Design of Experiments: I, II

R. A. Bailey

Queen Mary
University of London

r.a.bailey@qmul.ac.uk

Avignon, 18 March 2010

Outline

- I: Basic tips
- II: My book *Design of Comparative Experiments*
- III: Factorial designs—Confounding and Fractions
- IV: Panel diagrams and skeleton ANOVA
- V: Two-phase experiments

Basic tips

Three principles of experimental design

- Replication

- Control

- Randomization

Three principles of experimental design

- Replication
  - Increased replication usually decreases variance.
  - Increased replication may increase variability.
  - Increased replication usually increases power.
  - Increased replication increases costs (monetary and other).
  - Beware of false replication.
- Control
- Randomization

Three principles of experimental design

- Replication
  - Increased replication usually decreases variance.
  - Increased replication may increase variability.
  - Increased replication usually increases power.
  - Increased replication increases costs (monetary and other).
  - Beware of false replication.
- Control
  - Group the experimental units into blocks of alike units.
  - Concurrent comparison with “do nothing”.
  - Concurrent comparison with at least one other treatment.
- Randomization

Three principles of experimental design

- Replication
  - Increased replication usually decreases variance.
  - Increased replication may increase variability.
  - Increased replication usually increases power.
  - Increased replication increases costs (monetary and other).
  - Beware of false replication.
- Control
  - Group the experimental units into blocks of alike units.
  - Concurrent comparison with “do nothing”.
  - Concurrent comparison with at least one other treatment.
- Randomization
  - Why do we randomize?
  - How do we randomize?

Replication

- If there is too much replication then the experiment may waste time and money. If animals are to be sacrificed, it is unethical to use too many.

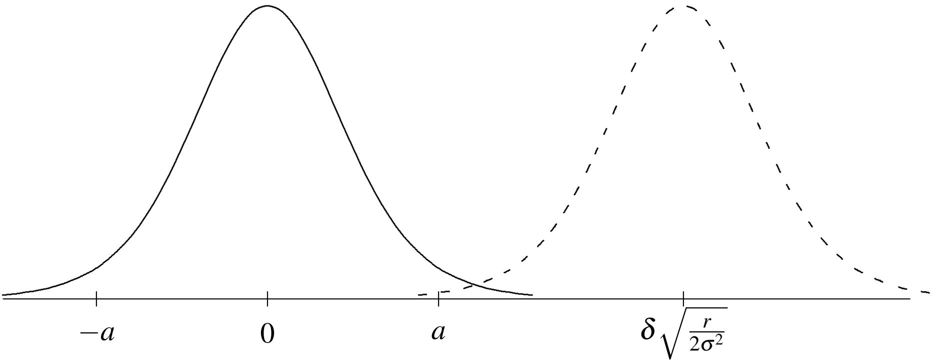
Replication

- If there is too much replication then the experiment may waste time and money. If animals are to be sacrificed, it is unethical to use too many.
- If there is too little replication then any genuine differences between treatments may be masked by the differences among the experimental units. An experiment which is too small to give any conclusions is also a waste of resources. It is also an unethical use of animals or people.

Replication

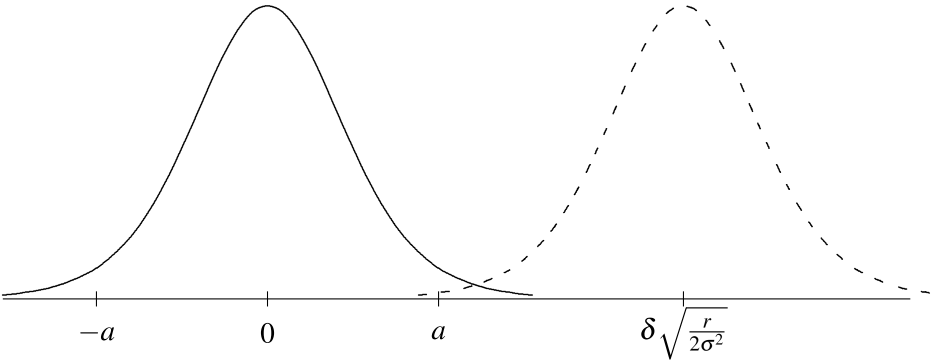
- If there is too much replication then the experiment may waste time and money. If animals are to be sacrificed, it is unethical to use too many.
- If there is too little replication then any genuine differences between treatments may be masked by the differences among the experimental units. An experiment which is too small to give any conclusions is also a waste of resources. It is also an unethical use of animals or people.
- Watch out for false replication.

Replication for power (two treatments, replication r)



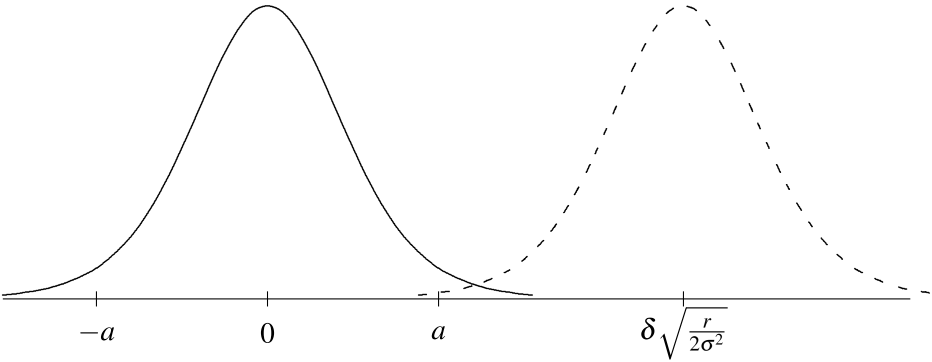
Solid curve defines the interval $[-a, a]$ used for the hypothesis test (where a depends on the significance level and on r).

Replication for power (two treatments, replication r)



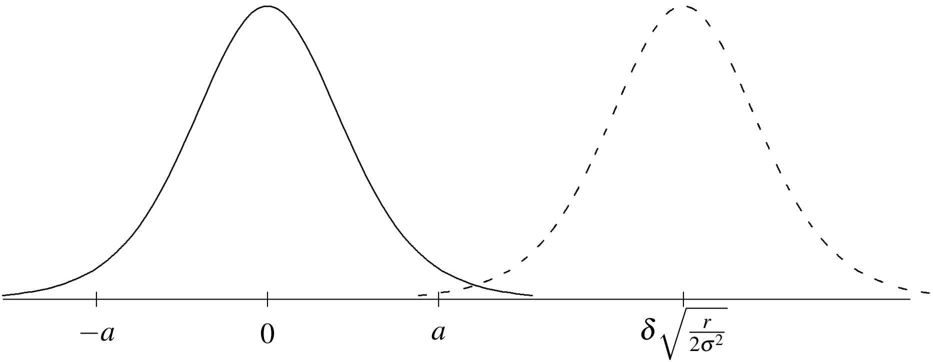
Solid curve defines the interval $[-a,a]$ used for the hypothesis test (where a depends on the significance level and on r).
Dashed curve gives the probability density of difference/s.e.d. if the real difference is δ and the variance per observation is σ^2.

Replication for power (two treatments, replication r)



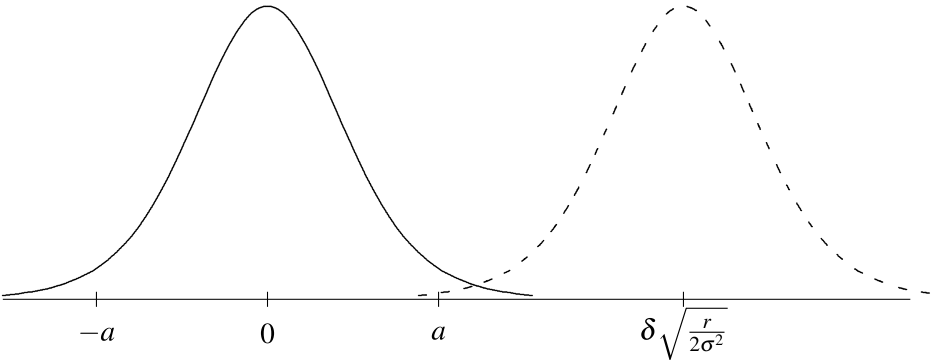
Solid curve defines the interval $[-a, a]$ used for the hypothesis test (where a depends on the significance level and on r).
Dashed curve gives the probability density of difference/s.e.d. if the real difference is δ and the variance per observation is σ^2.
Power for detecting a difference as big as δ is the proportion of the area under the dashed curve outside the interval $[-a, a]$.

Replication for power (two treatments, replication r)



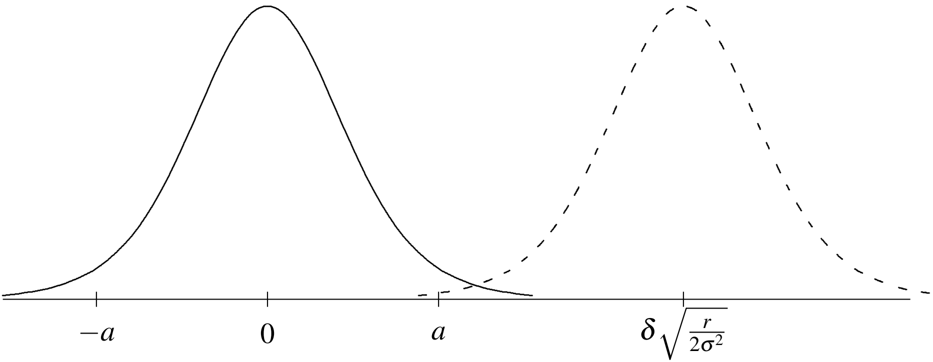
Power for detecting a difference as big as δ is the proportion of the area under the dashed curve outside the interval $[-a, a]$.

Replication for power (two treatments, replication r)



Power for detecting a difference as big as δ is the proportion of the area under the dashed curve outside the interval $[-a, a]$.
Increase $\delta \implies$ move dashed curve to right $\implies$ increase power.

Replication for power (two treatments, replication r)

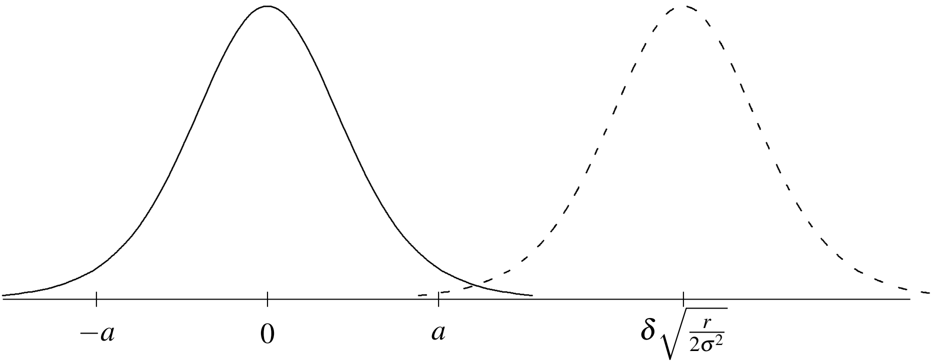


Power for detecting a difference as big as δ is the proportion of the area under the dashed curve outside the interval $[-a, a]$.

Increase $\delta \implies$ move dashed curve to right $\implies$ increase power.

Decrease $\sigma^2 \implies$ move dashed curve to right $\implies$ increase power.

Replication for power (two treatments, replication r)



Power for detecting a difference as big as δ is the proportion of the area under the dashed curve outside the interval $[-a,a]$.
Increase $\delta \implies$ move dashed curve to right $\implies$ increase power.
Decrease $\sigma^2 \implies$ move dashed curve to right $\implies$ increase power.
Increase $r \implies$ move dashed curve to right AND make both curves thinner $\implies$ increase power.

Diffusion of proteins

A post-doc added from 0 to 4 extra green fluorescent proteins to cells of *Escherichia coli*, adding 0 to 10 cells, 1 to 10 further cells, and so on. Then she measured the rate of diffusion of proteins in each of the 50 cells.

Diffusion of proteins

A post-doc added from 0 to 4 extra green fluorescent proteins to cells of *Escherichia coli*, adding 0 to 10 cells, 1 to 10 further cells, and so on. Then she measured the rate of diffusion of proteins in each of the 50 cells.

This is what she did.

Monday	Tuesday	Wednesday	Thursday	Friday
0000000000	1111111111	2222222222	3333333333	4444444444

Diffusion of proteins

A post-doc added from 0 to 4 extra green fluorescent proteins to cells of *Escherichia coli*, adding 0 to 10 cells, 1 to 10 further cells, and so on. Then she measured the rate of diffusion of proteins in each of the 50 cells.

This is what she did.

Monday	Tuesday	Wednesday	Thursday	Friday
0000000000	1111111111	2222222222	3333333333	4444444444

Are the perceived differences caused by differences in size?

Diffusion of proteins

A post-doc added from 0 to 4 extra green fluorescent proteins to cells of *Escherichia coli*, adding 0 to 10 cells, 1 to 10 further cells, and so on. Then she measured the rate of diffusion of proteins in each of the 50 cells.

This is what she did.

Monday	Tuesday	Wednesday	Thursday	Friday
0000000000	1111111111	2222222222	3333333333	4444444444

Are the perceived differences caused by differences in size?

Did she get better at preparing the samples as the week wore on?

Diffusion of proteins

A post-doc added from 0 to 4 extra green fluorescent proteins to cells of *Escherichia coli*, adding 0 to 10 cells, 1 to 10 further cells, and so on. Then she measured the rate of diffusion of proteins in each of the 50 cells.

This is what she did.

Monday	Tuesday	Wednesday	Thursday	Friday
0000000000	1111111111	2222222222	3333333333	4444444444

Are the perceived differences caused by differences in size?

Did she get better at preparing the samples as the week wore on?

Were there environmental changes in the lab that could have contributed to the differences?

Diffusion of proteins: continued

What she did.

Monday	Tuesday	Wednesday	Thursday	Friday
0000000000	1111111111	2222222222	3333333333	4444444444

Diffusion of proteins: continued

What she did.

Monday	Tuesday	Wednesday	Thursday	Friday
0000000000	1111111111	2222222222	3333333333	4444444444

Better to regard each day as a block.

Monday	Tuesday	Wednesday	Thursday	Friday
0011223344	0011223344	0011223344	0011223344	0011223344

Diffusion of proteins: continued

What she did.

Monday	Tuesday	Wednesday	Thursday	Friday
0000000000	1111111111	2222222222	3333333333	4444444444

Better to regard each day as a block.

Monday	Tuesday	Wednesday	Thursday	Friday
0011223344	0011223344	0011223344	0011223344	0011223344

There may still be systematic differences within each day,
so better still, randomize within each day.

Diffusion of proteins: continued

What she did.

Monday	Tuesday	Wednesday	Thursday	Friday
0000000000	1111111111	2222222222	3333333333	4444444444

Better to regard each day as a block.

Monday	Tuesday	Wednesday	Thursday	Friday
0011223344	0011223344	0011223344	0011223344	0011223344

There may still be systematic differences within each day,
so better still, randomize within each day.

Monday	Tuesday	Wednesday	Thursday	Friday
1040223134	2230110443	1421324030	4420013312	3204320411

Human-computer interaction

12 human subjects are each to undertake 4 tasks using computers to draw sketch maps. Response is number of correct features.

Human-computer interaction

12 human subjects are each to undertake 4 tasks using computers to draw sketch maps. Response is number of correct features.

Subjects make one sort of blocks.
Positions in time-order make another.

Human-computer interaction

12 human subjects are each to undertake 4 tasks using computers to draw sketch maps. Response is number of correct features.

Subjects make one sort of blocks.
Positions in time-order make another.

	Subject											
Time-order	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*
1	1	1	1	2	2	2	3	3	3	4	4	4
2	2	2	2	1	1	1	4	4	4	3	3	3
3	3	3	3	4	4	4	2	2	2	1	1	1
4	4	4	4	3	3	3	1	1	1	2	2	2

Human-computer interaction: an alternative interpretation

	Subject											
Experience	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*
1	1	1	1	2	2	2	3	3	3	4	4	4
2	2	2	2	1	1	1	4	4	4	3	3	3
3	3	3	3	4	4	4	2	2	2	1	1	1
4	4	4	4	3	3	3	1	1	1	2	2	2

Human-computer interaction: an alternative interpretation

	Subject											
Experience	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*
1	1	1	1	2	2	2	3	3	3	4	4	4
2	2	2	2	1	1	1	4	4	4	3	3	3
3	3	3	3	4	4	4	2	2	2	1	1	1
4	4	4	4	3	3	3	1	1	1	2	2	2

If we think that the 16 combinations of task with level of experience should give 16 different responses, then we cannot estimate them all from the above design, because the same combinations always occur with the same people.

	Subject											
Experience	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*
1	1	1	1	2	2	2	3	3	3	4	4	4
2	2	2	2	1	1	1	4	4	4	3	3	3
3	3	3	3	4	4	4	2	2	2	1	1	1
4	4	4	4	3	3	3	1	1	1	2	2	2

If we think that the 16 combinations of task with level of experience should give 16 different responses, then we cannot estimate them all from the above design, because the same combinations always occur with the same people.

	Subject											
Experience	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*
1	1	2	3	4	1	2	3	4	1	2	3	4
2	2	1	4	3	4	3	2	1	3	4	1	2
3	3	4	1	2	2	1	4	3	4	3	2	1
4	4	3	2	1	3	4	1	2	2	1	4	3

Why do we randomize?

It is to avoid

- systematic bias
  (for example, doing all the tests on treatment A in January
  then all the tests on treatment B in March)
- selection bias
  (for example, choosing the most healthy patients for the
  treatment that you are trying to prove is best)
- accidental bias
  (for example, using the first rats that the animal handler takes out
  of the cage for one treatment and the last rats for the other)
- cheating by the experimenter.

It also helps to justify the model.

Lanarkshire milk experiment: early 20th century

Treatments: extra milk rations or not.
These should have been randomized to the children within each school.
The teachers decided to give the extra milk rations to those children who were most undernourished.

A forestry experiment in a rectangle

7 varieties of guayule tree in a 5×7 rectangle, using a randomized complete-block design with the rows as blocks.

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

A forestry experiment in a rectangle

7 varieties of guayule tree in a 5×7 rectangle, using a randomized complete-block design with the rows as blocks.

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

A forestry experiment in a rectangle

7 varieties of guayule tree in a 5×7 rectangle, using a randomized complete-block design with the rows as blocks.

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

"Throw it away and re-randomize."

A forestry experiment in a rectangle

7 varieties of guayule tree in a 5×7 rectangle, using a randomized complete-block design with the rows as blocks.

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

"Throw it away and re-randomize."
For the 5×7 rectangle, the proportion of plans with no repeat in any column is only 0.000006.

Be honest with the statistician

"I didn't want to bother you with those details."

Constraints on the conduct of the experiment
should be incorporated into the design
(and therefore into the analysis),
not fudged in the randomization.

False replication

Three pesticides were compared for their side-effects on ladybirds.

A field was divided into three areas and one pesticide applied to each area. Ladybirds were counted on three samples from each area.

$$\begin{aligned} \text{Treatments} &= ? \\ \text{Experimental units} &= ? \\ \text{Observational units} &= ? \\ \text{Replication} &= ? \end{aligned}$$

False replication

Three pesticides were compared for their side-effects on ladybirds.

A field was divided into three areas and one pesticide applied to each area. Ladybirds were counted on three samples from each area.

$$\begin{aligned} \text{Treatments} &= \text{3 pesticides} \\ \text{Experimental units} &= ? \\ \text{Observational units} &= ? \\ \text{Replication} &= ? \end{aligned}$$

False replication

Three pesticides were compared for their side-effects on ladybirds.

A field was divided into three areas and one pesticide applied to each area. Ladybirds were counted on three samples from each area.

$$
\begin{aligned}
\text{Treatments} &= \quad \text{3 pesticides} \\
\text{Experimental units} &= \quad \text{3 areas} \\
\text{Observational units} &= \; ? \\
\text{Replication} &= \; ?
\end{aligned}
$$

False replication

Three pesticides were compared for their side-effects on ladybirds.

A field was divided into three areas and one pesticide applied to each area. Ladybirds were counted on three samples from each area.

$$\begin{array}{rcl} \text{Treatments} & = & \text{3 pesticides} \\ \text{Experimental units} & = & \text{3 areas} \\ \text{Observational units} & = & \text{9 samples} \\ \text{Replication} & = & ? \end{array}$$

False replication

Three pesticides were compared for their side-effects on ladybirds.

A field was divided into three areas and one pesticide applied to each area. Ladybirds were counted on three samples from each area.

$$\begin{aligned} \text{Treatments} &= \text{3 pesticides} \\ \text{Experimental units} &= \text{3 areas} \\ \text{Observational units} &= \text{9 samples} \\ \text{Replication} &= 1 \end{aligned}$$

My book *Design of Comparative Experiments*

Design of Comparative Experiments: Meaning?

NOT experiments to determine the exact value of g

Design of Comparative Experiments: Meaning?

NOT experiments to determine the exact value of g

BUT experiments to find out if A is better than B, and, if so, by how much.

Philosophy

The aim is to develop a coherent framework for thinking about the design and analysis of experiments

Philosophy

The aim is to develop a coherent framework for thinking about the design and analysis of experiments

BUT you cannot build a general theory until the reader has some pegs to hang it on.

Outline

Chapter 1 Forward Look
Show the reader that we are going to cover real experiments.
Get the reader thinking about experimental units, observational units, treatments

Outline

Chapter 1 Forward Look
Show the reader that we are going to cover real experiments.
Get the reader thinking about experimental units, observational units, treatments

Chapters 2–6 and 8
Gradually develop the main themes

Outline

Chapter 1 Forward Look
Show the reader that we are going to cover real experiments.
Get the reader thinking about experimental units, observational units, treatments

Chapters 2–6 and 8
Gradually develop the main themes

Chapters 7 and 9
Lighter uses of these ideas

Outline

Chapter 1 Forward Look
Show the reader that we are going to cover real experiments.
Get the reader thinking about experimental units, observational units, treatments

Chapters 2–6 and 8
Gradually develop the main themes

Chapters 7 and 9
Lighter uses of these ideas

Chapter 10 The Calculus of Factors
Unified general theory of orthogonal designs

Outline

Chapter 1 Forward Look
Show the reader that we are going to cover real experiments.
Get the reader thinking about experimental units, observational units, treatments

Chapters 2–6 and 8
Gradually develop the main themes

Chapters 7 and 9
Lighter uses of these ideas

Chapter 10 The Calculus of Factors
Unified general theory of orthogonal designs

Chapter 11
Incomplete-Block Designs

Chapters 12 and 13
Confounded and fractional factorial designs

Outline

Chapter 1 Forward Look
Show the reader that we are going to cover real experiments.
Get the reader thinking about experimental units, observational units, treatments

Chapters 2–6 and 8
Gradually develop the main themes

Chapters 7 and 9
Lighter uses of these ideas

Chapter 10 The Calculus of Factors
Unified general theory of orthogonal designs

Chapter 11
Incomplete-Block Designs

Chapters 12 and 13
Confounded and fractional factorial designs

Chapter 14 Backward Look
Putting it all together—
reflections that need most of the foregoing

Chapter 1 Forward Look

1. Stages in a statistically designed experiment—
   consultation, design, data collection, data scrutiny, analysis, interpretation

Chapter 1 Forward Look

1. Stages in a statistically designed experiment—
   consultation, design, data collection, data scrutiny, analysis, interpretation
2. The ideal and the reality—
   purpose, replication, local control, constraints, choice

Chapter 1 Forward Look

1. Stages in a statistically designed experiment—
   consultation, design, data collection, data scrutiny, analysis, interpretation
2. The ideal and the reality—
   purpose, replication, local control, constraints, choice
3. An example—
   coming up soon!

Chapter 1 Forward Look

1. Stages in a statistically designed experiment—
   consultation, design, data collection, data scrutiny, analysis, interpretation
2. The ideal and the reality—
   purpose, replication, local control, constraints, choice
3. An example—
   coming up soon!
4. Defining terms

Chapter 1 Forward Look

1. Stages in a statistically designed experiment—
   consultation, design, data collection, data scrutiny, analysis, interpretation
2. The ideal and the reality—
   purpose, replication, local control, constraints, choice
3. An example—
   coming up soon!
4. Defining terms
   - An experimental unit is the smallest unit to which a treatment can be applied.

Chapter 1 Forward Look

1. Stages in a statistically designed experiment—
   consultation, design, data collection, data scrutiny, analysis, interpretation
2. The ideal and the reality—
   purpose, replication, local control, constraints, choice
3. An example—
   coming up soon!
4. Defining terms
   - An experimental unit is the smallest unit to which a treatment can be applied.
   - A treatment is the entire description of what can be applied to an experimental unit.

Chapter 1 Forward Look

1. Stages in a statistically designed experiment—
   consultation, design, data collection, data scrutiny, analysis, interpretation
2. The ideal and the reality—
   purpose, replication, local control, constraints, choice
3. An example—
   coming up soon!
4. Defining terms
   - An experimental unit is the smallest unit to which a treatment can be applied.
   - A treatment is the entire description of what can be applied to an experimental unit.
   - An observational unit is the smallest unit on which a response will be measured.

Chapter 1 Forward Look

1. Stages in a statistically designed experiment—
   consultation, design, data collection, data scrutiny, analysis, interpretation
2. The ideal and the reality—
   purpose, replication, local control, constraints, choice
3. An example—
   coming up soon!
4. Defining terms
   - An experimental unit is the smallest unit to which a treatment can be applied.
   - A treatment is the entire description of what can be applied to an experimental unit.
   - An observational unit is the smallest unit on which a response will be measured.
5. Linear model

Calf-feeding experiment

Calves were housed in pens, with ten calves per pen. Each pen was allocated to a certain type of feed. Batches of this type of feed were put into the pen; calves were free to eat as much of this as they liked. Calves were weighed individually.

Feed D Pen 1 10 calves	Feed C Pen 2 10 calves	Feed D Pen 3 10 calves	Feed B Pen 4 10 calves
Feed B Pen 5 10 calves	Feed A Pen 6 10 calves	Feed A Pen 7 10 calves	Feed C Pen 8 10 calves

Calf-feeding experiment

Calves were housed in pens, with ten calves per pen. Each pen was allocated to a certain type of feed. Batches of this type of feed were put into the pen; calves were free to eat as much of this as they liked. Calves were weighed individually.

Feed D Pen 1 10 calves	Feed C Pen 2 10 calves	Feed D Pen 3 10 calves	Feed B Pen 4 10 calves
Feed B Pen 5 10 calves	Feed A Pen 6 10 calves	Feed A Pen 7 10 calves	Feed C Pen 8 10 calves

treatment $=$ type of feed

Calf-feeding experiment

Calves were housed in pens, with ten calves per pen. Each pen was allocated to a certain type of feed. Batches of this type of feed were put into the pen; calves were free to eat as much of this as they liked. Calves were weighed individually.

Feed D Pen 1 10 calves	Feed C Pen 2 10 calves	Feed D Pen 3 10 calves	Feed B Pen 4 10 calves
Feed B Pen 5 10 calves	Feed A Pen 6 10 calves	Feed A Pen 7 10 calves	Feed C Pen 8 10 calves

treatment = type of feed   experimental unit = pen

Calf-feeding experiment

Calves were housed in pens, with ten calves per pen. Each pen was allocated to a certain type of feed. Batches of this type of feed were put into the pen; calves were free to eat as much of this as they liked. Calves were weighed individually.

Feed D Pen 1 10 calves	Feed C Pen 2 10 calves	Feed D Pen 3 10 calves	Feed B Pen 4 10 calves
Feed B Pen 5 10 calves	Feed A Pen 6 10 calves	Feed A Pen 7 10 calves	Feed C Pen 8 10 calves

treatment = type of feed   experimental unit = pen

observational unit = calf

Running example

0	160	240
160	80	80
80	0	160
240	240	0
↑	↑	↑
Cropper	Melba	Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑	↑	↑
Melba	Cropper	Melle

Running example

0	160	240
160	80	80
80	0	160
240	240	0
↑	↑	↑
Cropper	Melba	Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑	↑	↑
Melba	Cropper	Melle

experimental unit $=$ observational unit $=$ plot

Running example

0	160	240
160	80	80
80	0	160
240	240	0
↑ Cropper	↑ Melba	↑ Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑ Melba	↑ Cropper	↑ Melle

experimental unit $=$ observational unit $=$ plot

treatment $=$ combination of cultivar and amount of fertilizer

Treatments in the running example

Treatments are all combinations of:

factor	levels
Cultivar (C)	Cropper, Melle, Melba
Fertilizer (F)	0, 80, 160, 240 kg/ha

How many treatments are there?

Treatments in the running example

Treatments are all combinations of:

factor	levels
Cultivar (C)	Cropper, Melle, Melba
Fertilizer (F)	0, 80, 160, 240 kg/ha

How many treatments are there?

Cultivar	Fertilizer			
	0	80	160	240
Cropper	$\surd$	$\surd$	$\surd$	$\surd$
Melle	$\surd$	$\surd$	$\surd$	$\surd$
Melba	$\surd$	$\surd$	$\surd$	$\surd$

Treatments in another example

Treatments are all combinations of:

factor	levels
Timing (T)	early, late
Fertilizer (F)	0, 80, 160, 240 kg/ha

How many treatments are there?

Treatments in another example

Treatments are all combinations of:

factor	levels
Timing (T)	early, late
Fertilizer (F)	0, 80, 160, 240 kg/ha

How many treatments are there?

Timing	Fertilizer			
	0	80	160	240
None	$\surd$			
Early		$\surd$	$\surd$	$\surd$
Late		$\surd$	$\surd$	$\surd$

Absolute basics.

First, some notation.

$$\omega = \text{plot} = \text{observational unit}$$

$$T(\omega) = \text{treatment on plot } \omega$$

$$Y_\omega = \text{response on plot } \omega$$

$$\mathrm{E}(Y_\omega) = \tau_{T(\omega)}$$

So if ω is the third plot with treatment 2 then $\mathrm{E}(Y_\omega) = \tau_2$.

Absolute basics.

First, some notation.

$$\omega = \text{plot} = \text{observational unit}$$

$$T(\omega) = \text{treatment on plot } \omega$$

$$Y_\omega = \text{response on plot } \omega$$

$$\mathrm{E}(Y_\omega) = \tau_{T(\omega)}$$

So if ω is the third plot with treatment 2 then $\mathrm{E}(Y_\omega) = \tau_2$.

Calling this response Y_{23}

- ignores the plots;
- encourages non-blindness;
- encourages operation by treatment instead of by inherent factors.

Chapter 2 Unstructured Experiments

- Completely randomized designs
- Why and how to randomize

- Completely randomized designs
- Why and how to randomize
  - How do we randomize? Write down a systematic plan. Then choose a random permutation (from a computer, or shuffle a pack of cards) and apply it to the systematic plan.

Chapter 2 Unstructured Experiments

- Completely randomized designs
- Why and how to randomize

Chapter 2 Unstructured Experiments

- Completely randomized designs
- Why and how to randomize
- The treatment subspace, Orthogonal projection,
- Linear model, Estimation, Matrix notation
- Sums of squares, Variance
- Replication: equal or unequal

Chapter 2 Unstructured Experiments

- Completely randomized designs
- Why and how to randomize
- The treatment subspace, Orthogonal projection,
- Linear model, Estimation, Matrix notation
- Sums of squares, Variance
- Replication: equal or unequal
- Allowing for the overall mean, Hypothesis testing

Source	SS	df	MS	VR
mean	107161.3513	1	107161.3513	13147.39
diets	117.8964	2	58.9482	7.23
residual	236.3723	29	8.1508	–
Total	107515.62	32		

- Completely randomized designs
- Why and how to randomize
- The treatment subspace, Orthogonal projection,
- Linear model, Estimation, Matrix notation
- Sums of squares, Variance
- Replication: equal or unequal
- Allowing for the overall mean, Hypothesis testing

Source	SS	df	MS	VR
mean	107161.3513	1	107161.3513	13147.39
diets	117.8964	2	58.9482	7.23
residual	236.3723	29	8.1508	–
Total	107515.62	32		

Fitting the grand mean as a submodel of the treatment space is a first taste of what we shall do many times with structured treatments: fit submodels and see what is left over.

Chapter 2 Unstructured Experiments

- Completely randomized designs
- Why and how to randomize
- The treatment subspace, Orthogonal projection,
- Linear model, Estimation, Matrix notation
- Sums of squares, Variance
- Replication: equal or unequal
- Allowing for the overall mean, Hypothesis testing

Chapter 2 Unstructured Experiments

- Completely randomized designs
- Why and how to randomize
- The treatment subspace, Orthogonal projection,
- Linear model, Estimation, Matrix notation
- Sums of squares, Variance
- Replication: equal or unequal
- Allowing for the overall mean, Hypothesis testing
- Replication for power

- Replication of control treatments

- Replication of control treatments
- Comparing new treatments in the presence of a control

- Replication of control treatments
- Comparing new treatments in the presence of a control
- Other treatment groupings
  Repeated splitting of groupings, obtaining nested submodels without the complication of understanding interaction.

Drugs at different stages of development

A pharmaceutical company wants to compare 6 treatments for a certain disease. There are are 3 different doses of formulation A, that has been under development for some time, and 3 different doses (not comparable with the previous 3) of a new formulation B, that has not been so extensively studied.

Drugs at different stages of development

A pharmaceutical company wants to compare 6 treatments for a certain disease. There are are 3 different doses of formulation A, that has been under development for some time, and 3 different doses (not comparable with the previous 3) of a new formulation B, that has not been so extensively studied.

response depends on treatment

all doses of B give same response   all doses of A give same response

response depends only on formulation

null model

zero

- Types of block

Chapter 4 Blocking

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments
  - Continuous gradients—
    do block, but choice of block boundary is somewhat arbitrary, so block size can be chosen

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments
  - Continuous gradients—
    do block, but choice of block boundary is somewhat arbitrary, so block size can be chosen
  - Blocking for trial management—
    different technicians or different harvest times should be matched to other blocking if possible, otherwise used as new blocks

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments
  - Continuous gradients—
    do block, but choice of block boundary is somewhat arbitrary, so block size can be chosen
  - Blocking for trial management—
    different technicians or different harvest times should be matched to other blocking if possible, otherwise used as new blocks
- Orthogonal block designs—treatment i occurs n_i times in every block

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments
  - Continuous gradients—
    do block, but choice of block boundary is somewhat arbitrary, so block size can be chosen
  - Blocking for trial management—
    different technicians or different harvest times should be matched to other blocking if possible, otherwise used as new blocks
- Orthogonal block designs—treatment i occurs n_i times in every block
- Models for block designs—block effects may be fixed or random

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments
  - Continuous gradients—
    do block, but choice of block boundary is somewhat arbitrary, so block size can be chosen
  - Blocking for trial management—
    different technicians or different harvest times should be matched to other blocking if possible, otherwise used as new blocks
- Orthogonal block designs—treatment i occurs n_i times in every block
- Models for block designs—block effects may be fixed or random
- Analysis when blocks have fixed effects

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments
  - Continuous gradients—
    do block, but choice of block boundary is somewhat arbitrary, so block size can be chosen
  - Blocking for trial management—
    different technicians or different harvest times should be matched to other blocking if possible, otherwise used as new blocks
- Orthogonal block designs—treatment i occurs n_i times in every block
- Models for block designs—block effects may be fixed or random
- Analysis when blocks have fixed effects
- Analysis when blocks have random effects

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments
  - Continuous gradients—
    do block, but choice of block boundary is somewhat arbitrary, so block size can be chosen
  - Blocking for trial management—
    different technicians or different harvest times should be matched to other blocking if possible, otherwise used as new blocks
- Orthogonal block designs—treatment i occurs n_i times in every block
- Models for block designs—block effects may be fixed or random
- Analysis when blocks have fixed effects
- Analysis when blocks have random effects
- Why use blocks?

- Types of block
  - Natural discrete divisions—
    do block, but block size may be less than the number of treatments
  - Continuous gradients—
    do block, but choice of block boundary is somewhat arbitrary, so block size can be chosen
  - Blocking for trial management—
    different technicians or different harvest times should be matched to other blocking if possible, otherwise used as new blocks
- Orthogonal block designs—treatment i occurs n_i times in every block
- Models for block designs—block effects may be fixed or random
- Analysis when blocks have fixed effects
- Analysis when blocks have random effects
- Why use blocks?
- Loss of power with blocking

Twelve treatments are all combinations of:

factor	levels
Cultivar (C)	Cropper, Melle, Melba
Fertilizer (F)	0, 80, 160, 240 kg/ha

Chapter 5 Factorial Treatment Structure

Twelve treatments are all combinations of:

factor	levels
Cultivar (C)	Cropper, Melle, Melba
Fertilizer (F)	0, 80, 160, 240 kg/ha

$$E(Y_\omega) = \tau_{C(\omega),F(\omega)}$$

$$E(Y_\omega) = \lambda_{C(\omega)} + \mu_{F(\omega)}$$

$$E(Y_\omega) = \lambda_{C(\omega)} \qquad E(Y_\omega) = \mu_{F(\omega)}$$

$$E(Y_\omega) = \kappa$$

$$E(Y_\omega) = 0$$

Twelve treatments are all combinations of:

factor	levels
Cultivar (C)	Cropper, Melle, Melba
Fertilizer (F)	0, 80, 160, 240 kg/ha

$$E(Y_\omega) = \tau_{C(\omega),F(\omega)}$$

←Interaction

$$E(Y_\omega) = \lambda_{C(\omega)} + \mu_{F(\omega)}$$

$$E(Y_\omega) = \lambda_{C(\omega)} \qquad E(Y_\omega) = \mu_{F(\omega)}$$

$$E(Y_\omega) = \kappa$$

$$E(Y_\omega) = 0$$

Twelve treatments are all combinations of:

factor	levels
Cultivar (C)	Cropper, Melle, Melba
Fertilizer (F)	0, 80, 160, 240 kg/ha

$$\begin{tikzcd}
 & E(Y_\omega) = \tau_{C(\omega),F(\omega)} \arrow[d, dash, "\leftarrow\text{Interaction}"] & \\
 & E(Y_\omega) = \lambda_{C(\omega)} + \mu_{F(\omega)} \arrow[dl, dash] \arrow[dr, dash] & \\
E(Y_\omega) = \lambda_{C(\omega)} \arrow[dr, dash] & & E(Y_\omega) = \mu_{F(\omega)} \arrow[dl, dash] \\
 & E(Y_\omega) = \kappa \arrow[d, dash] & \\
 & E(Y_\omega) = 0 &
\end{tikzcd}$$

Most books give a single model which has these six models as special cases

Twelve treatments are all combinations of:

factor	levels
Cultivar (C)	Cropper, Melle, Melba
Fertilizer (F)	0, 80, 160, 240 kg/ha

$$E(Y_\omega) = \tau_{C(\omega),F(\omega)}$$

←Interaction

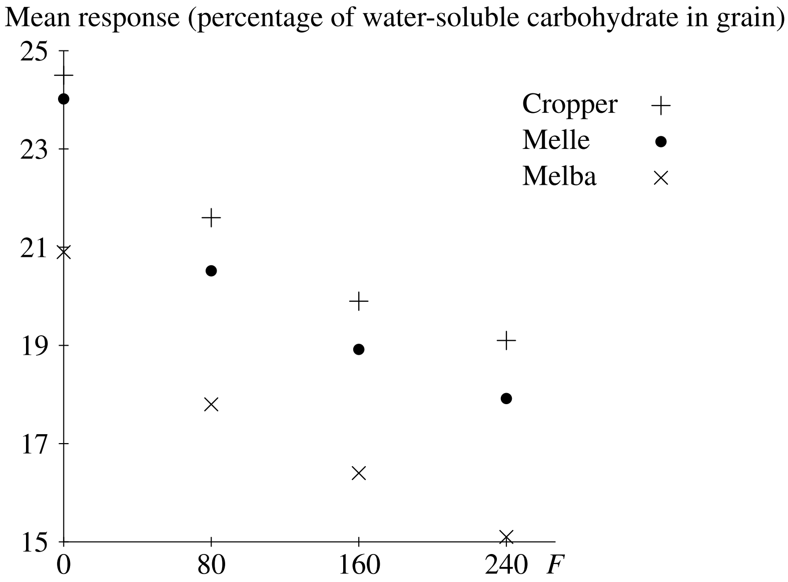
$$E(Y_\omega) = \lambda_{C(\omega)} + \mu_{F(\omega)}$$

$$E(Y_\omega) = \lambda_{C(\omega)} \qquad E(Y_\omega) = \mu_{F(\omega)}$$

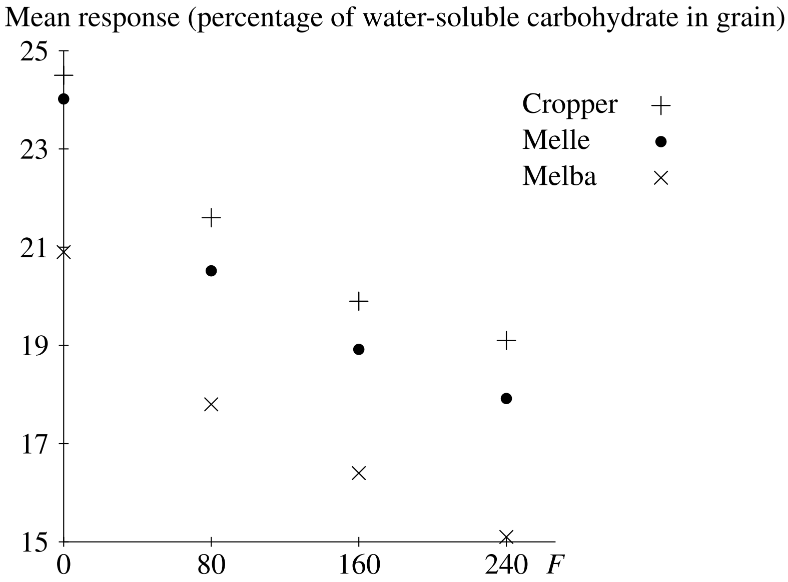
$$E(Y_\omega) = \kappa$$

$$E(Y_\omega) = 0$$

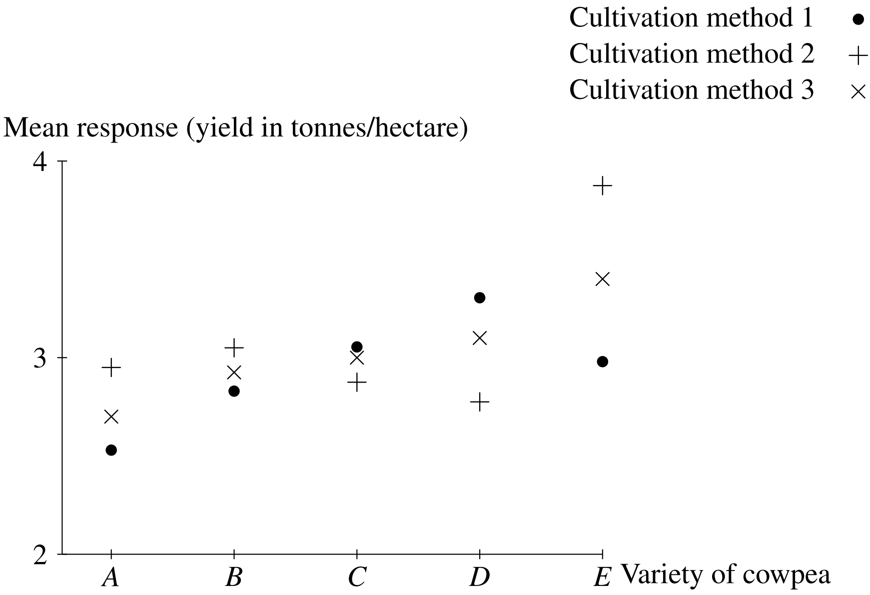
Most books give a single model which has these six models as special cases
but which also specializes to some inappropriate models, which your software may let you fit.

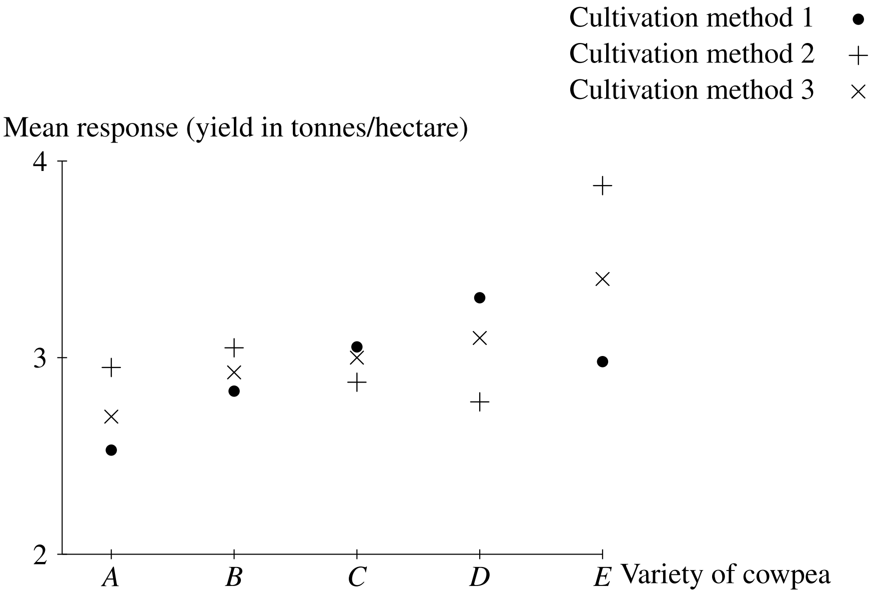


Mean response (percentage of water-soluble carbohydrate in grain)
25
23
21
19
17
15
0
80
160
240
F
Cropper
Melle
Melba



The difference between cultivars is (essentially) the same at each quantity of fertilizer—no interaction.


Cultivation method 1
Cultivation method 2
Cultivation method 3
Mean response (yield in tonnes/hectare)
4
3
2
A
B
C
D
E
Variety of cowpea



There is interaction between Variety and Cultivation Method.

Analysis of data (from factorial experiments)

1. Starting at the top of the model diagram, choose the smallest model that fits the data adequately.

Analysis of data (from factorial experiments)

1. Starting at the top of the model diagram, choose the smallest model that fits the data adequately.
2. Estimate the parameters of the chosen model.

Analysis of data (from factorial experiments)

1. Starting at the top of the model diagram, choose the smallest model that fits the data adequately.
2. Estimate the parameters of the chosen model.
3. There is no need to parametrize the other models.

Analysis of data (from factorial experiments)

1. Starting at the top of the model diagram, choose the smallest model that fits the data adequately.
2. Estimate the parameters of the chosen model.
3. There is no need to parametrize the other models.
4. Orthogonality $\Rightarrow$ different routes down the model diagram give consistent results.

- ...
- Three (or more) treatment factors
- Factorial experiments (benefits)
- Construction and randomization of factorial designs
- Factorial treatments plus control

Double blocking.
Wine-tasting example: treatments are 4 wines

	Judge							
Tasting	1	2	3	4	5	6	7	8
1								
2								
3								
4								

Double blocking.
Wine-tasting example: treatments are 4 wines

	Judge							
Tasting	1	2	3	4	5	6	7	8
1	*A*	*B*	*C*	*D*				
2	*D*	*A*	*B*	*C*				
3	*C*	*D*	*A*	*B*				
4	*B*	*C*	*D*	*A*				

a Latin square

Double blocking.
Wine-tasting example: treatments are 4 wines

	Judge							
Tasting	1	2	3	4	5	6	7	8
1	A	B	C	D	C	D	A	B
2	D	A	B	C	D	C	B	A
3	C	D	A	B	A	B	C	D
4	B	C	D	A	B	A	D	C
	a Latin square				and another			

Double blocking.
Wine-tasting example: treatments are 4 wines

	Judge							
Tasting	1	2	3	4	5	6	7	8
1	A	B	C	D	C	D	A	B
2	D	A	B	C	D	C	B	A
3	C	D	A	B	A	B	C	D
4	B	C	D	A	B	A	D	C
	a Latin square				and another			

Randomize the (order of) the 4 rows
Randomize the (order of) the 8 columns

- Applications of previous ideas

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.
  - Many designs are completely randomized.

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.
  - Many designs are completely randomized.
- Issues peculiar to such experiments

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.
  - Many designs are completely randomized.
- Issues peculiar to such experiments
  - Need for placebo

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.
  - Many designs are completely randomized.
- Issues peculiar to such experiments
  - Need for placebo
  - Sequential randomization to an unknown number of patients

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.
  - Many designs are completely randomized.
- Issues peculiar to such experiments
  - Need for placebo
  - Sequential randomization to an unknown number of patients
  - Ethical issues

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.
  - Many designs are completely randomized.
- Issues peculiar to such experiments
  - Need for placebo
  - Sequential randomization to an unknown number of patients
  - Ethical issues
    - Best for this patient or best for the trial?

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.
  - Many designs are completely randomized.
- Issues peculiar to such experiments
  - Need for placebo
  - Sequential randomization to an unknown number of patients
  - Ethical issues
    - Best for this patient or best for the trial?
  - Analysis by intention to treat

Chapter 7 Experiments on People and Animals

- Applications of previous ideas
  - A crossover trial with no carry-over effects is a row-column design.
  - A trial using matched pairs for a new drug and placebo is a block design.
  - Treatment centres may be regarded as blocks.
  - Many designs are completely randomized.
- Issues peculiar to such experiments
  - Need for placebo
  - Sequential randomization to an unknown number of patients
  - Ethical issues
    - Best for this patient or best for the trial?
  - Analysis by intention to treat
    - One mouthwash is more effective at preventing gum disease than another, but also more unpleasant, so some subjects may give up taking it.

Chapter 8 Small Units inside Large Units

Feed D Pen 1 10 calves	Feed C Pen 2 10 calves	Feed D Pen 3 10 calves	Feed B Pen 4 10 calves
Feed B Pen 5 10 calves	Feed A Pen 6 10 calves	Feed A Pen 7 10 calves	Feed C Pen 8 10 calves

Chapter 8 Small Units inside Large Units

Feed D Pen 1 10 calves	Feed C Pen 2 10 calves	Feed D Pen 3 10 calves	Feed B Pen 4 10 calves
Feed B Pen 5 10 calves	Feed A Pen 6 10 calves	Feed A Pen 7 10 calves	Feed C Pen 8 10 calves

Stratum	Source	Degrees of freedom	
mean	mean		1
pens	feed	3	
	residual	4	
	total		7
calves	calves		72
Total			80

Feed D Pen 1 10 calves	Feed C Pen 2 10 calves	Feed D Pen 3 10 calves	Feed B Pen 4 10 calves
Feed B Pen 5 10 calves	Feed A Pen 6 10 calves	Feed A Pen 7 10 calves	Feed C Pen 8 10 calves

Stratum	Source	Degrees of freedom
mean	mean	1
pens	feed	3
	residual	4 no matter how many calves per pen
	total	7
calves	calves	72
Total		80

Modification

The 4 feeds consist of all combinations of

- 2 types of hay, which must be put in whole pens
- 2 types of anti-scour treatment, which are given to calves individually.

Modification

The 4 feeds consist of all combinations of

- 2 types of hay, which must be put in whole pens
- 2 types of anti-scour treatment, which are given to calves individually.

Hay 2 Pen 1 5 calves A1 5 calves A2	Hay 2 Pen 2 5 calves A1 5 calves A2	Hay 2 Pen 3 5 calves A1 5 calves A2	Hay 1 Pen 4 5 calves A1 5 calves A2
Hay 1 Pen 5 5 calves A1 5 calves A2	Hay 1 Pen 6 5 calves A1 5 calves A2	Hay 1 Pen 7 5 calves A1 5 calves A2	Hay 2 Pen 8 5 calves A1 5 calves A2

Treatment factors in different strata

Stratum	Source	Degrees of freedom	
mean	mean		1
pens	hay	1	
	residual	6	
	total		7
calves	anti-scour	1	
	hay $\wedge$ anti-scour	1	
	residual	70	
	total		72
Total			80

Treatment factors in different strata

Stratum	Source	Degrees of freedom
mean	mean	1
pens	hay	1
	residual	6
	total	7
calves	anti-scour	1
	hay $\wedge$ anti-scour	1
	residual	70
	total	72
Total		80

Residual df for hay increase from 4 to 6, so power increases.

Treatment factors in different strata

Stratum	Source	Degrees of freedom
mean	mean	1
pens	hay	1
	residual	6
	total	7
calves	anti-scour	1
	hay $\wedge$ anti-scour	1
	residual	70
	total	72
Total		80

Residual df for hay increase from 4 to 6, so power increases.
Anti-scour and the interaction have smaller variance
(between calves within pens rather than between pens)
and substantially more residual df, so power increases.

(Classic) split-plot designs

Like the last one, but arrange the pens in complete blocks.

Using Latin squares for

- row-column designs
- two treatment factors with n levels each, in n blocks of size n, if it can be assumed that there is no interaction
- three treatment factors with n levels each, in n^2 experimental units, if it can be assumed that there is no interaction

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

Rows	Columns	Letters
Tasting	Judge	Wine
Field	Variety	Fertilizer
Factor 1	Factor 2	Factor 3

A factor F is a function for which we are more interested in knowing whether $F(\alpha) = F(\beta)$ than in knowing the value $F(\alpha)$.

A factor F is a function for which we are more interested in knowing whether $F(\alpha) = F(\beta)$ than in knowing the value $F(\alpha)$.

Let $\Omega =$ the set of observational units, and F a factor on Ω.
F-class containing α $= F[[\alpha]] = \{\omega \in \Omega : F(\omega) = F(\alpha)\}$.

A factor F is a function for which we are more interested in knowing whether $F(\alpha) = F(\beta)$ than in knowing the value $F(\alpha)$.

Let $\Omega =$ the set of observational units, and F a factor on Ω.
F-class containing $\alpha \; = F[[\alpha]] = \{\omega \in \Omega : F(\omega) = F(\alpha)\}$.

$F \equiv G$ (F is aliased with G) if every F-class is also a G-class.

A factor F is a function for which we are more interested in knowing whether $F(\alpha) = F(\beta)$ than in knowing the value $F(\alpha)$.

Let $\Omega =$ the set of observational units, and F a factor on Ω.
F-class containing α $= F[[\alpha]] = \{\omega \in \Omega : F(\omega) = F(\alpha)\}$.

$F \equiv G$ (F is aliased with G) if every F-class is also a G-class.

$F \prec G$ (F is finer than G) if every F-class is contained in a G-class but $F \not\equiv G$.

$F \preceq G$ if $F \prec G$ or $F \equiv G$.

A factor F is a function for which we are more interested in knowing whether $F(\alpha) = F(\beta)$ than in knowing the value $F(\alpha)$.

Let $\Omega =$ the set of observational units, and F a factor on Ω.
F-class containing α $= F[[\alpha]] = \{\omega \in \Omega : F(\omega) = F(\alpha)\}$.

$F \equiv G$ (F is aliased with G) if every F-class is also a G-class.

$F \prec G$ (F is finer than G) if every F-class is contained in a G-class but $F \not\equiv G$.

$F \preceq G$ if $F \prec G$ or $F \equiv G$.

The universal factor U has just one class.

A factor F is a function for which we are more interested in knowing whether $F(\alpha) = F(\beta)$ than in knowing the value $F(\alpha)$.

Let $\Omega =$ the set of observational units, and F a factor on Ω.
F-class containing $\alpha\ = F[[\alpha]] = \{\omega \in \Omega : F(\omega) = F(\alpha)\}$.

$F \equiv G$ (F is aliased with G) if every F-class is also a G-class.

$F \prec G$ (F is finer than G) if every F-class is contained in a G-class but $F \not\equiv G$.

$F \preceq G$ if $F \prec G$ or $F \equiv G$.

The universal factor U has just one class.

The equality factor E has one class per observational unit.

Running example

0	160	240
160	80	80
80	0	160
240	240	0
↑	↑	↑
Cropper	Melba	Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑	↑	↑
Melba	Cropper	Melle

Running example

0	160	240
160	80	80
80	0	160
240	240	0
↑	↑	↑
Cropper	Melba	Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑	↑	↑
Melba	Cropper	Melle

$E = \text{plot} \prec \text{strip} \prec \text{field} \prec U$

Running example

0	160	240
160	80	80
80	0	160
240	240	0
↑ Cropper	↑ Melba	↑ Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑ Melba	↑ Cropper	↑ Melle

$E = \text{plot} \prec \text{strip} \prec \text{field} \prec U$

$\text{strip} \prec \text{cultivar}$

Infimum of two factors

Given two factors F and G,
the factor $F \wedge G$ is defined by

$$(F \wedge G)[[\omega]] = F[[\omega]] \cap G[[\omega]].$$

Running example

0	160	240
160	80	80
80	0	160
240	240	0
↑	↑	↑
Cropper	Melba	Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑	↑	↑
Melba	Cropper	Melle

$$\text{cultivar} \wedge \text{fertilizer} = \text{treatment}$$

Running example

0	160	240
160	80	80
80	0	160
240	240	0
↑	↑	↑
Cropper	Melba	Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑	↑	↑
Melba	Cropper	Melle

$$\text{cultivar} \wedge \text{fertilizer} = \text{treatment}$$
$$\text{field} \wedge \text{cultivar} = \text{strip}$$

Supremum of two factors

Given two factors F and G,
the factor $F \vee G$ is the finest factor whose classes are unions of F-classes and unions of G-classes.

Supremum of two factors

Given two factors F and G,
the factor $F \vee G$ is the finest factor whose classes are unions of F-classes and unions of G-classes.

If you try to fit F and G in a linear model,
you will get into trouble unless you fit $F \vee G$ first.

Running example

Cropper	Melba	Melle
0	160	240
160	80	80
80	0	160
240	240	0

Melba	Cropper	Melle
160	80	0
0	160	80
240	0	240
80	240	160

$\text{field} \vee \text{fertilizer} = U$

Running example

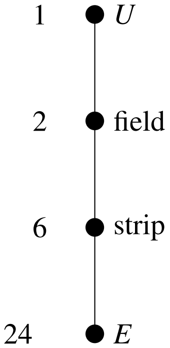
0	160	240
160	80	80
80	0	160
240	240	0
↑	↑	↑
Cropper	Melba	Melle

160	80	0
0	160	80
240	0	240
80	240	160
↑	↑	↑
Melba	Cropper	Melle

$\text{field} \vee \text{fertilizer} = U$

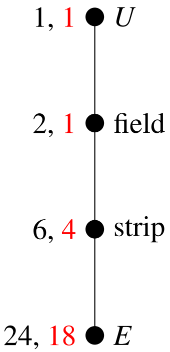
$\text{strip} \vee \text{treatment} = \text{cultivar}$

Hasse diagram for factors on the observational units



How many of each are there?

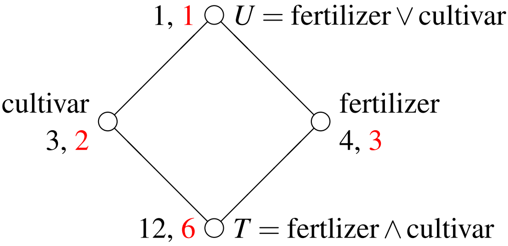
Hasse diagram for factors on the observational units



How many of each are there?

Degrees of freedom calculated by subtraction

Hasse diagram for factors on the treatments



Factorial treatments plus control

dose	type				
	Z	S	K	M	N
none	$\surd$				
single		$\surd$	$\surd$	$\surd$	$\surd$
double		$\surd$	$\surd$	$\surd$	$\surd$

Factorial treatments plus control

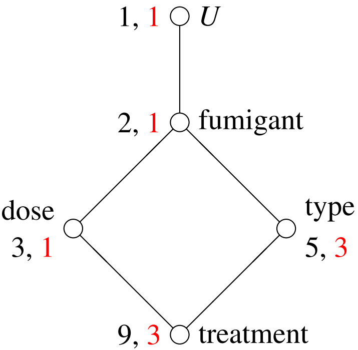
dose	type				
	Z	S	K	M	N
none	$\surd$				
single		$\surd$	$\surd$	$\surd$	$\surd$
double		$\surd$	$\surd$	$\surd$	$\surd$

$$\text{dose} \vee \text{type} = \text{fumigant}$$

Factorial treatments plus control

dose	type				
	Z	S	K	M	N
none	$\surd$				
single		$\surd$	$\surd$	$\surd$	$\surd$
double		$\surd$	$\surd$	$\surd$	$\surd$

$$\text{dose} \vee \text{type} = \text{fumigant}$$



Hence a complete theory for orthogonal designs,
including the location of treatment subspaces in the correct strata.

This covers everything so far,
and there are many further examples.

Blocks are incomplete if

- the block size is less than the number of treatments
- no treatment occurs more than once in any block.

Blocks are incomplete if

- the block size is less than the number of treatments
- no treatment occurs more than once in any block.

Balanced incomplete-block designs and square lattice designs.

Blocks are incomplete if

- the block size is less than the number of treatments
- no treatment occurs more than once in any block.

Balanced incomplete-block designs and square lattice designs.

Inserting a control treatment in every block.

If the number of blocks is equal to the number of treatments, algorithm to arrange the blocks as the columns of a row-column design in such a way that each treatment occurs once per row.

Combining the above.

Characters	Treatments								
A	0	0	0	1	1	1	2	2	2
B	0	1	2	0	1	2	0	1	2
$A+B$	0	1	2	1	2	0	2	0	1
$A+2B$	0	2	1	1	0	2	2	1	0
$2A+B$	0	1	2	2	0	1	1	2	0
$2A+2B$	0	2	1	2	1	0	1	0	2
$2A$	0	0	0	2	2	2	1	1	1
$2B$	0	2	1	0	2	1	0	2	1
I	0	0	0	0	0	0	0	0	0

Characters	Treatments								
A	0	0	0	1	1	1	2	2	2
B	0	1	2	0	1	2	0	1	2
$A+B$	0	1	2	1	2	0	2	0	1
$A+2B$	0	2	1	1	0	2	2	1	0
$2A+B$	0	1	2	2	0	1	1	2	0
$2A+2B$	0	2	1	2	1	0	1	0	2
$2A$	0	0	0	2	2	2	1	1	1
$2B$	0	2	1	0	2	1	0	2	1
I	0	0	0	0	0	0	0	0	0

$A \equiv 2A$ main effect of A

$B \equiv 2B$ main effect of B

$A+B \equiv 2A+2B$ 2 degrees of freedom for the A-by-B interaction

$A+2B \equiv 2A+B$ 2 degrees of freedom for the A-by-B interaction, orthogonal to the previous 2

Characters	Treatments								
A	0	0	0	1	1	1	2	2	2
B	0	1	2	0	1	2	0	1	2
$A+B$	0	1	2	1	2	0	2	0	1
$A+2B$	0	2	1	1	0	2	2	1	0
$2A+B$	0	1	2	2	0	1	1	2	0
$2A+2B$	0	2	1	2	1	0	1	0	2
$2A$	0	0	0	2	2	2	1	1	1
$2B$	0	2	1	0	2	1	0	2	1
I	0	0	0	0	0	0	0	0	0

$A \equiv 2A$	main effect of A
$B \equiv 2B$	main effect of B
$A+B \equiv 2A+2B$	2 degrees of freedom for the A-by-B interaction
$A+2B \equiv 2A+B$	2 degrees of freedom for the A-by-B interaction, orthogonal to the previous 2

For 3 blocks of size 3, can alias blocks with any character.

A factorial design is a fractional replicate if not all possible combinations of the treatment factors occur.

A fractional replicate can be useful if there are a large number of treatment factors to investigate and we can assume that some interactions are zero.

Chapter 9 constructed some fractional replicate designs from Latin squares.

Here we use characters to give us more types of fractional replicate.

Includes quantile plots for analysis.

Chapter 14 Backward Look

1. Randomization
2. Factors such as time, sex, age and breed—
   Are they treatment factors or plot factors?
3. Writing a protocol
4. …

Examples, Questions and Exercises

Not all the examples are agricultural.

Almost all of the examples in this book are real.
On the other hand, almost none of them is the whole truth.

Each chapter ends with questions for discussion:
there is no single correct answer.

There are more general exercises at the end.

Sources of all these are given, as far as possible.

A question from Chapter 1

Several studies have suggested that drinking red wine gives some protection against heart disease, but it is not known whether the effect is caused by the alcohol or by some other ingredient of red wine. To investigate this, medical scientists enrolled 40 volunteers into a trial lasting 28 days. For the first 14 days, half the volunteers drank two glasses of red wine per day, while the other half had two standard drinks of gin. For the remaining 14 days the drinks were reversed: those who had been drinking red wine changed to gin, while those who had been drinking gin changed to red wine. On days 14 and 28, the scientists took a blood sample from each volunteer and measured the amount of inflammatory substance in the blood.

Identify the experimental units and observational units. How many are there of each? What is the plot structure?

What are the treatments? What is the treatment structure?

A question from Chapter 5

A group of people researching ways to reduce the risk of blood clotting are planning their next experiments. One says:

> *We know that aspirin thins the blood. Let's experiment with the quantity of aspirin. We could enrol about 150 healthy men into the trial, give 50 of them one aspirin tablet per day for a year, another 50 one and a half aspirin tablets a day, and the final 50 will get two aspirin tablets per day.*
> *When we have decided which quantity is best, we can run another trial to find out if there is any difference between taking the aspirin after breakfast or after dinner.*

How do you reply?

A question from Chapter 11

A horticulture research institute wants to compare nine methods of treating a certain variety of houseplant while it is being grown in a greenhouse in preparation for the Christmas market. One possibility is to ask twelve small growers to test three treatments each in separate chambers in their greenhouses. A second possibility is to ask three large commercial growers to test nine methods each, also in separate greenhouse chambers.

1. Construct a suitable design for the first possibility.
2. Randomize this design.
3. If the plots stratum variance is the same in both cases, which design is more efficient?
4. Compare the designs in terms of likely cost, difficulty and representativeness of the results.

Design of the Month

Rest and exercise

The following quotation is taken from page 18 of the *New Scientist* of 28 July 2007.

> "Less pain, more gain" may become the new mantra for gym junkies. Taking a break during your workout may result in you burning more fat than the same amount of exercise without a break, according to a report from Kazushige Goto of the University of Tokyo, Japan, and his colleagues (*Journal of Applied Physiology*). They studied seven men with an average age of 25. On different days, the men did no exercise, exercised on stationary bikes for 1 hour, or exercised at the same intensity for two half-hour periods separated by a 20-minute rest.

1. What were the experimental units and how many of them were there?
2. What were the treatments and how many of them were there?
3. How do you think the experiment was designed?
4. Can you improve on the design?

Design of the Month
Preventing diabetes in the local Bangladeshi community

The following quotation is taken from the May 2007 issue of *The Bulletin* of Queen Mary, University of London.

> Colleagues ... are piloting how to best identify those at highest risk in the local Bangladeshi community using GP registers. They will then test the feasibility of interventions to prevent diabetes. In a factorial design, usual care provided by the GP will be compared to a behavioural change programme to improve lifestyle run by 'lay tutors' attached to Social Action for Health; medication strategies and a mix of lifestyle education and medication.

1. What are the observational units?
2. What are the experimental units?
3. What are the treatments and how many of them are there?
4. Draw the Hasse diagram for the factors on the treatments.
5. How should this trial be randomized?

Design of the Month
The greening of healthcare

The following quotation is taken from page 32 of the *New Scientist* of 22 December 2007.

> We also compared the effects of running on a treadmill while runners were faced with one of four views, which we classified as rural pleasant, rural unpleasant, urban pleasant and urban unpleasant. There was also a control group who had no view at all, as in most gyms. “Rural pleasant” was the winner, with improved psychological outcomes and substantially reduced blood pressure, while the “urban unpleasant” view came bottom. Runners with “no view” fared better than those viewing gritty urban scenes.

1. What were the observational units?
2. What were the experimental units?
3. What were the treatments and how many of them were there?
4. Draw the Hasse diagram for the factors on the treatments.