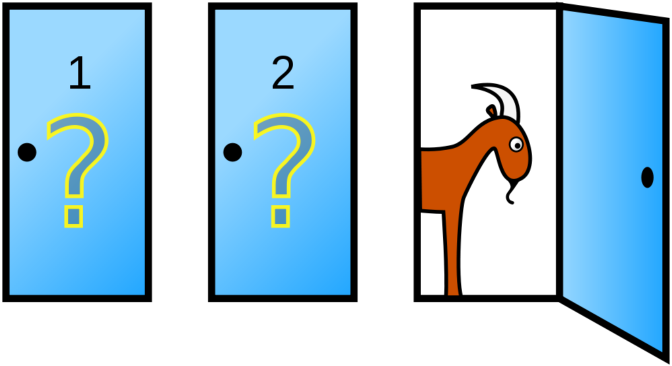
Conditional Probability, Independence, Bayes' Theorem

18.05 Spring 2018



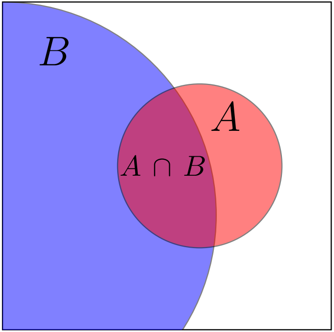
Slides are Posted

Don't forget that after class we post the slides including solutions to all the questions.

Conditional Probability

'the probability of A given B'.

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



$A = A \cap B$ $\quad B$

HHH	HHT	THH	THT
HTH	HTT	TTH	TTT

Conditional probability: Abstractly and for coin example

Table/Concept Question
(Work with your tablemates.)

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) $1/5$

answer: 1. (b) $1/8$. 2. (d) $1/5$.
Counting we find $|A| = 5$, $|B| = 8$ and $|A \cap B| = 1$. Since all sequences are equally likely

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{|A \cap B|}{|B|} = 1/8. \quad P(B|A) = \frac{|B \cap A|}{|A|} = 1/5.$$

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 *"Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality.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."**

Interesting

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

because P(farmer and shy) so low intuitivly

Discussion on next slide.

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

Discussion of Shy Steve

Discussion: Most people say that it is more likely that Steve is a librarian than a farmer. BUT for every male librarian in the United States there are about sixty male farmers. When this is explained, most people who chose librarian switch their solution to farmer. Suppose...

$$P(\text{shy}|\text{librarian}) = .8, \qquad P(\text{shy}|\text{farmer}) = .2$$

Says a librarian is four times as likely as a farmer to be shy).
Among 72,000,000 US male workers...

$$P(\text{librarian}) = .0005, \quad P(\text{farmer}) = .030, \quad P(\text{shy}) = .4$$

Says a US male is sixty times as likely to be a farmer as a librarian.

$$P(\text{farmer}|\text{shy}) = .015, \qquad P(\text{librarian}|\text{shy}) = .001$$

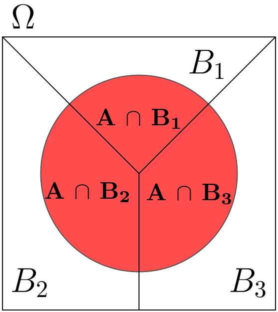
(CHECK THESE CALCULATIONS!) Conclusion is that a shy man is fifteen times as likely to be a farmer as a librarian. Farmer/lib = 60

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

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



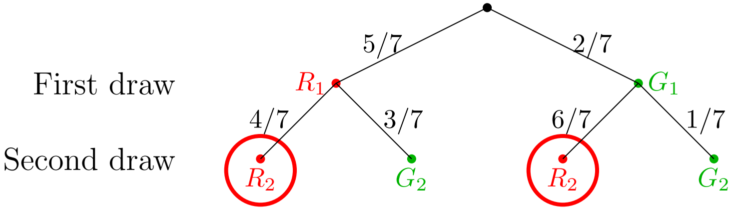
Trees

- 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?

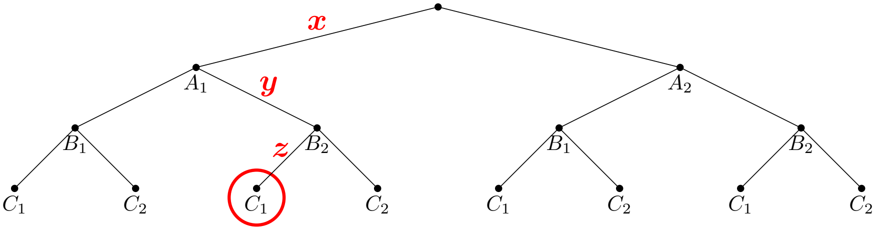


Solution

1. The law of total probability gives $P(R_2) = \frac{5}{7} \cdot \frac{4}{7} + \frac{2}{7} \cdot \frac{6}{7} = \frac{32}{49}$

2. Bayes' rule gives $P(R_1|R_2) = \frac{P(R_1 \cap R_2)}{P(R_2)} = \frac{20/49}{32/49} = \frac{20}{32}$

Concept Question: Trees 1

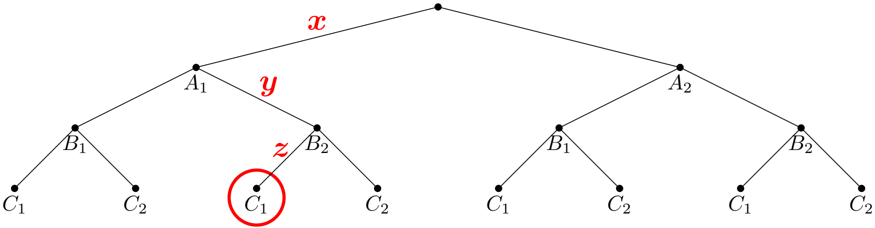


1. The probability x represents

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

answer: (a) $P(A_1)$.

Concept Question: Trees 2

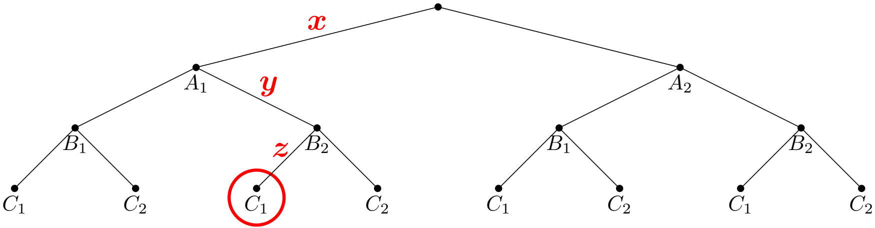


2. The probability y represents

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

answer: (c) $P(B_2|A_1)$.

Concept Question: Trees 3

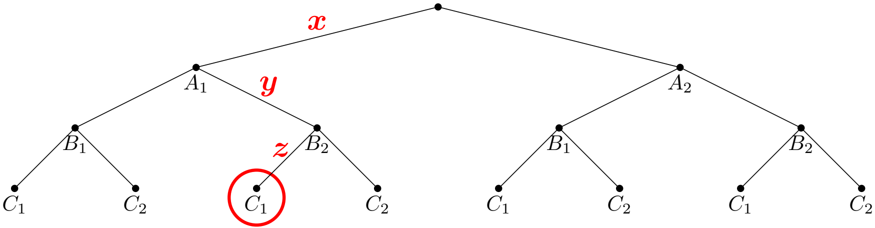


3. The probability z represents

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

answer: (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) $C_1 | B_2 \cap A_1$.

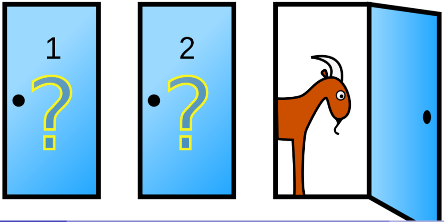
answer: (c) $A_1 \cap B_2 \cap C_1$.

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



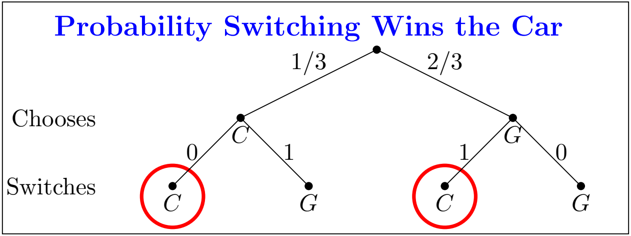
Board question: Monty Hall

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.

answer: Switch. $P(C|switch) = 2/3$

It's easiest to show this with a tree representing the switching strategy: First the contestant chooses a door, (then Monty shows a goat), then the contestant switches doors.



The (total) probability of C is $P(C|switch) = \frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = \frac{2}{3}$.

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. !!!

answer: (c). *(Explanation on next slide)*

Solution

$P(A) = 1/6$, $P(A|B) = 1/5$. Not equal, so not independent.

$P(A) = 1/6$, $P(A|C) = 1/6$. Equal, so independent.

Notice that knowing B, removes 6 as a possibility for the first die and makes A more probable. So, knowing B occurred changes the probability of A.

But, knowing C does not change the probabilities for the possible values of the first roll; they are still $1/6$ for each value. In particular, knowing C occured does not change the probability of A.

Could also have done this problem by showing

$$P(B|A) \neq P(B) \text{ or } P(A \cap B) \neq P(A)P(B).$$

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?

Solution on next slides.

One solution
(This is a base rate fallacy problem)

We are given:

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

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

$$P(\text{evil} \mid \text{alarm}) = \frac{P(\text{alarm} \mid \text{evil})P(\text{evil})}{P(\text{alarm})}$$

$$= \frac{P(\text{alarm} \mid \text{evil})P(\text{evil})}{P(\text{alarm} \mid \text{evil})P(\text{evil}) + P(\text{alarm} \mid \text{nice})P(\text{nice})}$$

$$= \frac{(0.99)(0.0001)}{(0.99)(0.0001) + (0.01)(0.9999)}$$

$$\approx 0.01$$

Squirrels continued

Summary:

$$\begin{aligned} \text{Probability a random test is correct} &= 0.99 \\ \text{Probability a positive test is correct} &\approx 0.01 \end{aligned}$$

These probabilities are not the same!

Alternative method of calculation:

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

Evil Squirrels Solution

answer: (a) This is the same solution as in the slides above, but in a more compact notation. Let E be the event that a squirrel is evil. Let A be the event that the alarm goes off. By Bayes' Theorem, we have:

$$\begin{aligned} P(E \mid A) &= \frac{P(A \mid E)P(E)}{P(A \mid E)P(E) + P(A \mid E^c)P(E^c)} \\ &= \frac{.99\frac{100}{1000000}}{.99\frac{100}{1000000} + .01\frac{999900}{1000000}} \\ &\approx .01. \end{aligned}$$

(b) No. The alarm would be more trouble than its worth, since for every true positive there are about 99 false positives.

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.)

<u>**answer:**</u> If the number rolled is 1-6 then $P(\text{six-sided}) = 4/7$.
If the number rolled is 7 or 8 then $P(\text{six-sided}) = 0$.

Explanation on next page

Dice Solution

This is a Bayes' formula problem. For concreteness let's suppose the roll was a 4. What we want to compute is $P(\text{6-sided}|\text{roll 4})$. But, what is easy to compute is $P(\text{roll 4}|\text{6-sided})$.
Bayes' formula says

$$\begin{aligned} P(\text{6-sided}|\text{roll 4}) &= \frac{P(\text{roll 4}|\text{6-sided})P(\text{6-sided})}{P(4)} \\ &= \frac{(1/6)(1/2)}{(1/6)(1/2) + (1/8)(1/2)} = 4/7. \end{aligned}$$

The denominator is computed using the law of total probability:

$$P(4) = P(4|\text{6-sided})P(\text{6-sided}) + P(4|\text{8-sided})P(\text{8-sided}) = \frac{1}{6} \cdot \frac{1}{2} + \frac{1}{8} \cdot \frac{1}{2}.$$

Note that any roll of 1,2,...6 would give the same result. A roll of 7 (or 8) would give clearly give probability 0. This is seen in Bayes' formula because the term $P(\text{roll 7}|\text{6-sided}) = 0$.