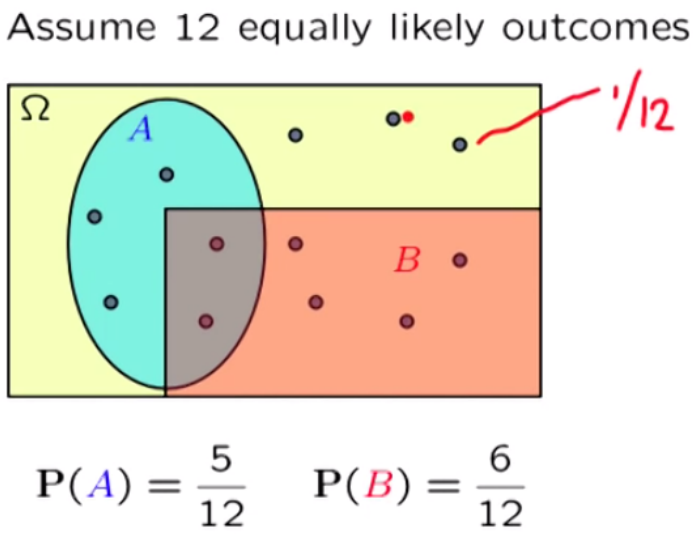
Unit 2 Conditioning and Independence
Conditioning
- revising a model based on new information
Independence
- assume outcomes are not related to each other to simplify complex models

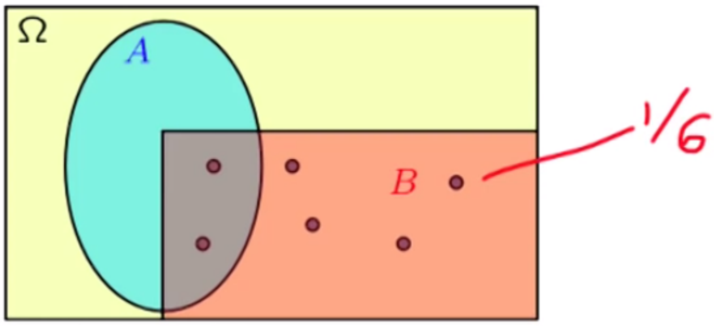
Conditional probability, 3 important rules
- multiplication rule
- total probability theorem
- Bayes' rule - foundation of inference theory

Use new information to revise a model



Probability of A given B has happened
P(A|B)

If told B occurred:



$$\mathbf{P}(A \mid B) = \frac{2}{6} = \frac{1}{3} \qquad \mathbf{P}(B \mid B) = 1$$

Definition of conditional probability

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

defined only when $\mathbf{P}(B) > 0$

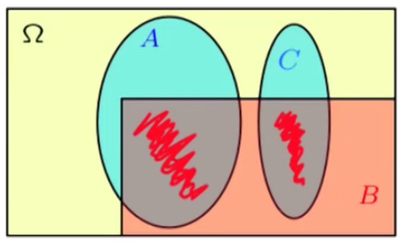
Conditional probabilities share properties of ordinary probabilities i.e. the same axioms apply
(for disjoint A and C below)

Conditional probabilities share properties of ordinary probabilities

$\mathbf{P}(A \mid B) \geq 0$ assuming $\mathbf{P}(B) > 0$

$$\mathbf{P}(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

$$\mathbf{P}(B \mid B) = \frac{P(B \cap B)}{P(B)} = 1$$



If $A \cap C = \varnothing$, then $\mathbf{P}(A \cup C \mid B) = \mathbf{P}(A \mid B) + \mathbf{P}(C \mid B)$

$$= \frac{P((A \cup C) \cap B)}{P(B)} = \frac{P((A \cap B) \cup (C \cap B))}{P(B)} = \frac{P(A \cap B) + P(C \cap B)}{P(B)} =$$

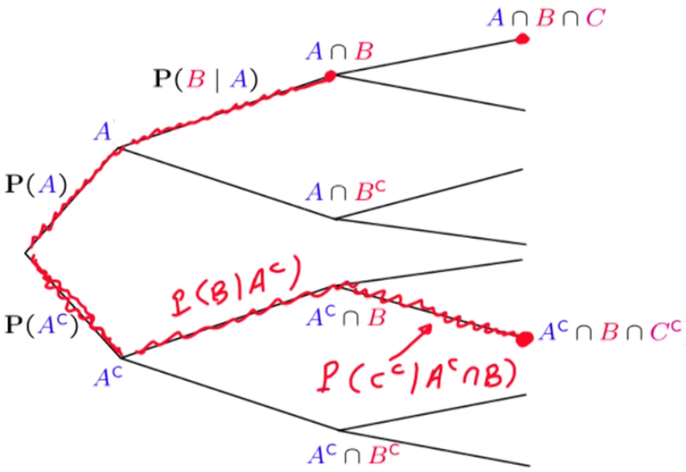
$$= P(A \mid B) + P(C \mid B)$$

also finite
countable additivity

The multiplication rule

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

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



$$\mathbf{P}(\underline{(A^c \cap B)} \cap \underline{C^c}) =$$
$$= P(A^c \cap B)\, P(C^c \mid A^c \cap B)$$
$$= P(A^c) \cdot P(B \mid A^c)\, P(C^c \mid A^c \cap B)$$

$$P(A_1 \cap A_2 \cap \cdots \cap A_n)$$
$$= P(A_1) \prod_{i=2}^{n} P(A_i \mid A_1 \cap \cdots \cap A_{i-1})$$

General form for multiplication rule is a product of the probability of A[1] with the probability that all previous events A[i] up to i=n have already occurred

1. $\mathbf{P}(A \cap B \cap C^c) = \mathbf{P}(A \cap B)\ \mathbf{P}(C^c \mid A \cap B)$

   True ✔ **Answer:** True

2. $\mathbf{P}(A \cap B \cap C^c) = \mathbf{P}(A)\ \mathbf{P}(C^c \mid A)\ \mathbf{P}(B \mid A \cap C^c)$

   True ✔ **Answer:** True

3. $\mathbf{P}(A \cap B \cap C^c) = \mathbf{P}(A)\ \mathbf{P}(C^c \cap A \mid A)\ \mathbf{P}(B \mid A \cap C^c)$

   False ✖ **Answer:** True

4. $\mathbf{P}(A \cap B \mid C) = \mathbf{P}(A \mid C)\ \mathbf{P}(B \mid A \cap C)$

   False ✖ **Answer:** True

Solution:

1. True. This is the usual multipication rule applied to the two events $A \cap B$ and C^c.
2. True. This is the usual multiplication rule.
3. True. This is because

$$\mathbf{P}(C^c \cap A \mid A) = \frac{\mathbf{P}(C^c \cap A \cap A)}{\mathbf{P}(A)} = \frac{\mathbf{P}(C^c \cap A)}{\mathbf{P}(A)} = \mathbf{P}(C^c \mid A).$$

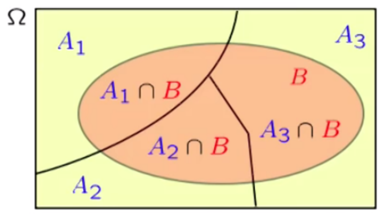
   So, this statement is equivalent to the one in part 2.

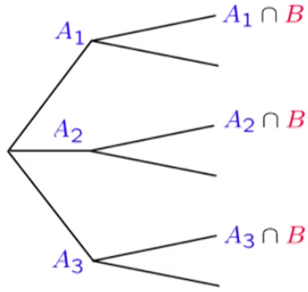
4. True. This is the usual multiplication rule $\mathbf{P}(A \cap B) = \mathbf{P}(A)\,\mathbf{P}(B \mid A)$, applied to a model/universe in which event C is known to have occurred.

Total Probability Theorem

- disjoint events that cover all possible outcomes (partition)
- if there were infinite partitions then it would be an infinite sum across all scenarios

Total probability theorem





- Partition of sample space into $A_1, A_2, A_3, \ldots$
- Have $\mathbf{P}(A_i)$, for every i
- Have $\mathbf{P}(B \mid A_i)$, for every i

$$\mathbf{P}(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3)$$
$$= P(A_1)P(B \mid A_1) + \cdots + \cdots$$

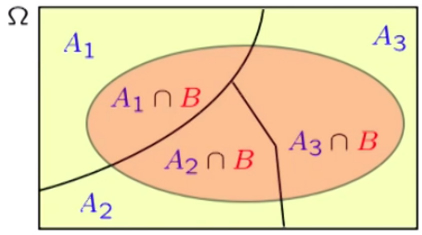
$\sum_i P(A_i) = 1$ weights

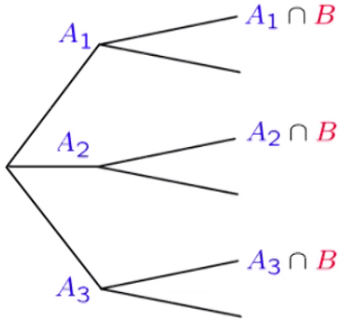
$$\boxed{\mathbf{P}(B) = \sum_i \mathbf{P}(A_i)\,\mathbf{P}(B \mid A_i)}$$

weighted average of $P(B \mid A_i)$

Bayes' Rule

Bayes' rule





- Partition of sample space into A_1, A_2, A_3
- Have $\mathbf{P}(A_i)$, for every i initial "beliefs"
- Have $\mathbf{P}(B \mid A_i)$, for every i

revised "beliefs," given that B occurred:

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

$$\boxed{\mathbf{P}(A_i \mid B) = \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\sum_j \mathbf{P}(A_j)\mathbf{P}(B \mid A_j)}}$$

- Bayesian inference
  - initial beliefs $\mathbf{P}(A_i)$ on possible causes of an observed event B
  - model of the world under each A_i: $\mathbf{P}(B \mid A_i)$

$$A_i \xrightarrow[\mathbf{P}(B \mid A_i)]{\text{model}} B$$

  - draw conclusions about causes

$$B \xrightarrow[\mathbf{P}(A_i \mid B)]{\text{inference}} A_i$$