

日期:

$$p(x|y, \theta) = \left(\frac{y_1}{x_1} \right) \left(\frac{\frac{1}{2}}{\frac{1}{2} + \frac{\theta}{4}} \right)^{x_1} \left(\frac{\frac{\theta}{4}}{\frac{1}{2} + \frac{\theta}{4}} \right)^{y_1 - x_1}$$

$$Q(\theta|\theta^{(k)}) = E \left(\log p(x, y|\theta) \mid T, \theta^{(k)} \right)$$

$$p(x, y|\theta) = \left(\frac{1}{2} \right)^{x_1} \left(\frac{\theta}{4} \right)^{x_2} \left(\frac{1-\theta}{4} \right)^{x_3} \left(\frac{1-\theta}{4} \right)^{x_4} \left(\frac{\theta}{4} \right)^{x_5} \left(\frac{x_1+x_2+x_3+x_4+x_5}{x_1, x_2, x_3, x_4, x_5} \right)$$

set $\log P(x, y|\theta) \triangleq c_1 + x_1 \log \left(\frac{1}{2} \right) + (x_2 + x_4) \log \left(\frac{\theta}{4} \right) + (x_3 + x_5) \log \left(\frac{1-\theta}{4} \right)$

$$\text{where } c_1 = \log \left(\frac{(x_1+x_2+x_3+x_4+x_5)!}{x_1! \cdot x_2! \cdot x_3! \cdot x_4! \cdot x_5!} \right)$$

substitute into $Q(\theta|\theta^{(k)})$, we have

$$\begin{aligned} Q(\theta|\theta^{(k)}) &= E_{\theta} \left[c_1 + x_1 \log \left(\frac{1}{2} \right) + (x_2 + x_4) \log \left(\frac{\theta}{4} \right) + (x_3 + x_5) \log \left(\frac{1-\theta}{4} \right) \mid T, \theta^{(k)} \right] \\ &= E(c_1 \mid T, \theta^{(k)}) + E(x_1 \log \left(\frac{1}{2} \right) \mid T, \theta^{(k)}) + \\ &\quad E \left((y_1 - x_1 + y_4) \log \left(\frac{\theta}{4} \right) + (y_2 + y_3) \log \left(\frac{1-\theta}{4} \right) \mid T, \theta^{(k)} \right) \end{aligned}$$

We find that 1st and 2nd term in RHS is no relevant with θ

so we consider 3rd term only.

$$\begin{aligned} \text{3rd term} &= (y_1 + y_4) \log \left(\frac{\theta}{4} \right) + (y_2 + y_3) \log \left(\frac{1-\theta}{4} \right) - E(x_1 \mid T, \theta^{(k)}) \cdot \log \left(\frac{\theta}{4} \right) \\ &= (y_1 + y_4) \log \left(\frac{\theta}{4} \right) + (y_2 + y_3) \log \left(\frac{1-\theta}{4} \right) - \left(\frac{2}{2+\theta^{(k)}} y_1 \right) \log \left(\frac{\theta}{4} \right) \end{aligned}$$

$$\text{if } \frac{\partial Q(\theta|\theta^{(k)})}{\partial \theta} = \frac{y_1 + y_4}{\theta} + \frac{y_2 + y_3}{\theta - 1} - \frac{\frac{2}{2+\theta^{(k)}} y_1}{\theta} = 0,$$

$$\frac{\partial Q(\theta|\theta^{(k)})}{\partial \theta} \Big|_{\theta^{(k+1)}} = 0, \quad \theta^{(k+1)} \text{ is equal to } M(\theta^{(k)}).$$

$$\text{that is } M(\theta^{(k)}) = \frac{\frac{\theta^{(k)}}{2+\theta^{(k)}} y_1 + y_4}{\frac{\theta^{(k)}}{2+\theta^{(k)}} y_1 + y_2 + y_3 + y_4}$$

$$= \frac{\theta^{(k)} \cdot y_1 + (2+\theta^{(k)}) y_4}{\theta^{(k)} y_1 + (2+\theta^{(k)}) (y_2 + y_3 + y_4)}$$

$$= \theta^{(k+1)}.$$

日期:

$$\theta^{(k+1)} = M(\theta^{(k)}) = \frac{230\theta^{(k)} + 110}{285\theta^{(k)} + 220}$$

$$= \boxed{\frac{46\theta^{(k)} + 22}{57\theta^{(k)} + 44}} \quad \square$$

2. single linkage: $\text{dist}(c_1, c_2) = \min_{x \in c_1, y \in c_2} \text{dist}(x, y)$

step 1: $\boxed{\text{level} = 0.47}$,

set $c_1 = \{x_1\}, c_2 = \{x_2, x_5\}, c_3 = \{x_3\}, c_4 = \{x_4\}$

	c_1	c_2	c_3
c_2	2.33		
c_3	3.15	1.30	
c_4	1.90	1.50	3.70

step 2:

$\boxed{\text{level} = 1.30}$,

$c_1 = \{x_1\}, c_5 = \{x_2, x_3, x_5\}, c_4 = \{x_4\}$

	c_1	c_5
c_5	2.33	
c_4	1.90	1.50

step 3:

$\boxed{\text{level} = 1.50}$

$c_1 = \{x_1\}, c_6 = c_4 \cup c_5 = \{x_2, x_3, x_4, x_5\}$

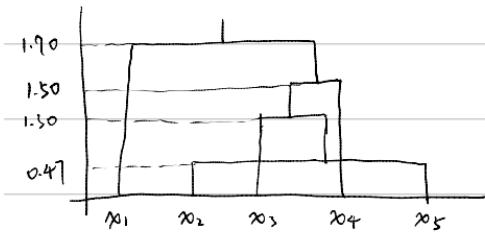
	c_1
c_6	1.90

step 4:

$\boxed{\text{level} = 1.90}$

日期:

cluster tree:



3. (a) $\text{support}\{\{e\}\} = \frac{8}{10} = \boxed{\frac{4}{5}}$,

$$\text{support}\{\{b, d\}\} = \frac{2}{10} = \boxed{\frac{1}{5}},$$

$$\text{support}\{\{b, d, e\}\} = \frac{2}{10} = \boxed{\frac{1}{5}}$$

(b) $\text{confidence}(\{\{b, d\}\} \rightarrow \{\{e\}\}) = \frac{\text{supp}\{\{b, d, e\}\}}{\text{supp}\{\{b, d\}\}} = \boxed{}$

$$\text{confidence}(\{\{e\}\} \rightarrow \{\{b, d\}\}) = \frac{\text{supp}\{\{b, d, e\}\}}{\text{supp}\{\{e\}\}} = \boxed{\frac{1}{4}}$$

not symmetric

(c) confidence $> 50\%$.

$$\text{confidence}(\{\{a\}\} \rightarrow \{\{b, e\}\}) = \frac{\text{supp}\{\{a, b, e\}\}}{\text{supp}\{\{a\}\}} = \frac{3}{7} < 50\%$$

$$\text{confidence}(\{\{b, e\}\} \rightarrow \{\{a\}\}) = \frac{3}{5} > 50\%$$

$$\text{confidence}(\{\{a, b\}\} \rightarrow \{\{e\}\}) = \frac{3}{4} > 50\%$$

$$\text{conf}(\{\{e\}\} \rightarrow \{\{a, b\}\}) = \frac{3}{8} < 50\%$$

日期: /

$$\text{conf}(\{a, e\} \rightarrow \{b\}) = \frac{3}{6} \geq 50\%$$

$$\text{conf}(\{b\} \rightarrow \{a, e\}) = \frac{3}{6} \geq 50\%$$

We Find Rules :

$$\boxed{\begin{array}{l} \{b\} \rightarrow \{a, e\}, \quad \{a, e\} \rightarrow \{b\} \\ \{a, b\} \rightarrow \{c\}, \quad \{b, e\} \rightarrow \{a\} \end{array}}$$