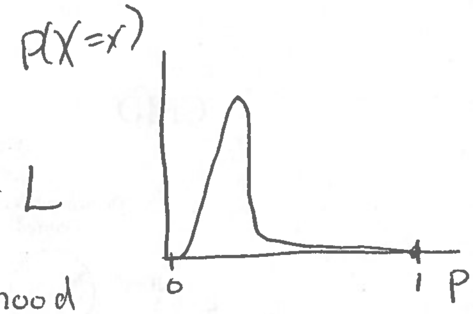
Lecture 2

1/15/2016

Needle stick example

$N = 51$ patients

$x = 3$ get hepatitis



$$P(X=x) = \binom{N}{x} p^x (1-p)^{N-x} = L$$

likelihood function

$$\log P(X=x) = \underbrace{\log \binom{N}{x}}_{\text{drop b/c doesn't influence where the mean is}} + \log p^x + \log (1-p)^{N-x}$$

abuse of notation!

$$l(p) = x \log p + (N-x) \log (1-p)$$

$$\frac{dl}{dp} = x \frac{1}{p} + (N-x) \frac{1}{1-p}(-1) = 0$$

Take derivative and set to 0 to find maximum likelihood est. of p

$$\Rightarrow p = \frac{x}{N}$$

The empirical relative frequency corresponds to the maximum likelihood!

NB: Max. likelihood is not in general unbiased. In this case it is!

As sample gets bigger

$p(|\hat{p} - p| > \epsilon) \to 0$ · Consistency

ϵ – tolerance

bias - if you repeat the experiment on avg you will on avg get the true p (even if sample is small)

Sidenote on Expectation

$$EX = \sum_x x \, P(X=x)$$

$$E[g(X)] = \sum_x g(x) P(X=x) \quad \text{expectation of transforms of } X$$

eg.

$$E[X^2] = \sum_x x^2 P(X=x) \quad \text{second moment}$$

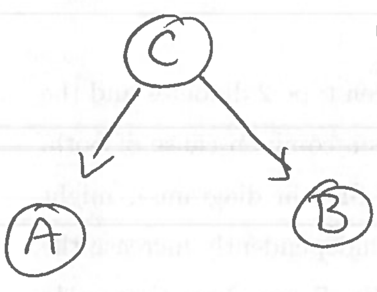
Conditional Independence

A and B ~~and~~ are conditionally ind given C if

$$P(AB|C) =$$

?

Bayes Nets

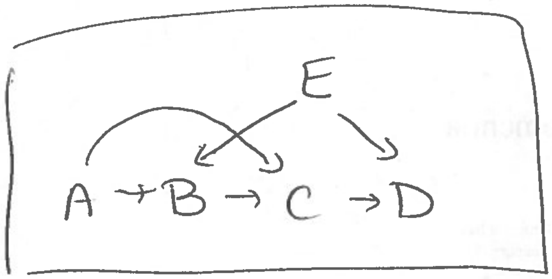


C influ A
" B

$$P(ABC) = P(A|BC)\, P(BC)$$
$$= P(A|BC)\, P(B|C) P(C)$$

= $P(A|C)$ if cond. indep.

Topological Sort of DAGs



Conditionally indep of all nondescendants given your parents

AEBCD

Verma graph

$$P(ABCDE) = P(D|ABCE)\,P(C|ABE)\,P(B|AE)\,P(E|A)\,P(A)$$

$$= P(D|CE)\,P(C|AB)\,P(B|AE)\,P(E)\,P(A)$$

$$\boxed{X_0 \longrightarrow X_1 \longrightarrow X_2 \longrightarrow X_3}$$

Topological Sort $X_0\, X_1\, X_2\, X_3$

eg. prev. at $t=0$ → prev at $t=1$ → prev at $t=2$ → prev at $t=3$

$$P(X_0 X_1 X_2 X_3) = P(X_3|X_1X_2X_3)\,P(X_2|X_1X_0)\,P(X_1|X_0)\,P(X_0)$$

$$= P(X_3|X_2)\,P(X_2|X_1)\,P(X_1|X_0)\,P(X_0)$$

If n Xs - Chain rule for Bayes networks

$$P(X_0 \ldots X_n) = \prod_i P(X_i|\text{Parents}(X_i))$$

This is a Markov Chain

Causal DAGs are examples of Bayes Networks
(You may ~~write~~ use Bayes Networks for other stuff too)

Random Sample of Prevalences

Hidden Markov Model

$$
\begin{tikzcd}
y_0 & y_1 & y_2 & y_3 \\
X_0 \arrow[u] \arrow[r] & X_1 \arrow[u] \arrow[r] & X_2 \arrow[u] \arrow[r] & X_3 \arrow[u]
\end{tikzcd}
$$

Sample of prevalences

prevalences at $t=0,1,2,3$

Needle stick Example - time evolving needle contamination

X_τ 0,1 ~~con~~ contamination status of needle

b is prob of decontamination

$P(X_\tau = 1)$ prob cont. at time τ

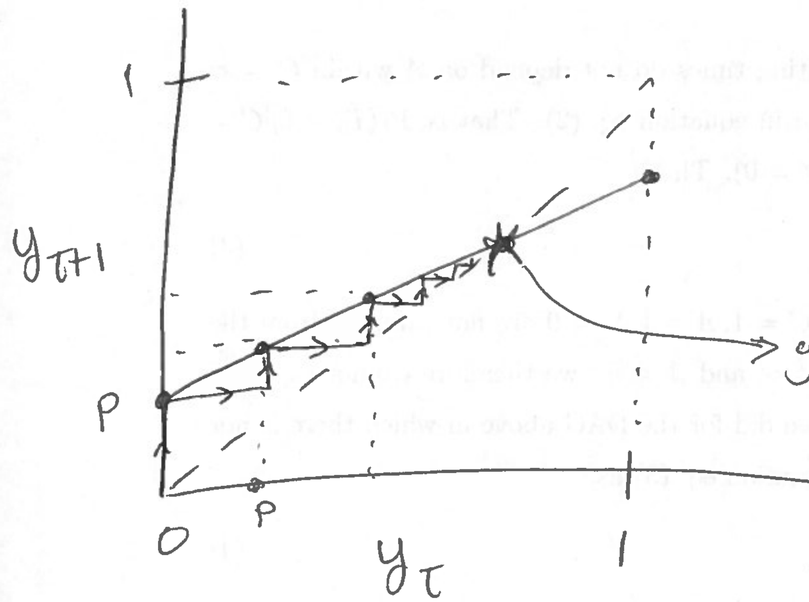
$$P(X_{\tau+1} = 1) = P(X_\tau = 1)\,P(X_{\tau+1} = 1 \mid X_\tau = 1)$$

$$+\, P(X_\tau = 0)\,\underbrace{P(X_{\tau+1} = 1 \mid X_\tau = 0)}_{\substack{\text{prob used on} \\ \text{a "contaminated"} \\ \text{person} \\ p}}$$

$$= \cancel{P(X_\tau = 1)}\ \underbrace{P(X_\tau)}_{=y_\tau}(1-b+bp) + (1 - P(X_\tau = 1))\,p$$

$$y_{t+1} = p + y_\tau(1-b)(1-p)$$

Linear Difference Equation



graphically we work out way up to this pt

pt where crosses 45° line

Cobb web Diagram

Start uninf. $y_0 = 0$

$$\Rightarrow y_1 = p$$

$$y_2 = p + p(1-b)(1-p)$$

Convergence of a Markov Chain to an Equilibrium * point (where crosses 45° line)

$$y = p + y\underbrace{(1-b)(1-p)}_{k}$$

$$y = p + yk \Rightarrow y = \frac{p}{1-(1-b)(1-p)}$$

Ensemble – large collection of needles

Con * This is an example of a dynamical system.

At equilibrium – flow of $u \to c$ needles equals the flow of $c \to u$

<u>Detailed Balance</u> – flows = for all pairs of variables (Pay attention for Markov Chains

Conditional Expectation

$$E[Y \mid X = x] = \sum_y y \, P(Y \mid X = x)$$

		$x=0$	$x=1$
y	0	0.4	0.2
	1	0.1	0.3

$P(X=0, Y=0) = 0.4$

$$E[Y \mid X = 0] = 0 \cdot 0.4 + 1 \cdot 0.1 / 0.5$$

$$= 0 \, \frac{P(Y=0, X=0)}{P(X=0)} + 1 \, \frac{P(Y=1, X=0)}{P(X=0)}$$

x	
0	0.2
1	0.6

$E[Y \mid X]$ is a transform of X!

$$E[g(X)] = \sum_x g(x) P(X = x)$$

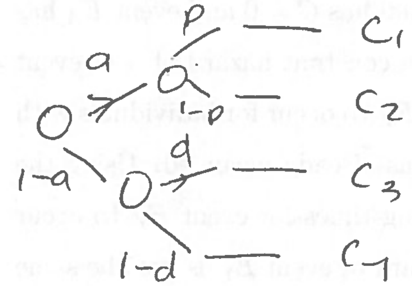
$$E[\underbrace{E[Y \mid X]}_{\text{This is a RV}}] = \sum_x E[Y \mid X = x] \, P(X = x)$$

$$= \sum_x \sum_y P(Y = y \mid X = x) \, P(X = x)$$

$$= \sum_y y \sum_x P(X = x, Y = y)$$

$$= \sum_y P(Y=y) = EY$$

$$E[E[Y|X]] = E[Y]$$ Adam's Law or Law of Iterated Expectation



* cost effectiveness analysis uses this Law

- a = prob antibiotic
- p = prob of complications given antibiotic treatment
- d = prob of death given no treatment

C's costs associated w/ each outcome

Conditional Variance

$\text{var}(Y|X)$ also a transform of X!

$$\text{var}(Y|X=x) = E[Y^2|X=x] - (E[Y|X=x])^2$$

$$\text{var}(Y|X) = E[Y^2|X] - (E[Y|X])^2$$

$$E[\text{var}(Y|X)] = E[E[Y^2|X]] - E[(E[Y|X])^2]$$

$$= EY^2 - E[(E(Y|X))^2]$$

$$\mathrm{var}(E[Y|X]) = E[(E[Y|X])^2] - (E[E[Y|X]])^2$$

$$= E[(E[Y|X])^2] - (EY)^2$$

using 2 examples above:

$$\underbrace{E[\mathrm{var}(Y|X)]}_{\text{variablity explained by } Y} + \underbrace{\mathrm{var}(E[Y|X])}_{\text{variability explained by } X} = EY^2 - (EY)^2 = \mathrm{var}\,Y$$

Can decompose the var Y over X using this.