

Exercise

Data Science

Statistics and Probabilities

JProf. Dr. Claudia Wagner

- Do first babies arrive late?
- Work with data from NSFG survey to explore this question
- **Answer the open questions as homework**

- In repeated independent tests with the same actual probability p of a particular outcome in each test, the chance that the fraction of times that outcome occurs differs from p converges to zero as the number of trials goes to infinity

Europa



USA



- Simulate the process of a fair roulette wheel
- We always bet 1 EUR on the same number
- Expected return after 100 spins?

```
100 spins of Fair Roulette  
Expected return betting 13 = -28.0%
```

```
100 spins of Fair Roulette  
Expected return betting 13 = 80.0%
```

```
100 spins of Fair Roulette  
Expected return betting 13 = 8.0%
```



- Expected return after 1.000.000 spins?

1000000 spins of Fair Roulette
Expected return betting 13 = 1.2968%

1000000 spins of Fair Roulette
Expected return betting 13 = 0.0836%

1000000 spins of Fair Roulette
Expected return betting 13 = 0.062%

- What if the Roulette has two additional zeros?



- On August 18, 1913, at the casino in Monte Carlo, black came up a record 26 times in succession⁵
- Everyone wanted to use this chance since the probability of observing such as sequence is $0,5^{26} = 0,00000014$
- What is the probability of red when previous 25 rolls were black?

ROLL DIE

$$X=(S,P)$$

S is a finite set of values

P: $S \rightarrow [0,1]$, whereby $\sum_s P(s) = 1$

Example:



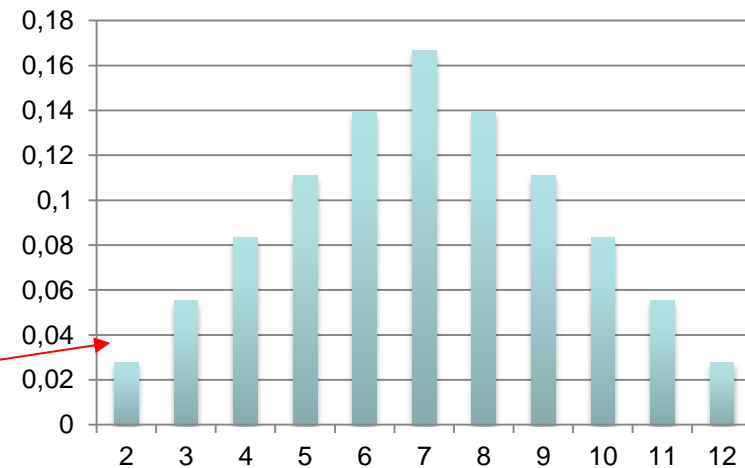
$$S=\{2,3,4,5,6,7,8,9,10,11,12\}$$

What does this mean?

Probability Mass Function



Two Dices



Implement roll dice function



What is the probability of getting two and five SIX in a row?

- Give exact solution and empirical estimate

Note: When **events are independent** of each other, the probability of all of the events occurring is equal to a **product of the probabilities** of each of the events occurring.

When are events independent?

Actual **probability of 66** = 0.027777778

100 trials:

- Estimated Probability of 66 = 0.02

10000 trials:

- Estimated Probability of 66 = 0.029

Actual **probability of 66666** = 0.0001286

10000 trials and 100 trials:

- Estimated Probability of 66666 = 0.0

- It takes a lot of trials to get a good estimate of the frequency of occurrence of a rare event.
- One should not confuse the sample probability with the actual probability
- There was really no need to do this by simulation, since there is a perfectly good closed form answer.
- But simulations are often useful when we do not have a closed form answer



- An ongoing process where the next state might depend on both the previous states and some random element
- It is not deterministic
 - ◆ You use seed function to make it deterministic for debugging
- The rollDie() function has a random component

- What is the probability of observing **3 SIX** when rolling 4 dices (**no order**)?

Number of ways to choose an (unordered) subset of k elements from a set of n elements

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- $4 * 1/6^3 * 5/6^1 = 0.015$



What is the probability of observing **3 SIX in a row** when rolling a dice 4 times (**order**)?

- $2 * (1/6)^3 * (5/6) = 0.0077$
- **$n-k+1$** ways to select **k consecutive items** out of **n items** where $n > k$
- $4-3+1 = 2$
- ***666 and 666***

- What is the probability of observing **2 SIX in a row** when rolling a dice 4 times (**order**)?
- $3 * (1/6)^2 * (5/6)^2$
- $n-m+1 = 4-2+1=3$
- $*22*$, $22**$, $**22$

- What is the probability of observing **2 SIX in a row** when rolling a dice 4 times (**no order**)?
- $4!/2!*2! = 6$
- $6 * (1/6)^2 * (5/6)^2$

QUESTIONS