The dice games of Chevalier de Mere

Margarita Geleta

27th february 2018

Calculating the probability of winning in each game

Game 1:

1. **Experiment:** Rolling a die four times.
2. **Event of interest:** A = *{a 6 appears}*
3. **Probability of the event:**

$$Pr(A) = \frac{6^4 - 5^4}{6^4} = \frac{671}{1296} \approx 0.5177.$$

In the Sample Space Ω we have all the possible outcomes of the experiment, which in this case are 6^4 as we roll 4 times a die with 6 possible results. As we are interested in A, we should count how many outcomes have **at least** one 6. To calculate this, it would be easier to count how many results do not have any 6 ($\bar{A}$) and make the following difference: $\Omega - \bar{A} = A$, and we obtain the set in which we are interested. We obtain the probabilty by dividing the number of favourable outcomes which is the cardinal number of A and the total number of outcomes (Ω).

Game 2:

1. **Experiment:** Rolling two dice 24 times.
2. **Event of interest:** B = *{a 6 appears in both dice}*
3. **Probability of the event:**

$$Pr(B) = \frac{36^{24} - 35^{24}}{36^{24}} \approx 0.4914.$$

In the Sample Space Ω we have all the possible outcomes of the experiment, which in this case are 36^{24} as we roll 24 times two dice with 6 possible results each one (in total, 36 possible results). As we are interested in B, we should count how many outcomes have two 6's. To calculate this, it would be easier to count how many results do not have two 6's: $36 - 1 = 35$, we exclude the outcome of getting two 6's.

With these results we can clearly see that the probability of winning in game 1 is slightly higher than in game 2.

Simulation of the games:

Game 1:

We define a function for the first experiment: *Roll* is our experiment. The size of sample is 4 as we want to roll a die four times, with possible results from 1 to 6 (1:6). Boolean expressions of TRUE and FALSE in R are interpreted as 0's and 1's, so we can compute the sum of these four rolls and stating the condition of the sum of these rolls we can count if there was *at least* one 6. And we return that boolean result.

```r
experimentA <- function(){
  roll <- sample(1:6, size = 4, replace = TRUE)
  condition <- sum(roll == 6) > 0
  return(condition)
}
```

Next, we use *replicate* to repeat that experiment N = 100 times, and we calculate the Relative frequency of A.

```
happensA <- replicate(100, experimentA())
fr.A <- sum(happensA)/length(happensA)
Pr.A <- 671/1296
```

Game 2:

We define a function for the second experiment: *first.die* and *second.die* are our dice which take part in our experiment. The size of each sample is 24 as we roll each die 24 times, with possible results from 1 to 6 (1:6). As in *game 1*, with the condition we count the sum of boolean expressions. We return it at the end of the code block.

```
experimentB <- function(){
  first.die <- sample(1:6, size = 24, replace = TRUE)
  second.die <- sample(1:6, size = 24, replace = TRUE)
  condition <- sum((first.die == second.die) & (first.die == 6)) > 0
  return(condition)
}
```

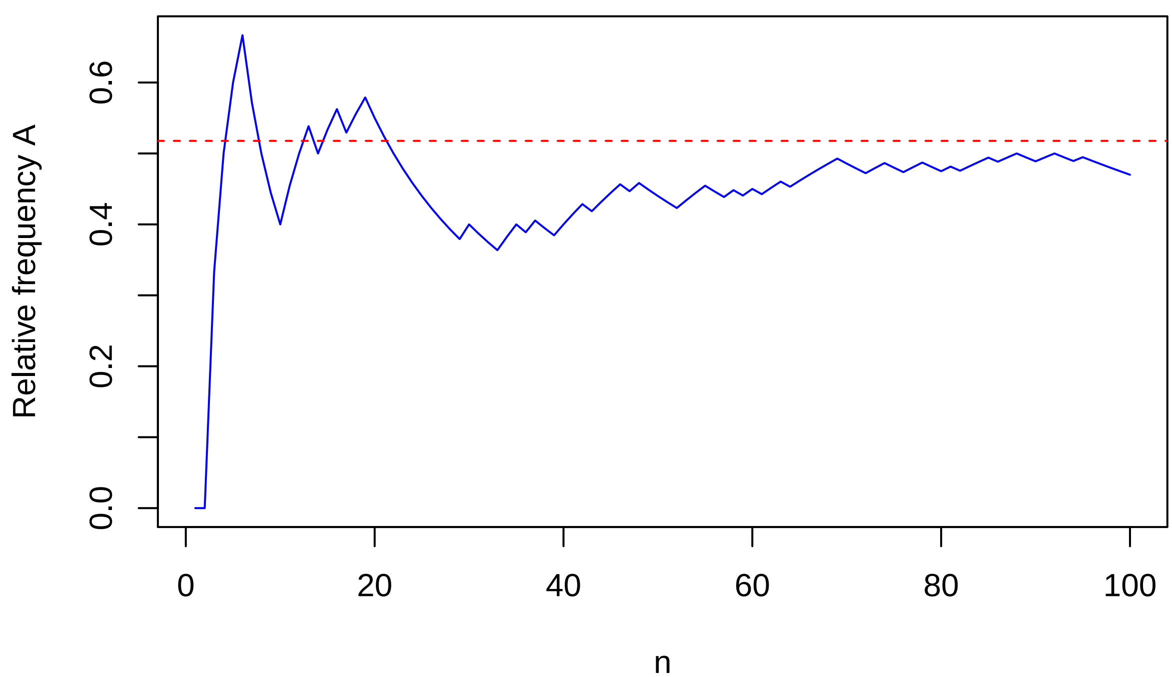
Then, we use *replicate* to repeat that experiment N = 100 times, and we calculate the Relative frequency of B.

```
happensB <- replicate(100, experimentB())
fr.B <- sum(happensB)/length(happensB)
Pr.B <- 0.4914038761
```

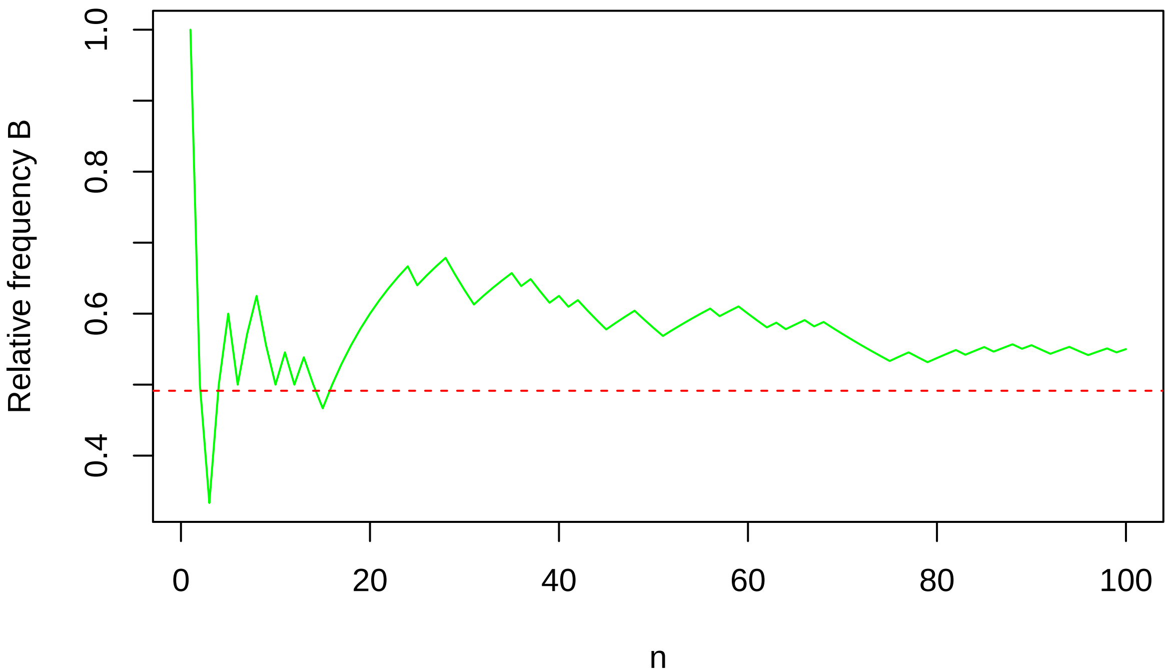
Graphing the evolution:

Let's see the evolution of the relative frequencies as n increases:

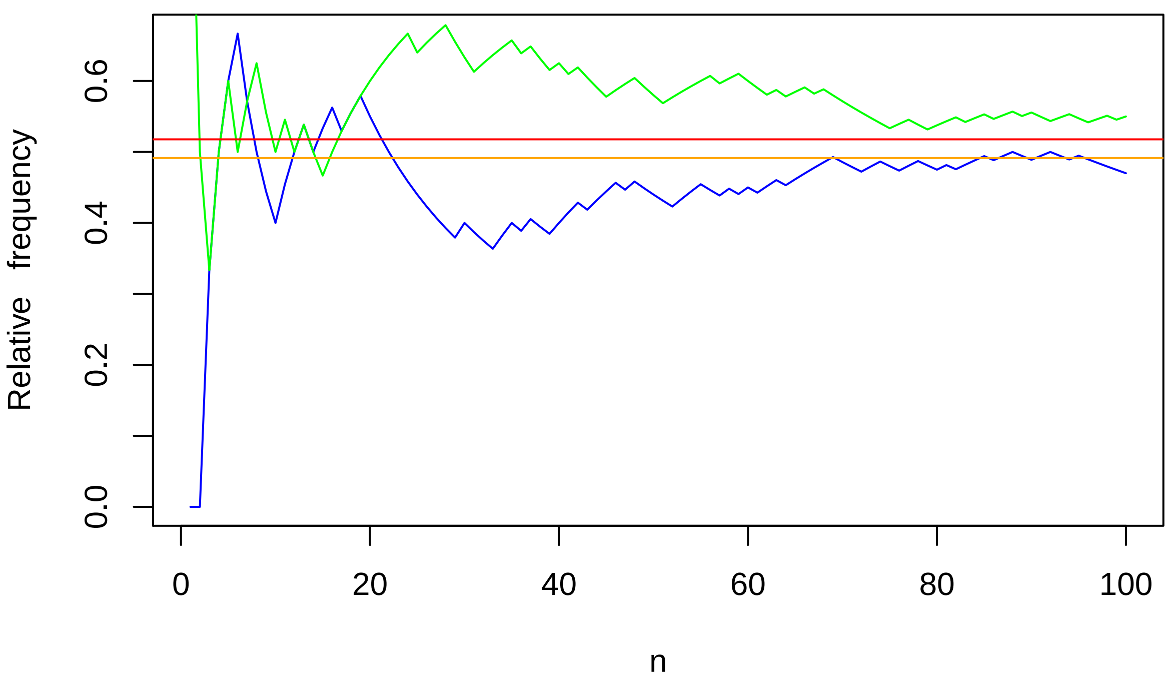
```
## Game 1
n <- 1:length(happensA)
cum.fr.A <- cumsum(happensA)/n
plot(1:length(happensA), cum.fr.A, type = "l",col = 'blue', xlab = "n", ylab = "Relative frequency A")
abline(h = 671/1296, col = 'red', lty = 2)
```



```
## Game 2
n <- 1:length(happensB)
cum.fr.B <- cumsum(happensB)/n
plot(1:length(happensB), cum.fr.B, type = "l",col = 'green', xlab = "n", ylab = "Relative frequency B")
abline(h = 0.4914038761, col = 'red', lty = 2)
```

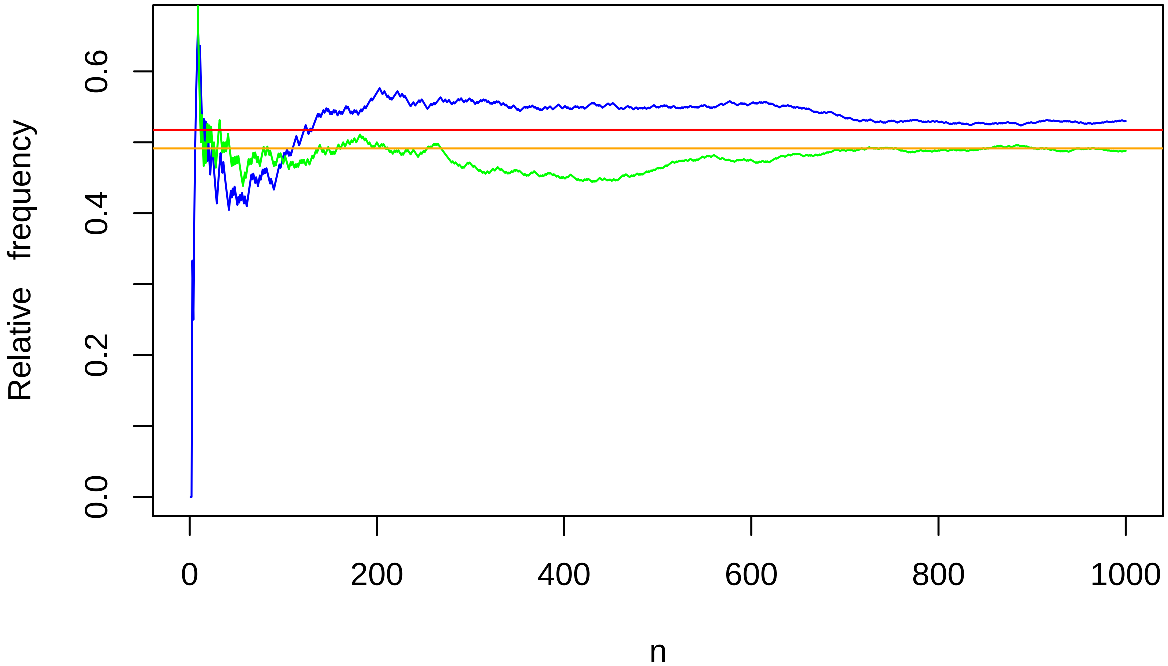


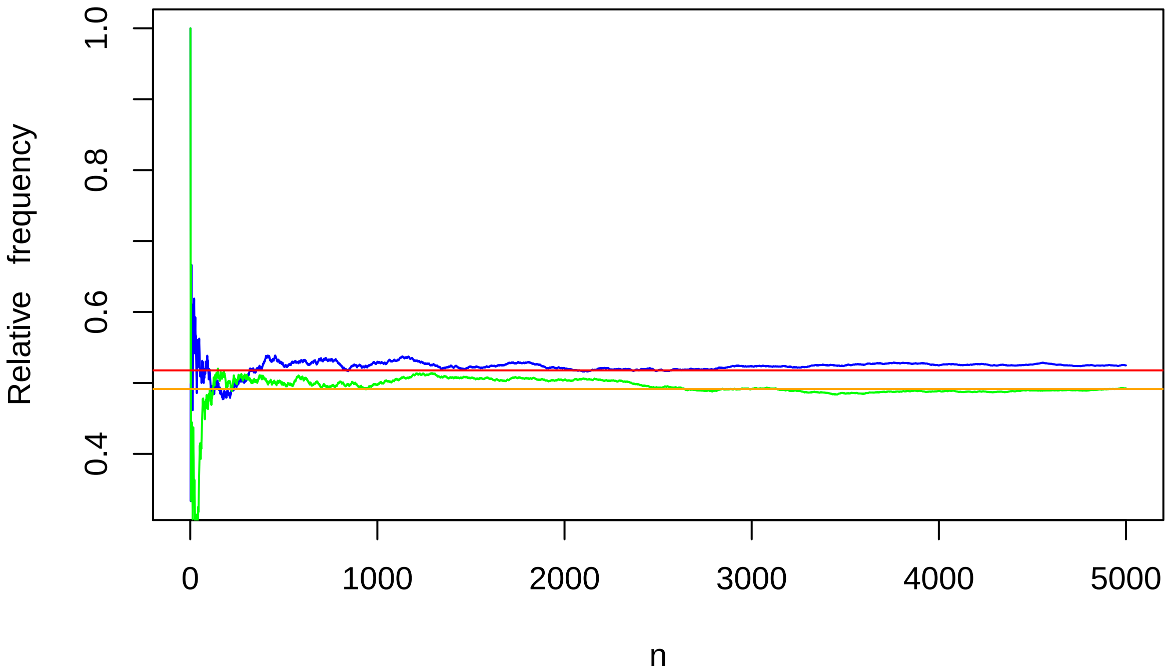
How many times should we play in order to see visually that the probability in winning in A is higher than in B? To see that, we are going to graph everything in one plot. The blue line is the Relative frequency of A, the green one - of B, the red line shows the probability of the event A and the orange line, the probabilty of the event B.



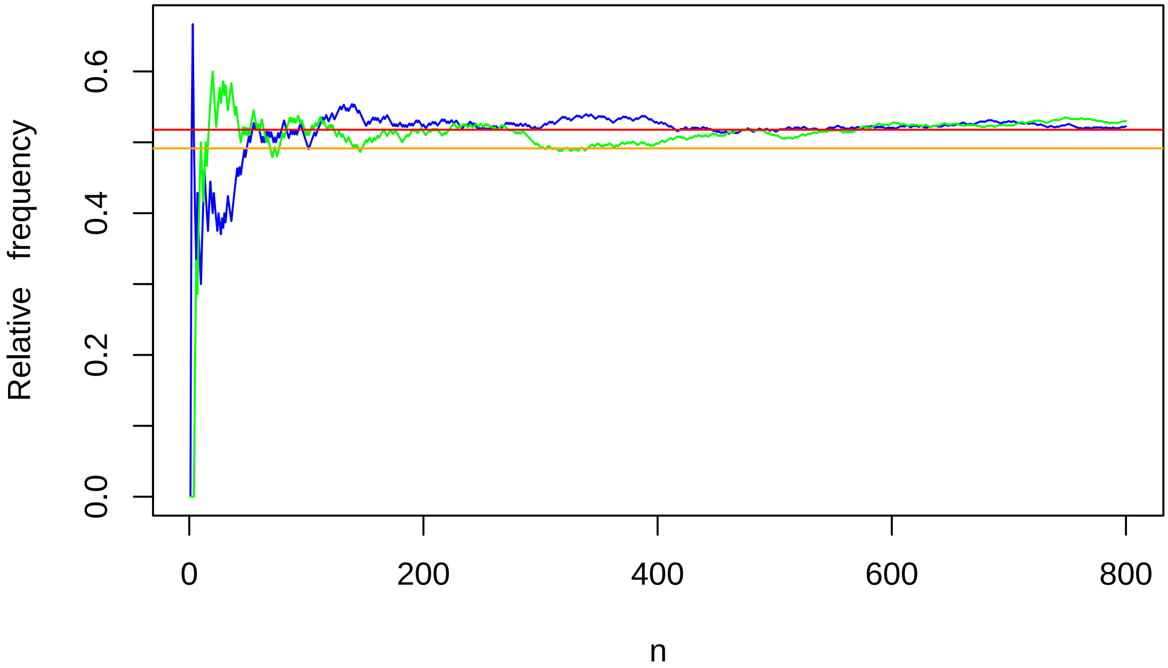
We can see that with N = 100 it is not clear whether A or B is more favourable. Tries with N = 1000 and N

= 5000:





For both values, just looking at the plots we can be certain that the probability of A is higher than of B. Yet, for instance, with N = 500 it is not clear yet. Let's try a value between N = 500 and N = 1000:



We see that while we do not reach N = 1000, sometimes we cannot be sure whether A or B is more favourable. So, Chevalier de Mere must have seen more than 1000 rounds to conclude that A is favourable than B.