Probability

Polychrome marble statue depicting Tyche holding the infant Plutus in her arms, 2nd century CE, Istanbul Archaeological Museum

SML201: Introduction to Data Science, Spring 2020

Michael Guerzhoy

Quantifying uncertainty

- Events:
  - The roll of a die: {1, 2, 3, 4, 5, 6}
  - Outcome of a coin flip: {H, T}
- Any of the following are events:
  - I rolled 2
  - I rolled an even number
  - I rolled one of 3 or 5
  - The coin came up heads
- If we know how dice and coins work, we can assign probabilities to events

Probability

- For a fair coin, what is the probability of heads?
  - 50%
  - Two equally likely outcomes, the probability of each is 50%
- What is the probability of rolling 1 using a fair die?
  - 1/6
- What is the probability of rolling an even number or a 5 using a fair die?
  - {2, 4, 6, 5}
  - 4/6

What *is* probability?

- Long-term relative frequency (LTRF): if I toss a coin a large number of times, it will come up heads half the time
  - 10000012/ 2000000 is approximately 0.5
  - P(H) = ½
- What is the probability that Trump will win in Nov 2020?
  - More difficult to think in term of LTRF
    - Can’t repeat 2020 2000000 times!
  - Can still elicit a subjective probability: how much would you be willing to bet that the even will happen?
  - More of a philosophical issue. It’s (according to some philosophers) fine to proceed with a naïve and inconsistent understanding of probability

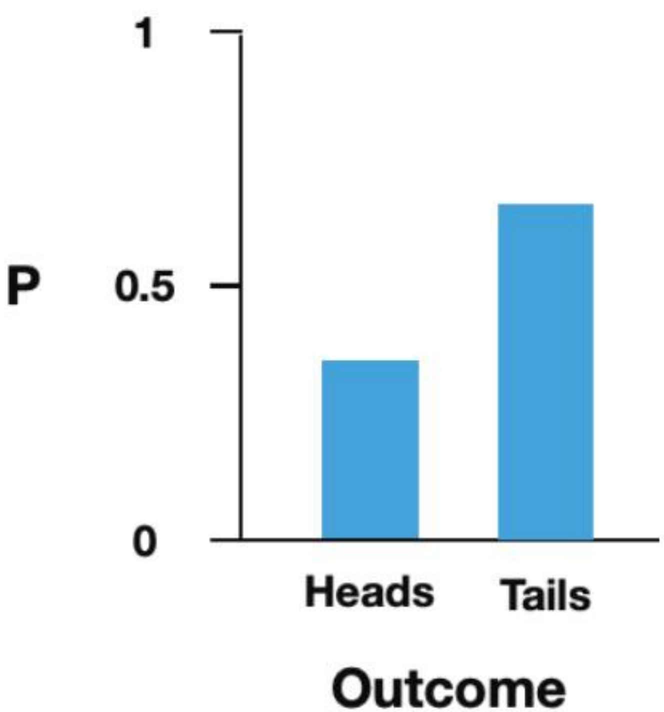
What *is* Probability?

- A number between 0 and 1 that indicates how likely an event is
- If we have an event space that covers all the possible outcomes, the probability of at least one of the possible outcomes happening is 1
- P(event occurs) + P(event does not occur) = 1

(Switch to R)

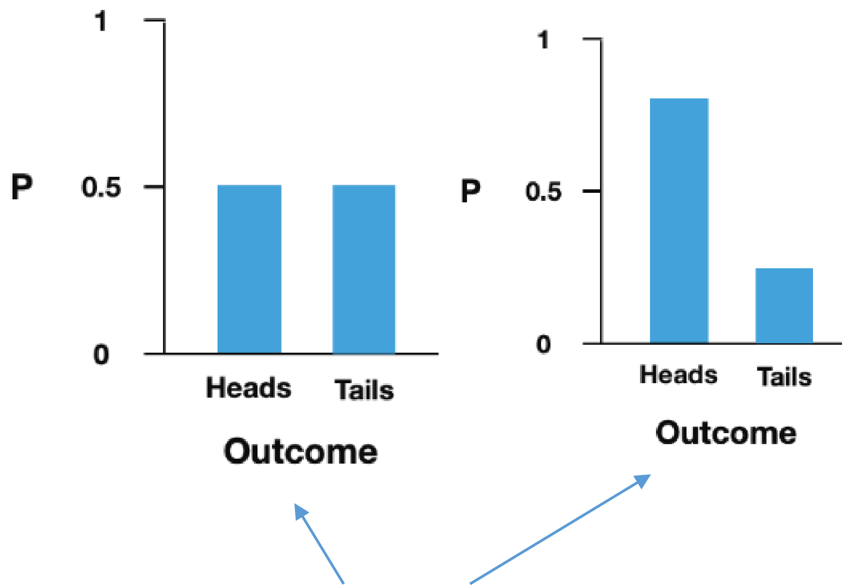
Probability Mass Function

- Probability Mass Function (PMF): computes the probabilities of each



- $0 \leq P \leq 1$
- $P(E_1) + P(E_2) + \cdots + P(E_k) = 1$

  If the events are non-overlapping and at least of them has to happen



different coins

Bernoulli Random Variable

- The outcome of a coin toss for a (possibly) biased coin
- P(H) = p, P(T) = 1-p
- Fair coin: p = 0.5
- Weighted coin: $p \neq 0.5$

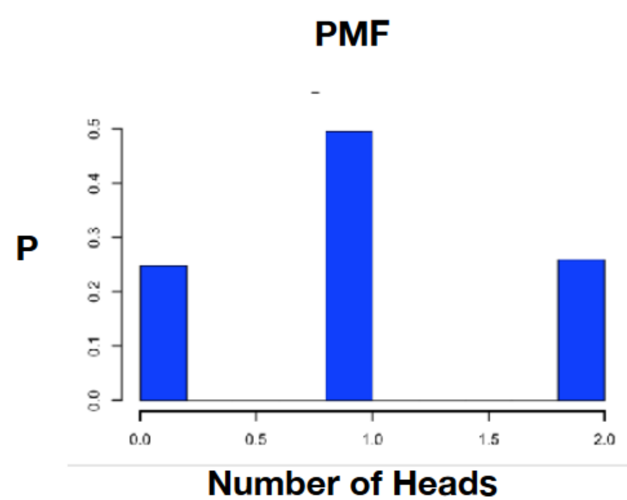
Switch to R

Binomial Random Variable

- *size* coin tosses, with a weighted coin with P(H) = p. The outcome is the number of times the coin came up heads
- For a fair coin, and two tosses

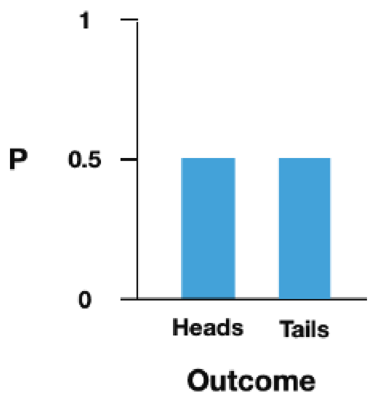
N = 2

Number of Heads	Probability
0	1/4
1	1/2
2	1/4

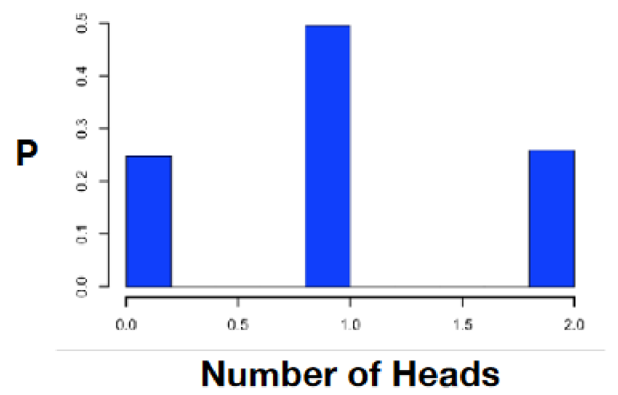


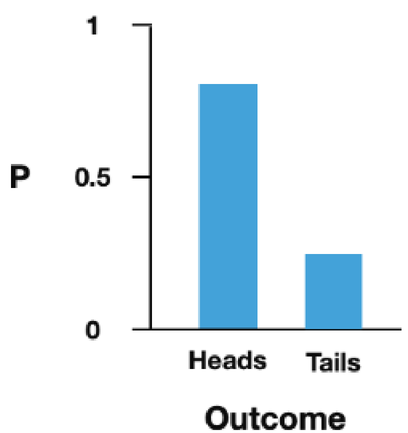
Num of tosses = 2

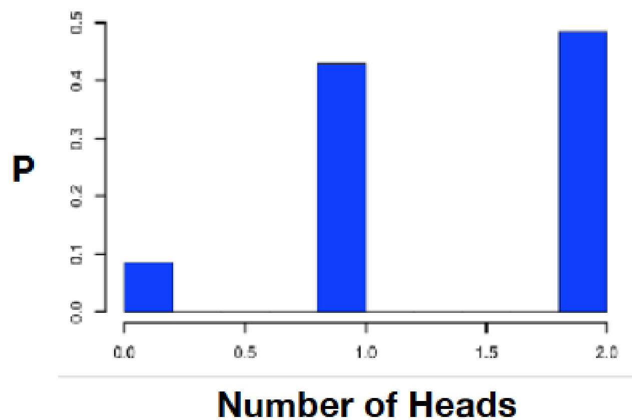
Bernoulli



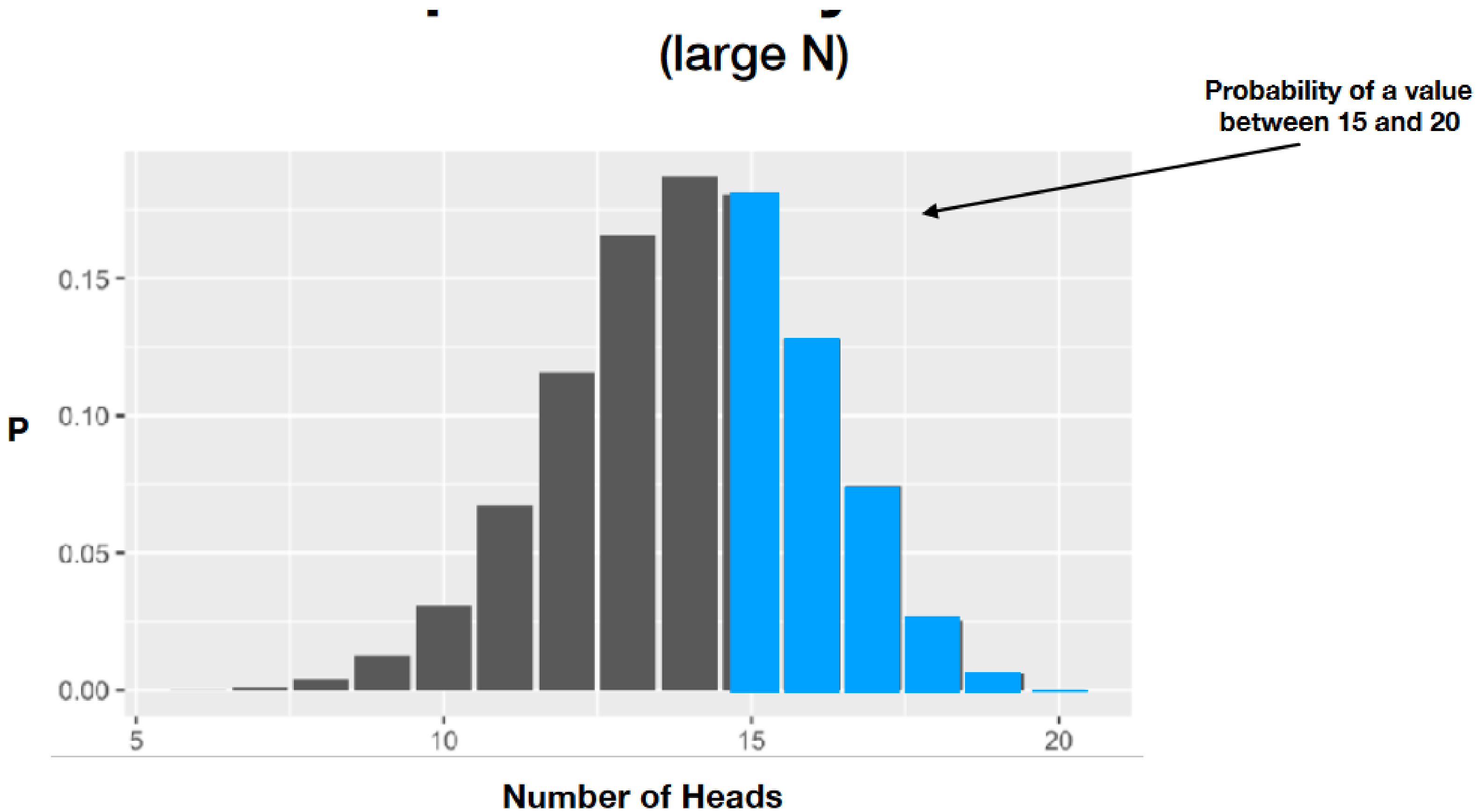
Binomial







Probabilities of events, again

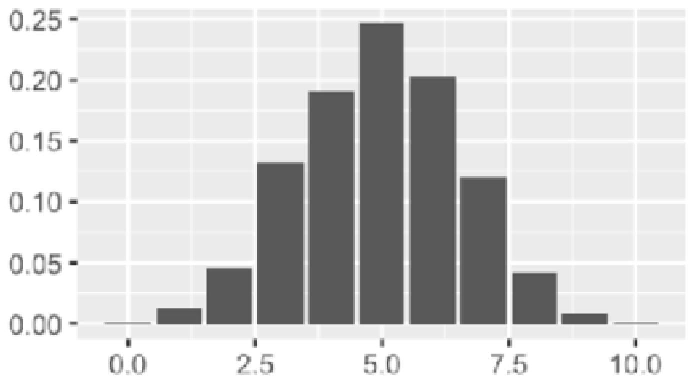


Can sum the probabilities of individual outcomes to get the probability of a range of outcomes

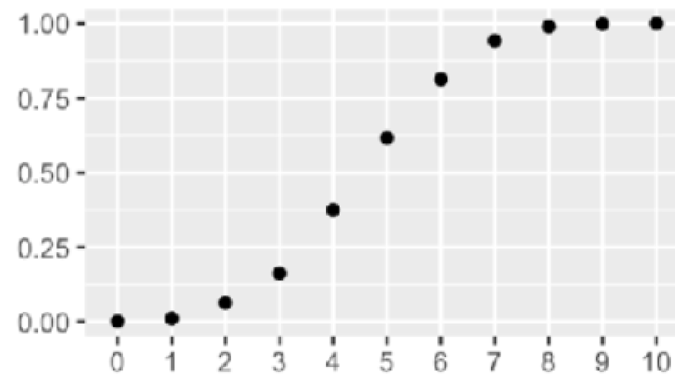
Cumulative Mass Function

- The CMF computes the probability that the outcome is *at most* x

PMF



CMF

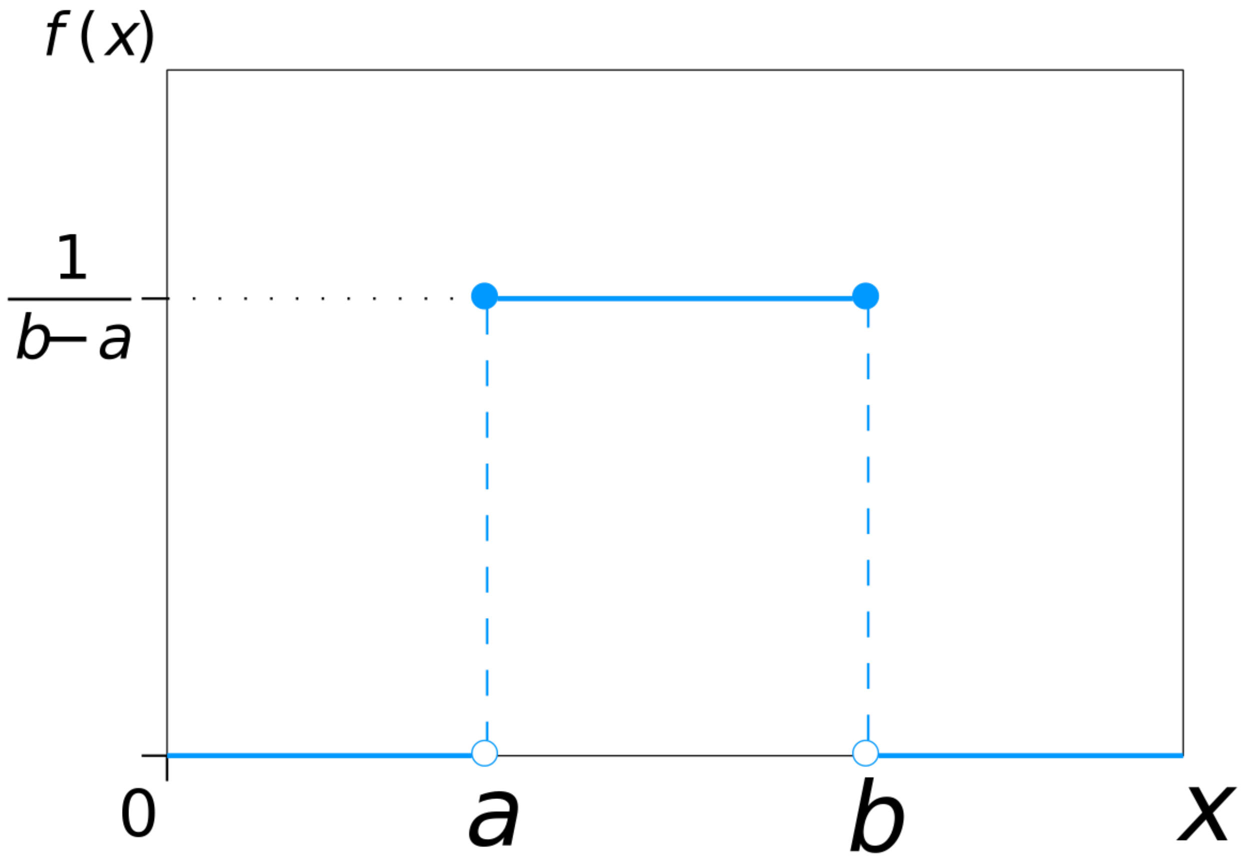


- CMF(x) = PMF(0)+PMF(1)+….+PMF(x)
- CMF(max(x)) = 1
- The CMF is always non-decreasing
  - Why?
- CMF(x): $P(X \leq x)$

Continuous distributions

- Suppose we randomly pick a real number between 0 and 1, with each real number being equally likely
  - Philosophical question: how would you even *do* that?
- There are so many real numbers that the probability of picking any particular one is literally zero
  - But the probability of picking a number between 0.123 and 0.124 is not zero – it's 0.01
    - Why?
- So we can't have a PMF (all the probabilities are zero). But we can have a Probability Density Function, which works somewhat the same way

Uniform distribution on [a, b]: PDF



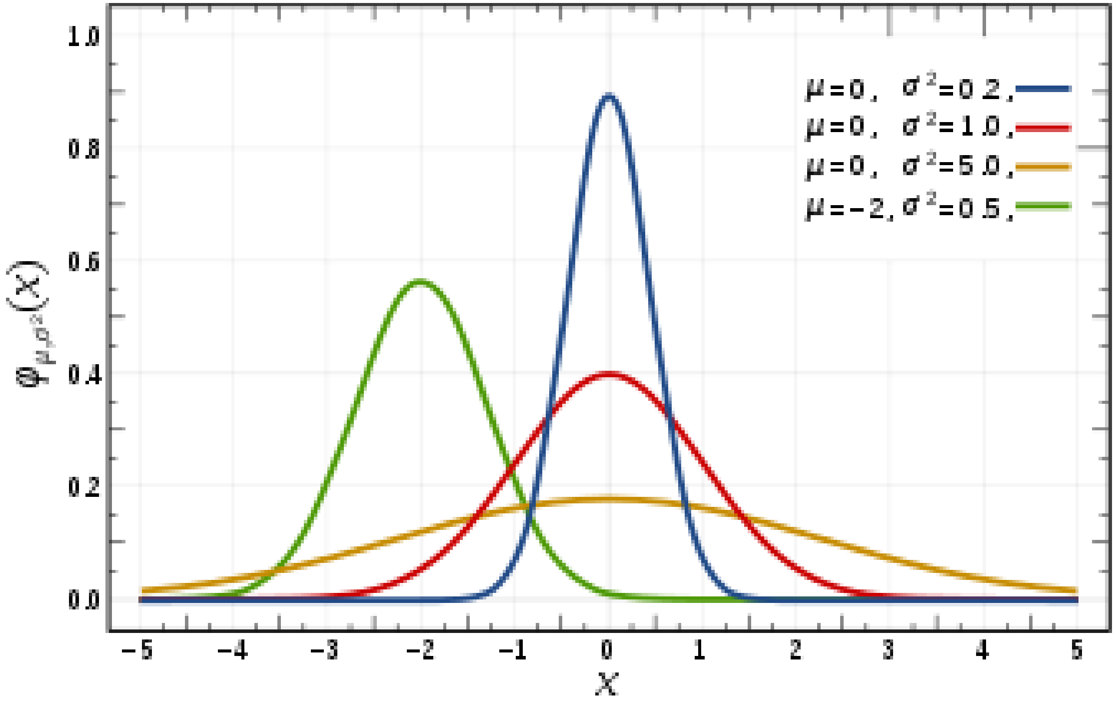
Uniform distribution on [a, b]: CDF

?

Gaussian distribution: PDF

Mean: μ (mu): the most likely value of the outcome
Standard deviation: σ : how spread out the outcomes are

$$N(\mu, \sigma)$$



Gaussian distribution: cdf

- ?