PHẦN 1. TOÁN CAO CẤP

Trong phần này, sinh viên sẽ được học những kiến thức cơ bản và cần thiết nhất của Toán cao cấp nhằm phục vụ cho việc học những học phần tiếp theo như những vấn đề về đại số tuyến tính bao gồm ma trận, các phép toán cơ bản về ma trận, hệ phương trình tuyến tính và cách giải hệ phương trình tuyến tính có sử dụng các phép biến đổi sơ cấp trên ma trận; đạo hàm, cách tính và một số bài toán ứng dụng trong thực tế của đạo hàm; nguyên hàm và tích phân, cách tính và một số ứng dụng của tích phân xác định trong đời sống kinh tế, xã hội.

Chương 1

Đại số tuyến tính

1.1. Ma trận và các phép toán cơ bản của ma trận

Ma trận là bảng số hình chữ nhật và được sử dụng để lưu trữ thông tin và làm việc với chúng. Ma trận có rất nhiều ứng dụng trong kỹ thuật, trong đời sống, trong kinh tế, kỹ thuật, vật lý, cơ học, công nghệ thông tin, thuyết mật mã, ... Chẳng hạn, một công ty kinh doanh 3 mặt hàng gồm áo, quần và kính. Công ty có hai cửa hàng A và B. Giả sử số lượng hàng bán được trong 1 tháng là: cửa hàng A: 100 áo, 120 quần, 300 kính và cửa hàng B: 125 áo, 100 quần, 250 kính. Sắp xếp dữ liệu này ở dạng bảng:

	áo	quần	kính
A	100	120	300
B	125	100	250

Ta có thể viết lại bảng trên dưới dạng $T_1 = \begin{pmatrix} 100 & 120 & 300 \\ 125 & 100 & 250 \end{pmatrix}$.

Khi đó T_1 ở trên chính là một ma trận.

1.1.1. Các khái niệm cơ bản về ma trận

Định nghĩa 1.1.1. Một bảng số gồm $m \times n$ phần tử được xếp thành m hàng, n cột được gọi là một *ma trận cỡ* $m \times n$, ký hiệu

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Ký hiệu rút gọn: $A = (a_{ij})_{m \times n}$, trong đó a_{ij} biểu thị phần tử ở hàng i, cột j của ma trận A (với $i = 1, 2, \dots m; j = 1, 2, \dots n$).

Khi $m = n$, ta có một ma trận vuông với n hàng, n cột gọi là *ma trận vuông cấp n*.

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

Các phần tử $a_{11}, a_{22}, \dots, a_{nn}$ gọi là các phần tử chéo. Đường thẳng xuyên qua các phần tử chéo gọi là đường chéo chính của ma trận.

Ví dụ 1.1.2. $A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 2 & 3 \end{pmatrix}$ là một ma trận cỡ 3×2. Ta có thể viết $A = (a_{ij})_{3\times 2}$ với $a_{11} = 1; a_{12} = 0; a_{21} = -1; a_{22} = 1; a_{31} = 2; a_{32} = 3$.

Chú ý 1.1.3. Ta chỉ xét các ma trận thực, tức là các ma trận với $a_{ij} \in R$.

Định nghĩa 1.1.4. *Ma trận tam giác* là ma trận vuông có tất cả các phần tử nằm về một phía của đường chéo chính đều bằng không. Có 2 loại ma trận tam giác, ma trận tam giác trên (ma trận A), ma trận tam giác dưới (ma trận B).

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix};\ B = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Định nghĩa 1.1.5. *Ma trận đường chéo* là ma trận có tất cả các phần tử nằm ngoài đường chéo chính đều bằng không.

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

Định nghĩa 1.1.6. *Ma trận đơn vị* là ma trận đường chéo với các phần tử trên đường chéo chính đều bằng 1.

Ma trận đơn vị được ký hiệu là: I (hoặc E)

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Định nghĩa 1.1.7. *Ma trận không* là ma trận mà tất cả các phần tử đều bằng không. Ma trận không được ký hiệu là O.

Ví dụ 1.1.8. Ma trận $O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ là ma trận không cỡ 3×2.

Định nghĩa 1.1.9. *Ma trận bậc thang* là ma trận thoả mãn điều kiện sau đây:

(1) Nếu có hàng không (tức là tất cả các phần tử đều bằng không) thì các hàng khác không luôn ở trên các hàng không.

(2) Trên hai hàng khác không thì phần tử khác không đầu tiên ở hàng dưới bao giờ cũng ở bên phải cột chứa phần tử khác không đầu tiên ở hàng trên.

Ví dụ 1.1.10. Trong các ma trận sau đâu là ma trận bậc thang?

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad B = \begin{pmatrix} -1 & 0 & -5 & -5 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}; \quad C = \begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & 2 \\ 0 & 0 & 3 \end{pmatrix};$$

$$D = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 3 & 2 \\ 0 & 2 & 4 & 5 \end{pmatrix}; \quad E = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

Giải: Các ma trận A, B, C là ma trận bậc thang.

Định nghĩa 1.1.11. Hai ma trận A và B được gọi là *bằng nhau* nếu chúng có cùng cỡ và các phần tử ở cùng vị trí bằng nhau, tức là: $A = (a_{ij})_{m\times n}; B = (b_j)_{m\times n}$ và $a_{ij} = b_{ij}, \forall i, j$. Ký hiệu hai ma trận bằng nhau là: $A = B$.

Ví dụ 1.1.12. Ta có $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$ nghĩa là $a = 0; b = 1; c = 1; d = 2$.

1.1.2. Các phép toán cơ bản của ma trận

a. Phép cộng hai ma trận

Định nghĩa 1.1.13. Cho hai ma trận $A = (a_{ij}); B = (b_{ij})$ có cùng cỡ $m \times n$, tổng của chúng, kí hiệu $A + B$ là một ma trận cũng có cỡ là $m \times n$ và được xác định bởi phép cộng các phần tử tương ứng ở cùng vị trí. Tức là,

$$A + B = (a_{ij} + b_{ij})_{m\times n}$$

Ví dụ 1.1.14. $\begin{pmatrix} 3 & -5 & 2 \\ 1 & 0 & -1 \\ 2 & 7 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & -2 & -2 \end{pmatrix} = \begin{pmatrix} 4 & -5 & 2 \\ 3 & -1 & -1 \\ 5 & 5 & 2 \end{pmatrix}$.

Tính chất 1.1.15. Phép cộng ma trận có các tính chất sau

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

(2) $(A + B) + C = A + (B + C)$.

(3) $A + 0 = 0 + A = A$.

(4) $A + (-A) = (-A) + A = 0$.

b. Phép nhân ma trận với một số

Định nghĩa 1.1.16. Cho $A = (a_{ij})$ là một ma trận cỡ $m \times n$, c là một hằng số tùy ý, thì tích của ma trận A và hằng số c là một ma trận cA cũng có cỡ là $m \times n$ và được xác định bằng cách đem số đó nhân với từng phần tử của ma trận. Tức là,

$$cA = c(a_{ij}) = (ca_{ij})$$

Ví dụ 1.1.17. Ta có $2\begin{pmatrix} 3 & 4 \\ 1 & 5 \end{pmatrix} = \begin{pmatrix} 2.3 & 2.4 \\ 2.1 & 2.5 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 2 & 10 \end{pmatrix}$.

Tính chất 1.1.18. Phép nhân ma trận với một số có các tính chất sau:

(1) $k(A + B) = kA + kB$.

(2) $(k + h)A = kA + hA$.

(3) $k(hA) = (kh)A$.

(4) $1A = A$.

(5) $0A = 0$.

Chú ý 1.1.19. $-B = (-1)B = (-b_{ij})_{m\times n}$ với $B = (b_{ij})_{m\times n}$.
Khi đó ta định nghĩa phép trừ hai ma trận: Nếu A và B là hai ma trận có cùng cỡ thì $A - B = A + (-B)$.

c. Phép nhân hai ma trận

Định nghĩa 1.1.20. Cho A là một ma trận cỡ $m \times p$ và B là một ma trận cỡ $p \times n$ thì tích hai ma trận $C = A.B$ là một ma trận cỡ $m \times n$ và mỗi phần tử c_{ij} của ma trận tích được tính bởi công thức

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{p} a_{ik}b_{kj},\ i = 1, \ldots, m, j = 1, \ldots, n.$$

Ví dụ 1.1.21.

$$\begin{pmatrix} 2 & -1 \\ 2 & 4 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2.1 + (-1)(-1) & 2.2 + (-1).0 & 2.3 + (-1).1 \\ 2.1 + 4.(-1) & 2.2 + 4.0 & 2.3 + 4.1 \\ 1.1 + 3.(-1) & 1.2 + 3.0 & 1.3 + 3.1 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 5 \\ -2 & 4 & 10 \\ -2 & 2 & 6 \end{pmatrix}.$$

Ví dụ 1.1.22. Một cửa hàng kinh doanh 3 mặt hàng: áo, quần, kính. Giả sử số lượng hàng bán được trong một tháng là: 100 áo, 120 quần, 300 kính. Ta sắp xếp số liệu trên dưới dạng ma trận

$$T_1 = \begin{pmatrix} 100 & 120 & 300 \\ 125 & 100 & 250 \end{pmatrix}.$$

Giả sử tháng thứ hai bán được: $T_2 = \begin{pmatrix} 130 & 80 & 240 \\ 115 & 150 & 150 \end{pmatrix}\cdot$ Khi đó, lượng hàng bán được trong hai tháng: $T = T_1 + T_2 = \begin{pmatrix} 230 & 200 & 540 \\ 240 & 250 & 400 \end{pmatrix}\cdot$

Giả sử tiền lãi trong tháng 1: áo 15 ngàn, quần 30 ngàn, kính 10 ngàn $L_1 = \begin{pmatrix} 15 \\ 30 \\ 10 \end{pmatrix}$.

Vậy lợi nhuận trong tháng 1 của cửa hàng là:

$$T_1L_1 = \begin{pmatrix} 100 & 120 & 300 \\ 125 & 100 & 250 \end{pmatrix} \begin{pmatrix} 15 \\ 30 \\ 10 \end{pmatrix} = \begin{pmatrix} 8100 \\ 7375 \end{pmatrix}.$$

Giả sử tiền lãi trong tháng 2: áo 25 ngàn, quần 35 ngàn, kính 17 ngàn $L_2 = \begin{pmatrix} 25 \\ 35 \\ 17 \end{pmatrix}$.

Vậy lợi nhuận trong hai tháng của cửa hàng là:

$$T_1L_1 + T_2L_2 = \begin{pmatrix} 8100 \\ 7375 \end{pmatrix} + \begin{pmatrix} 10130 \\ 10675 \end{pmatrix} = \begin{pmatrix} 18230 \\ 18050 \end{pmatrix}.$$

Chú ý 1.1.23. (1) Trong nhiều trường hợp ta có thể tính được AB nhưng lại không tính được BA. Ta chỉ thực hiện được cả tích AB và BA khi số cột của ma trận này bằng số hàng của ma trận kia và ngược lại. Đặc biệt khi A và B đều là ma trận vuông cấp n thì ta tính được cả tích AB và BA.
(2) Phép nhân hai ma trận không có tính chất giao hoán, tức là AB chưa chắc bằng BA.
(3) Có những ma trận $A \neq 0, B \neq 0$ nhưng $AB = 0$.

Ví dụ 1.1.24. Cho hai ma trận $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$; $B = \begin{pmatrix} 2 & -6 \\ -1 & 3 \end{pmatrix}$. Khi đó, ta có tích hai ma trận là $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$; $BA = \begin{pmatrix} -10 & -20 \\ 5 & 10 \end{pmatrix}$.

Tính chất 1.1.25. Phép nhân hai ma trận có tính chất sau:

(1) Tính chất kết hợp: $A(BC) = (AB)C$.

(2) Tính chất phân phối với phép cộng:

$$(B + C)A = BA + CA;\ A(B + C) = AB + AC.$$

(3) $k(BC) = (kB)C = B(kC)$.

(4) Với mọi A là ma trận vuông cấp n, I là ma trận đơn vị cùng cấp thì:

$$IA = AI = A;\ \underbrace{A \dots A}_{k} = A^k \text{ (quy ước} A^0 = I).$$

d. Phép chuyển vị ma trận

Định nghĩa 1.1.26. Ma trận chuyển vị của ma trận $A = (a_{ij})$ cỡ $m \times n$ là ma trận A^T có cỡ $n \times m$ thu được từ ma trận A bằng cách đổi hàng thành cột. Tức là, cột thứ i của ma trận A^T là hàng thứ i của ma trận A với mọi i. Phép toán biến ma trận A thành ma trận chuyển vị A^T được gọi là *phép chuyển vị ma trận.*

Ví dụ 1.1.27. Cho ma trận $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \end{pmatrix} \Rightarrow A^t = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix}$.

Tính chất 1.1.28. Phép chuyển vị có tính chất sau:
(1) $(A^T)^T = A$.
(2) $(A+B)^T = A^T + B^T$.
(3) $(kA)^T = kA^T (k \in R)$.
(4) $(AB)^T = B^T A^T$.

1.2. Hệ phương trình tuyến tính

Ví dụ 1.2.1. Hãy tìm số gà và chó trong bài toán sau.

Vừa gà vừa chó
Bó lại cho tròn
Ba mươi sáu con
Một trăm chân chẵn.

Nếu đặt số gà là x, số chó là $y, (x > 0, y > 0)$ khi đó ta thu được hệ phương trình sau

$$\begin{cases} x + y = 36 \\ 2x + 4y = 100 \end{cases}$$

Hệ trên có là hệ phương trình tuyến tính không? Bạn đọc dễ dàng giải được hệ phương trình trên để trả lời câu hỏi của bài toán ban đầu.

Ví dụ 1.2.2. (1) Phương trình tuyến tính tổng quát trong R^2 có dạng $ax + by = c$, trong đó a, b, c thuộc tập số thực R.
(2) Phương trình tuyến tính tổng quát trong R^3 có dạng $ax + by + cz = d$, trong đó a, b, c, d là các hệ số thuộc tập số thực R.

Định nghĩa 1.2.3. *Một phương trình tuyến tính n* biến $x_1, x_2, ...x_n$ là một phương trình có thể viết dưới dạng $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ trong đó $a_1, a_2, ..., a_n$ và b là các hằng số thuộc tập số thực R.

Ví dụ 1.2.4. Các phương trình nào sau đây là phương trình tuyến tính?

$$(a)\ 3x - 4y = -1; \qquad (b)\ \sqrt{2}x + \frac{\pi}{4}y - z = 1;$$
$$(c)\ \frac{x}{y} + z = 2. \qquad (d)\ \sqrt{2x} + \frac{\pi}{4}y - z = 0;$$

Định nghĩa 1.2.5. *Nghiệm của một phương trình tuyến tính* $a_1x_1 + a_2x_2 + \cdot + a_nx_n = b$ là một bộ số thực $(s_1, s_2, ...s_n)$ thỏa mãn phương trình đó, tức là khi thay $x_1 = s_1, x_2 = s_2, ..., x_n = s_n$ vào phương trình tuyến tính ta được một đẳng thức luôn đúng.

1.2.1. Dạng tổng quát của hệ phương trình tuyến tính

Định nghĩa 1.2.6. *Một hệ phương trình tuyến tính* là một hệ gồm các phương trình đều là phương trình tuyến tính với cùng các biến. Dạng tổng quát của hệ phương trình tuyến tính gồm m phương trình tuyến tính với n ẩn có dạng

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.1}$$

trong đó $x_1, x_2, \ldots, x_n$ là các ẩn. a_{ij} là các hệ số của phương trình thứ i của ẩn x_j; b_i là vế phải của phương trình thứ i.

Khi $m = n$ ta có một hệ vuông với n phương trình n ẩn.

Khi các $b_i = 0\ \forall i$, hệ phương trình đó gọi là hệ thuần nhất.

Định nghĩa 1.2.7. *Nghiệm của hệ phương trình tuyến tính* là một bộ số thực $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ thỏa mãn mọi phương trình trong hệ. *Tập nghiệm của hệ phương trình tuyến tính* là tập hợp gồm tất cả các nghiệm của hệ. Giải hệ phương trình tuyến tính là đi tìm nghiệm của nó.

Ví dụ 1.2.8. Ta có hệ $\begin{cases} x_1 - 5x_2 + 4x_3 = 1 \\ 3x_1 + x_2 - 6x_3 = 3 \end{cases}$ là một hệ 2 phương trình 3 ẩn và hệ $\begin{cases} x_1 - x_2 = 1 \\ 5x_1 + 6x_2 = 8 \end{cases}$ là hệ 2 phương trình 2 ẩn.

1.2.2. Dạng ma trận của hệ phương trình tuyến tính

Xét hệ (1.1). Đặt $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$; $b = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}$; $x = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$;

$$\bar{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right).$$

Ta gọi A là ma trận hệ số; $\bar{A}$ là ma trận bổ sung; b là ma trận vế phải; x là ma trận ẩn của hệ phương trình tuyến tính (1.1).

Dạng ma trận của hệ phương trình tuyến tính (1.1)là: $Ax = b$.

Dạng ma trận của hệ phương trình tuyến tính thuần nhất là: $Ax = 0$.

Ví dụ 1.2.9. Giải hệ phương trình sau

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

Giải. Ta có dạng ma trận của hệ phương trình tuyến tính trên là

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ -5 \end{pmatrix}.$$

Đặt $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & -1 & -2 \end{pmatrix}$. Ta có det A $= -5 \neq 0$. Ta tính các phần phụ đại số của các phần tử trong ma trận A như sau:

$$c_{11} = \begin{vmatrix} 3 & 1 \\ -1 & -2 \end{vmatrix} = -5; \qquad c_{12} = -\begin{vmatrix} 2 & 1 \\ 1 & -2 \end{vmatrix} = 5; \qquad c_{13} = \begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = -5;$$

$$c_{21} = -\begin{vmatrix} 1 & 1 \\ -1 & -2 \end{vmatrix} = 1; \qquad c_{22} = \begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = -3; \qquad c_{23} = -\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = 2;$$

$$c_{31} = \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} = -2; \qquad c_{32} = -\begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} = 1; \qquad c_{33} = \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = 1.$$

Vậy $A^{-1} = \dfrac{-1}{5} \begin{pmatrix} -5 & 1 & -2 \\ 5 & -3 & 1 \\ -5 & 2 & 1 \end{pmatrix}$.

Nghiệm của phương trình đã cho là

$$X = A^{-1}.B = \frac{-1}{5} \begin{pmatrix} -5 & 1 & -2 \\ 5 & -3 & 1 \\ -5 & 2 & 1 \end{pmatrix} . \begin{pmatrix} 3 \\ 5 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

Vậy hệ phương trình tuyến tính ban đầu có nghiệm duy nhất $x_1 = 0; x_2 = 1; x_3 = 2$.

Nhận xét 1.2.10. Trong trường hợp hệ phương trình tuyến tính có ma trận hệ số là ma trận vuông và định thức của ma trận đó khác không thì ta có thể tìm nghiệm của hệ bằng cách giải phương trình ma trận tương ứng.

1.2.3. Cách giải hệ phương trình tuyến tính

a. Phương pháp khử Gauss–Jordan

Định nghĩa 1.2.11. Hai hệ phương trình được gọi là *tương đương* nếu chúng có cùng một tập hợp nghiệm.

Định lí 1.2.12. *Các phép biến đổi sau đây là các phép biến đổi tương đương trên hệ phương trình:*

(1) *Đổi chỗ hai phương trình của hệ;*

(2) *Nhân một phương trình của hệ với một hằng số $k \neq 0$;*

(3) *Cộng k lần một phương trình vào một phương trình khác.*

Ví dụ 1.2.13. Giải hệ phương trình sau

$$\begin{cases} x_1 - x_2 + x_3 = 1 & (1.3) \\ 2x_1 + x_2 + x_3 = 2 & (1.4) \\ 3x_1 + x_2 + 2x_3 = 3 & (1.5) \end{cases}$$

Ta xem xét sự tương ứng giữa các phép biến đổi tương đương của hệ phương trình và các phép biến đổi sơ cấp theo hàng của ma trận bổ sung qua bảng đối chiếu sau:

$\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 + x_2 + x_3 = 2 \\ 3x_1 + x_2 + 2x_3 = 3 \end{cases}$	$\left(\begin{array}{ccc\|c} 1 & -1 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 3 & 1 & 2 & 0 \end{array}\right)$
Nhân PT (1.3) với -2 rồi cộng vào PT (1.4) Nhân PT (1.3) với -3 rồi cộng vào PT (1.5)	Nhân hàng 1 với -2 rồi cộng vào hàng 2; Nhân hàng 1 với -3 rồi cộng vào hàng 3
$\Leftrightarrow \begin{cases} x_1 - x_2 + x_3 = 1 & (1.3') \\ 3x_2 - x_3 = 0 & (1.4') \\ 4x_2 - x_3 = -3 & (1.5') \end{cases}$	$\rightarrow \left(\begin{array}{ccc\|c} 1 & -1 & 1 & 1 \\ 0 & 3 & -1 & 0 \\ 0 & 4 & -1 & -3 \end{array}\right)$
Nhân PT (1.4') với -4 ; Nhân PT (1.5') với 3 rồi cộng lại đặt ở vị trí phương trình thứ 3, ta được	Nhân hàng 2 với -4, Nhân hàng 3 với 3 rồi cộng lại đặt ở vị trí hàng 3 có
$\Leftrightarrow \begin{cases} x_1 - x_2 + x_3 = 1 \\ 3x_2 - x_3 = 0 \\ x_3 = -9 \end{cases} \quad (I)$	$\rightarrow \left(\begin{array}{ccc\|c} 1 & -1 & 1 & 1 \\ 0 & 3 & -1 & 0 \\ 0 & 0 & 1 & -9 \end{array}\right) \quad (II)$

Nhìn vào bảng trên ta có nhận xét sau:

(1) Có sự tương ứng giữa các phép biến đổi tương đương của hệ phương trình tuyến tính với các phép biến đổi sơ cấp theo hàng ở ma trận bổ sung, chẳng hạn:

- Phép đổi chỗ hai hàng ứng với phép đổi vị trí hai phương trình trong hệ.

- Phép nhân 1 hàng với một số khác không ứng với phép nhân hai vế của một phương trình trong hệ với một số khác không.

- Phép cộng bội k của một hàng vào một hàng khác ứng với phép cộng bội k của một phương trình vào một phương trình khác.

Nói cách khác các phép biến đổi tương đương của hệ phương trình thực chất chỉ biến đổi trên bộ hệ số của các phương trình.

Do đó thay cho việc biến đổi trên hệ phương trình có ẩn cồng kềnh như trên ta có thể biến đổi trên ma trận bổ sung của hệ mà không làm thay đổi nghiệm của hệ.

(2) Ma trận bậc thang (II) thu được chính là ma trận bổ sung của hệ phương trình (I) ở cột bên trái. Mà hệ này giải rất đơn giản bằng phương pháp thế ngược từ dưới lên.

(3) Khi giải hệ phương trình tuyến tính bằng phương pháp khử Gauss–Jordan, ta chỉ sử dụng các phép biến đổi sơ cấp theo hàng, không sử dụng các phép biến đổi sơ cấp theo cột.

Quay lại Ví dụ 1.2.13 ở trên, từ $x_3 = -9$ thay vào PT (1.4) ta có $x_2 = -3$, thay x_2, x_3 vào PT(1.3) có $x_1 = 7$.

Vậy hệ phương trình đã cho có nghiệm là: $\begin{cases} x_1 = 7 \\ x_2 = -3 \\ x_3 = -9 \end{cases}$.

Định nghĩa 1.2.14. Ma trận A và B được gọi là tương đương theo hàng nếu sử dụng liên tiếp các phép biến đổi sơ cấp theo hàng đưa ma trận A về ma trận B.

Định lí 1.2.15. *Ma trận A và B được gọi là tương đương theo hàng nếu và chỉ nếu chúng có thể biến đổi về cùng một ma trận bậc thang bởi các phép biến đổi sơ cấp theo hàng của ma trận. Các phép biến đổi sơ cấp theo hàng của ma trận gọi là các phép biến đổi tương đương trên ma trận.*

• Các bước giải hệ phương trình tuyến tính bằng khử Gauss–Jordan.

Khi các phép biến đổi sơ cấp theo hàng của ma trận được áp dụng vào ma trận bổ sung của một hệ phương trình tuyến tính đưa ma trận này về dạng bậc thang thì ma trận bổ sung ban đầu tương đương với ma trận bậc thang thu được và hệ phương trình ban đầu tương đương với hệ phương trình nhận ma trận bậc thang làm ma trận bổ sung, hệ này giải rất đơn giản bằng phương pháp thế ngược. Do vậy để giải một hệ phương trình tuyến tính bất kì ta làm theo các bước sau:

Bước 1: Lập ma trận bổ sung $\bar{A}$.

Bước 2: Sử dụng các phép biến đổi sơ cấp theo hàng đưa $\bar{A}$ về dạng bậc thang B.

Bước 3: Kết luận.

Hệ phương trình ban đầu tương đương với hệ phương trình mà nhận ma trận bậc thang B làm ma trận bổ sung. Giải hệ này bằng phương pháp thế ngược từ dưới lên.

Ví dụ 1.2.16. Giải hệ phương trình sau:

$$\begin{cases} x_1 + 5x_2 + 3x_3 = 20 \\ 2x_1 - x_2 - x_3 = -3 \\ x_2 + 3x_3 = 11 \\ -4x_1 + 14x_2 + 12x_3 = 60 \end{cases}$$

Ta có

$$\bar{A} = \left(\begin{array}{cccc|c} 1 & 5 & 3 & 20 \\ 2 & -1 & -1 & -3 \\ 0 & 1 & 3 & 11 \\ -4 & 14 & 12 & 60 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 5 & 3 & 20 \\ 0 & -11 & -7 & -43 \\ 0 & 1 & 3 & 11 \\ 0 & 34 & 24 & 140 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 5 & 3 & 20 \\ 0 & 0 & 26 & 78 \\ 0 & 1 & 3 & 11 \\ 0 & 0 & -78 & -234 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|c} 1 & 5 & 3 & 20 \\ 0 & 0 & 26 & 78 \\ 0 & 1 & 3 & 11 \\ 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 5 & 3 & 20 \\ 0 & 0 & 3 & 11 \\ 0 & 0 & 26 & 78 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Từ đó suy ra

$$\begin{cases} x_1 + 5x_2 + 3x_3 = 20 \\ x_2 + 3x_3 = 11 \\ 26x_3 = 78 \\ 0 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = 1 \\ x_2 = 2 \\ x_3 = 3 \end{cases}.$$

Ví dụ 1.2.17. Giải hệ phương trình

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

Xét $\bar{A} = \left(\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ 0 & 3 & -3 & -6 \\ 0 & 2 & -2 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 5 \end{array}\right).$

Hệ phương trình đã cho tương tương với hệ

$$\begin{cases} x_1 - x_2 + 2x_3 = 3 \\ 0x_1 + x_2 - x_3 = -2 \\ 0x_1 + 0x_2 + 0x_3 = 5 \end{cases} \quad \Rightarrow \quad \text{hệ đã cho vô nghiệm.}$$

Ví dụ 1.2.18. Giải hệ phương trình

$$\begin{cases} 2x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 = 0 \\ 4x_1 + 4x_2 + 2x_3 + 3x_4 = 0 \end{cases}$$

Xét $\bar{A} = \left(\begin{array}{cccc|c} 2 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 4 & 4 & 2 & 3 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 \\ 0 & 0 & -2 & 3 & 0 \end{array}\right).$

Hệ phương trình đã cho tương đương với hệ

$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ \quad - x_2 - x_3 + x_4 = 0 \\ \quad\quad - 2x_3 + 3x_4 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = -k \\ x_2 = -\dfrac{1}{2}k \\ x_3 = \dfrac{3}{2}k \\ x_4 = k \in \mathbb{R} \end{cases}.$$

Vậy hệ phương trình đã cho có vô số nghiệm có dạng $\left(-k, -\dfrac{1}{2}k, \dfrac{3}{2}k, k\right)$, $k \in \mathbb{R}$.

Nhận xét 1.2.19.
(1) Ví dụ 1.2.16, hệ có nghiệm duy nhất và có $r(A) = r(\bar{A}) = 3 =$ số ẩn.
(2) Ví dụ 1.2.17, hệ phương trình đã cho vô nghiệm và có

$$r(A) = 2 \neq 3 = r(\bar{A}).$$

(3) Ví dụ 1.2.18, hệ phương trình đã cho có vô số nghiệm và có

$$r(A) = r(\bar{A}) = 3 < 4 = \text{ số ẩn của hệ}.$$

Nhận xét trên đã được khẳng định thông qua định lý sau đây.

b. Định lý Kronecker - Capelli

Định lí 1.2.20. *Hệ phương trình tuyến tính* (1.1) *có nghiệm khi và chỉ khi* $r(A) = r(\bar{A})$.

Như vậy,

(1) Nếu $r(A) \neq r(\bar{A})$ thì hệ (1.1) vô nghiệm.

(2) Nếu $r(A) = r(\bar{A}) = n$ thì hệ (1.1) có nghiệm duy nhất.

(3) Nếu $r(A) = r(\bar{A}) = s < n$ thì hệ (1.1) có vô số nghiệm.

Khi đó ta giải hệ (1.1) như sau: Vì $r(A) = r(\bar{A}) = s$ nên sẽ tồn tại một định thức con khác không cấp s. Các phần tử của định thức con đó nằm ở s phương trình là các hệ số của s ẩn gọi là s ẩn chính, các ẩn còn lại gọi là ẩn phụ. Cả hệ sẽ tương đương với hệ mới gồm s phương trình. Trong s phương trình này, ta chuyển các ẩn phụ sang vế phải và coi là tham số. Giải hệ s phương trình s ẩn chính ta sẽ tìm được nghiệm phụ thuộc vào vế phải và các ẩn phụ.

Ví dụ 1.2.21. Tìm m để hệ sau có vô số nghiệm

$$\begin{cases} 3x + y - 2z = m \\ 2x + 4y - z = -2 \\ 4x - 2y - 3z = 1 \end{cases}.$$

Xét ma trận bổ sung của hệ đã cho

$$\bar{A} = \left(\begin{array}{ccc|c} 3 & 1 & -2 & m \\ 2 & 4 & -1 & -2 \\ 4 & -2 & -3 & 1 \end{array}\right) \xrightarrow[H3: -4H1+3H3]{H_2: -2H_1+3H_2} \begin{bmatrix} 3 & 1 & -2 & m \\ 0 & 10 & 1 & -2m-6 \\ 0 & -10 & -1 & -4m+3 \end{bmatrix}$$

$$\xrightarrow{H_3+H_2\to H_3} \begin{bmatrix} 3 & 1 & -2 & m \\ 0 & 10 & 1 & -2m-6 \\ 0 & 0 & 0 & -6m-3 \end{bmatrix}.$$

Ở ma trận bậc thang cuối cùng ta thấy $r(A) = 2, r(\bar{A}) \geq 2$.
Để hệ vô số nghiệm thì $r(\bar{A}) = r(A) = 2$ tức $-6m - 3 = 0$ hay $m = -\frac{1}{2}\cdot$
Vậy với $m = -\frac{1}{2}$ thì hệ trên có vô số nghiệm.

c. Hệ phương trình tuyến tính thuần nhất

Xét hệ thuần nhất

$$(III) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases},$$

với ma trận hệ số của hệ là ma trận vuông cấp n

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

Dễ thấy $x_1 = 0, x_2 = 0, \dots, x_n = 0$ thỏa mãn hệ phương trình trên nên $0(0, 0, 0, \dots, 0)$ là một nghiệm của hệ thuần nhất.

• Nghiệm tầm thường và nghiệm không tầm thường

Nghiệm $0(0, 0, 0, \dots, 0)$ của hệ thuần nhất được gọi là *nghiệm tầm thường*. Nghiệm không phải là nghiệm tầm thường của hệ (III) được gọi là *nghiệm không tầm thường.*

Nếu $\det A \neq 0$ thì hệ trên là hệ Cramer nên có nghiệm duy nhất mà ta thấy hệ có nghiệm 0. Do đó nếu $\det A \neq 0$ hệ chỉ có nghiệm tầm thường. Vì vế phải của hệ là ma trận không nên khi sử dụng các phép biến đổi sơ cấp theo hàng đưa $\bar{A}$ về dạng bậc thang, cột vế phải cũng luôn bằng 0, vì thế ta luôn có $r(A) = r(\bar{A}) < n =$ số ẩn (nếu $\det A = 0$) và hệ có vô số nghiệm, tức ngoài nghiệm 0 hệ có nghiệm khác 0. Ta có định lý sau đây.

Định lí 1.2.22. *Hệ thuần nhất (III) có nghiệm không tầm thường khi và chỉ khi ma trận các hệ số của hệ là ma trận suy biến, tức là* $\det A = 0$.

Ví dụ 1.2.23. Xác định a để hệ sau có nghiệm không tầm thường

$$\begin{cases} ax - 3y + z = 0 \\ 2x + y + z = 0 \\ 3x + 2y - 2z = 0 \end{cases}.$$

Hệ có ma trận hệ số là

$$A = \begin{pmatrix} a & -3 & 1 \\ 2 & 1 & 1 \\ 3 & 2 & -2 \end{pmatrix}.$$

$$\det A = \begin{vmatrix} a & -3 & 1 \\ 2 & 1 & 1 \\ 3 & 2 & -2 \end{vmatrix} = -2a - 9 + 4 - 3 - 12 - 2a = -4a - 20.$$

Theo Định lý 1.2.22 thì hệ có nghiệm không tầm thường khi và chỉ khi $\det A = 0$, tức là $4a + 20 = 0 \Rightarrow a = -5$.

Chú ý 1.2.24. (1) Ta cũng có thể sử dụng phương pháp khử Gauss để giải hệ thuần nhất có số phương trình và số ẩn khác nhau. Nghiệm thu được sẽ phụ thuộc vào một hay nhiều tham số.
(2) Ngoài hai phương pháp chính trên, tuỳ thuộc vào từng đặc thù của hệ phương trình ta có thể có nhiều cách khác nhau để đơn giản lời giải.

BÀI TẬP CHƯƠNG 1

Bài 1: Thực hiện phép nhân hai ma trận:

a) $\begin{pmatrix} -3 & 4 \\ 11 & 6 \end{pmatrix} \begin{pmatrix} 10 & 5 \\ -2 & 7 \end{pmatrix}$ b) $\begin{pmatrix} 9 & 5 & -1 \\ 4 & 12 & 5 \end{pmatrix} \begin{pmatrix} 4 & 7 & \\ -3 & 50 & 6 \end{pmatrix}$

Bài 2: Giải hệ phương trình:

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

3) $\begin{cases} x + 3y - 2z = -3 \\ 2x + y + 3z = 6 \\ 3x - y + 4z = 11 \end{cases}$ 4) $\begin{cases} -2x + y - 3z = 1 \\ 3x - 4y + 2z = 3 \\ 5x + 2y + z = -2 \end{cases}$

5) $\begin{cases} 3x_1 + x_2 - 2x_3 + 5x_4 = 1 \\ x_1 - x_2 + 3x_3 + x_4 = 0 \\ 2x_1 + 3x_2 - 8x_3 + 3x_4 = 3 \end{cases}$ 6) $\begin{cases} 4x + y + z = 1 \\ 3x - 2y - z = 0 \\ -x + 5y + 2z = 0 \\ 7x + 7y + 4z = 2 \end{cases}$

Chương 2

Đạo hàm và một số ứng dụng

2.1. Hàm một biến

2.1.1. Các khái niệm cơ bản về hàm số một biến số

Trong các lĩnh vực khoa học, kinh tế, đời sống, chúng ta thường gặp các đại lượng đo được bằng số. Khi nghiên cứu quy luật thay đổi trị số của các đại lượng đó, người ta thường dùng chữ để ký hiệu số đo của chúng. Chẳng hạn, trong hình học, chúng ta thường dùng chữ r để ký hiệu số đo bán kính, chữ s để ký hiệu diện tích, chữ v để ký hiệu thể tích. Với mỗi hình cụ thể, r, s, v là các số thực và được gọi là *các biến số* hay gọi tắt là *biến*. Các biến số thường có mối quan hệ chi phối lẫn nhau: sự thay đổi giá trị của các biến số này kéo theo sự thay đổi giá trị của biến số kia theo một quy luật nhất định. Chẳng hạn, chúng ta thấy khi hàng hóa thay đổi thì lượng hàng hóa mà người sản xuất muốn bán ra thị trường và lượng hàng hóa mà người tiêu dùng sử dụng cũng thay đổi theo; khi vận tốc của một vật di chuyển trong một khoảng thời gian xác định thay đổi thì quãng đường đi được cũng thay đổi theo... Sự phụ thuộc của biến số này vào biến số khác thường được biểu diễn dưới dạng hàm số.

Định nghĩa 2.1.1. Cho $X \subseteq \mathbb{R}$ và $Y \subseteq \mathbb{R}$, ánh xạ $f : X \to Y$ được gọi là *hàm số* một biến số thực. Tập X được gọi là *miền xác định* của hàm f, ký hiệu là D_f và tập $f(X)$ là *miền giá trị* của hàm f, ký hiệu là R_f; $x \in D_f$ được gọi là *biến độc lập* hay *đối số* và $f(x) \in R_f$ được gọi là *biến phụ thuộc* hay *hàm số*. Ký hiệu hàm số: $y = f(x)$.

Hàm số có thể cho dưới dạng bảng, dạng đồ thị, sơ đồ Venn nhưng phổ biến nhất là hàm số cho bằng biểu thức hoặc phương trình.

Ví dụ 2.1.2. Phương trình $y - x^2 = 0 \Leftrightarrow y = x^2$ (có thể ký hiệu $y = f(x) = x^2$) biểu diễn y như một hàm số của x. Trong hàm số này, x là biến số độc lập, y là hàm số biến số x. Khi thay x bởi giá trị cụ thể, chẳng hạn, thay $x = 3 \Rightarrow y(3) = f(3) = 9$. Chúng ta nói giá trị của hàm số f tại 3 là 9.

Ví dụ 2.1.3. Phương trình nào dưới đây biểu diễn y là một hàm số của x? Tìm miền xác định của hàm số đó?

$$(a)\ x + y = 1; \qquad (b)\ x^2 + y^2 = 1; \qquad (c)\ x^2 + y = 1.$$

Giải:
(a) $x + y = 1 \Leftrightarrow y = 1 - x$ là một hàm số vì mỗi giá trị của x xác định duy nhất một giá trị của y. MXĐ: $D = \mathbb{R}$.
(b) $x^2 + y^2 = 1 \Leftrightarrow y = \pm\sqrt{1 - x^2}$ không là một hàm số vì tồn tại những giá trị của x mà với mỗi giá trị đó xác định 2 giá trị tương ứng của y.
(c) $x^2 + y = 1 \Leftrightarrow y = 1 - x^2$ là một hàm số vì mỗi giá trị của x xác định duy nhất một giá trị của y. MXĐ: $D = \mathbb{R}$.

Định nghĩa 2.1.4. Cho $X \subseteq \mathbb{R}, Y \subset \mathbb{R}, Z \subseteq \mathbb{R}$, cho hai hàm số f, g xác định như sau

$$\begin{array}{rcrcl} g & : & X & \longrightarrow & Y \\ & & x & \longmapsto & y = g(x) \end{array}; \qquad \begin{array}{rcrcl} f & : & Y & \longrightarrow & Z \\ & & y & \longmapsto & z = f(y) = f(g(x)) \end{array}.$$

Hàm số h được gọi là *hàm số hợp* của hàm số f và hàm số g, ký hiệu là $h(x) = f(g(x))$ hay $h(x) = (f \circ g)(x) = f(g(x)), \forall x \in X$.

Ví dụ 2.1.5.
(a) Hàm số $y = \sqrt{x^2 + 1}$ là hàm hợp của hai hàm số $u = x^2 + 1$ và $y = \sqrt{u}$.
(b) Hàm số $y = \sin(2x + 1)$ là hàm hợp của hai hàm số $y = \sin u$ và $u = 2x + 1$.

Định nghĩa 2.1.6. Cho hàm số $y = f(x)$, với miền xác định là X, miền giá trị là $Y = f(X)$. Nếu với mỗi giá trị $y_0 \in Y$ tương ứng với duy nhất một giá trị $x_0 \in X$ sao cho $f(x_0) = y_0$ thì tương ứng đó cũng xác định một hàm số từ tập Y vào tập X, hàm số này được gọi là *hàm số ngược* của hàm f và được ký hiệu là f^{-1}.

Như vậy, hàm số f được cho dưới dạng

$$\begin{array}{rcrcl} f & : & X & \longrightarrow & Y \\ & & x & \longmapsto & y = f(x) \end{array}$$

có hàm số ngược f^{-1} là

$$\begin{array}{rcrcl} f^{-1} & : & Y & \longrightarrow & X \\ & & y & \longmapsto & x = f^{-1}(y). \end{array}$$

Ví dụ 2.1.7. (a) Hàm số $y = \log_a x$ là hàm ngược của hàm $y = a^x$.
(b) Hàm số $y = \arcsin x$ là hàm ngược của hàm $y = \sin x$.

• Các hàm số sơ cấp cơ bản:

1. Hàm hằng: $f(x) = C, \forall x \in \mathbb{R}$;
2. Hàm số lũy thừa: $f(x) = x^\alpha, (\alpha = const)$;
3. Hàm số mũ: $f(x) = a^x (a > 0, a \neq 1)$;
4. Hàm số logarit: $f(x) = \log_a x (a > 0, a \neq 1)$;
5. Các hàm số lượng giác: $f(x) = \sin x; f(x) = \cos x; f(x) = \tan x; f(x) = \cot x$;
6. Các hàm số lượng giác ngược: $f(x) = \arcsin x; f(x) = \arccos x; f(x) = \arctan x;$ $f(x) = \operatorname{arccot} x$.

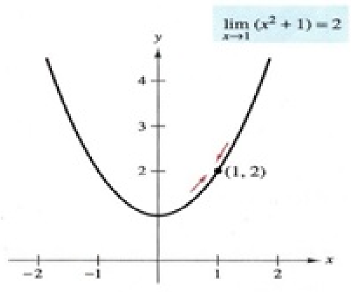
• **Các hàm số sơ cấp** là các hàm số được tạo lên từ các hàm số sơ cấp cơ bản bởi một số hữu hạn các phép toán sơ cấp, trong đó các phép toán sơ cấp đối với hàm số là các phép toán cộng, trừ, nhân, chia, phép lấy hàm hợp các hàm số sơ cấp cơ bản được thực hiện giống như đối với các biểu thức đại số và các phép toán đó được gọi là các *phép toán sơ cấp đối với hàm số*.

2.1.2. Giới hạn của hàm số

Lý thuyết giới hạn đề cập đến xu hướng biến thiên của biến phụ thuộc y khi biến độc lập x tiến dần đến một điểm x_0 cố định, tức là khi khoảng cách $|x - x_0|$ thu hẹp một cách tùy ý. Ta gọi đó là *quá trình x tiến đến x_0* và viết $x \to x_0$. Khi xét giới hạn của hàm số $y = f(x)$ khi $x \to x_0$, ta giả thiết rằng hàm số được xác định trong các khoảng $(a, x_0) \cup (x_0, b)$, còn tại chính điểm x_0 hàm số có thể xác định hoặc không xác định. Quá trình $x \to x_0$ được xem xét với giả thiết $x \neq x_0$.

Một phương pháp xem xét giới hạn của hàm số $f(x)$ khi $x \to x_0$ là "dẫn" biến độc lập x theo một dãy số x_n có giới hạn bằng x_0 (x_n được lấy từ miền xác định của hàm số và $x_n \neq x_0$) và xét giới hạn của dãy các giá trị tương ứng của hàm số: $y_n = f(x_n)$.

Ví dụ 2.1.8. Tính giới hạn $\lim\limits_{x\to 1}(x^2 + 1)$.
Giải: Đặt $f(x) = x^2 + 1$. Từ đồ thị của hàm số trong hình 2.1, ta thấy rằng khi $x \to 1$ từ hai phía thì $f(x) \to 2$ và ta có thể viết $\lim\limits_{x\to 1}(x^2 + 1) = 2$.



Hình 2.1:

Lập bảng tính các giá trị ta cũng nhận được cùng kết luận. Khi x tiến dần đến 1 từ bên trái và bên phải thì $f(x)$ cũng tiến dần đến 2 từ hai phía.

$x \to 1$ $\qquad\qquad$ $x \to 1$

x	0,9	0,99	0,999	1	1,001	1,01	1,1
$f(x)$	1,81	1,98	1,998	2	2,002	2,02	2,21

$f(x) \to 2$ $\qquad\qquad$ $f(x) \to 2$

Định nghĩa 2.1.9. Cho điểm $x_0 \in (a, b)$ và hàm số $f(x)$ xác định trên khoảng (a, b) hoặc $(a, b) \setminus x_0$. Ta nói rằng $f(x)$ *có giới hạn* là số A (hữu hạn) khi x dần đến x_0 nếu với bất kỳ dãy $x_n \to x_0; x_n \in (a, b)$ thì $\lim\limits_{n\to\infty} f(x_n) = A$. Ký hiệu: $\lim\limits_{x\to x_0} f(x) = A$ hoặc $f(x) \to A$ khi $x \to x_0$.

Định nghĩa nêu trên cũng áp dụng cho các trường hợp khi $x \to -\infty; x \to +\infty; A$ là một số hữu hạn hay $A = \pm\infty$. Với x_0 là số thực thì giới hạn của hàm số khi $x \to x_0$ còn được gọi là *giới hạn tại điểm* x_0.

Ví dụ 2.1.10. Tính giới hạn của hàm số $\dfrac{4x^2 - 16}{x - 2}$ khi $x \to 2$.

Giải: Tại $x = 2$ hàm số trên không xác định.

Giả sử x_n là một dãy số bất kỳ thỏa mãn $x_n \neq 2$ và $x_n \to 2$ khi $n \to +\infty$ ta có

$$\lim_{x\to 2} f(x_n) = \lim_{x\to 2} \frac{4x_n^2 - 16}{x_n - 2} = \lim_{x\to 2} \frac{4(x_n - 2)(x_n + 2)}{x_n - 2} = \lim_{x\to 2} 4(x_n + 2) = 16.$$

Vậy $\lim\limits_{x\to 2} f(x) = 16$.

• Các tính chất của giới hạn

Định lí 2.1.11. *Khi $x \to x_0$ hoặc $x \to \infty$ trong cùng quá trình đó mà $f_1(x) \to L_1; f_2(x) \to L_2$ thì:*

(1) $cf_1(x) \to cL_1$ *với c là hằng số;*

(2) $[f_1(x) \pm f_2(x)] \to L_1 \pm L_2$*;*

(3) $[f_1(x).f_2(x)] \to L_1.L_2$*;*

(4) $\left[\dfrac{f_1(x)}{f_2(x)}\right] \to \dfrac{L_1}{L_2}$, *với* $L_2 \neq 0$.

Chú ý 2.1.12.

(1) Nếu $P_n(x)$ là một đa thức bậc n đối với x thì $\lim\limits_{x\to x_0} P_n(x) = P_n(x_0)$.

(2) Nếu $R(x) = \dfrac{a_0 + a_1x + \ldots + a_nx^n}{b_0 + b_1x + \ldots + b_mx^m} = \dfrac{P_n(x)}{Q_m(x)}$ thì $\lim\limits_{x\to x_0} R(x) = \dfrac{P_n(x_0)}{Q_m(x_0)}$ với $Q_m(x_0) \neq 0$ ($R(x)$ là một phân thức hữu tỉ).

(3) Các dạng vô định: $\dfrac{0}{0}; \dfrac{\infty}{\infty}; 0.\infty; \infty - \infty; 1^\infty; \ldots$ Khi gặp các dạng vô định đó, muốn biết cụ thể của giới hạn phải tìm cách biến đổi để làm mất các dạng đó gọi là khử các dạng vô định.

Ví dụ 2.1.13. Tính các giới hạn sau.

(a) $\lim\limits_{x\to 1} \dfrac{x^n - 1}{x^m - 1} = \lim\limits_{x\to 1} \dfrac{(x-1)(1 + x + \ldots + x^{n-1})}{(x-1)(1 + x + \ldots + x^{m-1})} = \dfrac{n}{m}$.

(b) $\lim\limits_{x\to 0} \dfrac{\sqrt[3]{1+x} - \sqrt[5]{1+x}}{x} = \lim\limits_{x\to 0} \dfrac{(\sqrt[3]{1+x} - 1) + (1 - \sqrt[5]{1+x})}{x} = \dfrac{1}{3} - \dfrac{1}{5} = \dfrac{2}{15}$.

(đặt $y = \sqrt[3]{1+x} \Rightarrow x = y^3 - 1 \ldots$ rồi áp dụng ví dụ trên).

Định nghĩa 2.1.14. (1) Khi $x \to x_0$ từ những giá trị bên trái $x_0 (x < x_0)$ mà hàm $f(x)$ dần tới một số xác định A, thì số A gọi là *giới hạn trái của* $f(x)$ *tại* x_0. Ký hiệu $f(x_0 - 0) = \lim\limits_{x\to x_0^-} f(x) = A$.

(2) Khi $x \to x_0$ từ những giá trị bên phải $x_0 (x > x_0)$ mà hàm $f(x)$ dần tới một số xác định B, thì số B gọi là *giới hạn phải của* $f(x)$ *tại* x_0. Ký hiệu $f(x_0 + 0) = \lim\limits_{x\to x_0^+} f(x) = B$.

Vậy $f(x)$ chỉ có giới hạn tại x_0 khi và chỉ khi tồn tại giới hạn trái A và tồn tại giới hạn phải B và $A = B$.

Ví dụ 2.1.15. Tìm giới hạn của hàm số $y = \dfrac{|x|}{x} + x$ khi $x \to 0$.

Giải: Ta có $y = \begin{cases} 1+x & \text{khi } x > 0 \\ x-1 & \text{khi } x < 0 \end{cases}$

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+}(1+x) = 1;$$

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-}(-1+x) = -1.$$

Vậy hàm số trên không có giới hạn tại $x = 0$.

Định nghĩa 2.1.16. (1) Hàm số $f(x)$ được gọi là *vô cùng bé* (VCB) khi $x \to x_0$ nếu $\lim\limits_{x\to x_0} f(x) = 0$;

(2) Hàm số $g(x)$ được gọi là *vô cùng lớn* (VCL) khi $x \to x_0$ nếu $\lim\limits_{x\to x_0} f(x) = +\infty$.

(Số x_0 có thể là hữu hạn hoặc vô hạn).

• So sánh các VCB

(1) Cho $f_1(x)$ và $f_2(x)$ là hai VCB khi $x \to x_0$ ta nói rằng $f_1(x)$ *có bậc cao hơn* $f_2(x)$ nếu $\lim\limits_{x\to x_0} \dfrac{f_1(x)}{f_2(x)} = 0$. Khi đó ta cũng nói rằng $f_2(x)$ *có bậc thấp hơn* $f_1(x)$ trong quá trình $x \to x_0$.

(2) $f_1(x)$ *cùng bậc* với $f_2(x)$ nếu $\lim\limits_{x\to x_0} \dfrac{f_1(x)}{f_2(x)} = c \neq 0$.

(3) Đặc biệt nếu $c = 1$ thì ta nói rằng $f_1(x)$ *tương đương* với $f_2(x)$ khi $x \to x_0$ và viết $f_1(x) \sim f_2(x), x \to x_0$.

Định lí 2.1.17. *Nếu $f_1(x), f_2(x), g_1(x), g_2(x)$ là những VCB $x \to x_0$, nếu $f_1(x) \sim f_2(x)$ và $g_1(x) \sim g_2(x)$ thì*

$$\lim_{x\to x_0} \frac{f_1(x)}{g_1(x)} = \lim_{x\to x_0} \frac{f_2(x)}{g_2(x)}.$$

Chú ý 2.1.18. Nếu $\lim\limits_{x\to x_0} f(x) = A$ thì ta có thể viết $f(x) = A + \alpha(x)$ trong đó $\alpha(x)$ là VCB khi $x \to x_0$.

• Một số giới hạn cơ bản

(1) $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$ (2) $\lim\limits_{x\to 0} \dfrac{\log_a(1+x)}{x} = \log_a e$

(3) $\lim\limits_{x\to 0}(1+x)^{\frac{1}{x}} = \lim\limits_{x\to\infty}(1+\dfrac{1}{x})^x = e$ (4) $\lim\limits_{x\to 0} \dfrac{a^x - 1}{x} = \ln_a$

(5) $\lim\limits_{x\to 0} \dfrac{(1+x)^m - 1}{x} = m$

Ví dụ 2.1.19. Tính giới hạn sau $\lim\limits_{x\to 0} \dfrac{(1+\sin x)^2 - 1}{\sin x}$.

Giải: $\lim\limits_{x\to 0} \dfrac{(1+\sin x)^2 - 1}{\sin x} = \lim\limits_{x\to 0} \dfrac{1 + 2\sin x + \sin^2 x - 1}{\sin x} = \lim\limits_{x\to 0}(2 + \sin x) = 2.$

2.1.3. Sự liên tục của hàm số

Định nghĩa 2.1.20. Cho $f(x)$ là một hàm xác định trong miền D, nói rằng $f(x)$ *liên tục tại* x_0 nếu:

(1) $x_0 \in D$;

(2) Tồn tại giới hạn của hàm số $f(x)$ tại điểm x_0, tức là tồn tại $\lim\limits_{x\to x_0} f(x)$;

(3) Giới hạn của hàm số tại x_0 bằng giá trị của hàm số tại điểm đó, tức là

$$\lim_{x\to x_0} f(x) = f(x_0).$$

Định nghĩa 2.1.21. (1) Hàm $f(x)$ được gọi là *liên tục bên phải* (hoặc *bên trái*) tại điểm x_0 nếu

$$f(x_0+0) = \lim_{x\to x_0^+} f(x) = f(x_0)(\text{ hoặc } f(x_0-0) = \lim_{x\to x_0^-} f(x) = f(x_0)).$$

(2) Hàm $f(x)$ được gọi là *liên tục trong khoảng* (a,b) nếu $f(x)$ liên tục tại mọi $x \in (a,b)$.

Ví dụ 2.1.22. Hàm $f(x) = |x|$ liên tục tại $x = 0$ vì $\lim\limits_{x\to 0} |x| = 0$.

Chú ý 2.1.23.

(1) Các đa thức là những hàm số liên tục; phân thức hữu tỷ là hàm liên tục trừ các không điểm của đa thức mẫu số; các hàm số lượng giác liên tục trong miền xác định của nó.

(2) Nếu hàm $f(x)$ liên tục tại điểm x_0 thì $\lim\limits_{x\to x_0} f(x) = f\left(\lim\limits_{x\to x_0} x\right)$.

(3) Nếu hàm $f(x)$ không liên tục tại điểm x_0 thì nó gián đoạn tại đó và điểm x_0 được gọi là điểm gián đoạn. Từ đó ta thấy hàm $f(x)$ gián đoạn tại x_0 khi hoặc x_0 không thuộc miền xác định của $f(x)$; hoặc x_0 thuộc miền xác định của $f(x)$ nhưng $\lim\limits_{x\to x_0} f(x) \neq f(x_0)$; hoặc không có giới hạn $\lim\limits_{x\to x_0} f(x)$.

(4) Có 2 loại điểm gián đoạn:

Loại 1: $f(x)$ gián đoạn tại x_0 nhưng tại x_0 có giới hạn phải (hữu hạn) và giới hạn trái (hữu hạn). Khi đó hiệu của 2 giới hạn đó được gọi là bước nhảy.

Loại 2: Các trường hợp còn lại.

Chú ý 2.1.24. Nếu hàm $f(x)$ liên tục tại x_0 thì:

(1) $\lim\limits_{x\to x_0} \sqrt{f(x)} = \sqrt{\lim\limits_{x\to x_0} f(x)}$;

(2) $\lim\limits_{x\to x_0} e^{f(x)} = e^{\lim\limits_{x\to x_0} f(x)}$;

(3) $\lim\limits_{x\to x_0} u(x)^{v(x)} = \lim\limits_{x\to x_0} e^{v(x)\ln u(x)} = e^{\lim\limits_{x\to x_0} v(x)\ln u(x)}$.

Ví dụ 2.1.25. Tính giới hạn $\lim\limits_{x\to +\infty} \left(\dfrac{x^2-1}{x^2+1}\right)^{x^2}$.

Giải: Ta có:

$$\lim_{x\to+\infty}\left(\frac{x^2-1}{x^2+1}\right)^{x^2} = \lim_{x\to+\infty}\left(1-\frac{2}{x^2+1}\right)^{x^2} = \lim_{x\to+\infty}\left\{\left(1-\frac{2}{x^2+1}\right)^{-\frac{x^2+1}{2}}\right\}^{\frac{-2x^2}{x^2+1}} = \frac{1}{e^2}.$$

2.1.4. Đạo hàm của hàm số một biến số

Đạo hàm được Newton phát minh ra giúp ông giải quyết được bài toán xác định vận tốc, gia tốc chất điểm và ở đây ông đã đưa ra một ý nghĩa rất tổng quát: *Đạo hàm cho chúng ta biết được tốc độ biến thiên (tốc độ thay đổi) của một hàm số*. Điều này rất quan trọng và nó có rất nhiều ứng dụng trong nhiều ngành khoa học, kinh tế, xã hội. Với đạo hàm, bất cứ ở đâu có sự thay đổi, ở đó chúng ta sẽ biết được nó thay đổi như thế nào: liệu đại lượng đó đang tăng hay đang giảm hay đang không thay đổi, nếu là đang tăng vậy tăng nhanh hay tăng chậm... Nếu bạn là nhà kinh tế và muốn biết tốc độ tăng trưởng kinh tế nhằm đưa ra những quyết định đầu tư chứng khoán đúng đắn hoặc nếu bạn là nhà hoạch định chiến lược và muốn có những thông tin liên quan đến tốc độ gia tăng dân số ở từng vùng miền hoặc nếu bạn là nhà hóa học và muốn xác định được tốc độ phản ứng hóa học nào đó, hay nhà vật lý muốn tính toán vận tốc, gia tốc của một chuyển động... Đạo hàm sẽ là thứ mà chúng ta cần, đầu tiên bạn cần có hàm số mô tả đại lượng đang được quan tâm, và sau đó chỉ cần đạo hàm nó.

Trong chương này sẽ nhắc lại một số kiến thức về đạo hàm và tìm hiểu một số bài toán ứng dụng của đạo hàm trong kinh tế cũng như trong sản xuất nông lâm nghiệp.

Định nghĩa 2.1.26. Giả sử hàm số $y = f(x)$ xác định trên khoảng (a, b) và điểm x_0 thuộc khoảng đó. Giới hạn hữu hạn (nếu có) của tỷ số $\dfrac{f(x) - f(x_0)}{x - x_0}$ khi x tiến dần đến x_0 được gọi là *đạo hàm* của hàm số đã cho tại điểm x_0, ký hiệu $f'(x_0)$ hoặc $y'(x_0)$ nghĩa là

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}\cdot$$

Đặt $\Delta x = x - x_0; \Delta y = f(x) - f(x_0) = f(x_0 + \Delta x) - f(x_0)$ thì

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}, \tag{2.1}$$

trong đó Δx được gọi là *số gia của đối số* và Δy được gọi là *số gia của hàm số*.

Chú ý 2.1.27. Dùng liên hệ giữa giới hạn và vô cùng bé có thể biểu diễn hệ thức định nghĩa đạo hàm dưới dạng $f(x_0 + \Delta x) - f(x_0) = f'(x_0)\Delta x + 0(\Delta x)$ trong đó $0(\Delta x)$ là một VCB bậc cao hơn Δx khi $\Delta x \to 0$. Từ đó ta thấy nếu $f(x)$ có đạo hàm tại $x_0 \in (a, b)$ thì $f(x)$ liên tục tại x_0. Nhưng điều ngược lại chưa chắc đã đúng.

Ví dụ 2.1.28. Tìm đạo hàm của hàm số $y = \ln x$ $(x > 0)$ bằng định nghĩa.
Giải: Ta có

$$\Delta y = \ln(x + \Delta x) - \ln x = \ln\left(\frac{x + \Delta x}{x}\right) = \ln\left(1 + \frac{\Delta x}{x}\right)$$

do đó

$$\frac{\Delta y}{\Delta x} = \frac{\ln\left(1 + \dfrac{\Delta x}{x}\right)}{\Delta x} = \frac{1}{x}\frac{\ln\left(1 + \dfrac{\Delta x}{x}\right)}{\dfrac{\Delta x}{x}} \to \frac{1}{x} \text{ khi } \Delta x \to 0.$$

Vậy $(\ln x)' = \dfrac{1}{x}(x > 0)$.

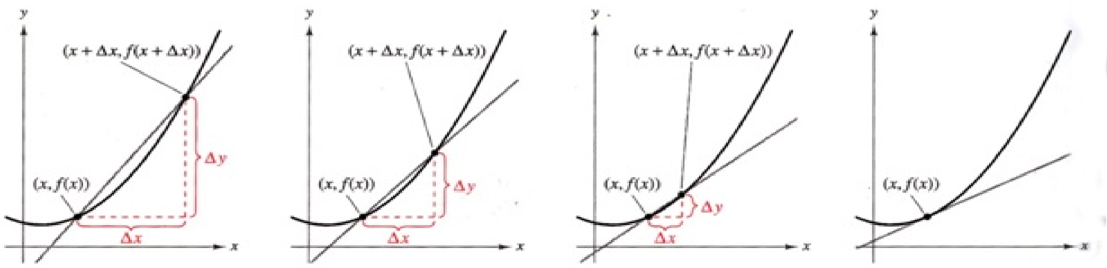
Ví dụ 2.1.29. Hàm $f(x) = |x|$ liên tục tại $x_0 = 0$ nhưng không có đạo hàm tại điểm đó vì không tồn tại giới hạn $\lim\limits_{\Delta x \to 0} \dfrac{f(0 + \Delta x) - f(0)}{\Delta x} = \lim\limits_{\Delta x \to 0} \dfrac{|\Delta x|}{\Delta x}$.

• Ý nghĩa hình học của đạo hàm của hàm số một biến số

Cho hàm số $y = f(x)$, một điểm $M(x, f(x))$ cố định thuộc đồ thị hàm số. Với mỗi điểm $N(x + \Delta x, f(x + \Delta x))$ thuộc đồ thị hàm số thì *hệ số góc* hay còn gọi là *độ dốc (slope)* của cát tuyến MN được xác định bởi

$$k_m = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Giả sử tồn tại giới hạn hữu hạn $k_0 = \lim\limits_{\Delta x \to 0} k_m = \lim\limits_{\Delta x \to 0} \dfrac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x_0)$ thì giới hạn đó được gọi là *độ dốc của hàm số* $y = f(x)$. Đường thẳng đi qua $M(x, f(x))$ và có hệ số góc k_0 là *vị trí giới hạn* của cát tuyến MN khi N di chuyển dọc theo đồ thị đến M và đường thẳng đó được gọi là *tiếp tuyến của đồ thị hàm số tại điểm* M, điểm M được gọi là *tiếp điểm.*



Như vậy:

Đạo hàm của hàm số $y = f(x)$ tại điểm x là hệ số góc của tiếp tuyến của đồ thị hàm số tại điểm đó.

• Các quy tắc tính đạo hàm

Định lí 2.1.30. *Cho $f(x)$ và $g(x)$ là hai hàm xác định trên (a, b); giả sử $f(x)$ và $g(x)$ có đạo hàm tại $x \in (a, b)$ khi đó:*

(1) $(f \pm g)'(x) = f'(x) \pm g'(x)$.

(2) $(cf)'(x) = cf'(x)$ *(với c là hằng số).*

(3) $(f.g)'(x) = f'(x).g(x) + f(x).g'(x)$.

(4) $\left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x).g(x) - f(x).g'(x)}{g^2(x)}$, *trong đó* $g(x) \neq \emptyset, \quad \forall x \in (a, b)$.

Ví dụ 2.1.31. Tính đạo hàm của hàm số $y = \dfrac{2x^2 - 3x + 1}{x}$.

Giải: Ta có $y' = \dfrac{(4x - 3)x - (2x^2 - 3x + 1)}{x^2} = \dfrac{2x^2 - 1}{x^2}$.

• Đạo hàm của hàm hợp và hàm ngược

Định lí 2.1.32. *Xét hàm hợp $y = y[u(x)]$. Nếu hàm $y = y(u)$ có đạo hàm đối với u, và $u = u(x)$ có đạo hàm đối với x thì hàm hợp $y = y[u(x)]$ có đạo hàm đối với x, và* $y'_x = y'_u.u'_x$.

Định lí 2.1.33. *Nếu hàm* $x = \varphi(y)$ *tại* y_0 *có đạo hàm* $\varphi'(y_0) \neq 0$ *và nếu hàm ngược của nó* $y = f(x)$ *liên tục tại điểm* $x_0 = \varphi(y_0)$ *thì tồn tại đạo hàm* $f'(x_0) = \dfrac{1}{\varphi'(y_0)}$.

Ví dụ 2.1.34. Tính đạo hàm của hàm số $(\ln|x|)', (x \neq 0)$.
Giải: Ta có, $(\ln|x|)', (x \neq 0)$ là một hàm số hợp nên

$$(\ln|x|)' = \frac{1}{|x|} \cdot (|x|)' = \begin{cases} \dfrac{1}{|x|} & \text{khi} \quad x > 0 \\ -\dfrac{1}{|x|} & \text{khi} \quad x < 0 \end{cases}$$

Vậy $(\ln|x|)' = \dfrac{1}{|x|}, x \neq 0$.

Ví dụ 2.1.35. Tính đạo hàm của hàm số $y = \arctan x$.
Giải: Hàm số $y = \arctan x$ là hàm ngược của hàm $x = \tan y \left(-\dfrac{\pi}{2} < y < \dfrac{\pi}{2}\right)$ nên $x'_y = \dfrac{1}{\cos^2 y} = 1 + \tan^2 y = 1 + x^2$.

• Đạo hàm một phía

Cho hàm số $y = f(x)$ xác định trong khoảng (a, b). Nếu tồn tại giới hạn hữu hạn $\lim\limits_{\Delta x \to 0^+} \dfrac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$ thì giới hạn đó được gọi là *đạo hàm bên phải của* f *tại* x_0. Ký hiệu $f'_+(x_0)$.

Nếu tồn tại giới hạn hữu hạn $\lim\limits_{\Delta x \to 0^-} \dfrac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$ thì giới hạn đó được gọi là *đạo hàm bên trái của* f *tại* x_0. Ký hiệu $f'_-(x_0)$.

Vậy, nếu hàm $f(x)$ có $f'_+(x_0)$ và $f'_-(x_0)$ và $f'_-(x_0) = f'_+(x_0)$ hàm $f(x)$ có đạo hàm tại x_0; nếu $f'_-(x_0) \neq f'_+(x_0)$ thì $f(x)$ không có đạo hàm tại x_0.

• Đạo hàm trên một khoảng - đoạn

Hàm số $y = f(x)$ xác định trên miền D được gọi là *có đạo hàm trên khoảng* $(a, b) \subset D$ nếu nó có đạo hàm tại mọi điểm trên khoảng (a, b).

Hàm số $y = f(x)$ xác định trên miền D được gọi là *có đạo hàm trên đoạn* $[a, b] \subset D$ nếu nó có đạo hàm tại mọi điểm trên khoảng (a, b) và có đạo hàm phải tại điểm a, đạo hàm trái tại điểm b.

• Bảng đạo hàm một số hàm sơ cấp

Hàm số y	Đạo hàm y'
c	0
x^α	$\alpha x^{\alpha-1}$
$\frac{1}{x}$	$-\frac{1}{x^2}$
a^x	$a^x \ln a \ a > 0$
$\log_a \lvert x\rvert$	$\frac{1}{x\ln a}$
$\ln \lvert x\rvert$	$\frac{1}{x}$

Hàm số y	Đạo hàm y'
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\frac{1}{\cos^2 x}$
$\cot x$	$-\frac{1}{\sin^2 x}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
$\operatorname{arccot} x$	$-\frac{1}{1+x^2}$

Ví dụ 2.1.36. Tính đạo hàm của các hàm số sau:

$$(a)\ y = 5e^{2x} - 2.3^x + \sqrt{x^3} \qquad (b)\ y = x.\sin x + \ln(x+1) - e^{4x}$$

Giải:

(a) Ta có $y' = 5(2x)'e^{2x} - 2.3^x.\ln 3 + \frac{x^{3/2-1}3/2-1}{=}10e^{2x} - 23^x \ln 3 + 2\sqrt{x}.$

(b) Ta có $y' = \sin x + x.\cos x + \frac{1}{x+1} - 4.e^{4x}.$

2.1.5. Một số bài toán ứng dụng của đạo hàm

Đạo hàm của hàm số một biến số được dùng để xem xét tỷ lệ biến đổi của giá trị hàm số y theo đối số x. Cho hàm số $y = f(x)$ xác định trên đoạn $[a, b]$ thì tỷ số

$$\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

được gọi là *tỷ lệ biến đổi trung bình* của y theo x trên đoạn $[x, x+\Delta x]$.

Và giới hạn của tỷ lệ biến đổi trung bình trên đoạn $[x, x+\Delta x]$, khi $\Delta x \to 0$

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0}\frac{f(x+\Delta x) - f(x)}{\Delta x} = f'(x)$$

được gọi là *tỷ lệ biến đổi tức thời*.

Ví dụ 2.1.37. Cho hàm số $y = f(x) = x^3$. Hãy ước lượng $f'(2)$ dựa vào tỷ lệ biến đổi tức thời của hàm số tại $x = 2$.

Giải: Vì $f'(2)$ là đạo hàm hay chính là tốc độ thay đổi của hàm số $f(x) = x^3$ tại điểm $x = 2$. Ta xét tỷ lệ biến đổi trung bình của y theo x trên khoảng $2 \le x \le 2,001$, chúng ta thấy rằng tốc độ trung bình của sự thay đổi trên $2 \le x \le 2,001$ là

$$\frac{2,001^3 - 2^3}{2,001 - 2} = \frac{8,0120 - 8}{0,001} \approx 12,006.$$

Như vậy, tốc độ thay đổi của $f(x)$ tại $x = 2$ khoảng $12,006$, vì vậy chúng ta ước tính $f'(2) = 12,006$.

• Trong lĩnh vực vật lý

Nếu y là hàm số chỉ quãng đường hay khoảng cách theo thời gian, x là biến số chỉ thời gian, thì tỷ lệ biến đổi tức thời chính là vận tốc tức thời (thường gọi chung là vận tốc) của một vật tại một thời điểm.

Vận tốc trung bình của một vật mà di chuyển trong khoảng thời gian xác định được đo bởi

$$\begin{aligned} v_{tb} &= \frac{\text{Quãng đường vật đi được}}{\text{Thời gian để đi được quãng đường trên}} \\ &= \frac{\Delta s}{\Delta t} = \frac{s(t) - s(t_0)}{t - t_0}. \end{aligned}$$

Khi đó *vận tốc tức thời* của vật tại 1 thời điểm (đặc trưng cho mức độ chuyển động nhanh hay chậm của vật tại 1 thời điểm đó) là giới hạn hữu hạn

$$v(t_0) = \lim_{t \to t_0} \frac{s(t) - s(t_0)}{t - t_0} = s'(t_0).$$

Gia tốc tức thời của vật tại 1 thời điểm (đặc trưng cho tốc độ biến thiên của vận tốc) là giới hạn hữu hạn

$$a(t) = \lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t} = v'(t) = s''(t).$$

Ví dụ 2.1.38. Một vật rơi tự do từ độ cao $100m$, bỏ qua lực cản của không khí, gọi h là độ cao của vật sau khoảng thời gian t (giây) và được tính bởi

$$h = -16t^2 + 100.$$

Tìm vận tốc tức thời của vật rơi tại thời điểm $t = 1$.
Giải: Ta có vận tốc tức thời của vật rơi tại thời điểm $t = 1$ chính là đạo hàm (tỷ lệ biến đổi tức thời) của hàm h theo thời gian t. Tức là: $v(t) = h'(t) = -32t$. Khi $t = 1$ thì $v(1) = -32(m/s)$ chính là vận tốc tức thời của vật rơi tại thời điểm 1 giây.

• Trong lĩnh vực kinh tế và nông nghiệp

Nếu $y = f(x)$ là hàm số biểu diễn ảnh hưởng giữa các nhân tố trong lĩnh vực kinh tế, các nhà kinh tế học đã sử dụng các khái niệm *lợi nhuận biên, doanh thu biên, chi phí biên,...* như là tỷ lệ biến đổi tức thời của tổng lợi nhuận, tổng doanh thu, tổng chi phí khi x đơn vị sản phẩm được sản xuất hay được bán ra. Nếu ký hiệu P là hàm tổng lợi nhuận, R là hàm tổng doanh thu, C là hàm tổng chi phí thì ta có:

$$\begin{aligned} P &= R - C, \\ \text{Lợi nhuận biên } &= \lim_{\Delta x \to 0} \frac{\Delta P}{\Delta x} = P'(x), \\ \text{Doanh thu biên } &= \lim_{\Delta x \to 0} \frac{\Delta R}{\Delta x} = R'(x), \\ \text{Chi phí biên } &= \lim_{\Delta x \to 0} \frac{\Delta C}{\Delta x} = C'(x). \end{aligned}$$

Như vậy, tại mỗi mức sản xuất nhất định, lợi nhuận biên, doanh thu biên, chi phí biên,... lần lượt cho biết xấp xỉ lượng lợi nhuận, lượng doanh thu, lượng chi phí tăng thêm khi sản xuất hay bán ra một đơn vị sản phẩm.

Ví dụ 2.1.39. Chi phí sản xuất t tạ phân bón của một nhà máy được cho bởi hàm số: $C = f(t)(\$)$. Biểu thức $f'(20) = 100$ có ý nghĩa gì?
Giải: $f'(20) = 100$ chính là chi phí biên khi nhà máy đó sản xuất 20 tạ phân bón. Tức là nhà máy sản xuất thêm một đơn vị sản phẩm tiếp theo (21 tạ phân bón) thì chi phí tăng thêm 100\$.

Ví dụ 2.1.40. Doanh thu (bằng \$) từ sản xuất q đơn vị sản phẩm được cho bởi hàm số $R(q) = 1000q - 3q^2$. Tìm $R(125)$ và $R'(125)$; đưa ra các đơn vị và nêu ý nghĩa của các biểu thức trên.
Giải: Ta có $R(125) = 1000.125 - 3.125^2 = 78125(\$)$.

Vì $R'(q) = 1000 - 6q$ nên $R'(125) = 1000 - 6.125 = 250(\$)$.

Nếu 125 đơn vị được sản xuất, doanh thu là 78125\$.

Nếu sản xuất thêm một đơn vị sản phẩm từ 125 sản phẩm lên 126 sản phẩm thì doanh thu tăng khoảng 250\$.

Một trong các ứng dụng quan trọng của đạo hàm là dùng để tìm giá trị lớn nhất, giá trị nhỏ nhất của hàm số trên một đoạn/khoảng. Nếu hàm số đã cho là mô hình hóa của các lĩnh vực kinh tế, nông, lâm, ngư nghiệp hay các lĩnh vực của đời sống xã hội thì tìm giá trị lớn nhất, nhỏ nhất của hàm số chính là tìm phương án tối ưu của bài toán thực tế, chẳng hạn như tìm mức sản xuất nhất định để lợi nhuận đạt cao nhất, tìm phương án để chi phí thấp nhất,...

Ví dụ 2.1.41. Khi nuôi cá thí nghiệm trong hồ, một nhà sinh vật học thấy rằng: Nếu trên mỗi đơn vị diện tích của mặt hồ có n con cá thì trung bình mỗi con cá sau một vụ cân nặng được đo bởi hàm số $P(n) = 480n - 20n^2(gam)$. Hỏi phải thả bao nhiêu cá trên một đơn vị diện tích của mặt hồ để sau một vụ sản lượng cá thu hoạch được là lớn nhất?
Giải: Ta có $D = R$; $\quad P' = 480 - 40n$; $\quad P' = 0 \Leftrightarrow n = 12$. Vì n là số cá được thả nên $n \geq 0$. Ta có bảng biến thiên

n	0		12		$+\infty$
$P'(n)$		$+$	0	$-$	
$P(n)$			2880		
	0	$\nearrow$		$\searrow$	$-\infty$

Vậy nếu thả 12 con trên một đơn vị diện tích thì sau một vụ sản lượng cá thu hoạch được là lớn nhất.

Ví dụ 2.1.42. Một người trồng cây ăn quả trên một mảnh vườn của nhà ước tính rằng nếu người đó trồng 16 cây thì trung bình năng suất là 80 quả táo trên cây. Nhưng vẫn với diện tích vườn đó, cứ trồng thêm mỗi một cây táo thì năng suất sẽ giảm đi 4

quả/cây. Hỏi nên trồng bao nhiêu cây trên mảnh vườn đó để tổng số táo thu được là lớn nhất?
Giải: Gọi x là số táo được trồng thêm. Vì trồng thêm mỗi một cây thì năng suất giảm đi 4 quả/cây nên nếu trồng thêm x cây thì năng suất giảm đi $4x$ quả trên cây. Vậy năng suất thực tế khi trồng $16 + x$ cây là $80 - 4x$. Gọi R là số táo thu được trên mảnh vườn đó thì

$$R = (16 + x)(80 - 4x) = -4x^2 + 16x + 1280, \quad (x \geq 0).$$

Ta có $R' = -8x + 16$; $\quad R' = 0 \Leftrightarrow x = 2$.

x	0		2		$+\infty$
$R'(x)$		$+$	0	$-$	
$R(x)$			1296		
		$\nearrow$		$\searrow$	
	0				$-\infty$

Nhìn vào bảng biên thiên ta thấy nếu ta trồng thêm 2 cây tức là mảnh vườn đó trồng 18 cây thì thu hoạch được nhiều táo nhất là 1296 quả.

2.1.6. Đạo hàm cấp cao của hàm số một biến

Định nghĩa 2.1.43. Giả sử $f(x)$ có đạo hàm tại mọi x thuộc $(a, b), f'(x) = g(x)$ cũng là một hàm số xác định trên (a, b). Nếu $g(x)$ có đạo hàm tại $x_0 \in (a, b)$ khi đó $g'(x) = [f'(x)]'$ được gọi là đạo *hàm cấp hai của hàm $f(x)$ tại x_0*. Ký hiệu: $f''(x)$.

Tổng quát, đạo hàm của của đạo hàm cấp $n - 1$ được gọi là *đạo hàm cấp n*. Ký hiệu: $f^{(n)} = [f^{(n-1)}(x)]'$.

Ví dụ 2.1.44. Tính đạo hàm cấp n của hàm $y = \sin x$.
Giải: Ta có:

$$\begin{aligned} y' &= \cos x = \sin\left(x + \frac{\pi}{2}\right) \\ y'' &= \cos\left(x + \frac{\pi}{2}\right) = \sin\left(x + 2\frac{\pi}{2}\right) \\ y''' &= \cos\left(x + 2\frac{\pi}{2}\right) = \sin\left(x + 3\frac{\pi}{2}\right) \end{aligned}$$

Tổng quát: $y^{(n)} = \sin\left(x + n\dfrac{\pi}{2}\right)\cdot$

• **Công thức Leibniz**

Giả sử $u = u(x), v = v(x)$ đều có đạo hàm đến cấp n khi đó hàm $y = u.v$ cũng có đạo hàm cấp n và

$$y^{(n)} = (uv)^{(n)} = \sum_{i=0}^{n} C_n^i u^{(n-i)} v^{(i)}.$$

Ví dụ 2.1.45. Tính đạo hàm cấp 10 của hàm số $y = (2x^2 - 1)e^{2x}$.
Giải: Đặt $u = 2x^2 - 1$ và $v = e^{2x}$. Khi đó,

$$\begin{array}{llll} u' = 4x; & u'' = 4; & u''' = \cdots u^{(10)} = 0. & \\ v' = 2e^{2x}; & v'' = 2^2e^{2x}; \ldots & & v^{(10)} = 2^{10}e^{2x}. \end{array}$$

Áp dụng công thức Leibniz ta được

$$\begin{aligned} y^{(10)} &= C_{10}^0(2x^2 - 1).2^{10}e^{2x} + C_{10}^1 4x.2^9e^{2x} + C_{10}^2 4.2^8e^{2x} \\ &= (2x^2 - 1).2^{10}e^{2x} + 10.2x.2^{10}e^{2x} + 45.2^{10}e^{2x} \\ &= 2^{10}e^{2x}(2x^2 + 20x + 44) \end{aligned}$$

2.1.7. Vi phân của hàm số một biến số

Xét hàm $f(x)$ liên tục tại x_0. Cho x một số gia Δx thì $f(x)$ có số gia $\Delta f = f(x_0 + \Delta x) - f(x_0) \to 0$ khi $\Delta x \to 0$. Nghĩa là Δf là VCB đồng thời với Δx. Nói chung việc tính Δf khá phức tạp. Bởi thế ta hãy tìm VCB tương đương với Δf, nhưng đơn giản hơn Δf.

Chẳng hạn xét hàm $f(x) = x^2$. Ta có $\Delta f = (x_0 + \Delta x)^2 - x_0^2 = 2x_0\Delta x + (\Delta x)^2$.

Vì $(\Delta x)^2$ là VCB cấp cao hơn Δx nên nếu $x_0 \neq 0$ thì $\Delta f \sim 2x_0\Delta x$. Ta thấy VCB $2x_0\Delta x$ đơn giản hơn Δf, nó chỉ là tuyến tính đối với Δx. Người ta gọi nó là vi phân của hàm $f(x) = x^2$ tại x_0.

Định nghĩa 2.1.46. Vi phân
Xét hàm số $y = f(x)$ xác định tại x_0 và lân cận x_0. Cho x số gia tuỳ ý Δx sao cho $x_0 + \Delta x$ vẫn thuộc lân cận đó. Nếu số gia $\Delta f = f(x_0 + \Delta x) - f(x_0)$ của hàm số có dạng

$$\Delta f = A.\Delta x + \alpha(\Delta x),$$

trong đó A chỉ phụ thuộc x_0 chứ không phụ thuộc $\Delta x, \alpha(\Delta x)$ là VCB cấp cao hơn Δx thì ta nói rằng $f(x)$ *khả vi tại* x_0 và biểu thức $A.\Delta x$ được gọi là *vi phân của hàm* $f(x)$ *tại* x_0. Ký hiệu df. Vậy

$$df = A.\Delta x.$$

• Liên hệ giữa vi phân và đạo hàm

Nếu hàm $f(x)$ khả vi tại x_0 thì nó có đạo hàm tại x_0 và $f'(x_0) = A$;

Nếu $f(x)$ có đạo hàm tại x_0 thì nó khả vi tại x_0 và $df = f'(x_0)\Delta x$. Vậy, công thức tính vi phân của hàm $f(x)$ tại x là $df = f'(x)\Delta x$.

Đặc biệt nếu xét hàm số $f(x) = x$, thì $dx = 1.\Delta x$, nghĩa là $\Delta x = dx$. Vậy

$$df = f'(x)dx \quad \text{hoặc} \quad f'(x) = \frac{df}{dx}\cdot$$

Chú ý 2.1.47. Để ký hiệu đạo hàm của hàm số ta có thể sử dụng một trong các cách ký hiệu sau đây: $y'(x); \dfrac{dy}{dx}; \dfrac{d}{dx}(f(x))$.

• Tính bất biến của biểu thức vi phân

Xét hàm hợp $y = f(x); x = g(t)$ trong đó t là biến độc lập khi đó $y = f[g(t)]$. Theo công thức tính vi phân và theo quy tắc lấy đạo hàm hàm hợp ta có

$$dy = \{f[g(t)]\}'_t dt = f'(x).g'(t)dt. \text{ Vì } dx = g'(t)dt \text{ nên } dy = f'(x)dx.$$

Vậy dạng của vi phân của hàm $y = f(x)$ không thay đổi dù x là biến độc lập hay là hàm khả vi của một biến độc lập khác. Người ta gọi đó là *tính bất biến* của biểu thức vi phân (cấp một).

Ví dụ 2.1.48. Cho y là hàm của biến x thông qua biến t như sau:

$$x = x(t); \quad y = y(t).$$

Hãy tính y'_x?

Giải: Do tính bất biến của biểu thức vi phân ta có $dy = y'dx$ hay $y'_x = \dfrac{dy}{dx}\cdot$

Nhưng theo đầu bài thì $dy = y'(t)dt; dx = x'(t)dt.$

Vậy $y'_x = \dfrac{dy}{dx} = \dfrac{y'(t)dt}{x'(t)dt} = \dfrac{y'(t)}{x'(t)}\cdot$

Điều cần chú ý là dấu đạo hàm ở vế phải chỉ đạo hàm đối với t.

• **Quy tắc tính vi phân**

(1) $d(f+g) = df + dg;$

(2) $d(cf) = cdf;$

(3) $d(fg) = df.g + f.dg;$

(4) $d\left(\dfrac{f}{g}\right) = \dfrac{df.g - f.dg}{g^2}$ (nếu $g(x) \neq 0$).

• **Ứng dụng của vi phân vào phép tính gần đúng**

Theo định nghĩa đạo hàm $f'(x_0) = \lim\limits_{\Delta x \to 0} \dfrac{\Delta f}{\Delta x}\cdot$ Với giá trị Δx đủ nhỏ ta có

$$\frac{\Delta f}{\Delta x} \approx f'(x_0) \text{ hay } \Delta f(x) \approx f'(x_0)\Delta x = df(x_0).$$

Vậy ta có công thức tính gần đúng

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x.$$

Ví dụ 2.1.49. Tính gần đúng các biểu thức sau:

$$(a)\ \sqrt[4]{17} \qquad (b)\ \sin 46^o.$$

Giải: (a) Xét hàm $f(x) = \sqrt[4]{x} \quad \Rightarrow f'(x) = \dfrac{1}{4\sqrt[4]{x^3}}\cdot$

Áp dụng công thức tính gần đúng $f(x_0 + \Delta x) = \sqrt[4]{x_0 + \Delta x} \approx \sqrt[4]{x_0} + \dfrac{1}{4\sqrt[4]{x_0^3}}\Delta x.$

Chọn $x_0 = 16; \Delta x = 1 \quad \Rightarrow \sqrt[4]{17} \approx \sqrt[4]{16} + \dfrac{1}{4\sqrt[4]{16^3}} \cdot 1 \approx 2,031$

(b) Xét hàm $f(x) = \sin x \quad \Rightarrow f'(x) = \cos x;$

Chọn $x_0 = 45^o = \frac{\pi}{4}; \Delta x = 1^o = \frac{\pi}{180}\cdot$

Vậy $\sin 46^o = \sin\left(\frac{\pi}{4} + \frac{\pi}{180}\right) \approx \sin\frac{\pi}{4} + \cos\frac{\pi}{4}\cdot\frac{\pi}{180} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\cdot 0,017 \approx 0,7194.$

• **Vi phân cấp cao**

Định nghĩa 2.1.50. Xét hàm $f(x)$ khả vi tại mọi x thuộc khoảng nào đó. Vi phân của vi phân cấp một được gọi là *vi phân cấp hai của hàm số*. Ký hiệu $d^2y = d(dy)$.

Một cách tổng quát ta có: Vi phân của vi phân cấp $(n-1)$ được gọi là *vi phân cấp* n. Ký hiệu

$$d^n y = d(d^{(n-1)}y) = y^{(n)}dx^n \quad \text{(nếu } x \text{ là biến độc lập).}$$

Chú ý 2.1.51. Vi phân cấp cao không có tính bất biến. Cụ thể là xét hàm $y = f(x)$ nếu x không phải là biến độc lập mà x là hàm của một biến độc lập khác (chẳng hạn x= $\varphi(t)$) thì công thức trên không còn đúng nữa.Vì khi đó $dx \neq 0 (dx = x'_t dt)$. Khi đó

$$d^2y = d(y'_x.dx) = dy'_x.dx + y'_x.d(dx).$$

Bằng cách dùng tính bất biến của dạng vi phân (cấp 1) thì $dy'_x = y''_{x^2}dx$. Vậy

$$d^2y = y''_{x^2}dx^2 + y'_x d^2x.$$

Ví dụ 2.1.52. Tính vi phân cấp hai của hàm số $f(x) = x^2$.
Giải: Nếu x là biến độc lập thì: $df = 2xdx$ và $d^2f = 2(dx)^2$.

Nếu x không phải là biến độc lập mà x là hàm của biến độc lập khác chẳng hạn $x = t^2$, khi đó $f = t^4$ và $df = 4t^3dt; d^2f = 12t^2(dt)^2$. Theo trên ta có $d^2f = 2(dx)^2$. Nếu thế $dx = 2tdt$ vào thì

$$d^2f = 2(2tdt)^2 = 8t^2(dt)^2 \neq 12t^2(dt)^2.$$

2.2. Hàm số hai biến số

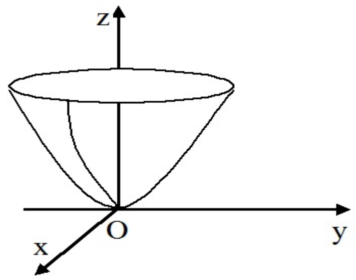
Định nghĩa 2.2.1. Hàm hai biến
Xét tích đề các $\mathbb{R}^2$ và tập hợp $D \subset \mathbb{R}^2$. Ta gọi ánh xạ $f : D \to \mathbb{R}$ là một *hàm hai biến số* xác định trên D, D được gọi là *miền xác định* của hàm f.

Nếu đặt $z = f(x, y)$ thì $f : (x, y) \mapsto z = f(x, y)$ trong đó x, y được gọi là các *biến độc lập*, z được gọi là *biến phụ thuộc*.

Ví dụ 2.2.2. Hàm $z = x^2 + y^2$ có đồ thị là paraboloid tròn xoay. Miền xác định của nó là toàn bộ mặt phẳng xOy.

2.2.1. Giới hạn và tính liên tục của hàm hai biến

Định nghĩa 2.2.3. Ta nói rằng *dãy điểm* $\{M_n(x_n, y_n)\}$ *dần tới điểm* $M_0(x_0, y_0) \subset \mathbb{R}^2$ *khi* $n \to \infty$ nếu $\lim\limits_{n\to\infty} x_n = x_0$ và $\lim\limits_{n\to\infty} y_n = y_0$ hay $\lim\limits_{n\to\infty} d(M_n, M_0) = 0$.



Định nghĩa 2.2.4. Giả sử $z = f(M) = f(x, y)$ là hàm số xác định trong một lân cận V nào đó của M_0 (có thể trừ tại M_0). Ta nói hàm $f(x, y)$ có *giới hạn* là L khi $M(x, y)$ dần tới $M_0(x_0, y_0)$ nếu mọi dãy $\{M_n(x_n, y_n)\} \subset V$ dần đến M_0 ta đều có $\lim\limits_{n\to\infty} f(x_n, y_n) = L$. Khi đó ta viết

$$\lim_{\substack{x\to x_0\\ y\to y_0}} f(x, y) = L.$$

Chú ý 2.2.5. Các định lý về giới hạn tổng, tích, thương đối với hàm số một biến số cũng đúng với hàm số nhiều biến số.

Ví dụ 2.2.6.

(a) Chứng minh rằng không tồn tại giới hạn sau $\lim\limits_{\substack{x\to 0\\ y\to 0}} \dfrac{2xy}{x^2 + y^2}$.

Theo Định nghĩa 2.2.4 ta chỉ cần chọn ra 2 dãy $M_n(x_n, y_n)$ và $M'_n(x'_n, y'_n)$ mà $(x_n, y_n) \to (0, 0); (x'_n, y'_n) \to (0, 0)$ nhưng giới hạn hàm tương ứng lại khác nhau.

Chọn $\{x_n, y_n\} = \left\{\dfrac{1}{n}, \dfrac{1}{n}\right\} \xrightarrow{n\to\infty} (0; 0)$. Ta có

$$\lim_{\substack{x\to 0\\ y\to 0}} \frac{2xy}{x^2 + y^2} = \lim_{n\to\infty} \frac{\frac{2}{n^2}}{\frac{2}{n^2}} = 1.$$

$$\{x'_n, y'_n\} = \left\{\frac{1}{n}, 0\right\} \xrightarrow{n\to\infty} (0; 0). \text{ Ta có}$$

$$\lim_{\substack{x\to 0\\ y\to 0}} \frac{2xy}{x^2 + y^2} = \lim_{n\to\infty} \frac{0}{\frac{1}{n^2}} = 0.$$

Vậy hàm số trên không có giới hạn.

(b) Tìm $\lim\limits_{\substack{x\to +\infty\\ y\to +\infty}} \left(\dfrac{xy}{x^2 + y^2}\right)^{x^2}$.

Ta có $x^2 + y^2 \geq 2xy \Rightarrow \dfrac{xy}{x^2 + y^2} \leq \dfrac{xy}{2xy} = \dfrac{1}{2} \Rightarrow 0 < \left(\dfrac{xy}{x^2 + y^2}\right)^{x^2} \leq \left(\dfrac{1}{2}\right)^{x^2}$.

Nhưng $\left(\dfrac{1}{2}\right)^{x^2} \xrightarrow{x\to +\infty} 0$. Vậy $\lim\limits_{\substack{x\to +\infty\\ y\to +\infty}} \left(\dfrac{xy}{x^2 + y^2}\right)^{x^2} = 0$.

Định nghĩa 2.2.7. Hàm hai biến liên tục Giả sử hàm số $f(M) = f(x, y)$ xác định trong miền D, $M_0(x_0, y_0)$ và $M_0(x_0 + \Delta x, y_0 + \Delta y)$ là những điểm thuộc D. Đặt $\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$.
(1) Hàm số $f(x, y)$ được gọi là *liên tục tại* $M_0(x_0, y_0)$ nếu $\lim\limits_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \Delta f = 0$.
(2) Hàm số $f(M) = f(x, y)$ gọi là *liên tục trên* D nếu nó liên tục tại mọi điểm thuộc D.

Chú ý 2.2.8.
(1) Hàm số nhiều biến số liên tục cũng có những tính chất như hàm số một biến số liên tục.
(2) Nếu hàm $f(x, y)$ không liên tục tại M_0 thì ta nói nó gián đoạn tại điểm đó. Vậy hàm $f(x, y)$ gián đoạn tại M_0 nếu:

+ Hoặc nó không xác định tại M_0;

+ Hoặc nó xác định tại M_0 nhưng tại điểm đó không tồn tại giới hạn của $f(M)$ khi M dần tới M_0;

+ Hoặc nó tồn tại giới hạn khi M dần tới M_0 nhưng $\lim\limits_{M \to M_0} f(M) \neq f(M_0)$.

Ví dụ 2.2.9. Xét tính liên tục của hàm số

$$f(x, y) = \begin{cases} x. \sin \dfrac{1}{y} & \text{khi} \quad y \neq 0, \\ 0 & \text{khi} \quad y = 0. \end{cases}$$

Giải: Ta thấy hàm số liên tục tại mọi $y \neq 0, x$ bất kỳ.

+ Ta xét tính liên tục của hàm số tại điểm $(0, 0)$. Có

$$0 \leq \left| x. \sin \frac{1}{y} \right| \leq |x| \xrightarrow{x \to 0} 0.$$

Vậy $\lim\limits_{\substack{x \to 0 \\ y \to 0}} x. \sin \dfrac{1}{y} = 0$. Mà $f(0, 0) = 0$ nên hàm số liên tục tại $(0, 0)$.

+ Xét tính liên tục của hàm số tại $(a, 0)$ với $a \neq 0$.

Có $f(a, 0) = 0$. Nhưng $\lim\limits_{\substack{x \to a \\ y \to 0}} x. \sin \dfrac{1}{y} \neq 0$. Vậy hàm số không lên tục tại mọi điểm dạng $(a, 0)$ với $a \neq 0$.

2.2.2. Đạo hàm của hàm số hai biến số

Định nghĩa 2.2.10. Đạo hàm riêng
Cho $z = f(x, y)$ xác định trên một miền mở $D \subset \mathbb{R}^2 . M_0(x_0, y_0) \in D$. Cho x_0 một số gia Δx sao cho $(x_0 + \Delta x, y_0) \in D$. Khi đó số gia của hàm số là

$$\Delta_x f = f(x_0 + \Delta x, y_0) - f(x_0, y_0)$$

gọi là *số gia riêng của hàm số tại* (x_0, y_0) *theo* x.

Tương tự, $\Delta_y f = f(x_0, y_0 + \Delta y) - f(x_0, y_0)$ gọi là *số gia riêng của hàm số tại* (x_0, y_0) *theo* y.

Nếu tồn tại $\lim\limits_{\Delta x \to 0} \dfrac{\Delta_x f}{\Delta x}(x_0, y_0)$ hữu hạn thì ta nói giới hạn đó là *đạo hàm riêng của hàm số theo biến* x *tại* (x_0, y_0). Ký hiệu: z'_x hay $\dfrac{\partial z}{\partial x}$ hay $f'_x(x_0, y_0)$ hay $\dfrac{\partial f(x_0, y_0)}{\partial x}\cdot$

Tương tự, nếu tồn tại $\lim\limits_{\Delta y \to 0} \dfrac{\Delta_y f}{\Delta y}(x_0, y_0)$ hữu hạn thì giới hạn đó gọi là *đạo hàm riêng của hàm số theo biến* y *tại* (x_0, y_0). Ký hiệu: z'_y hay $\dfrac{\partial z}{\partial y}$ hay $f'_y(x_0, y_0)$ hay $\dfrac{\partial f(x_0, y_0)}{\partial y}\cdot$

Chú ý 2.2.11. Khi tính đạo hàm riêng theo một biến thì ta coi biến còn lại là hằng số.

Ví dụ 2.2.12. Tính đạo hàm riêng của các hàm hai biến sau.

$$(a)\ z = x^y \qquad (b)\ z = \arctan \frac{y}{x}\cdot$$

Giải: (a) Ta có: $\dfrac{\partial z}{\partial x} = yx^{y-1}$; $\quad \dfrac{\partial z}{\partial y} = x^y \ln x$.

(b) Ta có:

$$\frac{\partial z}{\partial x} = \frac{\left(\frac{y}{x}\right)'_x}{1 + \left(\frac{y}{x}\right)^2} = \frac{-\frac{y}{x^2}}{\frac{x^2+y^2}{x^2}} = -\frac{y}{x^2+y^2};$$

$$\frac{\partial z}{\partial y} = \frac{\left(\frac{y}{x}\right)'_y}{1 + \left(\frac{y}{x}\right)^2} = \frac{\frac{1}{x}}{\frac{x^2+y^2}{x^2}} = \frac{x}{x^2+y^2}\cdot$$

Định nghĩa 2.2.13. Đạo hàm của hàm số hợp

Cho hàm số $z = f(x, y)$ có các đạo hàm riêng liên tục $z'_x; z'_y$ và $x = \varphi(t); y = \psi(t)$ có các đạo hàm $x'(t); y'(t)$. Khi đó hàm số hợp $f(\varphi(t), \psi(t))$ cũng có các đạo hàm theo t và

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial t}\cdot$$

Định lí 2.2.14. *Nếu* $z = f(x, y)$ *có các đạo hàm riêng liên tục* $z'_x; z'_y$ *và* $x = \varphi(t, v); y = \psi(t, v)$ *có các đạo hàm riêng* $x'_t; x'_v; y'_t; y'_v$ *thì hàm số hợp* $f(\varphi(t, v), \psi(t, v))$ *cũng có các đạo hàm theo* t *và* v.

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial t};$$

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial v}\cdot$$

Ví dụ 2.2.15. Tìm đạo hàm của hàm số hợp sau: $z = u + v^2$ với $\begin{cases} u = x^2 + \sin y, \\ v = \ln(x+y). \end{cases}$

Giải: Ta có: $z'_x = z'_u.u'_x + z'_v.v'_x = 1.2x + 2v \cdot \dfrac{1}{x} + y = 2x + \dfrac{2\ln(x+y)}{x+y},$

$$z'_y = z'_u.u'_y + z'_v.v'_y = 1.\cos y + 2v \cdot \frac{1}{x} + y = \cos y + \frac{2\ln(x+y)}{x+y}\cdot$$

2.2.3. Vi phân toàn phần và ứng dụng để tính gần đúng

Định nghĩa 2.2.16. Vi phân toàn phần

Cho hàm $z = f(x, y)$ xác định trong miền $D.(x_0, y_0) \in D$. Cho x_0 một số gia $\Delta x, y_0$ một số gia Δy sao cho $(x_0 + \Delta x, y_0 + \Delta y) \in D$. Khi đó biểu thức $\Delta f = f(x_0+\Delta x, y_0+\Delta y) - f(x_0, y_0)$ gọi là *số gia toàn phần của hàm số $f(x, y)$ tại $M_0(x_0, y_0)$*. Nếu nó được biểu diễn dưới dạng

$$\Delta f = A.\Delta x + B.\Delta y + O(\rho),$$

trong đó A, B chỉ phụ thuộc vào $(x_0, y_0).O(\rho)$ là vô cùng bé bậc cao hơn ρ với $\rho = \sqrt{\Delta x^2 + \Delta y^2}$ thì ta nói hàm số $u = f(x, y)$ *khả vi* tại (x_0, y_0). Còn biểu thức $A.\Delta x + B.\Delta y$ gọi là *vi phân toàn phần của hàm số tại* (x_0, y_0).

Ký hiệu: $df = A.\Delta x + B.\Delta y$.

Định lí 2.2.17. *Nếu hàm số $z = f(x, y)$ có các đạo hàm riêng ở lân cận của điểm (x_0, y_0) và nếu các đạo hàm riêng ấy liên tục tại M_0 thì $f(x, y)$ khả vi tại M_0 và ta có*

$$dz = f'_x \Delta x + f'_y \Delta y.$$

• Ứng dụng vi phân để tính gần đúng

Từ định nghĩa ta thấy rằng vi phân toàn phần df chỉ khác số gia toàn phần Δf một vô cùng bé bậc cao hơn $\rho = \sqrt{\Delta x^2 + \Delta y^2}$. Do đó khi $\Delta x, \Delta y$ có trị số tuyệt đối khá bé ta có thể xem $df \approx \Delta f$, tức là

$$f(x_0 + \Delta x, y_0 + \Delta y) \approx f(x_0, y_0) + f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y.$$

Công thức này gọi là công thức tính gần đúng.

Ví dụ 2.2.18. Tính gần đúng giá trị $\arctan \dfrac{1,02}{0,95}\cdot$

Giải: Xét hàm số $u = \arctan \dfrac{y}{x}$. Ta cần tính $u = f(x_0 + \Delta x, y_0 + \Delta y)$ với $x_0 = 1; \Delta x = -0,05; y_0 = 1; \Delta y = 0,02$.

Ta có: $f'_x = \dfrac{-y}{x^2+y^2}; f'_y = \dfrac{x}{x^2+y^2}\cdot$ Vậy theo công thức tính gần đúng ta có

$$f(1 - 0,05; 1 + 0,02) \approx f(1;1) + \frac{1.0,02 + 1.0,05}{2} = \frac{\pi}{4} + 0,035 = 0,82(rad).$$

2.2.4. Đạo hàm và vi phân cấp cao

Định nghĩa 2.2.19. Đạo hàm cấp cao
Cho hàm số $u = f(x, y)$ xác định và khả vi trong miền D. Các đạo hàm riêng f'_x, f'_y là những đạo hàm riêng cấp một. Các đạo hàm riêng của các đạo hàm riêng cấp một (nếu tồn tại) gọi là những *đạo hàm riêng cấp hai.*

Tùy theo thứ tự lấy đạo hàm đối với các biến số x và y ta có:

$$f''_{x^2}(x, y) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \tag{2.2}$$

$$f''_{xy}(x, y) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \tag{2.3}$$

$$f''_{yx}(x, y) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \tag{2.4}$$

$$f''_{y^2}(x, y) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \tag{2.5}$$

Chú ý 2.2.20.
(1) Công thức (2.2) và (2.5) được gọi là công thức đạo hàm vuông, công thức (2.3) và (2.4) được gọi là công thức đạo hàm chữ nhật.
(2) Các đạo hàm riêng cấp cao hơn cũng được định nghĩa tương tự.

Ví dụ 2.2.21. Tính các đạo hàm riêng cấp cao của hàm số $z = f(x, y) = x^2y^3 + x^4$.
Giải: Ta lần lượt tính các đạo hàm riêng cấp 1 và cấp 2 như sau.

$$\begin{aligned} z'_x &= 2xy^3 + 4x^3; & z'_y &= 3x^2y^2. \\ z''_{x^2} &= 2y^3 + 12x^2; & z''_{y^2} &= 6x^2y. \\ z''_{xy} &= 6xy^2; & z''_{yx} &= 6xy^2. \end{aligned}$$

Định lí 2.2.22. ***Định lý Schwarz***
Nếu trong một lân cận D nào đó của $M_0(x_0, y_0)$, hàm số $z = f(x, y)$ có các đạo hàm riêng $f''_{yx}; f''_{xy}$ và nếu các đạo hàm ấy liên tục tại M_0 thì $f''_{yx} = f''_{xy}$ tại M_0.

Định nghĩa 2.2.23. Vi phân cấp cao
Giả sử hàm số $z = f(x, y)$ xác định và khả vi trong miền D. Khi đó vi phân toàn phần của nó $dz = \dfrac{\partial z}{\partial x}dx + \dfrac{\partial z}{\partial y}dy$ nếu tồn tại cũng là một hàm số của x, y. Vi phân toàn phần của dz nếu tồn tại được gọi là *vi phân toàn phần cấp 2 của hàm z*. Ký hiệu

$$d^2z = d(dz) = d\left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right) \cdot$$

Tiếp tục như vậy người ta định nghĩa vi phân cấp cao hơn:

$$\begin{aligned} &d^3z = d(d^2z); \\ &\cdots\cdots\cdots \\ &d^nz = d(d^{n-1}z). \end{aligned}$$

• Biểu thức vi phân toàn phần cấp 2
Giả sử x, y là những biến số độc lập, khi đó $dx = \Delta x, dy = \Delta y$ đó là những hằng số không phụ thuộc x, y. Giả sử d^2z tồn tại. Ta có:

$$\begin{aligned} d^2z &= \left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right)'_x dx + \left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right)'_y dy \\ &= \frac{\partial^2 z}{\partial x^2}dx^2 + \left(\frac{\partial^2 z}{\partial x \partial y} + \frac{\partial^2 z}{\partial y \partial x}\right)dxdy + \frac{\partial^2 z}{\partial y^2}dy^2. \end{aligned}$$

Giả thiết thêm rằng $f''_{yx}; f''_{xy}$ liên tục thì $f''_{yx} = f''_{xy}$. Khi đó

$$d^2z = f''_{x^2}dx^2 + 2f''_{xy}dxdy + f''_{y^2}dy^2. \qquad (*)$$

Chú ý 2.2.24. Nếu x, y không phải là những biến số độc lập thì công thức (*) không còn đúng nữa. Giả sử x, y là các hàm của các biến độc lập s, t. Khi ấy dx, dy không phải là những hằng số nữa, mà phụ thuộc vào s, t. Do đó:

$$\begin{aligned} d^2z = d(dz) &= d(f'_x dx + f'_y dy) \\ &= d(f'_x)dx + f'_x d(dx) + d(f'_y)dy + f'_y d(dy) \\ &= f''_{x^2}dx^2 + 2f''_{xy}dxdy + f''_{y^2}dy^2 + f'_x d^2x + f'_y d^2y. \end{aligned}$$

Vậy vi phân toàn phần cấp cao của hàm nhiều biến số không có dạng bất biến.

Ví dụ 2.2.25. Cho hàm $z = \arctan\dfrac{x}{y}$. Tính $d^2z(0, 1)$.

Giải: Trước tiên ta cần tính các đạo hàm riêng cấp hai của hàm số đã cho. Cụ thể:

$$\begin{aligned} &z'_x = \frac{y}{x^2+y^2}; \quad z'_y = -\frac{x}{x^2+y^2}; \quad z''_{x^2} = \frac{-2xy}{(x^2+y^2)^2}; \quad z''_{y^2} = \frac{2xy}{(x^2+y^2)^2}; \\ &z''_{xy} = \frac{1}{x^2+y^2} - \frac{2y^2}{(x^2+y^2)^2} = \frac{x^2-y^2}{(x^2+y^2)^2} = z''_{yx}. \end{aligned}$$

Suy ra, $d^2z = \dfrac{-2xy}{(x^2+y^2)^2}dx^2 + \dfrac{2xy}{(x^2+y^2)^2}dy^2 + 2\dfrac{x^2-y^2}{(x^2+y^2)^2}dxdy.$
Vậy $d^2z(0, 1) = -2dxdy.$

BÀI TẬP CHƯƠNG 2

Bài 1: Tính đạo hàm cấp 1 của các hàm số sau đây.

$$\begin{aligned} &1)\ y = (x^7 - 5x^2)^3; \quad 2)\ y = (x^2+1)(5-3x^2); \quad 3)\ y = \frac{2x}{x^2-1}; \\ &4)\ y = \frac{1+x}{\sqrt{1-x}}; \qquad 5)\ y = \sqrt{2-5x-x^2}; \qquad 6)\ y = x.\cot x. \end{aligned}$$

Bài 2: Tính đạo hàm cấp n của các hàm số cho dưới đây.

$$\begin{aligned} &1)\ y = (3x^2+2x-1)e^{3x}; \quad 2)\ y = (2x^3-x)e^{-2x}; \quad 3)\ y = (2x^2-4x)\sin 2x; \\ &4)\ y = (x^2+1)e^{2x}; \qquad 5)\ y = (2x^2-1)e^{-2x}; \quad 6)\ y = (x^2-3x)e^{3x}. \end{aligned}$$

Bài 3: Chi phí sản xuất x đơn vị sản phẩm của một loại mặt hàng được cho bởi hàm số $C(x) = 2500 + 12q(\$)$.

a) Tìm chi phí trung bình sản xuất 100 sản phẩm? 1000 sản phẩm?

b) Tìm chi phí biên để sản xuất 100 sản phẩm? 1000 sản phẩm?

Bài 4: Một người nông dân cần quây 3 chuồng nuôi bò liền nhau có cùng diện tích là $15m^2$ bằng dây thép gai. Hỏi người nông dân nên quây chuồng có kích thước như thế nào để vừa đủ yêu cầu về diện tích mỗi chuồng mà tốn ít dây thép nhất?

Bài 5: Một người chăn nuôi bò sữa có $200m$ rào để quây hai chuồng bò bằng nhau hình chữ nhật. Hỏi người đó nên quây chuồng có kích thước như thế nào để diện tích mỗi chuồng là lớn nhất?

Bài 6: Tổng doanh thu (\$) khi sản xuất x đơn vị sản phẩm của một nhà máy được cho bởi hàm số sau

$$R = -x^3 + 450x^2 + 52500x \quad (x > 0).$$

Hỏi nhà máy nên đưa ra mức sản suất là bao nhiêu sản phẩm để có được doanh thu lớn nhất?

Bài 7: Tổng chi phí (\$) khi sản xuất x đơn vị sản phẩm của một nhà máy được cho bởi hàm số sau

$$C(x) = x^3 - 12x^2 + 60x \quad (x > 0).$$

Hỏi nhà máy nên đưa ra mức sản suất là bao nhiêu sản phẩm để chi phí là nhỏ nhất?

Bài 8: Sự lây lan của virut có thể được mô hình bởi

$$N = -t^3 + 12t^2, \quad 0 \leq t \leq 12.$$

Với N số lượng người bị nhiễm (tính bằng hàng trăm người), t là thời gian tính bằng tuần.

a) Theo anh (chị) dự đoán tối đa có bao nhiêu người bị nhiễm virut trên?

b) Virut sẽ lây lan nhanh nhất vào thời điểm nào?

Bài 9: Khi rác thải đổ xuống ao, sự phân hủy của rác thải tiêu hao oxy. Mức oxy có trong ao khi rác thải bị oxy hóa được mô hình bởi

$$O = \frac{t^2 - t + 1}{t^2 + 1}; \quad t \geq 0.$$

Với t là thời gian tính bằng tuần. Khi nào mức oxy là thấp nhất? Mức đó là bao nhiêu?

Bài 10: Tác dụng (E) của một loại thuốc giảm đau sau khi vào dòng máu t giờ được cho bởi

$$E = \frac{1}{27}(9t + 3t^2 - t^3), \quad 0 \leq t \leq 4,5.$$

Tìm tỷ lệ tác dụng trung bình của E trong khoảng thời gian từ 1 đến $2h$ và tỷ lệ tác dụng tức thời tại thời điểm $t = 2$ giờ.

Bài 11: Sự phát triển của một loài vi khuẩn được mô hình bởi hàm số

$$P = 500\left(1 + \frac{4t}{50 + t^2}\right),$$

với t là thời gian tính bằng (h). Tìm tỷ lệ tăng trưởng của số lượng vi khuẩn tại thời điểm $t = 2$.

Bài 12: Một công ty vừa ước lượng rằng chi phí (tính bằng \$) cho x đơn vị sản phẩm được sản xuất $(x > 0)$ được mô hình bởi hàm số $C = 800 + 0,04x + 0,0002x^2$. Hỏi công ty nên đưa ra mức sản xuất là bao nhiêu sản phẩm để mức chi phí trung bình cho một sản phẩm là nhỏ nhất?

Bài 13: Lợi nhuận thu được từ việc bán x cái đồng hồ báo thức được mô hình bởi hàm số

$$P = 0,0002x^3 + 10x \quad (\$).$$

a) Tìm lợi nhuận biên cho mức sản xuất 50 chiếc.

b) Lợi nhuận thực tế tăng lên bao nhiêu khi tăng mức sản xuất từ 50 đến 51 chiếc. So sánh con số đó với lợi nhuận biên ở trên rồi rút ra kết luận.

Bài 14: Tính các đạo hàm riêng cấp một của các hàm hai biến sau.

1) $f(x,y) = \sin(2x^2y^3) + e^{xy}$; 2) $f(x,y) = \cos(x^2y^4) + e^{x^2}$;
3) $f(x,y) = e^{x^2y} - 3x^4y^5$; 4) $f(x,y) = e^{x^3y^3} - 5x^4y^2$;
5) $f(x,y) = e^{x^3y^3} - \cos(5x^2)$; 6) $f(x,y) = x.e^{x^2y}$.

Bài 15: Tìm vi phân toàn phần của các hàm số hai biến số dưới đây.

1) $f(x,y) = \cos(xye^{xy})$; 2) $f(x,y) = \sin(xye^{xy})$; 3) $f(x,y) = \ln(x\sqrt{x+y})$.

Chương 3

Tích phân và một số ứng dụng

Từ Chương 2, ta biết rằng nếu hàm số $f(x)$ khả vi trong khoảng (a, b) thì có đạo hàm trong khoảng (a, b) và ta hoàn toàn tính được đạo hàm của hàm số trong khoảng đó. Một bài toán đặt ra là nếu cho trước một hàm số $f(x)$ xác định trong khoảng (a, b) thì liệu có tồn tại một hàm số $F(x)$ khả vi trong khoảng (a, b) và $F'(x) = f(x)$ với mọi x thuộc khoảng (a, b), và nếu hàm $F(x)$ như vậy tồn tại thì ta sẽ tìm hàm đó như thế nào? Chương này nhằm trả lời câu hỏi đó. Bên cạnh đó, chúng ta sẽ tìm hiểu một số ứng dụng quan trọng của bài toán trên trong nhiều lĩnh vực trong thực tế như kinh tế, nông nghiệp và một số ngành khoa học khác.

3.1. Tích phân bất định

3.1.1. Nguyên hàm của hàm số

Định nghĩa 3.1.1. Nguyên hàm

Hàm $F(x)$ được gọi là *nguyên hàm* của hàm $f(x)$ nếu tại mọi điểm x thuộc miền xác định của f ta đều có $F'(x) = f(x)$.

Nếu $F(x)$ là một nguyên hàm của $f(x)$ thì $F(x) + C$, với C là hằng số cũng là một nguyên hàm của $f(x)$.

Ví dụ 3.1.2. Ta có các hàm số $F(x) = x^3$; $\quad G(x) = x^3 + 2$; $\quad H(x) = x^3 + 0,1$ đều là các nguyên hàm của hàm số $f(x) = 3x^2$ vì đạo hàm của chúng đều bằng $3x^2$.

Định lí 3.1.3. *Giả sử $F(x)$ có đạo hàm trong (a, b) và $F(x)$ là nguyên hàm của $f(x)$ với mọi x. Khi đó, ta có các khẳng định sau:*

(1) Với mọi hằng số $C, F(x) + C$ cũng là một nguyên hàm của $f(x), \forall x \in (a, b)$.

(2) Ngược lại, mọi nguyên hàm của $f(x), \forall x \in (a, b)$ đều có dạng $F(x) + C$.

3.1.2. Tích phân bất định

Định nghĩa 3.1.4. Nếu $f(x)$ có một nguyên hàm là $F(x)$ thì nó có một họ các nguyên hàm là $F(x) + C$ với C là một hằng số tùy ý và họ nguyên hàm đó được gọi là *tích phân bất định của hàm $f(x)$*.

Ký hiệu: $\int f(x)dx$, trong đó: $\int$ là dấu tích phân; x là biến lấy tích phân; $f(x)$ là hàm dưới dấu tích phân; $f(x)dx$ là biểu thức dưới dấu tích phân.
Như vậy theo định nghĩa $\int f(x)dx = F(x) + C$.

Ví dụ 3.1.5. Tính các tích phân sau: (a) $\int x^5 dx$; (b) $\int \sin x dx$.

Giải: (a) Ta có $\int x^5 dx = \dfrac{x^6}{6} + C$ vì $\left(\dfrac{x^6}{6} + C\right)' = x^5$.

(b) $\int \sin x dx = -\cos x + C$ vì $(\cos x)' = \sin x$.

• **Các tính chất của tích phân bất định**

(1) $\left[\int f(x)dx\right]' = f(x)$; $\quad d\int f(x)dx = f(x)dx$;

(2) $dF(x) = F(x) + C$ hay $\int \dfrac{dF(x)}{dx}dx = \int F'(x)dx = \int dF(x) = F(x) + C$;

(3) $\int kf(x)dx = k\int f(x)dx$ với k là một hằng số;

(4) $\int (f(x) + g(x))dx = \int f(x)dx + \int g(x)dx$;

(5) Nếu có $\int f(x)dx = F(x) + C$ và $u = \varphi(x)$ thì $\int f(u)du = F(u) + C$.

• **Bảng các tích phân bất định của một số hàm cơ bản**

1) $\int 0dx = C$;	9) $\int \dfrac{dx}{1+x^2} = \arctan x + C$;
2) $\int x^\alpha dx = \dfrac{x^{\alpha+1}}{\alpha+1} + C \quad (\alpha \neq -1)$;	10) $\int \dfrac{dx}{\sqrt{1-x^2}} = \arcsin x + C$;
3) $\int \dfrac{1}{x}dx = \ln\|x\| + C$;	11) $\int \dfrac{dx}{a^2-x^2} = \dfrac{1}{2a}\ln\left\|\dfrac{a+x}{a-x}\right\| + C$;
4) $\int a^x dx = \dfrac{a^x}{\ln a} + C$;	12) $\int \dfrac{f'(x)}{\sqrt{f(x)}}dx = 2\sqrt{f(x)} + C$;
5) $\int \cos x dx = \sin x + C$;	13) $\int \dfrac{dx}{\sqrt{x^2+a}} = \ln(x + \sqrt{x^2+a}) + C$;
6) $\int \sin x dx = -\cos x + C$;	14) $\int \dfrac{dx}{a^2+x^2} = \dfrac{1}{a}\arctan\dfrac{x}{a} + C$;
7) $\int \dfrac{dx}{\cos^2 x} = \tan x + C$;	15) $\int \dfrac{dx}{\sqrt{a^2-x^2}} = \arcsin\dfrac{x}{a} + C$.
8) $\int \dfrac{dx}{\sin^2 x} = -\cot x + C$;	

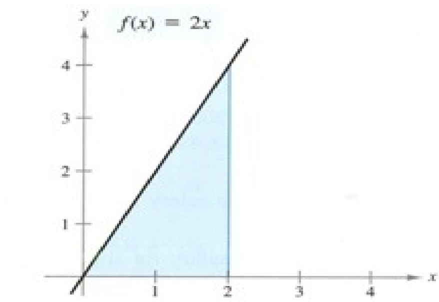
Chú ý 3.1.6. Từ định nghĩa nguyên hàm ta thấy phép tính đạo hàm và phép tính nguyên hàm là hai phép tính ngược nhau nên từ bảng đạo hàm của một số hàm cơ bản ta suy ra được bảng nguyên hàm của một số hàm cơ bản, chẳng hạn các tích phân từ 1 đến 10 trong bảng. Các tích phân từ 11 đến 15 ta có thể xây dựng dựa vào tính chất của tích phân bất định và các tích phân đã biết.

3.2. Tích phân xác định

Từ các kiến thức đã được học ở hình học phẳng, chúng ta biết diện tích là một số đo bằng số của miền được giới hạn. Với các hình đơn giản, như miền hình tam giác, miền hình chữ nhật, miền hình thang,... tức là các miền giới hạn bởi các đoạn thẳng, hay miền hình tròn, chúng ta có thể tính diện tích bằng các công thức tính trong hình học.

Ta cùng tìm hiểu một số ví dụ sau đây:

Ví dụ 3.2.1. (a) Tính diện tích S của miền tam giác giới hạn bởi các đường $f(x) = 2x$, $x = 0$ và $x = 2$.
(b) So sánh diện tích S và hiệu $F(2) - F(0)$ với $F(x)$ là nguyên hàm bất kỳ của $f(x)$ trên đoạn $[0, 2]$.



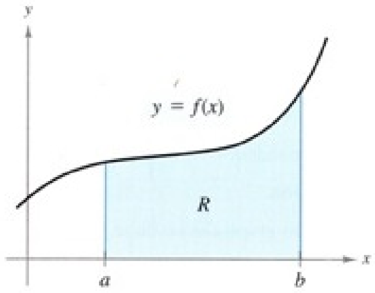
Hình 3.1:

Giải: (a) Nhìn hình vẽ 3.1 có $S = \dfrac{1}{2}ah = \dfrac{1}{2} \cdot 4.2 = 4$.
(b) Ta có $F(x) = x^2 \Rightarrow F(2) - F(0) = 2^2 - 0^2 = 4$. Vậy $S = F(2) - F(0)$.

Trong chương này, chúng ta được học thêm một phương pháp tìm diện tích của các hình trên, đó là dùng *tích phân xác định*. Bằng tích phân xác định, ta còn có thể tìm diện tích của các miền giới hạn bởi các đường cong kín bất kỳ, hay các miền không phải chỉ giới hạn bởi các đoạn thẳng như miền R được chỉ ra ở hình vẽ 3.2.

3.2.1. Diện tích của hình thang cong và tích phân xác định

Định nghĩa 3.2.2. Cho hàm số $f(x)$ liên tục và lấy giá trị dương trên đoạn $[a, b]$. Hình phẳng giới hạn bởi đường cong $y = f(x)$, trục hoành và hai đường thẳng $x = a, x = b$ được gọi là *hình thang cong* (Hình 3.2).



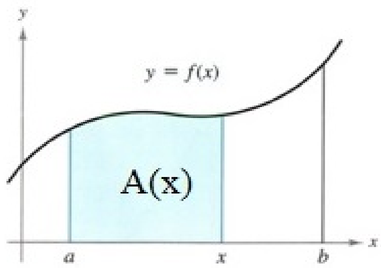
Hình 3.2:

Bài toán: Cho $f(x)$ là hàm số liên tục, đồng biến và nhận giá trị dương trên đoạn $[a,b]$. Chứng minh rằng diện tích A của hình thang cong giới hạn bởi đồ thị của hàm số f, trục hoành và hai đường thẳng $x=a, x=b$ được xác định bởi

$$A = F(b) - F(a).$$

Trong đó $F(x)$ là một nguyên hàm bất kỳ của $f(x)$ trên đoạn $[a,b]$.
Chứng minh.

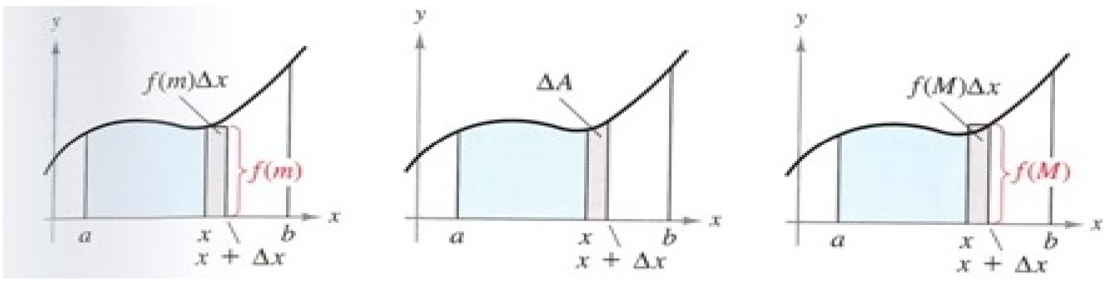
Giả sử $A(x),\ a \leq x \leq b$ (Hình 3.3) là hàm số biểu thị diện tích của miền giới hạn bởi đồ thị hàm số $y=f(x)$, trục hoành, đường thẳng $x=a$ và đường thẳng đi qua hoành độ x và vuông góc với trục Ox tại điểm x. Để tìm hiểu mối quan hệ giữa $A(x)$ và $f(x)$, cho x một số gia Δx, diện tích của phần ứng với số gia Δx ký hiệu là ΔA. Gọi $f(m)$ và $f(M)$ tương ứng là giá trị nhỏ nhất và giá trị lớn nhất của hàm f trên đoạn $[x, x+\Delta x]$.



Hình 3.3:

Nhìn hình vẽ 3.4, ta nhận thấy:

$$\begin{aligned} f(m).\Delta x \leq \Delta A \leq f(M).\Delta x &\Leftrightarrow f(m) \leq \frac{\Delta A}{\Delta x} \leq f(M) \\ &\Leftrightarrow \lim_{\Delta x \to 0} f(m) \leq \lim_{\Delta x \to 0} \frac{\Delta A}{\Delta x} \leq \lim_{\Delta x \to 0} f(M) \\ &\Leftrightarrow f(x) \leq A'(x) \leq f(x). \end{aligned}$$



Hình 3.4:

Do đó $A'(x) = f(x)$ vì $F'(x) = f(x) \Rightarrow A(x) = F(x) + C$. Vì $A(a) = 0$ nên $C = -F(a)$. Vì thế $A(x) = F(x) - F(a) \Rightarrow A(b) = F(b) - F(a)$.

Vậy diện tích của hình thang cong giới hạn bởi đồ thị của hàm số f, trục hoành và hai đường thẳng $x = a, x = b$ được xác định bởi

$$A = F(b) - F(a).$$

Trong đó $F(x)$ là một nguyên hàm bất kỳ của $f(x)$ trên đoạn $[a, b]$.

Định nghĩa 3.2.3. Cho hàm số f liên tục trên K và a, b là hai số bất kỳ thuộc K. Nếu F là một nguyên hàm của f trên K thì hiệu số $F(b) - F(a)$ được gọi là *tích phân xác định của f từ a đến b* và ký hiệu là $\int\limits_a^b f(x)dx$ hoặc $\int\limits_a^b f(x)dx = F(x)\Big|_a^b = F(b) - F(a)$,

trong đó a, b được gọi là các *cận lấy tích phân*, a là *cận dưới*, b là *cận trên*, f là *hàm số dưới dấu tích phân*, $f(x)dx$ là *biểu thức dưới dấu tích phân*, x là *biến số lấy tích phân*.

Chú ý 3.2.4. (1) $\int\limits_a^b f(x)dx$ là một số không phụ thuộc vào việc chọn nguyên hàm $F(x)$ trong họ các nguyên hàm của f. Thật vậy,

$$\int\limits_a^b f(x)dx = (F(x) + C)\Big|_a^b = (F(b) + C) - (F(a) + C) = F(b) - F(a).$$

(2) Tích phân xác định không phụ thuộc vào biến lấy tích phân mà chỉ phụ thuộc vào hàm số dưới dấu tích phân và cận lấy tích phân, tức là

$$\int\limits_a^b f(x)dx = \int\limits_a^b f(t)dt.$$

(3) Với định nghĩa tích phân xác định, diện tích A của hình thang cong giới hạn bởi đồ thị của hàm số $y = f(x)$ (hàm số liên tục, không âm trên đoạn $[a, b]$), trục hoành và hai đường thẳng $x = a, x = b$ là: $A = \int\limits_a^b f(x)dx$.

Quay trở lại Ví dụ 3.2.1 ta thấy diện tích của miền tam giác đó là

$$S = \int\limits_0^2 2xdx = x^2\Big|_0^2 = 2^2 - 0^2 = 4.$$

Ví dụ 3.2.5. Tính diện tích hình thang cong giới hạn bởi đồ thị hàm số $y = x^2 - 1$; trục hoành và hai đường thẳng $x = 1; x = 2$.

Giải: Ta có $S = \int\limits_1^2 (x^2 - 1)dx = \left(\dfrac{x^3}{3} - x\right)\Bigg|_1^2 = \left(\dfrac{2^3}{3} - 2\right) - \left(\dfrac{1}{3} - 1\right) = \dfrac{4}{3}\cdot$

• **Tính chất của tích phân xác định**

(1) $\int\limits_a^b kf(x)dx = k\int\limits_a^b f(x)dx, \quad k \in \mathbb{R};$

(2) $\int\limits_a^b (f(x) + g(x))dx = \int\limits_a^b f(x)dx + \int\limits_a^b g(x)dx;$

(3) Với 3 số $a < b < c$ ta có: $\int\limits_a^b f(x)dx = \int\limits_a^c f(x)dx + \int\limits_c^b f(x)dx;$

(4) $\int\limits_a^a f(x)dx = 0;$

(5) $\int\limits_a^b f(x)dx = -\int\limits_b^a f(x)dx.$

3.2.2. Các phương pháp tính tích phân xác định

(a) **Phương pháp tính trực tiếp:** Sử dụng các biến đổi đại số để đưa tích phân cần tính về những dạng tích phân đã có nguyên hàm cơ bản.

Ví dụ 3.2.6. Tính tích phân $\int\limits_0^2 |2x - 1|dx.$

Giải: Ta có $|2x - 1| = \begin{cases} 2x - 1, & x \geq \dfrac{1}{2}; \\ -(2x - 1), & x < \dfrac{1}{2}\cdot \end{cases}$

Nên $\int\limits_0^2 |2x - 1|dx = \int\limits_0^{\frac{1}{2}} -(2x - 1)dx + \int\limits_{\frac{1}{2}}^2 (2x - 1)dx = \left(-x^2 + x\right)\Big|_0^{\frac{1}{2}} + \left(x^2 - x\right)\Big|_{\frac{1}{2}}^2 = \dfrac{5}{2}\cdot$

Ví dụ 3.2.7. Tính tích phân $I = \int\limits_1^4 -3\sqrt{x}dx.$

Giải: $I = -3\int\limits_1^4 x^{\frac{1}{2}}dx = -3\cdot\frac{2}{3}\cdot x^{\frac{3}{2}}\Big|_1^4 = -14.$

(b) **Phương pháp đổi biến số**

Cách 1: Đổi biến $x = \varphi(t)$. Xét $I = \int\limits_a^b f(x)dx$ với $f(x)$ liên tục trên $[a, b]$.

Định lí 3.2.8. *Giả sử phép đổi biến $x = \varphi(t)$ thỏa mãn các điều kiện:*

(1) $\varphi(t)$ có đạo hàm liên tục trên $[\alpha, \beta]$;

(2) $\varphi(\alpha) = a$ và $\varphi(\beta) = b$;

(3) Khi t biến thiên trong $[\alpha, \beta]$ thì x biến thiên trong $[a, b]$.

Khi đó: $$\int\limits_a^b f(x)dx = \int\limits_\alpha^\beta f[\varphi(t)]\varphi'(t)dt.$$

Ví dụ 3.2.9. Tính tích phân $I = \int\limits_0^a x^2\sqrt{a^2 - x^2}dx \quad (a > 0).$

Giải: Đặt $x = a\sin t \Rightarrow dx = a\cos t dt$; khi $x = 0 \Rightarrow t = 0; x = a \Rightarrow t = \frac{\pi}{2}\cdot$

Vậy $I = \int\limits_0^{\frac{\pi}{2}} (a\sin t)^2(a\cos t)^2 dt = \frac{a^4}{4}\int\limits_0^{\frac{\pi}{2}} \sin^2 2t dt = \frac{a^4}{8}\left(t - \frac{1}{4}\sin 4t\right)\Big|_0^{\frac{\pi}{2}} = \frac{a^4\pi}{16}\cdot$

Ví dụ 3.2.10. Tính tích phân $I = \int\limits_0^2 \frac{xdx}{1 + 4x^2}\cdot$

Giải: Đặt $u = 1 + 4x^2 \Rightarrow du = 8xdx.$

Đổi cận lấy tích phân, ta được: $x = 0 \Rightarrow u = 1; x = 2 \Rightarrow u = 17.$

Vậy $I = \int\limits_1^{17} \frac{du}{8u} = \frac{1}{8}\ln u\Big|_1^{17} = \frac{1}{8}\ln 17.$

Cách 2: Đổi biến $t = \varphi(x)$.

Định lí 3.2.11. *Nếu phép đổi biến $t = \varphi(x)$ thỏa mãn các điều kiện:*

(1) $\varphi(x)$ biến thiên đơn điệu trên $[a, b]$ và có đạo hàm liên tục;

(2) $f(x)dx$ trở thành $g(t)dt$ với $g(t)$ liên tục trên $[\varphi(a), \varphi(b)]$.

Khi đó, $$\int\limits_a^b f(x)dx = \int\limits_{\varphi(a)}^{\varphi(b)} g(t)dt.$$

Chú ý 3.2.12. (1) Khi dùng công thức trên cần lưu ý tính đơn điệu của $\varphi(x)$ trên $[a, b]$ nếu không rất có thể xảy ra $\varphi(a) = \varphi(b)$ mặc dù $a \neq b$.
(2) Khi tính tích phân xác định bằng phương pháp đổi biến ta phải đổi cận lấy tích phân theo biến mới.

Ví dụ 3.2.13. Tính tích phân $I = \int\limits_0^{\frac{\pi}{2}} \frac{\cos x}{1+\sin^2 x} dx$.

Giải: Đặt $t = \sin x$ khi đó $\sin x$ biến thiên đơn điệu trên $\left[0, \frac{\pi}{2}\right]$; $dt = \cos x dx$.
$x = 0 \Rightarrow t = 0; x = \frac{\pi}{2} \Rightarrow t = 1$.

Vậy $I = \int\limits_0^1 \frac{dt}{1+t^2} = \arctan t\Big|_0^1 = \frac{\pi}{4}$.

(c) **Phương pháp tích phân từng phần:** Giả sử $u = u(x); v = v(x)$ đều có đạo hàm liên tục trên $[a, b]$. Khi đó

$$\int\limits_a^b u dv = uv\Big|_a^b - \int\limits_a^b v du.$$

Ví dụ 3.2.14. Tính tích phân $I = \int\limits_0^1 \ln(1+x) dx$.

Giải: Đặt $u = \ln(1+x) \Rightarrow du = \frac{dx}{1+x}; dv = dx \Rightarrow v = x$.

$$I = x\ln(1+x)\Big|_0^1 - \int\limits_0^1 \frac{xdx}{1+x} = \ln 2 - \int\limits_0^1 \left(1 - \frac{1}{1+x}\right) dx$$
$$= \ln 2 - (x - \ln(1+x))\Big|_0^1 = -1 + 2\ln 2.$$

Ví dụ 3.2.15. Tính tích phân $I = \int\limits_0^1 xe^x dx$.

Giải: Đặt $u = x \Rightarrow du = dx; dv = e^x dx \Rightarrow v = e^x$.
Vậy $I = xe^x\Big|_0^1 - \int\limits_0^1 e^x dx = e - e^x\Big|_0^1 = e - e + 1 = 1$.

3.2.3. Một số ứng dụng của tích phân xác định

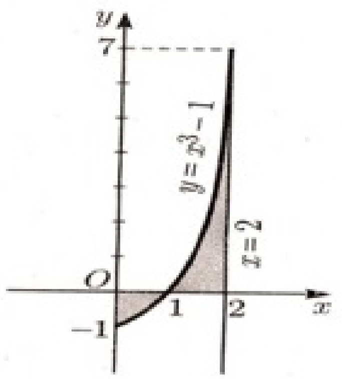
(a) **Ứng dụng tích phân xác định trong hình học**

Ta nhắc lại các công thức tính diện tích hình phẳng sau đây đã được học rất kỹ ở sách giáo khoa Giải tích 12.

Trường hợp 1: Nếu hàm số $y = f(x)$ liên tục trên đoạn $[a, b]$ thì diện tích S của hình phẳng giới hạn bởi đồ thị hàm số $y = f(x)$, trục hoành và hai đường thẳng $x = a, x = b$ là

$$S = \int\limits_a^b |f(x)| dx. \tag{3.1}$$

Ví dụ 3.2.16. Tính diện tích S của hình phẳng giới hạn bởi đồ thị hàm số $y = x^3 - 1$, đường thẳng $x = 2$ trục tung và trục hoành.



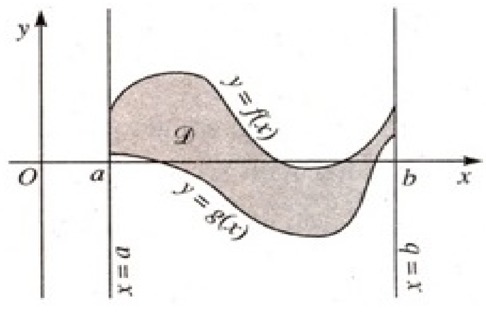
Hình 3.5:

Giải: Nhìn hình vẽ 3.5 ta thấy, $f(x) = x^3 - 1 \leq 0, \forall x \in [0, 1]$; $\quad f(x) \geq 0, \forall x \in [1, 2]$.
Ta có

$$S = \int_0^2 |x^3 - 1| dx = \int_0^1 (1 - x^3) dx + \int_1^2 (x^3 - 1) dx = \frac{7}{2}.$$

Trường hợp 2: Để tính diện tích S của hình phẳng giới hạn bởi đồ thị các hàm số $y = f(x); y = g(x)$ liên tục trên đoạn $[a, b]$ và hai đường thẳng $x = a, x = b$ ta có công thức

$$S = \int_a^b |f(x) - g(x)| dx. \tag{3.2}$$



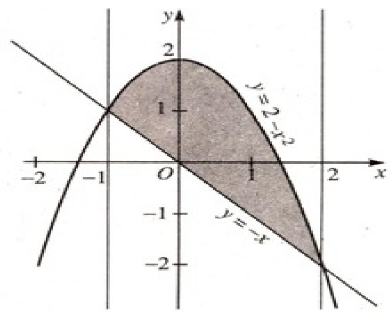
Hình 3.6:

Chú ý 3.2.17. (1) Tương tự bằng cách coi x là hàm số của biến y, diện tích S của hình phẳng giới hạn bởi các đường cong $x = g(y), x = h(y)$ (g, h là các hàm liên tục trên đoạn $[c, d]$) và hai đường thẳng $y = c; y = d$ là

$$S = \int_c^d |g(y) - h(y)| dy. \tag{3.3}$$

(2) Để tính diện tích của các hình phức tạp hơn ta phải chia thành các hình đơn giản đã biết cách tính diện tích.

Ví dụ 3.2.18. Tính diện tích S của hình phẳng giới hạn bởi đồ thị hàm số $y = 2 - x^2; y = -x$.



Hình 3.7:

Giải: Phương trình hoành độ giao điểm của hai đồ thị hàm số là:

$$2 - x^2 = -x \Leftrightarrow x = -1 \text{ hoặc } x = 2.$$

Ta có $S = \int\limits_{-1}^{2} (2 + x - x^2)dx = \frac{9}{2}\cdot$

(b) **Ứng dụng tích phân xác định trong vật lý**

Bài toán về quãng đường đi được của một vật: Giả sử một vật chuyển động có vận tốc thay đổi theo thời gian $t, v = f(t), (0 < t < T)$. Chứng minh rằng quãng đường vật di chuyển được trong khoảng thời gian từ thời điểm $t = a$ đến thời điểm $t = b(0 < a < b < T)$ là

$$S = F(b) - F(a).$$

Trong đó F là một nguyên hàm bất kỳ của f trên khoảng $(0, T)$.

Gọi $S(t)$ là hàm số chỉ quãng đường vật di chuyển được cho tới thời điểm t. Quãng đường vật di chuyển được trong khoảng thời gian từ thời điểm $t = a$ cho đến thời điểm $t = b$ là $S = S(b) - S(a)$. Mặt khác, ta biết rằng $S'(t) = v(t) = f(t)$, do đó $S(t)$ là một nguyên hàm của f và $S(t) = F(t) + C$. Vậy

$$S = S(b) - S(a) = [F(b) + C] - [F(a) + C] = F(b) - F(a) = \int\limits_a^b f(x)dx.$$

Ví dụ 3.2.19. Một vật đang chuyển động với vận tốc $10m/s$ thì tăng tốc với gia tốc $a(t) = 3t + t^2(m/s^2)$. Tính quãng đường vật đi được trong khoảng thời gian $10s$ kể từ lúc bắt đầu tăng tốc.

Giải: Vận tốc tăng lên của vật là $v(t) = a'(t) = 3 + 2t$, vận tốc thực tế của vật sau khi tăng tốc là $v(t) = 3 + 2t + 10 = 2t + 13$. Vậy quãng đường mà vật đi được trong khoảng thời gian $10s$ kể từ lúc bắt đầu tăng tốc là: $S = \int\limits_0^{10} (2t + 13)dt = (t^2 + 13t)\Big|_0^{10} = 230m.$

(c) **Ứng dụng tích phân xác định trong đời sống - kinh tế**

Các tình huống trong thực tế chẳng hạn như sự tăng lên, sự phân hủy của vật chất, sự tăng trưởng, sự sụt giảm về dân số thì tốc độ tăng trưởng, tốc độ sụt giảm của chúng theo thời gian t tỷ lệ với số lượng vật chất, hay dân số ở thời điểm đó. Trong mô hình đơn giản nhất, mối quan hệ này được biểu diễn bởi phương trình $\dfrac{dy}{dt} = ky \Leftrightarrow y = Ce^{kt}$. Trong phương trình trên C là giá trị ban đầu, k là hệ số tỷ lệ, t chỉ thời gian và y là một hàm số theo t chỉ số lượng chất, dân số, ... tại thời điểm t. Mô hình trên được gọi là *mô hình tăng trưởng mũ*. Mô hình tăng trưởng mũ là không bị giới hạn.

Ví dụ 3.2.20. Một tổ chức quản lý động vật hoang dã của một bang ở Mỹ đã thả 100 con hươu vào một khu vực bảo tồn. Trong suốt 5 năm đầu, quần thể đã tăng lên 432 con. Tổ chức quản lý này đã tin tưởng rằng số lượng cá thể của quần thể này có thể được mô hình bởi hàm tăng trưởng logarit, với giới hạn 2000 con. Viết mô hình tăng trưởng lôgarit cho quần thể này. Sau đó sử dụng mô hình này để lập bảng chỉ ra kích thước của quần thể hươu trong vòng 30 năm tới.

Giải: Gọi y là số lượng hươu trong t năm. Mô hình lôgarit có nghĩa là tốc độ thay đổi của số lượng cá thể trong quần thể tỷ lệ với y và $2000 - y$, phương trình có dạng

$$\frac{dy}{dt} = ky(2000 - y), \quad 100 \le y \le 2000.$$

Giải phương trình trên, ta được

$$y = \frac{2000}{1 + be^{-kt}}. \tag{3.4}$$

Giả thiết cho $y = 100$ khi $t = 0$ nên thay vào phương trình (3.4) có $b = 19$.

$$y = 432 \text{ khi } t = 5 \text{ nên có } 432 = \frac{2000}{1 + 19e^{-k.5}} \Leftrightarrow k \approx 0,33106.$$

Vậy mô hình tăng trưởng cho quần thể là $y = \dfrac{2000}{1 + 19e^{-0,33106t}}$.

Số lượng cá thể trong quần thể trong 5 năm đầu được chỉ ra bởi bảng sau:

Thời gian t	0	5	10	15	20	25	30
Kích thước của quần thể	100	432	1181	1766	1951	1990	1998

Ví dụ 3.2.21. Lợi nhuận biên cho sản phẩm của một công ty kinh doanh được cho bởi mô hình $\dfrac{dP}{dx} = -0,0005x + 12,2$ trong đó P là doanh thu và x là số sản phẩm đã bán. Lợi nhuận sẽ tăng lên bao nhiêu khi bán hàng tăng từ 100 đến 110 đơn vị sản phẩm.

Giải: Sự thay đổi lợi nhuận của sản phẩm khi bán hàng tăng từ 100 đến 110 đơn vị chính là tích phân xác định của hàm P trên đoạn $[100, 110]$. Do đó ta có lợi nhuận cần tìm là:

$$\int_{100}^{110} \frac{dP}{dx}dx = \int_{100}^{110} (-0,0005x + 12,2)dx = (-0,00025x^2 + 12,2x)\Big|_{100}^{110} = 121,48\$.$$

3.3. Tích phân suy rộng

3.3.1. Tích phân suy rộng với cận vô hạn

Định nghĩa 3.3.1. Khoảng lấy tích phân là $[a; +\infty)$.

Cho hàm $f(x)$ xác định trên $[a; +\infty)$, khả tích trên mọi đoạn $[a, b](a \leq x \leq b < +\infty)$ tức là tồn tại tích phân với mọi x thuộc $[a, b]; (b > a)$. Ta gọi giới hạn

$$\lim_{b\to+\infty} \int_a^b f(x)dx \tag{3.5}$$

là *tích phân suy rộng của hàm $f(x)$ trong* $[a; +\infty)$. Ký hiệu:

$$\int_a^{+\infty} f(x)dx. \tag{3.6}$$

+ Nếu (3.5) tồn tại (hữu hạn) thì ta nói tích phân suy rộng là *hội tụ*.

$$\int_a^{+\infty} f(x)dx = \lim_{b\to+\infty} \int_a^b f(x)dx. \tag{3.7}$$

+ Nếu (3.5) không tồn tại thì ta nói $\displaystyle\int_a^{+\infty} f(x)dx$ là *phân kỳ*.

Chú ý 3.3.2. (1) Tương tự ta định nghĩa:

$$+)\ \int_{-\infty}^{a} f(x)dx = \lim_{c\to-\infty} \int_c^a f(x)dx.$$

$$+)\ \int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{a} f(x)dx + \int_a^{+\infty} f(x)dx.$$

Ta thấy $\displaystyle\int_{-\infty}^{+\infty} f(x)dx$ chỉ hội tụ khi và chỉ khi cả hai tích phân ở vế phải đều hội tụ.

(2) Để tính tích phân (3.6) ta làm theo thứ tự như sau: Tính tích phân

$\int\limits_a^b f(x)dx = F(b) - F(a)$ (kết quả là một hàm số theo biến b) $\to$ Tính $\lim\limits_{b\to+\infty}(F(b) - F(a))$.
(3) Nếu tính tích phân suy rộng bằng phương pháp đổi biến số, ta tiến hành đổi cận bình thường lưu ý là khi $x \to +\infty$ thì lấy giới hạn của hàm biến mới xem biến mới tiến đến đâu.
(4) Tích phân suy rộng với cần vô hạn còn được gọi là tích phân suy rộng loại một.

Ví dụ 3.3.3. Xét sự hội tụ của tích phân $I = \int\limits_a^{+\infty} \dfrac{dx}{x^\alpha} \quad (\alpha > 0,\ a > 0)$.

Giải: Nếu $\alpha = 1$ thì $I = \int\limits_a^{+\infty} \dfrac{dx}{x} = \lim\limits_{b\to+\infty} \ln x\Big|_a^b = +\infty$.

Nếu $\alpha \neq 1$. Ta có $\int\limits_a^b \dfrac{dx}{x^\alpha} = \left(\dfrac{x^{1-\alpha}}{1-\alpha}\right)\Bigg|_a^b = \dfrac{1}{1-\alpha}\left(b^{1-\alpha} - a^{1-\alpha}\right)$.

Suy ra $I = \lim\limits_{b\to+\infty} \dfrac{1}{1-\alpha}\left(b^{1-\alpha} - a^{1-\alpha}\right) = \begin{cases} +\infty & \text{khi} \quad \alpha < 1 \\ -\dfrac{a^{1-\alpha}}{1-\alpha} & \text{khi} \quad \alpha > 1 \end{cases}$.

Vậy $\int\limits_a^{+\infty} \dfrac{dx}{x^\alpha}$ hội tụ khi $\alpha > 1$ và phân kỳ khi $\alpha \leq 1$.

Ví dụ 3.3.4. Tính tích phân suy rộng $I = \int\limits_{-\infty}^{+\infty} \dfrac{dx}{(1+x^2)^2}$.

Giải: Ta biến đổi tích phân I như sau:

$$I = \lim_{b\to+\infty} \int\limits_{-b}^{0} \frac{dx}{(1+x^2)^2} + \lim_{b\to+\infty} \int\limits_{0}^{b} \frac{dx}{(1+x^2)^2} = 2\lim_{b\to+\infty} \int\limits_{0}^{b} \frac{dx}{(1+x^2)^2}$$

$$= 2\lim_{b\to+\infty} \left(\frac{1}{2}\arctan x + \frac{x}{2(1+x^2)}\right)\Bigg|_0^b = \frac{\pi}{2}.$$

Chú ý 3.3.5. (1) Nếu $\int\limits_a^{+\infty} |f(x)|dx$ hội tụ thì $\int\limits_a^{+\infty} f(x)dx$ cũng hội tụ khi đó ta nói rằng $\int\limits_a^{+\infty} f(x)dx$ hội tụ tuyệt đối.

(2) Nếu $\int\limits_a^{+\infty} f(x)dx$ hội tụ nhưng $\int\limits_a^{+\infty} |f(x)|dx$ phân kỳ thì ta nói $\int\limits_a^{+\infty} f(x)dx$ bán hội tụ (hội tụ không tuyệt đối).

3.3.2. Tích phân suy rộng với hàm không giới nội

Định nghĩa 3.3.6. Cho hàm $f(x)$ khả tích trên $[a, b-\varepsilon]$ với $\varepsilon > 0$ bé tùy ý và không liên tục tại b (tức là $\lim\limits_{x\to b-0} f(x) = \infty$). Ta gọi giới hạn

$$\lim_{\varepsilon\to 0^+} \int\limits_a^{b-\varepsilon} f(x)dx \tag{3.8}$$

là *tích phân suy rộng của* $f(x)$ *trên* $[a, b]$. Ký hiệu:

$$\int\limits_a^b f(x)dx. \tag{3.9}$$

+ Nếu giới hạn (3.8) hữu hạn ta nói tích phân (3.9) *hội tụ* và

$$\int\limits_a^b f(x)dx = \lim_{\varepsilon\to 0^+} \int\limits_a^{b-\varepsilon} f(x)dx.$$

+ Nếu giới hạn (3.8) không tồn tại ta nói tích phân (3.9) *phân kỳ*. Điểm b khi đó được gọi là *điểm bất thường* (hay *điểm kỳ dị*).

Chú ý 3.3.7. (1) Tương tự ta định nghĩa cho tích phân suy rộng của hàm $f(x)$ khả tích trên mọi đoạn $[a+\varepsilon, b], \varepsilon > 0$, nhỏ tùy ý và không giới nội khi $x \to a+0$:

$$\int\limits_a^b f(x)dx = \lim_{\varepsilon\to 0^+} \int\limits_{a+\varepsilon}^{b} f(x)dx.$$

(2) Nếu $f(x)$ không giới nội khi $x \to x_0$ với $x_0 \in (a, b)$ ta chia đoạn $[a, b]$ thành 2 khoảng $[a, x_0]$ và $[x_0, b]$ rồi áp dụng cách tính trên.
(3) Nếu $f(x)$ có điểm gián đoạn vô cực tại $x = a$ (hoặc $x = b$; hoặc tại một số hữu hạn điểm trên $[a, b]$) và $F(x)$ là hàm thỏa mãn hai điều kiện:

+ $F(x)$ là nguyên hàm của $f(x)$ tại những điểm x trên $[a, b]$ trừ những điểm $f(x)$ gián đoạn vô cực;

+ $F(x)$ liên tục trên $[a, b]$; thì các tích phân trên vẫn được tính theo công thức Newton-Leibnitz.
(4) Tích phân suy rộng với hàm không giới nội còn được gọi là tích phân suy rộng loại hai.

Ví dụ 3.3.8. Tính tích phân $I = \int\limits_0^2 \dfrac{\sqrt{2+x}}{\sqrt{2-x}}dx$.

Giải: Dễ thấy, khi $x \neq 2$ thì $f(x) = \dfrac{\sqrt{2+x}}{\sqrt{2-x}} = \dfrac{2+x}{\sqrt{4-x^2}} = \dfrac{2}{\sqrt{4-x^2}} + \dfrac{x}{\sqrt{4-x^2}}$ và

$$\int \frac{\sqrt{2+x}}{\sqrt{2-x}}dx = 2\arcsin\frac{x}{2} - \sqrt{4-x^2} + C.$$

Ta thấy $F(x) = \arcsin\dfrac{x}{2} - \sqrt{4-x^2}$ là một nguyên hàm của $f(x)$ trong $[0,2)$ nhưng $F(x)$ liên tục trên toàn đoạn $[0,2]$. Do đó

$$I = \left(\arcsin\frac{x}{2} - \sqrt{4-x^2}\right)\Bigg|_0^2 = 2+\pi.$$

Ví dụ 3.3.9. Tính tích phân $I = \displaystyle\int\limits_1^2 \frac{dx}{x\ln x}\cdot$

Giải: Ta thấy, hàm số $f(x) = \dfrac{1}{x\ln x}$ gián đoạn vô cực tại $x=1$ và hàm số $F(x) = \ln(\ln x)$ là một nguyên hàm của $f(x)$ trên $(1,2]$ nhưng $F(x)$ không liên tục trên $[1,2]$. Do đó

$$I = \lim_{\varepsilon\to 0}\int\limits_{1+\varepsilon}^2 \frac{dx}{x\ln x} = \lim_{\varepsilon\to 0}[\ln(\ln 2) - \ln(\ln(1+\varepsilon))] = \infty.$$

Vậy tích phân trên là phân kỳ.

BÀI TẬP CHƯƠNG 3

Bài 1: Tính các tích phân

1) $\displaystyle\int \frac{\sqrt{1+x^2}+\sqrt{1-x^2}}{\sqrt{1-x^4}}dx$; 2) $\displaystyle\int x\sin x dx$; 3) $\displaystyle\int \frac{2x-1}{\sqrt{x^{\frac{4}{3}}}}dx$;

4) $\displaystyle\int x^2\sqrt[3]{x^3-2}dx$; 5) $\displaystyle\int \frac{-4xdx}{(1-2x^2)^2}$; 6) $\displaystyle\int x^3(3x^4+1)^2dx$;

7) $\displaystyle\int \frac{3xdx}{\sqrt{4-x^2}}$; 8) $\displaystyle\int \frac{x^2+1}{\sqrt{x^3+3x+4}}dx$; 9) $\displaystyle\int \frac{dx}{x\ln^2 x}$.

Bài 2: Tốc độ biến thiên của số lượng vi khuẩn theo một đơn vị thời gian t được đo bởi

$$\frac{dP}{dt} = \frac{3000}{1+0,25t}.$$

Trong đó, t là thời gian tính bằng đơn vị ngày. Khi $t=0$ thì số vi khuẩn $P = 1000$.

a) Viết phương trình mô tả số lượng vi khuẩn theo thời gian t;

b) Số vi khuẩn là bao nhiêu sau 3 ngày;

c) Sau bao lâu số vi khuẩn sẽ lên đến 12 000 con.

Bài 3: Doanh thu biên cho việc bán một sản phẩm được mô hình bởi

$$\frac{dR}{dx} = 50 - 0,02x + \frac{100}{x+1}\cdot$$

Với x là số lượng hàng hóa đã bán.

a) Tìm hàm doanh thu R biết khi $x = 0 \Rightarrow R = 0$;

b) Tìm tổng doanh thu khi bán được 1500 sản phẩm;

c) Phải bán được bao nhiêu sản phẩm để tổng doanh thu đạt 60230$.

Bài 4: Mức lương trung bình cho một người quản lý (S\$) ở Mỹ được thay đổi với tỷ lệ

$$\frac{dS}{dt} = 2621,7e^{0,07t}.$$

Với $t = 5$ tương ứng với năm 1995. Năm 2001, mức lương trung bình cho người quản lý đã là 118496$.

a) Tìm hàm số mô tả mức lương trung bình của người quản lý mỗi năm.

b) Năm 1999, mức lương trung bình của người quản lý là bao nhiêu?

Bài 5: Do sự cung cấp thiếu oxy nên cá hồi trong hồ đang bị chết dần. Tỷ lệ thay đổi của số lượng cá hồi trong hồ được đo bởi

$$\frac{dP}{dt} = -125e^{\frac{-t}{20}}.$$

Với t là thời gian tính bằng ngày. Khi $t = 0$ thì số cá hồi trong hồ là 2500.

a) Viết phương trình mô tả số lượng cá hồi theo thời gian t;

b) Số lượng cá hồi còn là bao nhiêu sau 15 ngày.

Bài 6: Một vườn ươm cây xanh thường bán một loại cây bụi sau 5 năm trồng và chăm sóc. Tỷ lệ phát triển của cây sau 5 năm được đo bởi

$$\frac{dh}{dt} = \frac{17,6t}{\sqrt{17,6t^2+1}}.$$

Với t là thời gian tính bằng năm, h là chiều cao của cây tính bằng cm. Biết mầm cây trước khi đem ươm cao $6cm$.

a) Tìm hàm số mô tả chiều cao của cây;

b) Khi cây được đem bán thì chúng cao bao nhiêu?

Bài 7: Chi phí biên cho việc sản xuất x đơn vị sản phẩm được mô hình bởi

$$\frac{dC}{dx} = 32 - 0,04x.$$

Biết chi phí để sản xuất một đơn vị sản phẩm là 50 000 đồng. Tìm tổng chi phí để sản xuất 200 sản phẩm.

Bài 8: Lợi nhuận biên cho một loại sản phẩm được mô hình bởi

$$\frac{dP}{dx} = -0,0005x + 12,2.$$

a) Lợi nhuận tăng lên bao nhiêu khi bán hàng tăng từ 100 đến 101 đơn vị sản phẩm.

b) Lợi nhuận tăng lên bao nhiêu khi bán hàng tăng từ 100 đến 110 đơn vị sản phẩm.

Bài 9: Tính các tích phân:

1) $\int\limits_0^1 \left(x + \frac{1}{x}\right)^2 dx$; 2) $\int\limits_0^1 \left(e^{2x} + \frac{3}{x+1}\right) dx$; 3) $\int\limits_1^3 \frac{x^3 - x^2 + x}{x} dx$;

4) $\int\limits_{\frac{1}{e}}^{e} |\ln x| dx$; 5) $\int\limits_0^1 \frac{2x^2}{x^3+1} dx$; 6) $\int\limits_{10}^{12} \frac{4x+2}{x^2+x-2} dx$;

7) $\int\limits_0^{\frac{\pi}{2}} \frac{\cos x dx}{1 + \sin x}$; 8) $\int\limits_{\frac{\pi}{4}}^{\frac{\pi}{3}} \frac{dx}{\sin 2x}$; 9) $\int\limits_0^{\frac{\pi}{2}} \frac{\sin x + \cos x}{\sqrt[3]{\sin x - \cos x}} dx$.

Bài 10: Tính các tích phân suy rộng:

1) $\int\limits_{\frac{2}{\pi}}^{+\infty} \frac{1}{x^2} \sin \frac{1}{x} dx$; 2) $\int\limits_0^{+\infty} \frac{\arctan x}{(1+x^2)^{\frac{3}{2}}} dx$; 3) $\int\limits_0^{+\infty} \frac{x^2+1}{x^4+1} dx$;

4) $\int\limits_{a^2}^{+\infty} \frac{dx}{x\sqrt{1+x^2}}$; 5) $\int\limits_{\sqrt{2}}^{+\infty} \frac{dx}{x\sqrt{x^2-1}}$; 6) $\int\limits_0^{+\infty} x e^{-x^2} dx$.

Mục lục

Tài liệu tham khảo

[1] L. Edwards, *Calculus an applied approach seventh edition*, Houghton Mifflin Company, New York, 2006.

[2] Trần Đức Long, Nguyễn Đình Sang, Hoàng Quốc Toàn, *Giáo trình giải tích tập 1, 2*, Nhà xuất bản Đại học quốc gia, Hà Nội, 2001.

[3] Nguyễn Duy Thuận, Phí Mạnh Ban, Nông Quốc Chinh, *Đại số tuyến tính*, Nhà xuất bản Đại học Sư phạm, Hà Nội, 2003.

[4] Lê Đình Thúy, *Toán cao cấp cho các nhà kinh tế*, Nhà xuất bản Thống kê, Hà Nội, 2005.

[5] Nguyễn Đình Trí, *Toán học cao cấp tập 1, 2,3*, Nhà xuất bản giáo dục, Hà Nội, 2001.

Phần 2. Xác suất

Chương 1

Khái niệm và các phép toán

1.1. Giải tích tổ hợp

1.1.1. Qui tắc cộng, quy tắc nhân

a. Quy tắc cộng

Nếu một công việc được chia ra thành k trường hợp để thực hiện, trường hợp một có n_1 cách thực hiện xong công việc, trường hợp hai có n_2 cách thực hiện xong công việc, . . ., trường hợp k có n_k cách thực hiện xong công việc và không có một cách thực hiện nào ở trường hợp này lại trùng với một cách thực hiện ở trường hợp khác.

Khi đó ta có: $n = n_1 + n_2 + \ldots + n_k$ cách thực hiện công việc.

b. Qui tắc nhân

Giả sử một công việc nào đó được chia thành k giai đoạn. Có n_1 cách thực hiện giai đoạn thứ nhất, n_2 cách thực hiện giai đoạn thứ hai, . . ., n_k cách thực hiện giai đoạn thứ k.

Khi đó ta có: $n = n_1.n_2..n_k$ cách thực hiện công việc.

Ví dụ 1.1.1. Giả sử để đi từ A đến C ta bắt buộc phải đi qua điểm B. Có 3 đường khác nhau để đi từ A đến B và có 2 đường khác nhau để đi từ B đến C.

Vậy có n = 3.2 = 6 cách khác nhau để đi từ A đến C

1.1.2. Chỉnh hợp

a. Chỉnh hợp không lặp

Định nghĩa 1.1.2. *Một chỉnh hợp chập k của* n *phần tử* là một nhóm sắp thứ tự gồm k phần tử khác nhau lấy từ n phần tử đã cho ($k \le n$).

Ký hiệu và công thức: $A_n^k = \dfrac{n!}{(n-k)!} = n(n-1)\ldots(n-k+1)$

Như vậy: Hai chỉnh hợp chập k của n phần tử khác nhau nếu:

- Hoặc chúng có ít nhất một phần tử khác nhau;
- Hoặc chúng gồm k phần tử như nhau nhưng sắp xếp theo thứ tự khác nhau.

Ví dụ 1.1.3. Một lớp phải học 10 môn, mỗi ngày phải học 2 môn. Hỏi có bao nhiêu cách sắp xếp thời khóa biểu trong một ngày.
Giải: Vì mỗi cách sắp xếp thời khóa biểu trong ngày là việc ghép hai môn trong số 10 môn, các cách sắp xếp này khác nhau do có ít nhất một môn khác nhau hoặc chỉ do thứ tự sắp xếp trước sau giữa hai môn. Vì thế mỗi cách sắp xếp ứng với một chỉnh hợp chập 2 từ 10 phần tử.

Tức là có: $A_{10}^2 = 10.9 = 90$ (cách).

b. Chỉnh hợp có lặp

Định nghĩa 1.1.4. *Một chỉnh hợp có lặp chập k của n phần* tử là một nhóm sắp thứ tự gồm k phần tử lấy từ n phần tử đã cho trong đó mỗi phần tử có thể có mặt đến n lần trong nhóm tạo thành.

Ký hiệu và công thức: $\overline{A_n^k} = n^k$

Ví dụ 1.1.5. Để đăng ký xe máy người ta dùng 4 chữ số từ 0,1,....,9 cho một sêri. Hỏi mỗi sêri có thể đăng ký được bao nhiêu xe.
Giải: Số xe máy được đăng ký trong một sêri chính là chỉnh hợp có lặp chập 4 của 10 (trừ đi 1 vì trong thực tế không dùng 4 chữ số 0)

$$\overline{A_{10}^4} - 1 = 10^4 - 1 = 9999 \text{ (xe).}$$

Ví dụ 1.1.6. Để truyền tin bằng tín hiệu mooc-xơ gồm hai kí hiệu chấm (.) và vạch (-), người ta mã hóa mỗi chữ cái của bảng chữ cái thành một nhóm có thứ tự gồm không quá 4 kí hiệu. Biết rằng cùng một kí hiệu có thể có mặt nhiều lần trong nhóm có thứ tự tạo thành. Hỏi có thể mã hóa được bao nhiêu chữ cái ?
Giải: Một nhóm có thứ tự gồm k kí hiệu $(1 \le k \le 4)$ tạo nên chính là một chỉnh hợp lặp chập k từ 2 phần tử đã cho. Vì vậy số chữ cái mã hóa được là:

$$\overline{A_2^1} + \overline{A_2^2} + \overline{A_2^3} + \overline{A_2^4} = 1^2 + 2^2 + 2^3 + 2^4 = 30$$

Như vậy nếu bảng chữ cái của một thứ tiếng nào đó gồm không quá 30 chữ thì ta có thể mã hóa theo cách trên.

1.1.3 Hoán vị

Định nghĩa 1.1.7. *Hoán vị* là một chỉnh hợp không lặp chập n của n phần tử. Hay hoán vị của n phần tử là một nhóm có thứ tự gồm đủ mặt n phần tử đã cho.
Vậy các hoán vị của n phần tử chỉ khác nhau bởi thứ tự sắp xếp giữa các phần tử đó.

Kí hiệu và công thức: $P_n = n(n-1)(n-2)....1 = n!$

Ví dụ 1.1.8. Có bao nhiêu cách xếp 5 người ngồi vào một chiếc ghế dài gồm 5 chỗ.
Giải: Số cách xếp 5 người vào một ghế dài gồm 5 chỗ chính là hoán vị của 5 phần tử, nên ta có

$$P_5 = 5! = 5.4.3.2.1 = 120 \text{ (cách).}$$

Ví dụ 1.1.9. Có bao nhiêu cách xếp n đại biểu ngồi quanh một bàn tròn.
Giải: Do các chỗ ngồi quanh một bàn tròn không có phần tử thứ nhất và phần tử cuối cùng nên đại biểu thứ nhất được ngồi tự do. Các đại biểu còn lại có số cách chọn vị trí ngồi lần lượt là: (n-1),(n-2),...,1.

Vậy cách xếp n đại biểu ngồi quanh bàn tròn là: $(n-1)!$ (cách).

1.1.4. Tổ hợp

Định nghĩa 1.1.10. *Một tổ hợp chập* k *của* n *phần tử* $(k \le n)$ là một nhóm không phân biệt thứ tự gồm k phần tử khác nhau lấy ra từ n phần tử đã cho**.**

Ký hiệu và công thức: $C_n^k = \dfrac{n!}{k!(n-k)!}$

Vậy: Mỗi tổ hợp gồm các phần tử khác nhau, hai tổ hợp khác nhau là do các phần tử chứa trong chúng khác nhau chứ không phải là do bộ khác nhau về thứ tự của các phần tử trong đó.

Một vài tính chất của tổ hợp:

+ $C_n^k = C_n^{n-k}$
+ $C_n^0 = C_n^n = 1$ (Quy ước: $0! = 1$)
+ $C_n^1 = C_n^{n-1} = n$
+ $C_{n+1}^k = C_n^k + C_n^{k-1}$

Ví dụ 1.1.11. 10 đội bóng thi đấu với nhau theo thể thức đấu vòng. Hỏi phải tổ chức bao nhiêu trận đấu?

Giải: Mỗi trận đấu ứng với một nhóm gồm 2 phần tử từ 10 đội (không phân biệt thứ tự). Vì vậy phải tổ chức tất cả:

$$C_{10}^2 = \frac{10.9}{2} = 45 \text{ trận đấu.}$$

Chú ý:

** Để nhận dạng một hoán vị của* n *phần tử ta thường dùng các dấu hiệu đặc trưng sau:*

+ n phần tử đều phải có mặt;
+ Mỗi phần tử chỉ xuất hiện đúng một lần;
+ Có thứ tự giữa các phần tử.

** Ta sử dụng khái niệm chỉnh hợp khi gặp tình huống:*

+ Phải chọn k phần tử từ n phần tử;
+ Sắp thứ tự k phần tử đó.

** Ta sử dụng khái niệm tổ hợp khi gặp tình huống*:

+ Cần chọn ra từ một tập có n phần tử một tập con có k phần tử;
+ Lưu ý rằng trong tập con k phần tử đó ta không quan tâm đến thứ tự của các phần tử.

1.2. Phép thử và biến cố

1.2.1. Phép thử và biến cố

Định nghĩa 1.2.1. Khi thực hiện một nhóm các điều kiện nào đó ta nói rằng đã thực hiện một *phép thử*. Hiện tượng được xét trong phép thử gọi là *biến cố* (hay *sự kiện*)

Ví dụ 1.2.2. Tung một con xúc xắc là thực hiện một phép thử. Hiện tượng xúc xắc xuất hiện mặt 3 chấm; xúc xắc xuất hiện mặt 6 chấm; xúc xắc xuất hiện mặt có số chấm lớn hơn 6; xúc xắc xuất hiện mặt có số chấm nhỏ hơn 7 là các biến cố.

Ví dụ 1.2.3. Tung một đồng xu là thực hiện một phép thử. Hiện tượng: đồng xu xuất hiện mặt sấp; đồng xu xuất hiện mặt ngửa là các biến cố.

Phân loại phép thử: 2 loại

+ *Phép thử lặp*: Là phép thử được thực hiện trong những điều kiện như nhau.

+ *Phép thử không lặp*: Là những phép thử được thực hiện trong những điều kiện khác nhau.

Phân loại biến cố: 3 loại

+ *Biến cố ngẫu nhiên*: là biến cố có thể xảy ra hoặc không xảy ra khi thực hiện một phép thử.

Ký hiệu: **A, B, C,...**

Ví dụ 1.2.4. Trong ví dụ 1.2.2. ở trên, biến cố "xúc xắc xuất hiện mặt ba chấm", "xúc xắc xuất hiện mặt 6 chấm" là các biến cố ngẫu nhiên.

+ *Biến cố chắc chắn*: là biến cố nhất định sẽ xảy ra khi thực hiện một phép thử.

Ký hiệu: **U** (hoặc Ω)

Ví dụ 1.2.5. Trong ví dụ 1.2.2. ở trên, biến cố "xúc xắc xuất hiện mặt có số chấm nhỏ hơn 7" là biến cố chắc chắn.

+ *Biến cố không thể có*: là biến cố nhất định không xảy ra khi thực hiện một phép thử.

Ký hiệu: **V** (hoặc Φ)

Ví dụ 1.2.6. Trong ví dụ 1.2.2. ở trên, biến cố "xúc xắc xuất hiện mặt có số chấm lớn hơn 6" là biến cố không thể có.

1.2.2. Quan hệ giữa các biến cố

a. Hợp (tổng) của các biến cố

Định nghĩa 1.2.7. Biến cố A được gọi là *hợp* của các biến cố $A_1, A_2, ..., A_n$ nếu A xảy ra khi và chỉ khi ít nhất một trong các biến cố A_i $(i = 1,...,n)$ xảy ra. Và ta viết:

$A = A_1 \cup A_2 \cup ... \cup A_n$

Ví dụ 1.2.8. Tung một con xúc xắc, gọi A_i là biến cố "xuất hiện mặt i chấm" (i=1,2,...6); gọi A là biến cố "xuất hiện mặt có số chấm chẵn". Khi đó ta có $A = A_2 \cup A_4 \cup A_6$.

b. Giao (tích) của các biến cố

Định nghĩa 1.2.9. Biến cố B được gọi là *giao* của các biến cố $A_1, A_2, ..., A_n$ nếu B xảy ra khi và chỉ khi tất cả các biến cố A_i $(i = 1,...,n)$ xảy ra. Và ta viết:

$B = A_1 \cap A_2 \cap ... \cap A_n$

Ví dụ 1.2.10. Một mạch điện gồm 2 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 2 bị cháy khi điện quá tải"; C là biến cố "mạch điện bị ngắt khi điện quá tải", vậy $C = A \cap B$ và C chỉ xảy ra khi cả 2 biến cố A và B cùng đồng thời xẩy ra.

c. Biến cố xung khắc

Định nghĩa 1.2.11. Hai biến cố A và B được gọi là *xung khắc* với nhau nếu chúng không đồng thời xẩy ra trong một phép thử.

Như vậy, nếu A và B xung khắc thì: $A \cap B = V$

Ví dụ 1.2.12. Tung một con xúc xắc, gọi A là biến cố "xuất hiện mặt 2 chấm"; B là biến cố "xuất hiện mặt 3 chấm", khi đó A và B là xung khắc nhau.

d. Biến cố đối lập

Định nghĩa 1.2.13. Biến cố không xảy ra biến cố A được gọi là biến cố *đối lập* của A. Kí hiệu biến cố đối lập của biến cố A là $\overline{A}$

Như vậy ta có: $\begin{cases} A \cup \overline{A} = U \\ A \cap \overline{A} - V \end{cases}$

Ví dụ 1.2.14. Bắn một viên đạn vào bia, biến cố "bắn trúng bia" và biến cố "bắn trượt bia" là hai biến cố đối lập.

e. Hệ đầy đủ các biến cố

Định nghĩa 1.2.15. Các biến cố $A_1, A_2, ..., A_n$ được gọi là *hệ đầy đủ* các biến cố nếu trong kết quả của phép thử sẽ xẩy ra một và chỉ một trong các biến cố đó.

Nghĩa là ta có $\begin{cases} A_i A_j = V \qquad (\forall i \neq j) \\ \sum\limits_{i=1}^{n} A_i = U \end{cases}$

Ví dụ 1.2.16. Gieo một con xúc xắc. Gọi A_i là biến cố "xúc xắc xuất hiện mặt i chấm" (i=1,2,...6) thì các biến cố $A_1; A_2; A_3; A_4; A_5; A_6$ tạo nên một hệ đầy đủ các biến cố.

Nếu khả năng xẩy ra các biến cố đó là như nhau ta gọi là hệ đầy đủ đồng khả năng.

1.2.3. Biến cố sơ cấp và không gian các biến cố sơ cấp

Một biến cố ngẫu nhiên được gọi là phức hợp nếu nó có thể biểu diễn được dưới dạng hợp của hai biến cố không đồng nhất với nó. Một biến cố không là phức hợp được gọi là biến cố sơ cấp (Nói cách khác: Các kết quả có thể có khi một phép thử được thực hiện gọi là các biến cố sơ cấp-hoặc các biến cố cơ bản). Vậy một biến cố phức hợp có thể xuất hiện theo nhiều cách khác nhau. Biến cố sơ cấp chỉ xuất hiện theo một cách duy nhất. Các biến cố sơ cấp từng đôi xung khắc. Tập hợp mọi biến cố sơ cấp của một phép thử được gọi là *không gian các biến cố sơ cấp.*

1.3. Các định nghĩa về xác suất

Mọi biến cố ngẫu nhiên đều giống nhau ở chỗ chúng không chắc chắn, nhưng khả năng xảy ra của mỗi biến cố lại có thể khác nhau. Với mỗi một biến cố ngẫu nhiên, người ta dùng một con số để đặc trưng cho khả năng xảy ra của biến cố đó nhiều hay ít, số đó được gọi là *xác suất* của biến cố A.

Ký hiệu: *P(A)* (P là viết tắt từ chữ *Probability*)

1.3.1. Định nghĩa cổ điển về xác suất

a. Định nghĩa 1.3.1. Giả sử trong một phép thử có n kết quả đồng khả năng có thể xảy ra. Khi đó *xác suất* xuất hiện biến cố A trong một phép thử là tỷ 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.

Nghĩa là: $p(A)=\dfrac{m}{n}$, trong đó: m là số kết cục thuận lợi cho A; n là tổng số các kết cục duy nhất đồng khả năng có thể xảy ra.

b. Tính chất*:*

(a) $0 \le P(A) \le 1$.

(b) $P(U)=1$.

(c) $P(V)=0$.

Ví dụ 1.3.2. Trong một lô xổ số có 100 vé trong đó có 8 vé có thưởng. Mua ngẫu nhiên 5 vé. Tính xác suất để trong 5 vé đã mua có 2 vé trúng thưởng.

Giải: Gọi A là biến cố "trong 5 vé đã mua có 2 vé có thưởng"

Số cách mua 5 vé là: $n=C_{100}^5$

Số kết cục thuận lợi cho A là $m=C_8^2.C_{92}^3$.

Vậy $P(A)=\dfrac{C_8^2.C_{92}^3}{C_{100}^5}$

Ví dụ 1.3.3. Một người khi gọi điện thoại quên mất 2 số cuối cùng của số điện thoại và chỉ nhớ được rằng chúng khác nhau. Tìm xác suất để quay ngẫu nhiên một lần được đúng số cần gọi.

Giải: Gọi A là biến cố "quay ngẫu nhiên một lần được đúng số cần gọi".

Số kết cục đồng khả năng là: $n=A_{10}^2=90$;

Số kết cục thuận lợi cho A là: $m=1$.

Vậy $P(A)=\dfrac{1}{90}$

Ví dụ 1.3.4. Trên 7 tấm bia kích thước như nhau trên mỗi tấm có ghi các chữ cái : 1 tấm ghi chữ H, 2 tấm ghi chữ O; 1 tấm ghi chữ C; 1 tấm ghi chữ A; 1 tấm ghi chữ T và 1 tấm ghi chữ N. Tính xác suất để xếp ngẫu nhiên 7 tấm bìa đó thành hàng ngang đọc được chữ "HOCTOAN".

Giải: Gọi A là biến cố "đọc thành chữ HOCTOAN" .

Số kết cục đồng khả năng $n=7!$

Số kết cục thuận lợi $m=1.2!.1\ldots1=2$.

Vậy $P(A) = \dfrac{2}{7!}$

Ví dụ 1.3.5. Một hộp chứa 7 cầu trắng và 3 cầu đen cùng kích thước. Rút ngẫu nhiên cùng lúc 4 cầu. Tính xác suất để trong bốn cầu rút được có:

a/ 2 cầu đen

b/ Ít nhất 2 cầu đen

c/ Toàn cầu trắng

Giải:

Ta thấy rút ngẫu nhiên cùng lúc 4 trong 10 cầu có n = C_{10}^4 cách

a/ Gọi A là sự kiện "trong 4 cầu rút ra có 2 cầu đen"

Số kết cục thuận lợi A là m = $C_3^2.C_7^2$

$$\text{Vậy P(A)} = \frac{C_3^2.C_7^2}{C_{10}^4} = 0{,}3$$

b/ Gọi B là sự kiện "trong 4 cầu rút ra có ít nhất 2 cầu đen"

Số kết cục thuận lợi B là m = $C_3^2.C_7^2 + C_7^1.C_3^3$

$$\text{Vậy P(B)} = \frac{C_7^2.C_3^2 + C_7^1.C_3^3}{C_{10}^4} = \frac{1}{3}$$

c/ Gọi C là sự kiện "4 cầu rút ra toàn cầu trắng"

Số kết cục thuận lợi C là m = C_7^4

$$\text{Vậy P(C)} = \frac{C_7^4}{C_{10}^4} = \frac{1}{6}$$

c. Ưu điểm và hạn chế của định nghĩa cổ điển về xác suất

+ Ưu điểm là tìm xác suất của biến cố ta không phải tiến hành phép thử. (Phép thử chỉ tiến hành một cách giả định).

+ Hạn chế: Nó đòi hỏi là số kết cục duy nhất đồng khả năng có thể xảy ra trong phép thử phải là hữu hạn. Hạn chế lớn nhất của định nghĩa cổ điển là trong thực tế nhiều khi không thể biểu diễn kết quả của phép thử dưới dạng tập hợp các kết cục duy nhất và đồng khả năng.

1.3.2. Định nghĩa thống kê về xác suất

a. Định nghĩa tần suất

Định nghĩa 1.3.6. *Tần suất* xuất hiện biến cố A 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.

Kí hiệu: $f(A) = \dfrac{m}{n}$

Ví dụ 1.3.7. Để nghiên cứu khả năng xuất hiện mặt sấp khi tung một đồng xu, người ta tiến hành tung đồng xu nhiều lần và thu được kết quả sau:

Người làm thí nghiệm	Số lần tung (n)	Số lần được mặt sấp (m)	Tần suất f(A)=m/n
Buffon	4040	2048	0,5069
Pearson	12000	6019	0,5016
Pearson	24000	12012	0,5005

b. Định nghĩa xác suất theo thống kê:

Định nghĩa 1.3.8. 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.

c. Ưu điểm và hạn chế của định nghĩa xác suất theo thống kê

+ Ưu điểm: Nó không đòi hỏi những điều kiện áp dụng như với định nghĩa cổ điển. Nó hoàn toàn dựa trên các quan sát thực tế để làm cơ sở kết luận về xác suất xảy ra của một biến cố.

+ Hạn chế: Chỉ áp dụng được đối với các hiện tượng ngẫu nhiên mà tần suất của nó có tính ổn định. Để xác định một cách tương đối chính xác giá trị của xác suất ta phải tiến hành trên thực tế một số đủ lớn các phép thử.

1.3.3. Nguyên lý xác suất lớn và xác suất nhỏ

Trong nhiều bài toán thực tế ta thường gặp các biến cố có xác suất rất nhỏ, tức là gần bằng không. Trong trường hợp đó ta không thể cho rằng những biến cố này sẽ không xảy ra khi thực hiện một phép thử. Thậm chí một biến cố có xác suất bằng không vẫn chưa chắc chắn đã là biến cố không thể có, tức là vẫn có thể xảy ra.

Qua quan sát người ta thấy rằng các biến cố có xác suất nhỏ gần như sẽ không xảy ra khi tiến hành một phép thử. Trên cơ sở đó có thể đưa ra "Nguyên lý thực tế không thể có của các biến cố có xác suất nhỏ" sau đây: *Nếu một biến cố có xác suất rất nhỏ thì thực tế có thể cho rằng trong một phép thử biến cố đó sẽ không xảy ra.* (Lưu ý: việc qui định mức xác suất được coi là rất nhỏ sẽ tùy thuộc vào từng bài toán cụ thể).

Một xác suất khá nhỏ mà với nó có thể cho rằng biến cố thực tế sẽ không xảy ra được gọi là mức ý nghĩa.

Tương tự như vậy ta có thể đưa ra "Nguyên lý thực tế chắc chắn xảy ra của các biến cố có xác suất lớn" như sau: *Nếu biến cố ngẫu nhiên có xác suất gần bằng 1 thì thực tế có thể cho rằng biến cố đó sẽ xảy ra trong một phép thử*. (Việc qui định một mức xác suất đủ coi là lớn tùy thuộc vào từng bài toán cụ thể).

1.4. Các định lý cơ bản về xác suất

1.4.1. Định lý cộng xác suất

Định lý 1.4.1. Nếu A và B là 2 biến cố bất kì thì :

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

Định lý 1.4.2. Nếu A, B và C là 3 biến cố *bất kì thì*:

P(A +B + C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).

Hệ quả 1.4.3.

(a) Xác suất của tổng hai biến cố xung khắc bằng tổng xác suất của các biến cố đó: $P(A+B)=P(A)+P(B)$

(b) Xác suất của tổng các biến cố xung khắc từng đôi một bằng tổng xác suất của các biến cố đó: $P\left(\sum_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)$

(c) Nếu các biến cố $A_1, A_2, ..., A_n$. tạo nên một hệ đầy đủ các biến cố thì tổng xác suất của chúng bằng 1, tức là: $\sum_{i=1}^{n} P(A_i)=1$.

(d) Tổng xác suất của hai biến cố đối lập nhau bằng 1, tức là:

$$P(A)+P(\overline{A})=1$$

Ví dụ 1.4.4. Trong một thùng đựng 30 quả cầu gồm 10 đỏ, 5 xanh và 15 trắng. Rút hú họa một quả. Hãy tính xác suất xuất hiện quả đỏ hoặc xanh.

Giải: Gọi A là sự kiện "xuất hiện quả màu đỏ", B là sự kiện "xuất hiện quả màu xanh". Rõ ràng các sự kiện A, B xung khắc nên:

$$P(A+B)=P(A)+P(B)=\frac{10}{30}+\frac{5}{30}=0{,}5$$

Ví dụ 1.4.5. Trong một cỗ bài có 52 quân (trong đó có 4 quân K). Lấy ngẫu nhiên ra 4 quân. Tính xác suất để trong 4 quân lấy ra có:

a/ Đúng 2 quân là K.

b/ Ít nhất 2 quân là K.

c/ Không quá 2 quân là K.

Giải: Gọi A_i là biến cố "lấy được đúng i quân là K", (i=0,1,2,3,4).

a/ Gọi A là sự kiện "lấy được đúng 2 quân là K", ta có

$$P(A)=P(A_2)=\frac{C_4^2.C_{48}^2}{C_{52}^4}\approx 0{,}025$$

b/ Gọi B là sự kiện "trong 4 quân lấy ra có ít nhất 2 quân là K"

$$B=A_2+A_3+A_4$$

$$P(B)=P(A_2)+P(A_3)+P(A_4)$$

$$=\frac{C_4^2.C_{48}^2}{C_{52}^4}+\frac{C_4^3.C_{48}^1}{C_{52}^4}+\frac{C_4^4}{C_{52}^4}$$

$$=0{,}025+0{,}0007+0{,}0000036$$

$$=0{,}026$$

c/ Gọi C là sự kiện "trong 4 quân lấy ra có không quá 2 quân là K"

$$C=A_0+A_1+A_2 \qquad \Rightarrow \qquad \overline{C}=A_3+A_4$$

$P(\overline{C})=0{,}0007$. Vậy $P(C)=1-P(\overline{C})\approx 0{,}9993$

1.4.2. Định lý nhân xác suất

a. Định nghĩa xác suất có điều kiện

Định nghĩa 1.4.6. Xác suất của biến cố A được tính với điều kiện biến cố B đã xảy ra gọi là *xác suất có điều kiện* của A.

Ký hiệu: $P\left(A/B\right)$

Ví dụ 1.4.7. 5 người rút 5 thăm trong đó có 2 vé đi xem đá bóng. Trước lúc bắt thăm, xác suất rút được vé của anh A (cũng như của anh B) là $P(A)=\frac{2}{5}$. Nếu cho biết thêm điều kiện trước đó B đã rút được một vé thì xác suất để A rút được vé là 1/4. Rõ ràng sự xuất hiện của B (anh B rút được vé) đã thay đổi khả năng rút được vé của A.

Vậy $P(A)=\frac{2}{5}$; $\quad P(A/B)=\frac{1}{4}$; $\quad P(A/\overline{B})=\frac{2}{4}$

b. Định lý nhân xác suất

Định lý 1.4.8. *Xác suất của tích 2 biến cố A và B bằng tích xác suất của một trong 2 biến cố đó với xác suất có điều kiện của biến cố còn lại.*

$$P(AB)=P(A).P\left(B/A\right)=P(B).P(A/B)$$

Hệ quả 1.4.9.

(a) Nếu $P(B)>0$ thì $\quad P\left(A/B\right)=\dfrac{P(AB)}{P(B)}$

(b) $P(A_1A_2..A_n)=P(A_1).P\left(A_2/A_1\right)...P\left(A_n/A_1.A_2...A_{n-1}\right)$

Ví dụ 1.4.10. Trong một bình kín có 10 quả cầu kích thước như nhau trong đó có 6 quả trắng và 4 quả đen lấy ngẫu nhiên liên tiếp không hoàn lại 2 lần mỗi lần 1 quả. Tìm xác suất để:

a/ Cả 2 quả đều là trắng

b/ 2 quả cùng một mầu.

c/ Có ít nhất 1 quả mầu trắng.

Giải:

Gọi A_i là biến cố "quả lấy lần thứ i có mầu trắng", (i=1,2).

a/ Gọi A là biến cố "cả 2 quả đều là mầu trắng".

$$A=A_1A_2 \Rightarrow P(A)=P(A_1A_2)=P(A_1).P(A_2/A_1)=\frac{6}{10}.\frac{5}{9}=\frac{1}{3}$$

b/ Gọi B là sự kiện "2 quả lấy ra cùng mầu", ta có: $\quad B=A_1A_2+\overline{A_1}\,\overline{A_2}$

$$P(B)=P(A_1).P\left(A_2/A_1\right)+P(\overline{A_1}).P\left(\overline{A_2}/\overline{A_1}\right)$$

$$=\frac{6}{10}\frac{5}{9}+\frac{4}{10}.\frac{3}{9}=\frac{7}{15}$$

c/ Gọi C là sự kiện "trong 2 quả lấy ra có ít nhất 1 quả mầu trắng", thì $\overline{C}=\overline{A_1}.\overline{A_2}$

$$P(\overline{C}) = P(\overline{A_1}).P\left(\overline{A_2}\big/\overline{A_1}\right) = \frac{4}{10}.\frac{3}{9} = \frac{2}{15}$$

Vậy : $P(C) = 1 - P(\overline{C}) = \dfrac{13}{15}$

Ví dụ 1.4.11. Trong một lớp có n học sinh, trong đó có m nữ $(m \le n)$ gọi liên tiếp (không gọi lại) 3 lần mỗi lần 1 học sinh. Tìm xác suất để cả 3 học sinh được gọi đều là nữ.

Giải: Gọi A_i là sự kiện "học sinh được gọi lần thứ i là nữ", (i=1,2,3)

A là sự kiện "cả 3 học sinh được gọi là nữ".

$$A = A_1A_2A_3$$

$$P(A) = P(A_1).P\left(A_2\big/A_1\right).P\left(A_3\big/A_1A_2\right)$$

$$= \frac{m}{n}.\frac{m-1}{n-1}.\frac{m-2}{n-2}$$

Ví dụ 1.4.12. Một người có 3 con gà mái, 2 con gà trống nhốt chung trong một lồng. Một người đến mua, người bán gà bắt ngẫu nhiên ra một con. Người mua chấp nhận mua con gà đó.

a/ Tìm xác suất để người đó mua được con gà mái.

b/ Người thứ hai đến mua, người bán gà lại bắt ngẫu nhiên ra một con. Tìm xác suất người thứ hai mua được gà trống.

c/ Xác suất này sẽ bằng bao nhiêu nếu người bán gà quên mất rằng con gà bán cho người thứ nhất là gà trống hay mái?

Giải: Gọi B_i là sự kiện "người thứ i mua được gà mái", (i =1, 2).

a/ $P(B_1) = 3/5 = 0,6$.

b/ Vì người thứ hai mua sau khi người thứ nhất đã mua xong cho nên:

$$P\left(\overline{B_2}\big/B_1\right) = \frac{2}{4} = 0,5.$$

c/ Ta có $\overline{B_2} = \left(B_1 \cup \overline{B_1}\right)\overline{B_2} = B_1\overline{B_2} \cup \overline{B_1}\,\overline{B_2}$.

Vậy $P(\overline{B_2}) = P(B_1)P\left(\overline{B_2}\big/B_1\right) + P(\overline{B_1})P\left(\overline{B_2}\big/\overline{B_1}\right) = \dfrac{3}{5}.\dfrac{2}{4} + \dfrac{2}{5}.\dfrac{1}{4} = 0,4$.

c. Sự kiện độc lập

Định nghĩa 1.4.13. Hai biến cố A và B gọi là *độc lập* với nhau nếu việc xảy ra hay không xảy ra của biến cố này không làm thay đổi xác suất xảy ra của biến cố kia. Trong trường hợp ngược lại thì 2 biến cố đó được gọi là phụ thuộc. Nếu A và B là độc lập với nhau thì:

+ $P\left(A\big/B\right) = P(A)$;

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

+ Các cặp $\left(\overline{A}, B\right); \left(A, \overline{B}\right), \left(\overline{A}, \overline{B}\right)$ cũng độc lập.

Tổng quát:

+ Các biến cố $A_1, A_2,...A_n$ gọi là độc lập từng đôi với nhau nếu mỗi cặp 2 trong n biến cố đó độc lập với nhau.

+ Các biến cố gọi là độc lập toàn phần với nhau nếu mỗi biến cố độc lập với tổ hợp bất kỳ các biến cố còn lại.

Ví dụ 1.4.14. Một phòng điều trị có 3 bệnh nhân bệnh nặng với xác suất cần cấp cứu trong cùng một giờ của các bệnh nhân tương ứng là 0,7 ; 0,8 và 0,9. Tìm xác suất sao cho trong cùng một giờ :

a/ Có 2 bệnh nhân cần cấp cứu.

b/ Có ít nhất 1 bệnh nhân không cần cấp cứu.

Giải: Gọi A_i là sự kiện "bệnh nhân thứ i cần cấp cứu", (i=1,2,3). Theo giả thiết ta có: $P(A_1) = 0{,}7; \quad P(A_2) = 0{,}8; \; P(A_3) = 0{,}9$

a/ Gọi A là sự kiện "có 2 bệnh nhân cần cấp cứu"

$A = A_1A_2\overline{A_3} + A_1\overline{A_2}A_3 + \overline{A_1}A_2A_3$ Vì A_i là độc lập nên

$$P(A) = P(A_1).P(A_2).P(\overline{A_3}) + P(A_1).P(\overline{A_2}).P(A_3) + P(\overline{A_1}).P(A_2).P(A_3)$$
$$= 0{,}7.0{,}8.0{,}1 + 0{,}7.0{,}2.0{,}9 + 0{,}3.0{,}8.0{,}9$$
$$= 0{,}398$$

b/ Gọi B là sự kiện "có ít nhất 1 bệnh nhân không cần cấp cứu", vậy $\overline{B}$ là sự kiện "không có bệnh nhân không cần cấp cứu".

$$\overline{B} = A_1.A_2.A_3 \qquad \Rightarrow \qquad P(\overline{B}) = 0{,}7.0{,}8.0{,}9$$
$$P(B) = 1 - P(\overline{B}) = 0{,}496$$

Ví dụ 1.4.15. Hai máy bay cùng ném bom một mục tiêu, mỗi máy bay ném 1 quả bom với xác suất trúng mục tiêu tương ứng là 0,7 và 0,8. Tìm xác suất để mục tiêu trúng bom.

Giải: Gọi A_i là sự kiện "máy bay thứ i bắn trúng mục tiêu", (i=1,2).

A là sự kiện "mục tiêu trúng bom"

$A = A_1 + A_2$. Vì $A_1; A_2$ là không xung khắc và độc lập nên

$$P(A) = P(A_1) + P(A_2) - P(A_1A_2)$$
$$= 0{,}7 + 0{,}8 - 0{,}7.0{,}8 = 0{,}94$$

Cách khác: Vì $A_1; A_2$ là không xung khắc và độc lập nên:

$$P(A) = 1 - P(\overline{A_1}.\overline{A_2}) = 1 - 0{,}3.0{,}2 = 0{,}94$$

1.4.3. Công thức xác suất toàn phần - Công thức Bayes

Bài toán 1.4.16. Giả sử trong một phép thử có một hệ đầy đủ các biến cố $A_1, A_2,...A_n$, B là một biến cố nào đó. B chỉ xảy ra khi và chỉ khi một trong các biến cố $A_i (i = \overline{1,n})$ của hệ trên xảy ra. Biết các xác suất $P(A_i)$ và $P\left(B/A_i\right)$. Hãy tìm $P(B)$.

Giải: Theo giả thiết B xảy ra khi và chỉ khi một trong các biến cố A_i của hệ trên xảy ra ta có: $B = A_1B \cup A_2B \cup .... \cup A_nB$. Vì A_i lập thành hệ đầy đủ các biến cố nên các biến cố A_i xung khắc từng đôi, do đó A_iB cũng xung khắc từng đôi và:

$$P(B) = P(A_1).P\left({}^{B}\!/\!_{A_1} \right) + P(A_2).P\left({}^{B}\!/\!_{A_2} \right) + \ldots + P(A_n).P\left({}^{B}\!/\!_{A_n} \right).$$

Hay: $P(B) = \sum_{i=1}^{n} P(A_i).P\left({}^{B}\!/\!_{A_i} \right)$.

Công thức này được gọi là *công thức xác suất toàn phần* (hay *công thức xác suất đầy đủ)*.

Cùng với giả thiết của bài toán trên nhưng thêm điều kiện là phép thử đã được thực hiện và sự kiện B đã xảy ra và do vậy phải có duy nhất một sự kiện A_k nào đó xảy ra từ đó ta có:

$$P\left({}^{A_k}\!/\!_{B} \right) = \frac{P(A_k).P\left({}^{B}\!/\!_{A_k} \right)}{\sum_{i=1}^{n} P(A_i).P\left({}^{B}\!/\!_{A_i} \right)}$$. Đây là *công thức Bayes*.

Chú ý 1.4.17. Công thức xác suất đầy đủ thường được gọi là công thức xác suất tiên nghiệm. Công thức Bayes được xác định sau khi phép thử đã tiến hành và biến cố B đã xảy ra do đó được gọi là công thức xác suất hậu nghiệm.

Ví dụ 1.4.17. Một trại lợn nhận lợn giống từ 3 cơ sở theo tỷ lệ 25%; 35% và 40%. Biết tỷ lệ lợn giống không đủ tiêu chuẩn ở mỗi cơ sở lần lượt là: 5%; 4% và 2%. Bắt ngẫu nhiên một con lợn của trại.

a/ Tìm xác suất để ta bắt được con lợn đủ tiêu chuẩn.

b/ Giả sử ta bắt được con lợn không đủ tiêu chuẩn. Theo anh (chị) con lợn đó có khả năng thuộc cơ sở nào nhất.

Giải: Gọi A_i là sự kiện "bắt được lợn giống của cơ sở thứ i", (i = 1..3)

A là sự kiện "bắt được lợn đủ tiêu chuẩn"

$A = AA_1 + AA_2 + AA_3$

a/ $P(A) = P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + P(A_3)P(A/A_3) = 0{,}9655$

b/ $P(\overline{A}) = 1 - P(A) = 0{,}0345$

Theo Bayes: $P(A_1/\overline{A}) = (0{,}25.0{,}05)/0{,}0345 = 0{,}362$

$P(A_2/\overline{A}) = (0{,}35.0{,}04)/0{,}0345 = 0{,}406$

$P(A_3/\overline{A}) = (0{,}4.0{,}02)/0{,}0345 = 0{,}229$

Vậy con lợn không đủ tiêu chuẩn đó khả năng thuộc cơ sở 2 cung cấp

Ví dụ 1.4.18. Một thiết bị gồm ba loại linh kiện: Loại I chiếm 35%, loại II chiếm 25%, loại III chiếm 40% tổng số linh kiện của toàn thiết bị. Xác suất hư hỏng sau khoảng thời gian làm việc nào đó của các loại linh kiện tương ứng là: 15%; 25% và

5%. Máy đang hoạt động bỗng bị hỏng, Hãy tính xác suất để từng loại linh kiện bị hỏng (giả thiết các loại linh kiện không cùng hỏng đồng thời).
Giải: Gọi A_i là sự kiện "kiểm tra linh kiện loại i", (i = 1..3)
A là sự kiện "máy bị hỏng".

$$A = AA_1 + AA_2 + AA_3$$

$$P(A) = P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + P(A_3)P(A/A_3)$$

$$= 0{,}35.0{,}15 + 0{,}25.0{,}25 + 0{,}4.0{,}05 = 0{,}135$$

Theo Bayes: $P(A_1/A) = (0{,}35.0{,}15)/0{,}135 = 0{,}389$

$$P(A_2/A) = (0{,}25.0{,}25)/0135 = 0{,}463$$

$$P(A_3/A) = (0{,}4.0{,}05)/0{,}135 = 0{,}148$$

1.4.4. Công thức Bernoulli

a. Dãy phép thử Bernoulli : Xét một dãy n phép thử độc lập, mỗi phép thử chỉ có 2 biến cố A hoặc $\overline{A}$ với $P(A) = p; P(\overline{A}) = 1 - p = q$ không đổi, không phụ thuộc vào thứ tự phép thử. Dãy phép thử đó gọi là *dãy phép thử Bernoulli*.

b. Công thức Bernoulli: Với dãy phép thử Bernoulli ta có bài toán.

Bài toán 1.4.19. Tìm xác suất sao cho trong n phép thử đó sự kiện A xuất hiện đúng k lần không phân biệt thứ tự .

Bài toán 1.4.20. Tìm xác suất sao cho trong n phép thử đó sự kiện A xuất hiện từ k_1 đến k_2 lần không phân biệt thứ tự.

Giải: Gọi B là biến cố trong n phép thử Bernoulli đó biến cố A xuất hiện đúng k lần không phân biệt thứ tự. Ta thấy biến cố B có thể xuất hiện theo nhiều cách khác nhau. Giả sử ta xét một trường hợp của B là biến cố A xuất hiện ở k phép thử đầu còn (n-k) phép thử tiếp theo xuất hiện $\overline{A}$. Vậy

$$P(A\ldots A\overline{A}\ldots\overline{A}) = P(A\ldots A).P(\overline{A}\ldots\overline{A})$$
$$= p^k.q^{n-k}$$

Ta có C_n^k cách để A xuất hiện với cùng một xác suất như trên.

Vậy: $P_n(k) = P(B) = C_n^k.p^k.q^{n-k}$. Đây là công thức Bernoulli.

Hoàn toàn tương tự như trên Gọi H là biến cố trong n phép thử Bernoulli biến cố A xuất hiện từ k_1 đến k_2 lần không phân biệt thứ tự.

$$H = \bigcup_{k=k_1}^{k_2} B_k.$$

Vì các B_k xung khắc từng đôi một do đó:

$$P_n(k_1, k_2) = P(H) = \sum_{k=k_1}^{k_2} C_n^k.p^k q^{n-k} .$$

Ví dụ 1.4.21. Xác suất tiêu thụ điện năng trong mỗi ngày không vượt quá quy định tại một xí nghiệp là p = 0,75. Tính xác suất sao cho trong 6 ngày liên tiếp có 4 ngày lượng điện năng tiêu thụ không vượt quá mức quy định.
Giải: Bài toán thoả mãn dãy phép thử Bernoulli
Ta có n = 6, k = 4, p = 0,75, q = 0,25
Vậy: $P_6(4) = C_6^4(0,75)^4.(0,25)^2 = 0,3$

Ví dụ 1.4.22. Bắn liên tiếp 5 viên đạn vào một mục tiêu xác suất trúng đích của mỗi viên là 0,2 .

a/ Tính xác suất để có đúng 2 viên trúng;

b/ Để mục tiêu bị tiêu diệt cần phải có từ 3 viên trúng mục tiêu trở lên. Tính xác suất để mục tiêu bị tiêu diệt.

Giải: Bài toán thoả mãn dãy phép thử Bernoulli, do đó

a/ $P_5(2) = C_5^2 P^2 q^3 = C_5^2.(0,2)^2.(0,8)^3 = 0,2048$

b/ $p_5(3,5) = C_5^3 p^3 q^2 + C_5^4 p^4 q + C_5^5 p^5 = 0,0575$

Ví dụ 1.4.23. Một bác sĩ chữa bệnh có xác suất chữa khỏi là 0,8 có người nói rằng cứ 5 người đến chữa thì có chắc chắn 4 người khỏi bệnh, người khác lại cho rằng trong 10 người đến chữa có chắc chắn 8 người khỏi bệnh. Điều đó có đúng không.
Giải: Cả 2 người khẳng định đều sai, vì bài toán trên thoả mãn dãy phép thử Bernoulli nên xác suất xảy ra trong các trường hợp là:

$$P_5(4) = C_5^4.(0,8)^4.0,2 = 0,4096$$
$$P_{10}(8) = C_{10}^8.(0,8)^8.(0,2)^2 = 0,3018$$

c. Giá trị khả năng nhất

Trị số của $P_n(k)$ nói chung phụ thuộc vào k (với n cố định), ta cần tìm trị số $k = k_0$ sao cho $P_n(k)$ đạt giá trị lớn nhất. Số k_0 được gọi là số lần xuất hiện chắc chắn nhất (có khả năng nhất) của biến cố A trong n phép thử đã cho. Qua khảo sát $P_n(k)$ ta có kết quả sau:

+ Nếu (np + p - 1) là số nguyên thì $P_n(k)$ đạt cực đại tại 2 giá trị $k_0 = np + p - 1$ và $k_0 = np + p$

+ Nếu (np + p - 1) là số không nguyên thì $P_n(k)$ đạt giá trị lớn nhất tại $k_0 = [np + p - 1] + 1$

Ví dụ 1.4.24. Tỷ lệ mắc một loại bệnh A ở một vùng là 10%. Trong đợt khám bệnh cho vùng đó người ta đã khám 100 người. Tìm xác suất để trong 100 người có:

a/ 6 người bị bệnh A;

b/ 95 người không bị bệnh A;

c/ Ít nhất 1 người bị bệnh A;

d/ Tìm số người bị bệnh A có khả năng nhất? Tính xác suất tương ứng.

Giải: Bài toán thỏa mãn dãy phép thử Bernoulli với P(A) = 0,10 do đó:

a/ $P_{100}(6) = C_{100}^6.0,1^6.0,9^{94}$;

b/ $P_{100}(95) = C_{100}^{95}.0,9^{95}.0,1^5$;

c/ $P_{100}(k \geq 1) = 1 - P_{100}(k - 0) = 1 - C_{100}^{0}.0{,}1^{0}.0{,}9^{100} = 1 - 0{,}9^{100}$

d/ Theo bài ra ta có np + p - 1 = 100.0,1 + 0,1 - 1 = 9,1. Vậy số người bị bệnh A có khả năng nhất khi khám 100 người là 10 người và

$$P_{100}(10) = C_{100}^{10}.0{,}1^{10}.0{,}9^{90}$$

BÀI TẬP CHƯƠNG 1

Dạng 1: Công thức xác suất cổ điển

1. Thang máy của một tòa 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.

2. Xếp ngẫu nhiên 4 khách lên 9 toa tầu hỏa. Tìm xác suất để:

a/ 4 người lên toa đầu.

b/ 4 người lên cùng một toa.

c/ 4 người lên 4 toa khác nhau.

3. Có 2 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 ra 1 sản phẩm. Tính xác suất để:

a/ Lấy được 1 chính phẩm;

b/ Lấy được ít nhất 1 chính phẩm.

c/ Lấy được 2 chính phẩm.

4. Có hai chuồng lợn giống, chuồng 1 có 7 con cái và 3 con đực, chuồng 2 có 6 con cái và 4 con đực. Bắt ngẫu nhiên từ mỗi chuồng ra một con. Tính xác suất để:

a/ Cả 2 con bắt ra đều là con cái.

b/ Bắt được một con cái và một con đực.

c/ Bắt được ít nhất một con đực.

5. Một kĩ sư nông nghiệp có hai hộp hạt giống cùng loại: Hộp 1 có 12 hạt giống trong đó 8 hạt đủ tiêu chuẩn, hộp 2 có 12 hạt giống trong đó có 9 hạt đủ tiêu chuẩn. Chọn ngẫu nhiên từ mỗi hộp ra 1 hạt giống. Tìm xác suất để trong hai hạt lấy ra:

a/ Có một hạt đủ tiêu chuẩn, một hạt không đủ tiêu chuẩn.

b/ Lấy được ít nhất 1 hạt đủ tiêu chuẩn.

c/ Lấy được 2 hạt đủ tiêu chuẩn.

6. Trong một hòm đựng 8 chi tiết là chính phẩm và 5 chi tiết là phế phẩm. Lấy đồng thời ra 3 chi tiết. Tính xác suất để:

a/ Cả 3 chi tiết lấy ra là chính phẩm.

b/ Trong 3 chi tiết lấy ra có 2 chính phẩm.

c/ Trong 3 chi tiết lấy ra có ít nhất 1 chính phẩm.

7. Trong một lớp học có 15 học sinh nam và 10 học sinh nữ. Gọi ngẫu nhiên 4 học sinh lên bảng làm bài tập. Tính xác suất để:

a/ có 2 học sinh nam.

b/ Có ít nhất 2 học sinh nam.

c/ Có cả nam và nữ.

8. Một hộp đựng 7 quả cầu trắng và 8 quả cầu đen cùng kích cỡ. Lấy ngẫu nhiên ra 4 quả cầu. Tìm xác suất để:

a/ Trong 4 quả lấy ra có 3 quả trắng?

b/ Có 4 quả cùng mầu?

c/ Có ít nhất 1 quả mầu đen?

9. Trong một hộp bút có 10 chiếc bút bi cùng kích cỡ, trong đó có 6 chiếc bút mực đen và 4 chiếc bút mực xanh. Lấy ngẫu nhiên ra 3 chiếc bút. Tìm xác suất trong 3 chiếc lấy ra có:

a/ 2 chiếc bút mực xanh?

b/ ít nhất 2 chiếc bút mực xanh:

c/ 2 chiếc cùng mầu:

10. Một chiếc hộp đựng 6 quả cầu trắng, 4 quả cầu đỏ và 2 quả cầu đen. Chọn ngẫu nhiên ra 6 quả cầu. Tìm xác suất trong 6 quả lấy ra có:

a/ 3 quả trắng, 2 quả đỏ và 1 quả đen?

b/ 4 quả đỏ?

c/ Không có quả nào mầu trắng?

Dạng 2: Công thức xác suất tổng, công thức xác suất đầy đủ, Bayss, Bernouly

11. Một nhà máy sản xuất bóng đèn. Máy A sản xuất 25% số bóng đèn ,máy B sản xuất 35% số bóng đèn,còn máy C sản xuất 40% số bóng đèn.Tỉ lệ sản phẩm hỏng của các máy tương ứng là 5% (máy A),4% (máy B) và 2% (máy C).

a/ Lấy ngẫu nhiên một bóng đèn.Tìm xác suất để gặp bóng đèn xấu.

b/ Khi lấy ngẫu nhiên một bóng đèn ta được bóng đèn tốt. Tìm xác suất để bóng tốt lấy được đó do máy B sản xuất.

12. Một dự án trồng cây lâm nghiệp nhận giống cây trồng từ 3 cơ sở sản xuất giống cây trồng. Trung bình cơ sở 1 cung cấp 35%, cơ sở 2 cung cấp 40%, cơ sở 3 cung cấp 25% tổng số giống cây trồng của dự án. Trong đó khoảng 90% cây giống do cơ sở 1 cung cấp là đủ tiêu chuẩn, 85% cây giống do cơ sở 2 cung cấp là đủ tiêu chuẩn, 80% cây giống do cơ sở 3 cung cấp là đủ tiêu chuẩn. Lấy ngẫu nhiên một cây trồng của dự án để kiểm tra.

a/ Tính xác suất để cây trồng lấy ra đủ tiêu chuẩn.

b/ Giả sử cây lấy ra đủ tiêu chuẩn, theo anh (chị) cây đó có khả năng do cơ sở nào cung cấp.

13. Một trại lợn nhận lợn giống từ 3 cơ sở theo tỷ lệ 20%; 35% và 45%. Biết tỷ lệ lợn giống không đủ tiêu chuẩn ở mỗi cơ sở lần lượt là 2%; 3% và 4%. Bắt ngẫu nhiên một con lợn của trại.

a/ Tìm xác suất để bắt được con lợn đủ tiêu chuẩn.

b/ Giả sử bắt được con lợn không đủ tiêu chuẩn. Theo bạn con lợn đó có khả năng thuộc cơ sở nào nhất?

14. Trong một bệnh viện, tỷ lệ bệnh nhân các tỉnh như sau: Tỉnh A: 25%, tỉnh B: 35% và tỉnh C: 40%. Biết tỷ lệ bệnh nhân là kỹ sư của các tỉnh tương ứng là $2{,}5\%$; 3% và $4{,}5\%$. Chọn ngẫu nhiên một bệnh nhân.

a/ Tính xác suất để bệnh nhân đó là kỹ sư.

b/ Giả sử bệnh nhân được chọn không phải là kỹ sư. Theo bạn bệnh nhân đó có khả năng thuộc tỉnh nào nhất?

15. Có 3 cửa hàng I, II và III cùng kinh doanh sản phẩm Y. Tỷ lệ sản phẩm loại A trong 3 của hàng I, II, III lần lượt là 70%, 75% và 50%. Một khách hàng chọn ngẫu nhiên một cửa hàng và từ đó mua một sản phẩm.

a/ Tính xác suất để khách hàng đó mua được sản phẩm loại A.

b/ Giả sử khách hàng đã mua được sản phẩm loại A, theo bạn sản phẩm đó có khả năng thuộc cửa hàng nào?

16. Một cửa hàng bán máy tính với 40% máy tính của hãng IBM, 60% máy tính của hãng Acer. Biết rằng tỷ lệ máy sản xuất tại chính hãng IBM và Acer lần lượt là 0,8; 0,9. Một khách hàng mua máy tính tại cửa hàng.

a/ Tính xác suất để khách hàng mua được máy tính sản xuất tại chính hãng.

b/ Giả sử khách hàng mua được máy tính sản xuất tại chính hãng, theo bạn máy tính đó có khả năng do hãng nào sản xuất?

17. Có 20 kiện hàng mỗi kiện hàng có 10 sản phẩm. Trong số đó có 8 kiện loại 1, mỗi kiện hàng có 1 phế phẩm; 7 kiện hàng loại 2, mỗi kiện hàng có 2 phế phẩm và 5 kiện hàng loại 3, mỗi kiện có 3 phế phẩm. Lấy ngẫu nhiên một kiện hàng, rồi từ đó lấy ngẫu nhiên một sản phẩm.

a/ Tính xác suất để sản phẩm lấy ra là phế phẩm.

b/ Nếu lấy được sản phẩm là phế phẩm, theo bạn sản phẩm đó có khả năng thuộc kiện hàng loại nào nhiều hơn cả?

18. Trong một lớp học, tỷ lệ học sinh thích chơi game là 70%. Biết rằng nếu ham chơi game thì tỷ lệ học sinh đạt học lực khá là 30%, còn nếu không chơi game thì tỷ lệ học sinh đạt học lực khá là 60%. Gọi một học sinh lên bảng.

a/ Tính xác suất để học sinh đó có học lực khá.

b/ Giả sử học sinh đó có học lực khá. Tính xác suất để học sinh đó chơi game.

19. Ở một vùng dân cư cứ 100 người có 20 người hút thuốc lá. Biết rằng tỷ lệ người viêm họng trong số người hút thuốc lá là 65%, còn trong số người không hút thuốc là 35%. Khám ngẫu nhiên một người thì thấy anh ta viêm họng, tìm xác suất để người đó hút thuốc. Nếu người đó không viêm họng thì xác suất để người đó không hút thuốc là bao nhiêu.

20. Có 2 hộp như nhau đựng các mẫu hàng xuất khẩu. Hộp thứ nhất có 10 mẫu trong đó có 6 mẫu loại A và 4 mẫu loại B. Hộp thứ 2 có 10 mẫu trong đó có 3 mẫu loại A và 7 mẫu loại B. Chọn ngẫu nhiên 1 hộp và từ đó lấy ngẫu nhiên 1 mẫu.

a/ Tính xác suất để mẫu lấy ra là loại B.

b/ Giả sử mẫu lấy ra loại A. Hỏi mẫu đó có khả năng thuộc hộp loại nào nhiều hơn?

21. Trong 1 bệnh viện bỏng: 80% bệnh nhân bị bỏng do nóng, 20% bệnh nhân bị bỏng do hóa chất. Trong số những bệnh nhân bị bỏng do nóng thì có 30% bị biến chứng, còn với bỏng do hóa chất thì có 60% bị biến chứng. Từ tập bệnh án rút ngẫu nhiên ra 1 hồ sơ thấy đó là của bệnh nhân bị biến chứng. Tìm xác suất để bệnh nhân đó bị bỏng do hóa chất gây ra?

22. Có 20 hộp sản phẩm cùng loại, trong đó có 10 hộp của xí nghiệp I, 6 hộp của xí nghiệp II, 4 hộp của xí nghiệp III. Tỷ lệ sản phẩm tốt của các xí nghiệp tương ứng lần lượt là 50%, 65% và 75%. Lấy ngẫu nhiên ra một hộp và chọn ngẫu nhiên ra một sản phẩm.

a/ Tính xác suất để sản phẩm đó là tốt.

b/ Nếu sản phẩm đó là tốt, theo bạn sản phẩm đó có khả năng thuộc xí nghiệp nào là nhiều hơn cả?

23. Có 18 học sinh thi học sinh giỏi chia làm 4 nhóm: nhóm I có 5 học sinh, nhóm II có 7 học sinh, nhóm III có 4 học sinh và nhóm IV có 2 học sinh. Xác suất để một học sinh trong nhóm đạt giải tương ứng lần lượt là 0,8; 0,7; 0,6; 0,5.

a/ Tính xác suất để một học sinh bất kỳ đạt giải.

b/ Nếu học sinh đó đạt giải hãy tính xác suất để học sinh đó thuộc nhóm I?

24. Trong một làng tỷ lệ nam là 60% và nữ là 40%. Khả năng mắc bệnh bạch tạng ở nam là 0,6% và ở nữ là 0,35%. Gặp một người trong làng thấy người đó mắc bệnh. Tìm xác suất để người đó là nam? Nếu người đó không mắc bệnh xác suất để người đó là nam là bao nhiêu?

25. Hai máy cùng sản xuất một loại sản phẩm. Tỉ lệ phế phẩm của máy I là 3% của máy II là 2%.Từ một kho gồm 2/3 sản phẩm của máy I và 1/3 sản phẩm của máy II ta lấy một sản phẩm.Tính xác suất để:

a/ Sản phẩm lấy ra là tốt.

b/ Giả sử sản phẩm lấy ra là sản phẩm tốt. Tính xác suất để sản phẩm lấy ra là của máy I sản suất.

26. Có 10 sinh viên đi thi, trong đó có 3 sinh viên thuộc loại giỏi, 4 khá và 3 trung bình. Trong số 20 câu hỏi thi qui định thì sinh viên loại giỏi trả lời được tất cả, sinh viên khá trả lời được 16 câu, còn sinh viên trung bình chỉ trả lời được 10 câu. Gọi ngẫu nhiên 1 sinh viên và phát 1 phiếu thi có 4 câu hỏi thì anh ta trả lời được cả 4 câu hỏi. Tính xác suất để sinh viên đó thuộc loại khá.

Chương 2

Đại lượng ngẫu nhiên – Quy luật phân phối xác suất

2.1. Đại lượng ngẫu nhiên

2.1.1. Định nghĩa*:*

Định nghĩa 2.1.1. *Đại lượng ngẫu nhiên (hay 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 của nó với một xác suất tương ứng xác định.

Người ta thường dùng các chữ cái in X, Y, Z,.. hoặc X_1, X_2... để chỉ đại lượng ngẫu nhiên và các chữ cái thường x, y, z,...hoặc x_1, x_2,...để chỉ các giá trị có thể có của nó.

Đại lượng ngẫu nhiên được phân làm 2 loại: Đại lượng ngẫu nhiên rời rạc và đại lượng ngẫu nhiên liên tục.

\+ Đại lượng ngẫu nhiên rời rạc là đại lượng ngẫu nhiên mà các giá trị có thể có của nó lập nên một tập hợp hữu hạn hoặc vô hạn đếm được phần tử.

Ví dụ 2.1.2. Gọi X là số viên đạn bắn trúng bia khi bắn 3 viên, khi đó X là đại lượng ngẫu nhiên rời rạc nhận một trong các giá trị 0, 1, 2, 3.

\+ Đại lượng ngẫu nhiên gọi là liên tục nếu các giá trị có thể có của nó lấp đầy một khoảng trên trục số.

Ví dụ 2.1.3. Gọi X là khoảng cách từ điểm chạm của viên đạn đến tâm bia, thì X là đại lượng ngẫu nhiên liên tục.

Chú ý 2.1.4. Có thể nói rằng gần như tất cả các đại lượng chỉ về trọng lượng, độ dài đều là đại lượng ngẫu nhiên liên tục.

2.1.2. Các quy luật phân phối xác suất của đại lượng ngẫu nhiên

a. Bảng phân phối xác suất của đại lượng ngẫu nhiên

Người ta thường biểu thị qui luật phân phối xác suất của đại lượng ngẫu nhiên rời rạc dưới dạng bảng gồm 2 dòng: Dòng 1 ghi các giá trị mà đại lượng ngẫu nhiên có thể nhận được: x_1, x_2,...., x_n. Dòng 2 ghi các giá trị xác suất tương ứng.

X	x_1	x_2 ....x_i...........x_n
p	p_1	p_2. ....p_i..........p_n

với $\sum_{i=1}^{n} p_i = 1$.

Ví dụ 2.1.5. Một tổ sản xuất có 3 mô tơ chạy độc lập với nhau, với xác suất để mỗi mô tơ chạy tốt trong ngày là 0,7. Gọi X là số mô tơ chạy tốt trong ngày. Lập bảng phân phối xác suất của X.

Giải: X là đại lượng ngẫu nhiên rời rạc với các giá trị có thể có 0, 1, 2, 3. với xác suất tương ứng được tính theo công thức Bernoulli

$$P(X = i) = p_3(i) = C_3^i p^i q^{3-i}\text{, với p = 0,7; q = 0,3.}$$

Ta có bảng phân phối xác suất của đại lượng ngẫu nhiên X là:

X	0	1	2	3
P	0,027	0,189	0,441	0,343

Ví dụ 2.1.6. Một xạ thủ có 3 viên đạn được yêu cầu bắn lần lượt từng viên cho đến khi trúng mục tiêu hoặc hết cả 3 viên thì thôi. Tìm bảng phân phối xác suất của số đạn đã bắn, biết rằng xác suất bắn trúng đích của mỗi lần là 0,8.

Giải: Gọi X là số đạn đã dùng, ta có X là đại lượng ngẫu nhiên rời rạc gồm 3 giá trị là 1, 2, 3 với

P(X = 1) = 0,8; P(X = 2) = 0,2.0,8 = 0,16; P(X =3) = 0,2.0,2 = 0,04.

Vậy

X	1	2	3
P	0,8	0,16	0,04

b. Hàm phân phối xác suất

Định nghĩa 2.1.7. *Hàm phân phối xác suất* của đại lượng X ngẫu nhiên, ký hiệu là F(x), là xác suất để đại lượng ngẫu nhiên nhận giá trị nhỏ hơn x, với x là một số thực bất kỳ.

$$F(x) = P(X < x).$$

Chú ý 2.1.8. Nếu X là đại lượng ngẫu nhiên rời rạc thì $F(x) = \sum_{x_i < x} p_i$.

Ví dụ 2.1.9. Bắn liên tiếp 3 viên đạn vào bia, với xác suất trúng đích của mỗi viên là 0,4. Gọi X là số viên đạn trúng bia. Hãy tìm hàm phân phối xác suất của X.

Giải: X là đại lượng ngẫu nhiên rời rạc nhận 4 giá trị 0, 1, 2, 3, với xác suất được tính theo công thức Bernoulli

$$P_3(i) = C_3^i p^i q^{3-i}. \quad \text{(i = 1..3)}$$

Ta có bảng phân phối xác suất:

X	0	1	2	3
P	0,216	0,432	0,288	0,064

+ Nếu $x \le 0$, biến cố $(X < x)$ là biến cố không thể có, do đó $F(x) = 0$;

+ Nếu $0 < x \le 1$, biến cố $(X < x)$ sẽ xảy ra khi X = 0, do đó $F(x) = 0{,}216$;

+ Nếu $1 < x \le 2$, biến cố $(X < x)$ sẽ xảy ra khi X = 0 hoặc X = 1,

do đó $F(x) = 0{,}216 + 0{,}432 = 0{,}648$;

+ Nếu $2 < x \le 3$ biến cố $(X < x)$ sẽ xảy ra khi X = 0 hoặc X = 1, hoặc X = 2, do đó $F(x) = 0{,}936$;

+ Nếu x > 3 biến cố $(X < x)$ sẽ xẩy ra khi X = 0 ; X = 1; X = 2, hoặc X = 3, do đó $F(x) = 1$.

Vậy hàm phân phối xác suất của X là:

$$F(x) = \begin{cases} 0 & x \le 0 \\ 0{,}216 & 0 < x \le 1 \\ 0{,}648 & 1 < x \le 2 \\ 0{,}936 & 2 < x \le 3 \\ 1 & x > 3 \end{cases}$$

Ví dụ 2.1.10. Có 2 lồng nhốt gà: Lồng thứ nhất có 6 gà mái và 2 gà trống, lồng thứ 2 có 5 gà mái và 3 gà trống. Từ lồng thứ nhất bắt 2 gà bỏ sang lồng 2, rồi từ lồng 2 bắt ngẫu nhiên ra 2 con.

a/ Tìm qui luật phân phối xác suất chỉ số gà mái được bắt ra.

b/ Tìm hàm phân phối xác suất chỉ số gà mái được bắt ra.

Giải: Gọi A_i là sự kiện bắt được i gà mái từ lồng 1 vào lồng 2 (i = 0..2)

X là số gà mái có trong 2 con được bắt từ lồng 2: X = 0; 1; 2

X = A_0X + A_1X + A_2X

$$P(X) = P(A_0)P(X/A_0) + P(A_1)P(X/A_1) + P(A_2)P(X/A_2)$$

$$P(X=0) = \frac{C_2^2}{C_8^2}.\frac{C_5^2}{C_{10}^2} + \frac{C_6^1C_2^1}{C_8^2}.\frac{C_4^2}{C_{10}^2} + \frac{C_6^2}{C_8^2}.\frac{C_3^2}{C_{10}^2} = \frac{127}{1260} \approx 0{,}1008$$

Tương tự: $P(X=1) = \dfrac{628}{1260} \approx 0{,}4984$ $\quad P(X=2) = \dfrac{505}{1260} \approx 0{,}4008$

a/ Qui luật phân phối xác suất của X:

X	0	1	2
P	0,1008	0,4984	0,4008

b/ Hàm phân phối xác suất của X: $F(X) = \begin{cases} 0 & Khi \quad x \le 0 \\ 0{,}1008 & Khi \quad 0 < x \le 1 \\ 0{,}5992 & 1 < x \le 2 \\ 1 & x > 2 \end{cases}$

Các tính chất của hàm phân phối xác suất

Tính chất 2.1.11.

(a) $0 \le F(x) \le 1$;

(b) Hàm phân phối xác suất là hàm không giảm tức là với $x_2 > x_1$ thì $F(x_2) \ge F(x_1)$.

(c) $F(-\infty) = 0$; $F(+\infty) = 1$

Hệ quả 2.1.12.

(a) $P(a \le X < b) = F(b) - F(a)$

(b) Xác suất để đại lượng ngẫu nhiên liên tục X nhận một giá trị xác định bằng 0: $P(X = x) = 0$

(c) Đối với đại lượng ngẫu nhiên liên tục X ta có các đẳng thức: $P(a \le X \le b) = P(a \le X < b) = P(a < X \le b) = P(a < X < b)$

(d) Nếu đại lượng ngẫu nhiên X chỉ nhận giá trị trong $[a,b]$ thì với $x \le a, F(x) = 0$ và với $x > b, F(x) = 1$.

Ý nghĩa của hàm phân phối xác suất: Hàm phân phối xác suất phản ánh mức độ tập trung xác suất ở về phía bên trái một số thực x nào đó.

c. Hàm mật độ xác suất

Định nghĩa 2.1.13. *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.

$$f(x) = F'(x)$$

Các tính chất của hàm mật độ xác suất

Tính chất 2.1.14.

(a) $f(x) \ge 0, \quad \forall x$

(b) $P(a < X < b) = \int\limits_a^b f(x)dx$

(c) $F(x) = \int\limits_{-\infty}^{x} f(x)dx$

(d) $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$

Ví dụ 2.1.15. Đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất như sau:

$$f(x) = \begin{cases} a\cos x & x \in \left[-\frac{\pi}{2};\frac{\pi}{2}\right] \\ 0 & x \notin \left[-\frac{\pi}{2};\frac{\pi}{2}\right] \end{cases}$$

a/ Tìm hệ số a.

b/ Tìm hàm phân phối xác suất $F(x)$.

c/ Tìm $P(0 < X < \frac{\pi}{4})$

Giải:

a/ Theo tính chất của hàm mật độ xác suất ta có $a \ge 0$ và

$$\int\limits_{-\infty}^{+\infty} f(x)dx = \int\limits_{-\pi/2}^{\pi/2} a\cos x dx = 1 \Leftrightarrow a = \frac{1}{2}$$

b/ Để tìm $F(x)$ ta áp dụng tính chất: $F(x) = \int\limits_{-\infty}^{x} f(x)dx$.

+ Với $x < -\frac{\pi}{2}$; $F(x) = \int_{-\infty}^{x} 0dx = 0$

+ Với $-\frac{\pi}{2} < x < \frac{\pi}{2}$, ta có:

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{-\pi/2} 0dx + \int_{-\pi/2}^{x} \frac{1}{2}\cos x dx = \frac{1}{2}(\sin x + 1).$$

+ Với $x > \frac{\pi}{2}$ ta có

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{-\pi/2} 0dx + \int_{-\pi/2}^{\pi/2} \frac{1}{2}\cos x dx + \int_{\pi/2}^{x} 0dx = 1.$$

Vậy: $F(x) = \begin{cases} 0 & x < -\frac{\pi}{2} \\ \frac{1}{2}(\sin x + 1) & -\frac{\pi}{2} \le x \le \frac{\pi}{2} \\ 1 & x > \frac{\pi}{2} \end{cases}$

c/ $P(0 < X < \frac{\pi}{4}) = \int_{0}^{\pi/4} \frac{1}{2}\cos x dx = \frac{\sqrt{2}}{4}$.

Ví dụ 2.1.16. Cho hàm phân phối của biến ngẫu nhiên X có dạng:

$$F(x) = a + b\,arctgx \quad (-\infty < x < +\infty)$$

Tìm a,b và $f(x)$.

Giải: Áp dụng tính chất 3 của hàm phân phối xác suất

$$\begin{cases} \lim\limits_{x\to-\infty} (a + b\,arctgx) = 0 \\ \lim\limits_{x\to+\infty} (a + b\,arctgx) = 1 \end{cases} \Leftrightarrow \begin{cases} a - \frac{b\pi}{2} = 0 \\ a + \frac{b\pi}{2} = 1 \end{cases} \Rightarrow \quad a = \frac{1}{2}; \quad b = \frac{1}{\pi}$$

Vậy $F(x) = \frac{1}{2} + \frac{1}{\pi} arctgx$ và do đó $f(x) = F'(x) = \frac{1}{\pi(1 + x^2)}$.

Ví dụ 2.1.17. Đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x) = \frac{A}{e^x + e^{-x}}$$

a/ Hãy xác định hệ số A.

b/ Tìm hàm phân phối xác suất F(x).

c/ Tìm xác suất để trong 3 phép thử độc lập có 1 lần X nhận giá trị trong khoảng (-1; 1)

Giải:

a/ Áp dụng tính chất $\begin{cases} f(x) \geq 0 \\ \int_{-\infty}^{+\infty} f(x)dx = 1 \end{cases} \Leftrightarrow \int_{-\infty}^{+\infty} \frac{A}{e^x + e^{-x}} dx = 1 \Leftrightarrow A = \frac{2}{\pi}$

b/ $F(x) = \int_{-\infty}^{x} \frac{2/\pi}{e^x + e^{-x}} dx = \frac{2}{\pi} \lim_{b \to -\infty} \int_{b}^{x} \frac{de^x}{e^{2x} + 1} = \frac{2}{\pi} arctge^x$

c/ $P(-1 < X < 1) = F(1) - F(-1) = \frac{2}{\pi}(arctge - arctge^{-1}) = p$

$P_3(1) = C_3^1 p(1-p)^2$

Ý nghĩa của hàm mật độ xác suất: Hàm mật độ xác suất của đại lượng ngẫu nhiên X tại mỗi điểm x cho biết mức độ tập trung xác suất tại điểm đó.

2.2. Các tham số đặc trưng của đại lượng ngẫu nhiên

2.2.1. Kỳ vọng toán

a. Các định nghĩa

Định nghĩa 2.2.1. X là đại lượng ngẫu nhiên rời rạc nhận một trong các giá trị có thể có $x_1, x_2, ..., x_n$ với các xác suất tương ứng $p_1, p_2, ..., p_n$. *Kỳ vọng toán* của đại lượng ngẫu nhiên rời rạc X, ký hiệu E(X) được xác định $E(X) = \sum_{i=1}^{n} x_i p_i$

Định nghĩa 2.2.2. Nếu X là đại lượng ngẫu nhiên liên tục với hàm mật độ xác suất $f(x)$ và miền giá trị là $[a,b]$ thì $E(X) = \int_a^b x.f(x)dx$.

Ví dụ 2.2.3. Tìm kỳ vọng toán của đại lượng ngẫu nhiên rời rạc X có bảng phân phối xác suất như sau:

X	1	2	3
P	0,8	0,16	0,04

Giải:

$$E(X) = 1.0,8 + 2.0,16 + 3.0,04 = 1,24$$

Ví dụ 2.2.4. Tìm kỳ vọng toán của đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất như sau:

$$f(x) = \begin{cases} \frac{3}{4}\left(x^2 + 2x\right) & x \in (0,1) \\ 0 & x \notin (0,1) \end{cases}$$

Giải: $E(x) = \int_0^1 \frac{3}{4}\left(x^2 + 2x\right)xdx = \frac{11}{16}$

Các tính chất của kỳ vọng toán

Tính chất 2.2.5.

(a) $E(C) = C \quad (c = const)$.

(b) $E(CX) = CE(X) \quad (C = const)$

(c) $E(X+Y) = E(X) + E(Y)$

(d) $E(X.Y) = E(X).E(Y)$ Nếu X và Y độc lập.

Bản chất và ý nghĩa của kỳ vọng toán

Giả sử đối với đại lượng ngẫu nhiên X tiến hành n phép thử 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(với $\sum_{i=1}^{k} x_i = n$). Giá trị trung bình của đại lượng ngẫu nhiên X trong n phép thử là:

$$\overline{X} = \frac{x_1 n_1 + x_2 n_2 + ...+ x_k n_k}{n} = x_1 \frac{n_1}{n} + x_2 \frac{n_2}{n} + ...+ x_k \frac{n_k}{n}$$

$$\overline{X} = x_1 f_1 + x_2 f_2 + ...+ x_k f_k$$

Theo định nghĩa thống kê về xác suất với n đủ lớn ta có:

$$\overline{X} = x_1 p_1 + x_2 p_2 + ...+ x_k p_k = E(X)$$

Vậy: Kỳ vọng toán của đại lượng ngẫu nhiên 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.

Ví dụ 2.2.6. Có 5000 người xét nghiệm máu để tìm ký sinh trùng sốt rét. Tỷ lệ mắc bệnh ở địa phương theo thống kê là 10%. Có thể làm xét nghiệm theo hai phương pháp.

+ Phương pháp 1: Xét nghiệm từng người

+ Phương pháp 2: Lấy máu 10 người một trộn lẫn làm một xét nghiệm. Nếu kết quả xét nghiệm là âm tính (vô trùng) thì thấy qua 10 người không ai mắc bệnh. Nếu kết quả xét nghiệm là dương tính thì chứng tỏ trong 10 người đó có ít nhất một người mắc bệnh. Lúc đó phải làm thêm 10 xét nghiệm lẻ để phát hiện người có bệnh cụ thể. Hỏi làm theo cách nào lợi hơn.

Giải:

Theo phương pháp 1 thì phải làm 5000 xét nghiệm.

Theo phương pháp 2: Gọi X là số xét nghiệm phải làm đối với mỗi nhóm 10 người, X = 1 (Nếu kết quả là âm tính); X = 11 (Nếu kết quả là dương tính)

$$p_1 = P(X=1) = (1-0{,}1)^{10} = (0{,}9)^{10}$$

$$p_2 = P(X=11) = 1 - p_1 = 1 - (0{,}9)^{10}$$

X	1	11
P	$(0{,}9)^{10}$	$1-(0{,}9)^{10}$

$$E(X) = (0{,}9)^{10} + 11.\left(1-(0{,}9)^{10}\right) \approx 7{,}51$$

Tức là trung bình phải làm 7,51 ca xét nghiệm cho mỗi nhóm 10 người. Vậy theo cách 2 phải làm $7{,}51.500 = 3755$ xét nghiệm.

Kết kuận: Làm theo phương pháp 2 lợi hơn.

b. Phương sai

Định nghĩa 2.2.7. *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: $D(X) = E[X - E(X)]^2$.

+ Nếu X là đại lượng ngẫu nhiên rời rạc thì

$$D(X) = \sum_{i=1}^{n} [X_i - E(X)]^2 . p_i .$$

+ Nếu X là đại lượng ngẫu nhiên liên tục thì

$$D(X) = \int_{-\infty}^{+\infty} [X - E(X)]^2 f(x)dx$$

Trong thực tế khi có một mẫu cụ thể ta thường dùng các công thức:

+ Nếu X là đại lượng ngẫu nhiên rời rạc:

$$D(X) = E(X^2) - [E(X)]^2 = \sum_{i=1}^{n} x_i^2 p_i - \left(\sum_{i=1}^{n} x_i p_i \right)^2$$

+ Nếu X là đại lượng ngẫu nhiên liên tục:

$$D(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - \left[\int_{-\infty}^{+\infty} xf(x)dx \right]^2$$

Ví dụ 2.2.8. Trong một hộp kín có 17 quả cầu kích thước như nhau, trong đó có 9 quả màu trắng, 8 quả màu đen, lấy ngẫu nhiên ra 2 quả. Gọi X là số cầu đen được lấy ra:

a/ Lập dãy phân phối xác suất của X.

b/ Tìm hàm phân phối xác suất của X.

c/ Tính E(X), D(X).

Giải: Vì X là số cầu đen được lấy ra nên X là ĐLNN rời rạc nhận các giá trị 0;1;2, với xác suất tương ứng:

$$P(X = 0) = \frac{C_9^2}{C_{17}^2} = \frac{36}{136} \approx 0,265 \quad ; \quad P(X = 1) = \frac{C_9^1 C_8^1}{C_{17}^2} = \frac{72}{136} \approx 0,529$$

$$P(X = 2) = \frac{C_8^2}{C_{17}^2} = \frac{28}{136} \approx 0,206$$

a/ Dãy phân phối xác suất của X:

X	0	1	2
P	0,265	0,529	0,206

b/ Hàm phân phối $F(x)=\begin{cases} 0 & Khi \quad x \le 0 \\ 0{,}265 & Khi \ \ 0 < x \le 1 \\ 0{,}794 & Khi \ \ 1 < x \le 2 \\ 1 & Khi \quad x > 2 \end{cases}$

c/ E(X) = 1.0,529 + 2.0,206 = 0,941;

D(X) = $0{,}529 + 4.0{,}026 - (0{,}941)^2$ = 0,468

Ví dụ 2.2.9. Cho X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất

$$f(x)=\begin{cases} \frac{3}{4}x(2-x) & khi \ \ x \in [0;2] \\ 0 & x \notin [0;2] \end{cases}$$

a/ Tính E(X), D(X).

b/ Tính P{0,9 < X < 1,1} và P{X > 1,5}

Giải:

a/ E(X) = $\int_0^2 \frac{3}{4}x^2(2-x)dx = \frac{3}{4}(\frac{2x^3}{3}-\frac{x^4}{4})\Big|_0^2 = 1$

$$D(X)=\int_0^2 \frac{3}{4}x^3(2-x)dx - \left(\int_0^2 \frac{3}{4}x^2(2-x)dx\right)^2 = \frac{1}{5}$$

b/ P(1,5 < X < 2) = $\int_{1,5}^{2} \frac{3}{4}x(2-x)dx = \frac{3}{4}(x^2-\frac{x^3}{3})\Big|_{1,5}^{2} = 0{,}15625$

P(0,1 < X < 1,1) = $\int_{0,9}^{1,1} \frac{3}{4}x(2-x)dx = \frac{3}{4}(x^2-\frac{x^3}{3})\Big|_{0,9}^{1,1} = 0{,}1495$

Các tính chất của phương sai

Tính chất 2.2.10.

(a) D(C) = 0 (C = Const)

(b) $D(CX) = C^2D(X), (C = const)$

(c) $D(X+Y) = D(X) + D(Y)$, với X và Y độc lập.

Hệ quả 2.2.11. $D(X-Y) = D(X) + D(Y)$, với X và Y độc lập.

Bản chất và ý nghĩa của phương sai.

Phương sai chính là trung bình số học của bình phương các sai lệch giữa các giá trị có thể có của đại lượng ngẫu nhiên so với giá trị trung bình của các giá trị đó. Nó phản ánh mức độ phân tán của các giá trị của đại lượng ngẫu nhiên xung quanh giá trị trung tâm của nó là kỳ vọng toán.

c. Độ lệch tiêu chuẩn

Định nghĩa 2.2.12. Căn bậc hai dương của phương sai được gọi là *độ lệch tiêu chuẩn.*

Ký hiệu : $\sigma_X = \sqrt{D(X)} \quad \Rightarrow D(X) = \sigma_X^2$

Chú ý 2.2.13. Khi cần đánh giá mức độ phân tán của đại lượng ngẫu nhiên theo đơn vị đo của nó người ta thường tính độ lệch tiêu chuẩn chứ không phải là phương sai vì độ lệch tiêu chuẩn có cùng đơn vị đo với đại lượng ngẫu nhiên cần nghiên cứu.

2.3. Một số quy luật phân phối xác suất thông dụng

2.3.1. Quy luật không - một

Định nghĩa 2.3.1. Đại lượng ngẫu nhiên rời rạc X nhận một trong hai giá trị có thể có X = 0; 1 với các xác suất tương ứng được tính bằng công thức $p_x = p^x q^{1-x}$ với x = 0; 1 gọi là *phân phối theo qui luật không - một* với tham số là p.

Ký hiệu: A(p)

Các tham số đặc trưng của qui luật không - một

Nếu X là đại lượng ngẫu nhiên phân phối theo qui luật không - một thì:

E(X) = p; D(X) = pq

2.3.2. Phân phối nhị thức

Định nghĩa 2.3.2. Đại lượng ngẫu nhiên rời rạc X nhận một trong các giá trị có thể có 0, 1, 2, ..., n với các xác suất tương ứng được tính theo công thức Bernoulli gọi là *phân phối theo qui luật nhị thức* với các tham số là n và p.

Kí hiệu: B (n,p)

Ta có bảng phân phối xác suất của đại lượng ngẫu nhiên có phân phối nhị thức là

X	0	1	...	i	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 pq^{n-1}$		$C_n^i p^i q^{n-i}$		$C_n^n p^n q^0$

Các tham số đặc trưng của qui luật nhị thức

Nếu X là đại lượng ngẫu nhiên tuân theo qui luật phân phối nhị thức thì:

E (X) = np, D (X) = npq

Ví dụ 2.3.3. Xác suất để một người bắn trúng bia là 0,8 . Tìm số viên đạn trúng bia trung bình khi người ấy bắn 6 viên đạn.

Giải: Bài toán thoả mãn dãy phép thử Bernoulli. Gọi X là số viên đạn trúng bia thì X phân phối theo qui luật nhị thức với các tham số n = 6; p = 0,8.

E (X) = 6 . 0,8 = 4,8.

Vậy số viên đạn trung bình bắn trúng bia là 5.

2.3.3. Qui luật poisson

Định nghĩa 2.3.4. Đại lượng ngẫu nhiên rời rạc X nhận một trong các giá trị có thể có 0,1,2,.... với xác suất tương ứng được tính bởi $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$, k = 0, 1, 2.... được gọi là phân phối theo qui luật Poisson với tham số là λ ($\lambda = np$).

Ký hiệu: $P(\lambda)$

Phân phối Poisson xuất hiện trong dãy phép thử Bernoulli khi số phép thử khá lớn và xác xuất p khá bé.

Các tham số đặc trưng của qui luật Poisson

Nếu X là đại lượng ngẫu nhiên tuân theo qui luật Poisson thì:

$$E(X) = \lambda \qquad D(X) = \lambda$$

Ví dụ 2.3.5. Xác suất để trong khi vận chuyển mỗi chai rượu bị vỡ là 0,001. Người ta tiến hành vận chuyển 2000 chai rượu đến cửa hàng. Tìm số chai vỡ trung bình khi vận chuyển.

Giải: Bài toán thoả mãn dãy phép thử Bernoulli với n = 2000, P = 0,001 (khá nhỏ). Ta có λ = 2000 . 0,001 = 2 (không đổi).

Gọi X là số chai rượu bị vỡ khi vận chuyển thì X là đại lượng ngẫu nhiên phân phối theo qui luật Poisson.

Vậy số chai bị vỡ trung bình là: $E(X)=2$ (chai).

2.3.4. Phân phối chuẩn $N(a,\sigma^2)$

Định nghĩa 2.3.6. Đại lượng ngẫu nhiên liên tục X nhận các giá trị trong khoảng $(-\infty;+\infty)$ gọi là *phân phối theo qui luật chuẩn* (hay *phân phối chính qui*) với các tham số là a và σ^2 nếu hàm mật độ xác suất của nó có dạng:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}} \qquad (2.1)$$

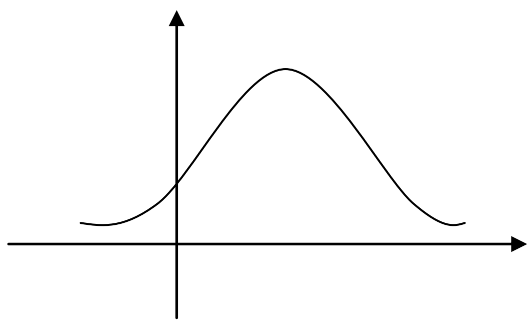
Ký hiệu : $N(a,\sigma^2)$.

Chú ý 2.3.7.

(a) a và σ^2 là hai tham số đặc trưng của phân phối chuẩn.

(b) Đồ thị của hàm (2.1) có dạng hình chuông (h1) nhận trục hoành làm đường tiệm cận, $f(x)>0$ với mọi x, $f_{max}=\frac{1}{\sigma\sqrt{2\pi}}$ đạt tại $x=a$.

Đồ thị hàm số có 2 điểm uốn với hoành độ $x=a\pm\sigma$ và tung độ $f=\frac{1}{\sigma\sqrt{2\pi e}}$



Khi a thay đổi thì dạng của đường cong $f(x)$ không thay đổi. Nó dịch chuyển sang phải nếu a giảm. Khi σ thay đổi thì dạng của đồ thị thay đổi theo. Nếu σ tăng lên thì đồ thị sẽ thấp xuống và phình ra còn khi σ giảm thì đồ thị sẽ cao lên và nhọn thêm.

Các tham số đặc trưng của qui luật chuẩn

$$E(X) = a;\ D(X) = \sigma^2$$

Từ đó ta thấy σ là thước đo độ tản mát các gía trị của biến ngẫu nhiên qua tâm phân phối là a.

Tính xác suất trong phân phối chuẩn

Giả sử $X : N(a,\sigma^2)$. Hãy tính $P(\alpha < X < \beta)$, trong đó (α,β) là khoảng cho trước tuỳ ý. Theo tính chất của hàm mật độ ta có:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}}dx$$

Đặt $\frac{x-a}{\sigma} = t \Rightarrow x = \sigma t + a \Rightarrow dx = \sigma dt$, do đó

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}}\int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}}dt = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right), \qquad (2.2)$$

với $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$ là hàm số Laplatxơ, (với các giá trị được tính sẵn trong bảng phụ lục 2).

Xét trường hợp đặc biệt khi (α,β) là khoảng đối xứng đối với a tức là khoảng $(a-\alpha; a+\alpha)$, ta có

$$P(|X-a| < \alpha) = 2\Phi\left(\frac{\alpha}{\sigma}\right), \qquad (2.3)$$

Ví dụ 2.3.8. Biết độ dài chi tiết X do một máy tự động sản xuất ra tuân theo qui luật chuẩn $N(20; 0,04)$. Tính $P(|X-20| < 0,3)$.

Giải: Theo công thức (2.3) ta có:

$$P(|X-20| < 0,3) = 2\Phi\left(\frac{0,3}{0,2}\right) = 2\Phi(1,5)$$
$$= 2 \,.\, 0,4332 = 0,8664$$

với 0,4332 là giá trị của hàm $\Phi(1,5)$ tra ở bảng hàm số Laplatxơ. Vậy có khoảng 87% chi tiết sản xuất ra có kích thước nằm trong khoảng (19,7; 20,3).

Ví dụ 2.3.9. Chiều cao của nam giới khi trưởng thành ở một vùng dân cư là đại lượng ngẫu nhiên phân phối chuẩn với a = 160cm với $\sigma = 6cm$. Một thanh niên bị coi là lùn nếu có chiều cao nhỏ hơn 155 cm.

a/ Tìm tỷ lệ thanh niên bị lùn ở vùng đó.

b/ Tìm xác suất để lấy ngẫu nhiên 4 người thì có ít nhất 1 người không bị lùn.

Giải:

a/ Gọi X là chiều cao của nam thanh niên khi trưởng thành ở vùng đó. Theo giả thiết $X : N(160\,;36)$. Tỷ lệ thanh niên bị lùn ở vùng đó là xác suất để lấy ngẫu nhiên 1 người thì người đó có kích thước nhỏ hơn 155 cm. Ta có:

$$P(0 < X < 155) = \Phi\left(\frac{155-160}{6}\right) - \Phi\left(\frac{0-160}{6}\right)$$
$$= \Phi\left(\frac{160}{6}\right) - \Phi\left(\frac{5}{6}\right)$$
$$= 0,4999 - 0,2967 = 0,2033$$

b/ Sử dụng công thức Bernoulli ta có:

$$P_4(1;4) = 1 - P_4(0) = 1 - C_4^0(1-0,2033)^0.(0,2033)^4 = 1 - 0,0017 = 0,9983$$

Phân phối chuẩn hoá

Trong trường hợp a=0 và $\sigma = 1$, ta có phân phối chuẩn hoá, khi đó hàm mật độ xác suất của đại lượng ngẫu nhiên X có dạng:

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

đây là hàm số Gauxơ với các giá trị của nó được tính sẵn trong bảng phụ lục 1.

Chú ý 2.3.10.

+ Đồ thị của hàm này nhận trục tung làm trục đối xứng;

+ Hàm phân phối có dạng $F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{u^2}{2}} du$ (Các giá trị của nó được tính ở bảng phụ lục 3).

Ký hiệu của phân phối chuẩn hoá là N(0,1).

Qui tắc Ba xich ma

Trong công thức $P(|X-a| < \alpha) = 2\Phi\left(\frac{\alpha}{\sigma}\right)$ nếu ta đặt $\alpha = 3\sigma$ thì:

$$P(|X-a| \le 3\sigma) = 2\Phi(3) = 2\,.\,0,49865 = 0,9973$$

Xác suất này rất gần 1 nên có thể coi hầu như chắc chắn biến cố $|X-a| \le 3\sigma$ sẽ xảy ra, từ đó ta có qui tắc: Nếu đại lượng ngẫu nhiên X có phân phối chuẩn $X : N(a,\sigma^2)$ thì hầu như chắc chắn rằng X sẽ nhận giá trị trong khoảng $(a-3\sigma; a+3\sigma)$.

2.3.5. Quy luật Student - T(n)

Giả sử U là đại lượng ngẫu nhiên phân phối chuẩn hóa, V là đại lượng ngẫu nhiên độc lập với U phân phối theo quy luật khi bình phương với n bậc tự do. Xét đại lượng ngẫu nhiên $T = \dfrac{U}{\sqrt{\dfrac{V}{n}}}$, đại lượng ngẫu nhiên T sẽ phân phối theo một quy luật phân phối xác suất gọi là *quy luật Student* với n bậc tự do.

Ký hiệu: T(n).

Cũng giống như phân phối chuẩn hóa, phân phối T đối xứng qua gốc O (tức là có trung bình bằng 0). Khi n bé thì phân phối T có đường cong mật độ "mập" hơn đường mật độ N(0, 1), nhưng khi n khá lớn nó rất gần với chuẩn hóa. Trong thực tế nếu n ≥ 30 thì đã có thể coi phân phối T và chuẩn hóa là như nhau.

2.4. Các định lý về giới hạn

2.4.1. Định lý 2.4.1. (Định lý Moavơlaplat) *Nếu trong mỗi phép thử độc lập biến cố A xuất hiện với xác suất* $P(A) = P, (0 < P < 1)$ *thì khi* $n \to \infty$ *ta có:*

$$\lim_{n\to\infty}\left[P_n(k) - \frac{1}{\sqrt{2\pi npq}}e^{-\frac{(k-np)^2}{2npq}}\right] = 0$$

Vậy với n khá lớn ta có: $P_n(k) \approx \dfrac{1}{\sqrt{npq}}\varphi(x)$, với $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ (hàm Gauxơ), $x = \dfrac{k-np}{\sqrt{npq}}$.

Ví dụ 2.4.1. Xác suất để một cây bị chết khi trồng là 0,2 Tính xác suất để khi trồng 400 cây có 80 cây bị chết.

Giải: Ta có n=400; k=80 ; P(A)=0,2 nên q = 0,8, nên $x = \dfrac{80 - 400.0{,}2}{\sqrt{400.0{,}8.0{,}2}} = 0$. Tra bảng Gaoxơ ta được $\varphi(0) = 0{,}3989$, vậy

$$P_{400}(80) = \frac{1}{\sqrt{400.0{,}2.0{,}8}}.0{,}3989 \approx 0{,}05$$

2.4.2. Định lý 2.4.2. (Định lý giới hạn tích phân) *Với các điều kiện như định lý trên ta có:*

$$\lim_{n\to\infty}\left[P_n(k_1,k_2) - \frac{1}{\sqrt{2\pi}}\int_{x_1}^{x_2} e^{-\frac{t^2}{2}}dt\right] = 0,$$ *trong đó* $x_i = \dfrac{k_i - np}{\sqrt{npq}}$, *i=1,2.*

Với n khá lớn ta có: $P_n(k_1,k_2) \approx \frac{1}{\sqrt{2\pi}}\int_{x_1}^{x_2} e^{-\frac{t^2}{2}}dt = \Phi(x_2) - \Phi(x_1)$, trong đó $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$ là hàm Laplatxơ.

2.4.3. Định lý 2.4.3. (Định lý Poisson) *Giả sử tiến hành n phép thử độc lập. Mỗi phép thử sự kiện A xuất hiện với xác suất P(A)=p. Nếu* $n \to \infty$ *mà* $p \to 0$ *sao cho* $np = \lambda = const$ *thì ta có:* $\lim\limits_{n\to\infty} P_n(k) = e^{-\lambda}.\frac{\lambda^k}{k!}$.

Vậy khi n khá lớn ta có: $P_n(k) \approx \frac{\lambda^k.e^{-\lambda}}{k!}$.

Chú ý 2.4.4. Trong trường hợp p rất gần 1 thì $P(\overline{A}) = 1 - p$ rất gần 0. Do đó để tính $P_n(k)$ ta chuyển sang tính xác suất cho n phép thử $\overline{A}$ xuất hiện n-k lần.

Ví dụ 2.4.5. Một công nhân đứng máy xe xợi gồm 800 ống xác suất để mỗi ống xợi bị đứt xợi trong 1 giờ là 0,005. Tính xác suất

a/ Trong 1 giờ có 3 ống xợi bị đứt.

b/ Trong 1 giờ có không quá 10 ống bị đứt xợi.

Giải:

a/ $n = 800; \quad p = 0{,}005; \quad \lambda = n.p = 4$

Vậy $P_{800}(3) = \frac{4^3.e^{-4}}{3!} \approx 0{,}1954$

b/ $P_{800}(0,10) = \sum\limits_{k=0}^{10} P_{800}(k) = 0{,}99716$

2.5. Đại lượng ngẫu nhiên hai chiều

2.5.1. Khái niệm về đại lượng ngẫu nhiên hai chiều

Ta ký hiệu đại lượng ngẫu nhiên hai chiều là (X, Y) trong đó X và Y được gọi là các thành phần của đại lượng ngẫu nhiên hai chiều thực chất mỗi thành phần lại là một đại lượng ngẫu nhiên một chiều. Vậy đại lượng ngẫu nhiên hai chiều thực chất là hệ hai đại lượng ngẫu nhiên X và Y được xét một cách đồng thời. Có hai loại đại lượng ngẫu nhiên:

+ Đại lượng ngẫu nhiên hai chiều là rời rạc nếu các thành phần của nó là rời rạc.

+ Đại lượng ngẫu nhiên hai chiều là liên tục nếu các thành phần của nó là liên tục.

2.5.2. Bảng phân phối xác suất của đại lượng ngẫu nhiên hai chiều.

Đối với đại lượng ngẫu nhiên hai chiều người ta cũng dùng bảng phân phối xác suất, hàm phân phối xác suất và hàm mật độ xác suất để thiết lập qui luật phân phối xác suất của chúng.

Bảng phân phối xác suất của đại lượng ngẫu nhiên hai chiều rời rạc liệt kê các giá trị có thể có của nó và các xác suất tương ứng. Nó có dạng sau đây:

Y \ X	x_1	x_2	. . .	x_i	. . .	x_n
y_1	$p(x_1,y_1)$	$p(x_2,y_1)$	. . .	$p(x_i,y_1)$	. . .	$p(x_n,y_1)$
y_2	$p(x_1,y_2)$	$p(x_2,y_2)$	. . .	$p(x_i,y_2)$	. . .	$p(x_n,y_2)$
. . .	. . .	. . .	. . .	. . .	. . .	. . .
y_j	$p(x_1,y_j)$	$p(x_2,y_j)$	. . .	$p(x_i,y_j)$	. . .	$p(x_n,y_j)$
. . .	. . .	. . .	. . .	. . .	. . .	. . .
y_m	$p(x_1,y_m)$	$p(x_2,y_m)$	. . .	$p(x_i,y_m)$	. . .	$p(x_n,y_m)$

trong đó x_i ($i = 1..n$) là các giá trị có thể có của thành phần X, y_j ($j = 1..m$) là các giá trị có thể có của thành phần Y; còn $p(x_i,y_j)$ là xác suất để đại lượng ngẫu nhiên hai chiều (X, Y) nhận giá trị (x_iy_j).

Chú ý 2.5.1. Các xác suất $p(x_i,y_j)$ phải thỏa mãn điều kiện

$$\begin{cases} p(x_i, y_j) \geq 0 \quad i = \overline{1;n} \quad j = \overline{1;m} \\ \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j) = 1 \end{cases}$$

Biết bảng phân phối xác suất của đại lượng ngẫu nhiên hai chiều bao giờ cũng có thể tìm được bảng phân phối xác suất của mỗi thành phần.

+ Bảng phân phối xác suất của thành phần X có dạng:

X	x_1	x_2	. . .	x_i	. . .	x_n
p	$p(x_1)$	$p(x_2)$	. . .	$p(x_i)$	. . .	$p(x_n)$

Trong đó: $p(x_i) = \sum_{j=1}^{m} p(x_i, y_j)$ $i = 1..n$ và $\sum_{i=1}^{n} p(x_i) = 1$.

+ Bảng phân phối xác suất của thành phần Y có dạng:

Y	y_1	y_2	. . .	y_j	. . .	y_m
p	$p(y_1)$	$p(y_2)$	. . .	$p(y_j)$	. . .	$p(y_m)$

Trong đó: $p(y_j) = \sum_{i=1}^{n} p(x_i, y_j)$ $j = 1..m$ và $\sum_{j=1}^{m} p(y_j) = 1$.

Ví dụ 2.5.2. Tìm bảng phân phối xác suất của các thành phần của đại lượng ngẫu nhiên hai chiều có bảng phân phối xác suất như sau:

Y \ X	x_1	x_2	x_3
y_1	0,10	0,30	0,20
y_2	0,06	0,18	0,16

Giải: Cộng các xác suất theo cột ta thu được các xác suất tương ứng với các giá trị của thành phần X: $p(x_1) = 0,16$; $p(x_2) = 0,48$; $p(x_3) = 0,36$.

Ta có bảng phân phối xác suất của thành phần X như sau:

X	x_1	x_2	x_3
p	0,16	0,48	0,36

Cộng các xác suất theo dòng ta có các xác suất tương ứng với các giá trị của thành phần Y: $p(y_1) = 0,6$; $p(y_2) = 0,4$.

Ta có bảng phân phối xác suất của thành phần Y như sau:

Y	y_1	y_2
p	0,6	0,4

BÀI TẬP CHƯƠNG 2

1. Một lô hàng gồm 7 sản phẩm trong đó có 3 phế phẩm. Chọn ngẫu nhiên ra 4 sản phẩm để kiểm tra. Gọi X là số sản phẩm tốt trong 4 sản phẩm lấy ra.

a/ Tìm quy luật phân phối xác suất của X.

b/ Tìm hàm phân phối xác suất.

c/ Tính E(X); D(X).

2. Kiểm tra vấn đáp hết môn cho 4 học sinh, mỗi học sinh chỉ được vào kiểm tra nếu người được kiểm tra trước đó đạt yêu cầu. Xác suất đạt yêu cầu khi kiểm tra của mỗi học sinh là 0,6. Lập bảng phân phối xác suất, tìm hàm phân phối xác suất, tính kỳ vọng và phương sai của số học sinh được vào kiểm tra.

3. Trong một chiếc hòm có 5 bóng đèn trong đó có 2 bóng tốt và 3 bóng hỏng. Lấy ngẫu nhiên ra 2 bóng để kiểm tra. Gọi X là số bóng tốt trong số 2 bóng được kiểm tra.

a/ Hãy lập dãy phân phối xác suất của X.

b/ Tìm hàn phân phối F(x)?

c/ Tìm E(X) và D(X)?

4. Một túi chứa 10 tấm thẻ đỏ và 6 tấm thẻ xanh. Chọn ra 3 tấm thẻ. Gọi X là số thẻ đỏ được lấy ra.

a/ Lập bảng phân phối xác suất của X?

b/ Tìm hàm phân phối xác suất F(x)?

c/ Tìm E(X) và D(X)?

5. Một thiết bị gồm 3 bộ phận hoạt động độc lập với nhau. Xác suất trong thời gian t các bộ phận bị hỏng tương ứng là 0,4; 0,2 và 0,3. Gọi X là số bộ phận bị hỏng.

a/ Tìm quy luật phân phối xác suất X .

b/ Tìm hàm phân phối F(x).

c/ Tính E(X); D(X).

6. Một xí nghiệp có hai ô tô vận tải hoạt động. Xác suất trong ngày làm việc các ô tô bị hỏng tương ứng là 0,1 và 0,2. Gọi X là số ô tô bị hỏng trong thời gian làm việc.

a/ Tìm quy luật phân phối xác suất của X .

b/ Tìm hàm phân phối xác suất.

c/ Tính E(X); D(X).

7. 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 và 0,5. Hỏi thời gian trung bình phải ngừng trên đường là bao nhiêu. Biết rằng mỗi khi gặp đèn đỏ người đó phải dừng mất 30 giây.

8. Trong phòng thí nghiệm có 3 nghiên cứu viên tiến hành 3 thí nghiệm độc lập về tế bào ung thư trong cùng một khoảng thời gian. Xác suất thực hiện thành công thí nghiệm của nghiên cứu viên thứ nhất là 0,75, nghiên cứu viên thú hai là 0,8 và nghiên cứu viên thứ ba là 0,6. Gọi X là số thí nghiệm thành công trong ba thí nghiệm.

a/ Lập bảng phân phối xác suất của X.

b/ Tính kỳ vọng và phương sai.

9. Có 3 xạ thủ bắn độc lập vào cùng một bia, mỗi xạ thủ bắn 1 viên đạn. Xác suất bắn trúng đích của mỗi xạ thủ là 0,6; 0,5 và 0,4. Gọi X là biến ngẫu nhiên chỉ số viên đạn bắn trúng bia.

a/ Lập bảng phân phối xác suất của biến ngẫu nhiên X.

b/ Tính kỳ vọng, phương sai của biến ngẫu nhiên X.

10. Một xạ thủ có 4 viên đạn. Xạ thủ đó bắn lần lượt từng viên cho đến khi trúng mục tiêu hoặc hết cả 4 viên thì thôi. Xác suất bắn trúng mục tiêu của mỗi viên đạn là 0,6. Gọi X là đại lượng ngẫu nhiên chỉ số viên đạn đã bắn.

a/ Lập bảng phân phối xác suất của X.

b/ Tính kỳ vọng, phương sai của X.

11. Cho hàm số: $f(x)=\begin{cases} \dfrac{x^3}{4}, & x\in(0;2) \\ 0, & x\notin(0;2) \end{cases}$

a/ Chứng minh hàm $f(x)$ là hàm mật độ xác suất của một đại lượng ngẫu nhiên liên tục X .

b/ Tính kỳ vọng, phương sai của đại lượng ngẫu nhiên có hàm mật độ xác suất nói trên.

c/ Tính xác suất để trong 3 phép thử độc lập có 1 lần X nhận giá trị trong $(1;3/2)$.

12. Biến ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x)=\begin{cases} \dfrac{6}{5}x(x-1) & Khi\ x\in[1;\ 2] \\ 0 & Khi\ x\notin[1;\ 2] \end{cases}$$

a/ Hãy tìm hàm phân phối $F(x)$.

b/ Tính $E(X)$.

c/ Tính xác suất P(0 <X <1.5).

13. Cho đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x)=\begin{cases} a(x^3+2x+1), & khi\, x\in[0;4] \\ 0, & khi\, x\notin[0;4] \end{cases}$$

a/ Tìm hệ số a?

b/ Tìm hàm phân phối $F(x)$?

c/ Tính $P(1<X<3)$?

14. Cho đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x)=\begin{cases} k\,x(4-x^2), & khi\ x\in(0;2) \\ 0, & khi\ x\notin(0;2) \end{cases}$$

a/ Tìm hệ số k?

b/ Tìm hàm phân phối $F(x)$?

c/ Tính $E(X)$?

15. Cho X là ĐLNN liên tục với hàm mật độ xác suất:

$$f(x)=\begin{cases} kx^2(1-x), & khi\, x\in[0;1] \\ 0, & khi\, x\notin[0;1] \end{cases}$$

a/ Tìm hệ số k?

b/ Tìm hàm phân phối $F(X)$?

c/ Tìm $P(-0,5<X<0,5)$

16. Cho đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x)=\begin{cases} k\,x^2(4-x), & khi\ x\in[0;4] \\ 0, & khi\ x\notin[0;4] \end{cases}$$

a/ Tìm hệ số k?

b/ Tìm hàm phân phối $F(x)$?

c/ Tính $E(X)$?

17. Cho đại lượng ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x)=\begin{cases} k(1-x^2), & khi\, x\in(-1;1) \\ 0, & khi\, x\notin(-1;1) \end{cases}$$

a/ Tìm hệ số k?

b/ Tìm hàm phân phối $F(x)$?

c/ Tính $E(X)$?

Phần 3. Thống kê

Chương 1

Cơ sở lý thuyết mẫu

1.1. Tổng thể và mẫu

1.1.1. Định nghĩa
a. Tổng thể
Định nghĩa 1.1.1. Toàn bộ tập hợp các phần tử đồng nhất theo một dấu hiệu nghiên cứu định tính hoặc định lượng nào đó được gọi là tổng thể nghiên cứu hay tổng thể (hay tập chính). Số lượng các phần tử của tổng thể được gọi là *kích thước của tổng thể* (thường được ký hiệu là N).

Với mỗi tổng thể ta không nghiên cứu trực tiếp tổng thể đó mà thông qua một hay nhiều dấu hiệu đặc trưng cho tổng thể đó, chúng được gọi là dấu hiệu nghiên cứu, các dấu hiệu này có thể là định tính hoặc định lượng.

b. Mẫu
Định nghĩa 1.1.2. Nếu từ tổng thể ta chọn ngẫu nhiên ra n phần tử thì tập hợp n phần tử này được gọi là mẫu kích thước n, khi đó ta sẽ tìm cách xem xét đánh giá mẫu đó rồi suy ra kết luận cho tổng thể.

c. Mẫu ngẫu nhiên:
Tiến hành n quan sát độc lập về biến ngẫu nhiên X nào đó gọi X_i là việc quan sát lần thứ i về biến ngẫu nhiên X. Khi đó $(X_1, X_2, \ldots, X_n)$ được gọi là mẫu ngẫu nhiên, n gọi là cỡ mẫu hay số lần quan sát. (mẫu ngẫu nhiên cỡ n thực chất là n biến ngẫu nhiên độc lập, cùng phân phối như biến ngẫu nhiên X)

Định nghĩa 1.1.3. Mẫu ngẫu nhiên kích thước n là tập hợp của n biến ngẫu nhiên độc lập $X_1, X_2, \ldots, X_n$ được thành lập từ biến ngẫu nhiên X và có cùng phân phối xác suất với X.

Mẫu ngẫu nhiên được ký hiệu W = $(X_1, X_2, \ldots, X_n)$. Lúc đó việc thực hiện một phép thử đối với mẫu ngẫu nhiên W chính là thực hiện một phép thử đối với mỗi thành phần của mẫu. Giả sử 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 $x_1, x_2, \cdots, x_n$ tạo thành một giá trị của mẫu ngẫu nhiên, hay còn gọi là một mẫu cụ thể, ký hiệu $w = (x_1, x_2, \cdots, x_n)$

Ví dụ 1.1.4. Gọi X là số chấm xuất hiện khi tung một con xúc xắc, X là ĐLNN với bảng phân phối xác suất như sau:

X	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Nếu tung con xúc xắc 3 lần và gọi X_i là số chấm xuất hiện trong lần tung thứ i (i = 1, 2, 3) thì ta có 3 ĐLNN độc lập có cùng qui luật phân phối xác suất với X. Vậy ta có mẫu ngẫu nhiên kích thước n = 3, W = (X_1, X_2, X_3) được xây dựng từ ĐLNN gốc X. Thực hiện một phép thử đối với mẫu ngẫu nhiên này tức là tung cụ thể 3 lần: Giả sử lần thứ nhất được 3 chấm, lần thứ hai được 4 chấm, lần thứ ba được 2 chấm thì ta thu được một mẫu cụ thể w =(3, 4, 2).

1.1.2. Phương pháp xây dựng mẫu

a. Nhận xét

Từ kết quả tập mẫu có được ta có thể suy ra các kết quả cho tổng thể bởi vậy bao giờ cũng có thể mắc phải sai lầm nhất định. Độ sai lệch lớn hay bé phụ thuộc vào phương pháp xây dựng mẫu và kích thước mẫu. Độ chính xác trong thống kê thường được gọi là độ tin cậy của kết luận: ký hiệu là γ. Gọi α là tỷ lệ sai sót (hay mức ý nghĩa) thì α = 1- γ.

Để có căn cứ vào thông tin của mẫu đưa ra những kết luận đủ chính xác về dấu hiệu nghiên cứu trong tổng thể, tức là phản ánh đúng đặc điểm của tổng thể theo dấu hiệu nghiên cứu đó. Để đảm bảo tính đại diện của mẫu và tiện cho việc mô hình hoá, mẫu được tạo lập với những giả thiết sau:

+ Lấy lần lượt từng phần tử vào mẫu, phương pháp này gọi là phương pháp đơn giản để phân biệt với cách lấy cùng một lúc nhiều phần tử vào mẫu.

+ Mỗi phần tử được lấy vào mẫu một cách hoàn toàn ngẫu nhiên, tức là mọi phần tử của tổng thể đều có thể được lấy vào mẫu với khả năng như nhau.

+ Các phần tử được lấy vào mẫu theo phương thức hoàn lại, tức là trước khi lấy phần tử thứ k thì trả lại tổng thể phần tử thứ (k - 1) mà ta đã nghiên cứu xong (k = 2..n).

Chú ý 1.1.5.

+ Trong việc lấy mẫu, do nhiều nguyên nhân khác nhau, sẽ không tránh khỏi các sai số trong số liệu mẫu. Vì vậy trước khi dùng các phương pháp thống kê để phân tích, sử lý ta cần loại bỏ các sai số không đáng có ở trong mẫu đã cho. Có 3 loại sai số

1/ Sai số thô: Là sai số sinh ra do vi phạm các điều kiện cơ bản của việc lấy mẫu, hoặc do sơ suất của người thực hiện. Chẳng hạn người kiểm tra cố ý chọn ra các sản phẩm tốt để kiểm tra khi đánh giá chất lượng, hoặc người kỹ thuật viên ghi nhầm kết quả thu được.

2/ Sai số hệ thống: Là sai số do không điều chỉnh chính xác dụng cụ hoặc không thống nhất giữa các kỹ thuật viên về cách xác định một đại lượng nào đó…, dẫn đến một loạt kết quả quan sát được bị lệch đi một tỷ lệ nhất định nào đó.

3/ Sai số ngẫu nhiên: Là sai số sinh ra do một số lớn các nguyên nhân mà tác dụng của chúng bé đến mức không tách riêng và tính riêng biệt cho từng nguyên nhân được.

Trong ba loại sai số trên, **sai số thô** và sai **số hệ thống** cần phát hiện sớm và khử bỏ ngay, còn sai số ngẫu nhiên thì không thể khử bỏ được trong mỗi lần quan sát. Do đó các kết quả quan sát được đưa vào sử lý bằng các phương pháp toán học ta sẽ giải quyết chúng chỉ chứa các sai số ngẫu nhiên.

+ Trong thực tế nếu kích thước của tổng thể khá lớn còn mẫu chỉ chiếm một phần rất nhỏ của tổng thể thì phương pháp lấy mẫu hoàn lại và không hoàn lại cho ta các kết quả sai lệch không đáng kể.

b. Một số phương pháp chọn mẫu

**** Mẫu ngẫu nhiên đơn:*** là loại mẫu được chọn trực tiếp từ danh sách đã được đánh số của tổng thể. Từ tổng thể kích thước N người ta dùng cách rút thăm đơn giản ra n phần tử của mẫu theo một bảng số ngẫu nhiên nào đó.

+ Ưu điểm: Phương pháp này có ưu điểm là cho phép thu được một mẫu có tính đại diện cao cho phép suy rộng các kết quả của mẫu cho tổng thể với một sai số xác định.

+ Nhược điểm: Phải có được toàn bộ danh sách của tổng thể nghiên cứu, mặt khác chi phí chọn mẫu sẽ khá lớn.

**** Mẫu ngẫu nhiên hệ thống*****:** Là loại mẫu ngẫu nhiên đã được đơn giản hoá trong cách chọn, trong đó chỉ có phần tử đầu tiên được chọn một cách ngẫu nhiên, sau đó dựa trên danh sách đã được đánh số của tổng thể để chọn ra các phần tử tiếp theo vào mẫu theo một thủ tục nào đó.

+ Nhược điểm của phương pháp này là dễ mắc sai số hệ thống khi danh sách của tổng thể không được sắp xếp một cách ngẫu nhiên mà lại theo một trật tự chủ quan nào đó.

**** Mẫu chùm:*** Tổng thể điều tra được phân chia ra thành nhiều chùm theo nguyên tắc:

+ Mỗi phần tử của tổng thể chỉ được phân vào một chùm.

+ Mỗi chùm cố gắng chứa nhiều phần tử khác nhau về dấu hiệu cần nghiên cứu sao cho nó có độ phân tán cao như của tổng thể.

+ Phân chia sao cho các chùm tương đối đồng đều nhau về qui mô.

+ Các chùm được chọn một cách ngẫu nhiên và tất cả các phần tử của chùm đó đều được chọn vào mẫu.

**** Mẫu phân tổ*****:** Trong chọn mẫu phân tổ trước hết người ta phân chia tổng thể ra thành các tổ có độ thuần nhất cao để chọn ra các phần tử đại diện cho từng tổ. Việc phân tổ có hiệu quả khi tổng thể nghiên cứu không thuần nhất theo dấu hiệu nghiên cứu. Sau khi đã phân tổ thì kích thước mẫu được phân bố cho mỗi tổ theo một qui tắc nào đó.

**** Mẫu nhiều cấp*****:** Nếu các phần tử của tổng thể phân tán quá rộng và thiếu thông tin về chúng, người ta thường chọn mẫu theo nhiều cấp, khi chọn nhiều cấp ta có nhiều loại đơn vị chọn mẫu ở mỗi cấp, thường được gọi là đơn vị chọn mẫu cấp 1, cấp 2,...Để chọn mẫu ở mỗi cấp chỉ cần có thông tin về phân bố của dấu hiệu ở cấp ấy là đủ. Việc chọn mẫu ở mỗi cấp có thể tiến hành theo phương pháp mẫu ngẫu nhiên đơn, mẫu ngẫu nhiên hệ thống, mẫu chùm hay mẫu phân tổ.

c. Các phương pháp sắp xếp mẫu ngẫu nhiên

Giả sử từ tổng thể của ĐLNN gốc X rút ra một mẫu ngẫu nhiên kích thước n, trong đó giá trị x_1 xuất hiện với tần số n_1, giá trị x_2 xuất hiện với tần số n_2,..., giá trị x_k xuất hiện với tần số n_k, lúc đó sau khi các x_i đã được sắp xếp theo trình tự tăng dần giá trị cụ thể của mẫu w có thể mô tả bằng bảng phân phối tần số thực nghiệm sau:

x_i	$x_1 \quad x_2 \; ... \; x_i \; ... \; x_k$
n_i	$n_1 \quad n_2 \; ... \; n_i \; ... \; n_k$

với $n_1 + n_2 + ... + n_k = n$. Gọi $f_i = \frac{n_i}{n}$, (i = 1...k)

Ta có bảng phân phối tần suất thực nghiệm:

x_i	$x_1 \quad x_2 \; ... \; x_i \; ... \; x_k$
f_i	$f_1 \quad f_2 \; ... \; f_i \; ... \; f_k$

với $f_1 + f_2 + ... + f_k = 1$

Ví dụ 1.1.6. Gặt ngẫu nhiên 100 điểm trồng lúa của một vùng thu được các số liệu được sắp xếp thành bảng sau (gọi là bảng phân phối tần số thực nghiệm).

Năng suất(tạ/ha)	21	24	25	26	28	32	34
Số điểm gặt tương ứng	10	20	30	15	10	10	5

Bảng phân phối tần suất thực nghiệm có dạng:

x_i	21	24	25	26	28	32	34
f_i	0,1	0,2	0,3	0,15	0,1	0,1	0,05

Chú ý 1.1.7. Nếu các phần tử của mẫu khá gần nhau khi đó ta có thể sắp xếp mẫu thành một dẫy phân phối thực nghiệm ghép lớp.

1.2. Các đặc trưng của mẫu ngẫu nhiên

Để nghiên cứu đại lượng ngẫu nhiên gốc X trong tổng thể, nếu chỉ rút ra một mẫu ngẫu nhiên W = $(X_1, X_2, \cdots, X_n)$ thì mới chỉ có được một vài kết luận sơ bộ và rời rạc về X, vì các giá trị X_i của mẫu có cùng qui luật phân phối xác suất với X song qui luật này lại thường chưa được xác định hoàn toàn. Song nếu tổng hợp các đại lượng ngẫu nhiên $X_1, X_2, \cdots, X_n$ này lại chúng sẽ bộc lộ những tính qui luật mới làm cơ sở để nhận định về đại lượng ngẫu nhiên gốc X trong tổng thể.

Việc tổng hợp mẫu W = $(X_1, X_2, \cdots, X_n)$ được thực hiện dưới dạng một hàm nào đó của các giá trị $X_1, X_2, \cdots, X_n$ của mẫu. Nó được gọi là ***thống kê***.

Ký hiệu là G. Vậy G = f($X_1, X_2, \cdots, X_n$)

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 qui luật phân phối xác suất nhất định và có các tham số đặc trưng tương ứng. Mặt khác, khi mẫu ngẫu nhiên nhận một giá trị cụ thể w = $(x_1, x_2, \ldots, x_n)$ thì G cũng nhận một giá trị cụ thể là

$$g = f(x_1, x_2, \ldots, x_n)$$

Sau đây ta sẽ nghiên cứu một số thống kê thông dụng nhất

1.2.1. Trung bình mẫu

Giả sử từ tổng thể của ĐLNN gốc X lập một mẫu ngẫu nhiên kích thước n.

$$W = (X_1, X_2, \ldots, X_n).$$

Trung bình mẫu là một thống kê ký hiệu là $\overline{X}$: $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Vì tập hợp mẫu W = ($X_1, X_2, \ldots, X_n$) được thực hiện dưới dạng một hàm nào đó của các giá trị $X_1, X_2, \ldots, X_n$ của mẫu nên nó là một thống kê do đó trung bình mẫu $\overline{X}$ cũng là một thống kê. Vì vậy nó là một ĐLNN tuân theo một qui luật phân phối xác suất nào đó với các tham số đặc trưng tương ứng. Khi mẫu nhận một giá trị cụ thể w = $(x_1, x_2, \ldots, x_n)$ thì trung bình mẫu cũng nhận một giá trị cụ thể ký hiệu là $\overline{x}$.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ hoặc } \overline{x} = \frac{1}{n}\sum_{i=1}^{k} m_i x_i .$$

Tính chất 1.2.1.

Nếu E(X) = a và D(X) = σ^2 thì E($\overline{X}$) = a và D($\overline{X}$) = $\frac{\sigma^2}{n}$

Chú ý 1.2.2. Nếu các x_i cách đều nhau một khoảng là h thì

$$\overline{x} = x_0 + \frac{h}{n}\sum_{i=1}^{k} m_i\left(\frac{x_i - x_0}{h}\right)$$

trong đó x_0 là giá trị tạo với m_i tương ứng đạt max

Ví dụ 1.2.3. Cho bảng phân phối tần số thực nghiệm (ví dụ 1). Hãy tính trung bình mẫu

Giải: Ta có

$$\overline{x} = \frac{1}{100}(21.10 + 24.20 + 25.30 + 26.15 + 28.10 + 32.10 + 34.5) = 26 \text{ (tạ/ha)}$$

Ví dụ 1.2.4. Cho mẫu

X	28	30	32	34	36	38	40	42
m_i	3	7	10	15	10	3	1	1

Hãy tìm kỳ vọng mẫu.

Giải: Nhận xét mẫu trên có kích thước n = 50 và các x_i cách đều nhau một khoảng h = 2; max m_i = 15 do đó chọn x_0 = 34 ta được

$$\bar{x} = 34 + \frac{2}{50}\left(3.(-3)+7.(-2)+10.(-1)+15.0+10.1+3.2+1.3+1.4\right) = 33{,}6$$

Vậy E($\bar{x}$) = 33,6.

1.2.2. Phương sai mẫu

Phương sai mẫu ký hiệu là $S^2(X)$.

Hoàn toàn tương tự như trung bình mẫu khi cho một mẫu cụ thể thì phương sai mẫu sẽ nhận một giá trị cụ thể. ký hiệu là $s^2(x) = s^2$

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

Trong thực tế để tiện cho việc tính toán người ta thường hay sử dụng công thức

$$s^2 = \frac{1}{n}\sum_{i=1}^{k} m_i x_i^2 - \left(\bar{x}\right)^2 \Leftrightarrow s^2 = \overline{x^2} - \left(\bar{x}\right)^2$$

Chú ý 1.2.5. Nếu các x_i cách đều nhau một khoảng là h thì

$$s^2 = \frac{h^2}{n}\sum_{i=1}^{k} m_i \left(\frac{x_i - x_0}{h}\right)^2 - (x_0 - \bar{x})^2$$

Tính chất 1.2.6. Nếu E(X) =a và D(X) = σ^2 thì E(S^2) = $\frac{n-1}{n}\sigma^2$

1.2.3. Phương sai điều chỉnh mẫu

Phương sai điều chỉnh mẫu ký hiệu S'^2

$$S'^2 = \frac{1}{n-1}\sum_{i-1}^{k} m_i (X_i - \overline{X})^2 = \frac{n}{n-1} S^2$$

Chú ý 1.2.7. Phương sai điều chỉnh mẫu S'^2 là một thống kê còn khi có mẫu cụ thể thì phương sai mẫu điều chỉnh s'^2 cũng là một số xác định.

1.2.4. Độ lệch tiêu chuẩn mẫu và độ lệch tiêu chuẩn điều chỉnh mẫu

+ Độ lệch tiêu chuẩn mẫu S = $\sqrt{S^2}$

+ Độ lệch tiêu chuẩn điều chỉnh mẫu $S' = \sqrt{S'^2}$

1.2.5. Tần suất mẫu

Giả sử từ tổng thể kích thước N trong đó có M phần tử mang dấu hiệu nghiên cứu, lấy ra một mẫu ngẫu nhiên kích thước n và trong đó thấy có X phần tử mang dấu hiệu nghiên cứu. Lúc đó tần suất mẫu là một thống kê ký hiệu là f, f = $\frac{X}{n}$.

Ví dụ 1.2.8. Cho mẫu thực nghiệm

X	65 70	70 75	75 80	80 85	85 90	90 95	95 100	100 105	105 110	110 115	115 120	120 125	125 130
m_i	1	0	2	5	8	16	18	17	16	9	5	2	1

Tìm $\overline{x}$, s^2, s.

Giải: Đặt $x_i = \frac{x_i + x_{i+1}}{2}$ (i = 1, ..., 13), các x_i cách đều nhau một khoảng h = 5 chọn $x_0 = 97,5$, đặt $t_i = \frac{x_i - x_0}{5}$ với i =1, ..., 13.

Ta lập lại bảng

STT	x_i	m_i	t_i	$m_i t_i$	t_i^2	$m_i t_i^2$
1	67,5	1	- 6	- 6	36	36
2	72,5	0	- 5	0	25	0
3	77,5	2	- 4	- 8	16	32
4	82,5	5	- 3	- 15	9	45
5	87,5	8	- 2	- 16	4	32
6	92,5	16	- 1	- 16	1	16
7	97,5	18	0	0	0	0
8	102,5	17	1	17	1	17
9	107,5	16	2	32	4	64
10	112,5	9	3	27	9	81
11	117,5	5	4	20	16	80
12	122,5	2	5	10	25	50
13	127,5	1	6	6	36	36
$\sum$		100		51		489

Nhìn vào bảng trên ta thấy:

$$\overline{x} = 97,5 + \frac{5}{100}..51 = 100,05$$

$$s^2 = \frac{25}{100}.489 - (97,5 - 100,05)^2 \approx 115,75$$

$$s = \sqrt{s^2} = \sqrt{115,75} \approx 10,76$$

Chương 2

Ước lượng tham số

Giả sử khi xét một ĐLNN X (X là một dấu hiệu định lượng cần nghiên cứu của một tổng thể) ta cần biết qui luật phân phối xác suất của nó. Bằng phương pháp phân tích lý thuyết ta giả sử biết được dạng phân phối xác suất của nó. Tuy nhiên các tham số đặc trưng của nó như kỳ vọng, phương sai,.. mà ta gọi chung là tham số lý thuyết θ lại chưa biết nên ta cần phải xác định θ. Việc tính chính xác θ là khó có thể thực hiện được mà ta chỉ có thể tính gần đúng. Việc tính gần đúng đó ta gọi là *ước lượng tham số* θ và dựa vào mẫu thực nghiệm đã có.

Phương pháp tiến hành: Từ tổng thể cần nghiên cứu rút ra một mẫu ngẫu nhiên kích thước n và dựa vào mẫu đó mà xây dựng một thống kê G dùng để ước lượng θ bằng cách này hay cách khác, có hai phương pháp sử dụng G để ước lượng θ là phương pháp ước lượng điểm và phương pháp ước lượng bằng khoảng tin cậy.

2.1. 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, vì bản thân θ là một số xác định. Thông thường giá trị được chọn là một thống kê G nào đó của mẫu ngẫu nhiên. Có nhiều cách chọn thống kê G khác nhau tạo nên những phương pháp ước lượng điểm khác nhau.

2.1.1. Phương pháp hàm ước lượng

a. Khái niệm

Giả sử cần ước lượng tham số θ của đại lượng ngẫu nhiên gốc X. Từ tổng thể lập mẫu ngẫu nhiên kích thước n: W = ($X_1, X_2, \ldots, X_n$). Chọn lập thống kê G = f($X_1, X_2, \ldots, X_n$) mà thực chất là một thống kê đặc trưng mẫu tương ứng với tham số θ cần ước lượng. Chẳng hạn, để ước lượng kỳ vọng toán a của ĐLNN gốc thì chọn thống kê trung bình mẫu $\overline{X}$, để ước lượng phương sai σ^2 của ĐLNN gốc thì chọn thống kê phương sai mẫu S^2…Nếu lập một mẫu cụ thể và tính được giá trị g = (x_1, $x_2, \ldots, x_n$) của thống kê G trên mẫu cụ thể đó thì ước lượng của θ là giá trị g vừa tính được.

Vì thống kê $G = f(X_1, X_2,\ldots,X_n)$ thực chất là hàm của các ĐLNN nên nó được gọi là hàm ước lượng của θ. Chất lượng của ước lượng không thể đánh giá qua một giá trị cụ thể của G. Vì như vậy chỉ có cách so sánh trực tiếp g và θ mà θ lại chưa biết. Do đó chỉ có thể đánh giá chất lượng của ước lượng thông qua bản thân thống kê $G = f(x_1, x_2,\ldots, x_n)$. Rõ ràng là có vô số cách chọn hàm f, tức là có vô số thống kê G có thể dùng làm hàm ước lượng của θ. Vì vậy cần đưa ra một tiêu chuẩn đẻ đánh giá chất lượng của thống kê G, từ đó lựa chọn được thống kê " xấp xỉ một cách tốt nhất" tham số cần ước lượng.

b. Các tiêu chuẩn lựa chọn hàm ước lượng

**** Ước lượng không chệch***

Giả sử thống kê G là ước lượng của tham số θ của ĐLNN gốc. Với k mẫu cụ thể rút ra từ tổng thể, thống kê G sẽ nhận k giá trị cụ thể tương ứng là $g_1,g_2,\ldots,g_k$. Nếu thống kê G là một ước lượng có dư của θ thì các giá trị $g_1,g_2,\ldots,g_k$ cũng đều sẽ lớn hơn θ và giá trị trung bình của chúng (tức là kỳ vọng toán của G) cũng lớn hơn θ: $E(G) > \theta$. Ngược lại, nếu thống kê G là một ước lượng thiếu của θ thì mọi giá trị $g_1, g_2,\ldots,g_k$ cũng đều sẽ nhỏ hơn θ nên $E(G) < \theta$.

Như vậy, việc sử dụng một thống kê mà kỳ vọng toán của nó khác với tham số cần ước lượng có thể dẫn đến sai số hệ thống (tất cả các giá trị của G đều lớn hơn hoặc nhỏ hơn θ). Để loại trừ sai số này hiển nhiên là cần yêu cầu $E(G) = \theta$. Dĩ nhiên yêu cầu trên không loại trừ được hoàn toàn các sai số, song như vậy các sai số khác dấu sẽ xuất hiện tương đối đều nhau, do đó các giá trị của G sẽ không bị lệch hẳn về một phía so với θ.

Định nghĩa 2.1.1. Thông kê G của mẫu được gọi là ước lượng không chệch của tham số θ của ĐLNN gốc X nếu $E(G) = \theta$.

Ngược lại $E(G) \neq \theta$. thì G được gọi là ước lượng chệch của θ.

Chú ý 2.1.2.

(a) G là ước lượng không chệch của tham số θ không có nghĩa là mọi giá trị của G đều trùng khít với θ mà chỉ có nghĩa: trung bình các giá trị của thống kê G bằng θ. Từng giá trị của G có thể sai lệch rất lớn so với θ.

(b) Trung bình mẫu $\overline{X}$ là ước lượng không chệch của kỳ vọng toán a của ĐLNN gốc [E(X) = a].

(c) Tần suất mẫu f là ước lượng không chệch của xác suất p của ĐLNN gốc [E(f) = p].

(d) Phương sai điều chỉnh mẫu S'^2 là ước lượng không chệch của phương sai σ^2 của ĐLNN gốc [$E(S'^2) = \sigma^2$].

Ví dụ 2.1.3. Giả sử trái cây của nông trường đã đóng thùng, mỗi thùng 10 trái. kiểm tra ngẫu nhiên 50 thùng ta thu được kết quả sau:

Số trái cây hỏng trong thùng k	0	1	2	3	4	5	6	7	8	9	10
Số thùng có k trái hỏng	0	2	3	7	20	6	4	7	0	0	1

a/ Tìm ước lượng cho tỷ lệ trái cây hỏng trong nông trường.

b/ Tìm ước lượng cho tỷ lệ trái cây hỏng trung bình trong mỗi thùng.

c/ Tìm ước lượng không chệch cho cho độ biến động tỷ lệ trái cây hỏng ở mỗi thùng,

Giải:

a/ Đây là bài toán ước lượng điểm cho tỷ lệ đám đông.

Tổng số trái điều tra là: n = 10x50 = 500 (trái)

Số trái hỏng phát hiện là: m = 1.2+2.3+3.7+4.20+5.6+6.4+7.7+10.1= 22.

Tỷ lệ hỏng trong mẫu là: $f = \frac{222}{500} = 0,444$.

Vậy tỷ lệ trái hỏng trong nông trường là vào khoảng 44,4%.

b/ Ta quan điểm đây là bài toán về số lượng chứ không phải là chất lượng. Mỗi thùng có tỷ lệ hỏng X_i và ta cần ước lượng điểm tỷ lệ hỏng trung bình ở mỗi thùng.

Ta có bảng sau: (h = 10; x_0 = 40; $t_i = \frac{x_i - 40}{10}$)

X_i	m_i	t_i	m_it_i	$m_it_i^2$
10	2	- 3	- 6	18
20	3	- 2	- 6	12
30	7	- 1	- 7	7
40	20	0	0	0
50	6	1	6	6
60	4	2	8	16
70	7	3	21	63
100	1	6	6	36
$\sum$	50		22	158

Nhìn vào bảng trên ta thấy:

$$\bar{x} = 40 + \frac{10}{50}.22 = 44,4.$$

Vậy tỷ lệ hỏng trung bình ở mỗi thùng là khoảng 44,4%.

c/ Ước lượng không chệch cho độ biến động tỷ lệ trái cây hỏng ở mỗi thùng chính là phương sai điều chỉnh mẫu s'^2

$$s^2 = \frac{100}{50}.158 - (40 - 44,4)^2 = 296,64$$

$$s'^2 = \frac{50}{49}.296,64 = 302,69$$

Ta dự đoán độ biến động của tỷ lệ hỏng giữa các thùng là vào khoảng 302

* ***Ước lượng hiệu quả:***

Như đã phân tích ở trên, dù G là ước lượng không chệch của θ thì từng giá trị cụ thể của G vẫn có thể sai lệch rất lớn so với θ, tức là phương sai D(G) vẫn có thể rất lớn. Lúc đó, nếu lấy một giá trị của G tìm được trên một mẫu cụ thể, chẳng hạn g_1 để ước lượng θ thì nó có thể sai lệch rất nhiều so với giá trị trung bình $\overline{G}$ tức là bản thân tham số θ cần ước lượng. Như vậy, nếu lấy một giá trị của G, chẳng hạn g_1 để ước lượng θ, thì có thể mắc sai số rất lớn. Còn nếu như đòi hỏi phương sai của G phải nhỏ thì có thể loại trừ sai số này trong ước lượng.

Định nghĩa 2.1.4. Thống kê G của mẫu được gọi là ước lượng hiệu quả của tham số θ của ĐLNN gốc X nếu nó có phương sai nhỏ nhất so với mọi thống kê khác được xây dựng trên cùng mâũ đó.

Như vậy, để xét xem G có phải là ước lượng hiệu quả của θ hay không, cần phải tìm được giá trị nhỏ nhất có thể có của phương sai các hàm ước lượng.

*** *Ước lượng vững*:**

Khi xét những mẫu có kích thước lớn thì nẩy sinh vấn đề là mẫu càng lớn thì thống kê G của mẫu phải càng gần tham số θ cần ước lượng.

Định nghĩa 2.1.5. Thống kê G của mẫu được gọi là ước lượng vững của tham số θ của ĐLNN gốc X nếu G hội tụ theo xác suất đến θ khi $n \to \infty$. Tức là với mọi ε dương bé tuỳ ý ta luôn có $\lim\limits_{x \to \infty} P(|G-\theta| < \varepsilon) = 1$

2.1.2. Phương pháp ước lượng hợp lý tối đa

Giả sử biết qui luật phân phối xác suất tổng quát của ĐLNN gốc X dưới dạng hàm mật độ f(x, θ). Đó cũng có thể là biểu thức xác suất nếu X là ĐLNN rời rạc, cần phải ước lượng tham số θ nào đó của X. Lập mẫu ngẫu nhiên kích thước n:

$$W = (X_1, X_2,. . .,X_n).$$

và xây dựng hàm của đối số θ tại một giá trị cụ thể của mẫu:

$$L(x_1, x_2, . . .,x_n, \theta) = f(x_1,\theta).f(x_2,\theta)\ldots f(x_n,\theta).$$

Hàm L được gọi là hợp lý của tham số θ, giá trị của hàm hợp lý chính là xác suất hay mật độ xác suất tại điểm $(x_1, x_2, . . .,x_n)$, còn giá trị của thống kê G tại điểm đó $g = f(x_1, x_2, . . .,x_n)$ được gọi là ước lượng hợp lý tối đa của θ nếu ứng với giá trị này của θ, hàm hợp lý đạt cực đại.

2.2. Phương pháp ước lượng bằng khoảng tin cậy

2.2.1. Khái niệm

Để ước lượng tham số θ của ĐLNN gốc X trong tổng thể, phương pháp này chủ trương từ một thống kê G nào đó của mẫu xây dựng một khoảng giá trị (G_1, G_2) sao cho với một xác suất cho trước, tham số θ sẽ rơi vào khoảng (G_1, G_2) đó. Chú ý rằng do G là ĐLNN nên khoảng (G_1, G_2) cũng là một khoảng ngẫu nhiên, còn θ lại là một số xác định nên phải nói chính xác hơn là khoảng (G_1, G_2) sẽ chứa đựng giá trị θ với một xác suất cho trước.

Định nghĩa 2.2.1. Khoảng (G_1, G_2) của thống kê G được gọi là *khoảng tin cậy* của tham số θ nếu với xác suất bằng (1 - α) cho trước thoả mãn điều kiện $P(G_1 < \theta < G_2) = (1 - \alpha)$; xác suất (1 - α) = γ được gọi là độ tin cậy của ước lượng, còn $I = G_2 - G_1$ được gọi là độ dài khoảng tin cậy.

***Phương pháp tiến hành*:**

Từ tổng thể lập mẫu ngẫu nhiên kích thước n: $W = (X_1, X_2,. . .,X_n)$ và từ đó xây dựng thống kê $G = f(X_1, X_2,. . .,X_n, \theta)$ sao cho qui luật phân phối xác suất của G không phụ thuộc vào các đối số của nó và hoàn toàn xác định. Lúc đó với độ tin cậy bằng (1 - α) cho trước có thể tìm được cặp giá trị α_1 và α_2 sao cho $\alpha_1 + \alpha_2 = \alpha$ và tương đương với chúng tìm được cặp giá trị $g\alpha_1$ và $g\alpha_2$ (thường là các phân vị của G) thoả mãn điều kiện $P(G < g_{\alpha_1}) = \alpha_1$ và $P(G > g_{\alpha_2}) = \alpha_2$. Từ đó suy ra $P(g_{\alpha_1} < G < g_{\alpha_2}) = 1 - (\alpha_1 + \alpha_2) = (1 - \alpha)$

Như vậy, với độ tin cậy (1 - α) ta đã xây dựng được khoảng tin cậy (g_{α_1}, g_{α_2}) cho G. Bằng các phép biến đổi tương đương bao giờ cũng có thể đưa công thức trên về dạng biểu thức tương đương $P(G_1 < \theta < G_2) = 1 - \alpha$ đó chính là khoảng tin cậy cần tìm.

Khi tiến hành một phép thử với mẫu ngẫu nhiên $W = (X_1, X_2, \ldots, X_n)$ ta thu được một mẫu cụ thể $w = (x_1, x_2, \ldots, x_n)$, do đó tính được giá trị của G_1 và G_2 ứng với mẫu cụ thể này là g_1 và g_2, lúc đó có kết luận là: Qua mẫu cụ thể w, với độ tin cậy 1 - α tham số θ của ĐLNN gốc X sẽ nằm trong khoảng (g_1, g_2) tức là ($g_1 < \theta < g_2$).

Nhận xét 2.2.2. Với độ tin cậy 1 - α cho trước ta có thể tìm được vô số cặp α_1, α_2 ($0 < \alpha_1 < 1$; $0 < \alpha_2 < 1$) thoả mãn $\alpha_1 + \alpha_2 = \alpha$ vì thế có vô số khoảng tin cậy tương ứng với độ tin cậy đã cho.

2.2.2. Ước lượng kỳ vọng toán của ĐLNN phân phối theo qui luật chuẩn

Giả sử trong tổng thể ĐLNN gốc X phân phối chuẩn $N(a, \sigma^2)$ nhưng chưa biết tham số a của nó. Để ước lượng a từ tổng thể lập một mẫu ngẫu nhiên kích thước n, $W = (X_1, X_2, \ldots, X_n)$. Để chọn thống kê G thích hợp ta xét các trường hợp sau:

a. Trường hợp đã biết phương sai σ^2 của ĐLNN gốc X

Ta chọn thống kê: $G = U = \dfrac{(\overline{X} - a)\sqrt{n}}{\sigma}$, (I)

Trong đó $\overline{X}$ là trung bình mẫu. Ta thấy thống kê U phân phối chuẩn hoá N(0,1). Do đó với độ tin cậy 1 - α cho trước tìm được cặp giá trị α_1 và α_2 sao cho $\alpha_1 + \alpha_2 = \alpha$. Từ đó tìm được hai phân vị chuẩn tương ứng là u_{α_1} và $u_{1-\alpha_2}$ thoả mãn điều kiện $P(U < u_{\alpha_1}) = \alpha_1$ và $P(U > u_{1-\alpha_2}) = \alpha_2$.

Từ đó $P(u_{\alpha_1} < U < u_{1-\alpha_2}) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$ Vì $u_{\alpha_1} = -u_{1-\alpha_1}$

Thay U từ biểu thức (I) vào biểu thức trên và giải ra theo ẩn a ta thu được

$$P\left(\overline{X} - \frac{\sigma}{\sqrt{n}} u_{1-\alpha_2} < a < \overline{X} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha_1}\right) = 1 - \alpha$$

Vậy với độ tin cậy bằng (1 - α) tham số a của ĐLNN gốc X sẽ nằm trong khoảng: $\left(\overline{X} - \dfrac{\sigma}{\sqrt{n}} u_{1-\alpha_2}; \overline{X} + \dfrac{\sigma}{\sqrt{n}} u_{1-\alpha_1}\right)$, (1)

Các trường hợp đặc biệt:

+ *Khoảng tin cậy đối xứng*: $\alpha_1 = \alpha_2 = \dfrac{\alpha}{2}$ thì khoảng tin cậy của a là:

$$(\overline{X} - \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} ;\ \overline{X} + \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}), \qquad (2)$$

Nếu ký hiệu $\varepsilon = \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}$ thì biểu thức trên có dạng: $(\overline{X} - \varepsilon ;\ \overline{X} + \varepsilon)$, (3)

ε được gọi là độ chính xác (độ sai số) 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.

+ ***Khoảng tin cậy bên phải***: Nếu $\alpha_1 = 0$, $\alpha_2 = \alpha$ thì $u_{1-\alpha_1} = +\infty$ do đó khoảng tin cậy bên phải của a là

$$(\overline{X} - \frac{\sigma}{\sqrt{n}} u_{1-\alpha} ;\ +\infty), \qquad (4)$$

Biểu thức trên dùng để ước lượng giá trị tối thiểu của a.

+ ***Khoảng tin cậy bên trái***: Nếu $\alpha_2 = 0$, $\alpha_1 = \alpha$ thì $u_{1-\alpha_2} = +\infty$ do đó khoảng tin cậy của a là:

$$(-\infty ;\ \overline{X} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha}), \qquad (5)$$

Biểu thức trên dùng để ước lượng giá trị tối đa của a.

Chú ý 2.2.3.

(a) Với cùng độ tin cậy $(1 - \alpha)$ thì độ dài khoảng tin cậy I sẽ là ngắn nhất khi khoảng tin cậy là đối xứng và $I = 2\varepsilon$.

(b) Từ các công thức trên ta thấy ba số ε, α, n luôn phụ thuộc vào nhau, nếu cho trước 2 số thì sẽ tìm được số còn lại, chẳng hạn cho trước α và ε thì ta sẽ tìm được kích thước tối thiểu n của mẫu cần phải điều tra là:

$$n \geq \left[\frac{\sigma^2 u^2_{1-\frac{\alpha}{2}}}{\varepsilon^2} \right] + 1 \quad \Leftrightarrow \quad n \geq \frac{4\sigma^2}{I^2} u^2_{1-\frac{\alpha}{2}} + 1, \quad (6)$$

(c) Các giá trị $u_{1-\frac{\alpha}{2}}$ và $u_{1-\alpha}$ được tra ở bảng phụ lục 3

(d) Các khoảng tin cậy nói trên vẫn đang còn là khoảng tin cậy ngẫu nhiên đối với một mẫu ngẫu nhiên $W = (X_1, X_2, \ldots, X_n)$. Thực hiện một phép thử đối với mẫu này thu được mẫu cụ thể $w = (x_1, x_2, \ldots, x_n)$, từ đó tìm được giá trị cụ thể $\overline{x}$ của trung bình mẫu. Lúc đó với độ tin cậy $(1 - \alpha)$ qua một mẫu cụ thể ta sẽ tìm được một khoảng tin cậy cụ thể của a. Vậy với một mẫu ngẫu nhiên cho phép xác định khoảng tin cậy ngẫu nhiên, còn mẫu cụ thể cho phép tìm được khoảng tin cậy cụ thể (bằng số) của a.

Ví dụ 2.2.4. Để xác định trọng lượng trung bình của các bao gạo trong kho người ta lấy ngẫu nhiên ra 100 bao và tìm được trọng lượng trung bình $\overline{x} = 36{,}06$ kg với $\sigma^2 = (0{,}28)^2$. Hãy tìm khoảng tin cậy đối xứng của trọng lượng trung bình của các bao gạo trong kho với độ tin cậy (1- α) = 0,99. Nếu giả thiết trọng lượng các bao gạo tuân theo luật phân phối chuẩn.

Giải: Đây là bài toán tìm khoảng tin cậy đối xứng của giá trị trung bình khi đã biết phương sai. Vậy từ mẫu cụ thể khoảng tin cậy đối xứng của a là

$$\left(\overline{x} - \frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\ ;\ \overline{x} + \frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\right)$$

với $u_{1-\frac{\alpha}{2}} = u_{0,995} = 2,576 \quad \Rightarrow \quad \varepsilon = \frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}} = \frac{0,28.2,576}{\sqrt{100}} \approx 0,072$

$$\overline{x} - \varepsilon = 36,06 - 0,072 = 35,988$$

$$\overline{x} + \varepsilon = 36,06 + 0,072 = 36,132$$

Vậy với độ tin cậy 99% thì trong lượng trung bình các bao gạo trong kho nằm trong khoảng (35,988 ; 36,132) kg

Ví dụ 2.2.5. Trọng lượng của một loại sản phẩm là một ĐLNN có phân phối chuẩn, với độ lệch tiêu chuẩn là 1. Cần phải điều tra một mẫu có kích thước là bao nhiêu để với độ tin cậy của ước lượng là 0,95 thì sai số cho phép không vượt quá 0,1.

Giải: Theo giả thiết ta có ε_0 = 0,1; σ =1; với 1 - α = 0,95 nên $u_{1-\frac{\alpha}{2}}$ = 1,96

Theo công thức (3.5) ta được n $\geq \left[\frac{1}{(0,1)^2}.(1,96)^2\right] + 1 = [384,16] + 1$

Vậy để đáp ứng các yêu cầu của đầu bài ta phải điều tra một mẫu có kích thước tối thiểu n = 385

b. Trường hợp chưa biết phương sai của ĐLNN gốc X

**** Kích thước mẫu nhỏ n $\leq$ 30.***

Ta chọn thống kê G = T = $\frac{\overline{X} - a}{S'}\sqrt{n}$

Trong đó S' là độ lệch tiêu chuẩn điều chỉnh mẫu. Ta thấy thống kê T tuân theo qui luật Student với (n - 1) bậc tự do. Với độ tin cậy 1 - α cho trước có thể tìm được cặp giá trị α_1 và α_2 sao cho $\alpha_1 + \alpha_2 = \alpha$, từ đó tìm được hai phân vị Student tương ứng là $t_{\alpha_1}^{(n-1)}$ và $t_{1-\alpha_2}^{(n-1)}$ thoả mãn điều kiện:

$$P(\,T < t_{\alpha_1}^{(n-1)}\,) = \alpha_1 \quad \text{và} \quad P(\,T > t_{1-\alpha_2}^{(n-1)}\,) = \alpha_2$$

Từ đó $P(t_{\alpha_1}^{(n-1)} < T < t_{1-\alpha_2}^{(n-1)}) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$

Vì $t_{\alpha_1}^{(n-1)} = -\,t_{1-\alpha_1}^{(n-1)}$ nên biểu thức trên có thể viết $P(-\,t_{1-\alpha_1}^{(n-1)} < T < t_{1-\alpha_2}^{(n-1)}) = 1 - \alpha$

Hoàn toàn tương tự như trên khoảng tin cậy của a với độ tin cậy 1 - α là:

$$\left(\overline{X} - \frac{S'}{\sqrt{n}}t_{1-\alpha_2}^{(n-1)}\ ;\ \overline{X} + \frac{S'}{\sqrt{n}}t_{1-\alpha_1}^{(n-1)}\right), \qquad (7)$$

Các trường hợp đặc biệt:

+ ***Khoảng tin cậy đối xứng***: $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ thì khoảng tin cậy của a là:

$$(\overline{X} - \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{(n-1)} ; \overline{X} + \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{(n-1)}), \qquad (8)$$

+ ***Khoảng tin cậy bên phải***: Khi $\alpha_1 = 0$, $\alpha_2 = \alpha$ thì khoảng tin cậy bên phải của a là:

$$(\overline{X} - \frac{S'}{\sqrt{n}} t_{1-\alpha}^{(n-1)}; +\infty), \qquad (9)$$

Biểu thức trên dùng để ước lượng giá trị tối thiểu của a.

+ ***Khoảng tin cậy bên trái***: Khi $\alpha_2 = 0$, $\alpha_1 = \alpha$ thì khoảng tin cậy bên trái của a là:

$$(-\infty; \overline{X} + \frac{S'}{\sqrt{n}} t_{1-\alpha}^{(n-1)}), \qquad (10)$$

Biểu thức trên dùng để ước lượng giá trị tối đa của a.

Chú ý 2.2.6.

(a) Tra bảng phụ lục 5 (phân vị Student $t_{\alpha}^{(n)}$) ta sẽ tìm được các giá trị $t_{1-\frac{\alpha}{2}}^{(n-1)}$ hoặc $t_{1-\alpha}^{(n-1)}$

(b) Nếu đặt $\varepsilon = \frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{(n-1)}$ thì (8) $\Leftrightarrow$ $(\overline{X} - \varepsilon; \overline{X} + \varepsilon)$, (8')

và độ dài khoảng tin cậy I sẽ là ngắn nhất khi khoảng tin cậy sẽ là đối xứng và

$$I = 2\varepsilon = 2\frac{S'}{\sqrt{n}} t_{1-\frac{\alpha}{2}}^{(n-1)}$$

(c) Nếu ấn định độ tin cậy và sai số cho phép không vượt quá ε_0 thì khi đó dung lượng mẫu cần thiết phải thoả mãn

$$n \geq \left[\frac{S'^2}{\varepsilon_0^2} \left(t_{1-\frac{\alpha}{2}}^{n-1} \right)^2 \right] + 1, \qquad (11)$$

(d) Với độ tin cậy khá lớn, để có khoảng tin cậy cụ thể của a, người ta lập mẫu cụ thể w = (x_1, x_2, . . ., x_n), từ đó tính được các giá trị $\overline{x}$ và s' cụ thể rồi thay chúng vào các công thức trên ta sẽ có khoảng tin cậy cụ thể bằng số phải tìm.

Ví dụ 2.2.7. Một giống lúa mới được gieo trong 10 miếng đất thí nghiệm có các điều kiện giống nhau, cho sản lượng tính theo cùng đơn vị như sau:

25,4; 28,0; 20,1; 27,4; 25,6; 23,9; 24,8; 26,4; 27,0; 25,4.

Hãy tìm khoảng tin cậy của sản lượng giống lúa trên với độ tin cậy 95%. Giả thiết sản lượng lúa là đại lượng ngẫu nhiên có phân phối chuẩn.

Giải: Gọi X là sản lượng của giống lúa trên, theo giả thiết X phân phối chuẩn. $N(a, \sigma^2)$ với a là tham số chưa biết cần ước lượng . Đây là bài toán tìm khoảng ước lượng đối xứng của sản lượng giống lúa khi chưa biết phương sai với mẫu nhỏ. (Sử dụng công thức 8')

Theo bài ra ta có: n =9 $\overline{x}$ = 25,4; s' = 2,238 γ = 1 - α = 0,95. Tra bảng ta được $t^{9}_{0,975}$ = 2,262, $\varepsilon = \frac{S'}{\sqrt{n}} t^{(n-1)}_{1-\frac{\alpha}{2}}$ = 1,601;

$\overline{x}$ - ε = 25,4 - 1,601 = 23,799 ; $\overline{x}$ + ε = 25,4 + 1,601 = 27,001

Vậy với độ tin cậy 95% sản lượng trung bình của giống lúa mới trên nằm trong khoảng (23,799, 27,001) hay 23,799 < a < 27,001

***Kích thước mẫu lớn n ≥ 30*:** Ta chọn thống kê G = T = $\frac{(\overline{X}-a)\sqrt{n}}{S'}$

Người ta đã chứng minh được rằng khi n → +∞ thì thống kê T sẽ phân phối xấp xỉ N(0,1) do đó với độ tin cậy cho trước và với n đủ lớn (n ≥ 30) ta sẽ xấp xỉ phân phối Student bằng phân phối chuẩn (Vẫn áp dụng các khoảng tin cậy như với mẫu nhỏ nhưng thay việc tra bảng phụ lục 5 bằng bảng phụ lục 3).

Ví dụ 2.2.8. Để ước lượng năng suất trung bình của giống lúa mới tại một vùng, người ta gặt ngẫu nhiên 100 thửa ruộng của vùng đó và thu được kết quả sau:

Năng suất X (tạ/ha)	40 42	42 44	44 46	46 48	48 50	50 52
Số thửa m_i	7	13	25	35	15	5

Biết năng suất là ĐLNN có phân phối chuẩn. Hãy ước lượng lượng năng suất trung bình của giống lúa mới ở vùng đó với độ tin cậy 95%.

Giải: Đây là bài toán tìm khoảng ước lượng đối xứng của trung bình đám đông khi chưa biết phương sai với mẫu có kích thước lớn. (áp dụng công thức (8'))

Theo bài ra ta có γ = 0,95 do đó $u_{1-\frac{\alpha}{2}} = u_{0,975}$ = 1,96. Ta phải tìm $\overline{x}$ và s',

Đặt $x_i = \frac{x_i + x_{1+1}}{2}$, chọn x_0 = 47 , h = 2 đặt $t_i = \frac{x_i - x_0}{h}$

STT	x	m_i	t_i	$m_i.t_i$	t_i^2	$m_i t_i^2$
1	41	7	-3	-21	9	63
2	43	13	-2	-26	4	52
3	45	25	-1	-25	1	25
4	47	35	0	0	0	0
5	49	15	1	15	1	15
6	51	5	2	10	4	20
Σ		100		-47		175

Nhìn vào bảng trên ta thấy:

$$\overline{x} = x_0 + \frac{h}{n}\sum_{i=1}^{6} m_i t_i = 47 + \frac{2}{100}.(-47) = 46,06$$

$$s^2 = \frac{h^2}{n}\sum_{i=1}^{6} m_i t_i^2 - (x_0 - \overline{x})^2 = \frac{4}{100}.175-(47-46,06)^2 = 6,1164$$

$s'^2 = \frac{100}{99}.6,1164 = 6,178 \Rightarrow s' = 2,486$

$\varepsilon = \frac{S'}{\sqrt{n}} u_{1-\frac{\alpha}{2}}$ = 2,486 .1,96/10 = 0,487

$\bar{x} - \varepsilon = 45,573$; $\bar{x} + \varepsilon = 46,547$

Vậy với độ tin cậy 95% năng suất trung bình của giống lúa mới tại vùng đó nằm trong khoảng (45,573; 46,547) tạ/ha.

Ví dụ 2.2.9. Chiều dài (X) loại sản phẩm A do một máy tự động sản xuất ra là một ĐLNN tuân theo luật phân phối chuẩn với độ lệch tiêu chuẩn $\sigma = 3$ cm. Để ước lượng chiều dài trung bình của loại sản phẩm nói trên với độ tin cậy 95% người ta tiến hành đo 36 sản phẩm.

a/ Tìm khoảng tin cậy đối xứng của chiều dài trung bình loại sản phẩm đó.

b/ Để ước lượng với độ tin cậy 99%, độ dài khoảng tin cậy đối xứng không vượt quá 0,6 cm thì phải đo bao nhiêu sản phẩm?

Giải: Gọi chiều dài trung bình mẫu là $\bar{x}$. Đây là bài toán ước lượng bằng khoảng tin cậy đối xứng của trung bình tổng thể của ĐLNN tuân theo luật phân phối chuẩn khi đã biết phương sai. Theo giả thiết n = 36; $\sigma = 3$; $\gamma = 0,95$ do đó

$\varepsilon = \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} = \frac{3}{6}.1,96 = 0,98.$

a/ Chiều dài trung bình của loại sản phẩm trên là

$(\bar{x} - 0,98 ; \bar{x} + 0,98)$

b/ Theo giả thiết $I \leq 0,6$ cm; $\varepsilon \leq 0,3$ cm; $\gamma = 0,99 \Rightarrow u_{1-\frac{\alpha}{2}} = 2,576.$

Áp dụng công thức 3.5 ta được $n \geq \left[\frac{(2,576)^2.3^2}{(0,3)^2}\right] + 1 = 664.$

Kết luận: Để đáp ứng các yêu cầu của đầu bài thì cần phải điều tra một mẫu có kích thước tối thiểu là 664 sản phẩm.

Ví dụ 2.2.10. Năng suất giống ngô A ở một vùng được báo lên qua 25 điểm thu hoạch và có kết quả sau:

Năng suất (tạ/ha) X	7	9	11	13	17
Số điểm thu hoạch m_i	2	7	12	3	1

Biết năng suất ngô của vùng đó là ĐLNN có phân phối chuẩn.

a/ Tìm khoảng tin cậy đối xứng của năng suất trung bình của giống ngô A của vùng đó với độ tin cậy 95%.

b/ Hãy tính năng suất trung bình tối thiểu của giống ngô A của vùng đó với độ tin cậy 95%

Giải:

a/ Theo giả thiết X là ĐLNN phân phối chuẩn $N(a,\sigma^2)$. Vậy năng suất trung bình chính là giá trị a. Đây là bài toán ước lượng bằng khoảng tin cậy đối xứng giá trị của tham số a của phân phối N(a,σ^2) khi chưa biết phương sai σ^2 của X với mẫu nhỏ. Ta có khoảng tin cậy đối xứng của a là:

$$\left(\overline{X} - \frac{S'}{\sqrt{n}} t^{(n-1)}_{1-\frac{\alpha}{2}} \; ; \; \overline{X} + \frac{S'}{\sqrt{n}} t^{(n-1)}_{1-\frac{\alpha}{2}} \right)$$

Qua 25 điểm báo lên tức là ta có một mẫu ngẫu nhiên kích thước n = 25, gọi X_i là năng suất của điểm thứ i (i = 1...25) ta có W = (X_1, X_2, . . ., X_{25}) từ đó $\overline{X} = \frac{1}{25}\sum_{i=1}^{25} X_i$. Với độ tin cậy 95% tra bảng phụ lục 5 ta được $t^{24}_{0,975}$ = 2,064.

Từ mẫu cụ thể ta sẽ tìm được $\overline{x}$ và s'. Ta thấy các x_i cách đều nhau một khoảng h = 2, chọn x_0 = 11, đặt $t_i = \frac{x_i - x_0}{2}$

X	m_i	t_i	m_it_i	t^2	m_it^2
7	2	-2	- 4	4	8
9	7	-1	- 7	1	7
11	12	0	0	0	0
13	3	1	3	1	3
17	1	3	3	9	9
$\sum$	25		-5		27

Nhìn vào bảng trên ta thấy:

$\overline{x}$ = 11 - 0,4 = 10,6; s^2 = 4,16 ; s' = 2,08 ; ε = 0,859.

Kết luận: Với độ tin cậy 95% năng suất trung bình của giống ngô A đó là:

(9,741 ; 11,459)

b/ Đây là bài toán tìm khoảng tin cậy bên phải của ước lượng a của X: N(a,σ^2) khi chưa biết phương sai với mẫu cụ thể n = 25

Khoảng tin cậy bên phải của a là $\left(\overline{X} - \frac{S'}{\sqrt{n}} t^{(n-1)}_{1-\alpha}; +\infty \right)$

Với độ tin cậy 95% thì $t^{24}_{1-\alpha}$ = 1,711 do đó $\overline{x} - \frac{S'}{\sqrt{n}} t^{n-1}_{1-\alpha}$ = 9,888.

Kết luận: Với độ tin cậy 95% năng suất trung bình tối thiểu của giống ngô A của vùng đó là a ≥ 9,888 tạ/ha.

Chú ý 2.2.11.

(a) Ta không được viết P (9,743 < a < 11,457) = 0,95 vì độ tin cậy gắn với độ tin cậy ngẫu nhiên chứ không gắn với một mẫu cụ thể. Hơn nữa do a là một hằng số nên nó chỉ có thể thuộc hoặc không thuộc khoảng (9,743 ; 11,457) tức là với một mẫu cụ thể thì biến cố (9,743 < a < 11,457) không phải là biến cố ngẫu nhiên. Hoặc nó là biến cố chắc chắn, hoặc nó là biến cố không thể có.

(b) Trong bài bài toán trên ở ý a, nếu ta tăng độ tin cậy từ 95% lên 99% thì giá trị tra bảng $t^{n-1}_{1-\frac{\alpha}{2}} = t^{24}_{0,995}$ = 2,797 do đó ε = $\frac{S'}{\sqrt{n}} t^{n-1}_{1-\frac{\alpha}{2}}$ = 1,164, khoảng tin cậy của a sẽ là (9,435 ; 11,764). Vậy nếu tăng độ tin cây lên mà giữ nguyên kích thước mẫu n thì giá trị của phân vị chuẩn cũng tăng theo do đó ε cũng tăng lên làm cho độ chính xác của ước lượng cũng giảm xuống.

Khi tăng kích thước mẫu n lên và giữ nguyên độ tin cậy cho trước thì ε giảm xuống tức là độ chính xác của ước lượng tăng lên.

2.2.3. Ước lượng của kỳ vọng toán của ĐLNN X không phân phối theo qui luật chuẩn:

Giả sử ở một tổng thể, dấu hiệu định lượng cần nghiên cứu nào đó được xem như ĐLNN X phân phối theo một qui luật nào đó khác qui luật chuẩn. X có kỳ vọng toán là a mà ta cần ước lượng. Để ước lượng a ta chọn thống kê:

G = U = $\dfrac{\overline{X}-a}{\sigma}\sqrt{n}$ (khi đã biết phương sai) hoặc

G = T = $\dfrac{\overline{X}-a}{S'}\sqrt{n}$ (khi chưa biết phương sai)

Người ta đã chứng minh khi kích thước của mẫu đủ lớn thì thống kê G được coi là có phân phối chuẩn hoá N(0,1). Do vậy để ước lượng a ta cần phải chọn mẫu có kích thước lớn và khi đó ta sẽ đi ước lượng a giống như ước lượng a của X có phân phối chuẩn.

2.3. Ước lượng khoảng tin cậy cho tỷ lệ

Giả sử ở một tổng thể của ĐLNN X có N phần tử, trong đó có M phần tử mang đặc tính A và N - M phần tử không mang đặc tính A (0 < M < N). Gọi P là tỷ lệ giữa số phần tử mang đặc tính A với toàn bộ số phần tử của tổng thể P = $\dfrac{M}{N}$, P chính là xác suất để một phần tử trong tổng thể mang đặc tính A. Việc tính chính xác P là gặp nhiều khó khăn, do đó ta đi ước cho tỷ lệ P với độ tin cậy cho trước (P không quá lớn hoặc quá bé)

Phương pháp tiến hành:

Từ tổng thể cần nghiên cứu ta lấy ra một mẫu kích thước n

W = (X_1, X_2, . . ., X_n), X_i là số phần tử mang đặc tính A ở lần thử thứ i. Các X_i (i = 1...n) là những ĐLNN độc lập với nhau và có phân phối 0 - 1 với tham số P. E(X_i) = P và D(X_i) = p(1 - p). Gọi f = $\dfrac{1}{n}\sum_{i=1}^{n} X_i = \overline{X} \Rightarrow$ f là tỷ lệ mang đặc tính A trong mẫu ngẫu nhiên được lấy ra.

Ta đã biết E(f) = P và D(f) = $\dfrac{p(1-p)}{n}$. Ta sử dụng thống kê U = $\dfrac{f-p}{\sqrt{p(1-p)}}\sqrt{n}$.

Ta thấy khi n đủ lớn (n > 30) f không quá bé và cũng không quá lớn thì U phân phối xấp xỉ N(0,1). Khi đó với độ tin cậy cho trước ta có thể tìm được hai phân vị chuẩn $u_{\frac{\alpha}{2}}$ và $u_{1-\frac{\alpha}{2}}$ thoả mãn điều kiện:

$P\left(u_{\frac{\alpha}{2}} < u < u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$. Do $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$ nên $P\left(-u_{1-\frac{\alpha}{2}} < u < u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$

hay $P\left(-u_{1-\frac{\alpha}{2}} < \frac{(f-p)\sqrt{n}}{\sqrt{p(1-p)}} < u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$. Giải hệ bất phương trình trên với ẩn p theo $u_{1-\frac{\alpha}{2}}$ ta được khoảng tin cậy đối xứng của P với độ tin cậy $1 - \alpha$ cho trước. Ta thấy khi n đủ lớn thì thống kê U cũng có phân phối xấp xỉ N(0,1). Như vậy với độ tin cậy $1 - \alpha$ thì:

+ Khoảng tin cậy đối xứng của P là

$$\left(f - \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\frac{\alpha}{2}};\ f + \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\right)$$

+ Khoảng tin cậy bên phải của P là

$$\left(f - \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\alpha};+\infty\right),$$

+ Khoảng tin cậy bên trái của P là

$$\left(-\infty;\ f + \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\alpha}\right)$$

Chú ý 2.3.1.

(a) Nếu đặt $\varepsilon = \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\frac{\alpha}{2}}$ thì độ dài khoảng tin cậy sẽ là ngắn nhất nếu nó là khoảng tin cậy đối xứng và $I = 2\varepsilon$. ε được gọi là độ chính xác hay sai số của ước lượng. Từ đó ta có thể suy ra:

+ Kích thước của mẫu cần phải điều tra đảm bảo cho việc ước lượng P có độ tin cậy $1 - \alpha$ và sai số cho phép không vượt quá ε_0 là:

$$N_1 \geq \left[\frac{f(1-f)}{\varepsilon_0^2}u^2_{1-\frac{\alpha}{2}}\right] + 1 \text{ (khi đã có mẫu định hướng)}$$

hoặc $$N_2 \geq \left[\frac{1}{4\varepsilon_0^2}u^2_{1-\frac{\alpha}{2}}\right] + 1 \quad \text{(khi chưa có mẫu định hướng)}$$

(b) Để sử dụng các công thức trên có kết quả chính xác hơn ta chú ý: n phải lớn, f không quá bé hoặc quá lớn. Trong thực hành thường áp dụng với $n \geq 100$; $0{,}1 \leq f \leq 0{,}9$; $nf \geq 10$; $n(1 - f) \geq 10$

Ví dụ 2.3.2. Kiểm tra ngẫu nhiên 400 sản phẩm do một máy sản xuất ra thấy có 20 sản phẩm là phế phẩm. Với độ tin cậy 95% hãy ước lượng tỷ lệ phế phẩm tối đa của máy đó.

Giải: Gọi P là tỷ lệ phế phẩm của máy đó. Như vậy P là cơ cấu của tập hợp sản phẩm do máy đó sản xuất theo dấu hiệu " phế phẩm". Đây là bài toán ước lượng tham số P của qui luật phân phối không - một A(P) bằng khoảng tin cậy bên trái.

Vậy khoảng tin cậy của P có dạng:

$$\left(-\infty; f+\frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{1-\alpha}\right).$$

Qua mẫu cụ thể ta có f = $\frac{20}{400}$ = 0,05. Với 1 - α = 0,95 $\Rightarrow u_{1-\alpha}$ = 1,645. Vậy với độ tin cậy 0,95, qua mẫu cụ thể này thì khoảng tin cậy của P là

$$\left(-\infty\quad;0{,}05+\frac{\sqrt{0{,}05.0{,}95}}{\sqrt{400}}.1{,}645\right)\text{ hay P < 0,0679.}$$

Kết luận: Với độ tin cậy 95% tỷ lệ phế phẩm tối đa của máy đó là 6,79%

Ví dụ 2.3.3. Để kiểm tra số cá trong một hồ, cơ quan quản lý đánh bắt 2000 con cá, đánh dấu rồi thả xuống hồ. Lần sau đánh bắt lại 400 con, được 80 con có dấu. Hãy ước lượng số cá trong hồ với độ tin cậy 95%

Giải: Gọi N là số cá có trong hồ (N phải nguyên, dương). Tỷ lệ cá bị đánh dấu là P = $\frac{2000}{N}$. Ta phải đi ước lượng P bằng khoảng tin cậy đối xứng. Từ mẫu cụ thể ta có f = $\frac{80}{400}$ = 0,2 (ta thấy: 0,1 < 0,2 < 0,9; nf > 10; n(1 - f) > 10), với độ tin cậy 95% tra bảng ta được $u_{1-\frac{\alpha}{2}}$ = 1,96 $\Rightarrow \varepsilon = \frac{\sqrt{0{,}2.0{,}8}}{\sqrt{400}}1{,}96 = 0{,}0392$. Vậy khoảng tin cậy đối xứng của P là: 0,2 - 0,0392 < P < 0,2 + 0,0392.

Hay $0{,}1608 < \frac{2000}{N} < 0{,}2392 \Leftrightarrow 8362 < N < 12438$

Kết luận: với độ tin cậy 95% số cá có trong hồ nằm trong khoảng từ 8362 đến 12438 con.

BÀI TẬP CHƯƠNG 2

1. Trọng lượng của một loại trứng gà được cho bởi bảng số liệu sau:

X-Trọng lượng (g)	25-30	30-35	35-40	40-45	45-50
Số quả	15	17	40	18	10

Bằng khoảng tin cậy đối xứng hãy ước lượng trọng lượng trung bình của loại trứng gà này với độ tin cậy 95%. Cho biết trọng lượng trứng gà là đại lượng ngẫu nhiên có phân phối chuẩn.

2. Kích thước của một loại sản phẩm do một máy tự động sản xuất ra là một đại lượng ngẫu nhiên tuân theo qui luật phân phối chuẩn. Sau khi kiểm tra 25 sản phẩm cụ thể ta thu được bảng số liệu sau:

Kích thước (cm)	20-22	22-24	24-26	26-28	30-32
Số sản phẩm	3	7	10	3	2

Hãy ước lượng kích thước trung bình của loại sản phẩm đó bằng khoảng tin cậy đối xứng với độ tin cậy 95%.

3. Để ước lượng năng suất trung bình của một giống lúa mới tại một vùng, người ta gặt ngẫu nhiên trên 50 thửa ruộng của vùng đó và thu được kết quả (tạ/ha):

Năng suất	57	58	59	60	61	62	63	64	65	66	67	68	70
Số thửa	2	3	2	6	4	4	8	6	4	3	4	3	1

Biết năng suất lúa là đại lượng ngẫu nhiên có phân phối chuẩn. Hãy ước lượng năng suất trung bình của giống lúa mới ở vùng đó với độ tin cậy 95%.

4. Để xác định giá trung bình đối với một loại hàng hoá trên thi trường, người ta điều tra ngẫu nhiên tại 100 cửa hàng thu được số liệu sau:

Giá (đồng)	83	85	87	89	91	93	95	97	99	101
Số cửa hàng	5	8	13	14	30	11	8	6	4	1

Với độ tin cậy 95% hãy ước lượng giá trung bình của loại hàng đó tại thời điểm đang xét. Biết rằng giá hàng hoá là đại lượng ngẫu nhiên tuân theo qui luật phân phối chuẩn.

5. Với độ tin cậy 95%, hãy ước lượng lượng xăng hao phí trung bình cho một loại xe ôtô chạy từ A đến B nếu chạy thử 30 lần trên đoạn đường này người ta ghi nhận được lượng xăng hao phí như sau:

Lượng xăng hao phí(lít)	9,6-9,8	9,8-10,0	10,0-10,2	10,2-10,4	10,4-10,6
Số lần tương ứng	3	5	10	8	4

Biết lượng xăng hao phí là ĐLNN tuân theo qui luật chuẩn.

6. Cân thử 100 quả trứng ta có kết quả sau:

X (g)	150	160	165	170	180	185
Số quả	4	20	25	30	15	6

Tìm khoảng ước lượng cho khối lượng trung bình của trứng với độ tin cậy 95%. Biết rằng khối lượng trúng là ĐLNN tuân theo quy luật phân phối chuẩn.

7. 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)	15-17	17-19	19-21	21-23	23-25	25-27
Số chi tiết máy tương ứng	1	3	4	12	3	2

Bằng khoảng tin cậy đối xứng hãy ước lượng thời gian gia công trung bình một chi tiết máy với độ tin cậy 95%. Giả thiết thời gian gia công chi tiết là ĐLNN tuân theo qui luật phân phối chuẩn.

8. Đo chỉ số mỡ sữa của 100 con bò lai Hà - Ấn F_1 ta được bảng số liệu sau:

Chỉ số mỡ sữa (X)	3,0-3,6	3,6-4,2	4,2-4,8	4,8-5,4	5,4-6,0	6,0-6,6	6,6-7,2
Số bò lai	2	8	30	35	15	7	3

Hãy ước lượng chỉ số mỡ sữa trung bình của giống bò lai trên với độ tin cậy 95%. Giả thiết chỉ số mỡ sữa là đại lượng ngẫu nhiên có phân phối chuẩn.

9. Đo áp lực X (tính bằng kg/cm^2) của 18 thùng chứa ta được bảng kết quả sau:

X	19,6	19,5	19,9	20,0	19,8	20,5	21,0	18,5	19,7
Số thùng	1	2	2	4	2	3	2	1	1

Với độ tin cậy 99% hãy tìm khoảng ước lượng đối xứng của áp lực trung bình của thùng trên. Biết rằng áp lực là ĐLNN có phân phối chuẩn.

10. 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 và thu được số liệu sau:

Giá (đồng) X	81	85	87	89	91	93	95	97	99	101
Số cửa hàng (m_i)	3	10	13	15	30	12	7	6	3	1

Với độ tin cậy 95% hãy ước lượng giá trung bình của loại hàng đó tại thời điểm đang xét bằng khoảng tin cậy đối xứng. Biết rằng giá của hàng hoá là đại lượng ngẫu nhiên tuân theo qui luật phân phối chuẩn.

11. Để xác định chiều cao trung bình của các cây bạch đàn, người ta tiến hành đo ngẫu nhiên 35 cây và có bảng số liệu:

Chiều cao (X-mét)	6,5-7,0	7,0-7,5	7,5-8,0	8,0-8,5	8,5-9,0	9,0-9,5
Số cây	2	4	10	11	5	3

Với độ tin cậy 95% có thể nói chiều cao trung bình của các cây đàn nằm trong khoảng nào. Giả thiết chiều cao của cây bạch đàn là ĐLNN tuân theo qui luật phân phối chuẩn.

12. Có số liệu về trọng lượng của loại trứng gà như ở bảng dưới đây. Bằng khoảng tin cậy đối xứng hãy ước lượng trọng lượng trung bình của loại trứng gà này với độ tin cậy 0,95. Giả thiết trọng lượng trứng gà là ĐLNN tuân theo qui luật phân phối chuẩn.

Trọng lượng (X-gam)	25-30	30-35	35-40	40-45	45-50
Số quả	2	3	10	8	2

13. Điều tra doanh số hàng tháng của 100 hộ kinh doanh một loại hàng, có bảng số liệu:

Doanh số (X-triệu đồng)	11,5	11,6	11,7	11,8	11,9	12,0
Số hộ tương ứng	10	15	20	30	15	10

Bằng khoảng tin cậy đối xứng hãy ước lượng doanh số trung bình hàng tháng của các hộ kinh doanh mặt hàng này với độ tin cậy 95%. Giả thiết doanh số là ĐLNN tuân theo qui luật phân phối chuẩn.

14. Đo độ chịu lực (kg/cm^2)của 200 mẫu bê tông người ta thu được kết quả trong bảng sau:

Độ chịu lực (X)	190-200	200-210	210-220	220-230	230-240	240-250
Số mẫubê tông	10	26	56	64	30	14

Hãy ước lượng độ chịu lực trung bình của bê tông với độ tin cậy 0,95. Biết rằng độ chịu lực của bê tông là ĐLNN tuân theo quy luật chuẩn.

15. Lấy 50 con sợi để xác định độ bền trung bình, ta có số liệu sau:

Độ bền (X-kg/cm^2)	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
Số con sợi	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 0,95. Giả thiết độ bền của sợi là ĐLNN tuân theo qui luật phân phối chuẩn.

16. Điều tra 365 điểm trồng lúa của một huyện có bảng số liệu:

Năng suất (X-ta/ha)	25	30	33	34	35	36	37	39	40
Số điểm trồng lúa	6	13	38	74	106	85	30	10	3

Với độ tin cậy 95% có thể nói năng suất lúa trung bình của huyện nằm trong khoảng nào. Giả thiết năng suất lúa là ĐLNN tuân theo qui luật phân phối chuẩn.

17. Đo đường kính của 20 chi tiết do một máy tiện sản xuất, ta có số liệu (tính bằng mm)

X	247	248	249	250	251	252	253	256	257	258	260
m_i	2	2	3	5	1	1	2	1	1	1	1

Giả thiết đường kính là ĐLNN có phân phối chuẩn

a/ Tìm khoảng ước lượng của độ dài trung bình của đường kính chi tiết với độ tin cậy 0,95.

b/ Các chi tiết có đường kính từ 249 đến 251 được coi là sản phẩm loại A. Hãy tìm khoảng ước lượng cho tỷ lệ loai A với độ tin cậy 0,95.

18. Tại một khu rừng nguyên sinh người ta đánh dấu 1000 con chim, sau đó thả chúng vào rừng. Một thời gian sau người ta bắt lại 200 con thấy có 40 có được đánh dấu. Với độ tin cậy 99% thử ước lượng số chim có trong khu rừng

Chương 3

Kiểm định giả thuyết thống kê

3.1. Khái niệm chung

3.1.1.Giả thuyết thống kê

Trong nhiều lĩnh vực đời sống kinh tế - xã hội chúng ta hay nêu ra các nhận xét khác nhau về các đối tượng quan tâm. Những nhận xét như vậy thường được coi là các giả thuyết, chúng cũng có thể đúng và cũng có thể sai. Vấn đề xác định đúng sai của một giả thuyết sẽ được gọi là **kiểm định**.

Trong thống kê chúng ta xuất phát từ một mẫu $x_1 \ldots x_n$ được chọn từ một tổng thể chưa biết phân phối hoặc biết được dạng phân phối nhưng chưa biết được tham số θ. Ta có thể phát biểu nhiều nhận xét khác nhau về các yếu tố chưa biết - đó là **giả thuyết thống kê**. Nếu tham số θ chưa biết và giả thuyết θ bằng giá trị cụ thể θ_0 được đưa ra, ta nói rằng có một giả thuyết đơn, nếu khác đi ta có giả thuyết phức.

Giả thuyết được đưa ra kiểm định được gọi là giả thuyết gốc và ký hiệu H_0; nó thường là giả thuyết đơn (trong các bài toán kiểm định tham số). Các giả thuyết khác với giả thuyết gốc gọi là đối thuyết ký hiệu H_1. Việc kiểm định một giả thuyết là đúng hay sai dựa trên thông tin mẫu được gọi là **kiểm định thống kê**

Định nghĩa 3.1.1. Giả thuyết thống kê là giả thuyết về dạng phân phối xác suất của ĐLNN, về các tham số đặc trưng của ĐLNN hoặc về tính độc lập của các ĐLNN.

3.1.2. Qui tắc kiểm định:

Nguyên tắc chung của kiểm định giả thuyết thống kê là dựa trên nguyên lý xác suất nhỏ: Một sự kiện có xác suất xuất hiện khác bé thì có thể coi rằng nó không xảy ra khi thực hiện một phép thử có liên quan đến sự kiện đó. Tuy nhiên trong thực tế, vấn đề phức tạp và tế nhị hơn nhiều.

a. Tiêu chuẩn kiểm định

Từ tổng thể rút ra mẫu, từ thông tin mẫu ta chọn một thống kê $G = f(X_1,.., X_n)$ có thể phụ thuộc vào tham số đã biết trong H_0. Nếu H_0 đúng thì qui luật của G phải hoàn toàn xác định. Một thống kê như vậy được gọi là tiêu chuẩn kiểm định.

Qui tắc kiểm định.

Nếu ta thành công trong việc chia miền xác định của tiêu chuẩn G thành 2 phần S_α và $\overline{S_\alpha}$, trong đó S_α là miền bác bỏ H_0 còn $\overline{S_\alpha}$ là miền chấp nhận H_0. Nếu G tính trên mẫu có giá trị thuộc miền S_α ta bác bỏ H_0 nếu ngược lại ta chấp nhận. Miền bác bỏ H_0 được gọi là miền tới hạn của tiêu chuẩn G.

Nếu dùng qui tắc trên có thể mắc 2 sai lầm

Sai lầm loại 1: Giả thuyết H_0 đúng mà ta lại bác bỏ

Sai lầm loại 2: Giả thuyết H_0 sai mà ta lại chấp nhận

Do giả thiết G có phân phối xác định khi H_0 đúng và nếu gọi α là xác suất để xảy ra sai lầm loại 1 thì $\alpha = P(G_{qs} \in S_\alpha / H_0$ đúng) (1). Trong đó G_{qs} chính là giá trị của G trên mẫu cụ thể đang xét. Tương tự nếu gọi β là xác suất phạm sai lầm loại 2 thì $\beta = P(G_{qs} \in \overline{S_\alpha} / H_0$ sai) (2).

Chú ý 3.1.2.

(a) Bác bỏ một giả thuyết chỉ có nghĩa là chấp nhận một giả thuyết khác chứ không có nghĩa giả thuyết bị bác bỏ là sai.

(b) Chấp nhận một giả thuyết chỉ có nghĩa là không chấp nhận các giả thuyết khác chứ không có nghĩa giả thuyết được chấp nhận là đúng.

(c) Ta mong muốn cả hai xác suất (1) và (2) càng bé càng tốt. Trong thực tế ta không thể đồng thời làm giảm cả 2 xác suất đó bởi vì cứ α giảm thì β tăng và ngược lại. Thông thường do sai lầm loại 1dễ kiểm soát và (1) dễ tính hơn nên người ta hay chọn trước α luôn là ngưỡng để xác suất phạm sai lầm loại 1 luôn nhỏ hơn α đủ bé đó. Các giá trị của α có thể là 0,1; 0,05; 0,01.. phụ thuộc vào yêu cầu của thực tế và nhà nghiên cứu. Giá trị α được gọi là mức ý nghĩa của qui tắc kiểm định

3.2. Kiểm định giả thuyết về giá trị trung bình

Giả sử từ tổng thể của ĐLNN X có phân phối chuẩn $N(a,\sigma^2)$ rút ra mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$. Căn cứ vào kinh nghiệm và đưa ra:

$H_0 : a = a_0$; $H_1: a \neq a_0$

hoặc $H_0 : a = a_0$; $H_1: a > a_0$

hoặc $H_0 : a = a_0$; $H_1: a < a_0$

3.2.1. Trường hợp đã biết phương sai

Chọn tiêu chuẩn kiểm định là thống kê $U = \dfrac{\overline{X} - a_0}{\sigma}\sqrt{n}$

Nếu H_0 đúng thì ta có $U = \dfrac{\overline{X} - a_0}{\sigma}\sqrt{n} = \dfrac{\overline{X} - a}{\sigma}\sqrt{n}$ và thống kê U phân phối chuẩn hoá N(0,1)

Từ mẫu cụ thể ta sẽ tìm được trung bình mẫu $\overline{x}$ và tính được giá trị quan sát của tiêu chuẩn kiểm định là U_{qs}

Với mức ý nghĩa α cho trước tra bảng (bảng phụ lục 3) ta sẽ tìm được các giá trị $u_{1-\frac{\alpha}{2}}$ hoặc $u_{1-\alpha}$

Tìm miền bác bỏ S:

**** Miền bác bỏ hai phía:*** (H_1: a ≠ a_0)

$$S=\left(-\infty; -u_{1-\frac{\alpha}{2}}\right)\cup\left(u_{1-\frac{\alpha}{2}}; +\infty\right)$$

Tức là nếu $|u_{qs}| \le u_{1-\frac{\alpha}{2}}$ thì chưa có cơ sở để bác bỏ H_0

Nếu $|u_{qs}| \ge u_{1-\frac{\alpha}{2}}$ thì bác bỏ H_0

**** Miền bác bỏ bên phải:*** H_1: a > a_0)

$$S = (u_{1-\alpha}; +\infty)$$

**** Miền bác bỏ bên trái:*** (H_1: a < a_0)

$$S = (-\infty; -u_{1-\alpha})$$

Ví dụ 3.2.1. Trọng lượng sản phẩm (X) do nhà máy sản xuất ra là một ĐLNN có phân phối chuẩn với độ lệch tiêu chuẩn σ = 2 kg và trọng lượng trung bình là 20kg. Nghi ngờ máy hoạt động không bình thường đã làm thay đổi trọng lượng trung bình của sản phẩm. Người ta cân thử 100 sản phẩm và thu được kêt quả sau:

Trọng lượngSP(X)	19	20	21	22	23
Số SP tương ứng	10	60	20	5	5

Với mức ý nghĩa α = 0,05 hãy cho kết luận về điều nghi ngờ trên?

Giải: Đây là bài toán kiểm định giả thuyết thống kê về giá trị trung bình a của đám đông có phân phối chuẩn khi biết phương sai.

Chọn giả thuyết: H_0: a = 20

Đối thuyết: H_1: a ≠ 20

Chọn tiêu chuẩn kiểm định u = $\dfrac{\overline{X}-20}{2}\sqrt{100}$

Trong đó $\overline{X}$ là trung bình mẫu ngẫu nhiên kích thước n = 100. Từ mẫu cụ thể trên ta tìm được $\overline{x}$ = 20,35 do đó giá trị quan sát của tiêu chuẩn kiểm định là u_{qs} = 1,75. Với mức ý nghĩa α = 0,05 thì $u_{1-\frac{\alpha}{2}} = u_{0,975}$ = 1,96.

Miền bác bỏ hai phía là: S = (- ∞ ; - 1,96) ∪ (1,96 ; +∞).

Ta thấy $u_{qs} \notin$ s do đó chưa có cơ sở để bác bỏ H_0 nghĩa là điều nghi ngờ trên là sai.

Ví dụ 3.2.2. Nếu máy móc hoạt động bình thường thì trọng lượng sản phẩm (X) là ĐLNN có phân phối chuẩn N(a,σ^2) với kỳ vọng toán a = 100 gam. Độ lệch tiêu chuẩn σ = 2 gam. Qua một thời gian sản xuất người ta nghi ngờ trọng lượng sản phẩm có xu hướng tăng lên, cân thử 100 sản phẩm thì trọng lượng trung bình sản phẩm của chúng là 100,4 gam. với mức ý nghĩa α = 0,05 hãy kết luận về điều nghi ngờ trên?

Giải: Đây là bài toán kiểm định giả thuyết thống kê về giá trị trung bình của tham số a của phân phối $N(a,\sigma^2)$ khi đã biết phương sai

Chọn giả thuyết: H_0: a = 100

Đối thuyết: H_1: a > 100

Chọn tiêu chuẩn kiểm định u = $\dfrac{\overline{X}-100}{2}\sqrt{100}$ trong đó $\overline{X}$ là trung bình mẫu ngẫu nhiên. Từ mẫu cụ thể $\bar{x}$ = 100,4 ta được u_{qs} = 2

với mức ý nghĩa α = 0,05 thì $u_{1-\alpha}$ = 1,645

Miền bác bỏ là S = (1,645 ; +∞). Ta thấy $u_{qs} \in$ S Vậy điều nghi ngờ trên là đúng tức là trọng lượng sản phẩm đã có xu hướng tăng lên.

3.2.2. Trường hợp chưa biết phương sai

Chọn tiêu chuẩn kiểm định là thống kê $T = \dfrac{\overline{X}-a_0}{S'}\sqrt{n}$

Nếu H_0 đúng thì ta có $T = \dfrac{\overline{X}-a_0}{S'}\sqrt{n} = \dfrac{\overline{X}-a}{S'}\sqrt{n}$

+ Nếu mẫu nhỏ thì thống kê T tuân theo qui luật Student với n-1 bậc tự do.

+ Nếu mẫu lớn thì phân phối Student tiến tới phân phối N(0,1) nên ta có thể xấp xỉ phân phối Student bằng phân phối chuẩn.

Từ mẫu cụ thể ta tìm được T_{qs}.

Với mức ý nghĩa α cho trước tra bảng phụ lục 5ta sẽ tìm được $t^{(n-1)}_{1-\frac{\alpha}{2}}$ hoặc $t^{(n-1)}_{1-\alpha}$ (đối với mẫu nhỏ) hoặc $u_{1-\frac{\alpha}{2}}$ hay $u_{1-\alpha}$ (đối với mẫu lớn). Từ đó sẽ tìm được miền bác bỏ S, so sánh và rút ra kết luận (tương tự như khi đã biết phương sai)

Ví dụ 3.2.3. Mức hao phí xăng (X) cho một loại xe ô tô chạy trên đoạn đường AB là ĐLNN có phân phối chuẩn với kỳ vọng toán là 50 lít. Do đường được tu sửa lại, người ta cho rằng mức hao phí xăng trung bình đã giảm xuống. Quan sát 30 chuyến xe chạy trên đoạn đường AB ta thu được bảng số liệu sau:

Mức hao phí X (lít)	48,5 49,0	49,0 49,5	49,5 50,0	50,0 50,5	50,5 51,0
Số chuyến	5	10	10	3	2

Với mức ý nghĩa α = 0,05, hãy kết luận về ý kiến nêu trên?

Giải: Đây là bài toán kiểm định giả thiết thống kê về giá trị trung bình của ĐLNN có phân phối chuẩn khi chưa biết phương sai với mẫu nhỏ.

Chọn giả thuyết: H_0: a = 50

Đối thuyết: H_1: a < 50

Chọn tiêu chuẩn kiểm định T = $\dfrac{\overline{X}-50}{S'}\sqrt{30}$

Từ mẫu cụ thể ta tìm được: $\overline{x}$ = 49,53 và s' = 0,55 thay các giá trị vừa tìm được vào thống kê trên ta được T_{qs} = - 4,68

Với mức ý nghĩa α = 0,05 tra bảng phụ lục 5 ta được $t_{0,95}^{(29)}$ =1,69

Miền bác bỏ bên trái là: S = $(-\infty ; -1,69)$. Ta thấy $T_{qs} \in S \Rightarrow$ bác bỏ H_0 nghĩa là điều nghi ngờ trên là đúng (mức hao phí xăng đã giảm xuống).

Chú ý 3.2.4. Nếu tổng thể của ĐLNN X không tuân theo qui luật phân phối chuẩn thì ta có thể tiến hành chọn mẫu có kích thước lớn (n > 30) khi đó ta có thể tiến hành kiểm định tương tự như tiến hành kiểm định đối với ĐLNN có phân phối chuẩn.

3.3. Kiểm định sự bằng nhau của hai kỳ vọng

Giả sử có hai tổng thể nghiên cứu, trong tổng thể thứ nhất ĐLNN gốc X_1 phân phối chuẩn N(a_1, σ_1^2), trong tổng thể thứ hai ĐLNN gốc X_2 phân phối N(a_2, σ_2^2). Nếu a_1 và a_2 chưa biết song có cơ sở để giả thiết rằng giá trị của chúng bằng nhau, người ta đưa ra giả thuyết thống kê: H_0: $a_1 = a_2$ và đối thuyết: H_1: $a_1 \neq a_2$ (hoặc H_1: $a_1 > a_2$, hoặc H_1: $a_1 < a_2$)

Để kiểm định giả thuyết trên, từ hai tổng thể nói trên rút ra hai mẫu ngẫu nhiên độc lập kích thước tương ứng là n_1 và n_2:

$W_1 = (X_{11}, X_{12}, \ldots, X_{1n})$

$W_2 = (X_{21}, X_{22}, \ldots, X_{2n})$

Ta xét hai trường hợp sau:

3.3.1. Trường hợp 1. Biết phương sai σ_1^2 và σ_2^2 của các ĐLNN gốc trong tổng thể

Chọn tiêu chuẩn kiểm định: G = U = $\dfrac{(\overline{X_1} - \overline{X_2}) - (a_1 - a_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$.

Ta thấy thống kê U phân phối N(0,1).

Nếu giả thuyết H_0 đúng thì thống kê U có dạng:

$$U = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \text{ và cũng phân phối N(0,1)}$$

Từ mẫu cụ thể ta sẽ tìm được các giá trị $\overline{x_1}$, và $\overline{x_2}$ cụ thể thay vào ta sẽ tìm được giá trị quan sát là U_{qs}.

Với mức ý nghĩa cho trước tra bảng ta sẽ tìm được giá trị $u_{1-\frac{\alpha}{2}}$ hoặc $u_{1-\alpha}$ từ đó ta tìm được miền bác bỏ S

**** Miền bác bỏ hai phía:*** S = $(-\infty ; -u_{1-\frac{\alpha}{2}}) \cup (u_{1-\frac{\alpha}{2}} ; +\infty)$.

**** Miền bác bỏ bên phải:*** S = ($u_{1-\alpha}$; +∞).

**** Miền bác bỏ bên trái:*** S = (- ∞ ; - $u_{1-\alpha}$).

So sánh U_{qs} với miền bác bỏ S rồi rút ra kết luận.

3.3.2. Trường hợp 2. Chưa biết phương sai σ_1^2 và σ_2^2 của các ĐLNN gốc trong tổng thể: Với kích thước mẫu n_1 và n_2 đủ lớn ta chọn tiêu chuẩn kiểm định là thống kê:

$$G = U = \frac{\left(\overline{X_1} - \overline{X_2}\right) - \left(a_1 - a_2\right)}{\sqrt{\dfrac{S_1'^2}{n_1} + \dfrac{S_2'^2}{n_2}}}$$

Ta thấy U phân phối xấp xỉ N(0,1). Nếu giả thuyết H_0 đúng, tiêu chuẩn kiểm định có dạng $U = \dfrac{\left(\overline{X_1} - \overline{X_2}\right)}{\sqrt{\dfrac{S_1'^2}{n_1} + \dfrac{S_2'^2}{n_2}}}$ và cũng phân phối N(0,1), Từ mẫu cụ thể ta sẽ tìm được các giá trị $\overline{x_1}$, và $\overline{x_2}$, $S_1'^2$, $S_2'^2$ cụ thể thay vào thống kê trên ta sẽ tìm được U_{qs}, vì vậy với mức ý nghĩa cho trước tra bảng ta sẽ tìm được giá trị $u_{1-\frac{\alpha}{2}}$ hoặc $u_{1-\alpha}$ từ đó ta tìm được miền bác bỏ S:

**** Miền bác bỏ hai phía:*** S = (- ∞ ; -$u_{1-\frac{\alpha}{2}}$) $\cup$ ($u_{1-\frac{\alpha}{2}}$; +∞)

**** Miền bác bỏ bên phải:*** S = ($u_{1-\alpha}$; +∞)

**** Miền bác bỏ bên trái:*** S = (- ∞ ; -$u_{1-\alpha}$)

So sánh U_{qs} với miền bác bỏ S rồi rút ra kết luận.

Ví dụ 3.3.1. Biết trọng lượng sản phẩm do hai máy sản xuất ra là ĐLNN có phân phối chuẩn với N(a_1,1) và N(a_2,1) với α = 0,05. Có thể xem trọng lượng trung bình của sản phẩm do hai máy sản xuất ra là như nhau hay không? nếu kiểm tra ngẫu nhiên n_1 = 25 sản phẩm do máy 1 sản xuất ra thu được $\overline{x_1}$ = 50 kg và n_2 = 20 sản phẩm do máy 2 sản xuất ra thu được $\overline{x_2}$ = 50,6 kg.

Giải: Đây là bài toán kiểm định sự bằng nhau của 2 kỳ vọng của 2 ĐLNN có phân phối chuẩn khi đã biết phương sai.

Chọn giả thuyết H_0: $a_1 = a_2$

Đối thuyết H_1: $a_1 \neq a_2$

Tiêu chuẩn kiểm định là: $U = \dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

Từ mẫu cụ thể ta tìm được giá trị quan sát $U_{qs} = \dfrac{50 - 50,6}{\sqrt{\dfrac{1}{25} + \dfrac{1}{20}}} = -2$

Với α = 0,05 tra bảng ta được $u_{1-\frac{\alpha}{2}}$ = $u_{0,975}$ = 1,96

Miền bác bỏ hai phía là: S = $(-\infty; -1,96) \cup (1,96; +\infty)$

Ta thấy U_{qs} thuộc miền bác bỏ. Vậy bác bỏ H_0

Kết luận: Không thể coi trọng lượng sản phẩm do 2 máy sản xuất ra là như nhau được.

Ví dụ 3.3.2. Nếu áp dụng biện pháp kỹ thuật thứ nhất thì khi điều tra ngẫu nhiên n_1 = 100 thửa ruộng trồng giống lúa A thu được năng suất trung bình $\overline{x_1}$ = 100 tạ/ha và $s'_1 = 10$ tạ/ha. Còn nếu áp dụng biện pháp kỹ thuật thứ 2 thì khi điều tra ngẫu nhiên n_2 = 50 thửa ruộng thu được năng suất trung bình $\overline{x_2}$ = 95 tạ/ha và $s'_1 = 9$ tạ/ha. Hãy kết luận với mức ý nghĩa α = 0,05 nếu áp dụng biện pháp kỹ thuật thứ nhất thì năng suất giống lúa A cao hơn thực sự so với kết quả áp dụng biện pháp kỹ thuật thứ hai không? giả thiết năng suất lúa tuân theo qui luật chuẩn.

Giải: Gọi X_1, X_2 tương ứng là năng suất của giống lúa A khi áp dụng biện pháp kỹ thuật thứ nhất và thứ hai, a_1, a_2 là năng suất trung bình tương ứng khi áp dụng các biện pháp đó.

Đây là bài toán kiểm định sự bằng nhau của hai kỳ vọng của hai ĐLNN có phân phối chuẩn khi chưa biết phương sai với mẫu lớn.

Chọn giả thuyết H_0: $a_1 = a_2$

Đối thuyết H_1: $a_1 > a_2$

Nếu H_0 đúng thì tiêu chuẩn kiểm định có dạng U = $\dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{S_1'^2}{n_1} + \dfrac{S_2'^2}{n_2}}}$

Từ mẫu cụ thể với các giá trị tương ứng tìm được thay vào thống kê trên ta tìm được giá trị quan sát U_{qs} = 3,09

Với α = 0,05 tra bảng ta được $u_{0,95}$ = 1,645 ta tìm được miền bác bỏ bên phải là: S = $(1,645, +\infty)$; $U_{qs} \in S \Rightarrow$ bác bỏ H_0.

3.4. Kiểm định giả thuyết của xác suất

3.4.1. Trường hợp một tổng thể

Giả sử trong tổng thể của ĐLNN X có xác suất xuất hiện biến cố A là p, nếu chưa biết p song có cơ sở để giả thuyết rằng giá trị của nó bằng p_0

Chọn giả thuyết H_0: $p = p_0$

Đối thuyết H_1: $p \neq p_0$ (hoặc $p > p_0$, hoặc $p < p_0$)

Từ tổng thể lập mẫu ngẫu nhiên kích thước n, với n đủ lớn ta chọn tiêu chuẩn kiểm định là thống kê G = U = $\frac{f-p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$

Nếu H_0 đúng thì U = $\frac{f-p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$ = $\frac{f-p}{\sqrt{p(1-p)}}\sqrt{n}$ phân phối xấp xỉ N(0,1)

Từ mẫu cụ thể ta tính được giá trị quan sát U_{qs} với mức ý nghĩa cho trước tra bảng ta sẽ tìm được giá trị $u_{1-\frac{\alpha}{2}}$ hoặc $u_{1-\alpha}$ từ đó ta tìm được miền bác bỏ S

**** Miền bác bỏ hai phía:*** S = $(-\infty; -u_{1-\frac{\alpha}{2}}) \cup (u_{1-\frac{\alpha}{2}}; +\infty)$

**** Miền bác bỏ bên phải:*** S = $(u_{1-\alpha};+\infty)$

**** Miền bác bỏ bên trái:*** S = $(-\infty; -u_{1-\alpha})$

So sánh U_{qs} với miền bác bỏ s rồi rút ra kết luận.

3.4.2. Trường hợp hai tổng thể

Giả sử trong tổng thể thứ nhất, ĐLNN gốc X_1 phân phối $A(p_1)$, ĐLNN gốc X_2 phân phối $A(p_2)$.

Nếu p_1 và p_2 chưa biết song có cơ sở để giả thuyết rằng giá trị của chúng bằng nhau ta đưa ra giả thuyết thống kê

H_0: $p_1 = p_2$

H_1: $p_1 \neq p_2$, (hoặc $p_1 < p_2$; hoặc $p_1 > p_2$)

Để kiểm định giả thuyết trên, từ tổng thể nói trên rút ra hai mẫu ngẫu nhiên độc lập kích thước tương ứng là n_1 và n_2:

$W_1 = (X_{11}, X_{12}, \ldots, X_{1n_1})$ và $W_2 = (X_{21}, X_{22}, \ldots, X_{2n_2})$

Chọn tiêu chuẩn kiểm định là G = U = $\frac{(f_1-f_2)-(P_1-P_2)}{\sqrt{f(1-f)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$, trong đó f_1 và f_2 là các tần suất mẫu tương ứng và f = $\frac{n_1 f_1 + n_2 f_2}{n_1+n_2}$

Với n_1 và n_2 khá lớn thì thống kê U nói trên sẽ phân phối xấp xỉ N(0,1)

Nếu giả thuyết H_0 đúng thì tiêu chuẩn kiểm định có dạng:

U = $\frac{f_1-f_2}{\sqrt{f(1-f)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ và vẫn phân phối xấp xỉ N(0,1).

Từ mẫu cụ thể ta tính được các giá trị cụ thể f_1, f_2 và f và giá trị quan sát U_{qs} với mức ý nghĩa cho trước tra bảng ta sẽ tìm được giá trị $u_{1-\alpha/2}$ hoặc $u_{1-\alpha}$ từ đó ta tìm được miền bác bỏ.

**** Miền bác bỏ hai phía:*** S = $(-\infty; -u_{1-\alpha/2}) \cup (u_{1-\alpha/2}; +\infty)$

**** Miền bác bỏ bên phải:*** S = $(u_{1-\alpha};+\infty)$

**** Miền bác bỏ bên trái:*** S = (- ∞ ; $-u_{1-\alpha}$)

So sánh U_{qs} với miền bác bỏ S rồi rút ra kết luận.

Ví dụ 3.4.1. Tỷ lệ phế phẩm do một máy tự động sản xuất ra là 5%. Kiểm tra ngẫu nhiên 300 sản phẩm thấy có 24 sản phẩm là phế phẩm. Từ đó có ý kiến cho rằng tỷ lệ phế phẩm do máy đó sản xuất có chiều hướng tăng lên. Hãy kết luận ý kiến nêu trên với mức ý nghĩa α = 0,05.

Giải: Gọi p là tỷ lệ sản phẩm là phế phẩm. Đây là bài toàn kiểm định giả thuyết thống kê về tỷ lệ

Chọn giả thuyết thống kê H_0: p = 0,05

Đối thuyết H_1: p > 0,05

Chọn thống kê $U = \dfrac{f-0,05}{\sqrt{0,05.0,95}}\sqrt{300}$, trong đó f là tần suất của mẫu ngẫu nhiên n = 300

Với mức ý nghĩa α = 0,05. tra bảng ta được $u_{0,95}$ = 1,645

Vậy miền bác bỏ bên phải là: S = (1,645 ; +∞)

Từ mẫu cụ thể ta tìm được f = 0,08 và U_{qs} = 2,38

Ta thấy $U_{qs} \in S \Rightarrow$ bác bỏ H_0 nghĩa là tỷ lệ phế phẩm do máy đó sản xuất ra đã tăng lên thật sự.

Ví dụ 3.4.2. Thống kê số tai nạn lao động tại hai xí nghiệp có các số liệu sau:

Xí nghiệp	Số công nhân	Số tai nạn lao động
I	200	20
II	800	120

Với mức ý nghĩa α = 0,05 hãy kết luận xem chất lượng của công tác bảo hộ lao động tại hai xí nghiệp trên có khác nhau không?

Giải: Gọi p_1, p_2 tương ứng là tỷ lệ bị tai nạn lao động của hai xí nghiệp I và II

Đây là bài toán kiểm định giả thuyết thống kê về tỷ lệ của hai tổng thể

Chọn giả thuyết H_0: $p_1 = p_2$

Đối thuyết H_1: $p_1 \neq p_2$

Chọn tiêu chuẩn kiểm định $U = \dfrac{f_1 - f_2}{\sqrt{f(1-f)\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$

Từ mẫu cụ thể ta tìm được f_1 = 0,1; f_2 = 0,15; f = 0,14; U_{qs} = - 1,85

với mức ý nghĩa α = 0,05 tra bảng ta được $u_{1-\frac{\alpha}{2}} = u_{0,975}$ = 1,96

Vậy miền bác bỏ hai phía là (-∞ ; - 1,96) ∪ (1,96 ; +∞)

Ta thấy $U_{qs} \notin S$

Kết luận: chưa có cơ sở để bác bỏ H_0 nghĩa là công tác bảo hộ lao động tại hai xí nghiệp là như nhau.

BÀI TẬP CHƯƠNG 3

1. Hàm lượng đường trung bình của một loại trái cây lúc đầu là 5%. Người ta chăm bón bằng một loại phân N và sau một thời gian kiểm tra một số trái cây được kết quả sau:

Hàm lượng X(%)	1-5	5-9	9-13	13-17	17-21	21-25	25-29	29-33	37-41
Số trái	51	47	39	36	32	8	7	3	2

Hãy cho kết luận về loại phân N trên với mức ý nghĩa 5%. Giả thiết hàm lượng đường của loại trái cây trên là ĐLNN tuân theo quy luật chuẩn.

2. Đo chỉ số mỡ sữa của 130 con bò lai Hà - Ấn F_1 ta được bảng số liệu sau:

Chỉ số mỡ sữa (X)	3,0 3,6	3,6 4,2	4,2 4,8	4,8 5,4	5,4 6,0	6,0 6,6	6,6 7,2
Số bò lai	2	8	35	43	22	15	5

Biết rằng chỉ số mỡ sữa trung bình của giống bò lai thuần chủng là 4,95. Với mức ý nghĩa 1% hãy cho kết luận về hiệu quả của việc lai giống.

3. Định mức thời gian hoàn thành một sản phẩm là 14 phút. Có cần thay đổi định mức không, nếu theo dõi thời gian hoàn thành một sản phẩm ở 25 công nhân ta thu được bảng số liệu sau:

Thời gian để SX 1 sản phẩm (phút)	10-12	12-14	14-16	16-18	20-22
Số công nhân tương ứng	3	6	10	4	2

Yêu cầu kết luận với mức ý nghĩa $\alpha = 0,05$ biết rằng thời gian hoàn thành một sản phẩm là đại lượng ngẫu nhiên có phân phối chuẩn.

4. Định mức cũ để sản xuất một sản phẩm là 20 phút. Nay do cải tiến kỹ thuật, người ta sản xuất thử 100 sản phẩm và thu được số liệu:

Thời gian sản xuất 1 sản phẩm (X - phút)	16-17	17-18	18-19	10-20	20-21	21-22
Số sản phẩm tương ứng	6	10	24	30	18	12

Với mức ý nghĩa $\alpha = 0,05$ có thể nói rằng việc cải tiến kỹ thuật giảm bớt thời gian sản xuất một sản phẩm hay không? Biết rằng thời gian sản xuất một sản phẩm là đại lượng ngẫu nhiên tuân theo qui luật phân phối chuẩn.

5. Mức hao phí xăng (X) cho một loại xe ôtô trên đoạn đường AB là đại lượng ngẫu nhiên phân phối chuẩn với kỳ vọng là 50 lít. Do đoạn đường được tu sửa lại, người ta cho rằng mức hao phí xăng trung bình đã giảm xuống. Quan sát 100 chuyến xe chạy trên đoạn đường AB thu được bảng số liệu:

Mức xăng hao phí (lít)	48,5-49,0	49,0-49,5	49,5-50,0	50,0-50,5	50,5-51,0
Số chuyến xe	15	17	40	18	10

Với mức ý nghĩa $\alpha = 0,05$, hãy kết luận về ý kiến nêu trên.

6. Kiểm tra các gói đường loại 1kg trong một siêu thị ta có kết quả:

Khối lượng (X-kg)	0,95	0,96	0,97	0,99	1,00	1,01	1,03	1,05
Số gói	19	30	32	8	2	3	5	1

Với mức ý nghĩa $\alpha = 0{,}05$ có thể kết luận việc đóng gói đảm bảo yêu cầu hay không. Biết rằng khối lượng các gói đường là đại lượng ngẫu nhiên tuân theo qui luật phân phối chuẩn.

7. Sản phẩm của một xí nghiệp đúc cho phép số khuyết tật trung bình của một sản phẩm là 3. Sau khi đổi mới thiết bị, kiểm tra ngẫu nhiên 36 sản phẩm kết quả thu được:

Số khuyết tật trên 1 sản phẩm	0	1	2	3	4	5	6
Số sảm phẩm	7	4	4	6	8	6	1

a/ Với độ tin cậy 95% hãy ước lượng số khuyết tật trung bình của mỗi sản phẩm sau khi đổi mới thiết bị.

b/ Kết luận về hiệu quả việc đổi mới thiết bị với mức ý nghĩa 5%

8. Kiểm tra chất lượng hai lô sản phẩm từ 2 cơ sở chuyển đến ta thấy: Trong 120 sản phẩm ở lô I có 70 sản phẩm loại A. Còn trong 150 sản phẩm ở lô II có 98 sản phẩm loại A. Hỏi với mức ý nghĩa 1% có thể coi hai nguồn hàng có cùng tỉ lệ hàng loại A hay không?

9. Điều tra về số người mắc bệnh bứu cổ ở một tỉnh phía Bắc thấy có 107 người bị bệnh trong 380 người đến khám. Trong khi ở một tỉnh miền Trung có 90 người trong số 310 người khám bệnh. Có thể kết luận về tỉ lệ mắc bệnh ở hai tỉnh trên là như nhau không, với mức ý nghĩa 5%.

Chương 4

Tương quan và hồi quy

Khi nghiên cứu sự phụ thuộc giữa hai đại lượng mà mỗi đại lượng chịu một sự phân tán ngẫu nhiên (sự tản mát kiểm tra được) ta dùng phương pháp phân tích tương quan. Phân tích tương quan không những phát hiện được mối quan hệ phụ thuộc giữa chúng mà còn "lượng hoá" được mối quan hệ này.

4.1. Hệ số tương quan

4.1.1. Phân tích hệ số tương quan

Định nghĩa 4.1.1. Giả sử X và Y là hai đại lượng ngẫu nhiên có D(X) > 0 và D(Y) > 0. Hệ số tương quan của hai đại lượng ngẫu nhiên X và Y, ký hiệu ρ(X,Y) được xác định như sau: $\rho(X,Y) = \dfrac{E[X - E(X)][Y - E(Y)]}{\sqrt{D(X)}\sqrt{D(Y)}}$

4.1.2. Tính chất của hệ số tương quan

(a) ρ(X,Y) = ρ(Y,X)

(b) ρ(X,Y) = 0 nếu X và Y độc lập với nhau

(c) $|\rho(X,Y)| \le 1$

Nhận xét 4.1.2.

- Ta có thể dùng ρ(X,Y) để đo mức độ phụ thuộc tuyến tính giữa hai ĐLNN. $|\rho|$ càng lớn thì sự phụ thuộc tuyến tính càng rõ. Đặc biệt nếu ρ =±1 thì sự phụ thuộc tuyến tính xẩy ra với xác suất bằng 1

- Nếu $|\rho|$ càng nhỏ thì mức độ phụ thuộc tuyến tính giữa X và Y càng ít. Đặc biệt khi ρ = 0 thì giữa X và Y không có quan hệ phụ thuộc tuyến tính. Trong trường hợp này ta nói X và Y không tương quan với nhau (chú ý: hai ĐLNN độc lập với nhau thì không tương quan, nhưng điều ngược lại chưa chắc đã đúng. Riêng đối với những ĐLNN phân phối chuẩn thì tính không tương quan và tính độc lập là tương đương nhau.)

4.1.3. Hệ số tương quan mẫu

Giả sử tiến hành n quan sát độc lập đối với cặp ĐLNN (X,Y) ta có mẫu ngẫu nhiên kích thước n sau đây: $\{(X_1,Y_1), (X_2,Y_2), \ldots,(X_n,Y_n)\}$

Hệ số tương quan mẫu của X và Y, ký hiệu r(X,Y) = r được xác định bằng công thức sau đây:

$$r = \frac{\frac{1}{n}\sum_i (X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\frac{1}{n}\sum_i (X_i-\overline{X})^2}\sqrt{\frac{1}{n}\sum_i (Y_i-\overline{Y})^2}} = \frac{\overline{XY}-\overline{X}.\overline{Y}}{S_X.S_Y}$$

Nếu trong mẫu cụ thể các x_i và y_i cách điều nhau một khoảng h_x và h_y thì ta dùng phép đổi biến với x_0 và y_0 được chọn thích hợp tuỳ theo bảng số liệu

đặt: $u_i = \frac{x_i - x_0}{h_x}$ và $v_i = \frac{y_i - y_0}{h_y}$, khi đó

$$r = \frac{\sum_i u_i v_i - n\overline{u}\,\overline{v}}{\sqrt{\sum_i u_i^2 - n(\overline{u})^2}\sqrt{\sum_i v_i^2 - n(\overline{v})^2}} = \frac{\overline{uv} - \overline{u}\,\overline{v}}{\sqrt{\overline{u^2}-(\overline{u})^2}\sqrt{\overline{v^2}-(\overline{v})^2}}$$

Chú ý 4.1.3. Sau khi đổi biến khi muốn trở về biến cũ thì:

$\overline{x} = h_x\overline{u} + x_0$; $s_x = h_x s_u$

$\overline{y} = h_y\overline{v} + y_0$; $s_y = h_y s_v$

Nhận xét 4.1.4.

(a) $|r| \le 1$

r > 0 thì X và Y là tương quan thuận

r < 0 thì X và Y là tương quan nghịch

(b) $0,7 \le r \le 1$ thì X và Y là tương quan thuận và mạnh

$-1 \le r \le -0,7$ thì X và Y là tương quan nghịch và mạnh

$|r| \le 0,3$ thì X và Y là tương quan yếu.

4.2. Hồi quy tuyến tính đơn giản

4.2.1. Phương trình hồi quy tuyến tính đơn giản của tổng thể

Khi có sự phụ thuộc tuyến tính giữa hai biến ngẫu nhiên tương đối chặt chẽ ta có thể hy vọng xấp xỉ biến này bởi một hàm tuyến tính của biến kia. Nghĩa là cần tìm biểu thức aX + b sao cho xấp xỉ Y tốt nhất theo nghĩa cực tiểu sai số bình phương trung bình $E(Y-aX-b)^2$.

Ta có: $E(Y-aX-b)^2 = E\{(Y-E(y))-a(X-E(X))+E(Y)-aE(X)-b\}^2$

$$= E(Y-E(Y))^2 + a^2E(X-E(X))^2 + (E(Y)-aE(X)-b)^2 - 2aE\{(Y-EY)(X-EX)\}^2$$

Vế phải sẽ đạt cực tiểu nếu và chỉ nếu tam thức bậc hai theo a: $a^2D(X) - 2a\rho\sqrt{DX}\sqrt{DY} + DY$ đạt cực tiểu và số hạng $(EY-aEX-b)^2 = 0$. Do đó ta chọn b = E(Y) - aE(X) còn a là tọa độ đỉnh của tam thức bậc hai:

$$a = -\frac{-2\rho\sqrt{D(X)}\sqrt{D(Y)}}{2D(X)} = \rho\sqrt{\frac{D(Y)}{D(X)}}$$

Khi đó giá trị nhỏ nhất của vế phải chính là giá trị của tam thức bậc hai theo a tại đỉnh của nó:

$$\min E(Y - aX - b)^2 = \rho^2 \frac{DY}{DX} DX - 2\rho\sqrt{\frac{DY}{DX}}\rho\sqrt{DX}\sqrt{DY} + DY = DY(1-\rho^2)$$

Vậy biểu thức aX + b cần tìm chính là $\rho\sqrt{\frac{DY}{DX}}X + EY - \rho\sqrt{\frac{DY}{DX}}EX$

Phương trình đường hồi qui bình phương trung bình tuyến tính của Y theo X là:

$$Y = \rho\sqrt{\frac{DY}{DX}}X + EY - \rho\sqrt{\frac{DY}{DX}}EX \Leftrightarrow Y = \rho\sqrt{\frac{DY}{DX}}(X - EX) + EY \quad (1)$$

Sai số bình phương trung bình khi dùng đường hồi qui trung bình tuyến tính để xấp xỉ Y là: $\sigma^2_{y/x} = DY(1-\rho^2)$ (2)

Sai số này càng nhỏ khi $|\rho|$ càng gần 1 tức là mức độ phụ thuộc tuyến tính giữa hai biến càng chặt.

Tương tự phương trình đường hồi qui bình phương trung bình tuyến tính của X theo Y là:

$$X = \rho\sqrt{\frac{DX}{DY}}Y + EX - \rho\sqrt{\frac{DX}{DY}}EY \Leftrightarrow X = \rho\sqrt{\frac{DX}{DY}}(Y - EY) + EX \quad (3)$$

Sai số bình phương trung bình khi dùng đường hồi qui trung bình tuyến tính để xấp xỉ X là: $\sigma^2_{x/y} = DX(1-\rho^2)$ (4)

Chú ý 4.2.1. Khi có mẫu ngẫu nhiên (x_i, y_i); i = 1..n ta xây dựng đường hồi qui trung bình tuyến tính thực nghiệm bằng cách thay trong (1) và (3) EY bởi $\overline{Y}$; EX bởi $\overline{X}$ và $\rho\sqrt{\frac{DY}{DX}}$ bởi $r\frac{s_y}{s_x}$; $\rho\sqrt{\frac{DX}{DY}}$ bởi $r\frac{s_x}{s_y}$.

4.2.2. Phương trình đường hồi quy tuyến tính thực mẫu

- Phương trình đường hồi qui bình phương trung bình tuyến tính thực nghiệm của Y theo X là

$$y = r\frac{S_y}{S_x}\left(x - \overline{x}\right) + \overline{y}$$

- Phương trình đường hồi qui bình phương trung bình tuyến tính thực nghiệm của X theo Y là

$$x = r\frac{S_x}{S_y}\left(y - \overline{y}\right) + \overline{x}$$

* Ước lượng sai số bình phương trung bình

$$s^2_{y/x} = s^2_y(1-r^2);\; s^2_{x/y} = s^2_x(1-r^2)$$

Ví dụ 4.2.1. Tính hệ số tương quan và phương trình đường hồi qui trung bình tuyến tính thực nghiệm của Y đối với X dựa vào số liệu cho trong bảng tương quan sau:

x_i	1	2	2	3	3	4
y_i	3	4	5	5	6	7
n_i	3	2	1	2	1	1

Giải:

Ta thiết lập lại bảng số liệu:

x_i	y_i	n_i	n_ix_i	n_iy_i	$n_ix_i^2$	$n_iy_i^2$	$n_ix_iy_i$
1	3	3	3	9	3	27	9
2	4	2	4	8	8	32	16
2	5	1	2	5	4	25	10
3	5	2	6	10	18	50	30
3	6	1	3	6	9	36	18
4	7	1	4	7	16	49	28
		$\sum 10$	$\sum 22$	$\sum 45$	$\sum 58$	$\sum 219$	$\sum 111$

Nhìn vào bảng số liệu trên ta có:

$\bar{x} = 2{,}2$; $\bar{y} = 4{,}5$; $\overline{x^2} = 5{,}8$; $\overline{y^2} = 21{,}9$; $\overline{xy} = 11{,}1$; $\overline{xy} - \bar{x}\bar{y} = 1{,}2$

$s_x^2 = 5{,}8 - (2{,}2)^2 = 0{,}96$; $s_y^2 = 21{,}9 - (4{,}5)^2 = 1{,}645$; $s_x s_y = 1{,}26$

Hệ số tương quan mẫu là: $r = \dfrac{1{,}2}{1{,}26} = 0{,}952$

Phương trình hồi qui tuyến tính thực nghiệm của Y theo X là:

$$y = r\frac{s_y}{s_x}\left(x - \bar{x}\right) + \bar{y} = 0{,}952\frac{1{,}28}{0{,}98}\left(x - 2{,}2\right) + 4{,}5 = 1{,}24x + 1{,}76$$

Ví dụ 4.2.2.

Ở một vùng có nghề phụ thủ công, quan sát 10 gia đình về 2 tiêu thức: Số trẻ em dưới 16 tuổi (X) và thu nhập thêm bằng nghề phụ (Y đơn vị nghìn đồng) thu được số liệu sau:

Gia đình	A	B	C	D	E	G	H	I	K	L
Số trẻ em dưới 16 tuổi (X)	3	5	2	4	4	4	6	1	3	3
Thu nhập Y (Nghìn đồng)	58	89	72	71	68	64	98	49	59	62

a/ Hãy khảo sát mối tương quan giữa hai tiêu thức trên

b/ Xây dựng đường hồi qui bình phương trung bình tuyến tính thực nghiệm của thu nhập theo số trẻ em

c/ Ước lượng sai số bình phương trung bình

Giải: a/ Ta thấy ở đây m_i =1 với mọi i. Đặt $u_i = x_i - 4$ ta có bảng số liệu sau:

x_i	y_i	u_i	u_i^2	y_i^2	u_iy_i
3	58	-1	1	3364	-58
5	89	1	1	7921	89
2	72	-2	4	5184	-144
4	71	0	0	5041	0
4	68	0	0	4624	0
4	64	0	0	4096	0
6	98	2	4	9604	196
1	49	-3	9	2401	-147
3	59	-1	1	3481	-59
3	62	-1	1	3844	-62
$\sum$	690	-5	21	49560	-185

Từ bảng trên ta thấy:

$$\overline{u} = -5/10 = -0{,}5 \qquad \Rightarrow \quad \overline{u}^2 = 0{,}25 \qquad \Rightarrow \quad \overline{x} = 3{,}5$$

$$\overline{y} = 690/10 = 69 \qquad \overline{y}^2 = 4761$$

$$\overline{u^2} = 21/10 = 2{,}1 \qquad \overline{y^2} = 49560/10 = 4956$$

$$\overline{uy} = -185/10 = -18{,}5$$

$$s_u = \sqrt{2{,}1 - 0{,}25} \approx 1{,}36 = s_x \qquad s_y = \sqrt{4956 - 4761} \approx 13{,}964$$

$$\Rightarrow \qquad r = \frac{-18{,}5 - (-0{,}5)\times 69}{1{,}36 \times 13{,}964} \approx 0{,}842$$

Điều đó chứng tỏ có sự phụ thuộc tuyến tính chặt chẽ giữa thu nhập và số trẻ em. Sự phụ thuộc là đồng biến.

b/ Phương trình đường hồi qui bình phương tuyến tính thực nghiệm của thu nhập theo số trẻ là:

$$y = 0{,}842 \times \frac{13{,}964}{1{,}36}(x - 3{,}5) + 69$$

$$y = 8{,}6446x + 39{,}256$$

c/ Ước lượng sai số bình phương trung bình $s^2_{y/x} = 195 \times (1 - 0{,}842^2) = 56{,}745$

Ví dụ 4.2.3. Nghiên cứu mối liên hệ giữa X là số tiền đầu tư cho việc phòng bệnh tính theo đầu người và Y là tỷ lệ người mắc bệnh ở 50 địa phương, ta thu được bảng tương quan thực nghiệm sau đây: (X: đơn vị nghìn đồng; Y: tính theo phần trăm).

X \ Y	2	2,5	3	3,5	4
100				2	3

200			3	6	2
300		4	6	3	
400	1	6	4	1	
500	6	3			

a/ Tìm hệ số tương quan tuyến tính.

b/ Tìm phương trình hồi qui tuyến tính của Y đối với X qua mẫu trên.

c/ Nếu năm sau đầu tư cho phòng bệnh là 600.000 đồng/người thì tỷ lệ mắc bệnh khoảng bao nhiêu phần trăm?

Giải: Đặt $u_i = \dfrac{x_i - 300}{100}$; $v_j = \dfrac{y_j - 3}{0,5}$

Ta có bảng số liệu sau:

V \ U	-2	-1	0	1	2	n_i	$n_i u_i$	$n_i u_i^2$
-2				2 / -4	3 / -12	5	-10	20
-1			3 0	6 / -6	2 / -4	11	-11	11
0		4 0	6 0	3 0		13	0	0
1	1 / -2	6 / -6	4 0	1 / 1		12	12	12
2	6 / -24	3 / -6				9	18	36
n_i	7	13	13	12	5	50	9	79
$n_i v_i$	-14	-13	0	12	10	-5		
$n_i v_i^2$	28	13	0	12	20	73		-63

$\sum n_i u_i = 9; \sum n_j v_j = -5; \quad \sum n_i u_i^2 = 79; \quad \sum n_j v_j^2 = 73; \quad \sum nuv = -63$

$\overline{u} = 0,18; \ \overline{v} = -0,1; \ \overline{u^2} = 1,58; \ \overline{v^2} = 1,46; \ \overline{uv} = -1,26; \quad s_u = 1,244; s_v = 1,204$

a/ $r = \dfrac{\overline{uv} - \overline{u}.\overline{v}}{s_u s_v}$ = - 0,829

b/ $\overline{x} = 318; \quad \overline{y} = 2,95; \quad s_x = 124,4; \quad s_y = 0,602$

Phương trình hồi qui $y = r\dfrac{s_y}{s_x}(x - \overline{x}) + \overline{y}$ = - 0,004x + 4,222

c/ y (x=600) = - 0,004.600 + 4,222 = 1,822 %

BÀI TẬP CHƯƠNG 4

1. Cho bảng tương quan thực nghiệm 2 chiều: (Từ bài ý 1 đến bài ý 22)

a/ Hãy tìm hệ số tương quan mẫu?

b/ Viết phương trình đường hồi qui tuyến tính thực nghiệm của Y theo X

1/

Y \ X	100	200	300	400	500
26	8	6			
30	2	10	4		
34		4	26	6	
38			5	10	7
42				4	8

2/

Y \ X	100	200	300	400	500
20	8	6			
30	2	10	4		
40		4	26	6	
50			5	10	7
60				4	8

3/

Y \ X	100	200	300	400	500
26	8	6			
30	2	10	4		
34		4	26	6	
38			5	10	7
42				4	8

4/

Y \ X	50	100	150	200	250
100				4	4
110		2	6	1	1
120	1	4	2		
130	3		1		1

5/

Y \ X	50	100	150	200	250
200				4	4
210		2	6	1	1
220	1	4	2		
230	3		1		1

6/

Y \ X	50	60	70	80	90

100				4	4
110		2	6	1	1
120	1	4	2		
130	3		1		1

7/

Y \ X	50	100	150	200	250
100				4	4
110		2	6	1	1
120	1	4	2		
130	3		1		1

8/

Y \ X	10	20	30	40	50	60
15	5	7				
25		20	23			
35			30	47	2	
45			10	11	20	6
55				9	7	3

9/

Y \ X	10	11	12	13
5				2
4			1	2
3		2	2	
2		1	2	
1	2	1		

10/

Y \ X	10	20	30	40	50	60
25	5	7				
35		20	23			
45			30	47	2	
55			10	11	20	6
65				9	7	3

11/

Y \ X	10	20	30	40

5				2
4			1	2
3		2	2	
2		1	2	
1	2	1		

12/

Y \ X	24	27	30	33	36
120	1	3			
125		2	6	1	
130		1	5	5	
135		1	6	7	2
140			1	4	2
145				1	1
150					1

13/

X \ Y	25	28	31	34	37
50	1	3			
55		2	6	1	
60		1	5	5	
65		1	6	7	2
70			1	4	2
75				1	1
80					1

14/

X \ Y	20	30	40	50	60
120	1	3			
130		2	6	1	
140		1	5	5	
150		1	6	7	2
160			1	4	2
170				1	1
180					1

2. Kiểm tra hai môn toán và vật lý một nhóm 10 sinh viên được chọn ngẫu nhiên từ một lớp ta có kết quả sau:

Điểm toán (X)	7	6	7	10	4	5	7	8	8	9
Điểm vật lý (Y)	6	7	7	9	5	3	8	9	6	7

a/ Hãy tìm hệ số tương quan mẫu?

b/ Viết phương trình đường hồi qui tuyến tính thực nghiệm của Y theo X.

3. Số vi khuẩn Y sinh sản sau X giờ được ghi lại trong bảng sau qua một thí nghiệm:

Thời gian (X)	0	1	2	3	4	5	6	7	8
Số vi khuẩn (Y)(triệu)	30	32	35	40	48	52	58	62	69

a/ Hãy tính hệ số tương quan mẫu.

b/ Tìm phương trình đường hồi qui tuyến tính thực nghiệm của Y theo X.

4. Để nghiên cứu về lượng Protein chứa trong hạt lúa mỳ người ta tiến hành điều tra trên 10 thửa ruộng và được kết quả sau:

Năng suất X	9,0	9,5	10,0	10,5	11,0	11,5	12,0	12,5	13,0	13,5
Tỉ lệ Protein Y	10,0	10,2	11,0	10,5	12,0	12,2	12,5	12,6	12,7	12,8

a/ Hãy tính hệ số tương quan mẫu.

b/ Tìm phương trình đường hồi qui tuyến tính thực nghiệm của Y theo X.

TÀI LIỆU THAM KHẢO

[1] Nguyễn Đình Trí, *Toán học cao cấp tập 1*, Nhà xuất bản giáo dục, năm 2004.

[2] Nguyễn Đình Trí, *Toán học cao cấp tập 2*, Nhà xuất bản giáo dục, năm 2004.

[3] Nguyễn Đình Trí, *Toán học cao cấp tập 3*, Nhà xuất bản giáo dục, năm 2002.

[4] Đào Hữu Hồ, Xác suất thống kê, Nhà xuất bản Đại học quốc gia Hà Nội, 2007.

[5] Nguyễn Văn Cao, Giáo trình Lý thuyết Xác suất & Thống kê toán, Trường Đại học Kinh tế quốc dân, Nhà xuất bản Thống kê, 2005.

Phụ lục 2

Bảng giá trị của hàm $\Phi(u)=\frac{1}{\sqrt{2\pi}}\int_0^u e^{-t^2/2}dt$ (Hàm Láp-la-xơ)

u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$
1	2	3	4	5	6
0,00	0,0000	0,41	0,1591	0,82	0,2939
0,01	0,0040	0,42	0,1628	0,83	0,2967
0,02	0,0080	0,43	0,1664	0,84	0,3995
0,03	0,0120	0,44	0,1700	0,85	0,3023
0,04	0,0160	0,45	0,1736	0,86	0,3051
0,05	0,0199	0,46	0,1772	0,87	0,3078
0,06	0,0239	0,47	0,1808	0,88	0,3106
0,07	0,0279	0,48	0,1844	0,89	0,3133
0,08	0,0319	0,49	0,1879	0,90	0,3159
0,09	0,0359	0,50	0,1915	0,91	0,3186
0,10	0,0398	0,51	0,1950	0,92	0,3412
0,11	0,0438	0,52	0,1985	0,93	0,3238
0,12	0,0478	0,53	0,2019	0,94	0,3264
0,13	0,0517	0,54	0,2054	0,95	0,3289
0,14	0,0557	0,55	0,2088	0,96	0,3315
0,15	0,0596	0,56	0,2123	0,97	0,3340
0,16	0,0636	0,57	0,2157	0,98	0,3365
0,17	0,0675	0,58	0,2190	0,99	0,3389
0,18	0,0714	0,59	0,2224	1,00	0,3413
0,19	0,0753	0,60	0,2257	1,01	0,3438
0,20	0,0793	0,61	0,2291	1,02	0,3461
0,21	0,0832	0,62	0,2324	1,03	0,3485
0,22	0,0871	0,63	0,2357	1,04	0,3508
0,23	0,0910	0,64	0,2389	1,05	0,3531
0,24	0,0948	0,65	0,2422	1,06	0,3554
0,25	0,0984	0,66	0,2454	1,07	0,3577
0,26	0,1026	0,67	0,2486	1,08	0,3599
0,27	0,1064	0,68	0,2517	1,09	0,3621
0,28	0,1103	0,69	0,2549	1,10	0,3643
0,29	0,1141	0,70	0,2580	1,11	0,3665
0,30	0,1179	0,71	0,2611	1,12	0,3686
0,31	0,1217	0,72	0,2612	1,13	0,3708
0,32	0,1255	0,73	0,2673	1,14	0,3729
0,33	0,1293	0,74	0,2703	1,15	0,3749
0,34	0,1331	0,75	0,2734	1,16	0,3770
0,35	0,1368	0,76	0,2764	1,17	0,3790
0,36	0,1406	0,77	0,2794	1,18	0,3810
0,37	0,1443	0,78	0,2823	1,19	0,3830
0,38	0,1480	0,79	0,2852	1,20	0,3849
0,39	0,1517	0,80	0,2881	1,21	0,3869
0,40	0,1554	0,81	0,2910	1,22	0,3883
u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$

1	2	3	4	5	6
1,23	0,3907	1,68	0,4535	2,26	0,4881
1,24	0,3925	1,69	0,4545	2,28	0,4887
1,25	0,3944	1,70	0,4554	2,30	0,4893
1,26	0,3962	1,71	0,4564	2,32	0,4898
1,27	0,3980	1,72	0,4573	2,34	0,4904
1,28	0,3997	1,73	0,4582	2,36	0,4909
1,29	0,4015	1,74	0,4591	2,38	0,4913
1,30	0,4032	1,75	0,4599	2,40	0,4918
1,31	0,4049	1,76	0,4608	2,42	0,4922
1,32	0,4066	1,77	0,4616	2,44	0,4927
1,33	0,4082	1,78	0,4625	2,46	0,4931
1,34	0,4099	1,79	0,4633	2,48	0,4934
1,35	0,4115	1,80	0,4641	2,50	0,4938
1,36	0,4131	1,81	0,4649	2,52	0,4941
1,37	0,4147	1,82	0,4656	2,54	0,4945
1,38	0,4162	1,83	0,4664	2,56	0,4948
1,39	0,4177	1,84	0,4671	2,58	0,4951
1,40	0,4192	1,85	0,4678	2,60	0,4953
1,41	0,4207	1,86	0,4686	2,62	0,4956
1,42	0,4222	1,87	0,4693	2,64	0,4959
1,43	0,4236	1,88	0,4699	2,66	0,4961
1,44	0,4251	1,89	0,4706	2,68	0,4963
1,45	0,4265	1,90	0,4713	2,70	0,4965
1,46	0,4279	1,91	0,4719	2,72	0,4967
1,47	0,4292	1,92	0,4726	2,74	0,4969
1,48	0,4306	1,93	0,4732	2,76	0,4971
1,49	0,4319	1,94	0,4738	2,78	0,4973
1,50	0,4332	1,95	0,4744	2,80	0,4974
1,51	0,4345	1,96	0,4750	2,82	0,4976
1,52	0,4357	1,97	0,4756	2,84	0,4977
1,53	0,4370	1,98	0,4761	2,86	0,4979
1,54	0,4382	1,99	0,4767	2,88	0,4980
1,55	0,4394	2,00	0,4772	2,90	0,4981
1,56	0,4406	2,02	0,4783	2,92	0,4982
1,57	0,4418	2,04	0,4793	2,94	0,4984
1,58	0,4429	2,06	0,4803	2,96	0,4985
1,59	0,4441	2,08	0,4812	2,98	0,4986
1,60	0,4452	2,10	0,4921	3,00	0,49865
1,61	0,4463	2,12	0,4830	3,20	0,49931
1,62	0,4474	2,14	0,4838	3,40	0,49966
1,63	0,4484	2,16	0,4846	3,60	0,499841
1,64	0,4495	2,18	0,4854	3,80	0,499928
1,65	0,4505	2,20	0,4861	4,00	0,499968
1,66	0,4515	2,22	0,4868	4,50	0,499997
1,67	0,4525	2,24	0,4875	5,00	0,499997

Phụ lục 3

Bảng giá trị phân vị U_α

α	U_α	α	U_α	α	U_α	α	U_α
0,50	0,00	0,75	0,674	0,95	1,645	0,975	1,96
0,51	0,025	0,76	0,706	0,951	1,655	0,976	1,977
0,52	0,030	0,77	0,739	0,952	1,665	0,977	1,995
0,53	0,075	0,78	0,772	0,953	1,675	0,978	2,014
0,54	0,100	0,79	0,806	0,954	1,685	0,979	2,034
0,55	0,126	0,80	0,842	0,955	1,695	0,980	2,054
0,56	0,151	0,81	0,878	0,956	1,706	0,981	2,075
0,57	0,176	0,82	0,915	0,957	1,717	0,982	2,097
0,58	0,202	0,83	0,954	0,958	1,728	0,983	2,120
0,59	0,228	0,84	0,994	0,959	1,739	0,984	2,144
0,60	0,253	0,85	1,036	0,960	1,751	0,985	2,170
0,61	0,279	0,86	1,080	0,961	1,762	0,986	2,197
0,62	0,305	0,87	1,126	0,962	1,774	0,987	2,226
0,63	0,332	0,88	1,175	0,963	1,787	0,988	2,257
0,64	0,358	0,89	1,227	0,964	1,799	0,989	2,290
0,65	0,385	0,90	1,282	0,965	1,812	0,990	2,326
0,66	0,412	0,905	1,311	0,966	1,825	0,991	2,366
0,67	0,440	0,910	1,341	0,967	1,837	0,992	2,409
0,68	0,468	0,915	1,372	0,968	1,852	0,993	2,457
0,69	0,496	0,920	1,405	0,969	1,866	0,994	2,512
0,70	0,524	0,925	1,440	0,970	1,881	0,995	2,576
0,71	0,553	0,930	1,476	0,971	1,896	0,996	2,652
0,72	0,583	0,935	1,514	0,972	1,911	0,997	2,748
0,73	0,613	0,940	1,555	0,973	1,927	0,998	2,878
0,74	0,643	0,945	1,598	0,974	1,943	0,999	3,090

Với $\alpha < 0,5$ suy ra giá trị $U_\alpha = -U_{1-\alpha}$

Phụ lục 5

Phân vị student $t_{\alpha}^{(n)}$

n \ α	0,90	0,95	0,975	0,99	0,995
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,183	4,541	5,842
4	1,533	2,132	2,776	3,767	4,601
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,335
9	1,383	1,833	2,262	2,821	3,325
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,974
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771
28	1,313	1,701	2,048	2,467	2,763
29	1,311	1,699	2,045	2,462	2,756
30	1,310	1,697	2,042	2,457	2,750