

Fall 10-28-2016

Statistics of FA Cup Soccer

Jonathan Heagy

Valparaiso University, jonathan.heagy@valpo.edu

Follow this and additional works at: <http://scholar.valpo.edu/fires>

Recommended Citation

Heagy, Jonathan, "Statistics of FA Cup Soccer" (2016). *Fall Interdisciplinary Research Symposium*. Paper 20.
<http://scholar.valpo.edu/fires/20>

This Poster Presentation is brought to you for free and open access by the Office of Sponsored and Undergraduate Research at ValpoScholar. It has been accepted for inclusion in Fall Interdisciplinary Research Symposium by an authorized administrator of ValpoScholar. For more information, please contact a ValpoScholar staff member at scholar@valpo.edu.

Statistics of FA Cup Soccer

Jonathan Heagy

Advisor: Professor Patrick Sullivan



Abstract

The FA Cup Soccer tournament in England is the oldest professional soccer tournament in the world, making it a very fascinating event for soccer fans all over the world. What makes this tournament especially interesting, however, is the randomization of the seeing in each round. Typical tournaments are seeded so teams are ranked by ability, however the FA Cup randomizes each round, so you never know who will play who next. We took particular interest in the FA Cup starting with the round where 64 teams remain; the Third Round Proper. Here, the teams from the top two English soccer leagues, the Championship and Premier leagues, enter into the tournament. Our goal was to investigate some of the interesting statistics that can be drawn from this style of tournament. Some of the main questions we worked with involved finding the probabilities of certain amounts of teams from each league advancing to each round starting with the Third Round Proper. Using the FA Cup website, we were able to collect data from 1996-2016 that allowed us to calculate the winning percentage of teams from each of our three categories; Premier League teams, Championship League teams, and lower league teams, against the other two categories of team. From here, we developed equations which allowed us to calculate the expected value for the number of teams advancing to each round from each league.

Data Collection

To collect our data, we used the FA Cup tournament website; thefa.com/thefacup. The results from previous years of the tournament were posted for each round. From year to year, teams are bumped up to higher leagues or relegated to lower leagues based on how well they performed that season. The FA Cup did not tell us what league each team was playing in a given year, so we used previously collected data from Professor Sullivan to look up where each team played each season. We then recorded all of our data for how each league performed against the other two and recorded it in an Excel spreadsheet.

Calculated Probabilities

$$\sum_{k=0}^{\min(m, 20-m)} \frac{2^{20-2k} \binom{32}{k, 20-2k, 12+k}}{\binom{64}{20}} * \binom{20-2k}{m-k} * .81^{m-k} * (1-.81)^{20-m-k}$$

Above is the equation we developed in order to calculate the probability distribution for the number of Premier League teams advancing to the round of 32. This equation uses the probability of their being k games where two Premier League teams are matched against each other multiplied by the number of Premier League winners and losers from the games that matched a Premier League team and a non-Premier league team. The .81 comes from the probability of a Premier League team defeating a team from a different league, which we discovered in our data collection. Shown below are the probability distributions for a certain number of Premier League teams advancing to each round.

Round of 32

0	1	2	3	4	5	6	7	8	9	
4.54·10 ⁻¹⁷	1.31·10 ⁻¹⁴	1.56·10 ⁻¹²	1.03·10 ⁻¹⁰	4.21·10 ⁻⁹	1.14·10 ⁻⁷	.0000021325	.0000280989	.0002654933	.0018175879	
10	11	12	13	14	15	16	17	18	19	20
.009068339	.033033778	.087695882	.168685661	.232672239	.226400934	.151591476	.067210562	.018525861	.002823468	.000178368

Round of 16

0	1	2	3	4	5	6	7	8	9
5.59·10 ⁻¹⁰	4.87·10 ⁻⁸	.0000017558	.0000346161	.000416997	.0032453405	.0168251534	.0590564565	.1412350693	.2297676149
10	11	12	13	14	15	16			
.251953771	.182968249	.085512772	.024592236	.004045988	.000333800	.000010130			

Round of 8

0	1	2	3	4	5	6	7	8
.0000011522	.0000613855	.0012494275	.0126494959	.0694477570	.2110083143	.3452698321	.2771240329	.0831886025

Round of 4

0	1	2	3	4
.000121742	.005120891	.069106430	.352909715	.572741222

Championship Round

0	1	2
.002757873	.108416883	.888825244

Probabilities Continued

From here, we were able to calculate the expected probability for how often a Premier League team wins the FA Cup Championship. By adding together the probability that two Premier League teams make the championship and the probability of one Premier League team making the championship multiplied by the chance that the Premier League team wins that game, we get: .888825244 + (.81) .108416883 = .9766429194, or about a 97.7% chance that a Premier League team wins the FA Cup Championship.

Calculated Expected Values

Using our probability distributions, we were able to calculate expected values for the exact number of Premier League teams we would expect to advance to each round. Below are these expected values compared to the average number of Premier League teams that advanced to the same round which we took from our data.

	Round of 32	Round of 16	Round of 8	Round of 4	Final Round
Expected	14.33	9.67	6.05	3.49	1.89
From Data	14.65	9.35	6.1	3.25	1.9

Beginning Future Work

Once we completed looking at Premier League teams vs. all other competition, we began to look at the Cup by grouping the teams into 3 categories, Premier, Championship, and Other, as opposed to just Premier and All other. This proved more difficult due to the added variable, however we were able to get expected values for the number of teams from each league advancing to the round of 32:

Premier: 20(.5(19/63)+.79(24/63)+.84(20/63))=14.37

Championship: 24(.5(23/63)+.21(20/63)+.68(20/63))=11.16

Other: 20(.5(19/63)+.32(24/63)+.16(20/63))=6.47

References

thefa.com/thefacup

Special Thanks To:

Professor Sullivan
MSEED
Valparaiso University