CSE 150. Assignment 1

Out: *Wed Apr 9*
Due: *Wed Apr 16* (in class)
Reading: Russell & Norvig, Chapter 13; Korb & Nicholson, Chapter 1.

1.1 Kullback-Leibler distance

Consider two discrete probability distributions, p_i and q_i, with $\sum p_i = \sum q_i = 1$. The Kullback-Leibler (KL) distance between these distributions is defined as:

$$\mathrm{KL}(p,q) = \sum_i p_i \log(p_i/q_i).$$

(a) By sketching graphs of $\log x$ and $x-1$, verify the inequality

$$\log x \leq x - 1,$$

with equality if and only if $x=1$. Confirm this result by differentiation of $\log x - (x-1)$. (Note: all logarithms in this problem are *natural* logarithms.)

(b) Use the previous result to prove that $\mathrm{KL}(p,q) \geq 0$, with equality if and only if the two distributions p_i and q_i are equal.

(c) Provide a counterexample to show that the KL distance is not a symmetric function of its arguments:

$$\mathrm{KL}(p,q) \neq \mathrm{KL}(q,p).$$

Despite this asymmetry, it is still common to refer to $\mathrm{KL}(p,q)$ as a measure of distance. Many algorithms in machine learning are based on minimizing KL distances between probability distributions.

1.2 Conditional independence [RN 13.10]

Show that the following three statements about random variables X, Y, and Z are equivalent:

$$\begin{aligned} P(X,Y|Z) &= P(X|Z)P(Y|Z) \\ P(X|Y,Z) &= P(X|Z) \\ P(Y|X,Z) &= P(Y|Z) \end{aligned}$$

You should become fluent with all these ways of expressing that X is conditionally independent of Y given Z.

1.3 Creative writing

Attach events to the binary random variables X, Y, and Z that are consistent with the following patterns of commonsense reasoning. You may use different events for the different parts of the problem.

(a) Accumulating evidence:

$$P(X\!=\!1) < P(X\!=\!1|Y\!=\!1) < P(X\!=\!1|Y\!=\!1, Z\!=\!1)$$

(b) Explaining away:

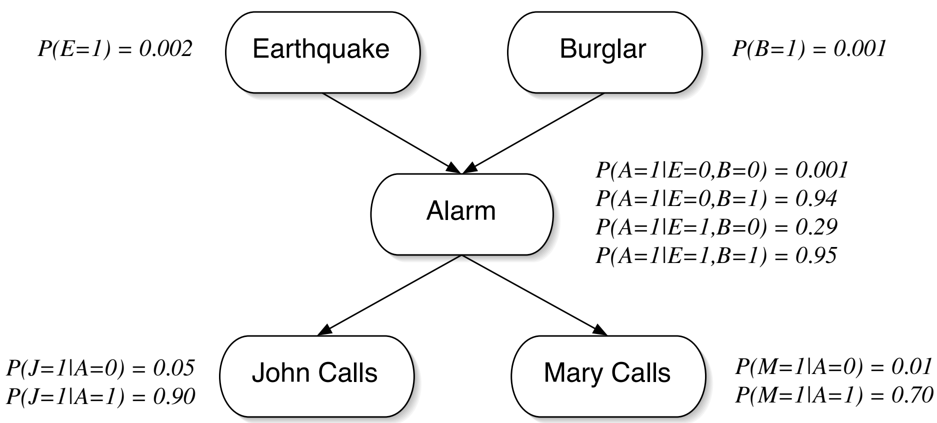
$$\begin{aligned} P(X\!=\!1|Y\!=\!1) &> P(X\!=\!1), \\ P(X\!=\!1|Y\!=\!1, Z\!=\!1) &< P(X\!=\!1|Y\!=\!1) \end{aligned}$$

(c) Conditional independence:

$$\begin{aligned} P(X\!=\!1, Y\!=\!1) &\neq P(X\!=\!1)P(Y\!=\!1) \\ P(X\!=\!1, Y\!=\!1|Z\!=\!1) &= P(X\!=\!1|Z\!=\!1)P(Y\!=\!1|Z\!=\!1) \end{aligned}$$

1.4 Probabilistic inference

Recall the probabilistic model that we described in class for the binary random variables $\{E = \text{Earthquake}, B = \text{Burglary}, A = \text{Alarm}, J = \text{JohnCalls}, M = \text{MaryCalls}\}$. We also expressed this model as a belief network, with the directed acyclic graph (DAG) and conditional probability tables (CPTs) shown below:



Compute numeric values for the following probabilities, exploiting relations of conditional independence as much as possible to simplify your calculations. Show your work.

(a) $P(E\!=\!1, B\!=\!0|A\!=\!1)$
(b) $P(A\!=\!1|J\!=\!0, M\!=\!0)$
(c) $P(A\!=\!1|E\!=\!1, B\!=\!0, J\!=\!0, M\!=\!0)$
(d) $P(J\!=\!1, M\!=\!1|B\!=\!1)$
(e) $P(B\!=\!1|J\!=\!1, M\!=\!1)$
(f) $P(B\!=\!1|J\!=\!1, M\!=\!1, E\!=\!1)$