Bayes Nets:

Directed, No Cycles

PDF = \prod P(\text{node} | \text{parents})

nodes
D-separation on arbitrary graph

Two nodes are independent if there is no valid path between them.
(all paths are d-separated)

Two nodes are dependent if any path between them is not d-separated

Naive Bayes:

\[ P(A, B, C) = P(A) \cdot P(B|A) \cdot P(C|B) \]

\[ P(C) = P(C | A, B, C) + P(C | \neg A, B, C) + P(C | \neg A, \neg B, C) + P(A, \neg B, C) \]
Conditional Distributions

\[
P(A|C) = \frac{P(A, C)}{P(C|A)P(A) + P(C|-A)P(-A)}
\]

\[
= \frac{P(A, C)}{P(A, C) + P(-A, C)}
\]

A slightly better way ...

A variable elimination

\[
P(B, J, M) = \?
\]

\[
= \sum_a \sum_e P(B)P(e)P(a|B,e)P(J|a)P(M|a)
\]

\[
= P(B) \sum_e \left[ P(e) \sum_a P(a|B,e)P(J|a)P(M|a) \right]
\]
Rewrite probability statements as functions...

\[ P_A(a, b, e) = P(a | B e) \]
\[ f_j(a) = P(j | a) \]
\[ f_M(a) = P(M | a) \]

\[ P(B, J, M) = P(B) \sum_{e} P(e) \sum_{a} f_j(a) f_M(a) f_A(a, B, e) \]

\[ f_{A, J, M}(B, e) = \sum_{a} f_j(a) f_M(a) f_A(a, B, e) \]

\[ = P(B) \sum_{e} P(e) f_A(B, e) \]

\[ f_{E, A, J, M}(B) \sum_{e} P(e) f_{A, J, M}(B, e) \]

\[ P(B, J, M) = P(B) \sum_{e} P(e) f_{E, A, J, M}(B) \]

BAM! Variables eliminated!
Expand this back out...

\[
(f_j(a) + \ldots) + (f_j(a) + \ldots)
\]

\[
(f_{\text{ABCDE}}(j) + \ldots) + (f_{\text{ABCDE}}(j)
\]

\(\approx 2^5\) calculations

\(\approx 2^5\) calculations that I don't have to do

\(\smile\)

Except... I still had to do those \(2^5\) calculations in the first place

\(\nabla\)
what to do when variable elimination fails?

Sample!

If we have all conditional probability parameters...

Given some values from the graph, we can sample the rest

We should always have "free variables," nodes w/o parents = easy to sample

\[ P(A) \begin{cases} \alpha = 1 & 0.3 \\ \alpha = 0 & 0.7 \end{cases} \]
Naive sampling
- Start w/ free nodes
- Sample rest

\[
P(C=1) \approx \frac{\#C=1}{\text{total samples}} = \frac{\sum I(c=1)}{\text{total samples}}
\]

\[
P(A=1) = 0.000001
\]

Naive sampling will always pick \( A=0 \)

\[\longrightarrow\] Weighted sampling

- Pick some arbitrary observations
- Calculate weight of assignment

For example, \( w = P(A=0) P(B=1) \)
- Sample the rest
  Instead of counting, increment
  by the weight

\[
P(c) = \frac{\sum w_i I(c=1)}{\sum w_i}
\]
Set $A = 1$, $B = 0$

$$W = P(A = 1) P(B = 0)$$

- sample $C$ a bunch

$$P(C = 1 | A = 1, B = 0) = \frac{\sum w_i I(C = 1)}{\sum w_i}$$

MCMC = Markov Chain Monte Carlo

- random walk
- iterate through nodes, resampling based on other values
\[ P(A) \quad ? \quad P(B|A) \quad ? \quad P(C|B) \quad ? \quad P(C|B) \]

\[ \quad A \rightarrow B \rightarrow C \]

Wanted \( P(C)? \) or \( P(C|A)? \)
- Need to estimate \( P(B|A) \)
- Get some observations
- E.M.
- Profit

→ More on this with HMMs later