

Xác Suất Thống Kê

Ngày 10 tháng 5 năm 2014

PHẦN II. LÝ THUYẾT THỐNG KÊ (Statistical theory)

Chương V. MẪU THỐNG KÊ VÀ ƯỚC LƯỢNG THAM SỐ

§1. Lý thuyết mẫu

§2. Ước lượng khoảng

.....

§1. LÝ THUYẾT MẪU

1.1. Tổng thể và Mẫu

- Tập hợp tất cả phần tử là các đối tượng mà ta nghiên cứu được gọi là *tổng thể*. Số phần tử của tổng thể được gọi là *kích thước* của tổng thể (thường rất lớn).

➤ Chương 5. Mẫu thống kê & Ước lượng tham số

- Từ tổng thể ta chọn ra n phần tử thì n phần tử đó được gọi là một *mẫu có kích thước n (cỡ mẫu)*.
- Mẫu được chọn ngẫu nhiên một cách khách quan được gọi là *mẫu ngẫu nhiên*.
- Có hai cách lấy mẫu:
 - *Mẫu có hoàn lại*: phần tử vừa quan sát xong được trả lại cho tổng thể trước khi quan sát lần sau.
 - *Mẫu không hoàn lại*: phần tử vừa quan sát xong không được trả lại cho tổng thể.

Khi mẫu có kích thước lớn thì ta không phân biệt mẫu có hoàn lại hay không hoàn lại.

➤ Chương 5. Mẫu thống kê & Ước lượng tham số

- **Mẫu định tính** là mẫu mà ta chỉ quan tâm đến các phần tử của nó có tính chất A nào đó hay không.
- **Mẫu định lượng** là mẫu mà ta quan tâm đến các yếu tố về lượng (như chiều dài, cân nặng,...) của các phần tử có trong mẫu.
- Gọi X_1, X_2, \dots, X_n là những kết quả quan sát. Ta xem như đã quan sát n lần, mỗi lần ta được một biến ngẫu nhiên X_i ($i = 1, \dots, n$).

Do ta thường lấy mẫu trong tổng thể có rất nhiều phần tử nên X_1, X_2, \dots, X_n được xem là độc lập và có cùng phân phối xác suất.

➤ Chương 5. Mẫu thống kê & Ước lượng tham số

1.2. Sắp xếp mẫu dựa vào số liệu thực nghiệm

a) Sắp xếp theo dạng bảng

VD 1. Kiểm tra ngẫu nhiên 50 sinh viên. Ta sắp xếp điểm số X thu được theo thứ tự tăng dần và số sinh viên n có điểm tương ứng vào bảng như sau:

X (điểm)	2	4	5	6	7	8	9	10
n (số SV)	4	6	20	10	5	2	2	1

b) Sắp xếp theo dạng khoảng

VD 2. Đo chiều cao X (cm) của $n = 100$ thanh niên. Vì chiều cao khác nhau nên để tiện việc sắp xếp, người ta chia chiều cao thành nhiều khoảng.

➤ Chương 5. Mẫu thống kê & Ước lượng tham số

Các thanh niên có chiều cao trong cùng 1 khoảng được xem là cao như nhau. Khi đó, ta có bảng số liệu ở dạng khoảng như sau:

X	148-152	152-156	156-160	160-164	164-168
n	5	20	35	25	15

Khi cần tính toán, người ta chọn *số trung bình* của mỗi khoảng để đưa số liệu trên về dạng bảng:

X	150	154	158	162	166
n	5	20	35	25	15

Chú ý

Đối với trường hợp số liệu được cho dưới dạng liệt kê thì ta sắp xếp lại ở dạng bảng.

➤ Chương 5. Mẫu thống kê & Ước lượng tham số**1.3. Các đặc trưng mẫu**

Xét một mẫu ngẫu nhiên (X_1, X_2, \dots, X_n) , ta có các đặc trưng mẫu như sau

a) Trung bình mẫu

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

Để đơn giản, ta dùng ký hiệu $\bar{X} = \bar{X}_n$.

b) Phương sai mẫu

- Phương sai mẫu

$$\hat{S}^2 = \hat{S}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

➤ Chương 5. Mẫu thống kê & Ước lượng tham số

- Phương sai mẫu hiệu chỉnh

$$S^2 = S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

- Trong tính toán cụ thể, ta sử dụng công thức

$$S^2 = \frac{n}{n-1} \left[\overline{(X^2)} - (\bar{X})^2 \right] = \frac{n}{n-1} \hat{S}^2$$

trong đó $\overline{(X^2)} = \frac{1}{n} \sum_{i=1}^n X_i^2$.

c) Tỷ lệ mẫu

Xét mẫu định tính với các biến ngẫu nhiên X_i ($i = 1, \dots, n$) có phân phối Bernoulli $B(1; p)$:

➤ Chương 5. Mẫu thống kê & Ước lượng tham số

$$X_i = \begin{cases} 0, & \text{nếu phần tử không có tính chất } A \\ 1, & \text{nếu phần tử có tính chất } A. \end{cases}$$

Nếu mẫu có m phần tử có tính chất A thì tỉ lệ mẫu là

$$F = F_n = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{m}{n}$$

d) Liên hệ giữa đặc trưng của mẫu và tổng thể

Các đặc trưng mẫu \bar{X} , S^2 , F là các thống kê dùng để nghiên cứu các đặc trưng μ , σ^2 , p tương ứng của tổng thể. Từ luật số lớn ta có:

$$F \rightarrow p, \bar{X} \rightarrow \mu, S^2 \rightarrow \sigma^2 \text{ (theo xác suất).}$$

Dùng máy tính bỏ túi để tính đặc trưng mẫu**SỬ DỤNG MÁY TÍNH BỎ TÚI ĐỂ TÍNH CÁC ĐẶC TRƯNG CỦA MẪU****1. Số liệu đơn (không có tần số)**

VD 1. Cho mẫu có cỡ mẫu là $n = 5$:
12; 13; 11; 14; 11.

a) Máy fx 500 – 570 MS

- **Xóa bộ nhớ:** **SHIFT** → **MODE** → **3** → **=** → **=**
- **Vào chế độ thống kê nhập dữ liệu:**
 - **MODE** → **2** (chọn SD đối với fx500MS);
 - **MODE** → **MODE** → **1** (chọn SD đối với fx570MS).
- **Nhập các số:**
12 M+ 13 M+ 11 M+ 14 M+ 11 M+