MT4614 Design of Experiments

R. A. Bailey
University of St Andrews

21 January 2022

Section 1.5: Linear Model

The response on observational unit ω is a random variable Y_ω with observed value y_ω.

We assume that

$$Y_\omega = Z_\omega + \tau_i \qquad \text{if } T(\omega) = i,$$

where

- τ_i is a constant depending on treatment i;
- Z_ω is a random variable depending on ω.

The probability space is the set of occasions and uncontrolled conditions under which the experiment might be carried out.

If $\alpha \neq \beta$ then Z_α and Z_β are different random variables on the same probability space, so they have a joint distribution.
Here are some common assumptions on the joint distribution of $\{Z_\omega : \omega \in \Omega\}$.

(i) The simple textbook model says that the Z_ω are independent, identically distributed $N(0, \sigma^2)$.

(ii) The fixed-effects model says that the Z_ω are independent $N(\mu_\omega, \sigma^2)$,
where μ_ω depends on how ω fits into the plot structure.

(iii) The random-effects model says that the Z_ω have identical distributions, and the correlation between Z_α and Z_β depends on how α and β are related in the plot structure.

(iv) The randomization model says that the Z_ω have identical distributions, and the correlation between Z_α and Z_β depends on the method of randomization.

Break for non-technical stuff

All information about MT4614, including timetable,
summary of material covered, problem sheets (when available),
data files (when needed), is on the web page
`http://www-groups.mcs.st-and.ac.uk/~rab/MT4614/`
There is a direct link to this from MMS.

I encourage you all to write out the notes of each lecture in your own handwriting, with your own comments and explanations added. This is really the only way that the material will get into your brain.

If you have any questions, do not hesitate to email me at
`rab24@st-andrews.ac.uk`
(but do not expect replies during weekends).

Any questions on this lecture so far?

2.1 Completely randomized designs

Suppose that

$$\text{experimental unit} = \text{observational unit} = \text{plot},$$

and it has been decided to give treatment i replication r_i, so that

$$\sum_{i=1}^{t} r_i = N.$$

(i) Number the plots $1, 2, \ldots, N$.

(ii) Allocate treatment 1 to plots $1, 2, \ldots, r_1$,
treatment 2 to plots $r_1 + 1, \ldots, r_1 + r_2$,
and so on.

(iii) Choose a random permutation of $\{1, 2, \ldots, N\}$ and apply it to the design.

(iv) Re-number the plots in natural order.

Section 2.1: How do I get a random permutation?

- Shuffle a pack of cards three times.
- Get a sequence of random numbers from your phone or calculator,
  put 1 in the position of the smallest,
  2 in the position of the next smallest, and so on.
- Get a sequence of random numbers from
  a table of random numbers in a book of statistical tables,
  by starting at a random place in the table.
- Use the online tool called Research Randomizer.
- Use other software, such as R or GAP.

Example

$t = 3, \mathcal{T} = \{A, B, C\}, N = 13, r_A = 5, r_B = r_C = 4.$

1	2	3	4	5	6	7	8	9	10	11	12	13
A	A	A	A	A	B	B	B	B	C	C	C	C

Research Randomizer gives

4	11	1	6	12	8	3	2	9	5	10	7	13
A	C	A	B	C	B	A	A	B	A	C	B	C

Re-number the plots in natural order.

1	2	3	4	5	6	7	8	9	10	11	12	13
A	C	A	B	C	B	A	A	B	A	C	B	C

Section 2.1: Layout sheet and data sheet

Layout

Plot	Treatment
1	A
2	C
3	A
4	B
5	C
6	B
7	A
8	A
9	B
10	A
11	C
12	B
13	C

Data sheet

Plot	Response
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	

The data sheet is a blinded version of the layout: the person who records the data is not biased by knowing the treatments.

Section 2.2: The treatment subspace

Let $V = \mathbb{R}^\Omega =$ the set of all real vectors whose coordinates are indexed by plots in Ω.

Put $V_T = \{\mathbf{v} \in V : v_\alpha = v_\beta \text{ whenever } T(\alpha) = T(\beta)\}$.

V_T is called the treatment subspace;

vectors in V_T are called treatment vectors.

$$\dim(V_T) = |\mathcal{T}| = t.$$

If

$$\mathbf{v} \in V_T \text{ and } \sum_{\omega \in \Omega} v_\omega = 0$$

then $\mathbf{v}$ is a treatment contrast.

Ending with more non-technical stuff

The next lecture uses some properties of variance and covariance of random vectors. In case you have forgotten this material, you can find some notes about it on the web page.