A.2 Conditional Probability - Outline

Die Rolling Example

Example

Suppose you roll a fair 6-sided die. *Given that the result is odd,* what is the probability that it is 3 or less?

$$\text{P[roll 3 or less} \mid \text{odd]} = \frac{\text{\# of ways to roll 3 or less and odd}}{\text{\# ways to roll an odd number}} = \frac{\#\{1,3\}}{\#\{1,3,5\}} = \frac{2}{3}$$

$$\text{P[roll 3 or less} \mid \text{odd]} = \frac{\text{\# of ways to roll 3 or less and odd / 6}}{\text{\# ways to roll an odd number / 6}} = \frac{\text{P[roll 3 or less and odd]}}{\text{P[odd]}} = \frac{2/6}{3/6} = \frac{2}{3}$$

Variations on SOA #6

A public health researcher examines the medical records of a group of 635 men who died in 1999 and discovers that 160 of the men died from causes related to heart disease. Moreover, 275 of the 635 men had at least one parent who suffered from heart disease, and, of these 275 men, 95 died from causes related to heart disease.

Determine the probability that a man randomly selected from this group died of causes not related to heart disease and that neither of his parents suffered from heart disease.

HD	No HD		
95	180	At least 1 parent with heart disease	275
65	295	Neither parent with heart disease	360
160	475		

Answer: $\boxed{295/635}$

Variations on SOA #6

A public health researcher examines the medical records of a group of 635 men who died in 1999 and discovers that 160 of the men died from causes related to heart disease. Moreover, 275 of the 635 men had at least one parent who suffered from heart disease, and, of these 275 men, 95 died from causes related to heart disease.

Determine the probability that a man randomly selected from this group died of causes not related to heart disease, given that neither of his parents suffered from heart disease.

HD	No HD		
95	180	At least 1 parent with heart disease	275
65	295	Neither parent with heart disease	360
160	475		

Answer: $\boxed{295/360}$

Definitions

Definition (Conditional Probability)

The conditional probability of A given B is

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{P[AB]}{P[B]}$$

Note that rearranging terms gives

$$P[B] \cdot P[A \mid B] = P[AB] = P[A] \cdot P[B \mid A]$$

Definition (Independence)

A and B are *independent* if $P[AB] = P[A] \cdot P[B]$.

Intuitively, this means that $P[A] = P[A \mid B]$ and $P[B] = P[B \mid A]$ so knowing if A or B occured gives no information on whether or not the other event occured.

Variation on SOA #12; F.00.28

The blood pressure (high, low, or normal) and heartbeats (regular or irregular) of a random sample of patients are measured. Of the patients,

1. 36% have high blood pressure and 16% have low blood pressure.
2. 21% have an irregular heartbeat.
3. Of those with an irregular heartbeat, one-third have high blood pressure. $(1/3) \cdot 0.21 = 0.07$
4. Of those with normal blood pressure, one-eighth have an irregular heartbeat. $(1/8) \cdot 0.48 = 0.06$

What portion have a regular heartbeat and low blood pressure?

$\boxed{0.08}$	0.42	0.29	0.79	regular
0.08	0.06	0.07	0.21	irregular
0.16	0.48	0.36		
low	normal	high		

Key words

“Given that a person is in college, there is an 80% chance that they use Facebook”

means P[Facebook | college] = 0.80.

But they might not always use the word “given.”

“80% of college students use Facebook”

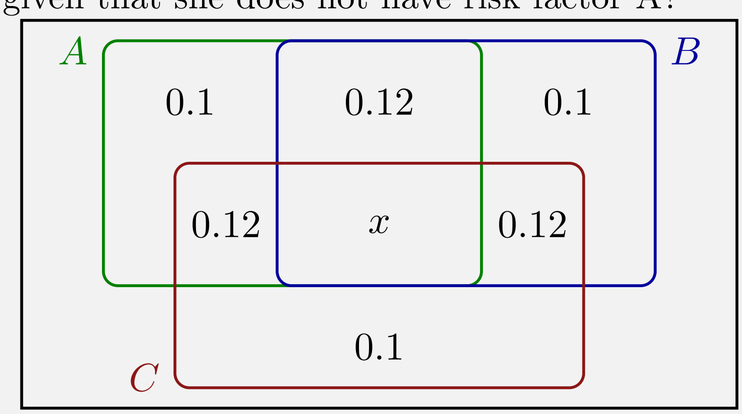
also means P[Facebook | college] = 0.80.

as does “College students have an 80% chance of using Facebook, while non-college students . . . ”

Language that restricts possible outcomes to one group/case means condition on that case.

SOA #13; S.01.09

An actuary is studying the prevalence of three health risk factors, denoted by A, B, and C, within a population of women. For each of the three factors, the probability is 0.1 that a woman in the population has only this risk factor (and no others). For any two of the three factors, the probability is 0.12 that she has exactly these two risk factors (but not the other). The probability that a woman has all three risk factors, given that she has A and B, is 1/3.
What is the probability that a woman has none of the three risk factors, given that she does not have risk factor A?



SOA #13; S.01.09

The probability that a woman has all three risk factors, given that she has A and B, is 1/3.

What is the probability that a woman has none of the three risk factors, given that she does not have risk factor A?

$$\begin{aligned}
\mathrm{P}[\text{all } 3 \mid AB] &= \frac{\mathrm{P}[\text{all } 3 \text{ and } AB]}{\mathrm{P}[AB]} \\
\frac{1}{3} &= \frac{x}{0.12 + x} \\
x &= 0.06 \\
\mathrm{P}[\text{none} \mid A'] &= \frac{\mathrm{P}[\text{none}]}{\mathrm{P}[A']} \\
&= \frac{1 - 3 \cdot 0.1 - 3 \cdot 0.12 - 0.06}{1 - 0.10 - 2 \cdot 0.12 - 0.06} \\
&= \frac{0.28}{0.60} = \boxed{7/15}
\end{aligned}$$

A.2 Conditional Probability - Outline

Key Idea

The definition of conditional probability was that

$$\mathrm{P}[B \mid A] = \frac{\mathrm{P}[AB]}{\mathrm{P}[A]}$$

We can use that to find $\mathrm{P}[AB]$. Clearing the denominator gives

$$\mathrm{P}[A] \cdot \mathrm{P}[B \mid A] = \mathrm{P}[AB]$$

This equation can be thought of as a sequence of events: first we need A to occur, and then second we need B to also occur, taking into account the fact that we know about A.

Probability of a flush

What is the probability of being dealt a flush after being dealt five cards from a standard deck?
(A flush is at least 5 cards from the same suit, and a standard deck has 4 suits each with 13 cards)
We want: $\mathrm{P}[\text{all 5 cards have the same suit}]$.
Method 1:

$$\begin{aligned}\mathrm{P}[\text{all 5 cards have the same suit}] &= \mathrm{P}[\text{all 5 are spades}] \\ &+ \mathrm{P}[\text{all 5 are hearts}] + \mathrm{P}[\text{all 5 are diamonds}] \\ &+ \mathrm{P}[\text{all 5 are clubs}] = 4 \cdot \mathrm{P}[\text{all 5 are spades}] \\ &= 4 \cdot \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} \cdot \frac{10}{49} \cdot \frac{9}{48} = 0.198\%\end{aligned}$$

Method 2: $\quad \dfrac{52}{52} \cdot \dfrac{12}{51} \cdot \dfrac{11}{50} \cdot \dfrac{10}{49} \cdot \dfrac{9}{48} = 0.198\%$

Probability of a flush

If exactly three of the first 5 cards dealt are spades, what is the probability of being dealt a flush in the first 7 cards?

$$
\begin{aligned}
P[\text{next 2 cards are spades}] \\
&= P[\text{6th card is a spade}] \cdot P[\text{7th is a spade} \mid \text{6th is a spade}] \\
&= \frac{13-3}{52-5} \cdot \frac{13-4}{52-6} \\
&= \frac{10}{47} \cdot \frac{9}{46} \\
&= 0.0416 = \boxed{4.16\%}
\end{aligned}
$$

Probability of a flush

If exactly four of the first 5 cards dealt are spades, what is the probability of being dealt a flush in the first 7 cards?

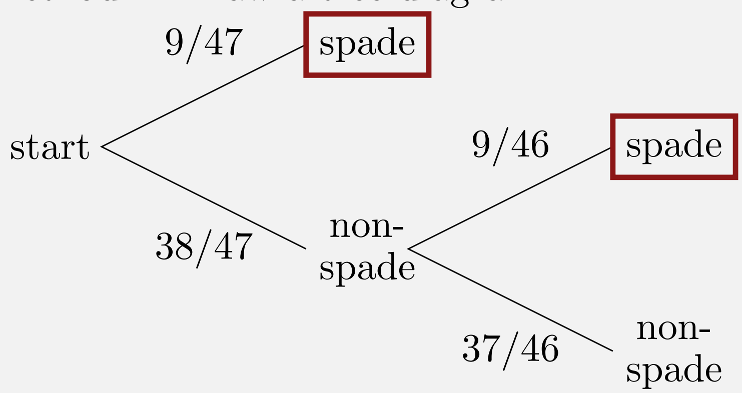
We want P[at least one of next two is a spade]

Method 1: P[at least one of next two is a spade] =
$1 - P[\text{neither is a spade}] = 1 - \frac{38}{47} \cdot \frac{37}{46} = 1 - 0.65 = 0.35$

Probability of a flush

If exactly four of the first 5 cards dealt are spades, what is the probability of being dealt a flush in the first 7 cards?

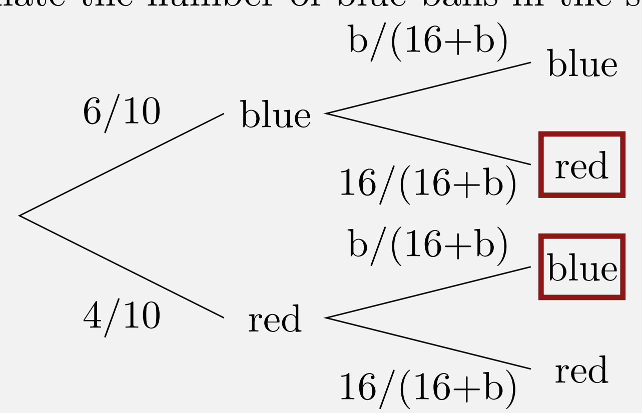
Method 2: Draw a tree diagram



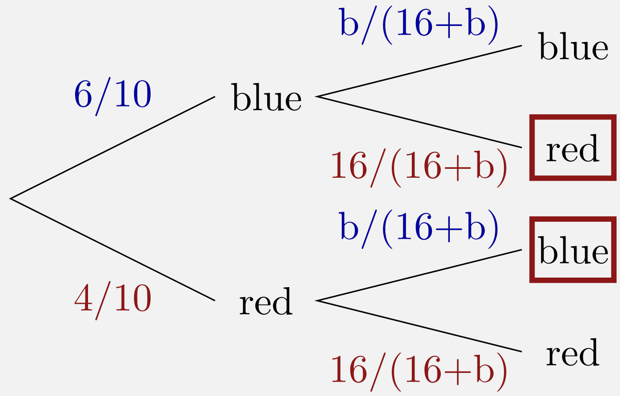
So $P[\text{flush}] = \frac{9}{47} + \frac{38}{47} \cdot \frac{9}{46} = \boxed{0.35}$

Urn Problem

An urn contains 10 balls: 4 red and 6 blue. A second urn contains 16 red balls and an unknown number of blue balls. A single ball is drawn from each urn. The probability that both balls are different colors is 0.528.
Calculate the number of blue balls in the second urn.



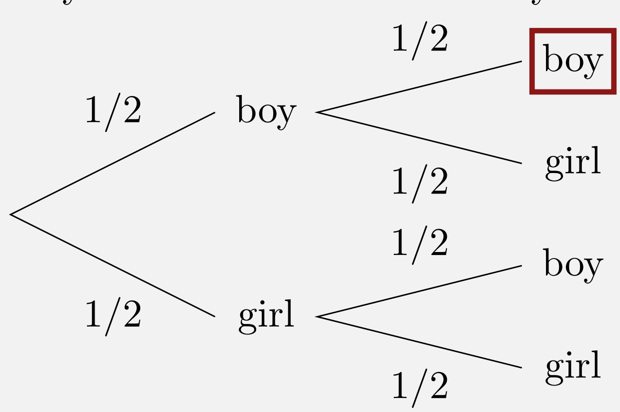
Urn Problem



$$0.528 = \frac{6}{10} \cdot \frac{16}{16+b} + \frac{4}{10} \cdot \frac{b}{16+b}$$
$$= \frac{96+4b}{160+10b}$$
$$b = \boxed{9}$$

Two Child Problem

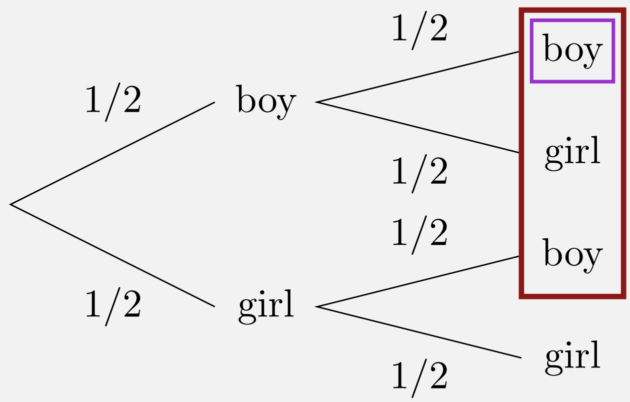
A family has two children, and they are not twins. What is the probability that both children are boys?



Answer: $\boxed{1/4}$

Two Child Problem

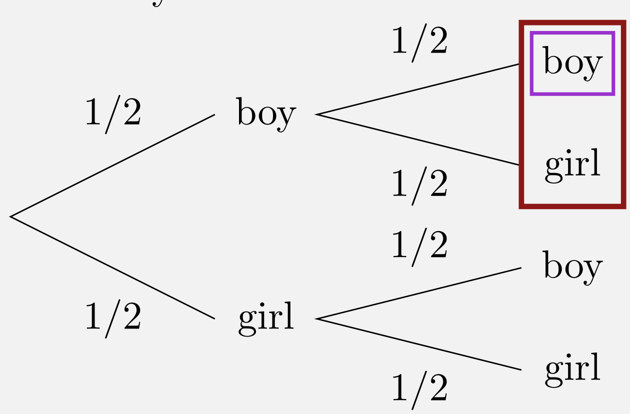
A family has two children, and they are not twins. Given that at least one of the children is a boy, what is the probability that both children are boys?



P[2 boys | at least 1 boy] = (1/4)/(3/4) = 1/3

Two Child Problem

A family has two children, and they are not twins. Given that the oldest child is a boy, what is the probability that both children are boys?



P[2 boys | oldest is a boy] = (1/4)/(2/4) = 1/2

A.2 Conditional Probability - Outline

Bayes' Theorem

We often want $\mathrm{P}[A \mid B]$ and are given $\mathrm{P}[B \mid A]$

To find it, we can use

$$\mathrm{P}[A \mid B] = \frac{\mathrm{P}[AB]}{\mathrm{P}[B]} = \frac{\mathrm{P}[A] \cdot \mathrm{P}[B \mid A]}{\mathrm{P}[B]}$$

Often, most of the work will be in finding $\mathrm{P}[B]$

Variation on SOA #19; S.03.08

An auto insurance company insures drivers of all ages. An actuary compiled the following statistics on the company's insured drivers:

Age of Driver	Probability of Accident	Portion of Company's Insured Drivers
16-20	0.06	0.08
21-30	0.03	0.15
31-65	0.02	0.49
66-99	0.04	0.28

A randomly selected driver that the company ensures has an accident. Calculate the probability that the driver was 31-65.

Variation on SOA #19; S.03.08

Age of Driver	Probability of Accident	Portion of Company's Insured Drivers
16-20	0.06	0.08
21-30	0.03	0.15
31-65	0.02	0.49
66-99	0.04	0.28

$$\mathrm{P}[\text{age 31-65} \mid \text{accident}] = \frac{\mathrm{P}[\text{age 31-65, accident}]}{\mathrm{P}[\text{accident}]}$$

$$= \frac{\mathrm{P}[\text{age 31-65}] \cdot \mathrm{P}[\text{accident} \mid \text{age 31-65}]}{\mathrm{P}[\text{accident}]} = \frac{(0.49)(0.02)}{\mathrm{P}[\text{accident}]}$$

$$\text{and } \mathrm{P}[\text{accident}] = \sum_{\text{age groups}} \mathrm{P}[\text{accident, age group}]$$

$$= (.08)(.06) + (.15)(.03) + (.49)(.02) + (.28)(.04) = 0.0303$$

$$\text{Combining both parts gives } \frac{0.49 \cdot 0.02}{0.0303} = \boxed{0.3234}$$

The Law of Total Probability

Theorem (Law of Total Probability)

If $A_1, A_2 \ldots, A_k$ *are disjoint and*
$P[A_1] + P[A_2] + \cdots + P[A_k] = 1$ *then*
$P[B] = P[BA_1] + P[BA_2] + \cdots + P[BA_k]$
$P[B] = P[A_1] \cdot P[B \mid A_1] + \cdots + P[A_k] \cdot P[B \mid A_k]$

The sets $A_1, \ldots, A_k$ are called a partition of the sample space. We will often refer to them as a list of all possible cases.

In the previous example, the age groups were the A_i, and B was the event of an accident.

Bayes' Theorem

Theorem (Bayes' Theorem)

Suppose $A_1, \ldots, A_k$ *are a partition of the sample space. Then*

$$\begin{aligned} P[A_1 \mid B] &= \frac{P[A_1 B]}{P[B]} \\ &= \frac{P[A_1] \cdot P[B \mid A_1]}{\sum_{i=1}^{k} P[BA_i]} \end{aligned}$$

$$P[A_1 \mid B] = \frac{P[A_1] \cdot P[B \mid A_1]}{\sum_{i=1}^{k} P[A_i] \cdot P[B \mid A_i]}$$

You can think of the final denominator as
$$\sum_{\text{cases}} \mathrm{P}[\text{case}] \cdot \mathrm{P}[B \mid \text{case}]$$

Variation on SOA #20; S.01.6

An insurance company issues life insurance policies in three separate categories: standard, preferred, and ultra-preferred. Of the company's policyholders, 50% are standard, 40% are preferred, and 10% are ultra-preferred. The probability of dying in the next year is 0.010 for each standard policyholder, 0.005 for each preferred policyholder, and 0.001 for each ultra-preferred policyholder.
A policyholder dies in the next year. What is the probability that the deceased policyholder was standard?

Let S denote someone who is standard.

$$\begin{aligned} \mathrm{P}[S \mid \text{died}] &= \frac{\mathrm{P}[\text{S and died}]}{\mathrm{P}[\text{died}]} \\ &= \frac{0.50 \cdot 0.010}{0.50 \cdot 0.010 + 0.40 \cdot 0.005 + 0.10 \cdot 0.001} \\ &= \frac{0.005}{0.0071} = \boxed{70.4\%} \end{aligned}$$

Taxicab Problem

Taxicabs in Crobuzon are all either green or blue. On Tuesday, a taxicab got into an accident. A witness to the accident thought that the cab involved was blue, and further tests showed that the witness has an 80% chance of correctly identifying the color of a taxicab, independently of its color.
If 100% of the taxicabs on the streets on Tuesday were green, what was the probability that the taxicab involved in the accident was blue?

Since none of the taxicabs were blue that night, clearly we know that the witness is wrong and the cab was not blue. So the answer is 0% (and in particular, not 80%)!

Taxicab Problem

Taxicabs in Crobuzon are all either green or blue. On Tuesday, a taxicab got into an accident. A witness to the accident thought that the cab involved was blue, and further tests showed that the witness has an 80% chance of correctly identifying the color of a taxicab, independently of its color.

If 85% of the taxicabs on the streets on Tuesday were green, what was the probability that the taxicab involved in the accident was blue?

$$\begin{aligned}
&\mathrm{P}[\text{cab was blue } | \text{witness said it was blue}] \\
&\qquad = \frac{\mathrm{P}[\text{cab was blue and witness said blue}]}{\mathrm{P}[\text{witness said blue}]} \\
&\qquad = \frac{0.15 \cdot 0.80}{0.15 \cdot 0.80 + 0.85 \cdot 0.20} \\
&\qquad = \frac{0.12}{0.29} = \boxed{41\%}
\end{aligned}$$