



# XÁC SUẤT THỐNG KÊ VÀ ỨNG DỤNG

## LÝ THUYẾT MẪU NGẪU NHIÊN

Buổi học 6

# LÝ THUYẾT MẪU NGẪU NHIÊN

---

- 1. TỔNG THỂ – MẪU – THỐNG KÊ SUY DIỄN**
- 2. PHƯƠNG PHÁP CHỌN MẪU**
- 3. CÁC THAM SỐ CỦA TỔNG THỂ**
- 4. CÁC THAM SỐ CỦA MẪU**
- 5. CÁC LUẬT PHÂN PHỐI CỦA MẪU**
- 6. PHÂN PHỐI TẦN SỐ – PHÂN PHỐI TẦN SUẤT**
- 7. PHƯƠNG PHÁP TÍNH CÁC THAM SỐ MẪU  
CHO DỮ LIỆU ĐƯỢC NHÓM**

# TỔNG THỂ VÀ MẪU

- Ta muốn rút ra một kết luận có giá trị về các cá thể hay vật thể trong một nhóm lớn.
- Thay vì phải khảo sát toàn bộ nhóm, được gọi là *tổng thể*, điều này khó thực hiện, nên ta chỉ có thể khảo sát trên một phần nhỏ của tổng thể này, được gọi là *mẫu*.
- Mục đích suy diễn một sự việc nào đó của tổng thể từ kết quả tìm được trên mẫu, được gọi là *suy diễn theo thống kê*.
- Quá trình lấy các phần tử từ tổng thể được gọi là *chọn mẫu*.

## TỔNG THỂ VÀ MẪU

- **VÍ DỤ 5.1.** Ta muốn rút ra một kết luận về chiều cao (trọng lượng) của 12.000 sinh viên (tổng thể) bằng cách chỉ khảo sát 100 sinh viên (mẫu) được chọn từ tổng thể này.
- **VÍ DỤ 5.2.** Ta muốn rút ra một kết luận về tỷ lệ con bu-long bị hỏng do một nhà máy sản xuất trong suốt tuần lễ (6 ngày làm việc), bằng cách mỗi ngày khảo sát 20 con tại các thời điểm khác nhau. Trong trường hợp này, tất cả các con bu-long được sản xuất trong tuần là tổng thể, khi đó 120 con bu-long được chọn tạo thành một mẫu.

## TỔNG THỂ VÀ MẪU

- Tổng thể thường chỉ sự đo lường hơn là để chỉ các cá thể hay các vật thể. Trong ví dụ 5.1, ta có tổng thể là chiều cao (trọng lượng) của 12.000 sinh viên, trong khi ở ví dụ 5.4 ta có tổng thể là màu của 200 bi trong hộp.
- Tổng thể có thể là hữu hạn hoặc vô hạn, số phần tử trong tổng thể được gọi là *kích thước tổng thể*, ký hiệu là N. Tương tự, số phần tử trong mẫu được gọi là *kích thước mẫu*, ký hiệu là n, thường là hữu hạn. Trong ví dụ 5.1,  $N = 12.000$ ,  $n = 100$ , trong ví dụ 5.3,  $N$  vô hạn,  $n = 60$ .

# CHỌN MẪU CÓ HOÀN LẠI VÀ KHÔNG HOÀN LẠI

- Chọn mẫu để mỗi phần tử của tổng thể có thể được chọn nhiều lần thì được gọi là *chọn mẫu có hoàn lại*, chọn mẫu để mỗi phần tử của tổng thể chỉ được chọn nhiều nhất một lần thì được gọi là *chọn mẫu không hoàn lại*.
- Chọn mẫu có hoàn lại từ một tổng thể hữu hạn trên lý thuyết có thể được xem như mẫu vô hạn. Trong thực hành, chọn mẫu từ tổng thể hữu hạn có kích thước lớn thì có thể được xem như chọn mẫu từ tổng thể vô hạn.

# MẪU NGẪU NHIÊN

---

- Việc rút ra các kết luận đáng tin cậy liên quan đến tổng thể là tùy thuộc vào chọn mẫu có đúng đắn, có đủ đại diện hay không. Một trong những vấn đề quan trọng của suy diễn thống kê là cách chọn mẫu.
- Một phương pháp để thực hiện lấy mẫu đối với tổng thể hữu hạn có kích thước nhỏ là chọn mẫu ngẫu nhiên, bằng cách rút thăm hay dùng bảng số ngẫu nhiên.
- Do tổng thể được suy diễn từ một mẫu nên ta phải sử dụng đến ngôn ngữ xác suất cho bất kỳ kết luận nào về tổng thể.

## CÁC THAM SỐ CỦA TỔNG THỂ

- Một tổng thể được xem là tường minh khi ta biết luật phân phối xác suất  $f(x)$  của ĐLNN X (tính chất được quan tâm của tổng thể). Chẳng hạn, trong ví dụ 5.1 Nếu X là ĐLNN với các giá trị là chiều cao (hoặc trọng lượng) của 12.000 sinh viên thì X có luật phân phối xác suất  $f(x)$ .
- Nếu X có phân phối chuẩn thì ta nói tổng thể có phân phối chuẩn. Tương tự, nếu X có phân phối nhị thức thì ta nói tổng thể có phân phối nhị thức.

# CÁC THAM SỐ CỦA TỔNG THỂ

- Hàm  $f(x)$  của tổng thể được xác định thì các tham số của tổng thể đó cũng sẽ được xác định, chẳng hạn  $\mu$  và  $\sigma$  trong trường hợp phân phối chuẩn hoặc  $p$  trong trường hợp phân phối nhị thức. Tất cả các số đặc trưng này được gọi là các tham số của tổng thể.
- Nếu luật phân phối xác suất  $f(x)$  của tổng thể chưa biết. Ví dụ, có thể có một vài lý do nào đó cho rằng tổng thể có dấu hiệu của phân phối chuẩn. Trong trường hợp đó có thể ta chưa biết một hoặc cả hai giá trị  $\mu$  và  $\sigma$ , vì vậy ta có khuynh hướng chọn suy diễn thống kê để đưa ra các giá trị của chúng.

# CÁC THAM SỐ CỦA MẪU

- Chọn mẫu NN từ tổng thể, ta sẽ sử dụng các phần tử mẫu này để tính các giá trị tham số mẫu, phục vụ cho ước lượng và kiểm định các tham số của tổng thể.
- Để minh họa, xem ví dụ 5.1, trong đó X là ĐLNN gồm các chiều cao khác nhau. Để chọn một mẫu có kích thước 100, trước tiên ta chọn NN một SV trong tổng thể 12.000 SV, gọi  $x_1$  là giá trị của ĐLNN  $X_1$ . Tương tự, chọn phần tử thứ hai của mẫu, gọi  $x_2$  là giá trị của ĐLNN  $X_2$ . Tiếp tục quá trình này cho đến  $X_{100}$ , để đơn giản, ta giả sử việc chọn mẫu trong trường hợp này là có hoàn lại.

# CÁC THAM SỐ CỦA MẪU

- Trong trường hợp tổng quát, một mẫu có kích thước n sẽ được mô tả bởi các giá trị  $x_1, x_2, \dots, x_n$  của các ĐLNN  $X_1, X_2, \dots, X_n$ .  
Trong trường hợp chọn mẫu có hoàn lại thì  $X_1, X_2, \dots, X_n$  độc lập, các ĐLNN có phân phối giống nhau và có hàm xác suất là  $f(x)$ .  
Hàm xác suất đồng thời là

$$P(X=x_1, X=x_2, \dots, X=x_n) = f(x_1)f(x_2)\dots f(x_n)$$

- Mọi số đặc trưng thu được từ mẫu nhằm ước lượng tham số của tổng thể thì được gọi là *thống kê mẫu*.

# CÁC THAM SỐ CỦA MẪU

- Tùy vào mỗi tham số của tổng thể sẽ có một thống kê được tính từ mẫu.
- Thông thường phương pháp để đạt được thống kê này từ mẫu tương tự như phương pháp thu được tham số từ tổng thể hữu hạn.
- Một trong những bài toán quan trọng của lý thuyết mẫu là quyết định cách thức thiết lập thống kê mẫu thích hợp để ước lượng tham số của tổng thể tốt nhất.
- Ta sẽ sử dụng các mẫu tự  $\mu$  và  $\sigma$  cho giá trị của tham số tổng thể, các mẫu tự  $\bar{x}$ ,  $s$ , ... cho giá trị của thống kê mẫu tương ứng.

# CÁC LUẬT PHÂN PHỐI CỦA MẪU

---

- Một thống kê mẫu được tính từ các ĐLNN  $X_1, X_2, \dots, X_n$  là một hàm của các ĐLNN trên. Luật phân phối xác suất của thống kê mẫu thường được gọi là *luật phân phối mẫu*.
- Ta có thể quan tâm đến tất cả các mẫu có cùng kích thước  $n$  được chọn ra từ một tổng thể, với mỗi mẫu ta có thể tính thống kê tương ứng. Theo cách này ta có thể thu được luật phân phối của thống kê mẫu.
- Với luật phân phối mẫu, ta có thể tính *trung bình, phương sai, độ lệch chuẩn* ...

# TRUNG BÌNH MẪU

- Cho một mẫu  $(X_1, X_2, \dots, X_n)$  gồm các ĐLNN độc lập, có cùng luật phân phối. *Trung bình mẫu* là một ĐLNN được định nghĩa như sau  $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$

- Nếu  $x_1, x_2, \dots, x_n$  là các giá trị thu được trong mẫu cụ thể có kích thước là n thì

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

- **VÍ DỤ 5.5.** Cho một mẫu có kích thước là 5 có các giá trị là 7, 9, 1, 6, 2, thì trung bình mẫu là

$$\bar{x} = \frac{7+9+1+6+2}{5} = 5$$

# PHÂN PHỐI CỦA TRUNG BÌNH MẪU

- **Định lý 5-1:** Trung bình của phân phối trung bình mẫu, ký hiệu là  $\mu_{\bar{X}}$ , được cho bởi biểu thức sau

$$E(\bar{X}) = \mu_{\bar{X}} = \mu$$

trong đó  $\mu$  là trung bình của tổng thể.

- **Định lý 5-2:** Nếu tổng thể là vô hạn và chọn mẫu ngẫu nhiên hoặc nếu tổng thể là hữu hạn và chọn mẫu có hoàn lại thì phương sai của phân phối trung bình mẫu, ký hiệu  $\sigma_{\bar{X}}^2$ , được cho bởi biểu thức

$$E[(\bar{X} - \mu)^2] = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$$

trong đó  $\sigma^2$  là phương sai của tổng thể.

# PHÂN PHỐI CỦA TRUNG BÌNH MẪU

- **Định lý 5-3:** Nếu tổng thể có kích thước N, chọn mẫu không hoàn lại và kích thước mẫu là  $n \leq N$  thì  $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \left( \frac{N-n}{N-1} \right)$

Chú ý rằng khi  $N \rightarrow \infty$  thì định lý 5-3 trở thành định lý 5-2.

- **Định lý 5-4:** Nếu tổng thể có phân phối chuẩn với trung bình  $\mu$  và phương sai  $\sigma^2$ . Một mẫu được chọn từ tổng thể này thì trung bình mẫu cũng có phân phối chuẩn với trung bình  $\mu$  và phương sai là  $\sigma^2/n$ .

# PHÂN PHỐI CỦA TRUNG BÌNH MẪU

□ **Định lý 5-5:** Giả sử tổng thể có luật phân phối với trung bình  $\mu$  và phương sai  $\sigma^2$ , không nhất thiết phải là phân phối chuẩn. Một mẫu được chọn từ tổng thể này thì ĐLNN được chuẩn hóa liên kết với  $\bar{X}$ , được cho bởi

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

xấp xỉ với phân phối chuẩn tắc, nghĩa là

$$\lim_{n \rightarrow \infty} P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du$$

# PHÂN PHỐI TỶ LỆ MẪU

- Giả sử một tổng thể hữu hạn và có phân phối nhị thức với tham số  $p$  và  $q = 1 - p$ .
- Ví dụ, tổng thể là tất cả các lần tung của một đồng xu, trong đó xác suất của biến cố mặt ngửa xảy ra trong mỗi lần tung là  $p = \frac{1}{2}$ .
- Xét các mẫu có kích thước  $n$  được chọn ra từ tổng thể này, với mỗi mẫu ta xác định một thống kê là tỷ lệ  $f$  của số lần thành công (tỷ lệ xuất hiện mặt ngửa). Từ đó ta thu được luật phân phối tỷ lệ mẫu

$$\mu_f = P \quad \sigma_f = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}$$

# PHÂN PHỐI HIỆU VÀ TỔNG CỦA MẪU

- Giả sử có 2 tổng thể. Từ tổng thể thứ nhất, chọn một mẫu  $n_1$ , tính thống kê  $S_1$ , được  $\mu_{S_1}$ ,  $\sigma_{S_1}$ . Tương tự, chọn một mẫu  $n_2$  từ tổng thể thứ hai, tính thống kê  $S_2$ , được  $\mu_{S_2}$ ,  $\sigma_{S_2}$ .
- Ta có thể thu được phân phối hiệu,  $S_1 - S_2$ , của thống kê mẫu. Trung bình  $\mu_{S_1-S_2}$  và độ lệch chuẩn  $\sigma_{S_1-S_2}$  của phân phối mẫu này cho bởi

$$\mu_{S_1-S_2} = \mu_{S_1} - \mu_{S_2} \quad \sigma_{S_1-S_2} = \sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2}$$

- Phân phối tổng mẫu của thống kê  $S_1$  và  $S_2$  có trung bình và độ lệch chuẩn là

$$\mu_{S_1+S_2} = \mu_{S_1} + \mu_{S_2}, \quad \sigma_{S_1+S_2} = \sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2}$$

# PHƯƠNG SAI MẪU

Gọi  $X_1, X_2, \dots, X_n$  là một mẫu ngẫu nhiên có kích thước  $n$  thì ĐLNN của *phương sai mẫu* được định nghĩa

$$S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n}$$

Trong định lý 5-1 ta thấy  $E(X) = \mu$ , và rất đẹp nếu ta cũng có  $E(S^2) = \sigma^2$ . Tuy nhiên nó có khuynh hướng là

$$E(S^2) = \mu_{s^2} = \frac{n-1}{n} \sigma^2$$

Khi giá trị kỳ vọng của thống kê tương ứng bằng tham số tổng thể thì ta gọi thống kê đó là một *ước lượng không chêch*.

# PHƯƠNG SAI MẪU

Biểu thức  $E(S^2)$  rất gần với  $\sigma^2$  chỉ khi giá trị n lớn ( $n \geq 30$ ). Ước lượng không chệch được xác định là

$$\hat{S}^2 = \frac{n}{n-1} S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n-1}$$

sao cho  $E(\hat{S}^2) = \sigma^2$

Vì lý do này, một vài nhà thống kê chọn định nghĩa phương sai mẫu là  $\hat{S}^2$ . Tuy nhiên ta sẽ dùng định nghĩa phương sai mẫu  $S^2$  vì các kết quả về sau sẽ đơn giản hơn.

# PHƯƠNG SAI MẪU

□ **VÍ DỤ 5.6.** Cho một mẫu có kích thước là 5 có các giá trị là 7, 9, 1, 6, 2, thì phương sai mẫu là

$$s^2 = \frac{(7-5)^2 + (9-5)^2 + (1-5)^2 + (6-5)^2 + (2-5)^2}{5} = 9,2$$

khi đó ước lượng không chêch là

$$\hat{s}^2 = \frac{5}{4} s^2 = \frac{(7-5)^2 + (9-5)^2 + (1-5)^2 + (6-5)^2 + (2-5)^2}{4} = 11,5$$

□ Nếu chọn mẫu từ một tổng thể vô hạn hoặc chọn mẫu có hoàn lại từ tổng thể hữu hạn có kích thước N thì đó

$$E(S^2) = \mu_{s^2} = \left( \frac{N}{N-1} \right) \left( \frac{n-1}{n} \right) \sigma^2$$

□ Khi  $N \rightarrow \infty$ ,  $E(S^2) = ((n-1)/n)\sigma^2$

# PHÂN PHỐI CỦA PHƯƠNG SAI MẪU

□ **Định lý 5-6:** Nếu mẫu ngẫu nhiên có kích thước n được chọn từ một tổng thể có phân phối chuẩn thì ĐLNN

$$\frac{nS^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{\sigma^2}$$

có phân phối chi-bình phương  $n-1$  bậc tự do

□ **Định lý 5-7:** Nếu các mẫu ngẫu nhiên có kích thước n được chọn từ một tổng thể có phân phối chuẩn thì thống kê

$$T = \frac{\bar{X} - \mu}{S / \sqrt{n}} = \frac{\bar{X} - \mu}{S / \sqrt{n-1}}$$

có phân phối Student với  $n - 1$  bậc tự do.

# PHÂN PHỐI TỶ LỆ CỦA PHƯƠNG SAI MẪU

Thay vì xét phân phối hiệu của phương sai mẫu, ta chú ý đến thống kê  $S_1^2/S_2^2$ .

**Định lý 5-8:** Cho hai mẫu ngẫu nhiên độc lập có kích thước lần lượt là  $m$  và  $n$ , được chọn từ hai tổng thể có phân phối chuẩn với phương sai lần lượt là  $\sigma_1^2$  và  $\sigma_2^2$ . Nếu phương sai của các mẫu ngẫu nhiên lần lượt là  $S_1^2$  và  $S_2^2$  thì thống kê

$$F = \frac{mS_1^2 / (m-1)\sigma_1^2}{nS_2^2 / (n-1)\sigma_2^2} = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}$$

có phân phối F với  $m - 1$ ,  $n - 1$  bậc tự do.

# PHÂN PHỐI TUẦN SUẤT

- Một mẫu có kích thước lớn thì rất khó khăn quan sát các tính chất khác nhau hoặc tính các thống kê, do đó cần tổ chức hoặc phân nhóm dữ liệu thô.
- Giả sử có một mẫu gồm các chiều cao của 100 sinh viên nữ ở Đại học XYZ. Ta sắp xếp dữ liệu thành từng lớp và xác định số cá thể thuộc vào mỗi lớp, được gọi là *tần số của lớp*.
- Kết quả việc sắp xếp này (xem Bảng 5-2), được gọi là *phân phối tần số* hay *bảng phân phối tần số*.

# PHÂN PHỐI TẦN SỐ

Chiều cao của 100 sinh viên nữ ở Đại học XYZ được thể hiện trong bảng 5-2. Hình 5-1 là biểu diễn tần số chiều cao của mẫu gồm 100 sinh viên nữ.

Table 5-2 Heights of 100 Male Students at XYZ University

Height (inches)	Number of Students
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8
TOTAL	100

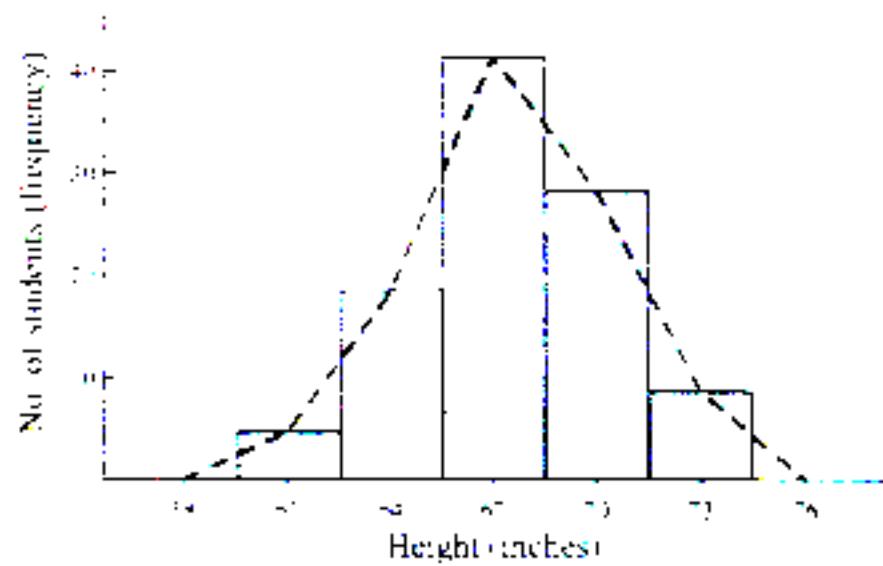
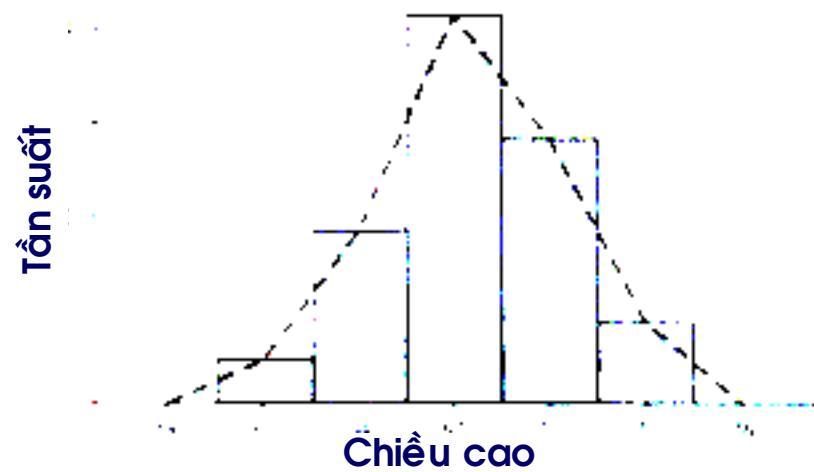


Fig. 5-1

# PHÂN PHỐI TẦN SUẤT

Nếu trong bảng 5-2 ta ghi nhận tần suất hay  
tỷ lệ bách phân chiều cao sinh viên

Chiều cao	Tần suất
60 - 62	5/100
63 - 65	18/100
66 - 68	42/100
69 - 71	27/100
72 - 74	8/100
Tổng số	1



Tổng diện tích các hình chữ nhật bằng 1  
Các tần suất được coi là các xác suất thực nghiệm, nên phân phối tần suất được hiểu là các phân phối xác suất thực nghiệm.

# PHƯƠNG PHÁP TÍNH THAM SỐ MẪU

Có 2 cách mô tả mẫu ngẫu nhiên

① Mô tả bằng bảng phân phối tần số

$x_i$	$x_1$	$x_2$	...	$x_j$	...	$x_k$
$n_i$	$n_1$	$n_2$	...	$n_j$	...	$n_k$

trong đó:  $\sum_{i=1}^k n_i = n$

$x_i$	$n_i$
$x_1$	$n_1$
$x_2$	$n_2$
.	.
.	.
$x_k$	$n_k$
Tổng	$n$

② Mô tả bằng bảng phân phối tần suất

$x_i$	$x_1$	$x_2$	...	$x_j$	...	$x_k$
$f_i$	$f_1$	$f_2$	...	$f_j$	...	$f_k$

trong đó:  $\sum_{i=1}^k f_i = 1$

# PHƯƠNG PHÁP TÍNH THAM SỐ MẪU

1. Trung bình mẫu
2. Phương sai mẫu

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

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i n_i \quad (1)$$

$$\Rightarrow \hat{s}^2 = \frac{n}{n-1} s^2 \quad (2)$$

Cách tính nhanh phương sai có điều chỉnh

Từ bảng phân phối tần số tính  $\sum x_i^2 n_i$

- ① Áp dụng công thức:  $\hat{s}^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n x_i^2 n_i - n(\bar{x})^2 \right] \quad (3)$
- ② Lập bảng tính theo công thức

$$\hat{s}^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n x_i^2 n_i - \frac{1}{n} \left( \sum_{i=1}^n x_i n_i \right)^2 \right] \quad (4)$$

3. Độ lệch chuẩn mẫu

$$s = \sqrt{s^2} \quad \text{hay} \quad \hat{s} = \sqrt{\hat{s}^2}$$

(5)

# PHƯƠNG PHÁP TÍNH THAM SỐ MẪU

**VÍ DỤ.** Tính điểm trung bình và phương sai từ một mẫu gồm 50 sinh viên như sau

$x_i$	4	5	7	9
$n_i$	10	15	13	12

**Cách 1.**

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i n_i = \frac{1}{50} (4 \times 10 + 5 \times 15 + 7 \times 13 + 9 \times 12) = 6,28$$

$$\hat{s}^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n x_i^2 n_i - n (\bar{x})^2 \right]$$

$$\sum x_i^2 n_i = (4^2 \times 10 + 5^2 \times 15 + 7^2 \times 13 + 9^2 \times 12) = 2.144$$

$$\hat{s}^2 = \frac{1}{49} \left[ 2.144 - 50 (6,28)^2 \right] = 3,5118$$

# PHƯƠNG PHÁP TÍNH THAM SỐ MẪU

Cách 2.

$x_i$	$n_i$	$n_i \cdot x_i$	$n_i \cdot x_i^2$
4	10	40	160
5	15	75	375
7	13	91	637
9	12	108	972
Tổng	50	314	2144

Từ kết quả tính toán ở bảng trên, ta có:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i n_i = \frac{314}{50} = 6,28$$

$$\begin{aligned}\hat{s}^2 &= \frac{1}{n-1} \left[ \sum_{i=1}^n x_i^2 n_i - \frac{1}{n} \left( \sum_{i=1}^n x_i n_i \right)^2 \right] = \frac{1}{49} \left[ 2144 - \frac{1}{50} \cdot (314)^2 \right] \\ &= 3,5118\end{aligned}$$

# PHƯƠNG PHÁP TÍNH THAM SỐ MẪU

**VÍ DỤ.** Chọn một mẫu NN gồm 40 đầu tư ngắn hạn tại thành phố trong năm 2000.

Kỳ hạn thanh toán (tháng)	30–39	40–49	50–59	60–69	70–79	80–89	90–99
Số đầu tư	3	1	8	10	7	7	4

(a) Tính giá trị trung bình và độ lệch chuẩn (có điều chỉnh) kỳ hạn thanh toán của 40 đầu tư ngắn hạn trên.

(b) Các đầu tư có kỳ hạn thanh toán dưới 60 tháng là các loại đầu tư kém hiệu quả, tính tỷ lệ các đầu tư kém hiệu quả của mẫu trên.

# PHƯƠNG PHÁP TÍNH THAM SỐ MẪU

## Cách 1. Đưa bảng về dạng điểm

Kỳ hạn thanh toán (tháng)	34,5	44,5	54,5	64,5	74,5	84,5	94,5
Số đầu tư	3	1	8	10	7	7	4

### (a) Trung bình và phương sai (có điều chỉnh)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i n_i = \frac{1}{40} (34,5 \times 3 + 44,5 \times 1 + \dots + 94,5 \times 4) = 68$$

$$\sum x_i^2 n_i = (34,5^2 \times 3 + 44,5^2 \times 1 + \dots + 94,5^2 \times 4) = 269,48$$

$$\hat{s}^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n x_i^2 n_i - n (\bar{x})^2 \right] = \frac{1}{39} [195.470 - 40 (68)^2] = 269,48$$

$$\Rightarrow \hat{s} = \sqrt{269,487} = 16,42$$

$$(b) Tỷ lệ mẫu \quad f = \frac{k}{n} = \frac{3+1+8}{40} = 0,3$$

# PHƯƠNG PHÁP TÍNH THAM SỐ MẪU

Cách 2:  
Lập bảng

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i n_i = \frac{2720}{40} = 68$$

Kỳ hạn	$x_i$	$n_i$	$x_i n_i$	$x_i^2 n_i$
30-39	34,5	3	103,5	3570,75
40-49	44,5	1	44,5	1980,25
50-59	54,5	8	436	23762
60-69	64,5	10	645	41602,5
70-79	74,5	7	521,5	38851,75
80-89	84,5	7	591,5	49981,75
90-99	94,5	4	378	35721
	40	2720	195470	

$$\hat{s}^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n x_i^2 n_i - \frac{1}{n} \left( \sum_{i=1}^n x_i n_i \right)^2 \right] = \frac{1}{39} \left[ 195.470 - \frac{(2720)^2}{40} \right]$$

$$= 269.487$$

$$f = \frac{k}{n} = \frac{3+1+8}{40} = 0,3 \Rightarrow \hat{s} = \sqrt{269,487} = 16,42$$

# THE END

---

