Homework #2

Group Members: Marjan Rezvani & Kieran Yuen 9/28/2020

```
setwd("/Users/kieranyuen/Documents/Econometrics")
```

PP#1

Doing only one trial with our dice and judging it as unfair when a 6 comes up, judging it as fair when any number between 1-5 come up:

If the dice were truly fair, the probability that we would erroneously judge it as unfair is the Type I Error, which is 1/6 (~16.67%).

Oppositely, if the dice were truly unfair, the probability that it would be judged to be fair would be a Type II Error. We would not be able to assess what the Type II Error is for two reasons: 1) we are only performing one trial and 2) we do not know how much the factor that made this dice unfair (i.e. how much do 1 or 2 or 3 shaved corners really affect the result of the roll). So the best way to detect if we would make a Type II Error is by adding more trials. Adding more trials should reveal to us if we have a really unfair dice (i.e. the number 6 comes up 90% of the rolls).

PP#2

Group's decision rule: a dice would be judged fair if the resulting rolls match the expected frequencies of a dice roll, which is each side of the dice has a 1/6 chance to land. | # | expected | actual | |—|———-|——| | 1 | 5 | 5 | | 2 | 5 | 5 | | 3 | 5 | 5 | | 4 | 5 | 5 | | 5 | 5 | 5 | | 6 | 5 | 5 |

If a dice is rolled 30 times, the likely number of resulting 6's should match the probability that each of the numbers of a fair dice should come up, which is 1/6 * 30 = 5.

How unusual is it to get 1,2 or 3 more or less than the expected number of 5 rolls?

```
pr_atMost_n_6 <- function(n, total_num_trial=30){
  pr = c()
  i=0
```

```
  while (i<=n){
    pr=c(pr, choose(total_num_trial,i)*5^(total_num_trial-i)/6^total_num_trial
    print(i)
    i= i+1
    }
  return(sum(pr))
}

pr_atMost_n_6(1,30) # probability of at-most 1 six
```

```
## [1] 0
## [1] 1

## [1] 0.02948904
```

```
pr_atMost_n_6(2,30) # probability of at-most 2 six
```

```
## [1] 0
## [1] 1
## [1] 2

## [1] 0.1027904
```

```
pr_atMost_n_6(3,30) # probability of at-most 3 six
```

```
## [1] 0
## [1] 1
## [1] 2
## [1] 3

## [1] 0.2396195
```

If the dice were fair, what is the chance it could be judged as unfair? This would be a Type I Error.

(Pr(number:6:resulting) = 1 - p(number:6:not:resulting)) (= 1 - \frac{5}{6}^{30} ≈ 0.00421 ≈ 0.421%)

PP#3

Group's decision rule: any deviation away from the expected roll proportions of 1/6 for each number of the dice will mean the die could possible be judged as "unfair." But just because a dice rolls even one more or one less than expected does not mean it should be judged as unfair. There is virtually no truly 100% fair dice due to the anatomical alterations to a dice from its manufacturing process to the dice losing a microscopic piece of it's anatomy due to it being rolled on a hard surface. So assuming that there will be some degree of deviation away from the expected number of rolls, we can test whether this deviation is statistically significant by using a chi-square test for goodness of fit. This chi-square test will find the difference between our actual rolls versus the expected rolls and calculate a test statistic. The resulting p-value from this hypothesis test will tell us if our actual rolls deviated enough away from the expected to be statistically significant enough to reject the null hypothesis that the dice is fair. So our boundaries for our decisions will be set by our level of alpha which we will set at the standard level used by many in academia, 0.05. Perhaps Las Vegas regulations have set a stricter level of significance (i.e. 0.01) for their dice since there is more at stake if casinos are using weighted dice.

Designing of EP#1

Assessing the fairness of a 6-sided dice is not the easy task we initially thought it to be. Initially you assume that a dice is fair because when you throw it it spins and tumbles and it appears to be completely random. But you don't know if the dice you're throwing has been altered in someway to make certain numbers appear more often than others.

Going through the three Possible Protocols (PP) we learned that you cannot assess the fairness simply by looking for the number "6" to appear (or not appear) and that we certainly need to increase the number of trials so we can get closer to the true distribution of a 6-side dice roll, a uniform distribution. Increasing the number of trials also reduces the chance of a Type II error just in case we actually are testing a truly unfair die. We also learned that even if the actual rolls that we do don't match up perfectly to the expected uniform distribution rolls of 16.67% (1 out of 6) for each number, the dice can still avoid being labeled as an unfair dice as long as the deviation from the expected proportion of rolls is not statistically significant.

In designing our experiment protocol, we used a physical dice that we got from a board game. It is a 6-sided stamped/pitted dice. This means that the plastic material of the dice was removed/pitted to make the numbers of the dice. The size of the pits are equal among all the numbers (i.e. the 1-side's "pit" is not larger or smaller than any of the other pits on the dice). This means there is more weight on the 1-side than there is on the 2-side, which has more weight than the 3-side, all leading to the 6-side which has the least amount of weight since it has the most amount of plastic die material pitted out. This unequal weighting on each side means this is what they may call a "weighted-die" in Las Vegas. Therefore, we believe this dice may possibly land more often on the sides with more weight on it than the sides with less weight (if the difference in weighting is big enough) resulting in the larger numbers being rolled more often than the smaller numbers.

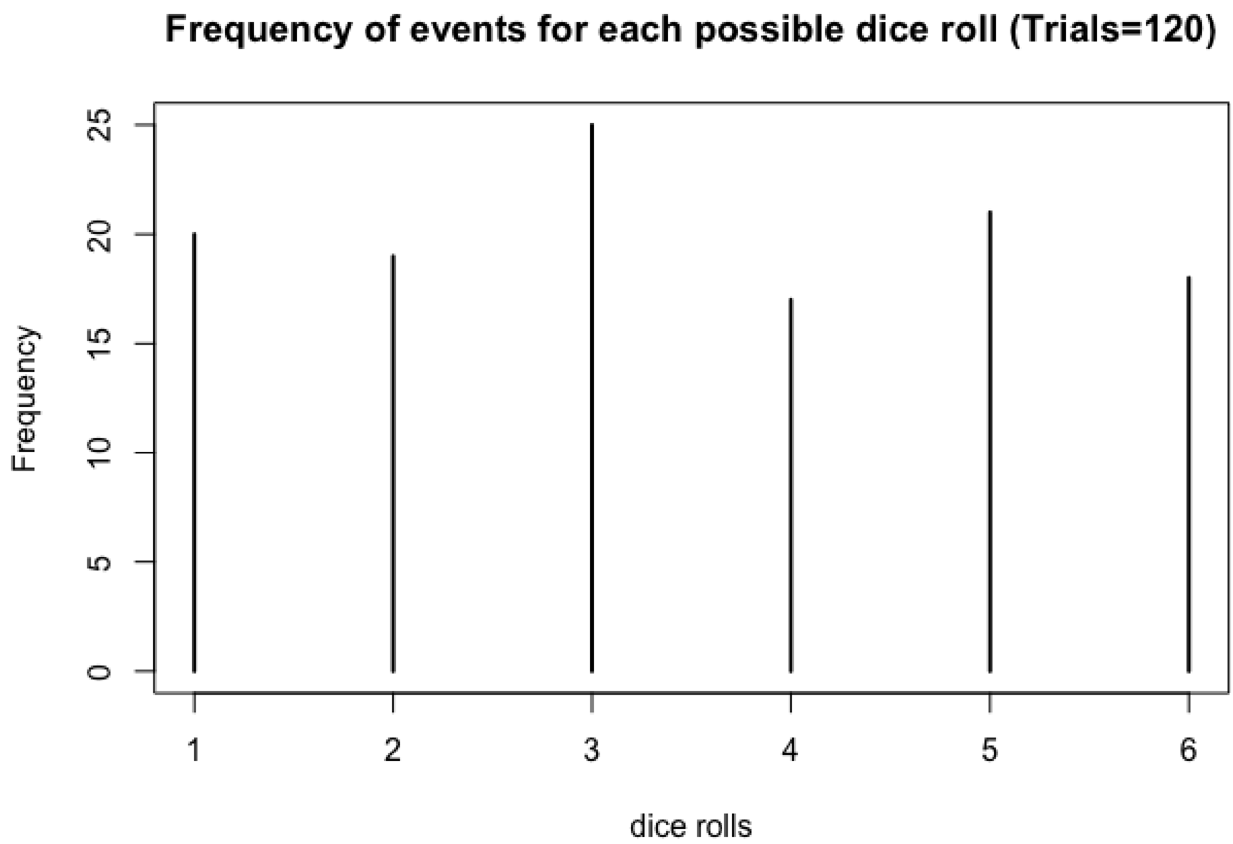
We have decided to roll our dice a number of 120 times as we believe this is a sufficient number of trials to get as close as we can to the expected uniform distribution of a 6-sided dice given the time constraints. This also allows each side of the die to come up an equal number of times compared to if we did 100 trials.

We rolled our physical die 120 times:

```
"Experiment Rolls"=c(2, 5, 4, 6, 3, 1, 2, 5, 1, 5, 2, 4, 6, 4, 4, 1, 4, 2, 6,
```

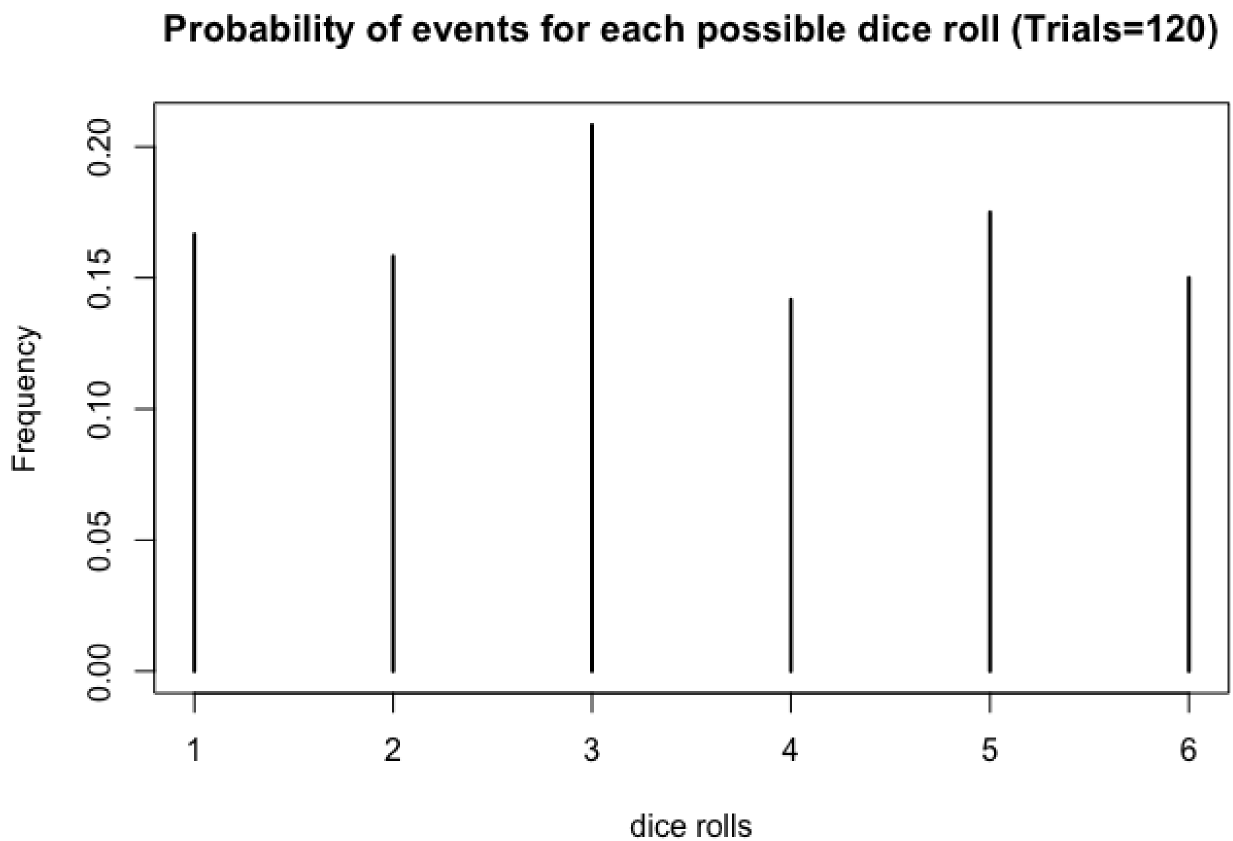
```
library(plyr)
count(`Experiment Rolls`)
```

```
##   x freq
## 1 1   20
## 2 2   19
## 3 3   25
## 4 4   17
## 5 5   21
## 6 6   18
```




```
prob <- table(`Experiment Rolls`) / length(`Experiment Rolls`)
```

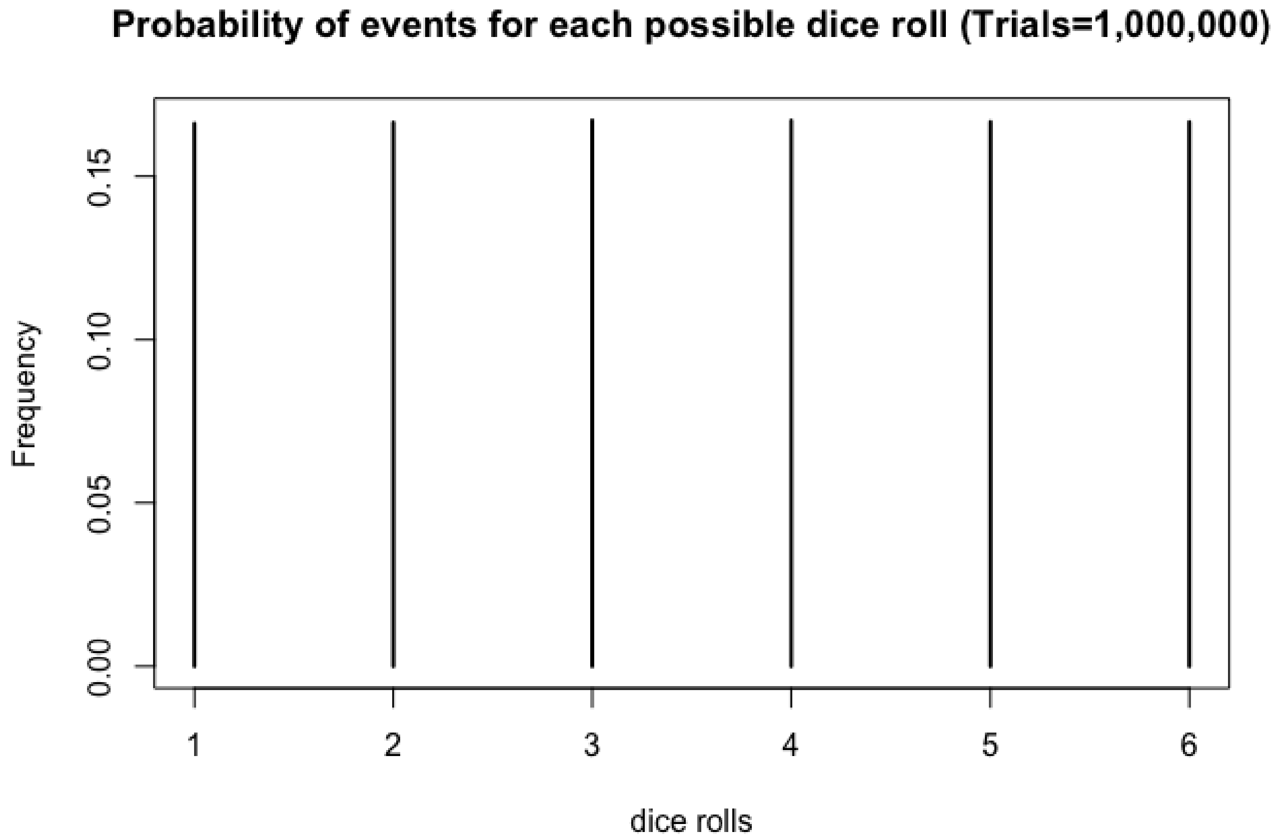
```
plot(prob,
     xlab = 'dice rolls',
     ylab = 'Frequency',
     main = 'Probability of events for each possible dice roll (Trials=120)')
```



Looking over our plots of the results above, we noticed the resulting distribution is not a "perfect" uniform distribution but perhaps if we increased our number of trials that we may get closer and closer to that "perfect" uniform distribution. Or perhaps if we reperformed this same experiment 9 more times and averaged the results to see if that would get us closer. According to the Law of Large Numbers, we would get closer.

In fact, running this quick simulation in R of 1,000,000 rolls of a 6-side die show us that the true distribution of a this random number generator in R is representative of a uniform distribution:

```
"OneMillionRolls" <- sample(1:6,1000000,replace = TRUE)
prob2 <- table(`OneMillionRolls`) / length(`OneMillionRolls`)
plot(prob2,
     xlab = 'dice rolls',
     ylab = 'Frequency',
     main = 'Probability of events for each possible dice roll (Trials=1,000,0
```



We can still assess the fairness of this dice by comparing our actual rolls to the proportion of rolls that we expected by using the Chi-Square Goodness of Fit test.

Chi square Goodness of Fit Test

```
chi2 <- chisq.test(table(`Experiment Rolls`))
chi2
```

```
##
##  Chi-squared test for given probabilities
##
## data:  table(`Experiment Rolls`)
## X-squared = 2, df = 5, p-value = 0.8491
```

The results show a p-value of 0.8491 which is much higher than the standard level of significance used of 0.05 (alpha = 0.05). The sum of our chi-square test statistic of "2" is telling us that the differences between the actual and expected values is small. In fact, looking up the critical value based on the alpha of 0.05 and 5 degrees of freedom, we get 11.070, which is much higher than our test statistic of 2. So we can be 95% confident that we should not reject the null hypothesis and that this physical dice that we performed this 1 experiment on is a “fair” dice.

These results do not reflect what we initially believed that our pitted/stamped die would come up with rolls of the higher numbers due to the difference in weight. Perhaps this is because the pits make such a slight difference that it did not show up in the Chi-square test. This is a dice that is made for a board game so it must have been manufactured to be as close to fair as the company could create. A dice found in a Las Vegas craps table is probably more fair but the difference between the two might be very minuscule.

Of course, this is said with the caveat that this experiment needs to be performed at least 99 more times so that we can see if we our estimate of 5 false positives is a good estimate. So in EP#2, we would perform this 120 trial experiment 99 more times. But since this idea sounds like the costs outweigh the benefits, we would most likely move it to a random number generator in R. We will also look into creating a simulated “weighted die” in R as it does appear to be possible based on a quick Google search.