Inference from joint distributions

Typically, we are interested in
the posterior joint distribution of the **query variables** $\mathbf{Y}$
given specific values $\mathbf{e}$ for the **evidence variables** $\mathbf{E}$

Let the **hidden variables** be $\mathbf{H} = \mathbf{X} - \mathbf{Y} - \mathbf{E}$

Then the required summation of joint entries is done by summing out the hidden variables:

$$\mathbf{P}(\mathbf{Y}|\mathbf{E}=\mathbf{e}) = \alpha \mathbf{P}(\mathbf{Y},\mathbf{E}=\mathbf{e}) = \alpha \sum_{\mathbf{h}} \mathbf{P}(\mathbf{Y},\mathbf{E}=\mathbf{e},\mathbf{H}=\mathbf{h})$$

The terms in the summation are joint entries because $\mathbf{Y}$, $\mathbf{E}$, and $\mathbf{H}$ together exhaust the set of random variables

Obvious problems:
1) Worst-case time complexity $O(d^n)$ where d is the largest arity
2) Space complexity $O(d^n)$ to store the joint distribution
3) How to find the numbers for $O(d^n)$ entries???

Independence

Two random variables A B are (absolutely) **independent** iff
$P(A|B) = P(A)$
or $P(A,B) = P(A|B)P(B) = P(A)P(B)$
e.g., A and B are two coin tosses

If n Boolean variables are independent, the full joint is
$\mathbf{P}(X_1, \ldots, X_n) = \prod_i \mathbf{P}(X_i)$
hence can be specified by just n numbers

Absolute independence is a very strong requirement, seldom met

Conditional independence

Consider the dentist problem with three random variables:
$Toothache$, $Cavity$, $Catch$ (steel probe catches in my tooth)

The full joint distribution 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) $P(Catch|Toothache, Cavity) = P(Catch|Cavity)$
i.e., $Catch$ is **conditionally independent** of $Toothache$ given $Cavity$

The same independence holds if I haven't got a cavity:
(2) $P(Catch|Toothache, \neg Cavity) = P(Catch|\neg Cavity)$

Conditional independence contd.

Equivalent statements to (1)

(1a) $P(Toothache|Catch, Cavity) = P(Toothache|Cavity)$ **Why**??

(1b) $P(Toothache, Catch|Cavity) = P(Toothache|Cavity)P(Catch|Cavity)$ **Why**

Full joint distribution can now be written as
$\mathbf{P}(Toothache, Catch, Cavity) = \mathbf{P}(Toothache, Catch|Cavity)\mathbf{P}(Cavity)$
$= \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)$
i.e., 2 + 2 + 1 = 5 independent numbers