

# NCAA Tournament Pool: Leveling The Brackets

March 16 2010, By Chris Gorski

---



March Madness -- featuring a 65-team field -- takes over the U.S. sporting world each spring.

That whirlwind of cubicle activity greeting office drones reporting to work this week is not a frenzy to finish last week's sales reports.

Those buzzing copiers and intensely-focused workers parked at their computers are in the midst of a country-wide cram session, completing brackets for the 2010 NCAA Men's [Basketball](#) Championship tournament pools. It's that rite of spring during which the field of 65 takes over the sports landscape.

Most pools use simple scoring systems that award one point for picking the winner of a first round matchup, 2 points for correctly choosing

second round winners, then 4, 8, 16, and 32 for the subsequent rounds. However, those familiar systems consider each game in a given round equal, emphasizing the final few games. They may not sufficiently reward those whose picks display the most extensive basketball knowledge.

"That system makes a certain amount of sense if you assume that each of the two teams in any particular game has a 50 percent chance of winning," said Ted Gooley, a biostatistician at the Fred Hutchinson Cancer Research Center in Seattle.

Teams are assigned to four regions that are ranked and seeded from 1 - 16. The better teams are assigned what are called the higher seeds (numerically smaller, like 1 or 2), and the lower-ranked teams with lower seeds (the highest numbers). For the first round, the 1 seed plays the 16 seed, the 2 seed plays the 15 seed, and so on.

Gooley's love of the tournament -- and frustration with simple scoring systems -- led him to borrow a technique from his professional research called logistic regression. He used it to develop his alternative scoring system.

"Why not base a scoring system on essentially the likelihood that a particular seed wins a particular game?" said Gooley. "There are clearly many games in the tournament that are far from a 50-50 proposition."

Gooley's system analyzes each NCAA Men's Basketball tournament from 1985 to 2009 and determines the probability of a particular seed winning a particular game. He found that the higher-seeded team has won 75 percent of all games.

"The way I would assign the points would be 1 divided by 25 percent for the lower seeded team, or 4 points, and for the higher seeded team, you

take 1 divided by 0.75, which is 1.33 points," said Gooley. "I wanted a little motivation for people to pick upsets, which always happen."

The mathematics is a little more complicated than this, because Gooley's model looks at each possible match up, how often it has occurred, and creates additional terms to provide the best fit of the model to the historical data. For example, a 13 seed team could play a 3 seed in a regional final, but several unlikely events would have to happen. The 10 seed differential is also a rarity, but there have been plenty of games between teams 9 and 11 seeds apart. By smoothing the data between those two points, Gooley formed an estimate of appropriate point values.

Gooley also found that his model improved if he considered not just the difference between seeds, but also what the highest seed was, because a 1 seed team has been more likely to beat a 5 seed team than a 7 seed team has been to beat an 11 seed.

Taking those components, Gooley developed a system that awards points for each seed winning each game. An 11 seed team that reached the Final Four -- like George Mason University did in 2006 -- would earn 170 points for the regional final win. But that has only happened twice in 25 years.

That's a lot of points, and more than the entrants that have won the last three pools Gooley organized with this system. The rationalization is that someone with the foresight to correctly predict that rare event should win. He designed the system that given an infinite number of tournaments, the average score for the brackets should be around 63 points, or 1 point for each game.

There's math in those simple systems, too. Tim Chartier, a mathematician at Davidson College in N.C., has developed his own mathematical methods for predicting tournament outcomes, one of

which beat out 97 percent of the 4.6 million entrants in ESPN's Tournament Challenge last year. He developed a ranking system that uses regular season game scores to assess the relative quality of teams. If the team he ranks 43 were scheduled to face the team ranked 45, no matter what the seeds, he picks the team at 43.

As a fan of basketball and math, he appreciates Gooley's efforts. "For true sports analysts and people who really look carefully I do see it as a really intriguing system," said Chartier.

He was especially interested to know if there were people who were able to consistently outperform the mean of 63 points, which would indicate some sort of advanced basketball knowledge, or exceptionally advanced luck.

"I love what one of my friends said. He said, 'look, you're trying to model and understand the behavior of 18 - 22 year old guys,'" said Chartier. "There's an unpredictable component."

Gooley feels that his system emphasizes basketball insight, that if a fan can identify that a certain team has a collection of players that will present difficulty to another, more accomplished, higher seeded team, they should be rewarded with more points than for automatically advancing a 1 seed over a 16. "I think this is the most logical way to assign points," he said.

Basketball fans need not worry that a massive changeover to this advanced scoring system would leave them unable to compete with the mathematically and statistically-inclined, however.

"Based on my performance in the last three pools, I think it's pretty clear my next area of work needs to be to focus on how to win pools other than how to score them," said Gooley.

(c) 2010 Inside Science News Service

Citation: NCAA Tournament Pool: Leveling The Brackets (2010, March 16) retrieved 24 April 2024 from <https://phys.org/news/2010-03-ncaa-tournament-pool-brackets.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.