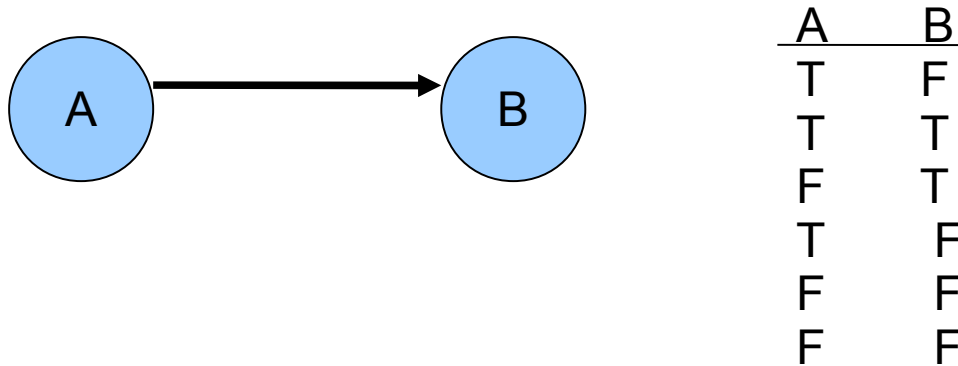Question-1:
Let us consider the network structure shown below. Calculate maximum likelihood (ML),
maximum a posteriori (MAP) and Bayesian estimates of the parameter for Pr(b |a). 8 samples
are available (see the table). We also know that the prior distribution of Pr(b | a) is Beta(3,3).
Hint: MAP takes the point of maximum of posterior distribution, however Bayesian estimate
takes the expected (mean) value of posterior distribution.

A	B
T	F
T	T
F	T
T	F
F	F
F	F

Question-2: Our data is n observations with 3 explanatory variable and one response variable.
The model is that

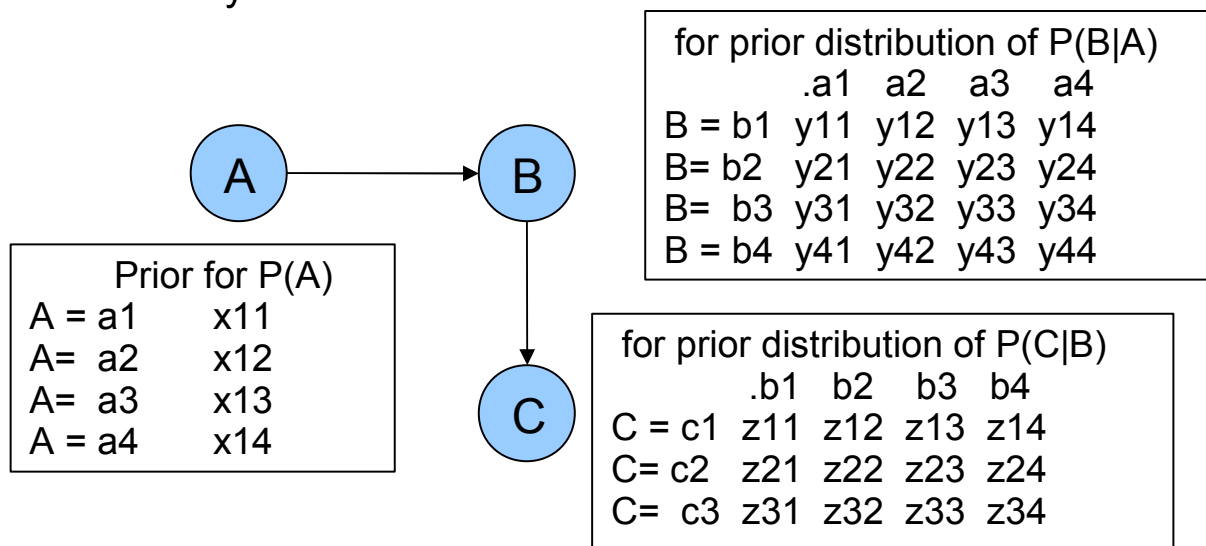$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$$

what is the MLE of β_1, β_2 and β_3 when ϵ_i are independent with mean= 0 and normal distribution.
The standard deviation for noise ϵ_i is σ.

Question-3

Given a Bayesian network with nodes A, B & C. Where A can take 4 values {a1,a2, a3, a4} and B can take 4 values {b1, b2, b3, b4} and C can take 3 values {c1, c2, c3}. The parameters for prior probability distributions have been mentioned. We are given 3 data-sets. What will be the parameters of posterior probability after learning over D1 and D2 for P(B|C) and P(C|A). What will be the final probability after learning for D1, D2 & D3 for P(B|C) and P(C|A).
Hint : They will follow Drichlet distribution.

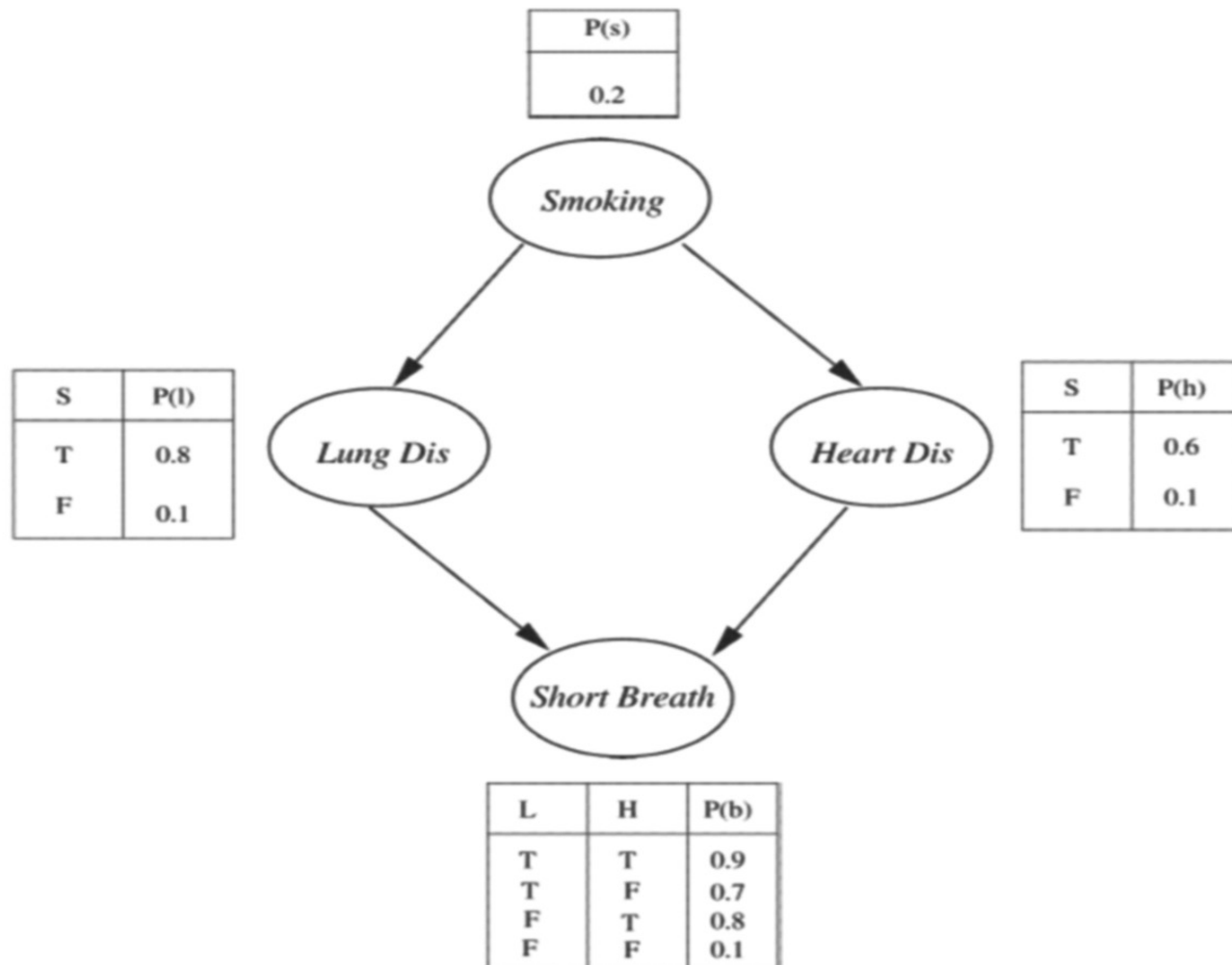for prior distribution of P(B|A)

	.a1	a2	a3	a4
B = b1	y11	y12	y13	y14
B= b2	y21	y22	y23	y24
B= b3	y31	y32	y33	y34
B = b4	y41	y42	y43	y44

A → B

B → C

Prior for P(A)

A = a1	x11
A= a2	x12
A= a3	x13
A = a4	x14

for prior distribution of P(C|B)

	.b1	b2	b3	b4
C = c1	z11	z12	z13	z14
C= c2	z21	z22	z23	z24
C= c3	z31	z32	z33	z34

Data D1:

A	B	C
.a1	b1	c1
.a2	b2	c2
.a2	b2	c3
.a3	b4	c2
.a4	b2	c1
.a3	b1	c2
.a2	b2	c2
.a1	b1	c1

Data D2

A	B	C
.a2	b1	c1
.a1	b4	c2
.a2	b2	c3
.a3	b4	c3
.a1	b2	c1
.a2	b2	c2
.a2	b2	c3
.a4	b4	c1

Data D3

A	B	C
.a1	b4	c3
.a3	b1	c2
.a2	b2	c3
.a1	b3	c2
.a4	b2	c1
.a3	b3	c1
.a1	b4	c2
.a2	b1	c1

Question -4: Given the following Bayesian network. Perform Markov chain Monte Carlo using Gibbs sampling to find the probability for smoking for a person who has short Breadth i.e. find P(smoking | shortBreadth) and P (¬ smoking | shortBreadth) using MCMC based Gibbs sampling. Start with all the nodes as false and go till 6 rounds start counting after 2nd round. What would be you answer if you go till 10 rounds start counting from 4th round. Your random numbers are : 0.13, 0.51, 0.6, 0.7, 0.3, 0.56, 0.67, 0.65, 0.47, 0.34, 0.87, 0.85, 0.43, 0.56, 0.76, 0.67, 0.83, 0.2, 0.3, 0.4, 0.54, 0.39, 0.45, 0.78, 0.8.

P(s)
0.2

Smoking

S	P(l)
T	0.8
F	0.1

Lung Dis

Heart Dis

S	P(h)
T	0.6
F	0.1

Short Breath

L	H	P(b)
T	T	0.9
T	F	0.7
F	T	0.8
F	F	0.1

Question-5:
Consider a Bayesian network with structure $X \leftarrow Y \rightarrow Z$ where each of X, Y , and Z take values from the finite sets X: (x1, x2) , Y:(y1, y2) and Z:(z1, z2) respectively. This network has the following parameters where

. π_y = P(Y = y1)

$R_{x1,y1}$ = P(X = x1 | Y = y1)   & $R_{x1,y2}$ = P(X = x1 | Y = y2)

$L_{z1,y1}$ = P(Z = z1 | Y = y1)   & $R_{z1,y2}$ = P(Z = z1 | Y = y2)

Give both the hard and soft EM algorithms for this model, given that we have a table of observations with missing values

```
   X    Y    Z
.  x1   y1   z1
.  x1   y1   z1
.  ?    y2   z1
.  x1   y2   z1
.  x2   y1   ?
.  x1   y2   z2
.  x1   y1   z1
.  x2   y2   z2
.  x1   ?    z1
```

Hint: in class we learned soft EM for Bayesnet where we guess probabilities of missing values for both cases (y1 and y2). However for hard EM we will take maximum value guess from prior distribution Beta(a, b)