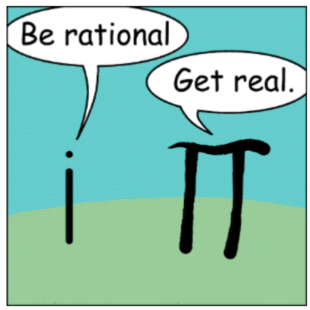
CSE 473

Lecture 20
(Chapters 13 & 14)

Probabilistic Inference



http://jokesprank.com/blog/wp-content/uploads/2010/12/math-cartoon.gif

Today's Outline

- Probabilistic Inference
- Conditional Independence
- Bayesian Networks

Recall: Prior Probability

Prior or unconditional probabilities of propositions
e.g., $P(Cavity = true) = 0.2$ and $P(Weather = sunny) = 0.72$
correspond to belief prior to arrival of any (new) evidence

Recall: Joint Probability

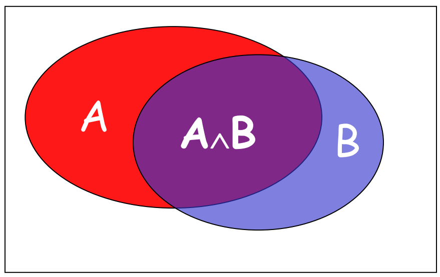
Joint probability distribution for a set of r.v.s gives the probability of every atomic event on those r.v.s
$\mathbf{P}(Weather, Cavity)$ = a 4×2 matrix of values:

$Weather =$	$sunny$	$rain$	$cloudy$	$snow$
$Cavity = true$	0.144	0.02	0.016	0.02
$Cavity = false$	0.576	0.08	0.064	0.08

We will see later how any question can be answered by the joint distribution

Conditional Probability

- P(*A* | *B*) is the probability of *A* given *B*
- Assumes that *B* is the only info known.
- Defined as: $P(A|B) = \frac{P(A,B)}{P(B)} = \frac{P(A \wedge B)}{P(B)}$



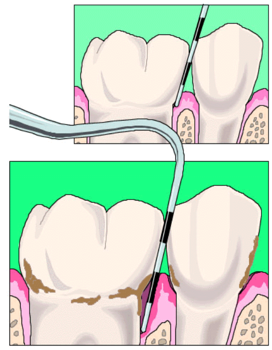
Conditional Probability Examples

- P(*Cavity* = *true* | *Toothache* = *true*) = probability of *cavity* given *toothache*
- Notation for conditional distribution:

  P(*Cavity* | *Toothache*) = 2-element vector of 2-element vectors (2 Pr values given Toothache is true and 2 Pr values given Toothache is false)
- If we know more, e.g., *Cavity* = *true*, then we have

  P(*cavity* | *toothache, cavity*) = 1
- New evidence may be irrelevant, allowing simplification:
- P(*cavity* | *toothache, sunny*) = P(*cavity* | *toothache*) = 0.8

Dilemma at the Dentist's



What is the probability of a cavity given a toothache?
What is the probability of a cavity given the probe catches?

Probabilistic Inference by Enumeration

Start with the joint distribution:

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

For any proposition ϕ, sum the atomic events where it is true:
$P(\phi) = \Sigma_{\omega:\omega\models\phi} P(\omega)$

P(toothache)= .108+.012+.016+.064
= .20 or 20%

Inference by Enumeration

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ cavity	**.016**	**.064**	**.144**	**.576**

P(toothache∨cavity) = ?

.20 + .108 + .012 + .072 + .008 - (.108+.012)

= .28

Inference by Enumeration

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ cavity	**.016**	**.064**	**.144**	**.576**

Can also compute conditional probabilities:

$$P(\neg cavity | toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)}$$

$$= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

Problems with Enumeration

- Worst case time: $O(d^n)$

  where d = max arity of random variables

  e.g., d = 2 for Boolean (T/F)

  and n = number of random variables
- Space complexity also $O(d^n)$
  - Size of joint distribution
- Problem: Hard/impossible to estimate all $O(d^n)$ entries of joint for large problems

Do we need to compute all $O(d^n)$ possible entries of joint distribution?

Independence

- Variables *A* and *B* are *independent* iff:

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

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

These two constraints are logically equivalent

Therefore, if *A* and *B* are independent:

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

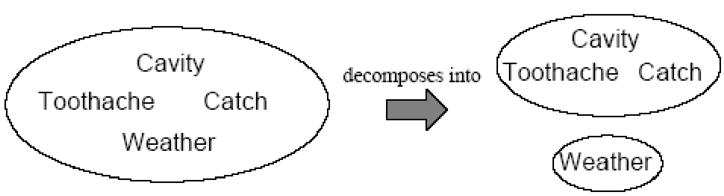
i.e., $P(A \wedge B) = P(A)P(B)$

13

Why is independence useful?

14

Independence



2*2*2*4=32 values

$$\mathbf{P}(Toothache, Catch, Cavity, Weather)$$
$$= \mathbf{P}(Toothache, Catch, Cavity)\mathbf{P}(Weather)$$

Only 2*2*2 + 4 = 12 values needed

32 entries reduced to 12; for n independent biased coins, $2^n \rightarrow n$

Complete independence is powerful but rare.
What to do if it doesn't hold?

Conditional Independence

Joint distribution:

$\mathbf{P}(Toothache, Cavity, Catch)$ has $2^3 - 1 = 7$ independent entries

If I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache:

(1) $P(catch|toothache, cavity) = P(catch|cavity)$

The same independence holds if I haven't got a cavity:

(2) $P(catch|toothache, \neg cavity) = P(catch|\neg cavity)$

$Catch$ is *conditionally independent* of $Toothache$ given $Cavity$:

$\mathbf{P}(Catch|Toothache, Cavity) = \mathbf{P}(Catch|Cavity)$

Instead of 7 entries in the joint distribution, only need 5 (why?)

Conditional Independence II

Given:

$\mathbf{P}(Catch \mid Toothache, Cavity) = \mathbf{P}(Catch \mid Cavity)$

Joint probability distribution:

$\mathbf{P}(Catch, Toothache, Cavity)$

$= \mathbf{P}(Catch \mid Toothache, Cavity)\mathbf{P}(Toothache, Cavity)$

$= \mathbf{P}(Catch \mid Toothache, Cavity)\mathbf{P}(Toothache \mid Cavity)\mathbf{P}(Cavity)$

$= \mathbf{P}(Catch \mid Cavity)\mathbf{P}(Toothache \mid Cavity)\mathbf{P}(Cavity)$

2 + 2 + 1

= 5 independent numbers

Power of Cond. Independence

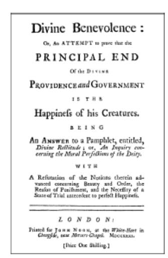
- Often, conditional independence can reduce the storage complexity of the joint distribution from exponential to linear!!
- Conditional independence is the most basic & robust form of knowledge in uncertain environments.

19

Thomas Bayes

Reverand Thomas Bayes
Nonconformist minister
(1702-1761)

- Publications:
- *Divine Benevolence, or an Attempt to Prove That the Principal End of the Divine Providence and Government is the Happiness of His Creatures* (1731)
- *An Introduction to the Doctrine of Fluxions* (1736)
- *An Essay Towards Solving a Problem in the Doctrine of Chances* (1764)

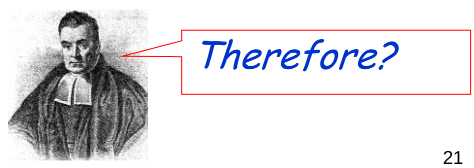

Divine Benevolence:
Or, An ATTEMPT to prove that the
PRINCIPAL END
Of the DIVINE
PROVIDENCE and GOVERNMENT
IS THE
Happiness of his Creatures.
BEING
An ANSWER to a Pamphlet, entitled, *Divine Rectitude; or, An Inquiry concerning the Moral Perfections of the Deity.*
WITH
A Refutation of the Notions therein advanced concerning Beauty and Order, the Reason of Punishment, and the Necessity of a State of Trial antecedent to perfect Happiness.
LONDON:
Printed for JOHN NOON, at the *White-Hart* in *Cheapside*, near *Mercers-Chapel*. MDCCXXXI.
[Price One Shilling.]

Recall: Conditional Probability

- P(*x* | *y*) is the probability of *x* given *y*
- Assumes that *y* is the only info known.
- Defined as:

$$P(x \mid y) = \frac{P(x, y)}{P(y)}$$

$$P(y \mid x) = \frac{P(y, x)}{P(x)} = \frac{P(x, y)}{P(x)}$$



21

Bayes' Rule

$$P(x, y) = P(x \mid y)P(y) = P(y \mid x)P(x)$$

i.e.

$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{P(y)}$$

What is this useful for?

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

22

Bayes' rule is used to Compute *Diagnostic* Probability from *Causal* Probability

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

E.g. let M be meningitis, S be stiff neck

P(M) = 0.0001,
P(S) = 0.1,
P(S|M)= 0.8 (note: these can be estimated from patients)

$$P(M \mid S) = \frac{P(S \mid M)P(M)}{P(S)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008$$

Note: posterior probability of meningitis still very small!

(But chance of M did increase from 0.0001 to 0.0008)

Next Time

- Bayesian Networks
- To Do
  - Project 3
  - Read Chapter 14