CSE 150. Assignment 1 *Spring 2018*

Out: *Tue Apr 10*
Due: *Tue Apr 17 (in class)*
Supplementary reading: RN, Ch 13; KN, Ch 1.

1.1 Probabilistic reasoning

Suppose that each day when driving to campus, there is a 3% chance that I have car trouble (e.g., a flat tire), and that when I have car trouble, there is a 97% chance that I will be late. On the other hand, even when I do not have car trouble, there remains a 2% chance that I will be late for other reasons (e.g., traffic).

Suppose that one day I am late to campus. What is the probability that I had car trouble? Use Bayes rule to compute your answer, and *show your work*.

1.2 Conditioning on background evidence

It is often useful to consider the impact of specific events in the context of general background evidence, rather than in the absence of information. Denoting such evidence by E, prove the following variants of Bayes rule and marginalization:

$$\text{(a)} \quad P(X|Y,E) = \frac{P(Y|X,E)P(X|E)}{P(Y|E)}$$

$$\text{(b)} \quad P(X|E) = \sum_y P(X,Y\!=\!y|E)$$

1.3 Conditional independence

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

$$\begin{aligned} \text{(i)} \quad P(X,Y|E) &= P(X|E)P(Y|E) \\ \text{(ii)} \quad P(X|Y,E) &= P(X|E) \\ \text{(iii)} \quad P(Y|X,E) &= P(Y|E) \end{aligned}$$

In other words, show that each one of these statements implies the other two. You should become fluent with all these ways of expressing that X is conditionally independent of Y given E.

1.4 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) Explaining away:

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

(b) Accumulating evidence:

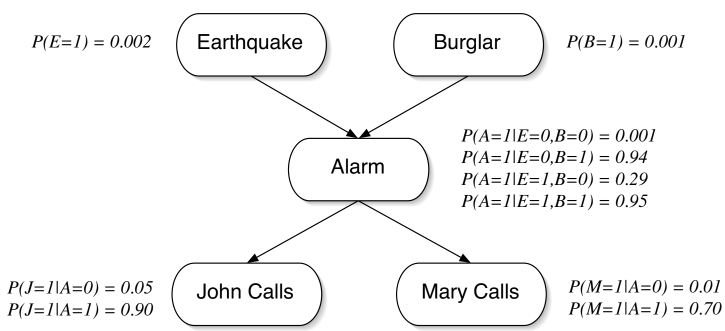
$$P(Z\!=\!1) < P(Z\!=\!1|Y\!=\!1) < P(Z\!=\!1|X\!=\!1,Y\!=\!1)$$

(c) Conditional independence:

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

1.5 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 marginal and conditional independence as much as possible to simplify your calculations. You may re-use numerical results from lecture, but otherwise show your work. Be careful not to drop significant digits in your answer.

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

Consider your results in (b) versus (a), (d) versus (c), and (f) versus (e). Do they seem consistent with commonsense patterns of reasoning?

1.6 Kullback-Leibler distance

Often it is useful to measure the difference between two probability distributions over the same random variable. For example, as shorthand let

$$\begin{aligned} p_i &= P(X\!=\!x_i|E), \\ q_i &= P(X\!=\!x_i|E') \end{aligned}$$

denote the conditional distributions over the random variable X for different pieces of evidence $E \neq E'$. Note that both distributions satisfy $\sum_i p_i = \sum_i q_i = 1$. 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 $\log z$ and $z-1$, verify the inequality

$$\log z \leq z - 1,$$

with equality if and only if $z\!=\!1$. Confirm this result by differentiation of $\log z - (z\!-\!1)$. (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. *Hint: substitute* $z = q_i/p_i$ *into the previous inequality.*

(c) Provide a simple numerical 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 between probability distributions.

1.7 Mutual information

Closely related to the Kullback-Leibler distance, the mutual information $I(X,Y)$ between two discrete random variables X and Y is defined as

$$I(X,Y) = \sum_{x,y} P(x,y) \log\left[\frac{P(x,y)}{P(x)P(y)}\right],$$

where the sum is over all possible values of the random variables X and Y. As you will see here, the mutual information provides a *quantitative* measure of conditional dependence.

(a) Prove that the mutual information $I(X,Y)$ is nonnegative. (*Hint:* use the result from the previous problem.)

(b) State a sufficient condition for the mutual information $I(X,Y)$ to vanish.