

# *Predicting Outcomes of NCAA Basketball Tournament Games*

Aaron Cripe  
March 12, 2008  
Math Senior Project

# Project Goals

- Collect data from NCAA Men's Division I Basketball Tournaments 2004-2007
- Compute the relative strength for all the rating systems of the favored team for each game in the tournament.
- Compute relative strength by both ratio and difference to test the accuracy of the predictions of each rating system.
- Are all five rating systems statistically significant predictors of outcomes?
- Does one rating system predict the outcome better than the others?
- Is the relative strength a good predictor of the probability that the favored team wins that particular game?

# Rating Systems

- RPI – Ratings Percentage Index
- Pomeroy
- Sagarin
- Massey
- Tournament Seeds

# Comparing the Systems

	<b>Margin of Victory</b>	<b>Home Court</b>	<b>Strength of Schedule</b>	<b>Strength of Opponents Schedule</b>	<b>Record (win and losses)</b>	<b>Range of Ratings</b>
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

# Tournament Seeds

- The team's actual seed is determined by a committee that has access to the other rating systems.
  
- Does that mean that the tournament seeds should be a better predictor than the other systems?

# Any Problems?

- Massey 2005
- Pomeroy 2007 (Instead of 252, we had 189)
- Ties?
  - One for RPI
  - Two for seeds

# 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%

# Statistically Significant?

- Hypothesis Test for One Sample (RPI)

$H_0 : p = .50$       System is ineffective <default assumption>

$H_A : p > .50$       System is efficient in predicting

where  $p$  is the population proportion of correct prediction

After calculating we obtain a z-score = 6.62. Any z-score  $> 3$  will result in a p-value that is approaching zero. Thus our p-value  $< 0.05$ , so we can conclude our alternate hypothesis.



# Which one is better?

- Hypothesis Test for two samples (Seeds and RPI)

$$H_0 : p_1 = p_2 \quad \text{No difference between seeds and RPI}$$

$$H_A : p_1 \neq p_2 \quad \text{There is a difference between the two}$$

After calculating we obtain a z-score = .58. Using a statistical table we obtain a p-value = .562 . Thus our p-value is not less than 0.05, so we can conclude that our alternate hypothesis is false. There is not significant difference between the two, seeds and RPI.

# Predicting Game Outcomes Using Relative Strength

- Logistic Regression Model

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

Log of odds is a linear function of relative strength

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

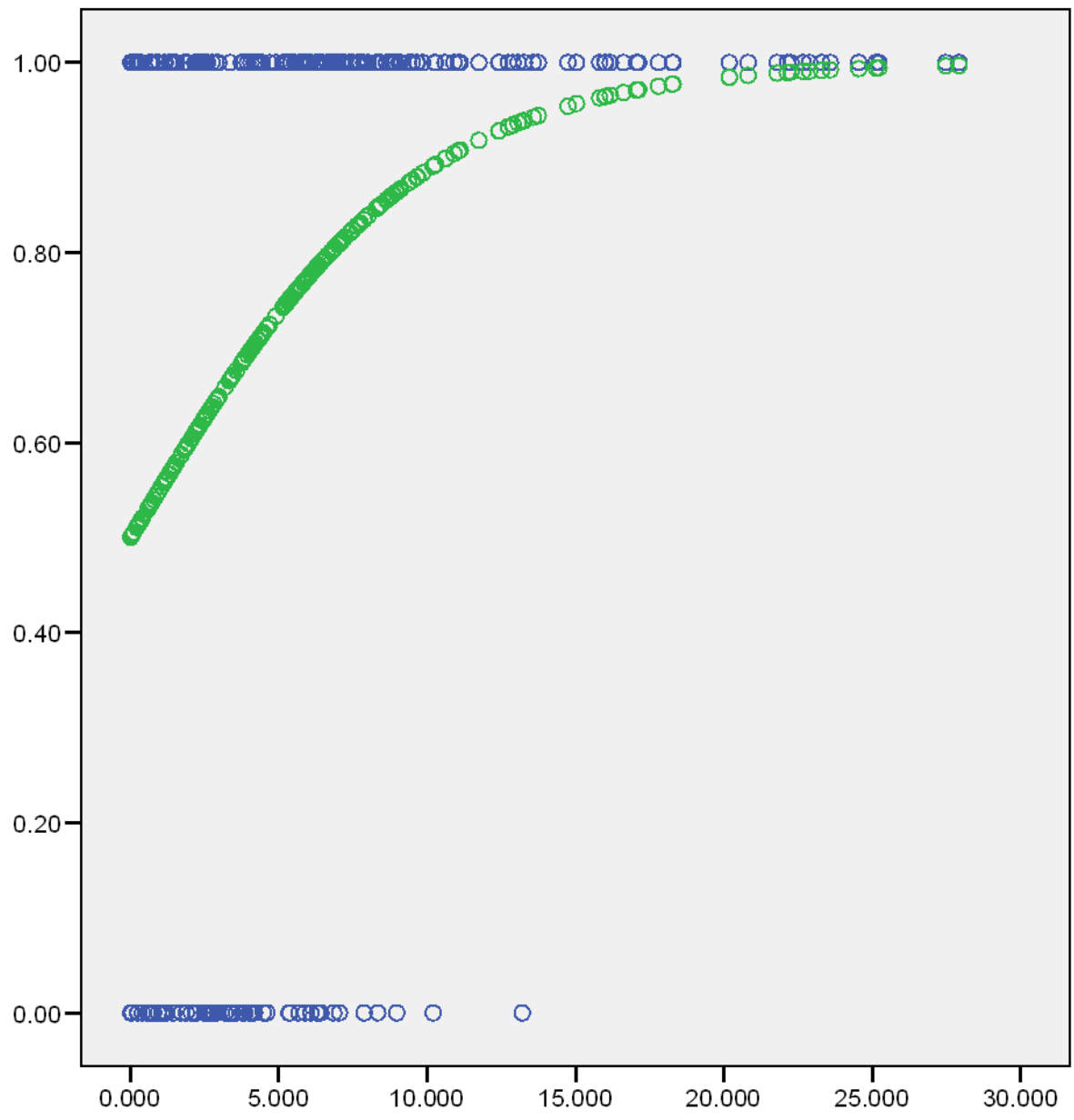
Predicting the probability of winning

$$\beta_0 = 0$$

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

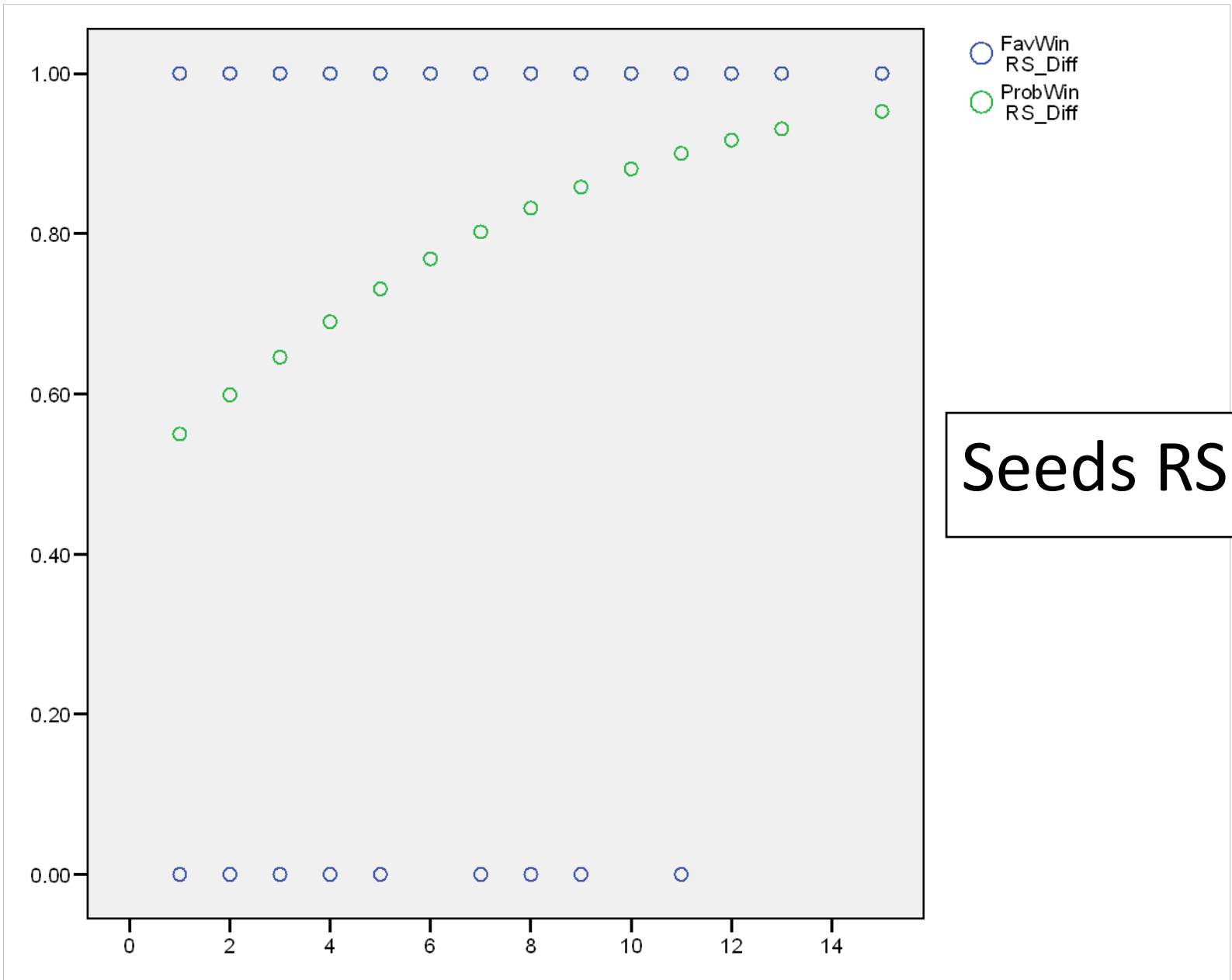
# Logistic Curve





- FavWin RS\_Diff
- ProbWin RS\_Diff

Sagarin RS Df



Seeds RS Df

# Example

- Sagarin RS-Df

2004 Elite 8

Connecticut vs. Alabama

Connecticut's Rating = 90.74

Alabama's Rating = 84.46

Relative Strength of the Favored Team = 6.28

\*Look at our Graph\*

Connecticut had approximately an 80% chance of winning according to Sagarin's RS\_Df.

# 2004 National Champions





	$\beta_1$	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

# Top 3

- Sagarin\_Df-  $p(x) = \frac{e^{.206x}}{1 + e^{.206x}}$

- Seeds\_Df-  $p(x) = \frac{e^{.200x}}{1 + e^{.200x}}$

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

# What happens when the relative strength increases 1 unit?

- $\ln \frac{p}{1-p} = \beta_1 x$  where relative strength =  $x$
- Then by raising both sides to the power of  $e$ , and we obtain:  
 $\frac{p}{1-p} = e^{\beta_1 x}$  where relative strength =  $x$
- Thus when we adding one to the relative strength we obtain:

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

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

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

$$\frac{p2}{1-p2} = e^{\beta_1} * \frac{p1}{1-p1}$$

# Conclusions

- All 5 rating systems were statistically significant.
- Not one rating systems more accurate than the others.
- Was the prediction about seeds correct?
- Ratio vs. Difference
- Top 3
  - Sagarin RS\_Df
  - Pomeroy RS\_Df
  - Seeds RS\_Df