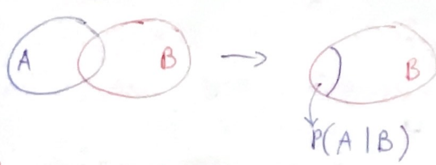
Unit 2: Conditional Probability and Bayes' Rule

Conditioning Probability

If A and B are events with $P(B) > 0$, then the conditional probability of A given B, $P(A \mid B)$ is:

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

Intuition:



Theorem 2.3.1

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

Applying theorem 2.3.1 repeatedly, we can generalize to the intersection of n events

Theorem 2.3.2

$$P(A_1, A_2, \dots A_n) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3)\,P(A_2 \mid A_1, A_2) \dots P(A_n \mid A_1, \dots A_{n-1})$$

We are now ready to introduce the 2 main theorems about conditional probability: Baye's rule and Law of Total ~~Prop~~ Probability

Theorem 2.3.3 (Bayes' rule)

$$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)}$$

Theorem 2.3.4 (Law of Total Probability ~~LOTB~~ LOTP)

Let $A_1, A_2 \dots A_n$ be a partition of the sample space S (i.e, the A_i are disjoint events and their union is S), with $P(A_i) > 0$ for all i. Then:

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i)\,P(A_i)$$

Example ~~2.3.7~~ 2.3.6

You have 1 fair coin and 1 biased coin with Heads 3/4. You pick 1 coin and flip it 3 times. What is the probability that the coin you picked is the fair one?

- A is events of the coin lands heads 3 times
- F is event that the chosen coin is fair one
- We need to find ~~P(A|F~~ $P(F \mid A)$, but its easier to find $P(A \mid F)$ and $P(A \mid F^c)$, therefore:

$$P(F \mid A) = \frac{P(A \mid F)\,P(F)}{P(A)} = \frac{P(A \mid F)\,P(F)}{P(A \mid F)\,P(F) + P(A \mid F^c)\,P(F^c)}$$

$$\approx 0.23$$

Example 2.3.7 (Testing for a rare disease)

Jimmy being tested for a disease, which affects 1% of population, the test resalt is positive. Let D be the event that Jimmy has the disease and T be the event that he tests positive.

The test is 95% accurate, which means $P(T|D)$ and $P(T^c|D^c)$ is 0.95

Find the probability that Jimmy has the disease, given the evidence provided by the test result.

$$P(D|T) = \frac{P(T|D)\ P(D)}{P(T)} = \frac{P(T|D)\ P(D)}{P(T|D)P(D) + P(T|D^c)\ P(D^c)}$$

$$\approx 0.16$$

Note: $P(D|T)$ is a balance between $P(D)$ and $P(T)$

With extra conditioning

Theorem 2.4.1 (Bayes' rule with extra conditioning)

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

Theorem 2.4.2 (LOTP with extra conditioning)

$$P(B|E) = \sum_{i=1}^{n} P(B|A_i, E)\ P(A_i|E)$$

Independent events

Events A and B are independent if

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

If $P(A) > 0$ and $P(B) > 0$, then:

$$P(A|B) = P(A) \quad \text{and} \quad P(B|A) = P(B)$$

Independence vs disjoint

- Independence events give no information about each other, they can happen at the same time. $P(A \cap B) = P(A).P(B)$
- Disjont event give information about each other, they ~~can~~ cannot happen ~~at the~~ at the same time. Since $P(A \cap B) = 0$

Pairwise independence and (complete) independence

$$\text{Complete} \begin{cases} \left.\begin{array}{l} P(A \cap B) = P(A)\ P(B) \\ P(A \cap C) = P(A)\ P(C) \\ P(B \cap C) = P(B)\ P(C) \end{array}\right\} \text{pairwise} \\ P(A \cap B \cap C) = P(A)\ P(B)\ P(C) \end{cases}$$

Example 2.5.4 (pairwise independence doesn't imply independence)

Consider 2 coin tosses, A is event that the first is Head, B is event that the second is Head, C is event that both has same results

- Knowing what happen with A or B separately doesn't give any information about C
- Knowing what happen ~~with~~ happens both A and B give us information about C. ~~To~~ A and B imply C by definition

Example 2.5.6 (conditional independence doesn't imply independence)

Return to example 2.5.6, suppose we have chosen a coin but not sure if it is a fair or a biased one. And we flip the coin a number of times. Conditional on choosing the fair coin, the coin tosses are independent. Conditional on choosing the biased coin, the tosses are also independent.

However, the coin tosses are not are not unconditionally independent. Because if we don't know which coin we have chosen, then observing the sequence of tosses give us information about the coin, and in turn predict future tosses.

Example 2.5.7 (independence doesn't imply conditionall independence)

Suppose A and B are the only two ppl who ever call me. A and B are clearly independence.

However, conditional on the phone ringing (R), if A is calling meaning B is not calling and vice versa. $P(A|R) < 1 = \cancel{P(A)}$
$= P(A|B^c, R)$

Therefore, A and B^c are not conditionally independent.

Conditioning as a problem-solving tool

Conditioning allow us to split a problem into sub-problems. We can ~~say,~~ "if condition on E and E^c, then combine them using LOTP

Strategy 1: Condition on what you wish you know

Example 2.6.1

There are 3 doors, 1 of which has a car behind it, Monty always reveal the first door without a car. The contestant is given the option to switch door before open, should he?

5

Condition on wish that we know where the car is.

. Let C_i be the event that the car behind door $i = 1, 2, 3$:

$$P(\text{get car}) = P(\text{get car}|C_1)\frac{1}{3} + P(\text{get car}|C_2)\frac{1}{3} + P(\text{get car}|C_3)\frac{1}{3}$$

. Suppose we switch, then if car is behind door 1 ~~and~~ fail, we switch to door 2, given Monty has already opened door 3 without a car, then switching will succeed:

$$P(\text{get car}) = 0.\frac{1}{3} + 1\frac{1}{3} + 1\frac{1}{3} = \frac{2}{3}$$

So switching works $\frac{2}{3}$ of the time.

Strategy 2: Condition on the first step

Example 2.6.5

A single Bobo, after 1 minute can either die, stay the same, or duplicate. What is the probability that Bobo's population will eventually die out?

. Let D be the event that the population die out

. Let B_i be the event that Bobo turns into i Bobo

Then: $P(D) = P(D|B_0)\frac{1}{3} + P(D|B_1)\frac{1}{3} + P(D|B_2)\frac{1}{3}$

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

$\Rightarrow$ Solve the above equation, $P(D) = 1$. So the population will eventually die out.

Example 2.6.4

2 gambler A and B make sequence of \$1 bets. Each bet, A has p chance of winning, B has $q = 1-p$ chance of winning. A starts with i dollars and B starts with $N-i$ dollars. When A gains, B loses. What is the probbility that A wins the game?

. Let p_i be event that A wins the game, so $p_0 = 0$ means A loses and $p_N = 1$ means A wins

6

Condition on outcome of the first round, we have:

$p_i = P(W \mid A \text{ starts at } i, \text{win})\, p + P(W \mid A \text{ starts at } i, \text{lose})\, q$

P(W)

$= P(W \mid A \text{ starts at } i+1)\, p + P(W \mid A \text{ starts at } i-1)\, q$

$= p_{i+1} \cdot p + p_{i-1} \cdot q$

$$\Rightarrow p_i = \begin{cases} \dfrac{1-\left(\frac{q}{p}\right)^i}{1-\left(\frac{q}{p}\right)^M} & \text{if } p \neq \frac{1}{2} \\ \dfrac{i}{M} & \text{if } p = \frac{1}{2} \end{cases}$$

Example 2.6.5 (Simpson's paradox)

Dr. Hibert

	Heart	Band-aid
Success	70	10
Failure	20	0

Dr. Nick

	Heart	Band-aid
Success	2	81
Failure	8	9

Simpson's paradox happens when

$P(A \mid B, C) < P(A \mid B^c, C)$

$P(A \mid B, C^c) < P(A \mid B^c, C^c)$

but

$P(A \mid B) > P(A \mid B^c)$

In this example, A is succes surgery, B is Dr. Nick perform, C is heart surgery. Here even the Dr. Nick ~~he~~ perfom worse than Dr. Hilbert in both types of surgery, his overal stats are still better!

LOTP can tell us ~~whey~~ why:

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

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

Although: $\begin{cases} P(A \mid C, B) < P(A \mid C, B^c) \\ P(A \mid C^c, B) < P(A \mid C^c, B^c) \end{cases}$

The weights $P(C \mid B)$ and $P(C^c \mid B)$ can flip the balance. Since Dr. Nick perform much more Band-aid $P(C \mid B) \uparrow$, lead to increase overall performance $P(A \mid B) \uparrow$

Practice Problems

Problem 1

There is an email spam filter. Suppose that 80% of email is spam. In 10% of spam emails, the phrase "free money" is used, whereas this phrase is only used in 1% of non-spam emails. A new email with "free money" arrived, what probability is it a spam?

Let $\begin{cases} F \text{ be event that email has "Free money"} \\ S \text{ be event that email is Spam} \end{cases}$

$$P(S \mid F) = \frac{P(F \mid S)\, P(S)}{P(F)} = \frac{P(F \mid S)\, P(S)}{P(F \mid S) P(S) + P(F \mid S^c)\, P(S^c)}$$

$$= \frac{0.1 \times 0.8}{0.1 \times 0.8 + 0.01 \times 0.2} = 0.975$$

Problem 2

The screen for phone is manufactured by 3 companies A, B, C. With the proportions of 0.5, 0.3, 0.2 respectively. Their defective probabilities are 0.01, 0.02, 0.03 respectively. Given the screen on the phone is defective, what probability that company A did it?

Let $\begin{cases} F \text{ be event of Defective} \\ A, B, C \text{ be events that company manufactured} \end{cases}$

$$P(A \mid D) = \frac{P(D \mid A)\, P(A)}{P(D)} = \frac{P(D \mid A)\, P(A)}{P(D \mid A)\, P(A) + P(D \mid B)\, P(B) + P(D \mid C)\, P(C)}$$

$$= \frac{0.01 \times 0.5}{0.01 \times 0.5 + 0.02 \times 0.3 + 0.03 \times 0.2}$$

$= 0.294$

Problem 3

A family has 3 children, named A, B, and C

a) ~~Fo~~ Is the event "A is older than B" independent of "A is older than C"

No they are not independent events, because if we consider the sequence of birth; given $A > B$:

$$\left.\begin{array}{l} ABC \\ ACB \end{array}\right\} \text{A older than C}$$

$$CAB \rightarrow \text{A younger than C}$$

So knowing $A > B$ (older), give us information about whether A is older than C.

To make this more intuitive, think of extreme case with 100 children ~~A_0~~ $A_1, \dots A_{100}$. Given $A_1 > A_2 \dots A_{99}$, then the only way for A_{100} to be older than A_1 is $A_{100} A_1 \dots A_{99}$ sequence.

b) Find Probability that A is older than B, given A older than C

$$P(A>B \mid A>C) = \frac{P(A>B, A>C)}{P(A>C)} = \frac{1/3}{1/2} = \frac{2}{3}$$

Problem 4

Consider the Monty Hall problem, except that Monty enjoys open door 2 more than door 3, if he has a choices, he would open door 2 with ~~probl~~ $p = 3/4$

a) Find the probability that the strategy of always switching succeeds

$$P(W) = P(W \mid C_1)P(C_1) + P(W \mid C_2)P(C_2) + P(W \mid C_3)P(C_3)$$
$$= 0 \cdot \tfrac{1}{3} + 1 \cdot \tfrac{1}{3} + 1 \cdot \tfrac{1}{3}$$
$$= 2/3$$

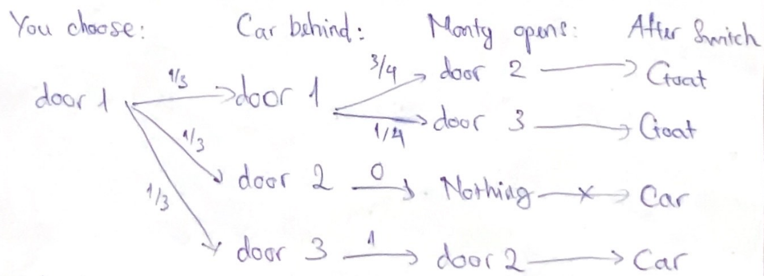
b) Find the probability that the strategy of always switching succeeds, given Monty opens door 2

Let D_i be event that Monty opens Door i

We are looking for ~~$P(D$~~ $P(W \mid D_2)$, which is the same as $P(C_3 \mid D_2)$

$$\cdot\ P(C_3 \mid D_2) = \frac{P(D_2 \mid C_3) \cdot P(C_3)}{P(D_2)}$$
$$= \frac{P(D_2 \mid C_3)\, P(C_3)}{P(D_2 \mid C_1)\, P(C_1) + P(D_2 \mid C_2)\, P(C_2) + P(D_2 \mid C_3)\, P(C_3)}$$
$$= \frac{1 \times 1/3}{4/3 \cdot 1/3 + 0 \cdot 1/3 + 1 \cdot 1/3} = \frac{4}{7}$$

Visualizing this:



c) Find the probability that the strategy of always switching succeeds, given Moty opens door 3

$$.\ P(C_2 \mid D_3) = \frac{P(D_3 \mid C_2) \cdot P(C_2)}{P(D_3)} = \frac{P(D_3 \mid C_2) \cdot P(C_2)}{P(D_3 \mid C_1)P(C_1) + P(D_3 \mid C_2)P(C_2) + P(D_3 \mid C_3)P(C_3)}$$
$$= 0.8$$

Homework Problems

Problem 1

Fred is answering a multiple choice problem with n options.

Let K be event that Fred know the answer

R be event that Fred gets the problem right

Suppose that if he knows the answer then he'll get the problem right. But if he doesn't know the answer then he'll guess randomly.

Let $P(K) = p$.

a) Find $P(K|R)$. $P(K|R) = \frac{P(R|K) \cdot P(K)}{P(R)}$

$$= \frac{P(R|K) \cdot P(K)}{P(R|K) \cdot P(K) + P(R|K^c) \cdot P(K^c)} = \frac{\cancel{1} \cdot p}{\cancel{1/n}} = \frac{1 \cdot p}{1 \cdot p + \frac{1}{n} \cdot (1-p)}$$

$$= \frac{p}{p + \frac{(1-p)}{n}}$$

b) When (if ever) does $P(K|R) = p$?

Given that $P(K|R) = \frac{p}{p + \frac{(1-p)}{n}}$

$\Rightarrow$ With these extreme cases: $n = 1$, $p = 0$, $p = 1$.

Problem 2

A hat contains 100 coins, where 99 are fair and 1 is always head. A coin is chosen, ~~fli~~ flipped 7 times and lands head all 7 times. What is the probability that the chosen coin is double-headed (always heads)?

Let F is event fair coin chosen

F^c is event double-headed coin chosen

H_i is event that lands head i times

$$P(F^c | H_7) = \frac{P(H_7 | F^c) \cdot P(F^c)}{P(H_7|F^c) P(F^c) + P(H_7|F) \cdot P(F)}$$

$$= \frac{1 \cdot 0.01}{1 \cdot 0.01 + 0.5^7 \cdot 0.99} = 0.564$$

Problem 3

Jimmy takes a series of n tests. Let D be event that he has the disease, $p = P(D)$ be the prior probability that he has the disease, and $q = 1 - p$. Let T_j be the event that he tests positive on the j^{th} test.

a) Assume the test results are conditionally independent given Jimmy's disease status. Let $a = P(T_j | D)$, $b = P(T_j | D^c)$. Find the posterior probability that Jimmy has the disease, given that he tests positive on exactly k out of n tests.

$$P(D | X = k) = \frac{P(X = k | D) \cdot P(D)}{P(X = k)} = \frac{a^k (1-a)^{n-k} \cdot p}{a^k (1-a)^{n-k} \cdot p + b^k (1-b)^{n-k} \cdot q}$$

b) There is a gene G that makes all the tests results positive. Assume $P(G) = \frac{1}{2}$, D and G are independent. Let $a_0 = P(T_j | D, G^c)$, $b_0 = P(T_j | D^c, G^c)$. Find probability that Jimmy has the disease, given that he tests positive all n times?

Let T be event that Jimmy tests positive all n times.

$$P(D|T) = \frac{P(T|D) \cdot P(D)}{P(T)} = \frac{P(T|D) \cdot p}{P(T|D) \cdot p + P(T|D^c) \cdot q}$$

Add extra conditioning on if Jimmy has the gene G:

$$\cancel{P(D,T}\ P(T|D) = P(T|D,G).P(G|D) + P(T|D,G^c).P(G^c|D) = \frac{1}{2} + \frac{a_0^n}{2}$$

$$P(T|D^c) = P(T|D^c,G).P(G|D^c) + P(T|D^c,G^c).P(G^c|D^c) = \frac{1}{2} + \frac{b_0^n}{2}$$

Thus: $$P(D|T) = \frac{p(1+a_0^n)}{p(1+a_0^n) + q(1+b_0^n)}$$

Problem 4

An election with 2 candidates A and B. Every voter is invited to do a poll. Let A be event that voter voted A, W be event that the voter willing to participate the poll. We know that $P(W|A) = 0.7$, $P(W|A^c) = 0.3$. The final poll said that 60% of respondents say they voted for A. Find $P(A)$, true ~~po~~ proportion of ppl who voted for A.

$$P(A|W) = \frac{P(W|A).P(A)}{P(W)} = \frac{P(W|A).P(A)}{P(W|A).P(A) + P(W|A^c)P(A^c)}$$

$$\Leftrightarrow 0.6 = \frac{0.7\ P(A)}{0.7\ P(A) + 0.3(1-P(A))}$$

$$\Rightarrow P(A) = 0.39$$