Spring 2019

Dice Mythbusters

C. Warren Campbell
Western Kentucky University, warren.campbell@wku.edu

William P. Dolan
Western Kentucky University and Carol Martin Gatton Academy of Mathematics and Science, william.dolan048@topper.wku.edu

Dice Mythbusters

C. Warren Campbell
Kenneth E. and Irene S. Hall Professor of Civil Engineering
School of Engineering and Applied Sciences
Western Kentucky University
warren.campbell@wku.edu

William P. Dolan
Carol Martin Gatton Academy of Mathematics and Science
Western Kentucky University

Dice Mythbusters

Abstract:

All dice are unfair because they cannot be manufactured with absolute precision. However, some dice are more unfair than others. Each year hundreds of millions of dice are sold worldwide. Dice commonly used in role playing games are 4-sided (D4), 6-sided (D6), 8-sided (D8), 10-sided (D10), 12-sided (D12), and 20-sided (D20). Most of the plastic dice are manufactured using plastic mold injection and rock tumbler methods. This method can result in dimensional inaccuracies in the dice and sometimes density inhomogeneities. In 3000-roll tests of fourteen D20 dice only three tested fair. Eight of the fourteen were plastic dice and only two tested fair. The remaining six were metal dice and of those only one tested fair so manufacturing issues are not restricted to plastic dice. In a running chi square test it was shown that for an unfair die the chi square statistic varies linearly with the number of rolls. The chi square statistic for an unfair die will always trend linearly to infinity while there is zero probability that a fair die will trend to infinity. The fair die chi square statistic will oscillate around the number of sides minus 1 (the degrees of freedom). An expression for the slope of the running chi square statistic was derived for unfair dice for both the situation of dimensional inhomogeneities and density inhomogeneities. Using simulations of billions of rolls of unfair dice, a probability distribution for the number of rolls beyond which the chi square statistic stays above the 95 percent critical value was obtained and is fitted well with a 2-parameter gamma distribution.

Background:

Hasbro sells about 50 million copies of the game Yahtzee each year. Each copy has 5 dice so that 250 million dice are produced for a single game. Thousands of people regularly play tabletop games including board games and role playing games (RPGs). The largest tabletop gaming conference in North America is Gen Con with more than 60,000 gamers attending each year. Video of gamers playing RPGs are live streamed regularly and recorded sessions can be viewed on Youtube and Twitch. World championships are held in many games with the winners receiving as much as $50,000.

There is much anecdotal evidence for unfair or cursed dice. *Big Bang Theory* and *Table Top* (Youtube) star Will Wheaton has the reputation for playing with cursed dice. There is even the somewhat humorous claim that dice he touches become cursed (Geek and Sundry [2018]). At least one person has observed 54 of Wheaton's dice rolls on Table Top (CritRoleStats [2016]). From these observations the chi square value is 42.3 compared to the critical value at the 95 % level of confidence of 30.14. The probability of a fair die achieving this value or greater is 0.0016. The chances that the die he used was fair is less than 1 in 600. Geek and Sundry (2016) also lists 15 of the craziest dice superstitions many of which relate to cursed dice. The point is that there are many examples of gamer's concerns about cursed dice on the Internet and links to some of these are given in Appendix B.

Gamers are concerned about the fairness of dice as is indicated by dozens of Youtube videos showing how to test dice. This concern is justified. Casino D6 dice are machined to tolerances of a few ten thousandths of an inch. The putty used for the pips is the same density as the body of the dice. However, many of the dice used by gamers which include D4s, D6s, D8s, D10s, D12s, and D20s are manufactured with plastic mold injection, painting of the full body of the dice, then removal of the paint using rock tumblers. This process rounds the edges of the dice and create random irregularities in dimensions. Also, plastic mold injection sometimes creates bubbles and dimensional inhomogeneities within the dice.

Dice can be unfair because of dimensional inhomogeneities (Haowei and Yang, 2013) involving either faces not the same size or inconsistent diameters. A shorter diameter results in 2 faces which will be rolled more often. Think of rolling a brick. It will land more often in a position with the center of gravity lower, and less often in an orientation with the center of gravity the highest. Reimer, et al. (2014) studied the effect of dimensional inhomogeneities in D6 dice and the effect of the method of rolling the dice. They found significant differences in rolling dice from a cup onto felt versus dropping them from a 1 meter height onto a steel plate. They also obtained a good fit to unfair dice probabilities using a Gibbs distribution. This effect on dice unfairness has been recognized at least since the writings of Isaac Newton circa 1665. Obreschkow (2006) performed dynamic simulations of cuboidal dice. Labby (2009) created an automated dice rolling machine that recreated Weldon's 1894 experiment of rolling 12 D6 dice 26,306 times. All of these are strong indications of interest in unfair dice.

There are numerous Internet citations concerning cursed or unfair dice. Some of these are provided in Appendix B. These just emphasize concerns about, and interest in unfair dice. The Awesome Dice Blog (2012) was interested enough to roll a Game Science and a Chessex D20 10,000 times each. Though they did not present the chi square statistic, the value for the Game Science die was 150 and for the Chessex was 261. The Game Science die had a burr where the die was removed from the sprue. The burr is on the 7 side and the opposite side was the 14 side. More than half of the value of the chi square statistic came from the low roll rate for 14. These numbers should be compared to the 95 percent critical chi square value of 30.14. So the null hypothesis of fairness for these two dice would be rejected.

There are many misconceptions concerning dice unfairness on the Internet. One Youtube video tests dice by doing 100 rolls of several D20s (Fisher [2015]) and then drew conclusions of relative unfairness of the dice. He claims that one die is more unfair than another because its chi square value after 100 rolls is higher than some of the other dice. Figure 1 shows the fallacy of drawing these types of conclusions. Each point in the figure is the chi square statistic for 100 rolls of a fair D20 die. The simulation of 100 rolls was done 10,000 times. To the right of the figure is the chi square distribution showing the peak occurring at $\chi^2 = 17.$ Because of the spread of the distribution, drawing conclusions based on single 100-roll tests of different dice is not statistically justified.

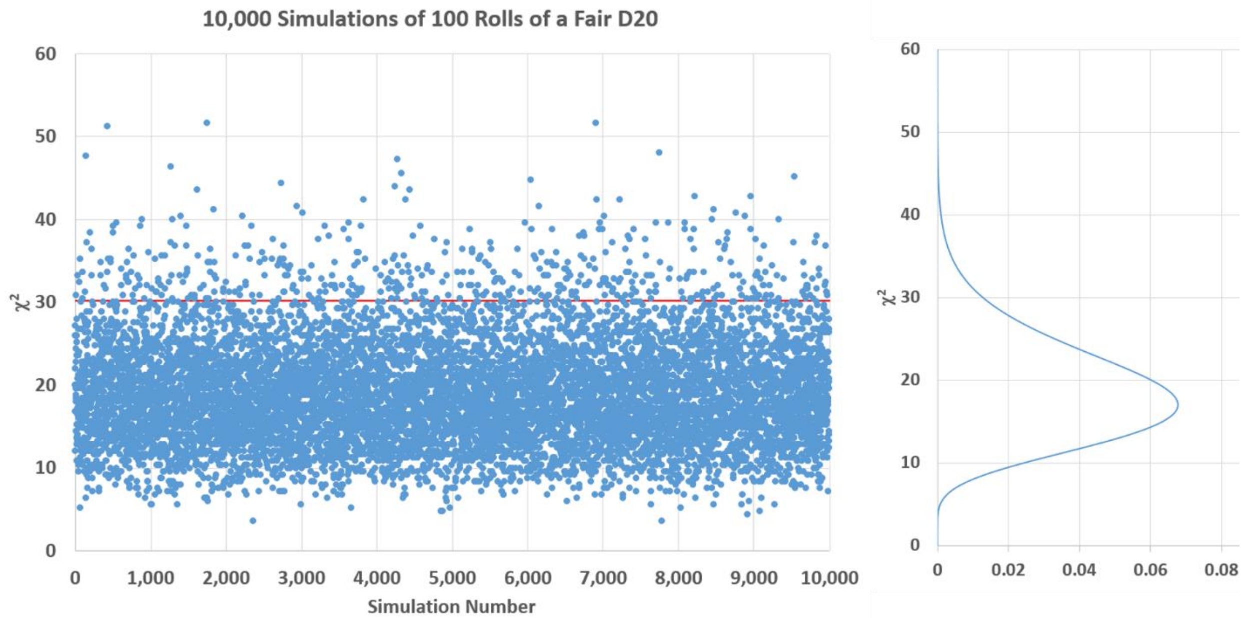

Figure 1. Ten thousand simulations of 100 roll tests of a fair D20 die.

The approach we used for the analysis is a running chi square statistic. That is, the statistic is calculated for one roll, for 2 rolls, and so on. When the statistic is plotted versus the number of rolls this gives an immediate and clear picture of the unfairness of the die.

All dice are unfair. Dice cannot be machined or molded perfectly, therefore all dice are unfair. The only question is how unfair they are. Figure 2 shows the running chi square statistic for simulation of 10,000 rolls of 3 dice of different levels of unfairness. All three dice are unfair, but the most fair of the dice does not cross the critical line in 10,000 rolls. The second die with medium unfairness crosses the critical line after 753 rolls and stays above it. The most unfair of the three crosses the critical line at 207 rolls and stays above it. Note that none of the three would have been detected as unfair after 100 rolls, even though the third die had a probability of 0.07 of rolling a 1 or a 20 and a probability of rolling 2 through 19 of 0.0478. This die has a 40 % deviation from 0.05 which is the probability of rolling a 1 or a 20 with a fair die. This is a very unfair die, but could not be detected as unfair in 100 rolls using the chi square test. The most fair of the dice had a probability of rolling a 1 or a 20 of 0.055 or a 10 percent error. Extending the best fit line out it would not have crossed the critical chi square value until more than 20,000 rolls.

We physically roll tested 14 D20 dice. Of the 14, only 3 tested fair in 3000 rolls. Seventy-eight percent of the D20s tested unfair after 3000 rolls. Eight of the 14 dice were plastic and the other six were metal. Two of the plastic dice tested fair and both were manufactured by Game Science. Two other Game Science dice tested unfair. Five of the 6 metal dice tested unfair. One of the unfair dice was manufactured by Level Up Dice, one by an unknown manufacturer, one by Metallic Dice Games, one by Die Hard Dice and one by Gravity Dice. The one metal die testing fair was a Die Hard Dice Precision Aluminum die. We tested four Chessex D20s and all four tested unfair. All of the results of testing are provided in Appendix A. In each figure the horizontal red line indicates the value of the chi square statistic that a fair die would exceed only 5 % of the time or 1 time in 20. From these results we conclude that it is very difficult to manufacture fair plastic or metal D20 dice.

The running chi square statistic of the 20 plots provided in Appendix illustrate some interesting behavior. Looking at Figures A.15, A.16, and A.20 for more than 1000 rolls these dice look fair. Then the statistic begins to grow and cross the critical line. Figure A.20 is particularly interesting. From 600 to 1800 rolls, it looks like a fair die. This was surprising because it was signioficantly rounded in the rock tumbler process. Then suddenly just beyond 1800 rolls the statistic climbs sharply ending at 56.55 at 3000 rolls.

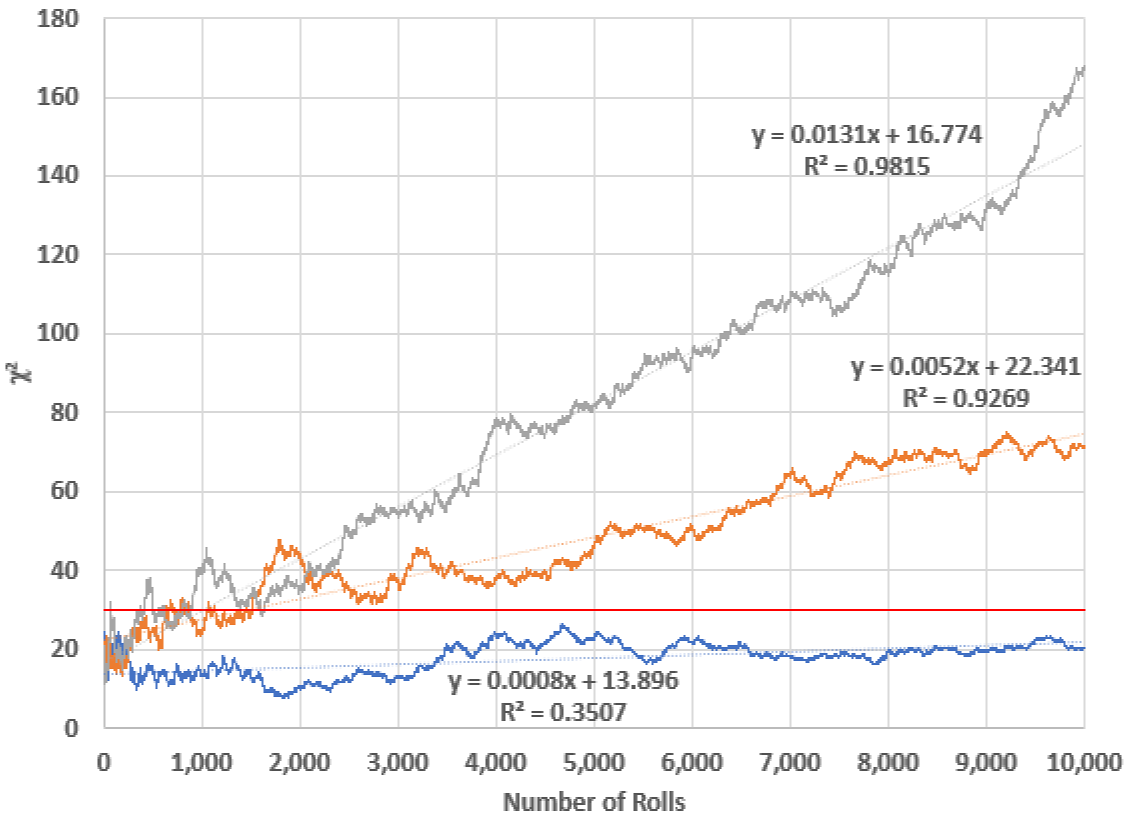

Figure 2. Running chi square statistic for dice with $p_1 = p_{20} = 0.055$, 0.060, and 0.070.

<u>Analysis:</u>

The chi square statistic for an s-sided die is given by Equation 1. Here O_i is the observed number of times side i was rolled and E_i is the number of times that side i would be expected to be rolled, that is n/s. A revised form of the Kolmogorov-Smirnov statistic was also checked but proved to be less sensitive to dice unfairness. For the purposes of this paper, a running chi square statistic was used because unfair dice can be spotted immediately by plotting the statistic as a function of the number of rolls of the die.

$$\chi^2 = \sum_{i=1}^{s} \frac{(O_i - E_i)^2}{E_i} \qquad (1)$$

The chi square statistic is only asymptotically chi square distributed for fair dice. For unfair dice rolls, the chi square statistic is not chi square distributed. The statistic is actually a discrete random variable when applied to dice. For example, if you roll a D3 die 3 times the only possible values of the chi square statistic are 0, 2, and 6. In general the maximum value of chi square obtained with a perfectly loaded die is $n \cdot (s-1)$ where n is the number of rolls and s is the number of sides of the die. The chi square distribution is a function of a real number on the interval $[0, \infty)$. That means there is a finite, nonzero probability for any finite real number interval, even those beyond $n \cdot (s-1)$. The chi square statistic is only asymptotically chi square distributed and only for nonexistent fair dice.

All dice are unfair because they cannot be made with absolute precision. Las Vegas requires casino dice to only be accurate to 0.0005 in. However, some dice are more unfair than others. The unfairness arises from 2 causes, 1) dimensional inaccuracies, and 2) density variations. Dice cannot be manufactured to zero tolerance. This inaccuracy causes a die to roll unfair. Think of a brick. It will land most often in the orientation with the lowest center of gravity, less often on the side, and most infrequently end up.

Density variation can be caused by bubbles in the dice. Plastic mold injection is the most common manufacturing method for gamer's polyhedral dice. When the injection mold is first turned on the mold is cold and the plastic hot. This can lead to bubbles in the dice (Figures 3 and 4). In Figure 3 the bubble is under the 6-8-18-17-14 corner. In the image you are looking at 5 copies of the same bubble refracted through the five sides adjoining the corner. This is a Game Science die sold at a discount because of the bubble. This bubble would cause an increase in rolls of 6, 8, 18, 17, and 14 and a decrease in the number of rolls for the opposite corner (3, 13, 15, 7, and 4).

Figure 3. Die with bubble under a corner.

Figure 4 shows a bubble under the 8 face. This would lead to an increase in the rolls of 8 and a decrease in the rolls of 13.

Figure 4. Bubble under the 8 face.

For a fair die, the probability $p=\frac{1}{s}$, p is the expected probability of rolling a particular number where s is the number of sides of the die. If there is a bubble under the center of one face then the actual probability of rolling that face is $p+\delta$ where δ is the deviation from the fair probability p. Similarly, the probability of the opposite face is reduced correspondingly to $p-\delta$. For this case, applying the law of large numbers, the chi square statistic is given in the limit by Equation 2.

$$\lim_{n\to\infty}\chi^2=\frac{\left[n\cdot\left(p+\delta\right)-n\cdot p\right]^2}{n\cdot p}+\frac{\left[n\cdot\left(p-\delta\right)-n\cdot p\right]^2}{n\cdot p}=\frac{2\cdot\delta^2}{p}\cdot n=2\cdot s\cdot\delta^2\cdot n \qquad (2)$$

Figure 5 shows the approach of an unfair die to the real die face probabilities as n becomes large.

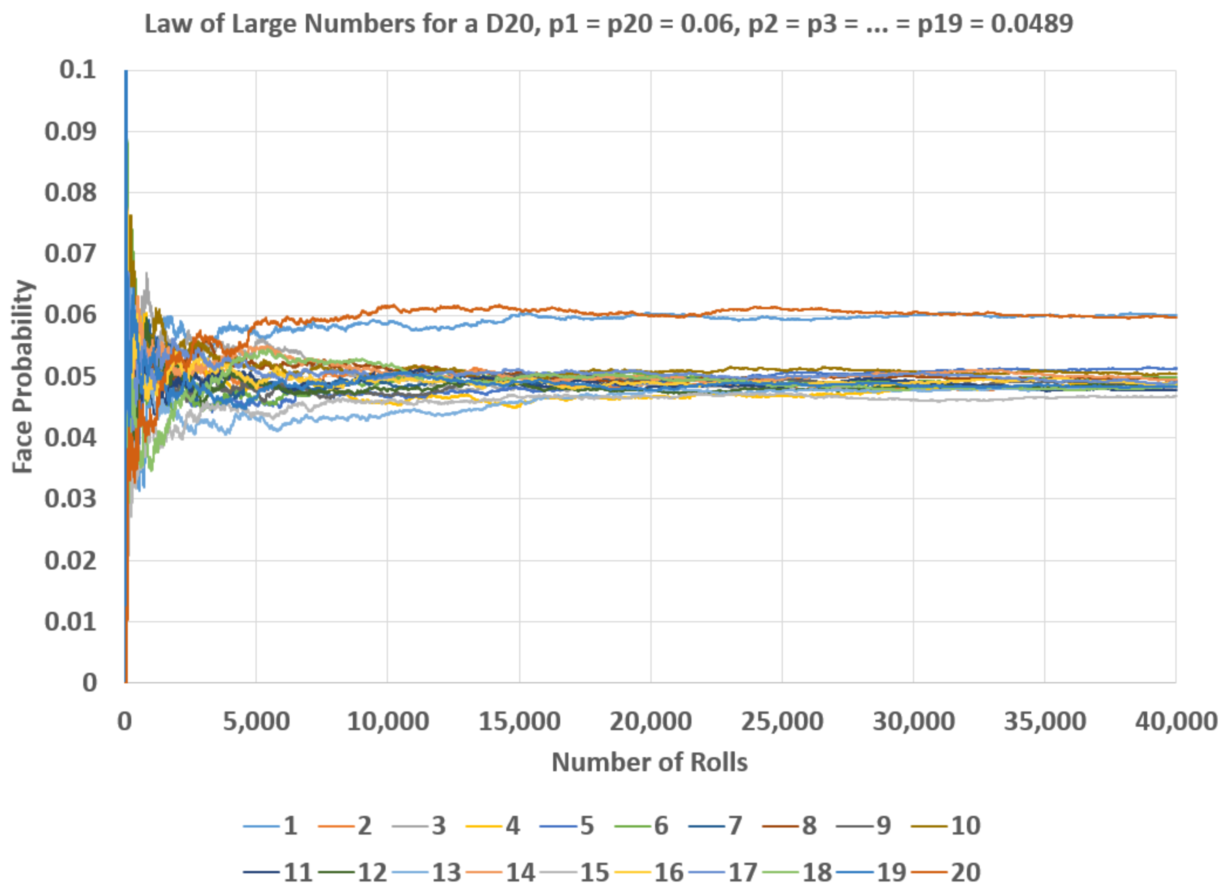

Figure 5. Law of large numbers for an unfair D20 die.

Similarly, for a bubble under a corner of a die, the chi square statistic is given by Equation 3.

$$\lim_{n\to\infty} \chi^2 = \frac{10\cdot\delta^2}{p}\cdot n = 10\cdot s\cdot\delta^2\cdot n \qquad (3)$$

In both cases of a corner bubble and a bubble under a face the value of chi square increases as the number of rolls n. The two equations seem to imply that the slope of the error versus number of rolls curve gets steeper for dice with larger numbers of sides. This is true but the effect is offset by the fact that the critical chi square value increases almost linearly with the number of degrees of freedom $(s-1)$.

The number of rolls on average to detect an unfair die is found by setting the right hand sides of Equations 2 and 3 equal to the critical chi square value and solving for n. As an example, consider a D20 die with a bubble in a corner and $\delta = 0.005$ which is 10 percent of p. It would require more than 6000 rolls on average for the chi square statistic to exceed the critical value.

This begs the question of how unfair does a die have to be to be detected as unfair in 3000 rolls? Setting Equation 3 equal to the critical chi square value and solve for δ. The result is 0.007 or about 14 percent of p.

For dimensional unfairness, if an increase δ occurs in two sides of the die, and the reduction in probability spread evenly over the other $s-2$ sides, the resulting variation of the chi square statistic is given by Equation 4.

$$\lim_{n\to\infty}\chi^2=\frac{2\cdot s}{(s-2)\cdot p}\cdot\delta^2\cdot n=\frac{2\cdot s^2}{s-2}\cdot\delta^2\cdot n \qquad (4)$$

So for a D20 die the term that is expressed as a fraction would be 44.4. Then for $\delta=0.005$ it would on average require more than 27,000 rolls to detect that the die was unfair. Similarly, for a die to be detected as unfair in 3000 rolls on average $\delta=0.015$ would be required. This would be 30 percent of p.

These analyses show that the chi square statistic increases linearly with n. This implies that all dice are unfair. No die can be made absolutely accurately with perfectly accurate dimensions and uniform density. Therefore, all real dice are unfair, but some are more unfair than others.

Data Collection:

Twenty different dice were hand rolled with a dice tower 3000 times each. These included dice from different manufacturers, made of different materials, namely plastic and metal, and different numbers of sides. Four D6 dice were tested. One was a casino die and another a precision backgammon die. These were both machined to high tolerances and have pips made of material with the same density as the body of the die. These both tested fair in 3000 rolls. We manufactured the other two D6s from aluminum. One was a distorted cuboid with the same dimensions as the cuboid tested by Reimer et al. (2014). As expected, the distorted cuboid tested unfair and the cube tested fair.

Six metal D20 dice from different manufacturers were rolled 3000 times. Five of the six tested unfair at the 95 percent level of confidence. The one that tested fair was a Diehard Precision Aluminum D20. The ones that tested unfair were one from Metallic Dice Games, a die made of heavy metal from an unknown manufacturer, a Level Up red cut die, a Die Hard Dice heavy metal (not aluminum) die, and a Gravity Dice aluminum die. Eight plastic D20 dice were tested. Of the eight only two tested fair after 3000 rolls. The two that tested fair were the Game Science dice, though two other Game Science dice tested unfair. None of the three Chessex dice tested fair.

A D7 die also tested unfair after 3000 rolls. One of the tools we use is a running chi square statistic which is demonstrated with the D7 die in Figure 7. The running chi square plotted in the figure is characteristic of an unfair die. Figure 7 shows the chi square statistic for a D20 fair to 3000 rolls. The horizontal line is the critical chi square value in both figures (95 percent confidence level). The running chi square plots for each die roll tested are provided in Appendix A.

Fickett (2016) also tested several D20 and D6 dice from various manufacturers. Though he did not give chi square values for the dice, he did provide enough information to calculate the chi square values for 23 of the dice he tested. The results are summarized in Table 1. The light pink indicates dice that tested unfair. From the table, 4 out of 5 Chessex D20 dice tested unfair, both Crystal Caste D20s tested highly unfair, the 3 Koplow D20s tested unfair, 4 of 5 Wiz Dice D20s tested unfair, 1 of 2 Crystal Caste D6s tested unfair, and 1 of 3 Koplow D6s tested unfair. Game Science dice are sold with a small, rough area on face 7 where the die is taken off the sprue. Game Science also sells a small file that can be used to remove the rough spot. With the rough spot, the black Game Science die tested unfair; with it removed it tested fair.

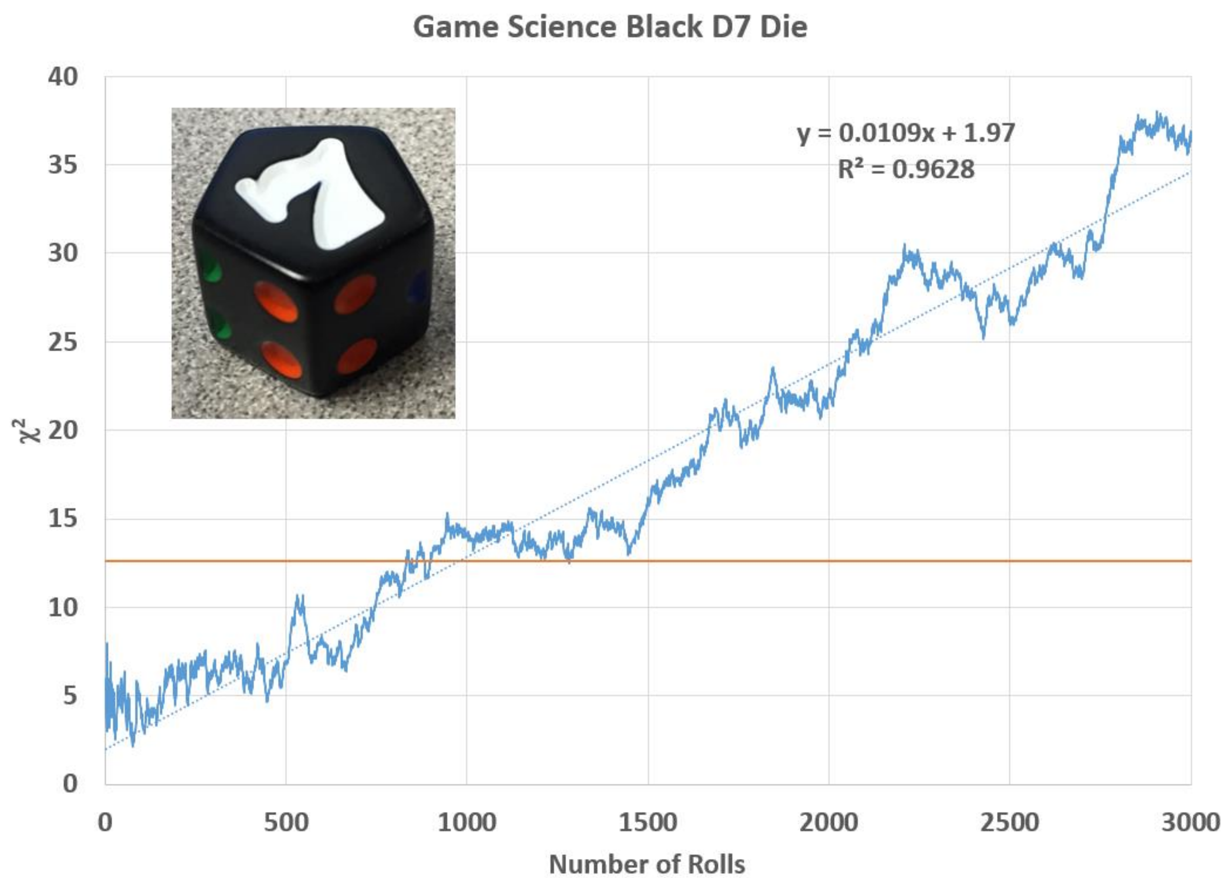

Figure 6. Running chi square for 3000 rolls of a D7 die.

From our data and Fickett's data we conclude that it is easier to manufacture a relatively fair D6 die than it is a D20.

Table 1. Fickett (2016) test data for several dice makers

n	Maker	Die	Description	χ^2	χ^2_{crit}
1	Chessex	D20	Yellow	84.03166	30.14353
2	Chessex	D20	Red/Orange	73.71509	30.14353
3	Chessex	D20	Purple/Gray	87.61713	30.14353
4	Chessex	D20	Gemini Copper Steel	58.36087	30.14353
5	Chessex	D20	Marbled Green	27.23641	30.14353
6	Crystal Caste	D20	Translucent Orange	262.972	30.14353
7	Crystal Caste	D20	Clear Black	331.2959	30.14353
8	Game Science	D20	White	43.61846	30.14353
9	Game Science	D20	Black before trim	46.33333	30.14353
10	Game Science	D20	Black after trim	14.26824	30.14353
11	Koplow	D20	Blue	60.34622	30.14353
12	Koplow	D20	Green 1	44.03166	30.14353
13	Koplow	D20	Green 2	55.20127	30.14353
14	Wiz Dice	D20	Yellow	78.38021	30.14353
15	Wiz Dice	D20	Solid Blue	29.72976	30.14353
16	Wiz Dice	D20	Translucent Blue	116.3542	30.14353
17	Wiz Dice	D20	Opaque Purple	155.3108	30.14353
18	Wiz Dice	D20	Translucent Blue	298.5174	30.14353
19	Crystal Caste	D6	Crystal Translucent Orange	1058.416	11.0705
20	Crystal Caste	D6	Clear Black	2.222777	11.0705
21	Koplow	D6	D6 a	10.52	11.0705
22	Koplow	D6	D6 b	7.688	11.0705
23	Koplow	D6	D6 c	12.96	11.0705

Figure 7. D20 die fair to 3000 rolls

Simulations:

We simulated billions of dice rolls to obtain the probability distribution of the number of rolls beyond which the chi square statistic stays above the 95 percent critical value. In each case we simulated 40,000 rolls and calculated the point beyond which the running chi square statistic exceeded the 95 percent critical value and stayed above it to the end of the 40,000 rolls. This was repeated 12,000 times for each level of unfairness. The probability distribution was obtained and was found to be skewed positively in every case. The simulated distribution was fitted well with a 2-parameter gamma distribution. Figure 8 shows the quality of fit for 3 different scenarios.

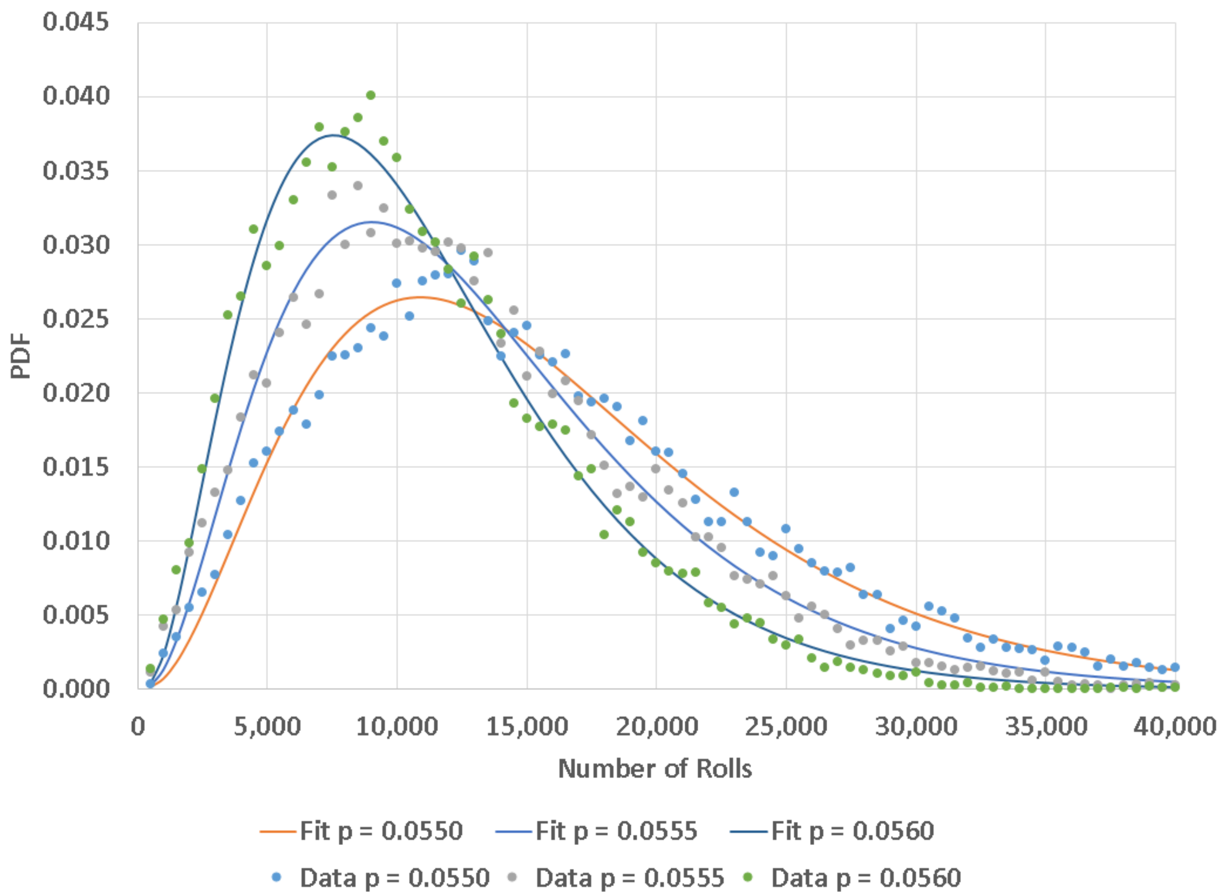

Figure 8. Simulated probability distribution and fit for a D20 die.

Figures 9 through 11 summarize the fitted probability distributions for D20 weighted dice. These are based on billions of dice rolls for different levels of unfairness. The probability of a fair die roll for each face is 0.05. The legend on the bottom of the chart shows the simulated probability fit. These charts correspond to the case when a bubble is just below one face so that the probability is greater than 0.05 for that face and is correspondingly reduced for the opposite face.

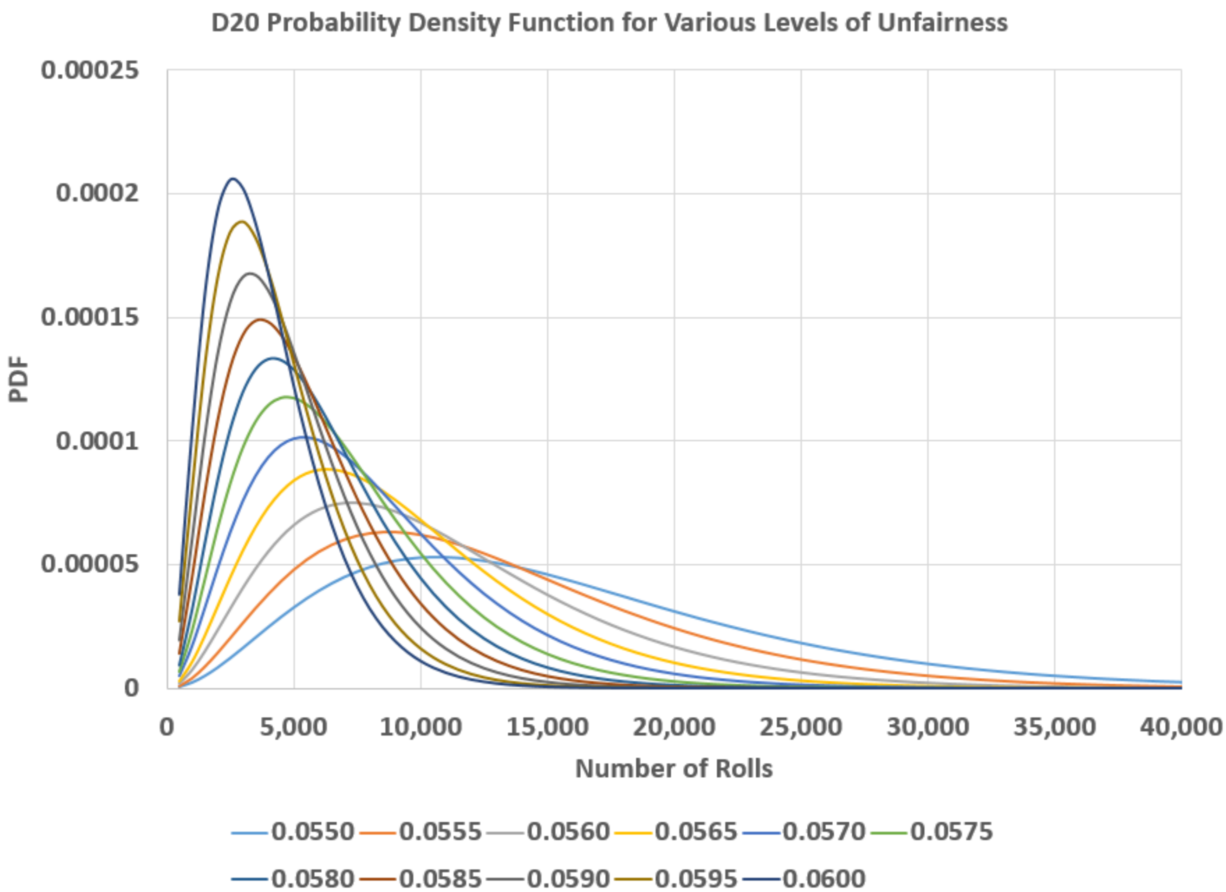

Figure 9. Fitted probability distributions characteristic of a D20 die with a shorter diameter.

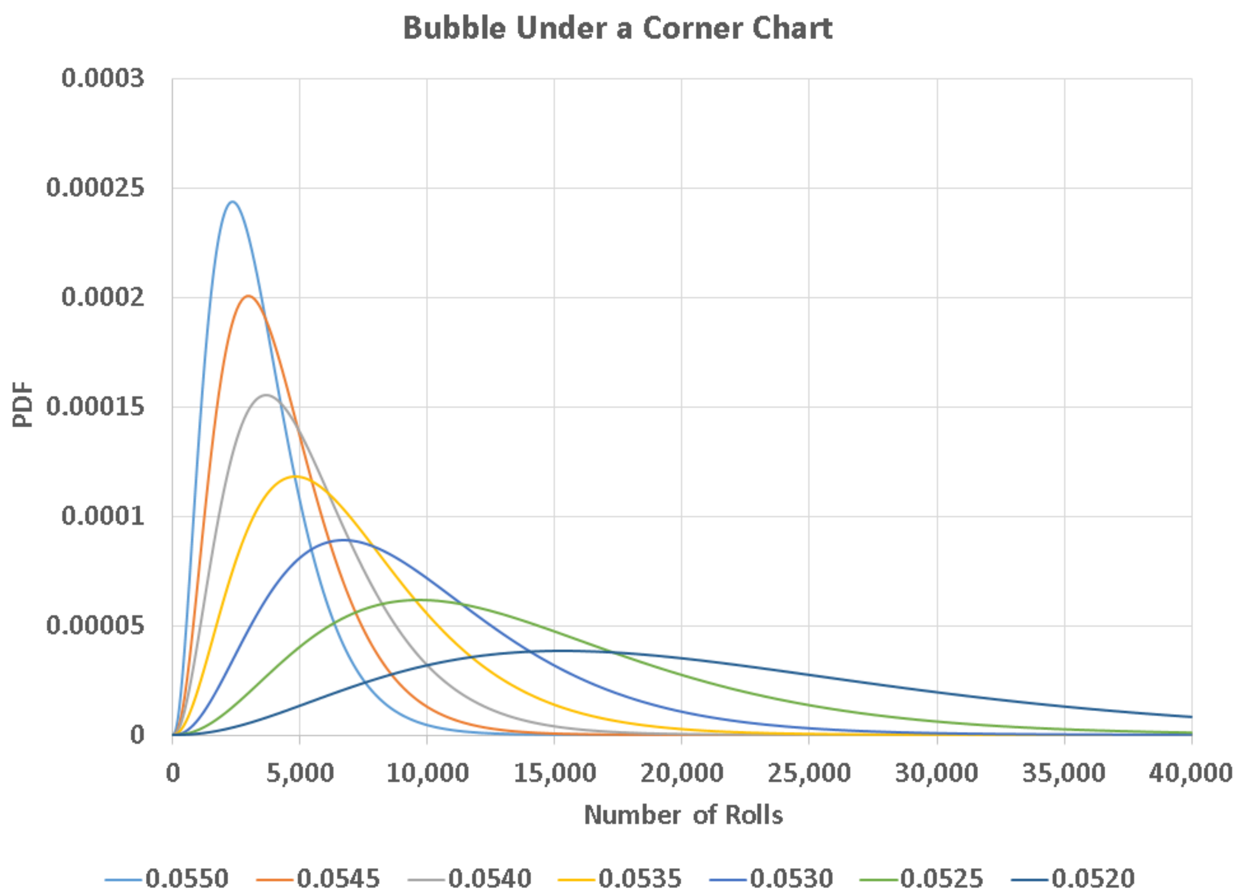

Figure 10. Probability distributions characteristic of a D20 die with a bubble under a corner

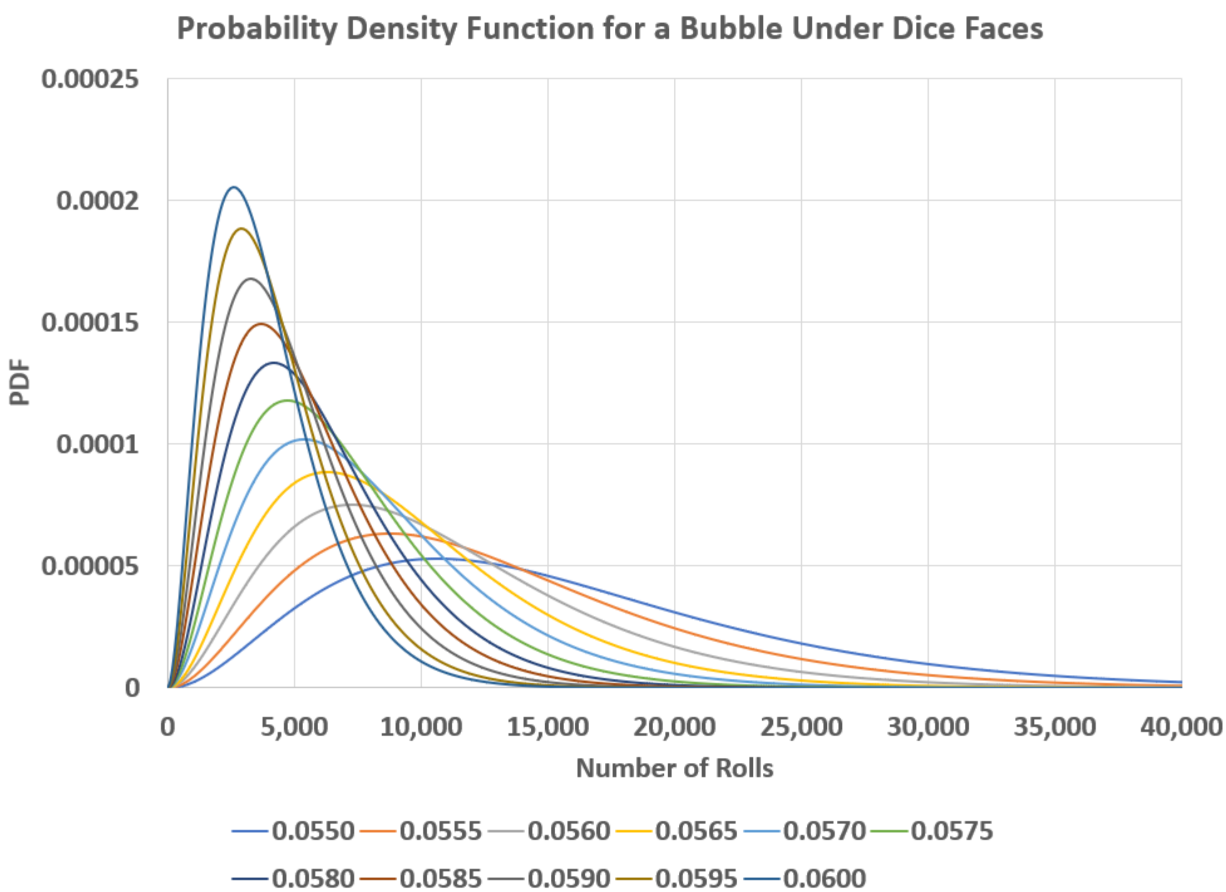

Figure 11. Probability distribution for bubble under a face of a D20 die

In a Youtube video, Fisher (2015) rolled five different D20s 100 times each and drew conclusions of unfairness based on the chi square value he obtained for each die. In fact the probability is quite high that a fair die will have a chi square value higher than an unfair die after 100 rolls. The probability distribution that this will occur is given by Equation 5 below. At $\Delta\chi^2 = 0$ the equation is undefined so you take the limit as $\Delta\chi^2 \to 0$ to get the value of the probability density there.

$$p\left(\Delta\chi^2, \nu\right) = \frac{2^{-\nu} \cdot \sqrt{\pi} \cdot |\Delta\chi^2|^{\frac{1}{2}\cdot(\nu-1)} \cdot BesselK\left[\frac{|\Delta\chi^2|}{2}, \frac{1}{2}\cdot(\nu-1)\right] \cdot \mathrm{Csc}\left(\frac{\pi\cdot\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}-1\right)\cdot\Gamma\left(\frac{\nu}{2}\right)^2} \quad (5)$$

The probability density functions described by equation 5 are plotted in Figure 12 for degrees of freedom $\nu = 5, 7, 9, 11, \text{and} 19$ which corresponds to that for D6, D8, D10, D12, and D20 dice. The x-axis in the figure is the difference in the fair die chi square value and the unfair die chi square value. This assumes a linear trend in the chi square statistic for an unfair die. Figure 13 shows the cumulative probability. A fair D20 die will have a 40 % probability of having a higher χ^2 value than an unfair die when the trend in the statistic for the unfair die gives a $|\Delta\chi^2| = 2.3$. From Equation 4 this corresponds to $p_1 = p_{20}$ greater than 0.07 or a 40 percent error in p for 100 rolls of a D20. For 3000 rolls it corresponds to 0.054 or an 8 % error. More rolls permit a significant improvement in the ability to detect unfair dice.

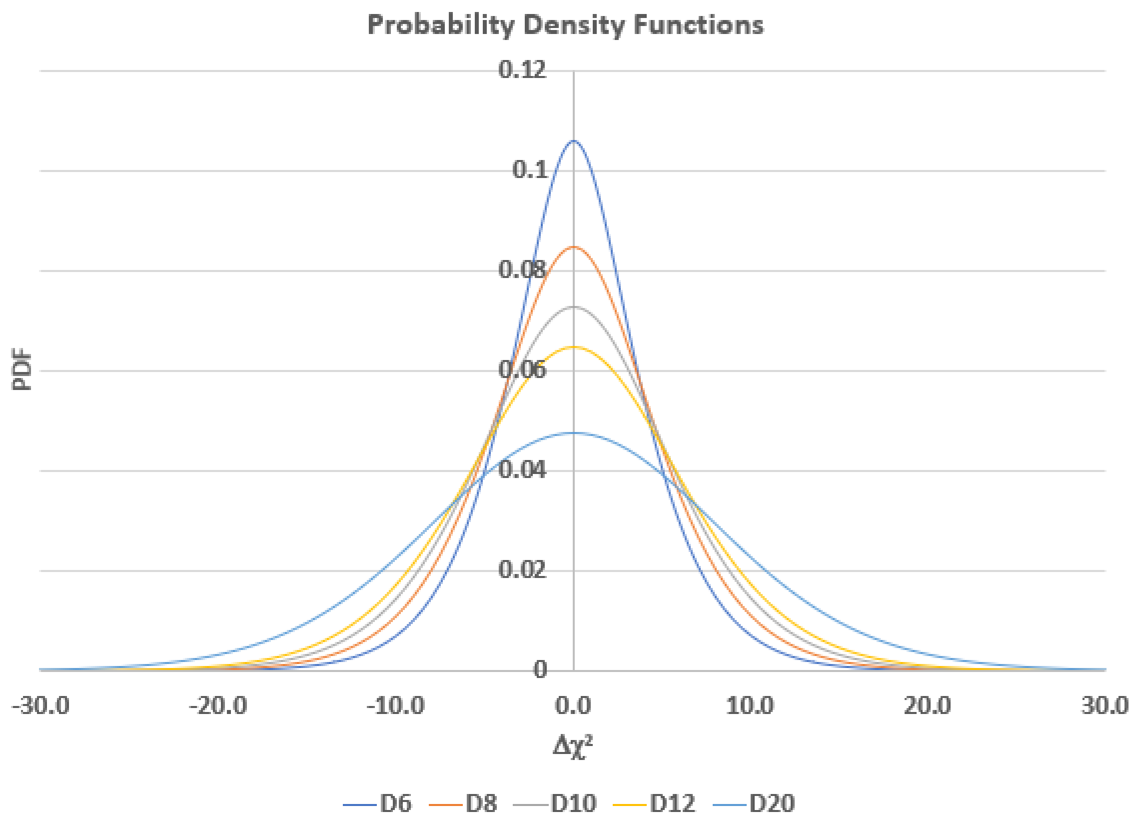

Figure 12. Probability densities that a fair die will have a higher χ^2 statistic than an unfair die

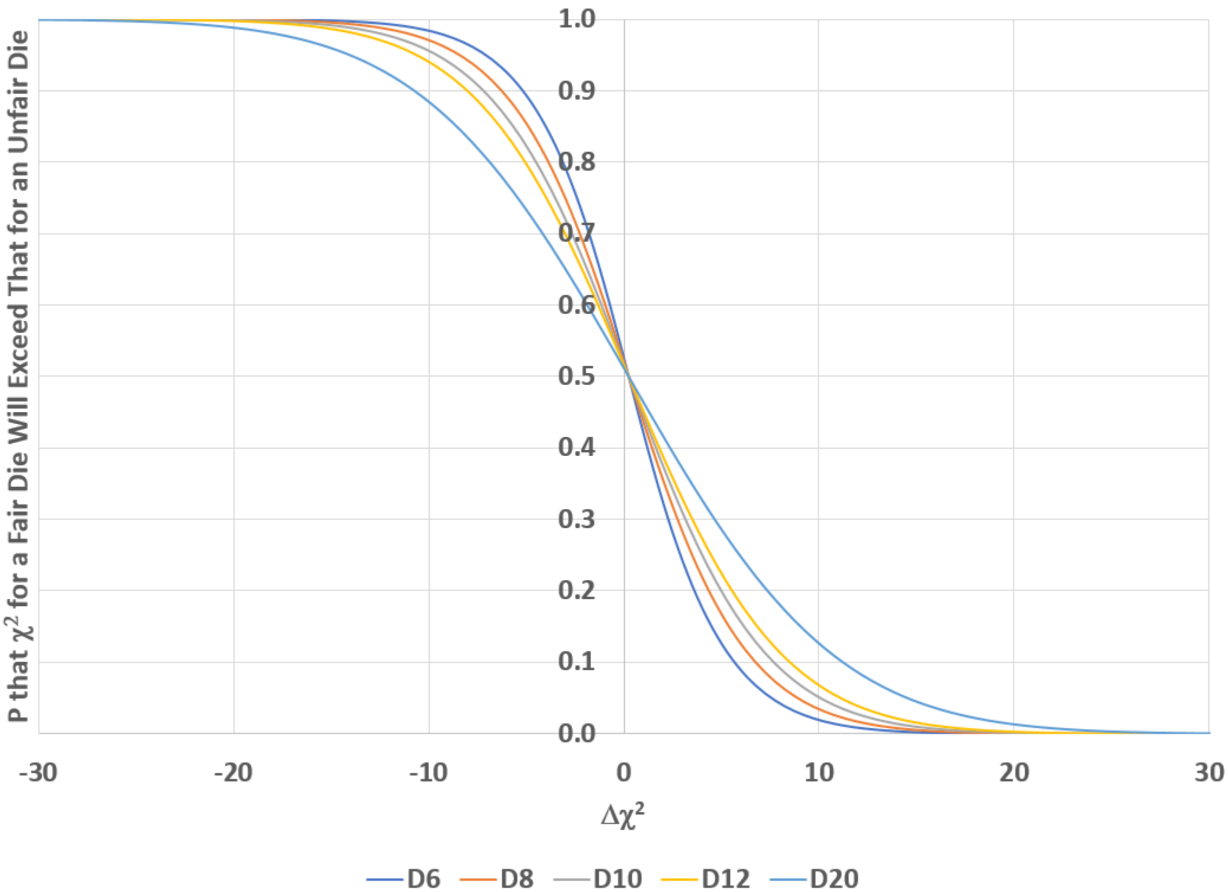

Figure 13. Cumulative probability that a fair die will have a higher χ^2 value than an unfair die

Finally, the chi square statistic for unfair dice does not follow a chi square distribution. The exact distribution of the statistic for fair and unfair dice is given by the multinomial distribution. The statistic for unfair dice is definitely not chi square distributed. This is illustrated clearly in Figure 14 which shows the exact discrete distribution for a highly unfair D3 die and the chi square distribution for 2 degrees of freedom. The D3 die probabilities are 0.45 for landing on 1, 0.1 for landing on a 2, and 0.45 for landing on a 3.

Calculating the exact probability distribution function for 3000 rolls of an unfair or fair D20 die may be difficult because of the exhaustive calculations that must be made. However, the chi square statistic is practically zero outside a hypersphere centered at $n_1 = n_2 = ... = n_{20} = 150$. The constant 150 is the radius of the hypersphere with 20 dimensions $(n_1, n_2, ..., n_{20})$. For $p_1 = p_{20} = 0.06$ outside a radius of about 480 the probabilities for any outcome are virtually zero as calculated with Excel. So the probabilities outside that hypersphere do not have to be calculated. This will simplify the calculations significantly. Put another way, all probabilities are virtually zero (less than 10^{-300}) when $\chi^2 > 1540$.

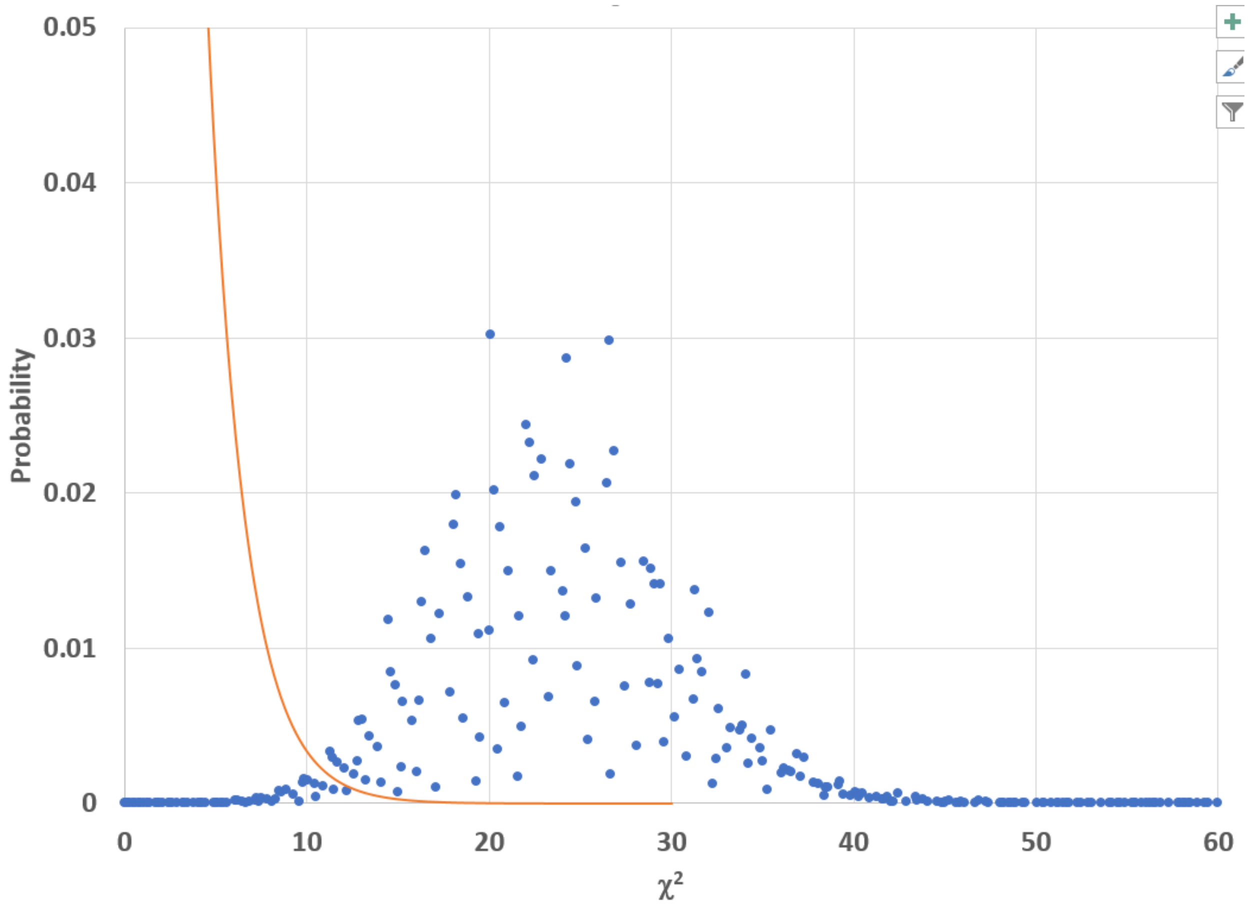

Figure 14. Exact discrete unfair D3 die 90 roll probability distribution and chi square distribution

Dimensional Effects

Our working hypothesis was that differences in dice diameters would affect the chi square statistic. Specifically, the larger the maximum difference in diameters, the higher the value of the chi square statistic. Figure 15 shows that over a certain range this is true. The chart is hard to interpret. However, if you take the dice with diameter differences less than 0.15 mm we get the chart in Figure 16. The correlation is quite high. It is even higher if we take the one additional point with a diameter difference of 0.27 mm. The implication is clear for dice manufacturers; aim for dimensional accuracy. This conclusion is from a very limited sample. If the results hold for dice in general, we could conclude that an accuracy less than 0.12 mm implies a high probability that the die tests fair in 3000 rolls. The implication would also be clear for gamers. Measure your dice and use the ones with the smallest diameter differences.

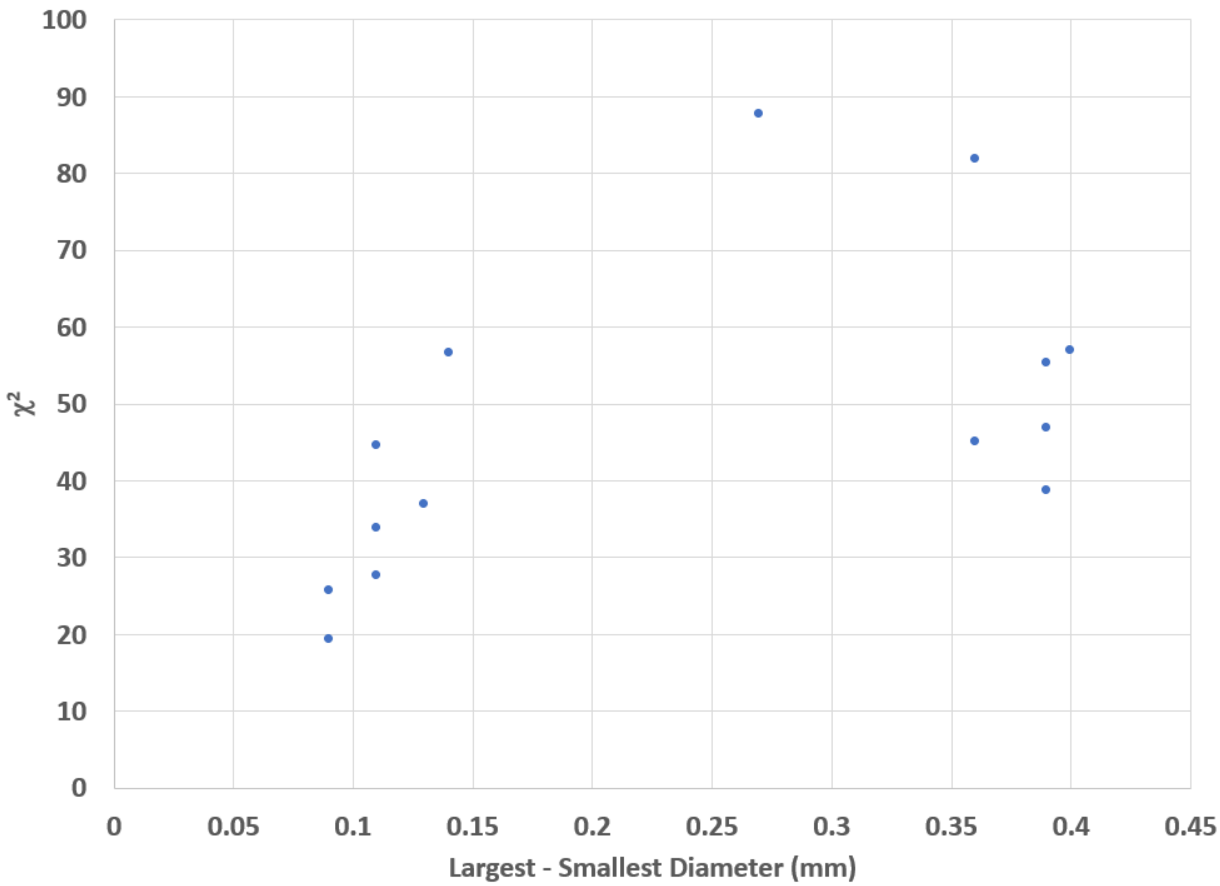

Figure 15. Chi Square Statistic Versus the Largest Diameter Difference for D20 Dice

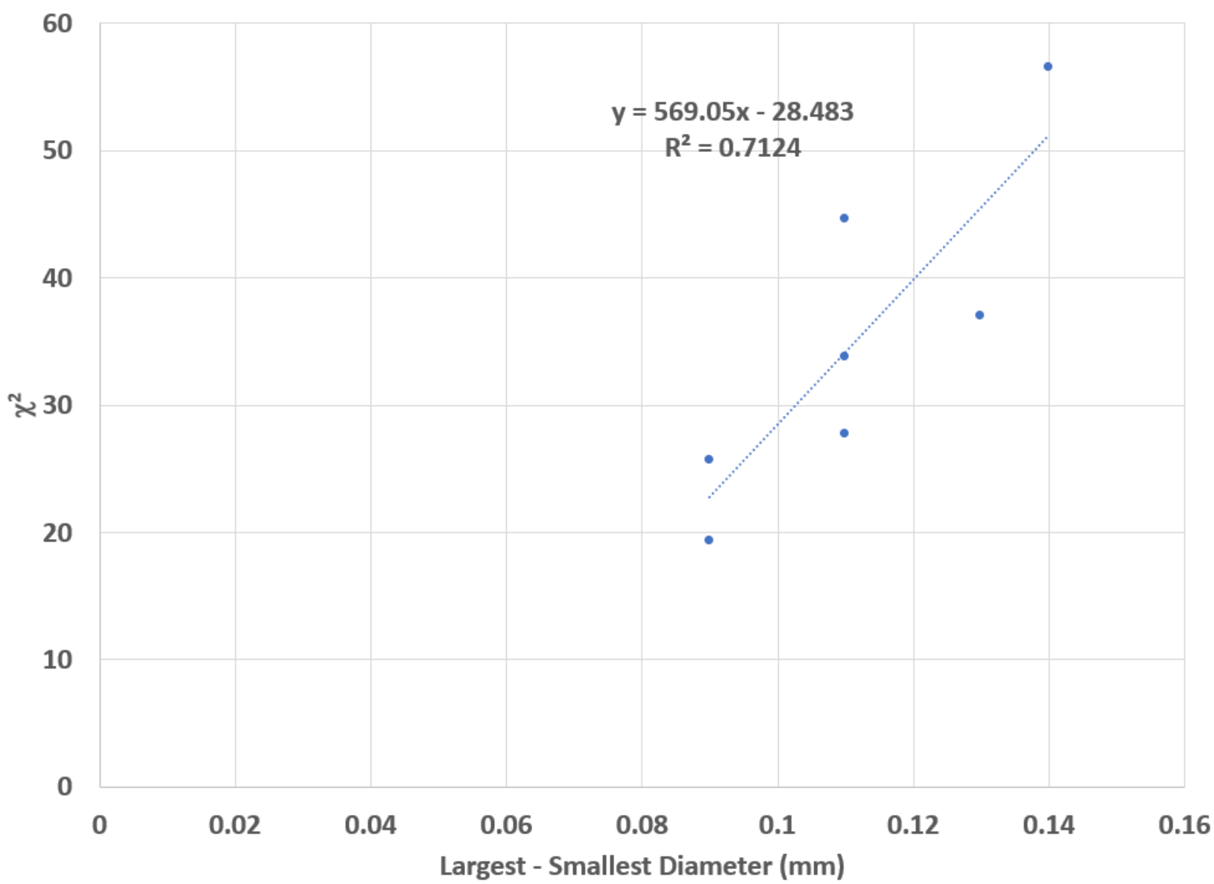

Figure 16. Chi Square Statistic for Dice with Maximum diameter differences Less than 0.14 mm.

Float Tests

Each of the plastic dice were also float tested to see if there was any significant correlation between the face that floated up most frequently and the face that turned up most frequently in the 3000 roll tests. The process used was to shake a die in a plastic container and dump it into a sugar solution. All three opaque Game Science dice floated in a solution with only a small amount of sugar. The translucent Game Science die and three of the Chessex dice floated in the solution with a moderate amount of sugar. The green-brown Chessex die required about 2 parts of sugar to one-part water. Not all of the sugar dissolved so stirring was required between each test so that the undissolved sugar would not settle out allowing the dense die to sink. The amount of sugar is important because the more sugar, the more viscous the mixture. The more viscous and denser the solution, the slower the time for the die to settle into an equilibrium position.

We observed dice that floated with corners up, with an edge up, and with one face up. Figure 17 shows a die floating with a corner up. Figure 18 shows a die floating with an edge up, and Figure 19 shows one with a face up. Sometimes a single die would exhibit all three behaviors.

We could plot face frequencies in the 3000-roll tests to face frequencies in the 26-float tests for each die. Figure 20 Shows a typical plot. The chart shows there is practically no correlation between roll face frequencies and float test face frequencies. This is not just for one die. It was a consistent result for all

8 dice. The r^2 value varied from 10^{-5} to 0.0791. In a typical Youtube float test, the die is flicked while in the water. We do not believe that is a valid test. Shaking them outside the solution and dropping into the solution should give more valid results. We believe the lack of correlation between dice rolls and float tests occurs because dynamically they are very different processes. In a float test, buoyancy plays a major role and is absent in an actual dice roll. A float test may or may not indicate if a die is unbalanced, but it has no value in predicting which faces are more or less likely to come up in a dice roll.

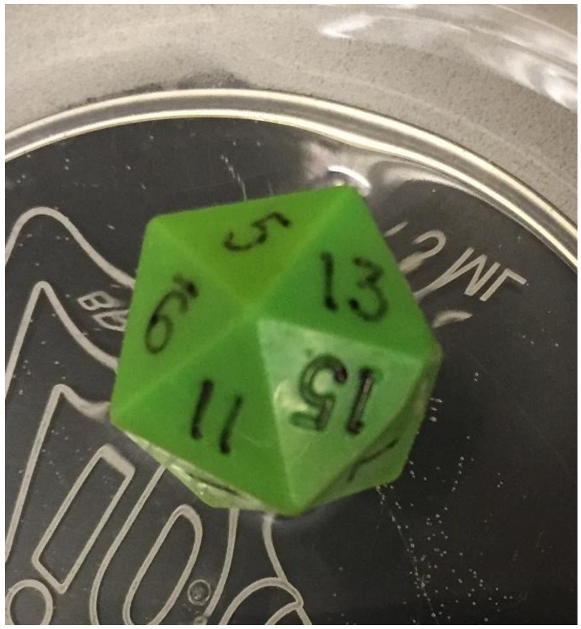

Figure 17. Die Floating with a Corner Up

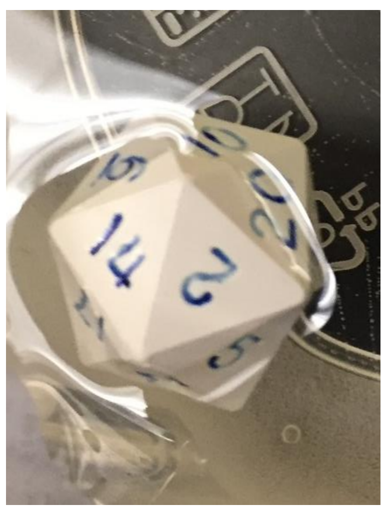

Figure 18. Die Floating Edge Up

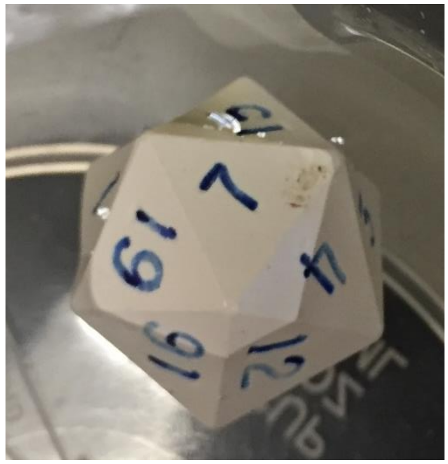

Figure 19. Die Floating Face Up

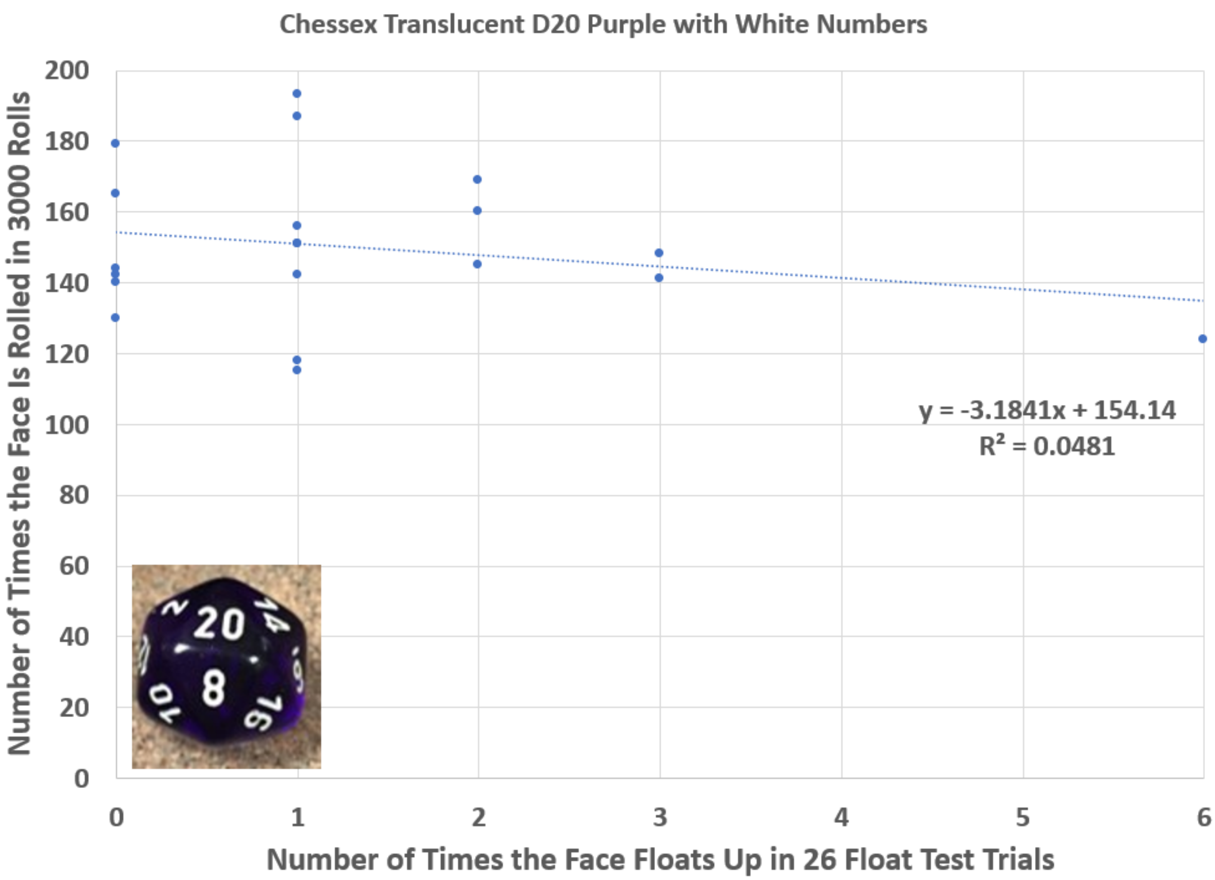

Figure 20. Plot of 3000-Roll Face Frequencies to 26-Float Test Face Frequencies.

Conclusions

1. All dice are unfair because they cannot be precisely manufactured with uniform density and dimensional correctness. However, the number of rolls required to detect the unfairness of a casino die manufactured to a few ten thousandths of an inch tolerance is very high.
2. It is very difficult to manufacture D20s that will test fair in 3000 rolls. This includes both plastic and metal dice.
3. Though our sample was small, there appears to be some merit to measuring the diameters of D20 dice; a die with a maximum diameter difference of less than 0.12 mm has a good chance of testing fair in 3000 rolls.
4. Float testing dice may tell if the die is unbalanced but it will not tell which side or sides will roll with higher or lower probabilities.
5. The manner in which a die is rolled makes a large difference in the face probabilities. It varies with surface and the manner in which it is rolled, that is, dice cup to felt, drop from a height to a hard surface, or rolling in a dice tower.
6. Rolling D20 dice 100 times and using the obtained chi square values to judge relative fairness is statistically unjustifiable.
7. A D20 die must be highly unfair to be detected in a few hundred rolls.
8. The exact probability distributions for unfair dice are not asymptotically a chi square distribution.

References:

Awesome Dice Blog (2012).

CritRoleStats (2016). “Wheatoning,” https://www.critrolestats.com/blog/2017/2/19/wheatoning.

Fisher, Daniel (2015). “Dice Tower vs. Table Top which is better?”, https://www.youtube.com/watch?v=5HXYOgjp3go

Geek and Sundry (2018). Destruction of Will Wheaton Touched Die, Facebook Page, https://www.facebook.com/geekandsundry/videos/khary-payton-vs-the-cursed-d20/1933456880040096/.

Geek and Sundry (2016). “15 of the Craziest Superstitions about Dice,” https://geekandsundry.com/15-of-the-craziest-superstitions-about-dice/.

Fickett, Mark, (2016). “Dice Roller,” http://www.markfickett.com/stuff/artPage.php?id=389.

Haowei, Liang and Yang, Xiaochen (2013). “The Study of Probabilities and Dynamics of Unfair Dice,” http://www.yau-awards.science/wp-content/uploads/2017/11/%E7%89%A9%E7%90%861140%E6%A2%81%E7%9A%93%E5%A8%81%E3%80%81%E6%9D%A8%E7%AC%91%E5%B0%98.pdf, 78 pp.

Labby, Zacariah (2009), “Weldon’s Dice, Automated,” *Chance,* Vol. 22, No. 4, pp. 6-13.

Reimer, Wolfgang, Stoyan, Dietrich, Obreschkow, Danail (2014). *Metrika*, Vol. 77, pp. 247-256.

Zocchi, Louis (2015). "Louis Zocchi of Gamescience at Gen Con 2015," https://www.youtube.com/watch?v=MRzg_M8pQms.

Appendix A. Running Chi Square Statistic for Hand Rolled Dice

The red line in each graph is the critical χ^2 value.

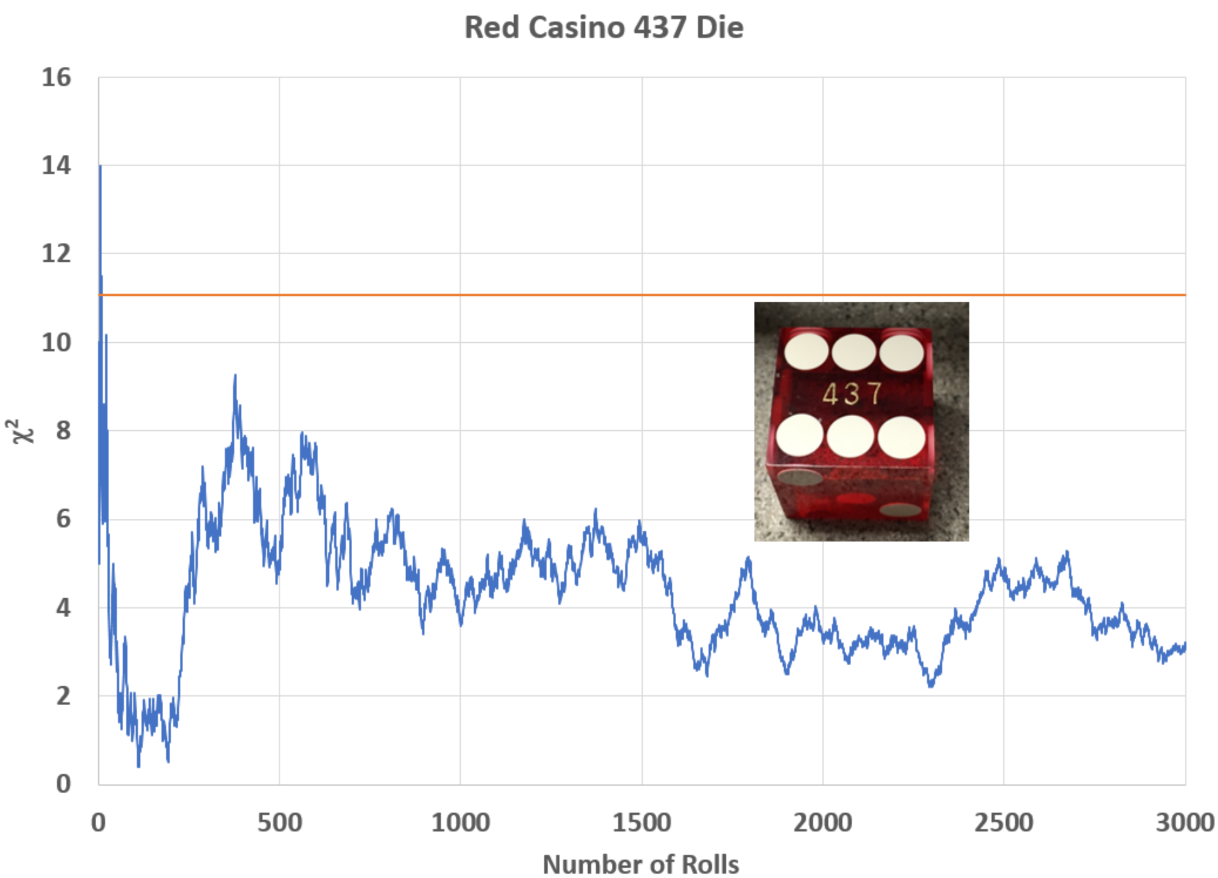

Figure A.1 Precision Casino Die

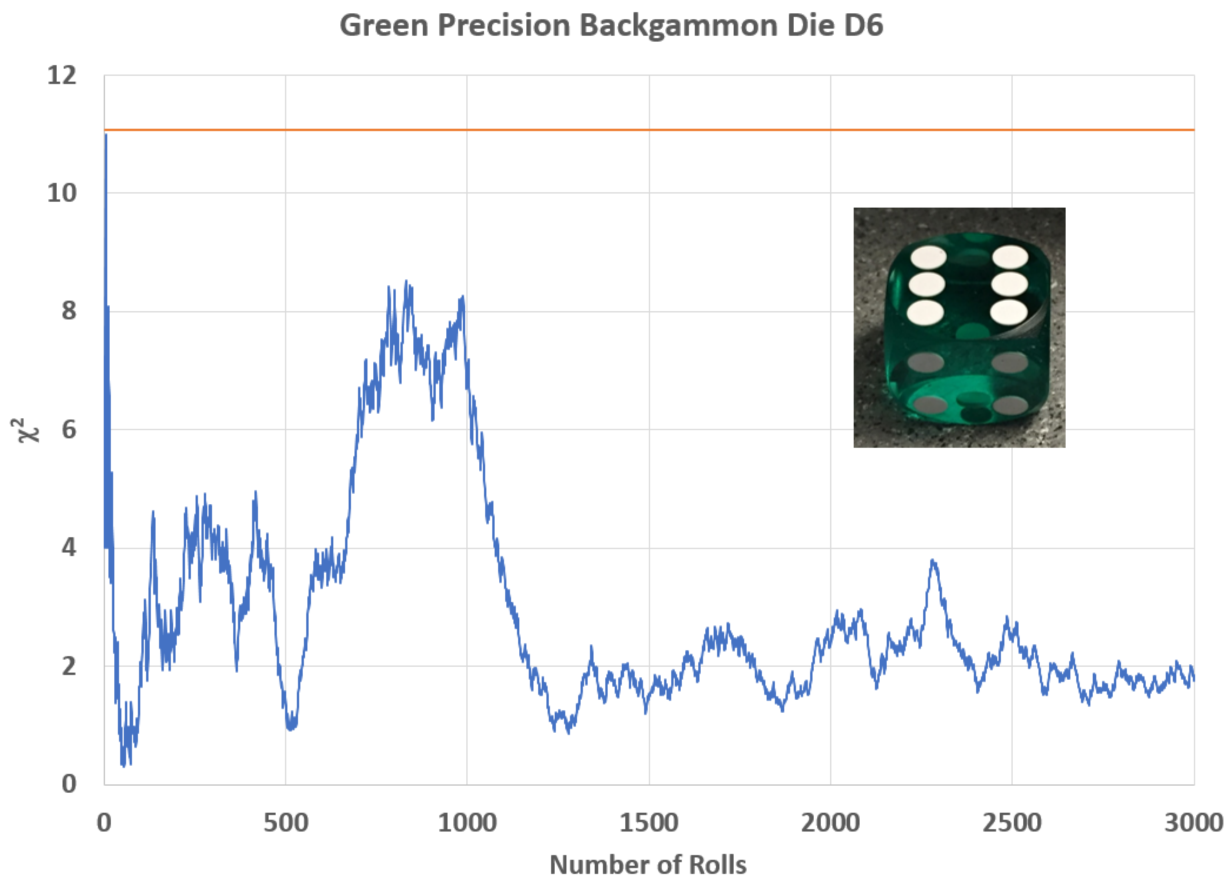

Figure A.2 Precision Backgammon Die

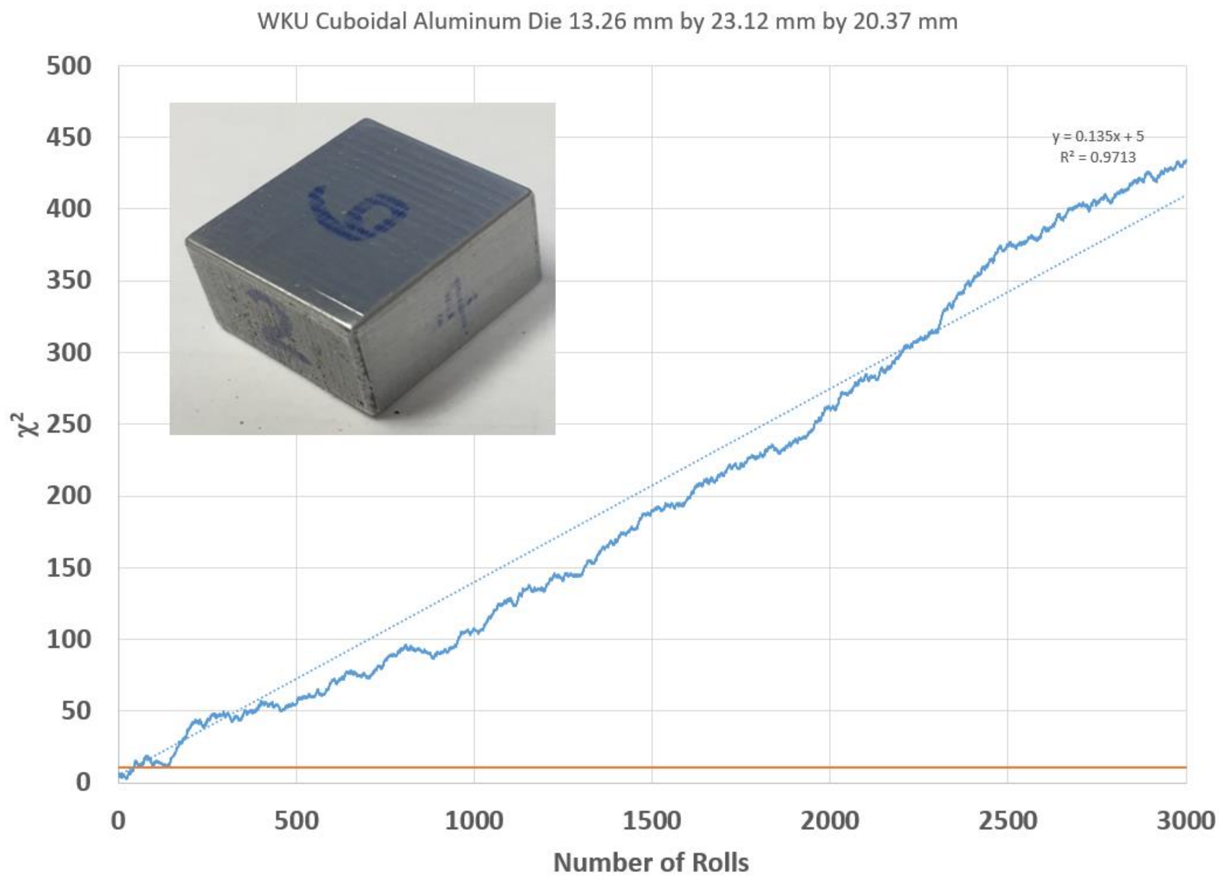

Figure A.3 WKU Machined D6 Cuboid with the Same Dimensions as Die in Reimer Paper

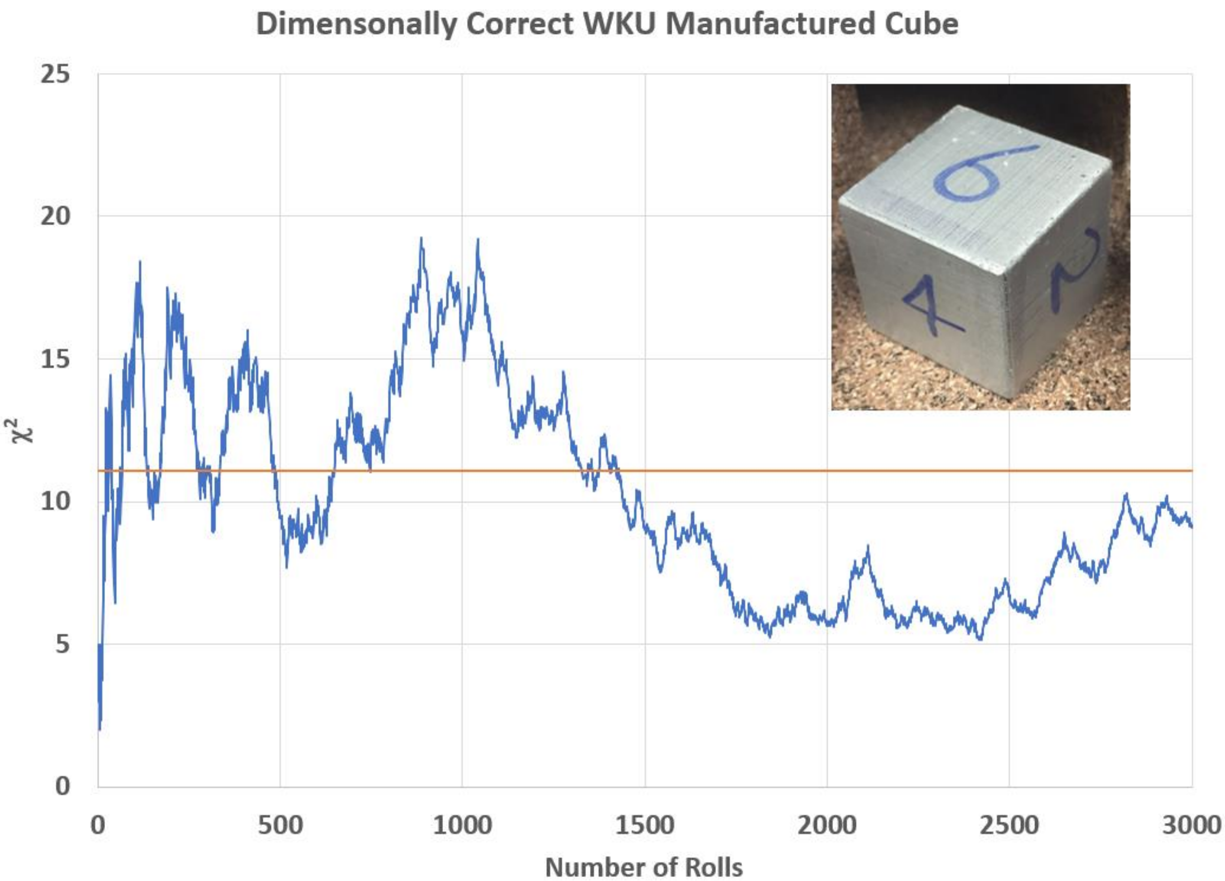

Figure A.4 WKU D6 Machined Cube Die

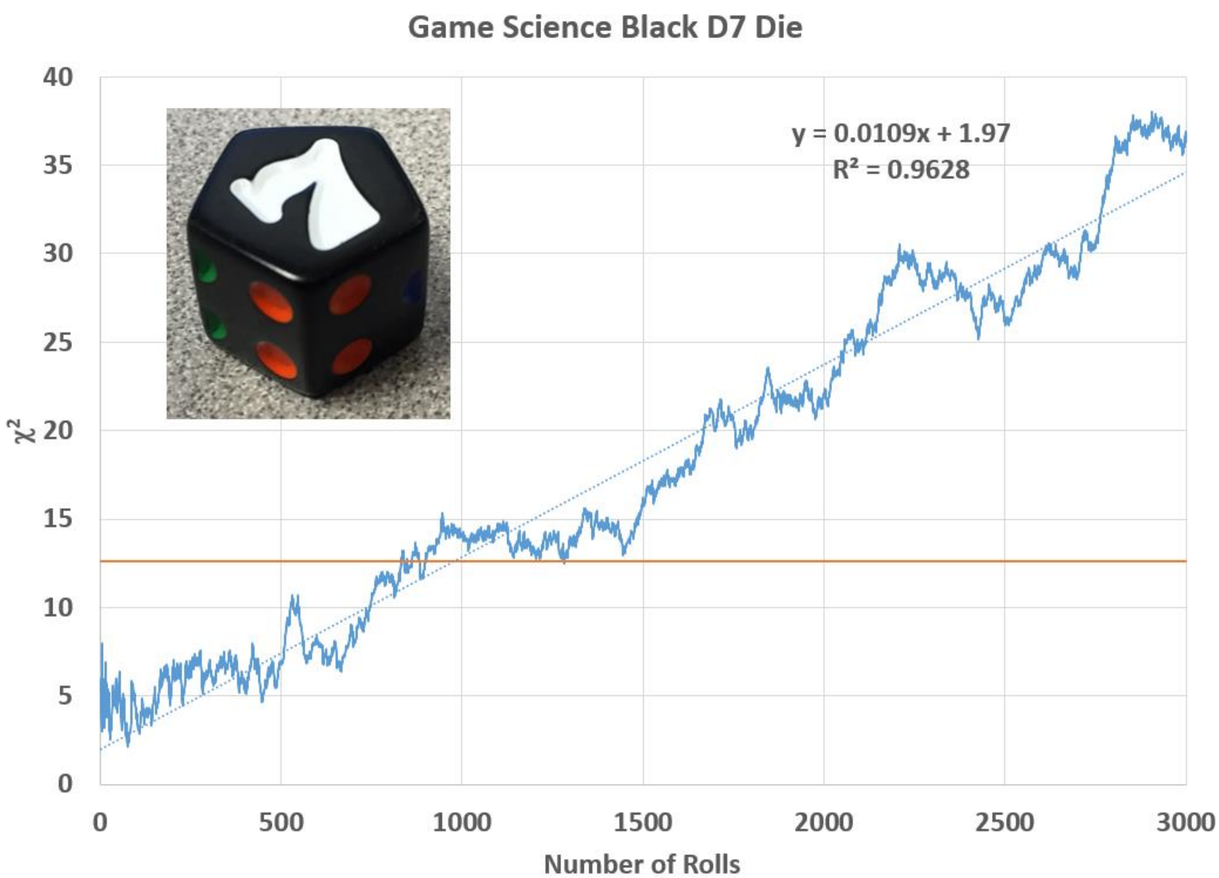

Figure A.5 Game Science D7 Die

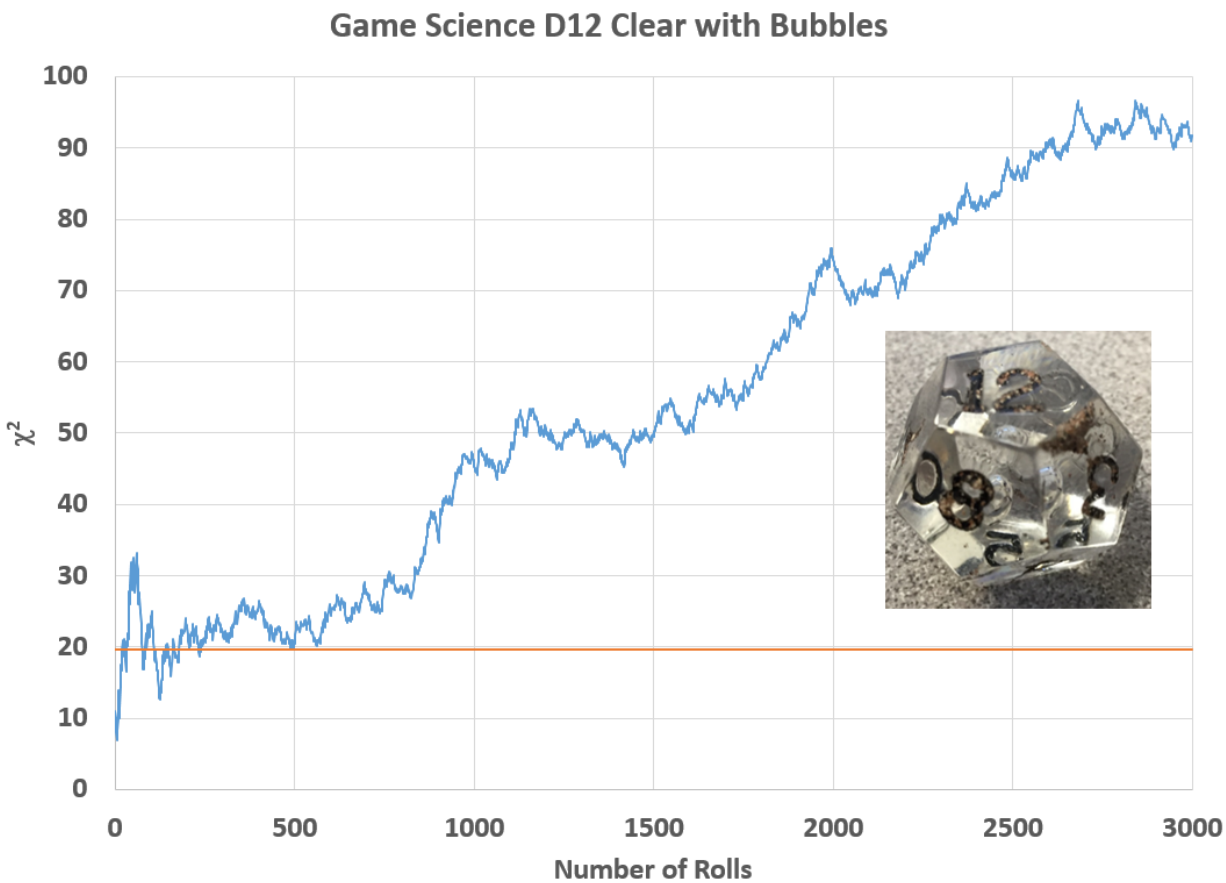

Figure A.6 Game Science D12 Die with a Large Bubble.
Note that Game Science sells these at a discounted price.

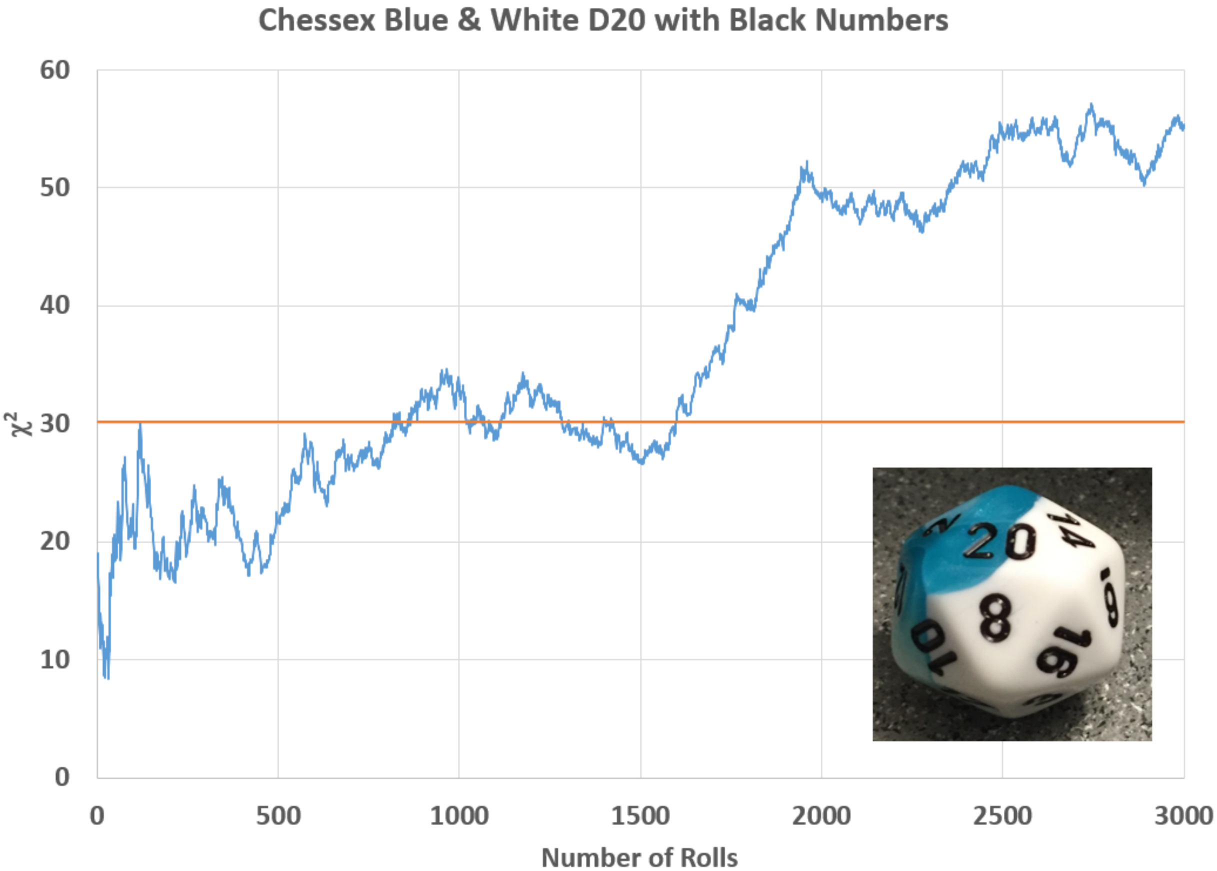

Figure A.7 Chessex D20 Die

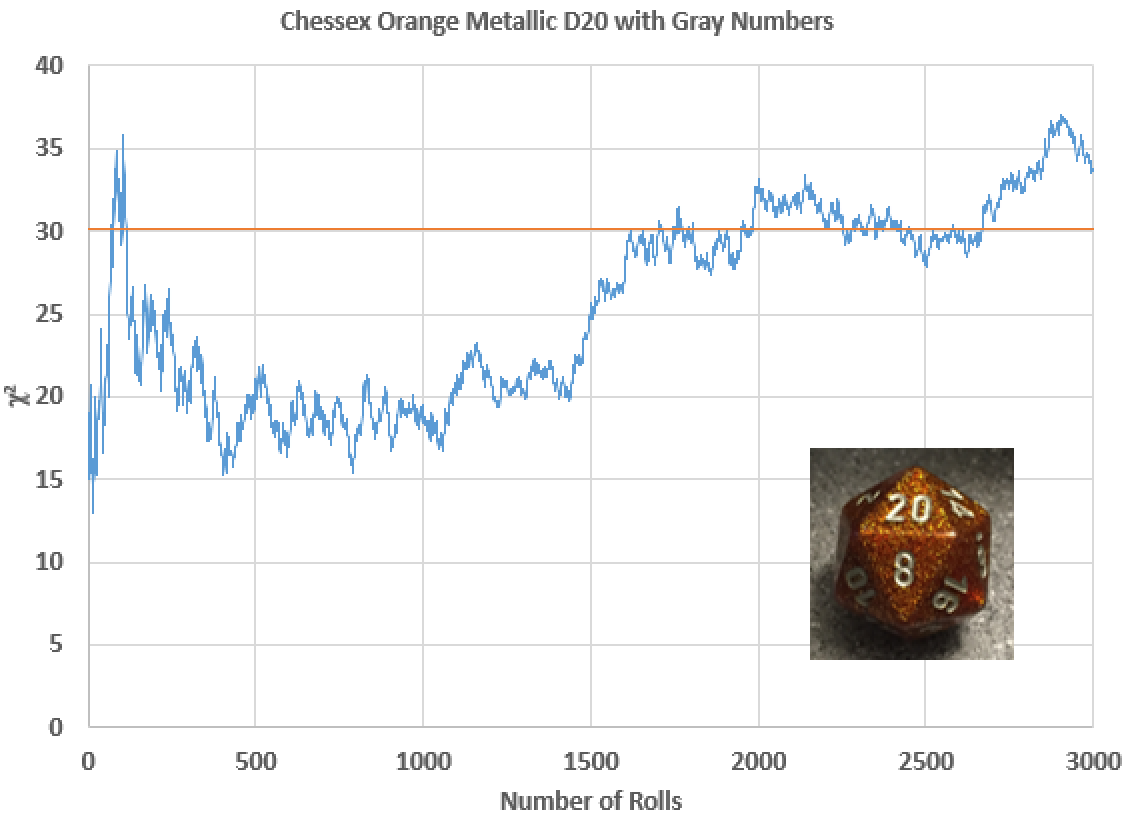

Figure A.8 Chessex D20 Die

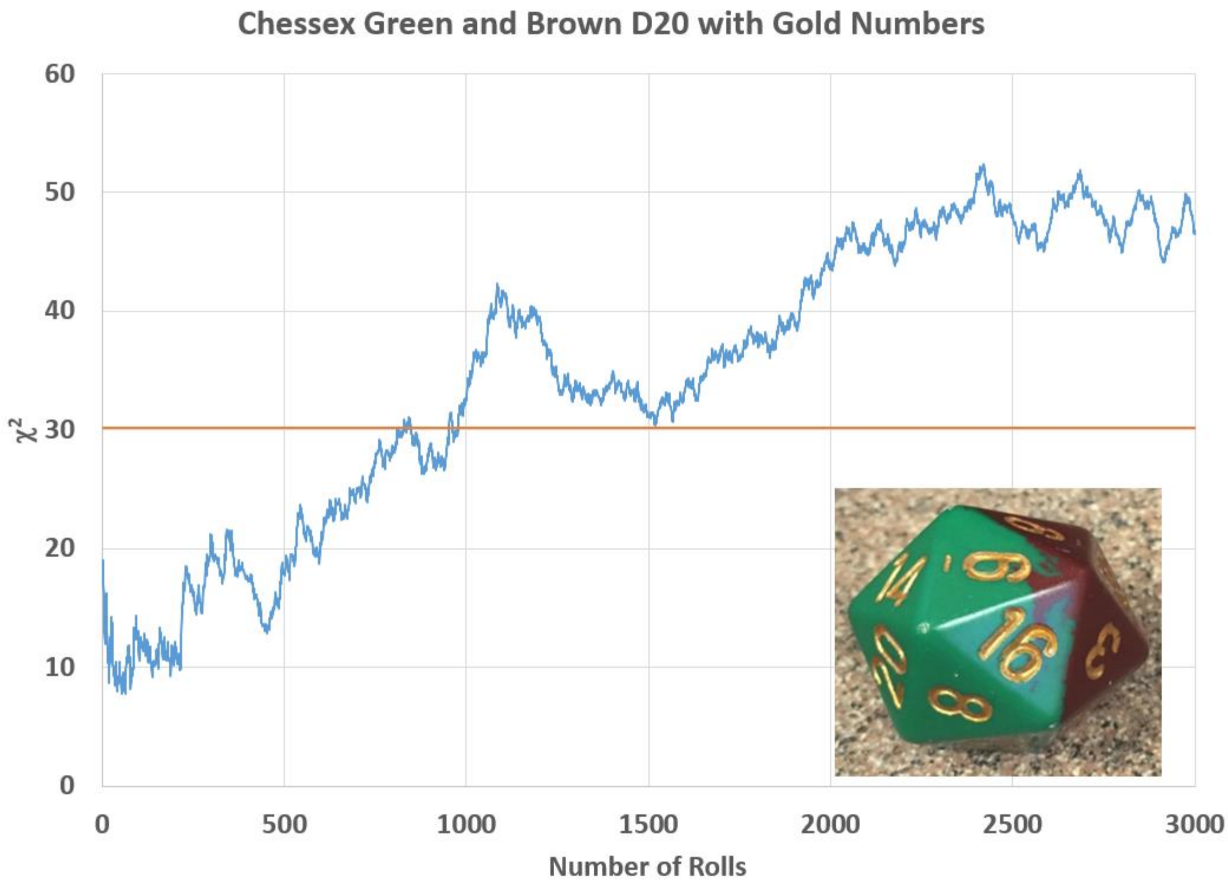

Figure A.9 Chessex D20 Die

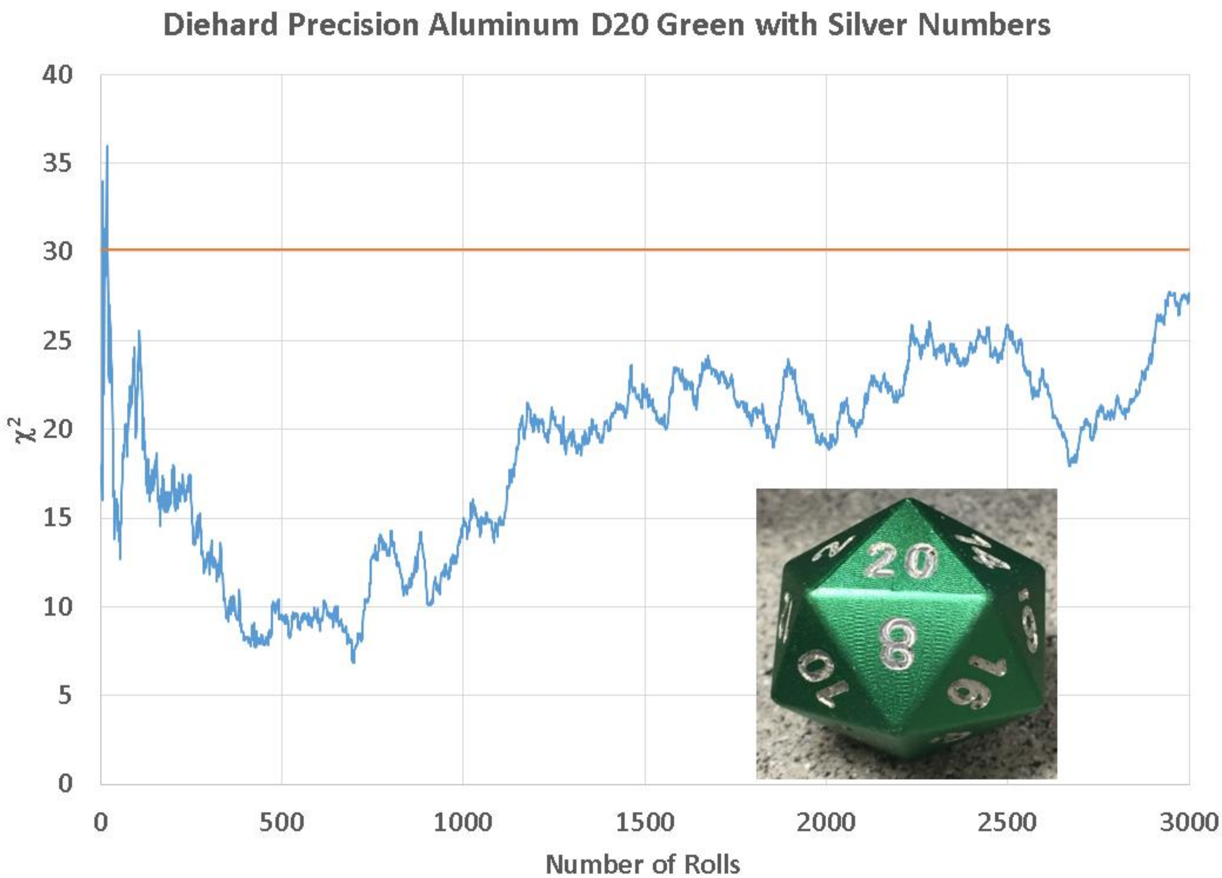

Figure A.10 Diehard Precision Aluminum D20 Die

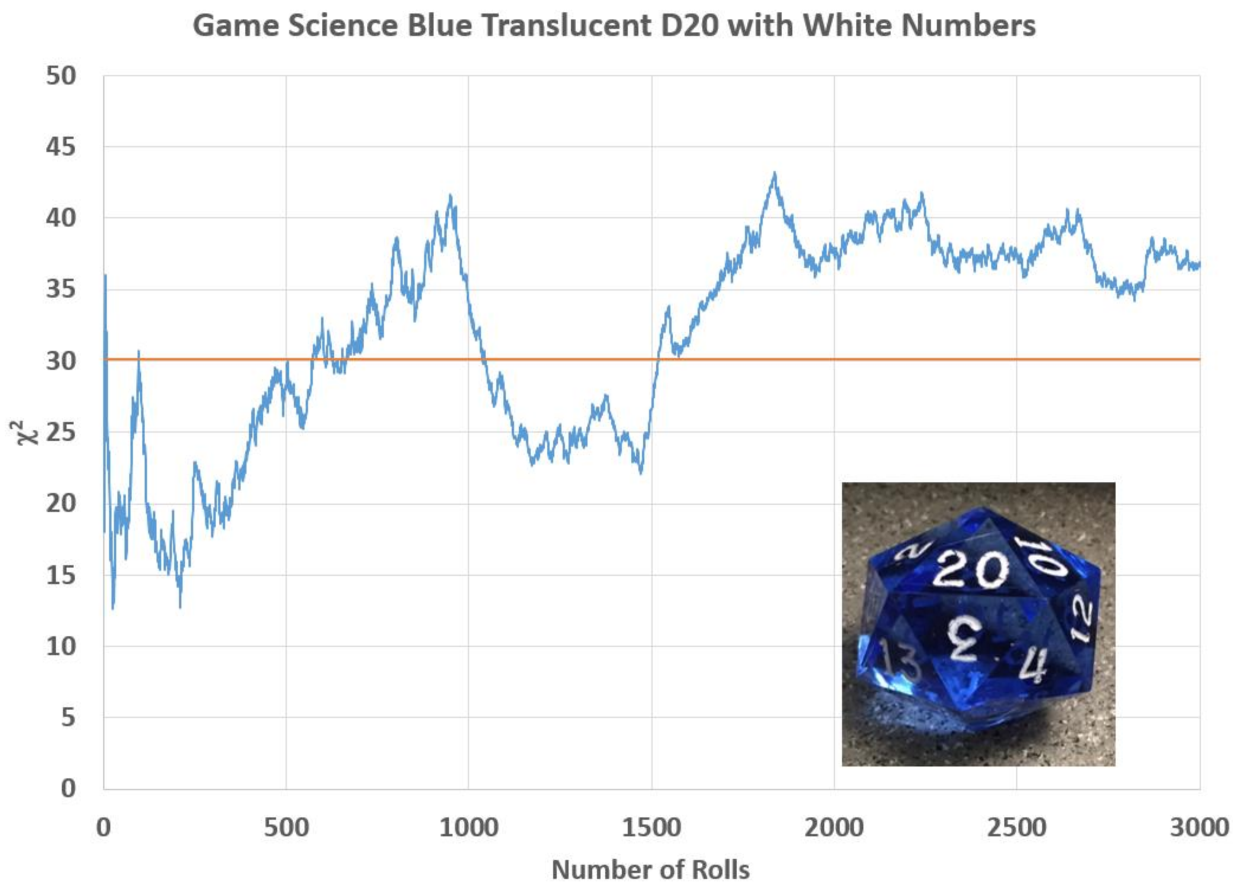

Figure A.11 Game Science D20 Die

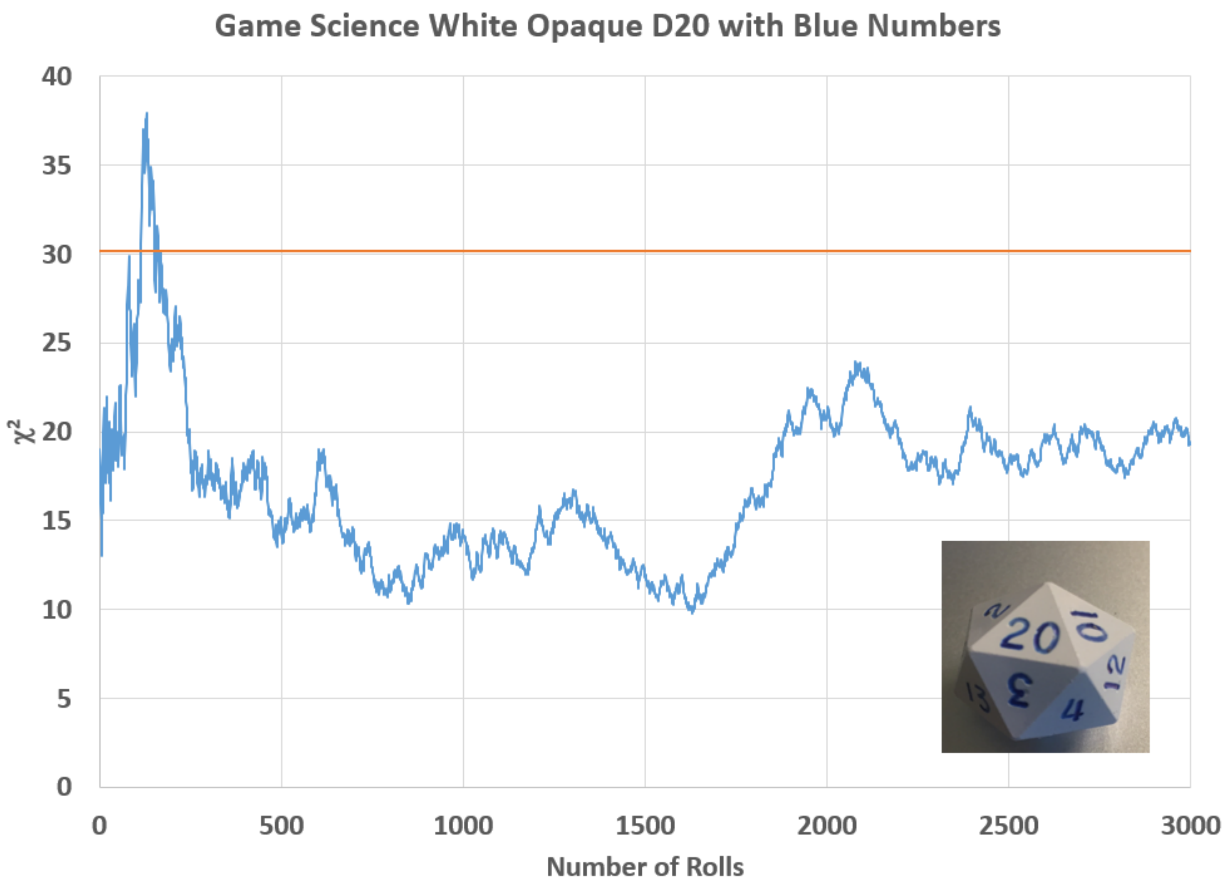

Figure A.12 Game Science D20 Die

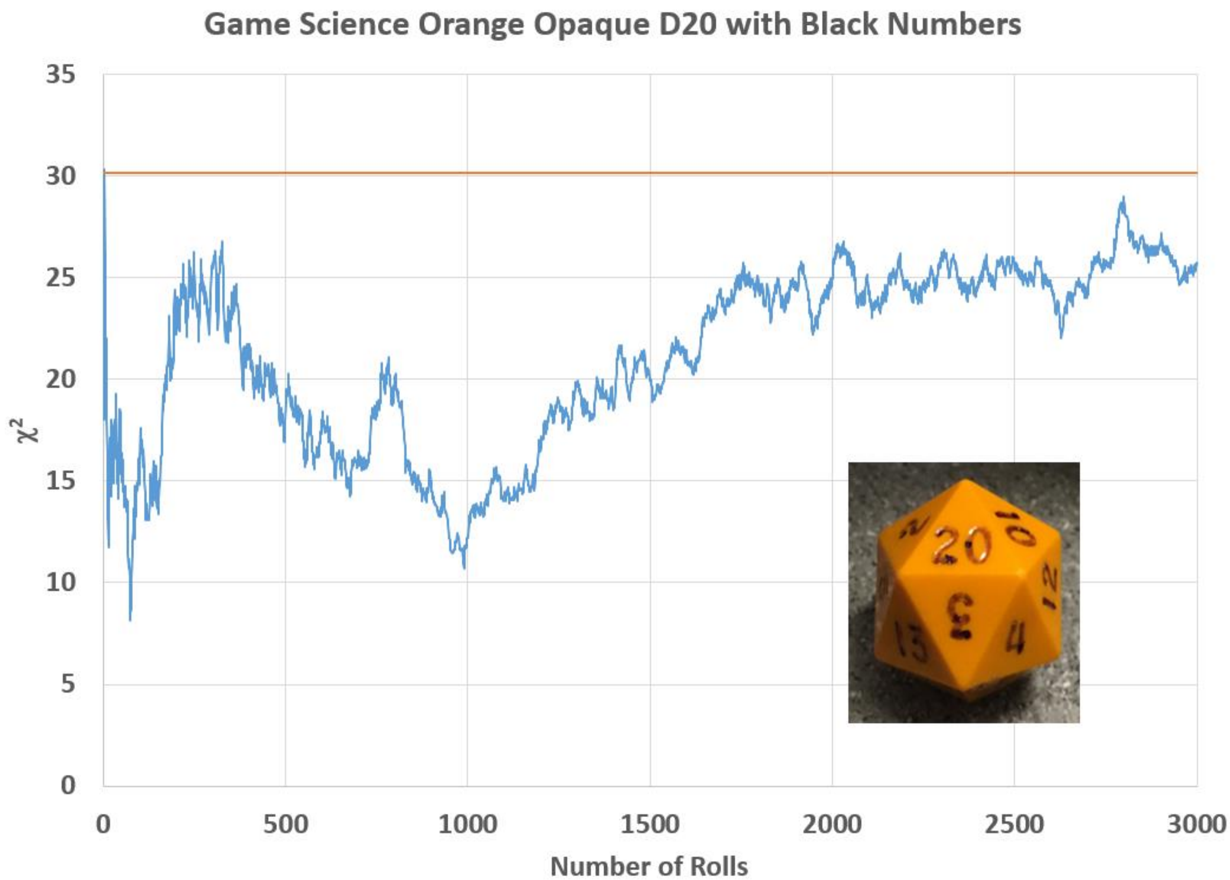

Figure A.13 Game Science D20 Die

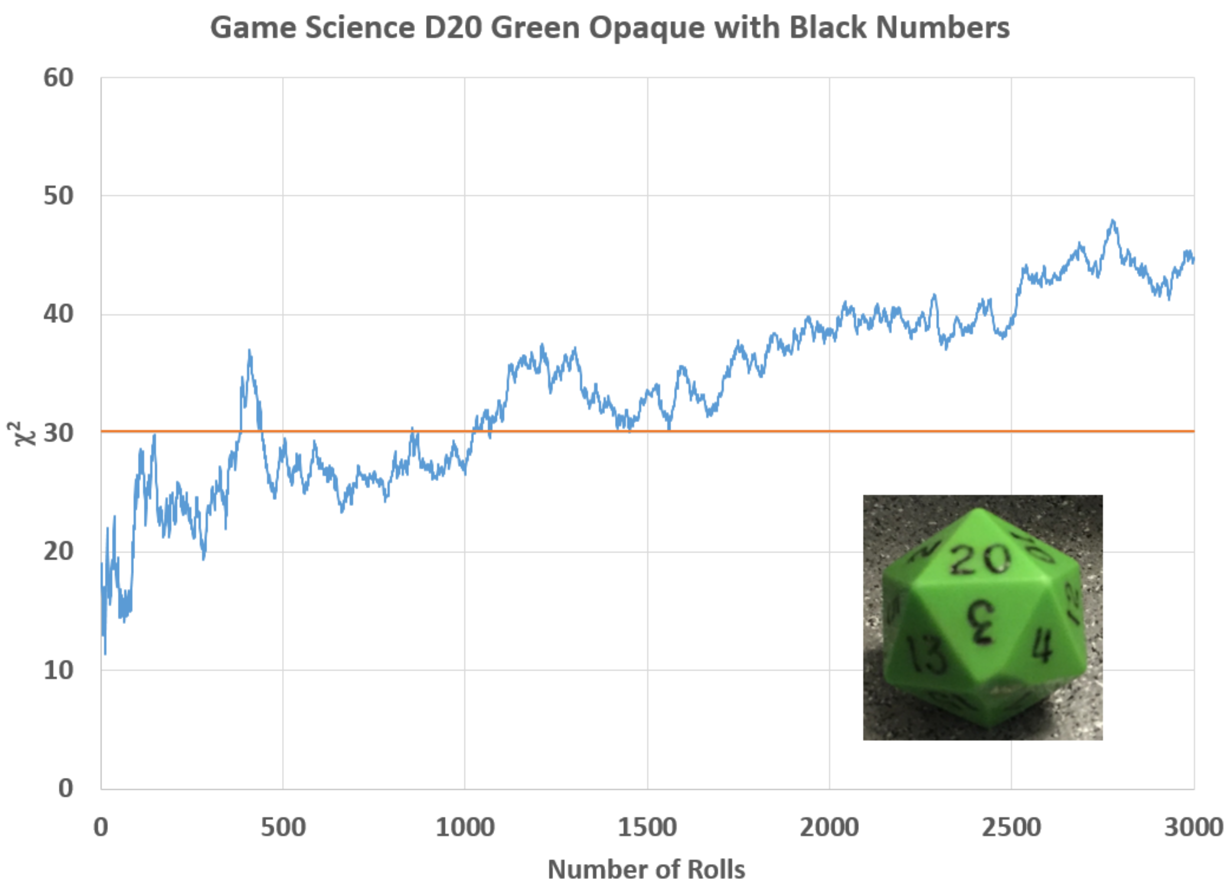

Figure A.14 Game Science D20 Die

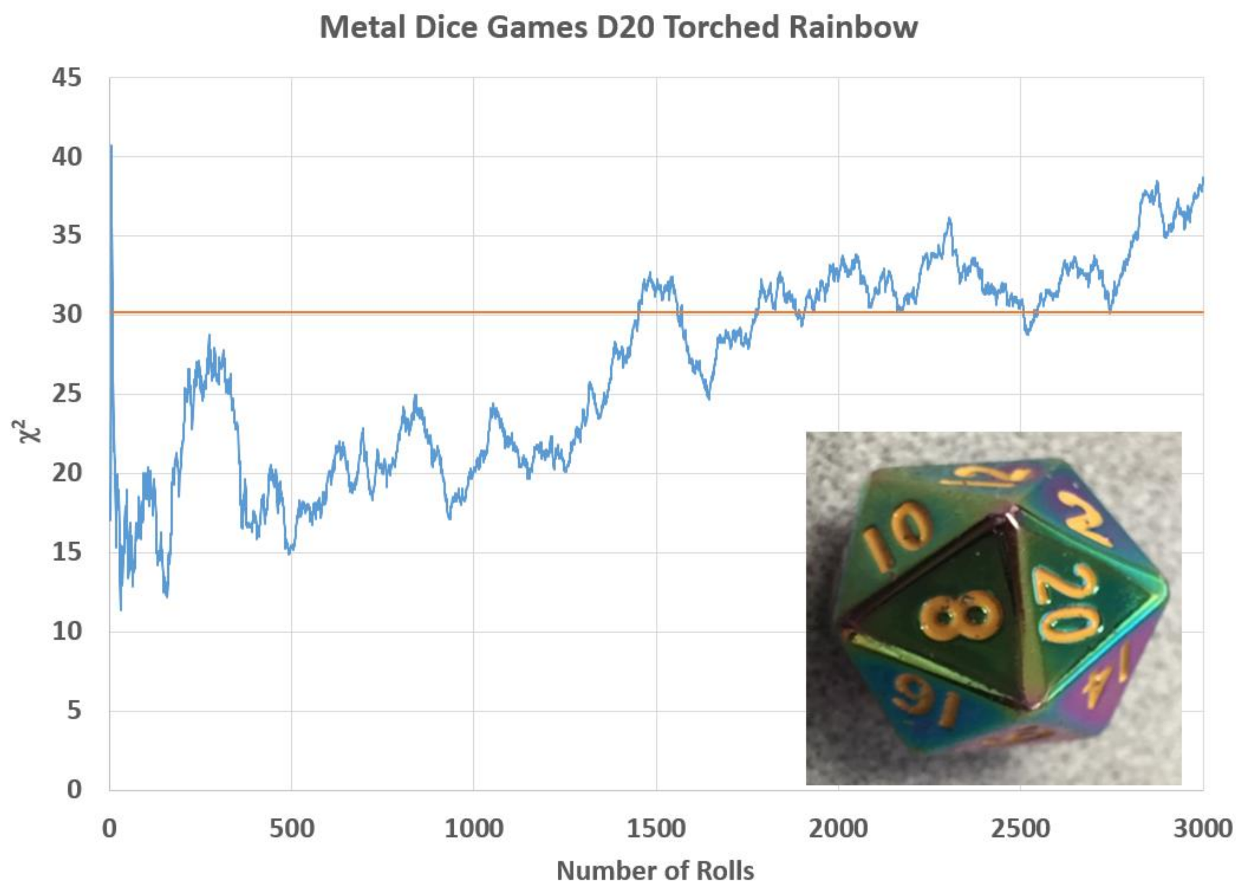

Figure A.15 Metal Dice Games Torched Rainbow D20

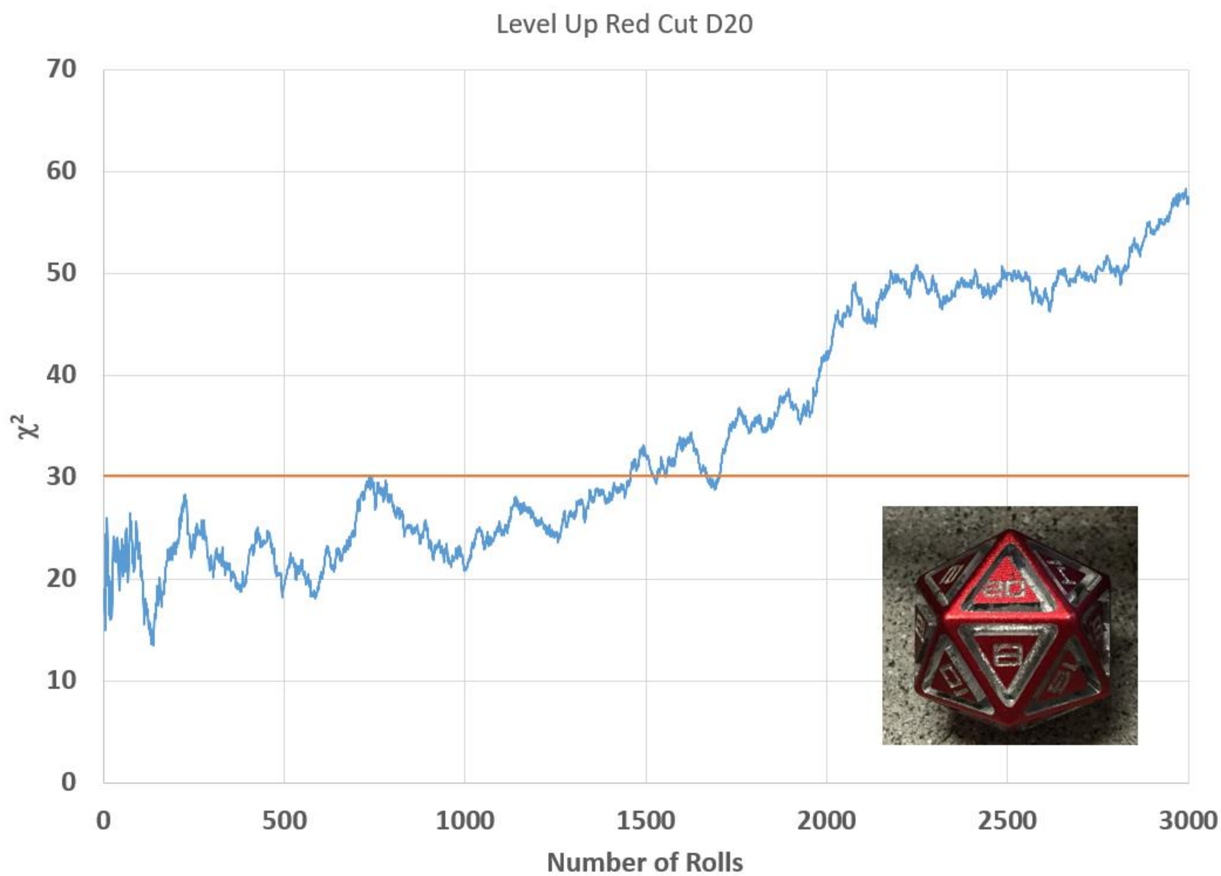

Figure A.16 Level Up Cut Red Aluminum D20 Die

Note: This is the most expensive and the worst die I own. At times to determine the roll you have to turn the die over and read the opposite side. The computer cutting of the numbers is awful and some numbers are almost impossible to read.

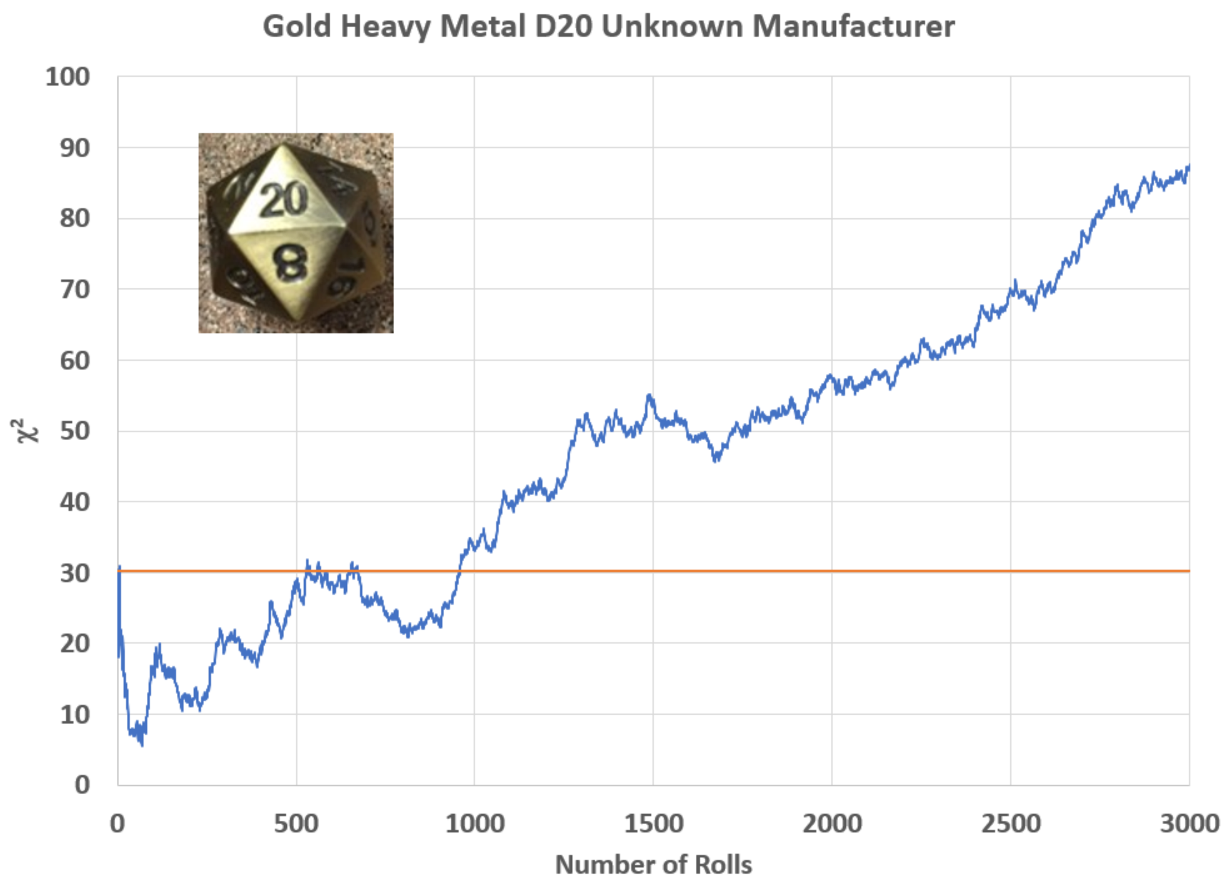

Figure A.17 D20 made of a heavy metal from an unknown manufacturer

Figure A.18 Die Hard Dice D20 Heavy Metal Black with Red Numbers

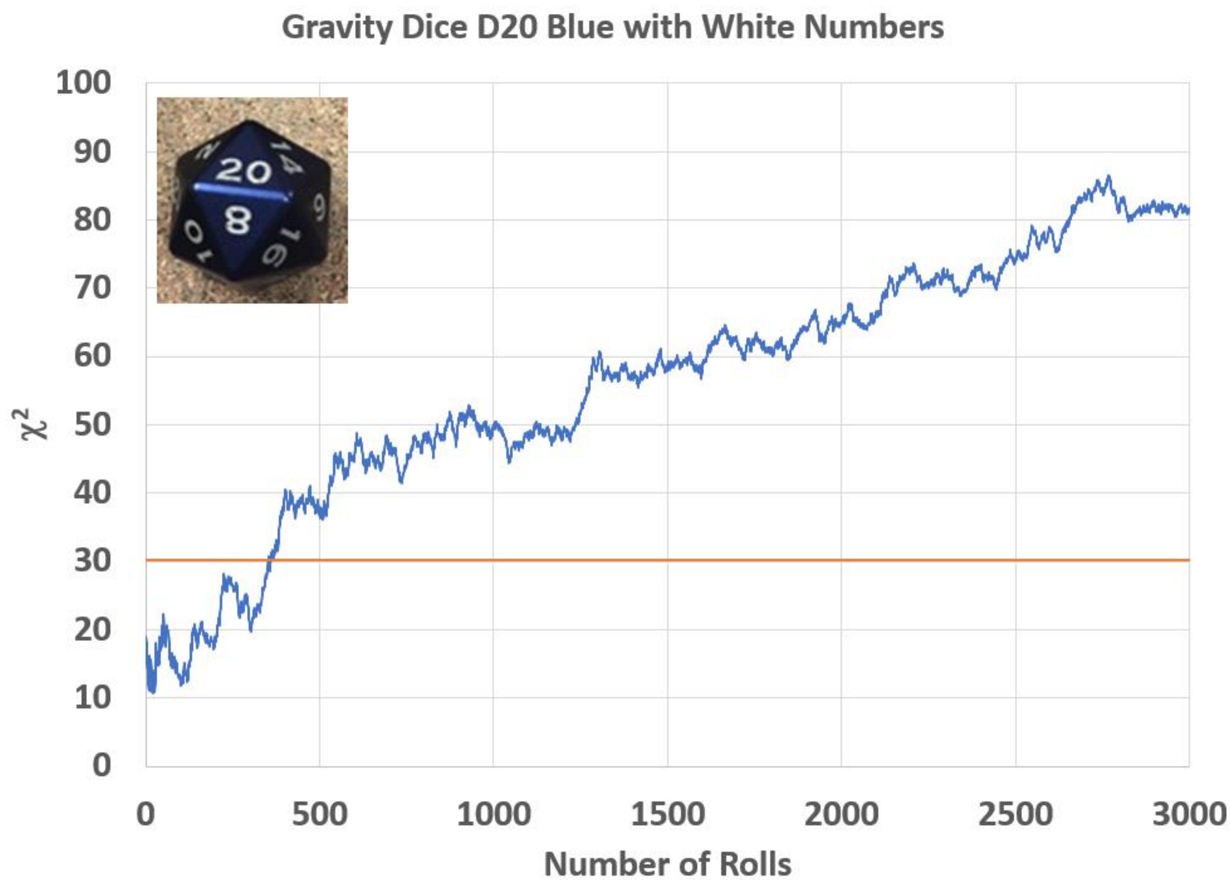

Figure A.19 Gravity Dice D20 Aluminum Navy with White Numbers

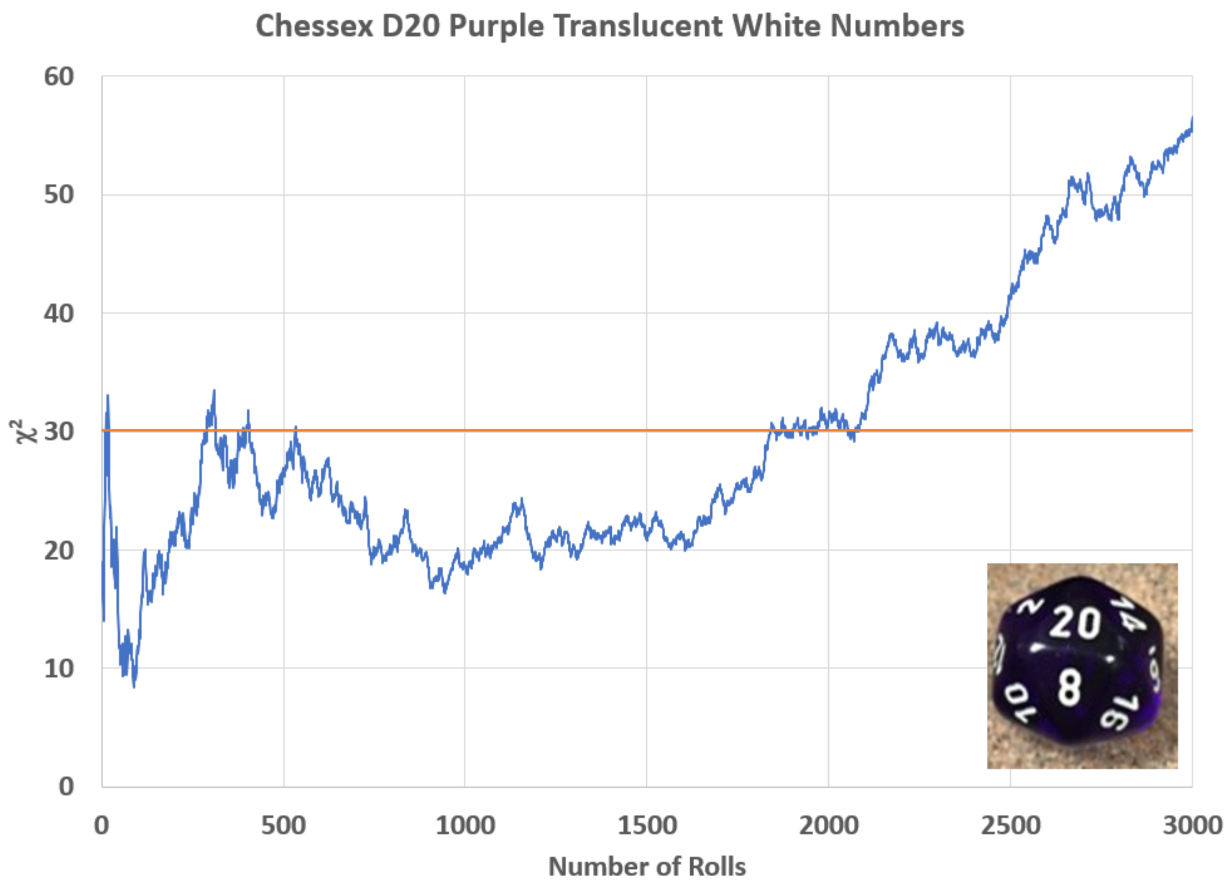

Figure A.20 Chessex Translucent Purple with White Numbers

Appendix B> Dice Testing Links

Note: Water solution float testing is not going to work on metal dice. You might be able to drop them into high tubes of water, but getting them out would be an issue. These tests do not tell you how unfair a die is.

Dice Float Test Links

https://www.youtube.com/watch?v=M3Y0loGqarI
https://www.youtube.com/watch?v=7_6jxOurHis
https://www.youtube.com/watch?v=_HhFz7fsFKk
https://www.youtube.com/watch?v=BOgnRrGoIi8
https://www.youtube.com/watch?v=tD2uy_jEUmo
https://www.youtube.com/watch?v=g21EAFWPsD4
https://www.youtube.com/watch?v=tSPJXPkVvGI
https://www.youtube.com/watch?v=7_6jxOurHis&t=13s
https://www.youtube.com/watch?v=E11v5QVhVfU
https://www.youtube.com/watch?v=Yxf1UYpkmuE
https://www.youtube.com/watch?v=bMFWZ-auNa4
https://www.youtube.com/watch?v=alZtaTUdBkU
https://www.youtube.com/watch?v=spp-M6Ur1-I
https://www.youtube.com/watch?v=RHKxdun-klg
https://www.youtube.com/watch?v=BaWT3gqWqDU
https://www.youtube.com/watch?v=fhoiFclRIeU
https://www.youtube.com/watch?v=i5IMo-Bu_4o
https://www.youtube.com/watch?v=t9t55XEAP2Q
https://www.youtube.com/watch?v=7Q6fIDk7b_c
https://www.youtube.com/watch?v=hlIqcigqbW4
https://www.youtube.com/watch?v=Yxf1UYpkmuE
https://www.youtube.com/watch?v=WBUimeE2gYc
https://www.youtube.com/watch?v=g21EAFWPsD4
https://www.youtube.com/watch?v=0-1qA-alRt8
https://www.youtube.com/watch?v=3aWeNnDGAlQ
https://www.youtube.com/watch?v=i8z6vTrU9d8
https://www.youtube.com/watch?v=89nU6Ac4njk
https://www.youtube.com/watch?v=IUlbD71RsII
https://www.youtube.com/watch?v=WIGlwM5MzxQ
https://www.youtube.com/watch?v=7Q6fIDk7b_c
https://www.kickstarter.com/projects/1874304670/prism-dice-handmade-from-bulletproof-glass/posts/1528387?lang=de
https://www.reddit.com/r/rpg/comments/3fpe0z/gamer_tests_15_sets_of_dice_for_balance_claims/

On line discussion of the effectiveness of the float test

20 Faces of Fate Reference to Golf Ball Testing and Daniel Fisher Test

Forbes Magazine reference to the 10,000 roll Chessex versus Game Science test

How casino dice are tested

10,000 Roll Test of Chessex Versus Game Science

Float Test Versus Chi Square – This one has a slight mistake, but it is interesting data.

Fate dice testing with random walk and chi square

Chi square and KS tests of dice fairness

Interesting D6 Roller and chi square analysis

D6 chi square test

Simple Rolling Machine for D20s

Simple D20 Rolling Machine and Analysis

Ethics of using lucky dice thread

Testing metal dice thread

Kevin Cook's web site. Kevin has the largest collection of dice in the world.

Persi Diaconis on Shaved Dice

Will Wheaton Cursed Dice

Will Wheaton Observed Dice Rolls

Geek and Sundry 15 of the Craziest Dice Superstitions

The Brothers Murph Cursed Dice and Dice Rituals

Cursed Dice, what can I do thread

Terminally Incoherent more dice superstitions

Cursed dice gifs

Lifting your dice curse

Louis Zocchi on dice

Lou Zocchi 2

Lou Zocchi 3

Numberphile Persi Diaconis Mathematical Fairness of Dice Part I

Numberphile Persi Diaconis Mathematical Fairness of Dice Part 2

Daniel Fisher Splits Open 2 D20s

Daniel Fisher on Bubbles in Dice

Daniel Fisher Recommends a Dice Tower

Daniel Fisher Again Recommends a Dice Tower

Testing by rolling many different dice: Note, he rolled a number of D6 Chessex dice. When you roll several they will average out fair. This would be valid only if each die was absolutely identical. They will not be because of the manufacturing process, but when you average several the results will look fair.

Does a D7 roll fair? Note: Appears to be a Game Science D7. He rolled it 1310 times and came up with a chi square value of 22.33 as opposed to a chi square critical value (95% level of confidence) of 12.59. Clearly unfair. The p value is 0.001055 or only ~ 1/1000 chance that a fair die would come up with this value of the chi square statistic. He did not calculate chi square, but gave enough information so we could. He was unable to draw a conclusion but we can. It is HIGHLY probable that the die is unfair, just as the one we rolled 3000 times tested unfair.

Discussion of fair dice, mostly D6s Note: Incorrectly states that square edges make dice roll less fair. Also incorrectly states smaller dice are better. Prefers Chessex 12 mm translucent dice. He states they are fair. He makes some correct statements, but some statements are not supported by statistics. Refers to Daniel Fishers float test.

Testing a D48 Note: She tests a D48 with more than 13,000 rolls. She gives a bar chart of the outcomes of the rolls. Numbers are given for some of the sides. The approximate chi square value is 620 which gives a p of zero. Definitely an unfair die. She does not give the chi square value. I calculated it with the approximate values of some of the numbers estimated off her graph.