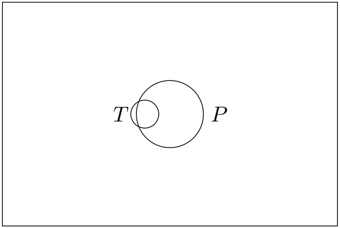
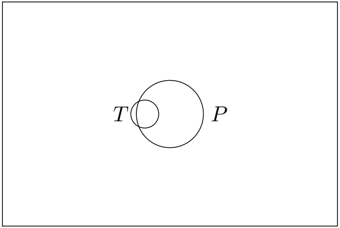
MAT1830

Lecture 22: Conditional probability and Bayes' theorem

Your friend believes that Python coding has become more popular than AFL in Melbourne. She bets you \$10 that the next person to pass you on the street will be a Python programmer. You feel confident about this bet. However, when you see a man in a "Hello, world!" t-shirt approaching, you don't feel so confident any more. Why is this?

We can think about this with a diagram. The rectangle represents the set of people in Melbourne, the circle P is the set of Python coders, and the circle T is the set of "Hello, world!" t-shirt owners.





Initially, you feel confident because the circle P takes up a small proportion of the rectangle. But when you learn that your randomly selected person is in the circle T, you feel bad because the circle P covers almost all of T. In mathematical language, the probability that a random Melbournian is a Python coder is low, but the probability that a random Melbournian is a Python coder given that they own a "Hello, world!" t-shirt is high.

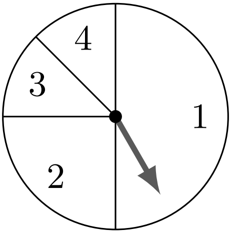
22.1 Conditional probability

Conditional probabilities measure the likelihood of an event, given that some other event occurs.

For events A and B, the *conditional probability of A given B* is
$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

This definition also implies that

$$\Pr(A \cap B) = \Pr(A|B)\Pr(B).$$



Example. The spinner from the last lecture is spun. Let A be the event that the result was at least 3 and B be the event that the result was even. What is $\Pr(A|B)$?

$$\begin{aligned}\Pr(A \cap B) = &\ \Pr(4) = \tfrac{1}{8} \\ \Pr(B) = &\ \Pr(2) + \Pr(4) = \tfrac{1}{4} + \tfrac{1}{8} = \tfrac{3}{8}\end{aligned}$$

Thus,

$$\Pr(A|B) = \tfrac{\Pr(A \cap B)}{\Pr(B)} = (\tfrac{1}{8})/(\tfrac{3}{8}) = \tfrac{1}{3}.$$

Example An ordered pair is selected uniformly at random from $\{1,2,3,4\} \times \{1,2,3,4\}$.

Let A be the event that its first coordinate is 3 or 4.

Let B be the event that the sum of its coordinates is 6.

There are 16 equally likely outcomes:

$$\begin{array}{cccc}
(1,1) & (1,2) & (1,3) & (1,4) \\
(2,1) & (2,2) & (2,3) & (2,4) \\
(3,1) & (3,2) & (3,3) & (3,4) \\
(4,1) & (4,2) & (4,3) & (4,4)
\end{array}$$

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{2/16}{3/16} = \frac{2}{3}$$

$$\Pr(B \mid A) = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{2/16}{8/16} = \frac{1}{4}$$

Example. A binary string of length 6 is generated uniformly at random. Let A be the event that the first bit is a 1 and B be the event that the string contains two 1s. What is $\Pr(A|B)$?

There are 2^6 strings in our sample space. Now $A \cap B$ occurs when the first bit is 1 and the rest of the string contains 1 one. There are $\binom{5}{1}$ such strings and so $\Pr(A \cap B) = \binom{5}{1}/2^6$. Also, there are $\binom{6}{2}$ strings containing two 1s and so $\Pr(B) = \binom{6}{2}/2^6$. Thus,

$$\Pr(A|B) = \tfrac{\Pr(A \cap B)}{\Pr(B)} = \tbinom{5}{1}/\tbinom{6}{2} = \tfrac{1}{3}.$$

Flux Exercise

A normal die is rolled. Let A be the event that the result is an odd number. Let B be the event that the result is a square number.

A. $\Pr(A \mid B) = 1/2$ and $\Pr(B \mid A) = 1/4$

B. $\Pr(A \mid B) = 1/2$ and $\Pr(B \mid A) = 1/3$

C. $\Pr(A \mid B) = 1/6$ and $\Pr(B \mid A) = 1/6$

D. $\Pr(A \mid B) = 2/3$ and $\Pr(B \mid A) = 1/4$

Answer B. We have $A = \{1, 3, 5\}$, $B = \{1, 4\}$ and $A \cap B = \{1\}$. So

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{1/6}{1/3} = 1/2$$

and

$$\Pr(B \mid A) = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{1/6}{1/2} = 1/3.$$

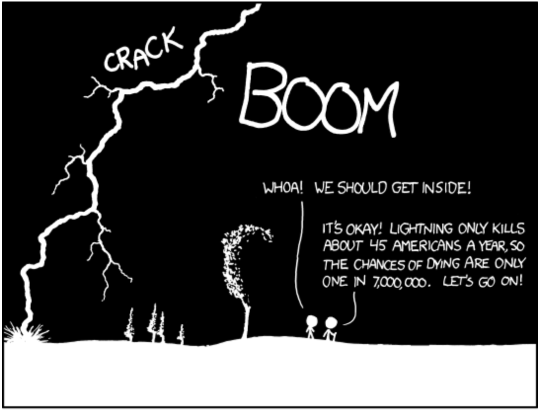
Question

22.1 An integer is selected uniformly at random from the set $\{1, 2, \ldots, 15\}$. What is the probability that it is divisible by 5, given that it is odd?

Answer

Let $A = \{5, 10, 15\}$ be the event that the number is divisible by 5.
Let $B = \{1, 3, 5, 7, 9, 11, 13, 15\}$ be the event that the number is odd.
So $A \cap B = \{5, 15\}$.

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{2/15}{8/15} = \frac{1}{4}.$$



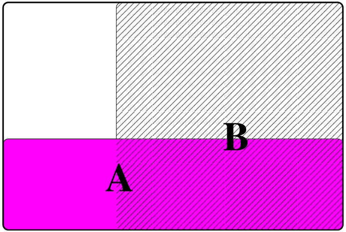
THE ANNUAL DEATH RATE AMONG PEOPLE WHO KNOW THAT STATISTIC IS ONE IN SIX.

22.2 Independence again

Our definition of conditional probability gives us another way of defining independence. We can say that events A and B are independent if

$$\Pr(A) = \Pr(A|B).$$

This makes sense intuitively: it is a formal way of saying that the likelihood of A does not depend on whether or not B occurs.



Flux Exercise

The previous flux question considered the events A (odd number) and B (square number) and we calculated that $\Pr(A) = 1/2$, $\Pr(B) = 1/3$, $\Pr(A \mid B) = 1/2$ and $\Pr(B \mid A) = 1/3$.

A. Event A is independent of Event B, and vice versa.
B. The events A and B are NOT independent because a die roll of 1 is included in both.
C. The event A is independent of B but the event B is NOT independent of A.
D. The event B is independent of A but the event A is NOT independent of B.

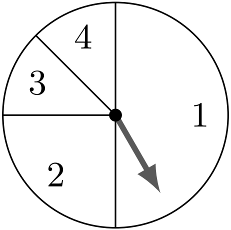
Answer A. Because
$\Pr(A \mid B) = 1/2 = \Pr(A)$ and
$\Pr(B \mid A) = 1/3 = \Pr(B)$.

(Either one of these is enough to show that the events are independent. Note that answers C and D are red herrings; the situations they describe are impossible regardless of what the events A and B are.)

22.3 Independent repeated trials

Generally if we perform exactly the same action multiple times, the results for each trial will be independent of the others. For example, if we roll a die twice, then the result of the first roll will be independent of the result of the second.

For two independent repeated trials, each from a sample space S, our overall sample space is $S \times S$ and our probability function will be given by $\Pr((s_1, s_2)) = \Pr(s_1)\Pr(s_2)$. For three independent repeated trials the sample space is $S \times S \times S$ and the probability function $\Pr((s_1, s_2, s_3)) = \Pr(s_1)\Pr(s_2)\Pr(s_3)$, and so on.



Example. The spinner from the previous example is spun twice. What is the probability that the results add to 5?
A total of 5 can be obtained as (1,4), (4,1), (2,3) or (3,2). Because the spins are independent:

$$\Pr((1,4)) = \Pr((4,1)) = \tfrac{1}{2} \times \tfrac{1}{8} = \tfrac{1}{16}$$
$$\Pr((2,3)) = \Pr((3,2)) = \tfrac{1}{4} \times \tfrac{1}{8} = \tfrac{1}{32}$$

So, because (1,4), (4,1), (2,3) and (3,2) are mutually exclusive, the probability of the total being 5 is $\frac{1}{16} + \frac{1}{16} + \frac{1}{32} + \frac{1}{32} = \frac{3}{16}$.

Question

22.2 A standard die is rolled twice. What is the probability that the first roll is a 1, given that the sum of the rolls is 6?

Answer

Let A be the event that the first roll is 1.
Let B be the event that the sum of the rolls is 6
(which happens as $5+1, 4+2, 3+3, 2+4, 1+5$).

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{1/36}{5/36} = \frac{1}{5}.$$

22.4 Bayes' theorem

Bayes' theorem gives a way of calculating the conditional probability of an event A given an event B when we already know the probabilities of A, of B given A, and of B given $\overline{A}$.

Bayes' theorem. For the events A and B,

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B|A)\Pr(A) + \Pr(B|\overline{A})\Pr(\overline{A})}.$$

Note that the denominator above is simply an expression for $\Pr(B)$. The fact that

$$\Pr(B) = \Pr(B|A)\Pr(A) + \Pr(B|\overline{A})\Pr(\overline{A})$$

is due to the *law of total probability*.

To prove Bayes' theorem, we need to show that

$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\Pr(A)}{\Pr(B \mid A)\Pr(A) + \Pr(B \mid \overline{A})\Pr(\overline{A})}.$$

$$\begin{aligned}
\frac{\Pr(B \mid A)\Pr(A)}{\Pr(B \mid A)\Pr(A) + \Pr(B \mid \overline{A})\Pr(\overline{A})} &= \frac{\Pr(B \cap A)}{\Pr(B \cap A) + \Pr(B \cap \overline{A})} \\
&= \frac{\Pr(B \cap A)}{\Pr(B)} \\
&= \Pr(A \mid B)
\end{aligned}$$

22.4 Bayes' theorem

Bayes' theorem gives a way of calculating the conditional probability of an event A given an event B when we already know the probabilities of A, of B given A, and of B given $\overline{A}$.

Bayes' theorem. For the events A and B,

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B|A)\Pr(A) + \Pr(B|\overline{A})\Pr(\overline{A})}.$$

Note that the denominator above is simply an expression for $\Pr(B)$. The fact that

$$\Pr(B) = \Pr(B|A)\Pr(A) + \Pr(B|\overline{A})\Pr(\overline{A})$$

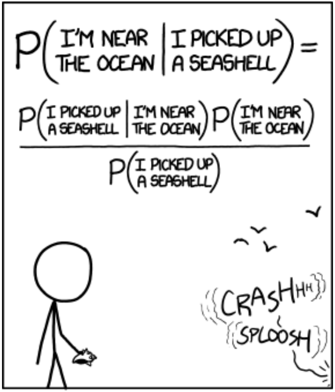
is due to the *law of total probability*.

22.5 Bayes' theorem examples

Example. Luke Skywalker discovers that some porgs have an extremely rare genetic mutation that makes them powerful force users. He develops a test for this mutation that is right 99% of the time and decides to test all the porgs on Ahch-To. Suppose there are 100 mutant porgs in the population of 24 million. We would guess that the test would come up positive for 99 of the 100 mutants, but also for 239 999 non-mutants.

We are assuming that the conditional probability of a porg testing positive given it's a mutant is 0.99. But what is the conditional probability of it being a mutant given that it tested positive? From our guesses, we would expect this to be $\frac{99}{99+239999} \approx 0.0004$. Bayes' theorem gives us a way to formalise this:

$$\begin{aligned}
\Pr(M|P) &= \frac{\Pr(P|M)\Pr(M)}{\Pr(P|M)\Pr(M)+\Pr(P|\overline{M})\Pr(\overline{M})} \\
&= \frac{\frac{100}{24000000}\times 0.99}{\frac{100}{24000000}\times 0.99+(1-\frac{100}{24000000})\times 0.01} \\
&= \frac{99}{99+239999} \\
&\approx 0.0004.
\end{aligned}$$



STATISTICALLY SPEAKING, IF YOU PICK UP A SEASHELL AND *DON'T* HOLD IT TO YOUR EAR, YOU CAN PROBABLY HEAR THE OCEAN.

Example. A binary string is created so that the first bit is a 0 with probability $\frac{1}{3}$ and then each subsequent bit is the same as the preceding one with probability $\frac{3}{4}$. What is the probability that the first bit is 0, given that the second bit is 0?

Let F be the event that the first bit is 0 and let S be the event that the second bit is 0. So $\Pr(F) = \frac{1}{3}$. If F occurs then the second bit will be 0 with probability $\frac{3}{4}$ and so $\Pr(S|F) = \frac{3}{4}$. If F does not occur then the second bit will be 0 with probability $\frac{1}{4}$ and so $\Pr(S|\overline{F}) = \frac{1}{4}$. So, by Bayes theorem,

$$\begin{aligned}
\Pr(F|S) &= \frac{\Pr(F)\Pr(S|F)}{\Pr(F)\Pr(S|F)+\Pr(\overline{F})\Pr(S|\overline{F})} \\
&= \frac{\frac{1}{3}\times\frac{3}{4}}{\frac{1}{3}\times\frac{3}{4}+\frac{2}{3}\times\frac{1}{4}} \\
&= (\tfrac{1}{4})/(\tfrac{5}{12}) \\
&= \tfrac{3}{5}.
\end{aligned}$$

Question A binary string is created so that the 1st bit is 0 with probability $\frac{1}{3}$ and then each subsequent bit is the same as the preceding one with probability $\frac{3}{4}$. What is the probability that the 1st bit is 0, given that the 2nd bit is 0?

Answer

Let F be the event that the first bit is 0.
Let S be the event that the second bit is 0.

$\Pr(F) = \frac{1}{3}$
$\Pr(S|F) = \frac{3}{4}$ because if F occurs then the 2nd bit will be 0 with probability $\frac{3}{4}$
$\Pr(S|\overline{F}) = \frac{1}{4}$ because if $\overline{F}$ occurs then the 2nd bit will be 0 with probability $\frac{1}{4}$

$$\begin{aligned}
\Pr(F|S) &= \frac{\Pr(F)\Pr(S|F)}{\Pr(F)\Pr(S|F) + \Pr(\overline{F})\Pr(S|\overline{F})} \\
&= \frac{\frac{1}{3} \times \frac{3}{4}}{\frac{1}{3} \times \frac{3}{4} + \frac{2}{3} \times \frac{1}{4}} \\
&= (\tfrac{1}{4})/(\tfrac{5}{12}) \\
&= \tfrac{3}{5}
\end{aligned}$$

Flux Exercise

A bag contains three black marbles and two white marbles and they are randomly selected and removed, one at a time until the bag in empty. Use Bayes' theorem to calculate the probability that the first marble selected is black, given that the second marble selected is black.

A. 3/4
B. 1/2
C. 2/5
D. 3/5.

Answer

Let A be the event that the first marble is black.
Let B be the event that the second marble is black.

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B|A)\Pr(A) + \Pr(B|\overline{A})\Pr(\overline{A})} = \frac{\frac{1}{2} \times \frac{3}{5}}{\frac{1}{2} \times \frac{3}{5} + \frac{3}{4} \times \frac{2}{5}} = \frac{1}{2}.$$