CSE 150 - Lecture 2 - 1/6/10

Motivation

* Modeling Uncertainty
  1) inherent randomness in world
  2) gross statistical description of world, which is complex & deterministic
  3) probability: guardian of common sense reasoning

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 the 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
  $P(X=x_i | Y=y_j)$ probability that $X=x_i$ given $Y=y_j$
  In general: $P(X=x_i|Y=y_j) \neq P(X=x_i)$
  Ex: conditional dependence
  weather W, $\{w_1 = \text{sunny}, w_2 = \text{rainy}\}$
  $P(W=\text{sunny}) = 0.9$
  $P(W=\text{sunny} | M=\text{jan}) = 0.8$
  $P(W=\text{sunny} | M=\text{aug}) = 0.98$

  probability can change in either direction

Ex: conditional independence

Day of week D, $\{d_1 = \text{sun}, d_2 = \text{mon}, \ldots\}$

$P(W = \text{rain} \mid D = \text{tues}) = P(W = \text{rain})$

Also true:

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

(ii) $\sum_i P(X = x_i \mid Y = y_j) = 1$ note: sum is over i, not over j !!!

- "Joint" probabilities

$P(X = x_i, Y = y_j)$ = prob that $X = x_i$ <u>and</u> $Y = y_j$

- Product Rule: from conditional probabilities to joint

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

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

- Generalized Product Rule

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

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

Ex: A = wake up on time

B = eat breakfast

C = hit traffic

D = arrive on time at UCSD

- Marginalization: from joint distribution to marginal distribution

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

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

Probs on LHS (over some subset of variables) are called "marginal" probabilities.

* Shorthand Notation

(i) implied universality

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

implies that equalities hold for all possible assignments $X = x_i, Y = y_j$

(ii) implied assignment

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

omit assignment when context is unambiguous

Ex: product rule $P(a,b,c,d,\dots) = P(a)P(b|a)P(c|a,b)P(d|a,b,c)\dots$

* $\boxed{\text{Bayes Rule}}$ relates conditional probabilities to other " "

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

if you observe an effect, you can infer the cause

Y

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

DIAGNOSIS $\in$ {cancer, healthy}

TEST $\in$ {pos, neg}

* Probabilities

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

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

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

We want to find:

$$P(\text{cancer} \mid \text{pos}) = \frac{\overset{(0.90)}{P(\text{pos} \mid \text{cancer})}\ \overset{(0.01)}{P(\text{cancer})}}{\underset{(0.207)}{P(\text{pos})}}$$

$\leftarrow$ need marginalization to find this

marginalization:

$$P(\text{pos}) = \sum_{\substack{\text{Diagnosis} \in \\ \{\text{cancer, healthy}\}}} P(\text{Test} = \text{pos}, \text{Diagnosis})$$

$$= \sum_{\substack{\text{cancer,} \\ \text{healthy}}} P(\text{pos} \mid \text{Diagnosis}) P(\text{Diagnosis}) \qquad \text{product rule}$$

$$= P(\text{pos} \mid \text{cancer}) P(\text{cancer}) + P(\text{pos} \mid \text{healthy}) P(\text{healthy})$$

$$= (0.90)(0.01) + (0.20)(0.99) = 0.207$$

Bayes Rule:

$$P(\text{cancer} \mid \text{pos}) = \frac{P(\text{pos} \mid \text{cancer}) P(\text{cancer})}{P(\text{pos})} = \frac{0.9 \times 0.01}{0.207} = \underline{0.043}$$

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

Often useful to reason in context of background knowledge.

Consider events X and Y, and background evidence E.

(i) conditionalized version of product rule

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

ordinary product rule

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

(ii) conditionalized version of Bayes rule

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

with background evidence $\quad P(X|Y,E) = \dfrac{P(Y|X,E)\,P(X|E)}{P(Y|E)}$

* Conditional independence statements

The following three statements are equivalent:

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

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

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

any one of these stetements implies the other two (HW)

* Kullback - Leibler Divergence (KL)

How to measure difference between two distributions?

Let $p_i = P(X=x_i|E)$
$\quad q_i = P(X=x_i|E')$ conditioned on different evidence $E \neq E'$

Define: $KL(p,q) = \sum_i p_i \log(p_i/q_i)$

Properties of KL "distance"

(i) $KL(p,q) \geq 0$, ~~with equality~~ vanishing only if $p_i = q_i$ for all i

(ii) $KL(p,q) \neq KL(q,p)$ in general