

CHƯƠNG III: PHƯƠNG PHÁP TỔNG HỢP THỐNG KÊ

NỘI DUNG

III.1

Phân tổ thống kê

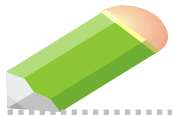
III.2

Bảng thống kê

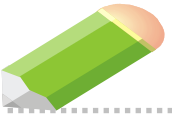
III.3

Đồ thị thống kê

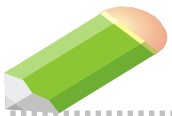
III.1. Phân tổ thống kê



Khái niệm, ý nghĩa và nhiệm vụ



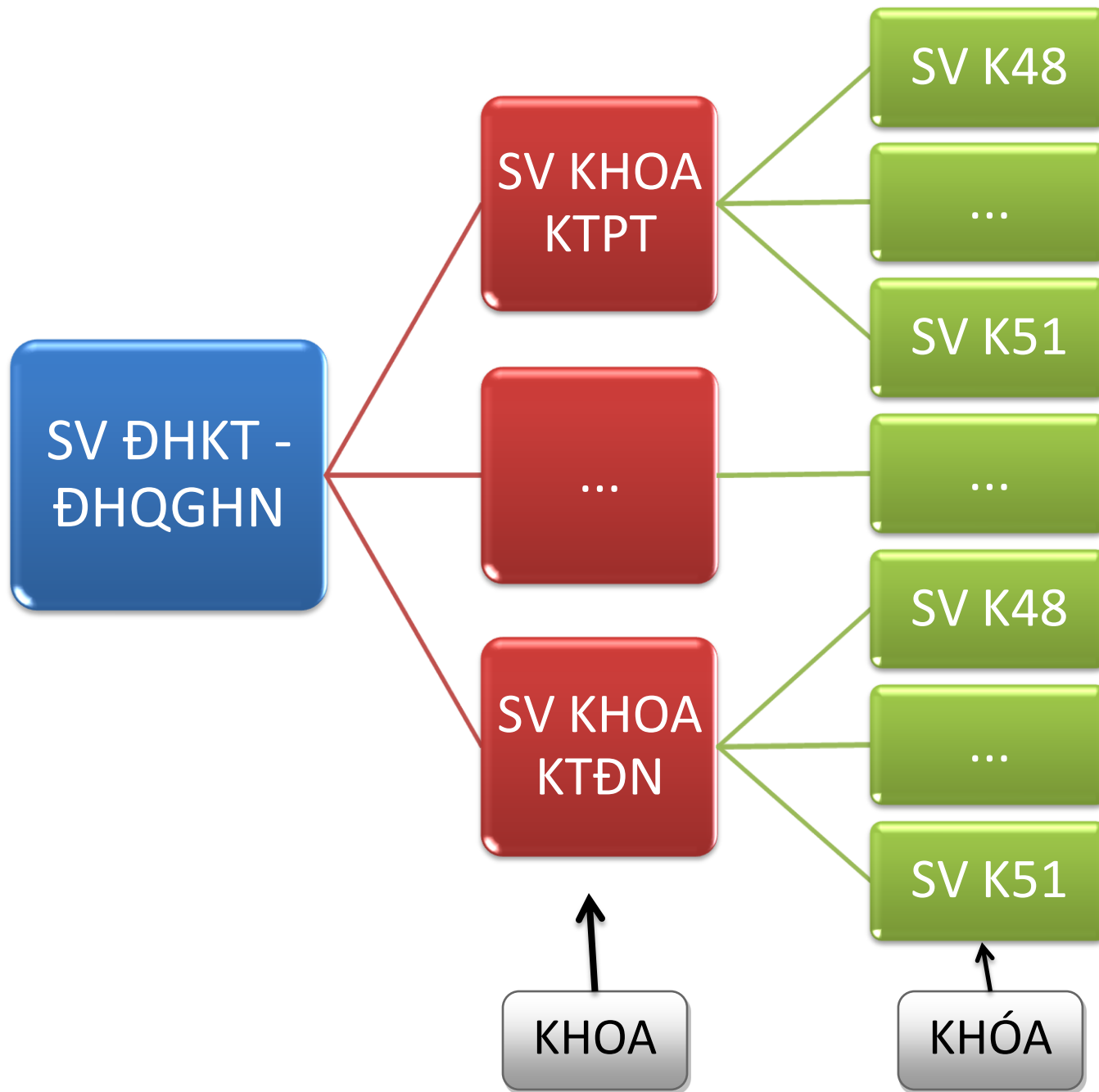
Các loại phân tổ thống kê



Các bước tiến hành phân tổ thống kê

Khái niệm phân tổ thống kê

Phân tổ thống kê là căn cứ vào một (hay một số) tiêu thức nào đó để tiến hành phân chia các đơn vị của hiện tượng nghiên cứu thành các tổ (và các tiểu tổ) có tính chất khác nhau



Ý nghĩa phân tổ thống kê

Có ý nghĩa trong cả quá trình nghiên cứu thống kê

- *Giai đoạn điều tra thống kê*: chia các đơn vị điều tra thành những bộ phận có đặc điểm tính chất khác nhau, là cơ sở lựa chọn các đơn vị điều tra thực tế
- *Giai đoạn tổng hợp thống kê*: là phương pháp cơ bản để tiến hành tổng hợp thống kê, giúp chỉnh lý, hệ thống hóa tài liệu thu thập
- *Giai đoạn phân tích thống kê*: là cơ sở để vận dụng các phương pháp phân tích thống kê khác

Nhiệm vụ phân tổ thống kê

- Phân chia thành các loại hình KTXH, từ đó thấy được đặc trưng của từng loại hình, sự vận động và phát triển của hiện tượng. Thường phân tổ theo tiêu thức thuộc tính -> [Phân tổ phân loại](#)
- Biểu hiện kết cấu của hiện tượng nghiên cứu: cho biết các bộ phận và tỷ trọng của bộ phận trong tổng thể -> [Phân tổ kết cấu](#)
- Nghiên cứu mối liên hệ giữa các tiêu thức (nguyên nhân – kết quả). Thường tiến hành phân tổ theo tiêu thức nguyên nhân, sau đó trong mỗi tổ sẽ tính các trị số bình quân của tiêu thức còn lại -> [Phân tổ liên hệ](#)



Phân tổ phân loại	Khu vực và thành phần kinh tế	Năm 2004 (DN)	Năm 2005 (DN)	Năm 2006 (DN)	Năm 2007 (DN)
	Nhà nước	4596	4086	3706	3494
	Ngoài nhà nước	84003	105167	123392	147316
	Có vốn đầu tư nước ngoài	3156	3697	4220	4961

Phân tổ kết cấu	Khu vực và thành phần kinh tế	Năm 2004 (DN)	Năm 2005 (DN)	Năm 2006 (DN)	Năm 2007 (DN)
	Nhà nước	5%	3,6%	2,8%	2,2%
	Ngoài nhà nước	91,6%	93,1%	94%	94,6%
	Có vốn đầu tư nước ngoài	3,4%	3,3%	3,2%	3,2%

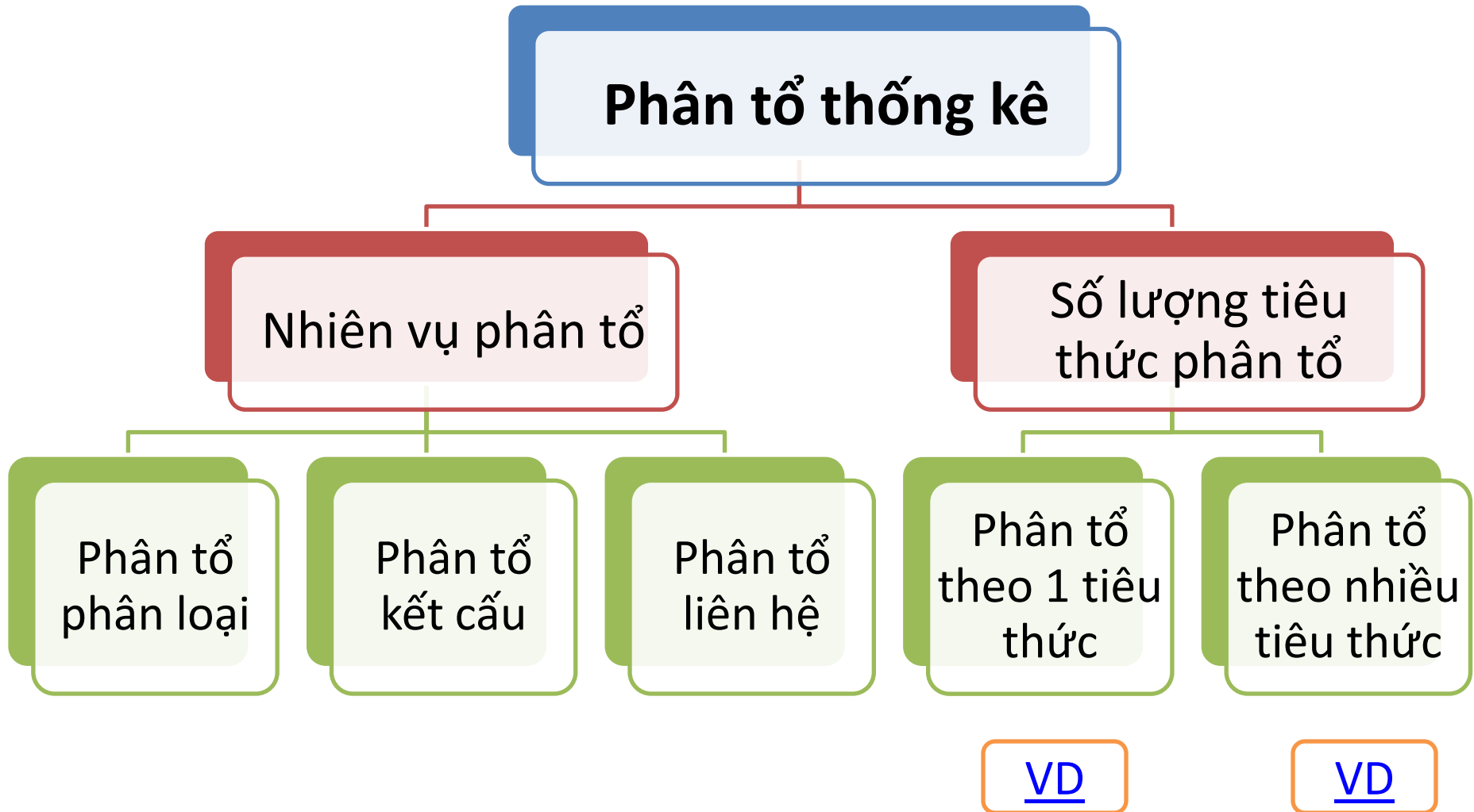


Phân tổ liên hệ	Tuổi nghề (năm)	< 5	5 - 10	10 - 25	15 - 20	≥ 20
	Năng suất lao động bình quân năm (tấn)	63	82	92	100	105

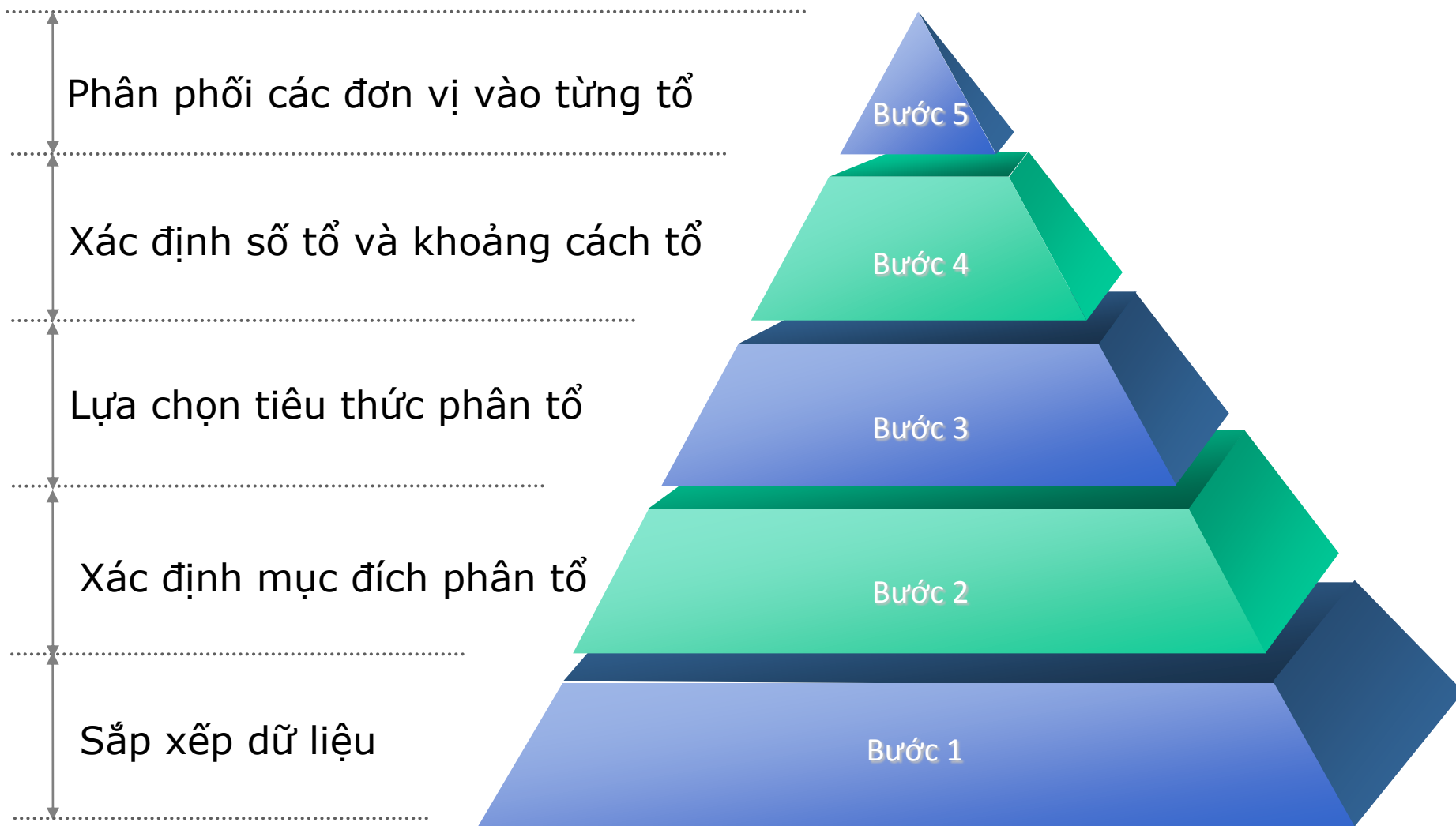
Phân tổ theo nhiều tiêu thức	Trình độ kỹ thuật	Tuổi nghề (năm)	Số công nhân
	Đã được đào tạo	<10	55
		10 – 20	55
		≥ 20	10
	Chưa được đào tạo	<10	40
		10 – 20	30
		≥ 20	10



Các loại phân tổ thống kê



Các bước phân tổ thống kê



Sắp xếp dữ liệu

- **Khái niệm:** Là sự sắp xếp theo một trật tự nhất định.
- **Cách sắp xếp:**
 - + Tiêu thức số lượng: tăng hoặc giảm dần
 - + Tiêu thức thuộc tính: theo vần a, b, c hoặc một trật tự logic
- **Tác dụng:**
 - + Cho sự nhận xét sơ bộ ban đầu về đặc điểm phân phối của hiện tượng
 - + Là cơ sở để phân tổ thống kê

Kỹ thuật sơ đồ thân lá

VD: Kết quả điểm Tiếng Anh của 30 sinh viên như sau:

50	65	72	81	42
37	91	82	77	68
78	65	72	87	25
37	98	76	55	53
27	95	86	71	60
50	100	67	62	75

<u>Thân</u>	<u>Lá</u>
2	5 7
3	7 7
4	2
5	0 5 3 0
6	5 8 5 0 7 2
7	2 7 8 2 6 1 5
8	1 2 7 6
9	1 8 5
10	0

R Xác định mục đích phân tổ

Trả lời câu hỏi: phân tổ để làm gì?



Lựa chọn tiêu thức phân tổ

⊕ ***Tiêu thức phân tổ*** là tiêu thức được chọn làm căn cứ để tiến hành phân tổ thống kê.

⊕ **Căn cứ lựa chọn tiêu thức phân tổ:**

- ✓ Dựa vào mục đích nghiên cứu và điều kiện thực tế
- ✓ Căn cứ vào đặc điểm, tính chất và điều kiện lịch sử cụ thể.

R Xác định số tổ và khoảng cách tổ

⊕ Phân tổ theo tiêu thức thuộc tính

- ✓ Tiêu thức có ít biểu hiện: 1 biểu hiện = 1 tổ
- ✓ Tiêu thức có nhiều biểu hiện: ghép một số biểu hiện thành một tổ theo nguyên tắc các tổ nhỏ ghép lại với nhau phải giống nhau (hoặc gần giống nhau) về tính chất, về giá trị sử dụng, về loại hình...

Xác định số tổ và khoảng cách tổ

⊕ Phân tổ theo tiêu thức số lượng

✓ **Phân tổ không có khoảng cách tổ:** 1 lượng biến = 1 tổ (lượng biến rời rạc, số lượng các lượng biến ít). ([VD](#))

✓ **Phân tổ có khoảng cách tổ:** mỗi tổ sẽ bao gồm một phạm vi lượng biến, với hai giới hạn: giới hạn trên và giới hạn dưới.

Giới hạn dưới: lượng biến nhỏ nhất để hình thành tổ đó.

Giới hạn trên: lượng biến lớn nhất mà nếu vượt quá giới hạn đó thì hình thành một tổ mới.

Khoảng cách tổ (h) = giới hạn trên – giới hạn dưới ([VD](#))

Các hình thức phân tổ có khoảng cách tổ

	Trường hợp áp dụng	Khoảng cách tổ
Có khoảng cách tổ bằng nhau (<u>VD</u> , <u>VD thực hành</u>)	Hiện tượng biến thiên tương đối đồng nhất.	$h = \frac{X_{\max} - X_{\min}}{n}$ Trong đó: h : Khoảng cách tổ $X_{\max}$: Giới hạn trên $X_{\min}$: Giới hạn dưới n : Số tổ cần chia
Có khoảng cách tổ không bằng nhau (<u>VD</u>)	Tất cả các hiện tượng.	Phụ thuộc mục đích nghiên cứu, đặc điểm của hiện tượng.
Phân tổ mở (<u>VD</u>)	Kết quả thu được là tổ đầu tiên không có giới hạn dưới hoặc tổ cuối cùng không có giới hạn trên.	Quy ước là khoảng cách tổ của tổ gần nhất.



Ví dụ

DN có 500 lao động

Lương_max=10 triệu đồng

Lương_min=1 triệu đồng

Chia thành 6 tổ đều nhau.

$$h = \frac{10-1}{6} = 1.5 \text{ (triệu đồng)}$$

Tiền lương (triệu đồng)
1.0 - 2.5
2.5 - 4.0
4.0 - 5.5
5.5 - 7.0
7.0 - 8.5
8.5 - 10.0

Quy ước: Nếu một đơn vị điều tra có trị số lượng biến vừa là giới hạn trên của 1 tổ, vừa là giới hạn dưới của tổ liền kề sau đó thì đơn vị điều tra thuộc vào tổ đứng sau.



Dãy số phân phối

☆ Khái niệm:

Là kết quả của phân tổ thống kê, bao gồm 2 thành phần chính là biểu hiện và số đơn vị của tổng thể tương ứng với biểu hiện đó (tần số của biểu hiện) ([VD](#))

☆ Phân loại:

- **Dãy số thuộc tính:** là kết quả của phân tổ theo tiêu thức thuộc tính.
- **Dãy số lượng biến:** là kết quả của phân tổ theo tiêu thức số lượng.

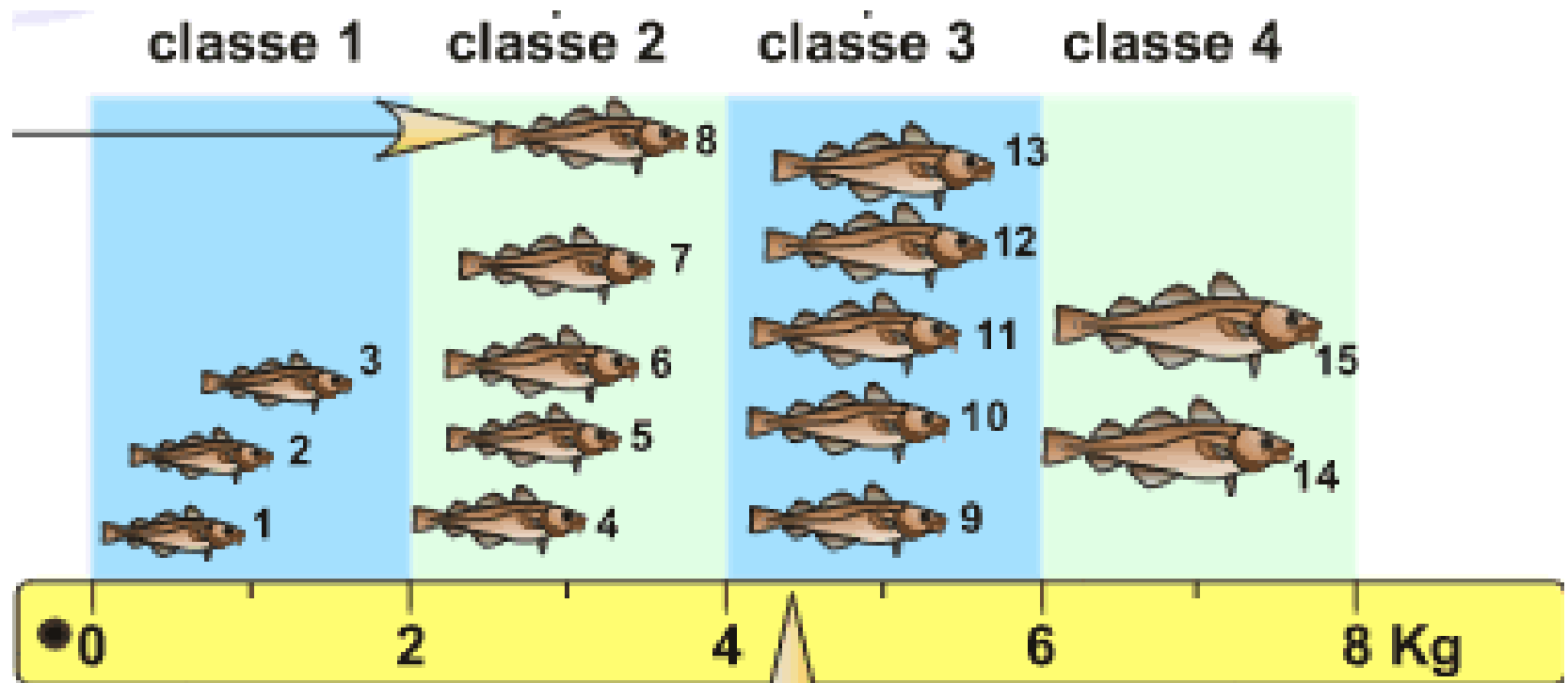
Dãy số lượng biến

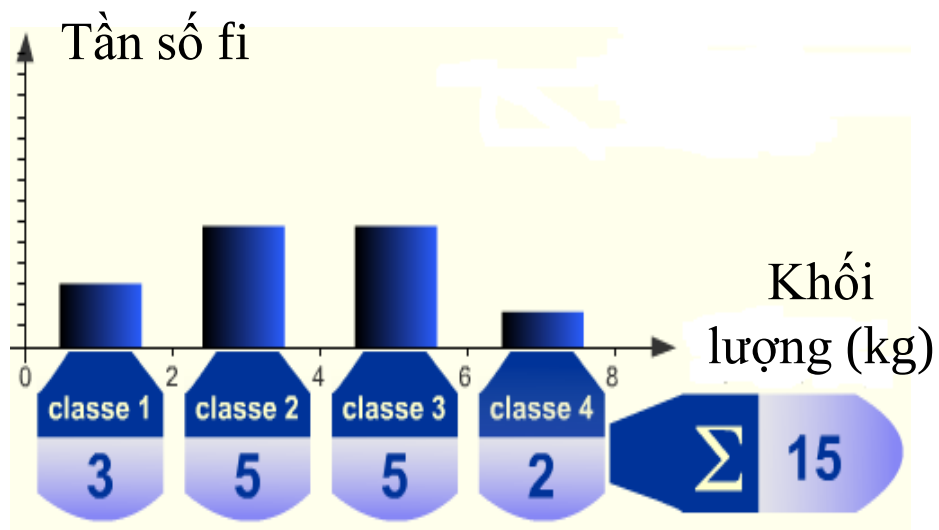
<u>Lượng biến</u> (x_i)	Tần số (f_i)	Tần suất ($d_i = f_i / \sum_1^n f_j$)	Tần số tích lũy ($S_i = \sum_1^i f_j$)
x_1	f_1	d_1	f_1
x_2	f_2	d_2	$f_1 + f_2$
...	...	...	
x_n	f_n	d_n	$f_1 + f_2 + \dots + f_{n-1} + f_n$

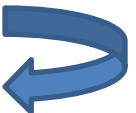
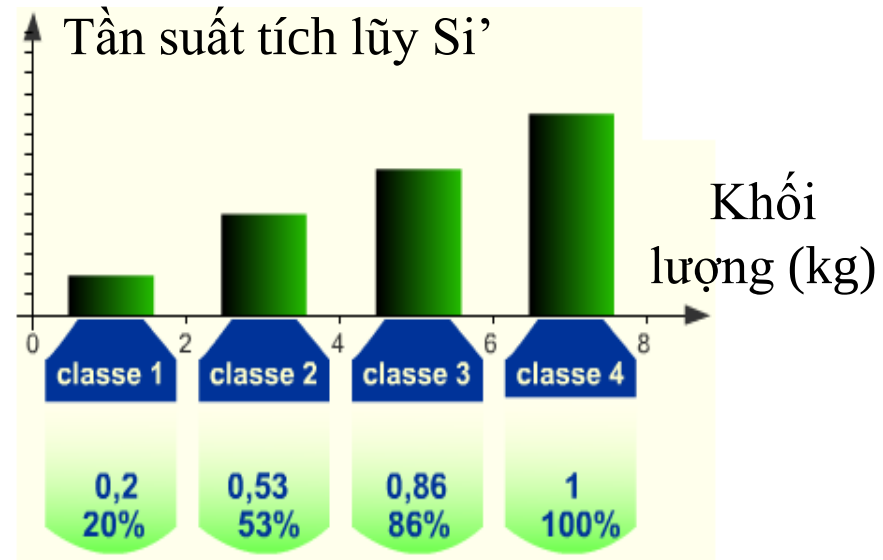
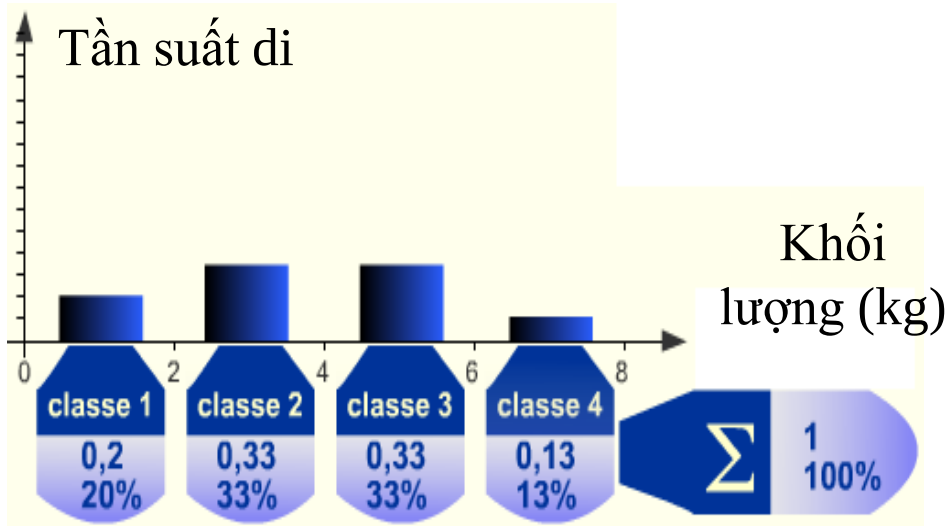
+ Nhận xét: $\sum_1^n f_j =$ số đơn vị của tổng thể; $\sum_1^n d_j = 1$. [VD](#)

+ Khi phân tổ có khoảng cách tổ không bằng nhau, để đánh giá mức độ tập trung từng tổ, sử dụng mật độ phân phối: $m_i = f_i/h_i$. [VD](#)

Giả sử người ta đánh bắt được 15 con cá hồi đại dương

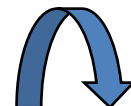






Xác định lượng biến (x_i) từ một tổ cho trước

Kiểu tổ	Kí hiệu	Điều kiện áp dụng	x_i
Tổ chỉ có 1 lượng biến (a) (VD1 , VD2)	a	Mọi tổ	a
Tổ có giới hạn trên (a) và giới hạn dưới (b) (VD)	a – b	Mọi tổ	$(a + b)/2$
Tổ mở (VD)	< a	Tổ liền kề sau nó có kí hiệu dạng a – c	$a - (c - a)/2$
	>= b	Tổ liền kề trước nó có kí hiệu dạng d – b	$b + (b - d)/2$



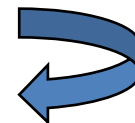
VD: Xác định dãy số lượng biến NSLĐ của 40 công nhân biết:

NSLĐ (sp) (x_i)	Số LĐ (ng) (f_i)	d_i (%)	S_i (ng)
50	3	7.5	3
55	5	12.5	8
60	10	25	18
65	12	30	30
70	7	17.5	37
72	3	7.5	40
Tổng	40	100	

Câu hỏi:

- Con số 12,5 trong cột d_i có ý nghĩa gì?
- Con số 18 trong cột S_i nói lên điều gì?
- Người đứng thứ 20 có NSLĐ là bao nhiêu?

$$d_i = f_i / \sum_1^n f_j \quad S_i = \sum_1^i f_j$$



VD: Điểm Tiếng Anh của sinh viên trong 1 lớp như sau:

2	5	7	9	10
6	6	3	4	5
7	5	7	8	7
8	2	6	7	6
7	7	8	9	5
3	6	8	9	6

Điểm	Số SV
< 5	5
>= 5	25

↓
Phân tổ mở

xi	Điểm	Số SV
4	< 5	5
6	5 - 7	10
8	7 - 9	11
10	>= 9	4

Không có khoảng
cách tổ



Điểm	Số SV
2	2
3	2
4	1
5	4
6	6
7	7
8	4
9	3
10	1

Có khoảng
cách tổ



Điểm	Số SV	mi
2 - 5	5	1,7
5 - 10	25	5

↓
Không bằng nhau

Điểm	Số SV
2 - 4	4
4 - 6	5
6 - 8	13
8 - 10	8

↓
Bằng nhau

