Homework 2

Fareha Sameen

9/28/2020

`Group Members: Neshma, Stan, and Hertz`

#Possible Protocol 1 (PP1): Roll the dice, if the outcome is a 6, then the dice is not fair. If the outcome is any number besides 6, then it is fair.

#Analyze PP1: What is the probability of the dice being unfair if the dice was fair? And what is the probability of the dice being fair if it was unfair?

#Rolling the dice for the first time gave a one. Since, the outcome wasn't 6, the dice was fair.

Dice = c("1", "2", "3", "4", "5", "6")

sample(x = dice, size = 1, replace = TRUE) [1] "1"

#Then, we rolled the tampered dice to see whether the outcomes alter.

We rolled a 5 with the tampered dice to make it fair. The probability of this dice being unfair is 1/6 or 16.67%. The dice would only be unfair if the outcome was 6.

If the dice was unfair, the probability of the rolls being fair is still 5/6 or 83.3%, because although we tempered with the dice, the dice still has 6 sides, so the probability is not effected. Rolling a 6 is unfair, while rolling 1, 2, 3, 4, 5, on the same dice are fair and have the probability of 5/6 or 83.3%.

#PP2: Roll the dice 30 times. Group can specify a decision rule to judge that dice is fair or unfair. Consider the stats question: if fair dice are rolled 30 times, what is likely number of 6 resulting? How unusual is it, to get 1 more or less than that? How unusual is it, to get 2 more or less? 3? #Analyze PP2 including the question: if the dice were fair, what is the chance it could be judged as unfair?

The distribution of events are written as "number on die:outcomes". The outcomes were 1:1, 2:6, 3:6, 4:4, 5:6, 6:7.

As seen in the tampered dice with the outcome of 2:6, two rolls can't equal a six in order for it to be fair. For the PP, the dice is fair.

Running the code in R provides the same rule regarding fairness of the dice. If a fair dice is rolled 30 times, the expected value of getting a 6 would be 6/30 or 20%. It is possible for the 6 to appear more or less times than expected due to the randomness of the dice.

dice = c("1", "2", "3", "4", "5", "6") sample(x = dice, size = 30, replace = TRUE) outcomes <- c(2,1,1,6,1,5,5,1,4,2,6,3,2,2,6,3,4,2,3,3,2,2,3,4,1,2,6,6,2,3) count1 <- length(which(outcomes == 1)) 1 count2 <- length(which(outcomes == 2)) 6 count3 <- length(which(outcomes == 3)) 7 count4 <- length(which(outcomes == 4)) 5 count5 <- length(which(outcomes == 5)) 9 count6 <- length(which(outcomes == 6)) 2

We can say that a dice is unfair if the outcomes of rolling the dice are coming out to be the same number a shocking amount of times. For example, if more than 50% of the outcome of the rolls is the same number then we can assume that the dice is unfair. In this scenario, we see that our value that has the highest percentage is 2 at just about 30% of this distribution. Since it is less than the 50% or more, I can predict that this is fair.

#EP1: What is a reasonable number of times to roll your experiment dice? (given time available in class, etc) If you roll the experiment dice a certain number of times and see a particular outcome, then you would conclude it is fair or not. How confident would you be, in any of those conclusions? Note that while previous Possible Protocols emphasized counting just the number of times a 6 came up, you might consider other outcomes such as the average value of the dice when rolled or the distribution of other outcomes (what number is on opposite face from 6? Do you suppose that number might be disproportionately represented if dice were not fair?).

By rolling a dice 100 times for 5 rounds (500 times) while recording the experimental probability of rolling any given number for each sample, we will be able to tell the fairness of the dice.

We will find the average of the experimental probability of the 5 separate trials. For the average of the population, the difference shouldn't be more than 10%. The average of the experimental probability of all 5 samples should be between 15.03% and 18.37%, or between 75 and 92 for any number on the dice during the 500 rolls. sample(x = dice, size = 500, replace = TRUE) a<-sample(1:6,size=500,replace = T) print(a) [1] “4” “5” “2” “4” “4” “4” “6” “5” “3” “2” “6” “2” “5” “1” “6” “5” “3” “3” “2” [20] “6” “6” “1” “1” “1” “2” “1” “1” “2” “2” “6” “2” “5” “1” “4” “2” “5” “4” “3” [39] “4” “6” “2” “4” “4” “4” “5” “6” “1” “2” “4” “1” “3” “4” “3” “4” “5” “1” “2” [58] “5” “2” “2” “3” “4” “1” “6” “1” “2” “1” “5” “2” “4” “3” “1” “2” “5” “2” “5” [77] “4” “6” “4” “1” “2” “3” “5” “3” “4” “2” “5” “2” “1” “2” “4” “6” “2” “1” “6” [96] “4” “6” “2” “6” “5” “1” “5” “5” “3” “6” “3” “3” “6” “4” “4” “1” “6” “1” “6” [115] “2” “4” “5” “4” “6” “5” “4” “6” “1” “1” “6” “3” “1” “5” “4” “1” “1” “4” “2” [134] “4” “3” “1” “3” “3” “2” “1” “4” “5” “6” “2” “6” “4” “3” “5” “1” “1” “5” “5” [153] “6” “4” “5” “4” “1” “6” “1” “4” “6” “1” “1” “6” “1” “4” “3” “2” “4” “6” “4” [172] “2” “3” “6” “3” “4” “4” “2” “1” “4” “3” “1” “2” “4” “6” “4” “1” “2” “2” “5” [191] “4” “2” “3” “2” “4” “3” “2” “3” “3” “3” “1” “6” “3” “1” “6” “2” “2” “4” “2” [210] “5” “1” “4” “6” “4” “2” “6” “1” “6” “1” “2” “4” “2” “5” “1” “4” “3” “2” “6” [229] “2” “3” “5” “2” “4” “5” “6” “1” “1” “5” “1” “1” “3” “4” “5” “5” “1” “2” “5” [248] “4” “6” “1” “1” “2” “3” “1” “4” “1” “3” “4” “3” “3” “4” “1” “5” “5” “6” “6” [267] “3” “4” “2” “2” “2” “4” “4” “3” “2” “2” “6” “6” “2” “2” “6” “3” “4” “1” “5” [286] “3” “4” “5” “3” “2” “6” “1” “6” “5” “1” “5” “5” “6” “1” “5” “1” “5” “4” “6” [305] “1” “2” “4” “6” “1” “3” “2” “3” “4” “4” “4” “5” “1” “3” “6” “2” “6” “2” “6” [324] “1” “2” “3” “3” “6” “4” “4” “1” “4” “4” “1” “4” “4” “3” “3” “3” “5” “1” “4” [343] “1” “2” “1” “2” “4” “2” “4” “6” “4” “6” “5” “6” “1” “3” “2” “2” “6” “4” “6” [362] “3” “5” “5” “5” “3” “3” “6” “4” “1” “3” “1” “3” “4” “4” “5” “4” “1” “4” “5” [381] “6” “5” “1” “3” “2” “1” “2” “2” “5” “5” “5” “2” “2” “2” “4” “2” “6” “5” “2” [400] “5” “2” “5” “6” “2” “5” “5” “2” “3” “5” “5” “5” “5” “1” “4” “6” “1” “2” “3” [419] “2” “5” “3” “1” “1” “4” “3” “1” “4” “3” “1” “5” “5” “5” “4” “5” “4” “3” “5” [438] “3” “6” “4” “5” “3” “3” “1” “3” “3” “3” “3” “4” “2” “3” “1” “1” “1” “4” “1” [457] “4” “2” “3” “5” “6” “4” “1” “2” “6” “2” “6” “2” “1” “1” “3” “3” “4” “2” “5” [476] “6” “6” “4” “5” “4” “6” “4” “2” “3” “3” “5” “5” “5” “3” “5” “5” “4” “5” “3” [495] “1” “4” “3” “4” “2” “5”

table(a) 1 2 3 4 5 6 75 99 91 76 93 66

library(ggplot2) qplot(a,binwidth=1) mean(a) [1] 3.52 median(a) [1] 3 var(a) [1] 2.999599 sqrt(var(a)) [1] 1.731935 summary(a) Min. 1st Qu. Median Mean 3rd Qu. Max. 1.00 2.00 3.00 3.52 5.00 6.00 sd(a,na.rm=TRUE) [1] 1.731935 hist(a)

The fairness of the dice is based on not one single number appearing significantly more compared to the frequency of other values. “‘