Example of treatment structure.

E1.12. Simple fungicide

– treatment : 3 doses of fungicide

full spray / half / no → one factor, 3 levels

2 treatments (full spray / half) 1 control (no)

Ex1.13. Fungicide factorial

– fungicide could be sprayed

early / mid-season / late / any combinations of them

early	mid	late
yes +	yes +	yes +
no −	no −	no −

early	mid	late
+	+	+
+	+	−
+	−	+
−	+	+
+	−	−
−	+	−
−	−	+
−	−	−

full factorial design ← 3 factors, each 2 levels ←

	spray midseason		no mid spray-season	
	spray late	no late spray	spray late	no late spray
spray early	√	√	√	√
no early spray	√	√	√	√

Ex1.14. Fungicide factorial plus control

– doses : full spray / half / no

– time : early / mid / late

– Spray is applied only once.

– Q : how many treatments? what's their structure?

2 factors
each 3 levels
9 treatments

	early	mid-season	late
full spray	√	√	√
half spray	√	√	√
no spray	√	√	√

not make sense

7 treatments

	early	mid-season	late	n/a
full spray	√	√	√	
half spray	√	√	√	
no spray				√

one factor 3 levels

one factor 2 levels

1 control

one factor, 7 levels

6 treatments 1 control

Ex 1.15. Oilseed rape

- expt on the control of disease
- compare 4 "new" chemicals, A, B, C, D
  with { (1) no treatment
  (2) standard chemical X
- each of the new chemicals could be applied early / late
- Q: how many treatments?
  Q: What's their structure?

one factor 2 levels

	early	late	both	n/a
none				√
A	√	√		
B	√	√		
C	√	√		
D	√	√		
X			√	

one factor, 4 levels

another control

one control

10 treatments

8 of them: full factorial

2 of them: control

early: Yes / No; late: Yes / No; new chemical: A, B, C, D

$2\times2\times4=16$

$16-4=12$

$12+2=\underline{\underline{14}}$

Q: What if new chemicals could be applied late / early / any combination of them?

notation for treatment
- a treatment: lower-case Latin letters, such as $i, j, \cdots$
- whole set of treatments: $\mathcal{T}$
- # of treatments $= |\mathcal{T}| = t$

notation for plots
- a plot: lower-case Greek letters such as $\alpha, \beta, \gamma, \omega$
- whole set of plots: Ω
- # of plots $= |\Omega| = N$

Def: treatment structure: meanful ways of dividing the set $\mathcal{T}$ of treatments

Def: plot structure: meanful way of dividing the set Ω of plots

Treatment tructure

- Unstructured
- Several new treatments plus control [Ex 1.12]
- All combinations of two (three, …) factors [Ex 1.8, 1.10, 1.13]
- All combinations of two (three, …) factors, plus control [Ex 1.14]
- Increasing doses of a quantitative factor

Plot structure

- Unstructured
- Experimental units containing observational units [Ex 1.2]
- Blocks: dividing the experimental units into homogeneous blocks [Ex 1.9: housewives]
- Blocks containing subblocks containing plots [Ex 1.5]
- Blocks containing experimental units containing observational units
- Two (several) different sorts of blocks, neither containing the other [Ex 1.7]

Q: What's the difference between factors (treatments) & blocks?

Milliken & Johnson (1992), *Analysis of Messy Data*, Vol.1, Chapter 4

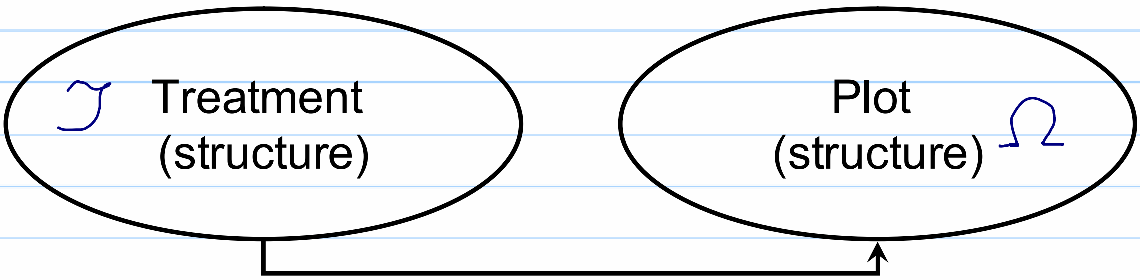
Treatment structure

- One-way
- Two-way
- Factorial arrangement
- Fractional factorial arrangement
- Factorial arrangement with one/more controls

Plot (design) structure

- Completely randomized
- Randomized complete block
- Latin square
- Incomplete block
- Various combinations and generalization

Any type of treatment structure can occur with any type of plot structure



allocation of treatments to plots
and randomization

<u>Def</u>: The <u>design</u> is the allocation of treatments to plots

- each plot can receive only one treatment
- a treatment could be applied to several plots
- therefore, the design is a function T

$$T: \Omega \to \mathcal{T}$$

- plot ω is allocated treatment $T(\omega)$

Def: The plan or layout is the design translated into actual plots.

- Some randomization is usually involved in this translation process.

A DOE should include 4 parts:

1. choice of treatments (structure)
2. choice of plots (structure)
3. design: $T: \Omega \to \mathcal{T}$
4. plan/layout.

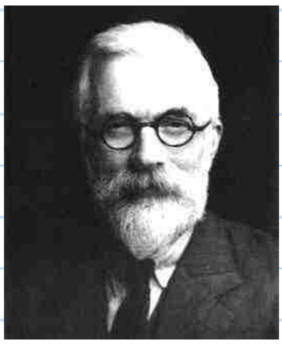
1-18

Why need randomization?

When assigning the treatments to the plots

- Fundamental difficulty: variability among the plots; no two plots are exactly the same.
- Different responses may be observed even if the same treatment is assigned to the plots.
- Systematic assignments may lead to bias.

R. A. Fisher worked at the Rothamsted Experimental Station in the United Kingdom to evaluate the success of various fertilizer treatments.



1-19

Fisher found the data from experiments going on for decades to be basically worthless because of poor experimental design.

- Fertilizer had been applied to a field one year and not in another in order to compare the yield of grain produced in the two years. BUT
  - It may have rained more, or been sunnier, in different years.
  - The seeds used may have differed between years as well.

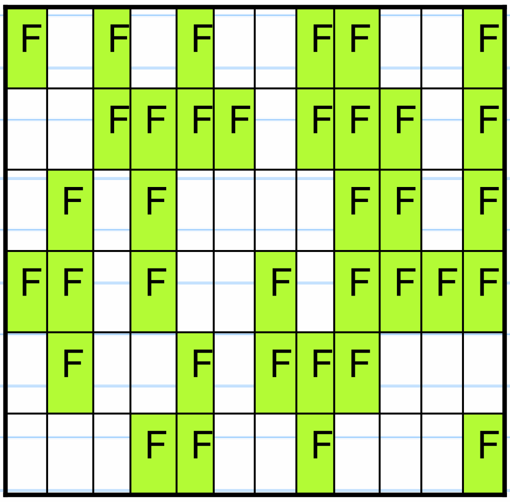
- Or fertilizer was applied to one field and not to a nearby field in the same year. BUT
  - The fields might have different soil, water, drainage, and history of previous use.

➔ Too many factors affecting the results were "uncontrolled."

controllable ⇒ block

Fisher's solution: Randomization 隨機化

- In the same field and same year, apply fertilizer to randomly spaced plots within the field.
- This averages out the effect of variation within the field in drainage and soil composition on yield, as well as controlling for weather, etc.



Randomization prevents any particular treatment from receiving more than its fair share of better plots, thereby may eliminate potential systematic bias. Some treatments may still get lucky, but if we assign many plots to each treatment, then the effects of chance will average out.

In addition to guarding against potential systematic biases, randomization also provides a basis for doing statistical inference. (⇒ Randomization model)

Block what you can, randomize what you cannot