ƯỚC LƯỢNG

Thống kê

Xác suất

Tổng thể:

* KÍch thước lớn.
* Có những quan sát không thể tiến hành trên tổng thể.

Số lượng cá thể của mẫu ~ kích thước mẫu.

Chọn mẫu:

* Mỗi cá thể của mẫu được chọn ngẫu nhiên.
* Chọn mẫu: hoàn lại/ không hoàn lại.

X=(X1,X2,..,Xn)

Xi: Biến ngẫu nhiên.

Xi: có cùng phân phối xác suất.

Vì Xi là biến ngẫu nhiên ⇒ X là biến ngẫu nhiên.

X có cùng phân phối XS vơi Xi.

Với 1 bộ giá trị cụ thể của mẫu: mẫu cụ thể.

(1.57, 1.62, 1.45, 1.75, 1.68,…, 1.64)

* Biểu diễn mẫu cụ thể:
* Bảng phân phối tần số:
* Mẫu cụ thể:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| X | X1 | X2 | X3 | X4 | X5 | X6 | X7 |
|  | 10 | 9 | 15 | 10 | 14 | 10 | 15 |

* **Bảng phân phối tần số**:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X | 9 | 10 | 14 | 15 |
| ni | 1 | 3 | 1 | 2 |

* Bảng phân tổ thống kê:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| X | x1-x2 | x2-x3 | x3-x4 | .. | xk-xk+1 |
| ni | n1 | n2 | n3 | .. | nk |

xk-xK+1 = (xk,xk+1]

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| X | 1.40 -1.50 | 1.50-1.60 | 1.60-1.70 | 1.70-1.80 | 1.80-1.90 |
| ni | 30 | 45 | 100 | 15 | 10 |

* Chuyển đổi từ bảng phân tổ thống kê thành Bảng phân phối tần số:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| X | 1.40 -1.50 | 1.50-1.60 | 1.60-1.70 | 1.70-1.80 | 1.80-1.90 |
| ni | 30 | 45 | 100 | 15 | 10 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| X | (1.40 +1.50)/2 | (1.50+1.60)/2 | (1.60+1.70) | (1.70+1.80) | (1.80+1.90) |
| ni | 30 | 45 | 100 | 15 | 10 |

Thống kê:

Đn: Cho 1 mẫu ngẫu nhiên X=(X1,X2,..,Xn). Một hàm:

T = f(X) là một thống kê.

⇒ T là 1 biến ngẫu nhiên.

Một số thống kê thường gặp (tham số thống kê)

Xét mẫu: X=(X1,X2,..,Xn).

* Giá trị trung bình:

Trung bình mẫu cụ thể:

* Phương sai:

Phương sai mẫu:

Phương sai mẫu điều chỉnh:

Phương sai mẫu điều chỉnh của mẫu cụ thể:

* Độ lệch chuẩn:

Sd = sqrt(S2)

Nếu mẫu cụ thể biểu diễn bằng bảng phân phối tần số:

* Trung bình mẫu:
* Phương sai mẫu:

Chú ý:

* E() =
* var () =
* var(αX) = α2 var(X)

1. **ƯỚC LƯỢNG ĐIỂM**

* Ước lượng tham số thống kê bằng 1 giá trị.

Nếu giá trị thống kê (thực): θ.

Xét mẫu X=(X1,X2,..,Xn) ⇒ = T(X) ước lượng θ (ước lượng mẫu).

* Nếu E() = θ: Ước lượng không chệch.
* Đại lượng: E() - θ = Bias() : Độ chệch.

Ước lượng không chệch ⇔ Bias() = 0.

Có thể có nhiều ước lượng thống kê cho θ ⇒ Đánh giá tính hiệu quả của ước lượng.

* Ước lượng không chệch tốt hơn.
* Nếu có 2 ước lượng không chệch của θ là thì

var () < var () ⇒ ước lượng là tốt hơn.

Ví dụ:

Một Kỹ sư ước lượng năng suất trung bình của 1 qui trình hóa học. Quan sát 1 mẫu mẫu thực nghiệm có kích thước là 3 về năng suất qui trình hóa học X = (X1,X2,X3). Ước lượng giá trị trung bình:

Cách 1:

Cách 2:

=

Cách thức ước lượng nào là hiệu quả hơn?

* Kiểm tra các ước lượng có bị chệch không?

E() = E( = .

E() = E( = .

Cả 2 ước lượng đều không chệch.

* Kiểm tra var của các ước lượng.

var() =var(

var() = var(

=

Vậy: var() = var() =

Ước lượng là hiệu quả hơn.

Bài tập:

Xét mẫu X = (X1,X2), ước lượng giá trị trung bình:

1. Chứng tỏ các ước lượng là không chệch.
2. Ước lượng nào hiệu quả hơn.

Bài làm:

E) = E( = .

E() = E( = .

Cả 2 ước lượng đều không chệch.

* Kiểm tra var của các ước lượng.

var() =var(

var() = var(

=

Vậy: var() = var() =

Ước lượng là hiệu quả hơn.