

# CHƯƠNG 3

## CÁC THAM SỐ ĐO LƯỜNG THỐNG KÊ

# Các tham số đo lường thống kê

Đo mức độ đại biểu

Số bq cộng

Số bq nhân

Mốt

Trung vị

Đo độ biến thiên

Khoảng biến thiên

Phương sai

Độ lệch tiêu chuẩn

Hệ số biến thiên

# I – Các tham số đo mức độ đại biểu

# 1 – Ý nghĩa của các tham số đo mức độ đại biểu

- Nêu lên đặc điểm chung của hiện tượng KT- XH số lớn.
- So sánh các hiện tượng không cùng qui mô
- Nghiên cứu quá trình biến động qua thời gian, quan sát xu hướng phát triển cơ bản của hiện tượng.
- Chiếm vị trí quan trọng trong việc vận dụng các phương pháp phân tích và dự đoán TK.

*Chú ý: Các tham số chỉ có ý nghĩa khi được tính ra từ tổng thể đồng chất.*

## 2 – Các tham số đo mức độ đại biểu

## 2.1 - Số bình quân cộng

(Bình quân số học – arithmetic mean)

a/ Điều kiện vận dụng : Các lượng biến của tiêu thức có quan hệ tổng.

b/ Công thức chung:

$$\text{Số bình quân cộng} = \frac{\text{Tổng các lượng biến của tiêu thức nghiên cứu}}{\text{Tổng số đơn vị của tổng thể}}$$

Cụ thể:

- TH các đơn vị không được phân tổ

$$\bar{x} = \frac{\sum x_i}{n}$$

Bình quân cộng  
giản đơn

- TH các đơn vị được phân tổ :

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

Bình quân cộng  
gia quyền

Chú ý:

- Nếu trong CT, quyền số nói lên tầm quan trọng của từng lượng biến đổi với toàn bộ tổng thể, số bình quân đó gọi là số bq có trọng số.



- Nếu quyền số là tỷ trọng mỗi tổ chiếm trong tổng thể:

$$\bar{x} = \frac{\sum x_i d_i}{100}$$

Nếu di tính bằng %

$$\bar{x} = \sum x_i d_i$$

Nếu di tính bằng lần

- TH các lượng biến có tần số bằng nhau, dùng CT số bình quân cộng giản đơn

- TH dãy số lượng biến có khoảng cách tổ

VD 1 : Tính NSLĐ bq của CN 1 DN biết

NSLĐ (c/giờ)	Số CN (người)
20-30	10
30-40	15
40-50	28
50-60	32
60-70	10
70-80	5

[Slide 22](#)

[Slide 30](#)

TH dãy số lượng  
biến có khoảng  
cách tổ mở, khi  
tính trị số giữa  
phải căn cứ vào  
khoảng cách tổ  
gần chúng nhất  
để tính.

Lượng biến	Trị số giữa
< 500	450
500 – 600	550
.....	
800 – 1000	900
> 1000	1100

- TH chỉ biết từng lượng biến ( $x_i$ ) và tổng các lượng biến  $M_i$  ( $M_i = x_i \cdot f_i$ ):

$$\bar{x} = \frac{\sum M_i}{\sum x_i}$$

Số bình quân điều  
hoà gia quyền

- Nếu  $M_1 = M_2 = \dots = M_n$

$$\bar{x} = \frac{n}{\sum \frac{1}{x_i}}$$

Số bình quân  
điều hoà giản đơn

# Bài tập

1. Có số liệu về tình hình xuất khẩu của một doanh nghiệp như sau:

Đợt XK	Giá xuất khẩu (USD/tấn)	KLXK (tấn)
1	250	200
2	280	300
3	300	500

Xác định giá xuất khẩu bình quân, khối lượng XK bình quân và giá trị XK bình quân 3 đợt.

# Bài tập

2. Một DN trong 3 năm bỏ ra một số tiền như nhau để mua NVL với giá NVL mỗi năm như sau:

Năm 1: 1 trđ/tấn

Năm 2: 1,5 trđ/tấn

Năm 3: 1,7 trđ/tấn

Xác định giá NVL bình quân trong 3 năm?

3. Một người đi từ A đến B với tốc độ bình quân 30 km/h và quay trở về (từ B đến A) cũng với con đường đó với tốc độ bình quân 60 km/h. Vậy tốc độ bq mà người đó đi cả 2 lượt là?

## 2.2 - Số bình quân nhân

(Bình quân hình học – geometric mean)

a/ Điều kiện vận dụng : Các lượng biến có QH tích số.

b/ CT:

- Số bq nhân giản đơn

$$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

- Số bq nhân gia quyền

$$\bar{x} = \sqrt[f_1 + f_2 + \dots + f_n]{x_1^{f_1} \cdot x_2^{f_2} \cdot \dots \cdot x_n^{f_n}}$$

VD :

Một doanh nghiệp trong 10 năm có các tốc độ phát triển như sau:

- 5 năm đầu có tốc độ phát triển mỗi năm là 115%
- 2 năm tiếp theo có tốc độ phát triển mỗi năm là 112%
- 3 năm cuối có tốc độ phát triển mỗi năm là 120%,

Tính tốc độ phát triển bình quân của doanh nghiệp trong 10 năm nói trên.



# Đặc điểm chung của số bình quân

- Tất cả các lượng biến đều tham gia vào tính toán → Chịu ảnh hưởng bởi những lượng biến đột xuất (quá lớn hoặc quá nhỏ).
- San bằng chênh lệch giữa các lượng biến.

## 2.3 - Một (mode) – $M_0$

### a/ KN

- **Đối với dãy số không có khoảng cách tổ:**  
Một là lượng biến hoặc biểu hiện được gặp nhiều nhất trong dãy số phân phối.

Cách xác định  $M_0$

Xác định lượng biến hoặc biểu hiện có tần số lớn nhất trong dãy số phân phối, đó chính là  $M_0$ .

VD 3:	$\underline{x}_i$	$\underline{f}_i$
	21	5
	25	8
	30	15
	32	22
	35	30
	40	25
	42	26

$$M_0 =$$

- Đối với dãy số có khoảng cách tử (Chỉ có ở dãy số lượng biến) :

Một là lượng biến trên đó chứa mật độ phân phối lớn nhất, tức là xung quanh lượng biến đó tập trung tần số nhiều nhất.

## Cách xác định $M_0$ của dãy số có khoảng cách tổ

B1 : Xác định tổ chứa  $M_0$

+ Nếu các tổ có khoảng cách tổ bằng nhau: Tổ nào có tần số lớn nhất là tổ chứa  $M_0$ .

+ Nếu các tổ có khoảng cách tổ không bằng nhau cần tính mật độ phân phối  $D_i$  ( $D_i = f_i/h_i$ ). Tổ nào có mật độ phân phối lớn nhất là tổ chứa  $M_0$ .

B2 : Tính giá trị gần đúng của  $M_0$  theo công thức:

$$M_0 = x_{M_0 \min} + h_{M_0} \cdot \frac{f_{M_0} - f_{M_0-1}}{(f_{M_0} - f_{M_0-1}) + (f_{M_0} - f_{M_0+1})}$$

$$M_0 = x_{M_0 \min} + h_{M_0} \cdot \frac{D_{M_0} - D_{M_0-1}}{(D_{M_0} - D_{M_0-1}) + (D_{M_0} - D_{M_0+1})}$$

VD : [Slide 10](#)

- Chú ý : Trường hợp dãy số phân phối có các tần số xấp xỉ bằng nhau hoặc có quá nhiều điểm tập trung thì không nên tính một.

b/ Đặc điểm của  $M_0$  :

- + Dễ xác định và có khả năng xác định nhanh
- + Không bị ảnh hưởng bởi các giá trị đột xuất (quá lớn hoặc quá nhỏ) của dãy số phân phối), vì vậy kém nhạy bén với sự biến thiên của tiêu thức.

c/ Tác dụng:

- + Dùng để bổ sung hoặc thay thế số bình quân trong TH tính số bq gặp khó khăn.
- + Dùng nhiều trong lý thuyết phục vụ đám đông.



## 2.4 – Trung vị (Median) – $M_e$

(Chỉ dùng với dãy số lượng biến)

a/ KN

Trung vị là lượng biến của đơn vị đứng vị trí chính giữa trong dãy số lượng biến, chia số đơn vị trong dãy số thành 2 phần bằng nhau.

VD 4:

- Dãy số : 20    21    25    27    30

Me =

- Dãy số:    xi    fi

21    2

23    6

25    1

26    3

30    1

Me = ?

Dãy số : 27    25    21    30    28

Me = ?

- Chú ý:

- + Trung vị là lượng biến của đơn vị đứng ở vị trí chính giữa chứ không phải lượng biến đứng chính giữa.
- + Khi xác định trung vị, dãy số phải được sắp xếp theo thứ tự nhất định (từ nhỏ đến lớn hoặc ngược lại).

## b/ Cách xác định trung vị

- Xác định đơn vị đứng ở vị trí chính giữa
- + Nếu số đơn vị tổng thể là số lẻ ( $n = 2m + 1$ ) thì đơn vị đứng ở vị trí chính giữa là đơn vị thứ  $m + 1$ .
- + Nếu số đơn vị tổng thể là số chẵn ( $n = 2m$ ) thì đơn vị đứng ở vị trí chính giữa là đơn vị thứ  $m$  và  $m + 1$

- Tính trung vị:

+ Đối với dãy số không có khoảng cách tổ, trung vị là lượng biến của đơn vị đứng ở vị trí chính giữa

Nếu số đơn vị tổng thể là số lẻ :  $M_e = x_{m+1}$

Nếu số đơn vị tổng thể là chẵn :

$$M_e = (x_m + x_{m+1}) : 2$$

+ Đối với dãy số có khoảng cách tổ, cần qua 2 bước

B1 : Xác tổ chứa trung vị : là tổ chứa lượng biến của đơn vị đứng ở vị trí chính giữa .

B2 : Tính trung vị theo công thức (giả định phân phối đều đặn):

$$M_e = x_{M_{e \min}} + h_{M_e} \cdot \frac{\frac{\sum f_i}{2} - S_{M_{e-1}}}{f_{M_e}}$$

Tính cho VD [Slide 10](#)

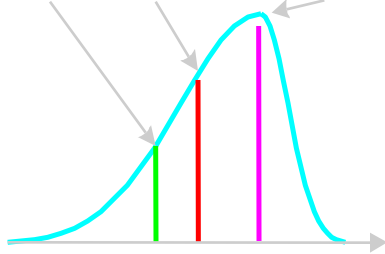
c/ Tác dụng của  $M_e$  :

- Bổ sung hoặc thay thế số bình quân khi cần thiết.
- Khi kết hợp với số bq cộng, một, trung vị có thể nêu lên đặc trưng của dãy số phân phối, cụ thể:

+

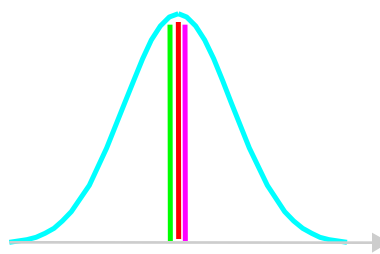
**Lệch trái**

Mean Median Mode



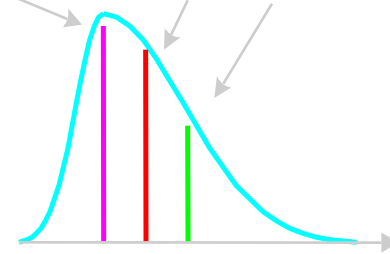
**Đối xứng**

Mean = Median = Mode



**Lệch phải**

Mode Median Mean



- Trung vị được ứng dụng nhiều trong công tác kỹ thuật và phục vụ công cộng (vì  $\sum |x_i - M_e| f_i = \min$ ).

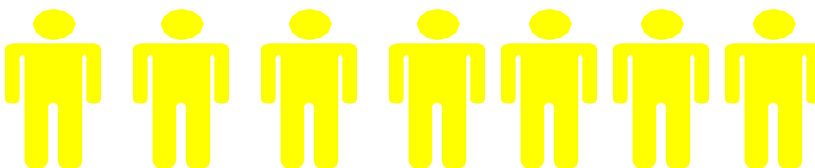
# Trong các tham số đo mức độ đại biểu, tham số nào đo mức độ đại biểu tốt nhất?

VD :

6000 \$ 

2000 \$ 

300 \$ 

100 \$ 

... Người lao động cho rằng mức lương thấp, phần lớn chỉ đạt 100\$/tháng.

... Chủ doanh nghiệp nói rằng mức lương khá cao, bình quân đạt 1000\$/tháng!



## II – Các tham số đo độ biến thiên của tiêu thức

# 1 – Ý nghĩa của độ biến thiên tiêu thức

- Đánh giá trình độ đại biểu của số bình quân
- Cho thấy độ phân tán, đánh giá độ đồng đều giữa các lượng biến trong tổng thể
- Kiểm tra chất lượng sản phẩm.
- Dùng nhiều trong các nghiên cứu thống kê khác

## 2 – Các chỉ tiêu đo độ biến thiên của tiêu thức

## 2.1 - Khoảng biến thiên ( R ) – (Range)

a/ KN : Là chênh lệch giữa lượng biến lớn nhất và lượng biến nhỏ nhất của tiêu thức.

b/ CT :  $R = X_{\max} - X_{\min}$

VD 5 : Tổ 1 : 45   50   55   60   65    $R_1 = ?$

Tổ 2: 51   53   55   57   59    $R_2 = ?$

c/ ưu điểm : Tính toán đơn giản, cho NX nhanh về độ biến thiên của tổng thể.

Nhược điểm: Cho NX không chính xác khi có các lượng biến đột xuất (quá lớn hoặc quá nhỏ).

## 2.2 – Phương sai ( $\sigma^2$ ) – (Variance)

a/ KN : Là số bình quân cộng của bình phương các độ lệch giữa lượng biến với bình quân các lượng biến đó.

b/ Công thức :

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - (\bar{x})^2 \quad (a)$$

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{\sum f_i} = \frac{\sum x_i^2 \cdot f_i}{\sum f_i} - (\bar{x})^2 \quad (b)$$

## VD5: Xác định thu nhập bq và phương sai về thu nhập từng tổ phân xưởng và NX

Tổ 1		Tổ 2	
Thu nhập (1000 đ)	Số CN (người)	Thu nhập (1000 đ)	Số CN (người)
1200	3	1500	6
1500	8	1800	11
2100	10	2100	7
2200	6	2200	4
2500	3	2500	2

c/ Tác dụng :

- Biểu hiện độ biến thiên tiêu thức
- Dùng nhiều trong phân tích thống kê như tính hệ số tương quan, xác định cỡ mẫu điều tra...

d/ Nhược điểm:

- Khuếch đại sai số
- Đơn vị tính toán không phù hợp.

## 2.3 - Độ lệch tiêu chuẩn ( $\sigma$ ) (Standard deviation)

a/ KN : Là căn bậc hai của phương sai

b/ Tác dụng:

- Là một trong những chỉ tiêu hoàn thiện nhất để đo độ biến thiên tiêu thức của một tổng thể hoặc so sánh độ biến thiên của các tổng thể cùng loại
- Dùng nhiều trong các phân tích thống kê.
- Cho biết sự phân phối của các lượng biến trong một tổng thể (dựa vào định lý Chebyshev)



Theo định lý Chebyshev: có ít nhất  $(1 - 1/k^2)\%$  số các lượng biến nằm trong khoảng

$(\bar{x} \pm k\sigma)$  với  $k$  là một số bất kỳ lớn hơn 1, nghĩa là **với 1 phân phối bất kỳ** có:

75% số các lượng biến nằm trong khoảng  $\bar{x} \pm 2\sigma$

89% số các lượng biến nằm trong khoảng  $\bar{x} \pm 3\sigma$

## 2.4 - Hệ số biến thiên (V) (Coefficient of variation).

a/ TH sử dụng :

- Giá trị bình quân của 2 tổng thể đưa ra so sánh khác nhau nhiều.
- So sánh độ biến thiên của 2 hiện tượng khác nhau (đơn vị tính khác nhau).

b/ Khái niệm: Là số tương đối được tính bằng cách so sánh giữa độ lệch tiêu chuẩn với số bình quân cộng.

( Trường hợp không có số bình quân cộng có thể thay bằng Một)

c/ Công thức :

$$V = \frac{\sigma}{x} \times 100 \text{ (\%)}$$

$$V = \frac{\sigma}{M_o} \times 100 \text{ (\%)}$$

Chú ý:

- Khi so sánh 2 hiện tượng phải sử dụng cùng 1 công thức.
- TH dùng V để đánh giá tính chất đại biểu của số bình quân, nếu V vượt quá 40% thì tính chất đại biểu của số bình quân quá thấp, không nên sử dụng số bình quân đó.