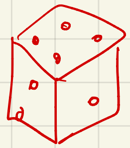
Example:

$A = \{2, 4, 6\}$

$S = \{1, 2, 3, 4, 5, 6\}$

$B = \{3, 6\}$



Uniform Prob. Space

$C = \{1, 2, 3, 4, 5\}$

$P(A) = 1/2 \qquad P(B) = 1/3 \qquad P(C) = 5/6$

$P(A, B) = 1/6 = P(A) \cdot P(B)$ (A & B are independent)

A : event that we get even number

B : event that we get a multiple of 3

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

Given that we rolled a multiple of 3, the prob. that it's even is still 1/2

$$P(A|C) = \frac{P(A, C)}{P(C)} = \frac{2/6}{5/6} = 2/5$$

$$P(B|C) = 1/6$$

$$P(A, B | C) = P(A, B, C) / P(C) = 0$$

$$P(A, B | C) \neq P(A|C)\, P(B|C)$$

Conditioned on C, A & B are not independent anymore

Independence can be tricky!

- It's possible that A, B, C and pairwise independent, but $P(A, B, C) \neq P(A)P(B)P(C)$

- It's also possible that $P(A, B, C) = P(A)P(B)P(C)$ but the events are not pairwise independent.

K-wise independence:

$A_1, A_2, \ldots, A_k$ are k-wise independent if for every subset $S \subset \{1, 2, \ldots, k\}$

$$P\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} P(A_i)$$

When we say events are independent that's what we usually mean.

Coin Example: Assume we toss a coin twice

$$S = \{HH, HT, TH, TT\}$$

Assume coin is fair, and tosses are independent.

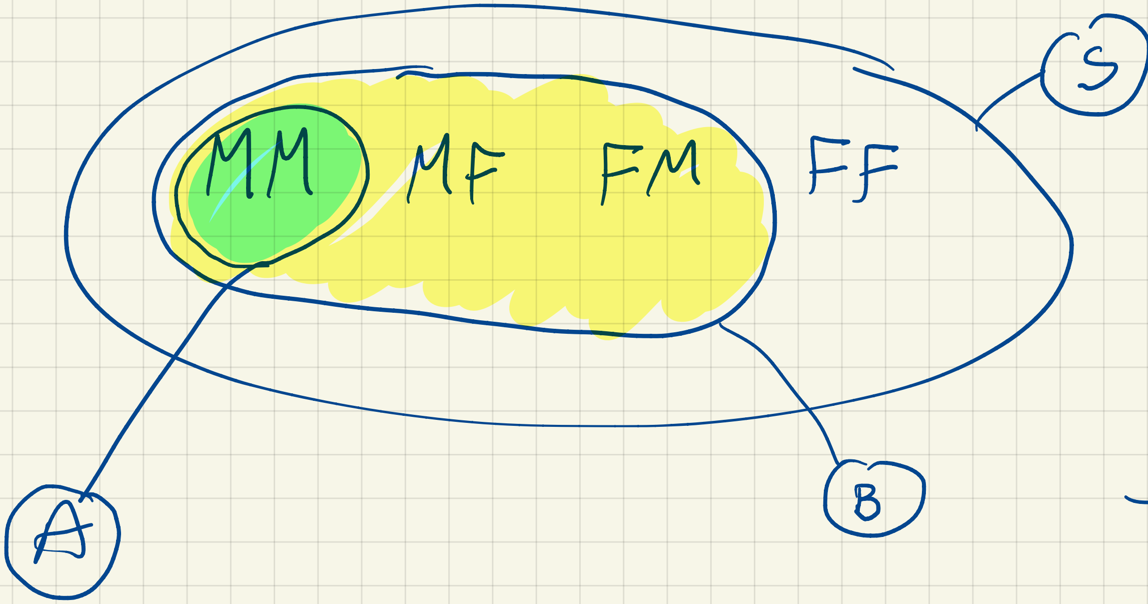
$$P(HH) = P(H)P(H) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Define $A = \{HH, HT, TH\}$ $\quad B = \{HH, TT\}$

$P(A) = 3/4 \quad P(B) = 1/2 \quad P(A, B) = 1/4 \quad P(A,B) \neq P(A)P(B)$

$P(B|A) = 1/3 \quad P(A|B) = 1/2$

King's Paradox: The King comes from a family of two children. What is the prob. that the king's sibling is a male?



King's Paradox

$$P(A|B) = \frac{1/4}{3/4} = 1/3$$

A: King's sibling is male

B: There is at least one male

There is "implicit" conditioning in the way the question is framed.

(There is at least one male!)

Another example:

FAIR

$P(H) = P(T) = \frac{1}{2}$

BIASED

$P(H) = 1$

- Pick one of the coins with prob. $\frac{1}{2}$

- Toss coin 3 times.

Question: If we observe more Heads then tails what is the probability that coin is Fair.

A: Coin is Fair

B: we see more Heads

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

$$B = \{ \underline{FHHH}, \underline{FHHT}, \underline{FHTH}, \underline{FTHH}, BHHH \}$$

$$A = \{ \underline{FHHH}, \underline{FHHT}, \underline{FHTH}, \underline{FTHH}, FHTT, FTHT, FTTH, FTTT \}$$

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}}{\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{2}} = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{2}} = \frac{1}{3}$$

There is some subtle issues to be clarified later regarding how I obtained $\frac{1}{16}$. (below)

$$P(FHHH) = P(F, H_1, H_2, H_3)$$

$$= \underbrace{P(F)}_{\frac{1}{2}} \underbrace{P(H_1|F)}_{\frac{1}{2}} \underbrace{P(H_2|H_1, F)} P(H_3|H_1, H_2, F)$$

(use chaining rule) next slide

Conditioned on the coin, all tosses are independent

$\Downarrow$

$$P(H_2|H_1, F) = P(H_2|F) = \frac{1}{2}$$

Multiplication Rule for Conditioning:

$$P(A|B) = P(A,B)/P(B) \implies P(A,B) = P(B)\,P(A|B)$$

$$P(A,B) = P(A) \cdot P(B|A)$$

What if I have n events:

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

example for $n=4$:

$$P(A_1,A_2,A_3,A_4) = P(A_1) \cdot P(A_2|A_1)\,P(A_3|A_1,A_2)\,P(A_4|A_1,A_2,A_3)$$

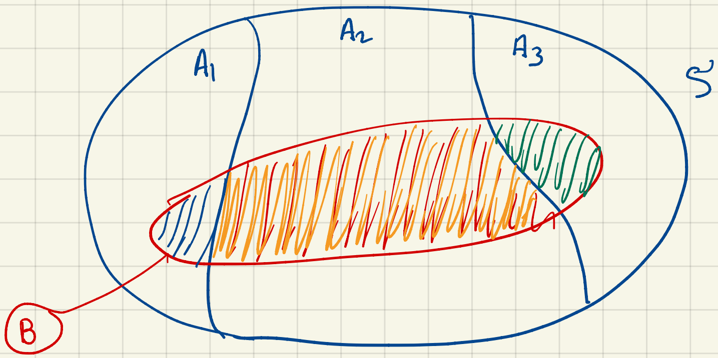
Independence when Conditioned on coin.

$$\underline{\underline{P(H_1 \mid H_2, F)}} = \frac{P(H_1, H_2, F)}{P(H_2, F)} =$$

$$\frac{P(F)\,P(H_1, H_2 \mid F)}{P(H_2 \mid F)\,P(F)} = \frac{P(H_1, H_2 \mid F)}{P(H_2 \mid F)} =$$

$$= \frac{P(H_1 \mid F)\,P(H_2 \mid F)}{P(H_2 \mid F)} = \underline{\underline{P(H_1 \mid F)}}$$

Law of total probability



A's are

1) exclusive $\quad A_i \cap A_j = \emptyset$

2) $\bigcup A_i = S \quad$ (exhaustive)

So the A's partition S

$$B = \underline{(A_1 \cap B)} \cup \underline{(A_2 \cap B)} \cup \underline{(A_3 \cap B)}$$

exclusive

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B)$$

$$= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)$$

$$P(B) = \sum_i P(B|A_i)P(A_i)$$

example: $P(HHH) = P(\underline{HHH}|F)P(F) + P(\underline{HHH}|B)P(B)$

$$= \frac{1}{2}\frac{1}{2}\frac{1}{2}\cdot\frac{1}{2} + 1.1.1.\frac{1}{2} = \frac{1}{16} + \frac{1}{2} = 9/16.$$

Bayes Rule

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

Posterior

Likelihood

prior

$$\boxed{P(A_i|B)} = \frac{P(B|A_i)\,\boxed{P(A_i)}}{P(B) = \sum_i P(B|A_i)\,P(A_i)} \qquad \text{(Bayes Rule)}$$

Usual Setting:

- A's are not observable, but they are governed by some probabilistic law

  e.g. $P(A_i)$ is known (prior)

- B can be observed, we know how B behaves conditioned on A.

Bayes: Given what we observe, what can we say about what we don't know?