Random Variables

Random variables

Defined by functions mapping outcomes to values

By choice, whatever we are interested in

Typically denoted by uppercase letters (e.g., X)

Generic values are corresponding lower case letters

Shorthand: P(x) = P(X=x)

Value "type" is arbitrary (typically categorical or real)

Example (from K&F)

Outcomes are student grades (A,B,C)

Random variable G=f_{GRADE}(student)

$$P(A) \equiv P(G = A) \equiv P(\{w \in \Omega : f_{GRADE}(w) = A\})$$

Basic terminology and rules

Conditional probability

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

Bayes

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

 $P(X | Y) \propto P(Y | X) P(X)$

(when Y is constant, i.e., evidence)

Joint Distributions

Joint distribution of random variables

$$P(X,Y) \equiv P(X=x,Y=y) \equiv P(\{w \in \Omega : X(w) = x \text{ and } Y(w) = y\})$$

Conditional definition, Bayes rule, chain rule all apply.

Marginal distributions ("sum rule")

$$P(X) = \sum_{Y} P(X, Y)$$

Chain (product) rule (two variable case of chain rule)

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

Normalization

Often we will deal with quantities or functions which are proportional to probabilities (OK if we just want argmax)

$$p(x) \propto P(X = x)$$

To get probabilities we normalize: $P(X = x) = \frac{p(x)}{\sum_{i=0}^{n} p(x_i)}$

Example: $P(X|Y) \propto P(X,Y)$

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

Independence

$$X \perp Y \iff P(X|Y) = P(X) \text{ or } P(Y)=0$$

$$X \perp Y \iff P(X,Y) = P(X)P(Y)$$

Conditional Independence

$$X \perp Y \mid Z \iff P(X,Y|Z) = P(X \mid Z)P(Y \mid Z)$$

Probabilistic Queries

Bold face because these are vectors of variables

Organize variables into

Evidence (observed), E

Query (what you want to know), Y

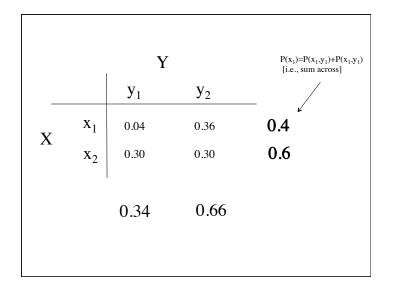
Hidden (leftover), X (for completeness)

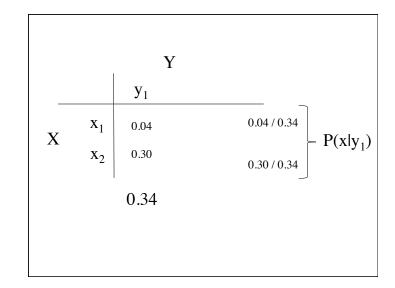
Generic Query: P(Y|E)

MAP Query: $MAP(\mathbf{W} \mid \mathbf{e}) = \text{Argmax } P(w, e)$

Example

$$\begin{array}{c|ccccc} & & & Y & & & \\ & & y_1 & & y_2 & & \\ \hline X & x_1 & & 0.04 & & 0.36 & & \\ X & x_2 & & 0.30 & & 0.30 & & \end{array}$$





		Y		
		\mathbf{y}_1	\mathbf{y}_2	_
X	\mathbf{x}_1	0.04	0.36	0.4
	\mathbf{x}_2	0.30	0.30	0.6
	١	0.34	0.66	
Arg max $P(X,Y)$ is (x_1, y_2)		Arg max $P(X)$ is (x_2) Arg max $P(Y)$ is (y_2)		
rg m	ax P(X,Y	(Arg	max P(X), Arg	max P(Y))

Discrete Distributions (Bernoulli)

$$x \in \{0,1\}$$
 (e.g., 1 is "heads" and 0 is "tails")

$$p(x=1|\mu)=\mu$$

$$Bern(x \mid \mu) = \mu^{x} (1 - \mu)^{(1-x)}$$

Code for sampling a Bernoulli

if (a<u) return heads
else return tails</pre>

Discrete Distributions (Binomial)

How likely it is that we get m "heads" in N tosses?

$$Bin(m|N,\mu) = \binom{N}{m} \mu^m (1-\mu)^{N-m}$$

where
$$\begin{pmatrix} N \\ m \end{pmatrix} \equiv \frac{N!}{(N-m)!m!}$$