

Chương III

Các tham số thống kê

I. Các tham số đo độ tập trung

1. Khái niệm, đặc điểm, điều kiện vận dụng

2. Các loại tham số

- Số bình quân cộng
- Số bình quân nhân
- Mốt (Mode)
- Trung vị (Median)
- Phân vị

II. Các tham số đo độ biến thiên tiêu thức

1. □ nghĩa

2. Các tham số đo độ biến thiên tiêu thức

- Khoảng biến thiên
- Độ trải giữa
- Độ lệch tuyệt đối
- Phương sai
- Độ lệch tiêu chuẩn
- Hệ số biến thiên

I. Số bình quân

1. Số bình quân trong thống kê

a) Khái niệm, đặc điểm của số bình quân

■ Khái niệm

Số bình quân là trị số biểu hiện mức độ đại biểu theo một tiêu thức nào đó của một hiện tượng bao gồm nhiều đơn vị cùng loại.

a) Khái niệm, đặc điểm của số bình quân

■ Đặc điểm

- Số bình quân mang tính tổng hợp và khái quát
- San bằng mọi chênh lệch giữa các đơn vị về trị số của tiêu thức nghiên cứu

b) Điều kiện vận dụng số bình quân

- Chỉ được tính số bình quân cho một tổng thể bao gồm các đơn vị cùng loại
- Số bình quân cần được tính ra từ tổng thể có nhiều đơn vị

❖ Tác dụng của số bình quân

- Số bình quân được sử dụng để phản ánh đặc điểm chung về mặt lượng của hiện tượng kinh tế xã hội số lớn trong điều kiện thời gian, không gian cụ thể
- Số bình quân được sử dụng để so sánh các hiện tượng không cùng quy mô.
- Số bình quân còn được sử dụng trong nghiên cứu các quá trình biến động qua thời gian
- Số bình quân có vị trí quan trọng trong việc vận dụng các phương pháp phân tích thống kê

2. Các loại số bình quân

2.1 Số bình quân cộng

a) Điều kiện vận dụng số bình quân cộng là các lượng biến phải có quan hệ tổng với nhau

Công thức tổng quát:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Quan hệ giữa các lượng biến như thế nào thì được coi là quan hệ tổng?

- Thu nhập CN1 T8/03 so với T7/03 bằng 1,2 lần
- Thu nhập CN2 T8/03 so với T7/03 bằng 1,2 lần
- Tổng 2 giá trị 1,2 lần và 1,1 lần bằng 2,3 lần?
- Thu nhập CN1 tháng 8/03 là 2tr VDN
- Thu nhập CN2 tháng 8/03 là 1tr VDN
- Tổng 2 giá trị trên: 3 tr VND là tổng thu nhập của hai công nhân trong tháng 8/03

Các trường hợp vận dụng cụ thể

- Trường hợp các đơn vị không được phân tổ $\rightarrow$ sử dụng công thức tổng quát
- CT số bình quân cộng giản đơn:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

▪ Trường hợp dãy số đã được phân tổ

- Dãy số đã được phân tổ không có khoảng cách tổ; bao gồm các thành phần: lượng biến, tần số và/hoặc tần suất tương ứng

Ví dụ: thu nhập của tổ CN T2/04 (*triệu VND*)

1.5	1.5	1.0	1.5	1.0	2.0	1.0	2.0
1.5	2.5	1.0	0.6	1.5	1.5	1.5	1.0
2.0	1.5	1.5	2.0	0.6	1.0	2.0	1.5
1.0	1.0	0.6	1.5	2.5	1.0	0.6	1.0
0.6	1.0	1.0	1.0	1.5	1.0	1.0	2.0

Ví dụ

Dãy số sau khi phân tổ

Mức thu nhập (tr\$)	0,6	1,0	1,5	2,0	2,5
Số lượng công nhân (người)	5	15	12	6	2

Nhận xét

Lượng biến $x_1 = 0,6$ (tr) có tần số $f_1 = 5$ có nghĩa là số lần xuất hiện của nó trong tổng thể là 5 lần

Do vậy tổng giá trị của các lượng biến x_1 không phải là 0,6 (tr\$) mà phải là $0,6 * 5 = 3,0$ (tr\$)

x_i (tr\$)	0,6	1,0	1,5	2,0	2,5
f_i (ng-ời)	5	15	12	6	2
Σx_i (tr\$)	3,0	15,0	18,0	12,0	5,0

➤ Dãy số đã được phân tổ không có khoảng cách tổ

■ Thu nhập bình quân:

$$\bar{x} = \frac{\sum_{i=1}^n \sum x_i}{n} = \frac{3 + 15 + 18 + 12 + 5}{6 + 15 + 12 + 6 + 2} = \frac{53}{40} = 1,325 \text{ (tr \$)}$$

■ Công thức tổng quát: (CT bình quân gia quyền với f_i là quyền số)

$$\bar{x} = \frac{\sum_{i=1}^n x_i \times f_i}{\sum_{i=1}^n f_i}$$

Các biến thể của CT bình quân gia quyền

- Khi quyền số là tần suất d_i (%)

$$\bar{x} = \frac{\sum_{i=1}^n x_i d_i}{100}$$

Tại sao?

$$\sum d_i = 100$$

- Khi quyền số là tần suất d_i (lần)

$$\bar{x} = \sum_{i=1}^n x_i d_i$$

Tại sao?

$$\sum d_i = 1$$

Ví dụ: tính giá thành sản xuất bình quân

	Giá thành sản xuất (\$/sp)	Tỷ lệ (%)
Tháng 4/03	155	58
Tháng 5/03	156	32
Tháng 6/03	158	10

Tính giá thành sản xuất bình quân

$$\bar{z} = \frac{\sum_{i=1}^n z_i d_i}{100}$$

$$\bar{z} = \frac{155 * 58 + 156 * 32 + 158 * 10}{100} = 155,62 (\$ / sp)$$

➤ Dãy số lượng biến có khoảng cách tổ

Xét ví dụ:

Tài liệu thống kê khối lượng lương thực bình quân đầu người tại 1 địa phương năm 1995

Khối lượng lương thực bình quân (kg/người)	Số người (người)
400 — 500	100
500 — 600	300
600 — 700	450
700 — 800	800
800 — 900	300
900 — 1000	50

Các bước tiến hành

- Bước 1: tính trị số giữa của từng tổ theo công thức

$$X_i = \frac{X_{i \min} + X_{i \max}}{2}$$

$x_i \min \div x_i \max$	x_i
400 ÷ 500	450
500 ÷ 600	550
600 ÷ 700	650
700 ÷ 800	750
800 ÷ 900	850
900 ÷ 1000	950

Các bước tiến hành

- Bước 2: xác định giá trị của số bình quân bằng công thức bình quân gia quyền

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i}$$

x_i	f_i	$x_i f_i$
450	100	45000
550	300	165000
650	450	292500
750	800	600000
850	300	255000
950	50	47500

Xác định số bình quân cộng cho VD trên

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} = \frac{450 * 100 + 550 * 300 + 650 * 450 + 750 * 800 + 850 * 300 + 950 * 50}{100 + 300 + 450 + 800 + 300 + 50}$$

$$\bar{x} = \frac{1405000}{2000} = 702,5 (kg / ng)$$

Chú ý

- Đối với những dãy số có khoảng cách tổ mở, chúng ta phải căn cứ vào khoảng cách tổ cũng như trị số giữa của tổ sát cạnh đó để tính.

x_i :min — x_i :max (kg)	Dưới 500	500 - 600	600 - 700	700 - 800	800 - 900	900 trở lên
f_i (ng)	100	300	450	800	300	100
x_i	450	550	650	750	850	950

- Biết x_i và tổng các lượng biến $M_i (= x_i f_i)$

Ví dụ:

X N	Sản lượng (sp)	NSLĐ bình quân (sp/CN)
A	21250	425
B	32400	432
C	32550	434

Cách xác định
NSLĐ bình quân

- C1: trước tiên xác định f_i qua M_i và x_i
Sau đó sử dụng CT bình quân gia quyền
- C2: tính trực tiếp, sử dụng CT bình quân cộng điều hoà

■ Công thức bình quân điều hoà:

$$\bar{x} = \frac{\sum_{i=1}^n M_i}{\sum_{i=1}^n \frac{M_i}{x_i}}$$

■ Ví dụ:

$$\bar{x} = \frac{21250 + 32400 + 32550}{\frac{21250}{425} + \frac{32400}{432} + \frac{32550}{434}}$$

$$\Rightarrow \bar{x} = \frac{86200}{50 + 75 + 75} = 431 \text{ (sp / CN)}$$

Bài tập áp dụng

- Hai CN cùng sản xuất 1 loại sp:
 - CN1: làm 2 phút đ-ợc 1 sp
 - CN1: làm 6 phút đ-ợc 1 sp
- Tính thời gian bình quân làm ra 1 sp của 2 CN trên trong các điều kiện sau:
 - a) Cả 2 CN cùng làm việc trong 8 giờ
 - b) CN1 làm 40%, CN2 làm 60% tổng số TG LĐ

Thời gian làm ra 1 sp: x_i

Số sp mỗi CN sản xuất được: f_i

Thời gian sản xuất: $M_i = x_i f_i$

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

a) $M_1 = 8 \times 60$; $x_1 = 2$

$M_2 = 8 \times 60$; $x_2 = 6$

$$\bar{x} = \frac{\frac{60 \times 8}{2} + \frac{60 \times 8}{6}}{\frac{60 \times 8}{2} + \frac{60 \times 8}{6}} = 3 (ph / sp)$$

b) Gọi tổng thời gian sản xuất là T

$\rightarrow M_1 = T \times 40\%$

$M_2 = T \times 60\%$

$$\bar{x} = \frac{\frac{T \times 60\%}{2} + \frac{T \times 40\%}{6}}{\frac{T \times 60\%}{2} + \frac{T \times 40\%}{6}} = 3,33 (ph/sp)$$

Xác định giá, lượng và tỷ giá bình quân

Có tài liệu về tình hình XNK của CT X tháng 8/03

Đợt	1	2	3
Giá xuất khẩu (USD/t)	180	186	185
Lượng xuất khẩu (t)	2200	1800	2000
Tỷ giá USD/VND	15530	15520	15540

Công thức nào sẽ được sử dụng?

$\bar{p}$

$$\bar{p} = \frac{\sum_{i=1}^n p_i q_i}{\sum_{i=1}^n q_i}$$

$\bar{q}$

$$\bar{q} = \frac{\sum_{i=1}^n q_i}{n}$$

$\bar{r}$

$$\bar{r} = \frac{\sum_{i=1}^n r_i p_i q_i}{\sum_{i=1}^n p_i q_i}$$

Tình hình xuất khẩu của công ty X

Giá xuất khẩu bình quân tháng (USD/t)	183.467
Khối l- ượng xuất khẩu bình quân mỗi đợt hàng trong tháng (t)	2000
Tỷ giá xuất khẩu bình quân tháng (USD/VND)	15530,32

2.2 Số bình quân nhân

- Điều kiện vận dụng:

khi các lượng biến có quan hệ tích với nhau

Quan hệ giữa các lượng biến là quan hệ tích khi nhân các lượng biến lại với nhau, thu được kết quả là giá trị có ý nghĩa

VD: Thu nhập của ông B bằng 1,5 lần thu nhập của ông A, còn thu nhập của ông C bằng 1,1 lần thu nhập của ông B $\rightarrow$ thu nhập của ông C bằng $1,1 * 1,5$ thu nhập của ông A

Có tài liệu về tình hình doanh thu của
Công ty A qua các năm (đv: %)

DT 99' so với DT 98'	DT 00' so với DT 99'	DT 01' so với DT 00'	DT 02' so với DT 01'
100	105	115	110

- Các lượng biến liên nhau có quan hệ tích với nhau

Tình hình doanh thu của Công ty A

$$t_1 = \frac{DT_{99}}{DT_{98}} = 100 \%$$

$$t_2 = \frac{DT_{00}}{DT_{99}} = 105 \%$$

$$t_3 = \frac{DT_{01}}{DT_{00}} = 115 \%$$

$$t_4 = \frac{DT_{02}}{DT_{01}} = 110 \%$$

Quan hệ tích giữa các lượng biến

$$t_1 \times t_2 = \frac{DT_{99}}{DT_{98}} \times \frac{DT_{00}}{DT_{99}} = \frac{DT_{00}}{DT_{98}} = 100 \% \times 105 \%$$

$$t_1 \times t_2 \times t_3 = \frac{DT_{99}}{DT_{98}} \times \frac{DT_{00}}{DT_{99}} \times \frac{DT_{01}}{DT_{00}} = \frac{DT_{01}}{DT_{98}}$$

$$t_1 \times t_2 \times t_3 \times t_4 = \frac{DT_{99}}{DT_{98}} \times \frac{DT_{00}}{DT_{99}} \times \frac{DT_{01}}{DT_{00}} \times \frac{DT_{02}}{DT_{01}} = \frac{DT_{02}}{DT_{98}}$$

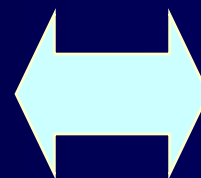
Tình hình doanh thu của Công ty A

$$t_1 \times t_2 \times t_3 \times t_4 = \frac{DT_{99}}{DT_{98}} \times \frac{DT_{00}}{DT_{99}} \times \frac{DT_{01}}{DT_{00}} \times \frac{DT_{02}}{DT_{01}} = \frac{DT_{02}}{DT_{98}}$$

- Tốc độ phát triển DT trong cả giai đoạn 1998 — 2002 là $T = \prod t_i$
- Tốc độ phát triển bình quân về DT trong giai đoạn đó chính là số bình quân nhân

Công thức số bình quân nhân

$$\bar{t} = \sqrt[n]{t_1 \times t_2 \times \dots \times t_n}$$



$$\bar{t} = \sqrt[n]{\prod_{i=1}^n t_i}$$

Tốc độ phát triển DT bình quân của CT X
trong giai đoạn 1998 — 2002:

$$\bar{t} = \sqrt[4]{100 \times 105 \times 115 \times 110} = 107,35 (\%)$$

- Trong trường hợp dãy số lượng biến đã phân
tổ và xuất hiện tần số $\neq 1$, ta có thể áp dụng
CT bình quân nhân gia quyền

Có tài liệu theo dõi về tốc độ phát triển DT
của 1 doanh nghiệp qua 10 năm ('93 ÷ '03)

Năm	1994	1995	1996	1997	1998
t (%)	110	125	115	110	110
Năm	1999	2000	2001	2002	2003
t (%)	110	110	115	125	115

Nếu tính số bình quân nhân theo CT giản đơn:

$$\bar{x}_t = \sqrt[10]{110 \times 125 \times 115 \times 110 \times 110 \times 110 \times 110 \times 115 \times 125 \times 115}$$

→ Cách này quá dài dòng và phức tạp khi số lượng biến nhiều

→ Nên phân tổ số liệu

Tốc độ phát triển DT (%)	110	115	125	x_i
Số năm có tốc độ tương ứng	5	3	2	f_i

Sử dụng CT bình quân nhân gia quyền

$$\bar{t} = \sum f_i \sqrt[n]{\prod_{i=1}^n t_i^{f_i}}$$

$$\bar{t} = 5 + 3 + 2 \sqrt[10]{110^5 \times 115^3 \times 125^2}$$

$$\bar{t} = 10 \sqrt[10]{3.827163119} \times 10^{20} = 114 \text{ (\%)}$$

2.3 Mốt (Mode - Mo)

■ KN

— *Với dãy số không có khoảng cách tổ*, Mo là lượng biến hoặc biểu hiện có tần số lớn nhất

— *Với dãy số có khoảng cách tổ*, Mo là lượng biến có mật độ phân phối lớn nhất (xung quanh đó tập trung nhiều đơn vị tổng thể nhất)

■ Tác dụng

— Biểu hiện mức độ phổ biến nhất

— Không san bằng, bù trừ chênh lệch giữa các lượng biến

— Không chịu ảnh hưởng của các lượng biến đột xuất

— Có nhiều ứng dụng thực tế

■ Phương pháp tính Mo

- ✧ TH1: Dãy số phân tử không có khoảng cách tử: Lượng biến có tần số lớn nhất chính là Mo
- ✧ TH2: Dãy số phân tử có khoảng cách tử:
 - ↳ Xác định vị trí của Mo
 - ↳ Xác định giá trị gần đúng của Mo

⇒ Xác định vị trí của Mo

- ✓ Các tổ có khoảng cách tổ đều nhau
→ Tổ có tần số (tần suất) lớn nhất là tổ chứa Mo
- ✓ Các tổ có khoảng cách tổ không đều nhau
→ Tổ có mật độ phân phối lớn nhất là tổ chứa Mo

⇒ Xác định giá trị gần đúng của M_o

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

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

Với:

- $x_{M_o \min}$: giới hạn dưới của tổ chứa M_o
- h_{M_o} : khoảng cách tổ của tổ chứa M_o
- $f_{M_o}(D_{M_o})$: tần số (mật độ) của tổ chứa M_o
- $f_{M_o-1}(D_{M_o-1})$: tần số (mật độ) của tổ liền trước tổ chứa M_o
- $f_{M_o+1}(D_{M_o+1})$: tần số (mật độ) của tổ liền sau tổ chứa M_o

Nhận xét tình hình lương thực tại địa phương

Khối lượng lương thực bình quân (kg/người)	Số người (người)
400 — 500	10
500 — 600	30
600 — 700	45
700 — 800	80
800 — 900	30
900 — 1000	5

Xác định Mo?

Nhận xét

- Dãy số phân phối có khoảng cách tổ đều nhau $\rightarrow$ xác định Mo dựa vào tần số
- Vị trí của Mo là tổ thứ 4 (khối lượng 700-800 kg) vì $f_4 = 80$ (max)

$X_i \text{ min} - X_i \text{ max}$ (kg/người)	f_i (người)
400 — 500	10
500 — 600	30
600 — 700	45
700 — 800	80
800 — 900	30
900 —	5

Áp dụng CT1 tính giá trị gần đúng của M_o

$$x_{M_o} \min = 700$$

$$h_{M_o} = 100$$

$$f_{M_o} = 80$$

$$f_{M_o-1} = 45$$

$$f_{M_o+1} = 30$$

$$M_o = 700 + 100 \times \frac{80 - 45}{(80 - 45) + (80 - 30)} = 741,2 (kg / ng)$$

2.4 Trung vị □ Me (Median)

■ KN

Trung vị là long biến của đơn vị đứng ở vị trí chính giữa trong dãy số long biến

■ Tính chất

- Trung vị phân chia dãy số long biến thành hai phần có số long đơn vị tổng thể bằng nhau.
- Tổng các độ chênh lệch tuyệt đối giữa các long biến với trung vị là một trị số nhỏ nhất (so với số bình quân hay Mo)

2.4 Trung vị □ Me (Median)

■ Tác dụng

- Trung vị không san bằng, bù trừ chênh lệch giữa các lượng biến → dùng trung vị để bổ sung hoặc thay thế số bình quân cộng
- Tính chất 2 được ứng dụng trong nhiều công tác kỹ thuật và phục vụ công cộng

2.4 Trung vị □ Me (Median)

- Phương pháp xác định trung vị:
 - **Bước 1:** Xác định vị trí chính giữa (vị trí của đơn vị đứng ở vị trí chính giữa).
vị trí trung vị là đơn vị thứ $(n+1)/2$

2.4 Trung vị □ Me (Median)

- **Bước 2:** Xác định trung vị:
- + Đối với dãy số lượng biến không có khoảng cách tổ
 - Nếu số đơn vị là lẻ thì $Me = x_{m+1}$
 - Nếu số đơn vị là chẵn thì

$$Me = \frac{x_m + x_{m+1}}{2}$$

Ví dụ: Theo dõi DT của 10 cửa hàng thuộc công ty tm trong tháng 2/03 (đv: tỷ VND)

0,5	1,6	0,8	1,1	0,3
0,9	2,1	1,2	1,3	1,4

Dãy số sắp xếp

0,3	0,5	0,8	0,9	1,1
1,2	1,3	1,4	1,6	2,1

2.4 Trung vị $\square$ Me (Median)

- + Đối với dãy số lượng biến có khoảng cách tổ
- Xác định tổ chứa trung vị (tổ chứa đơn vị đứng ở vị trí chính giữa)
- Xác định GT gần đúng của trung vị theo CT

$$Me = x_{Me \min} + h_{Me} \frac{\frac{\sum f_i}{2} - S_{Me-1}}{f_{Me}}$$

Nhận xét tình hình lương thực tại địa phương

Khối lượng lương thực bình quân (kg/người)	Số người (người)
400 — 500	10
500 — 600	30
600 — 700	45
700 — 800	80
800 — 900	30
900 — 1000	5

Xác định Me?

Nhận xét

- Vị trí của Me:
 x_{100} và x_{101}
- Tổ chứa Me là tổ thứ 4

Khối lượng lương thực bình quân (kg/người)	Số người (người)	S_i
400 — 500	10	10
500 — 600	30	40
600 — 700	45	85
700 — 800	80	165
800 — 900	30	195
900 —	5	200

1000

Áp dụng CT tính giá trị gần đúng của Me

$$x_{Me\min} = 700$$

$$h_{Me} = 100$$

$$f_{Me} = 80$$

$$S_{Me-1} = 85$$

$$Me = 700 + 100 \times \frac{\frac{200}{80} - 85}{80} = 718,75 \text{ (kg / ng)}$$

2.5 Phân vị

- Phân vị chia dãy số thành n phần có số lượng đơn vị tổng thể bằng nhau.
- Trong thực tế người ta hay dùng:
 - Tứ phân vị
 - Thập phân vị
 - Bách phân vị

Tứ phân vị

- Tài liệu phân tổ không có khoảng cách tổ:
 - Q_1 : tứ phân vị thứ nhất: là lượng biến đứng ở vị trí thứ $(n+1)/4$
 - Q_2 : tứ phân vị thứ hai: chính là trung vị: là lượng biến đứng ở vị trí thứ $2(n+1)/4$
 - Q_3 : tứ phân vị thứ ba: là lượng biến đứng ở vị trí thứ $3(n+1)/4$

Tứ phân vị

- Nếu $n+1$ không phải là bội số của 4, tứ phân vị được xác định bằng cách cộng thêm vào:

Ví dụ:

$$Q_1 = x_3 + 1/4(x_4 - x_3)$$

$$Q_1 = 6 + 1/4(8-6) = 6,5$$

1	3	6	8	10	14
16	18	22	24	26	30

Tứ phân vị

- Tài liệu phân tổ có khoảng cách tổ:

$$Q_1 = X_{Q_1 \min} + h_{Q_1} \times \frac{\frac{1}{4} \sum f - S_{Q_1-1}}{f_{Q_1}}$$

$$Q_3 = X_{Q_3 \min} + h_{Q_3} \times \frac{\frac{3}{4} \sum f - S_{Q_3-1}}{f_{Q_3}}$$

- □p dụng tính Q1 và Q3 cho VD

Bài tập

Thị trường	Số lượng HGD (trgđ)	Bquan (\$)	M _o (\$)	M _e (\$)
A	2,50	280	275	297
B	1,75	305	307	309
C	0,95	367	358	360
D	1,00	335	310	340
E	1,35	353	350	352

Trung vị là mức độ của đơn vị đứng ở vị trí chính giữa trong DSLB và phân chia dãy số thành 2 phần có số lượng đơn vị tổng thể bằng nhau

Thị trường	Số lượng HGĐ (trgđ)	M_e (\$)	Số lượng ĐVTT có giá trị lớn hơn M_e
A	2,50	197	1.250.000
B	1,75	209	875.000
C	0,95	360	475.000
D	1,00	340	500.000
E	1,35	352	675.000

II. Độ biến thiên tiêu thức

1. Ý nghĩa của độ biến thiên tiêu thức

- Đánh giá trình độ đại biểu của số bình quân
- Phản ánh đặc trưng của dãy số như đặc trưng về phân phối, về kết cấu, về tính chất đồng đều của tổng thể.
- Phản ánh chất lượng công tác và nhịp điệu hoàn thành kế hoạch chung của tổng thể cũng như của từng bộ phận
- Độ biến thiên tiêu thức còn được dùng trong nhiều trường hợp nghiên cứu thống kê khác

2. Các tham số đo độ biến thiên tiêu thức

- Khoảng biến thiên
- Độ trải giữa
- Độ lệch tuyệt đối
- Phương sai
- Độ lệch tiêu chuẩn
- Hệ số biến thiên

2.1 Khoảng biến thiên - R^*

- Khoảng biến thiên là độ lệch giữa lượng biến lớn nhất và lượng biến nhỏ nhất trong dãy số lượng biến
- $R^* = X_{\max} - X_{\min}$

Xét ví dụ: NSLD của 2 tổ CN (5 người/tổ) (đv:sp/h)

Tổ 1	40	50	60	70	80	$\bar{x}_1 = 60$	$R^* = 40$
Tổ 2	58	59	60	61	62	$\bar{x}_2 = 60$	$R^* = 4$

Nhận xét về ưu, nhược điểm của R^*

- Ưu điểm

Dễ tính toán, xác định

- Nhược điểm

Chỉ liên quan đến X_{max} và X_{min} mà không tính tới các lượng biến khác trong DS $\rightarrow$ không toàn diện, dễ dẫn đến sai số

2.2 Độ lệch tuyệt đối

- Độ lệch tuyệt đối trung bình là số bình quân cộng của các độ lệch tuyệt đối giữa lượng biến với số bình quân của các lượng biến đó

$$\bar{d}_x = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

$$\bar{d}_x = \frac{\sum_{i=1}^n |x_i - \bar{x}| \times f_i}{\sum_{i=1}^n f_i}$$

2.3 Độ trải giữa

- Độ trải giữa (R_Q) là chênh lệch giữa tứ phân vị thứ 3 và tứ phân vị thứ nhất
- CT: $R_Q = Q_3 - Q_1$

2.4 Phương sai - σ^2

- Phương sai là số bình quân cộng của bình phương các độ lệch giữa lượng biến với số bình quân của các lượng biến đó.
- Công thức 1 (trường hợp DS không phân tổ)

$$\sigma_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2$$

- Công thức 2 (trường hợp DS đã phân tổ — mỗi lượng biến x_i có tần số xuất hiện là f_i)

$$\sigma_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \times f_i}{\sum_{i=1}^n f_i} = \frac{\sum_{i=1}^n x_i^2 \times f_i}{\sum_{i=1}^n f_i} - \bar{x}^2$$

Ví dụ

Tổ 1	40	50	60	70	80	$\bar{x}_1 = 60$	
Tổ 2	58	59	60	61	62	$\bar{x}_2 = 60$	

$$\sigma_{x_1}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^n x_i^2}{n} - \frac{(\sum_{i=1}^n x_i)^2}{n} = \frac{19000}{5} - 3600 = 200$$

$$\sigma_{x_2}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^n x_i^2}{n} - \frac{(\sum_{i=1}^n x_i)^2}{n} = \frac{18010}{5} - 3600 = 2$$

Xác định ph- ơng sai:

Khối l- ợng l- ợng thực bình quân (kg/ng- ười)	Số ng- ười (ng- ười)
400 — 500	10
500 — 600	30
600 — 700	45
700 — 800	80
800 — 900	30
900 — 1000	5

Nhận xét về ưu, nhược điểm của σ^2

■ Ưu điểm

Trong công thức tính toán đã bao gồm tất cả các đơn vị tổng thể $\rightarrow$ toàn diện hơn R^*

■ Nhược điểm

- Khuếch đại sai số
- Đơn vị tính toán không đồng nhất

2.5 Độ lệch tiêu chuẩn - σ

■ Độ lệch tiêu chuẩn là khai phương của phương sai.

■ Công thức:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2}$$

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 \times f_i}{\sum_{i=1}^n f_i}} = \sqrt{\frac{\sum_{i=1}^n x_i^2 \times f_i}{\sum_{i=1}^n f_i} - \bar{x}^2}$$

Ví dụ

Tổ 1	40	50	60	70	80	$\bar{x}_1 = 60$	$\sigma^2 = 200$
Tổ 2	58	59	60	61	62	$\bar{x}_2 = 60$	$\sigma^2 = 2$

$$\sigma_{x_1} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} = \sqrt{200} = 14,142 \text{ (sp / h)}$$

$$\sigma_{x_2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} = \sqrt{2} = 1,4142 \text{ (sp / h)}$$

Nhận xét về ưu, nhược điểm của σ

■ Ưu điểm

- Trong công thức tính toán đã bao gồm tất cả các đơn vị tổng thể $\rightarrow$ toàn diện hơn R^*
- Không khuếch đại sai số $\rightarrow$ tốt hơn σ^2

$\rightarrow$ Độ lệch tiêu chuẩn là chỉ tiêu đo độ biến thiên khá toàn diện

■ Nhược điểm:

Không so sánh được độ biến thiên giữa 2 đại lượng khác loại

2.6 Hệ số biến thiên

- Hệ số biến thiên được sử dụng khi giá trị bình quân của hai tổng thể so sánh khác nhau nhiều hoặc so sánh hai hiện tượng khác nhau
- Công thức:

$$V_x = \frac{\sigma}{x} \times 100 \%$$

$$V_x = \frac{\sigma}{Mo} \times 100 \%$$

$$V_x = \frac{\sigma}{Me} \times 100 \%$$

Ví dụ

	SV Nam	SV nữ
Chiều cao bình quân (cm)	167	156
Trọng lượng bình quân (kg)	57	44
σ_{CC} (cm)	10	11
σ_{TL} (kg)	9	5
V_{CC} (%)	5,98	7,05
V_{TL} (%)	15,79	11,36

Bài tập

Có tài liệu về mối liên hệ giữa NSLD (sp/h) với thu nhập (tr\$) tại 1 tổ công nhân như sau:

NSLD (sp/h)	22	22	24	26	28	29	35	36	40	45
Thu nhập (tr \$)	1,0	1,1	1,3	1,4	1,6	1,6	1,8	1,8	2,0	2,1

Yêu cầu:

Xác định NSLD và thu nhập bình quân của công nhân tổ đó
So sánh trình độ đại biểu của 2 số bình quân trên

■ NSLD bình quân:

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{10} = \frac{307}{10} = 30,7 (sp / h)$$

■ Thu nhập bình quân

$$\bar{y} = \frac{\sum_{i=1}^{10} y_i}{10} = \frac{15,7}{10} = 1,57 (tr \$)$$

So sánh trình độ đại biểu của 2 số bình quân

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{10} x_i^2}{10} - \left(\bar{x}\right)^2} = 7,524 \text{ (sp / h)}$$



$$V_x = \frac{7,524}{30,7} = 24,51 \%$$

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{10} y_i^2}{10} - \left(\bar{y}\right)^2} = 0,349 \text{ (tr \$)}$$



$$V_y = \frac{0,349}{1,57} = 22,26 \%$$

Thu nhập bình quân từng tháng của CN 3 doanh nghiệp A, B, (*đv:100.000VND*)

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T 10	T 11	T 12
A	9	6	7	10	10	8	7	12	9	8	11	17
B	9	10	7	8	11	8	9	7	10	11	11	13
C	9	8	10	11	9	10	8	9	9	10	10	11