

Cơ bản về thiết kế thí nghiệm

Đỗ Trọng Hợp

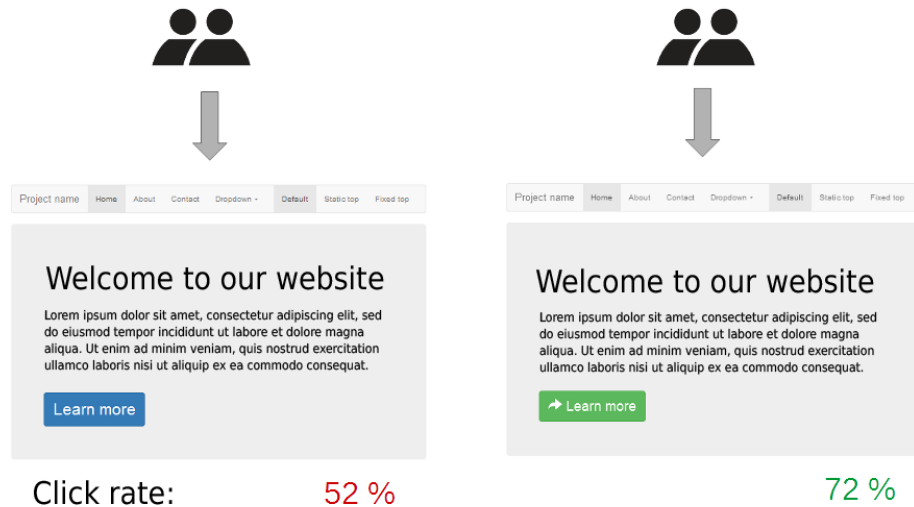
Khoa Khoa Học và Kỹ Thuật Thông Tin

Đại học Công Nghệ Thông Tin tp. Hồ Chí Minh

Giới thiệu

- Mục đích của thí nghiệm là để kiểm tra giả thuyết
- Kết luận được đưa ra thông qua việc so sánh
- Ví dụ
 - Agriculture experiment in which the effect of different fertilizers are compared
 - Medical experiments in which the experimental units are patients and the treatments evaluated might be a new medication, perhaps at different dosage levels, and a placebo
 - Industrial experiments in which different product designs or manufacturing processes are to be compared
 - Market research experiments in which the experimental units are consumers and the treatments are different advertising presentations
 - Education experiments in which the experimental units are groups of children and the treatments are different teaching materials or methods
 - A/B testing used by marketers, designers, software engineers to compare which variant of the tested item is more effective

A/B testing



Control

[Esurance® Auto Insurance - You Could Save 28% with Esurance.](#) Ads
www.esurance.com/California
Get Your Free Online Quote Today!

Treatment

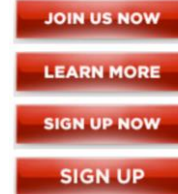
[Esurance® Auto Insurance - You Could Save 28% with Esurance.](#) Ads
www.esurance.com/California
Get Your Free Online Quote Today!
[Get a Quote](#) · [Find Discounts](#) · [An Allstate Company](#) · [Compare Rates](#)

- Tổng quát

- 2 phiên bản (A và B) của 1 thực thể được so sánh. 2 phiên bản này có mọi yếu tố giống nhau trừ yếu tố đang cần kiểm tra
- Những trang lớn như LinkedIn, Facebook, và Instagram đều sử dụng A/B testing để nâng cao user experiences.

Obama fundraising website a/b testing

The Winner - 40.6% in sign-up rate - \$60m



Combinations (24)		Page Sections (2)		Download: XML CSV TSV Print			
Relevance Rating ?	Variation	Est. conv. rate ?	Chance to Beat Orig. ?	Observed Improvement ?	Conv.Visitors ?		
5 / 5	Button	Original	7.51% ± 0.2%	—	—	5851 / 77858	
		Learn More	8.91% ± 0.2%	100%	18.6%	6927 / 77729	
		Join Us Now	7.62% ± 0.2%	73.5%	1.37%	5915 / 77644	
		Sign Up Now	7.34% ± 0.2%	13.7%	-2.38%	5660 / 77151	
5 / 5	Media	Original	8.54% ± 0.2%	—	—	4425 / 51794	
		Family Image	9.66% ± 0.2%	100%	13.1%	4996 / 51696	
		Change Image	8.87% ± 0.2%	92.2%	3.85%	4595 / 51790	
		Barack's Video	7.76% ± 0.2%	0.04%	-9.14%	3992 / 51427	
		Sam's Video	6.29% ± 0.2%	0.00%	-26.4%	3261 / 51864	
		Springfield Video	5.95% ± 0.2%	0.00%	-30.3%	3084 / 51811	

Combinations (24)		Page Sections (2)		Download: XML CSV TSV Print			
Disable	All Combinations (24)	Key: Winner Inconclusive Loser					
<input type="checkbox"/> Combination	Status	Est. conv. rate	Chance to Beat Orig.	Observed Improvement	Conv.Visitors		
Original	Enabled	8.26% ± 0.5%	—	—	1088 / 13167		
★ Top high-confidence winners. Run a follow-up experiment							
<input type="checkbox"/> Combination 11	Enabled	11.6% ± 0.6%	100%	40.6%	1504 / 12947		
<input type="checkbox"/> Combination 7	Enabled	10.3% ± 0.6%	100%	24.0%	1340 / 13073		
<input type="checkbox"/> Combination 3	Enabled	9.80% ± 0.6%	99.7%	18.7%	1277 / 13025		
<input type="checkbox"/> Combination 10	Enabled	9.23% ± 0.6%	95.9%	11.7%	1203 / 13031		
<input type="checkbox"/> Combination 8	Enabled	9.03% ± 0.6%	91.6%	9.28%	1178 / 13046		
<input type="checkbox"/> Combination 9	Enabled	8.77% ± 0.6%	81.8%	6.10%	1111 / 12672		
<input type="checkbox"/> Combination 6	Enabled	8.64% ± 0.5%	75.3%	4.58%	1108 / 12822		

Thuật ngữ và khái niệm

- **Yếu tố (Factor)**

- Là mọi loại biến số độc lập trong thí nghiệm
- Giá trị của biến số này được thiết lập bởi thí nghiệm
- Có thể là biến số định lượng hoặc định tính, gián đoạn hoặc liên tục.
- Có thể có hoặc không có ảnh hưởng đến kết quả.

- **Kết quả (Response)**

- Là kết quả thu đo được của thí nghiệm
- Có thể gọi là biến số phụ thuộc (vào các yếu tố được thiết lập) của thí nghiệm

- **Yếu tố chính (Primary factor)**

- Là yếu tố cần tìm hiểu ảnh hưởng của nó lên kết quả

- Ví dụ:

- Thí nghiệm tác dụng của 2 loại phân bón A và B lên sản lượng lúa
 - Có 1 yếu tố chính là phân bón.
- Thí nghiệm tác dụng của 3 loại phân bón A, B, C và nhiệt độ lên sản lượng lúa
 - Có 2 yếu tố chính là (1) phân bón và (2) nhiệt độ
- Thí nghiệm tác dụng của 3 loại phân bón A, B, C lên sản lượng lúa. Trong thí nghiệm, nhiệt độ được giữ ở mức cố định.
 - Có 1 yếu tố chính là phân bón. Nhiệt độ là yếu tố cố định (constant factor).
- Thí nghiệm tác dụng của 3 loại phân bón A, B, C lên sản lượng lúa. Trong thí nghiệm, nhiệt độ ở mức ngẫu nhiên.
 - Có 1 yếu tố chính là phân bón. Nhiệt độ là yếu tố ngẫu nhiên (random factor).

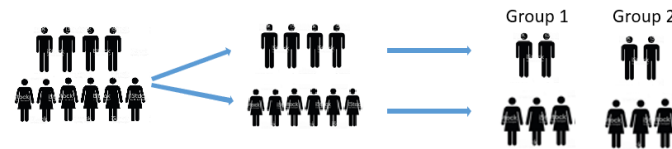
Tìm các yếu tố chính trong các thí nghiệm sau

- **So sánh tác dụng của 2 loại thuốc A và B lên lượng đường máu.** Đối tượng thí nghiệm gồm 100 người nam nữ lẫn lộn. Chọn 50 người ngẫu nhiên thử thuốc A và 50 còn lại thử thuốc B.
- 1 yếu tố chính là loại thuốc. Giới tính là yếu tố ngẫu nhiên (random factor).
- **So sánh tác dụng của 2 loại thuốc A và B lên lượng đường máu.** Đối tượng thí nghiệm gồm 100 nam giới. Chọn 50 người ngẫu nhiên thử thuốc A và 50 còn lại thử thuốc B.

- 1 yếu tố chính là loại thuốc. Giới tính là yếu tố cố định (constant factor).

- **So sánh tác dụng của 2 loại thuốc A và B lên lượng đường máu.** Đối tượng thí nghiệm gồm 40 nam và 60 nữ. 50 đối tượng (20 nam và 30 nữ) thử thuốc A. 50 đối tượng còn lại thử thuốc B.

- 1 yếu tố chính là loại thuốc. Giới tính là yếu tố khối (blocking factor).



- **So sánh tác dụng của 2 loại thuốc A và B và lượng nước uống mỗi ngày lên lượng đường máu.** Đối tượng thí nghiệm gồm 100 người nam nữ lẫn lộn. Chọn 50 người ngẫu nhiên thử thuốc A. 50 người còn lại thử thuốc B. Trong mỗi nhóm, chọn 25 người ngẫu nhiên cho uống nước 1L/ngày, 25 người còn lại uống 2L/ngày.

- 2 yếu tố chính là loại thuốc và lượng nước uống. Giới tính là yếu tố ngẫu nhiên.

Thuật ngữ và khái niệm

- Mức (Level)
 - Là loại hình hoặc giá trị của 1 yếu tố.
- Ví dụ:
 - Yếu tố loại thuốc: 2 level gồm loại A và loại B
 - Yếu tố nhiệt độ: vô số level là các giá trị liên tục của nhiệt độ
 - Yếu tố lượng giới tính: 2 level gồm nam và nữ
 - Yếu tố mật độ trồng lúa: 3 level gồm 10, 20, 40 khóm/m²

Thuật ngữ và khái niệm

- **Nghiệm thức (Treatment)**
 - Là những mức của một yếu tố chính được thử nghiệm để so sánh.
- **Ví dụ**
 - So sánh tác dụng thuốc. Mỗi nhóm bệnh nhân được thử (treated) 1 loại thuốc khác nhau. Mỗi loại thuốc là một nghiệm thức (treatment)
 - Thí nghiệm ảnh hưởng của tốc độ lên độ mòn bánh xe. Mỗi xe được thử 1 tốc độ. Mỗi tốc độ là 1 nghiệm thức.
 - So sánh năng suất giống lúa. Mỗi ô ruộng được trồng 1 giống lúa. Mỗi giống lúa là một nghiệm thức.

Chỉ ra các nghiệm thức

- So sánh tác dụng của 3 loại phân bón A, B, C lên năng suất lúa.
 - Mỗi loại phân bón là 1 nghiệm thức.
- So sánh tác dụng của 2 loại thuốc lên lượng đường máu bệnh nhân. Đối tượng thử gồm 50 người nam và 50 người nữ.
 - Yếu tố chính là loại thuốc. Mỗi loại thuốc là 1 treatment.
- So sánh 2 phương pháp dạy. Có 10 lớp. Mỗi lớp dạy 1 pp.
 - Yếu tố chính là pp dạy. Mỗi pp là 1 treatment.
- So sánh 3 mật độ trồng lúa. Có 10 ô ruộng. Mỗi ô trồng 1 mật độ.
 - Yếu tố chính là mật độ trồng. Mỗi mức mật độ là 1 treatment.

Thuật ngữ và khái niệm

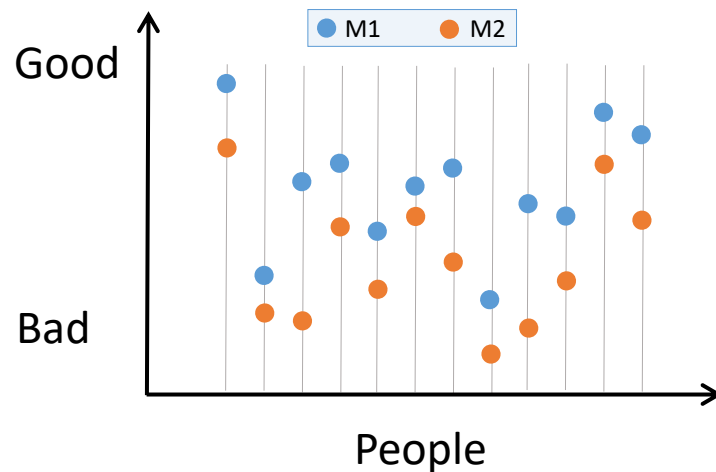
- Đơn vị thí nghiệm (Experiment unit)
 - Là đối tượng nhận các treatment.
 - Mỗi đối tượng có thể nhận 1 treatment khác nhau.
- Ví dụ
 - Thí nghiệm thử thuốc. Mỗi bệnh nhân thử thuốc là 1 đơn vị thí nghiệm.
 - So sánh năng suất 2 giống lúa. Có 10 ô ruộng, 5 ô trồng giống A, 5 ô trồng giống B. Mỗi ô ruộng là 1 đơn vị thí nghiệm.
 - So sánh tác dụng 2 loại thức ăn. Có 10 ao cá. 5 ao thử thức ăn loại A, 5 ao thử loại B. Mỗi ao cá là 1 đơn vị thí nghiệm.
 - So sánh tác dụng 2 loại thuốc tiêm. Có 10 chuồng gia súc. Mỗi con (có mã số riêng) tiêm 1 loại. Mỗi con là 1 đvtn.
 - So sánh phương pháp dạy. Có 10 lớp, 5 lớp thử pp offline, 5 lớp thử pp online. Mỗi lớp là 1 đvtn.
 - So sánh 2 loại vật liệu làm giày. Có 10 người, mỗi người thử 1 đôi giày làm từ 1 loại vật liệu. Mỗi người là 1 đvtn.
 - So sánh 2 loại vật liệu làm giày. Có 10 người, mỗi người mang 2 chiếc giày làm từ 2 loại vật liệu. Mỗi chân là 1 đvtn.

Tìm đvtn trong các thí nghiệm sau

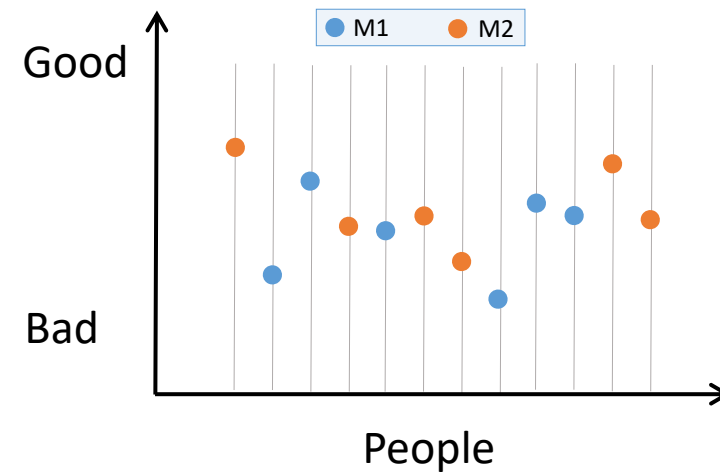
- So sánh 2 phương pháp đọc. Có 10 lớp, mỗi lớp 10 hs. Mỗi lớp học 1 pp đọc. Điểm thi của mỗi hs được ghi nhận để so sánh 2 pp.
 - Mỗi lớp là 1 đvtn. (nếu hs là đvtn thì mỗi hs có thể học 1 pp khác nhau, thực tế mọi học sinh trong 1 lớp chỉ học cùng 1 pp)
- Thí nghiệm 2 loại đồng phục thể dục. Có 10 lớp, mỗi lớp có 10 hs. Mỗi hs mặc 1 đồng phục. Độ thoải mái do mỗi hs đánh giá sẽ được ghi nhận để so sánh.
 - Mỗi hs là 1 đvtn. (2 hs trong cùng lớp có thể mặc 2 loại đồng phục khác nhau)
- So sánh tác dụng 2 loại phân bón lên trọng lượng quả bắp. Các cây bắp được trồng trong 10 ô ruộng, mỗi ô thử 1 loại phân bón. Trọng lượng các quả bắp được cân để so sánh.
 - Mỗi ô ruộng là 1 đvtn. (nếu mỗi cây bắp là 1 đvtn thì mỗi cây có thể được bón 1 loại phân khác nhau, thực tế không thể)
- So sánh tác dụng 2 loại phân bón lên trọng lượng quả bắp. Có 100 cây bắp được trồng cách xa nhau. Mỗi cây bón 1 loại phân. Trọng lượng của 100 quả bắp được cân để so sánh tác dụng phân bón.
 - Mỗi cây bắp là 1 đvtn.

Tại sao việc chọn đvtn quan trọng

- So sánh vật liệu sx giày



1 bàn chân là 1 đvtn



1 người là 1 đvtn

- Do cảm nhận mỗi người khác nhau. Nếu người là đvtn, mỗi người chỉ mang 1 loại giày, kết quả sẽ có độ dao động lớn hơn

Thuật ngữ và khái niệm

- Đơn vị thu thập dữ liệu (Measurement unit or Response unit)
 - Là thực thể mà kết quả được thu thập từ đó
 - Đvttdl có thể khác hoặc cùng là đơn vị thí nghiệm
- Ví dụ
 - Thí nghiệm thử thuốc. Mỗi bệnh nhân thử thuốc là 1 đơn vị thí nghiệm. Mỗi bệnh nhân được đo mức đường máu để so sánh. Mỗi bệnh nhân là 1 đvtpdl.
 - So sánh 2 loại thức ăn cho cá. Có 10 ao, mỗi ao nuôi 1000 con và cho ăn 1 loại thức ăn. Mỗi ao cá là 1 đvtn. Đo trọng lượng của 200 con cá bắt từ 10 ao để so sánh. 200 con cá này là các đvttdl.
 - So sánh tác dụng phân bón lá. Mỗi ô ruộng bón 1 loại phân. Đo kích thước của 1000 chiếc lá để so sánh. Mỗi ô ruộng là 1 đvtn. 1000 chiếc lá là các đvttdl.

Chỉ ra đvtn và đvttdl

- So sánh 2 phương pháp đọc. Có 10 lớp, mỗi lớp 10 hs. Mỗi lớp học 1 pp đọc. Điểm thi trung bình của mỗi lớp được ghi nhận để so sánh 2 pp.
 - Mỗi lớp là 1 E unit. Mỗi lớp là 1 M unit.
- Thí nghiệm 2 loại đồng phục thể dục. Có 10 lớp, mỗi lớp có 10 hs, mỗi hs mặc 1 loại. Độ thoải mái của mỗi hs được ghi nhận để so sánh.
 - Mỗi hs là 1 E unit. Mỗi hs là 1 M unit.
- So sánh tác dụng 2 loại phân bón lên trọng lượng quả bắp. Các cây bắp được trồng trong 10 ô ruộng, mỗi ô thử 1 loại phân bón. Trọng lượng của 100 quả bắp được cân để so sánh.
 - 10 ô ruộng là 10 E unit. 100 quả bắp là 100 M unit.
- So sánh tác dụng 2 loại phân bón lên trọng lượng quả táo. Có 20 cây táo được trồng cách xa nhau. Mỗi cây bón 1 loại phân. Trọng lượng của 100 quả táo từ 20 cây được cân để so sánh tác dụng phân bón.
 - 20 cây táo là 20 E unit. 100 quả táo được hái xuống là 100 M unit.
- So sánh tác dụng 2 loại phân bón lên sản lượng táo. Có 20 cây táo trồng cách xa nhau. Mỗi cây bón 1 loại phân. Tổng trọng lượng quả của mỗi cây sẽ được cân để so sánh.
 - 20 cây táo là 20 E unit. 20 cây táo là 20 M unit.

Chọn Measurement unit

- So sánh 2 giống bắp. Có 10 ruộng, mỗi ruộng trồng 1 giống. Hái 100 quả từ 10 ruộng để so sánh.

- Mỗi ruộng là 1 E unit. Mỗi trái bắp là 1 M unit.



- Cây phía ngoài nhận nhiều ánh sáng hơn ở giữa nên ra trái to hơn.
- Kết quả thu được dao động nhiều hơn.

- So sánh 2 giống bắp. Có 10 ruộng, mỗi ruộng trồng 1 giống. Trọng lượng trung bình của các trái bắp ở mỗi ruộng được so sánh.

- Mỗi ruộng là 1 E unit. Mỗi ruộng là 1 M unit.

- Sự khác nhau của cây ngoài và giữa không ảnh hưởng đến trọng lượng trung bình mỗi ruộng.
- Kết quả thu được dao động ít hơn.

Thuật ngữ và khái niệm

- Kết quả thay thế (surrogate response)
 - Kết quả thu thập thay thế cho kết quả chính (primary response)
- Ví dụ
 - Nghiên cứu về tác dụng của thuốc kéo dài sự sống sau khi bị đột quỵ. Primary response là số năm bệnh nhân sống được sau khi bị đột quỵ. Surrogate response có thể là tỷ lệ bệnh nhân vẫn sống sau 5 năm.
 - Nghiên cứu về thời gian training của thuật toán. Primary response là thời gian để đạt độ chính xác mong muốn 99%. Surrogate response có thể là thời gian để đạt được độ chính xác 90%.

Obama fundraising website a/b testing

Chỉ ra primary factor, levels, experiment unit, measurement unit, treatment, response



- 2 yếu tố chính: loại media và loại button
- Yếu tố media có 6 levels. Button có 4 levels.
- Experiment unit là 1 section người dùng vào website.
- Treatment là 1 combination treatment của [media;button]
- Measurement unit là 1 section người dùng vào website.
- Response là kết quả người dùng **có** hoặc **không** đăng ký.

Nguyên tắc của tkpttn

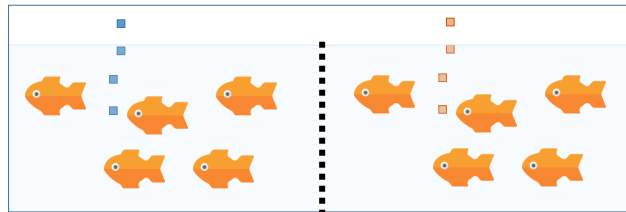
- Replication
- Randomization
- Blocking

Replication

- Thí nghiệm 2 loại thức ăn cho cá. Có 1 bể cá ngăn làm đôi, mỗi bên thả 1 con cá, mỗi con ăn 1 loại thức ăn. Trọng lượng mỗi con cá được đo để so sánh.



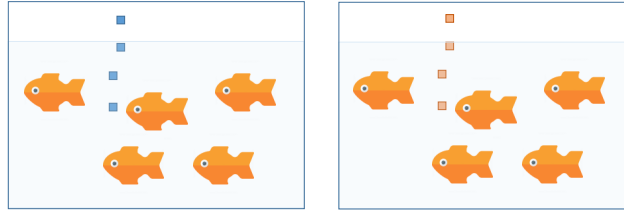
- Vấn đề: mỗi con cá có gene phát triển khác nhau nên không thể kết luận về tác dụng của thức ăn
- Replication principle:** the experiment should be repeated more than once. Thus, each treatment is applied in **many experimental units** instead of one. By doing so the statistical accuracy of the experiments is increased.



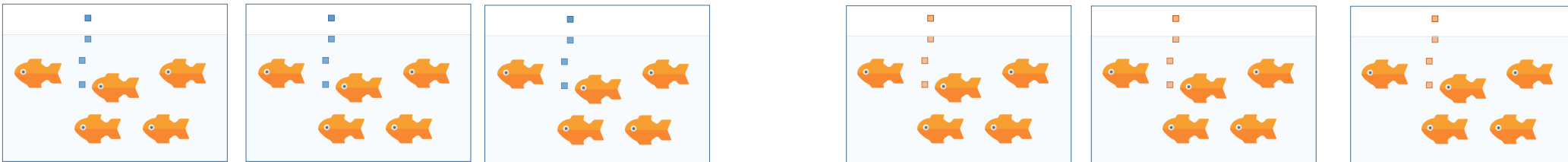
- Để đáp ứng nguyên tắc replication, mỗi loại thức ăn (treatment) sẽ được thử với nhiều con cá (experiment unit).

Replication

- Thí nghiệm 2 loại thức ăn cho cá. Có 2 bể cá, mỗi bể được thử một loại thức ăn. Trọng lượng cá trung bình mỗi bể được đo để so sánh.



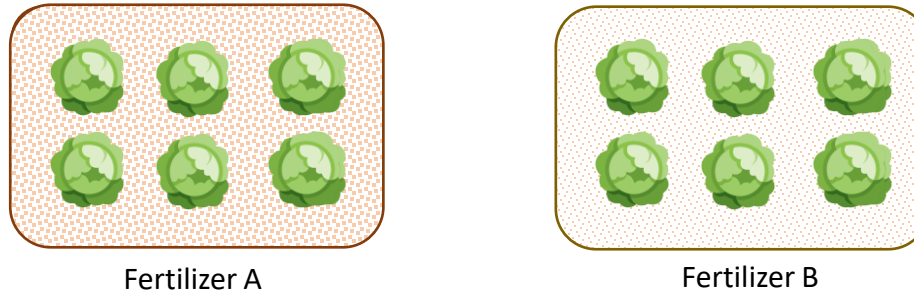
- Vấn đề: mỗi bể có thể có môi trường khác nhau nên cá ở hai bể có thể phát triển khác nhau. Do đó không thể kết luận về tác dụng của thức ăn
- Replication principle:** the experiment should be repeated more than once. Thus, each treatment is applied in **many experimental units** instead of one. By doing so the statistical accuracy of the experiments is increased.



- Để đáp ứng nguyên tắc replication, mỗi loại thức ăn (treatment) sẽ được thử với nhiều bể (experiment unit).

Replication \neq from subsampling within an experimental unit

- Subsampling within 1 E unit or repeated measurements on the same E unit.
 - Mỗi treatment chỉ được áp dụng cho đúng 1 E unit. Mỗi E unit gồm nhiều M unit.



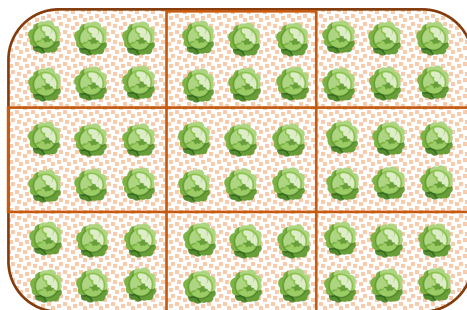
- Ví dụ: so sánh 2 loại phân bón. Có 2 ô ruộng, mỗi ô dùng 1 loại phân bón. Mỗi ô trồng 100 cây bắp cải. Lấy ngẫu nhiên 10 cây từ mỗi ô để so sánh.
 - Mỗi treatment (mỗi loại phân bón) chỉ được áp dụng cho 1 E unit (1 ô ruộng). Mỗi E unit có nhiều M unit (bắp cải). Số mẫu lấy nhỏ hơn số M unit trong mỗi E unit (subsampling).
 - Không thỏa nguyên tắc replication. Tuy lấy nhiều bắp cải để cân nhưng mỗi loại bắp cải tương ứng với 1 loại phân bón, tức là chỉ có 1 ô ruộng cho mỗi phân bón. Chính vậy nên kết quả thí nghiệm không thể khác biệt.

Replication đúng cách

- Mỗi loại phân bón được áp dụng cho nhiều E unit.
 - Cách 1: dùng thêm nhiều ô ruộng. Mỗi ô ruộng (1 E unit) bón 1 loại phân.



- Cách 2: chia mỗi ô ruộng ra nhiều phần. Mỗi phần (1 E unit) bón ngẫu nhiên 1 loại phân.



Ví dụ về sai nguyên tắc replication

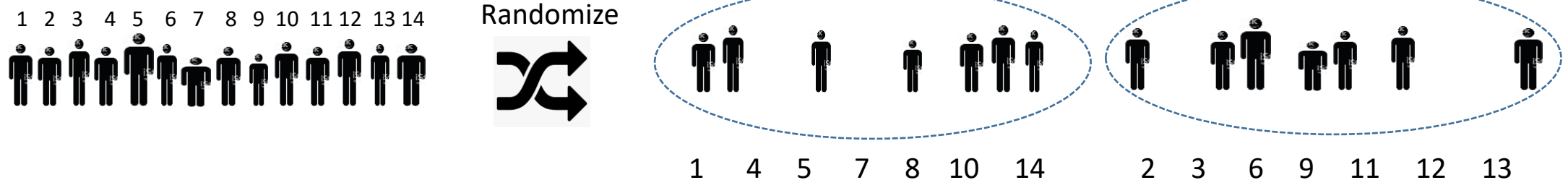
- So sánh 2 dây chuyền sản xuất bia. Sản xuất thử 2 lô bia, mỗi lô dùng 1 dây chuyền. Nhiều lon bia sẽ được đánh giá để so sánh 2 dây chuyền.
 - Mỗi dây chuyền (1 treatment) chỉ được áp dụng cho đúng 1 lô bia (1 E unit). Các lon bia (M unit) ứng với mỗi treatment chỉ thuộc 1 E unit -> sai nguyên tắc
 - Sửa lại: Sản xuất nhiều lô (nhiều E unit), mỗi lô dùng 1 trong 2 dây chuyền.
- So sánh độ chính xác 2 thuật toán định vị. Cho chạy thử tại 2 địa điểm. Mỗi lần chạy trong 1 tiếng. Đo độ chính xác tại 1000 thời điểm ngẫu nhiên trong thời gian này để so sánh kết quả.
 - Mỗi thuật toán (1 treatment) chỉ được áp dụng tại đúng 1 địa điểm (1 E unit). 1000 thời điểm đo (1000 M unit) chỉ thuộc 1 E unit -> sai nguyên tắc
 - Sửa lại: Đo tại nhiều địa điểm (nhiều E unit).
- So sánh 2 giáo trình. Thử trên 2 lớp, mỗi lớp (50 hs) dùng 1 giáo trình. Điểm trung bình của mỗi lớp được tính để so sánh.
 - Mỗi giáo trình (1 treatment) chỉ được áp dụng cho đúng 1 lớp (1 E unit) -> sai nguyên tắc
 - Sửa lại: Mỗi giáo trình cho dạy thử tại nhiều lớp (nhiều E unit) với giáo viên dạy khác nhau.

Randomization

- So sánh tác dụng thuốc tăng lực. 1 nhóm thanh niên được chia làm 2 nhóm, 1 nhóm thử thuốc thật, 1 nhóm thử thuốc giả.



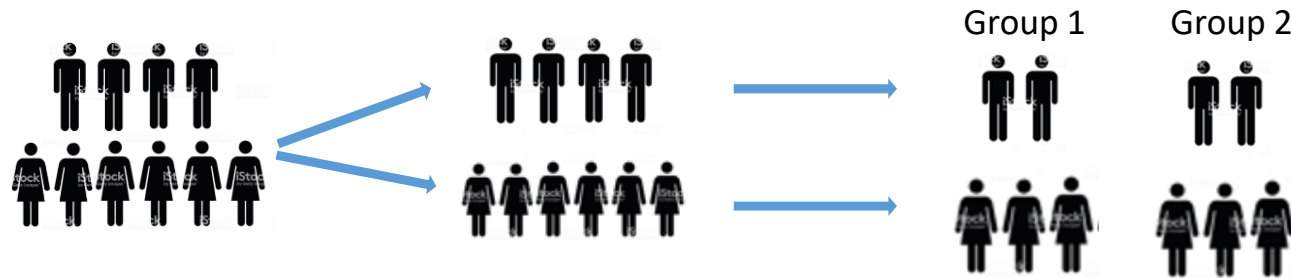
- Vấn đề: chọn đối tượng cho 2 nhóm thế nào để tránh thiên vị (bias)
- **Nguyên tắc randomization:** treatment cần được áp dụng ngẫu nhiên cho các Experiment unit



- Đánh số cho mỗi ứng viên. Dùng máy tính chọn ngẫu nhiên 7 số cho mỗi nhóm. Lấy các ứng viên tương ứng với số.

Blocking

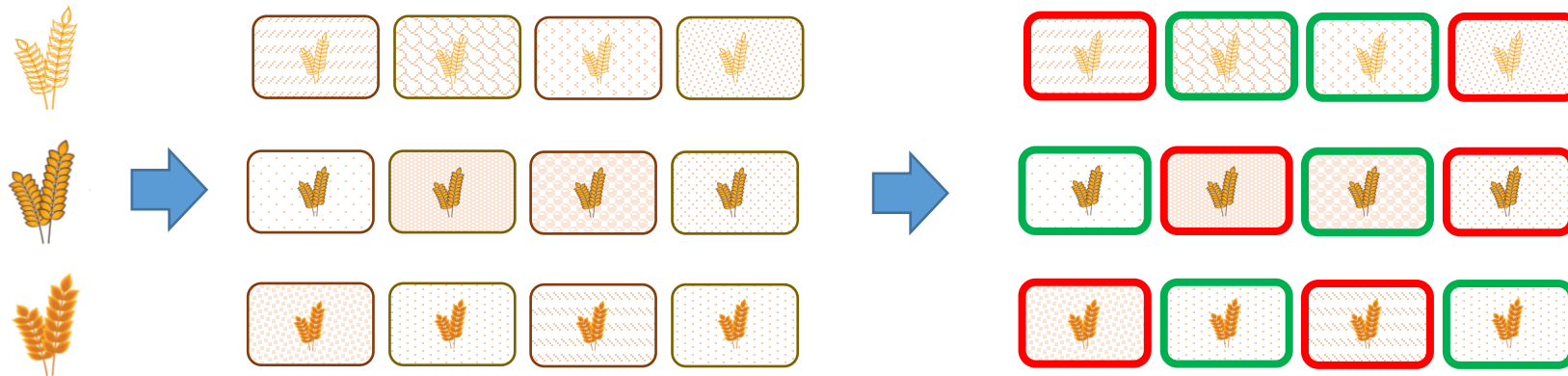
- Thí nghiệm thuốc tăng lực. Đối tượng là 1 nhóm nam nữ lẫn lộn. Chia ngẫu nhiên thành 2 nhóm để thử thuốc thật và thuốc giả để so sánh.
- Thường nam chạy nhanh hơn nữ. Kết quả có thể bị ảnh hưởng nếu 1 nhóm có nhiều nam hơn đáng kể.



- Sửa lại thiết kế: tách nam và nữ thành 2 nhóm (block), sau đấy chọn ngẫu nhiên những thành viên nam và nữ từ 2 block này vào 2 group thử thuốc.
- Nguyên tắc blocking: tách các E unit thành các khối (block) dựa theo yếu tố (factor) có khả năng ảnh hưởng đến kết quả. Sau đó mới áp dụng ngẫu nhiên các treatment lên các E unit trong mỗi khối.

Ví dụ về blocking

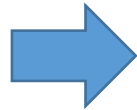
- So sánh 2 loại phân bón. Áp dụng thử lên 3 giống lúa. Trồng thử lên 12 ô ruộng. Mỗi ô trồng 1 loại lúa và dùng 1 loại phân bón.
- Phân bón là yếu tố chính. Giống lúa là yếu tố blocking. Mỗi ô ruộng là 1 E unit.
- Chia 12 ô ruộng thành 3 block dựa theo 3 giống lúa. Mỗi giống lúa trồng trên 4 ô của 1 block. Mỗi ô này sẽ được áp dụng ngẫu nhiên 1 trong 2 loại phân.



Matched pair (trường hợp đặc biệt của blocking)

- Là trường hợp đặc biệt của blocking. Được sử dụng khi thí nghiệm có 2 mức treatment và các E unit có thể được nhóm theo cặp dựa theo yếu tố khối. Sau đó với mỗi cặp, các E unit được áp dụng ngẫu nhiên 1 loại treatment.
- Ví dụ: So sánh 2 pp dạy A và B. Học lực của hs là yếu tố tạo khối vì nó ảnh hưởng đến kết quả. Các học sinh được nhóm theo cặp gần nhau nhất về điểm số hiện tại. Sau đấy mỗi học sinh được áp dụng ngẫu nhiên 1 pp dạy.

100	98
97	95
89	86
80	77
75	74
70	68
67	64
61	60



100	98
97	95
89	86
80	77
75	74
70	68
67	64
61	60

• Mỗi cặp gồm 2 E unit có các đặc tính (điểm số học lực) gần nhau nhất

Ví dụ thêm về matched pair

- Thử nghiệm 2 loại xăng. Có 10 chiếc xe. Do mỗi chiếc xe có hoạt động không hoàn toàn giống nhau nên có thể ảnh hưởng đến kết quả.
 - Xe là yếu tố blocking. Áp dụng thiết kế matched pair, mỗi xe được thử 2 lần với 2 loại xăng.
 - Mỗi lần thử là 1 E unit. Như vậy, mỗi xe có thể tạo thành 1 cặp E unit giống hệt nhau.
- Thử nghiệm 2 loại thuốc. Có 100 bệnh nhân ở các độ tuổi khác nhau. Độ tuổi là yếu tố blocking.
 - Chia 100 bệnh nhân thành 50 cặp có độ tuổi gần nhau nhất. Áp dụng ngẫu nhiên 2 loại thuốc cho mỗi cặp.
 - Mỗi bệnh nhân là 1 E unit. Mỗi cặp bệnh nhân là 1 cặp E unit có đặc điểm gần nhau nhất.