Solutions

Question 1

P(A=y,B=n,C=y) = P(A=y,B=n,C=y,D=y) + P(A=y,B=n,C=y,D=n) = 0.002+0.300=0.302

P("all four RVs have the same value") = P(A=y,B=y,C=y,D=y) + P(A=n,B=n,C=n,D=n) = 0.201+0.003 = 0.204

P(A=y v B=y v C=y v D=y) = 1 - P(A=n,B=n,C=n,D=n) = 1-0.003 = 0.997

P("two variables have value y and two variables have value n") = 0.009+0.007+0.3+0.105+0.005+0.01 = 0.436

P(A=y|B=n) = P(A=y,B=n) / P(B=n) = (0.002+0.007+0.3+0.004)/(0.002+0.007+0.3+0.004+0.01+0.109+0.151+0.003) = 0.53412969

P(C=n,D=n | A=y, B=y) = P(A=y,B=y,C=n,D=n) / P(A=y,B=y) = 0.009/(0.201+0.038+0.034+0.009) = 0.03191489

P("C and D have same value" | B=y) = (P(B=y,C=y,D=y) + P(B=y,C=n,D=n)) / P(B=y) = (0.201+0.010+0.009+0.012)/(0.201+0.038+0.034+0.009+0.010+0.105+0.005+0.012) = 0.56038647

P(C=y | A=y v B=y v D=y) = P(C=y & (A=y v B=y v D=y)) / P(A=y v B=y v D=y)

where
P(A=y v B=y v D=y) = P(A=y) + P(A=n,B=y) + P(A=n,B=n,D=y) =
(0.201+0.038+0.034+0.009+0.002+0.007+0.3+0.004)+(0.010+0.105+0.005+0.012)+(0.010+0.109)
= 0.846
and
P(C=y & (A=y v B=y v D=y)) = P(C=y,A=y) + P(C=y,A=n,B=y) + P(C=y,A=n,B=n,D=y) =
(0.201+0.034+0.002+0.300) + (0.010+0.005) + 0.010 = 0.562
so
P(C=y | A=y v B=y v D=y) = 0.562/0.846 = 0.66430260

Question 2

We need to show that P(A,B)=P(A)*P(B).
From the table, we derive the joint distribution of A and B as

P(A,B)	B=y	B=n
A=y	0.07	0.63
A=n	0.03	0.27

and the marginals of A and B as
P(A=y)=0.7
P(A=n)=0.3
P(B=y)=0.1
P(B=n)=0.9

Multiplying the marginals gives

P(A)*P(B)	B=y	B=n
A=y	0.7*0.1	0.7*0.9
A=n	0.3*0.1	0.3*0.9

which is indeed identical to P(A,B), thus showing independence of A and B.

Question 3

To show that B and C are not independent, we need to show that P(B)*P(C) is not equal to P(B,C). Consider B=y and C=y. P(B=y)=0.075+0.075=0.15, P(C=y)=0.0675+0.7075 = 0.775 and P(B=y,C=y) = 0.0975, but 0.15*0.775=0.11625. Thus, B and C are not independent.

To show that B and C are conditionally independent given A, we need to show that P(B|A)*P(C|A)=P(B,C|A). This is true, as can be seen from the following tables:

	A=y		A=n	
P(B,C|A)	B=y	B=n	B=y	B=n
C=y	0.09	0.81	0.12	0.28
C=n	0.01	0.09	0.18	0.42

P(B|A)	A=y	A=n
B=y	0.1	0.3
B=n	0.9	0.7

P(C|A)	A=y	A=n
C=y	0.9	0.4
C=n	0.1	0.6

	A=y		A=n	
P(B|A)*P(C|A)	B=y	B=n	B=y	B=n
C=y	0.1*0.9	0.9*0.9	0.3*0.4	0.7*0.4
C=n	0.1*0.1	0.9*0.1	0.3*0.6	0.7*0.6

Question 4

a)

ML parameters for full joint

A	B	C	L	P(A,B,C,L)
y	y	y	0	0
y	y	y	1	0
y	y	n	0	2/10
y	y	n	1	0
y	n	y	0	1/10
y	n	y	1	0
y	n	n	0	0
y	n	n	1	1/10
n	y	y	0	1/10
n	y	y	1	0
n	y	n	0	1/10
n	y	n	1	1/10
n	n	y	0	0
n	n	y	1	3/10
n	n	n	0	0
n	n	n	1	0

b) P(A=y)=0.4, P(A=n)=0.6, P(C=y)=P(C=n)=P(B=y)=P(B=n)=P(D=y)=P(D=n)=0.5

c) P(L=0)=P(L=1)=0.5

P(A|L)	A=y	A=n
L=0	3/5	2/5
L=1	1/5	4/5

P(B|L)	B=y	B=n
L=0	4/5	1/5
L=1	1/5	4/5

P(C|L)	C=y	C=n
L=0	2/5	3/5
L=1	2/5	3/5

d) Probabilities of the individual examples for the three different models:

	A	B	C	L	a)	b)	c)
11	y	n	n	0	0	0.05	0.036
12	n	y	y	1	0	0.075	0.032
13	y	y	n	0	0.2	0.05	0.144
14	y	y	y	0	0	0.05	0.096
15	y	y	n	1	0	0.05	0.012
16	n	n	n	1	0	0.075	0.192
17	y	y	n	0	0.2	0.05	0.144
18	y	y	n	1	0	0.05	0.012
19	y	y	y	1	0	0.05	0.008
20	n	y	y	1	0	0.075	0.032

The likelihood of the data is the product of the individual probabilities, which is
For model a) : 0
For model b) : 3.2958984375e-13
For model c) : 1.01445409544602e-12

Model a) is clearly useless: we have seen way too few examples to directly estimate the parameters of the full joint distribution, and therefore have lots of zero entries. For both models b) and c), we have seen enough data to avoid zero entries, and they thus assign a positive likelihood to the test data, which is good. Model c) assigns slightly higher likelihood to the test data than model b), so seems to be a better fit, but given the limited amount of data, we cannot be sure whether this really is the case. Generally, model b) makes the strongest assumptions, which may be the hardest to justify, and c) is probably a good compromise given the extremely limited information we have.