6.3

(a)

$$P(Y=1|X) = \sum_{Z\in\{0,1\}^n} P(Y=1,Z|X)$$

On the right side:

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X) = \sum_{Z\in\{0,1\}^n} P(Y=1|Z,X)P(Z|X)$$

……………………………………………………………………………………product rule

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X) = \sum_{Z\in\{0,1\}^n} P(Y=1|Z)P(Z|X)$$

……………………………………………………………………………d-separation #1

Expand $X = X_1, X_2, X_3 \ldots. X_n$, $Z = Z_1, Z_2, Z_3 \ldots. Z_n$, we can get

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X)$$

$$= \sum_{Z_1,Z_2\ldots Z_n} P(Y=1|Z_1,Z_2\ldots Z_n)P(Z_1,Z_2\ldots Z_n|X_1,X_2\ldots X_n)$$

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X)$$

$$= \sum_{Z_1,Z_2\ldots Z_n} P(Y=1|Z_1,Z_2\ldots Z_n)\prod_{i=1}^{n} P(Z_i|X_1,X_2\ldots X_n)$$

……………………………………………………………………………d-separation #3

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X) = \sum_{Z_1,Z_2\ldots Z_n} P(Y=1|Z_1,Z_2\ldots Z_n)\prod_{i=1}^{n} P(Z_i|X_i)$$

……………………………………………………………………………d-separation #3

Where

$$P(Y=1|Z_1,Z_2\ldots Z_n)=\begin{cases}1 & Z_i=1\ for\ any\ i\\ 0 & Z_i=0\ \ for\ all\ i\end{cases}$$

$$P(Z_i=1|X_i)=\begin{cases}0 & X_i=0\\ p_i & X_i=1\end{cases}$$

Group Z's situation, namely $Z_i=1\ for\ any\ i$ and $Z_i=0\ \ for\ all\ i$

Then we can get

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X)$$

$$=\sum_{Z_i=1\ for\ any\ i} P(Y=1|Z_1,Z_2\ldots Z_n)\prod_{i=1}^{n} P(Z_i|X_i)$$

$$+\sum_{Z_i=0\ \ for\ all\ i} P(Y=1|Z_1,Z_2\ldots Z_n)\prod_{i=1}^{n} P(Z_i|X_i)$$

For the first term $P(Y=1|Z_1,Z_2\ldots Z_n)=1$, while for the second term $P(Y=1|Z_1,Z_2\ldots Z_n)=0$, so we can get

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X)=\sum_{Z_i=1\ for\ any\ i}\prod_{i=1}^{n} P(Z_i|X_i)$$

Apparently,

$$\sum_{Z_1,Z_2\ldots Z_n}\prod_{i=1}^{n} P(Z_i|X_i)=\sum_{Z_1,Z_2\ldots Z_n} P(Z_1,Z_2\ldots Z_n|X_1,X_2\ldots X_n)=1$$

And

$$\sum_{Z_1,Z_2\ldots Z_n}\prod_{i=1}^{n} P(Z_i|X_i)$$

$$=\sum_{Z_i=1\ for\ any\ i}\prod_{i=1}^{n} P(Z_i|X_i)+\sum_{Z_i=0\ \ for\ all\ i}\prod_{i=1}^{n} P(Z_i|X_i)$$

So we can get

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X) = \sum_{Z_i=1\ for\ any\ i} \prod_{i=1}^{n} P(Z_i|X_i)$$

$$= 1 - \sum_{Z_i=0\ \ for\ all\ i} \prod_{i=1}^{n} P(Z_i|X_i)$$

$$= 1 - \prod_{i=1,\ \ Z_i=0\ \ for\ all\ i}^{n} P(Z_i|X_i) = 1 - \prod_{i=1,}^{n} P(Z_i = 0|X_i)$$

Where

$$P(Z_i = 1|X_i) = \begin{cases} 0 & X_i = 0 \\ p_i & X_i = 1 \end{cases}$$

So we can get

$$P(Z_i = 0|X_i) = \begin{cases} 1 & X_i = 0 \\ 1 - p_i & X_i = 1 \end{cases}$$

We can conclude this as

$$P(Z_i = 0|X_i) = (1 - p_i)^{x_i}$$

Take it into $\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X)$, we can get

$$\sum_{Z\in\{0,1\}^n} P(Y=1,Z|X) = 1 - \prod_{i=1,}^{n} P(Z_i = 0|X_i) = 1 - \prod_{i=1,}^{n} (1-p_i)^{x_i}$$

$$= P(Y=1|X) = left\ side$$

Thus, we know that:

$$P(Y=1|X) = \sum_{Z\in\{0,1\}^n} P(Y=1,Z|X)$$

(b)

$$L = \frac{1}{T}\sum_{t=1}^{T} logP(Y = y^{(t)}|X = \vec{x}^{(t)})$$

$$P(Z_i = 1, X_i = 1|X = x, Y = y) = I(x_i, 1)P(Z_i = 1|X = x, Y = y)$$

Where

$$P(Z_i = 1|Y = y, X = x) = \frac{P(Y = y|Z_i = 1, X = x)P(Z_i = 1|X = x)}{P(Y = y|X = x)}$$

$$P(Z_i = 1|Y = y, X = x) = \frac{P(Y = y|Z_i = 1, X = x)P(Z_i = 1|X_i = x_i)}{P(Y = y|X = x)}$$

……………………………………………………………………………………Bayes rule

When y=0, we have

$$P(Y = y|Z_i = 1, X = x) = P(Y = 0|Y = 1, X = x) = 0$$

$$P(Y = y|X = x) = \prod_{i=1,}^{n}(1 - p_i)^{x_i} \neq 0$$

So

$$P(Z_i = 1|Y = y, X = x) = 0$$

When y=1, we have

$$P(Y = y|Z_i = 1, X = x) = P(Y = 0|Y = 1, X = x) = 1$$

$$P(Z_i = 1|X_i = x_i) = \begin{cases} 0 & x_i = 0 \\ p_i & x_i = 1 \end{cases} \quad = x_i p_i$$

$$P(Y = y|X = x) = 1 - \prod_{i=1,}^{n}(1 - p_i)^{x_i}$$

So we can get

$$P(Z_i = 1|Y = y, X = x) = \frac{x_i p_i}{1 - \prod_{i=1,}^{n}(1-p_i)^{x_i}} = \frac{y x_i p_i}{1 - \prod_{i=1,}^{n}(1-p_i)^{x_i}}$$

Since when y=0 also fit this equation, we can conclude that

$$P(Z_i = 1|Y = y, X = x) = \frac{y x_i p_i}{1 - \prod_{i=1,}^{n}(1-p_i)^{x_i}} = \frac{y x_i p_i}{1 - \prod_j (1-p_i)^{x_j}}$$

$$P(Z_i = 1, X_i = 1|X = x, Y = y) = I(x_i, 1)\frac{y x_i p_i}{1 - \prod_j (1-p_i)^{x_j}}$$

When $x_i = 0$

$$P(Z_i = 1, X_i = 1|X = x, Y = y) = 0$$

When $x_i = 1$

$$P(Z_i = 1, X_i = 1|X = x, Y = y) = \frac{y x_i p_i}{1 - \prod_j (1-p_i)^{x_j}}$$

And $x_i = 0$ also fit this equation, so we can get

$$P(Z_i = 1, X_i = 1|X = x, Y = y) = \frac{y x_i p_i}{1 - \prod_j (1-p_i)^{x_j}}$$

(c)

According the definition:

$$P(Z_i = 1|X_i = 1) = p_i$$

In another way, according the EM algorithm formula:

$$P(X_i = x) <- \frac{1}{T}\sum_t P(X_i = x|V_t = v_t)$$

$$P(X_i = x|pa_i = \pi) <- \frac{\sum_t P(X_i = x, pa_i = \pi|V_t = v_t)}{\sum_t P(pa_i = \pi|V_t = v_t)}$$

We can get:

$$P(Z_i = 1|X_i = 1) < -\frac{\sum_t P(Z_i = 1, X_i = 1|X = x^{(t)}, Y = y^{(t)})}{\sum_t P(X_i = 1|X = x^{(t)}, Y = y^{(t)})}$$

$$P(Z_i = 1|X_i = 1) < -\frac{\sum_t P(Z_i = 1, X_i = 1|X = x^{(t)}, Y = y^{(t)})}{\sum_t P(X_i = 1|X = x^{(t)}, Y = y^{(t)})}$$

Where

$$\sum_t P\left(X_i = 1\middle|X = x^{(t)}, Y = y^{(t)}\right) = \sum_t I({x_i}^{(t)}, 1) = T_i$$

So we can get:

$$P(Z_i = 1|X_i = 1) < -\frac{1}{T_i}\sum_t P(Z_i = 1, X_i = 1|X = x^{(t)}, Y = y^{(t)})$$

$$p_i < -\frac{1}{T_i}\sum_t P(Z_i = 1, X_i = 1|X = x^{(t)}, Y = y^{(t)})$$

(d)

$$P(Y = 1|X) = 1 - \prod_{i=1,}^{n}(1 - p_i)^{x_i}$$

$$P(Y = 0|X) = \prod_{i=1,}^{n}(1 - p_i)^{x_i}$$

We can conclude that:

$$P(Y = y|X) = \left(1 - \prod_{i=1,}^{n}(1 - p_i)^{x_i}\right)^y \left(\prod_{i=1,}^{n}(1 - p_i)^{x_i}\right)^{1-y}$$

Thus,

$$L = \frac{1}{T}\sum_{t=1}^{T} logP(Y = y^{(t)}|X = \vec{x}^{(t)})$$

$$= \frac{1}{T}\sum_{t=1}^{T} \log\left(\left(1 - \prod_{i=1,}^{n}(1-p_i)^{x_i}\right)^{y} \left(\prod_{i=1,}^{n}(1-p_i)^{x_i}\right)^{1-y}\right)$$

$$= \frac{1}{T}\sum_{t=1}^{T}\left(y^{(t)} * \log\left(1 - \prod_{i=1,}^{n}(1-p_i)^{x_i{}^{(t)}}\right) + (1-y^{(t)}) * \log\left(\prod_{i=1,}^{n}(1-p_i)^{x_i{}^{(t)}}\right)\right)$$

In conclusion, with the result from (b),(c), and this(d), we have

E:

$$P(Z_i = 1, X_i = 1|X = x, Y = y) = \frac{yx_ip_i}{1 - \prod_j(1-p_i)^{x_j}}$$

M:

$$p_i < -\frac{1}{T_i}\sum_t P(Z_i = 1, X_i = 1|X = x^{(t)}, Y = y^{(t)})$$

$$T_i = \sum_t P\left(X_i = 1\middle|X = x^{(t)}, Y = y^{(t)}\right)$$

Log:

$$L = \frac{1}{T}\sum_{t=1}^{T}\left(y^{(t)} * \log\left(1 - \prod_{i=1,}^{n}(1-p_i)^{x_i{}^{(t)}}\right) + (1-y^{(t)}) * \log\left(\prod_{i=1,}^{n}(1-p_i)^{x_i{}^{(t)}}\right)\right)$$

Mistake:

$$M = Count\left(y^{(t)} = 0, P\left(y^{(t)} = 1 \middle| \vec{x}^{(t)}\right) \geq 0.5\right)$$
$$+ Count\left(y^{(t)} = 1, P\left(y^{(t)} = 1 \middle| \vec{x}^{(t)}\right) \leq 0.5\right)$$

According to the result of java program, we can get

```
"C:\Program Files\Java\jdk1.8.0_281\bin\java.exe" ...
iteration 0___175___-0.9580854082157914
iteration 1___56___-0.49591639407753635
iteration 2___43___-0.40822081705839114
iteration 4___42___-0.3646149825001877
iteration 8___44___-0.3475006162087826
iteration 16___40___-0.33461704895854844
iteration 32___37___-0.32258140316749784
iteration 64___37___-0.3148266983628559
iteration 128___36___-0.3111558472151897
iteration 256___36___-0.310161353474076
```

iteration	Number of mistakes M	Log-likelihood L
0	175	-0.9580854082157914
1	56	-0.49591639407753635
2	43	-0.40822081705839114
4	42	-0.3646149825001877
8	44	-0.3475006162087826
16	40	-0.33461704895854844
32	37	-0.32258140316749784
64	37	-0.3148266983628559
128	36	-0.3111558472151897
256	36	-0.310161353474076

(e)

```
"C:\Program Files\Java\jdk1.8.0_281\bin\java.exe" ...
iteration 0___175___-0.9580854082157914
iteration 1___56___-0.49591639407753635
iteration 2___43___-0.40822081705839114
iteration 4___42___-0.3646149825001877
iteration 8___44___-0.3475006162087826
iteration 16___40___-0.33461704895854844
iteration 32___37___-0.32258140316749784
iteration 64___37___-0.3148266983628559
iteration 128___36___-0.3111558472151897
iteration 256___36___-0.310161353474076
```