

CỜ MẪU NGHIÊN CỨU

ƯỚC TÍNH CỠ MẪU

Dựa vào 3 yếu tố:

Sai sót α (0,01 hoặc 0,05) và β (0,1 hoặc 0,2)

Sự khác biệt muốn có

Độ lệch chuẩn (cho kết cục là biến số)

Trị số Z theo α hoặc β (test 2 đuôi)

α (β)	$Z_{2\alpha} / Z_{2\beta}$ (test 2 đuôi)
0,01	2,58
0,05	1,96
0,10	1,28
0,20	1,04

Cỡ mẫu cho nghiên cứu 1 nhóm

Ước tính chỉ số trung bình:

Ví dụ 1. Theo kết quả điều tra trị số huyết áp tâm thu của người lớn bình thường ở Việt Nam là $114 \pm 10\text{mmHg}$. Như vậy cỡ mẫu là bao nhiêu với khoảng tin cậy 95%, với sai sót $\alpha=0,05$ và sai sót $\beta=0,2$ (lực mẫu=0,8). Cỡ mẫu được tính theo công thức:

δ : là sai số mong muốn: 1 mm Hg

$$n = \frac{(Z_{\alpha} + Z_{\beta})^2 \sigma^2}{\delta^2}$$

$$n = \frac{(1,96 + 1,04)^2 \cdot 100}{1^2} = 900$$

Ví dụ 2. Tính cỡ mẫu để ước tính chiều cao đàn ông Việt Nam với sai số δ trong vòng 1cm. Biết rằng độ lệch chuẩn trong các nghiên cứu trước đây là 4,6cm.

$$n = \frac{(1,96 + 1,04)^2 \cdot 4,6^2}{1^2} = 190$$

Cỡ mẫu cho nghiên cứu 1 nhóm

Ước tính tỉ lệ :

Các nghiên cứu trước cho tỉ lệ SDD khoảng 20% ($p=0,20$)

Muốn điều tra tỉ lệ SDD với sai số ε là 2% ($\varepsilon=0,2$)

$$SE = \sqrt{\frac{p(1-p)}{n}}$$

$$\varepsilon = 1,96 \times SE$$

$$n = 1,96^2 \times \frac{p(1-p)}{\varepsilon^2}$$

$$n = 1,96^2 \cdot \frac{0,2(1-0,2)}{0,02^2} = 1536$$

Cỡ mẫu cho nghiên cứu 2 nhóm

Nghiên cứu RCT:

Kết cục là biến tỉ lệ:

$$n = \frac{\{Z_{2\alpha} \sqrt{2pq} + Z_{2\beta} \sqrt{p_1q_1 + p_2q_2}\}^2}{(p_1 - p_2)^2}$$

$\alpha=0,05$ thì $Z_{2\alpha}=1,96$

$p_1=0,80$ và $q_1=1-p_1=0,20$

$p_2=0,90$ và $q_2=1-p_2=0,10$

$\beta=0,20$ thì $Z_{2\beta}=1,04$

$p = \frac{p_1 + p_2}{2} = \frac{0,80 + 0,90}{2} = 0,85$ và $q = 1 - p = 0,15$

$$n = \frac{\{1.96 \sqrt{0.255} + 1.04 \sqrt{0.16 + 0.09}\}^2}{(0.9 - 0.8)^2} = 277$$

Cỡ mẫu cho nghiên cứu 2 nhóm

Nghiên cứu RCT:

Kết cục là biến số:

So sánh 2 thuốc điều trị THA, cho biết SD : 10mmHg và thuốc mới giảm HA tâm thu hơn thuốc cũ là 10 mmHg

$$n = 2 \left[\frac{(Z_{2\alpha} + Z_{2\beta}) \sigma}{\delta} \right]^2$$

$\delta = 10$ mmHg và $\sigma = 10$ mmHg. Với $\alpha=0,05$ thì $Z_{2\alpha}=1,96$; $\beta=0,10$ thì $Z_{2\beta}=1,28$.

$$n = 2 \left[\frac{(1,96 + 1,28) \cdot 10}{10} \right]^2 = 21$$

Cỡ mẫu cho nghiên cứu 2 nhóm

Nghiên cứu Cohort:

Ước tính RR là bao nhiêu?

p_1 : tỉ lệ bệnh nhóm không PN

p_2 : tỉ lệ bệnh nhóm PN : $p_2 = p_1 \cdot RR$

Ví dụ: nhóm chứng có tỉ lệ phát sinh (incidence) bệnh tim bẩm sinh là 8 phần ngàn (0,008), thì $p_2 = 2 \times 0,008 = 0,016$

Với $p_1 = 0,008$ thì $q_1 = 1 - p_1 = 0,992$

$p_2 = 0,016$ thì $q_2 = 0,984$

$p = (p_1 + p_2)/2 = 0,012$ và $q = 1 - p = 0,988$

Với $\alpha = 0,05$ thì $Z_{2\alpha} = 1,96$ và $\beta = 0,10$ thì $Z_{2\beta} = 1,28$

$$n = \frac{\{Z_{2\alpha} \sqrt{2pq} + Z_{2\beta} \sqrt{p_1q_1 + p_2q_2}\}^2}{(p_1 - p_2)^2}$$

$$n = \frac{\{1,96 \sqrt{0,0237} + 1,28 \sqrt{0,0079 + 0,0157}\}^2}{(-0,008)^2} = 3884$$

Cỡ mẫu cho nghiên cứu 2 nhóm

Nghiên cứu bệnh-chứng:

Ước tính OR là bao nhiêu?

p1: tỉ lệ bệnh nhóm chứng

p2: tỉ lệ bệnh nhóm bệnh

$$\text{OR} = \frac{p_2}{1-p_2} / \frac{p_1}{1-p_1}$$

Thế OR=2 và p₁=0,20 ta tính được p₂=0,333

$$\text{Và } p = \frac{p_1 + p_2}{2} = 0,266 \text{ và } q = 0,734$$

$$n = \frac{\{Z_{2\alpha} \sqrt{2pq} + Z_{2\beta} \sqrt{p_1q_1 + p_2q_2}\}^2}{(p_1 - p_2)^2}$$

$$n = \frac{\{1.96 \sqrt{0.39} + 1.28 \sqrt{0.16 + 0.22}\}^2}{(0.33 - 0.2)^2} = 228$$

ƯỚC TÍNH CỠ MẪU

Phân tích sống sót

$$n = \frac{2 (Z_{2\alpha} + Z_{2\beta})^2}{\log_e(\theta)^2} \quad (5)$$

Trong nghiên cứu này, biến kết cục được tính theo thời gian sống sót trung bình nên nhà nghiên cứu phải ước tính được θ là tỉ số giữa thời gian sống sót trung bình giữa nhóm thực nghiệm (ví dụ μ_E là 3 năm) và nhóm chứng (ví dụ μ_C là 2 năm). Như vậy $\theta = \mu_E/\mu_C = 3/2 = 1,5$

Với $\alpha=0,05$ thì $Z_{2\alpha}=1,96$

$\beta=0,20$ thì $Z_{2\beta}=1,04$

Thế vào công thức (5) ta có:

$$n = \frac{2 (1,96 + 1,04)^2}{\log_e(1,5)^2} = 110$$

ƯỚC TÍNH CỠ MẪU

Phân tích tương quan

$$n = 3 + \frac{4C(\alpha, \beta)}{\left[\log \left(\frac{1+r}{1-r} \times \frac{1-r_0}{1+r_0} \right) \right]^2}$$

$C(\alpha, \beta)$ theo sai sót α và β theo bảng dưới đây:

α	$\beta=0.10$ (Lực mẫu 0.90)	$\beta=0.20$ (Lực mẫu 0.80)
0.01	16.74	13.33
0.05	10.51	7.85

ƯỚC TÍNH CỠ MẪU

Phân tích tương quan

$$n = 3 + \frac{4C(\alpha, \beta)}{\left[\log \left(\frac{1+r}{1-r} \times \frac{1-r_0}{1+r_0} \right) \right]^2}$$

Ví dụ: Một nghiên cứu về liên quan giữa lượng đường huyết (mg/dL) và tỉ lệ % HbA1C trong hồng cầu. Các nghiên cứu trước đây cho biết mối tương quan giữa 2 trị số này vào khoảng 60% ($r = 0,60$). Tính cỡ mẫu cần cho nghiên cứu này.

Với $\alpha = 0,05$ và $\beta = 0,20$ thì $C=7,85$

Và $r = 0,60$. Ta tính được cỡ mẫu là:

$$n = 3 + \frac{31.4}{\left[\log \left(\frac{1.6}{0.4} \right) \right]^2} = 35$$