NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

American Statistical Association

The Statistics Teacher Network

www.amstat.org/education/stn/index.html

Number 70 ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability **Fall 2006**

Dear Readers,

We open this year with reviews of some materials you may not have examined. We have reprinted blackline masters of some activities from these materials that you can use immediately. In addition to these reviews and activities, we are including two articles that explore game theory and probability with the hope that you will be able to adapt some of the ideas for incorporation into your curriculum this year.

Best wishes for a productive 2006-2007 school year.

Beth Lazerick
Murray Siegel

Book Reviews

STATISTICS Concepts and Controversies, 6th edition, David Moore, William Notz, 2006, 561 pp., $53.00, ISBN 0-7167-8636-2; W.H. Freeman & Co., www.whfreeman.com, (800-446-8923)

This classic book has been around for quite a few years and the latest edition has maintained its relevance. The authors consistently use examples and exercises that present real problems that are timely and will be of interest to students. Some excellent examples are environmental factors and cancer; exit polls vs. pre-election polls; and bone marrow transplants.

The opening narration for chapter 2 is a wonderful example. A small town newspaper asked its readers to call in and express an opinion on the local ambulance service. The newspaper was pleased that 3763 calls were made, demonstrating a real interest in the subject. Further investigation revealed that 638 calls were made from the offices of the ambulance company or from the homes of company executives. This chapter is devoted to methods of sampling.

The introduction to chapter 9 (Do Numbers Make Sense) relates the story of an automobile dealer who defrauded General Motors with numbers that should have immediately raised eyebrows at GM. Unfortunately, nobody assessed the data until six billion dollars had been loaned to the dealer.

Exercise 26 in chapter 15 asks the student to consider what lurking variables might cause a College Board study to show a relationship between taking algebra and geometry in high school and success in college. The introduction to that chapter recounts and questions the supposed relationship between who wins the NFL Super Bowl and the rise and fall of the stock market.

There is a potential hazard embedded in the text's use of relevant and interesting problems. Some of the topics may be troubling to administrators, especially in high schools. Same-sex marriage, gun control, sexual behavior & AIDS, gambling, drinking alcoholic beverages and abortion are investigated. If these subjects would be troubling at your school, the book would still be an excellent resource for the teacher, providing many useful examples of the importance of understanding how to use and how to abuse the study of data and chance.

This book could be used for a non-AP statistics course in high school or for a college introductory course for students who are not "mathy". This book should be in the personal library of every teacher of AP Statistics and would be a useful teaching asset for teachers at any level who include data analysis or probability in their curriculum.

Also In This Issue...

Baseball Math, **Christopher Jennison, 2005,104 pp., $12.95, ISBN 1-59647-007-0;**

Football Math, **Jack Coffland, David Coffland, 2005, 118 pp., $12.95, ISBN 1-59647-008-9;**

Racing Math, **Barbara Gregorich, Christopher Jennison, 2006, 106 pp., $12.95, ISBN 1-59647-060-7;**

Fourscore and 7, **Betsy Franco, 1999, 137 pp., $12.95, ISBN 1-59647-000-3.**

Good Year Books, www.goodyearbooks.com, (888-511-1530)

Baseball Math, **Football Math**, and **Racing Math** are paperbound books that contain multiple activities and projects that will challenge students in grades 4-8 to use computational skills as well as data analysis and probability concepts. Each book uses a different sport as background for the activities and students are more likely to become actively engaged in learning if the subject matter is of interest to them. Much of the material and many of the activities could easily be adapted for younger or older students. Activities include many opportunities to use and interpret data from graphs and charts. Many projects require long-term exploration of real data from news media and/or the Internet. Teachers will find these books include motivating material that will entice some of the more reluctant mathematicians because of the subject matter.

Four Score and 7 is a volume that helps teachers and students explore mathematical ideas in the context of American history. In " Eureka! The California Gold Rush" students experiment with probability using dice and spinners. In "I Do Solemnly Swear...! Presidents of the United States" students create and interpret bar graphs. Other activities incorporate primes and composites, fractions and whole number computational skills.

Teachers can reproduce portions of the text for use within their classrooms. Examples of activities are provided in this newsletter on pages 6 and 7.

Groundworks, Reasoning with Data and Probability, Grades 1-7, **Jenny Tsankova, Carol Findell, Carole Greenes, Barbara Irvin, 2006, 96 pp. each, $34.62/volume, ISBN 1-4045-3197-1 (grade 1), Creative Publications/The Wright Group, *www. WrightGroup.com*, 1-800-593-4418.**

Groundworks, Reasoning with Data and Probability is a series of seven paperbound books each containing 12 sets of problems covering different aspects of data analysis and probability at first through seventh grade levels. Five "big ideas" are addressed: Interpreting data displays, organizing data, describing data, counting and probability. Each set of problems begins with introductory material for teachers, followed by six blackline masters with an answer key.

These volumes are well laid out with clear and uncluttered activity sheets for students. A management chart for tracking students' progress is also included.

Stats Modeling the World, **2nd Edition, David Bock, Paul Velleman, Richard DeVeaux, 2007, 801 pp., $117.33, ISBN 0-13-187621-X.**

Stats Modeling the World (MTW) is designed to specifically meet the needs of AP Statistics classes. David Bock has been an AP Statistics teacher and exam reader. Paul Velleman is a significant contributor to the AP Statistics listserv. He has the knowledge and communications ability to clearly articulate difficult concepts. The text is chock full of colorful graphs, photographs and amusing cartoons that make reading a pleasant task.

The text material is written to be read and comments are injected that anticipate questions that the reader may wish to ask. One would have to say that MTW was written to be helpful. An example occurs at the conclusion of Chapter 5 that covers numerical description of data. A section labeled "What Can Go Wrong" provides advice to the student that will help reduce the chance of problems when summarizing data. Comments such as "Don't compute numerical summaries for a categorical variable" and "Make A Picture" are included.

There are some helpful features that are found throughout the text. In every chapter a step-by-step process is included using headings of "Think", "Show" and "Tell". TI Tips are incorporated to provide easy to follow explanations of how to use the TI-83 Plus and TI-84 Plus. Also, a number of chapters include Math Boxes, which are explanations of necessary mathematical concepts. For example, in chapter 17, which investigates probability models, there is a Math Box that explains the derivation of the formulas for the mean and standard deviation of the binomial distribution.

The authors are concerned that students have a complete understanding of the conceptual material. They have added a "just checking" feature in each chapter that allows students to work a simple problem to check for understanding. Answers are provided for the "just checking" problem at the end of the chapter. For example, Chapter 22 (Comparing Two Proportions) has a "just checking" problem that addresses a public broadcasting station's attempt to assess the difference between two methods of seeking contributions. A confidence interval for the difference between success rates is provided and students are asked to interpret the interval and to reach a conclusion based on that interval. Incorrect answers would warn the student that a review of the material was necessary.

The authors state that this edition of their text was crafted to be clear and accessible. It appears to this reviewer that their goal has been realized.

Statistics in Action, Ann Watkins, Richard Scheaffer, George Cobb, 2004, 751 pp., $69.95, ISBN 1-55953-313-7

Any teacher who has been teaching data analysis/probability at any level should be familiar with the authors of this text. Dick Scheaffer and Ann Watkins have been significant contributors to the Quantitative Literacy project and each was a founding guru of Advanced Placement Statistics. George Cobb is an expert on statistics education and is well known for his work on experimental design.

The text opens with an investigation of age discrimination in the work place. Without formulas or a pre-determined sequence of methods, the student is asked to judge the situation simply using the numerical data that is provided in the problem. This investigation introduces the course and provides real motivation for taking this course.

As you flip through the chapters you will note that there are three types of problems in the book. Those labeled with a "P" are practice activities. Those labeled "D" are meant to provide opportunities for discussion. The "E" problems are exercises designed to reinforce conceptual learning.

Real data are used in every chapter. One section is devoted to developing the formula for correlation, airline data is used to visualize the meaning of correlation. A scatter plot of mishandled baggage vs. the percent of on-time arrivals is displayed, with each point in the plot representing a particular airline. The plot is divided into quadrants and the meaning of an airline's point being located in a specific quadrant is discussed.

Sometimes the real data is actually gathered by students using an activity. In the development of the sampling distribution of sums and differences, students roll a die two times and record the sum and difference of the two rolls. This is repeated 100 times and the distribution of the 100 sums and 100 differences is analyzed. In another activity that focuses on inference on means, the effect of skewness on a confidence interval is investigated using a random sampling of a skewed set of data. The data are brain weights for 68 species of animals. This distribution is highly skewed towards greater values due to the presence of a small number of larger mammals in the set of data. Students gather random samples of size five. Each sample yields a 95% confidence interval. The percent of the confidence intervals that capture the mean computed for the sixty-eight animals.

The authors of this book were intimately involved in the development of the AP Statistics curriculum. Thus it should be very useful for teachers of this AP course. ■

Showing Pennies

Derek Webb
Bemidji State University
Bemidji, MN 56601-2699
dwebb@bemidjistate.edu

Is **THIS** a fair game? Showing Pennies is a simple game played by two people that lends itself to many interesting statistics and probability analyses and discussions including hypothesis testing, the expected value, and what is meant by a "fair game." The level of material can be tailored to fit a 7th through 12th grade class. Two people (Ralph and Mary) play the game. Mary and Ralph each have their own penny and one play of the game consists of them showing the other their penny. Score is kept as follows: If both Mary and Ralph show heads then Ralph gets 3 points. If both Mary and Ralph show tails then Ralph gets 1 point. If the coins don't match then Mary gets 2 points. The game is repeated, score is kept, and the highest score wins.

Before having students play the game, the teacher can pose the following hypotheses: The game is fair vs. the game is not fair? This will generate discussion of what a "fair game" really is and may lead to defining and discussing an expected value. Students should then decide on an experiment for testing the hypotheses and carry out the experiment.

Most students will start playing the game by flipping their pennies. In time, some will switch to what they consider to be an advantageous strategy. After a while, the teacher should have the students take a break from playing the game and discuss the initial hypotheses. By now, most students will say that the game is not fair (and it is not!), but there will probably be many ideas presented as to why the game is not fair. The teacher will want to remind the students to support their ideas with the data that they collected. Students should also be encouraged to share their strategies so that they realize that more than one strategy exists.

At this point, the teacher can encourage the students to update their hypotheses about the game based on their experiences. The teacher can direct students to focus on specific hypotheses such as "The best strategy for Mary is to always show tails." or "Ralph should show heads 2/3 of the games he plays." Depending on the level of the class, simulation or theory (or both) can be used to explore these hypotheses.

Simulation Approach

If a simulation approach is used, each student should decide upon a particular strategy and then figure out how to implement it using random number generators like spinners, dice, or calculators. For example, if Ralph wanted to show heads 2/3 of the games he plays, he could create a spinner with 3 equal sized outcomes, 2 tail outcomes and 1 head outcome. The spinner would then be used to determine if Ralph should show tails or heads to Mary for any particular game. Calculators and dice can be used in a similar way. Having students come up with their own methods for randomly deciding whether to show heads or tails in a particular game is an important part of learning about probability associated with carrying out a simulation.

Once strategies are decided upon students should play the game a number of times and keep track of the points. An interesting discussion point regarding using simulation for hypothesis testing is: "how many games should be played to test the hypothesis?" This is not an easy question to answer but does lead to good classroom discussion. After they have conducted the simulation, students can then summarize their findings using descriptive statistics, tables, or charts. Students should be given time to discuss their hypotheses concerning strategy in light of the simulation.

Theoretical Approach

The theoretical approach reveals a quite interesting result. Define the random variable X to be the number of points that Mary wins every time she plays the game. Let p = the probability that Mary shows a tail and let q = the probability that Ralph shows a tail. The discrete distribution of X is:

X	-3	-1	2	2
Probability of X	$(1-p)(1-q)$	pq	$(1-p)q$	$p(1-q)$

The expected value of X is $E[X] = (-3)(1-p)(1-q) + (-1)pq + 2(1-p)q + 2p(1-q)$ and is easily shown to be 0 for $p = q = 1/2$, a fair game. So, if students are flipping coins (or randomly choosing to show tails or heads) then the game is statistically fair because the expected value of number of points won is 0. The same argument could be made in terms of defining the random variable Y to be the number of points that Ralph wins every time he plays the game. The expected value of Y will also be 0. Students can try this calculation for themselves.

The expression for the expected value of X can be simplified to yield $E[X] = -3 - 8pq + 5p + 5q$. Students can choose different values of p and q and compute the expected value. Note that Mary wants the $E[X]$ to be greater than 0 and Ralph wants the $E[X]$ to be less than zero. At this point, students could create a table containing values of p and q ranging from 0 to 1 and the resultant expected values. Students can then theoretically see how good their strategies are.

The interesting aspect to this game is seen by further rearranging the expected value to be $E[X] = -3 + q(5-8p) + 5p$. Now, let $p = 5/8$. Then $E[X] = -3 + 25/8 = 1/8$. The expected value is positive (1/8) and does not depend on the value of q. This implies that the best strategy for Mary is to show tails 5 games out of 8. If Mary does this, then there is no way for Ralph to counter. In fact, it does not matter what Ralph does, Mary wins, on average, 1/8 of a point every time a game is played. In other words, in 8 games, Mary is up 1 point.

In conclusion, this simple coin game can be used to illustrate basic concepts of probability including simulation and the expected value. Students have the opportunity to develop their own strategies for playing the game and can quantitatively analyze those strategies. An optimal strategy exists and algebra can be used to find it. ■

Reference

Edward Spellman presented this coin game in the "Ask Marilyn" column of *Parade Magazine*, April 7, 2002.

A Statistical Analysis of MEGA MILLIONS

Sanderson M. Smith, Santa Barbara City College, Santa Barbara, California

MEGA MILLIONS was introduced by the California Lottery in June, 2005. California joins the other participating states: Georgia, Illinois, Maryland, Massachusetts, Michigan, New Jersey, New York, Ohio, Texas, Virginia, and Washington. There are nine ways to win at least some money for your "investment" of $1. The record jackpot to date is $363 million dollars.

The rules for MEGA MILLIONS are relatively simple. There are two number sets:

Set A: The integers from 1 to 56 inclusive.

Set B: The integers from 1 to 46 inclusive.

A player spends $1 for the privilege of selecting five numbers from Set A, and one number, called the MEGA number, from Set B. All prize payout amounts are pari-mutual. That is, the amount of money varies from week to week, depending upon the sales levels and the number of winners. The jackpot prize is obtained when one matches all five numbers and the MEGA number. A jackpot match occurs when one's number selections match a set of five numbers from A and one number from B chosen by random selection by the Lottery Commission.

There are nine ways to win in MEGA MILLIONS. Probabilities and odds associated with the game are displayed in the table below. Here are some notes and observations relating to the table:

* When playing, two events are involved. One involves choosing five numbers from Set A, and

Statistical Analysis of MEGA MILLIONS

EVENT	SYMBOLIC PROBABILITY	NUMERICAL PROBABILITY	ODDS: 1 in --
Match FIVE and MEGA	$({}_5C_5)({}_{51}C_0)(1/46)/{}_{56}C_5$	5.69114597E-9	175,511,536
Match FIVE and NO MEGA	$({}_5C_5)({}_{51}C_0)(45/46)/{}_{56}C$	2.561015687E-7	3,904,701
Match FOUR and MEGA	$({}_5C_4)({}_{51}C_1)(1/46)/{}_{56}C_5$	1.451242222E-6	689,065
Match FOUR and NO MEGA	$({}_5C_4)({}_{51}C_1)(45/46)/{}_{56}C_5$	6.530590001E-5	15,313
Match THREE and MEGA	$({}_5C_3)({}_{51}C_2)(1/46)/{}_{56}C_5$	7.256211112E-5	13,781
Match THREE and NO MEGA	$({}_5C_3)({}_{51}C_2)(45/46)/{}_{56}C_5$	0.0032652950	306
Match TWO and MEGA	$({}_5C_2)({}_{51}C_3)(1/46)/{}_{56}C_5$	0.0011851811	844
Match ONE and MEGA	$({}_5C_1)({}_{51}C_4)(1/46)/{}_{56}C_5$	0.0071110869	141
Match NONE and MEGA	$({}_5C_0)({}_{51}C_5)(1/46)/{}_{56}C_5$	0.0133688434	75
		TOTAL PROBABILITY OF WINNING SOMETHING = **0.0250699874**	0.0250699874 = 1/(1/0.0250699874) = **1/39.8883327** Overall odds of winning are **1 in 39.89** (Agrees with what is printed on Mega Million ticket.)

the second consists of picking one number from Set B. The two events are statistically independent, so related probabilities can be multiplied, as is done in the table.

* The probability that a player matches the MEGA number is 1/46, and the probability that the MEGA number is not matched is 45/46.
* The symbolism nCr represents the number of different sets of r objects that can be selected from a set of n objects. For instance, consider the set {red,white,blue,green}. If I want to choose two colors from this set, there are ${}_4C_2$=6 ways this can be done. The six possible sets are {red,white}, {red,blue}, {red,green}, {white,blue}, {white,green}, and {blue,green}. Many calculators (the TI-83, for instance), have menus that include nCr.
* A number such as 5.69114597E-9 is 5.69114597/10^9, which is 0.00000000569114597. (This number does appear in the table.)

It is often helpful to represent very small probabilities in the form 1/x using the algebraic identity

x = 1/(1/x), **if x is not zero**.

For instance, the probability of hitting the jackpot is 0.00000000569114597 (see table). Using the algebraic identity, this can be written as

0.00000000569114597 =
1/(1/0.00000000569114597) = 1/175,711,536

In other words, the odds of winning the jackpot are 1 in 175,711,536.

[Note on odds terminology: When we see odds 1:5, this means that in 6 trials we expect 1 success and 5 failures. For instance, in the roll of a single die, the probability of getting a 4 is 1/6. The odds favoring this event are 1:5 which is read “1 TO 5.” The California Lottery Commission used the word IN instead of TO. The Commission would express the odds as “1 IN 6.” This is the notation used in the following table where the odds are expressed as they appear on MEGA MILLION tickets.] Below is a statistical analysis of MEGA MILLIONS.

To introduce a time perspective, suppose you had a desire to purchase tickets covering all 175,511,536 possibilities to be sure you won the jackpot. If you could fill in one ticket per second, it would take you 175,511,536 seconds = 2,925,192 minutes = 48,753 hours = 2,031 days = 5.56 years to complete the task.

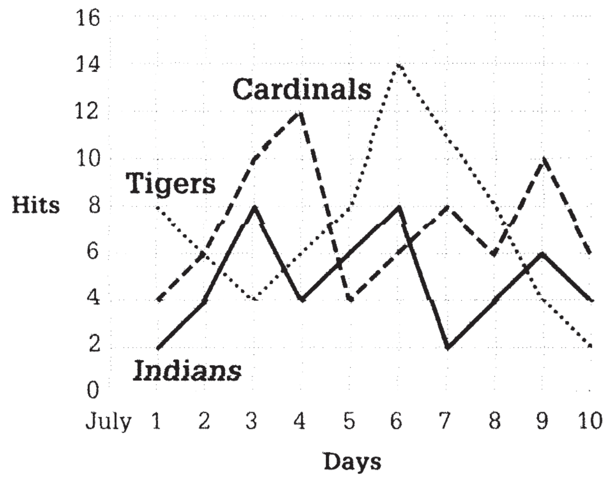
As the table indicates, the probability of hitting the jackpot is quite small. But people have won the jackpot in the past, and people will win it in the future.

“Oh, many a shaft at random sent
Finds mark the archer little meant.”

Sir Walter Scott, *The Lord of the Isles* ■

Hit Parade

The Tigers, Cardinals, and Indians are keeping track to see which team makes *the most hits in a ten-day period. This line graph shows the total number of* hits each team made.



1. What was the first day each team was able to make six hits?

   Tigers ____________________

   Cardinals ____________________

   Indians ____________________

2. On what day did the Tigers and Cardinals have the same number of hits?

   ____

3. On what day were the most number of hits made?

   ____

4. On what day were the fewest number of hits made?

   ____

5. What was the average number of hits made by the Tigers during the ten-game period?

   ____

6. Which team made the most hits during the ten-game period?

   ____

7. Which team improved its hit total the most from one day to the next?

   ____

8. Which team was most consistent throughout the ten-game period? *(In other words, which team had the least variation between its high and low totals?)*

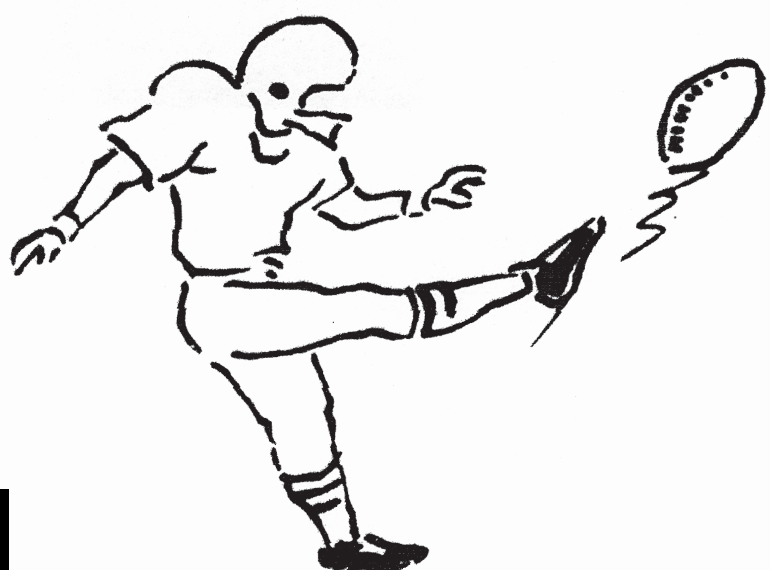
   ____

Kicking Statistics for the Season

Collect as many statistics as you can on several kickers who are kicking this year. Keep the statistics for an entire year. (You may want to keep statistics for several kickers, as it is always possible that someone will be hurt and not kick very much during a year.) Consider keeping the following statistics:

a. *Number of kickoffs*
b. Number of kickoffs kicked into the end zone for a touchback
c. Number of kickoffs returned for a touchdown
d. Average length of kickoff returns (This is important: Kickers who kick the ball high usually cause the other team to make shorter returns.)
e. *Number of field goals attempted and* made
f. Number of point after touchdowns attempted and made
g. Number of field goals blocked
h Number of points after touchdowns blocked
i. Length of all field goal kicks attempted and made
j. Any other statistic you think is important

Then show your statistics to a friend *and discuss who is the best kicker.*



Statistics Teacher Network
c/o American Statistical Association
732 North Washington Street
Alexandria, VA 22314-1943
USA

Non-Profit Organization
U.S. Postage
PAID
Permit No. 361
Alexandria, VA

Let us know your Zip +4!

Dear Subscribers,

We are looking for good articles for the spring issue of *STN*. We may want to highlight successful classroom practices. If you have had a great experience teaching probability and/or data analysis, let us hear from you.

To subscribe to the *STN Newsletter*, please send your email and snail mail addresses to *Rebecca@amstat.org*. ALL NEW SUBSCRIBERS will receive the newsletter electronically unless they specify otherwise. IF YOU ARE CURRENTLY RECEIVING A PAPER COPY AND ARE WILLING TO RECEIVE YOUR NEWSLETTER ELECTRONICALLY, please let Rebecca know by sending her your current address and email address also. In this way we can expand the subscriber base and decrease costs.

Thanks,

Beth Lazerick, *steditor@aol.com*
Murray Siegel, *siegel@gssm.k12.sc.us*