

Bài giảng 11: So sánh 2 tỉ lệ

Nguyễn Văn Tuấn

Garvan Institute of Medical Research, Australia
Đại học Tôn Đức Thắng, Việt Nam

Chúng ta học ...

Chúng ta đã học

- Cách tính 1 tỉ lệ

Chúng ta sẽ học

- Phương pháp so sánh 2 tỉ lệ

Nhưng tỉ lệ đến từ nhiều dạng khác nhau

Nhiều dạng dữ liệu khác nhau ...

- Nghiên cứu lâm sàng đối chứng ngẫu nhiên
- Nghiên cứu cắt ngang (thiết diện), cross-sectional study
- Nghiên cứu bệnh chứng

Zoledronate và gãy xương

Table 2. Rates of Fracture and Death in the Study Groups.*

Variable	Placebo	Zoledronic Acid	Hazard Ratio (95% CI)	P Value
Fracture — no. (cumulative %)				
Any	139 (13.9)	92 (8.6)	0.65 (0.50–0.84)	0.001
Nonvertebral	107 (10.7)	79 (7.6)	0.73 (0.55–0.98)	0.03
Hip	33 (3.5)	23 (2.0)	0.70 (0.41–1.19)	0.18
Vertebral	39 (3.8)	21 (1.7)	0.54 (0.32–0.92)	0.02
Death — no. (%)	141 (13.3)	101 (9.6)	0.72 (0.56–0.93)	0.01

* Rates of clinical fracture were calculated by Kaplan–Meier methods at 24 months and therefore are not simple percentages. There were 1062 patients in the placebo group, and 1065 in the zoledronic acid group. Because of variable follow-up, the number and percentage of patients who died are provided on the basis of 1057 patients in the placebo group and 1054 patients in the zoledronic acid group in the safety population.

Randomized controlled clinical trial

Placebo n = 1062, Zoledronate n = 1065

Length of follow-up: 3 years

Lyles KW, et al. Zoledronic acid and clinical fractures and mortality after hip fracture. *N Engl J Med* 2007;357. DOI: 10.1056/NEJMoa074941

Smoking và lung cancer

	Lung Cancer	Controls
Smokers	647	622
Non-smokers	2	27

R Doll and B Hill. BMJ 1950; ii:739-748



Sir Richard Doll (1912 – 2005)

http://en.wikipedia.org/wiki/Richard_Doll

Is there an association between smoking and lung cancer?

Mortality in the Titanic incident



Class	Dead	Survived	Total
I	123	200 (62%)	323
II	158	119 (43%)	277
III	528	181 (26%)	709
Total	809	500 (38%)	1309

<http://lib.stat.cmu.edu/S/Harrell/data/descriptions/titanic3info.txt>

Is there an association between passenger class and death?

So sánh 2 nhóm z-test

Binomial distribution – phân bố nhị phân

Population (true) proportion:
 π



Sample proportion:
 p

Lí thuyết:

- $p = (\text{cases} / \text{tổng số}) = x / N$
 p là ước số khách quan của π
- **Độ lệch chuẩn (standard deviation)** của p : $s = \sqrt{\frac{p(1-p)}{N}}$
- **KTC 95% của π is:** $p \pm 1.96 \times S$

So sánh 2 nhóm: Sample và population

	Sample (mẫu)		Population (quần thể)	
	Group 1	Group 2	Group 1	Group 2
N	n_1	n_2	Infinite	Infinite
Xác suất outcome	p_1	p_2	$\pi_1 = ?$	$\pi_2 = ?$
Hiệu số	$d = p_1 - p_2$		$\delta = \pi_1 - \pi_2$	
Tình trạng	Known		Unknown	

Mục tiêu: dùng dữ liệu của mẫu để suy luận cho quần thể

Phân tích sao sánh 2 nhóm

	Sample (mẫu)	
	Nhóm 1	Nhóm 1
N	n_1	n_2
Xác suất outcome	p_1	p_2
Độ lệch chuẩn	s_1	s_2

Hiệu số ảnh hưởng

$$d = p_1 - p_2$$

Độ lệch chuẩn của d

$$s = \sqrt{s_1^2 + s_2^2}$$

$$z \text{ test} = d / s$$

$$\text{KTC95\% of } d = d \mp 1.96s$$

Hiệu quả chống gãy xương của zoledronic acid

	Placebo	Zoledronic acid
Số bệnh nhân	1062	1065
Số ca gãy xương	139	92
Không gãy xương	923	973
Tỉ lệ gãy xương	0.131	0.086
Độ lệch chuẩn	0.0103	0.0086

Hiệu số ảnh hưởng

$$d = 0.131 - 0.086 = \mathbf{0.045}$$

Độ lệch chuẩn của d

$$s = \sqrt{0.0103^2 + 0.0086^2} = \mathbf{0.013}$$

KTC95%

$$0.045 \pm 1.96 * s = \mathbf{0.018 \text{ đến } 0.071}$$

Z test

$$z = 0.045 / 0.013 = 3.30$$

P-value

$$2 * (1 - \text{pnorm}(3.30)) = 0.0009$$

Diễn giải d và KTC95%

- Nếu zoledronic acid không có hiệu quả
 - $d = 0$
 - KTC95% của d dao động từ âm đến dương
- Nhưng kết quả cho thấy
 - $d \neq 0$
 - KTC95% của d đều dương
- Do đó, zoledronic acid có hiệu quả giảm nguy cơ gãy xương

Phương pháp Bayes

So sánh 2 tỉ lệ

	Drug A	Drug B
Có ảnh hưởng phụ	11	5
Không có ảnh hưởng phụ	9	15

- Hai nhóm có nguy cơ *side effects* giống nhau?

Phân tích cổ điển

- Gọi $X \sim \text{Binomial}(n_1, \pi_1)$ and $p_1 = X / n_1$
- Gọi $Y \sim \text{Binomial}(n_2, \pi_2)$ and $p_2 = Y / n_2$
- Giả thuyết $\mu_1 = \mu_2$

- Z-test

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

Phân tích cổ điển

	Drug A	Drug B
Có ảnh hưởng phụ	11	5
Không có ảnh hưởng phụ	9	15
Tỉ lệ p	0.55	0.25
Độ lệch chuẩn	0.111	0.097

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$
$$z = \frac{0.111 - 0.097}{\sqrt{(0.111)^2 + (0.097)^2}} = 2.03$$

```
> 2*(1-pnorm(2.03))  
[1] 0.04235654
```


Dùng R

```
prop.test(x=c(11, 5), n=c(20,20), correct=FALSE)
```

```
> prop.test(x=c(11, 5), n=c(20,20), correct=T)
```

2-sample test for equality of proportions without
continuity correction

data: c(11, 5) out of c(20, 20)

X-squared = 3.75, df = 1, p-value = 0.05281

alternative hypothesis: two.sided

95 percent confidence interval:

0.01094684 0.58905316

sample estimates:

prop 1 prop 2

Khác biệt có ý nghĩa thống kê!

Phương pháp chính xác: ExactCldiff

```
library(ExactCldiff)
```

```
BinomCI(20, 20, 11, 5)
```

```
> BinomCI(20, 20, 11, 5)
```

```
$estimate
```

```
[1] 0.3
```

```
$ExactCI
```

```
[1] -0.02405 0.57023
```

Khác biệt *không* có ý nghĩa thống kê!

Phân tích Bayesian

- Thông tin tiền định cho μ_1 và μ_2 : $\text{Beta}(\alpha_1, \beta_1)$ và $\text{Beta}(\alpha_2, \beta_2)$
- Thông tin hậu định:

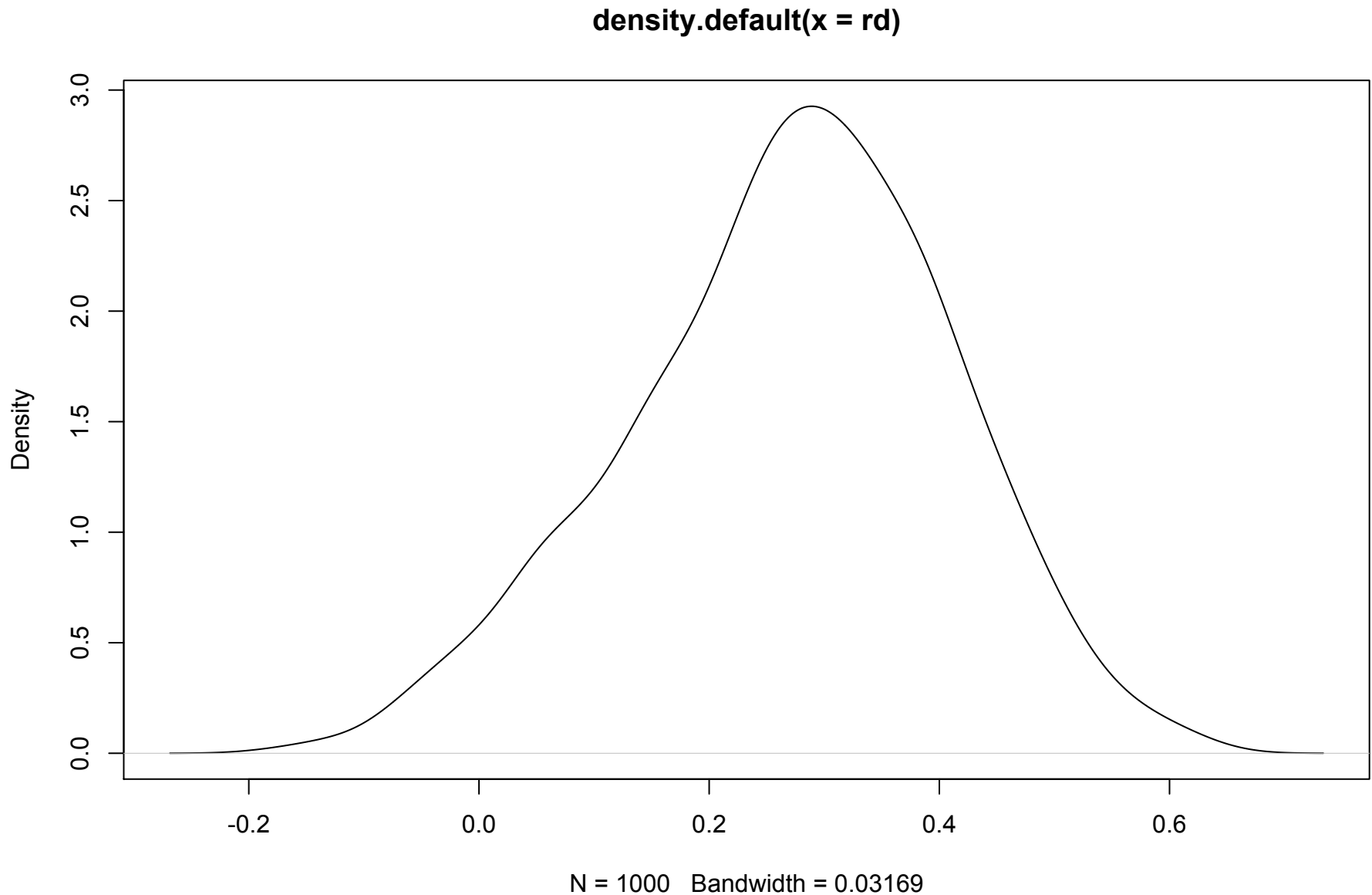
$$\pi(p_1, p_2) \propto p_1^{x+\alpha_1-1}(1-p_1)^{n_1+\beta_1-1} \times p_2^{y+\alpha_2-1}(1-p_2)^{n_2+\beta_2-1}$$

Chúng ta có thể mô phỏng bằng R

R analysis

```
x = 11; n1 = 20; alpha1 = 1; beta1 = 1  
y = 5; n2 = 20; alpha2 = 1; beta2 = 1  
p1 = rbeta(1000, x + alpha1, n1 - x + beta1)  
p2 = rbeta(1000, y + alpha2, n2 - y + beta2)  
rd = p1 - p2  
  
plot(density(rd))  
quantile(rd, c(.025, 0.5, 0.975))  
mean(rd) # Tìm trung bình của hiệu số  
median(rd) # Tìm trung vị của hiệu số
```

Phân bố của hiệu số (rd)



Khoảng tin cậy 95%

```
> quantile(rd, c(.025, 0.5, 0.975))  
      2.5%      50%      97.5%  
-0.02531479  0.27981356  0.52593816
```

Kết luận: Chưa đủ bằng chứng để nói tỉ lệ biến chứng giữa hai nhóm khác nhau

So sánh 2 tỉ lệ

- Phương pháp cổ điển: z-test
- Phương pháp chính xác: **ExactCldiff**
- Phương pháp Bayes: mô phỏng
- Các phương pháp có thể cho ra kết quả khác nhau.