Statistical Design and Analysis of Experiments
Part One

Lecture notes

Fall semester 2007

Henrik Spliid

Informatics and Mathematical Modelling

Technical University of Denmark

Foreword

The present collection af lecture notes is intended for use in the courses given by the author about the design and analysis of experiments. Please respect that the material is copyright protected.

The material relates to the textbook: D.C. Montgomery, Statistical Design and Analysis, 6th ed., Wiley.

The notes have been prepared as a supplement to the textbook and they are primarily intended to present the material in both a much shorter and more precise and detailed form. Therefore long explanations and the like are generally left out. For the same reason the notes are not suited as stand alone texts, but should be used in parallel with the textbook.

The notes were initially worked out with the purpose of being used as slides in lectures in a design of experiments course based on Montgomery's book, and most of them are still in a format suited to be used as such.

Some important concepts that are not treated in the textbook (especially orthogonal polynomials, Duncan's and Newman-Keuls multiple range tests and Yates' algorithm) have been added and a number of useful tables are given, most noteworthy, perhaps, the expected mean square tables for all analysis of variance models including up to 3 fixed and/or random factors.

A strict mathematical presentation is not intended, but by indicating some of the exact results and showing examples and numerical calculations it is hoped that a little deeper understanding of the different ideas and methods can be achieved.

In all circumstances, I hope these notes can inspire and assist the student in studying and learning a number of the most fundamental principles in the wonderful art of designing and analyzing scientific experiments.

The present version is a revision of the previous (2003) notes. Some of the material is reorganized and some additions have been made (sample size calculations for analysis of variance models and a simpler calculation of expectations of mean squares (2005)).

July 2004

A moderate revision has been made in January 2006 in which, primarily, the page references have been changed to the 6th edition of Montgomery's textbook.

January 2006

A larger revision was undertaken in August 2006. The format is now landscape. A number of slides I considered less important have been taken out. I hope this has clarified the subjects concerned.

August 2007 :

A major revision was carried out. No new material, but (hopefully) better organized. In part 11 a new and very easy way of computing expected mean squares (EMS) is introduced.

Henrik Spliid

August 2007

List of contents

List of contents, cont.

List of contents, cont.

List of contents, cont.

List of contents, cont.

List of contents, cont.

Design of Experiments (DoE)

What is DoE?

Ex: Hardening of a metallic item

Variables that may be of importance: Factors

1: Medium (oil, water, air or other)

2: Heating temperature

3: Other factors ?

Dependent variables: Response

1: Surface hardness

2: Depth of hardening

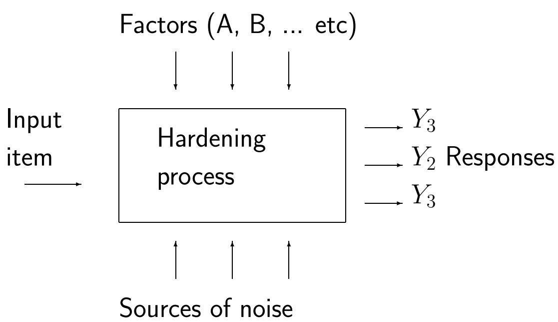
3: Others ?

Sources of variation (uncertainty)

1: Uneven usage of time for heating

2: All items not completely identical

3: Differences in handling by operators



Mathematical model

$$\boxed{Y = f(A, B, \ldots) + E}$$

How do we study the function $f(.)$.

The 25% rule.

Design of Experiments

Model of process determines	temperatures, heating time, etc. Factors in general based on a priori knowledge)
Laboratory resources decide	Number of measurements Practical execution Handling and staff
Conclusions wanted	How are data to be analyzed Which factors are important Which sources of uncertainty are important Estimation of effects and uncertainties

Demands:
You must have a reasonable <u>model idea</u> and you must have some idea about the <u>sources of uncertainty</u>.

Aims:
1) To <u>identify a good model</u>,
2) estimate its parameters,
3) assess the uncertainties of the experiment in general, and
4) assess the uncertainty of the estimates of the model in particular.

A weighing problem

Three items A B C

Standard weighing experiment:

Measurement	(1)	a	b	c
Meaning	No item	with A	with B	with C

$$\begin{array}{lrcl} \text{Model for} & (1) & = & \mu + E_1 \\ \text{responses} & \text{a} & = & \mu + A + E_2 \\ & \text{b} & = & \mu + B + E_3 \\ & \text{c} & = & \mu + C + E_4 \end{array}$$

$\mu =$ offset (zero reading) of weighing device

$A =$ weight of item A $\quad B =$ weight of item B $\quad C =$ weight of item C

E_1, E_2, E_3 and E_4 are the 4 measurement errors

The "natural" estimates of A, B and C are

$$\widehat{A} = a - (1)$$

and the corresponding for B and C

An alternative experiment:

(1)	ac	bc	ab
No item	with A and C	with B and C	with A and B

The alternative weighing design

$$\begin{array}{lrcl} \text{Model of} & (1) & = & \mu + E_5 \\ \text{responses} & \text{ac} & = & \mu + A + C + E_6 \\ & \text{bc} & = & \mu + B + C + E_7 \\ & \text{ab} & = & \mu + A + B + E_8 \end{array}$$

$$\widehat{A}^* = \frac{-(1) + ac - bc + ab}{2} = A + \frac{\text{4 errors}}{2}$$

Which design is preferable and why?

$$\mathrm{Var}\{\widehat{A}\} = 2\sigma_E^2$$

$$\mathrm{Var}\{\widehat{A}^*\} = \frac{4\sigma_E^2}{2^2} = \sigma_E^2$$

Conclusion

The alternative design is preferable because

1) The two designs both use 4 measurements

but

2) The second design is (much) more precise than the first design.

The reason for this is that

In the first design not all measurements are used to estimate all parameters, which is the case in the second design.

This is a basic property of (most) good designs.

Some repetition of elementary statistics

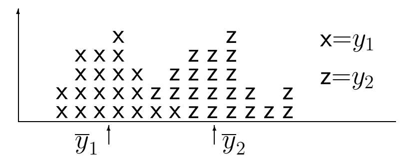
Table 2-1: Portland cement strength		
Observation number	Modified Mortar	Unmodified mortar
1	16.85	17.50
2	16.40	17.63
:	:	:
10	16.57	18.15

Factor: Types of mortar with 2 levels
Response: Strength of cement

The experiment represents a comparative (not absolute) study (it assesses differences between types of mortar).

Two treatments: the t-test can be applied

Two distributions to compare



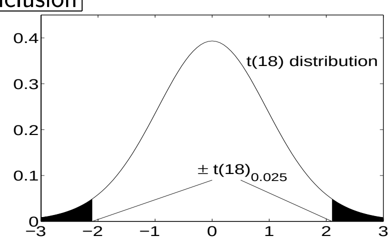
Model: $Y_{ij} = \mu_i + E_{ij} = \mu + \tau_i + E_{ij}; i = \{1, 2\}$ with $\tau_1 + \tau_2 = 0$

Test of $H_0 : \mu_1 = \mu_2 \Longleftrightarrow \tau_1 = \tau_2 = 0$

$$t = \frac{(\overline{Y}_1 - \overline{Y}_2) - (\mu_1 - \mu_2)}{s\sqrt{1/n_1 + 1/n_2}}$$

$$s^2 = s^2_{pooled} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 - 1 + n_2 - 1}$$

The t-test and the conclusion



Example p. 36: $\overline{Y}_1 = 16.76$, $\overline{Y}_2 = 17.92$, $s^2 = 0.284^2$

$$\mu_1 = \mu_2 \Rightarrow t = \frac{16.76 - 17.92}{0.284\sqrt{1/10 + 1/10}} = -9.13$$

The difference is strongly significant

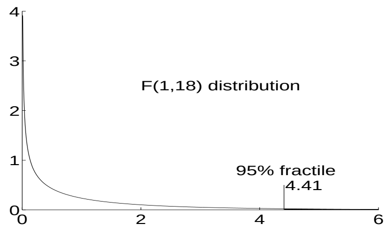
Analysis of variance for cement data

Two levels $\sim$ Two treatments.

The test (of the hypothesis of no difference between treatments) can be formulated as an analysis of variance (one-way model):

Source of variation	SSQ	df	s^2	F value
Between treatments	6.7048	2−1	6.7048	82.98
Within treatments	1.4544	18	0.0808	
Total Variation	8.1592	20−1		

$\boxed{\text{The reference distribution is an F-distribution:}}$



The t-test and the one-way analysis of variance with two treatments give the same results.

The F-value in the analysis of variance is the t-value squared:

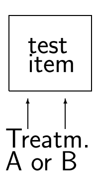
$$t^2(f) \sim F(1, f)$$

$\boxed{\text{Conclusions to formulated:}}$

Point estimates for	μ_1 and μ_2
	$\mu_1 - \mu_2$
	σ_E^2
Confidence intervals for	μ_1 and μ_2
	$\mu_1 - \mu_2$
	σ_E^2

and a suitable verbal formulation of the obtained result

$\boxed{\text{Two alternative experimental designs}}$



$\boxed{\text{Design I}}$: 20 items used

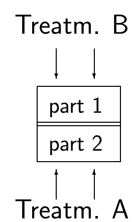
Method A	Method B
$Y_{1,A}$	$Y_{1,B}$
$Y_{2,A}$	$Y_{2,B}$
:	:
$Y_{10,A}$	$Y_{10,B}$

Allocation of treatments to items by randomization

The method of analysis?

Answer: One-way analysis of variance (or t-test)

$\boxed{\text{An alternative design using blocks (items)}}$



$\boxed{\text{Design II}}$: 10 items used

Item	Method A	Method B
1	$Y_{1,A}$	$Y_{1,B}$
2	$Y_{2,A}$	$Y_{2,B}$
:	:	:
10	$Y_{10,A}$	$Y_{10,B}$

Allocation of treatments to the two parts by randomization

The proper mathematical model is a two-way analysis of variance model.

Formulate the two models for designs I and II.

Which design is preferred? Why?

$\boxed{\text{Detailed mathematical models}}$

$$\boxed{\text{Design I}}: \begin{array}{l} Y_{i,A} = \mu_A + E_{i,A} + U_{i,A} \\ Y_{i,B} = \mu_B + E_{i,B} + U_{i,B} \end{array}$$

$$\mathrm{Var}\{\overline{\mathrm{Y}}_{\mathrm{A}} - \overline{\mathrm{Y}}_{\mathrm{B}}\} = \frac{2\sigma_{\mathrm{E}}^2 + 2\sigma_{\mathrm{U}}^2}{\mathrm{n}}$$

$$\boxed{\text{Design II}}: \begin{array}{l} Y_{i,A} = \mu_A + E_i + U_{i,A} \\ Y_{i,B} = \mu_B + E_i + U_{i,B} \end{array}$$

$$Y_{i,A} - Y_{i,B} = D_i = \mu_A - \mu_B + U_{i,A} - U_{i,B}$$

$$\mathrm{Var}\{\overline{\mathrm{Y}}_{\mathrm{A}} - \overline{\mathrm{Y}}_{\mathrm{B}}\} = \frac{2\sigma_{\mathrm{U}}^2}{\mathrm{n}}$$

Conclusion: Design II eliminates the variation <u>between</u> items.

Design II is preferable. The analysis is a paired t-test or a two-way analysis of variance with 2 treatments and 10 blocks.

$\boxed{\text{Analysis of variance example}}$

Sequence of measurements				
Factor is % cotton				
15%	20%	25%	30%	35%
1	6	11	16	21
2	7	12	17	22
3	8	13	18	23
4	9	14	19	24
5	10	15	20	25
The table displays a systematic sequence of measurements				

What are the problems with this design?

An alternative design: Randomized sequence

Factor is % cotton				
15%	20%	25%	30%	35%
7 (15)	12 (8)	14 (5)	19 (11)	7 (24)
7 (1)	17 (9)	18 (2)	25 (22)	20 (10)
15 (4)	12 (23)	18 (18)	22 (13)	16 (20)
11 (21)	18 (12)	19 (14)	19 (7)	15 (17)
9 (19)	18 (16)	19 (3)	23 (25)	11 (6)
The table displays both the data and the random sequence of measurements in (.)				

What is achieved by randomizing the sequence?

Mathematical model for randomized design

$$Y_{ij} = \mu + \tau_j + E_{ij}$$

Factor is % cotton						
	15%	20%	25%	30%	35%	sum
	7	12	14	19	7	
	7	17	18	25	20	
	15	12	18	22	16	
	11	18	19	19	15	
	9	18	19	23	11	
Sum	49	77	88	108	54	376
Complete randomization assumed						

$$SSQ_{tot} = 7^2 + 7^2 + 15^2 + \ldots + 11^2 - \frac{376^2}{25} = 636.96$$

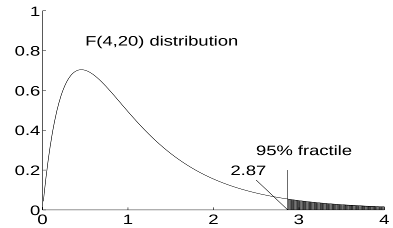
$$SSQ_{treatm} = \frac{49^2 + 77^2 + 88^2 + 108^2 + 54^2}{5} - \frac{376^2}{25} = 475.76$$

$$SSQ_{resid} = SSQ_{tot} - SSQ_{treatm} = 161.20$$

$$f_{tot} = N - 1 = 25 - 1 = 24$$
$$f_{treatm} = a - 1 = 5 - 1 = 4$$
$$f_{resid} = a(n - 1) = 5(5 - 1) = 20$$

ANOVA table for cotton experiment

Source	SSQ	f	s^2	EMS	F-value
Cotton	475.76	4	118.94	$\sigma_E^2 + 5\phi_\tau$	14.76
Residual	161.20	20	8.06	σ_E^2	
Total	636.96	24			



Conclusion: Since $14.76 >> 2.87$ the percentage of cotton is of importance for the strength measured.

Model identified:

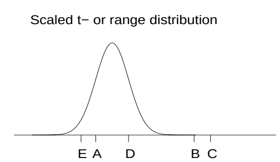
$$Y_{ij} = \mu + \tau_j + E_{ij}$$

Parameter	μ	τ_j	σ_E^2
Estimate	$\overline{Y}_{..}$	$\overline{Y}_{.j} - \overline{Y}_{..}$	s_E^2
Value from data	15.04	-5.24 0.36 2.56 6.56 -4.24	$8.06 = 2.84^2$

Design without or with structure - how to analyse after ANOVA

Design without structure				
A	B	C	D	E
Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}
Y_{21}	Y_{22}	Y_{23}	Y_{24}	Y_{25}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Y_{n1}	Y_{n2}	Y_{n3}	Y_{n4}	Y_{n5}

Design with structure		
A=control	B_1	B_2
Y_{11}	Y_{12}	Y_{13}
Y_{21}	Y_{22}	Y_{23}
$\vdots$	$\vdots$	$\vdots$
Y_{n1}	Y_{n2}	Y_{n3}

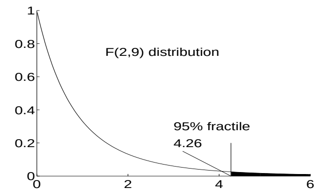


Natural comparisons?
Use orthogonal contrasts (two !)
How can they be constructed?

Important example of orthogonal contrasts

Design with structure		
A=Control	Tablet=B_1	Inject=B_2
24.0	11.0	23.0
29.0	18.5	21.0
32.1	29.0	18.8
28.0	16.0	16.8
113.1	74.5	79.6
Present method	Two alternative methods	

ANOVA table for drug experiment				
Source	SSQ	f	s^2	F-value
Treatm.	219.85	3-1	119.93	4.34
Residual	227.84	9	25.3	
Total	447.69	12-1		



$F(2,9)_{0.05} = 4.26$, such that the variation between treatments is (just) significant at the 5% significance level.

What now? We can suggest reasonable contrasts:

$$C_{A-B} = 2 \cdot T_A - (T_{B_1} + T_{B_2}) = 72.1$$

$$SSQ_{A-B} = \frac{C_{A-B}^2}{4 \cdot (2^2 + (-1)^2 + (-1)^2)} = 216.60 \quad , \quad f = 1$$

$$C_{B_1-B_2} = 0 \cdot T_A + T_{B_1} - T_{B_2} = -5.1$$

$$SSQ_{B_1-B_2} = \frac{C_{B_1-B_2}^2}{4 \cdot (0^2 + 1^2 + (-1)^2)} = 3.25 \quad , \quad f = 1$$

Splitting up the variance between treatments in two parts:

Detailed ANOVA table for drug experiment				
Source	SSQ	f	s^2	F-value
Between A and B: A−B	216.60	1	216.60	8.56
Between the two B's: B_1-B_2	3.25	1	3.25	0.13
Residual	227.84	9	25.3	
Total	447.69	12-1		

$F(1,9)_{0.05} = 5.12$, such that A−B is significant, but B_1-B_2 is far from.

The variation between all three treatments has been split up in variation between A and the B's and variation between the two B's.

The B's are probably not (very) different while A has significantly higher response than the B's.

Some 'patterns' leading to orthogonal contrasts

Design I	A	B_1	B_2
Contrasts	$2T_A$	$-T_{B_1}$	$-T_{B_2}$
		T_{B_1}	$-T_{B_2}$

Design II	A_1	A_2	B_1	B_2
Contrasts	T_{A_1}	$+T_{A_2}$	$-T_{B_1}$	$-T_{B_2}$
	T_{A_1}	$-T_{A_2}$		
			T_{B_1}	$-T_{B_2}$

Design III	A	B_1	B_2	B_3
Contrast	$3T_{A_1}$	$-T_{B_1}$	$-T_{B_2}$	$-T_{B_3}$
(artificial)		$2T_{B_1}$	$-T_{B_2}$	$-T_{B_3}$
(artificial)			T_{B_2}	$-T_{B_3}$

In the design III example the SSQ's from the two artificial contrasts $[2T_{B_1} - T_{B_2} - T_{B_3}]$ and $[T_{B_2} - T_{B_3}]$ add up to the variation between the three B's. An ANOVA table could in principal look like

Source	SSQ	f	s^2	F-value
A−B	SSQ_{A-B}	1		
Between B's	SSQ_B	2		
Residual	SSQ_{res}	N-1-3		
Total	SSQ_{tot}	N-1		

Patterns in two-way factorial designs

Factor A	Factor B B1	B2
A1	T_{11}	T_{12}
A2	T_{21}	T_{22}

Totals	T_{11}	T_{12}	T_{21}	T_{22}	Effect
Coeffi-	-1	-1	$+1$	$+1$	A main
cients	-1	$+1$	-1	$+1$	B main
	$+1$	-1	-1	$+1$	AB interaction

A 3 × 2 design

Factor A	Factor B B1	B2
Control (C)	T_{01}	T_{02}
A1	T_{11}	T_{12}
A2	T_{21}	T_{22}

Totals	T_{01}	T_{02}	T_{11}	T_{12}	T_{21}	T_{22}	Effect
Main	-2	-2	$+1$	$+1$	$+1$	$+1$	A-C
effects			-1	-1	$+1$	$+1$	A
	-1	$+1$	-1	$+1$	-1	$+1$	B
Inter-	$+2$	-2	-1	$+1$	-1	$+1$	(A-C)×B
actions			$+1$	-1	-1	$+1$	A×B

The two last contrasts correspond to interactions. They are easily constructed by multiplication of the coefficients of the corresponding main effects. All 5 contrasts are orthogonal.

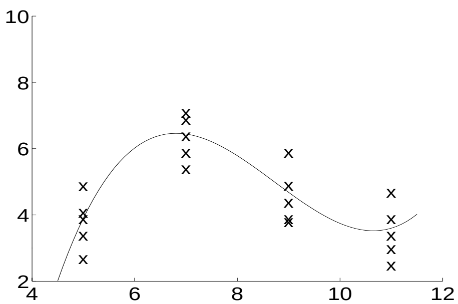
Polynomial effects in ANOVA

	Concentration			
	5%	7%	9%	11%
	3.5	6.0	4.0	3.1
	5.0	5.5	3.9	4.0
	2.8	7.0	4.5	2.6
	4.2	7.2	5.0	4.8
	4.0	6.5	6.0	3.5
Sum	19.5	32.2	23.4	18.0

Model : $Y_{ij} = \mu + \tau_j + E_{ij}$

ANOVA of response				
Source	SSQ	d.f.	s²	F
Concentration	24.35	4−1	8.1167	12.41
Residual	10.46	16	0.6538	(sign)
Total	34.81	20−1		

Plot of data and approximating 3. order polynomium:



Polynomial estimation in ANOVA

Possible empirical function as a polynomial:

$$Y_{ij} = \beta_0 + \beta_1 \cdot x_j + \beta_2 \cdot x_j^2 + \beta_3 \cdot x_j^3 + E_{ij}$$

With 4 x-points a polynomial of degree (4−1)=3 can be estimated using standard (polynomial) regression analysis.

Alternative (reduced) models:

$$Y_{ij} = \beta_0 + \beta_1 \cdot x_j + \beta_2 \cdot x_j^2 + E_{ij}$$

$$Y_{ij} = \beta_0 + \beta_1 \cdot x_j + E_{ij}$$

$$Y_{ij} = \beta_0 + E_{ij} \quad \text{(ultimately)}$$

By the general regression test method these models can be tested successively in order to identify the proper order of the polynomial.

An alternative method to identify the necessary (statistically significant) order of the polynomial is based on orthogonal polynomials. The technique uses the concept of ortogonal regression and it is much similar to the orthogonal contrast technique.

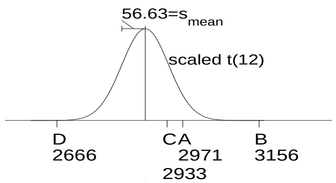
The technique is shown in the supplementary section I.

Exercise 3-1

Tensile strength			
A	B	C	D
3129	3200	2800	2600
3000	3300	2900	2700
2865	2975	2985	2600
2890	3150	3050	2765

ANOVA for mixing experiment				
Source	SSQ	df	s^2	F
Methods	489740	3	163247	12.73
Residual	153908	12	12826	
Total	643648	15		

How can we try to group the treatments?



$$s_{mean} = s_{residual}/\sqrt{n_{mean}} = \sqrt{12826}/\sqrt{4} = 56.63.$$

Which averages are possibly significantly different ?

LSD: Least Significant Difference

For example A versus B:

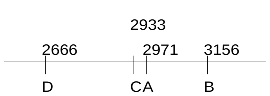
$$\frac{\overline{Y}_A - \overline{Y}_B}{s_{res}\sqrt{1/n_A + 1/n_B}} \sim t(f_{res})$$

$$|\overline{Y}_A - \overline{Y}_B| < s_{res}\sqrt{1/n_A + 1/n_B} \times t(f_{res})_{0.025}$$

Here $n_A = n_B = 4$, $s_{res} = 113.25$, $f_{res} = 12$

$$\boxed{|\overline{Y}_A - \overline{Y}_B| > 113.25\sqrt{1/4 + 1/4} \times 2.179 = 174.5 \quad ?}$$

$$\begin{aligned}
|A - B| &= |3156 - 2971| = 185 && \text{significant} \\
|A - C| &= |2971 - 2933| = 38 && \text{not significant} \\
|A - D| &= |2971 - 2666| = 305 && \text{significant} \\
|B - C| &= |3156 - 2933| = 223 && \text{significant} \\
|B - D| &= |3156 - 2666| = 490 && \text{significant} \\
|C - D| &= |2933 - 2666| = 223 && \text{significant}
\end{aligned}$$



Conclusion ? All pairs $\sim$ multiple testing - any problems ?

Newman - Keuls Range Test

Sort averages increasing: $\overline{Y}_{(1)}, \overline{Y}_{(2)}, \overline{Y}_{(3)}, \overline{Y}_{(4)}$

$$\boxed{\text{Range} = \overline{Y}_{(4)} - \overline{Y}_{(1)}}$$

Table VII (gives q_α) : Criterion

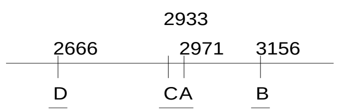
$$\overline{Y}_{(4)} - \overline{Y}_{(1)} > s_{mean} \cdot q_\alpha(4, f_{res}) \quad ?$$

$s_{mean} = s_{res}/\sqrt{n_{mean}} = 113.25/\sqrt{4} = 56.63$
$q_{0.05}(4, 12) = 4.20$

Range including 4: $LSR_4 = 4.20 \cdot 56.63 = 237.8$
Range including 3: $LSR_3 = 3.77 \cdot 56.63 = 213.5$
Range including 2: $LSR_2 = 3.08 \cdot 56.63 = 174.4$

B - D:	3156 - 2666	=	$490 > 237.8$ (LSR_4)	sign.
B - C:	3156 - 2933	=	$223 > 213.5$ (LSR_3)	sign.
B - A:	3156 - 2971	=	$185 > 174.4$ (LSR_3)	sign.
A - D:	2971 - 2666	=	$305 > 213.5$ (LSR_3)	sign.
A - C:	2971 - 2933	=	$38 < 174.4$ (LSR_2)	not s.

Conclusion:



$\boxed{\text{Duncans Multiple Range Test}}$

Sort averages increasing: $\overline{Y}_{(1)}, \overline{Y}_{(2)}, \overline{Y}_{(3)}, \overline{Y}_{(4)}$

$$\boxed{\text{Range} = \overline{Y}_{(4)} - \overline{Y}_{(1)}}$$

Criterion (from special table find r_α) :

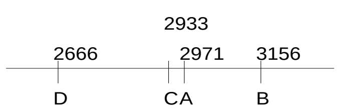
$$\overline{Y}_{(4)} - \overline{Y}_{(1)} > s_{mean} \cdot r_\alpha(4, f_{res}) \quad ?$$

$s_{mean} = s_{res}/\sqrt{n_{mean}} = 113.25/\sqrt{4} = 56.63$
$r_{0.05}(4, 12) = 3.33$

Range including 4: $LSR_4 = 3.33 \cdot 56.63 = 188.6$
Range including 3: $LSR_3 = 3.23 \cdot 56.63 = 182.9$
Range including 2: $LSR_2 = 3.08 \cdot 56.63 = 174.4$

B - D:	3156 - 2666	=	$490 > 188.6$ (LSR_4)	sign.
B - C:	3156 - 2933	=	$223 > 182.9$ (LSR_3)	sign.
B - A:	3156 - 2971	=	$185 > 174.4$ (LSR_3)	sign.
A - D:	2971 - 2666	=	$305 > 182.9$ (LSR_3)	sign.
A - C:	2971 - 2933	=	$38 < 174.4$ (LSR_2)	not s.

Conclusion is the same as for Newman -Keuls here:



$\boxed{\text{Newman - Keuls \& Duncans test}}$

Works alike, but use different types of range distributions. For example:

Duncan	Newman - Keuls
$r(6, 12)_{0.05} = 3.40$	$q(6, 12)_{0.05} = 4.75$
$r(5, 12)_{0.05} = 3.36$	$q(5, 12)_{0.05} = 4.51$
$r(4, 12)_{0.05} = 3.33$	$q(4, 12)_{0.05} = 4.20$
$r(3, 12)_{0.05} = 3.23$	$q(3, 12)_{0.05} = 3.77$
$r(2, 12)_{0.05} = 3.08$	$q(2, 12)_{0.05} = 3.08$
More significances	More conservative

A grouping of averages that is significant according to Newman - Keuls test is more reliable

No structure on treatments $\Longrightarrow$
Use Newman Keuls or Duncans test
(LSD method not recommendable)

Structure on treatments $\Longrightarrow$ Use contrast method or fx Dunnetts test (below)

Dunnetts test

	Control	Alternative Treatments		
	A	B	C	D
Parameters	μ_A	μ_B	μ_C	μ_D

H_0: $\mu_A = \mu_B = \mu_C = \mu_D$

H_1: One or more of (μ_B, μ_C, μ_D) different from μ_A

Example: Exercise 3-1 with A as control (fx).

Two sided criterion:

$$|\overline{Y}_A - \overline{Y}_B| > s_{res}\sqrt{1/n_A + 1/n_B} \cdot d(4-1, 12)_{0.05}$$

$d(3,12)_{0.05}$(two sided) $= 2.68 \Longrightarrow$

critical difference $= \sqrt{12826}\sqrt{1/4 + 1/4} \cdot 2.68 = 214.7$

One sided criterion:

$$\overline{Y}_A - \overline{Y}_B > s_{res}\sqrt{1/n_A + 1/n_B} \cdot d(4-1, 12)_{0.05}$$

$d(3,12)_{0.05}$(one sided) $= 2.29 \Longrightarrow$

critical difference $= \sqrt{12826}\sqrt{1/4 + 1/4} \cdot 2.29 = 183.5$

More reliable (and correct) than LSD if relevant

The fixed (deterministic) effect ANOVA model

4 treatments			
Filter	Clean	Heat	Nothing
x	x	x	x
x	x	x	x
x	x	x	x
x	x	x	x

Model for response:
$Y_{ij} = \mu + \tau_j + E_{ij}$

The 4 treatment effects are deterministic
(μ and τ_j are constants)

Assumptions: $\Sigma_j \tau_j = 0$ and $E_{ij} \in N(0, \sigma_E^2)$

The random effect ANOVA model (see chapter 13 in 6th ed. of book)

Example: choose 4 batches among a large number of possible batches and measure some response (purity for example) on these batches:

4 batches			
B-101	B-309	B-84	B-211
x	x	x	x
x	x	x	x
x	x	x	x
x	x	x	x

Model for response:
$Y_{ij} = \mu + B_j + E_{ij}$

The 4 batch effects are random variables (B_j are random variables)

Assumptions: $B_j \in N(0, \sigma_B^2)$ and $E_{ij} \in N(0, \sigma_E^2)$

σ_E^2 and σ_B^2 are called variance components:

They are the variances **within** and **between** (randomly chosen) batches, respectively.

Fixed effect model: $Y_{ij} = \mu + \tau_j + E_{ij}$

ANOVA for fixed effect model					
Source	SSQ	df	s^2	EMS = E$\{s^2\}$	F
Methods	SSQ_τ	f_τ	s_τ^2	$\sigma_E^2 + n \cdot \phi_\tau$	s_τ^2/s_E^2
Residual	SSQ_E	f_E	s_E^2	σ_E^2	
Total	SSQ_{tot}	f_{tot}			

$\phi_\tau = \Sigma_j \, \tau_j^2/(a-1)$, and $\widehat{\tau}_j = \overline{Y}_{.j} - \overline{Y}_{..}$

Fixed (deterministic) effects: temperature, concentration, treatment, etc.

Random effect model: $Y_{ij} = \mu + B_j + E_{ij}$

ANOVA for random effect model					
Source	SSQ	df	s^2	EMS = E$\{s^2\}$	F
Batches	SSQ_B	f_B	s_B^2	$\sigma_E^2 + n \cdot \sigma_B^2$	s_B^2/s_E^2
Residual	SSQ_E	f_E	s_E^2	σ_E^2	
Total	SSQ_{tot}	f_{tot}			

$\sigma_B^2 = \mathrm{V}\{B\}$, and $\widehat{\sigma}_B^2 = (s_B^2 - s_E^2)/n$

Random effects: batches, days, persons, experimental rounds, litters of animals, etc.

Example 13-1, p 487, typical example of random effect model

Looms			
1	2	3	4
98	91	96	95
97	90	95	96
99	93	97	99
96	92	95	98

Model for tensile strength:
$Y_{ij} = \mu + L_j + E_{ij}$

The 4 looms are randomly chosen with effects L_j (being random variables)

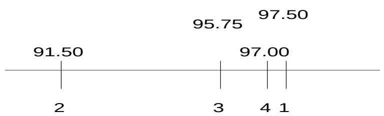
Assumptions: $L_j \in N(0, \sigma_L^2)$ and $E_{ij} \in N(0, \sigma_E^2)$

One-way ANOVA for loom example

ANOVA for variation between looms					
Source	SSQ	df	s^2	$E\{s^2\}$	F
Looms	89.19	3	29.73	$\sigma_E^2 + 4 \cdot \sigma_L^2$	15.65
Residual	22.75	12	1.90	σ_E^2	
Total	111.94	15			

$F(3,12)_{0.05} = 3.49 << 15.65 \Longrightarrow$ significance!

$\widehat{\sigma}_E^2 = 1.90 = 1.38^2$
$\widehat{\sigma}_L^2 = (29.73 - 1.90)/4 = 6.96 = 2.64^2$



How do we further analyze this result?

Newman-Keuls or Duncans test on looms

First: $s_Y = \sqrt{1.90} = 1.38 \Longrightarrow s_{\overline{Y}} = \sqrt{1.90/4} = 0.69$

Example: Newman - Keuls test:

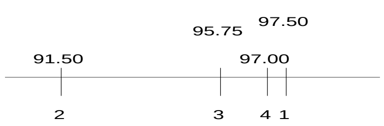
Find least significant ranges ($q(.,.)$) from studentized range table and multiply with standard deviation of group means:

$$\begin{array}{rcll} & & & \text{LSR} \\ q_{0.05}(4,12) & = & 4.20 \rightarrow \times s_{\overline{Y}} & = \ 2.90 \\ q_{0.05}(3,12) & = & 3.77 \rightarrow \times s_{\overline{Y}} & = \ 2.60 \\ q_{0.05}(2,12) & = & 3.08 \rightarrow \times s_{\overline{Y}} & = \ 2.13 \end{array}$$

Compare group means:

The smallest and the largest first and continue if difference is significant.
Then next largest versus smallest, etc.:

$$\begin{array}{rcl} |97.50 - 91.50| & = & 6.00 \ > \ 2.90 : \text{significant} \\ |97.50 - 95.75| & = & 1.75 \ < \ 2.60 : \text{not significant} \\ |91.50 - 97.00| & = & 5.50 \ > \ 2.60 : \text{significant} \\ |91.50 - 95.75| & = & 4.25 \ > \ 2.13 : \text{significant} \end{array}$$



Conclusion: loom no 2 is significantly different from the other looms

Confidence interval for σ_L^2

Interval for σ_L^2/σ_E^2 can be constructed

$$\boxed{\text{Lower} < \sigma_L^2/\sigma_E^2 < \text{Upper}}$$

$$\text{Lower} = \left[\frac{s_L^2}{s_E^2} \times \frac{1}{F(a-1, N-a)_{\alpha/2}} - 1\right]\frac{1}{n}$$

$$\text{Upper} = \left[\frac{s_L^2}{s_E^2} \times F(N-a, a-1)_{\alpha/2} - 1\right]\frac{1}{n}$$

Looms: Lower $= [15.65/4.47 - 1]/4 = 0.625$

Upper $= [15.65 \cdot 14.34 - 1]/4 = 55.85$

An alternative:

$$\boxed{\frac{\text{Lower}}{1+\text{Lower}} < \frac{\sigma_L^2}{\sigma_L^2+\sigma_E^2} < \frac{\text{Upper}}{1+\text{Upper}}}$$

Choice of sample size

i	A	B	C
1	y_{11}	y_{12}	y_{13}
2	y_{21}	y_{22}	y_{23}
:	:	:	:
n	y_{n1}	y_{n2}	y_{n3}

Problem : Choose sample size n with k treatment/groups

Fixed effect model : $Y_{ij} = \mu + \tau_j + E_{ij},\ \Sigma_i \tau_i = 0$

Requirements: 1) Know or assume σ_E^2

2) Which τ's are of interest to detect

3) How certain do we want to be to detect

Random effect model : $Y_{ij} = \mu + B_j + E_{ij},\ V(B) = \sigma_B^2$

Requirements: 1) Know or assume σ_E^2

2) Which σ_B^2 is of interest to detect

3) How certain do we want to be to detect

The textbook has graphs for both cases pp. 613-620. Below, after the examples based on the textbook, some mere general results are presented.

Example fixed effect model

Assume (based on previous knowledge) : $\sigma_E^2 \simeq 1.5^2$

Interesting values for τ (fx) : $\{-2.00, 0.00, +2.00\}$

Criterion: $P\{detection\} \geq 0.80$ (for example)

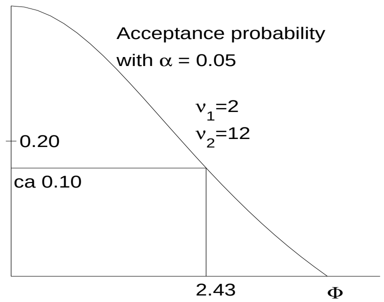
Try $n = 5$ (to start with)

Compute $\Phi^2 = (n\,\Sigma_j\,\tau_j^2)/(a \cdot \sigma_E^2)$

$= 5 \cdot (2^2 + 0^2 + 2^2)/(3 \cdot 1.5^2) = 5.92$

Compute $\Phi = \sqrt{5.92} = 2.43$

Read off graph page 613: $\nu_1 = a - 1 = 3 - 1 = 2$
$\nu_2 = a(n-1) = 3(5-1) = 12$



The graph shows, that $n = 5$ is enough

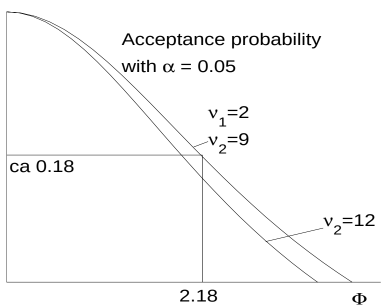
Will 4 be enough?

Compute $\Phi^2 = (n\,\Sigma_j\,\tau_j^2)/(a \cdot \sigma_E^2)$

$$= 4 \cdot (2^2 + 0^2 + 2^2)/(3 \cdot 1.5^2) = 4.74$$

Compute $\Phi = \sqrt{4.74} = 2.18$

Read off graph page 613: $\nu_1 = a - 1 = 3 - 1 = 2$
$\nu_2 = a(n-1) = 3(4-1) = 9$



The graph shows, that with $n = 4$ and testing with level of significance $\alpha = 0.05$ the probability of acceptance is about 18%.

The probability of rejection (detection of significant τ's) is about 82%.

$n = 4$ is thus enough.

$\boxed{\text{Example random effect model}}$

Assume (based on previous knowledge) : $\sigma_E^2 \simeq 1.5^2$

Interesting values (for example) for $\sigma_B^2 : 2.0^2$

Criterion: $P\{detection\} \geq 0.90$ (for example).

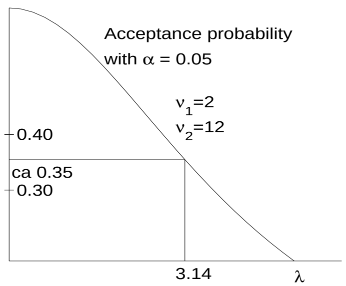
Try $n = 5$ (to start with)

Compute $\lambda = \sqrt{\frac{\sigma_E^2 + n \cdot \sigma_B^2}{\sigma_E^2}} = \sqrt{\frac{1.5^2 + 5 \cdot 2.0^2}{1.5^2}} = 3.14$

Read off graph page 617 :

$$\nu_1 = a - 1 = 3 - 1 = 2$$
$$\nu_2 = a(n-1) = 3(5-1) = 12$$

Note: The degrees of freedom labeling is wrong - for the $\alpha = 0.05$ curves. It should be as shown for the $\alpha = 0.01$ curves and for all graphs with $\nu_1 \geq 4$.



The graph shows, that $n = 5$ is not enough

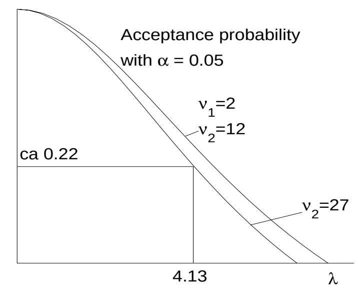
Will 10 be enough?

$$\lambda = \sqrt{\frac{\sigma_E^2 + n \cdot \sigma_B^2}{\sigma_E^2}} = \sqrt{\frac{1.5^2 + 10 \cdot 2.0^2}{1.5^2}} = 4.33$$

Read off graph page 617:

$$\nu_1 = a - 1 = 3 - 1 = 2$$
$$\nu_2 = a(n-1) = 3(10-1) = 27$$

Note: Remember the degrees of freedom labeling again!



The graph shows, that with $n = 10$ and testing with level of significance $\alpha = 0.05$ the probability of acceptance is still about 0.22 (it should be max. 0.10).

$n = 10$ is thus not enough. The graph p. 617 shows, that for $\lambda = 5.2$ the acceptance probability $\simeq$ 0.10 . It will require about $n = 15$ for $\sigma_E^2 = 1.5^2$ and $\sigma_B^2 = 2^2$.

In the supplementary part III the exact determination of sample size is described for bth deterministic and random effects models.

Block designs - one factor and one blocking criterion

Sources of uncertainty (noise)

Day-to-day variation

Batches of raw material

Litters of animals

Persons (doing the lab work)

Test sites or alternative systems

Treatment	A	B	C
Batch	B-X	B-V	B-II
Data	Y_{11}	Y_{12}	Y_{13}
	Y_{21}	Y_{22}	Y_{23}
	:	:	:
	Y_{n1}	Y_{n2}	Y_{n3}

One factor and one block, but they vary in the same way!

Mathematical model : $Y_{ij} = \mu + \tau_j + B_j + E_{ij}$

Is the model correct ?

How can we analyze it ?

What can and what cannot be concluded ?

Is there a problem ?

Confounding ?

The index for the factor and the block is the same:

100% confounding.

Alternative to confounded design

Treatment	A	B	C
Data	Y_{11} (B-II)	Y_{12} (B-XI)	Y_{13} (B-IV)
	Y_{21} (B-IX)	Y_{22} (B-I)	Y_{23} (B-VI)
	:	:	:
	Y_{n1} (B-III)	Y_{n2} (B-XX)	Y_{n3} (B-IIX)

In the design the batches used for the individual measurements are shown in parentheses

The batches are selected randomly

Mathematical model : $Y_{ij} = \mu + \tau_j + B_{ij} + E_{ij}$

How can this model be analyzed ?

What does the randomization do with respect to the mean and variance of Y_{ij} ?

Compared to the above design: any problems solved ?

Have any new problems been introduced ?

Can the second design be improved even more (how) ?

Examples of factors

Concentration of active compound in experiment: (2%,4%,6%,8%)

Electrical voltage in test circuit (10 volt, 12 volt, 14 volt)

Load in test of strength: (10 kp/m^2, 15 kp/m^2, 20 kp/m^2)

Alternative catalysts: (A, B, C, D)

Alternative cleaning methods: (centrifuge treatm., filtration, electrostatic removal)

Gender of test animal: (♀ , ♂)

Examples of blocks

Batches of raw material: (I, II, III, IV) (can be of limited size)

Collections of experiments conducted simultaneously (dates fx): (22/2-1990, 29/3-1990, 24/12-1990)

Groups of participants in an indoor climate experiment: (Test−team 1, Test−team 2, Test−team 3)

Litters of test animals: (Litter 1, Litter 2, Litter 3, Litter 4)

Position in test equipment: (position 1, position 2, position 3)

Design with inadequate confounding - schematic:

Thermometers (= experimental condition = block) are I, II and III.

Treatments	A	B	C
Data	25 (II)	16 (I)	19 (III)
	24 (II)	15 (I)	20 (III)
	24 (II)	17 (I)	20 (III)
Total	73	48	59

$$Y_{ij} = \mu + \alpha_j + T_j + E_{ij}$$

$$SSQ_{treat} = \frac{73^2+48^2+59^2}{3} - \frac{180^2}{9} = 104.67$$

$$SSQ_{tot} = (25^2 + 16^2 + \cdots + 20^2) - \frac{180^2}{9} = 108.00$$

$$SSQ_{resid} = SSQ_{tot} - SSQ_{treat}$$

Analysis of variance table for 100% confounded design

One way ANOVA					
Source of var.	SSQ	d.f.	s^2	EMS = E$\{s^2\}$	F-test
Between treatm.	104.67	$3-1$	52.335	$\sigma^2 + 3\phi_\alpha + 3\sigma_T^2$	94.30
Uncertainty	3.33	$3(3-1)$	0.555	σ^2	
Total	108.00	$3 \cdot 3 - 1$			

What can (or cannot) be concluded ?

Design with thermometers randomized

Thermometers are randomized (I, II, $\cdots$, X)

Treatments	A	B	C
Data	26 (X)	20 (IV)	25 (II)
	20 (II)	15 (V)	20 (VI)
	22 (I)	19 (III)	22 (V)
Total	68	54	67

$$Y_{ij} = \mu + \alpha_j + (T_{ij} + E_{ij})$$

$$SSQ_{treat} = \frac{68^2+54^2+67^2}{3} - \frac{189^2}{9} = 40.67$$

$$SSQ_{tot} = (26^2 + 20^2 + \cdots + 22^2) - \frac{189^2}{9} = 86.00$$

$$SSQ_{resid} = SSQ_{tot} - SSQ_{treat} = 45.34$$

Analysis of variance table for completely randomized design

One way ANOVA					
Source of var.	SSQ	d.f.	s^2	EMS	F-test
Between treatm.	40.67	$3-1$	20.34	$\sigma^2+\sigma_T^2+3\phi_\alpha$	2.69
Uncertainty	45.34	$3(3-1)$	7.56	$\sigma^2+\sigma_T^2$	
Total	86.00	$3\cdot 3-1$			

What can (or cannot) be concluded ?

Balanced block design

Thermometers are balanced (complete) blocks

	Treatments			
Thermometers	A	B	C	Total
I	25	18	21	64
II	21	15	19	55
III	22	18	20	60
Total	68	51	60	179

$$Y_{ij} = \mu + \alpha_j + T_i + E_{ij}$$

$$SSQ_{treat} = \frac{68^2+51^2+60^2}{3} - \frac{179^2}{9} = 48.22$$

$$SSQ_{therm} = \frac{64^2+55^2+60^2}{3} - \frac{179^2}{9} = 13.56$$

$$SSQ_{tot} = (25^2+18^2+\cdots+20^2) - \frac{179^2}{9} = 64.89$$

$$SSQ_{resid} = SSQ_{tot} - SSQ_{treat} - SSQ_{therm} = 3.11$$

Analysis of variance table for completely balanced block design

Two way ANOVA					
Source of var.	SSQ	d.f.	s^2	EMS	F-test
Between treatm.	48.22	$3-1$	24.11	$\sigma^2+3\phi_\alpha$	30.99
Between therm.	13.56	$3-1$	6.78	$\sigma^2+3\sigma_T^2$	(8.71)
Uncertainty	3.11	$8-2-2$	0.778	σ^2	
Total	64.89	8			

What can now be concluded ?

Latin square design

Laboratory technicians (labor : (1) (2) (3)) are balanced against both treatments and thermometers

	Treatments			
Thermometers	A	B	C	Total
I	27 (2)	20 (3)	21 (1)	68
II	21 (1)	18 (2)	20 (3)	59
III	24 (3)	17 (1)	22 (2)	63
Total	72	55	63	190

Labor−totals : (1)=59, (2)=67, (3)=64

$$Y_{ijk} = \mu + \alpha_j + T_i + L_k + E_{ijk}$$

Analysis of variance table for Latin square design

Two blocking criteria completely balanced block design :

$$SSQ_{treat} = 48.22 \text{ , } SSQ_{therm} = 13.56 \text{ , } SSQ_{tot} = 72.89$$

$$SSQ_{labor} = \frac{59^2+67^2+64^2}{3} - \frac{190^2}{9} = 10.89$$

$$SSQ_{resid} = 0.22$$

Latin square ANOVA					
Source of var.	SSQ	d.f.	s^2	EMS	F-test
Between treatm.	48.22	2	24.11	$\sigma^2 + 3\phi_\alpha$	219.19
Between therm.	13.56	2	6.78	$\sigma^2 + 3\sigma_T^2$	(61.68)
Between labor.	10.89	2	5.45	$\sigma^2 + 3\sigma_L^2$	(49.54)
Uncertainty	0.22	2	0.11	σ^2	
Total	72.89	8			

Triple balanced design = Graeco-Latin square

	Treatments			
Thermometers	A	B	C	Total
I	28 (2)(z)	21 (3)(y)	23 (1)(x)	72
II	23 (1)(y)	20 (2)(x)	20 (3)(z)	63
III	25 (3)(x)	18 (1)(z)	22 (2)(y)	65
Total	76	59	65	200

Labor−totals : (1)=64, (2)=70, (3)=66
Batch−totals : (x)=68, (y)=66, (z)=66

$$Y_{ijkr} = \mu + \alpha_j + T_i + L_k + B_r + E_{ijkr}$$

Analysis of variance table for Graeco-Latin square design

$$SSQ_{batch} = \frac{68^2+66^2+66^2}{3} - \frac{200^2}{9} = 0.89$$

$$SSQ_{treat} = 49.56,\ SSQ_{therm} = 14.89,\ SSQ_{labor} = 6.22$$

$$SSQ_{tot} = 71.56,\ SSQ_{resid} = 0.00$$

Graeco-Latin square ANOVA					
Source of var.	SSQ	d.f.	s^2	EMS	F-test
Between treatm.	49.56	2	24.78	$\sigma^2 + 3\phi_\alpha$	?
Between therm.	14.89	2	7.45	$\sigma^2 + 3\sigma_T^2$	?
Between labor.	6.22	2	3.11	$\sigma^2 + 3\sigma_L^2$	?
Between batches	0.89	2	0.45	$\sigma^2 + 3\sigma_B^2$	?
Uncertainty	0	0		(σ^2)	
Total	71.56	8			

The example shows the principle, but of course, since there is no residual variance no tests can be carried out. An external variance estimate could be used if available

Comments to slides 3.3 to 3.15

All the examples are created by numerical simulation using the corresponding models.

3.3: The variation of treatments is very significant, but it cannot be determined whether it is treatments or thermometers that cause it. If the experiment is repeated at a later occasion we will presumably again find a significant, but probably different treatment effect (since thermometers would be 3 other thermometers). The experiment is not reproducible and may lead to false conclusions.

3.4: The confounding treatments/thermometers is broken. However the variation between thermometers is causing a large uncertainty variance. The treatments are estimated with correct mean, but with a large variance. The treatments are not significant.

3.5: The thermometers are now balanced out of the ANOVA, and the estimate of the treatment effect has correct mean plus a small variance. The treatment effect is significant. Note, that essentially the SSQ for treatments is as in the randomized design, but the SSQ for the residual is now free of the variation between thermometers and, thus, much smaller.

3.6: In the Latin square the same principle as used for thermometers is now used for the laboratory technicians. Variation between technicians is eliminated from the residual variance, causing improved precision (however again loosing 2 degrees of freedom for the residual variance).

3.7: The same design principle (balance) is used to eliminate variation between batches from the residual variation.

The (important) two-period cross-over design (page 142)

Patient	Period 1	Period 2
1	A	B
2	B	A
3	B	A
4	A	B
:	:	:
:	:	:
2n-1	A	B
2n	B	A

$$Y_{ijk} = \mu + \tau_i + per_j + P_k + E_{ijk}$$

ANOVA for the two period cross over design:

Two period crossover ANOVA, example with 2n=20 patients					
Source of var.	SSQ	d.f.	s^2	EMS	F-test
Treatments (τ_i)	28.15	1	28.15	$\sigma^2 + 20\phi_\tau$	4.54
Periods (per_j)	2.45	1	2.45	$\sigma^2 + 20\phi_{per}$	0.40
Patients (P_k)	915.80	19	48.20	$(\sigma^2 + 2\sigma_P^2)$	(7.77)
Uncertainty	116.60	18	6.20	σ^2	
Total	1063.03	39			

The design consists of $R = n$ Latin squares repeated with different persons in all squares and identical periods (1 or 2).

The analysis of this design can take other forms if *residual effects* are suspected (effect from A on B different from the effect from B on A).

A little more about Latin squares

Table 4-8 page 136					
Batches of raw material	Operators 1	2	3	4	5
1	A	B	C	D	E
2	B	C	D	E	A
3	C	D	E	A	B
4	D	E	A	B	C
5	E	A	B	C	D
Treatments A, B, C, D, E					
A standard Latin square					

$$Y_{ijk} = \mu + \tau_i + B_j + O_k + E_{ijk}$$

ANOVA of Latin square example

ANOVA					
Source of var.	SSQ	d.f.	s^2	EMS	F-test
Treatments	330.00	4	82.50	$\sigma^2 + 5\phi_\tau$	7.73
Batches	68.00	4	17.00	$(\sigma^2 + 5\sigma_B^2)$	(1.59)
Operators	150.00	4	37.50	$(\sigma^2 + 5\sigma_O^2)$	(3.51)
Uncertainty	128.00	12	10.67	σ^2	
Total	676.00	24			

Interpretation of result of ANOVA

What has been achieved by using this design ?

Replication of Latin squares

	O_1	O_2	O_3		O_1	O_2	O_3		O_1	O_2	O_3
B_1	A	B	C	B_1	B	C	A	B_1	C	A	B
B_2	B	C	A	B_2	A	B	C	B_2	B	C	A
B_3	C	A	B	B_3	C	A	B	B_3	A	B	C
		R_1				R_2				R_3	

3 squares with identical operators and batches

	O_1	O_2	O_3		O_4	O_5	O_6		O_7	O_8	O_9
B_1	A	B	C	B_1	B	C	A	B_1	C	A	B
B_2	B	C	A	B_2	A	B	C	B_2	B	C	A
B_3	C	A	B	B_3	C	A	B	B_3	A	B	C
		R_1				R_2				R_3	

3 squares with 9 operators and 3 batches

	O_1	O_2	O_3		O_4	O_5	O_6		O_7	O_8	O_9
B_1	A	B	C	B_4	B	C	A	B_7	C	A	B
B_2	B	C	A	B_5	A	B	C	B_8	B	C	A
B_3	C	A	B	B_6	C	A	B	B_9	A	B	C
		R_1				R_2				R_3	

3 squares with 9 operators and 9 batches

Which design is probably the most precise (with smallest residual variance)? – Answer: The third design, but why?

How are the three different designs analyzed ?

In the supplementary part 6 the detailed ANOVA tables are indicated for each of the three cases.

To block or not to block ? Example 4.1

Type of tip	Test item (block) 1	2	3	4
A	9.3	9.4	9.6	10.0
B	9.4	9.3	9.8	9.9
C	9.2	9.4	9.5	9.7
D	9.7	9.6	10.0	10.2

$$Y_{ij} = \mu + t_i + B_j + E_{ij}$$

ANOVA for block design (data scaled 10:1)					
Source	SSQ	df	s^2	EMS	F
Type of tip (t)	38.50	3	12.83	$\sigma^2 + 4\phi_t$	14.44
Test item (B)	82.50	3	27.50	$\sigma^2 + 4\sigma_B^2$	(30.94)
Residual	8.00	9	0.89	σ^2	
Total	129.00	15			

Two alternative designs - which one is best?

Type of tip	Test item (block) 1	2	3	4
A	x x	x x	x x	x x
B	x x	x x	x x	x x
C	x x	x x	x x	x x
D	x x	x x	x x	x x

4 blocks of size 8. Double measurements for each treatment within the blocks.

Round 1	Test item (block) 1	2	3	4	5	6	7	8
A	x	x	x	x	x	x	x	x
B	x	x	x	x	x	x	x	x
C	x	x	x	x	x	x	x	x
D	x	x	x	x	x	x	x	x

8 blocks of size 4. One measurements for each treatment within the blocks.

The second design is preferable. It is more precise, because the blocks are smaller (variance within blocks is smaller).

Randomization is easier to do correct in small blocks and experimental circumstances are easier to keep constant.

ANOVA test of additivity

a treatments	b blocks 1	2	:	b
A (1)	y y	y y	:	y y
B (2)	y y	y y	:	y y
:	:	:	:	:
D (a)	y y	y y	:	y y

Basic model for block design with n measurements pr combination (the general case):

$$Y_{ijk} = \mu + t_i + B_j + E_{ijk}$$

where $i = \{1, a\}$, $j = \{1, b\}$ and $k = \{1, n\}$

If $n > 1$ start with

$$Y_{ijk} = \mu + t_i + B_j + TB_{ij} + E_{ijk}$$

and test the TB_{ij} term (two way ANOVA with interaction term) against E_{ijk} term

If accepted, reduce model to 'ideal model' and analyze as usual (two way ANOVA without interaction term)

If rejected, use TB_{ij} term to test the factor

Choice of sample size

a treat-ments	b blocks: 1	2	:	b
A (1)	y y	y y	:	y y
B (2)	y y	y y	:	y y
:	:	:	:	:
D (a)	y y	y y	:	y y

If: $Y_{ijk} = \mu + t_i + B_j + E_{ijk}$

$$F_{treat} = s^2_{treat}/s^2_E \sim F(\nu_1, \nu_2)$$

$$\Phi^2_t = (bn \cdot \Sigma_i t_i^2)/(a \cdot \sigma^2_E)$$

$$\nu_1 = a - 1 \text{ and } \nu_2 = abn - a - b + 1$$

If: $Y_{ijk} = \mu + t_i + B_j + TB_{ij} + E_{ijk}$

$$F_{treat} = s^2_{treat}/s^2_{TB} \sim F(\nu_1, \nu_2)$$

$$\Phi^2_t = (bn \cdot \Sigma_i t_i^2)/(a \cdot (\sigma^2_E + n\sigma^2_{TB}))$$

$$\Phi^2_t \simeq (b \cdot \Sigma_i t_i^2)/(a \cdot \sigma^2_{TB})$$

$$\nu_1 = a - 1 \text{ and } \nu_2 = (a-1)(b-1)$$

Incomplete block designs

Testing of 4 alternative extraction methods. Extraction of one production lasts 3 hours $\Longrightarrow$ only 3 methods can be tested on 1 day:

Design	Fine material	Normal material	2% additive Fine mat.	2% additive Norm. mat.
Day 1	X	X		X
Day 2	X	X		X
Day 3	X		X	X
Day 4		X	X	X
Day 5	X	X	X	
Day 6		X	X	X

$$Y_{ij} = \mu + \tau_j + D_i + E_{ij}$$

Is the design adequate.

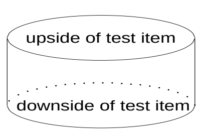
How could we improve the design. Which requirements should be made for the design.

1 day is an incomplete block: block size = 3

Small blocks, why ?

The smallest block size = 2

Example: Test item that can be treated on two sides with surface hardening



The intra-block variation is small for small blocks:
That is a physical fact that the experimenter can utilize: use small blocks!

A balanced incomplete block design

Four treatments, A, B, C and D. Two treatments per block.

4 treatments with block size 2	A	B	C	D
Item 1	X	X		
Item 2	X		X	
Item 3	X			X
Item 4		X	X	
Item 5		X		X
Item 6			X	X

Problem: If systematic difference between upside and downside treatment results. Can that be handled ? How ?

Other 'classical' examples of incomplete blocks

World Championship in football: 16 teams participate in 4 groups of 4 teams. In one group of 4 only 2 teams can be on the field at the same time (1 match = 1 block of size 2). 6 matches per group needed.

World Championship in speedway with 12 participants: Groups of 4 drivers compete at the same time.

Bridge tournament with 10 teams. In one 'round' 5 tables are used each with 2 teams. How many rounds are needed so that all 10 teams meet each other once.

Football tournament with 10 teams. In one 'round' 5 matches are played each with 2 teams. How many rounds are needed so that all 10 teams meet each other once.

How is the advantage of 'home matches' handled in practice.

A heuristic design (an inadequate design)

	Treatments			
Day	A	B	C	D
I	X	X	X	
II	X	X	X	
III		X	X	X
IV		X	X	X

$$Y_{ij} = \mu + \alpha_j + D_i + E_{ij}$$

α_j is the fixed factor effect (deterministic quantity)

D_i is the block effect. A random variable.

Estimate
fx : $\widehat{\alpha}_B - \widehat{\alpha}_A = (Y_{12} + Y_{22})/2 - (Y_{11} + Y_{21})/2$
or : $\widehat{\alpha}_B - \widehat{\alpha}_A = \overline{Y}_B - \overline{Y}_A$

Which one is best ? Depends on σ_E^2 and σ_D^2 .

Estimate
fx : $\widehat{\alpha}_B - \widehat{\alpha}_A = (Y_{12} + Y_{22})/2 - (Y_{11} + Y_{21})/2$
and : $\widehat{\alpha}_D - \widehat{\alpha}_A = (Y_{34} + Y_{44})/2 - (Y_{11} + Y_{21})/2$

Which one is the most precise ? Always $\widehat{\alpha}_B - \widehat{\alpha}_A$

Can the design be balanced, so that all comparisons are equally precise and independent of the actual blocks used ? (Yes)

Incomplete balanced block designs and some definitions

	Treatments			
Day	A	B	C	D
I	X	X		X
II	X	X	X	
III	X		X	X
IV		X	X	X

$$Y_{ij} = \mu + \alpha_j + D_i + E_{ij}$$

k	=	3	=	block size
a	=	4	=	number of treatments (some times called 't')
b	=	4	=	number of blocks
r	=	3	=	number of times each treatment is tried
λ	=	2	=	number of times any two treatments are in the same block = r·(k-1)/(a-1)
N	=	12	=	Total number of measurements = k·b = a·r

Exercise:

Design	Fine material	Normal material	2% additive Fine mat.	2% additive Norm. mat.
Day 1	X	X		
Day 2	X		X	
Day 3	X			X
Day 4		X	X	
Day 5		X		X
Day 6			X	X
Find k, a, b, r, λ and N for this design				

Data from incomplete balanced block design

Blocks (days)	Treatments A	B	C	D	T$_{i\cdot}$
I	52	–	75	57	184
II	–	87	86	53	226
III	54	68	69	–	191
IV	50	78	–	61	189
$T_{\cdot j}$	156	233	230	171	790

$N = t \cdot r = b \cdot k$
$t = a = 4$ (treatments)
$b = 4$ (blocks)
$k = 3$ (block size)
$r = 3$ (repeat. treat.)
$\lambda = 2$ (pairs in one block)

$Y_{ij} = \mu + \alpha_j + B_i + E_{ij}$

$\Sigma_{j=1}^{t} \alpha_j = 0$, $\Sigma_{i=1}^{b} B_i = 0$, $\mathsf{Var}\{E_{ij}\} = \sigma^2$

Computations for balanced incomplete block design

$Q_j = T_{\cdot j} - \frac{1}{k}\Sigma_{i=1}^{b} n_{ij} T_{i\cdot} \qquad n_{ij} = \begin{cases} 0 \text{ if cell } (i,j) \text{ is empty} \\ 1 \text{ if cell } (i,j) \text{ is not empty} \end{cases}$

$SSQ_\alpha = k \cdot \frac{Q_1^2 + Q_2^2 + \cdots + Q_t^2}{\lambda \cdot t}$

$SSQ_{blocks} = \frac{T_1^2 + T_2^2 + \cdots + T_b^2}{k} - \frac{T_{\cdot\cdot}^2}{N}$

$SSQ_{resid} = SSQ_{tot} - SSQ_\alpha - SSQ_{blocks}$

Expectations and variances of computed quantities - estimation

$\mathsf{E}\{Q_j\} = \frac{\lambda t}{k} \cdot \alpha_j \Rightarrow$ the estimate $\hat{\alpha}_j = \frac{Q_j}{\lambda t} \cdot k$

$\mathsf{Var}\{Q_j\} = \frac{\lambda(t-1)}{k} \cdot \sigma^2 \Rightarrow \mathsf{Var}\{\hat{\alpha}_j\} = \frac{t-1}{t^2} \cdot \frac{k}{\lambda}$

Treatment difference estimate is $\hat{\alpha}_i - \hat{\alpha}_j$ and

$\mathsf{Var}\{Q_i - Q_j\} = \frac{2\lambda t}{k} \cdot \sigma^2 \Rightarrow \mathsf{Var}\{\hat{\alpha}_i - \hat{\alpha}_j\} = \frac{2k}{\lambda t} \cdot \sigma^2$

$\hat{\mu} = \overline{Y}_{\cdot\cdot}$, $\mathsf{Var}\{\overline{Y}_{\cdot\cdot}\} = \sigma^2/N$

Treatment mean estimate: $[\hat{\mu} + \hat{\alpha}_j] = \overline{Y}_{\cdot\cdot} + \frac{Q_j}{\lambda t} \cdot k$

Variance of treatment mean estimate: $\mathsf{Var}\,[\hat{\mu} + \hat{\alpha}_j] = \sigma^2 \frac{k}{\lambda t}$

After-ANOVA tests based on the Q-values

Range-test (fx Newman-Keuls test and table VII):

$$\frac{\hat{\alpha}_{(max)}-\hat{\alpha}_{(min)}}{\hat{\sigma}\sqrt{k/\lambda t}} = \frac{Q_{(max)}-Q_{(min)}}{s_{resid}\sqrt{\lambda t/k}} \in q(\text{“number”}, f_{resid})$$

Contrast-test procedure:

$$contrast = [C] = Q_1 \cdot c_1 + Q_2 \cdot c_2 + \cdots + Q_t \cdot c_t$$

$$SSQ_{contrast} = \frac{k \cdot [C]^2}{\lambda t(c_1^2 + c_2^2 + \cdots + c_t^2)}$$

Analysis of data - example

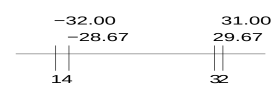
$\lambda = 2$, a = t = 4 , b = 4 , r = 3 , k = 3

$$\begin{aligned}
Q_1 &= 156 - \tfrac{1}{3}(184+191+189) &= -32.00 \\
Q_2 &= 233 - \tfrac{1}{3}(226+191+189) &= 31.00 \\
Q_3 &= 230 - \tfrac{1}{3}(184+226+191) &= 29.67 \\
Q_4 &= 171 - \tfrac{1}{3}(184+226+189) &= -28.67 \\
&\quad\quad\quad\quad\quad\quad\quad\quad \text{Sum} &= 0.00
\end{aligned}$$

$$\begin{aligned}
SSQ_{tot} &= \Sigma_{ij}\, y_{ij}^2 - T_{..}^2/N &= 1949.67 \\
SSQ_{treat} &= \frac{3 \cdot (32.00^2+31.00^2+29.67^2+28.67^2)}{2 \cdot 4} &= 1382.73 \\
SSQ_{blocks} &= \frac{184^2+226^2+191^2+189^2}{3} - \frac{790^2}{12} &= 369.67 \\
SSQ_{resid} &= SSQ_{tot} - SSQ_{treat} - SSQ_{blocks} &= 197.27
\end{aligned}$$

$$\begin{aligned}
f_{tot} &= 12 - 1 = 11 \\
f_{treat} &= 4 - 1 = 3 \\
f_{blocks} &= 4 - 1 = 3 \\
f_{resid} &= 11 - 3 - 3 = 5
\end{aligned}$$

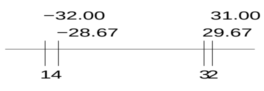
Newman-Keuls test for treatments using Q's



$$s_Q^2 = s_E^2 \cdot \frac{\lambda t}{k} = \frac{197.25}{5} \cdot \frac{2 \cdot 4}{3} = 105.20 = 10.26^2$$

$$\begin{aligned}
q(p,5)_{0.05} &= 3.64 \quad 4.60 \quad 5.22 \\
LSR = s_Q \cdot q(p,5)_{0.05} &= 37.25 \quad 47.20 \quad 53.56
\end{aligned}$$

$$
\begin{aligned}
|2-1| &\Rightarrow |31.00-(-32.00)| &&= 63.00 > 53.56 && \text{sign.}\\
|2-4| &\Rightarrow |31.00-(-28.67)| &&= 59.67 > 47.20 && \text{sign.}\\
|2-3| &\Rightarrow |31.00-29.67| &&= 1.33 < 37.25 && \text{not sign.}\\
&-\,-\,-\\
|3-1| &\Rightarrow |29.67-(-32.00)| &&= 61.67 > 47.20 && \text{sign.}\\
|3-4| &\Rightarrow |29.67-(-28.67)| &&= 58.33 > 37.25 && \text{sign.}\\
&-\,-\,-\\
|4-1| &\Rightarrow |28.67-(32.00)| &&= 3.33 < 37.25 && \text{not sign.}
\end{aligned}
$$



The Youden square (incomplete Latin square)

Construction of a Youden square design

Blocks (days)	Treatments A1	A2	B1	B2	$T_{i\cdot}$
I	?	–	?	?	
II	–	?	?	?	
III	?	?	?	–	
IV	?	?	–	?	
$T_{\cdot j}$					

Youden square design

Blocks (Days)	Treatments A1	A2	B1	B2	$T_{i\cdot}$
I	α	–	β	γ	
II	–	β	γ	α	
III	β	γ	α	–	
IV	γ	α	–	β	
$T_{\cdot j}$					

Data from Youden square experiment

Blocks (days)	Treatments A1	A2	B1	B2	$T_{i\cdot}$
I	52 (α)	–	75 (β)	57 (γ)	184
II	–	87 (β)	86 (γ)	53 (α)	226
III	54 (β)	68 (γ)	69 (α)	–	191
IV	50 (γ)	78 (α)	–	61 (β)	189
$T_{\cdot j}$	156	233	230	171	790

Analysis of Youden square

The data are the same as on slide 4.9 and the example primarily illustrates how the computations go.

$$T_\alpha = 52+53+69+78 = 252$$

$$T_\beta = 75+87+54+61 = 277$$

$$T_\gamma = 57+86+68+50 = 261$$

$$SSQ_{pos} = (252^2+277^2+261^2)/4 - 790^2/12 = 80.17$$

The other SSQ's : See slide 4.9 (the same data)

$SSQ_{treat} = 1382.73$

$SSQ_{blocks} = 369.67$

$SSQ_{tot} = 1949.67$

ANOVA for Youden Square

Source	SSQ	df	s^2	$E\{s^2\}$	F
Treatm.	1382.73	3	460.91	$\sigma^2 + c\cdot\phi_{treat}$	11.81
Blocks	369.67	3	123.22	–	–
Posit.	80.17	2	40.09	$\sigma^2 + 4\cdot\phi_{pos}$	(1.03)
Residual	117.10	3	39.03	σ^2	
Total	1949.67	12-1			

Example of computation for contrasts

Consider the Youden square slide 4.17.

k = 3, $\lambda = 2$, a = t = 4, b = 4, r = 3.

$$Q_{A1} = -32.00 \quad Q_{A2} = 31.00 \quad Q_{B1} = 29.67 \quad Q_{B2} = -28.67$$

$$\begin{array}{lllll} C_{A-B} & = & +Q_{A1} + Q_{A2} & -Q_{B1} - Q_{B2} & = -2.00 \\ C_{A1-A2} & = & +Q_{A1} - Q_{A2} & & = -63.00 \\ C_{B1-B2} & = & & +Q_{B1} - Q_{B2} & = 58.33 \end{array}$$

$$\begin{array}{lcccr} SSQ_{A-B} & = & (-2.00)^2 \cdot \frac{k}{\lambda t(1^2+1^2+1^2+1^2)} & = & 0.38 \\ SSQ_{A1-A2} & = & (-63.00)^2 \cdot \frac{3}{2 \cdot 4(1^2+1^2)} & = & 744.19 \\ SSQ_{B1-B2} & = & 58.33^2 \cdot \frac{3}{2 \cdot 4(1^2+1^2)} & = & 638.16 \\ & & \text{Sum} & = & 1382.73 \end{array}$$

The sums of squares for the 3 orthogonal contrasts add up to the total sum of squares between treatments.

Each of these sums of squares have 1 degree of freedom and can be tested against the residual sum of squares.

Tables of balanced incomplete block designs

A, B, C, ...	=	Treatments
a	=	Number of treatments (often called 't')
b	=	Number of blocks
r	=	Number replications of each treatment
k	=	Block size
N	=	ar = bk = total number of measurements
λ	=	number of times any two treatments occur in the same block =r(k−1)/(a−1)
α, β, γ, ...	=	'positions' within block for incomplete Latin squares (Youden squares)

Balanced designs for 'a' treatments with block size k = 2 consist of all possible combinations of two treatments giving a(a-1)/2 blocks of 2

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 2	a = 3	b = 3	r = 2	$\lambda = 1$	N = 6
Symmetrical. A Youden square					

Treat-ments	Blocks 1	2	3
A	α		β
B	β	α	
C		β	α

Blocksize	Treatments	Blocks	Replications	Pairings	Total design
k = 3	a = 4	b = 4	r = 3	$\lambda = 2$	N = 12
Symmetrical. A Youden square					

Treat-ments	Blocks 1	2	3	4
A	α	β	γ	
B	β	γ		α
C	γ		α	β
D		α	β	γ

4.24

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 3	a = 5	b = 10	r = 6	$\lambda = 3$	N = 30
All combinations of 3 treatments among 5					

Treat-	Blocks									
ments	1	2	3	4	5	6	7	8	9	10
A	x	x	x	x	x	x				
B	x	x	x				x	x	x	
C	x			x	x		x	x		x
D		x		x		x	x		x	x
E			x		x	x		x	x	x

4.25

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 3	a = 6	b = 10	r = 5	$\lambda = 2$	N = 30
10 out of 20 possible combinations of 3 treatments among 6 (reduced)					

Treat-	Blocks									
ments	1	2	3	4	5	6	7	8	9	10
A	x			x		x	x	x		
B		x			x		x	x	x	
C			x			x		x	x	x
D	x	x	x				x			x
E	x		x	x	x				x	
F		x		x	x	x				x

4.26

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 3	a = 7	b = 7	r = 3	$\lambda = 1$	N = 21
7 out of 35 possible combinations of 3 treatments among 7 (reduced) Symmetrical. Youden square.					

Treat-	Blocks						
ments	1	2	3	4	5	6	7
A	α	β	γ				
B			β	α	γ		
C	β			γ		α	
D			α			β	γ
E	γ				α		β
F		α			β	γ	
G		γ		β			α

4.27

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 3	a = 9	b = 12	r = 3	$\lambda = 1$	N = 36
12 out of 84 possible combinations of 3 treatments among 9 (reduced)					

Treat-	Blocks											
ments	1	2	3	4	5	6	7	8	9	10	11	12
A	x			x			x			x		
B	x				x			x			x	
C	x					x			x			x
D		x		x					x		x	
E		x			x		x					x
F		x				x		x		x		
G			x	x				x				x
H			x		x				x	x		
I			x			x	x				x	

4.28

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 4	a = 5	b = 5	r = 4	$\lambda = 3$	N = 20
Symmetrical. Youden square					

Treat-	Blocks				
ments	1	2	3	4	5
A	α	β	γ	δ	
B	β	γ	δ		α
C	γ	δ		α	β
D	δ		α	β	γ
E		α	β	γ	dy

4.29

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 4	a = 7	b = 7	r = 4	$\lambda = 2$	N = 28
7 out of 35 possible combinations of 4 treatments among 7 (reduced) Symmetrical. Youden square.					

Treat-	Blocks						
ments	1	2	3	4	5	6	7
A		α	β	γ			δ
B	α	β		δ		γ	
C	β		γ			δ	α
D		δ	α		γ	β	
E				β	δ	α	γ
F	γ		δ	α	β		
G	δ	γ			α		β

4.30

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 4	a = 8	b = 14	r = 7	$\lambda = 3$	N = 56
14 out of 70 possible combinations of 4 treatments among 8 (reduced)					

Treat-	Blocks													
ments	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	x	x	x	x					x	x	x			
B	x	x					x	x	x				x	x
C	x		x			x		x		x		x		x
D	x			x		x	x				x	x	x	
E					x	x	x	x	x	x	x			
F			x	x	x	x			x				x	x
G		x		x	x		x			x		x		x
H		x	x		x			x			x	x	x	

4.31

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 5	a = 6	b = 6	r = 5	$\lambda = 4$	N = 30
Symmetrical. Youden square.					

Treat-	Blocks					
ments	1	2	3	4	5	6
A	α	β	γ	δ	ϵ	
B	β	γ	δ	ϵ		α
C	γ	δ	ϵ		α	β
D	δ	ϵ		α	β	γ
E	ϵ		α	β	γ	δ
F		α	β	γ	δ	ϵ

4.32

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 5	a = 11	b = 11	r = 5	$\lambda = 2$	N = 30
11 out of 462 possible combinations of 5 treatments among 11 (reduced)					
Symmetrical. Youden square.					

Treat-ments	Blocks 1	2	3	4	5	6	7	8	9	10	11
A	α			ϵ		δ	γ	β			
B		α			ϵ		δ	γ	β		
C			α			ϵ		δ	γ	β	
D				α			ϵ		δ	γ	β
E	β				α			ϵ		δ	γ
F	γ	β				α			ϵ		δ
G	δ	γ	β				α			ϵ	
H		δ	γ	β				α			ϵ
I	ϵ		δ	γ	β				α		
J		ϵ		δ	γ	β				α	
K			ϵ		δ	γ	β				α

4.33

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 6	a = 7	b = 7	r = 6	$\lambda = 5$	N = 42
Symmetrical. Youden square.					

Treat-ments	Blocks 1	2	3	4	5	6	7
A	α		ϕ	ϵ	δ	γ	β
B	β	α		ϕ	ϵ	δ	γ
C	γ	β	α		ϕ	ϵ	δ
D	δ	γ	β	α		ϕ	ϵ
E	ϵ	δ	γ	β	α		ϕ
F	ϕ	ϵ	δ	γ	β	α	
G		ϕ	ϵ	δ	γ	β	α

4.34

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 6	a = 9	b = 12	r = 8	$\lambda = 5$	N = 72
12 out of 84 possible combinations of 6 treatments among 9 (reduced)					

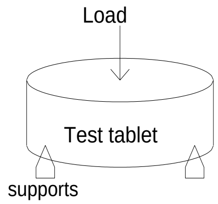
Treat-ments	Blocks 1	2	3	4	5	6	7	8	9	10	11	12
A		x	x		x	x		x	x		x	x
B		x	x	x		x	x		x	x		x
C		x	x	x	x		x	x		x	x	
D	x		x		x	x	x	x		x		x
E	x		x	x		x		x	x	x	x	
F	x		x	x	x		x		x		x	x
G	x	x			x	x	x		x	x	x	
H	x	x		x		x	x	x			x	x
I	x	x		x	x			x	x	x		x

4.35

Block size	Treatments	Blocks	Replications	Pairings	Total design
k = 6	a = 11	b = 11	r = 6	$\lambda = 3$	N = 66
11 out of 462 possible combinations of 6 treatments among 11 (reduced)					
Symmetrical. Youden square					

Treat-ments	Blocks 1	2	3	4	5	6	7	8	9	10	11
A		ϕ	ϵ		δ				γ	β	α
B	α		ϕ	ϵ		δ				γ	β
C	β	α		ϕ	ϵ		δ				γ
D	γ	β	α		ϕ	ϵ		δ			
E		γ	β	α		ϕ	ϵ		δ		
F			γ	β	α		ϕ	ϵ		δ	
G				γ	β	α		ϕ	ϵ		δ
H	δ				γ	β	α		ϕ	ϵ	
I		δ				γ	β	α		ϕ	ϵ
J	ϵ		δ				γ	β	α		ϕ
K	ϕ	ϵ		δ				γ	β	α	

An example with many issues



Response : Strength of tablet (load to breakage)

1 factor = Humidity in powder for tablets

Other possible factors: Load during production
Time of pressing in production
Size distribution in powder
etc

Practical problems

One test item can only be used once

The testing of one item takes time (is expensive)

How do we find a representative collection of independent tablets (randomization)

Other problems about sampling and preparation in lab

Design problems

Approximately, how large is the load to be measured
(select a suited apparatus to do the tests)

Which humidity percentages are relevant
(fx 5% $-$ 50%)

What is an interesting difference in load to detect
(fx $\Delta = \pm 12$ g)

The anticipated measurement uncertainty variance
(fx $\sigma^2 \simeq (15\text{ g})^2$) (measurement + tablet variation)

These questions must be answered before the experiment is started

Sources of uncertainty

Temperature in laboratory

Different handling by different operators

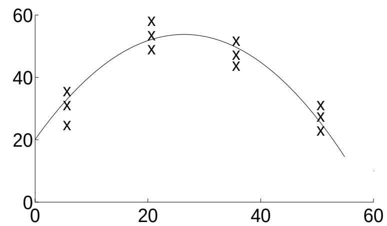
Day-to-day variation of measurement devices

Variation in the experimental setup (geometry)

The order of magnitude of these variations must be known or assessed before the experiment is started

Design must be based on knowledge

Guess (or study) how an experiment may turn out is a possible (good) way:



Plot how you think (or hope) the data will turn out.

Do you believe, that you will find what you are looking for: The optimal humidity for obtaining a high strength, fx.

How will the ANOVA look when I have collected the data?

Design	Humidity			
	5%	20%	35%	50%
Day 1	x	x	x	x
Day 2	x	x	x	x
...	...	...	...	...
Day b	x	x	x	x

Source of var.	SSQ	d.f. for design with b blocks
Humidity	SSQ_{treat}	$\nu_1 = a - 1 = 3$
Days (blocks)	SSQ_{blocks}	$\nu_3 = b - 1 = b - 1$
Residual	SSQ_{resid}	$\nu_2 = (a-1)(b-1) = 3(b-1)$
Total	SSQ_{tot}	$\nu_{tot} = ab - 1 = 4b - 1$

Determine the necessary number of days (blocks)

$$Y_{ij} = \mu + \tau_j + D_i + E_{ij}$$

Possible sizes of τ_j to detect				
$\tau_j =$	-12	$+12$	$+12$	-12
$x_j =$	5	20	35	50

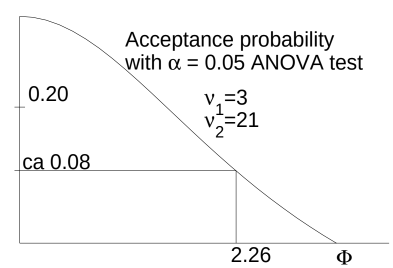
Compute $\Phi^2 = \frac{b}{a \cdot \sigma^2} \Sigma_j \tau_j^2 = \frac{b}{4 \cdot 15^2}(576) = 0.64 \cdot b$

$$\Phi = 0.80\sqrt{b}$$

Try fx $b = 8 \rightarrow \Phi = 2.26$, $\nu_1 = 3$, $\nu_2 = 3 \cdot 7 = 21$

Is there a reasonable probability to detect the prescribed differences?

Look up probability of acceptance page 648:

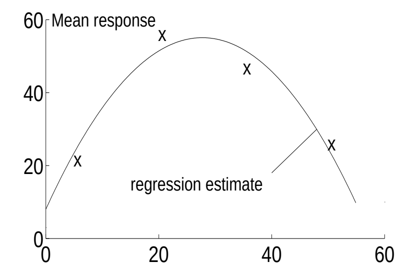


Looks reasonable. The chance of overlooking the above τ's is less than 10% (lucky punch). $b = 8$ could be worthwhile trying.

Analysis of the data from the experiment

Day	5%	20%	35%	50%	Sums
1					135
2					90
3	Data from				231
4	experiment				116
5					161
6	$SSQ_{resid} = 4910$				114
7					150
8					214
Sums	175	450	376	210	1211

Source of var.	SSQ	d.f.	s^2	EMS	F
Humidity	6496	3	2165	$\sigma^2 + 8 \cdot \phi_\tau$	9.3
Days (blocks)	4260	7	609	$(\sigma^2 + 4 \cdot \sigma_D^2)$	(2.6)
Residual	4910	21	234	σ^2	
Total	15666	31			



Estimates:

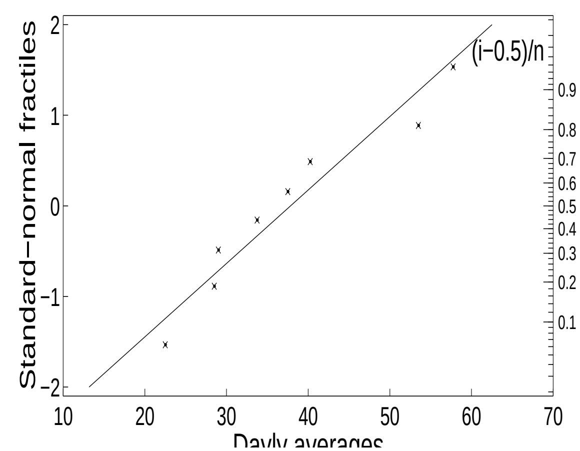
$$
\begin{array}{lllll}
\overline{y}_j = & 21.9 & 56.3 & 47.0 & 26.3 \\
\hat{\tau}_j = & -15.96 & 18.40 & 9.15 & -11.59 \\
\widehat{\mu} = & 37.84 & & & \\
\widehat{\sigma}^2 = & & 234 = 15.3^2 & & \\
(\widehat{\sigma}_D^2 = & (609 - 234)/4 = 93.8 = 9.7^2) & & &
\end{array}
$$

Regression function estimate $(5 \leq x_j \leq 50)$:

$$\widehat{Y}_{ij} = 8.0396 + 3.3946 \cdot x_j - 0.0612 \cdot x_j^2$$

Further analysis of days (blocks)

Draw 'normal probability plot' for daily averages, fx:



Newman-Keuls test for blocks:

$s_{mean,block} = s_{resid}/\sqrt{4} = \sqrt{234}/\sqrt{4} = 7.65$

$q(8,21)_{0.05} \simeq 4.75$, $LSR = 7.65 \cdot 4.75 = 36.34$

$\overline{Y}_{(8)} - \overline{Y}_{(1)} = 231/4 - 90/4 = 35.25 \Longrightarrow$ not sign.

Newman-Keuls test shows that the difference between the largest and the smallest block average is not unusually large (close to, however!). Thus no grouping of the days can hardly be identified.

Construction of the normal probability plot

Use as example the block averages computed from slide 5.9. $n = 8$ observations.

Averages sorted (x)	Order i	p = (i−0.5)/n	Normal quantile
22.50	1	0.0625	−1.53
28.50	2	0.1875	−0.89
29.00	3	0.3125	−0.49
33.75	4	0.4375	−0.16
37.50	5	0.5625	+0.16
40.25	6	0.6875	+0.49
53.50	7	0.8125	+0.89
57.75	8	0.9375	+1.53

Plot the quantiles against the averages.

Average of the averages is $\overline{x} = 37.84$

The standard deviation of the averages is $s_x = 12.33$

The suggested normal distribution is represented by the straight line through the point $\overline{x} = 37.84$ and slope $1/s_x = 1/12.33$.

One can draw the line through the two points
$(\overline{x} - 2s_x, -2)$ and $(\overline{x} + 2s_x, +2)$

The plot is shown on page 5.12. A normal probability scale is added to the right there.

The same problem with incomplete blocks

Only 3 measurements per day: block size = 3.

If the same precision as required in the previous example (8 complete blocks of size 4) is wanted the number of blocks of size 3 must be $b = 8 \cdot 4/3 \simeq 11$.

Choose fx 12 blocks organized as 3 balanced incomplete block designs or Youden squares.

Positions within a day are α, β or γ: with possible effects p_1, p_2, p_3

Day=block	5%	20%	30%	50%
1	α	γ	β	
2	γ	β		α
3	β		α	γ
4		α	γ	β
5		γ	α	β
6	α		β	γ
7	γ	β		α
8	β	α	γ	
9	β		γ	α
10		γ	α	β
11	α	β		γ
12	γ	α	β	

In this design k = 3, a = 4, b = 12, r = 9, $\lambda = 6$

Taking positions into account the model could be

$$Y_{ij} = \mu + \tau_i + B_j + p_k + E_{ijk}$$

Orthogonal polynomials

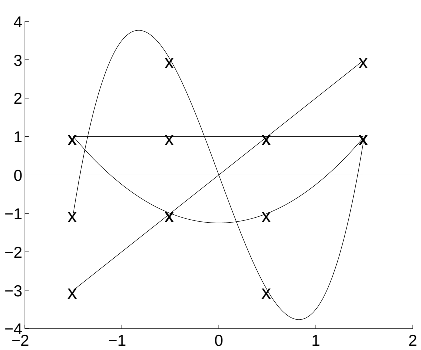
Data from slide 1.33 again:

	Concentration			
	5%	7%	9%	11%
	3.5	6.0	4.0	3.1
	5.0	5.5	3.9	4.0
	2.8	7.0	4.5	2.6
	4.2	7.2	5.0	4.8
	4.0	6.5	6.0	3.5
Sum	19.5	32.2	23.4	18.0

Model : $Y_{ij} = \mu + \tau_j + E_{ij}$

ANOVA of response				
Source	SSQ	d.f.	s^2	F
Concentration	24.35	4−1	8.1167	12.41
Residual	10.46	16	0.6538	(sign)
Total	34.81	20−1		

System of orthogonal polynomials, 4 points



Weights and expressions for orthogonal polynomials

4 - point polynomial weights

z	-1.5	-0.5	0.5	1.5
$P_0(z)$	1	1	1	1
$P_1(z)$	-3	-1	1	3
$P_2(z)$	1	-1	-1	1
$P_3(z)$	-1	3	-3	1

$$P_0(z) = 1$$

$$P_1(z) = \lambda_1 \cdot z \qquad , \lambda_1 = 2$$

$$P_2(z) = \lambda_2 \cdot [z^2 - \tfrac{(a^2-1)}{12}] \qquad , \lambda_2 = 1$$

$$P_3(z) = \lambda_3 \cdot [z^3 - z \cdot \tfrac{(3a^2-7)}{20}] \quad , \lambda_3 = 10/3$$

$$\sum_{j=1}^{a} P_\ell(z_j) \cdot P_m(z_j) = 0 \quad \text{for all } \ell \neq m$$

Orthogonal polynomials continued:

$$
\begin{aligned}
P_0(z) &= 1 \\
P_1(z) &= \lambda_1 \cdot z \\
P_2(z) &= \lambda_2 \cdot [z^2 - \tfrac{(a^2-1)}{12}] \\
P_3(z) &= \lambda_3 \cdot [z^3 - z \cdot \tfrac{(3a^2-7)}{20}] \\
P_4(z) &= \lambda_4 \cdot [z^4 - \tfrac{z^2}{14}(3a^2-13) + \tfrac{3}{560}(a^2-1)(a^2-9)] \\
P_5(z) &= \lambda_5 \cdot [z^5 - \tfrac{5z^3}{18}(a^2-7) + \tfrac{z}{1008}(15a^4 - 230a^2 + 407)]
\end{aligned}
$$

Weights for higher order orthogonal polynomials

a	Polynomial	x_1	x_2	x_3	x_4	x_5	x_6	x_7	λ
3	Linear	-1	0	1					1
	Quadratic	1	-2	1					3
4	Linear	-3	-1	1	3				2
	Quadratic	1	-1	-1	1				1
	Cubic	-1	3	-3	1				10/3
5	Linear	-2	-1	0	1	2			1
	Quadratic	2	-1	-2	-1	2			1
	Cubic	-1	2	0	-2	1			5/6
	Quartic	1	-4	6	-4	1			35/12
6	Linear	-5	-3	-1	1	3	5		2
	Quadratic	5	-1	-4	-4	-1	5		3/2
	Cubic	-5	7	4	-4	-7	5		5/3
	Quartic	1	-3	2	2	-3	1		7/12
	5th degr.	-1	5	-10	10	-5	1		21/10
7	Linear	-3	-2	-1	0	1	2	3	1
	Quadratic	5	0	-3	-4	-3	0	5	1
	Cubic	-1	1	1	0	-1	-1	1	1/6
	Quartic	3	-7	1	6	1	-7	3	7/12
	5th degr.	-1	4	-5	0	5	-4	1	7/20

The values x_1, x_2,...,x_a are equi-spaced with difference Δ_x between values and average value $\overline{x}$. Then

$$z_j = (x_j - \overline{x})/\Delta_x \quad ; \quad j = 1, 2, \ldots, a$$

Method:

$$\boxed{Y_{ij} = \beta_0 + \beta_1 \cdot x_j + \beta_2 \cdot x_j^2 + \beta_3 \cdot x_j^3 + E_{ij}}$$

Compute standardized variable $z_j = (x_j - \overline{x})/\Delta_x$, for $j = 1, 2, .., a$, where $\overline{x}$ is the mean of the x values, and Δ_x is the spacing between the x values.

Rewrite model using the orthogonal polynomials:

$$Y_{ij} = \alpha_0 + \alpha_1 \cdot P_1(z_j) + \alpha_2 \cdot P_2(z_j) + \alpha_3 \cdot P_3(z_j) + E_{ij}$$

Estimate coefficients, using treatment totals $T_{\cdot j}$:

$$\widehat{\alpha}_\ell = [\sum_{j=1}^{a} T_{\cdot j} \cdot P_\ell(z_j)]/(n \cdot \sum_{j=1}^{a} [P_\ell(z_j)]^2)$$

Compute sums of squares:

$$SSQ_\ell = [\sum_{j=1}^{a} T_{\cdot j} \cdot P_\ell(z_j)]^2/(n \cdot \sum_{j=1}^{a} [P_\ell(z_j)]^2)$$

Note that $C_\ell = [\Sigma_{j=1}^{a} T_{\cdot j} \cdot P_\ell(z_j)]$ for $\ell > 0$ is a <u>contrast</u>, and the contrasts are orthogonal. Therefore

$$\boxed{SSQ_{treatments} = \Sigma_{l=1}^{a-1} SSQ_\ell}$$

Numerical example from slide 6.1 (and 1.33)

Example $P_2 : C_2 = 19.5 - 32.2 - 23.4 + 18.0 = -18.1$

$SSQ_2 = (-18.1)^2/(5[1^2 + (-1)^2 + (-1)^2 + 1^2]) = 16.38$

$\widehat{\alpha}_2 = (-18.1)/(5[1^2 + (-1)^2 + (-1)^2 + 1^2]) = -0.905$

Computations for $a = 4$									
	$x_j =$	5	7	9	11	$\overline{x} = 8,\ \Delta_x = 2$			
	$z_j \rightarrow$	-1.5	-0.5	0.5	1.5	$z = (x-8)/2$			
ℓ	Totals $T_j \rightarrow$	19.5	32.2	23.4	18.0	C_ℓ	SSQ_ℓ	$\widehat{\alpha}_\ell$
0	$P_0 \rightarrow$	1	1	1	1	93.1	$-$	4.655
1	$P_1 \rightarrow$	-3	-1	1	3	-13.3	1.77	-0.133
2	$P_2 \rightarrow$	1	-1	-1	1	-18.1	16.38	-0.905
3	$P_3 \rightarrow$	-1	3	-3	1	24.9	6.20	0.249
Total							24.35	

ANOVA in detail with orthogonal polynomials

ANOVA of response				
Source	SSQ	d.f.	s^2	F
1. order polyn.	1.77	1	1.77	2.71
2. order polyn.	16.38	1	16.38	25.05
3. order polyn.	6.20	1	6.20	9.48 (sign)
Residual	10.46	16	0.6538	
Total	34.81	20−1		

The 3rd order term is significant. The polynomium probably has degree 3 (at least).

Test successively with higest order first. When a significant order is found the polynomial and all the lower order polynomials are retained in the model.

Contrasts with orthogonal polynomials

Factor X	Factor B B1	B2
10%	$T_{10,1}$	$T_{10,2}$
15%	$T_{15,1}$	$T_{15,2}$
20%	$T_{20,1}$	$T_{20,2}$

Totals	$T_{10,1}$	$T_{10,2}$	$T_{15,1}$	$T_{15,2}$	$T_{20,1}$	$T_{20,2}$	Effect
Main effects	-1	-1	0	0	$+1$	$+1$	X-linear, X_L
	$+1$	$+1$	-2	-2	$+1$	$+1$	X-quadr., X_Q
	-1	$+1$	-1	$+1$	-1	$+1$	B main
Inter-actions	$+1$	-1	0	0	-1	$+1$	$X_L \times B$
	-1	$+1$	$+1$	-1	-1	$+1$	$X_Q \times B$

The two last contrasts are constructed by multiplication of the coefficients of the corresponding main effects

Contrasts for two-factor orthogonal polynomials

Factor X	Factor Z 24^{o}C	30^{o}C
10%	$T_{10,24}$	$T_{10,30}$
15%	$T_{15,24}$	$T_{15,30}$
20%	$T_{20,24}$	$T_{20,30}$

Totals	$T_{10,24}$	$T_{10,30}$	$T_{15,24}$	$T_{15,30}$	$T_{20,24}$	$T_{20,30}$	Effect
Main effects	-1	-1	0	0	$+1$	$+1$	X_L
	$+1$	$+1$	-2	-2	$+1$	$+1$	X_Q
	-1	$+1$	-1	$+1$	-1	$+1$	Z_L
Inter-actions	$+1$	-1	0	0	-1	$+1$	$X_L \times Z_L$
	-1	$+1$	$+2$	-2	-1	$+1$	$X_Q \times Z_L$

One can then test all coefficients successively (fx e, d, c, b, a) in the model:

$$Y = \mu + a \cdot x + b \cdot x^2 + c \cdot z + d \cdot x \cdot z + e \cdot x^2 \cdot z + \epsilon$$

Orthogonal regression with x-factor at k levels - brief theory

Consider the balanced analysis of variance table :

x_1	x_2	...	x_k
Y_{11}	Y_{12}	...	Y_{1k}
Y_{21}	Y_{22}	...	Y_{2k}
:	:	:	:
Y_{n1}	Y_{n2}	...	Y_{nk}

where $k \geq 3$. We want to estimate (as an example):

$$Y_{ij} = \mu + \alpha_1 \cdot P_1(x_j) + \alpha_2 \cdot P_2(x_j) + E_{ij}$$

where $P_1(.)$ and $P_2(.)$ are some functions of the regression variable x (polynomials or any other functions).

Least squares estimation

The residual SSQ for the parameters $(\mu, \alpha_1, \alpha_2)$ is

$$SSQ_{res} = \sum_{j=1}^{k} \sum_{i=1}^{n} (Y_{ij} - \mu - \alpha_1 \cdot P_1(x_j) - \alpha_2 \cdot P_2(x_j))^2$$

We estimate the regression model such that SSQ_{res} is minimized (least squares),

and we therefore require that the partial derivatives are zero :

$$\partial(SSQ_{res})/\partial\mu = -2 \sum_{j=1}^{k} \sum_{i=1}^{n} (Y_{ij} - \mu - \alpha_1 \cdot P_1(x_j) - \alpha_2 \cdot P_2(x_j)) = 0$$

$$\partial(SSQ_{res})/\partial\alpha_1 = -2 \sum_{j=1}^{k} \sum_{i=1}^{n} P_1(x_j)(Y_{ij} - \mu - \alpha_1 \cdot P_1(x_j) - \alpha_2 \cdot P_2(x_j)) = 0$$

$$\partial(SSQ_{res})/\partial\alpha_2 = -2 \sum_{j=1}^{k} \sum_{i=1}^{n} P_2(x_j)(Y_{ij} - \mu - \alpha_1 \cdot P_1(x_j) - \alpha_2 \cdot P_2(x_j)) = 0$$

Use of orthogonal polynomials:

Choose the functions $P_1(.)$ and $P_2(.)$ such that

$$\sum_{j=1}^{k} P_1(x_j) = 0 \ , \ \sum_{j=1}^{k} P_2(x_j) = 0$$

$$\text{and} \ \sum_{j=1}^{k} P_1(x_j) P_2(x_j) = 0 \ \text{(i.e. orthogonal)}$$

then the solutions to the estimation equations are :

$$\widehat{\mu} = \sum_{j=1}^{k} \sum_{i=1}^{n} Y_i/(n \cdot k) = T_{..}/(n \cdot k) = \overline{Y}_{..}$$

$$\widehat{\alpha}_1 = \sum_{j=1}^{k} [T_{.j} \cdot P_1(x_j)]/(n \cdot \sum_{j=1}^{k} [P_1(x_j)]^2)$$

$$\widehat{\alpha}_2 = \sum_{j=1}^{k} [T_{.j} \cdot P_2(x_j)]/(n \cdot \sum_{j=1}^{k} [P_2(x_j)]^2)$$

where $T_{.j} = \Sigma_{i=1}^{n} Y_{ij}$ are the column totals.

If we introduce

$$SSQ(P_1) = \sum_{j=1}^{k} [T_{.j} \cdot P_1(x_j)]^2/(n \cdot \sum_{j=1}^{k} [P_1(x_j)]^2)$$

and similarly for $SSQ(P_2)$, it is easy also to show that

$$SSQ_{res} = \sum_{j=1}^{k} \sum_{i=1}^{n} (Y_{ij} - \mu)^2 - SSQ(P_1) - SSQ(P_2)$$

Note that, since $\Sigma_{j=1}^{k} P_1(x_j) = 0$, $\Sigma_{j=1}^{k} [T_{.j} \cdot P_1(x_j)]$ is a contrast with sum of squares $SSQ(P_1)$ which exactly is the part of the variation between the levels of x explained by the function $P_1(.)$ and similarly for $SSQ(P_2)$.

The example is easily generalized to more orthogonal functions than 2 (in fact to k-1 functions).

What if the model is a two-way model?

Experiment with an additive x on n batches					
Batch	$x_1 = 2\%$	$x_2 = 4\%$	$x_3 = 6\%$	$x_4 = 8\%$	$T_{1.}$
Batch 1	Y_{11}	Y_{12}	Y_{13}	Y_{14}	$T_{1.}$
Batch 2	Y_{21}	Y_{22}	Y_{13}	Y_{24}	$T_{2.}$
:		:	:	:	:
Batch n	Y_{n1}	Y_{n2}	Y_{n3}	Y_{n4}	$T_{n.}$
Totals	$T_{.1}$	$T_{.2}$	$T_{.3}$	$T_{.4}$	$T_{..}$

SSQ$_{Batches}$: $\Sigma_i T_{i.}^2/4 - T_{..}^2/(4n)$ (df=$n-1$)

SSQ$_{Additive}$: $\Sigma_j T_{.j}^2/n - T_{..}^2/(4n)$ (df=$4-1=3$)

SSQ$_{Total}$: $\Sigma_j \Sigma_j T_{ij}^2 - T_{..}^2/(4n)$ (df=$4n-1$)

SSQ$_{Residual}$: SSQ$_{Total}$ - SSQ$_{Batches}$ - SSQ$_{Additive}$

Split up the variance between concentration levels using orthogonal contrasts

Construct (4 - 1) orthogonal functions (fx polynomials), $P_1(x)$, $P_2(x)$ and $P_3(x)$, such that for all $\ell \neq m$

$$\sum_x P_\ell(x) = 0 \quad \text{and} \quad \sum_x P_\ell(x) \cdot P_m(x) = 0$$

then

$$SSQ_{Additive} = SSQ(P_1) + SSQ(P_2) + SSQ(P_3),$$

each with 1 degree of freedom

Sample size determination in general

Sample size in fixed effect model - exact method

The sample test quantity in the one-way ANOVA is

$$F_{sample} = \frac{n \cdot \Sigma_j (\overline{X}_{.j} - \overline{X}_{..})^2/(k-1)}{\Sigma_i \Sigma_j (\overline{X}_{ij} - \overline{X}_{.j})^2/(k(n-1))} \in F(k-1, k(n-1), \gamma^2(n))$$

where $F(.,.,.)$ denotes the non-central F-distribution with

$k-1$ and $k(n-1)$ degrees of freedom and non-centrality parameter

$$\gamma^2(n) = n \sum_i \tau_i^2/\sigma_E^2$$

which for $\gamma^2(n) = 0$ corresponds to the usual F-distribution.

Test $\gamma^2(n) = 0$ with level of significance α and require that the acceptance probability for a certain $\gamma^2(n) > 0$ is at most β (or that the power is at least $1-\beta$).

The p-critical value for the non-central F-distribution is (in the usual way) denoted by $F(\nu_1, \nu_2, \gamma^2(n))_p$ (p is upper tail probability).

The probability of acceptance is:

$$\beta(\gamma^2(n)) = P_r\{F_{sample} \leq F(\nu_1, \nu_2, 0)_\alpha\}$$

and our requirement is met if

$$F(\nu_1, \nu_2, 0)_\alpha \leq F(\nu_1, \nu_2, \gamma^2(n))_{1-\beta}$$

By trying different n values using $\nu_1 = k-1$ and $\nu_2 = k(n-1)$ and the corresponding $\gamma^2(n)$ the lowest n satisfying this inequality is the necessary sample size.

In order to do so a computer program is needed which can calculate the non-central F-distribution. All modern statistical programs can do it.

Take the example from slide 2.25 again: $\sigma_E^2 = 1.5^2$ and $\tau = [-2, 0, 2]$ and require a test with $\alpha = 0.05$ and probability of acceptance for this τ at most $\beta = 0.20$. Use $\gamma^2(n) = n \cdot \Sigma_i \tau_i^2 / \sigma_E^2 = n \cdot 8/2.25$:

n	ν_1	ν_2	$\gamma^2(n))$	$F(\nu_1, \nu_2, 0)_{0.05}$	$\beta(\gamma^2(n))$	$F(\nu_1, \nu_2, \gamma^2)_{0.80}$
2	2	3	7.11	9.55	0.711	1.97
3	2	6	10.67	5.14	0.392	3.16
4	2	9	14.22	4.26	0.185	4.44
5	2	12	17.78	3.89	0.079	5.77
6	2	15	21.33	3.68	0.031	7.15
7	2	18	24.89	3.55	0.012	8.55
8	2	21	28.44	3.47	0.004	9.98
9	2	24	32.00	3.40	0.001	11.43
10	2	27	35.56	3.35	0.000	12.90

The $\beta(\gamma^2(n))$ column is the probability of acceptance for the sample size n, $\sigma_E^2 = 1.5^2$ and $\tau = [-2, 0, 2]$. It decreases and must be at most 0.20 in our example.

The inequality is satisfied for $n \geq 4$; choose $n = 4$.

In the above example we also found by using the (not very detailed) graphs in the textbook, that we would need $n = 4$.

If, for example, $\beta \leq 0.10$ is required, $n = 5$ is chosen.

Sample size in random effect model - exact method

The test quantity is

$$F_{sample} = \frac{n \cdot \Sigma_j (\overline{X}_{.j} - \overline{X}_{..})^2/(k-1)}{\Sigma_i \Sigma_j (\overline{X}_{ij} - \overline{X}_{.j})^2/(k(n-1))} \in \lambda^2(n) \cdot F(k-1, k(n-1))$$

i.e. a usual F-distribution with scale parameter

$$\lambda^2(n) = (n \cdot \sigma_B^2 + \sigma_E^2)/\sigma_E^2$$

Test $\sigma_B^2 = 0$ with level of significance α and require that the acceptance probability for a certain σ_B^2 is at most β.

The p-critical value for the usual F-distribution is (as usual) denoted by $F(\nu_1, \nu_2)_p$ (p is upper tail probability).

The probability of acceptance is:

$$\beta(\lambda^2(n)) = P_r\{F_{sample} \leq F(\nu_1, \nu_2, 0)_\alpha\}$$

Our requirement is met if

$$F(\nu_1, \nu_2)_\alpha \leq \lambda^2(n) \cdot F(\nu_1, \nu_2)_{1-\beta} = \lambda^2(n)/F(\nu_2, \nu_1)_\beta$$

By trying different n values using $\nu_1 = k - 1$ and $\nu_2 = k(n-1)$ the lowest n satisfying this inequality is the necessary sample size.

With fx $\alpha = 0.05$ and $\beta = 0.10$ we can easily determine n using the standard 0.05-critical and the 0.10-critical values F-tables.

Take the example from slide 2.29 again: $\sigma_E^2 = 1.5^2$ and $\sigma_B^2 = 2.0^2$ and require a test with $\alpha = 0.05$ and probability of acceptance for $\sigma_B^2 = 2.0$ at most $\beta = 0.10$.

Use $\lambda^2(n) = (n \cdot \sigma_B^2 + \sigma_E^2)/\sigma_E^2$

n	ν_1	ν_2	$\lambda^2(n)$	$F(\nu_1,\nu_2)_{0.05}$	$\beta(\lambda^2(n))$	$F(\nu_2,\nu_1)_{0.10}$	$\frac{\lambda^2}{F(\nu_2,\nu_1)_{0.10}}$
10	2	27	18.78	3.35	0.163	9.45	1.99
11	2	30	20.56	3.32	0.148	9.46	2.17
12	2	33	22.33	3.28	0.136	9.46	2.36
13	2	36	24.11	3.26	0.126	9.46	2.55
14	2	39	25.89	3.24	0.117	9.47	2.74
15	2	42	27.67	3.22	0.110	9.47	2.92
16	2	45	29.44	3.20	0.103	9.47	3.11
17	2	48	31.22	3.19	0.097	9.47	3.30
18	2	51	33.00	3.18	0.092	9.47	3.48
19	2	54	34.78	3.17	0.087	9.47	3.67
20	2	57	36.56	3.16	0.083	9.47	3.86

The inequality is satisfied for $n \geq 17$; choose $n = 17$.

In the above example we found by using the (not very detailed) graphs in the textbook, that we would need about $n = 15$ which then was reasonably accurate.

The general fixed effect test sample size

The fixed effect test is generally carried out using

$$F = S_\tau^2 / S_2^2 \in F(\nu_\tau, \nu_2, \gamma^2(n_\tau))$$

where S_τ^2 is the mean square between treatments (τ_i, say) and S_2^2 is the proper test mean square.

The degrees of freedom are ν_τ and ν_2, respectively. In general τ_i may denote a fixed main effect or a fixed interaction effect.

In general the expected mean squares are of the form $E\{S_\tau^2\} = n_\tau \cdot \Sigma_i \tau_i^2/\nu_\tau + \omega^2$ and $E\{S_2^2\} = \omega^2$ where ω^2 is a linear combination of variances which depends on the design and the model chosen.

The constant n_τ is equal to the number of single measurements per level of the treatments or treatment combinations τ_i.

The non-centrality parameter is

$$\gamma^2(n_\tau) = n_\tau \cdot \sum_i \tau_i^2/\omega^2$$

Our requirement is, as above, met if

$$F(\nu_1, \nu_2, 0)_\alpha \leq F(\nu_1, \nu_2, \gamma^2(n_\tau))_{1-\beta}$$

By trying different n_τ and corresponding ν_1 and $\gamma^2(n_\tau)$ the lowest n_τ satisfying this inequality gives the necessary sample size.

In multi-factor and/or multilevel experiments the specification of a reasonable ω^2 may be difficult - not least because it depends on the design.

The general random effect test sample size

The random effect test is generally carried out using

$$F = S_B^2/S_2^2 \in \lambda^2(n_B)\cdot F(\nu_B,\nu_2)$$

where S_B^2 is the mean square between the levels of the random factor B and S_2^2 is the proper test mean square. The degrees of freedom are ν_B and ν_2, respectively.

In general the expected mean squares are of the form $E\{S_B^2\} = n_B\cdot\sigma_B^2 + \omega^2$ and $E\{S_2^2\} = \omega^2$ where ω^2 is a linear combination of variances which depends on the design chosen (may depend on n_B and the model, but does not include σ_B^2).

The constant n_B is equal to the number of single measurements per level of the random factor B.

The scale parameter

$$\lambda^2(n_B) = (n_B\cdot\sigma_B^2 + \omega^2)/\omega^2$$

Our requirement is met if

$$F(\nu_B,\nu_2)_\alpha \le \lambda^2(n_B)\cdot F(\nu_B,\nu_2)_{1-\beta} = \lambda^2(n_B)/F(\nu_2,\nu_B)_\beta$$

By trying different n_B values using $\nu_B = k-1$ and the corresponding ν_2 and $\lambda^2(n_B)$ the lowest n_B satisfying this inequality is the necessary sample size.

Again, in multi-factor and/or multilevel experiments the specification of a reasonable ω^2 may be difficult - not least because it depends on the design.

Repeated Latin squares and ANOVA

3 squares with identical operators (3) and batches (3)

$$\begin{array}{c|ccc|}
 & O_1 & O_2 & O_3 \\ \hline
B_1 & A & B & C \\
B_2 & B & C & A \\
B_3 & C & A & B \\ \hline
\multicolumn{4}{c}{R_1}
\end{array}
\qquad
\begin{array}{c|ccc|}
 & O_1 & O_2 & O_3 \\ \hline
B_1 & B & C & A \\
B_2 & A & B & C \\
B_3 & C & A & B \\ \hline
\multicolumn{4}{c}{R_2}
\end{array}
\qquad
\begin{array}{c|ccc|}
 & O_1 & O_2 & O_3 \\ \hline
B_1 & C & A & B \\
B_2 & B & C & A \\
B_3 & A & B & C \\ \hline
\multicolumn{4}{c}{R_3}
\end{array}$$

$$Y_{\nu ijk} = \mu + R_\nu + \tau_i + B_j + O_k + E_{\nu ijk}$$

Latin square ANOVA

Source of var.	SSQ	d.f.	s^2	EMS	F-test
Treatments	SSQ_τ	$3-1$	s_τ^2	$\sigma^2 + 9\phi_\tau$	F_τ
Replicates	SSQ_R	$3-1$	s_R^2	$(\sigma^2 + 9\sigma_R^2)$	(F_R)
Batches	SSQ_B	$3-1$	s_B^2	$(\sigma^2 + 9\sigma_B^2)$	(F_B)
Operators	SSQ_O	$3-1$	s_O^2	$(\sigma^2 + 9\sigma_O^2)$	(F_O)
Uncertainty	SSQ_E	18	s_E^2	σ^2	
Total	SSQ_{tot}	$27-1$			

Sums of squares are computed as usual - using sums:

$$SSQ_\tau = \sum_{i=1}^{3}\frac{T_{.i..}^2}{9} - \frac{T_{....}^2}{27}\ ,\ SSQ_R = \sum_{\nu=1}^{3}\frac{T_{\nu...}^2}{9} - \frac{T_{....}^2}{27}$$

$$SSQ_B = \sum_{j=1}^{3}\frac{T_{..j.}^2}{9} - \frac{T_{....}^2}{27}\ ,\ SSQ_O = \sum_{k=1}^{3}\frac{T_{...k}^2}{9} - \frac{T_{....}^2}{27}$$

$$SSQ_E = SSQ_{tot} - SSQ_R - SSQ_\tau - SSQ_B - SSQ_O$$

3 squares with 9 operators and 3 batches

	O_1	O_2	O_3		O_4	O_5	O_6		O_7	O_8	O_9
B_1	A	B	C	B_1	B	C	A	B_1	C	A	B
B_2	B	C	A	B_2	A	B	C	B_2	B	C	A
B_3	C	A	B	B_3	C	A	B	B_3	A	B	C
		R_1				R_2				R_3	

$$Y_{\nu ijk} = \mu + R_\nu + \tau_i + B_j + O(R)_{k(\nu)} + E_{\nu ijk}$$

Latin square ANOVA					
Source of var.	SSQ	d.f.	s^2	EMS	F-test
Treatments	SSQ_τ	$3-1$	s^2_τ	$\sigma^2 + 9\phi_\tau$	F_τ
Replicates	SSQ_R	$3-1$	s^2_R	$(\sigma^2 + 9\sigma^2_R)$	(F_R)
Batches	SSQ_B	$3-1$	s^2_B	$(\sigma^2 + 9\sigma^2_B)$	(F_B)
Operators	$SSQ_{O(R)}$	$3(3-1)$	$s^2_{O(R)}$	$(\sigma^2 + 3\sigma^2_{O(R)})$	$(F_{O(R)})$
Uncertainty	SSQ_E	14	s^2_E	σ^2	
Total	SSQ_{tot}	$27-1$			

$\boxed{\text{Sums of squares are computed as usual - using sums - again:}}$

$$SSQ_O = \sum_{\nu=1}^{3}\left[\sum_{k(\nu)} \frac{T^2_{\nu..k}}{3} - \frac{T^2_{\nu...}}{9}\right]$$

Note that now $SSQ_{O(R)}$ is computed <u>within</u> replicates and added up over the three replicates giving 2 degrees of freedom for each replicate. The summation over k is thus over the three values within the replicate ν.

SSQ_R , SSQ_τ , SSQ_B and SSQ_E as above.

3 squares with 9 operators and 9 batches

	O_1	O_2	O_3		O_4	O_5	O_6		O_7	O_8	O_9
B_1	A	B	C	B_4	B	C	A	B_7	C	A	B
B_2	B	C	A	B_5	A	B	C	B_8	B	C	A
B_3	C	A	B	B_6	C	A	B	B_9	A	B	C
		R_1				R_2				R_3	

$$Y_{\nu ijk} = \mu + R_\nu + \tau_i + B(R)_{j(\nu)} + O(R)_{k(\nu)} + E_{\nu ijk}$$

Latin square ANOVA					
Source of var.	SSQ	d.f.	s^2	EMS	F-test
Treatments	SSQ_τ	$3-1$	s^2_τ	$\sigma^2 + 9\phi_\tau$	F_τ
Replicates	SSQ_R	$3-1$	s^2_R	$(\sigma^2 + 9\sigma^2_R)$	(F_R)
Batches	$SSQ_{B(R)}$	$3(3-1)$	$s^2_{B(R)}$	$(\sigma^2 + 3\sigma^2_{B(R)})$	$(F_{B(R)})$
Operators	$SSQ_{O(R)}$	$3(3-1)$	$s^2_{O(R)}$	$(\sigma^2 + 3\sigma^2_{O(R)})$	(F_O)
Uncertainty	SSQ_E	10	s^2_E	σ^2	
Total	SSQ_{tot}	$27-1$			

$\boxed{\text{Sums of squares are computed as usual - using sums - again - again :}}$

$$SSQ_{B(R)} = \sum_{\nu=1}^{3}\left[\sum_{j(\nu)} \frac{T^2_{\nu.j.}}{3} - \frac{T^2_{\nu...}}{9}\right]$$

$$SSQ_{O(R)} = \sum_{\nu=1}^{3}\left[\sum_{k(\nu)} \frac{T^2_{\nu..k}}{3} - \frac{T^2_{\nu...}}{9}\right]$$

Now both $SSQ_{O(R)}$ and $SSQ_{B(R)}$ are computed within replicates.

SSQ_R , SSQ_τ , and SSQ_E as above.