3.2 Conditional probability

$$P(B|A) = \frac{P(A\cap B)}{P(A)} \quad \text{(defn)}$$

mult. rule: $P(A\cap B) = P(A)\,P(B|A)$

can extend. e.g.:

$$P(A\cap B\cap C\cap D)$$
$$= P(A)\,P(B|A)\,P(C|A\cap B)\,P(D|A\cap B\cap C)$$

More general mult. rule (Axiom IV):

$$P(A\cap B|C) = P(A|C)\,P(B|A\cap C)$$

Example 2a Joe is 80 percent certain that his missing key is in one of the two pockets of his hanging jacket, being 40 percent certain it is in the left-hand pocket and 40 percent certain it is in the right-hand pocket. If a search of the left-hand pocket does not find the key, what is the conditional probability that it is in the other pocket?

L = "Key is in left pocket."

R = " " right " "

$P(L) = 0.4$

$P(R) = 0.4$

$P(R|L^c) = ?$

$$P(R \mid L^c) = \frac{P(R \cap L^c)}{P(L^c)}$$

Since $R \Rightarrow L^c$, that means $R \subset L^c$.

So $R \cap L^c = R$

$P(R \cap L^c) = P(R) = 0.4$

$P(L^c) = 1 - P(L) = 1 - 0.4 = 0.6$

$$P(R \mid L^c) = \frac{0.4}{0.6} = \boxed{\frac{2}{3}}$$

Example 2b A coin is flipped twice. Assuming that all four points in the sample space ~~S~~ Ω $= \{(h,h), (h,t), (t,h), (t,t)\}$ are equally likely, what is the conditional probability that both flips land on heads, given that (a) the first flip lands on heads? (b) at least one flip lands on heads?

A = "The first flip lands on heads."

B = "At least one flip lands on heads."

C = "Both flips land on heads."

$\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$

P = equally likely

(a) $P(C \mid A) = ?$ (b) $P(C \mid B) = ?$

$A = \{(H,H), (H,T)\}$

$B = \{(H,H), (H,T), (T,H)\}$

$C = \{(H,H)\}$

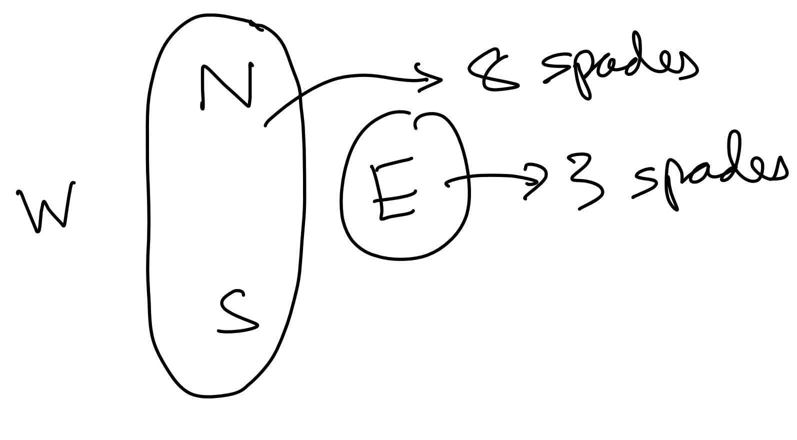
(a) $$P(C|A) = \frac{P(C \cap A)}{P(A)} = \frac{P(\{(H,H)\})}{P(A)}$$

$$= \frac{\left(\frac{1}{4}\right)}{\left(\frac{2}{4}\right)} = \boxed{\frac{1}{2}}$$

(b) $$P(C|B) = \frac{P(C \cap B)}{P(B)} = \frac{P(\{(H,H)\})}{P(B)}$$

$$= \frac{\left(\frac{1}{4}\right)}{\left(\frac{3}{4}\right)} = \boxed{\frac{1}{3}}$$

Example 2c In the card game bridge, the 52 cards are dealt out equally to 4 players—called East, West, North, and South. If North and South have a total of 8 spades among them, what is the probability that East has 3 of the remaining 5 spades?



[Consider N/S a single player to simplify the situation.]

F = "N/S get 8 spades."

E = "E gets 3 spades."

$\Omega = \{$all ways to choose 26 cards for N/S and 13 cards for E$\}$

P = equally likely

$P(E|F) = ?$

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

$$P(F) = \frac{|F|}{|\Omega|}$$

$$|\Omega| = \underbrace{\binom{52}{26}}_{\text{give N/S 26 cards}} \underbrace{\binom{26}{13}}_{\text{give E 13 of the remaining cards}}$$

$$|F| = \binom{13}{8}\binom{39}{18}\binom{26}{13}$$

give N/S 8 spades

give N/S 18 non-spades

give E 13 of the remaining cards

$$|E \cap F| = \binom{13}{8}\binom{39}{18}\binom{5}{3}\binom{21}{10}$$

give E 3 of the remaining spades

give E 10 of the remaining non-spades

$$P(F) = \frac{|F|}{|\Omega|} = \frac{\binom{13}{8}\binom{39}{18}\cancel{\binom{26}{13}}}{\binom{52}{26}\cancel{\binom{26}{13}}}$$

$$P(E \cap F) = \frac{|E \cap F|}{|\Omega|} = \frac{\binom{13}{8}\binom{39}{18}\binom{5}{3}\binom{21}{10}}{\binom{52}{26}\binom{26}{13}}$$

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

$$= \frac{\cancel{\binom{13}{8}}\cancel{\binom{39}{18}}\binom{5}{3}\binom{21}{10}}{\cancel{\binom{52}{26}}\binom{26}{13}} \cdot \frac{\cancel{\binom{52}{26}}}{\cancel{\binom{13}{8}}\cancel{\binom{39}{18}}}$$

$$= \frac{\binom{5}{3}\binom{21}{10}}{\binom{26}{13}} = \frac{5!}{3!\,2!} \cdot \frac{21!}{10!\,11!} \cdot \frac{13!\,13!}{26!}$$

$$= \frac{\cancel{5}\cdot\cancel{4}}{\cancel{2}} \cdot \frac{(\cancel{13}\cdot\cancel{12}\cdot\cancel{11})\cdot(13\cdot\cancel{12}^{3})}{\cancel{26}_{\cancel{2}}\cdot\cancel{25}_{5}\cdot\cancel{24}_{\cancel{2}}\cdot 23\cdot\cancel{22}_{\cancel{2}}}$$

$$= \frac{13\cdot 3}{5\cdot 23} = \boxed{\frac{39}{115}}$$

Example 2d Celine is undecided as to whether to take a French course or a chemistry course. She estimates that her probability of receiving an A grade would be $\frac{1}{2}$ in a French course and $\frac{2}{3}$ in a chemistry course. If Celine decides to base her decision on the flip of a fair coin, what is the probability that she gets an A in chemistry?

C = "Celine takes chemistry."

A = "Celine gets an A in her selected course."

$P(A|C^c) = \frac{1}{2}$

$P(A|C) = \frac{2}{3}$

$P(C) = \frac{1}{2}$

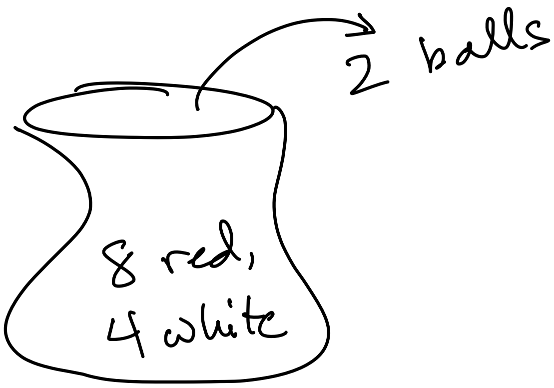
$P(C \cap A) = ?$

$$P(C \cap A) = P(C)\,P(A|C)$$

mult. rule

$$= \frac{1}{2} \cdot \frac{2}{3} = \boxed{\frac{1}{3}}$$

Example 2e Suppose that an urn contains 8 red balls and 4 white balls. We draw 2 balls from the urn without replacement. (a) If we assume that at each draw, each ball in the urn is equally likely to be chosen, what is the probability that both balls drawn are red? (b) Now suppose that the balls have different weights, with each red ball having weight r and each white ball having weight w. Suppose that the probability that a given ball in the urn is the next one selected is its weight divided by the sum of the weights of all balls currently in the urn. Now what is the probability that both balls are red?



R_1 = "First ball is red."

R_2 = "Second " "

$P(R_1 \cap R_2) = ?$

(a) $P(R_1) = \dfrac{8}{8+4} = \dfrac{8}{12} = \dfrac{2}{3}$

$P(R_2 | R_1) = \dfrac{7}{7+4} = \dfrac{7}{11}$

$$P(R_1 \cap R_2) = P(R_1)\,P(R_2 | R_1)$$

$$= \frac{2}{3} \cdot \frac{7}{11} = \boxed{\frac{14}{33}}$$

(b) $P(R_1) = \dfrac{8r}{8r + 4w}$

$P(R_2 | R_1) = \dfrac{7r}{7r + 4w}$

$$P(R_1 \cap R_2) = P(R_1)\,P(R_2 | R_1)$$

$$= \frac{{}^{2}\cancel{8}r}{{}^{2}\cancel{8}r + \cancel{4}w} \cdot \frac{7r}{7r + 4w}$$

$$= \boxed{\frac{14r^2}{(2r+w)(7r+4w)}}$$

Example 2g An ordinary deck of 52 playing cards is randomly divided into 4 piles of 13 cards each. Compute the probability that each pile has exactly 1 ace.

$A =$ "All four aces are in separate piles."

$P(A) = ?$

SOLN 1

$B =$ "The aces of spades, hearts & diamonds are in different piles."

$C =$ "The aces of spades and hearts are in different piles."

$D =$ "The ace of spades is in a pile."

$$A \Rightarrow B \Rightarrow C \Rightarrow D$$

i.e. $A \subset B \subset C \subset D$

$$\therefore \quad A = A \cap B \cap C \cap D$$

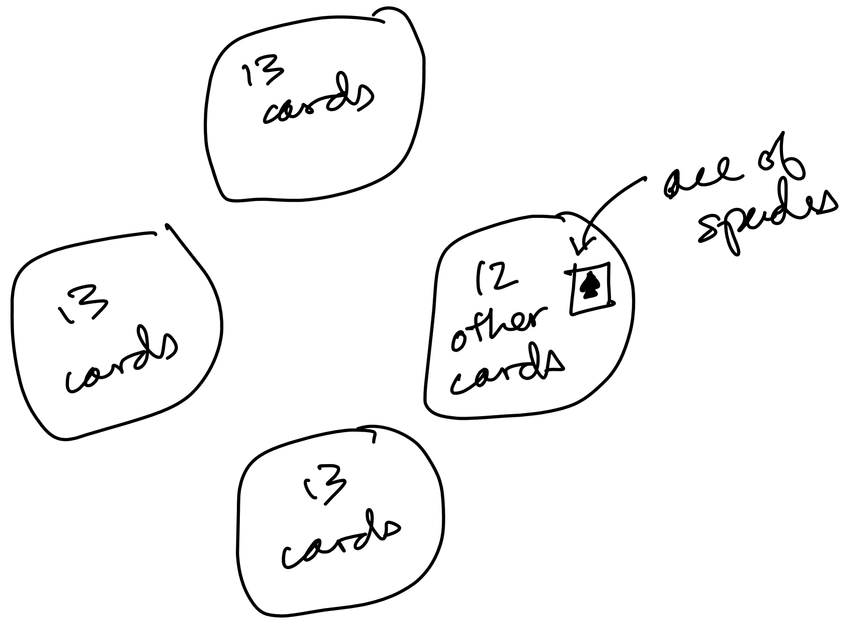
$$P(A) = P(A \cap B \cap C \cap D)$$

$$= P(D)\,P(C \mid D)\,P(B \mid \overset{C}{\cancel{C \cap D}})\,P(A \mid \overset{B}{\cancel{B \cap C \cap D}})$$

$P(A) = P(D)\, P(C|D)\, (B|C)\, P(A|B)$

$P(D) = 1$

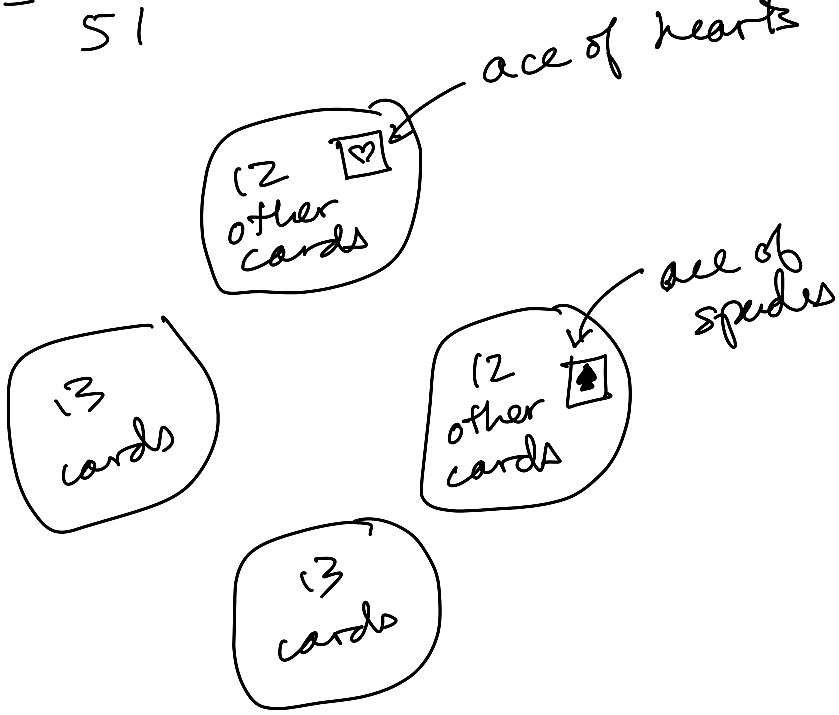
$P(C|D) = ?$



ace of hearts equally likely to be any of the other 51 cards. 39 of those are in a different pile.

$$P(C|D) = \frac{39}{51}$$

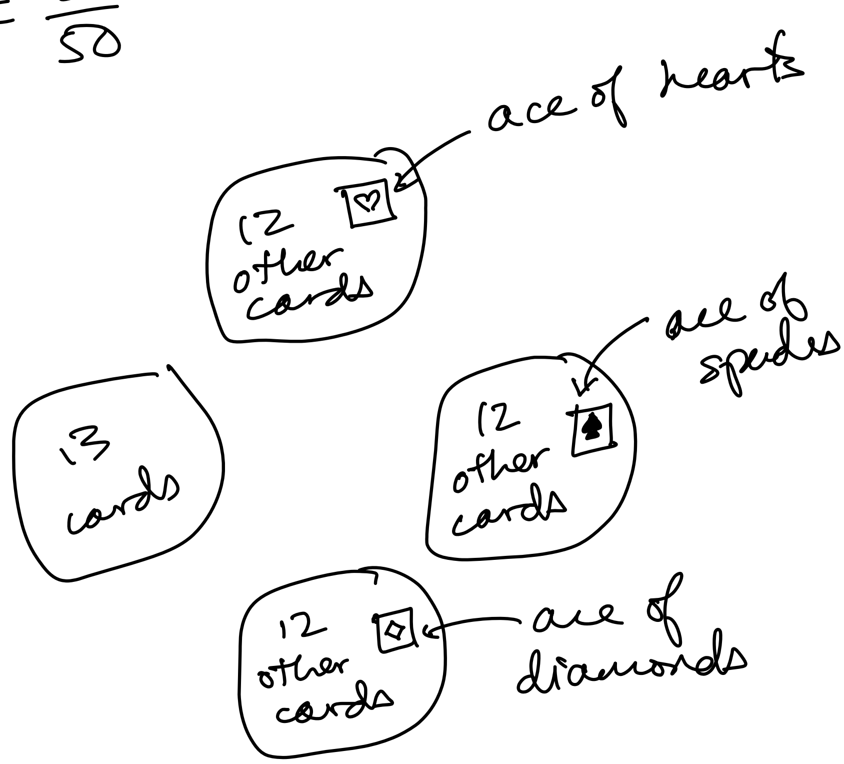
$P(B|C) = ?$



ace of diamonds equally likely to be any of the other 50 cards. 26 of those are in a different pile.

$$P(B|C) = \frac{26}{50}$$

$$P(A|B) = ?$$



ace of clubs equally likely to be any of the other 49 cards. 13 of those are in a different pile.

$$P(A|B) = \frac{13}{49}$$

$$P(A) = 1 \cdot \frac{\overset{13}{\cancel{39}}}{\underset{17}{\cancel{51}}} \cdot \frac{\overset{13}{\cancel{26}}}{\underset{25}{\cancel{50}}} \cdot \frac{13}{49} = \frac{13^3}{17 \cdot 25 \cdot 49}$$

$$= \boxed{\frac{2197}{20825}}$$

SOLN 2

(the previous soln illustrated the topic in this section, but here is another way)

$\Omega = \{$all ways to split 52 cards into 4 piles, numbered 1–4$\}$

$$P(A) = \frac{|A|}{|\Omega|}$$

$$|\Omega| = \binom{52}{13,13,13,13} = \frac{52!}{(13!)^4}$$

$$|A| = 4! \binom{48}{12,12,12,12} = \frac{4!\,48!}{(12!)^4}$$

4!: place each ace in its own pile

$\binom{48}{12,12,12,12}$: place the remaining cards into the piles

$$P(A) = \frac{4!\,48!\,(13!)^4}{(12!)^4\,52!} = \frac{\cancel{4}\cdot\cancel{3}\cdot\cancel{2}\cdot 13^{\cancel{4}\,3}}{\underset{17}{\cancel{52}}\cdot\cancel{51}\cdot\underset{25}{\cancel{50}}\cdot 49}$$

$$= \frac{13^3}{17\cdot 25\cdot 49} = \boxed{\frac{2197}{20825}}$$

3.3 Bayes' Formula

$$A = \underbrace{(A \cap B) \cup (A \cap B^c)}_{\text{mut. excl.}}$$

$$P(A) = P(A \cap B) + P(A \cap B^c) \quad \text{mult. rule}$$

$$P(A) = P(B)P(A|B) + P(B^c)P(A|B^c)$$

More generally...

$B_1, \ldots, B_n$ form a <u>partition</u> if they are mut. excl.

and $P(B_1 \cup B_2 \cup \cdots \cup B_n) = 1$

If $B_1, \ldots, B_n$ form a partition, then

$$\boxed{P(A) = \sum_{j=1}^{n} P(B_j)P(A|B_j)}$$

<u>Law of Total Probability</u>

---

odds of A: $O(A) = \dfrac{P(A)}{P(A^c)}$

conditional odds of A: $O(A|B) = \dfrac{P(A|B)}{P(A^c|B)}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \xrightarrow[\text{with odds}]{\text{rewrite}} O(A|B) = O(A) \cdot \underbrace{\frac{P(B|A)}{P(B|A^c)}}_{\text{likelihood ratio}}$$

Example 3a **(Part 1)**

An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. The company's statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability .4, whereas this probability decreases to .2 for a person who is not accident prone. If we assume that 30 percent of the population is accident prone, what is the probability that a new policyholder will have an accident within a year of purchasing a policy?

A = "The new guy has an accident in the 1^{st} year."

B = "The new guy is accident-prone."

$P(A|B) = 0.4$

$P(A|B^c) = 0.2$

$P(B) = 0.3$

$P(A) = ?$

$$P(A) = P(B)P(A|B) + P(B^c)P(A|B^c)$$
$$= 0.3(0.4) + 0.7(0.2)$$
$$= 0.12 + 0.14$$
$$= \boxed{0.26}$$

Example 3a **(Part 2)**

Suppose that a new policyholder has an accident within a year of purchasing a policy. What is the probability that he or she is accident prone?

A = "The new guy has an accident in the 1st year."

B = "The new guy is accident-prone."

$P(A|B) = 0.4$

$P(A|B^c) = 0.2$

$P(B) = 0.3$

$P(B|A) = ?$

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

$$\underset{\text{mult. rule}}{=} \frac{P(B)P(A|B)}{P(A)}$$

This "trick" effectively allowed us to switch "B|A" to "A|B".

P(B|A) is unknown, but P(A|B) is given.

$$= \frac{0.3(0.4)}{0.26} = \frac{12}{26} = \boxed{\frac{6}{13}}$$

from Part 1

Example 3c In answering a question on a multiple-choice test, a student either knows the answer or guesses. Let p be the probability that the student knows the answer and $1 - p$ be the probability that the student guesses. Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that he or she answered it correctly?

A = "Student knew the answer."

B = " " got the right answer."

From the teacher's POV:

$P(A) = p$ — prob. that A is true, given what the teacher knows, before grading the test

$P(B \mid A) = 1$

$P(B \mid A^c) = \frac{1}{m}$

$P(A \mid B) = ?$ — prob. that A is true, after grading and seeing the student got it right.

want to switch these around

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\overset{p}{\cancel{P(A)}}\ \overset{1}{\cancel{P(B \mid A)}}}{P(B)}$$

$$= \frac{p}{P(B)}$$

$$P(B) = P(A)P(B \mid A) + P(A^c)P(B \mid A^c)$$

$$= p \cdot 1 + (1-p) \cdot \frac{1}{m} = \frac{mp + 1 - p}{m}$$

$$P(A|B) = \frac{p}{\left(\frac{mp+1-p}{m}\right)} = \boxed{\frac{mp}{mp+1-p}}$$

Makes sense? (intuition check)

- p small, m moderate:

  $$P(A|B) \approx \frac{mp}{mp+1} \leftarrow \text{still small}$$

  If I knew the student was very unlikely to know the answer, a moderate number of alternatives would not provide a good indication.

- p close to 1:

  $$P(A|B) \approx \frac{m}{m} = 1$$

  If I knew the student was very likely to know the answer, the test won't tell me much.

- m very large:

  $$P(A|B) \approx \frac{mp}{mp} = 1$$

  If there are many alternatives and they get it right, then it's very likely they knew the answer.

Example 3d A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present. However, the test also yields a "false positive" result for 1 percent of the healthy persons tested. (That is, if a healthy person is tested, then, with probability .01, the test result will imply that he or she has the disease.) If .5 percent of the population actually has the disease, what is the probability that a person has the disease given that the test result is positive?

A = "The person has the disease."

B = "The person tests positive."

$P(B|A) = 0.95$

$P(B|A^c) = 0.01$

$P(A) = 0.5\% = 0.005$

$P(A|B) = ?$

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{\overset{0.005}{\cancel{P(A)}}\ \overset{0.95}{\cancel{P(B|A)}}}{P(B)}$$

$$= \frac{0.005(0.95)}{\underset{0.005}{\cancel{P(A)}}\ \underset{0.95}{\cancel{P(B|A)}} + \underset{0.995}{\cancel{P(A^c)}}\ \underset{0.01}{\cancel{P(B|A^c)}}}$$

$$= \frac{\left(\frac{1}{200}\right)\left(\frac{19}{20}\right)}{\frac{1}{200}\cdot\frac{19}{20} + \frac{199}{200}\cdot\frac{1}{100}}$$

$$P(A|B) = \frac{\left(\frac{1}{200}\right)\left(\frac{19}{20}\right)}{\frac{1}{200}\cdot\frac{19}{20} + \frac{199}{200}\cdot\frac{1}{100}} \cdot \frac{200(100)}{200(100)}$$

$$= \frac{19(5)}{19(5)+199}$$

$$= \frac{95}{95+199} = \boxed{\frac{95}{294}} \approx 32.3\%$$

Before doing the problem, most people think $P(A|B)$ will be much higher.

Example 3f At a certain stage of a criminal investigation, the inspector in charge is 60 percent convinced of the guilt of a certain suspect. Suppose, however, that a *new* piece of evidence which shows that the criminal has a certain characteristic (such as left-handedness, baldness, or brown hair) is uncovered. If 20 percent of the population possesses this characteristic, how certain of the guilt of the suspect should the inspector now be if it turns out that the suspect has the characteristic?

A = "The suspect is guilty."

B = "The criminal is left-handed."

I just picked one for concreteness.

$$P(A) = 0.6$$
$$P(B \mid A^c) = 0.2$$

These are probabilities from the investigator's POV, before discovering the handedness of the criminal

Assumption: if the suspect is innocent, the criminal is equally likely to be any other member of the population

$$P(A \mid B) = ?$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\overset{0.6}{\cancel{P(A)}}\, P(B \mid A)}{P(B)}$$

$$P(B \mid A) = 1$$

We must assume that it is known (or can be easily verified) that the suspect is left-handed.

$$P(B) = \underbrace{\cancel{P(A)}}_{0.6}\, \underbrace{\cancel{P(B \mid A)}}_{1} + \underbrace{\cancel{P(A^c)}}_{0.4}\, \underbrace{\cancel{P(B \mid A^c)}}_{0.2}$$

$$P(B) = 0.6 + 0.08 = 0.68$$

$$P(A \mid B) = \frac{0.6(1)}{0.68} = \frac{60}{68} = \boxed{\frac{15}{17}}$$

Example 3g In the world bridge championships held in Buenos Aires in May 1965, the famous British bridge partnership of Terrence Reese and Boris Schapiro was accused of cheating by using a system of finger signals that could indicate the number of hearts held by the players. Reese and Schapiro denied the accusation, and eventually a hearing was held by the British bridge league. The hearing was in the form of a legal proceeding with prosecution and defense teams, both having the power to call and cross-examine witnesses. During the course of the proceeding, the prosecutor examined specific hands played by Reese and Schapiro and claimed that their playing these hands was consistent with the hypothesis that they were guilty of having illicit knowledge of the heart suit. At this point, the defense attorney pointed out that their play of these hands was also perfectly consistent with their standard line of play. However, the prosecution then argued that as long as their play was consistent with the hypothesis of guilt, it must be counted as evidence toward that hypothesis. What do you think of the reasoning of the prosecution?

A = "They cheated."

E = "They played hand #7 according to the transcript submitted to the British bridge league."*

(I made up the hand number and the idea that there was a transcript in order to make this more concrete.)

Under what conditions should we call E "evidence toward" A?

The prosecutor says the condition is simply that A and E don't contradict one another.

In symbols, his condition is: $P(A \cap E) > 0$.
(i.e. it is possible that A and E are both true.)

In my opinion, the condition should be:

$$P(A|E) > P(A).$$ (the evidence make A more likely to be true)

This is equivalent to:

odds

$$O(A|E) > O(A)$$

$$O(A) \cdot \frac{P(E|A)}{P(E|A^c)} > O(A)$$

$$\frac{P(E|A)}{P(E|A^c)} > 1$$

$$P(E|A) > P(E|A^c)$$

In other words, it is only evidence toward them cheating if they would be more likely to play that way when they're cheating than when they're not.

Example 3k A plane is missing, and it is presumed that it was equally likely to have gone down in any of 3 possible regions. Let $1 - \beta_i$, $i = 1, 2, 3$, denote the probability that the plane will be found upon a search of the ith region when the plane is, in fact, in that region. (The constants β_i are called *overlook probabilities*, because they represent the probability of overlooking the plane; they are generally attributable to the geographical and environmental conditions of the regions.) What is the conditional probability that the plane is in the ith region given that a search of region 1 is unsuccessful?

A_i = "The plane is in Region i."

B_i = "A search of Region i succeeds."

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$$

$$1 - \beta_i = P(B_i \mid A_i)$$

$$P(A_i \mid B_1^c) = ?$$

$$P(A_i \mid B_1^c) = \frac{P(A_i \cap B_1^c)}{P(B_1^c)}$$

$$= \frac{\overset{\frac{1}{3}}{\cancel{P(A_i)}}\, P(B_1^c \mid A_i)}{P(B_1^c)}$$

$$P(B_1^c \mid A_i) = 1 - P(B_1 \mid A_i)$$

$$P(B_1 \mid A_1) = 1 - \beta_1$$

$$P(B_1 \mid A_2) = P(B_1 \mid A_3) = 0$$

$$P(B_1^c \mid A_i) = \begin{cases} \cancel{1-(1-\beta_1)} \ \beta_1 & \text{if } i=1, \\ \cancel{1-0} \ 1 & \text{if } i=2 \text{ or } 3 \end{cases}$$

$$P(B_1^c) = 1 - P(B_1)$$

$\{A_1, A_2, A_3\}$ is a <u>partition</u> of the sample space (exactly one of them is true)

$$P(B_1) = \underbrace{\cancel{P(A_1)}}_{\frac{1}{3}} \underbrace{\cancel{P(B_1 \mid A_1)}}_{1-\beta_1} + \underbrace{\cancel{P(A_2)}}_{\frac{1}{3}} \underbrace{\cancel{P(B_1 \mid A_2)}}_{0} + \underbrace{\cancel{P(A_3)}}_{\frac{1}{3}} \underbrace{\cancel{P(B_1 \mid A_3)}}_{0}$$

law of total probability

$$= \frac{1}{3}(1-\beta_1)$$

$$P(B_1^c) = 1 - \frac{1}{3}(1-\beta_1) = \frac{2}{3} + \frac{1}{3}\beta_1$$

Putting it together...

$$P(A_i \mid B_1^c) = \frac{\frac{1}{3} P(B_1^c \mid A_i)}{\frac{2}{3} + \frac{1}{3}\beta_1} \cdot \frac{3}{3}$$

$$= \frac{P(B_1^c \mid A_i)}{2 + \beta_1}$$

$$= \begin{cases} \dfrac{\beta_1}{2+\beta_1} & \text{if } i=1, \\ \dfrac{1}{2+\beta_1} & \text{if } i = 2 \text{ or } 3. \end{cases}$$

Makes sense? (intuition check)

- β_1 small (very hard to overlook plane in Region 1)

  $P(A_1 \mid B_1^c) \approx 0$ (if not found in Region 1, very unlikely to be there)

  $P(A_2 \mid B_1^c) \approx \frac{1}{2}$ (if not found in Region 1, roughly 50% to be in Region 2 or Region 3)

- B_1 close to one (very easy to overlook plane in Region 1)

$P(A_1|B_1^c) \gtrsim \frac{1}{3}$

(failed search tells us very little)

$P(A_2|B_1^c) \gtrsim \frac{1}{3}$

Example 3l Suppose that we have 3 cards that are identical in form, except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side black. The 3 cards are mixed up in a hat, and 1 card is randomly selected and put down on the ground. If the upper side of the chosen card is colored red, what is the probability that the other side is colored black?

A = "The upper side of the chosen card is red."

B = "The chosen card is the red-black card."

$P(B) = \frac{1}{3}$

$P(A|B) = \frac{1}{2}$ ← assuming both sides equally likely to be put upward.

$P(B|A) = ?$

$$P(B|A) = \frac{\overset{\frac{1}{3}}{\cancel{P(B)}}\ \overset{\frac{1}{2}}{\cancel{P(A|B)}}}{P(A)}$$

$P(A) = ?$

C = "The chosen card is the red-red card."

D = " " " black-black " "

$$P(C) = P(D) = \frac{1}{3}$$

$\{B, C, D\}$ is a partition
(exactly one of them is true)

$$P(A) = \underbrace{P(B)}_{\frac{1}{3}}\underbrace{P(A|B)}_{\frac{1}{2}} + \underbrace{P(C)}_{\frac{1}{3}}\underbrace{P(A|C)}_{1} + \underbrace{P(D)}_{\frac{1}{3}}\underbrace{P(A|D)}_{0}$$

$$= \frac{1}{6} + \frac{1}{3} = \frac{1}{2}$$

$$P(B|A) = \frac{\frac{1}{3}\left(\frac{1}{2}\right)}{\left(\frac{1}{2}\right)} = \boxed{\frac{1}{3}}$$

Example 3m A new couple, known to have two children, has just moved into town. Suppose that the mother is encountered walking with one of her children. If this child is a girl, what is the probability that both children are girls?

A = "The mother is seen walking with a daughter."

B = "The older child is a girl."

C = "The younger child is a girl."

$$\underbrace{P(B\cap C) = P(B\cap C^c) = P(B^c\cap C) = P(B^c \cap C^c) = \frac{1}{4}}_{\text{We're meant to take this for granted in this example}}$$

$P(B\cap C \mid A) = ?$

Let's name these events to simplify notation:

$D_1 = B\cap C$

$D_2 = B\cap C^c$

$D_3 = B^c\cap C$

$D_4 = B^c \cap C^c$

$$P(D_1 \mid A) = \frac{\overset{\frac{1}{4}}{\cancel{P(D_1)}}\ \overset{1}{\cancel{P(A\mid D_1)}}}{P(A)}$$

$\{D_1, D_2, D_3, D_4\}$ is a partition

$$P(A) = \overset{\frac{1}{4}}{\cancel{P(D_1)}}\ \overset{1}{\cancel{P(A\mid D_1)}} + \overset{\frac{1}{4}}{\cancel{P(D_2)}}\ P(A\mid D_2)$$

$$+ \underset{\frac{1}{4}}{\cancel{P(D_3)}}\ P(A\mid D_3) + \underset{\frac{1}{4}}{\cancel{P(D_4)}}\ \underset{0}{\cancel{P(A\mid D_4)}}$$

$$P(D_1 | A) = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{4}P(A|D_2) + \frac{1}{4}P(A|D_3)} \cdot \frac{4}{4}$$

$$= \boxed{\frac{1}{1 + P(A|D_2) + P(A|D_3)}}$$

This is as far as we can take it. To progress further, we need more assumptions.

If we assume $P(A|D_2) = P(A|D_3) = \frac{1}{2}$, then the answer is $\frac{1}{2}$.

If we know the mother loves girls and always prefers to walk with one, then maybe $P(A|D_2) = P(A|D_3) = 1$, and the answer is $\frac{1}{3}$.

Or maybe we know the older child is an adult and will only walk when he or she is visiting the mother. Maybe we also know that sons visit their mother with a higher frequency than daughters. Based on all that, maybe we're assuming $P(A|D_2) = \frac{1}{5}$ and $P(A|D_3) = \frac{1}{3}$. In that case, the answer would be $\frac{15}{23}$.

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

A_j = "The car is behind door #j."

B = "The host showed me a goat behind door #3."

$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$

These are probabilities from my POV, before the host shows me the goat.

Is $P(A_1|B) > P(A_2|B)$,

$P(A_1|B) = P(A_2|B)$, or

$P(A_1|B) < P(A_2|B)$?

$\{A_1, A_2, A_3\}$ is a partition

$$P(A_1|B) = \frac{\overset{\frac{1}{3}}{\cancel{P(A_1)}}\,P(B|A_1)}{\underset{\frac{1}{3}}{\cancel{P(A_1)}}\,P(B|A_1) + \underset{\frac{1}{3}}{\cancel{P(A_2)}}\,P(B|A_2) + \underset{\frac{1}{3}}{\cancel{P(A_3)}}\,\underset{0}{\cancel{P(B|A_3)}}} \cdot \frac{3}{3}$$

$$= \frac{P(B|A_1)}{P(B|A_1) + P(B|A_2)}$$

We're meant to assume the host always shows you a goat, never shows you what's behind your chosen door, and if you picked the car, he chooses one of the other doors with equal probability. So...

$$P(B|A_1) = \frac{1}{2}$$

$$P(B|A_2) = 1$$

$$P(B|A_3) = 0$$

$$\therefore \; P(A_1|B) = \frac{\frac{1}{2}}{\frac{1}{2}+1} = \frac{1}{3}$$

$$P(A_2|B) = \frac{\overset{\frac{1}{3}}{\cancel{P(A_2)}}P(B|A_2)}{\underset{\frac{1}{3}}{\cancel{P(A_1)}}P(B|A_1) + \underset{\frac{1}{3}}{\cancel{P(A_2)}}P(B|A_2) + \underset{\frac{1}{3}}{\cancel{P(A_3)}}\underset{0}{\cancel{P(B|A_3)}}} \cdot \frac{3}{3}$$

$$= \frac{P(B|A_2)}{P(B|A_1) + P(B|A_2)}$$

$$= \frac{1}{\frac{1}{2}+1} = \frac{2}{3}$$

So $P(A_1|B) < P(A_2|B)$.

$$\boxed{\text{You should switch}}$$

This is called the Monty Hall problem.
Whatever you do, don't listen to the explanation in the movie, 21. Those screenwriters don't know what they're talking about. (Big surprise.)

Notice I conditioned on

"The host showed me a goat behind door #3."

not

"There is a goat behind door #3."

Those are not the same statements. Yes, I know the second one is true, but you must always condition on everything you know.

Similarly, in Expl 3m, I conditioned on

"The mother is seen walking with a daughter."

not

"At least one of the children is a girl."

This kind of precision can make or break your solution. Be verbose. Write out all of your events, as statements, with a complete sentence.

Bayes' Thm: If $B_1, \dots, B_n$ form a partition, then

$$P(B_k \mid A) = \frac{P(B_k)P(A \mid B_k)}{\sum_{i=1}^{n} P(B_i)P(A \mid B_i)}$$

Don't memorize. Just understand the method in the previous expls

3.4 Independent Events

If $P(A) > 0$ and $P(B) > 0$, then $\underbrace{\text{TFAE}}_{\text{the following are equivalent}}$:

- $P(A \mid B) = P(A)$
- $P(B \mid A) = P(B)$
- $P(A \cap B) = P(A)P(B)$

In general, A and B are <u>independent</u> if $P(A \cap B) = P(A)P(B)$.

They are <u>dependent</u> otherwise.

<u>Expl</u>

Roll a die. A = "The result is prime."

B = "The result is less than 5."

$\Omega = \{1, \dots, 6\}$, P = equ. likely

$A = \{2, 3, 5\}$, $B = \{1, 2, 3, 4\}$, $A \cap B = \{2, 3\}$

$P(A) = \frac{3}{6} = \frac{1}{2}$, $P(B) = \frac{4}{6} = \frac{2}{3}$

$P(A)P(B) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

$P(A \cap B) = \frac{2}{6} = \frac{1}{3}$

So A and B are indep.

In general, if $P(A) = 0$ or $P(A) = 1$, then A is independent of everything.

$A, B,$ and C are independent if

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

For A, B, C, D to be indep., need

$$P(A \cap B \cap C \cap D) = P(A)P(B)P(C)P(D)$$

+ all combos of 3

+ all combos of 2

And so on...

$A_1, A_2, A_3, \ldots$ are indep. if $A_1, \ldots, A_n$ are indep. for every n

---

If $A_1, A_2, \ldots, A_7$ are indep., then

$\underbrace{A_1 \cap (A_4 \cap A_6)^c}$ and $\underbrace{A_2 \cup A_7}$ are indep.

built from different A_j's

Same principle applies to any indep. collection

Example 4f An infinite sequence of independent trials is to be performed. Each trial results in a success with probability p and a failure with probability $1 - p$. What is the probability that

(a) at least 1 success occurs in the first n trials;
(b) exactly k successes occur in the first n trials;
(c) all trials result in successes?

A_n = "The n^{th} trial is a success."

$A_1, A_2, \ldots$ indep.

$P(A_n) = p \quad \forall n$

(a) $\displaystyle\bigcup_{k=1}^{n} A_k$ = "A_1 or A_2 or $\cdots$ or A_n"

= "At least 1 success occurs in the 1st n trials"

$P\left(\displaystyle\bigcup_{k=1}^{n} A_k\right) = ?$

$$P\left(\bigcup_{k=1}^{n} A_k\right) = 1 - P\left(\left(\bigcup_{k=1}^{n} A_k\right)^c\right)$$

$$= 1 - P\left(\bigcap_{k=1}^{n} A_k^c\right)$$

indep

$$= 1 - \prod_{k=1}^{n} P(A_k^c)$$

$$= 1 - \prod_{k=1}^{n} (1 - \overset{p}{\cancel{P(A_k)}})$$

$$= \boxed{1 - (1-p)^n}$$

(b) A = "There are exactly k successes in the 1st n trials."

<u>Do with $k=3$, $n=4$ first as an example</u>

success failure

mut. excl., there are $\binom{4}{3}=4$ events here.

B_1 = "The 1st 4 trial results, in order, are SSSF."

B_2 = " " SSFS."

B_3 = " " SFSS."

B_4 = " " FSSS."

$$P(A) = P\left(\bigcup_{k=1}^{4} B_k\right) = \sum_{k=1}^{4} P(B_k)$$

$$P(B_1) = P(A_1 \cap A_2 \cap A_3 \cap A_4^c)$$

$$= \overset{p}{\cancel{P(A_1)}}\,\overset{p}{\cancel{P(A_2)}}\,\overset{p}{\cancel{P(A_3)}}\,\overset{1-p}{\cancel{P(A_4^c)}}$$

$$= p^3(1-p)$$

$$P(B_2) = \cdots = p \cdot p \cdot (1-p) \cdot p = p^3(1-p)$$

$$P(B_3) = \cdots = p \cdot (1-p) \cdot p \cdot p = p^3(1-p)$$

$$P(B_4) = \cdots = (1-p) \cdot p \cdot p \cdot p = p^3(1-p)$$

$$P(A) = \underbrace{\binom{4}{3}}_{\text{\# of } B_k\text{'s}} \underbrace{p^3(1-p)}_{\text{each one has this same probability}}$$

For general k and n, there are $\binom{n}{k}$ different events making up A; each one has probability $p^k(1-p)^{n-k}$.

$$\boxed{P(A) = \binom{n}{k} p^k (1-p)^{n-k}}$$

(c) E = "All trials are a success."

$$E = \bigcap_{n=1}^{\infty} A_n$$

indep.

$$\cancel{P\left(\bigcap_{n=1}^{\infty} A_n\right) = \prod_{n=1}^{\infty} P(A_n)}$$

can't do this for infinitely many events

$$B_m = \bigcap_{n=1}^{m} A_n$$

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$$

By cont. from above,

$$\lim_{m\to\infty} P(B_m) = P\Big(\underbrace{\bigcap_{m=1}^{\infty} B_m}_{=E}\Big)$$

$$P(B_m) = P\left(\bigcap_{n=1}^{m} A_n\right)$$

$$= \prod_{n=1}^{m} \underbrace{\cancel{P(A_n)}}_{p}$$

$$= p^m$$

$$P(E) = \lim_{m\to\infty} P(B_m)$$

$$= \lim_{m\to\infty} p^m = \boxed{\begin{cases} 1 & \text{if } p=1, \\ 0 & \text{if } 0 \le p < 1. \end{cases}}$$

Example 4g A system composed of n separate components is said to be a parallel system if it functions when at least one of the components functions. (See Figure 2.) For such a system, if component i, which is independent of the other components, functions with probability $p_i, i = 1, \ldots, n$, what is the probability that the system functions?

A_i = "Component i functions."

$A_1, A_2, \ldots, A_n$ indep.

$P(A_i) = p_i$

$\bigcup_{i=1}^{n} A_i$ = "At least one component functions."

= "The system functions."

$P\left(\bigcup_{i=1}^{n} A_i\right) = ?$

$$P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - P\left(\left(\bigcup_{i=1}^{n} A_i\right)^c\right)$$

$$= 1 - P\left(\bigcap_{i=1}^{n} A_i^c\right) \quad \text{indep}$$

$$= 1 - \prod_{i=1}^{n} P(A_i^c)$$

$$= \boxed{1 - \prod_{i=1}^{n} (1 - p_i)}$$

(can't simplify any further)

Example 4h Independent trials consisting of rolling a pair of fair dice are performed. What is the probability that an outcome of 5 appears before an outcome of 7 when the outcome of a roll is the sum of the dice?

A_n = "The n^{th} roll is a 5."

B_n = "The n^{th} roll is a 7."

C_n = "The n^{th} roll is neither 5 nor 7."

$C_n = A_n^c \cap B_n^c$

A = "A 5 occurs before the 1^{st} 7."

$P(A) = ?$

SOLN 1

E_n = "The 1^{st} $n-1$ rolls are neither 5 nor 7, and the n^{th} roll is a 5."

$E_1, E_2, E_3, \ldots$ mut. excl.

$$A = \bigcup_{n=1}^{\infty} E_n$$

$$P(A) = P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n)$$

countable additivity

$$P(E_n) = P(\underbrace{C_1 \cap C_2 \cap \cdots \cap C_{n-1} \cap A_n}_{\text{indep.}})$$

$$P(E_n) = P(C_1)P(C_2)\cdots P(C_{n-1})P(A_n)$$

For any k,

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$P(A_k) = \frac{4}{36} = \frac{1}{9}$$

roll a 5

$$P(B_k) = \frac{6}{36} = \frac{1}{6}$$

roll a 7

$$P(C_k) = P(A_k^c \cap B_k^c)$$

not indep.

$$= P((A_k \cup B_k)^c)$$

$$= 1 - P(A_k \cup B_k)$$

mut. excl.

$$= 1 - (\underbrace{P(A_k)}_{\frac{1}{9}} + \underbrace{P(B_k)}_{\frac{1}{6}})$$

$$= 1 - \frac{5}{18} = \frac{13}{18}$$

(or count the unhighlighted squares in the above grid)

$$P(E_n) = \left(\frac{13}{18}\right)^{n-1}\left(\frac{1}{9}\right)$$

$$P(A) = \sum_{n=1}^{\infty} P(E_n)$$

$$= \sum_{n=1}^{\infty} \left(\frac{13}{18}\right)^{n-1} \left(\frac{1}{9}\right)$$

$$= \frac{\left(\frac{1}{9}\right)}{1 - \frac{13}{18}} = \frac{1}{9} \cdot \frac{18}{5} = \boxed{\frac{2}{5}}$$

SOLN 2 $\quad A_1, B_1, C_1 \leftarrow$ partition

$$P(A) = P(A_1)P(A \mid A_1) + P(B_1)P(A \mid B_1) + P(C_1)P(A \mid C_1)$$

$A_1 \Rightarrow A$ so $P(A \mid A_1) = 1$

$B_1 \Rightarrow A^c$ so $P(A \mid B_1) = 0$

$C_1 \Rightarrow ?$

If the first roll is neither 5 nor 7, then the game effectively starts over.

$$\therefore \underbrace{P(A \mid C_1)}_{\text{prob. after rolling something other than 5 or 7}} = \underbrace{P(A)}_{\text{prob. at start of game}}$$

$$P(A) = \overset{\frac{1}{9}}{\cancel{P(A_1)}}\,\overset{1}{\cancel{P(A|A_1)}} + \overset{\frac{1}{6}}{\cancel{P(B_1)}}\,\overset{0}{\cancel{P(A|B_1)}}$$

$$+ \underset{\frac{13}{18}}{\cancel{P(C_1)}}\,\underset{P(A)}{\cancel{P(A|C_1)}}$$

$$P(A) = \frac{1}{9} + \frac{13}{18} P(A)$$

$$\left(1 - \frac{13}{18}\right) P(A) = \frac{1}{9}$$

$$P(A) = \frac{\left(\frac{1}{9}\right)}{1 - \frac{13}{18}} = \frac{1}{9} \cdot \frac{18}{5} = \boxed{\frac{2}{5}}$$

Example 4i Suppose there are n types of coupons and that each new coupon collected is, independent of previous selections, a type i coupon with probability p_i, $\sum_{i=1}^{n} p_i = 1$. Suppose k coupons are to be collected. If A_i is the event that there is at least one type i coupon among those collected, then, for $i \neq j$, find

(a) $P(A_i)$
(b) $P(A_i \cup A_j)$
(c) $P(A_i | A_j)$

Modern people might not understand this example.
Change "coupon" to "Pokemon card".
Each pack you buy has one card.
It's a type i card with probability p_i
(There are n possible types.)
You buy k packs.
A_i = "You get at least one type i Pokemon card."

(a) B_m = "The m^{th} pack has a type i card."

$B_1, \dots, B_k$ indep.

$P(B_m) = p_i$

$A_i = \bigcup_{m=1}^{k} B_m$

$$P(A_i) = P\left(\bigcup_{m=1}^{k} B_m\right) = 1 - P\left(\left(\bigcup_{m=1}^{k} B_m\right)^c\right)$$

$$= 1 - P\left(\bigcap_{m=1}^{k} B_m^c\right)$$

$$= 1 - \prod_{m=1}^{k} P(B_m^c) = \boxed{1-(1-p_i)^k}$$

(b) C_m = "The m^{th} pack has either type i or type j."

$C_1, \dots, C_k$ indep.

$P(C_m) = p_i + p_j$

$A_i \cup A_j = \bigcup_{m=1}^{k} C_m$

$$P(A_i \cup A_j) = P\left(\bigcup_{m=1}^{k} C_m\right)$$

$$= 1 - \prod_{m=1}^{k} P(C_m^c)$$

$$= \boxed{1-(1-p_i-p_j)^k}$$

(c) $$P(A_i \mid A_j) = \frac{P(A_i \cap A_j)}{P(A_j)}$$

$$P(A_i \cap A_j) = ?$$

$$P(A_i \cup A_j) = P(A_i) + P(A_j) - P(A_i \cap A_j)$$

$$1-(1-p_i-p_j)^k = 1-(1-p_i)^k + 1-(1-p_j)^k - P(A_i \cap A_j)$$

$$P(A_i \cap A_j) = 1-(1-p_i)^k - (1-p_j)^k + (1-p_i-p_j)^k$$

$$P(A_i \mid A_j) = \boxed{\frac{1-(1-p_i)^k - (1-p_j)^k + (1-p_i-p_j)^k}{1-(1-p_i)^k}}$$

HW: Ch. 3: 5, 10, 16, 20, 26, 30, 35, 43,
17, 21, 28, 32, 37, 45

44, 50, 53, 58(a), 70, 74, 78
46, 52, 58, 63(a), 74, 77, 81

mistake

Should have been:

Ch. 3: 5, 10, 17, 21, 28, 32, 37, 45
16, 20, 26, 30, 35, 43

46, 52, 58, 63(a), 74, 77, 81
44, 50, 53, 58(a), 70, 74, 78