1/9/14 Lecture 2

Motivation

* Modeling uncertainty

1. inherent randomness in world
2. gross statistical description of world (which is complex and deterministic)
3. probability as guardian of commonsense reasoning
4. emperical sucess 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 $\{m_1 = JAN, m_2 = FEB, \ldots, m_{12} = DEC\}$

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

sum over i, not j

* "Joint" probability

$P(X = x_i, Y = y_j)$ = prob that $X = x_i$ and $Y = y_j$

* 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

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

$P(C = c_k \mid A = a_i, B = b_j) P(D = d_l \mid A = a_i, B = b_j, C = c_k)$

$\ldots$

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 probabities to other conditional probabilities

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

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

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

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

$P(\text{pos}|\text{cancer}) = 1 - 0.1 = 0.9$ $\quad P(\text{neg}|\text{cancer}) = 0.1$

$P(\text{pos}|\neg\text{cancer}) = 0.2$ $\quad P(\text{neg}|\neg\text{cancer}) = 0.8$

We want to compute :

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

$\boxed{\text{Bayes Rule}}$

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

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

dang the sum

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

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

$$= 0.043 \approx 4.3\%$$

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

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

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

$$P(X,Y|E) = \frac{P(X,Y,E)}{P(Y,E)} \times \frac{P(Y,E)}{P(E)}$$ cross-terms cancel

two applications of product rule

$$\boxed{P(X,Y|E) = P(X|Y,E) \times P(Y|E)}$$

(ii) conditionalized version of Bayes rule

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

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

* Conditional independence statements

The following three statements are equivalent :

$$\left.\begin{array}{ll} \text{(i)} & P(X, Y \mid E) = P(X \mid E)\, P(Y \mid E) \\ \text{(ii)} & P(X \mid Y, E) = P(X \mid E) \\ \text{(iii)} & P(Y \mid X, E) = P(Y \mid E) \end{array}\right\} \text{any one of these implies the other two}$$