

Nguyễn Văn Minh Mẫn

# THỐNG KÊ và PHÂN TÍCH DỮ LIỆU TRONG KỸ THUẬT và CÔNG NGHIỆP

(BẢN LƯU HÀNH NỘI BỘ - Version  $\gamma$ , 2017)

NHÀ XUẤT BẢN

ĐẠI HỌC QUỐC GIA TP HỒ CHÍ MINH

ĐẠI HỌC QUỐC GIA TP HỒ CHÍ MINH  
TRƯỜNG ĐẠI HỌC BÁCH KHOA

Nguyễn V. M. Mẫn

**THỐNG KÊ và PHÂN TÍCH DỮ LIỆU**  
**TRONG KỸ THUẬT và CÔNG NGHIỆP**

(BẢN LƯU HÀNH NỘI BỘ - Version  $\gamma$ , 2017)

NHÀ XUẤT BẢN ĐẠI HỌC QUỐC GIA

TP HỒ CHÍ MINH - 2017

## THỐNG KÊ TRONG HÓA HỌC, MÔI TRƯỜNG, KỸ THUẬT và CÔNG NGHIỆP

**Thống kê trong Hóa học, Môi trường, Kỹ thuật và Công nghiệp** là một loạt sách giáo khoa cung cấp nền tảng và chi tiết các khái niệm, phương pháp thống kê và các bài toán nghiên cứu thực tiễn trong phát triển bền vững. Điểm chính là phát triển các giải pháp tựa *Khoa học Thống kê* trong các lĩnh vực cụ thể của khảo sát và nghiên cứu khoa học môi trường, sinh thái, cùng với các vấn đề liên quan như sản xuất công nghiệp và giao thông trong khung cảnh phát triển kinh tế bền vững.

Những quyển sách này cung cấp hỗ trợ thống kê cho các chuyên gia và nhà nghiên cứu trên nhiều lĩnh vực làm việc và nghiên cứu khác nhau; gồm khoa học môi trường, sinh thái, quản trị tài nguyên thiên nhiên, khoa học máy tính, nông nghiệp và sản xuất công nghiệp vv.

Loạt sách cũng cung cấp hỗ trợ cho sinh viên - bạn đọc theo học các khóa học về **Khoa học Dữ liệu - Dữ liệu lớn** liên quan tới các lĩnh vực trên (khi các tập dữ liệu thật được thu thập từ các lĩnh vực ấy), và để đáp ứng các nhu cầu thực tế hàng ngày lúc bạn cần ra quyết định tối ưu sử dụng dữ liệu thực tế khi mà các tập dữ liệu quan trắc được ngày càng có kích cỡ khổng lồ cũng như cấu trúc rất phức tạp.

✓ **Quyển 1: Thống kê và Phân tích Dữ liệu** - Ứng dụng trong Kỹ thuật và Công nghiệp

*Phân phối Xác suất, Thống kê Mô tả,*

*Thống kê Suy diễn, Hồi quy và Thiết kế Thí nghiệm*

**Quyển 2: Thống Kê Bayes trong Kỹ thuật**

*Ứng dụng cho Điều khiển Quá trình và Phân tích Độ Tin cậy Hệ thống*

*(Bayesian Statistics in Engineering with Applications in Process Control & Reliability Analysis)*

Trang này cố tình để trống.

# Lời Nói Đầu

---

**Khoa học Thống kê** (*Statistics, Statistical Science*) có thể được mô tả ngắn gọn là khoa học về giải quyết vấn đề và ra quyết định dựa trên dữ liệu khi ta biết rằng sự không chắc chắn hiện hữu trong dữ liệu ấy. Như thế, khoa học thống kê thực sự vượt xa sự tích lũy đơn thuần của dữ liệu.

Từ đầu thế kỷ XXI này dữ liệu đủ loại cấu trúc hiện đang ngày càng phổ biến trong xã hội, trong mọi ngành khoa học của nhân loại, thâm nhập vào các cơ quan khoa học, chính phủ và doanh nghiệp của mọi quốc gia, dù đã phát triển hay đang phát triển. Quyển sách được viết nhằm chuyển tải trung thực các kiến thức căn bản và tư duy thống kê thực tiễn tới các sinh viên năm cuối, đặc biệt khối kỹ thuật, học viên cao học và chuyên viên khu vực sản xuất công nghiệp, quản trị ...

Về cách sử dụng quyển sách, Phần I tới Phần IV dành cho sinh viên đại học, trong khi học viên cao học và nghiên cứu sinh được khuyến khích trang bị thêm Phần V gồm hai chương về **Thiết kế cách thu thập dữ liệu** và **Biến đổi dữ liệu**. Trong khung thức đó Phần I, gồm 3 chương đầu tiên cung cấp nền tảng về **Phân phối xác suất** (Chương 2) và **Thống kê mô tả** (Chương 3), sẽ hữu ích cho mọi phát triển về sau trong *Khoa học Thống kê*, không chỉ hạn chế trong quyển sách này. Chủ đề **Kiểm duyệt Dữ liệu** là rất quan trọng và cần thiết cho việc sử dụng mô hình thống kê trong kỹ thuật, sẽ được trình bày trong quyển sách tiếp theo về *Thống Kê Bayes trong Kỹ thuật - Điều khiển Quá trình và Phân tích Độ Tin cậy*. Một lưu ý quan trọng nữa là, do sử dụng phần mềm thống kê là không thể tránh khỏi khi đọc sách hay khi làm thực nghiệm, tôi chọn cách ghi số thập phân theo lối chung của các phần mềm tiên tiến Âu Mỹ, có nghĩa là ghi 1.999 trong sách là tương ứng với cách ghi 1,999 theo lối Việt Nam. Một thí dụ phổ biến khác là độ tin cậy 95% được ghi là 0.95 (như kết quả của phần mềm R) chứ không ghi lối Việt là 0,95.

Chúng tôi mong các bạn tìm thấy được niềm vui khi sử dụng được các phương pháp trong sách, dù là đang mày mò trong phòng thí nghiệm hay đang băn khoăn trước khi đưa ra một quyết định quan trọng dựa trên nhiều tập dữ liệu có cấu trúc quá đa dạng. Tác giả chân thành mong nhận được góp ý của các bạn để các quyển sau trong loạt tài liệu này hữu ích và thú vị hơn cho người đọc.

Nguyễn Văn Minh Mẫn

mannguyen@hcmut.edu.vn

Đại học Bách Khoa TP Hồ Chí Minh

Trang này cố tình để trống.

# Lời Cảm Ơn

---

Tác giả cảm ơn Trần Vĩnh Tân, một cựu thành viên của SAM Lab (2009-2011) tại Đại học Bách Khoa TP Hồ Chí Minh, người đã đóng nhiều hình vẽ biểu cảm, giúp tạo thuận lợi và gây cảm nhận thẩm mỹ cho bạn đọc nắm bắt tốt hơn các ý tưởng trừu tượng trong khoa học này. Tác giả trân trọng cảm ơn Tiến sỹ Nguyễn An Khương, giảng viên Đại học Bách Khoa TP Hồ Chí Minh, người đã dành thời gian quý báu đọc và góp ý các sai sót về suy luận toán cũng như hình thức của tài liệu này. Quyển sách được ra đời nhờ sự hỗ trợ tốt nhất của *Khoa Môi Trường và Tài Nguyên*, Đại học Bách Khoa TP Hồ Chí Minh. Đặc biệt xin gửi lời cảm ơn tới PGS. TS Lê Văn Trung, người đã cho phép tác giả sử dụng một số chất liệu và hình ảnh chuyên môn, cũng như quan tâm sâu sát tới công việc này.

Quyển sách này được chuẩn bị qua nhiều năm giảng dạy của tác giả ở các loại hình lớp và chủ đề khác nhau ở trong và ngoài ĐHQG Tp Hồ Chí Minh. Quan trọng hơn, một số chủ đề lý thuyết đã được đưa vào quyển sách do sự thu hút đặc biệt của tập dữ liệu quan trắc thực tế, qua trao đổi với các kỹ sư hay/ và lắng nghe từ các nhà nghiên cứu thực nghiệm, nảy sinh sau các seminar hay một cuộc trao đổi học thuật. Trong các ngữ cảnh đó, chính sự tương tác, giao tiếp ý tưởng với các bạn đồng nghiệp là yếu tố chính định hình cấu trúc cũng như nội dung của tài liệu này.

Một cách cụ thể, dữ liệu và câu hỏi nghiên cứu trong ví dụ mở đầu của Chương 6 có được qua trao đổi với Tiến sỹ Đỗ Thanh Nghị, Đại học Cần Thơ, dữ liệu về nghiên cứu trong sinh thái thực vật ở Chủ đề 23 là đóng góp của Tiến sỹ Nguyễn Thị Lan Thi, Khoa Sinh học của Đại học Khoa học TP Hồ Chí Minh, nhân dịp workshop SMA năm 2015, Tp Cần Thơ; dữ liệu và vấn đề nghiên cứu trong Mục 9.3 xuất phát từ sản xuất tại chi nhánh công ty Sanyo ở ngoại ô Sài Gòn, qua kỹ sư Võ Ngọc Thiên Ân... Các tập dữ liệu khác trong các ví dụ còn đến từ các học viên cao học ở khoa Khoa học Máy tính, khoa Môi trường và Tài nguyên, khoa Khoa học Ứng dụng của Đại học Bách Khoa TP Hồ Chí Minh ...

Sau nữa nếu không có sự ủng hộ tinh thần bền bỉ và hỗ trợ vật chất quý giá của PGS. Trần Đan Thư và PGS. Nguyễn Đình Thúc ở Đại học Khoa học Sài Gòn trong nỗ lực duy trì loạt seminar học thuật về mô hình hóa thống kê (SMA workshop) thì quyển sách này không thể ra đời. Tác giả chân thành cảm ơn quý đồng nghiệp qua sự cộng tác đó, và mong muốn nhận được đóng góp, lắng nghe ý kiến của các bạn trong tương lai.

Nguyễn Văn Minh Mẫn

mannguyen.hcmut@gmail.com

TP Hồ Chí Minh, tháng 2 - 2017

## Bảng Thuật ngữ Anh - Việt viết tắt

Từ viết tắt	Thuật ngữ tiếng Anh	Thuật ngữ tiếng Việt
ANOVA	Analysis of Variance	Phân tích phương sai
ANSI	American National Standard Institute	Viện Tiêu chuẩn quốc gia Mỹ
AQL	Acceptable Quality Level	Mức chất lượng chấp nhận được
ARL	Average Run Length	Chiều dài bước chạy trung bình
ASQ	American Society for Quality	Hội Chất lượng Hoa kỳ
BP	Bootstrap Population	Dân số Bootstrap
c.d.f.	cumulative distribution function	hàm phân phối tích lũy
CLT	Central Limit Theorem	Định lý giới hạn trung tâm
CNC	Computerized Numerically Controlled	Điều khiển số điện toán hóa
CPA	Circuit Pack Assemblies	Dây chuyền đóng gói mạch in
CUSUM	Cumulative Sum	Tổng tích lũy
DoE	Design of Experiments	Thiết kế thí nghiệm
EBD	Empirical Bootstrap Distribution	Phân phối Bootstrap thực nghiệm
EWMA	Exponentially Weighted Moving Average	Trung bình trượt có trọng và mũ hóa
GRR	Gage Repeatability and Reproducibility	Gage lặp lại và tái sinh được
i.i.d.	independent and identically distributed	độc lập và phân phối giống nhau
InfoQ	Information Quality	Chất lượng thông tin
IQR	Inter Quartile Range	Khoảng phân vị giữa
ISC	Short Circuit Current of Solar Cells (in Ampere)	Dòng mạch ngắn của các Tế bào năng lượng mặt trời
KS	Kolmogorov-Smirnov	Kiểm định Kolmogorov-Smirnov
LCL	Lower Control Limit	Cận kiểm soát dưới
LLN	Law of Large Numbers	Luật số lớn
LSE	Least Squares Estimators	Ước lượng bình phương tối thiểu
LQL	Limiting Quality Level	Mức chất lượng giới hạn
LSL	Lower Specification Limit	Cận thông số kỹ thuật dưới
LWL	Lower Warning Limit	Giới hạn cảnh báo dưới
m.g.f.	moment generating function	hàm sinh moment
MLE	Maximum Likelihood Estimator	Ước tính cơ may cực đại
MSD	Mean Squared Deviation	Độ lệch bình phương trung bình
MTBF	Mean Time Between Failures	Thời gian trung bình giữa các lỗi (thất bại)
MTTF	Mean Time To Failure	Thời gian trung bình tới khi gặp lỗi
OC	Operating Characteristic	Đặc trưng vận hành
p.d.f.	probability density function	hàm mật độ xác suất



Từ viết tắt	Thuật ngữ tiếng Anh	Thuật ngữ tiếng Việt
PERT	Project Evaluation and Review Technique	Kỹ thuật Đánh giá và Thẩm định dự án
PFA	Probability of False Alarm	Khả năng báo động sai
PPM	Defects in Parts Per Million	Số lỗi (khuyết tật) phần triệu
PSE	Practical Statistical Efficiency	Hiệu quả thống kê thực tế
QbD	Quality by Design	Chất lượng nhờ thiết kế
QQ-Plot	Quantile vs. Quantile Plot	Sơ đồ phân vị so với phân vị
RCBD	Randomized Complete Block Design	Thiết kế Khối đầy đủ ngẫu nhiên
RSWOR	Random Sample Without Replacement	Lấy mẫu ngẫu nhiên không hoàn lại
RSWR	Random Sample With Replacement	Lấy mẫu ngẫu nhiên có hoàn lại
SE	Standard Error	Lỗi chuẩn
SPC	Statistical Process Control	Kiểm soát quá trình một cách thống kê
SPRT	Sequential Probability Ratio Test	Kiểm định tỷ số xác suất tuần tự
SSE	Sum of Squares of Residuals	Tổng bình phương các số (thặng) dư
SSR	Sum of Squares Around the Regression	Tổng bình phương do phép hồi quy
SST	Total Sum of Squares	Tổng bình phương toàn bộ
STD	Standard Deviation	Độ lệch chuẩn
UCL	Upper Control Limit	Cận kiểm soát trên
USL	Upper Specification Limit	Cận thông số kỹ thuật trên
UWL	Upper Warning Limit	Giới hạn cảnh báo trên
WSP	Wave Soldering Process	Quá trình hàn sóng

Trang này cố tình để trống.

# Mục lục

---

<i>Lời nói đầu</i> . . . . .	iii
<i>Lời cảm ơn</i> . . . . .	v
<i>Bảng Thuật ngữ Anh - Việt viết tắt</i> . . . . .	viii
<b>Phần I: Xác suất và Thống kê mô tả</b> . . . . .	<b>1</b>
<b>Chương 1 Khoa học Thống kê và Dữ liệu</b> . . . . .	
<i>Số liệu thực có ý nghĩa trong Kỹ thuật không?</i> . . . . .	<b>3</b>
<b>1.1 Thống kê học là gì?</b> . . . . .	<b>3</b>
<b>1.2 Tổng thể, mẫu, tham số, biến và thống kê</b> . . . . .	<b>4</b>
<b>1.3 Dữ liệu kỹ thuật và đặc điểm</b> . . . . .	<b>5</b>
1.3.1 Đặc điểm dữ liệu môi trường- kỹ thuật . . . . .	5
1.3.2 Xử lý số liệu quan trắc . . . . .	5
1.3.3 Sử dụng dữ liệu quan trắc môi trường - sinh thái hay kỹ thuật để làm gì? . . . . .	6
<b>1.4 Xác suất cơ bản</b> . . . . .	<b>7</b>
1.4.1 Biến cố và không gian mẫu: Diễn tả hình thức của thí nghiệm . . . . .	8
1.4.2 Phép toán cơ bản với các biến cố – phép hợp, giao tập hợp . . . . .	9
1.4.3 Xác suất của biến cố . . . . .	11
1.4.4 Xác suất có điều kiện và sự độc lập của các biến cố . . . . .	13
1.4.5 Công thức Bayes và ứng dụng của nó . . . . .	16
<b>1.5 Sơ lược về biến ngẫu nhiên</b> . . . . .	<b>18</b>
1.5.1 Biến ngẫu nhiên rời rạc . . . . .	19
1.5.2 Biến ngẫu nhiên liên tục . . . . .	21
<b>1.6 Vấn đề thực tế sử dụng lý thuyết về biến ngẫu nhiên</b> . . . . .	<b>23</b>
<b>1.7 Giới thiệu R</b> . . . . .	<b>24</b>
<b>1.8 Tổng kết Chương 1</b> . . . . .	<b>25</b>
1.8.1 Bài tập Chương 1 . . . . .	25
1.8.2 Tóm lược . . . . .	27

**Chương 2 Xác suất và phân phối xác suất**

<i>Cơ sở của Thống kê và Khoa học thực nghiệm</i>	<b>33</b>
<b>2.1 Bài toán thực tiễn cần sử dụng phân phối xác suất</b>	<b>33</b>
<b>2.2 Các phân phối xác suất rời rạc</b>	<b>34</b>
2.2.1 Phân phối Bernoulli	34
2.2.2 Phân phối nhị thức	35
2.2.3 Phân phối Poisson	39
<b>2.3 Phân phối xác suất rời rạc trong Môi trường - Tài nguyên</b>	<b>43</b>
2.3.1 Quá trình Bernoulli và phân phối nhị thức	43
2.3.2 Quá trình Poisson	44
<b>2.4 Các phân phối xác suất liên tục</b>	<b>46</b>
2.4.1 Phân phối đều trên khoảng số thực	46
2.4.2 Phân phối Gauss và Gauss chuẩn tắc	46
2.4.3 Phân phối log-chuẩn	55
2.4.4 Phân phối mũ (exponential distribution)	56
2.4.5 Phân phối “Student” $T$	59
<b>2.5 Phân phối xác suất liên tục trong Kỹ thuật</b>	<b>61</b>
<b>2.6 Độ lệch chuẩn, phân vị, độ đo méo lệch và độ nhọn</b>	<b>62</b>
<b>2.7 Sự xấp xỉ mẫu lớn</b>	<b>64</b>
2.7.1 Luật số lớn	64
2.7.2 Định lý giới hạn trung tâm	66
2.7.3 Xấp xỉ một số phân phối bằng phân phối Gauss	66
<b>2.8 Tổng kết Chương 1 và 2</b>	<b>68</b>
2.8.1 Bài tập	69
2.8.2 Tóm lược khái niệm và thuật ngữ quan trọng	69
2.8.3 English Summary of Terms	70
2.8.4 Câu hỏi trắc nghiệm cuối chương	71

**Chương 3 Thống kê mô tả**

<i>Làm cho số liệu quan trắc có ý nghĩa</i>	<b>73</b>
<b>3.1 Giới thiệu</b>	<b>73</b>
<b>3.2 Hiện thị dữ liệu mẫu</b>	<b>74</b>
3.2.1 Bảng kê (Table)	74
3.2.2 Biểu đồ (Chart)	74
3.2.3 Sử dụng biểu đồ hay đồ thị	78
<b>3.3 Sai số, độ chuẩn xác và độ chính xác của các phép đo</b>	<b>78</b>

3.3.1	Sai số và phân loại	78
3.3.2	Hiển thị đồ thị sai số	80
3.3.3	Độ chuẩn xác và độ chính xác của các phép đo	80
3.4	Phân tích mô tả các giá trị mẫu dùng biến ngẫu nhiên	82
3.4.1	Tần suất của biến ngẫu nhiên rời rạc	82
3.4.2	Tần suất của biến ngẫu nhiên liên tục	87
3.4.3	Thống kê của mẫu được sắp thứ tự	90
3.4.4	Thống kê vị trí và thống kê phân tán	92
3.5	Các khoảng dự đoán	97
3.6	Kỹ thuật bổ sung cho phân tích dữ liệu thăm dò	98
3.6.1	Biểu đồ hộp	99
3.6.2	Thể hiện sự biến thiên thống kê và độ chính xác thống kê	101
3.6.3	Sơ đồ phân vị	101
3.6.4	Các thống kê bền vững cho vị trí và độ phân tán	102
3.7	Tổng kết Chương 3	104
3.7.1	Bài tập	104
3.7.2	Tóm lược	105
3.7.3	English Summary of Terms	106
3.7.4	Extra problems in English	108
3.8	Chủ đề ôn Phần I	110
3.8.1	Phần lý thuyết xác suất cơ bản	111
3.8.2	Phần phân phối xác suất	111
3.8.3	Phần thống kê mô tả	113

## Phần II: Suy diễn Thống kê

114

### Chương 4 Ước lượng thống kê

	Xác định giá trị của các tham số tổng thể	117
4.1	Giới thiệu	117
4.2	Ước lượng tham số đặc trưng của tổng thể	118
4.2.1	Mẫu và thống kê mẫu cho ước lượng tham số tổng thể	118
4.2.2	Về ước lượng tham số thống kê	119
4.3	Một số phương pháp ước lượng điểm	121
4.3.1	Ước lượng điểm nhờ tối đa hàm cơ may	122
4.3.2	Ước lượng tham số dùng phương trình moment (Tùy chọn)	124
4.4	Phân phối chi bình phương (chi-squared)	125
4.4.1	Tính chất phân phối chi bình phương	126
4.4.2	Phân phối của phương sai mẫu	126

4.5	<b>Khoảng tin cậy cho một tham số tổng thể</b>	129
4.5.1	Khoảng tin cậy cho trung bình $\mu$ khi biết $\sigma$	130
4.5.2	Khoảng tin cậy cho trung bình tổng thể khi chưa biết độ lệch chuẩn $\sigma$	131
4.5.3	Khoảng tin cậy cho phương sai (phân phối Gauss)	133
4.5.4	Khoảng tin cậy cho tỷ lệ $p$	135
4.6	<b>Khoảng dung sai</b>	137
4.6.1	Một ứng dụng tính khoảng dung sai trong kỹ thuật	138
4.6.2	Khoảng dung sai cho các phân phối Gauss	138
4.7	<b>Tổng kết Chương 4</b>	140
<b>Chương 5 Kiểm định giả thuyết</b>		
	Xác nhận khẳng định về các tham số tổng thể	143
5.1	<b>Giới thiệu</b>	143
5.1.1	Liệu chúng ta có thể khai thác hữu hiệu hơn nữa các dữ liệu thực?	143
5.1.2	Giả thuyết thống kê - Các khái niệm cơ bản	143
5.1.3	Kiểm định giả thuyết (một cách) thống kê	145
5.2	<b>Giá trị <math>P</math> (p-value) của kiểm định</b>	146
5.2.1	Sử dụng giá trị $P$ khi xét tham số $\theta = \mu$ (trung bình tổng thể)	147
5.2.2	Kết luận về giá trị $P$	148
5.3	<b>Kiểm định giả thuyết một mẫu thông dụng</b>	149
5.3.1	Thống kê $Z$ kiểm định trung bình tổng thể $\mu$ theo phân phối chuẩn	150
5.3.2	Thống kê $T$ kiểm định trung bình của phân phối Gauss, chưa biết $\sigma^2$	153
5.3.3	Kiểm định giả thuyết theo tỷ lệ 2 đại lượng trong 1 tổng thể	154
5.3.4	Kiểm định giả thuyết trên phương sai tổng thể	156
5.4	<b>Xác định cỡ mẫu khi kiểm định trên một tổng thể</b>	157
5.4.1	Tính sai lầm kiểm định loại II $\beta$	157
5.4.2	Xác định cỡ mẫu khi kiểm định cho trung bình tổng thể	158
5.5	<b>Kiểm định so sánh hai trung bình tổng thể</b>	159
5.5.1	So sánh 2 giá trị trung bình- Mẫu lớn $n \geq 30$	159
5.5.2	So sánh 2 giá trị trung bình- Mẫu nhỏ	159
5.6	<b>Kiểm định mức phù hợp mô hình với dữ liệu</b>	160
5.7	<b>Tổng kết Chương 5</b>	163
5.8	<b>Chủ đề ôn Phần II</b>	164
5.8.1	Kiểm định giả thuyết trong kỹ nghệ	164
5.8.2	Ước lượng sự khác biệt giữa hai trung bình	165
5.8.3	Câu hỏi trắc nghiệm Phần II	166

## Phần III: Hồi Quy và Phân tích Phương sai 167

### Chương 6 Hồi quy và Tương quan

<i>Các mối quan hệ giữa các biến</i>	169
<b>6.1 Giới thiệu</b>	169
6.1.1 Hồi quy trong khoa học thực nghiệm	169
6.1.2 Hiệp phương sai và hệ số tương quan mẫu	172
6.1.3 Hai trường hợp đặc biệt của hệ số tương quan	176
6.1.4 Dự báo thống kê (statistical prediction) dùng dữ liệu hiện trường	176
<b>6.2 Hồi quy tuyến tính đơn biến</b>	179
6.2.1 Phương pháp bình phương tối thiểu (hay nhỏ nhất)	179
6.2.2 Mô hình hồi quy và khoảng dự báo	186
<b>6.3 Hiệp phương sai và Hệ số tương quan đa biến</b>	187
<b>6.4 Phân tích hồi quy đa biến</b>	191
6.4.1 Dẫn nhập	191
6.4.2 Hồi quy hai biến	192
<b>6.5 Xây dựng mô hình: Thuật toán hồi quy từng bước</b>	197
<b>6.6 Phương pháp phân tích kết cục hữu hạn: Hồi quy logistic</b>	198
<b>6.7 Biểu diễn và phân tích đồ họa</b>	200
6.7.1 Biểu đồ phân tán	200
6.7.2 Biểu đồ hộp đa biến	203
<b>6.8 Tổng kết Chương 6</b>	206
6.8.1 Tóm lược Chương 6	206
6.8.2 English Summary of Terms	206
6.8.3 Bài tập Chương 6	206
6.8.4 Giải pháp cho Vấn đề 6	207

### Chương 7 Phân tích Phương sai

<i>Kiểm tra sự khác biệt giữa nhiều trung bình</i>	213
<b>7.1 Phân phối <math>F</math> cho biến tỉ lệ hai phương sai</b>	215
<b>7.2 Phân tích phương sai (ANOVA) đơn giản</b>	216
7.2.1 Kiểm định so sánh các trung bình	216
7.2.2 Phương pháp phân tích phương sai đơn giản	217
7.2.3 Về các giả định cần thiết cho Phân tích phương sai	220
<b>7.3 Bảng tiếp liên</b>	222
7.3.1 Cấu trúc của các bảng tiếp liên	222
7.3.2 Chỉ số liên kết trong bảng tiếp liên	226
<b>7.4 Tổng kết Chương 7</b>	231

7.4.1	Tóm lược Chương 7	231
7.4.2	Bài tập Chương 7	232
7.5	Câu hỏi trắc nghiệm Phần III	232
7.6	Chủ đề ôn Phần III: Xây dựng mô hình thực nghiệm	234
7.6.1	Mô hình thực nghiệm phi tuyến	234
7.6.2	Mô hình bồi lắng (sedimentation)	234
7.6.3	Phân tích bằng hồi quy tuyến tính	235

## Phần IV: Thiết kế và Phân tích Thí nghiệm 240

### Chương 8 Thiết kế Thí nghiệm Thống kê

	Tính toán khung thức thu thập dữ liệu	241
8.1	Giới thiệu về lịch sử và nội dung	242
8.1.1	Thế nào là thiết kế thí nghiệm?	242
8.1.2	Lịch sử và nội dung	242
8.2	Vai trò của Thiết kế Thí nghiệm Thống kê trong kỹ thuật	244
8.3	Các định nghĩa căn bản và Ứng dụng thực tiễn	245
8.4	Vài nguyên lý và công cụ cho thiết kế thí nghiệm thống kê	248
8.4.1	Phân khối và ngẫu nhiên hóa	248
8.4.2	Công cụ cơ bản	249
8.5	Mô hình cộng tính và không cộng tính	250
8.5.1	Thiết kế ngẫu nhiên hoàn toàn một yếu tố (CRD - Completely Randomized Design)	251
8.5.2	Trường hợp nhiều yếu tố	255
8.6	Phân tích các thiết kế khối hoàn chỉnh và ngẫu nhiên	256
8.6.1	Trường hợp nhiều khối, hai giải pháp cho mỗi khối: so sánh ghép cặp	256
8.6.2	Trường hợp nhiều khối, $t$ giải pháp cho mỗi khối	259
8.7	Tổng kết Chương 8	261
8.7.1	Tóm lược	261
8.7.2	English Summary of Terms	261
8.7.3	Bài tập Chương 8	261
8.8	Chủ đề ôn Chương 8	262

### Chương 9 Thiết kế và Phân tích Thí nghiệm Nhân tố

	Kết quả tối ưu khi có nhiều yếu tố ảnh hưởng	269
9.1	Thiết kế khối không đầy đủ và cân bằng (BIBD)	270
9.1.1	Các định nghĩa tổ hợp	270
9.1.2	Phân tích ANOVA cho một BIBD	271



<b>9.2 Thiết kế nhân tố đầy đủ và Phân tích ANOVA hai biến</b>	274
9.2.1 Cấu trúc của thiết kế nhân tố (yếu tố)	274
9.2.2 Phân tích ANOVA cho thiết kế đầy đủ 2 nhân tố (Two-way Anova)	275
<b>9.3 Thiết kế nhân tố trong sản xuất và kỹ nghệ chất lượng</b>	282
9.3.1 Quy trình dùng thiết kế thí nghiệm (DOE) trong công nghiệp	282
9.3.2 Điều khiển Chất lượng Thống kê	283
9.3.3 Thiết Kế Thí Nghiệm (DOE) và Nghiên cứu Khoa học	284
9.3.4 Nhu cầu đảm bảo chất lượng trong sản xuất công nghiệp	285
<b>9.4 Thiết kế nhân tố (FD- Factorial Designs)</b>	286
9.4.1 Khái niệm thiết kế nhân tố đầy đủ thuần	286
9.4.2 Sử dụng một thiết kế nhân tố đầy đủ $2^4$ trong sản xuất kỹ nghệ	286
9.4.3 Thiết kế nhân tố một phần thuần	289
9.4.4 Thiết kế nhân tố một phần hỗn hợp - Sản xuất đồ nội thất gỗ?	291
<b>9.5 Khảo sát các mặt đáp ứng</b>	292
9.5.1 Giới thiệu	292
9.5.2 Các thiết kế bậc hai	294
9.5.3 Vài thiết kế bậc hai cụ thể	296
<b>9.6 Tổng kết Chương 9</b>	301
<b>9.7 Toán học cho Thiết Kế Thí Nghiệm</b>	302
9.7.1 Khái niệm hình vuông Latin	302
9.7.2 Lý thuyết thiết kế tổ hợp	303
9.7.3 Ứng dụng của thiết kế tổ hợp	305
<b>9.8 Chủ đề nâng cao trong Thiết Kế Thí Nghiệm (DOE)</b>	306
9.8.1 Mạng trực giao	306
<b>9.9 Chủ đề ôn Phần IV</b>	309

## **Phần V: Thiết kế cách thu thập dữ liệu và Biến đổi dữ liệu** 310

### **Chương 10 Thiết kế cách thu thập dữ liệu**

<i>Làm sao bạn lấy mẫu chứa thông tin hữu ích?</i>	311
<b>10.1 Giới thiệu</b>	311
10.1.1 Tại sao kỹ sư cần thiết kế cách thu thập dữ liệu?	311
10.1.2 Vấn đề thực tế	312
<b>10.2 Lấy mẫu và Bài toán Ước lượng</b>	313
10.2.1 Các bước trong quy trình lấy mẫu cho nghiên cứu	313
10.2.2 Định nghĩa căn bản trong Lý thuyết lấy mẫu	314
10.2.3 Lấy mẫu ngẫu nhiên từ một tổng thể hữu hạn	315
10.2.4 Hàm xác suất cho việc lấy mẫu ngẫu nhiên	316

10.2.5 Phân bố lấy mẫu . . . . .	317
10.3 Ước lượng với các mẫu ngẫu nhiên đơn giản . . . . .	320
10.3.1 Khoảng tin cậy cho trung bình $\mu$ khi biết độ lệch chuẩn $\sigma$ . . . . .	321
10.3.2 Tính chất của trung bình và độ lệch chuẩn mẫu dưới cơ chế RSWR . . . . .	323
10.3.3 Tính chất trung bình mẫu và phương sai mẫu dưới cơ chế RSWOR . . . . .	325
10.4 Ước lượng trung bình sử dụng lấy mẫu phân tầng . . . . .	328
10.5 Tổng kết Chương 10 . . . . .	330
<b>Chương 11 Phép Biến đổi Dữ liệu</b>	
<i>Kỹ thuật khớp dữ liệu với giả định thống kê</i>	<b>333</b>
11.1 Tổng quan về phép biến đổi dữ liệu . . . . .	334
11.1.1 Biến đổi dữ liệu nhằm có phương sai đồng nhất . . . . .	335
11.1.2 Các phép biến đổi dữ liệu thông dụng . . . . .	335
11.1.3 Biến đổi BOX – COX . . . . .	337
11.2 Làm trơn dữ liệu . . . . .	339
11.2.1 Giới thiệu . . . . .	339
11.2.2 Biến đổi trung bình động (MA - Moving Average) . . . . .	339
11.2.3 Biến đổi trung bình trượt theo cấp độ mũ (trọng số mũ hóa) . . . . .	341
11.3 Áp dụng của biến đổi dữ liệu . . . . .	343
11.3.1 Áp dụng 1: Tính kích thước mẫu khi được cung cấp bán kính sai số . . . . .	343
11.3.2 Áp dụng 2: Tính độ chính xác ( <i>precision</i> ) của các biến ngẫu nhiên . . . . .	344
11.4 Khoảng tin cậy qua biến đổi dữ liệu . . . . .	346
11.4.1 Khoảng tin cậy cho trung bình $\mu$ khi chưa biết $\sigma$ hay cỡ mẫu nhỏ . . . . .	346
11.4.2 Khoảng tin cậy qua phép biến đổi . . . . .	347
<b>Chương 12 PHỤ LỤC và TÓM LƯỢC</b>	<b>349</b>
12.1 Moment và hàm sinh của biến ngẫu nhiên . . . . .	349
12.1.1 Kỳ vọng (trung bình) và các moment của các phân phối . . . . .	350
12.1.2 Hàm sinh moment . . . . .	352
12.2 Tổ hợp tuyến tính của nhiều biến ngẫu nhiên . . . . .	353
12.2.1 Kỳ vọng và phương sai của tổ hợp tuyến tính các biến ngẫu nhiên . . . . .	353
12.2.2 Kỳ vọng và phương sai của trung bình mẫu ngẫu nhiên . . . . .	353
12.3 Đảm bảo chất lượng thông tin khi phân tích dữ liệu . . . . .	354
12.4 Sự phân bố mẫu theo tỉ lệ và tối ưu . . . . .	356
12.4.1 Phân bố mẫu theo tỉ lệ . . . . .	356
12.4.2 Lấy mẫu phân tầng tối ưu . . . . .	359
12.5 Nghiên cứu năng lực quá trình dùng biểu đồ kiểm soát . . . . .	359

12.5.1	Các loại biểu đồ kiểm soát và cách dùng . . . . .	359
12.5.2	Phân loại biểu đồ kiểm soát theo kiểu dữ liệu . . . . .	360
12.5.3	Hai dạng thức nghiên cứu Năng lực Quá trình . . . . .	363
12.5.4	Các chỉ số năng lực quá trình $C_p$ và $C_{pk}$ . . . . .	363
12.5.5	Ước lượng các chỉ số năng lực quá trình . . . . .	364
12.6	Các chủ đề ôn của quyển sách . . . . .	366
Chỉ mục	. . . . .	372

Trang này cố tình để trống.

# Danh sách hình vẽ

---

1.1	Đồ thị chuỗi thời gian . . . . .	6
1.2	Luật xác suất tổng . . . . .	12
1.3	Hàm mật độ và hàm tích lũy xác suất của biến rời rạc. . . . .	20
1.4	Hàm tích lũy xác suất của $F(x) = 1 - e^{-x}$ . . . . .	23
2.1	Hàm mật độ của $B(50, p)$ với $p = .25, .50, .75$ . . . . .	38
2.2	Hàm mật độ Poisson khi $\lambda = 5, 10, 15$ . . . . .	40
2.3	Một số biểu đồ tần suất của phân phối Poisson. . . . .	42
2.4	Hàm mật độ của $N(\mu, \sigma)$ với $\mu = 10, \sigma = 1, 2, 3$ . . . . .	47
2.5	Hàm tích lũy xác suất $\Phi(x)$ của phân phối Gauss chuẩn . . . . .	50
2.6	Tính đối xứng của phân phối chuẩn (Gauss) . . . . .	51
2.7	Độ méo nhọn của phân phối $LN(\mu, \sigma)$ . . . . .	56
2.8	Hàm mật độ của $E(\beta)$ với $\beta = 1, 2, 3$ . . . . .	57
2.9	Hàm mật độ của $t[\nu]$ với $\nu = 5, 50$ . . . . .	60
2.10	Tham số đặc trưng cho khuynh hướng trung tâm và hình dạng đường cong hàm mật độ . . . . .	63
3.1	Xây dựng biểu đồ hình tròn nhằm thể hiện quy mô của hiện tượng . . . . .	75
3.2	Biểu đồ chuỗi thời gian thể hiện kim loại nặng thay đổi theo thời gian . . . . .	76
3.3	Sử dụng biểu đồ hay đồ thị . . . . .	77
3.4	Đồ thị sai số . . . . .	79
3.5	Phân biệt độ lệch (bias), độ chuẩn xác và độ chính xác. . . . .	81
3.6	Mẫu của 10 phép đo trên ba thiết bị khác nhau. . . . .	81
3.7	Biểu đồ cột của một phân phối tần số . . . . .	83
3.8	Hàm bước nhảy (bậc thang) của một tần suất tích lũy . . . . .	84
3.9	Biểu đồ cột về số nhược điểm trên tấm gốm . . . . .	85
3.10	Tần suất tích lũy của số lượng nhược điểm trên các tấm gốm . . . . .	86
3.11	Biểu đồ tần số và tần suất tích lũy xác suất của biến liên tục. . . . .	88

3.12 Tần suất tích lũy với các đường thẳng nội suy tuyến tính tại các tứ phân vị. . . . .	92
3.13 Các phân phối đối xứng và bất đối xứng . . . . .	94
3.14 Phân phối chuẩn, dốc và bằng phẳng . . . . .	95
3.15 Sơ đồ hộp của dữ liệu logarithm độ bền sợi . . . . .	100
3.16 Biểu đồ Box đầy đủ và thu gọn . . . . .	100
3.17 Hình vẽ phân vị của logarit độ bền sợi . . . . .	101
4.1 Sự liên quan giữa thu thập dữ liệu với phân tích dữ liệu - vai trò của ước lượng thống kê . . . . .	119
4.2 Khoảng tin cậy mô phỏng cho kỳ vọng của phân phối chuẩn $N(10, 1)$ với mẫu cỡ $n = 10$ . . . . .	128
4.3 Khoảng ước lượng có bán kính lỗi 1, 2 và 3 sai số chuẩn - khi ước lượng mẫu xấp xỉ Gauss . . . . .	129
4.4 Ước lượng khoảng tin cậy trung bình tổng thể dùng phân phối Student . . . . .	132
4.5 Phân vị trái và phải của khoảng ước lượng của phương sai $\sigma^2$ . . . . .	134
4.6 Hãy tưởng tượng phần màu đỏ là phần đại dương bị ô nhiễm trong dòng hải lưu trắng... . . . .	142
5.1 Quá trình học sinh ra dữ liệu - tri thức, cho phép con người giải quyết vấn đề. . . . .	144
5.2 Ba trường hợp tính giá trị $P$ theo Tính chất 9 . . . . .	147
5.3 Vùng tới hạn cho kiểm định $Z$ một phía . . . . .	150
5.4 Vùng tới hạn cho kiểm định $Z$ hai phía . . . . .	152
5.5 Sử dụng phân phối tham chiếu $\chi^2[n - 1]$ . . . . .	156
5.6 Sử dụng phân phối $N\left(\frac{\delta\sqrt{n}}{\sigma}, 1\right)$ tính $\beta$ . . . . .	158
5.7 Histogram của 100 trung bình mẫu. . . . .	162
6.1 Yếu tố nào ảnh hưởng nhất lên sự phát triển dân số rầy nâu (BPH)? . . . . .	170
6.2 Mô hình dẫn đến phương pháp bình phương tối thiểu (LSE) . . . . .	172
6.3 Các tổng bình phương lỗi, đường thẳng màu cam cho tổng $SSE$ nhỏ nhất . . . . .	173
6.4 Hệ số tương quan mẫu và diễn giải . . . . .	174
6.5 Các loại mô hình hồi quy . . . . .	176
6.6 Dữ liệu nước sông Thị Vải và sơ đồ phân tán của các biến đo lường . . . . .	178
6.7 Sai số và đường thẳng xấp xỉ . . . . .	180
6.8 Mối quan hệ của các giá trị ISC tại $t_1$ và $t_2$ . . . . .	182
6.9 Số dư so với các giá trị ISC dự đoán . . . . .	185
6.10 Các khoảng dự đoán cho các giá trị Res 3, cho trước các giá trị Res 7 . . . . .	187
6.11 Biểu đồ tán xạ đa biến của các phép đo trên các chốt tán nhôm . . . . .	190
6.12 Biểu đồ phân tán của $\hat{e}$ so với $\hat{Y}$ . . . . .	196
6.13 Biểu đồ tán xạ của hai biến ' $y$ -dev' so với ' $x$ -dev' . . . . .	201
6.14 Ma trận đa phân tán của tập dữ liệu <b>PLACE.csv</b> . . . . .	202
6.15 Biểu đồ tần suất của biến $\theta$ -dev trong tập dữ liệu <b>PLACE.csv</b> . . . . .	203

6.16	Biểu đồ tán xạ 3 chiều. . . . .	203
6.17	Biểu đồ hộp đa biến của $x$ so với số hiệu bảng mạch in (MINITAB). . . . .	204
6.18	Biểu đồ phân tán của $y$ -dev với $x$ -dev theo các biến mã hóa . . . . .	205
7.1	Hàm mật độ của $F(\nu_1, \nu_2)$ . . . . .	216
7.2	Biểu đồ hình thanh cho tỉ lệ lỗi của từng loại linh kiện. . . . .	224
7.3	Biểu đồ các giá trị lỗi dự đoán . . . . .	237
7.4	GS R.A. Fisher, nhà phát minh Thiết Kế Thí Nghiệm (DOE). . . . .	240
8.1	Hàm mật độ của $F(\nu_1, \nu_2)$ . . . . .	250
9.1	Ảnh hưởng của hệ số co giãn lên chu kỳ (Trục $y$ tương ứng với chu kỳ tính theo phút). . . . .	279
9.2	Biểu đồ tương tác của trọng lượng piston với hệ số co giãn (độ cứng). . . . .	281
9.3	An actual camera tuning system . . . . .	285
9.4	The actual data set before and after optimizing . . . . .	289
9.5	Hình vẽ tác động chính. . . . .	300
9.6	Đường đồng mức của mặt đáp ứng ở Bảng 9.14. . . . .	301
10.1	Quá trình học sinh ra dữ liệu - tri thức, cho phép con người giải quyết vấn đề. . . . .	312
10.2	Histogram của 1000 trung bình mẫu. . . . .	319
11.1	Phương sai đồng nhất cho các bộ dữ liệu khác nhau? . . . . .	334
11.2	Phép biến đổi dữ liệu cho phương sai đồng nhất . . . . .	336
11.3	Biến đổi logarit khi số liệu thỏa điều kiện $s^2 > \bar{y}$ . . . . .	337
11.4	Biến đổi Box-Cox . . . . .	338
11.5	Biến đổi trung bình động dữ liệu PCB . . . . .	340
11.6	Mô hình trung bình trượt theo cấp độ mũ so với trung bình trượt đơn giản MA . . . . .	342
12.1	Mô thức để phát hiện các nguyên nhân đặc biệt: tiếp theo. . . . .	361
12.2	Phân loại các sơ đồ điều khiển. . . . .	362
12.3	Biểu đồ hộp của $Y$ theo từng nhà sản xuất . . . . .	369

Trang này cố tình để trống.



# Danh sách bảng

---

1.1 Khoảng chu kỳ của piston (tính bằng giây) với các yếu tố kiểm soát được thiết lập ở mức tối thiểu . . . . .	24
2.1 Giá trị của hàm mật độ và hàm tích lũy xác suất của phân phối $B(30, 0.6)$ . . . . .	38
2.2 Phân phối nhị thức cho $np = 2$ và phân phối Poisson với $\lambda = 2$ . . . . .	41
2.3 Phân phối Poisson của các hạt bụi trong không khí . . . . .	45
2.4 Phân vị thứ $\alpha$ của hàm mật độ của $T$ . . . . .	60
3.1 Diện tích đất nông nghiệp ở miền Nam VN . . . . .	74
3.2 Biểu đồ của kim loại nặng trung bình hàng năm của sông SG, đo trong vòng 8 năm gần nhất. Nguồn: Sở TNMT SG . . . . .	76
3.3 Biểu đồ của BOD5 trung bình hàng tháng của sông SG, đo trong vòng 7 năm gần nhất. Nguồn: Sở TNMT SG . . . . .	77
3.4 Tần số và tần suất . . . . .	82
3.5 Tần suất tích lũy . . . . .	83
3.6 Tần suất của số nhược điểm trên các tấm gốm . . . . .	84
3.7 Mẫu của 100 giá trị logarithm độ bền sợi . . . . .	86
3.8 Tần số và tần số tích lũy của biến liên tục . . . . .	87
3.9 Mẫu của 100 giá trị logarithm độ bền sợi . . . . .	88
3.10 Phân phối tần số của biến logarithm độ bền sợi . . . . .	89
3.11 Tính phương sai mẫu . . . . .	93
5.1 Tần suất quan sát được và tần suất kỳ vọng của 100 trung bình mẫu . . . . .	162
5.2 Tần suất quan sát được và tần suất kỳ vọng của 100 chu kỳ . . . . .	163
6.1 Số lượng rầy và 11 yếu tố ảnh hưởng quan trắc được ở MDR (Nguồn: TS Đỗ Thanh Nghị, ĐH Cần Thơ) . . . . .	171
6.2 Nồng độ chất ô nhiễm không khí . . . . .	175
6.3 Dữ liệu chất xúc tác $X$ và đáp ứng $Y$ . . . . .	175

6.4	Các giá trị ISC của tế bào năng lượng ở ba điểm thời gian . . . . .	182
6.5	Giá trị của ISC và giá trị dự đoán tại $t_2$ . . . . .	184
6.6	Số liệu của DO và BOD trung bình hàng tháng của sông SG . . . . .	185
6.7	Các khoảng dự báo cho các giá trị ISC tại thời điểm $t_2$ . . . . .	186
6.8	Hiệp phương sai mẫu của các biến trong tập dữ liệu về các chốt tán nhôm . . . . .	189
6.9	Hệ số tương quan mẫu của các biến liên quan đến các chốt tán nhôm . . . . .	189
6.10	Hệ số tương quan mẫu của các biến đo trên các chốt tán nhôm, sau khi loại ra kết quả quan sát bất thường số 66 . . . . .	189
7.1	Bảng ANOVA đơn giản . . . . .	218
7.2	Bảng ANOVA đầy đủ theo một yếu tố . . . . .	219
7.3	Cách xử lý phân bón ảnh hưởng lên chiều cao cây cafe . . . . .	221
7.4	Bảng ANOVA một yếu tố phân bón . . . . .	222
7.5	Bảng tiếp liên của kết quả lắp đặt thêm vào theo từng thành phần . . . . .	223
7.6	Bảng tiếp liên cho lỗi lắp đặt . . . . .	224
7.7	Tỉ lệ thất bại (FPM) theo cấu trúc và theo thành phần . . . . .	225
7.8	Bảng tiếp liên cho số xy-lanh và nơi sản xuất . . . . .	226
7.9	Bảng tiếp liên của đường kính quay vòng và mức tiêu thụ nhiên liệu . . . . .	226
7.10	Tần số kì vọng và tần số quan sát được của đường kính quay vòng theo mức tiêu thụ nhiên liệu . . . . .	228
7.11	Bảng tiếp liên $2 \times 2$ cho phản hồi của khách hàng đối với Q3 và Q13 . . . . .	230
7.12	Chỉ số Cramer của Q1 – Q10 theo Q11 – Q13 . . . . .	231
7.13	Dữ liệu kiểm tra ống nghiệm tại các thời khắc khác nhau . . . . .	236
7.14	Bảng ANOVA mô hình lắng đọng chất rắn có 6 tham số . . . . .	238
8.1	Thiết kế CRD, nguồn [16] . . . . .	252
8.2	Thiết kế CRD - kết quả thực nghiệm, nguồn [16] . . . . .	252
8.3	Bảng ANOVA một yếu tố độ cứng gỗ (so sánh với Vấn đề 7). . . . .	254
8.4	Cách gán dấu và giá trị $\overline{D}$ . . . . .	258
8.5	Phân tích ANOVA cho RCBD . . . . .	259
8.6	Phân tích ANOVA cho dữ liệu lai ghép . . . . .	262
9.1	Phân tích ANOVA cho một BIBD . . . . .	272
9.2	Tập các chỉ số của các khối . . . . .	273
9.3	Giá trị của $Y_{il}$ , $i \in B_i$ . . . . .	273
9.4	Tập $T_j$ và các thống kê $W_j, W_j^*$ . . . . .	273
9.5	Phân tích ANOVA cho BIBD . . . . .	274
9.6	Các ảnh hưởng trung bình và lỗi chuẩn . . . . .	274

9.7	Phân tích ANOVA cho thiết kế gồm hai yếu tố . . . . .	277
9.8	Phân tích ANOVA hai chiều của dữ liệu chu kỳ piston . . . . .	280
9.9	Các mức nhân tố của thiết kế $2^3$ . . . . .	286
9.10	Kết quả thực nghiệm mô phỏng số nhờ thiết kế đầy đủ $2^3$ . . . . .	286
9.11	Eight factors, the number of levels and the level meanings . . . . .	291
9.12	Các nhân tố và mức chọn trong thí nghiệm mô phỏng piston . . . . .	298
9.13	Các nhân tố và mức chọn trong thí nghiệm mô phỏng piston . . . . .	299
9.14	Phân tích hồi quy của thí nghiệm piston với bốn yếu tố . . . . .	300
9.15	Bảng 2.2: Phân mảnh các nông trại . . . . .	305
9.16	Mảng trực giao trọng 2, ghi OA(12; $2^{11}$ ; 2) từ thiết kế nhân tố (đầy đủ) $2^{11}$ . . . . .	308
10.1	Bốn mẫu ngẫu nhiên có hoàn lại kích cỡ 20, chọn từ $\{1, 2, \dots, 100\}$ . . . . .	318
10.2	Các thống kê của phân phối mẫu . . . . .	321
11.1	Phép biến đổi dữ liệu cho phương sai đồng nhất . . . . .	335
12.1	Khoảng tin cậy cho $C_{pk}$ , ở mức độ tin cậy $1 - \alpha$ . . . . .	366

## Chương 7

## Phân tích Phương sai

*Kiểm tra sự khác biệt giữa nhiều trung bình*

[Nguồn [3]]

Hãy tưởng tượng các phổ màu trong bức ảnh trên thể hiện mật độ cây rừng ở một vùng bảo tồn sinh thái quốc gia (như Cúc Phương hay Nam Cát Tiên), khi ta hình dung mảng màu tối hơn thể hiện rằng rừng còn ít hơn, liệu ta có thể kết luận được rằng mật độ rừng là như nhau trên toàn miền?

Nội dung chính của chương này là nhằm kiểm tra sự khác biệt giữa nhiều trung bình tổng thể bằng nhóm phương pháp thống kê gọi là Phân tích Phương sai (*Analysis of Variance*, ANOVA). **Phân tích phương sai** là một tập các *mô hình thống kê* được sử dụng để phân tích sự khác biệt giữa các *trung bình nhóm* và các thủ tục liên quan (chẳng hạn như “sự biến đổi” bên trong mỗi nhóm và giữa các nhóm), được phát triển bởi Sir R.A. Fisher.

Mục 7.6 tổng kết các kỹ thuật trong hai chương 6 và 7 nhằm xây dựng mô hình hồi quy thực nghiệm, một vấn đề rất có ý nghĩa trong các lĩnh vực thực tiễn.

## 7.1 Phân phối $F$ cho biến tỉ lệ hai phương sai

Ta xem hai mẫu độc lập kích thước  $n_1$  và  $n_2$ , đã được lấy từ các quần thể Gauss có phương sai tương ứng là  $\sigma_1^2$  và  $\sigma_2^2$ . Gọi

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2, \text{ và } S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2$$

là phương sai của hai mẫu, có  $\bar{X}_1$  và  $\bar{X}_2$  là trung bình mẫu tương ứng.

Tỉ lệ

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}$$

có phân phối biểu thị bởi  $F[\nu_1, \nu_2]$ , với  $\nu_1 = n_1 - 1$  và  $\nu_2 = n_2 - 1$ . Phân phối này được gọi là phân phối  $F$  (Fisher) với  $\nu_1$  và  $\nu_2$  bậc tự do. Đồ thị của hàm mật độ của  $F[\nu_1, \nu_2]$  được cho trong Hình 7.1.

Giá trị kỳ vọng và phương sai của  $F[\nu_1, \nu_2]$  là

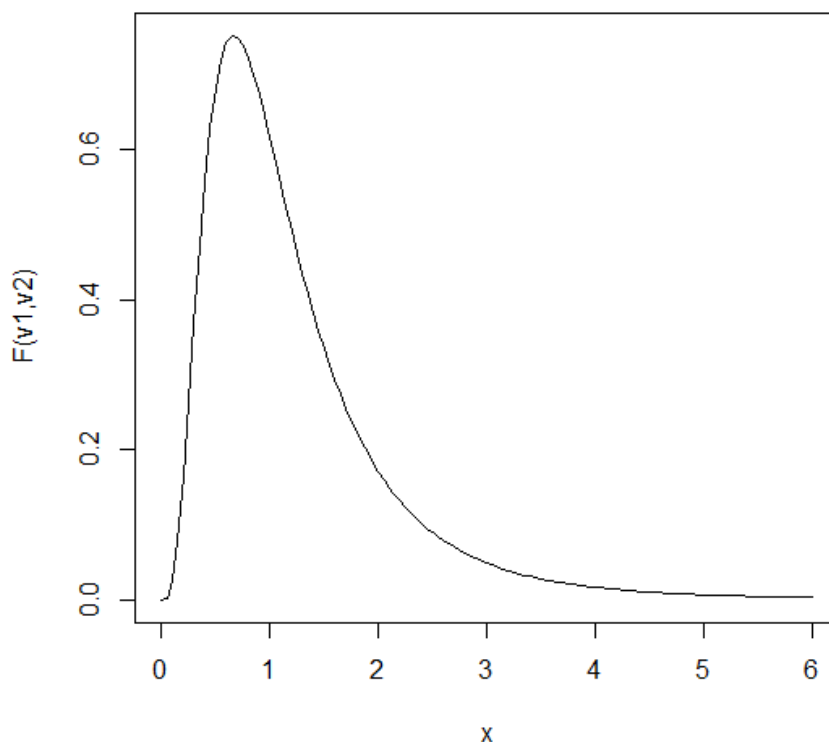
$$\begin{aligned} \mathbf{E}(F[\nu_1, \nu_2]) &= \nu_2/(\nu_2 - 2); & \text{khi } \nu_2 > 2 \\ \mathbf{V}(F[\nu_1, \nu_2]) &= \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}; & \text{khi } \nu_2 > 4. \end{aligned} \quad (7.1)$$

**Phân vị thứ  $1 - \alpha$  của  $F[\nu_1, \nu_2]$ ,  $F_{1-\alpha}[\nu_1, \nu_2]$  tính được bằng R.** Nếu muốn có **phân vị** (fractile) thứ  $\alpha$ ,  $F_\alpha[\nu_1, \nu_2]$  cho các giá trị  $\alpha < 0.5$ , ta có thể áp dụng mối quan hệ

$$F_{1-\alpha}[\nu_1, \nu_2] = \frac{1}{F_\alpha[\nu_2, \nu_1]}. \quad (7.2)$$

Thí dụ ta tính

$$F_{0.05}[15, 10] = 1/F_{0.95}[10, 15] = 1/2.54 = .3937.$$

Hình 7.1: Hàm mật độ của  $F(\nu_1, \nu_2)$ 

## 7.2 Phân tích phương sai (ANOVA) đơn giản

### 7.2.1 Kiểm định so sánh các trung bình

Khi các biến hồi quy  $x_1, x_2, \dots, x_k$  là các biến *định tính* (hay *phân loại, categorical*), và biến đáp ứng được quan tâm  $Y$  là biến *định lượng* (*quantitative*) thì các phương pháp hồi quy đa biến đã được thảo luận trong các phần trước **không thể sử dụng được nữa**. Các giá trị khác nhau thu được của các biến hồi quy thuộc vào các loại khác nhau.

Chẳng hạn, khi ta cần nghiên cứu đến quan hệ giữa

tốc độ phim ( $Y$ ) và

loại bột gelatine  $X$

được sử dụng trong quá trình chuẩn bị chất nhũ tương hóa để tráng phim, thì biến hồi quy  $X$  là *biến phân loại*, các giá trị của biến này là các loại gelatine khác nhau, được phân loại theo nhà sản xuất; trong khi  $Y$  là *định lượng*. Trong Ví dụ 26, tương tự ta sẽ thấy biến nhà cung cấp hộp chứa đĩa cứng là biến định tính còn số lần đóng mở  $L$  là định lượng.

- Khi ta có  $k$ ,  $k \geq 1$  biến định tính như thế, các tổ hợp của các mức phân loại khác nhau của  $k$  biến

được gọi là các **tổ hợp xử lý** hay **cách xử lý**, **phương án** hay **giải pháp** (*treatment combinations, treatments*)- một thuật ngữ được sử dụng phổ biến bởi các nhà khoa học thực nghiệm.

- Ở mỗi tổ hợp xử lý thứ  $i$ , ta có thể thực hiện nhiều ( $n_i$ ) quan sát. Kết quả của những quan sát này được coi là một mẫu ngẫu nhiên được chọn ra từ một tổng thể (vô hạn) gồm tất cả các kết quả quan sát có thể thực hiện được dưới tổ hợp xử lý cụ thể này.

Khi đó, mô hình thống kê của kết quả quan sát lần thứ  $j$  sẽ là

$$Y_{ij} = \mu_i + e_{ij}; \quad i = 1, \dots, k; \quad j = 1, \dots, n_i,$$

trong đó

- $\mu_i$  là trung bình của tổng thể tương ứng với tổ hợp xử lý thứ  $i$ ,  $k$  là số tổ hợp xử lý;
- $e_{ij}$ , ( $i = 1, \dots, k; j = 1, \dots, n_i$ ) được xem như là các biến ngẫu nhiên độc lập (các *sai số thực nghiệm*) với  $E[e_{ij}] = 0$ , và  $V[e_{ij}] = \sigma^2$ ,  $\forall (i, j)$ ; và sau chót
- $Y$  phải theo luật phân phối chuẩn.

Việc kiểm định so sánh các trung bình  $\mu_i$ , ( $i = 1, \dots, k$ ) sẽ cung cấp cho ta thông tin về những ảnh hưởng của các tổ hợp xử lý khác nhau. Phương pháp thực hiện phân tích này được gọi là phương pháp **Phân tích Phương sai (Analysis of Variance - ANOVA)**.

### 7.2.2 Phương pháp phân tích phương sai đơn giản

Trong mục này ta sẽ xây dựng cơ sở toán học cho ANOVA. Ở đây ta giả sử rằng các độ lệch  $e_{ij}$  đều có phân phối chuẩn, độc lập với nhau. Với tổ hợp xử lý (mẫu) thứ  $i$ , ta đặt

$$\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}, \quad i = 1, \dots, k \quad (7.3)$$

và

$$SSD_i = \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2, \quad i = 1, \dots, k. \quad (7.4)$$

Đặt  $\bar{\bar{Y}} = \frac{1}{N} \sum_{i=1}^k n_i \bar{Y}_i$  là trung bình tổng thể (của toàn bộ các đáp ứng quan sát).

#### Các đặc trưng

Phương pháp phân tích phương sai đơn giản ANOVA dựa trên sự phân tích tổng của các bình phương của các độ lệch so với  $\bar{\bar{Y}}$ ,

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{\bar{Y}})^2 = \sum_{i=1}^k SSD_i + \sum_{i=1}^k n_i (\bar{Y}_i - \bar{\bar{Y}})^2. \quad (7.5)$$

Ta kí hiệu vế trái là  $SST$  và các số hạng ở vế phải là  $SSW$  và  $SSB$ , tức là

$$SST = SSW + SSB \quad (7.6)$$

\* Khi đó,  $SST$ ,  $SSW$  và  $SSB$  là các dạng toàn phương đối xứng theo lần lượt các độ lệch

$$Y_{ij} - \bar{Y}, \quad Y_{ij} - \bar{Y}_i, \quad \text{và} \quad \bar{Y}_i - \bar{Y}.$$

\* Do bởi  $\sum_i \sum_j (Y_{ij} - \bar{Y}) = 0$  nên chỉ còn lại  $N - 1$  hàm tuyến tính

$$Y_{ij} - \bar{Y} = \sum_{i'} \sum_{j'} c_{i'j'} Y_{i'j'},$$

với

$$c_{i'j'} = \begin{cases} 1 - \frac{1}{N}, & i' = i, j' = j \\ -\frac{1}{N}, & \text{ngược lại} \end{cases}$$

là độc lập tuyến tính với nhau, với  $N = \sum_{i=1}^k n_i$ .

Vì lí do này ta nói

- dạng toàn phương  $SST$  có  $(N - 1)$  bậc tự do (d.f.).
- Tương tự  $SSW$  có  $(N - k)$  bậc tự do, bởi vì  $SSW = \sum_{i=1}^k SSD_i$  và do số bậc tự do của mỗi  $SSD_i$  là  $(n_i - 1)$ . Cuối cùng  $SSB$  có  $(k - 1)$  bậc tự do.
- Để ý rằng  $SSW$  là tổng toàn bộ của bình phương của các độ lệch **chỉ trong**  $k$  mẫu, và  $SSB$  là tổng bình phương độ lệch **giữa**  $k$  trung bình mẫu.

Chia một dạng toàn phương cho số bậc tự do của nó ta thu được thống kê trung bình bình phương (tức là phương sai, *mean-squared*). Ta tóm lược các thống kê này trong một bảng, gọi là **bảng ANOVA** (**ANOVA table**). Bảng 7.1 là bảng ANOVA cho kiểm định so sánh các trung bình của  $k$  tổ hợp xử lý.

**Bảng 7.1:** Bảng ANOVA đơn giản

Sự biến thiên	D.F.	S.S.	M.S.
Giữa các tổ hợp xử lý	$k - 1$	$SSB$	$MSB = SSB/(k - 1)$
Trong các tổ hợp xử lý	$N - k$	$SSW$	$MSW = SSW/(N - k)$
Toàn bộ (đã được hiệu chỉnh cho trung bình)	$N - 1$	$SST$	-

Tóm lại, trong bảng ANOVA, D.F. kí hiệu cho số bậc tự do, S.S. kí hiệu cho tổng bình phương của các độ lệch, và M.S. kí hiệu cho phương sai. Ta có

$$M.S. = \frac{S.S.}{D.F.} \quad (7.7)$$



Bây giờ ta chứng tỏ rằng

$$\mathbf{E}[MSW] = \sigma^2. \quad (7.8)$$

Thật vậy, theo mô hình trên, cộng với lý do rằng  $\{Y_{ij}, j = 1, \dots, n_i\}$  là một mẫu ngẫu nhiên có hoàn lại (RSWR) chọn ra từ tổng thể tương ứng với tổ hợp xử lý thứ  $i$ , nên ta có

$$\mathbf{E}\left[\frac{SSD_i}{n_i - 1}\right] = \sigma^2, \quad i = 1, \dots, k.$$

Vì  $MSW = \sum_{i=1}^k \nu_i \left(\frac{SSD_i}{n_i - 1}\right)$ , với  $\nu_i = \frac{n_i - 1}{N - k}$ ,  $i = 1, \dots, k$  nên ta có

$$\mathbf{E}[MSW] = \sum_{i=1}^k \nu_i \mathbf{E}\left[\frac{SSD_i}{n_i - 1}\right] = \sigma^2 \sum_{i=1}^k \nu_i = \sigma^2. \blacksquare$$

Một kết quả quan trọng nữa là

$$\mathbf{E}[MSB] = \sigma^2 \frac{1}{k - 1} \sum_{i=1}^k n_i \tau_i^2, \quad (7.9)$$

với  $\tau_i = \mu_i - \bar{\mu}$  ( $i = 1, \dots, k$ ) và  $\bar{\mu} = \frac{1}{N} \sum_{i=1}^k n_i \mu_i$ .

Vậy với giả thuyết ban đầu là  $H_0: \mu_1 = \dots = \mu_k$ , ta có  $\mathbf{E}[MSB] = \sigma^2$ .

Điều này gợi ý cho ta sử dụng thống kê  $F$  (xem Mục 7.1 về phân phối Fisher) để kiểm định  $H_0$

$$F = \frac{MSB}{MSW}. \quad (7.10)$$

Ta bác bỏ giả thuyết  $H_0$  với mức ý nghĩa  $\alpha$  nếu như

$$F > F_{1-\alpha}[k - 1, N - k].$$

**Bảng 7.2:** Bảng ANOVA đầy đủ theo một yếu tố

Sự biến thiên	D.F.	S.S.	M.S.	Thống kê $F$
Giữa các tổ hợp xử lý	$k - 1$	$SSB$	$MSB = SSB/(k - 1)$	$\frac{MSB}{MSW}$
Trong các tổ hợp xử lý	$N - k = k(n - 1)$	$SSW$	$MSW = SSW/(N - k)$ (hay $SSW/k(n - 1)$ )	
Toàn bộ	$N - 1$	$SST$	-	

### 7.2.3 Về các giả định cần thiết cho Phân tích phương sai

Tại thời điểm này, điều quan trọng là tóm tắt các giả định của phân tích phương sai. Như phân tích hồi quy trong Chương 6, một số giả định phải được đáp ứng để có thể kiểm tra sự bằng nhau của  $k$  trung bình tổng thể, bao gồm:

1. Mỗi tổng thể được phân phối chuẩn.
2. Các tổng thể có cùng phương sai  $\sigma^2$ .
3. Các quan sát trong một tổng thể là độc lập với các quan sát trong các tổng thể khác.
4. Các quan sát được lựa chọn ngẫu nhiên từ mỗi tổng thể.

Nếu các giả định được thỏa, phương sai trong nhóm ( $MSW$ ) là ước lượng không chệch của  $\sigma^2$ , xem Hệ thức 7.8. Do đó, sai số chuẩn của ước lượng cho bất kỳ  $k$  trung bình tổng thể là:

$$\text{S.E.}(\bar{y}_{i.}) = S_{\bar{y}_{i.}} = \sqrt{\frac{MSW}{n_i}}. \quad (7.11)$$

Khi có số lượng các quan sát bằng nhau cho mỗi nhóm,  $n_i = n, \forall i = 1, \dots, k$ , điều này đơn giản là

$$\text{S.E.}(\bar{y}_{i.}) = S_{\bar{y}_{i.}} = \sqrt{\frac{MSW}{n}}. \quad (7.12)$$

Vậy khoảng tin cậy cho trung bình  $\mu_i$  của tổng thể thứ  $i$  (bất kỳ trong  $k$  tổng thể) được cho bởi:

$$\mathbb{P}[\bar{y}_{i.} - t_{\alpha/2}[N - k] S_{\bar{y}_{i.}} < \mu_i < \bar{y}_{i.} + t_{\alpha/2}[N - k] S_{\bar{y}_{i.}}] = 1 - \alpha. \quad (7.13)$$

Bậc tự do của thống kê  $t$  trong khoảng tin cậy trên bằng với bậc tự do của trung bình bình phương trong nhóm  $MSW$ , nên khi các cỡ mẫu của các quan sát bằng nhau,  $n_i = n$  thì ta có bậc tự do là  $k(n - 1) = N - k$ .

- Khi phân tích phương sai là cân bằng, chiều rộng của các khoảng tin cậy của tất cả các trung bình  $\mu_i$  là như nhau.
- Khi phân tích phương sai là không cân bằng, khoảng tin cậy cho tất cả  $\mu_i$  được xây dựng với cùng giá trị thống kê  $t$ , nhưng chiều rộng của những khoảng này khác nhau dựa trên  $n_i$ .

### VẤN ĐỀ 7. [Ứng dụng trong Nông nghiệp.]

**Nghiên cứu ảnh hưởng khác biệt của loại phân bón lên cây trồng.**

Ba cách xử lý phân bón, bao gồm hai loại phân (hữu cơ và vô cơ) và một điều khiển (tức là không có phân bón), đã được thử nghiệm để điều tra tác động của chúng đến sự tăng trưởng chiều cao (tính bằng cm) của cây cafe 1 tuổi. Các dữ liệu ban đầu, các tổng số và trung bình được tóm tắt trong Bảng 7.3 sau.

Nhiệm vụ của chúng ta là so sánh sự bằng nhau của  $k = 3$  trung bình tổng thể chưa biết (chiều cao cây cafe  $Y$ ) sử dụng phân tích phương sai một nhân tố (phân bón) với mức ý nghĩa 0.05.

**Bảng 7.3:** Cách xử lý phân bón ảnh hưởng lên chiều cao cây cafe

	Cách xử lý phân bón			
	Điều khiển	Hữu cơ	Vô cơ	
	3.0	4.0	4.1	
	3.3	4.3	4.0	
	3.5	4.0	4.2	
	3.0	4.1	3.7	
Tổng	12.8	16.4	16.0	45.2
Trung bình	3.2	4.1	4.0	3.8

Bậc tự do cho ba tổng các bình phương ( $SST$ ,  $SSB$  và  $SSW$ ) là:

$$N = nk = k - 1 + k(n - 1) = 4 \cdot 3 = 12.$$

Ta xem

- tổng bình phương **giữa các nhóm** (nhóm - nhóm)  $SSB$  là tổng bình phương cách xử lý (*treatment*, còn ghi  $SS_{Treatments}$ ), và xem
- tổng bình phương **trong nhóm**  $SSW$  là tổng bình phương của lỗi thực nghiệm (hoặc đơn giản, tổng lỗi thực nghiệm  $SS_E$ ), bởi vì
- phân tích phương sai được sử dụng chủ yếu để phân tích dữ liệu được tạo ra bởi các thiết kế thí nghiệm, nơi mà các nhóm thường chỉ các cách xử lý và sự biến động trong nhóm là *lỗi thử nghiệm không giải thích được* (ngẫu nhiên hoàn toàn).

Xem Ví dụ 12.1 của [2]. Bảng phân tích phương sai là như sau:

Bảng 7.4: Bảng ANOVA một yếu tố phân bón

Sự biến thiên	D.F.	S.S.	M.S.	Thống kê $F$
Giữa các tổ hợp xử lý	$k - 1 = 2$	$SSB = 1.947$	$MSB = SSB/(k - 1) = 0.9735$	22.91
Trong các tổ hợp xử lý	$N - k = 9$	$SSW = 0.383$	$MSW = SSW/(N - k) = 0.0425$	
Toàn bộ	$N - 1 = 11$	$SST = 2.330$	-	

Giờ ta tính khoảng tin cậy 95% cho giá trị trung bình chưa biết của biến  $Y_3$ - chiều cao cây cafe khi ta xử lý bằng phân bón vô cơ. Ta có

$$\begin{aligned}\bar{y}_{3.} &= 4.0, & t_{0.025}[9] &= 2.26, & S_{\bar{y}_{3.}} &= \sqrt{\frac{MSW}{n}} = \sqrt{\frac{0.0425}{4}} = 0.1031 \\ \Rightarrow \mathbb{P}[4.0 - (2.26)(0.1031) < \mu_3 < 4.0 + (2.26)(0.1031)] &= 0.95 \\ \Rightarrow \mathbb{P}[3.37 < \mu_3 < 4.23] &= 0.95.\end{aligned}$$

Vậy trung bình tổng thể  $\mu_3 \in (3.77, 4.23)$  - mức phát triển chiều cao của cây cafe con được xử lý bằng phân bón vô cơ ở giữa 3.77 và 4.23 cm; với xác suất tin cậy 0.95.

Thống kê  $F = \frac{MSB}{MSW} = 22.91$  lớn hơn trị phân vị  $F_{1-\alpha; k-1, N-k} = F_{0.95; 2, 9} = 4.26$  ta bác bỏ  $H_0$ , chấp nhận  $H_1$ . Ta kết luận từ dữ liệu thực nghiệm trên rằng 3 trung bình tổng thể là khác nhau đáng kể (hay ít nhất hai trong số các trung bình  $\mu_1, \mu_2, \mu_3$  là khác nhau đáng kể).

## 7.3 Bảng tiếp liên

### 7.3.1 Cấu trúc của các bảng tiếp liên

Khi dữ liệu được phân loại ta thường trình bày tóm lược trong một bảng, trong đó thể hiện tần số của mỗi loại theo từng biến. Một bảng như thế được gọi là **bảng tiếp liên** (*contingency table*).

**Ví dụ 7.1.** Xét việc kiểm định một loại máy khi được lắp thêm vào một số linh kiện hay bộ phận khác. Lỗi đặt sai linh kiện của loại máy này đã được xem xét trong Ví dụ 6.9. Trong kiểm định này ta lắp thêm vào một lượng lớn ( $k = 9$ ) linh kiện khác nhau. Kết quả của mỗi phép thử (một lần lắp thêm) sẽ hoặc là “Thành công” (không bị lỗi), hoặc là “Thất bại” (bị lỗi). Vậy trong kiểm định này ta có hai biến phân loại: Biến “Linh kiện” (được lắp thêm vào) và biến “Kết quả” (của mỗi lần lắp thêm vào). Biến đầu tiên có chín loại sau đây:

## Phần IV: Thiết kế và Phân tích Thí nghiệm

### Chương 8: Thiết kế Thí nghiệm Thống kê

*Tính toán phương cách thu thập dữ liệu*

### Chương 9: Thiết kế và Phân tích Thí nghiệm Nhân tố

*Làm sao tìm ra kết quả tối ưu khi có nhiều yếu tố ảnh hưởng?*

Khi tiến hành một cuộc khảo sát hay thí nghiệm, ta có thể nghiên cứu cùng lúc nhiều yếu tố xem ảnh hưởng đồng thời của chúng lên kết quả. Trong trường hợp này, ta phải hạn chế sự khảo sát thực nghiệm lên một số hữu hạn khả năng của các yếu tố ấy. Các thí nghiệm như vậy được thực thi và nghiên cứu hoàn toàn ngược lại với cách hiểu thông thường khi ta thực thi hay khảo sát một yếu tố đơn lẻ tại một thời điểm, trong lúc giữ tất cả các yếu tố khác không thay đổi.

Chương 8 sẽ giới thiệu phương cách thu thập dữ liệu dựa trên các khung thức tổng quát mà có thể sử dụng nhiều lần trong các tình huống khác nhau. Các khung thức tổng quát được gọi chung là **Thí Nghiệm được Thiết Kế một cách Thống Kê** (*Statistically Designed Experiments*, hay gọn hơn là **Thiết Kế Thống Kê**, *Design of Experiments*, DOE).

Các thiết kế thống kê đã được sử dụng để gia tăng tốc độ học tập và làm giàu tri thức nhờ sự khai phá tiên phong của R.A. Fisher trong nửa đầu của thế kỷ XX. Một trong những đóng góp chính của Fisher là sự phát triển của các **thí nghiệm, thiết kế nhân tố** (*factorial experiments*), mà nhờ đó ta có thể nghiên cứu đồng thời nhiều nhân tố một lúc. Chương 9 trình bày phương pháp thiết kế nhân tố nhằm tìm ra kết quả tối ưu khi có nhiều yếu tố ảnh hưởng đồng thời lên kết quả.



Hình 7.4: GS R.A. Fisher, nhà phát minh Thiết Kế Thí Nghiệm (DOE).

Nhưng muốn xây dựng, tôn tạo một vườn hoa muôn sắc với cây xanh, hòn non bộ ... trong hình ở trang bên liệu bạn có cần các *thiết kế thống kê* hay không?

## Thiết kế Thí nghiệm Thống kê

*Tính toán khung thức thu thập dữ liệu*

---



## 8.1 Giới thiệu về lịch sử và nội dung

### 8.1.1 Thế nào là thiết kế thí nghiệm?

Quan sát một hiện tượng hoặc quá trình là kỹ năng không thể thiếu trong quá trình học hỏi của loài người. Galileo nhờ sử dụng chiếc kính viễn vọng mà khám phá ra thuyết nhật tâm, hay Newton nhờ quan sát một quả táo rơi mà nảy ra ý tưởng về lực hấp dẫn. Tuy nhiên, để hiểu được mối quan hệ nguyên nhân - kết quả giữa hiện tượng và các yếu tố tác động, chúng ta phải làm nhiều hơn là chỉ đứng quan sát - chúng ta phải thực hiện thay đổi các yếu tố tác động. Hay nói cách khác, chúng ta phải làm **thí nghiệm**. Quan sát một hệ thống hoặc tiến trình giúp chúng ta đi đến giả thuyết, còn để kiểm chứng sự đúng đắn của giả thuyết, chúng ta phải cần tới thí nghiệm.

Một ví dụ đơn giản về thí nghiệm, một kỹ sư vật liệu muốn tìm hiểu xem trong hai phương pháp làm nguội hợp kim nhôm, sử dụng dầu hoặc nước muối, phương pháp nào làm hợp kim nhôm đạt được độ cứng tốt nhất. Với mục tiêu như vậy, viên kỹ sư sẽ dùng nhiều mẫu nhôm khác nhau, thực hiện làm nguội từng mẫu nhôm bằng một trong hai phương pháp trên, sau đó dùng thiết bị đo độ cứng trên sản phẩm thí nghiệm. Trung bình độ cứng của các mẫu nhôm sẽ dùng để quyết định phương pháp nào là tốt hơn. Để thấy sẽ có những câu hỏi dành cho viên kỹ sư khi thực hiện thí nghiệm này:

1. Chúng ta có thực sự chỉ quan tâm đến hai phương pháp làm nguội này?
2. Còn yếu tố nào ảnh hưởng đến độ cứng của nhôm mà ta cần xem xét và kiểm soát trong thí nghiệm này (nhiệt độ của chất làm nguội chẳng hạn)?
3. Nên dùng bao nhiêu mẫu nhôm cho từng phương pháp?
4. Cách chọn mẫu nhôm cho từng phương pháp như thế nào? Việc đo độ cứng nên thực hiện theo thứ tự nào? Phương pháp phân tích dữ liệu nào nên được dùng sau khi đo đạc?
5. Trung bình độ cứng sau khi áp dụng từng phương pháp khác nhau đến mức nào thì được xem là khác nhau đáng kể?

Các câu hỏi này cần phải được trả lời trước khi tiến hành thí nghiệm, và **Thiết kế thí nghiệm thống kê** (*Design of Experiments – DOE*) có thể trả lời một cách thỏa đáng tất cả các câu hỏi trên. Hay nói cách khác, Thiết kế thí nghiệm thống kê là *quá trình lên kế hoạch thí nghiệm* sao cho có thể thu thập được dữ liệu phù hợp, để sau khi được phân tích bằng các phương pháp thống kê, có thể cho ra những kết luận đúng đắn và phù hợp với mục tiêu đã đề ra. Như vậy, một thí nghiệm bao giờ cũng gồm hai phần có liên hệ chặt chẽ với nhau: thiết kế thí nghiệm và phân tích dữ liệu. Cách thiết kế một thí nghiệm sẽ quyết định phương pháp thống kê nào cần được sử dụng để sau khi phân tích dữ liệu sẽ cho ra kết quả hợp lý.

### 8.1.2 Lịch sử và nội dung

Thiết Kế Thí Nghiệm Thống Kê (*DOE Design of Experiments or Experimental Designs*) ra đời từ đầu thế kỷ 20 khi Sir Ronald Aylmer Fisher (1890-1962), được coi là một trong những bậc tiền bối của Thống Kê Học hiện đại, khảo sát hiệu quả của các loại phân bón trên các mảnh đất khác nhau nhờ sử dụng khái niệm toán học gọi là *Latin squares*, xem chi tiết ở mục 9.7.1.



Làm việc ở Viện Nghiên cứu Nông nghiệp Rothamsted, Anh Quốc từ năm 1919, Fisher quan sát thấy kết quả vụ mùa trên các mảnh đất khác nhau không chỉ phụ thuộc vào phân bón mà còn vào vài yếu tố khác, như là độ màu mỡ và độ ẩm của đất, hay giống cây trồng nữa, nhưng không rõ là phụ thuộc ra sao. Fisher, sau đó làm việc ở UC London và Cambridge, nhà khoa học tiên phong đặt nền móng cho suy diễn xác suất (*likelihood inference*), phân tích phương sai (ANOVA), và đặc biệt là nền tảng lý thuyết cho DOE, đã sử dụng DOE nhằm tách bạch, phân biệt ảnh hưởng của phân bón và ảnh hưởng của các yếu tố khác trên kết quả vụ mùa, xem làm sao để có thu hoạch cao nhất.



GS George Edward Pelham Box

Trong những năm 50, nhà toán học R.C. Bose người gốc Ấn độ, đã mở rộng ý niệm Latin squares tới một lý thuyết hoàn chỉnh về các thiết kế tổ hợp (*combinatorial designs*), trình bày ở mục 9.7.2; và được phát triển mạnh mẽ trong những năm sau đó ở Hoa Kỳ, nhờ công các nhà toán học khác như C. R. Rao và S. S. Shrikhande. Ngày nay DOE được nghiên cứu sâu rộng ở các trường đại học và khu vực công nghiệp, và có ứng dụng thành công trong hầu hết các lĩnh vực kỹ thuật, sản xuất, y dược khoa, và công nghệ cao.

#### Một số hướng chính của DOE là

**Thiết kế nhân tố (*factorial designs*)** có biểu diễn toán học là ma trận, các cột là các yếu tố, mỗi yếu tố có các giá trị (*settings, levels*) khác nhau được sắp xếp theo trật tự xác định. Công cụ nghiên cứu chính là tổ hợp, đại số, lý thuyết số, và từ đầu thế kỷ 21 là hình học đại số (*algebraic geometry*) và đại số không giao hoán (trên đó hai phép toán đại số thông thường không đồng thời thỏa tính giao hoán).

**Thiết kế tối ưu (*optimal designs*)** cần kiến thức về đại số tuyến tính, lý thuyết ma trận và xác suất.

**Phương pháp luận mặt đáp ứng (*RSM- response surface methodology*)** sử dụng các loại thiết kế trên và lý thuyết xấp xỉ đa biến (xem từ Mục 9.5).



Toán học cho DOE rất phong phú và sâu sắc! Ta sẽ khảo sát sâu hơn sự liên quan này cùng với các hướng nghiên cứu và ứng dụng quan trọng của DOE trong các phần sau.

Ngoài Fisher, Bose và Rao, các khoa học gia tiên phong trong Thiết Kế Thí Nghiệm từ sau thế chiến II bao gồm George Box (sinh 1919), Douglas Montgomery, Genichi Taguchi (1924 – 2012), Giovanni Pistone, Henry Wynn, Andries Brouwer, Rosemary Bailey, John Borkowski, Vladimir Fedorov ... Các GS Henry Wynn (Anh quốc) cùng với Giovanni Pistone (Ý) đồng khai sinh ra lãnh vực Thống kê Đại số (*Algebraic Statistics*, từ năm 2000, nay ta gọi là trường phái Âu châu), dựa trên các ý tưởng nảy sinh khi tính toán thiết kế thí nghiệm.

## 8.2 Vai trò của Thiết kế Thí nghiệm Thống kê trong kỹ thuật

Sau khi nghiên cứu cẩn thận chương này, bạn sẽ có thể làm được các việc sau:

1. Thiết kế và thực hiện các thí nghiệm kỹ thuật liên quan đến một yếu tố duy nhất với một số tùy ý của các mức (cấp) chọn.
2. Hiểu cách phân tích phương sai được sử dụng để phân tích các dữ liệu thu lượm được từ những thí nghiệm.
3. Đánh giá tính đầy đủ của mô hình nhờ các sơ đồ thặng dư.
4. Sử dụng các thủ tục đa so sánh nhằm xác định sự khác biệt cụ thể giữa các trung bình.
5. Đưa ra quyết định về kích thước mẫu trong các thí nghiệm đơn yếu tố.
6. Hiểu được sự khác biệt giữa các *yếu tố cố định* và *yếu tố ngẫu nhiên*.
7. Ước tính các thành phần phương sai trong một cuộc thử nghiệm liên quan đến các yếu tố ngẫu nhiên.
8. Thiết kế và tiến hành các thí nghiệm liên quan đến thiết kế khối ngẫu nhiên hoàn chỉnh (RCBD).

Nói khái quát thì mỗi cuộc thí nghiệm liên quan đến một chuỗi các hoạt động lặp sau đây:

**Phỏng đoán** - giả thuyết (nghiên cứu) ban đầu rằng mà dẫn đến cuộc thí nghiệm.

**Tiến hành thử nghiệm** - các thí nghiệm cụ thể được thực hiện để điều tra - khẳng định - bác bỏ phỏng đoán.

**Phân tích** - phân tích thống kê các dữ liệu từ cuộc thử nghiệm.

**Kết luận** - những gì bạn đã học được về phỏng đoán ban đầu từ các thí nghiệm.

Thường thì các thử nghiệm cũng sẽ dẫn đến việc sửa đổi phỏng đoán, và một loạt thí nghiệm mới được thực hiện lại, và vv.

## 8.3 Các định nghĩa căn bản và Ứng dụng thực tiễn

Sau đây là những nguyên tắc chủ đạo mà chúng ta vận dụng trong việc thiết kế các thí nghiệm. Các nguyên tắc này được thiết kế để đảm bảo chất lượng thông tin (InfoQ) của việc nghiên cứu là cao, như đã giới thiệu ở Mục 12.3.

1. Các **mục tiêu** (*objectives*) của nghiên cứu cần phải được nêu ra tường minh, và các tiêu chí nên thiết lập để kiểm tra xem các mục tiêu đã đạt được chưa.
2. Các **biến đáp ứng** (*response variable*) cần được xác định rõ ràng để các mục tiêu nghiên cứu được diễn đạt đúng. Ở giai đoạn này cách đo lường, định lượng độ bất định nên được thiết lập (xem Chương 3).
3. Tất cả các nhân tố mà có thể ảnh hưởng đến các biến đáp ứng nên được liệt kê và xác định. Ta gọi đây là các **nhân tố (yếu tố) được kiểm soát** (*controllable factors*). Điều này đòi hỏi ta phải động não và có thể phải thảo luận chi tiết với các chuyên gia trong ngành.
4. Các loại **phép đo** (*measurements*) hoặc quan sát trên tất cả các biến cần được đặc tả.
5. Các **cấp độ** (mức chọn, *levels of controllable factors*) của các yếu tố được kiểm soát, mà cần được thử nghiệm, nên được xác định rõ.
6. Một **mô hình thống kê** cần được xây dựng liên quan đến mối quan hệ giữa các biến liên quan và các phân phối lỗi của chúng. Điều này có thể dựa trên kiến thức trước đây hoặc tìm kiếm thêm tài liệu tham khảo.
7. Một **cách bố trí thí nghiệm** (*experimental layout*) hoặc **mảng (bảng) thử nghiệm** (*experimental array*) nên được thiết kế sao cho việc suy luận từ các dữ liệu thu thập được sẽ là
  - (a) hợp lệ (đúng- *accurate* hay có ý nghĩa);
  - (b) chính xác (*precise*);
  - (c) có thể tổng quát hóa được (*generalizable*);
  - (d) dễ dàng tiến hành được.
8. Các **thử nghiệm được thực hiện**, nếu có thể, theo một **thứ tự ngẫu nhiên**, để tránh thiên vị bởi các yếu tố mà không được đưa vào xem xét.
9. Một **giao thức thực hiện** (*protocol of execution*) cần được chuẩn bị, kèm theo phương pháp phân tích. Lưu ý rằng, phương pháp phân tích và sự thu thập dữ liệu [phụ thuộc vào thiết kế](#).
10. **Việc thực thi thí nghiệm** nên tuân theo một cách cẩn thận giao thức nói trên với ghi chú thích hợp.
11. Các **kết quả của các thí nghiệm** cần được phân tích và báo cáo cẩn thận, đảm bảo sự chú thích thỏa đáng và cho phép truy xuất được nguồn gốc. Công nghệ hiện đại (như hàm `sweave` trong R) đảm bảo rằng dữ liệu, sự phân tích và các kết luận được tích hợp đầy đủ và tái sản sinh được.
12. **Thí nghiệm xác nhận** (hay khẳng định, *confirmatory*) cần được tiến hành để xác nhận các suy luận hay kết luận của các thí nghiệm chính.

Ta minh họa các nguyên tắc trên với ví dụ sau.

**Ví dụ 8.1. [Ứng dụng trong Công nghiệp - Sản xuất.]** (Xem thêm tài liệu [27, Chương 11])

Ta minh họa DOE qua một quá trình phức tạp trong sản xuất, với một số lượng lớn các yếu tố có thể ảnh hưởng đến các biến đầu ra. Hàn sóng của dây chuyền đóng gói vi mạch (*circuit pack assemblies*, CPA) là một quá trình hàn tự động mà, nếu được thực hiện trong một ngữ cảnh tối ưu, có thể nâng cao chất lượng và năng suất. Tuy nhiên, quá trình này bao gồm ba giai đoạn và liên quan tới nhiều biến. Ta phân tích các bước khác nhau mà cần thiết cho việc thiết kế thí nghiệm để tìm hiểu tác động của các yếu tố khác nhau lên trên quá trình. Ta sử dụng sự mô tả quá trình do Lin và Kacker (1989) đề xuất.

Nếu quá trình hàn mang lại kết quả tốt, ta có thể tiến hành ngay việc kiểm tra tự động các CPA. Đây là một khoản tiết kiệm lớn về chi phí lao động trực tiếp và làm tăng năng suất. Quá trình hàn sóng (*wave soldering process*, WSP) gồm ba giai đoạn.

- Trong giai đoạn I, được gọi là **tia trợ hàn** (*fluxing, flux formulation*), tia làm sạch mỗi hàn (*soldering flux*) chẳng những làm sạch bề mặt các khớp hàn, mà còn có thể bảo vệ nó chống lại việc bị ôxy hóa lại (*reoxidation*). Tia trợ hàn làm giảm sức căng bề mặt để giúp làm ướt mỗi hàn và hình thành mỗi hàn tốt hơn.
- Giai đoạn II của WSP là **lắp ráp mỗi hàn** (*soldering assembly*). Giai đoạn này được thực hiện trong một chuỗi các máy hàn sóng. Sau khi sấy sơ bộ sản phẩm, bên không chứa linh kiện thành phần của dây chuyền lắp ráp được nhúng vào trong một làn sóng hàn khoảng 1-2 giây. Mọi điểm hàn được hoàn thành khi các CPA thoát ra khỏi làn sóng. Sấy sơ bộ phải được làm dần dần. Sưởi ấm thỏa đáng là điều cần thiết để việc hàn hiệu quả. Tốc độ băng tải và góc của băng tải là rất quan trọng.
- Giai đoạn cuối cùng trong quá trình này, giai đoạn III, là công đoạn **tẩy rửa làm sạch** (*detergent cleaning*). Dây chuyền lắp ráp, đầu tiên được rửa trong dung dịch chất tẩy rửa, sau đó rửa lại bằng nước và cuối cùng là sấy khô bằng khí nóng. Nhiệt độ của dung dịch chất tẩy rửa được nâng lên để đạt được mức làm sạch hiệu quả và ngăn chặn việc tạo bọt quá mức. Nước rửa cũng được làm nóng để có được mức rửa hiệu quả.

Giờ chúng ta xét các bước thiết kế một cách chi tiết.

### 1. Các mục tiêu.

Xác định ảnh hưởng của nhiều yếu tố khác nhau trên chất lượng hàn sóng, và tối ưu hóa qui trình.

### 2. Các biến đáp ứng.

Có 4 biến đáp ứng:

- Mức kháng (*Insulation resistance*);
- Đặc tính làm sạch (*Cleaning characterization*);
- Hiệu quả hàn (*Soldering efficiency*);
- Khả năng nứt bề mặt hàn (*Solder mask cracking*).

### 3. Các nhân tố được kiểm soát.

Có tất cả 17 biến liên quan đến 3 giai đoạn của quá trình, như được cho ở bảng sau.

I. Yếu tố trợ hàn	II. Yếu tố lắp ráp mối hàn	III. Yếu tố tẩy rửa làm sạch
A. Kiểu của chất hoạt hoá ( <i>activator</i> )	G. Lượng chất hòa tan	M. Thiết lập chiều cao sóng
B. Lượng chất hoạt hoá	H. Lượng tia trợ hàn	N. Nồng độ chất tẩy rửa
C. Loại chất xúc tác bề mặt	I. Thời gian làm nóng trước	O. Nhiệt độ chất tẩy rửa
D. Lượng chất xúc tác bề mặt	J. Nhiệt độ hàn	P. Tốc độ băng tải làm sạch
E. Lượng chất chống oxy hóa	K. Tốc độ băng tải	Q. Nhiệt độ nước rửa
F. Loại chất hòa tan ( <i>solvent</i> )	L. Góc băng tải	

#### 4. Các loại phép đo, giá trị đo.

- (a) Kiểm mức kháng khoảng 30 phút, 1 và 4 ngày sau khi hàn ở các nhiệt độ
- $-35^{\circ}\text{C}$ , 90%RN, điện thế không chệch,
  - $-65^{\circ}\text{C}$ , 90%RN, điện thế không chệch.
- (b) Đặc tính làm sạch: lượng chất sót lại trên mạch (biến liên lục).
- (c) Hiệu quả hàn: kiểm tra hình ảnh xem (các khả năng) không có mối hàn, mối hàn không đủ, mối hàn tốt, mối hàn dư và các lỗi khác (các biến rời rạc).
- (d) Khả năng nứt bề mặt hàn: kiểm tra hình ảnh xem có bao nhiêu điểm nứt trên bề mặt (linh kiện, bản mạch được) hàn (là biến rời rạc).

Yếu tố	Số mức chọn	Yếu tố	Số mức chọn
A	2	J	3
B	3	K	3
C	2	L	3
D	3	M	2
E	3	N	2
F	2	O	2
G	3	P	3
H	3	Q	2
I	3		

#### 5. Mức chọn của các yếu tố được kiểm soát.

Xem ở bảng trên. Vậy ta có 10 nhân tố tam phân và 7 nhân tố nhị phân.

6. **Mô hình thống kê.** Các biến đáp ứng có liên quan đến các biến kiểm soát bởi các mô hình tuyến tính có chứa các tham số “ảnh hưởng chính” và “ảnh hưởng tương tác”, sẽ được giải thích ở Mục 8.5.

#### 7. Mảng thử nghiệm.

Một thí nghiệm nhân tố một phần, như sẽ được giải thích trong Mục 9.3, được thiết kế. Thiết kế này là cần thiết, bởi vì một thiết kế nhân tố đầy đủ (hoàn chỉnh) chứa tới  $3^{10}2^7 = 7.558.272$  tổ hợp mức chọn có thể của các yếu tố. Ta chọn ra có dụng ý một thiết kế một phần như một đa tập con (tập con mà phần tử có thể lập lại) của thiết kế nhân tố hoàn chỉnh theo cách thức mà cho phép ta suy luận hợp lý và ước lượng chính xác các thông số ta quan tâm.

## 8. Giao thức thực hiện.

Giả sử rằng ta quyết định thực hiện một thiết kế từng phần gồm  $3^3 2^2 = 108$  thử nghiệm ở các mức chọn nhất định của 17 yếu tố. Tuy nhiên, các thiết lập của các yếu tố cần có thời gian và không ai có thể thực hiện nhiều hơn 4 thử nghiệm một ngày. Các thử nghiệm sẽ kéo dài trong 27 ngày.

- Điều quan trọng là phải xây dựng các thiết kế sao cho những ảnh hưởng quan trọng, mà phải được ước tính, sẽ **không bị lẫn lộn** (*confounded*) với các sự khác biệt có thể có giữa các ngày (các khối). Thứ tự của các thử nghiệm trong mỗi ngày được ngẫu nhiên hóa, cũng như các thử nghiệm đó được gán cho hay sắp xếp vào những ngày khác nhau.
- **Ngẫu nhiên hóa** (*randomization*) là một thành phần quan trọng của thiết kế, được đưa vào kèm theo để nâng cao giá trị suy luận của nó. Giao thức thực hiện cần xác định rõ thứ tự thực hiện các thử nghiệm. ■

## 8.4 Vài nguyên lý và công cụ cho thiết kế thí nghiệm thống kê

### 8.4.1 Phân khối và ngẫu nhiên hóa

**Phân khối** và **ngẫu nhiên hóa** là các nguyên lý cơ bản trong việc hoạch định hay thiết kế các thí nghiệm, nhằm vào mục đích tăng độ chính xác của các kết quả và đảm bảo tính hợp lý và mạch lạc của các suy diễn. Phân khối được sử dụng để làm giảm sai sót. Một **khối** là một phần của thiết kế thống kê (là tập gồm tất cả các thí nghiệm, hay một phần các thí nghiệm) mà dự kiến các phần tử trong phần đó sẽ có tính đồng nhất nhiều hơn khi so với (các phần tử trong) toàn bộ thiết kế.

Ví dụ, nếu thí nghiệm được thiết kế để thử nghiệm tác dụng của lớp phủ polyester lên các vi mạch điện tử trên kết quả đầu ra hiện tại của chúng, sự biến đổi giữa các mạch có thể lớn hơn nhiều so với hiệu ứng của lớp phủ trên kết quả đầu ra hiện tại.

Để giảm bớt thành phần phương sai này, ta có thể phân khối theo vi mạch. Mỗi vi mạch sẽ được kiểm tra theo hai cách xử lý: không có lớp sơn phủ và có lớp sơn phủ. Ta đầu tiên thử nghiệm sản lượng hiện tại của một vi mạch mà không cần sơn. Sau đó ta sơn vi mạch, và kiểm tra lại. Sự so sánh trước và sau khi xử lý như vậy, của cùng một đơn vị, linh kiện, được gọi là **so sánh cặp** (*paired-comparison*).

Một trường hợp phân khối khác là từ ví dụ nổi tiếng của Box, Hunter và Hunter (1978, tr. 97), về các đôi giày của các cậu bé. Hai loại vật liệu đế giày phải được kiểm tra bằng cách xem xét đế của các đôi giày của  $n$  cậu bé, và đo lượng số lần mang giày trong một khoảng thời gian họ tích cực mang giày. Vì có sự dao động lớn giữa hoạt động của các cậu bé trai, nếu có  $m$  đôi có đế loại này và phần còn lại  $n - m$  đôi đế loại kia, ta sẽ không biết được rõ ràng rằng (bất kỳ) **sự khác biệt ở mức độ bào mòn** đế giày quan sát được là đơn thuần do sự khác biệt giữa các đặc tính của vật liệu đế, hay là do sự khác biệt giữa các cậu trai.

Bằng cách phân khối theo các đôi giày, ta có thể làm giảm sự biến đổi rất nhiều. Mỗi đôi giày được gán hai loại đế. Việc so sánh trong mỗi khối là không có biến đổi giữa các cậu trai. Hơn nữa, vì các bé trai sử dụng chân phải hoặc chân trái một cách khác nhau, ta nên dán các loại đế vào giày bên trái hoặc bên phải một cách ngẫu nhiên. Như vậy, các cách xử lý (hai loại đế) được phân bố một cách ngẫu nhiên trong mỗi khối. Các ví dụ khác của phân khối có thể là máy móc, ca sản xuất, các ngày trong tuần, các đốc công, v.v.

Tổng quát, nếu có  $t$  cách xử lý để so sánh, và có  $b$  khối, và nếu tất cả  $t$  cách xử lý có thể thực hiện được trong một khối duy nhất, ta gán tất cả  $t$  cách xử lý vào mỗi khối. Trình tự của việc áp dụng các cách xử lý trong mỗi khối sẽ được chọn ngẫu nhiên.

- Thiết kế này được gọi là một **thiết kế khối đầy đủ ngẫu nhiên** (RCBD, *randomized complete block design*). Ta sẽ xem phía sau cách thức một phân tích thỏa đáng sản lượng (i.e. biến đáp ứng) có thể, một cách hợp lệ, kiểm tra các ảnh hưởng của các cách xử lý lên biến ấy.
- Trường hợp không phải tất cả cách xử lý có thể được áp dụng trong mỗi khối, việc gán các cách xử lý cho các khối theo một nguyên lý cân bằng nào đó là đáng ước muốn.

Thiết kế như vậy, sẽ được thảo luận sau, được gọi là **thiết kế khối không đầy đủ cân bằng** (BIBD, *balanced incomplete block design*, BIBD).

Ngẫu nhiên hóa trong mỗi khối cũng quan trọng nữa, vì nhằm để **xác nhận giả định** rằng các thành phần lỗi trong mô hình thống kê là **độc lập**. Giả định này có thể không thỏa mãn (không được đảm bảo) nếu các cách xử lý không được phân bố ngẫu nhiên vào các đơn vị thí nghiệm trong mỗi khối. Tuy nhiên bạn thấy rằng nguyên lý ngẫu nhiên hóa hiển nhiên không nên áp dụng vào trong thiết kế cảnh quan, như Hình ?? chỉ ra, phải không?

## 8.4.2 Công cụ cơ bản

**Kết quả cần ghi nhớ 5.** Ta nhắc lại thống kê Fisher, mà sẽ rất hữu ích ở các phần sau.

Ta xem hai mẫu độc lập kích thước  $n_1$  và  $n_2$ , đã được lấy từ các quần thể Gauss có phương sai là  $\sigma_1^2$  và  $\sigma_2^2$  tương ứng, Gọi

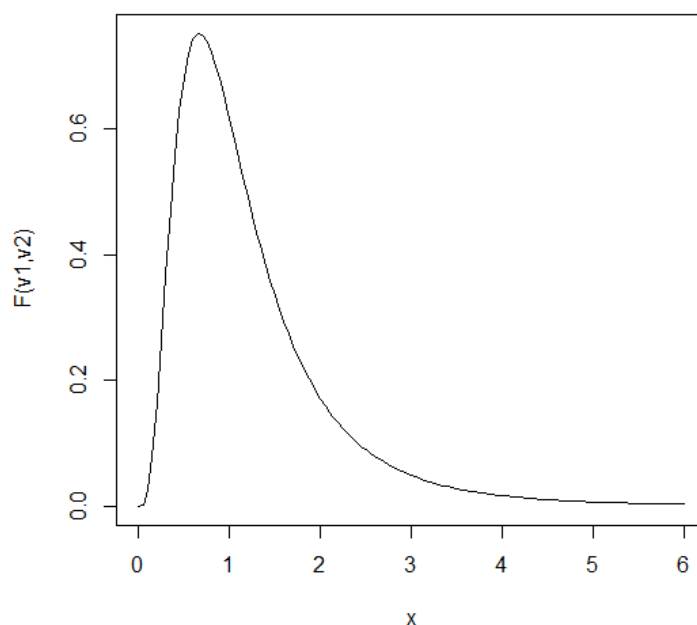
$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2, \text{ và}$$

$$S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2$$

là phương sai của hai mẫu, có  $\bar{X}_1$  và  $\bar{X}_2$  là trung bình mẫu tương ứng. Tỷ lệ

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}$$

có phân phối biểu thị bởi  $F[\nu_1, \nu_2]$ , với  $\nu_1 = n_1 - 1$  và  $\nu_2 = n_2 - 1$ . Phân phối này được gọi là phân phối  $F$  (Fisher) với  $\nu_1$  và  $\nu_2$  bậc tự do. Hình 8.1 cho ta đồ thị của hàm mật độ của  $F[\nu_1, \nu_2]$ .

Hình 8.1: Hàm mật độ của  $F(\nu_1, \nu_2)$ 

Giá trị kỳ vọng và phương sai của  $F[\nu_1, \nu_2]$  là

$$\mathbf{E}(F[\nu_1, \nu_2]) = v_2 / (v_2 - 2); \quad \text{khi } v_2 > 2$$

$$\mathbf{V}(F[\nu_1, \nu_2]) = \frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)}; \quad \text{khi } v_2 > 4.$$

Phân vị thứ  $1 - \alpha$  của  $F[\nu_1, \nu_2]$ ,  $F_{1-\alpha}[\nu_1, \nu_2]$  có thể được tính bằng MINITAB. Nếu muốn có được **phân vị** (fractile) thứ  $\alpha$ ,  $F_\alpha[\nu_1, \nu_2]$  cho các giá trị  $\alpha < 0.5$ , ta có thể áp dụng mối quan hệ

$$F_{1-\alpha}[\nu_1, \nu_2] = \frac{1}{F_\alpha[\nu_2, \nu_1]}.$$

Ví dụ ta tính  $F_{0.05}[15, 10] = 1/F_{0.95}[10, 15] = 1/2.54 = .3937$ .

## 8.5 Mô hình cộng tính và không cộng tính

Mười bảy yếu tố mà có thể ảnh hưởng đến kết quả lên WSP (quá trình hàn sóng) được liệt kê trong Ví dụ 8.1. Vài yếu tố trong số này, như kiểu của chất hoạt hoá (*activator types*, A), hoặc loại chất xúc tác bề mặt (C) là các biến phân loại (*categorical*). Số lượng các mức chọn được liệt kê cho các yếu tố này là 2. Đó là, nghiên cứu đã xét nhằm so sánh tác dụng của hai loại chất hoạt hoá và hai loại xúc tác bề mặt. Nếu các biến là liên tục, như số lượng chất hoạt hoá (B), ta có thể sử dụng một mô hình hồi quy tuyến tính để diễn đạt ảnh hưởng của các yếu tố trên các biến năng suất. Các mô hình này sẽ được thảo luận sau. Trong phần này ta trình bày các mô hình tuyến tính mà có giá trị sử dụng cho cả hai loại biến phân loại hoặc liên tục.

### 8.5.1 Thiết kế ngẫu nhiên hoàn toàn một yếu tố (CRD - *Completely Randomized Design*)

Để tiện giải thích, ta hãy bắt đầu với một trường hợp đơn giản, trong đó biến đáp ứng phụ thuộc vào một yếu tố duy nhất. Ký hiệu  $A$  cho yếu tố ấy, được áp dụng ở các mức chọn  $A_1, \dots, A_a$  khác nhau. Các mức này có thể là  $a$  loại, phạm trù, còn được gọi là **giải pháp** hay **cách xử lý** (*treatments*).

- Giả sử tại mỗi mức của  $A$ , ta tiến hành lặp đi lặp lại cùng một thí nghiệm  $n$  lần độc lập nhau (*replicas*). Gọi  $Y_{ij}$  ( $i = 1, 2, \dots, a$  và  $j = 1, 2, \dots, n$ ) là đáp ứng quan sát được tại lần lặp thứ  $j$  của mức  $A_i$ . Ta mô hình hóa biến ngẫu nhiên  $Y_{ij}$  bởi

$$Y_{ij} = \mu + \tau_i^A + e_{ij}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n, \quad (8.1)$$

trong đó  $\mu$  và  $\tau_1^A, \dots, \tau_a^A$  là các tham số chưa biết, thỏa mãn

$$\sum_{i=1}^a \tau_i^A = 0. \quad (8.2)$$

- Sai số (lỗi)  $e_{ij}$  (với mọi  $i = 1, 2, \dots, a$ ;  $j = 1, 2, \dots, n$ ) là các biến ngẫu nhiên độc lập sao cho

$$\mathbf{E}[e_{ij}] = 0 \text{ và } \mathbf{V}(e_{ij}) = \sigma^2. \quad (8.3)$$

Đặt  $\bar{Y}_i = \frac{1}{n} \sum_{j=1}^n Y_{ij}$ ,  $i = 1, 2, \dots, a$ . Các giá trị kỳ vọng của các trung bình này là

$$\mathbf{E}[\bar{Y}_i] = \mu + \tau_i^A, \quad i = 1, 2, \dots, a. \quad (8.4)$$

**Khái niệm 8.1.** Tham số  $\tau_i^A$  được gọi là **ảnh hưởng chính** (tác dụng chính, *main effect*) của nhân tố  $A$  ở mức  $A_i$ . Giá trị trung bình của tất cả  $N = a \times n$  quan sát (gọi là trung bình toàn thể) được cho bởi

$$\bar{\bar{Y}} = \frac{1}{a} \sum_{i=1}^a \bar{Y}_i. \quad (8.5)$$

Vì  $\sum_{i=1}^a \tau_i^A = 0$  nên ta có kỳ vọng của trung bình toàn thể

$$\mathbf{E}[\bar{\bar{Y}}] = \mu. \quad (8.6)$$

Thiết kế CRD cùng với số liệu thực nghiệm có thể biểu diễn tổng quát trong bảng sau đây



Bảng 8.1: Thiết kế CRD, nguồn [16]

Giải pháp	Quan sát					Tổng	Trung bình $\bar{Y}_i$
	1	2	...	$n-1$	$n$		
1	$y_{11}$	$y_{12}$	...	$y_{1n-1}$	$y_{1n}$	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
2	$y_{21}$	$y_{22}$	...	$y_{2n-1}$	$y_{2n}$	$y_{2\cdot}$	$\bar{y}_{2\cdot}$
⋮							
$a$	$y_{a1}$	$y_{a2}$	...	$y_{an-1}$	$y_{an}$	$y_{a\cdot}$	$\bar{y}_{a\cdot}$
						$y_{..} = \sum_{ij} y_{ij}$	$\bar{y}_{..} = \bar{\bar{y}}$

**VẤN ĐỀ 9. [Ứng dụng trong Môi trường - Công nghiệp.]****Thiết kế CRD cho sản xuất bao bì ‘xanh’**

Một nhà sản xuất sử dụng giấy làm túi hàng tạp hóa thân thiện với môi trường quan tâm tới việc cải thiện **sức bền** của sản phẩm. Các kỹ sư cho rằng độ bền sợi vải là một hàm theo lượng (nồng độ) gỗ cứng trong bột giấy và khoảng biến thiên hàm lượng được quan tâm thực tế là ở giữa 5% và 20%.

Một nhóm các kỹ sư chịu trách nhiệm nghiên cứu quyết định khảo sát bốn mức hàm lượng gỗ: 5%, 10%, 15%, và 20%. Họ quyết định tạo ra sáu mẫu thử tại mỗi mức hàm lượng, sử dụng một nhà xưởng thí điểm. Tất cả 24 mẫu túi hàng được thử nghiệm trên một máy thử độ bền sợi, theo một thứ tự ngẫu nhiên. Các dữ liệu từ thí nghiệm này được thể hiện trong bảng sau.

Bảng 8.2: Thiết kế CRD - kết quả thực nghiệm, nguồn [16]

Hàm lượng gỗ (%)	Quan sát						Tổng	Trung bình $\bar{Y}_i$
	1	2	3	4	5	6		
5	7	8	15	11	9	10	60	10.00
10	12	17	13	18	19	15	94	15.67
15	14	18	19	17	16	18	102	17.00
20	19	25	22	23	18	20	127	21.17
							$\sum_{ij} y_{ij} = 383$	$\bar{\bar{y}} = 15.96$

Thử nghiệm này là một CRD. Yếu tố độ cứng gỗ  $A$  có  $a = 4$  mức chọn mà ta gọi là cách xử lý

(*treatments*), mỗi cách xử lý được thử nghiệm  $n = 6$  lần và ta gọi là chúng là lần lặp (thực nghiệm, *replicates*), được tiến hành theo thứ tự ngẫu nhiên. Các trung bình ứng với mỗi cách xử lý thứ  $i$  là

$$\bar{Y}_i = \frac{1}{6} \sum_{j=1}^n Y_{ij}, \quad i = 1, 2, \dots, 4;$$

chính là các số ở cột chót của Bảng 8.2.

Ta có mô hình cho biến độ bền sợi vải là

$$Y_{ij} = \mu + \tau_i + e_{ij}, \quad i = 1, 2, \dots, 4, \quad j = 1, 2, \dots, 6, \quad (8.7)$$

các sai số  $e_{ij}$  là các biến ngẫu nhiên độc lập sao cho

$$\mathbf{E}[e_{ij}] = 0 \text{ và } \mathbf{V}(e_{ij}) = \sigma^2, \quad \forall i = 1, 2, \dots, 4; j = 1, 2, \dots, 6.$$

Do  $\mathbf{E}[e_{ij}] = 0$  các giá trị kỳ vọng của các trung bình này là

$$\mathbf{E}[\bar{Y}_i] = \mu + \tau_i, \quad i = 1, 2, \dots, 4; \quad (8.8)$$

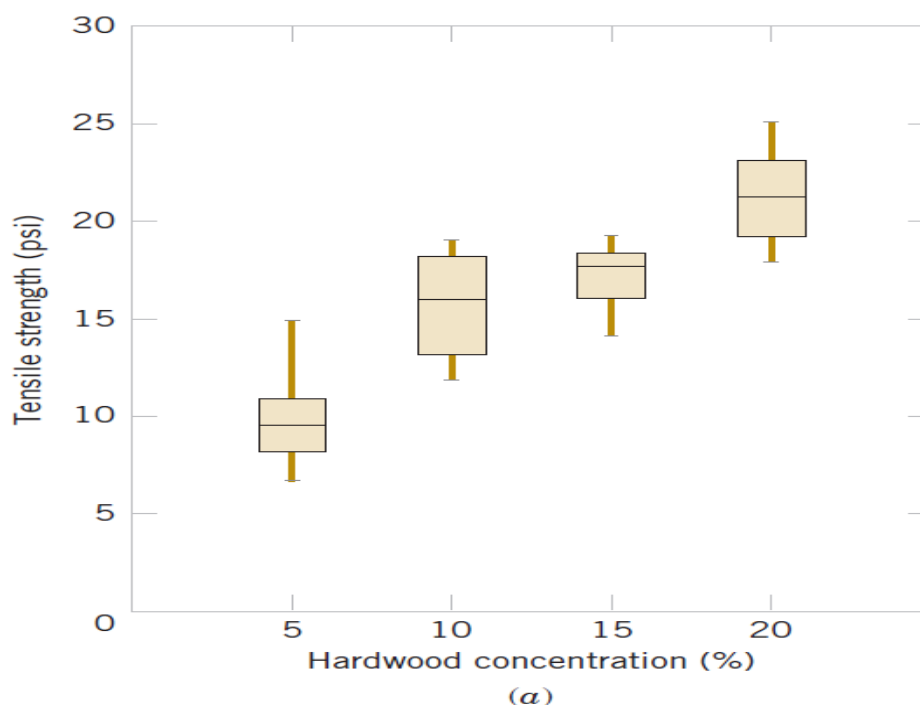
trong đó tham số  $\tau_i$  là **ảnh hưởng chính** của độ cứng gỗ  $A$  ở mức  $i$ .

Ta có thể sử dụng phân tích phương sai để kiểm tra giả thuyết rằng độ cứng gỗ khác nhau không ảnh hưởng đến độ bền sợi trung bình của túi hàng. Các giả thuyết là

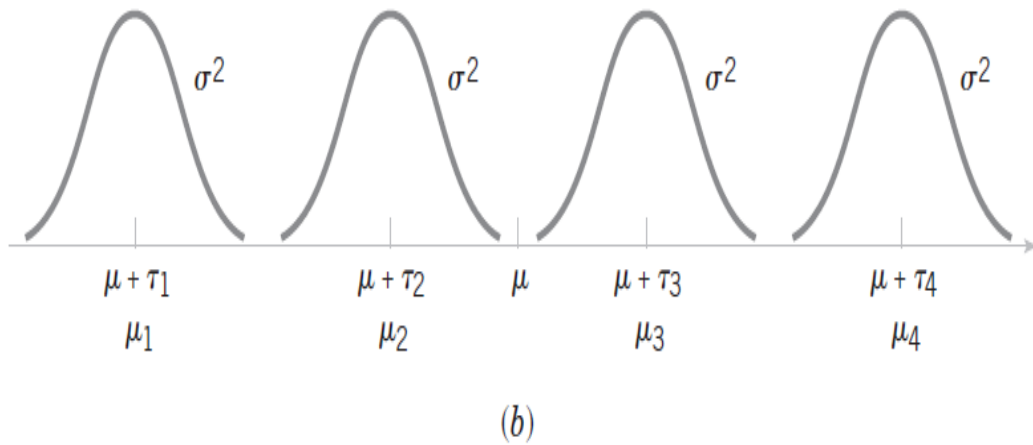
$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0,$$

$$H_1 : \tau_i \neq 0 \text{ với một } i \in \{1, 2, 3, 4\}.$$

Với các số liệu như trên ta có boxplot được cho trong hình (a) sau



Nếu giả định rằng 4 tổng thể ứng với 4 cách xử lý là Gauss và có cùng phương sai  $\sigma^2$  thì biểu đồ ảnh hưởng của các cách xử lý lên độ bền túi xách  $Y$  được cho như hình (b) sau



Tổng các bình phương của các độ lệch quan sát so với  $\bar{\bar{Y}}$ , của cách xử lý và lỗi ngẫu nhiên theo thứ tự là

$$SST = \sum_{i=1}^4 \sum_{j=1}^6 (y_{ij} - \bar{\bar{y}})^2 = SS_{Treatm} + SS_E = 512.96,$$

$$SS_{Treatm} = n \sum_{i=1}^4 (\bar{y}_i - \bar{\bar{y}})^2 = 382.79,$$

$$SS_E = SST - SS_{Treatm} = 512.96 - 382.79 = 130.17.$$

ANOVA được tóm tắt trong Bảng 8.3:

**Bảng 8.3:** Bảng ANOVA một yếu tố độ cứng gỗ (so sánh với Vấn đề 7).

Sự biến thiên	D.F.	S.S.	M.S.	Thống kê $F$
Cách xử lý độ cứng gỗ A	$a - 1 = 3$	$SS_{Treatm} = 382.79$	$MSB = \frac{SS_{Treatm}}{a - 1} = 127.60$	19.60
Lỗi ngẫu nhiên	$N - a = 20$	$SS_E = 130.17$	$MSW = \frac{SS_E}{N - a} = 6.51$	
Toàn bộ	$N - 1 = 23$	$SST = 512.96$	-	

Vì thống kê  $F$  từ dữ liệu là  $F_0 = 19.6 > f_{\alpha; a-1, N-a} = f_{0.01; 3, 20} = 4.94$  chúng ta bác bỏ  $H_0$  và kết luận với mức nghĩa 1% hàm lượng gỗ cứng trong bột giấy ảnh hưởng đáng kể lên độ bền túi xách trung bình.

### Thiết kế CRD một yếu tố - Khi các giải pháp không cân bằng

Giả sử giải pháp thứ  $i$  cần  $n_i$  lần thử lặp (các giải pháp không cân bằng), khi đó các tổng bình phương trong ANOVA là

$$\begin{aligned} SST &= \sum_{i=1}^a \sum_{j=1}^{n_i} y_{ij}^2 - \frac{(\bar{y})^2}{N}, \\ SS_{Treatm} &= \sum_{i=1}^a \frac{y_{i.}^2}{n_i} - \frac{(\bar{y})^2}{N}, \\ SS_E &= SST - SS_{Treatm}. \end{aligned}$$

### Thiết kế CRD một yếu tố - So sánh từng cặp trung bình

Ta gọi sai biệt ý nghĩa ít nhất (*least significant difference* - LSD), mọi  $n_i = n$ , là

$$LSD = t_{1-\alpha/2}[a(n-1)] \sqrt{\frac{2MS_E}{n}}. \quad (8.9)$$

Nếu các giải pháp không cân bằng thì gọi

$$LSD = t_{1-\alpha/2}[N-a] \sqrt{MS_E \left( \frac{1}{n_i} + \frac{1}{n_i} \right)}. \quad (8.10)$$

Trong VD trên ta kết luận với mức nghĩa 1% hàm lượng gỗ cứng trong bột giấy ảnh hưởng đáng kể lên độ bền túi xách trung bình, nhưng làm sao chọn ra hàm lượng nào để sản xuất? Dùng pp LSD ta có  $a = 4$  trung bình,  $n = 6$ ,  $MS_E = MSW = 6.51$ , và  $t_{0.975}[20] = 2.086$ . Trung bình các giải pháp (ứng với hàm lượng gỗ) là  $\bar{y}_{1.} = 10$  psi,  $\bar{y}_{2.} = 15.67$  psi,  $\bar{y}_{3.} = 17$  psi,  $\bar{y}_{4.} = 21.17$  psi, và giá trị

$$LSD = t_{0.975}[20] \sqrt{\frac{2MS_E}{n}} = 2.086 \sqrt{2(6.51)/6} = 3.07.$$

Vậy cặp trung bình mẫu giải pháp  $\bar{y}_{i.}, \bar{y}_{j.}$  nào có sai biệt hơn  $LSD = 3.07$  thì cặp trung bình tổng thể của các giải pháp tương ứng  $\mu_i, \mu_j$  là khác biệt nhau. Như trung bình mẫu giải pháp 4 vs. 1 cho  $\bar{y}_{4.} - \bar{y}_{1.} = 21.17 - 10.00 = 11.17 > 3.07$ , thì kết luận các giải pháp tương ứng  $\mu_4, \mu_1$  khác biệt là có ý nghĩa thống kê.

### 8.5.2 Trường hợp nhiều yếu tố

Nếu có hai yếu tố,  $A$  và  $B$ , có tương ứng  $a$  và  $b$  mức chọn, thì sẽ có cả thảy  $a \times b$  giải pháp tổ hợp ( $A_i, B_j$ ) (hay gọn hơn giải pháp - *treatment combinations*),  $i = 1, 2, \dots, a, j = 1, \dots, b$ . Ta cũng giả sử rằng  $n$  bản sao độc lập (*replicas*) được thực hiện tại mỗi một trong các giải pháp. Đáp ứng của bản sao thứ  $k$  của giải pháp ( $A_i, B_j$ ) được cho bởi

$$Y_{ijk} = \mu + \tau_i^A + \tau_j^B + \tau_{ij}^{AB} + e_{ijk}. \quad (8.11)$$

Các lỗi ngẫu nhiên  $e_{ijk}$  là các biến ngẫu nhiên độc lập sao cho

$$\mathbf{E}[e_{ijk}] = 0 \text{ và } \mathbf{V}[e_{ijk}] = \sigma^2, \quad (8.12)$$

với mọi  $i = 1, 2, \dots, a, j = 1, 2, \dots, b, k = 1, 2, \dots, n$ .

Ta giả định thêm rằng

$$\begin{aligned} \sum_{j=1}^b \tau_{ij}^{AB} &= 0, \quad i = 1, 2, \dots, a \\ \sum_{i=1}^a \tau_{ij}^{AB} &= 0, \quad j = 1, 2, \dots, b. \\ \sum_{i=1}^a \tau_i^A &= 0, \quad \sum_{j=1}^b \tau_j^B = 0. \end{aligned} \quad (8.13)$$

**Định nghĩa 8.1.**  $\tau_i^A$  là ảnh hưởng chính của nhân tố  $A$  ở mức chọn  $A_i$ ,  $\tau_j^B$  là ảnh hưởng chính của nhân tố  $B$  ở mức chọn  $B_j$ , và  $\tau_{ij}^{AB}$  là ảnh hưởng tương tác/phối hợp (*interaction effect*) của hai nhân tố  $A, B$  ở mức chọn tổ hợp  $(A_i, B_j)$ .

Nếu tất cả các tác động tương tác bằng không, mô hình rút gọn thành

$$Y_{ijk} = \mu + \tau_i^A + \tau_j^B + e_{ijk}. \quad (8.14)$$

Một mô hình như vậy được gọi là **cộng tính**. Nếu không phải tất cả các thành phần tương tác bằng không, mô hình được gọi là **không cộng tính**.

**Chú ý.** Mô hình này được khái quát hóa một cách dễ dàng để bao gồm một số lượng lớn các yếu tố. Như vậy, với ba yếu tố  $A, B$  và  $C$ , ta sẽ có

- \* ba thành phần ảnh hưởng chính  $\tau_i^A, \tau_j^B$  và  $\tau_k^C$ ; ba ảnh hưởng tương tác kép  $\tau_{ij}^{AB}, \tau_{ik}^{AC}$  và  $\tau_{jk}^{BC}$ ;
- \* và sau chót là một ảnh hưởng tương tác  $\tau_{ijk}^{ABC}$ .

Tổng quát hóa, nếu có tới  $p$  yếu tố, thì có tổng cộng  $2^p$  các thông số,

$$\mu, \tau_i^A, \tau_j^B, \dots, \tau_{ij}^{AB}, \tau_{ik}^{AC}, \dots, \tau_{ijk}^{ABC}, \dots$$

- Các tham số tương tác giữa hai yếu tố được gọi là **tương tác kép hay bậc 1**.
- Các tham số tương tác giữa ba yếu tố được gọi là **tương tác bậc 2**, và v.v.

Trong mô hình hóa thực tế ta thường được cho giả định rằng tất cả các tham số tương tác bậc cao hơn bậc 1 (như  $\tau_{ijk}^{ABC}$ ) là zero.

## 8.6 Phân tích các thiết kế khối hoàn chỉnh và ngẫu nhiên

### 8.6.1 Trường hợp nhiều khối, hai giải pháp cho mỗi khối: so sánh ghép cặp

Như trong ví dụ để giày (hay về tác dụng của lớp phủ polyester trên đầu ra vi mạch), có hai giải pháp được áp dụng cho mỗi một trong  $n$  khối.

Mô hình tuyến tính được viết là

$$Y_{ij} = \mu + \tau_i + \beta_j + e_{ij}, \quad (8.15)$$

trong đó  $\tau_i$  là ảnh hưởng của giải pháp thứ  $i$  và  $\beta_j$  là ảnh hưởng của khối thứ  $j$ , với các chỉ số  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, n$ .

- Lúc này  $e_{ij}$  là biến ngẫu nhiên độc lập mà đại diện cho các lỗi hoặc sai lệch ngẫu nhiên thực nghiệm. Ta giả sử rằng  $\mathbf{E}[e_{ij}] = 0$  và  $\mathbf{V}[e_{ij}] = \sigma_e^2$ .
- Vì ta quan tâm đến việc kiểm tra xem có phải hai giải pháp có hiệu ứng khác nhau không, việc phân tích được dựa trên các sự khác biệt trong khối, cho bởi

$$D_j = Y_{2j} - Y_{1j} = \tau_2 - \tau_1 + e_j^*, \quad j = 1, 2, \dots, n. \quad (8.16)$$

- Các lỗi ngẫu nhiên  $e_j^*$  là các biến ngẫu nhiên độc lập có kỳ vọng  $\mathbf{E}[e_j^*] = 0$  và phương sai  $\mathbf{V}[e_j^*] = \sigma_d^2$ , với mọi  $j = 1, 2, \dots, n$  trong đó

$$\sigma_d^2 = 2\sigma_e^2,$$

do bởi  $e_j^* = e_{2j} - e_{1j}$ .

- Ước lượng không chệch của  $\sigma_d^2$  là

$$S_d^2 = \frac{1}{n-1} \sum_{j=1}^n (D_j - \bar{D}_n)^2, \quad (8.17)$$

trong đó  $\bar{D}_n = \frac{1}{n} \sum_{j=1}^n D_j$ . Các giả thuyết cần được kiểm định là:

$$H_0 : \delta = \tau_2 - \tau_1 = 0, \text{ đối nghịch với}$$

$$H_1 : \delta \neq 0.$$

### Thống kê kiểm tra $t$ ( $t$ -test)

Thường được dùng nhất là  $t$ -test, trong đó  $H_0$  được kiểm nghiệm bằng cách tính toán thống kê

$$t = \frac{\sqrt{n} \bar{D}_n}{S_d}. \quad (8.18)$$

Nếu  $e_1^*, \dots, e_n^*$  là các biến ngẫu nhiên độc lập và có cùng phân phối chuẩn thì, dưới giả thuyết ban đầu  $H_0$ , thống kê  $t$  có một phân phối Student với  $(n-1)$  bậc tự do. Trong trường hợp này, với  $\alpha$  là mức ý nghĩa đã chọn thì  $H_0$  bị bác bỏ nếu

$$|t| > t_{1-\alpha/2}[n-1].$$

**Các kiểm tra ngẫu nhiên**

Một kiểm tra ngẫu nhiên cho sự so sánh cặp xây dựng một **phân phối tham khảo** (*reference distribution*) của tất cả các trung bình có thể có của các sai biệt, mà có thể thu được bằng cách gán ngẫu nhiên dấu + hoặc – vào giá trị của  $D_i$ . Khi đó kiểm tra ngẫu nhiên sẽ tính một sai biệt trung bình  $\bar{D}$  cho mỗi một trong  $2^n$  phép gán dấu.

Giá trị  $P$  của kiểm tra, cho đối thuyết hai phía  $H_1$ , xác định theo phân phối tham khảo là

$$P = \mathbb{P}[|\bar{D}| \geq |\bar{D} \text{ quan sát được}|].$$

**Bảng 8.4:** Cách gán dấu và giá trị  $\bar{D}$

Các dấu				$\bar{D}$
–1	–1	–1	–1	–0.55
1	–1	–1	–1	0
–1	1	–1	–1	–0.4
1	1	–1	–1	0.15
–1	–1	1	–1	–0.20
1	–1	1	–1	0.35
–1	1	1	–1	–0.05
1	1	1	–1	0.50
–1	–1	–1	1	–0.50
1	–1	–1	1	0.05
–1	1	–1	1	–0.35
1	1	–1	1	0.2
–1	–1	1	1	–0.15
1	–1	1	1	0.40
–1	1	1	1	0
1	1	1	1	0.55

Ví dụ, giả sử ta có bốn sai biệt, với giá trị 1.1, 0.3, –0.7, –0.1. Giá trị trung bình là  $\bar{D}_4 = 0.15$ . có  $2^4 = 16$  cách gán dấu vào  $D_i$  (xem Bảng 8.4). Dưới phân phối tham khảo, tất cả các trung bình có thể là có khả năng bằng nhau. Giá trị  $P$ , kết hợp với trị quan sát được  $\bar{D} = 0.15$ , là  $P = 12/16 = 0.75$ . Nếu số lượng các cặp (các khối)  $n$  là lớn thủ tục trở nên cồng kềnh, vì ta phải xác định tất cả  $2^n$  phép gán dấu. Nếu  $n = 20$  có tới  $2^{20} = 1.048.576$  phép gán như vậy.

Tuy nhiên ta có thể ước tính giá trị  $P$  bằng cách lấy một RSWR từ phân phối tham khảo này. Điều này có thể dễ dàng thực hiện bằng cách sử dụng chương trình R, được thực hiện như sau:

```
> X <- c(1.1, 0.3, -0.7, -0.1); M <- 200; set.seed(123)
> Di <- matrix(sample(x=c(-1,1), size=length(X)*M, replace=TRUE), nrow=M)
> Xi <- matrix(X, nrow=M, ncol=length(X), byrow=TRUE)
> sum(rowMeans(Di*Xi) >= mean(X))/M
[1] 0.355
```

### 8.6.2 Trường hợp nhiều khối, $t$ giải pháp cho mỗi khối

Như đã nói trước đây, các **thiết kế khối hoàn chỉnh ngẫu nhiên** (RCBD, *randomized complete block designs*) là những thiết kế trong đó mỗi khối chứa tất cả  $t$  giải pháp. Các giải pháp được gán cho các đơn vị thí nghiệm của mỗi khối một cách ngẫu nhiên. Gọi  $b$  là số lượng các khối.

- Mô hình tuyến tính cho những thiết kế này là

$$Y_{ij} = \mu + \tau_i + \beta_j + e_{ij}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, b, \quad (8.19)$$

với  $y_{ij}$  là kết quả của giải pháp  $i$  trong khối thứ  $j$ . Tác dụng (ảnh hưởng) chính của giải pháp thứ  $i$  là  $\tau_i$ , và tác dụng chính của khối thứ  $j$  là  $\beta_j$ .

- Ta giả định rằng những tác động là cộng tính (nghĩa là chúng không tương tác nhau). Theo giả thuyết này, mỗi giải pháp được thử chỉ một lần trong mỗi khối. Các khối khác nhau đóng vai trò của các bản sao. Tuy nhiên, vì các khối có thể có hiệu ứng cộng tính,  $\beta_j$ , ta phải điều chỉnh ảnh hưởng của những khối khi ước tính  $\sigma^2$ . Điều này được thực hiện trong Bảng ANOVA 8.5.

**Bảng 8.5:** Phân tích ANOVA cho RCBD

Nguồn sinh ra sự biến đổi	DF	SS	MS	E(MS)
Các giải pháp	$t - 1$	$SSTR$	$MSTR$	$\sigma^2 + \frac{b}{t-1} \sum_{i=1}^t \tau_i$
Các khối	$b - 1$	$SSBL$	$MSBL$	$\sigma^2 + \frac{t}{b-1} \sum_{j=1}^b \beta_j$
Lỗi	$(t-1)(b-1)$	$SSE$	$MSE$	$\sigma^2$
Tổng	$tb - 1$	$SST$	-	

Ta tiếp tục giả định rằng  $e_{ij}$  là biến ngẫu nhiên lỗi với  $\mathbf{E}(e_{ij}) = 0$  và  $\mathbf{V}(e_{ij}) = \sigma^2$ , cho tất cả  $(i, j)$ . Phân tích ANOVA cho mô hình này được trình bày trong bảng trên. Ở đây,

$$SST = \sum_{i=1}^t \sum_{j=1}^b (Y_{ij} - \bar{\bar{Y}})^2, \quad [\text{tổng bình phương tất cả lỗi}] \quad (8.20)$$

$$SSTR = b \sum_{i=1}^t (\bar{Y}_{i.} - \bar{\bar{Y}})^2, \quad [\text{tổng bình phương lỗi do giải pháp}] \quad (8.21)$$

$$SSBL = t \sum_{j=1}^b (\bar{Y}_{.j} - \bar{\bar{Y}})^2, \quad [\text{tổng bình phương lỗi do khối}] \quad (8.22)$$

và  $SSE = SST - SSTR - SSBL$  [tổng bình phương lỗi do tính ngẫu nhiên]. Ngoài ra

$$\bar{Y}_{i.} = \frac{1}{b} \sum_{j=1}^b Y_{ij}, \quad \bar{Y}_{.j} = \frac{1}{t} \sum_{i=1}^t Y_{ij} \quad (8.23)$$

và  $\bar{\bar{Y}}$  là trung bình toàn thể các kết xuất.



- Tầm quan trọng (thống kê) của **ảnh hưởng giải pháp** được kiểm tra bởi thống kê (Fisher)  $F$

$$F_t = \frac{MSTR}{MSE}. \quad (8.24)$$

- Tầm quan trọng của **ảnh hưởng khối** được kiểm tra bởi thống kê

$$F_b = \frac{MSBL}{MSE}. \quad (8.25)$$

- Những số liệu thống kê này được so sánh tương ứng với phân vị thứ  $(1 - \alpha)$  của phân phối  $F$ . Theo giả định  $\sum_{i=1}^t \tau_i = 0$  thì **các ảnh hưởng chính của các giải pháp** được ước tính bởi

$$\hat{\tau}_i = \bar{Y}_{i.} - \bar{\bar{Y}}, \quad i = 1, 2, \dots, t. \quad (8.26)$$

- Đây là ước lượng bình phương tối thiểu. Mỗi ước lượng  $\hat{\tau}_i$  như thế là một **ràng buộc tương phản tuyến tính** (*linear contrast*)

$$\hat{\tau}_i = \sum_{i'=1}^t c_{ii'} \bar{Y}_{i'}. \quad (8.27)$$

trong đó các hệ số ràng buộc

$$c_{ii'} = \begin{cases} 1 - \frac{1}{t} & \text{khi } i = i' \\ -\frac{1}{t} & \text{khi } i \neq i'. \end{cases} \quad (8.28)$$

Do vậy phương sai của **ảnh hưởng chính của giải pháp** thứ  $i$  là

$$\mathbf{V}[\tau_i] = \frac{\sigma^2}{b} \sum_{i'=1}^t c_{ii'}^2 = \frac{\sigma^2}{b} \left(1 - \frac{1}{t}\right), \quad i = 1, 2, \dots, t. \quad (8.29)$$

Một ước lượng không chệch của  $\sigma^2$  được cho bởi  $MSE = SSE/(t-1)(b-1)$ , là trung bình tổng bình phương lỗi ngẫu nhiên. Vậy khoảng tin cậy đồng thời cho  $\tau_i$  ( $i = 1, 2, \dots, t$ ), theo phương pháp Scheffé, là

$$\tau_i \pm S_\alpha \left( \frac{MSE}{b} \left(1 - \frac{1}{t}\right) \right)^{1/2}, \quad i = 1, 2, \dots, t \quad (8.30)$$

trong đó

$$S_\alpha = \left[ (t-1) F_{1-\alpha}[t-1, (t-1)(b-1)] \right]^{1/2}.$$

**Ví dụ 8.2.** Xem VD **Fabric Strength** có data trong Table 13-12 của Montgomery.

## 9.3 Thiết kế nhân tố trong sản xuất và kỹ nghệ chất lượng

### 9.3.1 Quy trình dùng thiết kế thí nghiệm (DOE) trong công nghiệp

Một cách tổng quát, mục tiêu chính của DOE là nhằm xác định một hàm  $\phi : D \rightarrow \mathbb{R}$  trên một thiết kế (nhân tố)  $D$ , trong đó

- $\phi$  là mô hình toán học của một đại lượng ta đang quan tâm trong nghiên cứu hay phát triển ứng dụng cụ thể, như là độ hiệu quả, chất lượng, sự trung thành của khách, độ cứng vòng bi xe (*efficiency, quality, customer loyalty, bearing hardness*),
- và ta muốn xác định  $\phi$  tường minh hay là tối ưu hóa nó.

Quy trình gồm nhiều bước (lặp) như sau:

**a/ Phát biểu mục tiêu** viết rõ mục đích của thí nghiệm hay dự án.

**b/ Chọn hàm đáp ứng** là về sự tư vấn, phải hỏi khách hàng (hay chính bạn) họ muốn ước lượng hay tối ưu hóa.

**c/ Phân tích dữ liệu tiền thí nghiệm** Tiến hành khai thác thông tin sơ bộ từ bản phỏng vấn khách hàng. Một cách toán học, là xác định các nhân tố cấu thành  $D$ , cũng như các giá trị thực nghiệm của hàm  $\phi$

**d/ Chọn yếu tố và đề xuất mô hình  $\phi$**  Bạn phải dùng sơ đồ (*flowchart*) để biểu đạt quá trình hay hệ thống, xác định tập các yếu tố mà bạn có thể điều khiển được (*controllable factors*) ảnh hưởng lên quá trình cùng với các nhân tố nhiễu (*noisy- uncontrollable factors*); dùng biểu đồ nhân quả nhằm liệt kê ra các yếu tố chính có tác động quan trọng trên đáp ứng  $\phi$ .

**e/ Chọn thiết kế thí nghiệm** từ tập các thiết kế đã có tùy theo mục tiêu nghiên cứu.

Toán học cực kỳ quan trọng ở bước này, vì: nhu cầu ứng dụng cần thiết kế có cỡ lớn hơn (có nhiều nhân tố hơn ảnh hưởng đến quá trình - hàm đáp ứng  $\phi$ , và số tổ hợp thực nghiệm ngày càng tăng hơn).

**f/ Tiến hành thí nghiệm** trong phòng thí nghiệm, nhà máy hay hiện trường ... Khâu này cần các máy móc, thiết bị (đôi khi rất đắt tiền), nhân lực ... nhằm thu thập dữ liệu thực nghiệm.

**g/ Phân tích dữ liệu** (data analysis/ interpretation) tính toán trên dữ liệu, hiểu các kết quả tính toán biểu diễn cho tính chất nào của quá trình hay đáp ứng. Bạn phải trừu tượng trước, trong giai đoạn thiết kế, các tính chất nào cần lượng hóa hay ước tính. Việc trừu tượng này giúp nhà thí nghiệm tự tin rằng họ đang thu thập đúng dữ liệu mà họ cần phân tích.

**h/ Kết luận** và đưa ra khuyến cáo. Việc này yêu cầu kiểm định các giả thuyết thống kê là đúng hay sai, dựa vào các phát biểu đã đưa ra trong bước a- d.

Toàn bộ qui trình nêu trên là một tiến trình lặp. Trong thực tế, nếu ngân sách cho phép, nhiều thí nghiệm hơn sẽ được thực hiện nhằm có kết quả chính xác hơn, hay do ta thay đổi các giả thuyết thống kê liên quan đến hàm  $\phi$ .

### 9.3.2 Điều khiển Chất lượng Thống kê

#### Tại sao là Điều khiển Chất lượng Thống kê- SQC: Statistical Quality Control?

Trong suốt Thế chiến 1, quá trình sản xuất đã trở nên phức tạp hơn, các đốc công đã bắt đầu kiểm soát một số lượng lớn công nhân (trong các dây chuyền sản xuất- *assembly lines*) nhằm đảm bảo chất lượng cho sản phẩm cần sản xuất ra với số lượng lớn. Tiếp cận này cho chất lượng trong sản xuất công nghiệp đã trở nên cấp thiết hơn ngay trước và trong Thế chiến 2, và đã dẫn đến khái niệm điều khiển chất lượng, còn gọi là Điều khiển Chất lượng Thống kê- *Statistical Quality Control*, hay SQC, do công của nhà thống kê *Walter A. Shewhart*, đề xuất từ năm 1924 tại Phòng Thí nghiệm Bell.

#### Khái niệm đảm bảo chất lượng (Quality assurance)

Đảm bảo chất lượng bao gồm mọi hoạt động từ thiết kế, phát triển, sản xuất, cài đặt, vận hành, dịch vụ và bảo trì. Hơn nữa, đảm bảo chất lượng gồm các quy định về chất lượng của các vật liệu thô, các dây chuyền, các sản phẩm và thành phần của sản phẩm; các dịch vụ liên quan đến sản xuất, như các khâu quản lý, sản xuất và kiểm định. Một trong các khung thức được sử dụng rộng khắp cho quản trị chất lượng là tiếp cận PDCA (*Planning, Doing, Checking, Acting*), Hoạch định- Thực thi- Kiểm tra- Hành động hay còn gọi là chu trình Shewhart (**Shewhart cycle**).

Ý tưởng chính của Shewhart (nhằm đảm bảo chất lượng cho sản phẩm công nghiệp) là việc sử dụng các khái niệm lấy mẫu (*sampling*) và sơ đồ điều khiển (*control charts*), nhờ đó các nhà kiểm tra có thể kiểm tra chỉ một phần thay vì 100% sản phẩm tạo ra, mà các kết luận vẫn trung thực. Trung thực một cách thống kê, nghĩa là tới một xác suất rất thấp nào đó bạn là người mua thiếu may mắn nhận được một sản phẩm có sơ sót!

#### Các khoa học gia tiên phong trong Điều khiển Chất lượng

Sau Thế chiến 2, ngoại trừ Hoa Kỳ, khả năng sản xuất của nhiều quốc gia hầu như không còn. Điều này khiến cho nền sản xuất Hoa Kỳ thiếu coi trọng các tiếp cận mới đối với chất lượng: họ vẫn tiếp tục áp dụng các ý niệm kiểm tra (*inspection*) và lấy mẫu, vốn khá thô sơ, nhằm phát hiện và loại bỏ sản phẩm bị lỗi từ dây chuyền sản xuất; dù một số khoa học gia Hoa Kỳ đã nhận thấy sự khiếm khuyết này và đề xuất một vài tiếp cận có tính cách mạng hơn.

Sự phát triển của kỹ thuật chất lượng phần lớn nhờ vào công *W. Edwards Deming* (1900- 1993). Nghịch lý là các quan điểm cách tân của ông, và các nhà tiên phong khác (như Genichi Taguchi, học trò nổi bật nhất của ông) lúc đầu bị thiếu coi trọng ngay tại Hoa Kỳ. Các ý tưởng mới của họ- có tính chất tổng hợp- vào kỹ thuật chất lượng lại được coi trọng và áp dụng ở Nhật Bản, giúp mang lại sự phát triển vượt bậc có tính chất thần kỳ cho nền kinh tế Nhật Bản trong những năm 60-70 của thế kỷ trước. Nhật Bản sau thế chiến 2 buộc phải tìm một cách thức mới phục hưng quốc gia, trong đó các hãng như Toyota nhận thấy không thể bì kịp với các đối thủ Mỹ giàu tiềm lực như General Motors trên các khía cạnh vốn, năng suất và sản lượng. Toyota đã thoát ra tình thế khó khăn ấy, nhờ sáng tạo nên một hệ thống sản xuất dựa trên sự tiết kiệm đa yếu tố như tốc độ và sự uyển chuyển trong các thiết kế, tuyệt đối giảm thiểu hư hại hay thao tác thừa trong các sản phẩm trung gian, theo đó cho phép khử bỏ sự thiếu hiệu quả, rút ngắn chu trình sản xuất... Ngày nay việc tiến hành thí nghiệm trong các công nghiệp hay khảo cứu khoa học nhằm gia tăng sự hiểu biết và tri thức của chúng ta về các tiến trình phức tạp khác nhau, bất luận đó là sản xuất xe hơi hay là một loại dược phẩm mới, xem [14, 6].



### 9.3.3 Thiết Kế Thí Nghiệm (DOE) và Nghiên cứu Khoa học

DOE và SQC, nhìn rộng hơn thuộc về một lĩnh vực gọi là Quản lý Chất lượng Hoàn toàn - *Total Quality Management*; trong đó chúng ta quan tâm đến ba tác vụ: hoạch định chất lượng, điều khiển chất lượng và cải tiến chất lượng. Các nội dung và phương pháp nhằm thực hiện được ba tác vụ này bao gồm:

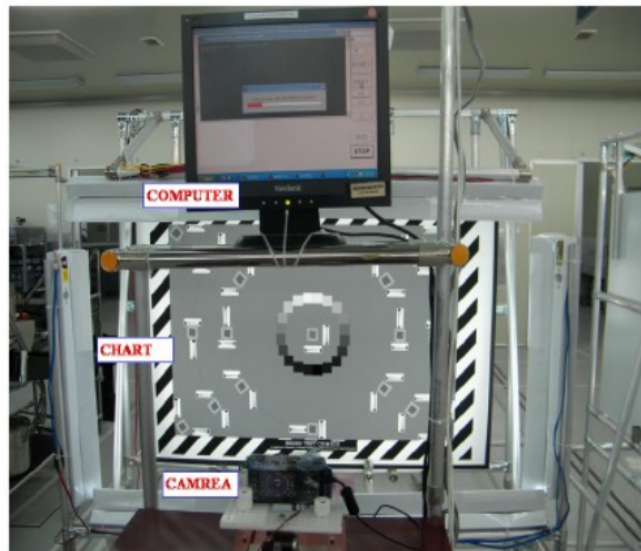
- hoạch định chất lượng: khâu tiên quyết trong sản xuất công nghiệp,
- điều khiển chất lượng (*quality control*): sử dụng các phương pháp DOE,
- cải tiến chất lượng: phải liên tục, dựa trên việc đo lường hiệu quả chính xác, trung thực, và rất coi trọng phản hồi, góp ý của khách hàng... (xem thêm [16, 6]).

#### Hiểu các tương tác chính trong quá trình

Các thí nghiệm trong công nghiệp, sản xuất ... thường được tiến hành qua nhiều phép thử hay kiểm tra nhằm sinh ra các kết quả có thể đo lường được. Đối với cải tiến chất lượng liên tục (*continuous improvement*) của sản phẩm hay quá trình, việc hiểu hành vi hay diễn biến của quá trình, độ lớn của sự sai biệt (*variability*) và ảnh hưởng của sự sai lệch này trên quá trình là điều căn bản. Trong môi trường kỹ thuật, kinh doanh hay khảo cứu khoa học, các thí nghiệm thường được tiến hành nhằm khảo sát, ước lượng một đại lượng hay xác nhận một tính chất nào đó. Sự khảo sát hàm ý việc hiểu dữ liệu (thu) từ quá trình. Sự ước lượng nói về việc xác định các ảnh hưởng của các biến quá trình (*process variables/factors/ parameters*) lên trên đặc tính hay hiệu năng của đầu ra hay sản phẩm. Sự xác nhận liên quan đến kiểm tra các kết quả đã dự đoán từ các thí nghiệm thực sự đã được tiến hành.

Trong các tiến trình này, các tương tác (*factor interactions*) giữa các thành tố (biến quá trình) là mối quan tâm chính của chúng ta, các kỹ sư và nhà quản lý. Ta cần hiểu và phân tích thỏa đáng các bài toán liên quan vấn đề ước lượng và tối ưu hóa các tác động tương tác lên kết quả của quá trình. Trong bài toán tối ưu hóa quá trình này, căn nguyên của vấn đề thường do sự tương tác giữa các thành tố hơn là do tác động riêng lẻ của từng thành tố (*main factor effects*) lên kết quả, hiệu năng của quá trình hay hệ thống. Các phương pháp và tư tưởng của DOE và RSM (*response surface methodology*, Mục 9.5) cung cấp giải pháp cho việc nghiên cứu các tương tác này. Trong phần kế chúng ta xem các khái niệm này được áp dụng ra sao trong sản xuất công nghiệp.

### 9.3.4 Nhu cầu đảm bảo chất lượng trong sản xuất công nghiệp



Hình 9.3: An actual camera tuning system

Trong kỹ thuật chất lượng, **đánh giá độ chính xác của tính năng sản phẩm** của là cần thiết về mặt kinh tế. đánh giá không chính xác có thể sẽ dẫn đến hai nguy cơ lớn:

- a / OK (tốt) sản phẩm được đánh giá là NG (trục trặc): sẽ yêu cầu sửa chữa không cần thiết dẫn đến một sự gia tăng chi phí; và
- b / sản phẩm NG được đánh giá là OK: sẽ dẫn đến việc bồi thường yêu cầu của khách hàng, và kết quả là, chi phí dịch vụ thêm.

Hai loại chẩn đoán sai là không mong muốn, và thường xảy ra ở những vùng chồng lấn (phủ lấp nhau) OK và NG. Do đó, việc thành lập các điều kiện cân bằng cho kết quả chính xác là rất quan trọng trong việc quản lý chất lượng. Bây giờ ta mô tả một ứng dụng thiết kế nhân tố cho vấn đề điều chỉnh máy ảnh kỹ thuật số (*digital camera tuning*), Hình 9.3 cho thấy một hệ thống thực tế. Mục đích là để giữ thông số độ phân giải <sup>1</sup> trong giới hạn tiêu chuẩn nhất định, đồng thời giảm thời gian thực hiện việc điều chỉnh. Chúng ta sẽ sử dụng vài loại thiết kế nhân tố cụ thể ở phần kế tiếp cho bài toán này.

<sup>1</sup>digital resolution, độ phân giải là các dữ liệu nhỏ nhất mà phần mềm máy tính vẫn có thể phân biệt nhau trên biểu đồ ảnh

## 9.4 Thiết kế nhân tố (FD- Factorial Designs)

### 9.4.1 Khái niệm thiết kế nhân tố đầy đủ thuận

Thiết kế nhân tố được dùng để tìm ra một mô hình hồi quy mô tả quan hệ giữa các nhân tố với (mà có ảnh hưởng tới) chất lượng sản phẩm. Giả sử ta có  $d$  tập  $Q_1, Q_2, \dots, Q_d$  lấy từ một trường số (như trường thực  $\mathbb{R}$ ); chúng được gọi là các tập nhân tố- *factor sets* hay chỉ là nhân tố- *factors*.

**Định nghĩa 9.2.** Thiết kế nhân tố (đầy đủ) tạo bởi  $d$  nhân tố  $Q_i$  là tập hợp tích  $D = Q_1 \times \dots \times Q_d \subset \mathbb{R}^d$ . Điểm thiết kế  $\mathbf{p} = (p_1, \dots, p_d)$  là một phần tử của  $D$ . Ngoài ra,  $r_i := |Q_i|$  chính là số lượng mức chọn (*levels*) của nhân tố thứ  $i$ .

**Bảng 9.9:** Các mức nhân tố của thiết kế  $2^3$

Factor	Low (0)	High (1)
Mix(ture) Ratio	45p	55p
Temp(erature)	1000C	1500C
Time period	30min	90min

Trong kỹ nghệ ta thường chọn  $r_i = 2$  hay 3, lúc đó ta có thiết kế nhân tố đầy đủ  $2^d$  hay  $3^d$ , xem thêm các tài liệu [14, 6, 27]. Chọn  $d = 3$  chẳng hạn, ta dùng *thiết kế nhân tố đầy đủ  $2^3$*  theo ba yếu tố nhị phân nhằm tìm ra quan hệ giữa các yếu tố  $x_1$  (tỉ lệ hỗn hợp), yếu tố  $x_2$  (nhiệt độ), yếu tố  $x_3$  (khoảng thời gian thí nghiệm), với hàm độ cứng gỗ  $Y$ . Các mức nhân tố và kết quả số mô phỏng thực nghiệm được cho trong các bảng 9.9 và 9.10. Mô hình hồi quy tuyến tính cho độ cứng gỗ  $Y$ , được tìm ra nhờ 8 thí nghiệm trên có dạng là:

$$Y = f(x_1, x_2, x_3) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3.$$

**Bảng 9.10:** Kết quả thực nghiệm mô phỏng số nhờ thiết kế đầy đủ  $2^3$

Run	Mix Ratio	Temp	Time	Y
1	45p (-)	100C (-)	30m (-)	8
2	55p (+)	100C (-)	30m (-)	9
3	45p (-)	150C (+)	30m (-)	34
4	55p (+)	150C (+)	30m (-)	52
5	45p (-)	100C (-)	90m (+)	16
6	55p (+)	100C (-)	90m (+)	22
7	45p (-)	150C (+)	90m (+)	45
8	55p (+)	150C (+)	90m (+)	56

### 9.4.2 Sử dụng một thiết kế nhân tố đầy đủ $2^4$ trong sản xuất kỹ nghệ

Phần tiếp sau đây minh họa rằng giải quyết vấn đề kỹ nghệ thực tiễn đòi hỏi phối hợp nhiều phương pháp phức tạp, trong đó việc dùng thiết kế nhân tố là cốt yếu. Nghiên cứu về ‘*Resolution Adjustment*

in *Digital Still Camera* được tiến hành tại Sanyo DI, một chi nhánh của Tổng công ty Sanyo tại Thành phố Hồ Chí Minh Việt Nam. Chúng tôi đã đề xuất một phương pháp kết hợp phương pháp luận mặt đáp ứng, tối ưu hóa đa mục tiêu với lý thuyết thiết kế nhân tố nhằm giải quyết một vấn đề Thiết kế Tham số (*Parameter Design*) phức tạp với hơn hai đáp ứng (xem chi tiết trong [30]). Giả sử rằng quá trình điều chỉnh độ phân giải ảnh của camera số trong công nghiệp phụ thuộc vào 4 yếu tố sau:

- Độ nghiêng của ống kính và biểu đồ theo trục dọc (trục tung), gọi là  $A$ , đơn vị  $^{\circ}$ ,
- Độ nghiêng của ống kính và biểu đồ để trục ngang (trục hoành), gọi là  $B$ , đơn vị  $^{\circ}$ .

Chúng ta sử dụng một máy đo góc để đo lường giá trị của các yếu tố  $A$  và  $B$ .

- Độ sáng trên biểu đồ bề mặt (gọi là  $C$ , đơn vị EV, đo bằng máy illuminometer), với

$$EV = \text{Giá trị phơi sáng (Exposure Value or Light Value)} = \log_2 \left[ \frac{F^2}{T} \right],$$

trong đó  $F$  là tốc độ đóng mở đồng tử ống kính (s, *lens iris speed*), và  $T$  là thời gian phơi sáng (s);

- Khoảng cách giữa ống kính và biểu đồ bề mặt ( $D$ , đơn vị cm).

Hai yếu tố không kiểm soát được là: mức rung (chấn động) của sàn nhà xưởng (tên là  $E$ , đơn vị dB), và độ cong vênh của biểu đồ bề mặt (tên là  $F$ , đơn vị  $^{\circ}$ ). Hai nhân tố này đều được coi là (yếu tố) nhiễu.

Theo các tiêu chuẩn công nghiệp của **Sanyo**, có 13 hàm đáp ứng (còn gọi là *điểm kiểm tra*) đáng quan tâm trong toàn bộ quy trình sản xuất (ví dụ như dữ liệu độ phân giải của ảnh, được đo bằng phần mềm máy tính) cho bài toán tối ưu hóa của chúng ta. Ta ký hiệu

1.  $f(A, B, C, D)$  là thời gian điều chỉnh của quá trình,
2.  $Y_i = y_i(A, B, C, D)$  là mô hình hồi quy cho giá trị trung bình của điểm kiểm tra  $i$ , và
3.  $\sigma_i = \sigma_i(A, B, C, D)$  là mô hình hồi quy cho độ lệch chuẩn của điểm kiểm tra  $i$ , với  $i = 1, 2, \dots, 13$ .

Mục tiêu mong muốn của chúng ta là tối ưu hóa đa mục tiêu sau đây:

$$\left\{ \begin{array}{ll} \text{Max } Y_i, & \text{với mỗi } i = 1, 2, \dots, 13, \\ \text{Min } \sigma_i(A, B, C, D), & \text{với mỗi } i = 1, \dots, 13, \text{ và} \\ \text{Min } f(A, B, C, D). \end{array} \right\}$$

Vì chưa có thông tin về các đáp ứng quá trình, lúc khởi đầu các kỹ sư chọn (theo kinh nghiệm) bộ giá trị  $(A, B, C, D) = (0, 0, 10, 95)$  và lặp thử nghiệm 10 lần, các điểm kiểm tra - đáp ứng thứ  $i$  (với  $i = 1, 2, \dots, 13$ ) cho ta kết quả như trong Hình 9.4.(a). Bạn thấy có nhiều vi phạm quy chuẩn công nghiệp, được tô màu đỏ trong hình này phải không?

**Câu hỏi.** Vậy làm sao ta có thể chọn ra bộ giá trị  $(A, B, C, D)$  tối ưu (theo nghĩa giảm thiểu số lượng vi phạm quy chuẩn và thời gian điều chỉnh quá trình  $f$ , mà tối đa hết các kỳ vọng  $Y_i$ , đồng thời tối thiểu hết các sai số  $\sigma_i$ ) để từ đó sản xuất hàng loạt máy ảnh số ra thị trường?



**Thiết kế nhân tố  $2^4$** 

Vì ta không xem xét các ảnh hưởng bậc hai của các yếu tố định lượng, ta sử dụng một thiết kế nhị phân thông thường  $2^4$  bao gồm các yếu tố A, B, C, D trên; với 4 lần lặp cho việc thu thập dữ liệu. Khi xem rằng các tương tác của nhiều hơn hai yếu tố là không đáng kể trong bất kỳ hàm đáp ứng trung bình  $Y$  theo bốn yếu tố  $X_k \in \{A, B, C, D\}$ , mô hình được lựa chọn của ta là

$$Y = \beta_0 + \sum_k \beta_k X_k + \sum_{k < l} \beta_{kl} X_k X_l + e.$$

Chúng ta sử dụng phần mềm thống kê **R** để phân tích dữ liệu đã thu được, cho tất cả các đáp ứng (điểm kiểm tra), từ  $Y_1, \sigma_1$  tới  $Y_{13}, \sigma_{13}$ , và hàm thời gian điều chỉnh  $f$ . Như một minh họa, để xác định điểm kiểm tra  $Y_1$ , vì mục đích chính là xác định xem các ảnh hưởng chính và các ảnh hưởng tương tác hai yếu tố nào là có ý nghĩa, sau khi sàng lọc ta sử dụng mã **R** như sau:

```
LinearModel.1 <- lm(Y1 ~ A+B+C+D+ A*B+ A*C+ A*D+ B*D, data=Design.withresp)
```

Tuy nhiên, xác định mô hình tuyến tính  $Y_i, \sigma_i$  và thời gian điều chỉnh  $f$  theo các nhân tố  $A, B, C, D$  chỉ là bước đầu tiên của giải pháp nhiều bước sau đây:

**Step 1.** Xây dựng các mô hình hồi quy của trung bình của tất cả các đáp ứng định lượng. Để tránh các lỗi hệ thống, ta sử dụng lặp 2 lần thiết kế nguồn  $2^4$  để xây dựng mô hình hồi quy của  $Y_i$ , và 2 lần thiết kế nguồn  $2^4$  để xây dựng mô hình hồi quy của độ lệch chuẩn  $\sigma_i$ .

**Step 2.** Chuyển đổi các hàm mục tiêu chất lượng (mô hình hồi quy của trung bình và mô hình hồi quy của độ lệch chuẩn) đến một phương trình bằng cách sử dụng **chỉ số năng lực** (capability index).

Cụ thể, đối với  $USL_i, LSL_i$ - các cận đặc điểm kỹ thuật trên và dưới của tính năng chất lượng thứ  $i$ , chúng ta tính toán các chỉ số năng lực quá trình  $CP_{k,i}$  (process capability indices, xem công thức 12.17 trong Mục 12.5.4, hay chi tiết hơn trong Chương 8 [22, 27]) theo công thức:

$$CP_{k,i} = \frac{\text{specification width}}{\text{process width}} = \text{Min} \left\{ \frac{USL_i - Y_i}{3\sigma_i}, \frac{Y_i - LSL_i}{3\sigma_i} \right\}$$

**Step 3.** Chuyển đổi chỉ số  $CP_{k,i}$  thành mức sigma nhờ phương trình

$$\text{Sigma Level} = 3CP_k + k\sigma, \text{ with } k = 1.5.$$

**Step 4.** Giải bài toán đa mục tiêu phi tuyến tính để xác định các giá trị của  $A, B, C, D$ , nhờ đó cung cấp các giá trị hàm đáp ứng tối ưu.



	Factor				Respose													
Rep#	A(°)	B(°)	C(EV)	D(cm)	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Y13	Time_Ad
1	0	0	10	95	69.68	55.73	53.42	53.39	55.38	50.12	50.38	50.13	49.14	46.08	45.32	42.75	39.96	244.10
2					75.06	49.79	48.25	49.91	57.22	46.14	51.15	46.07	43.96	40.02	39.35	44.69	43.30	236.35
3					72.91	52.32	52.51	51.78	52.67	44.08	47.13	49.83	48.20	41.58	40.95	40.06	42.16	248.75
4					75.87	56.78	53.89	57.48	55.35	48.35	45.26	44.75	50.51	47.01	46.34	42.80	46.12	240.17
5					68.75	52.92	50.84	48.93	56.23	50.47	50.44	51.90	46.55	43.21	42.61	43.65	38.68	247.35
6					71.08	52.16	49.75	56.59	49.37	44.79	48.73	50.87	44.42	42.36	41.69	36.70	40.33	223.18
7					68.30	54.62	55.22	53.30	55.69	48.08	49.80	43.57	50.92	44.89	44.22	43.08	39.55	238.10
8					74.63	56.29	55.24	53.58	53.04	50.43	49.97	50.38	45.97	46.49	45.82	40.44	44.88	229.78
9					70.74	47.80	49.80	49.98	56.19	52.59	50.55	48.56	49.86	44.08	43.33	43.61	38.93	241.46
10					75.41	52.23	57.23	56.06	51.92	44.67	47.73	50.69	51.94	42.46	41.79	39.39	44.66	226.10
Spec(≥)					70.00	50.00	50.00	50.00	50.00	45.00	45.00	45.00	45.00	40.00	40.00	40.00	40.00	—
ULS					75.87	56.78	57.23	57.48	57.22	52.59	51.15	51.90	51.94	47.01	46.34	44.69	46.12	248.75
LLS					68.30	47.80	48.25	48.93	49.37	44.08	45.26	43.57	43.96	40.02	39.35	36.70	38.68	223.18
Mean					72.24	53.06	52.62	53.10	54.31	47.97	49.11	48.68	48.15	43.82	43.14	41.72	41.86	237.53

(a)

	Factor				Respose													
Rep#	A(°)	B(°)	C(EV)	D(cm)	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Y13	Time Adj
1	2.252	3.878	8.969	94.93	70.08	54.23	53.18	52.65	54.68	49.35	49.31	49.25	48.39	45.38	44.69	42.52	41.35	235.30
2					74.35	53.13	52.24	51.48	55.89	47.35	50.16	47.35	45.87	41.36	42.35	41.89	42.35	238.23
3					75.92	51.18	53.13	50.79	53.84	46.73	48.32	48.97	49.16	42.38	41.09	41.38	43.05	239.45
4					74.70	55.45	49.19	56.34	54.38	47.86	49.35	46.39	48.93	46.04	45.72	39.94	45.38	239.34
5					71.04	51.48	51.09	51.48	54.85	49.39	44.89	50.13	49.35	42.97	43.56	43.28	42.38	239.86
6					70.14	53.28	52.34	54.75	51.78	48.32	49.76	49.89	47.38	43.09	42.63	42.56	41.21	237.22
7					69.87	56.15	54.26	52.38	53.61	47.97	48.65	46.72	50.01	43.76	43.58	41.78	40.56	240.16
8					73.57	53.17	55.78	49.79	51.35	51.06	47.38	51.38	48.39	45.68	44.67	40.58	43.79	236.43
9					71.54	51.75	52.34	52.34	55.49	50.78	49.78	49.45	49.38	44.72	45.89	42.38	43.59	238.75
10					74.31	53.24	55.78	56.06	53.26	48.72	46.35	48.72	50.87	41.89	42.35	41.87	45.08	235.15
Spec(≥)					70.00	50.00	50.00	50.00	50.00	45.00	45.00	45.00	45.00	40.00	40.00	40.00	40.00	---
ULS					75.92	56.15	55.78	56.34	55.89	51.06	50.16	51.38	50.87	46.04	45.89	43.28	45.38	240.16
LLS					69.87	51.18	49.19	49.79	51.35	46.73	44.89	46.39	45.87	41.36	41.09	39.94	40.56	235.15
Mean					72.55	53.31	52.93	52.81	53.91	48.75	48.40	48.83	48.77	43.73	43.65	41.82	42.87	237.99

(b)

Hình 9.4: The actual data set before and after optimizing

Kết luận, như bạn thấy qua số liệu cho trong Hình 9.4 rằng sau khi tiến hành thí nghiệm thì ta định vị được tọa độ tối ưu của 4 yếu tố là  $(A, B, C, D) = (2.252, 3.878, 8.969, 94.83)$ . Đồng thời sự kiện ít số liệu màu đỏ hơn trong Hình 9.4.(b) nói lên rằng các bộ giá trị  $(A, B, C, D)$  chọn xung quanh điểm tối ưu cho các đáp ứng ít vi phạm ràng buộc kỹ thuật (specifications) hơn.

### 9.4.3 Thiết kế nhân tố một phần thuần

Giả sử ta gộp các  $r_i$  trong Định nghĩa 9.2 thành dãy giảm dần  $s_1 > s_2 > \dots > s_m$  ( $m \leq d$ ) các số lượng mức chọn phân biệt (khác nhau) của các nhân tố tạo ra  $D$ , và giả sử thêm trong  $D$  có đúng  $a_i$  nhân tố mà có  $s_i$  mức; vậy  $a_1 + a_2 + \dots + a_m = d$ . Ta gọi  $s_1^{a_1} \cdot s_2^{a_2} \cdot \dots \cdot s_m^{a_m}$  là kiểu của thiết kế nhân tố (hỗn hợp, mixed factorial design)  $D$ . Ví dụ như thiết kế kiểu  $4 \cdot 2^3$  là  $D = \{(0, 0, 0, 0), (0, 0, 0, 1), \dots, (3, 1, 1, 1)\}$ .

**Câu hỏi.** Liệu chúng ta có thể ước đoán hàm  $Y$  dựa trên một tập con  $F$  của thiết kế đầy đủ  $D$  để tiết giảm ngân sách làm thực nghiệm không?

**Định nghĩa 9.3.** Giả sử  $D$  là thiết kế nhân tố hỗn hợp kiểu  $s_1^{a_1} \cdot s_2^{a_2} \cdot \dots \cdot s_m^{a_m}$ . Một thiết kế nhân tố một phần của  $D$  là một (đa) tập con  $F$  của  $D$ , nghĩa là cho phép một phần tử (thí nghiệm- run) xuất hiện nhiều lần trong  $F$ , ta gọi là bội (multiplicity) của phần tử ấy.

Ta gọi  $F$  là thiết kế nhân tố một phần có kiểu  $s_1^{a_1} \cdot s_2^{a_2} \cdot \dots \cdot s_m^{a_m}$ . Khi  $m = 1$  ta có  $F$  là thiết kế nhân tố một phần thuần kiểu  $s_1^{a_1}$ . Khi  $m > 1$  ta có  $F$  là thiết kế nhân tố một phần hỗn hợp kiểu  $s_1^{a_1} \cdot s_2^{a_2} \cdot \dots \cdot s_m^{a_m}$ .

**Ví dụ 9.3.** Một thiết kế nhân tố một phần  $H$  kiểu  $2^3$  có 4 phần tử, các cột  $H$  ứng với 3 yếu tố nhị phân của thiết kế nhân tố đầy đủ  $2^3$  gồm 8 thí nghiệm nêu trong Bảng 9.10 ở trên.

Thiết kế một phần  $H$  chính là ma trận Hadamard vuông cấp 4 sau (ma trận vuông chứa chỉ 1 và -1, trong đó mọi cặp dòng hay mọi cặp cột có tích vô hướng là 0). Ma trận Hadamard này [ghi tắt ma trận  $H$  từ đây] cung cấp đủ thông tin giúp nhà nghiên cứu ước lượng hoàn toàn các ảnh hưởng chính (*main effects*) của các yếu tố tỉ lệ hỗn hợp  $x_1$ , nhiệt độ  $x_2$  và khoảng thời gian thử nghiệm  $x_3$  (lên hàm độ cứng gỗ  $Y$ ) trong Ví dụ 1 nêu trên.

$$H = \begin{bmatrix} 1 & x_1 & x_2 & x_3 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}^T.$$

Khi  $m = 2 > 1$ , chẳng hạn ma trận sau đây là một thiết kế nhân tố một phần hỗn hợp kiểu  $4 \cdot 2^3$ , ký hiệu  $T$  là ký hiệu lấy ma trận chuyển vị.

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 3 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}^T$$

**Tại sao ta cần sử dụng FFD?** Nhằm cắt giảm chi phí làm thực nghiệm trong nghiên cứu khoa học hay/ và trong sản xuất công nghiệp. Ở ví dụ ma trận Hadamard, nếu ta dùng 4 thí nghiệm của  $H$  (thay vì dùng hết 8 thí nghiệm của  $2^3$ ) thì có thể xác định mô hình tuyến tính cho độ cứng gỗ  $y$ , rút gọn hơn  $Y$ , có dạng sau:

$$y = f(x_1, x_2, x_3) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

Nghĩa là, dùng ít nguồn lực hơn thì mô hình  $y$  chỉ cho phép ước lượng các ảnh hưởng chính thôi!

**Quan sát 9.1.** Bạn thử kiểm tra rằng:

Nếu một ma trận Hadamard cấp  $n > 2$  tồn tại thì  $n$  phải chia hết cho 4?

### Giải thuyết Hadamard

Liệu chiều đảo của quan sát trên đúng không? Đó chính là vấn đề mà Hadamard, nhà toán học lớn người Pháp trong thế kỷ 19 đã nêu, hiện vẫn là bài toán mở, thách thức giới toán học ở thế kỷ 21 này, rằng: Với bất kỳ số tự nhiên  $n$  mà chia hết cho 4, liệu tồn tại một Hadamard cỡ  $n$  hay không?

Với hầu hết các  $n$  (mà chia hết cho 4) luôn có ma trận  $H$  cỡ  $n$  được tìm ra bằng vài chục phương pháp toán học khác nhau dựa trên lý thuyết số, tổ hợp, mật mã, trường hữu hạn, đại số tuyến tính ... trong hơn 100 năm qua, nhưng một số trường hợp thực sự cứng đầu! Thực vậy, ma trận  $H$  cỡ 428 chỉ được tìm ra vào năm 2004 nhờ nỗ lực chung hơn 10 năm ròng rã của một nhóm nhà toán học và khoa học máy tính từ Canada và Iran, lãnh đạo bởi các GS Hadi Kharaghani và Behruz Tayfeh-Rezaie, sử dụng ý niệm toán gọi là *Turyn sequences*. Trước đó 19 năm, Hadamard cỡ 268 được tìm ra do công Kazue Sawade, nhà toán học Nhật bản. Và cho tới nay, 2017 câu hỏi **ma trận  $H$  cỡ 668 tồn tại hay không** vẫn còn là một thách thức cho chúng ta!

### 9.4.4 Thiết kế nhân tố một phần hỗn hợp - Sản xuất đồ nội thất gỗ?

**Bảng 9.11:** Eight factors, the number of levels and the level meanings

Factor	Description	#	Level					
			0	1	2	3	4	5
1 (A)	wood	6	pine	oak	birch	chestnut	poplar	walnut
2 (B)	glue	4	a (less adhesive)	b	c	d (most adhesive)		
3 (C)	moisture content	4	10%	20%	30%	40%		
4 (D)	processing time	2	1 h(our)	2h				
5 (E)	pretreatment	2	no	yes				
6 (F)	indenting of wood samples	2	no	yes				
7 (G)	pressure	2	1 pas(cal)	10 pas				
8 (H)	hardening conditions	2	no	yes				

Ta muốn nghiên cứu một đáp ứng  $Y$ , là độ cứng đồ nội thất gỗ. Để tối đa hóa độ cứng của sản phẩm mới, ta nghiên cứu ảnh hưởng kết hợp của các yếu tố sử dụng mô hình hồi quy tuyến tính. Trong nhiều tình huống thực tiễn ta có thể cần các yếu tố có số mức chọn khác nhau, như ta thấy có 8 yếu tố có số mức chọn theo thứ tự là 6, 4, 4, 2, 2, 2, 2, 2 trong Bảng 9.11 ở trên.

- Gọi  $N$  là số thí nghiệm trong thiết kế thực nghiệm; mỗi thí nghiệm sẽ được gán cho một kết hợp của các mức yếu tố.
- Gọi  $M := 6 \cdot 4^2 \cdot 2^5 = ?$  biểu thị số lượng kết hợp mức độ có thể có của các yếu tố  $A, B, C, D, E, F, G$  và  $H$ .
- Do ràng buộc ngân sách trong khảo sát thực nghiệm hay sản xuất ta cần một thiết kế thực nghiệm mà có  $N \leq M$ , mà vẫn cho phép ta nghiên cứu một số ảnh hưởng chính nhân tố (*main effects*) quan trọng, vài tương tác kép nhân tố (*two-factor interactions*) có ý nghĩa.

**The RAO Bound on the run-size of factorial designs**

Nếu kể hết ảnh hưởng chính nhân tố và tương tác kép (hai) nhân tố thì tổng số tham số (nhiều, ảnh hưởng chính và tương tác hai nhân tố) theo Rao là

$$1 + \sum_{i=1}^8 (s_i - 1) + \sum_{i,j=1, i < j}^8 (s_i - 1)(s_j - 1) = 5 + 3 + 3 + 5 \times 1 + 2 \times 5 \times 3 + 5 \times 5 \times 1 + 3 \times 3 + 2 \times 3 \times 5 + 10 = 120.$$

Nếu các tương tác được coi là không quan trọng, mô hình hồi quy tuyến tính chỉ chứa các tác dụng chính (tất cả là ?) có dạng dưới đây

$$Y = \theta_0 + \sum_{i=1}^5 \theta_{A_i} a^i + \sum_{j=1}^3 \theta_{B_j} b^j + \sum_{l=1}^3 \theta_{C_l} c^l + \theta_D d + \theta_E e + \dots + \theta_H h + \epsilon, \quad (9.31)$$

for  $a = 0, 1, 2, 3, 4, 5; b, c = 0, 1, 2, 3; d, e, f, g, h = 0 \text{ or } 1,$

tham số  $\theta_*$  là các hệ số hồi quy (*regression coefficients*),  $\epsilon$  là sai số ngẫu nhiên.

Nếu cần khảo cứu hết ảnh hưởng chính nhân tố và thêm các tương tác kép 2 nhân tố, theo kết quả của Rao trên kia thì ta cần có một thiết kế thực nghiệm mà có  $N = 120$  thí nghiệm, nhằm 'khớp' mô hình hồi quy tuyến tính

$$Y_1 = Y + \sum_{i=1}^5 \sum_{j=1}^3 \theta_{A_i B_j} a^i b^j; \sum_{j=1}^3 \sum_{l=1}^3 \theta_{B_j C_l} b^j c^l + \dots \quad (9.32)$$

Một thiết kế có  $N = 120$  thí nghiệm tồn tại hay không là một vấn đề rất mới, một lời giải đại số - có độ phức tạp cỡ mũ kép- có thể tìm thấy trong [20] và các tham khảo ở đó. Nhưng nếu các nhà công nghiệp chịu bỏ qua (không khảo cứu) vài tương tác kép 2 nhân tố và hài lòng với thiết kế có  $N = 96$  thí nghiệm thì ta có thể chỉ ra một thiết kế như vậy, theo đó họ có thể thử nghiệm để phân tích dữ liệu rồi sản xuất đại trà, xem thêm Mục 9.7 và tài liệu [21].

## 9.5 Khảo sát các mặt đáp ứng

### 9.5.1 Giới thiệu

Ta đã thấy trong các phần trên rằng vấn đề tối ưu hóa hàm đáp ứng là rất quan trọng trong **Cải tiến Quá trình** (*Process Improvement*). Thông thường cải tiến quá trình có liên quan đến việc tối ưu hóa của một hàm đáp ứng hoặc đặc tính chất lượng duy nhất, mà thường là quan trọng nhất đối với người tiêu dùng. Tuy nhiên, hầu hết các sản phẩm hay đặc trưng chất lượng là đa chiều, vì vậy người ta thường quan sát nhiều đáp ứng trong một tình huống thực nghiệm.

Phương pháp luận mặt đáp ứng (*Response Surface Methodology* - RSM) là kỹ thuật tối ưu hóa đa biến dựa trên các thiết kế thống kê nêu trên. Được giới thiệu từ 1950s bởi George Box và Wilson, khởi đầu sử dụng trong kỹ nghệ hóa học, ngày nay RSM là tập hợp phong phú các phương pháp toán học và thống kê hữu ích cho mô hình hóa và phân tích các đáp ứng của một quá trình bị ảnh hưởng bởi đa tác nhân. GS George Box, năm 1987 đã viết rằng "*essentially, all models are wrong, but some are useful*"

trong quyển ‘Empirical Model-Building and Response Surfaces’, p. 424, Wiley; nói lên tầm quan trọng, sự uyển chuyển và hữu ích của mô hình hóa thống kê (khi mô tả thế giới so với mô hình hóa toán học).

Từ quan điểm của các nhà công nghiệp quá trình, các phương pháp này giúp họ rút giảm sự biến thiên bất định (*variability*) của quá trình bằng cách chọn được các mức phù hợp của một số yếu tố điều khiển được, theo đó cực tiểu hóa sự biến thiên của hàm đáp ứng- kết xuất đang quan tâm, khi cho phép các yếu tố nhiễu (khó hay không điều khiển được) thay đổi tự do (nhưng không có ảnh hưởng lên kết xuất) [16].

### Vậy RSM là gì, một cách hình thức?

Đó là sự tối ưu hóa tựa xấp xỉ (*approximation-based optimization*), khi mà hàm xấp xỉ được gọi là mặt đáp ứng (*response surface*). Tổng quát, giả sử nhà khoa học hay kỹ sư (gọi chung là nhà thực nghiệm- *experimenter*) quan tâm đến một sản phẩm, quá trình hay hệ thống. Sản phẩm được đặc tả bởi hàm đáp ứng (chất lượng)  $Y$  mà phụ thuộc vào một số hữu hạn đại lượng - biến- yếu tố thử nghiệm điều khiển được (*controllable independent variables*).

Một cách toán học, mối quan hệ hàm giữa biến đáp ứng  $Y$  và các biến thử nghiệm  $\mathbf{x} = (x_1, x_2, \dots, x_k)$  với  $0 < k \in \mathbb{N}$  được biểu diễn bởi

$$Y = f(x_1, \dots, x_k) + e,$$

trong đó  $e$  là một biến ngẫu nhiên với trung bình không và phương sai hữu hạn  $\sigma^2$ .

**Khái niệm 9.1.** Tập các điểm  $\{f(x_1, \dots, x_k), x_i \in D_i, i = 1, \dots, k\}$ , trong đó  $D_1, \dots, D_k$  là miền thử nghiệm của các biến  $x_i$ , được gọi là một **mặt đáp ứng** (*response surface*). Ta gọi tập hợp tích  $D = D_1 \times \dots \times D_k$  là miền thử nghiệm của  $\mathbf{x} = (x_1, x_2, \dots, x_k)$ .

Hai loại mặt đáp ứng đã được thảo luận trước đây, gồm mặt **tuyến tính**

$$f(x_1, \dots, x_k) = \beta_0 + \sum_{i=1}^k \beta_i x_i \quad (9.33)$$

và mặt **bậc hai**

$$f(x_1, \dots, x_k) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j. \quad (9.34)$$

Các mặt đáp ứng có thể có dạng hàm phức tạp. Giả định rằng trong các miền địa phương ta đang quan tâm, các mặt đáp ứng có thể được xấp xỉ bởi các mô hình tuyến tính hoặc bậc hai.

Các nhà nghiên cứu quan tâm đến việc khảo sát hay khám phá bản chất của các mặt phản ứng, trong các miền quan tâm nhất định, nhằm mục tiêu dự đoán suất lượng  $Y$  trong tương lai, và đặc biệt để tối ưu hóa một quá trình, bằng cách chọn các giá trị  $\mathbf{x}$  nhằm tối đa hóa (hay là tối thiểu hóa) suất lượng trung bình (hoặc là tổn thất trung bình)  $E[Y]$ . Trong phần này ta trình bày các thiết kế đặc biệt dành cho việc thăm dò, khảo sát các mặt bậc hai, và dùng để xác định các miền (hay điều kiện) tối ưu. Các thiết kế cho các mô hình bậc hai được gọi là **thiết kế bậc hai** (*second order designs*). Ta hãy bắt đầu với lý thuyết về các thiết kế bậc hai và kết thúc với việc tối ưu hóa quá trình.

## Tài liệu tham khảo

---

- [1] Ali Jahan & Md Yusof Ismail and Rasool Noorossana *Multi Response optimization in DOE considering capability index in bounded objectives method*, Journal of Scientific and Industrial Research - India, **69** (2010), 11–16.
- [2] Antal Kozak, Robert A. Kozak, Christina L. Staudhammer, Susan B. Watts *Introductory Probability and Statistics Applications for Forestry and Natural Sciences*, CAB International (2008)
- [3] *Tranh sơn dầu được vẽ bởi các họa sĩ dân tộc thiểu số Úc*, Bảo tàng cư dân bản địa Úc châu.
- [4] Adrian Dumitrescu *Efficient Algorithms for Generation of Combinatorial Covering Suites* ISAAC 2003, LNCS 2906, pp. 300–308, 2003.
- [5] John J. Borkowski's Home Page, [www.math.montana.edu/jobocourses.html/](http://www.math.montana.edu/jobocourses.html/)
- [6] C.F. Jeff Wu, Michael Hamada *Experiments: Planning, Analysis and Parameter Design Optimization*, Wiley, 2000.
- [7] David M. Cohen, Siddhartha R. Dalal, Michael L. Fredman, and Gardner C. Patton *The AETG System: An Approach to Testing Based on Combinatorial Design*, IEEE Trans. on Soft. Engineering, Vol. 23, No. 7, July 1997.
- [8] S.R. Dalai and al., *Factor-covering designs for Testing Software*, *Technometrics* 40(3), 234-243, American Statistical Association and the American Society for Quality, 1998.
- [9] Douglas C. Montgomery, George C. Runger *Applied Statistics and Probability for Engineers*, Sixth Edition, (2014) John Wiley & Sons
- [10] M. F. Fecko and al., *Combinatorial designs in Multiple faults localization for Battlefield networks*, *IEEE Military Communications Conf.*, Vienna, 2001.
- [11] Glonek G.F.V. and Solomon P.J. *Factorial and time course designs for cDNA microarray experiments*, *Biostatistics* **5**, 89-111, 2004.
- [12] Hedayat, A. S., Sloane, N. J. A. and Stufken, J. *Orthogonal Arrays*, Springer-Verlag, 1999.
- [13] Harry M. Kaiser, Kent D. Messer. *Mathematical programming for agricultural, environmental, and resource economics*, (2011) John Wiley & Sons
- [14] Madhav, S. P., *Quality Engineering using robust design*, Prentice Hall, 1989.

- [15] Michael Baron, *Probability and Statistics for Computer Scientists*, 2nd Edition (2014), CRC Press, Taylor & Francis Group
- [16] R. H. Myers, Douglas C. Montgomery and Christine M. Anderson-Cook *Response Surface Methodology : Process and Product Optimization Using Designed Experiments*, Wiley, 2009.
- [17] Nguyen, V. M. Man, *An online service for computing mixed orthogonal arrays*, University of Technology, VNUHCM link [elearning.cse.hcmut.edu.vn/samgroup/OA.jsp](http://elearning.cse.hcmut.edu.vn/samgroup/OA.jsp).
- [18] Nguyễn Văn Minh Mẫn, *Thống Kê Bayes trong Kỹ thuật - Ứng dụng cho Điều khiển Quá trình và Phân tích Độ Tin cậy Hệ thống ĐHQG TPHCM*, (chuẩn bị) 2018
- [19] Nguyễn Văn Minh Mẫn, *Computer-Algebraic Methods for the Construction of Designs of Experiments*, Ph.D. thesis, 2005 , <http://www.mathdox.org/nguyen>
- [20] Man Nguyen and Scott H. Murray,  
*Algebraic Methods for Construction of Mixed Orthogonal Arrays*,  
Southeast Asian Journal of Sciences, Vol 1, No. 2 (2012) pp. 155-168; ISSN 2286-7724,  
(at [science.utcc.ac.th/sajs/wp-content/uploads/2013/06/3-MinhMan.pdf](http://science.utcc.ac.th/sajs/wp-content/uploads/2013/06/3-MinhMan.pdf))
- [21] Nguyen, Man V. M. *Some New Constructions of strength 3 Orthogonal Arrays*,  
the Memphis 2005 Design Conference Special Issue of the  
**Journal of Statistical Planning and Inference**, Vol 138, Issue 1 (Jan 2008) pp. 220-233.
- [22] Ron S. Kenett, Shelemyahu Zacks. *Thống kê Công nghiệp hiện đại với các ứng dụng viết trên R và MINITAB*, Nguyễn Văn Minh Mẫn dịch, VIASM, (2014), Nhà XB ĐH Bách Khoa Hà Nội.
- [23] Natarajan Gautam. *Operations Research and Management Science Handbook*, (2008), CRC Press
- [24] Nathabandu T. Kottegoda, Renzo Rosso. *Applied Statistics for Civil and Environmental Engineers*, 2nd edition (2008), Blackwell Publishing Ltd and The McGraw-Hill Inc
- [25] Natali Hritonenko, Yuri Yatsenko. *Mathematical Modeling in Economics, Ecology and the Environment*, 2nd edition (2013), Springer Optimization and Its Applications
- [26] Paul Mac Berthouex. L. C. Brown. *Statistics for Environmental Engineers*; 2nd edition (2002), CRC Press
- [27] Ron S. Kenett, Shelemyahu Zacks, *Modern Industrial Statistics with applications in R, MINITAB*, 2nd edition, (2014), Wiley
- [28] Sathees Kumar et al. *Land use change modeling using a Markov model and remote sensing*, **Journal of Geomatics, Natural Hazards and Risk**, Vol. 5, No. 2, 145–156, (2014)  
<http://dx.doi.org/10.1080/19475705.2013.795502>
- [29] Sloane N.J.A., <http://neilsloane.com/hadamard/index.html/>
- [30] Vo Ngoc Thien An, *Design of Experiment for Statistical Quality Control*, Master thesis, LHU, Vietnam (2011)



# Chỉ mục

---

- ý nghĩa thống kê, 145
- ý nghĩa thiết thực - công nghệ, 145
- Định lý
  - Chebyshev, 62
  - giới hạn trung tâm, 66
  - Pearson, 161
- độ chính xác, 80
- độ dốc, 181, 339
- độ lệch, 181
- độ lệch chuẩn, 62
  - $\alpha$  cắt, 104
  - mẫu, 93
- độ méo lệch, 62, 64
- độ méo lệch mẫu, 94
- độ nhọn, 62, 339
- độ nhọn mẫu, 94
- độ phân tán, 80
- độ phù hợp mô hình với dữ liệu, 118
- độc lập, 14
  - lẫn nhau, 15
  - theo cặp, 15
- đại lượng tổng thể, 304
- đưa vào dần, 199
- đa hồi quy, 193
- điểm
  - bất thường, 99
  - ngoại lai, 103
- ảnh hưởng
  - tương tác, 246
  - của các tương tác, 268
  - chính, 246, 250, 255
  - tương tác/phối hợp, 255
- ước lượng
  - điểm của tham số tổng thể, 120
  - không chệch, 120, 313
  - không thiên vị, 120
  - khoảng của tham số tổng thể, 120
  - nhất quán, 120, 313
- ước lượng
  - khả năng quan sát dữ liệu tối đa, 122
- tỉ lệ  $F$ , 215, 248
- bất đối xứng, 64
- Bất đẳng thức
  - Schwarz, 175
  - Schwarz, 190
- bất đẳng thức Chebyshev, 127
- bất thường, ngoại lai, 98
- bài toán lựa chọn mô hình, 199
- bản sao
  - độc lập, 254
- bảng
  - tiếp liên, 222
  - ANOVA, 218
- bảng phân phối xác suất, 20
- biến
  - ngẫu nhiên, 18, 337
  - đáp ứng, 194
  - định lượng, 202
  - định tính, 202
  - dự báo, 194
  - Gauss chuẩn tắc, 49
  - giải thích, 168, 194
  - hồi quy (giải thích hay dự báo), 194
  - mã hóa, 206
  - phân loại, 202
  - quan sát, 168
  - biến đáp ứng, 244



- biến đổi
  - trung bình động, 328
- biến cố, 9
- biến cố
  - chắc chắn, 9
  - không thể, 9
- biến cố rời nhau từng đôi, 11
- biến liên tục, 82
- biến ngẫu nhiên
  - liên tục, 21
  - nhị thức, 35
  - rời rạc, 34
  - chi-bình phương với d.f.  $n - 1$ , 126
  - liên tục, 19, 337
  - rời rạc, 19, 337
- biến rời rạc, 82
- biểu đồ
  - đa phân tán, 203
  - cột, 82
  - hộp, 98
  - hộp đa biến, 205
  - hộp-và-râu, 98
  - phân bố, 202
  - phân tán, 202
  - tán xạ, 202
  - tán xạ 3D, 204
  - tần số, 89
  - tần suất, 89
- cấp độ, mức chọn của nhân tố được kiểm soát, 244
- cách bố trí thí nghiệm, 244
- cách xử lý tổ hợp, 263
- cận dung sai
  - dưới, 138
- cỡ mẫu, 118
- chất lượng thông tin - InfoQ, 343
- Chỉ số
  - Cramer, 228
  - tiếp liên trung bình bình phương, 228
  - Tschuprow, 228
- chuẩn xác, 80
- dạng chuẩn hóa, 52
- danh sách, 306
- giá trị
  - bất thường, 103
  - bất thường, bất thường, 96
  - dự đoán, 195
- giả thuyết
  - gốc, hay vô hiệu  $H_0$ , 311
  - thống kê - statistical hypothesis, 143
- giả thuyết
  - gốc, đầu tiên, hay vô hiệu  $H_0$ , 144
- giả thuyết
  - đôi (thay thế) của  $H_0$ , 144
- giải pháp, cách xử lý, 250
- giải pháp/cách xử lý tổ hợp, 254
- giải thích được, 195
- giao hoán, 10
- giao thức thực hiện, 244
- hồi quy
  - đa biến, 193
  - logistic, 200
- hàm
  - mật độ xác suất, 22
  - Gauss mở rộng, 47
  - Gauss tổng quát, 47
  - khả năng (cơ may) quan sát dữ liệu, 122
- hàm
  - sinh moment, 35, 340
- hàm giá trị thực, 18, 337
- hàm mật độ
  - xác suất, 21
  - xác suất, 19
- hàm phân phối
  - xác suất, 20, 21
  - Gauss chuẩn tắc, 49
- hàm số
  - Gauss, 49
- hàm sinh moment (m.g.f.), 48
- hệ số
  - đa tương quan, 195
  - tương quan đa biến bình phương, 195
  - xác định, 183
- hệ số hồi quy, 181
- hệ số biến thiên, 96
- hệ số góc, 181

- hệ số tự do, 181
- hệ số tương quan mẫu, 176, 190
- hệ thống thông minh, 62
- hiệp phương sai mẫu, 189
- hiệu ứng
  - của các tương tác, 268
- hiệu quả thống kê, 343
- kết hợp, 10
- kích thước mẫu, 118
- Kỹ thuật về tính bất định, 61
- không bị lẫn lộn, 247
- không chuẩn xác, 80
- không liên kết, 227
- khả năng, năng lực của quá trình, 301
- khoảng
  - dung sai, 98
- khoảng tứ phân vị, 96
- Khoa học hành vi, 61
- khung dữ liệu, 85
- kiểm định
  - giả thuyết thống kê, 114
- kiểm định giả thuyết- *hypothesis testing*, 145
- kiểm tra - giám sát
  - trong quá trình, 301
- lắp ráp môi hàn, 245
- lấy mẫu
  - có hoàn lại, 305
  - không hoàn lại, 305
- lấy mẫu ngẫu nhiên có hoàn lại, 73
- lấy mẫu ngẫu nhiên không hoàn lại, 73
- lý thuyết xếp hàng, 58
- Lagrangian, 347
- log-khả năng, 122
- Luật yếu số lớn, 65
- mức ý nghĩa (thống kê) của kiểm định, 146
- mức độ tin cậy của khoảng, 128, 332
- mô hình
  - cộng tính, 255
  - kết cục hữu hạn, 200
  - không cộng tính, 255
  - thống kê, 118, 244
  - thống kê tham số, 121
  - tuyến tính, 181
- mặt đáp ứng, 285
  - bậc hai, 285
  - tuyến tính, 285
- mục tiêu
  - làm thí nghiệm, 244
  - ngiên cứu, 244
- mẫu, 73
- mẫu ngẫu nhiên, 118
  - có hoàn lại, RSWR, 37
  - không hoàn lại, 305
- mẫu ngẫu nhiên đơn giản, 305
- mẫu ngẫu nhiên phân theo tầng, 305
- ma trận
  - đa phân tán, 203
  - phân tán đa biến, 203
- ma trận phương sai-hiệp phương sai, 231
- miền bác bỏ, 146
- miền chấp nhận- *acceptance region*, 146
- moment, 338
- moments trung tâm, 338
- ngẫu nhiên có hoàn lại, 306
- ngẫu nhiên hóa, 247
- ngẫu nhiên không hoàn lại, 306
- nguyên lý
  - bình phương tối thiểu, 194
- nhân tố
  - (yếu tố) được kiểm soát, 244
- nhân tử tổng thể hữu hạn, 316
- nhân tử Lagrange, 347
- nhu cầu oxy sinh hóa - biochemical oxygen demand, 179
- phép đo, giá trị đo, 244
- phép thử, 7
- phân bố
  - tần số, 82
- phân bố tối ưu, 347
- phân hoạch, phân vùng, 10
- phân phối, 10
  - $F$  (Fisher) với  $\nu_1$  và  $\nu_2$  bậc tự do., 215, 248
  - $\chi^2$ , 125
  - đối xứng, 63
  - chuẩn, 47

- Gauss, 47
- lấy mẫu của một ước tính, 307
- nhị thức, 35, 38
- Poisson, 39
- tần số, 82
- tham khảo, 257
- xác suất, 20
- Gauss chuẩn tắc, 49
- mũ, 56
- nhị thức, 64
- Student hay  $t$  với bậc tự do  $\nu$ , 59
- phân phối Cauchy, 338
- phân phối chi bình phương, 126
- Phân tích
  - ANOVA, 265, 267
- phân tích
  - hồi quy, 173
  - hồi quy đa biến, 194
  - tương quan, 173
- phân tích phương sai, 217, 230
- phân vùng, 16
- phân vị, 215, 249
  - của phân phối mũ, 56
  - mẫu thứ  $p$ , 90
  - thứ  $p$  của phân phối, 62
  - thứ  $p$  của một biến ngẫu nhiên  $X$ , 37
- phân vị của phân phối Gauss chuẩn, 51
- phân vị,
  - bách phân vị của phân phối, 37
- phụ thuộc, 14
- phương sai mẫu, 119
- phương pháp
  - bình phương tối thiểu, 181
  - moment, 124
- phương pháp đa biến, 168
- phương pháp hồi quy tuyến tính, 168
- phương sai, 338
- phương sai mẫu, 93
- phương trình
  - chuẩn tắc - *normal equation*, 348
  - hồi quy, 181
- quá trình
  - hàn sóng, 245
- quần thể thống kê, 73, 117
- quan sát
  - bất thường, 102
- ràng buộc tương phản tuyến tính, 259
- RSWOR phân tầng, 319
- số trung vị, 63
- sự phân bố mẫu, 344
- sơ đồ phân vị, 100
- sơ đồ tương tác nhân tố, 273
- sai lầm (lỗi suy diễn- inference error), 145
- sai số, 181
  - chuẩn, 120
  - do dụng cụ đo, 81
  - do người đo, 81
- so sánh
  - cặp, 247
- suy diễn
  - thống kê, 98, 118
- tính bền vững, 102
- tính chất quên, 58
- tối đa mẫu, 90
- tối thiểu mẫu, 90
- tứ phân vị, 90
  - mẫu thứ ba, 91
  - mẫu thứ nhất, 91
- tác động
  - của các tương tác, 268
- tầm vực mẫu, 90
- tần số, 82
  - tích lũy, 82
  - tương đối, 82
  - tương đối tích lũy, 83
- tần suất, 82
  - có điều kiện, 227
  - tích lũy, 83
- tẩy rửa làm sạch, 245
- tổ hợp xử lý, 216
- tổng thể thống kê, 73, 117
- thí nghiệm, 7, 8
  - được thiết kế một cách thống kê, 239
  - nhân tố, 239
  - nhân tố đầy đủ, 268

- thống kê, 90
  - đủ, 123
  - bền vững, 103
  - chi - bình phương, 227
  - cho xu hướng trung tâm, 90
  - mẫu, 118
  - vị trí, 90
  - thứ *i*, 90
  - thứ tự, 90
- thứ tự ngẫu nhiên, 244
- thời gian
  - giữa 2 lần đến liên tiếp, 58
  - phục vụ, 58
- tham số, 19, 338
- thiết kế
  - không kỳ dị, 287
  - bậc hai, 285
  - nhân tố đầy đủ, 268
  - tổ hợp, 264
  - tổng hợp có trung tâm, 289
  - thí nghiệm (DOE), 239
  - trực giao, 287
  - xoay được, 288
- thiết kế khối
  - đầy đủ ngẫu nhiên, 248
  - hoàn chỉnh ngẫu nhiên, 258
  - không đầy đủ cân bằng, 248
  - không đầy đủ và cân bằng, 263
- thuật toán hồi quy từng bước, 199
- trợ dung, thổi sạch bề mặt, 245
- trung bình
  - hai mẫu độc lập, 215, 248
  - $\alpha$ -cắt, 102
  - hình học, 96
  - mẫu, 119
  - tổng thể, 338
  - mẫu, 92
- trung vị, 90
- vô hạn đếm được, 19
- xác suất
  - có điều kiện, 14
  - hậu nghiệm, 16
  - tiên nghiệm, 16
- tiên nghiệm, 16

