

Mục lục

1. Ma trận - Định thức - Hệ phương trình tuyến tính	5
1.1. Chuẩn bị	5
1.1.1. Tích Đề-các	5
1.1.2. Ánh xạ	5
1.2. Ma trận và các phép toán trên ma trận	6
1.2.1. Khái niệm ma trận.	6
1.2.2. Các phép toán trên ma trận.	8
1.3. Định thức	9
1.3.1. Định nghĩa định thức	9
1.3.2. Ma trận nghịch đảo	15
1.3.3. Hạng của ma trận	17
1.3.4. Hệ phương trình tuyến tính	19
Bài tập chương 1	26
2. Biến cố ngẫu nhiên và xác suất của nó	33
2.1. Phép thử và phân loại biến cố	33
2.1.1. Định nghĩa	33
2.1.2. Phân loại biến cố	33
2.2. Định nghĩa xác suất	34
2.2.1. Xác suất của biến cố	34
2.2.2. Định nghĩa cổ điển về xác suất	34
2.2.3. Định nghĩa hình học về xác suất	36
2.2.4. Định nghĩa thống kê về xác suất	37
2.3. Quan hệ giữa các biến cố	38
2.3.1. Tổng các biến cố	38
2.3.2. Tích các biến cố	38
2.3.3. Biến cố xung khắc	39
2.3.4. Nhóm đầy đủ các biến cố	39
2.3.5. Biến cố đối lập	39
2.4. Định lý cộng và nhân xác suất	40
2.4.1. Định lý cộng xác suất (trường hợp các biến cố xung khắc)	40
2.4.2. Định lý nhân xác suất	41
2.4.3. Định lý cộng xác suất (trường hợp tổng quát)	45
2.4.4. Định lý liên hệ cộng và nhân xác suất	46
2.5. Công thức Bernoulli	47
2.5.1. Các phép thử độc lập	47
2.5.2. Công thức Bernoulli	47
2.5.3. Số lần xuất hiện chắc nhất	48
2.5.4. Mở rộng công thức Bernoulli	49
2.6. Công thức đầy đủ và công thức Bayes	50
2.6.1. Công thức xác suất đầy đủ	50
2.6.2. Công thức Bayes	52

Bài tập chương 2	53
3. Đại lượng ngẫu nhiên và các quy luật phân phối xác suất	61
3.1. Định nghĩa và phân loại đại lượng ngẫu nhiên	61
3.1.1. Định nghĩa	61
3.1.2. Phân loại đại lượng ngẫu nhiên	61
3.2. Quy luật phân phối xác suất của đại lượng ngẫu nhiên	62
3.2.1. Bảng phân phối xác suất	62
3.2.2. Hàm phân phối xác suất	64
3.2.3. Hàm mật độ xác suất	66
3.3. Các tham số đặc trưng của đại lượng ngẫu nhiên	71
3.3.1. Kỳ vọng toán	71
3.3.2. Phương sai	75
3.3.3. Độ lệch tiêu chuẩn	78
3.3.4. Mốt	78
3.3.5. Trung vị	79
3.3.6. Phân vị	80
3.4. Một số quy luật phân phối xác suất thông dụng	80
3.4.1. Quy luật phân phối chuẩn $N(\mu, \sigma^2)$	80
3.4.2. Quy luật không - một $A(p)$	86
3.4.3. Quy luật nhị thức $B(n, p)$	87
3.4.4. Quy luật Poisson $P(\lambda)$	90
3.4.5. Quy luật siêu bội $M(N, n)$	92
3.4.6. Quy luật khi - bình phương $\chi^2(n)$	93
3.4.7. Quy luật Student $T(n)$	94
Bài tập chương 3	96
4. Mẫu ngẫu nhiên - Ước lượng tham số	103
4.1. Tổng thể nghiên cứu	103
4.1.1. Định nghĩa	103
4.1.2. Các phương pháp mô tả tổng thể	103
4.1.3. Các tham số đặc trưng của tổng thể	104
4.2. Mẫu ngẫu nhiên	105
4.2.1. Định nghĩa	105
4.2.2. Các phương pháp mô tả mẫu ngẫu nhiên	106
4.2.3. Đồ thị của phân phối thực nghiệm	107
4.3. Thống kê	109
4.3.1. Định nghĩa	109
4.3.2. Trung bình mẫu	109
4.3.3. Phương sai mẫu	110
4.3.4. Độ lệch tiêu chuẩn mẫu	112
4.3.5. Tần suất mẫu	113
4.3.6. Quy luật phân phối xác suất của một số thống kê đặc trưng mẫu	113
4.3.7. Ví dụ	115
4.4. Mẫu ngẫu nhiên hai chiều	115
4.4.1. Khái niệm	115
4.4.2. Phương pháp mô tả ngẫu nhiên hai chiều	116
4.4.3. Một số thống kê đặc trưng của mẫu ngẫu nhiên hai chiều	116
4.5. Ước lượng tham số của đại lượng ngẫu nhiên	118
4.5.1. Phương pháp ước lượng điểm	118
4.5.2. Phương pháp ước lượng bằng khoảng tin cậy	119

4.5.3. Khoảng tin cậy cho trung bình (Ước lượng kỳ vọng toán của đại lượng ngẫu nhiên phân phối theo quy luật chuẩn)	119
4.5.4. Khoảng tin cậy cho tỷ lệ (Ước lượng kỳ vọng toán của đại lượng ngẫu nhiên phân phối theo quy luật không - một)	126
4.5.5. Khoảng tin cậy cho phương sai (Ước lượng phương sai của đại lượng ngẫu nhiên phân phối theo quy luật chuẩn)	128
Bài tập chương 4	132
PHỤ LỤC	138
A. Giải tích tổ hợp	139
A.1. Các quy tắc đếm	139
A.1.1. Quy tắc cộng	139
A.1.2. Quy tắc nhân	139
A.2. Hoán vị - Chỉnh hợp - Tổ hợp	139
A.2.1. Chỉnh hợp (chỉnh hợp không lặp)	139
A.2.2. Chỉnh hợp lặp	140
A.2.3. Hoán vị	140
A.2.4. Tổ hợp	141
Bài tập phụ lục A	142
B. Sử dụng CNTT giải toán thống kê	143
B.1. Đối với máy tính điện tử cầm tay	143
B.1.1. Tính các đặc trưng của mẫu	143
B.1.2. Bài toán tìm hàm hồi quy	146
B.2. Dùng phần mềm Excel	150
B.2.1. Tính toán trong bài toán ước lượng	150
B.2.2. Tính toán các đặc trưng của mẫu	152
B.2.3. Các phân phối xác suất	153
C. Bảng tra	155
C.1. Bảng giá trị hàm mật độ của phân phối chuẩn hóa	156
C.2. Bảng giá trị hàm Laplace	157
C.3. Bảng phân vị chuẩn	158
C.4. Bảng phân vị Student	159
C.5. Bảng phân vị Khi - bình phương	160
TÀI LIỆU THAM KHẢO	161

Chương 1

Ma trận - Định thức - Hệ phương trình tuyến tính

1.1. Chuẩn bị

1.1.1. Tích Đề-các

★ **Định nghĩa 1.1** Cho họ gồm n tập $\{A_i\}_{i=1}^n$ (n là số nguyên dương). Tích Đề-các của họ đã cho là một tập, ký hiệu là $A_1 \times A_2 \times \cdots \times A_n$, mỗi phần tử của nó là một bộ có thứ tự gồm n thành phần (a_1, a_2, \dots, a_n) , trong đó $a_i \in A_i$ với $i = 1, 2, \dots, n$.

• **Ví dụ 1.1** Cho $A_1 = \{a, b, c\}$, $A_2 = \{1, 2\}$ khi đó:

$$A_1 \times A_2 = \{(a, 1); (a, 2); (b, 1); (b, 2); (c, 1); (c, 2)\}$$

$$A_2 \times A_1 = \{(1, a); (2, a); (1, b); (2, b); (1, c); (2, c)\}$$

Vậy $A_1 \times A_2 \neq A_2 \times A_1$.

Chú ý: Nếu $A_1 = A_2 = \cdots = A_n = A$, thay cho ký hiệu $A_1 \times A_2 \times \cdots \times A_n$ ta dùng ký hiệu A^n .

• **Ví dụ 1.2** $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n), x_i \in \mathbb{R}, i = 1, 2, \dots, n\}$.

1.1.2. Ánh xạ

★ **Định nghĩa 1.2** Cho hai tập khác rỗng X, Y . Một ánh xạ f từ X vào Y là một quy tắc cho phép với mỗi phần tử $x \in X$ **xác định duy nhất một** phần tử $y = f(x) \in Y$, ký hiệu: $f : X \rightarrow Y$ hoặc $y = f(x)$. Trong định nghĩa trên

- X được gọi là tập nguồn của ánh xạ f
- Y được gọi là tập đích của ánh xạ f
- $y = f(x)$ gọi là ảnh của x qua ánh xạ f , x gọi là tạo ảnh của $y = f(x)$
- Giả sử $A \subset X$, khi đó $f(A) = \{f(x) : x \in A\}$ gọi là ảnh của A qua ánh xạ f .
- Giả sử $B \subset Y$, Khi đó $f^{-1}(B) = \{x : y = f(x) \in B\}$ gọi là nghịch ảnh của B bởi f

★ **Định nghĩa 1.3** Cho $f : X \rightarrow Y$ là một ánh xạ

1. f là đơn ánh nếu $x_1, x_2 \in X$ và $x_1 \neq x_2$ thì $f(x_1) \neq f(x_2)$

2. f là toàn ánh nếu $f(X) = Y$
3. f là song ánh nếu f vừa là đơn ánh và vừa là toàn ánh

• **Ví dụ 1.3** 1. $y = e^x$ là một đơn ánh từ \mathbb{R} vào \mathbb{R}

2. $y = x^2$ là một toán ánh từ \mathbb{R} vào \mathbb{R}^+

3. $y = x$ là một song ánh từ \mathbb{R} vào \mathbb{R}

Chú ý: Nếu $X = Y$ ánh xạ $f : X \rightarrow X$ xác định bởi $y = f(x) = x$ được gọi là ánh xạ đồng nhất trên X , ký hiệu là id_X . Dễ thấy ánh xạ đồng nhất là một song ánh.

★ **Định nghĩa 1.4** Cho các ánh xạ $f : X \rightarrow Y$ và $g : Y \rightarrow Z$. Tích của ánh xạ g với ánh xạ f là một ánh xạ, ký hiệu $g.f$, xác định như sau:

$$(g.f) : X \rightarrow Z$$

$$(g.f)(x) = g[f(x)]$$

• **Ví dụ 1.4** Cho 2 ánh xạ $f, g : \mathbb{R} \rightarrow \mathbb{R}$ xác định bởi $y = f(x) = \frac{x^2}{x^2 + 1}$ và $y = g(x) = 2x + 1$.

$$\text{Khi đó } (g.f)(x) = \frac{2x^2}{x^2 + 1} + 1 = \frac{3x^2 + 1}{x^2 + 1}$$

★ **Định nghĩa 1.5** Giả sử $f : X \rightarrow Y$ là một ánh xạ. Nếu tồn tại ánh xạ $g : Y \rightarrow X$ sao cho

$$g.f = id_X \text{ và } f.g = id_Y$$

ta gọi g là ánh xạ ngược của f , f là ánh xạ ngược của g

• **Ví dụ 1.5** Cho $f : \mathbb{R} \rightarrow \mathbb{R}^+$ xác định bởi $y = f(x) = e^x$ và $g : \mathbb{R}^+ \rightarrow \mathbb{R}$ xác định bởi $y = g(x) = \ln(x)$. Dễ dàng kiểm tra rằng f và g là hai ánh xạ ngược của nhau.

1.2. Ma trận và các phép toán trên ma trận

1.2.1. Khái niệm ma trận.

★ **Định nghĩa 1.6** Cho $m, n \in \mathbb{N}^*$. Một ma trận thực cỡ $m \times n$ là một bảng chữ nhật gồm $m \times n$ số thực xếp thành m hàng, n cột. Số thực đứng ở hàng i cột j gọi là phần tử ij . Nếu ký hiệu phần tử này là a_{ij} thì một ma trận cỡ $m \times n$ có thể biểu diễn bởi:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Trong một số trường hợp ta còn dùng ký hiệu thu gọn $[a_{ij}]_{m \times n}$ để chỉ một ma trận m hàng, n cột.

• **Ví dụ 1.6** Cho $A = \begin{bmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \end{bmatrix}$. Đây là ma trận cỡ 2×3 có:

$$a_{11} = 1, a_{12} = 3, a_{13} = 5, a_{21} = 7, a_{22} = 9, a_{23} = 11.$$

Ma trận cột là ma trận cỡ $m \times 1$.

Ma trận hàng là ma trận cỡ $1 \times n$.

Ma trận không là ma trận mà mọi phần tử của nó đều bằng 0, ký hiệu là Θ hoặc $\Theta_{m \times n}$ nếu muốn chỉ rõ cỡ của ma trận.

Ma trận vuông cấp n là ma trận có số hàng bằng số cột và bằng n . Cho ma trận vuông cấp n , $A = [a_{ij}]_{n \times n}$, đường chéo chính của A là tập hợp các phần tử có dạng: $a_{11}, a_{22}, a_{33}, \dots, a_{nn}$.

Ma trận tam giác trên là ma trận vuông cấp n trong đó $a_{ij} = 0$ nếu $i > j$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

Ma trận tam giác dưới là ma trận vuông cấp n trong đó $a_{ij} = 0$ nếu $i < j$.

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Ma trận đường chéo là ma trận vuông cấp n trong đó $a_{ij} = 0$ nếu $i \neq j$.

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

Ma trận đơn vị là ma trận đường chéo có các phần tử thuộc đường chéo chính đều bằng 1, ma trận đơn vị cấp n thường được ký hiệu là I_n hoặc là I nếu không xảy ra hiểu lầm.

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Ma trận chuyển vị của ma trận A là ma trận ký hiệu là A^t , nhận được từ ma trận A bằng cách viết các hàng của A thành cột của A^t . Như vậy:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \Rightarrow A^t = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Hai ma trận A, B được gọi là bằng nhau nếu chúng có cùng cỡ nghĩa là $A = [a_{ij}]_{m \times n}$ thì $B = [b_{ij}]_{m \times n}$ và các phần tử ở các vị trí tương ứng bằng nhau cụ thể $a_{ij} = b_{ij}$ với mọi $i = 1, 2, \dots, m, j = 1, 2, \dots, n$, ký hiệu $A = B$.

1.2.2. Các phép toán trên ma trận.

a) Cộng hai ma trận cùng cỡ

★ **Định nghĩa 1.7** Cho hai ma trận cùng cỡ $A = [a_{ij}]_{m \times n}$, $B = [b_{ij}]_{m \times n}$. Tổng của A và B là một ma trận cùng cỡ $C = [c_{ij}]_{m \times n}$, ký hiệu là $C = A + B$, trong đó $c_{ij} = a_{ij} + b_{ij}$, $i = 1, \dots, m, j = 1, \dots, n$.

Như vậy muốn cộng hai ma trận cùng cỡ, ta cộng các phần tử cùng vị trí với nhau.

• Ví dụ 1.7

$$A = \begin{bmatrix} 1 & -2 & 4 \\ 0 & 5 & 7 \end{bmatrix}, B = \begin{bmatrix} 4 & -3 & -5 \\ 2 & 4 & 1 \end{bmatrix}.$$

$$\text{Khi đó: } A + B = C = \begin{bmatrix} 5 & -5 & -1 \\ 2 & 9 & 8 \end{bmatrix}$$

Tính chất: Các phép tính cộng trên các ma trận cùng cỡ có tính chất giống như các tính chất của phép cộng các số thực:

- Tính giao hoán $A + B = B + A$
- Tính kết hợp $(A + B) + C = A + (B + C)$
- $A + \Theta = A$
- $A = [a_{ij}]_{m \times n}$, \exists ma trận đối của ma trận A là $-A = [-a_{ij}]_{m \times n}$ thỏa mãn

$$A + (-A) = \Theta$$

b) Nhân ma trận với một số thực

★ **Định nghĩa 1.8** Cho ma trận $A = [a_{ij}]_{m \times n}$ và số thực k . Tích của A với số thực k là một ma trận cùng cỡ với ma trận A , ký hiệu là kA , xác định bởi công thức

$$kA = [ka_{ij}]_{m \times n}$$

• Ví dụ 1.8

$$2 \begin{bmatrix} 4 & 0 & 1 \\ -2 & 7 & 3 \\ -1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 0 & 2 \\ -4 & 14 & 6 \\ -2 & 4 & 2 \end{bmatrix}$$

Tính chất: Giả sử $k, h \in \mathbb{R}$ và A, B là các ma trận, ta có các tính chất sau:

- $k(A + B) = kA + kB$
- $k(hA) = khA$
- $(k + h)A = kA + hA$
- $1.A = A$

c) Nhân ma trận với ma trận

★ **Định nghĩa 1.9** Cho hai ma trận $A = [a_{ij}]_{m \times p}$ và $B = [b_{ij}]_{p \times n}$. Tích của ma trận A với ma trận B là một ma trận có cỡ là $m \times n$, ký hiệu là AB , xác định như sau:

$$AB = [c_{ij}]_{m \times n}, c_{ij} = \sum_{k=1}^p a_{ik} \cdot b_{kj} = a_{i1} \cdot b_{1j} + a_{i2} \cdot b_{2j} + \cdots + a_{ip} \cdot b_{pj}$$

• **Ví dụ 1.9** Cho

$$A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 3 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix}$$

Khi đó

$$AB = \begin{bmatrix} 1.2 + 2.3 & 1.0 + 2.1 \\ 0.2 - 1.3 & 0.0 - 1.1 \\ 3.2 + 1.3 & 3.0 + 1.1 \end{bmatrix} = \begin{bmatrix} 8 & 2 \\ -3 & -1 \\ 9 & 1 \end{bmatrix}$$

• **Ví dụ 1.10** Tính AB và BA nếu

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 4 & -8 \end{bmatrix}, B = \begin{bmatrix} 1 & 3 \\ -2 & 7 \\ 5 & 4 \end{bmatrix}$$

Ta có

$$AB = \begin{bmatrix} 9 & 3 \\ -48 & -4 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & 11 & -23 \\ -4 & 30 & -58 \\ 10 & 11 & -27 \end{bmatrix}$$

Nhận xét: Do $AB \neq BA$ nên phép nhân hai ma trận không có tính giao hoán.

Tính chất: Giả sử A, B, C là các ma trận và k là một số thực. Nếu các phép tính ở vế trái của đẳng thức dưới đây có nghĩa thì vế phải cũng có nghĩa và 2 vế bằng nhau

- $(AB)C = A(BC)$
- $A(B + C) = AB + AC$
- $(B + C)A = BA + CA$
- $k(AB) = (kA)B = A(kB)$
- $IA = AI = A$
- $\Theta A = A\Theta = \Theta$

1.3. Định thức

1.3.1. Định nghĩa định thức

1. Các phép biến đổi sơ cấp trên ma trận

Cho ma trận $A = [a_{ij}]_{m \times n}$. Ta gọi các phép biến đổi sau trên các hàng của A là các phép biến đổi sơ cấp trên hàng:

- Hoán vị 2 hàng của A .
- Nhân tất cả các phần tử của một hàng nào đó của A với cùng một số khác 0.
- Nhân tất cả các phần tử của một hàng nào đó của A với cùng một số rồi cộng vào các phần tử tương ứng của một hàng khác.

Chú ý: Các phép biến đổi sơ cấp trên cột cũng được định nghĩa tương tự.

• **Ví dụ 1.11** Cho ma trận $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$. Hãy dùng các phép biến đổi sơ cấp trên hàng

của A để đưa A về dạng tam giác trên.

Lời giải. Ta thực hiện dãy các phép biến đổi sơ cấp trên hàng để biến đổi dần ma trận A thành ma trận có dạng tam giác trên như sau:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \xrightarrow[\begin{smallmatrix} H_2-2H_1 \rightarrow H_2 \\ H_3-3H_1 \rightarrow H_3 \end{smallmatrix}]{\begin{smallmatrix} H_3-5H_2 \rightarrow H_3 \\ H_2-2H_1 \rightarrow H_2 \end{smallmatrix}} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & -5 & -7 \end{bmatrix} \xrightarrow{H_3-5H_2 \rightarrow H_3} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & 18 \end{bmatrix}.$$

2. Ma trận con A_{ij} của ma trận A

★ **Định nghĩa 1.10** Cho ma trận $A = [a_{ij}]_{m \times n}$. Ma trận con A_{ij} của A là ma trận thu được từ ma trận A bằng cách bỏ đi các phần tử nằm ở trên hàng i và các phần tử nằm trên cột j , cỡ của A_{ij} là $(m-1) \times (n-1)$.

• **Ví dụ 1.12** Cho ma trận $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Các ma trận con A_{ij} của A gồm:

$$A_{11} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}, A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}, A_{13} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix}, A_{22} = \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}, A_{23} = \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix},$$

$$A_{31} = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}, A_{32} = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}, A_{33} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}.$$

3. Định nghĩa định thức

★ **Định nghĩa 1.11** Cho ma trận vuông cấp n , $A = [a_{ij}]_{n \times n}$. Định thức của ma trận A được ký hiệu là: $\det(A)$, $|A|$ hoặc

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

xác định bằng phương pháp quy nạp như sau:

i) Với $n = 1$, $\det(A) = a_{11}$

ii) Với $n > 1$, định thức của ma trận A được định nghĩa thông qua các định thức của các ma trận con A_{ij} cấp $n-1$ của nó bằng công thức:

$$\det(A) = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det(A_{1j}).$$

Chú ý:

- 1) Định thức của ma trận vuông cấp n gọi là định thức cấp n .
- 2) Biểu thức định thức cấp 2, định thức cấp 3 và quy tắc Sarius

- Định thức cấp 2:

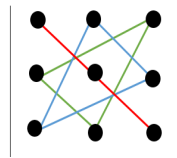
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

- Định thức cấp 3:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

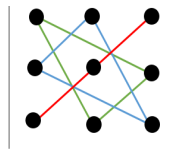
$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

- Quy tắc Sarius cho định thức cấp 3: Là tổng của 6 hạng tử, mà mỗi hạng tử là tích của 3 phần tử mà mỗi dòng, mỗi cột chỉ có một đại diện duy nhất.
 - Các số hạng mang dấu cộng: các số hạng là tích các phần tử nằm trên đường chéo chính hoặc tích các phần tử nằm trên các đỉnh của tam giác có một cạnh song song với đường chéo chính.



Hình 1.1: Quy tắc Sarius- các số hạng mang dấu cộng

- Các số hạng mang dấu trừ: các số hạng là tích các phần tử nằm trên đường chéo phụ hoặc tích các phần tử nằm trên các đỉnh của tam giác có một cạnh song song với đường chéo phụ.



Hình 1.2: Quy tắc Sarius- các số hạng mang dấu trừ

• **Ví dụ 1.13**

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = 1.3.2 + 2.1.3 + 2.1.3 - 3.3.3 - 2.2.2 - 1.1.1 = -18$$

4. **Tính chất của định thức:** Cho $A = [a_{ij}]_{n \times n}$ là ma trận vuông cấp n . Các hàng và cột của A ta cũng sẽ gọi là các hàng và các cột của $\det(A)$. Một hàng hoặc cột gồm toàn số 0 sẽ được gọi là một hàng không (tương ứng: cột không). Định thức có các tính chất dưới đây:

- **Tính chất 1:** Hoán vị 2 hàng (tương ứng 2 cột) của A thì $\det(A)$ đổi dấu.
- **Tính chất 2:** Nếu tất cả các phần tử của hàng i (tương ứng cột j) của định thức được nhân với cùng một số k thì giá trị của định thức mới nhận được bằng giá trị của định thức cũ nhân với k .
- **Tính chất 3:** Thêm vào một hàng (hoặc một cột) k lần một hàng (hoặc k một cột) khác thì định thức không đổi.
- **Tính chất 4:** Định thức của ma trận tam giác trên hoặc dưới bằng tích các phần tử trên đường chéo chính.
- **Tính chất 5:** $\det(A) = \det(A^t)$.
- **Tính chất 6:** Có thể tính định thức của ma trận bằng cách khai triển theo một hàng hoặc một cột tùy ý bởi các công thức sau :
 $\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$ (Công thức khai triển theo hàng thứ i),
 $\det(A) = \sum_{i=1}^n (-1)^{j+i} a_{ij} \det(A_{ij})$ (Công thức khai triển theo cột thứ j).
- **Tính chất 7:** Giả sử A, B là hai ma trận vuông cùng cấp thì

$$\det(AB) = \det(A)\det(B).$$

- **Tính chất 8:** Giả sử hàng i của A biểu diễn dưới dạng:

$$a_{ij} = b_{ij} + c_{ij}.$$

Gọi B là ma trận nhận được từ A bằng cách thay a_{ij} bằng b_{ij} , C là ma trận nhận được từ A bằng cách thay a_{ij} bằng c_{ij} . Khi đó

$$\det(A) = \det(B) + \det(C).$$

Phát biểu tương tự cũng đúng đối với cột.

- **Tính chất 9:** Nếu định thức có một trong các tính chất sau thì định thức bằng không:
 - + Có một hàng (hoặc một cột) bằng không;
 - + Có hai hàng (hoặc hai cột) tỷ lệ;
 - + Có một hàng (hoặc một cột) là tổ hợp tuyến tính của các hàng (hoặc của các cột) khác.

5. **Các ví dụ tính định thức nhờ các tính chất**

• **Ví dụ 1.14** Tính định thức

$$D = \begin{vmatrix} a & x & x & x \\ x & a & x & x \\ x & x & a & x \\ x & x & x & a \end{vmatrix}$$

Lời giải. Cộng tất cả các cột 2, 3, 4 vào cột 1, rút nhân tử chung của cột đầu trong định thức nhận được ta có :

$$D = \begin{vmatrix} a+3x & x & x & x \\ 3x+a & a & x & x \\ 3x+a & x & a & x \\ 3x+a & x & x & a \end{vmatrix} = (3x+a) \begin{vmatrix} 1 & x & x & x \\ 1 & a & x & x \\ 1 & x & a & x \\ 1 & x & x & a \end{vmatrix}$$

Nhân hàng 1 với (-1) rồi cộng lần lượt vào các hàng còn lại, ta được

$$D = (3x+a) \begin{vmatrix} 1 & x & x & x \\ 0 & a-x & 0 & 0 \\ 0 & 0 & a-x & 0 \\ 0 & 0 & 0 & a-x \end{vmatrix}$$

Áp dụng tính chất thứ 4: $D = (3x+a)(a-x)^3$.

•**Ví dụ 1.15** Giải phương trình sau:

$$\begin{vmatrix} 2x & -1 & -x & -x^2 \\ 1 & 2 & 3 & -4 \\ -4 & -1 & 2 & -4 \\ -2 & -1 & 1 & -1 \end{vmatrix} = 0.$$

Lời giải. Khai triển định thức trên theo hàng 1, ta được:

$$2x \begin{vmatrix} 2 & 3 & -4 \\ -1 & 2 & -4 \\ -1 & 1 & -1 \end{vmatrix} - (-1) \begin{vmatrix} 1 & 3 & -4 \\ -4 & 2 & -4 \\ -2 & 1 & -1 \end{vmatrix} + (-x) \begin{vmatrix} 1 & 2 & -4 \\ -4 & -1 & -4 \\ -2 & -1 & -1 \end{vmatrix} - (-x^2) \begin{vmatrix} 1 & 2 & 3 \\ -4 & -1 & 2 \\ -2 & -1 & 1 \end{vmatrix} = 0$$

$$\Leftrightarrow 7x^2 + 21x + 14 = 0 \Leftrightarrow x = -1, x = -2.$$

•**Ví dụ 1.16** Tính định thức

$$D = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 3 & 6 & 7 & -2 \\ 2 & 7 & -8 & 15 \\ -4 & -6 & 5 & -2 \end{vmatrix}$$

Lời giải. Lần lượt lấy $H_2 - 3H_1 \rightarrow H_2, H_3 - 2H_1 \rightarrow H_3$ và $H_4 + 4H_1 \rightarrow H_4$, ta được:

$$D = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 10 & -11 \\ 0 & 3 & -6 & 9 \\ 0 & 2 & 1 & 10 \end{vmatrix}$$

Tiếp theo, $H_3 : 3 \rightarrow H_3$ và $H_2 \leftrightarrow H_3$

$$D = -3 \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 10 & -11 \\ 0 & 2 & 1 & 10 \end{vmatrix}$$

Lấy $H_4 - 2H_2 \rightarrow H_4$

$$D = -3 \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 10 & -11 \\ 0 & 0 & 5 & 4 \end{vmatrix}$$

Tiếp tục thực hiện $H_3 \leftrightarrow H_4$ và $H_4 - 2H_3 \rightarrow H_4$, thu được định thức của ma trận tam giác trên, áp dụng tính chất 4, ta có định thức cần tìm

$$D = 3 \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & -19 \end{vmatrix} = 3 \cdot 1 \cdot 1 \cdot 5 \cdot (-19) = -285$$

•**Ví dụ 1.17** Tính định thức

$$D = \begin{vmatrix} 1 & 4 & 9 & 16 \\ 4 & 9 & 16 & 25 \\ 9 & 16 & 25 & 36 \\ 16 & 25 & 36 & 49 \end{vmatrix}$$

Lời giải. Thực hiện lần lượt các phép biến đổi: $H_4 - H_3 \rightarrow H_4, H_3 - H_2 \rightarrow H_3, H_2 - H_1 \rightarrow H_2$, ta được

$$D = \begin{vmatrix} 1 & 4 & 9 & 16 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 11 \\ 7 & 9 & 11 & 13 \end{vmatrix}$$

Tiếp theo, lấy $H_4 - H_3 \rightarrow H_4, H_3 - H_2 \rightarrow H_3$ ta thu được 2 hàng giống nhau, do đó

$$D = \begin{vmatrix} 1 & 4 & 9 & 16 \\ 3 & 5 & 7 & 9 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{vmatrix} = 0.$$

•**Ví dụ 1.18** Tính định thức

$$D = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & x \\ 4 & 9 & 16 & x^2 \\ 8 & 27 & 64 & x^3 \end{vmatrix}$$

Lời giải. Khai triển định thức D theo cột 4, ta thu được đa thức bậc 3 với ẩn x và hệ số cao nhất là

$$\begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 4 & 9 & 16 \end{vmatrix} = 2.$$

Cho x lần lượt nhận các giá trị $x = 2, x = 3, x = 4$ ta thấy D có 2 cột giống nhau, do đó $D = 0$ khi $x = 2, x = 3, x = 4$. Theo định lý Bezout ta phải có:

$$D = 2(x - 2)(x - 3)(x - 4).$$

1.3.2. Ma trận nghịch đảo

1. Định nghĩa

★ **Định nghĩa 1.12** Cho A là ma trận vuông cấp n . Nếu có ma trận B vuông cùng cấp sao cho:

$$AB = BA = I_n$$

thì ta nói A khả đảo và B được gọi là ma trận nghịch đảo của của ma trận A . Ký hiệu ma trận nghịch đảo của ma trận A là A^{-1} .

Như vậy $AA^{-1} = A^{-1}A = I_n$.

Khi A khả đảo ta nói A là ma trận không suy biến.

• **Ví dụ 1.19** Xét $A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 3/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$. Ta có $AB = BA = I_2$ nên theo định nghĩa $B = A^{-1}$.

2. Tính chất

◇ **Định lý 1.1** Ma trận nghịch đảo của ma trận vuông A nếu tồn tại thì duy nhất.

△. Giải sử B và B' là hai ma trận cùng thỏa mãn định nghĩa của ma trận nghịch đảo của ma trận A . Khi đó:

$$AB = BA = I, AB' = B'A = I$$

Vậy

$$B' = IB' = (BA)B' = B(AB') = BI = B.$$

□

◇ **Định lý 1.2** Nếu ma trận vuông A khả đảo thì $\det(A) \neq 0$.

△. Vì A khả đảo nên tồn tại A^{-1} và $AA^{-1} = I_n$. Áp dụng công thức tính định thức của tích hai ma trận ta có:

$$\det(AA^{-1}) = \det(I_n) \Rightarrow \det(A)\det(A^{-1}) = 1.$$

Vậy $\det(A) \neq 0$ và hơn nữa $\det(A^{-1}) = \frac{1}{\det(A)}$.

□

◇ **Định lý 1.3** Nếu ma trận vuông A cấp n có $\det(A) \neq 0$ thì A khả đảo và

$$A^{-1} = \frac{1}{\det(A)} C^t$$

ở đó $C = [c_{ij}]_{n \times n}$ là ma trận phụ hợp của ma trận A : $c_{ij} = (-1)^{i+j} \det(A_{ij})$.

△. Áp dụng công thức khai triển định thức theo hàng thứ i ta có

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$$

$$\Rightarrow a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in} = \det(A).$$

Hơn nữa, do định thức có hai hàng giống nhau thì bằng không nên suy ra

$$a_{k1}c_{i1} + a_{k2}c_{i2} + \cdots + a_{kn}c_{in} = \begin{cases} \det(A) & \text{nếu } k = n \\ 0 & \text{nếu } k \neq n \end{cases}$$

Do đó $AC^t = \det(A)I$.

Áp dụng công thức khai triển định thức theo cột và lập luận tương tự ta cũng có:

$$C^t A = \det(A) \cdot I$$

Như vậy

$$A \left(\frac{1}{\det(A)} C^t \right) = \left(\frac{1}{\det(A)} C^t \right) A = I$$

suy ra điều phải chứng minh. □

◇ **Định lý 1.4** Giả sử các ma trận vuông A, B là các ma trận khả đảo. Khi đó:

- (a) A^{-1} cũng khả đảo và $(A^{-1})^{-1} = A$;
- (b) $\forall m \in \mathbb{N}$, A^m cũng khả đảo và $(A^m)^{-1} = (A^{-1})^m$;
- (c) $\forall k \neq 0$ ta có kA cũng khả đảo và $(kA)^{-1} = \frac{1}{k}A^{-1}$;
- (d) AB cũng khả đảo và $(AB)^{-1} = B^{-1}A^{-1}$.

Δ. Bằng việc kiểm tra trực tiếp định nghĩa ma trận nghịch đảo, ta dễ dàng thu được các tính chất trên. □

3. Các ví dụ tính ma trận nghịch đảo bằng ma trận phụ hợp

● **Ví dụ 1.20** Công thức ma trận nghịch đảo của ma trận vuông cấp 2 khả đảo:

Giả sử $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ có $\det(A) = ad - bc \neq 0$, áp dụng công thức ta có

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}^t = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Chẳng hạn:

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}^{-1} = \frac{1}{1 \cdot 7 - 3 \cdot 5} \begin{bmatrix} 7 & -3 \\ -5 & 1 \end{bmatrix} = \begin{bmatrix} \frac{-7}{8} & \frac{3}{8} \\ \frac{5}{8} & \frac{-1}{8} \end{bmatrix}$$

● **Ví dụ 1.21** Tìm ma trận nghịch đảo của ma trận sau (nếu có):

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 3 \\ 0 & 3 & 1 \end{bmatrix}$$

Lời giải. Ta có $\det(A) = -3 \neq 0$ nên A khả đảo. Áp dụng công thức ta có:

$$\begin{aligned} c_{11} &= -7, & c_{12} &= -2, & c_{13} &= 6 \\ c_{21} &= 2, & c_{22} &= 1, & c_{23} &= -3 \\ c_{31} &= 1, & c_{32} &= -1, & c_{33} &= 0 \end{aligned}$$

Do đó

$$C = \begin{bmatrix} -7 & -2 & 6 \\ 2 & 1 & -3 \\ 1 & -1 & 0 \end{bmatrix} \Rightarrow C^t = \begin{bmatrix} -7 & 2 & 1 \\ -2 & 1 & -1 \\ 6 & -3 & 0 \end{bmatrix}$$

Vậy ma trận nghịch đảo của ma trận A là

$$A^{-1} = \frac{1}{-3} \begin{bmatrix} -7 & 2 & 1 \\ -2 & 1 & -1 \\ 6 & -3 & 0 \end{bmatrix} = \begin{bmatrix} \frac{7}{3} & \frac{-2}{3} & \frac{-1}{3} \\ \frac{2}{3} & \frac{-1}{3} & \frac{1}{3} \\ -2 & 1 & 0 \end{bmatrix}$$

• **Ví dụ 1.22** Tìm ma trận X , biết

$$\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} X = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & -4 \end{bmatrix}$$

Lời giải. Đặt

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & -4 \end{bmatrix}$$

Ta có $\det(A) = 5 \neq 0$ nên A khả đảo và

$$A^{-1} = \frac{1}{5} \begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix}$$

Phương trình ma trận trở thành

$$AX = B$$

Nhân vào bên trái cả hai vế A^{-1} ta được:

$$A^{-1}(AX) = A^{-1}B \Leftrightarrow (A^{-1}A)X = A^{-1}B \Leftrightarrow IX = A^{-1}B \Leftrightarrow X = A^{-1}B$$

Vậy ma trận X cần tìm là

$$X = \frac{1}{5} \begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & -4 \end{bmatrix} = \begin{bmatrix} \frac{-3}{5} & -1 & \frac{14}{5} \\ \frac{5}{5} & 0 & \frac{-2}{5} \end{bmatrix}$$

1.3.3. Hạng của ma trận

1. Định nghĩa

* **Định nghĩa 1.13** Cho A là ma trận cỡ $m \times n$. Ma trận con cấp p của A là ma trận có được từ A sau khi bỏ đi $m - p$ hàng và $n - p$ cột. Định thức của ma trận đó được gọi là định thức con cấp p của A .

* **Định nghĩa 1.14** Hạng của ma trận A là cấp cao nhất của định thức con khác không của ma trận A , ký hiệu là $r(A)$.

• **Ví dụ 1.23** Cho ma trận

$$A = \begin{bmatrix} 2 & 0 & 3 & 1 \\ 1 & 5 & -1 & 2 \\ 5 & 5 & 5 & 4 \end{bmatrix}$$

Các định thức con cấp 3 của ma trận A là:

$$\begin{vmatrix} 0 & 3 & 1 \\ 5 & -1 & 2 \\ 5 & 5 & 4 \end{vmatrix} = 0, \begin{vmatrix} 2 & 3 & 1 \\ 1 & -1 & 2 \\ 5 & 5 & 4 \end{vmatrix} = 0, \begin{vmatrix} 2 & 0 & 1 \\ 1 & 5 & 2 \\ 5 & 5 & 4 \end{vmatrix} = 0, \begin{vmatrix} 2 & 0 & 3 \\ 1 & 5 & -1 \\ 5 & 5 & 4 \end{vmatrix} = 0$$

nên $r(A) < 3$. Hơn nữa, có ít nhất một định thức con cấp 2 của A

$$\begin{vmatrix} 2 & 0 \\ 1 & 5 \end{vmatrix} \neq 0$$

nên ta có $r(A) = 2$.

★ **Định nghĩa 1.15** Ma trận bậc thang là ma trận có các đặc điểm sau:

- Các hàng bằng không ở dưới các hàng khác không.
- Đối với hai hàng khác không liên tiếp, phần tử khác 0 đầu tiên (kể từ trái sang) của hàng trên nằm bên trái phần tử khác 0 đầu tiên của hàng dưới

• **Ví dụ 1.24**

$$A = \begin{bmatrix} 2 & 0 & 3 & 1 \\ 0 & 5 & -1 & 2 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

là ma trận bậc thang có 3 hàng khác 0.

2. Tính chất của hạng ma trận A

- **Tính chất 1:** $r(A) = r(A^t)$.
- **Tính chất 2:** Hạng của ma trận bậc thang không đổi khi thực hiện các phép biến đổi sơ cấp.
- **Tính chất 3:** Hạng của ma trận bậc thang bằng số hàng khác không của ma trận đó.

Quy tắc thực hành tìm hạng của ma trận A : Để tìm hạng của ma trận A ta sử dụng 3 phép biến đổi sơ cấp đưa ma trận A về ma trận bậc thang B . Khi đó $r(A) = r(B) =$ Số hàng khác không của B .

3. Các ví dụ tính hạng của ma trận

• **Ví dụ 1.25** Tính hạng của ma trận sau:

$$A = \begin{bmatrix} 1 & -2 & 3 & 4 \\ 2 & -4 & 1 & 3 \\ -1 & 2 & 0 & -1 \end{bmatrix}$$

Lời giải.

$$A = \begin{bmatrix} 1 & -2 & 3 & 4 \\ 2 & -4 & 1 & 3 \\ -1 & 2 & 0 & -1 \end{bmatrix} \xrightarrow[\begin{matrix} H_2 - 2H_1 \rightarrow H_2 \\ H_3 + H_1 \rightarrow H_3 \end{matrix}]{\begin{matrix} H_3 + H_1 \rightarrow H_3 \\ H_2 - 2H_1 \rightarrow H_2 \end{matrix}} \begin{bmatrix} 1 & -2 & 3 & 4 \\ 0 & 0 & -5 & -5 \\ 0 & 0 & 3 & 3 \end{bmatrix} \xrightarrow{5H_3 + 3H_2 \rightarrow H_3} \begin{bmatrix} 1 & -2 & 3 & 4 \\ 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ma trận cuối cùng là ma trận bậc thang với 2 hàng khác 0. Vậy $r(A) = 2$.

Dạng ma trận của hệ phương trình tuyến tính

Đặt:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}; x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}; b = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$$

thì hệ phương trình tuyến tính m phương trình, n ẩn là hệ phương trình có dạng ma trận:

$$Ax = b$$

trong đó A là ma trận hệ số; x là ma trận ẩn; b là ma trận vế phải.**2. Giải hệ phương trình bằng phương pháp Cramer**

◇ **Định lý 1.5** (Định lý Cramer) Hệ phương trình tuyến tính $Ax = b$, với A là ma trận vuông không suy biến, có nghiệm duy nhất :

$$x_j = \frac{\det(A_j)}{\det(A)}, j = 1, 2, \dots, n$$

trong đó A_j là ma trận có được từ A sau khi thay cột thứ j bởi cột vế phải b .

Δ. Vì A không suy biến, $\det(A) \neq 0$ nên A có ma trận nghịch đảo:

$$A^{-1} = \frac{1}{\det(A)} C^t.$$

Từ phương trình $Ax = b$, nhân bên trái hai vế với A^{-1} ta có

$$A^{-1}(Ax) = A^{-1}b \leftrightarrow (A^{-1}A)x = A^{-1}b \leftrightarrow Ix = A^{-1}b \leftrightarrow x = A^{-1}b$$

Vậy $x = A^{-1}b$ là nghiệm của hệ phương trình. Do A^{-1}, b xác định nên nghiệm này là duy nhất.

Sử dụng biểu thức của ma trận A^{-1} ta suy ra

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}^t \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}$$

nghĩa là

$$x_j = \frac{c_{1j}b_1 + c_{2j}b_2 + \dots + c_{nj}b_n}{\det(A)} = \frac{\det(A_j)}{\det(A)}$$

với $j = 1, 2, \dots, n$ và ta có điều phải chứng minh. □

•**Ví dụ 1.27** Giải hệ phương trình sau:

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 3 \\ 2x_1 + 3x_2 + x_3 = -1 \\ 3x_1 + x_2 + 2x_3 = 5 \end{cases}$$

Lời giải. Ta có

$$D = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = -18; D_1 = \begin{vmatrix} 3 & 2 & 3 \\ -1 & 3 & 1 \\ 5 & 1 & 2 \end{vmatrix} = -19$$

Các phép biến đổi nói trên tương ứng với các phép biến đổi sơ cấp trên các hàng của ma trận bổ sung.

Phương pháp Gauss giải hệ phương trình tuyến tính:

- **Bước 1:** Lập ma trận \bar{A} bổ sung của hệ
- **Bước 2:** Dùng các phép biến đổi sơ cấp trên các hàng của ma trận bổ sung, đưa \bar{A} về ma trận bậc thang.
- **Bước 3:** Giải hệ phương trình bậc thang: Giải ngược từ phương trình cuối, sử dụng phương pháp thế ta có nghiệm cần tìm.

• **Ví dụ 1.29** Giải hệ phương trình bằng phương pháp Gauss

$$\begin{cases} 2x_1 + x_2 + x_3 = 2 \\ x_1 + 3x_2 + x_3 = 5 \\ x_1 + x_2 + 5x_3 = -7 \\ 2x_1 + 3x_2 - 3x_3 = 14 \end{cases}$$

Lời giải.

$$\begin{array}{ccc|c} 2 & 1 & 1 & 2 \\ 1 & 3 & 1 & 5 \\ 1 & 1 & 5 & -7 \\ 2 & 3 & -3 & 14 \\ \hline 1 & 1 & 5 & -7 \\ 1 & 3 & 1 & 5 \\ 2 & 1 & 1 & 2 & H_1 \leftrightarrow H_3 \\ 2 & 3 & -3 & 14 \\ \hline 1 & 1 & 5 & -7 \\ 0 & 2 & -4 & 12 & H_2 - H_1 \rightarrow H_2 \\ 0 & -1 & -9 & 16 & H_3 - 2H_1 \rightarrow H_3 \\ 0 & 1 & -13 & 28 & H_4 - 2H_1 \rightarrow H_4 \\ \hline 1 & 1 & 5 & -7 \\ 0 & 1 & -2 & 6 & H_2 : 2 \rightarrow H_2 \\ 0 & -1 & -9 & 16 \\ 0 & 1 & -13 & 28 \\ \hline 1 & 1 & 5 & -7 \\ 0 & 1 & -2 & 6 \\ 0 & 0 & -11 & 22 & H_3 + H_2 \rightarrow H_3 \\ 0 & 0 & -11 & 22 & H_4 - H_2 \\ \hline 1 & 1 & 5 & -7 \\ 0 & 1 & -2 & 6 \\ 0 & 0 & -11 & 22 \\ 0 & 0 & 0 & 0 & H_4 - H_3 \rightarrow H_4 \end{array}$$

Từ ma trận bậc thang cuối cùng ta nhận được hệ phương trình bậc thang tương đương

$$\begin{cases} x_1 + x_2 + 5x_3 = -7 \\ x_2 - 2x_3 = 6 \\ -11x_3 = 22 \end{cases}$$

Giải hệ bậc thang trên ta có nghiệm duy nhất

$$\begin{cases} x_1 = 1 \\ x_2 = 2 \\ x_3 = -2 \end{cases}$$

4. Giải và biện luận hệ phương trình bằng định lý Kronecker-Capelli

◇ **Định lý 1.6 (Định lý Kronecker-Capelli)** Xét hệ phương trình tuyến tính

$$Ax = b$$

với A là ma trận hệ số, \bar{A} là ma trận bổ sung tương ứng. Khi đó:

- Hệ phương trình vô nghiệm khi và chỉ khi $r(A) \neq r(\bar{A})$;
- Hệ phương trình có nghiệm duy nhất khi và chỉ khi $r(A) = r(\bar{A}) = \text{số ẩn}$;
- Hệ phương trình vô số nghiệm khi và chỉ khi $r(A) = r(\bar{A}) < \text{số ẩn}$.

•**Ví dụ 1.30** Giải hệ phương trình sau

$$\begin{cases} 2x_1 + 4x_2 + 5x_3 = 4 \\ 3x_1 + x_2 - 2x_3 = -2 \\ 4x_1 + 11x_2 + 7x_3 = 7 \end{cases}$$

Lời giải. Ta có

$$\begin{array}{ccc|c} 2 & 4 & 3 & 4 \\ 3 & 1 & -2 & -2 \\ 4 & 11 & 7 & 7 \\ \hline 2 & 4 & 3 & 4 \\ 0 & 10 & 13 & 16 & 2H_2 - 3H_1 \rightarrow H_2 \\ 0 & 3 & 1 & -1 & H_3 - 2H_1 \rightarrow H_3 \\ \hline 2 & 4 & 3 & 4 \\ 0 & 3 & 1 & -1 & H_3 \leftrightarrow H_2 \\ 0 & 0 & -29 & -58 & 10H_3 - 3H_2 \rightarrow H_3 \end{array}$$

Hệ đã cho tương đương với

$$\begin{cases} 2x_1 + 4x_2 + 5x_3 = 4 \\ 3x_2 + x_3 = -1 \\ -29x_3 = -58 \end{cases} \Leftrightarrow \begin{cases} x_1 = 1 \\ x_2 = -1 \\ x_3 = 2 \end{cases}$$

•**Ví dụ 1.31** Giải hệ phương trình sau

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 2 \\ -2x_1 + x_2 + 4x_3 = 1 \\ -x_1 + 3x_2 + x_3 = -4 \end{cases}$$

Lời giải. Ta có

$$\begin{array}{ccc|c} 1 & 2 & -3 & 2 \\ -2 & 1 & 4 & 1 \\ -1 & 3 & 1 & -4 \\ \hline 1 & 2 & -3 & 2 \\ 0 & 5 & -2 & 5 & H_2 + 2H_1 \rightarrow H_2 \\ 0 & 5 & -2 & -2 & H_3 + H_1 \rightarrow H_3 \\ \hline 1 & 2 & -3 & 2 \\ 0 & 5 & -2 & 5 \\ 0 & 0 & 0 & -7 & H_3 - H_2 \rightarrow H_3 \end{array}$$

Vì $r(A) = 2 \neq r(\bar{A}) = 3$ nên hệ phương trình đã cho vô nghiệm.

•**Ví dụ 1.32** Giải hệ phương trình sau

$$\begin{cases} x_1 - x_2 + 2x_3 - 2x_4 = 1 \\ -2x_1 - 3x_2 + x_3 + 4x_4 = -2 \\ -2x_1 + 2x_2 - 4x_3 + 4x_4 = -2 \\ 3x_1 + 2x_2 + x_3 - 6x_4 = 3 \end{cases}$$

Lời giải.

$$\begin{array}{cccc|c} 1 & -1 & 2 & -2 & 1 \\ -2 & -3 & 1 & 4 & -2 \\ -2 & 2 & -4 & 4 & -2 \\ 3 & 2 & 1 & -6 & 3 \\ \hline 1 & -1 & 2 & -2 & 1 \\ 0 & -5 & 5 & 0 & 0 & H_2 + 2H_1 \rightarrow H_2 \\ 0 & 0 & 0 & 0 & 0 & H_3 + 2H_1 \rightarrow H_3 \\ 0 & 5 & -5 & 0 & 0 & H_4 - 3H_1 \rightarrow H_4 \\ \hline 1 & -1 & 2 & -2 & 1 \\ 0 & -5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & H_4 + H_2 \rightarrow H_4 \end{array}$$

Do $r(A) = r(\bar{A}) = 2 < \text{số ẩn}$ nên hệ phương trình đã cho vô số nghiệm. Chọn x_4 và x_2 là hai ẩn tự do, ta có hệ phương trình tương đương

$$\begin{cases} x_1 + 2x_3 = 1 + x_2 + 2x_4 \\ x_3 = x_2 \end{cases} \leftrightarrow \begin{cases} x_1 = 1 - x_2 + 2x_4 \\ x_3 = x_2 \\ x_2, x_4 \in \mathbb{R} \end{cases}$$

•**Ví dụ 1.33** Giải và biện luận theo m số nghiệm của hệ phương trình sau :

$$\begin{cases} x_1 - x_2 - x_3 - 3x_4 + x_5 = 1 \\ x_1 + x_2 - 5x_3 - x_4 + 7x_5 = 2 \\ -x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 = 0 \\ -2x_1 + 5x_2 - 4x_3 + 9x_4 + 7x_5 = m \end{cases}$$

Lời giải. Xét ma trận bổ sung

$$\bar{A} = \left[\begin{array}{ccccc|c} 1 & -1 & -1 & -3 & 1 & 1 \\ 1 & 1 & -5 & -1 & 7 & 2 \\ -1 & 2 & 2 & 2 & 1 & 0 \\ -2 & 5 & -4 & 9 & 7 & m \end{array} \right]$$

Dùng các phép biến đổi sơ cấp về hàng đưa \bar{A} về ma trận bậc thang

$$\left[\begin{array}{ccccc|c} 1 & -1 & -1 & -3 & 1 & 1 \\ 0 & 2 & -4 & 2 & 6 & 1 \\ 0 & 0 & 6 & -4 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2m+1 \end{array} \right]$$

Biện luận:

+Nếu $2m + 1 = 0 \Leftrightarrow m = -\frac{1}{2}$ thì $r(A) = r(\bar{A}) = 3 < 5$ (số ẩn) \Leftrightarrow Hệ phương trình vô số nghiệm.

+Nếu $2m + 1 \neq 0 \Leftrightarrow m \neq -\frac{1}{2}$ thì $r(A) \neq r(\bar{A}) \Leftrightarrow$ Hệ phương trình vô nghiệm.

Bài tập chương 1

Bài 1.1. Tính AB và BA (nếu tồn tại), biết rằng:

a) $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 2 \end{bmatrix}, B = \begin{bmatrix} 0 & -1 \\ 2 & 3 \\ 4 & 1 \end{bmatrix}$

b) $A = \begin{bmatrix} -1 & 0 & 2 \\ 1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & -2 & 2 \\ 2 & 1 & 2 & -1 \\ 3 & -2 & 1 & 0 \end{bmatrix}.$

Bài 1.2. Thực hiện phép nhân hai ma trận sau:

a) $\begin{bmatrix} 1 & 1 & -2 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 1 & -2 & 1 \\ 0 & -1 & -4 \end{bmatrix}$ b) $\begin{bmatrix} -1 \\ -2 \\ -3 \\ -4 \end{bmatrix} [1 \ 2 \ 3 \ 4]$ c) $[2 \ -1 \ 1] \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$

d) $\begin{bmatrix} 2 & -1 & 3 \\ 5 & 4 & -1 \end{bmatrix} \begin{bmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & 1 \end{bmatrix}$ e) $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 1 \\ 1 & -5 & 6 \\ -1 & 4 & 3 \end{bmatrix}.$

Bài 1.3. Tìm ma trận X giao hoán với ma trận A , nghĩa là thỏa mãn hệ thức $XA = AX$, biết:

a) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$

Bài 1.4. Tìm ma trận X thỏa mãn điều kiện $AX = B$, biết:

a) $A = \begin{bmatrix} -2 & 3 \\ -3 & 4 \end{bmatrix}, B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$

b) $A = \begin{bmatrix} 5 & -4 \\ 4 & -3 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}, B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}.$

Bài 1.5. Tìm ma trận X thỏa mãn điều kiện:

a) $A^2 = \Theta$

b) $A^2 = I_2.$

Bài 1.6. Tìm:

a) $\begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix}^n$ b) $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}^n$ c) $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}^n$ d) $\begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}^{100}.$

Bài 1.7. Tính các định thức sau:

a) $\begin{vmatrix} 3 & 2 & 1 \\ 2 & 5 & 3 \\ 3 & 4 & 2 \end{vmatrix}, b) \begin{vmatrix} 4 & -3 & 5 \\ 3 & -2 & 8 \\ 1 & -7 & -5 \end{vmatrix}, c) \begin{vmatrix} 4 & 2 & -1 \\ 5 & 3 & -2 \\ 3 & 2 & -1 \end{vmatrix}$

d) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix}, e) \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}, f) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{vmatrix}$

$$g) \begin{vmatrix} \frac{3}{5} & -1 & \frac{-2}{5} & \frac{2}{5} \\ -1 & \frac{7}{4} & 1 & 1 \\ \frac{4}{3} & -3 & -1 & \frac{7}{3} \\ 1 & -3 & -1 & \frac{2}{3} \end{vmatrix}, h) \begin{vmatrix} -1 & 2 & -1 & 3 \\ \frac{2}{3} & \frac{-9}{2} & \frac{4}{5} & \frac{5}{7} \\ \frac{3}{7} & \frac{-12}{7} & \frac{5}{5} & \frac{2}{15} \\ \frac{1}{3} & \frac{-5}{2} & \frac{2}{5} & \frac{3}{2} \end{vmatrix}$$

Bài 1.8. Tính các định thức sau:

$$a) \begin{vmatrix} a & 3 & 0 & 5 \\ 0 & b & 0 & 2 \\ 1 & 2 & c & 3 \end{vmatrix}, b) \begin{vmatrix} 1 & 0 & 2 & a \\ 2 & 0 & b & 0 \\ 3 & c & 4 & 5 \\ d & 0 & 0 & 0 \end{vmatrix}, c) \begin{vmatrix} l & h & k & u & g \\ d & y & 0 & 0 & 0 \\ f & e & z & 0 & 0 \\ c & a & b & 0 & x \\ v & 0 & 0 & 0 & 0 \end{vmatrix}$$

Bài 1.9. Tính định thức sau:

$$\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$$

trong đó a, b, c là ba nghiệm của phương trình bậc ba $x^3 + px + q = 0$.

Bài 1.10. Tính các định thức sau:

$$a) \begin{vmatrix} 1 & x & x+1 & -x \\ x & 2 & x-1 & x+1 \\ 2 & 2+x & 2x & 1 \\ 3 & -x & 1 & x \end{vmatrix}, b) \begin{vmatrix} \cos x & 1 & 0 & 0 \\ 1 & 2\cos x & 1 & 0 \\ 0 & 1 & 2\cos x & 1 \\ 0 & 0 & 1 & 2\cos x \end{vmatrix}, c) \begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1+z & 1 \\ 1 & 1 & 1 & 1-z \end{vmatrix}$$

Bài 1.11. Chứng minh rằng:

$$D_n = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod_{i>j} (x_i - x_j)$$

Bài 1.12. Áp dụng kết quả bài tập số 6, hãy tính các định thức sau và để kết quả ở dạng tích:

$$a) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & x \\ 4 & 9 & 16 & x^2 \\ 8 & 27 & 64 & x^3 \end{vmatrix}, b) \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1+1 & x_2+1 & x_3+1 & x_4+1 \\ x_1^2+x_1 & x_2^2+x_2 & x_3^2+x_3 & x_4^2+x_4 \\ x_1^3+x_1^2 & x_2^3+x_2^2 & x_3^3+x_3^2 & x_4^3+x_4^2 \end{vmatrix}$$

Bài 1.13. Tìm ma trận nghịch đảo của ma trận sau nếu có:

$$a) \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 3 \\ 2 & 1 & 1 \end{bmatrix} \quad b) \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \quad c) \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & -3 & 2 \end{bmatrix} \quad d) \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

$$e) \begin{bmatrix} 1 & -2 & 1 & -1 \\ -1 & 4 & -2 & 3 \\ 2 & 0 & 1 & 3 \\ -2 & 6 & 0 & 5 \end{bmatrix} \quad f) \begin{bmatrix} 2 & -1 & 0 & 3 \\ 1 & 1 & 2 & -1 \\ -1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 1 \end{bmatrix}$$

Bài 1.14. Giải các phương trình ma trận sau

$$a) \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} X = \begin{bmatrix} 4 & -6 \\ 2 & 1 \end{bmatrix} \quad b) X \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 3 \\ 4 & 3 & 2 \\ 1 & -2 & 5 \end{bmatrix} \quad c) \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & 1 \\ -2 & 3 & 1 \end{bmatrix} X = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 0 & 2 & 2 \\ 1 & -2 & 2 & 0 \end{bmatrix}$$

Bài 1.15. Tìm ma trận nghịch đảo của ma trận $A = BC$ nếu:

$$B = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

Bài 1.16. Cho các ma trận

$$A = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 2 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 1 & -1 \\ 2 & 3 \end{bmatrix}, C = \begin{bmatrix} 1 & -2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix}$$

Tìm ma trận X biết $(XA)B = C$.

Bài 1.17. a) Hãy tính A^5 biết $A = BCB^{-1}$ với $B = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}, C = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

b) Hãy tính A^{2017} biết $A = BCB^{-1}$ với

$$B = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

c) Hãy tính A^{2018} biết $A = BCB^{-1}$ với

$$B = \begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Bài 1.18. Tìm hạng của các ma trận sau

$$a) \begin{bmatrix} 2 & -1 & 3 & -2 & 4 \\ 4 & -2 & 5 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{bmatrix} \quad b) \begin{bmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \\ 7 & 7 & 9 & 1 \end{bmatrix} \quad c) \begin{bmatrix} 4 & 3 & -5 & 2 & 3 \\ 8 & 6 & -7 & 4 & 2 \\ 4 & 3 & -8 & 2 & 7 \\ 4 & 3 & 1 & 2 & -5 \\ 8 & 6 & -1 & 4 & -6 \end{bmatrix}$$

Bài 1.19. Tìm m để hạng của ma trận sau là bé nhất

$$A = \begin{bmatrix} 3 & 1 & 1 & 4 & -1 \\ m & 4 & 10 & 1 & m+5 \\ 1 & 7 & 17 & 3 & m+5 \\ 2 & 2 & 4 & 3 & 1 \end{bmatrix}$$

Bài 1.20. Biện luận theo m hạng của ma trận A , biết:

$$A = \begin{bmatrix} 1 & -4 & m+21 & -6 \\ 1 & 1 & 3 & 2 \\ 1 & 2 & -1 & 3 \\ 2 & 3 & -1 & -1 \\ 3 & -1 & m+23 & -1 \end{bmatrix}$$

Bài 1.21. Giải các hệ phương trình tuyến tính sau bằng phương pháp Cramer

$$\begin{aligned}
 a) \begin{cases} x_1 - x_2 + 2x_3 = 1 \\ 3x_1 - 2x_2 + 5x_3 = 2 \\ -x_1 + x_2 - x_3 = 2 \end{cases} & \quad b) \begin{cases} x_1 - 2x_3 = 1 \\ -2x_1 + x_2 + 6x_3 = 11 \\ -x_1 + 5x_2 - 4x_3 = -4 \end{cases} \\
 c) \begin{cases} 4x_1 + x_2 + 2x_3 = 1 \\ x_1 - x_3 = 2 \\ 6x_1 + x_2 + 4x_3 = 3 \end{cases} & \quad d) \begin{cases} 3x_1 + 5x_2 - 2x_3 = 2 \\ 2x_1 - 7x_2 + 2x_3 = 12 \\ -x_1 + 5x_2 + 3x_3 = 9 \end{cases}
 \end{aligned}$$

Bài 1.22. Giải và biện luận các hệ phương trình tuyến tính sau bằng phương pháp Cramer

$$\begin{aligned}
 a) \begin{cases} x_1 + 3x_2 + (5-m)x_3 = 1 \\ 2x_1 + x_2 + 2x_3 = 0 \\ (3m-1)x_1 + (m+3)x_2 + 4x_3 = 2 \end{cases} \\
 b) \begin{cases} x_1 + (m-1)x_2 - 3x_3 = 1 \\ 2x_1 - 4x_2 + (4m-2)x_3 = -1 \\ 3x_1 + (m+1)x_2 - 9x_3 = 0 \end{cases}
 \end{aligned}$$

Bài 1.23. Giải các hệ phương trình tuyến tính sau bằng phương pháp Gauss

$$\begin{aligned}
 a) \begin{cases} x_1 + x_2 + 2x_4 = 5 \\ 2x_1 + 4x_2 - x_3 + 5x_4 = -1 \\ x_1 + 3x_2 + 5x_4 = -3 \\ 3x_1 + 7x_2 - 3x_3 + 9x_4 = -14 \\ x_1 + 4x_2 - 2x_3 + x_4 = -11 \end{cases} & \quad b) \begin{cases} 3x_1 - x_2 - x_3 + 2x_4 = 1 \\ x_1 - x_2 - 2x_3 + 4x_4 = 5 \\ x_1 + x_2 + 3x_3 - 6x_4 = -9 \\ 12x_1 - 2x_2 + x_3 - 2x_4 = -10 \end{cases} \\
 c) \begin{cases} 4x_1 + 2x_2 + x_3 - 3x_4 = 7 \\ x_1 - x_2 + x_3 + 2x_4 = 5 \\ 2x_1 + 3x_2 - 3x_3 + x_4 = 3 \\ 4x_1 + x_2 - x_3 + 5x_4 = 1 \end{cases} & \quad d) \begin{cases} x_1 - 3x_2 + 5x_3 = -21 \\ -3x_1 + 5x_2 + 6x_3 = 5 \\ 4x_1 - 3x_2 - 7x_3 = 6 \\ 2x_1 - 4x_2 - 3x_3 = 0 \end{cases}
 \end{aligned}$$

Bài 1.24. Giải và biện luận các hệ phương trình tuyến tính sau bằng phương pháp Gauss

$$\begin{aligned}
 a) \begin{cases} x_1 - x_2 + x_3 = 4 \\ x_1 + 4x_2 + 2x_3 = 1 \\ x_1 - x_2 + mx_3 = 2(m+1) \end{cases} & \quad b) \begin{cases} x_1 + x_2 + 2x_3 = 2 - m \\ x_1 - 2x_2 - x_3 = m + 1 \\ -x_1 + 5x_2 + 4mx_3 = -3m \\ x_1 + x_2 + 4x_3 = 7 \end{cases}
 \end{aligned}$$

Bài 1.25. Tìm điều kiện để các hệ phương trình tuyến tính thuần nhất sau có nghiệm không tầm thường

$$\begin{aligned}
 a) \begin{cases} x_1 - mx_2 + x_3 - (m+3)x_4 = 0 \\ 2x_1 + x_2 - 4x_3 + 7x_4 = 0 \\ mx_1 + 4x_2 + 2x_3 - mx_4 = 0 \\ x_1 - x_2 - mx_3 - 2(m^2+1)x_4 = 0 \end{cases} \\
 b) \begin{cases} -mx_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\ 2x_1 - mx_2 + 3x_3 + 4x_4 = 0 \\ 2x_1 + 3x_2 - mx_3 + 4x_4 = 0 \\ 2x_1 + 3x_2 + 4x_3 - mx_4 = 0 \end{cases}
 \end{aligned}$$

Đáp số bài tập chương 1

1. a) $AB = \begin{bmatrix} 16 & 8 \\ 0 & -10 \end{bmatrix}, BA = \begin{bmatrix} 0 & 4 & -2 \\ 2 & -8 & 12 \\ 4 & 4 & 14 \end{bmatrix}$
- b) $AB = \begin{bmatrix} 5 & -4 & 4 & -2 \\ 2 & 4 & 1 & 0 \\ 1 & -3 & -1 & 1 \end{bmatrix}, BA$ không tồn tại.
2. a) $\begin{bmatrix} 2 & 3 & 10 \\ 3 & 8 & -1 \end{bmatrix}$ b)-30, c) 1
- d) $\begin{bmatrix} 13 & -20 & 12 \\ 20 & -27 & 28 \end{bmatrix}$ e) $\begin{bmatrix} 3 & -3 & 16 \\ 1 & 6 & 13 \\ 6 & 8 & 12 \end{bmatrix}$
3. a) $X = \begin{bmatrix} a & b \\ \frac{3b}{2} & a + \frac{3b}{2} \end{bmatrix}$ với $a, b \in \mathbb{R}$,
- b) $X = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ với $a, b \in \mathbb{R}$.
4. a) $X = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix}$, b) $X = \begin{bmatrix} -11 & 6 \\ -14 & 7 \end{bmatrix}$, c) Không tồn tại.
5. a) $X = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ với a, b, c thỏa mãn $a^2 + bc = 0$,
- b) Ma trận X là một trong các ma trận dưới đây:
 $X = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, X = -I = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ và $X = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ với $a, b, c : a^2 + bc = 1$.
6. a) $\begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix}^n = \begin{bmatrix} \cos nx & -\sin nx \\ \sin nx & \cos nx \end{bmatrix}$
- b) $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & na \\ 0 & 1 \end{bmatrix}$
- c) $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}^n = \begin{bmatrix} a^n & 0 & 0 \\ 0 & b^n & 0 \\ 0 & 0 & c^n \end{bmatrix}$
- d) $\begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}^{100} = \begin{bmatrix} a^{100} & 100a^{99} & 50.99a^{98} \\ 0 & a^{100} & 100a^{99} \\ 0 & 0 & a^{100} \end{bmatrix}$.
7. a) -3, b) 100, c) 1, d) 1, e) 106, f) 12, g) $\frac{9}{10}$, h) $\frac{1}{35}$.
8. a) abcd, b) abcd, c) xyzuv.
9. Cộng hai hàng sau vào hàng đầu tiên, ta được một hàng gồm các phần tử là $a + b + c$. Theo Định lý Viét áp dụng cho phương trình bậc 3, phần tử này bằng 0, nên định thức bằng 0.
10. a) $x(x-1)(x+3)(x+2)$, b) $8\cos^4 x - 8\cos^2 x + 1$, c) $x^2 z^2$.

11. Lấy hàng $i + 1$ trừ đi hàng i , đưa nhân tử chung của mỗi hàng ra ngoài, sau đó sử dụng công thức truy hồi, ta được điều cần chứng minh.

12. a) $2(x - 2)(x - 3)(x - 4)$, b) $(x_2 - x_1)(x_3 - x_1)(x_4 - x_1)(x_3 - x_2)(x_4 - x_2)(x_4 - x_3)$.

13.

$$a) \frac{1}{2} \begin{bmatrix} -1 & -1 & 1 \\ 3 & 2 & -3 \\ -1 & 0 & 1 \end{bmatrix} \quad b) \begin{bmatrix} 1 & 1 & -4 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad c) \begin{bmatrix} 7 & -4 & -5 \\ 4 & -2 & -3 \\ -1 & 1 & 1 \end{bmatrix}$$

$$d) \frac{1}{4} \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix} \quad e) \frac{1}{2} \begin{bmatrix} 17 & 7 & -3 & 1 \\ 33 & 13 & -7 & 3 \\ 5 & 1 & -1 & 1 \\ -13 & -5 & 3 & -1 \end{bmatrix} \quad f) \frac{1}{2} \begin{bmatrix} 2 & 2 & 4 & -8 \\ 5 & 2 & 14 & -25 \\ -3 & -2 & -8 & 15 \\ 1 & 0 & 2 & -3 \end{bmatrix}$$

14.

$$a) \begin{bmatrix} 2 & -23 \\ 0 & 8 \end{bmatrix} \quad b) \begin{bmatrix} -3 & 2 & 0 \\ -4 & 5 & -2 \\ -5 & 3 & 0 \end{bmatrix} \quad c) \begin{bmatrix} 0 & 5 & 1 & 9 \\ 0 & 3 & 1 & 7 \\ 1 & -2 & 2 & 0 \end{bmatrix}$$

$$15. A^{-1} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$16. X = \begin{bmatrix} \frac{-23}{8} & \frac{2}{39} \\ \frac{13}{2} & \frac{13}{1} \\ \frac{3}{3} & \frac{3}{3} \end{bmatrix}$$

$$17. a) A^5 = \begin{bmatrix} 3197 & -1266 \\ 7385 & -2922 \end{bmatrix} \quad b) A^{2017} = A \quad c) A^{2018} = I.$$

18. a) 2; b) 3; c) 2.

19. $r(A)_{min} = 2$ khi $m = 0$.

- 20.
- Nếu $m = \frac{1}{2}$ thì $r(A) = 3$.
 - Nếu $m \neq \frac{1}{2}$ thì $r(A) = 4$.

$$21. a) x_1 = -3, x_2 = 2, x_3 = 3; b) x_1 = 1, x_2 = 1, x_3 = 2;$$

$$c) x_1 = \frac{3}{2}, x_2 = -4, x_3 = \frac{-1}{2}; d) x_1 = \frac{398}{139}, x_2 = \frac{22}{139}, x_3 = \frac{513}{139}.$$

22. a)

- $m = -3$ hệ phương trình vô số nghiệm

$$\left\{ (x_1 = \frac{-1}{5} + \frac{2}{5}x_3, x_2 = \frac{2}{5} - \frac{14}{5}x_3, x_3) \mid \forall x_3 \in \mathbb{R} \right\}$$

- $m = -1$ hệ phương trình vô nghiệm
- $m \neq -3, -1$ hệ phương trình có nghiệm duy nhất

$$x_1 = 0, x_2 = \frac{-1}{m+1}, x_3 = \frac{2}{m+1}$$

b)

- $m = -1 - \frac{3\sqrt{6}}{2}$ hoặc $-1 + \frac{3\sqrt{6}}{2}$ hệ phương trình vô nghiệm
- $m \neq -1 - \frac{3\sqrt{6}}{2}, \frac{3\sqrt{6}}{2}$ hệ phương trình có nghiệm duy nhất

$$\left\{ x_1 = \frac{3}{2(m-2)}, x_2 = \frac{9}{4(m^2 - m - 2)}, x_3 = \frac{2m^2 - 4m + 25}{4(m^2 - m - 2)} \right\}$$

23. a) $x_1 = \frac{29}{3}, x_2 = -2, x_3 = \frac{17}{3}, x_4 = \frac{-4}{3},$

b) $\left\{ (x_1 = \frac{-x_3 + 2x_4 - 4}{2}, x_2 = \frac{-5x_3 + 10x_4 - 14}{2}, x_3, x_4), \forall x_3, x_4 \in \mathbb{R} \right\}$

c) Hệ phương trình vô nghiệm

d) Hệ phương trình vô nghiệm

24. a)

- $m \neq 1$ hệ có nghiệm duy nhất : $x_1 = 1, x_2 = -1, x_3 = 2$
- $m = 1$ hệ vô số nghiệm $\left\{ (x_1 = \frac{17 - 6x_3}{5}, x_2 = -\frac{3 + x_3}{5}, x_3), \forall x_3 \in \mathbb{R} \right\}$

b)

- $m \neq -5, 1$ hệ phương trình vô nghiệm
- $m = -5$ hoặc $m = 1$ hệ có nghiệm duy nhất

$$x_1 = \frac{5}{2} + \frac{m}{2}, x_2 = -\frac{5}{6} - \frac{5m}{6}, x_3 = \frac{-13}{6} - \frac{7m}{6}$$

25. a) $m = 0$ hoặc $m = 1$; b) $m = -2, m = -3, m = -4$ hoặc $m = 9$

Chương 2

Biến cố ngẫu nhiên và xác suất của nó

2.1. Phép thử và phân loại biến cố

2.1.1. Định nghĩa

★ **Định nghĩa 2.1** Việc thực hiện một nhóm các điều kiện cơ bản để quan sát một hiện tượng nào đó có xảy ra hay không được gọi là thực hiện một phép thử, còn hiện tượng có thể xảy ra trong kết quả của phép thử đó được gọi là biến cố.

● **Ví dụ 2.1** Tung một con súc sắc xuống đất là một phép thử, còn việc lật lên một mặt nào đó là biến cố.

● **Ví dụ 2.2** Bắn một phát súng vào bia. Việc bắn súng là phép thử, còn việc trúng vào một miền nào đó của bia là biến cố.

2.1.2. Phân loại biến cố

Một biến cố chỉ có thể xảy ra khi một phép thử gắn liền với nó được thực hiện. Trong thực tế có thể xảy ra các loại biến cố sau đây:

- **Biến cố chắc chắn:** là biến cố nhất định sẽ xảy ra khi thực hiện một phép thử. Biến cố chắc chắn được ký hiệu là U .
- **Biến cố không thể có:** là biến cố nhất định không xảy ra khi thực hiện một phép thử. Biến cố không thể có được ký hiệu là V .
- **Biến cố ngẫu nhiên:** là biến cố có thể xảy ra hoặc không xảy ra khi thực hiện một phép thử. Các biến cố ngẫu nhiên được ký hiệu là A, B, C, \dots hoặc $A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_n$.

Tất cả các biến cố ta gặp trong thực tế đều thuộc về một trong ba loại biến cố kể trên. Tuy nhiên các biến cố ngẫu nhiên là các biến cố thường gặp hơn cả.

● **Ví dụ 2.3** Tung một con xúc xắc, xét các biến cố sau đây:

U = "Xuất hiện mặt có số chấm nhỏ hơn 7". U là biến cố chắc chắn.

V = "Xuất hiện mặt có 8 chấm". V là biến cố không thể có.

A = "Xuất hiện mặt có số chấm chẵn". A là biến cố ngẫu nhiên.

A_i = "Xuất hiện mặt i chấm", ($i = 1, 2, \dots, 6$). A_i là các biến cố ngẫu nhiên.

2.2. Định nghĩa xác suất

2.2.1. Xác suất của biến cố

★ **Định nghĩa 2.2** Xác suất của một biến cố là một con số đặc trưng cho khả năng khách quan xuất hiện biến cố đó khi thực hiện phép thử. Ký hiệu xác suất của biến cố A là $P(A)$.

Ta chú ý rằng, việc biến cố ngẫu nhiên xảy ra hay không xảy ra trong kết quả của phép thử là điều không thể đoán trước được, xác suất của một biến cố chỉ phản ánh khả năng khách quan xuất hiện biến cố, do những điều kiện của phép thử quy định chứ không tùy thuộc vào ý muốn chủ quan của con người.

2.2.2. Định nghĩa cổ điển về xác suất

a) Ví dụ mở đầu.

Giả sử thực hiện phép thử là tung một con súc sắc cân đối và đồng chất. Xét biến cố $A =$ "Xuất hiện mặt có số chấm chẵn". Ta sẽ xác định xác suất của biến cố A .

Khi tung một con súc sắc cân đối và đồng chất ta thấy có thể có 6 kết cục xảy ra là: xuất hiện các mặt 1 chấm, 2 chấm, ... , 6 chấm. Những kết cục này thoả mãn hai điều kiện: chúng duy nhất, tức là trong kết quả của phép thử chỉ xảy ra một và chỉ một kết cục trong số đó; hơn nữa chúng có khả năng xảy ra như nhau. Các kết cục thoả mãn hai điều kiện trên được gọi các kết cục duy nhất đồng khả năng.

Trong số 6 kết cục duy nhất đồng khả năng đó ta thấy chỉ có 3 kết cục mà nếu kết cục đó xảy ra thì biến cố A sẽ xảy ra, đó là những kết cục được mặt 2 chấm, 4 chấm, 6 chấm. Những kết cục làm cho biến cố xảy ra được gọi là các kết cục thuận lợi cho biến cố.

Như vậy ta thấy khả năng xảy ra của biến cố A là 3 phần 6, tức là 1 phần 2. Đó là cách xác định xác suất của biến cố theo quan điểm cổ điển.

b) Định nghĩa

★ **Định nghĩa 2.3** Xác suất xuất hiện biến cố A trong một phép thử là tỉ số giữa số kết cục thuận lợi cho A và tổng số các kết cục duy nhất đồng khả năng có thể xảy ra khi thực hiện phép thử đó.

Nếu ký hiệu: m là số kết cục thuận lợi cho biến cố A ; n là số kết cục duy nhất đồng khả năng của phép thử, ta có công thức tính xác suất của biến cố A như sau:

$$P(A) = \frac{m}{n}$$

c) Các tính chất của xác suất

- Xác suất của biến cố chắc chắn bằng một: $P(U) = 1$.
- Xác suất của biến cố không thể có bằng không: $P(V) = 0$.
- Xác suất của biến cố ngẫu nhiên là một số nằm trong khoảng giữa không và một:

$$0 < P(A) < 1$$

Như vậy, xác suất của một biến cố bất kỳ luôn thoả mãn điều kiện:

$$0 \leq P(A) \leq 1$$

d) Tính xác suất bằng định nghĩa cổ điển

•**Ví dụ 2.4** Một người khi gọi điện cho bạn quên mất 3 chữ số cuối và chỉ nhớ rằng chúng khác nhau. Tìm xác suất để quay ngẫu nhiên một lần được đúng số cần gọi.

Lời giải.

Gọi A là biến cố "Quay ngẫu nhiên một lần được ngay số cần gọi".

Số kết cục đồng khả năng là tất cả các cách lập nên một bộ 3 số khác nhau từ 10 số tự nhiên đầu tiên. Như vậy:

$$n = A_{10}^3 = 10 \cdot 9 \cdot 8 = 720.$$

Số kết cục thuận lợi cho biến cố A chỉ có một kết cục:

$$m = 1$$

Vì vậy theo định nghĩa cổ điển, xác suất của biến cố A là:

$$P(A) = \frac{m}{n} = \frac{1}{720}.$$

•**Ví dụ 2.5** Trong bình có a quả cầu trắng và b quả cầu đen. Lấy ngẫu nhiên một quả cầu. Tìm xác suất để lấy được một quả cầu trắng.

Lời giải.

Gọi A là biến cố lấy được cầu trắng.

Khi lấy ngẫu nhiên một quả cầu, ta có thể lấy được bất kỳ quả cầu nào trong số $a + b$ quả cầu trong bình, vì vậy số kết cục duy nhất đồng khả năng là:

$$n = a + b$$

Biến cố A sẽ xảy ra khi ta lấy được một trong số a quả cầu trắng, như vậy số kết cục thuận lợi là:

$$m = a$$

Từ đó theo định nghĩa cổ điển về xác suất, ta có:

$$P(A) = \frac{m}{n} = \frac{a}{a + b}$$

•**Ví dụ 2.6** Một hộp có 10 sản phẩm, trong đó có 6 chính phẩm và 4 phế phẩm. Lấy ngẫu nhiên từ hộp đó 3 sản phẩm. Tìm xác suất để:

- Cả 3 sản phẩm lấy ra đều là chính phẩm.
- Trong 3 sản phẩm lấy ra có đúng 2 chính phẩm.

Lời giải.

a) Gọi A là biến cố "Lấy được 3 chính phẩm".

Số kết cục đồng khả năng trong phép thử bằng số cách chọn 3 sản phẩm (phân biệt và không kể thứ tự) từ 10 sản phẩm, như vậy $n = C_{10}^3 = 120$.

Số kết cục thuận lợi cho A xảy ra là số cách chọn được 3 sản phẩm từ 6 chính phẩm, vậy $m = C_6^3 = 20$.

Do đó xác suất của biến cố A là: $P(A) = \frac{m}{n} = \frac{20}{120} = \frac{1}{6}$.

b) Gọi B là biến cố "Trong 3 sản phẩm lấy ra có đúng 2 chính phẩm".

Để biến cố B xảy ra ta phải thực hiện chọn theo 2 bước:

- Chọn 2 chính phẩm trong số 6 chính phẩm, số cách chọn là C_6^2 ;
- Chọn 1 phế phẩm trong số 4 phế phẩm, số cách chọn là C_4^1 .

Số kết cục thuận lợi cho biến cố B là số cách chọn cho biến cố B xảy ra:

$$m = C_6^2 \cdot C_4^1$$

Vậy xác suất của biến cố B là:

$$P(B) = \frac{m}{n} = \frac{C_6^2 \cdot C_4^1}{C_{10}^3} = \frac{1}{2}.$$

•**Ví dụ 2.7** Tung một con súc sắc hai lần. Tính xác suất để trong đó có đúng 1 lần xuất hiện mặt 6 chấm.

Lời giải.

Gọi A là biến cố: "Trong 2 lần tung súc sắc có đúng một lần xuất hiện mặt 6 chấm". Số kết cục đồng khả năng là số cách thiết lập được cặp 2 số lần lượt là số chấm xuất hiện của mỗi lần tung súc sắc. Nó bằng số cách chọn một bộ 2 số (có thể giống nhau và có quan tâm đến thứ tự) từ 6 số $\{1, 2, \dots, 6\}$, vậy:

$$n = \tilde{A}_6^2 = 6^2 = 36.$$

Số kết cục thuận lợi cho biến cố A xảy ra gồm 10 kết cục: 16, 26, 36, 46, 56, 61, 62, 63, 64, 65, do đó: $m = 10$.

Vậy xác suất của biến cố A là: $P(A) = \frac{m}{n} = \frac{10}{36} = \frac{5}{18}$.

2.2.3. Định nghĩa hình học về xác suất

★ **Định nghĩa 2.4** Xét một phép thử có vô hạn kết cục đồng khả năng. Giả sử ta có thể biểu thị tập hợp mọi kết cục này bởi một miền hình học Ω nào đó; và những kết cục thuận lợi cho biến cố A được biểu thị bởi miền con $G \subset \Omega$. Khi đó xác suất của biến cố A được tính bởi công thức:

$$P(A) = \frac{\text{độ đo } (G)}{\text{độ đo } (\Omega)}$$

Tùy theo Ω là đoạn thẳng, miền phẳng, mảnh mặt cong hay khối không gian mà độ đo được hiểu là độ dài, diện tích hay thể tích.

•**Ví dụ 2.8** Đường dây điện thoại ngầm nối hai trạm A, B bỗng nhiên bị đứt. Biết dây điện thoại đồng chất dài 800m chôn trong lòng đất có khả năng đứt tại mọi điểm như nhau, hãy tính xác suất dây đứt cách A không quá 100m.

Lời giải. Vì dây có thể đứt tại một điểm bất kỳ trên đoạn thẳng AB với khả năng như nhau nên có thể biểu thị tập hợp mọi kết cục đồng khả năng của phép thử bởi đoạn AB. Các kết cục thích hợp cho biến cố "dây đứt cách A không quá 100m" được biểu thị bởi đoạn AC. Do đó xác suất cần tìm bằng:

$$P(A) = \frac{\text{độ dài } AC}{\text{độ dài } AB} = \frac{100}{800} = \frac{1}{8}.$$

•**Ví dụ 2.9** Hai người bạn A và B hẹn gặp nhau tại một địa điểm xác định trong vòng từ 0 giờ đến 1 giờ. Mỗi người đến chỗ hẹn vào một thời điểm bất kỳ trong khoảng thời gian trên và chờ 20 phút, nếu không gặp người kia thì bỏ đi. Tính xác suất để họ gặp nhau.

Lời giải.

Gọi x (phút) là lúc đến của A, y (phút) là lúc đến của B.

Mọi kết cục đồng khả năng là một cặp số (x,y) :

$$0 \leq x \leq 60; 0 \leq y \leq 60.$$

Vậy miền kết cục đồng khả năng Ω là hình vuông OACB trên hình vẽ.

Các kết cục thuận lợi cho hai người gặp nhau là những cặp (x,y) sao cho: $|x - y| \leq 20$. Trên hình vẽ, tập hợp này ứng với miền con của hình vuông OACB gồm giữa các đường thẳng $y = x + 20$ và $y = x - 20$.

Vậy xác suất để hai người gặp nhau là:

$$P = \frac{\text{diện tích } (G)}{\text{diện tích } (\Omega)} = \frac{60^2 - 40^2}{60^2} = \frac{5}{9}.$$

2.2.4. Định nghĩa thống kê về xác suất**a) Tần suất xuất hiện của biến cố**

★ **Định nghĩa 2.5** Tần suất xuất hiện biến cố trong n phép thử là tỷ số giữa số phép thử trong đó biến cố xuất hiện và tổng số phép thử được thực hiện.

Như vậy, nếu ký hiệu số phép thử là n , số lần xuất hiện biến cố A là k , tần suất xuất hiện biến cố A là $f(A)$ thì:

$$f(A) = \frac{k}{n}.$$

● **Ví dụ 2.10** Khi kiểm tra ngẫu nhiên 80 sản phẩm do một máy sản xuất, người ta phát hiện ra 3 phế phẩm. Gọi A là biến cố "Xuất hiện phế phẩm". Vậy tần suất phát hiện phế phẩm bằng :

$$f(A) = \frac{3}{80}.$$

● **Ví dụ 2.11** Bắn 50 phát đạn vào bia thấy có 47 phát trúng. Gọi A là biến cố "Bắn trúng bia". Tần suất của việc bắn trúng bia bằng:

$$f(A) = \frac{47}{50}.$$

● **Ví dụ 2.12** Để nghiên cứu khả năng xuất hiện mặt sấp khi tung một đồng xu, người ta tiến hành tung một đồng xu nhiều lần và thu được kết quả sau đây:

Người làm thí nghiệm	Số lần tung (n)	Số lần được mặt sấp (k)	Tần suất $f(A) = \frac{k}{n}$
Buffon	4040	2048	0,5069
Pearson	12000	6019	0,5016
Pearson	24000	12012	0,5005

Ta thấy khi số phép thử tăng lên thì tần suất xuất hiện mặt sấp sẽ dao động ngày càng ít hơn xung quanh giá trị không đổi là 0,5. Điều đó cho phép hy vọng rằng khi số phép thử tăng lên vô hạn, tần suất sẽ hội tụ về giá trị 0,5.

Tính ổn định của tần suất là cơ sở để đưa ra định nghĩa thống kê về xác suất.

b) Định nghĩa thống kê về xác suất

★ **Định nghĩa 2.6** Xác suất xuất hiện biến cố A trong một phép thử là một số p không đổi mà tần suất f xuất hiện biến cố đó trong n phép thử sẽ hội tụ theo xác suất về p khi số phép thử tăng lên vô hạn.

$$f \xrightarrow[n \rightarrow \infty]{\text{Hội tụ theo xác suất}} p \text{ (ký hiệu là } f \xrightarrow{p} p \text{)}.$$

Như vậy về mặt thực tế, với số phép thử n đủ lớn ta có thể lấy $P(A) \approx f(A)$. Ta chú ý rằng ở đây tần suất f hội tụ theo xác suất về p chứ không phải hội tụ theo nghĩa thông thường của giải tích toán học. Điều đó có nghĩa là với mọi ε dương bé tùy ý ta luôn có:

$$\lim_{n \rightarrow \infty} P(|f - p| < \varepsilon) = 1.$$

2.3. Quan hệ giữa các biến cố

2.3.1. Tổng các biến cố

★ **Định nghĩa 2.7** Biến cố C được gọi là tổng của hai biến cố A và B, ký hiệu $C = A + B$, nếu C sẽ xảy ra khi có ít nhất một trong hai biến cố ấy xảy ra.

● **Ví dụ 2.13** Hai người cùng bắn vào một bia. Gọi A là biến cố "Người thứ nhất bắn trúng", B là biến cố "Người thứ hai bắn trúng", C là biến cố "Có người bắn trúng". Vậy $C = A + B$.

● **Ví dụ 2.14** Tung một con súc sắc. Gọi A_i là biến cố "Xuất hiện mặt i chấm" ($i = \overline{1, 6}$), B là biến cố "Xuất hiện ít nhất 5 chấm". Biến cố B sẽ xảy ra khi A_5 xảy ra hoặc A_6 xảy ra. Vậy $B = A_5 + A_6$.

★ **Định nghĩa 2.8** Biến cố A được gọi là tổng của n biến cố A_1, A_2, \dots, A_n , ký hiệu $\sum_{i=1}^n A_i$, nếu A sẽ xảy ra khi có ít nhất một trong n biến cố ấy xảy ra.

2.3.2. Tích các biến cố

★ **Định nghĩa 2.9** Biến cố C được gọi là tích của hai biến cố A và B, ký hiệu $C = A.B$, nếu C sẽ xảy ra khi và chỉ khi cả hai biến cố A và B cùng xảy ra.

● **Ví dụ 2.15** Hai người cùng bắn vào một bia. Gọi A là biến cố "Người thứ nhất bắn trúng", B là biến cố "Người thứ hai bắn trúng", C là biến cố "Cả hai người cùng bắn trúng". Vậy $C = A.B$.

● **Ví dụ 2.16** Một mạch điện gồm hai bóng đèn mắc song song.

Gọi A là biến cố "Bóng thứ nhất bị cháy khi điện quá tải", B là biến cố "Bóng thứ hai bị cháy khi điện quá tải", C là biến cố "Mạch điện bị ngắt khi điện quá tải". Rõ ràng biến cố C chỉ xảy ra khi cả hai biến cố A và B cùng đồng thời xảy ra. Vậy $C = A.B$.

★ **Định nghĩa 2.10** Biến cố A được gọi là tích của n biến cố A_1, A_2, \dots, A_n , ký hiệu $\prod_{i=1}^n A_i$, nếu A sẽ xảy ra khi và chỉ khi cả n biến cố nói trên xảy ra.

2.3.3. Biến cố xung khắc

★ **Định nghĩa 2.11** Hai biến cố A và B được gọi là xung khắc với nhau nếu chúng không thể đồng thời xảy ra trong một phép thử. Trường hợp ngược lại, nếu hai biến cố đó có thể cùng xảy ra trong một phép thử thì được gọi là không xung khắc.

● **Ví dụ 2.17** Hai người cùng bắn vào một bia. Gọi A là biến cố "Người thứ nhất bắn trúng", B là biến cố "Người thứ hai bắn trúng". Hai biến cố A và B là không xung khắc

● **Ví dụ 2.18** Một bình có 3 loại cầu là cầu trắng, cầu xanh và cầu đỏ. Lấy ngẫu nhiên từ bình đó một quả cầu. Gọi A là biến cố "Lấy được cầu trắng", B là biến cố "Lấy được cầu xanh". Các biến cố A và B là xung khắc với nhau.

★ **Định nghĩa 2.12** Nhóm n biến cố $\{A_1, A_2, \dots, A_n\}$ được gọi là xung khắc từng đôi nếu bất kỳ hai biến cố nào trong nhóm này cũng xung khắc với nhau.

Chú ý rằng việc nhận xét tính xung khắc của các biến cố chủ yếu dựa vào trực giác.

● **Ví dụ 2.19** Tung một con súc sắc. Gọi A_i là biến cố "Xuất hiện mặt i chấm" ($i = \overline{1,6}$). Các biến cố $\{A_1, A_2, \dots, A_n\}$ là xung khắc từng đôi.

2.3.4. Nhóm đầy đủ các biến cố

★ **Định nghĩa 2.13** Các biến cố $\{A_1, A_2, \dots, A_n\}$ được gọi là nhóm biến cố đầy đủ nếu trong kết quả của phép thử sẽ xảy ra một và chỉ một trong các biến cố đó. Nói cách khác, các biến cố $\{A_1, A_2, \dots, A_n\}$ sẽ tạo nên một nhóm đầy đủ nếu chúng xung khắc từng đôi với nhau và tổng của chúng là một biến cố chắc chắn.

● **Ví dụ 2.20** Tung một con súc sắc.

Gọi A_i là biến cố "Xuất hiện mặt i chấm" ($i = \overline{1,6}$). Các biến cố $\{A_1, A_2, \dots, A_n\}$ tạo nên nhóm đầy đủ các biến cố.

2.3.5. Biến cố đối lập

★ **Định nghĩa 2.14** Hai biến cố A và B được gọi là đối lập với nhau nếu chúng tạo nên một nhóm đầy đủ các biến cố. Ký hiệu biến cố đối lập của biến cố A là \overline{A} .

● **Ví dụ 2.21** Bắn một phát đạn vào bia.

Gọi A là biến cố "Bắn trúng bia", B là biến cố "Bắn trượt bia". Các biến cố A và B là hai biến cố đối lập nhau: $A = \overline{B}; B = \overline{A}$.

● **Ví dụ 2.22** Một hòm có a chính phẩm và b phế phẩm, lấy ngẫu nhiên n sản phẩm ($n \leq a + b$). Gọi A là biến cố "Lấy được ít nhất một chính phẩm". Khi đó biến cố đối lập của A là biến cố \overline{A} : "Trong n sản phẩm lấy ra không có chính phẩm nào (toàn phế phẩm)".

● **Ví dụ 2.23** Một xí nghiệp có 3 ô tô cùng hoạt động.

Gọi biến cố A_i là "ô tô thứ i bị hỏng", ($i = 1,2,3$). Hãy biểu diễn các biến cố sau theo biến cố A_1, A_2, A_3 .

A = "Chỉ ô tô thứ nhất hỏng"

B = "Có đúng 1 ô tô hỏng"

C = "Có đúng 2 ô tô hỏng"

D = "Có ô tô hỏng"

Lời giải.

$$A = A_1 \cdot \overline{A_2} \cdot \overline{A_3}$$

$$B = A_1 \cdot \overline{A_2} \cdot \overline{A_3} + \overline{A_1} \cdot A_2 \cdot \overline{A_3} + \overline{A_1} \cdot \overline{A_2} \cdot A_3$$

$$C = A_1 \cdot A_2 \cdot \overline{A_3} + \overline{A_1} \cdot A_2 \cdot A_3 + A_2 \cdot \overline{A_1} \cdot A_3$$

$$D = A_1 + A_2 + A_3.$$

2.4. Định lý cộng và nhân xác suất

2.4.1. Định lý cộng xác suất (trường hợp các biến cố xung khắc)

◇ **Định lý 2.1** Xác suất của tổng hai biến cố xung khắc bằng tổng xác suất của các biến cố đó. Như vậy, nếu A và B là hai biến cố xung khắc với nhau thì:

$$P(A + B) = P(A) + P(B)$$

Δ. Ta ký hiệu:

n - số kết cục đồng khả năng có thể xảy ra khi phép thử được thực hiện.

m_1 - số kết cục thuận lợi cho biến cố A .

m_2 - số kết cục thuận lợi cho biến cố B .

Do A và B xung khắc nhau do đó không thể có kết cục thuận lợi cho cả A và B đồng thời xảy ra. Vậy số kết cục thuận lợi cho A hoặc B xảy ra bằng $m_1 + m_2$.

Vì vậy:

$$P(A + B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A) + P(B).$$

□

▽ **Hệ quả 2.1** Xác suất của tổng các biến cố xung khắc từng đôi bằng tổng xác suất của các biến cố đó.

Tức là nếu A_1, A_2, \dots, A_n là các biến cố xung khắc từng đôi thì:

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$

▽ **Hệ quả 2.2** Nếu các biến cố A_1, A_2, \dots, A_n tạo nên một nhóm đầy đủ các biến cố thì tổng các xác suất của chúng bằng 1.

$$\sum_{i=1}^n P(A_i) = 1.$$

▽ **Hệ quả 2.3** Tổng xác suất của các biến cố đối lập nhau bằng 1.

$$P(A) + P(\bar{A}) = 1.$$

● **Ví dụ 2.24** Xác suất để một xạ thủ bắn bia trúng điểm 10 là 0,1; trúng điểm 9 là 0,2; trúng điểm 8 là 0,25 và ít hơn 8 điểm là 0,45. Xạ thủ ấy bắn một phát đạn. Tìm xác suất để xạ thủ được ít nhất 9 điểm.

Lời giải. Gọi A là biến cố "Xạ thủ bắn trúng điểm 10", B là biến cố "Xạ thủ bắn trúng điểm 9", C là biến cố "Xạ thủ được ít nhất 9 điểm". Vậy:

$$C = A + B$$

Vì A và B xung khắc nhau do đó theo định lý cộng xác suất ta có:

$$P(C) = P(A + B) = P(A) + P(B) = 0,1 + 0,2 = 0,3.$$

● **Ví dụ 2.25** Xác suất để sản phẩm sản xuất ra là chính phẩm bằng 0,9. Tìm xác suất sản phẩm sản xuất ra là phế phẩm.

Lời giải. Gọi A là biến cố "Sản phẩm sản xuất ra là phế phẩm", khi đó \bar{A} là biến cố "Sản phẩm sản xuất ra là chính phẩm", A và \bar{A} đối lập nhau, do đó:

$$P(A) = 1 - P(\bar{A}) = 1 - 0,9 = 0,1.$$

•**Ví dụ 2.26** Trong hòm có 12 chi tiết, trong đó có 5 chi tiết hỏng. Lấy ngẫu nhiên 4 chi tiết. Tính xác suất trong số 4 chi tiết lấy ra:

- có không quá một chi tiết hỏng;
- có chi tiết hỏng.

Lời giải.

Gọi A_i là biến cố "Trong 4 chi tiết lấy ra có đúng i chi tiết hỏng", $i = \overline{0,4}$.

a) Gọi A là biến cố "Trong 4 chi tiết lấy ra có không quá một chi tiết hỏng", vậy:

$$A = A_0 + A_1$$

Vì A_0 và A_1 xung khắc với nhau do đó:

$$P(A) = P(A_0) + P(A_1)$$

Dùng định nghĩa cổ điển về xác suất ta tính được:

$$P(A_0) = \frac{C_7^4}{C_{12}^4} = \frac{7}{99}, \quad P(A_1) = \frac{C_5^1 \cdot C_7^3}{C_{12}^4} = \frac{35}{99}$$

Vậy:

$$P(A) = \frac{7}{99} + \frac{35}{99} = \frac{14}{33}.$$

b) Gọi B là biến cố "Trong 4 chi tiết lấy ra có chi tiết hỏng", khi đó biến cố đối lập của B là "Trong 4 chi tiết lấy ra không có chi tiết hỏng" = A_0 . Vậy:

$$P(B) = 1 - P(\bar{B}) = 1 - P(A_0) = 1 - \frac{7}{99} = \frac{92}{99}.$$

2.4.2. Định lý nhân xác suất

a) Biến cố độc lập

★ **Định nghĩa 2.15** Hai biến cố A và B gọi là độc lập với nhau nếu việc xảy ra hay không xảy ra biến cố này không làm thay đổi xác suất xảy ra của biến cố kia và ngược lại.

Trong trường hợp việc biến cố này xảy ra hay không xảy ra làm cho xác suất xảy ra của biến cố kia thay đổi thì hai biến cố đó gọi là phụ thuộc nhau.

•**Ví dụ 2.27** Trong bình có 3 cầu trắng và 2 cầu đen. Lấy ngẫu nhiên 2 quả cầu theo theo phương thức có hoàn lại. Gọi A là biến cố "Lần thứ nhất lấy được cầu trắng", B là biến cố "Lần hai lấy được cầu trắng". Khi đó:

$$P(A) = \frac{3}{5}; \quad P(B) = \frac{3}{5}$$

Ta thấy xác suất lấy được cầu trắng lần thứ hai (biến cố B) không phụ thuộc gì vào kết quả lấy của lần trước (biến cố A), và tương tự, xác suất lấy được cầu trắng lần thứ nhất (biến cố A) không phụ thuộc gì vào kết quả lấy của lần thứ hai (biến cố B). Vậy hai biến cố A và B độc lập với nhau.

• **Ví dụ 2.28** Trong bình có 3 cầu trắng và 2 cầu đen. Lấy ngẫu nhiên 2 quả cầu theo phương thức không hoàn lại. Gọi A là biến cố "Lần thứ nhất lấy được cầu trắng", B là biến cố "Lần hai lấy được cầu trắng". Khi đó xác suất lấy được cầu trắng lần thứ nhất là:

$$P(A) = \frac{3}{5}$$

Tuy nhiên xác suất lấy được cầu trắng lần thứ hai sẽ phụ thuộc vào kết quả lấy của lần trước:

+ Nếu lần thứ nhất lấy được cầu trắng (A xảy ra) thì: $P(B) = \frac{2}{4}$;

+ Nếu lần thứ nhất lấy được cầu đen (A không xảy ra) thì: $P(B) = \frac{3}{4}$;

Vậy A và B phụ thuộc nhau.

Trong thực tế việc nhận xét tính độc lập hay phụ thuộc của các biến cố chủ yếu dựa vào trực giác. Ta cũng nhận thấy rằng tính độc lập của các biến cố có tính tương hỗ. Nếu A và B độc lập với nhau thì A và \bar{B} , \bar{A} và B , \bar{A} và \bar{B} cũng độc lập với nhau.

★ **Định nghĩa 2.16** Các biến cố A_1, A_2, \dots, A_n gọi là độc lập từng đôi với nhau nếu mỗi cặp hai trong n biến cố đó độc lập với nhau. Chẳng hạn ba biến cố A, B, C độc lập từng đôi với nhau nếu A độc lập với B , A độc lập với C , B độc lập với C .

• **Ví dụ 2.29** Tung một đồng xu ba lần, gọi A_i là biến cố "Được mặt sấp ở lần tung thứ i , $i = 1, 2, 3$. Rõ ràng là mỗi cặp hai trong ba biến cố đó độc lập với nhau.

Vậy A_1, A_2, A_3 độc lập từng đôi với nhau.

★ **Định nghĩa 2.17** Các biến cố A_1, A_2, \dots, A_n gọi là độc lập toàn phần với nhau nếu mỗi biến cố độc lập với một tổ hợp bất kỳ của các biến cố còn lại.

Chẳng hạn ba biến cố A, B, C độc lập toàn phần với nhau nếu A độc lập với B , A độc lập với C , A độc lập với tích BC , B độc lập với C , B độc lập với tích AC , C độc lập với tích AB .

Chú ý rằng nếu các biến cố độc lập toàn phần thì suy ra chúng độc lập từng đôi, nhưng điều ngược lại không đúng, nghĩa là nếu các biến cố độc lập từng đôi thì chúng chưa chắc độc lập toàn phần. Ta cũng quy ước rằng khi nói n biến cố độc lập thì có nghĩa là chúng độc lập toàn phần.

• **Ví dụ 2.30** Giả sử trong bình có 4 quả cầu, 1 quả sơn màu đỏ, 1 quả sơn màu xanh, 1 quả màu vàng, 1 quả sơn cả màu đỏ, xanh, vàng.

Lời giải.

Gọi A = "Lấy ngẫu nhiên 1 quả cầu từ bình được quả có màu đỏ";

B = "Lấy ngẫu nhiên 1 quả cầu từ bình được quả có màu xanh";

C = "Lấy ngẫu nhiên 1 quả cầu từ bình được quả có màu vàng";

$$P(A) = P(B) = P(C) = \frac{1}{2}.$$

Nếu biến cố A xảy ra, tức là đã lấy được quả có màu đỏ, thì vì trong 2 quả màu đỏ có 1 quả có màu sơn xanh nên xác suất của biến cố B vẫn là $P(B) = \frac{1}{2}$. Còn nếu A không xảy ra, thì trong

2 quả không có màu đỏ cũng chỉ có 1 quả màu sơn xanh nên $P(B) = \frac{1}{2}$. Vậy các biến cố A và B độc lập với nhau.

Tương tự có thể chứng tỏ A và C , B và C độc lập. Vậy ba biến cố A, B, C độc lập từng đôi.

Nếu biến cố A và B đã xảy ra, tức là quả cầu lấy ra có hai màu đỏ và xanh. Khi đó nó chỉ có thể là quả cầu có cả ba màu, vì vậy xác suất để lấy được quả cầu màu vàng là $P(C) = 1$. Nếu AB không xảy ra thì quả cầu lấy ra không phải quả có ba màu, vì vậy xác suất để lấy được quả vàng là $P(C) = \frac{1}{3}$. Vậy biến cố C không độc lập với AB . Do đó các biến cố A, B, C chỉ độc lập từng đôi chứ không độc lập toàn phần.

b) Xác suất điều kiện

★ **Định nghĩa 2.18** Xác suất của biến cố B được tính với điều kiện biến cố A đã xảy ra được gọi là xác suất có điều kiện của B và được ký hiệu là $P(B/A)$.

● **Ví dụ 2.31** Trong bình có 3 cầu trắng và 2 cầu đen. Lấy ngẫu nhiên 2 quả cầu theo theo phương thức có hoàn lại. Gọi A là biến cố "Lần thứ nhất lấy được cầu trắng", B là biến cố "Lần hai lấy được cầu trắng". Ta đã biết ở 2.27 là A và B độc lập. Ta có:

$$P(B/A) = \frac{3}{5}; \quad P(B/\bar{A}) = \frac{3}{5}.$$

● **Ví dụ 2.32** Trong bình có 3 cầu trắng và 2 cầu đen. Lấy ngẫu nhiên 2 quả cầu theo theo phương thức không hoàn lại. Gọi A là biến cố "Lần thứ nhất lấy được cầu trắng", B là biến cố "Lần hai lấy được cầu trắng". Ta đã biết ở 2.28 là A và B phụ thuộc. Ta có:

$$P(B/A) = \frac{2}{4}; \quad P(B/\bar{A}) = \frac{3}{4}.$$

Để có một nhận xét rằng nếu A và B độc lập thì:

$$P(A/B) = P(A/\bar{B}) = P(A); \quad P(B/A) = P(B/\bar{A}) = P(B).$$

c) Định lý nhân xác suất.

◇ **Định lý 2.2** Xác suất của tích hai biến cố A và B bằng tích xác suất của một trong hai biến cố đó với xác suất điều kiện của biến cố còn lại:

$$P(AB) = P(A).P(B/A) = P(B).P(A/B).$$

Δ. Ta ký hiệu:

n - số kết cục đồng khả năng có thể xảy ra khi phép thử được thực hiện.

m_1 - số kết cục thuận lợi cho biến cố A.

m_2 - số kết cục thuận lợi cho biến cố B.

k - số kết cục thuận lợi cho cả A và B đồng thời xảy ra.

Lúc đó:

$$P(AB) = \frac{k}{n}; \quad P(A) = \frac{m_1}{n}; \quad P(B) = \frac{m_2}{n}.$$

Với điều kiện biến cố A đã xảy ra thì số kết cục duy nhất đồng khả năng của phép thử đối với biến cố B là m_1 , trong đó có k kết cục thuận lợi cho biến cố B xảy ra, do đó:

$$P(B/A) = \frac{k}{m_1}.$$

$$\text{Như vậy: } P(AB) = \frac{k}{n} = \frac{m_1}{n} \cdot \frac{k}{m_1} = P(A).P(B/A).$$

Tương tự khi tính xác suất điều kiện đối với biến cố A, từ đó suy ra điều phải chứng minh. □

▽ **Hệ quả 2.4** Xác suất của tích n biến cố bằng tích xác suất của n biến cố đó, trong đó xác suất của mỗi biến cố tiếp theo đều được tính với điều kiện tất cả các biến cố trước đó đã xảy ra:

$$P(A_1.A_2...A_n) = P(A_1).P(A_2/A_1)...P(A_n/A_1...A_{n-1}).$$

▽ **Hệ quả 2.5** Nếu A và B độc lập thì $P(AB) = P(A).P(B)$.

▽ **Hệ quả 2.6** Xác suất của tích n biến cố độc lập toàn phần bằng tích các xác suất thành phần:

$$P(A_1.A_2...A_n) = P(A_1).P(A_2)...P(A_n).$$

d) Các ví dụ

•**Ví dụ 2.33** Trong bình có 3 cầu trắng và 2 cầu đen. Lấy ngẫu nhiên 2 quả cầu theo phương thức không hoàn lại. Gọi A là biến cố "Lần thứ nhất lấy được cầu trắng", B là biến cố "Lần hai lấy được cầu trắng". Ta đã biết ở 2.32 là A và B phụ thuộc và

$$P(A) = \frac{3}{5}; \quad P(B/A) = \frac{2}{4}P(B/\bar{A}) = \frac{3}{4}.$$

Khi đó xác suất để cả hai lần đều lấy được cầu trắng là:

$$P(AB) = P(A).P(B/A) = \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10}.$$

•**Ví dụ 2.34** Một xí nghiệp có 3 ô tô hoạt động độc lập. Xác suất để trong một ngày các ô tô bị hỏng tương ứng là 0,1; 0,2; 0,15.

- Tính xác suất để cả 3 ô tô đều hỏng.
- Tính xác suất để trong một ngày có đúng 1 ô tô hỏng.
- Tính xác suất để trong một ngày có ít nhất một ô tô hỏng.

Lời giải. Gọi A_i là biến cố "ô tô thứ i bị hỏng trong ngày"; ($i = \overline{1, 3}$).

Theo giả thiết: $P(A_1) = 0,1$; $P(A_2) = 0,2$; $P(A_3) = 0,15$.

Do đó: $P(\bar{A}_1) = 0,9$; $P(\bar{A}_2) = 0,8$; $P(\bar{A}_3) = 0,85$.

a) Gọi A là biến cố "cả 3 ô tô đều hỏng", ta có:

$$A = A_1.A_2.A_3$$

Vì các biến cố A_1, A_2, A_3 độc lập nên theo công thức nhân xác suất ta được:

$$P(A) = P(A_1).P(A_2).P(A_3) = 0,1.0,2.0,15 = 0,003.$$

b) Gọi B là biến cố "trong một ngày có đúng 1 ô tô hỏng", ta có:

$$B = A_1.\bar{A}_2.\bar{A}_3 + \bar{A}_1.A_2.\bar{A}_3 + \bar{A}_1.\bar{A}_2.A_3$$

Vì các biến cố $A_1.\bar{A}_2.\bar{A}_3, \bar{A}_1.A_2.\bar{A}_3$ và $\bar{A}_1.\bar{A}_2.A_3$ là xung khắc từng đôi và mỗi biến cố đó là tích của các biến cố độc lập toàn phần với nhau, do đó:

$$P(B) = A_1.\bar{A}_2.\bar{A}_3 + \bar{A}_1.A_2.\bar{A}_3 + \bar{A}_1.\bar{A}_2.A_3 \quad (2.1)$$

$$= P(A_1).P(\bar{A}_2).P(\bar{A}_3) + P(\bar{A}_1).P(A_2).P(\bar{A}_3) + P(\bar{A}_1).P(\bar{A}_2).P(A_3) \quad (2.2)$$

$$= 0,1.0,8.0,85 + 0,9.0,2.0,85 + 0,9.0,8.0,15 = 0,329. \quad (2.3)$$

$$(2.4)$$

c) Gọi C là biến cố "trong một ngày có ít nhất một ô tô hỏng", ta có:

$$C = A_1 + A_2 + A_3$$

Tuy nhiên các biến cố A_1, A_2, A_3 không xung khắc từng đôi, vì vậy không thể áp dụng định lý cộng xác suất đã học. Ta xét biến cố đối lập của C là biến cố \bar{C} = "tất cả các ô tô đều hoạt động tốt"

$$\bar{C} = \bar{A}_1.\bar{A}_2.\bar{A}_3$$

Theo công thức nhân xác suất:

$$P(\bar{A}_1.\bar{A}_2.\bar{A}_3) = P(\bar{A}_1).P(\bar{A}_2).P(\bar{A}_3) = 0,9.0,8.0,85 = 0,612.$$

Vậy xác suất cần tìm là: $P(C) = 1 - P(\bar{C}) = 1 - 0,612 = 0,388$.

• **Ví dụ 2.35** Tung một đồng xu 6 lần. Tính xác suất để số lần được mặt sấp nhiều hơn số lần được mặt ngửa.

Lời giải. Khi tung đồng xu 6 lần, có các trường hợp sau đây có thể xảy ra:

A - Số mặt sấp xuất hiện nhiều hơn số mặt ngửa;

B - Số mặt ngửa xuất hiện nhiều hơn số mặt sấp;

C - Số mặt sấp và số mặt ngửa xuất hiện bằng nhau;

Các biến cố A, B và C tạo nên một nhóm đầy đủ các biến cố, vì vậy:

$$P(A) + P(B) + P(C) = 1.$$

Vì mặt sấp và ngửa đối xứng nhau nên: $P(A) = P(B)$, từ đó ta có: $P(A) = \frac{1 - P(C)}{2}$.

Ta đi tìm xác suất của biến cố C. Khi tung đồng xu 6 lần số kết cục đồng khả năng là $n = 2^6$; để có đúng 3 lần xuất hiện mặt sấp và 3 lần xuất hiện mặt ngửa thì số kết cục thuận lợi là ra $m = C_6^3$, vậy:

$$P(C) = \frac{m}{n} = \frac{C_6^3}{2^6} = \frac{20}{64} = \frac{5}{16}.$$

Xác suất cần tìm là:

$$P(A) = \frac{1 - \frac{5}{16}}{2} = \frac{11}{32}.$$

2.4.3. Định lý cộng xác suất (trường hợp tổng quát)

◇ **Định lý 2.3** Xác suất của tổng hai biến cố không xung khắc bằng tổng xác suất các biến cố đó trừ đi xác suất của tích các biến cố đó.

$$P(A + B) = P(A) + P(B) - P(AB).$$

Δ. Giả sử:

n là số kết cục đồng khả năng có thể xảy ra trong phép thử;

m_1 là số kết cục thuận lợi cho biến cố A xảy ra;

m_2 là số kết cục thuận lợi cho biến cố B xảy ra;

k là số kết cục thuận lợi cho biến cố tích AB xảy ra;

Khi đó số kết cục thuận lợi cho ít nhất một trong hai biến cố A hoặc B xảy ra là $m_1 + m_2 - k$.

Như vậy: $P(A + B) = \frac{m_1 + m_2 - k}{n} = \frac{m_1}{n} + \frac{m_2}{n} - \frac{k}{n} = P(A) + P(B) - P(AB)$. □

▽ **Hệ quả 2.7** Nếu hai biến cố A, B phụ thuộc thì xác suất của tổng hai biến cố là:

$$P(A + B) = P(A) + P(B) - P(A).P(B/A)$$

▽ **Hệ quả 2.8** Nếu hai biến cố A, B độc lập thì xác suất của tổng hai biến cố là:

$$P(A + B) = P(A) + P(B) - P(A).P(B)$$

▽ **Hệ quả 2.9** Nếu hai biến cố A, B xung khắc thì xác suất của tổng hai biến cố là:

$$P(A + B) = P(A) + P(B)$$

▽ **Hệ quả 2.10** Xác suất của tổng n biến cố không xung khắc được xác định bằng công thức:

$$P\left(\sum_{i=1}^n A_i\right) = \sum_i P(A_i) - \sum_{i < j} P(A_i A_j) + \sum_{i < j < k} P(A_i A_j A_k) - \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n).$$

▽ **Hệ quả 2.11** Xác suất của tích n biến cố được xác định bằng công thức:

$$P\left(\prod_{i=1}^n A_i\right) = \sum_i P(A_i) - \sum_{i < j} P(A_i + A_j) + \sum_{i < j < k} P(A_i + A_j + A_k) - \dots + (-1)^{n-1} P(A_1 + A_2 + \dots + A_n).$$

• **Ví dụ 2.36** Xác suất để động cơ thứ nhất và thứ hai của máy bay trúng đạn lần lượt là 0,1 và 0,2, xác suất để phi công trúng đạn là 0,05; việc trúng đạn của các động cơ và phi công là độc lập với nhau. Tính xác suất máy bay rơi, biết máy bay sẽ rơi nếu cả hai động cơ trúng đạn hoặc phi công trúng đạn.

Lời giải. Gọi các biến cố:

A_1 - động cơ thứ nhất của máy bay trúng đạn;

A_2 - động cơ thứ hai của máy bay trúng đạn;

A_3 - phi công trúng đạn;

B - máy bay rơi. Ta có:

$$B = A_1 A_2 + A_3$$

các biến cố A_1, A_2, A_3 độc lập và không xung khắc, áp dụng định lý cộng xác suất cho hai biến cố $A_1 A_2$ và A_3 , sau đó áp dụng công thức nhân xác suất cho các biến cố độc lập ta được:

$$\begin{aligned} P(B) &= P(A_1 A_2) + P(A_3) - P(A_1 A_2 A_3) = P(A_1) \cdot P(A_2) + P(A_3) - P(A_1) \cdot P(A_2) \cdot P(A_3) \\ &= 0,1 \cdot 0,2 + 0,05 - 0,1 \cdot 0,2 \cdot 0,05 = 0,069. \end{aligned}$$

2.4.4. Định lý liên hệ cộng và nhân xác suất

◇ **Định lý 2.4** Xác suất của tổng n biến cố không xung khắc và độc lập toàn phần với nhau bằng một trừ đi tích xác suất của các biến cố đối lập với các biến cố đó.

$$P\left(\sum_{i=1}^n A_i\right) = 1 - \prod_{i=1}^n P(\bar{A}_i).$$

Δ. Ký hiệu $A = \sum_{i=1}^n A_i$, khi đó: $\bar{A} = \prod_{i=1}^n \bar{A}_i$. do đó:

$$P\left(\sum_{i=1}^n A_i\right) = P(A) = 1 - P(\bar{A}) = 1 - P\left(\prod_{i=1}^n P(\bar{A}_i)\right)$$

Theo giả thiết các biến cố A_i độc lập toàn phần nên các biến cố \bar{A}_i cũng độc lập toàn phần với nhau, áp dụng định lý nhân xác suất:

$$P\left(\prod_{i=1}^n \bar{A}_i\right) = \prod_{i=1}^n P(\bar{A}_i).$$

Vậy

$$P\left(\sum_{i=1}^n A_i\right) = 1 - \prod_{i=1}^n P(\bar{A}_i).$$

▽ **Hệ quả 2.12** Nếu A_1, A_2, \dots, A_n là các biến cố không xung khắc và độc lập toàn phần có cùng xác suất $P(A_1) = P(A_2) = \dots = P(A_n) = p$ thì:

$$P\left(\sum_{i=1}^n A_i\right) = 1 - (1 - p)^n.$$

• **Ví dụ 2.37** Phải tung một con súc sắc bao nhiêu lần để với xác suất không nhỏ hơn 0,5 có thể hy vọng rằng trong đó có ít nhất một lần được mặt 6 chấm?

Lời giải. Giả sử phải tung súc sắc n lần.

Gọi A_i là biến cố "Tung lần thứ i được mặt 6 chấm", ($i = \overline{1, n}$), gọi A là biến cố "Trong n lần tung có ít nhất một lần được mặt 6 chấm". Ta có:

$$A = A_1 + A_2 + \dots + A_n = \sum_{i=1}^n A_i$$

Các biến cố A_i là không xung khắc và độc lập toàn phần với nhau, và có:

$$P(A_1) = P(A_2) = \dots = P(A_n) = \frac{1}{6}$$

do đó:

$$P(A) = 1 - \left(1 - \frac{1}{6}\right)^n = 1 - \left(\frac{5}{6}\right)^n.$$

Theo giả thiết xác suất của A không nhỏ hơn 0,5, ta có bất phương trình sau:

$$1 - \left(\frac{5}{6}\right)^n \geq 0,5 \Leftrightarrow \left(\frac{5}{6}\right)^n \leq 0,5$$

$$\Leftrightarrow n \cdot \ln\left(\frac{5}{6}\right) \leq \ln 0,5$$

$$\Leftrightarrow n \geq \frac{\ln 0,5}{\ln\left(\frac{5}{6}\right)} \approx 3,7.$$

Như vậy $n \geq 4$, tức là phải tung ít nhất 4 lần.

2.5. Công thức Bernoulli

2.5.1. Các phép thử độc lập

★ **Định nghĩa 2.19** Các phép thử được gọi là độc lập với nhau nếu xác suất xảy ra một biến cố nào đó trong từng phép thử sẽ không phụ thuộc vào việc biến cố đó có xảy ra ở các phép thử khác hay không. Khi đó ta cũng gọi các phép thử độc lập như trên là một dãy phép thử độc lập.

Chẳng hạn tung nhiều lần một đồng xu sẽ tạo nên một dãy phép thử độc lập; lấy nhiều lần sản phẩm từ một lô sản phẩm theo phương thức có hoàn lại cũng tạo nên một dãy phép thử độc lập;...

2.5.2. Công thức Bernoulli

Bài toán được gọi là tuân theo lược đồ Bernoulli nếu thoả mãn các giả thiết sau:

- Tiến hành n phép thử độc lập;
- Trong mỗi phép thử chỉ có hai trường hợp: hoặc biến cố A xảy ra, hoặc biến cố A không xảy ra;
- Trong mỗi phép thử xác suất xảy ra của biến cố A đều bằng p ($0 < p < 1$), xác suất không xảy ra biến cố A đều bằng $q = 1 - p$.

◇ **Định lý 2.5** Giả sử bài toán tuân theo lược đồ Bemoulli, khi đó xác suất để trong n phép thử độc lập nói trên, biến cố A xuất hiện đúng k lần, ký hiệu là $P_n(k)$ hoặc p_k , được tính bằng công thức Bemoulli sau đây:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k} \quad (0 \leq k \leq n).$$

Δ. Gọi A_i là biến cố "Xảy ra biến cố A ở phép thử thứ i ", ($i = \overline{1, n}$); gọi B là biến cố "Trong n phép thử biến cố A xảy ra đúng k lần", ta có:

$$B = A_1 A_2 \dots A_k \overline{A_{k+1}} \dots \overline{A_n} + \dots + \overline{A_1} \overline{A_2} \dots \overline{A_{n-k}} A_{n-k+1} \dots A_n.$$

Tổng số các số hạng (mỗi số hạng là một tích n biến cố) chính là số cách chọn k phép thử trong n phép thử để biến cố A xảy ra, do đó số các số hạng của tổng trên là C_n^k . Trong mỗi số hạng trên, n biến cố là độc lập, do đó xác suất mỗi biến cố tích đều bằng $p^k \cdot q^{n-k}$. Vì các tích biến cố đó là xung khắc từng đôi với nhau nên:

$$P_n(k) = P(B) = C_n^k p^k q^{n-k}.$$

□

Nếu gọi B_k là biến cố "Trong n phép thử biến cố A xảy ra đúng k lần" thì dễ thấy $n + 1$ biến cố B_0, B_1, \dots, B_n tạo thành một nhóm đầy đủ, vì vậy:

$$p_0 + p_1 + p_2 + \dots + p_n = 1$$

● **Ví dụ 2.38** Trong phân xưởng có 7 máy hoạt động độc lập, xác suất để trong ca mỗi máy bị hỏng đều bằng 0,1. Tìm xác suất để trong ca đó:

- có đúng 2 máy hỏng;
- có không ít hơn 2 máy hỏng

Lời giải. Nếu coi sự hoạt động của mỗi máy là một phép thử, ta có 7 phép thử độc lập. Trong mỗi phép thử chỉ có hai trường hợp: hoặc máy hỏng, hoặc máy chạy tốt. Xác suất hỏng của mỗi máy đều bằng 0,1. Như vậy bài toán thoả mãn lược đồ Bernoulli.

a) Xác suất để trong ca có đúng 2 máy hỏng được tính bằng công thức Bernoulli như sau:

$$P_7(2) = C_7^2 \cdot (0,1)^2 \cdot (0,9)^5 \approx 0,124.$$

b) Xác suất để trong ca có không ít hơn 2 máy hỏng (tức là có nhiều hơn hoặc bằng 2 máy hỏng) là:

$$\begin{aligned} P_7(2) + P_7(3) + P_7(4) + \dots + P_7(7) &= 1 - P_7(0) - P_7(1) \\ &= 1 - C_7^0 \cdot (0,9)^7 - C_7^1 \cdot (0,1) \cdot (0,9)^6 \approx 1 - 0,8503 = 0,1497. \end{aligned}$$

2.5.3. Số lần xuất hiện chắc nhất

Giả sử bài toán tuân theo lược đồ Bernoulli, khi đó số n_o ứng với xác suất $P_n(n_o)$ lớn nhất trong số các $P_n(k)$, $k = \overline{1, n}$ là số lần xuất hiện chắc nhất của biến cố A . Ta có công thức sau đây để xác định n_o :

$$np - q \leq n_o \leq np + p$$

Δ. Vì n_o là giá trị ứng với xác suất lớn nhất so với các xác suất $P_n(k)$ khác nên ta có các bất đẳng thức sau:

$$\begin{cases} P_n(n_o) \geq P_n(n_o - 1) \\ P_n(n_o) \geq P_n(n_o + 1) \end{cases}$$

Thay các biểu thức của xác suất vào các bất đẳng thức trên theo công thức Bernoulli: $\begin{cases} C_n^{n_o} p^{n_o} q^{n-n_o} \geq C_n^{n_o-} \\ C_n^{n_o} p^{n_o} q^{n-n_o} \geq C_n^{n_o+} \end{cases}$

$$\Leftrightarrow \begin{cases} \frac{n!}{n_o!(n-n_o)!} p^{n_o} q^{n-n_o} \geq \frac{n!}{(n_o-1)!(n-n_o+1)!} p^{n_o-1} q^{n-n_o+1} \\ \frac{n!}{n_o!(n-n_o)!} p^{n_o} q^{n-n_o} \geq \frac{n!}{(n_o+1)!(n-n_o-1)!} p^{n_o+1} q^{n-n_o-1} \end{cases}$$

$$\Leftrightarrow \begin{cases} \frac{p}{n_o q} \geq \frac{q}{n-n_o+1} \\ \frac{p}{n-n_o} \geq \frac{q}{n_o+1} \end{cases} \Leftrightarrow \begin{cases} p(n-n_o+1) \geq q n_o \\ q(n_o+1) \geq p(n-n_o) \end{cases} \Leftrightarrow \begin{cases} n_o \leq np+p \\ n_o \geq np-q \end{cases} \quad \square$$

Ta chú ý rằng, vì n_o là một số nguyên, do đó có thể xảy ra hai trường hợp:

- Nếu $np+p$ là một số nguyên thì $np-q$ cũng là một số nguyên, và $P(np+p) = P(np-q)$, khi đó số lần xuất hiện chắc nhất nhận hai giá trị $n_o = np+p$ và $n_o = np-q$.
- nếu $np+p$ không là số nguyên thì số lần xuất hiện chắc nhất n_o là giá trị nguyên duy nhất nằm trong khoảng hai số $np-q$ và $np+p$.

• **Ví dụ 2.39** Xác suất để một hành khách chậm tàu là 0,02. Tìm số hành khách chậm tàu có khả năng xảy ra nhiều nhất trong một đoàn tàu 835 khách.

Lời giải. Bài toán thoả mãn lược đồ Bernoulli với $n = 835$ phép thử; đối với mỗi phép thử (việc đến tàu của mỗi hành khách) chỉ xảy ra hai trường hợp: chậm tàu (biến cố A) hoặc không chậm tàu; $p = P(A) = 0,02$. Số hành khách chậm tàu có khả năng xảy ra nhiều nhất là số lần xuất hiện chắc nhất n_o , ta có:

$$\begin{aligned} np - q &\leq n_o \leq np + p \\ \Leftrightarrow 835 \cdot 0,02 - 0,98 &\leq n_o \leq 835 \cdot 0,02 + 0,02 \\ \Leftrightarrow 15,72 &\leq n_o \leq 16,72 \end{aligned}$$

Vậy $n_o = 16$, tức là số hành khách chậm tàu có khả năng xảy ra nhiều nhất là 16 người.

2.5.4. Mở rộng công thức Bernoulli

Ta có thể tổng quát hoá lược đồ Bernoulli theo hướng như sau:

- Giả sử tiến hành n phép thử độc lập.
- Trong mỗi phép thử có m kết cục xung khắc A_1, A_2, \dots, A_m hợp thành nhóm đầy đủ.
- Trong mỗi phép thử, xác suất xuất hiện sự kiện A_i bằng p_i (không phụ thuộc chỉ số của phép thử), $p_1 + p_2 + \dots + p_m = 1$.

Ký hiệu $P_n(k_1, k_2, \dots, k_m)$ là xác suất để trong n phép thử đó:

biến cố A_1 xảy ra k_1 lần;

biến cố A_2 xảy ra k_2 lần;

...

biến cố A_m xảy ra k_m lần;

Ta có:

$$P_n(k_1, k_2, \dots, k_m) = C_n^{k_1} \cdot C_{n-k_1}^{k_2} \cdot \dots \cdot C_{n-k_1-k_2-\dots-k_{m-1}}^{k_m} \cdot p_1^{k_1} \cdot p_2^{k_2} \cdot \dots \cdot p_m^{k_m} = \frac{n!}{k_1! k_2! \dots k_m!} \cdot p_1^{k_1} \cdot p_2^{k_2} \cdot \dots \cdot p_m^{k_m}. \quad (2.5)$$

Thật vậy, gọi B là sự kiện "Trong n phép thử, biến cố A_i xảy ra đúng k_i lần". Biến cố B là tổng của các biến cố dạng:

$$\underbrace{A_1 \cdot A_1 \dots A_1}_{k_1 \text{ thừa số}} \cdot \underbrace{A_2 \cdot A_2 \dots A_2}_{k_2 \text{ thừa số}} \dots \underbrace{A_m \cdot A_m \dots A_m}_{k_m \text{ thừa số}}$$

Vì các biến cố A_i có thể được sắp xếp theo các thứ tự khác, sao cho đảm bảo có k_i biến cố A_i nên số các biến cố dạng trên là:

$$C_n^{k_1} \cdot C_{n-k_1}^{k_2} \dots C_{n-k_1-k_2-\dots-k_{m-1}}^{k_m} = \frac{n!}{k_1!(n-k_1)!} \cdot \frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} \dots \frac{(n-k_1-k_2-\dots-k_{m-1})!}{k_m!(n-k_1-k_2-\dots-k_m)!}$$

Vì mỗi biến cố dạng trên có cùng xác suất là $p_1^{k_1} \cdot p_2^{k_2} \dots p_m^{k_m}$ và các biến cố dạng này là xung khắc nên ta có công thức 2.5. \square

•**Ví dụ 2.40** Tại một ga xếp có 10 khách lên tàu gồm 3 toa. Xác suất một khách bất kỳ lên toa I, II và III theo thứ tự là 0,2 ; 0,5 và 0,3. Tính xác suất:

- Có đúng 3 khách lên toa I.
- Có 3 khách lên toa I, 5 khách lên toa II và 2 khách lên toa III;
- Có ít nhất 5 khách lên toa II và 4 khách lên toa III.

Lời giải. Coi sự kiện mỗi người khách lên tàu là thực hiện 1 phép thử, như vậy ta có dãy 10 phép thử độc lập.

a) Bài toán tuân theo lược đồ Bemoulli với $n = 10$; $p = 0,2$. Xác suất cần tìm là:

$$p_3 = P_{10}(3) = C_{10}^3 \cdot (0,2)^3 \cdot (0,8)^7 \approx 0,201.$$

b) Bài toán tuân theo lược đồ Bemoulli mở rộng với $n = 10$; $p_1 = 0,2$; $p_2 = 0,5$; $p_3 = 0,3$. Xác suất cần tìm là:

$$P_{10}(3, 5, 2) = C_{10}^3 \cdot C_7^5 \cdot C_2^2 (0,2)^3 \cdot (0,5)^5 \cdot (0,3)^2 = \frac{10!}{3!5!2!} (0,2)^3 \cdot (0,5)^5 \cdot (0,3)^2 = 0,0567.$$

c) Bài toán tuân theo lược đồ Bernoulli mở rộng tương tự phần b). Có 3 trường hợp có thể xảy ra:

- 6 khách lên toa II và 4 khách lên toa III;
- 5 khách lên toa II và 5 khách lên toa III;
- 1 khách lên toa I; 5 khách lên toa II và 4 khách lên toa III;

Vậy xác suất cần tìm là:

$$P_{10}(0, 6, 4) + P_{10}(0, 5, 5) + P_{10}(1, 5, 4) = C_{10}^6 \cdot (0,5)^6 \cdot (0,3)^4 + C_{10}^5 \cdot (0,5)^5 \cdot (0,3)^5 + C_{10}^1 \cdot C_9^5 \cdot 0,2 \cdot (0,5)^5 \cdot (0,3)^4 \approx 0,1095.$$

2.6. Công thức đầy đủ và công thức Bayes

2.6.1. Công thức xác suất đầy đủ

◇ **Định lý 2.6** Giả sử n biến cố H_1, H_2, \dots, H_n lập thành nhóm đầy đủ các biến cố; biến cố A có thể xảy ra đồng thời với một trong các biến cố H_1, H_2, \dots, H_n . Khi đó xác suất của biến cố A được tính bằng công thức sau đây:

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i).$$

Δ . Vì các biến cố H_1, H_2, \dots, H_n tạo nên một nhóm đầy đủ các biến cố nên:

$$A = (H_1 + H_2 + \dots + H_n) \cdot A = H_1A + H_2A + \dots + H_nA.$$

Vì các biến cố H_1, H_2, \dots, H_n xung khắc từng đôi, do đó các tích H_1A, H_2A, \dots, H_nA cũng xung khắc từng đôi, do đó:

$$P(A) = P(H_1A) + P(H_2A) + \dots + P(H_nA).$$

Áp dụng định lý nhân xác suất với các tích H_iA ta có:

$$P(A) = P(H_1)P(A/H_1) + \dots + P(H_n)P(A/H_n) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i)$$

□

•**Ví dụ 2.41** Có 2 hộp đựng sản phẩm. Hộp thứ nhất có 10 sản phẩm trong đó có 9 chính phẩm. Hộp thứ hai có 20 sản phẩm trong đó có 18 chính phẩm. Từ hộp thứ nhất lấy ngẫu nhiên một sản phẩm bỏ sang hộp thứ hai. Tìm xác suất để lấy ngẫu nhiên một sản phẩm từ hộp thứ hai được chính phẩm.

Lời giải. Gọi A là biến cố "Lấy được chính phẩm từ hộp thứ hai". Biến cố A có thể xảy ra đồng thời với một trong các biến cố sau đây tạo nên một nhóm đầy đủ các biến cố:

H_1 - sản phẩm bỏ từ hộp thứ nhất sang hộp thứ hai là chính phẩm;

H_2 - sản phẩm bỏ từ hộp thứ nhất sang hộp thứ hai là phế phẩm.

Xác suất để lấy được một chính phẩm từ hộp một là: $P(H_1) = \frac{9}{10}$

Xác suất để lấy được một phế phẩm từ hộp một là: $P(H_2) = \frac{1}{10}$

Nếu lấy một chính phẩm từ hộp một sang hộp hai thì xác suất lấy được một chính phẩm từ hộp hai là: $P(A/H_1) = \frac{19}{21}$.

Nếu lấy một phế phẩm từ hộp một sang hộp hai thì xác suất lấy được một chính phẩm từ hộp hai là: $P(A/H_2) = \frac{18}{21}$.

Do đó theo công thức đầy đủ ta có:

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) = \frac{9}{10} \cdot \frac{19}{21} + \frac{1}{10} \cdot \frac{18}{21} = \frac{9}{10}.$$

•**Ví dụ 2.42** Dây chuyền lắp ráp nhận được các chi tiết do hai máy sản xuất. Trung bình máy thứ nhất cung cấp 60% chi tiết, máy thứ hai cung cấp 40% chi tiết. Khoảng 90% chi tiết do máy thứ nhất sản xuất là đạt tiêu chuẩn, còn 85% chi tiết do máy thứ hai sản xuất là đạt tiêu chuẩn. Tính xác suất để lấy ngẫu nhiên từ dây chuyền được một sản phẩm đạt tiêu chuẩn.

Lời giải. Gọi A là biến cố "Chi tiết lấy từ dây chuyền đạt tiêu chuẩn". Biến cố A có thể xảy ra đồng thời với một trong hai biến cố sau đây tạo nên một nhóm đầy đủ các biến cố:

H_1 - Chi tiết do máy thứ nhất sản xuất;

H_2 - Chi tiết do máy thứ hai sản xuất.

Theo điều kiện bài toán:

$$\begin{aligned} P(H_1) &= 0,6; & P(A/H_1) &= 0,9 \\ P(H_2) &= 0,4; & P(A/H_2) &= 0,85 \end{aligned}$$

Do đó xác suất cần tìm là:

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) = 0,6 \cdot 0,9 + 0,4 \cdot 0,85 = 0,88.$$

2.6.2. Công thức Bayes

◇ **Định lý 2.7** Giả sử n biến cố H_1, H_2, \dots, H_n lập thành nhóm đầy đủ các biến cố; biến cố A có thể xảy ra đồng thời với một trong các biến cố H_1, H_2, \dots, H_n . Khi đó:

$$P(H_k/A) = \frac{P(H_k) \cdot P(A/H_k)}{\sum_{i=1}^n P(H_i) \cdot P(A/H_i)}; \quad (k = \overline{1, n})$$

Δ. Theo định lý nhân xác suất ta có: (với $k = \overline{1, n}$)

$$P(AH_k) = P(A) \cdot P(H_k/A) = P(H_k) \cdot P(A/H_k)$$

Từ đó:

$$P(H_k/A) = \frac{P(H_k) \cdot P(A/H_k)}{P(A)}$$

Thay $P(A)$ bằng công thức xác suất đầy đủ ta có:

$$P(H_k/A) = \frac{P(H_k) \cdot P(A/H_k)}{\sum_{i=1}^n P(H_i) \cdot P(A/H_i)}.$$

□

● **Ví dụ 2.43** Có 2 hộp đựng sản phẩm. Hộp thứ nhất có 10 sản phẩm trong đó có 9 chính phẩm. Hộp thứ hai có 20 sản phẩm trong đó có 18 chính phẩm. Từ hộp thứ nhất lấy ngẫu nhiên một sản phẩm bỏ sang hộp thứ hai. Sau đó lấy ngẫu nhiên một sản phẩm từ hộp hai thấy đó là chính phẩm, tính xác suất để sản phẩm lấy từ hộp một bỏ sang hộp hai cũng là một chính phẩm.

Lời giải. Ký hiệu các biến cố tương tự ví dụ 2.41:

A - Lấy được chính phẩm từ hộp thứ hai;

H_1 - sản phẩm bỏ từ hộp thứ nhất sang hộp thứ hai là chính phẩm;

H_2 - sản phẩm bỏ từ hộp thứ nhất sang hộp thứ hai là phế phẩm.

Ta có H_1, H_2 lập thành nhóm đầy đủ các biến cố và:

$$\begin{aligned} P(H_1) &= \frac{9}{10}; & P(A/H_1) &= \frac{19}{21} \\ P(H_2) &= \frac{1}{10}; & P(A/H_2) &= \frac{18}{21} \end{aligned}$$

Ta cần tính $P(H_1/A)$, theo công thức Bayes:

$$\begin{aligned} P(H_1/A) &= \frac{P(H_1) \cdot P(A/H_1)}{P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2)} \\ &= \frac{\frac{9}{10} \cdot \frac{19}{21}}{\frac{9}{10} \cdot \frac{19}{21} + \frac{1}{10} \cdot \frac{18}{21}} = \frac{57/70}{9/10} = \frac{19}{21}. \end{aligned}$$

● **Ví dụ 2.44** Dây chuyền lắp ráp nhận được các chi tiết do hai máy sản xuất. Trung bình máy thứ nhất cung cấp 60% chi tiết, máy thứ hai cung cấp 40% chi tiết. Khoảng 90% chi tiết do máy thứ nhất sản xuất là đạt tiêu chuẩn, còn 85% chi tiết do máy thứ hai sản xuất là đạt tiêu chuẩn. Lấy ngẫu nhiên từ dây chuyền một sản phẩm, thấy nó đạt tiêu chuẩn. Tìm xác suất để sản phẩm đó do máy thứ nhất sản xuất. Hỏi sản phẩm đó có khả năng do máy nào sản xuất ra lớn hơn?

Lời giải. Ký hiệu các biến cố tương tự ví dụ 2.42:

A - Chi tiết lấy từ dây chuyền đạt tiêu chuẩn;

H_1 - Chi tiết do máy thứ nhất sản xuất;

H_2 - Chi tiết do máy thứ hai sản xuất.

Ta có:

$$\begin{aligned} P(H_1) &= 0,6; & P(A/H_1) &= 0,9; \\ P(H_2) &= 0,4; & P(A/H_2) &= 0,85 \end{aligned}$$

Theo công thức xác suất đầy đủ:

$$P(A) = P(H_1).P(A/H_1) + P(H_2).P(A/H_2) = 0,6.0,9 + 0,4.0,85 = 0,88.$$

Xác suất để chính phẩm lấy ra do máy thứ nhất sản xuất là:

$$P(H_1/A) = \frac{P(H_1).P(A/H_1)}{P(A)} = \frac{0,6.0,9}{0,88} \approx 0,614.$$

Xác suất để chính phẩm lấy ra do máy thứ hai sản xuất là:

$$P(H_2/A) = \frac{P(H_2).P(A/H_2)}{P(A)} = \frac{0,4.0,85}{0,88} \approx 0,386.$$

Ta thấy $P(H_1/A) > P(H_2/A)$. Vậy sản phẩm đó có khả năng do máy thứ nhất sản xuất ra lớn hơn. newpage

Bài tập chương 2

Định nghĩa cổ điển về xác suất.

Bài 2.1. Gieo một con súc sắc đối xứng và đồng chất. Tính xác suất để được:

a) Mặt sáu chấm xuất hiện.

b) Mặt có số chấm chẵn xuất hiện.

Đáp số: a) 0,167; b) 0,5.

Bài 2.2. Gieo đồng thời hai con súc sắc. Tìm xác suất để được hai mặt có:

a) Tổng số chấm bằng 7.

b) Tổng số chấm nhỏ hơn 8.

c) ít nhất một mặt 6 chấm.

d) Giả sử gieo n con súc sắc đối xứng và đồng chất, tính xác suất để được tổng số chấm là $n + 1$.

Đáp số: a) 0,167; b) 0,583; c) 0,306.

Bài 2.3. Gieo đồng thời hai đồng xu. Tìm xác suất để được:

a) Hai mặt cùng sấp xuất hiện.

b) Một sấp một ngửa xuất hiện.

c) Có ít nhất một mặt sấp.

Đáp số: a) 0,25; b) 0,5; c) 0,75.

Bài 2.4. Có 100 tấm bìa hình vuông như nhau được đánh số từ 1 đến 100. Lấy ngẫu nhiên một tấm bìa, tìm xác suất:

a) Được tấm bìa không có chữ số 5.

b) Được một tấm bìa có số chia hết cho 2 hoặc cho 5.

Đáp số: a) 0,81; b) 0,6.

Bài 2.5. Một người gọi điện thoại cho bạn nhưng lại quên mất 3 chữ số cuối. Tìm xác suất để người đó quay số một lần được đúng số điện thoại của bạn nếu:

a) Người đó nhớ rằng 3 chữ số cuối khác nhau.

b) Người đó không nhớ gì thêm.

Đáp số: a) 0,001389; b) 0,001.

Bài 2.6. Mỗi vé xổ số có 5 chữ số. Tìm xác suất để một người mua 1 vé được vé:

- a) Có 5 chữ số khác nhau.
- b) Có 5 chữ số đều là lẻ.

Đáp số: a) 0,3024; b) 0,03125.

Bài 2.7. Một nhi đồng tập xếp chữ, em có các chữ N, Ê, H, G, H, N. Tìm xác suất để em đó trong khi sắp ngẫu nhiên được chữ NGHÊNH.

Đáp số: 0,0056.

Bài 2.8. Có 30 tấm bìa giống nhau được đánh số lần lượt từ 1 đến 30, trong đó có 15 tấm bìa có chữ "SU", 11 tấm bìa có chữ "ZU", 3 tấm bìa có chữ "KI", 1 tấm bìa có chữ "SUZUKI". Người chơi sẽ thắng cuộc nếu ghép được (hoặc chọn được) chữ SUZUKI. Tính xác suất người chơi thắng cuộc nếu người đó được chọn:

- a) 2 tấm bìa
- b) 3 tấm bìa
- c) 4 tấm bìa.

Đáp số: a) 0,067; b) 0,222; c) 0,368.

Bài 2.9. Trên giá sách có xếp ngẫu nhiên một tuyển tập của tác giả X gồm 12 cuốn. Tìm xác suất để các tập được xếp theo thứ tự hoặc từ trái sang phải hoặc từ phải sang trái.

Đáp số: 4,2.10⁻⁹.

Bài 2.10. Một lô hàng gồm 6 chính phẩm và 4 phế phẩm được chia ngẫu nhiên thành 2 phần bằng nhau. Tìm xác suất để mỗi phần đều có số chính phẩm như nhau.

Đáp số: 0,476.

Bài 2.11. Trong một hòm đựng 10 chi tiết đạt tiêu chuẩn và 5 chi tiết là phế phẩm, lấy ngẫu nhiên đồng thời 3 chi tiết. Tính xác suất: a) Cả 3 chi tiết lấy ra thuộc loại đạt tiêu chuẩn.

b) Trong số 3 chi tiết lấy ra có 2 chi tiết đạt tiêu chuẩn.

Đáp số: a) 0,264; b) 0,495.

Bài 2.12. Năm người A, B, C, D, E ngồi một cách ngẫu nhiên vào một chiếc ghế dài. Tìm xác suất để: a) C ngồi chính giữa.

b) A và B ngồi ở hai đầu ghế.

c) A và B ngồi cạnh nhau.

Đáp số: a) 0,2; b) 0,1; c) 0,4.

Bài 2.13. Có n người trong đó có m người trùng tên xếp một cách ngẫu nhiên thành một hàng ngang. Tìm xác suất để m người trùng tên đứng cạnh nhau.

Đáp số: $\frac{(n - m + 1)m!}{n!}$

Bài 2.14. Thang máy của một toà nhà 7 tầng xuất phát từ tầng một với 3 khách. Tìm xác suất để:

a) Tất cả cùng ra ở tầng bốn.

b) Tất cả cùng ra ở một tầng.

c) Mỗi người ra ở một tầng khác nhau.

Đáp số: a) 0,0046; b) 0,0278; c) 0,555.

Bài 2.15. Ba nữ nhân viên phục vụ A, B, c thay nhau rửa đĩa chén và giả thiết ba người này đều "khéo léo" như nhau. Trong một tháng có 4 chén bị vỡ. Tìm xác suất:

a) Chị A đánh vỡ 3 chén và chị B đánh vỡ 1 chén.

b) Một trong 3 người đánh vỡ 3 chén.

c) Một trong 3 người đánh vỡ cả 4 chén.

Đáp số: a) 0,049; b) 0,296; c) 0,037.

Bài 2.16. Có 10 khách bước ngẫu nhiên vào một cửa hàng có 3 quầy. Tìm xác suất để có

a) 3 người đến quầy số 1.

b) một quầy có 3 người và một quầy khác có 5 người.

Đáp số: a) 0,260; b) 0,256.

Định nghĩa hình học về xác suất

Bài 2.17. Một người chơi phi tiêu cắm ngẫu nhiên phi tiêu vào một điểm trong bia là hình vuông ABCD tâm O có cạnh 20cm. Tính xác suất để tiêu rơi vào một điểm:

- a) nằm trong hình tròn tâm oO bán kính 5cm
- b) nằm trên đường chéo AC
- c) không rơi vào tâm O.

Đáp số: a) 0,196; b) 0; c) 1.

Bài 2.18. Cho một đoạn thẳng và bẻ gãy ngẫu nhiên thành 3 đoạn. Tính xác suất để 3 đoạn đó tạo thành được một tam giác.

Đáp số: 0,25.

Định nghĩa thống kê về xác suất

Bài 2.19. Có thể xem xác suất sinh được cháu trai là bao nhiêu nếu quan sát 88200 trường hợp sinh trong năm ở một vùng thấy có 45600 cháu trai.

Đáp số: 0,517.

Bài 2.20. Tần suất xuất hiện biến cố viên đạn trúng đích là 0,85. Tìm số viên đạn bắn trúng đích của xạ thủ đó nếu người này bắn 200 viên đạn.

Đáp số: 170.

Mối quan hệ giữa các biến cố

Bài 2.21. Ba người cùng bắn vào một mục tiêu. Gọi A_k là biến cố người thứ k bắn trúng mục tiêu ($k = \overline{1,3}$). Hãy viết bằng ký hiệu các biến cố biểu thị rằng:

- a) Chỉ có người thứ nhất bắn trúng mục tiêu.
- b) Chỉ có một người bắn trúng mục tiêu.
- c) Chỉ có 2 người bắn trúng mục tiêu.
- d) Có người bắn trúng mục tiêu.

Bài 2.22. Kiểm tra theo thứ tự một lô hàng có 10 sản phẩm. Các sản phẩm đều thuộc một trong hai loại: Tốt hoặc Xấu. Ký hiệu \overline{A}_k ($k = \overline{1,10}$) là biến cố chỉ sản phẩm kiểm tra thứ k thuộc loại Xấu. Viết bằng ký hiệu các biến cố sau:

- a) Cả 10 sản phẩm đều Xấu.
- b) Có ít nhất một sản phẩm Xấu.
- c) Cả 6 sản phẩm kiểm tra đầu là Tốt, các sản phẩm còn lại là Xấu.
- d) Các sản phẩm kiểm tra theo thứ tự chẵn là tốt, còn các sản phẩm kiểm tra theo thứ tự lẻ là xấu.

Định lý cộng và nhân xác suất

Bài 2.23. Để được nhập kho, sản phẩm của nhà máy phải qua 3 phòng kiểm tra chất lượng, xác suất phát hiện ra phế phẩm ở các phòng theo thứ tự là 0,8; 0,9 và 0,99. Biết các phòng kiểm tra hoạt động độc lập, tính xác suất phế phẩm được nhập kho. **Đáp số:** 0,0002.

Bài 2.24. Xác suất để khi đo một đại lượng vật lý phạm sai số vượt quá tiêu chuẩn cho phép là 0,4. Thực hiện 3 lần đo độc lập. Tìm xác suất sao cho có đúng một lần đo phạm sai số vượt quá tiêu chuẩn cho phép.

Đáp số: 0,432.

Bài 2.25. Một hộp chứa 3 bi trắng, 7 bi đỏ và 15 bi xanh. Một hộp khác chứa 10 bi trắng, 6 bi đỏ và 9 bi xanh. Lấy ngẫu nhiên từ mỗi hộp một bi. Tìm xác suất để hai bi lấy ra là cùng màu.

Đáp số: 0,3312.

Bài 2.26. Hai người cùng bắn vào một mục tiêu. Khả năng bắn trúng của từng người là 0,8 và 0,9. Tìm xác suất: a) Chỉ có một người bắn trúng mục tiêu.

- b) Có người bắn trúng mục tiêu.
- c) Cả hai người bắn trượt.

Đáp số: a) 0,26; b) 0,98; c) 0,02.

Bài 2.27. Nhà máy có 3 ô tô hoạt động độc lập, xác suất để mỗi ô tô bị hỏng trong ngày làm việc lần lượt là 0,1; 0,2; 0,3. a) Tính xác suất trong ngày có chỉ ô tô thứ nhất hỏng.

- b) Tính xác suất trong ngày có đúng 1 ô tô hỏng.
- c) Tính xác suất trong ngày có ô tô hỏng.

d) Biết trong ngày làm việc có hai ô tô hỏng, tính xác suất 2 ô tô hỏng đó là ô tô thứ nhất và thứ hai.

Đáp số: a) 0,056; b) 0,398; c) 0,496; d) 0,152.

Bài 2.28. Trong hộp có n quả bóng bàn mới. Người ta lấy ra k quả để chơi ($k \leq \frac{n}{2}$) sau đó lại bỏ vào hộp. Tìm xác suất để lần sau lấy ra k quả để chơi thì lấy được toàn bóng mới.

Đáp số: $\frac{C_{n-k}^k}{C_n^k}$

Bài 2.29. Tín hiệu thông tin được phát 3 lần với xác suất thu được của mỗi lần là 0,4.

a) Tìm xác suất nguồn thu nhận được thông tin đó.

b) Nếu muốn xác suất thu được thông tin lên đến 0,9 thì phải phát bao nhiêu lần?

Đáp số: a) 0,784; b) 5 lần.

Bài 2.30. Hai người cùng bắn vào một mục tiêu, khả năng chỉ có một người bắn trúng là 0,38. Tìm xác suất bắn trúng của người thứ nhất, biết rằng khả năng bắn trúng của người thứ hai là 0,8

Đáp số: 0,7.

Bài 2.31. Trong 10 sản phẩm có 2 phế phẩm, lấy ngẫu nhiên 2 sản phẩm. Tính xác suất để cả 2 sản phẩm đều là phế phẩm trong trường hợp:

a) Lấy hoàn lại.

b) Lấy không hoàn lại.

Đáp số: a) 0,04; b) 0,022.

Bài 2.32. Xác suất để bắn một viên đạn trúng đích là 0,8. Hỏi phải bắn bao nhiêu viên đạn để với xác suất nhỏ hơn 0,4 có thể hy vọng rằng không có viên nào trượt?

Đáp số: 5 lần.

Bài 2.33. Phải gieo một con súc sắc trong bao nhiêu lần để với xác suất lớn hơn 0,5 có thể hy vọng có ít nhất một lần được mặt 6 chấm?

Đáp số: 4 lần.

Bài 2.34. Một nồi hơi được lắp 2 van bảo hiểm với xác suất hỏng của các van tương ứng là 0,1 và 0,2. Nồi hơi sẽ hoạt động an toàn khi có van không hỏng. Tìm xác suất để nồi hơi hoạt động:

a) An toàn;

b) Mất an toàn.

Đáp số: a) 0,98; b) 0,02.

Bài 2.35. Bắn liên tiếp vào một mục tiêu cho đến khi viên đạn đầu tiên trúng mục tiêu thì dừng. Tìm xác suất sao cho phải bắn đến viên thứ 6, biết rằng xác suất trúng mục tiêu của mỗi viên đạn là 0,2 và các lần bắn độc lập với nhau.

Đáp số: 0,066.

Bài 2.36. Một tủ kho có chùm chìa khoá gồm 9 chiếc trong đó chỉ có 1 chiếc mở được cửa kho. Anh ta thử ngẫu nhiên từng chìa khoá một, chiếc nào được thử thì không thử lại. Tính xác suất anh ta mở được cửa ở lần thứ 4.

Đáp số: 0,111.

Bài 2.37. Gieo một con súc sắc đối xứng và đồng chất. Gọi A là biến cố được mặt có số chẵn chấm, B là biến cố được mặt có số chấm là bội số của 3.

a) Xét xem hai biến cố này có xung khắc không?

b) Xét xem hai biến cố này có độc lập không?

Đáp số: a) không xung khắc; b) độc lập.

Bài 2.38. Gieo hai con súc sắc đối xứng và đồng chất. Gọi A là biến cố xuất hiện khi tổng số chấm thu được là lẻ; B là biến cố được ít nhất một mặt một chấm. Tính $P(AB)$, $P(A+B)$, $P(\overline{AB})$.

Đáp số: 0,167; 0,639; 0,833.

Bài 2.39. Có hai bóng đèn điện với xác suất hỏng tương ứng là 0,1 và 0,2 và việc chúng hỏng là độc lập với nhau. Tính xác suất để mạch không có điện do bóng hỏng nếu chúng mắc:

a) Nối tiếp.

b) Song song.

Đáp số: a) 0,28; b) 0,02.

Bài 2.40. Có hai lô hàng: Lô 1 có 90 chính phẩm và 10 phế phẩm;

Lô 2 có 80 chính phẩm và 20 phế phẩm;

Lấy ngẫu nhiên từ mỗi lô hàng một sản phẩm, tính xác suất để:

a) Lấy được một chính phẩm.

b) Lấy được ít nhất một chính phẩm.

Đáp số: a) 0,26; b) 0,98.

Bài 2.41. Một lô hàng có 9 sản phẩm. Mỗi lần kiểm tra chất lượng lấy ngẫu nhiên 3 sản phẩm, sau khi kiểm tra xong trả lại vào lô hàng. Tính xác suất để sau 3 lần kiểm tra lô hàng, tất cả các sản phẩm đều được kiểm tra.

Đáp số: 0,0028.

Bài 2.42. Sau khi ra khỏi phòng khách, mỗi người trong N người cùng số giày xỏ ngẫu nhiên vào một đôi giày trong bóng tối. Mỗi người chỉ có thể phân biệt chiếc giày phải với chiếc giày trái, còn không thể phân biệt giày của mình với giày của người khác. Tìm xác suất để:

a) Mỗi người khách đi đúng đôi giày của mình.

b) Mỗi người khách đi đúng 2 chiếc giày của cùng một đôi (không nhất thiết là đôi giày của người đó).

Đáp số: a) $\frac{1}{(N!)^2}$, b) $\frac{1}{N!}$

Bài 2.43. Hai cầu thủ bóng rổ, mỗi người ném bóng 2 lần, xác suất ném trúng đích của mỗi cầu thủ theo thứ tự là 0,6 và 0,7. Tính xác suất:

a) Số lần ném trúng rổ của cầu thủ thứ nhất nhiều hơn số lần ném trúng rổ của cầu thủ thứ hai.

b) Số lần ném trúng rổ của hai người như nhau.

c) Biết số lần ném trúng rổ của hai người như nhau, tính xác suất để mỗi người ném trúng đúng 1 quả.

Đáp số: a) 0,2268; b) 0,3924; c) 0,513.

Bài 2.44. Hai người chơi cờ thỏ thuận với nhau là ai thắng trước 3 ván thì sẽ thắng cuộc. Trận đấu bị gián đoạn khi người thứ nhất còn thiếu 1 ván thắng, người thứ hai còn thiếu 2 ván thắng. Vậy phải phân chia tiền đặt như thế nào là hợp lý nếu xác suất thắng mỗi ván của mỗi người đều bằng 0,5.

Đáp số: 0,75 : 0,25.

Công thức đầy đủ và công thức Bayes.

Bài 2.45. Hai máy cùng sản xuất một loại sản phẩm. Tỷ lệ phế phẩm của máy I là 3%, của máy II là 2%. Từ một kho gồm $\frac{2}{3}$ sản phẩm của máy I và $\frac{1}{3}$ sản phẩm của máy II ta lấy ra một số sản phẩm. Tính xác suất để sản phẩm đó là tốt.

Đáp số: 0,973.

Bài 2.46. Có hai xạ thủ loại I và tám xạ thủ loại II, xác suất bắn trúng đích của các loại xạ thủ theo thứ tự là 0,9 và 0,8.

a) Lấy ngẫu nhiên ra một xạ thủ và xạ thủ đó bắn một viên đạn. Tìm xác suất viên đạn đó trúng đích.

b) Nếu lấy ra hai xạ thủ và mỗi người bắn một viên thì khả năng cả hai viên đều trúng đích là bao nhiêu?

Đáp số: a) 0,82; b) 0,6722.

Bài 2.47. Có hai lô sản phẩm. Lô 1 gồm toàn chính phẩm, lô 2 có tỷ lệ phế phẩm và chính phẩm là $\frac{1}{4}$. Chọn ngẫu nhiên một lô, trong lô này lấy ngẫu nhiên 1 sản phẩm, thấy nó là chính phẩm,

rồi hoàn lại sản phẩm này vào lô. Hỏi rằng nếu lấy ngẫu nhiên (cũng từ lô đã chọn) một sản phẩm thì xác suất để sản phẩm này là phế phẩm bằng bao nhiêu?

Đáp số: 0,089.

Bài 2.48. Bắn ba phát đạn vào một máy bay với xác suất trúng tương ứng là 0,4; 0,5 và 0,7. Nếu

trúng một phát thì xác suất rớt máy bay là 0,2; nếu trúng 2 phát thì xác suất rớt máy bay là 0,6; còn nếu trúng cả 3 phát thì chắc chắn máy bay rơi. a) Tính xác suất máy bay rơi.

b) Giả sử máy bay đã bị bắn rơi, tính xác suất để nó trúng một phát.

Đáp số: a) 0,458; b) 0,1572.

Bài 2.49. Người ta dùng X quang để chẩn đoán bệnh A, tuy nhiên vẫn có sai sót. Nếu sau khi X quang kết luận là có bệnh thì chỉ đúng 99%, còn kết luận là không bị bệnh thì chỉ đúng được 80%. Biết tỷ lệ mắc bệnh A trên thực tế là 20,79%.

a) Tìm xác suất để một người khám bị chẩn đoán có bệnh.

b) Tìm tỷ lệ bệnh nhân không có bệnh bị chẩn đoán nhầm.

Đáp số: a) 0,01; b) 0,000126.

Bài 2.50. Có 2 lô hàng, lô 1 gồm 15 chính phẩm và 4 phế phẩm, lô 2 gồm 18 chính phẩm và 5 phế phẩm. Lấy ra 2 sản phẩm ở lô 1 và 3 sản phẩm ở lô 2. Sau đó chọn ngẫu nhiên 1 sản phẩm từ 5 sản phẩm ấy.

a) Tính xác suất để sản phẩm lấy ra sau cùng là chính phẩm.

b) Biết rằng sản phẩm lấy ra sau là chính phẩm. Hỏi sản phẩm ấy có khả năng thuộc lô nào nhiều hơn?

Đáp số: a) 0,785; b) lô 2.

Bài 2.51. Có hai lô hàng, lô 1 có 7 chính phẩm và 3 phế phẩm; lô 2 có 8 chính phẩm và 2 phế phẩm. Từ lô thứ nhất lấy ra hai sản phẩm, từ lô thứ hai lấy ra 3 sản phẩm; rồi trong số sản phẩm được lấy ra lại lấy tiếp ngẫu nhiên 2 sản phẩm. Tính xác suất để trong hai sản phẩm đó có ít nhất một chính phẩm.

Đáp số: 0,951.

Bài 2.52. Có hai lô sản phẩm, lô 1 có 16 chính phẩm và 4 phế phẩm; lô 2 có 15 chính phẩm và 5 phế phẩm.

a) Từ lô thứ nhất bỏ sang lô thứ hai một sản phẩm, sau đó từ lô thứ hai lấy ra một sản phẩm. Tìm xác suất để lấy được chính phẩm.

b) Từ lô thứ nhất bỏ sang lô thứ hai một sản phẩm, sau đó từ lô thứ hai bỏ sang lô thứ nhất một sản phẩm, sau đó từ lô thứ nhất lấy ra một sản phẩm. Tìm xác suất để lấy được chính phẩm.

Đáp số: a) 0,752; b) 0,798.

Bài 2.53. Xí nghiệp A sản xuất một loại sản phẩm với xác suất hỏng của mỗi sản phẩm bằng 0,05. Ở phân xưởng sản phẩm có thể được một trong ba nhân viên kiểm tra chất lượng với xác suất như nhau. Xác suất phát hiện sản phẩm hỏng của ba người đó lần lượt là 0,9; 0,85; 0,8. Nếu sản phẩm không bị loại ở phân xưởng thì được chuyển đến KCS của nhà máy và ở đó, sản phẩm hỏng sẽ được phát hiện với xác suất 0,95. Tìm xác suất để sản phẩm bị loại.

Đáp số: 0,049625.

Bài 2.54. Tỷ lệ người dân nghiện thuốc lá ở một vùng là 30%. Biết rằng tỷ lệ người bị viêm họng trong số người nghiện thuốc lá là 60%, còn tỷ lệ người bị viêm họng trong số người không hút thuốc lá là 40%.

a) Lấy ngẫu nhiên một người, biết rằng người đó viêm họng, tính xác suất người đó nghiện thuốc.

b) Nếu người đó không bị viêm họng, tính xác suất người đó nghiện thuốc.

Đáp số: a) 0,391; b) 0,222.

Bài 2.55. Một cỗ máy có ba bộ phận 1, 2, 3. Xác suất hỏng của các bộ phận trong thời gian làm việc theo thứ tự là 0,2; 0,4 và 0,3. Cuối ngày làm việc được biết rằng có hai bộ phận bị hỏng, tính xác suất để hai bộ phận hỏng đó là bộ phận 1 và 2.

Đáp số: 0,298.

Bài 2.56. Trong một bệnh viện, tỷ lệ bệnh nhân các tỉnh như sau: tỉnh A - 25%; tỉnh B - 35%; tỉnh C - 40%. Biết rằng tỷ lệ bệnh nhân là kỹ sư của các tỉnh là: tỉnh A - 2%; tỉnh B - 3%; tỉnh C - 3,5%.

a) Chọn ngẫu nhiên một bệnh nhân, tính xác suất để bệnh nhân đó là kỹ sư.

b) Chọn ngẫu nhiên 3 bệnh nhân, tính xác suất để trong đó có đúng 1 bệnh nhân là kỹ sư.

Đáp số: a) 0,0295; b) 0,0833.

Bài 2.57. Một người có ba chỗ ưa thích như nhau để câu cá. Xác suất để câu được cá ở những chỗ đó tương ứng là 0,6; 0,7 và 0,8. Biết rằng ở một chỗ người đó thả câu 3 lần và chỉ câu được một con cá. Tìm xác suất để cá được câu ở chỗ thứ nhất.

Đáp số: 0,503.

Bài 2.58. Máy bay đến oanh kích mục tiêu phải qua 3 tuyến phòng thủ, xác suất để mỗi tuyến phòng thủ tiêu diệt được máy bay đều bằng 0,8.

a) Tìm xác suất để máy bay rơi trước khi đến được mục tiêu.

b) Nếu muốn bảo vệ mục tiêu với xác suất không nhỏ hơn 99,99% thì phải tổ chức bao nhiêu tuyến phòng thủ?

c) Giả sử máy bay bị rơi, tìm xác suất để tuyến 1 bắn rơi.

Đáp số: a) 0,992; b) 6 tuyến; c) 0,806.

Công thức Becnoulli

Bài 2.59. Một cầu thủ nổi tiếng về đá phạt đền, xác suất đá vào gôn là $\frac{4}{5}$. Có người cho rằng cứ "sút" 5 quả thì chắc chắn có 4 quả vào lưới, điều khẳng định đó có đúng không? Tìm xác suất để trong 5 lần sút có đúng 4 lần bóng vào lưới.

Đáp số: không đúng; 0,4096.

Bài 2.60. Một nữ công nhân quản lý 12 máy dệt, xác suất để mỗi máy trong khoảng thời gian t cần đến sự chăm sóc của nữ công nhân là $\frac{1}{3}$. Tìm xác suất để trong khoảng thời gian t:

a) có máy cần đến sự chăm sóc của nữ công nhân.

b) có không nhiều hơn 2 máy cần đến sự chăm sóc của nữ công nhân.

c) Tìm số máy cần đến sự chăm sóc của nữ công nhân có khả năng xảy ra nhiều nhất và xác suất để xảy ra điều đó.

Đáp số: a) 0,992; b) 0,181; c) 4 máy; 0,238.

Bài 2.61. Xác suất để đường kính của một chi tiết bất kỳ rút từ một lô chi tiết là bé hơn, lớn hơn và nằm trong phạm vi cho phép tương ứng là 0,05; 0,10 và 0,85. Từ lô chi tiết đã cho rút ngẫu nhiên 100 chiếc. Tính xác suất để trong số đó có đúng 5 chi tiết mà đường kính bé hơn và 5 chi tiết mà đường kính lớn hơn phạm vi cho phép.

Đáp số: 0,00606.

Bài tập tổng hợp.

Bài 2.62. Hộp 1 có: 8 bi trắng, 10 bi đỏ. Hộp 2 có: 4 bi trắng, 16 bi đỏ.

a) Từ hộp 1 lấy ra 2 bi cho vào hộp 2 trộn đều sau đó từ hộp 2 lấy ra 2 bi. Tính xác suất để 2 bi lấy ra sau là bi đỏ.

b) Từ mỗi hộp lấy ra một bi sau đó trong 2 bi thu được lấy ngẫu nhiên ra một bi và thấy đó là bi trắng. Tính xác suất để viên bi ấy là của hộp 1.

Đáp số: a) 0,598; b) 0,690.

Bài 2.63. Có 2 hộp sản phẩm. Hộp thứ nhất chứa 8 sản phẩm tốt, 2 sản phẩm xấu. Hộp thứ 2 chứa 7 sản phẩm tốt, 5 sản phẩm xấu. Người ta chuyển 1 sản phẩm từ hộp thứ nhất sang hộp thứ 2, sau đó lấy từ hộp thứ 2 ra 2 sản phẩm.

a) Tính xác suất để 2 sản phẩm lấy từ hộp 2 chỉ có 1 sản phẩm tốt.

b) Tính xác suất trong 2 sản phẩm lấy từ hộp 2 có sản phẩm tốt.

Đáp số: a) 0,518; b) 0,859.

Bài 2.64. Sinh viên phải chọn học ít nhất 1 trong 3 môn tự chọn: Toán, Lý, Hoá. Biết có 60% sinh viên học Toán, 40% học Lý, 50% học Hoá, 28% học Toán và Lý, 21% học Lý và Hoá, 20% học Toán và Hoá.

a) Tính tỷ lệ sinh viên học cả 3 môn Toán, Lý, Hoá.

b) Chọn ngẫu nhiên 1 sinh viên trong lớp học Toán, tính xác suất để sinh viên đó học Hoá.

c) Chọn ngẫu nhiên 3 sinh viên, tính xác suất để trong đó có đúng 1 sinh viên chỉ chọn học duy nhất 1 môn Toán.

Đáp số: a) 0,19; b) 0,333; c) 0,441.

Bài 2.65. Trong một kho rượu số rượu loại A và rượu loại B bằng nhau. Người ta chọn ngẫu

nhien một chai rượu trong kho và đưa cho 5 người sành rượu nếm thử, giả sử mỗi người có xác suất đoán trúng là 75%. Có 4 người kết luận chai rượu loại A và 1 người kết luận chai rượu loại B. Tính xác suất để đó là chai rượu loại A.

Đáp số: 0,964.

Bài 2.66. Bia bắn được chia làm 2 vòng, xác suất bắn trúng vòng trong là 0,7 còn trúng vòng ngoài là 0,3. Tìm xác suất sao cho bắn 3 viên đạn thì được ít ra là 29 điểm, biết rằng nếu bắn trúng vòng trong thì được 10 điểm, trúng vòng ngoài thì được 9 điểm.

ĐS' 0,784.

Bài 2.67. Một người bắn 3 viên đạn. Xác suất để cả 3 viên trúng vòng 10 là 0,008, xác suất để 1 viên trúng vòng 8 là 0,15 và xác suất để 1 viên trúng vòng dưới 8 là 0,4. Tính xác suất để xạ thủ đạt ít nhất 28 điểm.

Đáp số: 0,0935.

Bài 2.68. Một nhân viên bán hàng mỗi năm đến bán ở công ty A ba lần. Xác suất lần đầu bán được hàng là 0,8. Nếu lần trước bán được hàng thì xác suất để lần sau bán được hàng là 0,9; còn nếu lần trước không bán được hàng thì xác suất để lần sau bán được hàng chỉ là 0,4. Tìm xác suất để:

- a) Cả ba lần đều bán được hàng;
- b) Có đúng hai lần bán được hàng.

Đáp số: a) 0,648; b) 0,176.

Bài 2.69. Một người bắn 3 viên đạn với xác suất trúng đích của mỗi viên là 0,6. Do yếu tố tâm lý nên: nếu viên thứ nhất trúng thì xác suất trúng của mỗi viên sau là 0,7; còn nếu cả hai viên đầu trúng thì xác suất trúng của viên thứ ba là 0,8; ngoài ra nếu viên thứ nhất trượt thì xác suất để cả hai viên sau trúng là 0,3. Tính xác suất:

- a) Cả 3 viên cùng trúng;
- b) Viên thứ hai và viên thứ ba trúng;
- c) Có ít nhất 1 viên trúng.

Chương 3

Đại lượng ngẫu nhiên và các quy luật phân phối xác suất

3.1. Định nghĩa và phân loại đại lượng ngẫu nhiên

3.1.1. Định nghĩa

★ **Định nghĩa 3.1** Đại lượng ngẫu nhiên (hay còn gọi là biến ngẫu nhiên) là đại lượng mà trong kết quả của phép thử sẽ nhận một và chỉ một trong các giá trị có thể có của nó với một xác suất tương ứng xác định.

Các đại lượng ngẫu nhiên được kí hiệu là X, Y, Z hoặc $X_1, X_2, \dots, X_n; Y_1, \dots, Y_m$ còn các giá trị có thể có của đại lượng ngẫu nhiên được kí hiệu là $x, x_1, x_2, \dots, x_n, y, y_1, \dots, y_m$.

Đại lượng X được gọi là ngẫu nhiên vì trước khi tiến hành phép thử ta chưa thể nói chắc chắn nó nhận giá trị bằng bao nhiêu mà chỉ có thể dự đoán với xác suất nhất định. Việc X nhận một giá trị nào đó ($X = x_1$) hoặc ($X = x_2$), ... ($X = x_n$), ($a < X < b$) là các biến cố ngẫu nhiên.

● **Ví dụ 3.1** Tung một con súc sắc. Gọi X là "Số chấm xuất hiện". X là đại lượng ngẫu nhiên vì trong kết quả của phép thử nó sẽ nhận một trong số các giá trị có thể có là 1, 2, 3, 4, 5, 6 với xác suất tương ứng là $p_k = P(X = k) = 1/6, k = \overline{1, 6}$.

● **Ví dụ 3.2** Bắn ngẫu nhiên một viên đạn vào bia. Gọi X là khoảng cách từ tâm bia đến điểm chạm vủa đạn vào bia. X cũng là một đại lượng ngẫu nhiên.

3.1.2. Phân loại đại lượng ngẫu nhiên

Chúng ta quan tâm đến hai loại đại lượng ngẫu nhiên sau đây:

- Đại lượng ngẫu nhiên được gọi là *rời rạc* nếu các giá trị có thể có của nó lập nên một tập hữu hạn hoặc đếm được.
- Đại lượng ngẫu nhiên được gọi là *liên tục* nếu các giá trị có thể có của nó lấp đầy một khoảng trên trục số.

Như vậy, đại lượng ngẫu nhiên sẽ là rời rạc nếu ta có thể liệt kê được tất cả các giá trị có thể có của nó. Còn đối với đại lượng ngẫu nhiên liên tục ta không thể liệt kê được tất cả các giá trị có thể có của nó.

•**Ví dụ 3.3** Một phân xưởng có 5 máy hoạt động. Gọi X là "Số máy hỏng trong một ca". X là đại lượng ngẫu nhiên rời rạc với các giá trị có thể có của nó là 0, 1, 2, 3, 4, 5.

•**Ví dụ 3.4** Gọi X là "Số người vào mua hàng tại một siêu thị trong một ngày". X là đại lượng ngẫu nhiên rời rạc với các giá trị có thể có của nó lập nên một tập đếm được: 0, 1, 2, 3, 4, 5, ...

•**Ví dụ 3.5** Phép thử là bắn ngẫu nhiên một phát súng vào bia hình tròn bán kính 40 cm. Gọi X là "Khoảng cách từ tâm bia đến điểm chạm của đạn vào bia". X là đại lượng ngẫu nhiên liên tục vì ta không thể kể ra được tất cả các giá trị có thể có của nó, ta chỉ có thể nói rằng tập hợp đó là khoảng $[0; 40]$ cm.

3.2. Quy luật phân phối xác suất của đại lượng ngẫu nhiên

★ **Định nghĩa 3.2** Bất kì một hình thức nào đó cho phép biểu diễn mối quan hệ giữa các giá trị có thể có của đại lượng ngẫu nhiên và các xác suất tương ứng đều được gọi là các quy luật phân phối xác suất của đại lượng ngẫu nhiên ấy.

Ta thường sử dụng ba phương pháp để thiết lập quy luật phân phối xác suất của đại lượng ngẫu nhiên là: bảng phân phối xác suất, hàm phân phối xác suất, hàm mật độ xác suất.

3.2.1. Bảng phân phối xác suất

Bảng phân phối xác suất chỉ dùng để thiết lập quy luật phân phối xác suất của đại lượng ngẫu nhiên rời rạc.

Giả sử đại lượng ngẫu nhiên rời rạc X có thể nhận các giá trị có thể có là x_1, x_2, \dots, x_n với các xác suất tương ứng là p_1, p_2, \dots, p_n . Bảng phân phối xác suất của đại lượng ngẫu nhiên rời rạc X có dạng:

X	x_1	x_2	...	x_i	...	x_n
P	p_1	p_2	...	p_i	...	p_n

Chú ý rằng để tạo nên một quy luật phân phối xác suất thì các xác suất p_i phải thỏa mãn điều kiện:

$$\begin{cases} 0 \leq p_i \leq 1 & \text{với } \forall i = \overline{1, n} \\ \sum_{i=1}^n p_i = 1 \end{cases}$$

Thông thường khi lập bảng phân phối xác suất người ta thường sắp xếp các giá trị có thể có theo thứ tự tăng dần; $x_1 < x_2 < \dots < x_n$.

•**Ví dụ 3.6** Tung một con súc sắc. Gọi X là "Số chấm xuất hiện". Hãy xây dựng quy luật phân phối xác suất của X .

Lời giải. Ta có X là đại lượng ngẫu nhiên rời rạc có thể nhận các giá trị có thể có là 1, 2, 3, 4, 5, 6 với các xác suất tương ứng đều bằng $1/6$. Vì vậy bảng phân phối xác suất của đại lượng ngẫu nhiên rời rạc X có dạng:

Nhận xét thấy $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$.

X	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

•**Ví dụ 3.7** Trong hộp có 10 sản phẩm trong đó có 6 chính phẩm. Lấy ngẫu nhiên 2 sản phẩm. Lập bảng phân phối xác suất của số chính phẩm được lấy ra.

Lời giải. Gọi X là số chính phẩm được lấy ra trong 2 sản phẩm. X là đại lượng ngẫu nhiên rời rạc có thể nhận các giá trị có thể có là 0, 1, 2. Xác suất để cả hai sản phẩm lấy ra đều là phế phẩm (không có chính phẩm) là:

$$P(X = 0) = \frac{C_4^2}{C_{10}^2} = \frac{2}{15}$$

Xác suất để hai sản phẩm lấy ra có đúng 1 chính phẩm là:

$$P(X = 1) = \frac{C_4^1 \cdot C_4^1}{C_{10}^2} = \frac{8}{15}$$

Xác suất để cả hai sản phẩm lấy ra đều là chính phẩm là:

$$P(X = 2) = \frac{C_6^2}{C_{10}^2} = \frac{1}{3}$$

Vì vậy bảng phân phối xác suất của đại lượng ngẫu nhiên rời rạc X có dạng:

X	0	1	2
P	$\frac{2}{15}$	$\frac{8}{15}$	$\frac{1}{3}$

$$\text{Nhận thấy: } \frac{2}{15} + \frac{8}{15} + \frac{1}{3} = 1.$$

•**Ví dụ 3.8** Xác suất để xạ thủ bắn trúng bia là 0,8. Xạ thủ được phát từng viên đạn để bắn đến khi trúng bia. Xây dựng quy luật phân phối xác suất của số viên đạn được phát.

Lời giải. Gọi X là "Số viên đạn xạ thủ được phát". X là đại lượng ngẫu nhiên rời rạc có thể nhận các giá trị có thể có là 1, 2, ..., k , ...

Xác suất để được phát một viên đạn là xác suất để bắn viên đạn đầu tiên trúng bia, do đó:

$$P(X = 1) = 0,8$$

Xác suất để được phát hai viên đạn là xác suất để xảy ra đồng thời hai biến cố: phát thứ nhất bắn trượt và phát thứ hai bắn trúng. Theo định lý nhân xác suất ta có:

$$P(X = 2) = 0,2 \cdot 0,8$$

Biến cố ($X = k$) là tích của k biến cố: $k - 1$ phát đầu bắn trượt và phát thứ k bắn trúng. Theo định lý nhân xác suất ta có:

$$P(X = k) = (0,2)^{k-1} \cdot 0,8$$

Như vậy bảng phân phối xác suất của X có dạng:

X	1	2	3	...	k	...
P	0,8	0,2 \cdot 0,8	$(0,2)^2 \cdot 0,8$...	$(0,2)^{k-1} \cdot 0,8$...

Tổng các xác suất là tổng cấp số nhân lùi vô hạn:

$$\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} (0,2)^{k-1} \cdot 0,8 = \frac{0,8}{1 - 0,2} = 1.$$

3.2.2. Hàm phân phối xác suất

a) Định nghĩa

★ **Định nghĩa 3.3** Hàm phân phối xác suất của đại lượng ngẫu nhiên X , ký hiệu $F(x)$, là xác suất để đại lượng ngẫu nhiên X nhận giá trị nhỏ hơn x , với x là số thực bất kỳ.

$$F(x) = P(X < x)$$

Hàm phân phối xác suất dùng để thiết lập quy luật phân phối xác suất cho cả đại lượng ngẫu nhiên rời rạc lẫn đại lượng ngẫu nhiên liên tục.

● **Ví dụ 3.9** Đại lượng ngẫu nhiên rời rạc X có bảng phân phối xác suất như sau:

X	1	3	4
P	0,1	0,5	0,4

Hãy xây dựng hàm phân phối xác suất của X và vẽ đồ thị.

Lời giải. Nếu $x \leq 1$, biến cố $(X < x)$ là biến cố không thể có, do đó:

$$F(x) = 0.$$

Nếu $1 < x \leq 3$, biến cố $(X < x)$ chỉ xảy ra khi $(X = 1)$, do đó

$$F(x) = P(X = 1) = 0,1.$$

Nếu $3 < x \leq 4$, biến cố $(X < x)$ sẽ xảy ra khi $(X = 1)$ hoặc khi $(X = 3)$, do đó

$$F(x) = P(X = 1) + P(X = 3) = 0,1 + 0,5 = 0,6.$$

Nếu $x > 4$, biến cố $(X < x)$ sẽ xảy ra khi $(X = 1)$, hoặc khi $(X = 3)$, hoặc khi $(X = 4)$ do đó

$$F(x) = P(X = 1) + P(X = 3) + P(X = 4) = 0,1 + 0,5 + 0,4 = 1.$$

Vậy hàm phân phối xác suất của X là:

$$F(x) = \begin{cases} 0 & \text{với } x \leq 1 \\ 0,1 & \text{với } 1 < x \leq 3 \\ 0,6 & \text{với } 3 < x \leq 4 \\ 1 & \text{với } x > 4 \end{cases}$$

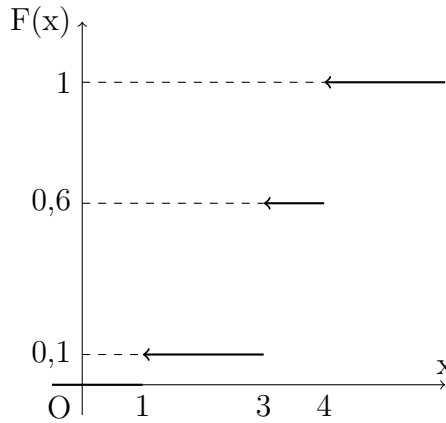
Đồ thị hàm $F(x)$ có dạng trong hình 3.1. **b) Tính chất của hàm phân phối xác suất.**

- Tính chất 1. Hàm phân phối xác suất luôn nhận giá trị trong khoảng $[0; 1]$:

$$0 \leq F(x) \leq 1.$$

Tính chất này suy ra trực tiếp từ định nghĩa hàm phân phối: $F(x) = P(X < x) \in [0; 1]$.

- Tính chất 2. Hàm phân phối xác suất của đại lượng ngẫu nhiên rời rạc là hàm bậc thang với số điểm gián đoạn bằng số giá trị có thể có của X .
Hàm phân phối xác suất của đại lượng ngẫu nhiên liên tục là hàm liên tục và khả vi hầu khắp nơi (số các điểm không khả vi là đếm được).



Hình 3.1:

- Tính chất 3. Hàm phân phối xác suất là hàm không giảm, tức là với $x_1 < x_2$ thì:

$$F(x_1) \leq F(x_2)$$

Δ. Giả sử $x_1 < x_2$, xét biến cố $(X < x_2)$. Biến cố này có thể phân tích thành hai biến cố xung khắc là $(X < x_1)$ và $(x_1 \leq X < x_2)$. Theo định lý cộng xác suất ta có:

$$P(X < x_2) = P(X < x_1) + P(x_1 \leq X < x_2)$$

Do đó:

$$F(x_2) - F(x_1) = P(X < x_2) - P(X < x_1) = P(x_1 \leq X < x_2)$$

Vì $P(x_1 \leq X < x_2)$ là một xác suất nên không âm, vì vậy ta có:

$$F(x_2) - F(x_1) \geq 0 \Leftrightarrow F(x_2) \geq F(x_1)$$

□

- Tính chất 4. Xác suất để đại lượng ngẫu nhiên liên tục X nhận một giá trị xác định bằng 0.

$$P(X = x) = 0$$

- Tính chất 5. Xác suất để đại lượng ngẫu nhiên X nhận giá trị trong khoảng $[a, b)$ bằng hiệu số giá trị của hàm phân phối xác suất ở hai đầu khoảng đó:

$$P(a \leq X < b) = F(b) - F(a)$$

Đặc biệt, nếu X là đại lượng ngẫu nhiên liên tục thì:

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = F(b) - F(a)$$

- Tính chất 6. Nếu đại lượng ngẫu nhiên X chỉ nhận giá trị trong khoảng $[a; b]$ thì:

$$F(x) = 0 \text{ với } \forall x \leq a$$

$$F(x) = 1 \text{ với } \forall x > b$$

Thật vậy, với $x \leq a$, biến cố $(X < x)$ là biến cố không thể có, do đó xác suất của nó bằng 0. Còn với $x > b$, biến cố $(X < x)$ là biến cố chắc chắn, do đó xác suất của nó bằng 1.

Từ tính chất trên ta có biểu thức giới hạn sau:

$$\lim_{x \rightarrow -\infty} F(x) = 0; \lim_{x \rightarrow +\infty} F(x) = 1$$

* **Chú ý:** Điều kiện để hàm số $F(x)$ là hàm phân phối xác suất của đại lượng ngẫu nhiên liên tục:

$$\begin{cases} F(x) \text{ là hàm liên tục} \\ F(x) \text{ là hàm tăng} \\ \lim_{x \rightarrow -\infty} F(x) = 0; \lim_{x \rightarrow +\infty} F(x) = 1. \end{cases}$$

c) **Ý nghĩa của hàm phân phối xác suất.**

Hàm phân phối xác suất phản ánh mức độ tập trung xác suất về phía bên trái một số thực X nào đó. Giá trị của hàm phân phối xác suất tại mỗi điểm X cho biết có bao nhiêu phần của một đơn vị xác suất phân phối trong khoảng $(-\infty; x)$.

• **Ví dụ 3.10** Đại lượng ngẫu nhiên liên tục X có hàm phân phối xác suất như sau:

$$F(x) = \begin{cases} 0 & \text{với } x \leq -1 \\ ax + b & \text{với } -1 < x \leq \frac{1}{3} \\ 1 & \text{với } x > \frac{1}{3} \end{cases}$$

a) Tìm các hệ số a, b .

b) Tìm xác suất để trong kết quả của phép thử, X nhận giá trị trong khoảng $[0, \frac{1}{3})$.

c) Tìm xác suất để trong 3 phép thử độc lập, có đúng một lần X nhận giá trị trong khoảng $(0, 1)$.

Lời giải.

a) Vì X là đại lượng ngẫu nhiên liên tục nên $F(x)$ là hàm liên tục, do đó nó phải liên tục tại $x = -1$ và $\frac{1}{3}$ vì vậy:

$$\begin{cases} -a + b = 0 \\ \frac{1}{3}a + b = 1 \end{cases} \Leftrightarrow \begin{cases} a = \frac{3}{4} \\ b = \frac{3}{4} \end{cases}$$

Khi đó $F(x)$ thoả mãn các điều kiện của hàm phân phối xác suất.

b) Theo tính chất của hàm phân phối, xác suất để X nhận giá trị trong khoảng $[0, \frac{1}{3})$ là:

$$P(0 \leq X < \frac{1}{3}) = F(\frac{1}{3}) - F(0) = 1 - \frac{3}{4} = \frac{1}{4}$$

c) Bài toán tuân theo lược đồ Bernoulli với $n = 3$ phép thử độc lập; xác suất để trong mỗi lần thử X nhận giá trị trong khoảng $(0, 1)$ là:

$$p = P(0 < X < 1) = F(1) - F(0) = 1 - \frac{3}{4} = \frac{1}{4}$$

Vậy xác suất cần tìm bằng:

$$C_3^1 p^1 \cdot q^2 = 3 \cdot \frac{1}{4} \cdot \left(\frac{3}{4}\right)^2 = \frac{27}{64}.$$

3.2.3. Hàm mật độ xác suất

a) **Định nghĩa.**

★ **Định nghĩa 3.4** Hàm mật độ xác suất của đại lượng ngẫu nhiên liên tục X , ký hiệu $f(x)$, là đạo hàm bậc nhất của hàm phân phối xác suất của đại lượng ngẫu nhiên đó.

$$f(x) = [F(x)]'$$

Khái niệm về hàm mật độ xác suất chỉ áp dụng đối với các đại lượng ngẫu nhiên liên tục.

b) Các tính chất của hàm mật độ xác suất.

- Tính chất 1. Hàm mật độ xác suất luôn không âm:

$$f(x) \geq 0, \text{ với } \forall x.$$

Δ. Hàm phân phối xác suất $F(x)$ là hàm không giảm, vì vậy đạo hàm bậc nhất của nó, $f(x) = [F(x)]'$, là hàm không âm. Về mặt hình học, đồ thị của hàm số $f(x)$ không nằm thấp hơn trục Ox . □

- Tính chất 2. Xác suất để đại lượng ngẫu nhiên liên tục X nhận giá trị trong khoảng $[a, b)$ bằng tích phân xác định của hàm mật độ xác suất trong khoảng đó:

$$P(a \leq X < b) = \int_a^b f(x)dx.$$

Δ. Theo tính chất của hàm phân phối ta có:

$$P(a \leq X < b) = F(b) - F(a)$$

Theo công thức Newton-Laibnitz:

$$F(b) - F(a) = \int_a^b F'(x)dx = \int_a^b f(x)dx.$$

Từ đó suy ra điều phải chứng minh. □

Vì X là đại lượng ngẫu nhiên liên tục nên từ tính chất trên ta có:

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = \int_a^b f(x)dx$$

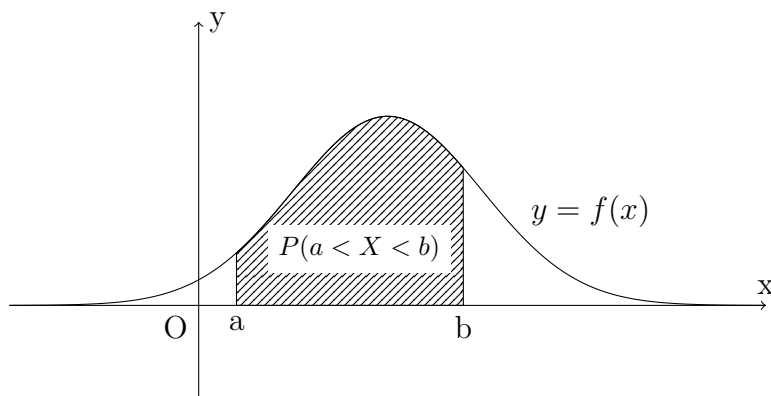
Về mặt hình học, kết quả trên có thể minh họa như sau: xác suất để đại lượng ngẫu nhiên liên tục X nhận giá trị trong khoảng (a, b) bằng diện tích hình thang cong giới hạn bởi trục Ox , đường cong $f(x)$ và các đường thẳng $X = a$; $X = b$.

- Tính chất 3. Hàm phân phối xác suất của đại lượng ngẫu nhiên liên tục X bằng tích phân suy rộng của hàm mật độ xác suất trong khoảng $(-\infty; x)$.

$$F(x) = \int_{-\infty}^x f(t)dt.$$

Δ. Từ tính chất 2, đặt $a = -\infty$, $b = x$, ta có:

$$F(x) = P(-\infty < X < x) = \int_{-\infty}^x f(t)dt$$

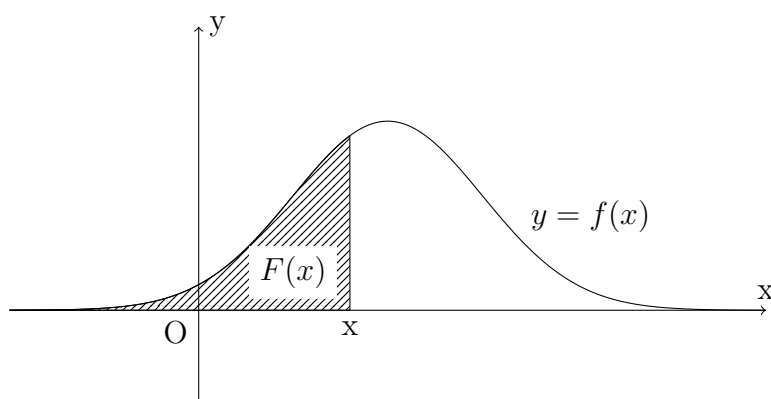


Hình 3.2:

□

Công thức trên cho phép tìm hàm phân phối xác suất của đại lượng ngẫu nhiên liên tục khi đã biết hàm mật độ xác suất của nó.

Về mặt hình học, giá trị của hàm phân phối xác suất $F(x)$ tại điểm a bằng diện tích của hình thang cong suy rộng giới hạn bởi trục Ox , đường cong $f(x)$ và đường thẳng $x = a$.



Hình 3.3:

- Tính chất 4. Tích phân suy rộng trong khoảng $(-\infty, +\infty)$ của hàm mật độ xác suất bằng 1:

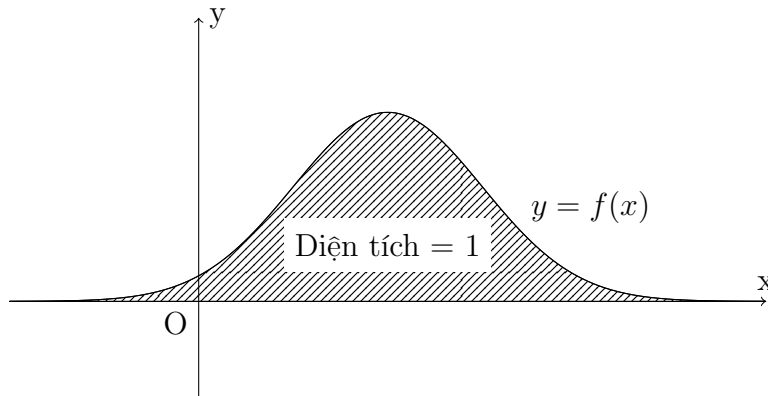
$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Δ. Từ tính chất 2, đặt $a = -\infty, b = +\infty$, và chú ý rằng biến cố $(-\infty < X < +\infty)$ là biến cố chắc chắn, ta có:

$$\int_{-\infty}^{+\infty} f(x)dx = P(-\infty < X < +\infty) = 1.$$

□

Về mặt hình học, điều đó có nghĩa là toàn bộ diện tích hình giới hạn bởi đường cong $f(x)$ và trục Ox bằng 1.



Hình 3.4:

Ta chú ý rằng điều kiện để một hàm số $f(x)$ có thể là hàm mật độ xác suất của một đại lượng ngẫu nhiên liên tục nào đó thì nó phải thỏa mãn hai điều kiện:

$$\begin{cases} f(x) \geq 0; \forall x \\ \int_{-\infty}^{+\infty} f(x)dx = 1 \end{cases}$$

• **Ví dụ 3.11** Đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất như sau:

$$f(x) = \begin{cases} a \cdot \cos x & \text{với } x \in \left[-\frac{\pi}{2}; \frac{\pi}{2}\right] \\ 0 & \text{với } x \notin \left[-\frac{\pi}{2}; \frac{\pi}{2}\right] \end{cases}$$

- Tìm hệ số a ,
- Tìm xác suất để đại lượng ngẫu nhiên X nhận giá trị trong khoảng $(0, \frac{\pi}{4})$,
- Tìm hàm phân phối $F(x)$.

Lời giải.

a) Điều kiện để $f(x)$ là hàm mật độ là:

$$\begin{cases} f(x) \geq 0; \forall x \\ \int_{-\infty}^{+\infty} f(x)dx = 1 \end{cases}$$

Ta có:

$$1 = \int_{-\infty}^{+\infty} f(x)dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} a \cdot \cos x dx = a \cdot \sin x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = 2a.$$

Từ đó $a = \frac{1}{2}$ thỏa mãn các điều kiện trên.

b) Theo tính chất của hàm mật độ xác suất:

$$P(0 < X < \frac{\pi}{4}) = \int_0^{\frac{\pi}{4}} f(x)dx = \int_0^{\frac{\pi}{4}} a \cdot \cos x dx = \frac{1}{2} a \cdot \sin x \Big|_0^{\frac{\pi}{4}} = \frac{\sqrt{2}}{4}.$$

c) Để tìm hàm phân phối xác suất ta sử dụng tính chất:

$$F(x) = \int_{-\infty}^x f(t)dt$$

$$\text{Với } x < -\frac{\pi}{2} : F(x) = \int_{-\infty}^x 0dt = 0.$$

$$\text{Với } -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} :$$

$$F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^{-\frac{\pi}{2}} 0dt + \int_{-\frac{\pi}{2}}^x \frac{1}{2} \cos t dt = \frac{1}{2}(\sin x + 1)$$

$$\text{Với } x > \frac{\pi}{2} :$$

$$F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^{-\frac{\pi}{2}} 0dt + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2} \cos t dt + \int_{\frac{\pi}{2}}^x 0dt = 1$$

Vậy hàm phân phối xác suất của X có dạng:

$$F(x) = \begin{cases} 0 & \text{với } x < -\frac{\pi}{2} \\ \frac{1}{2}(\sin x + 1) & \text{với } -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 1 & \text{với } x > \frac{\pi}{2} \end{cases}$$

• **Ví dụ 3.12** Hàm mật độ xác suất của đại lượng ngẫu nhiên X có dạng:

$$f(x) = \frac{1}{\pi(1+x^2)}; \forall x \in \mathbb{R}$$

Tìm xác suất để khi tiến hành 3 phép thử độc lập có đúng 2 lần X nhận giá trị trong khoảng $(-1, 1)$.

Lời giải. Xác suất để khi tiến hành một phép thử đại lượng ngẫu nhiên X nhận giá trị trong khoảng $(-1, 1)$ là:

$$P(-1 < X < 1) = \int_{-1}^1 f(x)dx = \int_{-1}^1 \frac{1}{\pi(1+x^2)} dx = \frac{1}{\pi} \arctan x \Big|_{-1}^1 = \frac{1}{2}.$$

Khi đó theo công thức Bernoulli, xác suất để trong 3 phép thử độc lập có 2 lần X nhận giá trị trong khoảng $(-1, 1)$ là:

$$p_2 = C_3^2 \left(\frac{1}{2}\right)^2 \cdot \left(1 - \frac{1}{2}\right) = \frac{3}{8}.$$

c) **Ý nghĩa của hàm mật độ xác suất**

Hàm mật độ xác suất của đại lượng ngẫu nhiên X tại điểm X cho biết mức độ tập trung xác suất tại điểm đó.

Thật vậy, theo định nghĩa của hàm mật độ xác suất ta có:

$$f(x) = F'(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{P(x \leq X < x + \Delta x)}{\Delta x}.$$

Như vậy hàm mật độ xác suất tại điểm X chính là giới hạn của mật độ xác suất trung bình trên đoạn $(x; x + \Delta x)$ khi $\Delta x \rightarrow 0$.

3.3. Các tham số đặc trưng của đại lượng ngẫu nhiên

3.3.1. Kỳ vọng toán

a) Định nghĩa

★ **Định nghĩa 3.5** Kỳ vọng toán của đại lượng ngẫu nhiên X , ký hiệu $M(X)$ hoặc $E(X)$, là một số được xác định như sau:

- Nếu X là đại lượng ngẫu nhiên rời rạc có bảng phân phối xác suất: thì kỳ vọng toán $M(X)$ được xác định bằng biểu thức:

X	x_1	x_2	...	x_i	...	x_n
P	p_1	p_2	...	p_i	...	p_n

$$M(X) = \sum_{i=1}^n x_i p_i.$$

- Nếu X là đại lượng ngẫu nhiên liên tục với hàm mật độ xác suất $f(x)$ thì kỳ vọng toán $M(X)$ được xác định bằng biểu thức:

$$M(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

• **Ví dụ 3.13** Cho X là đại lượng ngẫu nhiên rời rạc có bảng phân phối xác suất: Khi đó kỳ vọng

X	1	3	4
P	0,1	0,5	0,4

toán của đại lượng ngẫu nhiên X bằng:

$$M(X) = 1.0,1 + 3.0,5 + 4.0,4 = 3,2.$$

• **Ví dụ 3.14** Cho X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất như sau:

$$f(x) = \begin{cases} \frac{3}{4}(x^2 + 2x) & \text{với } x \in (0, 1) \\ 0 & \text{với } x \notin (0, 1) \end{cases}$$

Khi đó kỳ vọng toán của đại lượng ngẫu nhiên X bằng:

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^0 x \cdot 0 dx + \int_0^1 \frac{3}{4} x \cdot (x^2 + 2x) dx + \int_1^{+\infty} x \cdot 0 dx \\ &= \frac{3}{4} \int_0^1 (x^3 + 2x^2) dx = \frac{3}{4} \left(\frac{x^4}{4} + \frac{2x^3}{3} \right) \Big|_0^1 = \frac{11}{16}. \end{aligned}$$

b) Các tính chất của kỳ vọng toán

Các chứng minh trong phần này được thực hiện với đại lượng ngẫu nhiên rời rạc, nếu X là đại lượng ngẫu nhiên liên tục ta cũng có các chứng minh tương tự.

- Tính chất 1. Kỳ vọng toán của hằng số bằng chính hằng số đó. Như vậy nếu C là hằng số thì: $M(C) = C$.

Δ. Thật vậy, có thể coi C là đại lượng ngẫu nhiên rời rạc đặc biệt, với một giá trị có thể có là C và xác suất tương ứng bằng 1. Khi đó theo định nghĩa kỳ vọng toán ta có:

$$M(C) = 1.C = C$$

□

- Tính chất 2. Kỳ vọng toán của tích giữa một hằng số và đại lượng ngẫu nhiên bằng tích giữa hằng số đó và kỳ vọng toán của đại lượng ngẫu nhiên ấy.

$$M(CX) = C.M(X).$$

Δ. Thật vậy, giả sử X là đại lượng ngẫu nhiên rời rạc với bảng phân phối xác suất:

X	x_1	x_2	...	x_i	...	x_n
P	p_1	p_2	...	p_i	...	p_n

Khi đó đại lượng ngẫu nhiên CX có bảng phân phối xác suất như sau:

CX	Cx_1	Cx_2	...	Cx_i	...	Cx_n
P	p_1	p_2	...	p_i	...	p_n

Theo định nghĩa của kỳ vọng toán ta có:

$$M(CX) = Cx_1p_1 + \dots + Cx_np_n = C(x_1p_1 + \dots + x_np_n) = C.M(X)$$

□

- Tính chất 3. Kỳ vọng toán của tổng hai đại lượng ngẫu nhiên bằng tổng các kỳ vọng toán thành phần.

$$M(X + Y) = M(X) + M(Y).$$

Δ. Thật vậy, giả sử các đại lượng ngẫu nhiên rời rạc X và Y có các bảng phân phối xác suất như sau:

X	x_1	x_2	...	x_i	...	x_n
P	p_1	p_2	...	p_i	...	p_n

Y	y_1	y_2	...	y_i	...	y_m
P	q_1	q_2	...	q_i	...	q_m

Khi đó quy luật phân phối xác suất của tổng $X + Y$ như sau:

X+Y	$x_1 + y_1$	$x_1 + y_2$...	$x_i + y_j$...	$x_n + y_m$
P	p_{11}	p_{12}	...	p_{ij}	...	p_{nm}

trong đó p_{ij} là xác suất để tổng $X + Y$ nhận giá trị $x_i + y_j$.

Theo định nghĩa kỳ vọng toán ta có:

$$M(X + Y) = \sum_{i=1}^n \sum_{j=1}^m (x_i + y_j)p_{ij} = \sum_{i=1}^n \sum_{j=1}^m x_i p_{ij} + \sum_{i=1}^n \sum_{j=1}^m y_j p_{ij}$$

$$= \sum_{i=1}^n x_i \left(\sum_{j=1}^m p_{ij} \right) + \sum_{j=1}^m y_j \left(\sum_{i=1}^n p_{ij} \right) \quad (3.1)$$

Ta sẽ chứng minh rằng $\sum_{j=1}^m p_{ij} = p_i$. Thật vậy, biến cố $X = x_i$ sẽ xảy ra khi tổng $X + Y$ nhận các giá trị $x_i + y_1, x_i + y_2, \dots, x_i + y_m$. Do đó theo định lý cộng xác suất ta có:

$$\begin{aligned} p_i &= P(X = x_i) = P(X + Y = x_i + y_1) + P(X + Y = x_i + y_2) + \dots + P(X + Y = x_i + y_m) \\ &= p_{i1} + p_{i2} + \dots + p_{im} = \sum_{j=1}^m p_{ij}. \end{aligned}$$

Tương tự như vậy ta cũng có $\sum_{i=1}^n p_{ij} = q_j$.

Từ đó thay vào (3.1) ta có:

$$M(X + Y) = \sum_{i=1}^n x_i p_i + \sum_{j=1}^m y_j q_j = M(X) + M(Y).$$

□

Bằng phương pháp chứng minh quy nạp, ta cũng có kết quả sau đây:

Kỳ vọng toán của tổng n đại lượng ngẫu nhiên bằng tổng các kỳ vọng toán thành phần:

$$M(X_1 + X_2 + \dots + X_n) = M(X_1) + M(X_2) + \dots + M(X_n).$$

- Tính chất 4. Kỳ vọng toán của tích hai đại lượng ngẫu nhiên độc lập bằng tích các kỳ vọng toán thành phần:

$$M(X.Y) = M(X).M(Y).$$

Ở đó hai đại lượng ngẫu nhiên được gọi là độc lập với nhau nếu quy luật phân phối xác suất của đại lượng ngẫu nhiên này không phụ thuộc vào việc đại lượng ngẫu nhiên kia nhận giá trị bằng bao nhiêu. Tương tự, n đại lượng ngẫu nhiên được gọi là độc lập với nhau nếu các quy luật phân phối xác suất của một số bất kỳ các đại lượng ngẫu nhiên nào đó không phụ thuộc vào việc các đại lượng ngẫu nhiên kia nhận giá trị bằng bao nhiêu.

Δ. Giả sử các đại lượng ngẫu nhiên X và Y độc lập, có bảng phân phối xác suất như sau:

X	x_1	x_2	...	x_i	...	x_n
P	p_1	p_2	...	p_i	...	p_n

Y	y_1	y_2	...	y_i	...	y_m
P	q_1	q_2	...	q_i	...	q_m

Khi đó quy luật phân phối xác suất của tổng $X.Y$ như sau:

X.Y	$x_1.y_1$	$x_1.y_2$...	$x_i.y_j$...	$x_n.y_m$
P	p_1q_1	p_1q_2	...	p_iq_j	...	p_nq_m

Theo công thức tính kỳ vọng toán:

$$M(XY) = \sum_{i=1}^n \sum_{j=1}^m x_i y_j p_i q_j = \sum_{i=1}^n x_i p_i \sum_{j=1}^m y_j q_j = M(X).M(Y).$$

□

Bằng phương pháp chứng minh quy nạp, ta cũng có kết quả sau đây:

Kỳ vọng toán của tích n đại lượng ngẫu nhiên độc lập bằng tích các kỳ vọng toán thành phần:

$$M(X_1.X_2....X_n) = M(X_1).M(X_2)...M(X_n).$$

- Tính chất 5. Cho X là đại lượng ngẫu nhiên và $g(x)$ là hàm số liên tục, khi đó đại lượng ngẫu nhiên $g(X)$ có kỳ vọng toán $M[g(X)]$ được xác định như sau:
+) Nếu X là đại lượng ngẫu nhiên rời rạc có phân phối xác suất $p_i = P(X = x_i)$; $i = \overline{1, n}$:

$$M[g(X)] = \sum_{i=1}^n g(x_i)p_i.$$

+) Nếu X là đại lượng ngẫu nhiên liên tục có hàm mật độ $f(x)$:

$$M[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx.$$

c) Ý nghĩa của kỳ vọng toán

Kỳ vọng toán của đại lượng ngẫu nhiên X gần bằng trung bình số học của các giá trị quan sát của đại lượng ngẫu nhiên. Nó phản ánh giá trị trung tâm của phân phối xác suất của đại lượng ngẫu nhiên.

Thật vậy, giả sử đối với đại lượng ngẫu nhiên X tiến hành n phép thử độc lập trong đó có n_1 lần X nhận giá trị x_1 , n_2 lần X nhận giá trị x_2 , ..., n_k lần X nhận giá trị x_k . Giá trị trung bình của X trong n phép thử này là:

$$\bar{X} = \frac{x_1n_1 + x_2n_2 + \dots + x_kn_k}{n} = x_1\frac{n_1}{n} + x_2\frac{n_2}{n} + \dots + x_k\frac{n_k}{n}$$

ở đó $\frac{n_1}{n}, \dots, \frac{n_k}{n}$ chính là tần suất xuất hiện các giá trị x_1, x_2, \dots, x_k trong n phép thử trên, do đó:

$$\bar{X} = x_1f_1 + x_2f_2 + \dots + x_kf_k.$$

Theo định nghĩa thống kê về xác suất, khi $n \rightarrow \infty$ các tần suất trên sẽ hội tụ về xác suất tương ứng, do đó với n đủ lớn ta có:

$$\bar{X} \approx x_1p_1 + x_2p_2 + \dots + x_kp_k = M(X).$$

• **Ví dụ 3.15** Tung con súc sắc n lần. Tìm kỳ vọng toán của tổng số chấm thu được.

Lời giải. Gọi X_i , ($i = \overline{1, n}$) là số chấm thu được ở lần tung thứ i và X là tổng số chấm thu được sau n lần tung. Ta có:

$$X = X_1 + X_2 + \dots + X_n.$$

Theo tính chất của kỳ vọng toán:

$$M(X) = M(X_1) + M(X_2) + \dots + M(X_n).$$

Mỗi đại lượng ngẫu nhiên X_i đều có bảng phân phối xác suất như sau:

X_i	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Do đó:

$$M(X_i) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{7}{2}.$$

Vì vậy kỳ vọng toán của tổng số chấm thu được sau n lần tung là:

$$M(X) = \frac{7n}{2}.$$

3.3.2. Phương sai

a) Định nghĩa

★ **Định nghĩa 3.6** Phương sai của đại lượng ngẫu nhiên X , ký hiệu $D(X)$, là kỳ vọng toán của bình phương sai lệch của đại lượng ngẫu nhiên so với kỳ vọng toán của nó.

$$D(X) = M[X - M(X)]^2.$$

Như vậy nếu X là đại lượng ngẫu nhiên rời rạc thì phương sai được xác định bằng công thức:

$$D(X) = \sum_{i=1}^n [x_i - M(X)]^2 \cdot p_i$$

còn nếu X là đại lượng ngẫu nhiên liên tục thì phương sai được xác định bằng công thức:

$$\int_{-\infty}^{+\infty} [x - M(X)]^2 f(x) dx.$$

b) Cách tính phương sai

Theo định nghĩa phương sai và các tính chất của kỳ vọng toán ta có:

$$\begin{aligned} D(X) &= M[X - M(X)]^2 = M(X^2 - 2X.M(X) + [M(X)]^2) \\ &= M(X^2) + M(-2X.M(X)) + M([M(X)]^2) \\ &= M(X^2) - 2M(X).M(X) + M([M(X)]^2) \\ &= M(X^2) - [M(X)]^2 \end{aligned}$$

Như vậy đối với đại lượng ngẫu nhiên rời rạc, ta có thể tính phương sai bằng công thức:

$$D(X) = \left(\sum_{i=1}^n x_i^2 p_i \right) - [M(X)]^2 = \left(\sum_{i=1}^n x_i^2 p_i \right) - \left(\sum_{i=1}^n x_i p_i \right)^2$$

còn đối với đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$:

$$D(X) = \int_{-\infty}^{+\infty} x^2 f(x) dx - [M(X)]^2 = \left(\int_{-\infty}^{+\infty} x^2 f(x) dx \right) - \left(\int_{-\infty}^{+\infty} x f(x) dx \right)^2.$$

● **Ví dụ 3.16** Cho X là đại lượng ngẫu nhiên rời rạc có bảng phân phối xác suất:

X	1	3	4
P	0,1	0,5	0,4

Tìm phương sai $D(X)$.

Lời giải. Kỳ vọng toán của đại lượng ngẫu nhiên X bằng:

$$M(X) = 1 \cdot 0,1 + 3 \cdot 0,5 + 4 \cdot 0,4 = 3,2.$$

Kỳ vọng toán $M(X^2)$ bằng:

$$M(X^2) = 1^2 \cdot 0,1 + 3^2 \cdot 0,5 + 4^2 \cdot 0,4 = 11.$$

Phương sai của đại lượng ngẫu nhiên X bằng:

$$D(X) = M(X^2) - [M(X)]^2 = 11 - 3,2^2 = 0,76.$$

• **Ví dụ 3.17** Cho X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất như sau:

$$f(x) = \begin{cases} \frac{3}{4}(x^2 + 2x) & \text{với } x \in (0; 1) \\ 0 & \text{với } x \notin (0; 1) \end{cases}$$

Tìm phương sai $D(X)$.

Lời giải. Kỳ vọng toán của đại lượng ngẫu nhiên X bằng:

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^0 x \cdot 0 dx + \int_0^1 \frac{3}{4}(x^3 + 2x^2)dx + \int_1^{+\infty} x \cdot 0 dx \\ &= \frac{3}{4} \int_0^1 (x^3 + 2x^2)dx = \frac{3}{4} \left(\frac{x^4}{4} + \frac{2x^3}{3} \right) \Big|_0^1 = \frac{11}{16}. \end{aligned}$$

Kỳ vọng toán $M(X^2)$ bằng:

$$\begin{aligned} M(X^2) &= \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_{-\infty}^0 x^2 \cdot 0 dx + \int_0^1 \frac{3}{4}(x^4 + 2x^3)dx + \int_1^{+\infty} x^2 \cdot 0 dx \\ &= \frac{3}{4} \int_0^1 (x^4 + 2x^3)dx = \frac{3}{4} \left(\frac{x^5}{5} + \frac{2x^4}{4} \right) \Big|_0^1 = \frac{21}{40}. \end{aligned}$$

Phương sai của đại lượng ngẫu nhiên X bằng:

$$D(X) = M(X^2) - [M(X)]^2 = \frac{21}{40} - \left(\frac{11}{16} \right)^2 = \frac{67}{1280}.$$

c) Các tính chất của phương sai

- Tính chất 1. Phương sai của một đại lượng ngẫu nhiên là một giá trị xác định không âm; phương sai của một hằng số bằng không:

$$D(X) \geq 0, \forall X$$

$$D(C) = 0, C \text{ là hằng số.}$$

Tính chất này được suy trực tiếp từ định nghĩa phương sai.

- Tính chất 2. Phương sai của tích giữa một hằng số và một đại lượng ngẫu nhiên bằng tích giữa bình phương hằng số đó và phương sai của đại lượng ngẫu nhiên ấy:

$$D(CX) = C^2 \cdot D(X).$$

Δ. Thật vậy, theo định nghĩa phương sai:

$$\begin{aligned} D(CX) &= M(CX - M(CX))^2 = M(CX - CM(X))^2 \\ &= M[C^2(X - M(X))^2] = C^2 \cdot M[(X - M(X))^2] = C^2 \cdot D(X). \end{aligned}$$

□

$$M(X_i) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{7}{2} \quad (\forall i = \overline{1, n})$$

$$M(X_i^2) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6} \quad (\forall i = \overline{1, n})$$

$$D(X_i) = M(X_i^2) - [M(X_i)]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12} \quad (\forall i = \overline{1, n})$$

Vì vậy phương sai của tổng số chấm thu được sau n lần tung là:

$$D(X) = \frac{35n}{12}.$$

3.3.3. Độ lệch tiêu chuẩn

★ **Định nghĩa 3.7** Độ lệch tiêu chuẩn của đại lượng ngẫu nhiên X , ký hiệu σ_X , là căn bậc hai của phương sai:

$$\sigma_X = \sqrt{D(X)}$$

Ta thấy rằng đơn vị đo của độ lệch tiêu chuẩn bằng đơn vị đo của đại lượng ngẫu nhiên. Vì vậy độ lệch chuẩn được sử dụng để đánh giá mức độ phân tán của đại lượng ngẫu nhiên theo đơn vị đo của nó.

3.3.4. Mốt

★ **Định nghĩa 3.8** Mốt X , ký hiệu $mod(X)$, là giá trị của đại lượng ngẫu nhiên X có khả năng xuất hiện lớn nhất trong một lần cạn nào đó của nó.

Đối với đại lượng ngẫu nhiên rời rạc, $mod(X)$ là giá trị của X ứng với xác suất lớn nhất.

Đối với đại lượng ngẫu nhiên liên tục, $mod(X)$ là giá trị X làm cực đại hàm mật độ.

● **Ví dụ 3.19** Đại lượng ngẫu nhiên X có bảng phân phối xác suất:

X	-2	-1	0	3
P	0,1	0,7	0,05	0,15

Suy ra $mod(X) = -1$

● **Ví dụ 3.20** Đại lượng ngẫu nhiên X có hàm mật độ xác suất:

$$f(x) = \begin{cases} 0; & x \leq 0 \\ \frac{x}{2} \cdot e^{-\frac{x^2}{4}}; & x > 0 \end{cases}$$

Xác định $mod(X)$.

Lời giải. Với $x > 0$, ta có:

$$f'(x) = \frac{1}{2}e^{-\frac{x^2}{4}} - \frac{x^2}{4} \cdot e^{-\frac{x^2}{4}}$$

$$f'(x) = 0 \Leftrightarrow \frac{1}{4}e^{-\frac{x^2}{4}}(2 - x^2) = 0 \Leftrightarrow x = \pm\sqrt{2}$$

Kết hợp với điều kiện $x > 0$ suy ra $x = \sqrt{2}$.

Lập bảng biến thiên của $f(x)$

x	$-\infty$	$-\sqrt{2}$	0	$\sqrt{2}$	$+\infty$
$f'(x)$			+	0 -	
$f(x)$	0	-	-	> 0	0

↗ CD ↘

Ta được $mod(X) = \sqrt{2}$.

3.3.5. Trung vị

★ **Định nghĩa 3.9** Trung vị là giá trị của đại lượng ngẫu nhiên X , kí hiệu là $med(X)$, là giá trị của X thỏa mãn:

$$P(X < medX) \leq \frac{1}{2} \text{ và } P(X \leq medX) \geq \frac{1}{2}$$

Như vậy nếu X là đại lượng ngẫu nhiên liên tục thì:

$$P(X < medX) = P(X \geq medX) = \frac{1}{2}$$

và do đó $medX$ là nghiệm của phương trình $F(x) = \frac{1}{2}$.

• **Ví dụ 3.21** Đại lượng ngẫu nhiên rời rạc X có bảng phân phối xác suất:

X	-1	0	2	3	5	8
P	0,15	0,05	0,25	0,15	0,35	0,05

Khi đó $medX = 3$.

• **Ví dụ 3.22** Đại lượng ngẫu nhiên X có hàm phân phối xác suất:

$$F(x) = \begin{cases} 0 & \text{với } x \leq 0 \\ x^2 & \text{với } 0 < x \leq 1 \\ 1 & \text{với } x > 1 \end{cases}$$

Tìm $med(X)$.

Lời giải. Xét phương trình

$$F(x) = \frac{1}{2} \Leftrightarrow \begin{cases} 0 < x \leq 1 \\ x^2 = \frac{1}{2} \end{cases} \Leftrightarrow \begin{cases} 0 < x \leq 1 \\ x = \pm \frac{1}{\sqrt{2}} \end{cases} \Leftrightarrow x = \frac{1}{\sqrt{2}}$$

Vậy $medX = \frac{1}{\sqrt{2}}$.

3.3.6. Phân vị

★ **Định nghĩa 3.10** Phân vị mức α của X là giá trị x_α của X thỏa mãn $P(X < x_\alpha) = \alpha$.

Như vậy nếu đại lượng ngẫu nhiên X có hàm phân phối xác suất $F(x)$ thì phân vị mức α là nghiệm X của phương trình $F(x) = \alpha$.

• **Ví dụ 3.23** Đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x) = \begin{cases} e^{-x} & \text{với } x > 0 \\ 0 & \text{với } x \leq 0 \end{cases}$$

Tính các giá trị phân vị $x_{0,95}, x_{0,5}, x_{0,05}$.

Lời giải. Từ hàm mật độ ta suy ra hàm phân phối xác suất của X như sau:

$$F(x) = \begin{cases} 1 - e^{-x} & \text{với } x > 0 \\ 0 & \text{với } x \leq 0 \end{cases}$$

- Xét phương trình $F(x) = 0,95 \Leftrightarrow 1 - e^{-x} = 0,95 \Leftrightarrow e^{-x} = 0,05 \Leftrightarrow x = -\ln 0,05 = 2,996$.
Vậy $x_{0,95} = 2,996$.
- Xét phương trình $F(x) = 0,5 \Leftrightarrow 1 - e^{-x} = 0,5 \Leftrightarrow e^{-x} = 0,5 \Leftrightarrow x = -\ln 0,5 = 0,693$.
Vậy $x_{0,5} = 0,693$. Lưu ý rằng đây chính là $med(X)$.
- Xét phương trình $F(x) = 0,095 \Leftrightarrow 1 - e^{-x} = 0,095 \Leftrightarrow e^{-x} = 0,905 \Leftrightarrow x = -\ln 0,905 = 0,095$.
Vậy $x_{0,05} = 0,095$.

3.4. Một số quy luật phân phối xác suất thông dụng

3.4.1. Quy luật phân phối chuẩn $N(\mu, \sigma^2)$

Định nghĩa

★ **Định nghĩa 3.11** Đại lượng ngẫu nhiên liên tục X nhận giá trị trong khoảng $(-\infty, +\infty)$ được gọi là phân phối theo quy luật chuẩn với hai tham số là μ và σ^2 , nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x - \mu)^2}{2\sigma^2}}$$

Quy luật phân phối chuẩn với hai tham số μ và σ^2 được ký hiệu là $N(\mu, \sigma^2)$.

Khảo sát hàm mật độ $f(x)$:

- Hàm số xác định với $x \in \mathbb{R}$.

- Đạo hàm bậc nhất:

$$f'(x) = -\frac{x - \mu}{\sigma^3 \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x - \mu)^2}{2\sigma^2}}$$

Như vậy khi $x = \mu$ hàm số có cực đại bằng $\frac{1}{\sigma \cdot \sqrt{2\pi}}$.

- Đạo hàm bậc hai:

$$f''(x) = -\frac{1}{\sigma^3 \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x - \mu)^2}{2\sigma^2}} \cdot \left\{ 1 - \frac{(x - \mu)^2}{\sigma^2} \right\}$$

$$f''(x) = 0 \Leftrightarrow x = \mu \pm \sigma.$$

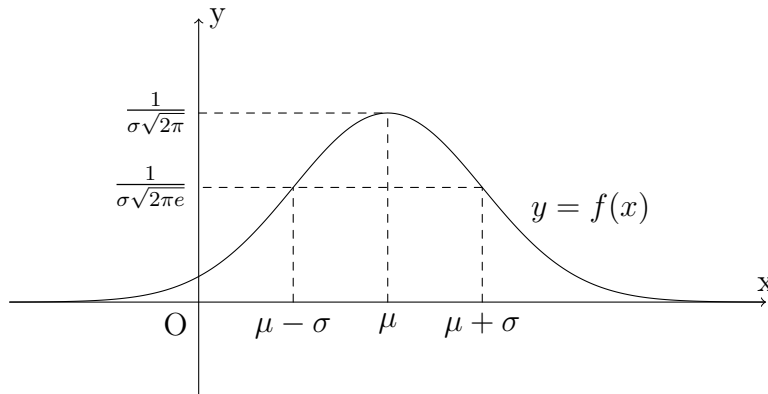
Như vậy hàm số có 2 điểm uốn: $\left(\mu - \sigma, \frac{1}{\sigma \cdot \sqrt{2\pi e}} \right)$ và $\left(\mu + \sigma, \frac{1}{\sigma \cdot \sqrt{2\pi e}} \right)$

- Giới hạn:

$$\lim_{x \rightarrow \pm\infty} f(x) = 0.$$

Như vậy trục Ox là tiệm cận ngang của đồ thị hàm số.

- Đồ thị của hàm số đối xứng qua đường thẳng $x = \mu$ và có dạng như sau:



Hình 3.5:

Các tham số đặc trưng của quy luật phân phối chuẩn

Cho X là đại lượng ngẫu nhiên phân phối chuẩn với hai tham số μ và σ^2 .

- a) Kỳ vọng toán của đại lượng ngẫu nhiên X chính là tham số μ :

$$M(X) = \mu.$$

Δ . Theo định nghĩa kỳ vọng toán của đại lượng ngẫu nhiên liên tục ta có:

$$M(X) = \int_{-\infty}^{+\infty} x f(x) dx = \frac{1}{\sigma \cdot \sqrt{2\pi}} \int_{-\infty}^{+\infty} x \cdot e^{-\frac{(x - \mu)^2}{2\sigma^2}} dx.$$

Thực hiện phép đổi biến số: $u = \frac{x - \mu}{\sigma}$. Khi đó cận tích phân không đổi và ta được:

$$\begin{aligned} M(X) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma u + \mu) \cdot e^{-\frac{u^2}{2}} du \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u \cdot e^{-\frac{u^2}{2}} du + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du. \end{aligned}$$

Tích phân thứ nhất bằng không do hàm số dưới dấu tích phân là hàm lẻ mà cận lấy tích phân lại đối xứng qua 0. Còn tích phân thứ hai là tích phân Poisson:

$$J = \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du = \sqrt{2\pi}.$$

Do đó $M(X) = \mu$ □

b) Độ lệch tiêu chuẩn của đại lượng ngẫu nhiên X chính là tham số σ :

$$\sigma_X = \sigma \text{ hay } D(X) = \sigma^2$$

Δ. Do $M(X) = \mu$, theo định nghĩa phương sai của đại lượng ngẫu nhiên liên tục:

$$D(X) = \int_{-\infty}^{+\infty} (x - M(X))^2 \cdot f(x) dx = \frac{1}{\sigma \cdot \sqrt{2\pi}} \int_{-\infty}^{+\infty} (x - \mu)^2 \cdot e^{-\frac{(x - \mu)^2}{2\sigma^2}} dx.$$

Thực hiện phép đổi biến số: $u = \frac{x - \mu}{\sigma}$. Khi đó cận tích phân không đổi và ta được:

$$\begin{aligned} D(X) &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u^2 \cdot e^{-\frac{u^2}{2}} du \\ &= -\frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u de^{-\frac{u^2}{2}} \\ &= -\frac{\sigma^2}{\sqrt{2\pi}} u \cdot e^{-\frac{u^2}{2}} \Big|_{-\infty}^{+\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du \quad (\text{tích phân từng phần}) \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \left(\lim_{x \rightarrow -\infty} \frac{u}{e^{\frac{u^2}{2}}} - \lim_{x \rightarrow +\infty} \frac{u}{e^{\frac{u^2}{2}}} \right) + \frac{\sigma^2}{\sqrt{2\pi}} \cdot \sqrt{2\pi} \quad (\text{tích phân Poisson}) \\ &= 0 + \sigma^2 = \sigma^2. \end{aligned}$$

* **Nhận xét:** Nếu đại lượng ngẫu nhiên X phân phối chuẩn với hai tham số μ và σ^2 thì đại lượng ngẫu nhiên $U = \frac{X - \mu}{\sigma}$ có kỳ vọng toán $M(U) = 0$ và độ lệch chuẩn $\sigma_U = 1$. □

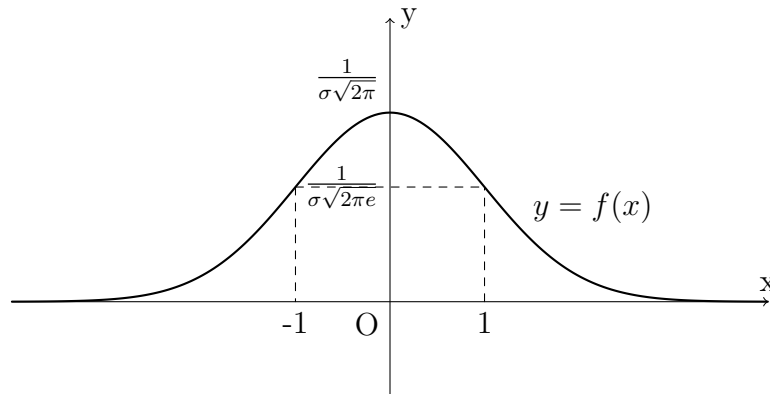
Quy luật phân phối chuẩn hoá

a) Định nghĩa

★ **Định nghĩa 3.12** Đại lượng ngẫu nhiên liên tục U nhận các giá trị trong khoảng $(-\infty, +\infty)$ được gọi là tuân theo quy luật phân phối chuẩn hoá nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Hàm mật độ $f(x)$ của quy luật phân phối chuẩn hoá có đồ thị đối xứng qua trục tung Oy như sau:



Hình 3.6:

Dễ thấy:

- Hàm số $f(x)$ là hàm số chẵn: $f(-x) = f(x)$.
- $\lim_{x \rightarrow +\infty} f(x) = 0$.

Các giá trị của hàm $f(x)$ được tính sẵn thành bảng (xem Phụ lục C.1).

b) Hàm phân phối xác suất của đại lượng ngẫu nhiên phân phối chuẩn hoá

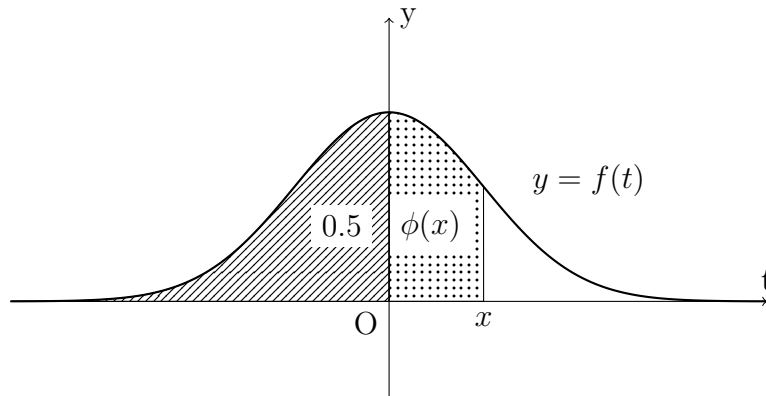
Từ hàm mật độ $f(x)$ ta có hàm phân phối xác suất $F(x)$ của đại lượng ngẫu nhiên phân phối chuẩn hoá có dạng:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

$$\text{Xét hàm số (tích phân Laplace): } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt.$$

Dễ thấy:

- $F(x) = 0,5 + \Phi(x)$.
- Hàm số $\Phi(x)$ là hàm số lẻ: $\Phi(-x) = -\Phi(x)$.
- $\lim_{x \rightarrow +\infty} \Phi(x) = 0,5$.



Hình 3.7:

Các giá trị của hàm $\Phi(x)$ được tính sẵn thành bảng (xem Phụ lục C.2).

c) Các tham số đặc trưng của đại lượng ngẫu nhiên phân phối chuẩn hoá

Giả sử U là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn hoá. Dễ thấy hàm mật độ của nó là hàm mật độ của quy luật phân phối chuẩn với hai tham số $\mu = 0$ và $\sigma^2 = 1$. Vì vậy các tham số đặc trưng của U là:

Kỳ vọng toán: $M(U) = 0$.

Phương sai: $D(U) = 1$; độ lệch tiêu chuẩn: $\sigma_U = 1$.

Phân phối chuẩn hoá được ký hiệu là $N(0, 1)$.

d) Phân vị chuẩn

★ **Định nghĩa 3.13** Phân vị chuẩn mức α ($0 \leq \alpha \leq 1$), ký hiệu u_α , là phân vị mức α của đại lượng ngẫu nhiên U có phân phối chuẩn hoá. Như vậy u_α thoả mãn điều kiện:

$$P(U < u_\alpha) = \alpha.$$

Từ hàm mật độ của phân phối chuẩn hoá ta có biểu thức xác định u_α như sau:

$$P(U < u_\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_\alpha} e^{-\frac{t^2}{2}} dt = \alpha.$$

Cho trước α , dựa vào biểu thức trên ta tính được phân vị u_α và ngược lại. Bảng tính sẵn giá trị u_α tương ứng với mức α cho trước được trình bày trong phần phụ lục (xem Phụ lục C.3).

Do tính chất đối xứng của hàm mật độ phân phối chuẩn hoá nên phân vị chuẩn có tính chất:

$$u_\alpha = -u_{1-\alpha}.$$

Các công thức tính xác suất của quy luật phân phối chuẩn

Cho đại lượng ngẫu nhiên liên tục X có phân phối chuẩn với hai tham số μ và σ^2 .

a) Xác suất để đại lượng ngẫu nhiên X nhận giá trị trong khoảng (a, b)

Ta có: $U = \frac{X - \mu}{\sigma}$ là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn hoá $N(0, 1)$, do đó:

$$\begin{aligned} P(a < X < b) &= P\left(\frac{a - \mu}{\sigma} < X < \frac{b - \mu}{\sigma}\right) \\ &= F\left(\frac{b - \mu}{\sigma}\right) - F\left(\frac{a - \mu}{\sigma}\right) \quad (F(x) \text{ là hàm phân phối chuẩn hoá}) \\ &= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right) \quad (\text{tích phân Laplace}) \end{aligned}$$

Như vậy ta thu được công thức:

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right) \quad (3.2)$$

Bảng giá trị của hàm số $\Phi(x)$ cho ở phụ lục C.2.

b) Xác suất của sự sai lệch giữa đại lượng ngẫu nhiên X và kỳ vọng toán của nó

Ta phải tính xác suất để xảy ra bất đẳng thức: $|X - \mu| < \epsilon$.

Sử dụng công thức (3.2) ta có:

$$\begin{aligned} P(|X - \mu| < \epsilon) &= P(\mu - \epsilon < X < \mu + \epsilon) \\ &= \Phi\left(\frac{\mu + \epsilon - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \epsilon - \mu}{\sigma}\right) = \Phi\left(\frac{\epsilon}{\sigma}\right) - \Phi\left(-\frac{\epsilon}{\sigma}\right) \\ &= \Phi\left(\frac{\epsilon}{\sigma}\right) + \Phi\left(\frac{\epsilon}{\sigma}\right) = 2 \cdot \Phi\left(\frac{\epsilon}{\sigma}\right) \end{aligned}$$

Như vậy ta thu được công thức:

$$P(|X - \mu| < \epsilon) = 2 \cdot \Phi\left(\frac{\epsilon}{\sigma}\right) \quad (3.3)$$

c) Quy tắc "3 - σ "

Nếu trong công thức (3.3) ta đặt $\epsilon = 3\sigma$ tức là bằng 3 lần độ lệch tiêu chuẩn của X thì ta có:

$$P(|X - \mu| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) = 2\Phi(3) = 2 \cdot 0,49865 = 0,9973$$

Điều đó cho thấy xác suất để đại lượng ngẫu nhiên phân phối chuẩn nhận giá trị trong khoảng $(\mu - 3\sigma, \mu + 3\sigma)$ lên tới 99,73%, còn xác suất để nó nhận giá trị ngoài khoảng đó chỉ có 0,27%. Xuất phát từ nguyên lý xác suất nhỏ, có thể cho rằng thực tế điều đó là không thể có.

Như vậy, nếu đại lượng ngẫu nhiên X phân phối chuẩn thì hầu như chắc chắn nó sẽ nhận giá trị trong khoảng $(\mu - 3\sigma, \mu + 3\sigma)$.

Mặt khác, nếu quy luật phân phối xác suất của đại lượng ngẫu nhiên được nghiên cứu chưa biết, song nó thoả mãn điều kiện của quy tắc "ba xích ma" thì có thể xem nó là đại lượng ngẫu nhiên phân phối chuẩn.

• Ví dụ 3.24 Kích thước của các chi tiết do một máy sản xuất ra là đại lượng ngẫu nhiên phân phối chuẩn với kích thước trung bình $\mu = 5\text{cm}$ và độ lệch tiêu chuẩn $\sigma = 0,9\text{cm}$.

- Tính tỷ lệ chi tiết có kích thước nằm trong khoảng từ 4 cm đến 7 cm.
- Tìm xác suất để lấy ngẫu nhiên 3 chi tiết có đúng 1 chi tiết có kích thước sai lệch so với trung bình không quá 1 cm.
- Hầu hết các chi tiết có kích thước như thế nào?

Lời giải. Gọi $X(\text{cm})$ là kích thước chi tiết do máy sản xuất. Theo giả thiết X phân phối chuẩn với hai tham số $\mu = 5\text{cm}$ và $\sigma = 0,9\text{cm}$.

a) Theo công thức (3.2) ta có:

$$\begin{aligned} P(4 < X < 7) &= \Phi\left(\frac{7 - 5}{0,9}\right) - \Phi\left(\frac{4 - 5}{0,9}\right) \\ &= \Phi(2,22) - \Phi(-1,11) = \Phi(2,22) + \Phi(1,11) \\ &= 0,48679 + 0,36650 = 0,85329. \end{aligned}$$

Vậy tỷ lệ chi tiết có kích thước nằm trong khoảng từ 4cm đến 7cm là 85,329%.

b) Theo công thức (3.3) ta có xác suất để lấy ngẫu nhiên 1 chi tiết có kích thước sai lệch so với trung bình không quá 1 cm là:

$$P(|X - \mu| < 1) = 2\Phi\left(\frac{1}{0,9}\right) = 2\Phi(1,11) = 2 \cdot 0,36650 = 0,733.$$

Vậy xác suất để lấy ngẫu nhiên 3 chi tiết có đúng 1 chi tiết có kích thước sai lệch so với trung bình không quá 1 cm bằng:

$$P_3(1) = C_3^1 \cdot (0,733)^1 \cdot (1 - 0,733)^2 \approx 0,157.$$

c) Theo quy tắc "3 - σ ", hầu hết chi tiết có kích thước trong khoảng:

$$(\mu - 3\sigma, \mu + 3\sigma) = (5 - 3 \cdot 0,9; 5 + 3 \cdot 0,9) = (2,3; 7,7)(cm).$$

Ứng dụng của quy luật chuẩn

Quy luật phân phối chuẩn là quy luật phân phối xác suất được áp dụng rất rộng rãi trong thực tế. Nhà toán học Nga Liapunốp đã chứng minh được rằng: Nếu đại lượng ngẫu nhiên X là tổng của một số lớn các đại lượng ngẫu nhiên độc lập và giá trị của mỗi đại lượng chỉ chiếm vị trí rất nhỏ trong tổng đó thì X sẽ có phân phối xấp xỉ chuẩn.

Trong thực tế ta gặp chính các đại lượng ngẫu nhiên như vậy. Chẳng hạn, kích thước của các chi tiết do các nhà máy sản xuất ra sẽ phân phối chuẩn nếu quá trình sản xuất diễn ra bình thường; năng suất của cùng một loại cây trồng tại các thửa ruộng khác nhau cũng phân phối chuẩn; năng suất lao động của các công nhân có cùng tay nghề và làm cùng một công việc như nhau cũng phân phối chuẩn; một số chỉ tiêu về sinh lý của những người cùng giới (chiều cao, cân nặng, vòng ngực, chiều dài cánh tay ...) cũng phân phối theo quy luật chuẩn. Sự nhận biết này cho phép lập kế hoạch sản xuất phù hợp năng suất lao động và nhu cầu người tiêu dùng...

3.4.2. Quy luật không - một $A(p)$

Ví dụ

• **Ví dụ 3.25** Giả sử trong bình có N quả cầu trong đó có M quả cầu trắng, $N - M$ quả cầu đen. Từ bình lấy ngẫu nhiên 1 quả cầu. Gọi X là số cầu trắng lấy được.

X là đại lượng ngẫu nhiên rời rạc nhận một trong hai giá trị 0; 1. Các xác suất tương ứng là:

$$P(X = 0) = \frac{N - M}{N}; \quad P(X = 1) = \frac{M}{N}$$

Nếu đặt $p = \frac{M}{N}$ và $q = 1 - p = \frac{N - M}{N}$ thì ta có:

$$P(X = 0) = q = p^0 \cdot q^1; \quad P(X = 1) = p = p^1 \cdot q^0$$

Do đó có thể viết quy luật phân phối xác suất của X như sau:

$$P(X = x) = q = p^x \cdot q^{1-x}; \quad \text{với } x = 0; 1, \text{ trong đó } q = 1 - p.$$

Định nghĩa

★ **Định nghĩa 3.14** Đại lượng ngẫu nhiên rời rạc X nhận một trong hai giá trị có thể có $X = 0; 1$ được gọi là phân phối theo quy luật không - một với tham số p ($0 < p < 1$) nếu các xác suất được tính bằng công thức:

$$P(X = x) = q = p^x \cdot q^{1-x}; \quad \text{với } x = 0; 1, \text{ trong đó } q = 1 - p.$$

Quy luật không - một với tham số p được ký hiệu là $A(p)$.

Các tham số đặc trưng của quy luật không - một

Gọi X là đại lượng ngẫu nhiên tuân theo quy luật không - một với tham số p . Ta có bảng phân phối xác suất của X như sau:

X	0	1
P	q	p

trong đó $q = 1 - p$.

Theo bảng phân phối xác suất của X ta tính được các tham số đặc trưng của X như sau:

- Kỳ vọng toán:

$$\begin{aligned} M(X) &= 0 \cdot q + 1 \cdot p = p \\ M(X^2) &= 0^2 \cdot q + 1^2 \cdot p = p. \end{aligned}$$

- Phương sai: $D(X) = M(X^2) - [M(X)]^2 = p - p^2 = pq$.

- Độ lệch chuẩn: $\sigma_X = \sqrt{D(X)} = \sqrt{pq}$.

3.4.3. Quy luật nhị thức $B(n, p)$ **Ví dụ**

• **Ví dụ 3.26** Giả sử trong bình có N quả cầu trong đó có M quả cầu trắng, $N - M$ quả cầu đen. Từ bình lấy ngẫu nhiên n quả cầu theo phương thức có hoàn lại. Gọi X là số cầu trắng lấy được.

Như vậy ta đã thực hiện n phép thử độc lập (mỗi phép thử là một lần lấy 1 quả cầu). Trong mỗi phép thử chỉ xảy ra hai trường hợp: hoặc lấy được cầu trắng (biến cố A) hoặc không lấy được cầu trắng (biến cố \bar{A}). Xác suất lấy được cầu trắng mỗi lần đều bằng $p = \frac{M}{N}$ và xác suất lấy được

cầu đen mỗi lần đều bằng $q = \frac{N - M}{N} = 1 - p$. Do đó theo công thức Bernoulli xác suất để lấy được k cầu trắng là:

$$P(X = k) = C_n^k \cdot p^k \cdot q^{n-k}, \quad k = \overline{0, n}, \quad q = 1 - p$$

Ta cũng dễ nhận thấy là nếu gọi X_k là số cầu trắng lấy được ở lần thứ k thì X_k tuân theo quy luật không - một $A(p)$ và $X = X_1 + X_2 + \dots + X_n$.

Định nghĩa

★ **Định nghĩa 3.15** Đại lượng ngẫu nhiên rời rạc X nhận một trong các giá trị có thể có $X = 0, 1, 2, \dots, n$ được gọi là phân phối theo quy luật nhị thức với các tham số là n và p nếu các xác suất tương ứng của nó được tính bởi công thức:

$$P(X = k) = C_n^k \cdot p^k \cdot q^{n-k}, \quad k = \overline{0, n}, \quad q = 1 - p$$

Quy luật nhị thức được ký hiệu là $B(n, p)$.

* **Nhận xét:** Giả sử bài toán tuân theo lược đồ Bemoulli: Thực hiện dãy n phép thử độc lập; trong mỗi phép thử chỉ có hai khả năng: biến cố A xảy ra hoặc biến cố \bar{A} không xảy ra; xác suất xảy ra biến cố A trong mỗi phép thử đều bằng p ; xác suất xảy ra biến cố \bar{A} trong mỗi phép thử là $q = 1 - p$. Khi đó nếu gọi X là số lần xuất hiện biến cố A trong dãy n phép thử trên thì X là đại lượng ngẫu nhiên phân phối theo quy luật nhị thức $B(n, p)$.

Các tham số đặc trưng của quy luật nhị thức

Giả sử X là đại lượng ngẫu nhiên phân phối theo quy luật nhị thức $B(n, p)$. Bảng phân phối xác suất của X như sau:

X	0	1	2	...	k	...	n
P	q^n	$C_n^1 p \cdot q^{n-1}$	$C_n^2 p^2 \cdot q^{n-2}$...	$C_n^k p^k \cdot q^{n-k}$...	p^n

a) Xác suất để đại lượng ngẫu nhiên X nhận giá trị trong khoảng $[k; k+h]$

Biến cố $(k \leq X \leq k+h)$ (k, h là các số tự nhiên; $k+h \leq n$) có thể tách ra thành tổng của $h+1$ biến cố xung khắc từng đôi là $(X=k), (X=k+1), \dots, (X=k+h)$. Do đó áp dụng định lý cộng xác suất ta có:

$$\begin{aligned} P(k \leq X \leq k+h) &= P(X=k) + P(X=k+1) + \dots + P(X=k+h) \\ &= p_k + p_{k+1} + \dots + p_{k+h} \quad \text{với } p_k = P(X=k) \end{aligned}$$

b) Kỳ vọng toán và phương sai của đại lượng ngẫu nhiên phân phối nhị thức

Từ bảng phân phối xác suất của X ta có:

$$M(X) = \sum_{k=0}^n k \cdot C_n^k p^k \cdot q^{n-k} \quad M(X^2) = \sum_{k=0}^n k^2 \cdot C_n^k p^k \cdot q^{n-k}$$

Xét tổng:

$$\sum_{k=0}^n C_n^k p^k \cdot q^{n-k} = (p+q)^n \quad (3.4)$$

Đạo hàm 2 vế của (3.4) theo p ta được:

$$\begin{aligned} \sum_{k=1}^n k \cdot C_n^k p^{k-1} \cdot q^{n-k} &= n(p+q)^{n-1} \\ \Rightarrow \sum_{k=0}^n k \cdot C_n^k p^k \cdot q^{n-k} &= np(p+q)^{n-1} \end{aligned} \quad (3.5)$$

Vế trái của (3.5) là $M(X)$, còn vế phải bằng np do $p+q=l$. Vì vậy $M(X) = np$.

Đạo hàm 2 vế của (3.5) theo p ta được:

$$\begin{aligned} \sum_{k=1}^n k^2 \cdot C_n^k p^{k-1} \cdot q^{n-k} &= n \cdot (p+q)^{n-1} + n \cdot (n-1) \cdot p \cdot (p+q)^{n-2} \\ \Rightarrow \sum_{k=0}^n k^2 \cdot C_n^k p^k \cdot q^{n-k} &= np(p+q)^{n-1} + n \cdot (n-1) \cdot p^2 \cdot (p+q)^{n-2} \\ \Rightarrow M(X^2) &= np + n(n-1)p^2 \end{aligned}$$

Do đó:

$$D(X) = M(X^2) - [M(X)]^2 = np + n(n-1)p^2 - n^2p^2 = np(1-p) = npq$$

Ở trên ta đã tính $M(X)$ và $D(X)$ trong trường hợp $n \geq 2$. Tuy nhiên nếu $n=1$ thì dễ thấy X tuân theo quy luật không - một $A(p)$ và các kết quả trên vẫn đúng.

Như vậy đại lượng ngẫu nhiên phân phối nhị thức $B(n, p)$ có:

- Kỳ vọng toán : $M(X) = np$.

- Phương sai : $D(X) = npq$.
- Độ lệch tiêu chuẩn : $\sigma_X = \sqrt{npq}$.
- Mod X: $np - q \leq Mod(X) \leq np + p$.

•**Ví dụ 3.27** Xác suất để một người bắn trúng bia là 0,8.

- a) Giả sử người đó bắn 12 phát đạn, tính số viên đạn trúng bia có khả năng xảy ra nhiều nhất và xác suất để xảy ra điều đó.
- b) Muốn số đạn trúng bia trung bình không nhỏ hơn 12 thì người đó phải bắn ít nhất bao nhiêu viên đạn?

Lời giải. Gọi X là số viên đạn mà người đó bắn trúng bia. X tuân theo quy luật nhị thức với $p = 0,8$.

a) Người đó bắn $n = 12$ viên đạn, khi đó X tuân theo quy luật nhị thức $B(n = 12; p = 0,8)$. Số viên đạn trúng bia có khả năng xảy ra nhiều nhất là $mod(X)$:

$$np - q \leq Mod(X) \leq np + p \Leftrightarrow 12 \cdot 0,8 - 0,2 \leq Mod(X) \leq 12 \cdot 0,8 + 0,8$$

$$\Leftrightarrow 9,4 \leq Mod(X) \leq 10,4 \Rightarrow Mod(X) = 10$$

Xác suất để người đó bắn trúng 10 viên đạn là:

$$P(X = 10) = C_{12}^{10} 0,8^{10} 0,2^2 \approx 0,2835$$

b) Giả sử người đó bắn n viên đạn, khi đó X tuân theo quy luật nhị thức $B(n; p = 0,8)$. Theo giả thiết:

$$M(X) = 12 \Leftrightarrow np = 12 \Leftrightarrow n = \frac{12}{0,8} = 15$$

Vậy muốn trung bình có 12 viên đạn trúng bia thì người đó phải bắn 15 viên đạn.

Công thức xấp xỉ Laplace

Giả sử X là đại lượng ngẫu nhiên phân phối nhị thức $B(n, p)$. Trong thực tế, nếu n khá lớn thì việc tính toán theo công thức Bemoulli sẽ gặp khó khăn, lúc đó có thể dùng quy luật chuẩn để thay thế cho quy luật nhị thức nếu thỏa mãn đồng thời hai điều kiện:

$$\begin{cases} n > 5 \\ \frac{|p - q|}{\sqrt{npq}} < 0,3 \end{cases}$$

Lúc đó thì đại lượng ngẫu nhiên X phân phối nhị thức $B(n, p)$ có thể coi như phân phối xấp xỉ chuẩn với kỳ vọng toán $\mu = np$ và phương sai $\sigma^2 = npq$. Từ đó ta có các công thức xấp xỉ Laplace như sau:

$$P(X = k) = C_n^k p^k \cdot q^{n-k} \approx \frac{1}{\sqrt{npq}} f\left(\frac{k - np}{\sqrt{npq}}\right)$$

$$P(a \leq X \leq b) \approx \Phi\left(\frac{b - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np}{\sqrt{npq}}\right)$$

Ở đó $f(x)$ là hàm mật độ của quy luật phân phối chuẩn hoá, và $\Phi(x)$ là tích phân Laplace:

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \text{ có giá trị được cho ở phụ lục C.1.}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt \text{ có giá trị được cho ở phụ lục C.2.}$$

• **Ví dụ 3.28** Xác suất để sản phẩm sau khi sản xuất không được kiểm tra chất lượng bằng 0,2. Giả sử trong kho của nhà máy có 400 sản phẩm vừa sản xuất ra.

- Tính xác suất để có 80 sản phẩm không được kiểm tra chất lượng.
- Tính xác suất để có từ 70 đến 100 sản phẩm không được kiểm tra chất lượng.
- Với xác suất 0,9545 hãy tính xem số lượng sản phẩm không được kiểm tra chất lượng nằm trong khoảng đối xứng nào xung quanh giá trị trung bình?

Lời giải. Gọi X là số sản phẩm không được kiểm tra chất lượng trong số 400 sản phẩm. Khi đó X tuân theo quy luật phân phối nhị thức $B(n = 400; p = 0,2)$. Tuy nhiên vì n khá lớn $n = 400 > 5$ và

$$\frac{|p - q|}{\sqrt{npq}} = \frac{|0,2 - 0,8|}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = 0,075 < 0,3$$

nên có thể coi X tuân theo quy luật phân phối chuẩn $N(\mu = np = 80; \sigma^2 = npq = 64)$.

a) Xác suất để có 80 sản phẩm không được kiểm tra chất lượng là:

$$P(X = 80) \approx \frac{1}{\sqrt{64}} f\left(\frac{80 - 80}{\sqrt{64}}\right) = \frac{1}{8} \cdot f(0) = \frac{1}{8} \cdot 0,39894 = 0,04987.$$

b) Xác suất để có từ 70 đến 100 sản phẩm không được kiểm tra chất lượng là:

$$\begin{aligned} P(70 \leq X \leq 100) &\approx \Phi\left(\frac{100 - 80}{\sqrt{64}}\right) - \Phi\left(\frac{70 - 80}{\sqrt{64}}\right) = \Phi(2,5) - \Phi(-1,25) \\ &= \Phi(2,5) + \Phi(1,25) = 0,49379 + 0,39435 = 0,88814. \end{aligned}$$

c) Ta phải tìm ϵ sao cho $P(|X - np| < \epsilon) = 0,9545$. Ta có:

$$\begin{aligned} P(|X - np| < \epsilon) &= P(np - \epsilon < X < np + \epsilon) = \Phi\left(\frac{np + \epsilon - np}{\sqrt{npq}}\right) - \Phi\left(\frac{np - \epsilon - np}{\sqrt{npq}}\right) \\ &= \Phi\left(\frac{\epsilon}{\sqrt{npq}}\right) - \Phi\left(\frac{-\epsilon}{\sqrt{npq}}\right) = 2\Phi\left(\frac{\epsilon}{\sqrt{npq}}\right) = 2\Phi\left(\frac{\epsilon}{8}\right). \end{aligned}$$

Do đó;

$$2\Phi\left(\frac{\epsilon}{8}\right) = 0,9545 \Leftrightarrow \Phi\left(\frac{\epsilon}{8}\right) = 0,47725 \Leftrightarrow \frac{\epsilon}{8} = 2,0 \Leftrightarrow \epsilon = 16$$

Vậy số lượng sản phẩm không được kiểm tra chất lượng nằm trong khoảng:

$$(np - \epsilon; np + \epsilon) = (80 - 16; 80 + 16) = (64; 96).$$

3.4.4. Quy luật Poisson $P(\lambda)$

Định nghĩa

★ **Định nghĩa 3.16** Đại lượng ngẫu nhiên rời rạc X nhận một trong các giá trị có thể có $X = 0, 1, \dots$ được gọi là phân phối theo quy luật Poisson với tham số $\lambda (\lambda > 0)$ nếu các xác suất tương ứng của nó được tính bởi công thức:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad k = 0, 1, \dots$$

Quy luật Poisson được ký hiệu là $P(\lambda)$.

X	0	1	2	...	k	...
P	$e^{-\lambda}$	$\frac{\lambda^1}{1!}e^{-\lambda}$	$\frac{\lambda^2}{2!}e^{-\lambda}$...	$\frac{\lambda^k}{k!}e^{-\lambda}$...

Các tham số đặc trưng của quy luật Poisson

a) Bảng phân phối xác suất của đại lượng ngẫu nhiên X phân phối theo quy luật Poisson có dạng:

Chú ý rằng:

$$e^\lambda = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \quad \text{và} \quad \lambda e^\lambda + e^\lambda = \sum_{k=1}^{+\infty} \frac{k \cdot \lambda^{k-1}}{(k-1)!}$$

ta có các kết quả sau đây:

b) Kỳ vọng toán của đại lượng ngẫu nhiên X phân phối $P(\lambda)$:

$$\begin{aligned} M(X) &= \sum_{k=0}^{+\infty} k \cdot P(X = k) = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \cdot \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} \\ &= \lambda \cdot e^{-\lambda} \cdot \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot e^\lambda = \lambda \end{aligned}$$

Tương tự ta cũng có:

$$\begin{aligned} M(X^2) &= \sum_{k=0}^{+\infty} k^2 \cdot P(X = k) = \sum_{k=0}^{+\infty} k^2 \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \cdot \sum_{k=1}^{+\infty} k \cdot \frac{\lambda^k}{(k-1)!} \\ &= \lambda \cdot e^{-\lambda} \cdot \sum_{k=1}^{+\infty} \frac{k \cdot \lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot (\lambda e^\lambda + e^\lambda) = \lambda^2 + \lambda \end{aligned}$$

c) Phương sai của đại lượng ngẫu nhiên X phân phối $P(\lambda)$:

$$D(X) = M(X^2) - [M(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Như vậy trong quy luật Poisson, cả kỳ vọng toán và phương sai đều bằng λ , đó là tính chất đặc biệt của quy luật Poisson.

d) Giá trị của X tương ứng với xác suất lớn nhất là:

$$\lambda - 1 \leq \text{mod}(X) \leq \lambda.$$

Kết quả này được chứng minh tương tự phần công thức Bemoulli:

$$\begin{cases} P(X = k) \geq P(X = k - 1) \\ P(X = k) \geq P(X = k + 1) \end{cases} \Leftrightarrow \begin{cases} \frac{\lambda^k}{k!} \geq \frac{\lambda^{k-1}}{(k-1)!} \\ \frac{\lambda^k}{k!} \geq \frac{\lambda^{k+1}}{(k+1)!} \end{cases} \Leftrightarrow \begin{cases} \frac{\lambda}{k} \geq 1 \\ 1 \geq \frac{\lambda}{k+1} \end{cases} \Leftrightarrow \begin{cases} k \leq \lambda \\ k \geq \lambda - 1 \end{cases}$$

Liên hệ giữa quy luật Poisson với quy luật nhị thức và quy luật chuẩn

Giả sử X là đại lượng ngẫu nhiên phân phối nhị thức $B(n, p)$. Trong trường hợp n khá lớn và p khá nhỏ ($p < 0, 1$), phân phối nhị thức được xấp xỉ bởi phân phối Poisson.

Thật vậy, giả sử tích np luôn luôn bằng một giá trị không đổi $np = \lambda$, khi đó xác suất của quy luật nhị thức có thể viết như sau:

$$\begin{aligned} p_k &= P(X = k) = C_n^k p^k q^{n-k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} p^k q^{n-k} \\ &= \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

Vì n lớn, do đó thay cho p_k ta tìm giới hạn $\lim_{n \rightarrow \infty} p_k$:

$$\lim_{n \rightarrow \infty} p_k = \lim_{n \rightarrow \infty} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{\frac{-n}{\lambda} \cdot \frac{\lambda(n-k)}{-n}} = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$$

Như vậy là trong trường hợp số phép thử n rất lớn, xác suất p rất nhỏ và tích $np = \lambda$ không đổi, các xác suất p_k của công thức Bernoulli có thể thay thế bằng công thức xấp xỉ Poisson.

Mặt khác, đối với quy luật Poisson thì quá trình hội tụ của nó về quy luật chuẩn sẽ diễn ra khi λ trở nên lớn hơn 20. Vì vậy nếu đại lượng ngẫu nhiên X phân phối theo quy luật Poisson song $\lambda > 20$ thì có thể xem X phân phối xấp xỉ chuẩn với $\mu = \lambda$ và $\sigma^2 = \lambda$.

• Ví dụ 3.29 Một máy dệt có 5000 ống sợi, xác suất để trong một phút một ống sợi bị đứt bằng 0,0002. Tìm xác suất để trong một phút có không quá 2 ống sợi bị đứt.

Lời giải. Gọi X là số ống sợi bị đứt trong một phút. Dễ thấy X là đại lượng ngẫu nhiên phân phối nhị thức $B(n = 5000, p = 0,0002)$. Tuy nhiên vì $n = 5000$ rất lớn, $p = 0,0002$ khá nhỏ và tích $np = 5000 \cdot 0,0002 = 1$ không đổi nên có thể coi X phân phối theo quy luật Poisson $P(\lambda = 1)$. Xác suất để trong một phút có không quá 2 ống sợi bị đứt là:

$$\begin{aligned} P(X \leq 2) &= p_0 + p_1 + p_2 = e^{-\lambda} \frac{\lambda^0}{0!} + e^{-\lambda} \frac{\lambda^1}{1!} + e^{-\lambda} \frac{\lambda^2}{2!} \\ &= e^{-1} \left(1 + \frac{1^1}{1!} + \frac{1^2}{2!}\right) = \frac{5}{2e} \approx 0,9197. \end{aligned}$$

3.4.5. Quy luật siêu bội $M(N, n)$

Ví dụ

• Ví dụ 3.30 Giả sử trong bình có N quả cầu, trong đó có M quả cầu trắng và $N - M$ quả cầu đen. Lấy ngẫu nhiên lần lượt n quả cầu theo phương thức không hoàn lại. Gọi X là số cầu trắng được lấy ra. Khi đó các phép thử không còn độc lập và ta không thể tính xác suất bằng công thức Bernoulli. Theo định nghĩa cổ điển về xác suất, xác suất để trong n quả cầu lấy ra có đúng k cầu trắng bằng:

$$P(X = k) = p_k = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n}; \quad (0 \leq k \leq M; \quad 0 \leq n - k \leq N - M).$$

Định nghĩa

★ **Định nghĩa 3.17** Đại lượng ngẫu nhiên rời rạc X nhận một trong các giá trị có thể có $X = 0, 1, 2, \dots$ được gọi là phân phối theo quy luật siêu bội với các tham số $N, M, n (n \leq N; M \leq N)$ nếu xác suất tương ứng của nó được tính bởi công thức:

$$P(X = k) = p_k = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n}; \quad \max(0; n - N + M) \leq k \leq \min(M; n).$$

Quy luật siêu bội được ký hiệu là $M(N, n)$.

Các tham số đặc trưng của quy luật siêu bội

Đại lượng ngẫu nhiên X phân phối theo quy luật siêu bội $M(N, n)$ có kỳ vọng toán:

$$M(X) = n \cdot \frac{M}{N} = np \quad \text{với } p = \frac{M}{N}$$

và phương sai:

$$D(X) = n \cdot \frac{M}{N} \cdot \frac{N-M}{N} \cdot \frac{N-n}{N-1} = npq \cdot \frac{N-n}{N-1}$$

Ta chú ý rằng khi số phép thử n là rất bé so với N thì phân phối siêu bội không khác so với phân phối nhị thức vì:

$$\lim_{\substack{n \\ N \rightarrow 0}} \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n} = C_n^k \left(\frac{M}{N} \right)^k \cdot \left(1 - \frac{M}{N} \right)^{n-k}.$$

3.4.6. Quy luật khi - bình phương $\chi^2(n)$

★ **Định nghĩa 3.18** Đại lượng ngẫu nhiên liên tục X được gọi là tuân theo quy luật phân phối khi- bình phương với n bậc tự do, ký hiệu $\chi^2(n)$, nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = \begin{cases} 0 & \text{với } x \leq 0 \\ \frac{1}{2^{n/2} \cdot \Gamma(n/2)} e^{-\frac{x}{2}} \cdot x^{\frac{n}{2}-1} & \text{với } x > 0 \end{cases}$$

trong đó $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ là hàm Gamma; $\Gamma(x) = (x-1)\Gamma(x-1)$.

Nếu n là một số nguyên thì $\Gamma(n) = (n-1)!$.

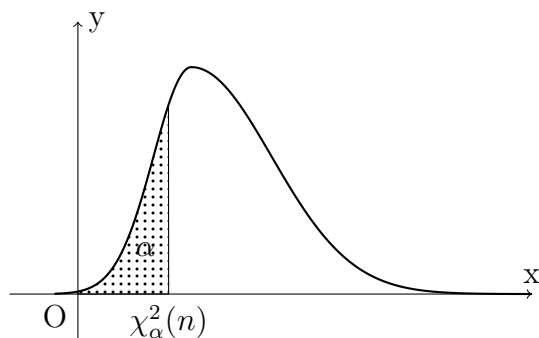
Nếu n là một số nguyên lẻ thì $\Gamma\left(\frac{n}{2}\right) = \left(\frac{n}{2}-1\right) \cdot \left(\frac{n}{2}-2\right) \dots \frac{3}{2} \cdot \frac{1}{2} \cdot \sqrt{\pi}$.

Có thể chứng minh được rằng nếu đại lượng ngẫu nhiên X phân phối theo quy luật khi - bình phương với n bậc tự do thì kỳ vọng toán: $M(X) = n$ và phương sai: $D(X) = 2n$.

Phân vị khi - bình phương mức α , ký hiệu là $\chi_\alpha^2(n)$ là giá trị của đại lượng ngẫu nhiên X tuân theo quy luật phân phối khi - bình phương với n bậc tự do thoả mãn điều kiện:

$$P[X < \chi_\alpha^2(n)] = \alpha$$

Giá trị của các phân vị này có thể tra ở phần phụ lục (xem Phụ lục C.5.), ý nghĩa của phân vị được thể hiện ở hình vẽ sau.



Hình 3.8:

Nếu có n đại lượng ngẫu nhiên độc lập $X_i, (i = \overline{1, n})$, cùng phân phối theo quy luật chuẩn hóa $N(0, 1)$ thì tổng bình phương các đại lượng đó:

$$X = \sum_{i=1}^n X_i^2$$

là đại lượng ngẫu nhiên phân phối theo quy luật khi - bình phương với n bậc tự do; còn nếu n đại lượng đó liên hệ với nhau bằng một hệ thức tuyến tính thì tổng bình phương X tuân theo quy luật khi - bình phương với $n - 1$ bậc tự do.

Khi số bậc tự do n tăng lên, quy luật khi - bình phương xấp xỉ với quy luật chuẩn.

3.4.7. Quy luật Student $T(n)$

★ **Định nghĩa 3.19** Đại lượng ngẫu nhiên liên tục X được gọi là tuân theo quy luật phân phối Student với n bậc tự do, ký hiệu $T(n)$, nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\Gamma\left(\frac{n-1}{2}\right)} \left[1 + \frac{x^2}{n-1}\right]^{-\frac{n}{2}}$$

trong đó $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ là hàm Gamma.

Đồ thị của hàm mật độ phân phối Student $T(n)$ có dạng như sau:

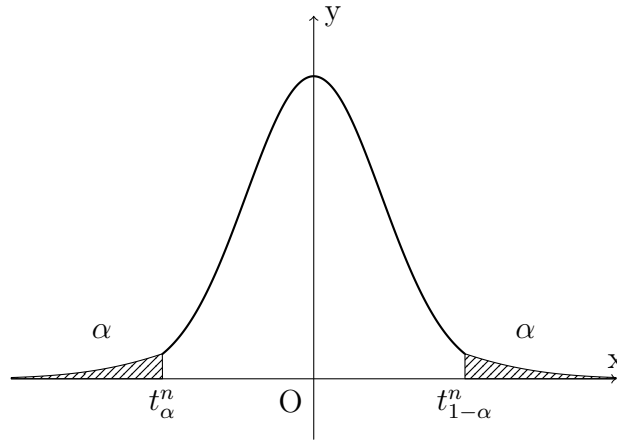
Có thể chứng minh được rằng nếu đại lượng ngẫu nhiên X tuân theo quy luật phân phối Student với n bậc tự do thì kỳ vọng toán: $M(X) = 0$ và phương sai: $D(X) = \frac{n}{n-2}$.

Phân vị Student mức α , ký hiệu là t_{α}^n là giá trị của đại lượng ngẫu nhiên X phân phối theo quy luật Student với n bậc tự do thoả mãn điều kiện:

$$P(X < t_{\alpha}^n) = \alpha$$

Giá trị của phân vị Student được cho trong phần phụ lục (xem Phụ lục C.4.), ý nghĩa của nó được thể hiện trên hình vẽ. Dễ thấy:

$$t_{\alpha}^n = -t_{1-\alpha}^n$$



Hình 3.9:

Giả sử U là đại lượng ngẫu nhiên phân phối chuẩn hoá $N(0, 1)$ và V là đại lượng ngẫu nhiên độc lập với U , phân phối theo quy luật khi - bình phương với n bậc tự do $\chi^2(n)$. Khi đó đại lượng ngẫu nhiên:

$$X = \frac{U}{\sqrt{\frac{V}{n}}}$$

sẽ phân phối theo quy luật Student với n bậc tự do $T(n)$.

Khi số bậc tự do n tăng lên, phân phối Student sẽ hội tụ nhanh về phân phối chuẩn hoá $N(0, 1)$. Do đó nếu n khá lớn ($n > 30$), có thể dùng phân phối chuẩn hoá thay thế cho phân phối Student.

Bài tập chương 3

Bài 3.1. Một thiết bị gồm 3 bộ phận hoạt động độc lập với nhau. Xác suất trong thời gian t các bộ phận bị hỏng lần lượt là 0,4; 0,2 và 0,3.

- Tìm quy luật phân phối xác suất của số bộ phận bị hỏng X .
- Thiết lập hàm phân phối xác suất của X .
- Tính xác suất trong thời gian t có không quá 2 bộ phận hỏng.
- Tìm $\text{mod}(X)$.

ĐS: a)

X	0	1	2	3
P	0,336	0,452	0,188	0,024

c) 0,976

d) 1

Bài 3.2. Một hộp có 3 quả cầu trắng và 2 quả cầu đen. Lấy ngẫu nhiên lần lượt từng quả cầu cho đến khi lấy được cầu trắng. Tìm quy luật phân phối xác suất của số cầu được lấy ra.

ĐS: **Bài 3.3.** Có hai lô sản phẩm, lô 1 có 8 chính phẩm và 2 phế phẩm; lô 2 có 7 chính phẩm và

X	1	2	3
P	0,6	0,3	0,1

3 phế phẩm. Từ lô thứ nhất lấy ngẫu nhiên ra 2 sản phẩm bỏ sang lô thứ hai, sau đó từ lô thứ 2 lấy ra 2 sản phẩm.

- Tìm quy luật phân phối xác suất của số chính phẩm được lấy ra.
- Xây dựng hàm phân phối xác suất của số chính phẩm được lấy ra.

ĐS:

X	0	1	2
P	0,0639	0,4388	0,4973

Bài 3.4. Bảng phân phối xác suất của biến ngẫu nhiên X như sau:

X	-5	2	3	4
P	0,4	0,3	0,1	0,2

- Tính $M(X)$; $D(X)$ và σ_X .
- Tìm $\text{mod}(X)$.

ĐS: a) 0,3; 15,21; 3,9; b) -5.

Bài 3.5. Cho X_1 và X_2 là hai đại lượng ngẫu nhiên độc lập có bảng phân phối xác suất như sau:

X_1	2	3	5
P	0,3	0,5	0,2

X_2	1	4
P	0,2	0,8

- Tính $M(X_1)$, $M(X_2)$, $D(X_1)$, $D(X_2)$;
- Tính $M(X_1 + X_2)$, $D(X_1 + X_2)$
- Tính $M(2X_1 - 3X_2)$, $D(2X_1 - 3X_2)$, $M(X_1^2 - X_2^3)$.

ĐS: a) 3,1; 3,4; 1,09; 1,44; b) 6,5; 2,53; c) -4; 17,32; -40,7.

Bài 3.6. Đại lượng ngẫu nhiên X có kỳ vọng toán $M(X) = a$ và $D(X) = b^2$.

a) Tính $M(X - M(X))$;

b) Tìm kỳ vọng toán và phương sai của đại lượng ngẫu nhiên $U = \frac{X - a}{b}$.

ĐS: a) 0; b) 0; 1.

Bài 3.7. Thực hiện ba lần bắn với xác suất lần bắn thứ nhất, thứ hai và thứ ba trúng đích lần lượt là $p_1 = 0,4; p_2 = 0,3; p_3 = 0,6$. Tìm kỳ vọng toán của số lần bắn trúng đích.

ĐS: 1,3.

Bài 3.8. Đại lượng ngẫu nhiên rời rạc X nhận 3 giá trị có thể có là $x_1 = 4$ với xác suất $p_1 = 0,5; X_2 = 0,6$ với xác suất $p_2 = 0,3$ và x_3 với xác suất p_3 . Tìm x_3 và p_3 biết $M(X) = 8$.

ĐS: 29,1; 0,2.

Bài 3.9. Đại lượng ngẫu nhiên rời rạc X nhận các giá trị có thể có là $x_1 = -1, x_2 = 0, x_3 = 1$. Tìm các xác suất tương ứng p_1, p_2 và p_3 biết rằng $M(X) = 0,1$ và $D(X) = 0,89$.

ĐS: 0,4; 0,1; 0,5.

Bài 3.10. Một người đi từ nhà đến cơ quan phải qua 3 ngã tư, xác suất để người đó gặp đèn đỏ ở các ngã tư tương ứng là 0,2; 0,4; 0,5. Hỏi thời gian trung bình phải ngừng trên đường là bao nhiêu, biết rằng mỗi khi gặp đèn đỏ người ấy phải đợi khoảng 2 phút.

ĐS: 2,2.

Bài 3.11. Đại lượng ngẫu nhiên X có bảng phân phối xác suất như sau:

X	1	4	8
P	0,3	0,1	0,6

a) Tính $P(|X - M(X)| < 4)$

b) Tính xác suất để trong 5 phép thử có ít nhất một lần $X < M(X)$.

ĐS: 0,7; 0,92224.

Bài 3.12. Cho X, Y là hai đại lượng ngẫu nhiên độc lập có bảng phân phối xác suất:

X	-1	0	2
P	x	0,2	y
Y	-1	1	
P	z	t	

a) Tìm x, y, z, t biết $M(X + Y) = 1,3; M(X^2 + Y^2) = 3,3$;

b) Lập bảng phân phối xác suất của $X^2 - 2X, XY, |X - Y|$;

c) Tìm số $p \in (0; 1)$ sao cho đại lượng ngẫu nhiên $pX^2 + (1 - p)Y^3$ có độ lệch tiêu chuẩn nhỏ nhất. Tìm độ lệch tiêu chuẩn nhỏ nhất đó.

ĐS: a) 0,3; 0,5; 0,2; 0,8; c) $p = 0,175; \sigma_{min} = 0,726$.

Bài 3.13. Cho đại lượng ngẫu nhiên rời rạc X nhận các giá trị $0, 1, 2, 3, \dots$ với xác suất tương ứng:

$$P[X = n] = \frac{3}{4}q^n$$

a) Tìm q ;

b) Tính xác suất trong 3 phép thử độc lập có lần $X < 3$;

c) Tính kỳ vọng toán $M(X)$ và phương sai $D(X)$.

ĐS: a) 0,25; b) $1 - 0,25^3$; c) $\frac{1}{3}; \frac{4}{9}$

Bài 3.14. Nhu cầu hàng năm về loại hàng A là đại lượng ngẫu nhiên X có hàm mật độ xác suất như sau (đơn vị: ngàn sản phẩm):

$$f(x) = \begin{cases} k(30 - x) & \text{với } x \in (0; 30) \\ 0 & \text{với } x \notin (0; 30) \end{cases}$$

- a) Tìm hệ số k,
 b) Tìm xác suất để nhu cầu về loại hàng đó không vượt quá 12 ngàn sản phẩm trong một năm.
 c) Tìm nhu cầu trung bình hàng năm về loại hàng đó.

ĐS: a) $\frac{1}{450}$; b) 0,64; c) 10.

Bài 3.15. Thời gian xếp hàng chờ mua hàng của khách là đại lượng ngẫu nhiên liên tục với hàm phân phối xác suất như sau: (đơn vị: phút)

$$F(x) = \begin{cases} 0 & \text{với } x \leq 0 \\ ax^3 - 3x^2 + 2x & \text{với } 0 < x \leq 1 \\ 1 & \text{với } x > 1 \end{cases}$$

- a) Tìm hệ số a.
 b) Tìm thời gian xếp hàng trung bình.
 c) Tìm xác suất để trong 3 người xếp hàng thì có không quá 2 người phải chờ quá 0,5 phút.
 ĐS: a) 2; b) 0,5; c) 0,875.

Bài 3.16. Cho X là đại lượng ngẫu nhiên liên tục có hàm phân phối xác suất sau đây:

$$F(x) = a + b \cdot \arctan x$$

- a) Tìm các hệ số a, b.
 b) Tìm $P(0 < X < l)$.
 c) Tìm giá trị x_1 thỏa mãn điều kiện $P(X > x_1) = \frac{3}{4}$.
 d) Tìm hàm mật độ xác suất của X.

ĐS: a) $\frac{1}{2}$; $\frac{1}{\pi}$; b) 0,25; c) -1; d) $\frac{1}{\pi(1+x^2)}$

Bài 3.17. Hàm phân phối xác suất của đại lượng ngẫu nhiên liên tục X có dạng:

$$F(x) = \begin{cases} 0 & \text{với } x \leq 0 \\ k \cdot \sin x & \text{với } 0 < x \leq \frac{\pi}{4} \\ 1 & \text{với } x > \frac{\pi}{4} \end{cases}$$

- a) Tìm hệ số a.
 b) Tìm hàm mật độ f(x).
 c) Tìm $M(X), D(X)$.
 ĐS: a) $\sqrt{2}$; c) 0,371; 0,0499.

Bài 3.18. Đại lượng ngẫu nhiên X có hàm mật độ xác suất:

$$f(x) = \begin{cases} a \cdot \cos x & \text{với } x \in \left(\frac{\pi}{2}, \frac{\pi}{2}\right) \\ 0 & \text{với } x \notin \left(\frac{\pi}{2}, \frac{\pi}{2}\right) \end{cases}$$

- a) Tìm hệ số a,
 b) Tìm $P(0 \leq X < \frac{\pi}{4})$.
 c) Tìm xác suất để trong 3 phép thử có lần $X \in \left[\frac{\pi}{3}, \frac{2\pi}{3}\right]$.

ĐS: a) 0,5; b) 0,354; c) 0,1878.

Bài 3.19. Thu nhập của dân cư một vùng là đại lượng ngẫu nhiên liên tục có hàm phân phối xác suất như sau:

$$F(x) = \begin{cases} 1 - \left(\frac{x_0}{x}\right)^\alpha & \text{với } x \geq x_0 \\ 0 & \text{với } x < x_0 \end{cases}$$

Hãy xác định mức thu nhập sao cho lấy ngẫu nhiên một người ở vùng đó thì thu nhập của người này vượt quá mức trên với xác suất 0,5.

ĐS: $\frac{1}{2\alpha} x_0$

Bài 3.20. Tỷ lệ mắc một loại bệnh trong một vùng dân cư là biến ngẫu nhiên liên tục có hàm mật độ:

$$f(x) = \begin{cases} \frac{1}{20} & \text{với } x \in (5; 25) \\ 0 & \text{với } x \notin (5; 25) \end{cases}$$

- a) Tính $P(|X - 10| > 2, 5)$.
 b) Tính tỷ lệ mắc bệnh trung bình và phương sai.

ĐS: a) 0,75; b) 15; 33,3.

Bài 3.21. Đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x) = \frac{k}{e^x + e^{-x}} \quad (-\infty < x < \infty)$$

ĐS: $\frac{2}{\pi}$

Bài 3.22. Hai kiện tướng bóng bàn ngang sức thi đấu với nhau. Hỏi thắng 2 trong 4 ván dễ hơn hay thắng 3 trong 6 ván dễ hơn?

ĐS: Thắng 2 trong 4 ván dễ hơn.

Bài 3.23. Một nữ công nhân quản lý 12 máy dệt. Xác suất để mỗi máy trong khoảng thời gian t cần đến sự chăm sóc của nữ công nhân là 0,3.

- a) Tính xác suất trong khoảng thời gian t có 4 máy cần đến sự chăm sóc của nữ công nhân.
 b) Tính xác suất trong khoảng thời gian t có từ 3 đến 6 máy cần đến sự chăm sóc của nữ công nhân.
 c) Tính số máy có khả năng cần đến sự chăm sóc của nữ công nhân nhiều nhất.

ĐS: a) 0,231; b) 0,709; c) 4 máy.

Bài 3.24. Việc sản xuất ra các sản phẩm được tiến hành độc lập. Hỏi phải sản xuất mỗi đợt bao nhiêu sản phẩm để trung bình có được 10 sản phẩm đạt tiêu chuẩn biết rằng xác suất được sản phẩm đạt tiêu chuẩn bằng 0,8.

ĐS: 13 sản phẩm.

Bài 3.25. Tiến hành kiểm tra chất lượng 900 chi tiết, xác suất để một chi tiết đạt tiêu chuẩn là 0,9.

- a) Tính xác suất để có 810 chi tiết đạt tiêu chuẩn.
 b) Tính xác suất để có từ 800 đến 825 chi tiết đạt tiêu chuẩn.
 c) Hãy tìm với xác suất 0,9544 xem số sản phẩm đạt tiêu chuẩn nằm trong khoảng đối xứng nào xung quanh giá trị trung bình?

ĐS: a) 0,0443; b) 0,81904; c) (792; 828).

Bài 3.26. Kích thước chi tiết do một máy sản xuất ra là đại lượng ngẫu nhiên phân phối chuẩn với $\mu = 5\text{cm}$ và $\sigma_X = 8,1\text{mm}$. a) Tìm xác suất để lấy ngẫu nhiên 1 chi tiết có kích thước từ 4cm đến 7cm.

b) Tìm xác suất để lấy ngẫu nhiên 1 chi tiết có kích thước sai lệch so với kỳ vọng toán không quá 2cm.

c) Tính xác suất để khi lấy ngẫu nhiên 5 chi tiết có 2 chi tiết có kích thước lớn hơn 5cm.

ĐS: a) 0,88389; b) 0,98648; c) 0,3125.

Bài 3.27. Kích thước chi tiết là biến ngẫu nhiên phân phối chuẩn với $\mu = 50\text{cm}$. Kích thước thực tế của các chi tiết không nhỏ hơn 32cm và không lớn hơn 67cm. Tìm xác suất để lấy ngẫu nhiên 2 chi tiết có chi tiết có kích thước nhỏ hơn 40cm.

ĐS: 0,09.

Bài 3.28. Máy bay bay dọc theo cầu dài 8m, rộng 4m và ném 2 quả bom. Biết rằng khoảng cách từ điểm rơi của quả bom đến trục đối xứng theo chiều dọc và chiều ngang của cầu là các biến ngẫu nhiên độc lập, cùng phân phối chuẩn với kỳ vọng toán bằng 0 và độ lệch chuẩn tương ứng là 4m và 6m. Tìm xác suất để cầu bị trúng bom.

ĐS: 0,3439.

Bài 3.29. Khi thâm nhập một thị trường mới, doanh nghiệp chỉ dự kiến rằng doanh số hàng tháng có thể đạt được phân phối chuẩn. Xác suất để đạt được doanh số nhỏ hơn 20 triệu là 0,1; trên 50 triệu là 0,2.

a) Tính doanh số hàng tháng trung bình của doanh nghiệp đó.

b) Tính xác suất để doanh nghiệp đạt được doanh số ít nhất là 40 triệu hàng tháng.

ĐS: a) 38,113; b) 0,44828

Bài 3.30. Trọng lượng một loại sản phẩm là biến ngẫu nhiên phân phối chuẩn với trung bình 25kg và độ lệch chuẩn 0,2kg.

a) Tính tỷ lệ sản phẩm có trọng lượng lớn hơn 25,3kg.

b) Tính xác suất để khi cân thử 25 sản phẩm thì trọng lượng trung bình của 25 sản phẩm này nhỏ hơn 25,01kg.

ĐS: a) 0,06681; b) 0,09871.

Bài 3.31. Theo thống kê dân số thì một người ở độ tuổi 40 sẽ sống thêm 1 năm nữa là 0,995. Một công ty bảo hiểm nhân thọ bán bảo hiểm 1 năm cho những người ở độ tuổi đó với giá là 10 ngàn và trong trường hợp người mua bảo hiểm bị chết thì số tiền bồi thường là 1 triệu. Hỏi lợi nhuận trung bình của công ty khi bán mỗi thẻ bảo hiểm loại này là bao nhiêu?

ĐS: 5 ngàn.

Bài 3.32. Nhu cầu về một loại thực phẩm ở một khu dân cư có bảng phân phối xác suất như sau:

Nhu cầu (kg)	50	60	70	80	90	100
Xác suất	0,15	0,25	0,35	0,12	0,07	0,06

a) Tính nhu cầu trung bình về loại thực phẩm đó ở khu dân cư nói trên;

b) Mỗi kg thực phẩm cửa hàng mua vào 45000đ, bán ra 60000đ. Giả sử cửa hàng nhập 70kg thực phẩm, hãy lập bảng phân phối xác suất của lượng sản phẩm được bán ra. Tính số tiền lãi trung bình của hàng thu được khi đó;

c) Nếu bị ế thì cửa hàng giảm giá bán ra chỉ 30000đ thì sẽ hết hàng. Hỏi cửa hàng nên nhập 70kg hay 80kg thì tốt hơn?

ĐS: a) 68,9; b) 720000; c) nên nhập 70kg.

Bài 3.33. Một chi tiết do máy tiện sản xuất ra là một đại lượng ngẫu nhiên $X \sim N(20mm, 4mm^2)$. Chi tiết được gọi là hợp quy cách nếu X dao động từ 18mm đến 22mm. Cho máy đó sản xuất 100 chi tiết, tính xác suất để có:

a) 50 chi tiết máy hợp quy cách;

b) Không quá 80 chi tiết máy hợp quy cách.

ĐS: a) 0,105; b) 0,994

Bài 3.34. Thời gian bảo hành một sản phẩm được quy định là 3 năm. Nếu bán được 1 sản phẩm thì cửa hàng lãi 150 ngàn song nếu sản phẩm bị hỏng trong thời gian bảo hành thì cửa hàng phải chi phí 500 ngàn cho việc bảo hành. Biết rằng tuổi thọ sản phẩm là đại lượng ngẫu nhiên phân phối chuẩn với tuổi thọ trung bình là 4,2 năm và độ lệch chuẩn là 1,8 năm.

a) Tìm số tiền lãi mà cửa hàng hy vọng thu được khi bán mỗi sản phẩm.

b) Nếu muốn số tiền lãi cho mỗi sản phẩm bán ra là 50 ngàn thì phải quy định thời gian bảo hành là bao nhiêu?

c) Nếu muốn tỷ lệ sản phẩm phải bảo hành không quá 10% thì phải quy định thời gian bảo hành là bao nhiêu?

ĐS: a) 24,285 ngàn; b) 2,688 năm; c) 1,896 năm.

Bài 3.35. Một người cân nhắc giữa việc mua cổ phiếu của 2 công ty A và B hoạt động trong 2 lĩnh vực độc lập nhau. Biết lãi suất cổ phiếu của 2 công ty là đại lượng ngẫu nhiên phân phối chuẩn với các tham số đặc trưng như sau:

	Kỳ vọng toán (%)	Độ lệch chuẩn (%)
Công ty A	11	4
Công ty B	10,4	2,6

a) Nếu người đó muốn đạt được lãi suất tối thiểu là 10% thì nên mua cổ phiếu của công ty nào?

b) Nếu người đó muốn hạn chế rủi ro bằng cách mua cổ phiếu của cả hai công ty thì nên mua theo tỷ lệ bao nhiêu để mức độ rủi ro về lãi suất là nhỏ nhất?

ĐS: a) Công ty A; b) A:B = 0,297:0,703.

Chương 4

Mẫu ngẫu nhiên - Ước lượng tham số

4.1. Tổng thể nghiên cứu

4.1.1. Định nghĩa

★ **Định nghĩa 4.1** Tập hợp toàn bộ các phần tử mang cùng một dấu hiệu nghiên cứu định tính hoặc định lượng nào đó được gọi là tổng thể nghiên cứu hay tổng thể.

Số lượng các phần tử của tổng thể được gọi là kích thước tổng thể, ký hiệu là N . Với mỗi tổng thể ta không nghiên cứu trực tiếp tổng thể mà thông qua một hay nhiều dấu hiệu đặc trưng cho tổng thể đó. Chúng được gọi là dấu hiệu nghiên cứu, ký hiệu là χ , các dấu hiệu này có thể là định tính hoặc định lượng.

• **Ví dụ 4.1** Muốn nghiên cứu tình hình đi làm thêm của sinh viên các trường Cao đẳng, Đại học trên toàn quốc ta tiến hành tổng điều tra sinh viên toàn quốc và phân tích số liệu theo số giờ làm thêm hàng tuần. Khi đó, tổng thể nghiên cứu là tập hợp tất cả sinh viên toàn quốc, dấu hiệu nghiên cứu X là thời gian làm thêm hàng tuần của sinh viên. Kích thước tổng thể N là số sinh viên của Việt Nam.

4.1.2. Các phương pháp mô tả tổng thể

Giả sử tổng thể kích thước N , dấu hiệu nghiên cứu định lượng χ nhận các giá trị x_1, x_2, \dots, x_k với các tần số tương ứng N_1, N_2, \dots, N_k (trong tổng thể có N_i từ mà dấu hiệu nghiên cứu χ nhận giá trị bằng x_i).

a) Bảng phân phối tần số.

Tổng thể có thể được mô tả bằng bảng phân phối tần số như sau:

Giá trị của χ	x_1	x_2	...	x_i	...	x_k
Tần số	N_1	N_2	...	N_i	...	N_k

Hiển nhiên:

$$\begin{cases} 0 \leq N_i \leq N, & \forall_i = \overline{1, k} \\ \sum_{i=1}^k N_i = N \end{cases}$$

b) Bảng phân phối tần suất.

Ký hiệu p_i , ($i = \overline{1, k}$) là tần suất của x_i :

$$p_i = \frac{N_i}{N} \quad i = \overline{1, k}$$

khi đó tổng thể còn được mô tả bởi bảng phân phối tần suất như sau:

Giá trị của χ	x_1	x_2	...	x_i	...	x_k
Tần số	p_1	p_2	...	p_i	...	p_k

Để thấy:

$$\begin{cases} 0 \leq p_i \leq 1 & \forall_i = \overline{1, k} \\ \sum_{i=1}^k p_i = 1 \end{cases}$$

c) Tần số tích lũy - Tần suất tích lũy.

Ký hiệu w_i ($i = \overline{1, k}$) là tần số tích lũy của x_i , tức là tổng số phần tử có giá trị nhỏ hơn x_i :

$$w_i = \sum_{x_j < x_i} = N_j$$

và $F(x_i)$ ($i = \overline{1, k}$) là tần số tích lũy của x_i :

$$F(x_i) = \frac{w_i}{N} = \sum_{x_j < x_i} p_i.$$

Việc mô tả tổng thể theo bảng phân phối tần suất và tần suất tích lũy cho thấy dấu hiệu định lượng χ có thể mô hình hóa bằng một đại lượng ngẫu nhiên rời rạc X . Điều này cũng đúng đối với tổng thể mang dấu hiệu nghiên cứu χ phân phối liên tục. Đại lượng ngẫu nhiên X dùng để mô hình hóa dấu hiệu nghiên cứu χ được gọi đại lượng ngẫu nhiên gốc, còn quy luật phân phối xác suất của nó là quy luật phân phối gốc.

4.1.3. Các tham số đặc trưng của tổng thể

a) Trung bình tổng thể.

Giả sử trong tổng thể kích thước N dấu hiệu định lượng χ nhận các giá trị x_1, x_2, \dots, x_N . Trung bình tổng thể, ký hiệu là m , là trung bình số học của các giá trị của χ trong tổng thể:

$$m = \frac{1}{N} \sum_{i=1}^N x_i.$$

Nếu tổng thể được mô tả bởi bảng phân phối tần số ở phần (3.1.2 a) thì trung bình tổng thể còn được tính bởi:

$$m = \frac{1}{N} \sum_{i=1}^k x_i N_i.$$

Nếu xem dấu hiệu nghiên cứu χ như đại lượng ngẫu nhiên thì trung bình tổng thể chính là kỳ vọng toán của đại lượng ngẫu nhiên đó.

b) Phương sai tổng thể.

Giả sử trong tổng thể kích thước N dấu hiệu định lượng χ nhận các giá trị x_1, x_2, \dots, x_N . Phương sai tổng thể, ký hiệu là σ^2 , là trung bình số học của bình phương các sai lệch giữa các giá trị của dấu hiệu χ trong tổng thể và trung bình tổng thể:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - m)^2.$$

Nếu tổng thể được mô tả bởi bảng phân phối tần số ở phần (3.1.2 a) thì phương sai tổng thể còn được tính bởi:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - m)^2 N_i.$$

Độ lệch tiêu chuẩn tổng thể là căn bậc hai của phương sai tổng thể:

$$\sigma = \sqrt{\sigma^2}.$$

Thực chất phương sai tổng thể chính là phương sai của dấu hiệu nghiên cứu trong tổng thể. Tương tự công thức tính phương sai của đại lượng ngẫu nhiên, ta toàn có thể tính phương sai tổng thể bằng công thức:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 - \left(\frac{1}{N} \sum_{i=1}^N x_i \right)^2 \quad \text{hay} \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^k x_i^2 N_i - \left(\frac{1}{N} \sum_{i=1}^k x_i N_i \right)^2.$$

c) Tần suất tổng thể.

Giả sử trong tổng thể kích thước N có M phần tử mang dấu hiệu nghiên cứu, còn $N - M$ phần tử không mang dấu hiệu đó. Khi đó tần suất tổng thể, ký hiệu p , là tỷ lệ phần tử mang dấu hiệu nghiên cứu trong tổng thể:

$$p = \frac{M}{N}.$$

Nếu xem dấu hiệu nghiên cứu χ là đại lượng ngẫu nhiên tuân theo quy luật không - một thì tần suất tổng thể p chính là kỳ vọng toán của đại lượng đó.

4.2. Mẫu ngẫu nhiên

4.2.1. Định nghĩa

Ta đã biết có thể coi dấu hiệu nghiên cứu χ của tổng thể như một đại lượng ngẫu nhiên χ có hàm phân phối xác suất là tần suất tích lũy của χ . Giả sử từ tổng thể lấy ngẫu nhiên n phần tử, và gọi X_i là giá trị của χ đo lường được trên phần tử thứ i . Khi đó có thể xem như đã thực hiện n phép thử độc lập đối với X và X_i là các đại lượng ngẫu nhiên độc lập cùng quy luật phân phối xác suất với X .

★ **Định nghĩa 4.2** Mẫu ngẫu nhiên kích thước n là tập hợp của n đại lượng ngẫu nhiên độc lập X_1, X_2, \dots, X_n được thành lập từ đại lượng ngẫu nhiên X và có cùng quy luật phân phối xác suất đối với X .

Mẫu ngẫu nhiên thường được ký hiệu là $W = (X_1, X_2, \dots, X_n)$.

Khi thực hiện một phép thử đối với mẫu ngẫu nhiên W ta sẽ thu được kết quả cụ thể: X_1 nhận giá trị x_1 , X_2 nhận giá trị x_2 , ..., X_n nhận giá trị x_n . Tập hợp n giá trị x_1, x_2, \dots, x_n tạo thành một giá trị của mẫu ngẫu nhiên, hay còn gọi là một mẫu cụ thể, ký hiệu là:

$$w = (x_1, x_2, \dots, x_n).$$

● **Ví dụ 4.2** Trong 1 cửa hàng bán đĩa, trên 1 kệ chứa các đĩa nhạc có các loại đĩa tương ứng với số tiền như sau:

Giá (ngàn đồng)	20	25	30	34	40
Số đĩa	35	10	25	17	13

Lấy ngẫu nhiên 1 đĩa nhạc trong giá.

Gọi X là giá tiền của đĩa nhạc này. $X = \{20; 25; 30; 34; 40\}$

Ta thấy X có quy luật phân phối xác suất như sau:

X	20	25	30	34	40
P	0,35	0,1	0,25	0,17	0,13

X đặc trưng cho dấu hiệu nghiên cứu là giá tiền của tổng thể là số đĩa trên giá.

Lấy ngẫu nhiên (có hoàn lại) 4 đĩa nhạc từ kệ.

Gọi X_i là giá tiền của đĩa nhạc thứ i lấy được, $i = 1 \div 4$

Ta thấy các X_i độc lập và có cùng quy luật phân phối xác suất giống như X .

Lập $W = (X_1, X_2, X_3, X_4)$, gọi là mẫu ngẫu nhiên.

Bây giờ ta xét xem giá cụ thể của từng đĩa lấy ra thấy như sau:

Đĩa 1: giá 20 ngàn đồng.

Đĩa 2: giá 30 ngàn đồng.

Đĩa 3: giá 20 ngàn đồng.

Đĩa 4: giá 40 ngàn đồng.

Lập $w = (x_1, x_2, x_3, x_4) = (20, 30, 20, 40)$ gọi là mẫu cụ thể.

4.2.2. Các phương pháp mô tả mẫu ngẫu nhiên

Giả sử từ tổng thể với đại lượng ngẫu nhiên gốc X rút ra một mẫu ngẫu nhiên kích thước n với mẫu ngẫu nhiên cụ thể: $w = (x_1, x_2, \dots, x_n)$. Giả sử trong mẫu cụ thể đó, giá trị x_1 xuất hiện n_1 lần, giá trị x_2 xuất hiện n_2 lần, ..., giá trị x_k xuất hiện n_k lần, và các x_i sắp xếp theo giá trị tăng dần. Khi đó giá trị mẫu cụ thể w có thể được mô tả như sau:

a) Bảng phân phối tần số thực nghiệm (đơn).

x_i	x_1	x_2	...	x_i	...	x_k
n_i	n_1	n_2	...	n_i	...	n_k

Hiển nhiên:

$$\begin{cases} 0 \leq n_i \leq n, & \forall i = \overline{1, k} \\ \sum_{i=1}^k n_i = n \end{cases}$$

b) Bảng phân phối suất số thực nghiệm (đơn).

Ký hiệu f_i , ($i = \overline{1, k}$) là tần suất xuất hiện giá trị x_i trong mẫu:

$$f_i = \frac{n_i}{n} \quad i = \overline{1, k}$$

khi đó mẫu ngẫu nhiên cụ thể còn được mô tả bởi bảng phân phối tần suất như sau:

x_i	x_1	x_2	...	x_i	...	x_k
f_i	f_1	f_2	...	f_i	...	f_k

Dễ thấy:

$$\begin{cases} 0 \leq f_i \leq 1, & \forall i = \overline{1, k} \\ \sum_{i=1}^k f_i = 1 \end{cases}$$

c) Bảng phân phối tần số thực nghiệm (ghép lớp).

Trong trường hợp dấu hiệu nghiên cứu có phân phối liên tục thì mẫu ngẫu nhiên còn được mô tả bởi bảng phân phối tần số ghép lớp:

x_i	$x_0 - x_1$	$x_1 - x_2$...	$x_{i-1} - x_i$...	$x_{k-1} - x_k$
n_i	n_1	n_2	...	n_i	...	n_k

Trong bảng trên, có n_1 phần tử của mẫu w nhận giá trị trong khoảng $[x_0, x_1)$, n_2 phần tử nhận giá trị trong khoảng $[x_1, x_2)$, ..., có n_k phần tử nhận giá trị trong khoảng $[x_{k-1}, x_k)$. Hiển nhiên:

$$\begin{cases} 0 \leq n_i \leq n, & \forall i = \overline{1, k} \\ \sum_{i=1}^k n_i = n \end{cases}$$

Tương tự ta cũng có bảng phân phối tần suất ghép lớp như sau:

x_i	$x_0 - x_1$	$x_1 - x_2$...	$x_{i-1} - x_i$...	$x_{k-1} - x_k$
f_i	f_1	f_2	...	f_i	...	f_k

Tuy nhiên nếu chú ý rằng đại lượng ngẫu nhiên gốc X phân phối liên tục thì có thể coi trong mẫu ngẫu nhiên cụ thể, trường hợp dấu hiệu nghiên cứu nhận giá trị đúng bằng x_i không xảy ra.

d) Hàm phân phối thực nghiệm.

Ký hiệu w_i ($i = \overline{1, k}$) là tần số tích lũy của x_i , tức là tổng số phần tử có giá trị nhỏ hơn x_i trong mẫu cụ thể w :

$$w_i = \sum_{x_j < x_i} = n_j$$

và $F^*(x_i)$ ($i = \overline{1, k}$) là tần số tích lũy của x_i :

$$F^*(x_i) = \frac{w_i}{n} = \sum_{x_j < x_i} f_j.$$

Khi đó $F^*(x)$ là một hàm của x và được gọi là hàm phân phối thực nghiệm của mẫu.

4.2.3. Đồ thị của phân phối thực nghiệm

a) Đa giác tần số.

Xét mẫu ngẫu nhiên cụ thể được mô tả bởi các bảng phân phối tần số thực nghiệm và tần suất thực nghiệm ở phần (3.2.2. a,b).

- Đa giác tần số là một đường gấp khúc mà các đoạn thẳng của nó nối các điểm $(x_1, n_1), (x_2, n_2), \dots, (x_k, n_k)$ trên mặt phẳng.
- Đa giác tần suất là một đường gấp khúc mà các đoạn thẳng của nó nối các điểm $(x_1, f_1), (x_2, f_2), \dots, (x_k, f_k)$ trên mặt phẳng.

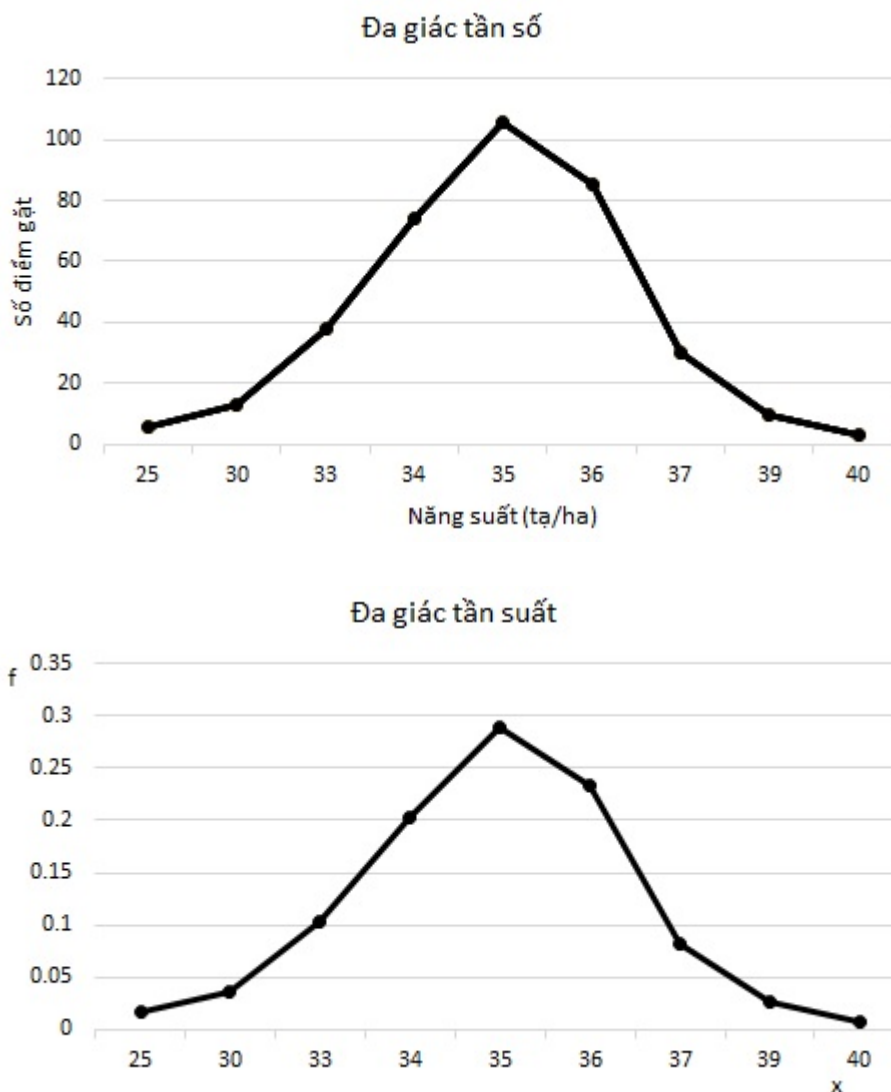
•**Ví dụ 4.3** Gặt ngẫu nhiên 365 điểm trồng lúa của một huyện ta được các số liệu được sắp xếp thành bảng sau:

Năng suất (tạ/ha)	25	30	33	34	35	36	37	39	40
Số điểm gặt	6	13	38	74	106	85	30	10	3

Như vậy giá trị của mẫu đã được mô tả dưới dạng bảng phân phối tần số thực nghiệm. Còn bảng phân phối tần suất thực nghiệm có dạng:

x_i	25	30	33	34	35	36	37	39	40
f_i	0,016	0,036	0,104	0,203	0,290	0,233	0,082	0,027	0,008

Đa giác tần số và đa giác tần suất của phân phối thực nghiệm này như sau:



b) Biểu đồ tần số.

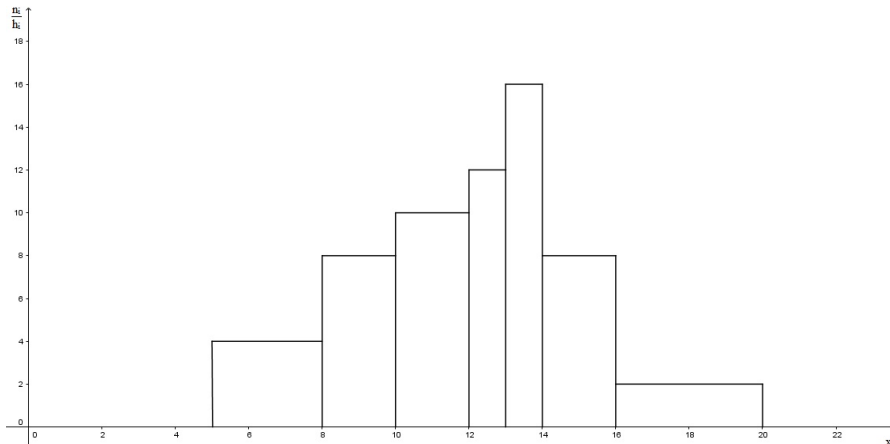
Xét mẫu ngẫu nhiên cụ thể được mô tả bởi các bảng phân phối tần số thực nghiệm và tần suất thực nghiệm ở phần (3.2.2. c).

Để vẽ biểu đồ tần số ta chia đoạn $[x_0, x_k]$ chứa tất cả các giá trị quan sát được của mẫu thành các đoạn nhỏ ứng với $[x_{i-1}, x_i]$. Biểu đồ tần suất là một bậc thang tạo nên bởi nhiều hình chữ nhật có đáy bằng $h_i = x_i - x_{i-1}$ và chiều cao bằng $\frac{n_i}{h_i}$. Lúc đó diện tích hình chữ nhật thứ i đúng bằng $h_i \cdot \frac{n_i}{h_i} = n_i$ là tần số ứng với đoạn thứ i . Vì vậy tổng diện tích của tất cả các hình chữ nhật sẽ bằng kích thước mẫu n .

Tương tự biểu đồ tần suất là một hình bậc thang được tạo nên bởi các hình chữ nhật có đáy bằng $h_i = x_i - x_{i-1}$ và chiều cao bằng $\frac{f_i}{h_i}$. Lúc đó diện tích hình chữ nhật thứ i bằng f_i và tổng diện tích cả biểu đồ bằng 1.

•**Ví dụ 4.4** Vẽ biểu đồ tần số của phân phối thực nghiệm cho ở bảng sau:

x_i	5 - 8	8 - 10	10 - 12	12 - 13	13 - 14	14 - 16	16 - 20
n_i	12	16	20	12	16	16	8
$\frac{n_i}{h_i}$	4	8	10	12	16	8	2



4.3. Thống kê

4.3.1. Định nghĩa

★ **Định nghĩa 4.3** Giả sử khi nghiên cứu đại lượng ngẫu nhiên gốc X trong tổng thể ta có mẫu ngẫu nhiên $W = (X_1, X_2, \dots, X_n)$. Thống kê là một hàm của các giá trị X_1, X_2, \dots, X_n , ký hiệu là G . Như vậy:

$$G = f(X_1, X_2, \dots, X_n).$$

Về thực chất thống kê là một hàm của các đại lượng ngẫu nhiên, do đó bản thân nó cũng sẽ là một đại lượng ngẫu nhiên tuân theo một quy luật phân phối xác suất nhất định và có các đặc trưng số tương ứng như trung bình $M(G)$, phương sai $D(G)$...

Khi mẫu ngẫu nhiên nhận một giá trị cụ thể $w = (x_1, x_2, \dots, x_n)$ thì thống kê G cũng nhận một giá trị cụ thể là $g = f(x_1, x_2, \dots, x_n)$.

4.3.2. Trung bình mẫu

a) Khái niệm. Giả sử từ đại lượng ngẫu nhiên gốc X trong tổng thể lập mẫu ngẫu nhiên kích thước $n : W = (X_1, X_2, \dots, X_n)$. Trung bình mẫu là một thống kê, ký hiệu là \bar{X} và là trung bình số học của các giá trị của mẫu:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

b) Tính chất

Nếu đại lượng ngẫu nhiên gốc X có kỳ vọng toán $M(X) = m$ và phương sai $D(X) = \sigma^2$ thì:

$$M(\bar{X}) = m; \quad D(\bar{X}) = \frac{\sigma^2}{n}.$$

Thật vậy, áp dụng các tính chất của kỳ vọng toán ta có:

$$M(\bar{X}) = M\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n M(X_i) = \frac{1}{n} \cdot n \cdot m = m.$$

Áp dụng các tính chất của phương sai với lưu ý X_i là các đại lượng ngẫu nhiên độc lập ta có:

$$D(\bar{X}) = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}.$$

c) Tính giá trị cụ thể.

Khi mẫu ngẫu nhiên nhận một giá trị cụ thể $w = (x_1, x_2, \dots, x_n)$ thì trung bình mẫu cũng nhận giá trị cụ thể bằng:

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

Nếu mẫu ngẫu nhiên được cho bởi mảng phân phối tần số:

x_i	x_1	x_2	...	x_i	...	x_k
n_i	n_1	n_2	...	n_i	...	n_k

 $\left(\sum_{i=1}^k n_i = n\right)$

thì giá trị cụ thể của trung bình mẫu được tính bởi:

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

4.3.3. Phương sai mẫu

a) Khái niệm. Giả sử từ đại lượng ngẫu nhiên gốc X trong tổng thể lập mẫu ngẫu nhiên $W = (X_1, X_2, \dots, X_n)$. Phương sai mẫu là một thống kê, ký hiệu S^2 , và là trung bình số học của bình phương các sai lệch giữa các giá trị của mẫu và trung bình mẫu:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Ngoài ra người ta còn dùng hai loại phương sai sau đây:

- Phương sai điều chỉnh mẫu, ký hiệu là S'^2 và được xác định bằng công thức:

$$S'^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

- Phương sai S^{*2} được xác định bằng biểu thức:

$$S^{*2} = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$$

trong đó m là trung bình tổng thể.

b) Tính chất.

- Phương sai mẫu có tính chất sau: Nếu đại lượng ngẫu nhiên gốc có kỳ vọng toán $M(X) = m$ và phương sai $D(X) = \sigma^2$ thì phương sai S^2 có kỳ vọng toán:

$$M(S^2) = \frac{n-1}{n} \sigma^2.$$

Thật vậy:

$$\begin{aligned}
 S^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n [(X_i - m) - (\bar{X} - m)]^2 \\
 &= \frac{1}{n} \sum_{i=1}^n \left[(X_i - m)^2 - 2(X_i - m)(\bar{X} - m) + (\bar{X} - m)^2 \right] \\
 &= \frac{1}{n} \sum_{i=1}^n (X_i - m)^2 - 2(\bar{X} - m) \frac{1}{n} \sum_{i=1}^n (X_i - m) + (\bar{X} - m)^2 \\
 &= \frac{1}{n} \sum_{i=1}^n (X_i - m)^2 - (\bar{X} - m)^2 \\
 \Rightarrow M(S)^2 &= \frac{1}{n} \sum_{i=1}^n M(X_i - m)^2 - M(\bar{X} - m)^2.
 \end{aligned}$$

Mặt khác:

$$\begin{aligned}
 M(X_i - m)^2 &= D(X_i) = D(X) = \sigma^2 \\
 M(\bar{X} - m)^2 &= D(\bar{X}) = \frac{\sigma^2}{n}.
 \end{aligned}$$

Do đó:

$$M(S^2) = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2.$$

- Giữa phương sai mẫu và phương sai điều chỉnh mẫu có mối liên hệ:

$$S'^2 = \frac{n}{n-1} S^2.$$

Do đó từ tính chất của phương sai mẫu S^2 có thể suy ra tính chất của S'^2 như sau:

$$M(S'^2) = \sigma^2.$$

Thật vậy:

$$M(S'^2) = M\left(\frac{n}{n-1} S^2\right) = \frac{n}{n-1} M(S^2) = \frac{n}{n-1} \frac{n-1}{n} \sigma^2 = \sigma^2.$$

- Đối với phương sai S^{*2} ta cũng có tính chất tương tự sau:

$$M(S^{*2}) = \sigma^2.$$

Thật vậy:

$$M(S^{*2}) = M\left[\frac{1}{n} \sum_{i=1}^n (X_i - m)^2\right] = \frac{1}{n} \sum_{i=1}^n M(X_i - m)^2 = \frac{1}{n} \sum_{i=1}^n D(X_i) = \sigma^2.$$

c) Tính giá trị cụ thể.

- Trên một mẫu cụ thể $w = (x_1, x_2, \dots, x_n)$, phương sai mẫu nhận một giá trị cụ thể bằng:

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

Nếu mẫu ngẫu nhiên được cho bởi bảng phân phối tần số:

x_i	x_1	x_2	...	x_i	...	x_k
n_i	n_1	n_2	...	n_i	...	n_k

 $\left(\sum_{i=1}^k n_i = n \right)$

thì giá trị cụ thể của phương sai mẫu được tính bởi:

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

Trên thực tế, để tiện cho việc tính toán người ta sử dụng công thức sau để tính giá trị của phương sai mẫu:

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

Thật vậy, ta có:

$$\begin{aligned} s^2 &= \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^k n_i (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\ &= \frac{1}{n} \sum_{i=1}^k n_i x_i^2 - 2\bar{x} \cdot \frac{1}{n} \sum_{i=1}^k n_i x_i + \bar{x}^2 = \frac{1}{n} \sum_{i=1}^k n_i x_i^2 - \bar{x}^2. \end{aligned}$$

- Giá trị cụ thể của phương sai điều chỉnh mẫu được tính bởi:

$$s'^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{hoặc} \quad s'^2 = \frac{1}{n-1} \sum_{i=1}^n n_i (x_i - \bar{x})^2.$$

Giá trị cụ thể s'^2 có thể tính thông qua giá trị s^2 như sau:

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

- Giá trị cụ thể của phương sai S^{*2} là:

$$s^{*2} = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 \quad \text{hoặc} \quad s^{*2} = \frac{1}{n} \sum_{i=1}^n n_i (x_i - m)^2.$$

4.3.4. Độ lệch tiêu chuẩn mẫu

Nếu lấy căn bậc hai của phương sai mẫu S^2 , của phương sai điều chỉnh mẫu S'^2 , của phương sai S^{*2} ta thu được các thống kê tương ứng gọi là độ lệch tiêu chuẩn mẫu S , độ lệch tiêu chuẩn điều chỉnh mẫu S' , độ lệch tiêu chuẩn S^* như sau:

$$\begin{aligned} S &= \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2} \\ S'^2 &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} \\ S^{*2} &= \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - m)^2}. \end{aligned}$$

Giá trị cụ thể của chúng trên một mẫu cụ thể là những số xác định, ký hiệu là s, s', s^* .

4.3.5. Tần suất mẫu

a) Khái niệm.

Giả sử từ tổng thể kích thước N , trong đó M phần tử mang dấu hiệu nghiên cứu, lấy ra một mẫu ngẫu nhiên kích thước n và trong đó thấy có Y phần tử mang dấu hiệu nghiên cứu. Lúc đó tần suất là một mẫu thống kê, ký hiệu là f và là tỷ số giữa số phần tử mang dấu hiệu nghiên cứu trong mẫu và kích thước mẫu:

$$f = \frac{Y}{n}.$$

Ta chú ý rằng, tần suất mẫu là một trường hợp riêng của trung bình mẫu \bar{X} khi xem dấu hiệu nghiên cứu trong tổng thể là đại lượng ngẫu nhiên X tuân theo quy luật không - một. Thật vậy, nếu gọi X_i là số phần tử mang dấu hiệu nghiên cứu khi lấy phần tử thứ i của mẫu thì ta có mẫu ngẫu nhiên $W = (X_1, X_2, \dots, X_n)$ với X_i tuân theo quy luật không - một và:

$$Y = X_1 + X_2 + \dots + X_n.$$

Khi đó:

$$f = \frac{X_1 + X_2 + \dots + X_n}{n} = \bar{X}.$$

b) Tính chất.

Nếu đại lượng ngẫu nhiên gốc X tuân theo quy luật $A(p)$ với kỳ vọng toán $M(X) = p$ và phương sai $D(X) = p(1-p)$ thì:

$$M(f) = p; \quad D(f) = \frac{p(1-p)}{n}.$$

Thật vậy, theo kết quả phần (3.2.2):

$$M(f) = M(\bar{X}) = M(X) = p$$

$$D(f) = D(\bar{X}) = \frac{D(X)}{n} = \frac{p(1-p)}{n}.$$

c) Tính giá trị cụ thể.

Giả sử trong mẫu ngẫu nhiên kích thước n có k phần tử mang dấu hiệu nghiên cứu, khi đó giá trị cụ thể của tần suất mẫu là:

$$f = \frac{k}{n}.$$

4.3.6. Quy luật phân phối xác suất của một số thống kê đặc trưng mẫu

Quy luật phân phối xác suất của các thống kê đặc trưng mẫu phụ thuộc chặt chẽ và quy luật phân phối xác suất của đại lượng ngẫu nhiên gốc X . Trong phần này ta chỉ dừng ở việc thừa nhận các kết quả. Bạn đọc có thể tìm thấy chứng minh các kết quả này ở các tài liệu đầy đủ hơn về thống kê toán.

a) Trường hợp đại lượng ngẫu nhiên gốc X tuân theo quy luật phân phối chuẩn

Giả sử dấu hiệu nghiên cứu trong tổng thể có thể xem như một đại lượng ngẫu nhiên X tuân theo quy luật chuẩn với kỳ vọng toán $M(X) = \mu$ và phương sai $D(X) = \sigma^2$.

Để xác định quy luật phân phối xác suất của các thống kê đặc trưng mẫu có thể sử dụng định lý sau:

◇ **Định lý 4.1** Nếu các đại lượng ngẫu nhiên X_1, X_2, \dots, X_n độc lập và cùng phân phối theo quy luật chuẩn thì mọi tổ hợp tuyến tính của các đại lượng ngẫu nhiên đó cũng phân phối theo quy luật chuẩn.

- Trung bình mẫu cũng phân phối theo quy luật chuẩn với kỳ vọng toán $M(\bar{X}) = \mu$ và phương sai $D(\bar{X}) = \frac{\sigma^2}{n}$.

Do đó nếu xây dựng thống kê

$$U = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{(\bar{X} - \mu) \sqrt{n}}{\sigma}$$

thì thống kê U sẽ phân phối theo quy luật chuẩn hoá $N(0, 1)$.

- Từ phương sai S^{*2} xét thống kê:

$$G = \frac{nS^{*2}}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2$$

khi đó thống kê G phân phối theo quy luật "Khi bình phương" với n bậc tự do $\chi^2(n)$.

- Từ phương sai S'^2 xây dựng thống kê:

$$X^2 = \frac{(n-1)S'^2}{\sigma^2}$$

thì thống kê X^2 phân phối theo quy luật "Khi bình phương" với $(n-1)$ bậc tự do $\chi^2(n-1)$.

- Từ thống kê U và χ^2 ở trên xây dựng tiếp thống kê:

$$T = \frac{U}{\sqrt{\frac{\chi^2}{n-1}}} = \frac{(\bar{X} - \mu) \sqrt{n}}{S'}$$

thì thống kê T sẽ phân phối theo quy luật Student với $(n-1)$ bậc tự do $T(n-1)$.

Chú ý rằng khi n khá lớn thì quy luật Student hội tụ về quy luật chuẩn hoá, do đó với $n > 30$ thực tế có thể xem thống kê T phân phối xấp xỉ $N(0, 1)$.

- b) Trường hợp đại lượng ngẫu nhiên gốc X tuân theo quy luật phân phối không - một
Giả sử dấu hiệu nghiên cứu trong tổng thể có thể xem như đại lượng ngẫu nhiên phân phối theo quy luật không - một với tham số $p(A(p))$.

- Trung bình mẫu chính là tần suất mẫu f với kỳ vọng toán $M(f) = p$ và phương sai $D(f) = \frac{p(1-p)}{n}$.

- Từ tần suất mẫu xây dựng thống kê:

$$U = \frac{f - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{(f - p) \sqrt{n}}{\sqrt{p(1-p)}}$$

thì thống kê U sẽ phân phối xấp xỉ chuẩn hoá $N(0, 1)$ khi n khá lớn.

- Nếu n đủ lớn thì từ thống kê U ở trên thay p bởi f và xây dựng thống kê:

$$U = \frac{(f - p) \sqrt{n}}{\sqrt{f(1-f)}}$$

thì lúc đó U cũng phân phối xấp xỉ $N(0, 1)$.

4.3.7. Ví dụ

•**Ví dụ 4.5** Nghiên cứu về mức tiêu thụ điện năng (kw) của một khu dân cư, ta nghiên cứu mức tiêu thụ điện năng của 5 hộ gia đình trong một khu dân cư đó thì thu được kết quả sau:
75; 123; 97; 145; 215

a. Tính trung bình số điện tiêu thụ của các hộ gia đình. Tìm phương sai, độ lệch tiêu chuẩn hiệu chỉnh của số điện tiêu thụ.

b. Nếu quy ước những hộ có mức tiêu thụ ít hơn hoặc bằng 123 kw thì được coi là tiêu thụ điện bình thường. Tìm tỷ lệ các hộ tiêu thụ điện bình thường.

Lời giải.

(bản kèm)

•**Ví dụ 4.6** Khi kiểm tra thể lực của một nhóm sinh viên, ta có kết quả về cân nặng như sau:

kg	45	50	55	60	65
Số sv	8	14	28	18	12

a. Tìm cân nặng trung bình của nhóm sinh viên trên. Phương sai, phương sai hiệu chỉnh, độ lệch tiêu chuẩn và độ lệch tiêu chuẩn hiệu chỉnh.

b. Tìm tỷ lệ sinh viên có cân nặng lớn hơn hoặc bằng 55 kg.

Lời giải.

(bản kèm)

•**Ví dụ 4.7** Khi kiểm tra thể lực của một nhóm sinh viên, ta có kết quả về cân nặng như sau:

kg	42,5-47,5	47,5-52,5	52,5-57,5	57,5-62,5	62,5-67,5
Số sv	8	14	28	18	12

a. Tìm cân nặng trung bình của nhóm sinh viên trên. Phương sai, phương sai hiệu chỉnh, độ lệch tiêu chuẩn và độ lệch tiêu chuẩn hiệu chỉnh.

b. Tìm tỷ lệ sinh viên có cân nặng lớn hơn hoặc bằng 52,5 kg.

Lời giải.

(bản kèm)

4.4. Mẫu ngẫu nhiên hai chiều

4.4.1. Khái niệm

Giả sử trên cùng một tổng thể phải nghiên cứu đồng thời hai dấu hiệu nghiên cứu, trong đó dấu hiệu thứ nhất có thể xem như đại lượng ngẫu nhiên X , còn dấu hiệu thứ hai xem như đại lượng ngẫu nhiên Y . Hệ hai đại lượng ngẫu nhiên X, Y làm thành đại lượng ngẫu nhiên hai chiều (X, Y) .

Từ tổng thể lấy ra mẫu kích thước n , tức là thực hiện n phép thử đối với đại lượng ngẫu nhiên (X, Y) . Giả sử trên phần tử thứ i của mẫu $i = \overline{1, n}$, đại lượng (X, Y) nhận giá trị (X_i, Y_i) , khi đó ta có n danh lượng ngẫu nhiên độc lập $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$. Nếu các đại lượng ngẫu nhiên X_1, X_2, \dots, X_n độc lập và cùng phân phối xác suất với X ; các đại lượng ngẫu nhiên Y_1, Y_2, \dots, Y_n độc lập và cùng phân phối xác suất với Y thì ta có mẫu ngẫu nhiên hai chiều, ký hiệu là:

$$W = [(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)].$$

Giả sử đại lượng ngẫu nhiên (X_i, Y_i) nhận giá trị $(x_i, y_i), i = \overline{1, n}$, ta thu được mẫu cụ thể là:

$$w = [(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)].$$

4.4.2. Phương pháp mô tả ngẫu nhiên hai chiều

Giả sử từ tổng thể rút ra một mẫu kích thước n , trong đó thành phần X nhận các giá trị $x_1, x_2, \dots, x_i, \dots, x_k$, còn thành phần Y nhận các giá trị $y_1, y_2, \dots, y_i, \dots, y_h$, trong đó có n_{ij} phần tử của mẫu nhận giá trị (x_i, y_i) . Khi đó giá trị cụ thể của mẫu w được mô tả bởi bảng phân phối tần số thực nghiệm sau:

X \ Y	Y						n_i
	y_1	y_2	\dots	y_i	\dots	y_h	
x_1	n_{11}	n_{12}	\dots	n_{1j}	\dots	n_{1h}	n_1
x_2	n_{21}	n_{22}	\dots	n_{2j}	\dots	n_{2h}	n_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_i	n_{i1}	n_{i2}	\dots	n_{ij}	\dots	n_{ih}	n_i
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_k	n_{k1}	n_{k2}	\dots	n_{kj}	\dots	n_{kh}	n_k
m_j	m_1	m_2	\dots	m_j	\dots	m_k	n

trong đó n_i là tổng số phần tử của mẫu mang giá trị $X = x_i$,
 m_j là tổng số phần tử của mẫu mang giá trị $Y = y_j$.

4.4.3. Một số thống kê đặc trưng của mẫu ngẫu nhiên hai chiều

Từ bảng phân phối tần số thực nghiệm của mẫu ngẫu nhiên hai chiều ta có thể rút ra:

1. Bảng phân phối thực nghiệm của X :

X	x_1	x_2	\dots	x_i	\dots	x_k
n_i	n_1	n_2	\dots	n_i	\dots	n_k

và tính được các thống kê đặc trưng của X :

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

2. Bảng phân phối thực nghiệm của Y :

Y	y_1	y_2	\dots	y_i	\dots	y_h
m_j	m_1	m_2	\dots	m_i	\dots	m_h

và tính được các thống kê đặc trưng của Y :

$$\bar{y} = \frac{1}{n} \sum_{j=1}^h m_j y_j; \quad s_Y^2 = \frac{1}{n} \sum_{j=1}^h m_j (y_j - \bar{y})^2 = \frac{1}{n} \sum_{j=1}^h m_j y_j^2 - \bar{y}^2.$$

3. Bảng phân phối có điều kiện của Y khi $X = x_i$:

$Y_{X=x_i}$	y_1	y_2	\dots	y_i	\dots	y_h
n_{ij}	n_{i1}	n_{i2}	\dots	n_{ij}	\dots	n_{ih}

và tính được các thống kê tương ứng:

$$\bar{y}_{X=x_i} = \frac{1}{n} \sum_{j=1}^h n_{ij} y_j; \quad s_{Y/X=x_i}^2 = \frac{1}{n} \sum_{j=1}^h n_{ij} (y_j - \bar{y})^2 = \frac{1}{n} \sum_{j=1}^h n_{ij} y_j^2 - \bar{y}^2.$$

4. Bảng phân phối có điều kiện của X khi $Y = y_i$:

$X_{Y=y_i}$	x_1	x_2	...	x_i	...	x_k
n_{ij}	n_{1j}	n_{2j}	...	n_{ij}	...	n_{kj}

và tính được các thống kê tương ứng:

$$\bar{x}_{Y=y_j} = \frac{1}{n} \sum_{i=1}^k n_{ij} x_i; \quad s_{X/Y=y_j}^2 = \frac{1}{n} \sum_{i=1}^k n_{ij} (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^k n_{ij} x_i^2 - \bar{x}^2.$$

• **Ví dụ 4.8** Thống kê năng suất một loại cây trồng Y (kg/ha) và lượng đầu tư cho cải tạo đất X (triệu đồng/ha) tại một số địa phương trong vòng 10 năm ta thu được bảng số liệu sau:

	Y	2	3	4	5	6	7
X							
20 - 22		1	3	5			
22 - 24			1	7	8	5	
24 - 26				1	4	3	1
26 - 28					1	2	1
28 - 30						1	1

• Từ bảng trên ta có bảng phân phối tần số của X :

X	2	3	4	5	6	7
Số địa phương (n_i)	1	5	13	13	11	3

Lập bảng tính:

x_i	n_i	$n_i x_i$	$n_i x_i^2$
2	1	2	4
3	5	15	45
4	13	52	208
5	13	65	325
6	11	66	396
7	3	21	147
	$n = \sum n_i = 46$	$\sum n_i x_i = 221$	$\sum n_i x_i^2 = 1125$

Từ đó có kết quả sau:

Lượng đầu tư trung bình: $\bar{x} = \frac{221}{46} \approx 4,804$ (triệu đồng/ha)

Độ phân tán của đầu tư:

$$s'_x = \sqrt{\frac{n}{n-1} \left(\frac{\sum n_i x_i^2}{n} - \bar{x}^2 \right)} = \sqrt{\frac{46}{45} \left(\frac{1125}{46} - 4,804^2 \right)} \approx 1,185 \text{ (triệu đồng/ha)}.$$

• Từ bảng trên ta có bảng phân phối tần số của Y :

Y	21	23	25	27	29
Số địa phương (n_i)	9	22	9	4	2

Tương tự phần trên ta cũng tính được:

$$\bar{y} \approx 23,609 \text{ (kg/ha)}; \quad s'_{y_A} \approx 1,337 \text{ (kg/ha)}.$$

• Tần suất mẫu của những địa phương nhóm A là: $f = \frac{k}{n} = \frac{12}{46} \approx 0,261$.

4.5. Ước lượng tham số của đại lượng ngẫu nhiên

Giả sử cần phải nghiên cứu dấu hiệu định lượng χ trong tổng thể, cụ thể là cần xác định các tham số đặc trưng của tổng thể như trung bình, phương sai, cơ cấu tổng thể theo dấu hiệu nghiên cứu. Nếu xem dấu hiệu nghiên cứu như đại lượng ngẫu nhiên X và giả sử đã biết dạng phân phối xác suất của nó thì vấn đề xác định các tham số đặc trưng của tổng thể sẽ quy về bài toán xác định các tham số đặc trưng của quy luật phân phối xác suất của đại lượng ngẫu nhiên X . Chẳng hạn nếu X là đại lượng ngẫu nhiên phân phối chuẩn thì để xác định trung bình tổng thể ta sẽ ước lượng kỳ vọng toán của X , còn để xác định phương sai tổng thể ta sẽ ước lượng phương sai của X .

Như vậy bài toán ước lượng tham số có thể phát biểu như sau: Cho đại lượng ngẫu nhiên X với quy luật phân phối xác suất đã biết song chưa biết tham số θ nào đó của nó, phải ước lượng (xác định một cách gần đúng) giá trị θ .

Vì tổng thể thường có kích thước lớn nên việc xác định tham số trực tiếp trên tổng thể gặp nhiều khó khăn, người ta giải quyết bài toán ước lượng tham số bằng phương pháp mẫu như sau: Từ tổng thể nghiên cứu rút ra một mẫu ngẫu nhiên kích thước n và dựa vào đó xây dựng một thống kê G dùng để ước lượng θ .

4.5.1. Phương pháp ước lượng điểm

Phương pháp ước lượng điểm chủ trương dùng một giá trị để thay thế cho tham số θ chưa biết của tổng thể. Từ tổng thể lập mẫu ngẫu nhiên kích thước $n : W = (X_1, X_2, \dots, X_n)$. Thông thường giá trị được chọn là một thống kê G nào đó của mẫu ngẫu nhiên. Có rất nhiều cách chọn thống kê G khác nhau tạo nên những phương pháp ước lượng điểm khác nhau. Trong giáo trình này trình bày một loại ước lượng điểm, đó là ước lượng không chệch như sau:

★ **Định nghĩa 4.4** Thống kê G của mẫu được gọi là ước lượng không chệch của tham số θ của đại lượng ngẫu nhiên gốc nếu $M(G) = \theta$.

Ngược lại, nếu $M(G) \neq \theta$ thì G được gọi là ước lượng chệch của θ .

Dựa vào các kết quả của phần (3.3.) có thể rút ra một số kết luận sau;

- Trung bình mẫu \bar{X} là ước lượng không chệch của kỳ vọng toán m của đại lượng ngẫu nhiên gốc:

$$M(\bar{X}) = m.$$

- Tần suất mẫu f là ước lượng không chệch của tần suất tổng thể p (tham số p của đại lượng ngẫu nhiên gốc tuân theo quy luật không - một):

$$M(f) = p.$$

- Phương sai điều chỉnh mẫu S'^2 và phương sai S^{*2} là các ước lượng không chệch của phương sai σ^2 của đại lượng ngẫu nhiên gốc:

$$M(S'^2) = \sigma^2; \quad M(S^{*2}) = \sigma^2.$$

Lưu ý rằng với k mẫu cụ thể rút ra từ tổng thể, G sẽ nhận k giá trị tương ứng là g_1, g_2, \dots, g_k . Khi đó G là ước lượng không chệch của θ không có nghĩa mọi giá trị g_i của G đều bằng θ mà chỉ có nghĩa: trung bình các giá trị của G bằng θ , từng giá trị g_i có thể sai lệch rất lớn so với θ . Tuy nhiên chọn ước lượng không chệch G để ước lượng cho tham số θ ta sẽ loại bỏ được sai số hệ thống: tất cả các giá trị của G đều lớn hơn hoặc nhỏ hơn θ .

4.5.2. Phương pháp ước lượng bằng khoảng tin cậy

Các phương pháp ước lượng điểm có một nhược điểm cơ bản là sai số của ước lượng có thể rất lớn, ngoài ra không đánh giá được khả năng mắc sai lầm khi ước lượng cũng như độ tin cậy của ước lượng bằng bao nhiêu. Do đó khi kích thước mẫu nhỏ người ta thường sử dụng phương pháp ước lượng bằng khoảng tin cậy.

Định nghĩa 5.

★ **Định nghĩa 4.5** Khoảng (G_1, G_2) của thông kê G được gọi là khoảng tin cậy của tham số θ nếu với xác suất bằng $1 - \alpha$ cho trước điều kiện sau đây được thỏa mãn:

$$P(G_1 < \theta < G_2) = 1 - \alpha.$$

$I = G_2 - G_1$ được gọi là độ dài khoảng tin cậy, còn xác suất $1 - \alpha$ được gọi là độ tin cậy của ước lượng (trên thực tế thường đòi hỏi độ tin cậy khá lớn. chẳng hạn $1 - \alpha = 0,95$).

4.5.3. Khoảng tin cậy cho trung bình

(Ước lượng kỳ vọng toán của đại lượng ngẫu nhiên phân phối theo quy luật chuẩn)

Giả sử trong tổng thể đại lượng ngẫu nhiên gốc X phân phối chuẩn $N(\mu, \sigma^2)$ nhưng chưa biết tham số μ của nó. Để ước lượng μ từ tổng thể ta lập mẫu ngẫu nhiên kích thước n : $W = (X_1, X_2, \dots, X_n)$.

Ta xét các trường hợp sau đây:

1. Đã biết phương sai σ^2 của đại lượng ngẫu nhiên gốc trong tổng thể.

Xét thống kê:

$$G = U = \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma}$$

trong đó \bar{X} là trung bình mẫu.

Ta đã biết thống kê U phân phối chuẩn hóa $N(0, 1)$. Do đó với các phân vị chuẩn u_{α_1} và $u_{1-\alpha_2}$ ta có:

$$P(u_{\alpha_1} < U < u_{1-\alpha_2}) = (1 - \alpha_2) - \alpha_1.$$

Từ đó suy ra:

$$\begin{aligned} P\left(u_{\alpha_1} < \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} < u_{1-\alpha_2}\right) &= 1 - (\alpha_1 + \alpha_2) \\ \Leftrightarrow P\left(\bar{X} - \frac{\sigma}{\sqrt{n}}u_{1-\alpha_2} < \mu < \bar{X} - \frac{\sigma}{\sqrt{n}}u_{\alpha_1}\right) &= 1 - (\alpha_1 + \alpha_2) \\ \Leftrightarrow P\left(\bar{X} - \frac{\sigma}{\sqrt{n}}u_{1-\alpha_2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}}u_{1-\alpha_1}\right) &= 1 - (\alpha_1 + \alpha_2) \quad (\text{vì } u_{\alpha_1} = -u_{1-\alpha_1}) \end{aligned}$$

Như vậy với độ tin cậy $1 - \alpha$ cho trước, ta có thể chọn cặp số α_1 và α_2 thỏa mãn điều kiện $\alpha_1 + \alpha_2 = \alpha$ khi đó khoảng tin cậy của tham số μ là:

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}}u_{1-\alpha_2}; \bar{X} + \frac{\sigma}{\sqrt{n}}u_{1-\alpha_1}\right).$$

a) Khoảng tin cậy đối xứng

Nếu lấy $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ thì khoảng tin cậy của μ là:

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}; \bar{X} + \frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\right).$$

Kí hiệu:

$$\varepsilon = \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}. \quad (4.1)$$

ε được gọi là độ chính xác của ước lượng, nó phản ánh mức độ sai lệch của trung bình mẫu so với trung bình tổng thể với xác suất $1 - \alpha$ cho trước.

Khi đó khoảng tin cậy đối xứng của μ có dạng:

$$(\bar{X} - \varepsilon; \bar{X} + \varepsilon).$$

b) Khoảng tin cậy một phía

- Khoảng tin cậy bên phải

Nếu lấy $\alpha_1 = 0$ và $\alpha_2 = \alpha$ thì $u_{1-\alpha_1} = +\infty$, do đó khoảng tin cậy của μ là:

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} u_{1-\alpha}; +\infty \right).$$

Biểu thức trên được sử dụng để ước lượng giá trị tối thiểu của μ .

- Khoảng tin cậy bên trái

Nếu lấy $\alpha_1 = \alpha$ và $\alpha_2 = 0$ thì $u_{1-\alpha_2} = +\infty$, do đó khoảng tin cậy của μ là:

$$\left(-\infty; \bar{X} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha} \right).$$

Biểu thức trên được dùng để ước lượng giá trị tối đa của μ .

c) Xác định kích thước tối thiểu

Với cùng độ tin cậy $1 - \alpha$ khoảng tin cậy nào có độ dài ngắn hơn sẽ tốt hơn. Trong trường hợp này độ dài khoảng tin cậy I sẽ ngắn nhất khi khoảng tin cậy là đối xứng. Lúc đó độ dài khoảng tin cậy đối xứng bằng:

$$I = 2\varepsilon = \frac{2\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}.$$

Do đó với độ tin cậy bằng $1 - \alpha$ cho trước, muốn độ dài khoảng tin cậy đối xứng không vượt quá giá trị I_0 cho trước thì kích thước mẫu n phải thỏa mãn điều kiện:

$$n \geq \frac{4\sigma^2}{I_0^2} u_{1-\frac{\alpha}{2}}^2.$$

Như vậy công thức xác định kích thước mẫu tối thiểu là:

$$n_{tt} = \left\lceil \frac{4\sigma^2}{I_0^2} u_{1-\frac{\alpha}{2}}^2 \right\rceil + 1 \quad \text{hoặc} \quad n_{tt} = \left\lceil \frac{4\sigma^2}{\varepsilon_0^2} u_{1-\frac{\alpha}{2}}^2 \right\rceil + 1.$$

Bài toán xác định kích thước tối thiểu theo công thức trên được đặt ra khi phải xác định kích thước mẫu cần điều tra để đáp ứng những yêu cầu chất lượng cho trước về độ tin cậy và độ chính xác của ước lượng.

d) Xác định độ tin cậy của ước lượng

Từ công thức 4.1 xác định độ chính xác của ước lượng, ta suy ra:

$$u_{1-\frac{\alpha}{2}} = \frac{\varepsilon\sqrt{n}}{\sigma}.$$

Từ phân vị $u_{1-\frac{\alpha}{2}}$, tra bảng phân vị chuẩn ta có thể tìm được $1 - \frac{\alpha}{2}$, từ đó suy ra $1 - \alpha$. Như vậy nếu biết trước độ chính xác (hoặc độ dài khoảng tin cậy đối xứng) của ước lượng thì có thể tìm được độ tin cậy của ước lượng.

Chú ý rằng các khoảng tin cậy đối xứng là khoảng tin cậy một phía trong các công thức trên chỉ là các khoảng tin cậy ngẫu nhiên. Muốn tìm khoảng tin cậy cụ thể ta phải thực hiện phép thử đối với mẫu ngẫu nhiên $W = (X_1, X_2, \dots, X_n)$ để thu được mẫu cụ thể $w = (x_1, x_2, \dots, x_n)$, từ đó tìm được giá trị cụ thể của trung bình mẫu \bar{x} . Khi đó với mẫu tin cậy $1 - \alpha$, khoảng tin cậy đối xứng cụ thể của μ (qua mẫu cụ thể trên) là:

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}; \bar{x} + \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right).$$

Tiến hành tương tự ta cũng thu được giá trị cụ thể của các khoảng tin cậy một phía.

•**Ví dụ 4.9** Trọng lượng một loại sản phẩm là đại lượng ngẫu nhiên phân phối theo quy luật chuẩn với độ lệch tiêu chuẩn là $\sigma = 1$ gam. Cân thử 25 sản phẩm loại này ta thu được kết quả sau:

Trọng lượng (gam)	18	19	20	21
Số sản phẩm tương ứng	3	5	15	2

- a) Với độ tin cậy là 0,95 hãy tìm khoảng tin cậy đối xứng của trọng lượng trung bình của loại sản phẩm nói trên.
- b) Với độ tin cậy là 98% hãy ước lượng trọng lượng trung bình tối đa của loại sản phẩm trên.
- c) Muốn ước lượng trọng lượng trung bình của loại sản phẩm trên với độ tin cậy là 98%, độ dài khoảng tin cậy đối xứng không quá 0,7gam thì cần kiểm tra ít nhất bao nhiêu sản phẩm?
- d) Muốn sử dụng kết quả cân thử 25 sản phẩm trên để ước lượng trọng lượng trung bình của loại sản phẩm trên với độ chính xác 0,399 gam thì đạt độ tin cậy bao nhiêu?

Lời giải.

Gọi X là trọng lượng sản phẩm, theo giả thiết X phân phối chuẩn với $\sigma = 1$. Vậy trọng lượng trung bình của sản phẩm chính là tham số μ .

a) Đây là bài toán ước lượng tham số μ bằng khoảng tin cậy đối xứng:

$$\mu \in \left(\bar{x} - \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}; \bar{x} + \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right).$$

Từ mẫu ngẫu nhiên kích thước $n = 25$ ở trên ta tích được trung bình mẫu.

$$\bar{x} = \frac{18.3 + 19.5 + 20.15 + 21.2}{25} = \frac{491}{25} = 19,64.$$

Từ độ tin cậy $1 - \alpha = 0,95$ suy ra $1 - \frac{\alpha}{2} = 0,975$. Tra bảng phân vị chuẩn ta có $u_{1-\frac{\alpha}{2}} = u_{0,975} = 1,96$. Độ chính xác của ước lượng:

$$\varepsilon = \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} = \frac{1}{\sqrt{25}} 1,96 = 0,392.$$

Vậy khoảng tin cậy đối xứng của μ là:

$$\mu \in (\bar{x} - \varepsilon; \bar{x} + \varepsilon) = (19,64 - 0,392; 19,64 + 0,392) = (19,148; 20,032).$$

b) Đây là bài toán ước lượng giá trị tối đa của μ

$$\mu \in \left(-\infty; \bar{x} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha} \right).$$

Từ độ tin cậy $1 - \alpha = 0,98$, tra bảng phân vị chuẩn ta có $u_{0,98} = 2,054$
 Giá trị tối đa của μ là:

$$\bar{x} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha} = 19,64 + \frac{1,2,054}{\sqrt{25}} = 20,0508.$$

c) Đây là bài toán xác định kích thước mẫu tối thiểu:

$$n \geq \frac{4\sigma^2}{I_0^2} u_{1-\frac{\alpha}{2}}^2.$$

Từ độ tin cậy $1 - \alpha = 0,98$, suy ra $1 - \frac{\alpha}{2} = 0,99$. Tra bảng phân vị chuẩn: $u_{0,99} = 2,326$

$$n \geq \frac{4 \cdot 1^2}{0,7^2} 2,326^2 \approx 44,2.$$

Vậy cần phải kiểm tra ít nhất 45 sản phẩm (cần kiểm tra thêm ít nhất 20 sản phẩm nữa).

d) Đây là bài toán xác định độ tin cậy:

$$u_{1-\frac{\alpha}{2}} = \frac{\varepsilon \sqrt{n}}{\sigma} = \frac{0,399 \sqrt{25}}{1} = 1,995.$$

Từ bảng phân vị chuẩn suy ra $1 - \frac{\alpha}{2} = 0,977$ do đó $1 - \alpha = 0,954$. Vậy độ tin cậy bằng 95,4%.

2. Chưa biết phương sai σ^2 của đại lượng ngẫu nhiên X và kích thước mẫu $n \leq 30$
 Chọn thống kê:

$$G = T = \frac{(\bar{X} - \mu)\sqrt{n}}{S'}$$

trong đó \bar{X} là trung bình mẫu, và S' là độ lệch tiêu chuẩn điều chỉnh mẫu.

Ta đã biết T phân phối theo quy luật Student với $n - 1$ bậc tự do. Vì vậy với các phân vị Student $t_{\alpha_1}^{n-1}$ và $t_{1-\alpha_2}^{n-1}$ ta có:

$$P(t_{\alpha_1}^{n-1} < T < t_{1-\alpha_2}^{n-1}) = (1 - \alpha_2) - \alpha_1.$$

Suy ra:

$$P\left(\bar{X} - \frac{S'}{\sqrt{n}} t_{1-\alpha_2}^{n-1} < \mu < \bar{X} + \frac{S'}{\sqrt{n}} t_{1-\alpha_1}^{n-1}\right) = 1 - (\alpha_1 + \alpha_2).$$

Chọn $\alpha_1 + \alpha_2 = \alpha$, khoảng tin cậy của μ với độ tin cậy $1 - \alpha$ là:

$$\left(\bar{X} - \frac{S'}{\sqrt{n}} t_{1-\alpha_2}^{n-1}; \bar{X} + \frac{S'}{\sqrt{n}} t_{1-\alpha_1}^{n-1}\right).$$

a) Khoảng tin cậy đối xứng

Chọn $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ ta có khoảng tin cậy đối xứng:

$$\left(\bar{X} - \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{n-1}; \bar{X} + \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{n-1}\right)$$

với độ chính xác:

$$\varepsilon = \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{n-1}$$

và độ dài khoảng tin cậy đối xứng:

$$I = 2\varepsilon = 2 \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{n-1}.$$

b) Khoảng tin cậy một phía

- Khoảng tin cậy bên phải

Chọn $\alpha_1 = 0, \alpha_2 = \alpha$ ta có khoảng tin cậy bên phải dùng để ước lượng giá trị tối thiểu của μ như sau:

$$\left(\bar{X} - \frac{S'}{\sqrt{n}} t_{1-\alpha}^{n-1}; +\infty \right).$$

- Khoảng tin cậy bên trái

Chọn $\alpha_1 = \alpha, \alpha_2 = 0$ ta có khoảng tin cậy bên trái dùng để ước lượng giá trị tối thiểu của μ như sau:

$$\left(-\infty; \bar{X} - \frac{S'}{\sqrt{n}} t_{1-\alpha}^{n-1} \right).$$

c) Xác định kích thước mẫu tối thiểu

Với độ tin cậy bằng $1 - \alpha$ cho trước, độ dài khoảng tin cậy đối xứng không vượt quá giá trị I_o cho trước, bài toán xác định kích thước mẫu tối thiểu được giải quyết bằng phương pháp mẫu kép sau đây:

Trước hết điều tra một mẫu sơ bộ kích thước $n \geq 2$: $W_1 = (X_1, X_2, \dots, X_n)$, từ đó tìm được phương sai điều chỉnh mẫu:

$$S'^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2; \quad \text{với} \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Giả sử để thỏa mãn độ chính xác và độ tin cậy cho trước, mẫu cần lập phải có kích thước N , khi đó giả sử lập mẫu thứ hai: $W_2 = (X_{n+1}, X_{n+2}, \dots, X_N)$.

Vì thống kê:

$$T = \frac{\left(\frac{1}{N} \sum_{i=1}^N X_i - \mu \right) \sqrt{n}}{S'}$$

cũng phân phối theo quy luật Student $n - 1$ bậc tự do nên phân vị Student $t_{1-\frac{\alpha}{2}}^{n-1}$ ta cũng có:

$$P\left(\sum_{i=1}^N X_i - \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{n-1} < \mu < \sum_{i=1}^N X_i + \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{n-1} \right) = 1 - \alpha.$$

Khi đó ước lượng có độ chính xác là:

$$\varepsilon_o = \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{n-1}.$$

Từ đó suy ra kích thước mẫu cần tìm:

$$N \geq \left(\frac{S'}{\varepsilon_o} t_{1-\frac{\alpha}{2}}^{n-1} \right)^2 \quad \text{hoặc} \quad N \geq \left(\frac{2S'}{I_o} t_{1-\frac{\alpha}{2}}^{n-1} \right)^2.$$

d) Xác định độ tin cậy của ước lượng

Từ mẫu ngẫu nhiên kích thước n , muốn ước lượng μ khoảng tin cậy đối xứng, với độ chính xác ε_o , thì độ tin cậy $1 - \alpha$ của ước lượng được suy ra từ biểu thức:

$$t_{1-\frac{\alpha}{2}}^{n-1} = \frac{\varepsilon_o \sqrt{n}}{S'}.$$

• **Ví dụ 4.10** Phân xưởng sản xuất tự động của nhà máy gang thép Thái Nguyên sản xuất 1 loại thép thành phẩm với chiều dài đặt trước bởi đơn đặt hàng là X (m) (X tuân theo quy luật phân phối chuẩn). Kiểm tra ngẫu nhiên 25 sản phẩm ta được bảng số liệu sau:

Chiều dài X (m)	1,0	1,1	1,2	1,3	1,4
Số sản phẩm (n_i)	3	6	9	5	2

- a) Bằng khoảng tin cậy đối xứng hãy ước lượng chiều dài trung bình của loại thép trên với độ tin cậy 95%.
- b) Muốn ước lượng chiều dài trung bình của loại thép đó với độ chính xác 5,63 (cm) thì đạt độ tin cậy bao nhiêu?

Lời giải.

- a) Đây là bài toán ước lượng bằng khoảng tin cậy đối xứng giá trị của tham số μ của phân phối $N(\mu, \sigma^2)$ khi chưa biết σ^2 của X , hơn nữa kích thước mẫu $n = 25 < 30$ nên bài toán được giải như sau:

Từ số liệu đã cho ta lập bảng:

x_i	n_i	$n_i x_i$	$n_i x_i^2$
1,0	3	3,0	3,00
1,1	6	6,6	7,26
1,2	9	10,8	12,96
1,3	5	6,5	8,45
1,4	2	2,8	3,92
	$n = \sum n_i = 25$	$\sum n_i x_i = 29,7$	$\sum n_i x_i^2 = 55,59$

Chiều dài trung bình mẫu: $\bar{x} = \frac{29,7}{25} = 1,188$ (m)

Độ lệch tiêu chuẩn mẫu điều chỉnh:

$$s'_x = \sqrt{\frac{n}{n-1} \left(\frac{\sum n_i x_i^2}{n} - \bar{x}^2 \right)} = \sqrt{\frac{25}{24} \left(\frac{55,59}{25} - 1,188^2 \right)} \approx 0,113 \text{ (m)}.$$

Từ độ tin cậy $1-\alpha = 0,95$ suy ra $1-\frac{\alpha}{2} = 0,975$. Tra bảng phân vị Student $t_{1-\frac{\alpha}{2}}^{n-1} = t_{0,975}^{24} = 2,064$.

Độ chính xác: $\epsilon = \frac{s'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{n-1} = \frac{0,113 \cdot 2,064}{\sqrt{25}} \approx 0,047$ (m).

Vậy khoảng tin cậy đối xứng của μ là:

$$1,188 - 0,047 < \mu < 1,188 + 0,047 \quad \text{hay} \quad 1,141 < \mu < 1,235 \text{ (m)}.$$

- b) Đây là bài toán xác định độ tin cậy của ước lượng, giả thiết đã cho độ chính xác $\epsilon = 5,63$ (cm) = 0,0563(m), và từ bảng tính phần a) ta đã có $s' \approx 0,113$ (m). Từ đó tính được:

$$t_{1-\frac{\alpha}{2}}^{n-1} = t_{1-\frac{\alpha}{2}}^{24} = \frac{\epsilon \sqrt{n}}{s'} = \frac{0,0563 \cdot \sqrt{25}}{0,113} \approx 2,491.$$

Từ bảng phân vị Student thấy $t_{1-\frac{\alpha}{2}}^{24} \approx 2,491 \approx t_{0,99}^{24}$. Vậy:

$$1 - \frac{\alpha}{2} = 0,99 \Rightarrow 1 - \alpha = 0,98.$$

3. Chưa biết phương sai σ^2 của đại lượng ngẫu nhiên gốc X và kích thước $n > 30$

Ta chọn thống kê

$$G = T = \frac{(\bar{X} - \mu)\sqrt{n}}{S'}$$

trong đó \bar{X} là trung bình mẫu, và S' là độ lệch tiêu chuẩn điều chỉnh mẫu.

Vì kích thước $n > 30$ nên thống kê T sẽ phân phối xấp xỉ chuẩn hóa $N(0, 1)$. Do đó tương tự các phần trên ta cũng có:

$$P(u_{\alpha_1} < T < u_{1-\alpha_2}) = (1 - \alpha_2) - \alpha_1.$$

Từ đó suy ra:

$$P\left(\bar{X} - \frac{S'}{\sqrt{n}}u_{1-\alpha_2} < \mu < \bar{X} + \frac{S'}{\sqrt{n}}u_{1-\alpha_1}\right) = 1 - (\alpha_1 + \alpha_2).$$

Như vậy với độ tin cậy là $1 - \alpha$ cho trước, ta có thể chọn cặp số α_1 và α_2 thỏa mãn điều kiện $\alpha_1 + \alpha_2 = \alpha$, khi đó khoảng tin cậy của tham số μ là:

$$\left(\bar{X} - \frac{S'}{\sqrt{n}}u_{1-\alpha_2}; \bar{X} + \frac{S'}{\sqrt{n}}u_{1-\alpha_1}\right).$$

a) Khoảng tin cậy đối xứng khi $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$:

$$\left(\bar{X} - \frac{S'}{\sqrt{n}}u_{1-\frac{\alpha}{2}}; \bar{X} + \frac{S'}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\right)$$

với độ chính xác:

$$\varepsilon = \frac{S'}{\sqrt{n}}u_{1-\frac{\alpha}{2}}$$

và độ dài khoảng tin cậy đối xứng:

$$I = 2\varepsilon = \frac{2S'}{\sqrt{n}}u_{1-\frac{\alpha}{2}}.$$

b) Khoảng tin cậy một phía

- Khoảng tin cậy bên phải

$$\left(\bar{X} - \frac{S'}{\sqrt{n}}u_{1-\alpha}; +\infty\right).$$

- Khoảng tin cậy bên trái

$$\left(-\infty; \bar{X} + \frac{S'}{\sqrt{n}}u_{1-\alpha}\right).$$

c) Xác định kích thước mẫu tối thiểu

Muốn ước lượng μ với độ tin cậy $1 - \alpha$, độ dài khoảng tin cậy đối xứng I_o thì kích thước mẫu N phải thỏa mãn:

$$N \geq \left(\frac{S'}{\varepsilon_o}u_{1-\frac{\alpha}{2}}\right)^2 \quad \text{hay} \quad N \geq \left(\frac{2S'}{I_o}u_{1-\frac{\alpha}{2}}\right)^2$$

ở đó S' là độ lệch tiêu chuẩn điều chỉnh của mẫu sơ bộ có kích thước n

d) Xác định độ tin cậy của ước lượng

Từ mẫu ngẫu nhiên kích thước n , muốn ước lượng μ bằng khoảng tin cậy đối xứng, với độ chính xác ε_o , thì độ tin cậy $1 - \alpha$ của ước lượng được suy ra từ biểu thức:

$$u_{1-\frac{\alpha}{2}} = \frac{\varepsilon_o\sqrt{n}}{S'}$$

4.5.4. Khoảng tin cậy cho tỷ lệ

(Ước lượng kỳ vọng toán của đại lượng ngẫu nhiên phân phối theo quy luật không - một)

Giả sử trong tổng thể kích thước N có M phần tử mang dấu hiệu nghiên cứu. Như vậy cơ cấu của tổng thể theo dấu hiệu nghiên cứu là:

$$p = \frac{M}{N}.$$

Nếu lấy ngẫu nhiên ra một phần tử, và gọi X là số phần tử mang dấu hiệu nghiên cứu được lấy ra thì X là đại lượng ngẫu nhiên phân phối theo quy luật không - một $A(p)$ với kỳ vọng toán $M(X) = p$ và phương sai $D(X) = \frac{p(1-p)}{n}$. Do đó bài toán ước lượng cơ cấu tổng thể chính là bài toán ước lượng kỳ vọng toán của đại lượng ngẫu nhiên tuân theo quy luật không - một.

Xét ngẫu nhiên kích thước n : $W = (X_1, X_2, \dots, X_n)$. Nếu kích thước mẫu khá lớn ta có thể chọn thống kê:

$$G = U = \frac{(f - p)\sqrt{n}}{\sqrt{f(1-f)}}$$

trong đó f là tần suất mẫu.

Với n khá lớn thống kê U phân phối xấp xỉ chuẩn hóa $N(0, 1)$. Do đó với các phân vị chuẩn u_{α_1} và $u_{1-\alpha_2}$ ta có:

$$P(u_{\alpha_1} < U < u_{1-\alpha_2}) = (1 - \alpha_2) - \alpha_1.$$

Chọn $\alpha_1 + \alpha_2 = \alpha$ ta có:

$$\begin{aligned} P\left(-u_{1-\alpha_1} < \frac{(f-p)\sqrt{n}}{\sqrt{f(1-f)}} < u_{1-\alpha_2}\right) &= 1 - \alpha \\ \Leftrightarrow P\left(f - \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\alpha_2} < p < f + \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\alpha_1}\right) &= 1 - \alpha. \end{aligned}$$

a) Khoảng tin cậy đối xứng

Chọn $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ ta có khoảng tin cậy đối xứng:

$$\left(f - \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\frac{\alpha}{2}}; f + \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\right)$$

với độ chính xác:

$$\varepsilon = \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\frac{\alpha}{2}}$$

và độ dài khoảng tin cậy đối xứng:

$$I = 2\varepsilon = \frac{2\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\frac{\alpha}{2}}.$$

b) Khoảng tin cậy một phía

- Khoảng tin cậy bên phải

Chọn $\alpha_1 = 0, \alpha_2 = \alpha$ ta có khoảng tin cậy bên phải dùng để ước lượng giá trị tối thiểu của p như sau:

$$\left(f - \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\alpha}; +\infty\right).$$

- Khoảng tin cậy bên trái

Chọn $\alpha_1 = \alpha, \alpha_2 = 0$ ta có khoảng tin cậy bên trái dùng để ước lượng giá trị tối thiểu của p như sau:

$$\left(-\infty; f + \frac{\sqrt{f(1-f)}}{\sqrt{n}} u_{1-\alpha} \right).$$

c) Xác định kích thước mẫu tối thiểu

Để đảm bảo độ tin cậy $1 - \alpha$ cho trước và độ dài khoảng tin cậy đối xứng không vượt quá I_o cho trước thì kích thước mẫu N phải đảm bảo:

$$N \geq \frac{4f(1-f)}{I_o^2} u_{1-\frac{\alpha}{2}}^2$$

trong đó f là tần suất mẫu sơ bộ kích thước n

d) Xác định độ tin cậy của ước lượng

Từ mẫu ngẫu nhiên kích thước n , muốn ước lượng p bằng khoảng tin cậy đối xứng, với độ chính xác ε , thì độ tin cậy $1 - \alpha$ của ước lượng được suy ra từ biểu thức:

$$u_{1-\frac{\alpha}{2}} = \frac{\varepsilon_o \sqrt{n}}{\sqrt{f(1-f)}}.$$

• **Ví dụ 4.11** Để điều tra số cá trong một hồ, cơ quan quản lý đánh bắt 2000 con cá, đánh dấu rồi thả xuống hồ. Lần sau bắt lại 400 con cá, được 80 con có dấu.

- Hãy ước lượng tỷ lệ cá bị đánh dấu trong hồ bằng khoảng tin cậy đối xứng với độ tin cậy 97%;
- Hãy ước lượng số cá tối thiểu trong hồ với độ tin cậy 95%;
- Muốn ước lượng tỷ lệ cá bị đánh dấu trong hồ với độ tin cậy 98%, độ chính xác không quá 0,04 thì cần đánh bắt bao nhiêu con cá nữa?
- Muốn ước lượng tỷ lệ cá bị đánh dấu trong hồ với độ dài khoảng tin cậy đối xứng không quá 5,76% thì đạt độ tin cậy bao nhiêu?

Lời giải.

Mẫu ngẫu nhiên có kích thước $n = 400$ (con cá), trong đó $k = 80$ con có dấu. Do đó tần suất mẫu là $f = \frac{k}{n} = \frac{80}{400} = 0,2$.

- Từ độ tin cậy $1 - \alpha = 0,97 \Rightarrow 1 - \frac{\alpha}{2} = 0,985$, tra bảng phân vị chuẩn $u_{0,985} = 2,17$;

Độ chính xác của ước lượng là: $\varepsilon = \frac{\sqrt{f(1-f)}}{\sqrt{n}} u_{1-\frac{\alpha}{2}} = \frac{\sqrt{0,2 \cdot 0,8 \cdot 2,17}}{\sqrt{400}} = 0,0434$.

Khoảng tin cậy đối xứng là:

$$p \in (0,2 - 0,0434; 0,2 + 0,0434) \quad \text{hay} \quad 0,1566 < p < 0,2434.$$

-

- Ước lượng tỉ lệ cá bị đánh dấu tối đa

Từ độ tin cậy $1 - \alpha = 0,95$, tra bảng phân vị chuẩn $u_{0,95} = 1,645$.

Độ chính xác một phía: $\varepsilon' = \frac{\sqrt{f(1-f)}}{\sqrt{n}} u_{1-\alpha} = \frac{\sqrt{0,2 \cdot 0,8 \cdot 1,645}}{\sqrt{400}} = 0,0329$.

Khoảng tin cậy bên trái là: $p \in (-\infty; 0,2 + 0,0329)$ hay $p < 0,2329$.

- Gọi số cá trong hồ là N . Ta có: $p = \frac{2000}{N} < 0,2329$ suy ra $N > 8587,4$.
Vậy số cá tối thiểu trong hồ là 8588 (con cá).

- Từ độ tin cậy $1 - \alpha = 0,98 \Rightarrow 1 - \frac{\alpha}{2} = 0,99$ tra bảng phân vị chuẩn $u_{0,99} = 2,326$;

$$n_{tt} \geq \frac{f(1-f)(u_{0,99})^2}{\varepsilon^2} = \frac{0,2 \cdot 0,8 \cdot 2,326^2}{0,04^2} \approx 541,03.$$

Vậy cần đánh bắt thêm ít nhất $542 - 400 = 142$ con cá nữa.

d) Độ dài khoảng tin cậy đối xứng $I_o = 5,76\%$, suy ra độ chính xác $\varepsilon = 2,88\% = 0,0288$

$$u_{1-\frac{\alpha}{2}} = \frac{\varepsilon\sqrt{n}}{\sqrt{f(1-f)}} = \frac{0,0288\sqrt{400}}{\sqrt{0,20,8}} = 1,44.$$

Từ bảng phân vị chuẩn ta thấy $1,44 = u_{0,925}$. Vậy $1 - \frac{\alpha}{2} = 0,925$ suy ra độ tin cậy: $1 - \alpha = 0,85$.

Ví dụ 11+. Một doanh nghiệp dự định đưa một sản phẩm mới vào tiêu thụ ở một vùng dân cư có 2500000 người. Nghiên cứu thị trường đối với 3500 người thấy có 1500 người sẵn sàng mua sản phẩm đó.

a. Với độ tin cậy 95%, hãy dự đoán thị phần tiềm năng của doanh nghiệp.

b. Dự đoán số lượng khách hàng tiềm năng mà doanh nghiệp hy vọng sẽ có được ở thị trường mới là bao nhiêu?

Đây là bài toán liên quan đến dự đoán thống kê. Hoạt động dự đoán này sẽ được giải quyết theo con đường lý thuyết, nghĩa là sinh viên phải vận dụng những kiến thức về ước lượng tham số để giải quyết vấn đề. Kết quả của dự đoán sẽ được thể hiện dưới dạng khoảng. Để dự đoán thị phần tiềm năng của doanh nghiệp, sinh viên phải tìm được khoảng tin cậy cho tỷ lệ p : tỷ lệ khách hàng

sẵn sàng mua sản phẩm, với điều kiện $n > 5$ và $\frac{|\sqrt{\frac{p}{p-1}} - \sqrt{\frac{1-p}{p}}|}{\sqrt{n}} < 0,3$. Khoảng tin cậy cho p là:

$\left(f - \frac{\sqrt{f(f-1)}}{\sqrt{n}} \phi_0^{-1} \left(\frac{y}{2} \right); f + \frac{\sqrt{f(f-1)}}{\sqrt{n}} \phi_0^{-1} \left(\frac{y}{2} \right) \right)$. Với tần suất mẫu $f = 0,428$; $\phi_0^{-1} \left(\frac{y}{2} \right) = 1,96$. Ta có: $0,412 < p < 0,444$.

Từ đó sinh viên có thể dự đoán được thị trường tiềm năng của doanh nghiệp chiếm 41,2% đến 44,4%.

Muốn dự đoán được số lượng khách hàng tiềm năng của doanh nghiệp, sinh viên phải tiến hành phân tích, so sánh, khái quát hóa để nhận ra tỷ lệ $\frac{N}{2500000}$ thuộc khoảng tin cậy p , với N là số lượng khách hàng tiềm năng.

Khi đó, các em sẽ lập được mối quan hệ; $0,412 < \frac{N}{2500000} < 0,444$.

Suy ra: $1030000 < N < 1110000$.

Vậy số lượng khách tiềm năng mà doanh nghiệp hy vọng sẽ có được ở thị trường mới từ 1030000 đến 1110000 người.

4.5.5. Khoảng tin cậy cho phương sai

(Ước lượng phương sai của đại lượng ngẫu nhiên phân phối theo quy luật chuẩn)

Giả sử trong tổng thể đại lượng ngẫu nhiên gốc X phân phối chuẩn $N(\mu, \sigma^2)$ nhưng chưa biết tham số σ^2 của nó. Để ước lượng σ^2 từ tổng thể ta lập mẫu ngẫu nhiên kích thước n : $W = (X_1, X_2, \dots, X_n)$.

Ta xét các trường hợp sau đây:

1. Đã biết kỳ vọng toán μ của đại lượng ngẫu nhiên gốc X

Xét thống kê:

$$G = \frac{nS^{*2}}{\sigma^2}; \quad \text{với} \quad S^{*2} = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2.$$

Thống kê G ở trên phân phối theo quy luật "khi bình phương" với n bậc tự do $\chi^2(n)$. Do đó với độ tin cậy $1 - \alpha$ cho trước có thể tìm được cặp giá trị α_1 và α_2 sao cho $\alpha_1 + \alpha_2 = \alpha$. Khi đó với hai phân vị "khi bình phương" n bậc tự do tương ứng $\chi_{\alpha_1}^{2(n)}$ và $\chi_{1-\alpha_2}^{2(n)}$ ta có:

$$P \left(\chi_{\alpha_1}^{2(n)} < \frac{nS^{*2}}{\sigma^2} < \chi_{1-\alpha_2}^{2(n)} \right) = (1 - \alpha_2) - \alpha_1$$

$$\Leftrightarrow P\left(\frac{nS^{*2}}{\chi_{1-\alpha_2}^{2(n)}} < \sigma^2 < \frac{nS^{*2}}{\chi_{\alpha_1}^{2(n)}}\right) = 1 - \alpha.$$

- Chọn $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ ta có khoảng tin cậy của σ^2 như sau:

$$\left(\frac{nS^{*2}}{\chi_{1-\frac{\alpha}{2}}^{2(n)}}; \frac{nS^{*2}}{\chi_{\frac{\alpha}{2}}^{2(n)}}\right).$$

Ta thấy rằng khoảng tin cậy này không đối xứng.

- Chọn $\alpha_1 = 0, \alpha_2 = \alpha$ ta có khoảng tin cậy bên phải dùng để ước lượng giá trị tối thiểu của σ^2 như sau:

$$\left(\frac{nS^{*2}}{\chi_{1-\alpha}^{2(n)}}; +\infty\right).$$

- Chọn $\alpha_1 = \alpha, \alpha_2 = 0$ ta có khoảng tin cậy bên trái dùng để ước lượng giá trị tối thiểu của σ^2 như sau:

$$\left(0; \frac{nS^{*2}}{\chi_{\alpha}^{2(n)}}\right).$$

• **Ví dụ 4.12** Mức hao phí nguyên liệu cho một đơn vị sản phẩm là đại lượng ngẫu nhiên phân phối chuẩn với trung bình là 20g. Để ước lượng mức độ phân tán của mức hao phí này người ta cân thử 25 sản phẩm và thu được kết quả sau:

Hao phí nguyên liệu (g)	19,5	20,0	20,5
Số sản phẩm tương ứng	5	18	2

Với độ tin cậy $1 - \alpha = 0,90$ hãy ước lượng σ^2 nếu $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$

Lời giải.

Gọi X là mức hao phí nguyên liệu cho một đơn vị sản phẩm, X tuân theo quy luật chuẩn $N(\mu = 20, \sigma^2)$. Đây là bài toán ước lượng phương sai của phân phối chuẩn khi đã biết μ .

Tra bảng phân bố Khi-bình phương:

$$\chi_{\frac{\alpha}{2}}^{2(n)} = \chi_{0,05}^{2(25)} = 14,6; \quad \chi_{1-\frac{\alpha}{2}}^{2(n)} = \chi_{0,95}^{2(25)} = 37,7.$$

Lập bảng tính s^{*2} :

x_i	n_i	$x_i - \mu$	$n_i(x_i - \mu)$	$n_i(x_i - \mu)^2$
19,5	5	-0,5	-2,5	1,25
20,0	18	0	0	0
20,5	2	0,5	1	0,5
	$n = \sum n_i = 25$			$\sum n_i(x_i - \mu)^2 = 1,75$

$$s^{*2} = \frac{1}{n} \sum n_i(x_i - \mu)^2 = \frac{1,75}{25} = 0,07.$$

Khoảng tin cậy của σ^2 là:

$$\sigma^2 \in \left(\frac{n \cdot s^{*2}}{\chi_{1-\frac{\alpha}{2}}^{2(n)}}; \frac{n \cdot s^{*2}}{\chi_{\frac{\alpha}{2}}^{2(n)}}\right) = \left(\frac{25 \cdot 0,07}{37,7}; \frac{25 \cdot 0,07}{14,6}\right) = (0,0464; 0,1198).$$

2. Chưa biết kỳ vọng toán μ của đại lượng ngẫu nhiên gốc X

Ta chọn thông kê:

$$G = \frac{(n-1)S'^2}{\sigma^2}$$

trong đó S'^2 là phương sai điều chỉnh mẫu.

Thông kê G ở trên phân phối theo quy luật "khi bình phương" với $n-1$ bậc tự do $\chi^2(n-1)$.

Do đó với độ tin cậy $1 - \alpha$ cho trước có thể tìm được cặp giá trị α_1 và α_2 sao cho $\alpha_1 + \alpha_2 = \alpha$. Khi đó với hai phân vị "khi bình phương" $n-1$ bậc tự do tương ứng $\chi_{\alpha_1}^{2(n-1)}$ và $\chi_{1-\alpha_2}^{2(n-1)}$ ta có:

$$P\left(\chi_{\alpha_1}^{2(n-1)} < \frac{(n-1)S'^2}{\sigma^2} < \chi_{1-\alpha_2}^{2(n-1)}\right) = (1 - \alpha_2) - \alpha_1$$

$$\Leftrightarrow P\left(\frac{(n-1)S'^2}{\chi_{1-\alpha_2}^{2(n-1)}} < \sigma^2 < \frac{(n-1)S'^2}{\chi_{\alpha_1}^{2(n-1)}}\right) = 1 - \alpha.$$

- Chọn $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ ta có khoảng tin cậy của σ^2 như sau:

$$\left(\frac{(n-1)S'^2}{\chi_{1-\frac{\alpha}{2}}^{2(n-1)}}; \frac{(n-1)S'^2}{\chi_{\frac{\alpha}{2}}^{2(n-1)}}\right).$$

- Chọn $\alpha_1 = 0, \alpha_2 = \alpha$ ta có khoảng tin cậy bên phải dùng để ước lượng giá trị tối thiểu của σ^2 như sau:

$$\left(\frac{(n-1)S'^2}{\chi_{1-\alpha}^{2(n-1)}}; +\infty\right).$$

- Chọn $\alpha_1 = \alpha, \alpha_2 = 0$ ta có khoảng tin cậy bên trái dùng để ước lượng giá trị tối thiểu của σ^2 như sau:

$$\left(0; \frac{(n-1)S'^2}{\chi_{\alpha}^{2(n-1)}}\right).$$

• **Ví dụ 4.13** Lãi suất cổ phiếu của một công ty trong vòng 5 năm qua là:

$$15\%; \quad 10\%; \quad 20\%; \quad 7\%; \quad 14\%;$$

a) Biết lãi suất cổ phiếu là biến ngẫu nhiên phân phối chuẩn, hãy ước lượng độ phân tán của lãi suất cổ phiếu của công ty đó với độ tin cậy 90%.

b) Hãy ước lượng độ phân tán tối đa của lãi suất cổ phiếu của công ty với độ tin cậy 99%.

Lời giải.

Gọi X là lãi suất cổ phiếu của công ty, X tuân theo quy luật chuẩn $N(\mu, \sigma^2)$. Đây là bài toán ước lượng σ^2 của X với μ chưa biết.

a) Từ độ tin cậy $1 - \alpha = 0,9$ suy ra $\frac{\alpha}{2} = 0,05$. Tra bảng phân vị Khi-bình phương:

$$\chi_{\frac{\alpha}{2}}^{2(n-1)} = \chi_{0,05}^{2(4)} = 0,711; \quad \chi_{1-\frac{\alpha}{2}}^{2(n-1)} = \chi_{0,95}^{2(4)} = 9,5.$$

Tính s'^2 : Từ giả thiết dễ tính được: $\sum x_i = 66$; $\sum x_i^2 = 970$. Do đó:

$$\bar{x} = \frac{66}{5} = 13,2(\%); \quad s'^2 = \frac{n}{n-1} \left(\frac{\sum x_i^2}{n} - \bar{x}^2 \right) = \frac{5}{4} \left(\frac{970}{5} - 13,2^2 \right) = 24,7.$$

Khoảng tin cậy của σ^2 là:

$$\sigma^2 \in \left(\frac{(n-1)s'^2}{\chi_{1-\frac{\alpha}{2}}^{2(n-1)}}; \frac{(n-1)s'^2}{\chi_{\frac{\alpha}{2}}^{2(n-1)}} \right) = \left(\frac{4.24,7}{9,5}; \frac{4.24,7}{0,711} \right) = (10,4; 138,959).$$

b) Từ độ tin cậy $1 - \alpha = 0,99$, tra bảng phân vị Khi-bình phương: $\chi_{\alpha}^{2(n-1)} = \chi_{0,01}^{2(4)} = 0,297$
 Khoảng tin cậy bên trái của σ^2 là:

$$\sigma^2 \in \left(0; \frac{(n-1)s'^2}{\chi_{\alpha}^{2(n-1)}} \right) = \left(0; \frac{4.24,7}{0,297} \right) = (0; 332,660).$$

Vậy $\sigma^2 < 332,66$ hay $\sigma < 18,239\%$.

Bài tập chương 4

Mẫu ngẫu nhiên và các tham số đặc trưng của nó

Bài 4.1. Cho 8 kết quả đo đạc về một đại lượng X bởi cùng một máy không có sai lầm hệ thống:

369;378;315;420;385;401;372;383

Hãy tính \bar{x} , s^2 , s' .

ĐS: $\bar{x} = 377,875$; $s^2 = 806,609$; $s' = 30,362$.

Bài 4.2. Theo dõi thời gian hoàn thành một sản phẩm (phút) ở hai nhóm công nhân, ta thu được số liệu sau:

Nhóm 1:

X_1 (phút)	42	44	45	58	60	64
Số người (n_i)	4	5	20	10	8	3

Nhóm 2:

X_2 (phút)	46	48	51
Số người (n_i)	2	40	8

Tính trung bình và độ lệch tiêu chuẩn mẫu điều chỉnh của hai mẫu cụ thể nói trên. Nêu nhận xét.

ĐS: $\bar{x}_1 = 50,8$; $s'_1 = 7,735$; $\bar{x}_2 = 48,4$; $s'_2 = 1,21$.

Bài 4.3. Đo chiều cao của 100 thanh niên từ 18 đến 22 tuổi ở tỉnh A, ta thu được bảng số liệu:

Chiều cao X (cm)	154-158	158-162	162-166	166-170	170-174	174-178	178-182
Số thanh niên (n_i)	10	14	26	28	12	8	2

Hãy tính chiều cao trung bình (\bar{x}) và độ lệch mẫu điều chỉnh (s').

ĐS: $\bar{x} = 166$; $s' = 5,82$.

Bài 4.4. Các kết quả đo độ bền các sợi chỉ được cho trong bảng số liệu dưới đây (đơn vị gam):

Độ bền của sợi (gam)	120-140	140-160	160-180	180-200	200-220	220-240	240-260	260-280
Số các sợi (n_i)	1	4	10	14	12	6	2	1

Hãy xác định độ bền trung bình, phương sai và độ lệch chuẩn mẫu điều chỉnh của mẫu trên.

ĐS: $\bar{x} = 195,2$; $s^2 = 821,96$; $s' = 28,8$.

Bài 4.5. Để nghiên cứu tuổi thọ (X) của một loại bóng đèn, người ta thắp thử 100 bóng và có số liệu sau:

Tuổi thọ (giờ)	Số bóng tương ứng (n_i)
1010-1030	2
1030-1050	3
1050-1070	8
1070-1090	13
1090-1110	25
1110-1130	20
1130-1150	12
1150-1170	10
1170-1190	6
1190-1210	1

a) Vẽ biểu đồ tần số và biểu đồ tần suất của bảng phân phối thực nghiệm trên.

b) Tính tuổi thọ trung bình và độ lệch tiêu chuẩn mẫu điều chỉnh của tuổi thọ loại bóng đèn nói trên.

ĐS: b) $\bar{x} = 1111,1$; $s' = 37$.

Bài 4.6. Sau khi cải tiến kỹ thuật, người ta thắp thử 100 bóng, kết quả là:

Tuổi thọ (giờ)	1150	1160	1170	1180	1190	1200
Số bóng (n_i)	10	15	20	30	15	10

- a) Vẽ đa giác tần số, đa giác tần suất của bảng phân phối tần số trên.
 b) So sánh trung bình và độ lệch tiêu chuẩn điều chỉnh của tuổi thọ loại bóng đèn nói trên trước và sau cải tiến kỹ thuật qua hai mẫu cụ thể được cho ở bài 5 và 6.

ĐS: b) $\bar{x} = 1175,5; s' = 14,3$.

Bài 4.7. Các số liệu của việc phân tích một số mẫu quặng sắt được cho ở bảng dưới đây, ở đó X (%) là hàm lượng oxyt sắt, Y (%) là hàm lượng tạp chất.

	Y	3 - 9	9 - 15	15 - 21	21 - 27	27 - 33
X						
35				5	4	2
45			1	6	3	
55	4	6	14	2		
65	5	8	3			

- a) Tính $\bar{x}, s'_x, \bar{y}, s'_y$;
 b) Những mẫu quặng sắt có hàm lượng tạp chất nhỏ hơn 21% được xếp và loại A, hãy tính các giá trị trung bình mẫu và độ lệch tiêu chuẩn mẫu hiệu chỉnh $\bar{x}, s'_x, \bar{y}, s'_y$ của quặng sắt loại A;
 c) Tính tần suất mẫu của các quặng sắt loại A.

ĐS: a) $\bar{x} = 52,460; s'_x = 10,313; \bar{y} = 16,095; s'_y = 5,983$;

b) $\bar{x}_A = 54,808; s'_{x_A} = 9,180; \bar{y}_A = 14,192; s'_{y_A} = 4,606$;

c) $f = 0,825$.

Giải: a, ***Từ bảng trên ta có bảng tần số của X**

X x_i	35	45	55	65
Số quặng sắt n_i	11	10	26	16

Lập bảng tính:

x_i	n_i	$n_i x_i$	$n_i x_i^2$
35	11	385	13475
45	10	450	20250
55	26	1430	78650
65	16	1040	67600
	$\sum n_i = 63$	$\sum n_i x_i = 3305$	$\sum n_i x_i^2 = 179975$

Từ đó có kết quả sau:

$$\bar{x} = \frac{1}{n} \sum n_i x_i = \frac{1}{63} \cdot 3305 = 52,46.$$

$$s'_x = \sqrt{\frac{n}{n-1} \left(\frac{\sum n_i x_i^2}{n} - \bar{x}^2 \right)} = \sqrt{\frac{63}{62} \left(\frac{177975}{63} - 52,46^2 \right)} = 10,31.$$

***Từ bảng trên ta có bảng tần số của Y**

Y (y_i)	6	12	18	24	30
Số quặng sắt (m_j)	9	15	28	9	2

Lập bảng tính:

y_j	m_j	$m_j y_j$	$m_j y_j^2$
6	9	54	324
12	15	180	2160
18	28	504	9072
24	9	216	5184
30	2	60	1800
	$\sum m_j = 63$	$\sum m_j y_j = 1014$	$\sum m_j y_j^2 = 18540$

Từ đó có kết quả sau:

$$\bar{y} = \frac{1}{n} \sum m_j y_j = \frac{1}{63} \cdot 1014 = 16,095.$$

$$S'_y = \sqrt{\frac{n}{n-1} \left(\frac{\sum m_j y_j^2}{n} - \bar{y}^2 \right)} = \sqrt{\frac{63}{62} \left(\frac{18540}{63} - 16,095^2 \right)} = 5,983.$$

b, *Từ bảng và các điều kiện, ta có bảng tần số của X của nhóm A:

X(x_{i_A})	35	45	55	65
Số quặng sắt (n_{i_A})	5	7	24	16

Lập bảng tính:

x_{i_A}	n_{i_A}	$x_{i_A} n_{i_A}$	$n_{i_A} x_{i_A}^2$
35	5	175	6125
45	7	315	14175
55	24	1320	72600
65	16	1040	67600
	$\sum n_{i_A} = 52$	$\sum n_{i_A} x_{i_A} = 2850$	$\sum n_{i_A} x_{i_A}^2 = 160500$

Từ đó có kết quả sau:

$$\bar{x}_A = \frac{1}{n} \sum x_A n_{i_A} = \frac{1}{52} \cdot 2850 = 54,81.$$

$$S'_{x_A} = \sqrt{\frac{n}{n-1} \left(\frac{\sum n_{i_A} x_A^2}{n} - \bar{x}_A^2 \right)} = \sqrt{\frac{52}{51} \left(\frac{160500}{52} - 54,81^2 \right)} = 9,17.$$

*Từ bảng và các điều kiện, ta có bảng tần số của Y của nhóm A:

Y(y_A)	6	12	18
Số quặng sắt (m_{j_A})	9	15	28

Lập bảng tính:

y_A	m_{j_A}	$y_A m_{j_A}$	$m_{j_A} y_A^2$
6	9	54	324
12	15	180	2160
18	28	504	9072
	$\sum m_{j_A} = 52$	$\sum y_A m_{j_A} = 738$	$\sum m_{j_A} y_A^2 = 11556$

Từ đó có kết quả sau:

$$\bar{y}_A = \frac{1}{n} \sum y_A m_{j_A} = \frac{1}{52} \cdot 738 = 14,192.$$

$$S'_{y_A} = \sqrt{\frac{n}{n-1} \left(\frac{\sum m_{j_A} y_A^2}{n} - \bar{y}_A^2 \right)} = \sqrt{\frac{52}{51} \left(\frac{11556}{52} - 14,192^2 \right)} = 4,607.$$

c, Tần suất quặng sắt loại A là:

$$f = \frac{5 + 1 + 6 + 4 + 6 + 14 + 5 + 8 + 3}{63} = 0,825.$$

Bài 4.8. Gọi X là "khối lượng sản phẩm sản xuất ra" và Y là "giá thành sản phẩm", điều tra ở 30 xí nghiệp cùng loại sản phẩm ta có kết quả sau:

	X	50	100	150	200	250
Y						
	100 - 120				4	4
	120 - 140		2	6	1	1
	140 - 160	1	4	2		
	160 - 180	3		1		1

- a) Tính trung bình mẫu và độ lệch tiêu chuẩn mẫu điều chỉnh $\bar{x}, s'_x, \bar{y}, s'_y$
 - b) Những xí nghiệp có $X > 50$ và $120 \leq Y < 160$ được xếp vào loại I, hãy tính các giá trị trung bình mẫu và độ lệch tiêu chuẩn mẫu điều chỉnh $\bar{x}, s'_x, \bar{y}, s'_y$ của các xí nghiệp loại I.
 - c) Tính tần suất mẫu của các xí nghiệp loại I.
- ĐS:** a) $\bar{x} = 155; s'_x = 66,111; \bar{y} = 136; s'_y = 21,107;$
 b) $\bar{x}_I = 140,625; s'_{x_I} = 41,708; \bar{y}_I = 137; s'_{y_I} = 10;$
 c) $f = 0,533.$

Ước lượng các tham số lượng đặc trưng của đại lượng ngẫu nhiên

Bài 4.9. Hao phí nguyên liệu cho một đơn vị sản phẩm là một đại lượng ngẫu nhiên tuân theo quy luật chuẩn $N(\mu, \sigma_X^2)$ với độ lệch tiêu chuẩn $\sigma_X = 0,03$. Người ta sản xuất thử 36 sản phẩm và thu được bảng số liệu sau:

Mức hao phí nguyên liệu (gam)	Số sản phẩm tương ứng (n_i)
19,5 - 19,7	8
19,7 - 19,9	8
19,9 - 20,1	18
20,1 - 20,3	2

Với độ tin cậy 95%, bằng khoảng tin cậy đối xứng hãy ước lượng mức hao phí nguyên liệu trung bình cho 1 đơn vị sản phẩm.

ĐS: (19,868; 19,8876).

Bài 4.10. Để định mức thời gian gia công một chi tiết máy, người ta theo dõi ngẫu nhiên quá trình gia công 25 chi tiết và thu được số liệu sau:

Thời gian gia công (phút)	Số chi tiết máy tương ứng (n_i)
15 - 17	1
17 - 19	3
19 - 21	4
21 - 23	12
23 - 25	3
25 - 27	2

Bằng khoảng tin cậy đối xứng hãy ước lượng thời gian gia công trung bình một chi tiết máy với độ tin cậy $1 - \alpha = 0,95$. Giả thiết thời gian gia công chi tiết là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn.

ĐS: (20,53; 22,51). **Bài 4.11.** Cân thử 25 bao bột mỳ, người ta tính được trọng lượng trung bình một bao là $\bar{x} = 40kg$, độ lệch tiêu chuẩn mẫu điều chỉnh $s' = 5kg$. Với độ tin cậy 95%, bằng khoảng tin cậy đối xứng hãy ước lượng trọng lượng trung bình của bao bột mỳ. Giả thiết trong

lượng của bao bột mì là đại lượng ngẫu nhiên tuân theo quy luật chuẩn.

ĐS: (37,936; 42,064).

Bài 4.12. Để ước lượng chiều dày trung bình của các tấm vật liệu do một xí nghiệp sản xuất, người ta tiến hành đo 5 tấm và thu được kết quả sau:

2,015mm; 2,025mm; 2,015mm; 2,020mm; 2,015mm

Dựa vào số liệu trên hãy ước lượng chiều dày trung bình các tấm vật liệu do một xí nghiệp sản xuất bằng khoảng tin cậy đối xứng với độ tin cậy $1 - \alpha = 0,95$. (Biết rằng chiều dày các tấm vật liệu là đại lượng tuân theo quy luật phân phối chuẩn).

ĐS: (2,006; 2,11).

Bài 4.13. Lấy 50 con sợi để xác định độ bền trung bình, ta có số liệu sau: (biết độ bền $X - kg/cm^2$ là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn).

X	0,6-0,8	0,8-1,0	1,0-1,2	1,2-1,4	1,4-1,6	1,6-1,8	1,8-2,0	2,0-2,2	2,2-2,4
n_i	1	2	7	10	11	9	6	3	1

Hãy ước lượng độ bền trung bình của loại sợi này bằng khoảng tin cậy đối xứng với hệ số tin cậy 0,98.

ĐS: (1,386; 1,614).

Bài 4.14. Để xác định giá trung bình đối với một loại hàng hoá trên thị trường, người ta điều tra ngẫu nhiên tại 100 cửa hàng thu được bảng số liệu sau đây:

Giá X (đồng)	83	85	87	89	91	93	95	97	99	101
Số cửa hàng (n_i)	6	7	12	15	30	10	8	6	4	2

Với độ tin cậy 97% hãy ước lượng giá trung bình của loại hàng đó tại thời điểm đang xét, biết rằng giá hàng hoá X là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn.

ĐS: (89,815; 91,625).

Bài 4.15. Chiều dài (X) một loại sản phẩm do một máy tự động sản xuất là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn với độ lệch tiêu chuẩn $\sigma_X = 3cm$.

Để ước lượng chiều dài trung bình của loại sản phẩm nói trên với độ tin cậy $1 - \alpha = 0,95$, người ta tiến hành đo 20 sản phẩm và thấy chiều dài trung bình của 20 sản phẩm đó là $\bar{x} = 50cm$.

a) Tìm khoảng tin cậy đối xứng của chiều dài trung bình loại sản phẩm đó.

b) Để ước lượng chiều dài trung bình loại sản phẩm đó với độ tin cậy $1 - \alpha = 0,99$, độ dài khoảng tin cậy đối xứng không vượt quá 0,6cm thì phải đo bao nhiêu sản phẩm?

ĐS: a) (48,685; 51,315); b) 664 sp.

Bài 4.16. Kiểm tra 100 bóng điện có tuổi thọ trung bình $\bar{x} = 3000$ (giờ). Nếu ước lượng tuổi thọ trung bình của loại bóng đèn đó với độ sai lệch không quá 3,92 giờ về giá trị tuyệt đối thì độ tin cậy của ước lượng là bao nhiêu? Kết luận về tuổi thọ trung bình của loại bóng điện đó nếu tuổi thọ bóng điện là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn với $\sigma = 20$ giờ.

ĐS: a) 95%; b) (2996,08; 3003,92).

Bài 4.17. Năng suất giống ngô A ở một vùng được báo lên qua 25 điểm thu hoạch và có kết quả sau:

Năng suất (X - tạ/ha)	7	9	11	13	17
Số điểm thu hoạch (n_i)	2	7	12	3	1

Với độ tin cậy 95%, hãy tính năng suất trung bình tối thiểu của vùng này, biết rằng năng suất ngô của vùng này là đại lượng ngẫu nhiên tuân theo quy luật chuẩn.

ĐS: $\mu \geq 9,888$.

Bài 4.18. Để định mức thời gian gia công một chi tiết máy người ta theo dõi ngẫu nhiên quá trình gia công một số chi tiết và thu được bảng số liệu sau:

Thời gian X (phút)	14	16	18	20	24
Số chi tiết (n_i)	4	12	22	8	4

Với độ tin cậy $1 - \alpha = 0,99$, hãy ước lượng thời gian gia công trung bình tối đa đối với loại chi tiết đó. Giả thiết thời gian gia công chi tiết là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn.

ĐS: $\mu \leq 18,797$.

Bài 4.19. Hãy ước lượng tỷ lệ chính phẩm của một nhà máy bằng khoảng tin cậy đối xứng với độ tin cậy 0,95, biết rằng kiểm tra 100 sản phẩm của nhà máy thấy có 10 phế phẩm.

ĐS: (0,8412; 0,9588).

Bài 4.20. Mở thử 200 hộp của một kho đồ hộp, người ta thấy có 8 hộp bị biến chất. Với độ tin cậy 97%, bằng khoảng tin cậy đối xứng, hãy ước lượng tỷ lệ đồ hộp bị biến chất ở kho đó.

ĐS: (0,01; 0,07).

Bài 4.21. Gieo thử 400 hạt giống thì thấy có 20 hạt không nảy mầm. Tỷ lệ hạt giống không nảy mầm tối đa là bao nhiêu? Hãy kết luận với độ tin cậy 98%.

ĐS: $p \leq 0,072$. **Bài 4.22.** Trong đợt vận động bầu cử tổng thống người ta phỏng vấn ngẫu nhiên 1600 cử tri thì được biết 960 người trong đó sẽ bỏ phiếu cho ứng cử viên A. Với độ tin cậy 95%, ứng cử viên A chiếm được tối thiểu bao nhiêu % số phiếu bầu?

ĐS: $p \geq 0,5716$.

Bài 4.23. Để điều tra số cá trong một hồ, cơ quan quản lý đánh bắt 2000 con cá, đánh dấu rồi thả xuống hồ. Lần sau bắt lại 400 con, được 80 con có dấu. Bằng khoảng tin cậy đối xứng, hãy ước lượng số cá trong hồ với độ tin cậy 97%.

ĐS: (8230; 12739).

Bài 4.24. Bắt 1000 thú hiêm tại một vùng, đánh dấu rồi thả ra. Lần sau bắt lại được 300 con, có 60 con có dấu. Hãy ước lượng số tối đa thú hiêm đó với độ tin cậy 95%.

ĐS: $N \leq 6173$.

Bài 4.25. Cần phải lập một mẫu ngẫu nhiên với kích thước là bao nhiêu để tỷ lệ phế phẩm của mẫu là 0,2, độ dài khoảng tin cậy đối xứng là 0,05 và độ tin cậy của ước lượng là 0,95?

ĐS: $N \geq 983$.

Bài 4.26. Tỷ lệ nảy mầm của một loại hạt giống là 90%. Cần ước lượng tỷ lệ nảy mầm của hạt giống đó với độ tin cậy 0,95 và độ dài khoảng tin cậy không quá 2% thì phải gieo bao nhiêu hạt?

ĐS: $N \geq 3456$.

Phụ lục A

Giải tích tổ hợp

A.1. Các quy tắc đếm

A.1.1. Quy tắc cộng

Nếu có m_1 cách chọn đối tượng x_1 , m_2 cách chọn đối tượng x_2, \dots, m_n cách chọn đối tượng x_n ; và cách chọn x_i không trùng cách chọn $x_j (i \neq j)$ thì có:

$$m_1 + m_2 + \dots + m_n$$

cách chọn một trong các đối tượng x_1, x_2, \dots, x_n .

A.1.2. Quy tắc nhân

Giả sử một phép chọn được thực hiện qua n bước liên tiếp. Bước 1 có m_1 cách thực hiện, bước 2 có m_2 cách, ..., bước n có m_n cách thực hiện. Khi đó số cách thực hiện phép chọn là:

$$m_1 \cdot m_2 \cdot \dots \cdot m_n.$$

A.2. Hoán vị - Chỉnh hợp - Tổ hợp

A.2.1. Chỉnh hợp (chỉnh hợp không lặp)

★ **Định nghĩa A.1** Cho tập hợp E gồm n phần tử. Chỉnh hợp chập k của n phần tử đã cho là một dãy có thứ tự gồm k ($0 < k \leq n$) phần tử khác nhau lấy từ n phần tử đã cho.

● **Ví dụ A.1** Cho tập hợp $E = a, b, c$. Các chỉnh hợp chập 2 của 3 phần tử của E là:

$$ab, ac, ba, bc, ca, cb$$

Vậy có 6 chỉnh hợp chập 2 của 3.

Công thức tính. Ký hiệu số chỉnh hợp chập k của n là A_n^k , ($0 < k \leq n$).

Để có một chỉnh hợp chập k ta có thể tiến hành chọn theo các bước như sau: Chọn phần tử đầu theo n cách, sau đó chọn phần tử thứ hai theo $n-1$ cách, chọn phần tử đứng thứ k theo $n-k+1$ cách. Vậy số chỉnh hợp chập k của n là:

$$A_n^k = \underbrace{n(n-1)\dots(n-k+1)}_{k \text{ thừa số}} = \frac{n!}{(n-k)!}.$$

•**Ví dụ A.2** Có bao nhiêu số tự nhiên có 3 chữ số khác nhau mà chữ số nào cũng là số lẻ?

Lời giải. Mỗi số tự nhiên thoả mãn yêu cầu là một bộ 3 chữ số khác nhau (có kể thứ tự) chọn từ 5 chữ số $\{1, 3, 5, 7, 9\}$. Vậy số các số tự nhiên thoả mãn là:

$$A_5^3 = 5.4.3 = 60 \text{ (số)}.$$

A.2.2. Chỉnh hợp lặp

★ **Định nghĩa A.2** Cho tập hợp E gồm n phần tử. Chỉnh hợp lặp chập k của n phần tử đã cho là một dãy có thứ tự gồm k ($k > 0$) phần tử khác nhau lấy từ n phần tử đã cho, trong đó mỗi phần tử có thể có mặt một hay nhiều lần trong dãy tạo thành.

•**Ví dụ A.3** Cho tập hợp $E = \{a, b, c\}$. Các chỉnh hợp lặp chập 2 của 3 phần tử của E là:

$$aa, ab, ac, ba, bb, bc, ca, cb, cc$$

Vậy có 9 chỉnh hợp lặp chập 2 của 3.

Công thức tính. Ký hiệu số chỉnh hợp lặp chập k của n là \tilde{A}_n^k , ($k > 0$).

Để có một chỉnh hợp lặp chập k ta có thể tiến hành chọn theo các bước như sau: Chọn phần tử đầu theo n cách, sau đó chọn phần tử thứ hai theo n cách, ..., chọn phần tử đứng thứ k theo n cách. Vậy số chỉnh hợp lặp chập k của n là:

$$\tilde{A}_n^k = \underbrace{n.n\dots n}_k \text{ thừa số} = n^k.$$

•**Ví dụ A.4** Để truyền tin bằng tín hiệu moóc-xơ gồm hai ký hiệu chấm (.) và vạch (-), người ta mã hoá mỗi chữ cái của bảng chữ cái thành một dãy có thứ tự gồm không quá 4 ký hiệu. Biết rằng cùng một ký hiệu có thể có mặt nhiều lần trong dãy có thứ tự tạo thành. Hỏi có thể mã hoá được bao nhiêu chữ cái?

Lời giải. Mỗi nhóm có thứ tự gồm k ký hiệu ($1 \leq k \leq 4$) tạo nên chính là một chỉnh hợp lặp chập k từ 2 phần tử đã cho (hai phần tử này là các ký hiệu (.) và (-)). Vì vậy số chữ cái mã hoá được là:

$$\tilde{A}_2^1 + \tilde{A}_2^2 + \tilde{A}_2^3 + \tilde{A}_2^4 = 2^1 + 2^2 + 2^3 + 2^4 = 30.$$

Như vậy nếu bảng chữ cái của một thứ tiếng nào đó gồm không quá 30 chữ cái thì ta có thể mã hoá theo cách trên.

A.2.3. Hoán vị

★ **Định nghĩa A.3** Cho tập hợp E gồm n phần tử. Hoán vị của n phần tử là một dãy có thứ tự gồm đủ n phần tử đã cho.

•**Ví dụ A.5** Cho tập hợp $E = \{a, b, c\}$. Các hoán vị của 3 phần tử của E là:

$$abc, acb, bac, bca, cab, cba$$

Vậy có 6 hoán vị của 3 phần tử.

Công thức tính. Ký hiệu số hoán vị của n phần tử là P_n .

Để thấy một hoán vị của n phần tử chính là một chỉnh hợp chập n của n phần tử. Vì vậy ta tính được số hoán vị của n phần tử là:

$$P_n = A_n^n = n(n-1)\dots 1 = n!$$

•**Ví dụ A.6** Một bàn có 5 học sinh ngồi. Khi đó có tất cả: $P_5 = 5! = 120$ cách sắp xếp chỗ ngồi.

A.2.4. Tổ hợp

★ **Định nghĩa A.4** Cho tập hợp E gồm n phần tử. Tổ hợp chập k từ n phần tử ($k \leq n$) là tập con gồm k phần tử (phân biệt, không kể thứ tự) của tập n phần tử đã cho.

• **Ví dụ A.7** Cho tập hợp $E = \{a, b, c\}$. Các tổ hợp chập 2 của 3 phần tử là:

$$\{a, b\}; \{b, c\}; \{a, c\}.$$

Vậy có 3 tổ hợp chập 2 của 3.

Công thức tính. Ký hiệu số tổ hợp chập k của n là C_n^k , ($0 \leq k \leq n$).

Giả sử từ n phần tử đã cho ta tạo nên C_n^k tổ hợp. Dem mỗi tổ hợp chập k này hoán vị theo mọi cách ta sẽ được $k!$ chỉnh hợp chập k . Mặt khác, hai chỉnh hợp chập k tạo ra từ hai tổ hợp chập k phân biệt là khác nhau. Vậy số chỉnh hợp chập k của n phần tử đã cho là:

$$A_n^k = k!.C_n^k$$

Vì vậy:

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!.(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1}$$

Tính chất thường gặp về tổ hợp:

$$\begin{aligned} C_n^k &= C_n^{n-k} \\ C_n^k &= C_{n-1}^k + C_{n-1}^{k-1}. \end{aligned}$$

• **Ví dụ A.8** Có 10 đội bóng thi đấu với nhau theo thể thức đấu vòng tròn một lượt. Hỏi phải tổ chức bao nhiêu trận đấu?

Lời giải. Mỗi trận đấu ứng với một nhóm gồm 2 đội (phân biệt, không kể thứ tự) từ 10 đội. Vì vậy phải tổ chức tất cả:

$$C_{10}^2 = \frac{10.9}{2.1} = 45 \text{ (trận đấu).}$$

Bài tập phụ lục A

Bài A.1. Một nhóm sinh viên gồm 3 sinh viên nữ và 4 sinh viên nam. Hỏi có bao nhiêu cách chọn từ nhóm sinh viên đó:

- 3 sinh viên vào Ban chấp hành Đoàn.
- 3 sinh viên vào Ban cán sự lớp, trong đó có 1 lớp trưởng, 1 lớp phó phụ trách học tập, 1 lớp phó phụ trách đời sống.
- 2 sinh viên nam và 2 sinh viên nữ đi tập văn nghệ.
- 3 sinh viên đi dự Đại hội, trong đó có ít nhất 1 sinh viên nữ.

Đáp số: a) 35; b) 210; c) 18; d) 31.

Bài A.2. Biển số xe ô tô của Hải Phòng có dạng: 16 - 2 chữ cái - 4 chữ số.

- Hỏi trên lý thuyết có thể lập được bao nhiêu biển xe? (các chữ cái được lấy từ bảng chữ cái tiếng Anh gồm 26 chữ cái).
- Trong đó có bao nhiêu biển xe mà tổng 4 chữ số là một số có tận cùng là 9?

Đáp số: a) $26^2 \cdot 10^4$; b) 676000.

Bài A.3. Có bao nhiêu cách xếp chỗ cho 5 người A, B, C, D, E:

- vào một chiếc ghế dài sao cho A, B ở hai đầu ghế?
- vào một chiếc ghế dài sao cho A, B, C luôn ở cạnh nhau?
- vào một bàn tròn sao cho A, B luôn ở cạnh nhau?

Đáp số: a) 12; b) 36; c) 12.

Bài A.4. Có 20 điểm phân biệt trong mặt phẳng trong đó có 8 điểm thẳng hàng. Nối 20 điểm đó lại với nhau. Hỏi:

- có bao nhiêu đường thẳng phân biệt?
- có bao nhiêu tam giác?

Đáp số: a) 163; b) 1084.

Bài A.5. Cho phương trình: $x + y + z = 12$

- Tìm số nghiệm nguyên dương của phương trình sao cho x, y, z đôi một phân biệt.
- Tìm số nghiệm nguyên dương của phương trình.
- Tìm số nghiệm nguyên không âm của phương trình.

Đáp số: a) 42; b) 55; c) 91.

Bài A.6. Cho tập hợp $A = \{0, 1, 2, 3, 4, 5, 6\}$. Từ A có thể tạo ra bao nhiêu:

- số tự nhiên chẵn 4 chữ số?
- số tự nhiên chẵn 4 chữ số đôi một phân biệt.
- số tự nhiên 7 chữ số, trong đó chữ số 2 có mặt đúng 2 lần, số 3 có mặt đúng 3 lần, các chữ số khác có mặt không quá 1 lần?

Đáp số: a) 1176; b) 420; c) 3960.

Bài A.7. Thầy giáo có 12 cuốn sách đôi một khác nhau trong đó có 5 cuốn sách Toán, 4 cuốn sách tiếng Anh và 3 cuốn sách Triết học. Thầy muốn lấy ra 6 cuốn và đem tặng cho 6 sinh viên khác nhau mỗi em một cuốn.

- Giả sử thầy giáo chỉ muốn tặng các cuốn sách thuộc hai thể loại Toán và tiếng Anh, hỏi có tất cả bao nhiêu cách tặng?
- Giả sử thầy giáo muốn rằng sau khi tặng sách xong, mỗi một trong ba thể loại đều còn lại ít nhất một cuốn. Hỏi có tất cả bao nhiêu cách tặng?

Đáp số: a) 60480; b) 579600.

Bài A.8. Một bàn dài có hai dãy ghế đối diện nhau, mỗi dãy gồm 6 ghế. Người ta muốn xếp chỗ ngồi cho 6 sinh viên lớp A và 6 sinh viên lớp B vào bàn nói trên. Hỏi có bao nhiêu cách xếp trong mỗi trường hợp sau:

- Bất cứ 2 sinh viên nào ngồi cạnh nhau hoặc đối diện nhau thì khác lớp nhau.
- Bất cứ 2 sinh viên nào ngồi đối diện nhau thì khác lớp nhau.

Đáp số: a) 1036800; b) 33177600.

Phụ lục B

Sử dụng CNTT giải toán thống kê

Tất cả các SV ở năm thứ nhất đều đã được trang bị những kiến thức cơ bản của Tin học. Việc giải các bài toán bằng máy tính điện tử (MTĐT) cũng đã trở nên quen thuộc đối với các em.

Trong quá trình ngồi trên ghế nhà trường, SV không chỉ tiếp thu các tri thức thông qua bài giảng của GV mà họ còn có thể tự mình tìm kiếm, kiến tạo ra những tri thức mới thông qua sách báo, tài liệu, mạng internet... Việc GV dành ra thời gian để hướng dẫn cho SV giải các bài toán thống kê bằng cách sử dụng các công cụ của CNTT sẽ tăng thêm khả năng hứng thú, tính tò mò trong khoa học, kích thích hoạt động tìm tòi khám phá của SV, từ đó dẫn tới việc hình thành kiến thức và kỹ năng mới.

Sau đây chúng tôi trình bày cách thức thể hiện của một số công cụ CNTT hỗ trợ cho việc giải các bài toán thống kê, kết hợp cùng các ví dụ tương ứng.

B.1. Đối với máy tính điện tử cầm tay

B.1.1. Tính các đặc trưng của mẫu

•**Ví dụ B.1** Gặt ngẫu nhiên 365 điểm trồng lúa của một huyện ta có bảng số liệu sau:

X	25	30	33	34	35	36	37	39	40
n	6	13	38	74	106	85	30	10	3

Tính $\bar{x}, S^2, \hat{S}^2, S, \hat{S}$.

Đối với Casio FX570ES

Nhập các số liệu

- Ấn các phím: Shift Mode 4 1
- Ấn các phím: Mode 3
- Ấn các phím: 1 1 - Var

Hiện ra bảng:

x	freq

Dùng phím Replay để di chuyển qua lại giữa 2 cột x và freq

Nhập cột x:

25 =
 30 =
 33 =
 34 =
 35 =
 36 =
 37 =
 39 =
 40 =

Nhập cột freq: Dùng phím Replay di chuyển qua cột freq, dòng 25

6 =
 13 =
 38 =
 74 =
 106 =
 85 =
 30 =
 10 =
 3 =

Nhập xong ấn phím AC Xem kết quả:

- Ấn các phím: Shift 1 5 1 = $\rightarrow n = 365$ (cỡ mẫu).
- Ấn các phím: Shift 1 5 2 = $\rightarrow \bar{X} = 34,795$ (tạ/ha).
- Ấn các phím: Shift 1 5 4 = $\rightarrow \hat{S} = 2,072$ (độ lệch chuẩn mẫu đã hiệu chỉnh).
- Ấn các phím: Shift 1 5 3 = $\rightarrow S = 2,069$ (độ lệch chuẩn mẫu chưa hiệu chỉnh).

* **Chú ý:** khi nhập số liệu ta có thể dùng các phím để:

- Xoá 1 dòng dữ liệu: di chuyển đến dòng cần xoá và ấn Del
- Chèn thêm 1 dòng dữ liệu: di chuyển đến dòng cần chèn và nhấn các phím
Shift 1 3 1 Ins
- Xoá toàn bộ nội dung đã nhập
Shift 1 3 2 Del - A

Đối với các loại máy $fx - 500MS$, $fx - 350MS$, $fx - 350TL$, $fx - 570MS$

Nhập các số liệu:

- Ấn các phím: MODE 2 để chuyển máy sang chế độ SD
hoặc MODE MODE 2 để chuyển máy sang chế độ SD
- Ấn các phím: 25 Shift , 6 data (Phím M_+)
- Ấn các phím: 30 Shift , 13 data
- Ấn các phím: 33 Shift , 38 data

- Ấn các phím: 34 Shift , 74 data
- Ấn các phím: 35 Shift , 106 data
- Ấn các phím: 36 Shift , 85 data
- Ấn các phím: 37 Shift , 30 data
- Ấn các phím: 39 Shift , 10 data
- Ấn các phím: 40 Shift , 3 data

Xem kết quả:

- Ấn các phím: Shift S - VAR 1 = $\rightarrow \bar{X} = 34,795$
- Ấn các phím: Shift S - VAR 2 = $\rightarrow \hat{S} = 2,072$
- Ấn các phím: Shift S - VAR 3 = $\rightarrow S = 2,069$

* **Chú ý:** Xoá SD bằng lệnh: Shift MODE 2 =
Đối với các loại máy $fx - 220MS$, $fx - 500A$, $fx - 95$, $fx - 82Super$
 Nhập số liệu:

- Ấn các phím: MODE . (Dấu chấm) để đưa máy về chế độ SD
- Ấn các phím: Shift Sac 26×6 data (Phím M_+)
- Ấn các phím: 30×13 data
- Ấn các phím: 33×38 data
- Ấn các phím: 34×74 data
- Ấn các phím: 35×106 data
- Ấn các phím: 36×85 data
- Ấn các phím: 37×30 data
- Ấn các phím: 39×10 data
- Ấn các phím: 40×3 data

Xem kết quả:

- Ấn các phím: Shift $n \rightarrow 365$
- Ấn các phím: Shift $\bar{X} \rightarrow 34,795$
- Ấn các phím: Shift $\hat{S} \rightarrow 2,072$
- Ấn các phím: Shift $S \rightarrow 2,069$

* **Chú ý:** Xoá SD bằng lệnh: MODE 0

B.1.2. Bài toán tìm hàm hồi quy

Vào **REG** ấn:

MODE **3** (máy fx 500MS, 95MS)

Máy khác

MODE **MODE** **2** (fx 570MS)

- Màn hình hiện:

Lin	Log	Exp →
1	2	3

▶ ↑ ↓ ◀

← Pwr	Inv	Quad
1	2	3

- Ấn số tương ứng ta sẽ và chức năng muốn chọn.

1 (**Lin**) (tuyến tính) $y = A + Bx$

2 (**Log**) (logarit) $y = A + B \ln x$

3 (**Exp**) (mũ) $y = Ae^{bx}$ ($\ln y = \ln A + Bx$)

▶ **1** (**Pwr**) (lũ thừa) $y = Ae^b$ ($\ln y = \ln A + B \ln x$)

▶ **2** (**Inv**) (nghịch đảo) $y = A + B.1/x$

▶ **3** (**Quad**) (bậc hai) $y = A + Bx + Cx^2$

Trước khi tính toán phải ấn

SHIFT **CLR** **1** (**Sel**) **=**

để xoá bộ nhớ của thống kê.

- Nhập dữ liệu theo cú pháp: <dữ liệu>'<dữ liệu y>**DT**.

- Các kết quả theo dữ liệu đã nhập được gọi theo bảng sau:

Giá trị cần gọi	Ấn
$\sum x^2$	SHIFT S.SUM 1
$\sum x$	SHIFT S.SUM 2
n	SHIFT S.SUM 3
$\sum y^2$	SHIFT S.SUM ► 1
$\sum y$	SHIFT S.SUM ► 2
$\sum xy$	SHIFT S.SUM ► 3
$\sum x$	SHIFT S.SUM ► ► 1
$\sum x^2y$	SHIFT S.SUM ► ► 2
$\sum x^4$	SHIFT S.SUM ► ► 3
\bar{x}	SHIFT S.VAR 1
$x\sigma_n$	SHIFT S.VAR 2
$x\sigma_{n-1}$	SHIFT S.VAR 3
\bar{y}	SHIFT S.VAR ► 1
$y\sigma_n$	SHIFT S.VAR ► 2
$y\sigma_{n-1}$	SHIFT S.VAR ► 3
Hệ số A	SHIFT S.VAR ► ► 1
Hệ số B	SHIFT S.VAR ► ► 2

- Trừ $y = A + Bx + Cx^2$

Gọi	Ấn
Hệ số tương quan r	SHIFT S.VAR ► ► 3
Số dự đoán \hat{x}	SHIFT S.VAR ► ► ► 1
Số dự đoán \hat{y}	SHIFT S.VAR ► ► ► 2

- Riêng với $y = A + Bx + Cx^2$ thì theo bảng sau:

Gọi	Ấn
Hệ số C	SHIFT S.VAR ▶ ▶ 3 =
Số dự đoán \hat{x}_1	SHIFT S.VAR ▶ ▶ ▶ 1 =
Số dự đoán \hat{x}_2	SHIFT S.VAR ▶ ▶ ▶ 2 =
Số dự đoán \hat{y}	SHIFT S.VAR ▶ ▶ ▶ 3 =

- Các giá trị này có thể dùng như các biến trong các biểu thức tính.

* Hồi quy tuyến tính $y = A + Bx$

• Ví dụ B.2 Áp suất theo nhiệt độ trong bảng sau:

Nhiệt độ	10°C	15°C	20°C	25°C	30°C
Áp suất	1003hpa	1005hpa	1010hpa	1011hpa	1014hpa

Hãy dùng Hồi quy tuyến tính $y = A + Bx$ để tính A, B và hệ số tương quan r , áp suất ở 18°C. Tìm nhiệt độ khi áp suất 1000hpa, hệ số tới hạn r^2 và số hiệp biến $\left(\frac{\sum xy - n\bar{x}\bar{y}}{n-1}\right)$.

Vào Mode **RED**

Ấn 1 (**Lin**)

SHIFT **CLR** **1** (**Sel**) **=** (xoá nhớ)

10 **•** 1003 **DT** n = REG
1

(Khi ấn **DT** dữ liệu được nhập và màn hình hiện giá trị của n).

Ấn tiếp

15 **'** 1005 **DT**
 20 **'** 1010 **DT** 25 **'** 1011 **DT**
 30 **'** 1014 **DT**

Hệ số $A = 997.4$

SHIFT **S.VAR** **▶** **▶** **1** **=**

Hệ số $B = 0.56$

SHIFT **S.VAR** **▶** **▶** **2** **=**

Hệ số tương quan $r = 0.982607368$

SHIFT **S.VAR** **▶** **▶** **3** **=**

Tìm áp suất ở 18°C = 1007.48

18 **SHIFT** **S.VAR** **▶** **▶** **▶** **2** **=**

Nhiệt độ ở 1000 hpa = 4.642857143

1000 **SHIFT** **S.VAR** **▶▶▶** **1** **=**

$r^2 = 0.965517241$

SHIFT **S.VAR** **▶▶** **3** **x²** **=**

Số hiệp biến = 35

(**SHIFT** **S.SUM** **▶** **3** **-** **SHIFT** **S.SUM** **3** **x** **SHIFT**
S.VAR **1** **x** **SHIFT** **S.VAR** **▶** **1** **)** **÷** **(** **SHIFT**
S.SUM **3** **-** **1** **)** **=**

* Hồi quy bậc hai: $A + Bx + Cx^2$.

• Ví dụ B.3 Cho bảng sau

x_i	29	50	74	103	118
y_i	1.6	23.5	38.0	46.4	48

Theo công thức hồi quy bậc hai hãy tìm các hệ số A, B, C sau đó tìm \hat{y} với $x = 16$ và \hat{x} với $y = 20$.

Ở Mode RED

Ấn **▶ 3** (Quad)

SHIFT **CLR** **1** (Sel) **=** (xoá nhớ thống kê)

29 **'** 1.6 **DT** 50 **'** 23.5 **DT**

74 **'** 38.0 **DT** 103 **'** 46.4 **DT**

118 **'** 48.0

Tính hệ số $A = -35.59856934$

SHIFT **S.VAR** **▶▶** **1** **=**

Tính hệ số $B = 1.4959856934$

SHIFT **S.VAR** **▶▶** **2** **=**

Tính hệ số $C = -6.71629667 \times 10^{-3}$

SHIFT **S.VAR** **▶▶** **3** **=**

Tính \hat{y} khi $x_i = 16$ (=13.38291067)

16 **SHIFT** **S.VAR** **▶▶▶** **3** **=**

Tính \hat{x}_1 khi $y_i = 20$ (=47.14556728)

20 **SHIFT** **S.VAR** **▶▶▶** **1** **=**

Tính \hat{x}_2 khi $y_i = 20$ (=175.5872105)

20 **SHIFT** **S.VAR** **▶▶▶** **2** **=**

Chú ý về nhập dữ liệu

- Ấn **DT DT** để nhập dữ liệu hai lần.
- Dùng phím **SHIFT** ; để nhập nhiều dữ liệu giống nhau. Ví dụ nhập "20, 30" năm lần thì ấn **20 ' 30 SHIFT ; 5 DT**.
- Kết quả được gọi không cần thứ tự như bảng trên.
- Vẫn có phần chú ý như ở **SD**.
- Xem thêm phần tính $P(t), Q(t), R(t)$ ở phần bổ sung cho máy Casio fx 570 MS.

- Các số nhớ **A, B, C, D, E, F, X, Y** không dùng được thống kê **SD** và hồi quy **RED**. Khi vào các chương trình này, các giá trị số nhớ được gán trước bị xoá hết.

B.2. Dùng phần mềm Excel

B.2.1. Tính toán trong bài toán ước lượng

•**Ví dụ B.4** Bảng số liệu sau cho độ lệch X của 200 chi tiết máy với kích thước chuẩn. Với mức ý nghĩa $\alpha = 0,05$ hãy kiểm tra giả thiết độ lệch X là biến ngẫu nhiên có phân phối chuẩn $N(a, \sigma^2)$.

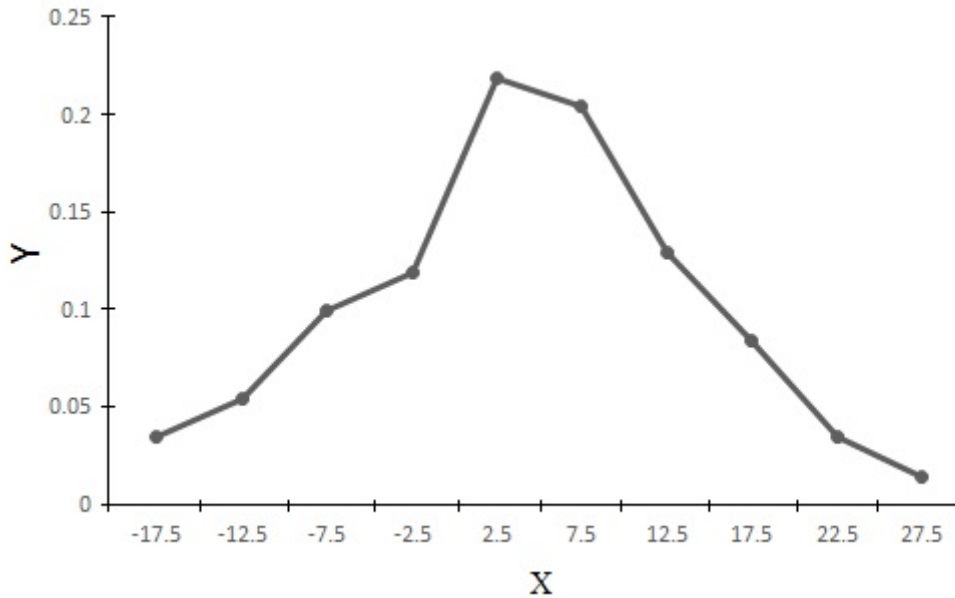
Thứ tự	Khoảng (x_{i-1}, x_i)	n_i	Thứ tự	Khoảng (x_{i-1}, x_i)	n_i
1	(-20, -15)	7	6	(5, 10)	41
2	(-15, -10)	11	7	(10, 15)	26
3	(-10, -5)	20	8	(15, 20)	17
4	(-5, 0)	24	9	(20, 25)	7
5	(0, 5)	44	10	(25, 30)	3

Lời giải. Giả thiết $H_0 : X \sim N(a, \sigma^2)$, đối thiết $H_1 : X \sim N(a, \sigma^2)$

+) Lập bảng Excel ước lượng các tham số $a = \bar{X}$ và $\sigma = S'$ và vẽ đa giác tần suất:

TT	A	B	C	D	E	F	G
1	TT	(x_{i-1}, x_i)	x_i	n_i	f_i	$n_i \cdot x_i$	$n_i (x_i - \bar{X})^2$
2	1	(-20, -15)	-17.5	7	0.035	-122.5	3250.818
3	2	(-15, -10)	-12.5	11	0.055	-137.5	3012.928
4	3	(-10, -5)	-7.5	20	0.1	-150	2668.05
5	4	(-5, 0)	-2.5	24	0.12	-60	1029.66
6	5	(0, 5)	2.5	44	0.22	110	105.71
7	6	(5, 10)	7.5	41	0.205	307.5	488.0025
8	7	(10, 15)	12.5	26	0.13	325	1856.465
9	8	(15, 20)	17.5	1	0.085	297.5	3075.343
10	9	(20, 25)	22.5	7	0.035	157.5	2382.818
11	10	(25, 30)	27.5	3	0.015	82.5	1649.708
12				200	1	810	19519.5
13			$\bar{X} = 810/200 =$			4.05	
14			$S' = \sqrt{19519.5/199} =$				9.903936

Chọn cột E làm trục tâm, cột C làm trục hoành ta có đa giác tần suất. Đồ thị cho ta hình dạng của hàm mật độ của phân phối chuẩn.



+) Thực hiện chia khoảng: Vì $n_{10} = 3 < 5$ nên ta ghép khoảng 10 vào khoảng 9, vậy ta có bảng:

Thứ tự	Khoảng (x_{i-1}, x_i)	n_i	Thứ tự	Khoảng (x_{i-1}, x_i)	n_i
1	< -15	7	6	$(5, 10)$	41
2	$(-15, -10)$	11	7	$(10, 15)$	26
3	$(-10, -5)$	20	8	$(15, 20)$	17
4	$(-5, 0)$	24	9	$(20, 25)$	10
5	$(0, 5)$	44			

Như vậy ta có $k = 9, r = 2$.

Lập bảng Excel tính Q_n và $\chi_{6,0.05}^2$, trong bảng Excel cột B ta chỉ ghi các cận dưới của các khoảng.

TT	A	B	C	D	E
1			n_i	p_i	$(n_i - np_i)^2 / np_i$
2	1	-15	7	0.02721	0.44604605
3	2	-10	11	0.050794	0.06966124
4	3	-5	20	0.102413	0.01136968
5	4	0	24	0.160879	2.07749498
6	5	5	44	0.196912	0.5413946
7	6	10	41	0.187795	0.31523907
8	7	15	26	0.139551	0.13073038
9	8	20	17	0.080798	0.04371067
10	9		10	0.053648	0.04960009
11			200	$Q_n =$	3.68524676
12				$\chi_{6,0.05}^2$	12.5915774

Để tính các xác suất $p_i (i = \overline{1, 9})$ ta thực hiện các lệnh sau:

D2 = NORMDIST(B2,4.05,9.903936,TRUE)

D3 = NORMDIST(B3,4.05,9.903936,TRUE)-NORMDIST(B3,4.05,9.903936,TRUE)

Sau đó dùng Autofill từ D4 đến D9

D10 = 1-NORMDIST(B9,4.05,9.903936,TRUE)

Ta cũng có thể tra bảng hàm Laplace tính các xác suất P_i trên các khoảng

$$(x_{i-1}, x_i) (i = \overline{1, 9}) \text{ với: } P_i = \phi\left(\frac{x_i - 4,05}{9,9039}\right) - \phi\left(\frac{x_{i-1} - 4,05}{9,9039}\right).$$

Vậy $Q_n = 3,68524776 < \chi_{6,0.05}^2 = 12,59$ chấp nhận H_0 , hay X có phân phối chuẩn $N(4,3;94,26)$.

B.2.2. Tính toán các đặc trưng của mẫu

•**Ví dụ B.5** Do chiều cao của 100 sinh viên năm thứ nhất ta được bảng thống kê (tính bằng cm).

X	150-154	154-158	158-162	162-166	166-170	170-174	174-178
SV	10	14	24	28	12	8	4

Chọn $x_0 = 160$ và lập bảng tính Excel ta được:

TT	A	B	C	D	E
1	X	n_i	$x_i - x_0$	$n_i(x_i - x_0)$	$n_i(x_i - x_0)^2$
2	152	10	-8	-80	640
3	156	14	-4	-56	224
4	160	24	0	0	0
5	164	28	4	112	448
6	168	12	8	96	768
7	172	8	12	96	1152
8	176	4	16	64	1024
9		n = 100		B = 232	C = 4256
10	$\bar{X} = \frac{B}{n} + x_0 =$			162.32	
11	$S^2 = \frac{C}{n} - (\bar{X} - x_0)^2 =$			37.1176	
12	S =			6.09734	

•**Ví dụ B.6** Hai máy A và B cùng sản xuất một chi tiết có độ dài X. Để kiểm tra độ chính xác của mỗi máy, định kỳ 6 tháng một lần người ta lấy mỗi máy một lô gồm 1000 sản phẩm và kiểm tra độ dài của chúng. Nếu độ lệch trung bình của mỗi lô vượt quá 5 mm thì máy cần duy tu lại. Bảng thống kê một lần của hai máy được cho dưới đây:

Độ dài	1222	1226	1230	1234	1238	1242	1246	1250	1254
	1226	1230	1234	1238	1242	1246	1250	1254	1256
n(A)	5	20	80	190	360	300	30	10	5
m(B)	70	90	110	140	170	170	140	90	20

Hãy kiểm tra xem máy nào đã cần duy tu lại.

Lời giải. Chọn $x_0 = 1236$ và lập bảng tính theo Excel:

X	x_i	$x_i - x_0$	$n_i A$	$m_i B$	$n_i (x_i - x_0)$	$m_i (x_i - x_0)$	$n_i (x_i - \bar{A})^2$	$m_i (x_i - \bar{B})^2$	
1222-1226	1224	-12	5	70	-60	-840	720	10080	
1226-1230	1228	-8	20	90	-160	-720	1280	5760	
1230-1234	1232	-4	80	110	-320	-440	1280	1760	
1234-1238	1236	0	190	140	0	0	0	0	
1238-1242	1240	4	360	170	1440	680	5760	2720	
1242-1246	1244	8	300	170	2400	1360	19200	10880	
1246-1250	1248	12	30	140	360	1680	4320	20160	
1250-1254	1252	16	10	90	160	1440	2560	23040	
1254-1258	1256	20	5	20	100	400	2000	8000	
			1000	1000	B1=3920	B2=3560	C1=37120	C2=82400	
$\bar{A} = \frac{B_1}{n} + x_0 =$				1239.92	$\bar{B} = \frac{B_2}{n} + x_0 =$				1239.56
$S_A^2 = \frac{C_2}{n} - (\bar{A} - x_0)^2 =$				21.7536	$S_A'^2 = \frac{n}{n-1} \cdot S_A^2 =$				21.77538
$S_B^2 = \frac{C_2}{n} - (\bar{B} - x_0)^2 =$				69.73	$S_B'^2 = \frac{n}{n-1} \cdot S_B^2 =$				69.7962
$S_A' =$				4.6664	$S_B' =$				8.3544

Như vậy máy A không cần duy tu còn máy B phải duy tu.

B.2.3. Các phân phối xác suất

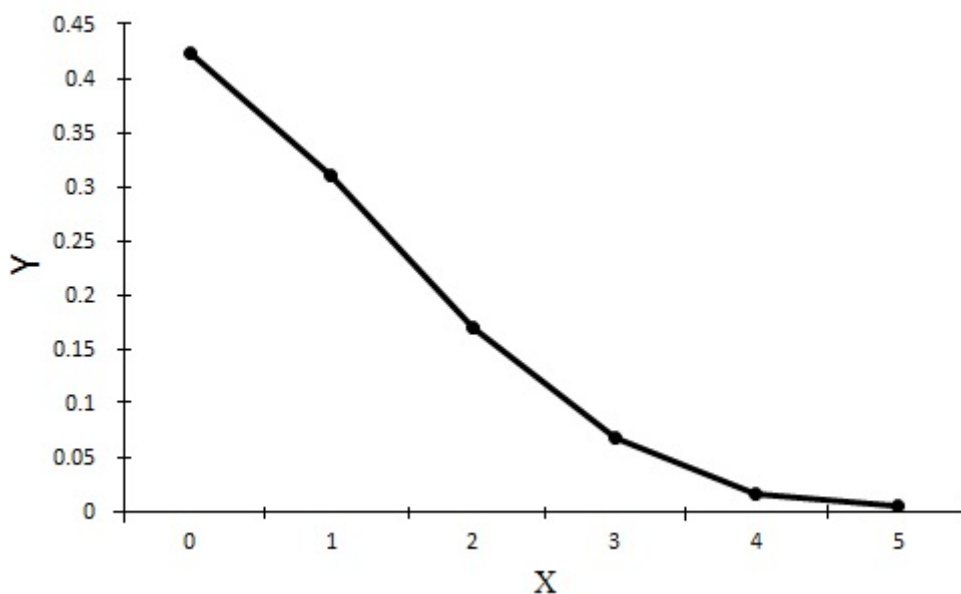
•**Ví dụ B.7** Quan sát số lượng ô tô vào xúc tại một máy xác trong khoảng thời gian định sẵn tại một mỏ than người ta thu được bảng kết quả sau:

Số ô tô	0	1	2	3	4	5
Số khoảng thời gian	75	55	30	12	3	1

Hãy kiểm tra xem số liệu trên có tuân theo quy luật phân phối Poát-Xông $p_m = \frac{\lambda^m}{m!} e^{-\lambda}$ không? Trong đó P_m là xác suất trong khoảng thời gian trên có m ô tô xuất hiện tại máy xúc.

Lời giải. Lập bảng Excel ước lượng tham số $\lambda = \bar{X}$ và vẽ đa giác tần suất, gọi số ô tô là X, số khoảng thời gian là n_i :

X	0	1	2	3	4	5	
n_i	75	55	30	12	3	1	176
f_i	0.426136	0.3125	0.170455	0.068182	0.017045	0.005682	1
	$\lambda \approx \bar{X} =$			0.954545			



Đa giác tần suất có dạng hàm mật độ của phân phối Poát-Xông. Các xác suất $p_m = \frac{1}{m!}e^{-0.621212}$ lập bảng tính Q_n ta được:

TT	A	B	C	D
1	X	n_i	p_i	$(n_i - np_i)^2 / np_i$
2	0	75	0.384987	0.774082
3	1	55	0.367488	1.448107
4	2	30	0.175392	0.024461
5	3	12	0.055806	0.482998
6	4	3	0.013317	0.183674
7	5	1	0.002542	0.682274
8		176	$Q_n =$	3.595594
9			$\chi_{4,0.05}^2 = CHIINV(0.05, 4) =$	9.487728

Để tính các p_i trong Excel ta thực hiện:

C2 = POISSON(A2,0.954545,FALSE)

Sau đó Autofill từ C3 đến C7.

Số bậc tự do $k = 6$, $r = 1$ ta có $k - r - 1 = 6 - 1 - 1 = 4$.

$$\chi_{4,0.05}^2 = 9,487728 > Q_n = 3,595594.$$

Vậy bảng số liệu thu được tuân theo phân phối Poát-Xông.

Trong phần này chúng tôi đã giới thiệu cách sử dụng một số máy tính điện tử cá nhân thông dụng và phần mềm Excel để giải một số bài toán trong thống kê. Với sự phát triển mạnh mẽ của công nghệ thông tin, ngày càng nhiều thế hệ máy tính và phần mềm mới ra đời, đó sẽ là một thuận lợi lớn cho SV. Do đó yêu cầu sử dụng được MTĐT các nhân và các phần mềm hỗ trợ giải toán là một điều cần thiết. Các công cụ này sẽ giúp chúng ta giải quyết nhanh chóng các bài toán nói chung và thống kê nói riêng, góp phần nâng cao tính chính xác và kịp thời; đồng thời tiết kiệm thời gian tính toán cho một bài toán, rèn luyện tư duy thuật giải.

TÀI LIỆU THAM KHẢO

- [1] Hoàng Văn Hùng, *Bài giảng Toán Cao cấp*, Nhà xuất bản Đại học Hàng Hải Việt Nam - Tài liệu lưu hành nội bộ, 2010.
- [2] Lê Đình Thúy, *Toán cao cấp cho các nhà kinh tế (tập một, hai, ba)*, Nhà xuất bản Đại học Kinh tế quốc dân, 2010.
- [3] Lê Tuấn Hoa, *Đại số tuyến tính qua các ví dụ và bài tập*, Nhà xuất bản Đại học Quốc gia Hà Nội, 2005.
- [4] Nguyễn Đình Dương - Nguyễn Thị Đỗ Hạnh, *Bài giảng Đại số*, Nhà xuất bản Đại học Hàng Hải Việt Nam - Tài liệu lưu hành nội bộ, 2014.
- [5] Nguyễn Đình Trí - Nguyễn Văn Đĩnh - Nguyễn Hồ Quỳnh, *Toán học cao cấp (tập một, hai, ba) - Đại số và hình học giải tích*, Nhà xuất Giáo dục, 2001.
- [6] Nguyễn Đình Trí - Nguyễn Văn Đĩnh - Nguyễn Hồ Quỳnh, *Bài tập toán học cao cấp (tập một, hai, ba) - Đại số và hình học giải tích*, Nhà xuất Giáo dục, 2001.
- [7] Nguyễn Hữu Việt Hưng, *Đại số tuyến tính*, Nhà xuất Đại học Quốc gia Hà Nội, 2001.
- [8] Nguyễn Cao Văn, Trần Thái Ninh, *Lí thuyết xác suất và thống kê toán*, Nhà xuất bản giáo dục, Hà Nội, 2002.
- [9] Lê Ngọc Lăng (chủ biên), Nguyễn Chí Bảo, Trần Xuân Hiền, Nguyễn Phú Trường, *Ôn thi học kỳ và thi vào giai đoạn 2 tập 1*, NXB Giáo dục, Hà Nội, 1997.
- [10] Lê Ngọc Lăng (chủ biên), Nguyễn Chí Bảo, Trần Xuân Hiền, Nguyễn Phú Trường, *Ôn thi học kỳ và thi vào giai đoạn 2 tập 2*, NXB Giáo dục, Hà Nội, 1997.
- [11] Tống Đình Quỳ, *Giáo trình Xác suất thống kê*, Nhà xuất bản giáo dục, Hà Nội, 1999.
- [12] Đào Hữu Hồ, Nguyễn Văn Hữu, Hoàng Hữu Như, *Thống kê toán học*, Nhà xuất bản Đại học quốc gia Hà Nội, Hà Nội, 2004.