Testing The Utility Of The Pythagorean Expectation Formula On Division One College Football: An Examination And Comparison To The Morey Model

Cary A. Caro, Xavier University of Louisiana, USA
Ryan Machtmes, Independent Research Consultant, USA

ABSTRACT

The Pythagorean Expectation Formula was the impetus for the statistical revolution of Major League Baseball. The formula, introduced by Bill James, has been used by baseball statisticians to forecast the number of wins a team should have given the total number of runs scored versus those allowed. Since its use in baseball, the formula has been applied to the NFL, the NBA, and the NHL. This study examines if the original formula, as introduced by James, can be fitted for and used to retrospectively predict winning percentage for NCAA Division I football teams. Residual analysis helps the authors conclude that the Pythagorean Expectation Formula provides an accurate prediction of the expected winning percentage for a team given its scoring offense and scoring defense production. Given the formulas predictive ability, coaches and athletic directors can now examine the achievement of their teams and make decisions about filling potential vacancies at college football programs.

Keywords: Pythagorean Expectation Formula; Bill James; Division One College Football; Morey Model

INTRODUCTION

The Pythagorean Theorem Win/Loss formula, or the Pythagorean Expectation (PE), was first introduced by Bill James in the 1980's to estimate the winning percentage for Major League baseball teams at a particular point in the season. The winning percentage formula is based on the number of runs produced and those allowed. The model is presented as:

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

Statisticians have been able to examine the number of games a team "should have won" at a given point in the season. They have also been able to statistical estimate the winning percentage the team should have achieved at the end of the season (Miller, 2007). Many argue that the formula was the impetus for baseball's "Sabermetricians" movement, where, most notably, the Oakland Athletics adopted statistical principles that revolutionized their approach to front office management (Lewis, 2003). The success that the A's experienced was captured in *Moneyball*. The *Moneyball* strategy was simply, "winning games at a minimum cost" (Hakes & Sauer, 2004). This strategy helped the A's, a small-market MLB team, remain competitive against large-market teams with seemingly endless revenue streams.

Starting with the PE formula, the A's then rigorously analyzed baseball statistics in an effort to determine which statistics had the greatest relationship to offensive scoring. Their efforts concluded that on-base percentage and slugging percentage were the best indicators of offensive success; thus, these statistics were used to shape their approach to the MLB draft, to the free-agency market, and it drove many of their personnel management decisions

(Lewis, 2003). However, as Hakes and Sauer (2006) conclude, the A's strategy was about inefficiencies in the economic system of baseball, and the value placed on particular athletes. Simply, the A's were able to exploit the economics of the sport, by finding undervalued, and thus, inexpensive, player attributes.

Since its introduction, statisticians have studied the formula in an effort to find the optimal exponential value for baseball. This value has been reported in a range from 1.82 to 2 (Miller, 2007; Cochran, 2008; Davenport & Woolner, 1999; Cha, Glatt, & Sommers, 2006). Alternative functions for the PE formula have also been explored. For example, Morey applied the PE formula to professional football and found that the optimal exponential value was 2.37 (Morey, 2003). This application has since been used to uncover over and under-achieving teams and forecast their improvement for the next season (Schatz, 2003). Oliver (2004) applied the formula to basketball and found that an exponential value of 14 was most appropriate. Morey (2003), in a publication for STATS, Inc., reported the optimal exponential value in basketball to be 13.91. Cochran and Blackstock (2009) found that an exponential value of 1.927 was the best application for hockey; however, Dayaratna and Miller (2012) concluded that exponents slightly above 2 were also appropriate. Finally, Rosenfeld, Fisher, Adler, and Morris (2005) used Pythagorean Formula full-length game and overtime-game ratios to explore expected outcomes in overtime football, basketball, and baseball games.

Vogel (2012) begins to examine the use of the PE formula to college football. Using Morey's 2.37 exponential value, Vogel examined twenty years of Nebraska football and reports a value of "luck" for each year in his study. The original PE formula, with an exponential value of 2 has not been applied to college football. Using the previous studies as a guide, this study examines the application of James' Pythagorean Win-Loss formula to college football. Using the earlier exponential finding of Morey (2003) in professional football as a guide, this study will utilize the original Pythagorean formula as presented by James, and compare the predicted winning percentages for each formula. This is done for several reasons. First, and most prominent, we seek to examine if the PE formula is applicable to college football. In essence, we seek to examine the usefulness of applying the original PE formula to college football in order to establish the formula's effectiveness as a retrospective tool. Should the PE formula be proven to be an effective tool at predicting winning percentages, it could prove to be an effective tool for athletic directors to use when making coaching decisions.

Moreover, there are different rules that govern game play at the professional and collegiate levels. For example, and most notable, pass interference penalties in college football are marked as a 15 yard penalty from the original line of scrimmage. In the NFL, pass interference penalties are marked where the foul occurred. This rule difference has the potential to influence scoring for NFL teams, and thus influence the application of the PE formula.

Further, professional and collegiate games are timed differently. While the game clocks are the same in both sports, the clocks are managed differently in the last two minutes of each half. As an example, clock stoppages after first downs do not occur in the NFL, which potentially shortens the game, alters decisions made by head coaches, and influences scoring for NFL teams. Finally, and perhaps most dramatic, overtime rules vary between the two sports. The NFL recently altered its previous "sudden death" rule, and now, both teams get a possession if the team that wins the overtime coin toss scores a field goal instead of a touchdown. After that possession, the overtime reverts back to sudden death. In college football, each team receives one possession to begin overtime play. Should the score remain tied after the first possession, the game continues under the same rules until the third overtime. In the third overtime, teams are no longer allowed to attempt a point-after-touchdown field goal, and instead, must try for a two-point conversion. As a seven overtime game between Mississippi and Arkansas in 2001 exhibited, this rule has the potential to change the scoring of a game dramatically.

Given these differences, this study expands on the work of Vogel, who applied the PE formula to Nebraska football, to determine if the original Pythagorean Formula is an accurate and appropriate measure for forecasting the winning percentages of Division I college football teams. We do so while comparing the exponential value used by Morey (2003) to professional football (and later by Vogel to Nebraska football) to the original PE exponent of 2. Using the population of Division I teams from 2005 to 2011, this study uses residual analysis to examine the appropriateness and fit of the PE model. The applicability of an adjusted PE formula for several sports has already been established; however, research into the original formula's applicability as a forecasting tool in college football has been void in the sports academic literature. This study seeks to begin the conversation of research and grow the body of literature with respect to this topic.

METHODS

Statistics for the Division I teams participating in college football for the 2005 to 2011 seasons were obtained from the NCAA online database. Specifically, wins, losses, scoring offense, and scoring defense were collected. Winning ratios (WR) were then calculated for each of the teams using wins divided by the total number of games played (wins plus losses). These data were then used to develop and statistically assess the performance of a novel model [the Pythagorean Expectation formula (formally developed for professional baseball), modified for collegiate football through pertinent variable choice] to predict an expected winning ratio for a team for a season based upon a limited set of inputs, and compare predictive performance of this new model to that of a model established formally for professional football. Scoring Offense (SO) and Scoring Defense (SD) were selected as those mirrored the Pythagorean Expectation (PE) variables of runs scored and runs allowed.

The Pythagorean Expectation formula was modified for collegiate football through pertinent variable choice to reflect the collected data.

$$\frac{Scoring\ Offense^2}{Scoring\ Offense^2 + Scoring\ Defense^2}$$

The value of 2 for the exponent was chosen for simplicity in this formative study of this new model, as it is in reasonable keeping with previous work by Morey (2003) for professional football. The Predicted Winning Ratio (PWR) became the model-predicted response variable for the independent variables Scoring Offense and Scoring Defense. A predicted winning ratio (PWR) was calculated for each of the teams using their earned value of SO and SD for the given year. The Predicted Winning Ratio (PWR) became the model-predicted response variable for the independent variables Scoring Offense and Scoring Defense.

Model performance assessment was determined through calculation of simple variances for a Normal population. For this purpose, an observation was defined as being the total offensive points scored and total defensive points allowed for an individual NCAA Division I team, for a given year. For each observation, a PWR score was calculated for the season. The difference of this model-predicted percentage and the actual winning percentage for each observation was calculated and squared, creating a squared deviation that was summed over all teams per season to create a sum of squared deviations. This sum of squared deviations for a season was divided by the number of teams that played collegiate football that season, creating a simple population variance. Model performance could be assessed and compared with that of competing models using this methodology; the lower the value of this variance, the less the amount of total error that might be ordinarily expected in application of the developed model for purpose of predicting a team's winning ratio.

Population variances for Normal populations were deemed appropriate for this study as we calculated this variance based upon calculating a PWR for all NCAA Division I teams that fielded a team for the given year rather than a random sample of such teams. The distributions of these deviations from the model-predicted PWR score were assumed to be Normal, as it seemed reasonable as a first approximation to assume the PWR an unbiased estimator of the actual winning percentage for a team. Under this assumption, differences between this predicted value and the observed value should occur randomly towards a team's under- or over- performance for a season from this predicted value and decrease in occurrence with continuing extremity of such under- or over- performance with increasing magnitude of the difference. This methodology was also applied to an application of the Morey model developed previously for application to professional football to predict overall team performance for a given year.

RESULTS

A total of 835 teams, across seven years, were included in the study. There was an expansion from 118 teams in 2005 to 119 in 2006. This number remained constant until 2009 when one more team was added to bring the total to 120. There are a total of 835 teams included in this study across seven years.

The performance statistics for the model developed herein, based upon adapting the Pythagorean Expectation formula for baseball to collegiate football, are summarized in Table 1. Performance statistics for the existing model of Morey for professional football, as applied to the same data for NCAA Division I football teams, are provided in Table 2. Both models predict team performance for a season as an expected Win-Loss ratio based on the model inputs. Performance of the Pythagorean Expectation formula-based model was found to have a total variance each year between 2005 and 2011 between 0.00675 and 0.00871, while the Morey model for professional football as applied to collegiate football data was found to have a total variance for the given years between 0.00621 and 0.00862.

Table 1: Model Performance Statistics for Pythagorean Expectation Model

Year	Sum of Deviation^2	N	Residual Variance
2005	0.9228	118	0.007820
2006	0.9657	119	0.008115
2007	0.8115	119	0.006819
2008	0.8036	119	0.006753
2009	1.049	120	0.008699
2010	0.9617	120	0.008014
2011	1.0451	120	0.008709

Table 2: Model Performance Statistics for Morey Model

Year	Sum of Deviation^2	N	Residual Variance
2005	0.9697	118	0.008218
2006	1.0077	119	0.008468
2007	0.7400	119	0.006219
2008	0.8177	119	0.006872
2009	1.0067	120	0.008389
2010	0.9948	120	0.008290
2011	1.0345	120	0.008621

DISCUSSION

As shown in the analysis of the residual variances of applying both the adapted Pythagorean Expectation formula and the Morey model as originally developed for professional football to the population of all NCAA Division I football teams between the years 2005 and 2011, both models appear to predict with an error of magnitude 10^{-3} the expected winning performance of a given team for a given season. We believe then that coaches of NCAA football teams could utilize either model to their advantage for assessing their team's performance over the previous season, in terms of determining whether their team successfully met expectations or exceeded them, or failed to do so. Each model presents interesting questions for further study, and opportunities for further development. For purposes of this initial study, we have chosen to focus upon examining the utility of the adapted PE formula, as it is the more novel of the two models for application to NCAA Division I football.

The Pythagorean Expectation formula, as presented by Bill James, is a measure of how many games a baseball team should have won given their total offensive and defensive production. This study sought to determine how appropriate the PE formula would be at forecasting winning percentages in Division I college football. Based on the outcome of this study, the data suggests that PE provides a measure of expectation based on the statistical performance on the team. With a performance (residual variance) of order 10^{-3}, the PE formula can be acknowledged to be an accurate forecasting method for coaches and athletic directors to use when evaluating the performance of their football programs.

The PE forecast is important for football coaches as they reflect upon the season that was just completed. A coach can, retrospectively, compare the forecast wins that his team was expected to achieve and compare this to the wins actually achieved. This provides coaches an opportunity to examine various aspects of their program. Overall, it provides coaches a measure of achievement for the season. If the predicted PE winning percentage is less than the actual winning percentage, a coach can come to understand if his team has "overachieved" given the expectations of the PE model. Conversely, if the PE forecast is greater than the actual winning percentage, this can be an indicator of an "underachieving" team. Theoretically, this could then help coaches review decisions made throughout the course of the season, specific to scoring opportunities, fourth down attempts, turnovers, and clock management.

The concept of over/under achieving is also important for athletic directors when examining coaching decisions. For example, in 2010 South Florida (USF) had a predicted winning percentage of 0.619 and an actual winning percentage of 0.615. In the 2011 season, they had a predicted winning percentage of 0.647 and an actual winning percentage of 0.412. Finally, the predicted winning percentage for 2012 was 0.340; however, USF's actual winning percentage for the 2012 season was 0.250. Using USF as an example, the data gathered from applying the statistical model would show that Coach Holtz' teams at USF had underperformed over the last three years. An athletic director could use a history of underperformance to inform and further justify a potential action on a coaching decision (Coach Holtz was fired after the 2012 season).

The importance of maximizing possessions is illustrated using the PE model. The relationship between scoring offense, scoring defense, and winning is not a new one. At the end of the game, the team that has the most points is going to always be victorious; however, how a coach is able to weigh decisions about the pace and tempo of the game, decisions to go or to punt on fourth down, and to take field goal points instead of a fourth-down opportunity all factor into the points scored category. For example, in the 2013 season, Louisiana Tech, Oregon, Oklahoma State, and Baylor ranked one through four respectively in scoring offense; of those, only Oklahoma State was not in the top twenty in fourth-down conversion percentage. The argument can be made that these coaches understood the importance of extending a possession in order to have an opportunity to score a touchdown instead of attempting a field goal or punting the ball. Thus, taking advantage of opportunities to extend a possession can potentially lead to more scoring opportunities, and can impact a team's winning percentage. During the course of a season, a coach can input their current scoring totals and make decisions about the total number of points that are needed in order to achieve a specific winning percentage. This output can then influence in-game decisions, especially early in a game, before decisions become largely dictated by the score and the amount of time left in a game.

The PE model helps to uncover coaches whose teams consistently over- or under-perform. For example, a coach who's PWR is consistently greater than his AWR may be an individual who has all of the available resources to win (athletes, facilities, etc.) but is unable to actually win games. This could be due to poor in-game decisions, bad luck, a tough schedule, or a myriad of other variables. Conversely, a coach who consistently out-performs the PWR may be an individual who maximizes the potential of his athletes, who makes above average in-game decisions, or who is able to inspire and motivate his athletes to perform beyond their own expectations. This candidate may be better suited to fill a potential job vacancy.

CONCLUSIONS AND FUTURE RESEARCH

Building upon the work of Bill James and later Daryl Morey, this study aimed to examine the utility of the Pythagorean Expectation model to retrospectively forecast performance for Division I college football teams, and compare the performance of the adapted PE formula to that of the Morey model for professional football. Using residual analysis, the results indicate that both the adapted PE and Morey models provide valuable metrics for examining expected versus actual team performance. Either model provides a simplistic approach to examining winning percentages and expectations.

There are many avenues of potential research from this study. While the model variables have been identified, the supporting variables need to be addressed. Analysis into scoring offense and scoring defense must take place in order to better define the game-play statistics that lead to increased wins. With the myriad of statistical measures for everything from rushing offense to third down efficiency, researchers should take advantage of factor analysis or logistic regression to identify the game-play variables that most heavily impact points scored and points allowed. Sensitivity analysis of the exponents and coefficients used in conjunction with the model variables should also be addressed via analysis of the residual variances as done in this initial study, to further improve model performance.

As previously noted, the PE model is strictly a retrospective model and cannot forecast future performance; thus, the utility of the model is severely limited. Further, the simplicity of the PE formula cannot be overlooked. A measure that simply examines total points scored and total points allowed, while accurate, does not help to inform coaching decisions, recruiting, game-play, or other measures of the game. Thus, given the increase in the availability of football statistics, it would be worthwhile for the academic literature to explore the development of a model for forecasting future performance. Given the bevy of statistical measures, such a model would be an invaluable source of information for athletic directors and coaching as they approach their respective seasons. Nonetheless, the value of a statistical model to examine performance, much like that presented in *Moneyball* cannot be overlooked by coaches and athletic directors.

AUTHOR INFORMATION

Cary A. Caro, Ph.D., Xavier University of Louisiana, 1 Drexel Drive, New Orleans, LA 70125 USA. Professor Caro is assistant professor of management at Xavier University of Louisiana and managing editor of the *American Journal of Business Education*. E-mail: ccaro@xula.edu

Ryan Machtmes has a Masters in Applied Statistics. E-mail: machmer@gmail.com

REFERENCES

1. Cha, D., Glatt, D., and Sommers, P. (2006). An empirical test of Bill James' pythagorean formula, *Journal of Recreational Mathematics*, 2006, *35*(2).
2. Cochran, J.J. (2008). The optimal value and potential alternatives of Bill James' pythagorean method of baseball, *StatOR*, *8*(2).
3. Cochran, J., and Blackstock, R. (2009). Pythagoras and the National Hockey League, *Journal of Quantitative Analysis in Sports*, *5*(2), 11.
4. Davenport, C. and Woolner K. (1999). Revisiting the pythagorean theorem. Retrieved from Baseball Prospectus: http://www.baseballprospectus.com/article.php?articleid=342
5. Dayaratna, K. D., and Miller, S. (2012). The pythagorean won-loss formula and hockey: A statistical justification for using the classic baseball formula as an evaluative tool in hockey, Unpublished manuscript, University of Maryland.
6. Hakes, J. K., & Sauer, R. D. (2004). Measurement and evaluation of managerial efficiency in MLB, Unpublished manuscript, Clemson University.
7. Hakes, J. K., & Sauer, R. D. (2006). An economic evaluation of the Moneyball hypothesis, *The Journal of Economic Perspectives*, *20*(3), 173-185.
8. Lewis, M. (2003). *Moneyball*, W.W. Norton & Company, Inc., New York, NY.
9. Miller, S.J. (2007). A derivation of the pythagorean won-loss formula in baseball, *Chance Magazine*, *20*(1).
10. Morey, D. in Dewan, J., and Zminda D. (2003). STATS, Inc. Staff, STATS Basketball Scoreboard, STATS, Inc., 17.
11. Oliver, D. (2004). *Basketball on Paper*, Potomac Books, Dulles, VA.
12. Rosenfeld, J., Fisher, J., Adler, D., and Morris, C. (2005). Predicting Overtime with the Pythagorean Formula, *Journal of Quantitative Analysis in Sports*, *6*(2).
13. Schatz, A. (2003). Pythagoras on the gridiron. Retrieved from Football Outsiders: http://www.footballoutsiders.com/stat-analysis/2003/pythagoras-gridiron
14. Vogel, B., Nebraska and Football's Pythagorean Theorem. (2012) Retrieved from Hail Varsity: http://hailvarsity.com/2012/08/nebraska-and-footballs-pythagorean-theorem/

