Machine Intelligence

4. Background for Reasoning Under Uncertainty-probabilities

A Reminder of some basic probability concepts

Álvaro Torralba

AALBORG UNIVERSITET

Fall 2022

Thanks to Thomas D. Nielsen and Jörg Hoffmann for slide sources

Summary

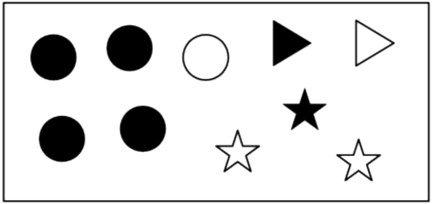
- The following slides recap some basic concepts in probability that you likely are already familiar with.
- During the lecture we will go quickly over these concepts, so please, go over these slides beforehand so that you don't get lost during the lecture.

Probability theory is built on the foundation of variables and worlds.

Worlds described by the variables:

- Filled: {true, false}
- Shape: {circle, triangle, star}

as well as position.



Probability measures

Ω: set of all possible worlds (for a given, fixed set of variables). A **probability measure over** Ω, is a function P, that assigns **probability values**

$$P(\Omega') \in [0, 1]$$

to subsets $\Omega' \subseteq \Omega$, such that

Axiom 1: $P(\Omega) = 1$.

Axiom 2: if $\Omega_1 \cap \Omega_2 = \emptyset$, then $P(\Omega_1 \cup \Omega_2) = P(\Omega_1) + P(\Omega_2)$.

Simplification for finite Ω

If all variables have a finite domain, then

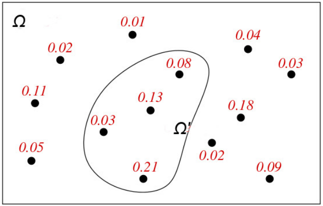
- Ω is finite, and
- a probability distribution is defined by assigning a probability value

$$P(\omega)$$

  to each individual possible world $\omega \in \Omega$.

For any $\Omega' \subseteq \Omega$ then

$$P(\Omega') = \sum_{\omega \in \Omega'} P(\omega)$$



$$P(\Omega') = 0.08 + 0.13 + 0.03 + 0.21 = 0.45$$

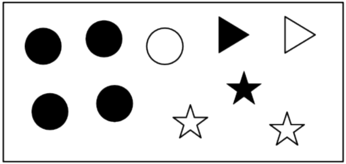
From now on, we will only consider variables with finite domains.

Probabilities of Propositions

A probability distribution over possible worlds defines probabilities for propositions α:

$$P(\alpha) = \sum_{\omega \in \Omega : \alpha \text{ is true in } \omega} P(\omega)$$

Example



Assume probability for each world is 0.1:

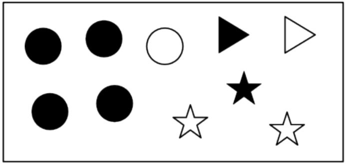
- $P(\mathit{Shape} = \mathit{circle}) =$

Probabilities of Propositions

A probability distribution over possible worlds defines probabilities for propositions α:

$$P(\alpha) = \sum_{\omega \in \Omega : \alpha \text{ is true in } \omega} P(\omega)$$

Example



Assume probability for each world is 0.1:

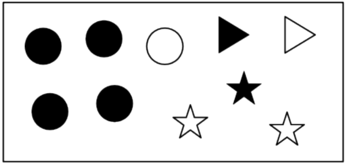
- $P(\mathit{Shape} = \mathit{circle}) = 0.5$
- $P(\mathit{Filled} = \mathit{false}) =$

Probabilities of Propositions

A probability distribution over possible worlds defines probabilities for propositions α:

$$P(\alpha) = \sum_{\omega \in \Omega : \alpha \text{ is true in } \omega} P(\omega)$$

Example



Assume probability for each world is 0.1:

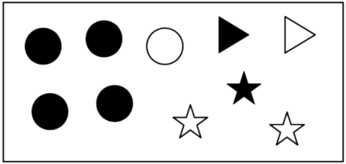
- $P(\mathit{Shape} = \mathit{circle}) = 0.5$
- $P(\mathit{Filled} = \mathit{false}) = 0.4$
- $P(\mathit{Shape} = c \wedge \mathit{Filled} = f) =$

Probabilities of Propositions

A probability distribution over possible worlds defines probabilities for propositions α:

$$P(\alpha) = \sum_{\omega \in \Omega : \alpha \text{ is true in } \omega} P(\omega)$$

Example



Assume probability for each world is 0.1:

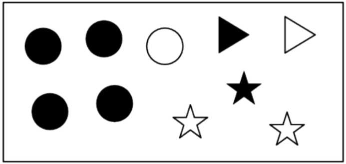
- $P(\mathit{Shape} = \mathit{circle}) = 0.5$
- $P(\mathit{Filled} = \mathit{false}) = 0.4$
- $P(\mathit{Shape} = c \wedge \mathit{Filled} = f) = 0.1$

Probabilities of Propositions

A probability distribution over possible worlds defines probabilities for propositions α:

$$P(\alpha) = \sum_{\omega \in \Omega : \alpha \text{ is true in } \omega} P(\omega)$$

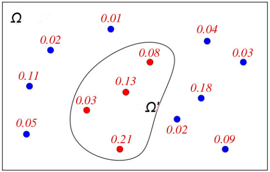
Example



Assume probability for each world is 0.1:

- $P(\mathit{Shape} = \mathit{circle}) = 0.5$
- $P(\mathit{Filled} = \mathit{false}) = 0.4$
- $P(\mathit{Shape} = c \wedge \mathit{Filled} = f) = 0.1$

Another example



$$\begin{aligned} P(\mathit{Color} = \mathit{red}) &= 0.08 + 0.13 + 0.03 + 0.21 \\ &= 0.45 \end{aligned}$$

Axiom

If $\mathcal{A}$ and $\mathcal{B}$ are disjoint, then $P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B})$.

Example

Consider a deck with 52 cards. If $\mathcal{A} = \{2, 3, 4, 5\}$ and $\mathcal{B} = \{7, 8\}$, then

$$P(\mathcal{A} \cup \mathcal{B}) =$$

Basic probability axioms

Axiom

If $\mathcal{A}$ and $\mathcal{B}$ are disjoint, then $P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B})$.

Example

Consider a deck with 52 cards. If $\mathcal{A} = \{2, 3, 4, 5\}$ and $\mathcal{B} = \{7, 8\}$, then

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) = \frac{4}{13} + \frac{2}{13} = \frac{6}{13}.$$

Basic probability axioms

Axiom

If $\mathcal{A}$ and $\mathcal{B}$ are disjoint, then $P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B})$.

Example

Consider a deck with 52 cards. If $\mathcal{A} = \{2, 3, 4, 5\}$ and $\mathcal{B} = \{7, 8\}$, then

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) = \frac{4}{13} + \frac{2}{13} = \frac{6}{13}.$$

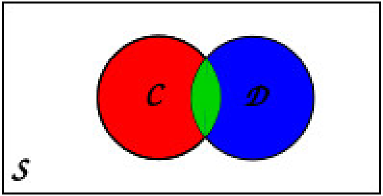
More generally

If $\mathcal{C}$ and $\mathcal{D}$ are not disjoint, then $P(\mathcal{C} \cup \mathcal{D}) = P(\mathcal{C}) + P(\mathcal{D}) - P(\mathcal{C} \cap \mathcal{D})$.

Example

If $\mathcal{C} = \{2, 3, 4, 5\}$ and $\mathcal{D} = \{\spadesuit\}$, then

$$P(\mathcal{C} \cup \mathcal{D}) =$$



Axiom

If $\mathcal{A}$ and $\mathcal{B}$ are disjoint, then $P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B})$.

Example

Consider a deck with 52 cards. If $\mathcal{A} = \{2, 3, 4, 5\}$ and $\mathcal{B} = \{7, 8\}$, then

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) = \frac{4}{13} + \frac{2}{13} = \frac{6}{13}.$$

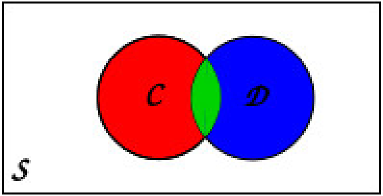
More generally

If $\mathcal{C}$ and $\mathcal{D}$ are not disjoint, then $P(\mathcal{C} \cup \mathcal{D}) = P(\mathcal{C}) + P(\mathcal{D}) - P(\mathcal{C} \cap \mathcal{D})$.

Example

If $\mathcal{C} = \{2, 3, 4, 5\}$ and $\mathcal{D} = \{\spadesuit\}$, then

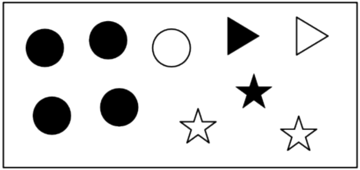
$$P(\mathcal{C} \cup \mathcal{D}) = \frac{4}{13} + \frac{1}{4} - \frac{4}{52} = \frac{25}{52}.$$



Definition

The **conditional probability** of proposition p given e is $P(p \mid e) = \frac{P(p \wedge e)}{P(e)}$

Example



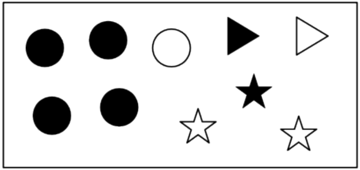
(probability for each world is 0.1)

- If we pick one at random, what is the probability that we get a circle?

Definition

The **conditional probability** of proposition p given e is $P(p \mid e) = \frac{P(p \wedge e)}{P(e)}$

Example



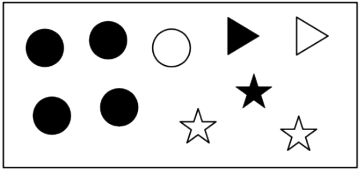
(probability for each world is 0.1)

- If we pick one at random, what is the probability that we get a circle?
  $P(\mathit{S{=}circ.}) = 0.5$
- And what if I tell you, you picked one that is not filled?

Definition

The **conditional probability** of proposition p given e is $P(p \mid e) = \frac{P(p \wedge e)}{P(e)}$

Example



(probability for each world is 0.1)

- If we pick one at random, what is the probability that we get a circle? $P(\mathit{S{=}circ.}) = 0.5$
- And what if I tell you, you picked one that is not filled?

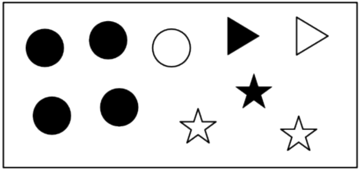
$$P(\mathit{S{=}circle} \mid \mathit{Fill} = f) = \frac{P(\mathit{S{=}circle} \wedge \mathit{Fill} = f)}{P(\mathit{Fill} = f)} = \frac{0.1}{0.4} = 0.25$$

- And what if I also tell you is a star?

Definition

The **conditional probability** of proposition p given e is $P(p \mid e) = \frac{P(p \wedge e)}{P(e)}$

Example



(probability for each world is 0.1)

- If we pick one at random, what is the probability that we get a circle? $P(\mathit{S{=}circ.}) = 0.5$
- And what if I tell you, you picked one that is not filled?

$$P(\mathit{S{=}circle} \mid \mathit{Fill} = f) = \frac{P(\mathit{S{=}circle} \wedge \mathit{Fill} = f)}{P(\mathit{Fill} = f)} = \frac{0.1}{0.4} = 0.25$$

- And what if I also tell you is a star?

$$P(\mathit{S{=}circle} \mid \mathit{Fill} = f \wedge \mathit{S{=}star}) = \frac{P(\mathit{S{=}star} \wedge \mathit{S{=}circle} \wedge \mathit{Fill} = f)}{P(\mathit{Fill} = f \wedge \mathit{S{=}star})} = \frac{0}{0.2} = 0$$

Definition. *Events a and b are independent if $P(a \wedge b) = P(a)P(b)$.*

Definition. *Events a and b are independent if $P(a \wedge b) = P(a)P(b)$.*

Proposition. *Given independent events a and b where $P(b) \neq 0$, we have $P(a \mid b) = P(a)$.*

Independence

Definition. *Events a and b are independent if $P(a \wedge b) = P(a)P(b)$.*

Proposition. *Given independent events a and b where $P(b) \neq 0$, we have $P(a \mid b) = P(a)$.*

Proof. By definition, $P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$,

Definition. *Events a and b are independent if $P(a \wedge b) = P(a)P(b)$.*

Proposition. *Given independent events a and b where $P(b) \neq 0$, we have $P(a \mid b) = P(a)$.*

Proof. By definition, $P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$, which by independence is equal to $\frac{P(a)P(b)}{P(b)} = P(a)$. ($\rightarrow$ Similarly, if $P(a) \neq 0$, we have $P(b \mid a) = P(b)$.)

Definition. *Events a and b are independent if $P(a \wedge b) = P(a)P(b)$.*

Proposition. *Given independent events a and b where $P(b) \neq 0$, we have $P(a \mid b) = P(a)$.*

Proof. By definition, $P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$, which by independence is equal to $\frac{P(a)P(b)}{P(b)} = P(a)$. ($\rightarrow$ Similarly, if $P(a) \neq 0$, we have $P(b \mid a) = P(b)$.)

Example: Two dice rolls are independent!

- $P(\mathit{Dice1} = 6 \wedge \mathit{Dice2} = 6) = 1/36$.