Medical Diagnosis Example

Suppose that a doctor knows that the disease meningitis causes the patient to have a stiff neck 50% of the time, the prior probability of a patient having meningitis is 1/50000, and the prior probability of any patient having a stiff nect is 1/20.

- Given:

$$P(\mathit{StiffNeck}|\mathit{Meningitis}) = 0.5$$
$$P(\mathit{Meningitis}) = 0.00002$$
$$P(\mathit{StiffNeck}) = 0.05$$

  Find:

$$P(\mathit{Meningitis}|\mathit{StiffNeck})$$

- Applying Bayes' rule:

$$\begin{aligned} & P(\mathit{Meningitis}|\mathit{StiffNeck}) \\ = & \frac{P(\mathit{StiffNeck}|\mathit{Meningitis})P(\mathit{Meningitis})}{P(\mathit{StiffNeck})} \\ = & \frac{0.5 \times 0.00002}{0.05} = 0.0002. \end{aligned}$$

- That is, we expect only one in 5000 patients with a stiff nect to meningitis.

Medical Diagnosis Example (CONT'D)

Suppose the doctor is also concerned with the possibility that the patient is suffering from whiplash given a stiff neck.

- Assume:

$$P(\mathit{StiffNeck}|\mathit{Whiplash}) = 0.8$$
$$P(\mathit{Whiplash}) = 0.001$$

- Then

$$\begin{aligned} & \frac{P(\mathit{Meningitis}|\mathit{StiffNeck})}{P(\mathit{Whiplash}|\mathit{StiffNeck})} \\ = & \frac{P(\mathit{StiffNeck}|\mathit{Meningitis})P(\mathit{Meningitis})}{P(\mathit{StiffNeck}|\mathit{Whiplash})P(\mathit{Whiplash})} \\ = & \frac{0.5 \times 0.00002}{0.8 \times 0.001} = \frac{1}{80}. \end{aligned}$$

- That is, whiplash is 80 times more likely than meningitis, given a stiff neck.

Normalization

In some cases, relative likelihood is sufficent for decision making, but when different actions yield radically different utilities, one needs exact values in order to make rational decisions.

- Instead of directly assessing the prior probability of the "symptoms", one can consider an exhaustive set of cases.
- For example, instead of determining $P(S)$ directly, use

$$\begin{aligned} P(M|S) &= \frac{P(S|M)P(M)}{P(S)} \\ P(\neg M|S) &= \frac{P(S|\neg M)P(\neg M)}{P(S)} \end{aligned}$$

- Using the fact that $P(M|S) + P(\neg M|S) = 1$, we obtain

$$P(S) = P(S|M)P(M) + P(S|\neg M)P(\neg M)$$

- That is,

$$P(M|S) = \frac{P(S|M)P(M)}{P(S|M)P(M) + P(S|\neg M)P(\neg M)}$$

- In general,

$$P(Y|X) = \alpha P(X|Y)P(Y)$$

  where α is a normalization constant needed to make the entries in the table $P(X|Y)$ sum to 1.

Normalization: Example

After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease, and that the test is 99% accurate. The good news is that this is a rare disease, striking only one in 10,000 people.

- Why is it good news that the disease is rare?
- Given:

$$P(Test|Disease) = 0.99$$
$$P(\neg Test|\neg Disease) = 0.99$$
$$P(Disease) = 0.0001$$
$$Test = True$$

- Determine $P(Disease|Test)$:

$$\begin{aligned} & P(Disease|Test) \\ &= \frac{P(Test|Disease)P(Disease)}{P(Test|Disease)P(Disease) + P(Test|\neg Disease)P(\neg Disease)} \\ &= \frac{0.99 \times 0.0001}{0.99 \times 0.0001 + 0.01 \times 0.9999} \\ &= 0.009804 \end{aligned}$$

- When the disease is much rarer that the test accuracy, a positive test result does not necessarily mean the disease is likely. A false positive reading remains much more likely.

Combining Evidence

- Find:

$$P(A|E_1, E_2)$$

- One approach: Absorbing all evidence at once.

$$P(A|E_1, E_2) = \frac{P(E_1, E_2|A)P(A)}{P(E_1, E_2)}$$

- Another approach: Absorbing evidence one piece at a time.
  - When only E_1 is available:

$$P(A|E_1) = P(A)\,\frac{P(E_1|A)}{P(E_1)}$$

  - When E_2 also becomes available:

$$\begin{aligned} P(A|E_1, E_2) &= P(A|E_1)\,\frac{P(E_2|E_1, A)}{P(E_2|E_1)} \\ &= P(A)\,\frac{P(E_1|A)}{P(E_1)}\,\frac{P(E_2|E_1, A)}{P(E_2|E_1)} \qquad (1) \end{aligned}$$

  - This process is order independent.

Combining Evidence: Example

- Suppose we have two conditional probabilities relating to cavities:

$$P(Cavity|Toothache) = 0.8$$

$$P(Cavity|Catch) = 0.95$$

- By Bayes' rule:

$$\begin{aligned} &P(Cavity|Toothache \wedge Catch) \\ &= \frac{P(Toothache \wedge Catch|Cavity)P(Cavity)}{P(Toothache \wedge Catch)} \end{aligned}$$

- For this to work, we need to know

$$P(Toothache \wedge Catch|Cavity)$$

- The process of Bayesian updating incorporates evidence one piece at a time.
- Beginning with $Toothache$, we have

$$P(Cavity|Toothache) = P(Cavity)\frac{P(Toothache|Cavity)}{P(Toothache)}$$

- When $Catch$ is observed, applying Bayes' rule:

$$P(Cavity|Toothache \wedge Catch)$$

$$= P(Cavity|Toothache)\frac{P(Catch|Toothache \wedge Cavity)}{P(Catch|Toothache)}$$

$$= P(Cavity)\frac{P(Toothache|Cavity)}{P(Toothache)}\frac{P(Catch|Toothache \wedge Cavity)}{P(Catch|Toothache)}$$

- Still need to find:

$$P(Catch|Toothache \wedge Cavity)$$

Conditional Independence

- Estimating $P(E_1, E_2|A)$ or $P(E_2|E_1, A)$ may be computationally unattractive, especially when many pieces of evidence have to be combined.
- Conditional independence can help.
- Independence:
  - Two random variables X and Y are *independent* if
    - $P(X|Y) = P(X)$, or equivalently
    - $P(Y|X) = P(Y)$.
  - Knowledge about one variable contains no information about the other.
- Conditional independence:
  - Two random variables X and Y are *conditionally independent* given a third variable Z if
    - $P(X|Y, Z) = P(X|Z)$, or equivalently
    - $P(Y|X, Z) = P(Y|Z)$.
  - Given background knowledge Z, knowledge about one variable contains no information about the other.

Conditional Independence: Example

- In the cavity example, cavity is the direct cause of both toothache and the probe catching in the tooth. Once the patient is known to have a cavity, we don't expect the probability of the probe catching to depend on the presence of a toothache; similarly, the probe catching is not going to change the probability that the cavity is causing a toothache.
- In probabilitic terms:

$$P(Catch|Cavity \wedge Toothache) = P(Catch|Cavity)$$

$$P(Toothache|Cavity \wedge Catch) = P(Toothache|Cavity)$$

- These equations express the conditional independence of $Toothache$ and $Catch$ given $Cavity$.
- Given conditional independence, we have:

$$P(Cavity|Toothache \wedge Catch)$$

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

$$= P(Cavity)\frac{P(Toothache|Cavity)P(Catch|Cavity)}{P(Catch \wedge Toothache)}$$

Making Use of Conditional Independence

- If E_1 and E_2 are conditionally independent given A, then

$$P(A|E_1, E_2) = \frac{P(E_1, E_2|A)P(A)}{P(E_1, E_2)} = \frac{P(E_1|A)P(E_2|A)P(A)}{P(E_1, E_2)}$$

or

$$\begin{aligned} P(A|E_1, E_2) &= P(A)\,\frac{P(E_1|A)}{P(E_1)}\,\frac{P(E_2|E_1, A)}{P(E_2|E_1)} \\ &= P(A)\,\frac{P(E_1|A)P(E_2|A)}{P(E_1, E_2)}. \end{aligned}$$

- $\frac{1}{P(E_1,E_2)}$ is the *normalization constant* and

$$\begin{aligned} P(E_1, E_2) &= P(E_1, E_2|A)P(A) + \\ &\quad P(E_1, E_2|\neg A)P(\neg A) \\ &= P(E_1|A)P(E_2|A)P(A) + \\ &\quad P(E_1|\neg A)P(E_2|\neg A)P(\neg A) \end{aligned}$$

The Meningitis Example

- Initially, $P(M) = 0.00002$.
- Updating for a stiff neck, we have:

$$P(M|S) = P(M)\frac{P(S|M)}{P(S)} = 0.00002 \times \frac{0.5}{0.05} = 0.0002.$$

- If the patient is found also to have a fever, $P(M)$ is updated again:

$$P(M|S, F) = P(M|S)\frac{P(F|M, S)}{P(F|S)} = P(M|S)\frac{P(F|M)}{P(F|S)}$$

- Assume that:

$$P(F|M) = 0.8$$

- It is not usually reasonable to assess $P(F|S)$ (because this denotes a relationship what is not causal!).
- Instead, we can use normalization:

$$P(\neg M|S, F) = P(\neg M|S)\frac{P(F|\neg M)}{P(F|S)}$$

- Assume that:

$$P(F|\neg M) = 0.02$$

- Determine $P(M|S, F)$:

$$P(M|S, F) = \alpha(0.0002 \times 0.8) = 0.00016$$

$$P(\neg M|S, F) = \alpha(0.9998 \times 0.02) = 0.019996$$

$$P(M|S, F) = \frac{0.00016}{0.00016 + 0.019996} = 0.007938$$

$$P(\neg M|S, F) = \frac{0.019996}{0.00016 + 0.019996} = 0.992062$$