The Mysteries of Dice

By: Jonathan Pak

March 23, 2018

Math 123

Professor Meshkat

Introduction:

Dice are small object with various amounts of sides that have numbers on them. The most common die today is a cubic shape with numbers ranging from one to six on each side. People believe that the die was first invented all the way back during the 6000 BC. They were not made to be used for board games as they are today but were used for such things as divine readings. It was until the 1600's where mathematical analysis was used to view the die. This is when dice become more sophisticated and seen more appearances in board games and gambling. Dice were made from a variety of materials such as bone, lead, crystal, ivory, silver, gold, glass, jade, brass, wood, marble, Bakelite, porcelain and other expensive and inexpensive materials. They can also come in various shapes and contain various amounts of sides. The six-sided die is the most common but there are many other shapes of dice with 4, 5, 7, 8, 10, 12, 16, and 20 or more sides. Most dice are made to be fair and precise to make sure there is no cheating during gambling games or unfairness in board games. However, dice made for board games are mass produced so the consistency and precision are not there. Are these dice as fair as companies claim they are?

Problem Statement 1: Do dice become more unfair as the number of sides increase?

Problem Statement 2: Do different surfaces change the fairness of dice?

Data Collection

The data was collected through rolling multiple sided dice on different surfaces in the villas. The data was collected throughout the quarter. There is an accurate representation of the population because the given rule for a Chi square test is that you need at least five observations per column. To make sure all columns had 5 instances of the number being rolled, I rolled the number of sides of the die times 10 to get the n for that die. In an ideal world, a population of 1000 rolls for each die would be ideal and allow greater accuracy. Due to time constraints, I was unable to achieve that goal.

Hypothesis Testing Techniques

For the Chi Square Test, the null hypothesis was the expected chance of landing on a side of the die. This equated to one divided by the number of sides the die had. All the alternative hypotheses were the probability of not equaling the null hypothesis. This caused all the Chi Square test to be two tailed. For the Z test, only the six-sided die was used so the null hypothesis was mu equaling 21 divided by 6. 21/6 is the sum all dots divided by the number of sides which is 6 for this die. The alternative hypothesis was mu not equaling 21/6 so the Z tests were also two tailed tests. I choose the Z test for the different surface problem because n (the number of rolls) was larger than 30 so a T test would no be appropriate.

For the Chi Square Test, I used the Excel function COUNTIF to count how many times I rolled on each difference side. I found the sum of these numbers to find my n. To find the Chi Square Value, I got the number of times I rolled on a side and subtracted that number by the expected number of rolls. Then, I squared the number obtained from the previous calculation. Finally, I divided that number by the expected number of rolls to get the Chi Square Value. To calculate the Z score for the Z-Test, I summed up all the numbers that I rolled for each different surface and divided by the n to get the average for each surface. I also calculated the standard deviation by using an Excel function STDEV. I divided that standard deviation by the square root of n to get the standard deviation of x. Then, I got the average I calculated earlier and subtracted that by the expected average in the null hypothesis. That number was then divided by the new standard deviation to get the Z-score.

Different Tests Used:

Chi Square Test:

D4:	**D6:**	**D8:**	**D10:**	**D12:**	**D20:**
H0: $p = .25$	**H0: $p = .167$**	**H0: $p = .125$**	**H0: $p = .10$**	**H0: $p = .083$**	**H0: $p = .05$**
H1: $p \neq .25$	**H1: $p \neq .167$**	**H1: $p \neq .125$**	**H1: $p \neq .10$**	**H1: $p \neq .083$**	**H1: $p \neq .05$**

Z-Test:

Table Top:	**Counter Top:**	**Carpet:**	**Box:**	**Game Board:**
H0: $\mu = 3.5$	**H0: $\mu = 3.5$**	**H0: $\mu = 3.5$**	**H0: $\mu = 3.5$**	**H0: $\mu = 3.5$**
H1: $\mu \neq 3.5$	**H1: $\mu \neq 3.5$**	**H1: $\mu \neq 3.5$**	**H1: $\mu \neq 3.5$**	**H1: $\mu \neq 3.5$**

D4:

Null Hypothesis: **H0: p = .25**

Alternative Hypothesis: **H1: p ≠ .25**

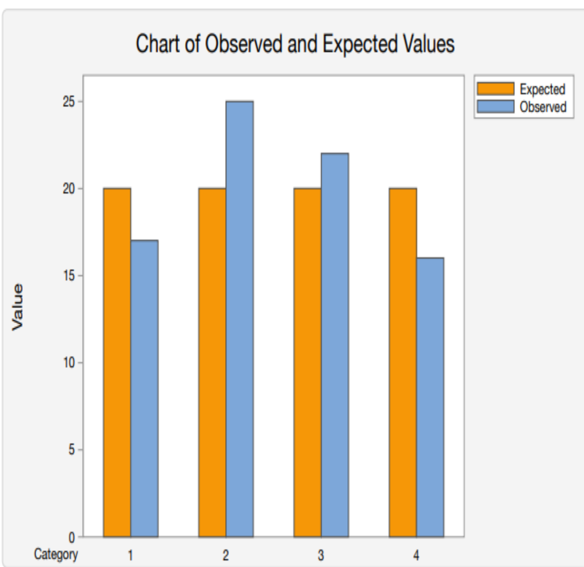
Chi-Square Goodness-of-Fit Test: Observed

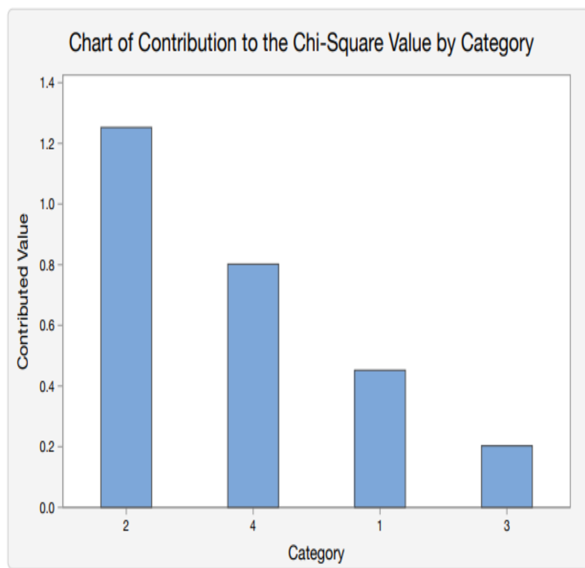
Observed and Expected Counts

Category	Observed	Test Proportion	Expected	Contribution to Chi-Sq
1	17	0.250000	20	0.45
2	25	0.250000	20	1.25
3	22	0.250000	20	0.20
4	16	0.250000	20	0.80

Chi-Square Test

N	DF	Chi-Sq	P-Value
80	3	2.70	0.4402





The Chi Square Value was 2.70 and we ended up getting a 0.4402 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.4402 > 0.025, we cannot reject the null hypothesis that the probability of landing on one of the sides is .25. Compared to the other Chi Square Tests, this was the only test to have an expected outcome of 20 rolls. I felt that 40 rolls as the n would be too low, so I decided to increase the total roll count to 80 to get more accurate data.

D6:

Null Hypothesis: **H0: p = .167**

Alternative Hypothesis: **H1: p ≠ .167**

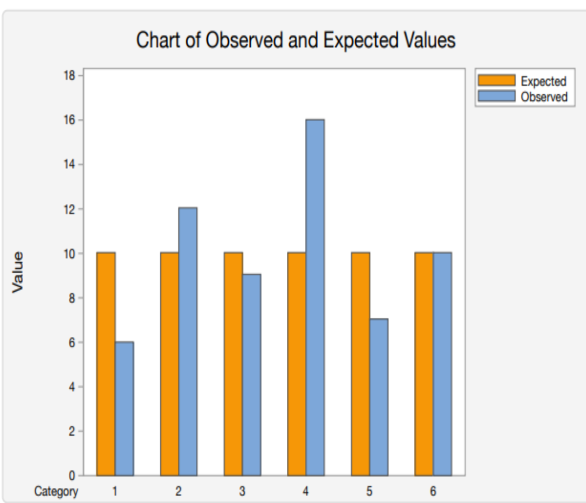
Chi-Square Goodness-of-Fit Test: Observed

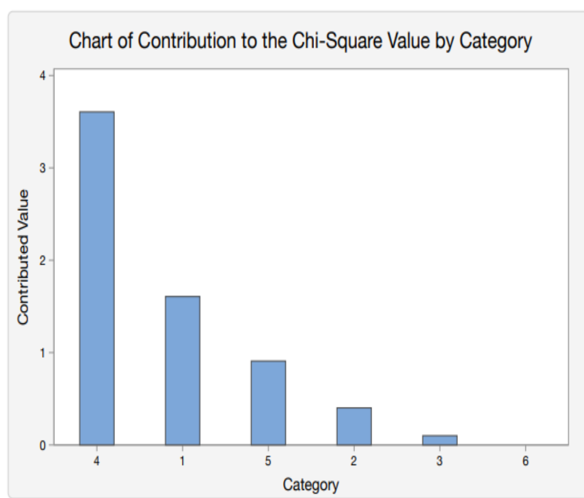
Observed and Expected Counts

Category	Observed	Test Proportion	Expected	Contribution to Chi-Sq
1	6	0.166667	10	1.60
2	12	0.166667	10	0.40
3	9	0.166667	10	0.10
4	16	0.166667	10	3.60
5	7	0.166667	10	0.90
6	10	0.166667	10	0.00

Chi-Square Test

N	DF	Chi-Sq	P-Value
60	5	6.60	0.2521





The Chi Square Value was 6.60 and we ended up getting a 0.2521 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.2521 > 0.025, we cannot reject the null hypothesis that the probability of landing on one of the sides is .167. The 6-sided die had one of the lower P-values compared to the other sided die. Most people would assume that the 6-sided die would be the fairest because it’s the most common die, but it seems that certain sides are more favorable. For example, it seems that this 6-sided landed on 4 very frequently.

D8:

Null Hypothesis: **H0: p = .125**

Alternative Hypothesis: **H1: p ≠ .125**

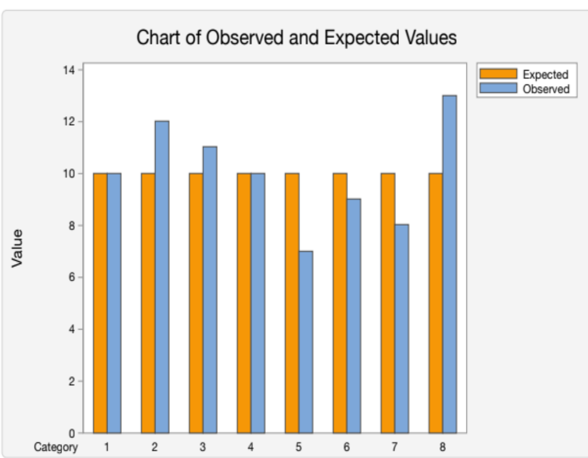
Chi-Square Goodness-of-Fit Test: Observed

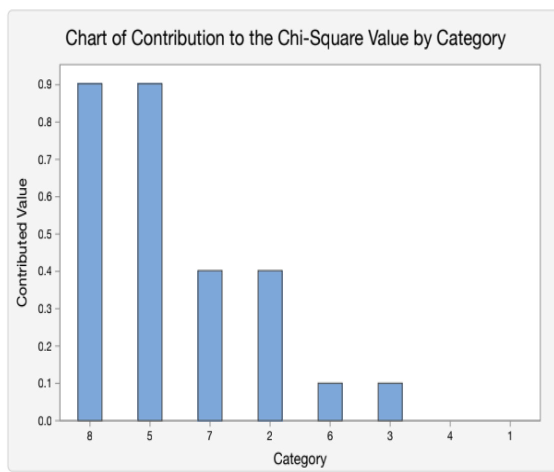
Observed and Expected Counts

Category	Observed	Test Proportion	Expected	Contribution to Chi-Sq
1	10	0.125000	10	0.00
2	12	0.125000	10	0.40
3	11	0.125000	10	0.10
4	10	0.125000	10	0.00
5	7	0.125000	10	0.90
6	9	0.125000	10	0.10
7	8	0.125000	10	0.40
8	13	0.125000	10	0.90

Chi-Square Test

N	DF	Chi-Sq	P-Value
80	7	2.80	0.9029





The Chi Square Value was 2.80 and we ended up getting a 0.9029 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.9029 > 0.025, we cannot reject the null hypothesis that the probability of landing on one of the sides is .125.This was one of the worst results out of all the testing. The P-Value was high since the Chi Square value is very low. This die seems to be fair since most of the rolls were close to the expected value or slightly higher/lower. It was funny and surprising to see that the D8 looked fairer than the D6.

D10:

Null Hypothesis: **H0: p = .10**

Alternative Hypothesis: **H1: p ≠ .10**

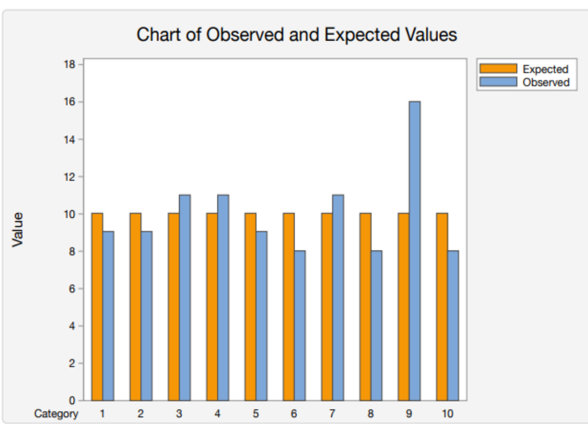
Chi-Square Goodness-of-Fit Test: Observed

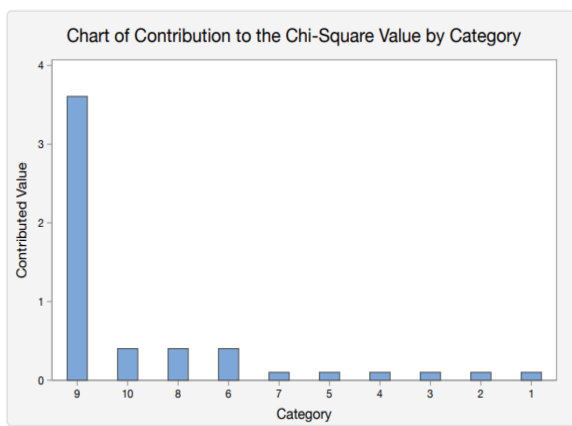
Observed and Expected Counts

Category	Observed	Test Proportion	Expected	Contribution to Chi-Sq
1	9	0.100000	10	0.10
2	9	0.100000	10	0.10
3	11	0.100000	10	0.10
4	11	0.100000	10	0.10
5	9	0.100000	10	0.10
6	8	0.100000	10	0.40
7	11	0.100000	10	0.10
8	8	0.100000	10	0.40
9	16	0.100000	10	3.60
10	8	0.100000	10	0.40

Chi-Square Test

N	DF	Chi-Sq	P-Value
100	9	5.40	0.7981





The Chi Square Value was 5.40 and we ended up getting a 0.7981 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.7981 > 0.025, we cannot reject the null hypothesis that the probability of landing on one of the sides is .10. Although the results weren't significant for the D10 test. We can conclude by looking at the data that the 9 side seems to be favored compared to all the other sides. All the other sides are relatively close to the expected and we can only assume that the 9 side might be shaped or weighted differently.

D12:

Null Hypothesis: **H0: p = .083**

Alternative Hypothesis: **H1: p ≠ .083**

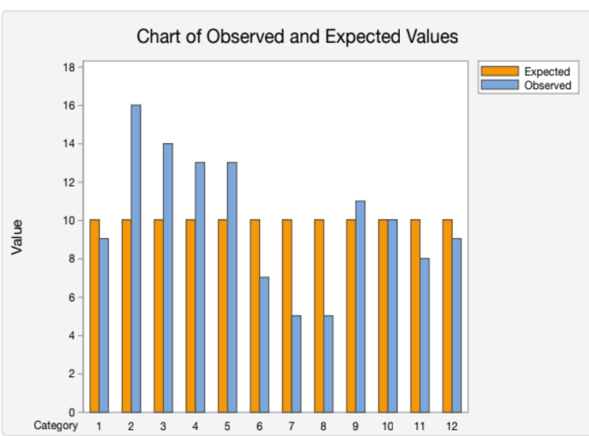
Chi-Square Goodness-of-Fit Test: Observed

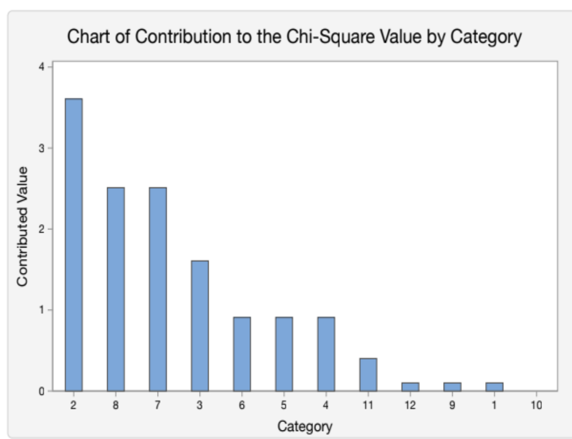
Observed and Expected Counts

Category	Observed	Test Proportion	Expected	Contribution to Chi-Sq
1	9	0.083333	10	0.10
2	16	0.083333	10	3.60
3	14	0.083333	10	1.60
4	13	0.083333	10	0.90
5	13	0.083333	10	0.90
6	7	0.083333	10	0.90
7	5	0.083333	10	2.50
8	5	0.083333	10	2.50
9	11	0.083333	10	0.10
10	10	0.083333	10	0.00
11	8	0.083333	10	0.40
12	9	0.083333	10	0.10

Chi-Square Test

N	DF	Chi-Sq	P-Value
120	11	13.60	0.2559





The Chi Square Value was 13.60 and we ended up getting a 0.2559 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.2559 > 0.025, we cannot reject the null hypothesis that the probability of landing on one of the sides is .083. This was also one of the better results for the Chi Square Tests. Although we still can't reject the null hypothesis, we can see from the data that certain sides of the die are favored while other sides are not. One can argue that the side with the 2-5 is slight weighted or the edges are shaped slightly different than usual causing this sort of results in our data.

D20:

Null Hypothesis: **H0: p = .05**

Alternative Hypothesis: **H1: p ≠ .05**

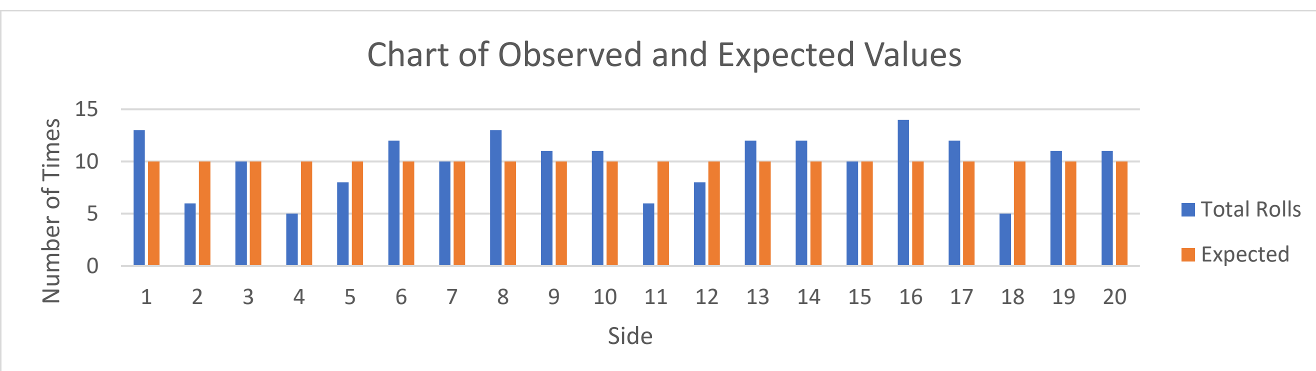
D 20

Side	Observed	Test Proportion	Expected	Contribution to Chi-Sq
1	13	0.05	10	$\frac{(13-10)^2}{10} = 0.9$
2	6	0.05	10	$\frac{(6-10)^2}{10} = 1.6$
3	10	0.05	10	$\frac{(10-10)^2}{10} = 0$
4	5	0.05	10	$\frac{(5-10)^2}{10} = 2.5$
5	8	0.05	10	$\frac{(8-10)^2}{10} = 0.4$
6	12	0.05	10	$\frac{(12-10)^2}{10} = 0.4$
7	10	0.05	10	$\frac{(10-10)^2}{10} = 0$
8	13	0.05	10	$\frac{(13-10)^2}{10} = 0.9$
9	11	0.05	10	$\frac{(11-10)^2}{10} = 0.1$
10	11	0.05	10	$\frac{(11-10)^2}{10} = 0.1$
11	6	0.05	10	$\frac{(6-10)^2}{10} = 1.6$
12	8	0.05	10	$\frac{(8-10)^2}{10} = 0.4$
13	12	0.05	10	$\frac{(12-10)^2}{10} = 0.4$
14	12	0.05	10	$\frac{(12-10)^2}{10} = 0.4$
15	10	0.05	10	$\frac{(10-10)^2}{10} = 0$
16	14	0.05	10	$\frac{(14-10)^2}{10} = 1.6$
17	12	0.05	10	$\frac{(12-10)^2}{10} = 0.4$
18	5	0.05	10	$\frac{(5-10)^2}{10} = 2.5$
19	11	0.05	10	$\frac{(11-10)^2}{10} = 0.1$
20	11	0.05	10	$\frac{(11-10)^2}{10} = 0.1$

Chi Square value = 0.9 + 1.6 + 0 + 2.5 + 0.4 + 0.4 + 0 + 0.9 + 0.1 + 0.1 + 1.6 + 0.4 + 0.4 + 0 + 1.6 + 0.4 + 2.5 + 0.1 + 0.1 = 14.4

df = 20 − 1 = 19

P value = 0.90 (11.651) > P-value (14.4) > 0.75 (14.562)



The Chi Square Value was 14.40 and we ended up getting a 0.7599 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.2559 > 0.025, we cannot reject the null hypothesis that the probability of landing on one of the sides is .05. Although it doesn't seem that the dice becomes unfair, I observed that one side of the die was landed on more frequently. 6,8,10 and 14 are all right next to each other and were landed on more than the expected. We can infer that the edge on that side is not precise of its weighted on that side of the die.

<u>Tabletop:</u>

Null Hypothesis: **H0: μ = 3.5**

Alternative Hypothesis **H1: μ ≠ 3.5**

1-Sample Z: data1

Descriptive Statistics

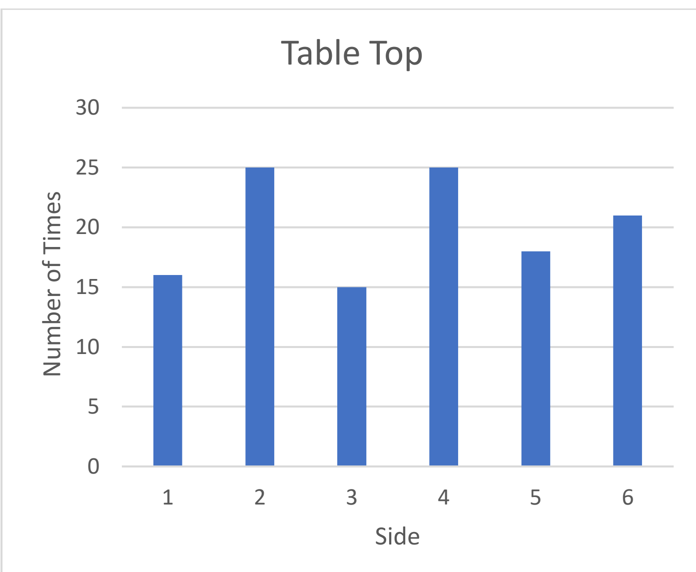
N	Mean	StDev	SE Mean	95% CI for μ
120	3.5583	1.6843	0.1500	(3.2644, 3.8523)

μ: mean of data1
Known standard deviation = 1.643

Test

Null hypothesis	H_0: μ = 3.5
Alternative hypothesis	H_1: μ ≠ 3.5

Z-Value	P-Value
0.39	0.6973



$$Z = \frac{3.5583 - 3.50}{\left(\frac{1.6843}{\sqrt{120}}\right)} = 0.39$$

$$P \text{ value} = 1 - 2(P(z > 0.39))$$
$$= 1 - 2(.1517)$$
$$= 1 - .3034$$
$$= .6966$$

The Z Value was 0.39 and we ended up getting a 0.6973 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.6973 > 0.025, we cannot reject the null hypothesis that the die is fair with a mean of 3.5. The mean was very close to the expected, so it wasn't surprising that the P-value would be high. This is also the most common surface used when playing board games, besides the actual game board itself, so you would hope that this surface wouldn't cause the die to become unfair.

Counter Top (Granite):

Null Hypothesis: **H0: μ = 3.5**

Alternative Hypothesis **H1: μ ≠ 3.5**

1-Sample Z: data1

Descriptive Statistics

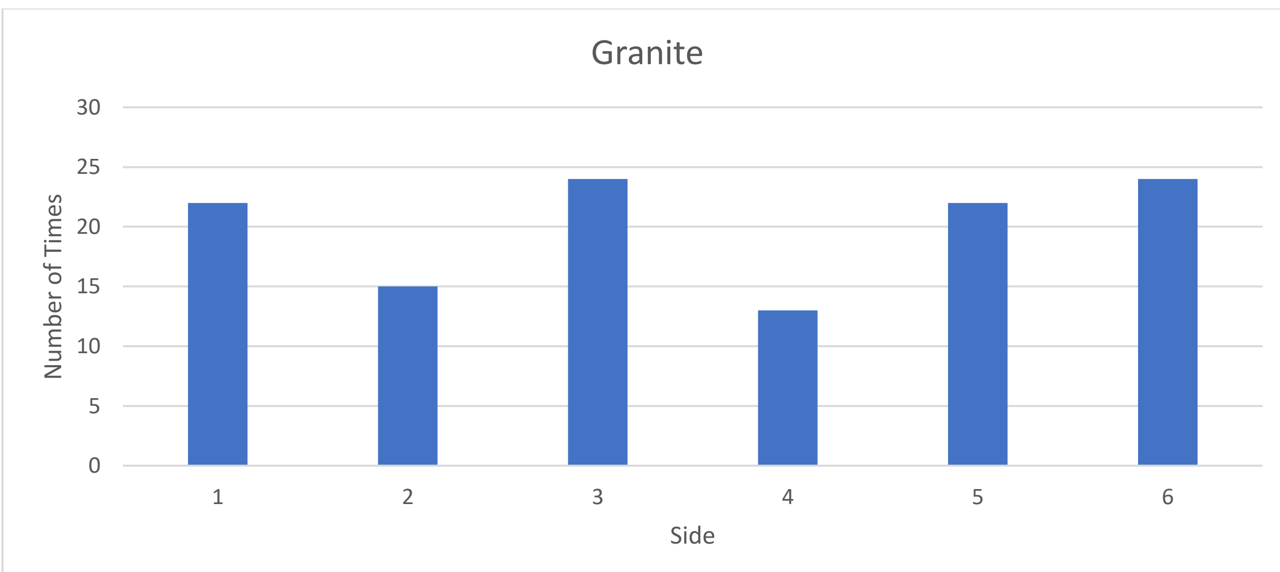
N	Mean	StDev	SE Mean	95% CI for μ
120	3.5833	1.7850	0.1629	(3.2640, 3.9027)

μ: mean of data1
Known standard deviation = 1.785

Test

Null hypothesis	H_0: μ = 3.5
Alternative hypothesis	H_1: μ ≠ 3.5

Z-Value	P-Value
0.51	0.6091



The Z Value was 0.51 and we ended up getting a 0.6091 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.6091 > 0.025, we cannot reject the null hypothesis that the die is fair with a mean of 3.5. This is a type of surface that would cause a lot of bouncing around and could potentially have made the die look unfair. I thought that with the results from the Chi Square Test on the 6-sided die and this surface would cause some bias. However, I was wrong, and this surface really didn't tinker with the results.

<u>Carpet:</u>

Null Hypothesis: **H0: μ = 3.5**

Alternative Hypothesis **H1: μ ≠ 3.5**

1-Sample Z: data1

Descriptive Statistics

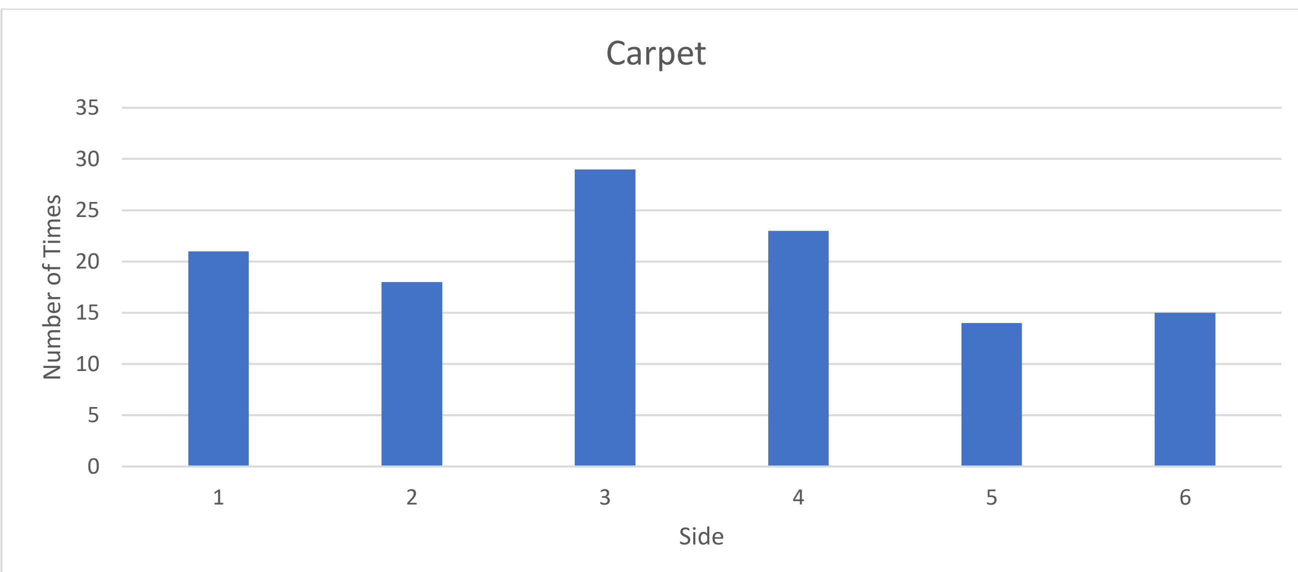
N	Mean	StDev	SE Mean	95% CI for μ
120	3.3000	1.6015	0.1462	(3.0135, 3.5865)

μ: mean of data1
Known standard deviation = 1.6015

Test

Null hypothesis	H_0: $\mu = 3.5$
Alternative hypothesis	H_1: $\mu \neq 3.5$

Z-Value	P-Value
-1.37	0.1713



The Z Value was -1.37 and we ended up getting a 0.1713 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.1713 > 0.025, we cannot reject the null hypothesis that the die is fair with a mean of 3.5. This test was close to rejecting the null hypothesis but couldn't conclude that. While collecting data for this test, I realized that the carpet prevented the die from bouncing/rolling around. Sometimes people play on the carpet because it's a convenient spot to gather around but one might not get the results they are seeking.

<u>Box:</u>

Null Hypothesis: **H0: μ = 3.5**

Alternative Hypothesis **H1: μ ≠ 3.5**

1-Sample Z: data1

Descriptive Statistics

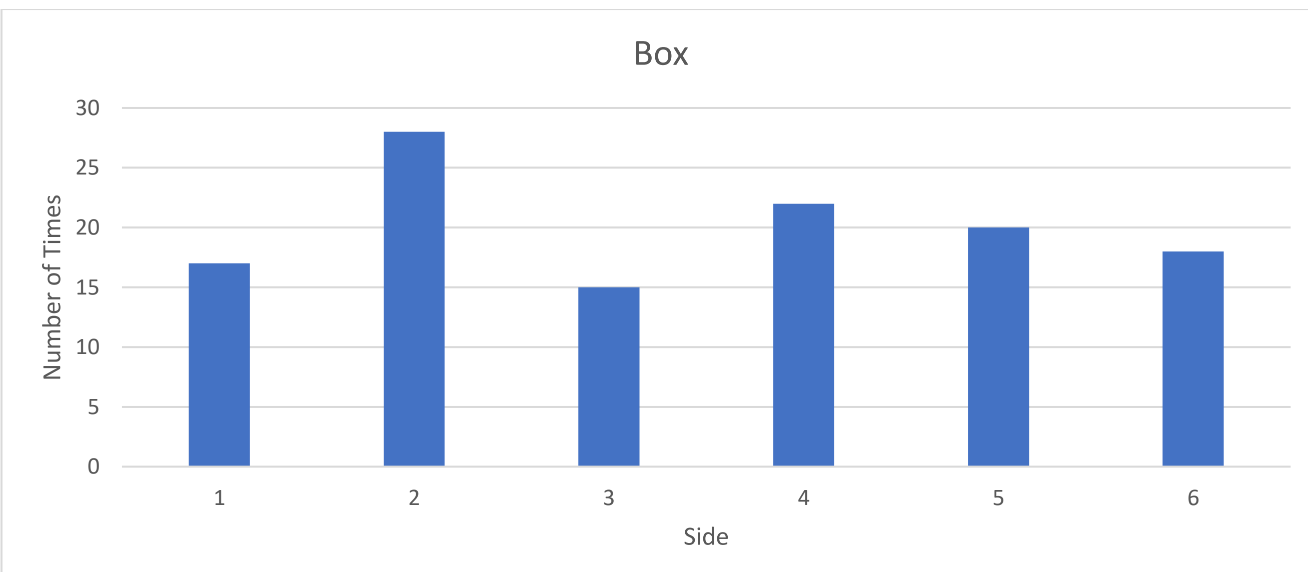
N	Mean	StDev	SE Mean	95% CI for μ
120	3.4500	1.6796	0.1533	(3.1495, 3.7505)

μ: mean of data1
Known standard deviation = 1.6796

Test

Null hypothesis	H_0: μ = 3.5
Alternative hypothesis	H_1: μ ≠ 3.5

Z-Value	P-Value
-0.33	0.7443



The Z Value was -0.33 and we ended up getting a 0.7443 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.7443 > 0.025, we cannot reject the null hypothesis that the die is fair with a mean of 3.5. I assumed that rolling dice in the game box would result in similar results with the carpet since these surfaces cause the die to not roll. However, that was not the case. It seems that rolling in a box is a suitable choice when playing a board game to make sure that the dice don’t roll too far from you.

Game Board:

Null Hypothesis: **H0: μ = 3.5**

Alternative Hypothesis **H1: μ ≠ 3.5**

1-Sample Z: data1

Descriptive Statistics

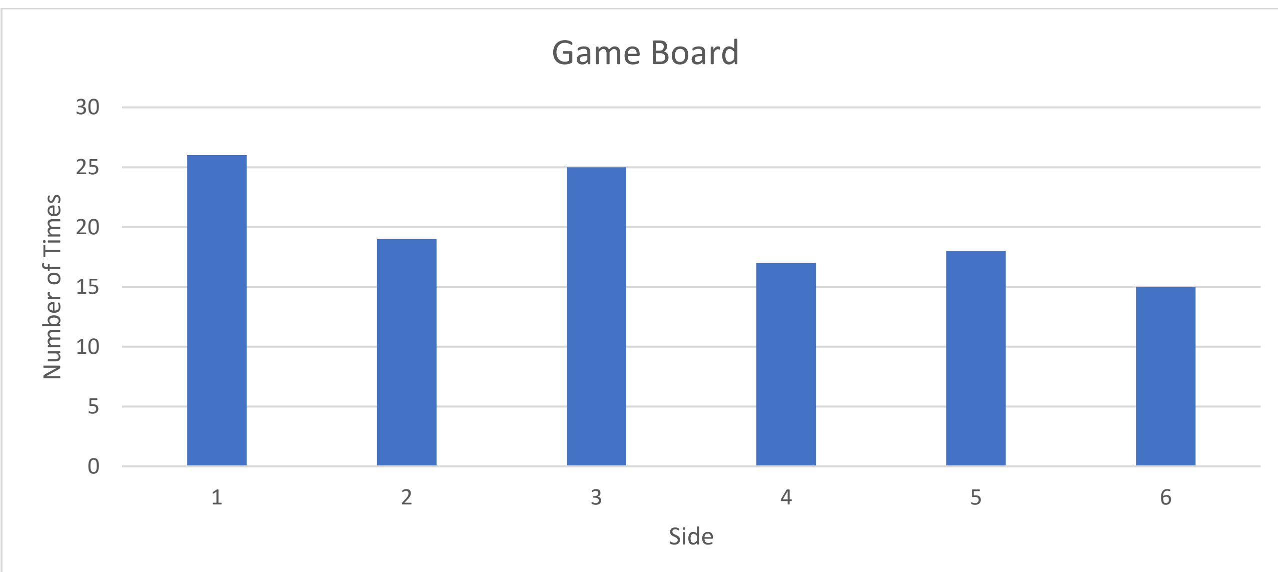
N	Mean	StDev	SE Mean	95% CI for μ
120	3.2250	1.6926	0.1545	(2.9222, 3.5278)

μ: mean of data1
Known standard deviation = 1.6926

Test

Null hypothesis	H_0: μ = 3.5
Alternative hypothesis	H_1: μ ≠ 3.5

Z-Value	P-Value
-1.78	0.0751



The Z Value was -1.78 and we ended up getting a 0.0751 P-value. The alpha level we are using is 0.05 so the alpha divided by two is 0.025. Since 0.0751 > 0.025, we cannot reject the null hypothesis that the die is fair with a mean of 3.5. This was the most surprising results since most try to roll their dice on the game board. I wanted to be fair by using a 0.05 alpha level but if I were to test with an alpha level of 0.10 then we could reject the null hypothesis. With these results, I can suggest others to not roll their dice on the game board since it could cause the dice to be unfair.

Statistically Significant

Statistically significant means that there is something other than randomness or luck that is influencing the results of our test. Depending on what alpha level used, there will be a different percent of confidence that a statistically significant result occurs. This is helpful to know so that further tests can be ran to found out what causes the change you are looking for. If I get a statistically significant result in my hypotheses test, it will show that some dice or not fair due to what characteristics I may be testing (such as # of sides or Surfaces).

Conclusion:

I was unable to reject any of the null hypotheses, but I was able to make some assumption and observations with the data I collected. For the testing on the number of sides, the results showed that the dice were fair. However, some dice have sides that could weigh more or are just not manufactured to perfection to be a fair dice. All the different sided dice had a favored side or number was rolled on much more than the expected. This could be due to n not being large enough (n usually must be around 1000 to get conclusive results) to look for the results we were looking for. It seems that as the number of sides on a die increases, there is not an increased likelihood of the die being unfair. For the different surfaces, I was close to rejecting the null hypothesis for two out of the five surfaces. The game board failed a 95% confidence test, but I can say with 80% confidence that rolling dice on a game board could cause them to be unfair. It was shocking to see that the game board, which is specifically made for people to roll dice on, could cause such unfairness to a game. I almost rejected the null hypothesis for the carpet as well but that is more justifiable because the dice don't roll as much on it. This could cause the dice to just fall flat on a number without any tumbling or rolling which should happen. Although I got such results with a sample size of 120, it would be interesting to see what the results turn out with a larger n. I cannot conclude that any of the surfaces cause dice to become unfair but there is an impact on the die depending on what surface you use. The table top seems to be a fair surface to use and can’t be blamed for dice being unfair when it might seem to be. Overall, all dice seem to be manufactured to be fair but not to the perfection they are supposed to be. This could be due to the mass production of these dice causing the quality to drop. However, these are all questions/assumptions that can’t be answered with the tests I have done. People can’t blame the dice all the time because maybe it’s just not their lucky day.