

inference

CSCI
373



let X be Xena's blood type.

let X_f, X_o be the genes Xena inherited from her parents



let Y be Yves's blood type.

let Y_f, Y_o be the genes Yves inherited from his parents



let Z be Zelda's blood type.

let Z_f, Z_o be the genes Zelda inherited from Xena and Yves



what is $P(Y=AB | Z=B)$, the probability that Yves is type AB, given we learn that Zelda is type B? assume we know how to compute joint probability $P(x_f, x_o, x, y_f, y_o, y, z_f, z_o, z)$

$$P(Y=AB | Z=B)$$

$$= \frac{P(Y=AB, Z=B)}{P(Z=B)}$$

$$= \frac{\sum_{x_\varphi} \sum_{x_{\sigma'}} \sum_{x} \sum_{y_\varphi} \sum_{y_{\sigma'}} \sum_{z_\varphi} \sum_{z_{\sigma'}} P(x_\varphi, x_{\sigma'}, x, y_\varphi, y_{\sigma'}, Y=AB, z_\varphi, z_{\sigma'}, Z=B)}{\sum_{x_\varphi} \sum_{x_{\sigma'}} \sum_{x} \sum_{y_\varphi} \sum_{y_{\sigma'}} \sum_{y} \sum_{z_\varphi} \sum_{z_{\sigma'}} P(x_\varphi, x_{\sigma'}, x, y_\varphi, y_{\sigma'}, y, z_\varphi, z_{\sigma'}, Z=B)}$$

$$P(Y=AB | Z=B)$$

$$y|x = \frac{P(Y=AB, Z=B)}{P(Z=B)}$$

$$= \frac{\sum_{x_\varphi} \sum_{x_{\sigma'}} \sum_x \sum_{y_\varphi} \sum_{y_{\sigma'}} \sum_{z_\varphi} \sum_{z_{\sigma'}} P(x_\varphi, x_{\sigma'}, x, y_\varphi, y_{\sigma'}, Y=AB, z_\varphi, z_{\sigma'}, Z=B)}{\sum_{x_\varphi} \sum_{x_{\sigma'}} \sum_x \sum_{y_\varphi} \sum_{y_{\sigma'}} \sum_y \sum_{z_\varphi} \sum_{z_{\sigma'}} P(x_\varphi, x_{\sigma'}, x, y_\varphi, y_{\sigma'}, y, z_\varphi, z_{\sigma'}, Z=B)}$$

$$P(Y=AB | Z=B)$$

$$y|x = \frac{P(Y=AB, Z=B)}{P(Z=B)}$$



$$= \frac{\sum_{x_q} \sum_{x_{\sigma'}} \sum_{x} \sum_{y_q} \sum_{y_{\sigma'}} \sum_{z_q} \sum_{z_{\sigma'}} P(x_q, x_{\sigma'}, x, y_q, y_{\sigma'}, Y=AB, z_q, z_{\sigma'}, Z=B)}{\sum_{x_q} \sum_{x_{\sigma'}} \sum_{x} \sum_{y_q} \sum_{y_{\sigma'}} \sum_{y} \sum_{z_q} \sum_{z_{\sigma'}} P(x_q, x_{\sigma'}, x, y_q, y_{\sigma'}, y, z_q, z_{\sigma'}, Z=B)}$$

$$P(Y=AB | Z=B)$$

$$y|x = \frac{P(Y=AB, Z=B)}{P(Z=B)}$$



$$= \sum_{x_q} \sum_{x_{\sigma'}} \sum_{x} \sum_{y_q} \sum_{y_{\sigma'}} \sum_{z_q} \sum_{z_{\sigma'}} P(x_q, x_{\sigma'}, x, y_q, y_{\sigma'}, Y=AB, z_q, z_{\sigma'}, Z=B)$$

$$\sum_{x_q} \sum_{x_{\sigma'}} \sum_{x} \sum_{y_q} \sum_{y_{\sigma'}} \sum_{y} \sum_{z_q} \sum_{z_{\sigma'}} P(x_q, x_{\sigma'}, x, y_q, y_{\sigma'}, y, z_q, z_{\sigma'}, Z=B)$$

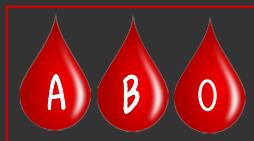
↑
how many terms are in this summation?

$$P(Y=AB | Z=B)$$

$$y|x = \frac{P(Y=AB, Z=B)}{P(Z=B)}$$



$$= \frac{\sum_{x_q} \sum_{x_{\sigma}} \sum_{\alpha} \sum_{y_q} \sum_{y_{\sigma}} \sum_{z_q} \sum_{z_{\sigma}} P(x_q, x_{\sigma}, \alpha, y_q, y_{\sigma}, Y=AB, z_q, z_{\sigma}, Z=B)}{\sum_{x_q} \sum_{x_{\sigma}} \sum_{\alpha} \sum_{y_q} \sum_{y_{\sigma}} \sum_y \sum_{z_q} \sum_{z_{\sigma}} P(x_q, x_{\sigma}, \alpha, y_q, y_{\sigma}, y, z_q, z_{\sigma}, Z=B)}$$



$$P(Y=AB | Z=B)$$

$$y|x = \frac{P(Y=AB, Z=B)}{P(Z=B)}$$



$$= \frac{\sum_{x_q} \sum_{x_{\sigma}} \sum_{x} \sum_{y_q} \sum_{y_{\sigma}} \sum_{z_q} \sum_{z_{\sigma}} P(x_q, x_{\sigma}, x, y_q, y_{\sigma}, Y=AB, z_q, z_{\sigma}, Z=B)}{\sum_{x_q} \sum_{x_{\sigma}} \sum_{x} \sum_{y_q} \sum_{y_{\sigma}} \sum_{y} \sum_{z_q} \sum_{z_{\sigma}} P(x_q, x_{\sigma}, x, y_q, y_{\sigma}, y, z_q, z_{\sigma}, Z=B)}$$

3 3 4 3 3 4 3 3

A, B, AB, O



$$P(Y=AB | Z=B)$$

$$y|x = \frac{P(Y=AB, Z=B)}{P(Z=B)}$$



$$= \frac{\sum_{x_f} \sum_{x_o} \sum_x \sum_{y_f} \sum_{y_o} \sum_{z_f} \sum_{z_o} P(x_f, x_o, x, y_f, y_o, Y=AB, z_f, z_o, Z=B)}{\sum_{x_f} \sum_{x_o} \sum_x \sum_{y_f} \sum_{y_o} \sum_y \sum_{z_f} \sum_{z_o} P(x_f, x_o, x, y_f, y_o, y, z_f, z_o, Z=B)}$$

$$= 3 \times 3 \times 4 \times 3 \times 3 \times 4 \times 3 \times 3$$

= 11664 terms

one approach to automating probability computations

query

$$P(Y=AB | Z=B)$$

joint
distribution

X_q	X_{σ}	X	Y_q	Y_{σ}	Y	Z_q	Z_{σ}	Z
A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	B
A	A	A	A	A	A	A	A	B
⋮								
0	0	0	0	0	0	0	0	0

$$P_{46656}$$

inference
algorithm

$$P(Y=AB | Z=B)$$

$$y|x = \frac{P(Y=AB, Z=B)}{P(Z=B)}$$

$$\text{total} = \frac{\sum_{x_q, x_{\sigma}, x} \sum_{y_q, y_{\sigma}, y} \sum_{z_q, z_{\sigma}, z} P(x_q, x_{\sigma}, x, y_q, y_{\sigma}, y=AB, z_q, z_{\sigma}, z=B)}{\sum_{x_q, x_{\sigma}, x} \sum_{y_q, y_{\sigma}, y} \sum_{z_q, z_{\sigma}, z} P(x_q, x_{\sigma}, x, y_q, y_{\sigma}, y, z_q, z_{\sigma}, z)}$$

. 18

what is a drawback of this approach?

One approach to automating probability computations

query

$$P(Y=AB \mid Z=B)$$

joint distribution

X_φ	X_σ	X	Y_φ	Y_σ	Y	Z_φ	Z_σ	Z	
A	A	A	A	A	A	A	A	A	P_1
A	A	A	A	A	A	A	A	B	P_2
									:
O	O	O	O	O	O	O	O	O	P

→

joint distributions are huge

inference algorithm

$$P(Y=AB | Z=B) = \frac{P(Y=AB, Z=B)}{\cdot}$$

$$\text{total} = \frac{\sum_{x_0} \sum_{y_0} \sum_{z_0} \sum_{x'} \sum_{y'} \sum_{z'} P(x_0, x_0', x, y_0, y_0', Y=AB, z_0, z_0', Z=B)}{\sum_{x_0} \sum_{y_0} \sum_{z_0} \sum_{x'} \sum_{y'} \sum_{z'} P(x_0, x_0', x, y_0, y_0', Y=z_0, z_0', Z=B)}$$

One approach to automating probability computations

query

$$P(Y=AB | Z=B)$$

joint distribution

46656

inference algorithm

$$P(Y=AB | Z=B) = \frac{P(Y=AB, Z=B)}{\cdot}$$

$$\text{total} = \frac{\sum_{x_0} \sum_{y_0} \sum_{x'} \sum_{y'} \sum_{z_0} \sum_{z'} P(x_0, x', x, y_0, y', z_0, z') | AB, z_0, z', Z = B)}{\sum_{x_0} \sum_{y_0} \sum_{x'} \sum_{y'} \sum_{z_0} \sum_{z'} P(x_0, x', x, y_0, y', z_0, z') | AB, z_0, z', Z = B)}$$

joint distributions are huge
what can we do about that?