NAME: KALPANA PRATAPANENI(A20448916)
COLLABORATOR(S): Tushar Nitave

CS 583 – Assignment 6

1. **Single-variable – Binomial**

You are given a single binary variable: X with domain {T, F}. $P(X = T) = \theta$. Here are the counts for a dataset D .

X	Counts
T	10
F	20

a. What is the MLE estimate for $P(X)$?

ANS:

X	**MLE**
T	10/30
F	20/30

b. Assuming a uniform prior (i.e., assuming $p(\theta) = 1$)

i. What is $P(X_{next}|D)$?

X	$P(X_{next}|D)$
T	11/32
F	21/32

ii. What is $p(\theta|D)$?

P(θ |D)=Beta(11, 21)

c. Assuming $p(\theta) \tilde{} Beta(2,3)$

i. What is $P(X_{next}|D)$?

X	**MLE**
T	12/35

F	23/35

ii. What is $p(\theta|D)$?
P(θ | D)=Beta(12, 23)

2. **Singe-variable – Multinomial**

You are given a single multinomial variable: X with domain {R, G, B}. $P(X) = \theta = \langle \theta_1, \theta_2, \theta_3 \rangle$.
Here are the counts for a dataset D.

X	Counts
R	10
G	20

(Note that if a count is zero for a case, it is not listed. In this case, X=B has zero count.)

a. What is the MLE estimate for $P(X)$?

X	**MLE**
R	10/30 = 0.33
G	20/30 = 0.67
B	0/30 = 0

b. Assuming a uniform prior (i.e., assuming $p(\theta) = 1$)
   i. What is $P(X_{next}|D)$?

X	$P(X_{next}|D)$
R	11/33 = 0.33
G	21/33 = 0.64
B	1/33 = 0.03

   ii. What is $p(\theta|D)$?

P(θ | D)=Beta(11, 21, 1)

c. Assuming $p(\theta) \tilde{} Dir(2,3,4)$

i. What is $P(X_{next} | D)$?

X	$P(X_{next} \| D)$
R	12/39
G	23/39
B	4/39

ii. What is $p(\theta|D)$?

P(θ | D)=Beta(12, 23, 4)

3. **Multiple variables**

We have three variables: X, Y, and Z. X and Z are binary with domain {T, F} and Y has three possible values: {R, G, B}. The Bayesian network has the following structure: $X \rightarrow Y \rightarrow Z$. Here are the counts for a dataset D. If a count is zero, it is not listed.

X	Y	Z	Counts
T	R	T	10
T	R	F	20
T	B	T	30
F	R	F	40
F	B	T	50

Note that we need to estimate $P(X)$, $P(Y | X)$, and $P(Z | Y)$ for this network.

a. What are the MLE estimates?

X	**Counts**	**MLE**
T	60	60/150
F	90	90/150

Y	**Y|X=T**	**MLE**
R	30	30/60
G	0	0
B	30	30/60

Y	**Y|X=F**	**MLE**
R	40	40/90
G	0	0
B	50	50/90

Z	**Z|Y=R**	**MLE**
T	10	10/70
F	60	60/70

Z	**Z|Y=G**	**MLE**
T	0	0
F	0	0

Z	**Z|Y=B**	**MLE**
T	80	80/80
F	0	0/80

b. Assuming a uniform prior and K2 approach to Bayesian estimation, what are the predictive probabilities for next X, Y|X, and Z|Y?

X	**Counts**	**$P(X_{next}\|D)$**
T	60	61/152
F	90	91/152

Y	**Y|X=T**	**$P(Y_{next}\|D)$**
R	30	(30+1)/63 = 31/63
G	0	(0+1)/63 = 1/63
B	30	(30+1)/63 = 31/63

Y	**Y|X=F**	**$P(Y_{next}\|D)$**
R	40	(40+1)/93 = 41/93
G	0	(0+1)/93 = 1/93
B	50	(50+1)/93 = 51/93

Z	**Z|Y=R**	**$P(Z_{next}\|D)$**
T	10	(10+1)/72 = 11/72
F	60	(60+1)/72 = 61/72

Z	**Z|Y=G**	**$P(Z_{next}\|D)$**

T	0	(0+1)/2 = 1/2
F	0	(0+1)/2 = 1/2

Z	**Z|Y=B**	**P(Z_{next}|D)**
T	80	(80+1)/82 = 81/82
F	0	(0+1)/82 = 1/82

c. Assuming a |D'| = 12, and P' is uniform, and a BDe approach to estimation, what are the predictive probabilities for next X, Y|X, and Z|Y?

X	**Counts**	**Imaginary Counts**	**P(X_{next}|D)**
T	60	6	66/162
F	90	6	96/162

Y	**Y|X=T**	**Imaginary Counts**	**P(Y_{next}|D)**
R	30	2	(30+2)/66 = 32/66
G	0	2	(0+2)/66 = 2/66
B	30	2	(30+2)/66 = 32/66

Y	**Y|X=F**	**Imaginary Counts**	**P(Y_{next}|D)**
R	40	2	(40+2)/96 = 42/96
G	0	2	(0+2)/96 = 2/96
B	50	2	(50+2)/96 = 52/96

Z	**Z|Y=R**	**Imaginary Counts**	**P(Z_{next}|D)**

T	10	2	(10+2)/74 = 12/74
F	60	2	(60+2)/74 = 62/74

Z	**Z|Y=G**	**Imaginary Counts**	**P(Z_{next}|D)**
T	0	2	(0+2)/4 = 2/4
F	0	2	(0+2)/4 = 2/4

Z	**Z|Y=B**	**Imaginary Counts**	**P(Z_{next}|D)**
T	80	2	(80+2)/84 = 82/84
F	0	2	(0+2)/84 = 2/84

4. **Missing data**

We have three variables: X, Y, and Z. All variables are binary with domain {0, 1}. The Bayesian network has the following structure: $X \rightarrow Y \rightarrow Z$. Assume we initialize these distributions as follows:

$$P(X) = \langle 0.4,\ 0.6 \rangle$$

$$P(Y \mid X = 0) = \langle 0.8,\ 0.2 \rangle$$

$$P(Y \mid X = 1) = \langle 0.2,\ 0.8 \rangle$$

$$P(Z \mid Y = 0) = \langle 0.7,\ 0.3 \rangle$$

$$P(Z \mid Y = 1) = \langle 0.3,\ 0.7 \rangle$$

Here is a sample dataset, where each row indicates an instance and? indicates a missing value for that variable for that instance.

X	Y	Z
0	0	?
0	?	1
?	1	0

?	?	0
?	1	?

a. Perform the expectation step of EM and calculate the necessary counts, i.e., counts(X), counts(Y, X), and counts(Z, Y).

<u>COUNTS(X):</u>

Row 3:

? 1 0:

P(X |Y=1, Z=0) = 1/Z * P(X) * P(Y/X=0) * P(Z=0 | Y=1)

P(X=0 |Y=1, Z=0) = 1/Z * 0.4*0.2*0.3 = 0.024/(0.024+0.144) = 0.143

P(X=1 |Y=1, Z=0) = 1/Z * 0.6*0.8*0.3 = 0.144/(0.024+0.144) = 0.857

Row 4:

? ? 0:

P(X|Z=0) = 1/Z * P(X)*P(Y|X)*P(Z=0|Y)

Eliminate Y:

X	Y	P(X|Y)P(Y|Z=0)
0	0	0.8*0.7 = 0.56
0	1	0.06
1	0	0.14
1	1	0.24

Tow(X) = <0.62, 0.38>

P(X=0|Z=0) = 1/Z * 0.4*0.62 = 0.248/(0.248+0.228) = 0.521

P(X=1|Z=0) =1/Z * 0.6*0.38 = 0.228/(0.248+0.228) = 0.479

Row 5:

? 1 ?:

P(X | Y=1) = 1/Z * P(X) * P(X|Y=1) * P(Y=1|Z)

Eliminating variable Z gives a vector of 1. So,

P(X | Y=1) = 1/Z * P(X) * P(X|Y=1)

P(X=0 | Y=1) = 0.4 * 0.2 = 0.08/(0.08 +0.48) = 0.143

P(X=1 | Y=1) = 0.6 * 0.8 = 0.48/(0.08 +0.48) = 0.857

Counts[X = 0] = 2 + 0.143 + 0.521 + 0.143 = 2.807

Counts[X = 1] = 0 + 0.857 + 0.479 + 0.857 = 2.193

<u>COUNTS[Y, X]:</u>

0 ? 1 - 2nd row

counts[Y=0, X=1] = 0

counts[Y=1, X=1] = 0 Since X = 0

Counts[Y=0, X = 0] = P(Y=0, X=0 | Z = 1)

= 1/Z * P(X = 0) * P(Y=0|X=0) * P(Z=1| Y=0)

= 0.4 * 0.8 * 0.3 = 0.096/(0.096+0.056) = 0.632

Counts[Y=1, X = 0] = P(Y=0, X=0 | Z = 1) = 0.4 * 0.2 * 0.7 = 0.056/(0.096+0.056) = 0.368

? 1 0 - 3rd row

counts[Y=0, X=0] = 0

counts[Y=0, X=1] = 0 Since Y = 1

counts[Y = 1, X = 0] = P(Y=1, X = 0 | Z=0)

= 1/Z * P(X=0) * P(Y=1|X=0) * P(Z=0|Y=1)

= 0.4 * 0.2 * 0.3

= 0.024/(0.024 + 0.144) = 0.143

counts[Y = 1, X = 1] = P(Y=1, X = 1 | Z=0)

= 1/Z * P(X=1) * P(Y=1|X=1) * P(Z=0|Y=1)

= 0.6 * 0.8 * 0.3

= 0.144/(0.024 + 0.144) = 0.857

? ? 0 - 4th row

Counts[Y=0, X=0] = P(Y, X | Z=0) = 1/Z * P(X=0) * P(Y=0|X=0) * P(Z=0|Y=0)

= 0.4 * 0.8 * 0.7 = 0.224/(0.224 + 0.024+0.084+0.144) = 0.471

Counts[Y=1, X=0] = P(Y, X | Z=0) = 1/Z * P(X=0) * P(Y=1|X=0) * P(Z=0|Y=1)

= 0.4 * 0.2 * 0.3 = 0.024/(0.224 + 0.024+0.084+0.144) = 0.050

Counts[Y=0, X=1] = P(Y, X | Z=0) = 1/Z * P(X=1) * P(Y=0|X=1) * P(Z=0|Y=0)

= 0.6 * 0.2 * 0.7 = 0.084/(0.224 + 0.024+0.084+0.144) = 0.176

Counts[Y=1, X=1] = P(Y, X | Z=0) = 1/Z * P(X=1) * P(Y=1|X=1) * P(Z=0|Y=1)

= 0.6 * 0.8 * 0.3 = 0.144/(0.224 + 0.024+0.084+0.144) = 0.303

? 1 ? - 5th row

Counts[Y=0, X=0] = 0

Counts[Y=0, X=1] = 0 Since Y=1

Counts[Y=1, X=0] = P(X=0) P(Y=1|X=0) = 0.4 * 0.2 = 0.08/(0.08+0.48) = 0.143

Counts[Y=1, X=1] = P(X=1) P(Y=1|X=1) = 0.6 * 0.9 = 0.48/(0.08+0.48) = 0.857

Finally,

Counts[Y=0, X=0] = 1+0.632+0+0.471+0 = 2.103

Counts[Y=0, X=1] = 0+0+0+0.176+0 = 0.176

Counts[Y=1, X=0] = 0+0.368+0.143+0.050+0.143 = 0.704

Counts[Y=1, X=1] = 0+0+0.857+0.303+0.857 = 2.017

<u>COUNTS[Z, Y]:</u>

0 0 ? - 1st row

Counts[Z=0, Y=1] = 0

Counts[Z=1, Y=1] = 0

Counts[Z, Y= 0] = 1/Z * P(Y=0/X=0) * P(Z/Y=0)

= 1/Z * <0.8 * 0.7, 0.8 * 0.3> = <0.56/(0.56+0.24), 0.24/(0.56+0.24)>

= <0.7, 0.3>

0 ? 1 - 2nd row:

Counts[Z=0, Y=0] = 0

Counts[Z=0, Y=1] = 0 Since Z=1

Counts[Z=1, Y] = P(Z=1 | X = 0)

= 1/Z * <0.4*0.8*0.3, 0.4*0.2*0.7>

= <0.096/(0.096+0.056), 0.056/(0.096+0.056)>

= <0.632, 0.368>

? ? 0 - 4th row

Counts[Z=1, Y=0] = 0, Counts[Z=1, Y=1] = 0 Since Z = 0

Counts[Z=0, Y] = <0.308/(0.308+0.168), 0.168/(0.308+0.168)>

= <0.647, 0.353>

? 1 ? - 5th row

Counts[Z=0, Y=0] = 0

Counts[Z=1, Y=0] = 0 Since Y=1

Counts<Z, Y=1> = 1/Z * P(Y=1/X), P(Z=1)

<0.168/(0.168 + 0.392), 0.392/(0.168 + 0.392)> = <0.3, 0.7>

Counts[Z=0, Y=0] = 0.7 + 0 + 0 + 0.647 + 0 **= 1.347**

Counts[Z=0, Y=1] = 0 + 0 + 1 + 0.353 + 0.3 **= 1.653**

Counts[Z=1, Y=0] = 0.3 + 0.632 + 0 + 0 + 0 **= 0.932**

Counts[Z=1, Y=1] = 0 + 0.368 + 0 + 0 + 0.7 **= 1.068**

b. Perform the maximization step of EM and reestimate P(X), P(Y|X), and P(Z|Y) using the counts from part a.

<u>P(X):</u>

P(X) = <2.807/(2.807+2.193), 2.193/(2.807+2.193)>

P(X) = <0.56, 0.44>

<u>P(Y|X):</u>

P(Y|X=0) = <#(Y=0, X=0)/#(X=0), #(Y=1, X=0)/#(X=0)>

P(Y|X=1) = <#(Y=0, X=1)/#(X=1), #(Y=1, X=1)/#(X=1)>

$P(Y|X=0) = \langle 2.103/2.807, 0.704/2.807\rangle = \langle 0.749, 0.251\rangle$

$P(Y|X=1) = \langle 0.176/2.193, 2.017/2.193\rangle = \langle 0.08, 0.92\rangle$

P(Z|Y):

$P(Z|Y=0) = \langle 1.347/(2.103+0.176), 0.932/(2.103+0.176)\rangle = \langle 0.59, 0.41\rangle$

$P(Z|Y=1) = \langle 1.653/(0.704+2.017), 1.068/(0.704+2.017)\rangle = \langle 0.61, 0.39\rangle$