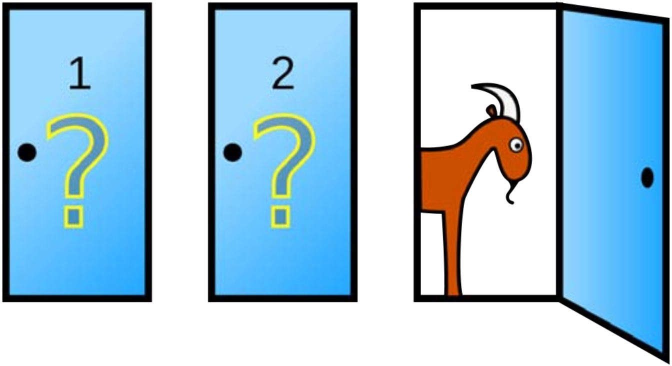
Conditional Probability, Independence, Bayes' Theorem



Sample Space Confusions

1. Sample space = *set* of all possible outcomes of an experiment.
2. The size of the set is **NOT** the sample space.
3. Outcomes can be sequences of numbers.

Examples.
1. Roll 5 dice: Ω = set of all sequences of 5 numbers between 1 and 6, e.g. $(1, 2, 1, 3, 1, 5) \in \Omega$.
The size $|\Omega| = 6^5$ is not a set.

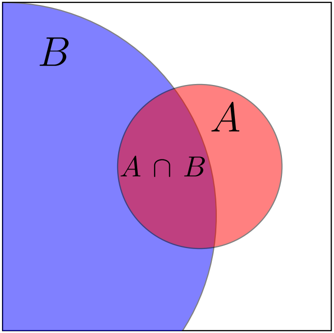
2. Ω = set of all sequences of 10 birthdays,
e.g. $(111, 231, 3, 44, 55, 129, 345, 14, 24, 14) \in \Omega$.
$|\Omega| = 365^{10}$

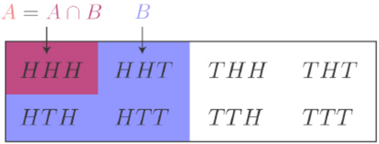
3. n some number, Ω = set of all sequences of n birthdays.
$|\Omega| = 365^n$.

Conditional Probability

'the probability of A given B'.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ provided } P(B) \neq 0.$$





Conditional probability: Abstractly and for coin example

Table/Concept Question

(Work with your tablemates, then everyone click in the answer.)

Toss a coin 4 times. Let

$A =$ 'at least three heads'

$B =$ 'first toss is tails'.

1. What is $P(A|B)$?

(a) $1/16$ (b) $1/8$ (c) $1/4$ (d) $1/5$

2. What is $P(B|A)$?

(a) $1/16$ (b) $1/8$ (c) $1/4$ (d) $1/5$

Table Question

> *"Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure and a passion for detail."**

What is the probability that Steve is a librarian?
What is the probability that Steve is a farmer?

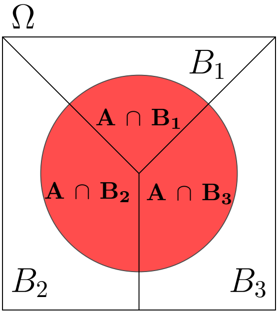
*From *Judgment under uncertainty: heuristics and biases* by Tversky and Kahneman.

Multiplication Rule, Law of Total Probability

Multiplication rule: $P(A \cap B) = P(A|B) \cdot P(B)$.

Law of total probability: If B_1, B_2, B_3 *partition* Ω then

$$\begin{aligned} P(A) &= P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) \\ &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \end{aligned}$$

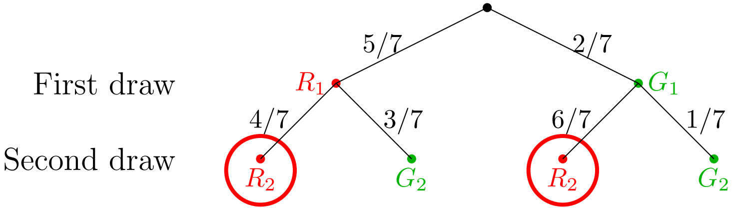


Trees

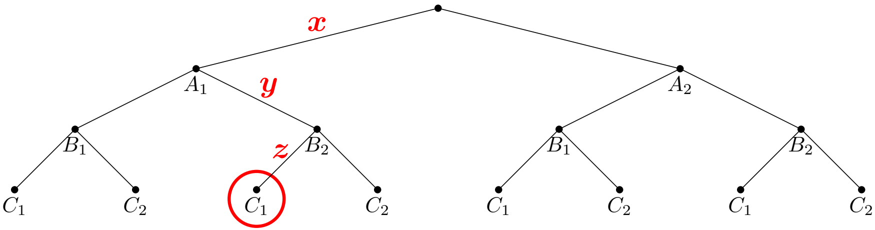
- Organize computations
- Compute total probability
- Compute Bayes' formula

Example. : Game: 5 red and 2 green balls in an urn. A random ball is selected and replaced by a ball of the other color; then a second ball is drawn.

1. What is the probability the second ball is red?
2. What is the probability the first ball was red given the second ball was red?



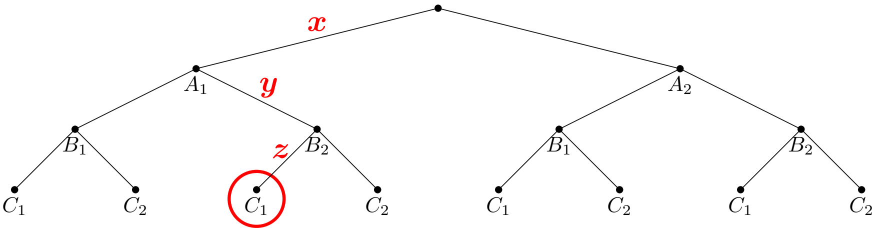
Concept Question: Trees 1



1. The probability x represents

(a) $P(A_1)$
(b) $P(A_1|B_2)$
(c) $P(B_2|A_1)$
(d) $P(C_1|B_2 \cap A_1)$.

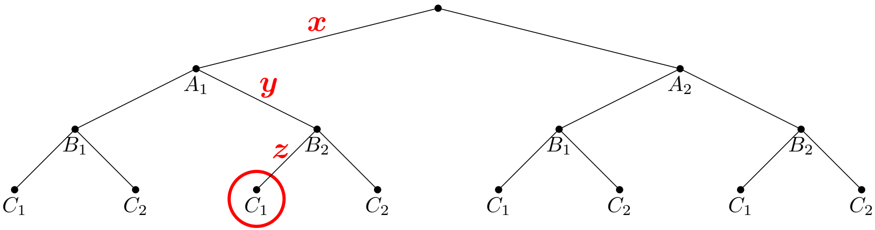
Concept Question: Trees 2



2. The probability y represents

(a) $P(B_2)$
(b) $P(A_1|B_2)$
(c) $P(B_2|A_1)$
(d) $P(C_1|B_2 \cap A_1)$.

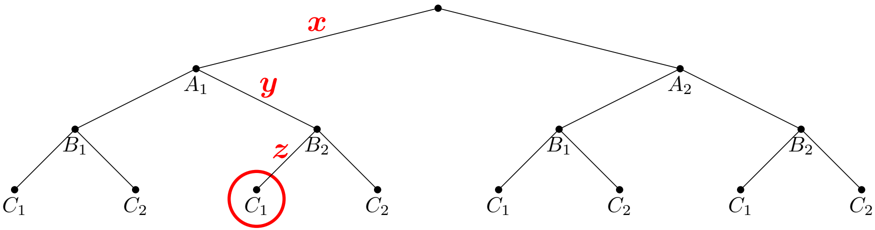
Concept Question: Trees 3



3. The probability z represents

(a) $P(C_1)$
(b) $P(B_2|C_1)$
(c) $P(C_1|B_2)$
(d) $P(C_1|B_2 \cap A_1)$.

Concept Question: Trees 4



4. The circled node represents the event

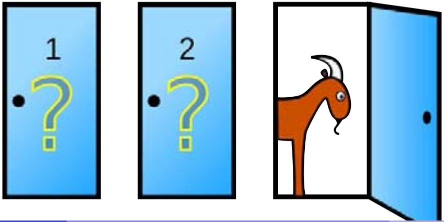
(a) C_1
(b) $B_2 \cap C_1$
(c) $A_1 \cap B_2 \cap C_1$
(d) $C_1 | B_2 \cap A_1$.

Let's Make a Deal with Monty Hall

- One door hides a car, two hide goats.
- The contestant chooses any door.
- Monty always opens a different door with a goat. (He can do this because he knows where the car is.)
- The contestant is then allowed to switch doors if she wants.

What is the best strategy for winning a car?

(a) Switch (b) Don't switch (c) It doesn't matter



Board question: Monty Hall

Organize the Monty Hall problem into a tree and compute the probability of winning if you always switch.

Hint first break the game into a sequence of actions.

Independence

Events A and B are independent if the probability that one occurred is not affected by knowledge that the other occurred.

$$\begin{aligned} \text{Independence} &\Leftrightarrow P(A|B) = P(A) \quad (\text{provided } P(B) \neq 0) \\ &\Leftrightarrow P(B|A) = P(B) \quad (\text{provided } P(A) \neq 0) \end{aligned}$$

(For any A and B)

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

Table/Concept Question: Independence

(Work with your tablemates, then everyone click in the answer.)

Roll two dice and consider the following events

- A = 'first die is 3'
- B = 'sum is 6'
- C = 'sum is 7'

A is independent of

(a) B and C (b) B alone

(c) C alone (d) Neither B or C.

Bayes' Theorem

Also called Bayes' Rule and Bayes' Formula.

Allows you to find $P(A|B)$ from $P(B|A)$, i.e. to 'invert' conditional probabilities.

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

Often compute the denominator $P(B)$ using the law of total probability.

Board Question: Evil Squirrels

Of the one million squirrels on MIT's campus most are good-natured. But one hundred of them are pure evil! An enterprising student in Course 6 develops an "Evil Squirrel Alarm" which she offers to sell to MIT for a passing grade. MIT decides to test the reliability of the alarm by conducting trials.

Evil Squirrels Continued

- When presented with an evil squirrel, the alarm goes off 99% of the time.
- When presented with a good-natured squirrel, the alarm goes off 1% of the time.

(a) If a squirrel sets off the alarm, what is the probability that it is evil?

(b) Should MIT co-opt the patent rights and employ the system?

One solution
(This is a base rate fallacy problem)

We are given:

$$P(\text{nice}) = 0.9999, \qquad P(\text{evil}) = 0.0001\ (\text{base rate})$$

$$P(\text{alarm} \mid \text{nice}) = 0.01, \quad P(\text{alarm} \mid \text{evil}) = 0.99$$

$$\begin{aligned}
P(\text{evil} \mid \text{alarm}) &= \frac{P(\text{alarm} \mid \text{evil}) P(\text{evil})}{P(\text{alarm})} \\
&= \frac{P(\text{alarm} \mid \text{evil}) P(\text{evil})}{P(\text{alarm} \mid \text{evil}) P(\text{evil}) + P(\text{alarm} \mid \text{nice}) P(\text{nice})} \\
&= \frac{(0.99)(0.0001)}{(0.99)(0.0001) + (0.01)(0.9999)} \\
&\approx 0.01
\end{aligned}$$

Squirrels continued

Summary:

$$\text{Probability a random test is correct} = 0.99$$

$$\text{Probability a positive test is correct} \approx 0.01$$

These probabilities are not the same!

Alternative method of calculation:

	Evil	Nice	
Alarm	99	9999	10098
No alarm	1	989901	989902
	100	999900	1000000

Washington Post, hot off the press

Annual physical exam is probably unnecessary if you're generally healthy

> *For patients, the negatives include time away from work and possibly unnecessary tests. "Getting a simple urinalysis could lead to a false positive, which could trigger a cascade of even more tests, only to discover in the end that you had nothing wrong with you." Mehrotra says.*

http://www.washingtonpost.com/national/health-science/annual-physical-exam-is-probably-unnecessary-if-youre-generall
2013/02/08/2c1e326a-5f2b-11e2-a389-ee565c81c565_story.html

Table Question: Dice Game

1. The Randomizer holds the 6-sided die in one fist and the 8-sided die in the other.
2. The Roller selects one of the Randomizer's fists and covertly takes the die.
3. The Roller rolls the die in secret and reports the result to the table.

Given the reported number, what is the probability that the 6-sided die was chosen? (Find the probability for each possible reported number.)