Sampling from the 9,223,372,036,854,775,808 Possible Brackets in the NCAA Men's Basketball Tournament using the Power Model

Sheldon H. Jacobson, Arash Khatibi, Douglas M. King

Department of Computer Science, University of Illinois at Urbana-Champaign

Introduction

- The NCAA D1 Men's Basketball Tournament is an annual single-elimination competition that attracts widespread attention in the United States.
- 64 teams (after the First Four) compete in four regions, each team assigned a seed number (1 to 16).
  - Smaller seed numbers represent stronger teams: Seed 1 is the strongest in each region.
- Generating a bracket is the process of picking the winners of all 63 games.
- Number of possible brackets: 9,223,372,036,854,775,808

- Objectives
  - Design a model to capture the probability mass function for all possible brackets.
  - Use the model to sample from this pool of brackets.

Power Model

- Models the relative strength of the two teams as a power function of their seed numbers.
- Estimates the associated parameters using historical tournament results since 1985 (the modern era).
- Let p denote the proportion of times seed s1 defeated seed s2 since 1985

$$\frac{p}{1-p} = \left(\frac{s_2}{s_1}\right)^{\alpha_j(s_1,s_2)}$$

- The Alpha value of seeds s1 and s2 in Round j is computed as

$$\alpha_j(s_1,s_2) = \frac{log(p/(1-p))}{log(s_2/s_1)}$$

- Probability that s1 defeats s2 in Round j

$$\frac{s_2^{\alpha_j(s_1,s_2)}}{s_1^{\alpha_j(s_1,s_2)} + s_2^{\alpha_j(s_1,s_2)}}$$

- Positive Alpha value: Larger winning probability for stronger seed.
- Negative Alpha value: Smaller winning probability for stronger seed.
- Alpha value of zero: Random pick.
- Alpha value of infinity: Always pick the stronger seed.
- Alpha value of one: Seeds provide a linear proportion of probability of winning (neutral).

Alpha Values

- Round of 64

Match up	Alpha (2012)	Alpha (2013)	Alpha (2014)	Alpha (2015)
(1,16)	3.00	3.00	3.00	3.00
(2,15)	1.62	1.43	1.36	1.38
(3,14)	1.36	1.16	1.14	1.13
(4,13)	1.11	1.10	1.10	1.13
(5,12)	0.79	0.76	0.69	0.62
(6,11)	1.14	1.10	1.12	1.08
(7,10)	1.16	1.12	1.18	1.23
(8,9)	-0.94	-0.61	-0.59	-0.28

- Round of 32

Match up	Alpha (2012)	Alpha (2013)	Alpha (2014)	Alpha (2015)
(1,8)	0.68	0.71	0.73	0.71
(1,9)	1.18	1.18	1.09	1.10
(2,7)	0.81	0.83	0.86	0.85
(2,10)	0.19	0.21	0.24	0.20
(3,6)	0.34	0.33	0.42	0.40
(3,11)	0.53	0.50	0.53	0.47
(4,5)	0.31	0.60	0.74	0.87
(4,12)	0.34	0.40	0.37	0.51
(5,13)	1.36	1.36	1.36	1.36
(6,14)	2.01	2.01	2.11	2.11
(7,15)	-	-	0.91	0.91
(10,15)	-	-	-	-
(11,14)	-	-	-	-
(12,13)	25.98*	17.32*	12.25*	12.25*
Other match ups	0.65	0.70	0.73	0.74

Alpha values of Round 3

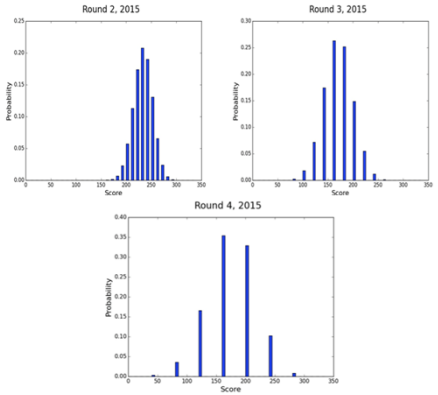
- Remaining Rounds
  - One Alpha value is used for all match ups, computed as the weighted average of different pairs of Alpha values

Round	Alpha (2012)	Alpha (2013)	Alpha (2014)	Alpha (2015)
4	0.82	0.86	0.83	0.80
5	0.27	0.22	0.03	0.05
6	0.43	0.48	0.48	0.48
7	1.35	1.44	1.45	1.45

Evaluation

- ESPN scoring system
  - 10 points for each correct pick in Round of 64.
  - Each correct pick worth twice its prior round: correct pick for National Champion is 320 points.
- Results based on one million generated brackets

year	Max score	Mean score	Max correct picks	Mean correct picks
2012	1760	784.10	51	35.72
2013	1680	672.60	52	34.23
2014	1420	578.88	50	34.58
2015	1730	833.93	54	37.91



Distribution of scores in the first three rounds of 2015

	Number of correct picks				
Tournament	0	1	2	3	4
2012	23.18	45.81	25.94	4.93	0.14
2013	44.24	52.50	3.20	0.05	0.01
2014	36.48	48.19	15.09	0.23	0.01
2015	15.44	39.95	34.39	10.07	0.15

Proportion of correct picks in Final Four

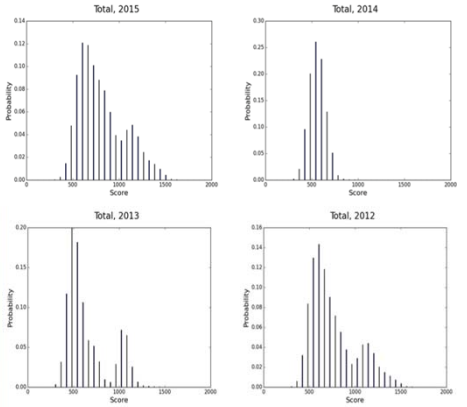
	Number of correct picks		
Tournament	0	1	2
2012	59.95	36.11	3.94
2013	67.81	31.91	0.28
2014	99.64	0.35	0.01
2015	51.43	40.27	8.30

	Number of correct picks	
Tournament	0	1
2012	79.63	20.37
2013	79.49	20.51
2014	99.99	0.01
2015	80.74	19.26

Proportion of correct picks in Rounds 6 and 7

Discussion

- A two-peak distribution of scores
  - Value of correct picks in later rounds
  - One-peak for 2014, where a seed 7 won the championship for the first time



Conclusions

- The Power Model is an intuitive model to sample from the large pool of possible brackets in the NCAA D1 Men's Basketball Tournament.
- The Alpha values summarize the performance history of each seed match up in a round.
- Generating a good bracket is more difficult for tournaments with many upsets or two few upsets in the early rounds.
- The results show a bell-shaped figure for each round and a two-peak distribution for the whole tournament

Acknowledgments

We would like to thank Bleacher Report for featuring our work in 2014. The Power Model is available at bracketodds.cs.illinois.edu

ILLINOIS