

Date

No.

Lê Quang Thọ - MSV: 181202289.  
CNTT1 - K59.

Câu 1.

X: 9.1 4.5 5.8 6.7 6.5 7 6.0 5.5 7  
7 8.5 8.6

Y: 8.5 5.5 7.5 8.5 6 6.5 6.8 9 7.2  
8.0 8.5

a, Với dãy X.

Mảng sau khi sắp xếp là:

4.5 5.5 5.5 6 6.5 6.7 7 7 7 8.5 8.6 9.1

$$\begin{aligned}
 + \text{Giá trị trung bình } \bar{X} &= \frac{x_1 + x_2 + \dots + x_n}{n} \\
 &= \frac{4.5 + 5.5 + \dots + 9.1}{12} \\
 &= \frac{81.7}{12} = 6.8
 \end{aligned}$$

$$+ \text{Trung vị} = \frac{6.7 + 7}{2} = 6.85$$

+ Mode = 7 (vì 7 xuất hiện nhiều nhất với tần suất là 3).

\* Với dãy Y

Mảng sau khi sắp xếp là:

5.5 6 6 6.5 6.8 7.2 7.5 8 8.5 8.5  
8.5 9



Date

No.

$$\begin{aligned}
 + \text{Giá trị trung bình } \bar{Y} &= \frac{y_1 + y_2 + \dots + y_n}{n} \\
 &= \frac{5.5 + 6 + \dots + 9}{12} \\
 &= \frac{88}{12} = 7.33
 \end{aligned}$$

$$+ \text{Trung vị} = \frac{7.2 + 7.5}{2} = 7.35$$

+ Mode = 8.5 (vì 8.5 xuất hiện với tần suất 3 lần, nhiều nhất)

b, Biên độ box plot của Y

$$Q_1 = \frac{6 + 6.5}{2} = 6.25$$

$$Q_2 = 7.35$$

$$Q_3 = \frac{8.5 + 8.5}{2} = 8.5$$

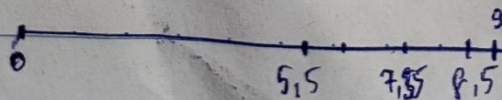
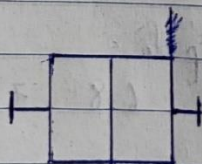
$$IQR = Q_3 - Q_1 = 8.5 - 6.25 = 2.25$$

$$\begin{aligned}
 \text{xét chiều ngược lại } Q_1 - 1.5 IQR &= 6.25 - 1.5 \times 2.25 \\
 &= 2.875
 \end{aligned}$$

$$\Rightarrow \text{Min} = 5.5$$

$$\begin{aligned}
 Q_2 + 1.5 IQR &= 8.5 + 1.5 \times 2.25 \\
 &= 11.875
 \end{aligned}$$

$$\Rightarrow \text{Max} = 9$$



HONG HA



c, chuẩn hoá y về dạng chuẩn z-score

$$+ \bar{y} = 7,33$$

$$+ s_y = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2} = \sqrt{\frac{1}{12} \sum_{i=1}^{12} (5,5^2 + 6^2 + \dots + 9^2) - 7,33^2}$$

$$= \sqrt{1,27} = 1,13$$

$$+ y'_1 = \frac{y_1 - \bar{y}}{s_y} = \frac{8,5 - 7,33}{1,13} = 1,085$$

Tương tự:

$$y'_2 = -1,62$$

$$y'_7 = -0,79$$

$$y'_3 = -1,18$$

$$y'_8 = -0,47$$

$$y'_4 = 0,15$$

$$y'_9 = 1,47$$

$$y'_5 = 1,085$$

$$y'_{10} = -0,12$$

$$y'_6 = -1,18$$

$$y'_{11} = 0,54$$

$$y'_{12} = 1,04$$

d, làm tròn giữ lại x

+ số thặng là 3

+ phân chia Equal-width

$$\text{Độ rộng} = \frac{\text{Max} - \text{Min}}{3} = \frac{9,1 - 4,5}{3} = 1,53$$

Sắp xếp:

4,5 5,3 5,5 6 6,5 6,7 7 7 7 8,5 8,6 9,1

Bin 1 [4,5; 6,03): 4,5, 5,3 5,5, 6

Bin 2 [6,03; 7,56): 6,5, 6,7, 7, 7, 7

Bin 3 [7,56; 9,1): 8,5, 8,6, 9,1

+ làm tròn giữ lại các bin đầu của x bằng phương pháp tròn thặng theo trung bình (bin means).



Date

No.

Tần suất bình của hỗn bin

bin 1. 5.325

bin 2. 6.84

bin 3. 8.73

Lần nín

→ bin 1. 5.325, 5.325, 5.325, 5.325

bin 2. 6.84, 6.84, 6.84, 6.84, 6.84

bin 3. 8.73, 8.73, 8.73

e, Xác định mối tương quan X và Y  
+ Tính hệ số tương quan.

$$r_{xy} = \frac{\sum_{i=1}^n x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{n \cdot s_x \cdot s_y}$$

$$\bar{x} = 6.8$$

$$\bar{y} = 7.3$$

$$s_x = 1.34$$

$$s_y = 1.13$$

$$\rightarrow r_{xy} = \frac{\sum_{i=1}^{12} x_i y_i - 12 \cdot 6.8 \cdot 7.3}{12 \cdot 1.34 \cdot 1.13}$$

$$= 0.73$$

→ X và Y tương quan dương



Date

No.

Câu 2.

a,  $\text{min sup} = 60\% \rightarrow \text{min sup count} = \lceil 60\% \times 5 \rceil = 3$ Top  $I = \{I_1, I_2, I_3, I_4, I_5, I_6, I_7\} \rightarrow 7 \text{ items}$ C<sub>1</sub>

item	Sub count
$\{I_1\}$	3
$\{I_2\}$	3
$\{I_3\}$	4
$\{I_4\}$	2
$\{I_5\}$	3
$\{I_6\}$	2
$\{I_7\}$	1

L<sub>1</sub>

item	Sub-count
$\{I_1\}$	3
$\{I_2\}$	3
$\{I_3\}$	4
$\{I_5\}$	3

C<sub>2</sub>

Item	Sup count
$\{I_1, I_2\}$	2
$\{I_1, I_3\}$	3
$\{I_1, I_5\}$	2
$\{I_2, I_3\}$	3
$\{I_2, I_5\}$	3
$\{I_3, I_5\}$	3

L<sub>2</sub>

Item	Sup-count
$\{I_1, I_3\}$	3
$\{I_2, I_3\}$	3
$\{I_2, I_5\}$	3
$\{I_3, I_5\}$	3

C<sub>3</sub>

Item	Sub-count
$\{I_2, I_3, I_5\}$	3

L<sub>3</sub>

Item	Sup-count
$\{I_2, I_3, I_5\}$	3



$\Rightarrow$  Tập phủ biến  $L = L_1 \cup L_2 \cup L_3$

$\{I_1\} \{I_2\} \{I_3\} \{I_5\} \{I_1, I_2\} \{I_2, I_3\}$   
 $\{I_2, I_5\} \{I_3, I_5\} \{I_2, I_3, I_5\}$

b) Xét tập  $\{I_1, I_3\}$

sinh tập con

$\{I_1\} \Rightarrow \{I_3\}$

confidence =  $3/3 = 100\%$

$\{I_3\} \Rightarrow \{I_1\}$

=  $3/4 = 75\%$

Xét tập  $\{I_2, I_3\}$

$\{I_2\} \Rightarrow \{I_3\}$

Confidence =  $3/3 = 100\%$

$\{I_3\} \Rightarrow \{I_2\}$

=  $3/4 = 75\%$

Xét tập  $\{I_2, I_5\}$

$\{I_2\} \Rightarrow \{I_5\}$

confidence =  $3/3 = 100\%$

$\{I_5\} \Rightarrow \{I_2\}$

=  $3/3 = 100\%$

Xét tập  $\{I_3, I_5\}$

Confidence =  $3/4 = 75\%$

$\{I_3\} \Rightarrow \{I_5\}$

=  $3/3 = 100\%$

$\{I_5\} \Rightarrow \{I_3\}$

Xét tập  $\{I_2, I_3, I_5\}$

Có các tập con

$I_2 \Rightarrow \{I_3, I_5\}$

confidence =  $3/3 = 100\%$

$I_3 \Rightarrow \{I_2, I_5\}$

=  $3/4 = 75\%$

$I_5 \Rightarrow \{I_2, I_3\}$

=  $3/3 = 100\%$

$\{I_2, I_3\} \Rightarrow I_5$

=  $3/3 = 100\%$

$\{I_2, I_5\} \Rightarrow I_3$

=  $3/3 = 100\%$

$\{I_3, I_5\} \Rightarrow I_2$

=  $3/3 = 100\%$



Date

No.

Các luật kết hợp có confidence  $\geq 75\%$  là  
luật kết hợp mạnh.

Vậy tất cả các luật kết hợp do trên đây là luật  
kết hợp mạnh.