Esssentials of Applied Data Analysis
IPSA-USP Summer School 2017

Handout - The Basics of Probability Theory - I

Leonardo Sangali Barone
leonardo.barone@usp.br

jan/17

Introduction to Probability - Part I

Basic Notions of probability, part I.

Probabilty

Probability is a formal model of uncertainty.

1. Objetive probability:
    - Classical - theory driven. Ex: dice or coin.
    - Empirical - observation driven. Ex: voting for winner in presidential election.
2. Subjective probability - belief driven. Ex: educated guess about the happening of an event.

Coins, dices, cards and legislators.

Let's see how probability works for coins, dices, cards and legislators in this handout.

- Toss a coin. What is the probability of getting a head?
- Toss a coin. What is the probability of getting a head?

$$P(Head) = P(Tail) = 1/2$$

How do we know it without actually tossing a coin?

First Definitions

- *Sample space (S)*: the set of all possible outcomes.
- *Outcome*: is an element of the sample space.
- *Event*: any collection of possible outcomes.
- The empty set $\varnothing$ and the sample space S are also events.
- *Probability of event A*:

$$P(A) = \frac{\text{Number of outcomes in event A}}{\text{Number of outcomes in the sample space}}$$

Axioms and theorems of probability (1)

- For every event A, $0 \leq P(A) \leq 1$
- $P(S) = 1$
- $P(\varnothing) = 0$

Coin

Toss a coin. What is the probability of getting a head?

Sample Space = $\{Head, Tail\}$ – 2 possible outcomes

Event A = $\{Head\}$ – 1 possible outcome

$$P(A) = \frac{\text{Number of outcomes in event A}}{\text{Number of outcomes in the sample space}} = \frac{1}{2}$$

Dice - quick exercise

Roll a 6-side dice.

What is the probability of getting a 5?

$$P(5) = ?$$

What is the probability of getting an even number?

$$P(even) = ?$$

What is the probability of getting a prime number?

$$P(prime) = ?$$

Dice - quick exercise - answers

Roll a 6-side dice.

What is the probability of getting a 5?

$$P(5) = 1/6$$

What is the probability of getting an even number?

$$P(even) = \frac{\#\{2,4,6\}}{6} = \frac{3}{6} = \frac{1}{2}$$

What is the probability of getting a prime number?

$$P(prime) = \frac{\#\{2,3,5\}}{6} = \frac{3}{6} = \frac{1}{2}$$

Random legislator - quick exercise

Choose a Legislative House of your choice, in any country/state/province/city in the world. Choose a political party and call it Party A. (A nice and not-up-to-date visualization of Brazilian Câmara dos Deputados can be found here. Let's get a random legislator from that House.

What is the probability of getting a legislator from the Party A?

$$P(\text{party A}) = ?$$

What is the probability of getting a woman?

$$P(woman) = ?$$

Question - classical or empirical?

What is the difference between the dice and the random legislator examples? Did we have to actually roll the dices to get calculate the probabilities of getting a 5, an even or prime number? Can I guess the probability of choosing at random a woman or a legislator from Party A without observing and counting legislators?

Let's get frustrated a little bit. Let's toss a coin 10 times (or toss 10 coins) and check how many heads we get. There "should" be 5 heads, rigth?

Now let's do it 100 times. And 1000. And 1000000.

```
*install new ado (aka package w/ new functions) using the command findit
findit heads

* toss 10 a fair coin
heads, flips(10)
```

See Figure 1

```
* toss 100 a fair coin
heads, flips(100)
```

See Figure 2

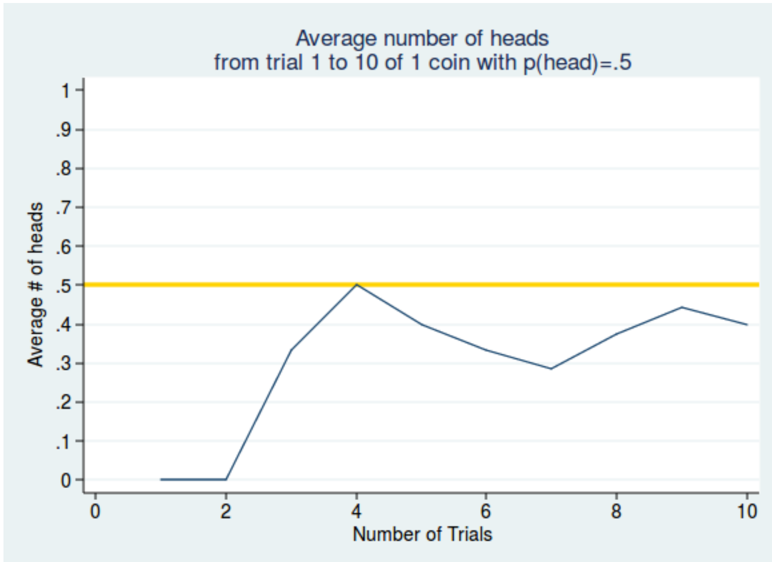

Figure 1: Simulation w/ fair coin - n = 10

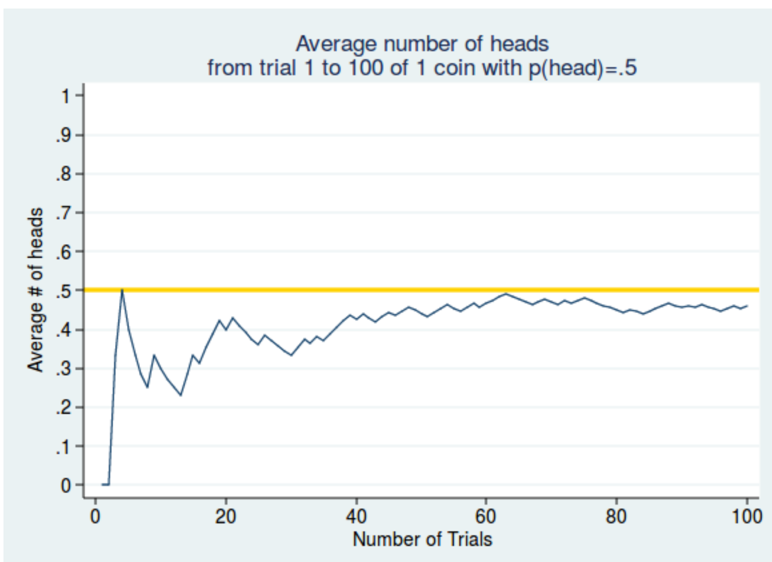

Figure 2: Simulation w/ fair coin - n = 100

```
* toss 1000 a fair coin
heads, flips(1000)
```

See Figure 3

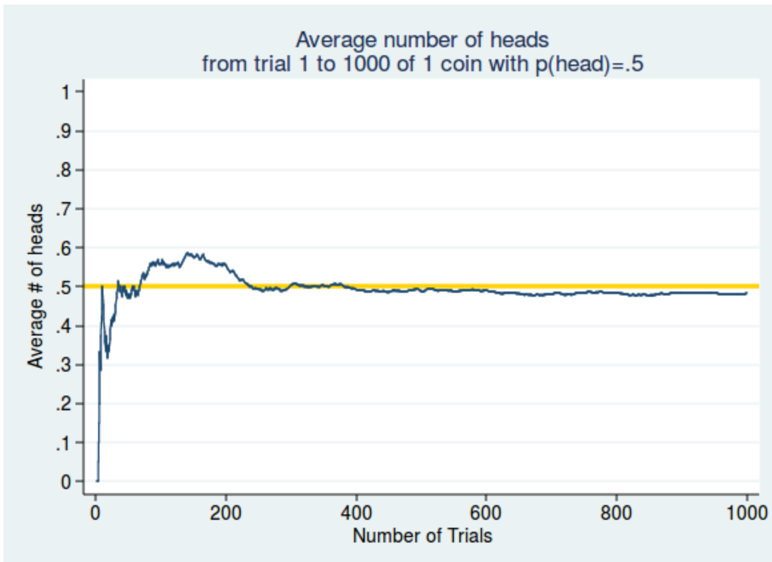

Figure 3: Simulation w/ fair coin - n = 1000

Try yourself with 1000000 coin tosses:

```
* toss 1000000 a fair coin
heads, flips(1000000)
```

What if we use a biased coin?

```
* toss 100 a biased coin w/ P(Head) = 0.3
heads, flips(100) prob(.3)
```

See Figure 4

```
* toss 1000 a biased coin w/ P(Head) = 0.3
heads, flips(1000) prob(.3)
```

See Figure 5

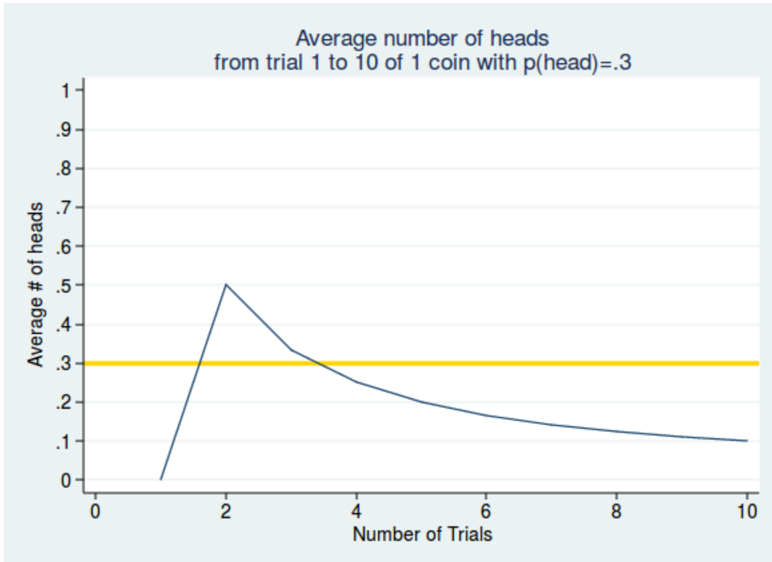

Figure 4: Simulation w/ biased coin - n = 10

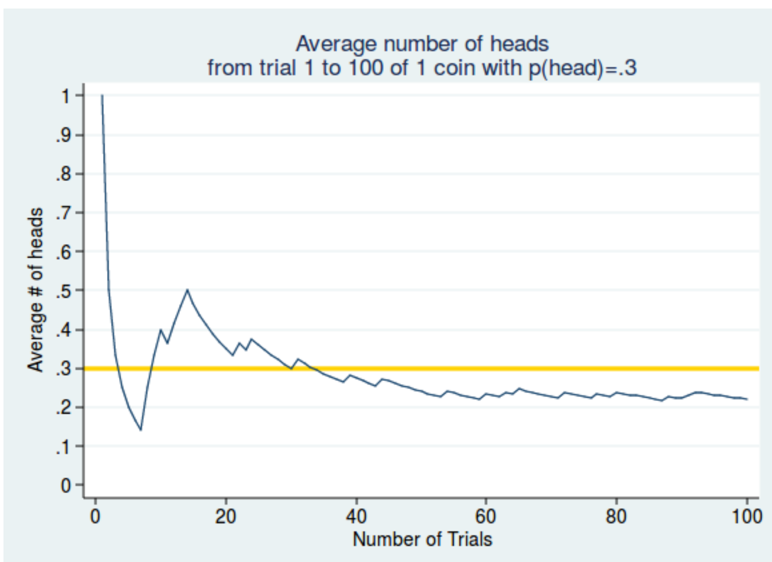

Figure 5: Simulation w/ biased coin - n = 100