A New, Even More Refined Method of Setting Par

Based on the Expected Score as Derived from the Scoring Distribution

By Steve West Disc Golf, LLC

June 25, 2018 – Updated June 29, 2018

A Prior Iteration: Errorless Par

Previously, I developed a method of setting par which was based on modelling the scores that would result from errorless play. The basic premise was that each throw had a certain chance of being errorless: meaning that it was not bad enough to increase the eventual score on the hole.

This one-parameter model was quite successful in replicating par values as set by expert TDs, or as commented on in videos when they disagreed with the TD.

However, that model only implicitly dealt with throws that were a result of error or were good enough to lower the expected score. It assumed high scores were a result of errors and low scores (below an arbitrary threshold) were not common enough to be expected.

Another defect of Errorless Par is that it requires a threshold value to be established independent of the calculation.

Errorless and Heroless Par

Now, I will look at a two-parameter model.

The development of the model starts with the idea that each hole has a single score which an expert disc golfer would be expected to make with errorless play under ordinary weather conditions.

If you take away all the scores that resulted from errors, and take away all the scores that resulted from "hero throws" that shaved a throw, then the scores that remains must be the score the expert was expected to make.

Working the other way, a scoring distribution starts its development as just one expected score. Unless something unexpected happens, all experts would get that score.

The two things that could happen to change the scoring distribution are: an error during a throw could cause the expert to expect to make more throws before finishing the hole, or an unexpectedly good "hero" throw could cause the expert to expect to finish the hole with fewer throws.

As the process is carried forward, a scoring distribution is built up. Given the right choice of two parameters – the probability of making an error, and the probability of making a hero throw - this distribution can look very much like a regular disc golf scoring distribution.

For example, on a hole where the expected score is 3:

After the first throw, the expert may be in a position where they expect to get a 2, or expect to get a 4, or most likely, still expect a 3. Say 10% of all throws result in an unrecoverable error, and 15% of all throws are hero throws that save a throw. Then after the first throw 10% of the players are now expecting a 4, while 75% of the players are still expecting a 3, and 15% are now expecting a 2.

After the second throw, because of errors, another 10% of those players who were still expecting a 3 will now expect a 4, and 10% of the players who were expecting a 4 will now expect a 5, and 10% of the players who were expecting a 2 will now go back to expecting a 3. Also, 15% of the players who were expecting 3s and 4s will get hero throws and expect 2s and 3s.

Real Life Example

The following is the typical scoring distribution (derived from hundreds of actual scoring distributions) for a hole that averages 3.0. The mode, median, and average are all 3, so most everyone would agree this is a par 3.

Score	Frequency
1	0.0%
2	24.3%
3	55.3%
4	17.1%
5	2.8%
6	0.4%
7	0.1%
8	0.0%
9	0.0%

And here is the scoring distribution the two parameter model calculates with a chance of error set at 10.4% and a chance of hero set at 14.7%.

Score	Before Any Throws	After 1 Throw	After Two Throws	After all Throws
1	0.0%	0.0%	0.0%	0.0%
2	0.0%	14.7%	24.3%	24.3%
3	100.0%	74.9%	59.1%	55.3%
4	0.0%	10.4%	15.5%	16.4%
4	0.0%	0.0%	1.1%	3.4%
5	0.0%	0.0%	0.0%	0.6%
6	0.0%	0.0%	0.0%	0.1%
7	0.0%	0.0%	0.0%	0.0%
8	0.0%	0.0%	0.0%	0.0%

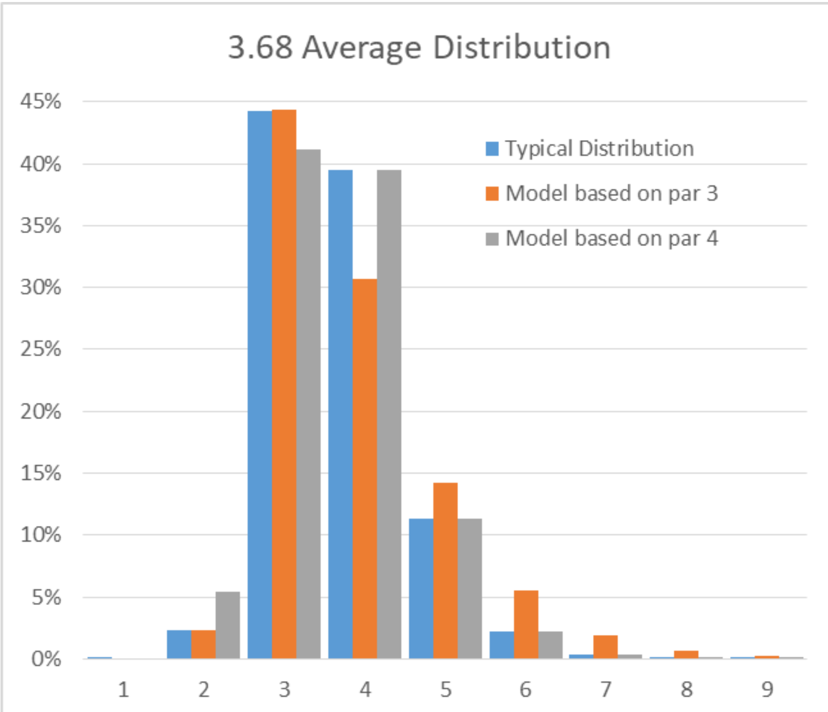
The final model distribution looks very much like the empirically derived typical distribution, so it would seem that the model may be acting very much like the underlying mechanisms which cause players to get scores other than the expected score.

However, so far, none of this could be used to set par. And, looking at only this obviously par 3 hole won't provide much insight into how to select from many potential par values.

For that, we need to look at a hole that is more ambiguous about its par.

Par ? Example

Here is a graph of the typical distribution of a hole that is arguably either par 3 or par 4. It averages 3.68. Also shown are the distributions of two of the best-fit (sum of absolute difference) two-parameter models. One starts with 100% expected 3s and the other starts with 100% expected 4s.



It is not obvious which model better replicates the distribution (except that the par 3 model better reproduces the frequency of 3s and the par 4 model better reproduces the frequency of 4s). So that doesn't give us a good way to choose which par is better.

However, when the model starts with 3 as the expected score, it solves for a 75% chance of each throw being normal (23.4% errors and 1.5% heroes). When the model starts with 4 as the expected score, it solves for only a 64% chance of each throw being normal (8.2% errors and 23.3% heroes).

It is possible to create fitted model scoring distributions based on other starting expected scores. For example, with 2 as the expected score, the errors are 46.0% and the heroes are 0.2%, for only 53.8% expected throws. With 5 as the expected score, the errors are 6.9% and the heroes are 49.4%, for only 44% expected throws.

How to Choose Which Par

I propose setting par to be whichever expected score generates the highest percent of normal throws when the model is fitted to the actual distribution.

- Since par is “expected”, it makes sense that the par which should be chosen is the one for which the model has the most “expected” throws.
- The pars that we know are “wrong” are also those that result in the fewest normal throws.

The two-parameter method thus has the distinct advantage of directly choosing a par without the need to reference any numbers that do not come from the actual scoring distribution.

Compare this to the one-parameter method which requires an arbitrary threshold, above which a distribution qualifies for a specific par.

Further Work

The two-parameter model presented here does not account for disastrous and miraculous throws that cost or save the player two or more throws. To be able to cover island holes, the model will need to include a probability of a throw that increases the expected score by two throws. Miraculous throws that save two throws are probably rare enough to be safely ignored, as the big end of any two-score swings would most likely feel more like a disaster than the lack of a miracle.

The problem of solving for the parameters is a difficult one. So, far I’ve done it by iteration. Perhaps it cannot be solved directly. A practical method needs to be found.

Then, the method needs to be tested against a lot of actual scoring distributions to see how it performs.

Update/Addendum June 29, 2018

- I will now call throws that reduce the expected score by one "Birdie quality". Likewise, "Bogey quality", "Eagle quality", and "Double Bogey quality".
- I reworked the spreadsheet and added parameters for the percentages of Double-birdie and –bogey throws.
- So, I need a name for the new method, as it now uses four parameters. Here, I will just use New Method.

Utah Open

I applied the new method to the scores from the Utah Open 2018. The new method was developed to account for the Bogey quality throws which exist within every scoring distribution, so it should differ more from other methods when there is a lot of punishment. The Utah Open is known for having a lot of one-throw punishment in the form of Hazards which are too near to the green to be avoidable.

Here are the course pars, Errorless method pars, and the new method pars:

Hole #	1	2	3	4	5	6	7	8	9
Feet	711	369	405	453	696	261	546	330	819
Average	4.03	2.91	2.98	3.19	3.86	2.64	3.40	2.93	3.74
Course Par	4	3	3	3	4	3	3	3	4
Errorless at 14/18ths	4	3	3	3	4	2	3	3	4
Errorless at 13/18ths	4	3	3	3	3	2	3	3	3
New Method Par	4	3	3	3	3	2	3	3	4
Hole #	10	11	12	13	14	15	16	17	18
Feet	285	303	348	524	238	651	408	375	207
Average	2.8	2.58	2.87	3.08	2.95	4.2	3.17	3.01	2.98
Course Par	3	3	3	3	3	4	3	3	3
Errorless at 14/18ths	3	3	3	3	3	4	3	3	3
Errorless at 13/18ths	3	3	3	3	3	4	3	3	3
New Method Par	3	3	3	3	3	4	3	3	2

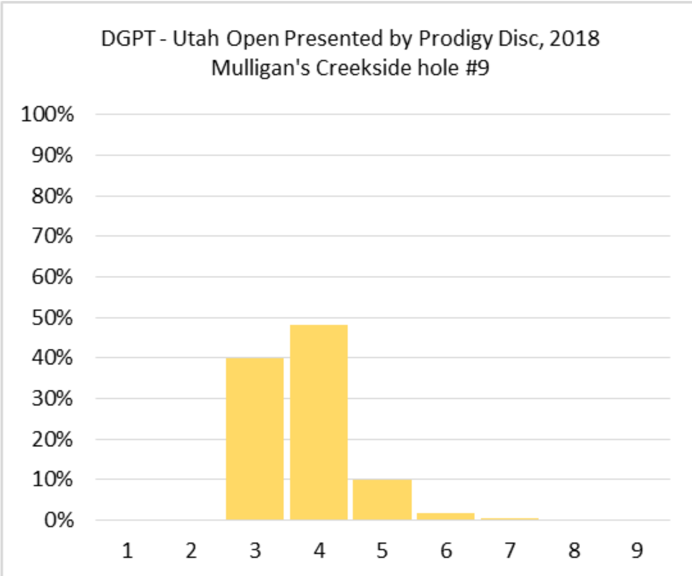
(Because Errorless par is dependent on an arbitrary threshold, I matched the new method's total par with the 13/18ths threshold.)

The total course par was 58, which averaged a 987 round rating. Errorless par at 57 averaged a 996 round rating, while errorless par at 13/18ths and the new method par averaged a 1013 round rating.

As I am using the scoring distributions for 1000-rated players to calculate par, it would make sense that a par round should be rated higher than 1000. The course offers the opportunity for punishment, so merely avoiding that punishment should get a score better than SSA.

One point for the new method (and, for tightening the threshold for errorless par).

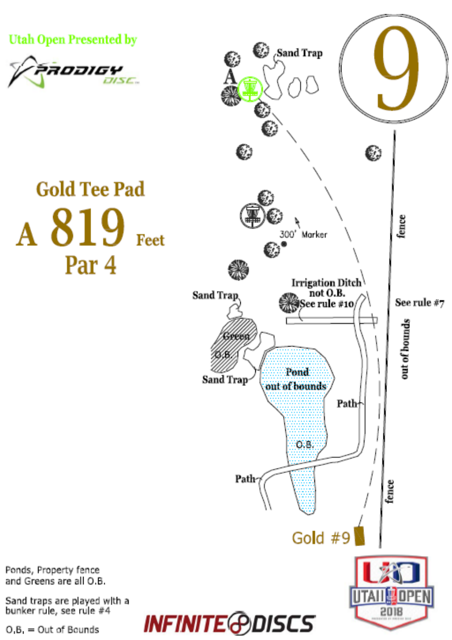
Hole 9



Histogram of the 1000-rated Scoring Distribution. Errorless par at 13/18ths would call this a par 3, and at 14/18ths a par 4. The difference is whether the number of 3s was enough to get over the threshold.

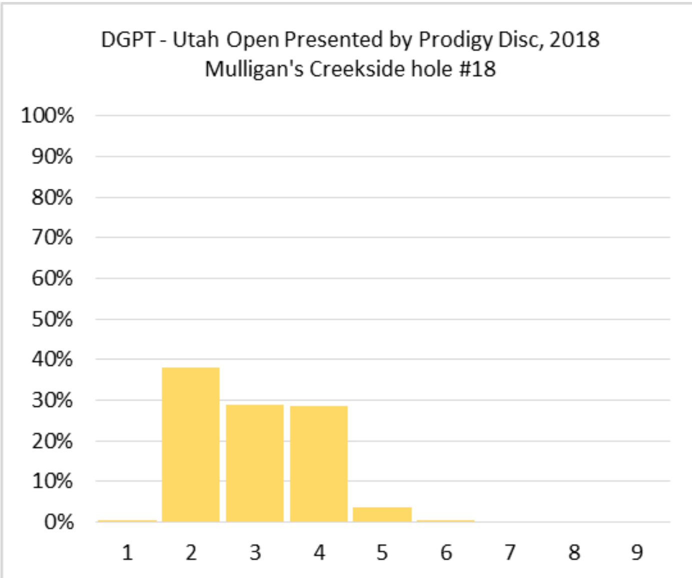
The new method calls this a par 4, with 18% of the throws being Birdie quality, and 5% being Bogey quality, with hardly any Doubles. (There are more than 18% threes because the thrower had three chances to make a birdie quality throw to get a 3.)

If this hole was actually a par 3, then this distribution tells us that there was no chance of a birdie quality throw, and a 25% chance of a bogey quality throw.



It seems more plausible that this hole does indeed offer opportunities for birdie quality throws, by simply throwing farther. Another point for the new method.

Hole 18

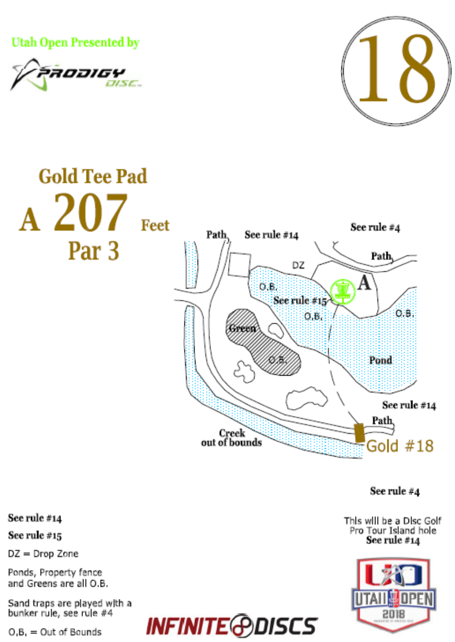


Errorless par would call this a par 3, because there aren't enough 2s.

The new method would call this a par 2 with a 38% chance of a bogey quality throw.

If this was a par 3, it would need to have a 25% chance of bogey quality throws AND a 33% chance of birdie quality throws. It does not seem likely that 33% of throws on a 207 foot hole were better then expected – unless they were all aces. The only non-ace way I can think of to have a birdie quality throw on this hole would be to make the putt from the DZ. Yet, there are not enough throws from the DZ (some fraction of the 25% chance of a bogey quality throw) to get to a chance of 33% of all throws.

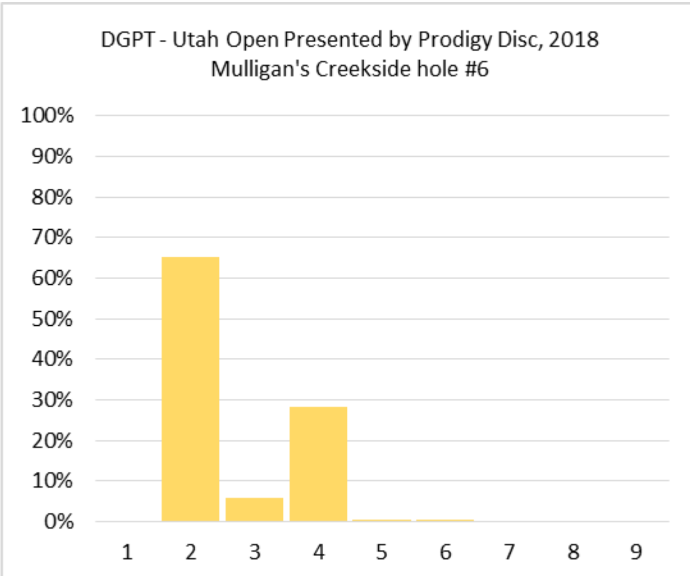
Also, when a hole has a small green that is surrounded by OB, it seems reasonable that there would be a higher percentage of bogey quality throws.



Another point for the new method.

Hole 6

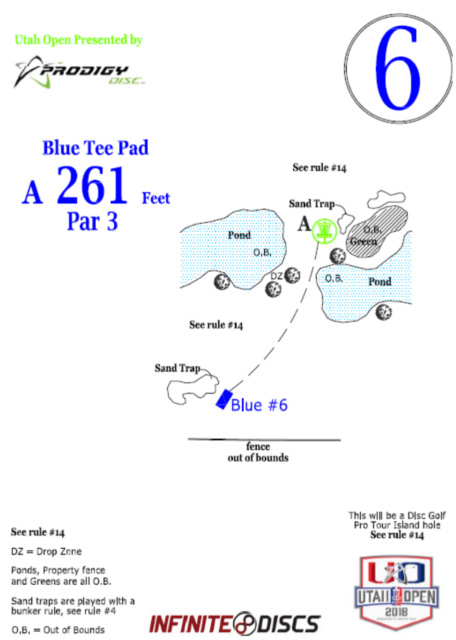
While both errorless and the new method agree this should be a par 2, the inflammatory aspects of calling any hole a par 2 warrant closer scrutiny.



No one is expecting a 3 on this hole. Other than the inflated average score, there is no reason to call it a par 3. If this were a par 3, there would be a 63% chance that each throw would be birdie quality, with a 15% chance of a bogey quality throw. That leaves only 20% of throws to be normal. Expecting most throws to be birdie quality makes no sense. I also doubt that the majority of 1000-rated playerr who made a bad throw felt it only cost them one throw.

As a par 2, 80% of throws are normal, with a 4% chance of making a bogey quality throw (usually missing the putt from on the green) and 15% chance of making a double bogey quality throw, with a very small chance of making a birdie quality throw.

We must also look at how this hole would act as a par 4 – the best candidate for the expected score if 2 is ruled out. As a par 4, 53% of throws would be normal. Most of the throws being normal is good. As a par 4, all it offers is a 43% chance of making a double birdie quality throw. So, if the label of par 2 is so repugnant that reality must be denied, then this hole could be called a really easy-to-eagle par 4, but not a par 3.



Conclusion of June 29 Update

The new method picked reasonable, defendable, pars in every case on this course. Further, it provides insights into the percentage of throws that maintain, increase, or descrease the final score.

Thanks to the Utah Open for the unauthorized use of their caddy book.