COMS 572: Homework #6

November 16, 2018 by 17:00pm

Professor Jin Tian

Le Zhang

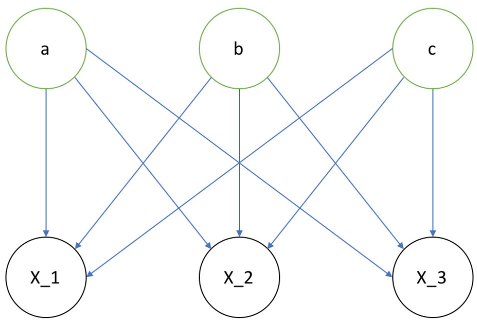
Problem 1

(20 pts.) We have a bag of three biased coins a, b, and c with probabilities of coming up heads of 20%, 60%, and 80%, respectively. One coin is drawn randomly from the bag (with equal likelihood of drawing each of the three coins), and then the coin is flipped three times to generate the outcomes X_1, X_2, and X_3.

a. Draw the Bayesian network corresponding to this setup and define the necessary CPTs.

Answer:

Bayesian network:



The variable "Coin" denotes the coins, we have the CPT for which coin we draw as below:

Coin	P(Coin)
a	1/3
b	1/3
c	1/3

CPT for X_i for each coin as below:

Coin	X_i	P(Coin)
a	head	0.2
b	head	0.6
c	head	0.8

b. Calculate which coin was most likely to have been drawn from the bag if the observed flips come out heads twice and tails once.

Answer:

Given information: $P(C|2H, 1T)$

With Bayesian fomula, we have:

$$P(C|2H, 1T) = P(2H, 1T|C)P(C)/P(2H, 1T)$$

Since $P(2H, 1T)$ and $P(C)$ are both independent of C, so the only part that determines the probability is $P(2H, 1T|C)$

For coin a, the possibility is:

$P(X_1 = H, X_2 = H, X_3 = tail|C = a) = P(X_1 = H|C = a)P(X_2 = H|C = a)P(X_3 = H|C = c)$, to calculate, it is $(1 - 0.2) \times 0.2 \times 0.2 = 0.032$. Remember we have same possibility for all 3 kind of orders, so it is $3 \times 0.032 = 0.096$.

Similarly, we have 0.432 for coin b and 0.384 for coin c.

As a result, **coin b** is the one that mostly likely have been drawn.

Problem 2

(40 pts.) Let H_x be a random variable denoting the handedness of an individual x, with possible values l or r. A common hypothesis is that left- or right-handedness is inherited by a simple mechanism; that is, perhaps there is a gene G_x, also with values l or r, and perhaps actual handedness turns out mostly the same (with some probability s) as the gene an individual possesses. Furthermore, perhaps the gene itself is equally likely to be inherited from either of an individuals parents, with a small nonzero probability m of a random mutation flipping the handedness.

A. Which of the three networks in Figure 14.20 claim that

$$P(G_{father}, G_{mother}, G_{child}) = P(G_{father})P(G_{mother})P(G_{child})?$$

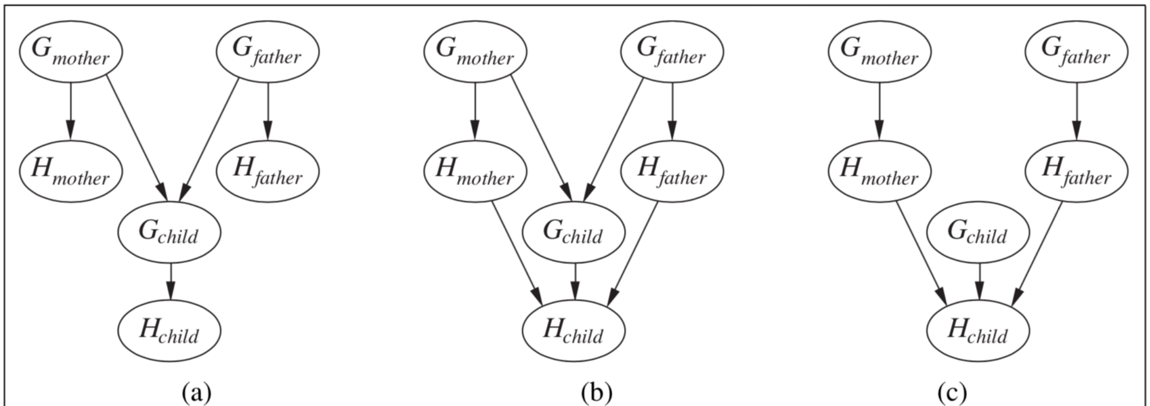


Figure 14.20 Three possible structures for a Bayesian network describing genetic inheritance of handedness.

Answer:

Graph (c) claims the statement.

B. Which of the three networks make independence claims that are consistent with the hypothesis about the inheritance of handedness?

Answer:

Graph (a) and (b). Because graph (c) implies that the gene of child is independent from gene of father and mother which conflict with the hypothesis.

C. Which of the three networks is the best description of the hypothesis?

Answer:

Graph (a) is the best description of the hypothesis. Because (b) and (c) implies that the left and right handedness are influenced by something other than genes which is conflict with the hypothesis.

D. Write down the CPT for the G_{child} node in network (a), in terms of s and m.

Answer:

G_{father}	l	l	r	r
G_{mother}	l	r	l	r
l	1-m	1/2	1/2	m
r	m	1/2	1/2	1-m

Problem 2 continued on next page...

E. Suppose that $P(G_{father} = l) = P(G_{mother} = l) = q$. In network (a), derive an expression for $P(G_{child} = l)$ in terms of m and q only, by conditioning on its parent nodes.

Answer:

$$\begin{aligned}
P(G_{child} = 1) &= \sum_{g_m, g_f} P(G_{child} = l | g_m, g_f) P(g_m, g_f) && (1)\\
&= \sum_{g_m, g_f} P(G_{child} = l | g_m, g_f) P(g_m) P(g_f) && (2)\\
&= (1 - m)q^2 + 0.5q(1 - q) + 0.5(1 - q)q + m(1 - q)^2 && (3)\\
&= q^2 - mq^2 + q - q^2 + m - 2mq + mq^2 && (4)\\
&= q + m - 2mq && (5)
\end{aligned}$$

F. Under conditions of genetic equilibrium, we expect the distribution of genes to be the same across generations. Use this to calculate the value of q, and, given what you know about handedness in humans, explain why the hypothesis described at the beginning of this question must be wrong.

Answer:

Genetic equilibrium implies that $P(G_{human} = l)$ are equal to all human. This means that the possibility of child, father, and mother are all the same. According to the result of previous question, we have:

$$q + m - 2mq = q$$

We can get $q = 0.5$ which means that the possibility of being left-handed and right-handed are the same. However, according to my own life experience, the number of left-handed people is much smaller than the number of right-handed people. Therefore, the hypothesis is wrong.