Avd. Matematisk statistik

KTH Matematik

BAYESIAN NETWORKS and STATISTICAL GENETICS :

Vt 2010

EXERCISE SET 2: due to 29th of April (in case there is a meeting)

1.

$$U = (X, Y, Z), p(U) = p(X \mid Z)\, p(Y \mid Z)\, p(Z)$$

X, Y, Z are binary.

We have data, four complete and independent observations $U^{(i)}$, $i = 1, 2, 3, 4$:

	X	Y	Z
$U^{(1)}$	0	1	0
$U^{(2)}$	1	1	0
$U^{(3)}$	0	0	1
$U^{(4)}$	0	1	0

By the word 'complete' we mean that there are no missing values in any of $U^{(i)}$.

$$p(U^{(1)}) \cdot p(U^{(2)}) \cdot p(U^{(3)}) \cdot p(U^{(4)})$$

$$p(U^{(i)}) = p\left(X^{(i)} \mid Z^{(i)}\right) p\left(Y^{(i)} \mid Z^{(i)}\right) p\left(Z^{(i)}\right)$$

We introduce the probabilities of success (conditioned on the state of the parent)

$$\theta_{x|z=0} = P(X = 1 \mid Z = 0), \theta_{x|z=1} = P(X = 1 \mid Z = 1)$$

and analogously

$$\theta_{y|z=0} = P(Y = 1 \mid Z = 0), \theta_{y|z=1} = P(Y = 1 \mid Z = 1)$$

$$\theta_z = P(Z = 1)$$

Find the maximum likelihood estimates of $\theta_{x|z=0}$, $\theta_{x|z=1}$, $\theta_{y|z=0}$, $\theta_{y|z=1}$, and θ_z.

2. C, E, M are binary (or here: yes− no). The simultaneous distribution is

$$p(C, E, M) = p(C \mid E, M)\, p(E)\, p(M)$$

We need three probability tables for this model.

$$P(E = \text{yes}) = 0.1, P(M = \text{yes}) = 0.2.$$

E and M are independent.

$$P(C = \text{yes} \mid M = \text{yes}, E = \text{yes}) = 1, P(C = \text{yes} \mid M = \text{no}, E = \text{yes}) = 1,$$

$$P(C = \text{yes} \mid M = \text{yes}, E = \text{no}) = 0.5, P(C = \text{yes} \mid M = \text{no}, E = \text{no}) = 0.$$

Suppose we know that the event $C = \text{yes}$ has occurred (or *is instantiated*).

Compute the posterior distribution for the two random variables M and E given $C = \text{yes}$ using Bayes' rule.