Chương 4

Bố trí thí nghiệm một nhân tố

Đối với kiểu thiết kế thí nghiệm một nhân tố, chúng ta xem xét 3 mô hình thiết kế sau:

1) Mô hình thí nghiệm hoàn toàn ngẫu nhiên
2) Mô hình thí nghiệm khối ngẫu nhiên
3) Mô hình thí nghiệm ô vuông La tinh

4.1. Kiểu thí nghiệm hoàn toàn ngẫu nhiên (Completely randomized Design - CRD)

4.1.1. Đặc điểm

Đây là phương pháp nghiên cứu cơ bản trong các nghiên cứu chăn nuôi - thú y. Thí nghiệm được thiết kế đơn giản và việc phân tích các dữ liệu của thí nghiệm cũng dễ dàng.

Đối với mô hình thí nghiệm này, các đơn vị thí nghiệm được bố trí một cách hoàn toàn ngẫu nhiên vào các nghiệm thức, hay nói một cách khác, mỗi động vật thí nghiệm đều có cơ hội được phân vào một nghiệm thức bất kỳ và chịu ảnh hưởng tác động của nghiệm thức đó. Chính vì vậy, mô hình thí nghiệm này đòi hỏi các động vật thí nghiệm phải đồng đều. Mô hình này chỉ xem xét ảnh hưởng của một yếu tố, ví dụ nghiên cứu ảnh hưởng của thức ăn đến tăng trọng, tồn dư thuốc kháng sinh trong cơ thể vật nuôi..., các yếu tố còn lại được cho là không có sai khác, ví dụ tất cả các động vật được chọn có cùng một lứa tuổi, tất cả các trại đều sử dụng các thức ăn như nhau...

Với những yêu cầu nêu trên, trong lĩnh vực chăn nuôi và thú y, mô hình này chỉ thực hiện có hiệu quả khi động vật có tính đồng đều cao và các điều kiện phi thí nghiệm được kiểm soát một cách dễ dàng và có tính ổn định cao.

4.1.2. Chất lượng động vật

Động vật thí nghiệm đòi hỏi phải có sự đồng đều cao, vì vậy trong quá trình chọn động vật thí nghiệm, cần phải lưu ý đến các yếu tố như: giống, nguồn gốc, giới tính, thành tích của bố mẹ…

Chọn động vật cùng một giống. Động vật được chọn ra phải tiêu biểu cho giống đó, không quá khác biệt về ngoại hình và đặc điểm sinh lý. Để đạt được sự đồng đều cao, chọn những động vật là anh em ruột, nửa ruột thịt hoặc những động vật có quan hệ họ hàng trong cùng một dòng, một gia đình. Với thí nghiệm bố trí theo cặp tốt nhất dùng những động vật sinh đôi cùng trứng. Tuy nhiên trong thực tế, xác định được 2 động vật sinh đôi cùng trứng là phức tạp và tốn kém. Có thể chọn những động vật không cùng dòng, họ nhưng có ngoại hình tương đối đồng đều và đặc tính ổn định.

Để có động vật đồng đều, chỉ chọn những động vật cùng tính biệt, đồng đều theo lứa tuổi, mức độ tăng trưởng, cùng thể chất, tình trạng sức khoẻ... Trong một số trường hợp cần thiết tiến hành những nghiên cứu kiểm tra một số chỉ tiêu hoá sinh, sinh lý.

4.1.3. Dung lượng mẫu cần thiết

Một trong những yếu tố quan trọng trong quá trình thiết kế thí nghiệm là xác định số đơn vị thí nghiệm cần thiết. Tăng số lượng sẽ làm tăng độ chính xác của ước tính, tuy nhiên khi số lượng tăng sẽ đòi hỏi nhiều không gian, thời gian và nguồn lực. Số lượng có thể bị hạn chế bởi các yếu tố tài chính và điều kiện thực tế.

Khi số lượng được sử dụng đủ lớn thì gần như sự sai khác nào cũng có ý nghĩa thống kê. Sự sai khác, mặc dù có ý nghĩa thống kê, nhưng có thể không có ý nghĩa thực tiễn. Ví dụ, thí nghiệm so sánh tăng trọng của lợn ở 2 khẩu phần. Sự chênh lệch về tăng trọng trung bình ngày giữa 2 khẩu phần vài gram không có ý nghĩa về mặt thực tiễn cũng không có ý nghĩa về kinh tế; mặc dù đây là một thí nghiệm được thiết kế với quy mô lớn và sự sai khác này có ý nghĩa thống kê.

Đối với trường hợp thí nghiệm có nhiều nghiệm thức có thể dùng các đường cong cho sẵn để xác định dung lượng mẫu cần thiết. Dung lượng mẫu sẽ phụ thuộc vào sự sai khác mong đợi giữa các nghiệm thức, mức sai lầm loại I (α) và mức sai lầm loại II (β). Để có thể sử dụng được các đường cong này ta cần phải xác định được giá trị ϕ^2. Giá trị này được tính theo công thức:

$$\phi^2 = \frac{n\sum_{i=1}^{a} d_i^2}{a\sigma^2}$$

Trong đó n = số động vật cần thiết cho một nghiệm thức

a = số nghiệm thức

d_i = sai khác mong đợi của nghiệm thức thứ i với μ

σ^2 = phương sai của tính trạng cần nghiên cứu

Để xác định được ϕ cần phải chọn các giá trị trung bình, ví dụ ta có $\mu_1, \mu_2, \ldots, \mu_a$ la các giá trị trung bình của từng nghiệm thức. Ta sẽ có $\mu = (1/a)\sum_{i=1}^{a}\mu_i$ và $d_i = \mu_i - \mu$.

Ví dụ 4.1: muốn thiết kế một thí nghiệm đế so sánh tăng trọng (g) của gà ở 4 khẩu phần. Các giá trị trung bình được chọn lần lượt là $\mu_1 = 71$, $\mu_2 = 79$, $\mu_3 = 80$ và $\mu_4 = 102$ với $\alpha = 0,05$ và $1 - \beta = 0,80$; biết $\sigma^2 = 35^2$. Cần bao nhiêu đơn vị thí nghiệm?

Ta có:

$\mu = (71 + 79 + 80 + 102) / 4 = 83$

$d_1 = 71 – 83,00 = - 12$

$d_2 = 79 – 83,00 = - 4$

$d_3 = 80 – 83,00 = - 3$

$d_4 = 102 – 83,00 = + 9$

$\sum_{i=1}^{4} d_i^2 = 530$, vậy ta có:

$$\phi^2 = \frac{n\sum_{i=1}^{a} d_i^2}{a\sigma^2} = \frac{n(530)}{4(35)^2} = 0{,}11n$$

Ta sẽ sử dụng đường cong với bậc tự do của nghiệm thức là $v_1 = a - 1 = 4 - 1 = 3$, của sai số ngẫu nhiên là $v_2 = N - a = na - a = a(n - 1) = 4(n - 1)$ và $\alpha = 0{,}05$ ở phần phụ lục.

Nếu ta thử với n = 24 thì sẽ có các giá trị $\phi^2 = 0{,}11\times 6 = 2{,}64$; $\phi = 1{,}62$ $v_2 = 4(24 - 1) = 92$. Dựa vào đường cong sẽ có $\beta = 0{,}23$. Bằng cách tương tự ta có:

n	ϕ^2	ϕ	$4(n-1)$	β	$1-\beta$
24	2,64	1,62	92	0,23	0,77
25	2,75	1,66	96	0,21	0,79
26	2,86	1,69	100	0,19	0,81
27	2,97	1,72	104	0,17	0,83
28	3,08	1,75	108	0,16	0,84

Để thoả mãn điều kiện của bài toán, ta cần chọn ít nhất 26 đơn vị thí nghiệm.

Để có thể sử dụng được đường cong cho sẵn, khó nhất đối với người thiết kế thí nghiệm là phải chọn ra các giá trị trung bình cho từng nghiệm thức để từ đó có thể xác định được dung lượng mẫu cần thiết. Có một cách tiếp cận khác đơn giản hơn để xác định dung lượng mẫu đó là chỉ cần xác định một giá trị d. Sự sai khác của 2 giá trị trung bình bất kỳ nếu vượt quá giá trị d thì giả thiết H_0 bị bác bỏ. Khi đó giá trị ϕ^2 được tính theo công thức rút gọn sau đây (xem mục 3.8.1):

$$\phi^2 = \frac{nd^2}{2a\sigma^2}$$

Để minh hoạ, ta có thể lấy ví dụ trên. Nếu chọn d = 33 gram ta sẽ có

$$\phi^2 = \frac{nd^2}{2a\sigma^2} = \frac{n(33)^2}{2(4)(35)^2} = 0{,}11n$$

Tương tự như trên, ta cần ít nhất 26 đơn vị thí nghiệm để thoả mãn điều kiện bài ra.

4.1.4. Ưu điểm và nhược điểm

Ưu điểm của mô hình này là thí nghiệm thiết kế đơn giản, chính vì vậy cho nên hạn chế được nhiều sai sót trong quá trình thu thập dữ liệu. Mô hình phân tích số liệu không phức tạp, kết quả phân tích đơn giản, dễ đọc và dễ hiểu.

Mô hình có lợi thế là thích nghi một cách dễ dàng với trường hợp các đơn vị thí nghiệm không đều nhau vì các nguyên nhân nào đó, ví dụ như số liệu bị khiếm khuyết do tác động của bệnh trong quá trình làm thí nghiệm.

Ngược lại, mô hình thí nghiệm hoàn toàn ngẫu nhiên thường không có hiệu quả cao, hiệu lực của thí nghiệm không lớn do sự không thuần nhất của các vật liệu thí nghiệm.

4.1.5. Cách bố trí

Chọn n đơn vị thí nghiệm, bắt thăm n_1 đơn vị để bố trí mức A_1, bắt thăm n_2 đơn vị để bố trí mức A_2, . . . , bắt thăm n_{k-1} đơn vị để bố trí mức A_{a-1}, n_a đơn vị còn lại bố trí mức A_a. Như vậy là *bắt thăm toàn bộ các đơn vị thí nghiệm để bố trí một cách hoàn toàn ngẫu nhiên các mức của nhân tố*. Cách bố trí ngẫu nhiên được trình bày chi tiết ở chương 3.

Ví dụ yếu tố thí nghiệm A có 4 nghiệm thức A_1, A_2, A_3 và A_4 với các 5 đơn vị thí nghiệm trong mỗi nghiệm thức. Như vậy toàn bộ số đơn vị thí nghiệm là 20 và giả sử số động vật này được đánh số từ 1 đến 20. Sau khi bố trí một cách ngẫu nhiên ta có thể được mô hình thiết kế thí nghiệm như sau:

A_1	A_2	A_3	A_4
6	11	19	2
1	8	17	18
9	7	13	12
4	14	16	5
20	10	3	15

Khi kết thúc thí nghiệm, số liệu có thể ghi lại để dễ dàng và thuận tiện cho việc tính toán như sau:

A_1		A_2		A_3		A_4	
6	x_{11}	11	x_{21}	19	x_{31}	2	x_{41}
1	x_{12}	8	x_{22}	17	x_{32}	18	x_{42}
9	x_{13}	7	x_{23}	13	x_{33}	12	x_{43}
4	x_{14}	14	x_{24}	16	x_{34}	5	x_{44}
20	x_{15}	10	x_{25}	3	x_{35}	15	x_{45}

Dưới dạng tổng quát với a nghiệm thức số lần lặp lại r ta có:

A_1	A_2	…	A_a
x_{11}	x_{21}	…	x_{a1}
x_{12}	x_{22}	…	x_{a2}
x_{13}	x_{23}	…	x_{a3}
…	…	…	…
x_{1r}	x_{2r}	…	x_{ar}

4.1.6. Phân tích số liệu

Với các thí nghiệm được bố trí đơn giản với 2 nghiệm thức. Tiến hành so sánh kết quả của 2 nghiệm thức bằng phép thử *t*. Nếu thí nghiệm bao gồm nhiều nghiệm thức, thì phân tích phương sai (ANOVA) là phù hợp nhất. Phép thử t và phân tích phương sai được trình bày chi tiết ở Chương 2.

4.1.6.1. Mô hình phân tích

$$x_{ij} = \mu + a_i + e_{ij} \qquad (i = 1, a;\ j = 1, r_i)$$

trong đó μ trung bình chung

ai chênh lệch do ảnh hưởng của mức i

e_{ij} sai số ngẫu nhiên; các e_{ij} độc lập, phân phối chuẩn N $(0,\sigma^2)$

4.1.6.2. Cách phân tích

Cách phân tích số liệu được trình bày chi tiết ở Chương 2. Lưu ý rằng, trong mô hình thí nghiệm hoàn toàn ngẫu nhiên có 2 nguồn biến động: 1) biến động giữa các nghiệm thức (SS_A) và 2) biến động do sai số ngẫu nhiên (SS_E); toàn bộ biến động của thí nghiệm (SS_{TO}) bằng tổng số các các biến động thành phần (SS_A và SS_E) hợp thành. Các nguồn biến động này có thể được tính như sau:

Tổng bình phương toàn bộ biến động

$$SS_{TO} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2 = \sum_{i=1}^{a}\sum_{j=1}^{n_i} x_{ij}^2 - G$$

Tổng bình phương do nhân tố

$$SS_A = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(\bar{x}_i - \bar{x}\ )^2 = \sum_{i=1}^{a}\frac{TA_i^2}{r_i} - G$$

Tổng bình phương do sai số

$$SS_E = SS_{TO} - SS_A = \sum_{i=1}^{t}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_{i.}\right)^2$$

Các bậc tự do $df_{TO} = n - 1$; $df_A = a - 1$; $df_E = n - a$

Các trung bình $MS_A = SS_A / df_A$; $MS_E = SSE / df_E$

$F_{TN} = MS_A / MS_E$; giá trị tới hạn $F_{(\alpha,dfA,dfE)}$

Kết luận:

Nếu $F_{TN} \le F_{(\alpha,dfA,dfE)}$ thì chấp nhận H_0, ngược lại thì bác bỏ H_0

Bảng phân tích phương sai

Nguồn biến động	df	SS	MS	F_{TN}	F
Nhân tố	a -1	SS_A	MS_A	MS_A / MS_E	$F(\alpha, df_A, df_E)$
Sai số	n - a	SS_E	MS_E		
Toàn bộ	n -1	SS_{TO}			

Ví dụ 4.2: Một thí nghiệm được tiến hành để so sánh mức độ tăng trọng của gà ở 4 khẩu phần ăn khác nhau. Chọn 20 con gà đồng đều nhau và phân một cách ngẫu nhiên vào một trong 4 khẩu phần. Như vậy ta có 4 nhóm động vật thí nghiệm, mỗi nhóm gồm 5 gà; kết quả thí nghiệm được ghi lại ở bảng sau (đơn vị tăng trọng tính theo g):

Khẩu phần 1	Khẩu phần 2	Khẩu phần 3	Khẩu phần 4
99	61	42	169
88	112	97	137
76	30	81	169
38	89	95	85
94	63	92	154

Đây là ví dụ về thí nghiệm được bố trí theo mô hình một nhân tố hoàn toàn ngẫu nhiên. Yếu tố thí nghiệm là *Khẩu phần* với 4 nghiệm thức (*Khẩu phần* 1, 2, 3 và 4).

Ta có bảng phân tích phương sai

Nguồn biến động	df	SS	MS	F_{TN}	$F_{(0,05;\,3;\,16)}$
Khẩu phần	3	16467	5489	6,65	3,24
Sai số ngẫu nhiên	16	13212	826		
Tổng biến động	19	29679			

Kết luận: Bác bỏ H_0, như vậy tăng trọng của gà ở 4 khẩu phần ăn không phải như nhau.

Sự sai khác nhỏ nhất có ý nghĩa (Least Significant Difference - LSD) đối với 2 mức A_i và A_j có số lần lặp n_i và n_j tính theo công thức:

$$LSD_\alpha = t_{(\alpha/2, dfE)} \times \sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Nếu chọn mức ý nghĩa $\alpha = 0,05$ $t(0,025;16) = 2,12$; $n_i = n_j = 5$ do đó khi so sánh các trung bình có thể dùng $LSD_{0,05} = 2,12 \times \sqrt{826 \times \frac{2}{5}} = 38,54$

So các trung bình:

(A_1) so với (A_2)	|79 - 71| = 8 < 38,54	Sai khác không có ý nghĩa
(A_1) so với (A_3)	|79 - 81,4| = 2,4 < 38,544	Sai khác không có ý nghĩa
(A_1) so với (A_4)	|79 - 142,8| = 63,8 > 38,54	Sai khác có ý nghĩa
(A_2) so với (A_3)	|71 - 81,4| = 10,4 < 38,54	Sai khác không có ý nghĩa
(A_2) so với (A_4)	|71 - 142,8| = 71,8 > 38,54	Sai khác có ý nghĩa
(A_3) so với (A_4)	|81,4 - 142,8| = 61,4 > 38,54	Sai khác có ý nghĩa

Ta có thể xây dựng một bảng có các chữ cái a, b, c... để thể hiện sự sai khác giữa các nghiệm thức theo các bước sau:

1) Sắp xếp các giá trị trung bình theo thứ tự giảm dần như sau:

Khẩu phần	Trung bình
1	79,00
2	71,00
3	81,40
4	142,80

Khẩu phần	Trung bình
4	142,80
3	81,40
1	79,00
2	71,00

2) Dựa vào kết quả so sánh để tạo các đường gạch chung cho các khẩu phần có giá trị trung bình bằng nhau; cụ thể như sau:

Khẩu phần	Trung bình		
		a	
4	142,80		b
3	81,40		
1	79,00		
2	71,00		

mỗi một đường thẳng tương ứng với một chữ cái (a, b, c...)

3) Từ bảng trên, ta có thể đặt các chữ cái bên cạnh các số trung bình và sắp xếp khẩu phần theo thứ tự tăng dần như ban đầu ta có như sau:

Khẩu phần	Trung bình
4	$142,80^{\mathbf{a}}$
3	$81,40^{\mathbf{b}}$
1	$79,00^{\mathbf{b}}$
2	$71,00^{\mathbf{b}}$

Khẩu phần	Trung bình
1	$79,00^{\mathbf{b}}$
2	$71,00^{\mathbf{b}}$
3	$81,40^{\mathbf{b}}$
4	$142,80^{\mathbf{a}}$

Việc so sánh hai trung bình theo LSD thường chỉ dùng để so sánh một số cặp trung bình mà trước khi thí nghiệm chúng ta đã có ý đồ so sánh. Nếu so sánh tất cả các cặp trung bình, hay còn gọi là kiểm định sự bằng nhau của tất cả các cặp trung bình (multiple comparisons) thì mức ý nghĩa không còn là α mà nhỏ đi nhiều, do đó các nhà nghiên cứu thống kê đã đề xuất nhiều cách kiểm định khác để đảm bảo mức ý nghĩa α như kiểm định Scheffé, Tukey, Bonferroni, Dunnett, kiểm định đa phạm vi (multiple range test) Duncan, Student- Newman - Keuls, . . .Trong các chương trình máy tính chuyên về thống kê còn có nhiều cách so sánh khác.

Thí dụ muốn so sánh theo Duncan (các lần lặp bằng nhau và gọi là r) phải sắp các trung bình từ nhỏ đến lớn. Khi so sánh hiệu số các trung bình thì, tuỳ theo các trung bình ở kề nhau hay cách nhau một trung bình, cách nhau hai trung bình, . . .mà dùng các ngưỡng so sánh khác nhau. Việc so sánh tiến hành như sau:

1) Tính sai số của trung bình $s_{\bar{x}i} = \sqrt{\dfrac{MS_E}{r}}$

2) Lấy giá trị r_p trong bảng Duncan ứng với bậc tự do df_E nhân với $s_{\bar{x}i}$ để có khoảng R_p.

3) So sánh hiệu $x_j - \bar{x}_i$ với R_p.

Nếu hai trung bình liền nhau thì lấy p = 2, cách nhau một thì p = 3, cách nhau hai thì p = 4, . . .

Nếu hiệu bé hơn hay bằng R_p thì sai khác không có ý nghĩa, ngược lại thì sai khác có ý nghĩa.

Trong thí dụ trên

(A_2)	(A_1)	(A_3)	(A_4)
71,0	79,0	(81,4)	(142,8)

$s_{\bar{x}i} \sqrt{\dfrac{826}{5}} = 12{,}853$ với bậc tự do df_E = 16

p	2	3	4
r_p	3,0	3,15	3,23
R_p	38,56	40,49	41,52

$(A_1) - (A_2) = 79{,}0 - 71{,}0 = 8 < R_2 = 38{,}56$ Sai khác không có ý nghĩa

$(A_3) - (A_2) = 81{,}4 - 71{,}0 = 10{,}4 < R_3 = 40{,}49$ Sai khác không có ý nghĩa

$(A_4) - (A_2) = 142{,}8 - 71 = 71{,}8 > R_4 = 41{,}52$ Sai khác có ý nghĩa

$(A_3) - (A_1) = 81{,}4 - 79{,}0 = 2{,}4 < R_2 = 38{,}56$ Sai khác không có ý nghĩa

$(A_4) - (A_1) = 142{,}8 - 79 = 63{,}8 > R_3 = 40{,}49$ Sai khác có ý nghĩa

$(A_4) - (A_3) = 142{,}8 - 81{,}4 = 61{,}4 > R_2 = 38{,}56$ Sai khác có ý nghĩa

Trong ví dụ này các kết luận không khác với so sánh theo LSD

4.2. Kiểu thí nghiệm khối ngẫu nhiên đầy đủ (Randomized complete block design - RCBD)

Như đã nêu trên, mô hình thiết kế thí nghiệm kiểu hoàn toàn ngẫu nhiên chỉ thực sự có hiệu quả khi toàn bộ động vật thí nghiệm có sự đồng đều cao và các điều kiện ngoại cảnh phải được kiểm soát dễ dàng. Trong thực tế, đặc biệt là trong chăn nuôi thú y rất khó có thể thoả mãn cùng một lúc các điều kiện đã nêu. Mô hình thiết kế thí nghiệm theo kiểu *khối ngẫu nhiên đầy đủ* được đưa ra nhằm hạn chế những khó khăn đó.

Nguyên tắc tạo khối là đạt được sự đồng đều tối đa trong một khối và sự khác nhau lớn nhất giữa các khối. Các khối được gọi là *đầy đủ* khi trong mỗi khối có đầy đủ các đại diện của các nghiệm thức và ngẫu nhiên khi các đơn vị thí nghiệm được bố trí một cách hoàn toàn *ngẫu nhiên* vào các nghiệm thức. Trong quá trình thí nghiệm, tất cả các đơn vị thí nghiệm trong cùng một khối nhận được tất các điều kiện như nhau ngoại trừ yếu tố thí nghiệm.

Trong chăn nuôi - thú y, khối có thể coi là các nhóm động vật cùng một giống, giới tính, tuổi, cùng khối lượng hoặc cũng có thể là nhóm động vật sinh ra cùng một bố, cùng lứa.

Một số lý do để chọn mô hình thí nghiệm khối ngẫu nhiên đầy đủ là:

a) Do không tìm được đủ n = $a \times b$ đơn vị thí nghiệm đồng đều do đó phải chọn b khối, mỗi khối có a đơn vị thí nghiệm để sắp xếp cho a mức của nhân tố.Ví dụ so sánh 6 công thức thí nghiệm, mỗi công thức lặp lại 5 lần. Giả sử ta không tìm được 30 con lợn đồng đều về khối lượng, do đó chọn 5 lô, mỗi lô 6 con đồng đều để bố trí 6 công thức.

b) Có thể có một nguồn biến động theo một hướng, thí dụ hướng nắng, hướng gió, hướng dốc, hướng chảy của nước ngầm, hướng thay đổi của chất đất, . . . khi đó phải bố trí các khối vuông góc với hướng biến động nhằm cân bằng tác động của biến động (vì mỗi công thức đều có mặt ở tất cả các khối, mỗi khối một lần).

4.2.1. Số khối cần thiết

Các kỹ thuật dùng để xác định dung lượng mẫu trong mô hình thiết kế thí nghiệm một nhân tố hoàn toàn ngẫu nhiên có thể được áp dụng trực tiếp đối với mô hình khối ngẫu nhiên đầy đủ. Các đường cong cho sẵn có thể được sử dụng với công thức:

$$\phi^2 = \frac{b\sum_{i=1}^{a} d_i^2}{a\sigma^2}$$

hoặc

$$\phi^2 = \frac{bd^2}{2a\sigma^2}$$

với b = số khối cần thiết.

Ví dụ ta chọn d = 0,76; α = 0,05; 1 - β = 0,8 ; số nghiệm thức a = 4 ; σ = 0,70; ta sẽ có

$$\phi^2 = \frac{bd^2}{2a\sigma^2} = \frac{b(1{,}72)^2}{2(4)(0{,}68)^2} = 0{,}8b$$

với các bậc tự do $v_1 = a$ -1 = 4 – 1 = 3 và $v_2 = (a - 1)(b - 1) = (4 - 1)(b - 1) = 3(b - 1)$

b	ϕ^2	ϕ	$3(b-1)$	β	$1-\beta$
3	2,40	1,55	6	0,60	0,40
4	3,20	1,79	9	0,30	0,70
5	4,00	2,00	12	0,20	0,80
6	4,80	2,19	15	0,12	0,88
7	5,60	2,37	18	0,08	0,92

Như vậy cần ít nhất 5 khối để thoả mãn điều kiện bài toán.

4.2.2. Ưu điểm và nhược điểm

Mô hình thí nghiệm kiểu khối ngẫu nhiên đầy đủ được thiết kế đơn giản gần như mô hình thí nghiệm kiểu hoàn toàn ngẫu nhiên. Mô hình thí nghiệm theo khối có thể được thiết kế với số nghiệm thức và với số lần lặp bất kỳ; nhưng đòi hỏi số lần lặp lại phải bằng nhau ở các nghiệm thức.

Mô hình thí nghiệm khối ngẫu nhiên đầy đủ chỉ thể hiện đầy đủ các ưu thế cho đến khi có một hay nhiều nghiệm thức hoặc khối bị loại bỏ, ví dụ có những số liệu bị khuyết trong quá trình thu thập hoặc trong quá trình phân tích.

So với mô hình thí nghiệm kiểu hoàn toàn ngẫu nhiên, mô hình khối ngẫu nhiên đầy đủ cho hiệu quả và độ chính xác cao hơn. Điều này được thể hiện rõ, với cùng một nguyên vật liệu thí nghiệm sẽ cho kết quả chính xác hơn hoặc với cùng một độ chính xác có thể giảm được nguyên vật liệu thí nghiệm. Độ chính xác của thí nghiệm tăng lên bởi vì biến động giữa các khối đã được loại bỏ trong quá trình phân tích và khả năng phát hiện được ảnh hưởng của các nghiệm thức tăng lên. Tuy nhiên với một số công thức thí nghiệm tương đối lớn (ví dụ nhiều hơn 20 nghiệm thức) và với các nguyên vật liệu có độ đồng đều thấp thì hiệu quả của mô hình bị giảm một cách đáng kể; khi đó mô hình khối không đầy đủ sẽ được áp dụng.

4.2.3. Cách bố trí thí nghiệm

Chọn b khối, mỗi khối có a đơn vị thí nghiệm, bắt thăm ngẫu nhiên để xếp a đơn vị thí nghiệm vào a công thức thí nghiệm trong khối 1, sau đó bắt thăm để xếp a công thức vào a ô trong khối 2, . . ., cuối cùng là bắt thăm cho khối b.

Ví dụ bố trí thí nghiệm với 4 nghiệm thức (A1, A2, A3 và A4) với 5 khối khác nhau (1, 2, 3, 4 và 5). Như vậy ta sẽ tạo ra 5 khối khác nhau đảm bảo sự đồng đều tối đa trong từng khối, mỗi khối có 4 đơn vị thí nghiệm (4 lần lặp lại) và kỹ thuật bắt thăm hoàn toàn ngẫu nhiên để phân 4 động vật thí nghiệm trong từng khối về với 4 công thức thí nghiệm.

Nếu động vật thí nghiệm được đánh số theo sơ đồ sau:

	Khối				
	1	2	3	4	5
Động vật thí nghiệm số	1	5	9	13	17
	2	6	10	14	18
	3	7	11	15	19
	4	8	12	16	20

Sau khi bố trí các đơn vị bằng cách bốc thăm ngẫu nhiên, sơ đồ thiết kế thí nghiệm có thể được trình bày theo sơ đồ:

	Khối				
Công thức	1	2	3	4	5
A1	1	8	11	14	18
A2	4	6	9	15	19
A3	2	7	10	16	17
A4	3	5	12	13	20

Số liệu thu được khi kết thúc thí nghiệm có thể được trình bày

	Khối									
Công thức	1		2		3		4		5	
A1	1	x_{11}	8	x_{12}	11	x_{13}	14	x_{14}	18	x_{15}
A2	4	x_{21}	6	x_{22}	9	x_{23}	15	x_{24}	19	x_{25}
A3	2	x_{31}	7	x_{32}	10	x_{33}	16	x_{34}	17	x_{35}
A4	3	x_{41}	5	x_{42}	12	x_{43}	13	x_{44}	20	x_{45}

Hay ở dạng tổng quát với a công thức và b khối

	Khối			
Công thức	1	2	…	b
A1	x_{11}	x_{12}	…	x_{1b}
A2	x_{21}	x_{22}	…	x_{2b}
…	…	…	…	…
Aa	x_{a1}	x_{a2}	…	x_{ab}

4.2.4. Phân tích số liệu

Phân tích phương sai (ANOVA) được sử dụng để phân tích số liệu. Trong mô hình thí nghiệm kiểu khối ngẫu nhiên đầy đủ có 3 nguồn biến động: 1) biến động giữa các khối (SS_K), 2) biến động giữa các nghiệm thức (SS_A) và 3) biến động do sai số ngẫu nhiên (SS_E); toàn bộ biến động của thí nghiệm (SS_{TO}) chính bằng tổng các các biến động thành phần (SS_K, SS_A và SS_E). Các nguồn biến động này có thể được trình bày qua mô hình phân tích dưới đây

4.2.4.1. Mô hình phân tích

$x_{ij} = \mu + a_i + b_j + e_{ij}$ i = 1,…,a; j = 1,…,b

μ là trung bình chung.

a_i là chênh lệch do ảnh hưởng của mức i của nhân tố, $\Sigma a_i = 0$

b_j là chênh lệch do ảnh hưởng của khối j , $\Sigma b_j = 0$

e_{ij} là sai số ngẫu nhiên; các e_{ij} độc lập, phân phối chuẩn $N(0,\sigma^2)$

4.2.4.2. Cách phân tích

Tính tổng bình phương toàn bộ SS_{TO}

$$SS_{TO} = \sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij} - \bar{x})^2$$

Tính tổng bình phương do nhân tố SS_A

$$SS_A = \sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}_{i.} - \bar{x})^2$$

Tính tổng bình phương do khối SS_K

$$SS_K = \sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}_{.j} - \bar{x})^2$$

Tính trung bình do sai số SS_E

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2$$

Cũng có thể tính nhanh các tổng bình phương như sau:

Tính tổng hàng (nghiệm thức) TA_i (i = 1, *a*), trung bình hàng (nghiệm thức) $\bar{x}_{i.}$

Tổng cột (khối) TK_j (j = 1, r), trung bình cột $\bar{x}_{.j}$

Tổng số quan sát $n = a \times b$.

Tổng toàn bộ các số liệu $ST = \Sigma\Sigma x_{ij}$, trung bình toàn bộ $\bar{x}$

Tính số hiệu chỉnh $G = ST^2 / n$

$$SS_{TO} = \sum_{i=1}^{a}\sum_{j=1}^{b} x_{ij}^2 - G$$

$$SS_A = \frac{1}{b}\sum_{i=1}^{a} TA_i^2 - G$$

$$SS_K = \frac{1}{a}\sum_{j=1}^{b} TK_j^2 - G$$

$SS_E = SS_{TO} - SS_A - SS_K$

Bậc tự do $df_{TO} = n - 1 = a \times b - 1$; $df_A = a - 1$; $df_K = r - 1$; $df_E = (a-1)(b-1)$

Các trung bình bình phương:

$MS_A = SS_A / df_A$; $MS_K = SS_K / df_K$; $MS_E = SSE / df_E$

Trong quá trình phân tích thường ít chú ý kiểm định khối mà chỉ tập trung kiểm định nhân tố. Giả thiết H_0 : "Các trung bình của các mức bằng nhau", đối thiết H_1: "Có ít nhất một cặp trung bình khác nhau"

Tính $F_{TN} = MS_A / MS_E$; so với giá trị tới hạn $F_{(\alpha, dfA, dfE)}$

Kết luận:

Nếu $F_{TN} \le F_{(\alpha, dfA, dfE)}$ thì chấp nhận H_0, ngược lại thì bác bỏ H_0

Dưới dạng tổng hợp ta có bảng phân tích phương sai

Nguồn biến động	df	SS	MS	F_{TN}	F tới hạn
Nhân tố	a-1	SS_A	MS_A	MS_A / MS_E	$F_{(\alpha, dfA, dfE)}$
Khối	b-1	SS_K	MS_K		
Sai số	(a-1)(b-1)	SS_E	MS_E		
Toàn bộ	ab -1	SS_{TO}			

Ví dụ 4.3: (Mead và cộng sự) Nghiên cứu số lượng tế bào lymphô ở chuột (×1000 tế bào mm^{-3} máu) được sử dụng 4 loại thuốc khác nhau (A, B, C và D; thuốc D là placebo) qua 5 lứa; số liệu thu được như sau:

	Lứa 1	Lứa 2	Lứa 3	Lứa 4	Lứa 5
Thuốc A	7,1	6,1	6,9	5,6	6,4
Thuốc B	6,7	5,1	5,9	5,1	5,8
Thuốc C	7,1	5,8	6,2	5,0	6,2
Thuốc D	6,7	5,4	5,7	5,2	5,3

Đây là mô hình thí nghiệm kiểu khối ngẫu nhiên đầy đủ với số công thức thí nghiệm a = 4, số khối chính bằng số lứa b = 5.

n = 4 x 5 = 20; ST = 119,3 ; G = $119{,}3^2$ / 20 =711,6245; $\Sigma\, x^2_{i\,j}$ = 720,51

(ΣTA^2_i) / r = 3567,35 / 5 = 713,47; (ΣTK^2_j) / a = 2872,11 / 4 = 718,0275

SS_{TO} = 720,51 – 711,6245 = 8,8855

SS_A = 713,47 – 711, 6245 = 1,8455

SS_K = 718,0275 – 711,6245 = 6,4030

SS_E = 8,8855 – 1,8455 – 6,4030 = 0,6370

Bảng phân tích phương sai

Nguồn	df	SS	MS	F_{TN}	$F_{(0,05;\,3;\,12)}$
Thuốc	3	1,8455	0,6152	11,59	3,49
Lứa	4	6,4030	1,6007		
Sai số	12	0,6370	0,0531		
Tổng số	19	8,8855			

Kết luận: Bác bỏ giả thiết H_0, điều này chứng tỏ khi sử dụng các loại thuốc khác nhau đã làm cho số lượng tế bào lymphô trong máu thay đổi.

Sai số thí nghiệm se$=\sqrt{MS_E}=\sqrt{0{,}0531}=0{,}2304$

Có thể sử dụng sai khác bé nhất có ý nghĩa ở mức 5% (LSD) để xác định sự sai khác có ý nghĩa thống kê của các cặp giá trị trung bình bất kỳ

$$LSD(0{,}05)=t_{dfE}^{(0{,}025)}\times\sqrt{\frac{MS_E\times 2}{b}}=2{,}179\times\sqrt{\frac{0{,}0531\times 2}{5}}=0{,}3176$$

Trung bình

(A) = 6,42

(B) = 5,72

(C) = 6,06

(D) =(5,66)

So (A) với (B) | 6,42 - 5,72 | = 0,70 > LSD Khác nhau có ý nghĩa

So (A) với (C) | 6,42 - 6,06 | = 0,36 > LSD Khác nhau có ý nghĩa

So (A) với (D) | 6,42 - 5,66 | = 0,76 > LSD Khác nhau có ý nghĩa

So (B) với (C) | 5,72 – 6,06 | = 0,34 > LSD Khác nhau có ý nghĩa

So (B) với (D) | 5,72 - 5,66 | = 0,06 < LSD Khác nhau không có ý nghĩa

So (C) với (D) | 6,06 - 5,66 | = 0,40 > LSD Khác nhau không có ý nghĩa

Sau khi so sánh ta có được các giá trị trung bình cùng với các chữ cái tương ứng thể hiện sự sai khác như sau:

A $6,42^{\mathbf{a}}$

B $5,72^{\mathbf{b}}$

C $6,06^{\mathbf{c}}$

D $5,66^{\mathbf{b}}$

Như vậy, các giá trị trung bình không có chữ giống nhau thì khác nhau (P < 0,05)

4.3. Khối ngẫu nhiên với nhiều đơn vị thí nghiệm ở một nghiệm thức và khối

4.3.1. Cách bố trí

Trong phần trước, đối với thí nghiệm khối ngẫu nhiên đầy đủ chỉ có một đơn vị thí nghiệm trong một tổ hợp (nghiệm thức × khối) và sai số ngẫu nhiên của mô hình chính bằng tương tác giữa nghiệm thức và khối. Chính vì vậy không thể kiểm tra được tác động tương tác giữa nghiệm thức và khối. Giải pháp duy nhất để kiểm tra tác động tương tác là tăng số đơn vị thí nghiệm trong mỗi tổ hợp (nghiệm thức × khối) lên ít nhất 2 đơn vị. Một lần nữa xem xét *a* nghiệm thức và *b* khối, nhưng trong mỗi tổ hợp (nghiệm thức × khối) có n đơn vị thí nghiệm. Như vậy số đơn vị thí nghiệm trong mỗi khối sẽ là (n × a) và được bố trí một cách ngẫu nhiên vào với các nghiệm thức đảm bảo mỗi nghiệm thức trong khối có n đơn vị thí nghiệm.

Ví dụ: Một thí nghiệm có 5 khối, 4 nghiệm thức và 8 đơn vị thí nghiệm trong từng khối; do đó sẽ có 2 đơn vị thí nghiệm trong một tổ hợp (nghiệm thức × khối). Sơ đồ thiết kế thí nghiệm được thể hiện như sau:

	Khối				
Công thức	1	2	3	4	5
A1	1	12	23	26	39
	7	11	18	31	37
A2	8	9	19	25	36
	6	15	20	32	38
A3	4	10	24	29	33
	5	16	17	27	40
A4	3	13	22	30	35
	2	14	21	28	34

Số liệu khi kết thúc thí nghiệm có thể được trình bày như sau:

	Khối				
Công thức	1	2	3	4	5
A1	x_{111} x_{112}	x_{121} x_{122}	x_{131} x_{132}	x_{141} x_{142}	x_{151} x_{152}
A2	x_{211} x_{212}	x_{221} x_{222}	x_{231} x_{232}	x_{241} x_{242}	x_{251} x_{252}
A3	x_{311} x_{312}	x_{321} x_{322}	x_{331} x_{332}	x_{341} x_{342}	x_{351} x_{352}
A4	x_{411} x_{412}	x_{421} x_{422}	x_{431} x_{432}	x_{441} x_{442}	x_{451} x_{452}

4.3.2. Mô hình phân tích

$x_{ijk} = \mu + a_i + b_j + a{\times}b_{ij} + e_{ij}$ i = 1,…,a; j = 1,…,b; k = 1,…,n

x_{ijk} là quan sát thứ *k* của khối thứ *j* và nghiệm thức thứ *i*

μ trung bình chung.

a_i chênh lệch do ảnh hưởng của mức i của nhân tố $\Sigma a_i = 0$

b_j chênh lệch do ảnh hưởng của khối j , $\Sigma b_j = 0$

$a{\times}b_{ij}$ chênh lệch do tương tác giữa nghiệm thức và khối

e_{ijk} sai số ngẫu nhiên; các e_{ijk} độc lập, phân phối chuẩn $N(0,\sigma^2)$

4.3.3. Cách phân tích

Trong mô hình này, các nguồn biến động bao gồm: 1) biến động giữa các khối (SS_K), 2) biến động giữa các nghiệm thức (SS_A), 3) biến động do ảnh hưởng tương tác (SS_{AK}) và 4) biến động do sai số ngẫu nhiên (SS_E); toàn bộ biến động của thí nghiệm (SS_{TO}) chính bằng tổng các các biến động thành phần (SS_K, SS_A, SS_{AK} và SS_E). Các nguồn biến động này có thể tính như sau:

Tính tổng bình phương toàn bộ SS_{TO}

$$SS_{TO} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(x_{ijk} - \bar{x})^2$$

Tính tổng bình phương do nhân tố SS_A

$$SS_A = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{x}_{i..} - \bar{x})^2 = bn\sum_{i=1}^{a}(\bar{x}_{i..} - \bar{x})^2$$

Tính tổng bình phương do khối SS_K

$$SS_K = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{x}_{.j.} - \bar{x})^2 = an\sum_{j=1}^{b}(\bar{x}_{.j.} - \bar{x})^2$$

Tính tổng bình phương do tương tác giữa nhân tố và khối SS_{AK}

$$SS_{AK} = n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}_{ij.} - \bar{x})^2 - SS_K - SS_A$$

Tổng bình phương do sai số $SS_E = SS_{TO} - SS_A - SS_K$

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(x_{ijk} - \bar{x}_{ij.}\right)^2$$

Có thể tính nhanh các tổng bình phương như sau:

$$SS_{TO} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} x_{ijk}^2 - G$$

$$SS_A = \frac{1}{bn}\sum_{i=1}^{a}\left(\sum_{j=1}^{b}\sum_{k=1}^{n} x_{ijk}\right)^2 - G$$

$$SS_K = \frac{1}{bn}\sum_{j=1}^{b}\left(\sum_{i=1}^{a}\sum_{k=1}^{n} x_{ijk}\right)^2 - G$$

$$SS_{AK} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\sum_{k=1}^{n} x_{ijk}\right)^2 - SS_K - SS_A - G$$

$$SS_E = SS_{TO} - SS_A - SS_K$$

Bậc tự do df_{TO} = abn -1; df_A = a -1; df_K = b - 1; df_{AK} = (a-1)(b-1); df_E = ab(n-1)

Các trung bình bình phương:

$MS_A = SS_A / df_A$; $MS_K = SS_K / df_K$; $MS_{AK} = SS_{AK} / df_{AK}$; $MS_E = se^2 = SSE / df_E$

Giả thiết đối với tương tác giữa nghiệm thức và khối; H_0: Không có tương tác giữa nghiệm thức và khối với đối thiết H_1: Có tương tác giữa nghiệm thức và khối.

Tính $F_{TN} = MS_{AK} / MSE$; so với giá trị tới hạn $F_{(\alpha, dfAK, dfE)}$; nếu $F_{TN} \le F_{(\alpha, dfAK, dfE)}$ thì chấp nhận H_0, ngược lại thì bác bỏ H_0

Giả thiết đối với yếu tố thí nghiệm; H_0 : "Các trung bình của các mức bằng nhau" với đối thiết H_1: "Có ít nhất một cặp trung bình khác nhau".

Tính $F_{TN} = MS_A / MS_E$; so với giá trị tới hạn $F_{(\alpha, dfA, dfE)}$; nếu $F_{TN} \le F_{(\alpha, dfA, dfE)}$ thì chấp nhận H_0, ngược lại thì bác bỏ H_0

Dưới dạng tổng hợp ta có bảng phân tích phương sai

Nguồn biến động	df	SS	MS	F_{TN}	F
Nhân tố	a-1	SS_A	MS_A	MS_A / MS_E	$F_{(\alpha, dfA, dfE)}$
Khối	b-1	SS_K	MS_K		
Nhân tố × Khối	(a-1)(b-1)	SS_{AK}	MS_{AK}	MS_{AK} / MS_E	$F_{(\alpha, dfAK, dfE)}$
Sai số	ab(n-1)	SS_E	MS_E		
Toàn bộ	abn -1	SS_{TO}			

Ví dụ 4.4: Một thí nghiệm được tiến hành để xác định ảnh hưởng của 3 công thức thí nghiệm (A1, A2 và A3) đến tăng trọng trung bình trên ngày (gram / ngày) của bê đực. Bê đực được cân và chia thành 4 khối dựa theo khối lượng bắt đầu thí nghiệm. Trong mỗi khối có 6 động vật thí nghiệm được chọn ra và được phân ngẫu nhiên về với các nghiệm thức. Như vậy toàn bộ số động vật thí nghiệm tham gia thí nghiệm là 4×3×2=24 bê. Số liệu thu thập sau khi kết thúc thí nghiệm như sau:

	Khối			
	I	II	III	IV
A1	826	864	795	850
	806	834	810	845
A2	827	871	729	860
	800	881	709	840
A3	753	801	736	820
	773	821	740	835

Tổng bình phương do nghiệm thức SS_A = 8025,58

Tổng bình phương do khối SS_K = 33816,83

Tổng bình phương do tương tác giữa khối và nghiệm thức SS_{AK} = 8087,42

Tổng bình phương do sai số SS_E = 2110,00

Bảng phân tích phương sai (ANOVA)

Nguồn biến động	df	SS	MS	F_{TN}	F
Nhân tố	2	8025,58	4012,79	22,82	$F_{(0,05,\ 2,\ 12)} = 3,89$
Khối	3	33816,83	11272,28		
Nhân tố × Khối	6	8087,42	1347,90	7,67	$F_{(0,05,\ 6,\ 12)} = 3,00$
Sai số	12	2110,00	175,83		
Toàn bộ	23	52039,83			

Như vậy, ở mức $\alpha = 0,05$; giả thiết H_0 bị bác bỏ đối với cả nghiệm thức và tương tác (nghiệm thức × khối). Điều này chứng tỏ rằng có ảnh hưởng của nghiệm thức và ảnh hưởng này khác nhau ở từng khối khác nhau. Hay nói một cách khác, ảnh hưởng của nghiệm thức khác nhau tuỳ thuộc vào khối lượng vào thời điểm bắt đầu thí nghiệm.

4.4. Kiểu thí nghiệm ô vuông La tinh

Ngoài kiểu bố trí hoàn toàn ngẫu nhiên và khối ngẫu nhiên đầy đủ còn hay dùng kiểu ô vuông La tinh trong thí nghiệm một nhân tố. Trong mô hình này nghiệm thức được bố trí vào các khối theo 2 hướng khác nhau, thường gọi là hàng và cột. Mỗi hàng và mỗi cột là một khối đầy đủ chứa tất cả các nghiệm thức.

Kiểu thí nghiệm này được lựa chọn khi khảo sát nhân tố trong hoàn cảnh có hai hướng biến động mà chúng ta muốn cân bằng, ví dụ theo dõi sản lượng sữa của các bò sữa ở các công thức thí nghiệm khác nhau và trong các giai đoạn tiết sữa khác nhau trong chu kỳ tiết sữa.

Mô hình này đặc biệt hữu ích đối với thí nghiệm có số lượng động vật bị hạn chế và sự đồng đều không cao. Ví dụ nghiên cứu sự biến đổi protein trong dạ cỏ bằng cách sử kỹ thuật lỗ dò dạ cỏ ở 4 động vật; 4 loại thức ăn (A, B, C và D) được tiến hành nghiên cứu, mỗi loại thức ăn chứa trong các túi nilon được đặt trong dạ cỏ của từng động vật trong các khoảng thời gian khác nhau.

Đặc điểm của cách bố trí này là mỗi mức của nhân tố có mặt một lần ở mỗi hàng và một lần ở mỗi cột, sự sắp xếp này là hoàn toàn ngẫu nhiên; ví dụ theo dõi lượng sữa của 4 con bò sữa trong 4 giai đoạn trong chu kỳ tiết sữa, khi cho ăn theo 4 công thức A_1, A_2, A_3, A_4.

Số nghiệm thức chính bằng số hàng và số cột còn số ô vuông cần thiết chính là bình phương của số nghiệm thức. Lưu ý rằng, tất cả các động vật tham gia thí nghiệm phải được giữ lại đến khi kết thúc thí nghiệm, nếu không trong quá trình xử lý số liệu sẽ gặp nhiều khó khăn.

Mô hình ô vuông La tinh thường được sử dụng với số nghiệm thức từ 4 đến 8, hay sử dụng nhất là mô hình 4×4 và ít sử dụng đối với mô hình lớn hơn 8×8.

4.4.1. Ưu điểm và nhược điểm của mô hình

Trong mô hình thí nghiệm này, hai hướng biến động được kiểm soát đồng thời, vì vậy mô hình này về cơ bản cho hiệu quả cao hơn so với mô hình thí nghiệm kiểu hoàn toàn ngẫu nhiên và khối ngẫu nhiên đầy đủ, đồng thời giảm được số động vật tham gia thí nghiệm cũng như khắc phục được sự kém đồng đều của động vật thí nghiệm.

Tuy nhiên, kiểu thí nghiệm này có những nhược điểm là số mức của hai hướng biến động phải chọn bằng nhau và bằng số mức của nhân tố, giả thiết rằng không có tương tác giữa các hướng với nhau và với nhân tố; thêm vào đó, số bậc tự do của sai số ngẫu nhiên tương đối nhỏ, nên các kiểm định F trong phân tích phương sai và các kiểm định về các trung bình kém chính xác.

4.4.2. Cách bố trí

Có *a* mức của nhân tố (A_1, A_2, . . . ,A_a). Chọn *a* mức của hướng biến động thứ nhất, gọi đó là a hàng. Chọn a mức của hướng biến động thứ hai, gọi đó là a cột. Chọn một sơ đồ ô vuông La tinh a × a để sau đó bắt thăm a mức của nhân tố vào các ô trong sơ đồ. Lưu ý rằng, cần phải tiến hành ngẫu nhiên hoá theo hàng hoặc theo cột cũng như bố trí các nghiệm thức trong các hàng và các cột phải tuân thủ theo nguyên tắc ngẫu nhiên.

Ví dụ bố trí thí nghiệm theo mô hình ô vuông La tinh 4 × 4, sơ đồ thiết kế thí nghiệm cơ bản có trong các bảng in sẵn hoặc có thể tự làm một cách đơn giản như sau. Hàng đầu viết các chữ cái a b c d; hàng thứ hai đẩy b lên đầu còn a chạy xuống cuối, hàng thứ ba đẩy c lên còn b chạy xuống cuối, . . . Cách này gọi tắt là xếp hàng vòng quanh, sau đó ta được

a	b	c	d
b	c	d	a
c	d	a	b
d	a	b	c

Bắt thăm ngẫu nhiên 4 thẻ có ghi các số 1, 2, 3, 4. Thí dụ được 3 4 1 2; như vậy chúng ta có tương ứng: a → A_3, b → A_4, c → A_1, d → A_2

A_3	A_4	A_1	A_2
A_4	A_1	A_2	A_3
A_1	A_2	A_3	A_4
A_2	A_3	A_4	A_1

Ta có một sơ đồ thiết kế thí nghiệm với 4 nghiệm thức. Các cột và hàng được biểu thị tương ứng với các giai đoạn và các động vật thí nghiệm như sau:

	Cột (Động vật)			
Hàng (Giai đoạn)	1	2	3	4
1	A_3	A_4	A_1	A_2
2	A_4	A_1	A_2	A_3
3	A_1	A_2	A_3	A_4
4	A_2	A_3	A_4	A_1

Nếu x_{ijk} là giá trị ở hàng thứ i, cột thứ j và ở nghiệm thức k; thì số liệu thu thập được từ mô hình có thể được trình bày dưới dạng tổng quát như sau:

	Cột (Động vật)			
Hàng (Giai đoạn)	1	2	3	4
1	$x_{11(3)}$	$x_{12(4)}$	$x_{13(1)}$	$x_{13(2)}$
2	$x_{21(14)}$	$x_{22(1)}$	$x_{23(2)}$	$x_{23(3)}$
3	$x_{31(1)}$	$x_{32(2)}$	$x_{33(3)}$	$x_{33(4)}$
4	$x_{41(2)}$	$x_{42(3)}$	$x_{43(4)}$	$x_{43(1)}$

4.4.3. Mô hình phân tích

$$x_{ijk} = \mu + h_i + c_j + a_k + e_{ijk} \quad (\text{i = 1, a; j = 1, k; k = 1, a})$$

x_{ijk} là quan sát ở hàng thứ i, cột thứ j và ở nghiệm thức k

μ trung bình chung.

h_i chênh lệch do ảnh hưởng của hàng i, $\Sigma h_i = 0$

c_j chênh lệch do ảnh hưởng của cột j, $\Sigma c_j = 0$

a_k chênh lệch do ảnh hưởng của mức k của nhân tố, $\Sigma a_k = 0$

e_{ijk} sai số ngẫu nhiên; giả sử các e_{ijk} độc lập, phân phối chuẩn $N(0,\sigma^2)$

4.4.4. Cách phân tích

Toàn bộ biến động được hợp thành từ các biến động thành phần hàng, cột, nghiệm thức và sai số ngẫu nhiên.

$$SS_{TO} = SS_H + SS_C + SS_A + SS_E$$

với các bậc tự do tương ứng

$$(a^2 - 1) = (a - 1) + (a - 1) + (a - 1) + (a - 2)(a - 1)$$

$$SS_H = a\sum_{i=1}^{a}\left(\bar{x}_i - \bar{x}\right)^2$$

$$SS_C = a\sum_{j=1}^{a}\left(\bar{x}_j - \bar{x}\right)^2$$

$$SS_A = a\sum_{k=1}^{a}\left(\bar{x}_k - \bar{x}\right)^2$$

$$SS_E = a\sum_{i=1}^{a}\sum_{j=1}^{a}\left(\bar{x}_{ij} - \bar{x}_i - \bar{x}_j - \bar{x}_k + 2\bar{x}\right)^2$$

$$SS_{TO} = \sum_{i=1}^{a}\sum_{j=1}^{a}\left(x_{ijk} - \bar{x}\right)^2$$

Bậc tự do $df_{TO} = a^2 - 1$; $df_H = a - 1$; $df_C = a - 1$; $df_A = a - 1$; $df_E = (a-1)(a-2)$

Các trung bình bình phương:

$MS_H = SS_A / df_H$; $MS_C = SS_C / df_C$; $MS_A = SS_A / df_A$; $MS_E = SS_E / df_E$

Giả thiết đối với yếu tố thí nghiệm; H_0 : "Các trung bình của các mức bằng nhau" với đối thiết H_1: "Có ít nhất một cặp trung bình khác nhau".

Tính $F_{TN} = MS_A / MS_E$; so với giá trị tới hạn $F_{(\alpha,\ dfA,\ dfE)}$; nếu $F_{TN} \le F_{(\alpha,\ dfA,\ dfE)}$ thì chấp nhận H_0, ngược lại thì bác bỏ H_0.

Kiểm định đối với hàng và cột thường ít được quan tâm đến vì không mang lại nhiều ý nghĩa, tuy nhiên cũng có thể làm tương tự như kiểm định đối với nghiệm thức.

Có thể tính nhanh các tổng bình phương như sau:

Tính tổng hàng TH_i, tổng cột TC_j, tổng theo từng mức của nhân tố TA_k, sau đó tính $n = a \times a$; Tổng toàn bộ giá trị số liệu trong bảng $ST = \Sigma x_{ij}$ hoặc $ST = \Sigma TH_i$

Tính tổng các giá trị số liệu bình phương $SST = \Sigma\Sigma x^2_{ij}$

Số điều chỉnh $G = ST^2 / n$

Tổng bình phương toàn bộ $SS_{TO} = SST - G$

Tổng bình phương do hàng $SS_H = \Sigma TH_i^2 / a - G$

Tổng bình phương do cột $SS_C = \Sigma TC_j^2 / a - G$

Tổng bình phương do nhân tố $SS_A = \Sigma TA_k^2 / a - G$

Tổng bình phương do sai số $SS_E = SS_{TO} - SS_A - SS_H - SS_C$

Bảng phân tích phương sai (ANOVA)

Nguồn biến động	df	SS	MS	F_{TN}	F tới hạn
Nhân tố	a-1	SS_A	MS_A	MS_A/MS_E	$F_{(\alpha, dfA, dfE)}$
Hàng	a-1	SS_H	MS_H	MS_H/MS_E	$F_{(\alpha, dfH, dfE)}$
Cột	a-1	SS_C	MS_C	MS_C/MS_E	$F_{(\alpha, dfC, dfE)}$
Sai số	(a-1)(a-2)	SS_E	MS_E		
Toàn bộ	a^2 - 1	SS_{TO}			

Ví dụ 4.5: Mead và cộng sự tiến hành nghiên cứu ảnh hưởng của thức ăn mùa đông đến sản lượng sữa theo mô hình ô vuông latinh. Có 4 khẩu phần ăn khác nhau (A_1, A_2, A_3, A_4), 4 giai đoạn thí nghiệm (1, 2, 3 và 4) mỗi giai đoạn kéo dài 3 tuần và có 4 động vật thí nghiệm (1, 2, 3 và 4). Mỗi bò ăn từng khẩu phần trong 3 tuần và mỗi bò tham gia ở cả 4 giai đoạn thí nghiệm. Sản lượng sữa chỉ được tính tổng cộng trong tuần thứ 3 của mỗi giai đoạn. Số liệu được ghi lại như sau (đơn vị tính pound)

		Bò (cột)				
		1	2	3	4	Tổng số
Giai đoạn (hàng)	1	A_1 192	A_2 195	A_3 292	A_4 249	928
	2	A_2 190	A_4 203	A_1 218	A_3 210	821
	3	A_3 214	A_1 139	A_4 245	A_2 163	761
	4	A_4 221	A_3 152	A_2 204	A_1 134	711
Tổng số		817	869	959	756	3221

Ta có bảng phân tích phương sai:

Nguồn biến động	df	SS	MS	F_{TN}	F
Khẩu phần	3	8608,70	2869,20	21,22	$F_{(0,05; 3; 6)} = 4,76$
Giai đoạn	3	6539,20	2179,20	16,12	
Bò	3	9929,20	3309,70	24,47	
Sai số	6	811,40	135,20		
Toàn bộ	15	25887,50			

Kết luận: Ở mức $\alpha = 0,05$ ta bác bỏ giả thiết H_0, tức là các khẩu phần ăn khác nhau đã làm ảnh hưởng đến sản lượng sữa.

Có thể dùng phương pháp LSD để so sánh sự khác nhau giữa từng cặp nghiệm thức như sau:

$$LSD = t(0,025;6)\times\sqrt{\frac{135,20\times 2}{4}} = 20,12$$

Các giá trị trung bình trước và sau khi so sánh:

Khẩu phần	Trung bình		Khẩu phần	Trung bình
A_1	170,80		A_1	$170,80^a$
A_2	188,00	→	A_2	$188,00^a$
A_3	217,00		A_3	$217,00^b$
A_4	229,50		A_4	$229,50^b$

Ngoài 3 kiểu thiết kế thí nghiệm đã nêu trên (Hoàn toàn ngẫu nhiên, Khối ngẫu nhiên đầy đủ và ô vuông La tinh) còn một số kiểu bố trí thí nghiệm một nhân tố phức tạp hơn như:

Khi mỗi khối không chứa đủ các mức của nhân tố (số ô trong một khối nhỏ hơn số mức a) thì có thể bố trí kiểu khối ngẫu nhiên cân đối không đủ (BIBD)

Khi có 3 hướng biến động thì có thể mở rộng kiểu ô vuông La tính thành ô vuông La tinh Hy lạp (Greco Latin square).

Khi bố trí ô vuông La tinh với số nghiệm thức ít thì số bậc tự do còn lại cho sai số ngẫu nhiên nhỏ do đó có thể lặp lại ô vuông La tinh để tăng bậc tự do cho sai số.

Trong các thí nghiệm về giống khi khảo sát ban đầu với số lượng các dòng (giống) quá lớn thì có thể chọn kiểu lưới ô vuông (Lattice design).

4.5. Bài tập

4.5.1: Một thí nghiệm được tiến hành nhằm nghiên cứu ảnh hưởng của 4 công thức thức ăn khác nhau (A, B, C và D) đến tăng trọng của bò BBB. Chọn 24 bò đồng đều và chia hoàn toàn ngẫu nhiên về với các công thức. Khối lượng (kg) khi kết thí nghiệm của 24 bò nêu trên thu được như sau:

A	B	C	D
456	365	502	457
436	400	476	456
432	375	487	467
463	387	499	487
454	408	476	469
453	355	453	432

Kết luận về ảnh hưởng của các công thức thức ăn đến tăng trọng của bò BBB.

4.5.2: Tăng trọng của gà ở 16 công thức thí nghiệm, các công thức khác nhau ở các mức axit amin. Mỗi giá trị trong bảng dưới đây là toàn bộ khối lượng (gram) của 3 gà cùng một lồng trong giai đoạn từ 10 đến 20 ngày tuổi. Có 6 khu chuồng khác nhau, các công thức thí nghiệm được phân về các lồng một cách hoàn toàn ngẫu nhiên trong cùng một khu chuồng có điều kiện tiểu khí hậu ở mức độ đồng đều cao nhất có thể. Kết luận về ảnh hưởng của axit amin đến tăng trọng của gà.

Công thức	Khu chuồng					
	1	2	3	4	5	6
A	125	95	121	92	80	87
B	201	169	152	174	141	128
C	251	216	209	231	226	230
D	332	323	310	317	320	291
E	224	170	176	193	163	153
F	294	290	268	279	274	267
G	206	187	172	180	170	147
H	298	237	281	291	267	184
I	116	101	103	146	94	80
J	135	137	129	138	121	131
K	171	160	156	207	171	144
L	262	277	233	249	213	221
M	165	155	135	165	145	124
N	222	196	184	200	164	167
O	180	156	187	187	162	157
P	247	264	211	247	222	229

4.5.3: Một thí nghiệm được tiến hành nhằm xác định ảnh hưởng của các loại thức ăn bổ sung khác nhau (A, B, C và D) đến lượng cỏ khô mà bê nuôi vỗ béo thu nhận được (kg/ngày). Thí nghiệm được thiết kế theo mô hình ô vuông la tinh với 4 động vật trong 4 giai đoạn, mỗi giai đoạn 20 ngày. Trong mỗi giai đoạn 10 ngày đầu được coi là giai đoạn thích nghi, 10 ngày tiếp theo là giai đoạn thí nghiệm để thu thập số liệu. Số liệu thu được ở bảng bên cạnh là khối lượng cỏ khô trung bình bê thu nhận được ở 10 ngày thí nghiệm. Hãy rút ra kết luận từ thí nghiệm nêu trên.

Giai đoạn	Bê			
	1	2	3	4
1	10,0 (B)	9,0 (D)	11,1 (C)	10,8 (A)
2	10,2 (C)	11,3 (A)	9,5 (D)	11,4 (B)
3	8,5 (D)	11,2 (B)	12,8 (A)	11 (C)
4	11,1 (A)	11,4 (C)	11,7 (B)	9,9 (D)

4.5.4: Giả sử, một thí nghiệm được thiết kế tương tự như ở bài tập 4.5.3, nhưng có có 2 ô vuông la tinh được thiết kế đồng thời và mỗi ô đều có 4 động vật thí nghiệm và 4 công thức thí nghiệm khác nhau. Số liệu ở ô vuông la tinh thứ nhất như trong bài tập 4.5.3, ở ô vuông la tinh thứ 2 như trong bảng bên. Hãy tiến hành phân tích để đưa ra kết luận và đưa ra nhận xét về mô hình thiết kế trong bài tập 4.5.3 và bài tập 4.5.4.

Giai đoạn	Bê			
	1	2	3	4
1	10,9 (C)	11,2 (A)	9,4 (D)	11,2 (B)
2	10,5 (B)	9,6 (D)	11,4 (C)	10,9 (A)
3	11,1 (A)	11,4 (C)	11,7 (B)	9,8 (D)
4	8,8 (D)	12,9 (B)	11,4 (A)	11,2 (C)