

Thiết kế thí nghiệm – tính cỡ mẫu

Huỳnh Kim Ân- Cty CP tư vấn XD điện 4
email: hhkiman@gmail.com

Mở đầu

1. Bất kỳ nghiên cứu nào cũng dựa vào mẫu; nếu:
Ít quá => độ chính xác không cao/Đôi khi vô nghĩa;
Nhiều quá => tốn thời gian, chi phí.
Mục đích của ước tính cỡ mẫu là tìm số lượng mẫu đại diện ít nhất để suy luận cho quần thể đạt độ chính xác cao nhất và đầy đủ nhất.
2. Câu hỏi là: Cần bao nhiêu mẫu cho một nghiên cứu ?.
(Tại sao trong thí nghiệm bê tông, tổ mẫu là 3 mẫu để XD giá trị trung bình ?) [4]

Sau đây sẽ trả lời cho câu hỏi trên.

Ba yếu tố ảnh hưởng cỡ mẫu [1], [2]

1. Sai sót người thiết kế chấp nhận: cụ thể là sai sót loại I và II (type I, type II)

Sai sót: thường 1 nghiên cứu chấp nhận type I khoảng 1% đến 5% ($\alpha = 0.01$ hay 0.05);

type II khoảng $\beta = 0.1 - 0.2$ (power=0.8-0.9).

2. Độ dao động của phép đo lường, độ chính xác của phép đo được thể hiện qua độ lệch chuẩn (σ)/ hoặc hệ số biến thiên (CoV);

3. Mức độ khác biệt (ảnh hưởng) của thí nghiệm mà thiết kế muốn biết.

Sai sót loại I và loại II

[2]; [3]

Khái niệm power :

Thống kê, là phương pháp khoa học tìm cái chưa biết (hoặc không quan sát/quan sát không đầy đủ);

Để tìm cái chưa biết, chỉ có thể tiến hành thí nghiệm trên một nhóm đối tượng (mẫu đại diện) để kết luận cho quần thể. Trước tiên đặt giả thuyết

H- (gt đảo) hiện tượng không tồn tại;

H+ (gt chính) hiện tượng tồn tại.

Dùng các phương pháp kiểm định thống kê đánh giá khả năng của giả thuyết trên dựa vào trị số P-Value; đơn giản có 2 giá trị: Có ý nghĩa hoặc không có ý nghĩa thống kê ($P < 0.05$ hoặc $P > 0.05$);

có 4 tình huống sau:

Sai sót loại I và loại II

4 tình huống trong việc xem xét một giả thuyết khoa học

Chú thích: s trong bảng này: Có nghĩa là significant (có ý nghĩa thống kê); ns: non-significant (không có ý nghĩa thống kê); H+ là giả thuyết đúng; H- là giả thuyết sai .

Kết quả kiểm định th. kê	Giả thuyết H	
	Đúng	Sai
Có ý nghĩa thống kê $P < 0.05$	a. Gọi Power: $P(s H+) = 1 - \beta$	c. Gọi Type I error $P(s H-) = \alpha$
Không có ý nghĩa thống kê $P > 0.05$	b. Type II error: $P(ns H+) = \beta$	d. Gọi True Negative $P(ns H-) = 1 - \alpha$

Như vậy, **power** là xác suất mà kết quả kiểm định thống kê có $P < 0.05$ với điều kiện giả thuyết H đúng. Ví dụ power = 0.8 (80%), nếu làm 100 thí nghiệm cứu thì sẽ có 80% (hay 80) kết quả có $P < 0.05$.

vài ví dụ tính cỡ mẫu về thí nghiệm bê tông

Vd1 tính cỡ mẫu để tìm 1 giá trị trung bình

VD1: Tính cỡ mẫu để xác định 1 chỉ số trung bình với sai số trong khoảng 2 (Δ); khoảng tin cậy 80% (tức $\alpha = 20\%$), power = 0.8 (tức $\beta = 0.2$). Biết trước được loại thí nghiệm này với $\sigma = 3.5$.

$$n = C/(\Delta/ \sigma)^2 \quad (1)$$

C: phụ thuộc type I, type II; tra bảng sau:

$\alpha =$ (Các giá trị thường dùng)	$\beta = 0.2$ Power = 0.80	$\beta = 0.1$ Power = 0.9	$\beta = 0.05$ Power = 0.95
0.1	6.15	8.53	10.79
0.05	7.85	10.51	13
0.01	13.33	16.74	19.84

Dùng R tính [2] hoặc tham khảo [1]

- Trong R có hàm `power.t.test`
- `power.t.test(delta= Δ , sd= σ , sig.level= α , power= $1-\beta$, type='one.sample')`.

VD1: Code trong R như sau

```
> power.t.test(delta=2, sd=3.5, sig.level=.2, power=.8,  
type='one.sample')
```

R cho kết quả:

One-sample t test power calculation

$n = 14.69539$ quy tròn 15 mẫu

delta = 2

sd = 3.5

sig.level = 0.2

power = 0.8

alternative = two.sided.

VD2: cỡ mẫu so sánh 2 số trung bình

Có 2 nhóm cấp phối, nhóm 1 là nhóm không can thiệp nhóm 2 là nhóm can thiệp bằng phụ gia cải thiện cường độ. *Tiêu chí đánh giá là hiệu quả của phụ gia làm tăng cường độ.*

Cỡ mẫu ? chứng minh cường độ nhóm 2 tăng 20% so nhóm 1 ?. biết cường độ TK là 15 Mpa, $\sigma = 3.5$ Mpa, sig.level=.2, power=.8

$$\Delta = (15 * 1.2) - 15 = 3 \text{ Mpa.}$$

Công thức (1): $n = 2C / (\Delta / \sigma)^2$ hoặc bằng R

```
> power.t.test(delta=3, sd=3.5, sig.level=0.2, power=0.8, type="two.sample")
```

ví dụ 2 (tt)

Tính trong R với code như sau:

➤ `power.t.test(delta=3, sd=3.5,
sig.level=0.2,power=0.8,type="two.sample")`

Chú ý: type = 'two.sample' : cho 2 nhóm mẫu.

R cho kết quả:

Two-sample t test power calculation

n = 12.71981 ~ 13 mẫu cho mỗi nhóm thí nghiệm.

delta = 3

sd = 3.5

sig.level = 0.2 (khoảng tin cậy 80%)

power = 0.8 (80% thí nghiệm có $p < 0.05$)

alternative = two.sided

*NOTE: n is number in *each* group*

tầm quan trọng tính cỡ mẫu

Tính cỡ mẫu cho 1 nghiên cứu là một bước cực kì quan trọng, vì nó có thể quyết định thành bại của nghiên cứu.

Trước khi ước tính cỡ mẫu nhà nghiên cứu cần phải biết trước (hay ít ra là có vài giả thiết cụ thể) về vấn đề mình quan tâm; để có 3 thông số ban đầu.

Trường hợp nghiên cứu hoàn toàn mới, chưa ai từng làm trước đó, có thể các thông số về độ ảnh hưởng và độ dao động đo lường sẽ không có, nhà nghiên cứu cần phải tiến hành một số mô phỏng (simulation) hay một nghiên cứu sơ khởi để có những thông số cần thiết. Cách ước tính cỡ mẫu bằng mô phỏng là một lĩnh vực nghiên cứu chuyên sâu, ***không đề cập ở đây.***

tài liệu tham khảo

- [1]: TS Chế Đình Lý: “*Thống kê và xử lý dữ liệu môi trường*” NXB - ĐHQG TPHCM – 2014, trang 73-78;
- [2]: Nguyễn Văn Tuấn: “*Phân tích dữ liệu với R*” NXB Tổng hợp TPHCM 10/2014;
- [3]: Người dịch: Nguyễn Văn Minh Mẫn “*Thống kê công nghiệp hiện đại với ứng dụng viết trên R, MiniTab và JMP*” NXB Bách khoa Hà Nội 1/2016; trang 148 – 149.
- [4]: TCVN 3118: *Bê tông nặng-Phương pháp xác định cường độ nén – Mục 2.1*

Cảm ơn các bạn theo dõi !