Seeding Statistics of Elimination Tournaments

Timothy Prescott*

10 July 2025

Contents

Abstract

We develop and provide Python code and a website to statistically analyze seedings in elimination tournaments. We are able to apply this code to fifty-eight thousand games to estimate the probability of an upset solely as a logistic function of the difference in seeding. We are also able to examine how well or poorly a team performs compared to its seeding. We conclude that the only team that is consistently underrated is \your_favorite_team, while the only team that is consistently overrated is \your_hated_rival.

Keywords— logistic regression, statistical analysis, elimination tournament, Python, data visualization

Introduction

A notable occurrence during the 2024 NCAA Division I Men's Basketball Tournament (commonly known as March Madness) was that #11 North Carolina

*University of North Dakota, `timothy.prescott@und.edu`

[0]The Version of Record of this manuscript has been published and is available in Scatterplot 10 July 2025 `https://www.tandfonline.com/10.1080/29932955.2025.2523666`

State reached the Final Four, beating teams seeded 6, 14, 2, and 4. Another notable occurrence was that the only upset in the first round of the women's tournament was by #11 Middle Tennessee. We will estimate that the probability of the first was 1.3%, while the probability of the second was 2.4%.

Bracket construction has been analyzed by Schwenk [1] and Seltzer and Simonson [2], while Wittry [3] examined upsets in the tournament. Making accurate upset predictions is popular and profitable (see Chartier [4] and the papers at Jacobson [5]), so that a driver of statistical research over the past few years has been to analyze teams' performance throughout the season in order to predict more accurately than others how each team will perform in the tournament.

We will develop and provide Python code to analyze how teams have performed over the years in these tournaments. Because a team's performance will vary from year to year, we will compare how a team performed with how they were seeded in a particular tournament (which we use to approximate how they were anticipated to perform). Because we provide the code for this analysis, the interested reader will be able to repeat or extend our analysis at their leisure.

NCAA Division I Basketball Tournament

The modern partition of NCAA men's basketball into three divisions began in 1974, with seeding beginning in 1978. For the years since, each tournaments' outcomes are in the Wikipedia page "[Year] NCAA Division I [men's|women's] basketball tournament" (the gender was added in 1982 when the NCAA began sponsoring the women's tournament). We are able to use Pywikibot [6] to automatically download these pages for further analysis (we will also cache these pages to reduce the network load). While Wikipedia renders a tournament as in Figure 1, the source for this bracket is

```
...
| RD1-seed01=1
| RD1-team01= '''[[2022-23 Alabama Crimson Tide men's basketball team|Alabama]]'''
| RD1-score01='''96'''
| RD1-seed02=16
| RD1-team02= [[2022-23 Texas A&M-Corpus Christi Islanders men's basketball team|
    Texas A&M-Corpus Christi]]
| RD1-score02=75

| RD1-seed03=8
| RD1-team03= '''[[2022-23 Maryland Terrapins men's basketball team|Maryland]]'''
| RD1-score03='''67'''
| RD1-seed04=9
| RD1-team04= [[2022-23 West Virginia Mountaineers men's basketball team|West
    Virginia]]
| RD1-score04=65

| RD1-seed05=5
| RD1-team05= '''[[2022-23 San Diego State Aztecs men's basketball team|San Diego
    State]]'''
| RD1-score05='''63'''
| RD1-seed06=12
| RD1-team06= [[2022-23 College of Charleston Cougars men's basketball team|
    Charleston]]
| RD1-score06=57
```

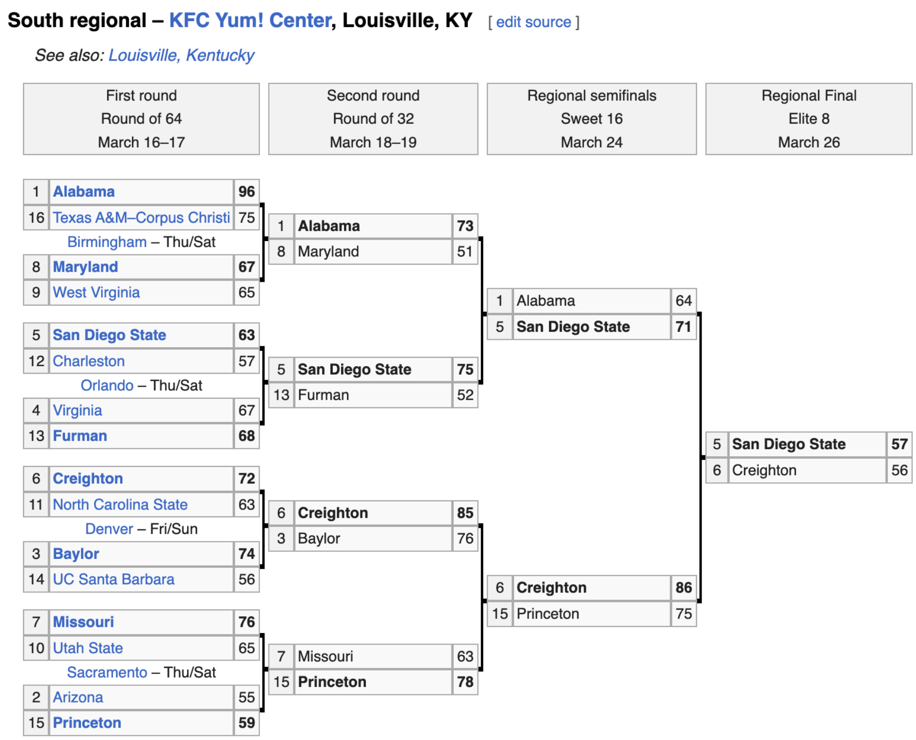
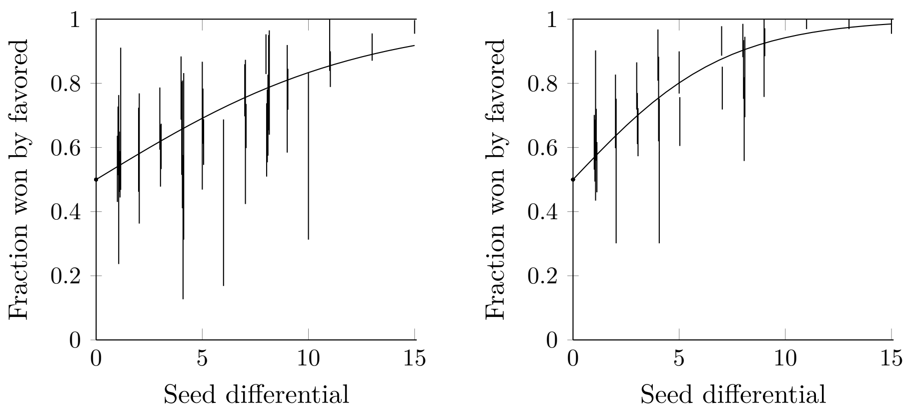


Figure 1: Tournament bracket as seen in Wikipedia [7]



(a) Men's. $\beta = 0.161$

(b) Women's. $\beta = 0.279$

Figure 2: Seed differential versus fraction of games that a favored team wins in the NCAA Division I basketball tournament, with 95% Wilson confidence intervals and the logistic best fit. Identical seed differentials have been slightly spread to avoid overlap.

```
| RD1-seed07=4
| RD1-team07=[[2022-23 Virginia Cavaliers men's basketball team|Virginia]]
| RD1-score07=67
| RD1-seed08= 13
| RD1-team08= '''[[2022-23 Furman Paladins men's basketball team|Furman]]'''
| RD1-score08='''68'''
...
```

Each line begins with "`| RD#-[seed|team|score]##`", where the first number indicates the round, and the second numbers the teams within that round; teams numbered $2n-1$ and $2n$ will play each other. Combining these results, we determine how often teams with a given seeding are victorious in 2833 (men's tournament) and 2372 (women's tournament) games through the year 2024. In Figure 2, we plot the fraction of games that the favored team wins as a function of the seed differential, along with their 95% confidence intervals (using the Wilson interval [8]).

We also plot the logistic best fits (determined using Python's Scikit [9]). Letting $Y = 1$ for victory, $Y = 0$ for defeat, and S denote the seed differential of a game, the average log-likelihood $\bar{\ell}$ is the expected value of $Y \ln p(S) + (1 - Y) \ln(1 - p(S))$. The logistic function is the function of the form $p(x) = (1 + \exp(-\beta x - \beta\mu))^{-1}$ that maximizes $\bar{\ell}$. We find that $\beta = 0.161$ for the men's tournament, and $\beta = 0.279$ for the women's. Because a win for a given seed differential corresponds to a loss for that negated seed differential, the logistic curves are symmetric about $(0, 0.5)$, so that $\mu = 0$, and we only plot the curve for nonnegative seed differentials. The confidence intervals do not inspire much confidence in this best fit, but we will be able to add more data as we proceed with our analysis.

Table 1: Fraction of time that the favored team wins in the second round, after one first round upset. (Minimum 10 games.)

Seeds	Men's (%)	Women's (%)
1 v 9	89 ± 6	93 ± 5
2 v 10	62 ± 11	84 ± 9
3 v 11	67 ± 11	69 ± 13
4 v 12	70 ± 13	82 ± 13
5 v 13	80 ± 15	
6 v 14	80 ± 16	
overall	75 ± 5	85 ± 5
Figure 2 prediction	78.4	90.3

A β that is 70% larger in the women's tournament than the men's tournament indicates that dramatic upsets are much less likely to occur. In the men's tournament, upsets with a seed differential of 7 or more occurred 238 times out of 1286 games, or 19% of the time. In the women's tournament, upsets with a seed differential of 7 or more occurred 76 times out of 953 games, or 8% of the time. A Fisher test shows that such upsets are more likely in the men's tournament ($p = 2.5 \times 10^{-13}$).

With $\beta = 0.279$ and $p(x) = (1 + \exp(-\beta x))^{-1}$, we are able to perform a calculation from the introduction. The probability P that there are no upsets in the first round of the women's tournament is $\prod_{i=1}^{8} p(2i-1)^4 \approx 0.29\%$. The probability that the only upset is a single seed k (losing to a seed $17-k$) is $P \cdot p(17-2k)^{-4} \cdot 4p(17-2k)^3(1-p(17-2k)) = 4P \cdot (1-p(17-2k))/p(17-2k)$. Therefore, the probability that there is at most one upset in the first round is

$$P + \sum_{k=1}^{8} 4P \frac{1-p(17-2k)}{p(17-2k)} \approx 2.4\%.$$

We are interested in the question of how a team performs in a game after an upset win. Are they tired from having over-exerted themselves, more experienced from having won a tournament game, or better than previously appreciated because we can now condition on them beating a supposedly superior opponent? As shown in [2], seed pairings must appear in known rounds. The seeds must add to 17 if and only if the teams meet in round one. Teams meet in the second round if and only if the seeds add to 9 or 25 or differ by 8 (corresponding to 0, 2, or 1 upsets in the first round). In Table 1, we examine the intervals from Figure 2 that correspond to the second round after one first round upset. The fraction of wins is slightly more in favor of an upset, but not significantly so.

In 2001, the men's tournament added a "play-in" game, where the last two teams to make the tournament played each other for the last spot. The play-in game expanded to eight teams for four spots in 2011. (The NCAA initially

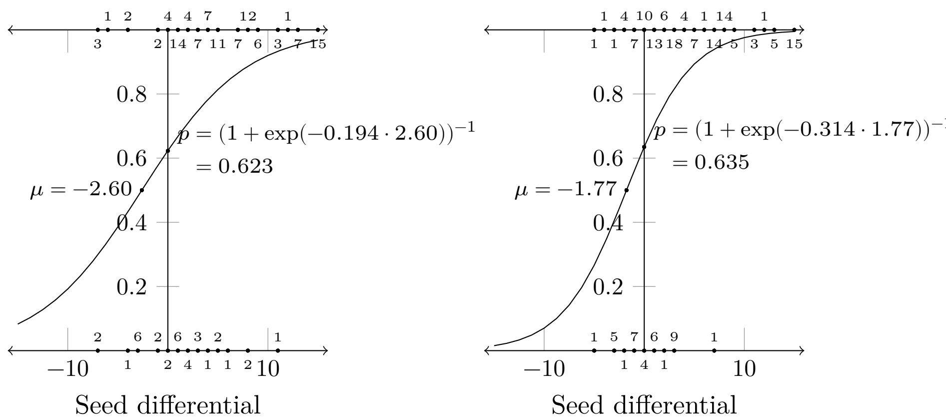
Table 2: Number of games won in first round in the men's tournament, by seed, where the worse seed was (not) the winner of a play-in game. Also significance of the statement "Play-in winners are more likely to have an upset in their first round."

Seed	7	10	6	11	5	12	4	13	3	14	1	16	Overall	
Play-in	1	1	9	9	3	1	0	1	1	0	35	1	49	13
Non-play-in	34	56	41	33	52	36	71	20	81	10	55	1	487	187
p	0.86		0.44		0.88		0.23		1		0.63		0.91	

referred to this as the first round, but later branded this the "First Four" so that "first round" would still refer to the round with 64 teams. We use the "later" terminology.) We are interested in the question of how the play-in winners perform in their first round game: are they tired from having played earlier in the week, more experienced from having played a tournament game, or better than previously appreciated because we can now condition on them beating a supposedly equivalent opponent? We are able to identify play-in games by looking for identically high seeded teams. We then record how the winning team performs in their first round game, and summarize according to their seed in Table 2. Each seed has 92 total games from 23 years and 4 games per seed per year. Of those 92, the first row of the table counts games played with play-in winners of that seed, while the second row counts games with non-play-in winners. The left column of each pair counts games where the favored team won, while the right column counts upsets. The final row gives the significance that play-in winners are more likely to have an upset. Given the high p values, we are unable to substantiate this claim. There is even less data for the women's tournament, which has only had 12 play-in games since they were added in 2022.

Each year, the bracket's unveiling is met with inevitable second guessing that a team should be seeded differently. Fans (and foes) of a particular team will often claim that team over (or under) performs in the tournament, or even that the team consistently receives (un)favorable bias from the selection committee.

We are able to examine these claims by recording how a team performs each year during the tournament. But while collecting this data, we quickly run into a problem: the millions of Wikipedia editors (unsurprisingly) have not entered the team names consistently. For example, eventually we see twelve of the twenty four possible configurations of S[o[uth[ern]]] Conn[ecticut] [St[ate]], only some of which use a period for the abbreviations, that will all become "Southern Connecticut". Even within a single bracket (presumably one editor), a team tends to have a more abbreviated name in later rounds. We will therefore devote a large portion of our code to normalizing the team names that appear by expanding most abbreviations and dropping unnecessary words, occasionally resulting in a team name that is slightly different than what the media and fans customarily use.



(a) North Carolina men. $\beta = 0.194$

(b) Tennessee women. $\beta = 0.314$

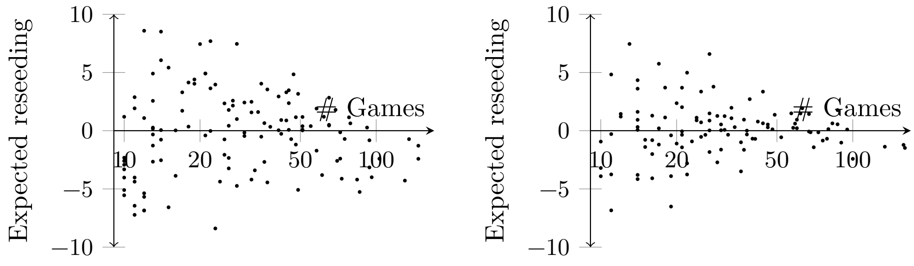
Figure 3: Wins and losses (with multiplicity) in the NCAA Division I basketball tournament, along with the logistic regression $(1 + \exp(-\beta(x - \mu)))^{-1}$.

Once team names have been normalized, we still have a problem of ambiguity. We will convert "USC" to "Southern California" or "South Carolina", often by looking for which mascot appears elsewhere on the page (a conference is also helpful). Similarly, "Saint John's" will become "Saint John's New York" or "Saint John's Minnesota" along with scores of other disambiguations. (Even after normalization and disambiguation, we still needed to correct many misspellings and other typos within Wikipedia. As a practical matter, we found it most useful to determine which state a team was from, and then examine why our code had trouble determining the state of a particular team.)

Within our sample, UNC (now North Carolina) has played the most men's tournament games at 139, while (the University of) Tennessee has played the most women's tournament games at 165. Plotting the seed differences of those games along with whether the game was a win (106 and 130 times) or a loss (33 and 35 times) along with the logistic regressions gives Figure 3.

The logistic function $(1 + \exp(-\beta(x - \mu)))^{-1}$ for $\beta = 0.194$ and $\mu = -2.60$ suggests that UNC performs about 2.60 seedings better than expected and that the log of their odds against an evenly seeded opponent is $-\beta\mu = 0.504$, while $\beta = 0.314$ and $\mu = -1.92$ suggests that Tennessee performs about 1.92 seedings better and has log odds of 0.603. The larger β also indicates that Tennessee is less likely to experience an upset (for better and worse).

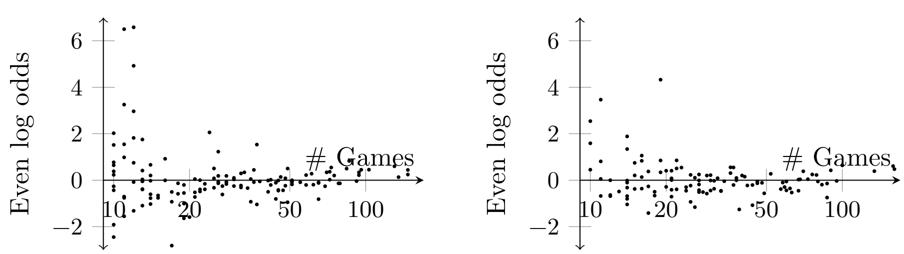
Of the 316 (303) teams that have participated in the men's (women's) tournament, 135 (110) have at least 10 games in the tournament and at least 1 win. Of these, only 7 men's (1 women's) teams have $\beta < 0.01$, causing a large μ. Repeating the calculation of μ for the remaining 128 (109) teams and plotting the number of games along with the team's μ gives us Figure 4, while plotting the number of games along with the team's log odds gives us Figure 5.



(a) Men's

(b) Women's

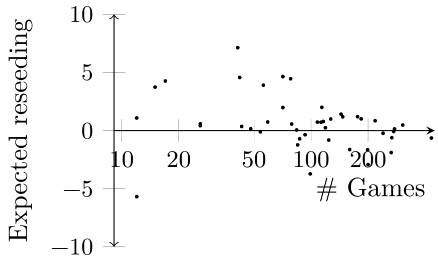
Figure 4: (Logarithmic) number of games versus expected reseeding in each NCAA Division I basketball tournament.



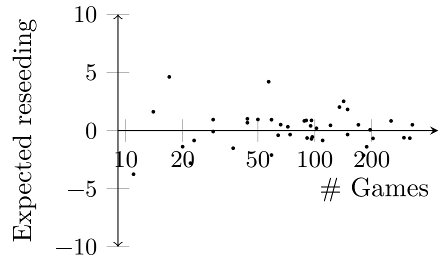
(a) Men's

(b) Women's

Figure 5: (Logarithmic) number of games versus logarithm of odds against an evenly matched opponent in each NCAA Division I basketball tournament.

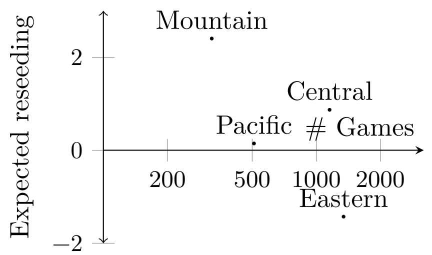


(a) Men's, omitting DE (0-7); HI (1-4); ND (2-5); SD (0-6); AK, ME, NH (all 0-0).

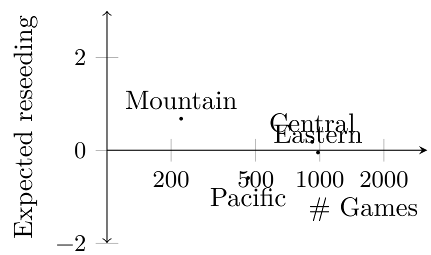


(b) Women's, omitting AK (0-0), DE (3-6), HI (1-8), ID (0-14), NH (0-7), ND (0-1), RI (2-7), VT (1-7), WY (0-2).

Figure 6: (Logarithmic) number of games versus expected reseeding in each NCAA Division I basketball tournament, grouped by state.



(a) Men's, omitting Alaska (0-0) and Hawaii (1-4).



(b) Women's, omitting Alaska (0-0) and Hawaii (1-8).

Figure 7: (Logarithmic) number of games versus expected reseeding in each NCAA Division I basketball tournament, grouped by timezone.

We observe a slight negative trend of reseeding for the teams that have a large number of games (the right most four men's teams are perennial contenders Kentucky, Kansas, Duke, and North Carolina, while the right three women's teams are Stanford, Connecticut, and Tennessee). Such teams have a large amount of victories against lower ranking teams, so that the logistic best fit is pulled up and toward a negative μ. This also corresponds to a slight positive trend for the log odds.

Having failed to identify any bias for or against a particular team, we consider the possibility that a *group* of teams may experience some bias. We repeat the calculation of μ (omitting $-\beta\mu$), now grouping teams by their state in Figure 6 and (slightly inaccurately) grouping the states by timezone in Figure 7. For the timezone plot, note that we include Arizona (Mountain Standard Time) in Mountain (Daylight) Time because the majority of the collegiate basketball season occurs outside of Daylight Saving Time. Also, we include Indiana in Eastern Time.

We now focus on our original example, that an 11 seed won games with seed differentials -5, 3, -9, and -7. For a game with seed differential s, we previously approximated the probability of the favored team winning by $p(s) = (1 + \exp(-\beta s))^{-1}$, where $\beta = 0.161$ for this tournament. Treating these as four independent games would give a probability of 0.89%. We will instead argue that the upsets in rounds 1 and 3 indicate that the team should have been seeded better, making the victories in rounds 2–4 and then round 4 more likely.

Returning to the 128 men's (109 women's) teams, if we discard 3 (2) outliers where $|\mu| > 16$, then we can weight the remaining 125 (107) values of μ by the number of games the team played to estimate that μ has a mean μ_0 of -0.2 (0.03) and a standard deviation σ of 3.2 (1.9). This indicates that we can instead assume that a better seeded team is correctly seeded, and that a worse seeded team should have had their seed adjusted downward by the normal random variable $N \sim \mathcal{N}(\mu_0, \sigma^2)$. In that case, the most likely seeding correction in the event of an upset would be the maximizer of the product of the probability density function for N and $1 - p(s - N)$, which is

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\Big\{-\frac{(N-\mu_0)^2}{2\sigma^2}\Big\} \frac{\exp(-\beta(s-N))}{1+\exp(-\beta(s-N))}.$$

Differentiating with respect to N and simplifying, this maximum occurs when N is the root of the nonlinear equation

$$\frac{\beta}{1+\exp(-\beta(s-N))} = \frac{N-\mu_0}{\sigma^2}.$$

With our values of β, μ_0, and σ, a numerical solver estimates that on $1 \le s \le 15$, this root is approximately $0.68 + \frac{s}{25}$ for the men's tournament and $0.65 + \frac{s}{42}$ for the women's tournament. This means that after an 11 seed defeats a 6 seed, we should instead model the 11 seed as if they had been seeded 10.12. Treating the last three games dependent on the first one (but independent from each other), we would act as if the seed differentials had been -5, 3.88, -8.12, and -6.12, bringing the probability up to 1.2%.

Repeating this analysis for two upsets, we would want to maximize

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\Big\{-\frac{(N-\mu_0)^2}{2\sigma^2}\Big\} \frac{\exp(-\beta(s_1-N))}{1+\exp(-\beta(s_1-N))} \frac{\exp(-\beta(s_2-N))}{1+\exp(-\beta(s_2-N))},$$

which happens when N solves

$$\frac{\beta}{1+\exp(-\beta(s_1-N))} + \frac{\beta}{1+\exp(-\beta(s_2-N))} = \frac{N-\mu_0}{\sigma^2}.$$

This root is approximately $1.43 + \frac{s_1+s_2}{24}$ for the men (and $1.17 + \frac{s_1+s_2}{38}$ for the women) on $1 \le s_1, s_2 \le 15$, so that after upsets of seed differentials -5 and -9, we should instead model the 11 seed as if they had been seeded 9. Repeating our calculation with seed differentials of -5, 3.88, -8.12, and -5, the probability is now the previously stated 1.3%.

Table 3: Average log-likelihood in the second round, after one first round upset, with and without reseeding. (Minimum 10 games.)

Seeds	Men's		Women's	
	Without	With	Without	With
1 v 9	−1.27	−1.20	−2.19	−2.04
2 v 10	−0.51	−0.45	−1.88	−1.76
3 v 11	−0.63	−0.59	−1.32	−1.22
4 v 12	−0.75	−0.70	−1.86	−1.76
5 v 13	−1.13	−1.10		
6 v 14	−1.18	−1.15		
overall	−0.85	−0.80	−1.83	−1.71

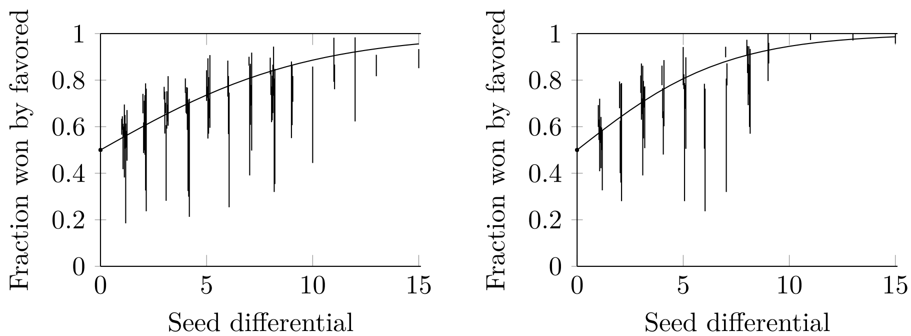
We will now revisit the probability of an upset in the second round after an upset in the first, examined in light of reseeding. For a logistic function $p(s) = (1+\exp(-\beta_0 - \beta_1 s))^{-1}$, recall that the average log-likelihood $\bar{\ell}$ is the expression $Y_k \ln p(S_k) + (1 - Y_k)\ln(1 - p(S_k))$ where S_k is the seed differential (positive for a favored team) and Y_k is 1 in victory and 0 in defeat. The expression is maximized when β_0 and β_1 are taken from the logistic best fit (note that this logistic function is slightly different than ours, where $\beta = \beta_1$ and $\mu = -\beta_0/\beta_1$, although $\mu = \beta_0 = 0$ because of symmetry in our case). In Table 3, we see the average log-likelihood for each second round matchup after a first round upset. Using the reseeding results in a better (less negative) average log-likelihood.

College Basketball Tournaments

Wikipedia has brackets for many other basketball tournaments, most notably 32 (32) different conference tournaments [with occasional changing names] often determining a champion for an automatic bid to the men's (women's) national tournament. There are also 7 (4) other tournaments with a national scope. We again begin by plotting seed differential against fraction of games won by the favored team in Figure 8, and then the number of games compared to how a team should be reseeded in Figure 9.

We also summarize the likelihood of upsets by calculating β for each tournament, and plot the results in Figure 10. Note that national conferences generally have lower β values, suggesting that upsets are more common. Within a conference tournament, most teams have already played each other, and a team's record (which all conferences use for seeding) is a useful estimate of how good the team is. In a national tournament, on the other hand, most teams have not played each other. In this case, tournament organizers resort to (sometimes objective) arbitrary methods to determine seedings, but these methods are not as useful at estimating how good the teams are.

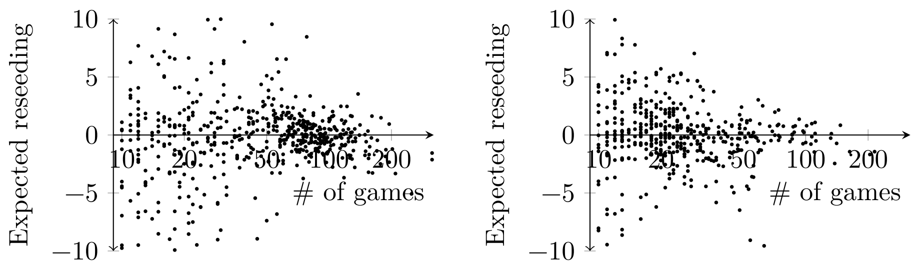
To examine one conference tournament slightly more in-depth, we plot in



(a) Men's. $\beta = 0.205$

(b) Women's. $\beta = 0.285$

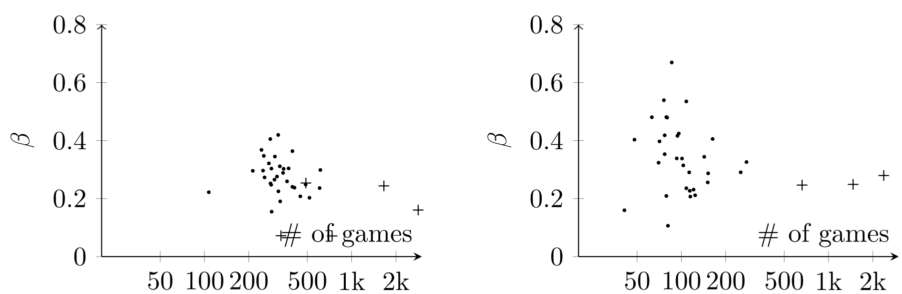
Figure 8: Seed differential versus fraction of games that a favored team wins in collegiate basketball tournaments, with 95% Wilson confidence intervals and the logistic best fit.



(a) Men's

(b) Women's

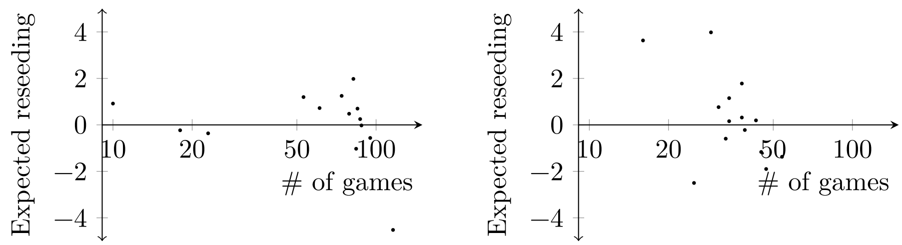
Figure 9: (Logarithmic) number of games versus expected reseeding in basketball tournaments.



(a) Men's

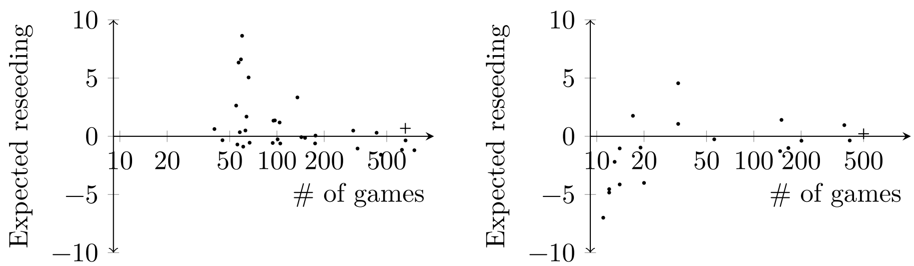
(b) Women's

Figure 10: (Logarithmic) number of games versus upset rate in basketball tournaments. National tournaments denoted with +.



(a) Men's (b) Women's

Figure 11: (Logarithmic) number of games versus expected reseeding in SEC basketball tournaments.



(a) Men's (b) Women's

Figure 12: (Logarithmic) number of games versus expected reseeding of conferences in national basketball tournaments. Unknown conferences have been consolidated into $+$.

Figure 11 the number of games versus the expected reseeding for the SEC, located at $(517, 0.20)$ in Figure 10a and $(253, 0.29)$ in Figure 10b. We again observe a negative reseeding bias for teams that have played the most games (Kentucky for the men and Tennessee for the women). This bias cannot be blamed on non-existent selection committees.

We would like to calculate conferences' expected reseeding to compare how conferences have performed in national tournaments over the years. Unfortunately, the teams in a conference have shifted, so that we cannot make a simple list of which teams go to which conference. Instead, for a particular year we will note which teams participated in which conference tournaments. Then in national tournaments we will record how conferences performed against each other in Figure 12. This approach does have the drawback that if a conference did not have a tournament in a particular year, then we are not able to identify any teams as coming from that conference, so that we group them into an unknown conference.

Wikipedia has many more tournaments to analyze. We tabulate the collegiate tournaments in Table 4, and note that we also can perform our calculations on brackets from the MLB, NBA, NFL, NHL, and WNBA. Because our analysis

Table 4: Tabulation of 172 collegiate tournaments with brackets in Wikipedia.

Sport	Men's		Women's	
	Conference	National	Conference	National
Basketball	32	8	32	5
Baseball/softball	22	4	9	2
Soccer	20	3	1	4
Ice hockey	14	5	0	1
Football	0	4	0	0
Volleyball	0	1	0	1
Field hockey	0	0	0	1
Lacrosse	0	1	0	0
Tennis	0	1	0	1
Total	88	27	42	15

does not uncover new information, we will not repeat it here.

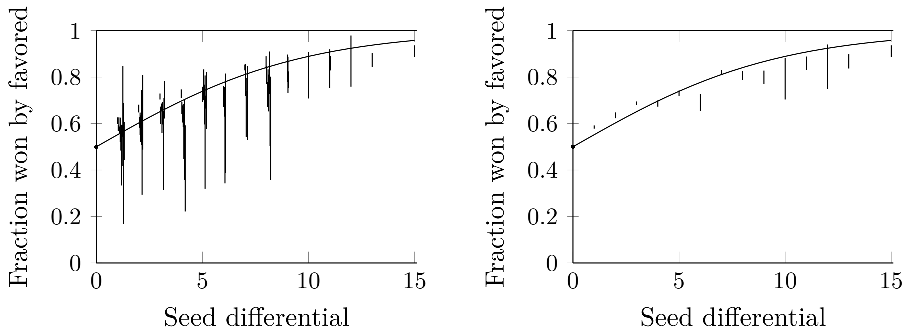
We do however provide one final plot of seed differential versus fraction of games won in Figure 13, using the entirety of our data. In Figure 13a, we keep identical seed differentials separate, as our habit. In Figure 13b, we coalesce the seed differentials to provide a single confidence interval for each seed differential.

Conclusions and Further Analysis

It is our hope that this analysis and its database will allow further exploration of tournament seeding. Consequently, the necessary Python code to reproduce our results is available at `https://github.com/teepeemm/bracket`. For the less programmatically minded, we provide an online interface at `https://sites.und.edu/timothy.prescott/bracket/` that can explore most of what we've done (counting wins and losses by seed differential for a particular team is more complicated).

Someone with experience in Python and probability or statistics would be able to extend this analysis in several ways. These possibilities are organized (we think) in order of decreasing complexity.

- Instead of collapsing seeds by only looking at the seed differential, can we find a probability estimation that also uses the average of the seeds? (Complicating this question is that the observed frequencies in Table 1 do not follow a clear pattern.) Does the formula become simpler if we use the better or worse team's seed instead of the average? Does this probability estimation depend on the tournament or the size of that tournament (for example, prior to 1985, the men's NCAA tournament had fewer than 64 teams)?



(a) Separated seed differentials. (b) Unified seed differentials.

Figure 13: Seed differential versus fraction of games that a favored team wins in their tournament, with 95% Wilson confidence intervals and the logistic best fit ($\beta = 0.208$).

- In the NCAA basketball tournaments, is there a significant difference for teams that qualified at large instead of teams that only qualified by virtue of winning their conference tournament? Does this question depend on the conference, and whether the conference is a traditional powerhouse conference or a “mid-major”? Can we predict a conference’s performance by tracking how many teams from that conference are in the tournament?
- Can we apply an intercept correction to teams with few wins in a tournament, and thereby provide more samples for our database, or will we be artificially inflating our numbers?
- Our choice to focus on μ instead of $\beta\mu$ means that a small β causes a large μ. Do we have notably different conclusions if we instead examine the logarithm of the odds against an evenly matched opponent, given by $-\beta\mu$?

Throughout the past century, conferences memberships have been in constant flux. Because that was usually a team or two at a time, however, it still makes sense to talk about a conference as a whole and how it performed. That is no longer true for the Pac 12, which in August 2024 exploded with 2, 4, and 4 teams going to the ACC, Big 10 and Big 12 conferences.

In all of our examination of tournaments, conferences, and teams, we must admit that most teams appear to be consistently correctly rated. Further exploration may show that \your_favorite_team is indeed underrated, or that \your_hated_rival is indeed overrated, but we must be careful to avoid torturing the data and p-hacking to arrive at these results. On the other hand, we now have tools to examine how teams have performed throughout the history of numerous tournaments.

Disclosure Statement

No potential competing interests were reported by the author.

References

1. Schwenk, A. J. What is the correct way to seed a knockout tournament? *Amer. Math. Monthly* **107,** 140–150. ISSN: 0002-9890. https://doi.org/10.2307/2589435 (2000).
2. Seltzer, S. & Simonson, S. Bracket challenge. *Math Horiz.* **31,** 20–23. ISSN: 1072-4117. https://doi.org/10.1080/10724117.2023.2285681 (2024).
3. Wittry, A. *Here's how to pick March Madness men's upsets, according to the data* https://www.ncaa.com/news/basketball-men/bracketiq/2018-03-13/heres-how-pick-march-madness-upsets-according-data (2024).
4. Chartier, T. Reflections on March Mathness. *SIAM News* **56,** 1, 4. ISSN: 1557-9573 (2023).
5. Jacobson, S. H. *Bracketodds* http://bracketodds.cs.illinois.edu/ (2025).
6. xqt. *Pywikibot* https://pypi.org/project/pywikibot/ (2024).
7. Wikipedia contributors. *2023 NCAA Division I men's basketball tournament — Wikipedia, The Free Encyclopedia* https://en.wikipedia.org/w/index.php?title=2023_NCAA_Division_I_men%27s_basketball_tournament&oldid=1285598021. [Online; accessed 4-June-2024]. 2024.
8. Brown, L. D., Cai, T. T. & DasGupta, A. Interval estimation for a binomial proportion. *Statist. Sci.* **16.** With comments and a rejoinder by the authors, 101–133. ISSN: 0883-4237. https://doi.org/10.1214/ss/1009213286 (2001).
9. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12,** 2825–2830 (2011).