DESIGN OF EXPERIMENTS

Chapter 9:

I. Observational studies and designed experiments
II. Experimental Units
III. Randomisation
IV. Completely Randomised Designs
V. Replication
VI. Blocking
VII. Randomised block designs
VIII. Latin square designs
IX. Controls
X. Balance

Observational studies

An **observational study** is one in which we observe and record what has already happened.

Observational studies:

- Measurements are taken, but there is no attempt to influence the response.
- Includes all surveys.
- Can show association, but cannot establish cause and effect.

Designed experiments

A **designed experiment** or **trial** is where the investigator intervenes and determines which experimental units receive which treatments.

Designed experiments can reasonably establish cause and effect.

Limitations of designed experiments:

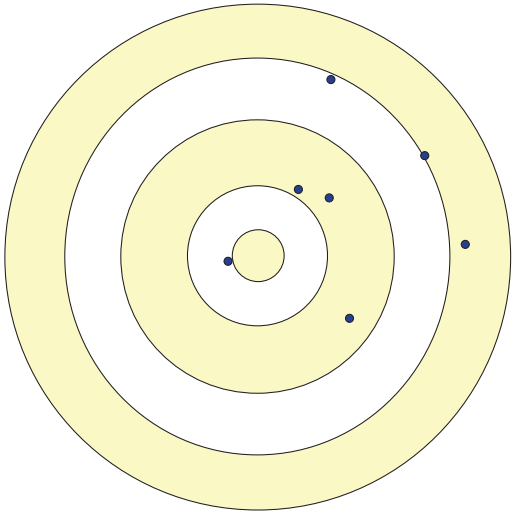
- Sometimes not feasible (usually due to ethical reasons);
- Limited to scientifically testable theories and hypotheses;

Experimental units

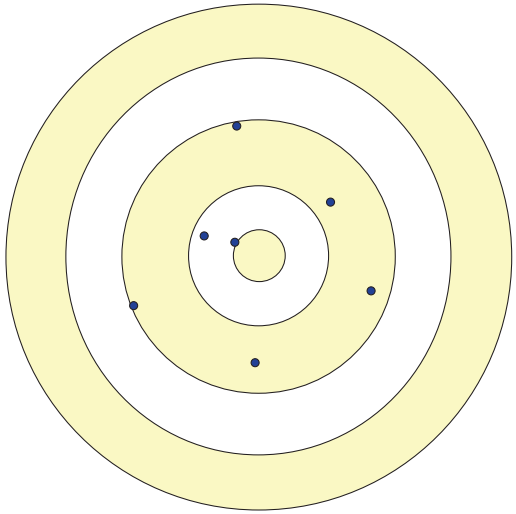
Experimental units: the objects or items to which the different treatments are allocated and applied. Some examples:

- a plot of land;
- a plant;
- a leaf;
- a cow;
- a flock of sheep;
- a block of cheese;
- a sealed package of pear slices;
- a vat of milk;
- a run of a manufacturing process;
- a person;
- an eye;
- a class of students.

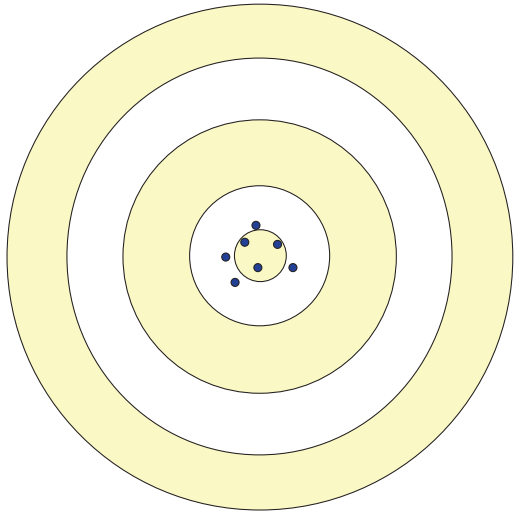
Principles of experimental design: validity and precision



Be aware of sources of error



Reduce bias (validity)



Reduce scatter (precision)

Good experimental design includes:

1. Randomisation (for validity);
2. Replication (for precision);
3. Blocking (in most experiments it improves precision).

Randomisation

Randomisation is the use of chance in allocating treatments to experimental units.

Randomisation:

- Prevents bias (ensures validity);
- Guards against confounding;
- Is not haphazard allocation;
- Is not systematic allocation.

Completely randomised designs

- The most basic design;
- Treatments are allocated to experimental units at random, with no restrictions.
- Every possible allocation of treatments to experimental units is equally likely.

Example: Experiment on the effect of four different pesticides on the yield of citrus trees. Twelve trees of the same variety in the same field are available.

Tree:	1	2	3	4	5	6	7	8	9	10	11	12
Pesticide:	3	1	4	3	4	1	2	4	2	3	1	2

Using R: `> sample(rep(1:4,3),12,replace=F)`

Replication

Replication is applying a treatment to more than one experimental unit.

- Needed because of *variation*;
- Enables the separation of systematic variation (due to treatments) and random variation (error).

Increased replication:

- Gives greater **precision**—narrower confidence intervals, smaller P-values, etc.;
- Also gives a more precise estimate of the error.

"False replication"

False replication occurs when experimental units are measured more than once but treated as separate experimental units.

Example:

Measuring trees within a multi-tree plot of a forest experiment and analysing the data as if each tree was an experimental unit.

Measuring experimental units more than once is fine, but it doesn't create more replicates!

Blocking

A **block** is a group of *similar* experimental units (in a way that is expected to affect the response variable).

The purpose of blocking (or matching, pairing, etc.) is to give greater precision, by identifying the contribution to variation from different blocks.
E.g. if you are interested in the effects of one of two fertilizers on tomato plants, and you believe that two varieties of tomatoes may react differently, then you create two blocks (Tomato type A and Tomato type B), and apply the different fertilizers in each of the two blocks. The goal is to separate the effects of Tomato type and fertilizer type.

Good blocking is usually cheaper than increased replication, both of which improve precision.

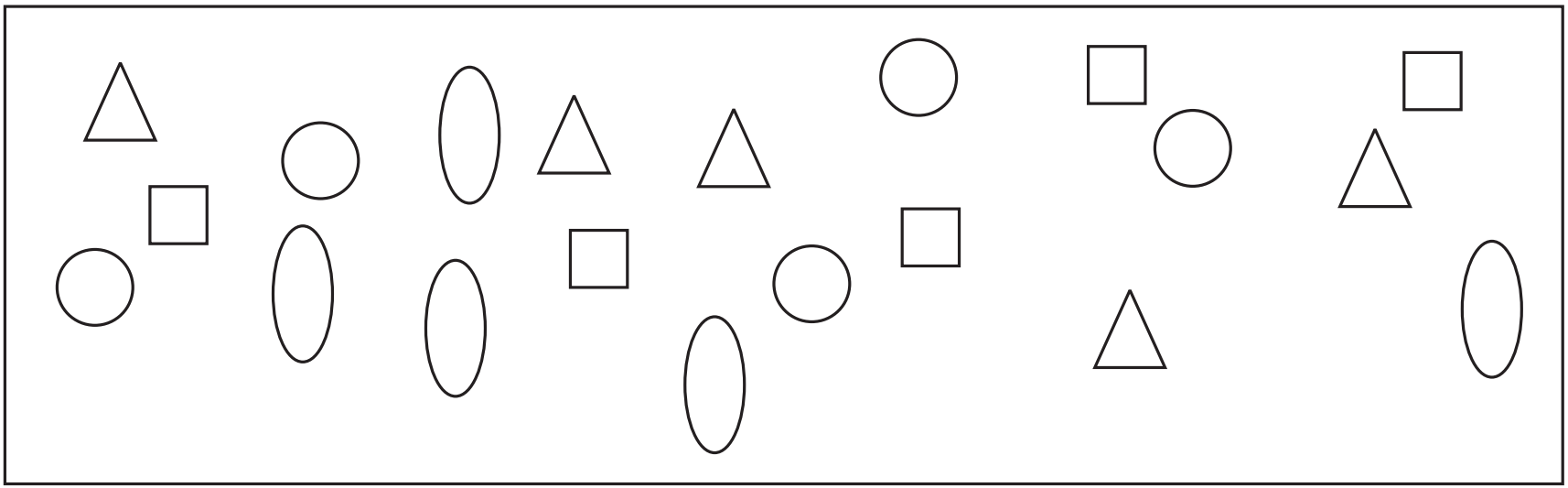
Blocking

Examples of blocks:

- adjacent plots in a field;
- pairs of identical twins;
- vehicles manufactured at the same time;
- litters of animals;
- periods of time, such as a day;
- batches of material, such as a vat of milk or a truckload of timber;
- groups of people matched by age, gender and ethnicity;
- trays of apples in cold storage.

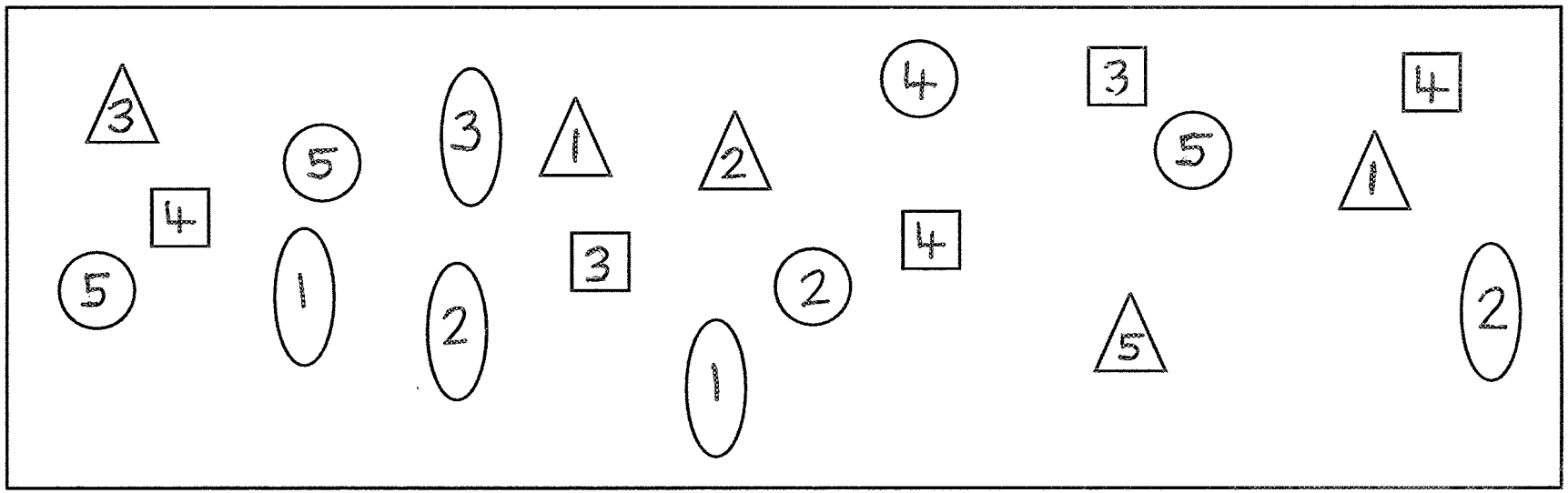
Blocking

Non-homogeneous groups:



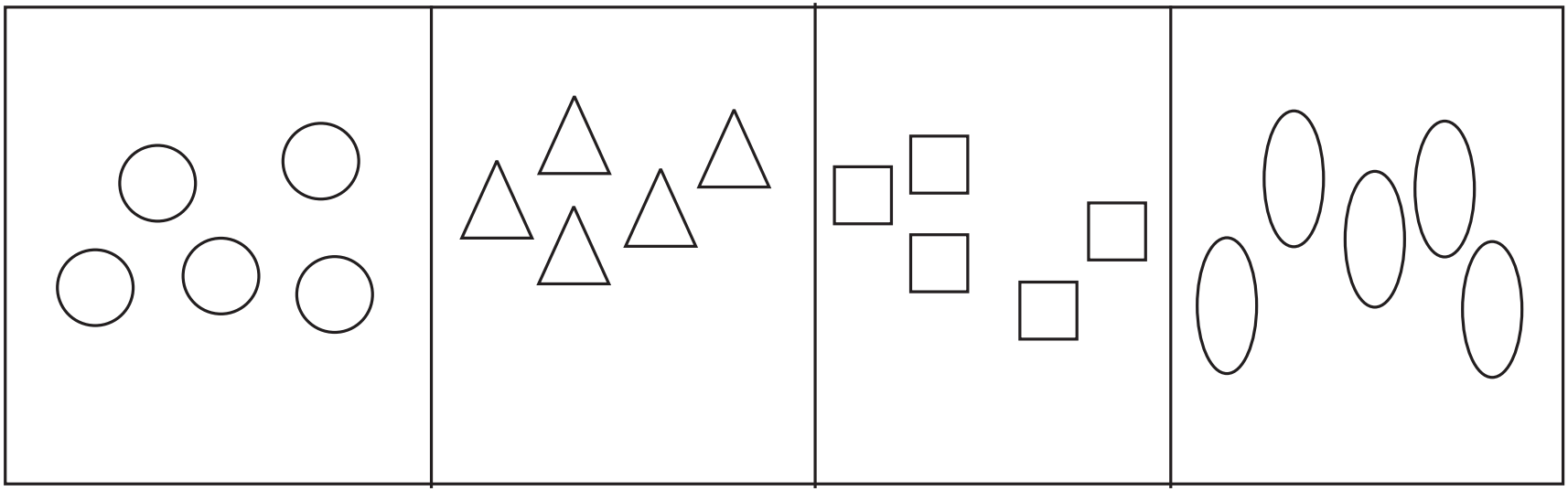
Blocking

Non-homogeneous groups – randomisation:



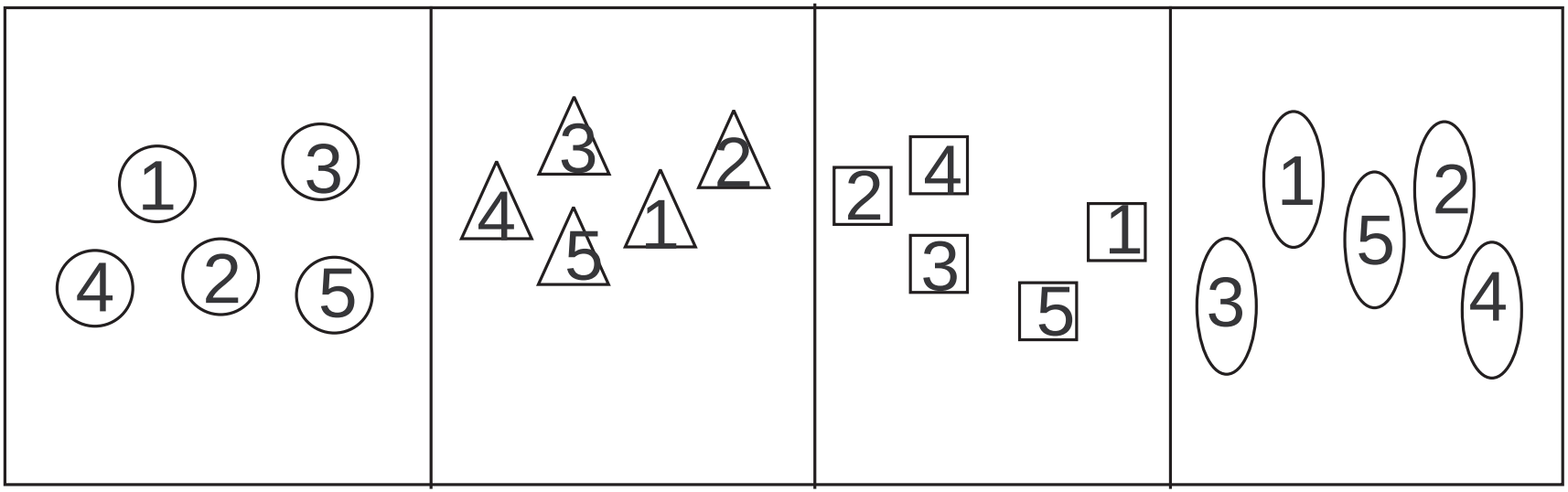
Blocking

Blocking into homogeneous groups:

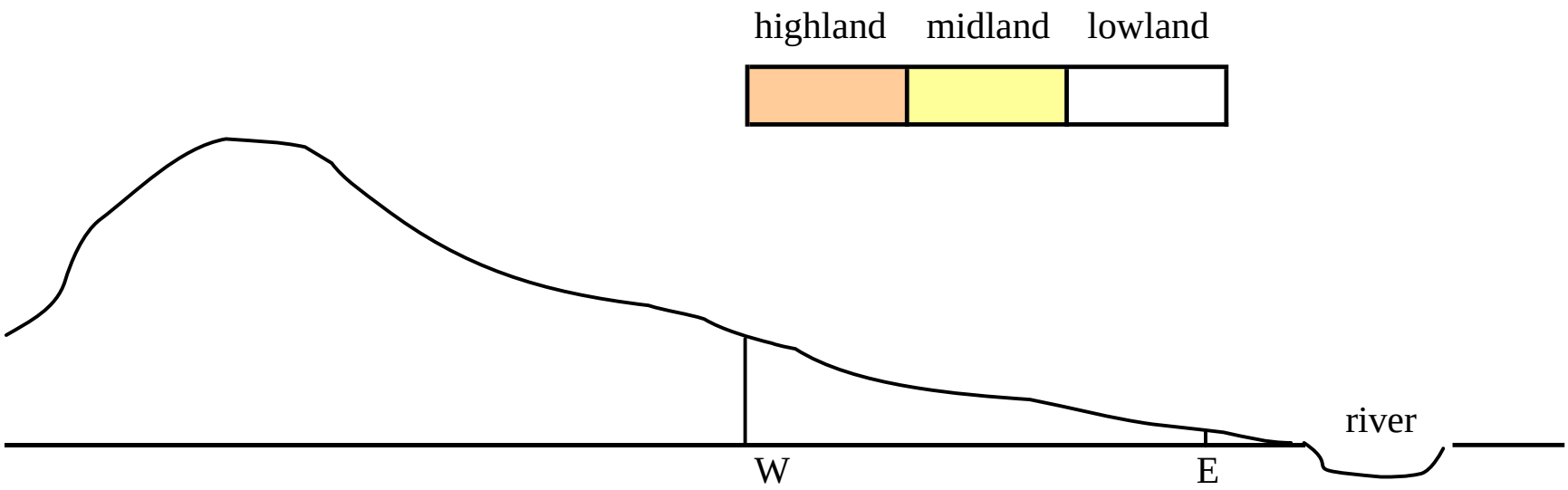


Blocking

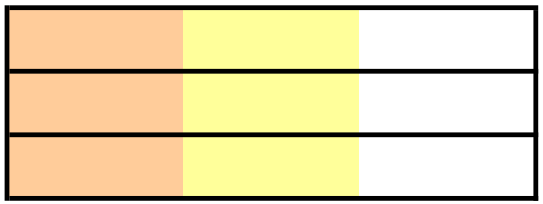
Randomisation within blocks:



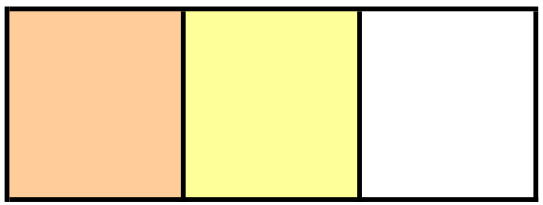
Blocking



Wrong blocking



Correct blocking



Randomised block designs

- Each treatment occurs the same number of times (usually once) in each block;
- If exactly once, we have a *randomised complete block design*;
- Treatments are randomly assigned to experimental units within each block separately;
- Other terminology: “stratified randomisation”, “matched pairs”.

Randomised block designs

Example: Experiment on the effect of four different pesticides on the yield of citrus trees. Twelve trees of the same variety in the same field are available, in 3 blocks of 4 trees each.

	Block 1				Block 2				Block 3			
Tree:	1	2	3	4	5	6	7	8	9	10	11	12
Pesticide:	1	4	2	3	4	3	2	1	2	3	1	4

Potato wireworm experiment

4 replications $\times$ 4 treatments

= 16 plots in a 4 $\times$ 4 square:

Randomised block design:

slope $\longrightarrow$

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

Potato wireworm experiment

Accounting for both slope and soil texture:

	slope $\longrightarrow$			
clay	3	1	2	4
$\downarrow$	1	4	3	2
	4	2	1	3
loam	2	3	4	1

Blocking in two directions: a Latin square.

Latin square designs

- Incorporate two blocking factors (often called rows and columns);
- There must be the same number of rows and columns as treatments;
- Each treatment must appear exactly once in each row and in each column.

Multiple Latin squares

- The number of levels of one of the blocking factors is an exact multiple of the number of levels of the other blocking factor and the treatment;
- The different squares should have separate randomisations of treatments to plots.

Example: Effect of 3 food supplements on milk yield of cows.
6 cows $\times$ 3 successive time periods: Two 3×3 Latin squares:

Cow	Period		
	I	II	III
1	A	B	C
2	B	C	A
3	C	A	B
4	B	A	C
5	A	C	B
6	C	B	A

Controls

- Necessary to sustain the argument of causation; experiment compares "what happened with what would have happened without the intervention";
- Experiments without them are often biased in favour of the treatment, because of the placebo effect;
- "Historical controls" are not sufficient;
- Can be no treatment, or an established treatment, or an "industry standard";
- Should match the treatments in as many ways as possible;
- Comparison with a placebo may be less useful than with a standard treatment;
- Should be included as an integral part of the experimental design and analysis.

Blind studies

- Experiments should be blind, if possible;
- It is even better if they are double-blind;
- Blinding is especially desirable if the subjects are people.

Balance

Balance means applying each treatment to the same number of experimental units.

Lack of balance can occur because of:

- experimental units being lost to the experiment;
- numbers of experimental units available exceeding a multiple of the number of treatments;
- deliberately replicating one treatment more than another because of its relative importance.

Balance

A balanced design has two advantages:

1. It simplifies the analysis.
2. It gives the most precise comparisons.

Example: $n_1 = 20, n_2 = 20$: $\sqrt{\frac{1}{20} + \frac{1}{20}} = 0.316.$

$n_1 = 30, n_2 = 10$: $\sqrt{\frac{1}{30} + \frac{1}{10}} = 0.365.$ Larger SEs & CIs.

But: $n_1 = 30, n_2 = 20$: $\sqrt{\frac{1}{30} + \frac{1}{20}} = 0.289.$

Better than $n_1 = n_2 = 20$.

Balance is good—*but increased replication is even better.*