Part1

1.
Using Product Rule, $P(A, B) = P(B) * P(A \mid B)$
$P(X=0,Y=0) = P(X=0)P(Y=0|X=0)= 0.3*0.3=0.09$
$P(X=0,Y=1)= P(X=0)P(Y=1|X=0)=0.3*0.7=0.21$
$P(X=1,Y=0)= P(X=1)P(Y=0|X=1)=0.7*0.8=0.56$
$P(X=1,Y=1)= P(X=1)P(Y=1|X=1)=0.7*0.2=0.14$

2.

X	Y	Z	P(x,y,z)
1	1	1	0.028
0	1	1	0.042
1	0	1	0.224
0	0	1	0.036
0	1	0	0.168
1	0	0	0.336
0	0	0	0.054
1	1	0	0.112

because Z is independent from X given Y, conditional independent rule
$Z \perp X|Y$,
which means
$P(Z|X,Y) = P(Z|Y)$ ----- Equation 1

To show that $P(B,C|A) = P(B|A)P(C|A,B)$ ----- Equation2
We can show $P(B,C|A) /P(B|A) = P(C|A,B)$
LHS = $(P(B,C,A)/ P(A))/(P(B,A)/ P(A))$
$= P(C,A,B)/P(A,B)$
$= P(C| A,B)$ =RHS

we can have $P(A,B,C) = P(A)P(B,C|A)$ --- product rule
$= P(A)P(B|A)P(C|A,B)$ --- From Equation2

=> $P(x,y,z) = P(X)P(Y|X)P(Z|X,Y)$
$= P(x)*P(Y|X)*P(Z|Y)$ ---From Equation1

3.
Using sum rule
$P(z=0)= P(x=0,y=1,z=0) + P(x=1,y=0,z=0) +P(x=0,y=0,z=0) +P(x=1,y=1,z=0)$
$=0.168+0.336+0.054+0.112=0.67$
$P(x=0,z=0)= P(x=0,y=1,z=0) + P(x=0,y=0,z=0) = 0.168+0.054=0.222$

if $P(z|x) == P(z)$
$P(z=0) =0.67$

P(z=0|x=0) =P(z=0,x=0)/P(x=0) = 0.222/0.3= 0.74
Because P(z=0) != P(z=0|x=0)
So X and Z are not independent

4.
P(A,B,C) = P(C)P(A,B|C)   --Using chain rule
P(X = 1, Y = 0, Z = 1) = P(Z=1)*P(X = 1, Y = 0|Z = 1)
P(X = 1, Y = 0|Z = 1) = P(X = 1, Y = 0, Z = 1)/ P(Z=1)
From above table,
P(X = 1, Y = 0, Z = 1)= 0.224
P(Z=1) =1-P(z=0) = 1-0.67 = 0.33
So,
P(X = 1, Y = 0|Z = 1)=0.224/0.33=0.679

(2)
P(A,B,C) =P(C)P(B|C)P(A|B,C) ----Using product rule and chain rule
P(X = 0,Y = 0, Z = 0) =P(Z=0)*P(Y=0|Z=0)*P(X = 0|Y = 0, Z = 0)

P(X = 0|Y = 0, Z = 0) = P(X = 0, Y = 0, Z = 0)/(P(Z=0)*P(Y=0|Z=0))

From above table
P(X = 0,Y = 0, Z = 0) = 0.054
P(Z=0)= 0.67
P(Y=0|Z=0) = P(x=0,y=0,z=0) + P(x=1,y=0,z=0)= 0.054+0.336=0.39
So,
P(X = 0|Y = 0, Z = 0) = 0.054 / (0.67*0.39)=0.207

Part2

Class	Spam	NonSpam
Total	51	149
F1=F	17	96
F1=T	34	53
F2=F	21	63
F2=T	30	86
F3=F	28	98
F3=T	23	51
F4=F	20	90
F4=T	31	59
F5=F	26	99
F5=T	25	50
F6=F	33	79
F6=T	18	70
F7=F	11	74
F7=T	40	75
F8=F	12	97
F8=T	39	52
F9=F	34	113
F9=T	17	36
F10=F	17	106

F10=T	34	43
F11=F	17	62
F11=T	34	87
F12=F	11	99
F12=T	40	50

1.

	Spam	NonSpam
P(Class)	51/200	149/200
P(F1=F|Class)	17/51	96/149
P(F1=T|Class)	34/51	53/149
P(F2=F|Class)	21/51	63/149
P(F2=T|Class)	30/51	86/149
P(F3=F|Class)	28/51	98/149
P(F3=T|Class)	23/51	51/149
P(F4=F|Class)	20/51	90/149
P(F4=T|Class)	31/51	59/149
P(F5=F|Class)	26/51	99/149
P(F5=T|Class)	25/51	50/149
P(F6=F|Class)	33/51	79/149
P(F6=T|Class)	18/51	70/149
P(F7=F|Class)	11/51	74/149
P(F7=T|Class)	40/51	75/149
P(F8=F|Class)	12/51	97/149
P(F8=T|Class)	39/51	52/149
P(F9=F|Class)	34/51	113/149
P(F9=T|Class)	17/51	36/149
P(F10=F|Class)	17/51	106/149
P(F10=T|Class)	34/51	43/149
P(F11=F|Class)	17/51	62/149
P(F11=T|Class)	34/51	87/149
P(F12=F|Class)	11/51	99/149
P(F12=T|Class)	40/51	50/149

2.

$P(S|D_{1.....10})$ = [0.0065, 0.5625, 0.5912, 0.0088, 0.383, 0.5518, 0.0102, 0.1328, 0.8302, 0.0295]

$P(nonSpam|D_{1.....10})$ = [0.9935, 0.4375, 0.4088, 0.9912, 0.617, 0.4482, 0.9898, 0.8672, 0.1698, 0.9705]

0 1 1 0 0 1 0 0 1 0

predicted class type from instance 1 to 10 on unlabelled dataset is:

nonSpam, Spam, Spam,nonSpam,nonSpam,Spam,nonSpam,nonSpam,Spam,nonSpam

3.

In reality, the spam emails have common features. Similarly, the non-spam emails do not have the common features of spam emails. For examples, most ads emails are spam emails. ads emails have common features, such as, “MILLION DOLLARS”, significant amounts of text in CAPS. Furthermore, " MILLION DOLLARS" and significant amount of test in CAPS are not conditional independent. But, in reality, non-spam emails are unlikely to have those features. That's why features are conditional independent for spam email is likely to be invalid assumption.

If the assumption is invalid, then two attributes actually not being independent, for example, if “MILLION DOLLARS" and significant amounts of text in CAPS are dependent, once Million DOLLARS is true, significant is more likely to be true. But if they are independent, the probability of Million DOLLARS is true regardless whether significant amounts of text in CAPS true or false. So, using naive bays algorithm would be underestimates the probability of spam emails, if the assumption is invalid.

Part3

Probability table of meeting variable:

meeting	probability
hasMeeting	0.7
nonMeeting	0.3

Probability table of lecture variable:

lecture	probability
hasLecture	0.6
nonLecture	0.4

Probability table of office variable:

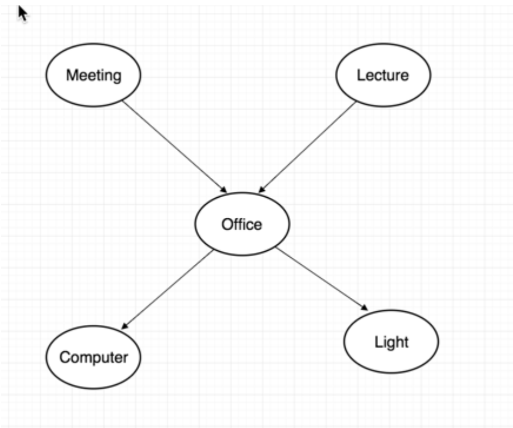
Meeting	Lecture	Office	Probability
HasMeeting	HasLecture	inOffice	P(in|H,H)=0.95
HasMeeting	NonLecture	inOffice	P(in|H,N)=0.75
nonMeeting	hasLecture	inOffice	P(in|N,H)=0.80
nonMeeting	nonLecture	inOffice	P(in|N,N)=0.06
HasMeeting	HasLecture	nonOffice	P(non|H,H)=0.05
hasMeeting	nonLecture	nonOffice	P(non|H,N)=0.25
nonMeeting	hasLecture	nonOffice	P(non|N,H)=0.20
nonMeeting	nonLecture	nonOffice	P(non|N,N)=0.94

Probability table of computer variable:

Computer	Office	Probability
LoggedOn	inOffice	P(L|in)=0.8
NonLogged	inOffice	P(N|in)=0.2
LoggedOn	nonOffice	P(L|non)=0.2
nonLogged	nonOffice	P(N|non)=0.8

Probability table of light variable:

Light	Office	Probability
On	inOffice	P(On|in)=0.5
Off	inOffice	P(Off|in)=0.5
On	nonOffice	P(On|non)=0.02
Off	nonOffice	P(Off|non)=0.98



2.
number of free parameters in BN is 1+1+4+2+2=10
3.
P(lectures=has,meeting=no, office=in, computer=on, light=off)
=P(lectures=has) P(meeting=no) P(office=in|meeting=no, lectures=has)*P(computer=on| office=in)*P(light=off| office=in)
=0.6*0.3*0.8*0.8*0.5=0.0576
4.
P(inOffice) = P(inOffice, H,N)+ P(inOffice, N,H)+ P(inOffice, H,H)+ P(inOffice, N,N)
= P(inOffice|H,N)P(H,N)+ P(inOffice|N,H)P(N,H)+ P(inOffice|H,H)P(H,H)+ P(inOffice|N,N)P(N,N)
= P(inOffice|H,N)P(H)P(N)+ P(inOffice|N,H)P(N)P(H)+ P(inOffice|H,H)P(H)P(H)+ P(inOffice|N,N)P(N)P(N)
=0.75*0.7*0.4+ 0.8*0.3*0.6+0.95*0.7*0.6+0.06*0.3*0.4=0.7602

5.
because computer and light has common cause - Office, Effect becomes independent once common cause is known, computer ⊥Light|Office
P(loggedOn, off|inOffice)=P(loggedOn|inOffice) P(off|inOffice) = 0.8*0.5=0.4

6.
P(Office=true)=0.7602 from question4
P(Office=false)= 1- P(Office=true)=0.2398
P(loggon)= P(loggon, office=true) + P(loggon, office=false)
=P(office=true)* P(loggon|office=true) + P(office=false)* P(loggon|office=false)
=0.7602*0.8+0.2398*0.2=0.65612
P(light=on) = P(light=on, office=true) + P(light=on, office=false)
=P(office=true)P(light=on|office=true)+P(office=false)P(light=on|office=false)
=0.7602 *0.5+0.2398*0.02=0.384896

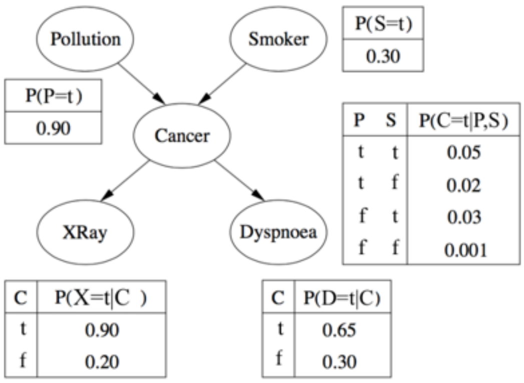
P(light=on,computer=loggon)=
P(light=on,computer=loggon,office=in)+P(light=on,computer=loggon,office=non)
= P(light=on| office=in) P(computer=loggon| office=in)
+ P(light=on| office=non) P(computer=loggon| office=non)
=0.5*0.8+0.02*0.2=0.404

P(light=on|computer=loggon)= P(light=on,computer=loggon)/ P(computer=loggon)
=0.404/0.65612=0.616

P(light=on|computer=loggon)> P(light=on)
0.616>0.384896

So, this will increase the student's belief that Rachel light is on, given that she is logged on

Part4



1(i)
Evidence variable: XRay
Hidden variables: Smoker, cancer, Dyspnoea
Query variables: Pollution
P(P=t|X=t)= P(P=t,X=t)/P(X=t)
P(S) Join P(P) ==> P(S,P)
P(C|S,P) join P(S,P) ==> P(C,S,P)

P(C,S,P) elimination ==> P(C,P)
P(X|C) Join P(C,P) Join P(D|C)= P(P,X,C,D)
P(P,X,C,D) elimination ==> P(P,X)

P(P)

t	0.9
f	0.1

P(S)

t	0.3
f	0.7

Join

P(S,P)

s	p	0.27
s	-p	0.03
-s	p	0.63
-s	-p	0.07

P(C|S,P)

c	s	p	0.05
c	s	-p	0.03
c	-s	p	0.02
c	-s	-p	0.001
-c	s	p	0.95
-c	s	-p	0.97
-c	-s	p	0.98
-c	-s	-p	0.999

Join

P(C,S,P)

c	s	p	0.0135
c	s	-p	0.0009
c	-s	p	0.0126
c	-s	-p	0.00007
-c	s	p	0.2565
-c	s	-p	0.0291
-c	-s	p	0.6174
-c	-s	-p	0.06993

P(C,S,P)

c	s	p	0.0135
c	s	-p	0.0009
c	-s	p	0.0126
c	-s	-p	0.00007
-c	s	p	0.2565
-c	s	-p	0.0291
-c	-s	p	0.6174
-c	-s	-p	0.06993

Elimation

P(C,P)

c	p	0.0261
c	-p	0.00097
-c	p	0.8739
-c	-p	0.09903

P(X|C)

c	x	0.9
c	-x	0.1
-c	x	0.2
-c	-x	0.8

Join

P(P,X,C)
Show as below

P(P,X,C)

p	x	c	cp,cx=0.0261*0.9=0.02349
p	x	-c	-cp,-cx=0.8739*0.2=0.17478
p	-x	c	cp,c-x=0.0261*0.9 =0.02349
p	-x	-c	-cp,-c-x=0.8739*0.8=0.69912
-p	x	c	c-p,cx=0.00097*0.9=0.000873
-p	x	-c	-c-p,-cx=0.09903*0.2=0.019806
-p	-x	c	c-p,c-x=0.00097*0.1=0.000097
-p	-x	-c	-c-p,-c-x=0.09903*0.8=0.079224

P(P,X,C) ----elimation to ---> P(P,X)

P(P,X)

P	X	0.19827
P	-X	0.72261
-P	X	0.020679
-P	-X	0.079321

(iii)
P(P=true,X=true) = 0.19827
P(X=true) = P(P=true,X=true)+ P(P=false,X=true)=0.19827+0.020679=0.218949
P(P=true|X=true) = P(P=true,X=true)/P(X=true) = 0.9056

2.
Pollution and smoker are independent of each other.
Because from the original bayesian network, there is no path between pollution and smoker and they do not have common parent node.

XRay and Dysponea are conditional independent, given Cancer
Because Effects becomes independent once common cause is known. There is no path between XRay and Dysponea. In addition, they have common parent node.

XRay and Smoker are conditional independent, given Cancer
Because Smoker is indirect cause of Xray, given Cancer.
there is no direct path between XRay and Smoker. In addition, Cancer is middle node of the flow between Somker and XRay

XRay and Pollution are conditional independent, given Cancer.
Because Pollution is indirect cause of Xray, given Cancer.
there is no direct path between XRay and Pollution. In addition, Cancer is middle node of the flow between Pollution and XRay

Dyspnoea and Smoker are conditional independent, given Cancer
Because Smoker is indirect cause of Dyspnoea, given Cancer.
Because there is no direct path between Dyspnoea and Smoker. In addition, Cancer is middle node of the flow between Somker and Dyspnoea

Dyspnoea and Pollution are conditional independent, given Cancer
Because Pollution is indirect cause of Dyspnoea, given Cancer.
there is no direct path between Dyspnoea and Pollution. In addition, Cancer is middle node of the flow between Pollution and Dyspnoea

3.
order as <Xray, Dyspnoea, Cancer, Smoker, Pollution>
Step1: Add node Xray
Step2: Add node Dyspnoea
-P(D|X) ==P(D) ? No,X->D

Step3: Add node Cancer
--P(C|D,X)=P(C)? No
--P(C|D,X)=P(C|X)? No
--P(C|D,X)=P(C|D)? No, so D->C and X->C
Step4: Add node Smoker
-P(S|D,X,C) = P(S)? No
-P(S|D,X,C) = P(S|C)? Yes, so C->S, no link from D or X to S
Step5: Add node Pollution
-P(P|D,X,C,S) = P(P) ? No
-P(P|D,X,C,S) = P(P|C) ? No
-P(P|D,X,C,S) = P(P|S) ? No
-P(P|D,X,C,S) = P(P|C,S) ? Yes, so C->P, S->P, no other links