M 630
15193

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC NÔNG LÂM THÁI NGUYÊN

ĐỖ THỊ NGỌC OANH (CHỦ BIÊN) - HOÀNG VĂN PHỤ
NGUYỄN THẾ HÙNG - HOÀNG THỊ BÍCH THẢO

Giáo trình
PHƯƠNG PHÁP THÍ NGHIỆM ĐỒNG RUỘNG

NHÀ XUẤT BẢN NÔNG NGHIỆP

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC NÔNG LÂM THÁI NGUYÊN

ĐỖ THỊ NGỌC OANH (Chủ biên) - HOÀNG VĂN PHỤ
NGUYỄN THẾ HÙNG - HOÀNG THỊ BÍCH THẢO

GIÁO TRÌNH
PHƯƠNG PHÁP THÍ NGHIỆM ĐỒNG RUỘNG

ĐẠI HỌC THÁI NGUYÊN
THƯ VIỆN
TRƯỜNG ĐẠI HỌC NÔNG LÂM
PHÒNG MƯỢN
15193

NHÀ XUẤT BẢN NÔNG NGHIỆP
HÀ NỘI - 2004

LỜI NÓI ĐẦU

*Phương pháp thí nghiệm đồng ruộng là hành trang cần thiết cho mọi kỹ sư nông nghiệp và những người làm nghiên cứu trong nông nghiệp. Với sự phát triển nhanh của khoa học kỹ thuật, đặc biệt là sự hỗ trợ của tin học, phương pháp nghiên cứu nông nghiệp có nhiều đổi mới. Để cung cấp cho sinh viên ngành trồng trọt những kiến thức cơ bản và cập nhật thành tựu khoa học mới vào nghiên cứu nông nghiệp chúng tôi biên soạn Giáo trình "**Phương pháp thí nghiệm đồng ruộng**".*

Giáo trình gồm có 4 nội dung chính:

1: Bố trí thí nghiệm - Chương 1 và 2

2: Xử lý số liệu- Chương 3 đến chương 5

3: Ứng dụng tin học trong xử lý và phân tích kết quả - Chương 6

4: Trình bày kết quả nghiên cứu - Chương 7

Phân công biên soạn giáo trình như sau:

Thạc sỹ Đỗ Thị Ngọc Oanh viết chương 1,2 3, 7

Tiến sỹ Hoàng Văn Phụ viết chương 5

Thạc sỹ Nguyễn Thế Hùng viết chương 4

Thạc sỹ Hoàng Thị Bích Thảo viết chương 6

Chúng tôi hy vọng rằng tài liệu sẽ giúp ích cho học tập và nghiên cứu của sinh viên và người làm công tác nghiên cứu khoa học nông nghiệp. Mặc dầu với sự cố gắng cao trong quá trình biên soạn và được nhiều đóng góp ý kiến của các đồng nghiệp, nhưng lần xuất bản đầu tiên này chắc chắn còn có nhiều hạn chế. Chúng tôi rất mong nhận được những đóng góp của bạn đọc để giáo trình ngày càng hoàn thiện hơn.

Các tác giả

BẢNG CHỮ VIẾT TẮT

Bảng chữ viết tắt nhằm giúp cho việc tra cứu và thống nhất giữa các thuật ngữ tiếng Việt và tiếng Anh. Sắp xếp theo a,b,c...

Viết tắt	Tiếng Anh	Tiếng Việt
μ, $\overline{X}$	Treatment mean, sample mean	Trung bình tổng thể, trung bình mẫu
δ, **s**	Standard deviation	Độ lệch chuẩn tổng thể, độ lệch chuẩn mẫu
ANOVA	Analysis of variance	Phân tích biến động
CF	Correction factor	Hệ số hiệu chỉnh
CI	Confidence interval	Khoảng tin cậy
CRD	Completed randomized design	Ngẫu nhiên hoàn chỉnh
CV	Coefficient variance	Hệ số biến động
df	Degree of freedom	Độ tự do
DMRT	Duncan's multiple range test	So sánh theo Dun can
F_t; Fb	Compute F; tabular F	F tính và F lý thuyết
G	Grand total	Tổng lớn
LS	Latin square	Ô vuông La tinh
LSD	Least significant difference	Sai khác nhỏ nhất có ý nghĩa
MS	Mean square	Trung bình tổng bình phương
ns	Non-significant	Không có ý nghĩa
r	- Coefficient of correlation - Replication	Hệ số tương quan Nhắc lại
R	Multiple correlation coefficient	Hệ số tương quan đa biến
R^2	Coeficience of determination	Hệ số tương quan xác định
RCBD	Randomized completed block design	Khối ngẫu nhiên hoàn chỉnh
SAS	Statistics Analysis System	Phần mềm SAS
t	Treatment	Công thức

MỤC LỤC

Lời nói đầu 3
Bảng chữ viết tắt 4

Chương 1. PHƯƠNG PHÁP NGHIÊN CỨU TRONG NÔNG NGHIỆP
1.1. Những vấn đề chung về Phương pháp nghiên cứu khoa học 9
1.1.1. Xác định vấn đề nghiên cứu 9
1.1.2. Xác định mục tiêu và mục đích nghiên cứu 10
1.1.3. Nghiên cứu tổng quan tài liệu 11
1.1.4. Xây dựng giả thuyết 11
1.1.5. Làm thí nghiệm và phân tích kết quả 12
1.1.6. Tổng hợp, kết luận và khuyến cáo 12
1.2. Các bước làm thí nghiệm 12
1.2.1. Xác định vấn đề nghiên cứu và mục đích thí nghiệm 13
1.2.2. Lựa chọn công thức thí nghiệm và thiết kế thí nghiệm 14
1.2.3. Chọn biến phụ thuộc 14
1.2.4. Thực hiện thí nghiệm 15
1.2.5. Phân tích kết quả 15
1.2.6. Kết luận và khuyến cáo 16
1.3. Nguyên tắc cơ bản của làm thí nghiệm 16
1.3.1. Nhắc lại 16
1.3.2. Ngẫu nhiên 17
1.3.3. Sai khác duy nhất 17
1.4. Phân tích sự sai khác trong kết quả thí nghiệm (phân tích biến động) 18
1.5. Nâng cao độ chính xác của thí nghiệm đồng ruộng 18
1.5.1. Tăng số lần nhắc lại 18
1.5.2. Kỹ thuật trong khi làm thí nghiệm 19
1.5.3. Kỹ thuật thiết kế thí nghiệm đồng ruộng 19
1.6. Thí nghiệm một nhân tố và thí nghiệm đa nhân tố 19
1.6.1. Thí nghiệm một nhân tố 19
1.6.2. Thí nghiệm đa nhân tố 20

1.7. Tổng thể và mẫu 21
1.8. Một số khái niệm trong phân tích thống kê 21
1.8.1. Số định tâm (Measure of central tendency) 21
1.8.2. Số phân tán (Measure of dispersion) 22

Chương 2. BỐ TRÍ THÍ NGHIỆM ĐỒNG RUỘNG VÀ THU THẬP SỐ LIỆU
2.1. Bố trí thí nghiệm đồng ruộng 25
2.1.1. Chia khối 25
2.1.2. Một số lưu ý khi bố trí thí nghiệm trên đồng ruộng 26
2.2. Bố trí thí nghiệm một nhân tố 26
2.2.1. Kiểu ngẫu nhiên hoàn toàn 26
2.2.2. Kiểu khối ngẫu nhiên hoàn chỉnh 30
2.2.3. Kiểu ô vuông la tinh 32
2.3. Bố trí thí nghiệm 2 nhân tố 35
2.3.1. Bố trí kiểu khối ngẫu nhiên hoàn chỉnh 35
2.3.2. Kiểu ô chính - ô phụ (Split-Plot Design) 37
2.3.3. Kiểu lô sọc (Strip plot design) (Phần đọc thêm) 40
2.4. Thu thập số liệu 41
2.4.1. Chọn mẫu 41
2.4.2. Đo đếm 42
2.4.3. Ghi chép số liệu: Số liệu thô và số liệu tinh 42
2.4.4. Một số quy tắc cần biết trong tính toán 43

Chương 3. PHÂN TÍCH BIẾN ĐỘNG
3.1. Ý nghĩa của phân tích biến động và hệ số biến động 44
3.1.1. Phân tích biến động 44
3.1.2. Hệ số biến động 45
3.2. Phân tích biến động cho thí nghiệm một nhân tố 46
3.2.1. Kiểu bố trí ngẫu nhiên hoàn toàn 46
3.2.2. Kiểu khối ngẫu nhiên hoàn chỉnh 49
3.2.3. Kiểu ô vuông la tinh 53
3.3. Tương tác trong thí nghiệm hai nhân tố 57
3.4. Phân tích biến động cho thí nghiệm hai nhân tố 61
3.4.1. Kiểu khối ngẫu nhiên hoàn chỉnh 61
3.4.2. Kiểu ô chính- ô phụ (split-plot design) 64

Chương 4. SO SÁNH SỐ TRUNG BÌNH

4.1. Ý nghĩa của so sánh số trung bình 72

4.2. So sánh cặp đôi 72

4.2.1. So sánh theo sai khác nhỏ nhất 73

4.2.2. So sánh Duncan: So sánh phân hạng dựa trên giá trị LSD 80

4.3. So sánh nhóm (Contrasts) 83

4.3.1. ý nghĩa của so sánh nhóm 83

4.3.2. Xác định hệ số so sánh trong so sánh nhóm 83

4.3.3. Cách so sánh nhóm và xu hướng 85

Chương 5. PHÂN TÍCH TƯƠNG QUAN

5.1. Khái niệm về tương quan và phân tích tương quan 88

5.2.1. Khái niệm về tương quan 88

5.1.2. Phân tích tương quan 90

5.2. Phân tích tương quan tuyến tính đơn giản 91

5.2.1. Tương quan tuyến tính đơn giản (simple linear regression) 91

5.2.2. Phân tích tương quan tuyến tính đơn giản 92

5.3. Tương quan tuyến tính đa biến 97

5.4. Tương quan phi tuyến tính đơn biến 100

5.5. Tương quan phi tuyến tính đa biến 101

5.6. Những sai lầm thường mắc khi phân tích tương quan 102

Chương 6. ỨNG DỤNG TIN HỌC TRONG XỬ LÝ VÀ TRÌNH BÀY KẾT QUẢ

6.1. Giới thiệu về phần mềm SAS 104

6.1.1. Ưu điểm của phần mềm SAS 104

6.1.2. Hướng dẫn sử dụng SAS 105

6.1.3. Cách nhập số liệu vào SAS 108

6.1.4. Một số thuật ngữ chính được sử dụng trong sử dụng SAS 109

6.1.5. Thực hành 110

6.2. Sử dụng SAS trong phân tích phương sai (ANOVA) 111

6.2.1. Các trình thủ tục để phân tích phương sai trong SAS 111

6.2.2. Cách đọc kết quả bảng phân tích phương sai của SAS 112

6.2.3. Phân tích phương sai cho thí nghiệm một nhân tố 113

6.2.4. Phân tích phương sai cho thí nghiệm 2 nhân tố 124

6.3. Sử dụng phần mềm SAS so sánh số trung bình 136
6.3.1. So sánh cặp đôi - LSD và DUNCAN 136
6.3.2. So sánh nhóm và so sánh theo xu hướng (Contrasts) 142
6.4. Phân tích tương quan và hồi quy tuyến tính 144
6.4.1. Trình thủ tục CORR 144
6.4.2. Trình thủ tục REG 145
6.5. Sử dụng Excel vẽ đồ thị và hàm tương quan 147
6.5.1. Sử dụng Excel vẽ đồ thị 147
6.5.2. Sử dụng Excel vẽ hàm tương quan 148

Chương 7. TRÌNH BÀY KẾT QUẢ NGHIÊN CỨU

7.1. Trình bày số liệu 150
7.1.1. Phương pháp lập bảng số liệu 150
7.1.2. Phương pháp vẽ đồ thị 151
7.1.3 . Trình bày kết quả phân tích thống kê 151
7.2. Bố cục và nội dung luận văn khoa học 151
7.3. Bảo vệ luận văn 156
7.3.1. Đề cương báo cáo bảo vệ luân văn 157
7.3.2. Thuyết trình báo cáo khoa học 157
Tài liệu tham khảo 158
Phụ lục: Các bảng số 159

Chương 1

PHƯƠNG PHÁP NGHIÊN CỨU TRONG NÔNG NGHIỆP

1.1. NHỮNG VẤN ĐỀ CHUNG VỀ PHƯƠNG PHÁP NGHIÊN CỨU KHOA HỌC

Khoa học là một nỗ lực nhằm phát hiện ra những quy luật của tự nhiên và vận dụng những kiến thức để mô tả, giải thích và tiên đoán những điều có thể xảy ra trong thiên nhiên. Mục đích của nghiên cứu khoa học là khám phá thế giới và cải tạo thế giới. Các nhà khoa học không phải là những người đi tìm những bằng chứng, những sự thật bất di bất dịch. Các quy luật, các giả thuyết, mô hình và các thuyết khoa học đều xây dựng trên cơ sở những điều mang tính chất thống kê chứ không phải xây dựng trên những giả thuyết cố định. Và tất cả những học thuyết đó luôn luôn được kiểm chứng và luôn bị thách thức bởi những phát kiến mới (Miller, 1996).

Theo Phương Kỳ Sơn (2001) "Phương pháp nghiên cứu là tổ hợp cách thức mà nhà khoa học sử dụng để tác động, khám phá đối tượng". Nghiên cứu khoa học là hoạt động phức tạp, chứa nhiều mâu thuẫn, nhiều trường phái, nhiều xu hướng. Quan điểm phù hợp với hiện thực, đem lại lợi ích cho cuộc sống con người sẽ tồn tại. Nghiên cứu khoa học mang tính mạo hiểm, nghiên cứu có thể thành công hoặc thất bại. "Sự thành công cho giá trị mới, sự thất bại không phải là tổn thất mà là sự trả giá của khoa học" (Phương Kỳ Sơn, 2001). Có một số loại hình nghiên cứu khoa học là: (1) Nghiên cứu cơ bản nhằm khám phá các quy luật trong tự nhiên; (2) Nghiên cứu ứng dụng là nghiên cứu dựa trên kết quả của nghiên cứu cơ bản để ứng dụng vào một ngành khoa học cụ thể; (3) Nghiên cứu triển khai - thực nghiệm nhằm áp dụng kết quả mà nghiên cứu ứng dụng đã đạt được vào thực tế sản xuất; (4) Nghiên cứu thăm dò là dạng nghiên cứu đặc biệt dựa trên cơ sở vận dụng các kết quả của nghiên cứu cơ bản hay ứng dụng để xác định hướng nghiên cứu; (5) Nghiên cứu mô tả nhằm mô tả, khám phá, phân tích, tổng hợp tài liệu thực tế và kết luận.

Quá trình nghiên cứu khoa học bắt đầu từ quan sát thực tế. Những vấn đề được giải thích bằng những giả thuyết đưa ra dựa trên kinh nghiệm. Tuy nhiên, giả thuyết chỉ có giá trị khi được kiểm chứng bằng thí nghiệm. Căn cứ vào kết quả thí nghiệm sẽ đưa ra những giải pháp để giải quyết vấn đề. Dưới đây là các bước của quá trình nghiên cứu khoa học.

1.1.1. Xác định vấn đề nghiên cứu

Xác định được vấn đề nghiên cứu là việc đầu tiên của quá trình nghiên cứu. Vấn đề nghiên cứu quyết định đến mục tiêu, phương pháp và nội dung nghiên cứu. Khi xác định vấn đề nghiên cứu cần tập trung các câu trả lời sau:

- Đề tài liên quan đến lĩnh vực nào?

- Vấn đề mắc phải là gì? Lẽ ra sự việc nên như thế nào?

- Trong các vấn đề đó, lĩnh vực nào chưa được đề cập đến?

Khi xác định vấn đề nghiên cứu cần làm rõ đối tượng nghiên cứu. Đối tượng nghiên cứu là bản chất sự vật hay hiện tượng cần xem xét và làm rõ trong quá trình nghiên cứu. Đồng thời cần xác định phạm vi nghiên cứu. Chúng ta không thể cùng một lúc nghiên cứu đối tượng một cách toàn diện mọi khía cạnh, trong mọi thời gian và địa điểm mà cần phải giới hạn trong một phạm vi nhất định. Phạm vi bao gồm quy mô đối tượng, không gian và thời gian. Lý do của đưa ra phạm vi còn do hạn chế về nguồn lực của người nghiên cứu.

Có nhiều cách để xác định và lựa chọn vấn đề nghiên cứu. Nhà nghiên cứu nên chọn những "tình huống có vấn đề" thuộc lĩnh vực mà mình quan tâm nhất. Dưới đây là những nguồn chỉ dẫn để có thể tìm ra những định hướng nghiên cứu:

1. Từ các báo cáo nghiên cứu khoa học: Thông thường các bài báo nghiên cứu đăng trong các tạp chí và các luận án có một phần được dùng để trình bầy về những hạn chế của kết quả nghiên cứu. Căn cứ vào những hạn chế của các nghiên cứu này, người ta có thể xác định được những chủ đề hoặc lĩnh vực cần phải nghiên cứu sâu hơn.

2. Các bài báo: Các chủ đề nghiên cứu cũng có thể được hình thành dựa trên cơ sở các bài báo.

3. Quan sát: Nhà nghiên cứu có thể xác định vấn đề nghiên cứu từ việc xem xét tình hình sản xuất thực tế.

4. Tham khảo ý kiến những nhà khoa học, chuyên gia.

Theo Phương Kỳ Sơn (2001) việc lựa chọn vấn đề nghiên cứu được xem xét theo các cấp độ sau: (1) Ý nghĩa khoa học, điều này thể hiện trên khía cạnh bổ sung hay làm rõ một số vấn đề lý thuyết đang tồn tại; (2) Ý nghĩa thực tiễn, thể hiện ở các mặt: xây dựng luận cứ cho các chương trình phát triển kinh tế xã hội. Giải đáp nhu cầu của sản xuất. (3) Tính cấp thiết, thể hiện mức độ ưu tiên giải đáp những nhu cầu lý thuyết và thực tiễn đã được xem xét. (4) Điều kiện đảm bảo cho việc hoàn thành nghiên cứu.

1.1.2. Xác định mục tiêu và mục đích nghiên cứu

Mục tiêu và mục đích là hai khái niệm quan trọng trong nghiên cứu khoa học. Mục tiêu là cái đích về mặt nội dung mà người nghiên cứu vạch ra để thực hiện. Mục tiêu là những điều cần làm rõ trong công việc nghiên cứu và là ý nghĩa khoa học của nghiên cứu. Câu hỏi đặt ra đối với mục tiêu là "làm cái gì". Mục đích là ý nghĩa thực tiễn của nghiên cứu, là đối tượng phục vụ của sản phẩm nghiên cứu. Câu hỏi đặt ra đối với mục đích là "để phục vụ cho cái gì" (Phương Kỳ Sơn, 2001).

Ví dụ với mục tiêu nghiên cứu là nâng cao năng suất lúa người ta có thể làm rất nhiều vấn đề, nhưng với mục đích nghiên cứu là "xác định mức phân đạm bón thích

hợp" thì đối tượng nghiên cứu cụ thể là mức phân đạm bón. Xác định mức phân đạm bón thích hợp chính là nhằm mục tiêu nâng cao năng suất lúa.

1.1.3. Nghiên cứu tổng quan tài liệu

1.1.3.1. Sự cần thiết của nghiên cứu tổng quan

Nghiên cứu tổng quan là xem xét, đánh giá các nghiên cứu trước có liên quan đến vấn đề, phương pháp, mô hình và các đề xuất của đề tài đang nghiên cứu. Mục đích của tổng quan tài liệu là giúp nhà nghiên cứu có được bức tranh tổng thể về vấn đề đang quan tâm nghiên cứu. Nghiên cứu tổng quan thiết lập cơ sở lý luận và thực tiễn cho vấn đề sẽ nghiên cứu. Nhờ nghiên cứu tổng quan người nghiên cứu rút ngắn được quãng đường tổ chức nghiên cứu, vì vậy tránh được các bước trùng lặp trong nghiên cứu và kết quả là tiết kiệm được nguồn lực trong tổ chức nghiên cứu. Cũng qua tổng quan tài liệu giúp cho nhà nghiên cứu chọn ra được phương pháp tốt nhất cho nghiên cứu của mình. Tổng quan sẽ giúp nhiều cho thảo luận kết quả nghiên cứu được và giúp độc giả so sánh, đánh giá được giá trị của kết quả nghiên cứu. Ngoài ra tổng quan tài liệu cũng tạo ra một diễn đàn thảo luận về những xu hướng hay phương pháp nghiên cứu tiếp theo.

Từ nghiên cứu tổng quan kết hợp với thực tế, nhà nghiên cứu phân tích và làm rõ vấn đề. Đó là cơ sở để xây dựng giả thuyết.

1.1.3.2. Cách nghiên cứu tổng quan

Xây dựng đề cương cho nghiên cứu tổng quan

Trước hết cần phải xác định rõ nội dung của nghiên cứu tổng quan. Người nghiên cứu dựa trên mục tiêu nghiên cứu để thiết lập đề cương cho nghiên cứu tổng quan. Đề cương này được xây dựng phù hợp với các nội dung sẽ tiến hành nghiên cứu. Đề cương sẽ bao gồm các khái niệm, các vấn đề sẽ bàn trong nghiên cứu mới, phương pháp sẽ áp dụng...

Thu thập các tài liệu có liên quan đến nội dung và phương pháp nghiên cứu.

Dựa trên đề cương viết tổng quan để thu thập các tài liệu, các nghiên cứu ở các nguồn thông tin khác nhau như sách, báo, tạp chí. Cách thông thường nhất là copy các nghiên cứu đó, thu đủ các thông tin về tên tạp chí, số và năm xuất bản, tác giả, số thứ tự của trang. Tốt nhất mỗi một nội dung liên quan nên bố trí thành thư mục riêng để tránh nhầm lẫn. Nên viết tóm tắt các ý chính của bài ngay từ đầu thư mục

1.1.4. Xây dựng giả thuyết

Trong quá trình nghiên cứu, sau khi phân tích các tài liệu thu được do quan sát, khảo sát, sưu tầm và hệ thông hoá tài liệu, ta đề xuất xây dựng một giả thuyết. Giả thuyết là điều được đặt ra một cách giả định về mối quan hệ giữa các chỉ tiêu nghiên cứu. Các giả thuyết phải cụ thể, có thể kiểm tra được bằng thí nghiệm.

Ví dụ, khi ta làm thí nghiệm so sánh một giống lúa mới lai tạo với giống Bao Thai đang được dùng đại trà trong sản xuất, sẽ có 2 giả thuyết đặt ra, đó là giả thuyết

Ho: Năng suất của 2 giống lúa như nhau; Giả thuyết H1 là năng suất của 2 giống lúa khác nhau. Để chứng minh giả thuyết phải làm thí nghiệm.

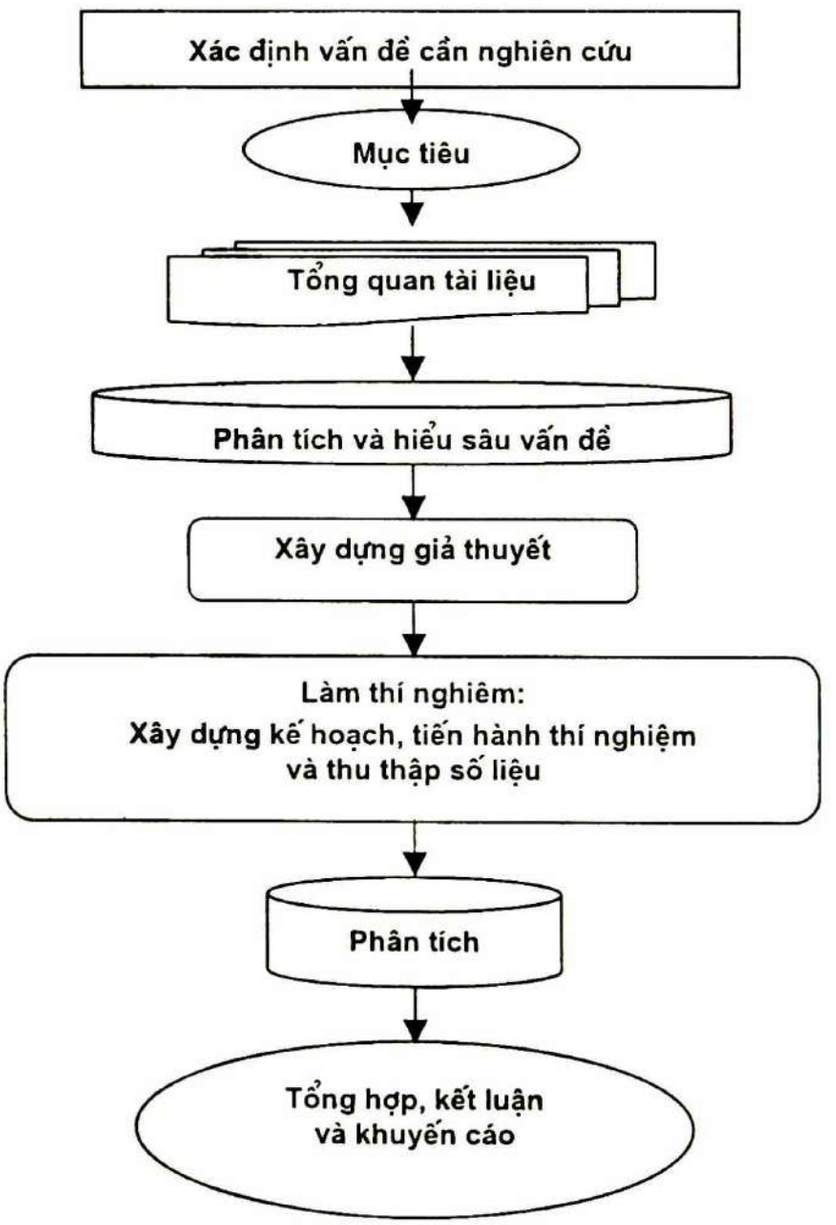
1.1.5. Làm thí nghiệm và phân tích kết quả

Làm thí nghiệm và phân tích kết quả để kiểm chứng giả thuyết đã nêu (sẽ được đề cập chi tiết ở phần sau).

1.1.6. Tổng hợp, kết luận và khuyến cáo

Việc cuối cùng của quá trình nghiên cứu là tổng hợp, viết báo cáo trình bày và kết luận về vấn đề nghiên cứu. Căn cứ vào mục đích thí nghiệm để đưa ra kết luận về vấn đề nghiên cứu và từ đó đưa ra những khuyến cáo. Khuyến cáo cần nêu rõ kết quả nghiên cứu có áp dụng được vào thực tế không hay cần tiếp tục nghiên cứu.

Có thể tóm tắt các bước của nghiên cứu khoa học như sau:



Sơ đồ 1.1. Các bước của nghiên cứu khoa học

Những vấn đề cần giải quyết trong mỗi bước của nghiên cứu khoa học được tóm tắt ở bảng sau:

Bảng 1.1. Những vấn đề cần giải quyết trong các bước của nghiên cứu khoa học

Các bước của nghiên cứu khoa học	Vấn đề cần giải quyết/câu hỏi cần trả lời
Xác định vấn đề cần nghiên cứu	Thực tế có vấn đề gì cần giải quyết?
Xác định mục tiêu nghiên cứu	Nghiên cứu để làm gì?
Tổng quan tài liệu	Những kết quả đã có về vấn đề nghiên cứu? Điều gì cần phải làm rõ hơn?
Phân tích và hiểu sâu vấn đề	Vấn đề nghiên cứu đã được giải quyết đến đâu? Nên tiếp tục như thế nào?
Xây dựng giả thuyết	Đưa ra giả thuyết để giải thích vấn đề.
Làm thí nghiệm	Tạo và thu thập số liệu để kiểm tra giả thuyết đã nêu
Phân tích kết quả thí nghiệm	Dùng số liệu thí nghiệm để chứng minh làm rõ vấn đề: chấp nhận hay bác bỏ giả thuyết
Tổng hợp, kết luận và khuyến cáo	Viết báo cáo trình bày về vấn đề nghiên cứu và rút ra kết luận. Thực tế có áp dụng được không hay phải nghiên cứu nữa?

1.2. CÁC BƯỚC LÀM THÍ NGHIỆM

Từ sơ đồ 1.1. cho thấy rõ làm thí nghiệm là một bước của quá trình nghiên cứu khoa học nhằm kiểm tra giả thuyết.

Có thể định nghĩa về thí nghiệm như sau: Thí nghiệm là một quá trình tạo ra số liệu. Hoặc, thí nghiệm là quá trình theo dõi một hay nhiều chỉ tiêu dưới tác động của một hay nhiều nhân tố nhằm phát hiện, hay kiểm định một hay một số vấn đề nào đó. Ví dụ, có giả thuyết là mức phân đạm bón cho lúa hiện nay chưa thoả mãn yêu cầu của cây lúa, nếu bón tăng sẽ đạt năng suất cao hơn. Vì vậy, nhà nông học làm thí nghiệm bón các mức đạm khác nhau cho lúa. Để kiểm định giả thuyết trên, quá trình làm thí nghiệm cần theo dõi sự thay đổi về sinh trưởng, phát triển của cây lúa và năng suất lúa ở các mức phân khác nhau. Có 6 bước cơ bản của làm thí nghiệm như sau:

Bảng 1.2. Các bước làm thí nghiệm

Các bước làm thí nghiệm	Vấn đề cần giải quyết
Xác định vấn đề nghiên cứu và mục đích làm thí nghiệm.	Trả lời câu hỏi: thực tế có vấn đề gì? Nghiên cứu để làm gì?
Lựa chọn công thức thí nghiệm và thiết kế thí nghiệm.	Thí nghiệm về cái gì? Bao nhiêu công thức? Cách bố trí trên đồng ruộng?.
Chọn biến phụ thuộc (xác định chỉ tiêu theo dõi)	Cần thu thập những số liệu gì để chứng minh cho giả thuyết?
Thực hiện thí nghiệm	Tạo và thu thập số liệu
Phân tích kết quả	Xử lý số liệu và phân tích số liệu.
Kết luận và khuyến cáo	Giả thuyết được chấp nhận hay bác bỏ, nên áp dụng trong thực tế như thế nào?

1.2.1. Xác định vấn đề nghiên cứu và mục đích thí nghiệm

Vấn đề nghiên cứu phải xuất phát từ thực tiễn sản xuất và đời sống. Khi xác định vấn đề nghiên cứu cần phải xác định:

Nghiên cứu nhằm giải quyết vấn đề gì?

Những chỉ tiêu cơ bản của vấn đề nghiên cứu là gì?

Cần phải làm rõ lý do của thí nghiệm, đó chính là mục đích thí nghiệm. Kết quả mong đợi của người nghiên cứu là khám phá cái mới, đánh giá, so sánh và phân tích. Mục đích nghiên cứu thường có những đặc điểm sau:

Thu được những thông tin về đối tượng nghiên cứu

Tìm ra các giải pháp giải quyết vấn đề cho phù hợp

Hoàn thiện thêm về phương pháp hay mô hình cần cho những nghiên cứu tiếp theo.

Khi xác định vấn đề nghiên cứu và mục đích thí nghiệm cần xác định phạm vi nghiên cứu. Nơi tiến hành thí nghiệm. Cần biết rõ nơi đó có đại diện cho các vùng khác không? Tính áp dụng của mô hình sau khi nghiên cứu kết thúc như thế nào?

1.2.2. Lựa chọn công thức thí nghiệm và thiết kế thí nghiệm

Lựa chọn công thức thí nghiệm bao gồm lựa chọn nhân tố thí nghiệm và mức độ của nhân tố thí nghiệm. Nhân tố thí nghiệm là nhân tố mà theo lý thuyết nó giải quyết được vấn đề cần nghiên cứu. Ví dụ: theo lý thuyết, nguyên nhân của năng suất lúa thấp trên đất lầy thụt là thiếu lân. Vậy để giải quyết vấn đề này cần phải làm thí nghiệm bón lân cho lúa trên đất lầy thụt. Trong đó, nhân tố thí nghiệm là lân. Việc tiếp theo là xác định nên bón lân ở mức độ nào, 20 P_2O_5/ha hay 50 P_2O_5/ha. Công việc đó được gọi là lựa chọn mức độ của nhân tố thí nghiệm.

Mỗi mức độ của nhân tố thí nghiệm được gọi là một công thức thí nghiệm. Như vậy, có thể hiểu công thức thí nghiệm là một kỹ thuật mà ta cần xác định ảnh hưởng của nó và so sánh với kỹ thuật khác. Ở ví dụ trên ta có thể có 3 công thức thí nghiệm là bón lân ở mức là 0, 20 và 50 P_2O_5/ ha. Căn cứ để xác định công thức thí nghiệm là kết quả nghiên cứu tổng quan, kinh nghiệm của người nghiên cứu và thực tiễn sản xuất.

Khi lựa chọn các mức độ của nhân tố thí nghiệm người ta thường chọn thêm một công thức là kỹ thuật đang được dùng phổ biến trong sản xuất. Công thức đó được gọi là công thức đối chứng. Một loại công thức đối chứng cũng thường được sử dụng là công thức 0 của nhân tố thí nghiệm như công thức 0 P_2O_5/ha ở ví dụ trên. Mục đích của việc đưa công thức đối chứng vào trong thí nghiệm là để dễ so sánh và có tính thuyết phục. Loại công thức đối chứng 0 thường được sử dụng để tính toán hiệu quả của kỹ thuật được thí nghiệm.

Khi xác định được các công thức thí nghiệm cần quyết định địa điểm thí nghiệm, thời gian thí nghiệm, vật liệu thí nghiệm và cách bố trí thí nghiệm trên đồng

ruộng. Công việc đó gọi là thiết kế thí nghiệm. Thiết kế của thí nghiệm được căn cứ vào mục tiêu, mục đích nghiên cứu và đối tượng thí nghiệm.

Bố trí thí nghiệm bao gồm việc quyết định số lần nhắc lại, kích thước ô và kiểu bố trí thí nghiệm. Căn cứ cho việc bố trí thí nghiệm là đối tượng nghiên cứu, nơi nghiên cứu và loại hình nghiên cứu.

1.2.3. Chọn biến phụ thuộc

Biến phụ thuộc là những đặc điểm của đối tượng nghiên cứu mà theo giả thuyết chúng sẽ thay đổi ứng với mỗi mức độ của nhân tố thí nghiệm. Những đặc điểm đó có thể là định tính hoặc định lượng.

Chọn biến phụ thuộc là chọn những đặc điểm cần quan sát của đối tượng nghiên cứu trong quá trình thí nghiệm để có những số liệu chứng minh cho giả thuyết của thí nghiệm.

Biến thường được chọn trong thí nghiệm đồng ruộng là năng suất cây trồng. Nhưng để giải thích về sự thay đổi năng suất thì cần phải có những số liệu quan sát về cây trồng, đất và điều kiện ngoại cảnh khác để trả lời các câu hỏi như:

Thực vật phản ứng với điều kiện ngoại cảnh và các biện pháp kỹ thuật như thế nào? Ví dụ, nền phân bón khác nhau dẫn đến tỷ lệ bệnh khác nhau không?

Nguyên nhân của sự khác nhau về năng suất giữa các công thức là gì? Ví dụ, năng suất các giống lúa khác nhau do số nhánh đẻ khác nhau hay khối lượng hạt khác nhau?

Mục tiêu và các giả thuyết nghiên cứu là những chỉ dẫn trong việc quyết định cần phải thu thập những loại số liệu nào. Căn cứ vào mục đích thí nghiệm người nghiên cứu quyết định phương pháp thu thập số liệu.

Với ví dụ về bón lân cho lúa trên đất lầy thụt ở trên thì biến phụ thuộc có thể là năng suất lúa và khả năng chống đổ của cây.

1.2.4. Thực hiện thí nghiệm

Thực hiện thí nghiệm bao gồm việc bố trí thí nghiệm trên đồng ruộng theo thiết kế và thu thập số liệu.

Thu thập số liệu là việc theo dõi và ghi chép sự thay đổi của những biến phụ thuộc trong quá trình thí nghiệm. Ngoài ra còn cần ghi chép mọi sự thay đổi có liên quan đến thí nghiệm như diễn biến thời tiết...

1.2.5. Phân tích kết quả

Phân tích kết quả bao gồm việc xử lý số liệu và phân tích số liệu. Những số liệu thu thập được trong quá trình theo dõi thí nghiệm được gọi là số liệu thô. Số liệu thô tự nó không nêu bật lên được bản chất hay tính quy luật của hiện tượng. Tuỳ theo mục đích nghiên cứu và đặc trưng của số liệu, nhà nghiên cứu cần sắp xếp, xử lý số liệu theo phương pháp phù hợp.

Xử lý số liệu là chuyển hoá, sắp xếp số liệu thô phù hợp cho việc phân tích, trình bày kết quả. Nó bao gồm cả việc kiểm tra tính chính xác của số liệu. Trong nghiên cứu nông nghiệp, kết luận khái quát được đưa ra dựa trên kết quả quan sát cụ thể thu được. Vì vậy phương pháp xử lý số liệu được dùng trong thí nghiệm đồng ruộng là phương pháp thống kê.

Phương pháp thống kê là phương pháp phân tích vấn đề trên cơ sở ngẫu nhiên. Cơ sở của nó là phương pháp quy nạp. Nói cách khác, phương pháp thống kê là rút ra kết luận chung cho tổng thể từ những mẫu được lấy một cách ngẫu nhiên đại diện cho tổng thể. Ví dụ, khi điều tra chiều cao của học sinh tiểu học người ta không thể đo tất cả học sinh mà chỉ chọn ngẫu nhiên một số để đo và lấy kết quả trung bình đo được làm kết quả.

Phân tích kết quả nhằm chứng minh giả thuyết. Vì vậy kỹ thuật trình bày và phân tích kết quả rất quan trọng nhằm làm rõ vấn đề và thể hiện năng lực của người nghiên cứu. Tuỳ theo mục đích nghiên cứu, đặc tính của số liệu để chọn phương pháp phân tích phù hợp. Thường có hai nhóm phân tích là: phương pháp thống kê mô tả (tóm tắt và mô tả đặc tính của số liệu dưới dạng bảng nhiều chiều; Sử dụng các thông số thống kê như số trung bình, sai khác nhỏ nhất...). Phương pháp ước lượng mô hình (phân tích hồi quy).

Yêu cầu của trình bầy kết quả thu được là phải phù hợp với mục tiêu và nội dung nghiên cứu. Khi thảo luận có thể kết hợp so sánh với những kết quả nghiên cứu liên quan để làm rõ thêm về nhận định kết quả nghiên cứu.

1.2.6. Kết luận và khuyến cáo

Từ kết quả thu được, qua phân tích đưa ra kết luận và khuyến cáo về vấn đề nghiên cứu. Kết luận phải dựa trên mục tiêu nghiên cứu. Kết luận chỉ ra liệu kết quả tìm được có khẳng định hay bác bỏ với giả thuyết nghiên cứu. Từ kết quả đưa ra những khuyến cáo cho sản xuất.

1.3. NGUYÊN TẮC CƠ BẢN CỦA LÀM THÍ NGHIỆM

Đối tượng và địa bàn nghiên cứu của thí nghiệm đồng ruộng phụ thuộc rất nhiều vào điều kiện tự nhiên. Sự thay đổi của những điều kiện đó nằm ngoài sự kiểm soát của con người nhưng lại ảnh hưởng đến kết quả thí nghiệm. Do vậy nguyên tắc bố trí thí nghiệm đồng ruộng là làm sao hạn chế ảnh hưởng của những yếu tố tự nhiên đó. Ba nguyên tắc cơ bản cần tuân thủ trong thí nghiệm đồng ruộng là: Nhắc lại, ngẫu nhiên và sai khác duy nhất.

1.3.1. Nhắc lại

Nhắc lại nghĩa là công thức thí nghiệm được lặp lại trong một thí nghiệm. Với thí nghiệm bón lân cho lúa trên đất lầy thụt nêu trên, nếu số lần nhắc lại là 3 có nghĩa là các công thức 0, 20 và 50 P_2O_5 phải được lặp lại 3 lần. Số lần nhắc lại tuỳ theo yêu cầu của thí nghiệm, thường nhắc lại ít nhất từ 2 lần trở lên. Thí nghiệm yêu cầu độ chính xác cao thì cần nhắc lại nhiều hơn thí nghiệm yêu cầu độ chính xác thấp hơn. Thí nghiệm khảo nghiệm giống với số lượng nhiều có thể chỉ cần làm 2 lần nhắc lại.

Tại sao phải làm nhắc lại. Ví dụ sau đây sẽ minh họa cho vấn đề này. Phương pháp nghiên cứu khối lượng búp chè (g/búp) của Viện nghiên cứu chè (1998) quy định chọn 3 điểm đại diện, mỗi điểm lấy 100 g búp, đếm số búp trong 100 g để tính khối lượng búp chè. Kết quả cuối cùng là trung bình của 3 lần cân trên. Tuy cùng một giống chè, một vườn chè (cùng chế độ chăm sóc), do mỗi lần cân 100 g búp khác nhau nên kết quả khối lượng 1 búp chè của 3 lần cân khác nhau. Sự khác nhau đó gọi là là sai số thí nghiệm. Sai số này là do 3 điểm lấy búp khác nhau chứ không phải do giống khác nhau hay vườn chè khác nhau. Sai số này cần phải được ước lượng để làm cơ sở cho việc xác định sự khác nhau trong số liệu quan sát được có thực sự là do yếu tố thí nghiệm không. Sai số thí nghiệm được ước lượng từ kết quả của nhắc lại.

Như vậy, vai trò quan trọng của nhắc lại là để xác định sai số thí nghiệm. Ngoài ra nhắc lại còn tăng độ chính xác cho thí nghiệm, đặc biệt là với loại vật liệu thí nghiệm kém đồng đều.

Do có nhắc lại nên số ô thí nghiệm nhiều hơn số công thức thí nghiệm. Nếu thí nghiệm có 2 công thức, với 3 lần nhắc lại cần phải có số ô thí nghiệm n = 2 x 3 = 6 ô.

Ô thí nghiệm: là một đơn vị của thí nghiệm mà trên đó bố trí các công thức thí nghiệm. Diện tích của các ô trong một thí nghiệm phải bằng nhau.

1.3.2. Ngẫu nhiên

Để áp dụng phương pháp phân tích thống kê cần phải tuân thủ nguyên tắc ngẫu nhiên. Trong thí nghiệm đồng ruộng, thực hiện nguyên tắc ngẫu nhiên có nghĩa là các công thức thí nghiệm được bố trí ngẫu nhiên và mẫu được lấy ngẫu nhiên. Nguyên tắc ngẫu nhiên trong thí nghiệm còn giúp ta tránh được những thiên vị do con người gây nên.

Thí nghiệm ngẫu nhiên có 2 đặc tính: Không biết chắc hậu quả nào sẽ xảy ra nhưng biết được các hậu quả có thể xảy ra. Ví dụ, khi gieo một hạt xúc xắc ta không biết chắc mặt nào sẽ xuất hiện; nhưng ta biết được có 6 trường hợp xảy ra (xúc xắc có 6 mặt là 1,2,3,4,5 và 6).

Với thí nghiệm bón lân cho lúa trên đất lầy thụt thì các công thức 0; 20 và 50 P_2O_5 phải được bố trí ngẫu nghiên nghĩa là bất kỳ ô nào trong thí nghiệm. Với việc nghiên cứu khối lượng búp chè nêu trên thì 3 điểm đại diện phải được lấy bất kỳ chỗ nào.

1.3.3. Sai khác duy nhất

Thí nghiệm đồng ruộng thường so sánh kết quả của kỹ thuật này với kỹ thuật kia. Để so sánh được thì yêu cầu chỉ có nhân tố so sánh khác nhau còn nhân tố khác phải đồng đều. Nguyên tắc đó được gọi là nguyên tắc sai khác duy nhất. Nhân tố được so sánh chính là nhân tố thí nghiệm, còn các nhân tố khác gọi là nhân tố phi thí nghiệm. Ví dụ, trong thí nghiệm so sánh giống lúa, giữa các ô thí nghiệm chỉ có giống là khác nhau còn các điều kiện khác - nhân tố phi thí nghiệm (hoặc nền thí nghiệm) như phân bón, nước và chăm sóc khác phải đồng đều.

ĐẠI HỌC THÁI NGUYÊN
THƯ VIỆN
TRƯỜNG ĐẠI HỌC NÔNG LÂM
NG MƯỢN
151193

Do đó yêu cầu đối với các ô thí nghiệm trong một thí nghiệm là có các điều kiện phi thí nghiệm càng giống nhau càng tốt. Chúng ta cũng cần hiểu chính xác từ *phải đồng đều* trên đây. Phải đồng đều có nghĩa là chúng ta cần tạo ra sự đồng đều của các yếu tố phi thí nghiệm bằng cách tạo một nền thí nghiệm như nhau, sử dụng kỹ thuật khối nếu cần thiết v.v. Tuy nhiên, chúng ta cũng không bao giờ đạt được sự đồng đều tuyệt đối giữa các ô thí nghiệm. Những sai khác do không đồng đều của yếu tố phi thí nghiệm (sai số thí nghiệm) sẽ được xử lý thông qua xử lý thống kê.

1.4. PHÂN TÍCH SỰ SAI KHÁC TRONG KẾT QUẢ THÍ NGHIỆM (PHÂN TÍCH BIẾN ĐỘNG)

Số liệu thu thập được trong quá trình thí nghiệm cho thấy các công thức khác nhau cho kết quả khác nhau. Có 2 nhân tố gây nên sự sai khác giữa các công thức thí nghiệm. Đó là sai khác gây ra do nhân tố thí nghiệm và nhân tố phi thí nghiệm (yếu tố phi thí nghiệm). Vì vậy, nếu chỉ dựa vào kết quả trung bình của mẫu quan sát có thể dẫn tới các kết luận sai lầm.

Để làm rõ hơn chúng ta xem xét ví dụ sau: Thí nghiệm so sánh 2 giống lúa với 3 lần nhắc lại có kết quả như sau: Năng suất trung bình của giống A là 5 tấn/ha, của giống B là 5,5 tấn/ha. Chúng ta có thể kết luận giống B có năng suất cao hơn giống A được không? Sự khác nhau đó có phải thực sự là do khác về giống hay là do nhân tố phi thí nghiệm gây ra, bởi vì có 2 nguyên nhân đưa đến sự sai khác của kết quả trên:

- Do khác nhau về giống (nhân tố thí nghiệm): Giống A khác giống B

- Do khác nhau về điều kiện như nước, ánh sáng, sâu bệnh...(nhân tố phi thí nghiệm).

Vấn đề này được giải thích như sau: Nếu như các ô thí nghiệm đều được cấy cùng một giống thì năng suất thu được trên mỗi ô cũng có sự sai khác. Sai khác đó gây ra là do điều kiện ở các ô không hoàn toàn giống nhau và có thể nói sai khác đó là do ngẫu nhiên.

Như vậy khi đã có kết quả thí nghiệm ta cần phải phân tích sự sai khác trong kết quả thí nghiệm là do nhân tố nào. Phương pháp phân tích thống kê cho phép xác định được sai khác do mỗi nguyên nhân (nguồn) khác nhau. Đó là lý do cần áp dụng phương pháp phân tích thống kê trong xử lý số liệu thu thập được. Phân tích sai khác được gọi là phân tích phương sai hoặc phân tích biến động.

Từ kết quả phân tích sai khác ta kết luận được chấp nhận hay bác bỏ giả thuyết. Với ví dụ 2 giống lúa ở trên sẽ có 2 giả thuyết trước khi làm thí nghiệm là Ho: năng suất 2 giống lúa như nhau và H1: năng suất 2 giống lúa khác nhau. Nếu như kết quả phân tích thống kê cho thấy sai khác là do yếu tố ngẫu nhiên thì ta kết luận 2 giống lúa có năng suất như nhau. Có nghĩa là ta chấp nhận giả thuyết Ho.

1.5. NÂNG CAO ĐỘ CHÍNH XÁC CỦA THÍ NGHIỆM ĐỒNG RUỘNG

Với đối tượng nghiên cứu là cây trồng vì vậy kết quả thí nghiệm đồng ruộng phụ nhiều vào điều kiện phi thí nghiệm và phụ thuộc vào kỹ năng của người làm thí nghiệm. Để nâng cao độ chính xác của thí nghiệm đồng ruộng có thể có một số cách sau.

1.5.1. Tăng số lần nhắc lại

Về nguyên tắc độ chính xác của thí nghiệm tăng lên khi ta tăng số lần nhắc lại. Tuy nhiên khi tăng số lần nhắc lại diện tích thí nghiệm tăng vì thế độ đồng đều giữa các ô giảm. Số lần nhắc lại nhiều hay ít phụ thuộc vào độ đồng đều giữa các ô thí nghiệm và yêu cầu về độ chính xác của thí nghiệm. Thường số lần nhắc lại với thí nghiệm đồng ruộng là từ 4 - 8, trong đó 4 lần nhắc lại được dùng rộng rãi nhất.

1.5.2. Kỹ thuật trong khi làm thí nghiệm

Kỹ thuật trong khi làm thí nghiệm thể hiện trình độ và tính cẩn thận. Đó là đảm bảo sự đồng đều khi thực hiện bất cứ một công việc gì trong thí nghiệm. Người làm thí nghiệm cần phải nhận thức rõ rằng không có bất cứ một kỹ thuật xử lý số liệu nào có thể nâng cao độ chính xác của tní nghiệm từ những số liệu thu được của một thí nghiệm được thực hiện một cách cẩu thả. Sự biến động do cẩu thả trong thực hiện thí nghiệm không phải là nguồn biến động ngẫu nhiên (yếu tố phi thí nghiệm) vì nó không tuân theo nguyên tắc ngẫu nhiên.

Có thể làm rõ về trình độ và tính cẩn thận khi thực hiện thí nghiệm ở một số trường hợp sau. Đó là việc đảm bảo đồng đều về mật độ cấy, bón phân và thực hiện tất cả các kỹ thuật khác trong thí nghiệm về giống lúa. Chú ý: Khi làm một công đoạn kỹ thuật nào đó trong thí nghiệm mà không thể đảm bảo độ đồng đều cho cả thí nghiệm thì cần đảm bảo đồng đều cho từng lần nhắc lại (chỉ có ý nghĩa trong thí nghiệm bố trí theo khối ngẫu nhiên hoàn chỉnh). Ví dụ: Hoàn thành cấy hay gặt cho một lần nhắc lại trong ngày....

1.5.3. Kỹ thuật thiết kế thí nghiệm đồng ruộng

Do điều kiện tự nhiên nơi làm thí nghiệm không đồng nhất nên có thể dùng kỹ thuật thiết kế để tách được sự sai khác do yếu tố ngẫu nhiên gây ra. Kỹ thuật thiết kế thường được áp dụng là kỹ thuật khối. Cách chia dựa vào sự biến đổi của ruộng thí nghiệm. Nơi đất tương đối đồng nhất thì đưa vào một khối. Ví dụ thí nghiệm trên đất dốc thì chia khối theo đường đồng mức. Kiểu thiết kế hay dùng trong thí nghiệm đồng ruộng là khối ngẫu nhiên hoàn chỉnh và ô vuông la tinh....Phần này sẽ được đề cập kỹ hơn ở chương 2.

1.6. THÍ NGHIỆM MỘT NHÂN TỐ VÀ THÍ NGHIỆM ĐA NHÂN TỐ

Nhân tố thí nghiệm là nhân tố thay đổi trong thí nghiệm để ta xác định ảnh hưởng của nó. Có thể phân loại thí nghiệm theo số nhân tố thí nghiệm như sau.

1.6.1. Thí nghiệm một nhân tố

Thí nghiệm chỉ có một nhân tố thay đổi gọi là thí nghiệm một nhân tố. Trong thí nghiệm một nhân tố số công thức bằng mức độ của nhân tố thí nghiệm. Thí nghiệm một nhân tố có những ưu điểm và nhược điểm sau. Ví dụ thí nghiệm các mức lân bón cho lúa trên đất lầy thụt nêu trên là thí nghiệm một nhân tố vì chỉ có mức lân là nhân tố thay đổi.

Ưu điểm

Đơn giản dễ làm, số lượng công thức thường không lớn

Dễ xử lý số liệu

Nhược điểm

Thí nghiệm một nhân tố thường được coi là hạn chế vì phản ứng của thực vật với một tác động nào đó có thể phụ thuộc nhiều vào mức độ tác động của nhân tố khác. Vì vậy nó có những nhược điểm sau:

Phạm vi ứng dụng hạn chế. Như ví dụ các mức lân nêu trên thì kết quả chỉ đúng với giống lúa được dùng trong thí nghiệm chứ chưa chắc đúng với giống lúa khác.

Không đánh giá được tương tác giữa các yếu tố. Trong ví dụ trên ta không đánh giá được quan hệ của giống với phân bón. Liệu các giống có phản ứng khác nhau với các mức phân bón khác nhau không?

Không tìm được những tổ hợp ưu việt nhất giữa các mức của các nhân tố thí nghiệm với nhau. Với ví dụ trên có thể có câu hỏi là giống nào cho năng suất cao nhất với mức phân nào?

1.6.2. Thí nghiệm đa nhân tố

Thí nghiệm đa nhân tố là thí nghiệm có nhiều nhân tố thay đổi. Thí nghiệm có hai nhân tố thay đổi gọi là thí nghiệm hai nhân tố. Số công thức thí nghiệm trong thí nghiệm đa nhân tố là số tổ hợp của sự kết hợp các mức của các nhân tố với nhau. Ví dụ thí nghiệm 2 nhân tố gồm 4 giống lúa (V1, V2, V3, V4) trên 3 nền phân bón khác nhau (cao, trung bình, thấp) sẽ có số công thức thí nghiệm là 12 (4x3 = 12).

Đối tượng của khoa học nông nghiệp là các sinh vật sống chịu ảnh hưởng đồng thời của nhiều yếu tố cho nên những thí nghiệm đa nhân tố rất quan trọng trong khoa học nông nghiệp. Thí nghiệm đa nhân tố có những ưu nhược điểm sau:

Ưu điểm

Phạm vi ứng dụng kết quả thí nghiệm rộng vì có nhiều công thức được thử nghiệm.

Đánh giá được tương tác giữa các nhân tố bởi vì các công thức là sự phối hợp giữa các mức của nhân tố thí nghiệm. Thí nghiệm 2 nhân tố nêu trên có các tổ hợp giữa phân và giống, do vậy sẽ xác định được sự tương tác giữa các nhân tố.

Xác định được các tổ hợp tối ưu giữa các nhân tố thí nghiệm bởi vì cho kết quả của các tổ hợp. Ví dụ, kết quả thí nghiệm trên cho thấy giống V1 cho năng suất cao nhất ở mức phân bón cao; giống V2 cho năng suất cao nhất ở mức phân bón trung bình. Như vậy có nghĩa ta đã tìm ra 2 tổ hợp phân và giống tốt nhất đó là V1 với mức phân cao và V2 với mức phân trung bình. Điều này rất quan trọng trong thực tiễn sản xuất với những điều kiện khác nhau. Ta có thể khuyến cáo giống V2 cho miền núi nơi không có điều kiện thâm canh cao. Nơi có điều kiện thâm canh tốt khuyến cáo trồng giống V1.

Nhược điểm

Thí nghiệm đa nhân tố thường phức tạp, nhiều công thức do đó yêu cầu diện tích thí nghiệm lớn. Với yêu cầu diện tích lớn khó tìm khu đất đồng đều để làm thí nghiệm.

Khi số lượng nhân tố thí nghiệm càng tăng thì mức độ phức tạp của thí nghiệm càng cao và càng nhiều mối tương tác giữa các nhân tố xảy ra gây khó khăn trong việc xử lý số liệu.

Vì vậy, việc chọn thí nghiệm đa nhân tố hay thí nghiệm một nhân tố phụ thuộc vào giả thuyết và mục đích tiến hành thí nghiệm. Ví dụ nếu ta thực sự muốn kiểm tra sự khác nhau giữa các giống nhưng không quan tâm nhiều đến chế độ phân bón vì ở nơi làm thí nghiệm thường bón phân một lượng nhất định nào đó thì nên làm thí nghiệm một nhân tố là giống và dùng mức bón phân phổ biến nhất ở nơi ta làm thí nghiệm.

1.7. TỔNG THỂ VÀ MẪU

Tổng thể là tập hợp của tất cả các đối tượng mà ta quan tâm nghiên cứu. Số phần tử của tổng thể được ký hiệu là N. Ví dụ tất cả các cây lúa của một công thức nào đó trong thí nghiệm về giống lúa là tổng thể. Mẫu là tập hợp con của tổng thể. Mẫu gồm một số hữu hạn (n) phần tử. Số n được gọi là cỡ mẫu. Như vậy mẫu được chọn ra từ tổng thể.

Tại sao phải chọn mẫu? Do không thể theo dõi khảo sát được toàn bộ tổng thể. Mẫu nhỏ hơn tổng thể nên việc thu thập và tổng hợp số liệu từ mẫu nhanh chóng và ít tốn kém hơn. Hơn nữa nhiều khi do phương pháp thu thập số liệu nên các phần tử bị biến đổi hoặc phá huỷ như trong trường hợp quan sát khối lượng rễ cây.

Có nhiều cách chọn mẫu khác nhau nhưng nguyên tắc quan trọng nhất là mẫu phải phản ánh trung thực tổng thể. Muốn vậy phải chọn mẫu đại diện cho tổng thể bằng cách chọn "*ngẫu nhiên*" từ trong tổng thể. Ngoài ra, sự chọn mẫu còn phải đảm bảo một đặc tính khác nữa là sự "*độc lập*" giữa các phần tử trong mẫu. Nghĩa là kết quả của việc chọn phần tử này không ảnh hưởng đến việc chọn các phần tử khác. Một mẫu thoả mãn được hai điều kiện trên được gọi là "*mẫu ngẫu nhiên*".

Để đảm bảo nguyên tắc trên, trong thí nghiệm đồng ruộng thường áp dụng chọn mẫu ngẫu nhiên hoặc máy móc. Ngẫu nhiên là lấy bất kỳ, còn máy móc là lấy theo một quy định được định trước. Độ lớn của mẫu tuỳ thuộc vào các yếu tố như: Phương pháp chọn; Mức độ sai khác trong tổng thể; Độ chính xác trông đợi của kết quả; Mức độ tin cậy thống kê yêu cầu.

1.8. MỘT SỐ KHÁI NIỆM TRONG PHÂN TÍCH THỐNG KÊ

1.8.1. Số định tâm (Measure of central tendency)

Số định tâm là số nói lên đặc điểm chung của nhóm dữ liệu. Dưới đây giới thiệu một số loại số định tâm: Số trung bình, số trung vị, mốt.

1.8.1.1. Số trung bình (Mean)

Số trung bình của tổng thể (Population mean): μ

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Trong đó:

N: số cá thể trong tổng thể

x_i là cá thể thứ i

i là số thứ tự của cá thể thứ i

Số trung bình của mẫu (Sample mean): $\overline{X}$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Trong đó:

x_i là cá thể thứ i

i là số thứ tự cá thể thứ i

n là số lượng cá thể

Số trung bình số học đôi khi bị ảnh hưởng bởi lượng biến quá cao hay quá thấp nên trị số của nó có thể không phản ánh được giá trị trung bình của tổng thể. Có thể kết hợp với tần số để nói lên đặc điểm của nhóm dữ liệu.

1.8.1.2. Số trung vị (Median)

Số trung vị là giá trị của số đứng ở vị trị giữa trong dãy số liệu. Gọi n là số giá trị quan sát được cách xác định số trung vị như sau:

Nếu n là số lẻ thì số trung vị là số có thứ tự (n +1)/2.

Nếu n là số chẵn thì số trung vị là số trung bình cộng của 2 số có thứ tự n/2 và $\frac{n}{2} + 1$

1.8.1.3. Số mốt (mode)

Mốt là trị số có tần số lớn nhất trong dãy số.

1.8.2. Số phân tán (Measure of dispersion)

Số phân tán dùng để thể hiện sự khác biệt giữa các số trong dãy số liệu đối với số định tâm.

1.8.2.1. Phương sai (Variance)

Phương sai là tổng bình phương của sai khác giữa cá thể thứ i với giá trị trung bình.

Phương sai của tổng thể (Population variance):

$$\delta^2 = \frac{\sum_{i=1}^{N}(xi-\mu)^2}{N} = \frac{\sum_{i=1}^{n} x_i^2}{N} - \mu$$

Phương sai của mẫu (Sample variance):

$$S^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}$$

Phương sai thể hiện sự biến động của quần thể. Có thể làm rõ khái niệm về phương sai bằng ví dụ sau.

Bảng 1.3. Số nhánh của 10 cây lúa từ 2 quần thể mẫu lúa

Quần thể 1							Quần thể 2						
	x_i				Σ	$\bar{x}$		x_i				Σ	$\bar{x}$
Cây số	1	2	3	4			Cây số	1	2	3	4		
Số nhánh	1	2	6	3	12	3	Số nhánh	3	4	3	2	12	3
$x_i - \bar{x}$	-2	-1	3	0			$x_i - \bar{x}$	0	1	0	-1		
$(x_i - \bar{x})^2$	4	1	9	0	14		$(x_i - \bar{x})^2$	0	1	0	1	2	

$$S^2 = \sum(x_i - \bar{x})^2 = 14\,;\ S^2 = \sum(x_i - \bar{x})^2 = 2$$

Phương sai đo độ phân tán của quần thể. Có nhiều trường hợp 2 quần thể có số trung bình bằng nhau nhưng phương sai khác nhau thì hai quần thể đó cũng khác nhau.

Ý nghĩa của phương sai nếu chọn giống bằng cách lai hữu tính quần thể 1 sẽ có cơ hội tạo ra nhiều giống mới, nhưng nếu giám định giống quần thể 2 có tính ổn định tốt hơn.

1.8.2.1. Độ lệch chuẩn (Standard deviation)

Độ lệch chuẩn của tổng thể (Population standard deviation)

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{N}\sum\left(x_i - \mu\right)^2}$$

Độ lệch chuẩn của mẫu (Sample standard deviation)

$$S = \sqrt{S^2} = \sqrt{\frac{1}{n-1}\sum\left(x_i - \bar{x}\right)^2}$$

Độ lệch chuẩn đo lường độ chính xác của mẫu được chọn. Nếu độ lệch chuẩn mẫu càng lớn thì độ tin cậy của việc dùng ước lượng từ mẫu để suy diễn cho tổng thể càng kém.

Ý nghĩa của độ lệch chuẩn σ *:*

Quy tắc phân bố hình chuông (Normal Distribution): Một tập hợp dữ liệu ngẫu nhiên có dạng phân bố hình chuông hay gọi là phân bố chuẩn. Đối với tập hợp dữ liệu có dạng phân bố hình chuông thì:

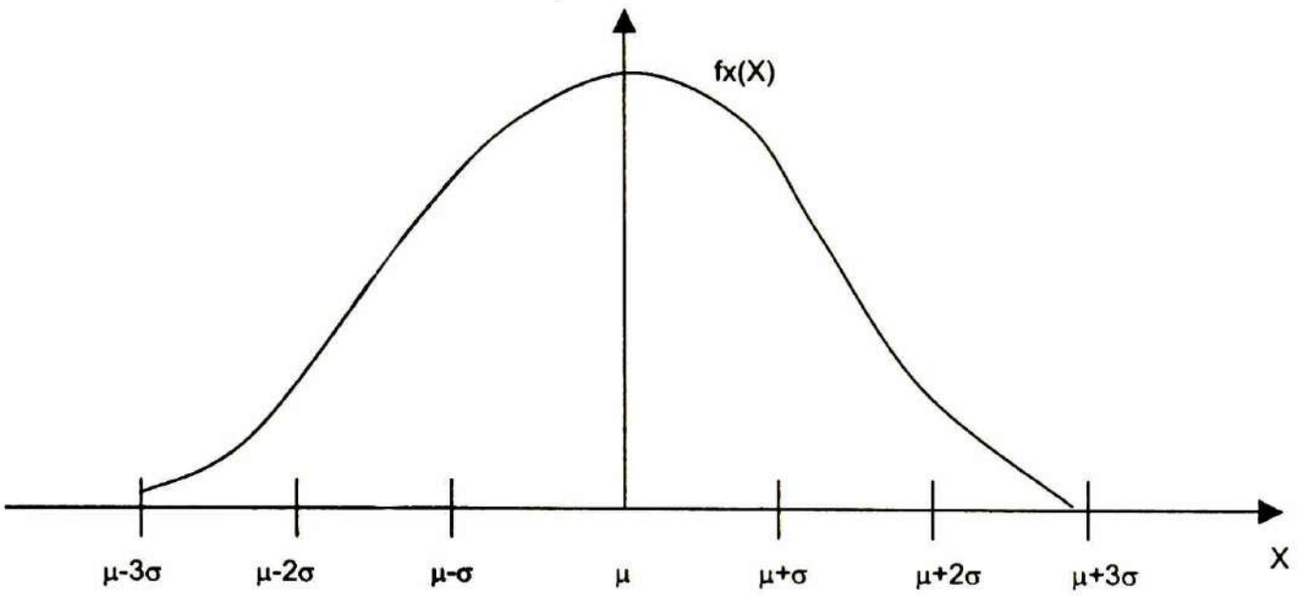
Khoảng 68% các giá trị nằm trong khoảng $\mu \pm \sigma$

Khoảng 95% các giá trị nằm trong khoảng $\mu \pm 2\sigma$

Khoảng 100% giá trị nằm trong khoảng $\mu \pm 3\sigma$

Đồ thị của một tập hợp dữ liệu phân bố hình chuông có dạng như sau:

fx: tần suất của X



Theo định nghĩa phương sai được tính như sau:

$$\sum\left(X_i - \overline{X}\right)^2 = \sum\left(X_i^2 - 2X_i\overline{X} + \overline{X^2}\right) = \sum X_i^2 - 2\sum X_i\overline{X} + n\overline{X^2} =$$

$$\sum X_i^2 - 2\overline{X}\sum X_i + n\overline{X}$$

Thay $\overline{X} = \dfrac{\sum X_i}{n}$ có

$$\sum X_i^2 - 2\frac{\sum X_i}{n}\sum X_i + n\left(\frac{\sum X_i}{n}\right)^2 = \sum X_i^2 - 2\frac{\left(\sum X_i\right)^2}{n} + \frac{(\sum X_i)^2}{n} =$$

$$\sum X_i^2 - \frac{(\sum X_i)^2}{n} = \sum X_i^2 - \frac{G^2}{n}$$

Ký hiệu $\sum$ có nghĩa là "tổng của", biểu thức $G = X_1 + X_2 + .. + X_n$ có thể được viết là $G = \sum_{i=1}^{n} X_i$ hay đơn giản là $G = \sum X$.

Chương 2

BỐ TRÍ THÍ NGHIỆM ĐỒNG RUỘNG VÀ THU THẬP SỐ LIỆU

2.1. BỐ TRÍ THÍ NGHIỆM ĐỒNG RUỘNG

Có rất nhiều yếu tố tự nhiên ảnh hưởng đến độ chính xác của kết quả thí nghiệm đồng ruộng như độ dốc, hướng của ánh sáng...nên cần phải xem xét bố trí các ô thí nghiệm phù hợp với điều kiện thực tế của đồng ruộng.

Bố trí thí nghiệm là cách xắp xếp các công thức thí nghiệm vào các ô thí nghiệm trên đồng ruộng. Yêu cầu của bố trí là tạo sự đồng đều tối đa giữa các ô và hạn chế sai khác do yếu tố phi thí nghiệm. Cách bố trí các công thức có ảnh hưởng quan trọng đến kết quả thí nghiệm. Ví dụ với thí nghiệm trên đất dốc, rõ ràng đất ở đỉnh đồi khác so với đất ở chân đồi. Do đó kết quả khác nhau giữa các công thức được bố trí ở đỉnh đồi với những công thức bố trí ở chân đồi có thể là do đất khác nhau chứ không phải do nhân tố thí nghiệm.

Kỹ thuật thường được dùng trong thí nghiệm đồng ruộng nhằm hạn chế ảnh hưởng của yếu tố phi thí nghiệm là chia khối. Cách chia khối thường dùng là: khối ngẫu nhiên hoàn chỉnh, ô vuông, ô chính - ô phụ,

2.1.1. Chia khối

Chia khối là kỹ thuật chia những khu vực của ruộng thí nghiệm có điều kiện tương đồng vào cùng một khối. Các ô thí nghiệm trong cùng một khối sẽ có độ đồng đều cao hơn so với những ô thí nghiệm ở khác khối. Nhờ vậy sai khác do yếu tố phi thí nghiệm gây ra giữa các ô trong mỗi khối giảm đến tối thiểu. Cách chia khối cần căn cứ vào hướng biến động. Nguyên tắc chung là các khối được bố trí theo hướng vuông góc với hướng biến động. Ví dụ, thí nghiệm trên đất dốc, khối sẽ được chia song song với đường đồng mức vì sự thay đổi của đất theo hướng từ đỉnh đồi đến chân đồi.

Những yếu tố biến động (yếu tố phi thí nghiệm) ảnh hưởng đến kết quả thí nghiệm có thể là: Sự không đồng nhất của đất trong các thí nghiệm về phân bón hoặc giống mà năng suất là chỉ tiêu khảo sát chính; Hướng di chuyển của côn trùng trong những thí nghiệm về thuốc trừ sâu mà ảnh hưởng của sâu bệnh là chỉ tiêu khảo sát chính; Độ dốc hay thế đất của khu thí nghiệm trong những nghiên cứu về ảnh hưởng của nước đến cây trồng.

Một số gợi ý cho chia khối

Khi nguồn biến thiên theo một hướng, chọn những khối dài và hẹp thẳng góc với hướng biến thiên

Khi nguồn biến thiên theo hai hướng với một hướng biến thiên mạnh hơn hướng kia, chỉ xét hướng biến thiên nhiều và bố trí như trên.

Khi nguồn biến thiên theo hai hướng đều mạnh bằng nhau và thẳng góc nhau, chọn một trong những cách sau: Bố trí khối ngẫu nhiên càng vuông càng tốt. Bố trí theo kiểu ô vuông la tinh.

Khi không xác định được rõ hướng biến thiên, nên bố trí khối vuông.

Quan niệm về khối cần phải hiểu rộng hơn. Khối có thể là: Thời gian: năm hoặc thời vụ; Không gian: các ruộng thí nghiệm khác nhau hoặc các khu khác nhau trong một ruộng thí nghiệm, hoặc các phòng thí nghiệm khác nhau.

Tuy nhiên cũng cần nhắc lại là kỹ thuật khối không cần thiết khi yếu tố phi thí nghiệm tương đối đồng đều. Khi đó sử dụng kỹ thuật khối không giảm được sai số ngẫu nhiên nhiều nhưng lại làm tăng F lý thuyết do làm giảm độ tự do của ngẫu nhiên.

2.1.2. Một số lưu ý khi bố trí thí nghiệm trên đồng ruộng

Chọn nền thí nghiệm: Chọn nền thí nghiệm tiên tiến và phù hợp với địa bàn sẽ áp dụng kết quả thí nghiệm. Ví dụ nền phân bón cho thí nghiệm giống phải đầy đủ (nhưng cũng không quá khác so với thực tiễn) thì giống mới có điều kiện thể hiện được tiềm năng của nó.

Cách ly: Yêu cầu cách ly giữa các ô để đảm bảo sự chính xác của các mức của nhân tố thí nghiệm. Ví dụ cách ly các ô thí nghiệm có mức phân bón khác nhau trong thí nghiệm bón phân cho lúa nước.

Hàng rào: Hàng rào thí nghiệm nên dùng loại cây cùng loại trong thí nghiệm để đảm bảo sự đồng đều giữa những ô bên ngoài và ô bên trong thí nghiệm.

Khoảng cách: Khoảng cách giữa các khối thuận tiện cho chăm sóc và theo dõi. Độ rộng tuỳ loại cây thí nghiệm, ví dụ với lúa, đậu, đỗ và rau nên 40-50 cm; với ngô 80-100 cm.

2.2. BỐ TRÍ THÍ NGHIỆM MỘT NHÂN TỐ

2.2.1. Kiểu ngẫu nhiên hoàn toàn (Completely Randomized Design -CRD)

2.2.1.1. Đặc điểm

Thí nghiệm bố trí theo kiểu ngẫu nhiên hoàn toàn có nghĩa là công thức được bố trí vào bất kỳ ô thí nghiệm nào. Đây là kiểu bố trí đơn giản. Các công thức được nhắc lại nhưng không nhóm thành khối. Ràng buộc duy nhất là các công thức phải có xác suất như nhau để được bố trí vào bất kỳ ô thí nghiệm nào.

Các nguồn biến động trong thí nghiệm bố trí kiểu ngẫu nhiên hoàn toàn

Biến động do công thức

Biến động do ngẫu nhiên

Ưu điểm

Thí nghiệm hoàn toàn linh động: Áp dụng dễ dàng cho bất kỳ một số lượng công thức hay một số lần nhắc lại nào (Số lượng công thức và số lần nhắc lại không bị hạn chế). Số lần nhắc lại cũng có thể không giống nhau giữa các công thức (mặc dù các công thức thường nên có số lần nhắc lại bằng nhau)

Xử lý thống kê đơn giản kể cả khi số lần nhắc lại không bằng nhau và trường hợp mất số liệu cũng không gây nhiều khó khăn trong xử lý số liệu.

Độ tự do (dfe) của sai số ngẫu nhiên lớn

Nhược điểm

Sai số thí nghiệm lớn nếu khu thí nghiệm không đồng đều vì sai số thí nghiệm bao gồm toàn bộ biến động giữa các ô thí nghiệm, chỉ trừ đi biến động do công thức.

Trường hợp áp dụng

Trong thực tế CRD chỉ áp dụng khi có sự đồng đều cao giữa các ô thí nghiệm, vì vậy nó thích hợp với các thí nghiệm trong phòng thí nghiệm và các thí nghiệm trong nhà kính

Cũng có thể áp dụng cho một số thí nghiệm nhỏ (ít công thức, ít lần nhắc). Trong những trường hợp này tăng tối đa độ tự do của sai số ngẫu nhiên (dfe) quan trọng hơn là làm giảm tổng bình phương của sai số ngẫu nhiên (SSe). Do vậy bố trí CRD trong những trường hợp này để tăng tối đa df là hợp lý

Đối với các thí nghiệm đồng ruộng nơi mà có độ biến động lớn giữa các ô thí nghiệm (ví dụ đất đai, độ dốc..) thì CRD ít được áp dụng.

2.2.1.2. Cách bố trí

Có nhiều cách để bố trí ngẫu nhiên các công thức thí nghiệm. Sau đây là cách bố trí ngẫu nhiên hoàn toàn theo Gomez (1984)

Minh họa qua bố trí thí nghiệm "So sánh năng suất của 4 giống lúa" có 4 công thức (t=4; t: treatment) ; 5 lần nhắc lại (r=5; r: replication).

Bước 1: Xác định số ô thí nghiệm n = (r)(t)= (5)(4)=20

Bước 2: Chia ruộng thí nghiệm ra làm 20 ô. Sau đó đánh số ô thí nghiệm từ 1-20 như trong sơ đồ 2.1.

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

Sơ đồ 2.1. Sơ đồ đánh số thứ tự các ô thí nghiệm

Bước 3: Bố trí công thức vào các ô thí nghiệm theo cách ngẫu nhiên như sau.

A_1: Chọn 1 số bất kỳ trong bảng số ngẫu nhiên (bằng cách nhắm mắt và đặt ngón tay vào 1 điểm bất kỳ). Theo dấu mũi tên ở bảng 2.1. số bất kỳ đó là số 69374.

Bảng 2.1. Bảng số ngẫu nhiên

14620	95430	12951	81953	17629
09724	85125	48477	42783	70473
56919	17803	95781	85069	61594
97310	78209	51263	52396	82680
07585	28040	26939	64531	70570
25950	85189	69374	37904	06759
82937	16405	81497	20863	94072
60819	27364	59081	72635	49180
59041	38475	03615	84093	49731
74208	69516	79530	47649	53046
39412	03642	87497	29735	14308
48480	50075	11804	24956	72182
95318	28749	49512	35408	21814
72094	16385	90285	72635	86259
63158	49753	84279	56496	30618
19082	73645	09182	73649	56823
15232	84146	87729	65584	83641
94252	77489	62434	20965	20247
72020	18895	84948	53072	74573
48392	06359	47040	05695	79799
37950	77387	35495	48192	84518
09394	59842	39573	51630	78548
34800	28055	91570	99154	39603
36435	75946	85712	06293	85621
28187	31824	52265	80494	66428

A_2: Từ số 69374 lấy 20 số liên tiếp (vì n = 20). Để đơn giản nên lấy 3 chữ số bất kỳ của những số này (có thể lấy 3 số giữa, 3 số đầu hoặc 3 số cuối). Sau đó, sắp xếp các số được chọn theo thứ tự xuất hiện của chúng trong bảng ngẫu nhiên như bảng dưới đây.

Bảng 2.2. Cách xếp thứ tự các số ngẫu nhiên đã chọn

Số ngẫu nhiên	Thứ tự	Số ngẫu nhiên	Thứ tự
937	1	918	11
149	2	772	12
908	3	243	13
361	4	494	14
953	5	704	15
749	6	549	16
180	7	957	17
951	8	157	18
018	9	571	19
427	10	226	20

A_3: xếp hạng các số đã chọn theo thứ tự từ nhỏ đến lớn hoặc ngược lại.

Bảng 2.3. Cách xếp hạng các số ngẫu nhiên đã chọn từ nhỏ đến lớn

Số ngẫu nhiên	Số thứ tự	Xếp hạng	Số ngẫu nhiên	Số thứ tự	Xếp hạng
937	1	17	918	11	16
149	2	2	772	12	14
908	3	15	243	13	6
361	4	7	494	14	9
953	5	19	704	15	12
749	6	13	549	16	10
180	7	4	957	17	20
951	8	18	157	18	3
018	9	1	571	19	11
427	10	8	226	20	5

A_4: Chia 20 số trên ra 4 nhóm (vì t = 4) mỗi nhóm có 5 số (vì r = 5) theo thứ tự các số xuất hiện trong bảng ngẫu nhiên. Cụ thể, nhóm 1 gồm 5 số thứ nhất (số thứ 1 đến thứ 5), nhóm 2 gồm 5 số tiếp theo (thứ 6 đến thứ 10), tiếp tục như vậy cho đến nhóm cuối cùng ta có bảng sau:

Bảng 2.4. Cách chia các số ngẫu nhiên cho các nhóm

Nhóm	Gồm các số có thứ tự xếp hạng là				
1	17	2	15	7	19
2	13	4	18	1	8
3	16	14	6	9	12
4	10	20	3	11	5

A_5: gán t công thức cho n ô thí nghiệm theo các nhóm đã được chia ở bước A_4 như sau:

Nhóm 1 gán cho công thức A. Nhóm 2 gán cho công thức B, nhóm 3 gán cho công thức C và nhóm 4 gán cho công thức D. Có nghĩa, công thức A được xếp vào

các ô thí nghiệm có số thứ tự như thứ tự xếp hạng của các số trong nhóm 1. Công thức B được xếp vào các ô có số thứ tự như thứ tự xếp hạng của các số trong nhóm 2. Tương tự như vậy cho đến cuối cùng. Kết quả, công thức A bố trí vào các ô thứ 17, 2, 15, 7; 19; công thức B bố trí vào các ô thứ 13, 4, 18, 1; 8; công thức C bố trí vào các ô thứ 16, 14, 6, 9, 12; công thức D bố trí vào các ô thứ 10, 20, 3, 11, 5 và ta có sơ đồ thí nghiệm bố trí theo kiểu ngẫu nhiên hoàn toàn như sau:

1 b	2 a	3 d	4 b
5 d	6 c	7 a	8 b
9 c	10 d	11 d	12 c
13 b	14 c	15 a	16 c
17 a	18 b	19 a	20 d

Sơ đồ 2.2: Sơ đồ thí nghiệm bố trí ngẫu nhiên hoàn toàn (t=4 (A,B,C,D); r=5)

2.2.2. Kiểu khối ngẫu nhiên hoàn chỉnh

(Randomized Complete Block Design - RCBD)

2.2.2.1. Đặc điểm

Bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh là các công thức thí nghiệm của mỗi lần nhắc lại được bố trí theo khối. Mỗi khối được coi là một lần nhắc lại và các công thức thí nghiệm được bố trí ngẫu nhiên trong một khối. Do đó, mỗi khối có đầy đủ các công thức nhưng mỗi công thức chỉ xuất hiện một lần trong một khối. Đây là kiểu thiết kế thí nghiệm cơ bản nhất và được áp dụng rộng rãi nhất trong nghiên cứu nông nghiệp. Nhờ kỹ thuật chia khối mà sự khác nhau giữa các ô trong một khối là nhỏ nhất.

Các nguồn biến động trong thí nghiệm bố trí kiểu khối ngẫu nhiên hoàn chỉnh bao gồm:

- Biến động do công thức
- Biến động do ngẫu nhiên
- Biến động theo khối (hay gọi là biến động do nhắc lại)

Ưu điểm

Độ chính xác cao hơn CRD do tác động của khối

Đơn giản và dễ hiểu hơn so với một số kiểu bố trí phức tạp khác.

Tăng khả năng phân biệt sự sai khác có ý nghĩa giữa các công thức thí nghiệm với nhau

Cho phép tiến hành với bất kỳ một số lượng công thức và bất kỳ một số lần nhắc lại nào.

Nhược điểm

Nếu hiệu quả của khối kém thì độ chính xác của thí nghiệm giảm so với CRD vì dfe trong RCBD bị giảm so với bố trí CRD

Độ đồng nhất trong một khối thường cũng bị giảm khi số lượng công thức quá lớn do kích thước của khối tăng.

Trường hợp áp dụng

RCBD rất thích hợp với thí nghiệm có số công thức không quá lớn và hướng biến động xác định được.

2.2.2.2. Cách bố trí

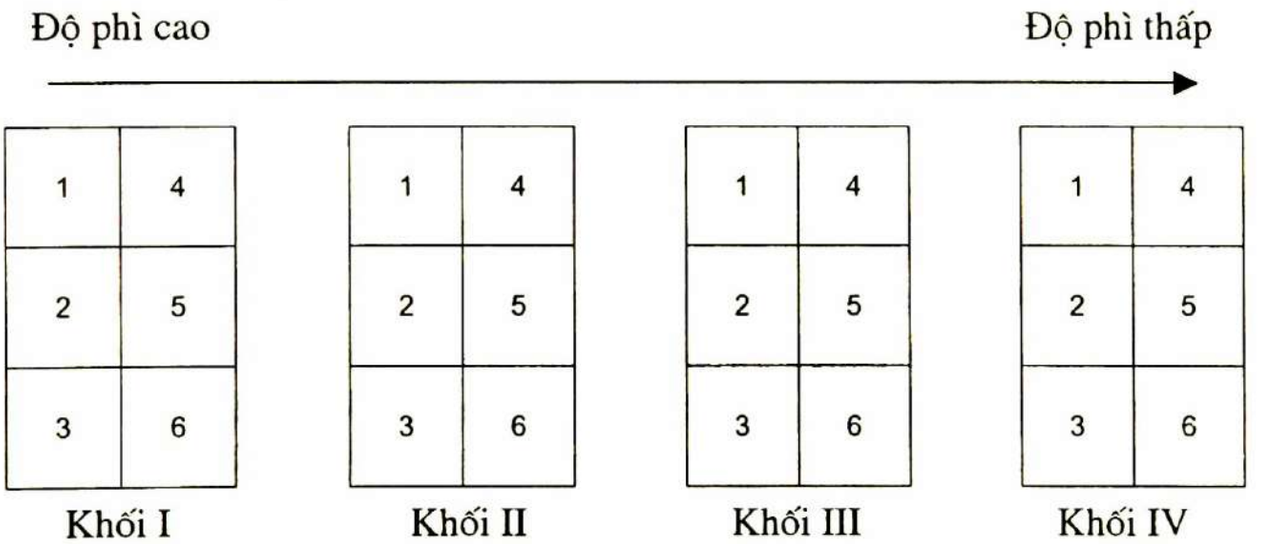
Chia ruộng thí nghiệm ra r khối, số khối bằng số lần nhắc lại. Khối được chia làm các ô, số ô bằng số công thức thí nghiệm. Sau đó, bố trí ngẫu nhiên các công thức vào các ô. Để minh họa chúng ta áp dụng cho trường hợp cụ thể sau.

Thí nghiệm "Nghiên cứu hiệu lực của 6 loại thuốc trừ sâu khác nhau đến rầy nâu hại lúa", có số công thức t = 6 (công thức: A, B, C, D, E, F) và số lần nhắc lại r = 4.

Bước 1:

Chia ruộng thí nghiệm ra 4 khối, chia mỗi khối thành 6 ô và đánh số thứ tự cho các khối và các ô như sơ đồ 2.3. sau.

Hướng biến động của độ phì của đất (hướng mũi tên)



Sơ đồ 2.3: Cách chia nơi thí nghiệm thành r khối và chia khối thành t ô

Bước 2: Bố trí ngẫu nhiên các công thức vào từng khối tương tự như phần 2.2.1

Bố trí cho khối I như sau: lấy 6 số ngẫu nhiên từ bảng số ngẫu nhiên 2.1. Xếp thứ tự và xếp hạng các số ngẫu nhiên đó theo thứ tự từ nhỏ đến lớn ta có kết quả ghi ở bảng sau:

Bảng 2.5. Cách xếp thứ tự và xếp hạng cho các số ngẫu nhiên được chọn

Số ngẫu nhiên	Thứ tự của số ngẫu nhiên (được gán cho công thức thứ)	Xếp hạng của số ngẫu nhiên (vị trí được đặt trong khối)
918	1	6
772	2	5
243	3	1
494	4	2
704	5	4
549	6	3

Gán các công thức với các số ngẫu nhiên theo cách: Thứ tự của số ngẫu nhiên chính là thứ tự của công thức, xếp hạng của số ngẫu nhiên chính là vị trí ô của công thức trong khối. Cụ thể trong khối I, công thức 1 đặt ở vị trí số 6; công thức 2 đặt ở vị trí số 5, công thức 3 đặt ở vị trí số 1, công thức 4 đặt ở vị trí số 2, công thức 5 đặt ở vị trí số 4 và công thức 6 đặt ở vị trí số 3. Tiếp tục như vậy cho các khối khác và kết quả được sơ đồ bố trí theo khối ngẫu nhiên hoàn chỉnh như sau.

1 C	4 E
2 D	5 B
3 F	6 A

Khối I

1 A	4 D
2 E	5 C
3 F	6 B

Khối II

1 F	4 A
2 D	5 B
3 C	6 B

Khối III

1 E	4 D
2 C	5 F
3 B	6 A

Khối IV

Sơ đồ 2.4: Thí nghiệm bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh (t=6; r=4)

2.2.3. Kiểu ô vuông la tinh (Latin Square Design - LS)

2.2.3.1. Đặc điểm

Thí nghiệm bố trí theo kiểu ô vuông la tinh là kiểu đặc biệt của kiểu khối ngẫu nhiên hoàn chỉnh. Trong kiểu này, khối được chia theo cả chiều ngang và chiều dọc. Khối chia theo chiều ngang được gọi là hàng và chia theo chiều dọc được gọi là cột. Các ô thí nghiệm được bố trí theo 2 hướng độc lập với nhau đó là theo hàng và theo cột. Yêu cầu của bố trí kiểu ô vuông la tinh là sự có mặt của tất cả các công thức thí nghiệm trong mỗi hàng và mỗi cột. Mỗi công thức chỉ xuất hiện một lần trong mỗi hàng và mỗi cột. Do đó kiểu này chỉ bố trí được khi số lần nhắc lại bằng số công thức.

Các ô thí nghiệm được nhóm theo hàng theo sự biến động của một yếu tố và bố trí theo cột theo sự biến động của một yếu tố khác. Nhờ vậy có thể tách sai số theo hàng và sai số theo cột ra khỏi sai số thí nghiệm.

- Các nguồn biến động trong thí nghiệm bố trí kiểu ô vuông la tinh bao gồm:
- Biến động do công thức
- Biến động do ngẫu nhiên
- Biến động theo cột
- Biến động theo hàng
- Ưu điểm

Đảm bảo được độ chính xác của thí nghiệm trong trường hợp khu thí nghiệm biến thiên theo hai chiều vuông góc với nhau

- Nhược điểm
- Với yêu cầu số công thức bằng số lần nhắc lại, nên nếu số công thức quá lớn sẽ không thể áp dụng được vì nó đòi hỏi số lần nhắc lại quá lớn. Trong thực tiễn, kiểu này thích hợp với số công thức nằm trong khoảng từ 4-8.
- Khi công thức tăng thì sai số thí nghiệm tăng lên.
- Nếu số công thức nhỏ (dưới 4) thì độ tự do của sai số sẽ rất nhỏ và ảnh hưởng đến việc ước lượng đúng sai số thí nghiệm (thường ước lượng sai số bị tăng lên)
- Nếu số công thức lớn thì diện tích khu thí nghiệm sẽ rất lớn, rất khó tìm được một khu đất đồng đều và đủ rộng để tiến hành thí nghiệm
- Phân tích số liệu rất phức tạp trong trường hợp mất số liệu.
- Trường hợp áp dụng
- Áp dụng khi khu thí nghiệm có 2 nguồn biến thiên vuông góc với nhau, đồng thời số công thức thí nghiệm nằm trong khoảng từ 4-8.
- Những trường hợp thí nghiệm chịu sự tác động của 2 yếu tố biến thiên vuông góc với nhau như: Khu vực thí nghiệm bố trí ở sườn núi vừa chịu ảnh hưởng của hướng gió, vừa chịu ảnh hưởng của độ cao; Thử nghiệm thuốc diệt côn trùng mà ở đó người ta dự đoán côn trùng sẽ di chuyển theo hướng thẳng góc với chiều thay đổi của độ phì đất.

2.2.3.2. Cách bố trí

Chia khu thí nghiệm thành khối theo phương ngang và dọc để tạo thành các hàng và các cột. Số hàng bằng số cột và bằng số công thức thí nghiệm. Từ phụ lục K, chọn sơ đồ phù hợp với số công thức thí nghiệm cần bố trí. Sau đó bố trí lại các hàng và các cột của sơ đồ đã chọn bằng cách lấy số ngẫu nhiên như ở phần 2.3.1.

Cách bố trí cụ thể được minh họa qua bố trí cho thí nghiệm "Nghiên cứu ảnh hưởng của các mức phân đạm tới năng suất sắn", có t = 5 (công thức A, B, C, D, E); r = 5.

Bước 1:

Chia ruộng thí nghiệm thành các khối theo 5 hàng và 5 cột, đánh số thứ tự cho hàng và cột như sơ đồ 2.5 sau:

Cột	1	2	3	4	5	
						Hàng 1
						Hàng 2
						Hàng 3
						Hàng 4
						Hàng 5

Sơ đồ 2.5. Sơ đồ chia khối theo kiểu ô vuông la tinh

Chọn sơ đồ cho thí nghiệm có 5 công thức từ phụ lục K, sơ đồ chọn như sau:

a	b	c	d	e	Hàng 1
b	c	d	e	a	Hàng 2
c	d	e	a	b	Hàng 3
d	e	a	b	c	Hàng 4
e	a	b	c	d	Hàng 5

Sơ đồ 2.6. Sơ đồ bố trí kiểu ô vuông la tinh cho thí nghiệm có 5 công thức

Bước 2:

Bố trí ngẫu nhiên các hàng bằng cách lấy ngẫu nhiên 5 số từ bảng 2.1 và xếp hạng các số này từ nhỏ đến lớn được kết quả như bảng sau.

Bảng 2.6: Cách xếp thứ tự và xếp hạng các số ngẫu nhiên đã chọn

Số ngẫu nhiên	Thứ tự (được gán cho hàng thứ)	Xếp hạng (được coi là vị trí của hàng trong sơ đồ 2.5)
628	(hàng) 1	3
846	(hàng) 2	4
475	(hàng) 3	2
902	(hàng) 4	5
452	(hàng) 5	1

Bố trí ngẫu nhiên từng hàng của sơ đồ 2.6. vào sơ đồ 2.5 theo cách: Gán số thứ tự của số ngẫu nhiên cho hàng. Ta có, thứ tự của số ngẫu nhiên chính là thứ tự của hàng, xếp hạng của số ngẫu nhiên chính là vị trí của hàng được đặt trong sơ đồ 2.5. Từ bảng 2.6, ta có hàng 1 đặt ở vị trí số 3; hàng 2 đặt ở vị trí số 4, hàng 3 đặt ở vị trí số 2, hàng 4 đặt ở vị trí số 5, hàng 5 đặt ở vị trí số 1. Ta có sơ đồ mới như sau:

c1	*c2*	*c3*	*c4*	*c5*
e	a	b	c	d
c	d	e	a	b
a	b	c	d	e
b	c	d	e	a
d	e	a	b	c

Sơ đồ 2.7. Sơ đồ bố trí ngẫu nhiên các hàng cho thí nghiệm bố trí kiểu LS

Tiếp theo, bố trí ngẫu nhiên cho các cột của sơ đồ 2.7 vào sơ đồ 2.5. Đầu tiên, đánh số thứ tự cho các cột của sơ đồ 2.7. Bố trí ngẫu nhiên cho 5 cột trên vào sơ đồ 2.5 bằng cách chọn ngẫu nhiên 5 số và xếp hạng từ nhỏ đến lớn được kết quả như trên. Ta có bảng sau.

Bảng 2.7. Xếp thứ tự và xếp hạng các số ngẫu nhiên đã chọn

Số ngẫu nhiên	Thứ tự (được gán cho cột thứ)	Xếp hạng (được coi là vị trí của cột trong sơ đồ 2.5.)
845	(cột) 1	4
785	(cột) 2	3
396	(cột) 3	1
856	(cột) 4	5
664	(cột) 5	2

Bố trí từng cột của sơ đồ 2.7 vào sơ đồ 2.5 bằng cách: Gán số thứ tự của cột theo thứ tự của số ngẫu nhiên trong bảng 2.7. Ta có, xếp hạng của số ngẫu nhiên trong bảng 2.7 chính là vị trí của cột được đặt trong sơ đồ 2.5. Từ bảng 2.7 ta có, cột 1 đặt ở vị trí số 4; cột 2 đặt ở vị trí số 3, cột 3 đặt ở vị trí số 1, cột 4 đặt ở vị trí số 5, cột 5 đặt ở vị trí số 2. Sơ đồ bố trí kiểu LS cho thí nghiệm có 5 công thức, 5 lần nhắc lại như sau:

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

Sơ đồ 2.8: Sơ đồ thí nghiệm bố trí theo kiểu ô vuông la tinh (t= 5; r = 5)

2.3. BỐ TRÍ THÍ NGHIỆM 2 NHÂN TỐ

2.3.1. Bố trí kiểu khối ngẫu nhiên hoàn chỉnh

(Randomized Complete Block Design)

2.3.1.1. Đặc điểm

Đây là thí nghiệm 2 nhân tố bố trí giống như kiểu khối ngẫu nhiên hoàn chỉnh của thí nghiệm 1 nhân tố. Số khối bằng số lần nhắc lại, trong một khối có đầy đủ các

công thức là các tổ hợp của hai nhân tố thí nghiệm. Các công thức được bố trí hoàn toàn ngẫu nhiên trong các khối. Kỹ thuật bố trí như lựa chọn hướng bố trí và hình dạng khối tương tự như trong phần thí nghiệm một nhân tố.

- Các nguồn biến động trong thí nghiệm bố trí kiểu khối ngẫu nhiên hoàn chỉnh của thí nghiệm 2 nhân tố bao gồm:

- Biến động do nhắc lại
- Biến động do nhân tố thứ 1
- Biến động do nhân tố thứ 2
- Biến động do ngẫu nhiên của nhân tố thứ 1
- Biến động do ngẫu nhiên của nhân tố thứ 2
- Biến động do tương tác 2 nhân tố

- Trường hợp áp dụng

Trường hợp áp dụng của thí nghiệm kiểu khối ngẫu nhiên hoàn chỉnh 2 nhân tố cũng giống như đối với thí nghiệm một nhân tố, chỉ khác là ở đây dùng cho 2 nhân tố.

2.3.1.2. Cách bố trí

Trước tiên ta xác định số công thức của thí nghiệm. Ví dụ thí nghiệm có 2 nhân tố là A (có a mức) và B (có b mức), số công thức thí nghiệm n = (a x b). Chia khu thí nghiệm ra làm r khối tương ứng với số lần nhắc lại. Tiếp theo, bố trí ngẫu nhiên các công thức vào các ô thí nghiệm cho từng khối một như trong kiểu bố trí khối ngẫu nhiên hoàn chỉnh cho thí nghiệm 1 nhân tố ở phần 2.3.2. Ví dụ sau minh họa về cách bố trí.

Thí nghiệm nghiên cứu phản ứng của 3 giống lúa (A1, A2, A3) với 5 mức phân N bón (N1,N2, N3, N4, N5) và 3 lần nhắc lại. Ta có: Giống là nhân tố A có a = 3 mức và N là nhân tố B có b = 4 mức, như vậy có tổng số công thức n= 3 x 4 = 12. Với 3 lần nhắc lại cần (12 x 3) = 36 ô. Cụ thể gồm các công thức sau:

A1n1, a1n2, a1n3, a1n4,

A2n1, a2n2, a2n3, a2n4,

A3n1, a3n2, a3n3, a3n4,

Sơ đồ bố trí:

Khối I

A1N1	A1N3	A3N3	A2N3
A3N2	A2N1	A1N2	A3N1
A2N2	A2N4	A3N4	A1N4

Khối II

A3N4	A1N4	A2N2	A2N4
A1N1	A1N3	A3N3	A2N3
A3N3	A2N3	A1N1	A1N3

Khối III

A1N3	A3N4	A2N4	A3N3
A2N3	A1N1	A2N3	A1N1
A1N4	A3N3	A1N3	A2N2

Sơ đồ 2.9. Sơ đồ thí nghiệm 2 nhân tố bố trí kiểu khối ngẫu nhiên hoàn chỉnh

2.3.2. Kiểu ô chính - ô phụ (Split-Plot Design)

2.3.2.1. Đặc điểm

Kiểu ô chính ô phụ là kiểu bố trí cho thí nghiệm 2 nhân tố, mà mỗi lần nhắc lại được chia theo khối chính, sau đó khối chính được chia thành những ô nhỏ gọi là ô phụ. Số khối chính bằng số công thức của nhân tố chính. Số ô phụ (trong mỗi khối chính) bằng số công thức của nhân tố phụ. Việc bố trí các công thức của nhân tố chính vào các khối chính và các công thức của nhân tố phụ vào các ô theo phương pháp ngẫu nhiên như phần 2.2.2.

Thí nghiệm bố trí kiểu này có 2 loại sai số thí nghiệm: Sai số của nhân tố chính lớn hơn sai số của nhân tố nhỏ. Bởi vì, trong một lần nhắc lại, các ô phụ có sự đồng đều cao hơn các khối chính. Do đó, tuỳ vào mục đích của thí nghiệm để chọn nhân tố chính và nhân tố phụ.

- Các nguồn biến động trong thí nghiệm bố trí kiểu ô chính ô phụ bao gồm

- Biến động do công thức của nhân tố chính
- Biến động do công thức của nhân tố phụ
- Biến động do ngẫu nhiên của ô chính
- Biến động do ngẫu nhiên của ô phụ
- Biến động do tương tác 2 nhân tố
- Biến động do nhắc lại

- Ưu điểm:

- Cho phép sử dụng có hiệu quả khi một nhân tố có nhiều công thức được phối hợp với một nhân tố có ít công thức
- Cho phép sử dụng hiệu quả khi có một nhân tố cần quan tâm nhiều hơn đến so với nhân tố khác. Do vậy ta có thể bố trí nhân tố cần được quan tâm nhiều hơn vào ô phụ để làm tăng độ chính xác của nhân tố đó.
- Cho phép ta đưa ra những công thức thử nghiệm mới vào thí nghiệm đang tiến hành
- Thích hợp cho thí nghiệm nhiều công thức hơn là kiểu khối ngẫu nhiên hoàn chỉnh.

- Nhược điểm

- Giảm độ chính xác khi so sánh giữa các mức của yếu tố bố trí trong ô lớn
- Tính toán, xử lý thống kê phức tạp hơn. Yêu cầu tính hai loại sai số để so sánh. Một sai số dùng để so sánh giữa các mức của yếu tố chính, một sai số dùng để so sánh giữa các mức của yếu tố phụ.
- Độ chính xác của nhân tố chính thấp dẫn đến sai khác lớn trở nên không có ý nghĩa, vì vậy ảnh hưởng của nhân tố chính nhiều khi trở thành không có ý nghĩa. Trong khi đó sự sai khác nhỏ của nhân tố phụ lại cũng có thể trở thành có ý nghĩa về mặt thống kê nhưng trong thực tế không có ý nghĩa.

- Trường hợp áp dụng

- Khi một nhân tố dễ dàng bố trí được trên ô nhỏ trong khi một nhân tố khác thì không. Điều này thường xảy ra trong các thí nghiệm giống và phân bón. Bố trí các giống khác nhau trong từng ô nhỏ, dễ dàng hơn việc bón các mức phân khác nhau vào các ô nhỏ cần phải cách ly tốt các ô. Trong trường hợp này, bố trí phân bón vào ô chính (ô lớn) và giống vào ô phụ (ô nhỏ) thì sẽ dễ dàng hơn. Tương tự như vậy, với trường hợp thí nghiệm 2 nhân tố có nhân tố là thuốc trừ sâu, chế độ nước... là những yếu tố yêu cầu cách ly cao. Bố trí nhân tố cần cách ly vào các ô lớn thì sẽ giảm công cách ly rất nhiều.
- Vì kiểu thí nghiệm ô chia nhỏ sẽ giảm độ chính xác đối với việc đánh giá yếu tố bố trí trong ô lớn nhưng lại tăng độ chính xác khi đánh giá yếu tố bố trí trong ô phụ. Do vậy kiểu bố trí này thường được áp dụng cho: Những thí nghiệm mà người nghiên cứu cần quan tâm nhiều đến một yếu tố hơn yếu tố khác; Những thí nghiệm mà có một yếu tố có mức ảnh hưởng lớn dễ quan trắc và một yếu tố có mức ảnh hưởng nhỏ khó quan trắc

2.3.2.2. Cách bố trí

Khác với kiểu khối ngẫu nhiên đầy đủ, trong kiểu chia ô chính ô phụ kích thước ô thí nghiệm và độ chính xác của số liệu thu được không giống nhau trên cả hai yếu tố, do đó việc gán yếu tố nào là chính hay phụ rất quan trọng.

- Những căn cứ để quyết định nhân tố chính, nhân tố phụ

- *Mức độ chính xác yêu cầu đối với nhân tố thí nghiệm*

  Nhân tố yêu cầu độ chính xác cao hơn được chọn là nhân tố phụ. Ví dụ: Thí nghiệm "5 giống ngô mới với 3 mức độ phân bón", nhà di truyền quan tâm đến việc so sánh giữa các giống, nên đã thiết kế bố trí các giống vào ô phụ. Ngược lại, nhà nông học quan tâm nhiều đến ảnh hưởng của việc bón phân hơn, bố trí các mức phân bón vào ô phụ.

- *Tiên đoán mức độ ảnh hưởng của các nhân tố thí nghiệm:*

  Bố trí nhân tố có thể gây nên biến động lớn vào ô chính và nhân tố gây biến động nhỏ vào ô phụ. Điều này gia tăng cơ hội tìm ra sự khác biệt giữa các mức độ của nhân tố có ảnh hưởng nhỏ hơn nhân tố kia.

- *Yêu cầu phân cách của nhân tố thí nghiệm*

Nhân tố yêu cầu phân cách cao bố trí vào ô chính. Ví dụ, trong thí nghiệm hai nhân tố có nhân tố các mức nước tưới, hoặc mức độ phân bón, nên bố trí nhân tố nước hoặc phân bón vào ô chính để giảm việc xây dựng các bờ ngăn các lô có mức nước hoặc mức phân bón khác nhau.

- Cách bố trí

- Chia khu thí nghiệm thành khối. Mỗi khối sẽ là một lần nhắc lại.
- Chia mỗi lần nhắc lại thành ra các ô chính, số ô bằng số mức của nhân tố A. Bố trí ngẫu nhiên a mức của nhân tố A vào các ô chính.
- Chia mỗi ô chính thành các ô nhỏ, số ô nhỏ bằng số mức của nhân tố B. Bố trí ngẫu nhiên b mức của nhân tố B vào các ô nhỏ.
- Để minh họa chúng ta thiết kế cho thí nghiệm "Nghiên cứu phản ứng của 4 giống lúa với 6 mức phân bón khác nhau" với 3 lần nhắc lại.
- Chọn phân bón là nhân tố chính, giống là nhân tố phụ với 3 lần nhắc lại ta cần tổng số ô thí nghiệm là 72 ô (6 x 4 x 3 =72)

Bước 1:

Chia khu thí nghiệm ra làm 3 khối (3 lần nhắc lại) như sơ đồ 2.9.

Bước 2:

Mỗi khối chia ra thành 6 ô chính (vì có 6 mức N) và đánh số thứ tự các ô. Tổng số có 18 ô chính. Bố trí các mức N vào mỗi khối như thủ tục ngẫu nhiên ở phần 2.3.2

Bước 3:

Chia mỗi ô chính ra làm 4 ô phụ (vì có 4 giống) và đánh số thứ tự các ô phụ. Bố trí các mức giống vào mỗi ô chính theo thủ tục ngẫu nhiên như phần 2.3.2. Tổng số có 18 x 6 =72 ô phụ, chính là 72 ô thí nghiệm. Kết quả ta có sơ đồ bố trí theo kiểu ô chính ô phụ cho thí nghiệm như sau.

N4	N3	N1	N0	N5	N2
v2	v1	v1	v2	v4	v3
v1	v4	v2	v3	v3	v2
v3	v2	v4	v1	v2	v1
v4	v3	v3	v4	v1	v4

Nhắc lại lần I

N1	N0	N5	N2	N4	N3
v1	v4	v3	v1	v1	v3
v3	v1	v4	v2	v4	v2
v2	v2	v1	v4	v2	v4
v1	v3	v2	v3	v3	v1

Nhắc lại lần II

N0	N1	N4	N5	N3	N2
v4	v3	v3	v1	v2	v1
v2	v4	v2	v3	v3	v4
v1	v1	v4	v2	v4	v2
v3	v2	v1	v4	v1	v3

Nhắc lại lần III

Sơ đồ 2.10: Sơ đồ bố trí kiểu ô chính ô phụ 4 giống lúa với 6 mức đạm

2.3.3. Kiểu lô sọc (Strip plot design) (Phần đọc thêm)

2.3.3.1. Đặc điểm

Thí nghiệm 2 nhân tố bố trí theo kiểu lô sọc thường áp dụng trong những trường hợp quan tâm đánh giá ảnh hưởng của tương tác giữa hai yếu tố hơn việc khảo sát ảnh hưởng chính của từng yếu tố.

Kiểu bố trí này có 3 loại ô với kích thước khác nhau:

- Lô ngang cho nhân tố thứ nhất, vì thế còn gọi nhân tố này là nhân tố theo phương ngang (horizontal factor).
- Lô dọc cho nhân tố thứ hai, vì thế còn gọi nhân tố này là nhân tố theo phương dọc (vertical factor).
- Ô tạo nên do sự giao nhau giữa các lô ngang và lô dọc cho sự tương tác giữa 2 nhân tố

• Nguồn biến động trong kiểu bố trí này bao gồm:

- Biến động do nhắc lại
- Biến động của nhân tố theo phương ngang
- Biến động của nhân tố theo phương dọc
- Biến động do sai số ngẫu nhiên của nhân tố A
- Biến động do sai số ngẫu nhiên của nhân tố B
- Biến động do tương tác A và B
- Biến động do sai số ngẫu nhiên.

• Ưu điểm

Ô dọc và ngang luôn vuông góc với nhau, tuy nhiên quan hệ về kích thước giữa chúng không giống như quan hệ ô chính ô phụ trong kiểu bố trí ô chính, ô phụ. Ô giao nhau giữa 2 lô là nhỏ nhất, vì vậy sự tương tác của ảnh hưởng chính của 2 nhân tố được tách riêng, nhờ vậy tăng sự chính xác trong đánh giá tương tác giữa 2 nhân tố.

2.3.3.2. Cách bố trí

Theo kiểu bố trí này khu thí nghiệm được chia thành khối, sau đó mỗi khối sẽ được chia thành a (số mức của nhân tố A) lô ngang. Bố trí ngẫu nhiên các mức của nhân tố A vào lô ngang. Tiếp tục, chia khối đó thành b (số mức của nhân tố B) lô dọc. Bố trí ngẫu nhiên mức của nhân tố B vào lô dọc. Ô giao nhau giữa lô ngang và lô dọc là ô thí nghiệm.

Việc bố trí cũng tuân theo các thủ tục ngẫu nhiên như đã giới thiệu ở các kiểu bố trí trước.

Dưới đây là sơ đồ minh họa kiểu bố trí lô sọc cho thí nghiệm 2 giống lúa (A và B) với 3 mức phân (P1; P2; P3) và 3 lần nhắc lại.

	A	B
P2		
P1		
P3		

Khối I

	A	B
P2		
P3		
P1		

Khối II

	A	B
P3		
P1		
P2		

Khối III

Sơ đồ 2.11. Sơ đồ bố trí thí nghiệm kiểu lô sọc

2.4. THU THẬP SỐ LIỆU

Thí nghiệm được thực hiện theo thiết kế mới chỉ hoàn thành một phần công việc của thí nghiệm. Điều quan trọng tiếp theo là thu thập số liệu. Kỹ thuật thu thập số liệu ảnh hưởng quan trọng đến kết quả thí nghiệm. Bởi vì, các kết luận đều được đưa ra dựa trên số liệu thu thập được. Cách thu thập số liệu rất khác nhau phụ thuộc vào đối tượng và địa điểm nghiên cứu. Có 3 vấn đề quan trọng trong thu thập số liệu là: Chọn mẫu, đo đếm và ghi chép.

2.4.1. Chọn mẫu

Do không thể và không nhất thiết phải khảo sát tất cả phần tử của tổng thể nên người ta chọn mẫu. Kết quả thu thập được từ mẫu sẽ đại diện cho tổng thể. Vì vậy việc chọn mẫu ảnh hưởng rất quan trọng đến kết quả thí nghiệm.

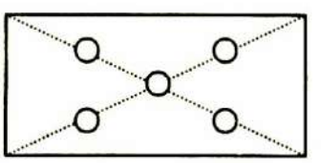
Nguyên tắc cơ bản của chọn mẫu là ngẫu nhiên. Ngẫu nhiên là nguyên tắc để áp dụng phương pháp phân tích thống kê trong xử lý số liệu. Ngẫu nhiên còn giúp loại bỏ hoàn toàn định hướng của người nghiên cứu.

Cách lấy mẫu cụ thể tuỳ thuộc vào đối tượng nghiên cứu và nội dung nghiên cứu. Vì vậy, với mỗi loại cây trồng, mỗi lĩnh vực có quy định cụ thể về cách lấy mẫu. Ví dụ trong nghiên cứu đậu đỗ quy định như sau: Quan trắc các đặc điểm của cây chọn 10 cây trên 2 hàng giữa luống. Mỗi hàng lấy 5 cây liên tục, trừ cây ngoài rìa. Sau đây là một số gợi ý về cách lấy mẫu để mẫu đại diện và mang tính ngẫu nhiên.

Lấy ngẫu nhiên bằng cách đánh số và gắp thăm

Chọn cây trung bình ở những chỗ điển hình của thí nghiệm

Chọn theo quy định máy móc theo đường chéo 5, 4, 3, 2 điểm, hoặc theo một khoảng cách nhất định.



Sơ đồ 2.12. Các vị trí lấy mẫu theo đường chéo 5 điểm

Xác định cỡ mẫu cũng quy định tuỳ thuộc vào đối tượng cụ thể và yêu cầu của thí nghiệm. Về nguyên tắc, mẫu nhỏ quá thì kết quả ước lượng từ mẫu kém chính xác và không dùng được. Ngược lại mẫu lớn quá đưa đến chi phí cao.

Một thông số để đo lường độ chính xác của mẫu được chọn là độ lệch chuẩn. Nếu độ lệch chuẩn mẫu càng lớn thì độ tin cậy của việc dùng ước lượng từ mẫu để suy diễn cho tổng thể càng kém.

2.4.2. Đo đếm

Việc đo đếm thực hiện theo quy trình của từng đối tượng nghiên cứu. Số liệu đo đếm theo từng ô thí nghiệm và phải được ghi riêng theo ô. Tuyệt đối không được để lẫn hay trộn kết quả của các lần nhắc lại của một công thức.

2.4.3. Ghi chép số liệu: Số liệu thô và số liệu tinh

Số liệu thô là số liệu đo đếm được ngoài đồng ruộng và ghi cụ thể kết quả của từng mẫu theo dõi và theo từng ô. Số liệu thô cần ghi theo ô, nhắc lại và công thức. Ví dụ để biết được chiều cao của cây ngô trong thí nghiệm người ta đo riêng từng ô thí nghiệm, mỗi ô đo 10 cây và ghi kết quả theo từng lần nhắc lại như sau:

Bảng 2.8: Chiều cao cây của 2 giống ngô (m) ngày đo 2/5/2003

Nhắc lại I

Cây số	1	2	3	4	5	6	7	8	9	10	TB
Giống A	0,9	0,8	0,8	1,2	1,1	0,6	0,9	1,0	0,7	1,1	0,9
Giống B	2,3	2,2	2,3	1,9	1,6	1,7	2,0	2,3	1,8	2,0	2,0

Nhắc lại II

Cây số	1	2	3	4	5	6	7	8	9	10	TB
Giống A	0,7	0,6	0,8	1,2	1,1	0,9	0,9	1,2	0,7	1,1	0,9
Giống B	2,2	1,7	2,3	1,6	1,6	1,7	1,9	2,5	1,8	2,0	1,9

Nhắc lại III

Cây số	1	2	3	4	5	6	7	8	9	10	TB
Giống A	0,7	0,8	1,2	1,2	1,1	0,6	0,9	1,1	0,7	1,1	1,0
Giống B	2,3	2,1	2,3	1,9	1,6	1,8	1,9	1,8	1,8	2,0	1,9

Các số liệu có được sau khi tính toán từ số liệu thô được gọi là số liệu tinh. Bảng 2.9 là kết quả tính trung bình chiều cao cây ngô của từng lần nhắc lại và của 3 lần nhắc lại. Khi xử lý thống kê ta dùng số liệu trung bình của các lần nhắc lại như bảng 2.9.

Bảng 2.9: Chiều cao cây trung bình của 2 giống ngô (m) ngày đo 2/5/2003

Giống	Nhắc lại I	Nhắc lại II	Nhắc lại III	Trung bình
A	0,9	0,9	1,0	0,9
B	2,0	1,9	1,9	1,9

Nhìn vào bảng số liệu trên rất khó nhận định kết quả. Giống nào cao hơn giống nào? Để trả lời được câu hỏi trên cần phải chuyển số liệu thô thành số liệu tinh. Trong báo cáo kết quả chỉ trình bày kết quả cuối cùng tức là trung bình của các lần nhắc lại như bảng 2.10 dưới đây.

Bảng 2.10. Chiều cao của các giống ngô (m)

Giống	Chiều cao cây (m)
A	0,9
B	1,9

2.4.4. Một số quy tắc cần biết trong tính toán (Phạm Chí Thành, 1988)

2.4.4.1. Con số có nghĩa

Những số được xem như chính xác là những số cần thiết cho việc diễn tả mà không kể số 0 để đặt dấu (,). Ví dụ con số chính xác 1/10 là 65,4. Con số 65,4 có 3 con số có nghĩa; số 4,530 có 4 con số có nghĩa; số 0,018 có 2 con số có nghĩa, mà 0,0 chỉ dùng để đặt dấu (,).

2.4.4.2. Phép tính gần đúng

Quy tắc: Trong các phép nhân, chia, khai căn thì kết quả cuối cùng không thể có nhiều con số có nghĩa hơn con số có nghĩa ở con số cần tính có ít nhất.

Ví dụ: 73,24 x 4,52 =331; kết quả này chỉ lấy đến 3 con số có nghĩa vì số 4,52 chỉ có 3 con số có nghĩa.

1,684: 0,023 = 72; kết quả này chỉ lấy đến 2 con số có nghĩa vì số 0,023 chỉ có 2 con số có nghĩa.

Quy tắc: Trong các phép cộng, trừ thì số có nghĩa (kể từ sau dấu (,) của kết quả cuối cùng không thể có nhiều hơn số có nghĩa (kể từ sau dấu (,) ở số hạng có ít số có nghĩa nhất.

Ví dụ: 3,16 + 2,7 = 5,9; kết quả này chỉ có 1 số có nghĩa kể từ sau dấu (,) vì số 2,7 chỉ có 1 số có nghĩa kể từ sau dấu (,).

83,12 -72 = 11 vì sau dấu (,) của 72 không có số có nghĩa.

Chương 3

PHÂN TÍCH BIẾN ĐỘNG

3.1. Ý NGHĨA CỦA PHÂN TÍCH BIẾN ĐỘNG VÀ HỆ SỐ BIẾN ĐỘNG

3.1.1. Phân tích biến động

Số liệu thu thập được trong quá trình làm thí nghiệm cho thấy các công thức khác nhau cho kết quả khác nhau. Phân tích biến động là dùng phương pháp thống kê để kiểm tra xem sự sai khác trong kết quả thí nghiệm có phải do nhân tố thí nghiệm hay không. Bằng phương pháp thống kê ta ước lượng được sai khác do các nguồn biến động khác nhau gây ra. Nếu như sai khác có ý nghĩa thì kết luận sự sai khác trong kết quả là do nhân tố thí nghiệm. Sai khác giữa các công thức trong thí nghiệm đồng ruộng được tính theo công thức của phương sai mẫu, vì vậy phân tích biến động (hay sai khác) còn gọi là phân tích phương sai.

Sai khác do yếu tố thí nghiệm gây ra được gọi là biến động do công thức và sai khác do yếu tố ngẫu nhiên gây ra gọi là biến động ngẫu nhiên. Yếu tố phi thí nghiệm được gọi là yếu tố ngẫu nhiên, do đó sai khác gây ra do yếu tố ngẫu nhiên được gọi là sai số ngẫu nhiên. Với những kiểu bố trí thí nghiệm khác nhau thì nguồn biến động cũng khác nhau, vì vậy cách phân tích biến động phụ thuộc vào cách bố trí thí nghiệm.

Để kiểm tra sai khác có ý nghĩa hay không ta so sánh sai khác do yếu tố thí nghiệm gây ra với sai khác do ngẫu nhiên gây ra bằng cách tính giá trị F thực nghiệm hay còn gọi là F tính. So sánh F thực nghiệm với F lý thuyết (tra từ bảng tính sẵn E) để kết luận chấp nhận hay bác bỏ giả thuyết của thí nghiệm. Hay có thể nói so sánh F thực nghiệm với F lý thuyết ta có thể kết luận sai khác trong kết quả giữa các công thức thí nghiệm có ý nghĩa hay không. F thực nghiệm được ký hiệu là F_t; F lý thuyết được ký hiệu là F_b. Khi so sánh:

Nếu $F_t > F_b$: Kết luận sai khác giữa các công thức có ý nghĩa. Có nghĩa sự khác nhau là do yếu tố thí nghiệm gây nên.

Nếu $F_t \leq F_b$: Kết luận sai khác giữa các công thức không có ý nghĩa. Có nghĩa các công thức khác nhau nhưng cho kết quả như nhau. Trong bảng phân tích phương sai đánh dấu "ns" vào giá trị của F_t (*n.s là viết tắt của not significant - không có ý nghĩa*).

Giá trị của F lý thuyết thay đổi tuỳ thuộc vào mức có ý nghĩa α. Mức có ý nghĩa α hay được dùng là $\alpha = 0,05$ và $\alpha = 0,01$. Từ mức có ý nghĩa ta biết được độ tin cậy.

Độ tin cậy = (100 - α). Như vậy, với mức có ý nghĩa α = 0,05 thì độ tin cậy là là 95% hay còn nói sai khác có ý nghĩa ở mức xác suất 95%. Với α = 0,01 thì độ tin cậy là 99% hay còn nói sai khác có ý nghĩa ở mức xác suất 99%. F bảng tra ở mức α = 0,05 ký hiệu là F_{05}; tra ở mức α = 0,01 ký hiệu là F_{01}. Khi so sánh:

Nếu $F_t > F_{01}$: Kết luận có sự sai khác rất chắc chắn giữa các công thức thí nghiệm. Có nghĩa sự sai khác giữa các công thức thí nghiệm ở mức độ tin cậy 99%. Và trong bảng phân tích phương sai đánh dấu dấu "**" vào giá trị của F_t

Nếu $F_{01} > F_t > F_{05}$: Kết luận có sự sai khác chắc chắn giữa các công thức thí nghiệm. Có nghĩa là sự sai khác giữa các công thức thí nghiệm ở mức độ tin cậy 95%. Và đánh dấu "*" vào giá trị của F_t trong bảng phân tích phương sai.

3.1.2. Hệ số biến động

Sau khi phân tích biến động cần tính hệ số biến động. Hệ số biến động được ký hiệu là cv (cv - coefficient of variation). cv thường gọi là sai số thí nghiệm. Hệ số biến động cv chỉ ra độ chính xác của thí nghiệm, nó chỉ ra sai số thí nghiệm chiếm bao nhiêu % số trung bình tổng. Vậy cv càng nhỏ thì nghiệm càng chính xác. Giá trị của cv có thể chấp nhận được khác nhau giữa các đối tượng nghiên cứu, địa bàn nghiên cứu. Hệ số biến động được tính theo công thức:

$$\text{Hệ số biến động cv} = \frac{\sqrt{MS_E}}{\overline{G}} \times 100$$

$$\text{Trung bình tổng } \overline{G} = \frac{G}{n}$$

Theo Gomez (1984), thường cv được chấp nhận đối với năng suất của lúa cấy thay đổi phụ thuộc nhân tố thí nghiệm như sau: 6-8% với thí nghiệm giống; 10-12% với thí nghiệm về phân bón; 13-15% với thí nghiệm về thuốc trừ cỏ và trừ sâu. Chỉ tiêu theo dõi khác nhau thì giá trị chấp nhận được của cv cũng khác nhau: Với năng suất lúa cv là 10%; Với số nhánh của lúa cv là 20%.

Quy ước các ký hiệu trong phân tích phương sai:

ANOVA - Analysis of Variation: Phân tích phương sai

CF - Correction Factor: Số hiệu chỉnh

df - Degree of freedom: độ tự do

F_t - Compute F - F: thực nghiệm hoặc F tính

F_b - Tabular F - F: bảng (tra từ bảng tính sẵn)

G- Grand: Tổng toàn bộ hoặc tổng lớn

MS - Mean Square: Trung bình bình phương

SS - Sum of Square: Theo định nghĩa, phương sai được tính bằng bình phương sự sai khác giữa x_i và $\bar{x}$ vì vậy tổng phương sai được gọi là tổng bình phương

SS_e - Error sum of square: Tổng bình phương của ngẫu nhiên

SS_t - Treatment sum of square: Tổng bình phương của công thức

TSS - Total sum of square: Tổng bình phương toàn bộ.

3.2. PHÂN TÍCH BIẾN ĐỘNG CHO THÍ NGHIỆM MỘT NHÂN TỐ

3.2.1. Kiểu bố trí ngẫu nhiên hoàn toàn (Completely Randomized Design - CRD)

3.2.1.1. Trường hợp các công thức có lần nhắc lại bằng nhau

Thí nghiệm bố trí theo kiểu ngẫu nhiên hoàn toàn có 2 nguồn biến động là: biến động do công thức và biến động do ngẫu nhiên. Tổng biến động của thí nghiệm bố trí kiểu ngẫu nhiên hoàn toàn được tính theo công thức:

$TSS = SS_t + SS_e$

Trong đó:

SS_t là tổng bình phương biến động do công thức

SS_e là tổng bình phương biến động do ngẫu nhiên

TSS là tổng bình phương biến động toàn bộ

Bảng 3.1. Bảng phân tích biến động (ANOVA) cho thí nghiệm bố trí theo kiểu ngẫu nhiên hoàn toàn

Nguồn biến động	Độ tự do (df)	Tổng bình phương (SS)	Trung bình của bình phương (MS)	F_t	$F_{bảng}$
Công thức Ngẫu nhiên Tổng	df_t df_e Tdf	SS_t SS_e TSS	$MS_t=SS_t/t-1$ $MS_e=SS_e/t(r-1)$	$F_t= MS_t/MS_e$	F_{05} F_{01}

Tính độ tự do (ký hiệu là df) của từng nguồn biến động theo các công thức:

Tổng df $\quad$ Tdf= (r)(t)-1

df của công thức: $\quad df_t = t-1$

df của sai số ngẫu nhiên: $\quad df_e = t(r-1) = Tdf- df_t$

Tính số hiệu chỉnh và các tổng bình phương biến động theo các công thức:

Số hiệu chỉnh (CF- correction factor): $CF = \dfrac{G^2}{n}$

Tổng bình phương của toàn bộ: $TSS = \sum_{i=1}^{n} X_i^2 - CF$

Tổng bình phương do công thức: $SS_t = \frac{\sum_{i=1}^{t} T_i^2}{r} - CF$

Tổng bình phương do sai số ngẫu nhiên: $SS_e = TSS - SS_t$

Ghi chú:

$\sum_{i=1}^{n} X_i^2$ có nghĩa là bình phương từng giá trị của Xi sau đó cộng lại. Vì có n ô thí nghiệm nên ta có n giá trị X; hay là có X_1 đến X_n.

$\sum_{i=1}^{t} T_i^2$ có nghĩa là bình phương "tổng của mỗi công thức (T_i)" sau đó cộng lại. Vì có số công thức = t nên ta có t giá trị T; hay là có X_1 đến X_t

Các bước phân tích được minh họa bằng phân tích phương sai cho năng suất lúa của thí nghiệm "So sánh hiệu quả của các loại thuốc trừ rày nâu ảnh hưởng tới năng suất của lúa", bố trí kiểu ngẫu nhiên hoàn toàn. Kết quả thí nghiệm được trình bày ở bảng 3.2.

Thí nghiệm có:

7 công thức t = 7 (7 loại thuốc trừ rày)

4 lần nhắc lại (r = 4).

Tổng số ô: n = (t)(r) = 7 × 4 = 28.

Bước 1:

Nhóm số liệu theo công thức; Tính tổng theo công thức (ký hiệu là T - total); và tổng của toàn bộ (ký hiệu là G - grand) các giá trị quan sát được (như bảng 3.2).

Bảng 3.2. Kết quả thí nghiệm so sánh hiệu quả của các loại thuốc trừ rày nâu đến năng suất lúa (kg/ha)

Công thức	Năng suất (X_i)				Tổng CT (T_i)	Trung bình CT
1. Dol-Mix (1 kg)	2537	2069	2104	1797	8507	2127
2. Dol-Mix (2 kg)	3366	2591	2211	2544	10712	2678
3. DDT+BHC	2536	2459	2827	2385	10207	2552
4. Azodrin	2387	2453	1556	2116	8512	2128
5. Dimecron-Boom	1997	1679	1649	1859	7184	1796
6. Dimecron-Knap	1796	1704	1904	1320	6724	1681
7. Đối chứng	1401	1516	1270	1077	5264	1316
Tổng toàn bộ (G)					57110	
T.bình toàn bộ						2040

Nguồn: Gomez, 1984.

X_i là năng suất của các ô thí nghiệm, có 28 ô thí nghiệm vì vậy ta có 28 giá trị X.

Bước 2: Tính độ tự do - df của từng nguồn biến động:

Tổng df Tdf = (r)(t)-1 = (4)(7)-1 = 27

df của công thức: df_t = t-1 = 7-1 = 6

df của sai số ngẫu nhiên: df_e = t(r-1) = 7(4-1) = Tdf- df_t = 27-6 = 21

Bước 3: Tính số hiệu chỉnh và các tổng bình phương biến động:

$$CF = \frac{G^2}{n} = \frac{(57.110)^2}{(4)(7)} = 116.484.004$$

Tổng bình phương của toàn bộ:

$$TSS = \sum_{i=1}^{n} X_i^2 - CF = TSS = [(2537)^2 + (2069)^2 + ... + (1270)^2 + (1077)^2] - 116.484.004$$

$$= 7.577.412$$

Tổng bình phương do công thức:

$$SS_t = \frac{\sum_{i=1}^{t} T_i^2}{r} - CF = \frac{(8507)^2 + (10712)^2 + ... + (5264)^2}{4} - 116.484.004 = 5.587.174$$

Tổng bình phương của sai số ngẫu nhiên:

$$SS_e = TSS - SS_t = 7.577.412 - 5.587.174 = 1.990.238$$

Bước 4: Tính trung bình bình phương của các biến động ứng với độ tự do của nó.

Trung bình bình phương do công thức:

$$MS_t = \frac{SS_t}{t-1} = \frac{5.587.174}{6} = 931.196$$

Trung bình bình phương do sai số ngẫu nhiên:

$$MS_e = \frac{SS_e}{t(r-1)} = \frac{1.990.238}{(7)(3)} = 94.773$$

Bước 5:

Tính giá trị F để kiểm tra sự sai khác của thí nghiệm có ý nghĩa hay không

$$\text{Tính } F_{tính} = \frac{MS_t}{MS_e} = \frac{931.196}{94.773} = 9,83$$

Bước 6: Tra F lý thuyết

Tra F lý thuyết ở bảng E (phụ lục) với: $f_1 = df_t$ = (t-1) = 6 ; $f_2 = df_e$ = t(r-1) = 21

Ta có F_{05} = 2,57; F_{01} = 3,81

Bước 7: So sánh F_t với F_b

$F_t > F_{01}$ như vậy sự sai khác giữa các công thức có ý nghĩa. Có nghĩa là sự khác nhau về năng suất lúa là do nhân tố thí nghiệm gây nên. Sai khác ở mức có ý nghĩa.01 nên ta đánh ** vào F_t.

Bước 8:

Tính hệ số biến động:

$$\overline{G} = \frac{G}{n} = \frac{57110}{28} = 2040$$

$$cv = \frac{\sqrt{MS_e}}{\overline{G}}.100 = \frac{\sqrt{94773}}{2040}.100 = 15,1\ \%$$

Với yêu cầu cv của thí nghiệm về thuốc trừ cỏ và trừ sâu cho lúa là 13-15% thì thí nghiệm đảm bảo độ chính xác theo yêu cầu.

Bước 9: Lập bảng phân tích phương sai và ghi các giá trị tính được.

Bảng 3.3. Kết quả phân tích biến động của năng suất lúa trong thí nghiệm thuốc trừ rày của thí nghiệm bố trí kiểu ngẫu nhiên hoàn toàn

Nguồn biến động	Độ tự do (df)	Tổng bình phương (SS)	Trung bình của bình phương (MS)	F_t	$F_{bảng}$	
					5%	1%
Công thức	6	5.587.174	931.196	9,83**	2,57	3,81
Ngẫu nhiên	21	1.990.238	94.773			
Tổng	27	7.577.412				

cv = 15,1%; **: có ý nghĩa ở 1%;

Kết luận: $F_t > F_{01}$ các loại thuốc trừ rày khác nhau cho năng suất lúa khác nhau rất chắc chắn, ở mức xác suất 99%.

3.2.1.2. Trường hợp các công thức có lần nhắc lại khác nhau

Có những trường hợp số lần nhắc lại của các công thức khác nhau, ví dụ như hỏng vài ô do điều kiện thời tiết hoặc gia súc phá. Trong trường hợp này cách phân tích phương sai vẫn như trường hợp có số lần nhắc lại như nhau. Khi đó n và x_i chính là số quan sát thực tế có được. Ví dụ, với thí nghiệm ghi ở bảng 3.2. nếu mất 4 số liệu thì n = 24 và tổng các x_i sẽ tính được từ 24 giá trị chứ không phải là 28 như thiết kế ban đầu.

3.2.2. Kiểu khối ngẫu nhiên hoàn chỉnh

Thí nghiệm bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh có 3 nguồn biến động là: biến động do công thức, biến động do ngẫu nhiên và biến động theo khối.

Tổng biến động của thí nghiệm bố trí kiểu khối ngẫu nhiên hoàn chỉnh tính theo công thức:

$$TSS = SS_t + SS_c + SS_r$$

Trong đó

SS_r là tổng bình phương biến động của khối (theo nhắc lại)

SS_t là tổng bình phương biến động của công thức

SS_c tổng bình phương biến động của ngẫu nhiên

Bảng 3.4. Phân tích biến động (ANOVA) cho thí nghiệm bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh

Nguồn biến động	df	SS	MS	$F_{tính}$	$F_{bảng}$
Nhắc lại (khối)	df_r	SS_r	$MS_r=SS_r/r-1$	$F_r= MS_r/MS_e$	
Công thức	df_t	SS_t	$MS_t=SS_t/t-1$	$F_t= MS_t/MS_e$	
Ngẫu nhiên	df_e	SS_e	$MS_e=SS_e/(t-1)(r-1)$		
Tổng	Tdf	TSS			

Độ tự do, CF và các tổng bình phương biến động tính như phần 3.2.1.1. Riêng độ tự do và tổng bình phương biến động của nhắc lại được tính theo công thức:

Độ tự do của nhắc lại: $df_r= r-1$

$$SS_r = \frac{\sum_{j=1}^{r} R_j^2}{t} - CF$$

Ghi chú:

$\sum_{j=1}^{r} R_j^2$ có nghĩa là bình phương "tổng của mỗi lần nhắc lại (Ri)" sau đó cộng lại. Vì có số lần nhắc lại r = j nên ta có r giá trị R; hay là có R_1 đến R_j.

Các bước phân tích phương sai được minh họa qua phân tích phương sai cho năng suất của thí nghiệm "ảnh hưởng của lượng hạt giống gieo tới năng suất lúa". Thí nghiệm bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh. Kết quả thí nghiệm được trình bày ở bảng 3.5. Thí nghiệm có:

6 công thức (t = 6), (6 lượng hạt giống gieo)

4 lần nhắc lại (r = 4),

Tổng số ô: n = (t)(r) = 24

Bước 1:

Nhóm số liệu theo công thức và lần nhắc lại. Tính tổng của công thức (T) của nhắc lại (R) và tổng toàn bộ (G) như ở bảng 3.5.

Bảng 3.5. Năng suất lúa của giống IR8 với 6 lượng hạt giống gieo trong thí nghiệm RCB, 4 lần nhắc lại

Công thức (kg hạt/ha)	Năng suất (kg/ha) (X_i)				Tổng công thức (T)	Trung bình công thức
	I	II	III	IV		
1. (25)	5113	5398	5307	4678	20496	5124
2. (50)	5346	5952	4719	4264	20281	5070
3. (75)	5272	5713	5483	4749	21217	5304
4. (100)	5164	4831	4986	4410	19391	4848
5. (125)	4804	4848	4432	4748	18832	4708
7. (150)	5254	4542	4919	4098	18813	4703
Tổng n.lại (R)	30953	31284	29846	26947		
Tổng lớn (G)					119030	
TB tổng						4960

(Nguồn: Gomez, 1984)

Bước 2:

Tính độ tự do df của từng nguồn biến động:

Tổng df = rt-1 = 24-1 = 23

df của nhắc lại df_r = r-1 = 4-1 = 3

df của công thức df_t = t-1 = 6-1 = 5

df của sai số ngẫu nhiên df_e = (r-1)(t-1) = (rt-1)-(r-1)-(t-1) = (3)(5) = 15

Bước 3: Tính số hiệu chỉnh và các tổng bình phương biến động:

$$CF = \frac{G^2}{rt} = \frac{(119.030)^2}{(4)(6)} = 590.339.204$$

Tổng bình phương của toàn bộ:

$$TSS = \sum_{i=1}^{t}\sum_{i=1}^{r} X_{ij}^2 - CF$$

$$= \left[(5.113)^2 + (5.398)^2 + ... + (4.098)^2\right] - 590.339.204 = 4.801.068$$

$\sum_{i=1}^{t}\sum_{i=1}^{r} X_{ij}^2$ có nghĩa là tổng "bình phương của các X" tại công thức thứ i và nhắc lại thứ j *(i: số thứ tự của công thức; j: số thứ tự của nhắc lại)*

Tổng bình phương của nhắc lại:

$$SS_r = \frac{\sum_{j=1}^{r} R_j^2}{t} - CF$$

$$= \frac{(30.953)^2 + (31.284)^2 + (29.846)^2 + (26.947)^2}{6} - 590.339.204 = 1.944.361$$

Tổng bình phương của công thức:

$$SS_t = \frac{\sum_{i=1}^{t} T_i^2}{r} - CF$$

$$= \frac{(20.496)^2 + ... + (18.813)^2}{4} - 590.339.204 = 1.198.331$$

Tổng bình phương của ngẫu nhiên:

$$SS_e = TSS - SS_r - SS_t = 4.801.068 - 1.944.361 - 1.198.331 = 1.658.376$$

Bước 4:

Tính trung bình bình phương của các nguồn biến động ứng với độ tự do của nó.

Trung bình bình phương của nhắc lại:

$$MS_r = \frac{SS_r}{(r-1)} = \frac{1.944.361}{3} = 648.120$$

Trung bình bình phương của công thức:

$$MS_t = \frac{SS_t}{(t-1)} = \frac{1.198.331}{5} = 239.666$$

Trung bình bình phương của ngẫu nhiên:

$$MS_e = \frac{SS_e}{(r-1)(t-1)} = \frac{1.658.376}{15} = 110.558$$

Bước 5: Tính giá trị của F để kiểm tra sự sai khác

$$F_{tính} = \frac{MS_t}{MS_e} = \frac{239.666}{110.558} = 2,17$$

Bước 6: So sánh F tính và F bảng

Tra bảng E với: $f_1 = df_t = 5$; $f_2 = df_e = 15$;

Ta được: $F_{05} = 2,90$ và $F_{01} = 4,56$.

Vậy $F_t < F_b$: Kết luận sai khác giữa các công thức thí nghiệm không có ý nghĩa. Có nghĩa là sự khác nhau về năng suất lúa là do ngẫu nhiên chứ không phải do công thức thí nghiệm.

Bước 7: Tính hệ số biến động

$$cv = \frac{\sqrt{MS_e}}{\overline{G}} \times 100 = \frac{\sqrt{110.558}}{4.960} \times 100 = 6,7\%$$

Kết luận: Thí nghiệm đảm bảo chính xác.

Bước 8: Lập bảng phân tích biến động

Bảng 3.6. Kết quả phân tích biến động của năng suất lúa trong thí nghiệm 6 mật độ bố trí kiểu khối ngẫu nhiên hoàn chỉnh

Nguồn biến động	df	SS	MS	F_t	F_b	
					5%	1%
Nhắc lại	3	1.944.361	648.120	5,86	3,29	5,42
Công thức	5	1.198.331	239.666	$2,17^{ns}$	2,90	4,56
Ngẫu nhiên	15	1.658.376	110.558			
Tổng	23	4.801.068				

cv = 6,7%; (ns): not significant- không có ý nghĩa;

Kết luận: Các công thức gieo lượng hạt giống khác nhau không ảnh hưởng đến năng suất. Có nghĩa với những lượng hạt giống gieo khác nhau cho năng suất lúa như nhau. Vậy nhà nghiên cứu sẽ khuyến cáo sản xuất nên dùng 25 kg hạt/ha.

Đánh giá hiệu quả của chia khối

Mục đích của chia khối là làm tăng tối đa sự khác nhau giữa các khối, nhờ vậy giảm tối thiểu sự khác nhau giữa các ô trong một khối. Từ kết quả chúng ta có thể đánh giá việc chia khối đã áp dụng có đạt được mục đích không. Muốn vậy, cần tính F_r của khối và so sánh với F_b. Dưới đây là cách tính toán nhằm kiểm định hiệu quả của khối.

Tính F_r

$$F_r = \frac{MS_r}{MS}$$

So sánh F_r tính với F bảng.

Tra F ở bảng E với: $f_1 = r-1$; $f_2 = (r-1)(t-1)$

Nếu $F_r > F_b$ kết luận khối có hiệu quả làm giảm sai số ngẫu nhiên

Với số liệu của bảng 3.5 và kết quả của bảng 3.5 ta có:

$$F_r = \frac{648,120}{110,558} = 5,86$$

Tra F_b với: $f_1 = 3$ và $f_2 = 15$; có $F_{.05} = 3,29$ và $F_{.01} = 5,42$

$F_r = 5,68 > F_{.01}$.

Kết luận sự khác nhau giữa các khối có ý nghĩa rất chắc chắn. Điều đó có nghĩa việc chia khối đã làm tăng độ chính xác của thí nghiệm.

3.2.3. Kiểu ô vuông la tinh (Latin Square Design - LS)

Thí nghiệm bố trí kiểu ô vuông la tinh có 4 nguồn biến động là: biến động do công thức, do ngẫu nhiên, do hàng và do cột. Tổng phương sai (TSS) của thí nghiệm bố trí kiểu ô vuông la tinh được tính theo công thức:

$$TSS = SS_t + SS_r + SS_c + SS_e$$

Trong đó

SS_r là tổng bình phương biến động theo cột,

SS_c là tổng bình phương biến động theo hàng,

SS_t là tổng bình phương biến động do công thức,

SS_e là tổng bình phương biến động do ngẫu nhiên.

Bảng 3.7. Phân tích biến động (ANOVA) cho thí nghiệm bố trí theo kiểu ô vuông la tinh

Nguồn biến động	df	SS	MS	$F_{tính}$ [b]	$F_{bảng}$
Hàng	df_r	SS_r	$MS_r=SS_r/t-1$	$F_r= MS_r/MS_e$	
Cột	df_c	SS_c	$MS_c=SS_c/t-1$	$F_c= MS_c/MS_e$	
Công thức	df_t	SS_t	$MS_t=SS_t/t-1$	$F_t= MS_t/MS_e$	
Ngẫu nhiên	df_e	SS_e	$MS_e=SS_e/(t-1)(t-2)$		
Tổng	Tdf	TSS			

Cách tính độ tự do, CF và tổng bình phương biến động của công thức và ngẫu nhiên giống như cách tính của các kiểu bố trí khác, riêng độ tự do và tổng bình phương biến động của hàng và cột được tính như sau:

$$df_r = df_c = df_t$$

$$df_c = (t-1)(t-2) = Tdf - df_r - df_c - df_t$$

$$SS_r = \frac{\sum R_i^2}{t} - CF$$

$$SS_c = \frac{\sum C_i^2}{t} - CF$$

$$SS_e = TSS - SS_r - SS_c - SS_t$$

Trong đó:

df_r: Độ tự do của hàng

df_c: Độ tự do của cột

SS_r: Tổng bình phương biến động của hàng

SS_c: Tổng bình phương biến động của cột

Minh họa các bước phân tích phương sai qua phân tích phương sai của năng suất lúa trong thí nghiệm "So sánh năng suất các giống lúa". Thí nghiệm có 4 giống: A,B,C (đối chứng) và D; bố trí theo kiểu ô vuông la tinh. Kết quả trình bày ở bảng 3.8. Thí nghiệm có t = r = 4;

Bước 1: Lập bảng số liệu theo hàng, cột và công thức như bảng 3.8.

Bước 2:

Tính tổng của hàng (R), tổng của cột (C) và tổng toàn bộ (G). Tính tổng của công thức (T) và trung bình của công thức như sau:

Bảng 3.8: Kết quả thí nghiệm năng suất của 4 giống lúa bố trí kiểu LS

Số thứ tự hàng	Năng suất hạt (tấn/ha) (X_i)				Tổng hàng (R)
	Cột 1	Cột 2	Cột 3	Cột 4	
Hàng 1	1,640(B)	1,210(D)	1,425(C)	1,345(A)	5,620
2	1,475(C)	1,185(A)	1,400(D)	1,290(B)	5,350
3	1,670(A)	0,710(C)	1,665(B)	1,180(D)	5,225
4	1,565(D)	1,290(B)	1,655(A)	0,660(C)	5,170
Tổng cột (C)	6,350	4,395	6,145	4,475	
Tổng lớn (G)					21,365

Tính trung bình năng suất của các công thức:

Công thức	Tổng số	Trung bình
A	5,855	1,464
B	5,885	1,471
C	4,270	1,068
D	5,355	1,339

Bước 3: Tính độ tự do của nguồn biến động, với t là số công thức theo công thức:

$Tdf = t^2 - 1 = 16 - 1 = 15$

$df_r = df_c = df_t = t-1 = 4-1 = 3$

$df_e = (t-1)(t-2) = (4-1)(4-2) = Tdf - df_r - df_c - df_t = 15-3-3-3 = 6$

Bước 4: Tính số hiệu chỉnh và các tổng bình phương:

$$CF = \frac{G^2}{t^2} = \frac{(21,365)^2}{16} = 28,528952$$

Tổng bình phương của toàn bộ

$$TSS = \sum X_i^2 - CF = \left[(1,640)^2 + (1,210)^2 + ... + (0,660)^2\right] - 28,5289 = 1,4139$$

Tổng bình phương của hàng

$$SS_r = \frac{\sum R_i^2}{t} - CF$$

$$= \frac{(5,620)^2 + (5,350)^2 + (5,225)^2 + (5,170)^2}{4} - 28,528 = 0,0301$$

Tổng bình phương của cột

$$SS_c = \frac{\sum C_i^2}{t} - CF = \frac{(6,350)^2 + (4,395)^2 + (6,145)^2 + (4,475)^2}{4} - 28,528 = 0,827$$

Tổng bình phương của công thức

$$SS_t = \frac{\sum T_i^2}{t} - CF = \frac{(5,855)^2 + (5,885)^2 + (4,270)^2 + (5,355)^2}{4} - 28,528 = 0,426$$

Tổng bình phương của ngẫu nhiên

$$SS_e = TSS - SS_r - SS_c - SS_t = 1,413 - 0,030 - 0,827 - 0,426 = 0,129$$

Bước 5:

Tính trung bình bình của từng nguồn bằng cách chia tổng phương sai cho độ tự do tương ứng.

$$MS_r = \frac{SS_r}{t-1} = \frac{0,030}{3} = 0,010$$

$$MS_c = \frac{SS_c}{t-1} = \frac{0,827}{3} = 0,275$$

$$MS_t = \frac{SS_t}{t-1} = \frac{0,426}{3} = 0,142$$

$$MS_e = \frac{SS_e}{(t-1)(t-2)} = \frac{0,129}{(3)(2)} = 0,021$$

Bước 6: Tính F để kiểm tra sự sai khác của thí nghiệm

$$Ft = \frac{MS_t}{MS_e} = \frac{0,142}{0,021} = 6,59$$

Bước 7: So sánh F tính với F bảng.

Tra F ở bảng E với: $f_1 = df_t = t - 1 = 3$ và ; $f_2 = df_E = (t-1)(t-2) = 6$

Ta có: $F_{05} = 4,76$; $F_{01} = 9,78$

Vì $F_{b05} < F_{tính} < F_{b01}$, kết luận sai khác giữa các công thức thí nghiệm là có ý nghĩa. Có nghĩa là sự khác nhau về năng suất lúa là do nhân tố thí nghiệm gây nên, sai khác ở mức xác suất 95% nên ta đánh * vào F_t.

Bước 8: Tính hệ số biến động- sai số thí nghiệm

$$cv = \frac{\sqrt{MS_e}}{\bar{G}} \times 100 = \frac{\sqrt{0,021598}}{1,335} \times 100 = 11,0\%$$

Kết luận thí nghiệm đảm bảo độ chính xác theo yêu cầu.

Bước 9: Lập bảng phân tích phương sai

Bảng 3.9. Kết quả phân tích biến động của năng suất lúa trong thí nghiệm so sánh năng suất giống lúa bố trí theo kiểu ô vuông la tinh

Nguồn biến động	df	SS	MS	F tính[b]	F bảng 5%	F bảng 1%
Hàng	3	0,030154	0,010051	0,465	8,94	27,91
Cột	3	0,827342	0,275781	12,77	8,94	27,91
Công thức	3	0,426842	0,142281	6,59*	4,76	9,78
Ngẫu nhiên	6	0,129585	0,021598			
Tổng	15	1,413923				

[a]*cv*= 11,0% [b]*= có ý nghĩa ở mức 5%

Kết luận: 4 giống lúa cho năng suất khác nhau chắc chắn, ở mức xác suất 95%.

Đánh giá hiệu quả của chia khối theo hàng và cột

Tương tự như kiểu khối ngẫu nhiên hoàn chỉnh, chia khối theo hàng và cột cũng nhằm giảm sai số ngẫu nhiên. Nhờ vậy, tăng độ chính xác so với kiểu bố trí ngẫu nhiên hoàn toàn và khối ngẫu nhiên hoàn chỉnh. Để đánh giá ta cần phải tính F của hàng (F_r) và F của cột (F_c) sau đó so sánh với F bảng. Cách tính như sau:

$$F_r = \frac{MS_r}{MS_e} \quad F_c = \frac{MS_c}{MS_e}$$

F bảng tra ở bảng E với $f_1 = t-1 = 3$; $f_2 = (t-1)(t-2) = 6$

$F_{05} = 8,94$; $F_{01} = 27,91$

Với thí kết quả ghi ở bảng 3.9 ta có:

$$F_r = \frac{0,010051}{0,021598} = 0,465; \quad F_c = \frac{0,275781}{0,021598} = 12,77$$

So sánh cho thấy: $F_r < F_b$ và $F_c > F_b$

Kết luận: Chia khối theo cột đạt mục đích làm giảm sai số ngẫu nhiên, còn chia khối theo hàng không đạt được mục đích.

3.3. TƯƠNG TÁC TRONG THÍ NGHIỆM HAI NHÂN TỐ

Từ kết quả thí nghiệm ta biết được ảnh hưởng của nhân tố thí nghiệm đến biến phụ thuộc. Trong thí nghiệm đa nhân tố ảnh hưởng của nhân tố này có thể phụ thuộc vào mức độ của nhân tố kia và người ta gọi đó là sự tương tác giữa các nhân tố thí nghiệm. Ví dụ, phản ứng của các mức đạm bón khác nhau tới năng suất lúa còn phụ thuộc vào giống. Giống chịu thâm canh sẽ có năng suất tăng khi bón tăng lượng đạm, còn giống không chịu thâm canh thì khi bón tăng lượng đạm năng suất có thể giảm. Như vậy ảnh hưởng của đạm tới năng suất lúa còn phụ thuộc vào giống. Hai nhân tố được coi là có tương tác với nhau nếu ảnh hưởng của nhân tố này thay đổi khi mức độ của nhân tố kia thay đổi.

Do đó, đối với thí nghiệm hai nhân tố cần xác định được ảnh hưởng của sự tương tác. Minh họa cách xác định ảnh hưởng tương tác giữa các nhân tố bằng thí nghiệm 2 nhân tố có kết quả ghi ở bảng 3.10. Thí nghiệm có:

Hai nhân tố A và B. Đó là nhân tố giống lúa và mức đạm bón.

Mỗi nhân tố có 2 mức tương ứng là: a_0, a_1 và b_0, b_1. Đó là giống X và Y; mức đạm No và N1.

Ta có 4 tổ hợp công thức thí nghiệm là: a_0b_0, a_1b_0, a_0b_1, a_1b_1.

Mỗi nhân tố có hai 2 loại ảnh hưởng là: ảnh hưởng đơn (simple effect) và ảnh hưởng chung (main effect);

Ngoài ra còn có ảnh hưởng tương tác giữa 2 nhân tố.

Bảng 3.10. Hai giả định số liệu cho thí nghiệm 2 nhân tố (giống và đạm): có tương tác và không có tương tác

Giống	Năng suất lúa (tấn/ha)		
	0 kg N/ha (No)	60 kg N/ha (N_1)	Trung bình
	Không có tương tác		
X	1,0	3,0	2,0
Y	2,0	4,0	3,0
TB	1,5	3,5	
	Có tương tác		
X	1,0	1,0	1,0
Y	2,0	4,0	3,0
TB	1,5	2,5	

Bước 1

Tính ảnh hưởng đơn của nhân tố A bằng cách so sánh kết quả ở 2 mức của A với mỗi mức của B theo công thức:

Ảnh hưởng đơn của A ở mức $b_0 = a_1b_0 - a_0b_0$

Ảnh hưởng đơn của A ở mức $b_1 = a_1b_1 - a_0b_1$

Tương tự tính ảnh hưởng đơn của B ứng với mỗi mức của A

Ảnh hưởng đơn của B ở mức $a_0 = b_1a_0 - b_0a_0$

Ảnh hưởng đơn của B ở mức $a_1 = b_1a_1 - b_0a_1$

Từ kết quả của bảng 3.10, ta có:

Trường hợp không có tương tác

Ảnh hưởng đơn của 2 giống X và Y ở No = 2,0 -1,0 = 1,0 t/ha

Ảnh hưởng đơn của 2 giống X và Y ở N_1 = 4,0 - 3,0 = 1,0 t/ha

Ảnh hưởng đơn của đạm với giống X = 3,0 - 1,0 = 2,0 t/ha

Ảnh hưởng riêng của đạm với giống Y = 4,0 - 2,0 = 2,0 t/ha

Trường hợp có tương tác

Ảnh hưởng đơn của 2 giống X và Y ở No = 2,0 - 1,0 = 1,0 t/ha

Ảnh hưởng đơn của 2 giống X và Y ở N_1 = 4,0 - 1,0 = 3,0 t/ha

Ảnh hưởng đơn của đạm với giống X = 1,0 - 1,0 = 0,0 t/ha

Ảnh hưởng đơn của đạm với giống Y = 4,0 - 2,0 = 2,0 t/ha

Bước 2:

Tính ảnh hưởng chính của A bằng cách lấy trung bình của các ảnh hưởng đơn của A với các mức của B theo công thức

Ảnh hưởng chính của A = (1/2) (ảnh hưởng đơn của A ở b_0
+ ảnh hưởng đơn của A ở mức b_1)

$$= (1/2)\left[(a_1b_0 - a_0b_0) + (a_1b_1 - a_0b_1)\right]$$

Ảnh hưởng chính của B = (1/2) $\left[(a_0b_1 - a_0b_0) + (a_1b_1 - a_1b_0)\right]$

Với ví dụ ta có:

Trường hợp không có tương tác

Ảnh hưởng chính của giống = (1/2)[(2,0 - 1,0) + (4,0 - 3,0)] = 1,0 t/ha

Ảnh hưởng chính của đạm = (1/2) [(3,0 - 1,0) + (4,0 - 2,0)] = 2,0 t/ha

Trường hợp có tương tác

Ảnh hưởng chính của giống = (1/2)[(2,0-1.0) + (4,0-1,0)] = 2,0 t/ha

Ảnh hưởng chính của đạm = (1/2) [(1,0-1,0) + (4,0-2.0)] = 1,0 t/ha

Bước 3:

Tính ảnh hưởng của tương tác A và B (ký hiệu A x B) đó là sự chênh lệch của ảnh hưởng đơn của một nhân tố tại 2 mức của nhân tố kia theo công thức:

A × B = (1/2) (ảnh hưởng đơn của A ở b_1 -ảnh hưởng đơn của A ở b_0)

$$= (1/2)\ [(a_1b_1 - a_0b_1) - (a_1b_0 - a_0b_0)]$$

hoặc = (1/2) (ảnh hưởng đơn của B ở mức a_1 - ảnh hưởng đơn của B ở a_0

$$= (1/2)[(a_1b_1 - a_1b_0) - (a_0b_1 - a_0b_0)]$$

Với ví dụ ta có:

Giống x Đạm = 1/2 (ảnh hưởng của đơn giống ở mức N_1
- ảnh hưởng đơn của giống ở mức N_0)
= 1/2 (3,0 -1,0) = 1,0 t/ha

Hoặc

Giống x Đạm = 1/2 (ảnh hưởng đơn của N với giống Y

- ảnh hưởng đơn của N với giống X)

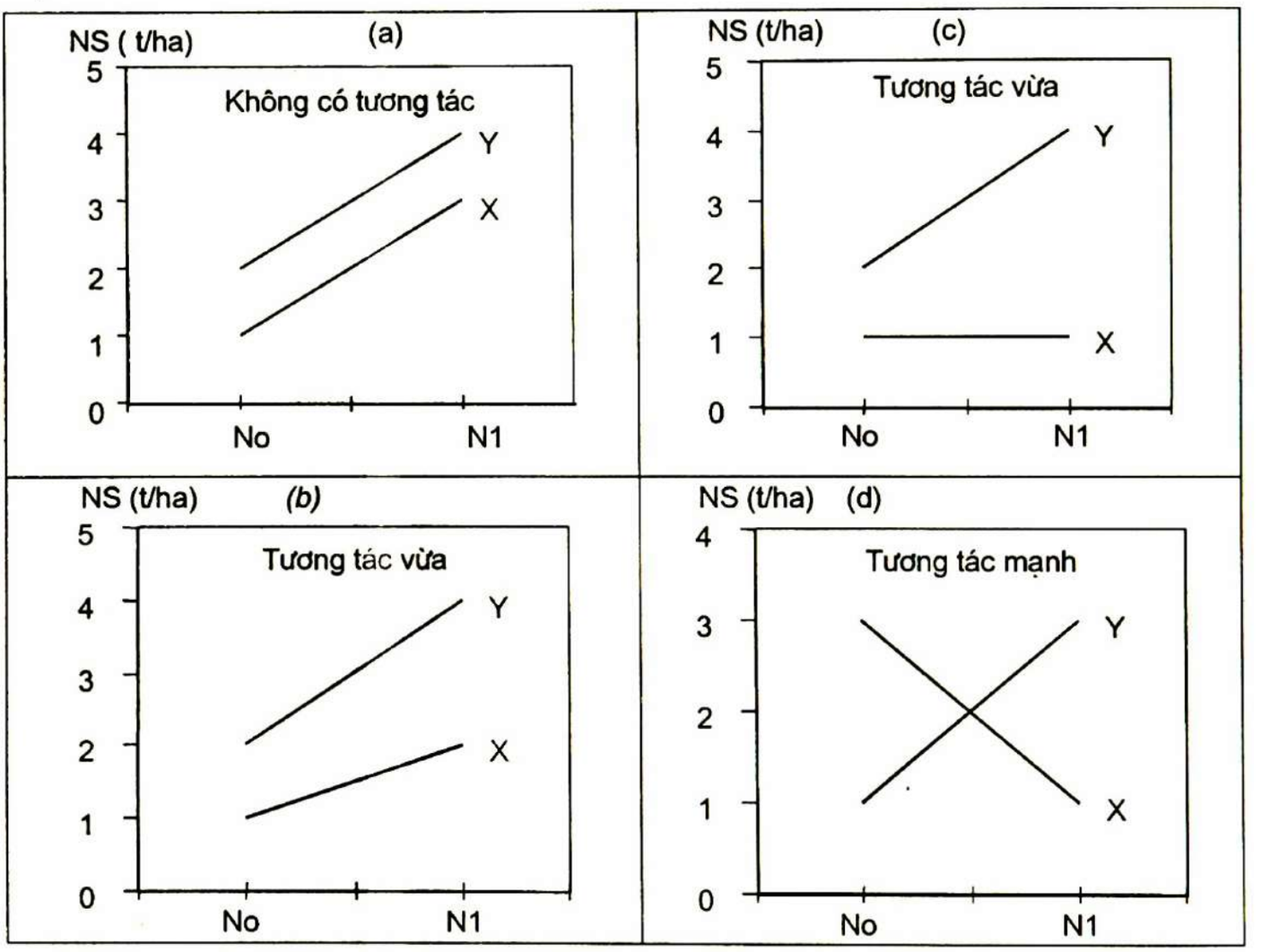
= 1/2 (2,0 -0,0) = 1,0 t/ha

Tính tương tác N và giống dựa trên số liệu thu được trong trường hợp không tương tác ta có:

Giống x đạm = 1/2 (1,0 - 1,0) = 0,0 t/ha

Hoặc: giống x đạm = 1/2 (2,0 - 2,0) = 0,0 t/ha

Sự phản ứng của 2 giống với các mức phân đạm khác nhau được thể hiện ở đồ thị 3.1.



Đồ thị biểu thị các mức độ tương tác giữa nhân tố A với nhân tố B

Nhận xét đồ thị:

Đồ thị 3.1.a cho thấy không có sự tương tác giữa 2 nhân tố.

Đồ thị 3.1.c cho thấy ảnh hưởng của tương tác là 1,0 t/ha.

Đồ thị 3.1.b và 3.1.d cho thấy ảnh hưởng của tương tác lớn hơn và nhỏ hơn 1,0 t/ha.

Đồ thị 3.1.b cũng cho thấy, cả 2 giống đều có phản ứng thuận với mức đạm bón, nhưng phản ứng của các giống không giống nhau. Giống Y (2 t/ha) phản ứng với đạm mạnh hơn giống X (1 t/ha).

Đồ thị 3.1.d cho thấy, giống X (2 t/ha) có tương quan thuận với đạm, giống Y có tương quan nghịch với đạm cho ảnh hưởng tương tác là 2 t/ha.

Từ số liệu tính toán và đồ thị cho thấy 3 điểm cần chú ý:

Tương tác giữa hai nhân tố chỉ có thể xác định được khi hai nhân tố đó cùng được thử trong cùng một thí nghiệm đa nhân tố.

Khi không có tương tác (như đồ thị 3.1.a) ảnh hưởng đơn của một nhân tố là như nhau ở tất cả các mức của nhân tố kia và bằng ảnh hưởng chính. Điều đó có nghĩa, khi không có tương tác kết quả của thí nghiệm 1 nhân tố có kết quả giống như thí nghiệm 2 nhân tố. Với ví dụ trên, ảnh hưởng đơn của giống ở mức N_0 và N_1 đều bằng nhau và bằng ảnh hưởng chính 1,0 t/ha. Trong trường hợp này ảnh hưởng của giống luôn bằng 1,0 dù:

- Hai giống được thử nghiệm trong các tổ hợp với 2 mức N trong cùng một thí nghiệm đa nhân tố.
- Hai giống đó được thử nghiệm ở mức N_0 trong thí nghiệm một nhân tố.
- Hai giống đó được thử nghiệm ở mức bón N_1 trong thí nghiệm một nhân tố

Khi có tương tác (như đồ thị 3.1.b, 3.1.c và 3.1.d) ảnh hưởng đơn của một nhân tố thay đổi khi mức độ của nhân tố kia thay đổi. Do đó, ảnh hưởng chính khác với ảnh hưởng đơn. Vì vậy, khi có tương tác giữa 2 nhân tố thì chỉ xem xét ảnh hưởng đơn, không xem xét ảnh hưởng chung và kết quả của thí nghiệm 1 nhân tố chỉ đúng ở điều kiện nền thí nghiệm của nó.

Ý nghĩa của nghiên cứu tương tác:

Khi có tương tác giữa hai nhân tố, cần đánh giá ảnh hưởng của các nhân tố dựa trên ảnh hưởng đơn (simple effect) chứ không dựa trên ảnh hưởng chung (main effect) của nhân tố đó. Ngược lại khi không có tương tác thì không cần dựa trên ảnh hưởng riêng mà chỉ cần đánh giá dựa trên ảnh hưởng chung. Như trong ví dụ trên, trường hợp không có tương tác, để đánh giá hiệu quả của việc bón đạm tính giá trị trung bình từ các giống và so sánh hai giá trị trung bình (tức là main effect). Nhưng trong trường hợp có tương tác ta không thể đánh giá dựa trên giá trị trung bình như vậy được mà phải xét riêng cho từng giống, tức là dựa trên ảnh hưởng đơn (simple effect). Đây cũng chính là một nguyên tắc quan trọng trong việc tính toán xử lý số liệu và trình bày số liệu trong các thí nghiệm đa nhân tố.

3.4. PHÂN TÍCH BIẾN ĐỘNG CHO THÍ NGHIỆM HAI NHÂN TỐ

3.4.1. Kiểu khối ngẫu nhiên hoàn chỉnh (RCBD)

Trong thí nghiệm 2 nhân tố bố trí kiểu khối ngẫu nhiên hoàn chỉnh có 6 nguồn biến động. Đó là biến động của công thức, của ngẫu nhiên, của khối, của nhân tố A,

của nhân tố B và biến động của tương tác giữa 2 nhân tố. Công thức tính tổng phương sai (TSS) của thí nghiệm hai nhân tố bố trí kiểu khối ngẫu nhiên hoàn chỉnh như sau:

$TSS = SS_r + SS_t + SS_a + SS_b + SS_{ab} + SS_c$

Trong đó

SS_r: Tổng bình phương biến động của khối

SS_t: Tổng bình phương biến động của công thức

SS_a: Tổng bình phương biến động của mức i của nhân tố A

SS_b: Tổng bình phương biến động của mức j của nhân tố B

SS_{ab}: Tổng bình phương biến động do tương tác của nhân tố A và B

SS_c: Tổng bình phương biến động do sai số ngẫu nhiên

Bảng phân tích phương sai và công thức tính trình bày ở bảng dưới đây.

Bảng 3.11. Bảng phân tích phương sai cho thí nghiệm hai nhân tố bố trí kiểu khối ngẫu nhiên hoàn chỉnh

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	F-tính	F-tra bảng
Khối (Rep)	r-1	$SSr = \Sigma R^2/ab - CF$			
Công thức (T)	ab-1	$SS_t = \Sigma T^2/r - CF$	$MS_t = SSt/ab-1$		f_1:df_t, f_2: df_e
Giống (A)	a-1	$SS_a = \Sigma A^2/rb - CF$	$MS_a = SSa/a-1$	MS_a/MS_e	f_1:df_a, f_2:df_e
Đạm (B)	b-1	$SS_b = \Sigma B^2/ra - CF$	$MS_b = SSb/b-1$	MS_b/MS_e	f_1:df_b, f_2: df_e
T. tác A x B	(a-1)(b-1)	$SS_{ab} = SS_t - SS_a - SS_b$	$MS_{ab} = SS_{ab}/(a-1)(b-1)$	MS_{ab}/MS_e	f_1:df_{ab}, f_2:df_e
Sai số	(r-1)(ab-1)	$SS_e = TSS - SS_r - SS_t$	$MS_e = SS_e/(r-1)(ab-1)$		
Toàn bộ	rab-1	$TSS = \Sigma x^2 - CF$ ($CF = G^2/rab$)			

Các bước phân tích phương sai được minh họa qua phân tích phương sai cho năng suất lúa thu được từ thí nghiệm "Nghiên cứu phản ứng của 3 giống lúa (kí hiệu V1, V2, V3) và 5 mức bón phân đạm (N1, N2, N3, N4, N5)" với 4 lần nhắc lại bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh. Số liệu thí nghiệm trình bày ở bảng 3.12.

Ta có:

Số mức của A (giống) ký hiệu là a: a = 3

Số mức của B (đạm) ký hiệu là b: b = 5

Số lần nhắc lại ký hiệu là r: r = 3

Bảng 3.12. Năng suất của 3 giống lúa thử nghiệm với 5 mức đạm bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh

Mức đạm	Năng suất lúa (tấn/ha)				Tổng của công thức
(kg/ha)	Nhắc lại 1	Nhắc lại II	Nhắc lại III	Nhắc lại IV	(T)
			V1		
No	3.852	2.606	3.144	2.894	12.496
N1	4.788	4.936	4.562	4.608	18.894
N2	4.576	4.454	4.884	3.924	17.838
N3	6.034	5.276	5.906	5.652	22.868
N4	5.874	5.916	5.984	5.518	23.292
			V2		
No	2.846	3.794	4.108	3.444	14.192
N1	4.956	5.128	4.150	4.990	19.224
N2	5.928	5.698	5.810	4.308	21.744
N3	5.664	5.362	6.458	5.474	22.958
N4	5.458	5.546	5.786	5.932	22.722
			V3		
No	4.192	3.754	3.738	3.428	15.112
N1	5.250	4.582	4.896	4.286	19.014
N2	5.822	4.848	5.678	4.932	21.280
N3	5.888	5.524	6.042	4.756	22.210
N4	5.864	6.264	6.056	5.362	23.546
Rep. total (R)	76.992	73.688	77.202	69.508	
Tổng toàn bộ (G)					297.390

Nguồn: Gomez, 1984

Bước 1: Thiết kế bảng để tính tổng A, tổng B như bảng 3.13.

Bảng 3.13. Tổng năng suất (ns) lúa tính theo nhân tố giống và đạm

Đạm (B)	Giống (A)			Tổng ns tính theo mức đạm (B)
	V1	V2	V3	
No	12.496	14.192	15.112	41.800
N1	18.894	19.224	19.014	57.132
N2	17.838	21.744	21.280	60.862
N3	22.868	22.958	22.210	68.036
N4	23.292	22.722	23.546	69.256
Tổng giống (A)	95.388	100.840	101.162	297.390

Áp dụng các công thức trong bảng 3.11 để tính các giá trị sau:

Bước 2: Tính độ tự do cho các nguồn biến động

Bước 3: Tính tổng bình phương của các nguồn biến động.

Bước 4: Tính trung bình bình phương của các nguồn biến động.

Bước 5: Tính F cho nhân tố A, B và tương tác A x B

Bước 6: Tra F bảng và so sánh F tính với F bảng

Bước 7: Tính cv

Bước 8: Thiết lập bảng kết quả phân tích phương sai và kết luận

Cách tính các giá trị tương tự như cách tính cho thí nghiệm một nhân tố và theo công thức bảng 3.11. Kết quả cụ thể ghi ở bảng 3.14.

Bảng 3.14. Kết quả phân tích phương sai cho năng suất của thí nghiệm 2 nhân tố (3 giống lúa với 5 mức phân đạm) bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	F-tính	F-tra bảng	
					5%	1%
Nhắc lại (khối)	3	2,599	0,866	5,74**	2,83	4,29
Công thức (T)	14	44,578	3,184	21,09**	1,94	2,54
Giống (A)	2	1,052	0,526	3,48*	3,22	5,15
Đạm (B)	4	41,234	10,308	68,26**	2,59	3,80
A x B	8	2,292	0,286	1,89ns	2,17	2,96
Ngẫu nhiên	42	6,353	0,151			
Toàn bộ	59	53,530				

Tính hệ số biến động- sai số thí nghiệm

$$cv = \frac{\sqrt{MS_e}}{\overline{G}} \times 100 = \frac{\sqrt{0,151}}{4,956} \times 100 = 7,8\%$$

Thí nghiệm đảm bảo độ chính xác.

Kết luận: Kết quả cho thấy không có sự tương tác giữa giống và phân. Điều đó có nghĩa giống khác nhau không chịu ảnh hưởng khác nhau của đạm và các mức đạm khác nhau không chịu ảnh hưởng của giống. Ảnh hưởng chính của giống và đạm đều có ý nghĩa. Hay các giống khác nhau và các mức đạm khác nhau cho năng suất khác nhau. Các giống khác nhau cho năng suất khác nhau ở mức xác suất 95%; Các mức đạm bón khác nhau cho năng suất khác nhau ở mức xác suất 99%.

3.4.2. Kiểu ô chính - ô phụ (split-plot design)

Với thí nghiệm bố trí kiểu ô chính - ô phụ cần chia ra phân tích cho ô chính và ô phụ. Thí nghiệm 2 nhân tố bố trí kiểu ô chính - ô phụ có 6 nguồn biến động. Đó là biến động của nhân tố a, nhân tố b; ngẫu nhiên của nhân tố a, ngẫu nhiên của nhân tố

b, ngẫu nhiên của tương tác ab và của nhắc lại. Tổng phương sai của kiểu bố trí này được tính theo công thức sau:

$TSS = SS_r + SS_a + SS_{ca} + SS_b + SS_{ab} + SS_{cb}$

Trong đó:

SS_r: Tổng bình phương biến động của nhắc lại

SS_a: Tổng bình phương biến động của mức i của nhân tố A

SS_b: Tổng bình phương biến động của mức j của nhân tố B

SS_{ab}: Tổng bình phương biến động do tương tác của nhân tố A và nhân tố B

SS_{ca}: Tổng bình phương biến động của sai số ngẫu nhiên của nhân tố A

SS_{cb}: Tổng bình phương biến động của sai số ngẫu nhiên của nhân tố B

Bảng phân tích phương sai và công thức tính trình bày ở bảng 3.15.

Bảng 3.15. Cách tính các giá trị trong phân tích phương sai của thí nghiệm bố trí kiểu ô chính ô phụ

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	F-tính	F-tra bảng
Khối (Rep)	r-1	$SS_r = \Sigma R^2/ab - CF$	$MS_r = SS_r/r-1$		
Nhân tố chính (A)	a-1	$SS_a = \Sigma A^2/rb - CF$	$MS_a = SS_a/(a-1)$	$Fa=MS_a/MS_{ea}$	$f1=df_a$; $f2 = df_{ea}$
Ngẫu nhiên a	(r-1) (a-1)	$SS_{ea}=\Sigma(RA)^2/b - CF - SS_r - SS_a$	$MS_{ea}=SS_{ea}/(r-1)(a-1)$		
Nhân tố phụ (B)	b-1	$SS_b = \Sigma B^2/ra - CF$	$MS_b=SS_b/b-1$	Fb=MSb/MSeb	$f1=df_b$; $f2 = df_{eb}$
Tương tác A x B	(a-1)(b-1)	$SS_{ab} = \Sigma(AB)^2/r - CF - SS_b - SS_a$	$MS_{ab}=SS_{ab}/(a-1)(b-1)$	$Fab=MS_{ab}/MS_{eb}$	$f1=df_{ab}$; $f2 = df_{eb}$
Ngẫu nhiên (b)	a(r-1)(b-1)	SS_{ab}=TSS- (toàn bộ SS khác)	$MS_{eb}=SS_{eb}/a(r-1)(b-1)$		
Toàn bộ	rab-1	$TSS = \Sigma x^2 - CF$ ($CF = G^2/rab$)			

Minh họa cách phân tích phương sai bằng phân tích phương sai cho năng suất lúa của thí nghiệm "Nghiên cứu phản ứng của 4 giống lúa với các mức đạm", thí nghiệm bố trí theo kiểu ô chính - ô phụ. Kết quả thí nghiệm trình bày ở bảng 3.16. Thí nghiệm có:

2 nhân tố thí nghiệm: 4 giống lúa và 6 mức phân N; r = 3

Số mức của nhân tố ở ô chính (mức phân) a= 6; Số mức của nhân tố ở ô phụ là (giống lúa) b=4.

Bảng 3.16. Năng suất của thí nghiệm 2 nhân tố (4 giống lúa, 6 mức đạm; 3 lần nhắc lại) bố trí theo kiểu ô chính ô phụ

Giống	Năng suất kg/ha			
	Nhắc lại I	Nhắc lại II	Nhắc lại III	Tổng
	N_0 (0 kg N/ha)			
V_1 (IR8)	4.430	4.478	3.850	3.142
V_2 (IR5)	3.944	5.314	12.758	9.550
V_3 (C4-63)	3.464	2.944	3.660	4.836
V_4 (Peta)	4.126	4.482	12.918	13.444
Tổng	15.964	17.216	15.488	
N_1 (60 kg N/ha)				
V_1	5.418	5.166	6.432	5.556
V_2	6.502	5.858	17.016	16.328
V_3	4.768	6.004	5.586	4.652
V_4	5.192	4.604	17.946	13.444
Tổng	21.880	21.632	22.226	
N_2 (90 kg N/ha)				
V_1	6.076	6.420	6.704	6.014
V_2	6.008	6.127	19.200	17.982
V_3	6.244	5.724	6.642	4.146
V_4	4.546	5.744	18.777	14.448
Tổng	22.874	24.015	23.506	
N_3 (120 kg N/ha)				
V_1	6.462	7.056	6.680	6.370
V_2	7.139	6.982	0.198	8.042
V_3	5.792	5.880	6.564	3.638
V_4	2.774	5.036	0.685	1.448
Tổng	22.167	24.954	23.252	
N_4 (150 kg N/ha)				
V_1	7.290	7.848	7.552	6.320
V_2	7.682	6.594	2.690	0.062
V_3	7.080	6.662	6.576	20.852
V_4	1.414	1.960	2.766	6.140
Tổng	23.466	23.064	23.214	
N_5 (180 kgN/ ha)				
V_1	8.452	8.832	8.818	5.480
V_2	6.228	7.387	26.102	18.196
V_3	5.594	7.122	6.006	2.014
V_4	2.248	1.380	19.621	5.642
Tổng	22.522	24.721	22.318	

Nguồn (Gomez, 1984)

Bước 1: Lập 2 bảng tổng năng suất

Bảng tổng năng suất theo

- Số lần nhắc lại của nhân tố A
- Nhắc lại,
- Nhân tố A
- Toàn bộ.

Trong ví dụ này đó là bảng 3.17 - bảng năng suất theo mức phân và nhắc lại

Bảng 3.17: Kết quả năng suất lúa theo mức phân đạm ở các lần nhắc lại

(tính từ bảng 3.16) (kg/ha)

Mức đạm	Tổng năng suất theo nhắc lại của nhân tố A (RA)			Tổng NS theo đạm (A)-nhân tố chính
	Nhắc lại I	Nhắc lại II	Nhắc lại III	
N_o	15.964	17.218	15.488	48.670
N_1	21.880	21.632	22.226	65.738
N_2	22.874	24.015	23.506	70.395
N_3	22.167	24.954	23.252	70.373
N_4	23.466	23.064	23.214	69.744
N_5	22.522	24.721	22.318	69.561
Tổng của nhắc lại (R)	128.873	135.604	130.004	
Tổng toàn bộ (G)				394.481

Bảng năng suất theo:

- Nhân tố A với nhân tố B
- Nhân tố B.

Trong ví dụ này, đó là bảng 3.18 bảng năng suất theo các mức đạm với các giống

Bảng 3.18. Kết quả năng suất lúa theo mức đạm bón và giống *(tính từ bảng 3.15)*

Mức đạm	Tổng năng suất (AB) (kg/ha)			
	V_1	V_2	V_3	V_4
N_o	12.758	12.918	9.550	13.444
N_1	17.016	17.946	16.328	14.448
N_2	19..200	18.777	17.982	14.436
N_3	20.198	20.685	18.042	11.448
N_4	22.690	20.852	20.062	6.140
N_5	26.102	19.621	18.196	5.642
Tổng NS theo giống (B)	117.964	110.799	100.160	65.558

Bước 2: Tính độ tự do cho các nguồn biến động

$df_r = r - 1 = 2$

$df_a = a - 1 = 5$

$df_{ca} = (r - 1)(a - 1) = 10$

$df_b = b - 1 = 3$

$df_{ab} = (a-1)(b - 1) = 15$

$df_{cb} = a(r - 1)(b - 1) = 36$

$Tdf = rab-1 = 71$

Bước 3: Tính số hiệu chỉnh và các tổng bình phương của nhân tố chính

$$CF = \frac{G^2}{rab} = \frac{(394.481)^2}{(3)(6)(4)} = 2.161.323.047$$

Tổng bình phương của toàn bộ

$$TSS = \sum X^2 - CF$$

$$= \left[(4.430)^2 + ... + (2.014)^2\right] - 2.161.323.047 = 204.747.916$$

Tổng bình phương do nhắc lại

$$SS_r = \frac{\sum R_i^2}{ab} - CF$$

$$= \frac{(128.873)^2 + (135.604)^2 + (130.004)^2}{(6)(4)} - 2.161.323.047 = 1.082.577$$

Tổng bình phương do nhân tố A (đạm)

$$SS_a = \frac{\sum A^2}{rb} - CF$$

$$= \frac{(48.670)^2 + ... + (69.561)^2}{(3)(4)} - 2.161.323.047 = 30.429.200$$

Tổng bình phương do ngẫu nhiên A (đạm)

$$SS_{ca} = \frac{\sum (RA)^2}{b} - CF - SS_r - SS_a$$

$$= \frac{(15.964)^2 + ... + (22.318)^2}{(4)} - 2.161.323.047 - 1.082.577 - 30.429.200 = 1.419.678$$

Bước 4: Tính tổng bình phương của khối phụ B (theo giống):

Tổng bình phương của B (giống)

$$SS_b = \frac{\sum B^2}{ra} - CF$$

$$= \frac{(117.964)^2 + ... + (65.558)^2}{(3)(6)} - 2.161.323.047 = 89.888.101$$

Tổng bình phương của tương tác AB (đạm và giống)

$$SS_{ab} = \frac{\sum (AB)^2}{r} - CF - SS_b - SS_a$$

$$= \frac{(12.758)^2 + ... + (5.642)^2}{3} - 2.161.323.047 - 89.888.101 - 30.429.200 = 69.343.487$$

Tổng bình phương của sai số do B (giống)

SS_{cb} = Tổng SS - (tổng của tất cả SS khác)

= 204.747.961-(1.082.577 + 30.429.200 + 1.419.687 + 89.888.101 + 69.343.487)
= 12.584.873

Bước 5:

Tính trung bình bình phương cho từng nguồn biến động bằng cách chia tổng bình phương cho độ tự do tương ứng

Trung bình bình phương của nhắc lại

$$MS_r = \frac{SS_r}{r-1} = \frac{1.082.577}{2} = 541.228$$

Trung bình bình phương của nhân tố A

$$MS_a = \frac{SS_a}{a-1} = \frac{30.429.200}{5} = 6.085.840$$

Trung bình bình phương của ngẫu nhiên của nhân tố A

$$MS_{ca} = \frac{SS_{ca}}{(r-1)(a-1)} = \frac{1.419.678}{10} = 141.968$$

Trung bình bình phương của nhân tố B

$$MS_b = \frac{SS_b}{b-1} = \frac{89.888.101}{3} = 29.962.700$$

Trung bình bình phương của 2 nhân tố AB

$$MS_{a.b} = \frac{SS_{ab}}{(a-1)(b-1)} = \frac{69.343.487}{15} = 4.622.899$$

Trung bình bình phương của ngẫu nhiên của nhân tố B

$$MS_{cb}=\frac{SS_{cb}}{a(r-1)(b-1)}=\frac{12.584.873}{36}=349.580$$

Bước 6:

Tính giá trị F cho từng nhân tố cần kiểm tra, bằng cách chia MS của nhân tố đó cho trung bình bình phương ngẫu nhiên của nó:

$$F(A)=\frac{MS_a}{MS_{ca}}=\frac{6.085.840}{141.968}=42,87$$

$$F(B)=\frac{MS_b}{MS_{cb}}=\frac{29.962.700}{349.580}=85,71$$

$$F(A.B)=\frac{MS_{ab}}{MS_{cb}}=\frac{4.622.899}{349.580}=13,22$$

Bước 7: Với những nhân tố có giá trị F tính lớn hơn 1, tra F bảng ở bảng E với f_1=df của MS tử số và f_2=df của MS của mẫu số.

Với ví dụ trên có các giá trị F bảng như sau:

F (A) với df_1= 5 và df_2 =10 ta có F_{05}= 3,33; F_{01}= 5,64

F (B) với df_1= 3 và df_2 = 36 ta có F_{05}= 2,86; F_{01}= 4,38

F (AB) với df_1=15 và df_2 = 36 ta có F_{05}= 1.96; F_{01}= 2,58

Bước 8: Tính hệ số biến động cho khối chính và khối phụ.

$$cv(A)=\frac{\sqrt{MS_{ca}}}{\overline{G}}\times 100=\frac{\sqrt{141.968}}{5.479}\times 100=6,9\%$$

$$cv(B)=\frac{\sqrt{MS_{cb}}}{\overline{G}}\times 100=\frac{\sqrt{349.580}}{5.479}\times 100=10,8\%$$

Giá trị của cv(A) chỉ độ chính xác của nhân tố ở khối chính. Giá trị của cv(B) chỉ độ chính xác của nhân tố ở khối phụ. Giá trị của cv(B) được ước đoán là nhỏ hơn cv(A) vì như đã trình bày ở phần trước nhân tố của khối chính thường yêu cầu độ chính xác ít hơn so nhân tố của khối phụ. Tuy nhiên không phải lúc nào trường hợp dự đoán đó cũng đúng. Như ở ví dụ trên là trường hợp ngoài dự kiến.

Bước 9:

Lập bảng phân tích phương sai với các giá trị đã tính được từ bước 3 đến bước 8. So sánh giá trị F tính với F bảng như hướng dẫn ở CRD.

Bảng 3.19. Kết quả phân tích phương sai năng suất cho thí nghiệm bố trí theo kiểu ô chính- ô phụ có 4 giống lúa với 6 mức đạm *(kết quả trình bày ở bảng 3.15[a])*

Nguồn biến động	df	SS	MS	F tính[b]	F bảng	
					5%	1%
Nhắc lại	2	1.082.577	541.228			
Đạm (A)	5	30.429.200	6.085.840	42,87**	3,33	5,64
Sai số (a)	10	1.419.678	141.968			
Giống (B)	3	89.888.101	29.962.700	85,71**	2,86	4,38
A x B	15	69.343.487	4.622.899	13,22**	1,96	2,58
Sai số (b)	36	12.584.873	349.580			
Tổng số	71	204.747.916				

a: cv(a)=6,9%, cv(b)=10,8%
b**: có ý nghĩa ở mức 1%

Kết luận:

Các mức bón N khác nhau có ảnh hưởng rất khác nhau đến năng suất lúa (ở mức tin cậy 99%)

Các giống lúa khác nhau có ảnh hưởng rất khác nhau đến năng suất (ở mức tin cậy 99%)

Các giống lúa có năng suất khác nhau ở các mức đạm khác nhau hay hiệu lực của các mức N tới năng suất của các giống lúa khác thì khác nhau rất chắc chắn (ở mức tin cậy 99%).

Chương 4
SO SÁNH SỐ TRUNG BÌNH

4.1. Ý NGHĨA CỦA SO SÁNH SỐ TRUNG BÌNH

Kết luận từ kết quả phân tích biến động như trình bày ở chương ba mới chỉ cho biết các công thức có sự khác nhau có ý nghĩa hay không. Nhưng điều đó không có nghĩa là tất cả các công thức thí nghiệm đều khác nhau chắc chắn. Đồng thời kết luận đó cũng chưa chỉ ra được trong những công thức thí nghiệm thì công thức nào tốt nhất.

So sánh các số trung bình của công thức thí nghiệm phục vụ cho mục đích của thí nghiệm là tìm ra công thức thí nghiệm có hiệu quả nhất, như thí nghiệm so sánh hiệu quả của 6 loại thuốc trừ rầy nâu (mục 3.2.1). Để làm được việc này, ta phải so sánh số trung bình của các công thức với nhau. Nhưng vì số trung bình của các công thức được ước lượng thông qua mẫu do vậy sự sai khác của chúng phải được kiểm chứng bằng phương pháp thống kê. Dựa vào kết quả phân tích thống kê để kết luận là chúng có thực sự khác nhau hay không. Như vậy so sánh số trung bình là dùng phương pháp thống kê để xác định sự sai khác (chênh lệch) giữa trung bình của các công thức có ý nghĩa hay không.

Có nhiều cách so sánh khác nhau tuỳ vào mục đích nghiên cứu, nhưng có thể phân loại thành 2 nhóm chính: So sánh cặp đôi *(Pair Comparison)* và So sánh nhóm *(Group Comparison)*. So sánh cặp đôi có 2 phương pháp: So sánh sai khác nhỏ nhất có ý nghĩa - *Least Significant Difference Test (LSD) và* So sánh Duncan - *Duncan's Multiple Range Test (DMRT)*. So sánh nhóm có 4 phương pháp: So sánh giữa các nhóm công thức với nhau - *Between-group Comparison;* So sánh giữa các công thức trong nhóm - *Within-group Comparison;* So sánh xu thế - *Trend Comparison;* và so sánh nhân tố - *Factorial Comparison).*

Trong tài liệu này chúng tôi chỉ trình bày các phương pháp so sánh cặp đôi, so sánh giữa các nhóm và xu thế.

4.2. SO SÁNH CẶP ĐÔI

Trong so sánh cặp đôi, hai phương pháp so sánh theo sai khác nhỏ nhất (LSD) và Duncan về cơ bản là giống nhau, đó là so sánh các số trung bình của các công thức dựa trên giới hạn của "Sai khác có ý nghĩa". Điểm khác nhau cơ bản là phương pháp so sánh theo sai khác nhỏ nhất chỉ dùng một giá trị để so sánh mọi cặp công thức còn phương pháp Duncan sử dụng những giá trị khác nhau để so sánh cho những cặp công thức khác nhau.

Các thí nghiệm có số công thức nhỏ hơn hoặc bằng 6, hay thí nghiệm đã được định sẵn các cặp cần so sánh, ví dụ như so sánh các công thức trong thí nghiệm với công thức đối chứng thì người ta sử dụng phương pháp LSD. Nếu mục đích thí nghiệm là so sánh bất kỳ cặp đôi công thức nào với nhau ở thí nghiệm có số công thức lớn hơn 6 thì người ta sử dụng phương pháp DMRT (hay dễ gọi là phương pháp Duncan).

Trước khi thực hiện so sánh ta cần nghiên cứu kết quả phân tích phương sai (ANOVA). Nếu $F_t > F_b$ có nghĩa rằng sự sai khác giữa các công thức có ý nghĩa và ta thực hiện công việc so sánh.

4.2.1. So sánh theo sai khác nhỏ nhất - *Least Significant Difference Test* - LSD

4.2.1.1. Nguyên tắc so sánh

So sánh theo sai khác nhỏ nhất là so sánh sự chênh lệch của 2 công thức với giá trị sai khác nhỏ nhất có ý nghĩa. Sai khác nhỏ nhất có ý nghĩa (ký hiệu là LSD) là giá trị nhỏ nhất chỉ ra ranh giới giữa khác nhau có ý nghĩa và khác nhau không có ý nghĩa của bất kỳ cặp công thức nào.

Nói một cách ngắn gọn, LSD được dùng để đánh giá trung bình của hai công thức có khác nhau hay không. Nguyên tắc đánh giá: Hai công thức được coi là khác nhau nếu sự chênh lệch (sự sai khác) giữa chúng lớn hơn LSD. Có thể tóm tắt như sau:

Nếu sự sai khác > LSD kết luận 2 công thức khác nhau có ý nghĩa

Nếu sự sai khác ≤ LSD kết luận 2 công thức khác nhau không có ý nghĩa

Giá trị của LSD thay đổi phụ thuộc vào mức có ý nghĩa α mà ở đó LSD được tính. Mức có ý nghĩa α tương ứng với mức xác suất hay độ tin cậy $(1-\alpha)$. Có nghĩa là ứng với mức có ý nghĩa khác nhau thì có giá trị LSD khác nhau. Việc dùng LSD ở độ tin cậy nào là do yêu cầu của thí nghiệm. Thông thường người ta tính LSD ở độ tin cậy 95% hay 99% tức là $\alpha = 0,05$ hay 0,01.

4.2.1.2. Trường hợp áp dụng

So sánh theo LSD là cách so sánh đơn giản nhất, nó thích hợp cho so sánh những cặp theo ý định thiết kế. Nhưng cách so sánh này không thích hợp khi so sánh tất cả các cặp, đặc biệt khi số công thức thí nghiệm lớn. Bởi vì số cặp so sánh sẽ tăng rất nhanh khi số công thức tăng. Ví dụ, nếu số công thức là 5 thì số cặp so sánh là 10, số công thức là 10 thì số cặp so sánh là 45. Chỉ nên dùng so sánh trung bình của các công thức bằng LSD khi $F_t > F_b$ trong phân tích biến động.

4.2.1.3. Cách thực hiện

Bước 1: Tính sự sai khác (ký hiệu là d_{ij}) của 2 số trung bình cần so sánh

$$d_{ij} = \overline{X_i} - \overline{X_j}$$

Trong đó $\overline{X_i}$ và $\overline{X_j}$ là trung bình của công thức thứ i và j.

Bước 2:

Tính LSD tại độ mức có ý nghĩa α. Giống như phân tích phương sai, so sánh theo LSD thường chọn 2 mức có ý nghĩa α là 0,05 hoặc 0,01 (tương ứng với mức xác suất là 95% hoặc 99%). Công thức tính LSD như sau:

$$LSD_\alpha = (t_\alpha)(S_{\bar{d}})$$

Trong đó:

t_α là giá trị t lý thuyết, tra từ bảng C với độ tự do bằng $n = df_e$.

α mức có ý nghĩa

$S_{\bar{d}}$ *standard error* là sai khác chuẩn giữa các số trung bình

Cách tính S_d: Cách tính S_d phụ thuộc vào kiểu bố trí thí nghiệm.

Với các thí nghiệm một nhân tố (như bố trí kiểu ngẫu nhiên hoàn toàn, khối ngẫu nhiên hoàn chỉnh, ô vuông la tinh) chỉ có một sai số ngẫu nhiên thì sai số chuẩn $S_{\bar{d}}$ được dùng cho tất cả các cặp so sánh và được tính theo công thức:

$$S_{\bar{d}} = \sqrt{\frac{2s^2}{r}}$$

Trong đó:

r: số lần nhắc lại;

s^2 là MS_e (trung bình bình phương của sai số ngẫu nhiên)

Nên ta có thể viết $S_{\bar{d}} = \sqrt{\frac{2MS_e}{r}}$

Với thí nghiệm hai nhân tố thiết kế kiểu *ô chính - ô phụ (Split-plot design)* ta có thể có các cặp so sánh khác nhau ứng với các giá trị $S_{\bar{d}}$ như sau:

Bảng 4.1. Các cặp so sánh và cách tính sai số chuẩn cho thí nghiệm kiểu ô chính - ô phụ

Các cặp so sánh	Công thức tính $S_{\bar{d}}$
• So sánh 2 số trung bình của khối chính (Nhân tố A)	$\sqrt{\frac{2E_a}{rb}}$
• So sánh 2 số trung bình của khối phụ (Nhân tố B)	$\sqrt{\frac{2E_b}{ra}}$
• So sánh 2 số trung bình của khối phụ tại 1 mức của công thức ở khối chính	$\sqrt{\frac{2E_b}{r}}$
• So sánh 2 trung bình của khối chính tại 1 mức của công thức ở khối phụ	$\sqrt{\frac{2[(b-1)E_b + E_a]}{rb}}$

Chú ý: Có thể có những cặp so sánh khác nhưng chỉ có những cặp so sánh trên là có ý nghĩa thực tế.

Trong đó:

$E_a = MS_e$ của nhân tố chính a; a = số công thức của nhân tố chính,

$E_b = MS_e$ của nhân tố phụ b; b = số công thức của nhân tố phụ.

r = số lần nhắc lại.

Bước 3: So sánh sự sai khác của 2 số trung bình với giá trị của LSD_α.

Nếu $|d_{ij}| > LSD_\alpha$ thì số trung bình của công thức i và j khác nhau có ý nghĩa ở độ tin cậy (1- α).

Nếu $|d_{ij}| > LSD_{.01}$: Sai khác ở mức xác suất 99%, ký hiệu (**)

Nếu $|d_{ij}| > LSD_{.05}$: Sai khác ở mức xác suất 95%, ký hiệu (*)

Nếu $|d_{ij}| \leq LSD_\alpha$ thì trung bình của hai công thức i và j không sai khác nhau ở mức có ý nghĩa (1 - α).

4.2.1.4. So sánh cặp đôi cho thí nghiệm một nhân tố

Minh họa cách so sánh bằng so sánh các công thức của thí nghiệm "So sánh năng suất của 4 giống ngô A, B, C và D" bố trí theo kiểu ô vuông la tinh có kết quả như sau

Bảng 4.2: Năng suất của 4 giống ngô (tấn/ha)

	Cột 1	Cột 2	Cột 3	Cột 4	Tổng
Hàng 1	6 (A)	9 (D)	8 (C)	10 (B)	33
Hàng 2	5 (D)	3 (B)	4 (A)	9 (C)	21
Hàng 3	7 (C)	3 (A)	5 (B)	9 (D)	24
Hàng 4	4 (B)	6 (C)	7 (D)	5 (A)	22
Tổng	22	21	24	33	100

Tính năng suất trung bình của các công thức:

Công thức	Tổng số	Trung bình
A	18	4,5
B	22	5,5
C	30	7,5
D	30	7,5

Bảng 4.3. Kết quả phân tích biến động của năng suất ngô
(thí nghiệm bố trí kiểu ô vuông la tinh)

Nguồn biến động	df	SS	MS	F tính[b]	F bảng	
					5%	1%
Hàng	3	22,5	7,5			
Cột	3	22,5	7,5	11,5**	4,76	9,78
Công thức	3	27	9			
Ngẫu nhiên	6	5	0,8			
Tổng	15					

Như vậy từ kết quả phân tích biến động ta có kết luận chung là năng suất của các giống ngô tham gia thí nghiệm khác nhau chắc chắn ở mức có ý nghĩa $\alpha = 0,01$. Nhưng để có kết luận năng suất của hai giống cụ thể nào đó có khác nhau một cách chắc chắn hay không thì ta cần so sánh số trung bình theo phương pháp LSD_α với các bước như sau:

Bước 1: Tính sự chênh lệch giữa số trung bình của công thức thí nghiệm với công thức muốn so sánh :

$$d_{ij} = \overline{X_i} - \overline{X_j}$$

Trong đó $\overline{X_i}$ và $\overline{X_j}$ là trung bình của công thức thứ *i* và *j*.

Bước 2: tính LSD_α tại mức có ý nghĩa $\alpha = 0,05$ hoặc $\alpha = 0,01$

$$LSD_\alpha = (t_\alpha)\sqrt{\frac{2s^2}{r}} = (t_\alpha)\sqrt{\frac{2MS_e}{r}}$$

Trong ví dụ này, $MS_e = 0,8$; $df_e = 6$; $r = 4$.

Từ bảng C tra được giá trị t với độ tự do $n = df_e = 6$;

Ta có: $t_{05} = 2,447$; $t_{01} = 3,707$.

$$LSD_{.05} = 2,447\sqrt{\frac{2x0,8}{4}} = 1,55 \text{ tấn/ha}$$

$$LSD_{.01} = 3,707\sqrt{\frac{2x0,8}{4}} = 2,34 \text{ tấn/ha}$$

Bước 3: So sánh giá trị $|d_{ij}|$ tính ở bước 1 với LSD tính ở bước 2

$d_{DC} = 0 < LSD$: 2 công thức D và C khác nhau không có ý nghĩa, hay 2 công thức cho năng suất như nhau

$d_{DB} = 2 > LSD_{05}$: 2 công thức D và B khác nhau có ý nghĩa ở mức độ tin cậy 95%, hay 2 công thức cho năng suất khác nhau ở mức xác suất 95%.

$d_{DA} = 3 > LSD_{01}$: 2 công thức D và A khác nhau có ý nghĩa ở mức độ tin cậy 99%, hay 2 công thức cho năng suất khác nhau ở mức xác suất 99%.

Tương tự như vậy ta so sánh và kết luận cho các cặp công thức khác và lập bảng so sánh như sau.

Bảng 4.4. So sánh sự chênh lệch của các cặp công thức với LSD

Cặp công thức	$d_{ij} = \overline{X_i} - \overline{X_j}$ (Chênh lệch giữa các công thức)	Đánh giá mức độ sai khác
D-C	7,5 - 7,5 = 0	NS
D-B	7,5 - 5,5 = 2,0	*
D-A	7,5 - 4,5 = 3,0	**
C-B	7,5 -5,5 = 2,0	*
C-A	7,5 - 4,5 = 3,0	**
B-A	5,5 - 4,5 = 1,0	NS

$LSD_{.05}$ = 1,55 tấn/ha

$LSD_{.01}$ = 2,34 tấn/ha

*: Sai khác có ý nghĩa ở mức xác suất $\alpha = 0,05$

**: Sai khác có ý nghĩa ở mức xác suất $\alpha = 0,01$

NS: Sai khác không có ý nghĩa.

Thông thường để cho dễ quan sát và so sánh người ta thường sắp xếp các giá trị trung bình của công thức theo thứ tự từ lớn đến nhỏ, sau đó so sánh các số có giá trị gần kề nhau. Trung bình công thức lớn nhất được đánh dấu là a. Sau đó nếu các công thức tiếp theo mà khác nhau có ý nghĩa thì đánh dấu bằng chữ khác (b, c, d). Nếu số tiếp theo mà không khác số trước có ý nghĩa thì đánh dấu bằng cùng một chữ. Ví dụ kết quả so sánh của bảng 4.3 được trình bày như sau:

Bảng 4.5. Xếp hạng và phân nhóm các công thức

Trung bình	Công thức	So sánh và phân nhóm	
		Với $\alpha = 0,05$	Với $\alpha = 0,01$
7,5	D	a	a
7,5	C	a	a
5,5	B	b	ab
4,5	A	b	b

$LSD_{.05}$ = 1,55 tấn/ha

$LSD_{.01}$ = 2,34 tấn/ha

Nhìn vào kết quả so sánh ta có kết luận như sau:

Với độ tin cậy là 95%, công thức A với B, C với D không khác nhau. Nhưng công thức A và B khác công thức C và D một cách chắc chắn.

Cũng tương tự như vậy với độ tin cậy 99% thì công thức A không khác công thức B, công thức B, C và D không khác nhau, nhưng công thức A khác công thức C, D một cách chắc chắn.

4.2.1.5. So sánh cặp đôi cho thí nghiệm 2 nhân tố bố trí kiểu ô chính ô phụ

Như phần trên đã trình bày, với kiểu ô chính ô phụ có 2 nhân tố và 2 hệ thống sai số, do đó có 4 kiểu so sánh cặp như sau.

- So sánh 2 số trung bình của khối chính (TB qua các mức của khối phụ)
- So sánh 2 số trung bình của khối phụ (TB qua các mức của khối chính)
- So sánh 2 trung bình của khối phụ ở cùng 1 công thức của khối chính
- So sánh 2 trung bình của khối chính ở cùng công thức của khối phụ

Minh họa bằng so sánh kết quả năng suất lúa của thí nghiệm ở phần 3.3.2. Kết quả phân tích biến động năng suất của thí nghiệm như sau:

Bảng 4.6. Kết quả phân tích phương sai cho thí nghiệm bố trí theo kiểu ô chính ô phụ

Nguồn biến động	df	SS	MS	F tính[b]	F bảng	
					5%	1%
Nhắc lại	2	1.082.577	541.228			
Đạm (A)	5	30.429.200	6.085.840	42,87**	3,33	5,64
Sai số (a)	10	1.419.678	141.968			
Giống (B)	3	89.888.101	29.962.700	85,71**	2,86	4,38
A.B	15	69.343.487	4.622.899	13,22**	1,96	2,58
Sai số (b)	36	12.584.873	349.580			
Tổng số	71	204.747.916				

Với kết quả như trên ta có thể kết luận là các mức N, các giống tham gia thí nghiệm cho năng suất lúa khác nhau một cách chắc chắn ở mức tin cậy 99%. Các giống lúa khác cho năng suất khác nhau trên các mức N khác nhau. Tuy nhiên để so sánh hai mức N hay hai giống lúa bất kỳ với độ tin cậy xác định thì ta phải so sánh trung bình các công thức theo LSD. Cách làm như sau:

Bảng 4.7. Tính sai số chuẩn cho 4 cặp so sánh

Cặp so sánh	s_d^a		Độ tự do
So sánh trung bình của hai mức N (ô chính -A)	$\sqrt{\frac{2E_a}{rb}}$	$\sqrt{\frac{141968x2}{3x4}} = 153,82$	$df_a = 10$
So sánh trung bình của hai giống (nhân tố phụ-B)	$\sqrt{\frac{2E_b}{r}}$	$\sqrt{\frac{349580x2}{3x6}} = 197,08$	$df_b = 36$
So sánh trung bình của 2 giống ở cùng 1 mức N	$\sqrt{\frac{2E_b}{ra}}$	$\sqrt{\frac{349,58x2}{3}} = 482,76$	$df_b = 36$
So sánh trung bình của 2 mức N cho cùng một giống	$\sqrt{\frac{2[(b-1)E_b + E_a]}{rb}}$	$\sqrt{\frac{(3x349,58 + 141,968)x2}{3x4}} = 418,11$	--

Trong đó: E_a = MS của sai số ngẫu nhiên a, E_b = MS của sai số ngẫu nhiên b, r = số nhắc lại, a = số công thức của khối chính, b = số công thức của khối phụ.

Tra bảng C giá trị t_{05} với độ tự do df = 10 và 36 ta có giá trị tương ứng là 2,228 và 2,030. Như vậy lần lượt ta tính được:

- LSD_{05} để so sánh các mức N là: 153,8 × 2,228 = 343 kg/ha
- LSD_{05} để so sánh các giống là: 197,1 × 2,030 = 400 kg/ha
- LSD_{05} để so sánh hai giống ở cùng một mức N là: 482,8 × 2,030 = 980 kg/ha
- LSD_{05} để so sánh hai mức N cho cùng một giống là:

$14,09 \times t_{AB} = 418,1 \times 2,054 = 859$ kg/ha.

Chú ý: vì t_{AB} không tính được qua tra bảng mà tính theo phương pháp sau:

$t_{AB} = [(b-1)(E_b)(t_B) + E_a(t_A)]: [(b-1)(E_b) + E_a] =$

[(4-1) x 349580 x 2,030 + 141986 x 2,228] : [(4-3) x 349580 + 141968] = 2,054

Với năng suất lúa của các mức N và các giống được trình bày ở bảng 3.16 ta tính toán và được bảng sau:

Bảng 4.8. Năng suất lúa theo mức đạm và giống lúa trong thí nghiệm thiết kế theo kiểu ô chính- ô phụ (kg/ha)

Công thức		Các giống				Trung bình
		V1 (IR8)	V2 (IR5)	V3 (C463)	V4 (Peta)	
		(1)	(2)	(3)	(4)	(5)
N_0	(1)	4253	4306	3183	4481	4056
N_1	(2)	5672	5982	5443	4816	5478
N_2	(3)	6400	6259	5994	4812	5866
N_3	(4)	6733	6895	6014	3816	5864
N_4	(5)	7563	6951	6687	2047	5812
N_5	(6)	8701	6540	6065	1881	5797
Trung bình	(8)	6554	6156	5564	3642	

Với kết quả tính toán LSD_{05} trên đây ta có thể dùng so sánh các số trung bình công thức ở bảng 4.7 như sau:

- Dùng LSD_{05} = 343 kg/ha để so sánh năng suất các mức N bằng cách so sánh các số trung bình ở cột (5) bảng 4.7 ta thấy: các mức bón N từ N_1 đến N_5 cho năng suất lúa không khác nhau ở độ tin cậy 95% nhưng năng suất lúa ở các công thức này đều hơn hẳn năng suất lúa ở công thức N_0 ở độ tin cậy 95%.

- Dùng LSD_{05} = 400 kg/ha để so sánh năng suất của các giống, bằng cách so sánh các số trung bình ở hàng (8) bảng 4.7 cho thấy năng suất của giống IR8 và giống IR5 không khác nhau nhưng cả hai giống này lại cho năng suất khác giống C463 và Peta chắc chắn ở độ tin cậy 95%. Giống C463 có năng suất cao hơn giống Peta chắc chắn ở độ tin cậy 95%.

- Tương tự ta dùng LSD_{05} = 980 kg/ha để so sánh số trung bình của hai giống trong cùng một mức N (cùng một hàng của các hàng 1, 2, 3, 4, 5, 6). Và dùng LSD_{05} = 859 kg/ha để so sánh số trung bình hai mức N cho cùng một giống (cùng một cột của các cột 1, 2, 3, 4).

4.2.2. So sánh Duncan: So sánh phân hạng dựa trên giá trị LSD (Multiple range test - Little M. và Hills, 1978)

4.2.2.1. Nguyên tắc so sánh Duncan

So sánh từng cặp như trên rất khó trình bày kết quả tổng quát của tất cả các phép so sánh, nhất là khi ta muốn quan tâm đến tất cả các cặp so sánh. So sánh Duncan là so sánh phân hạng các công thức dựa trên giá trị LSD và khoảng cách giữa các công thức trong bảng xếp thứ tự số trung bình các công thức. Vì vậy, khác với kết quả của so sánh LSD là chỉ ra sự khác nhau của một cặp công thức bất kỳ thì kết quả của so sánh Duncan là xếp hạng các công thức.

So sánh Duncan không chỉ dựa vào duy nhất một giới hạn là giá trị LSD mà là một số giới hạn "Sai khác có ý nghĩa". Giá trị "sai khác có ý nghĩa" khác nhau đối với những cặp công thức khác nhau. Giá trị "sai khác có ý nghĩa" được dùng để so sánh trong Duncan được ký hiệu là D_p với cách tính như sau:

$$D_p = R_p \times LSD$$

Trong đó

D_p: Giá trị sai khác nhau nhỏ nhất của một khoảng cách nào đó

R_p : Giá trị bảng tính phụ thuộc vào độ tự do của ngẫu nhiên và p

p: Khoảng cách giữa 2 số trung bình được so sánh trong bảng xếp thứ tự. p = 2 + n. Lý giải, 2 số trung bình nếu đứng cạnh nhau (cách nhau 0 công thức) khoảng cách p = 2; nếu cách nhau 1 công thức thì p = 3 và nếu cách nhau n công thức thì p = 2 + n.

Tương tự như so sánh LSD, nguyên tắc đánh giá của Duncan là: Hai công thức được coi là khác nhau nếu sự chênh lệch (sự sai khác) giữa chúng lớn hơn Dp. Có thể tóm tắt như sau:

Nếu sự sai khác > Dp, kết luận 2 công thức khác nhau có ý nghĩa

Nếu sự sai khác ≤ Dp, kết luận 2 công thức không khác nhau

4.2.2.2. Cách so sánh Duncan

Minh họa cách so sánh bằng so sánh năng suất của các công thức trong thí nghiệm "Nghiên cứu ảnh hưởng của 5 mức phân đạm tới năng suất sắn" có r = 4; bố trí theo khối ngẫu nhiên hoàn chỉnh với kết quả ghi ở bảng 4.9 sau:

Bảng 4.9. Năng suất sắn ở các công thức bón N khác nhau (tạ/ha)

Công thức	Nhắc lại				Tổng	Trung bình
	I	II	III	IV		
1	47,8	46,9	45,3	44,1	184,1	46,0
2	53,7	50,3	50,6	48,0	202,6	50,7
3	46,7	42,0	43,4	40,7	172,8	43,2
4	48,0	47,0	45,9	45,7	186,6	46,7
5	41,8	40,0	43,0	41,6	166,4	41,6

Biết MS_e = 1,63 với df = (r-1)(t-1) = 4 × 3 = 12, tra bảng C ta có t_{01} = 3,3055; t_{05} = 2,197

Bước 1:

Tính LSD

$$LSD_{05} = t\sqrt{\frac{2s^2}{r}} = 2,197\sqrt{\frac{2MS_e}{r}} = 2,197\sqrt{\frac{2(1,63)}{4}} = 1,98$$

Bước 2:

Tính giá trị D cho các cặp so sánh có khoảng cách p khác nhau. Có 5 số trung bình trong dãy so sánh vì vậy p có các giá trị là 2, 3, 4 và 5. Với n = df_e = 12, tra bảng A.4 ở mức 5% có các giá trị R như sau:

Bảng 4.10. Tính giá trị D tại các khoảng cách p

P	2	3	4	5
R	1,00	1,05	1,08	1,1
D = R x LSD	1,98	2,08	2,14	2,18

Bước 3: Xếp thứ tự các số trung bình lớn đến nhỏ và đánh giá sự sai khác bằng cách so sánh sự chênh lệch giữa chúng với giá trị D tương ứng.

Bảng 4.11. Xếp thứ tự các số trung bình theo giá trị từ lớn đến nhỏ

Công thức	2	4	1	3	5
Trung bình	50,7	46,7	46,0	43,2	41,6

(1) So sánh số trung bình lớn nhất (T2) với các số trung bình nhỏ hơn

- So sánh với số trung bình nhỏ nhất (T5) với *D* tại p = 5 (vì 2 số trung bình T2 và T5 cách nhau 3 công thức T4, T1, T3 nên p = 2 + 3 = 5)

50,7- 41,6 = 9,1> 2,18 vậy 2 công thức khác nhau có ý nghĩa.

- So sánh với số trung bình bé thứ 2 (T3) dùng D với p = 2 + 2 = 4

50,7- 43,2 = 7,5 > 2,14 vậy 2 công thức khác nhau chắc chắn.

- So sánh với số trung bình bé thứ 3 (T1) dùng D với p = 2 + 1 = 3

50,7- 46,0 = 4,7 > 2,08 vậy 2 công thức khác nhau chắc chắn.

- So sánh với số trung bình T4 dùng D với p = 2 (2 số cạnh nhau)

50,7 - 46,7 = 4,0 > 1,98 vậy 2 công thức khác nhau chắc chắn

(2) So sánh số trung bình lớn thứ 2 (T4) với các số trung bình nhỏ hơn

- So với số trung bình nhỏ nhất, dùng D với p = 4

46,7 - 41,6 = 5,1 > D_4 = 2,14 vậy 2 công thức khác nhau có ý nghĩa

- So với số trung bình nhỏ thứ 2 dùng D với p = 3

46,7 - 43,2 = 3,5 > D_3 = 2,0 vậy 2 công thức khác nhau có ý nghĩa

- So với số trung bình nhỏ thứ 3 dùng D với p = 2

46,7 - 46,0 = 0,72 < D_2 = 1,98 vậy 2 công thức không khác nhau

Tương tự

(3) So sánh số trung bình lớn thứ 3 (T1) với các số nhỏ hơn,

46,0 - 41,6 = 4,4 > D_3 = 2,08 vậy 2 công thức khác nhau có ý nghĩa

46,0 - 43,2 = 2,8 > D_2 = 1,98 vậy 2 công thức khác nhau có ý nghĩa

(4) So sánh số trung bình lớn thứ 4 với số trung bình nhỏ hơn

43,2 - 41,6 = 1,62 < D_2 = 1,98 vậy 2 công thức không khác nhau

Bước 4:

Lập bảng kết quả so sánh bằng xếp hạng các công thức không khác nhau vào nhóm có cùng chữ cái như bảng 4.12.

Bảng 4.12. Kết quả so sánh số trung bình các công thức theo Duncan

Công thức	2	4	1	3	5
Trung bình	50,7	46,7	46,0	43,2	41,6
Xếp hạng	a	b	b	c	c

Ghi chú: Những số trung bình có cùng chữ cái là không khác nhau ở mức $\alpha = 5\%$.

Kết luận: Dùng lượng đạm khác nhau có ảnh hưởng khác nhau đến năng suất sắn, tuy nhiên giữa công thức 5 với 3 và công thức 1 với 4 không có sự khác nhau về năng suất ở mức xác suất 95%.

4.3. SO SÁNH NHÓM (CONTRASTS)

4.3.1. Ý nghĩa của so sánh nhóm

Chúng ta đôi khi không chỉ muốn so sánh cặp công thức mà muốn so sánh giữa các nhóm công thức. Ví dụ so sánh nhóm các giống cũ với nhóm các giống mới, so sánh nhóm công thức bón phân hữu cơ với nhóm công thức bón phân hoá học. Hơn nữa, nhiều tác giả (Steel, 1997; Little và Hills, 1978) cho rằng: Trong phân tích biến động, đặc biệt trong những thí nghiệm bao gồm nhiều công thức, kết luận từ phân tích biến động có thể chỉ là kết luận chung cho hầu hết các trường hợp nhưng có thể bỏ qua những sai khác có ý nghĩa khác. Có nghĩa là tuy kết luận chung từ so sánh với F lý thuyết cho thấy không sai khác nhưng khi dùng so sánh nhóm thì có khác nhau. Như vậy liệu ta có thể tìm ra những sai khác có ý nghĩa đó không?

Thực tế hiện nay phương pháp so sánh nhóm chưa được sử dụng một cách rộng rãi do vai trò quan trọng của nó chưa được đánh giá một cách đúng mức. Theo Little and Hill (1978) thì phương pháp này có một số ưu điểm nổi bật như sau: (1) Cho phép nhà nghiên cứu có câu trả lời trực tiếp cho những vấn đề quan tâm; (2) Phương pháp tính toán đơn giản; (3) Có thể dùng để kiểm tra tổng bình phương biến động của công thức.

Muốn so sánh nhóm cần tính tổng bình phương (SS) cho từng nhóm công thức muốn so sánh và dùng F để kiểm tra sự sai khác có ý nghĩa hay không. Việc đầu tiên là xác định hệ số so sánh sau đó tính SS để tìm giá trị F.

Dưới đây giới thiệu hai cách so sánh nhóm: So sánh giữa các nhóm và so sánh theo xu hướng. So sánh xu hướng thường dùng với những công thức thí nghiệm về số lượng để tìm xu hướng biến động của biến phụ thuộc theo biến độc lập. Ví dụ như tìm xu hướng biến động của năng suất lúa khi ta bón N với các liều lượng khác nhau (0, 30, 60, 90 kg/ha).

4.3.2. Xác định hệ số so sánh trong so sánh nhóm

Để so sánh nhóm trước tiên ta phải xác định nhóm công thức muốn so sánh và từ đó xác định hệ số so sánh. Hai nhóm công thức so sánh luôn đối lập nhau, ví dụ so sánh nhóm các giống lúa truyền thống với nhóm các giống lúa mới; So sánh ảnh hưởng của các công thức bón phân khoáng với các công thức bón phân hữu cơ.... Xác định hệ số so sánh dựa theo một số nguyên tắc đơn giản như sau:

- Nguyên tắc 1: Nếu số công thức ở hai nhóm mà tương đương nhau thì đơn giản là gán hệ số +1 cho các công thức của nhóm này và hệ số -1 cho các công thức của nhóm kia. Việc nhóm nào nhận giá trị dương hay âm không quan trọng. Nguyên tắc chung nhất là tổng các hệ số của hai nhóm trong một lần so sánh bao giờ cũng bằng 0.

- Nguyên tắc 2: Nếu so sánh hai nhóm có số lượng công thức khác nhau thì gán hệ số cho nhóm này bằng số công thức của nhóm kia và ngược lại, nhưng trái dấu. Ví dụ so sánh hai nhóm: Nhóm một có 2 công thức và nhóm hai có 3 công thức thì hệ số của các công thức nhóm một sẽ là: +3 và hệ số của các công thức nhóm hai sẽ là -2.

- Nguyên tắc 3: Giảm hệ số của hai nhóm xuống số nguyên nhỏ nhất có thể. Ví dụ nếu so sánh hai nhóm: Nhóm 1 có 2 công thức và nhóm 2 có 4 công thức thì hệ số của hai nhóm sẽ là: +4 và -2. Hệ số như vậy có thể giảm xuống còn: +2 và -1.

- Nguyên tắc 4: Nếu muốn so sánh tương tác giữa hai nhóm thì chỉ việc nhân hệ số so sánh của hai nhóm đó lại với nhau.

Ví dụ minh họa cách xác định hệ số so sánh:

Thí nghiệm "So sánh năng suất của ba giống lúa: Bao thai (B), Khang dân (K) Tạp giao 1 (T) và trên nền không phân (P0), phân hoá học (Ph), phân hữu cơ (Pc)" . Như vậy ta sẽ có 9 công thức như sau: B0, BPh, BPc, K0, KPh, KPc, T0, TPh, TPc. Tuỳ theo mục đích nhà nghiên cứu mà có thể đặt những câu hỏi để xác định những nhóm so sánh và lập thành bảng hệ số so sánh (bảng 4.12) cụ thể như sau:

1. Năng suất của các công thức có giống lúa thuần (Bao thai và Khang dân) có khác với năng suất các công thức có giống lúa lai Tạp giao 1 không (T so với B + K)?. Ta có 2 nhóm, nhóm 1 lúa thuần có 6 công thức và nhóm 2 lúa lai có 3 công thức. Theo nguyên tắc 2 thì nhóm 1 sẽ có hệ số là +3 và nhóm 2 là -6. Giảm ước theo nguyên tắc 3 sẽ còn: +1 và -2 (xem bảng 4.12).
2. Năng suất khi bón phân hữu cơ và phân hoá học có khác nhau không (Ph so với Pc)?. Ta có 2 nhóm so sánh có số công thức bằng nhau (là 3) và theo nguyên tắc 1 thì các công thức sẽ nhận giá trị +1 và -1, còn các công thức không tham gia so sánh là 0 như bảng 4.12.
3. Bón phân năng suất có khác so với không bón phân (0P so với P)?. Ta có 2 nhóm để so sánh là nhóm 1 không bón phân (có 3 công thức) và nhóm 2 có bón phân (có 6 công thức). Tương tự như câu hỏi 1 ta cũng sẽ gán giá trị +2 và -1 cho các công thức của hai nhóm.
4. Năng suất lúa lai và lúa thuần biến đổi thế nào trong môi trường có phân và không phân? Với câu hỏi này ta chỉ cần nhân hệ số so sánh của câu hỏi 1 với câu hỏi 3, xem hệ số ở bảng 4.13.
5. Năng suất lúa lai và lúa thuần biến đổi thế nào khi bón phân hoá học và bón phân hữu cơ? với câu hỏi này ta chỉ cần nhân hệ số so sánh của câu hỏi 1 với câu hỏi 2 và xem hệ số ở bảng 4.12.

Bảng 4.13. Hệ số so sánh cho các nhóm so sánh

Nhóm so sánh	Công thức								
	B0	BPh	BPc	K0	KPh	KPc	T0	TPh	TPc
1. T vs. B + K	+1	+1	+1	+1	+1	+1	-2	-2	-2
2. Ph vs. Pc	0	-1	+1	0	-1	+1	0	-1	+1
3. 0P vs. P	+2	-1	-1	+2	-1	-1	+2	-1	-1
4. (T vs. B + K) vs. (0P vs. P)	+2	-1	-1	+2	-1	-1	-4	+2	+2
5.(T vs. B+K) vs. (Ph vs. Pc)	0	-1	+1	0	-1	+1	0	+2	-2

vs: so với

Trong so sánh nhóm thì ta tự xác định hệ số so sánh, còn trong so sánh theo xu hướng thì ta tra hệ số so sánh trong bảng A11 của Little và Hills (1978) như ví dụ dưới đây.

4.3.3. Cách so sánh nhóm và xu hướng

Áp dụng cho thí nghiệm "Nghiên cứu về ảnh hưởng của lượng hạt giống lúa gieo tới năng suất lúa" có số liệu như bảng 4.13.

Bảng 4.14. Năng suất lúa của giống IR8 với 6 lượng hạt giống gieo

Công thức (kg hạt/ha)	Năng suất (kg/ha) (X_i)				Tổng của công thức (T)	Trung bình công thức
	I	II	III	IV		
25	5113	5398	5307	4678	20496	5124
50	5346	5952	4719	4264	20281	5070
75	5272	5713	5483	4749	21217	5304
100	5164	4831	4986	4410	19391	4848
125	4804	4848	4432	4748	18832	4708
150	5254	4542	4919	4098	18813	4703
Tổng nhắc lại (R)	30953	31284	29846	26947		
Tổng lớn (G)					119030	
TB tổng						4960

Nguồn: Gomez, 1984

Thí nghiệm có kết quả phân tích biến động như sau:

Bảng 4.15. Phân tích biến động của năng suất lúa

Nguồn biến động	df	SS	MS	$F_{tính}$ [b]	$F_{bảng}$	
					5%	1%
Nhắc lại	3	1.944.361	648.120			
Công thức	5	1.198.331	239.666	$2,17^{ns}$	2,90	4,56
Ngẫu nhiên	15	1.658.376	110.558			
Tổng	23	4.801.068				

Kết luận: Ảnh hưởng của các lượng hạt giống gieo khác nhau tới năng suất lúa là không có ý nghĩa. Tuy nhiên có một vài câu hỏi nhà nghiên cứu muốn tìm hiểu như sau.

Năng suất lúa của những công thức có lượng hạt giống gieo ít hơn 100 kg có tương đương với năng suất ở những công thức có lượng giống gieo lớn hơn hoặc bằng 100 kg/ha không? Để trả lời được câu hỏi này ta cần so sánh 2 nhóm công thức: nhóm 1: Những công thức có lượng hạt gieo ít hơn 100 kg/ha đó là các công thức: 25, 50 và 75 kg/ha; nhóm 2: Những công thức có lượng hạt gieo từ 100 kg/ha trở lên đó là các công thức: 100, 125, 150 kg/ha. Như vậy mỗi nhóm có 3 công thức và theo nguyên tắc 1 thì hệ số so sánh sẽ là -1 và +1 (xem bảng 4.15).

Khi tăng lượng giống gieo thì năng suất lúa có thay đổi không? Nếu năng suất lúa tăng (hay giảm) thì quan hệ giữa năng suất (y) và lượng giống gieo (x) theo phương trình bậc 1 ($y = ax + b$) hay theo phương trình bậc hai ($y = ax^2 + bx + c$). Đó là so sánh xu hướng, ta cần tra hệ số cho cách so sánh này ở A.11 với số công thức n = 6 và chọn c1 (linear - Phương trình bậc 1) và c2 (quadratic - Phương trình bậc 2). Kết quả trình bày ở bảng 4.15.

Các bước thực hiện

Bước 1: xác định hệ số so sánh

Từ kết quả của bảng 4.14 và tra hệ số so sánh (với n= 6) từ bảng A11 ta có:

Bảng 4.16. Hệ số cho so sánh nhóm và so sánh xu hướng (Ci)

So sánh	Công thức					
	25	50	75	100	125	150
Tổng của công thức (Ti)	20496	20281	21217	19391	18832	18813
1. Các công thức của nhóm 1 so với các công thức của nhóm 2	-1	-1	-1	+1	+1	+1
2. Tuyến tính (bậc một)	-5	-3	-1	+1	+3	+5
3. Bậc hai	+5	-1	-4	-4	-1	+5

Bước 2: Tính SS

Công thức tính: $SS = \sum \frac{(C_i T_i)^2}{rx \sum C_i^2}$

Trong đó:

C_i: Hệ số so sánh;

T_i: Tổng của công thức i và

R: là số lần nhắc lại.

Như vậy ta có tổng bình phương biến động của mỗi phép so sánh là:

So sánh 2 nhóm 1 và 2:

$$SS1=\frac{\left[(-1)x20496+(-1)x20281+(-1)x21217+(1x19391)+(1x18832)+(1x18813)\right]^2}{4x\left[(-1)^2+(-1)^2+(-1)^2+1^2+1^2+1^2\right]}$$

= 1024240, 2 với độ tự do bằng 1.

So sánh xu hướng: Quan hệ phụ thuộc giữa năng suất với lượng giống gieo quan hệ theo phương trình bậc một.

$$SS2=\frac{\left[(-5)x20496+(-3)x20281+(-1)x21217+(1x19391)+(3x18832)+(5x18813)\right]^2}{4x\left[(-5)^2+(-3)^2+(-1)^2+1^2+3^2+5^2\right]}$$

= 760034,8 với độ tự do bằng 1.

So sánh xu hướng: Quan hệ phụ thuộc giữa năng suất với lượng giống gieo quan hệ theo phương trình bậc hai α

$$SS3=\frac{\left[(5x20496)+(-1)x20281+(-4)x21217+(-4)x19391+(-1)x18832+(5x18813)\right]^2}{4x\left[5^2+(-1)^2+(-4)^2+(-4)^2+(-1)^2+5^2\right]}$$

= 74404,8 với độ tự do bằng 1.

Từ những kết quả tính toán trên ta lập bảng biến động mới như sau

Bảng 4.17. Phân tích biến động của năng suất lúa

Nguồn biến động	df	SS	MS	$F_{tính}$ [b]	$F_{bảng}$	
					5%	1%
Nhắc lại (rep.)	3	1.944.361	648.120			
Công thức (treatment)	5	1.198.331	239.666	$2,17^{ns}$	2,90	4,56
(25 -75 vs 100 -150)	1	1.024.240	1.024.240	$9,26^{**}$	4,54	8,68
Phương trình bậc 1	1	760.034	760.034	$6,87^{*}$	4,54	8,68
Phương trình bậc 2	1	74.404	74.404	$0,67^{ns}$.	4,54	8,68
Ngẫu nhiên	15	1.658.376	110.558			
Tổng	23	4.801.068				

Căn cứ vào kết quả phân tích phương sai, ta có kết luận sau:

- Với lượng hạt giống lúa gieo từ 25 -75 kg/ha có năng suất cao hơn chắc chắn so với lượng hạt giống lúa gieo là 100-150 kg/ha ở độ tin cậy 99%.

- Khi tăng lượng hạt giống lúa gieo từ 25 tới 150 kg thì năng suất lúa giảm theo kiểu phương trình tuyến tính bậc một với độ tin cậy là 99%.

Rõ ràng ta có thể nhận thấy rằng nếu sử dụng so sánh nhóm để phân tích kết quả của thí nghiệm thì ta có thể có những kết luận rất cụ thể theo đúng ý muốn của nhà nghiên cứu.

Chương 5

PHÂN TÍCH TƯƠNG QUAN

5.1. KHÁI NIỆM VỀ TƯƠNG QUAN VÀ PHÂN TÍCH TƯƠNG QUAN

5.2.1. Khái niệm về tương quan

Trong tự nhiên có nhiều mối quan hệ phức tạp ảnh hưởng lẫn nhau. Với trồng trọt tương quan thường gặp là: Các quan hệ giữa các đặc tính của cây với nhau; Quan hệ giữa phản ứng của cây với nhân tố thí nghiệm; Quan hệ giữa phản ứng của cây với môi trường (như sản lượng chè có quan hệ với lượng mưa trong năm). Nếu biết được những mối quan hệ đó giúp ta tiên đoán được những gì sẽ xảy ra và có quyết định hợp lý, tiết kiệm và đem lại hiệu quả cao. Một hàm toán học biểu diễn mối quan hệ đó gọi là hàm tương quan.

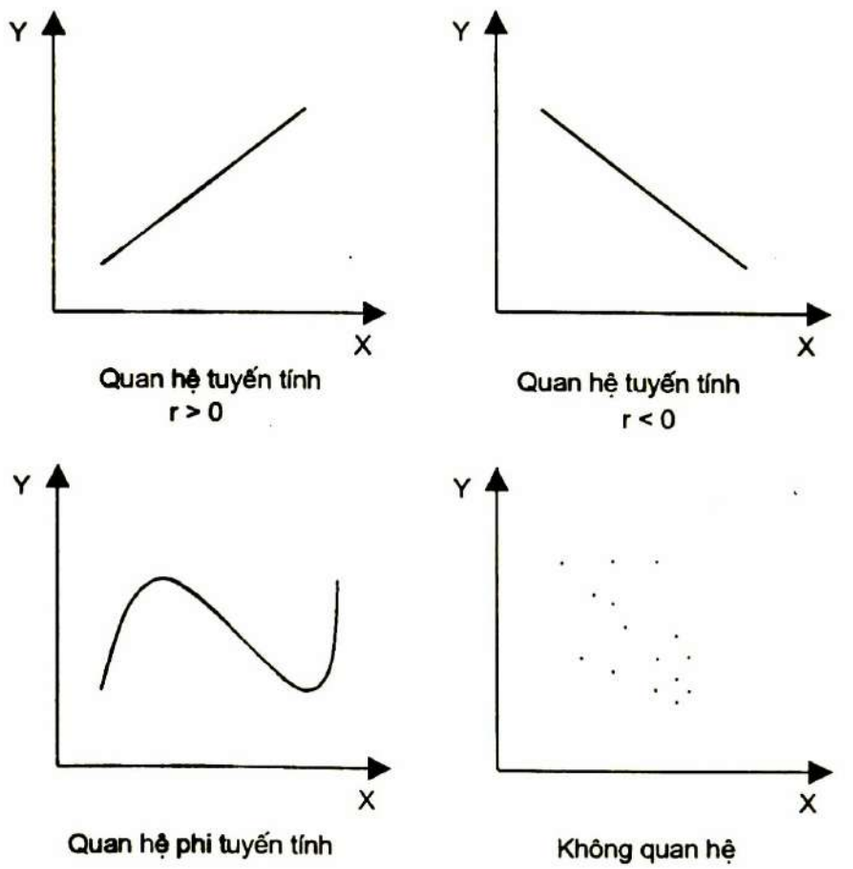
Có thể phân loại tương quan theo:

Số biến tham gia: Quan hệ đơn (single) hay đa biến (multiple)

Loại quan hệ:

Tuyến tính: Quan hệ theo đường thẳng (linear)

Phi tuyến tính: Quan hệ theo đường cong (curvelinear hay non-linear). Quan hệ theo đường cong có thể là theo các hàm khác nhau như hàm mũ, logarit, v.v.



Đồ thị 5.1. Đồ thị thể hiện các mối quan hệ giữa biến X và Y (r là hệ số tương quan)

Để xác định mối quan hệ giữa các biến qua đồ thị (như biến X và Y ở sơ đồ 5.1), ta vẽ đồ thị phân tán của 2 biến đó. Đồ thị phân tán của biến Y đối với biến X là tập hợp các điểm $M(x_i,y_i)$ trong hệ toạ độ vuông góc. Dựa vào đồ thị phân tán ta có thể xác định được dạng của hàm quan hệ giữa 2 biến X và Y. Tuy nhiên nếu là hàm đa biến thì không thể dùng phương pháp này vì với hàm đa biến ta không thể đưa lên đồ thị được (tham khảo Gomez, 1984).

Cách nhanh nhất để xác định dạng quan hệ giữa 2 biến X và Y là vẽ đồ thị bằng Excel (excel cho phép vẽ tất cả những dạng đồ thị theo yêu cầu và kèm theo nó là hệ số tương quan r). Việc quyết định chọn hàm nào phụ thuộc vào r nên chọn hàm có hệ số tương quan r lớn nhất. (Phần này sẽ được đề cập và hướng dẫn ở phần sau và khi thực hành chương 6).

Ví dụ phân loại hàm tuyến tính theo số biến độc lập,

Quan hệ tuyến tính đơn giản có một biến: $Y = a + bX$

Quan hệ tuyến tính đa biến: $Y = a + bX + cZ + dF +...$

Quan hệ phi tuyến tính đơn biến: $Y = f(X)$

Quan hệ phi tuyến tính đa biến: $Y = f(X, Z, F...)$

Trong đó Y là biến phụ thuộc, còn lại là biến độc lập

Tuỳ theo số mũ của biến độc lập đường biểu diễn hàm có dạng khác nhau như sau:

Bảng 5.1. Phân loại hàm tương quan theo số mũ của biến

Số mũ	Tên hàm	Tên đường biểu diễn
1	Tuyến tính	Đường thẳng
2	Bậc 2	Parabol
3	Bậc 3	Parabol bậc 3
4	Bậc 4	Parabol bậc 4
5	Bậc 5	Parabol bậc 5

Dưới đây là một số dạng quan hệ hàm thường gặp:

Hàm tuyến tính: Ví dụ quan hệ năng suất ngô (Y) và mức bón đạm (X):

$$Y = a + bX$$

Hàm Exponential: Ví dụ hàm sinh trưởng thực vật (Y) sau số ngày mọc (X):

$$Y = ae^{bX}$$

Hàm Asymptotic (dạng đặc biệt của Exponential với $b < 1$). Ví dụ hàm phân huỷ tàn dư thực vật (Y) theo thời gian (X):

$$Y = a +(1/2)^X$$

Hàm Polynomials: Ví dụ hàm năng suất thu hoạch của đậu xanh (Y) sau ngày thu hoạch đầu tiên (X):

$$Y = a + bX + cX^2 + dX^3 + ...$$

Chú ý khi phân tích tương quan, chọn hàm càng đơn giản (nhưng phù hợp với sự phân bố của kết quả quan sát) thì càng tốt.

5.1.2. Phân tích tương quan *(regression analysis)*

Phân tích tương quan đề cập đến mối quan hệ giữa các biến. Phân tích tương quan bao gồm xác định hàm tương quan và hệ số tương quan. Hàm tương quan mô tả mối quan hệ ảnh hưởng của một hoặc nhiều biến độc lập lên một biến phụ thuộc, nhờ vậy ta có thể dự đoán được giá trị của biến phụ thuộc qua giá trị biến độc lập. Vì chúng ta không nghiên cứu toàn bộ tổng thể mà chỉ nghiên cứu thông qua quan sát mẫu, do đó phân tích tương quan ước lượng tham số của các biến trong hàm tương quan qua mẫu (tập hợp những giá trị quan sát được) và xác định độ tin cậy của các tham số đó. Độ tin cậy cho biết hệ số ước lượng có biểu thị đúng mối quan hệ giữa các biến xảy ra trong tự nhiên hay không. Phân tích hệ số tương quan là đánh giá hàm ước lượng có phản ánh đúng quy luật của sự kiện hay không.

Chú ý: Trong phân tích tương quan cần phải xác định quan hệ giữa biến phụ thuộc và biến độc lập phải là có ý nghĩa. Ví dụ như trong thí nghiệm phản ứng của năng suất lúa với đạm, rõ ràng năng suất phụ thuộc vào mức đạm. Nhưng năng suất hạt và hàm lượng protein không có quan hệ phụ thuộc rõ ràng.

5.2. PHÂN TÍCH TƯƠNG QUAN TUYẾN TÍNH ĐƠN GIẢN

5.2.1. Tương quan tuyến tính đơn giản (simple linear regression)

Trong đồ thị phân tán, nếu các điểm $M(x_i , y_i)$ qui tụ xung quanh một đường thẳng (D) ta nói hai biến ngẫu nhiên Y và X có một sự tương quan tuyến tính. Đường thẳng (D) được gọi là đường hồi quy tuyến tính. Vì thế phân tích tương quan còn được gọi là phân tích hồi quy.

Mô hình hồi quy tuyến tính có dạng sau:

$$Y_j = \alpha + \beta X_i + e_i$$

Trong đó:

Y_i : Giá trị của biến phụ thuộc Y trong lần quan sát thứ i;

X_i: Giá trị của biến độc lập X trong lần quan sát thứ i

α : Giao điểm giữa trục tung với đường biểu diễn quan hệ X và Y;

e_i : Giá trị đối với sự dao động ngẫu nhiên hay sai số trong lần quan sát thứ i.

β : Thông số diễn tả độ dốc của đường hồi quy hay diễn tả sự thay đổi của giá trị trung bình của biến phụ thuộc Y khi biến độc lập X thay đổi một đơn vị.

Có thể định nghĩa tương quan tuyến tính như sau: Mối quan hệ giữa 2 biến bất kỳ là theo đường thẳng nếu sự thay đổi là một hằng số. Ví dụ với đường Y= 1+ 2X thì Y luôn tăng 2 đơn vị ứng với mỗi một đơn vị tăng lên của X. Cụ thể khi x tăng từ 0 đến 1 thì y tăng từ 3 lên 5.

Chúng ta có thể ước lượng các tham số (α, β) của phương trình hồi quy tuyến tính của tổng thể bằng cách sử dụng số liệu mẫu ngẫu nhiên thu thập được, khi đó các tham số được ký hiệu là a, b và hàm hồi quy đơn giản có dạng:

$$Y = a + bx$$

Trong đó:

- Y là ước lượng của giá trị trung bình của Y đối với biến X đã biết.
- a: Giá trị ước lượng của α, là giao điểm của Y với trục tung
- b: Giá trị ước lượng của β, là độ dốc của đường hồi quy

Dựa vào phương trình hồi quy ta có thể dự đoán Y khi biết X. Nếu X thay đổi mà Y không thay đổi thì X và Y không có tương quan.

5.2.2. Phân tích tương quan tuyến tính đơn giản

Phân tích tương quan tuyến tính là ước lượng giá trị của hệ số a và b trong hàm tương quan tuyến tính và kiểm tra mức độ có ý nghĩa của chúng. Số liệu yêu cầu để áp dụng cho phân tích là phải có n cặp số liệu của Y và X. Với điều kiện n > 2.

Bước 1:

Tính giá trị trung bình của biến X($\overline{X}$) và biến Y($\overline{Y}$), tổng hiệu chỉnh của $\sum x^2$, $\sum y^2$ và $\sum xy$ của biến X và Y.

$$\overline{X} = \frac{\sum X}{n}$$

$$\overline{Y} = \frac{\sum Y}{n}$$

$$\sum x^2 = \sum_{i=1}^{n} (X_i - \overline{X})^2$$

$$\sum y^2 = \sum_{i=1}^{n} (Y_i - \overline{Y})^2$$

$$\sum xy = \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})$$

Trong đó (X_i và Y_i) là cặp thứ i trong các giá trị của X và Y.

Minh họa bằng phân tích hệ số tương quan giữa năng suất lúa với mức đạm bón trong thí nghiệm "Nghiên cứu phản ứng của năng suất lúa với mức bón đạm (N) khác nhau" được trình bày ở bảng 5.2.

Bảng 5.2: Tính một số giá trị để tìm các tham số trong tương quan tuyến tính một biến số trong quan hệ giữa năng suất lúa với mức đạm bón

Mức đạm kg/ha (X)	Năng suất kg/ha(Y)	Sai khác so với trung bình		Bình phương của sai khác		(x)(y)
		$X=\left(x_i-\overline{x}\right)$	$Y=\left(y_i-\overline{y}\right)$	x^2	y^2	
0	4.230	-75	-1640,75	5.625	2.692.061	123.056
50	5.442	-25	-428,75	625	183.827	10.719
100	6.661	25	790,25	625	624.495	19.756
150	7.150	75	1279,25	5.625	1.636.481	95.944
Tổng						
300	23.483	0	0,00	12.500	5.136.864	249.475
Trung bình						
75	5.870,75					

Nguồn: Gomez, 1984

Thí nghiệm có n = 4 cặp giá trị của năng suất Y và mức đạm X.

Bước 2:

Ước lượng các hệ số a và b trong hàm tương quan tuyến tính.

$$b = \frac{\sum xy}{\sum x^2} = \frac{249.475}{12.500} = 19,96$$

$$a = \overline{Y} - b\overline{X} = 5.870,75 - (19,96)(75) = 4.374$$

Ta có hàm tương quan tuyến tính ước lượng mối quan hệ giữa năng suất lúa và mức bón đạm như sau:

$$Y = a + bX = 4.374 + 19,96X \text{ ; Trong khoảng } 0 \le X \le 150$$

$0 \le X \le 150$ có nghĩa là phương trình chỉ đúng trong khoảng mức đạm từ 0 - 150N, đó là trong khoảng các mức N đã thử nghiệm. Nếu ở ngoài khoảng trên cần phải làm thí nghiệm chứ không thể dùng hàm trên để ước lượng.

Bước 3: Vẽ đồ thị quan hệ Y và X.

Vẽ đồ thị biểu diễn hàm tuyến tính ước lượng được ở bước 2 trên đó xác định các giá trị quan sát được. Cách vẽ:

Đánh dấu những giá trị quan sát được trên đồ thị. Trong ví dụ trên có 4 điểm quan sát được đó là toạ độ của các điểm Xi và Yi.

Dùng đường hồi quy ước lượng giá trị Y tương ứng với giá trị của X_{min} và X_{max}.

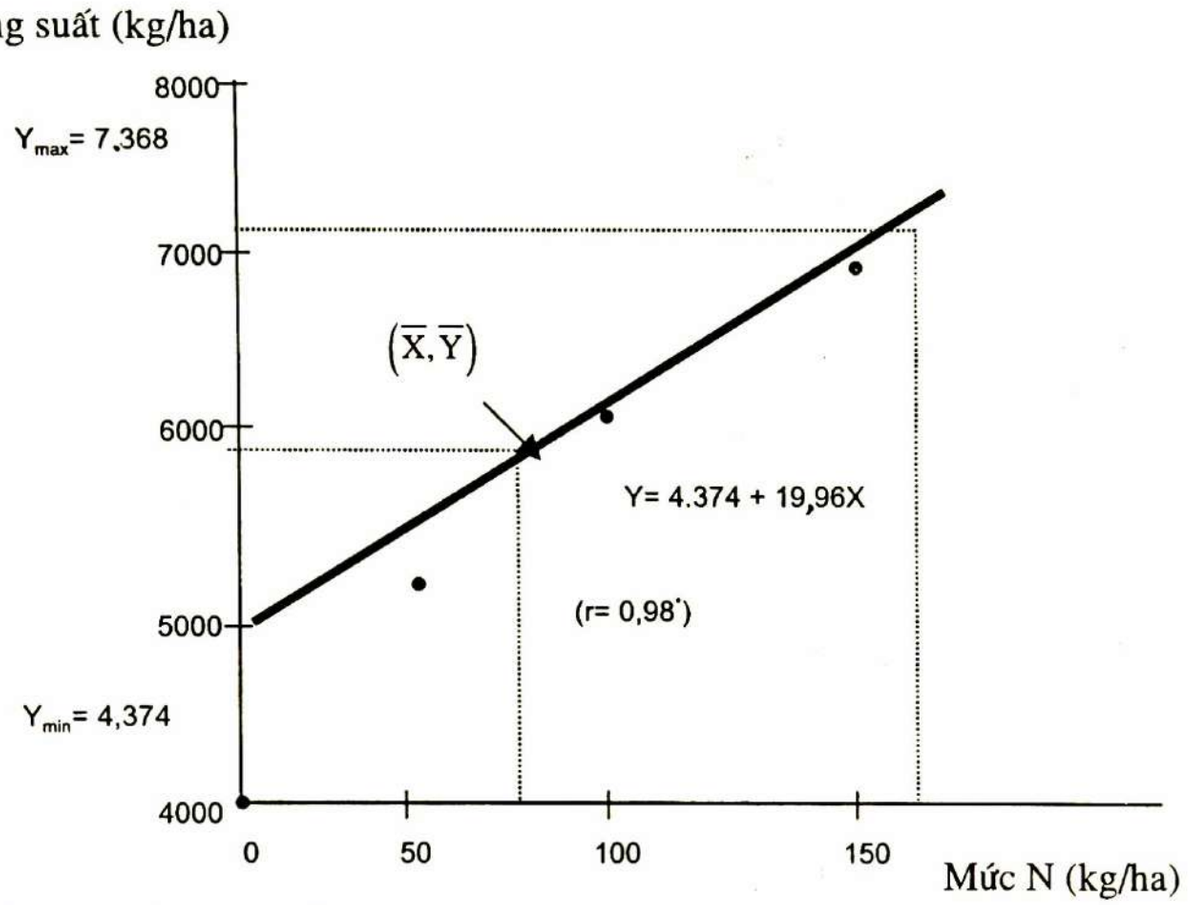
$$Y_{min} = a + bX_{min};\ Y_{max} = a + bX_{max}$$

Với X_{min}= 0 kg N/ha và X_{max}= 150 kgN/ha. Y ước lượng được tính như sau:

$$Y_{min} = 4.374 + 19,96\ (0) = 4.374 \text{ kg}$$

$$Y_{max} = 4.374 + 19,96\ (150) = 7.368 \text{ kg/ha}$$

Xác định 2 toạ độ $M_1(X_{min}, Y_{min})$ và $M_2(X_{max}, Y_{max})$ và kẻ đường thẳng qua 2 toạ độ trên.



Đồ thị 5.2: Đường thẳng ước lượng quan hệ giữa năng suất Y và mức đạm X

Ghi chú: Dấu • đánh dấu những điểm quan sát được, đó chính là các điểm có toạ độ là các giá trị của x và y trong bảng 5.2.

Đồ thị phải đạt yêu cầu sau:

- Đường thẳng phải đi qua 2 điểm X_{min} và X_{max}
- Đường thẳng phải đi qua 2 điểm $(\overline{X}, \overline{Y})$ trong đó $\overline{X}$ và $\overline{Y}$ là giá trị trung bình của X và Y
- Độ dốc của đường thẳng là "b"
- Nếu kéo dài đường thẳng phải gặp trục Y ở giá trị của "a".

Bước 4: Kiểm định ý nghĩa của hệ số β để xem hàm có ý nghĩa hay không.

Tính bình phương của sai khác $s^2_{y.x}$

$$s^2_{y.x} = \frac{\sum y^2 - \frac{(\sum xy)^2}{\sum x^2}}{n-2}$$

Với các giá trị đã tính ở bảng 5.1. ta có

$$s^2_{y.x} = \frac{5.136.864 - \frac{(249.475)^2}{12.500}}{4-2} = 78.921$$

Tính $t_{b\text{-tính}}$: $t_{b\text{-tính}} = \frac{b}{\sqrt{\frac{s^2_{y.x}}{\sum x^2}}} = \frac{19,96}{\sqrt{\frac{78.921}{12.500}}} = 7,94$

So sánh $t_{b\text{-tính}}$ với $t_{b\text{-bảng}}$: tra từ bảng C với độ tự do = (n - 2) = 4-2 = 2; Trong đó n là số cặp so sánh; 2 là số biến ở mô hình tuyến tính một biến.

Nếu $t_{b\text{-tính}} > t_{b\text{-bảng}}$ có nghĩa là hệ số ước lượng β # 0 có ý nghĩa, hay có nghĩa là Y có quan hệ tuyến tính với X. Giá trị $t_{b\text{-bảng}}$ ở mức 5% hay mức 1% chỉ mức độ có ý nghĩa.

Trong ví dụ này, $t_{b\text{-bảng}}$ ở mức 5% = 4,303 và $t_{b\text{-bảng}}$ ở mức 1% = 9,925.

Vậy $t_{b\,(\alpha=01)} > t_{b\text{-tính}} > t_{b\,(\alpha=05)}$

Có nghĩa là với mức đạm bón trong khoảng từ 0 kg/ha đến 150 kg/ha quan hệ tuyến tính giữa năng suất lúa và mức bón đạm tin cậy ở mức 95 %.

Bước 5:

Tính toán khoảng tin cậy (100-α)% (Confidence Interval - CI) cho β.

$$CI = b \pm t_\alpha \sqrt{\frac{s^2_{y.x}}{\sum x^2}}$$

t_α là $t_{bảng}$ tra ở bảng C với độ tự do (n - 2) ở mức có ý nghĩa α.

Trong ví dụ này:

$$CI\ (95\%) = b \pm t_{.05}\sqrt{\frac{s^2_{y.x}}{\sum x^2}} = 19,96 \pm 4.303\sqrt{\frac{78.921}{12.500}} = 19,96 \pm 10,81 = (9,15 \div 30,77)$$

Có nghĩa 95% khả năng là cứ bón thêm mỗi một kg đạm/ha (trong khoảng 0 kg đến 150 kg đạm) năng suất lúa sẽ tăng từ 9,15 đến 30,77 kg/ha.

5.2.3. Phân tích hệ số tương quan tuyến tính đơn giản (r) - *linear correlation coefficient*

Phân tích hệ số tương quan là ước lượng hệ số tương quan r và kiểm định hệ số đó có ý nghĩa hay không. Hệ số tương quan r đo mức độ quan hệ của 2 biến X và Y trong quan hệ tuyến tính.

$|r| > 0,8$ tương quan mạnh

$|r| = 0,4 - 0,8$ tương quan trung bình

$|r| < 0,4$ tương quan yếu

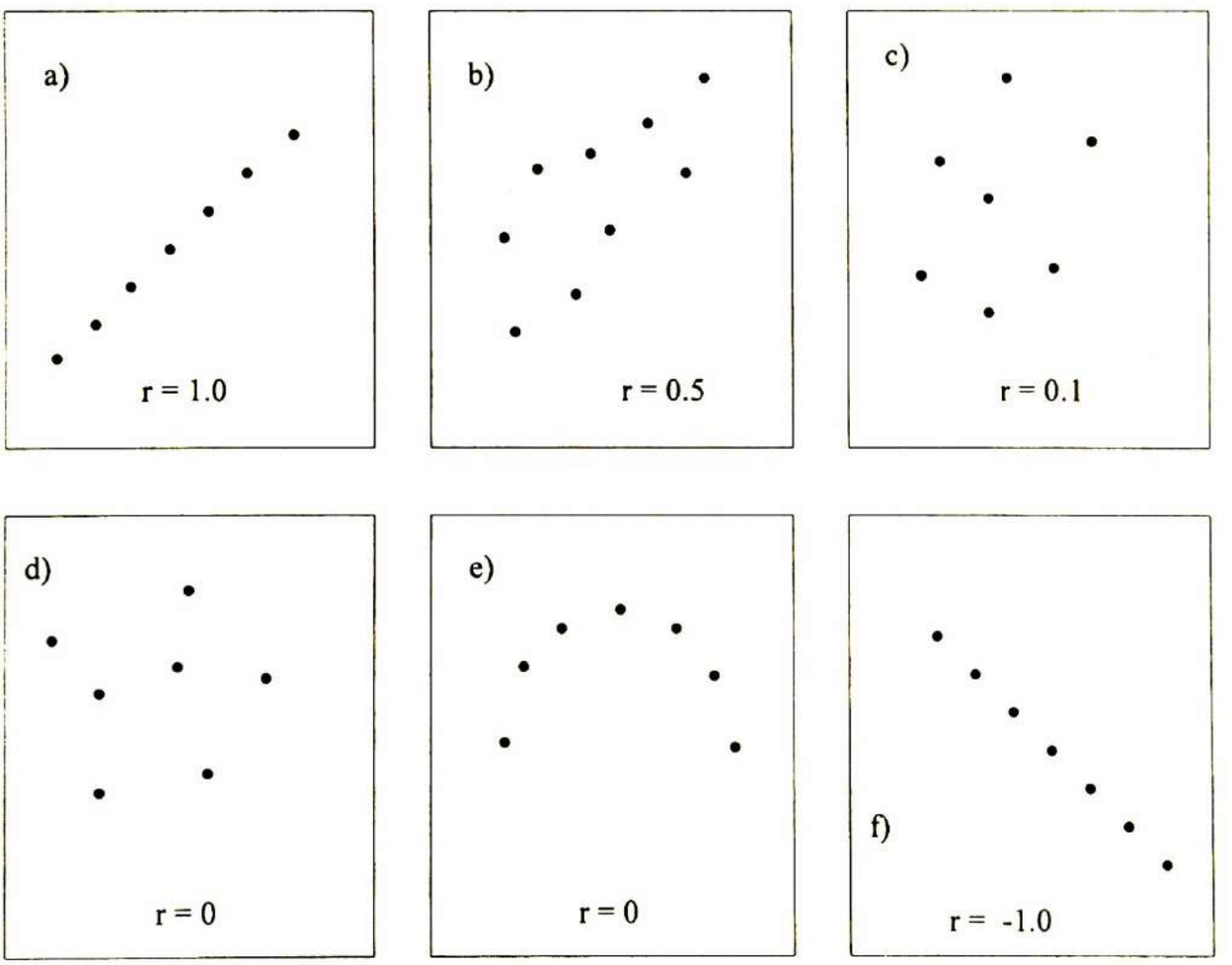
Giá trị tuyệt đối của r càng lớn thì biến Y có quan hệ càng chặt chẽ với biến X. Giá trị của hệ số r nằm trong khoảng -1 đến +1

$0 < r \leq 1$: gọi là tương quan tuyến tính thuận (X tăng, Y tăng)

$-1 < r \leq 0$: gọi là tương quan tuyến tính nghịch (X tăng, Y giảm)

Giá trị của r còn cho biết bao nhiêu % của sự sai khác của một biến là do quan hệ tuyến tính với biến khác. Ví dụ với giá trị r = 0,98, có thể hiểu 96 % (r^2. 100 = $(0,98)^2$.100 = 96 %) của sự sai khác của biến Y được giải thích bằng tương quan tuyến tính với biến X.

Chú ý trong trường hợp tính toán hệ số tương quan tuyến tính (r rất nhỏ), điều đó không có nghĩa là Y không có quan hệ với X, mà chỉ kết luận là Y không có quan hệ tuyến tính với X. Ví dụ nếu như trong 2 biến có quan hệ bậc 2 thì khi theo quan hệ tuyến tính r sẽ rất nhỏ. Đó là lý do khi dùng excel xác định quan hệ giữa các biến ta sử dụng r để quyết định chọn hàm nào phù hợp.



Đồ thị 5.3. Đồ thị biểu diễn giá trị của r với một số trường hợp có quan hệ khác nhau giữa 2 biến X và Y. Dấu • là các giá trị quan sát thực tế.

Các bước ước lượng và kiểm tra hệ số tương quan giữa X và Y của tương quan tuyến tính:

Bước 1: Tính giá trị $\overline{X}$ và $\overline{Y}$ và $\sum\overline{X^2}$ và $\sum\overline{Y^2}$ và $\sum xy$ như bước 1 phần 5.2

Bước 2: Tính hệ số tương quan hồi quy tuyến tính đơn biến (r):

$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Với ví dụ có kết quả ở bảng 5.1.

$$r = \frac{249.475}{\sqrt{(12.500)(5.136.864)}} = 0,985$$

Bước 3: Kiểm định hệ số r, để xác định r có ý nghĩa hay không

So sánh giá trị tuyệt đối của $r_{tính}$ với giá trị $r_{bảng}$ trong bảng H (phần phụ lục) với độ tự do (n = 2) ở mức 5 % và 1%. Ta có:

$$r_{bảng\ .05} = 0,950;\ r_{bảng01} = 0,990$$

Như vậy: $r_{bảng.05} < r_{tính} < r_{bảng\ .01}$

Kết luận: Hệ số tương quan này có ý nghĩa ở mức xác suất 95%.

Giá trị r = 0,985 chỉ ra rằng $r^2 = (100)(0,985)^2 = 97\%$ sai khác trong năng suất trung bình của lúa được giải thích do quan hệ tuyến tính của mức bón đạm khác nhau. Hệ số r có giá trị cao (0,98) cũng chỉ ra đường tuyến tính ước lượng rất gần với các giá trị quan sát thực tế. Điều đó được thể hiện trên đồ thị, trong khoảng bón phân đạm từ 0-150 kg/ha, đường tuyến tính ước lượng gần như trùng với số liệu về năng suất quan sát được.

5.3. TƯƠNG QUAN TUYẾN TÍNH ĐA BIẾN

Phần trên ta mới xét tới quan hệ tương hỗ giữa 2 biến. Nhưng trong tự nhiên cũng như trong nông nghiệp một yếu tố phụ thuộc bị ảnh hưởng bởi nhiều yếu tố độc lập với nhau. Ví dụ năng suất lúa có quan hệ phụ thuộc vào tổng nhánh đẻ hữu hiệu và tỉ lệ hạt chắc/bông...

Một mô hình toán biểu thị tương quan tuyến tính với số biến độc lập từ 2 trở lên được gọi là tương quan tuyến tính đa biến. Trong dạng quan hệ này được giả sử là tất cả các biến độc lập có quan hệ tuyến tính với biến phụ thuộc và độc lập với nhau.

Mô hình hàm tổng quát là:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k$$

Các số liệu quan sát của các biến độc lập và biến phụ thuộc phải được đo đếm cùng một thí nghiệm. Số liệu để áp dụng cho phân tích này được sắp xếp theo ma trận như sau:

Bảng 5.3. Cách ghi số liệu cho phân tích tương quan tuyến tính đa biến

Số quan sát	Giá trị đo đếm				
	Y	X_1	X_2	....	X_k
1	Y_1	X_{11}	X_{21}	....	X_{k1}
2	Y_2	X_{12}	X_{22}	....	X_{k2}
....	....	....	....	....	....
N	Y_n	X_{1n}	X_{2n}	....	X_{kn}

Ví dụ: Phân tích quan hệ tuyến tính giữa năng suất (Y) với chiều cao cây (X_1) và số nhánh đẻ của lúa (X_2) (Bảng 5.2).

Hàm hồi quy có dạng: $Y = a + b_1X_1 + b_2X_2$

Trước đây, tính toán phân tích tương quan tuyến tính đa biến là công việc rất phức tạp và dễ nhầm lẫn. Hiện nay, với sự trợ giúp của máy tính điện tử, công việc trở nên đơn giản. Có rất nhiều phần mềm thống kê giúp cho phân tích như SX3, STAT, SPSS, , SAS...

Ví dụ cách ước lượng các hệ số trong phân tích tương quan tuyến tính đa biến cho thí nghiệm giống lúa có kết quả ở bảng 5.4.

Bước 1: Tính $\sum x_1^2$; $\sum x_2^2$; $\sum x_1y$; $\sum x_2y$; $\sum x_1x_2$ và $\sum y^2$

Bảng 5.4: Phân tích mối quan hệ giữa năng suất lúa (Y) với chiều cao cây (X_1) và số nhánh đẻ của lúa (X_2)

Giống lúa	Năng suất (kg/ha) (Y)	Chiều cao cây (cm) (X_1)	Nhánh đẻ/khóm (X_2)
1	5.755	110,5	14,5
2	5.939	105,4	16,0
3	6.010	118,1	14,6
4	6.545	104,5	18,2
5	6.730	93,6	15,4
6	6.750	84,1	17,6
7	6.899	77,8	17,9
8	7.862	75,6	19,4
Trung bình	6.561	96,2	16,7
$\sum x_1^2$ = 1.753.72	$\sum x_2^2$ = 23,22	$\sum x_1y$ = -65.194	
$\sum x_2y$ = 7.210	$\sum x_1x_2$ = 7.210	$\sum y^2$ = 3.211.504	

Bước 2: Tính hệ số "b" trong trường hợp 2 biến theo công thức sau

$$b_1\sum x_1^2 + b_2\sum x_1x_2 = \sum x_1y$$

$$b_1\sum x_1x_2 + b_2\sum x_2^2 = \sum x_2y$$

Vậy: $$b_1 = \frac{(\sum x_2^2)(\sum x_1 y) - (\sum x_1 x_2)(\sum x_2 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$= \frac{(23,22)(-65.194) - (-156,65)(7.210)}{(1.753,72)(23,22) - (-156,65)^2} = -23,75$$

$$b_2 = \frac{(\sum x_1^2)(\sum x_2 y) - (\sum x_1 x_2)(\sum x_1 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$= \frac{(1.753,72)(7.210) - (-156,65)(-65.194)}{(1.753,72)(23,22) - (-156,65)^2} = 150,27$$

Bước 3: Tính hệ số a:

$$a = \overline{Y} - b_1\overline{X_1} - b_2\overline{X_2}$$

$$= 6.561 - (-23,75)(96,2) - (150,27)(16,7) = 6.336$$

Vậy ta có mô hình tuyến tính 2 biến được ước lượng là:

$Y = 6.336 - 23,75X_1 + 150,27X_2$

Bước 4: Tính hệ số xác định R^2 (*coeffiecient of determination*) và hệ số tương quan đa biến R *(multiple correlation coefficient)*:

R^2 biểu thị mức độ đóng góp của hàm tương quan với k biến độc lập lên biến phụ thuộc Y; R^2 thường được biểu thị bằng %. $R = \sqrt{R^2}$ được gọi là hệ số tương quan đa biến.

Tổng bình phương gây bởi hàm được ước lượng (SSR)

$$SSR = \sum_{i=1}^{k}(b_1)(\sum x_i y)$$

$$= (-23,75)(-65.149)+(150,27)(7.210) = 2.631.804$$

Tổng bình phương do sai số (SSE)

$$SSE = \sum y^2 - SSR$$

$$= 3.211.504 - 2.631.804 = 579.700$$

Hệ số xác định *(Coeffiecient of Determination* - R^2):

$$R^2 = \frac{SSR}{\sum y^2} = \frac{2.631.804}{3.211.504} = 0,82$$

Vậy trong thí nghiệm này, 82 % biến động của biến trong năng suất (Y) được giải thích bởi hàm quan hệ tuyến tính với chiều cao cây và số nhánh đẻ/khóm.

Bước 5: Kiểm tra độ tin cậy của hệ số R^2

Tính giá trị F: $F_{tính} = \frac{SSR/k}{SSE/(n-k-1)} = \frac{2.631.804/2}{579.700/(8-2-1)} = 11,35$

So sánh $F_{tính}$ với $F_{bảng}$ trong bảng E với độ tự do $f_1=k; f_2=(n-k-1)$.

Nếu: $F_{tính} > F_{bảng}$ kết luận R^2 có ý nghĩa, nghĩa là $R^2 \neq 0$,

Tra bảng với $f_1=2; f_2=5$, ta được $R^2_{05} = 5,79$; và $R^2_{01} = 13,27$

Vậy: $R^2_{05} < R^2_{tính} < R^2_{01}$. Có nghĩa là hàm ước lượng trên có ý nghĩa và tin cậy ở mức 95%, hay ảnh hưởng tuyến tính phối hợp của chiều cao cây và số nhánh đẻ/khóm có ý nghĩa đến năng suất lúa.

5.4. TƯƠNG QUAN PHI TUYẾN TÍNH ĐƠN BIẾN

Tương quan phi tuyến tính là tương quan không tuyến tính. Để phân tích trước tiên phải tuyến tính hoá hàm phi tuyến tính bằng cách chuyển đổi chúng về dạng tuyến tính.

Ví dụ: Chuyển hàm $Y = ab^X$ (*) sang dạng tuyến tính như sau

Nếu ta logarit hoá 2 vế của phương trình ta có:

$Log(Y) = Log(a) + X.Log(b)$

Đặt $Log(Y) = Y'$; $Log(a) = a'$; $Log(b) = b'$. Ta có: $Y' = a' + b'X$

Như vậy ta đã đưa hàm phi tuyến tính về dạng tuyến tính và phân tích theo dạng tuyến tính. Sau ước lượng được các tham số ta phải chuyển ngược lại để có hàm dạng (*) như trên.

Ví dụ ứng dụng với hàm Exponential: Hàm sinh trưởng thực vật sau mọc:

$Y = ae^{bX}$

Đặt $Y' = \ln(Y)$ và $a' = \ln(a)$. Logarit (cơ số e) hai vế của phương trình ta có.

$Y' = a' + bX$

Sau đây là ví dụ sử dụng tuyến tính hoá để phân tính quan hệ phi tuyến tính.

Bảng 5.5: Quan hệ phi tuyến tính giữa tỉ lệ truyền ánh sáng trong quần thể lúa và hệ số diện tích lá - giống lúa IR8

Mẫu đo đếm	Tỉ lệ ánh sáng (Y)	Hệ số diện tích lá (X)	Y' = Ln(Y)
1	75,0	0,50	4,31749
2	72,0	0,60	4,2766
3	42,0	1,80	3,73767
4	29,0	2,05	3,36730
5	27,0	2,80	3,29584
6	10,0	5,45	2,30259
7	9,0	5,60	2,19722
8	5,0	7,20	1,60944
9	2,0	8,75	0,69315
10	2,0	9,60	0,69315
11	1,0	10,40	0,00000
12	0,9	12,00	-0,10536
Trung bình	$\overline{Y}$ =22,9	$\overline{X}$ =5,60	$\overline{Y'}$ =2,19876
$\sum y'^2$ = 28,77936		$\sum x_2$ =175,40500	$\sum xy'$ = -70,76124

Nguồn: Gomez, 1984.

Hàm quan hệ của chúng có dạng:

$$Y = \alpha e^{\beta X}$$

Bước 1

Chuyển đổi tuyến tính hoá hàm $Y = \alpha e^{\beta X}$ bằng ln của 2 vế ta có $\ln Y = \ln \alpha e^{\beta X}$

$$\ln \alpha e^{\beta X} = \ln \alpha + \ln e^{\beta X} = \ln \alpha + \beta X \ln e = \ln \alpha + \beta X$$

Ta có $\ln Y = \ln \alpha + \beta X$ đặt $Y' = Ln(Y)$ và $\alpha' = Ln(\alpha)$, ta có hàm:

$$Y' = \alpha' + \beta X$$

Hàm ước lượng tương ứng là: $Y' = a' + bX$

Bước 2

Tính các giá trị chuyển đổi cho tất cả các giá trị quan sát của mỗi biến đã chuyển đổi ở bước 1. Như tính các $Y' = \ln Y$. Kết quả trình bày ở bảng 5.5.

Bước 3

Áp dụng thủ tục cho tương quan tuyến tính một biến ta ước lượng tham số α' (tham số a') và β (tham số b).

$$b = \frac{\sum xy'}{\sum x^2} = \frac{-70,76124}{175,40500} = -0,40342$$

$$a' = \overline{Y} - b\overline{X}$$

$$a' = 2,19876 - (-0,40342)(5,60) = 4,45791$$

Bước 4

Tiếp theo ta tính a bằng antilog cơ số a':

a = antilog của a'= antilog 4,45791 = 86,31

Thay vào ta có hàm ước lượng quan hệ giữa tỉ lệ truyền ánh sáng (Y) trong quần thể lúa và hệ số diện tích lá (X) - giống lúa IR8:

$$Y = 86,31e^{-0,40342X}$$

Ví dụ 2: Hàm quan hệ năng suất thu hoạch của lúa và mức bón đạm. Hàm quan hệ là hàm Polynomials dưới đây:

$$Y = a + b_1X_1 + b_2X^2$$

Thủ tục tuyến tính hoá cũng tương tự như trên.

Đặt $Z_1 = X$ và $Z_2 = X^2$ ta có: $Y = a + b_1Z_1 + b_2Z_2$

5.5. TƯƠNG QUAN PHI TUYẾN TÍNH ĐA BIẾN

Trường hợp quan hệ phi tuyến tính giữa biến phụ thuộc với nhiều biến độc lập ta gọi là *Tương quan phi tuyến tính đa biến*. Loại hàm này có nhiều dạng khác nhau.

Hàm có ít nhất 2 biến độc lập quan hệ phi tuyến tính với biến phụ thuộc:

$$Y = \alpha + \beta_1X_1 + \beta_2X_1^2 + \beta_3X_2 + \beta_4X_2^2$$

Hàm có hai biến độc lập tương tác với nhau:

$$Y = \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_1X_2$$

Khi phân tích cũng tương tự như trường hợp phi tuyến tính đơn biến, trước khi phân tích ta cũng phải tuyến tính hoá hàm phi tuyến tính bằng cách chuyển đổi chúng về dạng tuyến tính.

Ví dụ: $Y = \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_1X_2$

Đặt $X_3 = X_1X_2$ ta có: $Y = \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_3$

5.6. NHỮNG SAI LẦM THƯỜNG MẮC KHI PHÂN TÍCH TƯƠNG QUAN

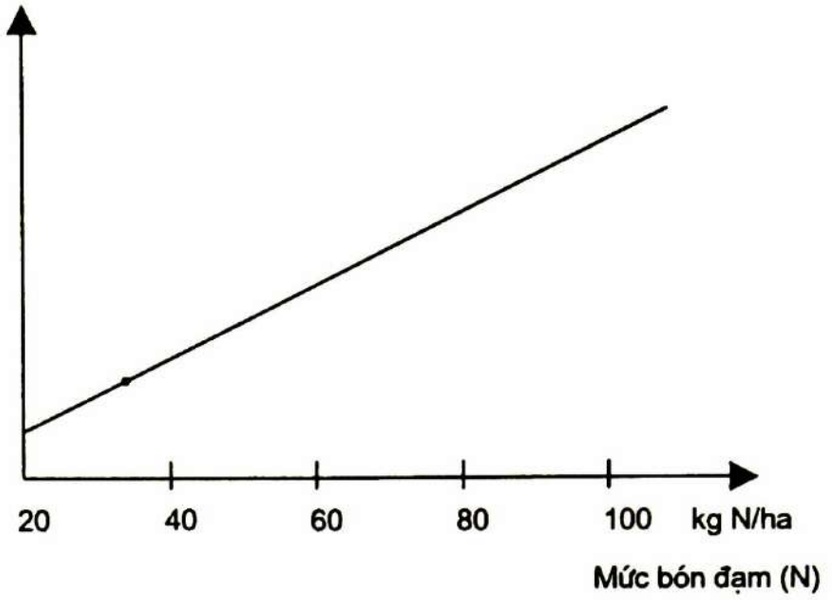
Phân tích tương quan là một công cụ rất hữu ích để tìm hiểu mối quan hệ giữa các biến sinh học, môi trường... với nhau. Do đó nó được áp dụng ngày càng rộng rãi trong nghiên cứu nông nghiệp. Tuy nhiên, nhiều người đã mắc phải sai lầm trong phân tích cũng như hiểu sai kết quả nghiên cứu. Sau đây là một số sai lầm thường mắc phải.

Phân tích tương quan không theo mục đích hay không có quan hệ đúng giữa các biến.

Phân tích tương quan có thể làm và cho kết quả với bất cứ tập hợp số liệu nào và từ bất kỳ nguồn nào. Nhưng, bản chất của phân tích tương quan là sử dụng công cụ thống kê toán học để ước lượng một mô hình toán học biểu thị mối quan hệ khách quan giữa các biến của một sự kiện hay hiện tượng. Nếu nhà khoa học không hiểu rõ bản chất của mối quan hệ này mà đem các tập hợp số liệu ra phân tích, thì kết quả ước lượng sẽ sai. Thực chất có thể không có mối quan hệ giữa các biến nhưng ta lại kết luận là chúng có quan hệ với nhau. Ví dụ, nhiệt độ ấm thì sâu bệnh thường phát triển mạnh. Tuy nhiên nhiệt độ ấm cũng là điều kiện cho cây trồng phát triển tốt. Nếu tính tương quan trong trường hợp này có thể cho kết quả trái ngược là sâu bệnh càng nhiều thì cây trồng càng sinh trưởng mạnh. Hay nói cách khác, kết luận từ việc sử dụng phân tích tương quan trong nông nghiệp không phải thuần tuý từ những thuật toán mà đòi hỏi người xử lý phải vững về chuyên môn nông nghiệp để có thể chọn biến (độc lập và phụ thuộc) phù hợp và lý giải kết quả của phân tích tương quan.

Kết luận mở rộng từ hàm hồi quy.

Ví dụ trong thí nghiệm bón đạm. Công thức bắt đầu bón từ 20 kg/ha và kết thúc là 100 kg/ha. Nếu dùng hàm tương quan ước lượng từ thí nghiệm này để suy luận năng suất lúa trong trường hợp bón 120 kg/ha là sai. Bởi hàm này được ước lượng trong khoảng mức bón từ 20-100 kg/ha.



Sử dụng số liệu từ các lần nhắc lại riêng rẽ.

Nếu làm như vậy là ta đưa sai số thí nghiệm vào đánh giá mối quan hệ giữa các biến, kết quả sẽ thiếu chính xác.

Hiểu sai về phân tích tương quan.

Có trường hợp số liệu phân bố theo nhóm nếu ta kết luận là có quan hệ, có thể mắc sai lầm mặc dầu hệ số tương quan (r) là lớn v.v...

Chương 6

ỨNG DỤNG TIN HỌC TRONG XỬ LÝ VÀ TRÌNH BÀY KẾT QUẢ

6.1. GIỚI THIỆU VỀ PHẦN MỀM SAS

6.1.1. Ưu điểm của phần mềm SAS

Hiện nay trên thế giới có rất nhiều phần mềm khác nhau được sử dụng để xử lý thống kê số liệu trong nông nghiệp như: SAS, IRRISTAT, MSTATC, SPSS... Trong đó, phần mềm SAS (statistical analysis system) là phần mềm chuyên dùng cho phân tích xử lý thống kê số liệu do SAS Institute Inc., Cary, NC, USA xây dựng. Đây là một trong những phần mềm xử lý số liệu được sử dụng rộng rãi nhất hiện nay. Phần mềm SAS có các ưu điểm nổi bật sau:

SAS có khả năng liên kết dữ liệu với các phần mềm quản lý số liệu mạnh và thông dụng như EXCEL ACCESS... .

Đây là một tiện ích rất có ý nghĩa của phần mềm SAS bởi trên thực tế các bộ số liệu trước khi đưa vào xử lý thường được nhập và lưu trữ bằng chương trình quản lý số liệu thông dụng là EXCEL. Do vậy sự liên kết giữa SAS và Excel giúp cho việc quản lý số liệu và xử lý số liệu rất tiện lợi và hiệu quả.

SAS là ngôn ngữ lập trình bao gồm các bộ lệnh mạnh trong xử lý thống kê.

Ta có thể dễ dàng lập ra và lưu trữ những chương trình chuẩn xử lý thống kê số liệu cho từng kiểu bố trí thí nghiệm. Như vậy, khi muốn xử lý thống kê cho bất kỳ một thí nghiệm nào chỉ cần gọi chương trình tương ứng của kiểu đó và sửa lại tên tệp, biến cho phù hợp với thí nghiệm rồi chạy chương trình mà không phải thực hiện một loạt các thao tác xử lý như trong các phần mềm khác. Chính vì vậy mà SAS có thể cho phép cùng một lúc xử lý hàng loạt các thí nghiệm một cách nhanh chóng.

Kết quả phân tích số liệu trong SAS dễ dàng được copy ra các phần mềm khác, ví dụ như phần mềm soạn thảo Winword.

SAS cho phép chỉ định từng loại sai số cho từng nguồn biến động.

Với những thí nghiệm phức tạp có nhiều loại sai số thí nghiệm yêu cầu việc kiểm tra ảnh hưởng của các yếu tố nên cần biết loại sai số khác nhau. SAS rất phù hợp với các yêu cầu trên.

SAS đưa ra các giá trị CV% và LSD của thí nghiệm.

6.1.1.3. Một số chức năng và lệnh chính thường sử dụng trong SAS

FILE

Open: Mở một tệp có sẵn.

Save: Cất một chương trình hay một kết quả xử lý ...

Save as: Cất tệp với một tên khác.

Print: in ra máy.

Print setup: Định dạng trang in.

EDIT

Undo: Huỷ bỏ thao tác vừa thực hiện.

Cut: Cắt đi một phần được chọn.

Copy: Copy phần được lựa chọn.

Paste: Dán phần được copy.

Clear selected text: Xoá phần được lựa chọn.

Clear text: Xoá sạch màn hình của cửa sổ hiện hành để làm những công việc tiếp khác.

OPTIONS

Fonts: Chọn kiểu chữ theo ý muốn.

Edit tools: Bố trí sửa đổi lại các thanh công cụ.

LOCALS

Submit: Chạy chương trình khi nháy chuột vào đây.

Recall text: Gọi lại chương trình để sửa chữa nếu cần thiết.

GLOBAL

Program Editor: Cửa sổ soạn thảo chương trình.

Log: Cửa sổ hiển thị các lỗi của chương trình.

Output: Cửa sổ hiển thị kết quả của chương trình.

Chú ý: Các cửa sổ này cũng có thể được mở từ thực đơn dọc WINDOW.

WINDOW

Minimize all windows: Thu nhỏ tất cả các cửa sổ.

Program Editor: Soạn thảo các chương trình.

Log: Hiển thị các lỗi của chương trình.

Output: Hiển thị kết quả của chương trình.

HELP: Các trợ giúp trong SAS.

6.1.1.4. Quy định viết các câu lệnh trong SAS

- Câu lệnh của SAS bắt đầu bằng một từ khoá và kết thúc bằng một dấu ';'.
- Câu lệnh có cấu trúc tự do, có thể bằng chữ in hoa hay chữ thường.
- Các câu lệnh có thể kéo dài nhiều dòng hoặc vài câu lệnh cũng có thể viết trên một dòng. Cần đảm bảo giữa các câu lệnh được cách nhau bởi một dấu chấm phẩy (;).
- Các trường phải cách nhau ít nhất một dấu cách.

6.1.1.5. Cấu trúc dữ liệu trong SAS

- Dữ liệu trong SAS được tổ chức theo bảng ngang trong đó mỗi cột của bảng là một biến (hay một trường).

Ví dụ một số biến:

Season: Mùa.

Trt: Công thức.

Rep: Nhắc lại.

Yield: Năng suất.

Bảng 6.1. Cấu trúc dữ liệu trong SAS

Các quan trắc	Các biến			
	Season	Trt	Rep	Yield
Quan trắc 1	Spring	1	1	5.381
Quan trắc 2	Spring	1	2	5.380
Quan trắc 3	Spring	1	3	5.682
....				
Quan trắc n	Summer	1	2	5.392

- Mỗi biến phải có một tên biến.

- Có hai loại biến:

Biến ký tự: Ví dụ biến mùa (season), tên công thức (trtname).

Biến kí tự cần được chỉ định cho SAS bằng dấu đô la ($).

Ví dụ season $.

Biến số: Khối lượng, %N...

Chú ý: Các biến có độ dài từ 1-200 kí tự.

- Các hàng trong bảng được gọi là các quan trắc (observation) hay còn gọi là bản ghi (record). Số hàng có thể nhiều bao nhiêu tuỳ ý (không có giới hạn). Số quan trắc là ứng với mỗi lần lấy mẫu để khảo sát.

6.1.1.6. Tên của tệp và các biến trong SAS

- Qui định đối với đặt tên trong SAS.

- Tên dài tối đa 8 kí tự.

- Tên không được có dấu cách.

- Bắt đầu bằng một kí tự từ A đến Z.

Tiếp tục bằng các số, kí tự hoặc dấu gạch dưới.

Ví dụ:	**Tên đúng**	**Tên sai**	**Lỗi**
	Expt1	1exprt	bắt đầu bằng số
	Pheight	plantheight	Quá 8 kí tự
	G_yield	Grain yield	Có dấu cách

6.1.2. Hướng dẫn sử dụng SAS

6.1.2.1. Khởi động SAS

Để khởi động SAS có thể dùng một trong những cách sau:

Cách1: Nháy đúp chuột vào biểu tượng SAS trên màn hình.

Cách 2: Chọn Start/ Program/the SAS system/the SAS for window.

6.1.1.2. Các cửa sổ làm việc của SAS

Màn hình SAS có 3 cửa sổ làm việc chính là program editor window, log WINDOW và out put WINDOW. Ba cửa sổ này có mối liên kết chặt chẽ với nhau ví dụ khi chương trình được chạy trong cửa sổ program editor thì kết quả của nó được thể hiện ở cửa sổ output. Các dòng lệnh và các thông điệp của cửa sổ program editor cũng được ghi ở trong cửa sổ log. Chức năng cụ thể của từng cửa sổ như sau:

Program editor window: Là cửa sổ soạn thảo chương trình và sửa chương trình. Chương trình được nhập vào cửa sổ này từ bàn phím hoặc được gọi lại khi chương trình đã được lưu hay có thể copy từ bất cứ một tệp bên ngoài nào.

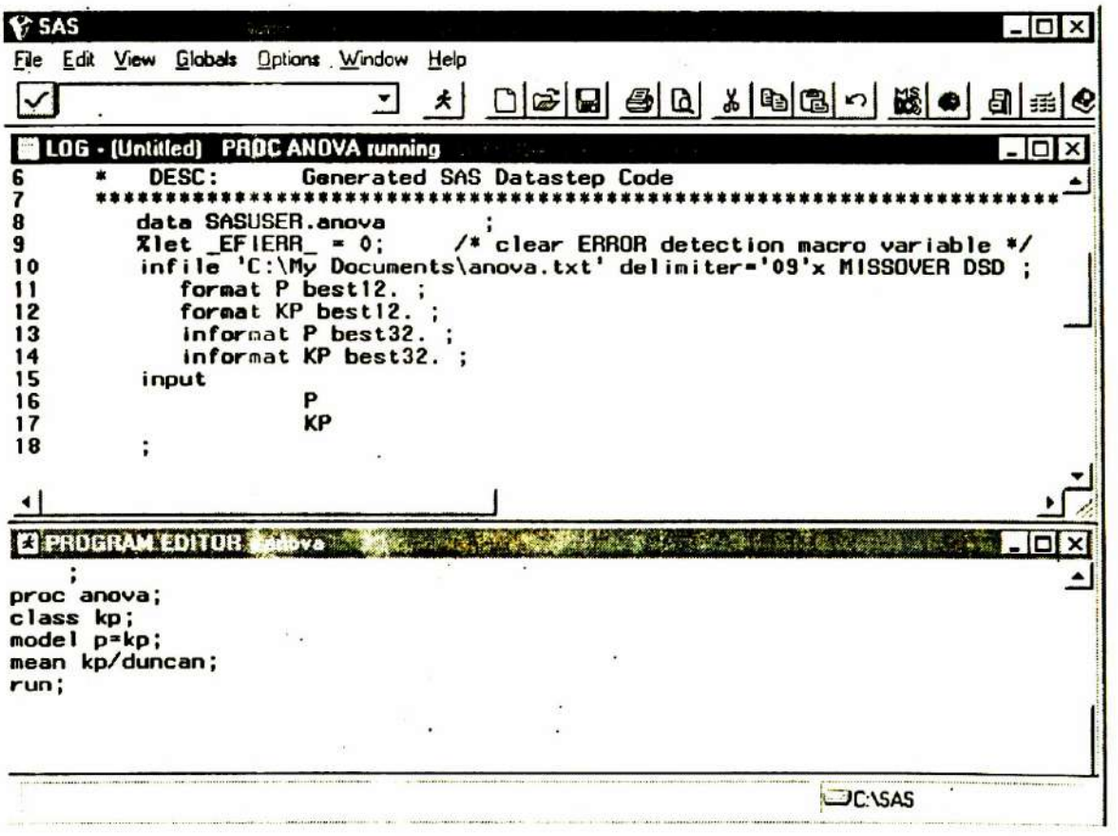
Log window: Là cửa sổ thông báo các lỗi của chương trình . Vì vậy khi chương trình mắc lỗi phải kích hoạt cửa sổ này để xem các lỗi của chương trình.

Out put window: Là cửa sổ thể hiện kết quả chạy chương trình. Khi chương trình chạy không mắc lỗi, kết quả sẽ được đưa ra ở cửa sổ này.

Ta có thể đóng mở, phóng to hay thu nhỏ các cửa sổ tuỳ ý thông qua hộp CONTROL BOX (hộp điều khiển) nằm ở góc trái trên của mỗi cửa sổ.

Chú ý: Cửa sổ và các lệnh được in mờ là tạm thời không thực hiện được.

Dưới đây là màn hình của 2 cửa sổ của SAS



Màn hình SAS với 2 cửa sổ LOG và PROGRAM EDITOR

6.1.3. Cách nhập số liệu vào SAS

SAS cho phép nhập số liệu trực tiếp vào SAS bằng câu lệnh CARDS hoặc đọc từ một file số liệu bên ngoài vào. Tuy nhiên tiện lợi hơn cả là hãy lưu trữ file số liệu trên EXCEL Ở dạng **prn file** (formatted text [space delimited]) và truy cập file đó từ phần mềm SAS. Mặc dù dùng lệnh CARDS trong SAS cũng cho phép nhập số liệu trực tiếp trên SAS nhưng điều này sẽ không thực tế đối với những bộ số liệu lớn.

Để truy cập được file số liệu từ phần mềm SAS cần nhập số liệu trên EXCEL như sau: Nhập số liệu theo bảng trong đó các biến được ghi theo cột, mỗi biến phải có một tên biến. Các hàng trong bảng được gọi là các quan trắc (tương ứng với một đơn vị thí nghiệm).

Ví dụ: Nhập số liệu cho một thí nghiệm kiểu ô vuông la tinh (phần 3.2.4). Thí nghiệm so sánh năng suất của 4 giống lúa A, B, C, D. Kết quả thí nghiệm như sau:

Bảng 6.2. Kết quả thí nghiệm so sánh giống lúa bố trí theo kiểu ô vuông la tinh

STT hàng	Năng suất hạt (tấn/ha) (Xi)			
	Cột 1	Cột 2	Cột 3	Cột 4
Hàng 1	1,640 (B)	1,210 (D)	1,425 (C)	1,345 (A)
Hàng 2	1,475 (C)	1,185 (A)	1,400 (D)	1,290 (B)
Hàng 3	1,670 (A)	0,710 (C)	1,665 (B)	1,180 (D)
Hàng 4	1,565 (D)	1,290 (B)	1,665 (A)	0,660 (C)

Nguồn: Gomez, 1984

Số liệu nhập vào Excel gồm 4 cột tương ứng với 4 biến : Hàng (row), cột (column); công thức (trt) và năng suất (yield) như sau:

Bảng 6.3. Cách nhập số liệu bảng 6.2 vào excel để chạy SAS

Row	Column	trt	Yield
1	1	B	1.640
1	2	D	1.210
1	3	C	1.425
1	4	A	1.345
2	1	C	1.475
2	2	A	1.185
2	3	D	1.400
2	4	B	1.290
3	1	A	1.670
3	2	C	0.710
3	3	B	1.665
3	4	D	1.180
4	1	D	1.565
4	2	B	1.290
4	3	A	1.655
4	4	C	0.660

Trường hợp biến kí tự quá dài gây khó khăn cho việc nhập số liệu, nên mã hoá chúng ở dạng số. Trong ví dụ này nếu mã hoá các công thức: 1-A; 2-B; 3-C; 4-D thì bảng số liệu trên được nhập như sau:

Bảng 6.4. Cách nhập số liệu bảng 6.2. vào excel để chạy SAS

Row	Column	trt	Yield
1	1	2	1.640
1	2	4	1.210
1	3	3	1.425
1	4	1	1.345
2	1	3	1.475
2	2	1	1.185
2	3	4	1.400
2	4	2	1.290
3	1	1	1.670
3	2	3	0.710
3	3	2	1.665
3	4	4	1.180
4	1	4	1.565
4	2	2	1.290
4	3	1	1.655
4	4	3	0.660

Chú ý: Sau khi nhập xong, số liệu phải lưu (save) ở dạng ***prn file (formatted text [space delimited])*** và phải đóng file trước khi thực hiện xử lý số liệu bằng chương trình SAS.

6.1.4. Một số thuật ngữ chính được sử dụng trong sử dụng SAS

- *Plot:* Ô thí nghiệm.

- *Treatment (trt)*: Công thức thí nghiệm.

- *Experiment unit*: Đơn vị thí nghiệm, là một ô trong khu thí nghiệm.

- *Replication (Rep)*: Lần nhắc lại của các công thức thí nghiệm.

- *Block*: Một khối thí nghiệm bao gồm nhiều ô thí nghiệm. Đối với một số loại thí nghiệm một block chính là một Rep.

- *Experiment material*: (Vật liệu thí nghiệm) là các yếu tố phi thí nghiệm (không phải là yếu tố cần quan trắc trong thí nghiệm).

- *Sample*: Mẫu thu thập cần xử lý thống kê.

- *Variable* (biến) là một cột trong tập tin số liệu nhập. Biến có thể đặt tên tuỳ ý sao cho dễ nhớ và phù hợp với kết quả cần xử lý. Thí dụ biến có thể là công thức, lần lặp lại, năng suất...

- *Source of variation*: Nguồn biến động.

- *Degree of freedom(df)*: Độ tự do của dãy số liệu.

- *Sum of squares (SS)*: Tổng bình phương.

- *Mean square (MS)*: Trung bình bình phương.

- *Error (E)*: Sai số. MSe: Trung bình bình phương của sai số .

- *Standard deviation* (Sd): Độ lệch chuẩn.

- *Standard error_(STD Dev):* Sai số chuẩn.

- *Variance*: Phương sai.

- *Covariance*: Hiệp phương sai.

- *Coefficienct of variance (CV)*: Hệ số biến động.

- *Intercept*: Hệ số tự do (giá trị của biến phụ thuộc khi các biến độc lập có giá trị 0).

- *Probability (Pr):* Xác suất.

- *F-value (F-tính)*: Giá trị tính của hàm phân bố xác suất ứng với một mức ý nghĩa nào đó.

- *Factor*: Yếu tố ảnh hưởng đến thí nghiệm.

- *Main plot factor*: Yếu tố trên ô chính.

- *Subplot factor*: Yếu tố ảnh hưởng trên ô phụ.

- *Vertical factor*: Yếu tố được xét đến trong các lô bố trí theo phương dọc.

- *Horizontal factor*: Yếu tố được xét trong các lô bố trí theo phương ngang.

- *Interaction*: (tương tác): Tác dụng tương hỗ.

- *Vertical strip plot*: Lô sọc đứng (băng dọc).

- *Horizontal strip plot*: Lô sọc ngang (băng ngang).

- *Grand total (G)*: Tổng toàn bộ.

- *Grand mean* ($\overline{X}$): Trung bình chung.

- *Yield* : Năng suất.

- *LSD*: Least Significant Difference: Sai khác nhỏ nhất có ý nghĩa.

6.1.5. Thực hành

6.1.4.1. Kích hoạt cửa sổ thao tác

- Chọn cửa sổ thao tác từ thực đơn WINDOW hoặc GLOBALS.

- Di chuyển chuột từ cửa sổ này sang cửa sổ khác bằng cách di chuột đến cửa sổ cần chọn hoặc:

Ấn F6: Di chuyển chuột đến cửa sổ LOG.

Ấn F5: Di chuyển chuột đến cửa cổ PROGRAM EDITOR.

Ấn F7: Di chuyển chuột đến cửa sổ OUTPUT.

6.1.4.2. Soạn thảo chương trình

6.1.4.3. Chạy chương trình: Chạy chương trình bằng các cách sau:

- Nháy chuột phải/chọn LOCALS/Chọn SUBMIT.
- Nhấn F8.
- Nháy biểu tượng hình người đang chạy trên thanh công cụ.

Nếu chương trình mắc lỗi, các lỗi sẽ được hiển thị màu đỏ trong của sổ LOG.

Nếu chương trình không mắc lỗi, kết quả được hiển thị trong cửa sổ OUTPUT. Di chuyển chuột đến cửa sổ này và dùng các thanh cuốn dọc và ngang hay ấn PAGE UP và PAGE DOWN để xem kết quả.

6.1.4.4. Gọi lại chương trình

Để gọi lại chương trình thực hiện một trong các các cách sau:

Đưa chuột vào cửa sổ PROGAM EDITOR và ấn F4.

Đưa chuột vào của sổ PROGRAM EDITOR và chọn Recall text trong thực đơn LOCALS.

Chú ý: Trước khi chạy lại chương trình, cần vào cửa sổ LOG chọn CLEAR TEXT để xóa những lỗi cũ đã được thông báo trong cửa sổ này. Như vậy cửa sổ sẽ chỉ hiển thị các thông báo của lần chạy mới nhất.

6.2. SỬ DỤNG SAS TRONG PHÂN TÍCH PHƯƠNG SAI (ANOVA)

Phần này hướng dẫn xử lý thống kê số liệu của các thí nghiệm bằng phần mềm SAS. Trong đó, bộ số liệu của thí nghiệm được nhập bằng phần mềm EXCEL và lưu trữ ở dạng prn file.

6.2.1. Các trình thủ tục để phân tích phương sai trong SAS

Có 2 trình thủ tục để phân tích phương sai trong SAS.

1. PROC ANOVA.
2. PROC GLM.

PROC ANOVA: Dùng để phân tích phương sai cho những số liệu cân đối (Không mất số liệu, các công thức thí nghiệm có số lần lặp lại giống nhau).

PROC GLM: Dùng để phân tích phương sai cho những số liệu không cân đối (mất số liệu, số lần nhắc lại của các công thức thí nghiệm khác nhau), phân tích tương quan hồi quy, hay các phép so sánh riêng nào đó (CONTRAST).

Đối với những số liệu cân đối (balanned data), PROC ANOVA và PROC GLM cho kết quả phân tích phương sai giống nhau.

Cấu trúc chung của PROC ANOVA và PROC GLM

- *Khai báo chương trình chính:*

PROC GLM tuỳ chọn;

Tuỳ chọn được khai báo như sau: DATA là các số liệu cần xử lý;

Khi đó SAS sẽ xử lý các bộ số liệu được liệt kê trong tuỳ chọn. Còn nếu không có tuỳ chọn thì SAS xử lý bộ số liệu mới nhất.

- *Câu lệnh:* Một số câu lệnh thường sử dụng với PROC GLM và PROC ANOVA như sau.

CLASS (khai biến)

Câu lệnh CLASS liệt kê các biến sử dụng trong phân tích. Ví dụ các yếu tố, công thức và các khối có trong các nguồn biến động của bảng ANOVA của thí nghiệm, các biến này được phân cách với nhau bởi dấu cách.

Cách viết: Minh họa bằng ví dụ 1: thí nghiệm 2 nhân tố gồm đạm (kí hiệu là N) và giống (kí hiệu là Var) bố trí theo kiểu RCBD (Block) thì các biến được khai báo trong câu lệnh CLASS như sau:

CLASS Block N Var;

Chú ý: Các biến, khối phải được kí hiệu thống nhất cho cả chương trình, nếu không SAS sẽ thông báo lỗi là không tìm thấy biến.

MODEL (ghi Biến phụ thuộc = Biến độc lập/ Tuỳ chọn);

Biến phụ thuộc: là các biến cần được đánh giá sự thay đổi bởi ảnh hưởng của các tuỳ chọn và được cách nhau bởi dấu cách.

Biến độc lập/ tuỳ chọn: là các nguồn gây biến động theo bảng ANOVA của thí nghiệm và được phân cách với nhau bởi dấu cách. Tương tác giữa các yếu tố được biểu thị bằng dấu hoa thị (*).

Minh họa cách viết cho ví dụ 1 thì

+ Biến phụ thuộc là năng suất (kí hiệu là yield).

+ Biến tuỳ chọn bao gồm: N, Var, Block và N*var.

và dòng lệnh được khai báo như sau:

*MODEL yield=block N var N*var*

6.2.2. Cách đọc kết quả bảng phân tích phương sai của SAS

- *The SAS System:* Hệ thống SAS

- *Analysis of Variance Procedure*

Chương trình xử lý số liệu SAS, phân tích phương sai ANOVA (analysis of variance procedure).

- *Class:* Các biến.

- *Levels:* Các mức của biến.

- *Values:* Tên hoặc mã số gán cho các mức của biến.

- *Number of observations in data set:* Số lượng quan trắc của toàn bộ số liệu tương ứng với số ô thí nghiệm.

- *Dependent variable:* Biến phụ thuộc.

- *Sources:* Nguồn biến động.

- *DF:* Độ tự do.

- *Yieldmean:* Trung bình chung (▯ EMBED Equation.3 ▯ ▯).

- *F value:* F-tính.

- *Pr>F:* Mức ý nghĩa.

Giá trị Pr>F chỉ cho biết sự sai khác của nguồn biến động của thí nghiệm có ý nghĩa hay không.

Nếu Pr >F của một nguồn biến động < 0.05 và > 0.01: Sự sai khác của nguồn biến động đó ở mức ý nghĩa 0.05 (Hay nói cách khác nguồn biến động đó sai khác nhau chắc chắn ở mức tin cậy 95%) (Kí hiệu *).

Nếu Pr >F của một nguồn biến động < 0.01 : Sự sai khác của nguồn biến động đó ở mức ý nghĩa 0.01 (Hay nói cách khác các mức của nguồn biến động đó sai khác nhau chắc chắn ở mức tin cậy 99%).

Nếu Pr >F của một nguồn biến động > 0.05 : Sự sai khác giữa các mức của nguồn biến động đó không có ý nghĩa (kí hiệu NS).

Các giá trị khác như df, SS, MS và F giống như trong phân tích biến động trình bày ở chương 3.

6.2.3. Phân tích phương sai cho thí nghiệm một nhân tố

6.2.3.1. Kiểu ngẫu nhiên hoàn toàn (Completely Randomized Design-CRD)

Để minh hoạ, ta dùng số liệu của thí nghiệm "So sánh hiệu quả của các loại thuốc trừ rày nâu ảnh hưởng đến năng suất lúa" (phần 3.3.2). Thí nghiệm gồm 7 công thức và 4 lần nhắc lại cho kết quả như sau:

Bảng 6.5. Kết quả thí nghiệm
"So sánh hiệu quả của các loại thuốc trừ rày nâu ảnh hưởng đến năng suất lúa"

Công thức (Treatment)	Năng suất (Yield) (kg/ha)			
	I	II	III	IV
Dol-Mix (1 kg)	2537	2069	2104	1797
Dol-Mix (2 kg)	3366	2591	2211	2544
DDT (BHC	2536	2459	2827	2385
Azodrin	2387	2453	1556	2116
Dimecron-boom	1997	1679	1649	1859
Dimecron-Knap	1796	1704	1904	1320
Control (Đối chứng)	1401	1516	1270	1077

Để thuận lợi cho quá trình nhập số liệu, các công thức được mã hoá như sau: **1**) Dol-Mix (1 kg); **2**- Dol-Mix (2 kg); **3**- DDT γBHC; **4**- Azodrin; **5**-Dimecron-boom;

6-Dimecron-Knap; 7- control. Như vậy số liệu được nhập trên EXCEL và lưu vào file **C:\ CRD.prn** như sau:

Bảng 6.6. Cách nhập số liệu cho thí nghiệm bố trí kiểu ngẫu nhiên hoàn chỉnh (số liệu từ bảng 6.5)

Rep	Treatmen	Yield
1	1	2537
1	2	3366
1	3	2536
1	4	2387
1	5	1997
1	6	1796
1	7	1401
2	1	2069
2	2	2591
2	3	2459
2	4	2453
2	5	1679
2	6	1704
2	7	1516
3	1	2104
3	2	2211
3	3	2827
3	4	1556
3	5	1649
3	6	1904
3	7	1270
4	1	1797
4	2	2544
4	3	2385
4	4	2116
4	5	1859
4	6	1320
4	7	1077

Chương trình xử lý số liệu cho CRD trên SAS

Trong cửa sổ Program Editor của SAS, chương trình được soạn thảo như sau:

```
Data crd;
Infile 'c:\crd.prn' firstobs=2;
Input rep trt yield;
Proc anova;
Class trt;
Model yield=trt;
Run;
```

Chú giải các lệnh

Data crd	→ Đặt tên bộ số liệu trong SAS là CRD
Infile 'C:\crd.prn'firstobs = 2;	→ Chỉ ra số liệu được lấy từ Excel trong file crd.prn ở ổ C và số liệu bắt đầu được tính từ dòng thứ 2.
Input rep trt yield;	→ khai báo 3 biến (các biến này sẽ ***tương ứng với 3 cột:*** **Rep, treatment và yield** ***trên file EXCEL***).
Proc anova;	→ Phân tích biến biến động bằng trình thủ tục ANOVA.
Class trt;	→ Khai báo các biến độc lập là trt (công thức).
Model yield=trt;	→ Mô hình phân tích: Năng suất (yield) phụ thuộc vào biến công thức (trt).
Run;	→ Chạy chương trình.

Thực hiện chạy chương trình và có kết quả trong của sổ OUTPUT như sau:

The SAS System

Analysis of Variance Procedure

Class Level Information

Class Levels Values

TRT 7 1 2 3 4 5 6 7

Number of observations in data set = 28

The SAS System

Analysis of Variance Procedure

Dependent Variable: YIELD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	5587174.929	931195.821	9.83	0.0001
Error	21	1990237.500	94773.214		
Corrected Total	27	7577412.429			

R-Square	C.V.	Root MSE	YIELD Mean
0.737346	15.09346	307.8526	2039.643

The SAS System

Analysis of Variance Procedure

Dependent Variable: YIELD

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TRT	6	5587174.929	931195.821	9.83	0.0001

Đọc kết quả bảng phân tích phương sai (chú giải các thuật ngữ tiếng anh đã trình bày ở phần 6.2.2):

Hàng thứ 4 và 5 của bảng kết quả chỉ cho biết: biến ở đây là "TRT"- công thức với 7 mức và có các giá trị là 1, 2, 3, 4, 5, 6 và 7.

Hàng thứ 9 của bảng kết quả cho biết: Biến phụ thuộc là "YIELD"- năng suất, hay có thể hiểu biến được phân tích biến động ở đây là năng suất.

Ta thấy có 2 bảng ANOVA; bảng 1 là kết quả phân tích biến động trong đó chia làm 2 nguồn biến động bao gồm biến động của "error" - ngẫu nhiên và của "Model" là những biến động khác. Bảng 2 là kết quả phân tích biến động trong đó chỉ có nguồn biến động của "TRT"- công thức. Vì thí nghiệm bố trí theo kiểu ngẫu nhiên hoàn chỉnh chỉ có 2 nguồn biến động nên các giá trị của phân tích biến động của "Model" và của "TRT" ở 2 bảng có kết quả như nhau.

Bảng 2 có cho biết giá trị của F Value của "TRT"- công thức = 9,83 đó chính là Ft mà ta cần biết để xem sai khác có ý nghĩa hay không. So sánh kết quả này với bảng 3.3 ở chương 3 cho thấy 2 cách làm đều cho kết quả Ft như nhau.

Kết luận: Bảng kết quả mà SAS in ra đã có sẵn kết quả so sánh Ft với Fb. Đối chiếu với cách hướng dẫn đọc kết quả ở phần 6.2.2. ta thấy nguồn biến động "TRT" - công thức có (Pr > F) = 0,0001 < 0,01 vì vậy ta kết luận các công thức khác nhau ở mức xác suất 99%. Kết luận này giống như kết luận của phần 3.2.1.1 ở chương 3.

Hệ số biến động CV= 15,09346

Từ kết quả xử lý số liệu trên, bảng ANOVA của thí nghiệm được lập như sau:

Bảng 6.7. Kết quả phân tích phương sai cho thí nghiệm bố trí theo kiểu ngẫu nhiên hoàn toàn (số liệu từ kết quả của SAS)

Nguồn biến động (Source)	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	F-tính (F-value)	Mức ý nghĩa Pr>F
Công thức	6	5587174,9	931195,8	9,83 **	0,0001
Sai số ngẫu nhiên	21	1990237,5	94773,2		
Tổng	27	7.577412,4			

CV = 15,1 %

Đánh giá kết quả thí nghiệm: (Pr>F) <0.01 ta kết luận hiệu quả của các thuốc trừ rày nâu đối với lúa khác nhau chắc chắn ở mức ý nghĩa 0.01(hay nói cách khác là hiệu quả của các thuốc trừ rày nâu đối với lúa khác nhau chắc chắn ở mức tin cậy 99%).

Trường hợp CRD với các công thức có số lần nhắc lại không giống nhau

Do đặc điểm của CRD là các xử lý số liệu không phức tạp kể cả khi số lần nhắc lại giữa các công thức không bằng nhau (số lần nhắc lại của các công thức không bằng nhau có thể là do đặc thù của vật liệu thí nghiệm hoặc là các thí nghiệm có số lần nhắc lại bằng nhau nhưng do một vài đơn vị thí nghiệm bị mất hoặc bị hỏng trong quá trình thí nghiệm). Để xử lý số liệu trong những trường hợp này thì:

- Phần nhập số liệu trên EXCEL: những số liệu bị mất hoặc thiếu được nhập dưới dạng dấu chấm '.'.

- Phần lập trình xử lý số liệu trên SAS phải dùng PROC GLM thay vì PROC ANOVA và câu lệnh INFILE cần đưa thêm chữ 'missover' để chỉ định rằng số liệu không đầy đủ.

Ví dụ: Xử lý số liệu cho thí nghiệm kiểu CRD trên nhưng giả sử thí nghiệm bị mất một số ô (công thức 1 lần nhắc lại 1, công thức 3 lần nhắc 1, công thức 4 lần nhắc 2 và công thức 5 lần nhắc 3).

Số liệu được nhập trên EXCEL và lưu trong file **C:\CRDM.prn** như sau:

Bảng 6.8. Nhập số liệu cho thí nghiệm bố trí theo kiểu ngẫu nhiên hoàn toàn trong trường hợp mất số liệu

Rep	Treatmen	Yield
1	1	.
1	2	3366
1	3	.
1	4	2387
1	5	1997
1	6	1796
1	7	1401
2	1	2069
2	2	2591
2	3	2459
2	4	.
2	5	1679
2	6	1704
2	7	1516
3	1	2104
3	2	2211
3	3	2827
3	4	1556
3	5	.
3	6	1904
3	7	1270
4	1	1797
4	2	2544
4	3	2385
4	4	2116
4	5	1859
4	6	1320
4	7	1077

Chương trình SAS xử lý số liệu kiểu CRD trong trường hợp mất số liệu:

```
data crd;
infile 'C:\CRDM.prn' firstobs=2 missover;
input rep trt yield;
proc glm;
class trt;
model yield=trt;
run;
```

Thực hiện chạy chương trình ta có kết quả trong của sổ OUTPUT của trường hợp mất số liệu trên như sau:

The SAS System

General Linear Models Procedure

Class Level Information

Class Levels Values

TRT 7 1 2 3 4 5 6 7

Number of observations in data set = 28

NOTE: Due to missing values, only 24 observations can be used in this analysis.

The SAS System

General Linear Models Procedure

Dependent Variable: YIELD

Sum of

Source	DF	Squares	Square	F Value	Pr > F
Model	6	5121310.292	853551.715	9.09	0.0002
Error	17	1595854.667	93873.804		
Corrected Total	23	6717164.958			

R-Square	C.V.	Root MSE	YIELD Mean
0.762421	15.34019	306.3883	1997.292

The SAS System

General Linear Models Procedure

Dependent Variable: YIELD

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TRT	6	5121310.292	853551.715	9.09	0.0002
Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	6 5	121310.292	853551.715	9.09	0.0002

(*Chú ý:* với Proc GLM Type III SS sẽ được sử dụng thay vì Type I SS)

Từ kết quả xử lý số liệu trên, bảng ANOVA được lập như sau:

Bảng 6.8. Kết quả phân tích phương sai cho thí nghiệm bố trí kiểu ngẫu nhiên hoàn toàn trong trường hợp mất số liệu (số liệu lấy từ kết quả của SAS)

Nguồn biến động (Source)	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	F-tính (F-value)	Mức ý nghĩa Pr>F
Công thức	6	5121310,3	853551,7	9,09	0,0002
Sai số ngẫu nhiên	17	1595854,7	93873,8		
Tổng	23	6717165,0			

CV = 15,3 %

Độ tự do của công thức không thay đổi vì số lượng công thức không thay đổi. Độ tự do của sai số giảm, phản ánh số lượng quan trắc (observations) giảm, đó là do ảnh hưởng của mất số liệu (4 ô thí nghiệm bị mất số liệu). Chính độ tự do của sai số giảm dẫn đến độ chính xác của thí nghiệm giảm (CV% tăng so với trường hợp không mất số liệu).

6.3.2.2 Kiểu khối ngẫu nhiên hoàn chỉnh *(Randomied Complete block Design-RCBD).*

Để minh họa ta xử lý cho thí nghiệm "Nghiên cứu ảnh hưởng của lượng hạt giống gieo đến năng suất lúa" (phần 3.2.2). Thí nghiệm gồm 6 công thức (t=6), (6 lượng hạt giống gieo: 25, 50, 75, 100, 125, 150 kg/ha) và 4 lần nhắc lại (r=4). Kết quả thí nghiệm như sau:

Bảng 6.9. Năng suất lúa của các công thức gieo hạt giống khác nhau

Công thức(kg hạt/ha)	NS hạt (kg/ha)			
	I	II	III	IV
25	5113	5398	5307	4678
50	5346	5952	4719	4264
75	5272	5713	5483	4749
100	5164	4831	4986	4410
125	4804	4848	4432	4748
150	5254	4542	4919	4098

Thực hiện mã hoá các công thức thí nghiệm: 1- 25 (kg hạt/ha); 2-50 (kg hạt/ha); 3-75(kg hạt/ha); 4-100 (kg hạt/ha); 5-125; 6-150 (kg hạt/ha).

Tương tự như việc nhập số liệu với CRD nhưng chú ý đối với kiểu RCBD thì bắt buộc phải có cột Rep (nhắc lại) vì cột này là một biến trong xử lý thống kê số liệu thí nghiệm RCBD.

Số liệu được nhập trên Excel và lưu trữ vào file **C:\RCBD.prn** như sau:

Bảng 6.10. Cách nhập số liệu cho thí nghiệm bố trí kiểu khối ngẫu nhiên hoàn chỉnh

rep	Trt	yield
1	1	5113
1	2	5346
1	3	5272
1	4	5164
1	5	4804
1	6	5254
2	1	5398
2	2	5952
2	3	5713
2	4	4831
2	5	4848
2	6	4542
3	1	5307
3	2	4719
3	3	5483
3	4	4986
3	5	4432
3	6	4919
4	1	4678
4	2	4264
4	3	4749
4	4	4410
4	5	4748
4	6	4098

Chương trình xử lý số liệu cho RCBD trên SAS

```
data rcbd;
infile 'C:\rcbd.prn' firstobs=2;
input rep trt yield;
proc anova;
class rep trt;
model yield= rep trt;
run;
```

Chú ý: Do Rep (khối) cũng là một nguồn biến động trong kiểu RCB nên trong câu lệnh CLASS và MODEL phải có biến này.

Kết quả phân tích thống kê trong cửa sổ OUTPUT:

The SAS System

Analysis of Variance Procedure

Dependent Variable: YIELD

Sum of Mean

Source	DF	Squares	Square	F Value	Pr > F
Model	8	3142691.6667	392836.4583	3.55	0.0165
Error	15	1658376.1667	110558.4111		
Corrected Total	23	4801067.8333			

R-Square	C.V.	Root MSE	YIELD Mean
0.654582	6.704258	332.50325	4959.5833

Source	DF	Anova SS	Mean Square	F Value	Pr > F
REP	3	1944360.8333	648120.2778	5.86	0.0074
TRT	5	1198330.8333	239666.1667	2.17	0.1128

Đọc kết quả phân tích ANOVA:

Tương tự như phân tích của thí nghiệm bố trí kiểu ngẫu nhiên hoàn toàn ta cũng có 2 bảng kết quả ANOVA.

Bảng 1 cho phân tích biến động do ngẫu nhiên- "error"; bảng 2 cho phân tích biến động cho "REP"- nhắc lại và "TRT"- công thức.

Vì biến REP đã được khai báo trong phần lập trình nên trong phần kết quả của RCBD sẽ có 2 nguồn biến động là REP và TRT.

- Trong kết quả phân tích biến REP (khối) có (Pr>F) = 0,0074 (P<0,05), có nghĩa rằng việc bố trí theo khối đã có ý nghĩa làm giảm sai số thí nghiệm

- Vì *(Pr>F)* của biến công thức = 0,1128 > 0,05 ta kết luận rằng 6 tỷ lệ gieo hạt có ảnh hưởng không khác nhau đến năng suất. Vì không có sự khác nhau chắc chắn giữa các công thức nên ta không cần so sánh giữa các công thức và không cần tính LSD.

Hệ số biến động CV= 6,70

Kết quả này giống như kết quả phân tích của phần 3.2.2. ở chương 3.

6.2.3.3. Kiểu bố trí ô vuông la tinh

Minh họa phân tích biến động cho thí nghiệm "So sánh năng suất của các giống lúa A, B, C (đối chứng) và D" bố trí theo kiểu ô vuông la tinh (phần 3.2.3). Kết quả thí nghiệm như sau:

Bảng 6.11. Năng suất lúa trong thí nghiệm so sánh giống lúa bố trí theo kiểu ô vuông la tinh

STT hàng	Năng suất hạt (tấn/ha) (Xi)			
	Cột 1	Cột 2	Cột 3	Cột 4
Hàng 1	1,640 (B)	1,210 (D)	1,425 (C)	1,345 (A)
Hàng 2	1,475 (C)	1,185 (A)	1,400 (D)	1,290 (B)
Hàng 3	1,670 (A)	0,710 (C)	1,665 (B)	1,180 (D)
Hàng 4	1,565 (D)	1,290 (B)	1,665 (A)	0,660 (C)

Kiểu ô vuông la tinh có thêm các biến hàng và cột do vậy trong phần nhập số liệu và xử lý số liệu cũng phải có thêm 2 biến này.

Mã hoá các công thức A, B, C, D từ 1 đến 4. Số liệu nhập vào Excel gồm 4 cột : Hàng (row), Cột (column); công thức (trt) và năng suất (yield) như sau:

Số liệu được nhập trên EXCEL và lưu trữ trong file **C:\ LS.prn** như sau:

Bảng 6.12. Cách nhập số liệu cho thí nghiệm bố trí kiểu ô vuông la tinh

Row	Column	trt	Yield
1	1	2	1.64
1	2	4	1.21
1	3	3	1.425
1	4	1	1.345
2	1	3	1.475
2	2	1	1.185
2	3	4	1.4
2	4	2	1.29
3	1	1	1.67
3	2	3	0.71
3	3	2	1.665
3	4	4	1.18
4	1	4	1.565
4	2	2	1.29
4	3	1	1.655
4	4	3	0.66

Chương trình phân tích phương sai cho kiểu thí nghiệm ô vuông la tinh trong SAS.

```
data LS;
infile 'C:\daysas\LS.prn' firstobs=2;
input row column trt yield;
proc ANOVA;
class row column trt;
model yield= row column trt ;
run;
```

Chú ý: Sự khác nhau trong lập trình của ô vuông la tinh với hai kiểu bố trí CRD và RCBD là có thêm biến **row** (hàng) và biến **column** (cột) vì kiểu bố trí này là bố trí theo hàng và cột.

Kết quả xử lý thống kê:

The SAS System

Analysis of Variance Procedure

Class Level Information

Class Levels Values

ROW 4 1 2 3 4

COLUMN 4 1 2 3 4

TRT 4 1 2 3 4

Number of observations in data set = 16

The SAS System

Analysis of Variance Procedure

Dependent Variable: YIELD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	1.28433906	0.14270434	6.61	0.0161
Error	6	0.12958437	0.02159740		
Corrected Total	15	1.41392344			

R-Square	C.V.	Root MSE	YIELD Mean
0.908351	11.00570	0.1469605	1.3353125

Source	DF	Anova SS	Mean Square	F Value	Pr > F
ROW	3	0.03015469	0.01005156	0.47	0.7170
COLUMN	3	0.82734219	0.27578073	12.77	0.0051
TRT	3	0.42684219	0.14228073	6.59	0.0251

Đọc kết quả bảng ANOVA

Tương tự như phần trước ta có 2 bảng ANOVA; bảng 1 cho kết quả phân tích biến động của "error"- ngẫu nhiên. Bảng 2 cho kết quả phân tích biến động của các nguồn khác bao gồm "ROW" - hàng; "COLUMN"- cột và "TRT" - công thức.

Kết quả cho thấy biến động của "TRT" có (Pr >F) = 0,0251 <0,05 nên ta kết luận năng suất của 4 giống lúa khác nhau nhau chắc chắn ở mức tin cậy 95 %.

Trong thí nghiệm trên chia khối theo hàng không có ý nghĩa làm giảm sai số thí nghiệm vì biến động của hàng có (Pr>F) = 0,7 > 0,05). Chia khối theo cột có ý nghĩa cao làm giảm sai số thí nghiệm vì biến động của cột có (Pr>F) = 0,0051 <0,01.

Hệ số biến động CV= 11,005

Kết quả trên cũng tương tự như kết quả của mục 3.2.3. ở chương 3.

6.2.4. Phân tích phương sai cho thí nghiệm 2 nhân tố

Đối với các thí nghiệm 2 nhân tố trở lên xuất hiện thêm tương tác giữa các nhân tố. Tương tác giữa các nhân tố phản ánh ảnh hưởng của một nhân tố này trong mối quan hệ với các nhân tố khác. Tương tác giữa hai nhân tố A và B được kí hiệu là A × B.

6.2.4.1. Kiểu khối ngẫu nhiên hoàn chỉnh (Randomized Complete Block Design - RCBD)

Ví dụ minh họa: Phân tích biến động cho năng suất thu được từ thí nghiệm "Nghiên cứu ảnh hưởng của đạm (với 5 mức bón khác nhau kí hiệu N0, N1, N2, N3 và N4) đến năng suất của 3 giống lúa (kí hiệu V1, V2 và V3)", thí nghiệm bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh với 4 lần lặp lại (phần 3.3.2). Kết quả thí nghiệm như sau:

Bảng 6.13. Năng suất lúa của thí nghiệm hai nhân tố bố trí theo kiểu khối ngẫu nhiên hoàn chỉnh

Các mức N (kg/ha)	NS hạt (tấn/ha)			
	Rep I	Rep II	Rep III	Rep IV
		V1		
No	3,852	2,606	3,144	2,894
N1	4,788	4,936	4,562	4,608
N2	4,576	4,454	4,884	3,924
N3	6,034	5,276	5,906	5,652
N4	5,874	5,916	5,984	5,518
		V2		
No	2,846	3,794	4,108	3,444
N1	4,956	5,128	4,150	4,990
N2	5,928	5,698	5,810	4,308
N3	5,664	5,362	6,458	5,474
N4	5,458	5,546	5,786	5,932
		V3		
No	4,192	3,754	3,738	3,428
N1	5,250	4,582	4,896	4,286
N2	5,822	4,848	5,678	4,932
N3	5,888	5,524	6,042	4,756
N4	5,864	6,264	6,056	5,362

Mã hoá 5 mức đạm N0, N1, N2, N3 và N4 tương ứng từ 1 đến 5 và 3 giống lúa V1, V2 và V3 tương ứng từ 1 đến 3. Kết quả thí nghiệm được nhập trên EXCEL và lưu trữ trong file **C:\2-RCBD.prn** như sau:

Bảng 6.14. Cách nhập số liệu cho thí nghiệm 2 nhân tố bố trí kiểu khối ngẫu nhiên hoàn chỉnh

Rep	Var	N	Yield
1	1	1	3.852
1	1	2	4.788
1	1	3	4.576
1	1	4	6.034
1	1	5	5.874
1	2	1	2.846
1	2	2	4.956
1	2	3	5.928
1	2	4	5.664
1	2	5	5.458
1	3	1	4.192
1	3	2	5.25
1	3	3	5.822
1	3	4	5.888
1	3	5	5.864
2	1	1	2.606
2	1	2	4.936
2	1	3	4.454
2	1	4	5.276
2	1	5	5.916
2	2	1	3.794
2	2	2	5.128
2	2	3	5.698
2	2	4	5.362
2	2	5	5.546
2	3	1	3.754
2	3	2	4.582
2	3	3	4.848
2	3	4	5.524
2	3	5	6.264
3	1	1	3.144
3	1	2	4.562
3	1	3	4.884
3	1	4	5.906
3	1	5	5.984
3	2	1	4.108
3	2	2	4.15
3	2	3	5.81
3	2	4	6.458

Rep	Var	N	Yield
3	2	5	5.786
3	3	1	3.738
3	3	2	4.896
3	3	3	5.678
3	3	4	6.042
3	3	5	6.056
4	1	1	2.894
4	1	2	4.608
4	1	3	3.924
4	1	4	5.652
4	1	5	5.518
4	2	1	3.444
4	2	2	4.99
4	2	3	4.308
4	2	4	5.474
4	2	5	5.932
4	3	1	3.428
4	3	2	4.286
4	3	3	4.932
4	3	4	4.756
4	3	5	5.362

Chương trình phân tích phương sai cho thí nghiệm 2 nhân tố kiểu RCBD

```
data Factor;
infile 'C:\2-RCBD.prn' firstobs=2;
input rep var N yield;
proc anova;
class rep var N;
model yield= rep var N var*N;
run;
```

Chú ý: biến giống kí hiệu là var, biến đạm kí hiệu là N và tương tác giữa đạm và giống là var*N.

Kết quả phân tích

```
The SAS System
Analysis of Variance Procedure
Class Level Information
Class Levels Values
REP   4 1 2 3 4
VAR   3 1 2 3
N     5 1 2 3 4 5
```

Number of observations in data set = 60

The SAS System

Analysis of Variance Procedure

Dependent Variable: YIELD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	17	47.17807273	2.77518075	18.35	0.0001
Error	42	6.35279627	0.15125705		
Corrected Total	59	53.53086900			

R-Square	C.V.	Root MSE	YIELD Mean
0.881325	7.846622	0.3889178	4.9565000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
REP	3	2.59981673	0.86660558	5.73	0.0022
VAR	2	1.05278440	0.52639220	3.48	0.0400
N	4	41.23474533	10.30868633	68.15	0.0001
VAR*N	8	2.29072627	0.28634078	1.89	0.0867

Đọc kết quả phân tích ANOVA

Từ bảng 1 ta có kết quả phân tích biến động của "error" có dfe = 42; Sse = 6,35279627; MSe = 0,15125705;

Từ bảng 2 ta có kết quả phân tích biến động của "VAR"- giống có (Pr>F) = 0,04 < 0,05 vì thế kết luận các giống cho năng suất sai khác nhau có ý nghĩa ở mức tin cậy 95%; "N"- đạm có (Pr > F) = 0,0001< 0,01, vì thế kết luận các mức đạm khác nhau cho năng suất khác nhau có ý nghĩa ở mức xác suất 99%; "VAR*N"- tương tác giống và đạm có (Pr >F) = 0,0867 > 0,05 như vậy tương tác giữa đạm và giống không có ý nghĩa; có nghĩa là các giống có phản ứng với đạm như nhau.

Hệ số biến động CV= 7,8%

Từ kết quả phân tích trên bảng ANOVA được lập như sau:

Bảng 6.15. Kết quả phân tích phương sai cho thí nghiệm 2 nhân tố bố trí kiểu khối ngẫu nhiên hoàn chỉnh (số liệu từ kết quả của SAS)

Nguồn biến động (Source)	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	F-tính (F-value)	Mức ý nghĩa Pr>F
Nhắc lại (Rep)	3	2,5998	0,866	5,73 **	0,0022
Giống (A)	2	1,0528	0,526	3,48 *	0,0400
Đạm (B)	4	41,2347	10,308	68,15 **	0,0001
Var*N (A xB)	8	2,2907	0,286	1,89 ns	0,0867
Sai số ngẫu nhiên	42	6,3527	0,151		
Tổng	59	53,5308			

CV = 7,8 %

6.2.4.2. Kiểu ô phụ hay kiểu ô chính ô phụ (Split-plot design)

Ví dụ minh hoạ: Phân tích biến động (phương sai) cho thí nghiệm nghiên cứu phản ứng của 4 giống lúa (kí hiệu V1, V2, V3 và V4) với 6 mức bón đạm (kí hiệu No, N1, N2, N3, N4 và N5). Thí nghiệm bố trí theo kiểu ô chính ô phụ với 3 lần nhắc lại. Nhân tố đạm (N) được bố trí vào ô chính (A), nhân tố giống (Var) được bố trí vào ô phụ (B) (phần 3.3.3). Kết quả thí nghiệm như sau:

Bảng 6.16. Kết quả thí nghiệm 2 nhân tố bố trí kiểu ô chính ô phụ

Giống (Var)	Năng suất (Yield) (kg/ha)		
	I	II	III
	N_o (0 kg N/ha)		
V_1 (IR8)	4430	4478	3850
V_2 (IR5)	3944	5314	3660
V_3 (C4-63)	3464	2944	3142
V_4 (Peta)	4126	4482	4836
	N_1 (60 kg N/ha)		
V_1	5418	5166	6432
V_2	6502	5858	5586
V_3	4768	6004	5556
V_4	5192	4604	4652
	N_2 (90 kg N/ha)		
V_1	6076	6420	6704
V_2	6008	6127	6642
V_3	6244	5724	6014
V_4	4546	5744	4146
	N_3 (120 kg N/ha)		
V_1	6462	7056	6680
V_2	7139	6982	6564
V_3	5792	5880	6370
V_4	2774	5036	3638

Giống (Var)	Năng suất (Yield) (kg/ha)		
	I	II	III
	N_4 (150 kg N/ha)		
V_1	7290	7848	7552
V_2	7682	6594	6576
V_3	7080	6662	6320
V_4	1414	1960	2766
	N_5 (180 kgN/ ha)		
V_1	8452	8832	8818
V_2	6228	7387	6006
V_3	5594	7122	5480
V_4	2248	1380	2014

Mã hoá các mức của nhân tố thí nghiệm như sau:

4 giống (V1, V2, V3 và V4) tướng ứng với 4 số từ 1 đến 4.

6 mức bón đạm (No, N1, N2, N3, N4 và N5) tương ứng với 6 số từ 1 đến 6.

Số liệu được nhập trên EXCEL và lưu trữ trong file **C:\2-split.prn** như sau:

Bảng 6.17. Cách nhập số liệu cho thí nghiệm 2 nhân tố bố trí kiểu ô chính ô phụ

Rep	N	Var	Yield
1	1	1	4430
1	1	2	3944
1	1	3	3464
1	1	4	4126
1	2	1	5418
1	2	2	6502
1	2	3	4768
1	2	4	5192
1	3	1	6076
1	3	2	6008
1	3	3	6244
1	3	4	4546
1	4	1	6462
1	4	2	7139
1	4	3	5792
1	4	4	2774
1	5	1	7290
1	5	2	7682
1	5	3	7080
1	5	4	1414
1	6	1	8452
1	6	2	6228
1	6	3	5594
1	6	4	2248
2	1	1	4478
2	1	2	5314
2	1	3	2944

Rep	N	Var	Yield
2	1	4	4482
2	2	1	5166
2	2	2	5858
2	2	3	6004
2	2	4	4604
2	3	1	6420
2	3	2	6127
2	3	3	5724
2	3	4	5744
2	4	1	7056
2	4	2	6982
2	4	3	5880
2	4	4	5036
2	5	1	7848
2	5	2	6594
2	5	3	6662
2	5	4	1960
2	6	1	8832
2	6	2	7387
2	6	3	7122
2	6	4	1380
3	1	1	3850
3	1	2	3660
3	1	3	3142
3	1	4	4836
3	2	1	6432
3	2	2	5586
3	2	3	5556
3	2	4	4652
3	3	1	6704
3	3	2	6642
3	3	3	6014
3	3	4	4146
3	4	1	6680
3	4	2	6564
3	4	3	6370
3	4	4	3638
3	5	1	7552
3	5	2	6576
3	5	3	6320
3	5	4	2766
3	6	1	8818
3	6	2	6006
3	6	3	5480
3	6	4	2014

Chương trình phân tích phương sai cho kiểu ô chính ô phụ trong SAS

```
data split;
infile 'c:\2-splitG.prn' firstobs=2;
input rep N var yield;
Proc GLM;
class rep N var;
model yield=rep N rep*N var N*var;
test h=rep N e=rep*N;
test h=var N*var;
run;
```

Chú ý: Sự khác nhau giữa lập trình xử lý số liệu cho kiểu RCBD và cho kiểu ô chính ô phụ.

- Trong kiểu RCBD chỉ có một loại sai số ngẫu nhiên, vì vậy sai số này cũng chính là sai số của cả mô hình và SAS luôn ngầm định dùng sai số của cả mô hình để kiểm định các giả thuyết và trong kiểu RCBD ta không phải chỉ định sai số cho SAS.

- Trong kiểu ô chính ô phụ có 2 loại sai số ngẫu nhiên khác nhau tương ứng với 2 ô: ô chính (main plot) và ô phụ (subplot). Sai số dùng để kiểm định yếu tố ô chính (yếu tố A) thường được gọi là sai số a (kí hiệu Ea) và một loại sai số khác được gọi là sai số b (kí hiệu Eb) dùng để kiểm định yếu tố ô phụ (yếu tố B) và tương tác giữa yếu tố ô chính và yếu tố ô phụ (A x B). Vì vậy trong chương trình xử lý cho kiểu ô chia nhỏ ta có câu lệnh:

Test h=rep N e=rep*N: Chỉ định SAS dùng sai số a (rep*N) để kiểm định ảnh hưởng của khối (rep) và ảnh hưởng của đạm (N).

Các kiểm định còn lại: Không chỉ định sai số cho SAS, do vậy SAS sẽ dùng sai số ngầm định (pooled error), sai số này chính là sai số b (rep*N*var) để kiểm định.

Kết quả phân tích phương sai trong SAS

```
The SAS System
General Linear Models Procedure
Class Level Information
Class Levels Values
REP  3 1 2 3
N    6 1 2 3 4 5 6
VAR  4 1 2 3 4
Number of observations in data set = 72
```

The SAS System

General Linear Models Procedure

Dependent Variable: YIELD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	35	192163043.15	5490372.66	15.71	0.0001
Error	36	12584873.17	349579.81		
Corrected Total	71	204747916.32			

R-Square	C.V.	Root MSE	YIELD Mean
0.938535	10.79144	591.25275	5478.9028

Source	DF	Type I SS	Mean Square	F Value	Pr > F
REP	2	1082576.694	541288.347	1.55	0.2264
N	5	30429199.569	6085839.914	17.41	0.0001
REP*N	10	1419678.806	141967.881	0.41	0.9348
VAR	3	89888101.153	29962700.384	85.71	0.0001
N*VAR	15	69343486.931	4622899.129	13.22	0.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
REP	2	1082576.694	541288.347	1.55	0.2264
N	5	30429199.569	6085839.914	17.41	0.0001
REP*N	10	1419678.806	141967.881	0.41	0.9348
VAR	3	89888101.153	29962700.384	85.71	0.0001
N*VAR	15	69343486.931	4622899.129	13.22	0.0001

Tests of Hypotheses using the Type III MS for REP*N as an error term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
REP	2	1082576.694	541288.347	3.81	0.0588
N	5	30429199.569	6085839.914	42.87	0.0001

Đọc kết quả phân tích phương sai:

- Bảng ANOVA 2 cho biết sai số của "var" - giống với F = 85,71 và Pr > F = 0,0001

- Bảng ANOVA 3 cho biết sai số của tương tác "N*VAR"- đạm*giống với F = 13,22 và Pr > F = 0,0001.

- Bảng ANOVA 4 cho biết sai số của "N" với F = 42,87 và Pr > F = 0,0001

- Kết luận: Các giống lúa khác nhau cho năng suất khác nhau ở mức độ tin cậy 99%; Các mức đạm khác nhau cũng cho năng suất lúa khác nhau ở mức độ tin cậy 99%; Phản ứng của các giống lúa cũng khác nhau đối với những mức đạm khác nhau ở mức độ tin cậy 99%. Kết luận này tương tự với kết quả tính được của phần 3.4.2. ở chương 3.

Chú ý: Đọc kết quả phân tích phương sai của SAS :

rep*N là sai số (a) và sai số còn lại của mô hình (pooled error) chính là sai số (b). Sai số này cũng là sai số ngầm định để SAS kiểm định các giả thuyết.

Kiểu thí nghiệm này có hai loại sai số tương ứng cho 2 nhân tố. Vì vậy cũng có 2 loại CV tương ứng cho 2 loại nhân tố này. SAS chỉ đưa ra một loại CV. CV này được tính dựa trên sai số ngầm định do vậy CV này chính là CV của yếu tố ô phụ (CV(b)), CV còn lại là CV của yếu tố ô chính (CV(a)) ta phải tự tính theo công thức.

$$CV(a) = \frac{\sqrt{MSe(a)}}{\overline{X}} \times 100 = \frac{\sqrt{141967.881}}{5478.9028} \times 100 = 6,9\%$$

Giá trị của cv (a) biểu thị độ chính xác của yếu tố ô chính (yếu tố đạm).

Giá trị cv(b) biểu thị độ chính xác của yếu tố ô phụ (yếu tố giống).

***6.2.4.3. Thí nghiệm theo lô sọc** (Strip plot design)*

(Dành cho sinh viên tự tham khảo)

Thí nghiệm 2 nhân tố bố trí theo kiểu lô sọc thường áp dụng trong những trường hợp muốn đánh giá chính xác ảnh hưởng của tương tác giữa hai yếu tố hơn so với việc khảo sát ảnh hưởng của từng yếu tố. Thí nghiệm sẽ có 3 loại lô:

- Lô ngang dành cho yếu tố theo phương ngang gọi tắt là yếu tố ngang (horizontal factor).

- Lô dọc dành cho yếu tố theo phương dọc gọi tắt là yếu tố dọc (vertical factor).

- Lô nhỏ (giao của lô ngang và lô dọc) để theo dõi tương tác giữa hai yếu tố đó.

Để minh họa cho phương pháp phân tích phương sai kiểu thí nghiệm này, ta xét thí dụ: Phân tích năng suất thu được từ thí nghiệm của 6 giống lúa với 3 mức bón đạm khác nhau. Thí nghiệm với 3 lần nhắc lại và bố trí theo kiểu lô sọc, trong đó yếu tố giống (var) là yếu tố ngang (factor A) và đạm (N) là yếu tố dọc (factor B).

Kết quả thí nghiệm trình bày ở bảng 6.18.

Bảng 6.18. Năng suất của 6 giống lúa ở 3 mức bón N khác nhau trong thí nghiệm lô sọc với 3 lần nhắc lại

Mức đạm (N) (kg/ha)	Năng suất (yield) (kg/ha)		
	I	II	III
IR8 (V1)			
0 (N1)	2373	3958	4384
60 (N2)	4076	6431	4889
120 (N3)	7254	6808	8582
IR127-80 (V2)			
0	4007	5795	5001
60	5630	7334	7177
120	7053	8284	6297
IR305-4-12 (V3)			
0	2620	4508	5621
60	4676	6672	7019
120	7666	7328	8611
IR400-2-5 (V4)			
0	2726	5630	3821
60	4838	7007	4816
120	6881	7735	6667
IR665-58 (V5)			
0	4447	3276	4582
60	5549	5340	6011
120	6880	5080	6076
Peta (V6)			
0	2572	3724	3326
60	3896	2822	4425
120	1556	2706	3214

Nguồn: Gomez, 198

Sinh viên tự nhập số liệu trên EXCEL.

Chương trình xử lý trong SAS

```
data strip;
infile 'đường dẫn\tên tệp.prn' firstobs=2;
input rep var N yield;
proc glm;
class rep var N;
model yield=rep var rep*var N rep*N var*N;
test h=rep var e=rep*var;
test h=N e=rep*N;
run;
```

Kết quả xử lý

The SAS System

General Linear Models Procedure

Class Level Information

Class Levels Values

REP 3 1 2 3

VAR 6 1 2 3 4 5 6

N 3 1 2 3

Number of observations in data set = 54

The SAS System

General Linear Models Procedure

Dependent Variable: YIELD

Sum of Mean

Source	DF	Squares	Square	F Value	Pr > F
Model	33	158.77273161	4.81129490	11.69	0.0001
Error	20	8.23291722	0.41164586		
Corrected Total	53	167.00564883			

R-Square	C.V.	Root MSE	YIELD Mean
0.950703	12.12860	0.6415963	5.2899444

Source	DF	Type I SS	Mean Square	F Value	Pr > F
REP	2	9.22096233	4.61048117	11.20	0.0005
VAR	5	57.10020128	11.42004026	27.74	0.0001
REP*VAR	10	14.92261922	1.49226192	3.63	0.0069
N	2	50.67606144	25.33803072	61.55	0.0001
REP*N	4	2.97490789	0.74372697	1.81	0.1672
VAR*N	10	23.87797944	2.38779794	5.80	0.0004

Source	DF	Type III SS	Mean Square	F Value	Pr > F
REP	2	9.22096233	4.61048117	11.20	0.0005
VAR	5	57.10020128	11.42004026	27.74	0.0001
REP*VAR	10	14.92261922	1.49226192	3.63	0.0069
N	2	50.67606144	25.33803072	61.55	0.0001
REP*N	4	2.97490789	0.74372697	1.81	0.1672
VAR*N	10	23.87797944	2.38779794	5.80	0.0004

Tests of Hypotheses using the Type III MS for REP*VAR as an error term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
REP	2	9.22096233	4.61048117	3.09	0.0902
VAR	5	57.10020128	11.42004026	7.65	0.0034

Tests of Hypotheses using the Type III MS for REP*N as an error term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
N	2	50.67606144	25.33803072	34.07	0.0031

Chú ý: Kiểu ô dọc có 3 loại sai số thí nghiệm tương ứng với 3 loại ô thí nghiệm. Sai số a (rep*var) dùng để đánh giá yếu tố giống, sai số b (rep*N) dùng để đánh giá yếu tố đạm (N) và sai số c là sai số còn lại (pooled error) dùng để đánh giá tương tác N * var.

Sinh viên tự lập bảng ANOVA và đánh giá kết quả cho thí nghiệm trên.

6.3. SỬ DỤNG PHẦN MỀM SAS SO SÁNH SỐ TRUNG BÌNH

6.3.1. So sánh cặp đôi - LSD và DUNCAN

Dựa vào kết quả phân tích phương sai mà quyết định nên sử dụng phép so sánh nào cho hợp lý. Đối với thí nghiệm 2 nhân tố nhưng tương tác giữa các nhân tố không có ý nghĩa thì để đánh giá ảnh hưởng của từng nhân tố phải dựa trên ảnh hưởng chung của nhân tố đó (main effect).

Trong SAS câu lệnh "MEANS biến/lựa chọn" được dùng để so sánh ảnh hưởng chung của các yếu tố. Câu lệnh này phải đứng sau câu lệnh MODEL và đứng trước câu lệnh RUN trong các chương trình xử lý.

Biến: Kê khai các yếu tố hay công thức mà ta cần so sánh phân hạng.

Lựa chọn: Chọn kiểu so sánh DUNCAN hay LSD.

Minh họa bằng thực hiện so sánh cho thí nghiệm trên lúa kiểu CRD gồm 6 công thức sử dụng thuốc và một công thức đối chứng (công thức 7) với 4 lần nhắc lại (mục 6.2.3.1). Kết quả phân tích ANOVA của thí nghiệm này cho kết luận: Các loại thuốc

trừ rày nâu có ảnh hưởng khác nhau đến năng suất lúa chắc chắn ở mức tin cậy 99%. Để so sánh cụ thể xem những công thức sử dụng thuốc nào tốt hơn chắc chắn so với đối chứng ta có hai cách thực hiện sau.

6.3.1.1. So sánh cặp đôi dùng LSD

Cách viết chương trình như sau:

- Gọi lại chương trình SAS đã viết để phân tích phương sai.

- Chèn thêm câu lệnh "***means trt/LSD;***"câu lệnh này có nghĩa hãy thực hiện so sánh phân hạng các công thức sử dụng LSD.

- Nếu muốn tính LSD ở mức α cụ thể nào đó thì nhập giá trị này bằng lựa chọn ALPHA= giá trị cần chọn. Nếu không chỉ định giá trị α cụ thể SAS sẽ lấy giá trị α ngầm định là 0.05.

- Để tính cả $LSD_{.05}$ và $LSD_{.01}$ cho việc so sánh giữa các công thức thì cần thêm 2 câu lệnh sau vào chương trình:

```
means trt/LSD;
means trt/LSD ALPHA=0.01;
```

Chương trình xử lý của SAS, có so sánh dùng LSD cho thí nghiệm trên:

```
Options nodates ps=76 ls=80;
Data crd;
Infile 'a:\crd.prn' firstobs=2;
Input rep trt yield;
Proc anova;
Class trt;
Model yield=trt;
Means trt/LSD;
Means trt/LSD ALPHA=0.01;
Run;
```

Lúc này kết quả xử lý thống kê sẽ có thêm các phần sau (đây chính là kết quả của 2 câu lệnh **Means trt/LSD**;và **Means trt/LSD ALPHA=0.01**;)

```
T tests (LSD) for variable: YIELD
NOTE: This test controls the type I comparisonwise error rate not the
experimentwise error rate.
    Alpha= 0.05 df= 21 MSE= 94773.21
        Critical Value of T= 2.08
Least Significant Difference= 452.7
```

Means with the same letter are not significantly different.

T	Grouping	Mean	N	TRT
	A	2678.0	4	2
	A			
B	A	2551.8	4	3
B				
B	C	2128.0	4	5
B	C			
B	C	2126.8	4	1
	C			
	C	1796.0	4	5
	C			
D	C	1681.0	4	6
D	C			
D	C	1316.0	4	7

T tests (LSD) for variable: YIELD

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.01 df= 21 M

T	Grouping	Mean	N	TRT
	A	2678.0	4	2
	A			
	A	2551.8	4	3
	A			
B	A	2128.0	4	4
B	A			
B	A	2126.8	4	1
B				
B	C	1796.0	4	5
B	C			
B	C	1681.0	4	6
	C			
	C	1316.0	4	7

Đọc và đánh giá kết quả so sánh

T tests (LSD) for variable: YIELD: Kiểm định LSD cho biến YIELD (năng suất)

Alpha= 0.01: mức 0,01

Least Significant Difference= 616.34: LSD= 616.34

T Grouping: Nhóm

Mean: Trung bình

N: Số lượng, ví dụ có 4 số của công thức 2 (tương ứng với 4 lần nhắc lại)

TRT: Công thức

T Grouping	Mean	N	TRT
A	2678.0	4	2

Ta hiểu là: công thức 2 có 4 lần nhắc lại trung bình là 2678 và thuộc nhóm A.

Theo kết quả trên thì $LSD_{0.05}$ là 452,7 và $LSD_{.01}$ là 616,34. Giá trị trung bình của các công thức được xắp xếp từ lớn tới nhỏ và được xếp hạng theo thứ tự kí tự (A, B...). Những giá trị trung bình nào có ít nhất một kí tự giống nhau thì sự khác biệt giữa chúng không có ý nghĩa.

So sánh ở mức 0,01 ta thấy:

Công thức 2, 3, 4 và 1 có năng suất tương đương nhau vì ở cùng nhóm A

Công thức 4, 1, 5, 6 có năng suất tương đương nhau vì cùng ở nhóm B

Công thức 5, 6 và 7 có năng suất tương đương nhau vì ở cúng nhóm C.

Tương tự như vậy ta kết luận cho so sánh ở mức 0,05.

6.3.1.2. So sánh DUNCAN (DMRT)

Tương tự như so sánh dùng LSD, muốn sử dụng so sánh DUNCAN (DMRT) ta làm như trên, chỉ khác câu lệnh. Thay vì dùng LSD thì câu lệnh tương ứng sẽ là ***means trt/DUNCAN.***

Với ví dụ trên, ta viết chương trình so sánh phân hạng DUNCAN ở mức α bằng 0,05 và 0,01 như sau:

```
Options nodates ps=76 ls=80;
Data crd;
Infile 'a:\crd.prn' firstobs=2;
Input rep trt yield;
Proc anova;
Class trt;
```

```
Model yield=trt;
Means trt/DUNCAN;
Means trt/DUNCAN alpha=0.01;
Run;
```

Kết quả so sánh dùng DMRT:

The SAS System

Analysis of Variance Procedure

Duncan's Multiple Range Test for variable: N

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha= 0.05 df= 21 MSE= 94773.21

Number of Means 2 3 4 5 6 7

Critical Range 452.7 475.3 489.7 499.7 507.2 512.9

Means with the same letter are not significantly different.

Duncan	Grouping	Mean	N	TRT
	A	2678.0	4	2
	A			
B	A	2551.8	4	3
B				
B	C	2128.0	4	5
B	C			
B	C	2126.8	4	1
	C			
	C	1796.0	4	5
	C			
D	C	1681.0	4	6
D	C			
D	C	1316.0	4	7

The SAS System

Analysis of Variance Procedure

Duncan's Multiple Range Test for variable: YIELD

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha= 0.01 df= 21 MSE= 94773.21

Number of Means	2	3	4	5	6	7
Critical Range	616.3	642.9	660.4	673.3	683.1	691.0

Means with the same letter are not significantly different.

Duncan	Grouping	Mean	N	TRT
	A	2678.0	4	2
	A			
	A	2551.8	4	3
	A			
B	A	2128.0	4	4
B	A			
B	A	2126.8	4	1
B				
B	C	1796.0	4	5
B	C			
B	C	1681.0	4	6
	C			
	C	1316.0	4	7

Như vậy kết quả so sánh DMRT trong trường hợp này cũng tương tự như so sánh dùng LSD.

Cách đọc kết quả giống như so sánh cặp đôi.

6.3.1.3. So sánh cặp đôi cho thí nghiệm đa nhân tố

Đối với thí nghiệm đa nhân tố, trước tiên phải xem xét tương tương tác giữa các nhân tố. Nếu tương tác giữa các nhân tố có ý nghĩa (tức là ảnh hưởng của một yếu tố sẽ thay đổi theo các mức độ của yếu tố khác) thì việc so sánh đánh giá phải dựa trên ảnh hưởng đơn(simple effect) của từng nhân tố thí nghiệm. Ngược lại, khi tương tác không có ý nghĩa thì việc so sánh đánh giá phải dựa trên ảnh hưởng chung (main effect) của từng nhân tố thí nghiệm.

- *So sánh đánh giá dựa trên ảnh hưởng chung*

Cách thực hiện so sánh tương tự như đối với thí nghiệm 1 nhân tố .

Minh họa bằng thực hiện so sánh giữa các giống lúa trong thí nghiệm 2 nhân tố kiểu RCBD gồm 3 giống và 5 mức bón đạm (mục 6.2.4.1). Kết quả phân tích ANOVA của thí nghiệm này cho thấy cả giống (var) và đạm (N) đều có ảnh hưởng chắc chắn đến năng suất nhưng tương tác giữa giống và đạm không có ý nghĩa ám chỉ rằng các giống phản ứng giống nhau với đạm. Vì vậy ta có thể so sánh các giống với nhau hay các mức đạm với nhau dựa trên ảnh hưởng chung. So sánh giữa các giống được thực hiện bằng sử dụng LSD hoặc DUNCAN vì giống là nhân tố không thể đo đếm được (non-quantitative factor). So sánh giữa các mức N có thể dùng LSD hoặc DUNCAN nhưng vì N là nhân tố có thể đo đếm được (quantitative factor) nên nếu đánh giá bằng phương pháp so sánh xu hướng (trend comparison) thì sẽ hợp lý hơn (xem chi tiết ở mục 4.3.2).

Câu lệnh thêm vào trong chương trình xử lý trên SAS để thực hiện so sánh giữa các giống là ***means var/LSD***;

Kết quả của câu lệnh thêm vào này như sau:

The SAS System

Analysis of Variance Procedure

T tests (LSD) for variable: YIELD

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 42 MSE= 0.151257

Critical Value of T= 2.02

Least Significant Difference= 0.2482

Means with the same letter are not significantly different.

T Grouping	Mean	N	VAR
A	5.0581	20	3
A			
A	5.0420	20	2
B	4.7694	20	1

- *So sánh đánh giá dựa trên ảnh hưởng đơn.*

SAS không tính được trực tiếp các LSD tương tác, tuy nhiên dựa vào các sai số thí nghiệm đã được tính trong SAS ta có thể tính toán dễ dàng các LSD tương tác này theo các công thức ở phần lý thuyết để thực hiện các phép so sánh hợp lý.

6.3.2. So sánh nhóm và so sánh theo xu hướng (Contrasts)

So sánh nhóm là thực hiện việc so sánh giữa các nhóm công thức với nhau theo các mục đích nghiên cứu cụ thể. Ví dụ so sánh giữa các nhóm giống cũ với các

nhóm giống mới, các công thức diệt cỏ bằng tay với các công thức diệt cỏ bằng thuốc hoá học...

So sánh theo xu hướng: Thường áp dụng cho các nhân tố có thể đo đếm được (quantitive factor) để tìm hiểu xu hướng biến động của hàm theo biến số. Ví dụ khi nghiên cứu ảnh hưởng của đạm đến năng suất người ta không chỉ muốn so sánh giữa các liều lượng bón đạm khác nhau mà còn muốn ước lượng được mức bón hợp lý nhất. Vì vậy tìm hiểu hàm tương quan giữa đạm và năng suất sẽ hợp lý hơn để đáp ứng mục đích này.

Cả hai loại so sánh trên đều được thực hiện trong SAS bằng câu lệnh CONTRASTS, câu lệnh này phải đứng sau câu lệnh MODEL trong PROC GLM .

Cấu trúc của câu lệnh CONTRAST bao gồm:

"***tên nhóm so sánh***": Cần được đặt trong dấu nháy.

Biến: Tên của biến cần thực hiện so sánh này, ví dụ công thức, các loại thuốc trừ sâu, các giống .... Biến phải có mặt trong câu lệnh MODEL.

Giá trị: Kê khai hệ số so sánh của phép so sánh đó.

Minh họa bằng thực hiện việc so sánh nhóm và so sánh theo xu hướng cho thí nghiệm nghiên cứu ảnh hưởng của lượng hạt giống lúa gieo tới năng suất lúa (mục 4.3.2).

- So sánh nhóm: So sánh xem có sự khác nhau hay không giữa nhóm các công thức có lượng hạt giống gieo ít hơn 100 kg/ha với nhóm các công thức có số lượng hạt giống gieo từ 100 kg trở lên.
- Thực hiện so sánh theo xu hướng giữa lượng hạt giống gieo (x) và năng suất lúa (y) để xác định xem mối tương quan giữa các biến đó là bậc 1, bậc 2 hay bậc 3.

Hệ số so sánh được xác định ở bảng sau 6.19.

Bảng 6.19. Hệ số so sánh nhóm và so sánh xu hướng

So sánh	Công thức (kg hạt/ha)					
	25	50	75	100	125	150
< 100kg so với >100 kg (C1)	-1	-1	-1	+1	+1	+1
Tuyến tính (linear)	-5	-3	-1	+1	+3	+5
Bậc 2 (Quadratic)	+5	-1	-4	-4	-1	+5
Bậc 3 (Cubic)	-5	+7	+4	-4	-7	+5

Để thực hiện được 4 contrast trên, phải thêm 4 câu lệnh contrast vào chương trình xử lý cho thí nghiệm này.

Chương trình xử lý SAS có so sánh nhóm và xu hướng được lập trên SAS như sau:

```
data rcbd;
infile 'C:\rcbd.prn' firstobs=2;
input rep trt yield;
proc GLM;
class rep trt;
model yield= rep trt ;
contrast "C1" trt -1 -1 -1 1 1 1;
contrast "linear" trt -5 -3 -1 1 3 5;
contrast "quadratic" trt 5 -1 -4 -4 -1 5;
contrast "cubic" trt -5 7 4 -4 -7 5;
run;
```

Kết quả của những câu lệnh CONTRAST thêm vào như sau:

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
C1	1	1024240.1667	1024240.1667	9.26	0.0082
linear	1	760034.8000	760034.8000	6.87	0.0192
quadratic	1	74404.7619	74404.7619	0.67	0.4249
cubic	1	113301.4222	113301.4222	1.02	0.3274

Đọc và đánh giá kết quả phân tích so sánh nhóm và so sánh theo xu hướng từ kết quả phân tích trên:

Contrast C1 có mức ý nghĩa < 0,01: Gieo lúa với lượng hạt giống từ 25-75 kg/ha cho năng suất cao hơn chắc chắn so với lượng hạt giống lúa gieo từ 100-150 kg/ha.

Trong phân tích theo xu hướng chỉ có thành phần truyến tính (linear) có mức ý nghĩa <0,05. Do vậy ta kết luận trong phạm vi gieo hạt từ 25 đến 150 kg/ha thì năng suất lúa giảm theo hàm tuyến tính bậc một với độ tin cậy 95%.

6.4. PHÂN TÍCH TƯƠNG QUAN VÀ HỒI QUY TUYẾN TÍNH

6.4.1. Trình thủ tục CORR

Trình thủ tục này tính toán hệ số tương quan (correlation) giữa các biến.

Cấu trúc chung của Proc CORR.

Proc CORR *lựa chọn*;

Một số lựa chọn sử dụng với Proc CORR.

Data kê khai tên bộ số liệu trong SAS cần phân tích tương quan;

Var kê khai các biến phụ thuộc;

With kê khai các biến độc lập cần tính hệ số tương quan với biến phụ thuộc.

Chú ý: Nếu không có lựa chọn WITH thì SAS sẽ tính hệ số tương quan của tất cả các biến lần lượt với nhau, còn nếu có lựa chọn này thì SAS chỉ tính tương quan của các biến độc lập (được kê khai trong lựa chọn WITH) đối với các biến phụ thuộc (được khai trong lựa chọn VAR).

6.4.2. Trình thủ tục REG

Trình này tính toán phương trình tương quan của biến độc lập với các các biến phụ thuộc.

Cấu trúc chung của Proc REG.

Proc REG;

Model biến phụ thuộc= biến độc lập;

run;

Ví dụ: Tính tương quan tuyến tính giữa năng suất (yield) với chiều cao cây (height) và số nhánh đẻ (tiller) của 8 giống lúa (phần 5.4). Số liệu thí nghiệm như sau:

Giống	Năng suất (kg/ha) (Y)	Chiều cao cây (cm) (X_1)	Đẻ nhánh (số nhánh đẻ/khóm) (X_2)
1	5755	110,5	14,5
2	5939	105,4	16
3	6010	118,1	14,6
4	6545	104,5	18,2
5	6730	93,6	15,4
6	6899	84,1	17,6
7	7862	77,8	17,9
8	6561	75,6	19,4

Số liệu được nhập trên EXCEL và lưu trữ trong file **C:\Mlinear.prn** như sau:

Var	Yield	Height	Tiller
	Y	X1	X2
1	5755	110,5	14,5
2	5939	105,4	16
3	6010	118,1	14,6
4	6545	104,5	18,2
5	6730	93,6	15,4
6	6899	84,1	17,6
7	7862	77,8	17,9
8	6561	75,6	19,4

Lập trình trên SAS như sau:

```
Data Mlinear;
infile 'C:\Mlinear1.prn' firstobs=3;
input Var yield height tiller;
Proc CORR;
   var yield;
   with height tiller;
Proc reg;
   var yield height tiller;
   model yield=height tiller;
   run;
```

Kết quả phân tích

Kết quả của Proc CORR

The SAS System

Correlation Analysis

2 'WITH' Variables: HEIGHT TILLER

1 'VAR' Variables: YIELD

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
HEIGHT	8	96.200000	15.828185	769.600000	75.600000	118.100000
TILLER	8	16.700000	1.821303	133.600000	14.500000	19.400000
YIELD	8	6561.250000	677.337593	52490	5755.000000	7862.000000

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 8

	YIELD
HEIGHT	-0.86871
	0.0051
TILLER	0.83498
	0.0099

Kết quả của Proc REG

The SAS System

Model: MODEL1

Dependent Variable: YIELD

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	2632048.1527	1316024.0764	11.356	**0.0138**
Error	5	579455.34729	115891.06946		
C Total	7	3211503.5			

Root MSE	340.42777	R-square	**0.8196**
Dep Mean	6561.25000	Adj R-sq	0.7474
C.V.	5.18846		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	**6335.596495**	942.9309576	2.153	0.0839
HEIGHT	1	**-23.748104**	12.89549159	-1.842	0.1249
TILLER	1	**150.312641**	112.06936787	1.341	0.2375

Đánh giá kết quả:

Từ kết quả phân tích trên, phương trình tương quan của năng suất (yield) với chiều cao cây (height) và số nhánh hữu hiệu (tiller) như sau:

Yield = 6336 -23.75*Height +150.31*Tiller

Hệ số tương quan R^2 = 0,82 và phương trình tương quan trên có ý nghĩa ở mức tin cậy 95%.

6.5. SỬ DỤNG EXCEL VẼ ĐỒ THỊ VÀ HÀM TƯƠNG QUAN

Kết quả thí nghiệm thường được trình bày theo hai dạng chính là bảng biểu và đồ thị. EXCEL là một trong các phần mềm hỗ trợ cho trình bày kết quả bằng đồ thị và phân tích tương quan tiện dụng nhất hiện nay, đặc biệt là trong phân tích tương quan một biến. Phần này giới thiệu phương pháp trình bày kết quả bằng đồ thị và phân tích tương quan trên phần mềm EXCEL.

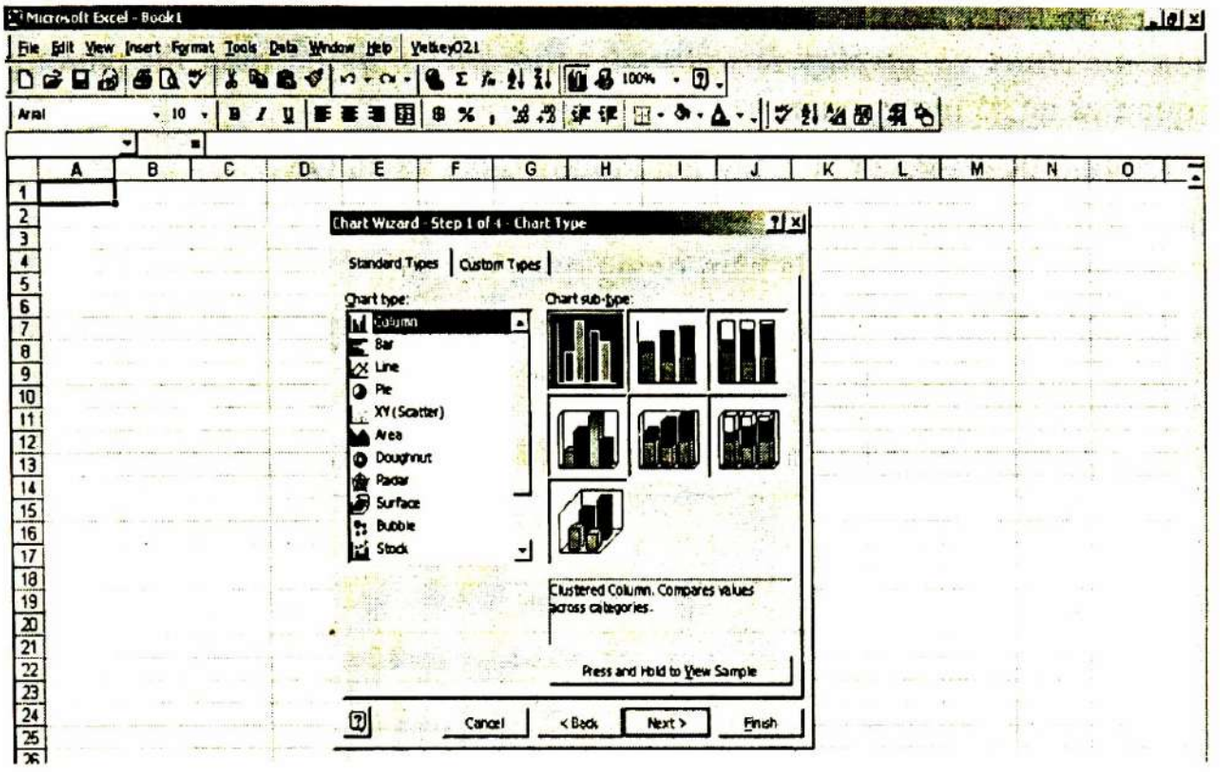
6.5.1. Sử dụng excel vẽ đồ thị

Bước 1. Nhập số liệu cần tạo đồ thị bao gồm các tiêu đề cột, dòng và số liệu.

Bước 2. Lựa chọn vùng dữ liệu cần vẽ đồ thị.

Bước 3. Nháy chuột vào biểu tượng ChartWizard trên Standard Toolbar hoặc chọn lệnh Chart trong thực đơn Insert. Hộp thoại Chart

Wizard xuất hiện trên màn hình.



Bước 4. Chọn kiểu đồ thị: Excel cung cấp 15 nhóm đồ thị, mỗi nhóm bao gồm nhiều kiểu phụ. Tuỳ theo nội dung muốn thể hiện mà lựa chọn kiểu đồ thị cho phù hợp. Nếu muốn kết hợp phân tích tương quan thì cần lựa chọn kiểu đồ thị **Scatter compairs pair of value** trong nhóm đồ thị **XY Scatter.**

Bước 5. Xác định dãy số liệu để tạo đồ thị:

Chọn **Data Range:** Trong bước này, excel yêu cầu kiểm tra lại vùng dữ liệu vẽ đồ thị. Thông thường không cần thực hiện bước này vì excel đã căn cứ vào vùng số liệu đã lựa chọn ở bước 2 để tự xác định.

Chọn kiểu **Series:** Xác định các bộ số liệu cần thể hiện trong đồ thị và khai báo số liệu ứng với các trục.

Bước 6. Khai báo tiêu đề của đồ thị và của các trục.

Bước 7. Xác định vị trí đặt đồ thị.

Bước 8. Nháy chuột vào phím **finish** để kết thúc việc tạo đồ thị.

6.5.2. Sử dụng excel vẽ hàm tương quan

Ưu điểm của dùng excel phân tích tương quan là excel vẽ đồ thị nhanh và kèm theo đồ thị còn cho biết giá trị R. Nhờ vậy ta yêu cầu excel vẽ tất cả các dạng đồ thị. Căn cứ vào giá trị của R để xác định hàm nào phù hợp nhất với số liệu của chúng ta. Các thủ tục để thực hiện phân tích tương quan như sau:

Bước 1. Thực hiện việc tạo đồ thị của bộ số liệu cần phân tích tương quan với dạng đồ thị là **Scatter compairs pair of value** trong nhóm đồ thị **XY Scatter.**

Bước 2. Đưa chuột vào vùng đồ thị, ấn chuột phải và chọn ADD TRENDLINE.

Trong hộp thoại Add trendline, trong mục OPTION đánh dấu vào 2 ô: ***display equation on chart và display R-squared value on chart*** để hiển thị phương trình tương quan và hệ số tương quan.

Trong mục TYPE chọn dạng hàm tương quan và ấn OK.

Bước 3. Lặp lại bước 2 với các dạng hàm tương quan khác nhau. Hàm tương quan có trị số của R^2 lớn nhất là hàm tương quan cần tìm.

Ví dụ: Vẽ đồ thị trình bày mối tương quan giữa đạm và năng suất của giống lúa A. Thí nghiệm có kết quả được trình bày ở bảng 6.20.

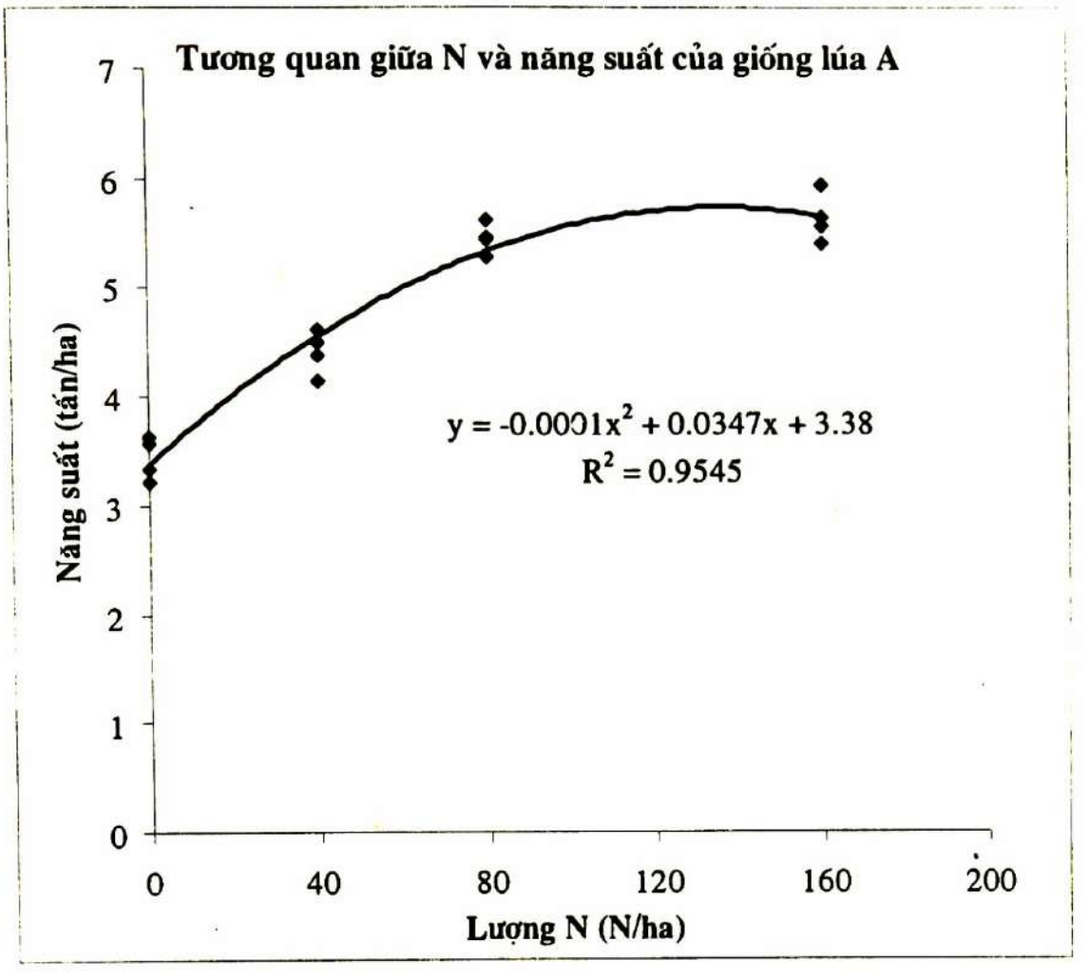
Bảng 6.20. Năng suất của giống lúa A và các mức bón đạm

Các mức đạm (kg/ha)	Năng suất (tấn/ha)			
	I	II	III	IV
0	3,22	3,63	3,34	3,56
40	4,15	4,50	4,38	4,61
80	5,62	5,44	5,47	5,28
160	5,93	5,57	5,64	5,41

Nhập số liệu trên EXCEL như sau:

Mức đạm	NS lúa
0	3.22
0	3.63
0	3.34
0	3.56
40	4.15
40	4.50
40	4.38
40	4.61
80	5.62
80	5.44
80	5.47
80	5.28
160	5.93
160	5.57
160	5.64
160	5.41

Các dấu ◊ là tọa độ của các điểm quan sát được ứng với giá trị của x và y. R^2 = 0,9545 cho thấy hàm $y = 0,0001x^2 + 0,0347x + 3,38$ phù hợp với những số liệu quan sát. Hay phản ánh tốt mối quan hệ giữa mức đạm và giống thí nghiệm.



Đồ thị biểu diễn ảnh hưởng của các mức đạm đến năng suất lúa

Chương 7
TRÌNH BÀY KẾT QUẢ NGHIÊN CỨU

Báo cáo kết quả nghiên cứu là một công việc hệ trọng, vì đây là cơ sở để đánh giá những cố gắng của nhà nghiên cứu, đồng thời đó là những bút tích để lại cho đồng nghiệp. Kết quả nghiên cứu có khả năng thuyết phục người quan tâm hay không phụ thuộc rất nhiều vào cách trình bày báo cáo. Một báo cáo được đánh giá tốt phải có kết cấu chặt chẽ, khoa học, ngắn gọn, số liệu chính xác và trình bày dễ hiểu. Trình bày kết quả nghiên cứu thể hiện qua cách trình bày số liệu, phân tích nhận xét số liệu, báo cáo trình bày và đánh giá kết quả nghiên cứu.

Chất lượng nghiên cứu được đánh giá theo những yêu cầu cơ bản sau đây: Có những kết luận và đề xuất mới mẻ có ý nghĩa về khoa học và thực tiễn. Tỏ rõ được năng lực nghiên cứu khoa học độc lập của người nghiên cứu. Thể hiện một kiến thức sâu rộng về lý thuyết trong lĩnh vực chuyên môn nghiên cứu và những vấn đề liên quan.

7.1. TRÌNH BÀY SỐ LIỆU

Việc trình bày kết quả nghiên cứu khoa học, hợp lý sẽ dễ làm cho việc nhận xét và đánh giá kết quả nghiên cứu. Nó còn làm cho người đọc dễ theo dõi và dễ thuyết phục được họ. Cách trình bày số liệu thường dùng là bảng số liệu và đồ thị.

7.1.1. Phương pháp lập bảng số liệu

Ưu điểm: Cho thấy sự khác nhau chính xác về lượng giữa các biến

Nhược điểm: Nhiều khi phức tạp

Yêu cầu: Đơn giản dễ hiểu, phản ánh được mục đích phân tích

Loại bảng:

Một chiều

Hai chiều

Ba chiều

Bốn chiều

Cấu tạo một bảng số liệu bao gồm:

- Số thứ tự bảng: Để người đọc dễ theo dõi, đánh số thứ tự bảng biểu hay đồ thị nên được đánh theo chương tương ứng. Ví dụ bảng thứ 3 của chương 4 nên đánh số là: Bảng 4.3.

- Tên bảng: Viết tóm tắt nhưng đầy đủ nội dung các số liệu trong bảng và ghi đơn vị nếu như đơn vị đo cho các số liệu trong bảng giống nhau.

- Thân bảng gồm các hàng ngang và cột dọc. Không nên có nhiều cột dọc (chỉ nên 5-6 cột). Nếu bảng có nhiều cột nên tách ra thành các bảng nhỏ như 3.4a; 3.4b.

- Phần nội dung trong bảng là phần ghi số liệu. Khi ghi chép chú ý: Thống nhất đơn vị và số lẻ trong cột. Số liệu phải viết thẳng hàng số lẻ và dấu ","

7.1.2. Phương pháp vẽ đồ thị

Mục đích của đồ thị để dễ nhận ra xu hướng chung mà nhìn vào bảng số liệu khó nhận thấy.

Ưu điểm: Tiện cho so sánh, đơn giản, bản chất của số liệu dễ thấy

Nhược điểm: Đồ thị không cho biết rõ về số lượng cụ thể.

Yêu cầu: Đồ thị cần rõ, chính xác.

Các loại đồ thị: Có nhiều loại đồ thị khác nhau như cột, đường, hình tròn, khối, hai chiều, ba chiều, tuỳ vào tính chất của số liệu để chọn loại đồ thị nào cho thích hợp. Có một số loại đồ thị hay dùng như sau:

- Đồ thị cột: Thường biểu thị diễn biến theo thời gian; so sánh

- Đồ thị đường: Thường biểu thị xu hướng và so sánh xu hướng giữa các công thức với nhau

- Đồ thị hình tròn: Thường biểu thị % của các phần trong một tổng thể.

Vẽ đồ thị cần chú ý:

Ghi rõ nội dung và đơn vị của trục hoành và trục tung.

Ghi rõ số thứ tự của đồ thị.

Tên đồ thị phải đơn giản, đủ ý và nên đặt ở dưới đồ thị.

7.1.3 . Trình bày kết quả phân tích thống kê

Chúng ta không thể đưa hết thông tin phân tích thống kê vào báo cáo vì vậy chúng ta chỉ chọn lọc những thông tin cơ bản như mức độ tin cậy của bảng số liệu, giá trị LSD (nếu so sánh số trung bình). Kết quả chi tiết nên trình bày trong phần phụ lục của báo cáo nghiên cứu thí nghiệm để tạo điều kiện cho độc giả tham khảo và đánh giá.

7.2. BỐ CỤC VÀ NỘI DUNG LUẬN VĂN KHOA HỌC

Đối với sinh viên đại học, báo cáo kết quả nghiên cứu của quá trình thực tập tốt nghiệp được gọi là "Luận văn khoa học". Luận văn khoa học là chuyên khảo về một vấn đề khoa học hoặc công nghệ do một người viết nhằm các mục đích: rèn luyện

phương pháp và kỹ năng nghiên cứu; thể nghiệm kết quả một giai đoạn học tập; Bảo vệ trước hội đồng chấm luận văn để đạt được văn bằng tốt nghiệp đại học (Phương Kỳ Sơn, 2001).

Hình thức và kết cấu của luận văn (Phương Kỳ Sơn, 2001)

Luận văn được trình bày trên giấy khổ A4, đánh máy một mặt. Soạn thảo trên winword với khổ chữ 13 hoặc 14. Sắp xếp kết cấu và bố cục có thể như sau:

1. Bìa: Gồm bìa chính và bìa phụ hoàn toàn giống nhau và được viết theo thứ tự từ trên xuống dưới như sau:

- Tên trường, khoa, bộ môn nơi hướng dẫn sinh viên làm luận văn.
- Tên đề tài, in bằng chữ lớn
- Tên tác giả
- Địa danh và tháng, năm bảo vệ luận án

2. Trang ghi lời cảm ơn: Trong trang này tác giả ghi lời cảm ơn đối với cơ quan đỡ đầu luận văn, cá nhân có nhiều công lao đối với công trình nghiên cứu và đã hướng dẫn, giúp đỡ tác giả.

3. Mục lục: Mục lục thường được đặt ở phía đầu luận văn, tiếp sau bìa phụ.

4. Ký hiệu và viết tắt: Liệt kê theo thứ tự vần chữ cái những ký hiệu và chữ viết tắt trong luận văn để người đọc tiện tra cứu.

5. Lời nói đầu: Lời nói đầu cho biết một cách vắn tắt lý do và bối cảnh của đề tài, ý nghĩa lý thuyết và thực tiễn của đề tài, kết quả đạt được và những vấn đề tồn tại, những dự kiến sau công trình nghiên cứu.

6. Đánh số cho các chương và mục: có nhiều cách đánh số khác nhau. Phương pháp đánh số dễ theo dõi có thể tham khảo phương pháp được sử dụng cho giáo trình này. Đó là cách dùng các số đếm thông thường, những phần nhỏ sau có kèm theo số thứ tự của mục lớn trên nó (xem phần mục lục).

7. Tóm tắt: Để người đọc với thời gian ngắn có thể nắm được nội dung của báo cáo. Yêu cầu: Không dài quá 300 từ (đối với bài đăng trên một số tạp chí quốc tế không vượt quá 200 từ hoặc 150 từ). Bài tóm tắt phải thể hiện được các nội dung sau: Vấn đề nghiên cứu; Mục tiêu nghiên cứu; Phương pháp áp dụng; Kết quả thu được; Những kết luận hay đề nghị

8. Phần nội dung: gồm các phần như sau

PHẦN 1: ĐẶT VẤN ĐỀ

1.1. Đặt vấn đề

- Báo cáo về vấn đề gì.
- Nêu sự cần thiết phải nghiên cứu vấn đề đó. Có nhiều lý do cho vấn đề nghiên cứu như: Chưa có tài liệu về vấn đề đó; muốn tiến hành phân tích sâu hơn, rộng hơn;

những kỹ thuật hiện tại kém hiệu quả. Nghiên cứu nhằm giải quyết vấn đề khoa học hay thực tiễn gì, giá trị khoa học và thực tiễn ra sao.

- Mục tiêu và mục đích nghiên cứu: Mục tiêu nghiên cứu là các vấn đề nào cần đặt ra trong nghiên cứu. Mục đích phải cụ thể, rõ ràng, có thể đạt được.

1.2. Nêu giả thuyết

Các giả thiết thường gắn bó chặt chẽ với mục tiêu nghiên cứu cần phải giải quyết. Giả thuyết trong nghiên cứu cần phải được đặt ra một cách cụ thể, rõ ràng và đơn giản. Cần có những luận cứ để có thể chứng minh là khẳng định hay bác bỏ. Cần liên hệ với cơ sở lý thuyết.

1.3. Tổng quan vấn đề nghiên cứu

Kết quả của tổng quan tài liệu là trình bày vấn đề đã được nghiên cứu ở trong và ngoài nước đến đâu. Dựa vào đề cương để viết, tránh liệt kê theo tác giả hay nghiên cứu. Khi viết tổng quan cho một nghiên cứu, điều quan trọng là phải phân tích và đánh giá một kết quả nghiên cứu trước đó.

Việc phân tích và đánh giá một nghiên cứu trước nên tuân theo trật tự và trả lời các câu hỏi sau: Nghiên cứu đó đề cập đến vấn đề gì?. Trước tiên đọc báo cáo về nghiên cứu đó, sau đó tóm tắt nội dung lại bằng một câu rõ ràng. Trả lời được câu hỏi này, nghiên cứu sẽ thành công trong bước đầu của sự thảo luận về nghiên cứu đó. Mục đích của nghiên cứu đó là gì? Bạn có thể trả lời một số câu hỏi sau nếu phù hợp:

Trong nghiên cứu đó có nhiều giả thuyết không? Nếu đúng vậy, những giả thuyết đó là gì? Và kết quả nghiên cứu đã khẳng định hay bác bỏ các giả thuyết. Nếu không có giả thuyết, vậy bài báo có dụng ý gì? Bài báo đưa ra kết luận gì? Số liệu nghiên cứu nào đã khẳng định kết luận đó? Ngoài ra có những nội dung nào mà kết quả nghiên cứu không khẳng định?

Các cơ sở lý thuyết hay phương pháp luận nào đã được dùng trong nghiên cứu này? Liệu ta có định nghĩa được chúng không? Liệu ta có thể trình bầy chúng trong một câu hay đoạn văn ngắn bằng lời văn của mình không? Ở nhiều nghiên cứu, nhiều khái niệm đã được trình bầy, đọc và tóm tắt lại bằng lời văn của chính mình ở phần giới thiệu.

Số liệu đã được thu thập bằng phương pháp nào? Hãy trả lời tất cả các câu hỏi sau: Mẫu được chọn như thế nào? Kỹ thuật nghiên cứu nào đã được dùng?. Những ưu điểm và nhược điểm của các phương pháp đó là gì?

Nghiên cứu đã đóng góp gì về chủ đề nghiên cứu và phương pháp luận? Hãy trả lời tất cả các câu hỏi sau: Thành công nhất ở chỗ nào về nội dung và về phương pháp. Điểm dừng lại và hạn chế của nghiên cứu đó là gì?

Cách trích dẫn tài liệu: Tài liệu tham khảo trích dẫn theo số thứ tự của tài liệu ở danh mục tài liệu tham khảo của luận văn. Số thứ tự đó được đặt trong ngoặc vuông. Trong trích dẫn phải đưa nguồn tài liệu bao gồm: Tác giả và và năm viết. Có 2 cách trích dẫn là trích nguyên văn và lấy ý như sau:

- Ghi nguyên văn phải đưa vào dấu "...". Ví dụ, "Năng suất chè của Thái Nguyên thấp. Lý do là giống chủ yếu được trồng hiện nay là giống trung du địa phương với tỷ lệ trên 70%" (Sở Nông Nghiệp Thái Nguyên, 2002).

- Lấy ý không cần đưa vào dấu "...". Ví dụ: theo Sở Nông nghiệp Thái Nguyên cơ cấu giống chè ở Thái Nguyên năm 2002 có 70% là giống trung du địa phương. Đó là lý do năng suất chè của Thái Nguyên còn thấp (Sở Nông Nghiệp Thái Nguyên, 2002).

Tùy từng trường hợp mà quyết định cách trích dẫn. Cách hay được dùng là lấy ý chứ không lấy nguyên văn.

Trích dẫn kèm theo nguồn gốc của tài liệu để tạo sự tin cậy của người đọc, tôn trọng tác giả và còn giúp người đọc có thể tìm đọc nguyên gốc nếu như họ muốn tìm hiểu sâu thêm. Hơn nữa khi trích dẫn tài liệu của những nhà khoa học có tiếng, những tài liệu có giá trị sẽ tăng việc thuyết phục người đọc và chứng tỏ sự hiểu biết rộng trong tổng quan vấn đề nghiên cứu của tác giả.

Tổng quan tài liệu trong và ngoài nước trên các khía cạnh sau:

- Tình hình, tầm quan trọng của vấn đề nghiên cứu.
- Kết quả đã đạt được liên quan đến vấn đề nghiên cứu.
- Những tranh luận liên quan đến vấn đề nghiên cứu.
- Những định hướng triển vọng trong tương lai
- Những phương pháp nghiên cứu hay thí nghiệm đã được áp dụng.

PHẦN 2. PHƯƠNG PHÁP NGHIÊN CỨU

Phần này mô tả phương pháp nghiên cứu đã được sử dụng. Đối với một nghiên cứu thí nghiệm đồng ruộng, trình bày phương pháp nghiên cứu gồm các phần sau:

2.1. Thời gian và địa điểm nghiên cứu

2.2. Đối tượng nghiên cứu

Số công thức, lần nhắc lại, kích thước ô thí nghiệm

Kiểu bố trí thí nghiệm

Điều kiện thí nghiệm: Đặc điểm đất đai, chế độ nước, phân bón, chăm sóc....

2.3. Nội dung và phương pháp nghiên cứu

Hệ thống chỉ tiêu nghiên cứu

Phương pháp thu thập số liệu và quan trắc. Số liệu thu thập bao gồm số liệu đã công bố gọi là số liệu thứ cấp; số liệu phải đo đếm quan sát mới có gọi là số liệu sơ cấp. Với số liệu sơ cấp phải nêu rõ cách quan sát, thu thập.

Phương pháp phân tích

Phương pháp phân tích thống kê

2.4. Giới hạn phạm vi nghiên cứu

PHẦN 3: KẾT QUẢ VÀ THẢO LUẬN

Phần này cần đưa ra các số liệu về kết quả nghiên cứu kèm theo những phân tích về các số liệu đó.

3.1. Trình bày kết quả nghiên cứu

Gồm những kết quả đạt được trong quá trình nghiên cứu. Kết quả được trình bày qua các dạng văn viết tả cảnh, bảng biểu, đồ thị hay họa đồ v.v..

Tuỳ theo mục đích nghiên cứu, loại thông tin mà có thể chọn phương pháp phân tích này hay phương pháp phân tích khác. Thường có các loại phương pháp phân tích sau:

Phương pháp thống kê

Mục đích: Phương pháp này dùng để mô tả. Gồm có thống kê mô tả và thống kê so sánh. Các chỉ tiêu phân tích thống kê đơn giản như số trung bình, tần suất, số tối đa, tối thiểu, số tuyệt đối, số tương đối dùng để so sánh theo thời gian, không gian.

Phương pháp ước lượng mô hình (hàm tương quan)

Mục đích: giúp cho người nghiên cứu ước lược được mối quan hệ giữa các biến. Khi dùng hàm tương quan cần phải xác định rõ đâu là biến độc lập và đâu là biến phụ thuộc. Sau đó xác định dạng của mô hình tương quan. Xác định các phần mềm để ước lượng các mô hình.

3.2. Thảo luận

Thảo luận kết quả đạt được qua việc so sánh với mục tiêu, yêu cầu đặt ra và so sánh với các kết quả nghiên cứu khác trước đó (có trích dẫn để chứng minh). Khái quát hoá có thể nâng lên tầm lý luận hay xây dựng giả thuyết mới. Thảo luận cũng có thể đưa ra khuyến cáo áp dụng kết quả nghiên cứu. Phát hiện những vấn đề chưa được giải quyết làm định hướng cho nghiên cứu tiếp theo.

Chú ý: Sự sinh trưởng phát triển của cây trồng, đối tượng nghiên cứu của thí nghiệm đồng ruộng còn phụ thuộc vào nền thí nghiệm. Nên khi đánh giá kết quả thí nghiệm hay đưa ra kết luận phải luôn gắn với điều kiện cụ thể nơi đã làm thí nghiệm.

PHẦN 4. KẾT LUẬN VÀ ĐỀ NGHỊ

Tóm tắt những nhận xét và kết quả của thí nghiệm và những đề xuất với sản xuất và nghiên cứu tiếp theo.

Kết luận phải dựa trên mục tiêu của nghiên cứu đã đề ra. Liệu kết quả tìm được có khẳng định hay bác bỏ giả thuyết nghiên cứu. Qua nghiên cứu tác giả có thể có một số đề nghị liên quan đến vấn đề nghiên cứu.

9. Danh mục tài liệu tham khảo

Tài liệu tham khảo gồm những sách, ấn phẩm, tạp chí đã đọc và được trích dẫn hoặc được sử dụng ý tưởng trong luận văn. Những tham khảo đó phải được chỉ rõ nguồn tài liệu trong luận văn.

Bảng chỉ dẫn tài liệu tham khảo được đặt ở cuối của báo cáo, trước phần phụ lục. Đối với một chuyên gia có kinh nghiệm thường chỉ cần qua sự tìm hiểu tài liệu tham

khảo cũng đủ để đánh giá đề tài, để biết tác giả luận án có biết sử dụng những nguồn tài liệu quan trọng hay không. Danh mục tài liệu tham khảo không phải là một yêu cầu hình thức mà là một bộ phận của bản luận văn.

Các tài liệu tham khảo được sắp xếp riêng theo ngôn ngữ (Việt, Anh...). Tên tài liệu nước ngoài giữ nguyên văn không dịch. Đối với tài liệu trong nước xếp thứ tự ABC theo tên tác giả. Đối với tài liệu nước ngoài xếp thứ tự ABC theo họ (chữ đầu tiên của tên).

Có nhiều cách ghi thông tin của danh mục tài liệu tài liệu tham khảo khác nhau, theo hệ thống thư viện Mỹ (The Harvard System) trình bày như sau: Viết tên tác giả, năm xuất bản, tên tài liệu, nhà xuất bản, nơi xuất bản, trang tham khảo.

Ví dụ:

Thomas M. Little and F. Jackson Hills, 1978. Agricultural Experimental: Design and Analysis. John Wiley & Sons Inc. USA. p. 103-105.

Phạm Chí Thành, 1988. Giáo trình phương pháp thí nghiệm đồng ruộng. Nhà xuất bản Nông nghiệp, Hà Nội. p. 35

10. Phụ lục

Gồm các thông tin, bảng biểu, mẫu câu hỏi hay tranh ảnh... để minh họa hay chứng minh kết quả trong báo cáo.

7.3. BẢO VỆ LUẬN VĂN

Việc bảo vệ luận văn bao gồm mấy giai đoạn chủ yếu sau: chuẩn bị bài trình bày về đề tài nghiên cứu. Chuẩn bị tài liệu minh họa cho báo cáo. Chuẩn bị những câu trả lời căn cứ theo tinh thần nhận xét của những người nhận xét, của những người trong và ngoài hội đồng bảo vệ luận văn.

Bài trình bày nên chuẩn bị trên giấy bóng kính. Mỗi trang giấy bóng kính nên trình bày một vấn đề. Trình bày trên giấy bóng kính càng ít chữ càng tốt, để người nghe dễ theo dõi và nắm được. Vì vậy chọn những từ chính, quan trọng để viết trên giấy bóng kính.

Chuẩn bị những câu hỏi sau, trước khi trình bày: Trình bày nhằm mục đích gì? Ai sẽ nghe? Họ có phải là chuyên gia về lĩnh vực mình báo cáo không? Có những phương tiện gì phục vụ cho trình bày?

7.3.1. Đề cương báo cáo bảo vệ luân văn

Đề cương báo cáo là tóm tắt những kết quả cơ bản của đề tài nghiên cứu. Tóm tắt luận văn được trình bày như sau:

7.3.1.1. Phần mở đầu: viết rất ngắn gọn và súc tích các mục sau

- Tính cấp thiết của đề tài
- Mục đích nghiên cứu
- Đối tượng, phạm vi
- Phương pháp và nội dung nghiên cứu

7.3.1.2. Phần kết quả nghiên cứu

Nêu tóm tắt những kết quả đã thu được kèm theo nhận xét và đánh giá kết quả. Chú ý nên chọn cách trình bày dễ hiểu, dễ cho người nghe theo dõi.

7.3.1.3. Phần kết luận

Nêu những kết luận quan trọng nhất của đề tài nghiên cứu

Ý nghĩa quan trọng của đề tài nghiên cứu

Khuyến nghị quan trọng từ kết quả nghiên cứu

7.3.2. Thuyết trình báo cáo khoa học

7.3.2.1. Làm chủ thời gian

Thời gian trung bình cho mỗi báo cáo là 15- 20 phút. Vì vậy phần đặt vấn đề không quá 4 phút. Nên giành 3 phút cho phương pháp nghiên cứu, 11 phút cho kết quả và thảo luận và 2 phút cho kết luận.

Bạn cần làm chủ thời gian sao cho người nghe nắm được toàn bộ kết quả nghiên cứu và theo dõi cho chuẩn bị thảo luận.

7.3.2.3. Cách nói

Nói đơn giản dễ hiểu và bình tĩnh. Không nên đọc tất cả các từ đã viết trên giấy bóng kính. Không nên lặp ý. Nên đặt ra các câu hỏi để hướng sự chú ý của người nghe.

Nói to, rõ, rành mạch, nói đơn giản và diễn cảm. Trong giọng nói biểu lộ rõ sự tin tưởng, có sức thuyết phục.

Kết thúc, cám ơn sự lắng nghe của hội đồng và đề nghị những ý kiến đóng góp.

7.3.2.4. Thảo luận, trả lời câu hỏi của hội đồng

Lắng nghe các câu hỏi và trả lời từng vấn đề. Nếu không rõ câu hỏi nên hỏi lại.

TÀI LIỆU THAM KHẢO

1. Đỗ Kim Chung, 2003. *Phương pháp nghiên cứu kinh tế - xã hội phát triển nông thôn "* Chương trình nghiên cứu Việt Nam - Hà Lan. Trong "Cơ sở lý thuyết và thực tiễn phát triển nông thôn bền vững" - Tập bài giảng bồi dưỡng kiến thức của VNRP - Nhà xuất bản Nông Nghiệp
2. Nguyễn Tri Khiêm, 2003. *Một số phương pháp thu thập và phân tích số liệu.* Trong Chương trình nghiên cứu Việt Nam - Hà Lan "Cơ sở lý thuyết và thực tiễn phát triển nông thôn bền vững" - Tập bài giảng bồi dưỡng kiến thức của VNRP - Nhà xuất bản Nông Nghiệp.
3. Phạm Chí Thành, 1976. *Giáo trình phương pháp thí nghiệm đồng ruộng.* Nhà xuất bản Nông nghiệp
4. Phạm Chí Thành, 1988. *Giáo trình phương pháp thí nghiệm đồng ruộng.* Nhà xuất bản Nông nghiệp
5. Cao Hào Thi và Lê Chiến Thắng, 2000. *Thống kê* Chương trình Fulbright Việt Nam, 2003. Chương trình đào tạo kinh tế Fulbright
6. Phương Kỳ Sơn, 2001. *Phương pháp nghiên cứu khoa học.* Nhà xuất bản chính trị Quốc gia
7. Đỗ Văn Xê, 1994. *Giới thiệu về kinh tế lượng.* Đại học Cần Thơ
8. Nguyễn Thị Út, 2003. *Phương pháp xử lý số liệu trong nghiên cứu các vấn đề kinh tế - xã hội nông thôn.* Trong Chương trình nghiên cứu Việt Nam - Hà Lan "Cơ sở lý thuyết và thực tiễn phát triển nông thôn bền vững" - Tập bài giảng bồi dưỡng kiến thức của VNRP - Nhà xuất bản Nông nghiệp
9. Baker R.J., 2000. *Statistical Method in Life Sciences.* University of Saskatchewan, Canada
10. Kwanchai A. Gomez & Arturo A. Gomez, 1976. *Statistical Procedures for Agricultural Research.* A Wiley-Interscience New York
11. Little M. Thomas và Hills F. Jackson, 1978. *Agricultural Experimentation Design and Analysis.* John Wiley & Sons Inc., Canada.
12. Miller, 1996. *Sustaining the Earth* - Wordsworth Publication - USA
13. Steel, R. G. D., J. H. Torrie and D. A. Dickey, 1997. *Principle and procedures of Statistics, A biometrical Approach.* The McGraw - Hill Companies, Inc. USA.

PHỤ LỤC: CÁC BẢNG SỐ

Các bảng số dùng cho phân tích thống kê được trích từ Gomez, 1984. Riêng bảng A4, A5, A11, Fa và Fb trích từ Little M. Thomas và Hills F Jackson, 1978.

Bảng A: Bảng số ngẫu nhiên

14620	95430	12951	81953	17629
09724	85125	48477	42783	70473
56919	17803	95781	85069	61594
97310	78209	51263	52396	82680
07585	28040	26939	64531	70570
25950	85189	69374	37904	06759
82937	16405	81497	20863	94072
60819	27364	59081	72635	49180
59041	38475	03615	84093	49731
74208	69516	79530	47649	53046
39412	03642	87497	29735	14308
48480	50075	11804	24956	72182
95318	28749	49512	35408	21814
72094	16385	90285	72635	86259
63158	49753	84279	56496	30618
19082	73645	09182	73649	56823
15232	84146	87729	65584	83641
94252	77489	62434	20965	20247
72020	18895	84948	53072	74573
48392	06359	47040	05695	79799
37950	77387	35495	48192	84518
09394	59842	39573	51630	78548
34800	28055	91570	99154	39603
36435	75946	85712	06293	85621
28187	31824	52265	80494	66428

Bảng A4. Hệ số R dùng trong so sánh Duncan ở mức 5% ứng với các giá trị p và n khác nhau (n = độ tự do của ngẫu nhiên)

	p:															
n	2	3	4	5	6	7	8	9	10	12	14	16	18	20	50	100
4	1.00	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
5	1.00	1.03	1.04	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05
6	1.00	1.03	1.05	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06
7	1.00	1.04	1.06	1.07	1.07	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08
8	1.00	1.04	1.06	1.08	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09
9	1.00	1.04	1.07	1.08	1.09	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10
10	1.00	1.05	1.07	1.09	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10
11	1.00	1.05	1.08	1.09	1.10	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.12	1.12	1.12	1.12
12	1.00	1.05	1.08	1.09	1.10	1.11	1.12	1.12	1.12	1.12	1.12	1.12	1.13	1.13	1.13	1.13
13	1.00	1.05	1.08	1.09	1.10	1.11	1.12	1.12	1.13	1.13	1.13	1.13	1.13	1.13	1.13	1.13
14	1.00	1.05	1.08	1.10	1.11	1.12	1.13	1.13	1.14	1.14	1.14	1.14	1.15	1.15	1.15	1.15
15	1.00	1.05	1.08	1.10	1.12	1.12	1.13	1.14	1.14	1.14	1.15	1.15	1.15	1.15	1.15	1.15
16	1.00	1.05	1.08	1.10	1.12	1.12	1.13	1.14	1.14	1.15	1.15	1.16	1.16	1.16	1.16	1.16
17	1.00	1.05	1.08	1.10	1.12	1.13	1.13	1.14	1.15	1.15	1.16	1.16	1.16	1.16	1.16	1.16
18	1.00	1.05	1.08	1.10	1.12	1.13	1.13	1.14	1.15	1.15	1.16	1.16	1.17	1.17	1.17	1.17
19	1.00	1.05	1.08	1.10	1.12	1.13	1.14	1.15	1.15	1.16	1.16	1.17	1.17	1.17	1.17	1.17
20	1.00	1.05	1.08	1.10	1.12	1.13	1.14	1.15	1.15	1.16	1.17	1.17	1.17	1.18	1.18	1.18
22	1.00	1.05	1.08	1.10	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.18	1.18	1.18	1.18
24	1.00	1.05	1.08	1.10	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.18	1.19	1.19	1.19
26	1.00	1.05	1.08	1.10	1.12	1.13	1.15	1.15	1.16	1.17	1.18	1.19	1.19	1.19	1.19	1.19
28	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.16	1.17	1.18	1.19	1.19	1.20	1.20	1.20
30	1.00	1.05	1.08	1.11	1.12	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.20	1.20	1.20	1.20
40	1.00	1.05	1.08	1.11	1.13	1.14	1.15	1.16	1.17	1.19	1.20	1.20	1.21	1.21	1.21	1.21
60	1.00	1.05	1.09	1.11	1.13	1.14	1.16	1.17	1.18	1.19	1.20	1.21	1.22	1.23	1.23	1.23
100	1.00	1.05	1.09	1.11	1.14	1.15	1.16	1.18	1.19	1.20	1.21	1.22	1.23	1.24	1.26	1.26
∞	1.00	1.05	1.09	1.12	1.14	1.15	1.17	1.18	1.19	1.21	1.22	1.23	1.24	1.25	1.30	1.32

Bảng A5. Hệ số R dùng trong so sánh Duncan ở mức 1% ứng với các giá trị p và n khác nhau (n = độ tự do của ngẫu nhiên)

n	p:															
	2	3	4	5	6	7	8	9	10	12	14	16	18	20	50	100
3	1.00	1.03	1.04	1.05	1.07	1.08	1.08	1.09	1.09	1.09	1.10	1.11	1.13	1.13	1.13	1.13
4	1.00	1.04	1.06	1.08	1.09	1.09	1.11	1.11	1.12	1.12	1.14	1.14	1.15	1.15	1.15	1.15
5	1.00	1.05	1.07	1.08	1.10	1.11	1.12	1.13	1.14	1.16	1.16	1.18	1.18	1.19	1.19	1.19
6	1.00	1.05	1.08	1.09	1.11	1.12	1.14	1.15	1.15	1.16	1.18	1.18	1.20	1.20	1.20	1.20
7	1.00	1.05	1.08	1.10	1.12	1.13	1.15	1.16	1.17	1.17	1.19	1.19	1.21	1.21	1.21	1.21
8	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.17	1.18	1.20	1.20	1.22	1.22	1.22	1.23
9	1.00	1.06	1.08	1.10	1.12	1.14	1.16	1.17	1.17	1.20	1.20	1.22	1.24	1.24	1.24	1.24
10	1.00	1.06	1.09	1.11	1.13	1.15	1.16	1.17	1.18	1.20	1.21	1.22	1.24	1.24	1.24	1.24
11	1.00	1.05	1.09	1.11	1.13	1.14	1.15	1.17	1.17	1.19	1.20	1.22	1.23	1.23	1.23	1.23
12	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.17	1.19	1.20	1.21	1.21	1.22	1.22	1.22
13	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.21	1.21	1.21
14	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.20	1.20	1.20	1.20
15	1.00	1.05	1.08	1.10	1.11	1.13	1.14	1.15	1.16	1.18	1.18	1.19	1.20	1.20	1.20	1.20
16	1.00	1.05	1.08	1.10	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.20	1.20	1.20
17	1.00	1.05	1.08	1.10	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.19	1.19	1.19
18	1.00	1.05	1.08	1.10	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.19	1.19	1.19	1.19
19	1.00	1.05	1.07	1.09	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.19	1.19	1.19	1.19
20	1.00	1.05	1.07	1.09	1.11	1.13	1.14	1.15	1.15	1.17	1.18	1.18	1.19	1.19	1.19	1.19
22	1.00	1.05	1.07	1.09	1.11	1.12	1.14	1.15	1.15	1.17	1.17	1.18	1.19	1.19	1.19	1.19
24	1.00	1.05	1.07	1.09	1.11	1.12	1.13	1.15	1.15	1.17	1.17	1.18	1.19	1.19	1.20	1.20
26	1.00	1.05	1.07	1.09	1.11	1.12	1.13	1.15	1.15	1.17	1.18	1.18	1.19	1.19	1.20	1.20
28	1.00	1.04	1.07	1.09	1.11	1.12	1.13	1.14	1.15	1.17	1.18	1.18	1.19	1.19	1.21	1.21
30	1.00	1.04	1.07	1.09	1.11	1.12	1.13	1.14	1.15	1.17	1.18	1.19	1.19	1.20	1.21	1.21
40	1.00	1.04	1.07	1.09	1.11	1.12	1.14	1.14	1.15	1.17	1.18	1.19	1.20	1.20	1.23	1.23
60	1.00	1.04	1.07	1.09	1.11	1.12	1.14	1.15	1.15	1.17	1.18	1.19	1.20	1.20	1.24	1.24
100	1.00	1.04	1.07	1.09	1.11	1.12	1.14	1.15	1.15	1.17	1.18	1.19	1.20	1.21	1.25	1.25
∞	1.00	1.04	1.07	1.09	1.11	1.12	1.14	1.15	1.15	1.17	1.18	1.19	1.20	1.21	1.26	1.29

Bảng A11: Hệ số so sánh để tách nguồn biến động từ phương trình bậc 1 tới phương trình bậc 4 cho thí nghiệm có nhân tố thí nghiệm biến động đều

(Coefficients for fitting up to quartic curve to equally spaced data, and partitioning the sum of squares)

N = 3	C1	-1	0	1							
	C2	1	-2	1							
N= 4	C1	-3	-1	1	3						
	C2	1	-1	-1	1						
	C3	-1	3	-3	1						
N = 5	C1	-2	-1	0	1	2					
	C2	2	-1	-2	-1	2					
	C3	-1	2	0	-2	1					
	C4	1	-4	6	-4	1					
N = 6	C1	-5	-3	-1	1	3	5				
	C2	5	-1	-4	-4	-1	5				
	C3	-5	7	4	-4	-7	5				
	C4	1	-3	2	2	-3	1				
N = 7	C1	-3	-2	-1	0	1	2	3			
	C2	5	0	-3	-4	-3	0	5			
	C3	-1	1	1	0	-1	-1	1			
	C4	3	-7	1	6	1	-7	3			
N = 8	C1	-7	-5	-3	-1	1	3	5	7		
	C2	7	1	-3	-5	-5	-3	1	7		
	C3	-7	5	7	3	-3	-7	-5	7		
	C4	7	-13	-3	9	9	-3	-13	7		
N = 9	C1	-4	-3	-2	-1	0	1	2	3	4	
	C2	28	7	-8	-17	-20	-17	-8	7	28	
	C3	-14	7	13	9	0	-9	-13	-7	14	
	C4	14	-21	-11	9	18	9	-11	-21	14	
N = 10	C1	-9	-7	-5	-3	-1	1	3	5	7	9
	C2	6	2	-1	-3	-4	-4	-3	-1	2	6
	C3	-42	14	35	31	12	-12	-31	-35	-14	42
	C4	18	-22	-17	3	18	18	3	-17	-22	18

Bảng C: Xác xuất phân bố t - Distribution of t probability

n	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.5	0.02	0.01	0.001
1	0.158	0.325	0.510	0.272	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	0.142	0.289	0.445	0.617	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	0.137	0.277	0.424	0.584	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.134	0.271	0.414	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.132	0.267	0.408	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.131	0.265	0.404	0.553	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.130	0.263	0.402	0.549	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.130	0.262	0.399	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.129	0.261	0.398	0.543	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.129	0.260	0.397	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.129	0.260	0.396	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.128	0.259	0.395	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.128	0.259	0.394	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.128	0.258	0.393	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.128	0.258	0.393	0.536	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.128	0.258	0.392	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.128	0.257	0.392	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.127	0.257	0.392	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.127	0.257	0.391	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.127	0.257	0.391	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.127	0.257	0.391	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.127	0.256	0.390	0.532	0.686	0.858	1.061	1.321	1.117	2.074	2.508	2.819	3.792
23	0.127	0.256	0.390	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	0.127	0.256	0.390	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.127	0.256	0.390	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.850	2.787	3.725
26	0.127	0.256	0.390	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.790	2.779	3.707
27	0.127	0.256	0.389	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.127	0.256	0.389	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.127	0.256	0.389	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.127	0.256	0.389	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.126	0.255	0.388	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	0.126	0.254	0.387	0.527	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460
120	0.126	0.254	0.386	0.526	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	0.126	0.253	0.385	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

Bảng E: Phân bố F (5% và 1% in đậm) - Point for the Distribution of F

f_2	f_1																								f_2
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	254	254	254	1
	4052	**4999**	**5403**	**5625**	**5764**	**5859**	**5928**	**5981**	**6022**	**6056**	**6082**	**6106**	**6142**	**6169**	**6208**	**6234**	**6261**	**6286**	**6302**	**6323**	**6334**	**6352**	**6361**	**6366**	
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50	19.50	2
	98.49	**99.00**	**99.17**	**99.25**	**99.30**	**99.33**	**99.36**	**99.37**	**99.39**	**99.40**	**99.41**	**99.42**	**99.43**	**99.44**	**99.45**	**99.46**	**99.47**	**99.48**	**99.48**	**99.49**	**99.49**	**99.49**	**99.50**	**99.50**	
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.54	8.53	3
	34.12	**30.82**	**29.46**	**28.71**	**28.24**	**27.91**	**27.67**	**27.49**	**27.34**	**27.23**	**27.13**	**27.05**	**26.92**	**26.83**	**26.69**	**26.60**	**26.50**	**26.41**	**26.35**	**26.27**	**26.23**	**26.18**	**26.14**	**26.12**	
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.64	5.63	4
	21.20	**18.00**	**16.96**	**15.98**	**15.52**	**15.21**	**14.98**	**14.80**	**14.66**	**14.54**	**14.45**	**14.37**	**14.24**	**14.15**	**14.02**	**13.93**	**13.83**	**13.74**	**13.69**	**13.61**	**13.57**	**13.52**	**13.48**	**13.46**	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37	4.36	5
	16.26	**13.27**	**12.06**	**11.39**	**10.97**	**10.67**	**10.45**	**10.29**	**10.15**	**10.05**	**9.96**	**9.89**	**9.77**	**9.68**	**9.55**	**9.47**	**9.38**	**9.29**	**9.24**	**9.17**	**9.13**	**9.07**	**9.04**	**9.02**	
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67	6
	13.74	**10.92**	**9.78**	**9.15**	**8.75**	**8.47**	**8.26**	**8.10**	**7.98**	**7.87**	**7.79**	**7.72**	**7.60**	**7.52**	**7.39**	**7.31**	**7.23**	**7.14**	**7.09**	**7.02**	**6.99**	**6.94**	**6.90**	**6.88**	
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23	7
	12.25	**9.55**	**8.45**	**7.85**	**7.46**	**7.19**	**7.00**	**6.84**	**6.71**	**6.62**	**6.54**	**6.47**	**6.35**	**6.27**	**6.15**	**6.07**	**5.98**	**5.90**	**5.85**	**5.78**	**5.75**	**5.70**	**5.67**	**5.65**	
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93	8
	11.26	**8.65**	**7.59**	**7.01**	**6.63**	**6.37**	**6.19**	**6.03**	**5.91**	**5.82**	**5.74**	**5.67**	**5.56**	**5.48**	**5.36**	**5.28**	**5.20**	**5.11**	**5.06**	**5.00**	**4.96**	**4.91**	**4.88**	**4.86**	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71	9
	10.56	**8.02**	**6.99**	**6.42**	**6.06**	**5.80**	**5.62**	**5.47**	**5.35**	**5.26**	**5.18**	**5.11**	**5.00**	**4.92**	**4.80**	**4.73**	**4.64**	**4.56**	**4.51**	**4.45**	**4.41**	**4.36**	**4.33**	**4.31**	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.54	10
	10.04	**7.56**	**6.55**	**5.99**	**5.64**	**5.39**	**5.21**	**5.06**	**4.95**	**4.85**	**4.78**	**4.71**	**4.60**	**4.52**	**4.41**	**4.33**	**4.25**	**4.17**	**4.12**	**4.05**	**4.01**	**3.96**	**3.93**	**3.91**	
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40	11
	9.65	**7.20**	**6.22**	**5.67**	**5.32**	**5.07**	**4.88**	**4.74**	**4.63**	**4.54**	**4.46**	**4.40**	**4.29**	**4.21**	**4.10**	**4.02**	**3.94**	**3.86**	**3.80**	**3.74**	**3.70**	**3.66**	**3.62**	**3.60**	
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30	12
	9.33	**6.93**	**5.95**	**5.41**	**5.06**	**4.82**	**4.65**	**4.50**	**4.39**	**4.30**	**4.22**	**4.16**	**4.05**	**3.98**	**3.86**	**3.78**	**3.70**	**3.61**	**3.56**	**3.49**	**3.46**	**3.41**	**3.38**	**3.36**	
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.33	2.28	2.26	2.24	2.22	2.21	13
	9.07	**6.70**	**5.74**	**5.20**	**4.86**	**4.62**	**4.44**	**4.30**	**4.19**	**4.10**	**4.02**	**3.96**	**3.85**	**3.78**	**3.67**	**3.59**	**3.51**	**3.42**	**3.37**	**3.30**	**3.27**	**3.21**	**3.18**	**3.16**	

Bảng E (Tiếp theo)

f_2	f_1																								f_2
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13	14
	8.86	**6.51**	**5.56**	**5.03**	**4.69**	**4.46**	**4.28**	**4.14**	**4.03**	**3.94**	**3.86**	**3.80**	**3.70**	**3.62**	**3.51**	**3.43**	**3.34**	**3.26**	**3.21**	**3.14**	**3.11**	**3.06**	**3.02**	**3.00**	
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08	2.07	15
	8.68	**6.36**	**5.42**	**4.89**	**4.56**	**4.32**	**4.14**	**4.00**	**3.89**	**3.80**	**3.73**	**3.67**	**3.56**	**3.48**	**3.36**	**3.29**	**3.20**	**3.12**	**3.07**	**3.00**	**2.97**	**2.92**	**2.89**	**2.87**	
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02	2.01	16
	8.53	**6.23**	**5.29**	**4.77**	**4.44**	**4.20**	**4.03**	**3.89**	**3.78**	**3.69**	**3.61**	**3.55**	**3.45**	**3.37**	**3.25**	**3.18**	**3.10**	**3.01**	**2.96**	**2.98**	**2.86**	**2.80**	**2.77**	**2.75**	
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97	1.96	17
	8.40	**6.11**	**5.18**	**4.67**	**4.34**	**4.10**	**3.93**	**3.79**	**3.68**	**3.59**	**3.52**	**3.45**	**3.35**	**3.27**	**3.16**	**3.08**	**3.00**	**2.92**	**2.86**	**2.79**	**2.76**	**2.70**	**2.67**	**2.65**	
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.29	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	1.92	18
	8.28	**6.01**	**5.09**	**4.58**	**4.25**	**4.01**	**3.85**	**3.71**	**3.60**	**3.51**	**3.44**	**3.37**	**3.27**	**3.19**	**3.07**	**3.00**	**2.91**	**2.83**	**2.78**	**2.71**	**2.68**	**2.62**	**2.59**	**2.57**	
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.90	1.88	19
	8.18	**5.93**	**5.01**	**4.50**	**4.17**	**3.94**	**3.77**	**3.63**	**3.52**	**3.43**	**3.36**	**3.30**	**3.19**	**3.12**	**3.00**	**2.92**	**2.84**	**2.76**	**2.70**	**2.63**	**2.60**	**2.54**	**2.51**	**2.49**	
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84	20
	8.10	**5.85**	**4.94**	**4.43**	**4.10**	**3.87**	**3.71**	**3.56**	**3.45**	**3.37**	**3.30**	**3.23**	**3.13**	**3.05**	**2.94**	**2.86**	**2.77**	**2.69**	**2.63**	**2.56**	**2.53**	**2.47**	**2.44**	**2.42**	
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.20	2.15	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81	21
	8.02	**5.78**	**4.87**	**4.37**	**4.04**	**3.81**	**3.65**	**3.51**	**3.40**	**3.31**	**3.24**	**3.17**	**3.07**	**2.99**	**2.88**	**2.80**	**2.72**	**2.63**	**2.58**	**2.51**	**2.47**	**2.42**	**2.38**	**2.36**	
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78	22
	7.94	**5.72**	**4.82**	**4.31**	**3.99**	**3.76**	**3.59**	**3.45**	**3.35**	**3.26**	**3.18**	**3.12**	**3.02**	**2.94**	**2.83**	**2.75**	**2.67**	**2.58**	**2.53**	**2.46**	**2.42**	**2.37**	**2.33**	**2.31**	
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20	2.14	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76	23
	7.88	**5.66**	**4.76**	**4.26**	**3.94**	**3.71**	**3.54**	**3.41**	**3.30**	**3.21**	**3.14**	**3.07**	**2.97**	**2.89**	**2.78**	**2.70**	**2.62**	**2.53**	**2.48**	**2.41**	**2.37**	**2.32**	**2.28**	**2.26**	
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.76	1.74	1.73	24
	7.82	**5.61**	**4.72**	**4.22**	**3.90**	**3.67**	**3.50**	**3.36**	**3.25**	**3.17**	**3.09**	**3.03**	**2.93**	**2.85**	**2.74**	**2.66**	**2.58**	**2.49**	**2.44**	**2.36**	**2.33**	**2.27**	**2.23**	**2.21**	
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.72	1.71	25
	7.77	**5.57**	**4.68**	**4.18**	**3.86**	**3.63**	**3.46**	**3.32**	**3.21**	**3.13**	**3.05**	**2.99**	**2.89**	**2.81**	**2.70**	**2.62**	**2.54**	**2.45**	**2.40**	**2.32**	**2.29**	**2.23**	**2.19**	**2.17**	
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.10	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.72	1.70	1.69	26
	7.72	**5.53**	**4.64**	**4.14**	**3.82**	**3.59**	**3.42**	**3.29**	**3.17**	**3.09**	**3.02**	**2.96**	**2.86**	**2.77**	**2.66**	**2.58**	**2.50**	**2.41**	**2.36**	**2.28**	**2.25**	**2.19**	**2.15**	**2.13**	

Bảng E (Tiếp theo)

f_2	f_1																								f_2
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67	27
	7.68	**5.49**	**4.60**	**4.11**	**3.79**	**3.56**	**3.39**	**3.26**	**3.14**	**3.06**	**2.98**	**2.93**	**2.83**	**2.74**	**2.63**	**2.55**	**2.47**	**2.38**	**2.33**	**2.25**	**2.21**	**2.16**	**2.12**	**2.10**	
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65	28
	7.64	**5.45**	**4.57**	**4.07**	**3.76**	**3.53**	**3.36**	**3.23**	**3.11**	**3.03**	**2.95**	**2.90**	**2.80**	**2.71**	**2.60**	**2.52**	**2.44**	**2.35**	**2.30**	**2.22**	**2.18**	**2.13**	**2.09**	**2.06**	
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64	29
	7.60	**5.42**	**4.54**	**4.04**	**3.73**	**3.50**	**3.33**	**3.20**	**3.08**	**3.00**	**2.92**	**2.87**	**2.77**	**2.68**	**2.57**	**2.49**	**2.41**	**2.32**	**2.27**	**2.19**	**2.15**	**2.10**	**2.06**	**2.03**	
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62	30
	7.56	**5.39**	**4.51**	**4.02**	**3.70**	**3.47**	**3.30**	**3.17**	**3.06**	**2.98**	**2.90**	**2.84**	**2.74**	**2.66**	**2.55**	**2.47**	**2.38**	**2.29**	**2.24**	**2.16**	**2.13**	**2.07**	**2.03**	**2.01**	
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59	32
	7.50	**5.34**	**4.46**	**3.97**	**3.66**	**3.42**	**3.25**	**3.12**	**3.01**	**2.94**	**2.86**	**2.80**	**2.70**	**2.62**	**2.51**	**2.42**	**2.34**	**2.25**	**2.20**	**2.12**	**2.08**	**2.02**	**1.98**	**1.96**	
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57	34
	7.44	**5.29**	**4.42**	**3.93**	**3.61**	**3.38**	**3.21**	**3.08**	**2.97**	**2.89**	**2.82**	**2.76**	**2.66**	**2.58**	**2.47**	**2.38**	**2.30**	**2.21**	**2.15**	**2.08**	**2.04**	**1.98**	**1.94**	**1.91**	
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.06	2.03	1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.56	1.55	36
	7.39	**5.25**	**4.38**	**3.89**	**3.58**	**3.35**	**3.18**	**3.04**	**2.94**	**2.86**	**2.78**	**2.72**	**2.62**	**2.54**	**2.43**	**2.35**	**2.26**	**2.17**	**2.12**	**2.04**	**2.00**	**1.94**	**1.90**	**1.87**	
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53	38
	7.35	**5.21**	**4.34**	**3.86**	**3.54**	**3.32**	**3.15**	**3.02**	**2.91**	**2.82**	**2.75**	**2.69**	**2.59**	**2.51**	**2.40**	**2.32**	**2.22**	**2.14**	**2.08**	**2.00**	**1.97**	**1.90**	**1.86**	**1.84**	
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51	40
	7.31	**5.18**	**4.31**	**3.83**	**3.51**	**3.29**	**3.12**	**2.99**	**2.88**	**2.80**	**2.73**	**2.66**	**2.56**	**2.49**	**2.37**	**2.29**	**2.20**	**2.11**	**2.05**	**1.97**	**1.94**	**1.88**	**1.84**	**1.81**	
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49	42
	7.27	**5.15**	**4.29**	**3.80**	**3.49**	**3.26**	**3.10**	**2.96**	**2.86**	**2.77**	**2.70**	**2.64**	**2.54**	**2.46**	**2.35**	**2.26**	**2.17**	**2.08**	**2.02**	**1.94**	**1.91**	**1.85**	**1.80**	**1.78**	
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.92	1.88	1.81	1.76	1.72	1.66	1.63	4.58	1.56	1.52	1.50	1.48	44
	7.24	**5.12**	**4.26**	**3.78**	**3.46**	**3.24**	**3.07**	**2.94**	**2.84**	**2.75**	**2.68**	**2.62**	**2.52**	**2.44**	**2.32**	**2.24**	**2.15**	**2.06**	**2.00**	**1.92**	**1.88**	**1.82**	**1.78**	**1.75**	
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97	1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46	46
	7.21	**5.10**	**4.24**	**3.76**	**3.44**	**3.22**	**3.05**	**2.92**	**2.82**	**2.73**	**2.66**	**2.60**	**2.50**	**2.42**	**2.30**	**2.22**	**2.13**	**2.04**	**1.98**	**1.90**	**1.86**	**1.80**	**1.76**	**1.72**	
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.96	1.90	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45	48
	7.19	**5.08**	**4.22**	**3.74**	**3.42**	**3.20**	**3.04**	**2.90**	**2.80**	**2.71**	**2.64**	**2.58**	**2.48**	**2.40**	**2.28**	**2.20**	**2.11**	**2.02**	**1.96**	**1.88**	**1.84**	**1.78**	**1.73**	**1.70**	

Bảng E (Tiếp theo)

f_2	f_1																								f_2
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.95	1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44	50
	7.17	**5.06**	**4.20**	**3.72**	**3.41**	**3.18**	**3.02**	**2.88**	**2.78**	**2.70**	**2.62**	**2.56**	**2.46**	**2.39**	**2.26**	**2.18**	**2.10**	**2.00**	**1.94**	**1.86**	**1.82**	**1.76**	**1.71**	**1.68**	
55	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.97	1.93	1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41	55
	7.12	**5.01**	**4.16**	**3.68**	**3.37**	**3.15**	**2.98**	**2.85**	**2.75**	**2.66**	**2.59**	**2.53**	**2.43**	**2.35**	**2.23**	**2.15**	**2.06**	**1.96**	**1.90**	**1.82**	**1.78**	**1.71**	**1.66**	**1.64**	
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39	60
	7.08	**4.98**	**4.13**	**3.65**	**3.34**	**3.12**	**2.95**	**2.82**	**2.72**	**2.63**	**2.56**	**2.50**	**2.40**	**2.32**	**2.20**	**2.12**	**2.03**	**1.93**	**1.87**	**1.79**	**1.74**	**1.68**	**1.63**	**1.60**	
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.94	1.90	1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.42	1.39	1.37	65
	7.04	**4.95**	**4.10**	**3.62**	**3.31**	**3.09**	**2.93**	**2.79**	**2.70**	**2.61**	**2.54**	**2.47**	**2.37**	**2.30**	**2.18**	**2.09**	**2.00**	**1.90**	**1.84**	**1.76**	**1.71**	**1.64**	**1.60**	**1.56**	
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89	1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37	1.35	70
	7.01	**4.92**	**4.08**	**3.60**	**3.29**	**3.07**	**2.91**	**2.77**	**2.67**	**2.59**	**2.51**	**2.45**	**2.35**	**2.28**	**2.15**	**2.07**	**1.98**	**1.88**	**1.82**	**1.74**	**1.69**	**1.62**	**1.56**	**1.53**	
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.88	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32	80
	6.96	**4.88**	**4.04**	**3.56**	**3.25**	**3.04**	**2.87**	**2.74**	**2.64**	**2.55**	**2.48**	**2.41**	**2.32**	**2.24**	**2.11**	**2.03**	**1.94**	**1.84**	**1.78**	**1.70**	**1.65**	**1.57**	**1.52**	**1.49**	
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28	100
	6.90	**4.82**	**3.98**	**3.51**	**3.20**	**2.99**	**2.82**	**2.69**	**2.59**	**2.51**	**2.43**	**2.36**	**2.26**	**2.19**	**2.06**	**1.98**	**1.89**	**1.79**	**1.73**	**1.64**	**1.59**	**1.51**	**1.46**	**1.43**	
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.86	1.83	1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25	125
	6.84	**4.78**	**3.94**	**3.47**	**3.17**	**2.95**	**2.79**	**2.65**	**2.56**	**2.47**	**2.40**	**2.33**	**2.23**	**2.15**	**2.03**	**1.94**	**1.85**	**1.75**	**1.68**	**1.59**	**1.54**	**1.46**	**1.40**	**1.37**	
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22	150
	6.81	**4.75**	**3.91**	**3.44**	**3.14**	**2.92**	**2.76**	**2.62**	**2.53**	**2.44**	**2.37**	**2.30**	**2.20**	**2.12**	**2.00**	**1.91**	**1.83**	**1.72**	**1.66**	**1.56**	**1.51**	**1.43**	**1.37**	**1.33**	
200	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92	1.87	1.83	1.80	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19	200
	6.76	**4.71**	**3.88**	**3.41**	**3.11**	**2.90**	**2.73**	**2.60**	**2.50**	**2.41**	**2.34**	**2.28**	**2.17**	**2.09**	**1.97**	**1.88**	**1.79**	**1.69**	**1.62**	**1.53**	**1.48**	**1.39**	**1.33**	**1.28**	
400	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.78	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13	400
	6.70	**4.66**	**3.83**	**3.36**	**3.06**	**2.85**	**2.69**	**2.55**	**2.46**	**2.37**	**2.29**	**2.23**	**2.12**	**2.04**	**1.92**	**1.84**	**1.74**	**1.64**	**1.57**	**1.47**	**1.42**	**1.32**	**1.24**	**1.19**	
1000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.80	1.76	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08	1000
	6.66	4.62	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.26	2.20	2.09	2.01	1.89	1.81	1.71	1.61	1.54	1.44	1.38	1.28	1.19	1.11	
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	2.79	1.75	1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.28	1.24	1.17	1.11	1.00	∞
	6.64	**4.60**	**3.78**	**3.32**	**3.02**	**2.80**	**2.64**	**2.51**	**2.41**	**2.32**	**2.24**	**2.18**	**2.07**	**1.99**	**1.87**	**1.79**	**1.69**	**1.59**	**1.52**	**1.41**	**1.36**	**1.25**	**1.15**	**1.00**	

Bảng Fa: Trị số R_p dùng cho so sánh Duncan (5%) - Significant Studentized Factor (R) to multiply by LSD for testing mean at various ranging (p) 5% level; n = degree of freedom for error

n	p:															
	2	3	4	5	6	7	8	9	10	12	14	16	18	20	50	100
4	1.00	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
5	1.00	1.03	1.04	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05
6	1.00	1.03	1.05	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06
7	1.00	1.04	1.06	1.07	1.07	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08
8	1.00	1.04	1.06	1.08	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.09
9	1.00	1.04	1.07	1.08	1.09	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10
10	1.00	1.05	1.07	1.09	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10
11	1.00	1.05	1.08	1.09	1.10	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.12	1.12	1.12	1.12
12	1.00	1.05	1.08	1.09	1.10	1.11	1.12	1.12	1.12	1.12	1.12	1.12	1.13	1.13	1.13	1.13
13	1.00	1.05	1.08	1.09	1.10	1.11	1.12	1.12	1.13	1.13	1.13	1.13	1.13	1.13	1.13	1.13
14	1.00	1.05	1.08	1.10	1.11	1.12	1.13	1.13	1.14	1.14	1.14	1.14	1.15	1.15	1.15	1.15
15	1.00	1.05	1.08	1.10	1.12	1.12	1.13	1.14	1.14	1.14	1.15	1.15	1.15	1.15	1.15	1.15
16	1.00	1.05	1.08	1.10	1.12	1.12	1.13	1.14	1.14	1.15	1.15	1.16	1.16	1.16	1.16	1.16
17	1.00	1.05	1.08	1.10	1.12	1.13	1.13	1.14	1.15	1.15	1.16	1.16	1.16	1.16	1.16	1.16
18	1.00	1.05	1.08	1.10	1.12	1.13	1.13	1.14	1.15	1.15	1.16	1.16	1.17	1.17	1.17	1.17
19	1.00	1.05	1.08	1.10	1.12	1.13	1.14	1.15	1.15	1.16	1.16	1.17	1.17	1.17	1.17	1.17
20	1.00	1.05	1.08	1.10	1.12	1.13	1.14	1.15	1.15	1.16	1.17	1.17	1.17	1.18	1.18	1.18
22	1.00	1.05	1.08	1.10	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.18	1.18	1.18	1.18
24	1.00	1.05	1.08	1.10	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.18	1.19	1.19	1.19
26	1.00	1.05	1.08	1.10	1.12	1.13	1.15	1.15	1.16	1.17	1.18	1.19	1.19	1.19	1.19	1.19
28	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.16	1.17	1.18	1.19	1.19	1.20	1.20	1.20
30	1.00	1.05	1.08	1.11	1.12	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.20	1.20	1.20	1.20
40	1.00	1.05	1.08	1.11	1.13	1.14	1.15	1.16	1.17	1.19	1.20	1.20	1.21	1.21	1.21	1.21
60	1.00	1.05	1.09	1.11	1.13	1.14	1.16	1.17	1.18	1.19	1.20	1.21	1.22	1.23	1.23	1.23
100	1.00	1.05	1.09	1.11	1.14	1.15	1.16	1.18	1.19	1.20	1.21	1.22	1.23	1.24	1.26	1.26
∞	1.00	1.05	1.09	1.12	1.14	1.15	1.17	1.18	1.19	1.21	1.22	1.23	1.24	1.25	1.30	1.32

Bảng Fb: Trị số R_p dùng cho so sánh Duncan (1%)- Significant Studentized Factor (R) to multiply by LSD for testing mean at various ranging (p) 1% level; n = degree of freedom for error

n	p:															
	2	3	4	5	6	7	8	9	10	12	14	16	18	20	50	100
3	1.00	1.03	1.04	1.05	1.07	1.08	1.08	1.09	1.09	1.09	1.10	1.11	1.13	1.13	1.13	1.13
4	1.00	1.04	1.06	1.08	1.09	1.09	1.11	1.11	1.12	1.12	1.14	1.14	1.15	1.15	1.15	1.15
5	1.00	1.05	1.07	1.08	1.10	1.11	1.12	1.13	1.14	1.16	1.16	1.18	1.18	1.19	1.19	1.19
6	1.00	1.05	1.08	1.09	1.11	1.12	1.14	1.15	1.15	1.16	1.18	1.18	1.20	1.20	1.20	1.20
7	1.00	1.05	1.08	1.10	1.12	1.13	1.15	1.16	1.17	1.17	1.19	1.19	1.21	1.21	1.21	1.21
8	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.17	1.18	1.20	1.20	1.22	1.22	1.22	1.23
9	1.00	1.06	1.08	1.10	1.12	1.14	1.16	1.17	1.17	1.20	1.20	1.22	1.24	1.24	1.24	1.24
10	1.00	1.06	1.09	1.11	1.13	1.15	1.16	1.17	1.18	1.20	1.21	1.22	1.24	1.24	1.24	1.24
11	1.00	1.05	1.09	1.11	1.13	1.14	1.15	1.17	1.17	1.19	1.20	1.22	1.23	1.23	1.23	1.23
12	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.17	1.19	1.20	1.21	1.21	1.22	1.22	1.22
13	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.21	1.21	1.21
14	1.00	1.05	1.08	1.10	1.12	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.20	1.20	1.20	1.20
15	1.00	1.05	1.08	1.10	1.11	1.13	1.14	1.15	1.16	1.18	1.18	1.19	1.20	1.20	1.20	1.20
16	1.00	1.05	1.08	1.10	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.20	1.20	1.20
17	1.00	1.05	1.08	1.10	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.19	1.19	1.19
18	1.00	1.05	1.08	1.10	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.19	1.19	1.19	1.19
19	1.00	1.05	1.07	1.09	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.19	1.19	1.19	1.19
20	1.00	1.05	1.07	1.09	1.11	1.13	1.14	1.15	1.15	1.17	1.18	1.18	1.19	1.19	1.19	1.19
22	1.00	1.05	1.07	1.09	1.11	1.12	1.14	1.15	1.15	1.17	1.17	1.18	1.19	1.19	1.19	1.19
24	1.00	1.05	1.07	1.09	1.11	1.12	1.13	1.15	1.15	1.17	1.17	1.18	1.19	1.19	1.20	1.20
26	1.00	1.05	1.07	1.09	1.11	1.12	1.13	1.15	1.15	1.17	1.18	1.18	1.19	1.19	1.20	1.20
28	1.00	1.04	1.07	1.09	1.11	1.12	1.13	1.14	1.15	1.17	1.18	1.18	1.19	1.19	1.21	1.21
30	1.00	1.04	1.07	1.09	1.11	1.12	1.13	1.14	1.15	1.17	1.18	1.19	1.19	1.20	1.21	1.21
40	1.00	1.04	1.07	1.09	1.11	1.12	1.14	1.14	1.15	1.17	1.18	1.19	1.20	1.20	1.23	1.23
60	1.00	1.04	1.07	1.09	1.11	1.12	1.14	1.15	1.15	1.17	1.18	1.19	1.20	1.20	1.24	1.24
100	1.00	1.04	1.07	1.09	1.11	1.12	1.14	1.15	1.15	1.17	1.18	1.19	1.20	1.21	1.25	1.25
∞	1.00	1.04	1.07	1.09	1.11	1.12	1.14	1.15	1.15	1.17	1.18	1.19	1.20	1.21	1.26	1.29

Bảng H: Giá trị hệ số tương quan tuyến tính đơn giản r - Simple Linear Correlation Coefficients, r, at the 5% and 1% level of significance

d.f.[1]	5%	1%	d.f.	5%	1%
1	0.997	1.000	26	0.374	0.478
2	0.950	0.990	27	0.367	0.470
3	0.878	0.959	28	0.361	0.463
4	0.811	0.917	29	0.355	0.456
5	0.754	0.874	30	0.349	0.449
6	0.707	0.834	32	0.339	0.437
7	0.666	0.798	34	0.329	0.424
8	0.632	0.765	36	0.321	0.413
9	0.602	0.735	38	0.312	0.403
10	0.576	0.708	40	0.304	0.393
11	0.553	0.684	45	0.288	0.372
12	0.532	0.661	50	0.273	0.354
13	0.514	0.641	55	0.262	0.340
14	0.497	0.623	60	0.250	0.325
15	0.492	0.606	70	0.232	0.302
16	0.486	0.590	80	0.217	0.283
17	0.456	0.575	90	0.205	0.267
18	0.444	0.561	100	0.195	0.254
19	0.433	0.549	125	0.174	0.228
20	0.423	0.537	150	0.159	0.208
21	0.413	0.526	175	0.148	0.194
22	0.404	0.515	200	0.138	0.181
23	0.396	0.505	300	0.113	0.148
24	0.388	0.496	400	0.098	0.128
25	0.381	0.487	500	0.088	0.115

[1] df = n-2 ; n là cỡ mẫu

Bảng K: Các cách xắp xếp công thức trong kiểu thí nghiệm Ô vuông la tinh (LS)

3 x 3

A	B	C
B	C	A
C	A	B

4 x 4

1

A	B	C	D
B	A	D	C
C	D	B	A
D	C	A	B

2

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

3

A	B	C	D
B	D	A	C
C	A	D	B
D	C	B	A

4

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

5 x 5

A	B	C	D	E
B	A	E	C	D
C	D	A	E	B
D	E	B	A	C
E	C	D	B	A

6 x 6

A	B	C	D	E	F
B	F	D	C	A	E
C	D	E	F	B	A
D	A	F	E	C	B
E	C	A	B	F	D
F	E	B	A	D	C

7 x 7

A	B	D	C	E	F	G
B	C	D	E	F	G	A
C	D	E	F	G	A	B
D	E	F	G	A	B	C
E	F	G	A	B	C	D
F	G	A	B	C	D	E
G	A	B	C	D	E	F

8 x 8

A	B	C	D	E	F	G	H
B	C	D	E	F	G	H	A
C	D	E	F	G	H	A	B
D	E	F	G	H	A	B	C
E	F	G	H	A	B	C	D
F	G	H	A	B	C	D	E
G	H	A	B	C	D	E	F
H	A	B	C	D	E	F	G

9 x 9

A	B	S	D	E	F	G	H	I
B	C	D	E	F	G	H	I	A
C	D	E	F	G	H	I	A	B
D	E	F	G	H	I	A	B	C
E	F	G	H	I	A	B	C	D
F	G	H	I	A	B	C	D	E
G	H	I	A	B	C	D	E	F
H	I	A	B	C	D	E	F	G
I	A	B	C	D	E	F	G	H

10 x 10

A	B	C	D	E	F	G	H	I	J
B	C	D	E	F	G	H	I	J	A
C	D	E	F	G	H	I	J	A	B
D	E	F	G	H	I	J	A	B	C
E	F	G	H	I	J	A	B	C	D
F	G	H	I	J	A		B	C	D
G	H	I	J	A	B	C	D	E	F
H	I	J	A	B	C	D	E	F	G
I	J	A	B	C	D	E	F	G	H
J	A	B	C	D	E	F	G	H	I

11 x 11

A	B	C	D	E	F	G	H	I	J	K
B	C	D	E	F	G	H	I	J	K	A
C	D	E	F	G	H	I	J	K	A	B
D	E	F	G	H	I	J	K	A	B	C
E	F	G	H	I	J	K	A	B	C	D
F	G	H	I	J	K	A	B	C	D	E
G	H	I	J	K	A	B	C	C	D	F
H	I	J	K	A	B	C	D	E	F	G
I	J	K	A	B	C	D	E	F	G	H
J	K	A	B	C	D	E	F	G	H	I
K	A	B	C	D	E	F	G	H	I	J

Chịu trách nhiệm xuất bản

NGUYỄN CAO DOANH

Phụ trách bản thảo

BÍCH HOA - HOÀI ANH

Trình bày bìa

ĐỖ THỊNH

NHÀ XUẤT BẢN NÔNG NGHIỆP

D14, Phương Mai, Đống Đa - Hà Nội

ĐT: 8524501 - 8521940 FAX: 04.5760748

CHI NHÁNH NXB NÔNG NGHIỆP

58 Nguyễn Bỉnh Khiêm - Q.1, Tp. Hồ Chí Minh

ĐT: 08.8299521 - 8297157 FAX: 08.9101036

In 215 bản, khổ 19 × 27cm tại XN In 15 Bộ Công nghiệp. Giấy chấp nhận đăng l
kế hoạch xuất bản số 191/393 do Cục Xuất bản cấp ngày 11/4/04. In xong và n
lưu chiểu quý II/2004.

M 630/15193

63 - 630 / NN - 04 - 191(b)/393 - 03