Motivation

- Modeling uncertainty

  1. Inherent randomness in world, unreliable info (sensory systems: vision, sonar, speech, e
  2. Gross statistical description of world (which is complex + deterministic)
  3. Probability as guardian of common sense reasoning
  4. Empirical success of modern AI

Review of Probability

- Discrete random variable X (capitalized)

  Domain of possible values $\{x_1, x_2, \ldots, x_n\}$ (lower case)

  Ex: month M $\{m_1 = \text{Jan}, m_2 = \text{Feb}, \ldots, m_{12} = \text{Dec}\}$

- "Unconditional" or "prior" probabilities $P(X=x)$

  Basic axioms

(i) $P(X=x) \geq 0$ probability that event $X=x$ is true

(ii) $\sum_{i=1}^{n} P(X=x_i) = 1$

(iii) $P(X=x_i \text{ or } X=x_j) = P(X=x_i) + P(X=x_j)$ if $x_i \neq x_j$

Probabilities add for union of mutually exclusive events

- "Conditional" or "posterior" probabilities

  $P(X=x_i \mid Y=y_j)$ probability that $X=x_i$ given $Y=y_j$

  In general $P(X=x_i \mid Y=y_j) \neq P(X=x_i)$

Ex: conditional dependence

weather W, $\{w_1 = \text{sunny},\ w_2 = \text{rainy}\}$

$P(W = \text{sunny}) = 0.9$

$P(W = \text{sunny} \mid M = \text{Jan}) = 0.8$

$P(W = \text{sunny} \mid M = \text{Aug}) = 0.98$

Probability can change in either direction

Ex: conditional independence

day of week D $\{d_1 = \text{Sun},\ d_2 = \text{Mon}, \ldots, d_7 = \text{Sat}\}$

$P(W = \text{sunny} \mid D = \text{Tues}) = P(W = \text{sunny})$

$P(W = \text{sunny} \mid M = \text{Jan}, D = \text{Tues}) = P(W = \text{sunny} \mid M = \text{Jan})$

Also true:

(i) $P(X = x_i \mid Y = y_j) \geq 0$

(ii) $\displaystyle\sum_{i=1}^{n} P(X = x_i \mid Y = y_j) = 1$

sum over i, not j

- "Joint" probability

  $P(X = x_i, Y = y_j)$

- Product rule: from conditional prob to joint prob

  For all i, j: $P(X = x_i, Y = y_j) = P(X = x_i \mid Y = y_j) P(Y = y_j)$

  $P(X = x_i, Y = y_j) = P(Y = y_j \mid X = x_i) P(X = x_i)$

- Generalized product rule

$$P(A=a_i, B=b_j, C=c_k, D=d_\ell \ldots) = P(A=a_i)\,P(B=b_j \mid A=a_i)$$

$$P(C=c_k \mid A=a_i, B=b_j)\,P(D=d_\ell \mid A=a_i, B=b_j, C=c_k) \ldots$$

Easier to assess conditional probabilities (RHS) than joint probabilities (LHS)

Ex: A = earthquake

B = tsunami

C = power outage

D = reactor breach

Easier to reason about one event at a time...

- Marginalization: from joint distribution over several variables to the joint distribution over a subset of them

"marginal" probs

$$P(X=x_i) = \sum_j P(X=x_i, Y=y_j)$$

$$P(X=x_i, Y=y_j) = \sum_k P(X=x_i, Y=y_j, Z=z_k)$$

- Shorthand notation

(i) Implied universality

$$P(X,Y) = P(X|Y)P(Y) = P(Y|X)P(X)$$

implies that equality holds for all possible assignments $X=x_i$, $Y=y_j$ across LHS and RHS

(ii) Implied assignment

$$P(x, y, z) = P(X=x, Y=y, Z=z)$$

omit the assignment when context is unambiguous

Ex: product rule

$$P(a,b,c,d) = P(a)\,P(b|a)\,P(c|a,b)\,P(d|a,b,c)$$

- Bayes rule: relates conditional probabilities to other conditional probabilitie

$$P(X|Y) = \frac{P(Y|X)\,P(X)}{P(Y)}$$

if you observe an effect Y, you can infer the cause X.

$P(\text{effect}\,|\,\text{cause})$: causal direction
$P(\text{cause}\,|\,\text{effect})$: diagnostic direction

Ex: cancer diagnosis

Given: 1% population has cancer

Test has 10% false negative rate
Test has 20% false positive rate

You test positive. Do you have cancer?

- Random variables

Health $\in$ {cancer, $\neg$ cancer}
Test $\in$ {positive, negative}

- Probabilities

$P(\text{cancer}) = 0.01$

$P(\neg\text{cancer}) = 1 - 0.01 = 0.99$

$P(\text{pos} \mid \text{cancer}) = 1 - 0.1 = 0.9$

$P(\text{neg} \mid \text{cancer}) = 0.1$

$P(\text{pos} \mid \neg\text{cancer}) = 0.2$

$P(\text{neg} \mid \neg\text{cancer}) = 0.8$

We want to compute:

$$P(\text{cancer} \mid \text{pos}) = \frac{\overset{0.9}{P(\text{pos} \mid \text{cancer})} \times \overset{0.01}{P(\text{cancer})}}{\underset{0.207}{P(\text{pos})}}$$

$\boxed{\text{Bayes rule}}$

Use marginalization for denominator

$$P(\text{Test} = \text{pos}) = \sum_{\substack{\text{Health} \in \\ \{\text{cancer}, \neg\text{cancer}\}}} P(\text{Test} = \text{pos}, \text{Health})$$ marginalization

$$= \sum P(\text{pos} \mid \text{Health})\, P(\text{Health})$$ product rule

$$P(\text{pos}) = \overset{0.9}{P(\text{pos} \mid \text{cancer})}\, \overset{0.01}{P(\text{cancer})} + \underset{0.2}{P(\text{pos} \mid \neg\text{cancer})}\, \underset{0.99}{P(\neg\text{cancer})}$$ take the sum

$P(\text{pos}) = 0.207$

$$\Rightarrow \quad P(\text{cancer} \mid \text{pos}) = \frac{P(\text{pos} \mid \text{cancer})\, P(\text{cancer})}{P(\text{pos})} = \frac{0.9 \times 0.01}{0.207}$$

$$= 0.043 \approx 4.3\%$$

before test: $P(\text{cancer}) = 0.01 \quad (1\%)$

after test: $P(\text{cancer} \mid \text{pos}) = 0.043 \quad (4.3\%)$

Note: $P(\text{cancer} \mid \text{pos}) \ll P(\text{pos} \mid \text{cancer})$

- Conditioning on Background Evidence

Consider events X and Y and background evidence E.
Often useful to reason in context of background knowledge.

(i) conditionalized version of product rule

$$P(X,Y \mid E) = \frac{P(X,Y,E)}{P(E)}$$ product rule

$$P(X,Y \mid E) = \frac{P(X,Y,E)}{P(Y,E)} \times \frac{P(Y,E)}{P(E)}$$ cross-terms cancel (multiply by 1)

$$\downarrow \qquad\qquad \downarrow$$

$$\boxed{P(X,Y \mid E) = P(X \mid Y,E) \times P(Y \mid E)}$$ two applications of product rule

(ii) conditionalized version of Bayes rule

ordinary Bayes rule
$$P(X \mid Y) = \frac{P(Y \mid X)\, P(X)}{P(Y)}$$

conditionalized version
$$P(X \mid Y, E) = \frac{P(Y \mid X, E)\, P(X \mid E)}{P(Y \mid E)}$$

- conditional independence statements

the following three statements are equivalent:

(i) $P(X, Y \mid E) = P(X \mid E)\, P(Y \mid E)$

(ii) $P(X \mid Y, E) = P(X \mid E)$

(iii) $P(Y \mid X, E) = P(Y \mid E)$

any one of these implies the other two