Conditional Probability, Independence, Bayes Theorem
18.05 Spring 2014
Jeremy Orloff and Jonathan Bloom

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: $\Omega =$ 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. $\Omega =$ 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, $\Omega =$ set of all sequences of n birthdays.
$|\Omega| = 365^n$.

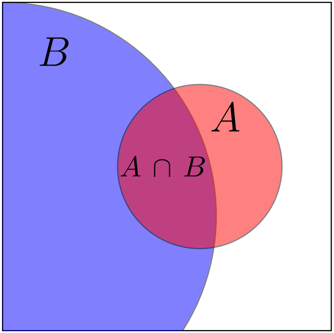
Slides are Posted

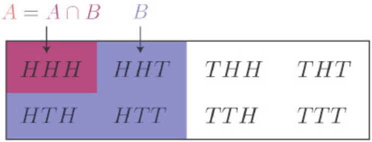
Don't forget that after class we post the slides including solutions to all the questions.

Conditional Probability

'the probability of A given B'.

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





Conditional probability: Abstractly and for coin example

Concept Question

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 *"Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality.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

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

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. Almost all people fail to consider that for every male librarian in the United States, there are more than fifty male farmers. When this is explained, most people who chose librarian switch their Solution to farmer.

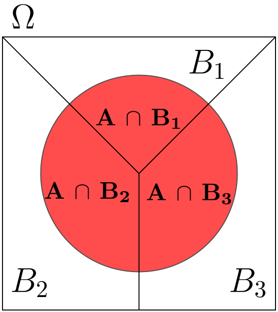
This illustrates how people often substitute representativeness for likelihood. The fact that a librarian may be likely to have the above personality traits does not mean that someone with these traits is likely to be a librarian.

Multiplication Rule, Law of Total Probability

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

LoTP: If B_1, B_2, B_3 *partition* Ω then

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3)$$
$$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)$$



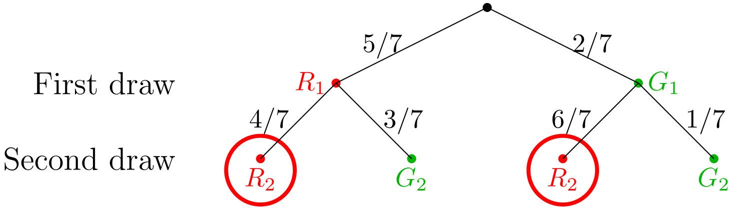
Trees

- 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.

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?

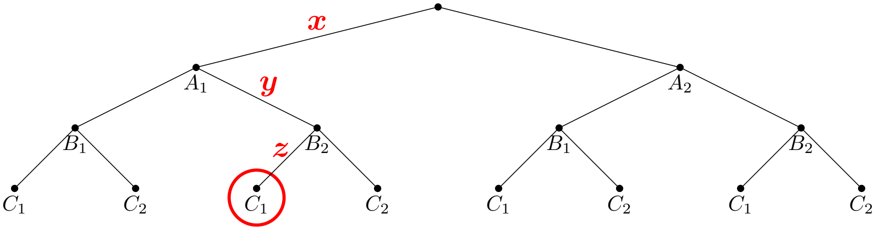


Solution

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

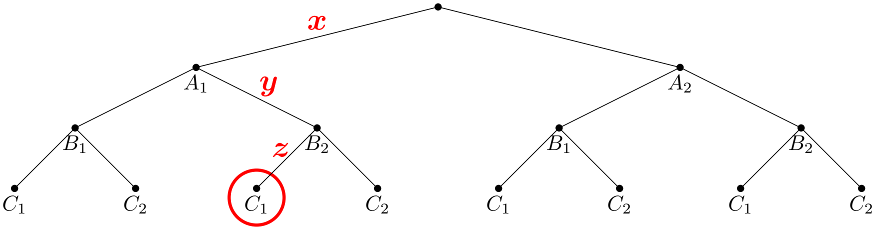


1. The probability x represents

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

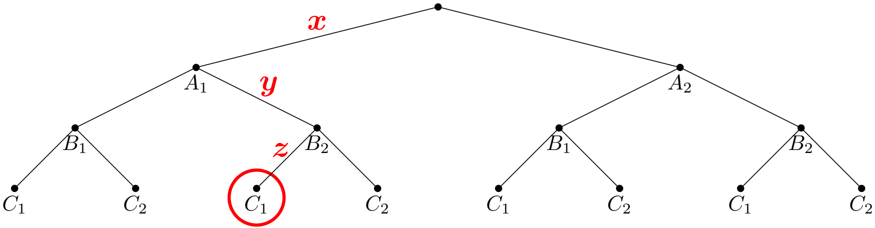


2. The probability y represents

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

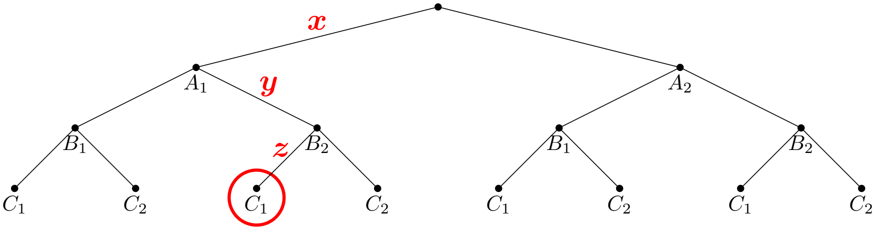


3. The probability z represents

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) $C_1|B_2 \cap A_1$.

<u>**answer:**</u> (c) $A_1 \cap B_2 \cap C_1$.

Concept question: Monty Hall

Let's Make a Deal

1. One door hides a car, two hide goats.
2. Contestant chooses door.
3. Monty (who knows where the car is) opens a different door with a goat.
4. Contestant can switch doors or keep her original choice.

What is the best strategy for winning a car?

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

Illustration of the Monty Hall problem removed due to copyright restrictions.

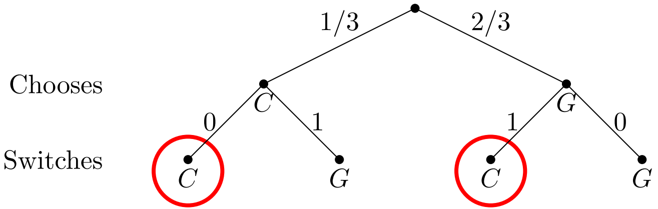
Board question: Monty Hall

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 process: contestant chooses (Monty shows a goat), contestant switches.



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.

$$\begin{aligned} A \text{ and } B \text{ are independent} \quad &\Leftrightarrow \; P(A \cap B) = P(A)P(B) \\ &\Leftrightarrow \; P(A|B) = P(A) \\ &\Leftrightarrow \; P(B|A) = P(B) \end{aligned}$$

Concept Question: Independence

Roll two dice.

$A =$ 'first die is 3'

$B =$ 'sum is 6'

A and B are independent

a) True   b) False

<u>**answer:**</u> False. $P(B) = 5/36$, $P(B|A) = 1/6$. Not equal $\Rightarrow$ not independent.

Could also show $P(A|B) \neq P(A)$ 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.

Photograph of squirrel removed due to copyright restrictions.

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

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

Summary:

$$\text{Probability a random test is correct} = \frac{99+989901}{1000000} = .99$$

$$\text{Probability a positive test is correct} = \frac{99}{10098} \approx .01$$

These probabilities are not the same!

Evil Squirrels Solution

answer: a) 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.

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-generally-2013/02/08/2c1e326a-5f2b-11e2-a389-ee565c81c565_story.html