The Pythagorean Expectation:

Softball Friendly?

By: Kristen Kirk

The Pythagorean Expectation

- What is the Pythagorean Expectation?

  The Pythagorean Expectation was first developed by Bill James to explain the relationship between wins and losses to runs scored and runs allowed.

  The equation is:

$$WP_P = \frac{RS^2}{RS^2 + RA^2}$$

The Pythagorean Expectation

- Bill James
- Steven J. Miller
  - Weibull distribution
  - Statistically independent

The Pythagorean Expectation and Softball—The Big Question

- The big question addressed in this project is….

Can the Pythagorean Expectation be carried over to softball?

Problems in Research

- What branch of softball should be used?
- Problems with research and available data

The Calculations

- First, the actual winning percentage of each team was determined by:

$$WP_A = \frac{Wins}{GamesPlayed}$$

Example: In 2006, the University of Tennessee won 61 games out of 73 games played. Their actual winning percentage is:

$$WP_A = \frac{61}{73} = 0.83562$$

The Calculations

- Next, the Pythagorean Expectation was used to determine the Pythagorean winning percentage.

Example: In 2006, the University of Tennessee scored 437 runs and only allowed 98 runs.

$$WP_P = \frac{437^2}{437^2 + 98^2} = 0.95212$$

So, the Pythagorean Expectation gave the winning percentage of 0.95212.

The Calculations

- There is a discrepancy between the actual and Pythagorean winning percentages. This leads one to wonder if the equation could be altered to give a more accurate winning percentage. What process should be used to alter the equation?

The Calculations

Finding the Residual: (for the University if Tennessee in 2006)

0.95212 – 0.83562 = 0.11650

Find the Sum of Squares:

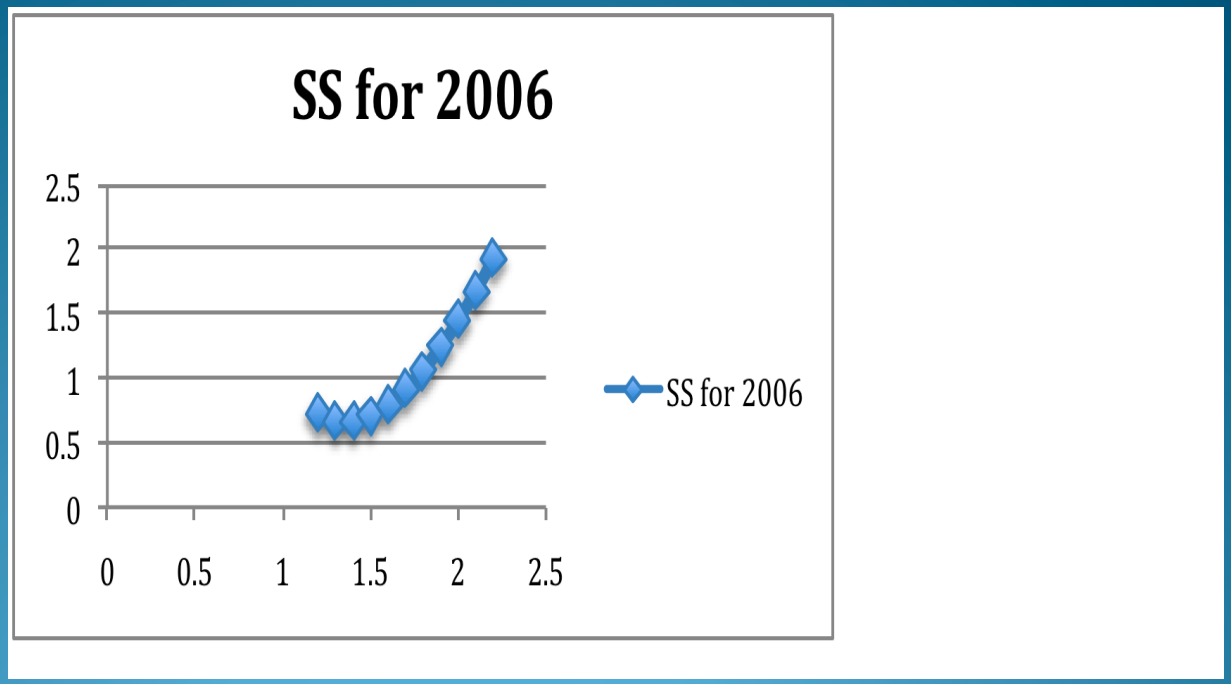
$(0.11650)^2 = 0.01357$

The Calculations

- To determine the best exponent, several steps were taken.
- Multiple exponents were used, ranging from 1.1 to 2.5.
- Pythagorean winning percentage, the residual, sum of squares for each team, and the total sum of squares were determined for each exponent.

The Calculations

- For each year, the sums of squares were graphed for each exponent. From these graphs, the general range of the ideal exponent was determined.



The Calculations

- After a smaller range of exponents was determined, like 1.3 to 1.4 for 2006, more specific exponents were used. In 2006, 1.31 to 1.39 were used. The same process was used as before.

The Calculations

- The exponent that gave the lowest sum of squares was determined to be the best exponent for that year. These exponents were compiled in a table, and the median exponent was found to be 1.31.

Table

Year	Exponents
2008	1.41
2007	1.39
2006	1.36
2005	1.31
2004	1.24
2003	1.27
2002	1.28

How well does it work?

- The Pythagorean Expectation may not be the most accurate predictor of winning percentages. The lowest sum of squares of the residual that was found was 0.59120 in 2007 with and exponent of 1.4. The highest sum of squares was a little over 3.1.
- Speculations and Hypotheses

Predicting with the Pythagorean Expectation

- Now, the question is raised: Does the actual winning percentage or the Pythagorean winning percentage best determine the winner of the Women's College World Series (WCWS)?

- Research

Calculations

- Using the actual and predicted winning percentages already determined, the possibility of victory was found for each team playing in each game.

Equations

- AP_H predicts the home team's actual possibility of winning.
- AP_V predicts the visiting team's actual possibility of winning.

- PP_H predicts the home team's predicted possibility of winning.
- PP_V predicts the visiting team's predicted possibility of winning.

Possibility of Victory

- In Game 1 of the 2008 WCWS, the University of Florida played Louisiana-Lafayette.
- Florida's WP_A= 0.93333 and WP_P= 0.87424
- Louisiana-Lafayette's WP_A= 0.77612 and WP_P= 0.63090
- Using the equations just shown, we find Florida's actual and predicted possibilities of winning.

- AP_H = 0.54589
- PP_H = 0.58084

Possibility of Victory

- Using the given equations, we find Louisiana-Lafayette's actual and predicted possibilities of victory.
- Recall:

  Florida's WP_A= 0.93333 and WP_P= 0.87424

  Louisiana-Lafayette's WP_A= 0.77612 and WP_P= 0.63090

- AP_V = 0.45401
- PP_V = 0.41916

Calculations

- What do these probabilities really show and how will they determine if one method is better at predicting the winners of the WCWS?

- Use the predictor deficit!! Find the predictor deficit for both actual possibility of victory and the predicted possibility of victory.

Determining the Best Method

- If the team with the higher possibility of victory wins, the predictor deficit will be less than one. If the team with the lower possibility of victory wins, the predictor deficit will be greater than one.

- The smaller the predictor deficit, the better the method predicts the winner.

Predictor Deficits

- Look back to the 2008 Florida vs. Louisiana-Lafayette Game 1. We will determine the predictor deficits for that game.

- Florida's Actual Predictor Deficit (APD_H) =

  1 - 0.54589 + 0.45401 – 0 = 0.90803

  Florida's Predicted Predictor Deficit (PPD_H) =

  1 - 0.58084 + 0.41916 – 0 = 0.83833

Predictor Deficits

- Now, we must determine Louisiana-Lafayette's predictor deficits.

- Louisiana-Lafayette's Actual Predictor Deficit (APD_V) =

  1 - 0.45401 + 0.54589 – 0 = 1.09197

  Louisiana-Lafayette's Predicted Predictor Deficit (PPD_V) =

  1 - 0.41916 + 0.41916 – 0 = 1.16167

Determining the Best Method

- To actually determine the best method, one needs to compare all the predicted deficits for both methods.

- Findings

Conclusions

- No clear answer to if the Pythagorean Expectation can be carried over completely to softball.
- No true answer to which method is best in predicting the winner of the WCWS.
- Hypotheses for these conclusions

Recommendations

- Continue project for several more years.
- Do not rule the Pythagorean Expectation out completely; follow over more time.

Want More Information??

- References:

"Pythagorean Expectation". 29 September 2008. http://en.wikipedia.org/wiki/Pythagorean_expectation.

"Archived Team-By-Team Final Statistics". 30 September 2008. http://web1.ncaa.org/stats/StatsSrv/careersearch?searchSport=WSB.

"Bill James". 29 September 2008. http://en.wikipedia.org/wiki/Bill_James.