Bayesian

2.2 Bayes Rule

Recall our previous problem where we had 2 coins, 1 biased and one fair, we chose one randomly and flipped it.

The next two problems naturally lead into BAYES RULE.

Q: Suppose we went back to our fair coin problem. I want to ask a new question:

Given that I saw a tails on my coin toss, whats the probability that I chose the fair coin?

That is I want

Chose Coin1 ? ? I saw a tails on my coin toss.

$$P(C1 \mid T)$$

given

C1 = Fair

C2 = Biased

The trick is to simply follow your nose.

"hard"

$$P(C1 \mid T) = \frac{P(C1, T)}{P(T)} = \frac{P(C1 \cap T)}{P(T)}$$

$$= \frac{P(T \cap C1)}{P(T)} = \frac{\overset{0.5}{P(T \mid C1)}\,\overset{1/2}{P(C1)}}{\underset{0.3}{P(T)}} = \frac{0.5\,(0.5)}{0.3}$$

$$= \frac{5}{6} = 0.83\ldots$$

$$= \frac{P(T|C1)P(C1)}{P(T)}$$

notice the "flip" here.

we did this a few minutes ago

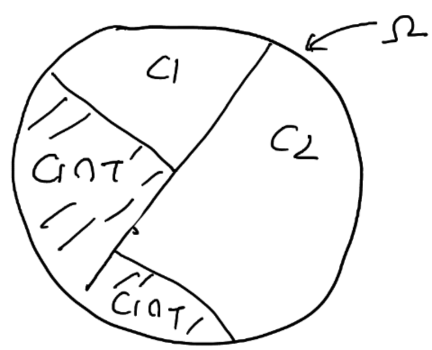
$$= \frac{P(T|C1)P(C1)}{P(T|C1)P(C1) + P(T|C2)P(C2)}$$

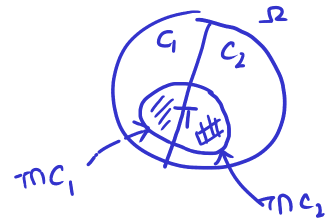
law of total prob.

$$= \frac{\frac{1}{2}\cdot\frac{1}{2}}{\frac{1}{2}\cdot\frac{1}{2} + 0.1\,\frac{1}{2}} = \frac{1}{1+0.2} = \frac{1}{1.2}$$

which is quite high. This makes sense because the probability of tails from the 2nd coin is quite low.

Note that we used an important law to determine $P(T)$:

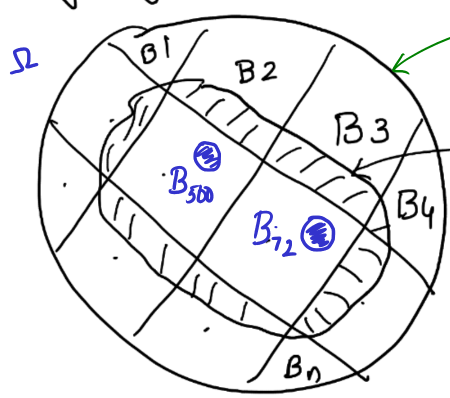




$$P(T) = P(T\cap C1) + P(T\cap C2)$$

$$= P(T|C1)P(C1) + P(T|C2)P(C2).$$

This law of total probability has the following generalization:



Partition: $\{B_1, \ldots, B_n\}$ $B_i \cap B_j = \phi$

$B_1 \sqcup B_2 \sqcup \cdots \sqcup B_n = \Omega$

add up the shaded regions

originally had only 2 events.

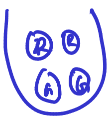
$$P(E) = \sum_{i=1}^{n} P(B_i \cap E) = P(B_1\cap E) + P(B_2 \cap E) + \cdots + P(B_n \cap E)$$

$$= \sum_{i=1}^{n} P(E|B_i)P(B_i)$$

This ties in nicely into strategy 1:

Break up complicated events into simpler ones.

2.12 (MORE ON BAYES THEOREM)

 U_1

We have 2 urns. First we choose an urn (uniformly)

U1 : 2G 2R (G = Green, R = Red)

 U_2

U2 : 2R 3Y

Then, you choose a ball from the urn. What's the prob. that it is red?

law of total prob.

$$P(R) = P(R \mid U1)\,P(U1) + P(R \mid U2)\,P(U2)$$

$$= \frac{2}{4} \cdot \frac{1}{2} + \frac{2}{5} \cdot \frac{1}{2}$$

Then suppose u picked a red ball, what's the prob it came from urn 1?

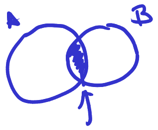
Bayes rule: As before, find $P(U1 \mid R)$.

Use the definition:

$$P(\underbrace{U1}_{\text{it came from urn 1}} \mid \underbrace{R}_{\text{red ball}}) = \frac{P(R \mid U1)\,P(U1)}{P(R)}$$

And that's it!

$$= \frac{\frac{2}{4} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{2}{4} + \frac{1}{2} \cdot \frac{2}{5}}$$



$U_1 \cap R = R \cap U_1$

$$P(U_1 | R) = \frac{P(U_1 \cap R)}{\underset{\text{found this}}{P(R)}} = \frac{P(R \cap U_1)}{P(R)} = \frac{\overset{2/4}{P(R | U_1)}\,\overset{1/2}{P(U_1)}}{P(R)}$$

$$= \frac{2/4 \cdot \frac{1}{2}}{\frac{2}{4} \cdot \frac{1}{2} + \frac{2}{5} \cdot \frac{1}{2}}$$

POLL

Given a red ball was chosen what's the prob

it came from urn 1?

U_1 : 2R 2G

U_2 : 2R 3Y

Open ended question on https://pollev.com/arjunkrishna250

Exercise to do in class. (or give in recitation)

$U_1 : 1R \quad 9G$
$U_2 : 2R \quad 8G$

5 urns 1 2 3 4 5

k^{th} urn has k Red $10-k$ Green. balls

Pick an urn, and pick a ball from it.
↳ uniformly at random

Find the probability that urn 1 was chosen given that a red ball was chosen.

U_i = Urn i is chosen

R = red ball is picked

We are looking for the following probability

$$P(\overset{\text{urn 1}}{U_1} \mid \overset{\text{red ball}}{R}) = \frac{P(U_1 \cap R)}{P(R)} = \frac{P(R \mid U_1)\, P(U_1)}{P(R)}$$

Poll

Do you want to try it yourself?

Solution:

First compute $P(R)$

$$P(R) = \sum_{k=1}^{5} P(R/U_k)\,P(U_k) = \sum_{k=1}^{5} \frac{k}{10}\cdot\frac{1}{5}$$

Then compute

$$\mathbb{P}(U_i \mid \text{Red})$$

$$\mathbb{P}(i^{th} \text{ urn is picked} \mid \text{red ball})$$

$$= \frac{\mathbb{P}(R \mid U_i)\,\mathbb{P}(U_i)}{\mathbb{P}(R)}.$$

$$\mathbb{P}(R) = \sum_{k=1}^{5} \frac{1}{5}\,\frac{k}{10}$$

$$\mathbb{P}(U_i \mid R) = \frac{\frac{i}{10}\cdot\cancel{\frac{1}{5}}}{\cancel{\frac{1}{5}}\sum_{k=1}^{5}\frac{k}{10}} = \frac{i}{\sum_{i=1}^{5} k}.$$

$$P(U_1 \mid R) = \frac{1}{15}.$$

pre-covid questions.

2.14 Suppose a test for a disease is 96% accurate. False positives 2% of the time. Percentage of population having it is 0.5%.
↳ the disease

A random person tests positive. What is the probab. that they have the disease?

POLL

How high do you think it's going to be?

> 80 %	< 20 %

Let us compute.

$A = \{\text{Person tests positive}\}$
$D = \{\text{Person has disease}\}$.

Experience will help you choose the right events.

$P(A|D) = \overbrace{0.96}^{\text{"96 \% accurate"}}$ $\quad P(A|D^c) = \overset{\text{"false positive rate"}}{0.02}$

rate of occurrence of disease → $P(D) = 0.005$

First step: write down what's given.

$$P(D \underset{\substack{\uparrow \\ \text{given}}}{|} A) = \; ? \qquad \frac{P(D \cap A)}{P(A)} = \frac{P(A \cap D)}{P(D)}$$

$P(A \cap D)$

0.96 0.005

$$\boxed{P(D|A) = \frac{P(A|D)\,P(D)}{P(A)}}$$

Law of total probability ?

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

$P(A) = P(A \cap D) + P(A \cap D^c)$

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

$(1 - 0.005)$

$$= \frac{0.96 \quad 0.005}{0.005 \quad 0.96 + 0.02 \quad 0.995}$$

$\approx 0.194 \approx 20\%$

There is only a 20% chance that a random person has the disease.

The key here is <u>RANDOM PERSON</u>

If your doctor makes you take the antigen test, then she is pretty certain that you have the disease or at least thinks its 50-50.

So $P(D) = \frac{1}{2}$ (say) "Dr. estimates that the prob. you have the diseases is 50-50"

$$P(D|A) = \frac{P(A|D)\,\overset{\nearrow\ 0.005 \text{ to } 0.5}{P(D)}}{P(A|D)P(D) + P(A|D^c)P(D^c)}$$

$$= \frac{0.96 \cdot \frac{1}{2}}{0.96 \cdot \frac{1}{2} + 0.02 \cdot \frac{1}{2}} = \frac{0.96}{0.98} \approx 98\%$$

So the accuracy of the test is <u>enhanced</u>

when you're already sort of sure that the person might have the disease.

Question: What happens if 15% of the population has the disease?

Lesson: Your sampling strategies HAVE to be different if there are low levels of the disease in the population versus high levels of disease.