Computing the Probability of Scoring a 2 in Disc Golf

Revised November 20, 2014

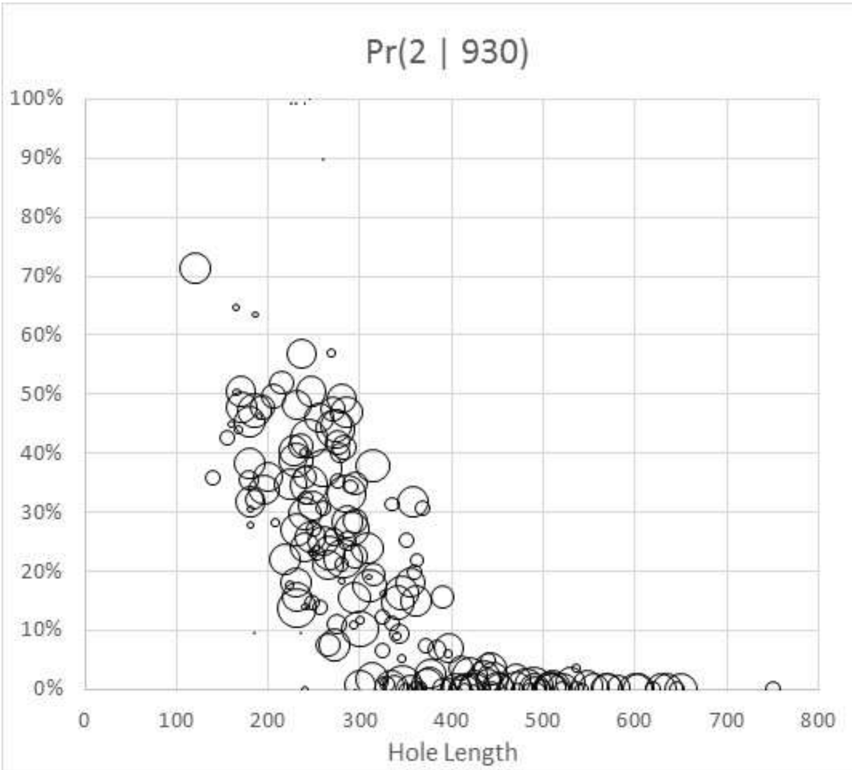
Steve West Disc Golf, LLC

Data: Scores from 2014 Am Worlds and related tournaments. The number of player-rounds contributing to the data for each hole was up to 500. The max was for 930-rated players for certain holes at Kaposia.

The average number of players contributing to the data for each hole by rating was as follows:

800	28.4
810	31.0
820	33.9
830	37.3
840	41.4
850	47.4
860	56.3
870	69.4
880	88.0
890	112.5
900	141.6
910	171.4
920	194.3
930	197.6
940	165.5
950	97.8
960	35.1
970	6.5
980	1.5
990	0.5
1000	0.0

For each hole, I computed the scoring distribution for each rating, and took the percentage of 2's (plus aces). Here is the plot for 930-rated players. Each circle is one hole. The size of the circle indicates the number of players contributing to the data for that hole.



Looking at this, 3 things become apparent:

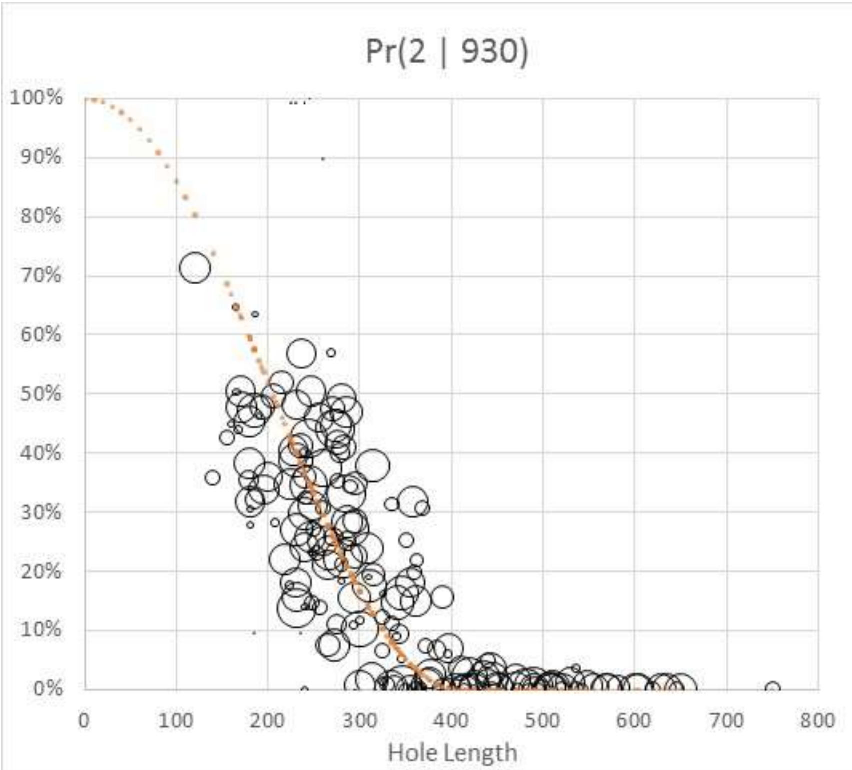
> There is a length beyond which a 2 becomes nearly impossible (the half-circles along the bottom are zeros).
>
> The probability of a 2 is not strictly linear with hole length.
>
> The probability of a 2 is not strongly tied to the length of the hole. Around the 250 foot length, for this rating there are holes where everyone scores a 2, and holes where no one scores a 2.

Thus, the function I choose to fit the data is

(1+COS(Pi * Length / Range)) / 2

The parameter "Range" can be thought of as the maximum length of hole for which a player of that Rating can think about getting a 2. Another way to think about it is that half of the Range is the length at which a player would expect to get a 2 about half the time.

I solved for the Range that reproduced the actual number of 2's recorded for each rating.



The fitted line represents the probability of a 2 for a hole of medium difficulty for a player with a 930 Rating.

This function gives an S-curve where the probability of a 2 approaches 100% as the length goes to zero, and where the probability gets smaller slower as the Length approaches the maximum Range.

This function is based on the Range for 930 rated players, a little over 400 feet. Higher ratings have bigger Ranges, because better players can score a 2 from farther out.

The pattern of Range by Rating is close to linear. So, I solved for the best fit line of Ranges by Rating. For a hole of medium difficulty, the Range can be found by

1.107173*Rating - 620

Here are samples:

Rating	Range
800	266
850	321
900	376
950	432
1000	487

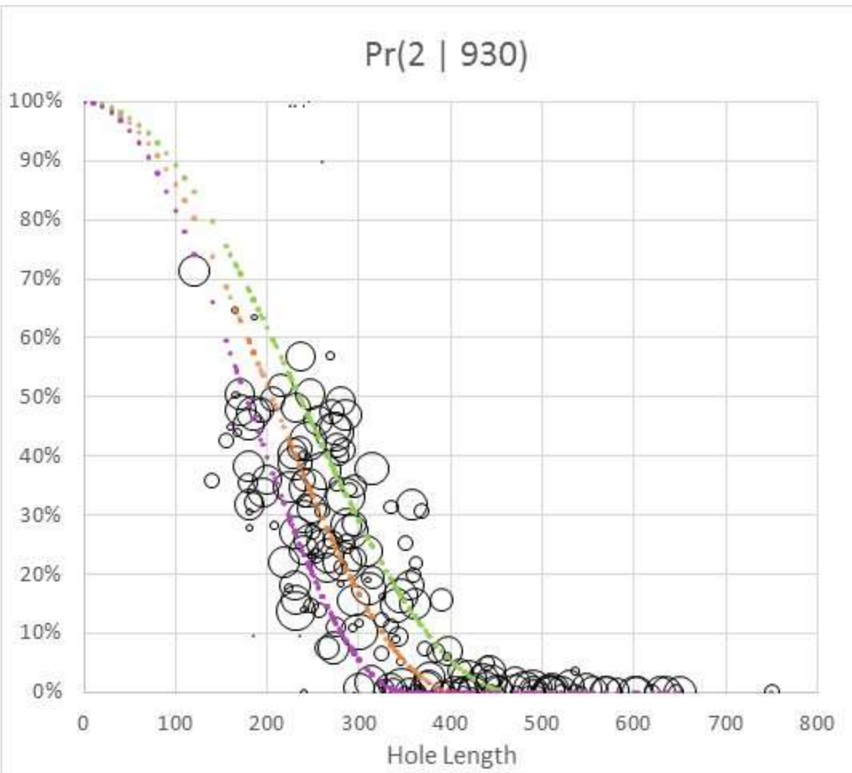
[Caution: There is very little data about players rated higher than 950. However, the smoothness of the patterns of parameters suggests that the formula could work fairly well for computing the probability of a 2 for the higher rated players.]

As is apparent from the first graph, the peculiar characteristics of the hole - other than length - are as important as length. The only thing to do is throw in the towel and add a parameter for hole difficulty. This covers all the characteristics of a hole other than length: elevation change, windiness, trees, mislabeled tee signs, etc.

I chose to define this parameter as how the hole relates to all other holes of that length – the percentile of difficulty.

To find the 25th percentile, I fitted a function to the middle of data from just the easiest half of the holes. (Those with probabilities higher than the function which fit all the holes.) For the 75th, I fit a line to the half of the holes that were more difficult.

Shown below are the lines for the 25th , 50th, and 75th percentiles of difficulty.



So, each Rating has a set of Ranges which varies by hole difficulty.

Example: find the probability of Red-level player (Rating of 850) scoring a 2 or better on a 250 foot hole which is judged to be somewhat easy (more difficult than 40% of holes of that length).

To compute the probability of a player getting a 2 on a hole, you need:

The Rating of the player,

The hole length,

A guess as to how difficult the hole is compared to all other holes of that length, and

These four constants:

ms	-0.503701
mi	233.85
cs	1.359024
ci	-736.93

Step 1. Using player Rating and the four constants, find the m and the c for the line which defines the Ranges by percentile for that Rating:

(ms * Rating) + mi = m =

(-0.503701 * 850 + 233.85 = -194.3

(cs * Rating) + ci = c

(1.359024 * 850) – 736.93 = 418.2

Step 2. Using the percentile and m and c from step 1, compute the Range for the Rating:

(%ile * m) + c = Range

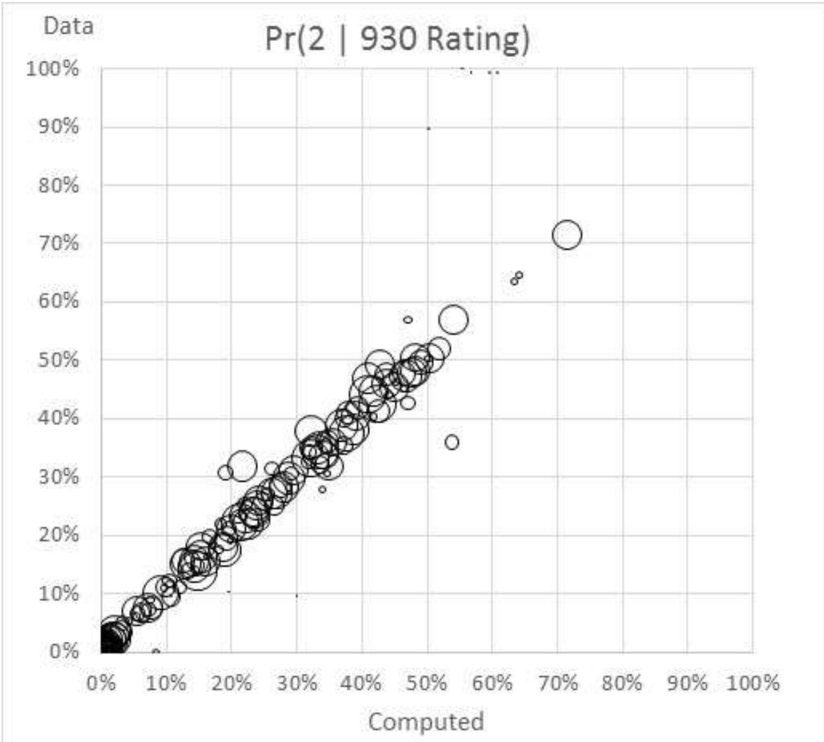
(40% * -194.3) + 418.2 = 340.5

Step 3. If hole length is longer than Range, the probability is zero. Otherwise, using the hole length and Range, compute the probability of a 2:

(1+COS(Pi * Length / Range)) / 2 = Pr(2)

(1+COS(Pi * 250 / 340.5)) / 2 = 16.4%.

To see how well this method could reproduce the observed probabilities, I ranked the holes by difficulty to assign a percentile to each hole. Then, using this percentile plus the hole length I computed the theoretical probability of a 2 for each hole. The following graph shows the computed theoretical probability (x-axis) to the observed data (y-axis).



This really just a test of whether extrapolating from the 75th and 25th percentiles works out at the edges of extremely weird holes. It seems to work OK.

The outlier hole that is so difficult (for its length) that it is "off the chart" is Hole 9 at Blue Ribbon Pines. This hole is so short that even if this hole is at the 99th percentile, it should still produce 57.9% 2's. Yet, only 35.9% of 930 rated players manage a 2. This hole can be played two ways: down a fairway with a 90 degree bend which severely restricts the length of the first throw, or a huge anhyzer up and over the trees. Only the second route gives much of a chance at 2. It would be easy to conclude that players may be missing out on many 2's by not knowing about the "secret" anhyzer route.

The hole that seems to be easier than should be possible (for its length) is hole 1 at the Valley (winter pin position). At 357 feet, even if the hole were at the 1st percentile of difficulty, there should be only 21.6% 2's. However, it actually produced 31.9%. It is not clear why this hole should be so much easier to 2 (for its length) than all other holes. It is downhill, but not extremely so. It is fairly wide-open, but has a tree right by the target in the hyzer flight path. When this data was taken, there was out of bounds all along the right side of the fairway. Perhaps the threat of OB corralled the throws toward the basket. Or, perhaps it is because this is the hole local players throw over and over again to warm up on.

We can conclude that the method is good at predicting the probability of a 2 for 98% of the holes. For the 2% at the tails, holes can get excessively more difficult or easier.

However, most of the holes in this data set are the product of a single design philosophy. Perhaps extra-difficult or extra-easy holes are not represented here. When more holes from other designers are added to the data set, perhaps the limits of what it takes to be at the 1st percentile and 99th percentile will expand to encompass hole 9 at BRP and hole 1 at The Valley.

Although the assumption of linear difficulty with percentile is not accurate at the tails, it does let us compute a theoretical maximum and minimum probability of a 2 for a given hole length and rating.

Here are the probabilities of scoring a 2 by rating and length for the easiest, average, and most difficult holes.

1st percentile (easiest)

Length \ Rating	700	750	800	850	900	950	1000	1050
50	87%	92%	95%	96%	97%	98%	98%	99%
100	55%	72%	81%	86%	90%	92%	94%	95%
150	20%	45%	61%	71%	78%	83%	86%	89%
200	1%	19%	39%	53%	63%	71%	76%	81%
250	0%	3%	18%	34%	47%	57%	65%	71%
300	0%	0%	5%	18%	32%	43%	52%	60%
350	0%	0%	0%	6%	18%	30%	40%	49%
400	0%	0%	0%	0%	7%	18%	28%	37%
450	0%	0%	0%	0%	1%	8%	17%	27%
500	0%	0%	0%	0%	0%	2%	9%	17%
550	0%	0%	0%	0%	0%	0%	3%	10%
600	0%	0%	0%	0%	0%	0%	0%	4%
650	0%	0%	0%	0%	0%	0%	0%	1%
700	0%	0%	0%	0%	0%	0%	0%	0%

50th percentile

Length \ Rating	700	750	800	850	900	950	1000	1050
50	76%	87%	92%	94%	96%	97%	97%	98%
100	28%	54%	69%	78%	84%	87%	90%	92%
150	0%	19%	40%	55%	66%	73%	78%	82%
200	0%	1%	14%	31%	45%	56%	64%	70%
250	0%	0%	1%	12%	25%	38%	48%	56%
300	0%	0%	0%	1%	10%	21%	32%	42%
350	0%	0%	0%	0%	1%	9%	18%	28%
400	0%	0%	0%	0%	0%	1%	8%	16%
450	0%	0%	0%	0%	0%	0%	1%	7%
500	0%	0%	0%	0%	0%	0%	0%	2%
550	0%	0%	0%	0%	0%	0%	0%	0%

99th Percentile (hardest)

Length \ Rating	700	750	800	850	900	950	1000	1050
50	47%	72%	83%	88%	92%	94%	95%	96%
100	0%	19%	43%	59%	70%	77%	82%	85%
150	0%	0%	8%	25%	41%	53%	62%	69%
200	0%	0%	0%	3%	15%	29%	40%	50%
250	0%	0%	0%	0%	1%	9%	20%	30%
300	0%	0%	0%	0%	0%	0%	6%	14%
350	0%	0%	0%	0%	0%	0%	0%	4%
400	0%	0%	0%	0%	0%	0%	0%	0%

We can also solve for the hole length that would produce a desired probability. For example, Close Range Par currently sets "Close Range" at the point where a player has a 90% chance of getting up and down in 2. Here are the hole lengths where a player has a 90% chance of scoring a 2:

Diffficulty \ Rating	700	750	800	850	900	950	1000	1050
1%	44	58	71	85	99	113	127	141
50%	32	43	54	66	77	88	100	111
99%	20	29	37	46	55	64	73	82

These are close to the hole lengths that CRP uses. (60, 70 , 80, 90, 100 for ratings 800 through 1000).

For 1000-rated players the Close Range length from CRP is right at its target for a hole of average difficulty.

For 800-rated players, the length of 60 feet is just 6 feet too long. 60 feet is the length where players would get a 2 90% of the time if the hole is easier than 66% of holes. Or, players on a hole of average difficulty would have only an 88% chance of getting a 2.

Adjustments to Effective Length

I selected 61 holes that have a significant probability of a 2 out of the Am Worlds holes for testing the adjustment factors.

Length alone accounted for 30% of the variation in probability of getting a 2.

I found that every foot of elevation up adds 42 inches to the effective length of the hole. However, every foot of elevation down subtracts only 7 inches from the effective length.

If a hole turns left, it is 7% easier, and if it turns right, it is 15% harder.

With those factors included, the computed effective length accounts for 44% of the variation in difficulty of 2ing.

I then solved for foliage factor, on a scale of 0 (no foliage) to 8 (pinball). The average for these holes was 2.5. Zero foliage holes are 24% easier, and 8 foliage holes are 51% harder. Or, about 9% harder for each step up in foliage.