CS 237: Probability in Computing

Wayne Snyder
Computer Science Department
Boston University

Lecture 3:

- Analyzing Problems using Decision Trees: The “Four-Step Method”
- Conditional Probability
- Independent vs Dependent Events

Review: Finite Equiprobable Probability Spaces

For finite and equiprobable probability spaces, we can calculate probabilities by simply counting:

$$P(A) = \frac{|A|}{|S|}$$

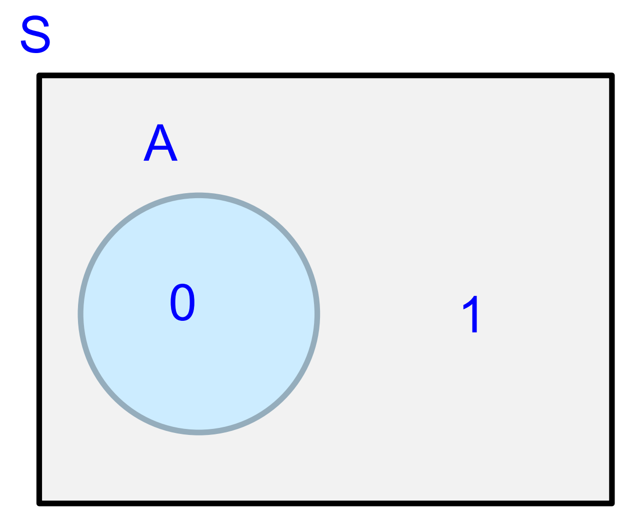
Recall: $|A|$ = cardinality (number of members) of A

Here, instead of area, we use the the number of elements – all have the same probability.

Example: Flip a coin, report how many heads are showing? Let A = "the coin lands with tails showing"

S = { 0, 1 }

P = { ½, ½ } **P(A) = 1/2**



Finite Equiprobable Probability Spaces

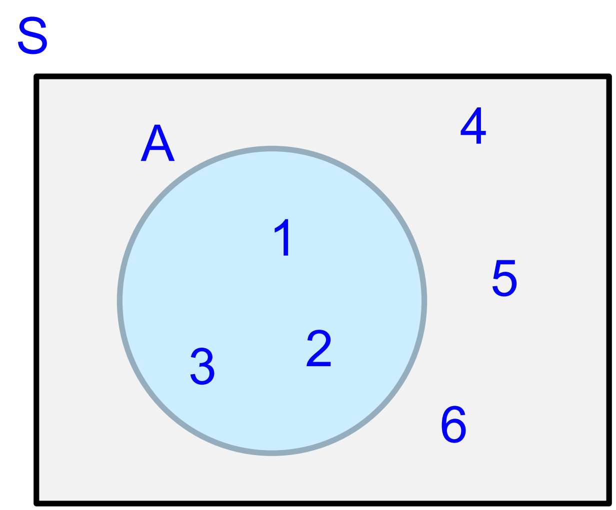
$$P(A) = \frac{|A|}{|S|}$$

Example: Roll a die, how many dots showing on the top face? Let A = "less than 4 dots are showing."

S = { 1, 2, ... , 6 }

P = { 1/6, 1/6, ... , 1/6 }

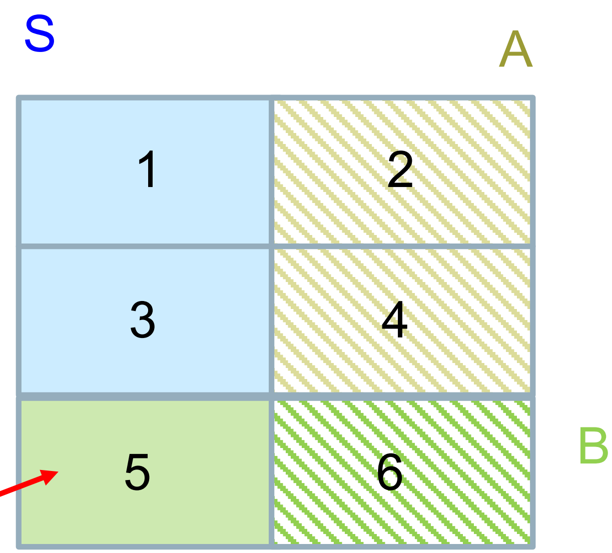
P(A) = 3/6 = 1/2



Review: Finite Equiprobable Probability Spaces

Of course, we can still think of probabilities as **areas**, as long as we make them all the same!

Example: Toss a die and output the number of dots showing. Let A = "there are an even number of dots showing" and B = "there are at least 5 dots showing."



Equiprobable: area of each elementary event is SAME:
1/6 = 0.16666...

P(A) = P(2) + P(4) + P(6)
= 1/6 + 1/6 + 1/6
= ½

P(B) = P(5) + P(6)
= 1/6 + 1/6
= 1/3

Review: Finite Non-Equiprobable Probability Spaces

For finite non-equiprobable spaces, you can not simply count cardinalities. Area is a good analogy for simple problems.

Example: Flip three fair coins and count the number of heads. Let A = "2 heads are showing" and B = "at most 2 heads are showing."

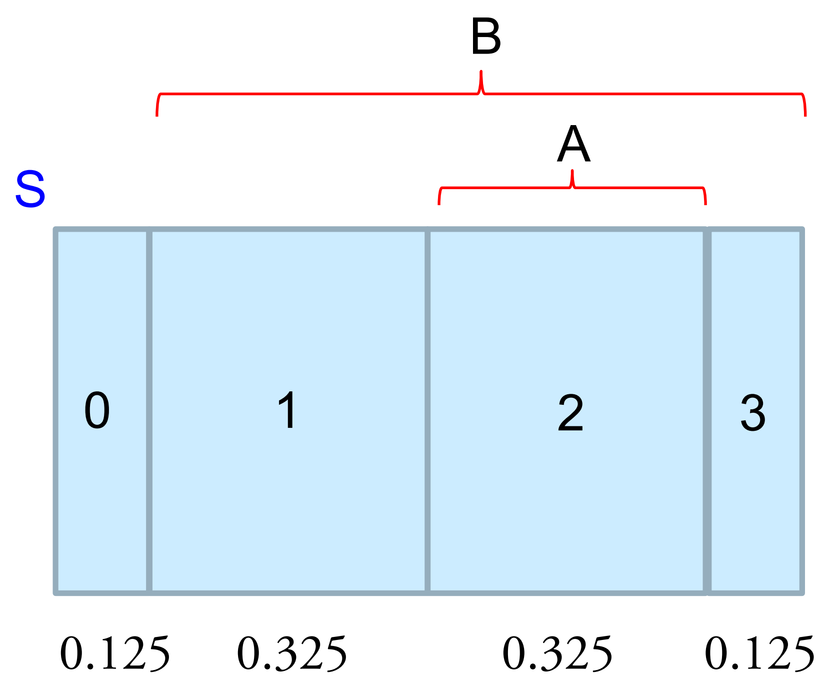
Pre-Sample:	**{ TTT,**	**TTH,**	**THT,**	**THH,**	**HTT,**	**HTH,**	**HHT,**	**HHH }**
# heads:	**0**	**1**	**1**	**2**	**1**	**2**	**2**	**3**

S = { 0, 1, 2, 3 }
P = { 1/8, 3/8, 3/8, 1/8 }

A = { 2 } **B = { 0, 1, 2 }**

P(A) = P(2)
= 3/8

P(B) = P(0) + P(1) + P(2)
= 1/8 + 3/8 + 3/8
= 7/8



Discrete Non-Equiprobable Probability Spaces

For discrete non-equiprobable probability spaces, we can't just count, we have to analyze each case.

Example: **F**lip an unfair coin twice, where the P(heads) = 0.6 and P(tails) = 0.4, and and count the number of heads showing. Let A = there is at least one head showing. Give S, P, and P(A).

Pre-sample space = { TT, TH, HT, HH }

S = { 0, 1, 2 }

P(0) = P(TT) = 0.4*0.4 = 0.16

P(1) = P(TH)+P(HT) = 0.4*0.6 + 0.6*0.4 = 0.24 + 0.24 = 0.48

P(2) = P(HH) = 0.6*0.6 = 0.36

Check: 0.16 + 0.48 + 0.36 = 1.00

Solving Probability Problems using Decision Trees and the "Four-Step" Method

To the board......

Conditional Probability: P(A | B)

Conditional probability problems occur when you have some additional information about the random experiment, but not all information—so the result is still random!

These are stated using two events, one of which is known to have happened, and one of which is unknown:

P(A | B) = The probability that A occurs, given that B HAS occurred.

Example: Suppose I toss two dice, one in front of you, and the other where you can not see the result. The one you can see shows 3 dots. What is the probability that more than 8 dots showed one both dice?

Or: P(A | B)

where:

A = "The total number of dots is more than 8"

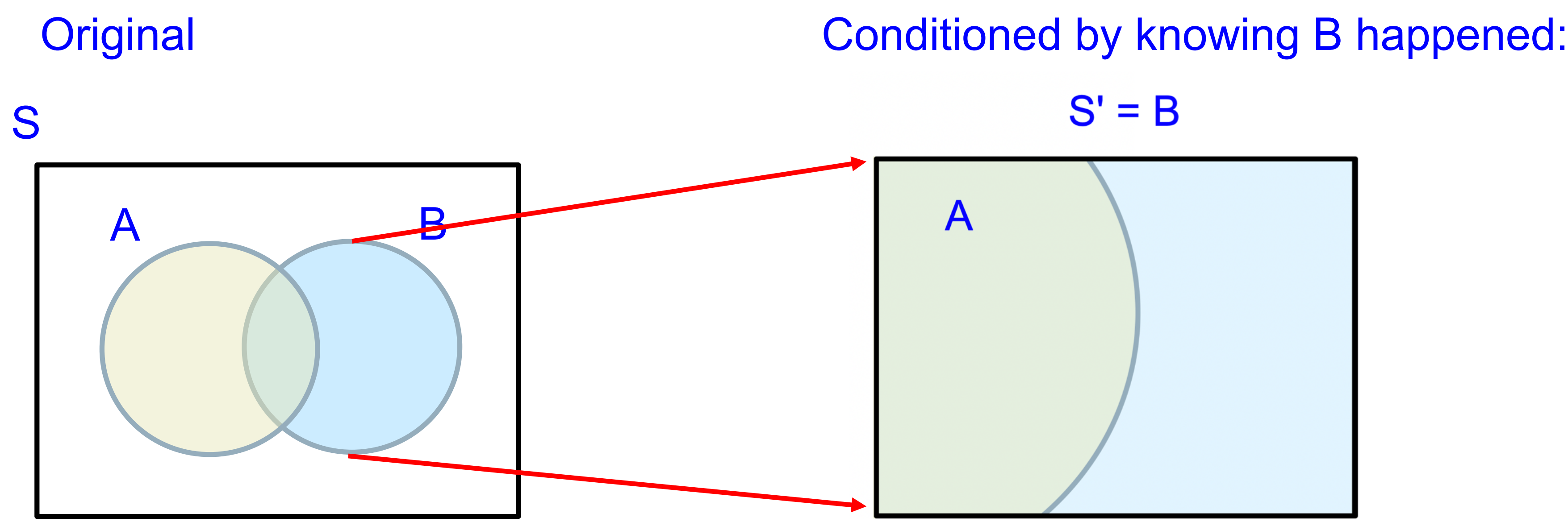
B = "The first die shows 3 dots"

Conditional Probability: P(A | B)

Example: Roll two dice. A = "the total dots is > 8" and B = "the first roll was 3"
What is P(A | B)?

The key to solving such problems is to realize that there are two probability spaces:

- the one before you know whether B has happened, and
- the one that has been "conditioned" by knowing that B has definitely happened, so the sample space has shrunk and the proportion representing event A may have changed:



Conditional Probability: P(A | B)

Example: Roll two dice. A = "the total # dots is > 8" and B = "the first roll was 3"
What is P(A | B)?

P(A) = ?

P(B) = ?

Second Roll

First Roll

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

A

First Roll

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

B

Conditional Probability: P(A | B)

Example: Roll two dice. A = "the total # dots is > 8" and B = "the first roll was 3"
What is P(A | B)?

Original

First Roll

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

A

Conditioned by knowing B happened

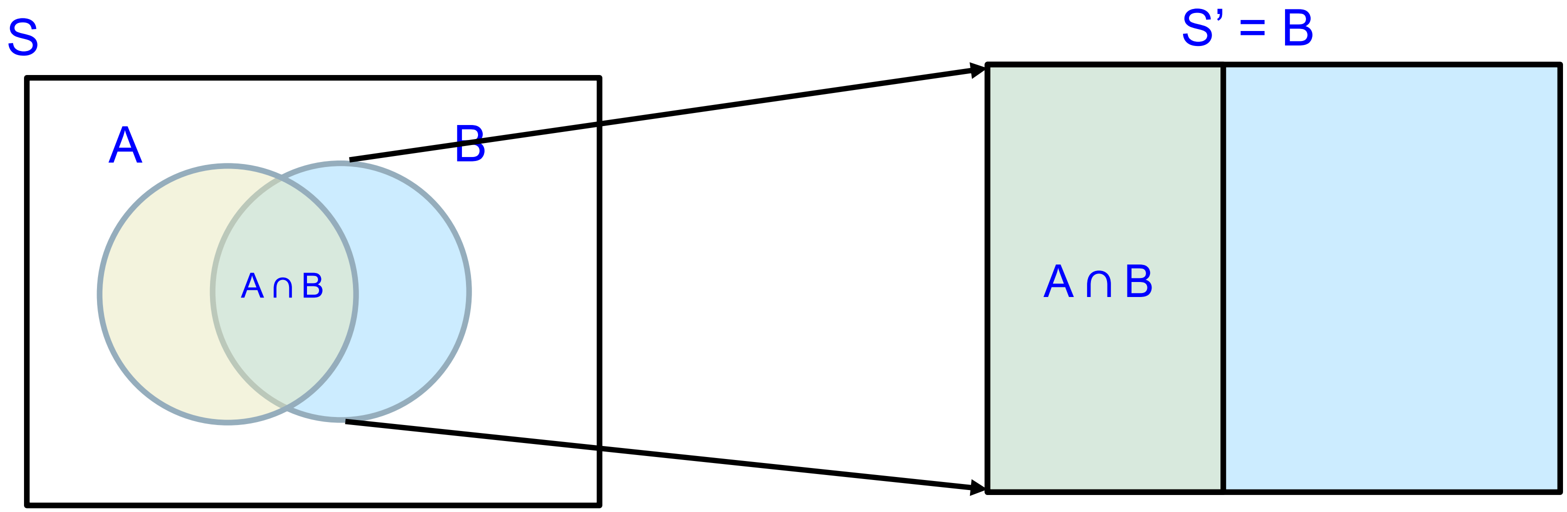
3	4	5	6	7	8	9

P(A | B) = 1/6

Conditional Probability

Conditioning the original sample space means changing the perspective: instead of finding the area of A inside S, we are finding the area of $A \cap B$ inside B:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$



Conditional Probability: P(A | B)

Example: Roll two dice. A = "the total # dots is > 8" and B = "the first roll was 3"
What is P(B | A)?

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{36}}{\frac{10}{36}} = \frac{1}{10}$$

Original

Second Roll (columns), First Roll (rows)

First Roll \ Second Roll	1	2	3	4	5	6	
1	2	3	4	5	6	7	
2	3	4	5	6	7	8	
3	4	5	6	7	8	9	B
4	5	6	7	8	9	10	
5	6	7	8	9	10	11	
6	7	8	9	10	11	12	

Conditioned by knowing A happened:

Second Roll (columns), First Roll (rows)

First Roll \ Second Roll	1	2	3	4	5	6
1						
2						
3						9
4					9	10
5				9	10	11
6			9	10	11	12

Independence and Dependence: Review

We say that two events A and B are independent if

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A \mid B) = P(A)$$

or, equivalently, and most importantly as we go forward:

$$P(A \cap B) = P(A) * P(B)$$

If two events are NOT independent, then they are dependent.

Example:

What is the probability of getting HHT when flipping three fair coins?

P(HHT) = P(H) * P(H) * P(T) = ½ * ½ * ½ = 1/8.

Note: Independence does not depend on physical independence, and dependence does not imply a causal relationship. However, it gives you some evidence!

Independence and Dependence: Review

We say that two events A and B are independent if

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A \mid B) = P(A)$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

or:

$$P(A \cap B) = P(A) * P(B)$$

Example:

Suppose in a particular city, 40% of the population is male, and 60% female, and 20% of the population smokes. If male smokers are 8% of the population, then are smoking and gender independent? That is, are the following two events independent?

A = Smoker

B = Male

Independence and Dependence: Review

We say that two events A and B are independent if

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A \mid B) = P(A)$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

or:

$$P(A \cap B) = P(A) * P(B)$$

Example:

Suppose in a particular city, 40% of the population is male, and 60% female, and 20% of the population smokes. If male smokers are 8% of the population, then are smoking and gender independent? That is, are the following two events independent?

YES. Check:

A = Smoker

$P(A \cap B) = 0.08 = 0.4 * 0.2 = P(A) * P(B)$

B = Male

Independence and Dependence: Review

Digression: Dependence does not imply causality!

The NEW ENGLAND JOURNAL of MEDICINE

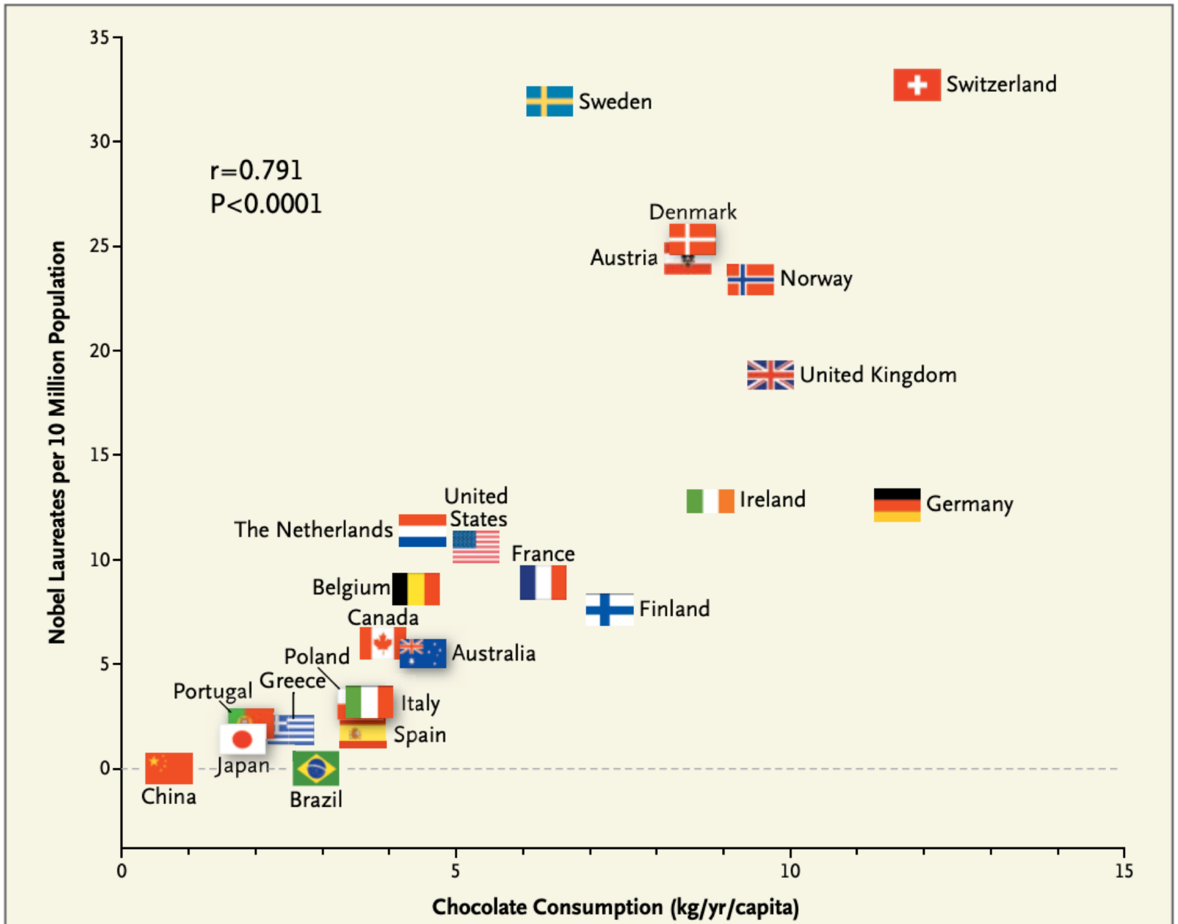


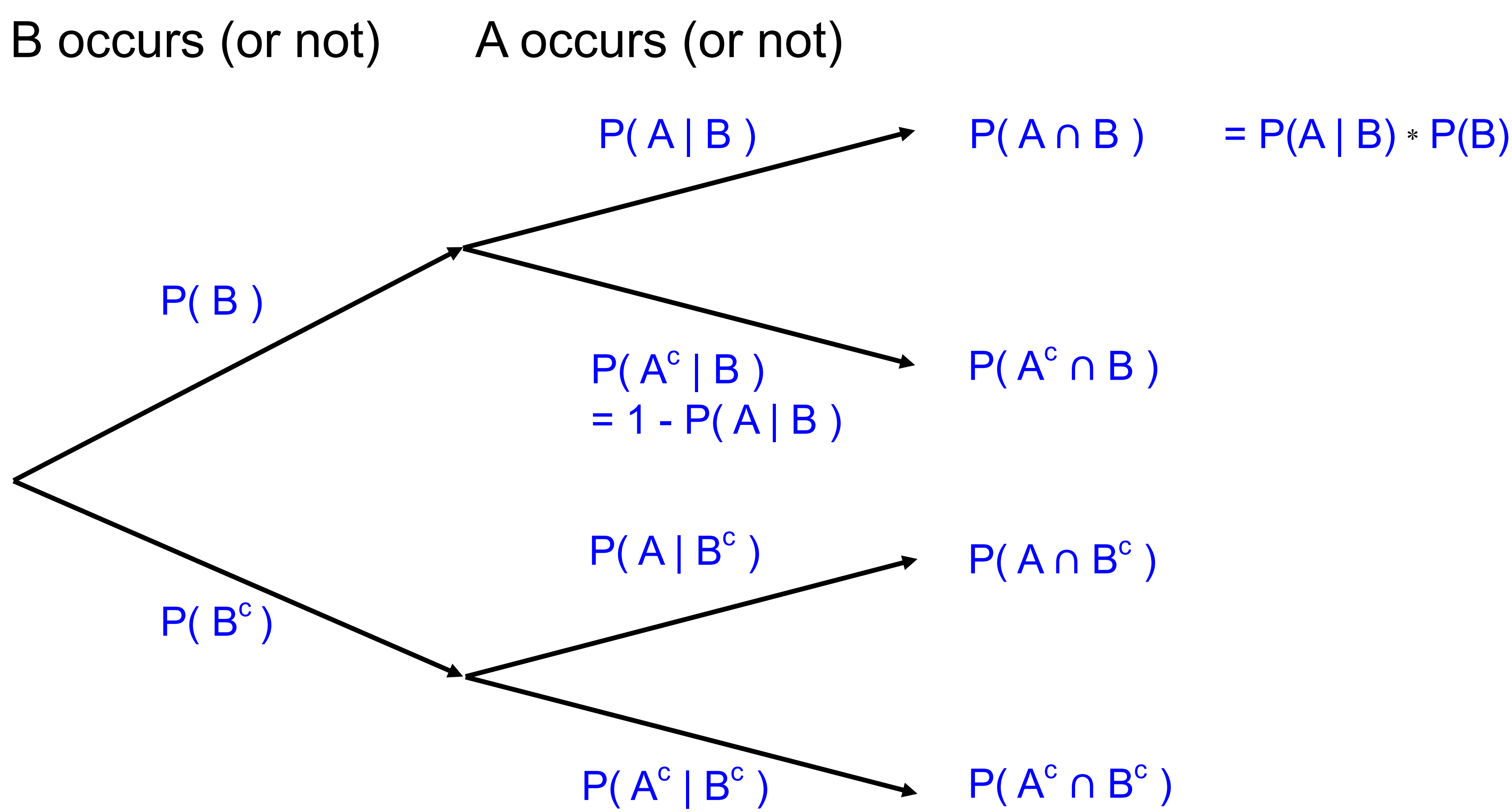
Figure 1. **Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.**

Independence and Dependence: Review

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

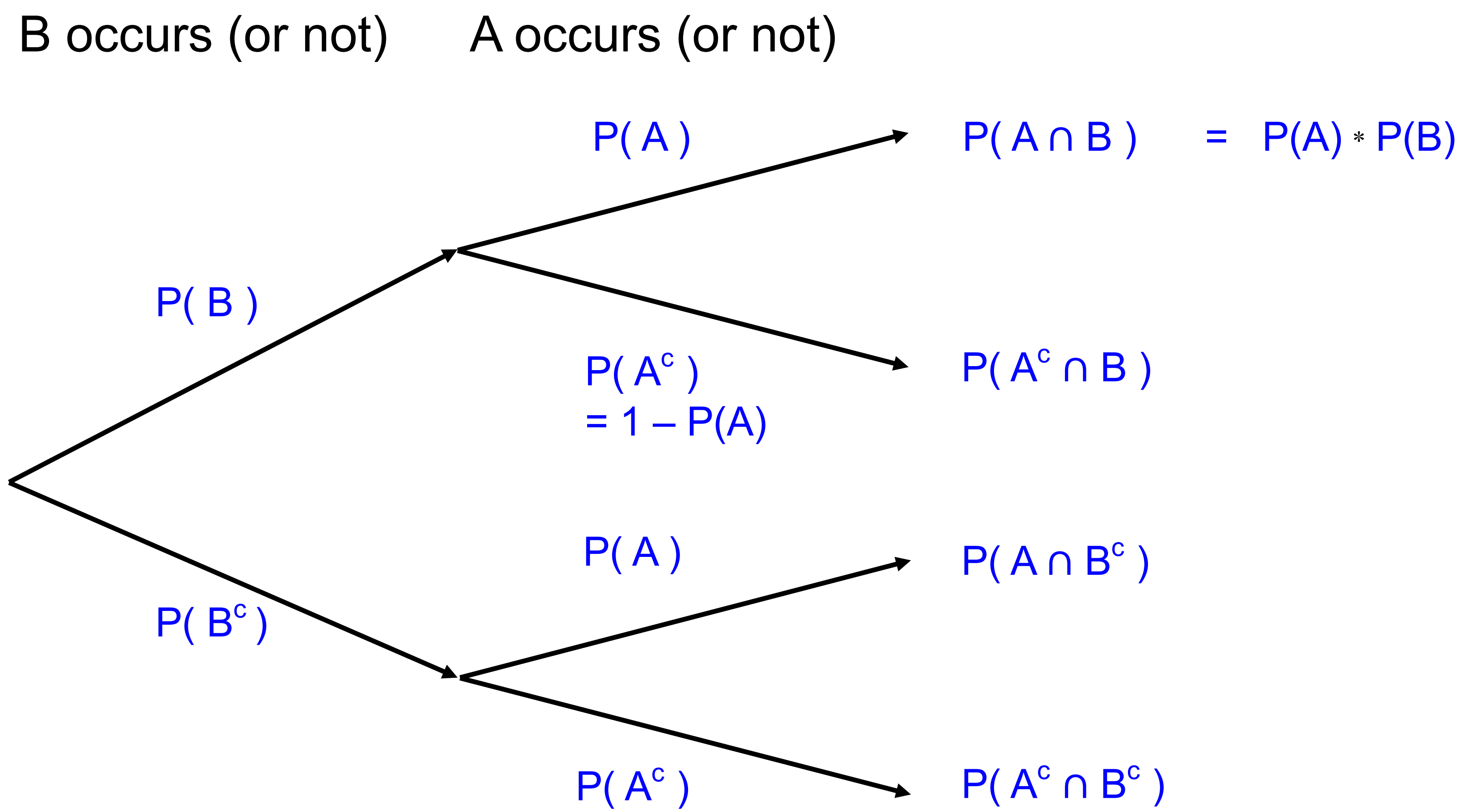
How does this relate to tree diagrams?

P(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?



Independence and Dependence: Review

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:



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?

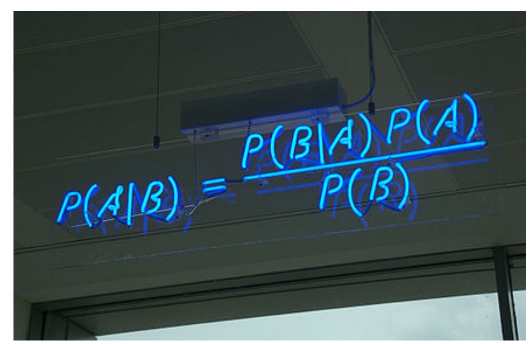
$$P(A\,|\,B) = \frac{P(A \cap B)}{P(B)} \qquad\qquad P(B\,|\,A) = \frac{P(B \cap A)}{P(A)}$$

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

$$P(B\,|\,A) * P(A) = P(B \cap A) = P(A\,|\,B) * P(B)$$

so:

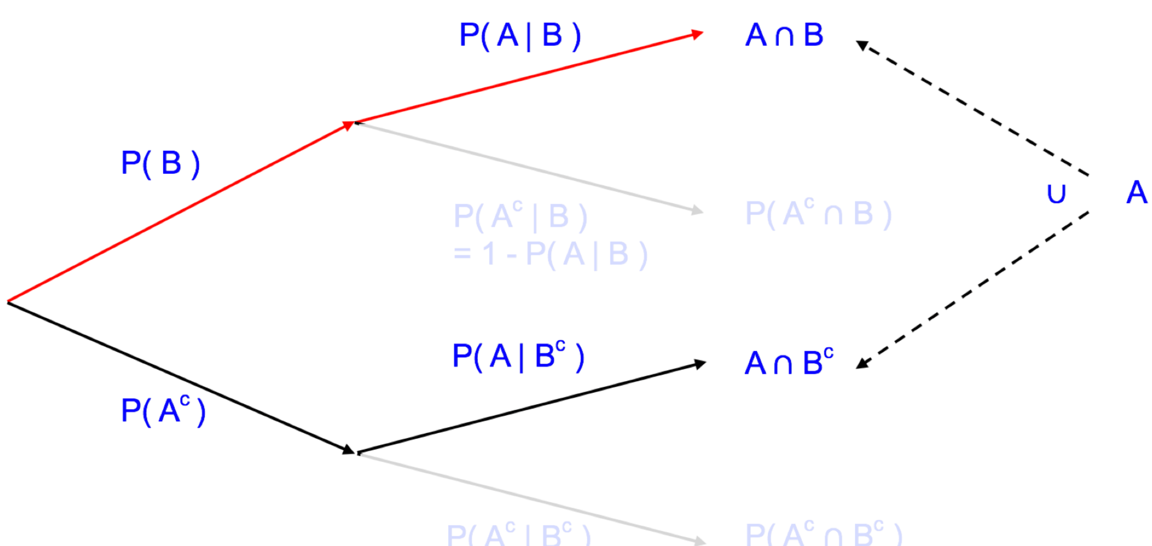
$$P(B\,|\,A) = \frac{P(A\,|\,B) * P(B)}{P(A)}$$



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:

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

So what percentage of A is due to A ∩ B ?

Same calculation as:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)}$$

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?"

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?"

This seems odd, because instead of reasoning forward from "causes to effects" we are reasoning backwards from "effects to causes" but really it is just different ways of phrasing the mathematical formulae. Time is not really relevant!