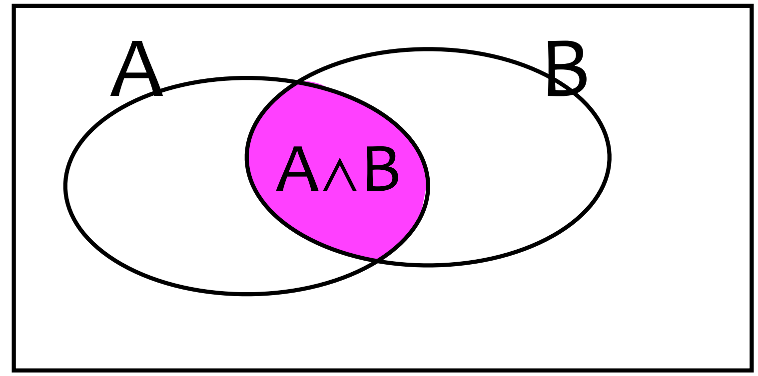
Probability of combined propositions

- $p(A \wedge B)$ = probability both A and B are true.

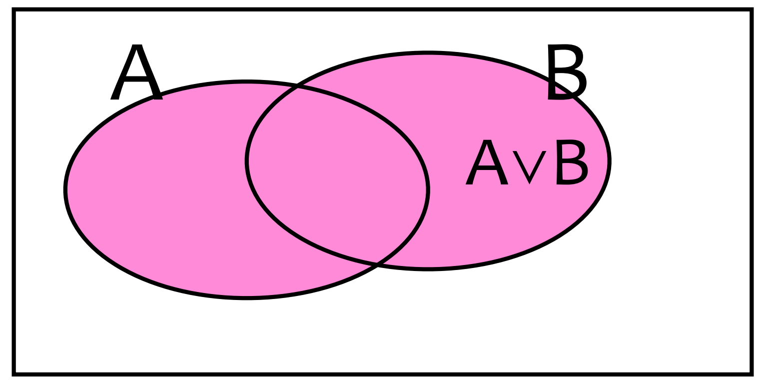


Conjunction of A and B

- $A \wedge B$ is a subset of A, so

$$0 \leq p(A \wedge B) \leq p(A)$$ Conjunction rule

- $A \vee B$ is a superset of A, so



Disjunction of A and B

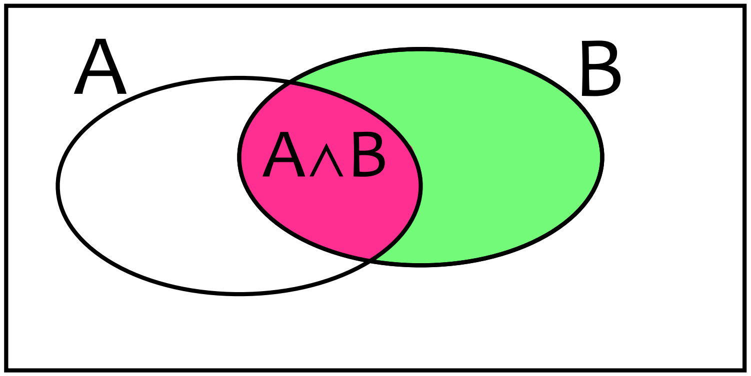
- $p(A) \leq p(A \vee B) \leq 1$ Disjunction rule

- $p(A \vee B) = p(A) + p(B) - p(A \wedge B)$ Relation between $\wedge$ and $\vee$

Conditional probability

- $p(A|B)$ means " probability of A given B is true"

- Definition: $p(A|B) = \dfrac{p(A \wedge B)}{p(B)}$



Probability tables

Gallup 2009

A \ **B**	T	F	
T	10	20	30
F	40	30	70
	50	50	

$p(A|B) = 10/50 = .2$

Party \ **Sex**	M	F	
D	27	31	58
R	23	19	42
	50	50	

$p(D|F) = 31/50 = .62$

$p(F|D) = 31/58 = .53$

Bayes' rule

- Conditional probability is the basis for inductive inference
- The conditional probability of a conclusion (hypothesis) H given premise (data) D is:

"Prior probability"
= belief in H before the data

$$p(H|D) = \frac{p(H)\,p(D|H)}{p(D)}$$

"Likelihood"
= fit of H to the data

"Posterior probability" = belief in H after the data

- This is called Bayes' rule and is incredibly useful for deciding how strongly to believe any inductive hypothesis on the basis of evidence and prior knowledge.
- Bayes' rule says: the posterior is proportional to the product of the prior and the likelihood (fit to the evidence).

Bayesian inference

- Given some data D and various hypotheses that might explain the data, Bayesian inference allows you to compute how strongly to believe H as a function of

  - the degree to which H fits the evidence (likelihood), and

  - the prior probability of H (how likely it was before the evidence).

- Given data D, Bayes' rule lets you calculate the posterior probability of H (the probability of H given the data D).

- Bayesian inference simply uses the mathematical laws of probability to decide what to believe

Bayesian inference: examples

- A random person. Is he/she a democrat?

  prior probability = $p(\mathrm{D})$ = 58% [58% of random people are democrats]

  Oh, she's a woman? Ok,

  $$p(\mathrm{D}|\mathrm{F}) = p(\mathrm{F}|\mathrm{D})p(\mathrm{D})/p(\mathrm{F})$$ *Bayes' rule*

  $$= (.53)(.58)/(.5) = .615$$

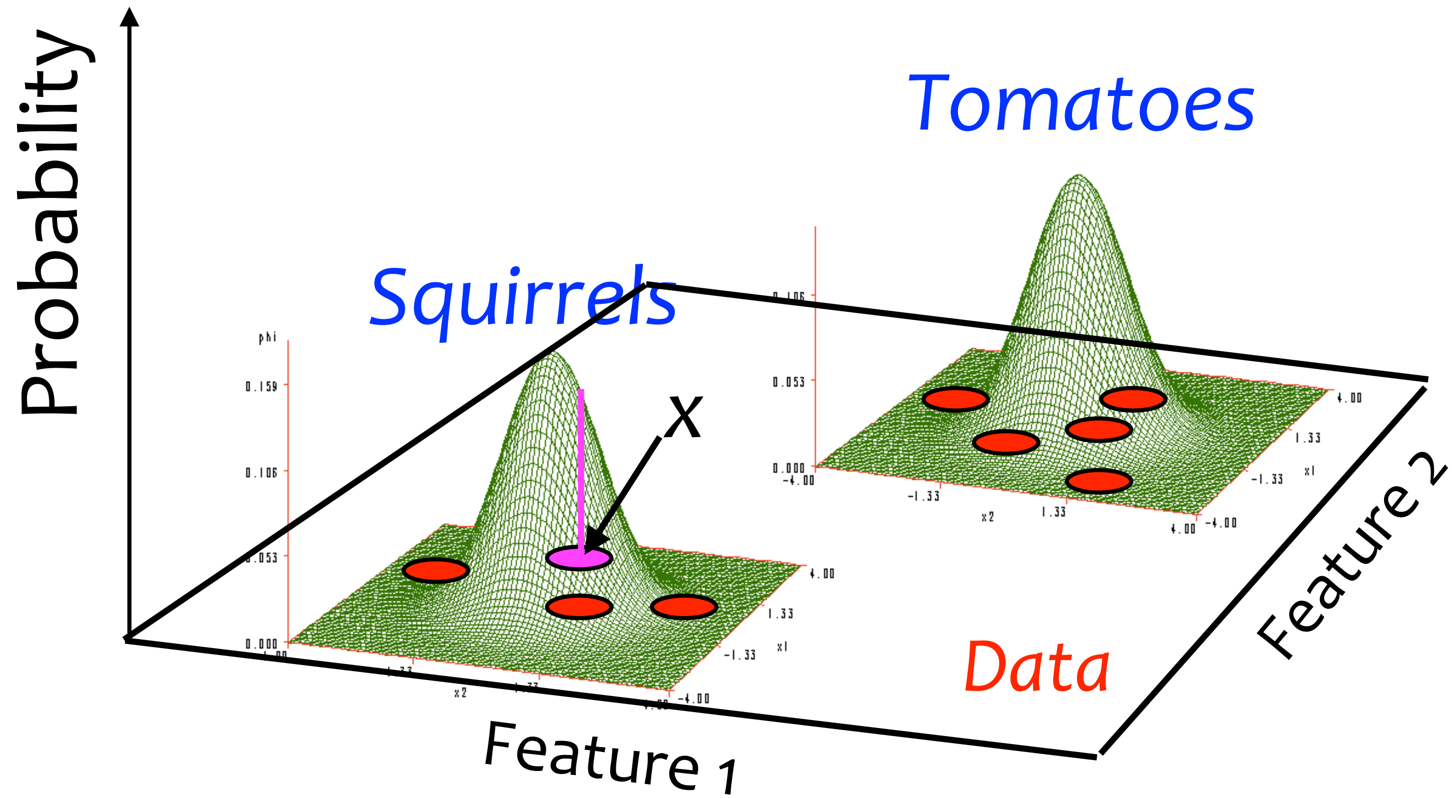
- A guy named Joe. Is he a symphony conductor?

  $p(\mathrm{SC}) = .0001$ [low prior — very people are symphony conductors]

  Oh, he likes music? p(SC|Likes music) =

  $p(\mathrm{LM}|\mathrm{SC})p(\mathrm{SC})/p(\mathrm{LM}) = (1)(.0001)/(1/2) = .0002$ (assuming 50% of people like music!)

Bayesian concept learning



$$p(\text{squirrel}|x) = \frac{p(x|\text{squirrel})p(\text{squirrel})}{p(x|\text{squirrel})p(\text{squirrel}) + p(x|\text{tomato})p(\text{tomato})}$$

prior of squirrel

prior of tomato

Conditional probability in reasoning

Deduction	Probabilistic reasoning
• Modus ponens	• "probabilistic modus ponens"
$A, A \rightarrow B;$	$A, p(B\|A)$ high;
B	$p(B)$ high
• Modus tollens	• "probabilistic modus tollens"
$A \rightarrow B, \sim B;$	$p(B\|A)$ low, B;
$\sim A$	$p(A)$ low

Bayesian inference and rationality

- Bayesian inference is the rational method for drawing inferences from experience—rational induction
- Bayesian inference is considered normative,

  - i.e. "objectively correct." This is really important!
- If people are "Bayesian", that means they form beliefs in a way that is optimal given the information available to them
- If not, people are irrational, which means that they form in a way that is incoherent or internally inconsistent.
- So: are people Bayesian?