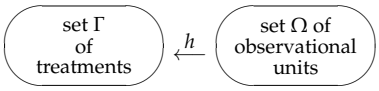
Multi-phase experiments: from design to analysis
Session 1: Concepts in experimental design and analysis

R. A. Bailey
University of St Andrews
QMUL (emerita)

Australasian Region of the International Biometric Society,
November 2015

Basic ideas of experimental design



A design consists of

- a set Γ of treatments;
- a set Ω of observational units;
- a function $h\colon \Omega \to \Gamma$ allocating treatment $h(\omega)$ to observational unit ω.

The allocation h is usually made by starting with a systematic design and then randomizing the observational units.

We usually try to choose h so as to minimize variances of estimators and maximize power for hypothesis tests, subject to any operational constraints.

Examples of structured sets of treatments

Structure	Example	Panel Diagram
unstructured	diets A, B, C, D, E	5 Diets
all combinations of levels of two treatment factors	Treatment A B C D E F Nitrogen 0 0 0 1 1 1 Phosphate 0 1 2 0 1 2	2 Nitrogen 3 Phosphate
… or of three treatment factors		3 Methods 5 Quantities 2 Times

Examples of structured sets of observational units

Structure	Example	Panel Diagram
unstructured		25 Plots
rectangle	Subject; Time	8 Subjects 4 Times

More examples of structured sets of observational units

Structure	Example	Panel Diagram
blocks	Cage 1 1 1 1 2 2 2 2 Rat 1 2 3 4 5 6 7 8	2 Cages 4 Rats in C
double nesting	6 fields, each of 3 plots, each split into 4 subplots	6 Fields 3 Plots in F 4 Subplots in P, F

Poset block structures

All the structured sets discussed in this session are poset block structures.

This means that each can be defined by a panel diagram, which shows

- a list of factors (CJB) or pre-factors (RAB) $F_1, \ldots, F_m$;
- for each F_i, its number n_i of levels;
- for each F_i, what it is nested in.

There is a set of $n_1 \times \cdots \times n_m$ objects, one for each combination of levels of $F_1, \ldots, F_m$.

"F_i is nested in F_j" means
"if two objects have the same level of F_i then this has no significance unless they have the same level of F_j".

The real factors (RAB) or generalized factors (CJB) are combinations of levels of none or more of $F_1, \ldots, F_m$ subject to the rule that if F_i is included and F_i is nested in F_j then F_j must be included.

Randomization: basic idea

5 Diets — 10 People

1. Systematic design: each treatment twice.

Person	1	2	3	4	5	6	7	8	9	10
Diet	*A*	*A*	*B*	*B*	*C*	*C*	*D*	*D*	*E*	*E*

2. Apply a random permutation to the 10 people.

Person	9	2	1	10	7	4	8	5	6	3
Diet	*E*	*A*	*A*	*E*	*D*	*B*	*D*	*C*	*C*	*B*

3. Relabel the observational units in standard order.

Person	1	2	3	4	5	6	7	8	9	10
Diet	*E*	*A*	*A*	*E*	*D*	*B*	*D*	*C*	*C*	*B*

Randomization: a rectangle

5 Methods — 5 Subjects, 5 Times

1. Systematic design: a Latin square with times as rows and subjects as columns.

	1	2	3	4	5
1	1	2	3	4	5
2	4	5	1	2	3
3	2	3	4	5	1
4	5	1	2	3	4
5	3	4	5	1	2

	1	2	3	4	5
1	1	2	3	4	5
3	2	3	4	5	1
4	5	1	2	3	4
5	3	4	5	1	2
2	4	5	1	2	3

	5	4	2	3	1
1	5	4	2	3	1
3	1	5	3	4	2
4	4	3	1	2	5
5	2	1	4	5	3
2	3	2	5	1	4

2. Randomize times; randomize subjects.
   Note: this can be achieved by a single permutation of Ω.
3. Relabel both times and subjects in standard order (not shown).

Randomization: blocks

2 Nitrogen, 3 Phosphate — 3 Blocks, 6 Plots in B

1. Systematic design: each treatment once per block.

Block	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3
Plot	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Nitrogen	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	1
Phosphate	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2

2. Randomize blocks; randomize plots within each block independently. (Achievable by a single permutation of Ω.)

Block	2	2	2	2	2	2	1	1	1	1	1	1	3	3	3	3	3	3
Plot	12	11	7	9	10	8	2	4	6	3	1	5	16	17	13	14	15	18
Nitrogen	1	1	0	0	1	0	0	1	1	0	0	1	1	1	0	0	0	1
Phosphate	2	1	0	2	0	1	1	0	2	2	0	1	0	1	0	1	2	2

3. Relabel both blocks and plots in standard order (not shown).

Decomposition into subspaces

set Γ of treatments — $\xleftarrow{h}$ — set Ω of observational units

vector space $V_\Gamma = \mathbb{R}^\Gamma$ — $V_\Gamma \to\!\!\leq V_\Omega$ — vector space $V_\Omega = \mathbb{R}^\Omega$

Decomposition $\mathcal{R}$ of V_Γ into orthogonal subspaces

Decomposition $\mathcal{Q}$ of V_Ω into orthogonal subspaces

Under orthogonality (this session), $\mathcal{R}$ further decomposes $\mathcal{Q}$.

- What happens if the design (allocation) is not orthogonal?
- What happens if there are 2 or more phases?

Expectation models

Let $\mathbf{Y} = (Y_\omega)_{\omega\in\Omega}$ be the vector of reponses.
Assume that $E(\mathbf{Y}) \in V_\Gamma$.

Expectation models are various subspaces of V_Γ, such as those corresponding to

- all expected values are the same;
- expected values depend only on the level of treatment factor A;
- the expected values are additive in the sense that they are the sum of something that depends only on the level of treatment factor A and something that depends only on the level of treatment factor B.

These lead to a decomposition of V_Γ into orthogonal subspaces.

Expectation model subspaces: Nitrogen and Phosphate

$E(\mathbf{Y}) \in V_N \iff$ there are constants α_i such that $E(Y_\omega) = \alpha_i$ whenever $N(\omega) = i$.

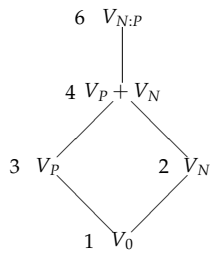
$E(\mathbf{Y}) \in V_P \iff$ there are constants β_j such that $E(Y_\omega) = \beta_j$ whenever $P(\omega) = j$.

$E(\mathbf{Y}) \in V_0 \iff$ there is a constant μ such that $E(Y_\omega) = \mu$ for all ω.

$E(\mathbf{Y}) \in V_N + V_P \iff$ there are constants α_i and β_j such that $E(Y_\omega) = \alpha_i + \beta_j$ if $N(\omega) = i$ and $P(\omega) = j$.

$E(\mathbf{Y}) \in V_{N:P} \iff$ there are constants λ_{ij} such that $E(Y_\omega) = \lambda_{ij}$ if $N(\omega) = i$ and $P(\omega) = j$.

Hasse diagram for expectation model subspaces



full model

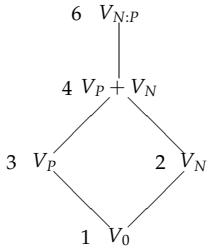
additive model

only factor N makes any difference

null model

These subspaces are all the sums of none or more of
V_N, V_P and $V_{N:P}$
(under the convention that the sum of none of them is V_0).
Abbreviate this as $N + P + N : P$.

Main effects and interaction



The interaction between factors N and P is the difference between the vector of fitted values in $V_{N:P}$ and the vector of fitted values in $V_P + V_N$.

The main effect of factor N is the difference between the vector of fitted values in V_N and the vector of fitted values in V_0.

The vector of fitted values in V_0 has the grand mean in every coordinate.

This decomposition is indicated by the formula $N + P + N : P$,
or Nitrogen + Phosphate + Nitrogen : Phosphate.

Example of treatment decomposition

Treatment	A	B	C	D	E	F
Nitrogen	0	0	0	1	1	1
Phosphate	0	1	2	0	1	2

vectors						explanation	name
1	1	1	1	1	1	constant	overall mean
1	1	1	−1	−1	−1	differences between levels of N	main effect of N
1	−1	0	1	−1	0	differences between levels of P	main effect of P
0	−1	1	0	−1	1		
1	−1	0	−1	1	0	differences not explained by main effects of N and P	N-by-P interaction
0	−1	1	0	1	−1		

Subspaces and orthogonal projectors

Define $\Omega \times \Omega$ matrices $\mathbf{J}$ and $\mathbf{J}_N$ by

$$\mathbf{J}(\alpha, \beta) = 1 \quad \text{for all } \alpha \text{ and } \beta;$$

$$\mathbf{J}_N(\alpha, \beta) = \begin{cases} 1 & \text{if } \alpha \text{ and } \beta \text{ have the same level of N} \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that the six treatments all have replication r.

name	notation	projector
mean	U_0	$\mathbf{R}_0 = (6r)^{-1}\mathbf{J}$
main effect of N	U_N	$\mathbf{R}_N = (3r)^{-1}\mathbf{J}_N - \mathbf{R}_0$
main effect of P	U_P	$\mathbf{R}_P = (2r)^{-1}\mathbf{J}_P - \mathbf{R}_0$
N-by-P interaction	U_{NP}	$\mathbf{R}_{NP} = r^{-1}\mathbf{J}_{NP} - \mathbf{R}_0 - \mathbf{R}_N - \mathbf{R}_P$

This extends to three or more factors.

Variance-covariance matrix

Let $\mathbf{V} = \mathrm{Cov}(\mathbf{Y})$.
Let G be the set of allowable permutations for randomization.
Randomization lets us assume that $\mathbf{V}(\alpha_1, \beta_1) = \mathbf{V}(\alpha_2, \beta_2)$ if there is any g in G such that $g(\alpha_1) = \alpha_2$ and $g(\beta_1) = \beta_2$.

All our examples are poset block structures.
One consequence is that $\mathbf{V}$ has such a nice pattern that we can identify its eigenspaces W_j and their orthogonal projector $\mathbf{Q}_j$.
These are labelled by the real factors.

The eigenspaces are called strata. The spectral form of $\mathbf{V}$ is

$$\mathbf{V} = \sum_j \eta_j \mathbf{Q}_j,$$

where the (usually unknown) eigenvalues η_j must be non-negative and are called spectral components of variance.

Example of variance-covariance matrix

3 Blocks
6 Plots in B

$$\begin{aligned}
\mathbf{V} &= \sigma^2\mathbf{I} + \rho_1\sigma^2(\mathbf{J}_B - \mathbf{I}) + \rho_2\sigma^2(\mathbf{J} - \mathbf{J}_B) \\
&= \sigma^2(1-\rho_1)\mathbf{I} + \sigma^2(\rho_1 - \rho_2)\mathbf{J}_B + \sigma^2\rho_2\mathbf{J} \\
&= \sigma^2(1-\rho_1)(\mathbf{I} - 6^{-1}\mathbf{J}_B) + \sigma^2(1 + 5\rho_1 - 6\rho_2)(6^{-1}\mathbf{J}_B - (18)^{-1}\mathbf{J}) \\
&\quad + \sigma^2(1 + 5\rho_1 + 12\rho_2)(18)^{-1}\mathbf{J} \\
&= \eta_{BP}\mathbf{Q}_{BP} + \eta_B\mathbf{Q}_B + \eta_0\mathbf{Q}_0 \\
&= \psi_{BP}\mathbf{I} + \psi_B\mathbf{J}_B + \psi_0\mathbf{J} \\
&= \psi_{BP}\mathbf{Q}_{BP} + (\psi_{BP} + 6\psi_B)\mathbf{Q}_B + (\psi_{BP} + 6\psi_B + 18\psi_0)\mathbf{Q}_0.
\end{aligned}$$

The randomization model demands that the
spectral components η_{BP}, η_B and η_0 be non-negative;
the usual mixed model demands that the
canonical components ψ_{BP}, ψ_B and ψ_0 be non-negative.

Abbreviation for variance-covariance matrix

3 Blocks
6 Plots in B

$$\begin{aligned}\mathbf{V} &= \eta_{BP}\mathbf{Q}_{BP} + \eta_B\mathbf{Q}_B + \eta_0\mathbf{Q}_0 \\ &= \psi_{BP}\mathbf{I} + \psi_B\mathbf{J}_B + \psi_0\mathbf{J}.\end{aligned}$$

η_0 and ψ_0 can never be estimated,
and so ψ_0 is often set to be zero, which gives

$$\begin{aligned}\mathbf{V} &= \eta_{BP}\mathbf{Q}_{BP} + \eta_B(\mathbf{Q}_B + \mathbf{Q}_0) \\ &= \psi_{BP}\mathbf{I} + \psi_B\mathbf{J}_B\end{aligned}$$

This is indicated by the formula

$$\text{Plots} + \text{Blocks}.$$

If you renumber the plots within each block, then this must be written as

$$\text{Blocks} : \text{Plots} + \text{Blocks}.$$

Orthogonality

Given the treatment subspaces U_i,
with their orthogonal projectors $\mathbf{R}_i$,
and the strata W_j,
with their orthogonal projectors $\mathbf{Q}_j$,
the design is orthogonal if each U_i is contained in a single W_j.

This means that $\mathbf{R}_i\mathbf{Q}_j = \mathbf{Q}_j\mathbf{R}_i = \mathbf{R}_i$
and $\mathbf{R}_i\mathbf{Q}_k = \mathbf{Q}_k\mathbf{R}_i = \mathbf{0}$ if $k \neq j$.

If $U_i \leq W_j$
then all the information about $\mathbf{R}_iE(\mathbf{Y})$ is contained in $\mathbf{Q}_j\mathbf{Y}$,
whose variance-covariance matrix $\eta_j\mathbf{Q}_j$ is effectively scalar.

If every $\mathbf{R}_i$ commutes with every $\mathbf{Q}_j$ then
the treatment subspace V_Γ can be decomposed
(usually by using pseudofactors)
to make the design orthogonal.

Skeleton analysis of variance: complete-block-design (1)

2 Nitrogen
3 Phosphate

3 Blocks
6 Plots in B

plots Ω	
source	df
Mean	1
Blocks	5
Plots[Blocks]	12

treatments Γ	
source	df
Mean	1
main effect of N	1
main effect of P	2
N-by-P interaction	2

Subspaces are called 'sources';
and there are various conventions for labelling them.

Skeleton analysis of variance: complete-block-design (2)

plots Ω	
source	df
Mean	1
Blocks	5
Plots[Blocks]	12

treatments Γ	
source	df
Mean	1
main effect of N	1
main effect of P	2
N-by-P interaction	2

$U_0 = W_0$ always. The systematic design in this case ensures that all the other treatment subspaces are contained in W_{BP}.

plots Ω		treatments Γ	
source	df	source	df
Mean	1	Mean	1
Blocks	5		
Plots[Blocks]	12	main effect of N	1
		main effect of P	2
		N-by-P interaction	2
		Residual	7

Skeleton analysis of variance: complete-block-design (3)

plots Ω		treatments Γ	
source	df	source	df
Mean	1	Mean	1
Blocks	5		
Plots[Blocks]	12	main effect of N	1
		main effect of P	2
		N-by-P interaction	2
		Residual	7

The skeleton analysis of variance shows

- which Ω-source the main effect of N is confounded with, hence the likely magnitude of the variance of the estimator of the contrast between two levels ($(2/9)\eta_{BP}$);
- ditto main effect of P, and the N-by-P interaction;
- the relevant residual term, and its df, hence the likely precision of the estimators of those variances, and information about the power of any hypothesis tests.

Estimation

Of course, the analysis of variance table is only one part of data analysis.
We usually want to estimate treatment effects.

When the design is orthogonal, all estimation of expectation parameters is by ordinary least squares.

For a treatment effect in stratum W_i,
the relevant standard error is a known multiple of $\sqrt{\eta_i}$;
and η_i is estimated by the residual mean square in this stratum.

Skeleton analysis of variance: Latin square

5 Methods

5 Subjects
5 Times

obs. units Ω	
source	df
Mean	1
Times	4
Subjects	4
T#S	16

treatments Γ	
source	df
Mean	1
Methods	4

The Latin square design ensures that $U_{\text{Methods}} \leq W_{TS}$.

obs. units Ω		treatments Γ	
source	df	source	df
Mean	1	Mean	1
Times	4		
Subjects	4		
T#S	16	Methods	4
		Residual	12

A split-plot design

3 Varieties of oats
4 Nitrogen quantities

6 Blocks
3 Whole-plots in B
4 Subplots in W, B

1. Constraint: varieties must be sown on whole-plots.
2. Systematic design:
   each variety on one whole-plot per block;
   each quantity of nitrogen on one subplot per whole-plot.
3. Randomization: randomize blocks;
   randomize whole-plots within each block independently;
   randomize subplots within each whole-plot independently.
   (This builds a single permuation of the subplots.)

A split-plot design: skeleton anova

3 Varieties of oats
4 Nitrogen quantities

6 Blocks
3 Whole-plots in B
4 Subplots in W, B

obs. units Ω		treatments Γ	
source	df	source	df
Mean	1	Mean	1
Blocks	5		
Whole-plots[B]	12	Varieties	2
		Residual	10
Subplots[B,W]	54	Nitrogen	3
		V#N	6
		Residual	45

Warning about notation

Sources in anova tables are often labelled by real factors.
Their interpretation depends on which other real factors are included.

N may indicate "the main effect of N"
which is not the same as "the factor N".

$N : P$ may indicate "the interaction between N and P"
or "the effect of N within levels of P;
neither of these is the same the real factor $N : P$.