

4.2.7.1 Lift

Association Rule : Tea \rightarrow Coffee

	coffee	$\overline{\text{coffee}}$	
Tea	150	50	200
$\overline{\text{Tea}}$	650	150	800
	800	200	1000

Q : Lift (Tea, coffee) = ?

independent?

positively correlated?

negatively correlated?

Solution 1

$$\textcircled{1} \text{ Lift (Tea, coffee)} = \frac{\text{confidence (Tea} \rightarrow \text{coffee)}}{\text{Support (Coffee)}}$$

$$\textcircled{2} \text{ Support (coffee)} = p(\text{coffee}) = \frac{800}{1000} = 0.8$$

$$\textcircled{3} \text{ confidence (Tea} \rightarrow \text{coffee)}$$

$$= \frac{p(\text{Tea} \cup \text{coffee})}{p(\text{Tea})}$$

← 所有U当
来看, 特色

$$= \frac{\frac{150}{1000}}{\frac{200}{1000}} = 0.75$$

	coffee	¬coffee	
Tea	150	50	200
¬Tea	650	150	800
	800	200	1000

$$\textcircled{4} \text{ Lift} = \frac{0.75}{0.8} = 0.9375 < 1$$

negatively correlated

Solution 2

confidence (Tea \rightarrow coffee)

$$= p(\text{coffee} | \text{Tea})$$

$$= \frac{150}{200}$$

	coffee	$\overline{\text{coffee}}$	
Tea	150	50	200
$\overline{\text{Tea}}$	650	150	800
	800	200	1000

$$p(\text{Tea}) = \frac{200}{1000} = 0.2 \quad p(\text{coffee}) = \frac{800}{1000} = 0.8$$

$$p(\text{coffee} | \overline{\text{Tea}}) = \frac{650}{800} = 0.8125$$

即使你忽略某人是否喝茶，喝咖啡的概率也是 80%
知道喝茶反而会降低喝咖啡的概率

$$\text{Lift}(X, Y) = \frac{\text{confidence}(X \rightarrow Y)}{\text{support}(Y)}$$

$$= \frac{p(X \cup Y)}{p(X)} \cdot \frac{1}{\text{support}(Y)}$$

$$= \frac{p(X \cup Y)}{p(X)} \cdot \frac{1}{p(Y)}$$

$$= \frac{p(Y|x)}{p(Y)}$$

$$\text{Lift (Tea} \rightarrow \text{coffee)} = \frac{p(\text{coffee} | \text{Tea})}{p(\text{coffee})}$$

$$= \frac{150/200}{0.8}$$

$$= \frac{0.75}{0.8}$$

$$= 0.9375 < 1$$

not interesting

负相关