

Directed Graphical Models

Skye Hersh
Modern Computational Statistics

November 13, 2017

Bishop, Chapter 8

8.3

Consider three binary variables $a, b, c \in \{0, 1\}$ having the joint distribution given in the table below. Show by direct evaluation that this distribution has the property that a and b are marginally dependent, so that $p(a, b) \neq p(a)p(b)$, but that they become independent when conditioned on c , so that $p(a, b|c) = p(a|c)p(b|c)$ for both $c = 0$ and $c = 1$.

a	b	c	$p(a, b, c)$
0	0	0	.192
0	0	1	.144
0	1	0	.048
0	1	1	.216
1	0	0	.192
1	0	1	.064
1	1	0	.048
1	1	1	.096

Solution.

We can decompose the discrete distribution $p(a, b)$ into four respective probabilities for the four combinations that the variables a and b can take on: that for when $(a = 0, b = 0)$; $(a = 0, b = 1)$; $(a = 1, b = 0)$; and when $(a = 1, b = 1)$.

As a and b are Bernoulli random variables, we can begin by finding the marginal probabilities for either possible outcome for either variable by itself, summing over the rows in the table wherein such outcome's condition is met, e.g.,

$$p(a = 0) = \sum_b \sum_c p(a = 0, b, c).$$

Thus, we find that

$$p(a = 0) = 0.192 + 0.144 + 0.048 + 0.216 \tag{1}$$

$$= 0.600. \tag{2}$$

According to the same principle, we have

$$p(a = 1) = 0.400$$

$$p(b = 0) = 0.592$$

$$p(b = 1) = 0.408.$$

(Sensibly, $p(a = 0) + p(a = 1) = p(b = 0) + p(b = 1) = 1$.) The joint probability distribution $p(a, b)$ comprises four discrete probabilities – one for each combination of the variables' binary outcomes – wherein, for example,

$$p(a = 0, b = 0) = \sum_c p(a = 0, b = 0, c).$$

From the table and the marginal probabilities worked out above, we have that

a	b	$p(a, b)$	$p(a)p(b)$
0	0	$0.192 + 0.144 = 0.336$	$0.60 \cdot 0.592 = 0.3552$
0	1	$0.048 + 0.216 = 0.264$	$0.60 \cdot 0.408 = 0.2448$
1	0	$0.192 + 0.064 = 0.256$	$0.40 \cdot 0.592 = 0.2368$
1	1	$0.048 + 0.096 = 0.144$	$0.40 \cdot 0.408 = 0.1632$

Immediately, we see that in every case, $p(a, b)p(a)p(b)$, as a and b are marginally dependent. On the other hand, we find that they become independent when conditioned on c . We know that

$$p(a|c) = \frac{p(a, c)}{p(c)},$$

$$p(b|c) = \frac{p(b, c)}{p(c)},$$

$$p(a, b|c) = \frac{p(a, b, c)}{p(c)}.$$

Noting that $p(c = 0) = 0.48$ and $p(c = 1) = 0.52$, we can compute the marginal, conditional probabilities accordingly:

a	b	c	$p(a c)$	$p(b c)$	$p(a, b c)$	$p(a c)p(b c)$
0	0	0	$\frac{0.240}{0.480} = 0.500$	$\frac{0.384}{0.480} = 0.800$	$\frac{.192}{0.480} = 0.400$	$0.500 \times 0.800 = 0.400$
0	0	1	$\frac{0.360}{0.520} \approx 0.692$	$\frac{0.208}{0.520} = 0.400$	$\frac{.144}{0.520} \approx 0.277$	$0.692 \times 0.400 \approx 0.277$
0	1	0	$\frac{0.240}{0.480} = 0.500$	$\frac{0.096}{0.480} = 0.200$	$\frac{.048}{0.480} = 0.100$	$0.500 \times 0.200 = 0.100$
0	1	1	$\frac{0.360}{0.520} \approx 0.692$	$\frac{0.312}{0.520} = 0.600$	$\frac{.216}{0.520} \approx 0.415$	$0.692 \times 0.600 = 0.415$
1	0	0	$\frac{0.240}{0.480} = 0.500$	$\frac{0.384}{0.480} = 0.800$	$\frac{.192}{0.480} = 0.400$	$0.500 \times 0.800 = 0.400$
1	0	1	$\frac{0.160}{0.520} \approx 0.307$	$\frac{0.208}{0.520} = 0.400$	$\frac{.064}{0.520} \approx 0.123$	$0.307 \times 0.400 \approx 0.123$
1	1	0	$\frac{0.240}{0.480} = 0.500$	$\frac{0.096}{0.480} = 0.200$	$\frac{.048}{0.480} = 0.100$	$0.500 \times 0.200 = 0.100$
1	1	1	$\frac{0.160}{0.520} \approx 0.308$	$\frac{0.312}{0.520} = 0.600$	$\frac{.096}{0.520} \approx 0.185$	$0.308 \times 0.600 \approx 0.185$

The latter two columns show the equivalence between $p(a, b|c)$ and $p(a|c)p(b|c)$.

□

8.4

Evaluate the distributions $p(a)$, $p(b|c)$, and $p(c|a)$ corresponding to the joint distribution given in the original table. Hence show by direct evaluation that $p(a, b, c) = p(a)p(c|a)p(b|c)$. Draw the corresponding directed graph.

Solution.

We have the distributions for $p(a)$ and $p(b|c)$ above. The distribution $p(c|a) = \frac{p(a,c)}{p(a)}$ can be written accordingly:

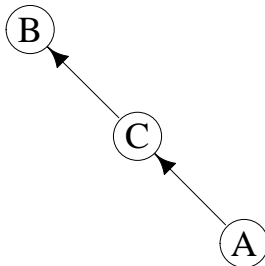
a	c	$p(c a)$
0	0	$\frac{0.240}{0.600} = 0.400$
0	1	$\frac{0.360}{0.600} = 0.600$
1	0	$\frac{0.240}{0.400} = 0.600$
1	1	$\frac{0.160}{0.400} = 0.400$

From here, we can complete the following table, demonstrating the equivalency of $p(a)p(c|a)p(b|c)$ and $p(a)p(c|a)p(b|c)$.

a	b	c	$p(a)p(c a)p(b c)$	$p(a)p(c a)p(b c)$
0	0	0	$0.600 \times 0.400 \times 0.800 = 0.192$	0.192
0	0	1	$0.600 \times 0.600 \times 0.400 = 0.144$	0.144
0	1	0	$0.600 \times 0.400 \times 0.200 = 0.048$	0.048
0	1	1	$0.600 \times 0.600 \times 0.600 = 0.216$	0.216
1	0	0	$0.400 \times 0.600 \times 0.800 = 0.192$	0.192
1	0	1	$0.400 \times 0.400 \times 0.400 = 0.064$	0.064
1	1	0	$0.400 \times 0.600 \times 0.200 = 0.048$	0.048
1	1	1	$0.400 \times 0.400 \times 0.600 = 0.096$	0.096

□

```
suppressMessages(library(igraph))
g <- graph_from_literal(A-->C-->B)
plot.igraph(g, edge.color = "black", vertex.color = "white",
  vertex.label = c("A", "C", "B"), vertex.size = 45, vertex.label.cex = 4,
  vertex.label.color = "black", edge.width = 1, edge.arrow.size = 2,
  edge.arrow.width = 1.2)
```



8.11

Consider the example of the car fuel system shown in Figure 8.21, and suppose that instead of observing the state of the fuel gauge G directly, the gauge is seen by the driver D who reports to us the reading on the gauge. This report is either that the gauge shows full $D = 1$ or that it shows empty $D = 0$. Our driver is a bit unreliable, as expressed through the following probabilities

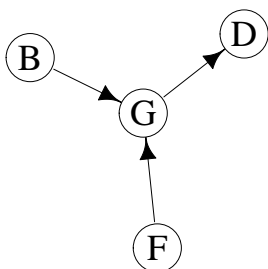
$$p(D = 1|G = 1) = 0.9$$

$$p(D = 0|G = 0) = 0.9.$$

Suppose that the driver tells us that the fuel gauge shows empty, in other words that we observe $D = 0$. Evaluate the probability that the tank is empty given only this observation. Similarly, evaluate the corresponding probability given also the observation that the battery is flat, and note that this second probability is lower. Discuss the intuition behind this result, and relate the result to Figure 8.54.

Solution. We can conceive of the problem in terms of a graph like this, wherein we've observed D :

```
g <- graph_from_literal(F-->G+-B, G-->D)
plot.igraph(g, edge.color = "black", vertex.color = "white",
            vertex.label = c("F","G","B", "D"), vertex.size = 45, vertex.label.cex = 4,
            vertex.label.color = "black", edge.width = 1, edge.arrow.size = 2,
            edge.arrow.width = 1.2)
```



We have the following prior probabilities regarding whether the battery is charged or flat and whether the tank is full or empty:

$$P(B = 1) = 0.9$$

$$P(F = 1) = 0.9.$$

And we know that, given the state of the fuel tank and the battery, the fuel gauge reads full with probabilities given by

$$p(G = 1|B = 1, F = 1) = 0.8$$

$$p(G = 1|B = 1, F = 0) = 0.2$$

$$p(G = 1|B = 0, F = 1) = 0.2$$

$$p(G = 1|B = 0, F = 0) = 0.1$$

We know that we have a prior $p(F = 0) = 0.1$. Following the graph, we can find the evidence term – the denominator for the posterior – like this

$$p(D = 0) = \sum_G \sum_F \sum_B p(D = 0|G)p(G|F, B)p(F)p(B) \quad (3)$$

$$= 0.352 \quad (4)$$

and the likelihood $p(D = 0|F = 0)$

$$= \sum_G \sum_B p(D = 0|G)p(G|B, F = 0)p(B) \quad (5)$$

$$= 0.748. \quad (6)$$

Thus, we can compute the posterior probability that the tank is empty

$$p(F = 0|D = 0) = \frac{0.1 \times 0.748}{0.352} \quad (7)$$

$$= 0.2125 \quad (8)$$

If we receive new information, and it turns out that the battery is definitely flat, we can recompute the posterior for the status of the fuel gauge. We can use the same prior $p(B = 0) = 0.1$. But now, we rearticulate the evidence and likelihood accordingly:

New evidence:

$$p(D = 0|B = 0) = \sum_G \sum_F p(D = 0|G)p(G|F, B = 0)p(F) \quad (9)$$

$$= 0.81 \quad (10)$$

New likelihood:

$$p(D = 0|B = 0, F = 0, G) = \sum_G p(D = 0|G)(G|B = 0, F = 0) \quad (11)$$

$$= 0.82 \quad (12)$$

And thus, the new posterior:

$$p(F = 0|D = 0, B = 0) = \frac{0.1 \times 0.82}{0.81} \quad (13)$$

$$\approx 0.101 \quad (14)$$

We see that the posterior probability given the information about the battery's status – specifically, that it is flat – is less than the posterior probability before we had this information. Intuitively, this is because the battery's being flat is likely to account, at least partly, for the gauge reading 0 – and thus, for the driver reporting 0 on the gauge.

□