Chapter 0. Introduction to Statistical Experimental Design

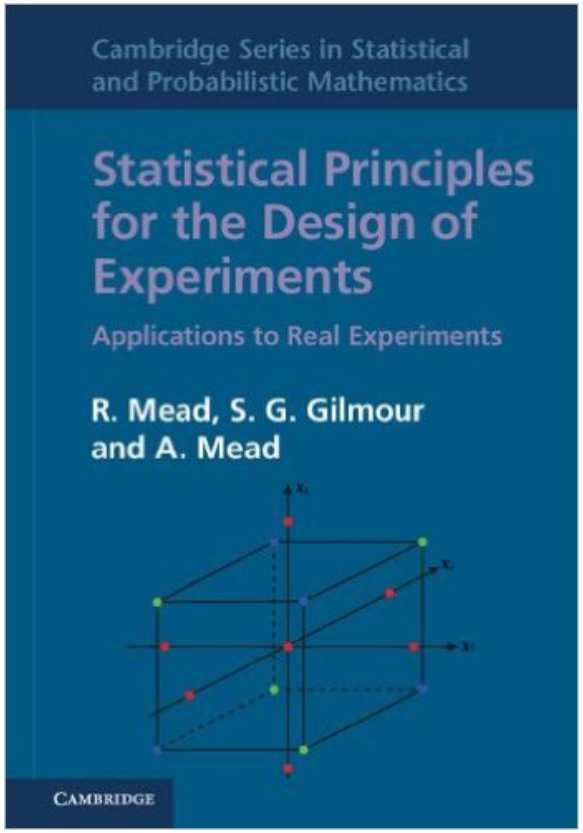
C.O.S. Sorzano
coss@cnb.csic.es

National Center of Biotechnology (CSIC)

October 14, 2016

Outline

References



R. Mead, S.G. Gilmour, A. Mead. Statistical Principles for the Design of Experiments: Applications to Real Experiments. Cambridge University Press (2012)

Outline

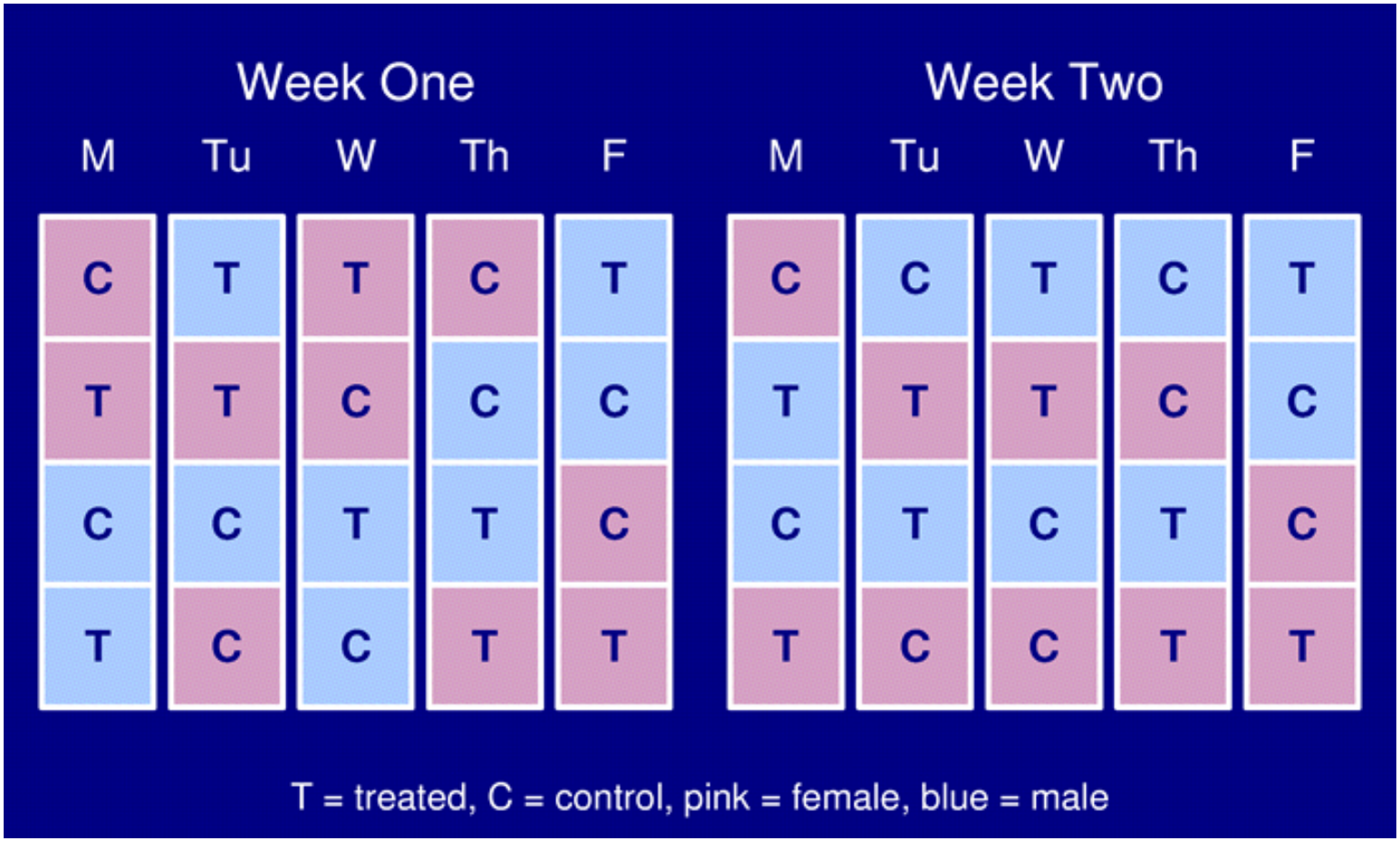
Why this course?

"In God we trust. All others must bring data".

W. Edwards Deming

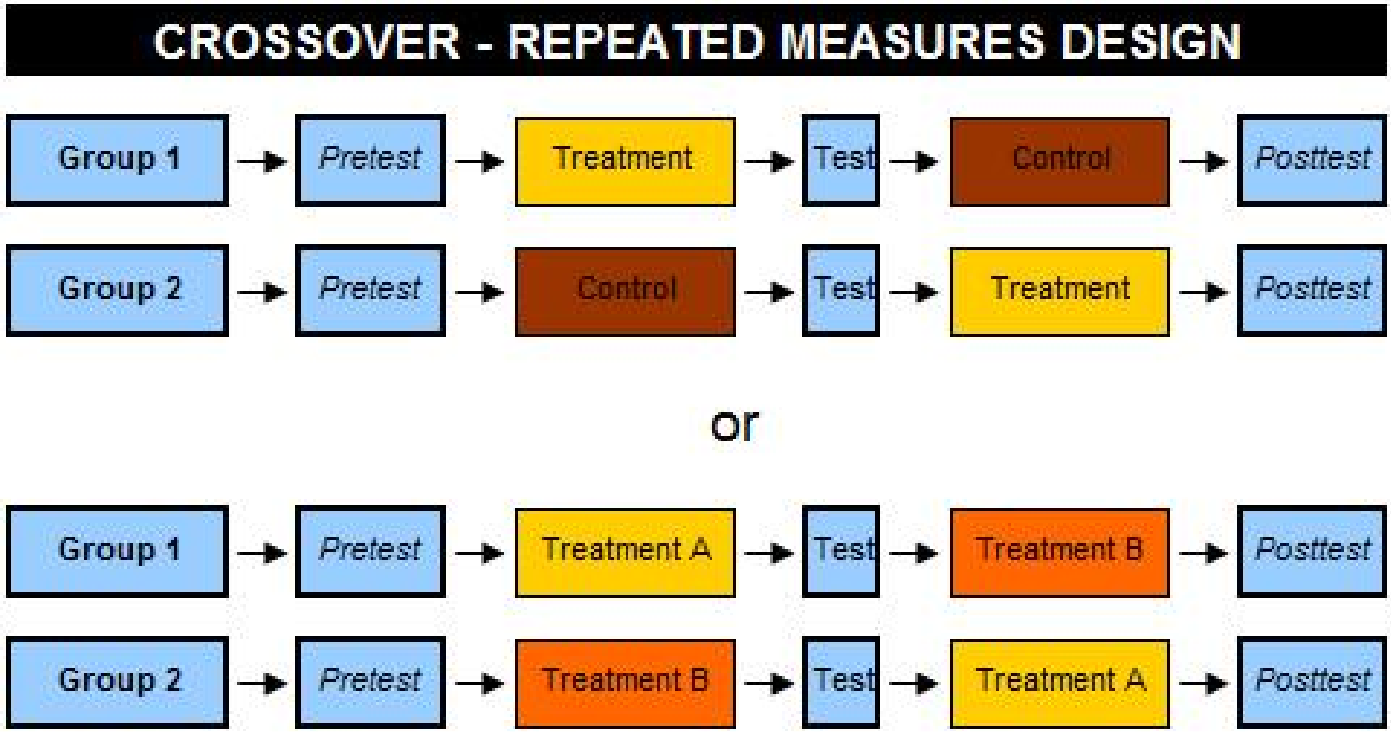
Why this course?

You need to study 40 animals (20 female, 20 male; 20 treated, 20 untreated). Only 4 animals/day can be processed, so you need 10 days to perform the experiment. What is the optimal way of distributing the animals?



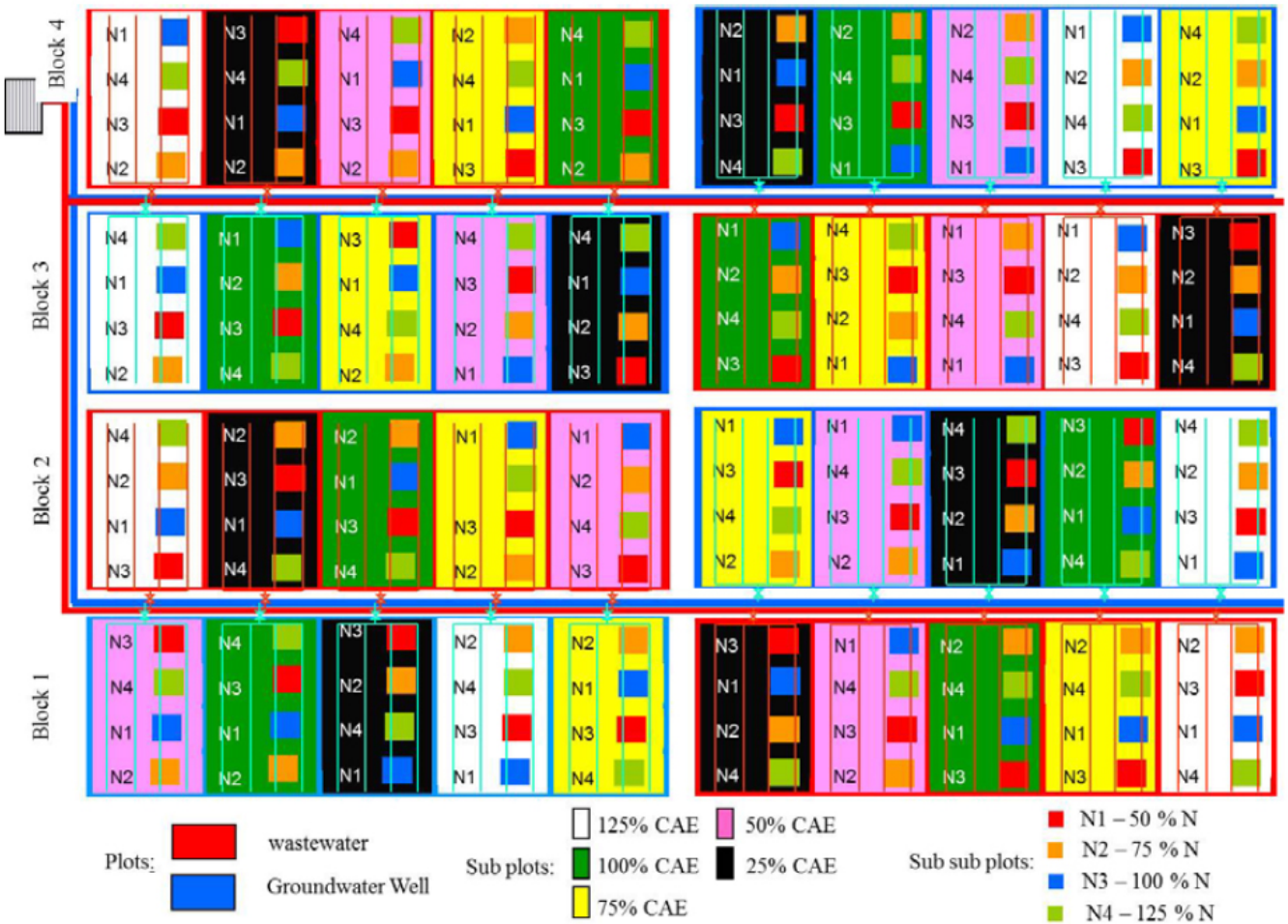
Why this course?

You are studying the effect of a new drug. How will you perform the experiment and how will you analyze the data?



Why this course?

You are looking for the optimal way of irrigating and fertilizing land to grow sunflowers. You have 5 different ways of irrigation (CAE) and 5 different levels of fertilizer.



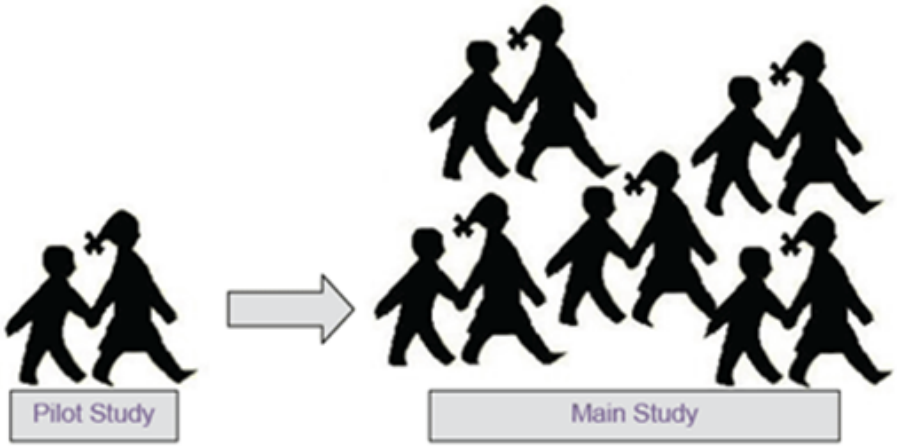
Why this course?

You cannot fix by analysis, what you have bungled by design.

Outline

2. Types of experiments

Types of experiments



Pilot experiments are small studies (1-20 experimental subjects) used to:

1. Test the logistics of a proposed larger study
2. Gain familiarity with the experimental material,
3. Ensure that treatments are not obviously excessively mild or severe
4. Check that staff are sufficiently well trained in the necessary procedures
5. Ensure that all steps in a proposed future experiment are feasible.
6. Gain some information on variability, although this will not usually be sufficiently reliable to form the basis of power analysis calculations of sample size.

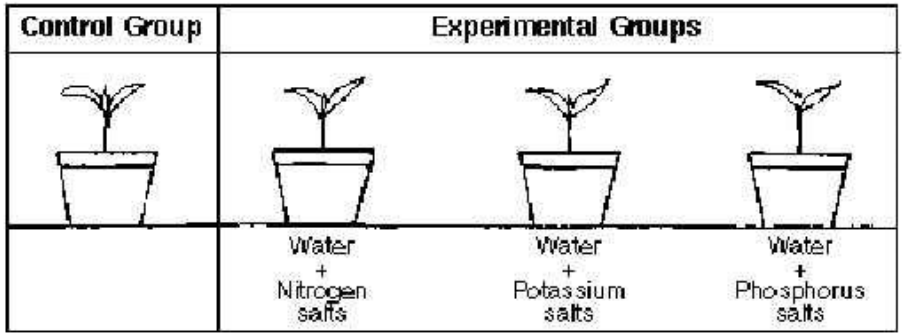
Types of experiments

Exploratory experiments can be used to generate data with which to develop hypotheses for future testing. They may "work" or "not work". They may have no clearly stated hypothesis ("let's see what happens if ..." is not a valid hypothesis on which to base an experiment).

Often they will measure many outcomes (characters). Picking out "interesting looking differences" (known as data snooping) and then doing a hypothesis test to see if the differences are statistically significant will lead to serious overestimation of the magnitude of a response and excessive numbers of false positive results. Such differences should always be tested in a controlled experiment where the hypothesis is stated *a priori* before the results are published.

Depending on the nature of the data, statistical analysis will often be done using an analysis of variance (ANOVA)

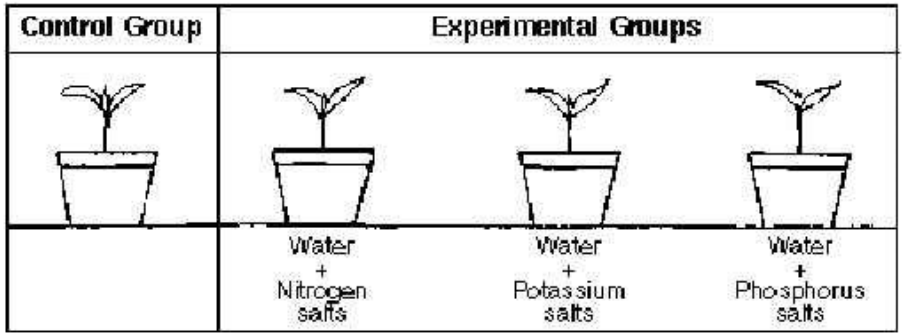
Types of experiments



Confirmatory experiments are used to test some relatively simple hypothesis stated *a priori*. This is the type of experiment mainly considered in this course. The basic principles are:

1. Experiments involve *comparisons* between two or more groups
2. Their aim is to *test a "null hypothesis"* that there is no difference among the groups for the specified outcome.
3. If the null hypothesis is rejected at a certain level of probability (often 5%) this means that the probability of getting a result as extreme as this or more extreme in the absence of a true effect is 5% (assuming also that the experiment has been properly conducted). So it is assumed that such a difference is likely to be the result of the treatment. But, it could be a false positive resulting from sampling variation.
4. Failure to reject the null hypothesis does not mean that the treatment has no effect, only that if there is a real effect this experiment failed to detect it. "Absence of evidence is not evidence of absence".
5. Experimental subjects need to be *independently replicated* because individuals (of whatever type) vary. Two subjects can normally be regarded as being independent if they can theoretically receive different treatments.

Types of experiments



6. Subjects need to be assigned to groups, held in the animal house and measured *at random* in order to minimise the chance of bias (a systematic difference between groups)
7. As far as possible the experimenter should be *"blind"* with respect to the treatment group in order to minimise bias.
8. The experiments need to be *powerful*, i.e. they should have a high probability of detecting an effect of clinical or scientific importance if it is present.
9. In many cases a *formal experimental design* such as a "completely randomised", "randomised block", "Latin square" etc. design will be used.
10. In most cases it is useful if the experiment has a *wide range of applicability*. In other words the results should hold true under a range of different conditions (different strains, both sexes, different diets, different environments etc.). At least some of these factors should be explored using factorial and randomised block designs.

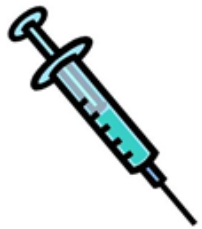
Outline

3. Experimental units

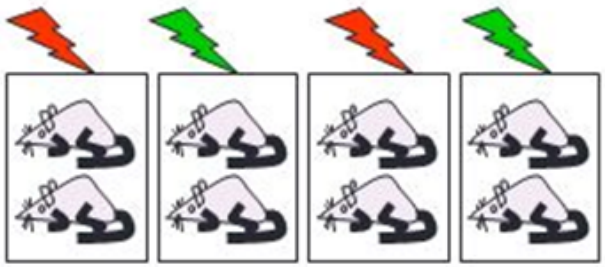
Experimental units

"The smallest division of the experimental material such that any two experimental units can receive different treatments"

In this study the animals are all housed in one cage and the treatment is given by injection.

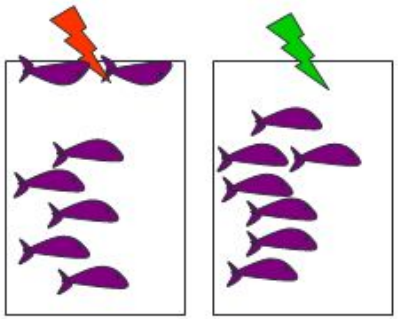


Any two animals can receive different treatments, so the animal is the experimental unit and the total number of subjects is N=8

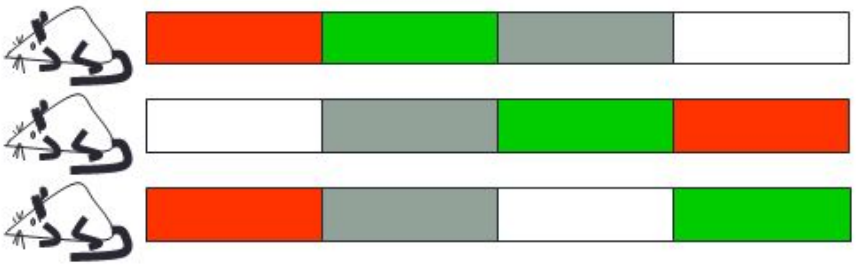


In this study the animals are housed two per cage and the treatment is given in the food or water. N=4

Experimental units



The experiment on the left has seven fish in each of two tanks. The left-hand tank has been treated with a test substance poured into the water and the right-hand has only the vehicle as a control. The aim is to measure the level of an enzyme in the fish. N=2.



In a crossover experiment an animal could be given a treatment for a period, then rested and given a different treatment for a period. It is assumed that the treatment doesn't alter the animal, so it has to be very mild. N=12.

Experimental units

In a teratology experiment the pregnant female is treated with the test compound or a placebo. The pregnant females are killed at about mid-gestation and the pups are weighed, measured and studied for abnormalities. N=2.

Strain WKY rats are sometimes used as a model of depression, whereas Wistar rats are not depressive. The goal is to see if there is a relationship between depression and pain sensitivity. So he obtains 10 rats of each strain, houses them two per cage for three weeks and tests them in random order using a standard test of pain threshold. N=2.

Outline

4. Experiment design

"There's a flaw in your experimental design.
All the mice are scorpios."

Experiment design

A good experiment must:

1. Have a clear specification of the aims of the experiment. The hypothesis to be tested needs to be clearly formulated *before* starting any detailed planning. It should be one which the experiment is capable of answering. It would be a serious error to look at the results of the experiment and then adjust the hypothesis to fit them!

Experiment design

2. Be unbiased. There should be no systematic differences between the treated and control groups apart from the effects of the treatment.

Bias may result in false positive results when the effects of some other factor are confounded (mixed with) the treatment effect.

Bias is minimised by

1. correct choice of the experimental unit
2. randomisation of the units to treatments and in the order in which subjects are housed and outcomes are measured
3. blinding where possible, using coded samples.

Experiment design

3. Be powerful. If the treatment really has an effect, there should be a high chance that it can be detected. Experiments which lack power will have too many false negative results. Power is increased by

1. Larger sample sizes
2. Good control of variability
3. Use of sensitive subjects.

However, large sample sizes cost animals and money so emphasis should be placed on the last two of these.

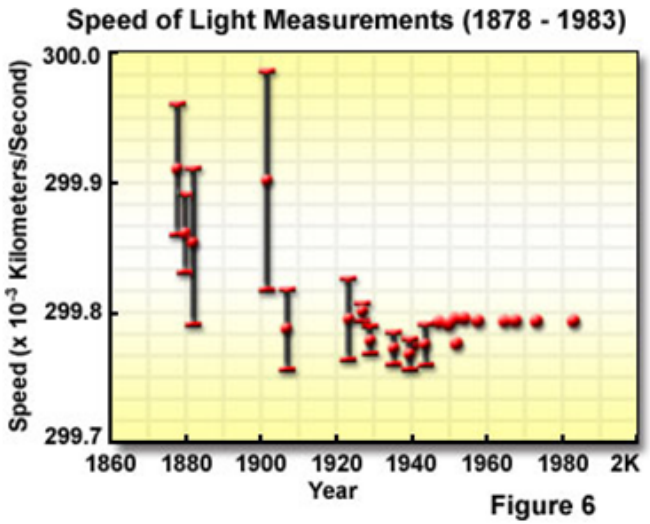


Figure 6

Experiment design

4. Have a wide range of applicability. An experiment where the results can only be replicated in some animal houses but not in others lacks generality. The range of applicability is explored using factorial and randomised block designs which can sample different situations.

Internal validity
It has a high probability of getting the correct answer. Basically, this means that it should be unbiased and powerful so that it is unlikely to produce either a false positive or a false negative result.

External validity
The results can be generalised to other conditions or situations. Note that it can not have high external validity unless it first has high internal validity.

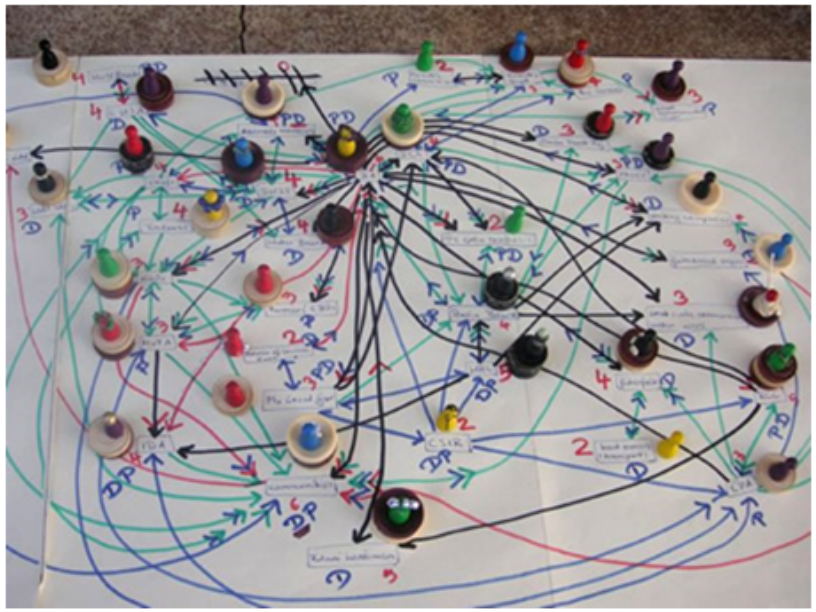
As an example, an experiment which uses only a single strain of mice may have high internal validity, but if the same results are not seen with other strains of mice, then it will have low external validity.

It is acceptable to do an experiment with high internal validity but no exploration of its external validity, provided it is made clear that the external validity is unknown. But note that in many cases randomised block and factorial designs can be done at little or no extra cost

Experiment design

5. Experiments should be simple. They should not be so complex that mistakes are made, the statistical analysis is excessively complex or they are impossible to interpret.

Clearly written protocols and stand operating procedures should be used. In some cases it may be necessary to work to "Good Laboratory Practice" standards



Experiment design

6. It should be possible to statistically analyse the result of an experiment. The statistical analysis and the experiment should be planned at the same time.

1. An investigator should never start an experiment without knowing how it is going to be analysed.
2. The results of each experiment should be analysed before starting the next one so that the findings from the first experiment can be taken into account.
3. The most powerful available statistical methods should be used, such as parametric rather than non parametric tests, where applicable.

Outline

5. Avoiding bias

J Chron Dis Vol. 32, pp. 51 to 63
Pergamon Press Ltd 1979. Printed in Great Britain

57 types of bias

BIAS IN ANALYTIC RESEARCH

DAVID L. SACKETT

INTRODUCTION

CASE-CONTROL studies are highly attractive. They can be executed c
cost, even when the disorders of interest are rare. Furthermore, the
case-control studies is becoming automated; strategies have been de
puter scanning' of large files of hospital admission diagnoses and pri
with more detailed analyses carried out in the same data set on an
As evidence of their growing popularity, when one original article was
from each issue of **The New England Journal of Medicine, The Lance**
of the American Medical Association for the years, 1956, 1966 and 19
reporting case-control analytic studies increased fourfold over those t
whereas the proportion reporting cohort analytic studies fell by half
tally, a general trend toward fewer study subjects but more study
noted [2].

Avoiding bias

Randomisation ensures that each experimental unit has an equal probability of receiving a particular treatment.

A successful randomisation does not allow to predict the group in advance.

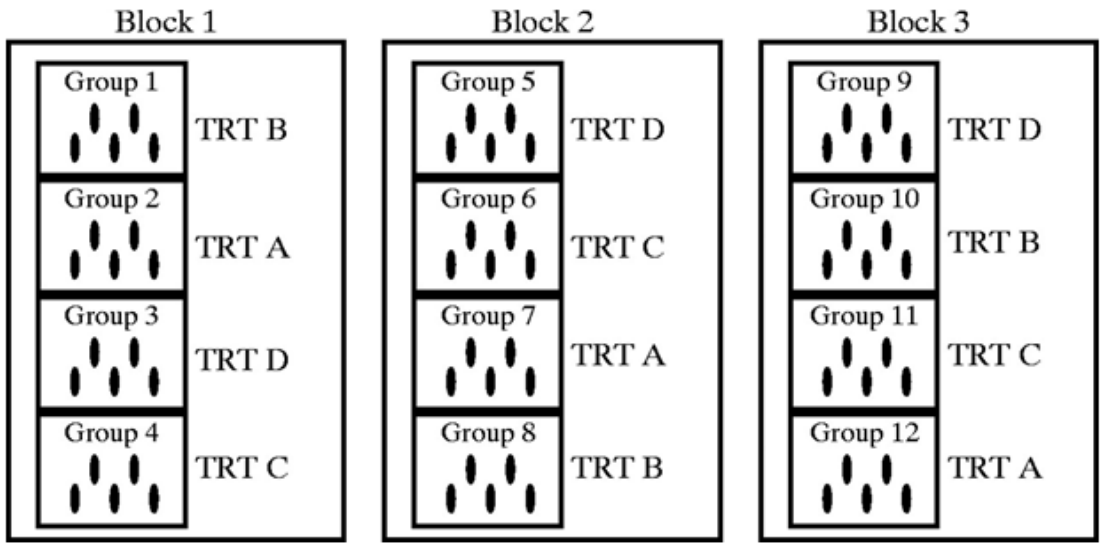
Original	=rand()	Sorted on =rand()		Animal number
A	0.527	A	0.067	1
A	0.100	A	0.100	2
A	0.067	A	0.122	3
A	0.122	C	0.210	4
B	0.665	B	0.248	5
B	0.875	C	0.265	6
B	0.478	B	0.478	7
B	0.248	A	0.527	8
C	0.210	C	0.628	9
C	0.628	B	0.665	10
C	0.265	B	0.875	11
C	0.895	C	0.895	12

Avoiding bias

Randomising a randomised block design

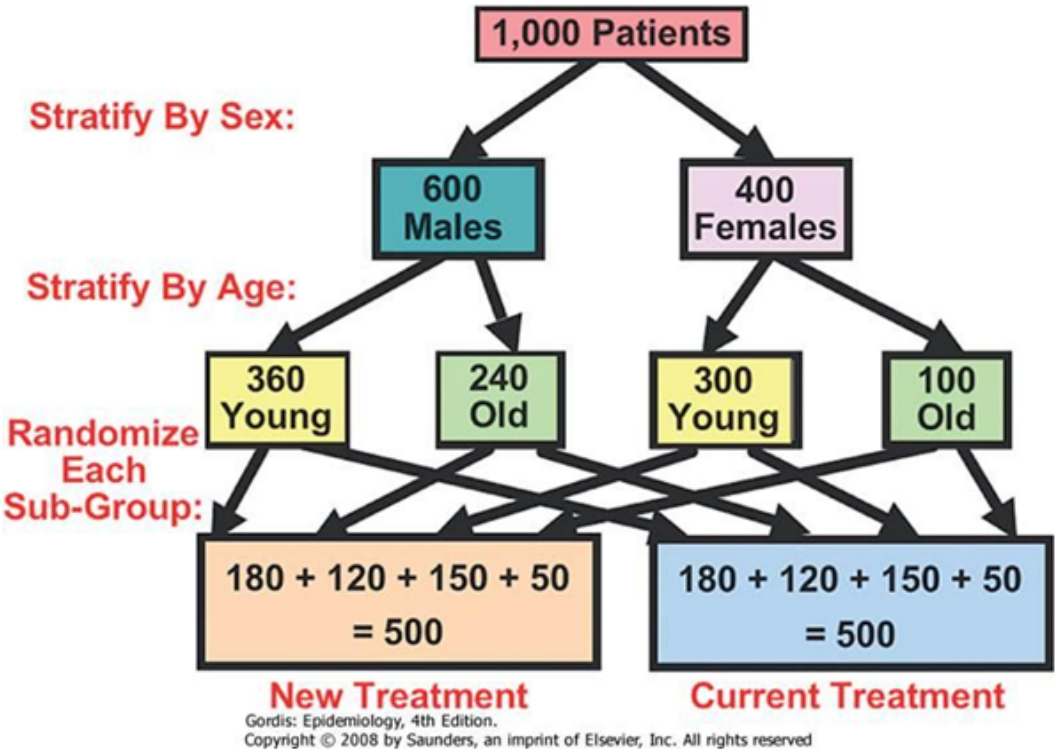
In a randomised block design the experiment is split up into a number of small parts or "blocks". Typically each block has one experimental unit of each treatment (but it may have more). So if there are four treatments, block size is four experimental units.

Animal	"=rand()"	Treatmen	Block
1	0.75	D	1
2	0.40	A	1
3	0.73	C	1
4	0.70	B	1
5	0.02	A	2
6	0.60	D	2
7	0.08	B	2
8	0.12	C	2
9	0.07	B	3
10	0.04	A	3
11	0.54	C	3
12	0.84	D	3
13	0.94	D	4
14	0.39	A	4
15	0.80	C	4
16	0.70	B	4



Avoiding bias

Randomising and blocking two variables

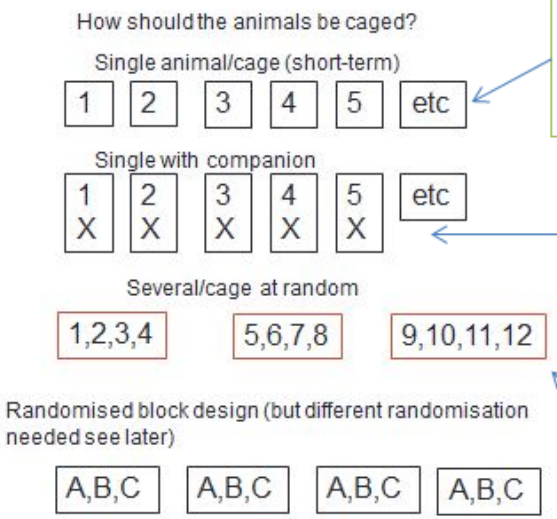


Avoiding bias

There is no one answer to the numbers of animals housed per cage.
It depends on species and the nature of the experiment.

Treatment (randomised)	Random number now sorted	Animal number
A	0.067	1
A	0.100	2
A	0.122	3
C	0.210	4
B	0.248	5
C	0.265	6
B	0.478	7
A	0.527	8
C	0.628	9
B	0.665	10
B	0.875	11
C	0.895	12

It is not a good idea to house **all the controls in one cage**, all of treatment A in a second cage etc. as then the cage becomes the experimental unit. There can be "**cage effects**" due to social interactions which could seriously bias the results (e.g. if all the controls are fighting, but the treated animals are not).



Single housing of mice and rats may be stressful and is strongly discouraged for welfare reasons. But male mice may fight, depending on the strain and husbandry conditions.

Very valuable animals such as those fitted with telemetry apparatus, or ones with a genetic modification are sometimes **housed with a companion** which is not part of the experiment.

Group housing poses problems if treatment is given in the food or water as the cage is then the experimental unit unless sophisticated apparatus is used so that each animal can have a different diet. This is sometimes done with farm animals.

Group housing may also be a problem if drug treatments are involved as rats and mice are coprophageous so control animals may consume metabolites of the test compound if animals of different treatment groups are housed together.

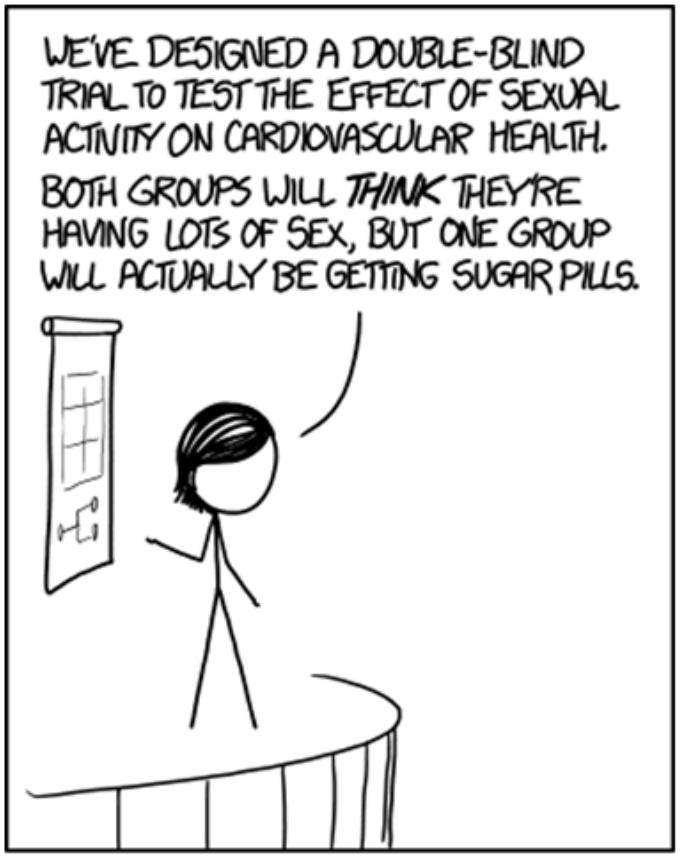
If animals receiving different treatments (or genetically modified and wild type animals) can be housed together, then a **randomised block design** might be used as shown at the bottom of the figure (above).

Avoiding bias

Blinding

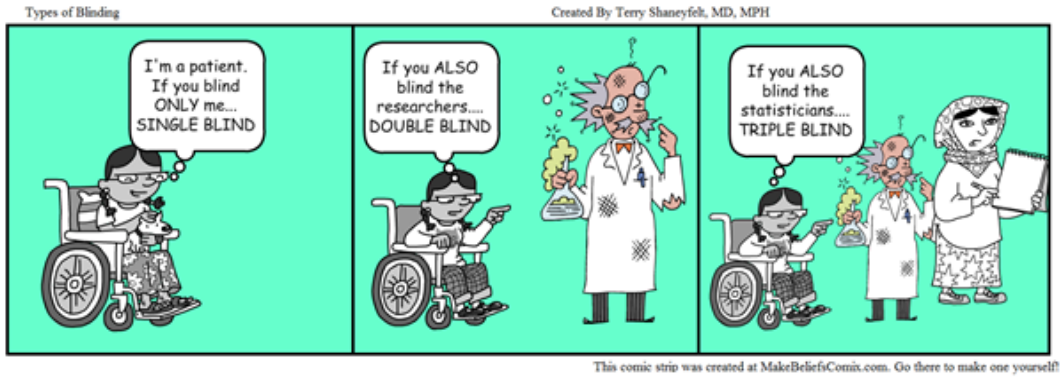
We usually have a vested interest in the outcome of our experiments. We might want to find "significant" differences between groups, or in some cases no significant differences (particularly if we are toxicologists). So, having done the randomisation, wherever possible use the animal numbers as codes to "blind" everyone to the treatment.

This is particularly important when making measurements, scoring histological sections or measuring behaviour. Blinding may be difficult in some cases such as when comparing two mouse strains which differ in coat colour.



Avoiding bias

Avoiding bias



BRIEF REPORTS

Emergency Medicine Animal Research: Does Use of Randomization and Blinding Affect the Results?

Vik Bebarta, MD, Dylan Luyten, MD, Kennon Heard, MD

290 animal studies were scored for blinding, randomisation and whether the outcome was positive or negative, as defined by authors. The results are shown below:

	Odds ratio
Blind/not blind	3.4 (95% CI 1.7-6.9)
Random/not random	3.2 (95% CI 1.3-7.7)
Both/neither	5.2 (95% CI 2.0-13.5)

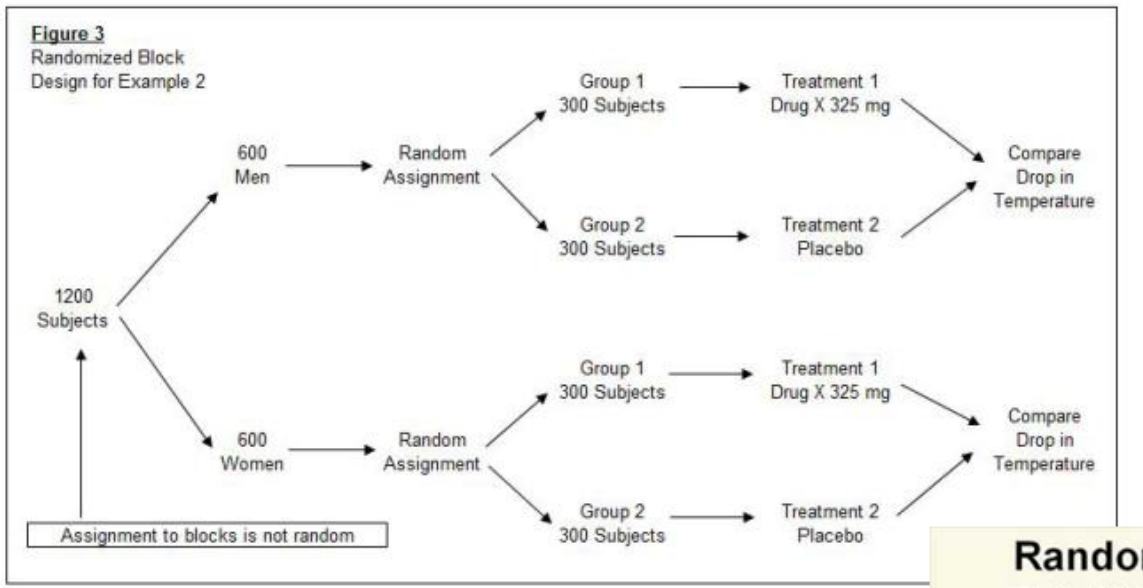
An odds ratio of one implies that blinding or randomisation was not associated with the outcome of an experiment. **These positive odds ratios show that on average studies which were not blinded and/or randomised produced excessive numbers of (presumably false) positive results**.

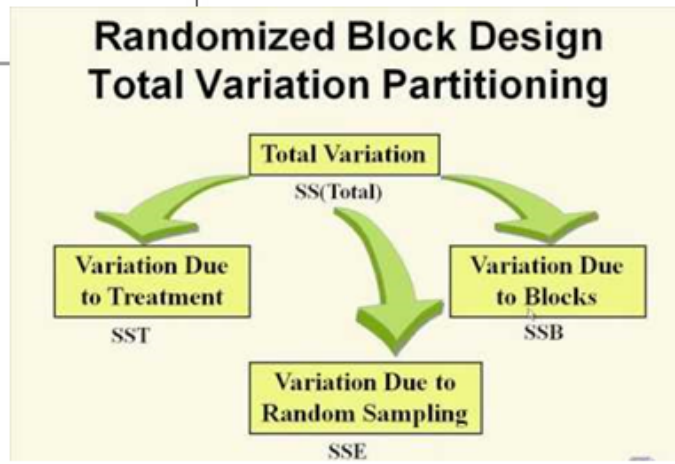
Outline

6. Reducing variance

Reducing variance

Randomized block designs

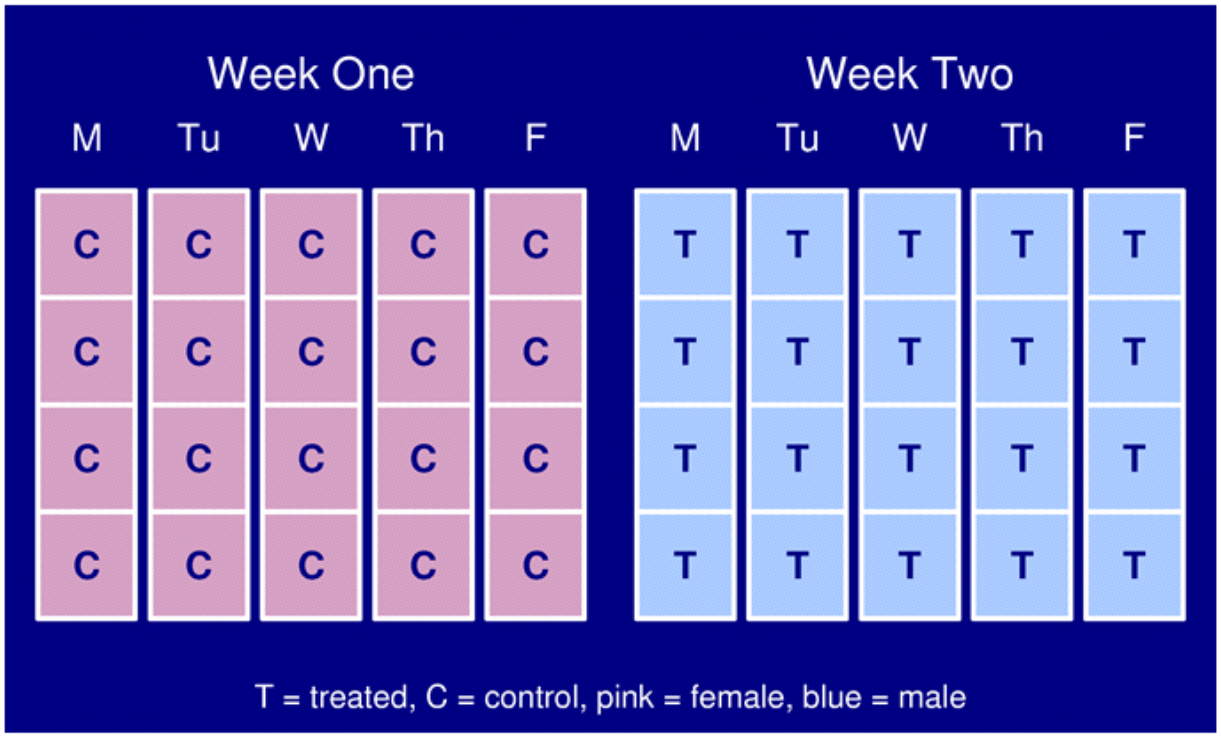




Reducing variance

Randomized block designs (blocking time)

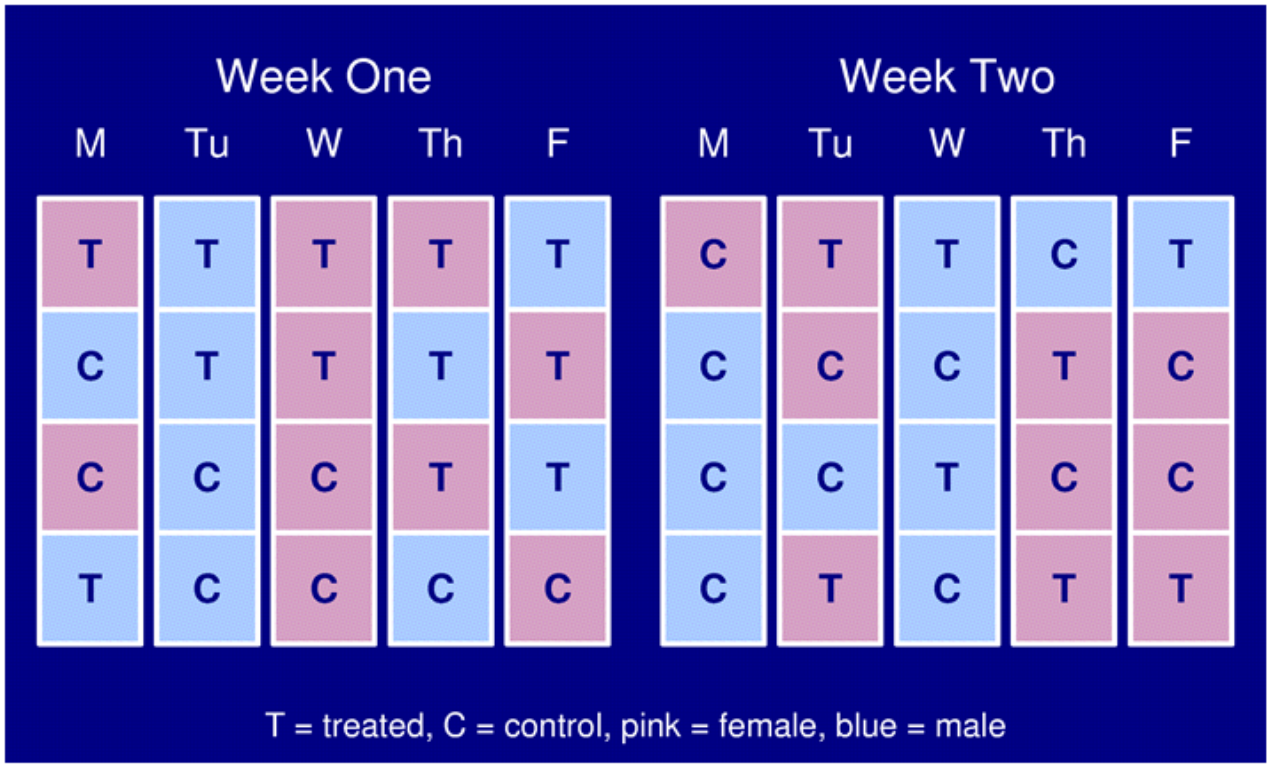
20 female, 20 male. 20 treated, 20 untreated.
Only 4 animals/day can be processed → 10 days



Reducing variance

Randomized block designs (blocking time)

20 female, 20 male. 20 treated, 20 untreated.
Only 4 animals/day can be processed → 10 days



Reducing variance

Randomized block designs (blocking time)

20 female, 20 male. 20 treated, 20 untreated.
Only 4 animals/day can be processed → 10 days

Week One					Week Two				
M	Tu	W	Th	F	M	Tu	W	Th	F
C	T	T	C	T	C	C	T	C	T
T	T	C	C	C	T	T	T	C	C
C	C	T	T	C	C	T	C	T	C
T	C	C	T	T	T	C	C	T	T

T = treated, C = control, pink = female, blue = male

Reducing variance

Randomized block designs

3) and Randomize the rest

2) Block what you cannot,

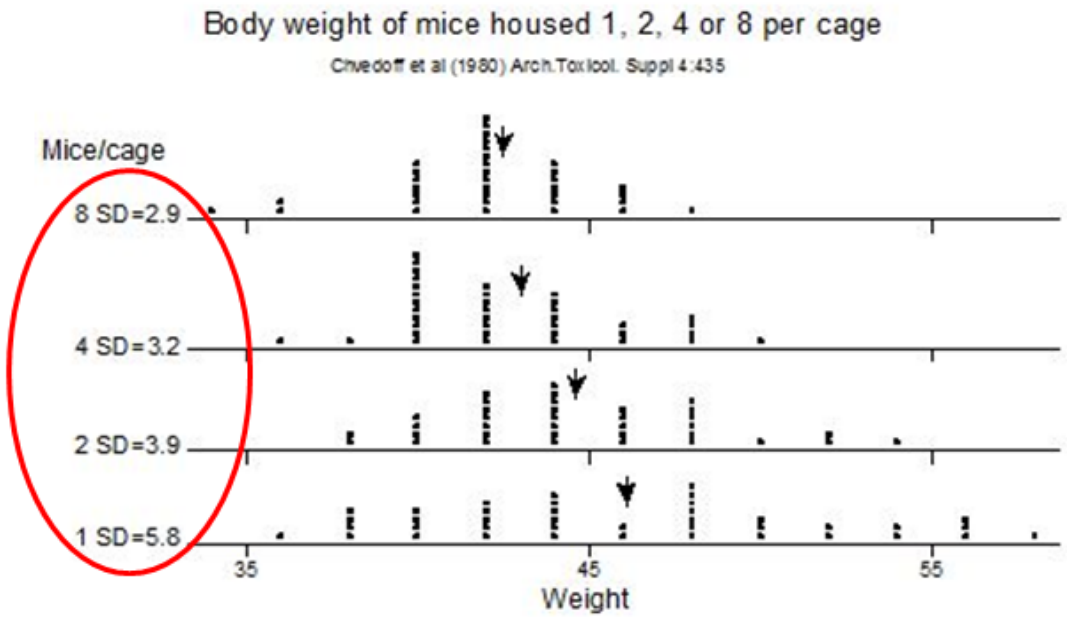
1) Control what you can,

- Randomize position in the shelf
- Randomize order of feeding
- Randomize time of treatment
- Randomize order of treatment
- Randomize ...

Reducing variance

Change experimental conditions

Chvedoff M et al (1980). Effects on mice of numbers of animal per cage: an 18-month study. (preliminary results). Archives of Toxicology, Supplement 4:435-438



Reducing variance

Choose subjects with less variability

Sleeping time under barbiturate anesthetic is sometimes used to indicate whether a test drug alters drug metabolising enzymes. All mice receive the barbiturate and half of them receive the test compound while the other are used as controls. A difference in sleeping time would indicate that the test substances alters drug metabolism.

The table below shows the number, mean and standard deviation of sleeping time in five inbred strains (A/N to SWR/HeN) and two outbred stocks (CFW and Swiss) of mice under hexobarbital anesthetic.

Strain	N	Mean	SD	No needed*	Power**
A/N	25	48	4	23	86
BALB/c	63	41	2	7	>99
C57BL/HeN	29	33	3	13	98
C3H/He	30	22	3	13	98
SWR/HeN	38	18	4	23	86
CFW	47	48	12	191	17
Swiss	47	43	15	297	13

79

* Power analysis: number needed in a two-sample t-test to detect a 4 min. change in the mean (2-sided) with a=0.05 and a power of 90%

** power of an experiment to detect a 4 min. change in the mean if the sample size is fixed at 20 mice/group

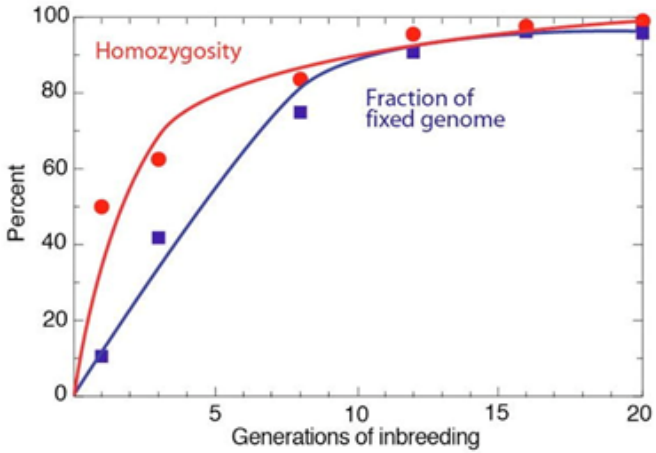
Data from Jay 1955 Proc Soc. Exp Biol Med 90:378

Reducing variance

Choose subjects with less variability

Inbred strain: homozygosity

Effect of inbreeding on homozygosity and allele fixation



Alleles
Strain1: a/a, b/b, c/c, ...
Strain2: A/A, B/B, C/C, ...

There is no genetic variability, all differences must be due to environment or treatment.

But, can it be extrapolated to the whole population? They reproduce poorly, they are not a model for all genetic diseases.

Reducing variance

Choose subjects with less variability

Hybrid F1: homozygosity

Alleles
Strain1: a/a, b/b, c/c, ...
+
Strain2: A/A, B/B, C/C, ...
=
Hybrid F1: a/A, b/B, c/C, ...

There is no genetic variability, all differences must be due to environment or treatment. Gain in hybrid vigor.

But, can it be extrapolated to the whole population?

Reducing variance

Choose subjects with less variability

Outbred stock: heterozygosity, but reduced genetic variability

There is limited genetic variability. More viable animals. Special care is taken to keep the genetic variability at a maximum within the colony.

But, can it be extrapolated to the whole population?

Reducing variance

Choose subjects with less variability

Mixed stock: heterozygosity, wide genetic variability

This is more similar to the "mouse species"

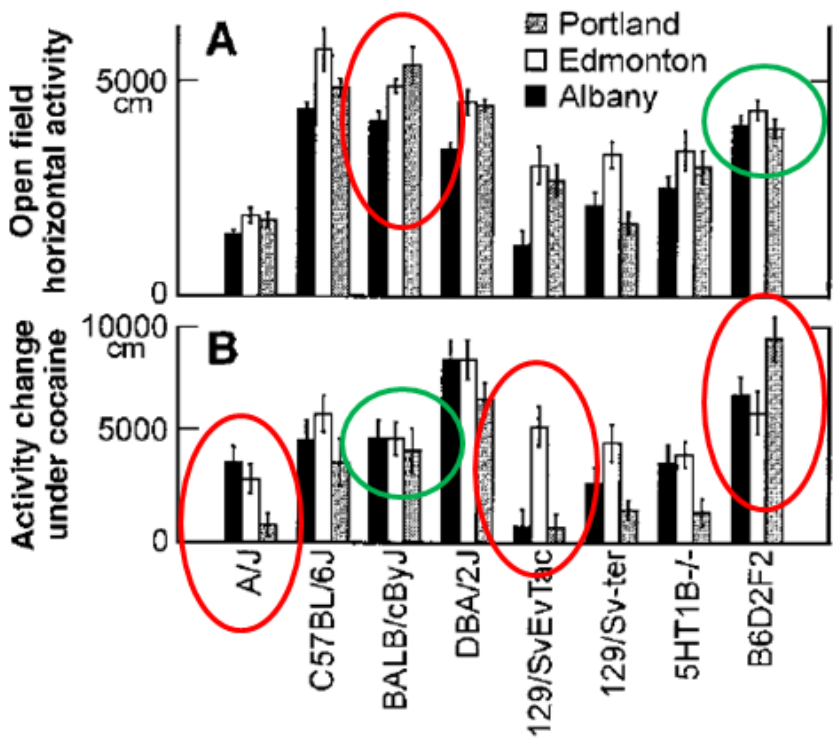
Reducing variance

Genetics of Mouse Behavior: Interactions with Laboratory Environment

John C. Crabbe,[1]* Douglas Wahlsten,[2] Bruce C. Dudek[3]

Strains of mice that show characteristic patterns of behavior are critical for research in neurobehavioral genetics. Possible confounding influences of the laboratory environment were studied in several inbred strains and one null mutant by simultaneous testing in three laboratories on a battery of six behaviors. Apparatus, test protocols, and many environmental variables were rigorously equated. Strains differed markedly in all behaviors, and despite standardization, there were systematic differences in behavior across labs. For some tests, the magnitude of genetic differences depended upon the specific testing lab. Thus, experiments characterizing mutants may yield results that are idiosyncratic to a particular laboratory.

Crabbe, J. C.; Wahlsten, D. & Dudek, B. C. Genetics of mouse behavior: interactions with laboratory environment. *Science*, **1999**, *284*, 1670-1672

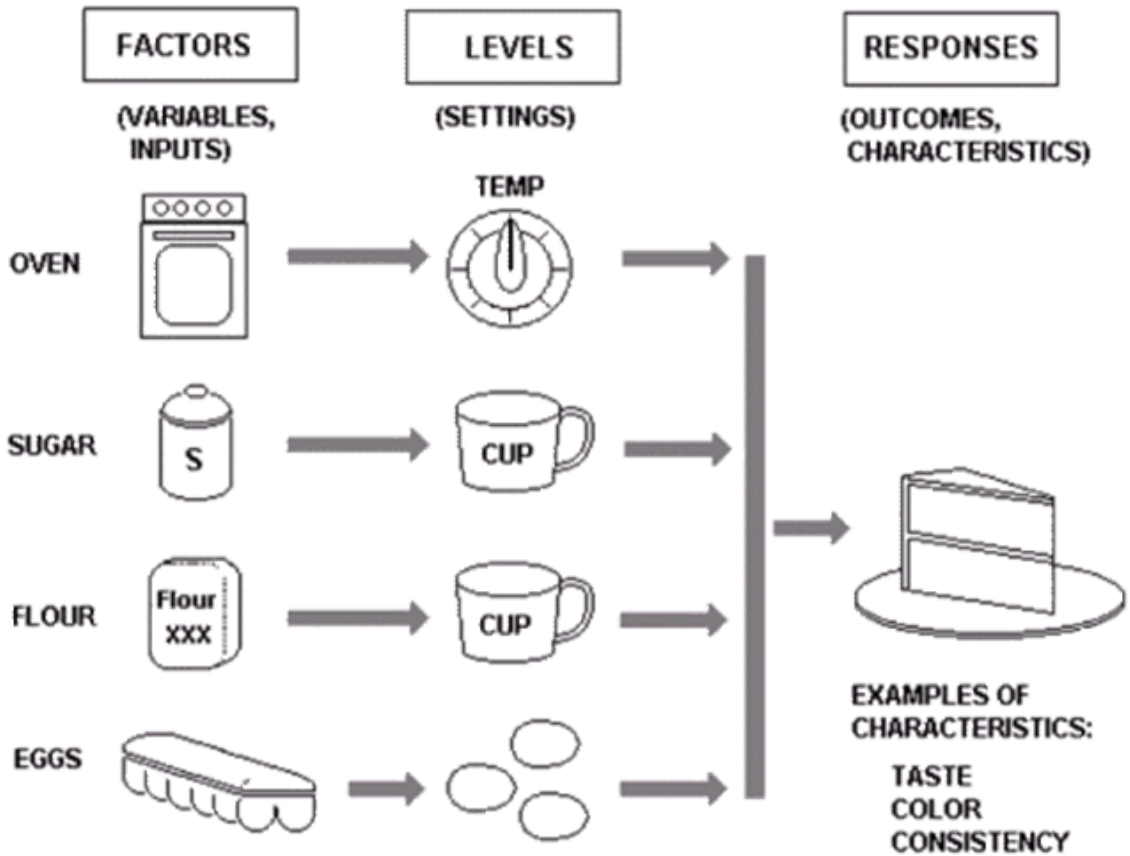


1. Same research team
2. Same inbred strains
3. Equally calibrated apparatus
4. Equated husbandry
5. Same testing protocols
6. Same age
7. Same starting time
8. Same protocol order

But significantly different results

Outline

7. Statistical Experimental Design



Statistical Experimental Design

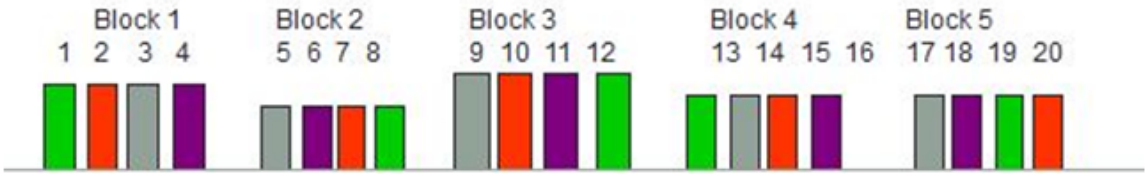
The completely randomised design

This is the simplest design. Each experimental unit is assigned to a treatment strictly at random without taking account of any individual characteristics. It is best used when relatively homogeneous experimental units are available. It can tolerate unequal numbers in each group and is perfectly adequate in many experimental situations. Following treatment investigators should (where possible) be blinded by using only the animal numbers when making measurements

If, for example, surgery is involved skill may increase, leading to a bias against gray. If the experiment needs to be split up, (e.g. if applying the treatments or if making the measurements takes several hours or days) then this can be done in any way as the subjects have already been randomised. However, if splitting the experiment up in this way is likely to introduce an unknown source of variation, then the design loses power. In such circumstances a randomised block design might be preferable.

Statistical Experimental Design

The randomised block design



In this design the experimental material is split up into a number of "mini-experiments", typically with one subject on each treatment. It is assumed that differences between treatments *are* of interest while differences between blocks, which are random effects are of *no* interest.

Subjects are matched using any criteria available at the time the experiment is started. This might be on size (as above), space (e.g. location within the animal house such as shelf level) or time (as in within-litter experiments, where litters are infrequent). Blocks can differ in several ways at the same time. For example, block 1 might be large animals held on the top shelf and processed on day 1.

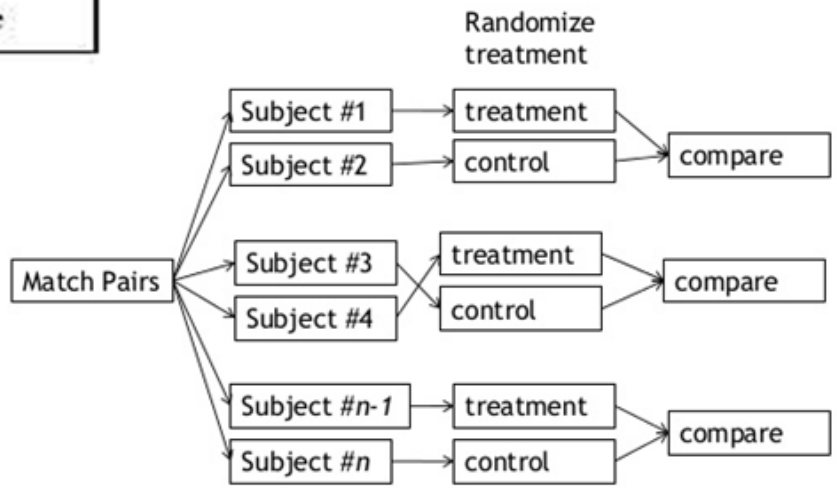
Although it is usual to have only a single experimental unit of each treatment in a block, it is possible to have two or more.

Statistical Experimental Design

IV = sleep		
	1. 2 hours sleep	2. 10 hours sleep
	Group A (10 students)	Group B (10 students, matched for age, gender, normal sleeping length)
DV	Reaction Time	Reaction Time

Matched-pairs design

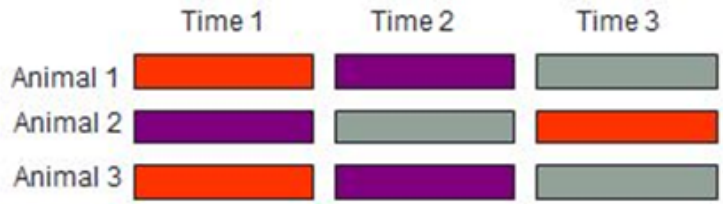
Pairs of individuals with similar characteristics are given two different treatments.



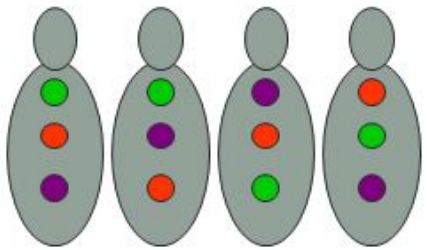
Statistical Experimental Design

Cross-over design

In which the experimental unit is the animal (or other entity) for a period of time. Each subject receives different treatments sequentially and it is assumed that the treatment does not permanently alter the subject. The blocking factor is time, with all animals being measured at each time.



Individual animals can be "blocks". In this case different treatments are applied to the shaved back of an animal. The experimental unit is an area of skin and it is assumed that the treatments do not interact with each other.

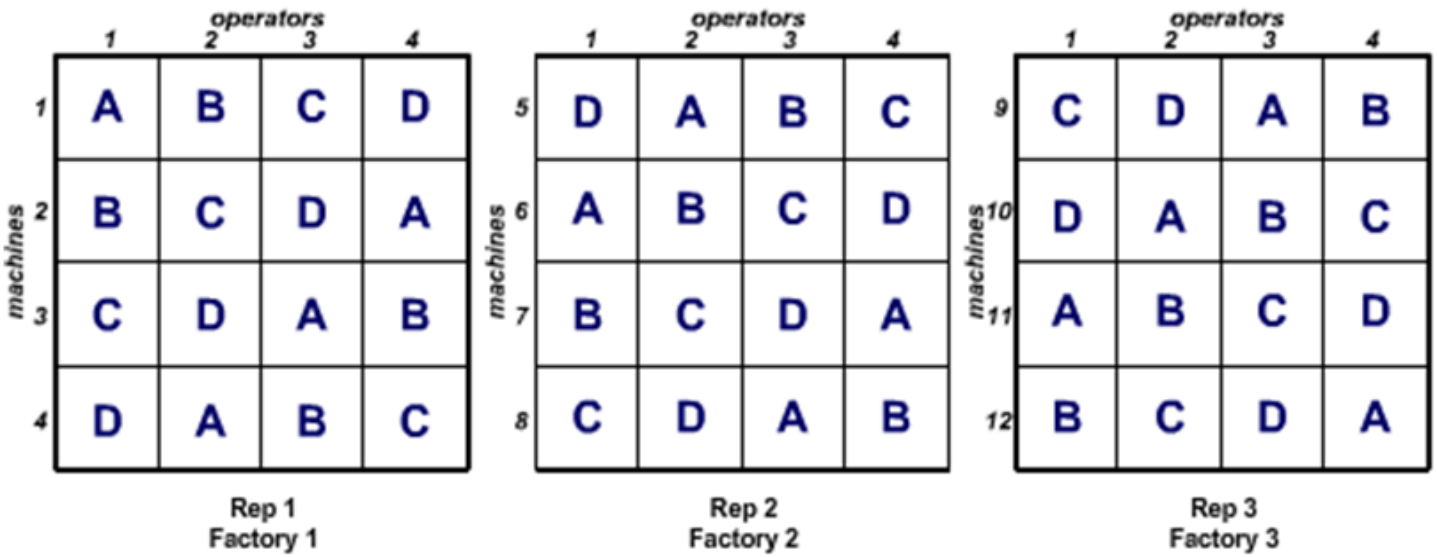


Blocks can be set up at different times (even weeks apart) and/or housed in different locations.

Statistical Experimental Design

Latin squares design

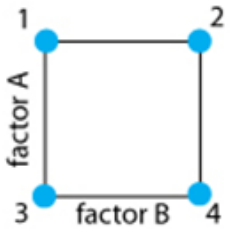
Capable of removing the effect of two blocking variables (e.g., operators and machines) and concentrate on the treatment (A,B,C,D). The following example is a replicated latin squares design.



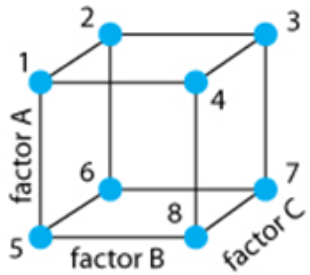
Statistical Experimental Design

Factorial design

Capable of analyzing the effect of two or more fixed effects variables (treatments)

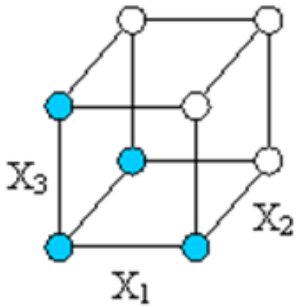


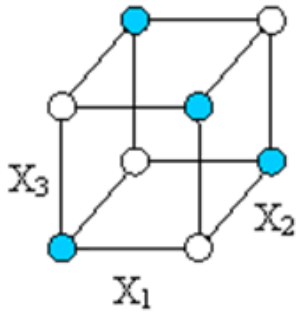
factor levels		
trial	A	B
1	+	−
2	+	+
3	−	−
4	−	+



factor levels			
trial	A	B	C
1	+	−	−
2	+	−	+
3	+	+	+
4	+	+	−
5	−	−	−
6	−	−	+
7	−	+	+
8	−	+	−

Full factorial



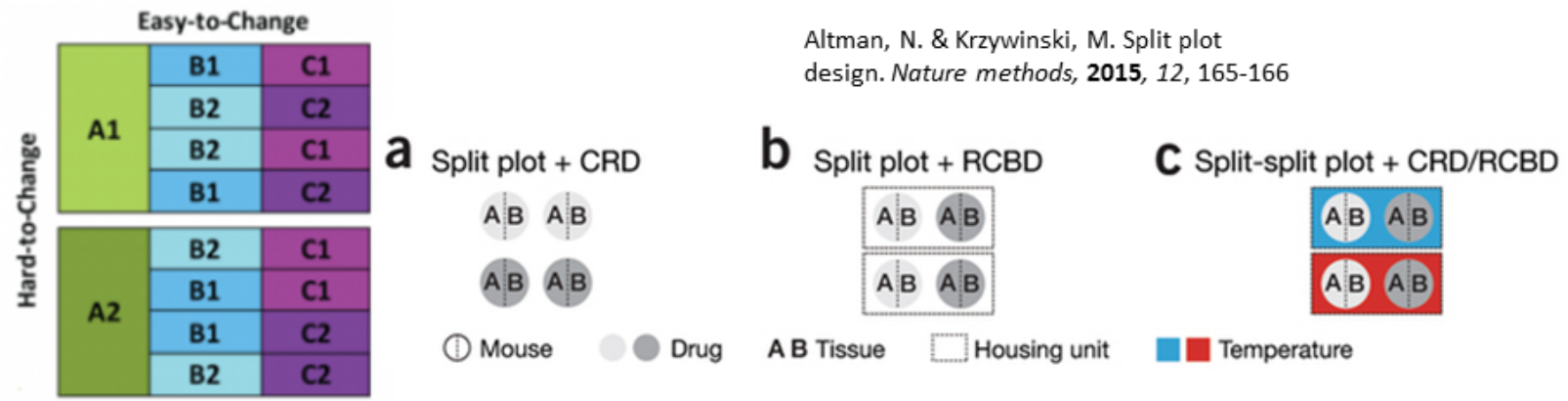


Fractional factorial

Statistical Experimental Design

Split-plot design

A factorial design in which some factors are easy to change, and some others are not. The hard-to-change factors are assigned to a whole-plot. Within this plot, the easy-to-change factors are analyzed in a factorial way.



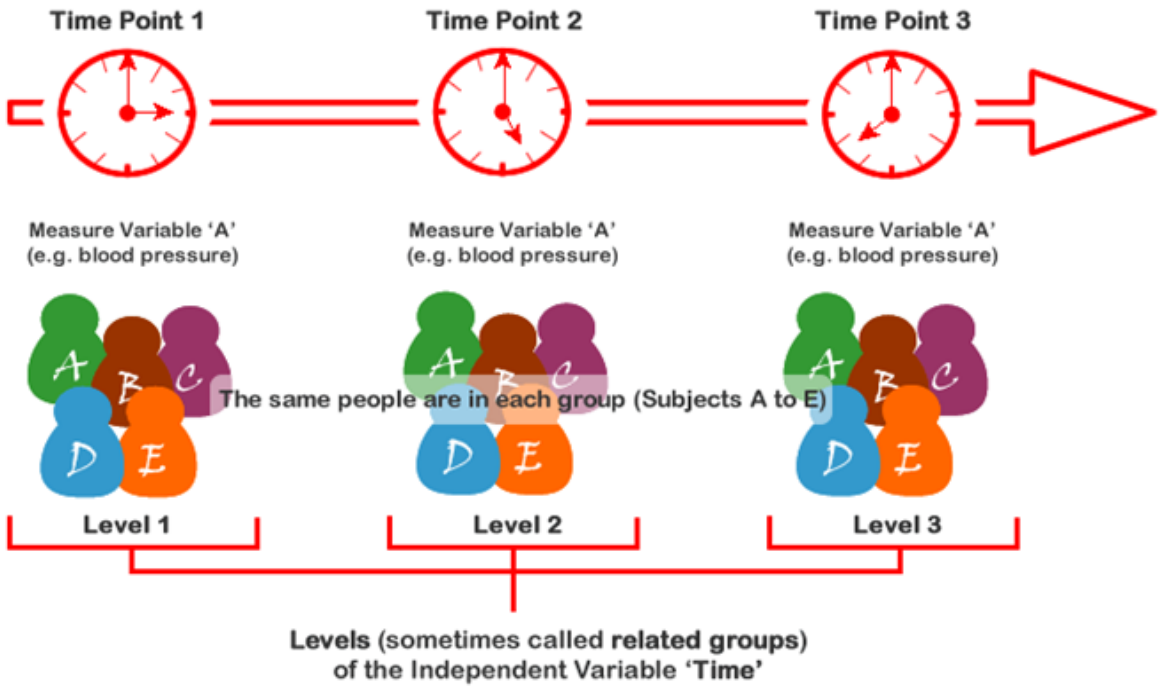
Altman, N. & Krzywinski, M. Split plot design. *Nature methods*, **2015**, *12*, 165-166

"How to Recognize a Split Plot Experiment" by Scott M. Kowalski and Kevin J. Potcner, Quality Progress, November 2003.

Statistical Experimental Design

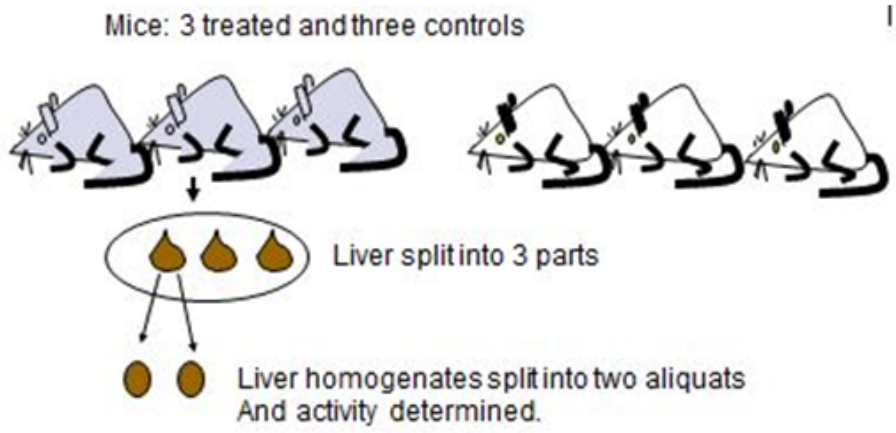
Repeated Measures design

In which each experimental unit is measured several times *without* different treatments being applied and time effects *are* of interest. Note that some authors use the term "repeated measures designs" for crossover experiments in which a subject receives different treatments over a period of time.



Statistical Experimental Design

Hierarchical design



In these designs more than one sample is taken from each experimental unit, and in some case the samples are sub-sampled, as illustrated, where the liver of each individual is split into three parts, homogenised and then determinations done on two aliquots from each part. The usual aim is to increase power by reducing measurement error. Sometimes the terms "technical replication" and "biological replication" are used. The former refers to replication of measurements on the same experimental unit.

These designs help to answer questions such as whether it is better to do more measurements on each experimental unit (which could be relatively inexpensive) or use more experimental units, if the aim is to increase power. In general if the measurements on each experimental unit are variable, then that is where there should be more replication. If they are similar, then more experimental units should be used (ethical considerations being taken into account).

Outline

Summary

Design Guidelines

Number of Factors	Comparative Objective	Screening Objective	Response Surface Objective
1	1-factor completely randomized design	——	——
2 - 4	Randomized block design	Full or fractional factorial	Central composite or Box-Behnken
5 or more	Randomized block design	Fractional factorial or Plackett-Burman	Screen first to reduce number of factors

Design Selection Guidelines

Summary

Mead's Resource equation:

$$T + B + E = N - 1$$

where

- N: Number of experimental units
- T: Number of treatments
- N: Number of blocks
- E: Number of degrees of freedom for the residual error. It must be between 10 and 20. Below 10, the experiment lacks of statistical power. Above 20, it may be a waste of resources.

Outline

Chapter 1. Basic designs

C.O.S. Sorzano
coss@cnb.csic.es

National Center of Biotechnology (CSIC)

October 14, 2016

Outline

Outline

Completely Randomized Design

Example 0

We are testing a new drug (X 325mg) for blood pressure versus a placebo on 1000 people. We divide the group of people in two equal groups of 500 people. Each person will be randomly assigned to the treatment or the placebo.

y_{11}	y_{21}
y_{12}	y_{22}
...	...
$y_{1,500}$	$y_{2,500}$

- $y_{1\cdot}, y_{2\cdot}$: Means of each one of the groups
- $y_{\cdot\cdot}$: Overall mean

Completely Randomized Design

The data (blood pressure) is supposed to be generated as

$$y_{jk} = \mu + t_j + \epsilon_{jk}$$

- μ is the average blood pressure of the whole population.
- t_1 and t_2 are the effects of the drug (t_1) and the placebo (t_2). It must be

$$\sum_j t_j = 0$$

- y_{jk} is the measurement observed for the k-th individual who has been given treatment j.
- ϵ_{jk} is the part of the observed measurement that cannot be explained by the average and the treatment.

Completely Randomized Design

$$y_{jk} = \mu + t_j + \epsilon_{jk}$$

- $y_{..}$: average of all observations

$$y_{..} = \frac{1}{n} \sum_{jk} y_{jk} \approx \mu$$

- $y_{j.}$: average of observations in treatment j

$$y_{j.} = \frac{1}{n_j} \sum_{k} y_{jk} \approx \mu + t_j$$

Completely Randomized Design

The total variation of the data is

$$
\begin{aligned}
SS &= \sum_{jk}(y_{jk} - y_{..})^2 = \sum_{jk}\left(y_{jk}^2 + y_{..}^2 - 2y_{jk}y_{..}\right) \\
&= \sum_{jk} y_{jk}^2 + \sum_{jk} y_{..}^2 - \sum_{jk} 2y_{jk}y_{..} = \sum_{jk} y_{jk}^2 + ny_{..}^2 - 2y_{..}\sum_{jk} y_{jk} \\
&= \sum_{jk} y_{jk}^2 + ny_{..}^2 - 2ny_{..}^2 = \sum_{jk} y_{jk}^2 - ny_{..}^2 \\
&= \sum_{jk} y_{jk}^2 - n\left(\frac{1}{n}\sum_{jk} y_{jk}\right)^2 = \sum_{jk} y_{jk}^2 - \frac{\left(\sum_{jk} y_{jk}\right)^2}{n} \boxed{= \sum_{jk} y_{jk}^2 - \frac{Y_{..}^2}{n}}
\end{aligned}
$$

Completely Randomized Design

The treatment effect is estimated as

$$\hat{t}_j = y_{j\cdot} - y_{\cdot\cdot} \approx (\mu + t_j) - \mu = t_j$$

and its associated variance

$$\boxed{SS_T = \sum_{jk} \hat{t}_j^2 = \left(\sum_j \frac{Y_{j\cdot}^2}{n_j} \right) - \frac{Y_{\cdot\cdot}^2}{n}}$$

Similarly, for the residuals

$$\hat{\epsilon}_{jk} = y_{jk} - y_{j\cdot} \approx (\mu + t_j + \epsilon_{jk}) - (\mu + t_j) = \epsilon_{jk}$$

the sum of squares of the residuals (within the treatments)

$$\boxed{SS_\epsilon = \sum_{jk} \hat{\epsilon}_{jk}^2 = \sum_{jk} y_{jk}^2 - \sum_j \frac{Y_{j\cdot}^2}{n_j}}$$

Completely Randomized Design

The sum of squares of all measurements can be decomposed into a sum of different components

$$\begin{array}{rcl} SS & = & SS_T + SS_\epsilon \\ \sum_{jk}(y_{jk} - y_{..})^2 & = & \sum_{jk}(y_{j.} - y_{..})^2 + \sum_{jk}(y_{jk} - y_{j.})^2 \end{array}$$

and similarly for the degrees of freedom

$$n - 1 = \sum_j (n_j - 1) + (t - 1)$$

Remind in our example, $n = 1000$ (=total population), $t = 2$ (two treatments: drug and placebo), and $n_1 = n_2 = 500$ (500 individuals in each treatment).

Completely Randomized Design

Normally this is presented in a table

Source	Sum of Squares (SS)	Degrees of freedom (df)	Mean squares (MS=SS/df)
Treatments	$SS_T = \sum_{jk}(y_{j\cdot} - y_{\cdot\cdot})^2$	$t-1$	$MS_T = \frac{SS_T}{df_t}$
Residuals	$SS_\epsilon = \sum_{jk}(y_{jk} - y_{j\cdot})^2$	$\sum_j(n_j - 1) = n - t$	$MS_\epsilon = \frac{SS_\epsilon}{df_\epsilon}$
Total	$SS = \sum_{jk}(y_{jk} - y_{\cdot\cdot})^2$	$n-1$	

If the residuals are normally distributed, then the Linear Model checks whether the treatments have a significant contribution explaining the variance through a F-Snedecor statistic with $t-1$ and $\sum_j(n_j - 1)$ degrees of freedom.

$$F = \frac{MS_T}{MS_\epsilon}$$

Completely Randomized Design

Example 1

Let us assume that the table in our case is

Source	SS	df	MS=SS/df
Treatments	256.88	1	256.88
Residuals	13600.28	998	13.61
Total	13857.16	999	

Note

$$\begin{aligned} 13857.16 &= 256.88 + 13600.28 \\ 999 &= 1 + 998 \end{aligned}$$

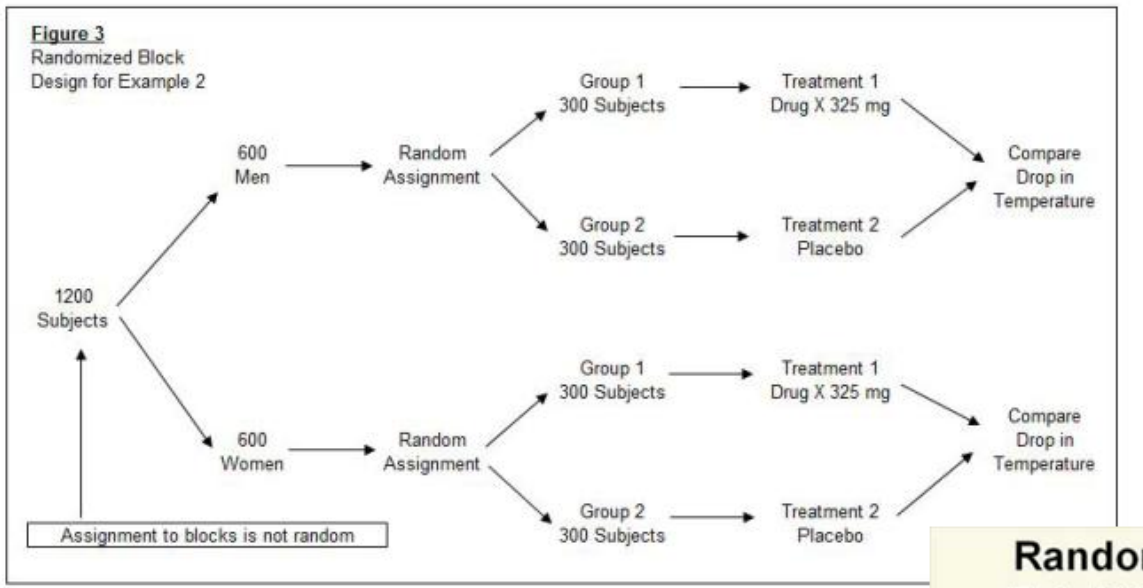
In this case

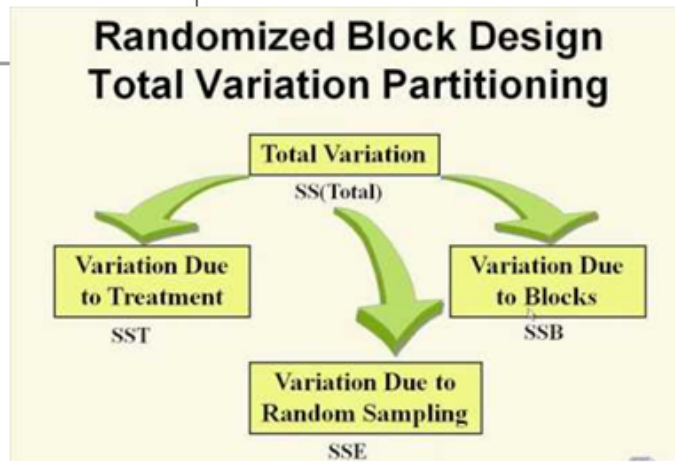
$$F = \frac{256.88}{13.61} = 18.87 \gg 3.85 = F_{0.95,1,998}$$

Outline

Reducing variance

Randomized block designs





Randomized Complete Block Design

Blocks are groups of experimental units that are formed to be as homogeneous as possible with respect to the block characteristics. The term block comes from the agricultural heritage of experimental design where a large block of land was selected for the various treatments, that had uniform soil, drainage, sunlight, and other important physical characteristics. Homogeneous clusters improve the comparison of treatments by randomly allocating levels of the treatments within each block. (SAS)

2 X 2

4 207	6 208	9 209	8 407	3 408	6 409
7 204	2 205	3 206	1 404	4 405	9 406
5 201	1 202	8 203	7 401	5 402	2 403
2 107	4 108	5 109	9 307	4 308	1 309
1 104	8 105	7 106	6 304	3 305	2 306
9 101	3 102	6 103	8 301	5 302	7 303

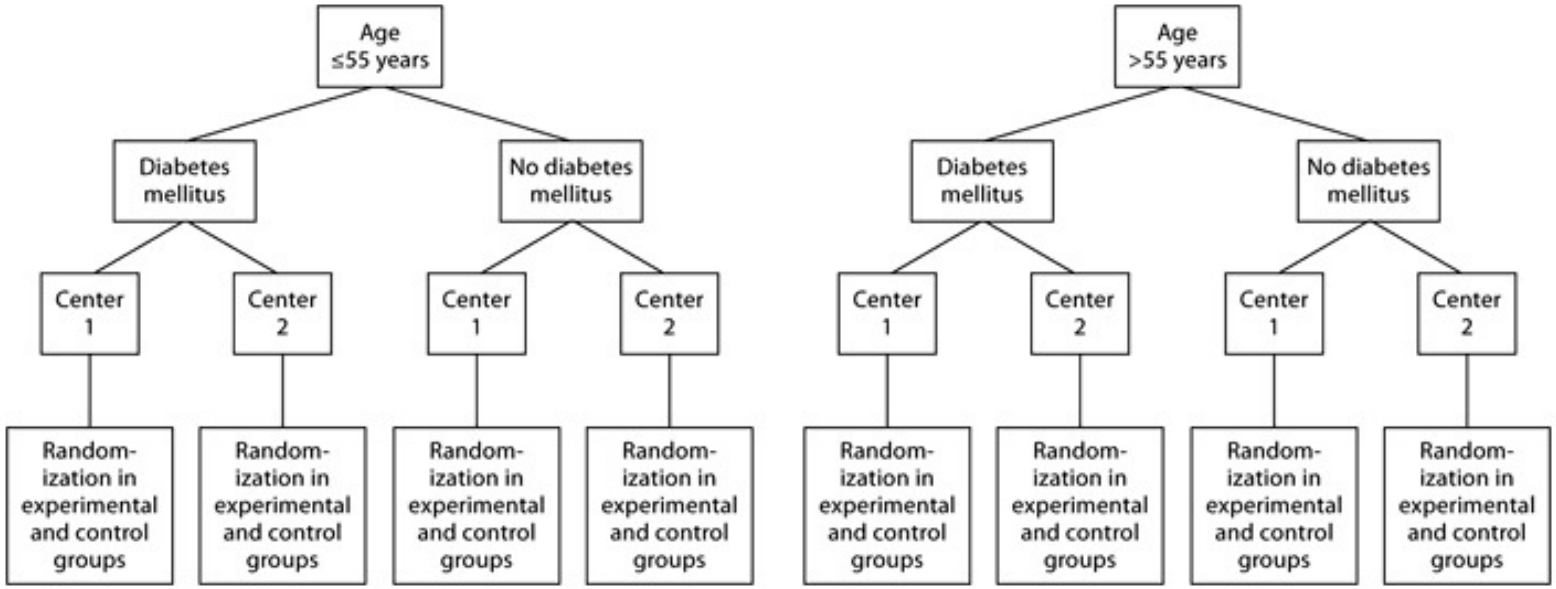
4 x 1

8 407	3 408	6 409
1 404	4 405	9 406
7 401	5 402	2 403
9 307	4 308	1 309
6 304	3 305	2 306
8 301	5 302	7 303
4 207	6 208	9 209
7 204	2 205	3 206
5 201	1 202	8 203
2 107	4 108	5 109
1 104	8 105	7 106
9 101	3 102	6 103

1 x 4

2 107	4 108	5 109	4 207	6 208	9 209	9 307	4 308	1 309	8 407	3 408	6 409
1 104	8 105	7 106	7 204	2 205	3 206	6 304	3 305	2 306	1 404	4 405	9 406
9 101	3 102	6 103	5 201	1 202	8 203	8 301	5 302	7 303	7 401	5 402	2 403

Randomized Complete Block Design



Within each block, experimental units must be randomly assigned to treatments. When several variables must be blocked, each combination (e.g. >55, Diabetes, Center 1) can be treated as a block. Alternatively, each block may be treated independently (we will see how later).

Randomized Complete Block Design

The data (blood pressure) is supposed to be generated as

$$y_{ijk} = \mu + b_i + t_j + \epsilon_{ijk}$$

- μ is the average blood pressure of the whole population.
- b_1 and b_2 are the differences in blood pressure between men (b_1) and women (b_2), the blocks. It must be
$$\sum_i b_i = 0$$
- t_1 and t_2 are the effects of the drug (t_1) and the placebo (t_2). It must be
$$\sum_j t_j = 0$$
- y_{ijk} is the measurement observed for the k-th individual of the i-th block who has been given treatment j.
- ϵ_{ijk} is the part of the observed measurement that cannot be explained by the average, block and treatment.

Randomized Complete Block Design

We now have the relationships

$$
\begin{aligned}
\hat{\mu} &= y_{...} \\
\hat{b}_i &= y_{i..} - y_{...} \approx (\mu + b_i) - \mu = b_i \\
\hat{t}_j &= y_{.j.} - y_{...} \approx (\mu + t_j) - \mu = t_j \\
\hat{\epsilon}_{ijk} &= y_{ijk} - y_{i..} - y_{.j.} + y_{...} = y_{ijk} - (\hat{\mu} + \hat{b}_i + \hat{t}_j) \\
&\approx (\mu + b_i + t_j + \epsilon_{ijk}) - (\mu + b_i) - (\mu + t_j) + \mu = \epsilon_{ijk} \\
SS &= \sum_{ijk}(y_{ijk} - y_{...})^2 = \sum_{ijk} y_{ijk}^2 - \frac{Y_{...}^2}{n} \\
SS_B &= \sum_{ijk} \hat{b}_i^2 \\
SS_T &= \sum_{ijk} \hat{t}_j^2 \\
SS_\epsilon &= \sum_{ijk} \hat{\epsilon}_{ijk}^2
\end{aligned}
$$

$$
\boxed{
\begin{aligned}
SS &= SS_B + SS_T + SS_\epsilon \\
n-1 &= (b-1) + (t-1) + (n-b-t+1)
\end{aligned}
}
$$

Randomized Complete Block Design

The table of the linear model becomes

Source	SS	df	MS=SS/df
Blocks	SS_B	$b-1$	$MS_B = \frac{SS_B}{df_B}$
Treatments	SS_T	$t-1$	$MS_T = \frac{SS_T}{df_T}$
Residuals	SS_ϵ	$n-b-t+1$	$MS_\epsilon = \frac{SS_\epsilon}{df_\epsilon}$
Total	SS	$n-1$	

If the residuals are Gaussian, we may test whether the contribution of the blocks or treatments are significant through the same F-Snedecor as before (pay attention to use the corresponding degrees of freedom).

Randomized Complete Block Design

Example 2

Let us assume that in our case it becomes

Source	SS	df	MS=SS/df
Blocks	1500.04	1	1500.04
Treatments	256.88	1	256.88
Residuals	12100.24	997	12.13
Total	13857.16	999	

Note

$$\begin{aligned} 13857.16 &= 1500.04 + 256.88 + 12100.24 \\ 999 &= 1 + 1 + 997 \end{aligned}$$

In this case

$$F = \frac{256.88}{12.13} = 21.17 \gg 3.85 = F_{0.95,1,997}$$

Randomized Complete Block Design

Example 3

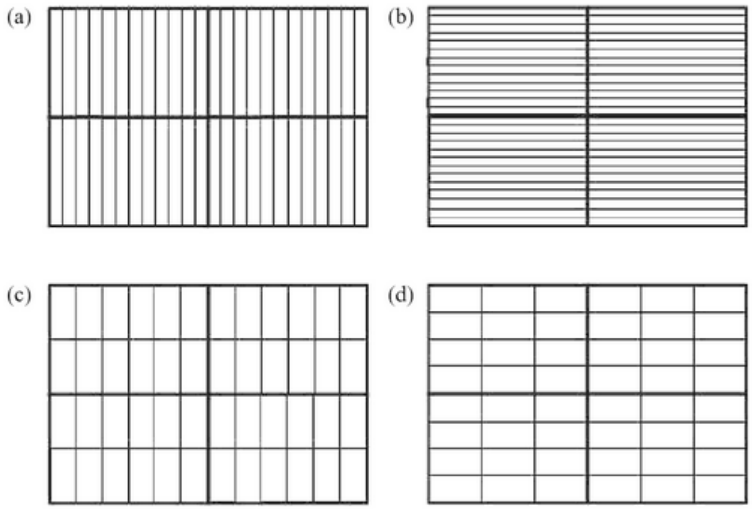
We want to analyze the optimal spacing (in terms of yield measured in kilos) between plants (10 treatments: 30×30, 30×24, 30×20, 30×15, 24×24, 24×20, 24×15, 20×20, 20×15, 15×15). To avoid possible land effects, we divide the land in 4 blocks, and within each block we randomly apply the 10 treatments.

We may compute the difference between many pairs of treatments, creating a problem of Type I error inflation by multiple testing. Instead, we may analyze the data converting the treatments to a numerical variable (area per plant, e.g. $30 \times 30 = 900$) and performing a regression analysis of yield versus area and making the hypothesis testing only on a single parameter, the slope.

Randomized Complete Block Design

- If there are clear variables to block, they should be blocked. Litters are normally chosen as blocks (and birth weight as covariate).

- If there are no obvious blocking variables, but we may create blocks, we may do as an "insurance" against possible patterns not yet identified.

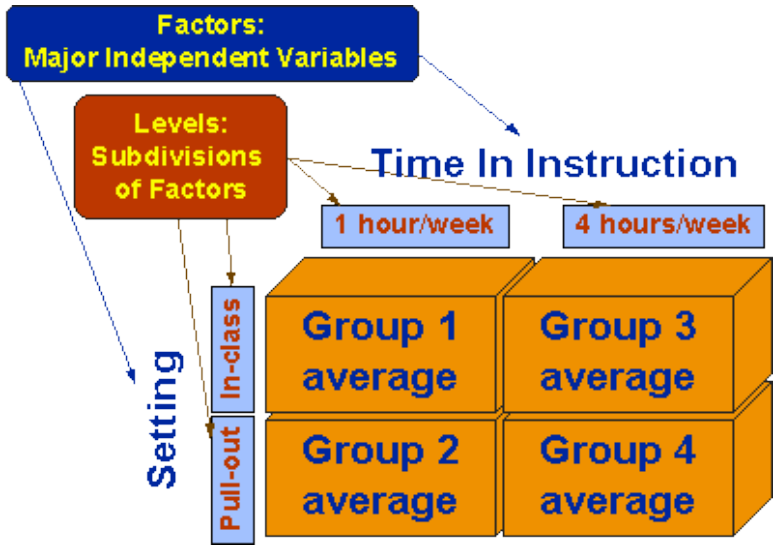


(e.g. 4 block, 12 treatments)

Outline

Factorial Design

Let's imagine a design where we have an educational program where we would like to look at a variety of program variations to see which works best. For instance, we would like to vary the amount of time the children receive instruction with one group getting 1 hour of instruction per week and another getting 4 hours per week. And, we'd like to vary the setting with one group getting the instruction in-class (probably pulled off into a corner of the classroom) and the other group being pulled-out of the classroom for instruction in another room.

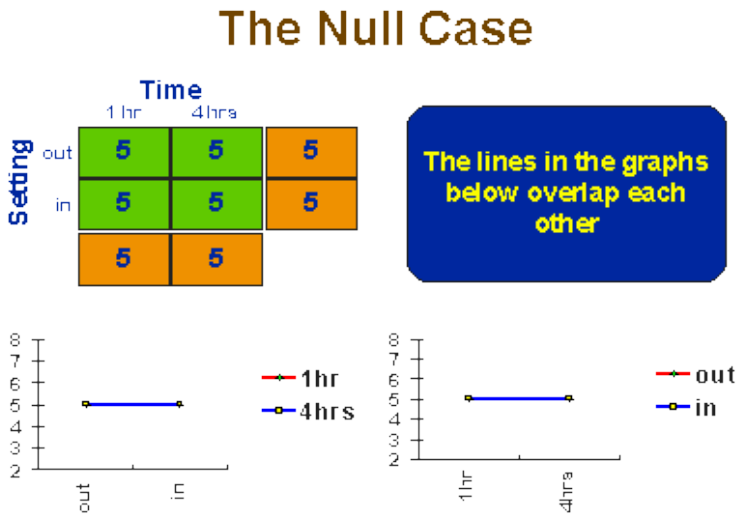


Factorial Design

The data is supposed to be generated as

$$y_{ijk} = \mu + p_i + q_j + \epsilon_{ijk}$$

Treatment variables are P(=amount of time) and Q (=setting).

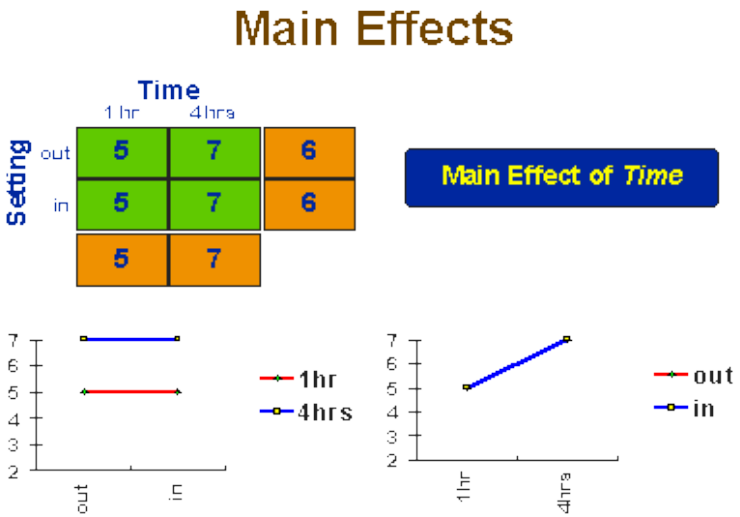


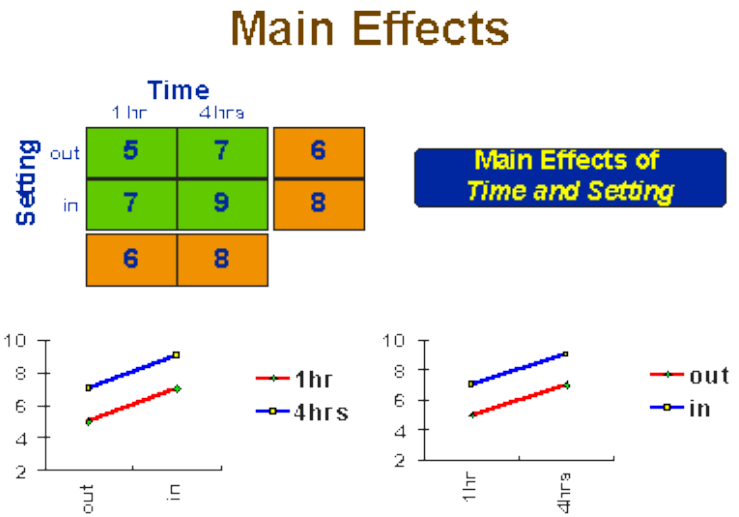
In case that there is no effect of any of the variables, we should not observe differences amongst the groups.

5+0+0	5+0+0	$q_1 = 0$
5+0+0	5+0+0	$q_2 = 0$
$p_1 = 0$	$p_2 = 0$	$\mu = 5$

Factorial Design

Main effects are the consistent differences observed for the levels of each one of the factors.





Outcome example if the amount of time has an effect but the setting does not.

6-1+0	6+1+0	$q_1 = 0$
6-1+0	6+1+0	$q_2 = 0$
$p_1 = -1$	$p_2 = 1$	$\mu = 6$

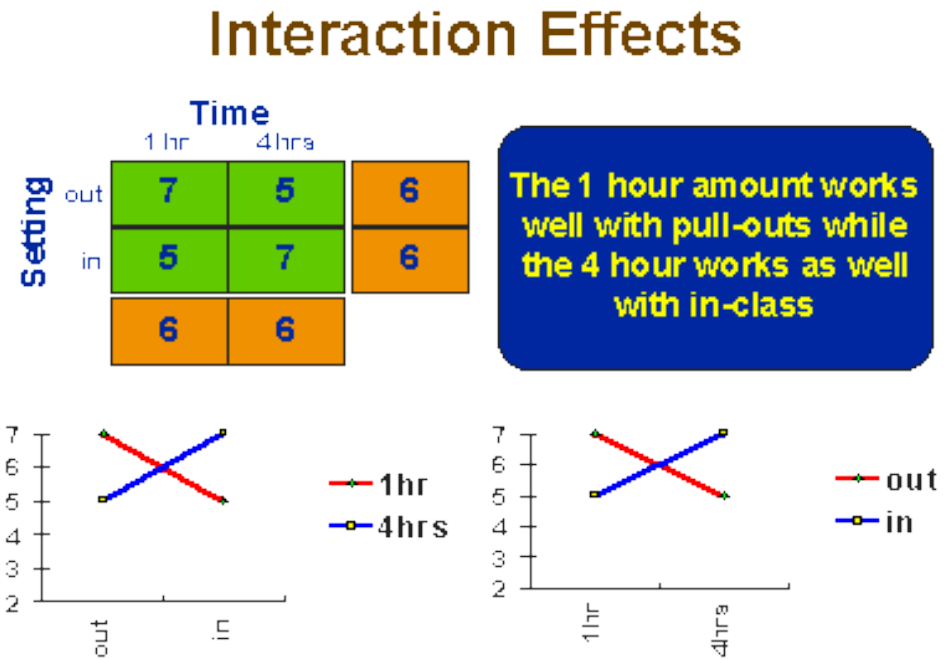
Outcome example if the amount of time and the setting have an effect.

7-1-1	7+1-1	$q_1 = -1$
7-1+1	7+1+1	$q_2 = 1$
$p_1 = -1$	$p_2 = 1$	$\mu = 7$

Factorial Design

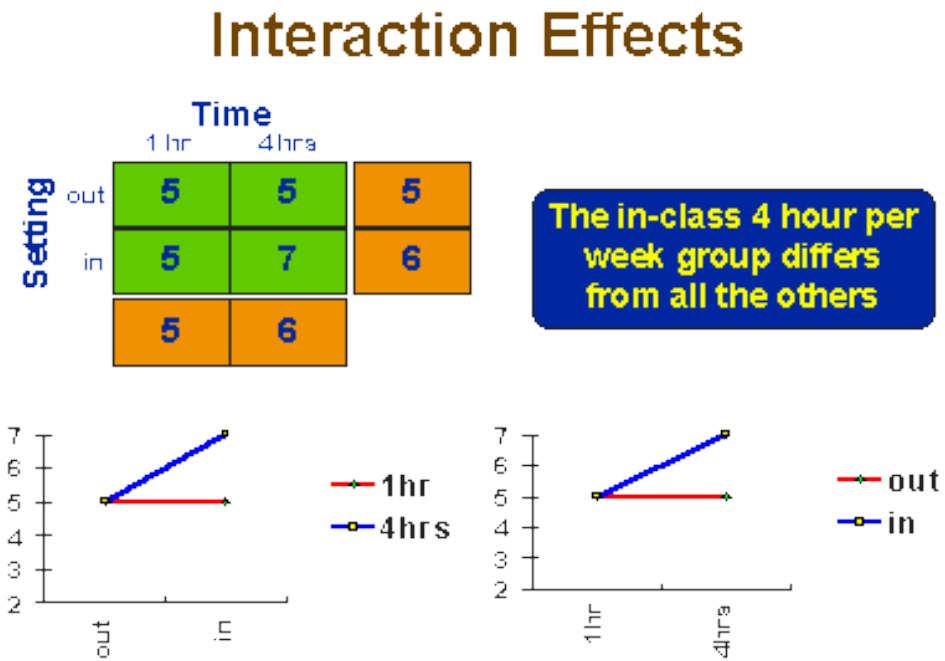
Interaction effects exist when differences on one factor depend on the level you are on another factor. The interactions are between factors and not between levels.

$$y_{ijk} = \mu + p_i + q_j + (pq)_{ij} + \epsilon_{ijk}$$



6+0+0+1$(= (pq)_{11})$	6+0+0-1$(= (pq)_{12})$	$q_1 = 0$
6+0+0-1$(= (pq)_{21})$	6+0+0+1$(= (pq)_{22})$	$q_2 = 0$
$p_1 = 0$	$p_2 = 0$	$\mu = 6$

Factorial Design



5.5-0.5-0.5+0.5$(= (pq)_{11})$	5.5+0.5-0.5-0.5$(= (pq)_{12})$	$q_1 = -0.5$
5.5-0.5+0.5-0.5$(= (pq)_{21})$	5.5+0.5+0.5+0.5$(= (pq)_{22})$	$q_2 = 0.5$
$p_1 = -0.5$	$p_2 = 0.5$	$\mu = 5.5$

Factorial Design

Given the linear model

$$y_{ijk} = \mu + p_i + q_j + (pq)_{ij} + \epsilon_{ijk}$$

The model constraints are

$$\sum_i p_i = \sum_j q_j = \sum_i (pq)_{ij} = \sum_j (pq)_{ij} = 0$$

and we may estimate each one of the components as

$$\hat{\mu} = y_{...} \qquad SS = \sum_{ijk}(y_{ijk} - \hat{\mu})^2 \qquad df = n - 1$$

$$\hat{p}_i = y_{i..} - y_{...} \qquad SS_P = \sum_{ijk} \hat{p}_i^2 \qquad df_P = p - 1$$

$$\hat{q}_j = y_{.j.} - y_{...} \qquad SS_Q = \sum_{ijk} \hat{q}_j^2 \qquad df_Q = q - 1$$

$$\widehat{(pq)}_{ij} = y_{ij.} - y_{i..} - y_{.j.} + y_{...} \qquad SS_{PQ} = \sum_{ijk} \widehat{(pq)}_{ij}^2 \qquad df_{PQ} = (p-1)(q-1)$$

$$\hat{\epsilon}_{ijk} = y_{ijk} - y_{ij.} \qquad SS_\epsilon = \sum_{ijk} \hat{\epsilon}_{ijk}^2 \qquad df_\epsilon = n - pq$$

Factorial Design

The analysis table may be represented as

Source	SS	df	MS=SS/df
P main effects	SS_P	$p-1$	$MS_P = \frac{SS_P}{df_P}$
Q main effects	SS_Q	$q-1$	$MS_Q = \frac{SS_Q}{df_Q}$
PQ interactions	SS_{PQ}	$(p-1)(q-1)$	$MS_{PQ} = \frac{SS_{PQ}}{df_{PQ}}$
Residuals	SS_ϵ	$n-pq$	$MS_\epsilon = \frac{SS_\epsilon}{df_\epsilon}$
Total	SS	$n-1$	

Factorial Design

Example 4

We are testing water uptake by amphibia. Frogs and toads (species factor S) are kept in most or dry conditions before the experiment (moisture factor M) and half of the animals are injected with a mammalian water balance hormone (hormone factor H). A full factorial experiment is performed with 2 animals per treatment combination (cell).

$$y_{ijkl} = \mu + s_i + m_j + h_k + (sm)_{ij} + (sh)_{ik} + (mh)_{jk} + \epsilon_{ijkl}$$

Source	SS	df	MS
Species	515.06	1	
Moisture	471.33	1	
Hormone	218.01	1	
SM	39.50	1	
SH	165.12	1	
MH	57.73	1	
SMH	43.43	1	
Error	276.05	8	$s^2 = 34.51$
Total	1786.33	15	

Source	SS	df	MS
Species	515.06	1	
Moisture	471.33	1	
Hormone	218.01	1	
SH	165.12	1	
Lack of fit	140.71	3	46.90
Error	276.05	8	$s^2 = 34.51$
Total	1786.33	15	

Factorial Design

Factors and blocks: Factors and blocks may be combined, the difference between a block and a factor is that it makes no sense to study the interaction of blocks

$$y_{ijkl} = \mu + b_i + p_j + q_k + (pq)_{jk} + \epsilon_{ijkl}$$

The model constraints are

$$\sum_i b_i = \sum_j p_j = \sum_k q_k = \sum_j (pq)_{jk} = \sum_k (pq)_{jk} = 0$$

and we may estimate each one of the components as

$\hat{\mu} = y....$	$SS = \sum_{ijkl}(y_{ijkl} - \hat{\mu})^2$	$df = n - 1$
$\hat{b}_i = y_{i...} - y....$	$SS_B = \sum_{ijkl} \hat{b}_i^2$	$df_B = b - 1$
$\hat{p}_j = y_{.j..} - y....$	$SS_P = \sum_{ijkl} \hat{p}_j^2$	$df_P = p - 1$
$\hat{q}_k = y_{..k.} - y....$	$SS_Q = \sum_{ijkl} \hat{q}_k^2$	$df_Q = q - 1$
$\widehat{(pq)}_{jk} = y_{.jk.} - y_{.j..} - y_{..k.} + y....$	$SS_{PQ} = \sum_{ijkl} \widehat{(pq)}_{jk}^2$	$df_{PQ} = (p-1)(q-1)$
$\hat{\epsilon}_{ijkl} = y_{ijkl} - y_{i...} - y_{.jk.} + y....$	$SS_\epsilon = \sum \hat{\epsilon}_{ijkl}^2$	$df_\epsilon = n - pq - b - 1$

Factorial Design

Advantages of factorial design:

- Interactions between factors can be estimated and their significance tested.
- Wider validity of main effects: they have been tested in many different cases (e.g. the effect of moisture have been tested with frogs and toads, and with and without hormone)
- Several experiments are done simultaneously: the variance of pairwise comparisons is minimal, as shown in the following experiment

Factorial Design

Example 5

Assume that we have resources for 24 observations and we assume that there is no interaction between factors

$$y_{ijkl} = \mu + s_i + m_j + h_k + \epsilon_{ijkl}$$

Three different experiment designs are considered:

1. One variable changes at a time
   - (Frogs,Dry,NoHormone) vs (Toad,Dry,NoHormone): 4 animals each
   - (Frogs,Dry,NoHormone) vs (Frogs,Wet,NoHormone): 4 animals each
   - (Frogs,Dry,NoHormone) vs (Frogs,Dry,Hormone): 4 animals each
2. Do not repeat (Frogs,Dry,NoHormone) in each comparison:
   - (Frogs,Dry,NoHormone): 6 animals
   - (Toads,Dry,NoHormone): 6 animals
   - (Frogs,Wet,NoHormone): 6 animals
   - (Frogs,Dry,Hormone): 6 animals
3. Factorial design (all possible combinations) with 3 animals each.

Factorial Design

Example 6(continued)

We now want to test if there is a difference induced by the hormone injection, for which we construct the statistic

$$\Delta h = h_0 - h_1$$

Its variance in the three experiments are

1. $\sigma^2_{\Delta h} = 2\frac{\sigma^2_\epsilon}{4}$
2. $\sigma^2_{\Delta h} = 2\frac{\sigma^2_\epsilon}{6}$
3. $\sigma^2_{\Delta h} = 2\frac{\sigma^2_\epsilon}{12}$

The factorial design yields the smallest variance for the comparison of any of its components.

Factorial design: Hold all factors constant except ~~the one~~ those whose effects we are investigating.

Outline

Non-orthogonal Designs

Example 6

We are testing 2 spray treatments (t_k) using 2 different concentrations of a chemical growth regulator. We also include a control spray without the chemical. We have 9 plots (3×3) for the experiment and we allow for row (r_i) and column (c_j) differences

$$y_{ijkl} = \mu + r_i + c_j + t_k + \epsilon_{ijkl}$$

Results are	A 3.72	B 3.39	C 2.95
	C 3.50	A 3.08	B 1.72
	B 4.18	C 4.36	A 0.81

This is a latin square and the analysis techniques are not the same as in the randomized complete block design (the reason is that in block designs, for each block (in our case row and column) we assume that we have all treatments, and this is not the case.

Example 6(continued)

The solution comes through Least Squares fitting

$$
\begin{aligned}
3.72 &= \mu + r_1 + c_1 + t_A \\
3.39 &= \mu + r_1 + c_2 + t_B \\
2.95 &= \mu + r_1 + c_3 + t_C \\
3.50 &= \mu + r_2 + c_1 + t_C \\
3.08 &= \mu + r_2 + c_2 + t_A \\
1.72 &= \mu + r_2 + c_3 + t_B \\
4.18 &= \mu + r_3 + c_1 + t_B \\
4.36 &= \mu + r_3 + c_2 + t_C \\
0.81 &= \mu + r_3 + c_3 + t_A
\end{aligned}
$$

Non-orthogonal Designs

Example 6(continued)

$$\mathbf{y} = A\theta$$

$$\begin{pmatrix} 3.72 \\ 3.39 \\ 2.95 \\ 3.50 \\ 3.08 \\ 1.72 \\ 4.18 \\ 4.36 \\ 0.81 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ r_1 \\ r_2 \\ r_3 \\ c_1 \\ c_2 \\ c_3 \\ t_A \\ t_B \\ t_C \end{pmatrix}$$

However we have not introduced yet the constraints

$$r_3 = -r_1 - r_2, c_3 = -c_1 - c_2, t_C = -t_A - t_B$$

Example 6(continued)

With the constraints, the LS problem becomes

$$\begin{pmatrix} 3.72 \\ 3.39 \\ 2.95 \\ 3.50 \\ 3.08 \\ 1.72 \\ 4.18 \\ 4.36 \\ 0.81 \end{pmatrix} = \begin{bmatrix} \mu & r_1 & r_2 & c_1 & c_2 & t_A & t_B \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & -1 & -1 & -1 & -1 \\ 1 & 0 & 1 & 1 & 0 & -1 & -1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & -1 & -1 & 0 & 1 \\ 1 & -1 & -1 & 1 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 0 \end{bmatrix} \begin{pmatrix} \mu \\ r_1 \\ r_2 \\ c_1 \\ c_2 \\ t_A \\ t_B \end{pmatrix}$$

Note that for any pair of factor, their corresponding columns in the design matrix are orthogonal

$$\langle \mu, r_i \rangle = \langle \mu, c_j \rangle = \langle \mu, t_k \rangle = \langle r_i, c_j \rangle = \langle r_i, t_k \rangle = \langle c_j, t_k \rangle = 0$$

Non-orthogonal Designs

Example 7

We are now given 3 extra plots (another row), which we employ to replicate the treatments and have better estimates.

Results are now

A 3.72	B 3.39	C 2.95
C 3.50	A 3.08	B 1.72
B 4.18	C 4.36	A 0.81
C 5.45	B 5.26	A 4.85

Non-orthogonal Designs

Example 7(continued)

$$\begin{pmatrix} 3.72 \\ 3.39 \\ 2.95 \\ 3.50 \\ 3.08 \\ 1.72 \\ 4.18 \\ 4.36 \\ 0.81 \\ 5.45 \\ 5.26 \\ 4.85 \end{pmatrix} = \begin{bmatrix} \mu & r_1 & r_2 & r_3 & c_1 & c_2 & t_A & t_B \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & -1 & -1 & -1 & -1 \\ 1 & 0 & 1 & 0 & 1 & 0 & -1 & -1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & -1 & -1 \\ 1 & 0 & 0 & 1 & -1 & -1 & 1 & 0 \\ 1 & -1 & -1 & -1 & 1 & 0 & -1 & -1 \\ 1 & -1 & -1 & -1 & 0 & 1 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 & 1 & 0 \end{bmatrix} \begin{pmatrix} \mu \\ r_1 \\ r_2 \\ r_3 \\ c_1 \\ c_2 \\ t_A \\ t_B \end{pmatrix}$$

Factor columns in the design matrix are no longer orthogonal (in particular $\langle c_j, t_k \rangle \neq 0$).

Non-orthogonal Designs

- Orthogonal designs are insensitive to the order in which the parameters are fitted. We may fit all of them at the same time (as shown), or
  1. fit first μ, produce a new experiment dataset removing the part we have already fitted (μ)
  2. fit then r_i and c_j, produce a new experiment dataset removing the part we have already fitted (μ, r_i, c_j)
  3. fit finally the treatments (t_k)
- Non-orthogonal designs depend on the order in which parameters are fitted (nothing terrible, but something to keep in mind).

Outline

Covariates

Researchers cannot control covariates, but can measure them and use them to increase the predictive power of the Linear Model.

Example 8

We suspect that the effect of the growth chemical depends on the ambient temperature, we extend the model with this covariate

$$y_{ijkl} = \mu + r_i + c_j + t_k + \beta T_{ijkl} + \epsilon_{ijkl}$$

T_{ijkl} is the ambient temperature measured when the spray was applied.

A 3.72 (T=28)	B 3.39 (T=22)	C 2.95 (T=23)
C 3.50 (T=24)	A 3.08 (T=25)	B 1.72 (T=26)
B 4.18 (T=20)	C 4.36 (T=22)	A 0.81 (T=26)

Covariates

Example 8(continued)

$$\mathbf{y} = A\boldsymbol{\theta}$$

$$\begin{pmatrix} 3.72 \\ 3.39 \\ 2.95 \\ 3.50 \\ 3.08 \\ 1.72 \\ 4.18 \\ 4.36 \\ 0.81 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 28 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 22 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 23 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 24 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 25 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 26 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 20 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 22 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 26 \end{pmatrix} \begin{pmatrix} \mu \\ r_1 \\ r_2 \\ r_3 \\ c_1 \\ c_2 \\ c_3 \\ t_A \\ t_B \\ t_C \\ \beta \end{pmatrix}$$

Contrasts

Example 9

Remind that our simplified parameter vector is

$$\boldsymbol{\theta} = (\mu, r_1, r_2, c_1, c_2, t_A, t_B)^T$$

We want to know whether there is a difference in the spray treatment

$$t_A - t_B = 0 = (0, 0, 0, 0, 0, 1, -1)^T \boldsymbol{\theta}$$

or if there are differences in the rows

$$\begin{aligned} r_1 - r_2 = 0 &= (0, 1, -1, 0, 0, 0, 0)^T \boldsymbol{\theta} \\ r_2 - r_3 = 0 &= r_2 - (-r_1 - r_2) = 2r_2 + r_1 \\ &= (0, 1, 2, 0, 0, 0, 0)^T \boldsymbol{\theta} \end{aligned}$$

In general, many interesting tests are of the form $\mathbf{c}^T \boldsymbol{\theta} = 0$.
If $\mathbf{1}^T \mathbf{c} = 0$, $\mathbf{c}$ is called a contrast.

Outline

Least squares

The linear model is of the form

$$\mathbf{y} = A\boldsymbol{\theta} + \boldsymbol{\epsilon}$$

and it assumes

$$\begin{aligned} E\{\boldsymbol{\epsilon}\} &= \mathbf{0} \\ \Sigma_{\boldsymbol{\epsilon}} &= \sigma_{\epsilon}^2 I \end{aligned}$$

Consequently

$$E\{\mathbf{y}\} = A\boldsymbol{\theta}$$

And the deviations from the expected value is the sum of squares

$$SS = (\mathbf{y} - A\boldsymbol{\theta})^T(\mathbf{y} - A\boldsymbol{\theta})$$

The minimizer of this Sum of Squares is

$$\hat{\boldsymbol{\theta}} = (A^T A)^{-1} A^T \mathbf{y}$$

Least squares

The covariance matrix of the fitting parameters (assuming that ϵ is a multivariate normal) is

$$Cov\{\hat{\boldsymbol{\theta}}\} = \sigma_\epsilon^2 (A^T A)^{-1}$$

If we diagonalize $A^T A$, then after some suitable rotation P

$$Cov\{P\hat{\boldsymbol{\theta}}\} = \begin{pmatrix} \frac{\sigma_\epsilon^2}{\lambda_1^2} & 0 & \dots & 0 \\ 0 & \frac{\sigma_\epsilon^2}{\lambda_2^2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{\sigma_\epsilon^2}{\lambda_M^2} \end{pmatrix}$$

being λ_1, λ_2, ... λ_M the Singular Values of the matrix A

The goal of the Experimental Design is to construct a matrix A such that: 1) $A^T A$ has a determinant as small as possible; or 2) the variance of a specific parameter is as small as possible. We would also like the matrix A to be well-conditioned (otherwise some parameter will be too variable).

Least squares

If in our experiment the most important test is of the form

$$c = \mathbf{c}^T \boldsymbol{\theta} = 0$$

we may design our experiment such that the variance of c is minimized

$$\mathrm{Var}\{c\} = \sigma_\epsilon^2 \mathbf{c}^T (A^T A)^{-1} \mathbf{c}$$

The goal of the Experimental Design is to construct a matrix A such that: ... or 3) the variance of a specific statistic is as small as possible.

Particular structures (Factorial Design, Completely Randomized Design, Randomized Complete Block Design) are "precooked" A constructions, which additionally allow very easy Least Squares fitting.

Outline

Chapter 2. Foundations revisited

C.O.S. Sorzano
coss@cnb.csic.es

National Center of Biotechnology (CSIC)

October 14, 2016

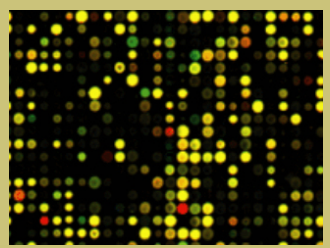
Outline

Outline

Experimental units

Experimental units are the smallest division of the experimental material such that any two experimental units can receive different treatments.
Each experimental unit gives a single observation.

Example 10



Microarrays are used to analyze the differences between the transcription of different genes. Depending on the technology 1, 2 or more samples can be hybridized to the array probes.

$$y_{ijk} = estrogen_i(Yes/No) + time_j(10h/48h) + \epsilon_{ijk}$$

Considering the gene as a treatment suggests that each spot is considered as the experimental unit

$$y_{ijkl} = estrogen_i + time_j + gene_k + \epsilon_{ijkl}$$

Experimental units

Example 11

A clinical trial for a new ointment is designed as follows. There will be 3 groups:

1. Group 1: use the new gel for 12 months.
2. Group 2: use first the new gel for 6 months and a placebo for 6 months.
3. Group 3: use first a placebo for 6 months and the new gel for 6 months.

The experimental unit are not the people in the experiment, but the period of 6 months of each person.

Experimental units

Example 12

In agricultural crop trials, the experimental unit cannot be each plant, but it is normally a plot. Plots must be large enough to be representative of large fields and remove the inter-plant variability, and small enough to be manageable and remove variability between soil differences. Long thin plots are normally preferrable.

Example 13

Trees are normally treated independently, so that they are the experimental unit.

In a clinical trial in which every patient is given a new drug or the best current treatment (or placebo), each patient is the experimental unit.

Experimental units

Example 14

In animal feeding experiments, the experimental unit is normally the pen (or cage), unless each animal can be fed independently of the rest. Experiments in which the whole group is the experimental unit are called cluster randomisation.

Example 15

Educational systems normally group children together in a way that each student cannot receive an individual treatment, the whole class is considered the experimental unit.

Experimental units

Example 16

Keith Smolkowski

Consider a study designed to test a parenting intervention that addresses child behavior at home. The parenting program teaches parents specific behaviors through classes with six to ten parents and two trainers. Assume that the content of the classes, the specific behaviors taught to parents, have been chosen through a program of research that has previously shown their efficacy with individual families. Do changes to a child's environment, through changes in his or her parents' behaviors, result in improved child behavior at home?

Experimental units

Example 16(continued)

Keith Smolkowski

This question assumes that the intervention will change parent behavior, and then it asks if those parenting behaviors influence children. Here the family represents the unit because the IV represents the change in the behavior of the parent(s) within a home and the DV accounts for the behavior of the child(ren). Individual children or parents would not do because they are not independent within a household. Parents influences each other, they influence children, and children influence parents.

Experimental units

Example 17

Keith Smolkowski

Does instruction in parenting skills change parents' behaviors at home, assuming those new parenting behaviors will lead to improved child behavior?

We have assumed that the parenting behaviors taught will lead to changes in child behavior, if implemented appropriately. The study, then, represents an attempt to show that a specific form of parent training, the parenting classes led by two trainers, can teach parents to master the skills and apply them at home. In particular, to generalize to any pair of sufficiently prepared trainers, the unit in this study must capture the parent trainers. Furthermore, parents within a class all meet at the same place, at the same time, and with the same pair of trainers. They influence each other, so groups of parents cannot be considered independent if they have the same instructors.

Experimental units

Example 18

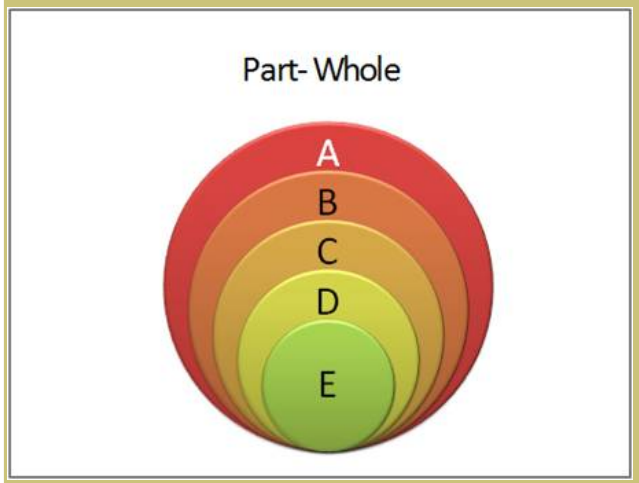
We want to detect a rare disease in a population. Blood samples are pooled into groups and the pooled sample is tested for the disease. If we cannot find it, none of the individuals in the pool has the disease. The experimental unit is the pool.

If we find the disease, we may analyze each blood sample individually to identify the person having it. The experimental unit is now the individual.

Pooling is a very effective way of cutting down costs.

Experimental units

Example 19



Using parts of a whole as experimental units reduces the variability of the measurements, e.g.,

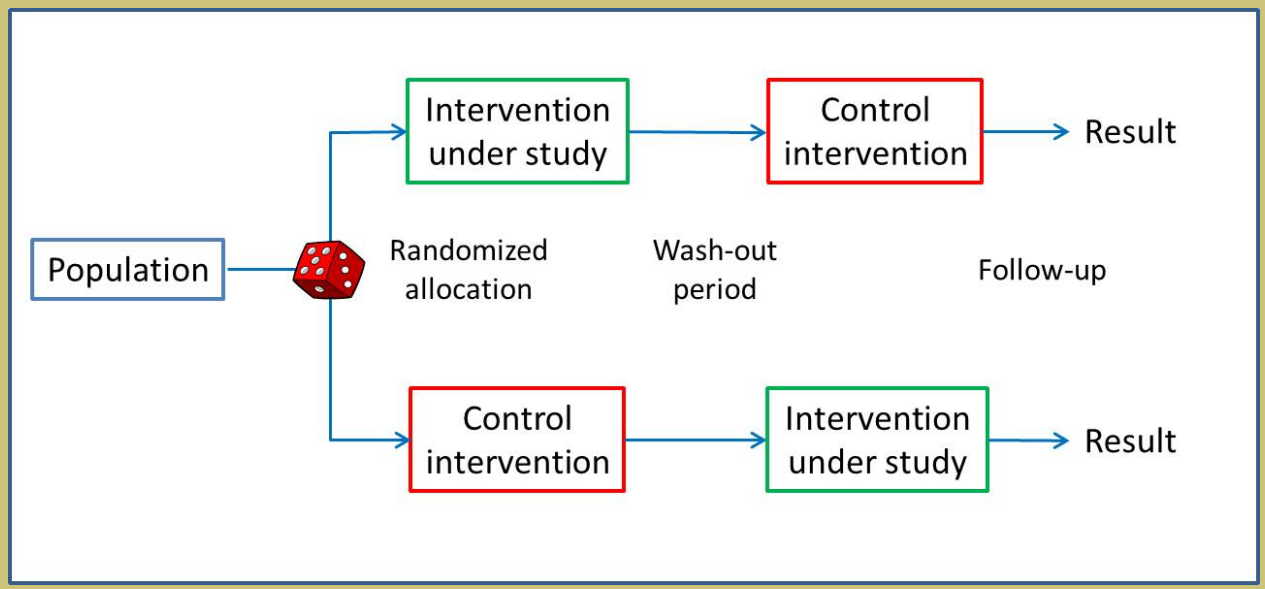
- Two leaves from the same tree may receive two different fungicide treatments.
- Two fruits from the same tree may receive two different storage treatments.
- Different parts of the same bake mix may receive different cooking treatments.
- Each eye of the same person may receive different surgical procedures.

Experimental units

Example 20

Crossover experiments

1. Group 1: use first the new gel for 6 months and a placebo for 6 months.
2. Group 2: use first a placebo for 6 months and the new gel for 6 months.



Experimental units

1. A trial in which a patient receives a sequence of different drugs cannot provide legitimate information for future treatment in which a patient will receive only a single drug. Does the washout period revert to the initial condition?
2. Once a patient has received one drug, that patient is changed, and therefore does not provide the same condition for the second drug. Does the washout period revert to the initial condition?
3. Treating a patient with a time-cocktail of drugs may provide a more dangerous situation for the patient, for which the dangers are difficult to predict. The experiment must be conceived as a whole, including a previous pharmacological study.
4. The order in which treatments are presented may affect the apparent benefits of different treatments. The order of treatments must be randomized from patient to patient.
5. It may be difficult to prevent the patient making judgements about which treatment she is receiving at a particular time and this may bias the results (whether or not these judgements are correct). Objective measurements are preferred.

Experimental units

Repeated measure design:

$$\text{Treatment} \rightarrow \text{Measure}(\text{Time}{=}t_1) \rightarrow \text{Measure}(\text{Time}{=}t_2) \rightarrow \text{Measure}(\text{Time}{=}t_3)$$

The experimental unit is the subject, and the different measures are observations at different times. This kind of designs are treated as split-plot designs.

Outline

Replication

Example 21

The following two tables show two extreme cases which should be avoided (having too few observations so that it is difficult to show the value of our experiment or so many that it was a waste of resources)

	SS	df	MS
Treatments	16	7	1.14
Errors	16	8	2
All	32	15	

$$F = \frac{1.14}{2} = 0.57$$

	SS	df	MS
Treatments	500	7	71.4
Errors	500	492	1.02
All	1000	499	

$$F = \frac{71.4}{1.02} = 70.28$$

Replicates will help in an ANOVA test to determine if at least one of the treatments makes a difference or not.

Replication

If we have a completely randomized design, that responds to the model

$$y_{ij} = \mu + t_i + \epsilon_{ij}$$

and there are n_i replicates of the treatment i and $n_{i'}$ of the treatment i', then the test to check whether one of the treatments is significantly better/worse than the other will use the statistic

$$\Delta t = t_i - t_{i'}$$

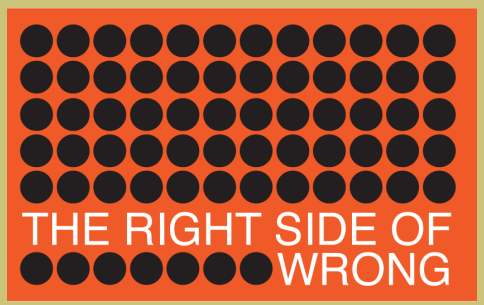
whose variance is

$$\sigma^2_{\Delta t} = \sigma^2_\epsilon \left(\frac{1}{n_i} + \frac{1}{n_{i'}} \right)$$

Replicates will help in pairwise tests to determine if one of the treatments is significantly different from another.

Replication

Example 22



- Giving a drug to two different people is a true replicate, but giving a drug twice to the same person is not.
- A microarray technical replicate is not a true replicate, but biological replicates are.
- In an animal feeding experiment, pigs within the same litter are not replicates. The experimental unit is the litter, and a true replicate is another litter.
- Automatic measurements on the same subject do not provide replicates, but time measures (see Repeated Measures Design).

Replication

Let us illustrate the effect of dealing with non-replicates as replicates.

Example 23

We want to determine the effect of 3 different teaching styles on student learning. To do so 9 classes are given one of the 3 teaching styles and a final assessment is performed. Depending on the time of the assessment along the day, there is some pattern so that in the afternoon students are more tired.

Example 23(continued)

The results of a single test are

Treatment 1			Treatment 2			Treatment 3		
A	B	C	D	E	F	G	H	I
27	43	38	41	30	47	46	34	50

whose ANOVA table is

	SS	df	MS
Treatments	81	2	40
Errors	447	6	74

$$F = \frac{40}{74} = 0.54$$

which is not too convincing.

Replication

Example 24

Assume that now we repeat the test on 4 consecutive days at the same time (to avoid the diurnal pattern)

Treatment 1			Treatment 2			Treatment 3		
A	B	C	D	E	F	G	H	I
27	43	38	41	30	47	46	34	50
25	43	36	43	35	42	48	37	44
30	46	37	44	31	46	46	38	52
31	44	41	45	35	48	45	35	49

whose ANOVA table is

	SS	df	MS
Treatments	288	2	144
Errors	1394	33	42

$$F = \frac{144}{42} = 3.4$$

much more convincing now, but we are essentially measuring 4 times the same thing, measures are not independent and they are not a true replicate.

Replication

Example 25

Assume that only one class can be tested at a time, so that each class is tested at a different time according to the pattern

	Time								
	8	9	10	11	12	13	14	15	16
Day 1	A	D	G	E	B	H	I	F	G
Day 2	H	E	I	A	F	C	D	B	G
Day 3	I	C	F	G	D	A	B	H	E
Day 4	F	B	E	H	C	I	G	D	A

whose ANOVA table is

	SS	df	MS
Treatments	322	2	162
Errors	2464	33	75

$$F = \frac{162}{75} = 2.2$$

The apparent effect has been wiped out by the diurnal pattern, and still that was incorrect because the 4 tests were not independent.

Replication

Example 25(continued)

If the tests were true replicates it would have been better to introduce a blocking variable *time*:

$$y_{ijk} = \mu + treatment_i + time_j + \epsilon_{ijk}$$

If results are not too conclusive, the solution is to apply the 3 methods to more than 3 classes each, not to take several tests on the same class. The experimental unit is the whole class, and true replicates are more classes.

Replication

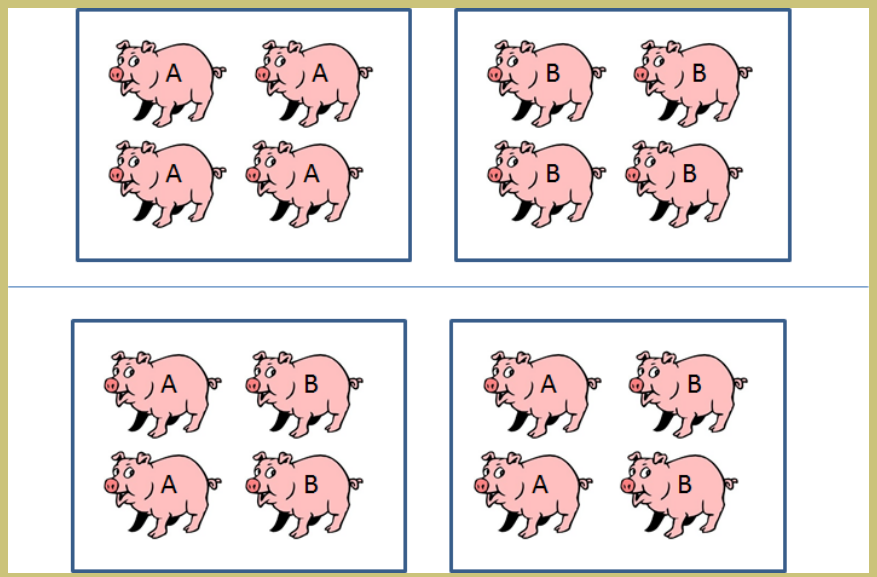
Example 26

The effectiveness of two fungicides is being tested. The fungicide may be applied on the upper or lower side of young or old leaves. This gives a total of 8 treatments (2 fungicides $\times$ 2 sides $\times$ 2 leave ages). The experimenter measures the amount of fungi in a small disc of the leave. Each treatment is replicated 10 times, obtaining $F = 500$ (which seems a little overkill).

But, fungicide 1 was applied to Tree 1, and fungicide 2 was applied to Tree 2. The 8 treatments were applied to 8 leaves, and the replicates were obtained by cutting 10 discs from the same leave. The differences could be due to the tree, and each treatment was applied only once. The 10 replicates are not true replicates, but they are measuring essentially the same thing (they come from the same leave).

Replication

Example 27



We are interested on the effect of hormones in animals. Two designs are proposed. The measure will be the weight of the animals when they are 6 months old.

Replication

Example 27(continued)

Both designs look fine (assuming that we will account for the pen effect). But the bottom design has a drawback. Let's say that A makes pigs to be more aggressive, and B more docile. So the extra A weight is not due to growth effect of A, but to the growth effect when they are fed with B animals. A animals grown alone would not have extra weight because they are all equally aggressive.

If there is an interaction between the two levels of a factor, this design is not able to detect it.

Replication

Example 28

We are interested in the effect of 3 temperature levels in the growth of greenhouse plants. Typically, there are only, at most 6 greenhouse sections with independent heating. In order to have a replicate, each treatment can be applied only to two sections. The experimental unit is the greenhouse section and two replicates is not much. Can we use individual plants as the experimental unit? Technically no.

Replication

Example 28(continued)

But since there is no other way of carrying out the research, we may if we assume that

- greenhouse sections do not affect the plant growth.
- the variation between plants are essentially due to plant-to-plant variation.
- there is no competitive variation (like in the pig case) between plants within a section, induced by the treatment.

Outline

Blocking

Example 29

The idea of blocking is to reduce the variance due to some known (uninteresting) factor, and letting most of the variance to come from the treatments. To compare A and B, we may design

	Treatments
Male	A B
Female	A B

These comparisons can be performed even if both treatments are not applied to the same block, as long as there is a common third treatment.

	Treatments
Male	A C
Female	B C

These comparisons are more accurate as the number of intermediaries increases:

	Treatments
Male	A C D
Female	B C D

Blocking (simple case)

Let us analyze the first example.

	Treatments
Block 1	A B
Block 2	A B

We assume that the data is generated according to

$$y_{ijk} = \mu + b_i + t_j + \epsilon_{ijk}$$

Let us also assume that the design is balanced and there are K replicates for each block-treatment combination.

$$\begin{aligned}
E\{Y_{1..}\} &= E\{Y_{1A.} + Y_{1B.}\} = K(\mu + b_1 + t_A) + K(\mu + b_1 + t_B) = 2K(\mu + b_1) \\
E\{Y_{2..}\} &= E\{Y_{2A.} + Y_{2B.}\} = K(\mu + b_2 + t_A) + K(\mu + b_2 + t_B) = 2K(\mu + b_2) \\
E\{Y_{...}\} &= E\{Y_{1..} + Y_{2..}\} = 4K\mu \\
E\{Y_{.A.}\} &= 2K(\mu + t_A) \\
E\{Y_{.B.}\} &= 2K(\mu + t_B)
\end{aligned}$$

Blocking (simple case)

We may write the least squares equations (taking into account the zero-mean constraints)

$$\begin{pmatrix} 4K & 0 & 0 \\ 2K & 2K & 0 \\ 2K & -2K & 0 \\ 2K & 0 & 2K \\ 2K & 0 & -2K \end{pmatrix} \begin{pmatrix} \mu \\ b_1 \\ t_A \end{pmatrix} = \begin{pmatrix} Y_{...} \\ Y_{1..} \\ Y_{2..} \\ Y_{.A.} \\ Y_{.B.} \end{pmatrix}$$

Or equivalently

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ b_1 \\ t_A \end{pmatrix} = \begin{pmatrix} Y_{...}/4K \\ Y_{1..}/2K - Y_{...}/4K \\ (Y_{1..} + Y_{2..} - Y_{...})/2K = 0 \\ Y_{.A.}/2K - Y_{...}/4K \\ (Y_{.A.} + Y_{.B.} - Y_{...})/2K = 0 \end{pmatrix}$$

Blocking

The estimates of the differents contributions of the model can be calculated as

$$\begin{aligned}\hat{\mu} &= Y_{...}/4K = y_{...} \\ \hat{b}_1 &= Y_{1..}/2K - Y_{...}/4K = y_{1..} - y_{...} \\ \hat{t}_A &= Y_{.A.}/2K - Y_{...}/4K = y_{.A.} - y_{...}\end{aligned}$$

For $\hat{b}_2 = -\hat{b}_1$ and $\hat{t}_B = -\hat{t}_A$. But it is also convenient to note the relationship

$$\frac{Y_{1..}}{2K} + \frac{Y_{2..}}{2K} - \frac{Y_{...}}{2K} = y_{1..} + y_{2..} - 2y_{...} \Rightarrow y_{...} = \frac{y_{1..} + y_{2..}}{2}, y_{1..} = 2y_{...} - y_{2..}$$

Consequently

$$\hat{b}_2 = -\hat{b}_1 = y_{...} - y_{1..} = y_{2..} - y_{...}$$

Similarly

$$\hat{t}_B = -\hat{t}_A = y_{...} - y_{.A.} = y_{.B.} - y_{...}$$

Blocking (simple case)

If we want to test whether there is a significant difference between the treatments A and B we will construct the statistic

$$\Delta_{AB} = \hat{t}_A - \hat{t}_B = y_{.A.} - y_{.B.}$$

whose variance is

$$\sigma^2_{\Delta_{AB}} = \sigma^2_{y_{.A.}} + \sigma^2_{y_{.B.}} = \frac{\sigma^2_\epsilon}{2K} + \frac{\sigma^2_\epsilon}{2K} = \frac{\sigma^2_\epsilon}{K}$$

The number of replicates needed for a two-sided hypothesis test with confidence level $1-\alpha$, power $1-\beta$ and effect size Δ must be

$$\Delta > \left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)\sigma_{\Delta_{AB}}$$

from where we can easily solve for the number of replicates

$$K > \left(\frac{\left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)\sigma_\epsilon}{\Delta}\right)^2$$

Blocking (more complicated)

Let us analyze the second example.

	Treatments
Block 1	A C
Block 2	B C

We assume that the data is generated according to

$$y_{ijk} = \mu + b_i + t_j + \epsilon_{ijk}$$

Let us also assume that the design is balanced and there are K replicates for each block-treatment combination.

$$\begin{aligned}
E\{Y_{1..}\} &= E\{Y_{1A.} + Y_{1C.}\} = K(\mu + b_1 + t_A) + K(\mu + b_1 + t_C) \\
&= 2K(\mu + b_1) + K(t_A + t_C) \\
E\{Y_{2..}\} &= E\{Y_{2B.} + Y_{2C.}\} = K(\mu + b_2 + t_B) + K(\mu + b_2 + t_C) \\
&= 2K(\mu + b_2) + K(t_B + t_C) \\
E\{Y_{...}\} &= E\{Y_{1..} + Y_{2..}\} = 4K\mu + Kt_C \\
E\{Y_{.A.}\} &= K(\mu + b_1 + t_A) \\
E\{Y_{.B.}\} &= K(\mu + b_2 + t_B) \\
E\{Y_{.C.}\} &= 2K(\mu + t_C)
\end{aligned}$$

Blocking (more complicated)

We may write the least squares equations (taking into account the zero-mean constraints)

$$\begin{pmatrix} 4K & 0 & -K & -K \\ 2K & 2K & 0 & -K \\ 2K & -2K & -K & 0 \\ K & K & K & 0 \\ K & -K & 0 & K \\ 2K & 0 & -2K & -2K \end{pmatrix} \begin{pmatrix} \mu \\ b_1 \\ t_A \\ t_B \end{pmatrix} = \begin{pmatrix} Y_{...} \\ Y_{1..} \\ Y_{2..} \\ Y_{.A.} \\ Y_{.B.} \\ Y_{.C.} \end{pmatrix}$$

Or equivalently

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ b_1 \\ t_A \\ t_B \end{pmatrix} = \begin{pmatrix} (y_{.A.} + y_{.B.} + 4y_{...})/6 \\ (4y_{1..} - y_{.A.} + y_{.B.} - 4y_{...})/2 \\ y_{1..} + y_{2..} - 2y_{...} = 0 \\ (4y_{.A.} - 6y_{1..} - 2y_{.B.} + 4y_{...})/3 \\ (6y_{1..} - 2y_{.A.} + 4y_{.B.} - 8y_{...})/3 \\ y_{.A.} + y_{.B.} + 2y_{.C.} - 4y_{...} = 0 \end{pmatrix}$$

Blocking (more complicated)

If we want to test whether there is a significant difference between the treatments A and B we will construct the statistic

$$\Delta_{AB} = \hat{t}_A - \hat{t}_B = 2y_{.A.} - 2y_{.B.} - 4y_{1..} + 4y_{...}$$

whose variance is

$$\sigma^2_{\Delta_{AB}} = 4\sigma^2_{y_{.A.}} + 4\sigma^2_{y_{.B.}} + 16\sigma^2_{y_{1..}} + 16\sigma^2_{y_{...}} = 4\frac{\sigma^2_\epsilon}{K} + 4\frac{\sigma^2_\epsilon}{K} + 16\frac{\sigma^2_\epsilon}{2K} + 16\frac{\sigma^2_\epsilon}{4K} = 20\frac{\sigma^2_\epsilon}{K}$$

For the comparison between A and C, we have

$$\Delta_{AC} = \hat{t}_A - \hat{t}_C = 2\hat{t}_A + \hat{t}_B = 2y_{.A.} - 2y_{1..}$$

whose variance is

$$\sigma^2_{\Delta_{AC}} = 4\sigma^2_{y_{.A.}} + 4\sigma^2_{y_{1..}} = 4\frac{\sigma^2_\epsilon}{K} + 4\frac{\sigma^2_\epsilon}{2K} = 6\frac{\sigma^2_\epsilon}{K}$$

Comparisons within the same block are more precise than amongst blocks.

Blocking (more complicated)

We will not analyze the third example.

	Treatments
Block 1	A C D
Block 2	B C D

But let us mention that the more treatments in common between Block 1 and Block 2 (in this case C and D), the smaller the variance of the statistics for the tests.

Blocking (incorrect design)

Let us analyze an incorrect design.

	Treatments
Block 1	A C
Block 2	B D

The design is incorrect because there is no way to distinguish the effect of the block from the treatments. Let us perform the same analysis as we did in the previous cases.

$$
\begin{aligned}
E\{Y_{1..}\} &= E\{Y_{1A.} + Y_{1C.}\} = 2K(\mu + b_1) + K(t_A + t_C) \\
E\{Y_{2..}\} &= E\{Y_{2B.} + Y_{2D.}\} = 2K(\mu + b_2) + K(t_B + t_D) \\
E\{Y_{...}\} &= E\{Y_{1..} + Y_{2..}\} = 4K\mu \\
E\{Y_{.A.}\} &= K(\mu + b_1 + t_A) \\
E\{Y_{.B.}\} &= K(\mu + b_2 + t_B) \\
E\{Y_{.C.}\} &= K(\mu + b_1 + t_C) \\
E\{Y_{.D.}\} &= K(\mu + b_2 + t_D)
\end{aligned}
$$

Blocking (incorrect design)

We may write the least squares equations (taking into account the zero-mean constraints)

$$\begin{pmatrix} 4K & 0 & 0 & 0 & 0 \\ 2K & 2K & K & 0 & K \\ 2K & -2K & -K & 0 & -K \\ K & K & K & 0 & 0 \\ K & -K & 0 & K & 0 \\ K & K & 0 & 0 & K \\ K & -K & -K & -K & -K \end{pmatrix} \begin{pmatrix} \mu \\ b_1 \\ t_A \\ t_B \\ t_C \end{pmatrix} = \begin{pmatrix} Y_{...} \\ Y_{1..} \\ Y_{2..} \\ Y_{.A.} \\ Y_{.B.} \\ Y_{.C.} \\ Y_{.D.} \end{pmatrix}$$

Or equivalently

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ b_1 \\ t_A \\ t_B \\ t_C \end{pmatrix} = \begin{pmatrix} y_{...} \\ 2y_{1..} - y_{.A.} - y_{...} \\ y_{1..} + y_{2..} - 2y_{...} = 0 \\ 2y_{.A.} - 2y_{1..} \\ 2y_{1..} - y_{.A.} + y_{.B.} - 2y_{...} \\ y_{.A.} + y_{.C.} - 2y_{1..} = 0 \\ 2y_{1..} + y_{.B.} + y_{.D.} - 4y_{...} = 0 \end{pmatrix}$$

Blocking and Orthogonality

Of the four designs in this Section the only one with an orthogonal matrix is

	Treatments
Block 1	A B
Block 2	A B

- Orthogonality keeps calculations very simple (which is good for manual calculation, but irrelevant for computers).
- It makes the estimates to be independent of the order in which they are fitted (although the variations are small).

Consequently non-orthogonality should not be considered as a major drawback of a design.
Balance (see next section) is a much more important issue.

Outline

Balanced Incomplete Block Designs

A design is balanced if:

- All treatments are applied the same number of times
- All pairs of treatments appear in the same number of blocks

For instance, the following design is balanced

	Treatments
Block 1	A B C
Block 2	A B D
Block 3	A C E
Block 4	A D F
Block 5	A E F
Block 6	B C F
Block 7	B D E
Block 8	B E F
Block 9	C D E
Block 10	C D F

- Each treatment is applied 5 times.
- Each pair (AB, AC, AD, AE, AF, BC, BD, BE, BF, CD, CE, CF, DE, DF, EF) appears 2 times.

Balanced Incomplete Block Designs

Example 30

6 treatments (A-F) are to be compared in 24 units, for which a natural blocking system gives 4 blocks of 6 units each. How can the treatments be allocated?

	Treatments
Block 1	A B C D E F
Block 2	A B C D E F
Block 3	A B C D E F
Block 4	A B C D E F

Any other allocation would repeat one treatment in one of the blocks so that pair comparisons cannot be performed in the same block making them less efficient.

If possible, apply all treatments in each block.

Balanced Incomplete Block Designs

Example 31

6 treatments (A-F) are to be compared in 30 units, for which a natural blocking system gives 6 blocks of 5 units each. How can the treatments be allocated?

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

Only 5 (instead of 6) treatments can be applied in a block, so that 1 treatment has to be skipped in each block. We may do so by removing F in the 1st block, E in the 2nd, D in the 3rd, ... Keeping a symmetric design will not favor any treatment comparison (all will have the same variance).

Keep the design as symmetric as possible.

Balanced Incomplete Block Designs

Example 32

6 treatments (A-F) are to be compared in 24 units, for which a natural blocking system gives 6 blocks of 4 units each. How can the treatments be allocated?

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

We now have to skip two treatments from each block. If this is done in a cycle, symmetry is better preserved (e.g., set of omissions (AB), (BC), (CD), (DE), (EF), (FA))

Keep the design as symmetric as possible (cycles help in this regard).

Balanced Incomplete Block Designs

Let us define these designs in general

v	No. Treatments (varieties)
b	No. Blocks
r_i	No. of blocks containing treatment i
	For a balanced design $r_i = r$ for all treatments
k	Size of the block
$\lambda_{ii'}$	No. of blocks containing treatments i and i'
	For a balanced design $\lambda_{ii'} = \lambda$ for all pairs

The designs are named (v, b, r, k, λ)-designs. A balanced design must fulfill:

$$\begin{aligned} bk &= vr \\ r(k-1) &= \lambda(v-1) \end{aligned}$$

The first equation simply states that the number of blocks times their size must be equal to the number of treatments and their repeats. $r - \lambda$ is the order of the design.

Balanced Incomplete Block Designs

Not all possible designs exist and there are different approaches to their construction

3.4.2.2 Existence table for BIBDs

Some of the most fruitful construction methods for BIBD are dealt with in separate sections, difference sets (page 167), finite geometry (page 170), Steiner triple systems (page 173), and Hadamard matrices (page 81). The table below gives all parameters for which BIBDs exist with $k \leq v/2$ and $b \leq 30$.

v	b	r	k	λ
6	10	5	3	2
6	20	10	3	4
6	30	15	3	6
7	7	3	3	1
7	14	6	3	2
7	21	9	3	3
7	28	12	3	4
8	14	7	4	3
8	28	14	4	6
9	12	4	3	1
9	18	8	4	3
9	24	8	3	2
10	15	6	4	2

v	b	r	k	λ
10	18	9	5	4
10	30	9	3	2
10	30	12	4	4
11	11	5	5	2
11	22	10	5	4
12	22	11	6	5
13	13	4	4	1
13	26	6	3	1
13	26	8	4	2
13	26	12	6	5
14	26	13	7	6
15	15	7	7	3

v	b	r	k	λ
15	30	14	7	6
16	16	6	6	2
16	20	5	4	1
16	24	9	6	3
16	30	15	8	7
19	19	9	9	4
21	21	5	5	1
21	30	10	7	3
23	23	11	11	5
25	25	9	9	3
25	30	6	5	1
27	27	13	13	6

Balanced Incomplete Block Designs

A necessary condition to be balanced is that the row and column sums of the incidence matrix are all equal

	Treatments
Block 1	A B C
Block 2	A B D
Block 3	A C E
Block 4	A D F
Block 5	A E F
Block 6	B C F
Block 7	B D E
Block 8	B E F
Block 9	C D E
Block 10	C D F

Block \ Treatment	A	B	C	D	E	F	
Block 1	1	1	1				3
Block 2	1	1		1			3
Block 3	1		1		1		3
Block 4	1			1		1	3
Block 5	1				1	1	3
Block 6		1	1			1	3
Block 7		1		1	1		3
Block 8		1			1	1	3
Block 9			1	1	1		3
Block 10			1	1		1	3
	5	5	5	5	5	5	

Balanced Incomplete Block Designs

However, this condition is not sufficient

	Treatments
Block 1	A C
Block 2	B D
Block 3	A C
Block 4	B D

Block \ Treatment	A	B	C	D	
Block 1	1		1		2
Block 2		1		1	2
Block 3	1		1		2
Block 4		1		1	2
	2	2	2	2	

The pair AC appears 2 times ($\lambda_{AC} = 2$), while AB or AD do not appear ($\lambda_{AB} = \lambda_{AD} = 2$).

Balanced Incomplete Block Designs (Cyclic design)

An easy way to design experiments is by starting with an initial block and adding 1 to each treatment modulo the number of treatments (this is called a cyclic design). For example, for 5 blocks of size 3 with 5 treatments we would have

	Treatments
Block 1 (initial)	A B D
Block 2	B C E
Block 3	C D A
Block 4	D E B
Block 5	E A C

Balanced Incomplete Block Designs (Lattice design=Youden squares)

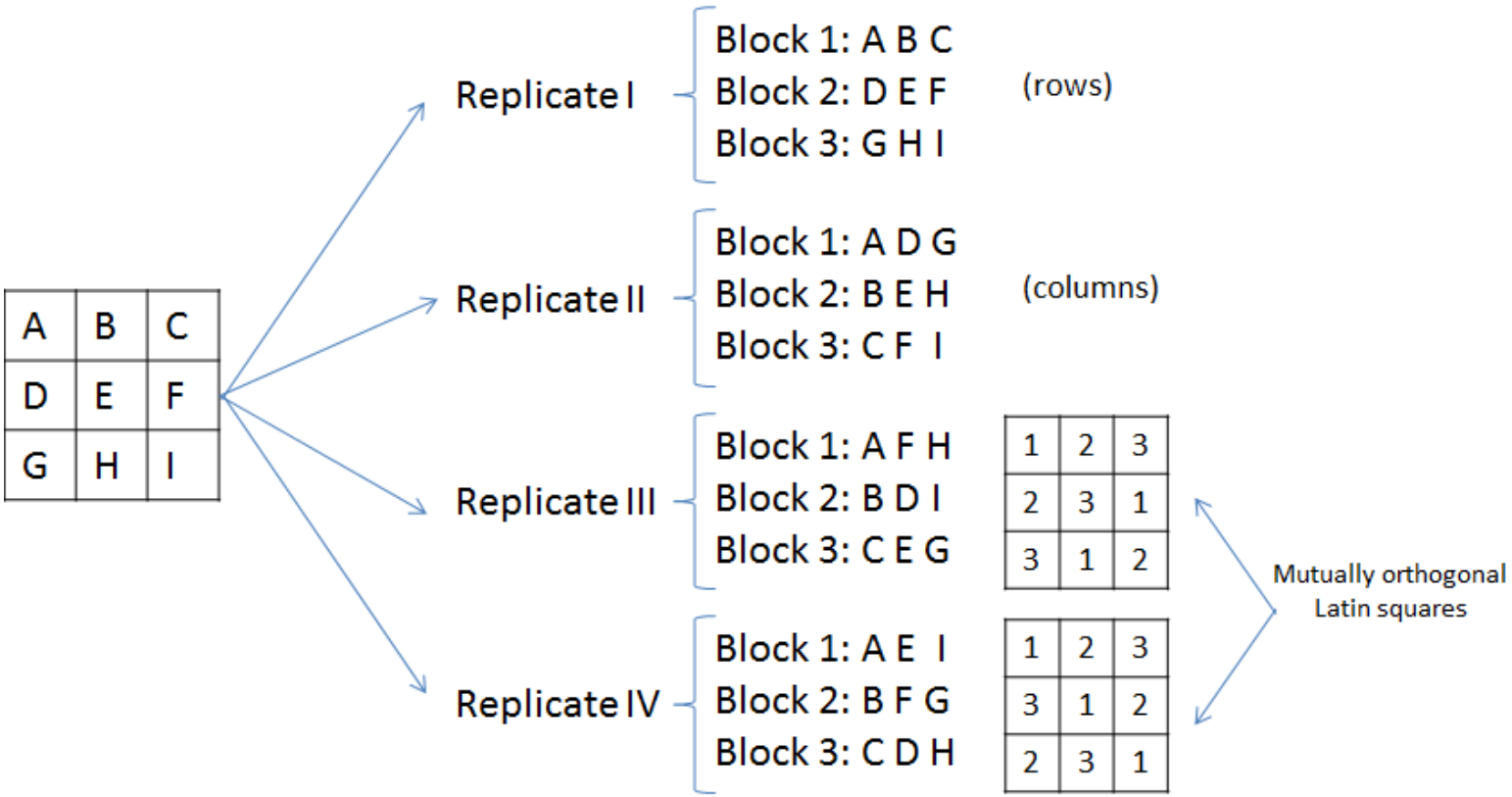
Another easy way to generate balanced incomplete designs are based on lattices (lattice design). For example, for 7 blocks of size 3 with 7 treatments, we construct a Latin square with 7 treatments (a Latin square is a square in which each treatment appears only once in each row and column). Then, we take 3 columns (not any 3 are valid) and construct the different blocks. These rectangles are called Youden squares.

A	B	C	D	E	F	G
B	C	D	E	F	G	A
C	D	E	F	G	A	B
D	E	F	G	A	B	C
E	F	G	A	B	C	D
F	G	A	B	C	D	E
G	A	B	C	D	E	F

A	B	D
B	C	E
C	D	F
D	E	G
E	F	A
F	G	B
G	A	C

Balanced Incomplete Block Designs (Lattice design)

If the number of treatments is large (and a perfect square, i.e., $v = x^2$), then we may use a different kind of designs also based on Latin squares. This design assumes that the experiment will be replicated several times. At each replication the block composition changes and different treatments are used in the same block.



Balanced Incomplete Block Designs (Many treatments)

Although outside of the scope of this course, for a large number of treatments (a few hundreds), the interested reader may look for

- Cubic lattice designs
- Alpha lattice designs for large-scale variety trials

Outline

Multiple blocking

Example 33

We are interested in the pattern of variation over time of a constituent of blood (=treatment). We need sampling blood from 9 chickens (=replication, 1st blocking variable) on 25 weekly occasions. Only 6 samples can be analyzed at a time, and there can be a substantial difference between batches of samples (=2nd blocking variable).

Latin squares

Example 34

We are interested in the wearing performance of 4 tyre brands (=treatment). There can be differences depending on the car (=1st blocking variable) and the position within the car (=2nd blocking variable). The organization of this experiment can be done through a Latin square design.

Position \ Car	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

	df
Treatments	$(t-1)=3$
Blocking 1	$(t-1)=3$
Blocking 2	$(t-1)=3$
Errors	$(t-1)(t-2)=6$
All	$t^2-1=15$

Since the number of degrees of freedom for the error is relatively low for a Latin square design, the experiment must be replicated several times with independent latin squares.

Latin squares

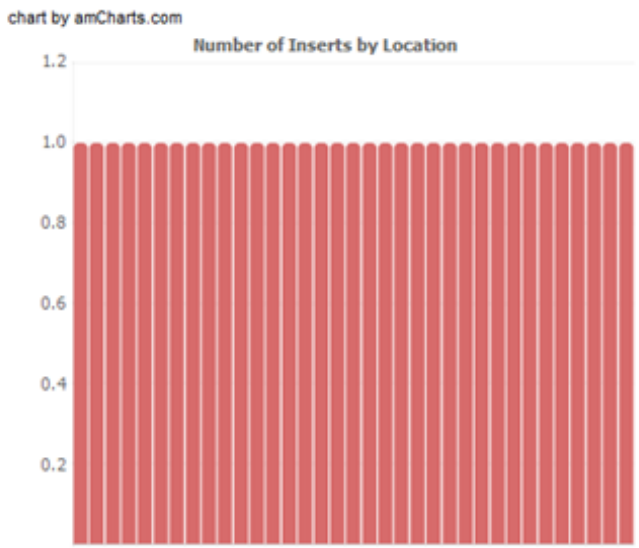
The Latin square model for the l replicate is

$$y_{ijkl} = \mu + r_{l(i)} + c_{l(j)} + t_{k(ijl)} + \epsilon_{ijkl}$$

Google: Latin square generator
(http://hamsterandwheel.com/grids/index2d.php)

Finished in 0.00179 seconds with 36 inserts attempted, 0 of which had to be replaced.
Square Size (2-15): 6 Generate (Will bail out after 10000 attempted inserts, successful or otherwise.)
Memory allocation - current:1024Kb - peak:1024Kb
Memory usage - current:949Kb - peak:972Kb

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



Latin squares

Replicates may share one of the blocking variables (Latin rectangle) ...

Example 35

For instance, the two replicates may be performed on the same cars

Position \ Car	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C
5	B	D	A	C
6	A	B	C	D
7	D	C	B	A
8	C	A	D	B

Latin squares

... or not

Example 36

For instance, the two replicates may be performed on different cars

Position \ Car	1	2	3	4	5	6	7	8
1	A	B	C	D				
2	B	C	D	A				
3	C	D	A	B				
4	D	A	B	C				
5					B	D	A	C
6					A	B	C	D
7					D	C	B	A
8					C	A	D	B

Sequences of experiments: Orthogonal Latin Squares

Sometimes we need to reuse the experimental units from one experiment to the next (fruit trees, agricultural plots, patients in a clinical trial, ...) The solution is to use two orthogonal latin squares (if both designs are superimposed each pair appears only once; not all Latin squares have orthogonal squares). Examples of these orthogonal designs are Graeco-Latin squares. The idea is to eliminate the long-term effects of the first experiment on the second experiment.

Example 37

Experiment 1: A, B, C, D

Experiment 2: $\alpha, \beta, \gamma, \delta$

Block1 \ Block2	1	2	3	4
1	Aα	Dδ	Bγ	Cβ
2	Cδ	Bα	Dβ	Aγ
3	Dγ	Aβ	Cα	Bδ
4	Bβ	Cγ	Aδ	Dα

Non-orthogonal row-and-column designs

Latin squares can successfully block two variables whose number of levels is equal between them and equal to the number of treatments (c=4 cars, r=4 positions and t=4 tyre brands). Row-and-column designs address those cases with different number of levels in each one of the blocks.

If the number of rows, r, or columns, c, is equal to t, we may use Youden squares

Example 38

A	B	C	D	E	F	G
B	C	D	E	F	G	A
C	D	E	F	G	A	B
D	E	F	G	A	B	C
E	F	G	A	B	C	D
F	G	A	B	C	D	E
G	A	B	C	D	E	F

A	B	D
B	C	E
C	D	F
D	E	G
E	F	A
F	G	B
G	A	C

Non-orthogonal row-and-column designs

However, more complicated patterns may appear: the number of rows and columns is not a multiple of the number of treatments or some combinations of blocks are unfeasible (some plots in a field are useless).

In the design of row-and-column designs:

1. The goal should be to achieve orthogonality in each one of the blocking variables.
2. Balance, if ortohogonality is not possible.
3. if balance is not possible, then the joint occurrences of treatments in rows and columns should be made as equal as possible.

Non-orthogonal row-and-column designs

Example 39

$t = 9$ treatments in $r = 5$ rows and $c = 7$ columns

Block1 \ Block2	1	2	3	4	5	6	7
1	I	B	A	G	C	D	H
2	A	F	G	E	B	C	D
3	G	I	B	H	F	E	A
4	C	H	I	D	E	G	F
5	E	D	F	A	H	B	C

Example 40

$t = 6$ treatments in $r = 6$ rows and $c = 8$ columns with useless cells

Block1 \ Block2	1	2	3	4	5	6	7	8
1	C	E	B	F	A		E	D
2	E	A	D		F	C	B	
3	F		E	A	D	B		C
4	D	C	A	B	E	F		
5	B	D	F	E		A	C	F
6	A	B		C			D	

Blocking time: Cross-over designs

The different treatments are applied in sequence to the same experimental unit (common in clinical trials).

- The aim to use the same subject is because by applying the more than 1 treatment to the same subject we remove inter-subject variability, gain in statistical precision, and reduce the number of subjects.
- However, the design assumes that there is no effect from one period to the next (washout period between treatments; what if the subject is cured by the first treatment?).

The experimental unit is redefined to be an observation for an individual subject (=1st blocking variable) in a short period of time (=2nd blocking variable).

Example 41

Period \ Subject	1	2	3	4	5	6	7	8	9	10	11	12
1	A	B	B	A	A	B	A	B	B	B	A	A
2	B	A	A	B	B	A	B	A	A	A	B	B

By having the same number of subjects for the two orderings (AB or BA), we remove the effects of treatment order.

But, the experiment (swapping treatment) is rather different from its future use (one treatment). And many objections should be expected

Blocking time: Cross-over designs

The different treatments are applied in sequence to the same experimental unit (common in clinical trials).

- The aim to use the same subject is because by applying the more than 1 treatment to the same subject we remove inter-subject variability, gain in statistical precision, and reduce the number of subjects.
- However, the design assumes that there is no effect from one period to the next (washout period between treatments; what if the subject is cured by the first treatment?).

The experimental unit is redefined to be an observation for an individual subject (=1st blocking variable) in a short period of time (=2nd blocking variable).

Example 42

Period \ Subject	1	2	3	4	5	6	7	8	9	10	11	12
1	A	B	B	A	A	B	A	B	B	B	A	A
2	B	A	A	B	B	A	B	A	A	A	B	B

By having the same number of subjects for the two orderings (AB or BA), we remove the effects of treatment order.

Blocking time: Cross-over designs

Example 43

The number of periods and treatments do not need to be the same and different sequences can be applied (e.g. ABB or BAA)

Period \ Subject	1	2	3	4	5	6	7	8	9	10	11	12
1	A	B	B	A	A	B	A	B	B	B	A	A
2	B	A	A	B	B	A	B	A	A	A	B	B
3	B	A	A	B	B	A	B	A	A	A	B	B

or ABB, BAA, ABA and BAB

Period \ Subject	1	2	3	4	5	6	7	8	9	10	11	12
1	A	B	A	B	B	B	A	A	B	B	B	A
2	B	A	B	A	A	A	B	B	A	A	A	B
3	B	A	A	B	B	A	A	B	B	A	B	B

Blocking time: Cross-over designs

But, the experiment (swapping treatment) is rather different from its future use (one treatment). And many objections should be expected

- Order effects: it is possible that the order in which treatments are administered may affect the outcome. An example might be a drug with many adverse effects given first, making patients taking a second, less harmful medicine, more sensitive to any adverse effect.
- Carry-over effects: can be avoided with sufficiently long washout periods and designs to eliminate 1st order, 2nd order, ... carryover effects.
- Learning effects: this is important where you have controls who are naive to the intended therapy. In such a case e.g. you cannot make a group (typically the group which learned the skill first) unlearn a skill such as yoga and then act as a control in the second phase of the study.

Blocking time: Cross-over designs

- If there are no carry-over effects, these designs are like row-and-column designs.
- If there are carry-over effects, the design has to be performed to remove 1st order (only from the treatment in the previous period) or higher order (from treatments in the two, three, ... previous periods) carry-over effects.

A design is

- uniform within sequences if each treatment appears the same number of times within each sequence (e.g. AB/BA is uniform in sequence, but ABA/BAB is not)
- uniform within periods if each treatment appears the same number of times within each period (this depends on the number of subjects assigned to each sequence)

Blocking time: Cross-over designs

A design is balanced with respect to 1st order carry-over effects if each treatment precedes any other treatment the same number of times. Latin squares (although not all of them) help to construct balanced designs.

Example 44

With $t = 4$ treatments

Sequence \ Period	1	2	3	4
ABCD	A	B	C	D
BCDA	B	C	D	A
CDAB	C	D	A	B
DABC	D	A	B	C

This design is not balanced
(A precedes B 3 times,
but does not precede C or D)

Sequence \ Period	1	2	3	4
ABCD	A	B	C	D
BDAC	B	D	A	C
CADB	C	A	D	B
DCBA	D	C	B	A

This design is balanced
(all pairs appear only once)

Blocking time: Cross-over designs

- If the number of treatments is even (e.g. $t = 4$), only 1 Latin square is needed to produce a balanced design.
- If the number of treatments is odd (e.g. $t = 3$), 2 Latin squares are needed to produce a balanced design.

Example 45

With $t = 3$ treatments

Sequence \ Period	1	2	3
ABC	A	B	C
BCA	B	C	A
CAB	C	A	B
ACB	A	C	B
BAC	B	A	C
CBA	C	B	A

This design is balanced
(all pairs appear 2 times)

Blocking time: Cross-over designs

A design is strongly balanced with respect to 1st order carry-over effects if each treatment precedes every other treatment (including itself) the same number of times.

Example 46

With $t = 4$ treatments

Sequence \ Period	1	2	3
ABB	A	B	B
BAA	B	A	A

This design is strongly balanced but it is not uniform within sequences.

Sequence \ Period	1	2	3	4
ABBA	A	B	B	A
BAAB	B	A	A	B
AABB	A	A	B	B
BBAA	B	B	A	A

This design is strongly balanced and uniform within sequences.

Blocking time: Cross-over designs

Let us analyze an example with carry-over effects

Sequence \ Period	1	2
AB	A	B
BA	B	A

The expected values at each one of the cells are

Sequence \ Period	1	2
AB	$\mu + t_A + o_{AB} + p_1$	$\mu + t_B + o_{AB} + p_2 + \lambda_A$
BA	$\mu + t_B + o_{BA} + p_1$	$\mu + t_A + o_{BA} + p_2 + \lambda_B$

being o_{AB} and o_{BA} the effect size corresponding to the ordering block, p_1 and p_2 the effect size corresponding to the period block and λ_A and λ_B the carry-over effects for having applied first A (λ_A) or B (λ_B).

Blocking time: Cross-over designs

Sequence \ Period	1	2
AB	$\mu + t_A + o_{AB} + p_1$	$\mu + t_B + o_{AB} + p_2 + \lambda_A$
BA	$\mu + t_B + o_{BA} + p_1$	$\mu + t_A + o_{BA} + p_2 + \lambda_B$

The mean estimate for each one of the treatments is normally performed by averaging the cells receiving that treatment

$$\begin{aligned}
\hat{y}_A &= \frac{y_{AB,1}+y_{BA,2}}{2} \\
&= \frac{(\mu+t_A+o_{AB}+p_1)+(\mu+t_A+o_{BA}+p_2+\lambda_B)}{2} \\
&= \mu + t_A + \frac{\lambda_B}{2} \\
\hat{y}_B &= \frac{y_{AB,2}+y_{BA,1}}{2} \\
&= \frac{(\mu+t_B+o_{AB}+p_2+\lambda_A)+(\mu+t_A+o_{BA}+p_2+\lambda_B)}{2} \\
&= \mu + t_B + \frac{\lambda_A}{2}
\end{aligned}$$

Treatments are aliased with the carry-over effects.

Blocking time: Cross-over designs

Let us repeat it with a strongly balaced design

Sequence \ Period	1	2	3
ABB	$\mu + t_A + o_{ABB} + p_1$	$\mu + t_B + o_{ABB} + p_2 + \lambda_A$	$\mu + t_B + o_{ABB} + p_3 + \lambda_B$
BAA	$\mu + t_B + o_{BAA} + p_1$	$\mu + t_A + o_{BAA} + p_2 + \lambda_B$	$\mu + t_A + o_{BAA} + p_3 + \lambda_A$

The mean estimate for each one of the treatments is normally performed by averaging the cells receiving that treatment

$$\begin{aligned}
\hat{y}_A &= \frac{y_{ABB,1}+y_{BAA,2}+y_{BAA,3}}{3} \\
&= \frac{(\mu+t_A+o_{ABB}+p_1)+(\mu+t_A+o_{BAA}+p_2+\lambda_B)+(\mu+t_A+o_{BAA}+p_3+\lambda_A)}{3} \\
&= \mu + t_A + \frac{o_{BAA}}{3} \\
\hat{y}_B &= \frac{y_{BAA,1}+y_{ABB,2}+y_{ABB,3}}{3} \\
&= \frac{(\mu+t_B+o_{BAA}+p_1)+(\mu+t_B+o_{ABB}+p_2+\lambda_A)+(\mu+t_B+o_{ABB}+p_3+\lambda_B)}{3} \\
&= \mu + t_B + \frac{o_{ABB}}{3}
\end{aligned}$$

Treatments are aliased with the treatment order effects.

Outline

Split-unit designs

We have an experiment with two factors. One of them requires large experimental units, while the other one small ones. Additionally, the second factor can be applied to a "small portions" of the experimental units of the first factor.

Example 47

We are investigating the effect of light and diet on the growth of mice.

- The experimental unit for the light factor is the whole room, all cages receive the same treatment (number of light hours)
- The experimental unit for the diet is the cage, all mice in the same cage receive the same treatment.

These designs are called split-unit designs

Split-unit designs

Let us call P the factor applied to large units and Q the factor applied to small units. Assume that a large unit receives the treatment p_j and it is allocated to the i-th block.

From the point of view of the large unit the observations should respond to the model

$$z_{ij} = \mu + b_i + p_j + \eta_{ij}$$

Assume that a small unit receives the treatment q_k.
From the point of view of the small unit

$$y_{ijk} = \mu + m_{ij} + q_k + (pq)_{jk} + \epsilon_{ijk}$$

where m_{ij} contains the main effects of the blocks and the P treatments and their interactions.

Split-unit designs

Both models can be integrated in a single model

$$y_{ijk} = \mu + b_i + p_j + \epsilon'_{ij} + q_k + (pq)_{jk} + \epsilon_{ijk}$$

with

$$\begin{aligned} z_{ij} &= y_{ij\cdot} \\ \eta_{ij} &= \epsilon'_{ij} + \epsilon_{ij\cdot} \\ m_{ij} &= b_i + p_j + \epsilon'_{ij} \end{aligned}$$

Split-unit designs

Example 48

We are investigating the effect of 5 irrigation systems (large unit factor, P), and 3 rice variants.

Block I

p_2q_1	p_5q_1	p_3q_2	p_1q_3	p_4q_1
p_2q_3	p_5q_3	p_3q_1	p_1q_1	p_4q_3
p_2q_2	p_5q_2	p_3q_3	p_1q_2	p_4q_2

Block II

p_3q_2	p_4q_3	p_2q_2	p_5q_1	p_1q_2
p_3q_3	p_4q_2	p_2q_1	p_5q_3	p_1q_1
p_3q_1	p_4q_1	p_2q_3	p_5q_2	p_1q_3

Block III

p_5q_1	p_1q_3	p_3q_3	p_2q_2	p_4q_3
p_5q_2	p_1q_2	p_3q_1	p_2q_3	p_4q_2
p_5q_3	p_1q_1	p_3q_2	p_2q_1	p_4q_1

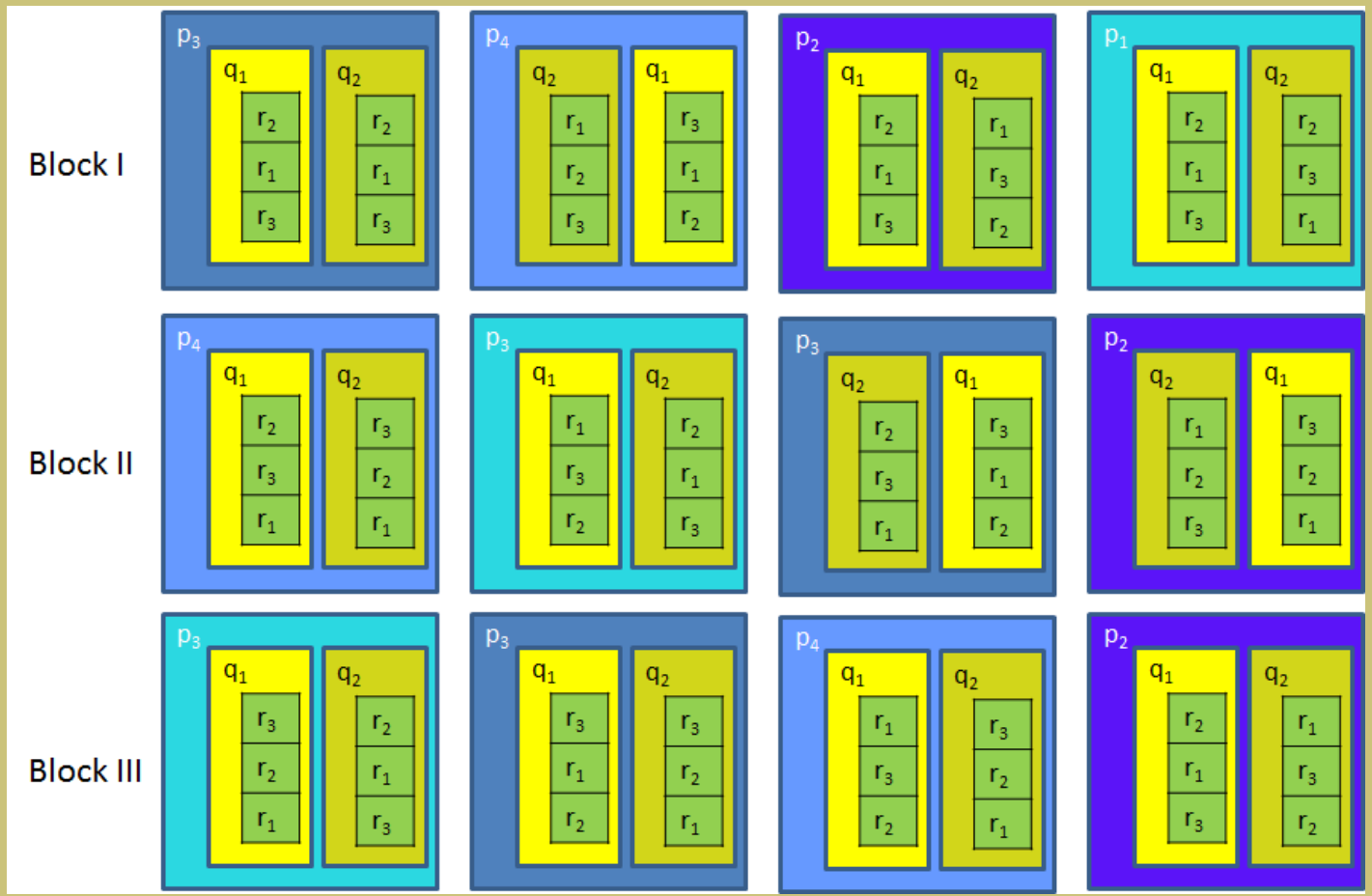
Block IV

p_2q_3	p_4q_1	p_1q_2	p_3q_1	p_5q_2
p_2q_1	p_4q_3	p_1q_1	p_3q_2	p_5q_3
p_2q_2	p_4q_2	p_1q_3	p_3q_3	p_5q_1

Split-split-unit designs

We may nest several variables requiring increasingly small experimental units like the design in the example below.

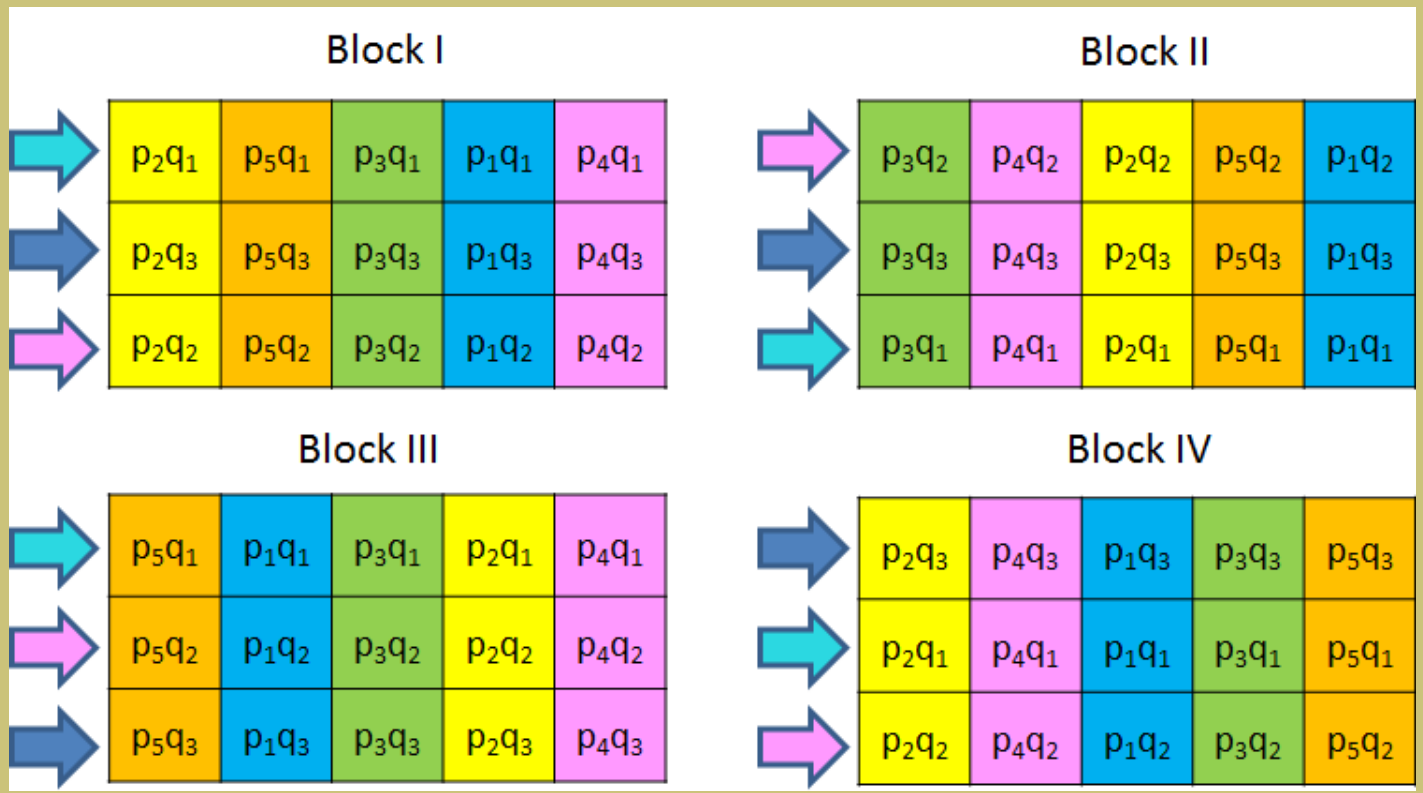
Example 49



Criss-cross designs

If both treatments require relatively large experimental units, we may apply one of them on the columns, and another one on the rows.

Example 50



Outline

Randomization

Example 51

8 students are divided in two groups, each student has the same probability of being in each group. Group 1 (A,B,C,D) receives a special training program, while Group 2 (E,F,G,H) receives the standard training. The scores in a test are

F	A	C	H	E	B	G	D
19	16	15	15	14	13	12	10

The probability of 4 students being at the top of the list by chance is p-val=$1/70$ ($= 1/C(8,4)$). The fact that we have randomly assigned students to both groups make the results "generalizable" to the whole population.

Randomization

Example 52

We give a test to 8 students. The scores are

F	A	C	H	E	B	G	D
19	16	15	15	14	13	12	10

We observe that the 4 top scores (F,A,C,H) correspond to females, while the 4 bottom scores (E,B,G,D) correspond to males. The fact that we have a *post-hoc* observation makes the result less reliable (case-studies). We might have found any other spurious pattern (the 4 older people, the 4 blond people, the 4 people born from Aries to Virgo, ...)

Randomization

- If all units are known at the beginning of the experiment, randomization can be performed simply by a random permutation (performed by a computer, not by a person)

F	A	C	H	E	B	G	D

- If units arrive sequentially, we may assign randomly the treatment depending on the number of already assigned units. We assign to Group 1 with probability

$$p_1 = \frac{4 - g_1}{8 - g_1 - g_2}$$

where we will have in total 8 units, 4 assigned to Group 1 and 4 to Group 2, and g_1 and g_2 are the number of units assigned until this moment to each one of the groups.

Randomization

- If units arrive sequentially, more complex schemes may be followed. The probability of being assigned to A, which initially is 0.5, is modified by the number of subjects in treatments A and B for each one of the characteristics.

Example 53

For instance, we are conducting a clinical trial with 2 treatments (A and B) in which we classify patients by age, sex and occupation. Assume that a new patient arrives with an age of 28 years, male, and occupation IV, and that the previous patients have been allocated as

	A	B	
Age			
<30	10	6	$\rightarrow 6/10$
30-50	12	12	
50-70	4	5	
>70	4	7	
Sex			
Male	17	14	$\rightarrow 14/17$
Female	13	16	
Occupation			
I	5	8	
II	9	13	
III	7	2	
IV	9	7	$\rightarrow 7/9$

The probability of being assigned to treatment A is

$$p_A = 0.5\frac{6}{10}\frac{14}{17}\frac{7}{9} = 0.19$$

Randomization

In clinical studies it is important that the patient (=single blind) AND the doctor (=double blind) do not know (cannot guess) which is the treatment being applied, because this may bias the results (doctors/patients tend to evaluate differently if they know that they have been given Treatment 1 instead of Treatment 2). Doctors should not be able to distinguish which patient is receiving which treatment.

For small number of treatments, blocks should contain more than 1 replicate of each treatment.

Outline

Chapter 3. Factorial designs

C.O.S. Sorzano
coss@cnb.csic.es

National Center of Biotechnology (CSIC)

October 14, 2016

Outline

Outline

Factorial design

Example 50

We want to find the optimal combination of number of meals and daily calories for a weight loss diet. We are thinking of 3 or 5 meals, and 1500, 1700 or 1900 calories. 5 individuals will be given all possible combinations:

	Meals	Calories
Treatment 1	3	1500
Treatment 2	3	1700
Treatment 3	3	1900
Treatment 4	5	1500
Treatment 5	5	1700
Treatment 6	5	1900

Factorial design

Factor	Set of treatments of a single type (e.g., No. Meals or Calories)
Level of a factor	Particular treatment from the set (e.g., 1500, 1700, 1900)
Experimental treatment	A combination of one level from each factor (e.g., 3 meals, 1500 calories)
Main effect	Comparison between levels of a single factor (e.g. 3 vs 5 meals)
Interaction	Comparison between levels of several factors (e.g. 3 meals and 1500 calories vs 3 meals and 1700 calories)

Factorial designs ...

- test the main effects of each factor with a variety of other levels improving the relevance of the study.
- allow estimating interactions between factors.
- reduce the number of samples with respect to the *change only one variable at a time*.

Incomplete factorial design

Designs must not be full factorial if some combinations of treatments make no sense. If we remove some of the combinations we increasingly loose orthogonality, but the loss may compensate for not performing nonsensical combinations

Example 51

	Factor A	Factor B	Factor C
~~Treatment 0~~	~~no~~	~~no~~	~~no~~
Treatment 1	no	no	yes
Treatment 2	no	yes	no
Treatment 3	no	yes	yes
Treatment 4	yes	no	no
Treatment 5	yes	no	yes
Treatment 6	yes	yes	no
~~Treatment 7~~	~~yes~~	~~yes~~	~~yes~~

Factorial design + Control

There can be designs similar to factorials but with an extra replicate control. The way we perform the control has effects on the analysis.

Example 52

Given $N_A = 3$ levels of Factor A and $N_B = 3$ of Factor B, we have two possible designs with an extra control.

Design I: $N_{replicates} = 6$, $(N_A N_B + Control)N_{replicates} = (3 \times 3 + 1) \times 6 = 60$ units
Design II: $N_{replicates} = 5$, $N_A(N_B + Control)N_{replicates} = 3 \times (3 + 1) \times 5 = 60$ units

	Design I	Design II
Replicates Control	6	15
Replicates A	18	15
Replicates AB	6	5
Variance Control-AB	$\frac{2\sigma^2}{6} = 0.33\sigma^2$	$\frac{\sigma^2}{15} + \frac{\sigma^2}{5} = 0.27\sigma^2$
Variance Control-A	$\frac{\sigma^2}{6} + \frac{\sigma^2}{18} = 0.22\sigma^2$	$\frac{2\sigma^2}{15} = 0.13\sigma^2$
Variance A_1-A_2	$\frac{2\sigma^2}{18} = 0.11\sigma^2$	$\frac{2\sigma^2}{15} = 0.13\sigma^2$
Variance AB_1-AB_2	$\frac{2\sigma^2}{6} = 0.33\sigma^2$	$\frac{2\sigma^2}{5} = 0.4\sigma^2$

Complicated "factorial" designs

Example 53

We want to compare two chemicals (E and O) versus a control. Chemical O is an oil requiring a surfactant (S_1 or S_2). The application can be performed with two different sprayers (SP_1 or SP_2). Three concentrations of the chemicals will be used. The anticipated comparisons will be: O vs E; O or E vs Control; Main effect of the surfactants; Main effect of the sprayers; Conc_1 vs Conc_2 vs Conc_3. The number of replicates could be

Chemical	Surfactant	Sprayer	Conc_1	Conc_2	Conc_3
O	S_1	SP_1	$\times 1$	$\times 1$	$\times 1$
O	S_1	SP_2	$\times 1$	$\times 1$	$\times 1$
O	S_2	SP_1	$\times 1$	$\times 1$	$\times 1$
O	S_2	SP_2	$\times 1$	$\times 1$	$\times 1$
E		SP_1	$\times 2$	$\times 2$	$\times 2$
E		SP_2	$\times 2$	$\times 2$	$\times 2$
Control		SP_1			$\times 5$
Control		SP_2			$\times 5$

Outline

2^k Two-level factorial designs

When each factor has only two levels, the design is called two-level k factorial design (2^k).

Example 54

We want to determine the effect of a mammalian water balance hormone in amphibia. Two species (=Factor P) are studied (toads=0, frogs=1), at two levels of hormone (=Factor Q, control=0, hormone=1), and two pre-experiment moisture conditions (=Factor R, wet=0, dry=1). We measure the percentage increase in weight after immersion in water for 2h. 2 replicates are studied for each treatment combination. The results are

Species(P)	Moisture(Q)	Hormone(R)	Results	
Toad	wet	control	2.31	-1.59
Toad	wet	hormone	28.37	14.16
Toad	dry	control	17.68	25.23
Toad	dry	hormone	28.39	27.94
Frog	wet	control	0.85	2.90
Frog	wet	hormone	3.82	2.86
Frog	dry	control	2.47	17.72
Frog	dry	hormone	13.71	7.38

2^k Two-level factorial designs

Let us call

- (1) the mean of the observations with no treatment applied ($P = Q = R = -1$)
- P the effect size of applying P=1
- p the mean of the observations that has P=1
- PQ the effect size of applying P=1 and Q=1
- pq the mean of the observations that have P=1 and Q=1
- ...

We may estimate the effect size of P as the difference between those observations with $P = 1$ and $P = -1$. But this can be done in many different ways

$$\begin{aligned}
\hat{P} &= p - (1) \\
&= \frac{p+pq}{2} - \frac{1+q}{2} \\
&= \frac{p+pq+pr}{3} - \frac{1+q+r}{3} \\
&= \frac{p+pq+pr+pqr}{4} - \frac{1+q+r+qr}{4} = \frac{1}{4}(p-1)(q+1)(r+1) \\
&= \frac{pq+pr+pqr}{3} - \frac{q+r+qr}{3} \\
&= \ldots
\end{aligned}$$

2^k Two-level factorial designs

The following table shows how to choose the signs to estimate the different effects. It is contructed by setting the signs for P, Q and R. Then the rest of columns (PQ, PR, QR, PQR) are simply the multiplication of the corresponding signs.

Estimate	μ	P	Q	PQ	R	PR	QR	PQR
(1)	+	-	-	+	-	+	+	-
p	+	+	-	-	-	-	+	+
q	+	-	+	-	-	+	-	+
pq	+	+	+	+	-	-	-	-
r	+	-	-	+	+	-	-	+
pr	+	+	-	-	+	+	-	-
qr	+	-	+	-	+	-	+	-
pqr	+	+	+	+	+	+	+	+

For instance to estimate PQR we would have

$$\widehat{PQR} = \frac{1}{4}\left(-1 + p + q - pq + r - pr - qr + pqr\right)$$

2^k Two-level factorial designs

Similarly

$$\begin{aligned}
\hat{P} &= \tfrac{1}{4}(p-1)(q+1)(r+1) = \tfrac{1}{2^{k-1}}(p-1)(q+1)(r+1) \\
\hat{Q} &= \tfrac{1}{4}(p+1)(q-1)(r+1) \\
\hat{R} &= \tfrac{1}{4}(p+1)(q+1)(r-1) \\
\widehat{PQ} &= \tfrac{1}{4}(p-1)(q-1)(r+1) \\
\widehat{QR} &= \tfrac{1}{4}(p+1)(q-1)(r-1) \\
\widehat{PR} &= \tfrac{1}{4}(p-1)(q+1)(r-1) \\
\widehat{PQR} &= \tfrac{1}{4}(p-1)(q-1)(r-1) \\
\hat{\mu} &= \tfrac{1}{8}(p+1)(q+1)(r+1) = \tfrac{pqr+pq+pr+qr+p+q+r+1}{2^k}
\end{aligned}$$

General formulas for k factor and r replicates for each treatment

$$\begin{aligned}
\widehat{effect} &= \tfrac{1}{2^{k-1}}(p_1 \pm 1)(p_2 \pm 1)...(p_k \pm 1) \\
\mathrm{Var}\{effect\} &= \tfrac{\sigma^2}{r2^{k-2}} \\
\mathrm{SS}\{effect\} &= r2^{k-2}(effect)^2
\end{aligned}$$

2^k Two-level factorial designs

In matrix form

$$\begin{pmatrix} 2^k\hat{\mu} \\ 2^{k-1}\hat{P} \\ 2^{k-1}\hat{Q} \\ 2^{k-1}\widehat{PQ} \\ 2^{k-1}\hat{R} \\ 2^{k-1}\widehat{PR} \\ 2^{k-1}\widehat{QR} \\ 2^{k-1}\widehat{PQR} \end{pmatrix} = \begin{pmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \\ -1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 & -1 & -1 & +1 & +1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ -1 & -1 & -1 & -1 & +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ -1 & +1 & +1 & -1 & +1 & -1 & -1 & +1 \end{pmatrix} \begin{pmatrix} 1 \\ p \\ q \\ pq \\ r \\ pr \\ qr \\ pqr \end{pmatrix}$$

or equivalently

$$\hat{\mathbf{y}} = U\mathbf{x}$$

U is an orthogonal matrix (the rows and columns of U are orthogonal to each other, $U^{-1} = \frac{1}{\det U} U^T$), so

$$\mathbf{x} = \frac{1}{\det U} U^T \hat{\mathbf{y}}$$

2^k Two-level factorial designs

The data generation model comes from this latter equation $\mathbf{x} = \frac{1}{\det U} U^T \hat{\mathbf{y}}$ and it can be expressed as

$$\begin{aligned} x_{ijkl} = \; & \mu \\ & +0.5\left((-1)^{i-1}P + (-1)^{j-1}Q + (-1)^{k-1}R\right) \\ & +0.5\left((-1)^{i+j-2}PQ + (-1)^{i+k-2}PR + (-1)^{j+k-2}QR\right) \\ & +0.5\left((-1)^{i+j+k-3}PQR\right) \\ & +\epsilon_{ijkl} \end{aligned}$$

Example 55

$$\begin{aligned} x_{toad,dry,control} = x_{010} = \; & \mu \\ & +0.5(-P + Q - R) \\ & +0.5(-PQ - PR - QR) \\ & +0.5(+PQR) \end{aligned}$$

$$\begin{aligned} x_{toad,dry,hormone} = x_{011} = \; & \mu \\ & +0.5(-P + Q + R) \\ & +0.5(-PQ + PR + QR) \\ & +0.5(-PQR) \end{aligned}$$

Example 56

We are interested in testing if there is a difference due to the hormone in toads with a dry period before getting immersed in water

$$\begin{aligned} c = x_{toad,dry,hormone} - x_{toad,dry,control} &= x_{011} - x_{010} \\ &= R - PR + QR - PQR \\ &= effect_{hormone} \\ &\quad - effect_{(toad,hormone)} \\ &\quad + effect_{(dry,hormone)} \\ &\quad - effect_{(toad,dry,hormone)} \end{aligned}$$

Each effect in c has a variance $\frac{\sigma^2}{r2^{k-2}} = \frac{\sigma^2}{2 \cdot 2^{3-2}} = \frac{\sigma^2}{4}$. So the variance of c depending on the model is

Full model	$4\frac{\sigma^2}{4}$
No 3rd order interactions ($PQR = 0$)	$3\frac{\sigma^2}{4}$
No 2nd order interactions ($PQ = PR = QR = PQR = 0$)	$\frac{\sigma^2}{4}$

Example 56(continued)

The model we choose has consequences in the analysis results

Model	Estimate of $x_{011} - x_{010}$	Std.Error of estimate	95% Confidence interval
Full model	6.7	5.88	$(-6.28, 20.28)$
No 3rd order	10	5.09	$(-1.76, 21.76)$
No 2nd order	7.38	2.94	$(0.59, 14.17)$

Factorial designs allow estimating many interactions. But **the simpler the model, the better**. The choice to remove interactions must be done before the experiment, never after seeing the experiment results.

$3^k, 4^k, \ldots$ High-level factorial designs

High-level factorial designs are possible, but the analysis gets more and more complicated. For example for the 3^2-factorial design

$$
\begin{aligned}
\hat{P}' &= \tfrac{1}{3}(p_2 - p_0)(q_2 + q_1 + q_0) \\
\hat{P}'' &= \tfrac{1}{6}(p_2 - 2p_1 + p_0)(q_2 + q_1 + q_0) \\
\hat{Q}' &= \tfrac{1}{3}(p_2 + p_1 + p_0)(q_2 - q_0) \\
\hat{Q}'' &= \tfrac{1}{6}(p_2 + p_1 + p_0)(q_2 - 2q_1 + q_0) \\
\hat{P}'\hat{Q}' &= \tfrac{1}{2}(p_2 - p_0)(q_2 - q_0) \\
\hat{P}'\hat{Q}' &= \tfrac{1}{4}(p_2 - p_0)(q_2 - 2q_1 + q_0) \\
\hat{P}''\hat{Q}' &= \tfrac{1}{4}(p_2 - 2p_1 + p_0)(q_2 - q_0) \\
\hat{P}''\hat{Q}' &= \tfrac{1}{8}(p_2 - 2p_1 + p_0)(q_2 - 2q_1 + q_0) \\
\hat{\mu} &= \tfrac{1}{9}(p_2 + p_1 + p_0)(q_2 + q_1 + q_0)
\end{aligned}
$$

Replication of factorial designs

Example 57

We may replicate a factorial design by simply repeating the sequence of experiments. However, repeating in the same order is not a good idea, randomisation is better (to avoid the influence of the order of treatments). For example, for a 2^3-factorial design we may perform:

Design run	Treatment	Experimental run
0	000	7
1	001	2
2	010	15
3	011	10
4	100	1
5	101	3
6	110	5
7	111	13
8	000	9
9	001	8
10	010	14
11	011	0
12	100	6
13	101	12
14	110	4
15	111	11

Factorial designs and single replicates

High-order interactions can be assimilated to the error, and single replicate factorial designs may be conceived.

Example 58

We are interested in the survival of *Salmonella typhimurium* under 3 experimental factors: 3 levels of sorbic acid (=Factor S), 6 levels of water activity (=Factor A), and 3 levels of pH (=Factor P). The data will be the log (density/ml) measured after 7 days after treatment started.

We have $3 \times 6 \times 3 = 54$ treatments, and we will use a single replicate for each treatment.

Example 58(continued)

The data analysis table would be

	SS	df	MS	F
Water activity (A)	81.57	5=(6-1)	16.31	$473 > F_{0.95,5,20}$
Sorbic acid (S)	2.76	2=(3-1)	1.38	$40 > F_{0.95,5,20}$
pH (P)	0.01	2=(3-1)	0.01	$0.2 < F_{0.95,2,20}$
AS	1.32	10=(6-1)(3-1)	0.13	$3.8 > F_{0.95,10,20}$
AP	0.45	10=(6-1)(3-1)	0.04	$1.3 < F_{0.95,10,20}$
SP	0.23	4=(3-1)(3-1)	0.06	$1.7 < F_{0.95,4,20}$
$ASP \approx$ **Error**	0.69	20=(6-1)(3-1)(3-1)	0.03	
Total	87.03	53		

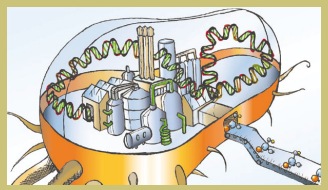
The problem with single replicate, factorial designs is that 1) it is difficult to use blocking, 2) due to the lack of replication, there is no possibility to construct an unbiased estimate of the noise.

Outline

Fractional replication

Example 59

We are interested in a cell line as biologics bioreactor, and we want to optimize production. We have identified 7 variables we may control (temperature, humidity, pH, O_2 concentration, CO_2 concentration, glucose concentration, aminoacid concentration). For each variable we have 2 possible values. There are $2^7 = 128$ possible treatments, but we can only afford 64. We do not foresee 3rd order interactions or higher. Can we perform this experiment?



The number of degrees of freedom needed to identify main effects and 2nd order interactions is

	df
Main effects	7
2nd Order Interactions	21=C(7,2)=7!/(2!5!)

So we need 28 samples plus sufficient replication for estimating the error. For instance, if we perform 64 experiments, there would be 37 df for the noise.

2^{k-1} Factorial design

Let us perform $1/2$ of a full 2^k factorial design. We need to find an appropriate subset and understand its consequences. Let us consider the full factorial design with 3 factors:

Treatment	P	Q	R	PQ	PR	QR	PQR
000	-	-	-	+	+	+	-
001	-	-	+	+	-	-	+
010	-	+	-	-	+	-	+
011	-	+	+	-	-	+	-
100	+	-	-	-	-	+	+
101	+	-	+	-	+	-	-
110	+	+	-	+	-	-	-
111	+	+	+	+	+	+	+

This matrix also defines how to estimate the different contributions. For instance

$$\begin{aligned} \hat{P} &= \tfrac{1}{4}(-y_{000} - y_{001} - y_{010} - y_{011} + y_{100} + y_{101} + y_{110} + y_{111}) \\ \widehat{PQR} &= \tfrac{1}{4}(-y_{000} + y_{001} + y_{010} - y_{011} + y_{100} - y_{101} - y_{110} + y_{111}) \end{aligned}$$

2^{k-1} Factorial design

We now choose 4 (one half) treatments that preserve column orthogonality amongst the treatments

Treatment	P	Q	R	PQ	PR	QR	PQR
001	-	-	+	+	-	-	+
010	-	+	-	-	+	-	+
100	+	-	-	-	-	+	+
111	+	+	+	+	+	+	+

Actually, the column for P is the same as the one for QR, meaning that when we compute

$$\widehat{P + QR} = -y_{001} - y_{010} + y_{100} + y_{111}$$

we are confounding P with QR, we cannot distinguish between the effect of boths, but we presume that the main effect of P is larger than the QR interaction. This is also called aliasing.
In this design there are other aliasings (Q with PR, R with PQ, and the mean (1) with PQR).

2^{k-1} Factorial design

The previous design is not the only 2^{k-1} we can do with $k =$. Here we show other two designs

Treatment	P	Q	R	PQ	PR	QR	PQR
000	-	-	-	+	+	+	-
011	-	+	+	-	-	+	-
101	+	-	+	-	+	-	-
110	+	+	-	+	-	-	-

In the design above the aliasings are exactly the same as before (P with QR, Q with PR, R with PQ, and (1) with PQR).

Treatment	P	Q	R	PQ	PR	QR	PQR
001	-	-	+	+	-	-	+
010	-	+	-	-	+	-	+
101	+	-	+	-	+	-	-
110	+	+	-	+	-	-	-

In the design above the aliasings are P with PQR, PQ with PR, and (1) with QR.

2^{k-1} Factorial design

Another way of constructing a 2^{3-1} design is by starting with a 2^2 design (2=3-1).

Treatment	P	Q	PQ
00	-	-	+
01	-	+	-
10	+	-	-
11	+	+	+

The we change PQ by R, knowing that we will be confounding R with PQ

Treatment	P	Q	R≡PQ
001	-	-	+
010	-	+	-
100	+	-	-
111	+	+	+

Then automatically other confoundings will be caused

Treatment	P	Q	R≡PQ	PR≡Q	QR≡P	PQR≡ (1)
001	-	-	+	-	-	+
010	-	+	-	+	-	+
100	+	-	-	-	+	+
111	+	+	+	+	+	+

2^{k-p} Factorial design

The design $R \equiv PQ$ can be written as $3 \equiv \pm 12$ meaning that the third column is the product of the first two (or minus the product of the first two). For a 2^{k-p} design we need p design equations, e.g., a 2^{8-3} design can be achieved with

$$\begin{aligned} 6 &\equiv \pm 345 \\ 7 &\equiv \pm 1245 \\ 8 &\equiv \pm 1235 \end{aligned}$$

If we multiply again by the 6th, 7th, 8th columns, then we have the equations

$$\begin{aligned} (1) &\equiv \pm 3456 \\ (1) &\equiv \pm 12457 \\ (1) &\equiv \pm 12358 \end{aligned}$$

That are called the generators of the design. The length of the shortest word amongst the generators is called the resolution. In our example length(3456)=4, so our design is of resolution IV

$$2^{8-3}_{IV} \text{ design}$$

2^{k-p} Factorial design and Resolution

Given the generator we may discover the rest of confounding terms associated to that equation:

$$(1) \equiv 3456, 3 \equiv 456, 4 \equiv 356, 5 \equiv 346, 6 \equiv 345, 34 \equiv 56, 35 \equiv 46, 36 \equiv 45$$

Resolution	Ability	Example
I	Not useful: an experiment of exactly one run only tests one level of a factor and hence can't even distinguish between the high and low levels of that factor	(1)≡1
II	Not useful: main effects are confounded with other main effects	(1)≡12
III	Estimate main effects, but these may be confounded with two-factor interactions	(1)≡123
IV	Estimate main effects unconfounded by two-factor interactions. Estimate two-factor interaction effects, but these may be confounded with other two-factor interactions	(1)≡1234
V	Estimate main effects unconfounded by three-factor (or less) interactions. Estimate two-factor interaction effects unconfounded by two-factor interactions. Estimate three-factor interaction effects, but these may be confounded with other two-factor interactions.	(1)≡12345
VI	Estimate main effects unconfounded by four-factor (or less) interactions. Estimate two-factor interaction effects unconfounded by three-factor (or less) interactions. Estimate three-factor interaction effects, but these may be confounded with other three-factor interactions.	(1)≡123456

Example 59 (continued)

We have identified 7 variables we may control, but we cannot afford more than 64 experiments. We do not foresee 3rd order interactions or higher. Can we perform this experiment?

We can do with even less (32 experiments): 2^{7-2}_{IV} with generators $6 \equiv 123$ and $7 \equiv 124$

... But I can afford up to 64 experiments, and I don't mind doing more than 32 to increase results accuracy!!!

Mirror-image foldover designs

Given a resolution III design we may increase its resolution to IV by mirroring it. All we have to do is to replicate the experiment and change the signs of all treatments

Example 61

For a 2^{5-2} fractional factorial design we have

	X_1	X_2	X_3	$X_4 = X_1X_2$	$X_5 = X_1X_3$		X_1	X_2	X_3	X_4	X_5
Run 1	-	-	-	+	+	Run 9	+	+	+	-	-
Run 2	-	-	+	+	-	Run 10	+	+	-	-	+
Run 3	-	+	-	-	+	Run 11	+	-	+	+	-
Run 4	-	+	+	-	-	Run 12	+	-	-	+	+
Run 5	+	-	-	-	-	Run 13	-	+	+	+	+
Run 6	+	-	+	-	+	Run 14	-	+	-	+	-
Run 7	+	+	-	+	-	Run 15	-	-	+	-	+
Run 8	+	+	+	+	+	Run 16	-	-	-	-	-
	Original design						Mirrored design				

Irregular fractions of 2^k Factorial designs

We may adopt a regression approach to the analysis of 2^k factorial designs. With the -1 and 1 encoding, the regression model would look like

$$\begin{aligned} y_{ijk} &= \beta_0 + \beta_p p_i + \beta_q q_j + \beta_{pq} p_i q_j + \beta_r r_k + \beta_{pr} p_i r_k + \beta_{qr} q_j r_k + \beta_{pqr} p_i q_j r_k \\ \mathbf{y} &= A\boldsymbol{\theta} \end{aligned}$$

Consider the fractional design

$$A = \begin{pmatrix} 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Treatment
001
010
100
111

$\longrightarrow$

$$A^T A = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 4 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 4 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 4 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 4 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Eigenvalues: $8(4), 0(4)$

Irregular fractions of 2^k Factorial designs

We now add two extra measurements (000 and 011)

Treatment
000
001
010
011
100
111

$\longrightarrow$

$$A = \begin{pmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$A^T A = \begin{pmatrix} 6 & -2 & 0 & 0 & 0 & 0 & 2 & 2 \\ -2 & 6 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 6 & -2 & 2 & 2 & 0 & 0 \\ 0 & 0 & -2 & 6 & 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 & 6 & -2 & 0 & 0 \\ 0 & 0 & 2 & 2 & -2 & 6 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 0 & 6 & -2 \\ 2 & 2 & 0 & 0 & 0 & 0 & -2 & 6 \end{pmatrix}$$

Eigenvalues: $8(6), 0(2)$

Irregular fractions of 2^k Factorial designs

Finally, we may compare it to the full factorial by adding two extra measurements (101 and 110)

Treatment
000
001
010
011
100
101
110
111

$\longrightarrow$

$$A = \begin{pmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$A^T A = \begin{pmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{pmatrix}$$

Eigenvalues: $8(8)$

Outline

Screening designs

Example 62

We are screening drugs and we expect that most of them do not have any effect. We also expect that there is no interaction between compounds. Can we screen many different compounds with as few runs as possible?

If very few effects are expected to have an effect, and we do not expect interactions, we may opt for a Resolution III design (they can estimate main effects, but confounded with 2nd order interactions).

Saturated designs

Example 63

Let us design an experiment for screening 7 factors with just $8 (= 2^3)$ treatments.

1) We start from a cyclic design to compare 7 treatments in 7 blocks of three units per block.

Block	1	2	3	4	5	6	7
Treatments	1	2	3	4	5	6	7
	2	3	4	5	6	7	1
	4	5	6	7	1	2	3

2) We now convert each block (i) to a factor, and put $+1$ if the treatment j was in block i. Finally add a run with all factors

Factor	1	2	3	4	5	6	7
Run 1	+	-	-	-	+	-	+
Run 2	+	+	-	-	-	+	-
Run 3	-	+	+	-	-	-	+
Run 4	+	-	+	+	-	-	-
Run 5	-	+	-	+	+	-	-
Run 6	-	-	+	-	+	+	-
Run 7	-	-	-	+	+	+	+
Run 8	+	+	+	+	+	+	+

Plackett-Burman designs

Plackett-Burman designs are also very popular for screening a large number of factors. They exist for a number of runs that is a multiple of 4 (20 in the example).

Example 64

Let us design an experiment for screening 16 factors with just 20 treatments.

Plackett - Burman Design

Factors:	16	Replicates:	1
Base runs:	20	Total runs:	20
Base blocks:	1	Total blocks:	1

Design Table

Run	Blk	A	B	C	D	E	F	G	H	J	K	L	M	N	O	P	Q
1	1	+	-	+	+	-	-	-	-	+	-	+	-	+	+	+	+
2	1	+	+	-	+	+	-	-	-	-	+	-	+	-	+	+	+
3	1	-	+	+	-	+	+	-	-	-	-	+	-	+	-	+	+
4	1	-	-	+	+	-	+	+	-	-	-	-	+	-	+	-	+
5	1	+	-	-	+	+	-	+	+	-	-	-	-	+	-	+	-
6	1	+	+	-	-	+	+	-	+	+	-	-	-	-	+	-	+
7	1	+	+	+	-	-	+	+	-	+	+	-	-	-	-	+	-
8	1	+	+	+	+	-	-	+	+	-	+	+	-	-	-	-	+
9	1	-	+	+	+	+	-	-	+	+	-	+	+	-	-	-	-
10	1	+	-	+	+	+	+	-	-	+	+	-	+	+	-	-	-
11	1	-	+	-	+	+	+	+	-	-	+	+	-	+	+	-	-
12	1	+	-	+	-	+	+	+	+	-	-	+	+	-	+	+	-
13	1	-	+	-	+	-	+	+	+	+	-	-	+	+	-	+	+
14	1	-	-	+	-	+	-	+	+	+	+	-	-	+	+	-	+
15	1	-	-	-	+	-	+	-	+	+	+	+	-	-	+	+	-
16	1	-	-	-	-	+	-	+	-	+	+	+	+	-	-	+	+
17	1	+	-	-	-	-	+	-	+	-	+	+	+	+	-	-	+
18	1	+	+	-	-	-	-	+	-	+	-	+	+	+	+	-	-
19	1	-	+	+	-	-	-	-	+	-	+	-	+	+	+	+	-
20	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Outline

Blocking causes confounding

Example 65

Consider a 2^3 full factorial design with factors A, B, C. We need 8 runs to perform the experiment. But we cannot run more than 4 experiments per day (and differences between days can be expected). The day of the experiment acts as a blocking variable.

$$y_{ijk} = b_{block(ijk)} + \beta_0 + \beta_A a_i + \beta_B b_j + \beta_{AB} a_i b_j + \beta_C c_k + \beta_{AC} a_i c_k + \beta_{BC} b_j c_k + \beta_{ABC} a_i b_j c_k$$

Treatment	A	B	AB	C	AC	BC	ABC	Day
(1)	-	-	+	-	+	+	-	Day 1
a	+	-	-	-	-	+	+	Day 2
b	-	+	-	-	+	-	+	Day 2
ab	+	+	+	-	-	-	-	Day 1
c	-	-	+	+	-	-	+	Day 2
ac	+	-	-	+	+	-	-	Day 1
bc	-	+	-	+	-	+	-	Day 1
abc	+	+	+	+	+	+	+	Day 2

Blocking causes confounding

Example 65(continued)

We now reorganize treatments in the same block together

Treatment	A	B	AB	C	AC	BC	ABC	Day
(1)	-	-	+	-	+	+	-	Day 1
ab	+	+	+	-	-	-	-	Day 1
ac	+	-	-	+	+	-	-	Day 1
bc	-	+	-	+	-	+	-	Day 1
a	+	-	-	-	-	+	+	Day 2
b	-	+	-	-	+	-	+	Day 2
c	-	-	+	+	-	-	+	Day 2
abc	+	+	+	+	+	+	+	Day 2

In Day 1 we only have - signs in ABC, while in Day 2 we only have + signs. This means that the ABC effect has been confounded with the blocks. For the rest of variables, each block contains the same number of + and - signs.

The only way of escaping from confounding is by replication.

Blocking causes confounding

Example 66

Assume that we cannot perform more than 2 experiments per day, and we decide to sacrifice BC interactions. Now the blocks may look like

Treatment	A	B	AB	C	AC	BC	ABC	Day
(1)	-	-	+	-	+	+	-	Day 1
bc	-	+	-	+	-	+	-	Day 1
ab	+	+	+	-	-	-	-	Day 2
ac	+	-	-	+	+	-	-	Day 2
a	+	-	-	-	-	+	+	Day 3
abc	+	+	+	+	+	+	+	Day 3
b	-	+	-	-	+	-	+	Day 4
c	-	-	+	+	-	-	+	Day 4

Blocking causes confounding

Once we decide to confound a treatment, other treatments get also confound. In the example we have decided to confound BC and ABC. However, any other treatment that can be reached by generalized interaction also gets confounded

Given any two interactions, the generalized interaction is obtained by multiplying the factors (in capital letters) and ignoring all the terms with an even exponent.

$$ABC \times BCD = AB^2C^2D = AD$$
$$AB \times BC \times ABC = A^2B^3C^2 = B$$

In our example

$$BC \times ABC = AB^2C^2 = A$$

A has also been confounded!!!

A better choice would have been AB and BC.

Blocking 2^k factorial designs

Let us show the procedure through an example

Example 67

We have an experiment with 5 two-level factors (A,B,C,D,E) and consequently there are $2^5 = 32$ treatments to be estimated. The runs need to be allocated in $8 = 2^3$ blocks of size $4 = 2^2$. We need to confound $7 = 2^3 - 1$ treatments. But these confounded treatments are not independent.

1) Choose 3 treatments to confound: AD, BE, ABC
2) Construct the remaining 4 treatments by generalized interaction:

$$AD \times BE = ABDE$$
$$AD \times ABC = BCD$$
$$BE \times ABC = ACE$$
$$AD \times BE \times ABC = CDE$$

Blocking 2^k factorial designs

Example 67(continued)

3) Write the treatments in the standard order

(1)	a	b	ab	c	ac	bc	abc
d	ad	bd	abd	cd	acd	bcd	abcd
e	ae	be	abe	ce	ace	bce	abce
de	ade	bde	abde	cde	acde	bcde	abcde

4) Construct a principal block:

- A treatment belongs to the principal block if it has an even number of letters in common with the generating, confounded treatments (AD, BE, ABC).
- If two treatments belong to the principal block, so does their generalized interaction.

The principal block is not unique. In this case we will use

Block 1
(1)
acd
bce
abde(=acd$\times$ bce)

Blocking 2^k factorial designs

Example 67(continued)

5) Construct the rest of blocks by multiplying the first block by the "head" of the columns in the standard table

(1)	a	b	ab	c	ac	bc	abc
d	ad	bd	abd	cd	acd	bcd	abcd
e	ae	be	abe	ce	ace	bce	abce
de	ade	bde	abde	cde	acde	bcde	abcde

Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8
(1)	a	b	ab	c	ac	bc	abc
acd	cd	abcd	bcd	ad	d	abd	bd
bce	abce	ce	ace	be	abe	e	ae
abde	bde	ade	de	abcde	bcde	acde	cde

Blocking mixed-level factorial design

Example 68

We want to test the effect in maize growth of

- 2 levels of nitrogen (N)
- 3 spatial arrangements (S)
- 2 management systems (M)
- 2 maize genotypes (G, if possible)

We have 4 blocks of 12 plots each. We want to estimate all main effects and the interactions NS, NM, NSM, GN, GS, GM.

Blocking mixed-level factorial design

Example 68(continued)

If the 2 genotypes are not tested we have 12 treatments that will be allocated in the 12 plots and replicated 4 times. The ANOVA table in this case would be

	df
Blocks	3
Nitrogen (N)	1
Spacings (S)	2
Management (M)	1
NS interactions	2
NM interactions	1
SM interactions	2
NSM interactions	2
Error	33
Total	47

33 df for the error is a lot, so we try to introduce the genotypes.

Blocking mixed-level factorial design

Example 68(continued)

1. To be able to estimate main effects of a factor (e.g. G), each level must appear the same number of times in each block and appear with each of the combinations of the rest of factors (e.g. NSM).

Block I	Block II
$n_1 s_1 m_1 g_1$	$n_1 s_1 m_1 g_2$
$n_1 s_2 m_2 g_2$	$n_1 s_2 m_2 g_1$
$n_1 s_3 m_1 g_1$	$n_1 s_3 m_1 g_2$
$n_1 s_1 m_2 g_2$	$n_1 s_1 m_2 g_1$
$n_1 s_2 m_1 g_1$	$n_1 s_2 m_1 g_2$
$n_1 s_3 m_2 g_2$	$n_1 s_3 m_2 g_1$
$n_2 s_1 m_1 g_1$	$n_2 s_1 m_1 g_2$
$n_2 s_2 m_2 g_2$	$n_2 s_2 m_2 g_1$
$n_2 s_3 m_1 g_1$	$n_2 s_3 m_1 g_2$
$n_2 s_1 m_2 g_2$	$n_2 s_1 m_2 g_1$
$n_2 s_2 m_1 g_1$	$n_2 s_2 m_1 g_2$
$n_2 s_3 m_2 g_2$	$n_2 s_3 m_2 g_1$

2 replicates of the following design will be performed.

It can be easily verified that N, S and M fulfill this condition. G is the most

compromised factor and looking at the table this condition is fulfilled.

Blocking mixed-level factorial design

Example 68(continued)

2. To be able to estimate 2nd order interactions (e.g. GN), each combination must appear the same number of times in each block and appear with each of the combinations of the rest of factors (e.g. SM).

Block I	Block II
$n_1 s_1 m_1 g_1$	$n_1 s_1 m_1 g_2$
$n_1 s_2 m_2 g_2$	$n_1 s_2 m_2 g_1$
$n_1 s_3 m_1 g_1$	$n_1 s_3 m_1 g_2$
$n_1 s_1 m_2 g_2$	$n_1 s_1 m_2 g_1$
$n_1 s_2 m_1 g_1$	$n_1 s_2 m_1 g_2$
$n_1 s_3 m_2 g_2$	$n_1 s_3 m_2 g_1$
$n_2 s_1 m_1 g_1$	$n_2 s_1 m_1 g_2$
$n_2 s_2 m_2 g_2$	$n_2 s_2 m_2 g_1$
$n_2 s_3 m_1 g_1$	$n_2 s_3 m_1 g_2$
$n_2 s_1 m_2 g_2$	$n_2 s_1 m_2 g_1$
$n_2 s_2 m_1 g_1$	$n_2 s_2 m_1 g_2$
$n_2 s_3 m_2 g_2$	$n_2 s_3 m_2 g_1$

$n_1 g_1$ and $n_1 g_2$ fulfill this condition as has been highlighted. $n_2 g_1$, $n_2 g_2$ also do as can be easily verified.

NS, NM, SM, GS and GM interactions also fulfill this criterion.

Blocking mixed-level factorial design

Example 68(continued)

3. To be able to estimate 3rd order interactions (e.g. NSM), each combination must appear the same number of times in each block and appear with each of the combinations of the rest of factors (e.g. G).

Block I	Block II
$n_1 s_1 m_1 g_1$	$n_1 s_1 m_1 g_2$
$n_1 s_2 m_2 g_2$	$n_1 s_2 m_2 g_1$
$n_1 s_3 m_1 g_1$	$n_1 s_3 m_1 g_2$
$n_1 s_1 m_2 g_2$	$n_1 s_1 m_2 g_1$
$n_1 s_2 m_1 g_1$	$n_1 s_2 m_1 g_2$
$n_1 s_3 m_2 g_2$	$n_1 s_3 m_2 g_1$
$n_2 s_1 m_1 g_1$	$n_2 s_1 m_1 g_2$
$n_2 s_2 m_2 g_2$	$n_2 s_2 m_2 g_1$
$n_2 s_3 m_1 g_1$	$n_2 s_3 m_1 g_2$
$n_2 s_1 m_2 g_2$	$n_2 s_1 m_2 g_1$
$n_2 s_2 m_1 g_1$	$n_2 s_2 m_1 g_2$
$n_2 s_3 m_2 g_2$	$n_2 s_3 m_2 g_1$

$n_1 s_1 m_1$ fulfills this condition as has been highlighted. The rest of NSM combinations also do as can be easily verified.

But $s_3 m_1 g_1$ does not.

Blocking mixed-level factorial design

Example 68(continued)

24 df for the error is enough and we have gain much analytical capabilities.

	df
Blocks	3
Nitrogen (N)	1
Spacings (S)	2
Management (M)	1
NS interactions	2
NM interactions	1
SM interactions	2
GN interactions	1
GS interactions	1
GM interactions	1
NSM interactions	2
NSG interactions	2
SMG interactions	2
Error	24
Total	47

Incompletely confounded designs

Classical designs have concentrated in completely confounded (e.g. NSM in the example above) or completely unconfouded effects (e.g. SMG in the example above). However, with computers we may have partially confounded parameters

Example 69(continued)

With the same experiment as above

Block I	Block II
$n_1 s_1 m_1 g_1$	$n_1 s_1 m_1 g_2$
$n_1 s_2 m_2 g_2$	$n_1 s_2 m_2 g_1$
$n_1 s_3 m_1 g_2$	$n_1 s_3 m_1 g_1$
$n_1 s_1 m_2 g_1$	$n_1 s_1 m_2 g_2$
$n_1 s_2 m_1 g_1$	$n_1 s_2 m_1 g_2$
$n_1 s_3 m_2 g_2$	$n_1 s_3 m_2 g_1$
$n_2 s_1 m_1 g_2$	$n_2 s_1 m_1 g_1$
$n_2 s_2 m_2 g_1$	$n_2 s_2 m_2 g_2$
$n_2 s_3 m_1 g_1$	$n_2 s_3 m_1 g_2$
$n_2 s_1 m_2 g_2$	$n_2 s_1 m_2 g_1$
$n_2 s_2 m_1 g_2$	$n_2 s_2 m_1 g_1$
$n_2 s_3 m_2 g_1$	$n_2 s_3 m_2 g_2$

The GNM effect is estimated as

$$\begin{aligned} GNM &= (g_2 - g_1)(n_2 - n_1)(m_2 - m_1)(s_1 + s_2 + s_3) \\ &= +(n_1 m_1 g_2 + n_1 m_2 g_1 + n2 m_1 g_1 + n_2 m_2 g_2)(s_1 + s_2 + s_3) \\ &\quad -(n_1 m_1 g_1 + n_1 m_2 g_2 + n2 m_1 g_2 + n_2 m_2 g_1)(s_1 + s_2 + s_3) \end{aligned}$$

GNM is not totally confounded with the blocks, it is only partially confounded.

Completely confounded mixed-level factorials

Consider an experiment with k factors. There is no restriction on the number of levels of each factor (2, 3, 4, ...). If the number of experiments is restricted to 1 full replicate of the factorial, then some interactions must be confounded with the blocks and will be inestimable.

The model of a blocked factorial experiment can be written as

$$\mathbf{y} = X\boldsymbol{\theta} + Z\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\theta}$ is the vector of estimable treatment effects, and $\boldsymbol{\beta}$ is the vector of block effects (confounded with the inestimable treatment effects).

The goal of the D_s-optimal design is to minimize the covariance matrix of the LS estimator, or equivalently maximize the determinant of

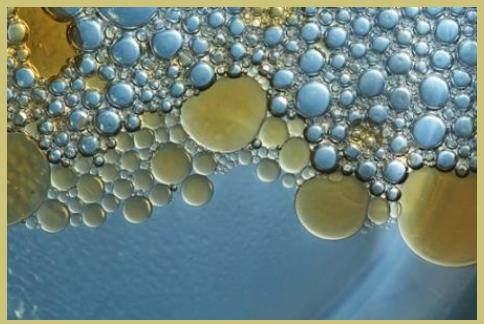
$$X^*, Z^* = \arg\max_{X,Z} X^T(I - Z(Z^T Z^{-1} Z))X$$

Lawson, Schaalje, Collings. Blocking Mixed-Level Factorials with SAS. J. Statistical Software, 32, 1 (2009)

Outline

Factorial designs for quantitative factors

Example 69



We are preparing a formulation for a drug that must be delivered as an emulsion. We may dissolve the drug in 3 compounds simltaneously. The goal is to determine the optimal concentration of each of the three compounds such that the amount released is maximized. We will study 3 levels of each of the 3 compounds.

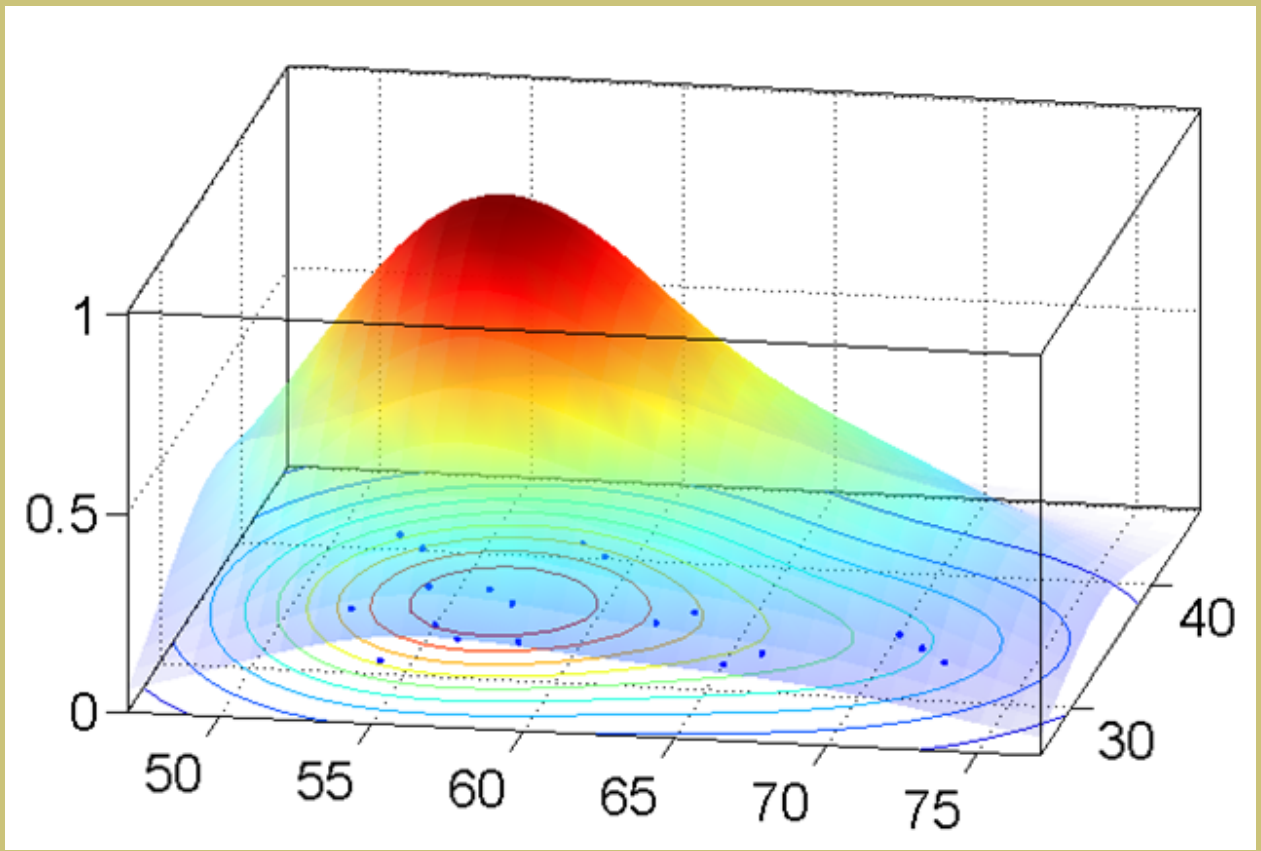
First we need to choose which function will be used to model the data

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$	Plane: Allows linear estimation
$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{12} X_1 X_2 \beta_{12} X_2^2 + \epsilon$	Quadratic: Allows linear estimation
$Y = A(1 - e^{-\beta_1 X_1 - \beta_2 X_2}) + \epsilon$	Asymptotic response high
$Y = e^{-\beta_1 X_1 - \beta_2 X_2} + \epsilon$	Asymptotic response low
$Y = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2)}} + \epsilon$	Logistic function
$\frac{1}{Y} = \beta_0 + \frac{\beta_1}{X_1} + \epsilon$	Michaelis-Menten
...	...

Factorial designs for quantitative factors

Example 70

The goal of the experimental design is to distribute the samples in X_1 and X_2 such that the observations, Y, obtained at this locations allow estimating with the "maximum precision" the coefficients describing the response surface.



Optimal designs

Our regression model, in general, will be of the form

$$Y = f(\mathbf{X}, \boldsymbol{\theta}) + \boldsymbol{\epsilon}$$

The regression will find the $\hat{\boldsymbol{\Theta}}$ that better fits the experimental data, and we should have an expected value of the covariance matrix of $\boldsymbol{\Theta}$

$$\Sigma_{\boldsymbol{\theta}}$$

The inverse of this matrix is called Fisher's Information Matrix

$$I_{\boldsymbol{\theta}} = \Sigma_{\boldsymbol{\theta}}^{-1}$$

This inverse depends solely on $\mathbf{X}$ (fixed by our experimental design) and σ_ϵ^2 (the experimental noise). So, by judiciously choosing the $\mathbf{X}$ values we should be able to minimize the uncertainty about the regression parameters.

Optimal designs

For linear models

$$\begin{aligned} Y &= X\boldsymbol{\theta} + \boldsymbol{\epsilon} \\ \Sigma_{\boldsymbol{\theta}} &= \sigma_\epsilon^2 (X^T X)^{-1} \\ I_{\boldsymbol{\theta}} &= \frac{1}{\sigma_\epsilon^2} X^T X \end{aligned}$$

The covariance of the predictions is given by

$$\Sigma_Y = X^T \Sigma_{\boldsymbol{\theta}} X$$

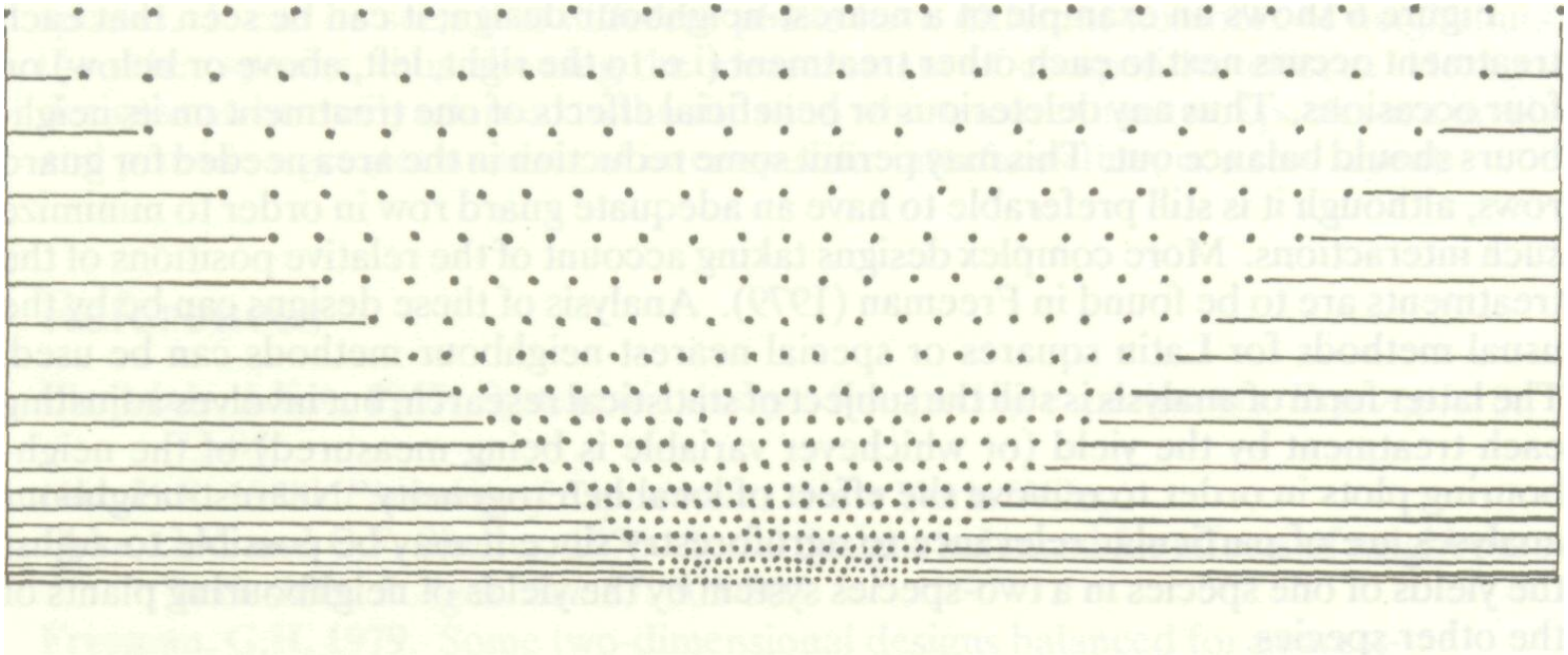
There are several optimization criteria

D-optimal	Maximize the determinant of $I_{\boldsymbol{\theta}}$
A-optimal	Minimize the trace of $\Sigma_{\boldsymbol{\theta}}$
T-optimal	Maximize the trace of $I_{\boldsymbol{\theta}}$
E-optimal	Maximize the minimum eigenvalue of $I_{\boldsymbol{\theta}}$
G-optimal	Minimize the maximum entry of Σ_Y
I-optimal	Minimize the trace of Σ_Y

None of them is necessarily better than the rest and it depends on our experimental objectives.

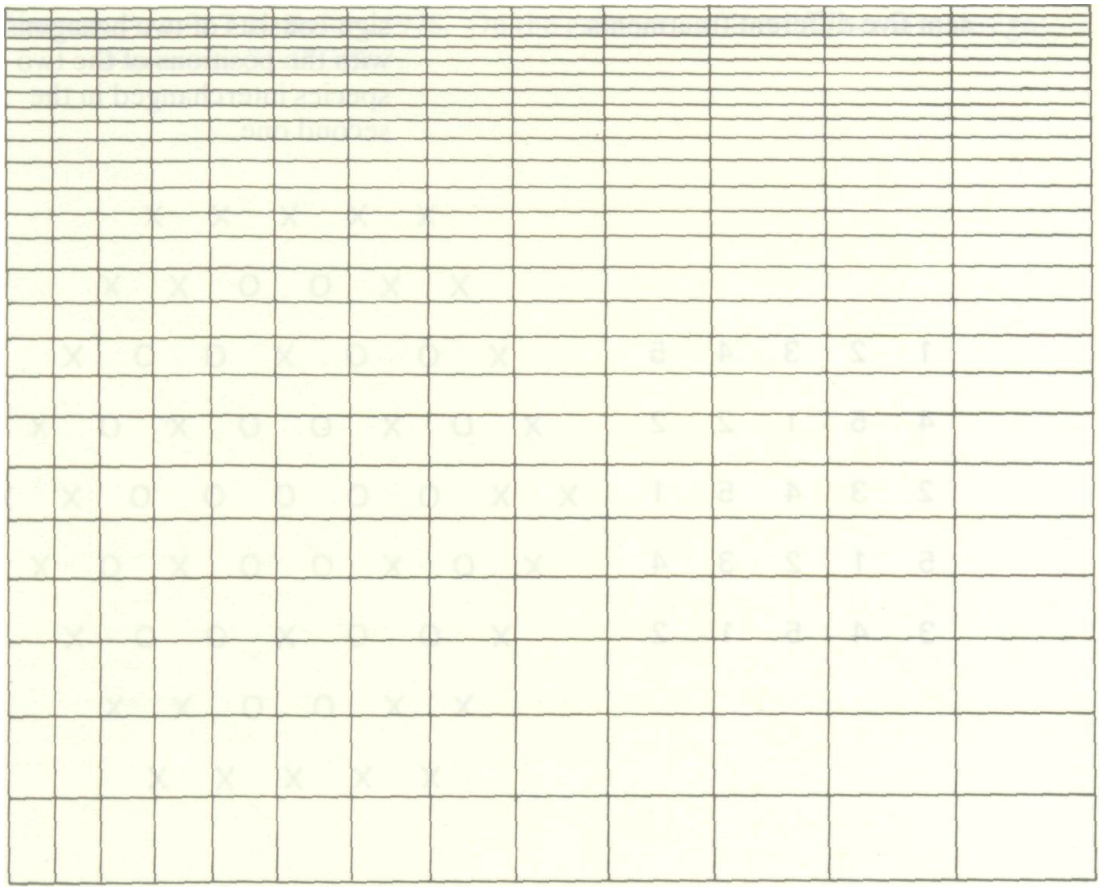
(Nelder) Systematic designs

Systematic designs aim at minimizing the effect of a gradient of an interfering variable. They are used in agricultural experiments.



(Nelder) Systematic designs

The following design blocks two orthogonal gradients.

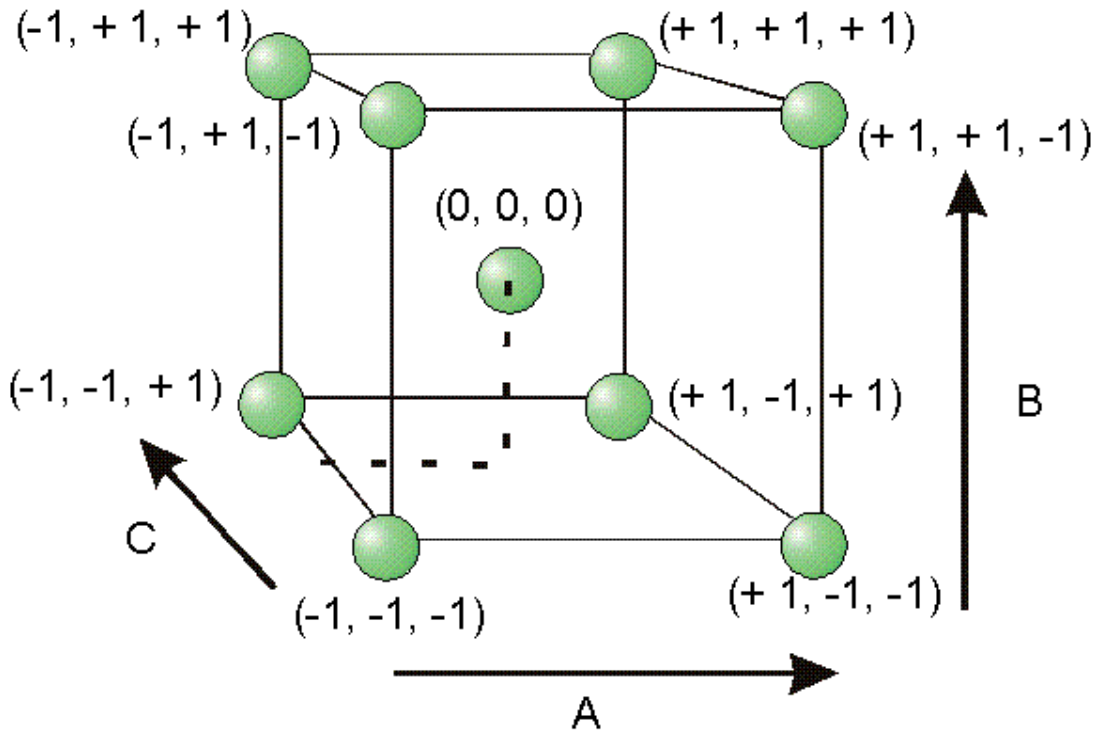


Linear models and 2^k factorial designs

If the model to be estimated includes only main effects and second order interactions

$$Y = \beta_0+\beta_1X_1+\beta_2X_2+\beta_{11}X_1^2+\beta_{12}X_1X_2+\beta_{22}X_2^2+\beta_3X_3+\beta_{13}X_1X_3+\beta_{23}X_2X_3+\beta_{33}X_3^2$$

then pure 2^k factorial designs cannot estimate the quadratic terms of the form X_i^2. Extra samples need to be added, converting each factor into a 3-level factor (3^k-factorial design).



Linear models and 2^k factorial designs

Centerpoints are added for

- To provide a measure of process stability and inherent variability.
- To check for curvature.

- Centerpoint runs should begin and end the experiment, ...
- ... and should be dispersed as evenly as possible throughout the design matrix.
- The centerpoint runs are not randomized! There would be no reason to randomize them as they are there as guardians against process instability and the best way to find instability is to sample the process on a regular basis.

As a rough guide, you should generally add approximately 3 to 5 centerpoint runs to a full or fractional factorial design.

(NIST Engineering statistics handbook)

Linear models and 2^k factorial designs

The full 3^k-factorial design allows estimating 3rd order interactions

$$\begin{aligned} Y = & \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \text{ (main effects)} \\ & + \beta_{11} X_1^2 + \beta_{12} X_1 X_2 + \beta_{22} X_2^2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 + \beta_{33} X_3^2 \text{ (2nd order)} \\ & + \beta_{111} X_1^3 + \beta_{112} X_1^2 X_2 + \beta_{113} X_1^2 X_3 + \beta_{122} X_1 X_2^2 + \beta_{123} X_1 X_2 X_3 \text{ (3rd order)} \\ & + \beta_{133} X_1 X_3^2 + \beta_{222} X_2^3 + \beta_{223} X_2^2 X_3 + \beta_{333} X_3^3 \text{ (3rd order)} \end{aligned}$$

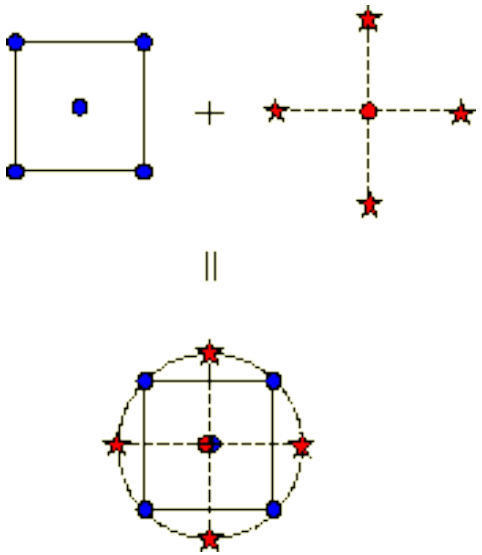
But the full factorial experimental quickly calls for many experiments, we may stay at the level of the quadratic function

k	Full 3^k	Quadratic terms
2	9	6
3	27	10
4	81	15
5	243	21
6	729	28

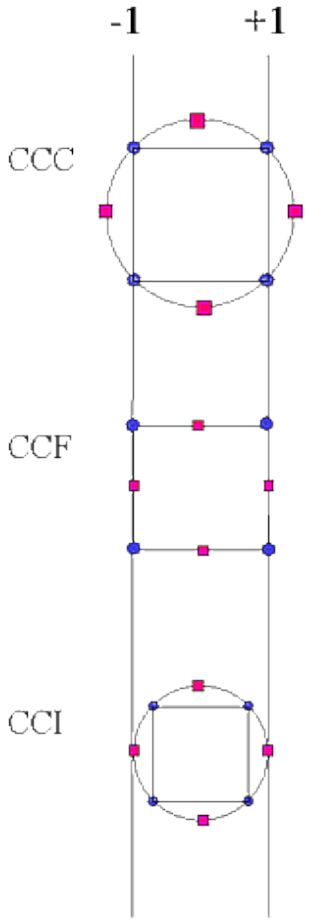
A fractional design is required. Typical designs are Box-Wilson central composite designs (CCC, CCI, or CCF) or Box-Behnken designs.

Box-Wilson Central Composite Designs

Central composite design complements a full 2^k design with middle point, with a star (axial observations).



CCC=Circumscribed
CCI=Inscribed
CCF=Face centered



The distance from the points in the star to the center is

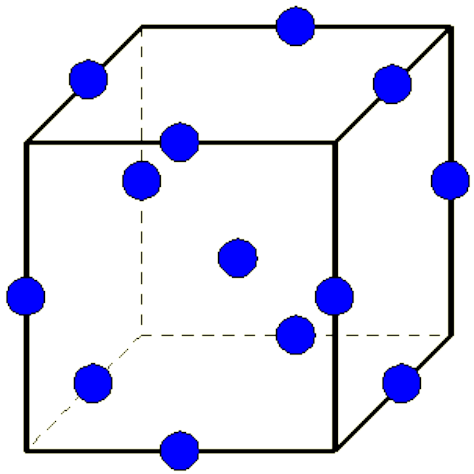
$$\alpha = (2^k)^{1/4}$$

An example with $k = 2$ blocks.

X_1	X_2
-1	-1
-1	1
1	-1
1	1
0	0
0	0
$-\sqrt{2}$	0
$\sqrt{2}$	0
0	$-\sqrt{2}$
0	$\sqrt{2}$
0	0
0	0

Box-Behnken Designs

These designs are like the full 2^k factorial with a middle sample, but the samples are at the edges of the cube (hypercube).



The advantage is that it requires fewer runs than the Box-Wilson designs.

CCF			Box-Behnken		
X_1	X_2	X_3	X_1	X_2	X_3
-	-	-	0	-	-
-	-	+	0	-	+
-	+	-	0	+	-
-	+	+	0	+	+
+	-	-	-	0	-
+	-	+	-	0	+
+	+	-	+	0	-
+	+	+	+	0	+
0	0	-	-	-	0
0	0	+	-	+	0
0	-	0	+	-	0
0	+	0	+	+	0
-	0	0	0	0	0 (3 repl)
+	0	0			
0	0	0(6 repl)			
20 runs			15 runs		

Blocking Response Surface Designs

- CCC allows blocking
- CCF does not. Box-Behnken allows blocking only in limited circumstances.

For a CCC design, $b = 2$ blocks are easily obtained by separating the full factorial design and the axial design

Block	X_1	X_2
1	-1	-1
1	-1	1
1	1	-1
1	1	1
1	0	0
1	0	0
2	$-\sqrt{2}$	0
2	$\sqrt{2}$	0
2	0	$-\sqrt{2}$
2	0	$\sqrt{2}$
2	0	0
2	0	0

Blocking Response Surface Designs

For a CCC design, $b = 3$ blocks the full factorial design is split in two and the axial design is not split.

Block	X_1	X_2	X_3
1	-1	-1	-1
2	-1	-1	1
2	-1	1	-1
1	-1	1	1
2	1	-1	-1
1	1	-1	1
1	1	1	-1
2	1	1	1
1	0	0	0
1	0	0	0
2	0	0	0
2	0	0	0
3	$-2^{3/4}$	0	0
3	$2^{3/4}$	0	0
3	0	$-2^{3/4}$	0
3	0	$2^{3/4}$	0
3	0	0	$-2^{3/4}$
3	0	0	$2^{3/4}$
3	0	0	0
3	0	0	0

Outline

Conclusions

C.O.S. Sorzano
coss@cnb.csic.es

National Center of Biotechnology (CSIC)

October 14, 2016

Outline

Outline

Experiment selection key

TABLE 3.1: Design Selection Guideline

Number of Factors	Comparative Objective	Screening Objective	Response Surface Objective
1	1-factor completely randomized design	–	–
2 - 4	Randomized block design	Full or fractional factorial	Central composite or Box-Behnken
5 or more	Randomized block design	Fractional factorial or Plackett-Burman	Screen first to reduce number of factors

NIST Handbook of Statistics

Experiment selection key

Experimental Design Selection Key

1. Do you want to test for differences among treatment means?

- **Yes**:
  - >> I know my ANOVA design, or
  - >> If you want help choosing your **ANOVA** design, go to **red arrow 2** below.
- **No, I want to explore relationships among variables**:
  - >> I know my **Regression** design, or
  - >> I want help choosing my **Regression** design (these pages under construction)

For ANOVA, you may choose among seven experimental designs via this key. The choice you make below will lead you to pages that enable you to then choose among seven treatment designs, with the choice to further refine the analysis with any combination of three specialized features (or no specialized features).

2. Are you blocking on a factor?

- **Yes** >> Go to **red arrow 3** below.
- **No** >> You have a Completely Randomized Design (What is CRD?)

3. Is your block too small to contain all treatments?

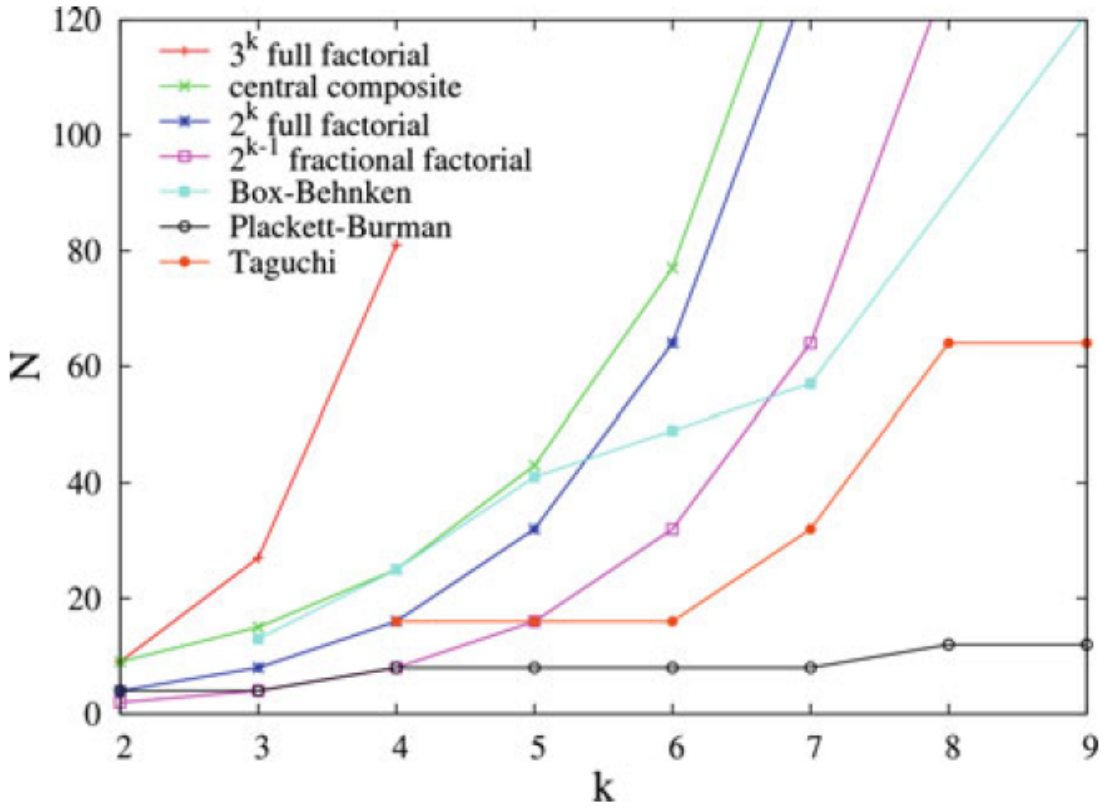
- **Yes** >> Go to **red arrow 4** below.
- **No** >> You have a Randomized Complete Block Design (What is RCBD?)

4. Does each block contain one and only one treatment?

- **Yes** >> Go to **red arrow 5** below.
- **No** >> You have an Incomplete Block Design (What is Incomplete Block Design?)

http://dawg.utk.edu/choose.htm

Experiment selection key



Number of experiments required by the DOE techniques

Cavazzutti, M. Optimization methods: From theory to design. Chap. 2

Experiment selection key

Table 2.14 DOE methods synoptic table

Method	Number of experiments	Suitability
RCBD	$N(L_i) = \prod_{i=1}^{k} L_i$	Focusing on a primary factor using blocking techniques
Latin squares	$N(L) = L^2$	Focusing on a primary factor cheaply
Full factorial	$N(L, k) = L^k$	Computing the main and the interaction effects, building response surfaces
Fractional factorial	$N(L, k, p) = L^{k-p}$	Estimating the main and the interaction effects
Central composite	$N(k) = 2^k + 2k + 1$	Building response surfaces
Box-Behnken	$N(k)$ from tables	Building quadratic response surfaces
Plackett-Burman	$N(k) = k + 4 - \mod\left(\frac{k}{4}\right)$	Estimating the main effects
Taguchi	$N(k_{in}, k_{out}, L) = N_{in} N_{out}$, $N_{in}(k_{in}, L)$, $N_{out}(k_{out}, L)$ from tables	Addressing the influence of noise variables
Random	chosen by the experimenter	Building response surfaces
Halton, Faure, Sobol	chosen by the experimenter	Building response surfaces
Latin hypercube	chosen by the experimenter	Building response surfaces
Optimal design	chosen by the experimenter	Building response surfaces

Cavazzutti, M. Optimization methods: From theory to design. Chap. 2

Conclusions

- Define the objectives of the experiment.
- Identify all sources of variation, including:
  - treatment factors and their levels
  - experimental units
  - blocking factors, noise factors, and covariates
- Choose appropriate rule for assigning the experimental units to the treatments. Remind:
  - Randomization
  - Orthogonality
  - Replication
  - Blocking
- Specify the measurements to be made.
- Run a pilot study if possible.
- Specify the model.
- Outline the analysis.
- Calculate the number of observations that need to be taken.

Conclusions

Outline