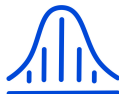


Giới thiệu về Thiết kế thí nghiệm (Design of Experiments)

Hoàng Văn Hà
University of Science, VNU - HCM
hvha@hcmus.edu.vn



VIASM
VIETNAM INSTITUTE FOR
ADVANCED STUDY IN MATHEMATICS

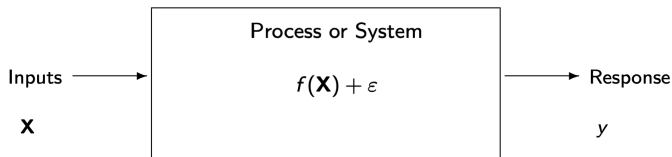


VASN
VIETNAM APPLIED STATISTICS NETWORK

- 1 Giới thiệu về thiết kế thí nghiệm
- 2 Phân tích phương sai một nhân tố (One-way ANOVA)
- 3 Kiểm tra các giả định của mô hình
- 4 Randomized Complete Block Design (RCBD)

Giới thiệu về thiết kế thí nghiệm

- Thiết kế thí nghiệm - **Design of experiment (DOE)** là một phương pháp thống kê có hệ thống và hiệu quả cho phép các nhà khoa học và kỹ sư nghiên cứu được mối quan hệ giữa nhiều biến đầu vào (còn được gọi là các nhân tố - **factors**) và các biến đầu ra quan tâm (biến đáp ứng - **responses**).



- Các câu hỏi quan tâm đối với quy trình trên: sự thay đổi của X ảnh hưởng như thế nào đến trung bình đáp ứng? Các mức nào của biến đầu vào sẽ cực đại hóa trung bình đáp ứng? Phối hợp các biến đầu vào như thế nào để cho kết quả với độ bất định nhỏ nhất?

Giới thiệu về thiết kế thí nghiệm

- Thiết kế thí nghiệm rất hữu ích trong nhiều lĩnh vực khoa học - kỹ thuật để giải quyết nhiều vấn đề quan trọng: khám phá ra những hiện tượng cơ bản mới có thể dẫn đến sự phát triển của những sản phẩm mới, quy trình sản xuất mới, hoặc cải tiến các quy trình và các sản phẩm hiện tại, ...
- Hầu hết các quy trình đều có thể được mô tả dưới dạng một số biến có thể kiểm soát được, như thời gian, nhiệt độ, nồng độ, ... Bằng việc sử dụng các thiết kế thí nghiệm, các kỹ sư có thể xác định tập con nào của các biến quy trình có ảnh hưởng nhiều nhất đến hiệu suất và chất lượng của quy trình. Các kết quả của một thí nghiệm như vậy có thể giúp:
 - cải thiện năng suất của quy trình,
 - giảm sự biến thiên của quy trình và đạt được mục tiêu cần có,
 - giảm thời gian thiết kế và phát triển,
 - giảm chi phí vận hành.

Giới thiệu về thiết kế thí nghiệm

Một số ứng dụng của thiết kế thí nghiệm thống kê trong thiết kế kỹ thuật (**engineering design**) như

- đánh giá và so sánh các cấu hình thiết kế cơ bản,
- đánh giá các vật liệu khác nhau,
- lựa chọn các tham số thiết kế để sản phẩm hoạt động tốt trong nhiều điều kiện khác nhau,
- xác định các tham số thiết kế quan trọng mà ảnh hưởng đến hiệu suất/chất lượng của sản phẩm.

Một số yếu tố quan trọng của thiết kế thí nghiệm

- **Randomization:** đây là yếu tố rất quan trọng của bất kỳ thí nghiệm nào, giúp ta tránh được việc rút ra những kết luận sai lệch.
- **Replication:** giảm độ bất định ([uncertainty](#)) và tăng độ chính xác bằng việc giảm đi sự biến thiên của thí nghiệm.
- **Blocking:** giảm độ bất định bằng việc loại bỏ đi các hiệu ứng của một số nhân tố phiền nhiễu ([nuisance factors](#)).
- **Factorial designs:** việc phân tích đồng thời nhiều nhân tố sẽ hiệu quả hơn một nhân tố và còn cho phép ta khảo sát được tương tác của các nhân tố khác nhau.
- **Confounding:** trong hai yếu tố ảnh hưởng lên một biến đáp ứng, ta không thể cô lập được tác động của một yếu tố này so với yếu tố khác (chẳng hạn xem [Simpson's paradox](#)). Ta có thể tránh được hiện tượng này với một thí nghiệm được thiết kế đúng cách.

Confounding

Ví dụ 1 (Simpson's paradox)

Người ta ghi nhận tỷ lệ phục hồi của 700 bệnh nhân mà đã được cung cấp thuốc điều trị, trong đó có 350 bệnh nhân chọn uống thuốc và 350 bệnh nhân không dùng thuốc. Kết quả nghiên cứu được cho bởi bảng sau:

	Drug	No drug
Men	81 out of 87 recovered (93%)	234 out of 270 recovered (87%)
Women	192 out of 263 recovered (73%)	55 out of 80 recovered (69%)
Combined data	273 out of 350 recovered (78%)	289 out of 350 recovered (83%)

Câu hỏi: bác sĩ có nên chỉ định cho bệnh nhân dùng thuốc hay không?

Các bước lập kế hoạch, thực hiện và phân tích một thí nghiệm

- 1 Nhận diện và phát biểu vấn đề
- 2 Lựa chọn các nhân tố, mức và khoảng giá trị
- 3 Lựa chọn biến đáp ứng
- 4 Xác định thiết kết
- 5 Thực hiện thí nghiệm
- 6 Thực hiện các phân tích thống kê
- 7 Rút ra các quyết định, và khuyến nghị

Phân tích phương sai một nhân tố: ví dụ

Ví dụ 2 (Tensile strength/cf. Montgomery, 13-2)

A manufacturer of paper used for making grocery bags is interested in improving the product's tensile strength. Product engineering believes that tensile strength is a function of the hardwood concentration in the pulp and that the range of hardwood concentrations of practical interest is between 5 and 20%. A team of engineers responsible for the study decides to investigate four levels of hardwood concentration: 5%, 10%, 15%, and 20%. They decide to make up six test specimens at each concentration level by using a pilot plant. All 24 specimens are tested on a laboratory tensile tester in random order. The data from this experiment are shown in Table below.

Hardwood Concentration (%)	Observations						Totals	Averages
	1	2	3	4	5	6		
5	7	8	15	11	9	10	60	10.00
10	12	17	13	18	19	15	94	15.67
15	14	18	19	17	16	18	102	17.00
20	19	25	22	23	18	20	<u>127</u>	<u>21.17</u>
							383	15.96

Bảng 1: Tensile Strength of Paper (psi)

Phân tích phương sai một nhân tố

- **Câu hỏi:** có sự khác biệt về độ đàn hồi (psi) giữa các bao bì có hàm lượng gỗ cứng trong bột gỗ ở 4 mức khác nhau hay không?
- Ví dụ trên đặt ra bài toán so sánh sự khác biệt giữa trung bình của nhiều nhóm khác nhau (≥ 3).
- Thiết kế thí nghiệm trong ví dụ trên được gọi là **Completely Randomized Design - CRD**:
 - mỗi đơn vị thí nghiệm (**experiment unit**) có cơ hội như nhau để nhận bất kỳ một trong các xử lý/liệu pháp (**treatments**),
 - việc phân chia các liệu pháp/xử lý cho các đơn vị thí nghiệm được thực hiện một cách độc lập.

Mô hình cho thí nghiệm 1 nhân tố

- Giả sử ta quan tâm đến việc so sánh k mức (**levels**) của một nhân tố (**factor**).
- N đơn vị thí nghiệm được chọn ngẫu nhiên để nhận một trong k liệu pháp (**treatments**), trong đó n_i đơn vị thí nghiệm được chọn để nhận liệu pháp i , $i = 1, \dots, k$.
- Dữ liệu ghi nhận được sẽ được biểu diễn giống như trong bảng 1, mỗi giá trị trong bảng 1, ký hiệu là y_{ij} , gọi là quan trắc (**observation**) thứ j được xử lý theo liệu pháp i .
- Dữ liệu trong bảng 1 có thể được biểu diễn theo mô hình tuyến tính sau

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad (1)$$

với $i = 1, 2, \dots, k$ và $j = 1, 2, \dots, n_i$. Trong đó:

- y_{ij} là quan trắc thứ j trong mẫu thứ i , μ là trung bình chung (**overall mean**),
- τ_i là hiệu ứng của liệu pháp thứ i ,
- ε_{ij} là thành phần sai số ngẫu nhiên.

Mô hình cho thí nghiệm 1 nhân tố

- Ta có thể biểu diễn mô hình (1) như sau

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad (2)$$

với $\mu_i = \mu + \tau_i$ là trung bình của liệu pháp thứ i .

- Các giả định về sai số của mô hình: sai số được giả sử là độc lập, tuân theo phân phối chuẩn với trung bình 0 và phương sai là hằng số σ^2 .
- Đối với mô hình với **các hiệu ứng cố định (fixed-effects model)**, ta giả sử tổng các hiệu ứng của các liệu pháp bằng 0, tức là

$$\sum_{i=1}^k \tau_i = 0. \quad (3)$$

Mô hình cho thí nghiệm 1 nhân tố

Dữ liệu cho mô hình (2) có thể được mô tả như bảng sau:

Treatment	Observations				Totals	Averages
1	y_{11}	y_{12}	\cdots	y_{1n_1}	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
2	y_{21}	y_{22}	\cdots	y_{2n_2}	$y_{2\cdot}$	$\bar{y}_{2\cdot}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
k	y_{k1}	y_{k2}	\cdots	y_{kn_k}	$y_{k\cdot}$	$\bar{y}_{k\cdot}$
					$y_{\cdot\cdot}$	$\bar{y}_{\cdot\cdot}$

Trong đó,

$$y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}, \quad \bar{y}_{i\cdot} = y_{i\cdot}/n_i, \quad i = 1, 2, \dots, k,$$

$$y_{\cdot\cdot} = \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}, \quad \bar{y}_{\cdot\cdot} = y_{\cdot\cdot}/N, \quad N = \sum_{i=1}^k n_i.$$

ANOVA 1 nhân tố: giả thuyết và đối thuyết

- Giả thuyết:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k,$$

trung bình của tất cả các liệu pháp xử lý bằng nhau, hay nói cách khác, không có sự khác biệt về trung bình giữa các nhóm.

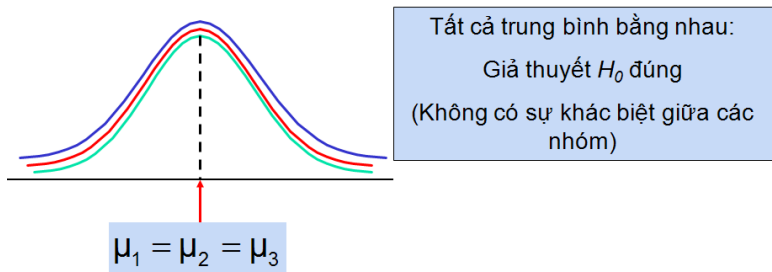
- Đối thuyết:

$$H_1 : \mu_i \neq \mu_j \quad \text{với ít nhất một cặp } 1 \leq i, j \leq k,$$

nghĩa là có tồn tại sự khác biệt giữa các nhóm.

ANOVA 1 nhân tố: giả thuyết và đối thuyết

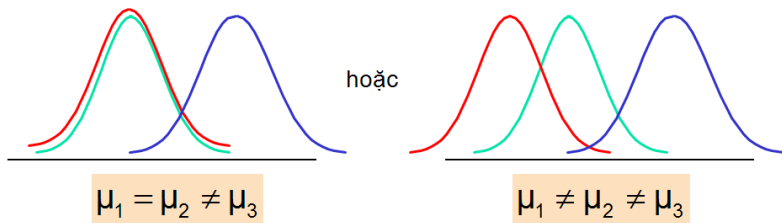
$$\begin{cases} H_0 : \mu_1 = \mu_2 = \dots = \mu_k, \\ H_1 : \mu_i \neq \mu_j \text{ với ít nhất một cặp } 1 \leq i, j \leq k \end{cases}$$



Giả thuyết và đối thuyết

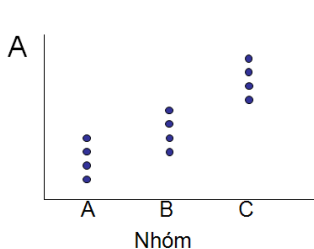
$$\begin{cases} H_0 : \mu_1 = \mu_2 = \dots = \mu_k, \\ H_1 : \mu_i \neq \mu_j \text{ với ít nhất một cặp } 1 \leq i, j \leq k \end{cases}$$

Giả thuyết H_0 sai
(Có sự khác biệt giữa các nhóm)

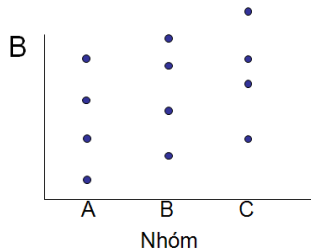


Phân tích sự biến thiên của dữ liệu

- Trong ANOVA, sự biến thiên trong dữ liệu là chìa khóa để kiểm tra sự bằng nhau của trung bình giữa các nhóm.



Biến thiên nhỏ trong các nhóm



Biến thiên lớn trong các nhóm

Tổng các bình phương (Sum of Squares)

Sự biến thiên toàn phần trong dữ liệu có thể phân chia thành hai thành phần như sau

$$SST = SSB + SSE, \quad (4)$$

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2. \quad (5)$$

Trong đó

- SST = Tổng bình phương toàn phần (Total Sum of Squares).
- SSB = Tổng bình phương giữa các nhóm (Sum of Squares Between Groups)
- SSE = Tổng bình phương sai số (Error Sum of Squares). Kí hiệu khác: SSW = Tổng bình phương bên trong các các nhóm (Within groups Sum of Squares).

Trường hợp các nhóm có cỡ mẫu bằng nhau, $n_1 = n_2 = \dots = n_k = n$:

$$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2. \quad (6)$$

Trung bình bình phương (Mean squares)

Trung bình bình phương (**mean squares**) là tổng các bình phương (**sum of squares**) chia cho bậc tự do tương ứng:

- Trung bình bình phương toàn phần

$$MST = \frac{SST}{N - 1}, \quad (7)$$

với $N = \sum_{i=1}^k n_i$ ($N = kn$ nếu cỡ mẫu của các nhóm bằng nhau).

- Trung bình bình phương giữa các nhóm

$$MSB = \frac{SSB}{k - 1}. \quad (8)$$

- Trung bình bình phương sai số

$$MSE = \frac{SSE}{N - k}. \quad (9)$$

Ước lượng của phương sai - MSE

- Khi ta chỉ có 1 mẫu y_1, y_2, \dots, y_n , phương sai tổng thể σ^2 được ước lượng bởi phương sai mẫu

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}.$$

- Đối với mô hình $y_{ij} = \mu_i + \varepsilon_{ij}$, vì tất cả các nhóm $i = 1, \dots, k$ đều có phương sai bằng nhau và bằng σ^2 , nên phương sai mẫu s_i^2 của bất kỳ nhóm thứ i nào đều có thể ước lượng phương sai chung σ^2 , với

$$s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2}{n_i - 1}, \quad i = 1, 2, \dots, k.$$

Ước lượng của phương sai - MSE

- Ta có thể gộp tất cả các $s_1^2, s_2^2, \dots, s_k^2$ để thu được một ước lượng tốt hơn cho σ^2 , cho bởi:

$$\begin{aligned}\hat{\sigma}^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2}{(n_1 - 1) + (n_2 - 1) + \dots + (n_k - 1)} \\ &= \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})}{N - k} = \frac{SSE}{N - k} = MSE.\end{aligned}$$

- Ta cũng dễ dàng kiểm tra được MSE là một ước lượng không chệch cho σ^2 , vì trong mỗi nhóm, s_i^2 là ước lượng không chệch cho σ^2 , ta có

$$\begin{aligned}\mathbb{E}(SSE) &= \mathbb{E} \left(\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 \right) = \mathbb{E} \left(\sum_{i=1}^k (n_i - 1)s_i^2 \right) \\ &= \sum_{i=1}^k (n_i - 1)\sigma^2 = (N - k)\sigma^2.\end{aligned}$$

Do vậy, $\mathbb{E}(MSE) = \mathbb{E}(SSE/(N - k)) = \sigma^2$.

Kiểm định ANOVA 1 nhân tố

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \dots = \mu_k, \\ H_1 : \mu_i \neq \mu_j \quad \text{với ít nhất một cặp } 1 \leq i, j \leq k \end{cases}$$

- Dưới $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$, ta mong rằng

$$\bar{y}_{1.} \approx \bar{y}_{2.} \approx \dots \approx \bar{y}_{k.},$$

tức là $\bar{y}_{i.} \approx \bar{y}_{.}$ với mọi liệu pháp i . Do đó, một giá trị lớn của $SSB = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_{i.} - \bar{y}_{.})^2$ là bằng chứng để có thể dẫn đến kết luận bác bỏ H_0 .

- Mặt khác, nếu σ^2 càng lớn \Rightarrow sự biến thiên lớn trong các nhóm ngay cả khi H_0 đúng. Do vậy, ta nên xem xét SSB/σ^2 thay vì SSB .
- σ^2 được ước lượng bởi MSE .

Thống kê kiểm định F

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \dots = \mu_k, \\ H_1 : \mu_i \neq \mu_j \quad \text{với ít nhất một cặp } 1 \leq i, j \leq k \end{cases}$$

- Thống kê kiểm định được đề xuất là

$$F_0 = \frac{MSB}{MSW} = \frac{SSB/(k-1)}{SSE/(N-k)} = \frac{\text{Sự biến thiên giữa các nhóm}}{\text{Sự biến thiên bên trong các nhóm}}. \quad (10)$$

Thống kê F_0 là tỷ số giữa phương sai ước lượng giữa các nhóm với phương sai ước lượng trong từng nhóm.

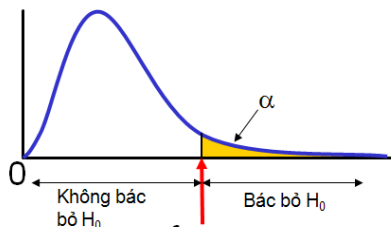
- Thống kê F_0 có phân phối F với hai bậc tự do $df_1 = k - 1$ và $df_2 = N - k$.

Bảng ANOVA một nhân tố

Nguồn của sự biến thiên	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Giữa các nhóm	<i>SSB</i>	$k - 1$	<i>MSB</i>	$F = \frac{MSB}{MSE}$
Trong từng nhóm	<i>SSE</i>	$N - k$	<i>MSE</i>	
Tổng	<i>SST</i>	$N - 1$		

- Bác bỏ H_0 khi:

$$F_0 > F_{\alpha; k-1, N-k} \quad (11)$$



ANOVA 1 nhân tố: ví dụ

VÍ DỤ 1. (tiếp theo)

Với $\alpha = 0.01$, ta kiểm định giả thuyết

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ H_1 : \mu_i \neq \mu_j \quad \text{với ít nhất một cặp } 1 \leq i, j \leq 4 \end{cases}$$

Tính các tổng bình phương liên quan, từ bảng 1 ta có

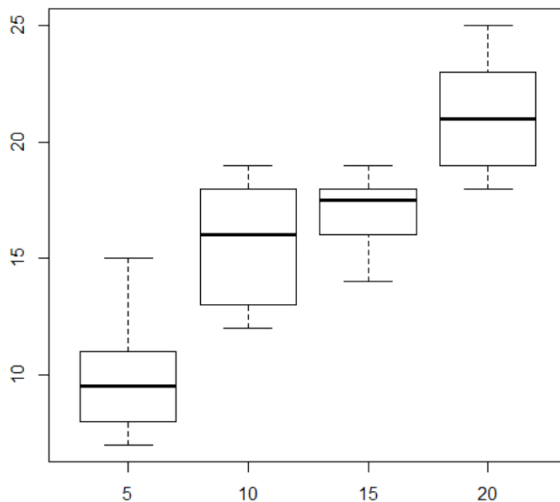
$$SST = \sum_{i=1}^4 \sum_{j=1}^6 y_{ij}^2 - \frac{y_{..}^2}{N} = (7^2 + 8^2 + \dots + 20^2) - \frac{383^2}{24} = 512.96,$$

$$SSB = \sum_{i=1}^4 \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N} = \frac{60^2 + 94^2 + 102^2 + 127^2}{6} - \frac{383^2}{24} = 382.79,$$

$$SSE = SST - SSB = 512.96 - 382.79 = 130.17.$$

ANOVA 1 nhân tố: ví dụ

Đồ thị boxplot:



ANOVA 1 nhân tố: ví dụ

- Tính giá trị thống kê kiểm định:

$$F_0 = \frac{MSB}{MSE} = \frac{SSB/(k-1)}{SSE/(N-k)} = \frac{382.79/3}{130.17/20} = 19.60.$$

Vì $F_{0.01;3,20} = 4.94$, nên ta bác bỏ H_0 và kết luận rằng hàm lượng gỗ cứng trong nguyên liệu có ảnh hưởng đến độ đàn hồi của các bao bì giấy.

- Kết luận dùng p -giá trị:

$$p = \mathbb{P}(F_{3,20} > 19.60) = 3.59 \times 10^{-6},$$

vì $p = 3.59 \times 10^{-6} \ll 0.01$, nên ta có đủ bằng chứng mạnh để bác bỏ H_0 .

So sánh bội sau ANOVA

- Khi giả thuyết $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ được bác bỏ trong ANOVA, ta biết có sự khác biệt giữa các nhóm (các liệu pháp khác nhau), nhưng chưa biết được cụ thể các nhóm nào gây ra sự khác biệt.
- Để xác định nhóm nào gây ra sự khác biệt, ta sử dụng các phương pháp so sánh bội (Multiple comparison methods). Một phương pháp so sánh bội đơn giản là phương pháp Least significant difference (LSD) của Fisher.
- Nội dung của phương pháp LSD là so sánh tất cả các cặp trung bình tương ứng với giả thuyết $H_0 : \mu_i = \mu_j$ (với mọi $i \neq j$), sử dụng thống kê t cho bởi

$$T = \frac{\bar{y}_{i\cdot} - \bar{y}_{j\cdot}}{\sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}. \quad (12)$$

So sánh bội sau ANOVA

- Giả sử với đối thuyết hai phía, cặp trung bình μ_i và μ_j gọi là khác nhau có ý nghĩa nếu

$$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD$$

với

$$LSD = t_{\alpha/2}^{N-k} \sqrt{MSW \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \quad (13)$$

trong đó N là tổng số đơn vị thí nghiệm, k là số liệu pháp, n_i và n_j lần lượt là số đơn vị thí nghiệm tương ứng với liệu pháp thứ i và j .

So sánh bội sau ANOVA: ví dụ

Thực hiện so sánh bội với kết quả phân tích ANOVA trong ví dụ 1, ta có số nhóm $k = 4$, $n = 6$, $MSE = 6.51$ và $t_{0.975}^{20} = 2.086$. Trung bình mẫu tương ứng với 4 nhóm được tính như sau:

$$\bar{y}_{1.} = 10.00, \quad \bar{y}_{2.} = 15.67, \quad \bar{y}_{3.} = 17.00, \quad \bar{y}_{4.} = 21.17.$$

Giá trị $LSD = t_{0.975}^{20} \sqrt{2MSE/n} = 2.086 \sqrt{2 \times 6.51/6} = 3.07$. So sánh các cặp trung bình như sau:

$$4 \text{ vs. } 1 = 21.17 - 10.00 = 11.17 > 3.07,$$

$$4 \text{ vs. } 2 = 21.17 - 15.67 = 5.50 > 3.07,$$

$$4 \text{ vs. } 3 = 21.17 - 17.00 = 4.17 > 3.07,$$

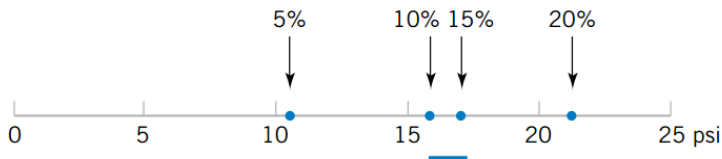
$$3 \text{ vs. } 1 = 17.00 - 10.00 = 7.00 > 3.07,$$

$$3 \text{ vs. } 2 = 17.00 - 15.67 = 1.33 < 3.07,$$

$$2 \text{ vs. } 1 = 15.67 - 10.00 = 5.67 > 3.07.$$

So sánh bội sau ANOVA: ví dụ

Từ phân tích trên, ta nhận thấy rằng ngoại trừ nhóm 2 và 3, trung bình của tất cả các nhóm còn lại đều khác nhau có ý nghĩa. Điều này chứng tỏ rằng hàm lượng gỗ cứng trong bột gỗ ở mức 10% và 15% đều cho ra những sản phẩm có độ đàn hồi xấp xỉ bằng nhau. Kết quả có thể được mô tả bởi hình vẽ sau:



Kiểm tra các giả định của mô hình

- Nhắc lại, các giả định về sai số của mô hình:
 - ❶ Các sai số độc lập với nhau
 - ❷ Sai số tuân theo phân phối chuẩn
 - ❸ Sai số có phương sai không đổi
- Các kỹ thuật kiểm tra thường sử dụng các giá trị dự đoán (**predicted values**) \hat{y}_{ij} và các thặng dư (**residuals**) $\hat{\epsilon}_{ij}$ (là các ước lượng của thành phần sai số không quan sát được).

Phân tích đồ thị thặng dư

❶ Kiểm tra phân phối chuẩn:

- Vẽ đồ thị histogram của các thặng dư.
- Vẽ đồ thị QQ-norm của các thặng dư.
- Dùng kiểm định Shapiro-Wilks hay Kolmogorov - Smirnov đối với các thặng dư.

❷ Kiểm tra phương sai là hằng số:

- Vẽ đồ thị thặng dư (residual plot): vẽ các $\hat{\epsilon}_{ij}$ tương ứng với các \hat{y}_{ij} ,
- Sử dụng kiểm định Levene.

❸ Kiểm tra tính độc lập:

- Vẽ các thặng dư $\hat{\epsilon}_{ij}$ theo các biến thời gian/không gian.
- Sử dụng kiểm định Durbin-Watson.
- Vẽ các thặng dư $\hat{\epsilon}_{ij}$ với các biến khác được quan tâm và đánh giá dạng (pattern) của đồ thị.

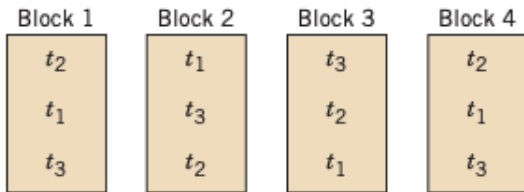
❹ Ngoài ra, việc phân tích thặng dư còn có thể giúp ta phát hiện ra các điểm **outliers** hay các quan trắc bất thường. Các điểm này thường cho kết quả với giá trị $|\hat{\epsilon}_{ij}|$ lớn.

Thiết kế khối (Block Designs)

- Một nhân tố phiền nhiễu (**nuisance factor**) là một nhân tố có ảnh hưởng đến biến đáp ứng y nhưng không phải là đối tượng quan tâm của người làm thí nghiệm.
- Khi thiết kế một thí nghiệm, ta cần phải xem xét đến các khả năng có thể có sự xuất hiện của các nhân tố phiền nhiễu:
 - Nếu nhân tố phiền nhiễu không biết, việc ngẫu nhiên hóa giúp tránh được việc rút ra kết luận chệch nhưng ảnh hưởng của nó có thể làm tăng phương sai.
 - Nếu ta xác định được nhân tố phiền nhiễu nhưng không thể kiểm soát nó, ta có thể sử dụng phân tích hiệp phương sai **Analysis of Covariance (ANCOVA)** để đo lường và loại bỏ hiệu ứng của các nhân tố này.
 - Nếu ta xác định được nhân tố phiền nhiễu và có thể kiểm soát được, ta sử dụng thiết kế khối (**block designs**).

Thiết kế khối (Block Designs)

- Một khối (**block**) là tập hợp các đơn vị thí nghiệm mà đồng nhất theo một nghĩa nào đó. Các đơn vị thí nghiệm trong cùng 1 khối sẽ có các giá trị đáp ứng tương tự nhau khi được áp dụng với cùng một liệu pháp.
- **Thiết kế khối (block designs)**: ngẫu nhiên hóa các đơn vị trong mỗi khối với các liệu pháp.



Thiết kế khối: ví dụ

Ví dụ 3 (Air Freshener Sale/cf. Dean & Voss, p. 427)

- ▶ Goal: comparing 4 price+display treatments on the sales of a brand of air fresheners.
 - ▶ Treatment A = high price + extra display
 - ▶ Treatment B = middle price + extra display
 - ▶ Treatment C = low price + extra display
 - ▶ Treatment D = middle price + no extra display
- ▶ conducted at 8 stores for 4 weeks and each treatment lasts for one week in each store
- ▶ Response: unit sales in a one-week period

Week	Store							
	1	2	3	4	5	6	7	8
1	B 31	A 23	C 12	D 3	A 10	C 30	B 23	D 14
2	A 19	D 16	B 14	C 4	B 21	D 25	C 17	A 14
3	D 15	C 30	A 12	B 6	C 12	A 47	D 5	B 3
4	C 16	B 27	D 5	A 11	D 12	B 38	A 13	C 6

Randomized Complete Block Design

- Xét b khối, mỗi khối có a đơn vị thí nghiệm.
- Gán ngẫu nhiên a liệu pháp vào các đơn vị thí nghiệm trong mỗi khối.
- Bảng số liệu cho RCBD được biểu diễn như sau:

Treatments	Blocks				Totals	Averages
	1	2	...	b		
1	y_{11}	y_{12}	...	y_{1b}	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
2	y_{21}	y_{22}	...	y_{2b}	$y_{2\cdot}$	$\bar{y}_{2\cdot}$
\vdots	\vdots	\vdots		\vdots	\vdots	\vdots
a	y_{a1}	y_{a2}	...	y_{ab}	$\bar{y}_{a\cdot}$	$\bar{y}_{a\cdot}$
Totals	$y_{\cdot 1}$	$y_{\cdot 2}$...	$y_{\cdot b}$	$\bar{y}_{\cdot\cdot}$	
Averages	$\bar{y}_{\cdot 1}$	$\bar{y}_{\cdot 2}$...	$\bar{y}_{\cdot b}$		$\bar{y}_{\cdot\cdot}$

- Ta xác định các biến có thể xem xét để tạo khối (**blocking**) bằng cách xác định nguồn của sự biến thiên (ví dụ như tuổi, giới tính, lịch sử khám bệnh, vị trí địa lý, ...), hoặc tạo khối theo không gian, thời gian, máy, ...

Mô hình cho RCBD

- Sử dụng mô hình tuyến tính

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \quad \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \end{cases} \quad (14)$$

trong đó

- y_{ij} : đáp ứng của đơn vị thí nghiệm nhận được liệu pháp i trong khối j
- μ : trung bình chung
- α_i : hiệu ứng của liệu pháp i
- β_j : hiệu ứng của khối j
- ε_{ij} : sai số, $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$
- Ràng buộc của mô hình

$$\sum_{i=1}^a \alpha_i = 0 \quad \text{và} \quad \sum_{j=1}^b \beta_j = 0. \quad (15)$$

Tổng các bình phương

- Quan trắc y_{ij} có thể được biểu diễn như sau:

$$\begin{aligned} y_{ij} &= \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) \\ &= \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\epsilon}_{ij}. \end{aligned}$$

- Tương tự như trong thiết kế CRD, tổng bình phương toàn phần $SST = \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{..})^2$ được phân tích như sau:

$$SST = SSG + SSB + SSE,$$

với

- $SSG = b \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2$: Sum of Squares between Groups.
- $SSB = a \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{..})^2$: Sum of Squares between Blocks.
- $SSE = \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$: Sum of Squares Error.
- Bậc tự do tương ứng:

$$ab - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1).$$

Trung bình bình phương (Mean squares)

- Các trung bình bình phương trong thiết kế CRBD cho bởi:

$$MSG = \frac{SSG}{a-1}, \quad MSB = \frac{SSB}{b-1}, \quad MSE = \frac{SSE}{(a-1)(b-1)}.$$

- Theo mô hình cho RCBD, $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$ với $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ và các ràng buộc $\sum_{i=1}^a \alpha_i = 0$ và $\sum_{j=1}^b \beta_j = 0$, ta có thể chứng tỏ rằng

$$\mathbb{E}(MSG) = \sigma^2 + \frac{b}{a-1} \sum_{i=1}^a \alpha_i^2,$$

$$\mathbb{E}(MSB) = \sigma^2 + \frac{a}{b-1} \sum_{j=1}^b \beta_j^2,$$

$$\mathbb{E}(MSE) = \sigma^2.$$

Ta thấy rằng MSE vẫn là một ước lượng không chệch cho σ^2 .

Kiểm định ANOVA cho hiệu ứng của liệu pháp

- Đối với liệu pháp (các nhóm):

$$\begin{cases} H_{0a} : \alpha_1 = \alpha_2 = \dots = \alpha_a = 0 \\ H_{1a} : \text{không phải tất cả } \alpha_i \text{ bằng } 0. \end{cases} \quad (16)$$

- Dưới H_{0a} , thống kê kiểm định:

$$F_{\text{trt}} = \frac{MSG}{MSE} \sim F_{a-b, (a-1)(b-1)}. \quad (17)$$

- Bác bỏ H_{0a} khi:

$$F_{\text{trt}} > F_{\alpha; a-1, (a-1)(b-1)}.$$

Bảng ANOVA 2 cho RCBD

Nguồn của sự biến thiên	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Giữa các nhóm	<i>SSG</i>	$a - 1$	<i>MSG</i>	$F_{\text{trt}} = \frac{MSG}{MSE}$
Giữa các khối	<i>SSB</i>	$b - 1$	<i>MSB</i>	$F_{\text{block}} = \frac{MSB}{MSE}$
Sai số	<i>SSE</i>	$(a - 1)(b - 1)$	<i>MSE</i>	
Tổng	<i>SST</i>	$ab - 1$		

Thống kê F_{block} được dùng để kiểm định cho hiệu ứng của các khối nhưng đây không phải là điều được quan tâm nên ta thường bỏ qua.

Tài liệu tham khảo

- ① Montgomery, Douglas C., and George C. Runger. *Applied statistics and probability for engineers*. John wiley & sons, 2010.
- ② Dean, Angela, and Daniel Voss, eds. *Design and analysis of experiments*. New York, NY: Springer New York, 1999.
- ③ Montgomery, Douglas C. *Design and analysis of experiments*. John wiley & sons, 2017.
- ④ Lawson, John. *Design and Analysis of Experiments with R*. Vol. 115. CRC press, 2014.