

Note: These are mostly unedited notes. For a summary of this work, see: <https://doug-osborne.com/darts/>

Part 2 - A Programmer Actually Plays Darts

At the end of [Part 1](#), I posed several potential issues with using a single bivariate normal distribution to represent an actual human throwing darts.

In this part, I will put these issues to the test by throwing a lot of darts. Of course, the throws are not worth much if they are not tracked. But first, here's a look at the electronic dartboard that I purchased:

The WIN.MAX Electronic Soft Tip Dartboard



The image shows a regulation dartboard above the dartboard I used for this experiment. While the size of the scoring sectors look about the same - there is one exception: the double bullseye is significantly larger for the electronic dartboard.

The other noticeable difference is that the eDb has a thick, black "spider" separating each scoring sector, while the regulation dartboard has a much thinner silver wire. The spider is designed to reduce bounce outs - darts that hit the spider usually end up in one of the holes adjacent to the spider.

The result is that the scoring sectors are larger than they appear - this makes the biggest difference for the smaller, high scoring sectors. My estimates are that the double, triple, and double bull sectors are about 33% larger than a regulation dartboard, with the single bull about 25% larger.

There are many other electronic dartboards available, some with different designs - but this design is possibly the most common, and also probably the closest to a regulation dartboard. Some, for example, having cartoonish sized double/triple rings:

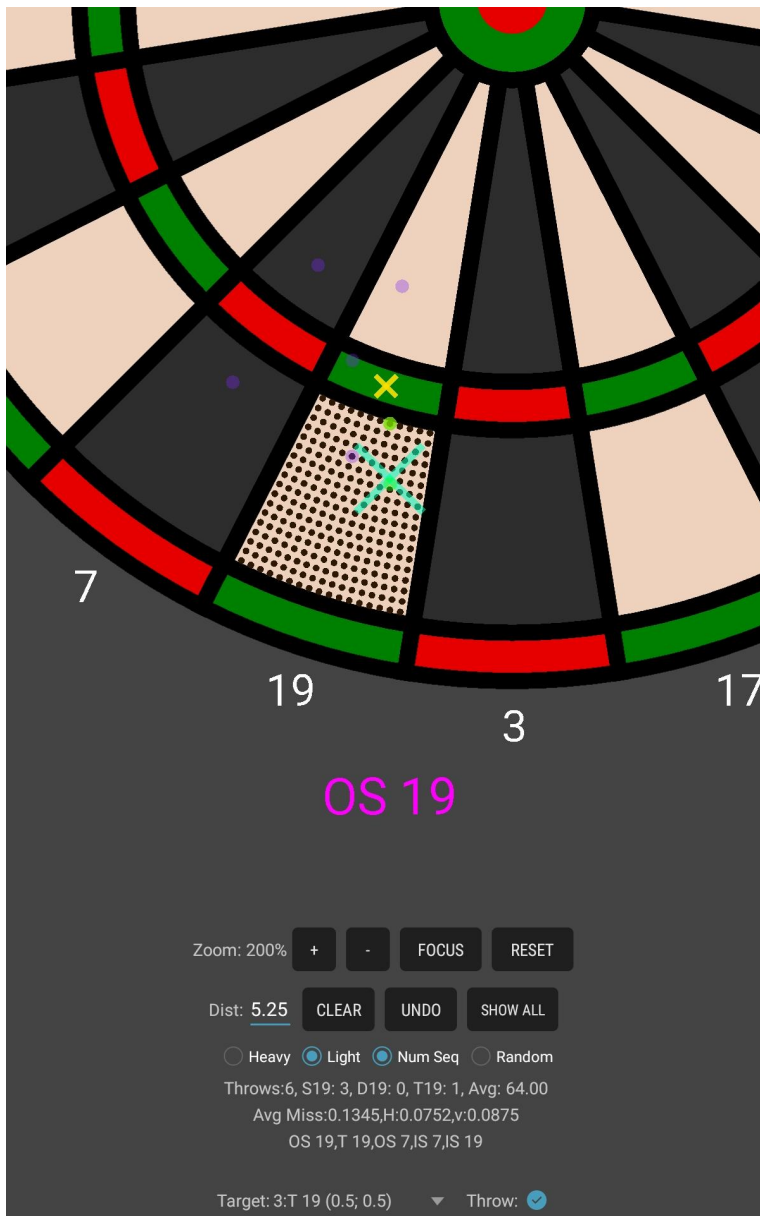


The Dart Tracker App

Initially, I considered tracking only the scores and sector for each throw. Ultimately I decided that tracking the location of the throws and not just the scores would be worth the initial investment in time, as it would require significantly fewer throws for an accurate result.

The decision was aided by my experience writing [puzzle games](#) for Android, which at least solved the technical detail of my computer being in a different room than my dartboard.

In the first version of the app, I estimated the location of each dart within each sector by eye. However, I found myself spending a lot of time looking back and forth between the dartboard and my app, so I decided to take advantage of a feature of the electronic dartboard:



The number of holes in each row and sector is exactly the same as my dartboard. While it might seem harder to find the exact hole corresponding to each throw, it became a lot easier for me than the version without the holes once I learned how to do it mostly without counting.

Let's Throw Some Darts

Before we begin, I will admit to a couple flaws to this study:

- Only one person (me) threw darts, though if anyone is interested I can share the dart tracking app.
- The researcher and test subject are the same person, and neither was blinded.
- I analyzed the data on a daily basis, which almost certainly influenced the results.

I threw a lot of darts - about 30,000 over a period of about 2 months. There were three main phases to the testing:

- The initial phase, where I mostly threw where I wanted while making changes to the app.
- The second phase, where I threw at the same targets as the first, but followed a plan.
- The third phase, where I threw at all 20 numbers on the dartboard in each session.

For the first two phases, I aimed at the four main targets that I did back in my fantasy dart leagues: the bullseye, the triple 20, the triple 19, and the triple 14.

I threw darts in 100 throw sessions, all at the same target, and from four distances:

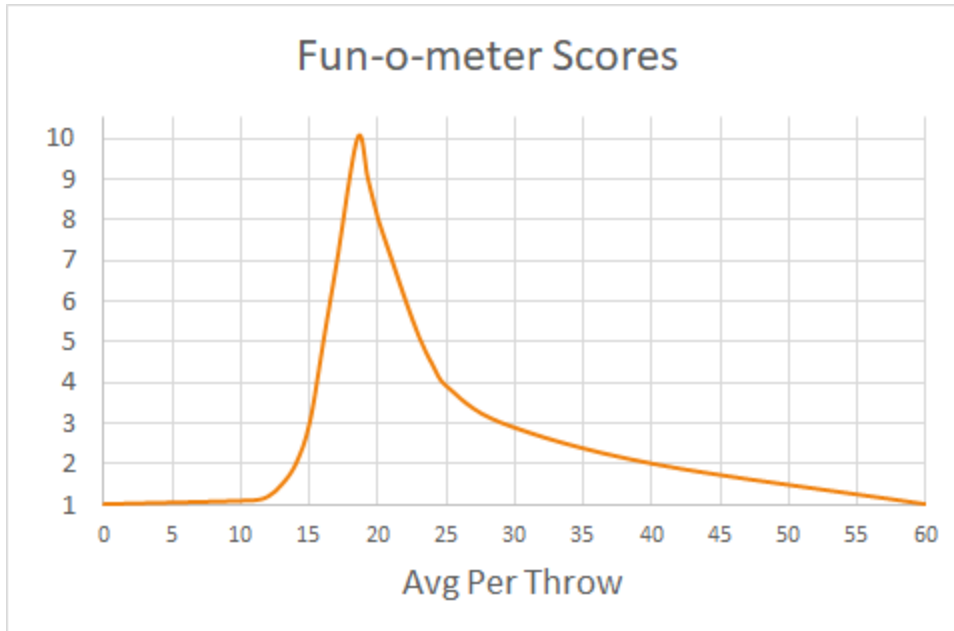
1. 7 feet 9 inches (about regulation distance)
2. 6 feet 4 inches
3. 5 feet 3 inches
4. 4 feet 6 inches (phase 1 only)

The following table shows how accurately I threw at each distance:

Dist (feet)	σ (mm)	Avg
4.5	16.4	22.7
5.25	21.3	18.3
6.33	26.9	16.2
7.75	33.7	14.8

The average is per throw as opposed to per turn to illustrate a following point.

I am not a particularly skilled dart player, which is why I mostly threw from closer than the regulation distance - where my average score is less than a point higher than the average of a random point on my dartboard (about 13.97). The following highly scientific data supports my decision to cheat:



Fun-o-Meter Scale



1 (No Fun)

5 (Meh)

10 (Super Fun)

As you can see, darts is most fun when your average score is a bit below 20 per throw. Why aren't higher averages more fun? For one, once your average is above 20, a single of anything is a "bad" throw, thus, you need to hit a triple, which is kind of hard to do, unless you're really good at darts. Triples are fun to hit for most people, but if you're really good at darts, triples are relatively boring, which makes darts no fun, unless you're so good that you make a lot of money from playing darts.

The above chart applies to all dartboard types and distances except for the regulation dartboard and distance, which has one significant difference:



Phase 2

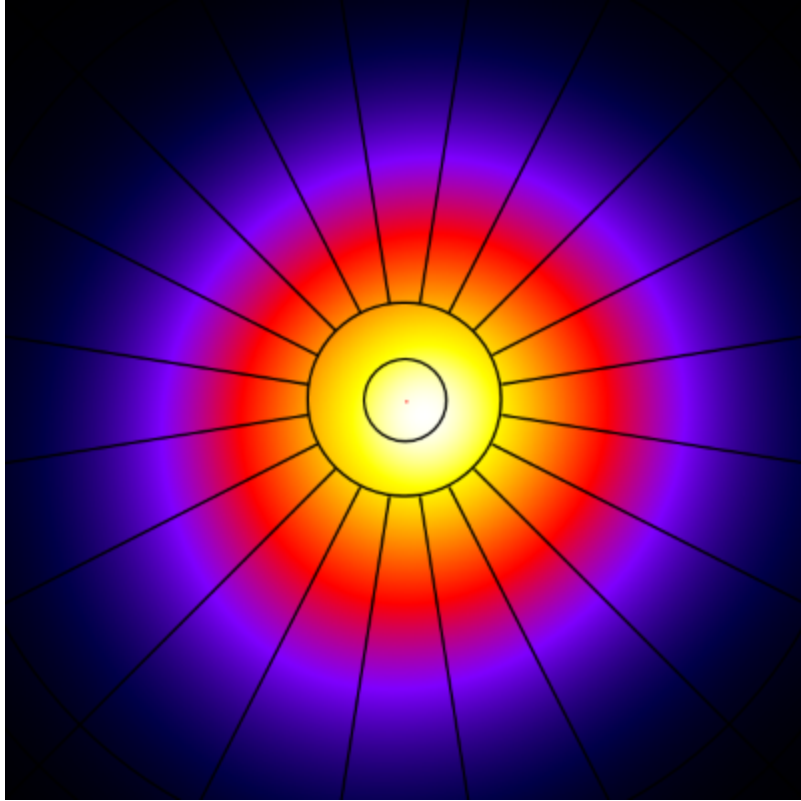
For this phase I threw for 10 sessions (1000 darts) at each of the four targets from each of the first 3 distances - a total of 12000 darts.

Creating a Pdf from Dart Throws

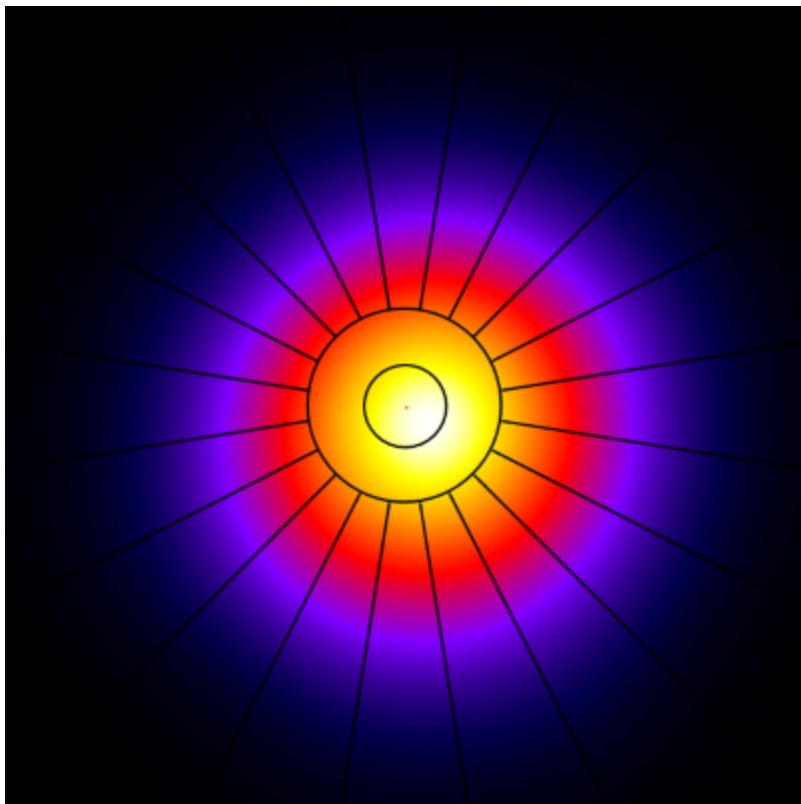
The function I wrote to create a probabilistic density function (pdf) from my throws uses the pdf from Part 1 (a Gaussian density function) to assign a weight to each of the throws for each point on the dartboard. As such, it will be Gaussian-like in appearance, and is designed so that a pdf based on a simulation of Pigfoote or Duck's throwing darts will eventually match their respective pdfs in part 1.

The following images show the pdf created by this function using my 4000 throws from each distance. The throws were first normalized to a single target (the bullseye):

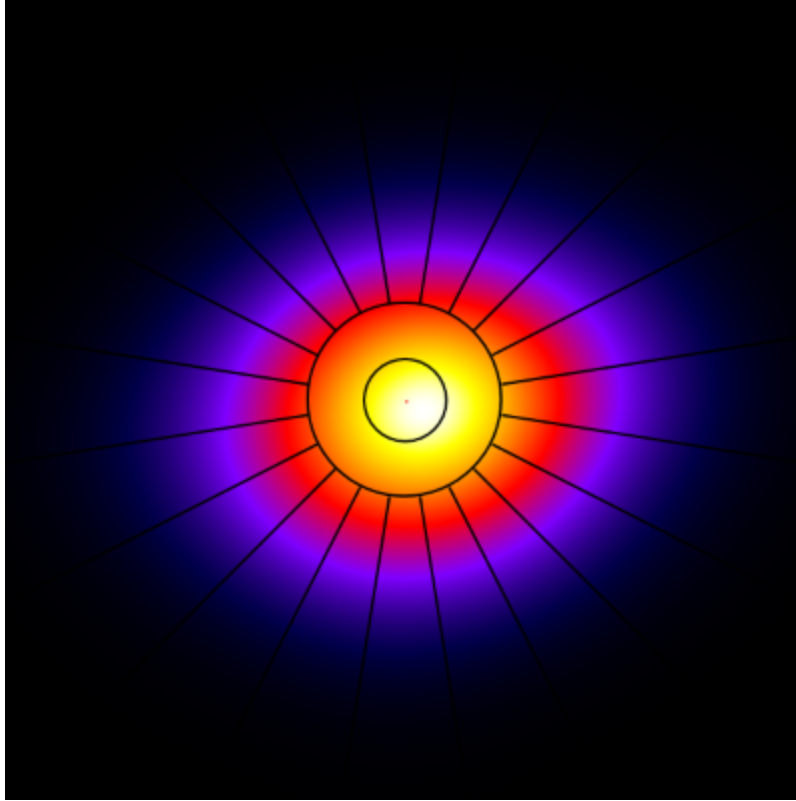
7.75 feet



6.33 feet



5.25 feet



The first two pdfs are boring-looking and generally circular - though you will notice that I tended to miss slightly to the right and below my target (strangely, less so from the farthest distance). The football shape of the 5.25 pdf does not seem to fit with the other two. Here is a table of numbers:

Dist	7.75	6.33	5.25
Avg	44.5	48.7	53.9
$x\sigma$	33.6	26.4	25.8
$v\sigma$	33.8	27.4	20.9
H/V	0.99	0.96	1.23

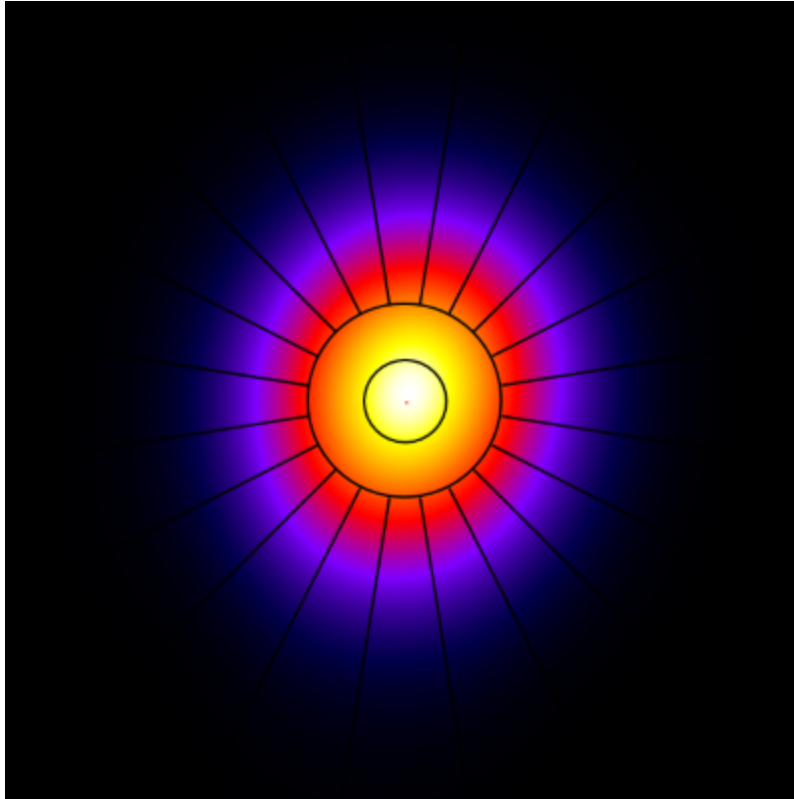
Moving from 6.33 to 5.25 feet, my accuracy barely improved horizontally (0.6 mm), but I was far more accurate vertically (6.5 mm), leading me to:

Unscientific Conclusion About My Dart Throwing That Might Not Apply To Anyone Else

#1 - My throwing looks like a football from shorter distances, and like someone who doesn't quite know how to draw a circle from longer distances.

It makes sense that moving further away would decrease one's vertical accuracy more than horizontal accuracy since the speed of a dart is a variable that increases in influence along with distance. This is exactly what I expected when I started this experiment before Phase 2. However, it is the *opposite* of what I thought I found out before phase 2.

My first few thousand dart throws in phase 1 were from 5.25 feet, with over 6000 total throws. Here is the pdf from those throws:



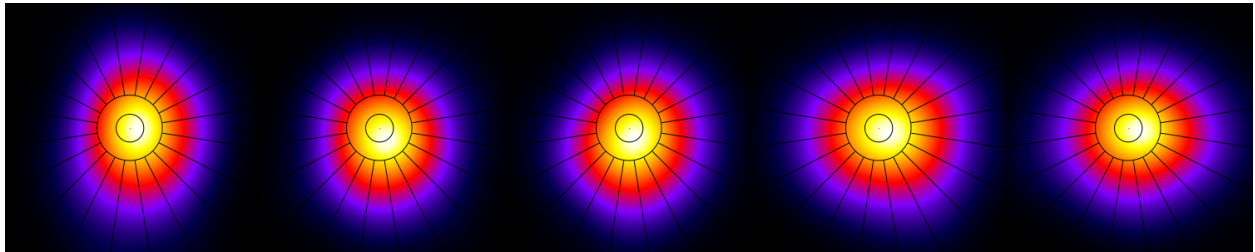
This would match my throwing in phase 2 from this distance reasonably well if I rotated the image clockwise, but I didn't. And while one might think that I would improve at throwing darts after some 10,000 throws, that wasn't the case either:

	Phase 1	Phase 2
Avg	56.0	53.9
$x\sigma$	19.4	25.8
$v\sigma$	23.3	20.9
H/V	0.83	1.23

What Happened to My Throwing?

For phase 2 I threw 3600 darts from 7.75 feet, 4000 from 6.33 feet, then 4000 from 5.25 feet (then the last 400 from 7.75 feet, which I had avoided because of the low fun-o-meter score).

The migration of my dart throwing (from a tall pdf to a wide one) seemed to occur mostly while I was throwing from the middle distance, which I did for a period of 5 days. From left to right, here are the pdfs from those 5 days:

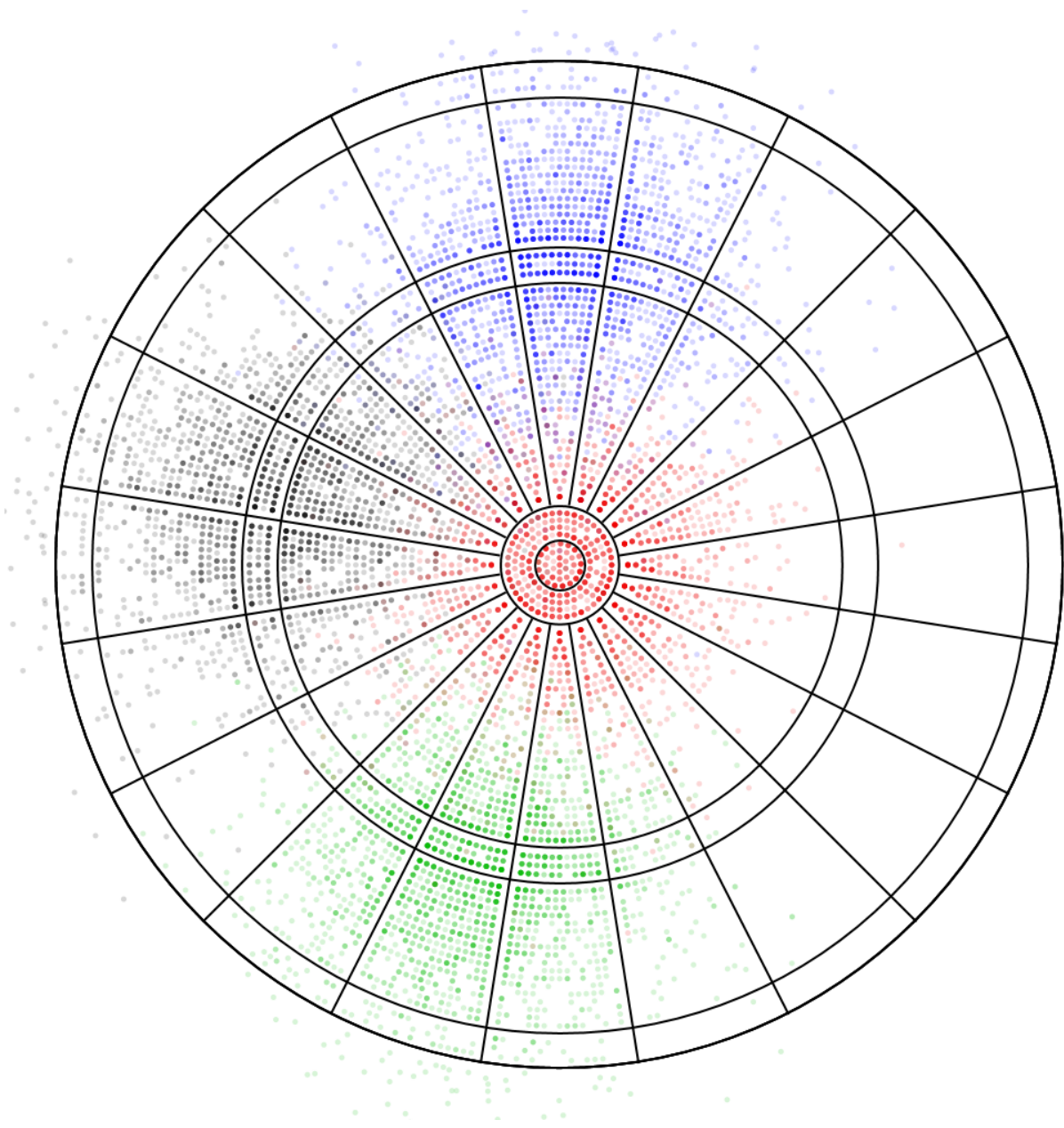


Unscientific Conclusion About My Dart Throwing That Might Not Apply To Anyone Else

#1 Which Negates The Other One - *I don't throw darts the same every day, but I am somewhat more likely to throw like I did yesterday than I did a month ago.*

So far, I've shown a few pretty images (that I claimed represent my dart throws), a few dartboards, and a screenshot of an app that I claimed to write - none of which proves that I actually threw any darts. Behold, the proof*:

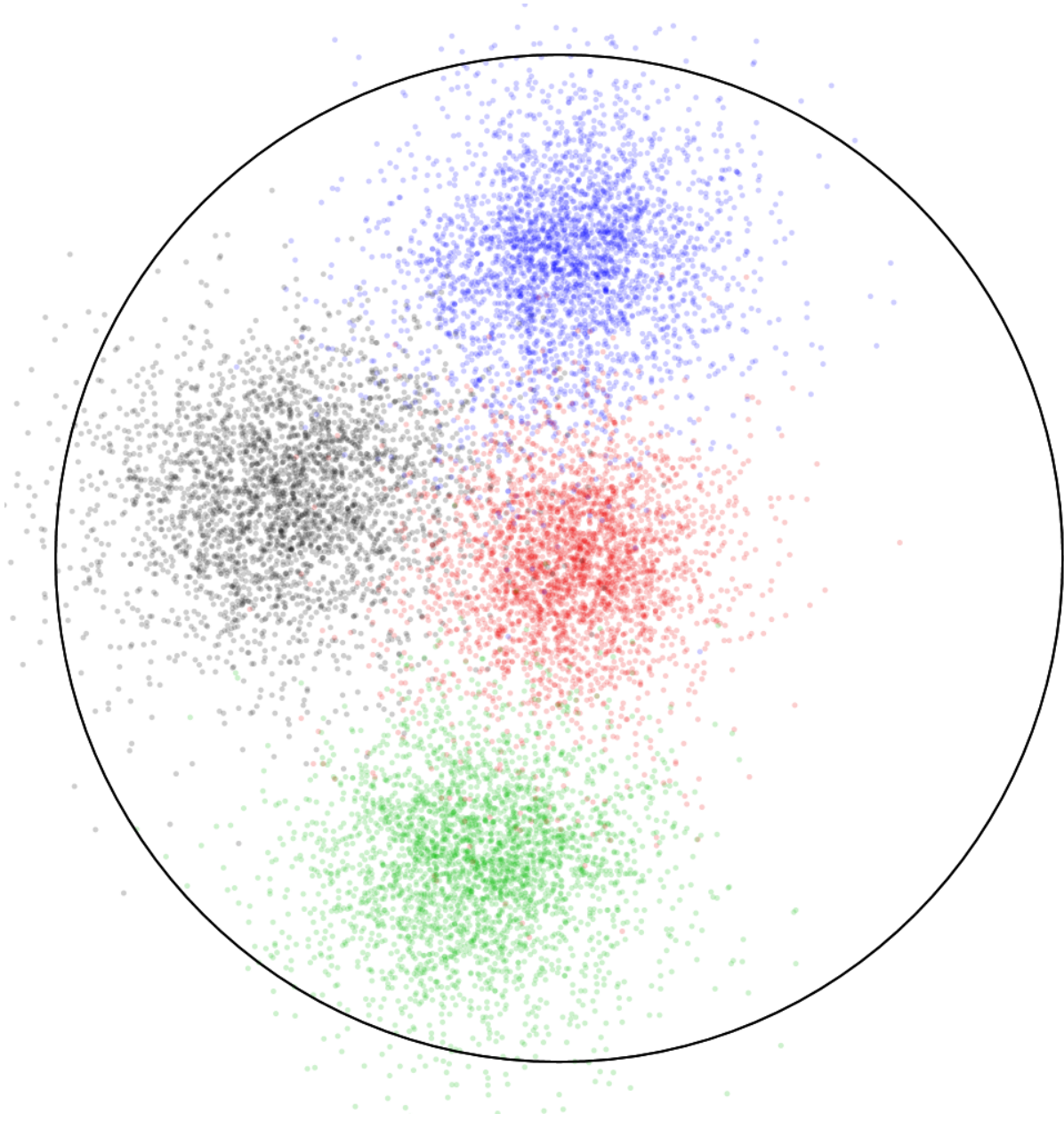
All 12,000 Phase Two Dart Throws



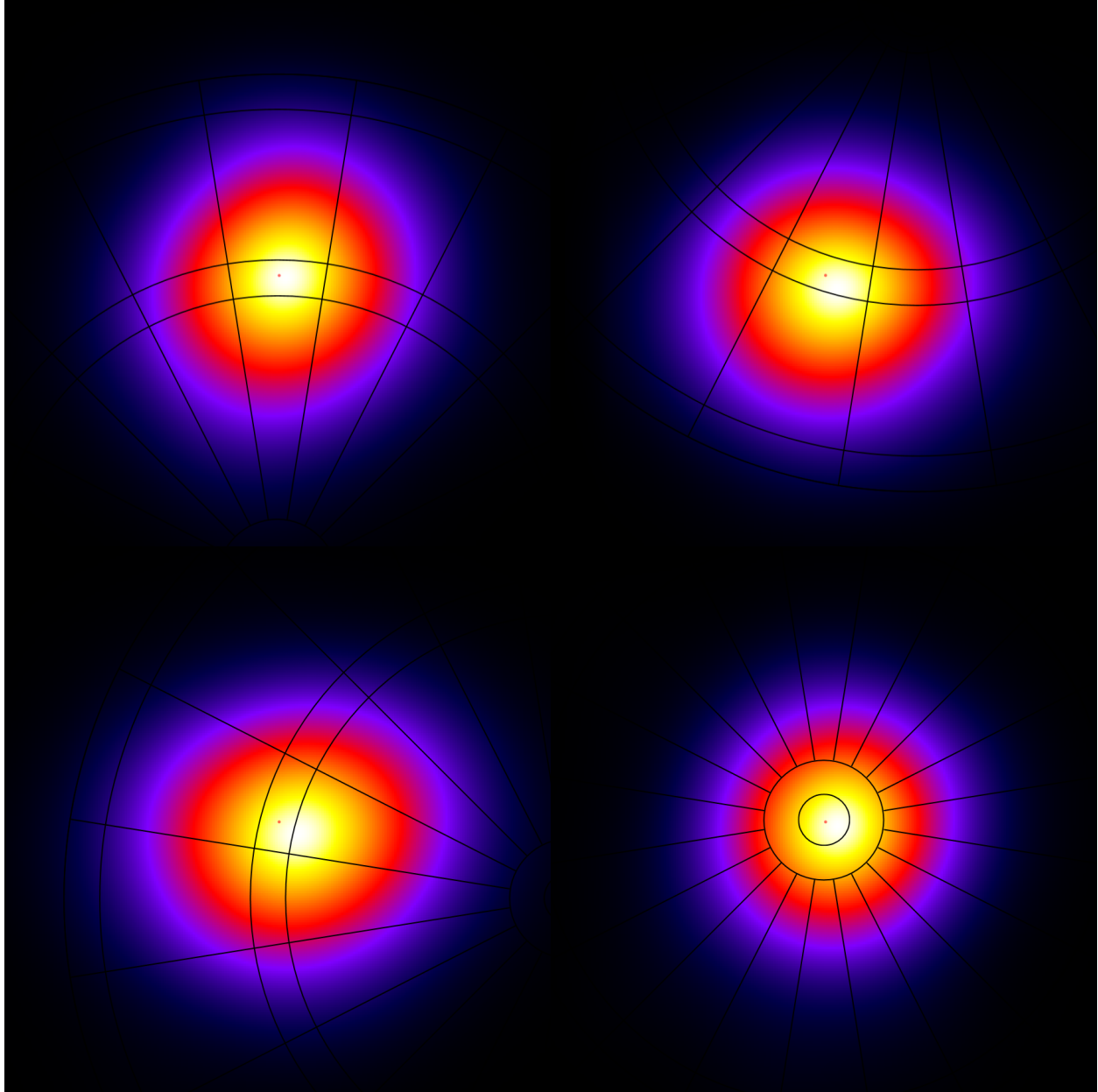
*- not actual proof.

Holes adjacent to the spider cover a larger area, which is why more darts land in them compared to holes in the middle of a sector. Of course my darts were not thrown at exactly the center of the holes - the dartboard did its best to make them end up in one, anyway.

I can make the throws look more like throws and less like my dartboard by randomly assigning a point for each dart that is covered by the area of the hole it landed in. That would look something like this:



If it looks like I didn't throw the same for each of the four targets, that is because I didn't. Here are the pdfs generated by my function for each of the targets:



Unscientific Conclusion About My Dart Throwing That Might Not Apply To Anyone Else
#2 - *I don't throw the same way at every target on the dartboard.*

Here is a table that compares how well I threw at each target:

Target	Avg	Pred	$x\sigma$	$v\sigma$
T20	50	47.7	28.3	29.2
T19	50.4	49.7	29.1	26.9
T14	46.2	46.9	31.9	28.1
Bull	49.6	49.5	25.6	27.2
Avg	49.1	48.5	28.7	27.9

The Pred column shows what the pdfs predict my scoring average should be for each target. Next, I will make an unscientific conclusion about my dart throwing that might not apply to anyone else (now abbreviated UCAMDTTMNATE) for each target:

UCAMDTTMNATE #3 - *I hit more triple 20s than I should.*

In 1000 turns/3000 throws aiming at the T20, I scored 2.3 points higher than the pdf generated from the throws thought I should. The difference is entirely explained by the number of triple 20s hit, which was higher than predicted across the board:

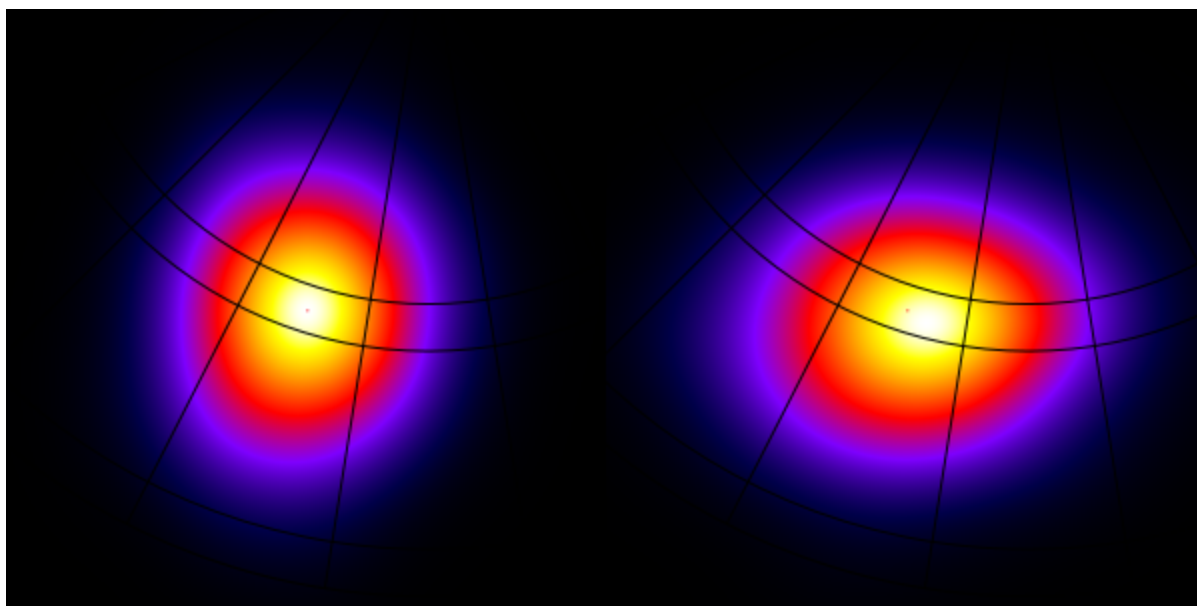
Dist	T20's	Pred
7.75	66	53
6.33	98	84
5.25	119	107
All	283	244

Math is telling me that 283+ triples should happen less than 1% of the time if the pdf is correct, so I will say that I might actually be better at hitting triple 20s than my program thinks I am.

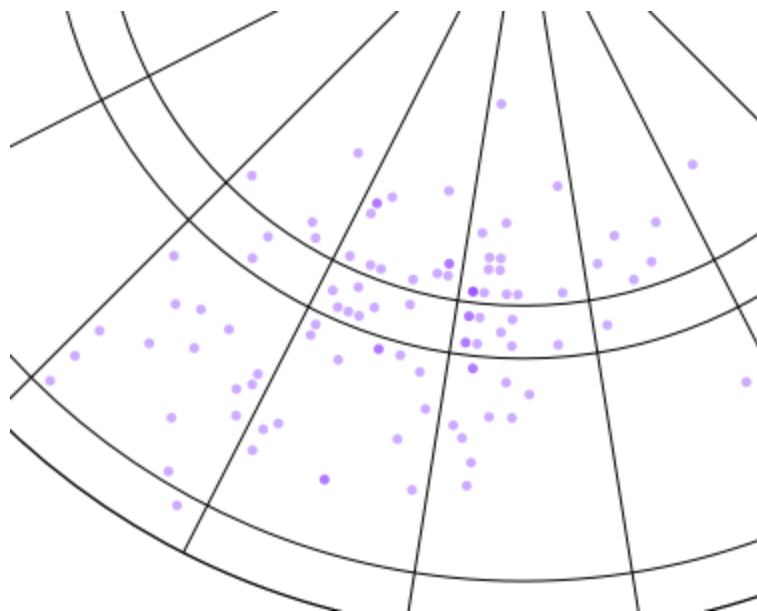
UCAMDTTMNATE #4 - *It's not that easy to keep your eyes on the target while throwing darts.*

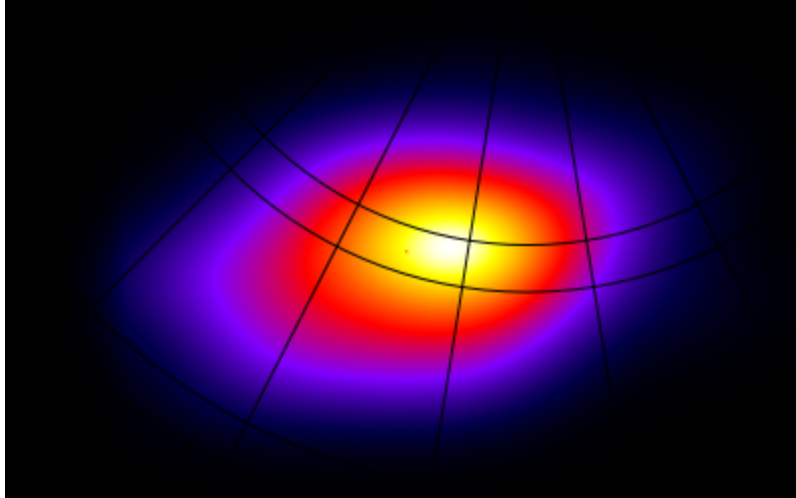
Although the T19 wins both the actual and predicted score contest, my throwing was less than ideal, which I attribute largely to letting misses to the "3" get into my head.

This was most evident in my throwing from 5.25 feet, the pdf of which looks quite a bit different than my phase 1 throws at the T19 from the same distance:



This was not a desired change. Some sessions were worse than others - the following shows perhaps my most frustrating session:





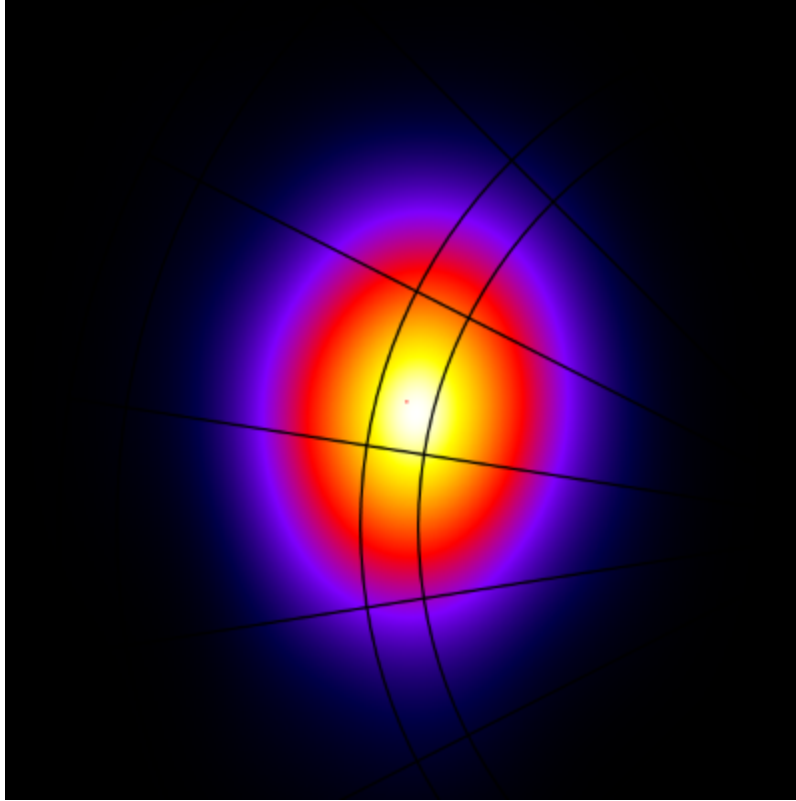
The thought going through my head while I was throwing the 100 darts in this session was “don’t hit another 3, don’t hit another 3.” The result was often that my eyes drifted to the 3, and I either hit a 3, or tried to make a last second adjustment and went too far left, hitting a 7 instead.

UCAMDTMNATE #5 - *The T14 scores poorly on the fun-o-meter; therefore I threw poorly when aiming at it.*

The fun-o-meter scores drop off significantly at a single throw average of 19-20 points because a single is no longer a good score for any target except the bull. The threshold is a lot lower when aiming at the T14, because 14 is less than both 19 and 20. Further, the funness of actually hitting the target is much lower since 3 times 14 is less than both 3 times 19 and 3 times 20. All in all, not a fun target to aim at.

When aiming at the T14, the goal isn’t just to try to hit 14s and T14s, it is to hit lots of triples in the area of the T14. Because there are no really low scores in the area, it is a low risk target, and in our fantasy dart leagues it was usually reserved for situations when whoever was throwing had a big lead.

Unfortunately, when my horizontal accuracy took a major downhill turn, triples stopped happening, which made the target even less fun and made me throw even worse as a result. This wasn’t always the case, as I threw far better and scored decently at the target during phase 1, when my throwing suited the triples goal well:

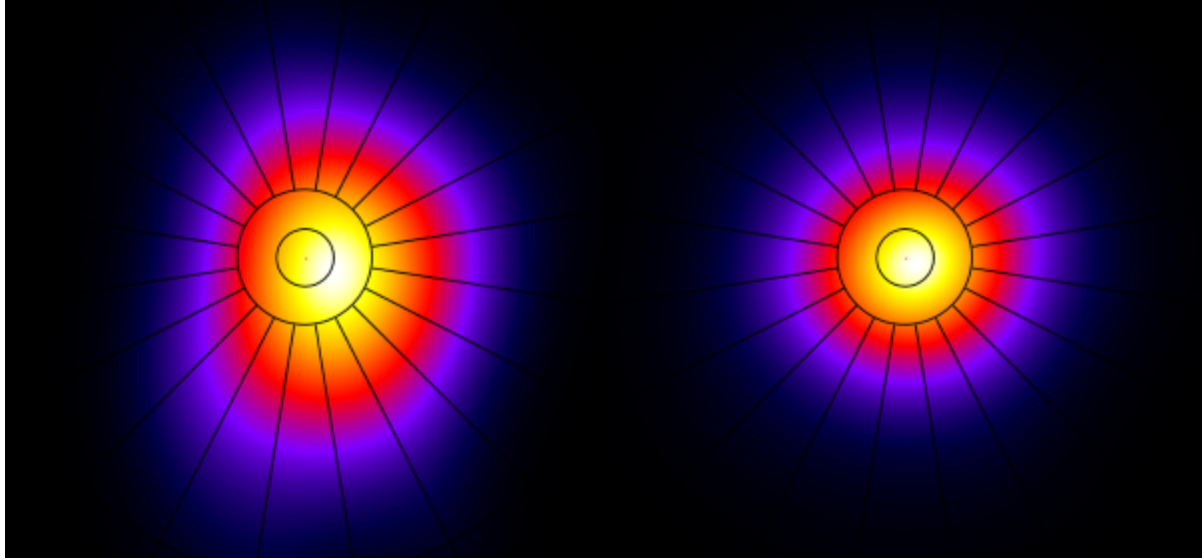


UCAMDTTMNATE #6 - *I throw better at targets that look like targets*

If a person who has never seen a dartboard see one without any of the numbers on it, where do you think they would aim? I'm guessing it would be the red circle at the center of the board.

My throwing was easily the most accurate when aiming at the bullseye. In combination with my dartboard's larger double and single bull zones compared to a regulation board, the bullseye becomes a viable scoring target (at least for me).

If not for one bad day throwing from the middle distance, the pdf for the bull would look a lot better. This shows the pdfs for these 500 throws and my other 2500 throws at the bull in phase 2:



I threw about as badly for the 5 sessions/500 throws from 6.3 feet than I did for my 1000 throws at the bull when I didn't cheat. As a result, I made by far the biggest improvement from 6.3 to 5.25 feet for the bull than the other three targets:

Target	Hits (6.3)	Hits (5.25)
T20	98	119
T19	95	112
T14	68	78
Bull	52	102

Phase 3

For phase 3, instead of throwing at a single target per session, I threw at all 20 numbers per each session. This was done in 120 throw sessions, with two consecutive turns (6 throws) at each target. The first 20 sessions I went in number sequence, starting with 1 for the 1st session, 2 for the 2nd session, 3 for the 3rd, and so on. The remaining 80 sessions the target order was chosen randomly by the dart tracker app.

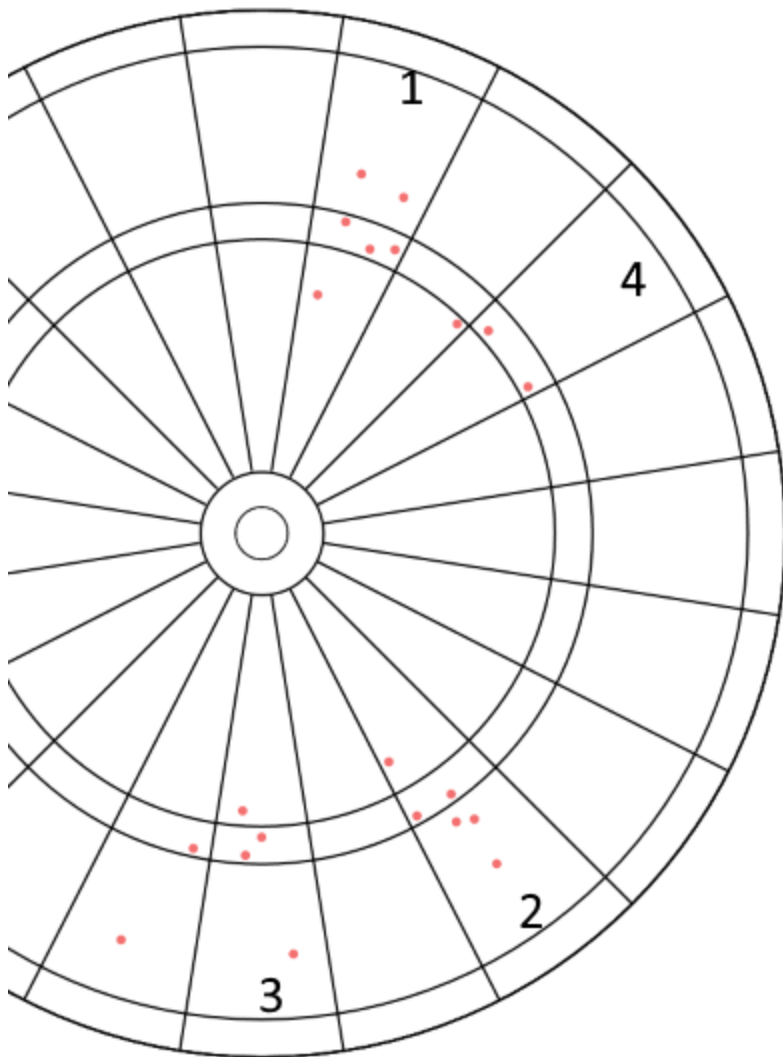
For each throw, I gave myself a score of 1 for hitting the single of the current target, 3 for the triple, 2 for the double, and 0 for everything else.

Half of my sessions were from 5.25 feet, with 25 of each from 6.33 and 4.25 feet (which replaced the longest distance from phase 2).

My first sessions were from the (new) middle distance. My expectations were smashed after just 21 throws, leading me to:

UCAMDTMNATE #7 - *There is no "warm-up" period when switching to a new target.*

Here are my first 21 throws of phase 3:



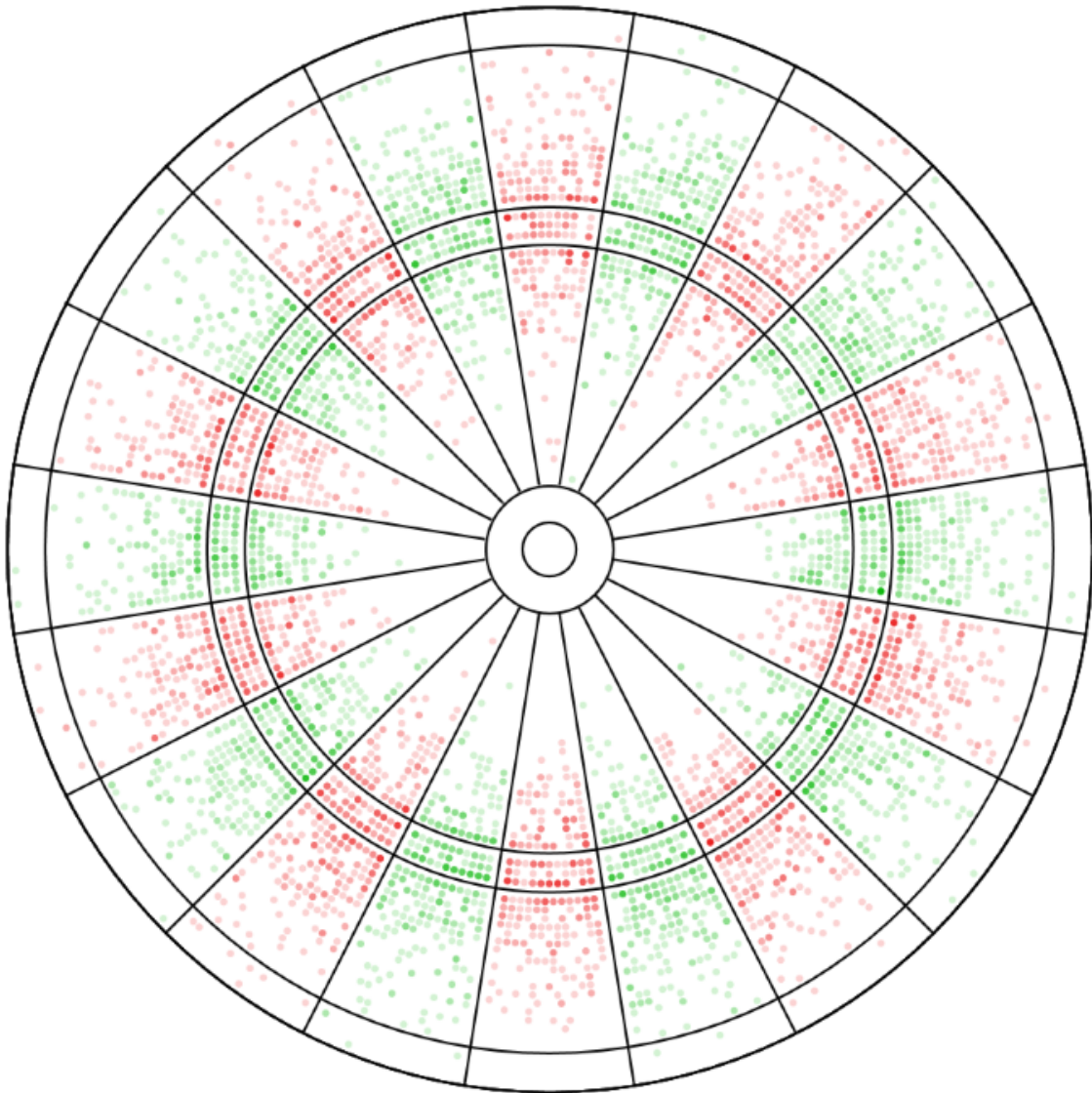
The 9 triples and 9 singles I hit were good for a **1.71** average per throw by phase 3's scoring. By comparison, I averaged about **0.74** aiming at the T19/T20 in phase 2 from this distance, with my best 21 consecutive throws averaging just **1.33** (a seemingly lucky 8 triples with just 4 singles).

My expectation going in was that I would throw relatively worse for phase 3 since I had to throw at 20 targets per session and not just one. I also thought that I would throw much better on the 2nd turn at a target than the 1st, which didn't happen.

Unfortunately, I never threw 21 consecutive darts from 5.25 feet as well as my first 21 darts during phase 3.

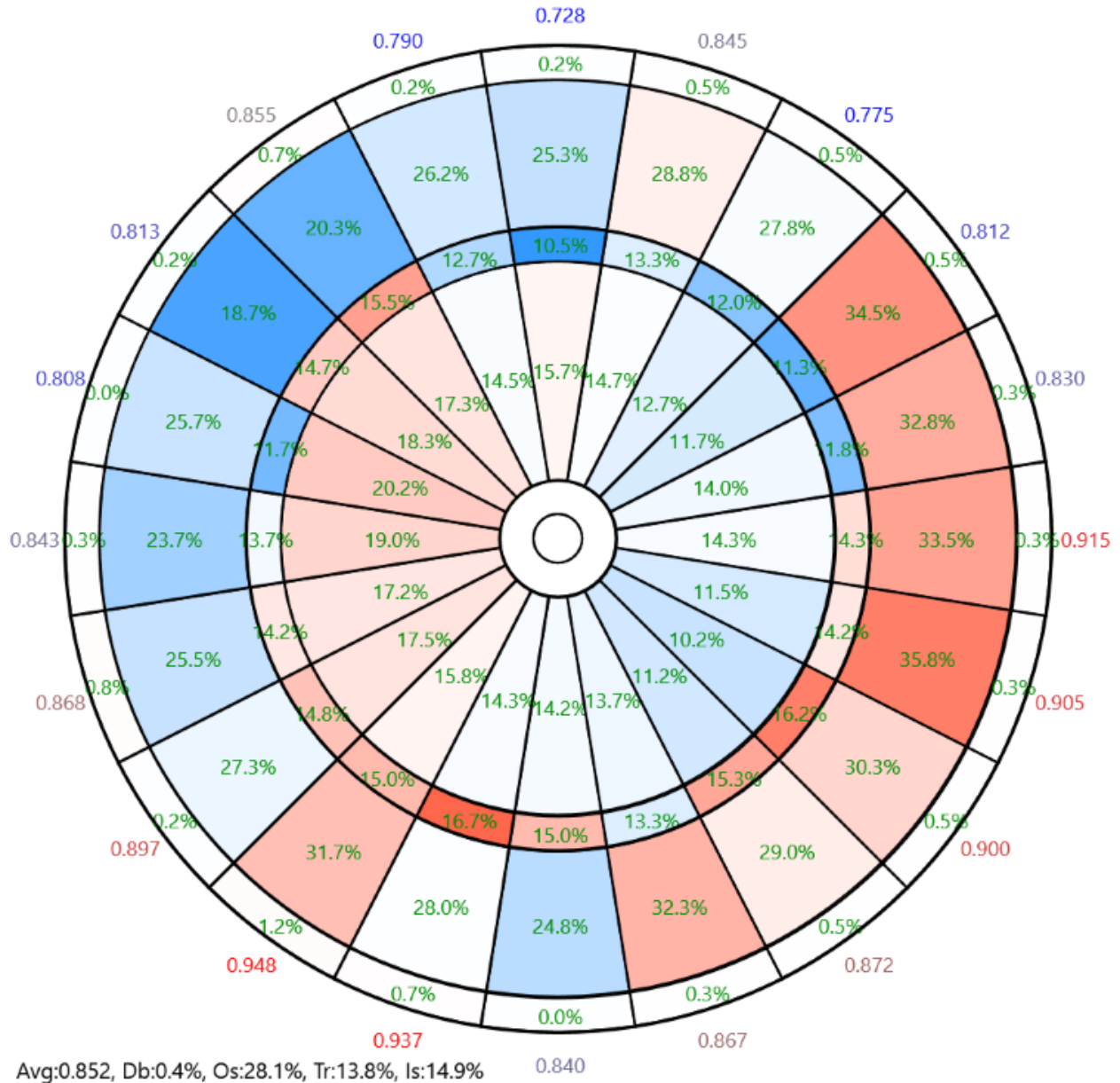
6,866 Throws

This shows the 6,866 out of 12,000 throws (600 at each target) I made in phase 3 that scored points:



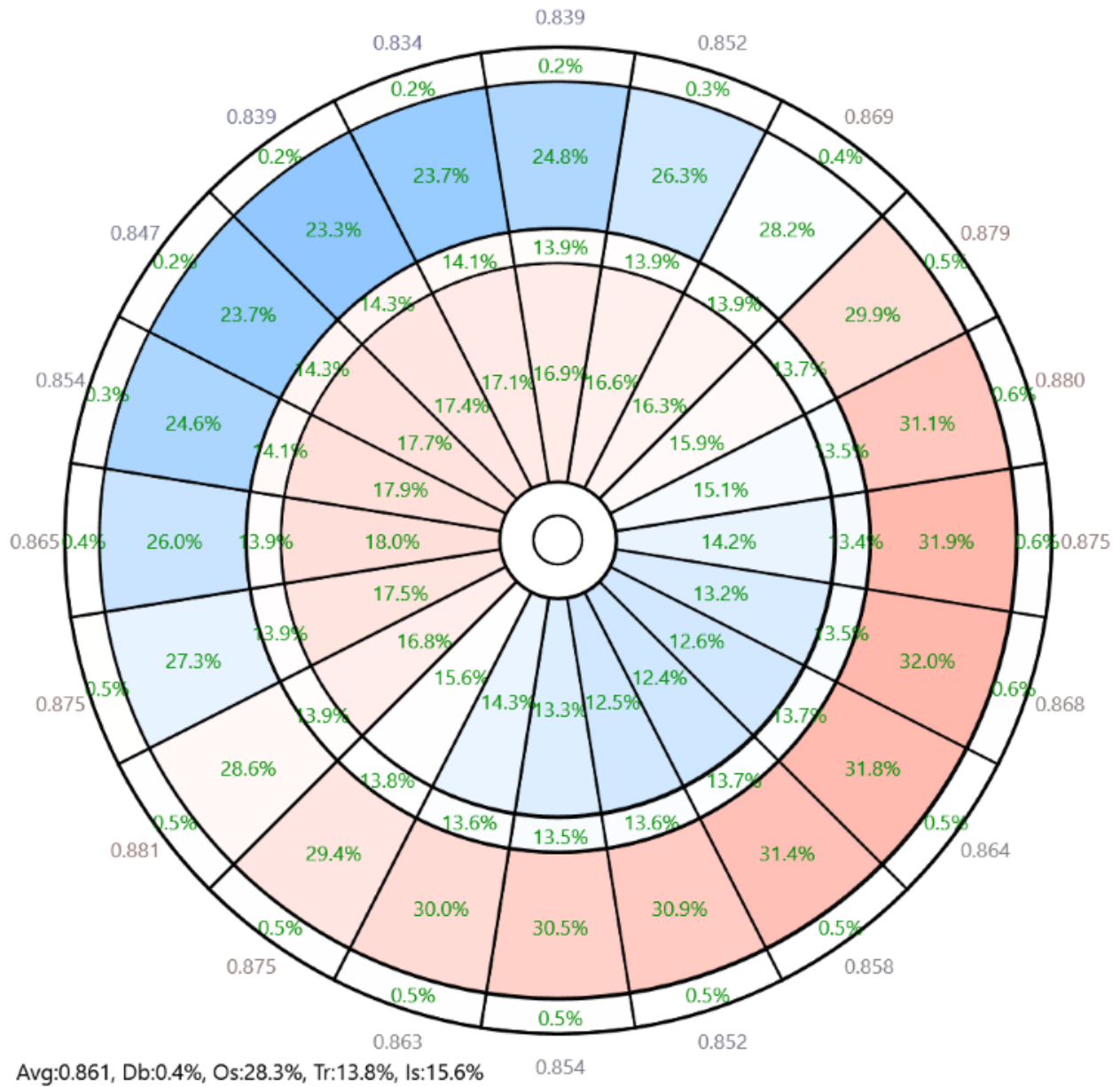
How I Scored For Each Sector

This chart (which I will refer to as a "dartboard wheel") breaks down my scoring for each sector:

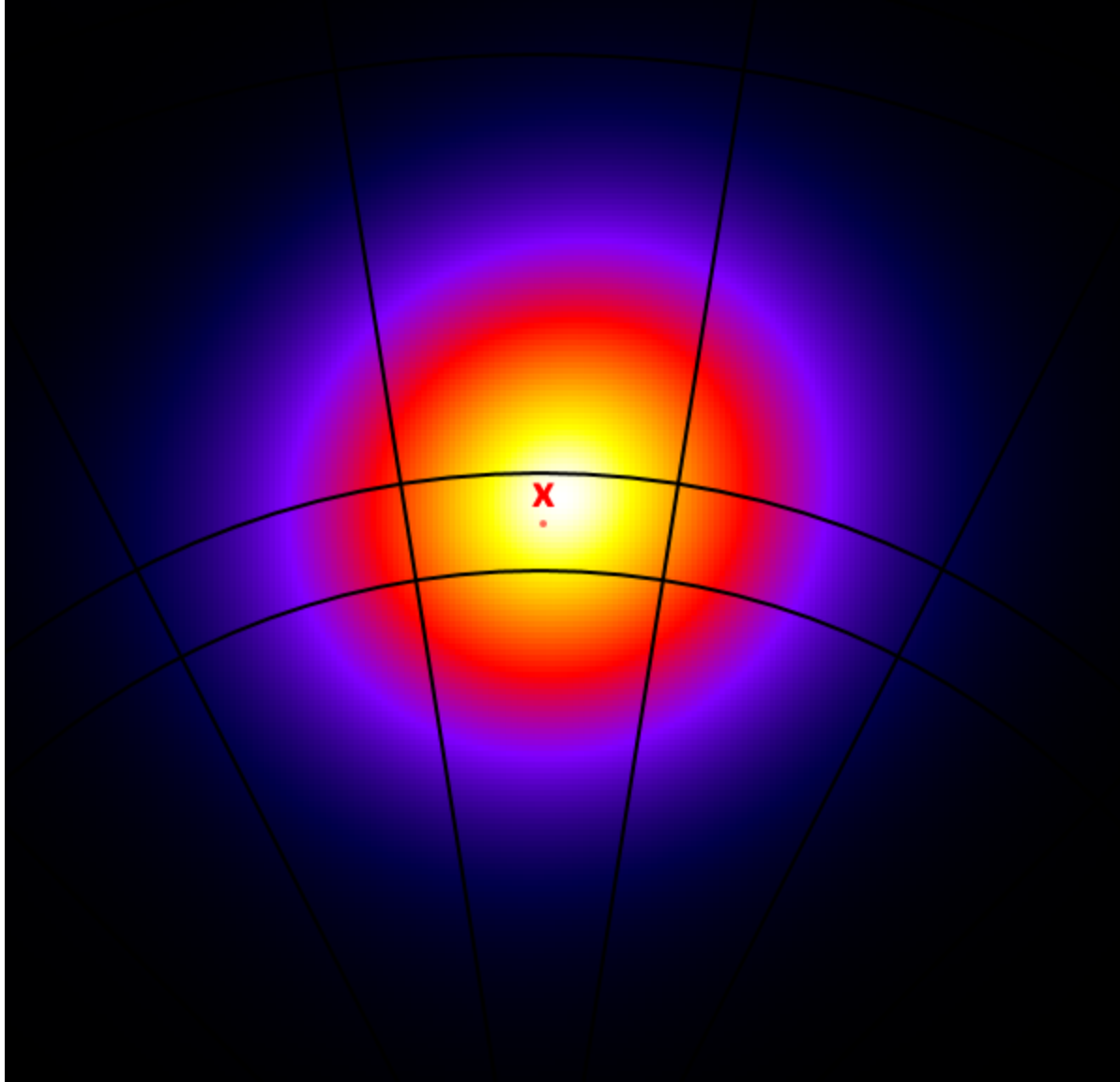


The coloring for each sector is scaled based on the points scored compared to my average for that sector, which is why there is a wider range of colors for the triple than the inner single sector despite a smaller range of percentages.

How My Generic Pdf Thinks I Should Score



This time, all 12,000 throws were used to create a generic pdf (image below), which was then applied to each of the 20 targets.



The red X is where I was apparently aiming based on my throws, while the small red circle is at the center of the triple 20. The X is there to illustrate a point - that my darts tend to land a bit below where I am aiming (also to the right, but the X wasn't necessary to show that).

My low/right throwing tendency explains quite a bit of the variance of the sectors in the above dartboard wheel, e.g., the outer sectors are redder low and to the right. The other significant factor is a bottom left to top right tilt in my pdf - which favors numbers that are angled in the same way.

Here is another table of numbers:

Target	Act	Pred	Gen	$x\sigma$	$v\sigma$
20	0.728	0.767	0.839	24.1	21.4
1	0.845	0.855	0.852	21.5	20.6
18	0.775	0.790	0.869	24.3	21.7
4	0.812	0.863	0.879	21.2	21.1
13	0.830	0.840	0.880	22.5	22.4
6	0.915	0.879	0.875	21.3	20.8
10	0.905	0.946	0.868	19.1	19.5
15	0.900	0.883	0.864	20.5	20.3
2	0.872	0.865	0.858	21.2	20.6
17	0.867	0.872	0.852	20.6	21.1
3	0.840	0.818	0.854	23.2	18.4
19	0.937	0.887	0.863	21.8	21.0
7	0.948	0.933	0.875	20.6	19.9
16	0.897	0.900	0.881	21.0	19.8
8	0.868	0.891	0.875	20.7	20.3
11	0.843	0.869	0.865	20.9	20.8
14	0.808	0.868	0.854	21.8	20.0
9	0.813	0.833	0.847	21.0	20.6
12	0.855	0.830	0.839	22.5	21.0
5	0.79	0.790	0.834	23.6	20.3
Avg	0.852	0.859	0.861	21.7	20.6
StDev	0.054	0.044	0.014	1.26	0.81

The targets start at the 20 and go clockwise around the board, which I felt was more useful visually than ordering them by number.

Time for some more conclusions:

UCAMDTTMNATE #8 - *Low is good, high is bad.*

All of the red cells in the first two columns in the above table are between 6 and 8. What isn't between 6 and 8? Any target in the upper half of the board. The only blue number on the low side is 3, which apparently I can only hit when I aim at the 19.

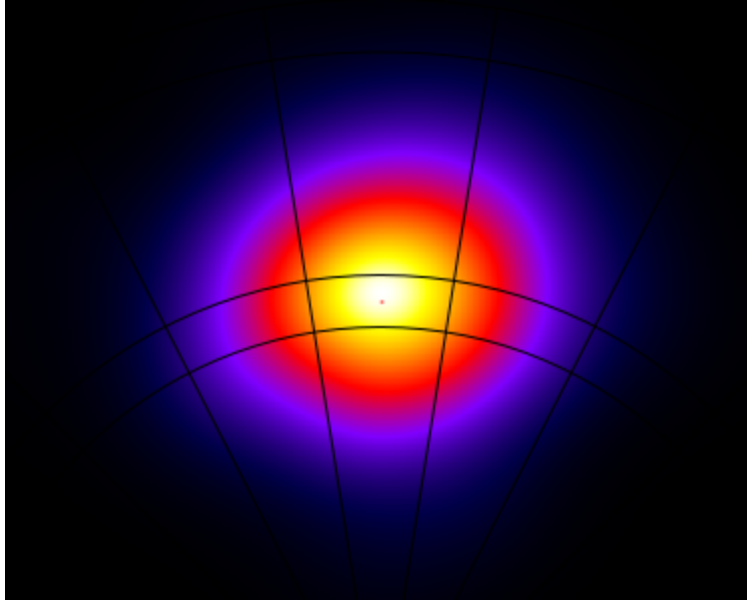
And while my generic pdf predicts I should score slightly better for low targets, it does not predict that my throwing would be more accurate, as shown by the following table:

	$x\sigma$	$v\sigma$
Bottom 5	21.5	20.2
Top 5	23.2	21

UCAMDTTMNATE #9 - *I will draw no conclusions about specific targets (except for the 3, which is evil) because 600 throws isn't enough.*

Nevertheless, here are a few of my best and worst targets and their pdfs:

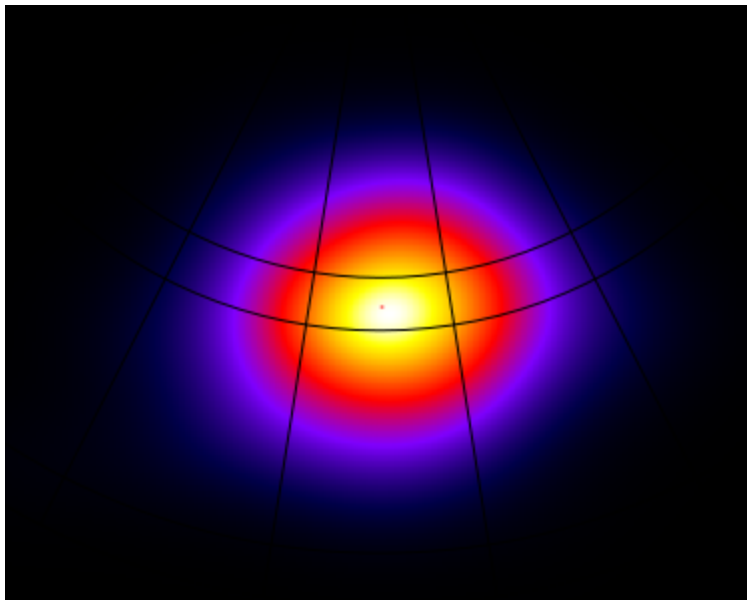
The Worst Scoring Target (Actual and Predicted): 20
(0.728, 0.767, 0.839, 24.1, 21.4)



The winner by a lot, and not a good one if you want to be good at regular darts. UCAMDTTMNATE #3 did not apply either, as the 63 triples I hit were 5 fewer than any other target, and 11 fewer than predicted by this pdf.

Widest Pdf: 3

(0.840, 0.818, 0.854, 23.2, 18.4)

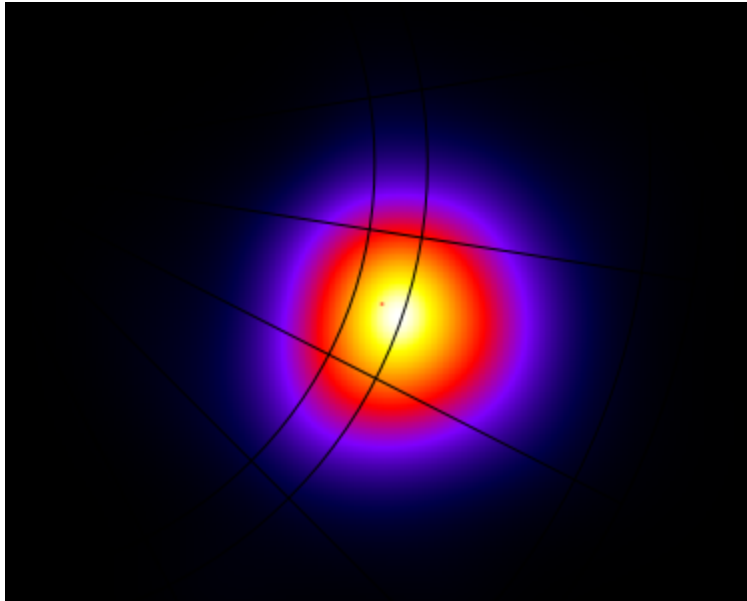


My vertical accuracy was the best aiming at the 3 ($y\sigma=18.4$ mm) by a lot, while my horizontal accuracy ($x\sigma = 23.2$ mm) was 4th worst. The 20 ranked 3rd in $x\sigma/y\sigma$ ratio. Unfortunately, these are the two targets where horizontal accuracy is most important. Strangely, the target with the

lowest $x\sigma/y\sigma$ ratio is 17, which is next to the 3. This is one of the reasons that the 3 (particularly the outer single) sticks out like a sore thumb in the first two dartboard wheels.

Best Predicted Score: 10

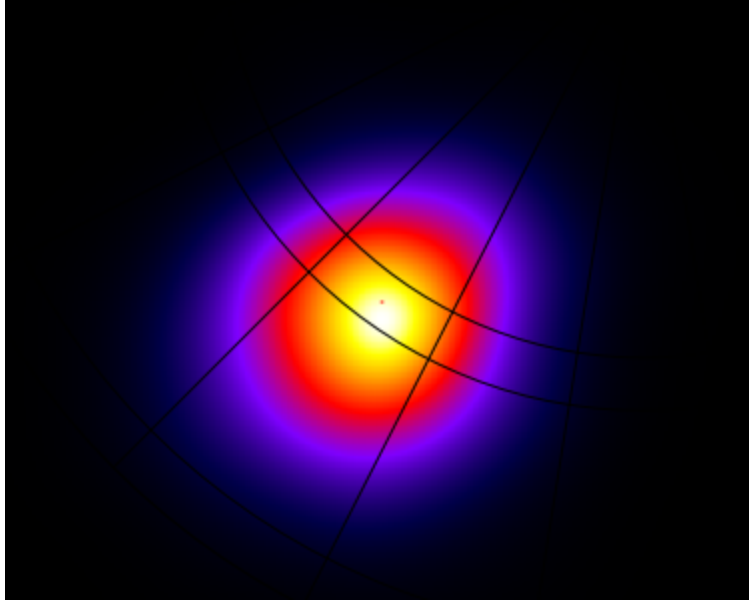
(0.905, 0.946, 0.868, 19.1, 19.5)



Notice that the pdf is a lot smaller than first two images, which is good because it means I didn't throw many darts in the black areas. The target ranked 1st in horizontal accuracy and 2nd in vertical accuracy.

Highest Scoring (Actual): 7

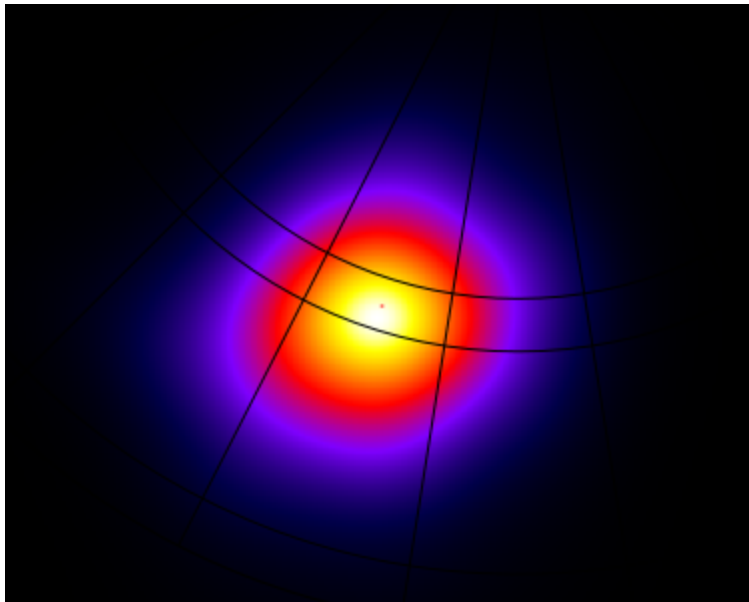
(0.948, 0.933, 0.875, 20.6, 19.9)



The 7 ranked in the top 5 in all 5 categories, and the shape of the pdf fits the target quite well, too.

Biggest Overachiever: 19

(0.937, 0.887, 0.863, 21.8, 21.0)



I scored about the same in phase 2 for the two targets that I also aimed at in phase 3, but they took wildly different paths this phase:

	19	20
Gen	0.863	0.839
Pred	0.887	0.767
Act	0.937	0.728
Gain	0.074	-0.111
Value*	4.2	-6.7
Avg (P2)	50.4	50.0
Avg (P3)**	62.2	51.9

*- value of the gain vs. my generic pdf. Per turn based on phase 2 scoring.

** - converted to phase 2 scoring. I threw from shorter distances in phase 3 than phase 2; some gain from phase 2 to 3 is expected.

Phase 2 vs. Phase 3

	5.25 (2)	5.25 (3)	6.5 (2)	6.5 (3)
$x\sigma$	25.8	22.9	26.4	24.5
$y\sigma$	20.9	19.9	27.4	25.7
Triples*	11.6%	12.8%	9.7%	9.1%
Singles*	38.7%	43.3%	37.6%	40.4%
P3 Avg*	0.738	0.821	0.671	0.683

*- based on T19 and T20 throws only (phase 2)

Although my throwing was never as good as the first 21 throws in phase 3, I was significantly more accurate than phase 2 from the common distances, though the difference is much smaller from 6.5 feet.

One theory I have is that the phase 3 scoring helped me focus on the target without thinking about the scores to the left and right. I also think I was getting quite bored by the end of phase 2 when I made all of my throws from 5.25 feet.

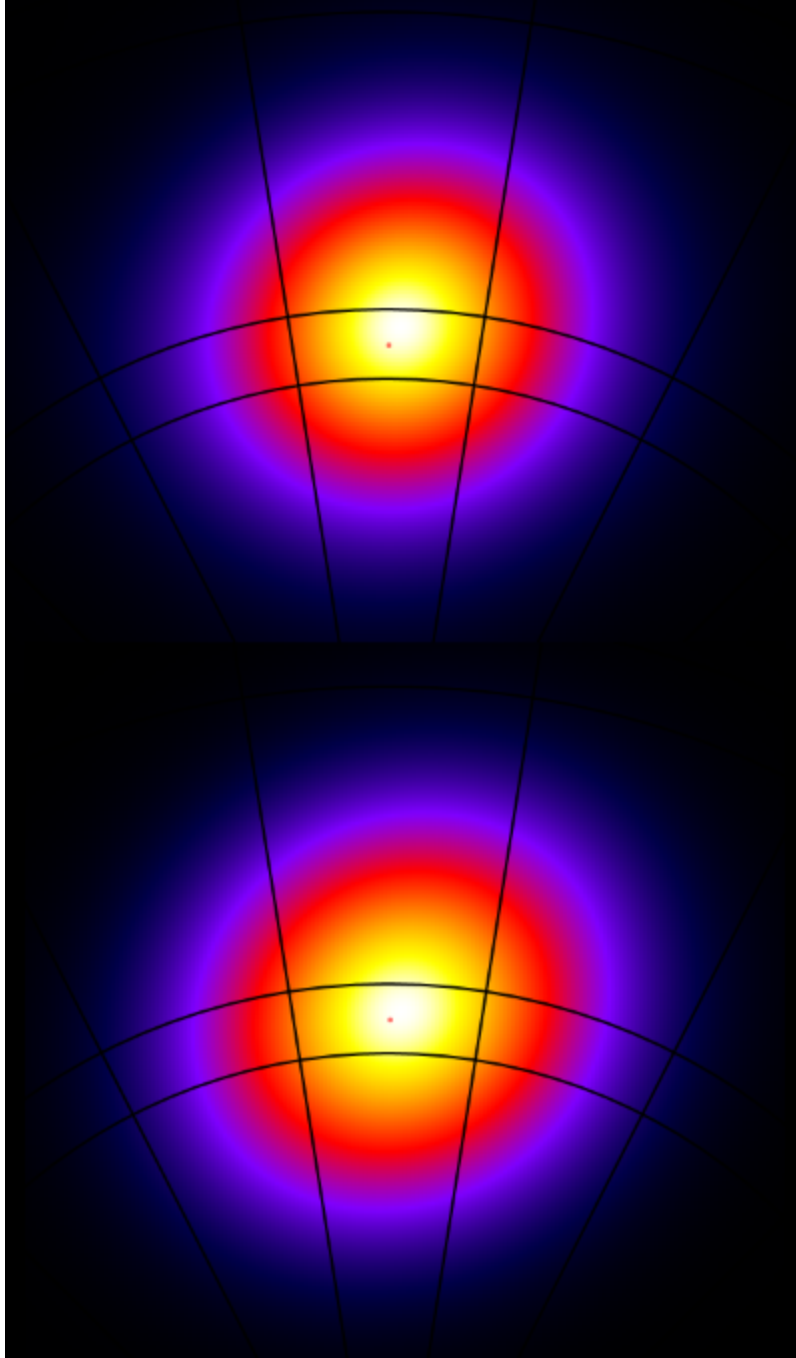
I have little doubt that if I reproduced all 12,000 throws for both phases, that the results would be significantly different (as they were from phase 1 to phase 2), so I will avoid making UCAMDTTMNATE #10.

Addressing the Questions at the End of Part 1

At the end of Part 1, I questioned whether several assumptions made apply to human dart players. Now I will rephrase each as a question and pass my indisputable judgment on each.

- 1. Does a human's dart throwing distribution follow a bivariate normal distribution (aka, Gaussian density function)?**

To answer this, I will show my generic pdf for phase 3 (top) along with a Gaussian density function (bottom) based on my phase 3 statistics:



My actual throwing doesn't quite fit the Gaussian model. For example, the rings are not shaped the same in the top image - the yellow ring is more circular than the blue/violet ring.

Verdict: Close, but not quite.

2. Are a player's throws independent of one another?

The evidence that my throwing varied from day to day (and to a lesser extent, from session to session) is convincing.

However, within a session I was unable to find sufficient evidence of "hot" or "cold" streaks (even though I certainly felt like I was hot or cold at times).

Verdict: Another loss for Part 1, but not in the way I expected.

3. Does a play throw exactly the same way at every target on the dartboard.?

Easily my biggest takeaway from Part 2 is that the target does in fact matter.

Verdict: Definitely Not

4. Does a player throw as well when frequently switching targets compared to throwing at a single target?

Conversely, phase 3 seemed to prove my theory wrong - though unlike #3, I question whether my own results apply to most other people.

Verdict: Yes (for me), Maybe for anyone else.

5. Can a player aim at specific points to the precision suggested by my work, the geek, and the statistician playing darts?

This is the hardest one to either prove or disprove. My subjective feeling is strongly negative, though I suspect target precision is a skill that varies considerably from player to player.

The best evidence I have that I couldn't aim at a target to Pigfoote's 0.5mm precision is that even when I knew my throws were generally low and to the right, I wasn't able to adjust my target to fix it.

Verdict: Probably not, but I can't prove it.

6. Should the average player aim at a target that produces a low score (e.g. the inner portion of the single 7) just because the math in Part 1 says they should?

While this wasn't intended to be a serious question, there is actually some proof that aiming at low scoring areas can affect one's accuracy. See: My throwing at the T14 from phase 2, for example.

Verdict: If you're not very good at darts, go ahead and throw at the target you like most. If some geek tells you to aim at the 7, you'll have a good response. In other words, no.

So, Can Any of Part 1 be Applied to a Human?

Despite all of these issues, the answer is “Yes”, but at a minimum, more information is required beyond a single pdf and a perfect solution will never happen.

The work in Part 1 is not dependent on a single probabilistic density function per player. In fact, my program uses this function only once for each player and each possible target on the dartboard, saving a table of dartboard sectors and percentages for each. For example, here is Pigfoote’s table for his optimal scoring target (Triple 19, $H=0.58$, $V = 0.81$):

Sector	Percent
OS 19	34.37%
IS 19	19.89%
T 19	11.98%
OS 7	9.35%
IS 7	7.07%
OS 3	5.75%
IS 3	4.83%
T 7	3.75%
T 3	2.43%
All Others	0.57%

Each “sweep” of the dartboard refers to these tables only, mapping the sectors to the appropriate values depending on the game, situation, and goal.

Thus, if I wanted to go through the process for myself, I could use the 20 PDFs per each number for targets in that number slice only - though I would probably be better off using some combination of the individual pdfs and an estimate based on my generic pdf and the location of the target. Since my throwing is subject to significant change over time, the process would have to be done continually to remain accurate, and I would need some mechanism to retrieve the

proper targets while playing for it to be useful (my own memory is quite probably not a viable option).

In summary, it would take a lot of effort for a relatively small gain, but it is possible.

The following invention would produce a much greater improvement in results, in my opinion:

Dartboard Goggles



The goggles would combine tracking and throwing prediction software with a target projection system: they would scan a dartboard, and through the goggles you would see only the target (a point, circle(s) or maybe an X) instead of the board. The challenge would be to make it look natural. For example, they would need motion sensors, and you would need to see your arm while throwing.

The goggles would provide three advantages:

1. Improved target selection. The software would predict the optimal target based on your predicted throwing, the variant of darts and the situation.
2. Improved target precision. A human darts player would have a hard time aiming at the exact optimal target point without a visual marker which the goggles would provide.
3. Improving throwing accuracy. During the phase 1 and 2, I threw much better at the bullseye than any other target (I didn't throw at the bull in phase 3).

Of course, if the goggles did exist they would almost surely be banned from professional darts. And rightly so - part of the fun of the game of Darts is that it is not *just* an accuracy competition, which is effectively what these goggles would turn the game into.

Revised Conclusion

When I finished working on this project, I felt like I had failed to achieve my goal of building a sophisticated statistical model that could make me a significantly better darts player.

While I initially hung onto the notion that the work from part 1 was still somewhat useful for human darts players, I no longer believe that statistical modeling on its own has much, if any, practical value.

I believe that human intuition and a knowledge of basic strategy can't be significantly improved upon without a way to visualize the target, such as dartboard goggles. While I was hoping for a mathematical solution to darts for a human, sometimes the solution isn't what we want it to be. For a human, the solution to darts is to see the dartboard more like a computer program. Then the mathematical solution from Part 1 with a statistical model can be applied to make the solution better for an individual human.