Probability in Computing

LECTURE 10

CS
237

Reminders

- HW 5 due on Thursday

Reading

- LLM 18.7, 18.8

Last time

- Conditional Probability
- Tree Diagrams
- Product Rule
- Independent Events

Today

- Independent Events
- Sequences of Independent Events
- Bayes' Rule

Independent Events

Definition: Independent Events

Two events A and B are independent if

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$$

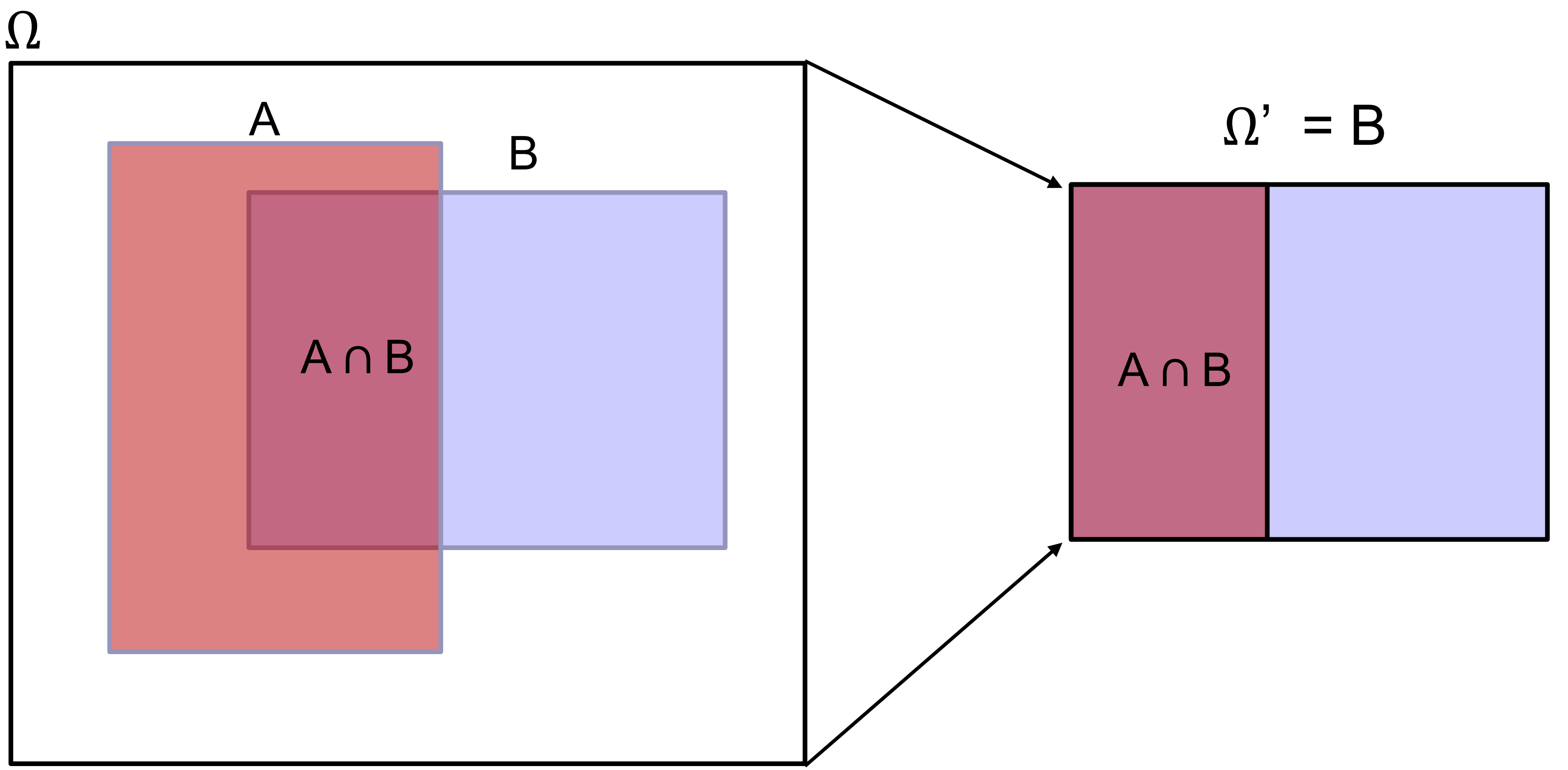
- Equivalent condition: $\Pr(A \mid B) = \Pr(A)$

  Alternatively, $\Pr(B \mid A) = \Pr(B)$
- Independence is symmetric:

  if A is independent of B then B is independent of A
- This definition can be used even when $\Pr(B) = 0$
  - An event B with $\Pr(B) = 0$ is independent of all events, including itself.

Independent Events

Conditioning the original sample space means changing the perspective: when A and B are independent, then Pr(A) does not change when reduce the sample space from Ω to B:



Independent Events

Google Colab Simulation

A = “C1 flips heads”

B = “C2 flips heads”

$$\Pr(A) = \frac{1}{2}, \Pr(B) = \frac{1}{2}$$

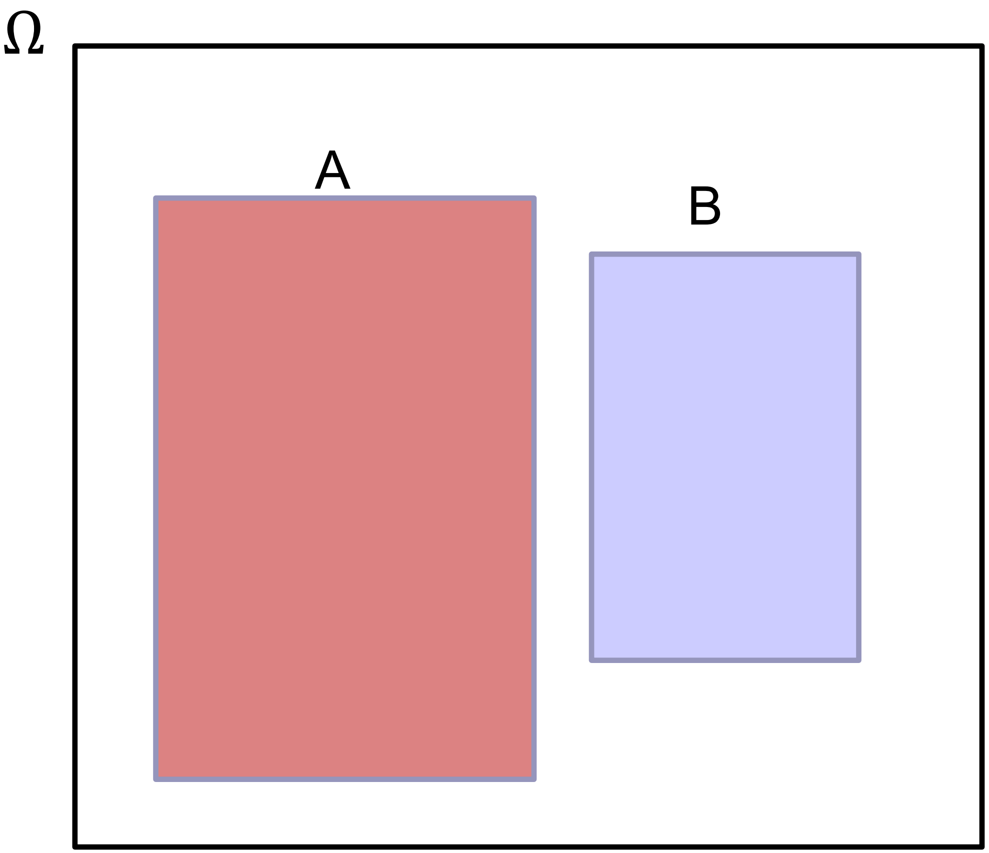
$$\Pr(A \cap B) = \frac{1}{4}$$

$$\Pr(A \mid B) = \frac{1/4}{1/2} = \frac{1}{2} = \Pr(A)$$

Independent Events

Be Careful: Independent is not the same as disjoint!

When events A and B are disjoint, Pr(A), Pr(B) > 0, if A happens then B is guaranteed not to happen and vice versa:



$$A \cap B = \emptyset$$

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0}{\Pr(B)} = 0$$

Independence

Let's try some examples.....

Experiment: roll two fair 4-sided dice

A = "first die is 1"

B = "sum of the two rolls is 5"

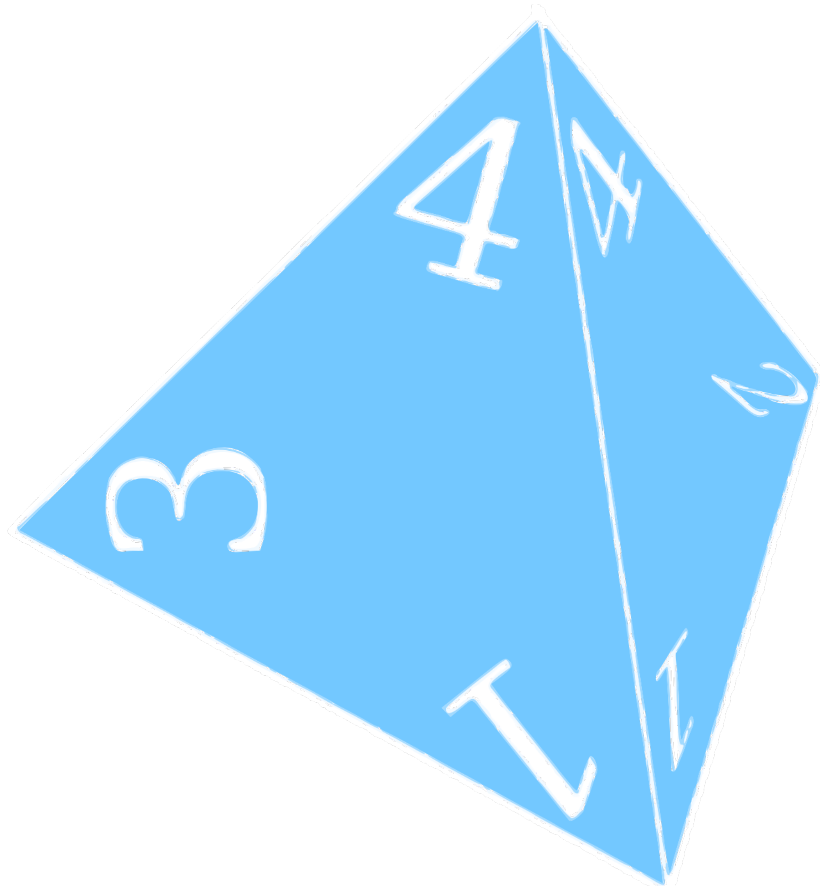
Are *A* and *B* independent?

Experiment: roll two fair 4-sided dice

A = "first die is 1"

B = "sum of the two rolls is ≤ 5"

Are *A* and *B* independent?



D4, Four Sided Die (Blue)
https://openclipart.org/

Top Hat question (Join Code: 033357)

Experiment: roll two fair 4-sided dice

A = "first die is 1"

B = "sum of the two rolls is 5"

Are *A* and *B* independent?

A. YES

B. NO

Top Hat question (Join Code: 033357)

Experiment: roll two fair 4-sided dice

A = "first die is 1"

B = "sum of the two rolls is ≤ 5"

Are *A* and *B* independent?

A. YES

B. NO

Top Hat question (Join Code: 033357)

Experiment: toss two fair coins

- A = “first toss is H”
- B = “both tosses give the same result”

Are *A* and *B* independent?

A. YES

B. NO

Top Hat question (Join Code: 033357)

Experiment: toss two biased coins with Pr(H)= 2/3

- A = "first toss is H"
- B = "both tosses give the same result"

Are *A* and *B* independent?

A. YES

B. NO

Top Hat question (Join Code: 033357)

Experiment: toss two biased coins with Pr(H)= 2/3

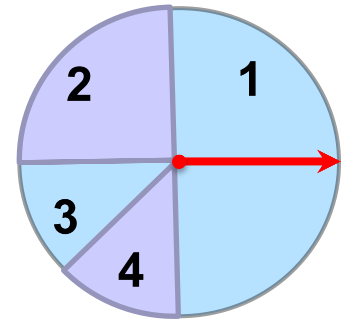
- A = "first toss is H" = {(HH,HT)}
- B = "both tosses give the same result" = {(HH,TT)}

Are *A* and *B* independent?

A. YES
B. NO

Independent Repeated Trials: Example

- Experiment: X = spin the dial of the spinner and observe the region where it stopped.
  - $\Omega = \{1,2,3,4\}$
  - $\Pr(1) = \frac{1}{2},\ \Pr(2) = \frac{1}{4},\ \Pr(3) = \Pr(4) = \frac{1}{8}$



- Now we spin the dial twice (X_1, X_2)
  - Sample space is $\Omega \times \Omega = \{(i,j) : 1 \leq i,j \leq 4\}$
  - $\Pr((i,j)) = \Pr(i) \cdot \Pr(j)$

also denoted Ω^2

Events ``1st spin is $i$'' and ``2nd spin is j'' are independent

- We can spin it n times
  - Each outcome is a sequence of n elements, each drawn from Ω

$$(X_1, X_2, \ldots X_n)$$

  - Sample space is Ω^n
  - $\Pr((s_1, \ldots, s_n)) = \Pr(s_1) \cdot \ldots \cdot \Pr(s_n)$

Technical note: We are overusing Pr
We are talking about two different probability functions (over Ω^n and over Ω)

Independent Repeated Trials

- Let Ω be a sample space for X with a probability function Pr and $n \in \mathbb{N}$.
- Let Ω^n denote the set of all length-n sequences of elements from Ω.
- Then n independent repeated trials of the original random experiment are represented by n random variables $(X_1, X_2, \ldots X_n)$ with
  - sample space Ω^n
  - probability function $\Pr\big((s_1, \ldots, s_n)\big) = \Pr(s_1) \cdot \ldots \cdot \Pr(s_n)$

Examples

- Roll a 6-sided die twice:
  - Sample space for one roll is $\Omega = \{1,2,3,4,5,6\}$ and $\Pr(s) = \frac{1}{6}$ for all $s \in \Omega$
  - Sample space for two rolls is Ω^2, and $\Pr((s_1, s_2)) = \frac{1}{36}$ for all $s_1, s_2 \in \Omega$

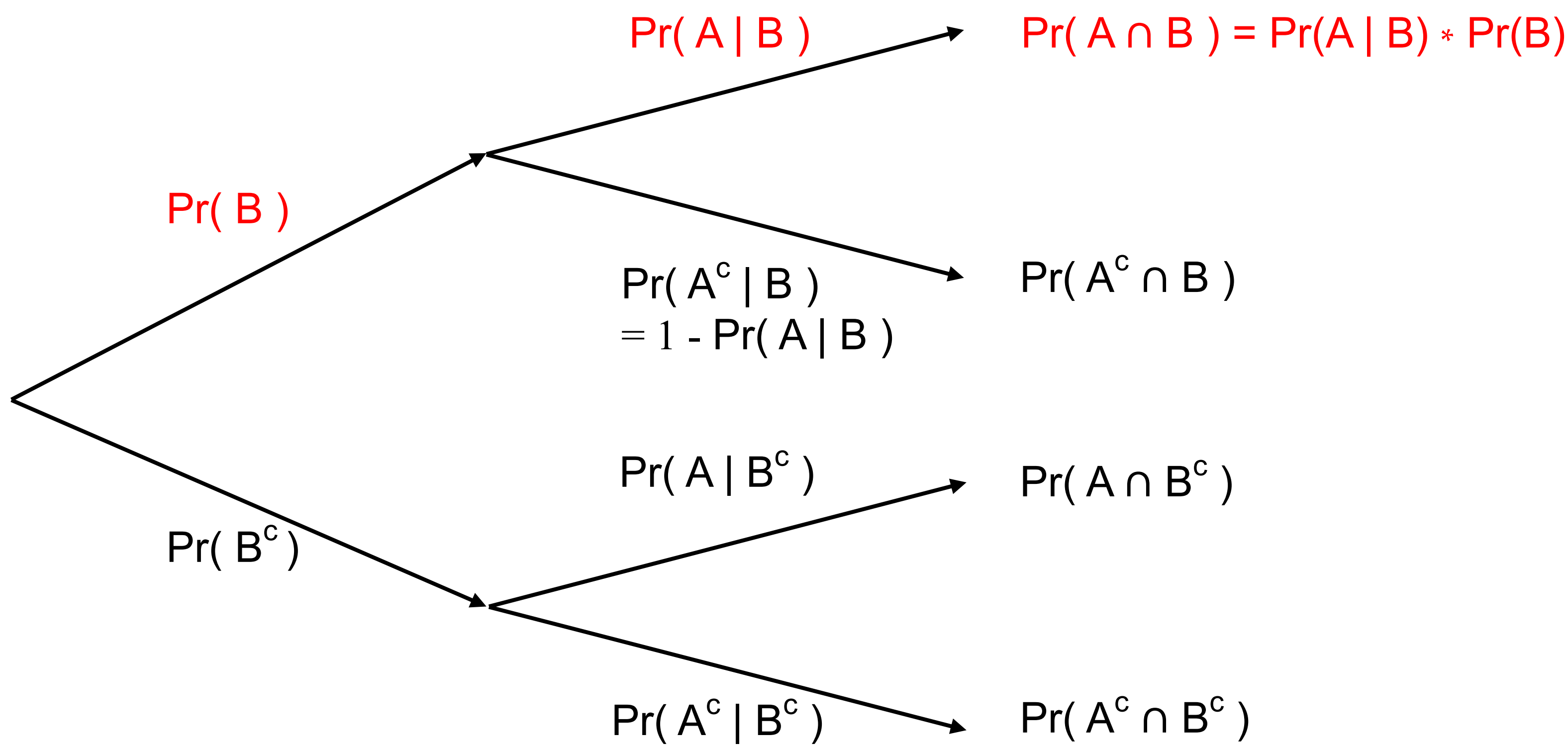
- Toss a fair coin until you see a head: $(X_1, X_2, \ldots, X_k, \ldots)$
  - Sample space for one toss is $\Omega = \{0,1\}$ and $\Pr(0) = \Pr(1) = \frac{1}{2}$
  - Sample space for one trial is $\Omega^{\mathbb{N}}$; an outcome in Ω^k has probability $\frac{1}{2^k}$.

CS 237

Independence in Tree Diagrams

Pr(A | B) considers an event B followed by an event A, and how the occurence of B affects the occurence of A. What are the labels on a tree diagram of this random experiment?

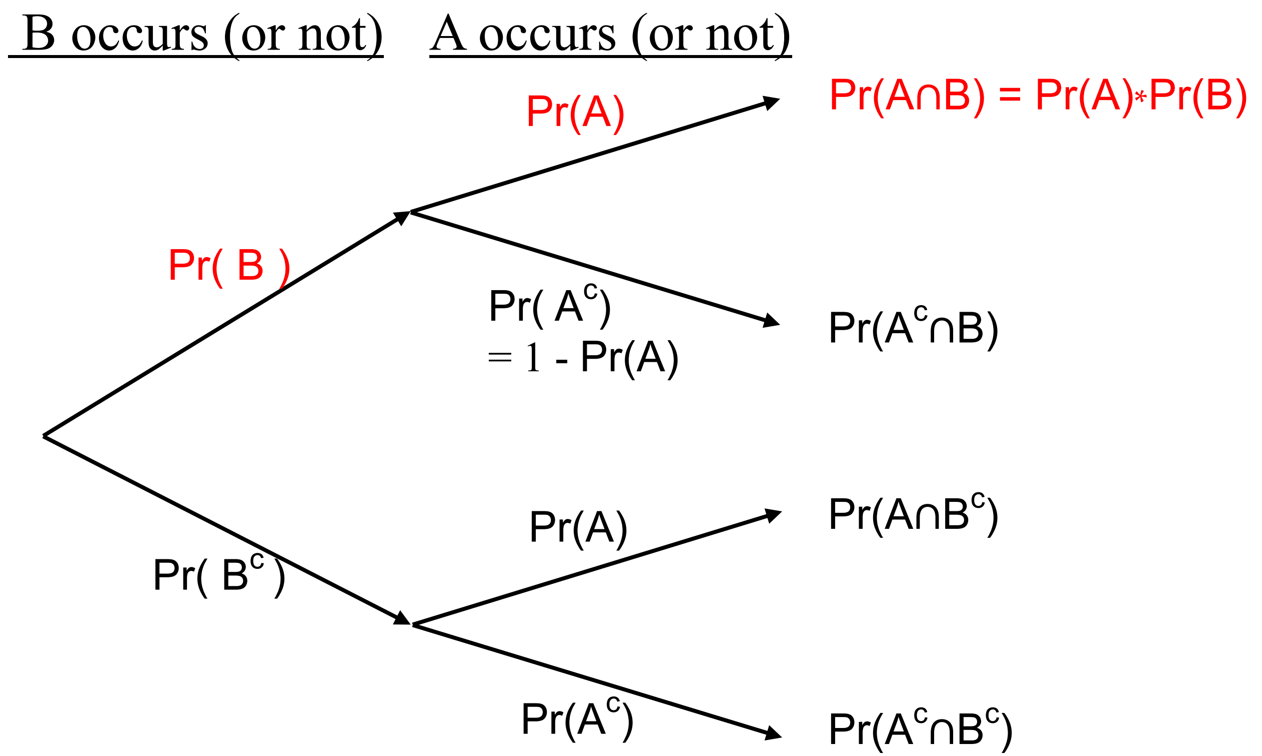
B occurs (or not) A occurs (or not)



Independence in Tree Diagrams

How does this relate to tree diagrams?

When the events are independent, then we have the familiar tree diagram in which we simply write the probabilities of the events on each arc:



Independence in Tree Diagrams

Example: Sampling with or without replacement

1. You draw 3 cards from a standard deck with replacement: what is the probability they are all Spades?

$$\frac{13}{52} \cdot \frac{13}{52} \cdot \frac{13}{52}$$

2. You draw 3 cards from a standard deck without replacement: what is the probability they are all Spades?

$$\frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50}$$

Dependence and Causality

Fun fact: Dependence does not imply causality!

The NEW ENGLAND JOURNAL of MEDICINE

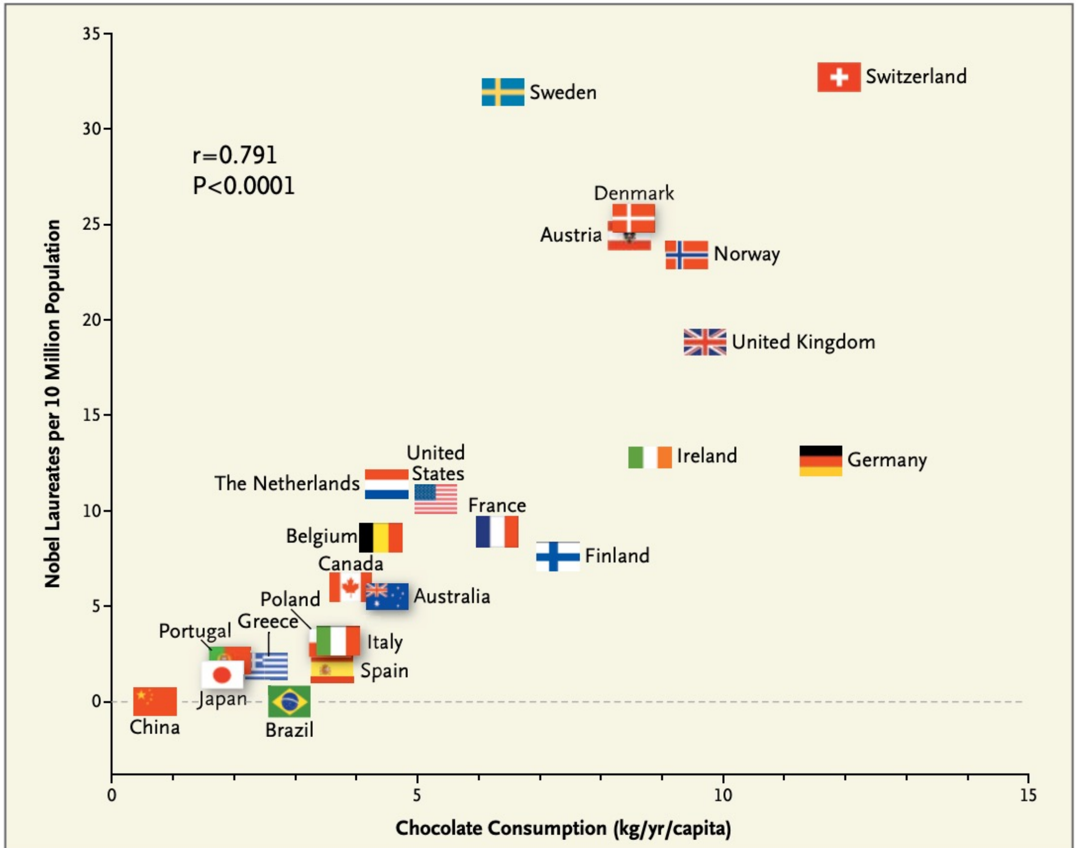


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

Dependence and Causality

ELSEVIER

Social Sciences & Humanities Open

Volume 2, Issue 1, 2020, 100082

Chocolate consumption and Noble laureates

Aloys Leo Prinz

Show more

+ Add to Mendeley   Share   Cite

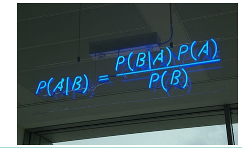
https://doi.org/10.1016/j.ssaho.2020.100082   Get rights and content

Highlights

- Chocolate consumption per capita is positively correlated with the stock of Nobel prizes per capita.
- A two-stage Heckman selection model is estimated.
- The correlation remains after control for scientific articles and R&D expenditures.
- It remains unclear whether the correlation is spurious or an indication for hidden variables.

CS 237

Bayes' Rule



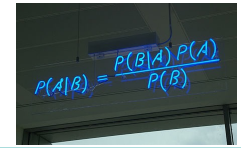
We can rearrange the conditional probability rule in a way that makes the sequence of the events irrelevant -- which happened first, A or B? Or did they happen at the same time? Does it matter?

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} \qquad \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

We can do a little algebra to define conditional probabilities in terms of each other:

CS 237

Bayes' Rule



We can rearrange the conditional probability rule in a way that makes the sequence of the events irrelevant -- which happened first, A or B? Or did they happen at the same time? Does it matter?

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} \qquad \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

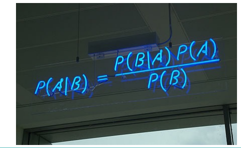
We can do a little algebra to define conditional probabilities in terms of each other:

$$\boxed{\Pr(B|A)} \cdot \boxed{\Pr(A)} = \Pr(B \cap A) = \boxed{\Pr(A|B)} \cdot \boxed{\Pr(B)}$$

so:

$$\Pr(B|A) = \frac{\Pr(A|B) \cdot \Pr(B)}{\Pr(A)}$$

Bayes' Rule



We can rearrange the conditional probability rule in a way that makes the sequence of the events irrelevant -- which happened first, A or B? Or did they happen at the same time? Does it matter?

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} \qquad \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

We can do a little algebra to define conditional probabilities in terms of each other:

$$\Pr(B|A) \cdot \Pr(A) = \Pr(B \cap A) = \Pr(A|B) \cdot \Pr(B)$$

so:

$$\Pr(B|A) = \frac{\Pr(A|B) \cdot \Pr(B)}{\Pr(A)}$$

Bayes' Rule

This has an interesting flavor, because we can ask about causes of outcomes:

A Priori Reasoning -- "I randomly choose a person and observe that he is male; what the probability that it is a smoker?"

"The first toss of a pair of dice is a 5; what is the probability that the total is greater than 8?"

Bayes' Rule

This has an interesting flavor, because we can ask about causes of outcomes:

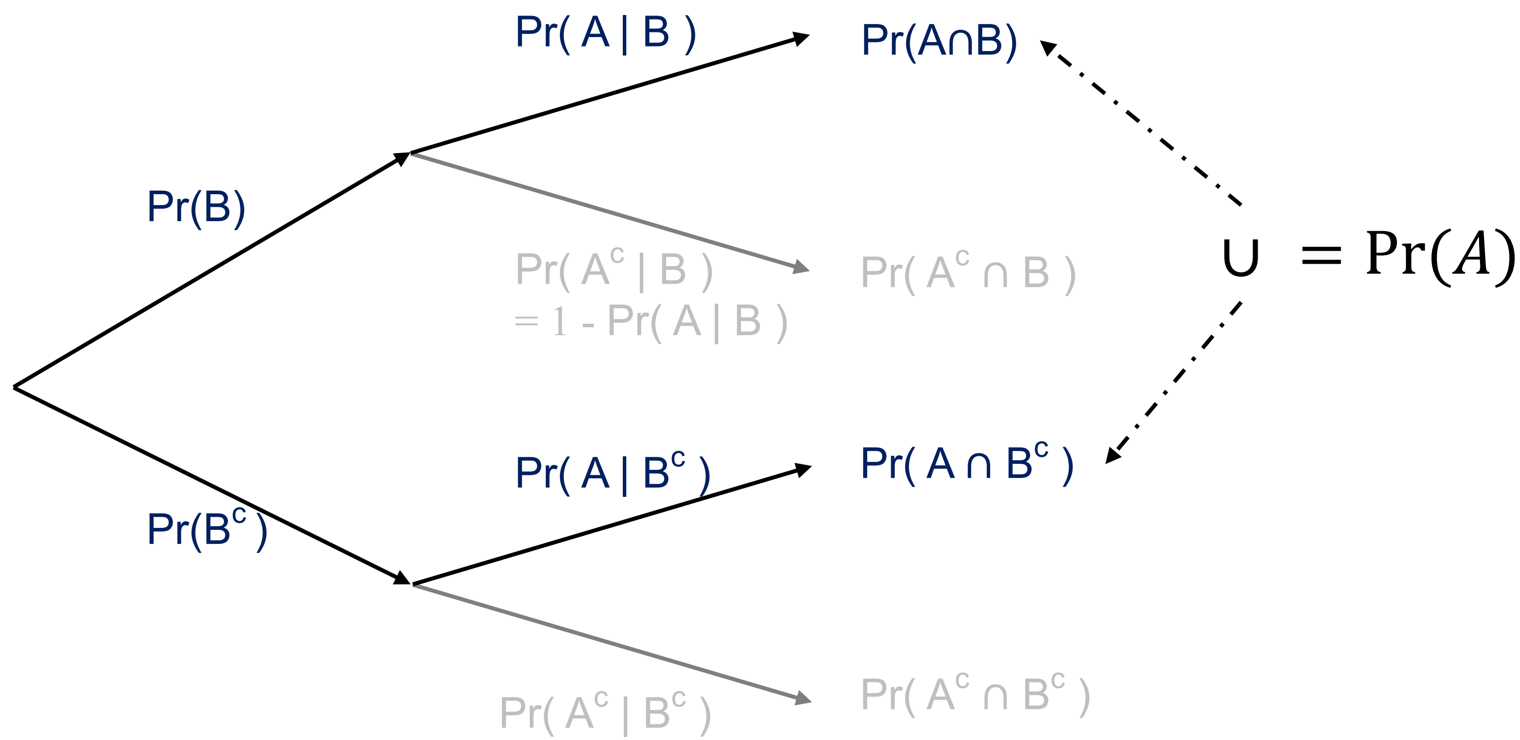
A Posteriori Reasoning -- "I find a cigarette butt on the ground, what is the probability that it was left by a man?"

"The total of a pair of thrown dice is greater than 8; what is the probability that the first toss was a 5?"

Bayes' Rule

The best way to understand this is to view it with a tree diagram!

P(B|A) = the probability that when A happens, it was "preceeded" by B:

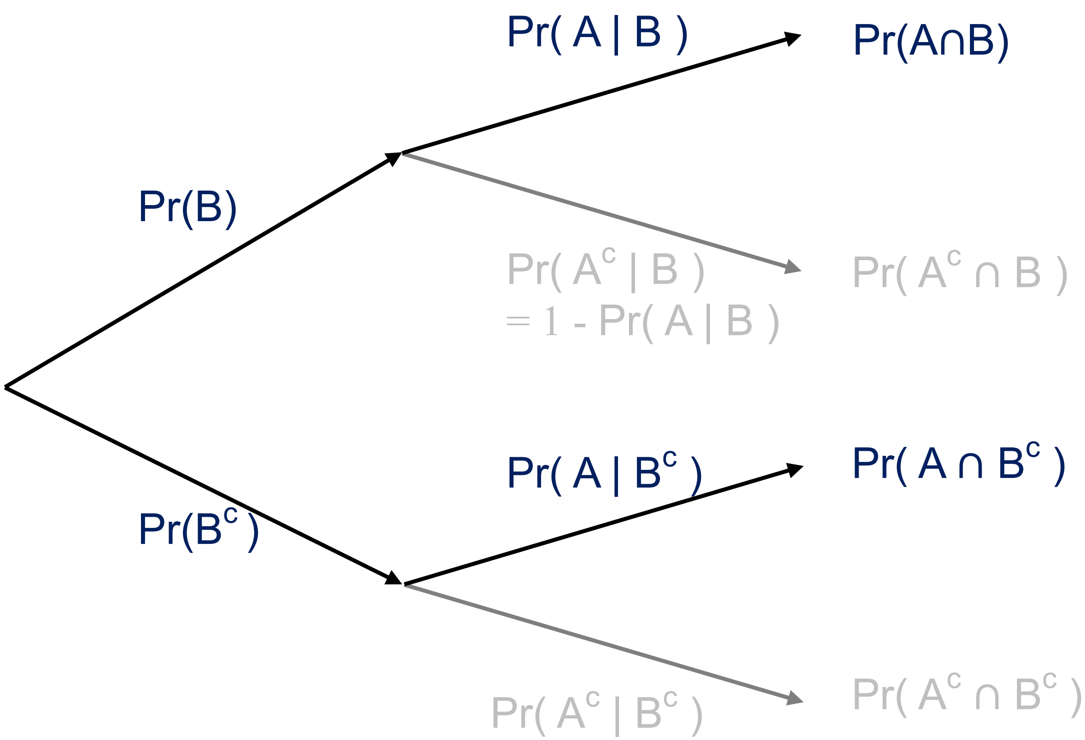


CS 237

Bayes' Rule

The best way to understand this is to view it with a tree diagram!

P(B|A) = the probability that when A happens, it was "preceeded" by B:



If A has happened, what is the probability that it did so on the path where B also occurred?

Note:

$\Pr(A) = \Pr(A \cap B) \cup \Pr(A \cap B^c)$

So what percentage of A is due to $A \cap B$?

Same calculation as:

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} = \frac{\Pr(A \cap B)}{\Pr(A)}$$

Bayes' Rule

Example: Suppose that 10% of female BU students smoke cigarettes and 20% of male BU students smoke cigarettes. Suppose that 60% of BU students are female and 40% male. I see a cigarette butt on the pavement. What is the probability it was thrown there by a female student?

Bayes' Rule

Example: We have two urns, A and B. A contains 2 red balls and 3 black balls. Urn B contains 2 red balls and 1 black ball. We select an urn uniformly at random and pick a ball and find it is red. What is the probability it came from Urn A?