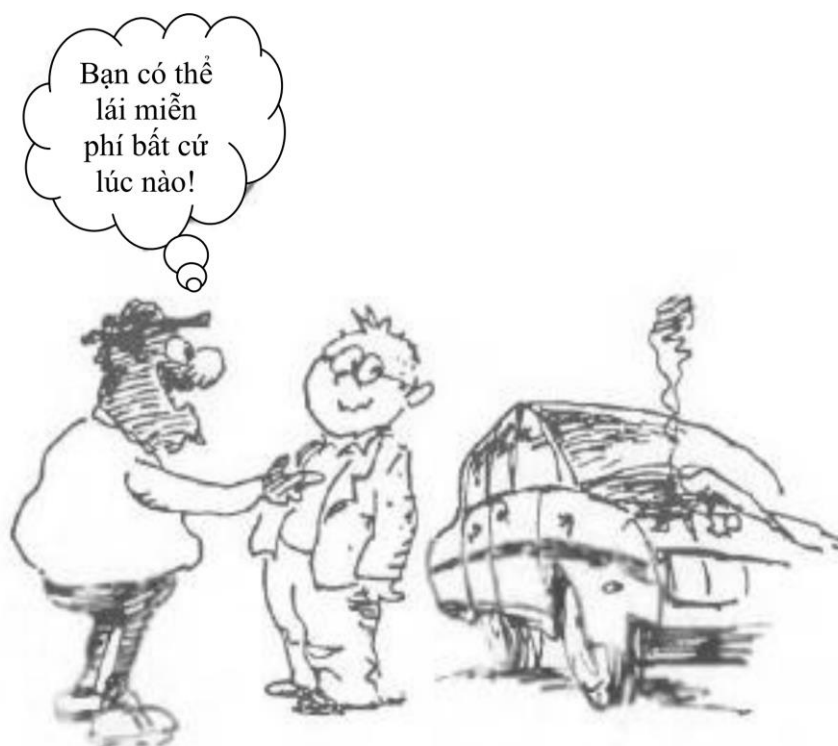


CHƯƠNG 10: THIẾT KẾ ĐIỀU TRA

(Tiếp theo)

Vai trò thiết kế trong một cuộc điều tra thường báo hiệu dấu hiệu thành công hay thất bại của cuộc điều tra đó. Trong ví dụ so sánh cặp, các nhà thống kê đã thay đổi quy tắc từ phân tích nhóm cặp thụ động sang các hoạt động thiết kế điều tra một cách chủ động.

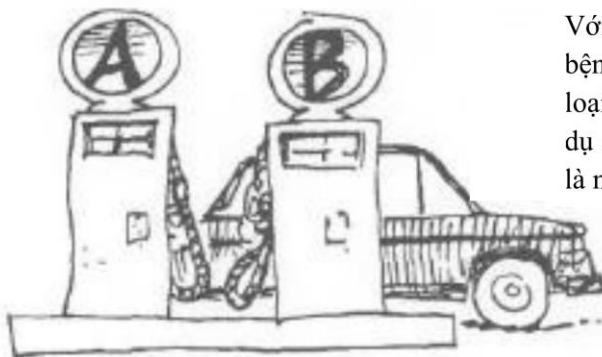
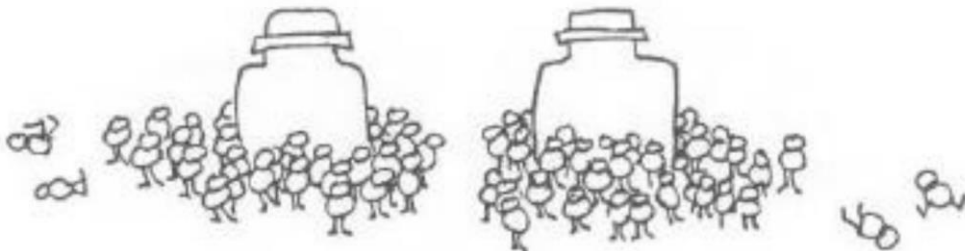


Trong chương này, chúng tôi sẽ giới thiệu các ý tưởng cơ bản của việc thiết kế thực nghiệm, cho phép chúng ta phân tích các số liệu một cách chi tiết thông qua phần mềm thống kê tiện ích.



Không có công thức nào trong chương này...
Rất tiếc!

Các yếu tố cần thiết của một thiết kế chính là các đơn vị thực nghiệm và các phép thử được gán cho từng đơn vị. Đối tượng thiết kế được dùng để so sánh với các phép thử.



Với những phép thử về thuốc, các bệnh nhân chính là các đơn vị còn loại thuốc chính là phép thử. Trong ví dụ đo tốc độ, các đơn vị thực nghiệm là những chiếc xe taxi, và các phép thử được so sánh với nhau là hai loại nhiên liệu A và B.

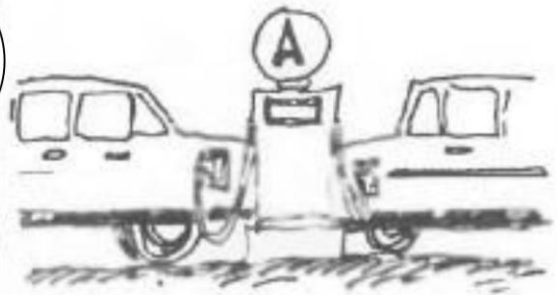
Ở các thí nghiệm nông nghiệp, đơn vị thực nghiệm thường được đánh dấu trong cùng một lĩnh vực, và các phép thử có thể áp dụng cho các loại lúa mỳ, thuốc trừ sâu, phân bón, vv... khác nhau.

Ngày nay, các ý tưởng thiết kế thực nghiệm được sử dụng một cách rộng rãi trong trong quá trình công nghiệp hóa, trong lĩnh vực y học và khoa học xã hội. Một vài thiết kế thực nghiệm sử dụng ba nguyên tắc cơ bản, đã được minh họa rõ trong ví dụ taxi của chúng ta.



Tôi luôn hiểu rằng lái taxi đã từng là một hình thức của khoa học xã hội...

Sự lặp lại: các phép thử tương tự được gán cho các đơn vị thực nghiệm khác nhau. Nếu không lặp lại, thì khó có thể định mức được sự biến thiên tự nhiên và sai số đo lường.



Kiểm soát cục bộ: liên quan tới một vài phương pháp tính toán và giảm thiểu sự biến thiên tự nhiên. Một trong những phương pháp đó là nhóm các đơn vị thực nghiệm thành các cụm. Ở ví dụ taxi, mỗi chiếc taxi sử dụng cả hai loại nhiên liệu và chúng ta gọi taxi chính là một cụm.

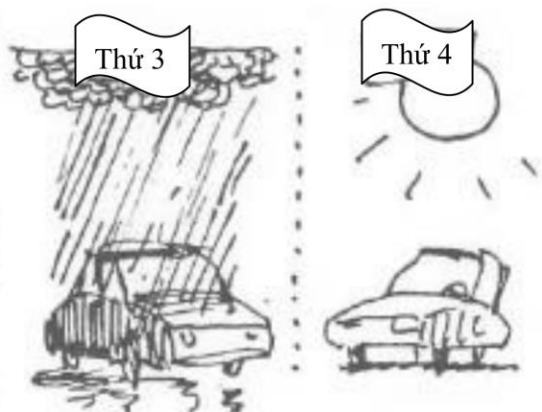


Hãy đưa tôi đi vòng quanh cụm!

Thưa bà, chúng ta đang ở TRONG cụm...



Sự ngẫu nhiên: Điều cốt lõi trong mọi nghiên cứu thống kê! Các phép thử phải được chỉ định ngẫu nhiên vào các đơn vị thực nghiệm. Với mỗi chiếc taxi, chúng ta định dùng loại nhiên liệu A hoặc B vào ngày Thứ 3 hoặc Thứ 4 bằng việc tung đồng xu. Nếu không thực hiện như vậy, các kết quả có thể bị sai do sự khác biệt giữa ngày Thứ 3 và ngày Thứ 4.



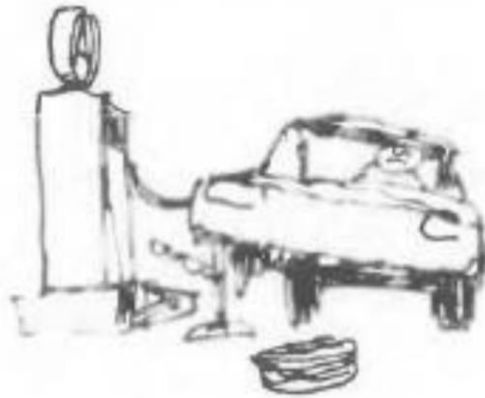
Thứ 3

Thứ 4

➤➤➤ HỌC THỐNG KÊ QUA TRUYỆN TRANH

Cũng như hai loại nhiên liệu bây giờ, giả sử rằng chúng ta muốn tìm ra ảnh hưởng của hai hãng sản xuất lốp xe. Chúng ta có 4 khả năng được sắp xếp theo bảng thiết kế hệ số 2 nhân 2. Hai yếu tố là nhiên liệu và lốp xe.

	Nhiên liệu A	Nhiên liệu B
Lốp A	<i>a</i>	<i>b</i>
Lốp B	<i>c</i>	<i>d</i>



Chúng ta có thể gán ngẫu nhiên 4 phép thử cho 4 ngày với mỗi chiếc taxi. Tất cả 4 phép thử (*a*, *b*, *c* và *d*) được lặp lại ở mỗi cụm (taxi). Điều này được gọi là một thiết kế cụm ngẫu nhiên hoàn toàn.

Giả sử tất cả các ngày trong tuần giống nhau, tuy nhiên thực tế chúng ta cũng có thể kiểm soát điều này, theo cách sau: chỉ sử dụng 4 chiếc taxi, và gán phép thử như bảng bên phải:

		Ngày			
		1	2	3	4
Xe taxi	1	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
	2	<i>b</i>	<i>c</i>	<i>d</i>	<i>a</i>
	3	<i>c</i>	<i>d</i>	<i>a</i>	<i>b</i>
	4	<i>d</i>	<i>a</i>	<i>b</i>	<i>c</i>

Chú ý: mỗi phép thử chỉ xuất hiện một lần duy nhất ở mỗi dòng và mỗi cột!



Một bảng 4×4 với 4 nhân tố khác nhau, mỗi nhân tố xuất hiện một lần ở mỗi cột và mỗi dòng, được gọi là một **ô vuông latin**. Trong thực nghiệm này, bốn ngày và bốn chiếc taxi cho bốn phép thử chính xác.

Hãy thử xem việc tiến hành thống kê bằng các số la mã!



Trong bước ngẫu nhiên, chọn ngẫu nhiên một ô vuông latin từ danh sách toàn bộ các ô vuông latin trong bảng.

Nếu bốn đơn vị vẫn chưa đủ, chúng ta có thể tăng số lượng các đơn vị thực nghiệm lên bằng cách lặp lại thiết kế thực nghiệm. Từ tám chiếc taxi ban đầu, chúng ta có thể chia chúng ra thành hai nhóm, mỗi nhóm gồm 4 chiếc sau đó lặp lại thiết kế với từng nhóm.



Ok... ô tô số 6 dùng loại nhiên liệu **B** và lớp **A** vào ngày Thứ **2**...
: Úi chà :

➤ ➤ ➤ HỌC THỐNG KÊ QUA TRUYỆN TRANH

Nói chung chúng tôi không đi sâu vào việc phân tích dữ liệu một cách chi tiết, nhưng sẽ giúp bạn hiểu được những thiết kế rắc rối như thế này được xử lý ra sao.

Bởi một nhà
thống kê nặng
300 pound!



Các thiết kế thực nghiệm được phân tích nhờ sự phân bố tổng biến thiên giữa các nguồn khác nhau. Ở ví dụ taxi, nguồn của biến là taxi, lớp chế tạo, loại nhiên liệu, ngày - và sai số ngẫu nhiên. Phân tích phương sai, Anova trong ngắn hạn, phân bố tổng biến thiên, sự phân bố tỷ lệ cho mỗi nguồn.

Trong chương tiếp theo, chúng tôi sẽ mô tả chi tiết việc phân tích một thiết kế phức tạp: thông qua một mô hình cụ thể trong hồi quy tuyến tính, bạn có thể tìm hiểu chi tiết về Anova và các con số định lượng...



(Còn nữa)

Biên dịch: Minh Ánh và các nghiên cứu viên, Viện Khoa học Thống kê