLG) statistic.

Table 2. Batting statistics for the 2000 American League Teams.
We focus on the use of a single batting measure, say AVG, in predicting a team’s runs scored per game. To do this, we

- explore the relationship between AVG and R/G using a scatterplot
- use a least-squares line to describe the linear relationship between AVG and R/G
- use a mean squared error criterion to judge the goodness of the fit

We repeat this process for each of the four batting statistics. What one discovers is that the traditional batting average (AVG) is a relatively poor predictor of runs scored and the OBP and OPS statistics are better predictors of runs.

5. Lectures in Probability

"Big League Baseball" (Discrete Probability)

In this class, we introduce probability by first discussing its interpretation (relative frequency and subjective viewpoints) and then computing probabilities for simple random experiments. The dice game “Big League Baseball” provides a nice illustration of an experiment with equally likely outcomes. This game is played with three dice; one red and two white. The red die determines the pitch result as shown in Table 3.

| Team      | R/G | AVG  | OBP  | SLG  | OPS  |
|-----------|-----|------|------|------|------|
| Anaheim   | 5.33| .280 | .352 | .472 | .824 |
| Baltimore | 4.90| .272 | .341 | .435 | .776 |
| Boston    | 4.89| .267 | .341 | .423 | .764 |
| Chicago   | 6.04| .286 | .356 | .470 | .826 |
| Cleveland | 5.86| .288 | .367 | .470 | .837 |
| Detroit   | 5.08| .275 | .343 | .438 | .781 |
| Kansas City | 5.43| .288 | .348 | .425 | .773 |
| Minnesota | 4.62| .270 | .337 | .407 | .744 |
| New York  | 5.41| .277 | .354 | .450 | .804 |
| Oakland   | 5.88| .270 | .360 | .458 | .818 |
| Seattle   | 5.60| .269 | .361 | .442 | .803 |
| Tampa Bay | 4.55| .257 | .329 | .399 | .728 |
| Texas     | 5.23| .283 | .352 | .446 | .798 |
| Toronto   | 5.31| .275 | .341 | .469 | .810 |

Table 3. Result of rolling the red die in “Big League Baseball."
If the ball is put in play, then one rolls two dice to determine the play outcome. Table 4 shows the outcomes.

Table 4. Result of rolling the two white dice in “Big League Baseball.”

| First die | Second die |
|-----------|------------|
|           | 1 2 3 4 5 6 |
| 1         | Single Out Out Out Out Error |
| 2         | Out Double Single Out Single Out |
| 3         | Out Single Triple Out Out Out |
| 4         | Out Out Out Out Out Out |
| 5         | Out Single Out Out Out Single |
| 6         | Error Out Out Out Single Home run |

This game motivates many questions for discussion:

- What is the chance that a pitch will be a strike?
- What is the probability that a ball in play will be a home run?
- What is the probability that the player gets on base?
- If a player gets a hit, what is the chance that it is a home run?

These questions introduce the concepts of finding probabilities for equally likely outcomes, computation of probabilities for mutually exclusive events, and conditional probability. I am careful to distinguish a hitter’s plate appearance profile (what can happen at a plate appearance) from a hitting profile (what type of hits does the player get).

"All-Star Baseball"

Once the students get familiar with the “Big League Baseball” game, they realize that it has limitations and isn’t really a good model for baseball competition. There is no distinction between players of different abilities - each player has the same chance of hitting a home run. The “All Star Baseball” game is a more sophisticated game that allows for different batting abilities. Each batter is represented by a
spinner where the areas of the batting events on the spinner correspond to the probabilities of the
different events. A spinner for Mike Schmidt is shown in Figure 3.

Figure 3. Spinner for Mike Schmidt constructed using career hitting statistics.

Each student in the class was given the project for constructing a spinner for a famous player (in Fall 2000 we looked at all-time All Star lineups of American and National Leaguers; in Spring 2001, we considered the 1927 Yankees and the 1975 Reds). The student was asked to

- find the hitting statistics from his or her player on the Web
- find the probabilities of each plate appearance event (out, single, double, triple, home run, walk) for the player
- compute the size of the regions on the spinner for each event (to make calculations easier, we subdivided the spinner into 36 equal areas and found the number of areas for each event)
- make the spinner like a colorful baseball card with interesting statistics and pictures
  We concluded this example by playing out a spinner game using the spinners constructed by the students. We made this activity fun by singing songs (National Anthem and Take Me Out to the Ball Game) and eating Cracker Jacks.

6. Lectures in Inference

"Ability and Performance" (An Introduction to Statistical Inference)

When we played the spinner game in class, we observed an interesting result - the team that was predicted to win actually lost. That raises the question: Is there a distinction between a team’s ability and their actual performance? We describe an ability of a team or a player as the power or skill to play baseball, and the performance as the actual baseball playing that we observe from day to day. The batting ability, say ability to get on-base, of a particular player can be represented by means of a spinner where the size of the on-base region is equal to $p$. The size of this region corresponds to a player’s unknown probability of getting on-base. Although we don’t know a player’s batting ability, or value of
we can learn about his ability by watching him bat. This discussion motivates the construction of a confidence interval for the on-base probability $p$.

To illustrate confidence intervals and the use of these intervals to make decisions about parameters, suppose one is interested in comparing the on-base proportions of Barry Bonds and Sammy Sosa in the 2001 baseball season. The on-base proportion OBP is defined to be the fraction of times the player gets on-base - one computes this by dividing the number of times on-base (found by summing hits (H), walks (BB), and hit-by-pitches (HBP)) by the number of plate appearances (found by summing at-bats (AB), BB, HBP, and sacrifice flies (SF)). In the expression below, $X$ denotes the number of times the player got on-base, and $PA$ denotes the number of plate appearances.

$$\text{OBP} = \frac{H + BB + HBP}{AB + BB + HBP + SF} = \frac{X}{PA}$$

Table 5 shows the basic hitting statistics for Bonds and Sosa for the 2001 season.

| Player        | PA  | AB  | H   | BB  | HBP | SF | OBP |
|---------------|-----|-----|-----|-----|-----|----|-----|
| Barry Bonds   | 664 | 476 | 156 | 177 | 9   | 2  | .515|
| Sammy Sosa    | 711 | 577 | 189 | 116 | 6   | 12 | .437|

We see that Bonds had an OBP that was 0.078 higher than Sosa’s OBP, which is perceived by baseball fans to be a big difference in the two players’ on-base performances. But did Bonds have a greater ability than Sosa to get on-base? To answer this question, we can define two parameters $p_B$ and $p_S$ that represent Bonds’ and Sosa’s respective probabilities of getting on-base. Based on the 2001 season statistics, can one say with some confidence that $p_B$ is greater than $p_S$?

We can answer this question by the use of confidence intervals. Letting $\hat{p} = X / PA$ denote the observed on-base proportion for a player, the standard 95% confidence interval for the underlying probability is given by

$$\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{PA}}$$

Using this formula, we compute the 95% intervals for Bonds and Sosa to be

$$0.477 < p_B < 0.553 \text{ and } 0.400 < p_S < 0.474.$$  

These intervals are graphed in Figure 4. The intervals do not overlap, so one can draw the conclusion that Bonds had a greater ability to get on-base in the 2002 season. However, most baseball fans would
regard these interval estimates to be unusually wide. One thing that is learned from this example is that one really doesn’t have good knowledge about a player’s on-base probability from a single season of data.

Figure 4. 95% confidence intervals for Bonds’ and Sosa’s on-base probabilities based on 2001 season data.

"Making Sense of Situational Statistics"

After we discuss the basic notions of statistical inference, we discuss several interesting baseball inferential questions. One of the most interesting issues is how to interpret the popular situational or breakdown statistics that are available for all players. (Albert and Bennett 2001, Chapter 4.) If the player is a hitter, then we know how he hits during home games and away games, how he bats during each month of the season, how he bats on grass and on artificial turf, and how he bats against individual pitchers. Baseball fans and even baseball managers typically overstate the significance of these statistics - for example, a player might be benched for a game because he is 1 for 10 against the starting pitcher on the opposing team.

One basic data structure for situational statistics is the performance of a group of hitters in two mutually exclusive situations. For example, one could look at 20 hitters and find their on-base percentages (OBP) for home games and away games.

The first step in understanding the significance of situational statistics is to explore the data. The observed situational effect

\[ \text{OBP(home) - OBP (away)} \]

is found for all players. When we graph these situational effects, we see a number of interesting things. Particular players have very large and very small effects - are these interesting effects meaningful?

We see if these observed situational effects are meaningful by proposing some simple probability
models for situational data. If we have 20 players, then there are 20 hitting probabilities \( p_1, \ldots, p_{20} \), that represent the on-base abilities of the players. The question is how these hitting probabilities change across the home vs. away situation. One model would say that the “true” situational effect is nonexistent - the player will have the same on-base probability for home games and away games. A slightly more complicated model would say that there is a situational bias. Playing at home may increase the on-base probability by a constant amount \( d \) for all players. Our basic method for doing inference is based on simulating situational data assuming our probability models and seeing how the simulated data compare to the actual situational data that we observed. What we discover is that most of the interesting observed situational effects that we see are simply due to chance variation and, if they exist, the true situational effects will tend to be small.

"Streakiness"

A second popular topic among baseball fans is the presence of the so-called “hot or cold hand.” During the baseball season, we will observe teams with long winning or losing streaks, or observe batters or pitchers with extended periods of success. Are these periods of observed streakiness meaningful? To most baseball fans, the answer is yes - if a player goes through a difficult stretch of hitting, writers and broadcasters will offer a variety of explanations for this hitting slump, implying that the player has a low batting ability.

One goal of this discussion is to clearly distinguish between real streaky ability and observed streakiness. With respect to ability, it is easiest to describe a player who is not streaky. If we are focusing on the event of getting on-base, then a player has true consistent (not streaky) ability if the probability of him getting on-base is always the same value. In contrast, a true streaky hitter has a more complicated probability structure. Perhaps this player is either “hot” or “cold” with respective on-base probabilities of \( p_H \) and \( p_C \) and he moves between these two hot and cold states according to a Markov Chain with given transition probabilities.

We next discuss ways of measuring streaky performance of a player or team. The basic data structure is the day-to-day hitting performance (for a batter) or day-to-day win/lose performance (for a team). From these data, some “streaky” statistics are

- moving averages using a suitable window width
- lengths of runs of good days and bad days
- the total number of runs in the sequence

Finally, we connect the discussion of consistent and streaky ability with the observed streakiness that we measure by the lengths of runs or the unusually large or small moving averages. We focus on the basic coin-tossing model where the probability of an event does not change across games. We simulate data from this consistent model, compute streaky statistics from the simulated data, and compare the values of these statistics with the data from the player who is thought to be streaky. What we learn is that genuine streakiness is very hard to detect statistically and even hitting or win/loss data from a truly consistent player or team can look very streaky. Chapter 5 of Albert and Bennett (2001) gives a more extensive discussion on the topic of detecting streakiness.

7. Discussion

This section contains responses to several arguments against offering an introductory baseball statistics course, and some observations based on our experience teaching this course for two semesters.
Argument 1: *All students aren't interested in baseball.*

Obviously, many students are not interested in baseball and wouldn’t find this course any more interesting or relevant than the standard statistics course. But at our university and many others, there is a large audience for this introductory course and it is easy to fill one class that is devoted to baseball. Also, there were students in the class who were not necessarily baseball fans, but were interested in learning more about the game and the associated statistics.

Argument 2: *Baseball (a game) and statistics (serious science) don't mix.*

Although baseball is a game, it is a serious business for the players, managers, and owners. A proper interpretation of baseball statistics is important for the enterprise of building a team and winning games.

Argument 3: *The course appeals mainly to one gender.*

It is true that more men are interested in baseball than women and this course tends to draw more men. But there is a large population of women who attend baseball games and there is likely a large group of women from the population of students who are taking introductory statistics. There were some women in the class who were not that familiar with the game but were receptive to learn.

Argument 4: *The students won't be able to think statistically in other settings.*

Because the goal of this particular introductory statistics course is to help the student become a better consumer of statistical information that is reported in the media, it would seem beneficial to expose the student to applications outside of the world of sports. Of course, the biggest challenge is for the student to actually learn the concept, such as the distinction between the population and the sample. If the students can learn the concept through the baseball application, then it would seem to be relatively easy to apply this concept to a non-sports setting.

Argument 5: *This course does not cover all of the topics that are typically discussed in a first course.*

The only topic that received little attention in this course was the issue of collecting data through samples and designed experiments. However, it would be possible to use baseball to discuss sampling and experimentation. Sampling can be used to summarize the large mass of historical baseball data, and experimentation has been used in baseball in the construction of equipment such as baseball and bats.

Was this course successful? The answer depends on one’s definition of success, but two things were obvious in our experience teaching this course. First, the course was fun for both the instructor and the students. The fact that the instructor enjoyed the course is important. The enthusiasm of the instructor about the baseball material seemed to have a positive impact on the learning of the material. Second, baseball provided an interesting context to learn about statistical thinking. In a student evaluation given at the end of the course, students overwhelmingly said that the course was “useful.” This comment doesn’t mean that the students will use what they learned about baseball in their future work. Rather, it meant that the students could make sense of the statistical material since it was taught from a baseball perspective. The positive experience in this class suggests that we should encourage alternative models for teaching statistics. We should explore ways or contexts to engage students so they can make more sense of statistical thinking.
References

Albert, J., and Bennett, J. (2001), *Curve Ball: Baseball, Statistics, and the Role of Chance in the Game*, New York: Copernicus Books.

Hogg, R. V. (1992), “Towards Lean and Lively Courses in Statistics”, in *Statistics in the Twenty-First Century*, eds. F. Gordon and S. Gordon, Washington, DC: Mathematical Association of America.

Moore, D. S. (1992), “Teaching Statistics as a Respectable Subject”, in *Statistics in the Twenty-First Century*, eds. F. Gordon and S. Gordon, Washington, DC: Mathematical Association of America.

Moore, D. S. (1993), “A Generation of Statistics Education: An Interview with Frederick Mosteller”, *Journal of Statistics Education* [Online], 1(1). (ww2.amstat.org/publications/jse/v1n1/moore.html)

Snell, J. L., and Finn, J. (1992), "A Course called Chance," *Chance*, 5, 12-16.

Sowey, E. R. (1995), “Teaching Statistics: Making It Memorable," *Journal of Statistics Education* [Online], 3(2). (ww2.amstat.org/publications/jse/v3n2/sowey.html)

Willett, J. B., and Singer, J. D. (1992), “Teaching Applied Statistics Using Real-World Data,” in *Statistics for the Twenty-First Century*, eds. F. Gordon and S. Gordon, Washington, DC: Mathematical Association of America.

Yilmaz, M. R. (1996), “The Challenge of Teaching Statistics to Non-Specialists," *Journal of Statistics Education* [Online], 4(1). (ww2.amstat.org/publications/jse/v4n1/yilmaz.html)

Zetterqvist, L. (1997), “Statistics for Chemistry Students: How to Make a Statistics Course Useful by Focusing on Applications,” *Journal of Statistics Education* [Online], 5(1). (ww2.amstat.org/publications/jse/v5n1/zetterqvist.html)

Jim Albert
Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, OH
USA
albert@bgnet.bgsu.edu