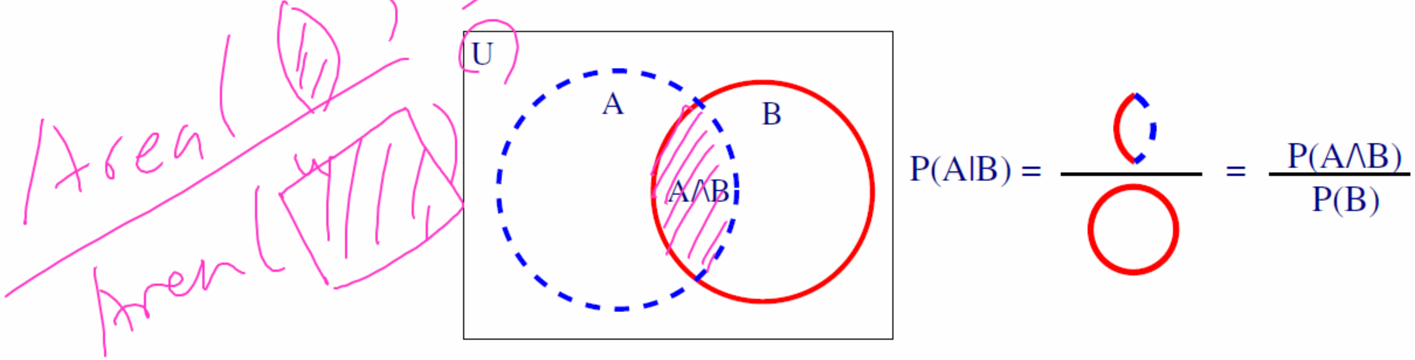
Conditional Probability



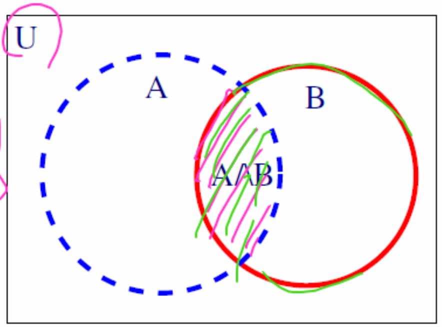
- Think about the **area** occupied by each event.
- The bounding rectangle U has an area of 1, thus

$$P(A) = \frac{\text{Area of } A}{\text{Area of } U} = \frac{\text{Area of } A}{1} = \text{Area of } A$$

- $P(A|B)$ means B now takes on the role of U. Within this limited event space, what is the probability of A.

$$\frac{\text{Area}(\ )}{\text{Area}(\ )} = P(A \wedge B)$$

Conditional Probability



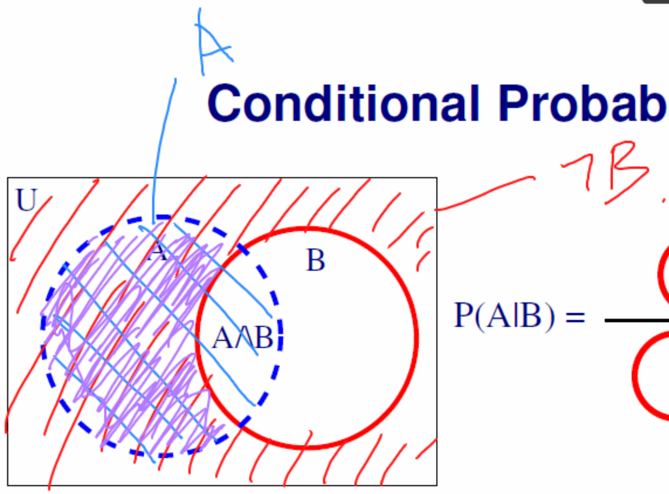
$$P(A|B) = \frac{\text{Area}(\ )}{\text{Area}(\ )} = \frac{P(A \wedge B)}{P(B)}$$

- Think about the **area** occupied by each event.
- The bounding rectangle U has an area of 1, thus

$$P(A) = \frac{\text{Area of } A}{\text{Area of } U} = \frac{\text{Area of } A}{1} = \text{Area of } A$$

- $P(A|B)$ means B now takes on the role of U. Within this limited event space, what is the probability of A.

Conditional Probability



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

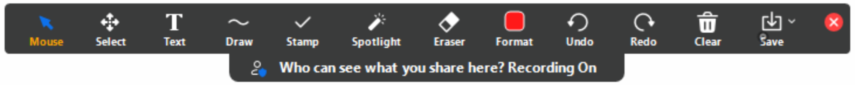
$$P(A|\neg B) =$$

$$\frac{P(A,\neg B)}{P(\neg B)}$$

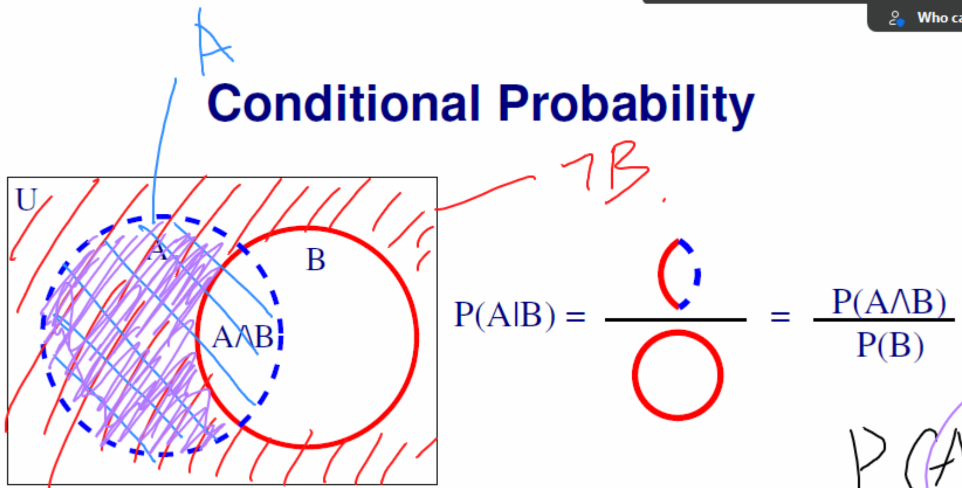
- Think about the **area** occupied by each event.
- The bounding rectangle U has an area of 1, thus

$$P(A) = \frac{\text{Area of } A}{\text{Area of } U} = \frac{\text{Area of } A}{1} = \text{Area of } A$$

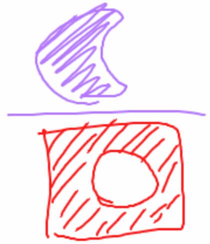
- $P(A|B)$ means B now takes on the role of U. Within this limited event space, what is the probability of A.



Conditional Probability



$P(A|\neg B) =$



$$\frac{P(A,\neg B)}{P(\neg B)}$$

- Think about the **area** occupied by each event.
- The bounding rectangle U has an area of 1, thus

$$P(A) = \frac{\text{Area of } A}{\text{Area of } U} = \frac{\text{Area of } A}{1} = \text{Area of } A$$

- $P(A|B)$ means B now takes on the role of U. Within this limited event space, what is the probability of A.

The Axioms of Probability

All axioms

1. All probabilities are between 0 and 1

$$0 \leq P(A) \leq 1$$

Boolean

2. For a valid proposition A ($\mathbf{T}$ under all interpretations): $P(A) = 1$, and for a inconsistent proposition A ($\mathbf{F}$ under all interpretations): $P(A) = 0$.

3. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

Other properties follow from these three axioms.

The Axioms of Probability

All axioms

1. All probabilities are between 0 and 1

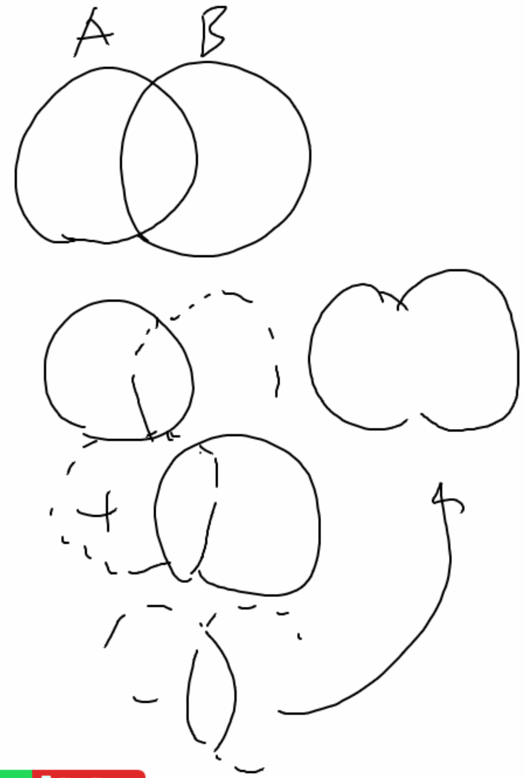
$$0 \leq P(A) \leq 1$$

Boolean

2. For a valid proposition A (**T** under all interpretations): $P(A) = 1$, and for a inconsistent proposition A (**F** under all interpretations): $P(A) = 0$.

3. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

Other properties follow from these three axioms.



So here you had this here you got that you have these

$A \vee \neg A = \text{True}$

Other Properties

$A \wedge \neg A = \text{False}$

- From the axioms,

$$
\begin{aligned}
P(A \vee \neg A) &= P(A) + P(\neg A) - P(A \wedge \neg A) \\
P(\mathbf{T}) &= P(A) + P(\neg A) - P(\mathbf{F}) \\
1 &= P(A) + P(\neg A) \; - 0 \\
P(\neg A) &= 1 - P(A) \quad - P(A)
\end{aligned}
$$

- More generally, the **sum** of probabilities $P(X = v)$ is 1, for all values v the random variable X can take:

$$\left[\sum_{v \in V} P(X = v) \right] = 1,$$

where V is the set of all possible values X can take.

Other Properties

- From the axioms,

$$
\begin{aligned}
P(A \vee \neg A) &= P(A) + P(\neg A) - P(A \wedge \neg A) \\
P(\mathbf{T}) &= P(A) + P(\neg A) - P(\mathbf{F}) \\
-P(A) + 1 &= \cancel{P(A)} + P(\neg A) - \cancel{P(A)} \\
P(\neg A) &= 1 - P(A)
\end{aligned}
$$

- More generally, the **sum** of probabilities $P(X = v)$ is 1, for all values v the random variable X can take:

$$\left[\sum_{v \in V} P(X = v) \right] = 1,$$

where V is the set of all possible values X can take.

P(A, B)

P(A, B)

A, B	P(A,B)
T T	0.2
T F	0.3
F T	0.1
F F	0.4

1.0

Joint Probability Distribution

For random variables $X_1, X_2, ..., X_n$,

- An **atomic event** is an assignment of particular values to each random variable.
- The **joint probability distribution** $\mathbf{P}(X_1, X_2, ..., X_n)$ completely specifies the probabilities of all **atomic events**.
- Thus,

$$\left[\sum_{(v_1, v_2, \ldots, v_n) \in \mathbf{V}} P(X_1 = v_1, X_2 = v_2, ..., X_n = v_n)\right] = 1,$$

where $\mathbf{V}$ is a set of all possible $n-$vectors that the vector $(X_1, X_2, .., X_n)$ can assume..

$P(A,B)$

$P(A,B)$

A, B	P(A,B)
T T	0.2
T F	0.3
F T	0.1
F F	0.4

1.0

Joint Probability Distribution

For random variables $X_1, X_2, ..., X_n$,

- An **atomic event** is an assignment of particular values to each random variable.
- The **joint probability distribution** $\mathbf{P}(X_1, X_2, ..., X_n)$ completely specifies the probabilities of all **atomic events**.
- Thus,

$$\left[\sum_{(v_1, v_2, ..., v_n) \in \mathbf{V}} P(X_1 = v_1, X_2 = v_2, ..., X_n = v_n)\right] = 1,$$

where $\mathbf{V}$ is a set of all possible $n-$vectors that the vector $(X_1, X_2, .., X_n)$ can assume..

In this case is a 2 dimensional. vector So these are the vectors

Decision Theory: Example

Decision theory = Probability theory + Utility theory

	$\sum$ Utility of Resulting State $\times$ Probability	Expected Utility
Action 1	$10 \times 0.2 + 0 \times 0.8$	2
Action 2	$1000 \times 0.001 + 0 \times 0.999$	1
Action 3	$5 \times 0.799 + 0 \times 0.201$	3.995

Action 3 has the maximum expected utility, thus action 3 will be carried out.

A, B		P(A, B)
T	T	?
T	F	?
F	T	?
F	F	?

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

Tooth	Cav	P(Tooth, Cav)
T	T	0.04
T	F	0.01
F	T	0.06
F	F	0.89

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Mouse Select Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On

A	B	P(A,B)
T	T	?
T	F	?
F	T	?
F	F	?

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

marginal prob.

Abbreviations: $C = Cavity, T = Toothache$

Tooth	Cav	P(Tooth, Cav)
T	T	0.04
T	F	0.01
F	T	0.06
F	F	0.89

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

These are called the marginal probabilities.

You are screen sharing Stop Share

$P(\text{Toothache}, \text{Cavity})$
$\parallel$
$P(T \wedge C)$
$+$
$+P(T \wedge \neg C)$

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Joint Probability Distribution: Example

$P(\text{Toothache}, \text{Cavity}) = P(T \wedge C) + P(T \wedge \neg C)$

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

$P(C) = P(C \wedge T) + P(C \wedge \neg T) = 0.04 + 0.06$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Okay, so that would be point 0 4 plus 0 point 0 6.

$P(\text{Toothache}, \text{Cavity})$
$\|$
$P(T \wedge C)$
$+ P(T \wedge \neg C)$

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

$p(c) = p(C \wedge T) + p(C \wedge \neg T)$
$= 0.04 + 0.06$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Joint Probability Distribution: Example

	Toothache	$\neg$Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
$\neg$Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$ $= P(C,T) = 0.04$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Joint Probability Distribution: Example

	Toothache	$\neg$Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
$\neg$Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$ $= P(C,T) = 0.04$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$ $= P(C,T) = 0.04$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Joint Probability Distribution: Example

	Toothache	$\neg$Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
$\neg$Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Friday

Action1: buy a lottery ticket.

Action2: go to dept. picnic, — raffle

x box

wii

0.01

Decision Theory: Example

Decision theory = Probability theory + Utility theory

	$\sum$ Utility of Resulting State $\times$ Probability	Expected Utility
Action 1	$10 \times 0.2 + 0 \times 0.8$	2
Action 2	$1000 \times 0.001 + 0 \times 0.999$	1
Action 3	$5 \times 0.799 + 0 \times 0.201$	3.995 ⟵

Action 3 has the maximum expected utility, thus action 3 will be carried out.

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

$p(c|T) = 0.8$

$p(\neg c|T) = \frac{p(\neg c \wedge T)}{p(T)}$

$= \frac{0.01}{\quad}$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

$P(C|T) = 0.8$

$P(\neg C|T) = \frac{P(\neg C \wedge T)}{P(T)}$

$= \frac{0.01}{0.05}$

$= \frac{1}{5} = \frac{2}{10} = 0.2$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

$P(C|T) = 0.8$
$+ \quad +$
$P(\neg C|T) = 0.2$
$\|$
$P(C|T) + P(\neg C|T) = 1.0$

Joint Probability Distribution: Example

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

$P(C|T) = 0.8$

$P(\neg C|T) = \frac{P(\neg C \wedge T)}{P(T)}$

$= \frac{0.01}{0.05}$

$= \frac{1}{5} = \frac{2}{10} = 0.2$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Okay.

Joint Probability Distribution: Example

$P(C|T) = 0.8$
$+$
$P(\neg C|T) = 0.2$
$=$

$p(c|T) + p(\neg c|T) = 1.0$

	Toothache	¬Toothache	Sum
Cavity	0.04	0.06	$P(C) = 0.1$
¬Cavity	0.01	0.89	$P(\neg C) = 0.9$
Sum	$P(T) = 0.05$	$P(\neg T) = 0.95$	$\sum = 1.0$

Abbreviations: $C = Cavity, T = Toothache$

- $P(C \vee T) = P(C) + P(T) - P(C \wedge T)$
  $= 0.1 + 0.05 - 0.04 = 0.11$
- $P(C|T) = \frac{P(C \wedge T)}{P(T)} = \frac{0.04}{0.05} = 0.8$
- $P(T|C) = \frac{P(C \wedge T)}{P(C)} = \frac{0.04}{0.1} = 0.5$

$P(C|T) = 0.8$

$P(\neg C|T) = \frac{P(\neg C \wedge T)}{P(T)}$

$= \frac{0.01}{0.05}$

$= \frac{1}{5} = \frac{2}{10} = 0.2$

In practice, writing a full joint probability table like this is impossible (or too much effort): for n boolean random variables, you need 2^n entries.

Right, so let's move on any questions

Bayes' Rule

- From $P(A|B) = \frac{P(A \wedge B)}{P(B)}$ and $P(B|A) = \frac{P(A \wedge B)}{P(A)}$, we get

$$P(A|B)P(B) = P(B|A)P(A)$$

and in turn from which we get the **Bayes' Rule**:

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

Again during probability, divide by the probability of that

Bayes' Rule

$\times P(B)$ $\times P(A)$

- From $P(A|B) = \frac{P(A \wedge B)}{P(B)}$ and $P(B|A) = \frac{P(A \wedge B)}{P(A)}$, we get

$P(B)$ $P(A)$

$$P(A|B)P(B) = P(B|A)P(A)$$

and in turn from which we get the **Bayes' Rule**:

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

Bayes' Rule

$=$

$\times P(\cancel{B})$ $\times P(\cancel{A})$

- From $P(A|B) = \frac{P(A \wedge B)}{P(\cancel{B})}$ and $P(B|A) = \frac{P(A \wedge B)}{P(\cancel{A})}$, we get

  $\times P(B)$ $\times P(A)$

$$P(A|B)P(B) = P(B|A)P(A)$$

and in turn from which we get the **Bayes' Rule**:

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

Bayes' Rule

- From $P(A|B) = \frac{P(A \wedge B)}{P(B)}$ and $P(B|A) = \frac{P(A \wedge B)}{P(A)}$, we get

$$\underbrace{P(A|B)P(B)}_{P(A)} = \underbrace{P(B|A)P(A)}_{P(A)}$$

and in turn from which we get the **Bayes' Rule**:

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

Bayes' Rule

- From $P(A|B) = \frac{P(A \wedge B)}{P(B)}$ and $P(B|A) = \frac{P(A \wedge B)}{P(A)}$, we get

$$P(A|B)P(B) = P(B|A)\cancel{P(A)}$$

P(A) P(A)

and in turn from which we get the **Bayes' Rule**:

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

Probability: Notations[a]

- **Random variable**: variable that can take on different values
  - $A, B, \ldots$: boolean values ($\mathbf{T}$ or $\mathbf{F}$).
  - $X, Y, \ldots$: numerical values or other multi-valued enumerations (1, 2, 0.5, Cloudy, Rainy, Sunny, ...)
- $P(X = v)$: **probability** of the variable X having value v.
  - This can be viewed as an *event*.
  - For boolean variables, $P(A)$ means $P(A = \mathbf{T})$, and $P(\neg A)$ means $P(A = \mathbf{F})$.
- $\mathbf{P}(X)$: **probability distribution**, a full list of probabilities for all possible values that X can take (note that $\mathbf{P}$ is in **bold**).

[a] All conventions follow Russel & Norvig

Then you have a probability, distribution.

Bayes' Rule

- From $P(A|B) = \frac{P(A \wedge B)}{P(B)}$ and $P(B|A) = \frac{P(A \wedge B)}{P(A)}$, we get

$$P(A|B)P(B) = P(B|A)\cancel{P(A)}$$

$$\underset{P(A)}{} \qquad \underset{P(A)}{}$$

and in turn from which we get the **Bayes' Rule**:

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

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)} = P(X,Y|E)$$

This rule follows from

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

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

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

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

$$P(Y|X,E) = \frac{P(X,Y|E)}{P(X|E)}$$

Note: $\mathbf{P}(Y|X,E) = \mathbf{P}(Y|\underbrace{(X,E)}).$

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)} = P(X,Y|E)$$

This rule follows from

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

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

$$P(Y|X,E) = \frac{P(X,Y|E)}{P(X|E)}$$

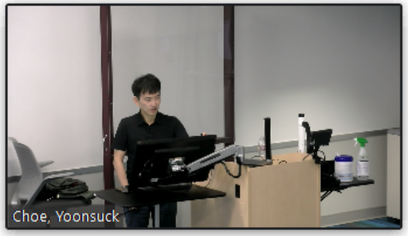
$\times\, P(X|E)$

$\times\, P(X|E)$

Note: $\mathbf{P}(Y|X,E) = \mathbf{P}(Y|\underbrace{(X,E)})$.

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)} = P(X,Y|E)$$



This rule follows from

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

$$\begin{aligned}
\mathbf{P}(A,B|E) &= \frac{\mathbf{P}(A,B,E)}{\mathbf{P}(E)} \\
&= \frac{\mathbf{P}(A|B,E)\mathbf{P}(B,E)}{\mathbf{P}(E)} \\
&= \mathbf{P}(A|B,E)\mathbf{P}(B|E)
\end{aligned}$$

$$P(Y|X,E) = \frac{P(XY|E)}{P(\cancel{X|E})}$$

$$\times\, P(X|E) \qquad \times\, p(\cancel{X|E})$$

$$P(Y|X,E)\,p(X|E) = p(XY|E)$$

Note: $\mathbf{P}(Y|X,E) = \mathbf{P}(Y|\underbrace{(X,E)})$.

So this is the left hand side, then that is the right hand side

Mouse Select Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On

Saved as PNG Show in Folder

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)} = P(X,Y|E)$$

This rule follows from

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

$$\begin{aligned}
\mathbf{P}(A,B|E) &= \frac{\mathbf{P}(A,B,E)}{\mathbf{P}(E)} \\
&= \frac{\mathbf{P}(A|B,E)\mathbf{P}(B,E)}{\mathbf{P}(E)} \\
&= \mathbf{P}(A|B,E)\mathbf{P}(B|E)
\end{aligned}$$

$$P(Y|X,E) = \frac{P(X,Y|E)}{P(\cancel{X|E})}$$

$$\times \; P(X|E) \qquad \times \; p(\cancel{X|E})$$

$$P(Y|X,E)\,p(X|E) = p(X,Y|E)$$

Note: $\mathbf{P}(Y|X,E) = \mathbf{P}(Y|\underbrace{(X,E)})$.

So this is the left hand side, then that is the right hand side

You are screening sharing Stop Share

Mouse Select Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On

$$P(Y|X,E) = \frac{P(Y,X,E)}{P(X,E)} = P(X|E)P(E) \quad (2)$$

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)}$$

$$P(Y,X,E) = P(X,Y,E) = \underbrace{P(X|Y,E)P(Y,E)}_{(1)}$$

This rule follows from

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

$$\begin{aligned}
\mathbf{P}(A,B|E) &= \frac{\mathbf{P}(A,B,E)}{\mathbf{P}(E)} \\
&= \frac{\mathbf{P}(A|B,E)\mathbf{P}(B,E)}{\mathbf{P}(E)} \\
&= \mathbf{P}(A|B,E)\mathbf{P}(B|E)
\end{aligned}$$

$$\frac{(1)\ P(X|Y,E)P(Y,E)}{(2)\ P(X|E)P(E)}$$

Note: $\mathbf{P}(Y|X,E) = \mathbf{P}(Y|\underbrace{(X,E)})$.

Given E. right. So this would be 2 and only one most

You are screen sharing Stop Share

$$P(Y|X,E) = \frac{P(Y,X,E)}{P(X,E)} = \frac{P(X|E)P(E)}{}\ ②$$

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)}$$

$$P(Y,X,E) = P(X,Y,E) = \underbrace{P(X|Y,E)P(Y,E)}_{①}$$

① $P(X|Y,E)P(Y,E)$ — $P(Y|E)P(\cancel{E})$

② $P(X|E)P(\cancel{E})$

$\Downarrow$

$P(X|Y,E)P(Y|E)$

This rule follows from

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

$$\begin{aligned}\mathbf{P}(A,B|E) &= \frac{\mathbf{P}(A,B,E)}{\mathbf{P}(E)} \\ &= \frac{\mathbf{P}(A|B,E)\mathbf{P}(B,E)}{\mathbf{P}(E)} \\ &= \mathbf{P}(A|B,E)\mathbf{P}(B|E)\end{aligned}$$

Note: $\mathbf{P}(Y|X,E) = \mathbf{P}(Y|(X,E))$.

$$P(Y|X,E) = \frac{P(Y,X,E)}{P(X,E)} = P(X|E)P(E) \quad ②$$

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)}$$

$$P(Y,X,E) = P(X,Y,E) = \underbrace{P(X|Y,E)P(Y,E)}_{①}$$

① $P(X|Y,E)\,P(Y,E)$ $\quad P(Y|E)P(E)$

② $P(X|E)\,\cancel{P(E)}$

$$\Downarrow$$

$$\frac{P(X|Y,E)P(Y|E)}{P(X|E)}$$

This rule follows from

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

$$\begin{aligned}
\mathbf{P}(A, B|E) &= \frac{\mathbf{P}(A, B, E)}{\mathbf{P}(E)} \\
&= \frac{\mathbf{P}(A|B, E)\mathbf{P}(B, E)}{\mathbf{P}(E)} \\
&= \mathbf{P}(A|B, E)\mathbf{P}(B|E)
\end{aligned}$$

Note: $\mathbf{P}(Y|X, E) = \mathbf{P}(Y|\underbrace{(X, E)})$.

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)}$$

This rule follows from

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

$$\begin{aligned}
\mathbf{P}(A,B|E) &= \frac{\mathbf{P}(A,B,E)}{\mathbf{P}(E)} \\
&= \frac{\mathbf{P}(A|B,E)\mathbf{P}(B,E)}{\mathbf{P}(E)} \\
&= \mathbf{P}(A|B,E)\mathbf{P}(B|E)
\end{aligned}$$

Note: $\mathbf{P}(Y|X,E) = \mathbf{P}(Y|\underbrace{(X,E)}_{})$.

And then just use the definition of conditional probability

Extended Bayes' Rule

$$\mathbf{P}(Y|X,E) = \frac{\mathbf{P}(X|Y,E)\mathbf{P}(Y|E)}{\mathbf{P}(X|E)}$$

This rule follows from

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

$$\begin{aligned}
\mathbf{P}(A,B|E) &= \frac{\mathbf{P}(A,B,E)}{\mathbf{P}(E)} \\
&= \frac{\mathbf{P}(A|B,E)\mathbf{P}(B,E)}{\mathbf{P}(E)} \\
&= \mathbf{P}(A|B,E)\mathbf{P}(B|E)
\end{aligned}$$

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

Note: $\mathbf{P}(Y|X,E) = \mathbf{P}(Y|\underbrace{(X,E)})$.

Then this is the first factor. So basically you can apply all of this tricks directly to that which is what I do

Examples

- Boolean:
  $P(Infected) = 0.01, P(\neg Infected) = 0.99.$
- Multi valued:
  $P(Dice = 1) = \frac{1}{6}, P(Dice = 2) = \frac{1}{6}, \ldots$
- Multi valued:
  $P(Weather = Sunny) = 0.7,$
  $P(Weather = Rainy) = 0.2, \ldots$

$P(X) + P(\neg X) = 1.0$

$6 \times \frac{1}{6} = 1.0$

So one over, 6 times 6

$$\begin{cases} P(T|D) = 0.99 \\ P(\neg T|\neg D) = 0.99 \end{cases}$$

want to know

$P(D|T)$

0.99

Example: Application of Bayes Rule

After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease, and that the test is 99% accurate (i.e. the probability of testing positive given that you have the disease is 0.99, as is the probability of testing negative given you don't have the disease). The good news is that this is a rare disease, striking only 1 in 10,000 people. Why is it good news that the disease is rare? What are the chances that you actually have the disease?

T: tested positive, $\neg T$: tested negative, D : have disease, $\neg D$: clean.

$P(D) = 0.001$

$\rightarrow P(D|T)$

Solution: Good News and Bad News (cont'd)

evidence

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

- $P(T|D) = 0.99$, $P(\neg T|\neg D)$, and $P(D) = 0.0001$ are given.
- From these, we can get $P(\neg T|D) = 0.01$, $P(T|\neg D) = 0.01$, and $P(\neg D) = 0.9999$.

Since $P(T|D)$ and $P(D)$ are give, we only need to calculate $P(T)$.

And I, after receiving some evidence, you have to update your belief to some other value.

Solution: Good News and Bad News (cont'd)

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

- $P(T|D) = 0.99$, $P(\neg T|\neg D)$, and $P(D) = 0.0001$ are given.

$P(\neg T|\neg D) = 0.99$

- From these, we can get $P(\neg T|D) = 0.01$, $P(T|\neg D) = 0.01$, and $P(\neg D) = 0.9999$.

Since $P(T|D)$ and $P(D)$ are give, we only need to calculate $P(T)$.

Solution: Good News and Bad News (cont'd)

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

- $P(T|D) = 0.99$, $P(\neg T|\neg D)$, and $P(D) = 0.0001$ are given.
- From these, we can get $P(\neg T|D) = 0.01$, $P(T|\neg D) = 0.01$, and $P(\neg D) = 0.9999$.

Since $P(T|D)$ and $P(D)$ are give, we only need to calculate $P(T)$.

Mouse Select Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On

Solution: Good News and Bad News (cont'd)

Observation [a]: $P(T) = P(T|D)P(D) + P(T|\neg D)P(\neg D)$.

Thus, $P(T) = 0.99 \times 0.0001 + 0.01 \times 0.9999 = 0.010098$,

and with this,

$$P(D|T) = \frac{0.99 \times 0.0001}{0.010098} = 0.0098,$$

which is slightly less than 1%.

Exercise: how accurate should the test be so that $P(D|T)$ is greater than 0.95 (i.e. 95%)?

---

[a] $P(T) = P(T \wedge D) + P(T \wedge \neg D)$.

$P(T|D)P(D)$ $P(T|\neg D)P(\neg D)$

Then from that you get P of t given, not the p of not d, which becomes that so you can click

You are screen sharing Stop Share

Solution: Good News and Bad News (cont'd)

Observation [a]: $P(T) = P(T|D)P(D) + P(T|\neg D)P(\neg D)$.

Thus, $P(T) = 0.99 \times 0.0001 + 0.01 \times 0.9999 = 0.010098$,

and with this,

$$P(D|T) = \frac{0.99 \times 0.0001}{0.010098} = 0.0098,$$

which is slightly less than 1%.

Exercise: how accurate should the test be so that $P(D|T)$ is greater than 0.95 (i.e. 95%)?

---

[a] $P(T) = P(T \wedge D) + P(T \wedge \neg D)$.

$P(T|D)P(D)$ $\quad$ $P(T|\neg D)P(\neg D)$

Solution: Good News and Bad News (cont'd)

Observation [a]: $P(T) = P(T|D)P(D) + P(T|\neg D)P(\neg D)$.

Thus, $P(T) = 0.99 \times 0.0001 + 0.01 \times 0.9999 = 0.010098$,

and with this,

$$P(D|T) = \frac{0.99 \times 0.0001}{0.010098} = 0.0098,$$

$1 - P(\neg T|\neg D) = 0.95$

which is slightly less than 1%.

Exercise: how accurate should the test be so that $P(D|T)$ is greater than 0.95 (i.e. 95%)?

---

[a] $P(T) = P(T \wedge D) + P(T \wedge \neg D)$.

$P(T|D)P(D)$ $\quad$ $P(T|\neg D)P(\neg D)$

Solution: Good News and Bad News (cont'd)

Observation [a]: $P(T) = P(T|D)P(D) + P(T|\neg D)P(\neg D)$.

Thus, $P(T) = 0.99 \times 0.0001 + 0.01 \times 0.9999 = 0.010098$,

and with this,

P(+|D)=0.99

P(T|D)　P(D)

$$P(D|T) = \frac{0.99 \times 0.0001}{0.010098} = 0.0098, < 0.01$$

1%

P(T)

which is slightly less than 1%.

Exercise: how accurate should the test be so that $P(D|T)$ is greater than 0.95 (i.e. 95%)?

[a] $P(T) = P(T \wedge D) + P(T \wedge \neg D)$.

Solution: Good News and Bad News (cont'd)

Observation [a]: $P(T) = \overbrace{P(T|D)P(D) + P(T|\neg D)P(\neg D)}^{0.9999}$.

Thus, $P(T) = \cancel{0.99}^{\,0.95} \times 0.0001 + \cancel{0.01}^{\,0.05} \times 0.9999 = 0.010098$,

and with this,

$P(+|D) = 0.95$

$$P(D|T) = \frac{\overset{P(T|D)\ 0.95}{\cancel{0.99}} \times \overset{P(D)}{0.0001}}{\underset{P(T)}{0.010098}} = 0.0098 < 0.01 \quad 1\%$$

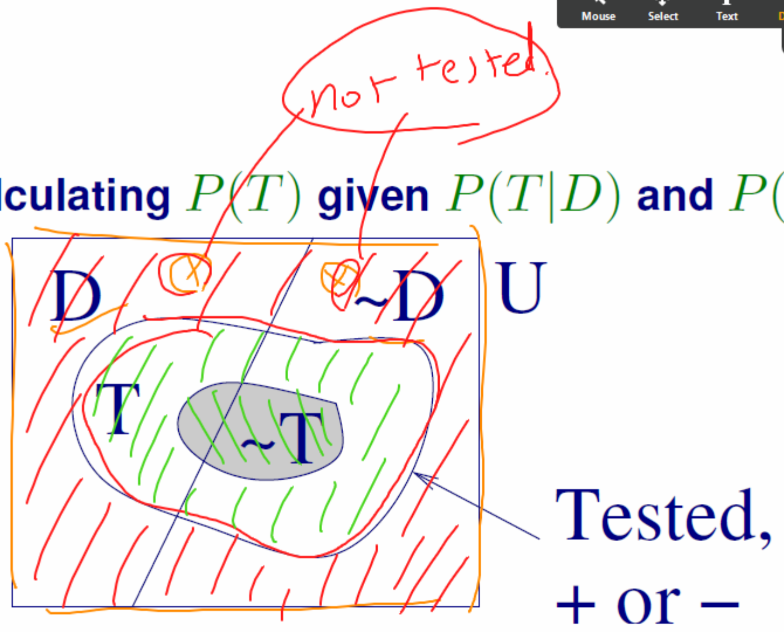
which is slightly less than 1%.

Exercise: how accurate should the test be so that $P(D|T)$ is greater than 0.95 (i.e. 95%)?

---

[a] $P(T) = P(T \wedge D) + P(T \wedge \neg D)$.

Calculating $P(T)$ given $P(T|D)$ and $P(D)$



$$
\begin{aligned}
P(T) &= P(T \land D) + P(T \land \neg D) \\
&= P(T|D)P(D) + P(T|\neg D)P(\neg D)
\end{aligned}
$$

- $\{D\} \cup \{\neg D\}$ completely account for the whole population, but $\{T\} \cup \{\neg T\}$ does not cover the whole population (because you **did not test everyone!**).

Logical Connectives and Conditional Probability

A	B	P(A,B)
T	T	
T	F	
F	T	
F	F	

$\Sigma = 1.0$

- Logical connectives can be used:

$$P(A \vee B), P(A \wedge \neg B), P(Cavity \wedge \neg Insured), \text{ etc.}$$

- Conditional Probability $P(A|B)$ (read *probability of* A *given* B):

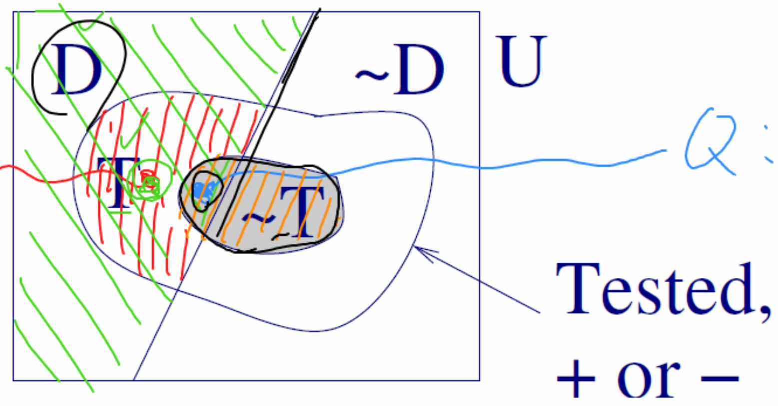
$$P(Cavity|Toothache) = 0.8$$

- As new evidence comes in, the conditional probability gets updated:

$$P(Cavity|\underbrace{Toothache \wedge BadBreath})$$

Calculating $P(T)$ given $P(T|D)$ and $P(D)$

Q
① D? – Yes
② T? – Yes



Q:
① D? – Yes
② T? – No
¬T: Yes
$P(\neg T | D)$

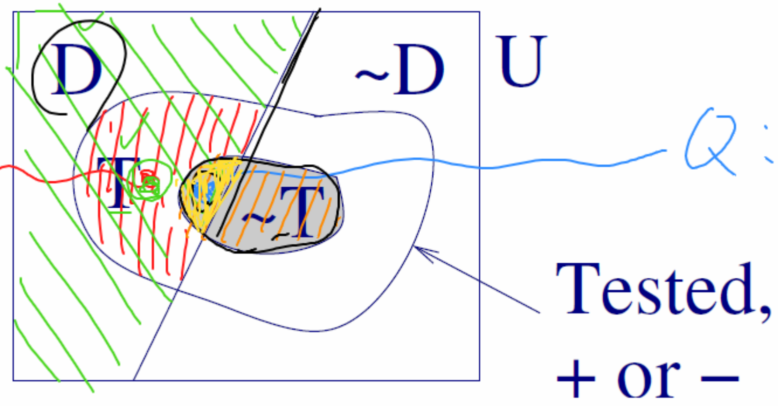
$$P(T) = P(T \wedge D) + P(T \wedge \neg D)$$
$$= P(T|D)P(D) + P(T|\neg D)P(\neg D)$$

- $\{D\} \cup \{\neg D\}$ completely account for the whole population, but $\{T\} \cup \{\neg T\}$ does not cover the whole population (because you **did not test everyone!**).

What is the probability of not testing positive? Given the disease

Calculating $P(T)$ given $P(T|D)$ and $P(D)$

Q
① D? – Yes
② T? – Yes



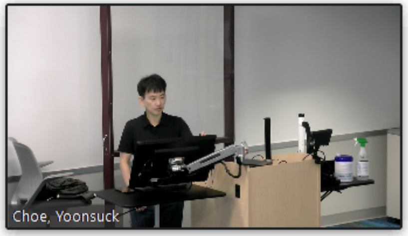
Q:
① D? – Yes
② T? – No
¬T: Yes

$P(\neg T|D)$

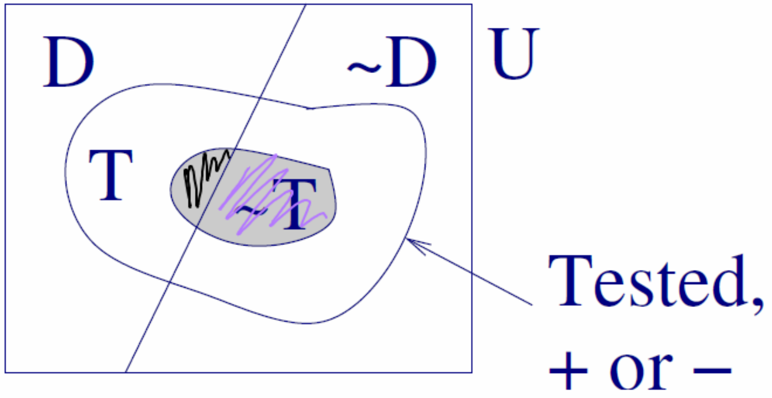
$$P(T) = P(T \wedge D) + P(T \wedge \neg D)$$
$$= P(T|D)P(D) + P(T|\neg D)P(\neg D)$$

- $\{D\} \cup \{\neg D\}$ completely account for the whole population, but $\{T\} \cup \{\neg T\}$ does not cover the whole population (because you **did not test everyone!**).

What is the probability of not testing positive? Given the disease? Now the Chris point 2 this particular region, everyone in there would have first narrative



Calculating $P(T)$ given $P(T|D)$ and $P(D)$



$p(\neg T) = p(\neg T, D)$

$+ p(\neg T, \neg D)$

$$\begin{aligned} P(T) &= P(T \wedge D) + P(T \wedge \neg D) \\ &= P(T|D)P(D) + P(T|\neg D)P(\neg D) \end{aligned}$$

- $\{D\} \cup \{\neg D\}$ completely account for the whole population, but $\{T\} \cup \{\neg T\}$ does not cover the whole population (because you **did not test everyone!**).

Sorry not t. and the and you want to add that part which is probably plus probability of not t, and not the

Mouse Select Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On

Calculating $P(T)$ given $P(T|D)$ and $P(D)$

Another way of deriving
$P(T) = P(T|D)P(D) + P(T|\neg D)P(\neg D)$:

From

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

$$P(\neg D|T) = \frac{P(T|\neg D)P(\neg D)}{P(T)}$$

and from $P(D|T) + P(\neg D|T) = 1$,

$$1 = \frac{P(T|D)P(D)}{P(T)} + \frac{P(T|\neg D)P(\neg D)}{P(T)}, \text{ thus}$$

$$P(T) = P(T|D)P(D) + P(T|\neg D)P(\neg D)$$

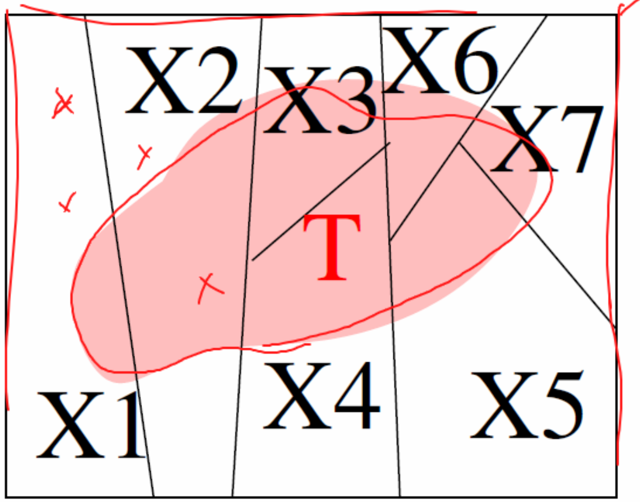
Skip.
(repetitive)

You can quickly find out can quickly follow through this okay so i'm gonna skip

You are screen sharing Stop Share

Calculating $P(T)$: General Case

$\cup$



$X \in \{x_1, x_2, x_3, \ldots x_7\}$

$\sum_{i=1}^{7} P(X = x_i) = 1.0$

More Generally, if $\left[\sum_{x\in\{x_1,x_2,\ldots,x_n\}} P(X = x)\right] = 1$ and events $X = x_m$ and $X = x_n$ are disjoint for all $m \neq n$,

$$P(T) = \sum_{x\in\{x_1,x_2,\ldots,x_n\}} P(T|\underbrace{X = \overset{T/F}{x}})P(X = x)$$

But if you have multiple possible values, then you can just quickly extended

What's The Big Deal?

- $P(T|D)$ may be easier to obtain: you can run the test on a small pool of known patients (say 100) at a hospital.
- $P(D|T)$ is much harder to obtain directly. Since the test makes 1 mistake out of 100 tests, if you run the test on 10,000 people, you'll get 100 false-positives, and one genuine patient who tests positive (consider that $P(T) = 0.010098$). So, just to get about 100 people testing positive, you have to run the tests on 10,000 people.
- $P(D)$ serves as a **prior** in this case. In many cases, the prior represents subjective **belief** of the person calculating the probability in case $P(D)$ is not directly measurable.

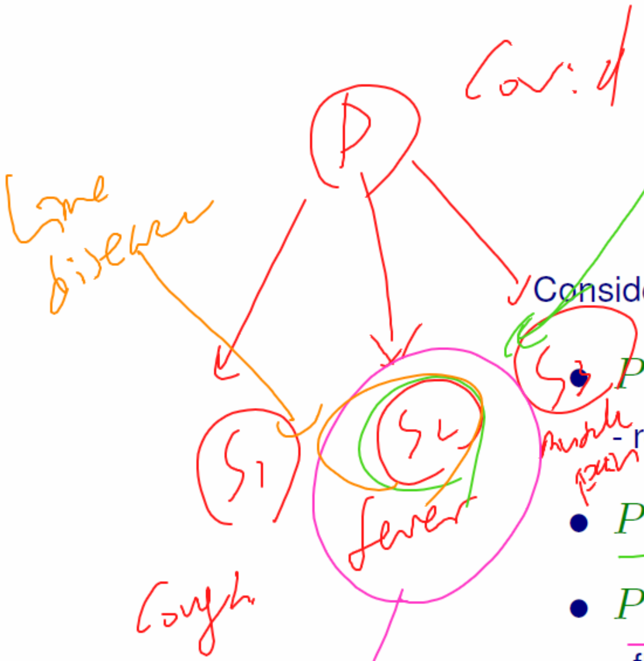
What's The Big Deal?

- $P(T|D)$ may be easier to obtain: you can run the test on a small pool of known patients (say 100) at a hospital.
- $P(D|T)$ is much harder to obtain directly. Since the test makes 1 mistake out of 100 tests, if you run the test on 10,000 people, you'll get 100 false-positives, and one genuine patient who tests positive (consider that $P(T) = 0.010098$). So, just to get about 100 people testing positive, you have to run the tests on 10,000 people.
- $P(D)$ serves as a **prior** in this case. In many cases, the prior represents subjective **belief** of the person calculating the probability in case $P(D)$ is not directly measurable.

Mouse Select Text Draw Stamp Spotlight Eraser Format Undo Redo Clear Save

Who can see what you share here? Recording On

Diagnostic vs. Causal Knowledge



Consider these probabilities:

- $P(Symptom|Disease)$: **causal knowledge**

  - relatively fixed.

- $P(Disease)$: somewhat variable.

- $P(Disease|Symptom)$: **diagnostic knowledge**

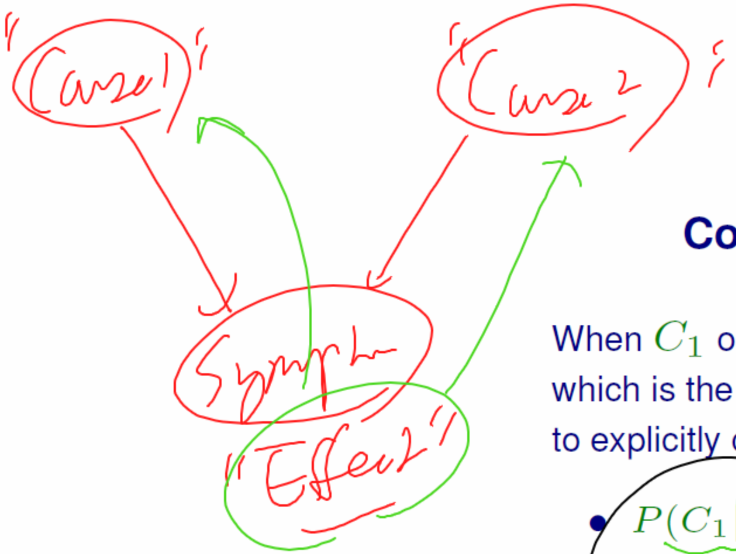
  - fluctuates as $P(Disease)$ change.

$P(Covid \mid fever)$

$P(Disease|Symptom)$ directly measured can be no longer accurate when $P(Disease)$ changes (e.g. an epidemic outburst), however the calculation based on Bayes' rule can be much more robust.

So, because there could be so many different factors. It might be difficult to compute this, or estimate this kind of probability, which is a diagnostic knowledge

You are screening sharing Stop Share



Observed.

Comparison of Conditional Probabilities

When C_1 or C_2 can cause phenomenon (or *effect*) E, to find out the which is the more probable cause of phenomenon E, we do not need to explicitly calculate $P(E)$:

C1 or C2 ?

- $P(C_1|E) = \frac{P(E|C_1)P(C_1)}{\cancel{P(E)}}$ ← Bayes Rule.
- $P(C_2|E) = \frac{P(E|C_2)P(C_2)}{\cancel{P(E)}}$
- From the above, we get:

$$\frac{P(C_1|E)}{P(C_2|E)} = \frac{P(E|C_1)P(C_1)}{P(E|C_2)P(C_2)} = \frac{a}{b}$$

The outcome will be the same right. So one way of doing that is to just find the ratio of this over that. so

Comparison of Conditional Probabilities

When C_1 or C_2 can cause phenomenon (or *effect*) E, to find out the which is the more probable cause of phenomenon E, we do not need to explicitly calculate $P(E)$:

- $P(C_1|E) = \frac{P(E|C_1)P(C_1)}{P(\cancel{E})}$
- $P(C_2|E) = \frac{P(E|C_2)P(C_2)}{P(\cancel{E})}$
- From the above, we get:

$$\frac{P(C_1|E)}{P(C_2|E)} = \frac{P(E|C_1)P(C_1)}{P(E|C_2)P(C_2)} = \frac{a}{b} > 1.$$

C_1 is more probable.

Logical Connectives and Conditional Probability

- Logical connectives can be used:

$$P(A \vee B), P(A \wedge \neg B), P(Cavity \wedge \neg Insured), \text{ etc.}$$

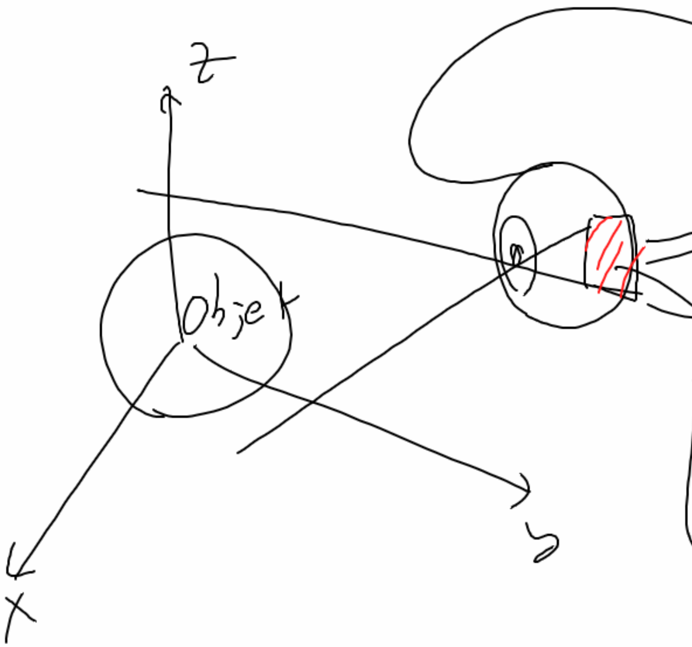
- Conditional Probability $P(A|B)$ (read *probability of* A *given* B):

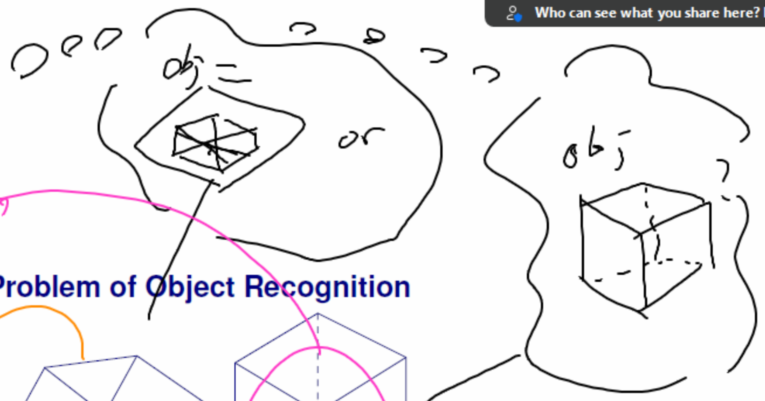
$$P(Cavity|Toothache) = 0.8$$

$P(Cavity) = 0.2$

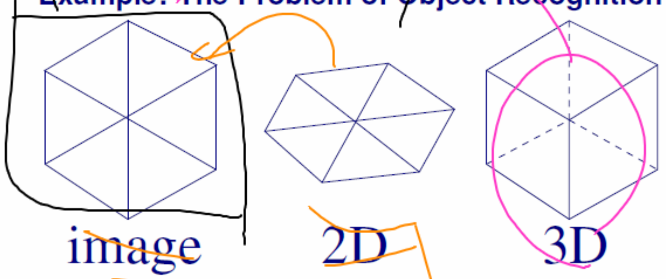
- As new evidence comes in, the conditional probability gets updated:

$$P(Cavity|\underbrace{Toothache \wedge BadBreath}) = 0.5$$





Example: The Problem of Object Recognition



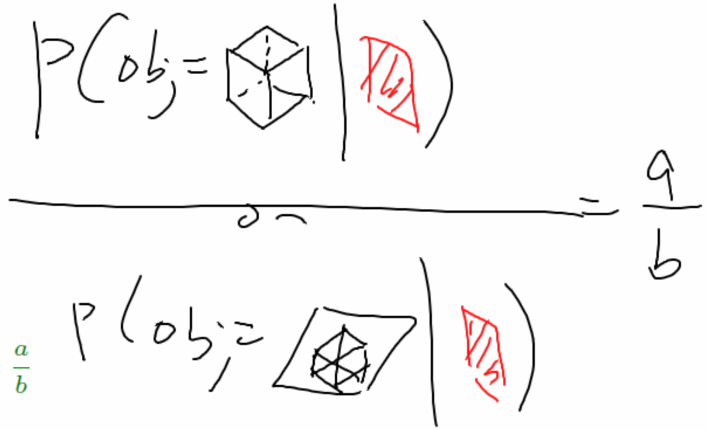
Given an image projected on the retina, what is the more likely cause? the 2D hexagon? or a transparent 3D cube? This is basically a computer vision problem.

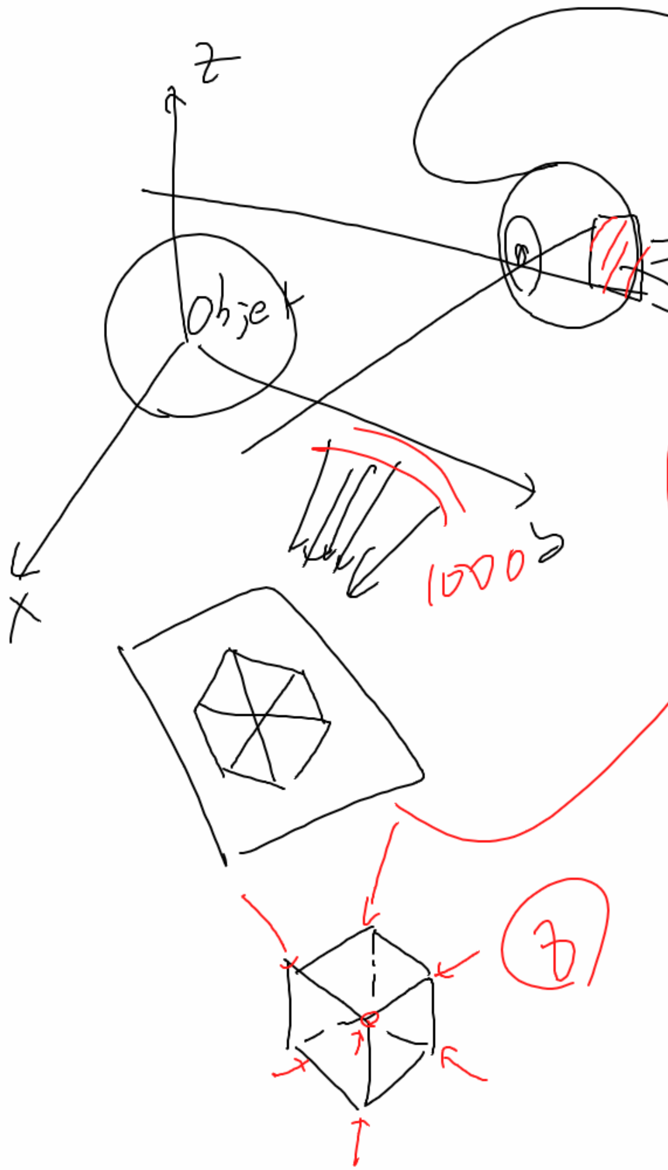
obj = Retina

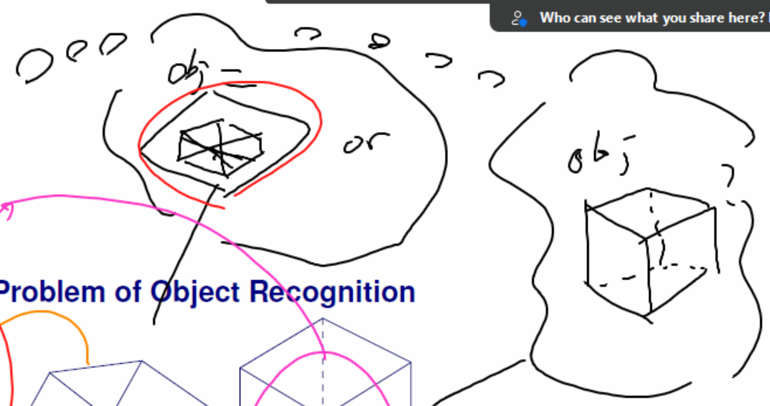
$$\frac{P(Hexagon|Image)}{P(Cube|Image)} = \frac{P(Image|Hexagon)P(Hexagon)}{P(Image|Cube)P(Cube)} = \frac{a}{b}$$

A probabilistic vision agent can make a decision based on such a ratio.

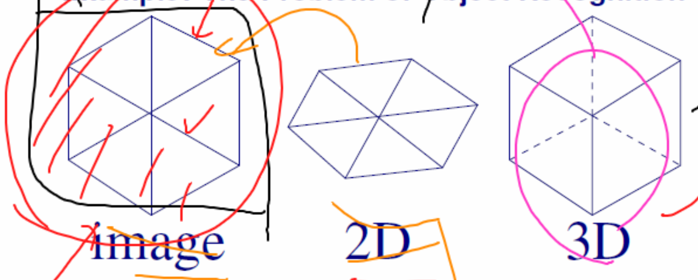
assume uniform prior:







Example: The Problem of Object Recognition



Given an image projected on the retina, what is the more likely cause? the 2D hexagon? or a transparent 3D cube? This is basically a computer vision problem.

obj = Retina =

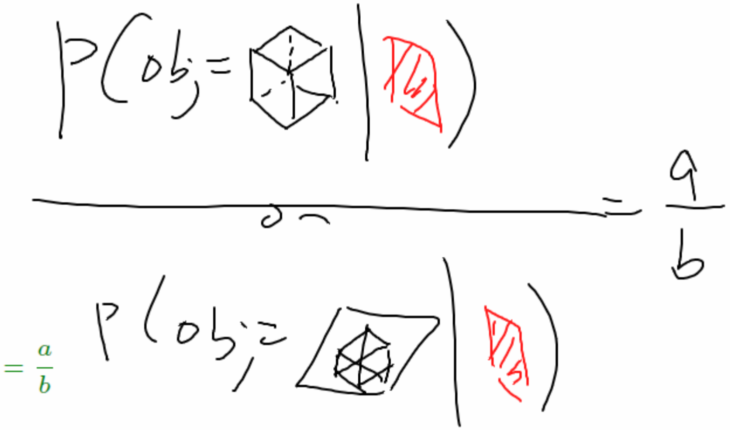
$$\frac{P(Hexagon|Image)}{P(Cube|Image)} = \frac{P(Image|Hexagon)P(Hexagon)}{P(Image|Cube)P(Cube)} = \frac{a}{b}$$

1000/n

8/n

A probabilistic vision agent can make a decision based on such a ratio.

assume uniform prior:



Logical Connectives and Conditional Probability

- Logical connectives can be used:

$$P(A \vee B), P(A \wedge \neg B), P(Cavity \wedge \neg Insured), \text{ etc.}$$

- Conditional Probability $P(A|B)$ (read *probability of A given B*):

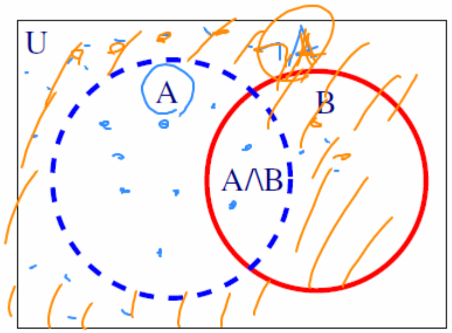
$$P(Cavity|Toothache) = 0.8$$

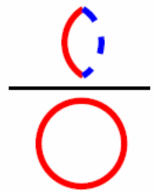
$P(Cavity) = 0.2$

- As new evidence comes in, the conditional probability gets updated:

$$P(Cavity|\underbrace{Toothache \wedge BadBreath}) = 0.5$$

Conditional Probability



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

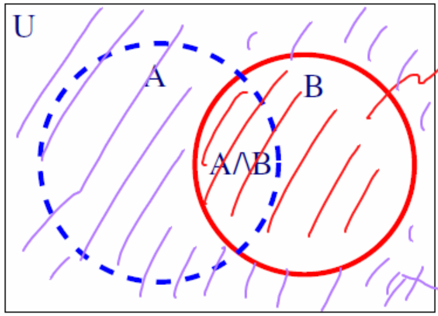
- Think about the **area** occupied by each event.
- The bounding rectangle U has an area of 1, thus

$$P(A) = \frac{\text{Area of } A}{\text{Area of } U} = \frac{\text{Area of } A}{1} = \text{Area of } A$$

- $P(A|B)$ means B now takes on the role of U. Within this limited event space, what is the probability of A.

So let me just probably use a different color. So anything in this region would have not a

Conditional Probability



B = true

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

B = False

- Think about the **area** occupied by each event.
- The bounding rectangle U has an area of 1, thus

$$P(A) = \frac{\text{Area of } A}{\text{Area of } U} = \frac{\text{Area of } A}{1} = \text{Area of } A$$

- $P(A|B)$ means B now takes on the role of U. Within this limited event space, what is the probability of A.

He is false.