CHAPTER 30

Solutions

1.1

$$p(x,y|z) = \frac{p(x,y,z)}{p(z)} = \frac{p(y|x,z)p(x,z)}{p(z)} = p(y|x,z)p(x|z) \tag{30.0.1}$$

$$p(x|y,z) = \frac{p(x,y,z)}{p(y,z)} = \frac{p(y|x,z)p(x,z)}{p(y,z)} = \frac{p(y|x,z)p(x|z)}{p(y|z)} \tag{30.0.2}$$

1.2

$$p(a \cup b) = p(a) + p(b) - p(a,b) \leq 1 \rightarrow p(a,b) \geq p(a) + p(b) - 1 \tag{30.0.3}$$

1.3

$$p(box = 1|redball) = \frac{p(box = 1, redball)}{p(redball)} \tag{30.0.4}$$

$$= \frac{p(redball|box = 1)p(box = 1)}{p(redball|box = 1)p(box = 1) + p(redball|box = 2)p(box = 2)} \tag{30.0.5}$$

$$= \frac{3/8}{3/8 + 2/7} = \frac{21}{37} \tag{30.0.6}$$

1.4 We want

$$p(\text{2 red in box}|\text{ 3 red drawn}) = \frac{p(\text{3 red drawn}|\text{2 red in box})p(\text{2 red in bag})}{p(\text{3 red drawn})}$$

A priori, there are 3 possibilities for what balls can be in the box

$$a: \text{ 2 white balls in box,} \qquad p(a) = 1/4 \tag{30.0.7}$$

$$b: \text{ 2 red balls in box,} \qquad p(b) = 1/4 \tag{30.0.8}$$

$$c: \text{ 1 red 1 white ball in box,} \qquad p(c) = 1/2 \tag{30.0.9}$$

In the use of Bayes' rule above we have

$$p(\text{3 red drawn}|\text{2 red in box})p(\text{2 red in box}) = 1 \times 1/4 \tag{30.0.10}$$

Also, using the shorthand for the above events:

$$p(\text{3 red drawn}) = p(\text{3 red drawn}|a)p(a) + p(\text{3 red drawn}|b)p(b) + p(\text{3 red drawn}|c)p(c) \tag{30.0.11}$$

$$= 0 + 1 \times 1/4 + 0 + 1/8 \times 1/2 \tag{30.0.12}$$

Hence the result is 4/5.

1.5 Simple approach in which we ignore the information that the person is the first to test positive:

$$p(terr = \mathsf{tr}|scan = \mathsf{tr}, inf) = \frac{p(scan = \mathsf{tr}|terr = \mathsf{tr}, inf)p(terr = \mathsf{tr}|inf)}{p(scan = \mathsf{tr}|terr = \mathsf{tr}, inf)p(terr = \mathsf{tr}|inf) + p(scan = \mathsf{tr}|terr = \mathsf{fa}, inf)p(terr = \mathsf{fa}|inf)} \tag{30.0.13}$$

$$= \frac{0.95 \times 0.01}{0.95 \times 0.01 + 0.05 \times 0.99} = 0.161 \tag{30.0.14}$$

More comprehensive answer:
Can interpret the problem as equivalent to computing the probability that, given my neighbour is declared to be the terrorist (the declaration process is such that this will necessarily be the first person to be declared a terrorist), that this person is indeed the terrorist.

$$
\begin{aligned}
&p(\text{neighbour is the terrorist}|\text{neighbour is declared the terrorist})\\
&= \frac{p(\text{neighbour is the terrorist, neighbour is declared the terrorist})}{p(\text{neighbour is declared the terrorist})}\\
&= \frac{p(\text{neighbour is correctly declared as the terrorist})}{p(\text{neighbour is declared the terrorist})}
\end{aligned}
$$

If we assume that:

$$
\begin{aligned}
&p(\text{someone is correctly declared as the terrorist})\\
&= \sum_{i=1}^{100} p(\text{person } i \text{ is correctly declared as the terrorist})\\
&= 100 \times p(\text{neighbour is correctly declared as the terrorist})
\end{aligned}
$$

and

$$
\begin{aligned}
&p(\text{someone is declared the terrorist})\\
&= \sum_{i=1}^{100} p(\text{person } i \text{ is declared the terrorist})\\
&= 100 \times p(\text{neighbour is declared as the terrorist})
\end{aligned}
$$

we can equivalently consider

$$
\begin{aligned}
&p(\text{neighbour is the terrorist}|\text{neighbour is declared the terrorist})\\
&= \frac{p(\text{someone is correctly declared as the terrorist})}{p(\text{someone is declared the terrorist})}
\end{aligned}
$$

(I think a better question would have been: "The police haul off the plane the first person for which the scanner tests positive. What is the probability that this person is a terrorist." In this setup, it is natural that we will not know the seating arrangement – I should change the question at some point.)

Let $\gamma = 0.95$ be the probability that the scanner makes the correct identification and $\bar{\gamma} = 1 - \gamma$.

p(someone is correctly declared as the terrorist):

$$
\begin{aligned}
&p(\text{someone is correctly declared as the terrorist})\\
&= \sum_{i=1}^{100} p(\text{someone is correctly declared as the terrorist}|\text{seating plan } i)p(\text{seating plan } i)
\end{aligned}
$$

Seating plan i places the terrorist in seat i. Without loss of generality, we assume that the testing process is that the person in seat 1 is scanned first, and then subsequently the person in seat 2, *etc.* When a declaration of having found a terrorist is made, the process stops. The terrorist is correctly identified when, for seating plan i, he is declared in the i^{th} position. This happens with probability γ^i (since all $i-1$ previous people tested must have been correctly identified as non-terrorists and the i^{th} person correctly identified as a terrorist. Summing these probabilities we obtain

$$
p(\text{someone is correctly declared as the terrorist}) = \frac{1}{100}\sum_{i=1}^{100}\gamma^i = \frac{\gamma}{100}\frac{1-\gamma^{100}}{1-\gamma} = 0.188875 \tag{30.0.15}
$$

p(someone is declared the terrorist):

For the case of $N = 4$ passengers we can represent the probability of making a declaration of a passenger in seat j being the terrorist given seating plan i as the i, j element of the matrix

$$
\begin{pmatrix}
\gamma & \bar{\gamma}^2 & \bar{\gamma}^2\gamma & \bar{\gamma}^2\gamma^2\\
\bar{\gamma} & \gamma^2 & \bar{\gamma}^2\gamma & \bar{\gamma}^2\gamma^2\\
\bar{\gamma} & \gamma\bar{\gamma} & \gamma^3 & \bar{\gamma}^2\gamma^2\\
\bar{\gamma} & \gamma\bar{\gamma} & \gamma^2\bar{\gamma} & \gamma^4
\end{pmatrix}
$$

For example, for the 2nd row, third column, this is the probability that, when the terrorist is really in seat 2, we declare the person in seat 3 as the terrorist. For this to happen, the scanner must have gotten the first person person correct (probability γ), the second person incorrect (probability $\bar{\gamma}$) and the third person incorrect (probability $\bar{\gamma}$). The other entries follow similarly.

The probability then that a person (in an unknown seat and unknown seating arrangement) is declared as a terrorist is the sum of all the entries in the corresponding 100×100 matrix, divided by 100 (the probability of a correct identification is the sum of the diagonal terms of this matrix). This has value 0.999688. One can work out an analytic expression for this by recognising that the upper and lower triangular matrices have banded diagonal structure).

Plugging these two values together we obtain the probability of our neighbour being the terrorist (given that he was the first to be declared to be the terrorist) to be $\frac{0.188875}{0.999688} = 0.18893$. This value is strikingly low – even though the scanner is 95% accurate, there is less than 20% probability that the first person identified by the scanner is actually the terrorist.

1.6 $2+2+1=5$. This is compared to the 7 parameters in a general three binary-variable distribution.

1.7 See `demoClouseau2.m`.

1.8

$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c) \tag{30.0.16}$$

so

$$p(a \cap (b \cup c)) = p(a \cap b) + p(a \cap c) - p(a \cap b \cap a \cap c) = p(a, b) + p(a, c) - p(a, b, c) \tag{30.0.17}$$

1.9

$$p(x|z) = \sum_y p(x, y|z) = \sum_y \frac{p(x, y, z)}{p(z)} = \sum_y \frac{p(x|y, z)p(y, z)}{p(z)} = \sum_y p(x|y, z)p(y|z) \tag{30.0.18}$$

Similar argument for second assertion, beginning with:

$$p(x|z) = \sum_{y,w} p(x, y, w|z) \tag{30.0.19}$$

1.10 The issue here is one of interpretation. Essentially, what the confidence argument is saying is that if we were able to repeat this experiment of visiting Berlin at some point during its wall's lifetime, and use the procedure above to predict the end time, then 95% of the time our predictions will be correct. However, this is an odd thing to want. Intuitively, we are really interested in

$$p(t_{end}|t_{now}) = \frac{p(t_{now}|t_{end})p(t_{end})}{p(t_{now})} \tag{30.0.20}$$

To manipulate this quantity, we must use a prior $p(t_{end})$. Perhaps a less passionate example is that you are given a uniformly generated random positive value from the interval $[1, M]$ with unknown maximum M. Your task is to estimate M based on this single observed value. For example, you observe the value 80. You then say, OK with 95% confidence I think M is between 82 and 3200. I then tell you the true answer. We then repeat this experiment, giving each time a possible range for M. In the limit of many repetitions, the true M each time will indeed lie in the predicted interval in 95% of the cases. However, this is really more a statement about the procedure than an interesting statement about M for any particular experiment. We really want to know something about the Berlin wall, not about the delta-t method. To answer that, we must use a prior.
See `www.cs.ucl.ac.uk/staff/D.Barber/publications/tipping-barber-future.pdf` for a more detailed discussion.

1.11 See `softxor.m`.

1.12 The point here is that if we wish to define a distribution $p(ham|KJ)$ for use with BRMLtoolbox, we are not explicitly given $p(ham = \mathsf{tr}|KJ = \mathsf{fa}) = \gamma$. However, we can figure out this probability using

$$\underbrace{p(ham = \mathsf{tr})}_{\alpha} = \underbrace{p(ham = \mathsf{tr}|KJ = \mathsf{tr})}_{0.9}\underbrace{p(KJ = \mathsf{tr})}_{\beta} + \underbrace{p(ham = \mathsf{tr}|KJ = \mathsf{fa})}_{\gamma}\underbrace{p(KJ = \mathsf{fa})}_{1-\beta} \tag{30.0.21}$$

where $\beta = 1/100000$, $\alpha = 1/2$. Solving this gives

$$\gamma = \frac{\alpha - 0.9\beta}{1 - \beta} \tag{30.0.22}$$

which can then be used to define the table entries required for BRMLtoolbox. See `hamburger.m`.

1.13 See `twoDiceExample.m`.

1.14 There will be $1000, 000/100 = 10000$ people each week that choose $3, 5, 7, 9$. If they win, they will each take home £$1000, 0000/10000 =$ £100. The probability that $3, 5, 7, 9$ (or any four numbers) arises each week is $4!/(9 \times 8 \times 7 \times 6)$. The expected winnings for someone choosing $3, 5, 7, 9$ on any given week is therefore £0.7937. Given that it costs £1 to enter each week, they will lose around 21 pence a week. On the other hand for the least popular number $1, 2, 3, 4$, if this number comes up, each player will win £$10, 000$, so that the average profit per week is £$79.37 - 1 =$ £78.37. The lottery is purely random, but there is 'skill' involved in maximising the amount of money one wins (contrary to some government definitions of a lottery for which skill should play no part in the 'success').

1.15 Probability of getting $0, 1, 2, 3, 4, 5$ correct matches is $44/120, 45/120, 20/120, 10/120, 0, 1/120$. The expected number of correct matches is $(45+2\times 20+3\times 10+5)/120 = 1$. Probability of at least 1 correct is 1 minus the probability of no correct matches $1-44/120 = 76/120$. The clever way to do this is using 'rencontres numbers'. The brute force enumeration approach is in `psychometry.m`.

1.17

1. A nice way to show this is to consider using 'dividers' that separate the 7 friends into their pizza choices. For example $oo|ooo|o|o$ would say that two have chosen pizza A, three pizza B, one pizza C and one pizza D. We are interested in the number of such partitions. There are then $7 + 3 = 10$ positions, each position containing either a friend o or a divider $|$. We can place the dividers in $10 \times 9 \times 8 = 720$ positions. Since the dividers have themselves $3 \times 2 = 6$ arrangements, the number of partitions is $720/6 = 120$.
2. Note that this is not $1/120$ – this would be the probability of the chef uniformly choosing one of the 120 partitions. However, he chooses a pizza at random, so that some partitions are more probable than others. For example, it is less likely that all friends chose pizza A, than two friends choosing pizza A, two friends pizza B, two friends pizza C and one friend pizza D. One needs therefore to compute these probabilities, $p_i, i = 1, \ldots, 120$. Since the chef and friends chose independently, the probability of agreement is given by $\sum_i p_i^2 \approx 0.0184$. See `pizza.m`.