

Phân tổ lại

- Khái niệm: lập ra một số tổ mới trên cơ sở các tổ cũ đã có sẵn, nhằm đáp ứng một mục đích nghiên cứu nào đó.
- Điều kiện vận dụng:
 - Với cách phân tổ cũ, việc so sánh giữa các tài liệu gặp khó khăn do không thống nhất về số tổ, khoảng cách tổ.
 - Cách phân tổ cũ chưa phản ánh rõ các loại hình KTXH.
 - Cách phân tổ cũ chưa hợp lý, không phản ánh đúng tình hình thực tế.

R Phương pháp phân tổ lại

Lập tổ mới bằng cách thay đổi các khoảng cách tổ của phân tổ cũ



Lập tổ mới theo tỷ trọng của mỗi tổ trong tổng thể

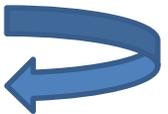


Thay đổi các khoảng cách tổ của phân tổ cũ

Doanh nghiệp A			Doanh nghiệp B		
Thâm niên (năm)	Tỷ lệ (%) về		Thâm niên (năm)	Tỷ lệ (%) về	
	Lao động	Tiền lương		Lao động	Tiền lương
< 5	35	26	< 2	18	13
5 - 10	30	30	2 - 5	22	22
10 - 15	20	24	5 - 7	26	27
15 - 20	10	12	7 - 10	20	20
>= 20	5	8	10 - 15	8	10
			15 - 20	4	5
			>= 20	2	3
Cộng	100	100	Cộng	100	100

Thay đổi các khoảng cách tổ của phân tổ cũ

Doanh nghiệp A			Doanh nghiệp B		
Thâm niên (năm)	Tỷ lệ (%) về		Thâm niên (năm)	Tỷ lệ (%) về	
	Lao động	Tiền lương		Lao động	Tiền lương
< 5	35	26	< 5	40	35
5 - 10	30	30	5 - 10	46	47
10 - 15	20	24	10 - 15	8	10
15 - 20	10	12	15 - 20	4	5
>= 20	5	8	>= 20	2	3
Cộng	100	100	Cộng	100	100



Lập tổ mới theo tỷ trọng của mỗi tổ trong tổng thể

Số HS (100 người)	Tỷ lệ % về	
	Số trường	Giáo viên
< 7	10	5
7 - 9	15	10,1
9 - 11	18	16,8
11 - 13	27	27,2
13 - 15	20	24,8
>= 15	10	16,1
Cộng	100	100

Cỡ trường	Tỷ lệ % về	
	Số trường	Giáo viên
Nhỏ	35	?1
Trung bình	50	?2
Lớn	15	?3
Cộng	100	

$$?1 = 5 + 10,1 + \frac{16,8}{18} * 10 = 24,433$$

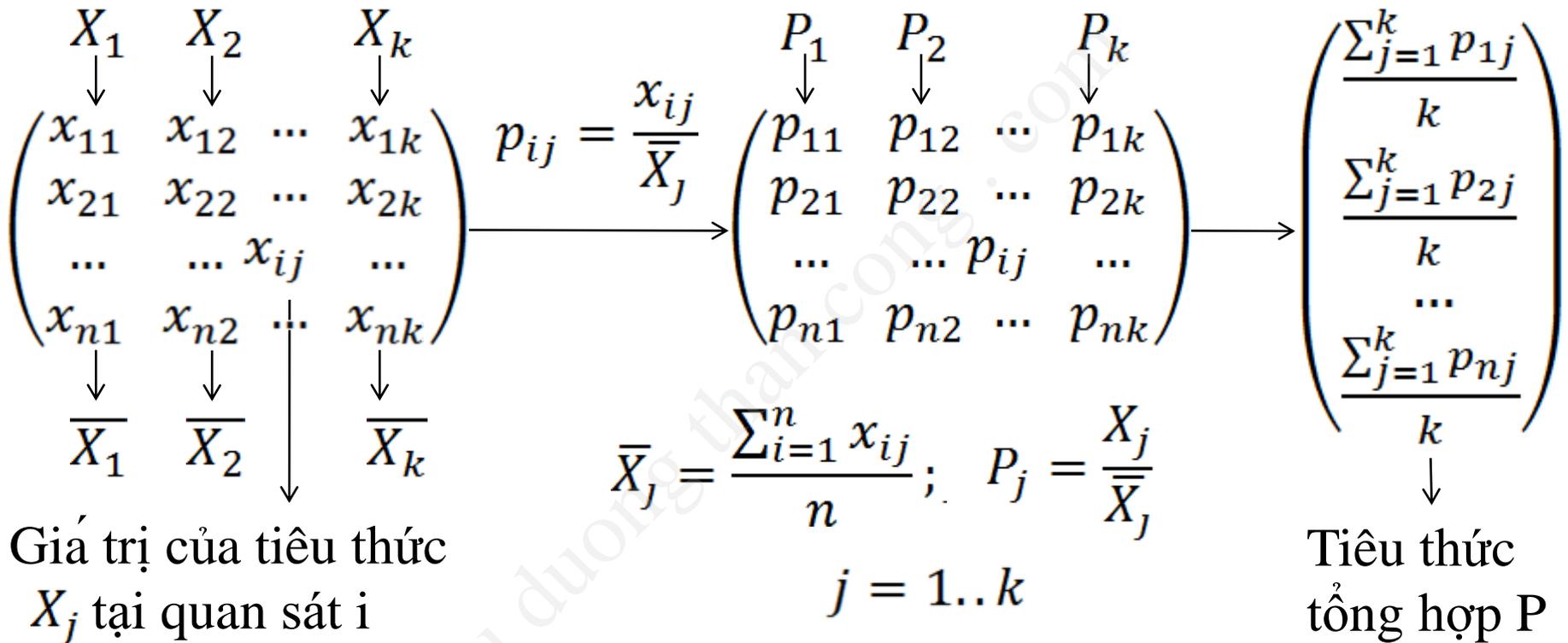
$$?2 = \frac{16,8}{18} * 8 + 27,2 + \frac{24,8}{20} * 15 = 53,267$$

$$?3 = \frac{24,8}{20} * 5 + 16,1 = 22,3$$

Phân tổ nhiều chiều

- Khái niệm: cùng lúc phân tổ theo nhiều tiêu thức có vai trò như nhau trong việc đánh giá hiện tượng.
- Tác dụng:
 - Nghiên cứu kết cấu tổng thể theo một số tiêu thức cơ bản có mối quan hệ với nhau.
 - Thay thế cho phân tổ kết hợp khi có nhiều hơn 2 tiêu thức nguyên nhân cùng tác động đến tiêu thức kết quả.
 - Xây dựng tài liệu đồng nhất của thông tin ban đầu.
 - Khả thi ngay cả khi các phương pháp phân tổ khác không thực hiện được.

Tiêu thức phân tử (nhiều chiều)



$$P = \frac{1}{k} \sum_{j=1}^k \frac{X_j}{\bar{X}_j} = \frac{1}{k} \sum_{j=1}^k P_j$$

Ý nghĩa: đưa các tiêu thức khác nhau về cùng một dạng có vai trò như nhau trong việc biểu hiện tính chất của hiện tượng.

STT	Giá trị thiết bị SX chủ yếu (tỷ đồng) X1	Số LĐ X2	Giá trị SX năm (tỷ đồng) Y	P1	P2	P
1	10,6	2730	191	2,657	2,084	2,37
2	0,6	200	10	0,15	0,153	0,152
3	5,9	3000	161	1,479	2,29	1,884
4	0,9	366	12	0,226	0,279	0,252
5	4,7	1210	62	1,178	0,924	1,051
6	3,5	990	30	0,877	0,756	0,816
7	0,8	880	27	0,201	0,672	0,436
8	4,3	960	48	1,078	0,733	0,905
9	7,3	1910	103	1,83	1,458	1,644
10	1,3	854	24	0,326	0,652	0,489
TB	3,99	1410	66,8	1	1	

Phương pháp phân tổ nhiều chiều

Chuẩn bị: Sắp xếp dữ liệu theo chiều tăng của tiêu thức tổng hợp P.

Tiến hành: 2 cách:

- C1: Đưa về bài toán phân tổ theo 1 tiêu thức P

Sau khi phân tổ sử dụng các hàm kiểm tra để kiểm tra tính đồng nhất và bền vững của từng tổ.

- C2: Sử dụng các hàm kiểm tra



Hàm kiểm tra tính đồng nhất

$$U(L) = \frac{n-1}{n(n-L)L} * \frac{[(n-L) \sum_{i=1}^L x_i - L \sum_{i=L+1}^n x_i]^2}{\sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2}$$

n : số đơn vị trong tổ

L : thứ tự lượng biến trong tổ $L = 1..(n-1)$

x_i : lượng biến của tiêu thức tại vị trí i $i = 1..n$



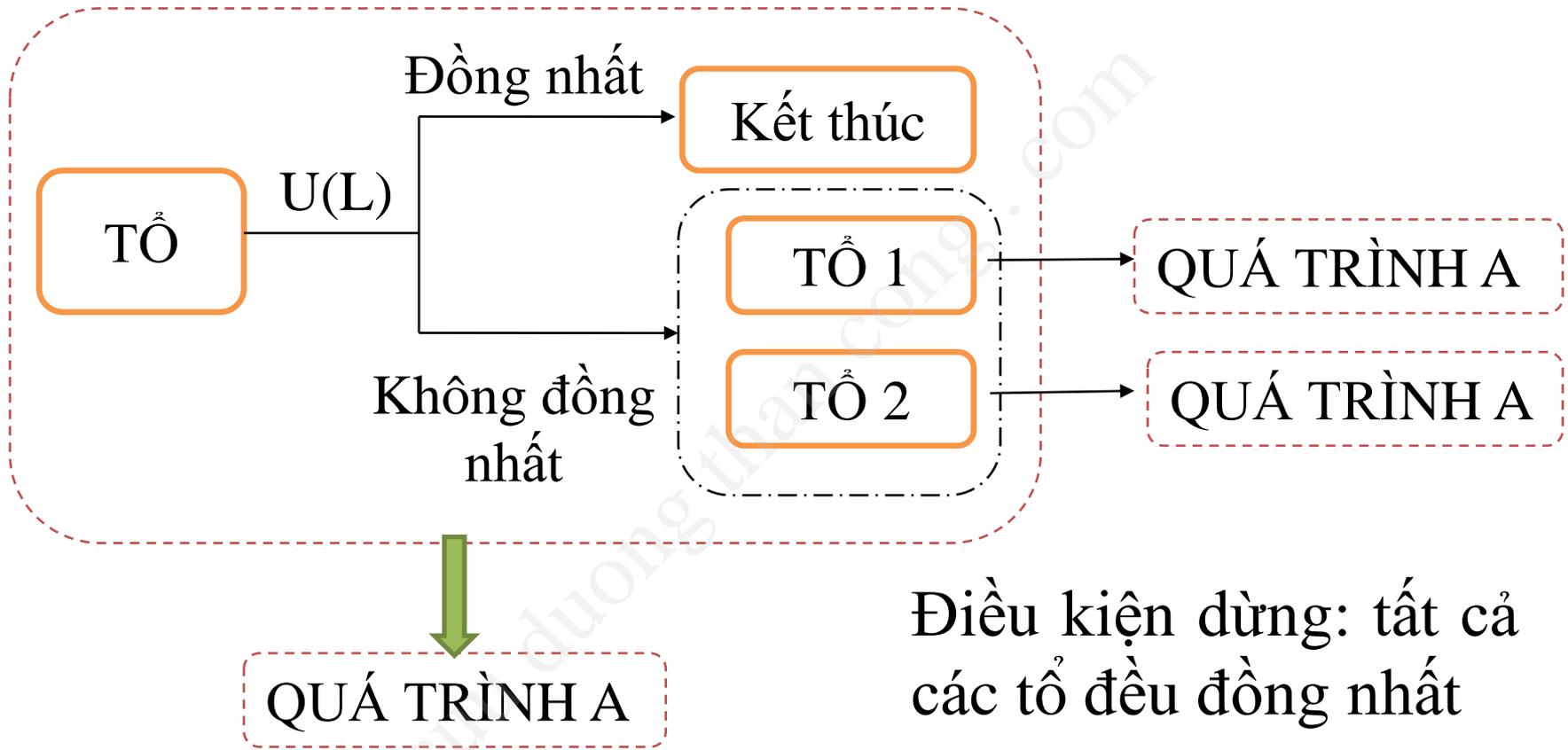
Hàm kiểm tra tính đồng nhất

$$\Leftrightarrow U(L) = \frac{(n - 1)L}{n - L} * \frac{(\overline{X^L} - \bar{X})^2}{\overline{X^2} - (\bar{X})^2}$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \quad \overline{X^2} = \frac{1}{n} \sum_{i=1}^n x_i^2 \quad \overline{X^L} = \frac{1}{L} \sum_{i=1}^L x_i$$

- Tổ đồng nhất $\Leftrightarrow \forall L \in \{1, \dots, n\}, U(L) \leq \chi_{\alpha;1}^2$
- Tổ không đồng nhất $\Leftrightarrow \exists L \in \{1, \dots, n\} / U(L) > \chi_{\alpha;1}^2$

Quá trình phân tổ nhiều chiều



Căn cứ phân tổ: Lấy quan sát thứ L sao cho $U(L)$ lớn nhất làm giới hạn của 2 tổ

STT	X	L	X ²	$\overline{X^L}$	U(L)
2	0,152	1	0,02296	0,152	1,446
4	0,252	2	0,06374	0,202	2,879
7	0,436	3	0,19021	0,28	4,017
10	0,489	4	0,23899	0,332	5,375
6	0,816	5	0,66661	0,429	5,893
8	0,905	6	0,81949	0,508	6,553
5	1,051	7	1,10419	0,586	7,234
9	1,644	8	2,70206	0,718	5,745
3	1,884	9	3,55091	0,848	3,772
1	2,37	10	5,61835	1	
TB	1		1,49775		

$$\alpha = 5\%$$

$$\chi_{\alpha;1}^2 = 3,841$$

$$\text{Max } U(L) = 7,234$$

$$\Leftrightarrow L = 7$$

Tổ 1: 7 doanh nghiệp đầu

Tổ 2: 3 doanh nghiệp còn lại

STT	X	L	X ²	$\overline{X^L}$	U(L)
2	0,152	1	0,023	0,152	1,879
4	0,252	2	0,064	0,202	3,523
7	0,436	3	0,19	0,28	4,193
10	0,489	4	0,239	0,332	5,127
6	0,816	5	0,667	0,429	3,674
8	0,905	6	0,819	0,508	2,152
5	1,051	7	1,104	0,586	
TB	0,586		0,444		

→ Tổ 1

$$\chi_{\alpha;1}^2 = 3,841$$

$$\text{Max } U(L) = 5,127$$

$$\Leftrightarrow L = 4$$

Tổ 1.1: 4 doanh nghiệp đầu

Tổ 1.2: 3 doanh nghiệp còn lại

STT	X	L	X ²	$\overline{X^L}$	U(L)
9	1,644	1	2,702	1,644	1,138
3	1,884	2	3,551	1,764	1,789
1	2,37	3	5,618	1,966	
TB	1,966		3,957		

→ Tổ 2

TỔ 1.1

STT	X	L	X^2	$\overline{X^L}$	U(L)
2	0,152	1	0,023	0,152	1,757
4	0,252	2	0,064	0,202	2,738
7	0,436	3	0,19	0,28	1,32
10	0,489	4	0,239	0,332	
TB	0,332		0,129		

TỔ 1.2

STT	X	L	X^2	$\overline{X^L}$	U(L)
6	0,816	1	0,667	0,816	1,243
8	0,905	2	0,819	0,861	1,718
5	1,051	3	1,104	0,924	
TB	0,924		0,863		



Hàm kiểm tra tính bền vững

$$U(k; k + 1) = \frac{n_k + n_{k+1} - 1}{n_k n_{k+1} (n_k + n_{k+1})} * \frac{(n_{k+1} \sum_k x_i - n_k \sum_{k+1} x_i)^2}{\sum_{k;k+1} x_i^2 - \frac{1}{(n_k + n_{k+1})} (\sum_{k;k+1} x_i)^2}$$

k : Số thứ tự của tổ n_k : Số đơn vị của tổ

$\sum_k x_i$: Tổng lượng biến của tổ thứ k

$\sum_{k;k+1} x_i$: Tổng lượng biến của tổ thứ k và $k+1$

$\sum_{k;k+1} x_i^2$: Tổng bình phương các lượng biến của tổ thứ k và $k+1$



Hàm kiểm tra tính bền vững

Đặt: $n = n_k + n_{k+1}$

$$U(k; k + 1) = \frac{(n - 1)n_k}{n - n_k} * \frac{(\overline{X}_k - \overline{X}_{k;k+1})^2}{\overline{X_{k;k+1}^2} - (\overline{X}_{k;k+1})^2}$$

n : Tổng số đơn vị của tổ thứ k và $k+1$

\overline{X}_k : Trung bình các lượng biến của tổ thứ k

$\overline{X_{k;k+1}}$: Trung bình các lượng biến của tổ thứ k và $k+1$

$\overline{X_{k;k+1}^2}$: Trung bình bình phương các lượng biến của tổ thứ k và $k+1$



Hàm kiểm tra tính bền vững

$$U(k; k + 1) \leq \chi_{\alpha; 1}^2 \Rightarrow \text{Gộp 2 tổ lại}$$

$$U(k; k + 1) > \chi_{\alpha; 1}^2 \Rightarrow \text{2 tổ độc lập khác nhau}$$

Ý nghĩa: dùng để kiểm tra tính bền vững của giới hạn giữa hai tổ gần nhau (các tổ được xác định là kết quả của quá trình kiểm tra tính đồng nhất hoặc phân tổ giản đơn). Đây là bước sau cùng của phân tổ nhiều chiều.

III.2. Bảng thống kê

Bảng thống kê là một hình thức trình bày các tài liệu thống kê một cách có hệ thống, hợp lý và rõ ràng, nhằm nêu lên các đặc trưng về mặt lượng của hiện tượng nghiên cứu.

Cấu thành bảng thống kê

Tên bảng thống kê (tiêu đề chung)

Phần giải thích Phần chủ đề	Các chỉ tiêu giải thích (tên cột)				
	(A)	(1)	(2)	(3)	(4)
Các chỉ tiêu giải thích (tên cột)					

Nguyên tắc trình bày bảng

- Quy mô của bảng thống kê không nên quá lớn.
- Các tiêu đề và tiêu mục trong bảng cần ghi chính xác, gọn và dễ hiểu.
- Các hàng và cột thường được ký hiệu bằng chữ hoặc bằng số.
- Các chỉ tiêu giải thích trong bảng cần được sắp xếp theo thứ tự hợp lý
- Phần ghi chú ở cuối bảng thống kê.

Cách ghi số liệu vào bảng thống kê

- + Nếu hiện tượng không có số liệu đó, thì trong ô sẽ ghi một dấu gạch ngang (-)
- + Nếu số liệu còn thiếu, sau này có thể bổ sung, thì trong ô có ký hiệu 3 chấm (...)
- + Trong một ô nào đó: hiện tượng không có liên quan đến chỉ tiêu đó, nếu viết số liệu vào ô đó sẽ vô nghĩa, thì trong ô có ký hiệu gạch chéo (x)
- + Các số liệu trong cùng một cột, có đơn vị tính toán giống nhau, phải ghi theo trình độ chính xác như nhau (số lẻ đến 0,1 hay 0,01,...) đơn vị tính phải ghi thống nhất theo quy định.

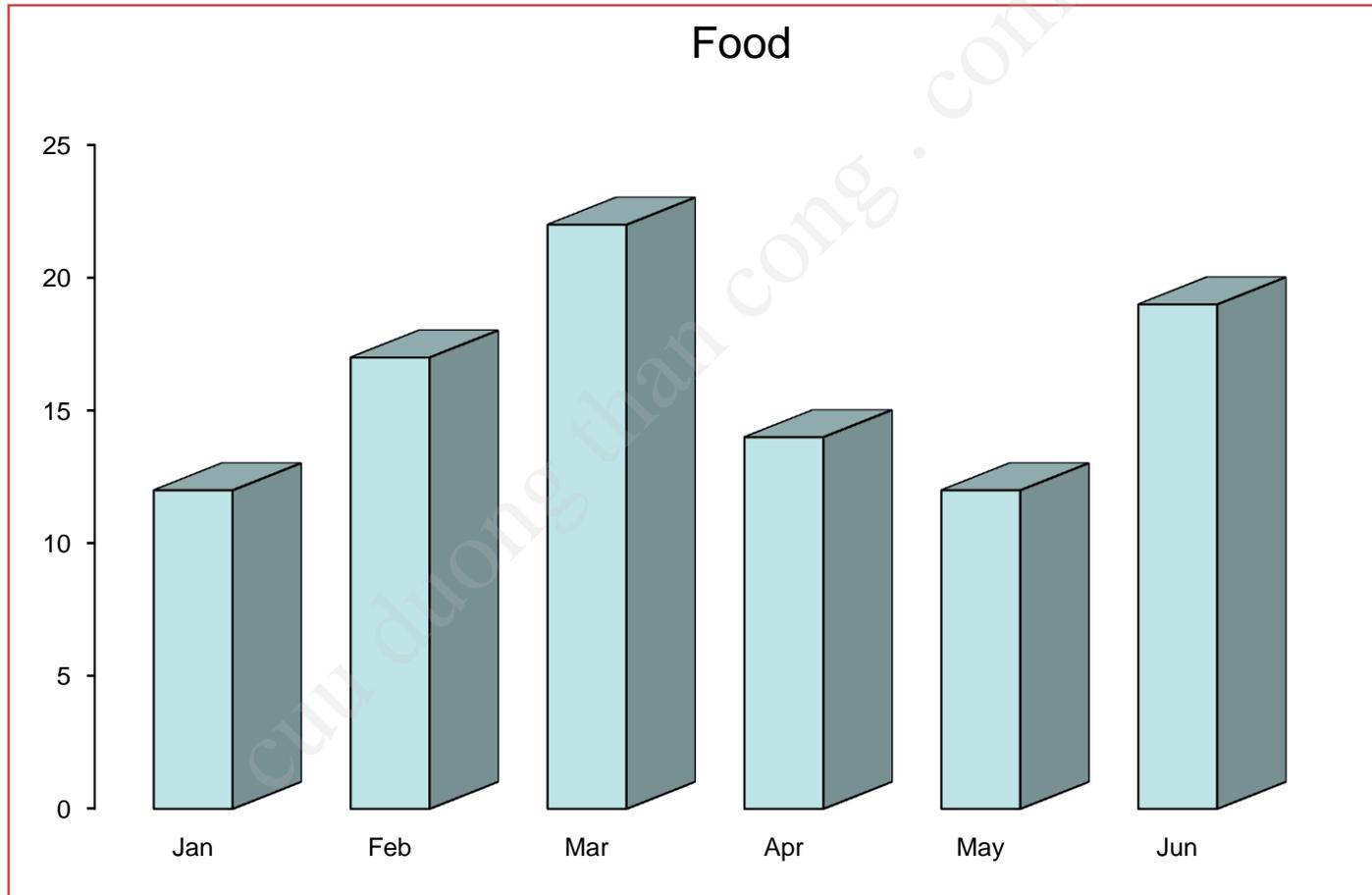
III.3. Đồ thị thống kê

- ✦ Là các hình vẽ hoặc đường nét hình học dùng để miêu tả có tính chất quy ước các tài liệu thống kê.
- ✦ Có tính quần chúng, có sức hấp dẫn và sinh động, làm cho người hiểu biết ít về thống kê vẫn lĩnh hội được vấn đề chủ yếu một cách dễ dàng, đồng thời giữ được ấn tượng sâu đối với người đọc.

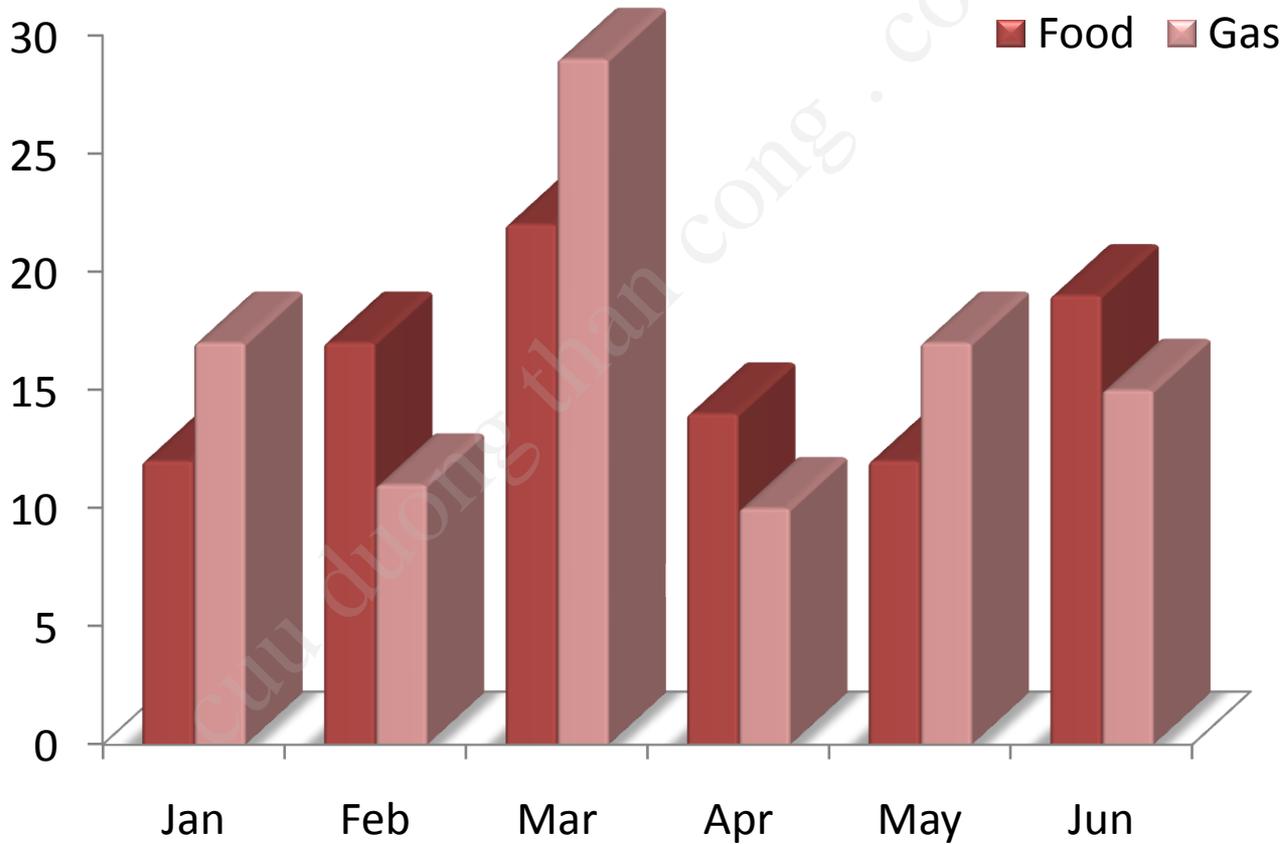
Nguyên tắc trình bày đồ thị

- Quy mô đồ thị vừa phải
- Lựa chọn loại đồ thị phù hợp
- Thang đo và tỷ lệ xích phải được xác định chính xác
- Đồ thị phải dễ xem và dễ hiểu

Đồ thị hình cột



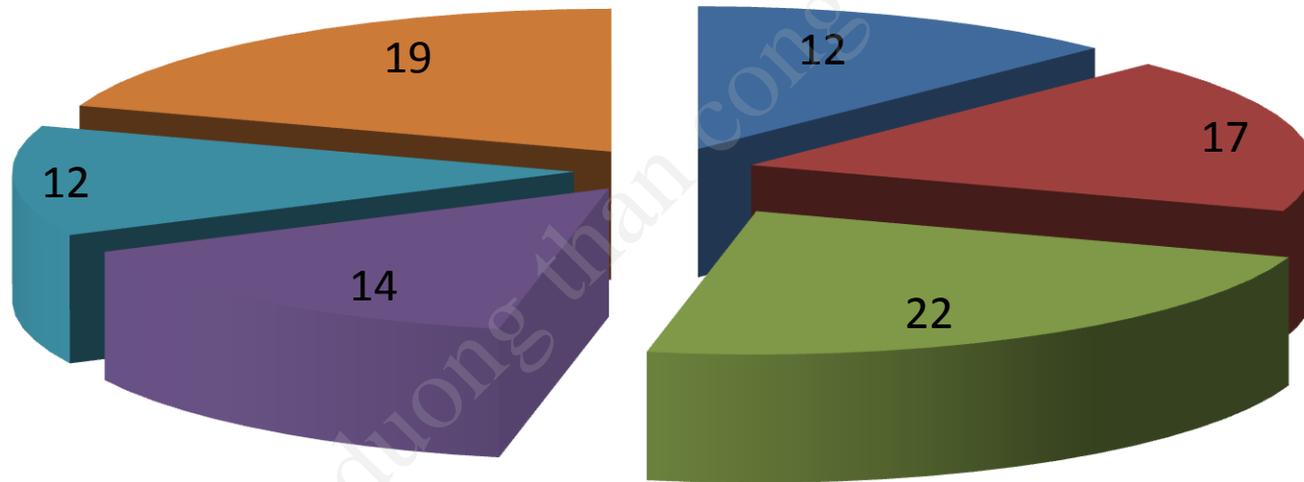
Đồ thị so sánh hình cột



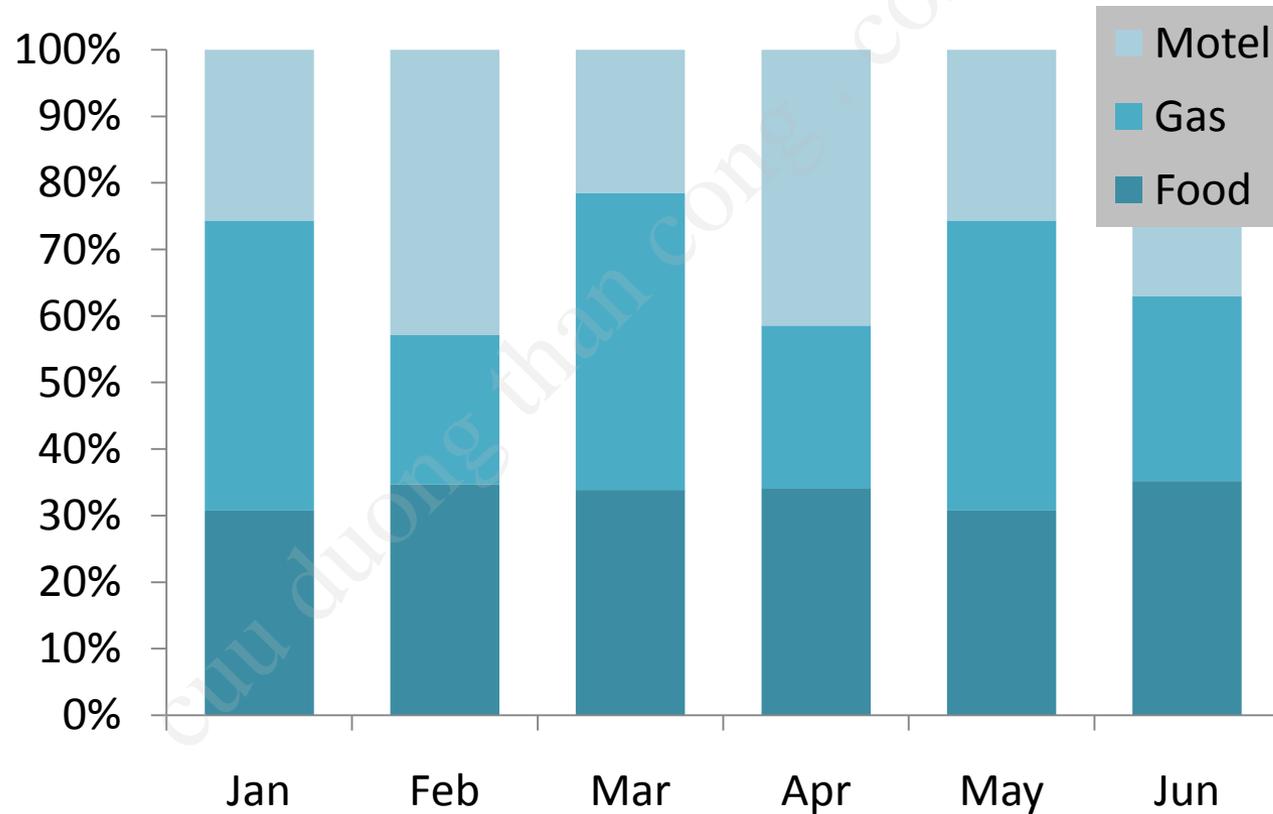
Đồ thị đường gấp khúc



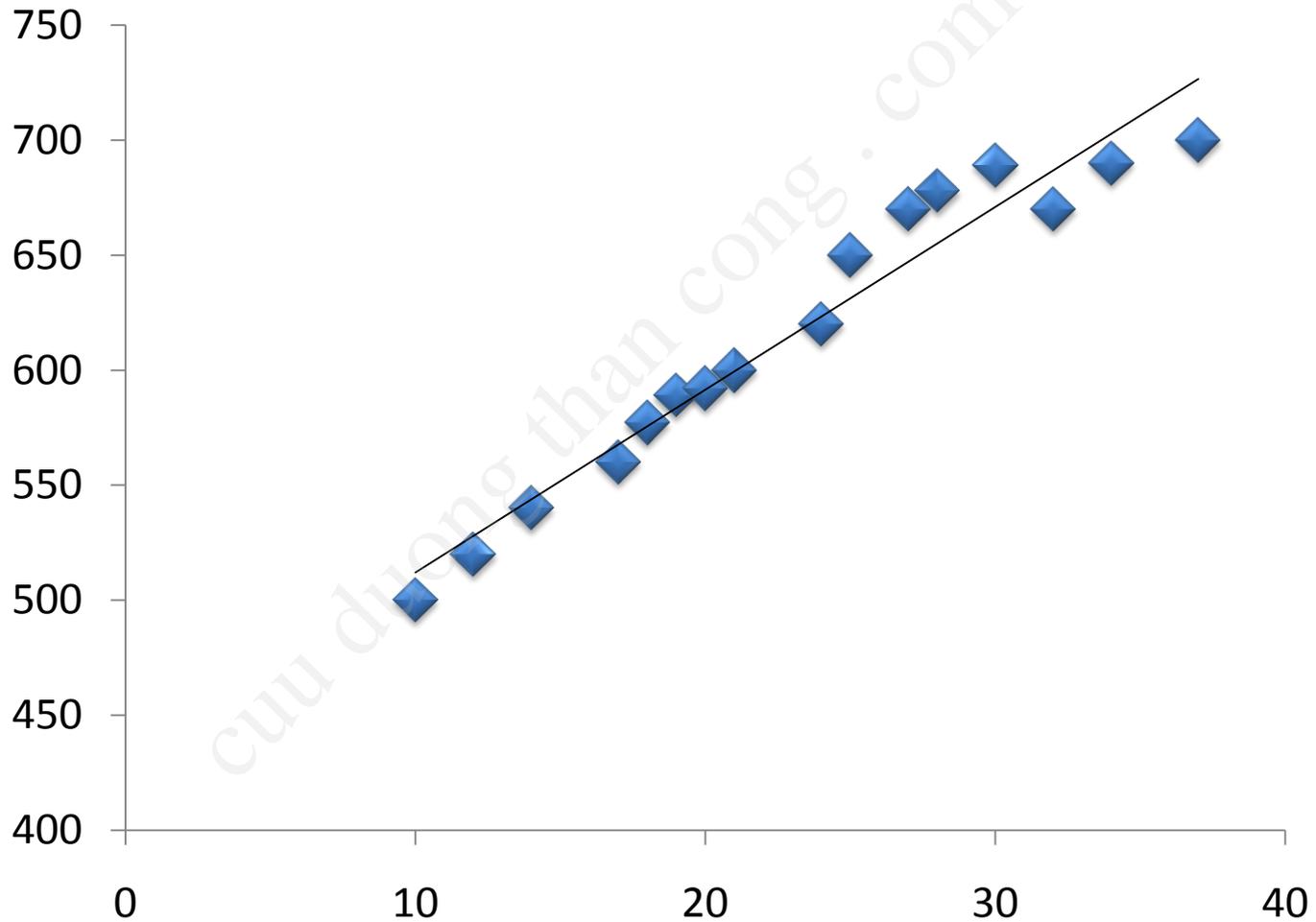
Đồ thị cơ cấu



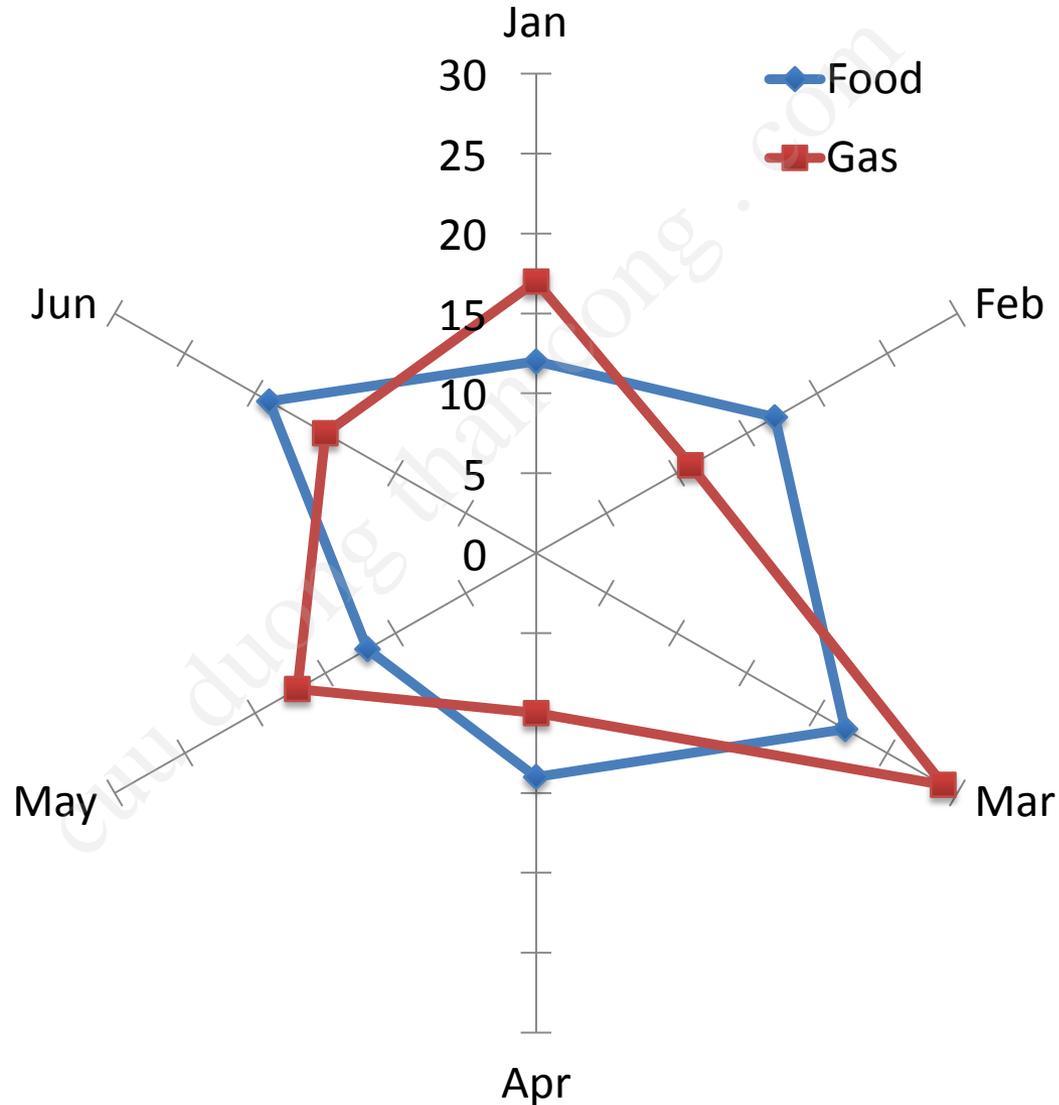
Đồ thị so sánh cơ cấu



Đồ thị liên hệ



Đồ thị hình “mạng nhện”



Bản đồ thống kê

