

NUMB3RS Activity: The Pythagorean Expectation Episode: "Hardball"

Topic: Algebra and Sabermetrics

Grade Level: 9 - 12

Objective: Apply and manipulate a formula for winning percentage.

Time: 15 minutes

Materials: TI-83 Plus/TI-84 Plus graphing calculator

Introduction

In "Hardball," an amateur mathematician uncovers an equation that identifies steroid abuse in baseball. One branch of analysis used to study baseball statistics is called Sabermetrics, which is named for SABR (Society for American Baseball Research). Charlie describes Sabermetrics as "a powerful form of analysis in baseball - because the physical nature of the game involves chance. The difference between a hit and an out can be millimeters or milliseconds. At the same time, there's a tremendous amount of data recorded throughout the season. Literally thousands of at bats. And when you have athletic situations involving chance repeated over and over again – statistical analysis can isolate and reveal human performance."

The *Pythagorean expectation* is a sabermetric formula used to estimate how many games a team should win based on the number of runs scored and the number of runs allowed. The basic formula for the *Pythagorean expectation* is shown below, where W is the winning percentage, S is the number of runs scored, and A is the number of runs allowed. The formula is named the Pythagorean expectation because of the resemblance to the Pythagorean Theorem, but there is no other mathematical connection. This formula for expected winning percentage was developed through empirical analysis.

$$W = \frac{S^2}{S^2 + A^2}$$

In this activity, students will solve for the individual variables given baseball statistics from the 2006 season. After becoming familiar with the formula, students will analyze and manipulate the equation. An extension activity offers the opportunity to explore alternate values for the exponent. A spreadsheet will make the cumbersome computations in the extension much easier.

Student Page Answers:

1a. 0.597 **2b.** .97 **2a.** 0.516 **2b.** 756 **2c.** 762 **3.** 949 **4a.** The value of W approaches one, which means the team is expected to win all of its games. **4b.** The lowest value of W is zero, which occurs when $S = 0$. **5.** The numerator and denominator of the fraction on the right hand side are each divided by S^2 (or, the right hand side of the equation is multiplied by $(1/S^2)/(1/S^2)$).

Name _____ Date _____

NUMB3RS Activity: The Pythagorean Expectation

In "Hardball," an amateur mathematician uncovers an equation that identifies steroid abuse in baseball. One branch of analysis used to study baseball statistics is called Sabermetrics, which is named for SABR (Society for American Baseball Research). Charlie describes Sabermetrics as "a powerful form of analysis in baseball - because the physical nature of the game involves chance. The difference between a hit and an out can be millimeters or milliseconds. At the same time, there's a tremendous amount of data recorded throughout the season. Literally thousands of at bats. And when you have athletic situations involving chance repeated over and over again – statistical analysis can isolate and reveal human performance."

The *Pythagorean expectation* is a sabermetric formula used to estimate how many games a team should win based on the number of runs scored and the number of runs allowed. The basic formula for the *Pythagorean expectation* is shown below, where W is the winning percentage, S is the number of runs scored, and A is the number of runs allowed. The formula is named the Pythagorean expectation because of the resemblance to the Pythagorean Theorem, but there is no other mathematical connection. This formula for expected winning percentage was developed through empirical analysis.

$$W = \frac{S^2}{S^2 + A^2}$$

Part 1: Applying the Formula

In the 2006 regular season, the Detroit Tigers played 162 games. The team scored 822 runs and allowed 675.

1.
 - a. Use the Pythagorean expectation formula to determine their expected winning percentage.
 - b. The Tigers won 95 games in the 2006 season. How many games did the Pythagorean expectation from Question 1a predict they would win?

In the 2006 regular season, the St. Louis Cardinals won 83 of 161 games and scored 781 runs (They played only 161 games because the September 17 game against the San Francisco Giants was rained out and not rescheduled.)

2.
 - a. Determine the Cardinals' actual winning percentage.
 - b. Using the actual winning percentage from Question 2a in the Pythagorean expectation formula, determine how many runs they should have allowed.
 - c. The Pythagorean expectation predicted a winning percentage of .512. Use this value to determine the actual number of runs the Cardinals allowed.

3. In the 2004 season, the Boston Red Sox had an expected winning percentage of .604 and allowed 768 runs. Use the Pythagorean expectation formula to determine how many runs they scored.

Part 2: Analyzing the Formula

4. Again, look at the formula for the expected winning percentage.

$$W = \frac{S^2}{S^2 + A^2}$$

- a. What happens to the value of W as A gets smaller? Why is this value of W an upper bound?
- b. What is the smallest possible value of W ? What must be true about the value of S in this case?
5. Mathematicians sometimes rewrite formulas to make analysis easier. One area of study for the *Pythagorean expectation* formula involves varying the exponent of 2. The formula shown below is equivalent to the original formula, but rewritten to reduce the number of powers. Explain how the original formula was rewritten to obtain this new form.

$$W = \frac{1}{1 + \left(\frac{A}{S}\right)^2}$$

The goal of this activity is to give your students a short and simple snapshot into a very extensive math topic. TI and NCTM encourage you and your students to learn more about this topic using the extensions provided below and through your own independent research.

Extensions

In the activity, we let $x = 2$ in the sabermetric formula below.

$$W = \frac{S^x}{S^x + A^x}$$

However, other values of x have proven more accurate. Use the data below to generate predicted winning percentages for different values of x . Fill in the blank rows with your favorite teams and consider using a spreadsheet to decrease time spent on computation.

| Team | Runs Scored | Runs Allowed | Actual Winning Percentage | W , when $x = 1$ | W , when $x = 1.3$ | W , when $x = 1.5$ | W , when $x = 1.8$ | W , when $x = 2.0$ |
|--------------|-------------|--------------|---------------------------|--------------------|----------------------|----------------------|----------------------|----------------------|
| Diamondbacks | 773 | 788 | | | | | | |
| Braves | 849 | 805 | | | | | | |
| Orioles | 768 | 899 | | | | | | |
| Red Sox | 820 | 825 | | | | | | |
| Cubs | 716 | 834 | | | | | | |
| Reds | 749 | 801 | | | | | | |
| Rockies | 813 | 812 | | | | | | |
| Astros | 735 | 719 | | | | | | |
| Twins | 801 | 683 | | | | | | |
| Phillies | 865 | 812 | | | | | | |
| Yankees | 930 | 767 | | | | | | |
| Blue Jays | 809 | 754 | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |

[Source: <http://www.baseball-reference.com>]

For each value of x , plot the expected winning percentage against the actual winning percentage. The graph that is closest to a straight line is the best value of x . After analyzing a large amount of data, mathematicians have settled on a value of 1.8.

Additional Resources

- You can look up your own baseball data to analyze at the Web site <http://www.baseball-reference.com>.
- The Web site for the Society for American Baseball Research is <http://www.sabr.org>.