

TRƯỜNG ĐẠI HỌC NÔNG NGHIỆP I HÀ NỘI
KHOA CHĂN NUÔI - THÚ Y

BÀI GIẢNG
PHƯƠNG PHÁP THÍ NGHIỆM
TRONG CHĂN NUÔI & THÚ Y
(PHẦN II)

Đỗ Đức Lực
Bộ môn Di truyền - Giống, Khoa Chăn nuôi - Thú y

Hà Nội - 2004

MỤC LỤC

1. Các khái niệm cơ bản và các bước tiến hành thí nghiệm.....	5
1.1. Giới thiệu	5
1.2. Mục đích.....	5
1.3. Yêu cầu của thí nghiệm	5
1.4. Các loại thí nghiệm.....	5
1.5. Một số khái niệm cơ bản	5
2. Các bước tiến hành lập kế hoạch thí nghiệm	6
2.1. Xác định mục đích nghiên cứu.....	6
2.2. Lựa chọn đối tượng cần nghiên cứu	6
2.3. Xác định các nguồn gây biến động	9
2.4. Lập sơ đồ thí nghiệm hoặc quan sát	10
2.5. Lựa chọn mô hình thống kê để phân tích số liệu.....	10
2.6. Tiến hành thí nghiệm.....	10
2.7. Thu thập số liệu.....	11
2.8. Phân tích số liệu.....	11
2.9. Viết báo cáo (phần này sẽ đề cập ở cuối kỳ).....	11
3. Bố trí thí nghiệm 1 nhân tố	12
3.1. Thí nghiệm kiểu hoàn toàn ngẫu nhiên.....	12
3.2. Thí nghiệm kiểu khối ngẫu nhiên đầy đủ	21
3.3. Thí nghiệm kiểu ô vuông La tinh.....	23
4. Bố trí thí nghiệm 2 nhân tố	29
4.1. Thí nghiệm kiểu hoàn toàn ngẫu nhiên.....	29
4.2. Bài kiểm tra số 5	32
4.3. Bài kiểm tra số 6	33
5. Hồi quy tuyến tính đơn giản.....	34
5.1. Giới thiệu	34
5.2. Ví dụ	34
5.3. Mô hình tuyến tính	35
5.4. Ước lượng các tham số β_0 và β_1	35
5.5. Kiểm định giả thuyết	38
6. Tương quan.....	41
6.1. Giới thiệu	41
6.2. Tính hệ số tương quan.....	41
6.3. Những ví dụ về sự tương quan	42
6.4. Tính toán trong Minitab.....	43
6.5. Mối quan hệ và hệ số tương quan.....	43
6.6. Mối quan hệ giữa tương quan và hồi quy	44
6.7. Bài kiểm tra số 7	45
7. Phụ lục.....	47
8. Tài liệu tham khảo	54
8.1. Tiếng Việt.....	54
8.2. Tiếng Anh	54
8.3. Tiếng Nga.....	54
8.4. Tiếng Pháp	54

Bài giảng môn học *Phương pháp thí nghiệm trong chăn nuôi thú y* được soạn riêng cho sinh viên chuyên ngành chăn nuôi & thú y, hệ chính quy. Bài giảng này bao gồm 2 phần; đây là phần II, bao gồm 2 chủ đề chính là *Bố trí thí nghiệm* và *Tương quan hồi quy*; phần I được in riêng với 2 chủ đề chính là *Tóm tắt dữ liệu* và *Ước lượng & Kiểm định giả thiết*.

Mặc dù có rất nhiều cố gắng trong quá trình biên soạn, xong không thể tránh được những thiếu sót. Tác giả rất mong sự góp ý của bạn đọc. Mọi ý kiến góp ý xin gửi theo địa chỉ sau đây:

Đỗ Đức Lực

Phòng 303 & 304

**Bộ môn Di truyền - Giống, Khoa Chăn nuôi - Thú y
Đại học Nông nghiệp I Hà Nội, Trâu Quỳ, Gia Lâm**

E-mail: dtghn@yahoo.co.uk

Điện thoại Bộ môn: 04 - 876 82 65

Giới thiệu chung

Trong khoá học *Phương pháp thí nghiệm trong chăn nuôi và thú y* sẽ đề cập đến 4 nội dung chính sau đây:

- Tóm tắt và mô tả số liệu
- Kiểm định giả thuyết
- Các nguyên tắc cơ bản và một số mô hình thiết kế thí nghiệm thường gặp trong chăn nuôi và thú y.
- Tương quan và hồi quy.

Khoá học sẽ cung cấp cho sinh viên chuyên ngành chăn nuôi thú y nắm được cách phân tích số liệu, các nguyên tắc bố trí một thí nghiệm và rút ra những kết luận từ việc phân tích số liệu.

Tổng số thời lượng của khoá học là 2 đơn vị học trình (30 tiết), trong đó phần lý thuyết 20 tiết và thực hành 10 tiết. Các bài thực hành được thực hiện tại Phòng máy tính Khoa Chăn nuôi - Thú y (Phòng 218 tầng 2).

Trong suốt khoá học sẽ có 5 bài kiểm tra; điểm số của mỗi bài kiểm tra được nhân với hệ số 0,1 nhưng chỉ lấy 4 bài có điểm số cao nhất để tính vào điểm cuối kỳ. Kết thúc khoá học sẽ có một bài thi cuối kỳ; điểm số của bài thi được nhân với hệ số 0,6. Điểm đánh giá của môn học chính là tổng số điểm của 4 bài kiểm tra và bài thi cuối kỳ sau khi đã nhân với các hệ số tương ứng. Học viên được sử dụng tài liệu trong quá trình làm bài kiểm tra hoặc bài thi.

1. Các khái niệm cơ bản và các bước tiến hành thí nghiệm

1.1. Giới thiệu

Mô hình thí nghiệm đóng một vai trò quan trọng như trong phân tích thống kê. Mô hình thí nghiệm phải được xây dựng đối với từng thí nghiệm cụ thể, phụ thuộc vào yêu cầu đặt ra, điều kiện và kết quả mong đợi của thí nghiệm.

Nếu ta có một mô hình thí nghiệm tốt thì cũng luôn luôn có một phương pháp phân tích tương ứng, chính xác và nhanh chóng; ngược lại chúng ta sẽ gặp nhiều khó khăn trong công việc xử lý số liệu, hoặc số liệu không thể phân tích được, hoặc là xử lý được nhưng không phản ánh kết quả một cách chính xác.

1.2. Mục đích

Bố trí thí nghiệm là lập kế hoạch, các bước tiến hành để thu thập số liệu cho vấn đề cần nghiên cứu để từ đó rút ra những kết luận chính xác với chi phí tối thiểu.

1.3. Yêu cầu của thí nghiệm

- Thí nghiệm phải mang tính chất điển hình
- Triệt để tôn trọng nguyên tắc sai khác duy nhất
- Thí nghiệm phải đạt được độ chính xác nhất định
- Thí nghiệm phải có khả năng diễn lại
- Thí nghiệm phải được tiến hành trên những vật liệu đã nắm rõ được tiền sử của chúng

1.4. Các loại thí nghiệm

Theo mức độ và quy mô ta có thể chia thí nghiệm thành:

- Thí nghiệm thăm dò
- Thí nghiệm chính thức
- Thí nghiệm thực hiện trong điều kiện sản xuất

Theo bản chất của thí nghiệm ta có thể chia thí nghiệm thành:

- **Thí nghiệm quan sát** là thí nghiệm được tiến hành dựa trên những yếu tố ta đã có để tiến hành thu thập, phân tích số liệu và đưa ra các kết luận. Ưu điểm của loại thí nghiệm này là ít tốn kém thời gian, công sức và chi phí hơn; nhưng hạn chế là chỉ tiến hành nghiên cứu được những yếu tố đã có và không kiểm soát hoặc chỉ kiểm soát được một phần không lớn được các yếu tố phi thí nghiệm.
- **Thí nghiệm bố trí** là thí nghiệm đòi hỏi phải có thời gian và địa điểm để tiến hành thí nghiệm. Ưu điểm và nhược điểm của loại thí nghiệm này thì hoàn toàn trái ngược với thí nghiệm quan sát.

1.5. Một số khái niệm cơ bản

- **Yếu tố** là một biến độc lập cần nghiên cứu ở nhiều mức độ khác nhau.
- **Mức** là số công thức thí nghiệm trong một yếu tố
- **Nghiệm thức** là tổ hợp giữa yếu tố và mức

- **Đơn vị thí nghiệm** là một đơn vị nghiên cứu trong thí nghiệm, hay nói cụ thể hơn đó chính là đơn vị bé nhất (một lần lặp lại) trong mỗi nghiệm thức được áp dụng.
- **Khối** là những đơn vị thí nghiệm có chung một hay nhiều đặc tính.
- **Ngẫu nhiên** là cách bố trí các đơn vị thí nghiệm vào các nghiệm thức hoàn toàn ngẫu nhiên.
- **Lặp lại** là số đơn vị thí nghiệm trong một nghiệm thức. Trong một nghiệm thức có thể có một hay nhiều lần lặp lại (đơn vị thí nghiệm)
- **Nhắc lại** là tiến hành thực hiện lại thí nghiệm đã tiến hành trước đó với các điều kiện tương tự
- **Nhóm đối chứng** là nhóm động vật được chọn ra trong quá trình bố thí nghiệm nhưng được nuôi dưỡng trong điều kiện hiện có.

2. Các bước tiến hành lập kế hoạch thí nghiệm

- Xác định mục đích nghiên cứu
- Lựa chọn đối tượng nghiên cứu
- Xác định các nguồn gây biến động đối với các thí nghiệm quan sát hoặc phải quản lý được các nguồn gây biến động đối với các thí nghiệm bố trí
- Xác định đơn vị quan sát đối với các thí nghiệm quan sát hoặc đơn vị thí nghiệm đối với các thí nghiệm bố trí.
- Lựa chọn cách quan sát hoặc lập sơ đồ thí nghiệm
- Lựa chọn mô hình thống kê để phân tích số liệu
- Tiến hành thí nghiệm
- Thu thập số liệu
- Phân tích số liệu
- Viết báo cáo

2.1. Xác định mục đích nghiên cứu

Để xác định được mục tiêu nghiên cứu ta cần phải giải đáp những câu hỏi sau:

- Tính cấp thiết của vấn đề nghiên cứu?
- Giả thiết nghiên cứu và các tham số cần ước tính?
- Mức độ ưu tiên của các vấn đề đặt ra?
- Kết quả mong đợi và mức độ chính xác của thí nghiệm?

2.2. Lựa chọn đối tượng cần nghiên cứu

- Chọn đối tượng nghiên cứu
- Cách chọn động vật thí nghiệm.
 - Nhóm động vật thí nghiệm tương tự về chất lượng sẽ được chọn ra (giống, nguồn gốc, giới tính...).
 - Nên chọn những động vật cùng một giống;

- Những động vật được chọn phải tiêu biểu cho loại giống đó; không quá khác biệt về ngoại hình và đặc điểm sinh lý so với đặc điểm chung của toàn đàn. Đối với một số thí nghiệm bố trí theo cặp tốt nhất dùng những động vật sinh đôi cùng trứng, cùng máu, nửa anh em theo cha (theo cùng một dòng hoặc họ).
 - Cuối cùng ta cũng có thể sử dụng những động vật không cùng dòng, họ nhưng tương đối tương tự nhau về ngoại hình và một số tính chất khác.
 - Hoặc các động vật thí nghiệm được chọn một cách hoàn toàn ngẫu nhiên từ quần thể
 - Tóm lại, để tạo ra các nhóm tương đối giống nhau ta cũng có thể chỉ chọn những động vật cùng giới, cùng lứa tuổi, cùng mức độ tăng trưởng, cùng thể chất, tình trạng sức khỏe... Trong một số trường hợp để chọn được những cặp tương tự chúng ta phải tiến hành nghiên cứu sơ bộ tới thành phần của máu, hô hấp...
- Số lượng đơn vị thí nghiệm
 - Cần bao nhiêu động vật? Cần phải đủ sao cho các đặc tính riêng biệt của từng cá thể không làm ảnh hưởng lên kết quả của thí nghiệm.
 - Điều gì sẽ xảy ra nếu số lượng động vật quá ít trong thí nghiệm? Độ tin cậy của kết quả thu được từ thí nghiệm sẽ không cao
 - Điều gì sẽ xảy ra nếu số lượng động vật quá nhiều trong thí nghiệm? Không phải lúc nào ta cũng cần số lượng động vật thí nghiệm quá lớn. Nếu quá lớn có thể gây ra nhiều khó khăn trong quá trình theo dõi đối với từng cá thể, tạo ra khó khăn khi muốn tạo ra các điều kiện đồng nhất, cho động vật ăn... chính là những lý do làm giảm độ chính xác về mặt kỹ thuật của thí nghiệm; ngoài ra còn tạo thêm nhiều khó khăn trong quá trình tính toán các chỉ tiêu theo dõi.
 - Những yếu tố nào làm ảnh hưởng đến số lượng động vật tham gia thí nghiệm?
 - Chất lượng của động vật tham gia thí nghiệm (giống, độ tuổi, thể trạng của động vật); càng đồng nhất về giống thì càng giảm được số động vật thí nghiệm và ngược lại
 - Mức độ chuẩn bị để đưa vào thí nghiệm (mức độ phát triển, chuẩn bị cân bằng)
 - Tính chất của thí nghiệm (thí nghiệm thăm dò hay mang tính quyết định)
 - Kết quả mong đợi của thí nghiệm (sự sai khác giữa các công thức thí nghiệm)
 - Nhiệm vụ đặt ra
 - Độ tuổi của con vật cũng đóng vai trò quan trọng trong quá trình chọn dung lượng mẫu. Dưới tác động của các yếu tố ngoại cảnh, động vật càng non thì mức độ biến động càng lớn (cả về mặt sinh lý và ngoại hình). Bảng dưới đây cho ta thấy số lượng động vật tham gia thí nghiệm cũng phụ thuộc rất nhiều vào độ tuổi.

Độ tuổi	Số lượng trong một nghiệm thức	Độ tuổi	Số lượng trong một nghiệm thức
Bê (đến 1 năm tuổi)	17	Bò đẻ lứa ba	13
Bê (1 đến 2 năm tuổi)	16	Bò đẻ lứa bốn	12
Bò đẻ lứa đầu	15	Bò đẻ lứa năm	11
Bò đẻ lứa hai	14	Bò đẻ lứa sáu	10

Trong quá trình thiết kế thí nghiệm cũng cần phải chú ý rằng, nhu cầu dinh dưỡng của động vật luôn thay đổi tùy theo vào độ tuổi của. Trong thí nghiệm của Kurilo, hàm lượng lizin được bổ sung vào khẩu phần ăn của lợn con nuôi vỗ béo thay đổi theo từng giai đoạn phát triển.

Giai đoạn	I	II	III
Trọng lượng cơ thể (kg)	35 - 60	61 - 100	101 - 135
Lizin (%) ¹	5,5	5,0	4,7

- Công thức tính số lượng động vật tham gia thí nghiệm? Không có một công thức nào có thể thỏa mãn đồng thời nhiều điều kiện kể trên; tuy nhiên các nhà khoa học cũng cố gắng đưa ra một số cách xác định dung lượng mẫu. Mitchel và Greendley đã đưa ra công thức để tính dung lượng mẫu như sau:

$$n = \left(\frac{1,849C_v \sqrt{2 + \frac{1}{2}C^2}}{100C} \right)^2$$

- Trong đó
- n - số lượng động vật thí nghiệm cần xác định
 - C_v - hệ số biến động của tính trạng cần nghiên cứu
 - C - Sự sai khác mong nhóm theo một tính trạng nghiên cứu giữa 2 nhóm

Số lượng động vật trong một nhóm (đối với đại gia súc và lợn, giả sử C_v = 17%)

Sự sai khác mong đợi về khối lượng giữa 2 nhóm (%)	Số lượng động vật cần thiết cho một nhóm	Sự sai khác mong đợi về khối lượng giữa 2 nhóm (%)	Số lượng động vật cần thiết cho một nhóm
50	1	12,5	13
40	2	10	20
30	3	7,5	36
20	5	5	80
17,5	7	2,5	317
15	9		

- Giáo sư Arandi đưa ra công thức để xác định dung lượng mẫu như sau:

$$n = 2K^2 \frac{\sigma^2}{D^2_A}$$

- Trong đó n - số lượng động vật

¹ % lizin được tính so với toàn bộ lượng protein thô trong khẩu phần

- σ^2 - phương sai
- D_A - sự sai khác mong đợi giữa 2 nhóm
- K - hệ số

Nếu muốn đạt được mức tin cậy $P = 0,95$ thì giá trị $K = 3,29$

$$n = 2 \times 3,29^2 \frac{\sigma^2}{D_A^2}$$

- Hoặc theo Pearson và Hartley (*không đề cập đến trong khoá học này*), thì ta có thể dùng các đường cong cho sẵn để xác định dung lượng mẫu cần thiết. Trong trường hợp này, dung lượng mẫu sẽ phụ thuộc vào sự sai khác mong đợi giữa các nghiệm thức, mức sai lầm loại I (α) và mức sai lầm loại II (β)

$$\phi^2 = \frac{n \sum_{i=1}^t \tau_i^2}{t \sigma^2}$$

- Trong đó n - số động vật cần thiết cho một nghiệm thức
- t - số nghiệm thức
- τ_i - sai khác mong đợi của nghiệm thức thứ i với μ
- σ^2 - Phương sai của quần thể cần nghiên cứu

Lưu ý rằng, trong một thí nghiệm có tất cả các điều kiện thuận thì số động vật trong một nhóm không thể ít hơn 6-8; trong điều kiện cho phép, số động vật tối thiểu nên thấp nhất là 12 con.

2.3. Xác định các nguồn gây biến động

- Các kiểu biến động - trong một nghiệm thức ta có thể bố trí 1 hoặc nhiều các đơn vị thí nghiệm và giữa các đơn vị thí nghiệm có sự khác nhau thường gọi là sự **biến động**. Nguồn biến động có thể trong quá trình bố trí thí nghiệm ta tác động lên đối tượng nghiên cứu hoặc là những nguồn biến động không thể kiểm soát được. Có các kiểu biến động sau đây:
 - Biến động kiểm soát được
 - Biến động có thể quan sát được
 - Biến động không thể kiểm soát được
- Một số vấn đề khác
 - Xác định loại biến nghiên cứu (định tính, định lượng...)
 - Lựa chọn các công thức thí nghiệm
 - Đơn vị đo
 - Các giá trị thập phân (sau dấu phẩy lấy bao nhiêu số)
 - Kế hoạch thực hiện (địa điểm, ngày, giờ...)
 - Phán đoán các trường hợp đủ do
 - Kế hoạch kiểm soát các công đoạn nghiên cứu.

2.4. Lập sơ đồ thí nghiệm hoặc quan sát

Tùy theo mô hình thí nghiệm hoặc quan sát mà ta có một sơ đồ thí nghiệm thích hợp. Sơ đồ thí nghiệm được vẽ trên giấy hoặc trên máy tính; bao gồm cách phân các đơn vị thí nghiệm vào các công thức khác nhau.

2.5. Lựa chọn mô hình thống kê để phân tích số liệu

Đối với mỗi một mô hình thí nghiệm, ta sẽ chọn ra một mô hình thống kê để phân tích số liệu.

- Với thí nghiệm chỉ có một lô duy nhất (thí nghiệm thức duy nhất), nhằm so sánh với giá trị của quần thể (so sánh một phương thức chăn nuôi mới với phương pháp hiện có hay một khẩu phần mới với khẩu phần hiện có của trại...); ta sử dụng phép thử z nếu biết phương sai quần thể hoặc phép thử t nếu không biết phương sai quần thể.
- Với thí nghiệm có 2 lô (2 thí nghiệm thức); ta sử dụng phép thử z nếu biết phương sai quần thể hoặc phép thử t nếu không biết phương sai quần thể.
- Với thí nghiệm từ 3 lô trở lên (3 thí nghiệm thức trở lên) ta dùng phép phân tích phương sai để so sánh.
- Đối với các thí nghiệm dữ liệu thu thập ở dạng biến định tính ta có thể dùng phép thử χ^2 hoặc phép thử z khi dung lượng mẫu lớn.
- Ngoài ra chúng ta còn có một số mô hình khác như phân tích hồi quy (logistic, binary...), thống kê phi tham số để phân tích số liệu cho phù hợp (sẽ không đề cập đến trong khoá học này).

2.6. Tiến hành thí nghiệm

- Chuẩn bị động vật thí nghiệm
- Bố động vật vào các công thức thí nghiệm

Bố trí động vật vào các công thức thí nghiệm phải theo các nguyên tắc sau đây:

- Ngẫu nhiên - Động vật bố trí vào các công thức thí nghiệm theo nguyên tắc hoàn toàn ngẫu nhiên. Chúng ta có thể sử dụng các phương pháp sau để phân động vật về các khẩu phần một cách ngẫu nhiên:
 - Tung đồng xu (xấp, ngửa)
 - Dùng các quân bài
 - Bảng số ngẫu nhiên (xem phụ lục)
 - Dùng máy tính
- Đồng đều - Đối với động vật trong cùng một nhóm sự sai khác về khối lượng không vượt quá ngưỡng 15% và giữa các nhóm khác nhau không quá 5%.
- Số lượng (tham khảo mục 1.2.2)
- Giai đoạn trước thí nghiệm
 - Giai đoạn cân bằng - nhiệm vụ chính của giai đoạn này là kiểm tra sự đồng đều của các nhóm nghiên cứu. Động vật ở giai đoạn này cho ăn cùng một chế độ và điều kiện chuồng trại như nhau. Thời gian của giai đoạn này phụ thuộc vào yếu tố nghiên cứu, nhưng thông thường không dưới 2 tuần (15 ngày). Trong trường hợp cần thiết chúng ta phải can thiệp để làm cân bằng 2 nhóm. Kiểm tra kỹ lưỡng tình trạng sức khỏe, khả năng cho sản phẩm, phản ứng đối với các yếu tố môi trường bên ngoài...Đặc biệt kiểm tra kỹ lưỡng tình trạng sức khỏe của từng

con vật để kịp thời phát hiện ra một số bệnh truyền nhiễm, bệnh mãn tính, rối loạn tiêu hoá, giun sán; những bệnh này ảnh hưởng rất lớn đến sự tăng trưởng. Trong giai đoạn này có thể chuyển con vật từ nhóm này qua nhóm khác, thậm chí có thể thay mới.

- Giai đoạn thích nghi - thường kéo dài trên một tuần (7 ngày). Mục đích chính của giai đoạn này là từng bước cho động vật làm quen với chế độ thí nghiệm và tránh không làm cho con vật bị ức chế, không làm thay đổi đột ngột điều kiện sống và chế độ ăn uống. Trong giai đoạn này không được chuyển con vật từ lô này qua lô khác cũng như không được thay thế chúng. Tiến hành theo dõi, ghi chép riêng biệt từng con nhưng không đưa các số liệu này vào quá trình xử lý số liệu nghiên cứu. Nên lưu ý rằng giai đoạn này có thể bỏ qua nếu trong giai đoạn cân bằng không có sự chuyển đổi con vật từ lô này qua lô khác, không có sự thay mới và các yếu tố nghiên cứu không đòi hỏi con vật nhiều để thích nghi.
- Giai đoạn thí nghiệm (giai đoạn chính)
 - Trong giai đoạn này không được chuyển động vật từ lô này qua lô khác cũng như không được thay thế chúng. Loại bỏ con vật chỉ trong trường hợp dúi do, Con vật bị loại thải cũng phải được ghi chép nguyên nhân dẫn đến loại thải.
- Giai đoạn sau thí nghiệm.
- Ví dụ mô hình bố trí thí nghiệm trường hợp đơn giản nhất

Giai đoạn trước thí nghiệm		Giai đoạn chính
Điều kiện chăn nuôi thực tế	Điều kiện chăn nuôi thực tế + từng bước đưa yếu tố nghiên cứu vào	Điều kiện chăn nuôi thực tế + yếu tố nghiên cứu

2.7. Thu thập số liệu

Tuỳ theo từng thí nghiệm cụ thể mà tiến hành thu thập số liệu. Có thể số liệu thu thập vào nhiều thời điểm khác nhau, cũng có thể thu thập ngay sau khi kết thúc thí nghiệm.

2.8. Phân tích số liệu

Đối với các mô hình thí nghiệm được thiết kế thoả mãn các điều kiện nêu trên thì không gặp nhiều khó khăn trong quá trình phân tích xử lý. Chú ý trong quá trình thu thập số liệu không thể tránh khỏi sự sai số hoặc có những số liệu không điển hình cho toàn bộ các quan sát.

- Các sai lầm hay mắc phải
 - Do con người gây ra
 - Do sai số của các dụng cụ
- Phát hiện các số liệu không điển hình
 - Ta có thể phát hiện các số liệu không điển hình theo nguyên tắc $\mu \pm 3\sigma$
 - Hoặc bằng cách mô tả số liệu để phát hiện ra những giá trị không bình thường
- Loại bỏ những giá trị không điển hình
 - Những số liệu không điển hình có thể loại bỏ trước khi tiến hành phân tích, nhưng cũng cần lưu ý rằng trước khi loại bỏ ta cần phải tìm hiểu nguyên nhân tại sao. Có thể chính những giá trị này sẽ cho ta biết một số thông tin quan trọng liên quan đến thí nghiệm.

2.9. Viết báo cáo (phần này sẽ đề cập ở cuối kỳ)

3. Bố trí thí nghiệm 1 nhân tố

3.1. Thí nghiệm kiểu hoàn toàn ngẫu nhiên

3.1.1. Giới thiệu

Đây là phương pháp nghiên cứu cơ bản và tổng hợp trong chăn nuôi, được sử dụng để bố trí thí nghiệm khi chỉ nghiên cứu một yếu tố thí nghiệm; ví dụ ta nghiên cứu ảnh hưởng của thức ăn đến tăng trọng, tồn dư thuốc kháng sinh trong cơ thể vật nuôi...

Ở đây chúng ta chỉ xem xét đến một yếu tố thí nghiệm còn các yếu tố phi thí nghiệm khác ta giả sử rằng không có sự sai khác có hệ thống giữa các đơn vị thí nghiệm. Ví dụ tất cả các con vật được chọn cùng một lứa tuổi, tất cả các trại đều sử dụng các thức ăn như nhau...

Đối với thí nghiệm này, các đơn vị thí nghiệm được bố trí một cách ngẫu nhiên vào các nghiệm thức (công thức thí nghiệm).

3.1.2. Ví dụ

Tiến hành một thí nghiệm để so sánh mức độ tăng trọng ở 3 khẩu phần ăn khác nhau (khẩu phần 1, 2, 3 tương ứng với các chữ cái A, B, C) trên 15 đơn vị thí nghiệm (động vật thí nghiệm).

Đây là một ví dụ về thí nghiệm có 1 yếu tố thí nghiệm (khẩu phần ăn) với $t = 3$ nghiệm thức (khẩu phần 1, 2 và 3) và $n = 5$ lần lặp lại trong mỗi nghiệm thức (khẩu phần)

3.1.3. Xây dựng sơ đồ thí nghiệm và bố trí động vật vào các công thức thí nghiệm

- **Vẽ sơ đồ thí nghiệm** - Đối với các thí nghiệm 1 nhân tố ta sẽ xây dựng một bảng có số cột bằng chính số nghiệm thức và số hàng bằng số động vật trong một nghiệm thức + 1.

Đối với ví dụ nêu ở phần trên ta có thể xây dựng bảng sơ đồ thí nghiệm như sau:

- **Đánh số cho từng động vật thí nghiệm** (nếu động vật chưa được đánh số) - Mục đích chính của việc đánh số là thuận tiện cho việc chia lô cũng như trong suốt quá trình theo dõi từng cá thể.

Ở ví dụ trên, giả sử 15 động vật được đánh số từ số 1 đến 15

- **Phân động vật về các nghiệm thức** theo phương thức hoàn toàn ngẫu nhiên. Ta sẽ tiến hành phân 15 động vật trên về 3 khẩu phần ăn khác nhau, mỗi khẩu phần sẽ có 5 động vật. Chúng ta có thể sử dụng các phương pháp sau để phân động vật về các khẩu phần một cách ngẫu nhiên:
 - Tung đồng xu (xấp, ngửa)
 - Bảng số ngẫu nhiên
 - Dùng máy tính

Ví dụ ta sử dụng bảng số ngẫu nhiên, thí nghiệm với 3 khẩu phần khác nhau (A, B, C)

- Từ bảng số ngẫu nhiên ta chọn dòng đầu tiên từ trên xuống: 8, 4, 8, 7, ...
- Số ngẫu nhiên từ 1 đến 3 chọn khẩu phần A, số ngẫu nhiên từ 4 đến 6 chọn khẩu phần B, số ngẫu nhiên từ 7 đến 9 chọn khẩu phần C và số ngẫu nhiên 0 bỏ qua.

Đơn vị thí nghiệm số	Số ngẫu nhiên	Cách xử lý	
1	8	C	
2	4	B	
3	8	C	
4	7	C	
5	1	A	
6	6	B	
7	1	A	
8	5	B	
9	3	A	
10	1	A	
11	9	C	
12	5	B	
13	1	A	(5A)
14	4	B	(5B)
15	2	C	(5C)

Sau đó bố trí động vật vào các khẩu phần tương ứng như sau:

Khẩu phần 1	Khẩu phần 2	Khẩu phần 3
5	2	1
7	6	3
9	8	4
10	12	11
13	14	15

• **Áp dụng Minitab để bốc thăm ngẫu nhiên**

```

MTB > SET C1
MTB > SAMPLE 15 C1 C2
MTB > SET C3
MTB > PRINT C1-C3
    
```

Calc > Make Patterned Data > Arbitrary Set of Numbers...
Calc > Random Data... / Sample from Columns...
Calc > Make Patterned Data > Arbitrary Set of Numbers...
Manip > Display Data...

Display Data

ROW	UNIT	RANDOM	GROUP
1	1	5	1
2	2	13	1
3	3	11	1
4	4	1	1
5	5	15	1
6	6	3	2
7	7	2	2
8	8	6	2
9	9	14	2
10	10	7	2
11	11	10	3
12	12	12	3
13	13	9	3
14	14	4	3
15	15	8	3

```
MTB > UNSTACK C2 C4-C6 ;           Manip > Unstack Columns...
SUBC> SUBS C3 .
MTB > NAME C4 'TREAT A' C5 'TREAT B' C6 'TREAT C'
MTB > PRINT C4-C6                 Manip > Display Data...
```

Display Data

ROW	TREAT A	TREAT B	TREAT C
1	5	3	10
2	13	2	12
3	11	6	9
4	1	14	4
5	15	7	8

các đơn vị thí nghiệm được chia thành các thí nghiệm khác nhau

- Đến đơn vị thứ 13 thì ta đã có đầy đủ 5 đơn vị thí nghiệm với khẩu phần A, và thứ 14 với B. Như vậy đơn vị thí nghiệm 15 phải nhận được khẩu phần C.

3.1.4. Mô hình phân tích

- Nếu số liệu có phân bố chuẩn và các phương sai đồng nhất thì
 - thí nghiệm có 2 nghiệm thức → phép thử t (t -test)
 - Thí nghiệm ≥ 3 nghiệm thức → phép phân tích phương sai 1 yếu tố
- Nếu số liệu không thoả mãn 2 điều kiện nêu trên chúng ta phải tiến hành biến đổi số liệu để phân tích hoặc dùng phép kiểm định phi tham số (sẽ không đề cập đến trong khoá học này).

Mô hình toán học mô tả các quan sát đối với thí nghiệm có t nghiệm thức và trong mỗi nghiệm thức có n_i quan sát và $N = n_1 + n_2 + \dots + n_t$

$y_{ij} = \mu + \varepsilon_{ij}$ hay tương đương với $y_{ij} = \mu_i + \tau_i + \varepsilon_{ij}$

- Trong đó
- $i = 1, 2, \dots, t$
 - $j = 1, 2, \dots, n_i$
 - y_{ij} - quan sát j ở nghiệm thức i
 - μ_i - giá trị trung bình của nghiệm thức i
 - μ - giá trị trung bình của toàn bộ các quan sát
 - τ_i - hiệu quả của nghiệm thức thứ i
 - ε_{ij} - sai số ngẫu nhiên của quan sát j ở nghiệm thức i

Bảng số liệu tổng quát đối với thí nghiệm một nhân tố (dạng tổng quát từ ví dụ 1.2.1)

	Nghiệm thức (t)			
	1	2	...	t
Các quan sát (n)	y_{11}	y_{21}	...	y_{t1}
	y_{12}	y_{22}	...	y_{t2}

	y_{1n1}	y_{2n2}	...	y_{nt}
Trung bình	-	-	...	-
	$y_{1.}$	$y_{2.}$...	$y_{t.}$
				$y_{..}$

Ta biết trong thí nghiệm này có 2 nguồn biến động (xem chi tiết trang 38, phần I của bài giảng)

- Biến động trong từng nhóm (ngẫu nhiên) có thể khái quát dưới dạng mô hình như sau:

$$SS_{\text{ngẫu nhiên}} = \sum_{i=1}^t \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_i \right)^2 .$$

- Biến động giữa các nhóm (nghiệm thức) có thể khái quát dưới dạng mô hình như sau:

$$SS_{\text{nghiệm thức}} = n \sum \left(\bar{y}_i - \bar{y}_{..} \right)^2 .$$

- Toàn bộ các biến động của thí nghiệm chính bằng tổng của *biến động ngẫu nhiên* và *biến động của nghiệm thức*; ta có thể khái quát bằng công thức sau:

$$\sum_{i=1}^t \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_{..} \right)^2 = n \sum \left(\bar{y}_i - \bar{y}_{..} \right)^2 + \sum_{i=1}^t \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_i \right)^2$$

Xây dựng cấu trúc của bảng phân tích phương sai

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	Giá trị F quan sát
Nghiệm thức	$t - 1$	$SS_{\text{nghiệm thức}}$	$SS_{\text{nghiệm thức}}/(t-1)$	$SS_{\text{nghiệm thức}}/(t-1)$
Sai số ngẫu nhiên	$N - t$	$SS_{\text{ngẫu nhiên}}$	$SS_{\text{ngẫu nhiên}}/(N-t)$	$SS_{\text{ngẫu nhiên}}/(N-t)$
Tổng biến động	$N - 1$	$SS_{\text{tổng số}}$		

Giá trị F lý thuyết được xác định ở bảng phân phụ lục với mức xác suất sai số α và bậc tự do $v_1 = t - 1$ và $v_2 = N - t$. Nếu $P \geq 0,05$ ta chấp nhận H_0 , và bác bỏ H_0 nếu $P < 0,05$.

Ví dụ (trang 42, phần I của bài giảng)

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

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

Lời giải:

- Lập bảng phân tích phương sai

Nguồn biến động	Bậc tự do (<i>df</i>)	Tổng bình phương (SS)	Trung bình bình phương (MS)	Giá trị <i>F</i> quan sát
Khẩu phần	3	16.467	5.489	6,65
Sai số ngẫu nhiên	16	13.212	826	
Tổng biến động	19	29.679		

Áp dụng phần mềm *Minitab* cũng cho ta kết quả tương tự

One-way ANOVA: P versus KP				Stat > ANOVA > One-way...	
Analysis of Variance for P					
Source	DF	SS	MS	F	P
KP	3	16467	5489	6.65	0.004
Error	16	13212	826		
Total	19	29679			
Individual 95% CIs For Mean Based on Pooled StDev					
Level	N	Mean	StDev	-----+-----+-----+-----	
KP1	5	79.00	24.47	(-----*-----)	
KP2	5	71.00	31.02	(-----*-----)	
KP3	5	81.40	22.88	(-----*-----)	
KP4	5	142.80	34.90	(-----*-----)	
-----+-----+-----+-----					
Pooled StDev =		28.74		70	105 140
MTB >					

Với $P = 0,004$, giả thuyết H_0 bị bác bỏ hay nói cách khác hoàn toàn có thể loại bỏ giả thuyết rằng tăng trọng trung bình giữa các khẩu phần ăn là bằng nhau.

- So sánh từng cặp các giá trị trung bình của nghiệm thức

Giả sử ta muốn so sánh tất cả các cặp của t nghiệm thức với nhau và giả thiết H_0 trong các phép thử này là $H_0 : \mu_1 = \mu_2 (i \neq j)$. Trong ví dụ trên ta có 6 cặp cần phải so sánh.

Có 4 phương pháp để tiến hành so sánh từng cặp các giá trị trung bình với nhau:

- Sự sai khác bé nhất có ý nghĩa (LSD)
- Phép so sánh Duncan
- Phép so sánh Newman - Keuls
- Phép so sánh Tukey

Trong khuôn khổ khoá học sẽ không đề cập đến cách tính toán cụ thể, phần mềm *Minitab* được sử dụng để tiến hành các phép thử này

Tukey's pairwise comparisons

Family error rate = 0.0500
Individual error rate = 0.0113

Critical value = 4.05

Intervals for (column level mean) - (row level mean)

	1	2	3
2	-44.0 60.0		
3	-54.4 49.6	-62.4 41.6	
4	-115.8 -11.8	-123.8 -19.8	-113.4 -9.4

Qua Output trên ta thấy *Minitab*:

- Đã thực hiện phép thử *Tukey* để so sánh từng cặp với nhau
- Sai số của toàn bộ 6 phép thử là $P = 0,05$
- Sai số của từng phép thử là $P = 0,0113$
- *Minitab* cho ta một ma trận 3×3 thể hiện từng cặp so sánh. Nếu hiệu số giữa khoảng tin cậy 95% của các nghiệm thức có chứa số không, điều này chứng tỏ trung bình của 2 tổng thể đó bằng nhau ($P \geq 0,05$) và ngược lại nếu khoảng này không chứa số không thì trung bình của 2 tổng thể không bằng nhau. Trong output trên ta thấy hiệu số của khoảng tin cậy 95% của cặp so sánh thứ nhất giữa kp1 và 2 là (-44 ; +60) có chứa số 0; chứng tỏ $\mu_1 = \mu_2$.

Bạn đọc có thể tìm hiểu chi tiết hơn ở phần I trang 42 - 51 về phân tích phương sai và so sánh cặp đôi từng nghiệm thức ở phần I trang 49 - 51. Để tiện theo phần so sánh cặp đôi được trình dưới đây:

Ví dụ: So sánh tăng trọng của chuột ở 4 khẩu phần ăn khác nhau (khẩu phần 1, 2, 3 và 4). Số chuột tham gia vào thí nghiệm vào từng khẩu phần là 7, 8, 6 và 8. Số liệu thu được trình bày ở bảng sau (% tăng trọng so với khối lượng cơ thể):

	1	2	3	4
	3,42	3,17	3,34	3,64
	3,96	3,63	3,72	3,93
	3,87	3,38	3,81	3,77
	4,19	3,47	3,66	4,18
	3,58	3,39	3,55	4,21
	3,76	3,41	3,51	3,88
	3,84	3,55		3,96
		3,44		3,91

Bài giải: (Dùng phần mềm *Minitab* để giải quyết).

Nhập số liệu vào *Minitab*, tính các tham số thống kê mô tả ta thu được kết quả sau:

Variable	N	Mean	Median	TrMean	StDev	SE Mean
1	7	3.8029	3.8400	3.8029	0.2512	0.0949
2	8	3.4300	3.4250	3.4300	0.1353	0.0478
3	6	3.5983	3.6050	3.5983	0.1675	0.0684
4	8	3.9350	3.9200	3.9350	0.1906	0.0674

Variable	Minimum	Maximum	Q1	Q3
1	3.4200	4.1900	3.5800	3.9600
2	3.1700	3.6300	3.3825	3.5300
3	3.3400	3.8100	3.4675	3.7425
4	3.6400	4.2100	3.7975	4.1250

Giả thiết $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ (bằng lời, bạn đọc tự nêu)

$H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$

So sánh sự đồng nhất của phương sai: $0,2512 / 0,1353 = 1,86 < 2$

Kiểm tra phân bố chuẩn: bằng cách kiểm tra phân bố chuẩn của sai số ngẫu nhiên (phần dư). Đây là một thí nghiệm mà số động vật tham gia vào từng công thức thí nghiệm hạn chế ($n_1 = 7, n_2 = 8, n_3 = 6$ và $n_4 = 8$), vì vậy ta không kiểm tra phân bố chuẩn của từng biến riêng biệt. dùng *Minitab* để kiểm tra phân bố chuẩn ta có $P = 0,55$

Phân tích phương sai

Analysis of Variance for P					
Source	DF	SS	MS	F	P
KP	3	1.1601	0.3867	10.73	0.000
Error	25	0.9012	0.0360		
Total	28	2.0613			

Individual 95% CIs For Mean Based on Pooled StDev					
Level	N	Mean	StDev		
1	7	3.8029	0.2512	-----+-----+-----+----- (-----*-----)	
2	8	3.4300	0.1353	(-----*-----)	
3	6	3.5983	0.1675	(-----*-----)	
4	8	3.9350	0.1906	(-----*-----)	

Pooled StDev =	0.1899	3.50	3.75	4.00
----------------	--------	------	------	------

Kết luận

Vì $P = 0,000 < 0,05$ ta bác bỏ H_0 và chấp nhận H_1 (bằng lời, bạn đọc tự nêu)

So sánh từng cặp.

Dùng menu *Comparisons* của *Minitab* ta có

Tukey's pairwise comparisons			
Family error rate = 0.0500			
Individual error rate = 0.0109			
Critical value = 3.89			
Intervals for (column level mean) - (row level mean)			
	1	2	3
2	0.1026 0.6431		
3	-0.0860 0.4951	-0.4504 0.1137	
4	-0.4024 0.1381	-0.7661 -0.2439	-0.6187 -0.0546

Nếu nhìn vào Ma trận trên ta thấy $\mu_1 \neq \mu_2, \mu_1 = \mu_3, \mu_1 = \mu_4, \mu_2 = \mu_3, \mu_2 \neq \mu_4, \mu_3 \neq \mu_4$.

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

Thực hiện theo các bước sau:

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

Khẩu phần	Trung bình		Khẩu phần	Trung bình
1	3,8029	→	4	3,9350
2	3,4300		1	3,8029
3	3,5983		3	3,5983
4	3,9350		2	3,4300

- Dựa vào ma trận đã nêu ở mục 6 để tạo các đường gạch chung cho các khẩu phần có giá trị trung bình bằng nhau; cụ thể như sau:

Khẩu phần	Trung bình	
4	3,9350	a
1	3,8029	b
3	3,5983	c
2	3,4300	

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

- Từ mục b, ta có thể đặt các chữ cái bên cạnh các số trung bình như sau:

Khẩu phần	Trung bình
4	3,9350 ^a
1	3,8029 ^{ab}
3	3,5983 ^{bc}
2	3,4300 ^c

- Sắp xếp khẩu phần theo thứ tự tăng dần như ban đầu (ở mục 6.a.) ta có

Khẩu phần	Trung bình
1	3,8029 ^{ab}
2	3,4300 ^c
3	3,5983 ^{bc}
4	3,9350 ^a

- Kiểm tra phân bố chuẩn của số liệu

Muốn thực hiện phép *phân tích phương sai* ta phải thoả mãn một loạt các điều kiện, trong đó có điều kiện số liệu quan sát *phải tuân theo phân bố chuẩn* $y \sim N(\mu, \sigma^2)$.

Ta có thể biểu diễn các số liệu trên đồ thị của từng nghiệm thức để xác định xem số liệu có phân bố chuẩn hay không; tuy nhiên điều này rất khó thực hiện khi số lượng đơn vị trong từng nghiệm thức bị hạn chế. Như trong ví dụ trên ta thấy trong mỗi nghiệm thức chỉ có 5 động vật. Để khắc phục hạn chế này ta đưa ra cách thử như sau:

Có thể mô tả số liệu dưới mô hình sau:

$$y_{ij} = \mu_i + \varepsilon_{ij} \rightarrow \varepsilon_{ij} = y_{ij} - \mu_i$$

Tức là ta thay phép kiểm tra số liệu từ $y \sim N(\mu, \sigma^2)$ thành kiểm tra $\varepsilon \sim N(0, \sigma^2)$.

Điều này có thể minh hoạ bằng output của *Minitab*.

Row	p	kp	RESI1	FITS1
1	99	1	20.0	79.0
2	88	1	9.0	79.0
3	76	1	-3.0	79.0
4	38	1	-41.0	79.0
5	94	1	15.0	79.0
6	61	2	-10.0	71.0
7	112	2	41.0	71.0
8	30	2	-41.0	71.0
9	89	2	18.0	71.0
10	63	2	-8.0	71.0

Row	p	kp	RESI1	FITS1
11	42	3	-39.4	81.4
12	97	3	15.6	81.4
13	81	3	-0.4	81.4
14	95	3	13.6	81.4
15	92	3	10.6	81.4
16	169	4	26.2	142.8
17	137	4	-5.8	142.8
18	169	4	26.2	142.8
19	85	4	-57.8	142.8
20	154	4	11.2	142.8

- Những hạn chế của mô hình thí nghiệm kiểu hoàn toàn ngẫu nhiên

Mô hình này đòi hỏi tất cả các đơn vị thí nghiệm phải tương tự nhau như trước khi tiến hành thí nghiệm. Tuy nhiên trong thực tế rất khó có thể thực hiện được điều kiện này và nếu có một sự không đồng nhất giữa các đơn vị thí nghiệm mà ta bỏ qua yếu tố này thì mắc phải 5 sai lầm sau:

- Trung bình bình phương của sai số ngẫu nhiên sẽ lớn
- Giá trị F thực nghiệm sẽ bé
- Giá trị P thực nghiệm sẽ lớn
- Ít cơ may hơn để phát hiện sự sai khác
- Đi đến kết luận thiếu chính xác

Điều này được thể hiện rõ qua sơ đồ bố trí các nghiệm thức vào chuồng nuôi

Hướng tây	1 C	2 B	3 C
	4 C	5 A	6 B
	7 A	8 B	9 A
	10 A	11 C	12 B
	13 A	14 B	15 C

Ta thấy ô chuồng ở hướng tây chỉ có các khẩu phần A và C; vì vậy không thể biết chắc chắn rằng mức tăng trọng là do khẩu phần gây nên hay là hướng của chuồng gây nên.

3.2. Thí nghiệm kiểu khối ngẫu nhiên đầy đủ

3.2.1. Giới thiệu

Tiến hành thiết kế thí nghiệm theo kiểu khối hoàn toàn ngẫu nhiên khi có sự sai khác có hệ thống giữa các đơn vị thí nghiệm. Ở đây tất cả các đơn vị thí nghiệm có chung một đặc tính được bố trí vào một nhóm thường được gọi là **khối**. Ở đây ta chỉ xem xét đến một yếu tố thí nghiệm duy nhất, còn yếu tố **khối** là yếu tố phi thí nghiệm và chỉ đưa vào để làm giảm biến động do yếu tố phi thí nghiệm gây ra.

3.2.2. Xây dựng sơ đồ thí nghiệm và bố trí động vật vào các công thức thí nghiệm

- **Vẽ sơ đồ thí nghiệm** - Xây dựng một bảng có số cột bằng chính số khối và số hàng bằng số nghiệm thức và trong mỗi khối bố trí ngẫu nhiên công thức thí nghiệm vào mỗi ô chuồng.

Đối với ví dụ nêu ở phần trên ta có thể xây dựng bảng sơ đồ thí nghiệm như sau:

Khối 1	Khối 2	Khối 3

- **Đánh số cho từng động vật thí nghiệm** (xem 1.3.1.3).
- **Phân động vật về các khối và nghiệm thức** (xem 1.3.1.3)

Chú ý rằng, trong mỗi khối các nghiệm thức được bố trí ngẫu nhiên vào các ô chuồng

Hướng tây

Khối 1	Khối 2	Khối 3
B	C	C
C	B	B
A	A	A

3.2.3. Mô hình phân tích

Đối với thí nghiệm kiểu này, phải sử dụng mô hình **phân tích phương sai** để phân tích số liệu

Mô hình toán học mô tả các quan sát với thí nghiệm có t nghiệm thức và b khối như sau

$$y_{ijk} = \mu + \varepsilon_{ijk} \text{ hay tương đương với } y_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk}$$

Trong đó $i = 1, 2, \dots, t$

$j = 1, 2, \dots, b$

$k = 1, 2, \dots, n_{ij}$

y_{ijk} - quan sát k ở nghiệm thức i và khối thứ j

μ_i - giá trị trung bình của nghiệm thức i

μ - giá trị trung bình của toàn bộ các quan sát

τ_i - tác động của nghiệm thức thứ i

β_j - tác động của khối thứ j

ε_{ijk} - sai số ngẫu nhiên của quan sát thứ k ở nghiệm thức i và khối thứ j

- Trong thí nghiệm kiểu khối hoàn toàn ngẫu nhiên có 3 nguồn biến động đó là
 - Biến động trong từng nhóm (ngẫu nhiên)

$$SS_{\text{ngẫu nhiên}} = \sum_{k=1}^{n_{ij}} \sum_{i=1}^t \sum_{j=1}^b \left(y_{ijk} - \bar{y}_{.j} - \bar{y}_{i.} - \bar{y}_{..} \right)^2$$

- Biến động giữa các nhóm (thí nghiệm)

$$SS_{\text{thí nghiệm}} = b \sum_{i=1}^t \left(\bar{y}_{i.} - \bar{y}_{..} \right)^2$$

- Biến động giữa các khối

$$SS_{\text{khối}} = t \sum_{j=1}^b \left(\bar{y}_{.j} - \bar{y}_{..} \right)^2$$

- Toàn bộ các biến động của thí nghiệm chính bằng tổng của *biến động ngẫu nhiên*, *biến động giữa các khối* và *biến động của thí nghiệm*; ta có thể khái quát bằng công thức sau:

$$\begin{aligned} & \sum_{k=1}^{n_{ij}} \sum_{i=1}^t \sum_{j=1}^b \left(y_{ijk} - \bar{y}_{..} \right)^2 = \\ & = b \sum_{i=1}^t \left(\bar{y}_{i.} - \bar{y}_{..} \right)^2 + t \sum_{j=1}^b \left(\bar{y}_{.j} - \bar{y}_{..} \right)^2 + \sum_{k=1}^{n_{ij}} \sum_{i=1}^t \sum_{j=1}^b \left(y_{ijk} - \bar{y}_{.j} - \bar{y}_{i.} - \bar{y}_{..} \right)^2 \end{aligned}$$

Xây dựng cấu trúc của bảng phân tích phương sai

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	Giá trị F quan sát
Thí nghiệm	t - 1	SS _{thí nghiệm}	SS _{thí nghiệm} /(t-1)	SS _{thí nghiệm} /(t-1)
				SS _{ngẫu nhiên} /(t-1)(b-1)
Khối	b - 1	SS _{khối}	SS _{khối} /(b-1)	
Sai số ngẫu nhiên	nbt - t - b + 1	SS _{ngẫu nhiên}	SS _{ngẫu nhiên} /(t-1)(b-1)	
Tổng biến động	N - 1	SS _{tổng số}		

Giá trị F lý thuyết được xác định ở bảng phân phụ lục với mức xác suất sai số α và bậc tự do $v_1 = t - 1$ và $v_2 = nbt - t - b + 1$. Nếu $P \geq 0,05$ ta chấp nhận H_0 , và bác bỏ H_0 nếu $P < 0,05$.

Ví dụ: (Mead và cộng sự) Nghiên cứu số lượng tế bào lymphô ở chuột ($\times 1000$ tế bào mm^{-3} máu) được sử dụng 4 loại thuốc khác nhau ở qua 5 lứa; số liệu thu được như sau:

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

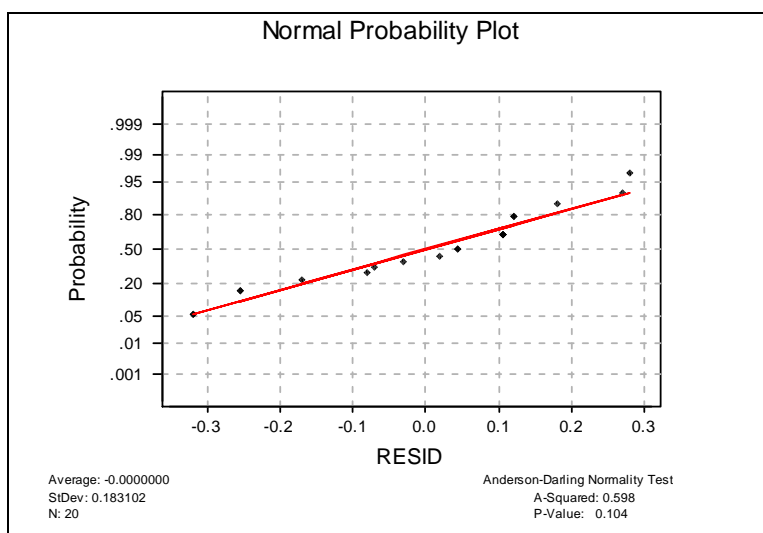
Áp dụng *Minitab*

Analysis of Variance for Tebao

Source	DF	SS	MS	F	P
lua	4	6.4030	1.6008	30.16	0.000
thuoc	3	1.8455	0.6152	11.59	0.001
Error	12	0.6370	0.0531		
Total	19	8.8855			

Qua kết quả bảng trên ta thấy có sự khác nhau về số lượng tế bào lymphô sau khi được xử lý bằng các loại thuốc khác nhau

Để so sánh sự sai khác của từng cặp công thức và kiểm tra phân bố chuẩn của số liệu (xem trang 14)



Kiểm tra phân bố chuẩn của số liệu $y \sim N(\mu, \sigma^2)$ thông qua việc kiểm tra phân sai số ngẫu nhiên $\varepsilon \sim N(0, \sigma^2)$ bằng phần mềm *Minitab*, thấy rằng điều kiện số liệu có phân bố chuẩn được thoả mãn.

3.3. Thí nghiệm kiểu ô vuông La tinh

3.3.1. Giới thiệu

Mô hình khối hoàn toàn ngẫu nhiên đã được giới thiệu nhằm khắc phục những hạn chế của mô hình hoàn toàn ngẫu nhiên, khi mà ta muốn có sự đồng nhất trong khối, ví dụ nhiệt độ như nhau trong một khu trường. Đôi khi chúng ta cần quan tâm đến sự gia tăng của cả 2 hướng, ví dụ trong mô hình *khối hoàn toàn ngẫu nhiên*, chúng ta thấy trong hàng thứ nhất chỉ có thí 2 nghiệm thức B và C và dòng thứ 3 chỉ có A.

Hướng tây

Khối 1	Khối 2	Khối 3
B	C	C
C	B	B
A	A	A

Đôi khi bố trí thí nghiệm do số lượng động vật tham gia thí nghiệm với số lượng bị hạn chế (thí nghiệm trên đại gia súc) dẫn đến tình trạng không đồng đều giữa các động vật; trong quá trình tiến hành thí nghiệm kéo dài điều kiện khí hậu cũng thay đổi theo; thêm vào đó các công thức thí nghiệm cũng ở nhiều mức độ khác nhau. Để khắc phục những hạn chế nêu trên ta hãy xem xét đến thiết kế thí nghiệm theo kiểu *Ô vuông Latin*.

Ví dụ: Thiết kế thí nghiệm kiểu ô vuông Latin với 4 nghiệm thức

C	A	D	B
B	D	A	C
A	B	C	D
D	C	B	A

Lưu ý rằng trong thí nghiệm kiểu ô vuông Latin:

- Mỗi nghiệm thức được áp dụng một lần trong một hàng
- Mỗi nghiệm thức được áp dụng một lần trong một cột
- Sự sắp xếp của các nghiệm thức trong mỗi dòng và mỗi cột là ngẫu nhiên
- Số ô cần thiết = (Số nghiệm thức)²
- Khi xem xét theo dòng hoặc theo cột thì chúng ta thấy tương tự như mô hình khối hoàn toàn ngẫu nhiên
- Số lượng động vật trong nhóm là bội số của số nghiệm thức, ví dụ số nghiệm thức là 3 thì số động vật trong nhóm là 3, 6, 9, 12,...
- Các nhóm được tạo ra theo phương pháp rút mẫu ngẫu nhiên
- Số hàng = số cột = số nghiệm thức
- Tất cả các động vật tham gia thí nghiệm phải được giữ lại đến hết thí nghiệm (nếu không trong quá trình xử lý số liệu sẽ gặp nhiều khó khăn)

Ví dụ: Sản lượng sữa của bò

- Yếu tố thí nghiệm: Khẩu phần ăn (4 khẩu phần)
- Kiểm soát đối với: Từng con bò (4 bò) + từng mùa trong năm (4 mùa)

	C	A	D	B
Mùa	B	D	A	C
	A	B	C	D
	D	C	B	A

Bò

Như vậy mỗi con bò sẽ nhận được tất cả 4 nghiệm thức (A, B, C và D). Đây là mô hình thí nghiệm rất kinh tế khi bị hạn chế bởi số lượng động vật .

3.3.2. Bố trí các nghiệm thức vào ô thí nghiệm đối với mô hình ô vuông Latinh

Giả sử ta cần có mô hình ô vuông Latinh 4×4 , tương ứng với $t = 4$ thí nghiệm A, B, C , và D .

Bước 1: Điền các nghiệm thức cần thiết vào ô Latinh bất kỳ. Các thí nghiệm này có thể điền vào theo một bố trí có hệ thống hoặc là được chọn một cách ngẫu nhiên. Một trong những kiểu bố trí thí nghiệm theo mô hình ô vuông Latinh có thể như sau:

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

Bước 2: Chọn ngẫu nhiên các thí nghiệm với các chữ cái ở trong ô vuông. Sử dụng bảng số ngẫu nhiên ta có thể nhận được

$$a \rightarrow C \quad b \rightarrow D \quad c \rightarrow A \quad d \rightarrow B$$

Sử dụng sự ngẫu nhiên này ta có

C	D	A	B
D	A	B	C
A	B	C	D
B	C	D	A

3.3.3. Mô hình phân tích

Mô hình mô tả các quan sát

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \varepsilon_{ijk}$$

Trong đó $i = 1, 2, \dots, t$

$j = 1, 2, \dots, t$

$k = 1, 2, \dots, t$

y_{ijk} - quan sát ở hàng thứ i cột thứ k và nghiệm thức thứ j

μ - giá trị trung bình của toàn bộ các quan sát

α_i - tác động của hàng thứ i

τ_j - tác động của nghiệm thức thứ j

β_j - tác động của cột thứ k

ε_{ijk} - sai số ngẫu nhiên của quan sát ở cột thứ k ở nghiệm thức j và hàng thứ j

$$SS_{\text{toàn bộ}} = SS_{\text{hàng}} + SS_{\text{cột}} + SS_{\text{nghiệm thức}} + SS_{\text{ngẫu nhiên}}$$

với bậc tự do $(p^2 - 1) = (p - 1) + (p - 1) + (p - 1) + (p - 2)(p - 1)$

- Trong thí nghiệm kiểu ô vuông latin có 4 nguồn biến động đó là

- Biến động trong từng nhóm (ngẫu nhiên)

$$SS_{\text{ngẫu nhiên}} = SS_{\text{toàn bộ}} - SS_{\text{hàng}} - SS_{\text{cột}} - SS_{\text{thí nghiệm}}$$

- Biến động giữa các nhóm (thí nghiệm)

$$SS_{\text{thí nghiệm}} = \frac{1}{t} \sum_{j=1}^t y_{.j}^2 - \frac{y_{\dots}^2}{N}$$

- Biến động giữa các hàng

$$SS_{\text{hàng}} = \frac{1}{t} \sum_{i=1}^t y_{i.}^2 - \frac{y_{\dots}^2}{N}$$

- Biến động giữa các cột

$$SS_{\text{cột}} = \frac{1}{t} \sum_{k=1}^t y_{.k}^2 - \frac{y_{\dots}^2}{N}$$

- Toàn bộ các biến động của thí nghiệm:

$$SS_{\text{toàn bộ}} = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{\dots}^2}{N}$$

Xây dựng cấu trúc của bảng phân tích phương sai

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	Giá trị F quan sát
Thí nghiệm	$t - 1$	$SS_{\text{thí nghiệm}}$	$SS_{\text{thí nghiệm}}/(t-1)$	$SS_{\text{thí nghiệm}}/(t-1)$
				$SS_{\text{ngẫu nhiên}}/(t-2)(t-1)$
Hàng	$t - 1$	$SS_{\text{hàng}}$	$SS_{\text{hàng}}/(t-1)$	
Cột	$t - 1$	$SS_{\text{cột}}$	$SS_{\text{cột}}/(t-1)$	
Sai số ngẫu nhiên	$(t-2)(t-1)$	$SS_{\text{ngẫu nhiên}}$	$SS_{\text{ngẫu nhiên}}/(t-2)(t-1)$	
Tổng biến động	$t^2 - 1$	$SS_{\text{tổng số}}$		

Giá trị F lý thuyết được xác định ở bảng phân phụ lục với mức xác suất sai số α và bậc tự do $v_1 = t - 1$ và $v_2 = (t-2)(t-1)$. Nếu $P \geq 0,05$ ta chấp nhận H_0 , và bác bỏ H_0 nếu $P < 0,05$.

Ví dụ: (Mead và cộng sự, tr. 72) Nghiên cứu ảnh hưởng của thức ăn mùa đông đến sản lượng sữa theo mô hình ô vuông latin. Có 4 khẩu phần ăn khác nhau (A, B, C, D), cho mỗi bò ăn từng khẩu phần trong 3 tuần. Sản lượng sữa chỉ được tính tổng cộng trong tuần thứ 3 của mỗi chu kỳ. Sản lượng sữa được ghi lại như sau (đơn vị tính pound)

		Bò				Tổng số
		1	2	3	4	
Giai đoạn	1	A 192	B 195	C 292	D 249	928
	2	B 190	D 203	A 218	C 210	821
	3	C 214	A 139	D 245	B 163	761
	4	D 221	C 152	B 204	A 134	711
Tổng số		817	869	959	756	3221

Trong ví dụ này ta thấy có các nguồn biến động sau đây:

- Do khẩu phần ăn gây nên (A, B, C, D)
- Do các giai đoạn gây nên (4 giai đoạn, mỗi giai đoạn kéo dài 3 tuần)
- Do các động vật gây nên (4 bò)
- Do các yếu tố ngẫu nhiên gây nên

Và ta thấy cách tính bậc tự do cũng khác. Chỉ với 4 bò nhưng số bậc tự do của ta được tính như sau:

- Bậc tự do của khẩu phần $df_{\text{khẩu phần}} = 4 - 1 = 3$
- Bậc tự do của giai đoạn $df_{\text{giai đoạn}} = 4 - 1 = 3$
- Bậc tự do của động vật $df_{\text{động vật}} = 4 - 1 = 3$
- Bậc tự do của sai số ngẫu nhiên $df_{\text{sai số}} = (4 - 2)(4 - 1) = 6$
- Bậc tự do của tổng biến động $df_{\text{tổng số}} = 4^2 - 1 = 15$

Ta có bảng phân tích phương sai:

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	Giá trị F quan sát
Giai đoạn	3	6539	2180	
Bò	3	9929	3310	
Khẩu phần	3	8608	2869	21,12**
Sai số ngẫu nhiên	6	811	135	
Tổng biến động	15	648428		

Từ bảng phân tích phương sai ta thấy các khẩu phần ăn khác nhau đã làm ảnh hưởng rõ rệt đến sản lượng sữa ($P < 0,01$).

Lưu ý: Các ký hiệu thường gặp trong thống kê ở các mức sau:

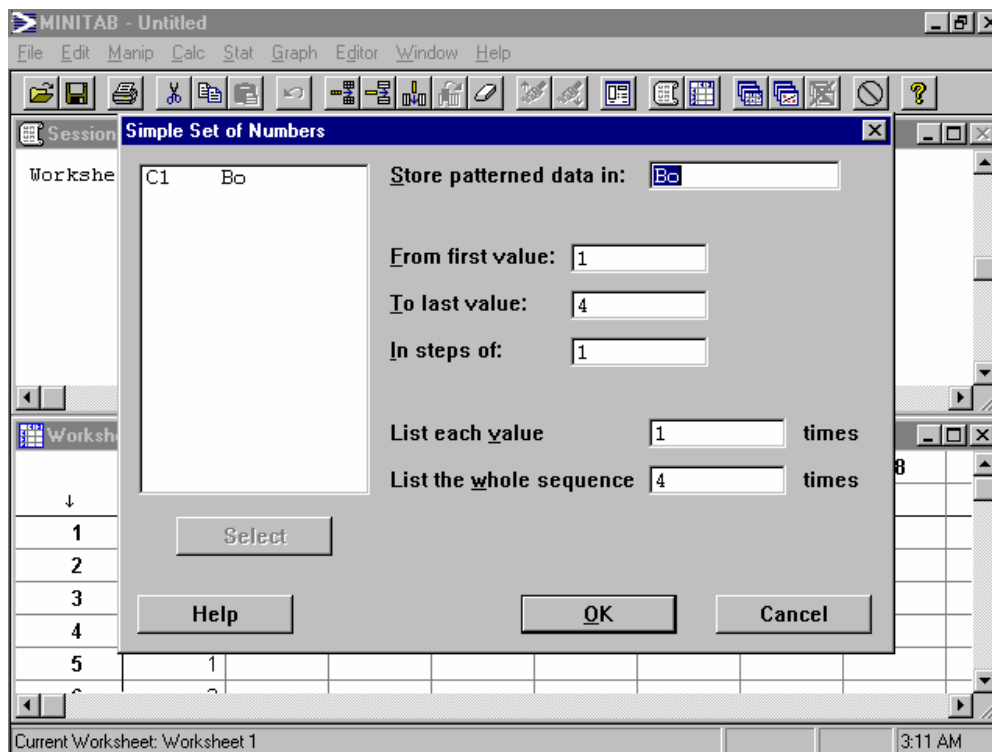
- * mức xác suất $P < 0,05$ (5%)
- ** mức xác suất $P < 0,01$ (1%)
- *** mức xác suất $P < 0,001$ (0,1%)

Để xử lý số liệu bằng các phần mềm thống kê ta phải trình bày số liệu như sau:

Sản lượng	Bò	Giai đoạn	Khẩu phần
192	1	1	A
195	2	1	B
292	3	1	C
249	4	1	D
190	1	2	B
203	2	2	D
218	3	2	A
210	4	2	C
214	1	3	C
139	2	3	A
245	3	3	D
163	4	3	B
221	1	4	D
152	2	4	C
204	3	4	B
134	4	4	A

Nếu sử dụng phần mềm *Minitab* để phân tích số liệu, bạn đọc có thể dùng menu tiện ích sau đây để điền các số ở cột **Bò** và **Giai đoạn** tự động hoá:

Calc → Make Patterned Data → Simple Set of Numbers...



4. Bố trí thí nghiệm 2 nhân tố

Ở phần trên chúng ta chỉ xem xét đến các thí nghiệm khi chỉ có 1 yếu tố thí nghiệm duy nhất tác động lên các đơn vị thí nghiệm. Tuy nhiên trong thực tế chúng ta thường xuyên bắt gặp những vấn đề phức tạp hơn, khi có 2 yếu tố thí nghiệm đồng thời tham gia. Ví dụ, nghiên cứu ảnh hưởng của mức protein và thức ăn tinh trong khẩu phần đến sản lượng sữa của bò cái lai hướng sữa vùng ven đô Hà Nội (Bùi Quang Tuấn).

Thí nghiệm 2 yếu tố được xây dựng trên cơ sở đã tiến các nghiên cứu đối với từng nghiệm thức riêng biệt (thí nghiệm 1 nhân tố).

Đối với những thí nghiệm với 2 yếu tố thí nghiệm ta dùng một số cách bố trí sau:

- Thí nghiệm kiểu hoàn toàn ngẫu nhiên
- Thí nghiệm kiểu khối ngẫu nhiên
- Thí nghiệm kiểu phân cấp.

Tuy nhiên trong khoá học này chúng tôi chỉ đề cập đến mô hình *thí nghiệm kiểu hoàn toàn ngẫu nhiên*.

4.1. Thí nghiệm kiểu hoàn toàn ngẫu nhiên

Trong thí nghiệm này tất cả các động vật được phân về các nghiệm thức hoàn toàn ngẫu nhiên; chính vì vậy có tên gọi *hoàn toàn ngẫu nhiên*.

Ví dụ: Nghiên cứu ảnh hưởng của mức protein (3 mức) và thức ăn tinh (2 mức). Với thí nghiệm này ta có số nghiệm thức (công thức thí nghiệm) là $t = 3 \times 2 = 6$; được minh hoạ như sau:

Mức thức ăn tinh	Mức protein		
	A	B	C
1	A1	B1	C1
2	A2	B2	C2

Cách chọn động vật và phân động vật về các ô thí nghiệm tương tự như đã nêu ở phần 1.3.

4.1.1. Mô hình phân tích

Giả sử, nhân tố A có a nghiệm thức, nhân tố B có b nghiệm thức và trong mỗi nghiệm thức có n quan sát; ta sẽ có mô hình quan sát sau:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

Trong đó $i = 1, 2, \dots, a$

$j = 1, 2, \dots, b$

$k = 1, 2, \dots, n$

y_{ijk} - quan sát thứ k ở nghiệm thứ i của nhân tố A và nghiệm thức thứ j của nhân tố B

μ - giá trị trung bình của toàn bộ các quan sát

α_i - tác động của nghiệm thức thứ i thuộc yếu tố A

β_j - tác động của nghiệm thức thứ k thuộc yếu tố B

ϵ_{ijk} - sai số ngẫu nhiên của quan sát thứ k ở nghiệm thức i thuộc yếu tố A và nghiệm thức thứ j thuộc yếu tố B

$$SS_{\text{toàn bộ}} = SS_{\text{yếu tố A}} + SS_{\text{yếu tố B}} + SS_{\text{tương tác AB}} + SS_{\text{ngẫu nhiên}}$$

Với bậc tự do $N - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1)$

- Trong thí nghiệm kiểu hoàn toàn ngẫu nhiên có 4 nguồn biến động đó là

- Biến động trong từng nhóm (ngẫu nhiên)

$$SS_{\text{ngẫu nhiên}} = SS_{\text{toàn bộ}} - SS_{\text{hàng}} - SS_{\text{cột}} - SS_{\text{nghiệm thức}}$$

- Biến động của nhân tố A

$$SS_{\text{nghiệm thức A}} = bn \sum_{i=1}^a \left(\bar{y}_{i..} - \bar{y}_{...} \right)^2$$

- Biến động của nhân tố B

$$SS_{\text{nghiệm thức B}} = an \sum_{j=1}^b \left(\bar{y}_{.j.} - \bar{y}_{...} \right)^2$$

- Biến động tương tác A&B

$$SS_{\text{tương tác AB}} = n \sum_{i=1}^a \sum_{j=1}^b \left(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...} \right)^2$$

- Toàn bộ các biến động của thí nghiệm:

$$SS_{\text{toàn bộ}} = \sum_i \sum_j \sum_k \left(y_{ijk} - \bar{y}_{ij.} \right)^2$$

Xây dựng cấu trúc của bảng phân tích phương sai

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	Giá trị F quan sát
Nghiệm thức A	$a - 1$	$SS_{\text{nghiệm thức A}}$	$SS_{\text{nghiệm thức A}}/(a-1)$	F_A
Nghiệm thức B	$b - 1$	$SS_{\text{nghiệm thức B}}$	$SS_{\text{nghiệm thức B}}/(b-1)$	F_B
Tương tác AB	$(a - 1)(b - 1)$	$SS_{\text{tương tác AB}}$	$SS_{\text{tương tác}}/(a-1)(b-1)$	F_{AB}
Sai số ngẫu nhiên	$ab(n - 1)$	$SS_{\text{ngẫu nhiên}}$	$SS_{\text{ngẫu nhiên}}/ab(n-1)$	
Tổng biến động	$abn - 1$	$SS_{\text{tổng số}}$		

Giá trị F lý thuyết được xác định ở bảng phân phụ lục với mức xác suất sai số $\alpha = 0,05$; với F_A ta có bậc tự $v_1 = a - 1$ và $v_2 = ab(n-1)$, F_B ta có $v_1 = b - 1$ và $v_2 = ab(n-1)$, F_{AB} ta có $v_1 = (a - 1)(b-1)$ và $v_2 = ab(n-1)$.

- Nếu $P \geq 0,05$ ta chấp nhận H_0 bác bỏ H_1 ,
- Nếu $P < 0,05$ ta bác bỏ H_0 và chấp nhận H_1 .

Ví dụ: Một thí nghiệm được tiến hành nhằm xác định ảnh hưởng của mức bổ sung kẽm và đồng vào khẩu phần ăn đến tăng trọng của gà con. Số liệu được trình bày ở bảng sau:

		Đồng							
		1				2			
Kẽm	1	21,6	22,4	25,7	22,3	23,7	22,0	26,7	25,0
		16,3	18,9	24,8	22,2	19,2	11,2	15,5	19,9
	2	21,2	26,3	28,1	23,0	20,9	22,3	26,7	21,5
		22,6	17,1	24,4	18,8	14,2	12,5	13,9	15,7

Trong ví dụ này ta thấy có 2 yếu tố thí nghiệm đó là *đồng* (ở 2 mức) và *kẽm* (ở 2 mức); trong mỗi khẩu có $n = 8$ con gà hay $N = n \times 4 = 32$ con

Xây dựng cấu trúc của bảng phân tích phương sai

Nguồn biến động	Bậc tự do (df)	Tổng bình phương (SS)	Trung bình bình phương (MS)	Giá trị F quan sát	Giá trị P
Kẽm	1	2,64	2,64	0,14	0,710
Đồng	1	65,55	65,55	3,50	0,072
Đồng*Kẽm	1	14,85	14,85	0,79	0,381
Sai số ngẫu nhiên	28	524,09	18,72		
Tổng biến động	31	607,13			

Qua bảng trên ta thấy giá trị P ở các phép thử đối với kẽm, đồng và tương tác giữa đồng và kẽm đều lớn hơn 0,05; điều này chứng tỏ rằng khi bổ sung thêm đồng và kẽm ở 2 mức nhưng đã không có tác dụng (không làm thay đổi mức tăng trọng ở gà).

4.2. Bài kiểm tra số 5

Một trung tâm nghiên cứu tiến hành thực hiện một thí nghiệm nhằm so sánh tỷ lệ mỡ sữa của bò ở 3 khẩu phần ăn khác nhau (khẩu phần kp1, kp2 và kp3). Trong thí nghiệm này, 30 con bò sữa 3 năm tuổi được chọn ra và phân về các khẩu phần ăn một cách hoàn toàn ngẫu nhiên (mỗi khẩu phần 10 con). Tỷ lệ mỡ sữa trung bình trong một chu kỳ tiết sữa 305 ngày (thời gian thí nghiệm) được ghi lại như sau:	STT	kp1	kp2	kp3
	1	4,25	4,42	3,88
	2	4,24	3,96	3,70
	3	4,09	4,05	3,99
	4	4,32	4,32	4,17
	5	4,52	4,28	4,42
	6	3,96	4,32	3,89
	7	4,40	4,09	3,81
	8	4,06	4,10	3,97
	9	4,48	3,86	3,92
	10	3,96	4,17	3,97

1. (1 điểm) Đây là mô hình thí nghiệm nào?

2. (1 điểm) Hãy cho biết yếu tố thí nghiệm?

3. (1 điểm) Có bao nhiêu nghiệm thức (công thức thí nghiệm), đó là những nghiệm thức nào?

4. (1 điểm) Có bao nhiêu nguồn biến động trong ví dụ nêu trên, đó là những nguồn biến động nào?

5. (6 điểm) Nếu là kỹ sư chăn nuôi, anh (chị) sẽ chọn khẩu phần nào để đưa vào áp dụng trong chăn nuôi bò sữa, vì sao? Biết rằng sản lượng sữa của bò ở 3 khẩu phần (kp1, kp2 và kp3) là như nhau.

4.3. Bài kiểm tra số 6

Một trung tâm nghiên cứu tiến hành thực hiện một thí nghiệm nhằm so sánh mức tăng trọng của lợn ở 5 khẩu phần ăn khác nhau (khẩu phần kp1, kp2, kp3, kp4 và kp5). Trong thí nghiệm này, 60 con lợn 21 ngày tuổi được chọn ra và phân về các khẩu phần ăn một cách hoàn toàn ngẫu nhiên (mỗi khẩu phần 12 con).

1. (1 điểm) Hãy cho biết yếu tố thí nghiệm?
2. (1 điểm) Có bao nhiêu nghiệm thức (công thức thí nghiệm), đó là những nghiệm thức nào?
3. (4 điểm) Vẽ sơ đồ thí nghiệm (bố trí động vật về các khẩu phần)
4. (4 điểm) Điền những cụm từ hoặc các giá trị thích hợp vào các ô có dấu (*) trong bảng phân tích phương sai sau đây

Nguồn biến động	Bậc tự do (<i>df</i>)	<i>SS</i>	<i>MS</i>	<i>F</i> <i>qs</i>
*	*	-	-	-
*	*	-	-	-
*	*	-	-	-

5. Hồi quy tuyến tính đơn giản

5.1. Giới thiệu

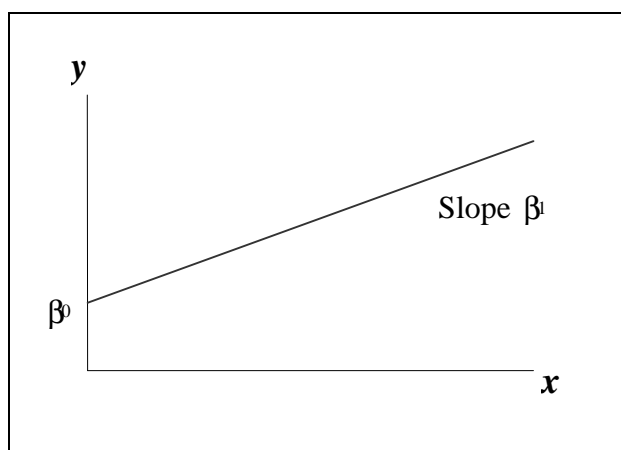
Giả sử ta có một mối quan hệ tuyến tính hoàn hảo giữa 2 biến x và y (Ví dụ x = nhiệt độ, y = áp suất của khí). Ta có thể biểu diễn mối quan hệ đó như sau:

$$y = \beta_0 + \beta_1 x$$

ở đây

β_0 = y -hệ số hồi quy (ở đó đường thẳng đi qua trục y); và

β_1 = độ dốc (slope) của đường hồi quy



Các tham số β_0 và β_1 có thể tính toán được từ 2 điểm bất kỳ, giả sử (x_1, y_1) và (x_2, y_2) :

$$\text{ta có } \beta_1 = \frac{y_2 - y_1}{x_2 - x_1} \text{ và } \beta_0 = y_1 - \beta_1 x_1 = y_2 - \beta_1 x_2$$

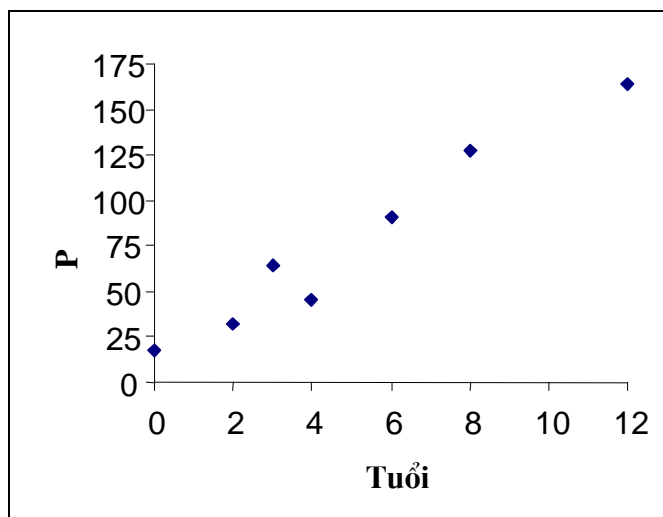
Vấn đề đặt ra: Trong sinh học cũng như trong các ngành khoa học về môi trường các số liệu thu được không tuân theo một liên hệ tuyến tính hoàn hảo như trên mà chỉ có các điểm phân bố xung quanh đường thẳng.

5.2. Ví dụ

Khối lượng (kg) và tuổi (tháng) của 7 con bê khác nhau được theo dõi như sau:

Tuổi (x)	Khối lượng (y)
0	18
2	32
3	64
4	45
6	91
8	127
12	164

Đây là mối quan hệ tuyến tính thuận, khi độ tuổi tăng thì khối lượng cũng tăng theo:



5.3. Mô hình tuyến tính

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

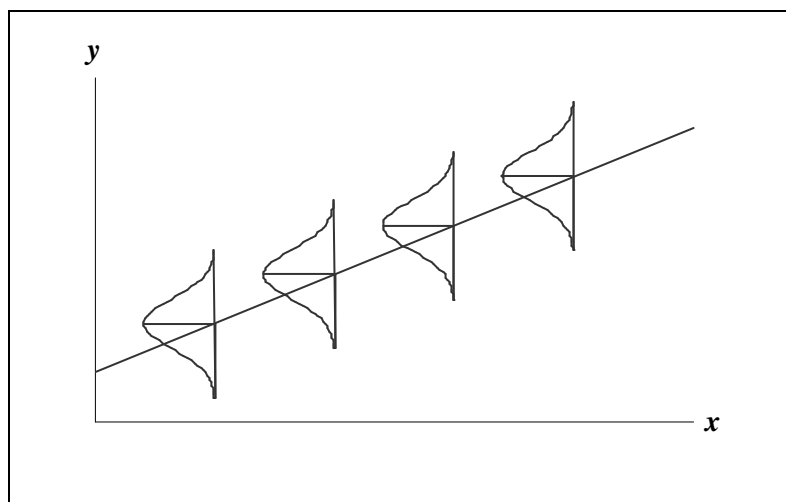
$$(\text{Khối lượng}) = \beta_0 + \beta_1 \times (\text{tháng tuổi}) + \varepsilon_i.$$

Số liệu từ các cặp tương ứng $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

- Giả thuyết:

- (1) số liệu có phân bố chuẩn
- (2) phương sai là một hằng số (σ^2)
- (3) mô hình tuyến tính phải tuân theo:

$$y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2) \text{ hoặc tương đương } \varepsilon_i \sim N(0, \sigma^2).$$



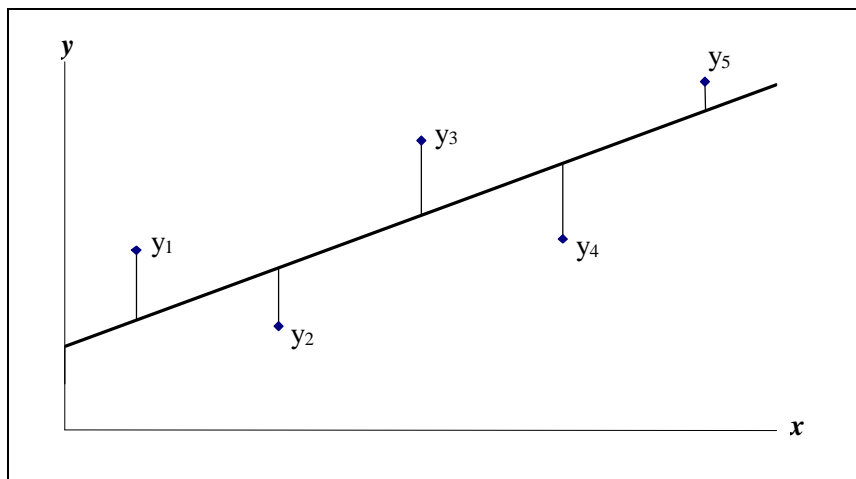
Như vậy đối với từng giá trị của x , y sẽ có phân bố chuẩn với trung bình là $\beta_0 + \beta_1 x_i$ và phương sai là σ^2 .

5.4. Ước lượng các tham số β_0 và β_1

Các tham số β_0 và β_1 không thể xác định được chính xác từ các số liệu thu thập được, nhưng ta có thể ước lượng được chúng thông qua các số liệu này. Đặt ước tính của β_0 bằng b_0 và β_1 bằng b_1

Mô hình được chuyển thành $\hat{y}_i = b_0 + b_1 x_i$ và \hat{y}_i là giá trị ước lượng lý tưởng của y_i dựa trên giá trị thực của y_i . Sai số ngẫu nhiên là $= y_i - \hat{y}_i = y_i - b_0 - b_1 x_i$ và sự sai khác của các quan sát sẽ nằm trên hoặc nằm dưới đường hồi quy.

Một trong những cách thông thường để ước tính giá trị tốt nhất cho β_0 và β_1 là chọn giá trị b_0 và b_1 sao cho phần dư bình phương là bé nhất so với đường hồi quy tuyến tính. Phương pháp này được gọi là phương pháp **ước tính bình phương bé nhất**.



Tổng bình phương sai số so với đường hồi quy =

$$= \sum_{i=1}^n \text{res}_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

Chúng ta có thể nhận thấy rằng lựa chọn giá trị b_0 và b_1 sao cho phần dư bình phương là tối thiểu khi:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad b_0 = \bar{y} - b_1 \bar{x}$$

Các bước tính toán cụ thể được trình bày ở bảng dưới đây:

Tuổi (x_i)	Khối lượng (y_i)	$X_i = x_i - \bar{x}$	$Y_i = y_i - \bar{y}$	X_i^2	Y_i^2	$X_i Y_i$
0	18	-5	-59,29	25	3514,80	296,43
2	32	-3	-45,29	9	2050,80	135,86
3	64	-2	-13,29	4	176,51	26,57
4	45	-1	-32,29	1	1042,37	32,29
6	91	1	13,71	1	188,08	13,71
8	127	3	49,71	9	2471,51	149,14
12	164	7	86,71	49	7519,37	607,00
35	541	0	0,00	98	16963,43	1261,00

$$\bar{x} = 5,00 \quad \bar{y} = 77,286$$

$$n = 7$$

Do đó, $b_1 = 1261 / 98 = 12,867$ $b_0 = 77,286 - 12,867 \times 5,00 = 12,949$

Do đó đường hồi quy tuyến tính bình phương bé nhất là $\hat{y} = 12,95 + 12,87x$

Giá trị hồi quy b_0 : Khi $x = 0$ (lúc sinh ra), trọng lượng trung bình là 12,95 kg.

Độ nhọn b_1 : Cứ thêm mỗi tháng tuổi thì trọng lượng của bê tăng 12,87 kg.

Phần dư và giá trị theo đường hồi quy

Tuổi (x_i)	Khối lượng (y_i)	Giá trị hồi quy $\hat{y}_i = 12,95 + 12,87x_i$	Phần dư $\text{res}_i = y_i - \hat{y}_i$	res_i^2
0	18	12,95	5,05	25,51
2	32	38,68	-6,68	44,67
3	64	51,55	12,45	154,98
4	45	64,42	-19,42	377,07
6	91	90,15	0,85	0,72
8	127	115,89	11,11	123,48
12	164	167,36	-3,36	11,27
35	541	541,00	0,00	737,70

Lưu ý rằng giá trị trung bình của phần dư bằng không

$$[\text{Phần dư so với đường hồi quy bình phương}] = \sum_{i=1}^7 \text{res}_i^2 = 737,70$$

Nếu lựa chọn b_0 và b_1 là các giá trị khác sẽ làm tăng phần dư này.

Giá trị theo đường hồi quy (\hat{y}_i) được dùng để ước tính khối lượng trung bình của bê đối với một ngày tuổi cho trước.

Có thể ước lượng khối lượng trung bình của bê 10 tháng tuổi như sau

$$12,95 + 12,87 \times 10 = 141,62 \text{ kg.}$$

Ước lượng σ^2

Ta quay trở lại với giả thiết đối với mô hình hồi quy:

$$y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2) \quad \text{hoặc tương đương với} \quad \varepsilon_i \sim N(0, \sigma^2).$$

trong đó σ^2 là phương sai của phương trình hồi quy.

Nó được ước tính như sau s^2 ($s_{Y.X}^2$)

$$s^2 = \text{Residual SS} / (n - 2)$$

$$= \sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - 2)$$

Đối với số liệu về khối lượng của bê: $s^2 = 79,70 / 5 = 147,54$, and $\therefore s = 12,15$ kg.

5.5. Kiểm định giả thuyết

5.5.1. Hệ số hồi quy

Giả thuyết $H_0 : \beta_0 = 0$ vs $H_0 : \beta_0 \neq 0$ (đường hồi quy đi qua gốc tọa độ)

$$\text{Kiểm định thống kê: } t = \frac{b_0}{\text{se}(b_0)} \quad df = n - 2 \quad \text{trong đó } \text{se}(b_0) = s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}}$$

Đối với ví dụ về khối lượng của bê:

$$\text{se}(b_0) = 12,15 \times \sqrt{\frac{273}{7 \times 98}} = 7,66$$

$$t = 12,95 / 7,66 = 1,69, \text{ với bậc tự do } df = 7 - 2 = 5.$$

$$P\text{-value: } P = 2 \times P(T_5 > 1,69) = 0,15$$

Như vậy giả thuyết H_0 được chấp nhận: kết luận rằng đường hồi quy đi qua gốc tọa độ.

Chú ý: khi $\beta_0 = 0$, thì mô hình được rút gọn như sau $y_i = \beta_1 x_i + \varepsilon_i$ có nghĩa là y ‘tỷ lệ’ với x .

5.5.2. Độ dốc

Giả thuyết $H_0 : \beta_1 = 0$ với $H_1 : \beta_1 \neq 0$ (độ dốc bằng không: không có quan hệ tuyến tính)

$$\text{Kiểm định thống kê: } t = \frac{b_1}{\text{se}(b_1)} \quad df = n - 2 \quad \text{trong đó } \text{se}(b_1) = \frac{s}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

Đối với số liệu về khối lượng của bê:

$$\text{se}(b_1) = \frac{12,15}{\sqrt{98}} = 1,23$$

$$t = 12,87 / 1,23 = 10,49, \text{ với bậc tự do } df = 7 - 2 = 5.$$

$$P\text{-value: } P = 2 \times P(T_5 > 10,49) = 0,00$$

Như vậy giả thuyết H_0 bị bác bỏ: Kết luận rằng khối lượng của bê tăng một cách có ý nghĩa với độ tuổi.

5.5.3. Bảng phân tích phương sai (ANOVA) đối với hồi quy

Cũng như trong phân tích phương sai (ANOVA), chúng ta cũng có thể chia sự biến động của số liệu (y) thành các thành phần được giải thích trong mô hình và thành phần không giải thích được:

Tổng bình phương (SS):

$$\begin{array}{l} \text{Tổng SS} = \text{SS hồi quy} + \text{SS phần dư} \\ \text{bậc tự do: } (n - 1) = 1 + (n - 2) \end{array}$$

Giá trị SS trong ví dụ được tính toán như sau:

$$\begin{aligned}
\text{Tổng SS} &= \sum_{i=1}^n (y_i - \bar{y})^2 = \Sigma(\text{Quan sát} - \text{Trung bình})^2 \\
&= (18 - 77.29)^2 + (32 - 77.29)^2 + \dots + (164 - 77.29)^2 \\
&= 16,963 \\
\text{SS hồi quy} &= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \Sigma(\text{Hồi quy} - \text{Trung bình})^2 \\
&= (12.95 - 77.29)^2 + (38.68 - 77.29)^2 + \dots + (167.36 - 77.29)^2 \\
&= 16,226 \\
\text{SS phần dư} &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \Sigma(\text{Quan sát} - \text{Hồi quy})^2 = \Sigma(\text{Phần dư})^2 \\
&= (18 - 12.95)^2 + (32 - 38.68)^2 + \dots + (164 - 167.36)^2 \\
&= 738
\end{aligned}$$

Chú ý rằng SS của hồi quy có thể xác định bằng sử dụng phương trình sau đây,

$$\text{SS hồi quy} = b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2 = 12.87^2 \times 98 = 16,226$$

cũng như trên, apart from some round off error.

Kết quả phân tích được trình bày ở bảng ANOVA

Nguồn biến động	Tổng bình phương (SS)	Bậc tự do (df)	TB bình phương (MS)
Hồi quy	Reg SS	1	Reg SS
Phần dư	Res SS	$n - 2$	Res SS / ($n - 2$)
Tổng	Tot SS	$n - 1$	

Với số liệu về bê, bảng ANOVA là

Nguồn biến động	Tổng bình phương (SS)	Bậc tự do (df)	TB bình phương (MS)
Hồi quy	16,226	1	16,226
Phần dư	738	5	147.5
Tổng	16,963	6	

Chúng ta cũng có thể xác định ý nghĩa của độ dốc với phương pháp thử F . Kiểm định thống kê:

$$F = \frac{\text{Regression MS}}{\text{Residual MS}} \quad df = 1, n - 2$$

Trong ví dụ về bê: $F = 16,226 / 147.5 = 110.0$ với $df = 1, 5$

So sánh với phân bố $F_{1,5}$, ta có $P = 0.00$.

Như vậy ta có giá trị P tương tự như phần kiểm định t như trên.

Đối với **hồi quy tuyến tính đơn giản**, ta có mối quan hệ chặt chẽ giữa t -test và F -test:

$$t^2 = F \quad (10.49^2 = 110.0) \quad \text{Chú ý bậc tự do bằng nhau (bằng 5)}$$

R^2 - Phần biến động được giải thích bằng mô hình. Với số liệu về bê, $R^2 = 16,226 / 16,963 = 0.957$, hay 96% biến động được giải thích bằng độ tuổi của bê.

Minitab example: Khối lượng (y) và tuổi (x) của 7 bê

```
MTB > NAME C1 'Tuoi' C2 'Khoi luong'
```

```
MTB > REGR C2 1 C1
```

Stat > Regression > Regression...

Regression Analysis

The regression equation is

Khoi luong = 12.9 + 12.9 Tuoi

Predictor	Coef	StDev	T	P
Constant	12.949	7.663	1.69	0.152
Tuoi	12.867	1.227	10.49	0.000

S = 12.15 R-Sq = 95.7% R-Sq(adj) = 94.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	16226	16226	109.97	0.000
Residual Error	5	738	148		
Total	6	16963			

Lưu ý:

Kiểm định, nếu trọng lượng có tương quan với độ tuổi, Chúng ta kiểm tra giả thuyết $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$. Khi chỉ có một biến ước tính x , Thì chúng ta có thể dùng t -test hoặc F -test để thực hiện phép thử. Cả 2 phương pháp đều cho ta giá trị P như nhau (bằng 0.000), và chú ý rằng

$$t^2 = (10.49)^2 = 109.97 = F.$$

Biến động của tăng trọng được tính toán theo lứa tuổi là

$$SS \text{ hồi quy} / \text{Tổng SS} = 16226 / 16963 = 0.957$$

Giá trị R^2 cho ta thấy giá trị hồi quy (R-sq=95.7%)

6. Tương quan

6.1. Giới thiệu

Chúng ta có thể sử dụng hệ số tương quan để xác định mức độ quan hệ tuyến tính giữa 2 biến. Hệ số tương quan có giá trị từ -1 đến +1. Nếu một biến có xu hướng tăng còn biến kia giảm thì hệ số tương quan là âm. Còn nếu cả hai biến có xu hướng cùng tăng thì hệ số tương quan là dương. Hệ số tương quan của quần thể được ký hiệu bằng ρ và r với mẫu. Mức độ tương quan có thể được kiểm định bằng phép thử từ 2 phía:

$H_0: \rho = 0$ versus $H_1: \rho \neq 0$ trong đó ρ là tương quan giữa 2 biến.

6.2. Tính hệ số tương quan

Đối với 2 biến x và y ,

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y}$$

trong đó \bar{x} và s_x là giá trị trung bình và độ lệch chuẩn của mẫu thứ nhất, \bar{y} và s_y là giá trị trung bình và độ lệch chuẩn của mẫu thứ 2. Chú ý rằng:

$r = 0 \Rightarrow$ không có mối quan hệ tuyến tính;

$r = +1 \Rightarrow$ quan hệ tuyến tính dương lý tưởng; và

$r = -1 \Rightarrow$ quan hệ tuyến tính âm lý tưởng;

Chúng ta có thể sử dụng ví dụ về tăng trọng của bê ở ví dụ hồi quy tuyến tính đơn giản để tính toán. Các số liệu về độ dốc (b_1) đã được tính toán trong phần hồi quy tuyến tính đơn giản (xem bảng tính ở phần này).

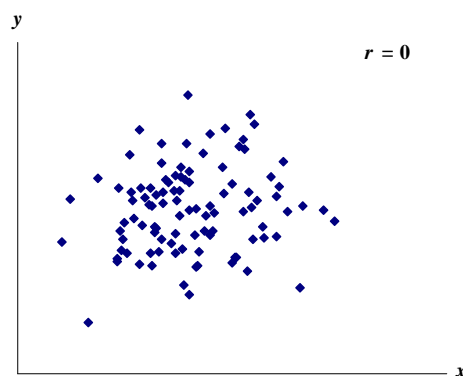
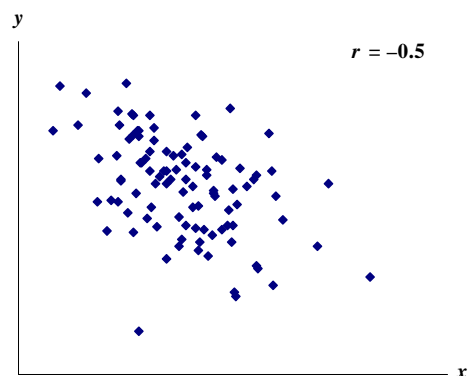
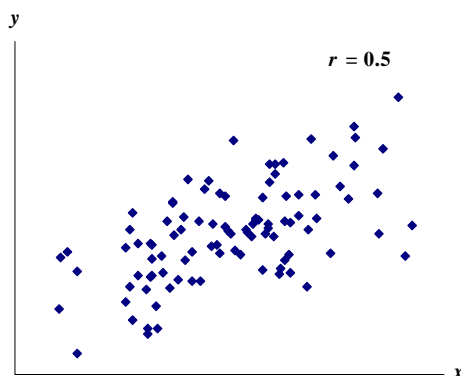
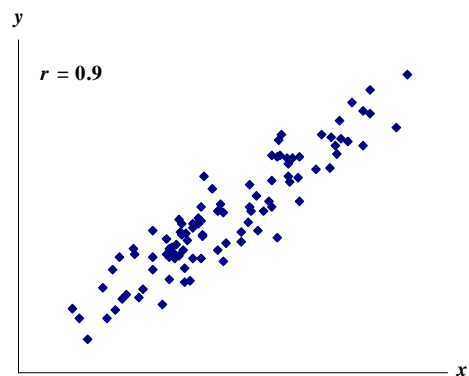
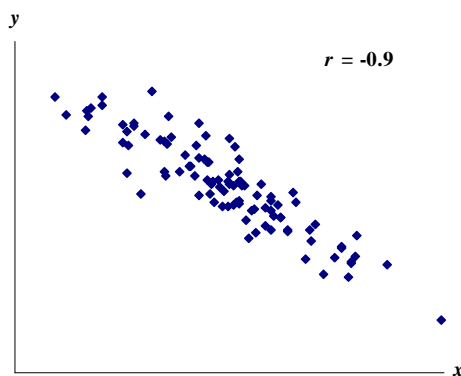
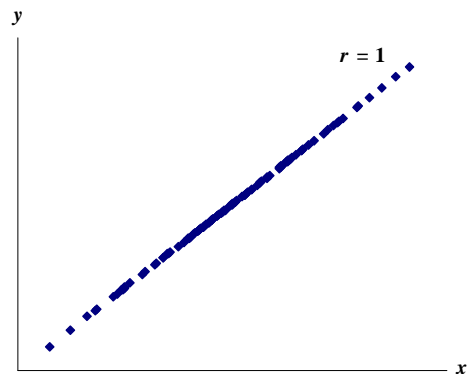
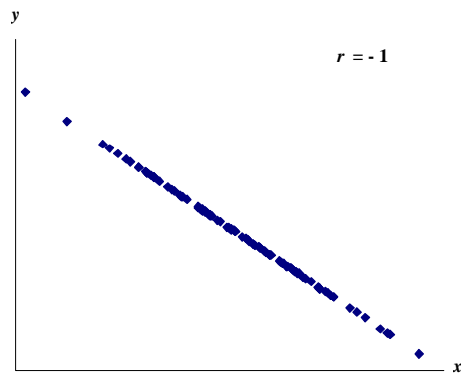
Tuổi (x_i)	Khối lượng (y_i)	$X_i = x_i - \bar{x}$	$Y_i = y_i - \bar{y}$	$X_i Y_i$
0	18	-5	-59,29	296,43
2	32	-3	-45,29	135,86
3	64	-2	-13,29	26,57
4	45	-1	-32,29	32,29
6	91	1	13,71	13,71
8	127	3	49,71	149,14
12	164	7	86,71	607,00
35	541	0	0,00	1.261,00
$\bar{x} = 5,00$	$\bar{y} = 77,286$	$s_x = 4,04$	$s_y = 53,2$	$n = 7$

Ta có:

$$r = 1261 / (6)(4,04)(53,2) = 0,978$$

Đây là sự tương quan rất chặt chẽ (giá trị tối đa là 1).

6.3. Những ví dụ về sự tương quan



6.4. Tính toán trong Minitab

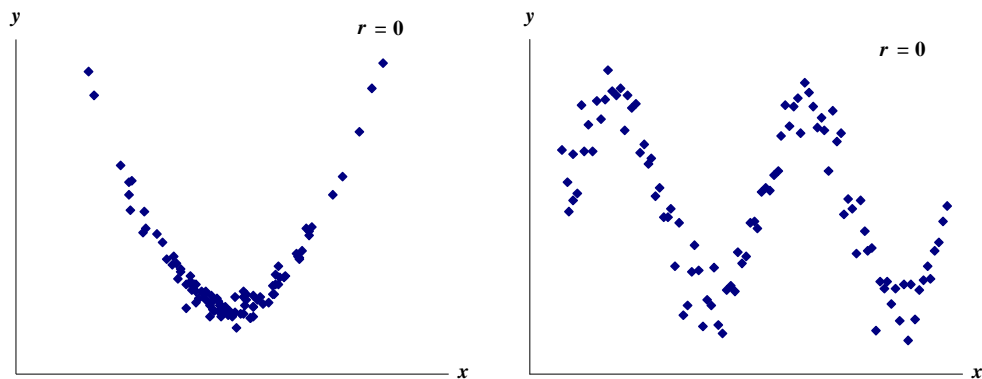
Chọn cột chứa biến mà bạn cần tính giá trị tương quan. Khi các bạn chọn 2 cột đồng thời thì *Minitab* sẽ tính hệ số tương quan cho 1 cặp. Khi bạn chọn nhiều hơn 2 cột, *Minitab* sẽ tính hệ số tương quan cho từng cặp có thể và cho ta một ma trận hệ số tương quan (trong một khối nếu có đủ chỗ trống để hiển thị trong 1 trang). Kiểm tra giá trị *P* đối với giả thuyết hệ số tương quan bằng 0 theo mặc định. Sử dụng ví dụ về tăng trọng của bê ta có:

```
MTB > Correlation 'Tuoi' 'Khoi luong'.      Stat>Basic Statistics>Correlations

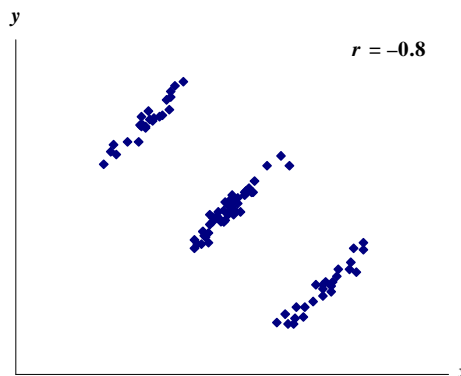
Pearson correlation of Tuoi and Khoi luong = 0.978
P-Value = 0.000
```

6.5. Mối quan hệ và hệ số tương quan

Chú ý rằng nếu $r \neq 0$, sẽ tồn tại một mối quan hệ giữa x và y (mặc dù không cần thiết phải có ý nghĩa). Tuy nhiên hai biến có thể không có tương quan tuyến tính ($r = 0$) nhưng chúng lại có mối quan hệ chặt chẽ không tuyến tính nào đó. Một số đám mây sinh học sau đây sẽ minh họa điều này:



Như vậy hệ số tương quan phải được sử dụng một cách thận trọng để xác định mối quan hệ: **luôn luôn kiểm tra mối quan hệ không tuyến tính đối với số liệu!** Dưới đây là một ví dụ minh chứng điển hình, khi mà 3 đường riêng biệt; mỗi đường đều có mối quan hệ dương giữa x và y nhưng về tổng thể là tương quan âm!



6.6. Mối quan hệ giữa tương quan và hồi quy

Như ta đã biết từ phân tích hồi quy ta có thể tính được R^2 bằng phần biến động của y từ biến x . Đây chính là bình phương của hệ số tương quan, r^2 , được trình bày dưới đây:

R^2 = Biến động đối với phương trình hồi quy / Toàn bộ các biến động

$$\begin{aligned} &= \frac{b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \\ &= \frac{\left[\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right]^2}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} \times \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \\ &= \frac{\left[\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right]^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \times \sum_{i=1}^n (y_i - \bar{y})^2} \\ &= r^2. \end{aligned}$$

Áp dụng công thức này với số liệu về tăng trọng của bê ta có:

Biến động đối với phương trình hồi quy / Toàn bộ các biến động = $16,226 / 16,963 = 0.957$

tương đương với $r^2 = 0.978^2 = 0.957$.

6.7. Bài kiểm tra số 7

Để xác định hệ số tương quan giữa các tính trạng năng suất sinh sản của 2 giống lợn nái ngoại (Landrace và Yoorkshire) nuôi tại các cơ sở giống ở miền Bắc Việt Nam; tiến hành rút một cách ngẫu nhiên thành tích của 25 nái từ số giống thu được số liệu như sau:

STT	scdr	sc21	poss	pssc	po21	p21con
1	13	10	13,5	1,23	32,0	3,20
2	8	8	11,2	1,40	41,6	5,20
3	13	12	18,2	1,40	72,0	6,00
4	8	6	9,0	1,13	20,0	3,33
5	14	8	14,0	1,40	42,0	5,25
6	8	8	11,5	1,44	36,0	4,50
7	10	10	14,5	1,45	60,0	6,00
8	9	8	10,8	1,20	40,0	5,00
9	12	12	19,0	1,58	65,0	5,42
10	10	10	10,5	1,05	38,2	3,82
11	11	10	12,0	1,09	62,0	6,20
12	10	8	14,0	1,40	47,0	5,88
13	10	10	16,0	1,60	56,0	5,60
14	10	10	14,0	1,40	51,0	5,10
15	12	10	15,5	1,29	42,0	4,20
16	11	10	12,9	1,29	63,8	6,38
17	10	10	13,0	1,30	46,0	4,60
18	12	9	14,0	1,27	43,0	4,78
19	12	10	14,6	1,22	52,0	5,20
20	9	8	12,5	1,39	34,0	4,25
21	10	9	17,6	1,76	50,0	5,56
22	10	9	13,7	1,37	40,3	4,48
23	11	10	14,0	1,27	51,0	5,10
24	10	8	13,0	1,30	46,3	5,79
25	8	8	9,0	1,13	36,6	4,58

Ghi chú:

scdr - Số con đẻ ra/ lứa

pssc - Khối lượng sơ sinh/ con

sc21 - Số con còn sống đến 21 ngày tuổi

po21 - Khối lượng 21 ngày tuổi/ lứa

poss - Khối lượng sơ sinh/ lứa

p21con - Khối lượng 21 ngày tuổi/ con

1) (2 điểm) Hãy điền các giá trị thích hợp vào bảng sau:

Chỉ tiêu	Đơn vị tính	n	$\bar{X} \pm S_{\bar{X}}$	Cv (%)
Số con đẻ ra/ lứa				
Số con còn sống đến 21 ngày tuổi				
Khối lượng sơ sinh/ lứa				
Khối lượng sơ sinh/ con				
Khối lượng 21 ngày tuổi/ lứa				
Khối lượng 21 ngày tuổi/ con				

2) (6 điểm) Điền các hệ số tương quan thích hợp vào bảng sau:

Chỉ tiêu	(1)	(2)	(3)	(4)	(5)	(6)
Số con đẻ ra/ lứa (1)						
Số con sống đến 21 ngày tuổi (2)						
Khối lượng sơ sinh/ lứa (3)						
Khối lượng sơ sinh/ con (4)						
Khối lượng 21 ngày tuổi/ lứa (5)						
Khối lượng 21 ngày tuổi/ con (6)						

3) (2 điểm) Anh (chị) thử tiến hành bình luận một hệ số tương quan trong bảng trên.

7. Phụ lục

BẢNG SỐ NGẪU NHIÊN

81	37	66	40	77	65	29	99	77	42	92	78	15	25	07	76	79	24	21	84
48	03	48	91	03	57	56	56	42	76	57	27	60	60	16	30	76	96	94	49
86	49	52	63	66	70	80	71	09	64	84	36	03	54	53	39	36	30	69	27
73	59	16	61	43	18	86	80	19	42	23	78	86	08	44	08	55	51	12	97
10	46	82	01	40	55	50	91	24	12	34	43	20	37	71	52	13	25	67	31
63	34	98	49	54	23	60	36	10	40	08	12	34	46	59	82	91	74	60	92
18	40	40	07	42	21	10	22	39	57	86	80	03	29	64	96	73	84	72	47
59	86	66	45	91	17	29	15	92	05	97	60	76	48	44	58	89	64	01	26
30	99	69	70	16	08	76	29	74	90	18	42	43	71	47	22	10	21	08	69
14	49	02	64	25	44	27	12	36	82	67	84	58	21	61	72	45	23	63	43
99	76	35	87	72	35	14	61	70	33	94	30	18	23	70	30	80	72	72	04
50	42	77	64	94	44	17	80	67	98	72	15	00	52	41	76	16	85	33	23
10	38	18	55	57	31	38	12	97	80	91	47	94	45	67	92	31	55	16	91
46	52	61	13	33	04	30	47	97	11	30	03	87	98	33	06	29	77	56	41
29	21	02	78	61	84	33	50	43	75	42	28	40	16	12	42	03	44	10	28
83	59	26	14	81	77	04	94	98	12	33	71	07	29	35	25	86	82	52	43
87	22	31	54	76	04	80	79	92	37	97	31	53	34	10	57	19	48	32	86
73	53	23	83	40	45	57	33	18	29	13	61	64	03	38	09	01	88	13	14
29	32	83	46	27	05	18	31	46	93	59	83	90	79	53	91	47	02	26	90
70	71	37	04	12	71	30	23	31	51	92	96	09	93	08	52	94	79	45	34
87	29	28	54	53	54	33	39	22	61	46	98	84	24	28	71	42	75	98	07
83	78	88	92	75	35	07	41	70	05	83	13	45	06	24	89	75	66	06	27
69	26	97	35	72	95	58	30	84	12	70	41	36	92	05	62	89	01	62	31
07	82	88	94	99	80	07	37	94	52	15	26	90	39	39	51	53	40	98	78
55	80	29	81	32	27	28	59	29	74	27	46	15	47	00	47	94	04	03	43
80	73	03	69	35	68	22	77	82	26	83	58	62	71	77	88	00	70	45	58
45	69	97	79	98	33	45	64	83	62	20	36	34	64	67	29	08	47	56	72
25	15	57	13	07	95	01	02	02	70	86	74	56	14	94	33	49	73	62	71
82	87	56	32	99	86	35	13	22	12	25	90	89	20	82	87	46	23	14	27
00	98	13	94	00	85	09	30	97	98	72	40	81	87	33	96	58	28	08	64
61	99	16	38	11	08	28	65	70	71	79	51	31	38	27	99	64	57	99	98
79	93	50	34	41	50	21	49	74	52	03	52	53	24	89	53	96	19	31	06
36	19	99	62	65	08	46	68	44	96	73	98	65	41	72	37	46	27	11	41
88	27	35	22	39	59	19	39	65	55	59	20	25	48	23	61	78	35	48	89
24	20	27	94	31	17	47	50	37	11	15	19	46	34	23	80	37	60	30	50
54	55	44	08	73	05	63	52	47	43	82	40	98	97	92	13	46	31	02	67
83	93	99	35	06	85	63	39	04	12	93	91	86	88	63	68	62	75	91	38
64	64	87	77	53	05	29	76	06	23	88	81	10	33	02	86	86	93	12	00
74	72	31	23	20	17	06	56	26	91	86	60	48	28	08	93	56	03	26	44
81	76	68	15	22	70	38	56	71	59	69	38	45	64	79	98	69	02	11	90

BẢNG XÁC SUẤT CỦA PHÂN BỐ TIÊU CHUẨN HOÁ

Các giá trị trong bảng là của phân bố chuẩn với trung bình bằng 0 và độ lệch chuẩn là 1. Ứng với mỗi giá trị z trong là giá trị P , $P(Z < z)$.

z	P	z	P	z	P	z	P
-4,00	0,00003	-1,50	0,0668	0,00	0,5000	1,55	0,9394
-3,50	0,00023	-1,45	0,0735	0,05	0,5199	1,60	0,9452
-3,00	0,0013	-1,40	0,0808	0,10	0,5398	1,65	0,9505
-2,95	0,0016	-1,35	0,0885	0,15	0,5596	1,70	0,9554
-2,90	0,0019	-1,30	0,0968	0,20	0,5793	1,75	0,9599
-2,85	0,0022	-1,25	0,1056	0,25	0,5987	1,80	0,9641
-2,80	0,0026	-1,20	0,1151	0,30	0,6179	1,85	0,9678
-2,75	0,0030	-1,15	0,1251	0,35	0,6368	1,90	0,9713
-2,70	0,0035	-1,10	0,1357	0,40	0,6554	1,95	0,9744
-2,65	0,0040	-1,05	0,1469	0,45	0,6736	2,00	0,9772
-2,60	0,0047	-1,00	0,1587	0,50	0,6915	2,05	0,9798
-2,55	0,0054	-0,95	0,1711	0,55	0,7088	2,10	0,9821
-2,50	0,0062	-0,90	0,1841	0,60	0,7257	2,15	0,9842
-2,45	0,0071	-0,85	0,1977	0,65	0,7422	2,20	0,9861
-2,40	0,0082	-0,80	0,2119	0,70	0,7580	2,25	0,9878
-2,35	0,0094	-0,75	0,2266	0,75	0,7734	2,30	0,9893
-2,30	0,0107	-0,70	0,2420	0,80	0,7881	2,35	0,9906
-2,25	0,0122	-0,65	0,2578	0,85	0,8023	2,40	0,9918
-2,20	0,0139	-0,60	0,2743	0,90	0,8159	2,45	0,9929
-2,15	0,0158	-0,55	0,2912	0,95	0,8289	2,50	0,9938
-2,10	0,0179	-0,50	0,3085	1,00	0,8413	2,55	0,9946
-2,05	0,0202	-0,45	0,3264	1,05	0,8531	2,60	0,9953
-2,00	0,0228	-0,40	0,3446	1,10	0,8643	2,65	0,9960
-1,95	0,0256	-0,35	0,3632	1,15	0,8749	2,70	0,9965
-1,90	0,0287	-0,30	0,3821	1,20	0,8849	2,75	0,9970
-1,85	0,0322	-0,25	0,4013	1,25	0,8944	2,80	0,9974
-1,80	0,0359	-0,20	0,4207	1,30	0,9032	2,85	0,9978
-1,75	0,0401	-0,15	0,4404	1,35	0,9115	2,90	0,9981
-1,70	0,0446	-0,10	0,4602	1,40	0,9192	2,95	0,9984
-1,65	0,0495	-0,05	0,4801	1,45	0,9265	3,00	0,9987
-1,60	0,0548	0,00	0,5000	1,50	0,9332	3,50	0,99977
-1,55	0,0606					4,00	0,99997

Một vài giá trị tới hạn của z :

P	0,80	0,90	0,95	0,975	0,99	0,995	0,999
z	0,842	1,282	1,645	1,960	2,326	2,576	3,090

BẢNG XÁC SUẤT CỦA PHÂN BỐ STUDENT (T)

Các giá trị trong bảng là của phân bố t . Cột thứ nhất là bậc tự do (df). Các cột còn lại cho ta các giá trị lý thuyết về kiểm định một hướng (phần trên); $P(T_{df} > t) = P$, hoặc 2 hướng; $P(T_{df} > t \text{ hoặc } T_{df} < -t) = P$ trong đó P là mức xác suất được thể hiện ở đầu cột.

df	P					
	0,10	0,05	0,025	0,01	0,005	0,001 (1 hướng)
	0,20	0,10	0,05	0,02	0,01	0,002 (2 hướng)
1	3,078	6,314	12,706	31,821	63,657	318,313
2	1,886	2,920	4,303	6,965	9,925	22,327
3	1,638	2,353	3,182	4,541	5,841	10,215
4	1,533	2,132	2,776	3,747	4,604	7,173
5	1,476	2,015	2,571	3,365	4,032	5,893
6	1,440	1,943	2,447	3,143	3,707	5,208
7	1,415	1,895	2,365	2,998	3,499	4,785
8	1,397	1,860	2,306	2,896	3,355	4,501
9	1,383	1,833	2,262	2,821	3,250	4,297
10	1,372	1,812	2,228	2,764	3,169	4,144
11	1,363	1,796	2,201	2,718	3,106	4,025
12	1,356	1,782	2,179	2,681	3,055	3,930
13	1,350	1,771	2,160	2,650	3,012	3,852
14	1,345	1,761	2,145	2,624	2,977	3,787
15	1,341	1,753	2,131	2,602	2,947	3,733
16	1,337	1,746	2,120	2,583	2,921	3,686
17	1,333	1,740	2,110	2,567	2,898	3,646
18	1,330	1,734	2,101	2,552	2,878	3,611
19	1,328	1,729	2,093	2,539	2,861	3,579
20	1,325	1,725	2,086	2,528	2,845	3,552
21	1,323	1,721	2,080	2,518	2,831	3,527
22	1,321	1,717	2,074	2,508	2,819	3,505
23	1,319	1,714	2,069	2,500	2,807	3,485
24	1,318	1,711	2,064	2,492	2,797	3,467
25	1,316	1,708	2,060	2,485	2,787	3,450
26	1,315	1,706	2,056	2,479	2,779	3,435
27	1,314	1,703	2,052	2,473	2,771	3,421
28	1,313	1,701	2,048	2,467	2,763	3,408
29	1,311	1,699	2,045	2,462	2,756	3,396
30	1,310	1,697	2,042	2,457	2,750	3,385
40	1,303	1,684	2,021	2,423	2,704	3,307
60	1,296	1,671	2,000	2,390	2,660	3,232
120	1,289	1,658	1,980	2,358	2,617	3,160
∞	1,282	1,645	1,960	2,326	2,576	3,090

BẢNG XÁC SUẤT CỦA PHÂN BỐ KHI BÌNH PHƯƠNG (χ^2)

Giá trị trong bảng là của phân bố χ^2 . Cột thứ nhất là bậc tự do (df). Các cột còn lại cho ta các giá trị lý thuyết ở phần đuôi; $P(\chi^2_{df} > x^2) = P$, trong đó P là mức xác suất thể hiện ở đầu cột.

df	P					
	0,10	0,05	0,025	0,01	0,005	0,001
1	2,71	3,84	5,02	6,63	7,88	10,83
2	4,61	5,99	7,38	9,21	10,60	13,82
3	6,25	7,81	9,35	11,34	12,84	16,27
4	7,78	9,49	11,14	13,28	14,86	18,47
5	9,24	11,07	12,83	15,09	16,75	20,51
6	10,64	12,59	14,45	16,81	18,55	22,46
7	12,02	14,07	16,01	18,48	20,28	24,32
8	13,36	15,51	17,53	20,09	21,95	26,12
9	14,68	16,92	19,02	21,67	23,59	27,88
10	15,99	18,31	20,48	23,21	25,19	29,59
11	17,28	19,68	21,92	24,73	26,76	31,26
12	18,55	21,03	23,34	26,22	28,30	32,91
13	19,81	22,36	24,74	27,69	29,82	34,53
14	21,06	23,68	26,12	29,14	31,32	36,12
15	22,31	25,00	27,49	30,58	32,80	37,70
16	23,54	26,30	28,85	32,00	34,27	39,25
17	24,77	27,59	30,19	33,41	35,72	40,79
18	25,99	28,87	31,53	34,81	37,16	42,31
19	27,20	30,14	32,85	36,19	38,58	43,82
20	28,41	31,41	34,17	37,57	40,00	45,31
21	29,62	32,67	35,48	38,93	41,40	46,80
22	30,81	33,92	36,78	40,29	42,80	48,27
23	32,01	35,17	38,08	41,64	44,18	49,73
24	33,20	36,42	39,36	42,98	45,56	51,18
25	34,38	37,65	40,65	44,31	46,93	52,62
26	35,56	38,89	41,92	45,64	48,29	54,05
27	36,74	40,11	43,19	46,96	49,65	55,48
28	37,92	41,34	44,46	48,28	50,99	56,89
29	39,09	42,56	45,72	49,59	52,34	58,30
30	40,26	43,77	46,98	50,89	53,67	59,70
40	51,81	55,76	59,34	63,69	66,77	73,40
50	63,17	67,50	71,42	76,15	79,49	86,66
60	74,40	79,08	83,30	88,38	91,95	99,61
80	96,58	101,88	106,63	112,33	116,32	124,84
100	118,50	124,34	129,56	135,81	140,17	149,45

Đối với trường hợp bậc tự do lớn ta có thể tính toán như sau, áp dụng phân bố chuẩn cho χ^2 , $z = \sqrt{2\chi^2} - \sqrt{2 \times df - 1}$, và so sánh giá trị z với “Bảng xác suất của phân bố tiêu chuẩn hoá”.

BẢNG XÁC SUẤT CỦA PHÂN BỐ FISHER

Trong bảng là giá trị của phân bố Fisher F . Bậc tự do (v_1) xác định vị trí của cột và bậc tự do (v_2) xác định vị trí của hàng. Các giá trị trong bảng là giá trị lý thuyết của phần đuôi trên; $P = (F_{v_1, v_2} > f) = P$, trong đó P là xác suất (0,10; 0,05; 0,01).

v_2	P	v_1																			
		1	2	3	4	5	6	7	8	9	10	11	12	15	20	24	30	40	60	120	∞
1	0,10	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,19	60,47	60,71	61,22	61,74	62,00	62,26	62,53	62,79	63,06	63,33
	0,05	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,0	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
	0,01	4052	4999	5404	5624	5764	5859	5928	5981	6022	6056	6083	6107	6157	6209	6234	6260	6286	6313	6340	6366
2	0,10	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,40	9,41	9,42	9,44	9,45	9,46	9,47	9,47	9,48	9,49
	0,05	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
	0,01	98,50	99,00	99,16	99,25	99,30	99,33	99,36	99,38	99,39	99,40	99,41	99,42	99,43	99,45	99,46	99,47	99,48	99,48	99,49	99,50
3	0,10	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,22	5,22	5,20	5,18	5,18	5,17	5,16	5,15	5,14	5,13
	0,05	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
	0,01	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,13
4	0,10	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,91	3,90	3,87	3,84	3,83	3,82	3,80	3,79	3,78	3,76
	0,05	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
	0,01	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,45	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56	13,46
5	0,10	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,28	3,27	3,24	3,21	3,19	3,17	3,16	3,14	3,12	3,10
	0,05	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
	0,01	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,96	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	0,10	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,92	2,90	2,87	2,84	2,82	2,80	2,78	2,76	2,74	2,72
	0,05	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
	0,01	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	0,10	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,68	2,67	2,63	2,59	2,58	2,56	2,54	2,51	2,49	2,47
	0,05	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
	0,01	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65
8	0,10	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,52	2,50	2,46	2,42	2,40	2,38	2,36	2,34	2,32	2,29
	0,05	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
	0,01	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,73	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	0,10	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,40	2,38	2,34	2,30	2,28	2,25	2,23	2,21	2,18	2,16
	0,05	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
	0,01	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31

v_2	P	1	2	3	4	5	6	7	8	9	10	11	12	15	20	24	30	40	60	120	∞
10	0,10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,30	2,28	2,24	2,20	2,18	2,16	2,13	2,11	2,08	2,06
	0,05	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
	0,01	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,77	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	0,10	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25	2,23	2,21	2,17	2,12	2,10	2,08	2,05	2,03	2,00	1,97
	0,05	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
	0,01	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	0,10	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21	2,19	2,17	2,15	2,10	2,06	2,04	2,01	1,99	1,96	1,93	1,90
	0,05	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
	0,01	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,22	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
15	0,10	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06	2,04	2,02	1,97	1,92	1,90	1,87	1,85	1,82	1,79	1,76
	0,05	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
	0,01	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
20	0,10	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94	1,91	1,89	1,84	1,79	1,77	1,74	1,71	1,68	1,64	1,61
	0,05	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
	0,01	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,29	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
24	0,10	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94	1,91	1,88	1,85	1,83	1,78	1,73	1,70	1,67	1,64	1,61	1,57	1,53
	0,05	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
	0,01	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,09	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
30	0,10	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82	1,79	1,77	1,72	1,67	1,64	1,61	1,57	1,54	1,50	1,46
	0,05	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
	0,01	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,91	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	0,10	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76	1,74	1,71	1,66	1,61	1,57	1,54	1,51	1,47	1,42	1,38
	0,05	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
	0,01	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,73	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
60	0,10	2,79	2,39	2,18	2,04	1,95	1,87	1,82	1,77	1,74	1,71	1,68	1,66	1,60	1,54	1,51	1,48	1,44	1,40	1,35	1,29
	0,05	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
	0,01	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,56	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	0,10	2,75	2,35	2,13	1,99	1,90	1,82	1,77	1,72	1,68	1,65	1,63	1,60	1,55	1,48	1,45	1,41	1,37	1,32	1,26	1,19
	0,05	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,87	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
	0,01	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,40	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
∞	0,10	2,71	2,30	2,08	1,94	1,85	1,77	1,72	1,67	1,63	1,60	1,57	1,55	1,49	1,42	1,38	1,34	1,30	1,24	1,17	1,00
	0,05	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00
	0,01	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,25	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

BẢNG GIÁ TRỊ 2½% PHÍA TRÊN CỦA PHÂN BỐ FISHER F

Giá trị trong bảng là của phân bố Fisher F . Bậc tự do (v_1) xác định vị trí của cột và bậc tự do (v_2) xác định vị trí của hàng. Các giá trị trong bảng là giá trị lý thuyết tại điểm 2,5%; $P(F_{v_1, v_2} > f) = 0,025$.

v_2	v_1																			
	1	2	3	4	5	6	7	8	9	10	11	12	15	20	24	30	40	60	120	∞
1	647,8	799,5	864,2	899,6	921,8	937,1	948,2	956,6	963,3	968,6	973,0	976,7	984,9	993,1	997,3	1001	1006	1010	1014	1018
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,41	39,43	39,45	39,46	39,46	39,47	39,48	39,49	39,50
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,37	14,34	14,25	14,17	14,12	14,08	14,04	13,99	13,95	13,90
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,79	8,75	8,66	8,56	8,51	8,46	8,41	8,36	8,31	8,26
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,57	6,52	6,43	6,33	6,28	6,23	6,18	6,12	6,07	6,02
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,41	5,37	5,27	5,17	5,12	5,07	5,01	4,96	4,90	4,85
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,71	4,67	4,57	4,47	4,41	4,36	4,31	4,25	4,20	4,14
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,24	4,20	4,10	4,00	3,95	3,89	3,84	3,78	3,73	3,67
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,91	3,87	3,77	3,67	3,61	3,56	3,51	3,45	3,39	3,33
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,66	3,62	3,52	3,42	3,37	3,31	3,26	3,20	3,14	3,08
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,47	3,43	3,33	3,23	3,17	3,12	3,06	3,00	2,94	2,88
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,32	3,28	3,18	3,07	3,02	2,96	2,91	2,85	2,79	2,72
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	3,01	2,96	2,86	2,76	2,70	2,64	2,59	2,52	2,46	2,40
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,72	2,68	2,57	2,46	2,41	2,35	2,29	2,22	2,16	2,09
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64	2,59	2,54	2,44	2,33	2,27	2,21	2,15	2,08	2,01	1,94
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,57	2,51	2,46	2,41	2,31	2,20	2,14	2,07	2,01	1,94	1,87	1,79
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39	2,33	2,29	2,18	2,07	2,01	1,94	1,88	1,80	1,72	1,64
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27	2,22	2,17	2,06	1,94	1,88	1,82	1,74	1,67	1,58	1,48
120	5,15	3,80	3,23	2,89	2,67	2,52	2,39	2,30	2,22	2,16	2,10	2,05	1,94	1,82	1,76	1,69	1,61	1,53	1,43	1,31
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05	1,99	1,94	1,83	1,71	1,64	1,57	1,48	1,39	1,27	1,00

8. Tài liệu tham khảo

8.1. Tiếng Việt

- Pascal Leroy, Frederic Farnir (1999). *Thống kê sinh học*. Tài liệu dịch từ nguyên bản tiếng Pháp; người dịch Đặng Vũ Bình. Đại học Nông nghiệp I Hà Nội.
- Phạm Chí Thành (1988). *Phương pháp thí nghiệm đồng ruộng*. Đại học Nông nghiệp I Hà Nội.
- Phan Hiếu Hiền (2001). *Phương pháp bố trí thí nghiệm*. Nhà xuất bản Nông Nghiệp.
- Chu Văn Mẫn, Đào Hữu Hồ (1999). *Thống kê sinh học*. Nhà xuất bản Khoa học và kỹ thuật.
- Nguyễn Văn Thiện (1997). *Phương pháp nghiên cứu trong chăn nuôi*. Nhà xuất bản Nông nghiệp.

8.2. Tiếng Anh

- R.C. Campbell (2000). *Statistics for Biologists*. Cambridge University Press.
- Aviva Petrie and Paul Watson (2001). *Statistics for veterinary and animal science*. Blackwell Science.
- R. Mead, R.N. Curnow and A.M. Hasted (1993). *Statistical methods in agriculture and experimental biology*. Chapman & Hall/Crc.
- W.G. Cochran and G.M. Cox (1966). *Experimental Designs*. Wiley International Edition.
- D.R.Cox (1958). *Planning of experiments*. Wiley International Edition.
- Robert R. Sokal, F. James Rohlf (2000). *Biometry*. W.H. Freeman and Company.
- Mick O'Neill, Peter Thomson (2002). *Third year biometry: Experimental design, Statistical modelling*. The University of Sydney.
- Peter Thomson, Frank Nicholas, Cris Moran (2002). *Genetics and biometry*. The University of Sydney.
- Douglas C. Montgomery (1996). *Design and analysis of experiments*. Wiley International Edition.
- Harold R. Lindman (1991). *Analysis of variance in experimental design*. Springer-Verlag.
- *Meet Minitab*, release 13 for Windows[®]. Minitab Inc.
- *Minitab user's guide 1*, release 13 for Windows[®]. Minitab Inc.
- *Minitab user's guide 2*, release 13 for Windows[®]. Minitab Inc.

8.3. Tiếng Nga

- Б.А. Доспехов (1985). *Методика полевого опыта*. Агропромиздат.
- А.И. Овсянников (1976). *Основы опытного дела в животноводстве*. Колос.

8.4. Tiếng Pháp

- Claustrioux J.J. (2002). *Expérimentation, concevoir pour analyser*. Gembloux, faculté universitaire des sciences agronomique.