

# XÁC SUẤT – THỐNG KÊ

## ÔN THI TUYỂN SINH SAU ĐẠI HỌC

---

### CHƯƠNG I. ÔN TẬP ĐẠI SỐ TỔ HỢP

#### 1. Quy tắc cộng

Giả sử một công việc  $V$  có thể thực hiện theo hai phương án  $V_1$  hoặc  $V_2$ , trong đó  $V_1$  có  $m_1$  cách thực hiện,  $V_2$  có  $m_2$  cách thực hiện và mỗi cách thực hiện  $V_1$  không trùng với bất kì cách thực hiện  $V_2$  nào. Khi đó số cách thực hiện công việc  $V$  là  $n = m_1 + m_2$ .

**Ví dụ 1.** Nhà An có 2 xe đạp, 3 xe máy. Khi đến trường An đi xe đạp hoặc xe máy. Hỏi An có bao nhiêu cách đi đến trường?

**Ví dụ 2.** Một bộ bài có 52 lá với 4 chất khác nhau. Hỏi có bao nhiêu cách lấy ra một lá cơ hoặc lá át? Trong ví dụ này có thể sử dụng quy tắc cộng (bằng cách lấy số cách chọn một lá át cộng với số cách chọn một lá cơ) được không? Tại sao?

#### 2. Quy tắc nhân

Giả sử một công việc  $V$  bao gồm hai công đoạn  $V_1$  và  $V_2$ , trong đó  $V_1$  có  $m_1$  cách thực hiện,  $V_2$  có  $m_2$  cách thực hiện và mỗi cách thực hiện  $V_1$  đều có  $m_2$  cách thực hiện  $V_2$ . Khi đó số cách thực hiện công việc  $V$  là  $n = m_1 \cdot m_2$ .

**Ví dụ 3.** Từ nhà Bình ( Q.3, TP. HCM) về quê ( Long Xuyên, An Giang) phải đi qua bến xe Miền Tây. Từ nhà ra bến xe, Bình đi xe buýt; từ bến xe về quê, Bình đi xe ô tô khách. Biết rằng có 3 xe buýt đi đến bến xe Miền Tây và từ đó có 5 xe khách về quê. Hỏi Bình có bao nhiêu cách về quê?

#### 3. Tổ hợp

Mỗi tập con gồm  $k$  phần tử khác nhau lấy ra từ tập hợp có  $n$  phần tử được gọi là một tổ hợp chập  $k$  của  $n$  phần tử đã cho. Số các tổ hợp chập  $k$  của  $n$  phần tử là  $C_n^k = \frac{n!}{k!(n-k)!}$ .

**Ví dụ 4.** Có 5 đội bóng thi đấu vòng loại. Mỗi trận đấu giữa các đội (gồm 2 phần tử lấy từ 5 phần tử) là một tổ hợp chập 2 của 5 phần tử đã cho. Vậy số trận đấu là  $C_5^2 = \frac{5!}{2!3!} = 10$ .

**Ví dụ 5.** Trong một cái hộp có 10 viên phấn trắng và 6 viên phấn màu. Lấy ra 5 viên phấn. Hỏi có bao nhiêu cách lấy được:

- |                             |                             |
|-----------------------------|-----------------------------|
| a) các viên phấn bất kì?    | b) 2 viên phấn màu?         |
| c) ít nhất 4 viên phấn màu? | d) ít nhất 1 viên phấn màu? |

### BÀI TẬP CHƯƠNG I

1. Một lô hàng có 40 sản phẩm, trong đó có 5 phế phẩm. Chọn ra 8 sản phẩm để kiểm tra. Hỏi có bao nhiêu cách chọn được

- a) các sản phẩm bất kì?      b) không quá 2 phế phẩm?      c) ít nhất 1 phế phẩm?

2. Người ta lấy ra 3 viên bi từ một cái hộp đựng 6 viên vi đỏ, 4 viên bi xanh, 5 viên bi vàng. Hỏi có bao nhiêu cách lấy ra

- |                                  |                                  |
|----------------------------------|----------------------------------|
| a) các viên bi tùy ý?            | b) 2 viên bi đỏ, 1 viên bi xanh? |
| c) các viên bi có màu khác nhau? | d) một viên màu vàng?            |
| e) nhiều nhất một viên màu đỏ?   | f) ít nhất một viên màu xanh?    |

# CHƯƠNG II. XÁC SUẤT

## II.1. PHÉP THỬ - BIẾN CỐ

### 1. Khái niệm

Hành động mà ta thực hiện là phép thử. Kết quả của phép thử được gọi là biến cố.

**Ví dụ 1.** a) Tung một viên phần lên cao, viên phần rơi xuống.

b) Một sinh viên đi thi môn Toán và đậu môn này, nhưng đi thi ngoại ngữ lại bị rớt.

c) Bóc một tờ lịch trong quyển lịch năm 2009, được tờ có ghi ngày 31-6-2009.

Hãy chỉ ra phép thử và biến cố trong từng ví dụ trên.

### 2. Phân loại biến cố

- Biến cố luôn luôn xảy ra trong phép thử được gọi là biến cố chắc chắn, kí hiệu là  $\Omega$ .

- Biến cố không bao giờ xảy ra được gọi là biến cố không thể, kí hiệu là  $\phi$ .

- Biến cố có thể xảy ra, hoặc không xảy ra trong phép thử được gọi là biến cố ngẫu nhiên, kí hiệu là  $A, B, \dots, C_1, C_2, \dots$

**Ví dụ 2.** Các biến cố ở ví dụ 1 là biến cố gì?

### 3. Các phép toán đối với biến cố

#### a) Tổng các biến cố

Cho hai biến cố  $A$  và  $B$ . Tổng của chúng là một biến cố  $C$  sao cho  $C$  xảy ra khi  **$A$  hoặc  $B$**  xảy ra, kí hiệu  $C = A + B$ .

#### b) Tích các biến cố

Cho hai biến cố  $A$  và  $B$ . Tích của chúng là một biến cố  $C$  sao cho  $C$  xảy ra khi  **$A$  và  $B$**  xảy ra, kí hiệu  $C = A.B$

#### c) Biến cố đối lập

Hai biến cố  $A$  và  $B$  được gọi là đối lập nhau nếu biến cố  $A$  xảy ra thì biến cố  $B$  không xảy ra và nếu  $A$  không xảy ra thì  $B$  phải xảy ra, kí hiệu  $B = \overline{A}$ .

**Ví dụ 3.** Một sinh viên thi hai môn Toán, Lý. Gọi  $A$  là biến cố sinh viên đó đậu Toán,  $B$  là biến cố sinh viên đó đậu Lý. Hãy viết các biến cố sau thành phép toán của  $A$  và  $B$ :

a) Sinh viên đó đậu ít nhất một môn.

b) Sinh viên đó đậu cả hai môn.

c) Sinh viên đó chỉ đậu môn Lý.

d) Sinh viên đó bị rớt cả hai môn.

e) Sinh viên đó chỉ đậu một môn.

g) Sinh viên đó đậu không quá một môn.

### 4. Các biến cố xung khắc nhau

Hai biến cố được gọi là xung khắc nhau nếu chúng không cùng xảy ra.

### 5. Các biến cố độc lập

Hai biến cố được gọi là độc lập nếu biến cố này xảy ra không ảnh hưởng đến biến cố kia và ngược lại.

### 6. Hệ biến cố đầy đủ

Hệ  $n$  biến cố  $A_1, A_2, \dots, A_n$  được gọi là **hệ đầy đủ** nếu luôn **có một và chỉ một** biến cố của hệ xảy ra trong phép thử.

**Ví dụ 4.** Hai biến cố  $A, \overline{A}$  tạo thành một hệ đầy đủ.

**Ví dụ 5.** Một sinh viên phải thi 4 môn. Gọi  $D_k$  ( $k = 0, 1, 2, 3, 4$ ) là số môn sinh viên đó đậu trong 4 môn đã thi. Ta có hai biến cố bất kì trong các biến cố này xung khắc với nhau và hệ 5 biến cố đó là hệ đầy đủ.

**Ví dụ 4.** Tung một súc sắc ( hình lập phương gồm 6 mặt có đánh số từ 1 đến 6). Gọi  $A_k$  là biến cố xuất hiện mặt  $k$ . Hãy nêu các biến cố xung khắc với nhau. Các biến cố nào tạo thành hệ đầy đủ?

## II.2. ĐỊNH NGHĨA XÁC SUẤT

### 1. Định nghĩa

Cho  $T$  là một phép thử và  $A$  là biến cố có thể xảy ra trong phép thử đó.

Giả sử  $T$  có  $n$  trường hợp có thể xảy ra, trong số đó có  $m$  trường hợp làm biến cố  $A$  xuất hiện. Khi đó tỉ số  $\frac{m}{n}$  được gọi là xác suất của biến cố  $A$ , kí hiệu là  $P(A)$ .

Vậy 
$$P(A) = \frac{m}{n}$$

**Ý nghĩa của xác suất:** Xác suất của một biến cố là một số đặc trưng cho khả năng xuất hiện biến cố đó trong phép thử, xác suất càng lớn, khả năng xuất hiện biến cố càng nhiều.

**Phương pháp tính xác suất bằng định nghĩa:**

Để tính xác suất của một biến cố, ta cần thực hiện các bước sau đây:

- *Gọi phép thử, tính số trường hợp có thể xảy ra (tính số cách thực hiện phép thử).*
- *Gọi tên biến cố cần tìm xác suất, tính số cách làm biến cố đó xuất hiện.*
- *Áp dụng công thức định nghĩa, tìm xác suất của biến cố đã cho.*

**Ví dụ 1.** Một chi đoàn có 30 sinh viên nam và 15 sinh viên nữ. Cần chọn ra 8 sinh viên tham gia chiến dịch mùa hè xanh. Tìm xác suất để trong nhóm chọn ra có 3 sinh viên nữ.

**Ví dụ 2.** Đề cương thi môn Triết có 70 câu hỏi. Một sinh viên chỉ ôn 40 câu. Cho biết đề thi tự luận gồm 3 câu thuộc đề cương và nếu sinh viên trả lời đúng ít nhất hai câu thì đậu. Tìm xác suất sinh viên đó đậu môn Triết.

**Ví dụ 3.** Tung 2 đồng tiền, mỗi đồng có một mặt sấp và một mặt ngửa. Tìm xác suất được

- a) 2 mặt đều sấp.                      b) 2 mặt đều ngửa.                      c) 1 mặt sấp và 1 mặt ngửa.

Trong ba biến cố trên, biến cố nào thường xảy ra nhiều hơn?

### 2. Các tính chất của xác suất

1) Với mọi biến cố  $A$  ta luôn có  $0 \leq P(A) \leq 1$

2)  $P(\phi) = 0$

3)  $P(\Omega) = 1$

4)  $P(\bar{A}) = 1 - P(A)$

**Ví dụ 4.** Một bộ bài có 52 lá với hai màu đen, đỏ và 4 chất: rô, cơ, píc, chuồn. Lấy ra 8 lá bài từ bộ bài đó. Tìm xác suất lấy được

- a) 3 lá màu đỏ;                      b) ít nhất 1 lá màu đỏ;  
c) 2 lá rô, 4 lá cơ;                      d) 1 lá píc, 2 lá rô, 3 lá chuồn;  
e) không quá 1 lá màu đen;                      f) ít nhất 1 lá rô.

**Ví dụ 5.** Một lô hàng có 50 sản phẩm, trong đó có 6 phế phẩm. Lấy ngẫu nhiên 10 sản phẩm từ lô hàng đó. Tìm xác suất lấy được

- a) 8 sản phẩm tốt;                      b) không quá 1 phế phẩm;  
c) Ít nhất 1 phế phẩm.

## II.3. CÁC CÔNG THỨC TÍNH XÁC SUẤT

### 1. Công thức cộng

Cho hai biến cố  $A, B$  và  $C = A + B$ . Cần tính xác suất của  $C$  theo xác suất của  $A$  và  $B$ .

a) Trường hợp hai biến cố  $A$  và  $B$  xung khắc:  $P(A+B) = P(A) + P(B)$  (1)

b) Trường hợp  $A$  và  $B$  không xung khắc:  $P(A+B) = P(A) + P(B) - P(AB)$  (2)

**Ví dụ 1.** Có 10 cái bút, trong đó có 4 bút đỏ, số còn lại là bút xanh. Lấy ngẫu nhiên 3 cái bút. Tìm xác suất lấy được

- a) 1 bút đỏ;                      b) 2 bút đỏ;                      c) 3 bút xanh;  
d) không quá 1 bút đỏ;                      e) ít nhất 1 bút đỏ.

**Ví dụ 2.** Trong hộp phấn có 10 viên phấn màu và 40 viên phấn trắng. Lấy ngẫu nhiên 5 viên phấn. Tìm xác suất lấy được

- a) 1 viên phấn màu;
- b) toàn phấn trắng;
- c) nhiều nhất 1 viên phấn màu;
- d) ít nhất 1 viên phấn màu.

**Ví dụ 3.** Một lớp học có 50 sinh viên, trong đó có 35 người đậu môn Toán, 28 người đậu môn Lý. Số sinh viên của lớp đậu cả hai môn này là 20. Gọi ngẫu nhiên một sinh viên của lớp. Tìm xác suất sinh viên đó đậu ít nhất một môn.

**Ví dụ 4.** Lớp 12A có 40 học sinh, trong đó có 20 em thích chơi bóng đá, 15 em thích bơi, trong số đó có 5 em thích cả đá bóng và bơi. Gọi ngẫu nhiên 1 em học sinh trong lớp. Tìm xác suất gọi được em không thích cả 2 môn thể thao này.

## 2. Công thức nhân

Cho hai biến cố A, B và  $C = AB$ . Cần tính xác suất của C theo xác suất của A và B.

a) Trường hợp hai biến cố A và B độc lập:  $P(AB) = P(A) P(B)$  (3)

b) Trường hợp hai biến cố A và B không độc lập:

$$P(AB) = P(A) P(B/A) \quad \text{hoặc} \quad P(AB) = P(B) P(A/B) \quad (4)$$

**Ví dụ 5.** Một sinh viên phải thi Toán và Lý. Cho biết xác suất đậu hai môn đó lần lượt là 0,9; 0,8. Hãy tính các xác suất sau đây:

- a) Sinh viên đó rớt Toán; rớt Lý;
- b) Sinh viên đó chỉ đậu Toán;
- c) Sinh viên đó đậu cả hai môn;
- d) Sinh viên đó chỉ đậu một môn;
- e) Sinh viên đó đậu không quá một môn.
- f) Sinh viên đó đậu ít nhất một môn.

**Ví dụ 6.** Một xạ thủ bắn hai viên đạn, xác suất bắn trúng từng viên lần lượt là 0,6 ; 0,7. Tìm xác suất anh ta bắn trúng

- a) cả hai viên;
- b) chỉ một viên;
- c) ít nhất một viên.

**Ví dụ 7.** Một cậu bé có 10 cái bút chì, trong đó có 3 bút chì màu. Cậu bé cho anh mình 2 cái bút, sau đó cho chị mình 1 cái bút. Tìm xác suất cậu bé còn lại

- a) toàn bút chì đen;
- b) 2 bút chì màu;
- c) ít nhất 1 bút chì màu.

**Ví dụ 8.** Hai bạn An và Bình rủ nhau mua vé số tại một quầy có 50 vé, trong đó có 4 vé sẽ trúng thưởng. An mua trước 2 vé, sau đó Bình chọn mua 1 vé. Tìm xác suất hai bạn đó mua được:

- a) 3 vé trúng.
- b) 1 vé trúng.
- c) Ít nhất 1 vé trúng.

## 3. Công thức xác suất đầy đủ

Cho hệ đầy đủ các biến cố  $A_1, A_2, \dots, A_n$  và B là biến cố xảy ra khi một trong các biến cố của hệ đó xảy ra. Khi đó xác suất của B được tính bởi công thức

$$P(B) = P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + \dots + P(A_n)P(B/A_n) \quad (5)$$

**Ví dụ 9.** Cho 3 cái hộp đựng bút hình dáng giống nhau. Hộp thứ nhất có 2 bút đỏ, 8 bút xanh. Hộp thứ hai có 4 bút đỏ, 6 bút xanh. Hộp thứ ba có 4 bút đỏ, 8 bút xanh. Lấy ngẫu nhiên một hộp, từ đó lấy ngẫu nhiên 3 cái bút. Tìm xác suất lấy được

- a) 1 bút đỏ;
- b) ít nhất một bút đỏ.

**Ví dụ 10.** Có 4 lô hàng, mỗi lô có 20 sản phẩm. Cho biết lô thứ nhất có 3 phế phẩm, lô thứ hai có 2 phế phẩm, lô thứ ba có 4 phế phẩm và lô thứ tư có 1 phế phẩm. Chọn ngẫu nhiên 1 lô hàng, từ đó lấy ngẫu nhiên 6 sản phẩm. Tìm xác suất lấy được

- a) 2 phế phẩm;
- b) ít nhất 1 phế phẩm.

**Ví dụ 11.** Có hai lô hàng đựng các thiết bị điện tử. Lô thứ nhất có 4 phế phẩm và 46 sản phẩm tốt. Lô thứ hai có 3 phế phẩm và 45 sản phẩm tốt. Từ lô thứ nhất lấy ra 2 sản phẩm bỏ sang lô thứ hai. Sau đó từ lô thứ hai lấy ra 5 sản phẩm. Tìm xác suất lấy được

a) 5 sản phẩm tốt;

b) ít nhất một phế phẩm.

**Ví dụ 12.** Có 2 chuồng gà: chuồng thứ nhất có 5 con gà trống và 10 con gà mái; chuồng thứ hai có 3 con gà trống và 8 con gà mái. Một con từ chuồng thứ nhất nhảy sang chuồng thứ hai, sau đó từ chuồng thứ hai người ta bắt ra 2 con gà. Tìm xác suất bắt được 1 con gà mái.

#### 4. Công thức xác suất giả thiết

Cho hệ đầy đủ các biến cố  $A_1, A_2, \dots, A_n$  và  $B$  là biến cố xảy ra khi một trong các biến cố của hệ đó xảy ra. Ta còn nói các biến cố thuộc hệ đầy đủ là giả thiết để  $B$  xảy ra.

Bây giờ ta giả sử biến cố  $B$  đã xảy ra và đi tìm xác suất để  $B$  xảy ra là do giả thiết  $A_k$ , kí hiệu  $P(A_k/B)$ . Ta có công thức  $P(A_k/B) = \frac{P(A_k)P(B/A_k)}{P(B)}$  ( $k = 1, 2, \dots, n$ ) (6)

Công thức này còn được gọi là công thức Bayes.

**Ví dụ 13.** Một nhà máy có ba phân xưởng cùng sản xuất một loại sản phẩm. Phân xưởng thứ nhất sản xuất 25%, phân xưởng thứ hai sản xuất 35%, còn phân xưởng thứ ba sản xuất 40% tổng số sản phẩm của cả nhà máy. Tỷ lệ phế phẩm của từng phân xưởng lần lượt là 1%; 1,5%; 2%. Lấy ngẫu nhiên một sản phẩm trong kho hàng của nhà máy.

a) Tìm xác suất lấy được phế phẩm.

b) Giả sử lấy được phế phẩm, tìm xác suất phế phẩm đó do phân xưởng thứ nhất sản xuất.

c) Nếu lấy được sản phẩm tốt, theo ý bạn, khả năng sản phẩm đó do phân xưởng nào sản xuất là nhiều nhất?

**Ví dụ 14.** Cho 3 cái hộp đựng bút hình dáng giống nhau. Hộp thứ nhất có 2 bút đỏ, 8 bút xanh. Hộp thứ hai có 4 bút đỏ, 6 bút xanh. Hộp thứ ba có 4 bút đỏ, 8 bút xanh. Lấy ngẫu nhiên một hộp, từ đó lấy ngẫu nhiên 2 cái bút.

a) Tìm xác suất lấy được hai bút khác màu.

b) Giả sử đã lấy được hai bút khác màu. Tìm xác suất đó là các bút của hộp thứ ba.

#### PHƯƠNG PHÁP TÍNH XÁC SUẤT BẰNG CÔNG THỨC

Để tính xác suất của một biến cố bằng công thức, ta cần thực hiện các bước sau đây:

- *Gọi tên biến cố cần tìm xác suất, phân tích nó thành phép toán đối với các biến cố khác*
- *Phân tích mối quan hệ giữa các biến cố tham gia vào phép toán: xung khắc hay không, độc lập hay không, có tạo thành hệ đầy đủ hay không...*
- *Chọn công thức tính xác suất của biến cố ban đầu theo xác suất của các biến cố đó.*
- *Tính xác suất của các biến cố tham gia vào phép toán, nếu cần.*
- *Tính xác suất của biến cố ban đầu.*

#### 5. Công thức Bernoulli

Giả sử: - Phép thử  $T$  lặp lại  $n$  lần

- Biến cố  $A$  có thể xuất hiện trong mỗi lần thử với xác suất không đổi  $P(A) = p$ .

Khi đó xác suất để biến cố  $A$  xuất hiện  $k$  lần trong  $n$  lần thử là

$$P_n(k, A) = C_n^k p^k (1 - p)^{n-k} \quad (k = 0, 1, \dots, n) \quad (7)$$

**Ví dụ 15.** Một sinh viên thi 5 môn với xác suất đậu từng môn là 0,7. Tìm xác suất anh ta

a) đậu 3 môn;

b) không đậu môn nào;

c) đậu ít nhất một môn;

d) đậu từ 1 đến 3 môn.

**Ví dụ 16.** Một xạ thủ đã bắn 6 viên đạn. Cho biết xác suất bắn trúng mục tiêu của mỗi lần bắn đều là 0,8. Tìm xác suất anh ta bắn trúng

a) 4 viên;

b) không quá 2 viên;

c) ít nhất một viên đạn.

**Ví dụ 17.** Cho biết tỷ lệ học sinh bị cận thị ở một trường THPT là 20%. Chọn ngẫu nhiên 12 học sinh của trường đó. Tìm xác suất có

- a) 3 học sinh bị cận;
- b) không quá 1 học sinh bị cận;
- c) có từ 2 đến 4 học sinh bị cận;
- d) ít nhất một học sinh bị cận.

## **BÀI TẬP CHƯƠNG II**

**1.** Hai người cùng bắn, mỗi người bắn một viên đạn vào tấm bia. Gọi  $N_i$  là biến cố người thứ  $i$  bắn trúng bia,  $i = 1, 2$ . Hãy biểu diễn các biến cố sau qua  $N_1, N_2$  :

- a) Chỉ có người thứ nhất bắn trúng bia.
- b) Có đúng một người bắn trúng.
- c) Cả hai người đều bắn trúng.
- d) Không có ai bắn trúng.
- e) Có ít nhất một người bắn trúng.
- f) Có không quá một người bắn trúng.

**2.** Lớp học môn xác suất có 64 sinh viên, trong đó có 15 sinh viên nữ. Chọn ngẫu nhiên một nhóm gồm 10 sinh viên. Tìm xác suất trong nhóm chọn ra có

- a) 4 sinh viên nữ;
- b) không quá 2 sinh viên nữ;
- c) ít nhất 1 sinh viên nữ.

**3.** Từ một cái hộp có 20 viên phấn, trong đó có 5 viên phấn màu, người ta lấy ra 6 viên phấn. Tìm xác suất lấy được

- a) 2 viên phấn màu;
- b) không quá một viên phấn màu;
- c) ít nhất một viên phấn màu.

**4.** Từ một lô hàng có 6 quả bóng đỏ, 9 quả bóng xanh và 5 quả bóng vàng người ta lấy ra 5 quả bóng. Tìm xác suất lấy được

- a) 2 quả bóng đỏ;
- b) 2 quả bóng vàng, 3 quả bóng xanh;
- c) Ít nhất 1 quả bóng đỏ;
- d) Không quá 1 quả bóng đỏ.

**5.** Có ba người, mỗi người bắn một viên đạn vào bia với xác suất bắn trúng lần lượt là 0,6 ; 0,7 ; 0,8. Tìm các xác suất sau đây:

- a) Chỉ có người thứ hai bắn trúng;
- b) Có đúng một người bắn trúng;
- c) Chỉ có người thứ ba bắn trúng;
- d) Có đúng hai người bắn trúng;
- e) Cả ba người đều bắn trúng;
- f) Không có ai bắn trúng.

**6.** Có ba cái hộp đựng bút. Hộp thứ nhất có 5 bút đỏ, 10 bút xanh. Hộp thứ hai có 3 bút đỏ, 7 bút xanh. Hộp thứ ba có 3 bút đỏ, 4 bút xanh. Từ hộp thứ nhất lấy ra 1 cái bút, từ hộp thứ hai lấy ra 2 cái, cùng bỏ vào hộp thứ ba. Tìm xác suất trong hộp thứ ba số bút đỏ nhiều hơn số bút xanh.

**7.** Có hai lô hàng, lô thứ nhất có 10 sản phẩm loại A, 2 sản phẩm loại B; lô thứ hai có 16 sản phẩm loại A, 4 sản phẩm loại B. Từ mỗi lô ta lấy ngẫu nhiên ra một sản phẩm. Sau đó, trong hai sản phẩm thu được lại lấy ra một sản phẩm. Tìm xác suất để sản phẩm lấy ra sau cùng là sản phẩm loại A.

**8.** Hai máy cùng sản xuất ra một loại chi tiết. Năng suất của máy thứ hai gấp đôi máy thứ nhất. Tỷ lệ chi tiết đạt tiêu chuẩn của máy thứ nhất là 65%, của máy thứ hai là 80%. Lấy ngẫu nhiên một chi tiết từ lô hàng do hai máy sản xuất.

- a) Tìm xác suất lấy được chi tiết đạt tiêu chuẩn.
- b) Nếu chi tiết đó là phế phẩm, tìm xác suất chi tiết đó do máy thứ hai sản xuất.

**9.** Ở một vùng cứ 100 người thì có 30 người hút thuốc lá. Biết tỷ lệ người bị viêm họng trong số người hút thuốc là 60%, còn trong số người không hút là 10%.

- a) Khám ngẫu nhiên một người. Tìm xác suất người đó bị viêm họng.
- b) Giả sử người được khám bị viêm họng. Tìm xác suất anh ta hút thuốc.
- c) Nếu người đó không bị viêm họng thì xác suất để anh ta hút thuốc bằng bao nhiêu ?

**10.** Có bốn nhóm xạ thủ tập bắn. Nhóm thứ nhất có 6 người, nhóm thứ hai có 7 người, nhóm thứ ba có 8 người và nhóm thứ tư có 4 người. Xác suất bắn trúng đích của mỗi người trong bốn nhóm đó lần lượt là 0,8 ; 0,7 ; 0,6 ; 0,5. Chọn ngẫu nhiên một xạ thủ.

a) Tìm xác suất anh ta bắn trúng đích.

b) Giả sử xạ thủ này bắn trượt. Hỏi khả năng người đó ở nhóm nào là nhiều nhất ?

11. Giả sử xác suất sinh con trai là 0,51. Một gia đình có 4 người con. Tìm xác suất để gia đình đó có a) hai con trai; b) không quá một con trai.

12. Một xạ thủ có xác suất bắn trúng đích ở mỗi lần bắn là 0,7. Anh ta đã bắn 5 lần, mỗi lần 1 viên đạn. Tìm xác suất có a) 3 viên trúng đích; b) không quá 3 viên trúng.

## CHƯƠNG III. ĐẠI LƯỢNG NGẪU NHIÊN VÀ CÁC PHÂN PHỐI XÁC SUẤT

### III.1. ĐỊNH NGHĨA và PHÂN LOẠI ĐẠI LƯỢNG NGẪU NHIÊN

#### 1. Khái niệm

Đại lượng cho tương ứng mỗi kết quả của phép thử với một số được gọi là đại lượng ngẫu nhiên (hay biến ngẫu nhiên) trên các kết quả của phép thử đó. Nói một cách khác, đại lượng ngẫu nhiên là đại lượng có giá trị thay đổi tùy theo phép thử.

**Ví dụ 1.** a) Số môn thi đậu của một sinh viên trong một học kì (khi phải thi 5 môn).

b) Nhiệt độ của phòng học trong một ngày đêm.

c) Số người đến giao dịch tại một ngân hàng trong một tháng.

d) Chiều cao của thanh niên Việt nam thường trong khoảng 155 cm đến 180 cm.

#### 2. Các loại đại lượng ngẫu nhiên

Đại lượng ngẫu nhiên được chia thành hai loại: rời rạc và liên tục.

- Đại lượng ngẫu nhiên  $X$  có dạng

$$X = \{x_1, x_2, \dots, x_n\} \text{ hoặc } X = \{x_1, x_2, \dots, x_n, \dots\}$$

được gọi là đại lượng ngẫu nhiên rời rạc.

- Đại lượng ngẫu nhiên có giá trị lấp đầy một khoảng  $(a, b)$  hay đoạn  $[a, b]$  nào đó được gọi là đại lượng ngẫu nhiên liên tục ( $a, b$  có thể hữu hạn hoặc vô hạn).

**Ví dụ 2.** Các đại lượng ngẫu nhiên cho ở ví dụ 1 là đại lượng gì?

#### 3. Phân phối xác suất

Để nghiên cứu đại lượng ngẫu nhiên  $X$  ta cần biết các giá trị có thể có của  $X$  và xác suất để nó nhận mỗi giá trị đó. Mối liên hệ giữa các giá trị có thể có của  $X$  và xác suất tương ứng được gọi là phân phối xác suất của đại lượng ngẫu nhiên  $X$ .

Đối với đại lượng ngẫu nhiên rời rạc ta có bảng phân phối xác suất. Trường hợp đại lượng ngẫu nhiên liên tục ta có hàm mật độ phân phối xác suất.

Dưới đây ta chỉ xét đại lượng ngẫu nhiên rời rạc nhận hữu hạn các giá trị.

##### a) Bảng phân phối xác suất

Cho  $X = \{x_1, x_2, \dots, x_n\}$  là một đại lượng ngẫu nhiên rời rạc.

Đặt  $p_i = P(X = x_i)$ ,  $i = 1, 2, \dots, n$ .

Khi đó bảng sau đây được gọi là bảng phân phối xác suất của  $X$ .

|   |       |       |     |       |
|---|-------|-------|-----|-------|
| X | $x_1$ | $x_2$ | ... | $x_n$ |
| P | $p_1$ | $p_2$ | ... | $p_n$ |

Bảng phân phối xác suất có các tính chất sau

$$(1) \quad 0 \leq P_i \leq 1$$

$$(2) \quad \sum_{i=1}^n p_i = 1$$

**Ví dụ 3.** Gọi X là số môn thi đậu của một sinh viên trong học kì phải thi 5 môn. Khi đó X nhận các giá trị: 0, 1, 2, 3, 4, 5. Giả sử X có bảng phân phối xác suất sau đây.

|   |      |      |     |      |      |   |
|---|------|------|-----|------|------|---|
| X | 0    | 1    | 2   | 3    | 4    | 5 |
| P | 0,05 | 0,15 | 0,3 | 0,35 | 0,15 | 0 |

Từ bảng ta có xác suất thi đậu 4 môn của sinh viên đó là 0,15; xác suất đậu cả 5 môn là 0. Trong các xác suất ta thấy  $P(X = 3)$  lớn nhất nên khả năng anh ta đậu 3 môn là nhiều nhất.

**Ví dụ 4.** Một xạ thủ được phép bắn 3 viên đạn. Gọi X là số viên đạn anh ta bắn trúng bia. Hãy lập bảng phân phối xác suất của X, biết xác suất bắn trúng mục tiêu của mỗi viên đạn đều là 0,8.

### b) Hàm phân phối xác suất

Cho X là đại lượng ngẫu nhiên rời rạc. Khi đó hàm số có dạng  $F(x) = P(X < x)$ ,  $x \in \mathbb{R}$ , được gọi là hàm phân phối xác suất của X.

Hàm phân phối xác suất có các tính chất sau đây:

- (1)  $F(x)$  là hàm không giảm; (2)  $0 \leq F(x) \leq 1$ ,  $\forall x \in \mathbb{R}$ ;  
 (3)  $\lim_{x \rightarrow -\infty} F(x) = 0$ ;  $\lim_{x \rightarrow +\infty} F(x) = 1$ ; (4)  $P(a \leq X < b) = F(b) - F(a)$ ,  $\forall a, b \in \mathbb{R}$ .

(5) Nếu X là đại lượng ngẫu nhiên liên tục thì  $F'(x) = f(x)$ ,  $\forall x \in \mathbb{R}$ .

Ngược lại, nếu  $F(x)$  là hàm số xác định trên  $\mathbb{R}$  và có các tính chất (1) – (3) thì  $F(x)$  là hàm phân phối xác suất của một đại lượng ngẫu nhiên nào đó.

Nếu X là đại lượng ngẫu nhiên rời rạc có bảng phân phối xác suất

|   |       |       |     |       |
|---|-------|-------|-----|-------|
| X | $x_1$ | $x_2$ | ... | $x_n$ |
| P | $p_1$ | $p_2$ | ... | $p_n$ |

với  $x_1 < x_2 < \dots < x_n$ , thì hàm phân phối xác suất của X là

$$F(x) = \begin{cases} 0 & , \text{nếu } x \leq x_1 \\ p_1 & , \text{nếu } x_1 < x \leq x_2 \\ \dots & \dots \\ p_1 + p_2 + \dots + p_{n-1} & , \text{nếu } x_{n-1} < x \leq x_n \\ 1 & , \text{nếu } x > x_n \end{cases}$$

**Ví dụ 6.** Cho X là đại lượng ngẫu nhiên rời rạc có bảng phân phối xác suất như sau

|   |      |      |     |
|---|------|------|-----|
| X | 1    | 2    | 4   |
| P | 0,25 | 0,45 | 0,3 |

Hàm phân phối xác suất của X có dạng

$$F(x) = \begin{cases} 0 & \text{khi } x \leq 1 \\ 0,25 & \text{khi } 1 < x \leq 2 \\ 0,7 & \text{khi } 2 < x \leq 4 \\ 1 & \text{khi } x > 4 \end{cases}$$

**Ví dụ 7.** Một sinh viên thi ba môn Toán, Lý, Hóa với xác suất đậu lần lượt là 0,6 ; 0,7 ; 0,8. Hãy tìm hàm phân phối xác suất của số môn anh ta đậu trong ba môn đó.

## III.2. CÁC ĐẶC TRƯNG CỦA ĐẠI LƯỢNG NGẪU NHIÊN RỜI RẠC

### 1. Kỳ vọng

Cho  $X$  là đại lượng ngẫu nhiên rời rạc có bảng phân phối xác suất là :

|     |       |       |         |       |
|-----|-------|-------|---------|-------|
| $X$ | $x_1$ | $x_2$ | $\dots$ | $x_n$ |
| $P$ | $p_1$ | $p_2$ | $\dots$ | $p_n$ |

Khi đó số  $E(X) = \sum_{i=1}^n x_i p_i$  được gọi là kỳ vọng của  $X$ .

Kỳ vọng của đại lượng ngẫu nhiên là trung bình theo xác suất các giá trị có thể nhận của đại lượng đó.

### 2. Phương sai

Số  $D(X) = E(X^2) - E^2(X)$ ,

trong đó  $E(X)$  là kỳ vọng của  $X$ ,  $E(X^2) = \sum_{i=1}^n x_i^2 p_i$  là kỳ vọng của  $X^2$ ,

được gọi là phương sai của đại lượng ngẫu nhiên  $X$ .

Phương sai còn được tính bởi công thức:  $D(X) = \sum_{i=1}^n (x_i - E(X))^2 p_i$

Phương sai là trung bình của bình phương sai số giữa  $X$  và trung bình theo xác suất của  $X$ .

### 3. Độ lệch chuẩn

Số  $\sigma(X) = \sqrt{D(X)}$  được gọi là độ lệch chuẩn của đại lượng ngẫu nhiên  $X$ .

**Ví dụ 1.** Tính kỳ vọng, phương sai và độ lệch chuẩn của đại lượng ngẫu nhiên  $X$ , biết bảng phân phối xác suất của nó là

|     |      |      |     |
|-----|------|------|-----|
| $X$ | 1    | 2    | 4   |
| $P$ | 0,25 | 0,45 | 0,3 |

**Giải.** Ta có

$$E(X) = 1 \cdot 0,25 + 2 \cdot 0,45 + 4 \cdot 0,3 = 2,35; E(X^2) = 1^2 \cdot 0,25 + 2^2 \cdot 0,45 + 4^2 \cdot 0,3 = 6,85;$$

$$D(X) = E(X^2) - E^2(X) = 6,85 - 2,35^2 = 1,3275; \sigma(X) = \sqrt{D(X)} = \sqrt{1,3275} = 1,1522.$$

Ta có thể tính phương sai bằng cách khác:

$$D(X) = (1 - 2,35)^2 \cdot 0,25 + (2 - 2,35)^2 \cdot 0,45 + (4 - 2,35)^2 \cdot 0,3 = 1,3275.$$

**Ví dụ 2.** Một sinh viên thi 4 môn, xác suất đậu từng môn là 0,6. Gọi  $X$  là số môn sinh viên đó đậu. Hãy lập bảng phân phối xác suất, hàm phân phối xác suất và tính kỳ vọng, phương sai, độ lệch chuẩn của  $X$ .

## BÀI TẬP CHƯƠNG III

**1.** Một xạ thủ được phát 3 viên đạn và được phép bắn lần lượt từng viên cho đến khi trúng mục tiêu thì dừng bắn. Biết xác suất bắn trúng từng viên đều là 0,8. Hãy lập bảng phân phối xác suất, hàm phân phối xác suất và tính kỳ vọng, phương sai, độ lệch chuẩn của số viên

- đạn                      a) trúng mục tiêu;                      b) anh ta đã sử dụng.
2. Một hộp chứa 10 viên phấn trắng và 6 viên phấn màu. Chọn ngẫu nhiên ra 3 viên phấn. Gọi  $X$  là số viên phấn màu lấy được. Hãy lập bảng phân phối xác suất, hàm phân phối xác suất và tính kì vọng, phương sai, độ lệch chuẩn của  $X$ .
3. Trong một hộp có 20 sản phẩm, trong đó có 5 phế phẩm. Lấy ngẫu nhiên 4 sản phẩm. Gọi  $X$  là số phế phẩm trong các sản phẩm lấy ra. Tìm hàm phân phối xác suất của  $X$  và tính kì vọng, phương sai, độ lệch chuẩn của  $X$ .
4. Trong một hộp có 10 sản phẩm, trong đó có 2 phế phẩm. Lấy ngẫu nhiên 3 sản phẩm. Gọi  $X$  là số phế phẩm trong các sản phẩm lấy ra. Hãy lập bảng phân phối xác suất và hàm phân phối xác suất của  $X$ .
5. Một sinh viên thi 4 môn, xác suất đậu từng môn là 0,7. Gọi  $X$  là số môn anh ta đậu. Hãy lập bảng phân phối xác suất và tính kì vọng, phương sai của  $X$ .

## CHƯƠNG IV. LÝ THUYẾT MẪU

### IV.1. MẪU VÀ PHƯƠNG PHÁP CHỌN MẪU

#### 1. Tổng thể và mẫu

- Tập hợp tất cả các phần tử mà ta quan tâm nghiên cứu trong một vấn đề nào đó được gọi là tổng thể. Số phần tử của tổng thể được gọi là kích thước của tổng thể.
- Tập con lấy ra từ tổng thể được gọi là mẫu. Số phần tử của mẫu được gọi là kích thước mẫu hay cỡ mẫu.

**Ví dụ 1.** Khi nghiên cứu về số cá trong một hồ thì tổng thể là toàn bộ các con cá trong hồ đó. Số cá có trong hồ chính là kích thước của tổng thể. Nếu ta bắt lên 20 con cá thì ta được một mẫu với kích thước là 20.

**Ví dụ 2.** Nghiên cứu về tỉ lệ chất kích thích trong bia Sài Gòn thì tổng thể là toàn bộ số chai bia và lon bia do nhà máy bia Sài Gòn sản xuất. Nếu ta lấy ra 50 chai để điều tra về chất kích thích trong bia thì ta được một mẫu kích thước 50.

#### 2. Phương pháp chọn mẫu

Có hai cách chọn các phần tử của tổng thể để lấy làm mẫu

- a) **Chọn không hoàn lại:** phần tử đã chọn sẽ bị loại ra khỏi tổng thể rồi mới chọn phần tử tiếp theo. Khi đó ta có mẫu không hoàn lại.
- b) **Chọn có hoàn lại:** phần tử đã chọn được trả lại tổng thể rồi mới chọn phần tử tiếp theo. Khi đó ta được mẫu có hoàn lại.

**Ví dụ 3.** Khi lấy chai bia ra điều tra thì người ta phải khai nó ra để phân tích thành phần hoá học. Dĩ nhiên sau đó ta không thể trả nó về tổng thể. Vậy, ở ví dụ 2 ta có mẫu không hoàn lại.

**Ví dụ 4.** Ở ví dụ 2 ta có hai cách bắt cá: bắt một lúc 20 con lên xem xét – đây cũng là mẫu không hoàn lại. Nhưng nếu ta bắt lần lượt từng con, xem xét xong thì thả lại vào hồ rồi mới bắt con tiếp theo (con bắt lần sau có thể trùng với con bắt lần trước). Làm như vậy 20 lần ta được mẫu có hoàn lại kích thước 20.

Thông thường, khi tổng thể có ít phần tử thì người ta hay chọn mẫu có hoàn lại. Trường hợp tổng thể có nhiều phần tử thì người ta xem hai cách chọn mẫu này như nhau, nghĩa là không phân biệt có hoàn lại hay không hoàn lại.

#### 3. Phân loại mẫu

a) **Mẫu định tính:** là mẫu mà ta quan tâm đến các phần tử trong mẫu có tính chất  $T$  nào đó hay không. Khi đó mẫu được cho ở dạng:

- Kích thước mẫu:  $n$
- Số phần tử có tính chất  $T$  của mẫu:  $k$

**Ví dụ 5.** Trước ngày bầu cử người ta phỏng vấn 125 cử tri thì thấy 74 người ủng hộ ứng cử viên A.

Đây là mẫu định tính lấy ra từ tổng thể tất cả cử tri. Kích thước mẫu  $n = 125$ , tính chất T là “ủng hộ ứng cử viên A”, số phần tử có tính chất T của mẫu là  $k = 74$ .

b) **Mẫu định lượng:** là mẫu mà ta quan tâm đến giá trị của từng phần tử trong mẫu. Khi đó mẫu được cho ở dạng :

- Kích thước mẫu:  $n$
- Giá trị của các phần tử:  $x_1, x_2, \dots, x_n$ .

**Ví dụ 6.** Điểm thi môn Toán của 10 sinh viên như sau: 5, 6, 8, 4, 2, 8, 9, 6, 5, 7.

Đây là mẫu định lượng với kích thước  $n = 10$ . Điểm thi của từng sinh viên chính là giá trị của các phần tử trong mẫu.

Nếu trong mẫu có một số phần tử có giá trị giống nhau thì người ta thường cho mẫu ở dạng thu gọn: giá trị  $x_1$  có tần số  $n_1$  (nghĩa là trong mẫu có  $n_1$  phần tử cùng có giá trị  $x_1$ ), giá trị  $x_2$  có tần số  $n_2$ , ..., giá trị  $x_m$  có tần số  $n_m$ , trong đó  $n_1 + n_2 + \dots + n_m = n$  là tổng số phần tử của mẫu. Khi đó người ta viết mẫu ở dạng bảng sau đây

|         |       |       |     |       |
|---------|-------|-------|-----|-------|
| Giá trị | $x_1$ | $x_2$ | ... | $x_m$ |
| Tần số  | $n_1$ | $n_2$ | ... | $n_m$ |

**Ví dụ 7.** Điểm thi môn Anh văn của một số sinh viên Trường ĐH như sau

|              |   |    |    |    |    |   |   |
|--------------|---|----|----|----|----|---|---|
| Điểm thi     | 3 | 4  | 5  | 6  | 7  | 8 | 9 |
| Số sinh viên | 4 | 10 | 15 | 12 | 10 | 7 | 2 |

Đây là mẫu định lượng ở dạng thu gọn. Giá trị của các phần tử chính là điểm thi, tần số của giá trị 3 là 4, tần số của giá trị 4 là 10, ..., tần số của giá trị 9 là 2. Kích thước mẫu  $n = 4 + 10 + 15 + 12 + 10 + 7 + 2 = 60$ .

Nếu  $n_i$  là số phần tử của mẫu nhận giá trị trong khoảng  $(a_i, b_i)$  thì ta xem chúng nhận giá trị chung là  $x_i = \frac{a_i + b_i}{2}$ .

**Ví dụ 8.** Chiều cao của một số thanh niên được cho trong bảng sau đây

| Chiều cao (cm) | Số người có chiều cao tương ứng |
|----------------|---------------------------------|
| 154 – 158      | 10                              |
| 158 – 162      | 14                              |
| 162 – 166      | 26                              |
| 166 – 170      | 28                              |
| 170 – 174      | 12                              |
| 174 – 178      | 8                               |
| 178 - 182      | 2                               |

Từ bảng này ta có bảng các giá trị đại diện như sau

|               |     |     |     |     |     |     |     |
|---------------|-----|-----|-----|-----|-----|-----|-----|
| Chiều cao     | 156 | 160 | 164 | 168 | 172 | 176 | 180 |
| Số thanh niên | 10  | 14  | 26  | 28  | 12  | 8   | 2   |

## IV.2. CÁC ĐẶC TRƯNG CỦA MẪU

### 1. Các đặc trưng của mẫu định tính

$$\text{Tần suất mẫu: } f = \frac{k}{n}$$

**Ví dụ 1.** Trước kỳ bầu cử, người ta phỏng vấn 1575 cử tri thì thấy có 1212 người trả lời là ủng hộ ứng cử viên X. Tìm tỷ lệ mẫu ủng hộ ứng cử viên X.

**Ví dụ 2.** Người ta bắt được 1200 con cá, đánh dấu rồi thả lại vào hồ nước. Sau một thời gian bắt lại 250 con thì thấy có 32 con bị đánh dấu. Hãy tìm tỷ lệ mẫu bị đánh dấu.

### 2. Các đặc trưng của mẫu định lượng

a) Trung bình mẫu: 
$$\bar{X} = \frac{1}{n} \sum_{i=1}^m x_i n_i$$

b) Trung bình mẫu bình phương: 
$$\overline{X^2} = \frac{1}{n} \sum_{i=1}^m x_i^2 n_i$$

c) Phương sai mẫu: 
$$\hat{S}^2 = \overline{X^2} - (\bar{X})^2$$

d) Phương sai mẫu hiệu chỉnh: 
$$S^2 = \frac{n}{n-1} \hat{S}^2$$

e) Độ lệch mẫu hiệu chỉnh: 
$$S = \sqrt{S^2}$$

**Ví dụ 3.** Điều tra năng suất lúa của một vùng, ta có bảng số liệu sau

|                                           |    |    |    |    |    |    |    |
|-------------------------------------------|----|----|----|----|----|----|----|
| Năng suất lúa (tạ/ha)                     | 41 | 44 | 45 | 46 | 48 | 52 | 54 |
| Diện tích có năng suất lúa tương ứng (ha) | 10 | 20 | 30 | 15 | 10 | 10 | 5  |

Hãy tính các đặc trưng mẫu.

**Ví dụ 4.** Để nghiên cứu nhu cầu mua gạo ở một thành phố, người ta tiến hành điều tra một số gia đình và ghi kết quả ở bảng sau đây:

| Nhu cầu<br>(kg/tháng) | Số gia đình | Nhu cầu | Số gia đình |
|-----------------------|-------------|---------|-------------|
| 30 – 35               | 45          | 55 – 60 | 182         |
| 35 – 40               | 68          | 60 – 65 | 151         |
| 40 – 45               | 103         | 65 – 70 | 115         |
| 45 – 50               | 179         | 70 – 75 | 94          |
| 50 – 55               | 208         | 75 – 80 | 55          |

a) Hãy tính các đặc trưng mẫu.

b) Tính tỉ lệ mẫu có nhu cầu trên 60 kg/tháng.

## BÀI TẬP CHƯƠNG IV

1. Điểm kiểm tra môn Toán của một số sinh viên được cho trong bảng sau.

| Điểm         | 0 | 1 | 2 | 3  | 4  | 5  | 6 | 7  | 8  | 9 | 10 |
|--------------|---|---|---|----|----|----|---|----|----|---|----|
| Số sinh viên | 1 | 2 | 5 | 10 | 20 | 48 | 3 | 22 | 10 | 5 | 2  |

a) Tính điểm kiểm tra trung bình và phương sai mẫu hiệu chỉnh của mẫu đó.

b) Tìm tỉ lệ mẫu có điểm kiểm tra dưới trung bình.

2. Đo chiều cao của một số thanh niên ở tỉnh Tiền Giang, ta thu được bảng sau đây

| Chiều cao (cm) | Số người có chiều cao tương ứng |
|----------------|---------------------------------|
| 154 – 158      | 10                              |
| 158 – 162      | 14                              |
| 162 – 166      | 26                              |
| 166 – 170      | 28                              |
| 170 – 174      | 12                              |
| 174 – 178      | 8                               |
| 178 – 182      | 2                               |

a) Hãy tính các đặc trưng mẫu.

b) Tính tỉ lệ mẫu có chiều cao trên 1,7m.

## CHƯƠNG V. ƯỚC LƯỢNG CÁC THAM SỐ CỦA ĐẠI LƯỢNG NGẪU NHIÊN

### V.1. KHÁI NIỆM

#### 1. Ước lượng tham số

Giả sử một đại lượng ngẫu nhiên  $X$  có phân phối xác suất đã biết nhưng các đặc trưng của nó phụ thuộc vào một (hoặc nhiều) tham số  $\theta$  chưa biết. Thông qua mẫu ngẫu nhiên, ta cần xác định giá trị gần đúng để thay thế cho tham số chưa biết đó.

**Ví dụ 1.** Ta biết một đại lượng ngẫu nhiên nào đó có phân phối chuẩn  $X \sim N(\mu, \sigma^2)$ , nhưng  $\mu$  và  $\sigma^2$  nhận giá trị nào thì chưa biết. Do đó cần tìm giá trị gần đúng của  $\mu$  và  $\sigma^2$ , thông qua một mẫu ngẫu nhiên, lấy từ tổng thể các giá trị của đại lượng ngẫu nhiên đó.

Muốn vậy, từ tổng thể ta lấy một mẫu ngẫu nhiên kích thước  $n$ , sau đó lập một đại lượng thống kê  $\hat{\theta}$  để thay thế cho  $\theta$ . Khi đó ta gọi  $\hat{\theta}$  là ước lượng của  $\theta$ .

#### 2. Ước lượng điểm

Giả sử  $X = (x_1, x_2, \dots, x_n)$  là một mẫu ngẫu nhiên kích thước  $n$  lấy từ tổng thể có kì vọng  $\mu$  và phương sai  $\sigma^2$ . Ta dùng một hàm nào đó của mẫu này để ước lượng cho  $\theta$ , kí hiệu là  $\hat{\theta} = F(x_1, x_2, \dots, x_n)$ . Khi đó, với một mẫu cụ thể thì hàm số này sẽ nhận một giá trị cụ thể biểu thị bởi một điểm trên trục số. Do đó ta gọi số  $\hat{\theta}$  là một ước lượng điểm của  $\theta$  và coi  $\hat{\theta}$  là giá trị gần đúng của  $\theta$ .

#### 3. Ước lượng không chệch

Ước lượng  $\hat{\theta}$  của  $\theta$  được gọi là ước lượng không chệch nếu kì vọng của  $\hat{\theta}$  là  $\theta$ , nghĩa là nếu  $E(\hat{\theta}) = \theta$ .

Với mọi mẫu ta luôn có:

Tỉ lệ mẫu  $f$  là ước lượng không chệch của tỉ lệ tổng thể  $P$ .

Trung bình mẫu  $\bar{X}$  là ước lượng không chệch của kì vọng của tổng thể  $\mu$ .

Phương sai mẫu hiệu chỉnh  $S^2$  là ước lượng không chệch của phương sai của tổng thể  $\sigma^2$ .

#### 4. Ước lượng khoảng

Khoảng  $(a, b)$  được gọi là khoảng ước lượng của  $\theta$  nếu ta coi  $\theta \in (c, d)$ .

Xác suất  $P[\theta \in (c, d)] = 1 - \alpha$  được gọi là độ tin cậy của ước lượng.

Nếu  $\hat{\theta}$  là một ước lượng không chệch của  $\theta$  thì khoảng ước lượng của  $\theta$  có dạng  $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$ , khoảng này được gọi là khoảng ước lượng đối xứng. Số  $\varepsilon > 0$  được gọi là độ chính xác (hay sai số) của ước lượng.

Nếu  $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$  là khoảng ước lượng đối xứng của  $\theta$  với độ tin cậy  $1 - \alpha$  thì xác suất  $P[\theta \in (\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)] = P(|\theta - \hat{\theta}| < \varepsilon) = 1 - \alpha$ .

## V.2. ƯỚC LƯỢNG GIÁ TRỊ TRUNG BÌNH CỦA TỔNG THỂ

### 1. Bài toán

Giả sử tổng thể có giá trị trung bình là  $\mu$  chưa biết. Ta cần ước lượng  $\mu$  với độ tin cậy  $1 - \alpha$  cho trước. Ta cũng giả thiết rằng ta đã có một mẫu gồm  $n$  quan sát được chọn từ tổng thể đó và đã tính được trung bình mẫu  $\bar{X}$ , độ lệch mẫu hiệu chỉnh  $S$ . Khi đó tùy từng trường hợp cụ thể, ta có phương pháp tìm khoảng ước lượng như sau.

### 2. Trường hợp tổng thể có phân phối chuẩn, đã biết phương sai $\sigma^2$

Độ chính xác được tính bởi công thức:  $\varepsilon = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ , trong đó  $Z$  là biến ngẫu nhiên có phân

phối chuẩn với  $\varphi(Z_{\frac{\alpha}{2}}) = \frac{1}{2} - \frac{\alpha}{2}$  và  $\varphi(Z)$  là hàm phân phối xác suất Laplace (có bảng giá trị

cho trước).

Khoảng ước lượng của  $\mu$  là  $(\bar{X} - \varepsilon, \bar{X} + \varepsilon)$ .

**Ví dụ 1.** Kết quả thu thập trong 15 ngày tại một công ty cho thấy trung bình một ngày có 267 trang tài liệu được chuyển đi bằng fax. Theo kinh nghiệm từ các văn phòng tương tự thì độ lệch tiêu chuẩn là 32 trang. Giả sử rằng số trang tài liệu chuyển bằng fax trong một ngày có phân phối chuẩn. Hãy ước lượng số trang tài liệu được chuyển trong một ngày của công ty với độ tin cậy 95%.

**Ví dụ 2.** Một mẫu gồm 16 sản phẩm được chọn từ tổng thể có phân phối chuẩn với phương sai là 25. Cho biết trung bình mẫu là 330 gam. Hãy xác định khoảng ước lượng cho trung bình của tổng thể với độ tin cậy 90%.

### 3. Trường hợp chưa biết phương sai của tổng thể

**a) Nếu cỡ mẫu  $n \geq 30$**  thì  $\varepsilon = Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$  (tổng thể có phân phối bất kì).

**Ví dụ 3.** Khảo sát 100 sinh viên chọn ngẫu nhiên trong trường thì thấy điểm trung bình môn Toán là 5,12 và phương sai mẫu hiệu chỉnh là 0,0676. Hãy ước lượng điểm trung bình môn Toán của sinh viên toàn trường với độ tin cậy 97%.

**Ví dụ 4.** Chọn ngẫu nhiên 49 sản phẩm từ một lô hàng. Kết quả kiểm tra cho thấy tuổi thọ trung bình của các sản phẩm đó là 485 giờ và độ lệch mẫu hiệu chỉnh là 36 giờ. Hãy ước lượng tuổi thọ trung bình của cả lô hàng với độ tin cậy 99%.

**b) Nếu cỡ mẫu  $n < 30$**  và tổng thể có phân phối chuẩn thì  $\varepsilon = T_{n-1, \frac{\alpha}{2}} \frac{S}{\sqrt{n}}$ , trong đó

$T_{n-1, \frac{\alpha}{2}}$  có phân phối Student với  $n - 1$  bậc tự do (tra bảng phân phối Student dòng  $n - 1$ , cột

$\frac{\alpha}{2}$ ).

**Ví dụ 5.** Chiều dài của một loại sản phẩm có phân phối chuẩn. Đo ngẫu nhiên 10 sản phẩm được chiều dài trung bình là 10,02 m, độ lệch mẫu hiệu chỉnh là 0,04 m. Tìm khoảng ước lượng chiều dài trung bình của loại sản phẩm này với độ tin cậy 95%.

**Ví dụ 6.** Một mẫu ngẫu nhiên gồm 16 khách hàng sử dụng dịch vụ ATM thuộc hệ thống của một ngân hàng được ghi nhận về thời gian (giây) thực hiện xong một dịch vụ: 65, 30, 40, 58, 26, 60, 75, 45, 50, 36, 76, 34, 38, 50, 44, 56. Giả sử thời gian thực hiện dịch vụ qua ATM có phân phối chuẩn. Hãy tìm khoảng ước lượng cho thời gian trung bình thực hiện dịch vụ qua ATM với độ tin cậy 99%.

### V.3. ƯỚC LƯỢNG TỈ LỆ

Giả sử tổng thể có hai loại phân tử, một trong hai loại có tính chất A nào đó. Ta cần ước lượng tỉ lệ P các phân tử có tính chất A của tổng thể với độ tin cậy  $1 - \alpha$  cho trước. Giả sử ta có một mẫu ngẫu nhiên gồm  $n$  phân tử của tổng thể ( $n \geq 30$ ) và  $f$  là tỉ lệ các phân tử có tính chất A trong mẫu.

Khi đó độ chính xác được tính bởi công thức  $\varepsilon = Z_{\frac{\alpha}{2}} \sqrt{\frac{f(1-f)}{n}}$ .

Khoảng ước lượng của P là  $(f - \varepsilon, f + \varepsilon)$ .

**Ví dụ 1.** Một nghiên cứu được thực hiện nhằm ước lượng tỉ lệ khách hàng sử dụng bánh kẹo nội địa. Kết quả điều tra ngẫu nhiên 100 khách hàng cho thấy có 34 khách hàng dùng bánh kẹo nội địa. Với độ tin cậy 95% hãy ước lượng tỉ lệ khách hàng sử dụng bánh kẹo nội địa.

**Ví dụ 2.** Điểm danh ngẫu nhiên 50 sinh viên thấy có 6 sinh viên bỏ học. Hãy ước lượng tỉ lệ sinh viên bỏ học với độ tin cậy 95,45%.

**Ví dụ 3.** Trước ngày bầu cử tổng thống, người ta phỏng vấn ngẫu nhiên 1800 cử tri thì thấy có 1180 người ủng hộ ứng cử viên A. Với độ tin cậy 95%, hỏi ứng cử viên đó thu được tối thiểu bao nhiêu phần trăm số phiếu bầu?

**Ví dụ 4.** Kiểm tra ngẫu nhiên 500 sản phẩm của một nhà máy thì thấy có 360 sản phẩm loại một. Hãy ước lượng tỉ lệ sản phẩm loại một tối thiểu của cả nhà máy với độ tin cậy 95%.

**Ví dụ 5.** Người ta bắt được 1500 con thú, đánh dấu rồi thả lại vào rừng. Sau một thời gian bắt lại 360 con thì thấy có 27 con bị đánh dấu. Hãy ước lượng số thú có trong rừng với độ tin cậy 99%.

### V.4. XÁC ĐỊNH CỠ MẪU

#### 1. Trường hợp ước lượng tỉ lệ

Từ công thức tính độ chính xác của bài toán ước lượng tỉ lệ, ta có  $n = Z_{\frac{\alpha}{2}}^2 \frac{f(1-f)}{\varepsilon^2}$ .

(Nếu kết quả tìm được không phải là số nguyên thì ta lấy phần nguyên của kết quả đó cộng với 1).

**Ví dụ 1.** Lấy ngẫu nhiên 200 sản phẩm trong kho hàng thấy có 25 phế phẩm.

a) Nếu muốn độ chính xác của ước lượng là  $\varepsilon = 0,035$  thì độ tin cậy là bao nhiêu?

b) Nếu muốn độ chính xác là 0,01; độ tin cậy 95% thì cần kiểm tra thêm bao nhiêu sản phẩm nữa?

**Ví dụ 2.** Cần một mẫu kích thước bao nhiêu để tỉ lệ mẫu là 0,2; độ chính xác là 0,01 và độ tin cậy là 95%?

**Ví dụ 3.** Nếu kích thước mẫu là 100; tỉ lệ mẫu là 0,2; độ chính xác là 0,2 thì độ tin cậy của ước lượng vào khoảng bao nhiêu?

#### 2. Trường hợp ước lượng giá trị trung bình

Từ công thức tính độ chính xác của bài toán ước lượng giá trị trung bình ta có

$$n = \frac{Z_{\alpha}^2 \sigma^2}{\varepsilon^2} \quad \text{hoặc} \quad n = \frac{Z_{\alpha}^2 S^2}{\varepsilon^2}.$$

**Ví dụ 4.** (Tiếp theo ví dụ 1, bài VI.2) Giả sử ta muốn ước lượng số lượng tài liệu chuyển bằng fax trong một ngày với độ tin cậy 99% và độ chính xác là 10 trang thì kích thước mẫu là bao nhiêu? Cho biết độ lệch tiêu chuẩn của tổng thể là 32 trang.

**Ví dụ 5.** Đo đường kính của 100 chi tiết do một máy sản xuất, ta được các số liệu sau:

| Đường kính (cm) | 9,75 | 9,80 | 9,85 | 9,90 |
|-----------------|------|------|------|------|
| Số chi tiết     | 5    | 37   | 42   | 16   |

- a) Với độ chính xác  $\varepsilon = 0,06$  hãy xác định độ tin cậy.  
b) Muốn độ chính xác là 0,03; độ tin cậy 95% thì cần kiểm tra thêm bao nhiêu chi tiết?

## BÀI TẬP CHƯƠNG V

- Một trường đại học thực hiện nghiên cứu về số giờ tự học của sinh viên trong một tuần. Chọn ngẫu nhiên 200 sinh viên cho thấy số giờ tự học trong tuần tính trung bình là 18,36 giờ. Với độ tin cậy 95%, hãy ước lượng số giờ tự học trung bình trong một tuần của sinh viên toàn trường. Cho biết độ lệch tiêu chuẩn của số giờ tự học của sinh viên là 3,92 giờ.
- Công ty điện thoại ở một thành phố muốn ước lượng thời gian trung bình của một cuộc điện đàm đường dài vào những ngày cuối tuần. Mẫu ngẫu nhiên 20 cuộc gọi đường dài vào những ngày cuối tuần cho thấy thời gian điện đàm trung bình là 14,8 phút, độ lệch mẫu hiệu chỉnh là 5,6 phút. Với độ tin cậy 90%, hãy ước lượng thời gian điện đàm trung bình của dân cư thành phố đó vào những ngày cuối tuần. Cho biết thời gian điện đàm là đại lượng ngẫu nhiên có phân phối chuẩn.
- Hãy ước lượng tỉ lệ sản phẩm tốt của một nhà máy với độ tin cậy 90,106%, biết rằng khi kiểm tra 400 sản phẩm thì thấy có 20 phế phẩm.
- Tại một vùng quê, thăm dò 100 thanh niên thấy có 45 người tốt nghiệp THPT. Hãy ước lượng tỉ lệ thanh niên tốt nghiệp THPT ở vùng đó với độ tin cậy 93%.
- Kiểm tra ngẫu nhiên 500 sản phẩm của một nhà máy thì thấy có 360 sản phẩm loại một. Hãy ước lượng tỉ lệ sản phẩm loại một tối thiểu của cả nhà máy với độ tin cậy 95%.
- Lấy ngẫu nhiên 400 hộp từ một kho đồ hộp để kiểm tra thì thấy có 20 hộp bị hỏng. Từ kết quả kiểm tra đó, hãy ước lượng tỉ lệ phế phẩm của kho hàng với độ tin cậy 95,45%.
- Tại một khu rừng nguyên sinh, người ta đeo vòng vào chân của 1200 con chim. Sau một thời gian bắt lại 250 con thì thấy 40 con có đeo vòng. Hãy ước lượng số chim trong khu rừng đó với độ tin cậy 99%.
- Muốn biết số cá có trong một hồ lớn, người ta bắt lên 2000 con, đánh dấu xong lại thả chúng xuống hồ. Sau đó người ta bắt lên 400 con thì thấy có 55 con bị đánh dấu. Với độ tin cậy 0,95 hãy ước lượng số cá trong hồ. Cho biết mỗi con cá có khối lượng trung bình 800 gam và mỗi kilôgam cá bán được 28000đ. Tính doanh thu tối thiểu khi bán hết số cá trong hồ.
- Cân thử 100 trái cây của một nông trường, ta có kết quả sau đây

| Khối lượng (g) | Số trái |
|----------------|---------|
| 35 – 55        | 3       |
| 55 – 75        | 10      |
| 75 – 95        | 25      |
| 95 – 115       | 35      |

|           |    |
|-----------|----|
| 115 – 135 | 20 |
| 135 – 155 | 6  |
| 155 - 175 | 1  |

- a) Hãy ước lượng khối lượng trung bình của trái cây ở nông trường với độ tin cậy 90,106%.  
b) Xem các trái cây có khối lượng không quá 95 gam là trái cây loại hai. Hãy ước lượng tỉ lệ trái cây loại hai với độ tin cậy 95,45%.

**10.** Đo đường kính của một số chi tiết do một máy sản xuất, ta có số liệu sau:

| Đường kính (mm) | Số chi tiết |
|-----------------|-------------|
| 19,80 – 19,85   | 3           |
| 19,85 – 19,90   | 5           |
| 19,90 – 19,95   | 16          |
| 19,95 – 20,00   | 28          |
| 20,00 – 20,05   | 23          |
| 20,05 – 20,10   | 14          |
| 20,10 – 20,15   | 7           |
| 20,15 – 20,20   | 4           |

- a) Hãy ước lượng đường kính trung bình của các chi tiết do máy sản xuất với độ tin cậy 95%.  
b) Hãy ước lượng tỉ lệ chi tiết đạt tiêu chuẩn với độ tin cậy 99%, biết rằng những chi tiết đạt tiêu chuẩn phải có đường kính từ 19,9 mm đến 20,1 mm.

**11.** Trong một kì thi của trung tâm ngoại ngữ, người ta chọn ngẫu nhiên 49 bài thi và đếm số lỗi viết sai chính tả. Kết quả cho thấy số lỗi trung bình của mỗi bài thi là 6,5 và độ lệch mẫu hiệu chỉnh là 3,44.

- a) Tìm khoảng ước lượng cho số lỗi sai trung bình của tất cả các bài thi với độ tin cậy 90%.  
b) Với độ tin cậy 95%, muốn sai số ước lượng cho số lỗi trung bình không quá 1,2 thì cần chọn bao nhiêu bài thi để kiểm tra?

**12.** Một cuộc nghiên cứu được thực hiện để tìm hiểu số giờ xem tivi trung bình trong một tuần của học sinh tiểu học. Kết quả điều tra cho thấy số giờ xem tivi trung bình là 15 giờ và độ lệch mẫu hiệu chỉnh là 6 giờ. Với độ tin cậy 99%, muốn sai số ước lượng không quá 1,2 giờ thì cần chọn bao nhiêu học sinh để điều tra?

## CHƯƠNG VI. KIỂM ĐỊNH GIẢ THUYẾT THỐNG KÊ

### VI.1. BÀI TOÁN KIỂM ĐỊNH

#### 1. Khái niệm

Các đặc trưng của mẫu ngoài việc sử dụng để ước lượng các đặc trưng của tổng thể còn được dùng để đánh giá xem một giả thuyết nào đó của tổng thể là đúng hay sai. Việc tìm ra kết luận để bác bỏ hay chấp nhận một giả thuyết được gọi là kiểm định giả thuyết.

**Ví dụ 1.** Một nhà sản xuất cho rằng khối lượng trung bình của một gói mì là 75 gam. Để kiểm tra điều này đúng hay sai, chọn ngẫu nhiên một số gói mì để kiểm tra và tính toán.

**Ví dụ 2.** Một xí nghiệp cho rằng tỉ lệ phế phẩm trong kho hàng là 5%. Để kiểm tra điều này đúng hay sai, chọn ngẫu nhiên một số sản phẩm để kiểm tra, tính toán.

#### 2. Giả thuyết $H_0$ và giả thuyết $H_1$

Giả sử tổng thể có đặc trưng  $\theta$  chưa biết. Với giá trị cụ thể  $\theta_0$  cho trước nào đó, ta cần kiểm định giả thuyết  $H_0 : \theta = \theta_0$ .

Giả thuyết  $H_1$  là kết quả ngược lại của giả thuyết  $H_0$ . Nếu  $H_0$  đúng thì  $H_1$  sai và ngược lại.  $H_1$  còn được gọi là giả thuyết đối.

Vậy cặp giả thuyết  $H_0$  và  $H_1$  được thể hiện trong trường hợp kiểm định sau đây

$$H_0 : \theta = \theta_0 ; H_1 : \theta \neq \theta_0 .$$

**Ví dụ 3.** Trong trường hợp ví dụ 1 ta có thể đặt giả thuyết:  $H_0 : \theta = 75 ; H_1 : \theta \neq 75$ .

### 3. Sai lầm loại một, sai lầm loại hai

Vì chỉ dựa trên một mẫu để kết luận các giá trị của tổng thể nên ta có thể phạm sai lầm khi đưa ra kết luận về giả thuyết  $H_0$ . Các sai lầm đó là:

**a) Sai lầm loại một:** Giả thuyết  $H_0$  đúng nhưng ta bác bỏ nó.

**b) Sai lầm loại hai:** Giả thuyết  $H_0$  sai nhưng ta chấp nhận nó.

Khi kiểm định, người ta mong muốn khả năng mắc sai lầm loại một không vượt quá một số  $\alpha$  cho trước, nghĩa là xác suất bác bỏ  $H_0$  khi nó đúng là  $\alpha$  thì xác suất chấp nhận nó là  $1 - \alpha$ . Ta gọi  $\alpha$  là mức ý nghĩa của kiểm định.

Ngược lại, đối với sai lầm loại hai, xác suất chấp nhận  $H_0$  khi nó sai được kí hiệu là  $\beta$ , suy ra xác suất bác bỏ  $H_0$  là  $1 - \beta$ . Ta gọi  $1 - \beta$  là độ giá trị của kiểm định.

Trong một bài toán kiểm định, nếu khả năng phạm sai lầm loại một giảm thì khả năng phạm sai lầm loại hai lại tăng lên. Do đó người ta thường chọn  $\alpha$  trong khoảng từ 1% đến 10%.

**Bảng tóm tắt hai loại sai lầm** và xác suất phạm sai lầm khi kiểm định giả thuyết

|              | Giả thuyết $H_0$ đúng                  | Giả thuyết $H_0$ sai                  |
|--------------|----------------------------------------|---------------------------------------|
| Không bác bỏ | Quyết định đúng, xác suất $1 - \alpha$ | Sai lầm loại hai, xác suất $\beta$    |
| Bác bỏ       | Sai lầm loại một, xác suất $\alpha$    | Quyết định đúng, xác suất $1 - \beta$ |

### 4. Các bước cần thực hiện trong bài toán kiểm định giả thuyết

Một bài toán kiểm định giả thuyết thường bao gồm 5 bước sau đây.

- **Bước 1.** Thiết lập giả thuyết  $H_0$  và  $H_1$ .
- **Bước 2.** Tính giá trị kiểm định. (Mỗi loại kiểm định có công thức riêng nhằm đánh giá giả thuyết).
- **Bước 3.** Chọn mức ý nghĩa  $\alpha$  và xác định miền bác bỏ giả thuyết  $H_0$  (nếu giá trị kiểm định nằm trong miền này thì  $H_0$  bị bác bỏ).
- **Bước 4.** Đưa ra kết luận về mặt thống kê, nghĩa là ở một mức ý nghĩa  $\alpha$  nào đó ta sẽ bác bỏ hay không thể bác bỏ giả thuyết  $H_0$ .
- **Bước 5.** Kết luận cuối cùng về nội dung bài toán nhằm trả lời một cách rõ ràng câu hỏi mà bài toán đặt ra (không dùng thuật ngữ thống kê).

## VI.2. KIỂM ĐỊNH GIẢ THUYẾT VỀ TRUNG BÌNH CỦA TỔNG THỂ

### 1. Bài toán

Giả sử ta có một mẫu gồm  $n$  quan sát được chọn từ tổng thể. Gọi  $\mu, \sigma^2, \bar{X}, S^2$  lần lượt là trung bình, phương sai của tổng thể; trung bình, phương sai mẫu hiệu chỉnh. Ta cần kiểm định giả thuyết:  $H_0 : \mu = \mu_0 ; H_1 : \mu \neq \mu_0$  ( $\mu_0$  là một giá trị cụ thể) với mức ý nghĩa  $\alpha$  cho trước. Khi đó việc kiểm định được thực hiện như sau.

### 2. Trường hợp đã biết phương sai của tổng thể

Giá trị kiểm định được tính bởi công thức:  $Z = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$ .

Quy tắc quyết định: Bác bỏ  $H_0$  ở mức ý nghĩa  $\alpha$  nếu  $|Z| > Z_{\frac{\alpha}{2}}$  (tức là chấp nhận  $H_0$  ở mức ý nghĩa  $\alpha$  nếu  $|Z| \leq Z_{\frac{\alpha}{2}}$ ).

Ta nhận thấy có sự liên hệ giữa ước lượng và kiểm định giả thuyết, cụ thể, giả thuyết  $H_0$  bị bác bỏ với mức ý nghĩa  $\alpha$  khi và chỉ khi khoảng tin cậy (với độ tin cậy  $1 - \alpha$ ) của  $\mu$  không chứa trị số  $\mu_0$ .

### 3. Trường hợp chưa biết phương sai của tổng thể

- Với mẫu có kích thước  $n \geq 30$  thì  $Z = \frac{\bar{X} - \mu_0}{S} \sqrt{n}$  và quy tắc quyết định như trường hợp đã biết phương sai của tổng thể.

- Với mẫu có kích thước  $n < 30$  và tổng thể có phân phối chuẩn thì giá trị kiểm định:

$$T = \frac{\bar{X} - \mu_0}{S} \sqrt{n}.$$

Quy tắc quyết định: Bác bỏ  $H_0$  ở mức ý nghĩa  $\alpha$  nếu  $|T| > T_{n-1, \frac{\alpha}{2}}$ , trong đó  $T_{n-1, \frac{\alpha}{2}}$  có phân

phối Student với  $n - 1$  bậc tự do (tra bảng phân phối Student dòng  $n - 1$ , cột  $\frac{\alpha}{2}$ ).

**Chú ý.** Trong tất cả các trường hợp trên, khi giả thuyết  $H_0 : \mu = \mu_0$  đã bị bác bỏ (tức là  $\mu \neq \mu_0$ ) thì

- Với  $\bar{X} > \mu_0$  ta kết luận  $\mu > \mu_0$ .
- Với  $\bar{X} < \mu_0$  ta kết luận  $\mu < \mu_0$ .

**Ví dụ 1.** Một máy đóng mì gói tự động quy định khối lượng trung bình là 75 gam, độ lệch tiêu chuẩn là 15 gam. Sau một thời gian sản xuất, kiểm tra 80 gói ta có khối lượng trung bình mỗi gói là 72 gam. Hãy kết luận về tình hình sản xuất với mức ý nghĩa 5%.

**Ví dụ 2.** Một nhà máy sản xuất đèn chụp hình cho biết tuổi thọ trung bình của sản phẩm là 100 giờ. Người ta chọn ngẫu nhiên 15 bóng thử nghiệm thấy tuổi thọ trung bình là 99,7 giờ, phương sai mẫu hiệu chỉnh là 0,15. Giả sử tuổi thọ của đèn có phân phối chuẩn. Cho kết luận về tình hình sản xuất của nhà máy với mức ý nghĩa 1%.

**Ví dụ 3.** Một loại đèn chiếu được nhà sản xuất quảng cáo có tuổi thọ trung bình là 65 giờ. Kết quả kiểm tra ngẫu nhiên 21 đèn cho thấy tuổi thọ trung bình là 62,5 giờ, độ lệch mẫu hiệu chỉnh là 3 giờ. Với mức ý nghĩa 1% có thể kết luận gì về lời quảng cáo đó?

**Ví dụ 4.** Một hãng sản xuất lốp xe ô tô tuyên bố sản phẩm của hãng có thể sử dụng 100 000 km, độ lệch tiêu chuẩn bằng 12 000 km. Một công ty vận tải mua 64 lốp xe, sau một thời gian sử dụng thấy độ bền trung bình là 98 500 km. Hãy kết luận về tuyên bố của hãng sản xuất lốp xe với mức ý nghĩa 5%.

**Ví dụ 5.** Năng suất lúa trung bình ở vụ trước là 4,5 tấn/ha. Vụ lúa năm nay người ta áp dụng biện pháp kỹ thuật mới cho toàn bộ diện tích trồng lúa trong vùng. Theo dõi 100 ha, ta có bảng năng suất lúa sau đây

| Năng suất (tạ/ha) | Diện tích (ha) | Năng suất | Diện tích |
|-------------------|----------------|-----------|-----------|
| 30 – 35           | 7              | 50 – 55   | 20        |
| 35 – 40           | 12             | 55 – 60   | 8         |
| 40 – 45           | 18             | 60 – 65   | 5         |
| 45 – 50           | 27             | 65 – 70   | 3         |

Với  $\alpha = 0,01$  hãy kết luận về biện pháp kỹ thuật mới.

### VI.3. KIỂM ĐỊNH GIẢ THIẾT VỀ TỶ LỆ

#### 1. Bài toán

Giả sử tổng thể gồm hai loại phân tử: có tính chất T và không có tính chất này. Ta đưa ra giả thiết:

$$H_0: \text{Tỷ lệ } P \text{ các phân tử có tính chất T của tổng thể là } P = p_0$$

Ta cần kiểm định xem giả thiết đó đúng hay sai với mức ý nghĩa  $\alpha$  cho trước.

#### 2. Phương pháp giải bài toán

- Lấy mẫu định tính gồm  $n$  phân tử của tổng thể, ta tính được tần suất mẫu  $f$ .
- Từ mức ý nghĩa  $\alpha$ , tra bảng hàm số Laplace, ta tìm được số  $Z_{\frac{\alpha}{2}}$  sao cho

$$\varphi\left(Z_{\frac{\alpha}{2}}\right) = \frac{1}{2} - \frac{\alpha}{2}.$$

- Tính giá trị kiểm định:  $Z = \frac{f - p_0}{\sqrt{p_0(1 - p_0)}} \sqrt{n}$

- So sánh  $Z$  với  $Z_{\frac{\alpha}{2}}$ : Nếu  $|Z| \leq z_{\frac{\alpha}{2}}$  thì ta chấp nhận  $H_0$ .

Nếu  $|Z| > z_{\frac{\alpha}{2}}$  thì ta bác bỏ  $H_0$ .

**Ví dụ 1.** Theo báo cáo, tỉ lệ hàng phế phẩm trong kho là 10%. Kiểm tra ngẫu nhiên 100 sản phẩm thấy có 8 phế phẩm. Hỏi báo cáo trên có đáng tin ở mức ý nghĩa 5% không?

**Ví dụ 2.** Trước đây tỉ lệ phế phẩm của một nhà máy là 5%. Năm nay người ta áp dụng một biện pháp kỹ thuật mới để sản xuất. Sau một thời gian, kiểm tra 800 sản phẩm thì thấy có 24 phế phẩm. Với mức ý nghĩa 1%, hãy đánh giá hiệu quả của biện pháp kỹ thuật đó.

### BÀI TẬP CHƯƠNG VI

1. Kiểm tra các gói đường loại 1 kg trong một siêu thị, ta có kết quả sau

|                 |      |      |      |      |      |      |      |
|-----------------|------|------|------|------|------|------|------|
| Khối lượng (kg) | 0,95 | 0,96 | 0,97 | 0,99 | 1,00 | 1,01 | 1,03 |
| Số gói          | 19   | 30   | 32   | 8    | 6    | 3    | 2    |

Với mức ý nghĩa  $\alpha = 0,05$  có thể kết luận việc đóng gói đảm bảo yêu cầu hay không?

2. Một vườn ươm cây giống quy định khi nào cây cao trung bình từ 1m trở lên thì đem ra trồng. Đo ngẫu nhiên 25 cây, được số liệu sau

|               |     |     |     |     |     |     |
|---------------|-----|-----|-----|-----|-----|-----|
| Chiều cao (m) | 0,8 | 0,9 | 1,0 | 1,1 | 1,2 | 1,3 |
| Số cây        | 1   | 2   | 9   | 7   | 4   | 2   |

Với mức ý nghĩa 5% có thể đem cây ra trồng được chưa? Cho biết chiều cao của cây có phân phối chuẩn.

3. Một mẫu gồm 49 quan sát được chọn ngẫu nhiên từ tổng thể có phân phối chuẩn. Kết quả cho thấy trung bình mẫu là 448 gam và phương sai mẫu hiệu chỉnh là 30,25. Với mức ý nghĩa 0,01, hãy kiểm định giả thuyết:  $H_0: \mu = 450$ ;  $H_1: \mu \neq 450$ .

4. Một mẫu gồm 64 quan sát được chọn ngẫu nhiên từ tổng thể có phân phối chuẩn. Kết quả cho thấy trung bình mẫu là 252 gam và phương sai mẫu hiệu chỉnh là 14,44. Hãy kiểm định giả thuyết:  $H_0: \mu = 250$ ;  $H_1: \mu \neq 250$  với mức ý nghĩa 0,1.

5. Chuyên viên kiểm tra chất lượng của một nhà máy sản xuất bóng đèn chọn ngẫu nhiên 16 sản phẩm, kết quả cho thấy tuổi thọ trung bình của các bóng đèn là 925 giờ và độ lệch mẫu hiệu chỉnh là 45 giờ.

a) Hãy ước lượng tuổi thọ trung bình của sản phẩm với độ tin cậy 90% .

b) Giám đốc nhà máy cho biết tuổi thọ của sản phẩm là 950 giờ thì có tin được không với mức ý nghĩa 5%?

6. Một công ty dược phẩm quan tâm đến mức độ tạp chất trong sản phẩm (thuốc viên), tối đa không được vượt quá 3%. Lấy ngẫu nhiên 100 viên thuốc từ một lô sản phẩm, mức độ tạp chất trung bình là 3,09%. Giả sử rằng mức độ tạp chất tuân theo phân phối chuẩn với độ lệch tiêu chuẩn là 0,5%.

a) Hãy ước lượng mức độ tạp chất trung bình của lô sản phẩm với độ tin cậy 95%.

b) Ở mức ý nghĩa 5% có thể nói rằng mức độ tạp chất trong sản phẩm là 3% được không?

7. Khối lượng quy định cho mỗi gói bánh được đóng gói tự động là 250 gam. Kiểm tra ngẫu nhiên 81 gói thì thấy khối lượng trung bình là 235 gam và độ lệch mẫu hiệu chỉnh là 36 gam. Với mức ý nghĩa 0,01 hãy kết luận về tình hình sản xuất.

8. Độ bền của một loại dây thép sản xuất theo công nghệ cũ là 150. Sau khi cải tiến kỹ thuật người ta lấy 100 sợi dây thép để thử thì thấy độ bền trung bình là 185 và  $S = 25$ . Với  $\alpha = 0,03$  hãy kết luận hiệu quả của việc cải tiến kỹ thuật.

9. Tỷ lệ người mắc bệnh tai mũi họng ở một thành phố là 6%. Trong lần kiểm tra sức khỏe ngẫu nhiên 300 người thì thấy có 24 người mắc bệnh tai mũi họng. Với  $\alpha = 0,01$  có thể cho rằng tỷ lệ người mắc bệnh đó có xu hướng tăng lên không ?

10. Khi điều trị bằng thuốc A, tỷ lệ bệnh nhân khỏi bệnh là 80%. Đổi sang thuốc B để điều trị cho 1100 người thì thấy có 920 người khỏi bệnh. Với mức ý nghĩa  $\alpha = 0,02$  có thể cho rằng thuốc B hiệu quả hơn thuốc A hay không ?

11. Một nhà máy sản xuất sản phẩm với tỷ lệ sản phẩm loại một lúc đầu là 20%. Sau khi áp dụng phương pháp sản xuất mới, kiểm tra ngẫu nhiên 500 sản phẩm thấy có 150 sản phẩm loại một. Cho kết luận về tác dụng của phương pháp sản xuất mới với mức ý nghĩa 1%.

12. Giả sử sản phẩm của một công ty sản xuất lốp xe ô tô đã chiếm được 42% thị trường. Trước sự cạnh tranh của các đối thủ và điều kiện thay đổi môi trường kinh doanh, Ban lãnh đạo công ty nghi ngờ thị trường của công ty đang bị giảm sút. Kiểm tra ngẫu nhiên 550 ô tô trên đường, kết quả cho thấy có 219 xe sử dụng lốp của công ty. Có thể kết luận gì về điều nghi ngờ trên với mức ý nghĩa 10%?