Predicting Outcomes of NCAA Basketball Tournament Games

Aaron Cripe

March 12, 2008

Math Senior Project

Introduction

All my life I have really enjoyed watching college basketball. When I was younger my favorite month of the year was March, solely because of the fact that the NCAA Division I Basketball Tournament took place in March. I would look forward to filling out a bracket every year to see how well I could predict the winning teams. I quickly realized that I was not an expert on predicting the results of the tournament games. Then I started wondering if anyone was really an expert in terms of predicting the results of the tournament games. For my project I decided to find out, or at least compare some of the top rating systems and see which one is most accurate in predicting the winner of the each game in the Men's Basketball Division I Tournament. For my project I compared five rating systems (Massey, Pomeroy, Sagarin, RPI) with the actual tournament seedings. I compared these systems by looking at the pre-tournament ratings and the tournament results for 2004 through 2007. The goal of my project was to determine which, if any, of these systems is the best predictor of the winning team in the tournament games.

Project Goals

Each system that I compared gave a rating to every team in the tournament. For my project I looked at each game and then compared the two team's ratings. In most cases the two teams had a different rating; however there were a couple of games where the two teams had the same rating which I will address later. Since the two teams had different ratings this created a favored team (which is the team that should win according to the rating system), and non-favored team or underdog (which is the team that should lose the game according to the rating). Differing rating systems might have a different favored team for each game. Of course there are going to be some games when every rating system has the same favored team, but there will also

be games when the systems could be split in terms of who the favored team is. For my project we computed the relative strength of the favored team versus the underdog. If a rating system has any predictive power then the higher the favorites' relative strength, the more likely they are to win. The relative strength of the two teams can be determined two different ways, ratio and difference. Relative strength ratio can be computed by taking the favored team's rating and dividing it by the non-favored team's rating. The relative strength difference can be computed by taking the favored team's rating and subtracting the non-favored team's rating. When the favored team has a much higher rating than the non-favored team, the relative strength will be large. This occurs in the tournament when a one seed plays a sixteen seed, two seed plays a fifteen seed, three seed plays a fourteen seed, or a four seed plays a thirteen seed. However, when the two teams have a similar rating the relative strength of the favored team is much smaller. This occurs in the tournament when an eight seed plays a nine seed, seven seed plays a ten seed, or a six seed plays and eleven seed. A lower relative strength means that the favored team is less likely to win than a favored team with a large relative strength. Another property of a useful rating system is that the favored team should win at least 50% of the time (because if you just pick one team at random there is still a 50% that team will win), even if the relative strength is extremely small. If a rating system has less than a 50% accuracy rate, then it has no ability to predict which team will win each game. We are hoping that we can find out that every system has a much higher accuracy rate than 50% because that would then tell me that there are valid methods for predicting the winner of each game. We also wish to determine if certain methods are better at predicting winners. Then we will develop models for each method to estimate the probability a favored team wins, given their relative strength versus their opponent.

NCAA Tournament Selection Process

Before going into further detail about each rating method that we studied it is important to understand the procedure for determining which teams make the NCAA tournament. There is a ten-member committee chosen by the NCAA made up of athletic directors and conference commissioners. However, the athletic director is not allowed to discuss his/her own team unless someone else directs a question to them. Also the conference commissioner is not allowed to discuss the teams in their conference but is entitled to update the committee about each team. Updates include any injuries, suspension, or anything else that may be relevant. The committee is solely responsible for choosing which teams make the tournament. The actual process is very detailed and extremely tedious. There are thirty teams that receive an "automatic bid" to the tournament. The only way to receive this automatic bid is to win your conference tournament. The Ivy League, however, does not have a conference tournament so the team that wins their conference season gets the "automatic bid." Thus the committee is in charge of determining the rest of the teams that make up the tournament, seeding the teams, and creating the brackets. The other teams that make the tournament receive one of the 34 at-large bids. They did not win their conference tournament however they are picked to be in the tournament by the committee. Most of these at-large bids come from the teams in the most prestigious conferences. It is less common for a small school to receive an at-large bid. The committee has access to any rating system of their choice when deciding which teams make the tournament. Also the national polls and coaches' poll play a large factor in deciding which teams make the tournament. The teams that are in the top twenty-five will almost always make the tournament. When deciding which teams receive the at-large bids the committee is supposed to take into account how good that team is at the start of the tournament to decide which teams should make it and which teams

should not. The committee is allowed to take into account any injuries of suspensions that any

given team may have at the start of the tournament. There are four (possibly more) aspects that

the committee cannot use to determine which teams get the at-large bids. Those four are:

geography of the schools, the coach of the school, the prestige of the school, or the style of the

play of the team. The committee is allowed to use anything else when determining which teams

make it. The committee is also in charge of giving each team a seed. The tournament is a 65

team tournament and two teams have to play in what is known as the "play in game" to decide

which team plays in the 64 team tournament. The 64-team set up consists of four regional

brackets where each section consists of 16 teams seeded from 1 to 16. The team with the one

seed is the "best" team in that region and the team with the sixteen seed is the "worst" team in

that region according to the committee. Like other single-elimination tournaments the top seed

playing the lowest seed, then the next best seed plays the second to last seed, etc. Each region

has one team go through without a loss and those winning teams make up what is known as the

Final Four. Then it is played just like any four-team tournament single-elimination, until one

team is deemed the National Champion. Before explaining the actual seeding process, we now

discuss the rating systems to be investigated (NCAA 1).

Rating Systems

RPI

Ratings Percentage Index or RPI as it is more commonly referred to, is one of the most

widely known rating systems. It has been rumored that it is the most influential system in

regards to the committee deciding which teams make the tournament. This system has been in

use since 1981. RPI consists of three main parts: 1) the team's winning percentage (25%), 2) the

team's opponents' winning percentage (50%), and 3) the winning percentage of those opponents'

opponents (25%). The third part represents the schedule strength of the teams that your team

plays. The RPI rating is determined as a weighted average of these three:

$$RPI = (.25 * (1)) + (.50 * (2)) + (.25 * (3))$$

The major criticism of RPI is that this system places to much emphasis upon strength of schedule

and unfairly advantages teams from major conferences. The teams in major conferences are

forced to play against certain schools, which usually are schools of high caliber. The smaller

schools are forced to play the schools in their conference that are usually also small schools.

Since RPI is so heavy looked at by the selection committee the smaller schools are trying to get

the bigger schools to play them during the non-conference portion of the season. Small schools

will go as far as even paying the bigger schools to allow them to play them. It is a big risk for a

big school to play a talented smaller school because a loss will adversely affect their RPI while a

win provides little benefit. Thus this is a weakness of the RPI system because not every team

has complete control of their schedule (Pomeroy(2) 1).

Since 2004 the RPI has added one more aspect into the systems' ratings. The idea of

home court advantage has been taken into consideration. For example, a home win now counts

as 0.6 win, while a road win counts as 1.4 wins. Conversely, a home loss equals 1.4 losses,

while a road loss counts as 0.6 losses. This change is one of the main reasons we only look at

the tournaments as far back as 2004 (Pomeroy(2) 1).

Pomeroy

The Pomeroy rating system is done by Ken Pomeroy. He states that his system is

designed to be purely a predictor. He claims that he does not rate the teams on how successful

their season has been, but instead by how strong he feels that team is if they played tonight, independent of injuries and emotional factors. Thus as you can imagine there are a lot of aspects about his system that he does not like to reveal to the public. However I will try and describe his system as I best understand it. Pomeroy looks at a game that the team won and then looks to see how much they won by. Similarly, for a loss he looks to see how much they lost by. There is more emphasis placed on a 20-point win than say a 5-point win. This system gives a higher rating to a team that lost a lot of close games against strong opposition than one that wins a lot of close games against weak opposition. Thus there could be a team in a major conference with a losing record that has a higher rating than a team in a small conference with a winning record. Pomeroy's system also takes consideration into home court advantage. To sum Pomeroy's system up he takes into account margin of victory, strength of opponent, schedule strength, offensive efficiency, defensive efficiency, and how good your opponents' offense and defense is. His system is very mathematical and very complex. Ken Pomeroy is one of the few who truly understands how his system works. One of the weaknesses of his system is that he keeps making changes on his system from year to year. He has not yet had a chance to compare his system over a long period of time because he makes little changes from year to year always trying to make his rating system better. This could be considered a strength of his system because each year he is trying to make his system a stronger predictor (Pomeroy(1) 1). .

Sagarin

Sagarin's system is done by American sports statistician Jeff Sagarin. His rating system has been used by the selection committee since 1984 to help determine which teams should or should not make the tournament and, if they do, what seed they should be. It is also a widely known fact that Sagarin does not like to reveal his exact methodology behind his system.

Sagarin's system is made up of two different systems which he then combines to form his actual rating system for college basketball. The first system that he calls "ELO Chess" (based on the ELO rating system used internationally to rank chess players). That system only takes into account wins and losses and has no reference to margin of victory. The other system he calls "Predictor" takes victory margin into account. This system is meant to predict the margin of victory for the stronger team at a neutral site by the difference in two teams' rating scores. The rating that I used is a combination of the two; however I was unable to find how he combines them. I know that he does not simply take the average of the two. I would assume that he weights one of the systems higher than the other. Also Sagarin's system favors teams that lose close games against stronger opponents versus teams that win close games against weak opponents. For both systems, teams gain higher ratings within the Sagarin system by winning games against stronger opponents, factoring in home-court advantage. Thus, in that aspect this system is similar to Pomeroy's. Even though Sagarin's system is very secretive and not a lot is known about it, it is a very widely used and respected system (Schatz 1).

Massey

The fourth rating system that we investigated is the Massey Rating System. This system is a computer rating system done by Kenneth Massey. No different than other rating systems, Massey does not like to divulge his exact methods to the public. However, we do know that Massey inputs the score, venue, and date of each game to come up with his ratings. Massey does not use any statistics such as rebounds or steals to come up with his ratings. He goes on to say that it is impossible to include all aspects of the game such as injuries, suspension, or how the ball bounces on any given night. However, Massey feels that the score of the game will usually include those various components. He believes that, in a sense, will include the score when

finding his ratings because those outside factors will affect the score of the game. Massey believes that his rating system is designed to measure the teams' past performance, and it is not an accurate predictor of future outcomes. Massey's rating system uses a rating scale where the highest team is around a 2.5 and the lowest teams are usually in the negatives. One difference that Massey's system has compared to the others is that he includes all the divisions of basketball and not just Division I (Massey 1).

Table to summarize the different rating systems.

	Margin of Victory	Home Court	Strength of Schedule	Strength of Opponents Schedule	Record (win and losses)	Range of Ratings
RPI	No	Yes	Yes	Yes	Yes	0.0 to 1.0
Pomeroy	Yes	Yes	Yes	No	Yes	0.0 to 100
Sagarin	Yes	Yes	Yes	No	Yes	0 to 100
Massey	Yes	Yes	No	No	Yes	-4.0 to 4.0

Figure 1

Tournament Seeds

The selection committee's seeding of each team can also be considered a type of rating systems. The selection procedure is extremely complex and not completely revealed to the public. However, the selection committee first decides which 65 teams make the tournament before they even think about seeds. Then they have to, in a sense, rank the teams from 1 to 65, this process is the most mysterious of the seeding selections. After the committee has the teams ranked in order than the process is quite simple and systematic. The teams that are ranked 1 through 4 will all receive 1 seeds, and teams 5 through 8 will all receive 2 seeds, and so on until

there are four teams of each seed from 1 to 16. Then they have to determine what teams go into what regional bracket, as discussed earlier. They take the number 1 overall seed (the team with the 1 seed when seeded 1 through 65) and put that team into a regional bracket. Then they place the other number 1 seeds in different regions. Thus if the Final Four occurred as the seedings predict the top number 1 seed would play the fourth number 1 seed, while the second number 1 seed would play the third number 1 seed, in the final four. Then they divide the four number 2 seeds into the regions in a similar way. They take the fourth number 2 seed and put them with the top number 1 seed, and the top number 2 seed with the fourth number 1 seed. They do this accordingly for the rest of the seeds, but each time alternating which region to ensure that every region is equal in difficulty according to the selection committee. While the selection committee is known for doing a good job when assigning seeds to each team, it is inevitable that there will still be upsets in every tournament. However, in all the years of the NCAA tournament a number 1 seed has never lost to a number 16 seed. There have been a couple number 2 seed schools to lose to a number 15 seeded team.

Goals, Problems, and Assumptions

We analyzed all five of these rating systems to determine which one is the most accurate predictor of the winner of the tournament games. The five rating systems sometimes make changes to keep the systems up to date. This caused me to only use the data as far back 2004 because in 2004 RPI rating system made some changes, taking home court advantage in to account. We wanted to make sure that the rating system was consistent from year to year. Also when we were looking at the Massey rating systems we noticed that in 2005 the range of the ratings was different than the range of the other years. For example the top team in 2005 had a rating of 6.5, while the top team in the other three years had a rating around 2.5. However, the

lowest team's ratings were really similar amongst all the years, which caused a much higher range in year 2005. This change in the range is a problem for our data because we are using the ratio and difference of the ratings. The ratio and differences would be drastically different in year 2005. Thus we decided that I needed to re-scale the 2005 Massey ratings to align the ratings with the rest of the years. After consulting with Kenneth Massey we were able to re-scale the data so it was comparable to the other years. This also occurred with Pomeroy's rating system. In 2007 he used a different method which resulted in different ratings than in his previous three years of data. This time we were unable to manual change the data to correspond with the other three years so our only option was to omit the 2007 data of Pomeroy. Now instead of 252 games of data for Pomeroy, we used 189 games. Also there was the opportunity for a team to have the same rating as its opponent. This is very possible for the tournament seeds, since there are four regional brackets with essentially the same seeds. However the only time this could occur would be during the Final Four, when each regional bracket has one team emerge and play the winners of the other regional brackets. When looking at all five rating systems over four years of tournament games we ran into three ties. Two out of the three ties occurred with the seeds and the other was in 2005 with the RPI rating system. We decided to simply not use those games in our study since neither team is considered to be the favorite. We omitted the data of that one specific game only for that specific rating system.

Rating Systems' Predicting Ability

Rating System	Success rate for predicting winner
Seeds	73.2%
Pomeroy	73.0%
Massey	72.6%
Sagarin	72.2%
RPI	70.9%

Figure 2

The table in Figure 2 provides the percentage of time that the favored team wins according to each rating system. It is calculated by taking the total number of games predicted correctly divided by the total number of games. If the rating systems cannot predict a winner with an accuracy higher than 50% that means that the system has no merit, because you can just simply flip a coin and have the same accuracy. The actual seeds rating system was the most accurate in terms of just predicting which team would win at 73.2% of the time. However, when looking at Figure 2 you can see that the five rating systems all are within 3% of one another which leads me to believe that there is not a lot of difference in the systems in terms of purely predicting the winner. You can use the hypothesis test for one sample proportion to determine whether or not a rating system can predict the winner with accuracy higher than 50%. We, let our null hypothesis to be the proportion predicted correctly equals .50. Our alternative hypothesis is the probability proportion greater than .50 (i.e. the rating system is effective at predicting winners). According to SPSS 15, RPI had the lowest percentage of predicting the winner at 70.9%. Thus for the hypothesis test we used RPI because if it passes the test it is

obvious that the other systems will too. A z-score is computed by comparing the sample proportion of .709 to .50 then we obtain a z-score of 6.62. Since z is greater than 3 we know the corresponding P-value is below .05. From this we can conclude that RPI is effective at predicting at a higher accuracy than 50%. As a result we can conclude that all of the other rating systems also are effective at predicting the winner.

Another question that needs to be addressed is whether one system is "better" than any other system at predicting the winner. For this we can use a hypothesis test comparing two sample proportions. We will use Seeds (because according to figure 2 it is the highest) as sample 1 and RPI (because according to Figure 2 it is the lowest) as sample 2. If the hypothesis test shows that there is no difference between these two systems then we can conclude that there is no difference between any of the systems. We will assume the true proportion p_1 and p_2 are equal to each other for our null hypothesis. Then for our alternative hypothesis we will say that they are not equal. If we calculate a p-value that exceeds .05 then we can conclude that our alternative hypothesis is false. A z-score comparing .732 and .709 results with $z = .575$, resulting in a p-value of .562 which is greater than .05. Thus, there is no evidence to conclude that any of the rating systems are different in terms of the likelihood of predicting the winner.

Predicting Game Outcomes using Relative Strength

The overall goal of my project is to use each rating system to develop a model for the probability that the favored team wins; furthermore, we will indentify which model makes the most accurate predictions of game outcomes. In order for us to find the probability of winning based on each rating system we used a logistic regression model. Logistic regression models the natural log of the odds as a linear function of some independent variable *x*:

1) $\ln \frac{p}{1-p} = \beta_0 + \beta_1 x$

Odds of winnings are defined as $\frac{p}{1-p}$ where p is the probability the favored team wins. We will

use relative strength as our predictor variable x. The parameters of the logistic model are β_0 and

β_1. Solving equation (1) for p, we obtain:

2) $p(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$

Here $p(x)$ is the probability that a team with relative strength x (compared to their opponent) will

win the game. The parameters of the model, will be found using SPSS 15. This model will form
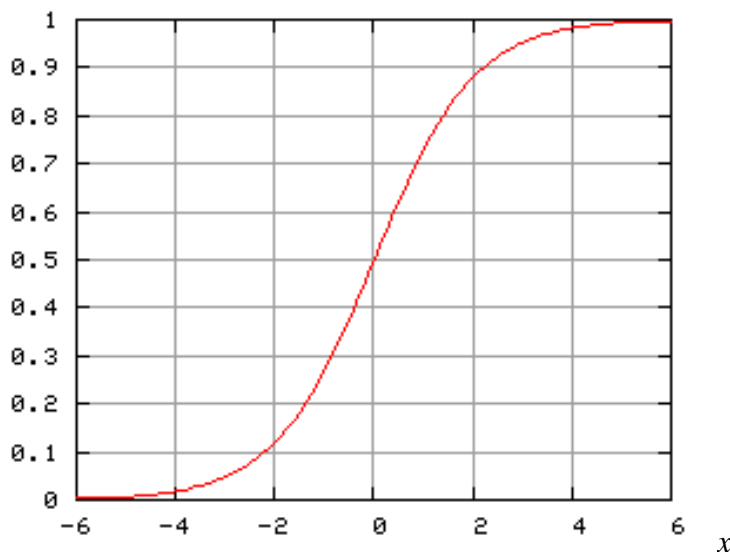
what is known as a logistic curve.

p(x)

Figure 3

The above picture is an example of a typical shape of a logistic curve. This shows as x increases

(that is relative strength) the probability of winning $p(x)$ approaches 1. As the relative strength

decreases the probability of winning approaches 0, and as relative strength approaches 0 (meaning no difference in the rating of the two teams) p(x) logically approaches .5 or 50%. For us the only part of Figure 3 that is relevant is the positive *x* values which represent the relative strength of the favored team. We will be comparing relative strength two different ways by difference and ratio. The relative strength by difference is found by subtracting the non-favored team's rating from the favored team's rating. Two teams that have the same rating would each have a relative strength of zero in terms of difference, as shown in Figure 1. However, for our investigation the relative strength *x* will be positive. The relative strength by ratio is found by dividing the favored team's rating by the non-favored team's rating. Thus, two teams with equal rating will have a relative strength of one rather than zero when using differences. When two teams have the same rating it impossible to label one of them the favored team. Thus, there is no way of predicting the winner better than by just simply flipping a coin. Each team has the same chance of winning the game. That is why it was important for us, in the project, to make sure that the modeled logistic curve goes through the point that represents a 50% chance of winning when two teams have the same rating. In order to insure this happens we can force β_0 to be zero, which is permitted in SPSS. The logistic model then simplifies to:

$$3) \quad p(x) = \frac{e^{\beta_1 x}}{1 + e^{\beta_1 x}}$$

SPSS will construct the function p(x) for each rating system using both the relative strength with ratio and with difference. We hope the data will show that the higher the relative strength the more likely the favored team is to win. Furthermore the logistic regression model will help me determine which rating system is the most accurate predictor of the probability of winning for the favored team (McCabe, Moore Chapter 16).

After analyzing the data from SPSS we were able to construct a table of the numbers that are particularly relevant to the project (see appendix 1). The table consists of β_1, the P-value, Cox and Snell R^2, and the percentage of games that the favored team won.

	β_1	Sig. (P-Value)	Cox and Snell R^2	% of games fav. team won
Sagarin RS_Df	.206	.000	.312	72.2
Sagarin RS_Rt	.930	.000	.202	72.2
RPI RS_Df	24.068	.000	.271	70.9
RPI RS_Rt	.864	.000	.181	70.9
Pomeroy RS_Df	.219	.000	.297	73.0
Pomeroy RS_Rt	.953	.000	.217	73.0
Massey RS_Df	3.078	.000	.278	72.6
Massey RS_Rt	.867	.000	.231	72.6
Seed RS_Df	.200	.000	.301	73.2
Seed RS_Rt	.472	.000	.293	73.2

Figure 4

The first concern we had was if any of these rating systems provide a statistically significant predicator of the winner. That is, the relative strength is useful when estimating the probability the favored team wins. In order to tell if the rating system is statistically significant we look at the associated p-value column. If the p-value is less than 0.05 then there is statistical evidence of a relationship between *p(x)* and relative strength. As you can see from the table, all five ratings systems provide significant results for relative strength using both differences (RS_Df) and ratios

(RS_Rt). The Cox and Snell R^2 shows how accurate the relative strength is in predicting *p(x)*.

If relative strength were 100% accurate in predicting the probability of the winner then the R^2

rating would be 1. The Cox and Snell R^2 is extremely important to the project because that

number shows how useful relative strength is, which exactly is the goal of the project. Thus, as

you can see from the table, Sagarin difference is the most accurate when analyzing the relative

strengths with the highest R^2 of .312. Roughly speaking, $R^2 = .312$ indicates, 31.2% of variation

in the favored team's chance of winning can be explained by knowing their relative strength
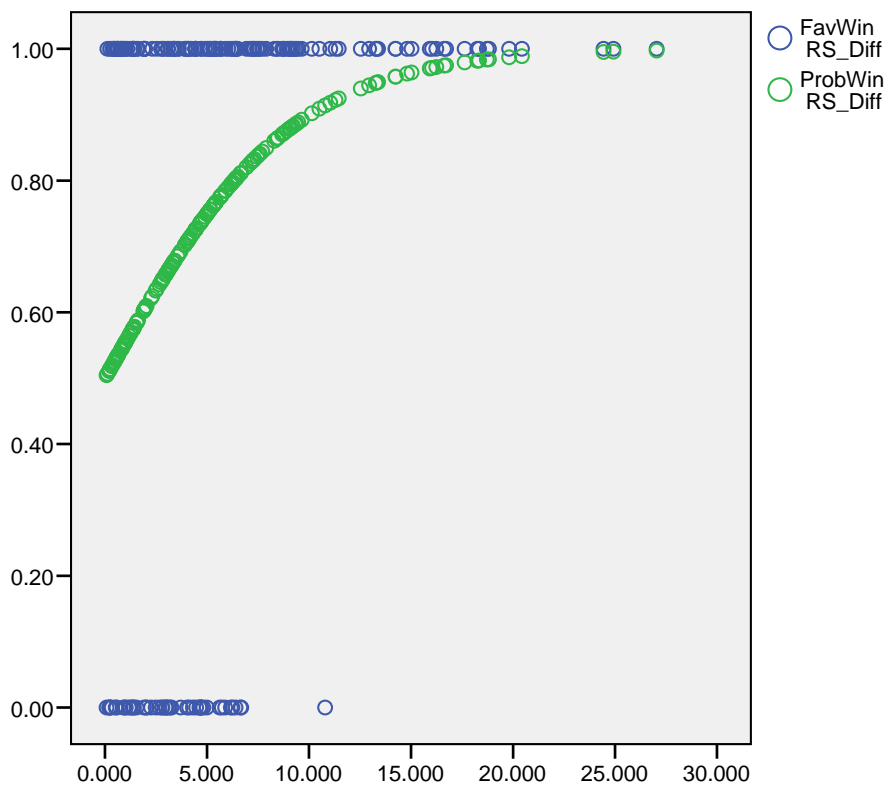
according to Sagarin's system.

Based on Figure 4 we see that Sagarin_Df ($R^2 = .312$), Seeds_Df ($R^2 = .301$), and

Pomeroy_Df ($R^2 = .297$) provide the most accurate predictions of the probability of winning.

Based on the 4 years of tournament data, that we examined, the best three models are:
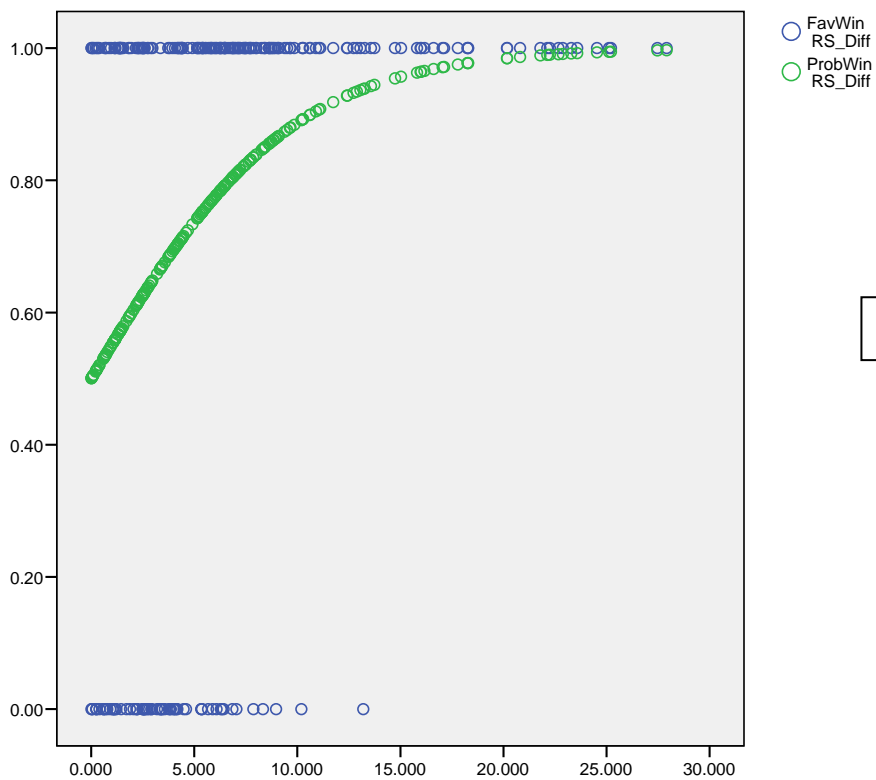
Sagarin_Df- $$p(x) = \frac{e^{.206x}}{1 + e^{.206x}}$$
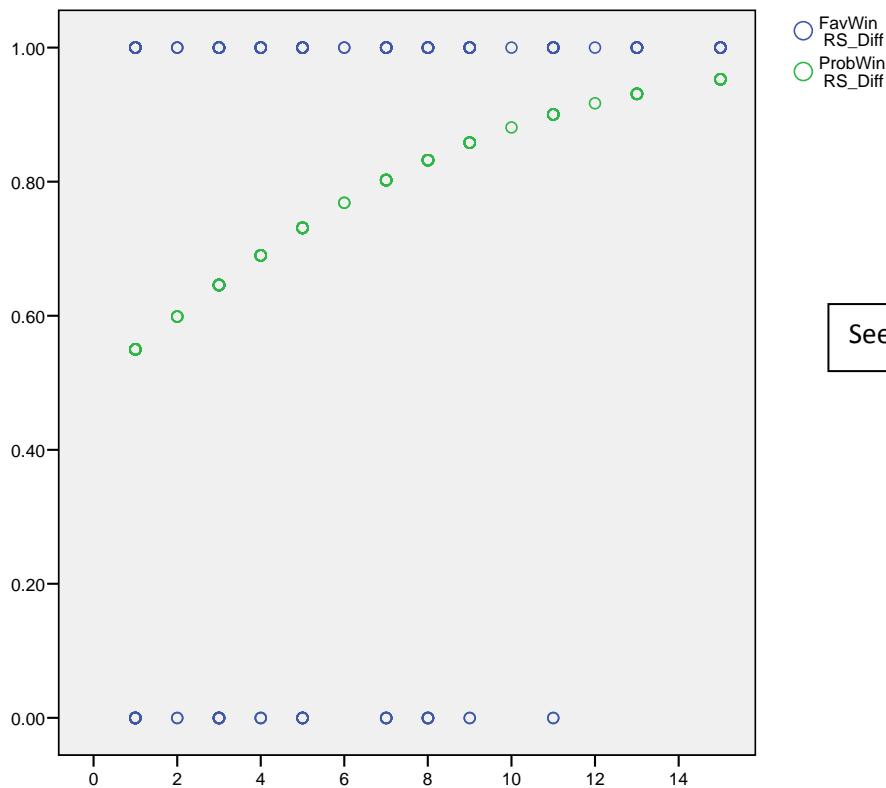
Seeds_Df- $$p(x) = \frac{e^{.200x}}{1 + e^{.200x}}$$

Pomeroy_Df- $$p(x) = \frac{e^{.219x}}{1 + e^{.219x}}$$

Pomeroy RS Df

Sagarin RS Df

These three graphs above are representations of the three most accurate predictors of the winning teams. The graphs show two types of scatter plot. The first is the points either on zero or one, where one is where the favored team won, and zero is where the favored team lost. The other curved graph is taking the data and plotting it on the logistic model. As you can see from the three graphs above, as relative strength increases so does a team's chance of winning. For example, consider 2004 Elite Eight Alabama vs. Connecticut. Connecticut is a 2 seed in the tournament and Alabama is an eight, and in the project we switched the ratings so the favored team would have a higher seed, thus Connecticut a rating of 15 and Alabama a rating of 9. Thus the relative strength using Seed_Df would 6 for Connecticut. Thus if you look at the graph of Seed_Df for $x = 6$, you can see that the probability of winning for Connecticut is just below 80%. That means that Connecticut has just below an 80% chance of winning the game

according to the Seeds_Df ratings. It turns out that Connecticut did, indeed, win the game and end up winning the National Championship in 2004.

In these models e^{β_1} represents the factor increase in odds of winning when relative strength increases 1 unit. This can be illustrated using a three-team tournament where team A has a rating of 4, team B a rating of 3, and team C a rating of 2. If team A plays team B (say in Game #1) then team A will have a relative strength Df of 1. Thus if we plug that into the formula 3 and use the Sagarin_Df model $\dfrac{e^{.206x}}{1+e^{.206x}}$ we obtain a $p(1) = .5513$. Then if we plug our $p(1)$ into our odds formula $\dfrac{p}{1-p}$ we result with 1.23. Which means that a team with relative Df of 1 according to Sagarin, has 1.23:1 odds of winning the game. Then if team A plays team C (say Game #2) then team A will have a relative strength Df of 2. Using formula 3 and Sagarin Df (β_1=.206) we obtain a $p(2) = .6016$. Then plugging that into the same odds formula we obtain odds of 1.51:1. Now we can calculate what happens to a team's chance of winning when their relative strength Df increases by 1. By dividing the higher odds by the lower odds we obtain 1.228, which means that when a team's relative strength Df increases by one their odds of winning increases by a factor of 1.228. To prove this we should look at the original formula:

$$\ln \frac{p}{1-p} = \beta_1 x \quad \text{where relative strength} = x$$

Then by raising both sides to the power of e, and we obtain:

$$\frac{p}{1-p} = e^{\beta_1 x} \quad \text{where relative strength} = x$$

Thus when we adding one to the relative strength we obtain:

$$\frac{p2}{1-p2} = e^{\beta_1(x+1)} \text{ , where } \frac{p2}{1-p2} \text{ represents the odds Team A wins Game \#2}$$

$$\frac{p2}{1-p2} = e^{\beta_1 x + \beta_1}$$

$$\frac{p2}{1-p2} = e^{\beta_1 x} * e^{\beta_1}$$

$$\frac{p2}{1-p2} = e^{\beta_1} * \frac{p1}{1-p1} \text{ , where } \frac{p1}{1-p1} \text{ is the odds of Team A winning Game \#1}$$

Now you can see that when you add one to the relative strength it results the odds of winning being multiplied by e^{β_1} . If we would calculate $e^{.206}$ (.206 is β_1 for Sagarin_Df) we obtain 1.228, which is exactly what was in our example when the relative strength Df increases by 1.

Conclusions

After comparing and analyzing all of the data of the NCAA tournament games for the last four years we can make some conclusions. All five rating systems do predict the winner better than 50% which was proven using the hypothesis tests. Also there is no clear difference between the five systems in terms of predicting outcomes. That was clearly shown using a two-sample hypothesis test. The other questions entering this project were about whether or not the actual numerical ratings of the teams actually matters in predicting their chance of winning in actual games. After using the logistic model we can determine that relative strength is a statistically significant predictor of the probability of winning for all five systems. Also we can conclude that both relative strength difference and relative strength ratio were both effective methods of determining relative strength. However, looking at the data it appears that the relative strength by difference is a better predictor than relative strength by ratio. By using the R^2 we can conclude that the three best predictors that provided the most accurate predictions were: Sagarin

Df, Seeds Df, and Pomeroy Df.   I was glad to see that the actual Seeds rating system does just as well, if not better, than other ratings system because the seeds is what is used in tournament. The committee has access to all of the other rating systems, so it makes sense that the seeds rating system is one of the most accurate.

Works Cited

Massey, Kenneth. (2006). "About the Ranking Comparision." 5, February 2008.

http://www.masseyratings.com/cf/aboutcomp.htm

McCabe, George. Moore, David. (2006). "Introduction to the practice of Statistics. Fifth Edition." W.H. Freeman and Company, New York. 5, February 2008.

National Collegiate Athletic Assosiation (NCAA). (2007). "Principles and Procedures for Establishing the Men's Bracket." 5, February 2008.

http://www.ncaa.com/basketball-mensbracket/article.aspx?id=92514

Pomeroy, Ken(1). (2006). "Rating Explanation." 5, February 2008.

http://www.kenpom.com/blog/index.php/weblog/ratings_explanation/

Pomeroy, Ken(2). (2008). "College Basketball Ratings Percentage Index." 5, February 2008. http://www.kenpom.com/rpi.php

Schatz, Amy. (2005). "Indiana Sports Guy Tackles Time Zones." The Wall Street Journal. 5, February 2008. http://www.post-gazette.com/pg/05292/591293.stm